
Senior Oracle DBA Servival Guide

By Feroz (Oracle Certified)

Oracle DBA Basics FAQs

1. What is an instance? Draw Architecture?
2. What is SGA?
3. What is PGA (or) what is pga_aggregate_target?
4. What are new memory parameters in Oracle 10g?
5. What are new memory parameters in Oracle 11g?
6. What are the mandatory background processes?
7. What are the optional background processes?
8. What are the new background processes in Oracle 10g?
9. How do you use automatic PGA memory management with Oracle 9i and above?
10. Explain two easy SQL optimizations?
11. What are the new features in Oracle 11gR1?
12. What are the new features in Oracle 11g R2?
13. What are the new features in Oracle 12c?
14. What process will get data from datafiles to DB cache?
15. What background process will writes data to datafiles?
16. What background process will write undo data?
17. What are physical components of Oracle database?
18. What are logical components of Oracle database?
19. Types of segment space management?
20. Types of extent management?
21. What are the differences between LMTS and DMTS?
22. What is a datafile?
23. What are the contents of control file?
24. What is the use of redo log files?
25. What are the uses of undo tablespace or redo segments?
26. How undo tablespace can guarantee retain of required undo data?
27. What is ORA-01555 - snapshot too old error and how do you avoid it?
28. What is the use/size of temporary tablespace?
29. What is the use of password file?
30. How to create password file?
31. How many types of indexes are there?
32. What is bitmap index & when it'll be used?
33. What is B-tree index & when it'll be used?
34. How you will find out fragmentation of index?
35. What is the difference between delete and truncate?
36. What's the difference between a primary key and a unique key?
37. What is the difference between schema and user?
38. What is the difference between SYSDBA, SYSOPER and SYSASM?
39. What is the difference between SYS and SYSTEM?
40. What is the difference between view and materialized view?
41. What are materialized view refresh types and which is default?
42. How fast refresh happens?
43. How to find out when was a materialized view refreshed?
44. What is materialized view log (type)?

45. What is atomic refresh in mvviews?
46. How to find out whether database/tablespace/datafile is in backup mode or not?
47. What is row chaining?
48. What is row migration?
49. What are different types of partitions?
50. What is local partitioned index and global partitioned index?
51. How you will recover if you lost one/all control file(s)?
52. Why more archivelogs are generated, when database is begin backup mode?
53. What UNIX parameters you will set while Oracle installation?
54. What is the use of inittrans and maxtrans in table definition?
55. What are differences between dbms_job and dbms_scheduler?
56. What are differences between dbms_scheduler and cron jobs?
57. Difference between CPU & PSU patches?
58. What you will do if (local) inventory corrupted [or] opatch lsinventory is giving error?
59. What are the entries/location of oraInst.loc?
60. What is the difference between central/global inventory and local inventory?
61. What is the use of root.sh & oraInstRoot.sh?
62. What is transportable tablespace (and across platforms)?
63. How can you transport tablespaces across platforms with different endian formats?
64. What is xtss (cross platform transportable tablespace)?\
65. What is the difference between restore point & guaranteed restore point?
66. What is the difference between 10g/11g OEM Grid control and 12c Cloud control?
67. What are the components of Grid control?
68. What are the new features of 12c Cloud control?
69. How to find if your Oracle database is 32 bit or 64 bit?
70. How to find opatch Version?
71. Which of the following does not affect the size of the SGA?
72. A set of Dictionary tables are created?
73. The order in which Oracle processes a single SQL statement is?
74. What are the mandatory datafiles to create a database in Oracle 11g?
75. In one server can we have different oracle versions?
76. How do sessions communicate with database?
77. Which SGA memory structure cannot be resized dynamically after instance startup?
78. When a session changes data, where does the change get written?
79. How many maximum no of control files we can have within a database?
80. System Data File Consists of?
81. What is the function of SMON in instance recovery?
82. Which action occurs during a checkpoint?
83. SMON process is used to write into LOG files?
84. Oracle does not consider a transaction committed until?
85. How many maximum DBWn (Db writers) we can invoke?
86. Which activity would generate less undo data?
87. What happens when a user issues a COMMIT?
88. What happens when a user process fails?
89. What are the free buffers in the database buffer cache?
90. When the SMON Process perform ICR?
91. Which dynamic view can be queried when a database is started up in no mount state?
92. Which two tasks occur as a database transitions from the mount stage to the open stage?

93. In which situation is it appropriate to enable the restricted session mode?
94. Which is the component of an Oracle instance?
95. Which process is involved when a user starts a new session on the database server?
96. In the event of an Instance failure, which files store command data NOT written to the datafiles?
97. When are the base tables of the data dictionary created?
98. Sequence of events takes place while starting a Database is?
99. The alert log will never contain information about which database activity?
100. Where can you find the non-default parameters when an instance is started?
101. Which tablespace is used as the temporary tablespace if TEMPORARY TABLESPACE is not specified for a user?
102. User SCOTT creates an index with this statement: CREATE INDEX emp_indx on employee (empno). In which tablespace would be the index created?
103. Which data dictionary view shows the available free space in a certain tablespace?
104. Which methods increase the size of a tablespace?
105. What does the command ALTER DATABASE . . . RENAME DATAFILE do?
106. Can you drop objects from a read-only tablespace?
107. SYSTEM TABLESPACE can be made off-line?
108. Data dictionary can span across multiple Tablespaces?
109. Multiple Tablespaces can share a single datafile?
110. All datafiles related to a Tablespace are removed when the Tablespace is dropped?
111. What is a default role?
112. Who is the owner of a role?
113. When granting the system privilege, which clause enables the grantee to further grant the privilege to other users or roles?
114. Which view will show a list of privileges that are available for the current session to a user?
115. Which view shows all of the objects accessible to the user in a database?
116. Which statement about profiles is false?
117. Which password management feature is NOT available by using a profile?
118. Which resource can not be controlled using profiles?
119. You want to retrieve information about account expiration dates from the data dictionary. Which view do you use?
120. It is very difficult to grant and manage common privileges needed by different groups of database users using roles?
121. Which data dictionary view would you query to retrieve a table's header block number?
122. When tables are stored in locally managed tablespaces, where is extent allocation information stored?
123. Which of the following three portions of a data block are collectively called as Overhead?
124. Can a tablespace hold objects from different schemes?
125. Which data dictionary view would you query to retrieve a table's header block number?
126. What is default value for storage parameter INITIAL in 10g if extent management is Local?
127. Using which package we can convert Tablespace from DMTS to LMTS?
128. Is it Possible to Change ORACLE Block size after creating database?
129. Locally Managed table spaces will increase the performance?
130. Index is a Space demanding Object?
131. What is a potential reason for a Snapshot too old error message?
132. An Oracle user receives the following error? ORA-01555 SNAPSHOT TOO OLD, What is the possible solution?
133. The status of the Rollback segment can be viewed through?
134. Explicitly we can assign transaction to a rollback segment?
135. Are uncommitted transactions written to flashback redologs?

136. Is it possible to do flashback after truncate?
137. Can we restore a dropped table after a new table with the same name has been created?
138. Which following command will clear database recyclebin?
139. What is the OPTIMAL parameter?
140. Flashback query time depends on?
141. Can we create spfile in shutdown mode?
142. Can we alter static parameters by using scope=both?
143. Can we take backup of spfile in RMAN?
144. . Does Drop Database command removes spfile?
145. Using which SQL command we can alter the parameters?
146. OMF database will improve the performance?
147. Max number of controlfiles that can be multiplexed in an OMF database?
148. Which environment variable is used to help set up Oracle names?
149. Which Net8 component waits for incoming requests on the server side?
150. What is the listener name when you start the listener without specifying an argument?
151. When is a request sent to a listener?
152. In which file is the information that host naming is enabled stored?
153. Which protocols can oracle Net 11g Use?
154. Which of the following statements about listeners is correct?
155. Can we perform DML operation on Materialized view?
156. Materialized views are schema objects that can be used to summarize pre compute replicate and distribute data?
157. Does a materialized view occupy space?
158. Can we name a Materialized View log?
159. How to improve sqlldr (SQL*Loader) performance?
160. By using which view can a normal user see public database link?
161. Can we change the refresh interval of a Materialized View?
162. Can we use a database link even after the target user has changed his password?
163. Can we convert a materialized view from refresh fast to complete?
164. A normal user can create public database link?
165. If we truncate the master table, materialized view log on that table?
166. What is the correct procedure for multiplexing online redo logs?
167. In which situation would you need to create a new control file for an existing database?
168. When configuring a database for ARCHIVELOG mode, you use an initialisation parameter to specify which action?
169. Which command creates a text backup of the control file?
170. You are configuring a database for ARCHIVELOG mode. Which initialization parameter should you use?
171. How does a DBA specify multiple control files?
172. Which dynamic view should a DBA query to obtain information about the different sections of the control file?
173. What is the characteristic of the control file?
174. Which statements about online redo log members in a group are true?
175. Which command does a DBA use to list the current status of archiving?
176. When performing an open database backup, which statement is NOT true?
177. Which task can a DBA perform using the export/import facility?
178. Why does this command cause an error?
179. Which import option do you use to create tables without data?

180. Which export option will generate code to create an initial extent that is equal to the sum of the sizes of all the extents currently allocated to an object?
181. Can I take 1 dump file set from my source database and import it into multiple databases?
182. Can we export a dropped table?
183. What is the default value for IGNORE parameter in EXP/IMP?
184. Why is Direct Path Export Faster?
185. Is there a way to estimate the size of an export job before it gets underway?
186. Can I monitor a Data Pump Export or Import job while the job is in progress?
187. If a job is stopped either voluntarily or involuntarily, can I restart it?
188. Does Data Pump support Flashback?
189. If the tablespace is Read Only, Can we export objects from that tablespaces?
190. Dump files exported using traditional EXP are compatible with DATAPUMP?
191. Before a DBA creates a transportable tablespace, which condition must be completed?
192. Can we transport tablespace from one database to another database which is having SYS owned objects?
193. What is default value for TRANSPORT_TABLESPACE Parameter in EXP?
194. How to find whether tablespace is created in that database or transported from another database?
195. Can we Perform TTS using EXPDP?
196. Can we Transport Tablespace which has Materialized View in it?
197. When would a DBA need to perform a media recovery?
198. Why would you set a data file offline when the database is in MOUNT state?
199. What is the cause of media failures?
200. Which of the following would not require you to perform an incomplete recovery?
201. In what scenario you have to open a database with reset logs option?
202. Is it possible taking consistent backup if the database is in NOARCHIVELOG mode?
203. Database is in Archivelog mode and Loss of unbackedup datafile is?
204. You should issue a backup of the control file after issuing which command?
205. The alert log will never contain specific information about which database backup activity) ?
206. . A tablespace becomes unavailable because of a failure. The database is running in NOARCHIVELOG mode?
What should the DBA do to make the database available?
207. How often does a read-only tablespace need to be backed up?
208. With the instance down, how would you recover a lost control file?
209. Which action does Oracle recommend after a DBA recovers from the loss of the current online redo-log?
210. Which command creates a text backup of the control file?
211. Which option is used in the parameter file to detect corruptions in an Oracle data block?
212. Your database is configured in ARCHIVELOG mode. Which backups cannot be performed?
213. You are using hot backup without being in archivelog mode, can you recover in the event of a failure?
214. Which following statement is true when tablespaces are put in backup mode for hot backups?
215. Can Consistant Backup be performed when the database is open?
216. Can we shut down the database if it is in BEGIN BACKUP mode?
217. Which data dictionary view helps you to view whether tablespace is in BEGIN BACKUP Mode or not?
218. Which command is used to allow RMAN to store a group of commands in the recovery catalog?
219. When using Recovery Manager without a catalog, the connection to the target database?
220. Work is done by Recovery Manager through?
221. You perform an incomplete database recovery using RMAN. Which state of target database is needed?
222. Is it possible to perform Transportable tablespace (TTS) using RMAN?
223. Which type of file does Not RMAN include in its backups?
224. When using Recovery Manager without a catalog, the connection to the target database should be made as?
225. RMAN online backup generates excessive Redo information?

226. Which background process will be invoked when we enable BLOCK CHANGE TRACKING?
227. Where should a recovery catalog be created?
228. How to list restore points in RMAN?
229. Without LIST FAILURE can we say ADVISE FAILURE in Data Recovery Advisor?
230. Import Catalog Command is used for?
231. Interfile backup parallelism does?
232. What is the difference between pfile and spfile. Where these files are located?
233. What will you do if pfile and spfile file is deleted? Can you start the database?
234. What is the difference between Static and Dynamic init.ora/spfile parameters?
235. What is the complete syntax to set DB_CACHE_SIZE in memory and spfile?
236. How do we configure multiple Buffer Cache in Oracle. Whats the benefit? Does setting multiple Cache requires Database Restart?
237. What is Oracle Golden Gate?
238. Can we create Tablespaces of multiple Block Sizes. If yes, what is the Syntax?
239. How do you calculate the size of oracle memory areas Buffer Cache, Log Buffer, Shared Pool, PGA etc?
240. What is OMF? What spfile parameters are used to configure OMF. What is the benefit?
241. What is Database Cloning? Why Cloning is needed? What are the steps to clone a database?
242. What is Oracle Streams?
243. There are 2 control files for a database. What will happen when 1 control file is deleted and you try to start database? How you will fix this problem?
244. What is Dynamic performance view and What is Data Dictionary Views. Give some examples of each?
245. You are working in database that does lot of Sorting , i.e SELECT queries use a lot of ORDER BY and GROUP BY? What Oracle memory area and Physical File/Tablespace you need to tune and How?
246. Why we upgrade a database. What are the steps to upgrade database. Any errors you got during upgrade?
247. What is MEMORY_TARGET not supported error. How do you fix it?
248. What are the steps to manually create a database?
249. A DBA ran a delete statement to delete all records on a table. The table has 50 Million rows? While Delete is running his SQLPLUS session terminate abnormally? What oracle will do internally?
250. What is Oracle Dataguard?
251. Can we change the DB_BLOCK_SIZE? if Yes. What are the steps?
252. Explain the Oracle Architecture?
253. What happens internally in Oracle when a User Connects and run a SELECT Query? What SGA areas and background processes are involved?
254. How do you create a tablespace, undo tablespace and temp tablespace. What are the Syntax?
255. As a HR user you logged in and Creating a EMP_BIG Table and inserting 10 lac rows? While inserting 10 lac rows you got error ORA-01688: unable to extend table EMP_BIG by 512 in tablespace HR_DATA? What are the two ways to fix this Tablespace error?
256. What are the steps to rename a database?
257. What is the syntax to create a user and roles?
258. What is the 3 init.ora parameter to manage UNDO? What is their usage?
259. What is/are Snapshot too old error? How do you fix it?
260. What is Undo Retention Gurantee? How do we set it? What is the Proc and Cons of setting it?
261. What are System Privileges and Object Privileges? Give some examples? What Data Dictionary view we use to check both?
262. What is PGA? What information is stored in PGA? What is PGA Tuning?
263. What are the steps to identify a slow running SQL and tune it?
264. What is all the preparation works a DBA need to do before installing Oracle?
265. Any error that you got during Oracle Installation and how did you fixes it?

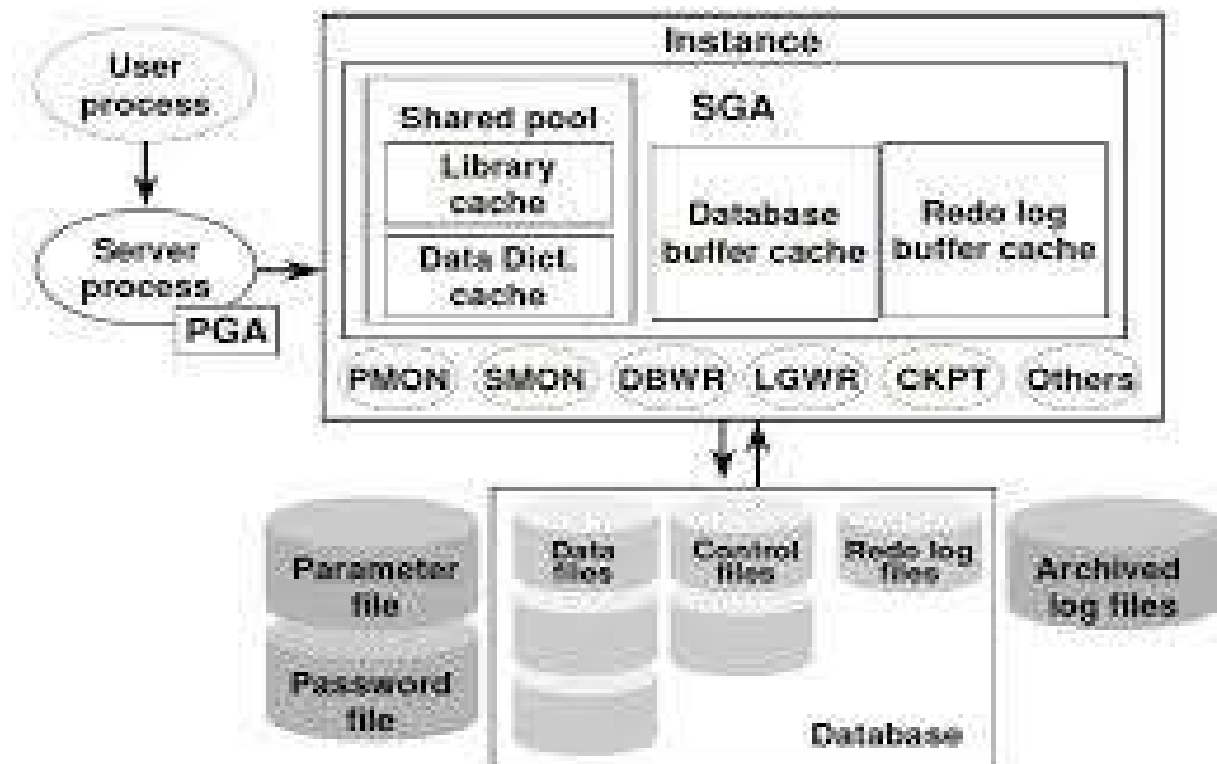
266. What is default tablespace and temporary tablespace?
267. Which privilege allows you to select from tables owned by other users?
268. What command we use to revoke system privilege?
269. How do we create a Role?
270. Difference between Non-Deferred and deferred constraint?
271. Difference between varchar and varchar2 data types?
272. In which language Oracle has been developed?
273. What is RAW datatype?
274. What is the use of NVL function?
275. Whether any commands are used for Months calculation? If so, what are they?
276. What are nested tables?
277. What is COALESCE function?
278. What is BLOB datatype?
279. How do we represent comments in Oracle?
280. What is DML?
281. What is the difference is between TRANSLATE and REPLACE?
282. How do we display rows from the table without duplicates?
283. What is the usage of Merge Statement?
284. What is NULL value in oracle?
285. What is USING Clause and give example?
286. What is key preserved table?
287. What is WITH CHECK OPTION?
288. What is the use of Aggregate functions in Oracle?
289. What do you mean by GROUP BY Clause?
290. What is a sub query and what are the different types of subqueries?
291. What is cross join?
292. What are temporal data types in Oracle?
293. How do we create privileges in Oracle?
294. What is VArray?
295. How do we get field details of a table?
296. What is the difference between rename and alias?
297. What is a View?
298. What is a cursor variable?
299. What are cursor attributes?
300. What are SET operators?
301. How can we delete duplicate rows in a table?
302. What are the attributes of Cursor?
303. Can we store pictures in the database and if so, how it can be done?
304. What is an integrity constraint?
305. What is an ALERT?
306. What is hash cluster?
307. What are the various constraints used in Oracle?
308. What is difference between SUBSTR and INSTR?
309. What is the parameter mode that can be passed to a procedure?
310. What are the different Oracle Database objects?
311. What are the differences between LOV and List Item?
312. What are privileges and Grants?
313. What is the difference between \$ORACLE_BASE and \$ORACLE_HOME?

314. What is the fastest query method to fetch data from the table?
315. What is the maximum number of triggers that can be applied to a single table?
316. How to display row numbers with the records?
317. How can we view last record added to a table?
318. What is the data type of DUAL table?
319. What is difference between Cartesian Join and Cross Join?
320. How to display employee records who gets more salary than the average salary in the department?
321. What is the difference between RMAN and a traditional hot backup?
322. What are bind variables and why are they important?
323. In PL/SQL, what is bulk binding, and when/how would it help performance?
324. Why is SQL*Loader direct path so fast?
325. What are the tradeoffs between many vs few indexes? When would you want to have many, and when would it be better to have fewer?
326. What is the difference between RAID 5 and RAID 10? Which is better for Oracle?
327. When using Oracle export/import what character set concerns might come up? How do you handle them?
328. Name three SQL operations that perform a SORT?
329. What is your favorite tool for day-to-day Oracle operation?
330. What is the difference between Truncate and Delete? Why is one faster? Can we ROLLBACK both? How would a full table scan behave after?
331. What is the difference between a materialized view (snapshot) fast refresh versus complete refresh? When is one better, and when the other?
332. What does the NO LOGGING option do? Why would we use it? Why would we be careful of using it?
333. Tell me about standby database? What are some of the configurations of it? What should we watch out for?
334. What do you know about privileges?

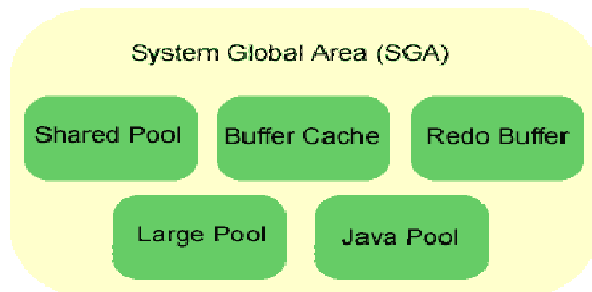
Answers

1. What is an instance? Draw Architecture?

SGA + [background processes](#).



2. What is SGA? Draw?



The System Global Area (SGA) is a group of shared memory areas that are dedicated to an Oracle “instance” (an instance is your database programs and RAM).

All Oracle processes use the SGA to hold information. The SGA is used to store incoming data (the data buffers as defined by the *db_cache_size* parameter), and internal control information that is needed by the database. You control the amount of memory to be allocated to the SGA by setting some of the Oracle “initialization parameters”. These might include *db_cache_size*, *shared_pool_size* and *log_buffer*.

In Oracle Database 10g you only need to define two parameters (*sga_target* and *sga_max_size*) to configure your SGA. If these parameters are configured, Oracle will calculate how much memory to allocate to the different areas of the SGA using a feature called Automatic Memory Management (AMM). As you gain experience you may want to manually allocate memory to each individual area of the SGA with the initialization parameters.

We have already noted that the SGA was sub-divided into several memory structures that each have different missions. The main areas contained in the SGA that you will be initially interested in have complicated names, but are actually quite simple:

- * The buffer cache (*db_cache_size*)
- * The shared pool (*shared_pool_size*)
- * The redo log buffer (*log_buffer*)

Let's look at these memory areas in more detail.

Note: AMM and dynamic Oracle memory management has measurable overhead.

Inside the Data Buffer Cache

The Buffer Cache (also called the database buffer cache) is where Oracle stores data blocks. With a few exceptions, any data coming in or going out of the database will pass through the buffer cache.

The total space in the Database Buffer Cache is sub-divided by Oracle into units of storage called "blocks". Blocks are the smallest unit of storage in Oracle and you control the data file blocksize when you allocate your database files.

An Oracle block is different from a disk block. An Oracle block is a logical construct -- a creation of Oracle, rather than the internal block size of the operating system. In other words, you provide Oracle with a big whiteboard, and Oracle takes pens and draws a bunch of boxes on the board that are all the same size. The whiteboard is the memory, and the boxes that Oracle creates are individual blocks in the memory.

Each block inside a file is determined by your `db_block_size` parameter and the size of your "default" blocks are defined when the database is created. You control the default database block size, and you can also define tablespaces with different block sizes. For example, many Oracle professionals place indexes in a 32k block size and leave the data files in a 16k block size.

Google: "oracle multiple block sizes"

When Oracle receives a request to retrieve data, it will first check the internal memory structures to see if the data is already in the buffer. This practice allows the server to avoid unnecessary I/O. In an ideal world, DBAs would be able to create one buffer for each database page, thereby ensuring that Oracle Server would read each block only once.

The `db_cache_size` and `shared_pool_size` parameters define most of the size of the in-memory region that Oracle consumes on startup and determine the amount of storage available to cache data blocks, SQL, and stored procedures.

Google: "oracle sga size"

The default size for the buffer pool (64k) is too small. We suggest you set this to a value of 1m when you configure Oracle.

The common components are:

Data buffer cache - cache data and index blocks for faster access.

Shared pool - cache parsed [SQL](#) and [PL/SQL](#) statements.

Dictionary Cache - information about data dictionary objects.

Redo Log Buffer - [committed](#) transactions that are not yet written to the redo log files.

JAVA pool - caching parsed [Java](#) programs.

Streams pool - cache [Oracle Streams](#) objects.

Large pool - used for [backups](#), [UGAs](#), etc.

Shared Pool:

The shared pool consists of the following areas:

Library cache includes the shared SQL area, private SQL areas, PL/SQL procedures and packages the control structures such as locks and library cache handles. Oracle code is first parsed, then executed, this parsed code is stored in the library cache, Oracle first checks the library cache to see if there is an already parsed and ready to execute form of the statement in there, if there is this will reduce CPU time considerably, this is called a **soft parse**, If Oracle has to parse it then this is called a **hard parse**. If there is not enough room in the cache Oracle will remove older parsed code, obviously it is better to keep as much parsed code in the library cache as possible. Keep an eye on missed cache hits which is an indication that a lot of hard parsing is going on.

Dictionary cache is a collection of database tables and views containing information about the database, its structures, privileges and users. When statements are issued Oracle will check permissions, access, etc and will obtain this information from its dictionary cache, if the information is not in the cache then it has to be read in from the disk and placed in to the cache. The more information held in the cache the less Oracle has to access the slow disks. The parameter `SHARED_POOL_SIZE` is used to determine the size of the shared pool, there is no way to adjust the caches independently, you can only adjust the shared pool size.

The shared pool uses a LRU (least recently used) list to maintain what is held in the buffer, see buffer cache for more details on the LRU.

You can clear down the shared pool area by using the following command

```
alter system flush shared_pool;
```

Buffer cache:

This area holds copies of read data blocks from the datafiles. The buffers in the cache contain two lists, the write list and the least used list (LRU). The write list holds dirty buffers which contain modified data not yet written to disk.

The LRU list has the following

- free buffers hold no useful data and can be reused
- pinned buffers actively being used by user sessions
- dirty buffers contain data that has been read from disk and modified but hasn't been written to disk

It's the database writers job to make sure that they are enough free buffers available to users session, if not then it will write out dirty buffers to disk to free up the cache.

There are 3 buffer caches

- Default buffer cache, which is everything not assigned to the keep or recycle buffer pools, DB_CACHE_SIZE
- Keep buffer cache which keeps the data in memory (goal is to keep warm/hot blocks in the pool for as long as possible), DB_KEEP_CACHE_SIZE.
- Recycle buffer cache which removes data immediately from the cache after use (goal here is to age out a blocks as soon as it is no longer needed), DB_RECYCLE_CACHE_SIZE.

The standard block size is determined by the DB_CACHE_SIZE, if tablespaces are created with a different block sizes then you must also create an entry to match that block size.

DB_2K_CACHE_SIZE (used with tablespace block size of 2k)

DB_4K_CACHE_SIZE (used with tablespace block size of 4k)

DB_8K_CACHE_SIZE (used with tablespace block size of 8k)

DB_16K_CACHE_SIZE (used with tablespace block size of 16k)

DB_32K_CACHE_SIZE (used with tablespace block size of 32k)

buffer cache hit ratio is used to determine if the buffer cache is sized correctly, the higher the value the more is being read from the cache.

hit rate = (1 - (physical reads / logical reads)) * 100

You can clear down the buffer pool area by using the following command

```
alter system flush buffer_cache;
```

Redo buffer:

The redo buffer is where data that needs to be written to the online redo logs will be cached temporarily before it is written to disk, this area is normally less than a couple of megabytes in size. These entries contain necessary information to reconstruct/redo changes by the INSERT, UPDATE, and DELETE, CREATE, ALTER and DROP commands.

The contents of this buffer are flushed:

- Every three seconds
- Whenever someone commits a transaction
- When its gets one third full or contains 1MB of cached redo log data.
- When LGWR is asked to switch logs

Use the parameter LOG_BUFFER parameter to adjust but be-careful increasing it too large as it will reduce your I/O but commits will take longer.

Large Pool:

This is an optional memory area that provide large areas of memory for:

- Shared Server - to allocate the UGA region in the SGA
- Parallel execution of statements - to allow for the allocation of inter-processing message buffers, used to coordinate the parallel query servers.
- Backup - for RMAN disk I/O buffers

The large pool is basically a non-cached version of the shared pool.

Use the parameter LARGE_POOL_SIZE parameter to adjust

Java Pool:

Used to execute java code within the database.

Use the parameter JAVA_POOL_SIZE parameter to adjust (default is 20MB)

Streams Pool:

Streams are used for enabling data sharing between databases or application environment.

Use the parameter STREAMS_POOL_SIZE parameter to adjust

3. What is PGA (or) what is pga_aggregate_target?



Session Information (runtime area):

PGA in an instance running with a shared server requires additional memory for the user's session, such as private SQL areas and other information.

Stack space (private sql area): The memory allocated to hold a sessions variables, arrays, etc and other information relating to the session.

Explanation-1:

Not all RAM in Oracle is shared memory. When you start a user process, that process has a private RAM area, used for sorting SQL results and managing special joins called "hash" joins. This private RAM is known as the Program Global Area (PGA). Each individual PGA memory area is allocated each time a new user connects to the database.

Google:"oracle pga sizing"

Oracle Database 10g will manage the PGA for you if you set the `pga_aggregate_target` parameter (we will discuss parameters and how they are set later in this book), but you can manually allocate the size of the PGA via parameters such as `sort_area_size` and `hash_area_size`. We recommend that you allow Oracle to configure these areas, and just configure the `pga_aggregate_target` parameter.

The PGA can be critical to performance, particularly if your application is doing a large number of sorts. Sort operations occur if you use `ORDER by` and `GROUP BY` commands in your SQL statements.

Explanation-2:

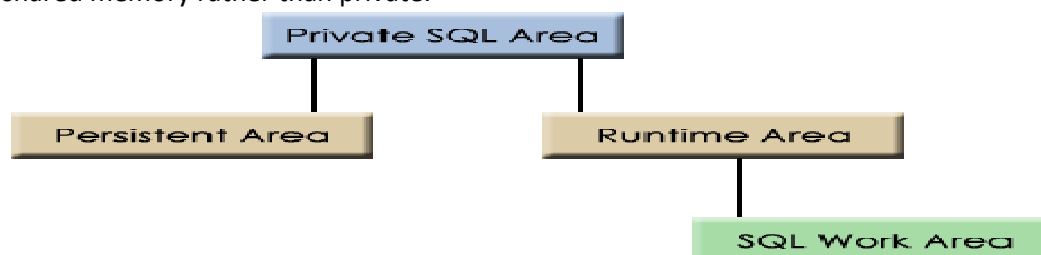
Apart from SGA (System Global Area), Oracle allocates global area specific and attached to each process and session. PGA stands for Process Global Area which is also known as Program Global Area. It is called global area because it keeps information which is required by all modules of Oracle Code. PGA keeps information specific to the server process upon which Oracle code acts. PGA keeps process specific information like operating system resources that a process is using, Oracle Shared Resources in SGA being used by a process, and stat information of the process. PGA also keeps information about oracle shared resources so that it can free those resources upon unexceptional death. PGA contains process specific information so it does not require any latch or locks to serialize the access. PGA contains other areas like UGA and CGA. Generally DBAs are unaware about these areas as separate and consider the information kept in these areas integrated into PGA.

Each session contains specific information like bind variables and runtime structures in a private SQL area. Whenever a session executes a statement, a private SQL area is assigned to that session. Even if multiple users are issuing the same statement using the same shared SQL area, each session will have its own dedicated private SQL area.

A private SQL area contains data such as bind information and runtime memory structures. Each user that submits the same SQL statement has his or her own private SQL area that uses a single shared SQL area. Thus, many private SQL areas can be associated with the same shared SQL area.

A private SQL area itself is divided into run-time area and persistent area. Persistent area contains information like bind variable and will be freed once the cursor is closed. Run-time area is allocated with the first step of the execute request and will be freed when execution is completed.

In a shared server private SQL area is allocated from the shared pool (if large pool is not configured), as it needs shared memory rather than private.



Automatic PGA Management: To reduce response times sorts should be performed in the PGA cache area (optimal mode operation), otherwise the sort will spill on to the disk (single-pass / multiple-pass operation) this will reduce

performance, so there is a direct relationship between the size of the PGA and query performance. You can manually tune the below to increase performance

- **sort_area_size** - total memory that will be used to sort information before swapping to disk
- **sort_area_retained_size** - memory that is used to retained data after a sort
- **hash_area_size** - memory that will would be used to store hash tables
- **bitmap_merge_area_size** - memory Oracle uses to merge bitmaps retrieved from a range scan of the index.

Starting with Oracle 9i there is a new to manage the above settings that is to let oracle manage the PGA area automatically by setting the parameter following parameters Oracle will automatically adjust the PGA area basic on users demand.

- **workarea_size_policy** - you can set this option to manual or auto (default)
- **pga_aggregate_target** - controls how much to allocate the PGA in total

Oracle will try and keep the PGA under the target value, but if you exceed this value Oracle will perform multi-pass operations (disk operations).

System Parameters	
workarea_size_policy	manual or auto (default)
pga_aggregate_target	total amount of memory allocated to the PGA
PGA/UGA amount used	select a.name, to_char(b.value, '999,999,999') value from v\$statname a, v\$mystat b where a.statistic# = b.statistic# and a.name like '%ga memory%';
Display if using memory or disk sorts	set autotrace traceonly statistics; set autotrace off;
Display background process PGA memory usage	select program, pga_used_mem, pga_alloc_mem, pga_max_mem from v\$process;

PGA and UGA

The PGA (Process Global Area) is a specific piece of memory that is associated with a single process or thread, it is not accessible by any other process or thread, note that each of Oracles background processes have a PGA area. The UGA (User Global Area) is your state information, this area of memory will be accessed by your current session, depending on the connection type (shared server) the UGA can be located in the SGA which is accessible by any one of the shared server processes, because a dedicated connection does not use shared servers the memory will be located in the PGA

Shared server - UGA will be part of the SGA

Dedicated server - UGA will be the PGA

Memory Area	Dedicated Server	Shared Server
Nature of session memory	Private	Shared
Location of the persistent area	PGA	SGA
Location of part of the runtime area for select statements	PGA	PGA
Location of the runtime area for DML/DDDL statements	PGA	PGA

Oracle creates a PGA area for each users session, this area holds data and control information, the PGA is exclusively used by the users session. Users cursors, sort operations are all stored in the PGA. The PGA is split in to two areas

4. What are new memory parameters in Oracle 10g?

SGA_TARGET and PGA_TARGET

SGA_TARGET:

SGA_TARGET specifies the total size of all SGA components. If the SGA_TARGET is set, then the following memory pools are automatically sized:

- Buffer cache (DB_CACHE_SIZE)
- Shared pool (SHARED_POOL_SIZE)
- Large pool (LARGE_POOL_SIZE)
- Java pool (JAVA_POOL_SIZE)

SGA_MAX_SIZE specifies the hard limit upto which the SGA_TARGET can dynamically grow. While executing DBCA, Oracle suggests that the estimated SGA_MAX_SIZE is to set aside 40% of memory. However, it should be set according to your requirement that depends on multiple factors such as no of concurrent users, volume of transactions and growth rate of database. Under normal operation, you can set the SGA_MAX_SIZE equals to the SGA_TARGET. Sometimes, we need to perform some extra-heavy batch processing jobs that leads to more SGA size. At this circumstance, you must have capability to adjust for peak loads. That is why, you set hard limit for your SGA_MAX_SIZE.

SGA_MAX_SIZE cannot be changed dynamically without bouncing the database whereas SGA_TARGET can be changed dynamically without bouncing the database.

If you try to modify SGA_MAX_SIZE dynamically, you will get an error of

ORA-02095: specified initialization parameter cannot be modified.

SGA_TARGET can never be greater than SGA_MAX_SIZE. If you try to set the SGA_TARGET to a value which is greater than that of SGA_MAX_SIZE, then Oracle will throw an error of

ORA-02097: parameter cannot be modified because specified value is invalid

ORA-00823: specified value of SGA_TARGET greater than SGA_MAX_SIZE.

If the SGA_MAX_SIZE is not set and the SGA_TARGET is set, then the SGA_MAX_SIZE takes the value of SGA_TARGET. If you set the SGA_MAX_SIZE greater than your server memory capacity and bounce the database, you will get an error of

ORA-27102 : out of memory

SVR4 Error : 12 : not enough space

Automatic memory management is introduced into Oracle 11g. This can be configured using a target memory size initialization parameter MEMORY_TARGET and a maximum memory size initialization parameter MEMORY_MAX_TARGET. Oracle Database then tunes to MEMORY_TARGET size, by distributing memory as needed between the system global area (SGA) and the instance program global area (instance PGA).

Relation between MEMORY_TARGET, SGA_TARGET and PGA_AGGREGATE_TARGET :

If **MEMORY_TARGET is set to a non-zero** value.

- If SGA_TARGET and PGA_AGGREGATE_TARGET are set, they will be considered the minimum values for the sizes of SGA and the PGA respectively. MEMORY_TARGET can take values from SGA_TARGET + PGA_AGGREGATE_TARGET to MEMORY_MAX_TARGET.
- If SGA_TARGET is set and PGA_AGGREGATE_TARGET is not set, we will still auto-tune both parameters. PGA_AGGREGATE_TARGET will be initialized to a value of (MEMORY_TARGET-SGA_TARGET).
- If PGA_AGGREGATE_TARGET is set and SGA_TARGET is not set, we will still auto-tune both parameters. SGA_TARGET will be initialized to a value of min(MEMORY_TARGET-PGA_AGGREGATE_TARGET, SGA_MAX_SIZE (if set by the user)) and will auto-tune subcomps.
- If neither is set, they will be auto-tuned without any minimum or default values. We will have a policy of distributing the total memory set by memory_target parameter in a fixed ratio to the the SGA and PGA during initialization. The policy is to give 60% for sga and 40% for PGA at startup.

If **MEMORY_TARGET is not set or set to set to 0 explicitly** (default value is 0 for 11g):

- If SGA_TARGET is set we will only auto-tune the sizes of the sub-components of the SGA. PGA will be autotuned independent of whether it is explicitly set or not. Though the whole SGA(SGA_TARGET) and the PGA(PGA_AGGREGATE_TARGET) will not be auto-tuned, i.e., will not grow or shrink automatically.
- If neither SGA_TARGET nor PGA_AGGREGATE_TARGET is set, we will follow the same policy as we have today; PGA will be auto-tuned and the SGA will not be auto-tuned and parameters for some of the sub-components will have to be set explicitly (for SGA_TARGET).
- If only MEMORY_MAX_TARGET is set, MEMORY_TARGET will default to 0 and we will not auto tune sga and pga. It will default to 10gR2 behavior within sga and pga.
- If sga_max_size is not user set, we will internally set it to MEMORY_MAX_TARGET.

In a text initialization parameter file, if you omit the line for MEMORY_MAX_TARGET and include a value for MEMORY_TARGET, the database automatically sets MEMORY_MAX_TARGET to the value of MEMORY_TARGET. If

you omit the line for MEMORY_TARGET and include a value for MEMORY_MAX_TARGET, the MEMORY_TARGET parameter defaults to zero. After startup, you can then dynamically change MEMORY_TARGET to a non-zero value, provided that it does not exceed the value of MEMORY_MAX_TARGET.

If you wish to monitor the decisions made by Automatic Memory Management following views can be useful

- V\$MEMORY_DYNAMIC_COMPONENTS has the current status of all memory components
- V\$MEMORY_RESIZE_OPS has a circular history buffer of the last 800 SGA resize requests

SGA_TARGET vs SGA_MAX_SIZE

SGA_MAX_SIZE

sga_max_size sets the maximum value for sga_target

If sga_max_size is less than the sum of db_cache_size + log_buffer + shared_pool_size + large_pool_size at initialization time, then the value of sga_max_size is ignored.

SGA_TARGET

This parameter is new with Oracle 10g. It specifies the total amount of SGA memory available to an instance. Setting this parameter makes Oracle distribute the available memory among various components - such as shared pool (for SQL and PL/SQL), Java pool, large_pool and buffer cache - as required.

This new feature is called Automatic Shared Memory Management. With ASMM, the parameters java_pool_size, shared_pool_size, large_pool_size and db_cache_size need not be specified explicitly anymore.

sga_target cannot be higher than sga_max_size.

SGA_TARGET is a database initialization parameter (introduced in Oracle 10g) that can be used for automatic SGA memory sizing.

Parameter description:

SGA_TARGET

Property Description

Parameter type Big integer

Syntax SGA_TARGET = integer [K | M | G]

Default value 0 (SGA autotuning is disabled)

Modifiable ALTER SYSTEM

Range of values 64 to operating system-dependent

Basic Yes

SGA_TARGET provides the following:

- Single parameter for total SGA size
 - Automatically sizes SGA components
 - Memory is transferred to where most needed
 - Uses workload information
 - Uses internal advisory predictions
 - STATISTICS_LEVEL must be set to TYPICAL
- By using one parameter we don't need to use all other SGA parameters like.
- DB_CACHE_SIZE (DEFAULT buffer pool)
 - SHARED_POOL_SIZE (Shared Pool)
 - LARGE_POOL_SIZE (Large Pool)
 - JAVA_POOL_SIZE (Java Pool)

SGA_TARGET And LOCK_SGA

SGA_TARGET is to tell how much memory Oracle can use for SGA.

LOCK_SGA is use to make sure that the contents from the SGA are not flushed, i.e data from the db buffer cache is not written back to disc. It is like to pin the contents of SGA.

SGA_TARGET is the minimum value of sga that is used on startup or the memory allocated to sga on startup.....

SGA_LOCK is a parameter that protect your sga to be pagged....

The lock_sga parameter is used to make the Oracle SGA region ineligible for swapping, effectively pinning the SGA RAM in memory. This technique is also known as "page fencing", using lock_sga=true to guarantee that SGA RAM is never sent to the swap disk during a page-out operation.

So, my question is what will be effect of "alter system flush" if LOCK_SGA is set to TRUE....

Logically the SGA can be considered one monolithic block of memory; Oracle knows what is in it but to the OS it is opaque. The entire SGA might be in memory or part of the SGA may be in memory and part may have been 'swapped' to disk by the OS.

In either case the OS does not know and does not care what is in the SGA. If it needs memory for other things it may

swap (page) part of large memory segments to disk and then if a 'memory' reference is made to a part that is on disk the OS will load it back into memory and may swap something else out to disk to make room for it.

LOCK_SGA ensures that all of the SGA is kept in memory and prevents any of it from being swapped to disk.

Flushing is an Oracle process that flushes the 'contents' of the SGA regardless of where the SGA is physically located.

The part in memory and any swapped parts will all be flushed. The flush process does not know, and does not care if all of the SGA is in memory or if part of it is swapped out.

They are two separate and distinct operations.

5. What are new memory parameters in Oracle 11g?

MEMORY_TARGET:

MEMORY_TARGET specifies the Oracle system-wide usable memory. The database tunes memory to the MEMORY_TARGET value, reducing or enlarging the SGA and PGA as needed.

In a text-based initialization parameter file, if you omit MEMORY_MAX_TARGET and include a value for MEMORY_TARGET, then the database automatically sets MEMORY_MAX_TARGET to the value of MEMORY_TARGET.

If you omit the line for MEMORY_TARGET and include a value for MEMORY_MAX_TARGET, the MEMORY_TARGET parameter defaults to zero. After startup, you can then dynamically change MEMORY_TARGET to a nonzero value, provided that it does not exceed the value of MEMORY_MAX_TARGET.

MEMORY_TARGET and MEMORY_MAX_TARGET

The Oracle documents state the following:

MEMORY_TARGET specifies the Oracle system-wide usable memory.

MEMORY_MAX_TARGET (...) decide on a maximum amount of memory that you would want to allocate to the database for the foreseeable future.

So my guess is, MEMORY_MAX_TARGET (static) is the maximum you can set MEMORY_TARGET (dynamic) to. A couple of days ago, I wanted to experiment a bit with these memory settings.

My [Oracle Enterprise Linux](#) (5.5) machine was set for MEMORY_MAX_TARGET=512M and MEMORY_TARGET=256M, but after starting the database, it showed the following:

```
SQL> startup pfile=init.ora
```

ORACLE instance started.

Total System Global Area 534462464 bytes

Fixed Size 2215064 bytes

Variable Size 473957224 bytes

Database Buffers 50331648 bytes

Redo Buffers 7958528 bytes

Database mounted.

Database opened.

Total SGA, 534462464 bytes? That's about 510M, certainly not what I had specified for MEMORY_TARGET...!?

Checking SGA in Enterprise Manager (yes, I use it sometimes), it showed 256M allocated for MEMORY_TARGET, containing SGA and PGA:

Memory Advisors

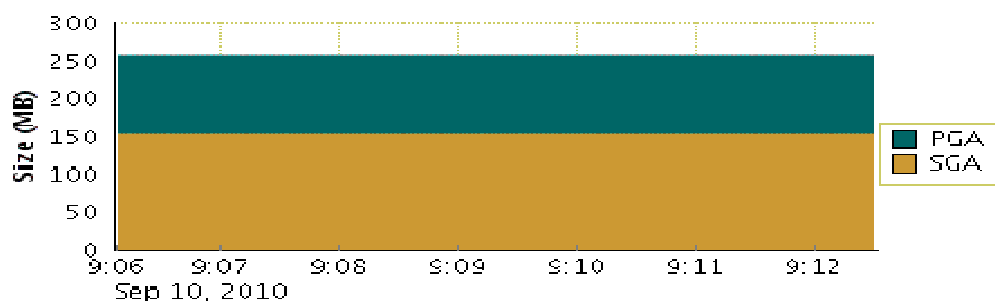
Automatic Memory Management **Enabled**

Total Memory Size

Maximum Memory Size

Allocation History

This chart shows the history of the components of the Memory.



AMM SGA/PGA sizes

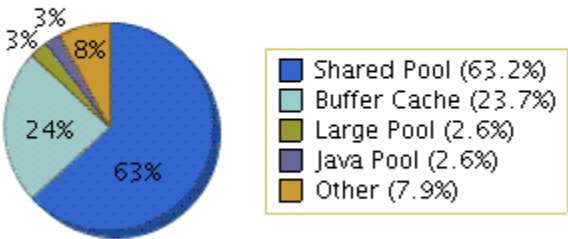
SGA was using 152M and PGA took the rest:

Current Allocation

Automatic Shared Memory Management Enabled

Total SGA Size (MB) 152

SGA Component	Current Allocation (MB)
Shared Pool	96
Buffer Cache	36
Large Pool	4
Java Pool	4
Other	12



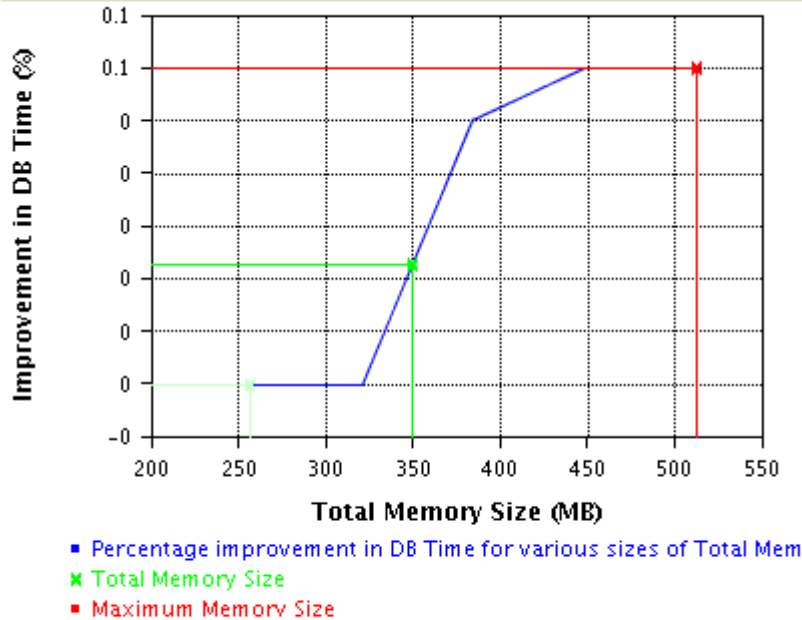
SGA size and contents

Also running 'select sum(bytes) from v\$sgastat' showed me the SGA is taking 152M. It seems 'show sga' shows the MEMORY_MAX_TARGET and 'Variable Size' included the memory it will not use.

Automatic Memory Memory Advisor

Oracle keeps track of memory usage and is able to advice about the MEMORY_TARGET size:

Memory Size Advice



Total Memory Size (MB) 349

You can click on the curve in the graph to set a new value. Total Memory Size cannot be greater than the Maximum Memory Size. First modify the Maximum Memory size (from the parent page) and then select a value of Total Memory up to the Maximum Memory size.

AMM Advice

Clicking the graph will update the MEMORY_TARGET parameter. One can also query V\$MEMORY_TARGET_ADVICE for this information:
MEMORY_SIZE MEMORY_SIZE_FACTOR ESTD_DB_TIME ESTD_DB_TIME_FACTOR VERSION

256	1	501	1	0
320	1.25	501	1	0
384	1.5	501	.9995	0
448	1.75	501	.9994	0
512	2	501	.9994	0

What is /dev/shm?

It is an in-memory mounted file system (tmpfs) and is very fast, but non-persistent when Linux is rebooted.

In Oracle 11g, it is used to hold SGA memory by storing the SGA structures in files with the same **granule size**. This granule size comes in 4M and 16M flavours, depending the MEMORY_MAX_TARGET smaller or larger than 1G. When these MEMORY_TARGET and MEMORY_MAX_TARGET parameters are set, oracle will create as much as $=(\text{MEMORY_MAX_TARGET} / \text{granule size})$ files. For instance, when MEMORY_MAX_TARGET set to 512M, it will create $512/4 = 128$ files (actually 129, the sneaky...).

The output of 'ls -la /dev/shm', will show you that not all the 128 files are taking the 4M of space:

```
shm> ls -la
total 151780
drwxrwxrwt 2 root root    2620 Sep 10 11:13 .
drwxr-xr-x 12 root root    3880 Sep 10 08:47 ..
-rw-rw---- 1 oracle oinstall 4194304 Sep 10 14:17 ora_ianh_3768323_0
-rw-rw---- 1 oracle oinstall 4194304 Sep 10 14:11 ora_ianh_3768323_1
-rw-rw---- 1 oracle oinstall    0 Sep 10 11:13 ora_ianh_3768323_10
-rw-rw---- 1 oracle oinstall 4194304 Sep 10 14:17 ora_ianh_3768323_100
(...)
-rw-rw---- 1 oracle oinstall 4194304 Sep 10 14:17 ora_ianh_3768323_127
-rw-rw---- 1 oracle oinstall 4194304 Sep 10 11:13 ora_ianh_3768323_128
-rw-rw---- 1 oracle oinstall    0 Sep 10 11:13 ora_ianh_3768323_13
-rw-rw---- 1 oracle oinstall    0 Sep 10 11:13 ora_ianh_3768323_14
-rw-rw---- 1 oracle oinstall    0 Sep 10 11:13 ora_ianh_3768323_15
-rw-rw---- 1 oracle oinstall    0 Sep 10 11:13 ora_ianh_3768323_16
-rw-rw---- 1 oracle oinstall    0 Sep 10 11:13 ora_ianh_3768323_17
-rw-rw---- 1 oracle oinstall    0 Sep 10 11:13 ora_ianh_3768323_18
-rw-rw---- 1 oracle oinstall    0 Sep 10 11:13 ora_ianh_3768323_19
-rw-rw---- 1 oracle oinstall 4194304 Sep 10 11:13 ora_ianh_3768323_2
```

Now this is the trick Oracle is using. When you add up all the files that do take 4M of space, it will never take more space than MEMORY_TARGET. Therefore, Oracle does not allocate more memory than the MEMORY_TARGET and the sum of these files might even be smaller than MEMORY_TARGET.

When you look at the SGA memory size using 'select ceil(sum(bytes)/(1024*1024*4)) from v\$sgastat', you will see it is near the sum of the files in /dev/shm (again, plus one...).

0 bytes in memory

When a file in /dev/shm is 0 bytes, it does not use memory. That memory is 'free' to other applications. Now this is Oracle's implementation of releasing memory back to the Linux OS, by cleaning up one or more of these in-memory files (will it do a 'cat /dev/null > ora_sid_number_id' ?).

Funny thing is, PGA is not stored in shared memory, because this is private memory. MEMORY_MAX_TARGET (used for SGA and PGA) is 'allocated' in /dev/shm, but PGA is not stored in /dev/shm. This means, when memory for PGA is allocated (and/or pga_aggregate_target is set), not all files in /dev/shm will get used!

Increase /dev/shm

If you increase the MEMORY_MAX_TARGET above the available /dev/shm space (df -h), you will receive:

ORA-00845: MEMORY_TARGET not supported on this system

If you have enough memory on your Linux machine, but /dev/shm is mounted to small by default, one can increase this amount of memory by changing /etc/fstab for permanent changes. The default is half of your physical RAM without swap.

For temporary changes to at least start the database, execute the following (change the 1500m to your environment):

```
> umount tmpfs
> mount -t tmpfs shmfs -o size=1500m /dev/shm
```

152M boundary

When I was playing around with these settings, it seems 152M is an initial minimal memory target.

If you start oracle with a pfile setting of lower than 152M, it fails to start and you will get the following message:

ORA-00838: Specified value of MEMORY_TARGET is too small, needs to be at least 152M

Remarks

- When I changed MEMORY_TARGET to 152M in my pfile, after the bounce the PGA was set to Manual Mode.

- Oracle will divide SGA/PGA as 60/40% when enough memory is available.
- The PGA_AGGREGATE_TARGET and SGA_TARGET are not ignored, but act as a minimum when set.
- When SGA_MAX_SIZE is set, it will act as a maximum; when it's not set it will show the MEMORY_MAX_TARGET value.
- /dev/shm must be mounted for at least 384M bytes (You are trying to use the MEMORY_TARGET feature. This feature requires the /dev/shm file system to be mounted for at least 402653184 bytes).

Conclusion

With Automatic Memory Management, one can set the upper limit of the total SGA and PGA to use. It is using an in-memory file structure, so it can give back unused memory to the Linux OS, unlike 10g, setting SGA_MAX_TARGET will just use all the memory specified.

On the other hand, when problems arise, one still needs to dive into the memory structures and tune. The 'automatic' feature added is memory distribution between SGA and PGA, and Oracle and OS.

6. What are the mandatory background processes?

DBWR LGWR SMON PMON CKPT RECO. (<http://satya-dba.blogspot.in/2009/08/background-processes-in-oracle.html>)

Background Processes in oracle

To maximize performance and accommodate many users, a multiprocess Oracle database system uses background processes. Background processes are the processes running behind the scene and are meant to perform certain maintenance activities or to deal with abnormal conditions arising in the instance. Each background process is meant for a specific purpose and its role is well defined.

Background processes consolidate functions that would otherwise be handled by multiple database programs running for each user process. Background processes asynchronously perform I/O and monitor other Oracle database processes to provide increased parallelism for better performance and reliability.

A background process is defined as any process that is listed in V\$PROCESS and has a non-null value in the pname column.

Not all background processes are mandatory for an instance. Some are mandatory and some are optional. Mandatory background processes are DBWn, LGWR, CKPT, SMON, PMON, and RECO. All other processes are optional, will be invoked if that particular feature is activated.

Oracle background processes are visible as separate operating system processes in Unix/Linux. In Windows, these run as separate threads within the same service. Any issues related to background processes should be monitored and analyzed from the trace files generated and the alert log.

Background processes are started automatically when the instance is started.

To find out background processes from database:

```
SQL> select SID,PROGRAM from v$session where TYPE='BACKGROUND';
```

To find out background processes from OS:

```
$ ps -ef | grep ora_ | grep SID
```

Mandatory Background Processes in Oracle

If any one of these 6 mandatory background processes is killed/not running, the instance will be aborted.

Database Writer (maximum 20):

Whenever a log switch is occurring as redo log file is becoming CURRENT to ACTIVE stage, oracle calls DBWn and synchronizes all the dirty blocks in database buffer cache to the respective datafiles, scattered or randomly.

Database writer (or Dirty Buffer Writer) process does multi-block writing to disk asynchronously. One DBWn process is adequate for most systems. Multiple database writers can be configured by initialization parameter DB_WRITER_PROCESSES, depends on the number of CPUs allocated to the instance. To have more than one DBWn only make sense if each DBWn has been allocated its own list of blocks to write to disk. This is done through the initialization parameter DB_BLOCK_LRU_LATCHES. If this parameter is not set correctly, multiple DB writers can end up contending for the same block list.

The possible multiple DBWR processes in RAC must be coordinated through the locking and global cache processes to ensure efficient processing is accomplished.

DBWn will be invoked in following scenarios:

- When the dirty blocks in SGA reaches to a threshold value, oracle calls DBWn.
- When the database is shutting down with some dirty blocks in the SGA, then oracle calls DBWn.
- DBWn has a time out value (3 seconds by default) and it wakes up whether there are any dirty blocks or not.
- When a checkpoint is issued.
- When a server process cannot find a clean reusable buffer after scanning a threshold number of buffers.

- When a huge table wants to enter into SGA and oracle could not find enough free space where it decides to flush out LRU blocks and which happens to be dirty blocks. Before flushing out the dirty blocks, oracle calls DBWn.
- [Oracle RAC](#) ping request is made.
- When Table DROPPed or TRUNCATEd.
- When tablespace is going to OFFLINE/READ ONLY/BEGIN BACKUP.

Log Writer (maximum 1) LGWR

LGWR writes redo data from redolog buffers to (online) redolog files, sequentially.

Redolog file contains changes to any datafile. The content of the redolog file is file id, block id and new content.

LGWR will be invoked more often than DBWn as log files are really small when compared to datafiles (KB vs GB). For every small update we don't want to open huge gigabytes of datafiles, instead write to the log file.

Redolog file has three stages CURRENT, ACTIVE, INACTIVE and this is a cyclic process. Newly created redolog file will be in UNUSED state.

When the LGWR is writing to a particular redolog file, that file is said to be in CURRENT status. If the file is filled up completely then a log switch takes place and the LGWR starts writing to the second file (this is the reason every database requires a minimum of 2 redolog groups). The file which is filled up now becomes from CURRENT to ACTIVE. Log writer will write synchronously to the redolog groups in a circular fashion. If any damage is identified with a redolog file, the log writer will log an error in the LGWR trace file and the alert log. Sometimes, when additional redolog buffer space is required, the LGWR will even write uncommitted redolog entries to release the held buffers. LGWR can also use group commits (multiple committed transaction's redo entries taken together) to write to redologs when a database is undergoing heavy write operations.

In RAC, each RAC instance has its own LGWR process that maintains that instance's thread of redo logs.

- LGWR will be invoked in following scenarios:
- LGWR is invoked whenever 1/3rd of the redo buffer is filled up.
- Whenever the log writer times out (3sec).
- Whenever 1MB of redolog buffer is filled (This means that there is no sense in making the redolog buffer more than 3MB).
- Shutting down the database.
- Whenever checkpoint event occurs.
- When a transaction is completed (either committed or rolled back) then oracle calls the LGWR and synchronizes the log buffers to the redolog files and then only passes on the acknowledgement back to the user. Which means the transaction is not guaranteed although we said commit, unless we receive the acknowledgement. When a transaction is committed, a System Change Number (SCN) is generated and tagged to it. Log writer puts a commit record in the redolog buffer and writes it to disk immediately along with the transaction's redo entries. Changes to actual data blocks are deferred until a convenient time (Fast-Commit mechanism).
- When DBWn signals the writing of redo records to disk. All redo records associated with changes in the block buffers must be written to disk first (The write-ahead protocol). While writing dirty buffers, if the DBWn process finds that some redo information has not been written, it signals the LGWR to write the information and waits until the control is returned.

Checkpoint (maximum 1) CKPT

Checkpoint is a background process which triggers the checkpoint event, to synchronize all database files with the checkpoint information. It ensures data consistency and faster database recovery in case of a crash.

When checkpoint occurred it will invoke the DBWn and updates the SCN block of the all datafiles and the control file with the current SCN. This is done by LGWR. This SCN is called checkpoint SCN.

Checkpoint event can be occurred in following conditions:

- Whenever database buffer cache filled up.
- Whenever times out (3seconds until 9i, 1second from 10g).
- Log switch occurred.
- Whenever manual log switch is done.
SQL> ALTER SYSTEM SWITCH LOGFILE;
- Manual checkpoint.
SQL> ALTER SYSTEM CHECKPOINT;
- Graceful shutdown of the database.

- Whenever BEGIN BACKUP command is issued.
- When the time specified by the initialization parameter LOG_CHECKPOINT_TIMEOUT (in seconds), exists between the incremental checkpoint and the tail of the log.
- When the number of OS blocks specified by the initialization parameter LOG_CHECKPOINT_INTERVAL, exists between the incremental checkpoint and the tail of the log.
- The number of buffers specified by the initialization parameter FAST_START_IO_TARGET required to perform roll-forward is reached.
- [Oracle 9i](#) onwards, the time specified by the initialization parameter FAST_START_MTTR_TARGET (in seconds) is reached and specifies the time required for a crash recovery. The parameter FAST_START_MTTR_TARGET replaces LOG_CHECKPOINT_INTERVAL and FAST_START_IO_TARGET, but these parameters can still be used.

System Monitor (maximum 1) SMON

If the database is crashed (power failure) and next time when we restart the database SMON observes that last time the database was not shutdown gracefully. Hence it requires some recovery, which is known as INSTANCE CRASH RECOVERY. When performing the crash recovery before the database is completely open, if it finds any transaction committed but not found in the datafiles, will now be applied from redolog files to datafiles.

If SMON observes some uncommitted transaction which has already updated the table in the datafile, is going to be treated as a in doubt transaction and will be rolled back with the help of before image available in [rollback segments](#). SMON also cleans up [temporary segments](#) that are no longer in use.

It also coalesces contiguous free extents in dictionary managed tablespaces that have PCTINCREASE set to a non-zero value.

In [RAC environment](#), the SMON process of one instance can perform instance recovery for other instances that have failed.

SMON wakes up about every 5 minutes to perform housekeeping activities.

5) Process Monitor (maximum 1) PMON

If a client has an open transaction which is no longer active (client session is closed) then PMON comes into the picture and that transaction becomes in doubt transaction which will be rolled back.

PMON is responsible for performing recovery if a user process fails. It will rollback uncommitted transactions. If the old session locked any resources that will be unlocked by PMON.

PMON is responsible for cleaning up the database buffer cache and freeing resources that were allocated to a process.

PMON also registers information about the instance and dispatcher processes with Oracle (network) listener.

PMON also checks the dispatcher & server processes and restarts them if they have failed.

PMON wakes up every 3 seconds to perform housekeeping activities.

In [RAC](#), PMON's role as service registration agent is particularly important.

Recoverer (maximum 1) RECO [Mandatory from Oracle 10g]

This process is intended for recovery in distributed databases. The distributed transaction recovery process finds pending distributed transactions and resolves them. All in-doubt transactions are recovered by this process in the distributed database setup. RECO will connect to the remote database to resolve pending transactions.

Pending distributed transactions are two-phase commit transactions involving multiple databases. The database that the transaction started is normally the coordinator. It will send request to other databases involved in two-phase commit if they are ready to commit. If a negative request is received from one of the other sites, the entire transaction will be rolled back. Otherwise, the distributed transaction will be committed on all sites. However, there is a chance that an error (network related or otherwise) causes the two-phase commit transaction to be left in pending state (i.e. not committed or rolled back). It's the role of the RECO process to liaise with the coordinator to resolve the pending two-phase commit transaction. RECO will either commit or rollback this transaction.

7. What are the optional background processes?

ARCH, MMAN, MMNL, MMON, CTWR, ASMB, RBAL, ARBx etc

Optional Background Processes in Oracle

Archiver (maximum 10) ARC0-ARC9

The ARCN process is responsible for writing the online redolog files to the mentioned archive log destination after a log switch has occurred. ARCN is present only if the database is running in archivelog mode and automatic archiving is enabled. The log writer process is responsible for starting multiple ARCN processes when the workload increases. Unless ARCN completes the copying of a redolog file, it is not released to log writer for overwriting.

The number of archiver processes that can be invoked initially is specified by the initialization parameter LOG_ARCHIVE_MAX_PROCESSES (by default 2, max 10). The actual number of archiver processes in use may vary based on the workload.

ARCH processes, running on primary database, select archived redo logs and send them to standby database. Archive log files are used for media recovery (in case of a hard disk failure and for maintaining an Oracle standby database via log shipping). Archives the standby redo logs applied by the managed recovery process (MRP).

In RAC, the various ARCH processes can be utilized to ensure that copies of the archived redo logs for each instance are available to the other instances in the RAC setup should they be needed for recovery.

Coordinated Job Queue Processes (maximum 1000) CJQ0/Jnnn

Job queue processes carry out batch processing. All scheduled jobs are executed by these processes. The initialization parameter JOB_QUEUE_PROCESSES specifies the maximum job processes that can be run concurrently. These processes will be useful in refreshing [materialized views](#).

This is the Oracle's dynamic job queue coordinator. It periodically selects jobs (from JOB\$) that need to be run, scheduled by the Oracle job queue. The coordinator process dynamically spawns job queue slave processes (J000-J999) to run the jobs. These jobs could be PL/SQL statements or procedures on an Oracle instance.

CQJO - Job queue controller process wakes up periodically and checks the job log. If a job is due, it spawns Jnnnn processes to handle jobs.

From [Oracle 11g release2](#), DBMS_JOB and DBMS_SCHEDULER work without setting JOB_QUEUE_PROCESSES. Prior to 11gR2 the default value is 0, and from 11gR2 the default value is 1000.

Dedicated Server

Dedicated server processes are used when MTS is not used. Each user process gets a dedicated connection to the database. These user processes also handle disk reads from database datafiles into the database block buffers.

LISTENER

The LISTENER process listens for connection requests on a specified port and passes these requests to either a distributor process if MTS is configured or to a dedicated process if MTS is not used. The LISTENER process is responsible for load balance and failover in case a RAC instance fails or is overloaded.

CALLOUT Listener

Used by internal processes to make calls to externally stored procedures

Lock Monitor (maximum 1) LMON

Lock monitor manages global locks and resources. It handles the redistribution of instance locks whenever instances are started or shutdown. Lock monitor also recovers instance lock information prior to the instance recovery process. Lock monitor co-ordinates with the Process Monitor (PMON) to recover dead processes that hold instance locks.

Lock Manager Daemon (maximum 10) LMDn

LMDn processes manage instance locks that are used to share resources between instances. LMDn processes also handle deadlock detection and remote lock requests.

Global Cache Service (LMS)

In an [Oracle Real Application Clusters](#) environment, this process manages resources and provides inter-instance resource control.

Lock processes (maximum 10) LCK0- LCK9

The instance locks that are used to share resources between instances are held by the lock processes.

Block Server Process (maximum 10) BSP0-BSP9

Block server Processes have to do with providing a consistent read image of a buffer that is requested by a process of another instance, in certain circumstances.

Queue Monitor (maximum 10) QMN0-QMN9

This is the advanced queuing time manager process. QMNn monitors the message queues. QMN used to manage Oracle Streams Advanced Queuing.

Event Monitor (maximum 1) EMN0/EMON

This process is also related to advanced queuing, and is meant for allowing a publish/subscribe style of messaging between applications.

Dispatcher (maximum 1000) Dnnn

Intended for multi threaded server (MTS) setups. Dispatcher processes listen to and receive requests from connected sessions and places them in the request queue for further processing. Dispatcher processes also pickup outgoing responses from the result queue and transmit them back to the clients. Dnnn are mediators between the client processes and the shared server processes. The maximum number of dispatcher process can be specified using the initialization parameter MAX_DISPATCHERS.

Shared Server Processes (maximum 1000) Snnn

Intended for multi threaded server (MTS) setups. These processes pickup requests from the call request queue, process them and then return the results to a result queue. These user processes also handle disk reads from database datafiles into the database block buffers. The number of shared server processes to be created at instance startup can be specified using the initialization parameter SHARED_SERVERS. Maximum shared server processes can be specified by MAX_SHARED_SERVERS.

Parallel Execution/Query Slaves (maximum 1000) Pnnn

These processes are used for parallel processing. It can be used for parallel execution of SQL statements or recovery. The Maximum number of parallel processes that can be invoked is specified by the initialization parameter PARALLEL_MAX_SERVERS.

Trace Writer (maximum 1) TRWR

Trace writer writes trace files from an Oracle internal tracing facility.

Input/Output Slaves (maximum 1000) Innn

These processes are used to simulate asynchronous I/O on platforms that do not support it. The initialization parameter DBWR_IO_SLAVES is set for this purpose.

Data Guard Monitor (maximum 1) DMON

The Data Guard broker process. DMON is started when Data Guard is started. This is broker controller process is the main broker process and is responsible for coordinating all broker actions as well as maintaining the broker configuration files. This process is enabled/disabled with the DG_BROKER_START parameter.

Data Guard Broker Resource Manager RSMO

The RSM process is responsible for handling any SQL commands used by the broker that need to be executed on one of the databases in the configuration.

Data Guard NetServer/NetSlave NSVn

These are responsible for making contact with the remote database and sending across any work items to the remote database. From 1 to n of these network server processes can exist. NSVn is created when a Data Guard broker configuration is enabled. There can be as many NSVn processes (where n is 0- 9 and A-U) created as there are databases in the Data Guard broker configuration.

DRCn

These network receiver processes establish the connection from the source database NSVn process. When the broker needs to send something (e.g. data or SQL) between databases, it uses this NSV to DRC connection. These connections are started as needed.

Data Guard Broker Instance Slave Process INSV

Performs Data Guard broker communication among instances in an Oracle RAC environment

Data Guard Broker Fast Start Failover Pinger Process FSFP

Maintains fast-start failover state between the primary and target standby databases. FSFP is created when fast-start failover is enabled.

LGWR Network Server process LNS

In Data Guard, LNS process performs actual network I/O and waits for each network I/O to complete. Each LNS has a user configurable buffer that is used to accept outbound redo data from the LGWR process. The NET_TIMEOUT attribute is used only when the LGWR process transmits redo data using a LGWR Network Server(LNS) process.

Managed Recovery Process MRP

In Data Guard environment, this managed recovery process will apply archived redo logs to the standby database.

Remote File Server process RFS

The remote file server process, in Data Guard environment, on the standby database receives archived redo logs from the primary database.

Logical Standby Process LSP

The logical standby process is the coordinator process for a set of processes that concurrently read, prepare, build, analyze, and apply completed SQL transactions from the archived redo logs. The LSP also maintains metadata in the database. The RFS process communicates with the logical standby process (LSP) to coordinate and record which files arrived.

Wakeup Monitor Process (maximum 1) WMON

This process was available in older versions of Oracle to alarm other processes that are suspended while waiting for an event to occur. This process is obsolete and has been removed.

Recovery Writer (maximum 1) RVWR

This is responsible for writing [flashback](#) logs (to [FRA](#)).

Fetch Archive Log (FAL) Server

Services requests for archive redo logs from FAL clients running on multiple standby databases. Multiple FAL servers can be run on a primary database, one for each FAL request.

Fetch Archive Log (FAL) Client

Pulls archived redo log files from the primary site. Initiates transfer of archived redo logs when it detects a gap sequence.

Data Pump Master Process DMnn

Creates and deletes the master table at the time of export and import. Master table contains the job state and object information. Coordinates the Data Pump job tasks performed by Data Pump worker processes and handles client interactions. The Data Pump master (control) process is started during job creation and coordinates all tasks performed by the Data Pump job. It handles all client interactions and communication, establishes all job contexts, and coordinates all worker process activities on behalf of the job. Creates the Worker Process

Data Pump Worker Process DWnn

It performs the actual heavy duty work of loading and unloading of data. It maintains the information in master table. The Data Pump worker process is responsible for performing tasks that are assigned by the Data Pump master process, such as the loading and unloading of metadata and data.

Shadow Process

When client logs in to an Oracle Server the database creates an Oracle process to service Data Pump API.

Client Process

The client process calls the [Data pump](#) API.

ARbx is configured by ASM_POWER_LIMIT.

8. What are the new background processes in Oracle 10g?

MMAN, MMON, MMNL, CTWR, ASMB, RBAL and ARbx

New Background Processes in Oracle 10g

Memory Manager (maximum 1) MMAN

MMAN dynamically adjust the sizes of the SGA components like buffer cache, large pool, shared pool and java pool and serves as SGA memory broker. It is a new process added to Oracle 10g as part of automatic shared memory management.

Memory Monitor (maximum 1) MMON

MMON monitors SGA and performs various manageability related background tasks. MMON, the Oracle 10g background process, used to collect statistics for the Automatic Workload Repository (AWR).

Memory Monitor Light (maximum 1) MMNL

New background process in Oracle 10g. This process performs frequent and lightweight manageability-related tasks, such as session history capture and metrics computation.

Change Tracking Writer (maximum 1) CTWR

CTWR will be useful in RMAN. Optimized incremental backups using [block change tracking](#) (faster incremental backups) using a file (named block change tracking file). CTWR (Change Tracking Writer) is the background process responsible for tracking the blocks.

ASMB

This ASMB process is used to provide information to and from cluster synchronization services used by [ASM](#) to manage the disk resources. It's also used to update [statistics](#) and provide a heart beat mechanism.

Re-Balance RBAL

RBAL is the ASM related process that performs rebalancing of disk resources controlled by [ASM](#).

Actual Rebalance ARbx

9. How do you use automatic PGA memory management with Oracle 9i and above?

Set the WORKAREA_SIZE_POLICY parameter to AUTO and set PGA_AGGREGATE_TARGET

[Explanation:](#)

Automated PGA Memory Management:

There are two different memory types in the Oracle PGA: not tunable and tunable. To configure the tunable are, there are several database parameters that can be used. These include sort_area_size, hash_area_size, bitmap_merge_area_size, and create_bitmap_area_size. In Oracle8i, you could set these parameters dynamically. However, it was difficult to tune them well. More memory was often allocated to a given session than was really needed. This results in wasted memory.

In 10G, PGA can be configured by setting the PGA_AGGREGATE_TARGET initialization parameter. To instruct the Oracle Database whether to tune PGA automatically, one needs to set WORKAREA_SIZE_POLICY to AUTO. If the value

of this parameter is set to MANUAL, that means work area size will be based on *_AREA_SIZE parameters like SORT_AREA_SIZE and HASH_AREA_SIZE. Note that this is not recommended in 10g. At any given time, the amount of memory available for the active work area is derived from the PGA_AGGREGATE_TARGET value. The value is set to PGA_AGGREGATE_TARGET - memory allocated for PGA by other sessions. Under automatic PGA memory management mode, the main goal of Oracle is to honor the PGA_AGGREGATE_TARGET limit set by the DBA, by controlling dynamically the amount of PGA memory allotted to SQL work areas. At the same time, Oracle tries to maximize the performance of all the memory-intensive SQL operations by maximizing the number of work areas that are using an optimal amount of PGA memory (cache memory). The rest of the work areas are executed in one-pass mode, unless the PGA memory limit set by the DBA with the parameter PGA_AGGREGATE_TARGET is so low that multi-pass execution is required to reduce even more the consumption of PGA memory and honor the PGA target limits.

To set the PGA initially, rule of thumb says to set the value at 20% of (80% of Total Physical memory) for OLTP and 50% of (80% of Total Available Memory) for DSS systems. Here we are taking 80% of Total Available Memory as the size of the SGA.

Three statistics have been added to the V\$SYSSTAT and V\$SESSTAT views that relate to automated PGA memory. These are:

Work Area Executions: Optimal Size Represents the number of work areas that had an optimal size, and no writes to disk were required.

Work Area Executions: One Pass Size Represents the number of work areas that had to write to disk, but required only one pass to disk.

Work Area Executions: Multipasses Size represents the number of work areas that had to write to disk using multiple passes. High numbers of this statistic might indicate a poorly tuned PGA.

New columns have been added to V\$PROCESS to help tune the PGA:

PGA_USED_MEM - reports how much PGA memory the process uses.

PGA_ALLOCATED_MEM - the amount of PGA memory allocated to the process.

PGA_MAX_MEM - the maximum amount of PGA memory allocated by the process

Finally, three new views are available to help the DBA extract information about the PGA:

V\$SQL_WORKAREA - provides information about SQL work areas.

V\$SQL_WORKAREA_ACTIVE - provides information on current SQL work area allocations.

V\$SQL_MEMORY_USAGE - displays current memory-use statistics.

10. Explain two easy SQL optimizations?

a. EXISTS can be better than IN under various conditions.

b. UNION ALL is faster than UNION (not sorting).

11. What are the new features in Oracle 11gR1?

12. What are the new features in Oracle 11g R2?

13. What are the new features in Oracle 12c?

14. What process will get data from datafiles to DB cache?

Server process

15. What background process will writes data to datafiles?

DBWR

16. What background process will write undo data?

DBWR

17. What are physical components of Oracle database?

Oracle database is comprised of three types of files. One or more datafiles, two or more redo log files, and one or more control files. Password file and parameter file also come under physical components.

18. What are logical components of Oracle database?

Blocks, Extents, Segments, Tablespaces

19. Types segment space management?

LMTS and DMTS

When Oracle allocates space to a segment (like a table or index), a group of contiguous free blocks, called an extent, is added to the segment. Metadata regarding extent allocation and unallocated extents are either stored in the data dictionary, or in the tablespace itself. Tablespaces that record extent allocation in the dictionary, are called dictionary managed tablespaces, and tablespaces that record extent allocation in the tablespace header, are called locally managed tablespaces.

```
SQL> select tablespace_name, extent_management, allocation_type from dba_tablespaces;
```

TABLESPACE_NAME	EXTENT_MAN	ALLOCATIO
-----------------	------------	-----------

SYSTEM	DICTIONARY	USER
SYS_UNDOTS	LOCAL	SYSTEM
TEMP	LOCAL	UNIFORM

Dictionary Managed Tablespaces (DMT):

Oracle use the data dictionary (tables in the SYS schema) to track allocated and free extents for tablespaces that is in "dictionary managed" mode. Free space is recorded in the SYS.FET\$ table, and used space in the SYS.UEF\$ table. Whenever space is required in one of these tablespaces, the ST (space transaction) enqueue latch must be obtained to do inserts and deletes against these tables. As only one process can acquire the ST enqueue at a given time, this often lead to contention.

Execute the following statement to create a dictionary managed tablespace:

```
SQL> CREATE TABLESPACE ts1 DATAFILE '/oradata/ts1_01.dbf' SIZE 50M
      EXTENT MANAGEMENT DICTIONARY
      DEFAULT STORAGE ( INITIAL 50K NEXT 50K MINEXTENTS 2 MAXEXTENTS 50 PCTINCREASE 0);
```

Locally Managed Tablespaces (LMT):

Using LMT, each tablespace manages its own free and used space within a bitmap structure stored in one of the tablespace's data files. Each bit corresponds to a database block or group of blocks. Execute one of the following statements to create a locally managed tablespace:

```
SQL> CREATE TABLESPACE ts2 DATAFILE '/oradata/ts2_01.dbf' SIZE 50M
      EXTENT MANAGEMENT LOCAL AUTOALLOCATE;
SQL> CREATE TABLESPACE ts3 DATAFILE '/oradata/ts3_01.dbf' SIZE 50M
      EXTENT MANAGEMENT LOCAL UNIFORM SIZE 128K;
```

Note the difference between AUTOALLOCATE and UNIFORM SIZE:

AUTOALLOCATE specifies that extent sizes are system managed. Oracle will choose "optimal" next extent sizes starting with 64KB. As the segment grows larger extent sizes will increase to 1MB, 8MB, and eventually to 64MB. This is the recommended option for a low or unmanaged environment.

UNIFORM specifies that the tablespace is managed with uniform extents of SIZE bytes (use K or M to specify the extent size in kilobytes or megabytes). The default size is 1M. The uniform extent size of a locally managed tablespace cannot be overridden when a schema object, such as a table or an index, is created.

Also not, if you specify, LOCAL, you cannot specify DEFAULT STORAGE, MINIMUM EXTENT or TEMPORARY

Advantages of Locally Managed Tablespaces:

- Eliminates the need for recursive SQL operations against the data dictionary (UEF\$ and FET\$ tables)
- Reduce contention on data dictionary tables (single ST enqueue)
- Locally managed tablespaces eliminate the need to periodically coalesce free space (automatically tracks adjacent free space)
- Changes to the extent bitmaps do not generate rollback information

Locally Managed SYSTEM Tablespace:

From Oracle9i release 9.2 one can change the SYSTEM tablespace to locally managed. Further, if you create a database with DBCA (Database Configuration Assistant), it will have a locally managed SYSTEM tablespace by default. The following restrictions apply:

- No dictionary-managed tablespace in the database can be READ WRITE.
- You cannot create new dictionary managed tablespaces
- You cannot convert any dictionary managed tablespaces to local

- Thus, it is best only to convert the SYSTEM tablespace to LMT after all other tablespaces are migrated to LMT.

Segment Space Management in LMT:

From Oracle 9i, one can not only have bitmap managed tablespaces, but also bitmap managed segments when setting Segment Space Management to AUTO for a tablespace. Look at this example:

```
SQL> CREATE TABLESPACE ts4 DATAFILE '/oradata/ts4_01.dbf' SIZE 50M
      EXTENT MANAGEMENT LOCAL
      SEGMENT SPACE MANAGEMENT AUTO;
```

Segment Space Management eliminates the need to specify and tune the PCTUSED, FREELISTS, and FREELISTS GROUPS storage parameters for schema objects. The Automatic Segment Space Management feature improves the performance of concurrent DML operations significantly since different parts of the bitmap can be used simultaneously eliminating serialization for free space lookups against the FREELISTS. This is of particular importance when using RAC, or if "buffer busy waits" are detected.

Convert between LMT and DMT:

The DBMS_SPACE_ADMIN package allows DBAs to quickly and easily convert between LMT and DMT mode. Look at these examples:

```
SQL> exec dbms_space_admin.Tablespace_Migrate_TO_Local('ts1');
PL/SQL procedure successfully completed.
SQL> exec dbms_space_admin.Tablespace_Migrate_FROM_Local('ts2');
PL/SQL procedure successfully completed.
```

20. Types of extent management?

Auto and Manual

AUTO: Specify AUTO if you want the database to manage the free space of segments in the tablespace using a bitmap. If you specify AUTO, then the database ignores any specification for PCTUSED, FREELIST, and FREELIST GROUPS in subsequent storage specifications for objects in this tablespace. This setting is called automatic segment-space management. Oracle strongly recommends that you create tablespaces with automatic segment-space management.

MANUAL: Specify MANUAL if you want the database to manage the free space of segments in the tablespace using free lists.

To determine the segment management of an existing tablespace, query the SEGMENT_SPACE_MANAGEMENT column of the DBA_TABLESPACES or USER_TABLESPACES data dictionary view.

Notes: If you specify AUTO segment management, then:

- If you set extent management to LOCAL UNIFORM, then you must ensure that each extent contains at least 5 database blocks.
- If you set extent management to LOCAL AUTOALLOCATE, and if the database block size is 16K or greater, then Oracle manages segment space by creating extents with a minimum size of 5 blocks rounded up to 64K.

Restrictions on Automatic Segment-space Management

- You can specify this clause only for a permanent, locally managed tablespace.
- You cannot specify this clause for the SYSTEM tablespace.

21. What are the differences between LMTS and DMTS?

Tablespaces that record extent allocation in the dictionary are called dictionary managed tablespaces, and tablespaces that record extent allocation in the tablespace header are called locally managed tablespaces.

22. What is a datafile?

Every Oracle database has one or more physical datafiles. Datafiles contain all the database data. The data of logical database structures such as tables and indexes is physically stored in the datafiles allocated for a database.

23. What are the contents of control file?

Database name, SCN, LSN, datafile locations, redo log locations, archive mode, DB Creation Time, RMAN Backup & Recovery Details, Flashback mode.

24. What is the use of redo log files? (<http://www.ordba.net/Tutorials/Redolog.htm>)

Explanation-1: Redo logs are transaction journals. Each transaction is recorded in the redo logs. Redo logs are used in a serial fashion with each transaction queuing up in the redo log buffers and being written one at a time into the redo logs. Redo logs as a general rule should switch about every thirty minutes. However, you may need to adjust the time up or down depending on the importance of your data. The rule of thumb is to size the redo logs such that you only lose the amount of data you can stand to lose should for some reason the online redo log become corrupt. With modern Oracle redo log mirroring and with disk array mirroring and various forms of online disk repair and replacement the occurrence of redo log corruptions has dropped to practically zero, so size based on the number of archive logs you want to apply should the database fail just before your next backup.

The LOG_BUFFER_SIZE and LOG_BUFFERS parameters control the redo log buffers. The LOG_BUFFER_SIZE should be set to reduce the number of writes required per redo log but not be so large that it results in an excessive IO wait time. Some studies have shown that sizing bigger than one megabyte rarely results in performance gains. Generally I size the LOG_BUFFER_SIZE such that it is equal to or results in an even divisor of the redo log size.

Monitor redo logs using the alert log, V\$LOGHIST, V\$LOGFILE, V\$RECOVERY_LOG and V\$LOG DPTs.

Explanation-2: Redo log files record changes made to the database and are used by Oracle for system crash recovery. Archiving of redo log files is necessary for hot (on-line) backups, and is mandatory for point-in-time recovery. Redo log files are created upon database creation and additional ones can be added by the DBA. To enable archive redo logging, the init.ora file must be modified, the database needs to be altered, and filesystem space is required. The following explains a little about redo logs, how archive logging can be enabled, and how backups can be performed.

Why redo log files:

Crash Recovery: Redo log files record changes made to the database. Databases can crash in many ways, such as a sudden power loss, a SHUTDOWN ABORT, or the death of an Oracle process. In these cases, redo log files can provide information about how to repair the database. During the ALTER DATABASE OPEN phase of database startup, the on-line redo log files are used for "crash recovery". This type of recovery is generally handled by Oracle and does not require DBA intervention.

Point-In-Time Recovery: Redo log files contain information that can be useful for broader types of recovery. Since they contain all the changes that brought the database to its current state, the redo logs can bring an old backup forward to any point in time. However, on-line redo log files are used in a circular fashion, so it is important to make a copy of each redo log file before it gets overwritten with new information. This can be done automatically with archive log mode.

Hot Backups: During a hot backup, a tablespace's writes are done in a special manner. During this time, tables residing on this tablespace can be modified, however, extra information about the change is written to the redo log files. After the tablespace backup is finished, normal on-line redo logging is resumed. Note that during a hot backup each datafile backup is from a different point in time. And, in some cases, the datafile itself could have been modified during the backup process. If all of these datafiles were restored, the database would be completely out of sync - each part would be from a different time. In this case, old copies of the on-line redo log files (archived redo logs) can be applied to each datafile to bring them all to a single point in time.

25. What are the uses of undo tablespace or redo segments?

Every Oracle database must have a method of maintaining information that is used to roll back, or undo, changes to the database. Such information consists of records of the actions of transactions, primarily before they are committed. Oracle refers to these records collectively as **undo**.

Undo records are used to:

- Roll back transactions when a ROLLBACK statement is issued

- Recover the database

- Provide read consistency

When a rollback statement is issued, undo records are used to undo changes that were made to the database by the uncommitted transaction. During database recovery, undo records are used to undo any uncommitted changes applied from the redo log to the datafiles. Undo records provide read consistency by maintaining the before image of the data for users who are accessing the data at the same time that another user is changing it.

Historically, Oracle has used rollback segments to store undo. Space management for these rollback segments has proven to be quite complex. Oracle now offers another method of storing undo that eliminates the complexities of managing rollback segment space, and enables DBAs to exert control over how long undo is retained before being overwritten. This method uses an undo tablespace. Both of these methods of managing undo space are discussed in this chapter.

You cannot use both methods in the same database instance, although for migration purposes it is possible, for example, to create undo tablespaces in a database that is using rollback segments, or to drop rollback segments in a database that is using undo tablespaces. However, you must shut down and restart your database in order to effect the switch to another method of managing undo.

Undo vs Rollback

Earlier releases of Oracle Database used rollback segments to store undo. Oracle9i introduced automatic undo management, which simplifies undo space management by eliminating the complexities associated with rollback segment management. Oracle strongly recommends (Oracle 9i and on words) to use undo tablespace (automatic undo management) to manage undo rather than rollback segments.

To see the undo management mode and other undo related information of database-

```
SQL> show parameter undo
```

NAME	TYPE	VALUE
undo_management	string	AUTO
undo_retention	integer	900
undo_tablespace	string	UNDOTBS1

Since the advent of Oracle9i, less time-consuming and suggested way is—using Automatic Undo Management—in which Oracle Database creates and manages rollback segments (now called “undo segments”) in a special-purpose undo tablespace. Unlike with rollback segments, we don’t need to create or manage individual undo segments—Oracle Database does that for you when you create the undo tablespace. All transactions in an instance share a single undo tablespace. Any executing transaction can consume free space in the undo tablespace, and when the transaction completes, its undo space is freed (depending on how it’s been sized and a few other factors, like undo retention). Thus, space for undo segments is dynamically allocated, consumed, freed, and reused—all under the control of Oracle Database, rather than manual management by someone.

Switching Rollback to Undo

1. We have to create an Undo tablespace. Oracle provides a function (10g and up) that provides information on how to size new undo tablespace based on the configuration and usage of the rollback segments in the system.

```
DECLARE
```

```
utbsiz_in_MB NUMBER;
```

```
BEGIN
```

```
utbsiz_in_MB := DBMS_UNDO_ADV.RBU_MIGRATION;
```

```
end;
```

```
/
```

```
CREATE UNDO TABLESPACE UNDOTBS
```

```
DATAFILE '/oradata/dbf/undotbs_1.dbf'
```

```
SIZE 100M AUTOEXTEND ON NEXT 10M
```

```
MAXSIZE UNLIMITED RETENTION NOGUARANTEE;
```

Note: In undo tablespace creation, “SEGMENT SPACE MANAGEMENT AUTO” can not be set

2. Change system parameters

```
SQL> alter system set undo_retention=900 scope=both;
```

```
SQL> alter system set undo_tablespace=UNDOTBS scope=both;
```

```
SQL> alter system set undo_management=AUTO scope=spfile;
```

```
SQL> shutdown immediate
```

```
SQL> startup
```

UNDO_MANAGEMENT is a static parameter. So database needs to be restarted.

26. How undo tablespace can guarantee retain of required undo data?

```
Alter tablespace undo_ts retention guarantee;
```

27. What is ORA-01555 - snapshot too old error and how do you avoid it?

The ORA-01555 is caused by Oracle read consistency mechanism. If you have a long running SQL that starts at 10:30 AM, Oracle ensures that all rows are as they appeared at 10:30 AM, even if the query runs until noon!

Oracles does this by reading the "before image" of changed rows from the online undo segments. If you have lots of updates, long running SQL and too small UNDO, the ORA-01555 error will appear.

From the docs we see that the ORA-01555 error relates to insufficient undo storage or a too small value for the *undo_retention* parameter:

ORA-01555: snapshot too old: rollback segment number string with name "string" too small

Cause: Rollback records needed by a reader for consistent read are overwritten by other writers.

Action: If in Automatic Undo Management mode, increase the setting of UNDO_RETENTION. Otherwise, use larger rollback segments.

You can get an ORA-01555 error with a too-small *undo_retention*, even with a large undo tables. However, you can set a super-high value for *undo_retention* and still get an ORA-01555 error. Also see these important notes on [commit frequency and the ORA-01555 error](#)

The ORA-01555 snapshot too old error can be addressed by several remedies:

- Re-schedule long-running queries when the system has less DML load.
- [Increasing the size of your rollback segment \(undo\) size](#). The ORA-01555 snapshot too old also relates to your setting for [automatic undo retention](#).
- Don't fetch between commits.

Avoiding the ORA-01555 error

- [Steve Adams](#) has good notes on avoiding the ora-1555 snapshot too old error:
- Do not run discrete transactions while sensitive queries or transactions are running, unless you are confident that the data sets required are mutually exclusive.
- Schedule long running queries and transactions out of hours, so that the consistent gets will not need to rollback changes made since the snapshot SCN. This also reduces the work done by the server, and thus improves performance.
- Code long running processes as a series of restartable steps.
- Shrink all rollback segments back to their optimal size manually before running a sensitive query or transaction to reduce risk of consistent get rollback failure due to extent deallocation.
- Use a large optimal value on all rollback segments, to delay extent reuse.
- Don't fetch across commits. That is, don't fetch on a cursor that was opened prior to the last commit, particularly if the data queried by the cursor is being changed in the current session.
- Use a large database block size to maximize the number of slots in the rollback segment transaction tables, and thus delay slot reuse.
- Commit less often in tasks that will run at the same time as the sensitive query, particularly in PL/SQL procedures, to reduce transaction slot reuse.
- If necessary, add extra rollback segments (undo logs) to make more transaction slots available.

Oracle ACE Steve Karam also has advice on avoiding the ORA-01555: Snapshot too old, rollback segment too small with UNDO sizing.

Question: I am updating 1 million rows on Oracle 10g, and I run it as batch process, committing after each batch to avoid undo generation. But in Oracle 10g I am told undo management is automatic and I do not need run the update as batch process.

Answer: Automatic undo management was available in 9i as well, and my guess is you were probably using it there. However, I'll assume for the sake of this writing that you were using manual undo management in 9i and are now on automatic.

Automatic undo management depends upon the UNDO_RETENTION parameter, which defines how long Oracle should try to keep committed transactions in UNDO segments. However, the UNDO_RETENTION parameter is only a suggestion. You must also have an UNDO tablespace that's large enough to handle the amount of UNDO you will be generating/holding, or you will get "ORA-01555: Snapshot too old, rollback segment too small" errors.

You can use the UNDO advisor to find out how large this tablespace should be given a desired UNDO retention, or look online for some scripts...just Google for: oracle undo size

Oracle 10g also gives you the ability to guarantee undo. This means that instead of throwing an error on SELECT statements, it guarantees your UNDO retention for consistent reads and instead errors your DML that would cause UNDO to be overwritten.

Now, for your original question...yes, it's easier for the DBA to minimize the issues of UNDO when using automatic undo management. If you set the UNDO_RETENTION high enough with a properly sized undo tablespace you shouldn't have as many issues with UNDO.

How often you commit should have nothing to do with it, as long as your DBA has properly set UNDO_RETENTION and has an optimally sized UNDO tablespace. Committing more often will only result in your script taking longer, more LGWR/DBWR issues, and the "where was I" problem if there is an error (if it errors, where did it stop?).

Lastly (and true even for manual undo management), if you commit more frequently, you make it more possible for ORA-01555 errors to occur. Because your work will be scattered among more undo segments, you increase the chance that a single one may be overwritten if necessary, thus causing an ORA-01555 error for those that require it for read consistency.

It all boils down to the size of the undo tablespace and the undo retention, in the end...just as manual management boiled down to the size, amount, and usage of rollback segments. Committing frequently is peroxidizing band-aid: it covers up the problem, tries to clean it, but in the end it just hurts and causes problems for otherwise healthy processes.

Oracle guru Joel Garry offers another great explanation of the machinations of the ORA-01555 error:

You have to understand, in general, ORA-01555 means something else is causing it to die - Oracle needs to be able to create a read-consistent view of the table for the query as it looked at the start of the query, and it is unable to because something has overwritten the undo necessary to create such a view. Since you have the same table over and over in your alert log, that probably means the something is the previous queries your monitoring software is making, not ever releasing the transaction.

Something like:

- 10AM query starts, never ends
- 11AM query starts, never ends
- Noon query starts, never ends
- 1PM query starts

Meanwhile, the undo needed from the 10AM query for the 1PM query gets overwritten, 1PM query dies with ORA-01555, since it needs to know what the table looked like before the 10AM query started mucking with it.

Also if the query is a loop with a commit in it, it can do the same thing without other queries, as eventually the next iteration requires looking back at it's own previous first generation, can't do it, and barfs.

Upping undo_retention may help, or may not, depending on the real cause. Also check `v$undostat`, you may still have information in there if this is ongoing (or may not, since by the time you check it the needed info may be gone).

28. What is the use/size of temporary tablespace?

Temporary tablespaces are used for special operations, particularly for sorting data results on disk. For SQL with millions of rows returned, the sort operation is too large for the RAM area and must occur on disk. The temporary tablespace is where this takes place.

Each database should have one temporary tablespace that is created when the database is created. You create, drop and manage tablespaces with create temporary tablespace, drop temporary tablespace and alter temporary tablespace commands, each of which is like it's create tablespace counterpart.

The only other difference is that a temporary tablespace uses temporary files (also called tempfiles) rather than regular datafiles. Thus, instead of using the datafiles keyword you use the tempfiles keyword when issuing a create, drop or alter tablespace command as you can see in these examples:

```
CREATE TEMPORARY TABLESPACE temp
```

```
TEMPFILE '/ora01/oracle/oradata/booktst_temp_01.dbf' SIZE 50m;
```

```
DROP TEMPORARY TABLESPACE temp INCLUDING CONTENTS AND DATAFILES;
```

Tempfiles are a bit different than datafiles in that you may not immediately see them grow to the size that they have been allocated (this particular functionality is platform dependent). Hence, don't panic if you see a file that looks too small.

Temporary Tablespace Group Overview

Oracle 10g first introduced "temporary tablespace group." A temporary tablespace group consists of only temporary tablespace, and has the following properties:

- It contains one or more temporary tablespaces.
- It contains only temporary tablespace.
- It is not explicitly created. It is created implicitly when the first temporary tablespace is assigned to it, and is deleted when the last temporary tablespace is removed from the group.

Temporary Tablespace Group Benefits

- Temporary tablespace group has the following benefits:
- It allows multiple default temporary tablespaces to be specified at the database level.
- It allows the user to use multiple temporary tablespaces in different sessions at the same time.
- It allows a single SQL operation to use multiple temporary tablespaces for sorting.

29. What is the use of password file?

Explanation-1:

As a DBA we must have used sqlplus “/as sysdba” to connect to database, atleast hundred times a day. Never bothered about the password to provide !!!

This is because we were using OS level authentication. We can change the configuration and make Oracle to ask for the password. Well, “/as sysdba” works fine if we are connecting to the host where the database is actually installed. For example I have installed a database as oracle01 user (which belongs to DBA group) on one of my host called “host1”. I telnet to host1 as oracle01 user and provide the password to login. Once I successfully login to the host, there ends the authentication part. Not for administering the database all I have to do is to use our famous command to connect to database – “sqlplus /as sysdba”.

The reason above thing work is because I was using Operating System level authentication. If I try to connect to same database **as sysdba** from some other host, I wont be able to connect. Because the authentication is done based on host login password. Since I haven’t logged into host, authentication will fail and connect as sysdba will fail. So for OS authentication its mandatory that you are always logged into the host where the oracle is installed (oracle database resides).

Authentication Type

There are 2 types of authentication:

- OS (Operating System) Authentication
- Password File Authentication

And yes the above one that i explained is OS level authentication. Lets see what is password file authentication.

Password File Authentication

In case of password file authentication, we create a password file for our database. ORAPWD is the utility for creating a password file. This utility is provided by oracle and comes when you install database. This binary is present in ORACLE_HOME/bin directory. Below is the usage for the same.

ORAPWD FILE=(file_name) password=(password) ENTRIES=(Entries)

Where file_name is the name and location of the password file. Usually we create password with name as ora(SID).pwd in ORACLE_HOME/dbs directory. So value for file_name becomes \$ORACLE_HOME/dbs/ora(sid).pwd
password - is the password you want to set for password file. Remember that this will become the password for sys user as well. Meaning that when you are connecting as sys user, you need to provide this password. (oracle will prompt for password in case of password file authentication).

Entries - This is the number of entries that password file can have. Be careful while providing this value as once you set this value, you cannot change it. You have to delete password file and recreate it, but its risky.

Example:

\$orapwd FILE=/u01/oracle/product/9.2.0/dbs/oraorcl.pwd PASSWORD=welcome1 ENTRIES=10

This will create a password file oraorcl.pwd in /u01/oracle/product/9.2.0/dbs directory.

After creating password file, how your database will know that you have created password file and you are supposed to use the same. This is done by INIT.ORA parameter REMOTE_LOGIN_PASSWORDFILE. This parameter can have 3 values (none - OS level authentication, shared/exclusive – password file authentication). So for using password file, you need to set the value of this parameter to either shared or exclusive.

What is the difference between SHARED and EXCLUSIVE?

If we set the value of REMOTE_LOGIN_PASSWORDFILE to SHARED in INIT.ORA file, then following is true.

- This file can be used for more then one database (shared file)
- One SYS user will be recognized by database. Meaning that you can login to database using SYS and **no other user holding sysdba responsibility**. However you can connect to database using SYSTEM or any other user but not the once holding sysdba responsibility.
- If we set the value of **REMOTE_LOGIN_PASSWORDFILE** to EXCLUSIVE in INIT.ORA file, then following is true.
- This file will be specific to one database only. Other database cannot use this file.
- Any user enter having sysdba responsibility, which is present in password file can be connected to database as sysdba from remote server.

So when using password file authentication remember to set the value of **REMOTE_LOGIN_PASSWORDFILE** to SHARED or EXCLUSIVE in INIT.ORA file. Also when using OS level authentication set the value if this parameter to NONE

Explanation-2: If the DBA wants to [start up](#) an Oracle [instance](#) there must be a way for Oracle to authenticate this DBA. That is if (s)he is allowed to do so. Obviously, his password can not be stored in the database, because Oracle

can not access the database before the instance is started up. Therefore, the authentication of the DBA must happen outside of the database. There are two distinct mechanisms to authenticate the DBA: using the **password file** or through the operating system.

The init parameter [remote_login_passwordfile](#) specifies if a password file is used to authenticate the DBA or not. If it set either to **shared** or **exclusive** a password file will be used.

Scenario:

QUICK REFERENCE

Step 1. Log on the database machine and create a password file:

For Unix (Shell)

```
orapwd file=$ORACLE_HOME/dbs/orapw password=password_for_sys
```

For Windows (Command Prompt)

```
orapwd file=%ORACLE_HOME%\database\PWDsid_name.ora  
password=password_for_sys
```

Step 2. Add the following line to init.ora in UNIX, or init.ora in Windows:

```
REMOTE_LOGIN_PASSWORDFILE=EXCLUSIVE
```

Step 3. Restart the Database and Test the Remote Login.

```
connect sys/password_for_sys@tns_name_of_db as sysdba
```

SYSDBA AUTHENTICATING APPROACHES

A SYSDBA authenticating approach is the method of verifying the identity of database administrators. On the layer of data dictionary, Oracle database administrators are authenticated using an account password like other users. In addition to the normal data dictionary, the following approaches are available to secure the authentication of administrators with the SYSDBA privilege:

- * Operating-System-based Authentication;
- * Password-File-based Authentication;
- * Strong and Centralized Authentication (from 11g on).

Operating-System-Based Authentication:

It means to authenticate database administrators by establishing a user group on the operating system, granting Oracle DBA privileges to that group, and then adding the database administrative users to that group. Users authenticated in this way can logon to the Oracle database as a SYSDBA without having to enter a user name or password (i.e. "connect / as sysdba"). On UNIX platform, the special user group is called the DBA group, and on Windows systems, it is called the ORA_DBA group.

Password-File-Based Authentication:

Oracle Database uses database-specific password files to keep track of the database users who have been granted the SYSDBA and SYSOPER privileges

Strong and Centralized Authentication:

This authenticating approach (from 11g on) is featured by a network-based authentication service, such as Oracle Internet Directory. It is recommended by Oracle for the centralized control of SYSDBA access to multiple databases. One of the following methods can be used to enable the Oracle Internet Directory server to authorize SYSDBA connections:

- * Directory Authentication;
- * Kerberos Authentication;
- * Secure Sockets Layer Authentication.

CONFIGURING STEPS

To use the password file authentication, you must configure the database to use a password file. To do so, you first need to create the password file, and then configure the database so that it knows to use it. Steps 1 to 3 require the local login to the database server.

Step 1: Create the Password File

To set a password file on the server-side, log on the server machine where the remote Oracle database resides. Create the database password file by using the Oracle utility "orapwd."

The Orapwd Command For Oracle 8.1.7 through 10g :

Usage: orapwd file=<filename> password=<password> [entries=<numusers>] where

- * file - (mandatory) The password filename (Refer to Notice 1);
- * password - (mandatory) The password for the sys user (Refer to Notice 3);
- * entries - (Optional) Maximum number of entries (user accounts) to permit in the file (Refer to Notice 2);

There are no spaces around the equal-to (=) character.

In UNIX:

For Shell :

```
orapwd file=$ORACLE_HOME/dbs/orapw password=change_on_install entries=30
```

For SQL* Plus :

```
host orapwd file=$ORACLE_HOME/dbs/orapw password=change_on_install entries=30
```

The above command creates a password file named "orapw" that allows up to 30 privileged users with different passwords.

In Windows:

For Command Prompt:

```
orapwd file=%ORACLE_HOME%\database\PWDorcl92.ora password=change_on_install entries=30
```

For SQL* Plus :

```
host orapwd file=%ORACLE_HOME%\database\PWDorcl92.ora password=change_on_install entries=30
```

The above command creates a password file named "PWDorcl92" that allows up to 30 privileged users with different passwords.

The Orapwd Command For Oracle 11g Release 1 :

Usage: orapwd file=<filename> [entries=<numusers>] [force={y|n}] [ignorecase={y|n}] [nosysdba={y|n}]

where

- * file - (mandatory) The password filename ;
- * entries - (Optional) Maximum number of entries (user accounts) to permit in the file;
- * force - (Optional) If y, permits overwriting an existing password file;
- * ignorecase - (Optional) If y, passwords are treated as case-insensitive;
- * nosysdba - (Optional) For Data Vault installations

There are no spaces around the equal-to (=) character.

The command, when executed, prompts for the SYS password and stores the password in the created password file.

Orapwd Command Examples:

In UNIX :

```
orapwd file=$ORACLE_HOME/dbs/orapw entries=30
```

Enter password: change_on_install

The above commands create a password file named "orapw" that has "change_on_install" as the password for the sys user and allows up to 30 privileged users with different passwords.

In Windows :

```
orapwd file=%ORACLE_HOME%\database\PWDorcl11.ora entries=30
```

Enter password: change_on_install

The above commands create a password file named "PWDorcl11" that has "change_on_install" as the password for the sys user and allows up to 30 privileged users with different passwords.

Step 2: Configure the Database to Use the Password File

By default, an Oracle database is not configured to use the password file. However, you'd better first verify the value of the parameter "remote_login_passwordfile" in initservice_name.ora, in UNIX, or init.ora, in Windows. If the value is "exclusive," continue with Step 3: Restart the Database. If the value is "shared," or if the line "REMOTE_LOGIN_PASSWORDFILE=EXCLUSIVE" is remarked off, continue with the procedure below: Stop the Database.

Use the SQLPlus show statement to check the parameter value:

```
SQL> show parameter password;
```

NAME	TYPE	VALUE
remote_login_passwordfile	string	EXCLUSIVE

Stop the database by stopping the services or using the SQLPlus shutdown immediate statement.

Add the following line to initservice_name.ora, in UNIX , or init.ora, in Windows

```
REMOTE_LOGIN_PASSWORDFILE=EXCLUSIVE
```

Step 4 : (Optional) Change the Password for the SYS User

```
SQL>PASSWORD sys;
```

Changing password for sys

New password: password

Retype new password: password

Step 5 : Verify Whether SYS Has the SYSDBA Privilege

Use the SQLPlus select statement to check the password file users:

```
SQL> select * from v$pwfile_users;
```

USERNAME	SYSDB	SYSOP
SYS	TRUE	TRUE

30. How to create password file?

```
$ orapwd file=orapwSID password=sys_password force=y nosysdba=y
```

31. How many types of indexes are there? (<http://www.orafaq.com/node/1403>)

Clustered and Non-Clustered

- 1.B-Tree index
- 2.Bitmap index
- 3.Unique index
- 4.Function based index
5. Implicit index and explicit index

Explicit indexes are again of many types like simple index, unique index, Bitmap index, Functional index, Organizational index, cluster index.

Explanation-1:

B*Tree Indexes: B*tree stands for balanced tree. This means that the height of the index is the same for all values thereby ensuring that retrieving the data for any one value takes approximately the same amount of time as for any other value. Oracle b*tree indexes are best used when each value has a high cardinality (low number of occurrences) for example primary key indexes or unique indexes. One important point to note is that NULL values are not indexed. They are the most common type of index in OLTP systems.

B*Tree Cluster Indexes: These are B*tree index defined for clusters. Clusters are two or more tables with one or more common columns and are usually accessed together (via a join).

```
CREATE INDEX product_orders_ix ON CLUSTER product_orders;
```

Hash Cluster Indexes: In a hash cluster rows that have the same hash key value (generated by a hash function) are stored together in the Oracle database. Hash clusters are equivalent to indexed clusters, except the index key is replaced with a hash function. This also means that here is no separate index as the hash is the index.

```
CREATE CLUSTER emp_dept_cluster (dept_id NUMBER) HASHKEYS 50;
```

Reverse Key Indexes: These are typically used in Oracle Real Application Cluster (RAC) applications. In this type of index the bytes of each of the indexed columns are reversed (but the column order is maintained). This is useful when new data is always inserted at one end of the index as occurs when using a sequence as it ensures new index values are created evenly across the leaf blocks preventing the index from becoming unbalanced which may in turn affect performance.

```
CREATE INDEX emp_ix ON emp(emp_id) REVERSE;
```

Bitmap Indexes: These are commonly used in datawarehouse applications for tables with no updates and whose columns have low cardinality (i.e. there are few distinct values). In this type of index Oracle stores a bitmap for each distinct value in the index with 1 bit for each row in the table. These bitmaps are expensive to maintain and are therefore not suitable for applications which make a lot of writes to the data.

For example consider a car manufacturer which records information about cars sold including the colour of each car. Each colour is likely to occur many times and is therefore suitable for a bitmap index.

```
CREATE BITMAP INDEX car_col ON cars(colour) REVERSE;
```

Partitioned Indexes: Partitioned Indexes are also useful in Oracle datawarehouse applications where there is a large amount of data that is partitioned by a particular dimension such as time.

Partition indexes can either be created as local partitioned indexes or global partitioned indexes. Local partitioned indexes means that the index is partitioned on the same columns and with the same number of partitions as the table. For global partitioned indexes the partitioning is user defined and is not the same as the underlying table.

Function-based Indexes: As the name suggests these are indexes created on the result of a function modifying a column value. For example

```
CREATE INDEX upp_ename ON emp(UPPER(ename));
```

The function must be deterministic (always return the same value for the same inputs).

Index Organised Tables: In an index-organised table all the data is stored in the Oracle database in a B*tree index structure defined on the table's primary key. This is ideal when related pieces of data must be stored together or data must be physically stored in a specific order. Index-organised tables are often used for information retrieval, spatial and OLAP applications.

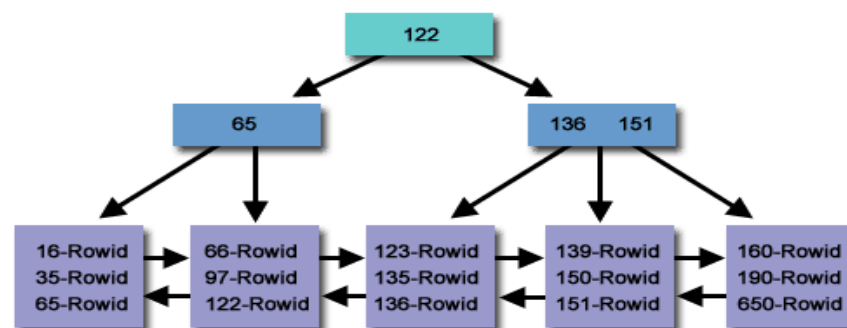
Domain Indexes: These indexes are created by user-defined indexing routines and enable the user to define his or her own indexes on custom data types (domains) such as pictures, maps or fingerprints for example. These type of index require in-depth knowledge about the data and how it will be accessed.

Oracle includes numerous data structures to improve the speed of Oracle SQL queries. Taking advantage of the low cost of disk storage, Oracle includes many new indexing algorithms that dramatically increase the speed with which Oracle queries are serviced. This article explores the internals of Oracle indexing; reviews the standard b-tree index, bitmap indexes, function-based indexes, and index-only tables (IOTs); and demonstrates how these indexes may dramatically increase the speed of Oracle SQL queries.

Oracle uses indexes to avoid the need for large-table, full-table scans and disk sorts, which are required when the SQL optimizer cannot find an efficient way to service the SQL query. I begin our look at Oracle indexing with a review of standard Oracle b-tree index methodologies.

The Oracle b-tree index: The oldest and most popular type of Oracle indexing is a standard b-tree index, which excels at servicing simple queries. The b-tree index was introduced in the earliest releases of Oracle and remains widely used with Oracle.

B-tree indexes are used to avoid large sorting operations. For example, a SQL query requiring 10,000 rows to be presented in sorted order will often use a b-tree index to avoid the very large sort required to deliver the data to the end user.



An Oracle b-tree index

Oracle offers several options when creating an index using the default b-tree structure. It allows you to index on multiple columns (concatenated indexes) to improve access speeds. Also, it allows for individual columns to be sorted in different orders. For example, we could create a b-tree index on a column called last_name in ascending order and have a second column within the index that displays the salary column in descending order.

```

create index
  name_salary_idx
on
  person
(
  last_name asc,
  salary desc);

```

While b-tree indexes are great for simple queries, they are not very good for the following situations:

- **Low-cardinality columns**—columns with less than 200 distinct values do not have the selectivity required in order to benefit from standard b-tree index structures.
- **No support for SQL functions**—B-tree indexes are not able to support SQL queries using Oracle's built-in functions. Oracle9i provides a variety of built-in functions that allow SQL statements to query on a piece of an indexed column or on any one of a number of transformations against the indexed column.

Prior to Oracle9i, the Oracle SQL optimizer had to perform time-consuming long-table, full-table scans due to these shortcomings. Consequently, it was no surprise when Oracle introduced more robust types of indexing structures.

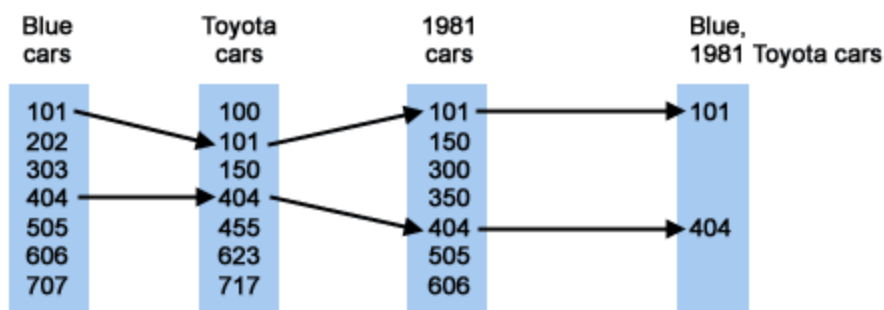
Bitmapped indexes: Oracle bitmap indexes are very different from standard b-tree indexes. In bitmap structures, a two-dimensional array is created with one column for every row in the table being indexed. Each column represents a distinct value within the bitmapped index. This two-dimensional array represents each value within the index multiplied by the number of rows in the table. At row retrieval time, Oracle decompresses the bitmap into the RAM data buffers so it can be rapidly scanned for matching values. These matching values are delivered to Oracle in the form of a Row-ID list, and these Row-ID values may directly access the required information.

The real benefit of bitmapped indexing occurs when one table includes multiple bitmapped indexes. Each individual column may have low cardinality. The creation of multiple bitmapped indexes provides a very powerful method for rapidly answering difficult SQL queries.

For example, assume there is a motor vehicle database with numerous low-cardinality columns such as car_color, car_make, car_model, and car_year. Each column contains less than 100 distinct values by themselves, and a b-tree index would be fairly useless in a database of 20 million vehicles. However, combining these indexes together in a query can provide blistering response times a lot faster than the traditional method of reading each one of the 20 million rows in the base table. For example, assume we wanted to find old blue Toyota Corollas manufactured in 1981:

```
select
  license_plat_nbr
from
  vehicle
where
  color = 'blue'
and
  make = 'toyota'
and
  year = 1981;
```

Oracle uses a specialized optimizer method called a bitmapped index merge to service this query. In a bitmapped index merge, each Row-ID, or RID, list is built independently by using the bitmaps, and a special merge routine is used in order to compare the RID lists and find the intersecting values. Using this methodology, Oracle can provide sub-second response time when working against multiple low-cardinality columns:



Oracle bitmap merge join

Function-based indexes: One of the most important advances in Oracle indexing is the introduction of function-based indexing. Function-based indexes allow creation of indexes on expressions, internal functions, and user-written functions in PL/SQL and Java. Function-based indexes ensure that the Oracle designer is able to use an index for its query.

Prior to Oracle8, the use of a built-in function would not be able to match the performance of an index.

Consequently, Oracle would perform the dreaded full-table scan. Examples of SQL with function-based queries might include the following:

```
Select * from customer where substr(cust_name,1,4) = 'BURL';
Select * from customer where to_char(order_date,'MM') = '01';
Select * from customer where upper(cust_name) = 'JONES';
Select * from customer where initcap(first_name) = 'Mike';
```

In Oracle, Oracle always interrogates the *where* clause of the SQL statement to see if a matching index exists. By using

function-based indexes, the Oracle designer can create a matching index that exactly matches the predicates within the SQL *where* clause. This ensures that the query is retrieved with a minimal amount of disk I/O and the fastest possible speed.

Once a function-based index is created, you need to create CBO statistics, but beware that there are numerous bugs and issues when analyzing a function-based index. See these important notes on [statistics and function-based indexes](#).

Index-only tables: Beginning with Oracle8, Oracle recognized that a table with an index on every column did not require table rows. In other words, Oracle recognized that by using a special table-access method called an index fast full scan, the index could be queried without actually touching the data itself.

Oracle codified this idea with its use of index-only table (IOT) structure. When using an IOT, Oracle does not create the actual table but instead keeps all of the required information inside the Oracle index. At query time, the Oracle SQL optimizer recognizes that all of the values necessary to service the query exist within the index tree, at which time the Oracle cost-based optimizer has a choice of either reading through the index tree nodes to pull the information in sorted order or invoke an index fast full scan, which will read the table in the same fashion as a full table scan, using sequential prefetch (as defined by the `db_file_multiblock_read_count` parameter). The multiblock read facility allows Oracle to very quickly scan index blocks in linear order, quickly reading every block within the index tablespace. Here is an example of the syntax to create an IOT.

```
CREATE TABLE emp_iot (  
    emp_id number,  
    ename varchar2(20),  
    sal number(9,2),  
    deptno number,  
    CONSTRAINT pk_emp_iot_index PRIMARY KEY (emp_id) )
```

ORGANIZATION index

TABLESPACE spc_demo_ts_01

PCTHRESHOLD 20 INCLUDING ename;

Index performance

Oracle indexes can greatly improve query performance but there are some important indexing concepts to understand.

- Index clustering
- Index block sizes

Indexes and blocksize: Indexes that experience lots of index range scans or index fast full scans (as evidenced by multiblock reads) will greatly benefit from residing in a 32k blocksize.

Today, most Oracle tuning experts utilize the multiple blocksize feature of Oracle because it provides buffer segregation and the ability to place objects with the most appropriate blocksize to reduce buffer waste. Some of the world record Oracle [benchmarks](#) use very large data buffers and multiple block sizes.

According to an [article](#) by Christopher Foot, author of the [OCP Instructors Guide for Oracle DBA Certification](#), larger block sizes can help in certain situations:

"A bigger block size means more space for key storage in the branch nodes of B-tree indexes, which reduces index height and improves the performance of indexed queries."

In any case, there appears to be evidence that block size affects the tree structure, which supports the argument that data blocks affect the structure of the tree.

Indexes and clustering: The CBO's decision to perform a full-table vs. an index range scan is influenced by the `clustering_factor` (located inside the `dba_indexes` view), `db_block_size`, and `avg_row_len`. It is important to understand how the CBO uses these statistics to determine the fastest way to deliver the desired rows.

Conversely, a high `clustering_factor`, where the value approaches the number of rows in the table (`num_rows`), indicates that the rows are not in the same sequence as the index, and additional I/O will be required for index range scans. As the `clustering_factor` approaches the number of rows in the table, the rows are out of sync with the index. Oracle MOSC Note:223117.1 has some great advice for tuning-down "db file sequential read" waits by table reorganization in row-order:

- If Index Range scans are involved, more blocks than necessary could be being visited if the index is unselective: by forcing or enabling the use of a more selective index, we can access the same table data by visiting fewer index blocks (and doing fewer physical I/Os).
- If the index being used has a large Clustering Factor, then more table data blocks have to be visited in order to get

the rows in each Index block: by rebuilding the table with its rows sorted by the particular index columns we can reduce the Clustering Factor and hence the number of table data blocks that we have to visit for each index block. This validates the assertion that the physical ordering of table rows can reduce I/O (and stress on the database) for many SQL queries.

Tip! In some cases Oracle is able to bypass a sort by reading the data in sorted order from the index. Oracle will even read data in reverse order from an index to avoid an in-memory sort.

32. What is bitmap index & when it'll be used?

- Bitmap indexes are preferred in Data warehousing environment. Refer Q31
- Preferred when cardinality is low.

33. What is B-tree index & when it'll be used?

- B-tree indexes are preferred in OLTP environment. Refer Q31
- Preferred when cardinality is high

34. How you will find out fragmentation of index?

- AUTO_SPACE_ADVISOR_JOB will run in daily maintenance window and report fragmented Indexes/Tables

SQL>ANALYZE INDEX VALIDATE STRUCTURE;

This populates the table 'INDEX_STATS'. It should be noted that this table contains only one row and therefore only one index can be analyzed at a time.

An index should be considered for rebuilding under any of the following conditions:

- * the percentage of deleted rows exceeds 30% of the total, i.e. if $\text{del_lf_rows} / \text{lf_rows} > 0.3$.
- * If the 'HEIGHT' is greater than 4.
- * If the number of rows in the index ('LF_ROWS') is significantly smaller than 'LF_BKLS' this can indicate a large number of deletes, indicating that the index should be rebuilt.

35. What is the difference between delete and truncate?

Truncate will release the space. Delete won't.

Delete can be used to delete some records. Truncate can't.

Delete can be rolled back.

Delete will generate undo (Delete command will log the data changes in the log file whereas the truncate will simply remove the data without it. Hence data removed by Delete command can be rolled back but not the data removed by TRUNCATE).

Truncate is a DDL statement whereas DELETE is a DML statement.

Truncate is faster than delete.

36. What's the difference between a primary key and a unique key?

Both primary and unique key enforce uniqueness of the column on which they are defined. But by default primary key creates a clustered index on the column, whereas unique key creates a non-clustered index by default. Primary key doesn't allow NULLs, but unique key allows one NULL only.

37. What is the difference between schema and user?

Schema is collection of user's objects.

38. What is the difference between SYSDBA, SYSOPER and SYSASM?

SYSOPER can't create and drop database.

SYSOPER can't do incomplete recovery.

SYSOPER can't change character set.

SYSOPER can't CREATE DISKGROUP; ADD/DROP/RESIZE DISK

SYSASM can do anything SYSDBA can do.

38. What is the difference between SYS and SYSTEM?

SYSTEM can't shutdown the database.

SYSTEM can't create another SYSTEM, but SYS can create another SYS or SYSTEM.

Explanation-1: In general, unless the documentation tells you, you will NEVER LOG IN as sys or system, they are our internal data dictionary accounts and not for your use. You will be best served by forgetting they exist.

Sysdba and sysoper are ROLES - they are not users, not schemas. The SYSDBA role is like "root" on UNIX or "Administrator" on Windows. It sees all, can do all. Internally, if you connect as sysdba, your schema name will appear to be SYS.

In real life, you hardly EVER need sysdba - typically only during an upgrade or patch.

Sysoper is another role, if you connect as sysoper, you'll be in a schema "public" and will only be able to do things granted to public AND start/stop the database. Sysoper is something you should use to startup and shutdown. You'll

use sysoper much more often than sysdba.

Do not grant sysdba to anyone unless and until you have absolutely verified they have the NEED for sysdba - the same with sysoper.

Explanation-2: The following administrative user accounts are automatically created when you install Oracle Database. They are both created with the password that you supplied upon installation, and they are both automatically granted the DBA role.

SYS

This account can perform all administrative functions. All base (underlying) tables and views for the database data dictionary are stored in the SYS schema. These base tables and views are critical for the operation of Oracle Database. To maintain the integrity of the data dictionary, tables in the SYS schema are manipulated only by the database. They should never be modified by any user or database administrator. You must not create any tables in the SYS schema. The SYS user is granted the SYSDBA privilege, which enables a user to perform high-level administrative tasks such as backup and recovery.

SYSTEM

This account can perform all administrative functions except the following:

- Backup and recovery
- Database upgrade

While this account can be used to perform day-to-day administrative tasks, Oracle strongly recommends creating named users account for administering the Oracle database to enable monitoring of database activity.

SYSDBA and SYSOPER System Privileges

SYSDBA and SYSOPER are administrative privileges required to perform high-level administrative operations such as creating, starting up, shutting down, backing up, or recovering the database. The SYSDBA system privilege is for fully empowered database administrators and the SYSOPER system privilege allows a user to perform basic operational tasks, but without the ability to look at user data.

The SYSDBA and SYSOPER system privileges allow access to a database instance even when the database is not open. Control of these privileges is therefore completely outside of the database itself. This control enables an administrator who is granted one of these privileges to connect to the database instance to start the database. You can also think of the SYSDBA and SYSOPER privileges as types of connections that enable you to perform certain database operations for which privileges cannot be granted in any other way. For example, if you have the SYSDBA privilege, then you can connect to the database using AS SYSDBA.

The SYS user is automatically granted the SYSDBA privilege upon installation. When you log in as user SYS, you must connect to the database as SYSDBA or SYSOPER. Connecting as a SYSDBA user invokes the SYSDBA privilege; connecting as SYSOPER invokes the SYSOPER privilege. Oracle Enterprise Manager Database Control does not permit you to log in as user SYS without connecting as SYSDBA or SYSOPER.

When you connect with the SYSDBA or SYSOPER privilege, you connect with a default schema, not with the schema that is generally associated with your user name. For SYSDBA this schema is SYS; for SYSOPER the schema is PUBLIC.

Explanation-3:

Difference between sys and system users:

(1) the data stored in the importance of the type;

[Sys] Oracle **data dictionary base tables and views** are stored in the sys user, these base tables and views for the operation of the oracle is critical, database maintenance, any user can not manually change

* Sys user has **dba, sysdba, sysoper roles** or permissions, the **highest users of Oracle permission.**

[System used to store **the second level** of internal data, such as some of the **characteristics** of the oracle **or tool management information.**

* System users with **ordinary dba role** permissions.

(2) Privileges.

System users **can only** landed **(as) normal identity** ORCL, unless you grant it sysdba system privileges or sysoper of system privileges.

Sys User **can** use **(as) SYSDBA or (as) the SYSOPER identity** Login ORCL, can not use normal.

Sys user login Oracle, perform **select * from V_\$ PWFILERS;**

Can query the user with sysdba privileges, such as:

SQL> select * from V_\$ PWFILERS;

the results are shown as follows:

USERNAME SYSDB sysop

SYS TRUE TRUE

Normal SYSDBA and sysoper,, three systems permission difference:

(1) normal, sysdba, sysoper What is the difference:

- 1) normal **ordinary users**
- 2) sysdba have **the highest system privileges**, after landing sys
- 3) sysoper mainly used to **start, shut down the database**, sysoper after landing the user is public
- 4) the sysdba and sysoper belongs to a **system privilege**, also known as the **Administrative privilege**, such as database permission to open the shut down like some system management level.

SYSDBA and SYSOPER specific permissions to the table below:

system normal as normal login, it is actually an **ordinary dba user**

If you are logged in as sysdba As a result, it is actually logged in as the **sys user** login information inside we can see it.

Principle: as sysdba connect to the database **objects created are actually generated in the sys**. Other users as well as sysdba login, but also as the sys user login.

See the following experiments:

SQL> create user strong identified by strong;

The user has been created.

SQL> conn strong / strong @ magick as sysdba;

Is connected.

SQL> show user;

USER is "SYS"

SQL> create table test (a int);

The table has been created.

SQL> select owner from dba_tables where table_name = 'test';

// Query from dba_tables the table (table_name) the owner (owner).

No rows selected // the oracle because when you create a table **automatically be converted to uppercase**, lowercase query does not exist;

SQL> select owner from dba_tables where table_name = 'TEST';

OWNER

SYS

40. What is the difference between view and materialized view?

Materialized views: Materialized views are disk based and update periodically base upon the query definition

Views: Views are virtual only and run the query definition each time they are accessed

Views are evaluating the data in the tables underlying the view definition at the time the view is queried. It is a logical view of your tables, with no data stored anywhere else. The upside of a view is that it will always return the latest data to you. The downside of a view is that its performance depends on how good a select statement the view is based on. If the select statement used by the view joins many tables, or uses joins based on non-indexed columns, this view can perform poorly.

Materialized views are similar to regular views, in that they are a logical view of your data (based on a select statement), however, the underlying query resultset has been saved to a table. The **upside** of this is that when you query a materialized view, you are querying a table, which may also be indexed. In addition, because all the joins have been resolved at materialized view refresh time, you pay the price of the join once (or as often as you refresh your materialized view), rather than each time you select from the materialized view. In addition, with query rewrite enabled, Oracle can optimize a query that selects from the source of your materialized view in such a way that it instead reads from your materialized view. In situations where you create materialized views as forms of aggregate tables, or as copies of frequently executed queries, this can greatly speed up the response time of your end user application The downside though is that the data you get back from the materialized view is only as up to date as the last time the materialized view has been refreshed.

Materialized views can be set to refresh manually, on a set schedule, or based on the database detecting a change in data from one of the underlying tables. Materialized views can be incrementally updated by combining them with materialized view logs, which act as change data capture sources on the underlying tables.

Materialized views are most often used in data warehousing / business intelligence applications where querying large fact tables with thousands of millions of rows would result in query response times that resulted in an unusable application.

Explanation-2:

View is logical, will store only the query, and will always gets latest data.

Mview is physical, will store the data, and may not get latest data.

41. What are materialized view refresh types and which is default?

Complete, fast, force (default)

COMPLETE	Refreshes by recalculating the materialized view's defining query.
FAST	Applies incremental changes to refresh the materialized view using the information logged in the materialized view logs, or from a SQL*Loader direct-path or a partition maintenance operation.
FORCE	Applies FAST refresh if possible; otherwise, it applies COMPLETE refresh.
NEVER	Indicates that the materialized view will not be refreshed with refresh mechanisms.

FAST:

Specify FAST to indicate the incremental refresh method, which performs the refresh according to the changes that have occurred to the master tables. The changes for conventional DML changes are stored in the materialized view log associated with the master table. The changes for direct-path INSERT operations are stored in the direct loader log.

If you specify REFRESH FAST, then the CREATE statement will fail unless materialized view logs already exist for the materialized view master tables. Oracle Database creates the direct loader log automatically when a direct-path INSERT takes place. No user intervention is needed.

For both conventional DML changes and for direct-path INSERT operations, other conditions may restrict the eligibility of a materialized view for fast refresh.

Materialized views are not eligible for fast refresh if the defining query contains an analytic function

COMPLETE:

Specify COMPLETE to indicate the complete refresh method, which is implemented by executing the defining query of the materialized view. If you request a complete refresh, then Oracle Database performs a complete refresh even if a fast refresh is possible.

FORCE:

Specify FORCE to indicate that when a refresh occurs, Oracle Database will perform a fast refresh if one is possible or a complete refresh if fast refresh is not possible. If you do not specify a refresh method (FAST, COMPLETE, or FORCE), then FORCE is the default.

42. How fast refresh happens?

43. How to find out when was a materialized view refreshed?

Query dba_mviews or dba_mview_analysis or dba_mview_refresh_times

```
SQL> select MVIEW_NAME, to_char(LAST_REFRESH_DATE, 'YYYY-MM-DD HH24:MI:SS') from dba_mviews;
```

(or)

```
SQL> select NAME, to_char(LAST_REFRESH, 'YYYY-MM-DD HH24:MI:SS') from dba_mview_refresh_times;
```

(or)

```
SQL> select MVIEW_NAME, to_char(LAST_REFRESH_DATE, 'YYYY-MM-DD HH24:MI:SS') from dba_mview_analysis;
```

44. What is materialized view log (type)?

45. What is atomic refresh in mviews?

From Oracle 10g, complete refresh of single materialized view can do delete instead of truncate. To force the refresh to do truncate instead of delete, parameter ATOMIC_REFRESH must be set to false.

ATOMIC_REFRESH = FALSE, mview will be truncated and whole data will be inserted. The refresh will go faster, and no undo will be generated.

ATOMIC_REFRESH = TRUE (default), mview will be deleted and whole data will be inserted. Undo will be generated.

We will have access at all times even while it is being refreshed.

```
SQL> EXEC DBMS_MVIEW.REFRESH ('mv_emp', 'C', atomic_refresh=FALSE);
```

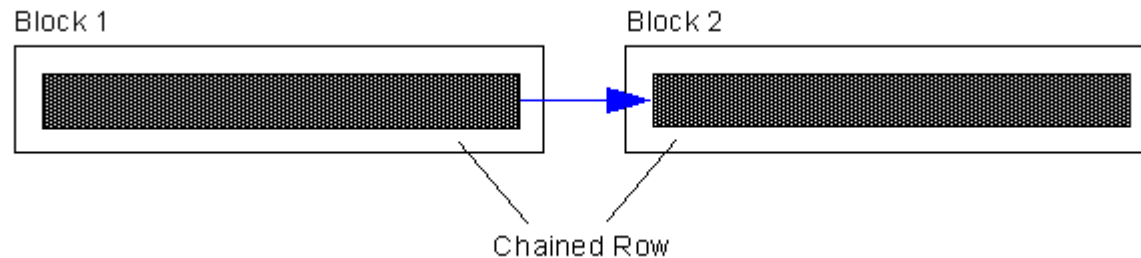
46. How to find out whether database/tablespace/datafile is in backup mode or not?

Query V\$BACKUP view.

47. What is row chaining?

Explanation-1: If the row is too large to fit into an empty data block in this case the oracle stores the data for the row in a chain of one or more data blocks, can occur when the row is inserted

Explanation-2: A row is too large to fit into a single database block. For example, if you use a 4KB blocksize for your database, and you need to insert a row of 8KB into it, Oracle will use 3 blocks and store the row in pieces. Some conditions that will cause row chaining are: Tables whose rowsize **exceeds** the blocksize. Tables with LONG and LONG RAW columns are prone to having chained rows. Tables with more than 255 columns will have chained rows as Oracle break wide tables up into pieces. So, instead of just having a forwarding address on one block and the data on another we have data on two or more blocks.



Chained rows affect us differently. Here, it depends on the data we need. If we had a row with two columns that was spread over two blocks, the query:

`SELECT column1 FROM table`

where column1 is in Block 1, would not cause any «table fetch continued row». It would not actually have to get column2, it would not follow the chained row all of the way out. On the other hand, if we ask for:

`SELECT column2 FROM table`

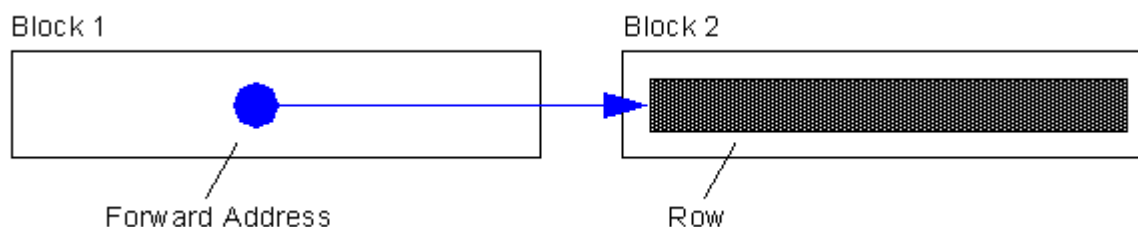
and column2 is in Block 2 due to row chaining, then you would in fact see a «table fetch continued row»

48. What is row migration?

Explanation-1: An update statement increases the amount of data in a row so that the row no longer fits in its data blocks. Now the oracle tries to find another free block with enough space to hold the entire row if such a block is available oracle moves entire row to new block.

Row Migration

Explanation-2: We will migrate a row when an update to that row would cause it to not fit on the block anymore (with all of the other data that exists there currently). A migration means that the entire row will move and we just leave behind the «forwarding address». So, the original block just has the rowid of the new block and the entire row is moved.



Full Table Scans are not affected by migrated rows

The forwarding addresses are ignored. We know that as we continue the full scan, we'll eventually get to that row so we can ignore the forwarding address and just process the row when we get there. Hence, in a full scan migrated rows don't cause us to really do any extra work -- they are meaningless.

Index Read will cause additional IO's on migrated rows

When we Index Read into a table, then a migrated row will cause additional IO's. That is because the index will tell us «goto file X, block Y, slot Z to find this row». But when we get there we find a message that says «well, really goto file A, block B, slot C to find this row». We have to do another IO (logical or physical) to find the row.

Scenario with Row migration and Row chaining:

Overview

If you notice poor performance in your Oracle database Row Chaining and Migration may be one of several reasons, but we can prevent some of them by properly designing and/or diagnosing the database.

Row Migration & Row Chaining are two potential problems that can be prevented. By suitably diagnosing, we can improve database performance. The main considerations are:

What is Row Migration & Row Chaining?

How to identify Row Migration & Row Chaining?

How to avoid Row Migration & Row Chaining?

Migrated rows affect OLTP systems which use indexed reads to read singleton rows. In the worst case, you can add an extra I/O to all reads which would be really bad. Truly chained rows affect index reads and full table scans.

Oracle Block

The Operating System Block size is the minimum unit of operation (read /write) by the OS and is a property of the OS file system. While creating an Oracle database we have to choose the «Data Base Block Size» as a multiple of the Operating System Block size. The minimum unit of operation (read /write) by the Oracle database would be this «Oracle block», and not the OS block. Once set, the «Data Base Block Size» cannot be changed during the life of the database (except in case of Oracle 9i). To decide on a suitable block size for the database, we take into consideration factors like the size of the database and the concurrent number of transactions expected.

The database block has the following structure (within the whole database structure)

Header

Header contains the general information about the data i.e. block address, and type of segments (table, index etc). It also contains the information about table and the actual row (address) which that holds the data.

Free Space

Space allocated for future update/insert operations. Generally affected by the values of **PCTFREE** and **PCTUSED** parameters.

Data

Actual row data.

FREELIST, **PCTFREE** and **PCTUSED**

While creating / altering any table/index, Oracle used two storage parameters for space control.

PCTFREE - The percentage of space reserved for future update of existing data.

PCTUSED - The percentage of minimum space used for insertion of new row data.

This value determines when the block gets back into the **FREELISTS** structure.

FREELIST - Structure where Oracle maintains a list of all free available blocks.

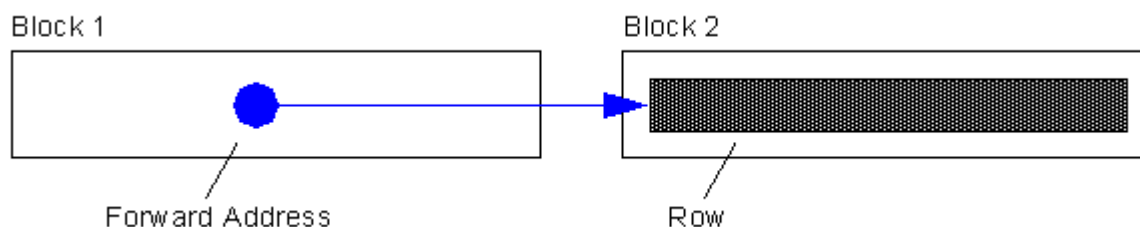
Oracle will first search for a free block in the **FREELIST** and then the data is inserted into that block. The availability of the block in the **FREELIST** is decided by the **PCTFREE** value. Initially an empty block will be listed in the **FREELIST** structure, and it will continue to remain there until the free space reaches the **PCTFREE** value.

When the free space reaches the **PCTFREE** value the block is removed from the **FREELIST**, and it is re-listed in the **FREELIST** table when the volume of data in the block comes below the **PCTUSED** value.

Oracle uses **FREELIST** to increase the performance. So for every insert operation, Oracle needs to search for the free blocks only from the **FREELIST** structure instead of searching all blocks.

Row Migration

We will migrate a row when an update to that row would cause it to not fit on the block anymore (with all of the other data that exists there currently). A migration means that the entire row will move and we just leave behind the «forwarding address». So, the original block just has the rowid of the new block and the entire row is moved.



Full Table Scans are not affected by migrated rows

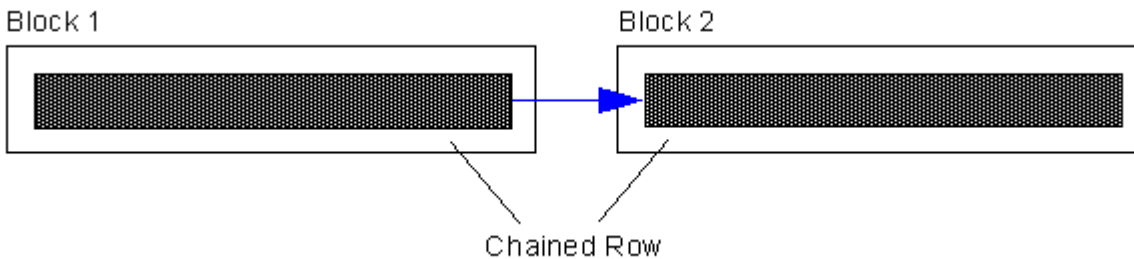
The forwarding addresses are ignored. We know that as we continue the full scan, we'll eventually get to that row so we can ignore the forwarding address and just process the row when we get there. Hence, in a full scan migrated rows don't cause us to really do any extra work -- they are meaningless.

Index Read will cause additional IO's on migrated rows

When we Index Read into a table, then a migrated row will cause additional IO's. That is because the index will tell us «goto file X, block Y, slot Z to find this row». But when we get there we find a message that says «well, really goto file A, block B, slot C to find this row». We have to do another IO (logical or physical) to find the row.

Row Chaining

A row is too large to fit into a single database block. For example, if you use a 4KB blocksize for your database, and you need to insert a row of 8KB into it, Oracle will use 3 blocks and store the row in pieces. Some conditions that will cause row chaining are: Tables whose rowsize **exceeds** the blocksize. Tables with LONG and LONG RAW columns are prone to having chained rows. Tables with more than 255 columns will have chained rows as Oracle break wide tables up into pieces. So, instead of just having a forwarding address on one block and the data on another we have data on two or more blocks.



Chained rows affect us differently. Here, it depends on the data we need. If we had a row with two columns that was spread over two blocks, the query:

```
SELECT column1 FROM table
```

where column1 is in Block 1, would not cause any «table fetch continued row». It would not actually have to get column2, it would not follow the chained row all of the way out. On the other hand, if we ask for:

```
SELECT column2 FROM table
```

and column2 is in Block 2 due to row chaining, then you would in fact see a «table fetch continued row»

Example

The following example was published by **Tom Kyte**, it will show row migration and chaining. We are using an 4k block size:

```
SELECT name,value
FROM v$parameter
WHERE name = 'db_block_size';
```

```
NAME          VALUE
```

```
-----
db_block_size 4096
```

Create the following table with CHAR fixed columns:

```
CREATE TABLE row_mig_chain_demo (
  x int PRIMARY KEY,
  a CHAR(1000),
  b CHAR(1000),
  c CHAR(1000),
  d CHAR(1000),
  e CHAR(1000)
);
```

That is our table. The CHAR(1000)'s will let us easily cause rows to migrate or chain. We used 5 columns a,b,c,d,e so that the total rowsize can grow to about 5K, bigger than one block, ensuring we can truly chain a row.

```
INSERT INTO row_mig_chain_demo (x) VALUES (1);
INSERT INTO row_mig_chain_demo (x) VALUES (2);
INSERT INTO row_mig_chain_demo (x) VALUES (3);
COMMIT;
```

We are not interested about seeing a,b,c,d,e - just fetching them. They are really wide so we'll suppress their display.

```
column a noprint
column b noprint
column c noprint
column d noprint
column e noprint
SELECT * FROM row_mig_chain_demo;
```

```
X
```

1
2
3

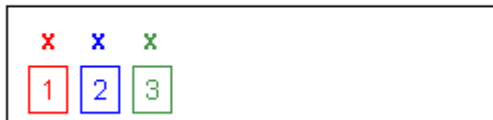
Check for chained rows:

```
SELECT a.name, b.value
FROM v$statname a, v$mystat b
WHERE a.statistic# = b.statistic#
AND lower(a.name) = 'table fetch continued row';
```

NAME	VALUE
table fetch continued row	0

Now that is to be expected, the rows came out in the order we put them in (Oracle full scanned this query, it processed the data as it found it). Also expected is the table fetch continued row is zero. This data is so small right now, we know that all three rows fit on a single block. No chaining.

Block 1



Demonstration of the Row Migration

Now, lets do some updates in a specific way. We want to demonstrate the row migration issue and how it affects the full scan:

```
UPDATE row_mig_chain_demo SET a = 'z1', b = 'z2', c = 'z3' WHERE x = 3;
COMMIT;
UPDATE row_mig_chain_demo SET a = 'y1', b = 'y2', c = 'y3' WHERE x = 2;
COMMIT;
UPDATE row_mig_chain_demo SET a = 'w1', b = 'w2', c = 'w3' WHERE x = 1;
COMMIT;
```

Note the order of updates, we did last row first, first row last.

```
SELECT * FROM row_mig_chain_demo;
```

X

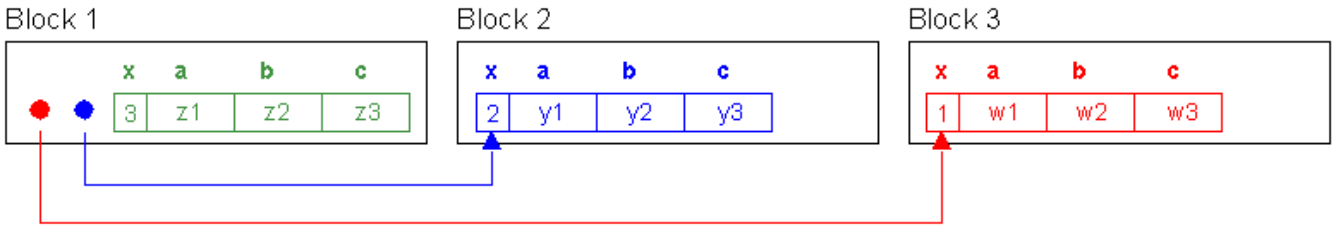
3
2
1

```
SELECT a.name, b.value
FROM v$statname a, v$mystat b
WHERE a.statistic# = b.statistic#
AND lower(a.name) = 'table fetch continued row';
```

NAME	VALUE
table fetch continued row	0

Interesting, the rows came out «backwards» now. That is because we updated row 3 first. It did not have to migrate, but it filled up block 1. We then updated row 2. It migrated to block 2 with row 3 hogging all of the space, it had to. We then updated row 1, it migrated to block 3. We migrated rows 2 and 1, leaving 3 where it started. So, when Oracle full scanned the table, it found row 3 on block 1 first, row 2 on block 2 second and row 1 on block 3 third. It ignored the head rowid piece on block 1 for rows 1 and 2 and just found the rows as it scanned the table. That is why the table fetch continued row is still zero. No chaining.

Migrated Rows



So, let's see a migrated row affecting the «table fetch continued row»:

```
SELECT * FROM row_mig_chain_demo WHERE x = 3;
```

```
X
-----
3
SELECT a.name, b.value
FROM v$statname a, v$mystat b
WHERE a.statistic# = b.statistic#
AND lower(a.name) = 'table fetch continued row';
NAME                                VALUE
-----
table fetch continued row            0
```

This was an index range scan / table access by rowid using the primary key. We didn't increment the «table fetch continued row» yet since row 3 isn't migrated.

```
SELECT * FROM row_mig_chain_demo WHERE x = 1;
```

```
X
-----
1
SELECT a.name, b.value
FROM v$statname a, v$mystat b
WHERE a.statistic# = b.statistic#
AND lower(a.name) = 'table fetch continued row';
NAME                                VALUE
-----
table fetch continued row            1
```

Row 1 is migrated, using the primary key index, we forced a «table fetch continued row».

Demonstration of the Row Chaining

```
UPDATE row_mig_chain_demo SET d = 'z4', e = 'z5' WHERE x = 3;
COMMIT;
```

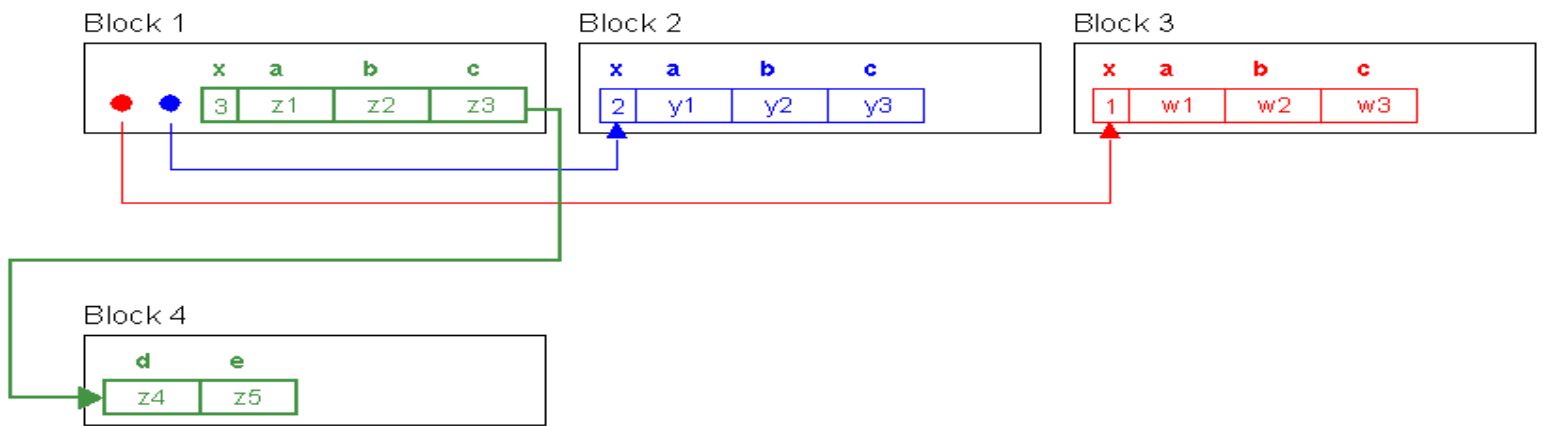
Row 3 no longer fits on block 1. With d and e set, the row size is about 5k, it is truly chained.

```
SELECT x,a FROM row_mig_chain_demo WHERE x = 3;
```

```
X
-----
3
SELECT a.name, b.value
FROM v$statname a, v$mystat b
WHERE a.statistic# = b.statistic#
AND lower(a.name) = 'table fetch continued row';
NAME                                VALUE
-----
table fetch continued row            1
```

We fetched column «x» and «a» from row 3 which are located on the «head» of the row, it will not cause a «table fetch continued row». No extra I/O to get it.

Chained Row 3



```
SELECT x,d,e FROM row_mig_chain_demo WHERE x = 3;
```

```
SELECT a.name, b.value
```

```
FROM v$statname a, v$mystat b
```

```
WHERE a.statistic# = b.statistic#
```

```
AND lower(a.name) = 'table fetch continued row';
```

```
NAME
```

```
VALUE
```

```
table fetch continued row
```

```
2
```

Now we fetch from the «tail» of the row via the primary key index. This increments the «table fetch continued row» by one to put the row back together from its head to its tail to get that data.

Now let's see a full table scan - it is affected as well:

```
SELECT * FROM row_mig_chain_demo;
```

```
X
```

```
-----
```

```
3
```

```
2
```

```
1
```

```
SELECT a.name, b.value
```

```
FROM v$statname a, v$mystat b
```

```
WHERE a.statistic# = b.statistic#
```

```
AND lower(a.name) = 'table fetch continued row';
```

```
NAME
```

```
VALUE
```

```
table fetch continued row
```

```
3
```

The «table fetch continued row» was incremented here because of Row 3, we had to assemble it to get the trailing columns. Rows 1 and 2, even though they are migrated don't increment the «table fetch continued row» since we full scanned.

```
SELECT x,a FROM row_mig_chain_demo;
```

```
X
```

```
-----
```

```
3
```

```
2
```

```
1
```

```
SELECT a.name, b.value
```

```
FROM v$statname a, v$mystat b
```

```
WHERE a.statistic# = b.statistic#
```

```
AND lower(a.name) = 'table fetch continued row';
```

```
NAME
```

```
VALUE
```

```
table fetch continued row
```

```
3
```

No «table fetch continued row» since we didn't have to assemble **Row 3**, we just needed the first two columns.

```
SELECT x,e FROM row_mig_chain_demo;
```

```

X
-----
3
2
1

```

```

SELECT a.name, b.value
FROM v$statname a, v$mystat b
WHERE a.statistic# = b.statistic#
AND lower(a.name) = 'table fetch continued row';

```

```

NAME                                     VALUE
-----
table fetch continued row                4

```

But by fetching for d and e, we incremented the «table fetch continued row». We most likely have only migrated rows but even if they are truly chained, the columns you are selecting are at the front of the table.

So, how can you decide if you have migrated or truly chained?

Count the last column in that table. That'll force to construct the entire row.

```
SELECT count(e) FROM row_mig_chain_demo;
```

```
COUNT(E)
```

```

-----
1

```

```

SELECT a.name, b.value
FROM v$statname a, v$mystat b
WHERE a.statistic# = b.statistic#
AND lower(a.name) = 'table fetch continued row';

```

```

NAME                                     VALUE
-----
table fetch continued row                5

```

Analyse the table to verify the chain count of the table:

```
ANALYZE TABLE row_mig_chain_demo COMPUTE STATISTICS;
```

```

SELECT chain_cnt
FROM user_tables
WHERE table_name = 'ROW_MIG_CHAIN_DEMO';

```

```

CHAIN_CNT
-----
3

```

Three rows that are chained. Apparently, 2 of them are migrated (Rows 1 and 2) and one is truly chained (Row 3).

Total Number of «table fetch continued row» since instance startup?

The **V\$SYSSTAT** view tells you how many times, since the system (database) was started you did a «table fetch continued row» over all tables.

```
sqlplus system/<password>
```

```

SELECT 'Chained or Migrated Rows = ' || value
FROM v$sysstat
WHERE name = 'table fetch continued row';

```

```
Chained or Migrated Rows = 31637
```

You could have 1 table with 1 chained row that was fetched 31'637 times. You could have 31'637 tables, each with a chained row, each of which was fetched once. You could have any combination of the above -- any combo.

Also, 31'637 - maybe that's good, maybe that's bad. it is a function of

how long has the database has been up

how many rows is this as a percentage of total fetched rows.

For example if 0.001% of your fetched are table fetch continued row, who cares!

Therefore, always compare the total fetched rows against the continued rows.

```
SELECT name,value FROM v$sysstat WHERE name like '%table%';
```


NAME	VALUE
-----	-----
table scans (short tables)	124338
table scans (long tables)	1485
table scans (rowid ranges)	0
table scans (cache partitions)	10
table scans (direct read)	0
table scan rows gotten	20164484
table scan blocks gotten	1658293
table fetch by rowid	1883112
table fetch continued row	31637
table lookup prefetch client count	0

How many Rows in a Table are chained?

The USER_TABLES tells you immediately after an ANALYZE (will be null otherwise) how many rows in the table are chained.

ANALYZE TABLE row_mig_chain_demo COMPUTE STATISTICS;

```
SELECT chain_cnt,
       round(chain_cnt/num_rows*100,2) pct_chained,
       avg_row_len, pct_free , pct_used
FROM user_tables
WHERE table_name = 'ROW_MIG_CHAIN_DEMO';
CHAIN_CNT PCT_CHAINED AVG_ROW_LEN PCT_FREE PCT_USED
-----
```

3	100	3691	10	40
---	-----	------	----	----

PCT_CHAINED shows 100% which means all rows are chained or migrated.

List Chained Rows

You can look at the chained and migrated rows of a table using the ANALYZE statement with the **LIST CHAINED ROWS** clause. The results of this statement are stored in a specified table created explicitly to accept the information returned by the LIST CHAINED ROWS clause. These results are useful in determining whether you have enough room for updates to rows.

Creating a CHAINED_ROWS Table

To create the table to accept data returned by an ANALYZE ... LIST CHAINED ROWS statement, execute the UTLCHAIN.SQL or UTLCHN1.SQL script in \$ORACLE_HOME/rdbms/admin. These scripts are provided by the database. They create a table named CHAINED_ROWS in the schema of the user submitting the script.

```
create table CHAINED_ROWS (
  owner_name      varchar2(30),
  table_name      varchar2(30),
  cluster_name    varchar2(30),
  partition_name  varchar2(30),
  subpartition_name varchar2(30),
  head_rowid      rowid,
  analyze_timestamp date
);
```

After a CHAINED_ROWS table is created, you specify it in the INTO clause of the ANALYZE statement.

ANALYZE TABLE row_mig_chain_demo LIST CHAINED ROWS;

```
SELECT owner_name,
       table_name,
       head_rowid
FROM chained_rows
```

OWNER_NAME	TABLE_NAME	HEAD_ROWID
-----	-----	-----
SCOTT	ROW_MIG_CHAIN_DEMO	AAAPVIAAFAAAAKiAAA
SCOTT	ROW_MIG_CHAIN_DEMO	AAAPVIAAFAAAAKiAAB

How to avoid Chained and Migrated Rows?

Increasing PCTFREE can help to avoid migrated rows. If you leave more free space available in the block, then the row has room to grow. You can also reorganize or re-create tables and indexes that have high deletion rates. If tables frequently have rows deleted, then data blocks can have partially free space in them. If rows are inserted and later expanded, then the inserted rows might land in blocks with deleted rows but still not have enough room to expand. Reorganizing the table ensures that the main free space is totally empty blocks.

The **ALTER TABLE ... MOVE** statement enables you to relocate data of a nonpartitioned table or of a partition of a partitioned table into a new segment, and optionally into a different tablespace for which you have quota. **This statement also lets you modify any of the storage attributes** of the table or partition, including those which cannot be modified using ALTER TABLE. You can also use the ALTER TABLE ... MOVE statement with the COMPRESS keyword to store the new segment using table compression.

1. ALTER TABLE MOVE

First count the number of Rows per Block before the ALTER TABLE MOVE

```
SELECT dbms_rowid.rowid_block_number(rowid) "Block-Nr", count(*) "Rows"
FROM row_mig_chain_demo
GROUP BY dbms_rowid.rowid_block_number(rowid) order by 1;
```

Block-Nr	Rows
2066	3

Now, **de-chain the table**, the ALTER TABLE MOVE rebuilds the row_mig_chain_demo table in a new segment, specifying new storage parameters:

```
ALTER TABLE row_mig_chain_demo MOVE
PCTFREE 20
PCTUSED 40
STORAGE (INITIAL 20K
NEXT 40K
MINEXTENTS 2
MAXEXTENTS 20
PCTINCREASE 0);
```

Table altered.

Again count the number of Rows per Block after the ALTER TABLE MOVE

```
SELECT dbms_rowid.rowid_block_number(rowid) "Block-Nr", count(*) "Rows"
FROM row_mig_chain_demo
GROUP BY dbms_rowid.rowid_block_number(rowid) order by 1;
```

Block-Nr	Rows
2322	1
2324	1
2325	1

2. Rebuild the Indexes for the Table

Moving a table changes the rowids of the rows in the table. This causes indexes on the table to be marked **UNUSABLE**, and DML accessing the table using these indexes will receive an **ORA-01502** error. The indexes on the table must be dropped or rebuilt. Likewise, any statistics for the table become invalid and new statistics should be collected after moving the table.

```
ANALYZE TABLE row_mig_chain_demo COMPUTE STATISTICS;
```

ERROR at line 1:

ORA-01502: index 'SCOTT.SYS_C003228' or partition of such index is in unusable state

This is the primary key of the table which must be rebuilt.

```
ALTER INDEX SYS_C003228 REBUILD;
```

Index altered.

```
ANALYZE TABLE row_mig_chain_demo COMPUTE STATISTICS;
Table analyzed.
SELECT chain_cnt,
       round(chain_cnt/num_rows*100,2) pct_chained,
       avg_row_len, pct_free , pct_used
FROM user_tables
WHERE table_name = 'ROW_MIG_CHAIN_DEMO';
CHAIN_CNT PCT_CHAINED AVG_ROW_LEN PCT_FREE PCT_USED
-----
1      33.33      3687      20      40
```

If the table includes LOB column(s), this statement can be used to move the table along with LOB data and LOB index segments (associated with this table) which the user explicitly specifies. If not specified, the default is to not move the LOB data and LOB index segments.

Detect all Tables with Chained and Migrated Rows

Using the CHAINED_ROWS table, you can find out the tables with chained or migrated rows.

1. Create the CHAINED_ROWS table
cd \$ORACLE_HOME/rdbms/admin
sqlplus scott/tiger
@utlchain.sql
2. Analyse all or only your Tables
SELECT 'ANALYZE TABLE ' || table_name || ' LIST CHAINED ROWS INTO CHAINED_ROWS;'
FROM user_tables
/

```
ANALYZE TABLE ROW_MIG_CHAIN_DEMO LIST CHAINED ROWS INTO CHAINED_ROWS;
ANALYZE TABLE DEPT LIST CHAINED ROWS INTO CHAINED_ROWS;
ANALYZE TABLE EMP LIST CHAINED ROWS INTO CHAINED_ROWS;
ANALYZE TABLE BONUS LIST CHAINED ROWS INTO CHAINED_ROWS;
ANALYZE TABLE SALGRADE LIST CHAINED ROWS INTO CHAINED_ROWS;
ANALYZE TABLE DUMMY LIST CHAINED ROWS INTO CHAINED_ROWS;
Table analyzed.
```

3. Show the RowIDs for all chained rows
This will allow you to quickly see how much of a problem chaining is in each table. If chaining is prevalent in a table, then that table should be rebuild with a higher value for PCTFREE

```
SELECT owner_name,
       table_name,
       count(head_rowid) row_count
FROM chained_rows
GROUP BY owner_name,table_name
/
OWNER_NAME      TABLE_NAME      ROW_COUNT
-----
SCOTT           ROW_MIG_CHAIN_DEMO      1
```

Conclusion:

Migrated rows affect OLTP systems which use indexed reads to read singleton rows. In the worst case, you can add an extra I/O to all reads which would be really bad. Truly chained rows affect index reads and full table scans. Row **migration** is typically caused by **UPDATE** operation

Row **chaining** is typically caused by **INSERT** operation.

SQL statements which are creating/querying these chained/migrated rows will degrade the performance due to more I/O work.

To diagnose chained/migrated rows use ANALYZE command , query V\$SYSSTAT view

To remove chained/migrated rows use higher PCTFREE using ALTER TABLE MOVE.

49. What are different types of partitions?

With Oracle8, Range partitioning (on single column) was introduced.

With Oracle8i, Hash and Composite (Range-Hash) partitioning was introduced.

With [Oracle9i](#), List partitioning and Composite (Range-List) partitioning was introduced.

With [Oracle 11g](#), Interval partitioning, Reference partitioning, Virtual column based partitioning, System partitioning and Composite partitioning [Range-Range, List-List, List-Range, List-Hash, Interval-Range, Interval-List, and Interval-Interval was introduced.

50. What is local partitioned index and global partitioned index?

A local index is an index on a partitioned table which is partitioned in the exact same manner as the underlying partitioned table. Each partition of a local index corresponds to one and only one partition of the underlying table.

A global partitioned index is an index on a partitioned or non partitioned tables which are partitioned using a different partitioning key from the table and can have different number of partitions. Global partitioned indexes can only be partitioned using [range partitioning](#).

51. How you will recover if you lost one/all control file(s)?

Lost one controlfile:

- a. Shut database
- b. Copy and rename the controlfile from the existing or mirror controlfile at os level 'OR'
- Remove the controlfile location from the pfile
- c. start the database

Lost of all controlfile: using the backup:

- a. shut the database (abort)
- b. startup the database in nomount state
- c. restore the controlfile from the autobackup
- d. open the database with resetlogs

Lost of all controlfile: without using the backup:

- a. create the controlfile manually with all the datafile locations
- b. mount the controlfile
- c. open the database with resetlogs

52. Why more archive logs are generated, when database is begin backup mode?

During begin backup mode datafile headers get freeze and as result row information cannot be retrieved as a result the entire block is copied to redo logs as a result more redo generated and more log switch and in turn more archive logs. Normally only details (change vectors) are logged to the redo logs. When in backup mode, Oracle will write complete changed blocks to the redo log files.

Mainly to overcome fractured blocks. Most of the cases Oracle block size is equal to or a multiple of the operating system block size.

e.g. Consider Oracle blocksize is 2k and OSBlocksize is 4k. so each OS Block is comprised of 2 Oracle Blocks. Now you are doing an update when your db is in backup mode. An Oracle Block is updating and at the same time backup is happening on the OS block which is having this particular DB block. Backup will not be consistent since the one part of the block is being updated and at the same time it is copied to the backup location. In this case we will have a fractured block, so as to avoid this Oracle will copy the whole OS block to redo logfile which can be used for recovery. Because of this redo generation is more.

53. What UNIX parameters you will set while Oracle installation?

shmmax, shmmni, shmall, sem

SHMMAX and SHMALL are two key shared memory parameters that directly impact's the way by which Oracle creates an SGA. Shared memory is nothing but part of Unix IPC System (Inter Process Communication) maintained by kernel where multiple processes share a single chunk of memory to communicate with each other.

While trying to create an SGA during a database startup, Oracle chooses from one of the 3 memory management models a) one-segment or b) contiguous-multi segment or c) non-contiguous multi segment. Adoption of any of these models is dependent on the size of SGA and values defined for the shared memory parameters in the linux kernel, most importantly SHMMAX.

So what are these parameters - SHMMAX and SHMALL?

SHMMAX is the **maximum size** of a **single shared memory segment** set in “**bytes**”.

```
silicon:~ # cat /proc/sys/kernel/shmmax
```

```
536870912
```

SHMALL is the **total size of Shared Memory Segments System wide** set in “**pages**”.

```
silicon:~ # cat /proc/sys/kernel/shmall
```

```
1415577
```

The key thing to note here is the value of SHMMAX is set in “**bytes**” but the value of SHMALL is set in “**pages**”.

What’s the optimal value for SHMALL?

As SHMALL is the total size of Shared Memory Segments System wide, it should always be less than the Physical Memory on the System and should also be less than sum of SGA’s of all the oracle databases on the server. Once this value (sum of SGA’s) hit the limit, i.e. the value of shmall, then any attempt to start a new database (or even an existing database with a resized SGA) will result in an “out of memory” error (below). This is because there won’t be any more shared memory segments that Linux can allocate for SGA.

```
ORA-27102: out of memory
```

```
Linux-x86_64 Error: 28: No space left on device.
```

So above can happen for two reasons. Either the value of shmall is not set to an optimal value or you have reached the threshold on this server.

Setting the value for SHMALL to optimal is straight forward. All you want to know is how much “Physical Memory” (excluding Cache/Swap) you have on the system and how much of it should be set aside for Linux Kernel and to be dedicated to Oracle Databases.

For e.g. Let say the Physical Memory of a system is 6GB, out of which you want to set aside 1GB for Linux Kernel for OS Operations and dedicate the rest of 5GB to Oracle Databases. Then here’s how you will get the value for SHMALL. Convert this 5GB to bytes and divide by page size. Remember SHMALL should be set in “pages” not “bytes”.

So here goes the calculation.

Determine Page Size first, can be done in two ways. In my case it’s 4096 and that’s the recommended and default in most cases which you can keep the same.

```
silicon:~ # getconf PAGE_SIZE
```

```
4096
```

or

```
silicon:~ # cat /proc/sys/kernel/shmmni
```

```
4096
```

Convert 5GB into bytes and divide by page size, I used the linux calc to do the math.

```
silicon:~ # echo "( 5 * 1024 * 1024 * 1024 ) / 4096 " | bc -l
```

```
1310720.000000000000000000000000
```

Reset shmall and load it dynamically into kernel

```
silicon:~ # echo "1310720" > /proc/sys/kernel/shmall
```

```
silicon:~ # sysctl -p
```

Verify if the value has been taken into effect.

```
silicon:~ # sysctl -a | grep shmall
```

```
kernel.shmall = 1310720
```

Another way to look this up is

```
silicon:~ # ipcs -lm
```

```
----- Shared Memory Limits -----
```

```
max number of segments = 4096          /* SHMMNI */
```

```
max seg size (kbytes) = 524288         /* SHMMAX */
```

```
max total shared memory (kbytes) = 5242880 /* SHMALL */
```

```
min seg size (bytes) = 1
```

To keep the value effective after every reboot, add the following line to /etc/sysctl.conf

```
echo "kernel.shmall = 1310720" >> /etc/sysctl.conf
```

Also verify if sysctl.conf is enabled or will be read during boot.

```
silicon:~ # chkconfig boot.sysctl
```

```
boot.sysctl on
```

If returns “off”, means it’s disabled. Turn it on by running

```
silicon:~ # chkconfig boot.sysctl on
```

boot.sysctl on

What's the optimal value for SHMMAX?

Oracle makes use of one of the 3 memory management models to create the SGA during database startup and it does this in following sequence. First Oracle attempts to use the one-segment model and if this fails, it proceeds with the next one which's the contiguous multi-segment model and if that fails too, it goes with the last option which is the non-contiguous multi-segment model.

So during startup it looks for **shmmax** parameter and compares it with the initialization parameter ***.sga_target**. If **shmmax > *.sga_target**, then oracle goes with **one-segment** model approach where the entire SGA is created within a single shared memory segment.

But the above attempt (one-segment) fails if SGA size otherwise ***.sga_target > shmmax**, then Oracle proceeds with the 2nd option – contiguous multi-segment model. Contiguous allocations, as the name indicates are a set of shared memory segments which are contiguous within the memory and if it can find such a set of segments then entire SGA is created to fit in within this set.

But if cannot find a set of contiguous allocations then last of the 3 option's is chosen – non-contiguous multi-segment allocation and in this Oracle has to grab the free memory segments fragmented between used spaces.

So let's say if you know the max size of SGA of any database on the server stays below 1GB, you can set shmmax to 1 GB. But say if you have SGA sizes for different databases spread between 512MB to 2GB, then set shmmax to 2Gigs and so on.

Like SHMALL, SHMMAX can be defined by one of these methods..

Dynamically reset and reload it to the kernel..

```
silicon:~ # echo "536870912" > /proc/sys/kernel/shmmax
```

```
silicon:~ # sysctl -p          -- Dynamically reload the parameters.
```

Or use sysctl to reload and reset ..

```
silicon:~ # sysctl -w kernel.shmmax=536870912
```

To permanently set so it's effective in reboots...

```
silicon:~ # echo "kernel.shmmax=536870912" >> /etc/sysctl.conf
```

Install doc for 11g recommends the value of shmmax to be set to "4GB – 1byte" or half the size of physical memory whichever is lower. I believe "4GB – 1byte" is related to the limitation on the 32 bit (x86) systems where the virtual address space for a user process can only be little less than 4GB. As there's no such limitation for 64bit (x86_64) bit systems, you can define SGA's larger than 4 Gig's. But idea here is to let Oracle use the efficient one-segment model and for this shmmax should stay higher than SGA size of any individual database on the system.

54. What is the use of initrans and maxtrans in table definition?

Initial and Maximum transactions allowed to read/write to the block concurrently

INITRANS specifies the number of DML transaction entries for which space is initially reserved in the data block header. Space is reserved in the headers of all data blocks in the associated segment. As multiple transactions concurrently access the rows of the same data block, space is allocated for each DML transaction's entry in the block. Once the space reserved by INITRANS is depleted, space for additional transaction entries is allocated out of the free space in a block, if available. Once allocated, this space effectively becomes a permanent part of the block header. The MAXTRANS parameter limits the number of transaction entries that can concurrently use data in a data block. Therefore, you can limit the amount of free space that can be allocated for transaction entries in a data block using MAXTRANS.

The INITRANS and MAXTRANS parameters for the data blocks allocated to a specific schema object should be set individually for each schema object based on the following criteria:

The space you would like to reserve for transaction entries compared to the space you would reserve for database data

The number of concurrent transactions that are likely to touch the same data blocks at any given time

For example, if a table is very large and only a small number of users simultaneously access the table, the chances of multiple concurrent transactions requiring access to the same data block is low. Therefore, INITRANS can be set low, especially if space is at a premium in the database.

Alternatively, assume that a table is usually accessed by many users at the same time. In this case, you might consider preallocating transaction entry space by using

a high INITRANS. This eliminates the overhead of having to allocate transaction entry space, as required when the object is in use. Also, allow a higher MAXTRANS so that no user has to wait to access necessary data blocks.

INITRANS and MAXTRANS are used when you expect multiple access to the same data block.

Every transaction which modifies a block must acquire an entry in the Interested Transaction List (ITL). Space for this list is defined by INITRANS. The ITL grows dynamically as needed by transactions up to the value MAXTRANS. It also shrinks back down to the setting for INITRANS.

INITRANS

The default value is 1 for tables and 2 for clusters and indexes.

MAXTRANS

The default value is an operating system-specific function of block size, not exceeding 255.

55. What are differences between dbms_job and dbms_scheduler?

Through dbms_scheduler we can schedule OS level jobs also.

56. What are differences between dbms_scheduler and cron jobs?

Through dbms_scheduler we can schedule database and os level jobs, but through cron we can't set.

57. Difference between CPU & PSU patches?

CPU - Critical Patch Update - includes only Security related patches.

PSU - Patch Set Update - includes CPU + other patches deemed important enough to be released prior to a minor (or major) version release.

58. What you will do if (local) inventory corrupted [or] opatch lsinventory is giving error?

What to do if my Global Inventory is corrupted?

If your global Inventory is corrupted, you can recreate global Inventory on machine using Universal Installer and attach already installed oracle home by option

-attachHome

```
./runInstaller -silent -attachHome -invPtrLoc $location_to_oralnst.loc ORACLE_HOME=Oracle_Home_Location  
ORACLE_HOME_NAME=Oracle_Home_Name CLUSTER_NODES={}
```

59. What are the entries/location of oralnst.loc?

/etc/oralnst.loc is pointer to central/local Oracle Inventory.

60. What is the difference between central/global inventory and local inventory?

Overview of Inventory

The inventory is a very important part of the Oracle Universal Installer. This is where OUI keeps all information regarding the products installed on a specific machine.

There are two inventories with the newer releases of OUI (2.x and higher):

- * The inventory in the ORACLE_HOME (Local Inventory)
- * The central inventory directory outside the ORACLE_HOME (Global Inventory)

At startup, the Oracle Universal Installer first looks for the key that specifies where the global inventory is located at (this key varies by platform).

- * /var/opt/oracle/oralnst.loc (typical)
- * /etc/oralnst.loc (AIX and Linux)
- * HKEY_LOCAL_MACHINE -> Software -> Oracle -> INST_LOC (Windows platforms)

If this key is found, the directory within it will be used as the global inventory location.

If the key is not found, the inventory path will default created as follows:

- * UNIX : ORACLE_BASE\oraInventory
- * WINDOWS : c:\program files\oracle\Inventory

If the ORACLE_BASE environment variable is not defined, the inventory is created at the same level as the first Oracle home. For example, if your first Oracle home is at /private/ORACLEHome1, then, the inventory is at /private/oraInventory.

With Oracle Applications 11i the inventory contains information about both the iAS and RDBMS ORACLE_HOMES
About the Oracle Universal Installer Inventory

The Oracle Universal Installer inventory is the location for the Oracle Universal Installer's bookkeeping. The inventory stores information about:

- * All Oracle software products installed in all Oracle homes on a machine
- * Other non-ORACLE_HOME specific products, such as the Java Runtime Environment (JRE)

Starting with Oracle Universal Installer 2.1, the information in the Oracle Universal Installer inventory is stored in Extensible Markup Language (XML) format. The XML format allows for easier diagnosis of problems and faster loading of data. Any secure information is not stored directly in the inventory. As a result, during deinstallation of some products, you may be prompted for required secure information, such as passwords.

By default, the Universal Installer inventory is located in a series of directories at /Program Files/Oracle/Inventory on Windows computers and in the /Inventory directory on UNIX computers.

Local Inventory

There is one Local Inventory per ORACLE_HOME. It is physically located inside the ORACLE_HOME at \$ORACLE_HOME/inventory and contains the detail of the patch level for that ORACLE_HOME.

The Local Inventory gets updated whenever a patch is applied to the ORACLE_HOME, using OUI.

If the Local Inventory becomes corrupt or is lost, this is very difficult to recover, and may result in having to reinstall the ORACLE_HOME and re-apply all patchsets and patches.

Global Inventory

The Global Inventory is the part of the XML inventory that contains the high level list of all oracle products installed on a machine. There should therefore be only one per machine. Its location is defined by the content of oraInst.loc. The Global Inventory records the physical location of Oracle products installed on the machine, such as ORACLE_HOMES (RDBMS and IAS) or JRE. It does not have any information about the detail of patches applied to each ORACLE_HOMES.

The Global Inventory gets updated every time you install or de-install an ORACLE_HOME on the machine, be it through OUI Installer, Rapid Install, or Rapid Clone.

Note: If you need to delete an ORACLE_HOME, you should always do it through the OUI de-installer in order to keep the Global Inventory synchronized.

61. What is the use of root.sh & oraInstRoot.sh?

Explanation-1:

Changes ownership & permissions of oraInventory

Creating oratab file in the /etc directory

In RAC, starts the clusterware stack

Note: Both the script should be run as root user

oraInstRoot.sh:

It is located in \$ORACLE_BASE/oraInventory

Usage:

- It creates the inventory pointer file (/etc/oraInst.loc), The file shows the inventory location and group it is linked to.
- Changing groupname of the oraInventory directory to oinstall group

root.sh:

It is located in \$ORACLE_HOME directory

Usage:

root.sh script performs many things, namely

- It changes or correctly sets the environment variables
- copying of few files into /usr/local/bin , the files are dbhome, oraenv, coraenv etc.
- creation of /etc/oratab file or adding database home and SID's entry into /etc/oratab file.

62. What is transportable tablespace (and across platforms)?

http://docs.oracle.com/cd/B28359_01/backup.111/b28270/rcmttbsb.htm

Overview

Oracle's *Transportable Tablespace* is one of those much awaited features that was introduced in Oracle8i (8.1.5) and is commonly used in Data Warehouses (DW). Using transportable tablespaces is much faster than using other utilities like export/import, SQL*Plus copy tables, or backup and recovery options to copy data from one database to another. This article provides a brief introduction into configuring and using transportable tablespaces.

Introduction to Transportable Tablespaces

Before covering the details of how to setup and use transportable tablespaces, let's first discuss some of the terminology and limitations to provide us with an introduction.

The use of transportable tablespaces are much faster than using export/import, SQL*Plus copy tables, or backup and recovery options to copy data from one database to another.

A transportable tablespace set is defined as two components:

All of the datafiles that make up the tablespaces that will be moved AND an export that contains the data dictionary information about those tablespaces.

COMPATIBLE must be set in both the source and target database to at least 8.1.

When transporting a tablespace from an OLTP system to a data warehouse using the Export/Import utility, you will most likely NOT need to transport TRIGGER and CONSTRAINT information that is associated with the tables in the

tablespace you are exporting. That is, you will set the TRIGGERS and CONSTRAINTS Export utility parameters equal to "N".

The data in a data warehouse is inserted and altered under very controlled circumstances and does not require the same usage of constraints and triggers as a typical operational system does.

It is common and recommended though that you use the GRANTS option by setting it to Y.

The TRIGGERS option is new in Oracle8i for use with the export command. It is used to control whether trigger information, associated with the tables in a tablespace, are included in the tablespace transport.

Limitations of Transportable Tablespaces:

The transportable set must be self-contained.

Both the source and target database must be running Oracle 8.1 or higher release.

The two databases do not have to be on the same release

The source and target databases must be on the same type of hardware and operating-system platform.

The source and target databases must have the same database block size.

The source and target databases must have the same character set.

A tablespace with the same name must not already exist in the target database.

Materialized views, function-based indexes, scoped REFs, 8.0 compatible advanced queues with multiple-recipients, and domain indexes can't be transported in this manner. (As of Oracle8i)

Users with tables in the exported tablespace should exist in the target database prior to initiating the import. Create the user reported by the error message.

Explanation: The metadata exported from the target database does not contain enough information to create the user in the target database. The reason is that, if the metadata contained the user details, it might overwrite the privileges of an existing user in the target database.

(i.e. If the user by the same name already exists in the target database)

By not maintaining the user details, we preserve the security of the database.

Using Transportable Tablespaces

In this section, we finally get to see how to use transportable tablespaces. Here is an overview of the steps we will perform in this section:

Verify that the set of source tablespaces are self-contained

Generate a transportable tablespace set.

Transport the tablespace set

Import the tablespaces set into the target database.

In this example, we will be transporting the tablespaces, "FACT1, FACT2, and FACT_IDX" from a database named DWDB to REPORTDB. The user that owns these tables will be "DW" and password "DW".

Verify Self-Contained Status with the DBMS_TTS Package

To verify that all tablespaces to transport are *self-contained*, we can use the TRANSPORT_SET_CHECK procedure within the DBMS_TTS PL/SQL Package. The first parameter to this procedure is a list of the tablespaces to transport. Keep in mind that all indexes for a table, partitions, and LOB column segments in the tablespace must also reside in the tablespace set. The second parameter to this procedure is a boolean value that indicates whether or not to check for referential integrity.

```
SQL> connect sys/change_on_install@dwdb as sysdba
```

```
SQL> exec DBMS_TTS.TRANSPORT_SET_CHECK('fact1, fact2', TRUE);
```

```
SQL> SELECT * FROM TRANSPORT_SET_VIOLATIONS;
```

```
VIOLATIONS
```

```
-----
```

```
Index DW.DEPT_PK in tablespace FACT_IDX enforces primary constraints of table D
```

```
W.DEPT in tablespace FACT1
```

```
Index DW.EMP_PK in tablespace FACT_IDX enforces primary constraints of table DW
```

```
.EMP in tablespace FACT1
```

OOOPS! As we can see from the above example, I forgot to include all tablespaces that will make this self-contained.

In this example, I forgot to include the FACT_IDX tablespace. Let's correct that:

```
SQL> exec DBMS_TTS.TRANSPORT_SET_CHECK('fact1, fact2, fact_idx', TRUE);
```

```
SQL> SELECT * FROM TRANSPORT_SET_VIOLATIONS;
```

```
no rows selected
```

Generate a Transportable Tablespace Set

To generate a *Transportable Tablespace Set*, you will need to perform the following:

Place all tablespace within the tablespace set in READ ONLY mode.

Use Export to gather tablespace data-dictionary information.

Copy datafiles and the export dump from the source location to the target location.

Place all tablespace within the tablespace set back to READ/WRITE.

```
% sqlplus "sys/change_on_install@dwdb as sysdba"
```

```
SQL> ALTER TABLESPACE fact1 READ ONLY;
```

```
SQL> ALTER TABLESPACE fact2 READ ONLY;
```

```
SQL> ALTER TABLESPACE fact_idx READ ONLY;
```

```
SQL> exit
```

```
% exp
```

```
  userid=\"sys/change_on_install@dwdb as sysdba\"
```

```
transport_tablespace=y
```

```
tablespaces=fact1, fact2, fact_idx
```

```
triggers=y
```

```
constraints=y
```

```
grants=y
```

```
file=fact_dw.dmp
```

```
% cp /u10/app/oradata/DWDB/fact1_01.dbf /u10/app/oradata/REPORTDB/fact1_01.dbf
```

```
% cp /u10/app/oradata/DWDB/fact2_01.dbf /u10/app/oradata/REPORTDB/fact2_01.dbf
```

```
% cp /u09/app/oradata/DWDB/fact_idx01.dbf /u09/app/oradata/REPORTDB/fact_idx01.dbf
```

```
% sqlplus "sys/change_on_install@dwdb as sysdba"
```

```
SQL> ALTER TABLESPACE fact1 READ WRITE;
```

```
SQL> ALTER TABLESPACE fact2 READ WRITE;
```

```
SQL> ALTER TABLESPACE fact_idx READ WRITE;
```

```
SQL> exit
```

Transport the Tablespace Set

To actually *transport* the tablespace, this is nothing more than copying (or FTP'ing) all tablespace set datafiles to be put in their proper location on the target database. In the section previous to this, we did that with the **cp** command in UNIX.

In some cases this would be necessary if the files were copied off to a staging area in the previous step.

Import the Tablespace Set

Before actually importing the tablespace(s) into the target database, you will need to ensure that all users that own segments in the imported tablespaces exist. For this example, the only user that owns segments in the exported tablespaces is DW. I will create this user:

```
% sqlplus "sys/change_on_install@reportdb as sysdba"
```

```
SQL> create user dw identified by dw default tablespace users;
```

```
SQL> grant dba, resource, connect to dw;
```

```
SQL> exit
```

We now use the Import utility to bring the tablespace set's data-dictionary information into the target database.

The two required parameters are TRANSPORT_TABLESPACE=Y and DATAFILES='...' as in the following example:

```
% imp
```

```
  userid=\"sys/change_on_install@reportdb as sysdba\"
```

```
transport_tablespace=y
```

```
datafiles='/u10/app/oradata/REPORTDB/fact1_01.dbf', '/u10/app/oradata/REPORTDB/fact2_01.dbf',
```

```
 '/u09/app/oradata/REPORTDB/fact_idx01.dbf'
```

```
file=fact_dw.dmp
```

Final Cleanup

When the tablespaces are successfully imported into the target database, they are in READ ONLY mode. If you intend to use the tablespaces for READ WRITE, you will need to manually alter them:

```
% sqlplus "sys/change_on_install@reportdb as sysdba"
```

```
SQL> ALTER TABLESPACE fact1 READ WRITE;
SQL> ALTER TABLESPACE fact2 READ WRITE;
SQL> ALTER TABLESPACE fact_idx READ WRITE;
SQL> exit
```

Explanation-2:

You can use the transportable tablespaces feature to move a subset of an Oracle database and "plug" it in to another Oracle database, essentially moving tablespaces between the databases. The tablespaces being transported can be either dictionary managed or locally managed. Starting with Oracle9i, the transported tablespaces are not required to be of the same block size as the target database's standard block size. Transporting tablespaces is particularly useful for:

Moving data from OLTP systems to data warehouse staging systems
 Updating data warehouses and data marts from staging systems
 Loading data marts from central data warehouses
 Archiving OLTP and data warehouse systems efficiently
 Data publishing to internal and external customers
 Performing Tablespace Point-in-Time Recovery (TSPITR)
 Moving data using transportable tablespaces can be much faster than performing either an export/import or unload/load of the same data, because transporting a tablespace only requires the copying of datafiles and integrating the tablespace structural information. You can also use transportable tablespaces to move index data, thereby avoiding the index rebuilds you would have to perform when importing or loading table data.

LIMITATIONS

Be aware of the following limitations as you plan for transportable tablespace use:

The source and target database must be on the same hardware platform. For example, you can transport tablespaces between Sun Solaris Oracle databases, or you can transport tablespaces between Windows NT Oracle databases. However, you cannot transport a tablespace from a Sun Solaris Oracle database to an Windows NT Oracle database. The source and target database must use the same character set and national character set. You cannot transport a tablespace to a target database in which a tablespace with the same name already exists. Transportable tablespaces do not support: Materialized views/replication Function-based indexes.

63. How can you transport tablespaces across platforms with different endian formats?

RMAN

64. What is xtss (cross platform transportable tablespace)?

A transportable tablespace allows you to quickly move a subset of an Oracle database from one Oracle database to another. However, in the previous release of Oracle server, you can only move a tablespace across Oracle databases within the same platform.

Oracle 10g is going one step further by allowing you to move tablespace across different platforms.

Benefits

One of the major benefits for organizations that hosts Oracle databases on different platforms is that data can now be moved between databases quickly, across different platforms. Using the new cross-platform transportable tablespaces method to move data is more efficient than the traditional method of export and import.

Supported Platforms and New Data Dictionary Views

Oracle Database 10g supports nine platforms for transportable tablespace.

A new data dictionary view, *v\$transportable_platform*, lists all nine supported platforms, along with platform ID and endian format.

PLATFORM_ID	PLATFORM_NAME	ENDIAN_FORMAT
1	Solaris[tm] OE (32-bit)	Big
2	Solaris[tm] OE (64-bit)	Big
3	HP-UX (64-bit)	Big
4	HP-UX IA (64-bit)	Big
5	HP Tru64 UNIX	Little
6	AIX-Based Systems (64-bit)	Big
7	Microsoft Windows NT	Little
8	Linux IA (32-bit)	Little
9	Linux IA (64-bit)	Little

Table 3.3: Supported platforms for transportable tablespaces.

The v\$database data dictionary view also adds two columns, platform ID and platform name:

```
SQL> select name, platform_id, platform_name
       2 from v$database;
NAME    PLATFORM_ID  PLATFORM_NAME
-----  -
```

```
GRID      2    Solaris[tm] OE (64-bit)
```

To transport a tablespace from one platform to another, datafiles on different platforms must be in the same endian format (byte ordering).

The pattern for byte ordering in native types is called endianness. There are only two main patterns, big endian and little endian. Big endian means the most significant byte comes first, and little endian means the least significant byte comes first. If the source platform and the target platform are of different endianness, then an additional step must be taken on either the source or target platform to convert the tablespace being transported to the target format. If they are of the same endianness, then no conversion is necessary and tablespaces can be transported as if they were on the same platform.

Be aware of the following limitations as you plan for transportable tablespace use:

- The source and target database must use the same character set and national character set.
- You cannot transport a tablespace to a target database in which a tablespace with the same name already exists. However, you can rename either the tablespace to be transported or the destination tablespace before the transport operation.
- The set should be self-containing

Convert Datafiles using RMAN

You do not need to convert the datafile to transport a tablespace from an AIX-based platform to a Sun platform, since both platforms use a big endian.

However, to transport a tablespace from a Sun platform (big endian) to a Linux platform (little endian), you need to use the CONVERT command in the RMAN utility to convert the byte ordering. This can be done on either the source platform or the target platform.

```
RMAN> CONVERT TABLESPACE 'USERS'
```

```
      TO PLATFORM = 'Linux IA (32-bit)'
```

```
      DB_FILE_NAME_CONVERT = '/u02/oradata/grid/users01.dbf', '/dba/recovery_area/transport_linux'
```

The limitation requiring transportable tablespaces to be transferred to the same operating system has been removed. However, to transport tablespaces across different platforms, both the source and target databases must be at least on Oracle Database 10g, be on at least version 10.0.1, and have the COMPATIBLE initialization parameter set to 10.0.

Transporting Tablespaces Between Databases: A General Procedure

Perform the following steps to move or copy a set of tablespaces.

- You must pick a self-contained set of tablespaces. Verify this using the dbms_tts.transport_set_check package.
- Next, generate a transportable tablespace set, using the Export utility.
- A transportable tablespace set consists of the set of datafiles for the set of tablespaces being transported and an Export file containing metadata information for the set of tablespaces.
- Transporting a tablespace set to a platform different from the source platform will require connection to the Recovery Manager (RMAN) and invoking the CONVERT command. An alternative is to do the conversion on the target platform after the tablespace datafiles have been transported.
- The final step is to plug in the tablespace - You use the Import utility to plug the set of tablespaces metadata, and hence the tablespaces themselves, into the target database.

If you are transporting these tablespaces to a different platform, use the v\$(platform) view to find the platform name. You can then use the Recovery Manager CONVERT command to perform the conversion.

Note - As an alternative to conversion before transport, the CONVERT command can be used for the conversion on the target platform after the tablespace set has been transported.

The limitation requiring transportable tablespaces to be transferred to the same operating system has been removed. However, to transport tablespaces across different platforms, both the source and target databases must be at least on Oracle Database 10g, be on at least version 10.0.1, and have the COMPATIBLE initialization parameter set to 10.0.

Transporting Tablespaces Between Databases: A General Procedure

Perform the following steps to move or copy a set of tablespaces.

- You must pick a self-contained set of tablespaces. Verify this using the `dbms_tts.transport_set_check` package.
- Next, generate a transportable tablespace set, using the Export utility.
- A transportable tablespace set consists of the set of datafiles for the set of tablespaces being transported and an Export file containing metadata information for the set of tablespaces.
- Transporting a tablespace set to a platform different from the source platform will require connection to the Recovery Manager (RMAN) and invoking the CONVERT command. An alternative is to do the conversion on the target platform after the tablespace datafiles have been transported.
- The final step is to plug in the tablespace - You use the Import utility to plug the set of tablespaces metadata, and hence the tablespaces themselves, into the target database.

If you are transporting these tablespaces to a different platform, use the `v$platform` view to find the platform name. You can then use the Recovery Manager CONVERT command to perform the conversion.

Note - As an alternative to conversion before transport, the CONVERT command can be used for the conversion on the target platform after the tablespace set has been transported.

65. What is the difference between restore point & guaranteed restore point?

Sometimes it is more efficient to roll-back changes in a database rather than do a point-in-time recovery. Flashback Database has the ability to rewind the entire database and changes that occurred within a given time window. The effects are similar to database point-in-time recovery.

Normal Restore Points

You can create restore points to enable you to Flashback the database to a particular point in time or SCN. You can think of it as a bookmark or alias that can be used with commands that recognize a RESTORE POINT clause as shorthand for specifying an SCN. In essence before you perform any operations that you may have to reverse you can create a normal restore point. The name of the restore point and SCN are then recorded within the control file. So basically creating a restore point eliminates the need to determine the current SCN before performing an operations or finding the proper one after the fact. You can use RESTORE POINTS to specify the target SCN in the following contexts:

RECOVER DATABASE and FLASHBACK DATABASE commands within RMAN

FLASHBACK TABLE in SQL*Plus

Guaranteed Restore Points (GRP)

A Guaranteed Restore Point can be used to perform a Flashback Database operation even if flashback logging is not enabled for your database. It can be used to revert a whole database to a known good state days or weeks ago, as long as there is enough disk space in flash recovery area to store the needed logs. Even effects of NOLOGGING operations like direct load inserts can be reversed using guaranteed restore points.

Limits to both types of Restore Points include shrinking a datafile or dropping a tablespace can prevent flashing back the affected datafiles to the restore point.

About Logging for Flashback Database and GRP

The logging for Flashback Database and guaranteed restore points is based upon capturing images of datafile blocks before changes are applied. These images can then be used to return datafiles to their previous state when a FLASHBACK DATABASE command is executed. The chief difference between normal flashback logging and GRP logging are related to when blocks are logged and whether the logs can be deleted in response to space pressure in FRA. If no files are eligible for deletion because of retention policy and GRP then the database will behave as if it has encountered a disk full condition and may halt.

In general it is more efficient to turn off logging for Flashback Database and use only guaranteed restore points if the primary need is to be able to return your database to a specific time in which the guaranteed restore point was created. In other words you don't have a need to restore to a point between the GRP and the current SCN of the database. And you don't have a reason to use any of the other "Flashback" technologies.

If Flashback Database is enabled and one or more guaranteed restore points is defined then the database performs normal flashback logging. This can cause some performance overhead and significant space pressure in the flash

recovery area. It keeps all the information it needs to ally FLASHBACK DATABASE to any time as far back as the earliest currently defined guaranteed restore point.

Create Guaranteed Restore Point

```
CREATE RESTORE POINT before_damage GUARANTEE FLASHBACK DATABASE;
```

To See Restore Points

```
SELECT SCN, RESTORE_POINT_TIME, NAME, PERSERVED FROM GV$RESTORE_POINT;
```

To FLASHBACK DATABASE

```
SHUTDOWN IMMEDIATE;
```

```
STARTUP MOUNT EXCLUSIVE;
```

```
FLASHBACK DATABASE TO RESTORE POINT before_damage;
```

```
ALTER DATABASE OPEN RESETLOGS;
```

To Drop Restore Points

```
DROP RESTORE POINT before_damage;
```

How to quickly restore to a clean database using Oracle's restore point

Applies to:

Oracle database – 11gR2

Problem:

Often while conducting benchmarking tests, it is required to load a clean database before the start of a new run. One way to ensure a clean database is to recreate the entire database before each test run, but depending on the size of it, this approach may be very time consuming or inefficient.

Solution:

This article describes how to use Oracle's flashback feature to quickly restore a database to a state that existed just before running the workload. More specifically, this article describes steps on how to use the 'guaranteed restore points'.

Restore point:

Restore point is nothing but a name associated with a timestamp or an SCN of the database. One can create either a normal restore point or a guaranteed restore point. The difference between the two is that guaranteed restore point allows you to flashback to the restore point regardless of the DB_FLASHBACK_RETENTION_TARGET initialization parameter i.e. it is *always* available (assuming you have enough space in the flash recovery area).

NOTE: In this article Flashback logging was not turned ON.

Guaranteed Restore point:

Prerequisites: Creating a guaranteed restore point requires the following prerequisites:

The user must have the SYSDBA system privileges

Must have created a flash recovery area

The database must be in ARCHIVELOG mode

Create a guaranteed restore point:

After you have created or migrated a fresh database, first thing to do is to create a guaranteed restore point so you can flashback to it each time before you start a new workload. The steps are as under:

```
$> su - oracle
```

```
$> sqlplus / as sysdba;
```

Find out if ARCHIVELOG is enabled

```
SQL> select log_mode from v$database;
```

If step 3 shows that ARCHIVELOG is not enabled then continue else skip to step 8 below.

```
SQL> shutdown immediate;
```

```
SQL> startup mount;
```

```
SQL> alter database archivelog;
```

```
SQL> alter database open;
```

```
SQL> create restore point CLEAN_DB guarantee flashback database;
```

where CLEAN_DB is the name given to the guaranteed restore point.

Viewing the guaranteed restore point

```
SQL> select * from v$restore_point;
```

Verify the information about the newly created restore point. Also, note down the SCN# for reference and we will refer to it as "reference SCN#"

Flashback to the guaranteed restore point

Now, in order to restore your database to the guaranteed restore point, follow the steps below:

```
$> su - oracle
```

```
$> sqlplus / as sysdba;
```

```
SQL> select current_scn from v$database;
```

```
SQL> shutdown immediate;
```

```
SQL> startup mount;
```

```
SQL> select * from v$restore_point;
```

```
SQL> flashback database to restore point CLEAN_DB;
```

```
SQL> alter database open resetlogs;
```

```
SQL> select current_scn from v$database;
```

Compare the SCN# from step 9 above to the reference SCN#.

NOTE: The SCN# from step 9 above may not necessarily be the exact SCN# as the reference SCN# but it will be close enough.

Normal restore point

A label for an SCN or time. For commands that support an SCN or time, you can often specify a restore point. Normal restore points exist in the circular list and can be overwritten in the control file. However, if the restore point pertains to an archival backup, then it will be preserved in the recovery catalog.

Guaranteed restore point

A restore point for which the database is guaranteed to retain the flashback logs for an Oracle Flashback Database operation. Unlike a normal restore point, a guaranteed restore point does not age out of the control file and must be explicitly dropped. Guaranteed restore points utilize space in the flash recovery area, which must be defined.

66. What is the difference between 10g/11g OEM Grid control and 12c Cloud control?

https://blogs.oracle.com/oem/entry/questions_and_answers_from_the

67. What are the components of Grid control?

Grid Control Configuration: http://docs.oracle.com/html/B12013_03/configs.htm

OMS (Oracle Management Server):

OMR (Oracle Management Repository):

OMS (Oracle Management Server):

OEM Agent: oracle Management Agent (Management Agent):

Grid Control Console:

OMS is a J2EE Web application that orchestrates with Management Agents to discover targets, monitor and manage them, and store the collected information in a repository for future reference and analysis. OMS also renders the user interface for the Grid Control console. OMS is deployed to the application server that is installed along with other core components of Grid Control.

OMR (Oracle Management Repository):

Management Repository is the storage location where all the information collected by the Management Agent gets stored. It consists of objects such as database jobs, packages, procedures, views, and tablespaces.

Technically, OMS uploads the monitoring data it receives from the Management Agents to the Management Repository. The Management Repository then organizes the data so that it can be retrieved by OMS and displayed in the Grid Control console. Since data is stored in the Management Repository, it can be shared between any number of administrators accessing Grid Control.

Management Repository is configured in Oracle Database. This Oracle Database can either be an existing database in your environment or a new one installed along with other core components of Grid Control.

OEM Agent: oracle Management Agent (Management Agent):

Management Agent is an integral software component that is deployed on each monitored host. It is responsible for monitoring all the targets running on those hosts, communicating that information to the middle-tier Oracle Management Service, and managing and maintaining the hosts and its targets.

Grid Control Console:

Grid Control Console is the user interface you see after you install Grid Control. From the Grid Control console, you can monitor and administer your entire computing environment from one location on the network. All the services within your enterprise, including hosts, databases, listeners, application servers, and so on, are easily managed from one central location.

68. What are the new features of 12c Cloud control?

69. How to find if your Oracle database is 32 bit or 64 bit?

Execute the command "file \$ORACLE_HOME/bin/oracle", you should see output like /u01/db/bin/oracle: ELF 64-bit MSB executable SPARCV9 Version 1

means you are on 64 bit oracle.

If your oracle is 32 bit you should see output like below

oracle: ELF 32-bit MSB executable SPARC Version 1

70. How to find opatch Version?

Opatch is utility to apply database patch, In order to find opatch version execute

"\$ORACLE_HOME/OPatch/opatch version"

71. Which procedure does not affect the size of the SGA?

Stored procedure

72. When Dictionary tables are created?

Once for the Entire Database creation.

73. The order in which Oracle processes a single SQL statement is?

Parse, execute and fetch

74. What are the mandatory datafiles to create a database in Oracle 11g?

SYSTEM, SYSAUX, UNDO

75. In one server can we have different oracle versions?

Yes

76. How do sessions communicate with database?

Server processes execute SQL received from user processes.

77. Which SGA memory structure cannot be resized dynamically after instance startup?

Log buffer

78. When a session changes data, where does the change get written?

To the data block in the cache, and the redo log buffer

79. How many maximum no of control files we can have within a database?

8

80. System Data File Consists of?

Metadata

Bigfile Tablespaces

Oracle lets you create bigfile tablespaces. This allows Oracle Database to contain tablespaces made up of single large files rather than numerous smaller ones. This lets Oracle Database utilize the ability of 64-bit systems to create and manage ultralarge files. The consequence of this is that Oracle Database can now scale up to 8 exabytes in size.

With Oracle-managed files, bigfile tablespaces make datafiles completely transparent for users. In other words, you can perform operations on tablespaces, rather than the underlying datafile. Bigfile tablespaces make the tablespace the main unit of the disk space administration, backup and recovery, and so on. Bigfile tablespaces also simplify datafile management with Oracle-managed files and Automatic Storage Management by eliminating the need for adding new datafiles and dealing with multiple files.

The system default is to create a smallfile tablespace, which is the traditional type of Oracle tablespace. The SYSTEM and SYSAUX tablespace types are always created using the system default type.

Bigfile tablespaces are supported only for locally managed tablespaces with automatic segment-space management. There are two exceptions: locally managed undo and temporary tablespaces can be bigfile tablespaces, even though their segments are manually managed.

An Oracle database can contain both bigfile and smallfile tablespaces. Tablespaces of different types are indistinguishable in terms of execution of SQL statements that do not explicitly refer to datafiles.

You can create a group of temporary tablespaces that let a user consume temporary space from multiple tablespaces.

A tablespace group can also be specified as the default temporary tablespace for the database. This is useful with bigfile tablespaces, where you could need a lot of temporary tablespace for sorts.

Benefits of Bigfile Tablespaces

- Bigfile tablespaces can significantly increase the storage capacity of an Oracle database. Smallfile tablespaces can contain up to 1024 files, but bigfile tablespaces contain only one file that can be 1024 times larger than a smallfile tablespace. The total tablespace capacity is the same for smallfile tablespaces and bigfile tablespaces. However, because there is limit of 64K datafiles for each database, a database can contain 1024 times more bigfile tablespaces than smallfile tablespaces, so bigfile tablespaces increase the total database

capacity by 3 orders of magnitude. In other words, 8 exabytes is the maximum size of the Oracle database when bigfile tablespaces are used with the maximum block size (32 k).

- Bigfile tablespaces simplify management of datafiles in ultra large databases by reducing the number of datafiles needed. You can also adjust parameters to reduce the SGA space required for datafile information and the size of the control file.
- They simplify database management by providing datafile transparency.

Considerations with Bigfile Tablespaces

- Bigfile tablespaces are intended to be used with Automatic Storage Management or other logical volume managers that support dynamically extensible logical volumes and striping or RAID.
- Avoid creating bigfile tablespaces on a system that does not support striping because of negative implications for parallel execution and RMAN backup parallelization.
- Avoid using bigfile tablespaces if there could possibly be no free space available on a disk group, and the only way to extend a tablespace is to add a new datafile on a different disk group.
- Using bigfile tablespaces on platforms that do not support large file sizes is not recommended and can limit tablespace capacity. Refer to your operating system specific documentation for information about maximum supported file sizes.
- Performance of database opens, checkpoints, and DBWR processes should improve if data is stored in bigfile tablespaces instead of traditional tablespaces. However, increasing the datafile size might increase time to restore a corrupted file or create a new datafile.

The SYSTEM Tablespace

- Every Oracle database contains a tablespace named SYSTEM, which Oracle creates automatically when the database is created. The SYSTEM tablespace is always online when the database is open.
- To take advantage of the benefits of locally managed tablespaces, you can create a locally managed SYSTEM tablespace, or you can migrate an existing dictionary managed SYSTEM tablespace to a locally managed format.
- In a database with a locally managed SYSTEM tablespace, dictionary managed tablespaces cannot be created. It is possible to plug in a dictionary managed tablespace using the transportable feature, but it cannot be made writable.
- Note:
- If a tablespace is locally managed, then it cannot be reverted back to being dictionary managed.

The SYSAUX Tablespace

- The SYSAUX tablespace is an auxiliary tablespace to the SYSTEM tablespace. Many database components use the SYSAUX tablespace as their default location to store data. Therefore, the SYSAUX tablespace is always created during database creation or database upgrade.
- The SYSAUX tablespace provides a centralized location for database metadata that does not reside in the SYSTEM tablespace. It reduces the number of tablespaces created by default, both in the seed database and in user-defined databases.
- During normal database operation, the Oracle database server does not allow the SYSAUX tablespace to be dropped or renamed. Transportable tablespaces for SYSAUX is not supported.
- Note:
- If the SYSAUX tablespace is unavailable, such as due to a media failure, then some database features might fail.

Undo Tablespaces

- Undo tablespaces are special tablespaces used solely for storing undo information. You cannot create any other segment types (for example, tables or indexes) in undo tablespaces. Each database contains zero or more undo tablespaces. In automatic undo management mode, each Oracle instance is assigned one (and only one) undo tablespace. Undo data is managed within an undo tablespace using undo segments that are automatically created and maintained by Oracle.
- When the first DML operation is run within a transaction, the transaction is bound (assigned) to an undo segment (and therefore to a transaction table) in the current undo tablespace. In rare circumstances, if the instance does not have a designated undo tablespace, the transaction binds to the system undo segment.
- Caution:
- Do not run any user transactions before creating the first undo tablespace and taking it online.

- Each undo tablespace is composed of a set of undo files and is locally managed. Like other types of tablespaces, undo blocks are grouped in extents and the status of each extent is represented in the bitmap. At any point in time, an extent is either allocated to (and used by) a transaction table, or it is free.
- You can create a bigfile undo tablespace.

Creation of Undo Tablespaces

A database administrator creates undo tablespaces individually, using the CREATE UNDO TABLESPACE statement. It can also be created when the database is created, using the CREATE DATABASE statement. A set of files is assigned to each newly created undo tablespace. Like regular tablespaces, attributes of undo tablespaces can be modified with the ALTER TABLESPACE statement and dropped with the DROP TABLESPACE statement.

Note: An undo tablespace cannot be dropped if it is being used by any instance or contains any undo information needed to recover transactions.

Assignment of Undo Tablespaces

You assign an undo tablespace to an instance in one of two ways:

- At instance startup. You can specify the undo tablespace in the initialization file or let the system choose an available undo tablespace.
- While the instance is running. Use ALTER SYSTEM SET UNDO_TABLESPACE to replace the active undo tablespace with another undo tablespace. This method is rarely used.

You can add more space to an undo tablespace by adding more datafiles to the undo tablespace with the ALTER TABLESPACE statement.

You can have more than one undo tablespace and switch between them. Use the Database Resource Manager to establish user quotas for undo tablespaces. You can specify the retention period for undo information.

Default Temporary Tablespace

When the SYSTEM tablespace is locally managed, you must define at least one default temporary tablespace when creating a database. A locally managed SYSTEM tablespace cannot be used for default temporary storage.

If SYSTEM is dictionary managed and if you do not define a default temporary tablespace when creating the database, then SYSTEM is still used for default temporary storage. However, you will receive a warning in ALERT.LOG saying that a default temporary tablespace is recommended and will be necessary in future releases.

How to Specify a Default Temporary Tablespace

Specify default temporary tablespaces when you create a database, using the DEFAULT TEMPORARY TABLESPACE extension to the CREATE DATABASE statement.

If you drop all default temporary tablespaces, then the SYSTEM tablespace is used as the default temporary tablespace.

You can create bigfile temporary tablespaces. A bigfile temporary tablespaces uses tempfiles instead of datafiles.

Note: You cannot make a default temporary tablespace permanent or take it offline.

Using Multiple Tablespaces

A very small database may need only the SYSTEM tablespace; however, Oracle recommends that you create at least one additional tablespace to store user data separate from data dictionary information. This gives you more flexibility in various database administration operations and reduces contention among dictionary objects and schema objects for the same datafiles.

You can use multiple tablespaces to perform the following tasks:

- Control disk space allocation for database data
- Assign specific space quotas for database users
- Control availability of data by taking individual tablespaces online or offline
- Perform partial database backup or recovery operations
- Allocate data storage across devices to improve performance

A database administrator can use tablespaces to do the following actions:

- Create new tablespaces
- Add datafiles to tablespaces
- Set and alter default segment storage settings for segments created in a tablespace
- Make a tablespace read only or read/write
- Make a tablespace temporary or permanent
- Rename tablespaces
- Drop tablespaces

Managing Space in Tablespaces

Tablespaces allocate space in extents. Tablespaces can use two different methods to keep track of their free and used space:

- Locally managed tablespaces: Extent management by the tablespace
- Dictionary managed tablespaces: Extent management by the data dictionary

When you create a tablespace, you choose one of these methods of space management. Later, you can change the management method with the DBMS_SPACE_ADMIN PL/SQL package.

Note: If you do not specify extent management when you create a tablespace, then the default is locally managed.

Locally Managed Tablespaces

A tablespace that manages its own extents maintains a bitmap in each datafile to keep track of the free or used status of blocks in that datafile. Each bit in the bitmap corresponds to a block or a group of blocks. When an extent is allocated or freed for reuse, Oracle changes the bitmap values to show the new status of the blocks. These changes do not generate rollback information because they do not update tables in the data dictionary (except for special cases such as tablespace quota information).

Locally managed tablespaces have the following advantages over dictionary managed tablespaces:

- Local management of extents automatically tracks adjacent free space, eliminating the need to coalesce free extents.
- Local management of extents avoids recursive space management operations. Such recursive operations can occur in dictionary managed tablespaces if consuming or releasing space in an extent results in another operation that consumes or releases space in a data dictionary table or rollback segment.

The sizes of extents that are managed locally can be determined automatically by the system. Alternatively, all extents can have the same size in a locally managed tablespace and override object storage options.

The LOCAL clause of the CREATE TABLESPACE or CREATE TEMPORARY TABLESPACE statement is specified to create locally managed permanent or temporary tablespaces, respectively.

Segment Space Management in Locally Managed Tablespaces

When you create a locally managed tablespace using the CREATE TABLESPACE statement, the SEGMENT SPACE MANAGEMENT clause lets you specify how free and used space within a segment is to be managed. Your choices are:

AUTO

This keyword tells Oracle that you want to use bitmaps to manage the free space within segments. A bitmap, in this case, is a map that describes the status of each data block within a segment with respect to the amount of space in the block available for inserting rows. As more or less space becomes available in a data block, its new state is reflected in the bitmap. Bitmaps enable Oracle to manage free space more automatically; thus, this form of space management is called automatic segment-space management.

Locally managed tablespaces using automatic segment-space management can be created as smallfile (traditional) or bigfile tablespaces. AUTO is the default.

MANUAL

This keyword tells Oracle that you want to use free lists for managing free space within segments. Free lists are lists of data blocks that have space available for inserting rows.

Dictionary Managed Tablespaces

If you created your database with an earlier version of Oracle, then you could be using dictionary managed tablespaces. For a tablespace that uses the data dictionary to manage its extents, Oracle updates the appropriate tables in the data dictionary whenever an extent is allocated or freed for reuse. Oracle also stores rollback information about each update of the dictionary tables. Because dictionary tables and rollback segments are part of the database, the space that they occupy is subject to the same space management operations as all other data.

Multiple Block Sizes

Oracle supports multiple block sizes in a database. The standard block size is used for the SYSTEM tablespace. This is set when the database is created and can be any valid size. You specify the standard block size by setting the initialization parameter DB_BLOCK_SIZE. Legitimate values are from 2K to 32K.

In the initialization parameter file or server parameter, you can configure subcaches within the buffer cache for each of these block sizes. Subcaches can also be configured while an instance is running. You can create tablespaces having any of these block sizes. The standard block size is used for the system tablespace and most other tablespaces.

Note: All partitions of a partitioned object must reside in tablespaces of a single block size.

Multiple block sizes are useful primarily when transporting a tablespace from an OLTP database to an enterprise data warehouse. This facilitates transport between databases of different block sizes.

81. What is the function of SMON in instance recovery?

It roles forward by applying changes in the redo log.

Shutdown Modes

A database administrator with SYSDBA or SYSOPER privileges can shut down the database using the SQL*Plus SHUTDOWN command or Enterprise Manager. The SHUTDOWN command has options that determine shutdown behavior. [Table 13-2](#) summarizes the behavior of the different shutdown modes.

Table 13-2 Shutdown Modes

Database Behavior	ABORT	IMMEDIATE	TRANSACTIONAL	NORMAL
Permits new user connections	No	No	No	No
Waits until current sessions end	No	No	No	Yes
Waits until current transactions end	No	No	Yes	Yes
Performs a checkpoint and closes open files	No	Yes	Yes	Yes

The possible SHUTDOWN statements are:

SHUTDOWN ABORT

This mode is intended for emergency situations, such as when no other form of shutdown is successful. This mode of shutdown is the fastest. However, a subsequent open of this database may take substantially longer because instance recovery must be performed to make the data files consistent.

Note: Because SHUTDOWN ABORT does not checkpoint the open data files, instance recovery is necessary before the database can reopen. The other shutdown modes do not require instance recovery before the database can reopen.

SHUTDOWN IMMEDIATE

This mode is typically the fastest next to SHUTDOWN ABORT. Oracle Database terminates any executing SQL statements and disconnects users. Active transactions are terminated and uncommitted changes are rolled back.

SHUTDOWN TRANSACTIONAL

This mode prevents users from starting new transactions, but waits for all current transactions to complete before shutting down. This mode can take a significant amount of time depending on the nature of the current transactions.

SHUTDOWN NORMAL

This is the default mode of shutdown. The database waits for all connected users to disconnect before shutting down.

How a Database Is Closed

The database close operation is implicit in a database shutdown. The nature of the operation depends on whether the database shutdown is normal or abnormal.

How a Database Is Closed During Normal Shutdown

When a database is closed as part of a SHUTDOWN with any option other than ABORT, Oracle Database writes data in the SGA to the data files and online redo log files. Next, the database closes online data files and online redo log files. Any offline data files of offline tablespaces have been closed already. When the database reopens, any tablespace that was offline remains offline.

At this stage, the database is closed and inaccessible for normal operations. The control files remain open after a database is closed.

How a Database Is Closed During Abnormal Shutdown

If a SHUTDOWN ABORT or abnormal termination occurs, then the instance of an open database closes and shuts down the database instantaneously. Oracle Database does not write data in the buffers of the SGA to the data files and redo log files. The subsequent reopening of the database requires instance recovery, which Oracle Database performs automatically.

How a Database Is Unmounted

After the database is closed, Oracle Database unmounts the database to disassociate it from the instance. After a database is unmounted, Oracle Database closes the control files of the database. At this point, the instance remains in memory.

How an Instance Is Shut Down

The final step in database shutdown is shutting down the instance. When the database instance is shut down, the SGA is removed from memory and the background processes are terminated.

In unusual circumstances, shutdown of an instance may not occur cleanly. Memory structures may not be removed from memory or one of the background processes may not be terminated. When remnants of a previous instance exist, a subsequent instance startup may fail. In such situations, you can force the new instance to start by removing

the remnants of the previous instance and then starting a new instance, or by issuing a SHUTDOWN ABORT statement in SQL*Plus or using Enterprise Manager.

Database writer (DBWn)

The database writer writes modified blocks from the database buffer cache to the datafiles. Oracle Database allows a maximum of 20 database writer processes (DBW0-DBW9 and DBWa-DBWj). The DB_WRITER_PROCESSES initialization parameter specifies the number of DBWn processes. The database selects an appropriate default setting for this initialization parameter or adjusts a user-specified setting based on the number of CPUs and the number of processor groups.

For more information about setting the DB_WRITER_PROCESSES initialization parameter, see the Oracle Database Performance Tuning Guide.

Log writer (LGWR)

The log writer process writes redo log entries to disk. Redo log entries are generated in the redo log buffer of the system global area (SGA). LGWR writes the redo log entries sequentially into a redo log file. If the database has a multiplexed redo log, then LGWR writes the redo log entries to a group of redo log files. See Chapter 10, "Managing the Redo Log" for information about the log writer process.

Checkpoint (CKPT)

At specific times, all modified database buffers in the system global area are written to the datafiles by DBWn. This event is called a checkpoint. The checkpoint process is responsible for signalling DBWn at checkpoints and updating all the datafiles and control files of the database to indicate the most recent checkpoint.

System monitors (SMON)

The system monitor performs recovery when a failed instance starts up again. In an Oracle Real Application Clusters database, the SMON process of one instance can perform instance recovery for other instances that have failed. SMON also cleans up temporary segments that are no longer in use and recovers dead transactions skipped during system failure and instance recovery because of file-read or offline errors. These transactions are eventually recovered by SMON when the tablespace or file is brought back online.

Process monitor (PMON)

The process monitor performs process recovery when a user process fails. PMON is responsible for cleaning up the cache and freeing resources that the process was using. PMON also checks on the dispatcher processes (described later in this table) and server processes and restarts them if they have failed.

Archiver (ARCn)

One or more archiver processes copy the redo log files to archival storage when they are full or a log switch occurs. Archiver processes are the subject of Chapter 11, "Managing Archived Redo Logs".

Recoverer (RECO)

The recoverer process is used to resolve distributed transactions that are pending because of a network or system failure in a distributed database. At timed intervals, the local RECO attempts to connect to remote databases and automatically complete the commit or rollback of the local portion of any pending distributed transactions. For information about this process and how to start it, see Chapter 33, "Managing Distributed Transactions".

Dispatcher (Dnnn)

Dispatchers are optional background processes, present only when the shared server configuration is used. Shared server was discussed previously in "Configuring Oracle Database for Shared Server".

Global Cache Service (LMS)

In an Oracle Real Application Clusters environment, this process manages resources and provides inter-instance resource control.

82. Which action occurs during a checkpoint?

Oracle flushes the dirty blocks in the database buffer cache to disk.

Explanation-1:

A **checkpoint** occurs when the [DBWR](#) (database writer) process writes all modified buffers in the [SGA](#) buffer cache to the database [data files](#). Data file headers are also updated with the latest checkpoint [SCN](#), even if the file had no changed blocks.

Checkpoints occur AFTER (not during) every [redo log](#) switch and also at intervals specified by initialization parameters.

Set parameter `LOG_CHECKPOINTS_TO_ALERT=TRUE` to observe checkpoint start and end times in the database alert log.

Checkpoints can be forced with the `ALTER SYSTEM CHECKPOINT;` command.

Explanation-2:

Checkpoint types can be divided as INCREMENTAL and COMPLETE.

Also COMPLETE CHECKPOINT can be divided further into PARTIAL and FULL.

In **Incremental Checkpoint**, checkpoint information is written to the controlfile. In the following cases:

1. Every three seconds.
2. At the time of log switch - Sometimes log switches may trigger a complete checkpoint, if the next log where the log switch is to take place is Active.

In **complete Checkpoint**, checkpoint information is written in controlfile, datafile header and also dirty block is written by DBWR to the datafiles.

Full Checkpoint

1. fast_start_mttr_target
2. Before Clean Shutdown
3. Some log switches may trigger a complete checkpoint, if the next log where the log switch is to take place is Active. This has more chance of happening when the Redo Log files are small in size and continuous transactions are taking place.
4. when the 'alter system checkpoint' command is issued

Partial Checkpoint happens in the following cases.

1. before begin backup.
2. before tablespace offline.
3. before placing tablespace in read only.
4. Before dropping tablespace.
5. before taking datafile offline.
6. When checkpoint queue exceeds its threshold.
7. before segment is dropped.
8. Before adding and removing columns from table.

Explanation-3:

a checkpoint is the act of flushing modified, cached database blocks to disk. Normally, when you make a change to a block -- the modifications of that block are made to a memory copy of the block. When you commit -- the block is not written (but the REDO LOG is -- that makes it so we can "replay" your transaction in the event of a failure).

Eventually, the system will checkpoint your modified blocks to disk.

there is no relationship between "checkpoint" and sid and instance recovery does not imply "checkpoint". a checkpoint reduces the amount of time it takes to perform instance recovery.

Explanation-3:

PURPOSE OF CHECKPOINTS

Database blocks are temporarily stored in Database buffer cache. As blocks are read, they are stored in DB buffer cache so that if any user accesses them later, they are available in memory and need not be read from the disk. When we update any row, the buffer in DB buffer cache corresponding to the block containing that row is updated in memory. Record of the change made is kept in redo log buffer. On commit, the changes we made are written to the disk thereby making them permanent. But where are those changes written? To the datafiles containing data blocks? No !!! The changes are recorded in online redo log files by flushing the contents of redo log buffer to them. This is called write ahead logging. If the instance crashed right now, the DB buffer cache will be wiped out but on restarting the database, Oracle will apply the changes recorded in redo log files to the datafiles.

Why doesn't Oracle write the changes to datafiles right away when we commit the transaction? The reason is simple. If it chose to write directly to the datafiles, it will have to physically locate the data block in the datafile first and then update it which means that after committing, user has to wait until DBWR searches for the block and then writes it before he can issue next command. This will bring down the performance drastically. That is where the role of redo logs comes in. The writes to the redo logs are sequential writes -- LGWR just dumps the info in redologs to log files sequentially and synchronously so that the user does not have to wait for long. Moreover, DBWR will always write in units of Oracle blocks whereas LGWR will write only the changes made. Hence, write ahead logging also improves performance by reducing the amount of data written synchronously. When will the changes be applied to

the datablocks in datafiles? The data blocks in the datafiles will be updated by the DBWR asynchronously in response to certain triggers. These triggers are called checkpoints.

Checkpoint is a synchronization event at a specific point in time which causes some / all dirty blocks to be written to disk thereby guaranteeing that blocks dirtied prior to that point in time get written.

Whenever dirty blocks are written to datafiles, it allows oracle

- **to reuse a redo log** : A redo log can't be reused until DBWR writes all the dirty blocks protected by that logfile to disk. If we attempt to reuse it before DBWR has finished its checkpoint, we get the following message in alert log :
Checkpoint not complete.

- **to reduce instance recovery time** : As the memory available to a database instance increases, it is possible to have database buffer caches as large as several million buffers. It requires that the database checkpoint advance frequently to limit recovery time, since infrequent checkpoints and large buffer caches can exacerbate crash recovery times significantly.

- **to free buffers for reads** : Dirtied blocks can't be used to read new data into them until they are written to disk. Thus DBWR writes dirty blocks from the buffer cache, to make room in the cache.

Various types of checkpoints in Oracle :

- Full checkpoint
- Thread checkpoint
- File checkpoint
- Parallel Query checkpoint
- Object checkpoint
- Log switch checkpoint
- Incremental checkpoint

Whenever a checkpoint is triggered:

- DBWR writes some /all dirty blocks to datafiles
- CKPT process updates the control file and datafile headers

FULL CHECKPOINT

- Writes block images to the database for all dirty buffers from all instances.
- Statistics updated
 - . DBWR checkpoints
 - . DBWR checkpoint buffers written
 - . DBWR thread checkpoint buffers written
- Caused by :
 - . Alter system checkpoint [global]
 - . Alter database begin backup
 - . Alter database close
 - . Shutdown [immediate]
- Controlfile and datafile headers are updated
 - . Checkpoint_change#

THREAD CHECKPOINT

- Writes block images to the database for all dirty buffers from one instance
- Statistics updated
 - . DBWR checkpoints
 - . DBWR checkpoint buffers written
 - . DBWR thread checkpoint buffers written
- Caused by :
 - . Alter system checkpoint local
- Controlfile and datafile headers are updated
 - . Checkpoint_change#

FILE CHECKPOINT

When a tablespace is put into backup mode or take it offline, Oracle writes all the dirty blocks from the tablespace to disk before changing the state of the tablespace.

- Writes block images to the database for all dirty buffers for all files of a tablespace from all instances
- Statistics updated
 - . DBWR checkpoints
 - . DBWR tablespace checkpoint buffers written

- . DBWR checkpoint buffers written
- Caused by :
 - . Alter tablespace xxx offline
 - . Alter tablespace xxx begin backup
 - . Alter tablespace xxx read only
- Controlfile and datafile headers are updated
- . Checkpoint_change#

PARALLEL QUERY CHECKPOINT

Parallel query often results in direct path reads (Full tablescan or index fast full scan). This means that blocks are read straight into the session's PGA, bypassing the data cache; but that means if there are dirty buffers in the data cache, the session won't see the most recent versions of the blocks unless they are copied to disk before the query starts – so parallel queries start with a checkpoint.

- Writes block images to the database for all dirty buffers belonging to objects accessed by the query from all instances.
- Statistics updated
- . DBWR checkpoints
- . DBWR checkpoint buffers written
- Caused by :
 - . Parallel Query
 - . Parallel Query component of Parallel DML (PDML) or Parallel DDL (PDDL)
- Mandatory for consistency
- Controlfile and datafile headers are updated
- . Checkpoint_change#

OBJECT CHECKPOINT

When an object is dropped/truncated, the session initiates an object checkpoint telling DBWR to copy any dirty buffers for that object to disk and the state of those buffers is changed to free.

- Writes block images to the database for all dirty buffers belonging to an object from all instances.
- Statistics updated
- . DBWR checkpoints
- . DBWR object drop buffers written
- Caused by dropping or truncating a segment:
 - . Drop table XXX
 - . Drop table XXX Purge
 - . Truncate table xxx
 - . Drop index xxx
- Mandatory for media recovery purposes
- Controlfile and datafile headers are updated
- . Checkpoint_change#

LOG SWITCH CHECKPOINT

- Writes the contents of the dirty buffers whose information is protected by a redo log to the database .
- Statistics updated
- . DBWR checkpoints
- . DBWR checkpoint buffers written
- . background checkpoints started
- . background checkpoints completed
- Caused by log switch
- Controlfile and datafile headers are updated
- . Checkpoint_change#

INCREMENTAL CHECKPOINT

Prior to Oracle 8i, only well known checkpoint was log switch checkpoint. Whenever LGWR filled an online logfile, DBWR would go into a frenzy writing data blocks to disks, and when it had finished, Oracle would update each data file header block with the SCN to show that file was updated up to that point in time.

Oracle 8i introduced incremental checkpointing which triggered DBWR to write some dirty blocks from time to time so as to advance the checkpoint and reduce the instance recovery time.

Incremental checkpointing has been implemented using two algorithms :

- Ageing algorithm
- LRU/TCH algorithm

AGEING ALGORITHM

This strategy involves writing changed blocks that have been dirty for the longest time and is called aging writes. This algorithm relies on the CKPT Q running thru the cache and buffers being linked to the end of this list the first time they are made dirty.

.The LRU list contains all the buffers – free / pinned / dirty. Whenever a buffer in LRU list is dirtied, it is placed in CKPT Q as well i.e. a buffer can simultaneously have pointers in both LRU list and CKPT Q but the buffers in CKPT Q are arranged in the order in which they were dirtied. Thus, checkpoint queue contains dirty blocks in the order of SCN# in which they were dirtied

Every 3 secs DBWR wakes up and checks if there are those many dirty buffers in CKPT Q which need to be written so as to satisfy instance recovery requirement..

If those many or more dirty buffers are not found,

DBWR goes to sleep

else (dirty buffers found)

.CKPT target RBA is calculated based on

- The most recent RBA
- log_checkpoint_interval
- log_checkpoint_timeout
- fast_start_mttr_target
- fast_start_io_target
- 90% of the size of the smallest redo log file

. DBWR walks the CKPT Q from the low end (dirtied earliest) of the redo log file collecting buffers for writing to disk until it reaches the buffer that is more recent than the target RBA. These buffers are placed in write list-main.

. DBWR walks the write list-main and checks all the buffers

– If changes made to the buffer have already been written to redo log files

. Move those buffers to write-aux list

else

. Trigger LGWR to write changes to those buffers to redo logs

. Move those buffers to write-aux list

. Write buffers from write-aux list to disk

. Update checkpoint RBA in SGA

. Delink those buffers from CKPT Q

. Delink those buffers from write-aux list

- Statistics Updated :

. DBWR checkpoint buffers written

- Controlfile updated every 3 secs by CKPT

. Checkpoint progress record

As sessions link buffers to one end of the list, DBWR can effectively unlink buffers from the other end and copy them to disk. To reduce contention between DBWR and foreground sessions, there are two linked lists in each working set so that foreground sessions can link buffers to one while DBWR is unlinking them from the other.

LRU/TCH ALGORITHM

LRU/TCH algorithm writes the cold dirty blocks to disk that are on the point of being pushed out of cache.

As per ageing algorithm, DBWR will wake up every 3 seconds to flush dirty blocks to disk. But if blocks get dirtied at a fast pace during those 3 seconds and a server process needs some free buffers, some buffers need to be flushed to the disk to make room. That's when LRU/TCH algorithm is used to write those dirty buffers which are on the cold end of the LRU list.

Whenever a server process needs some free buffers to read data, it scans the LRU list from its cold end to look for free buffers.

While searching

If unused buffers found

Read blocks from disk into the buffers and link them to the corresponding hash bucket

if it finds some clean buffers (contain data but not dirtied or dirtied and have been flushed to disk),

if they are the candidates to be aged out (low touch count)

Read blocks from disk into the buffers and link them to the corresponding hash bucket

else (have been accessed recently and should not be aged out)

Move them to MRU end depending upon its touch count.

If it finds dirty buffers (they are already in CKPT Q),

Delink them from LRU list

Link them to the write-main list (Now these buffers are in CKPT Q and write-main list)

The server process scans a threshold no. of buffers (`_db_block_max_scan_pct = 40`(default)). If it does not find required no. of free buffers,

It triggers DBWR to dirty blocks in write-mainlist to disk

. DBWR walks the write list-main and checks all the buffers

– If changes made to the buffer have already been written to redo log files

. Move those buffers to write-aux list

else

. Trigger LGWR to write changes to those buffers to redo logs

. Move those buffers to write-aux list

. Write buffers from write-aux list to disk

. Delink those buffers from CKPT Q and write-aux list

. Link those buffers to LRU list as free buffers

Note that

- In this algorithm, the dirty blocks are delinked from LRU list before linking them to write-main list in contrast to ageing algorithm where the blocks can be simultaneously be in both CKPT Q and LRU list.

- In this algorithm, checkpoint is not advanced because it may be possible that the dirty blocks on the LRU end may actually not be the ones which were dirtied earliest. They may be there because the server process did not move them to the MRU end earlier. There might be blocks present in CKPT Q which were dirtied earlier than the blocks in question.

I hope the information was useful. Thanks for your time.

Explanation-3:

A Checkpoint is a database event which synchronizes the modified data blocks in memory with the datafiles on disk. It offers Oracle the means for ensuring the consistency of data modified by transactions. The mechanism of writing modified blocks on disk in Oracle is not synchronized with the commit of the corresponding transactions.

A checkpoint has two purposes:

(1) to establish data consistency, and

(2) enable faster database recovery.

The checkpoint must ensure that all the modified buffers in the cache are really written to the corresponding datafiles to avoid the loss of data which may occur with a crash (instance or disk failure).

Depending on the number of datafiles in a database, a checkpoint can be a highly resource intensive operation, since all datafile headers are frozen during the checkpoint. Frequent checkpoints will enable faster recovery, but can cause performance degradation

Key Initialization parameters related to Checkpoint performance.

- `FAST_START_MTTR_TARGET`
- `LOG_CHECKPOINT_INTERVAL`
- `LOG_CHECKPOINT_TIMEOUT`
- `LOG_CHECKPOINTS_TO_ALERT`

`FAST_START_MTTR_TARGET`: It enables you to specify the number of seconds the database takes to perform crash recovery

of a single instance. Based on internal statistics, incremental checkpoint automatically adjusts the checkpoint target to meet the requirement of `FAST_START_MTTR_TARGET`. `V$INSTANCE_RECOVERY.ESTIMATED_MTTR` shows the current estimated mean time to recover (MTTR) in seconds. This value is shown even if `FAST_START_MTTR_TARGET` is not specified.

`LOG_CHECKPOINT_INTERVAL`: It influences when a checkpoint occurs, which means careful attention should be given to the setting of this parameter, keeping it updated as the size of the redo log files is changed. The checkpoint frequency is one of the factors which impacts the time required for the database to recover from an unexpected failure. Longer intervals between checkpoints mean that if the system crashes, more time will be needed for the

database to recover. Shorter checkpoint intervals mean that the database will recover more quickly, at the expense of increased resource utilization during the checkpoint operation

LOG_CHECKPOINT_TIMEOUT: The parameter specifies the maximum number of seconds the incremental checkpoint target should lag the current log tail. In another word, it specifies how long a dirty buffer in buffer cache can remain dirty. Checkpoint frequency impacts the time required for the database to recover from an unexpected failure. Longer intervals between checkpoints mean that more time will be required during database recovery.

LOG_CHECKPOINTS_TO_ALERT: It lets you log your checkpoints to the alert file. Doing so is useful for determining whether checkpoints are occurring at the desired frequency.

Relationship between Redologs and Checkpoint: A checkpoint occurs at every log switch. If a previous checkpoint is already in progress, the checkpoint forced by the log switch will override the current checkpoint. Maintain well-sized redo logs to avoid unnecessary checkpoints as a result of frequent log switches. The alert log is a valuable tool for monitoring the rate that log switches occur, and subsequently, checkpoints occur.

Checkpoint not complete: This message in alert log indicates that Oracle wants to reuse a redo log file, but the current checkpoint position is still in that log. In this case, Oracle must wait until the checkpoint position passes that log. When the database waits on checkpoints, redo generation is stopped until the log switch is done. This situation may be encountered if DBWR writes

too slowly, or if a log switch happens before the log is completely full, or if log file sizes are too small.

Explanation-4:

<http://www.oracleportal.org/knowledge-base/oracle-database/database-concepts/general/checkpoint.aspx>

83. SMON process is used to write into LOG files?

No

84. Oracle does not consider a transaction committed until?

The LGWR successfully writes the changes to redo

85. How many maximum DBWn (Db writers) we can invoke?

20

86. Which activity would generate less undo data?

INSERT

87. What happens when a user issues a COMMIT?

- 1) LGWR wakes up.
- 2) LGWR acquires the redo allocation latch and redo copy latch.
- 3) LGWR flushes the redologbuffer to the logfiles (2 memebtrs in parallel)
- 4) LGWR releases the redo latches
- 5) LGWR posts session A

The LGWR flushes the log buffer to the online redo log.

"When you issue a DML Oracle generates redo entries based on the changes and these entries are buffered in memory while the transaction is occurring.

When you issue a commit, oracle immediately writes this redo entries to disk along with redo for the commit. Oracle does not return from the commit until the redo has been completely written to disk.

The matter is, the redo information is written to the disk immediately and the session waits for the process to complete before return.

Asynchronous Commit:

But in Oracle 10g R2 oracle changed this concept (i.e.)

You can let the log writer write the redo information to disk in its own time, instead of immediately, and you can have the commit return to you before it's completed, instead of waiting."

Explanation-2:

Committing means that a user has explicitly or implicitly requested that the changes in the transaction be made permanent. An explicit request occurs when the user issues a COMMIT statement. An implicit request occurs after normal termination of an application or completion of a data definition language (DDL) operation. The changes made by the SQL statement(s) of a transaction become permanent and visible to other users only after that transaction commits. Queries that are issued after the transaction commits will see the committed changes.

You can name a transaction using the SET TRANSACTION ... NAME statement before you start the transaction. This makes it easier to monitor long-running transactions and to resolve in-doubt distributed transactions.

88. What happens when a user process fails?

PMON performs process recovery.

Explanation: The process monitor (PMON) performs process recovery when a user process fails. PMON is responsible for cleaning up the database buffer cache and freeing resources that the user process was using. For example, it resets the status of the active transaction table, releases locks, and removes the process ID from the list of active processes.

PMON periodically checks the status of dispatcher and server processes, and restarts any that have stopped running (but not any that Oracle Database has terminated intentionally). PMON also registers information about the instance and dispatcher processes with the network listener.

Like SMON, PMON checks regularly to see whether it is needed and can be called if another process detects the need for it.

Explanation-2: Types of Database failures:

Statement	A single database operation fails, such as a DML (Data Manipulation Language) statement - INSERT, UPDATE, and so on.
User process	A single database connection fails.
Network	A network component between the client and the database server fails, and the session is disconnected from the database.
User error	An error message is not generated, but the operation's result, such as dropping a table, is not what the user intended.
Instance	The database instance fails unexpectedly.
Media	One or more of the database files is lost, deleted, or corrupted.

Database Recovery

When a database fails to run or a media failure occurs, or any of database or schema objects are lost or corrupted, a recovery process is needed. For this, an understanding of various types of database failures is essential.

Database Failure Types

There are six general categories for database-related failures. Understanding what category a failure belongs in will help you to more quickly understand the nature of the recovery effort you need to use to reverse the effects of the failure and maintain a high level of availability and performance in your database. The six general categories of failures are as follows:

Statement Failures

Statement failures occur when a single database operation fails, such as a single INSERT statement or the creation of a table. In the list that follows are a few of the most common problems and their solutions when a statement fails. Although granting user privileges or additional quotas within a tablespace solves many of these problems, also consider whether there are any gaps in the user education process that might lead to some of these problems in the first place.

User Process Failures

The abnormal termination of a user session is categorized as a user process failure; any uncommitted transaction must be cleaned up. The PMON (process monitor) background process periodically checks all user processes to ensure that the session is still connected. If the PMON finds a disconnected session, it rolls back the uncommitted transaction and releases all locks held by the disconnected process. Causes for user process failures typically fall into one of these categories:

A user closes their SQL*Plus window without logging out.

The workstation reboots suddenly before the application can be closed.

The application program causes an exception and closes before the application can be terminated normally.

A user process times out and Oracle disconnects the session.

A small percentage of user process failures is generally no cause for concern unless it becomes chronic; it may be a sign that user education is lacking—for example, training users to terminate the application gracefully before shutting down their workstation.

Network Failures

Depending on the locations of your workstation and your server, getting from your workstation to the server over the network might involve a number of hops: you might traverse several local switches and WAN routers to get to the database. From a network perspective, this configuration provides a number of points where failure can occur. These types of failures are called network failures.

In addition to hardware failures between the server and client, a listener process on the Oracle server can fail or the network card on the server itself can fail. To guard against these kinds of failures, you can provide redundant network paths from your clients to the server, as well as additional listener connections on the Oracle server and redundant network cards on the server.

User Error Failures

Even if all your redundant hardware is at peak performance, and your users have been trained to disconnect from their Oracle sessions properly, users can still inadvertently delete or modify data in tables or drop an index. This is known as a user error failure. Although these operations succeed from a statement point of view, they might not be logically correct: the DROP TABLE command worked fine, but you really didn't want to drop that table!

If data was inadvertently deleted from a table, and not yet committed, a ROLLBACK statement will undo the damage. If a COMMIT has already been performed, you have a number of options at your disposal, such as using data in the undo tablespace for a Flashback Query or using data in the archived and online redo logs with the LogMiner utility, available as a command-line or GUI interface.

You can recover a dropped table using Oracle's recycle bin functionality: a dropped table is stored in a special structure in the tablespace and is available for retrieval as long as the space occupied by the table in the tablespace is not needed for new objects. Even if the table is no longer in the tablespace's recycle bin, depending on the criticality of the dropped table, you can use either tablespace point in time recovery (TSPITR) or Flashback Database Recovery to recover the table, taking into consideration the potential data loss for other objects stored in the same tablespace for TSPITR or in the database if you use Flashback Database Recovery.

If the inadvertent changes are limited to a small number of tables that have few or no interdependencies with other database objects, Flashback Table functionality is most likely the right tool to bring back the table to a point of time in the past.

Instance Failures

An instance failure occurs when the instance shuts down without synchronizing all the database files to the same system change number (SCN), requiring a recovery operation the next time the instance is started. Many of the reasons for an instance failure are out of your direct control; in these situations, you can minimize the impact of these failures by tuning instance recovery.

A few causes for instance failure:

- A power outage.
- A server hardware failure.
- Failure of an Oracle background process.
- Emergency shutdown procedures (intentional power outage or SHUTDOWN ABORT) .

In all these scenarios, the solution is easy: run the STARTUP command, and let Oracle automatically perform instance recovery using the online redo logs and undo data in the undo tablespace. If the cause of the instance failure is related to an Oracle background process failure, you can use the alert log and process-specific trace files to debug the problem. The EM Database Control makes it easy to review the contents of the alert log and any other alerts generated right before the point of failure.

Media Failures

Another type of failure that is somewhat out of your control is media failure. A media failure is any type of failure that results in the loss of one or more database files: datafiles, control files, or redo log files. Although the loss of other database-related files such as an init.ora file or a server parameter file (SPFILE) is of great concern, Oracle Corporation does not consider it a media failure.

The database file can be lost or corrupted for a number of reasons:

- Failure of a disk drive.
- Failure of a disk controller.
- Inadvertent deletion or corruption of a database file.

Following the best practices by adequately mirroring control files, redo log files, and ensuring that full backups and their subsequent archived redo log files are available will keep you prepared for any type of media failure

89. What are the free buffers in the database buffer cache?

Buffer that can be overwritten

<http://koenigocm.blogspot.in/2012/07/database-buffer-cache-architecture.html>

Explanation-1: Database Buffer cache is one of the most important components of System Global Area (SGA).

Database Buffer Cache is the place where data blocks are copied from datafiles to perform SQL operations. Buffer Cache is shared memory structure and it is concurrently accessed by all server processes.

Working of Database buffer Cache

Buffer Cache is organized into two lists

Write List

Write list contains dirty buffers. These are the data blocks which contain modified data and needed to be written to datafiles.

Least Recent Used (LRU) List

Buffers owned by LRU list are categorized into **Pinned**, **Clean**, **Free or Unused** and **Dirty buffers**. **Pinned** buffers are currently being used while **Clean** buffers are available for use. Although **Clean** buffers contain some data but it is sync with block content stored in datafiles, so there is no need to write these buffers to disk. Free buffers are empty and haven't been used yet. Dirty buffers are those which needed to be moved to write list.

When oracle server process requires a specific data block, it first searches it in Buffer cache. If it finds required block, it is directly accessed and this event is known as **Cache Hit**. If searching in Buffer cache fails then it is read from datafile on the disk and the event is called **Cache Miss**. If the required block is not found in Buffer cache then process needs a free buffer to read data from disk. It starts search for free buffer from least recently used end of LRU list. In process of searching, if user process finds dirty block in LRU list it shifts them to **Write List**. If the process can not find free buffers until certain amount of time then process signals DBWn process to write dirty buffers to disks.

By default accessed buffers are moved to most recently used end of the LRU list. Search for free buffers is initiated from least recently used end of LRU list, this means that recently accessed buffers are kept in cache for longer time. But when a Full table scan happens, oracle process puts the blocks of table to least recently used end of LRU list. This means that they are quickly re-acclaimed by oracle process. When a table is created, a storage parameter **Cache | NoCache | Cache Reads** is required. If a table is created with **Cache** parameter, then the data block of table are added to most recently used end inspite of full table scan.

Size of the Database Buffer Cache

Oracle allows different block size for different tablespaces. A standard block size is defined in **DB_BLOCK_SIZE** initialization parameter. System tablespace uses standard block size. **DB_CACHE_SIZE** parameter is used to define size for Database buffer cache. For example to create a cache of 800 mb, set parameter as below

DB_CACHE_SIZE=800M

If you have created a tablespace with block size different from standard block size, for example your standard block size is 4k and you have created a tablespace with 8k block size then you must create a 8k buffer cache as below.

DB_8K_CACHE_SIZE=256M

Keep Buffer Pool and Recycle Buffer Pool

Data required by oracle user process is loaded into buffer cache, if it is not already present in cache. Proper memory tuning is required to avoid repeated disk access for the same data. This means that there should be enough space in buffer cache to hold required data for long time. If same data is required in very short intervals then such data should be permanently pinned into memory. Oracle allows us to use multiple buffers. Using multiple buffers we can control that how long objects should be kept in memory.

Keep Buffer Pool

Data which is frequently accessed should be kept in Keep buffer pool. Keep buffer pool retains data in the memory. So that next request for same data can be entertained from memory. This avoids disk read and increases performance. Usually small objects should be kept in Keep buffer. **DB_KEEP_CACHE_SIZE** initialization parameter is used to create Keep buffer Pool. If **DB_KEEP_CACHE_SIZE** is not used then no Keep buffer is created. Use following syntax to create a Keep buffer pool of 40 MB.

DB_KEEP_CACHE_SIZE=40M

To pin an object in Keep buffer pool use **DBMS_SHARED_POOL.KEEP** method.

Recycle Buffer Pool

Blocks loaded in Recycle Buffer pool are immediately removed when they are not being used. It is useful for those objects which are accessed rarely. As there is no more need of these blocks so memory occupied by such blocks is made available for others data. For example if ASM is enabled then available memory can be assigned to other SGA components. Use following syntax to create a Recycle Buffer Pool

DB_RECYCLE_CACHE_SIZE=20M

Default Pool

If an object is not assigned a specific buffer pool then its blocks are loaded in default pool **DB_CACHE_SIZE** initialization parameter is used to create Default Pool. For more information on Default Pool visit following link, <http://exploreoracle.com/2009/03/31/database-buffer-cache/>

BUFFER_POOL value in storage clause of schema objects lets you to assign an object to a specific Buffer pool. Value of **BUFFER_POOL** can be **KEEP, RECYCLE** or **DEFAULT**.

90. When the SMON Process perform Instance Crash Recovery?

Only at the time of startup after abort shutdown

91. Which dynamic view can be queried when a database is started up in no mount state?

V\$INSTANCE

92. Which two tasks occur as a database transitions from the mount stage to the open stage?

The online data files & Redo log files are opened.

93. In which situation is it appropriate to enable the restricted session mode?

Exporting a consistent image of a large number of tables

94. What is the component of an Oracle instance?

The SGA

95. Which process is involved when a user starts a new session on the database server?

The Oracle server process

96. In the event of an Instance failure, which files store command data NOT written to the datafiles?

Online redo logs

97. When are the base tables of the data dictionary created?

When the database is created

98. Sequence of events takes place while starting a Database is?

Instance started, Database mounted & Database opened

99. The alert log will never contain information about which database activity?

Performing operating system restore of the database files

100. Where can you find the non-default parameters when an instance is started?

Alert log

101. Which tablespace is used as the temporary tablespace if TEMPORARY TABLESPACE is not specified for a user?

SYSTEM

102. User SCOTT creates an index with this statement: CREATE INDEX emp_idx on employee (empno). In which tablespace would be the index created?

SCOTT'S default tablespace

103. Which data dictionary view shows the available free space in a certain tablespace?

DBA_FREE_SPACE

104. Which method increase the size of a tablespace?

Add a datafile to a tablespace.

105. What does the command ALTER DATABASE . . . RENAME DATAFILE do?

It updates the control file.

106. Can you drop objects from a read-only tablespace?

Yes

107. SYSTEM TABLESPACE can be made off-line?

No

108. Data dictionary can span across multiple Tablespaces?

No

109. Multiple Tablespaces can share a single datafile?

No

110. All datafiles related to a Tablespace are removed when the Tablespace is dropped?

No

111. What is a default role?

A role automatically enabled when the user logs on.

112. Who is the owner of a role?

Nobody

113. When granting the system privilege, which clause enables the grantee to further grant the privilege to other users or roles?

WITH ADMIN OPTION

114. Which view will show a list of privileges that are available for the current session to a user?

SESSION_PRIVS

115. Which view shows all of the objects accessible to the user in a database?

ALL_OBJECTS

116. Which statement about profiles is false?

Profiles are assigned to users, roles, and other profiles.

117. Which password management feature is NOT available by using a profile?

Password change

118. Which resource can not be controlled using profiles?

PGA memory allocations

119. You want to retrieve information about account expiration dates from the data dictionary. Which view do you use?

DBA_USERS

120. It is very difficult to grant and manage common privileges needed by different groups of database users using roles?

No

121. Which data dictionary view would you query to retrieve a table's header block number?

DBA_SEGMENTS

122. When tables are stored in locally managed tablespaces, where is extent allocation information stored?

Corresponding tablespace itself

123. Which of the following three portions of a data block are collectively called as Overhead?

Table directory, row directory and data block header

124. Can a tablespace hold objects from different schemes?

Yes

125. Which data dictionary view would you query to retrieve a table's header block number?

DBA_SEGMENTS

126. What is default value for storage parameter INITIAL in 10g if extent management is Local?

40k

127. Using which package we can convert Tablespace from DMTS to LMTS?

DBMS_SPACE_ADMIN

128. Is it Possible to Change ORACLE Block size after creating database?

No

129. Locally Managed table spaces will increase the performance?

TRUE

130. Index is a Space demanding Object?

Yes

131. What is a potential reason for a Snapshot too old error message?

An ITL entry in a data block has been reused.

132. An Oracle user receives the following error? ORA-01555 SNAPSHOT TOO OLD, What is the possible solution?

https://blogs.oracle.com/db/entry/troubleshooting_ora_1555

<http://oracle-randolf.blogspot.in/2009/04/read-consistency-ora-01555-snapshot-too.html>

Increase the extent size of the rollback segments.

Explanation-1:

Oracle uses for its read consistency model a true multi-versioning approach which allows readers to not block writers and vice-versa, writers to not block readers. Obviously this great feature allowing highly concurrent processing doesn't come for free, since somewhere the information to build multiple versions of the same data needs to be stored.

Oracle uses the so called undo information not only to rollback on-going transactions but also to re-construct old versions of blocks if required. Very simplified when reading data Oracle knows the point in time (which corresponds to an internal counter called SCN, System Change Number) that data needs to be consistent with. In the default READ COMMITTED isolation mode this point in time is defined when a statement starts to execute. You could also say at the moment a statement starts to run its result is pre-ordained. When Oracle processes a block it checks if the block is "old" enough and if it discovers that the block content is too new (has been changed by other sessions but the current access is not supposed to see this updated content according to the point-in-time assigned to the statement execution) it will start to create a copy of the block and use the information available from the corresponding undo segment to re-construct an older version of the block. Note that this process can be iterative: If after re-constructing the older version of the block it's still not sufficiently old more undo information will be used to go further back in

time.

Since the undo information of transactions that have been committed is marked as re-usable Oracle is free to overwrite the corresponding undo data under certain circumstances (e.g. no more free space left in the UNDO tablespace). If now an older version of a block needs to be created but the corresponding undo information required to do so has been overridden, the infamous "ORA-01555 snapshot too old" error will be raised, since the required read-consistent view of the data can not be generated any longer.

In order to avoid this error starting from 10g on you only need to have a sufficiently large UNDO tablespace in automatic undo management mode so that the undo information required to create old versions of the blocks doesn't get overridden prematurely. In 9i you need to set the UNDO_RETENTION parameter according to the longest expected runtime of your queries and of course have sufficient space in the UNDO tablespace to allow Oracle to adhere to this setting.

So until now Oracle was either able to provide a consistent view of the data according to its read-consistency model, or you would get an error message if the required undo data wasn't available any longer.

Enter the SCN_ASCENDING hint: As already mentioned by [Martin Berger](#) and [Chandra Pabba](#) Oracle officially documented the SCN_ASCENDING hint for Oracle 11.1.0.7 in Metalink Note 6688108.8 (Enhancement: Allow ORA-1555 to be ignored during table scan).

Explanation-2:

The ORA-1555 errors can happen when a query is unable to access enough undo to build a copy of the data at the time the query started. Committed "versions" of blocks are maintained along with newer uncommitted "versions" of those blocks so that queries can access data as it existed in the database at the time of the query. These are referred to as "consistent read" blocks and are maintained using Oracle undo management.

Diagnosis:

Due to space limitations, it is not always feasible to keep undo blocks on hand for the life of the instance. Oracle Automatic Undo Management (AUM) helps to manage the time frame that undo blocks are stored. The time frame is the "retention" time for those blocks.

There are several ways to investigate the ORA-1555 error. In most cases, the error is a legitimate problem with getting to an undo block that has been overwritten due to the undo "retention" period having passed.

AUM will automatically tune up and down the "retention" period, but often space limitations or configuration of the undo tablespace will throttle back continuous increases to the "retention" period.

Explanation-3:

1. Problem:

Below are the settings for the undo tablespace:

undo_retention - 1200

undo_management - AUTO

The user encounters the following error in the job that is running in the database:

Ora-01555 snapshot too old error.

2. Impact: Medium to high because it would effect the long running queries due to insufficient undo tablespace thus impacting performance. Also could be a part of a batch process.

3. Solutions: The Ora-01555 snapshot too old error occurs when the undo tablespace storage space is smaller as compared to the space needed by long running queries. It could also occur because of inappropriate (too small) value of the undo_retention. Undo_retention specifies the time period (in seconds) till which the system retains undo ie. undo would be retained for at least the time specified in this parameter. The undo_retention parameter would be efficient if the current undo tablespace has enough space. If there is an active transaction which requires undo space and there is not enough available space, then the system reuses unexpired undo space. This causes some queries to fail with a snapshot too old error message. The underlying technology that undo supports is the Oracle read consistency mechanism.

Below are the remedies to address and remedy this error:

1. Reduce and delay extent reuse by increasing the size of the undo tablespace and the undo_retention parameter.
2. Try not to do a fetch between the commits. So, if a cursor was opened before the last commit don't fetch the cursor as it is still performing actions in the current sessions.
3. Don't perform frequent commits as this would reduce the size of the undo tablespace and also the queries would take more time.
4. Try to perform the long-running queries when the system has the least load of DML transactions.
5. Set a large value for the database block size (db_block_size) parameter to reduce and delay extent reuse.

6. Run separate transactions while the sensitive long-running queries are taking place only when it is very important and the transactions are not dependent on each other and do not prejudice each others performance.
7. Before you run long-running and sensitive sql queries make sure that you have sufficient and optimal undo tablespace. If you do not have sufficient undo tablespace manually resize it to prevent rollback failure thus preventing the error.
8. You can also calculate the size of the optimal undo_retention, undo tablespace and the db_block_size before hand.
9. You can manually manage the usage, size and the amount of the rollback segments.

133. The status of the Rollback segment can be viewed through?

DBA_ROLLBACK_SEG

134. Explicitly we can assign transaction to a rollback segment?

TRUE

135. Are uncommitted transactions written to flashback redologs?

Yes

136. Is it possible to do flashback after truncate?

No

137. Can we restore a dropped table after a new table with the same name has been created?

Yes

138. Which following command will clear database recyclebin?

Purge recyclebin

139. What is the OPTIMAL parameter?

length of a rollback segment.

140. Flashback query time depends on?

Undo_retention

141. Can we create spfile in shutdown mode?

Yes

142. Can we alter static parameters by using scope=both?

No

143. Can we take backup of spfile in RMAN?

Yes

144. Does Drop Database command removes spfile?

Yes

145. Using which SQL command we can alter the parameters?

Alter system

146. OMF database will Improve the performance?

No

147. Max number of controlfiles that can be multiplexed in an OMF database?

5

148. Which environment variable is used to help set up Oracle names?

TNS_ADMIN

149 Which Net8 component waits for incoming requests on the server side?

Listener

150. What is the listener name when you start the listener without specifying an argument?

LISTENER

151. When is a request sent to a listener?

After name resolution.

152. In which file is the information that host naming is enabled stored?

sqlnet.ora

153. Which protocols can oracle Net 11g Use?

TCP

154. Which of the following statements about listeners is correct?

Multiple listeners can share one network interface card.

155. Can we perform DML operation on Materialized view?

No

156. Materialized views are schema objects, that can be used to summarize pre compute replicate and distribute data?

True

157. Does a materialized view occupies space?

Yes

158. Can we name a Materialized View log?

No

159. How to improve sqlldr (SQL*Loader) performance?

160. By using which view can a normal user see public database link?

ALL_DB_LINKS

161. Can we change the refresh interval of a Materialized View?

YES

162. Can we use a database link even after the target user has changed his password?

Yes

163. Can we convert a materialized view from refresh fast to complete?

Yes

164. A normal user can create public database link?

False

165. If we truncate the master table, materialized view log on that table?

Will be dropped

166. What is the correct procedure for multiplexing online redo logs?

Issue the ALTER DATABASE. . . ADD LOGFILE MEMBER command.

167. In which situation would you need to create a new control file for an existing database?

When MAXLOGMEMBERS needs to be changed.

168. When configuring a database for ARCHIVELOG mode, you use an initialisation parameter to specify which action?

To Store Archive log Files

169. Which command creates a text backup of the control file?

ALTER DATABASE BACKUP CONTROLFILE TO TRACE;

170. You are configuring a database for ARCHIVELOG mode. Which initialization parameter should you use?

LOG_ARCHIVE_DEST

171. How does a DBA specify multiple control files?

By listing the files in the CONTROL_FILES parameter.

172. Which dynamic view should a DBA query to obtain information about the different sections of the control file?

V\$CONTROLFILE_RECORD_SECTION

173. What is the characteristics of the control file?

It must be updated at every log switch.

174. Which statements about online redo log members in a group are true?

All members in a group are the same size.

175. Which command does a DBA use to list the current status of archiving?

ARCHIVE LOG LIST;

176. When performing an open database backup, which statement is NOT true?

The database can be open but only in READ ONLY mode.

177. Which task can a DBA perform using the export/import facility?

Transport tablespaces between databases.

178. Why does this command cause an error?

exp system/manager inctype=full file=expdat.dnp

The full=y parameter needs to be specified.

179. Which import option do you use to create tables without data?

ROWS

180. Which export option will generate code to create an initial extent that is equal to the sum of the sizes of all the extents currently allocated to an object?

COMPRESS

181. Can I take 1 dump file set from my source database and import it into multiple databases?

Yes

181. EXP command is used?

To take Backup of the Oracle Database

182. Can we export a dropped table?

No

183. What is the default value for IGNORE parameter in EXP/IMP?

No

184. Why is Direct Path Export Faster?

This option By Passes the SQL Layer

185. Is there a way to estimate the size of an export job before it gets underway?

Yes

186. Can I monitor a Data Pump Export or Import job while the job is in progress?

Yes

187. If a job is stopped either voluntarily or involuntarily, can I restart it?

Yes

188. Does Data Pump support Flashback?

Yes

189. If the tablespace is Read Only, Can we export objects from that tablespaces?

Yes

190. Dump files exported using traditional EXP are compatible with DATAPUMP?

False

191. Before a DBA creates a transportable tablespace, which condition must be completed?

The target system is in the same operating system.

192. Can we transport tablespace from one database to another database which is having SYS owned objects?

No

193. What is default value for TRANSPORT_TABLESPACE Parameter in EXP?

No

194. How to find whether tablespace is created in that database or transported from another database?

Dba_tablespaces

195. Can we Perform TTS using EXPDP?

Yes

196. Can we Transport Tablespace which has Materialized View in it?

No

197. When would a DBA need to perform a media recovery?

When a data file is not synchronized with the other data files, redo logs, and control files.

198. Why would you set a data file offline when the database is in MOUNT state?

To allow for automatic data file recovery.

199. What is the causes of media failures?

There is a head crash on the disk containing a database file.

200. Which of the following would not require you to perform an incomplete recovery?

Instance failure

201. In what scenario you have to open a database with reset logs option?

All of the above

202. Is it possible take consistent backup if the database is in NOARCHIVELOG mode?

Yes

203. Database is in Archivelog mode and Loss of unbackedup datafile is?

Complete Online Recovery

204. You should issue a backup of the control file after issuing which command?

CREATE TABLESPACE

205. The alert log will never contain specific information about which database backup activity) ?

Performing an operating system backup of the database files.

206. A tablespace becomes unavailable because of a failure. The database is running in NOARCHIVELOG mode?

What should the DBA do to make the database available?

Restore the data files, redo log files, and control files from an earlier copy of a full database backup.

207. How often does a read-only tablespace need to be backed up?

Only once after the tablespace becomes read-only

208. With the instance down, how would you recover a lost control file?

Restore backup control file & recover using backup controlfile

209. Which action does Oracle recommend after a DBA recovers from the loss of the current online redo-log?

Back up the database

210. Which command creates a text backup of the control file?

ALTER DATABASE BACKUP CONTROLFILE TO TRACE;

211. Which option is used in the parameter file to detect corruptions in an Oracle data block?

DBVERIFY

212. Your database is configured in ARCHIVELOG mode. Which backups cannot be performed?

Online control file backups using the ALTER CONTROLFILE BACKUP command

213. You are using hot backup without being in archivelog mode, can you recover in the event of a failure?

No

214. Which following statement is true when tablespaces are put in backup mode for hot backups?

High Volume of REDO is generated

215. Can Consistant Backup be performed when the database is open?

No

216. Can we shutdown the database if it is in BEGIN BACKUP mode?

Yes

217. Which data dictionary view helps you to view whether tablespace is in BEGIN BACKUP Mode or not?

V\$backup

218. Which command is used to allow RMAN to store a group of commands in the recovery catalog?

CREATE SCRIPT

219. When using Recovery Manager without a catalog, the connection to the target database?

Can be a local or a remote connection.

220. Work is done by Recovery Manager through?

Operating system commands

221. You perform an incomplete database recovery using RMAN. Which state of target database is needed?

Mount

222. Is it possible to perform Transportable tablespace(TTS) using RMAN ?

Yes

223. Which type of file does Not RMAN include in its backups?

Online redo-logs

224. When using Recovery Manager without a catalog, the connection to the target database should be made as?

A user with SYSDBA privilege

225. RMAN online backup generates excessive Redo information?

Flase

226. Which background process will be invoked when we enable BLOCK CHANGE TRACKING?

CTWr

227. Where should a recovery catalog be created?

In the target database

228. How to list restore points in RMAN?

RC_RESTORE_POINT view

229. Without LIST FAILURE can we say ADVISE FAILURE in Data Recovery Advisor?

Yes

230. Import Catalog Command is used for?

To Merge Two diff catalogs

231. Interfile backup parallelism does?

Divide files into Multiple sections & take backup parallel

232. What is the difference between pfile and spfile. Where these files are located?

233. What will you do if pfile and spfile file is deleted? Can you start the database?

pfile or init.ora is a text file, hence setting any oracle init parameters in this file requires restarting database. With spfile we can dynamically set certain oracle init parameters without restarting the instance.

Example: alter system set DB_CACHE_SIZE=2G scope=both; both means memory and spfile both. Location of pfile/spfile is \$ORACLE_HOME/dbs

if init.ora/spfile is lost, we can manually create a pfile using any other database pfile. Edit the pfile as per the db_name, control_files etc.

And then start the database. Later on we can create spfile from pfile.

234. What is the difference between Static and Dynamic init.ora/spfile parameters?

Changing Oracle Static parameters requires instance restart to make them effective.

Dynamic parameters are immediately effective in running Oracle Instance and does not require restart.

235. What is the complete syntax to set DB_CACHE_SIZE in memory and spfile?

alter system set DB_CACHE_SIZE=2G scope=both;

236. How do we configure multiple Buffer Cache in Oracle. Whats the benefit? Does setting multiple Cache requires Database Restart?

We can set multiple Buffer Cache by setting DB_NK_CACHE_SIZE dynamic parameter in pfile or spfile. NK can be (2K, 4K, 8K, 16, 32K)

if db_block_size=8K then DB_8K_CACHE_SIZE is not allowed.

OLTP database has small transactions so they need small block size(2k,4k,8k) and hence 2k, 4k, 8k Cache Size.

Datawarehouse database work on big transaction that effect big tables hence we need bigger block size(8k,16k,32k)

If [a database](#) is mix having both OLTP and Datawarehouse needs we need to configure Multiple Block size and Also create Tablespace of different block size using **BLOCKSIZE** syntax.

Multiple Buffer Cache parameters are Dynamic database restart is not needed.

237. What is Oracle Golden Gate?

It a software used for Replicating data from one database to another. The source and target can be Microsoft Sql Server, Oracle , IBM DB2, Sybase, MYSQL running on any OS.

238. Can we create Tablespaces of multiple Block Sizes. If yes, what is the Syntax?

YES it is possible. We need to set Buffer Caches of corresponding block size, and create the tablespace with BLOCKSIZE syntax.

For example if we need Tablespace of 32K size we will use following steps:

alter system set db_32k_cache_size=2G scope=both;

create tablespace hr_data datafile '/u01/app/oracle/oradata/hrprd/hr_data01.dbf' size 1G BLOCKSIZE 32K;

239. How do you calculate the size of oracle memory areas Buffer Cache, Log Buffer, Shared Pool, PGA etc?

We allocate 70- 80% of Unix Server RAM to Oracle and then allocate 60-70% to Buffer Cache

20-30% to PGA and remaining to Shared Pool and Log Buffer.

240. What is OMF? What spfile parameters are used to configure OMF. What is the benefit?

OMF is oracle managed files and it is used to simplify the syntax for Datafile, Logfile, Tablespace and controlfile creation.

init.ora/spfile parameters to configure OMF are:

db_create_file_dest ,db_create_online_log_dest_n (n=1 to 5)

241. What is Database Cloning? Why Cloning is needed? What are the steps to clone a database?

Cloning is used to create dev and test database from production on a different machine. Refer to blog for complete steps.

242. What is Oracle Streams?

Oracle Streams is used to Replicate/Transfer Data from one Oracle Database to another Oracle Database.

243. There are 2 control files for a database. What will happen when 1 control file is deleted and you try to start database? How you will fix this problem?

If one Control file is missing out of 2, Oracle will complain when we start database. To fix this we need to modify CONTROL_FILES init.ora/spfile parameter and remove the entry for deleted control file. We can also copy control01.ctl to control02.ctl and then start the database and it will fix the error.

244. What is Dynamic performance view and What is Data Dictionary Views. Give some examples of each?

During Database operation, Oracle maintains a set of virtual tables/view that record current database activity.

These are called dynamic performance views because they are continuously updated while a database is open and in use.

These are also called V\$ views. GV\$ views used in RAC are same as V\$ views but

Dynamic performance view (v\$datafile, v\$controlfile, v\$sql, v\$transaction)

Data Dictionary Views (dba_users, dba_tablespaces, dba_sys_privs)

245. You are working in database that does lot of Sorting , i.e SELECT queries use a lot of ORDER BY and GROUP BY? What Oracle memory area and Physical File/Tablespace you need to tune and How?

We need bigger PGA and TEMP tablespace space to support excessive Sorting.

246. Why we upgrade a database. What are the steps to upgrade database. Any errors you got during upgrade?

Every few years an Oracle Database Version gets desupported by Oracle so we need to upgrade to newer Oracle version. Currently Oracle 9i is not supported by oracle. Also we need to upgrade to newer versions to use the new features/tools provided by newer Oracle version like 11gr1/11gr2.

We should use utlu112i.sql , utlu112s.sql and DBUA/catupgrd.sql to upgrade a database to 11gr2.

247. What is MEMORY_TARGET not supported error. How do you fix it?

This error occurs when linux shared memory or swap space is defined less. Increase its size to fix the error.

248. What are the steps to manually create a database?

Create init.ora/spfile

startup nomount

Run Create Database command to manually create database.

Refer to blog for exact steps.

249. A DBA ran a delete statement to delete all records on a table. The table has 50 Million rows? While Delete is running his SQLPLUS session terminate abnormally? What oracle will do internally?

When the session terminates PMON Process will rollback this transaction.

Next question- Which query/view you will use to monitor the Rollback/Undo that Oracle is doing
V\$TRANSACTION columns used_ublk and used_urec

250. What is Oracle Dataguard?

Dataguard is use to configure a Standby Database at a Remote location. Dataguard provides database protection in case of natural disaster(Earthquake, Flood) when complete Datacenter is lost and database is damaged. Business will using Standby database present at remote location.

251. Can we change the DB_BLOCK_SIZE? if Yes. What are the steps?

we can not change the db_block_size using any oracle commands. If it is required here are steps
export data from old database using Datapump(expdp) into a .dmp file.

Create a new databsae with db_block_size to any of the values (2k,4k,8k,16k,32K) as per your requirement.

Import data into new database using the using the .dmp file

252. Explain the Oracle Architecture?

Oracle consists of Instance and Physical Database.

Instance has SGA, PGA and Background Process.

Physical Database consists of Datafiles, Control files, Log files and Archive log files

253. What happens internally in Oracle when a User Connects and run a SELECT Query? What SGA areas and background processes are involved?

254. How do you create a tablespace, undo tablespace and temp tablespace. What are the Syntax?

Tablespace -> create tablespace ...

Undo Tablespace -> create undo tablespace..

Temp Tablespace -> create temporary tablespapce...

255. As a HR user you logged in and Creating a EMP_BIG Table and inserting 10 lac rows? While inserting 10 lac rows you got error ORA-01688: unable to extend table EMP_BIG by 512 in tablespace HR_DATA? What are the two ways to fix this Tablespace error?

1) Resize the existing tablespace datafile to add more space

2) Add new datafile to tablespace to add more space

256. What are the steps to rename a database?

Shutdown Immediate.

Startup mount

Then use the **NID command** to rename a database.

Refer to blog for exact steps.

257. What is the syntax to create a user and roles?

create user username identified by pass1 default tablespace hr_data temporary tablespace temp;

create role hr_read_role;

258. What are the 3 init.ora parameters to manage UNDO? What is their usage?

UNDO_TABLESPACE

UNOD_MANAGEMENT=AUTO/MANAUAL

UNDO_RETENTION

259. What is Snapshot too old error? How do you fix it?

Snapshot too old error occurs when a long running queries tries to read data that from Undo Tablespace which is already overwritten by some new Transactions.

To fix this error, We need to create proper size Undo Tablespace. Query the v\$UNDOSTAT for undo tablespace size recommendation.

Also we can set RETENTION GUARANTEE for a tablespace, But it is not recommended.

260. What is Undo Retention Guarantee. How do we set it. What is the Pros and Cons of setting it? When Retention Guarantee is set for Undo Tablespace, committed transactions are not overwritten for UNDO_RETENTION period. If we set this the new Transactions will fail if there is less space in Undo Tablespace and hence its very Risky and not recommended.

261. What are System Privileges and Object Privileges? Give some examples? What Data Dictionary view we use to check both?

System privileges are generic database privileges e.g CREATE TABLE, CREATE VIEW, CREATE SESSION

To see System privileges query : SELECT * FROM DBA_SYS_PRIVS

Object privileges are on specific database object/table e.g SELECT ON EMPLOYEE, DELETE ON EMPLOYEE

To see Object privileges query : SELECT * FROM DBA_TAB_PRIVS

262. What is PGA? What information is stored in PGA?. What is PGA Tuning?

PGA is process global are used to store sorting data, bind variable etc. PGA tuning is setting the proper size of PGA_AGGREGATE_TARGET init.ora/spfile parameter for better performance.

263. What are the steps to identify a slow running SQL and tune it?

- a) Monitor sessions to find slow running sql.
- b) Generate Explain Plan/SQL plan to find the root cause of slowness.
- c) Tune the sqls by Creating indexes or Using SQL Hints or by rewriting a Bad sql

264. What are all the preparation work a DBA need to do before installing Oracle?

Set linux kernel parameters.

Install Oracle recommended Linux packages.

For all steps Refer to Oracle Installation blog.

265. Any error that you got during Oracle Installation and how did you fix it?

Example of Oracle errors/warnings are: Kernel parameters not set, Linux packages missing, In sufficient Memory for Oracle.

266. What is default tablespace and temporary tablespace?

Default Tablespace : Place where a user creates objects if the user does not specify some other tablespace. Note that having a default tablespace does not imply that the user has the privilege of creating objects in that tablespace, nor does the user have a quota of space in that tablespace in which to create objects. Both of these are granted separately.

Temporary tablespace: This is a place where temporary objects, such as sorts and temporary tables, are created on behalf of the user by the instance. No quota is applied to temporary tablespaces.

267. Which privilege allows you to select from tables owned by other users?

The SELECT ANY TABLE privilege allows to select from tables owned by other users

268. What command we use to revoke system privilege?

Revoke select table from username

269. How do we create a Role?

A role is a named group of related privileges that are granted to users or to other roles. A DBA manages privileges through roles.

To create a role:

CREATE ROLE role_name;

OR

1. In Enterprise Manager Database Control, click the Server tab and then click Roles under the Security heading.
2. Click the Create button

270. Difference between Non-Deferred and deferred constraint?

Nondeferred constraints, also known as immediate constraints, are enforced at the end of every DML statement. A constraint violation causes the statement to be rolled back. If a constraint causes an action such as delete cascade, the action is taken as part of the statement that caused it. A constraint that is defined as nondeferrable cannot be changed to a deferrable constraint. For nondeferrable constraints, the primary key and unique key constraints need unique indexes; if the column or columns already have a non-unique index, constraint creation fails because those indexes cannot be used for a unique or primary key.

Deferred constraints are constraints that are checked only when a transaction is committed.

If constraint violations are detected at commit time, the entire transaction is rolled back. These constraints are most

useful when both the parent and child rows in a foreign key relationship are entered at the same time, as in the case of an order entry system in which the order and the items in the order are entered at the same time. For deferrable constraints, primary key and unique keys need non-unique indexes; if the column or columns already have a unique index on them, constraint creation fails because those indexes cannot be deferred.

271. Difference between varchar and varchar2 data types?

Varchar can store upto 2000 bytes and varchar2 can store upto 4000 bytes. Varchar will occupy space for NULL values and Varchar2 will not occupy any space. Both are differed with respect to space.

272. In which language Oracle has been developed?

Oracle has been developed using C Language.

273. What is RAW datatype?

RAW datatype is used to store values in binary data format. The maximum size for a raw in a table is 32767 bytes.

274. What is the use of NVL function?

The NVL function is used to replace NULL values with another or given value. Example is –

NVL(Value, replace value)

275. Whether any commands are used for Months calculation? If so, what are they?

In Oracle, months_between function is used to find number of months between the given dates. Example is –

Months_between(Date 1, Date 2)

276. What are nested tables?

Nested table is a data type in Oracle which is used to support columns containing multi valued attributes. It also hold entire sub table.

277. What is COALESCE function?

COALESCE function is used to return the value which is set to be not null in the list. If all values in the list are null, then the coalesce function will return NULL.

Coalesce(value1, value2,value3,...)

278. What is BLOB datatype?

A BLOB data type is a varying length binary string which is used to store two gigabytes memory. Length should be specified in Bytes for BLOB.

279. How do we represent comments in Oracle?

Comments in Oracle can be represented in two ways –

Two dashes(–) before beginning of the line – Single statement

/*— */ is used to represent it as comments for block of statement

280. What is DML?

Data Manipulation Language (DML) is used to access and manipulate data in the existing objects. DML statements are insert, select, update and delete and it won't implicitly commit the current transaction.

281. What is the difference between TRANSLATE and REPLACE?

Translate is used for character by character substitution and Replace is used substitute a single character with a word.

282. How do we display rows from the table without duplicates?

Duplicate rows can be removed by using the keyword DISTINCT in the select statement.

283. What is the usage of Merge Statement?

Merge statement is used to select rows from one or more data source for updating and insertion into a table or a view. It is used to combine multiple operations.

284. What is NULL value in oracle?

NULL value represents missing or unknown data. This is used as a place holder or represented it in as default entry to indicate that there is no actual data present.

285. What is USING Clause and give example?

The USING clause is used to specify with the column to test for equality when two tables are joined.

[sql]Select * from employee join salary using employee ID[/sql]

Employee tables join with the Salary tables with the Employee ID.

286. What is key preserved table?

A table is set to be key preserved table if every key of the table can also be the key of the result of the join. It guarantees to return only one copy of each row from the base table.

287. What is WITH CHECK OPTION?

The WITH CHECK option clause specifies check level to be done in DML statements. It is used to prevent changes to a view that would produce results that are not included in the sub query.

288. What is the use of Aggregate functions in Oracle?

Aggregate function is a function where values of multiple rows or records are joined together to get a single value output. Common aggregate functions are –

Average

Count

Sum

289. What do you mean by GROUP BY Clause?

A GROUP BY clause can be used in select statement where it will collect data across multiple records and group the results by one or more columns.

290. What is a sub query and what are the different types of subqueries?

Sub Query is also called as Nested Query or Inner Query which is used to get data from multiple tables. A sub query is added in the where clause of the main query.

There are two different types of subqueries:

Correlated sub query

A Correlated sub query cannot be as independent query but can reference column in a table listed in the from list of the outer query.

Non-Correlated subquery

This can be evaluated as if it were an independent query. Results of the sub query are submitted to the main query or parent query.

291. What is cross join?

Cross join is defined as the Cartesian product of records from the tables present in the join. Cross join will produce result which combines each row from the first table with the each row from the second table.

292. What are temporal data types in Oracle?

Oracle provides following temporal data types:

Date Data Type – Different formats of Dates

TimeStamp Data Type – Different formats of Time Stamp

Interval Data Type – Interval between dates and time

293. How do we create privileges in Oracle?

A privilege is nothing but right to execute an SQL query or to access another user object. Privilege can be given as system privilege or user privilege.

```
[sql]GRANT user1 TO user2 WITH MANAGER OPTION;[/sql]
```

294. What is VArray?

VArray is an oracle data type used to have columns containing multivalued attributes and it can hold bounded array of values.

295. How do we get field details of a table?

Describe <Table_Name> is used to get the field details of a specified table.

296. What is the difference between rename and alias?

Rename is a permanent name given to a table or a column whereas Alias is a temporary name given to a table or column. Rename is nothing but replacement of name and Alias is an alternate name of the table or column.

297. What is a View?

View is a logical table which based on one or more tables or views. The tables upon which the view is based are called Base Tables and it doesn't contain data.

298. What is a cursor variable?

A cursor variable is associated with different statements which can hold different values at run time. A cursor variable is a kind of reference type.

299. What are cursor attributes?

Each cursor in Oracle has set of attributes which enables an application program to test the state of the cursor. The attributes can be used to check whether cursor is opened or closed, found or not found and also find row count.

300. What are SET operators?

SET operators are used with two or more queries and those operators are Union, Union All, Intersect and Minus.

301. How can we delete duplicate rows in a table?

Duplicate rows in the table can be deleted by using ROWID.

302. What are the attributes of Cursor?

Attributes of Cursor are

%FOUND

Returns NULL if cursor is open and fetch has not been executed

Returns TRUE if the fetch of cursor is executed successfully.

Returns False if no rows are returned.

%NOT FOUND

Returns NULL if cursor is open and fetch has not been executed

Returns False if fetch has been executed

Returns True if no row was returned

%ISOPEN

Returns true if the cursor is open

Returns false if the cursor is closed

%ROWCOUNT

Returns the number of rows fetched. It has to be iterated through entire cursor to give exact real count.

303. Can we store pictures in the database and if so, how it can be done?

Yes, we can store pictures in the database by Long Raw Data type. This datatype is used to store binary data for 2 gigabytes of length. But the table can have only on Long Raw data type.

304. What is an integrity constraint?

An integrity constraint is a declaration defined a business rule for a table column. Integrity constraints are used to ensure accuracy and consistency of data in a database. There are types – Domain Integrity, Referential Integrity and Domain Integrity.

305. What is an ALERT?

An alert is a window which appears in the center of the screen overlaying a portion of the current display.

306. What is hash cluster?

Hash Cluster is a technique used to store the table for faster retrieval. Apply hash value on the table to retrieve the rows from the table.

307. What are the various constraints used in Oracle?

Following are constraints used:

NULL – It is to indicate that particular column can contain NULL values

NOT NULL – It is to indicate that particular column cannot contain NULL values

CHECK – Validate that values in the given column to meet the specific criteria

DEFAULT – It is to indicate the value is assigned to default value

308. What is difference between SUBSTR and INSTR?

SUBSTR returns specific portion of a string and INSTR provides character position in which a pattern is found in a string.

SUBSTR returns string whereas INSTR returns numeric.

309. What is the parameter mode that can be passed to a procedure?

IN, OUT and INOUT are the modes of parameters that can be passed to a procedure.

310. What are the different Oracle Database objects?

There are different data objects in Oracle –

Tables – set of elements organized in vertical and horizontal

Views – Virtual table derived from one or more tables

Indexes – Performance tuning method for processing the records

Synonyms – Alias name for tables

Sequences – Multiple users generate unique numbers

Tablespaces – Logical storage unit in Oracle

311. What are the differences between LOV and List Item?

LOV is property whereas list items are considered as single item. List of items is set to be a collection of list of items. A list item can have only one column, LOV can have one or more columns.

312. What are privileges and Grants?

Privileges are the rights to execute SQL statements – means Right to connect and connect. Grants are given to the object so that objects can be accessed accordingly. Grants can be provided by the owner or creator of an object.

313. What is the difference between \$ORACLE_BASE and \$ORACLE_HOME?

Oracle base is the main or root directory of an oracle whereas ORACLE_HOME is located beneath base folder in which all oracle products reside.

314. What is the fastest query method to fetch data from the table?

Row can be fetched from table by using ROWID. Using ROW ID is the fastest query method to fetch data from the table.

315. What is the maximum number of triggers that can be applied to a single table?

12 is the maximum number of triggers that can be applied to a single table.

316. How to display row numbers with the records?

Display row numbers with the records numbers -

```
[sql]Select rownum, <fieldnames> from table;[/sql]
```

This query will display row numbers and the field values from the given table.

317. How can we view last record added to a table?

Last record can be added to a table and this can be done by –

```
[sql]Select * from (select * from employees order by rownum desc) where rownum<2;[/sql]
```

318. What is the data type of DUAL table?

The DUAL table is a one-column table present in oracle database. The table has a single VARCHAR2(1) column called DUMMY which has a value of 'X'.

319. What is difference between Cartesian Join and Cross Join?

There are no differences between the join. Cartesian and Cross joins are same. Cross join gives cartesian product of two tables – Rows from first table is multiplied with another table which is called cartesian product.

Cross join without where clause gives Cartesian product.

320. How to display employee records who gets more salary than the average salary in the department?

This can be done by this query –

```
[sql]Select * from employee where salary>(select avg(salary) from dept, employee where dept.deptno = employee.deptno;[/sql]
```

321. What is the difference between RMAN and a traditional hot backup?

RMAN is faster, can do incremental (changes only) backups, and does not place tablespaces into hotbackup mode.

322. What are bind variables and why are they important?

With bind variables in SQL, Oracle can cache related queries a single time in the SQL cache (area). This avoids a hard parse each time, which saves on various locking and latching resources we use to check objects existence and so on.

BONUS: For rarely run queries, especially BATCH queries, we explicitly DO NOT want to use bind variables, as they hide information from the Cost Based Optimizer.

BONUS BONUS: For batch queries from 3rd party apps like peoplesoft, if we can't remove bind variables, we can use bind variable peeking!

323. In PL/SQL, what is bulk binding, and when/how would it help performance?

Oracle's SQL and PL/SQL engines are separate parts of the kernel which require context switching, like between unix processes. This is slow, and uses up resources. If we loop on an SQL statement, we are implicitly flipping between these two engines. We can minimize this by loading our data into an array, and using PL/SQL bulk binding operation to do it all in one go!

324. Why is SQL*Loader direct path so fast?

SQL*Loader with direct path option can load data ABOVE the high water mark of a table, and DIRECTLY into the datafiles, without going through the SQL engine at all. This avoids all the locking, latching, and so on, and doesn't impact the db (except possibly the I/O subsystem) at all.

325. What are the tradeoffs between many vs few indexes? When would you want to have many, and when would it be better to have fewer?

Fewer indexes on a table mean faster inserts/updates. More indexes mean faster, more specific WHERE clauses possibly without index merges.

326. What is the difference between RAID 5 and RAID 10? Which is better for Oracle?

RAID 5 is striping with an extra disk for parity. If we lose a disk we can reconstruct from that parity disk.

RAID 10 is mirroring pairs of disks, and then striping across those sets.

RAID 5 was created when disks were expensive. Its purpose was to provide RAID on the cheap. If a disk fails, the IO subsystem will perform VERY slowly during the rebuild process. What's more your likelihood of failure increases dramatically during this period, with all the added weight of the rebuild. Even when it is operating normally RAID 5 is slow for everything but reading. Given that and knowing databases (especially Oracle's redo logs) continue to experience write activity all the time, we should avoid RAID5 in all but the rare database that is MOSTLY read activity. Don't put redologs on RAID5.

RAID10 is just all around goodness. If you lose one disk in a set of 10 for example, you could lose any one of eight other disks and have no troubles. What's more rebuilding does not impact performance at all since you're simply making a mirror copy. Lastly RAID10 perform exceedingly well in all types of databases.

327. When using Oracle export/import what character set concerns might come up? How do you handle them?

Be sure to set NLS_LANG for example to "AMERICAN_AMERICA.WE8ISO8859P1". If your source database is US7ASCII, beware of 8-bit characters. Also be wary of multi-byte characters sets as those may require extra attention. Also watch export/import for messages about any "character set conversions" which may occur.

328. Name three SQL operations that perform a SORT?

- a. CREATE INDEX
- b. DISTINCT
- c. GROUP BY
- d. ORDER BY
- f. INTERSECT
- g. MINUS
- h. UNION
- i. UNINDEXED TABLE JOIN

329. What is your favorite tool for day-to-day Oracle operation?

Hopefully we hear some use of command line as the answer!

330. What is the difference between Truncate and Delete? Why is one faster? Can we ROLLBACK both? How would a full table scan behave after?

Truncate is nearly instantaneous, cannot be rolled back, and is fast because Oracle simply resets the HWM. When a full table scan is performed on a table, such as for a sort operation, Oracle reads to the HWM. So if you delete every single solitary row in 10 million row table so it is now empty, sorting on that table of 0 rows would still be extremely slow.

331. What is the difference between a materialized view (snapshot) fast refresh versus complete refresh? When is one better, and when the other?

Fast refresh maintains a change log table, which records change vectors, not unlike how the redo logs work. There is overhead to this, as with a table that has a LOT of indexes on it, and inserts and updates will be slower. However if you are performing refreshes often, like every few minutes, you want to do fast refresh so you don't have to full-table-scan the source table. Complete refresh is good if you're going to refresh once a day. Does a full table scan on the source table, and recreates the snapshot/mview. Also inserts/updates on the source table are NOT impacted on tables where complete refresh snapshots have been created.

332. What does the NO LOGGING option do? Why would we use it? Why would we be careful of using it?

It disables the logging of changes to the redo logs. It does not disable ALL LOGGING, however as Oracle continues to use a base of changes, for recovery if you pull the plug on the box, for instance. However it will cause problems if you are using standby database. Use it to speed up operations, like an index rebuild, or partition maintenance operations.

333. Tell me about standby database? What are some of the configurations of it? What should we watch out for?

Standby databases allow us to create a copy of our production db, for disaster recovery. We merely switch mode on the target db, and bring it up as read/write. Can setup as master->slave or master->master. The latter allows the former prod db to become the standby, once the failure cause is remedied. Watch out for NO LOGGING!! Be sure we're in archive log mode.

334. What do you know about privileges?

A privilege is a right to execute a particular type of SQL statement or to access another user's object

Privileges are divided into two categories:

System privileges: Each system privilege allows a user to perform a particular database operation or class of database operations. For example, the privilege to create tablespaces is a system privilege.

Object privileges: Object privileges allow a user to perform a particular action on a specific object, such as a table, view, sequence, procedure, function, or package. Without specific permission, users can access only their own objects.

Oracle ASM FAQ

Questions

1. What is the use of ASM (or) Why ASM preferred over file system? Benefits?
2. Describe about ASM architecture?
3. How does database connects to ASM Instance?
4. What are the init parameters related to ASM?
5. What is rebalancing (or) what is the use of ASM_POWER_LIMIT?
6. What is significance of re-balance power?
7. In what situations we need to re-balance the disk?
8. In what situation asm instance will automatically re-balance the disk?
9. Explain about disk group managements?
10. What are different types of redundancies in ASM & explain?
11. How to copy file to/from ASM from/to file system?
12. How to find out the databases, which are using the ASM instance?
13. What is Striping and Mirroring? What are different types of striping and Mirroring in ASM & their differences?
14. What are Diskgroup's and Failuregroups?
15. Can ASM be used as replacement for RAID?
16. What are the background processes in ASM?
17. What are the file types that ASM support and keep in disk groups?
18. How many ASM Diskgroups can be created under one ASM Instance?
19. What process does the rebalancing?
20. How does ASM provides Redundancy?
21. Can we change the Redundancy for Diskgroup after its creation?
22. Unable to open the ASM instance. What is the reason?
23. Can ASM instance and database (rdbms) be on different servers?
24. Can we see the files stored in the ASM instance using standard unix commands.
25. Can we use ASM for storing Voting Disk/OCR in a RAC instance?
26. Does ASM instance automatically rebalances and takes care of hot spots?
27. What is ASMLIB?
28. What is SYSASM role?
29. Can we use BCV to clone the ASM Diskgroup on same host?
30. Can we edit the ASM Disk header to change the Diskgroup Name?
31. Whats is Kfed?
32. Can we use block devices for ASM Disks?
33. Is it mandatory to use disks of same size and characteristics for Diskgroups?
34. Do we need to install ASM and Oracle Database Software in different ORACLE_HOME?
35. What is the maximum size of Disk supported by ASM?
36. I have created Oracle database using DBCA and having a different home for ASM and Oracle Database. I see that listener is running from ASM_HOME. Is it correct?
37. How does the database interact with the ASM instance and how do I make ASM go faster?
38. Do I need to define the RDBMS FILESYSTEMIO_OPTIONS parameter when I use ASM?
39. Why Oracle recommends two diskgroups?
40. We have a 16 TB database. I'm curious about the number of disk groups we should use; e.g. 1 large disk group, a couple of disk groups, or otherwise?
41. We have a new app and don't know our access pattern, but assuming mostly sequential access, what size would be a good AU fit?
42. Would it be better to use BIGFILE tablespaces, or standard tablespaces for ASM?
43. What is the best LUN size for ASM?
44. In 11g RAC we want to separate ASM admins from DBAs and create different users and groups. How do we set this up?
45. Can my RDBMS and ASM instances run different versions?
46. Where do I run my database listener from; i.e., ASM HOME or DB HOME?

47. How do I backup my ASM instance?
48. When should I use RMAN and when should I use ASMCMD copy?
49. I'm going to do add disks to my ASM diskgroup, how long will this rebalance take?
50. We are migrating to a new storage array. How do I move my ASM database from storage A to storage B?
51. Is it possible to unplug an ASM disk group from one platform and plug into a server on another platform (for example, from Solaris to Linux)?
52. How does ASM work with multipathing software?
53. Is ASM constantly rebalancing to manage "hot spots"?
54. Draw the Diagram that how database interacts with ASM when a request is to read or open a datafile.
55. Can my disks in a diskgroup can be varied size? For example one disk is of 100GB and another disk is of 50GB. If so how does ASM manage the extents?
56. What is Intelligent Data Placement?
57. What is ASM preferred Mirror read? How does it useful?
58. What is ACFS?
59. What is ADVN?
60. What is ASM Template?
61. Why Oracle recommends two diskgroups. Why?

Answers

1. What is the use of ASM (or) Why ASM preferred over file system? Benefits?

(http://www.datadisk.co.uk/html_docs/oracle/asm.htm)

Explanation-1:

ASM is a [volume](#) manager and a [file system](#) for Oracle database files that supports single-instance Oracle Database and Oracle Real Application Clusters (Oracle RAC) configurations. ASM is Oracle's recommended storage management solution that provides an alternative to conventional volume managers, file systems, and raw devices. ASM uses [disk group](#) to store datafiles; an ASM disk group is a collection of disks that ASM manages as a unit. Within a disk group, ASM exposes a file system interface for Oracle database files. The content of files that are stored in a disk group are evenly distributed, or striped, to eliminate hot spots and to provide uniform performance across the disks. The performance is comparable to the performance of raw devices.

You can add or remove disks from a disk group while a database continues to access files from the disk group. When you add or remove disks from a disk group, ASM automatically redistributes the file contents and eliminates the need for downtime when redistributing the content.

The ASM volume manager functionality provides flexible server-based mirroring options. The ASM normal and high redundancy disk groups enable two-way and three-way mirroring respectively. You can use external redundancy to enable a Redundant Array of Inexpensive Disks (RAID) storage subsystem to perform the mirroring protection function.

ASM also uses the Oracle Managed Files (OMF) feature to simplify database file management. OMF automatically creates files in designated locations. OMF also names files and removes them while relinquishing space when tablespaces or files are deleted.

ASM reduces the administrative overhead for managing database storage by consolidating data storage into a small number of disk groups. This enables you to consolidate the storage for multiple databases and to provide for improved I/O performance.

ASM files can coexist with other storage management options such as raw disks and third-party file systems. This capability simplifies the integration of ASM into pre-existing environments.

Oracle Enterprise Manager includes a wizard that enables you to migrate non-ASM database files to ASM. ASM also has easy to use management interfaces such as SQL*Plus, the ASMCMD command-line interface, and Oracle Enterprise Manager.

ASM provides striping and mirroring.

ASM is Logical Volume Manager; it's not just a set of File System.

ASM let you plug or unplug (add or remove) disk while Oracle Database is running by using simple SQL statement. I think it's a strong point for ASM.

Moreover, ASM load balances the I/O across disks so as to improve performance.

Explanation-2:

In Oracle Database 10g/11g there are two types of instances: database and ASM instances. The ASM instance, which is generally named +ASM, is started with the INSTANCE_TYPE=ASM init.ora parameter. This parameter, when set, signals the Oracle initialization routine to start an ASM instance and not a standard database instance. Unlike the standard database instance, the ASM instance contains no physical files; such as logfiles, controlfiles or datafiles, and only requires a few init.ora parameters for startup.

Upon startup, an ASM instance will spawn all the basic background processes, plus some new ones that are specific to the operation of ASM. The STARTUP clauses for ASM instances are similar to those for database instances. For example, RESTRICT prevents database instances from connecting to this ASM instance. NOMOUNT starts up an ASM instance without mounting any disk group. MOUNT option simply mounts all defined diskgroups

For RAC configurations, the ASM SID is +ASMx instance, where x represents the instance number.

Benefits-1:

Provides automatic load balancing over all the available disks, thus reducing hot spots in the file system

Prevents fragmentation of disks, so you don't need to manually relocate data to tune I/O performance

Adding disks is straight forward - ASM automatically performs online disk reorganization when you add or remove storage

Uses redundancy features available in intelligent storage arrays

The storage system can store all types of database files

Using disk group makes configuration easier, as files are placed into disk groups
ASM provides striping and mirroring (fine and coarse grain - see below)
ASM and non-ASM oracle files can coexist
ASM is free!!!!!!!!!!!!!!

Benefits-2:

ASM provides filesystem and volume manager capabilities built into the Oracle database kernel. With this capability, ASM simplifies storage management tasks, such as creating/laying out databases and disk space management. Since ASM allows disk management to be done using familiar create/alter/drop SQL statements, DBAs do not need to learn a new skill set or make crucial decisions on provisioning.

The following are some key benefits of ASM:

ASM spreads I/O evenly across all available disk drives to prevent hot spots and maximize performance.
ASM eliminates the need for over provisioning and maximizes storage resource utilization facilitating database consolidation.

Inherent large file support.

Performs automatic online redistribution after the incremental addition or removal of storage capacity.

Maintains redundant copies of data to provide high availability, or leverages 3rd party RAID functionality.

Supports Oracle Database as well as Oracle Real Application Clusters (RAC).

Capable of leveraging 3rd party multipath technologies.

For simplicity and easier migration to ASM, an Oracle database can contain ASM and non-ASM files.

Any new files can be created as ASM files whilst existing files can also be migrated to ASM.

RMAN commands enable non-ASM managed files to be relocated to an ASM disk group.

Enterprise Manager Database Control or Grid Control can be used to manage ASM disk and file activities.

Benefits-3:

Stripes files rather than logical volumes

Provides redundancy on a file basis

Enables online disk reconfiguration and dynamic rebalancing

Reduces the time significantly to resynchronize a transient failure by tracking changes while disk is offline

Provides adjustable rebalancing speed

Is cluster-aware

Supports reading from mirrored copy instead of primary copy for extended clusters

Is automatically installed as part of the Grid Infrastructure

2. Describe about ASM architecture?

Automatic Storage Management (ASM) instance

Instance that manages the diskgroup metadata

Disk Groups

Logical grouping of disks

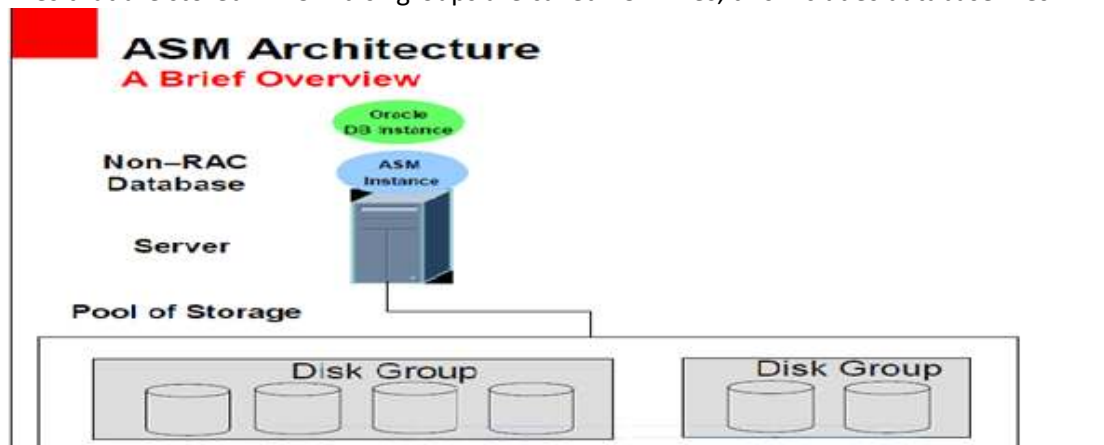
Determines file mirroring options

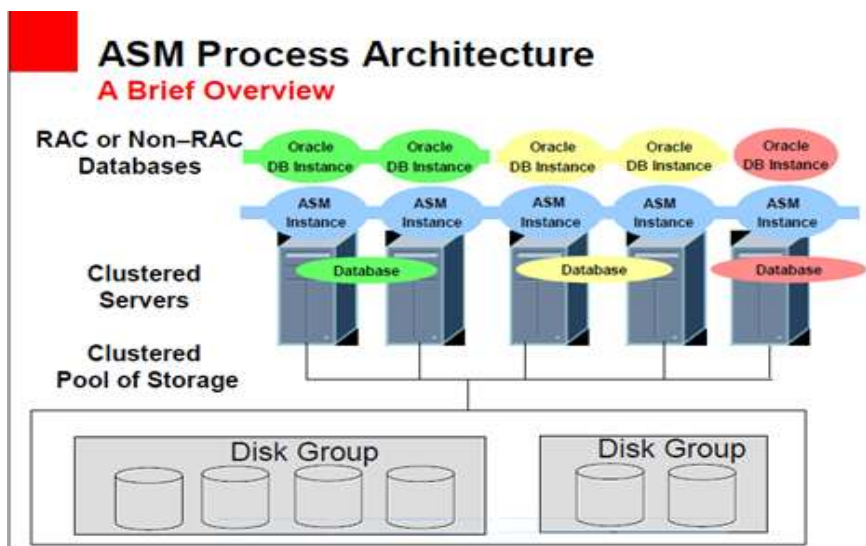
ASM Disks

LUNs presented to ASM

ASM Files

Files that are stored in ASM disk groups are called ASM files, this includes database files





Notes:

Many databases can connect as clients to single ASM instances

ASM instance name should only be +ASM only

One diskgroup can serve many databases

3. How does database connects to ASM Instance?

The database communicates with ASM instance using the ASMB (umbilicus process) process. Once the database obtains the necessary extents from extent map, all databases IO going forward is processed through by the database processes, bypassing ASM. Thus we say ASM is not really in the IO path. So, the question how do we make ASM go faster.....you don't have to.

4. What are the init parameters related to ASM?

The default parameter settings work perfectly for ASM. The only parameters needed for 11g ASM:

- PROCESSES*
- ASM_DISKSTRING*
- ASM_DISKGROUPS
- INSTANCE_TYPE

For Example:

INSTANCE_TYPE = ASM

ASM_POWER_LIMIT = 11

ASM_DISKSTRING = '/dev/rdisk/*s2', '/dev/rdisk/c1*'

ASM_DISKGROUPS = DG_DATA, DG_FRA

- ASM is a very passive instance in that it doesn't have a lot concurrent transactions or queries. So the memory footprint is quite small.
- Even if you have 20 dbs connected to ASM, the ASM SGA does not need to change. This is because the ASM metadata is not directly tied to the number of clients.
- The 11g MEMORY_TARGET (DEFAULT VALUE) will be more than sufficient.

The processes parameter may need to be modified. Use the formula to determine the approp value:

Processes = 40 + (10 + [max number of concurrent database file creations, and file extend operations possible])*n

Where n is the number of databases connecting to ASM (ASM clients).

The source of concurrent file creations can be any of the following:

- Several concurrent create tablespace commands.
- Creation of a Partitioned table with several tablespaces creations.
- RMAN backup channels.
- Concurrent archive logfile creations.

5. What is rebalancing (or) what is the use of ASM_POWER_LIMIT?

(<http://asm-support-guy.blogspot.in/2011/11/rebalancing-act.html>)

Explanation-1: Rebalancing a disk group moves data between disks to ensure that every file is equally spread across all of the disks in a disk group. ASM automatically initiates a rebalance after storage configuration changes, such as

when you add, drop, or resize disks. The power setting parameter determines the speed with which rebalancing operations occur.

ASM_POWER_LIMIT is dynamic parameter, which will be useful for rebalancing the data across disks.

Value can be 1(lowest) to 11 (highest).

Rebalancing and tuning:

ASM ensures that file [extents](#) are evenly distributed across all disks in a disk group. This is true for the initial file creation and for file resizes operations. That means we should always have a *balanced* space distribution across all disks in a disk group.

Rebalance operation:

Disk group rebalance is triggered automatically on ADD, DROP and RESIZE disk operations and on moving a file between hot and cold regions. Running rebalance by explicitly issuing ALTER DISKGROUP ... REBALANCE is called a manual rebalance. We might want to do that to change the rebalance power for example. We can also run the rebalance manually if a disk group becomes unbalanced for any reason.

The POWER clause of the ALTER DISKGROUP ... REBALANCE statement specifies the degree of parallelism of the rebalance operation. It can be set to a minimum value of 0 which halts the current rebalance until the statement is either implicitly or explicitly re-run. Higher values may reduce the total time it takes to complete the rebalance operation.

The ALTER DISKGROUP ... REBALANCE command by default returns immediately so that we can run other commands while the rebalance operation takes place in the background.

To check the progress of the rebalance operations we can query V\$ASM_OPERATION view.

Three phase power:

The rebalance operation has three distinct phases.

The First, ASM has to come up with the rebalance plan. That will depend on the rebalance reason, disk group size, number of files in the disk group, whether or not partnership has to be modified, etc. In any case this shouldn't take more than a couple of minutes.

The second phase is the moving or relocating the extents among the disks in the disk group. This is where the bulk of the time will be spent. As this phase is progressing, ASM will keep track of the number of extents moved, and the actual I/O performance. Based on that it will be calculating the estimated time to completion (GV\$ASM_OPERATION.EST_MINUTES). Keep in mind that this is an estimate and that the actual time may change depending on the overall (mostly disk related) load. If the reason for the rebalance was a failed disk(s) in a redundant disk group, at the end of this phase the data mirroring is fully re-established.

The third phase is disk(s) compacting (ASM version 11.1.0.7 and later). The idea of the compacting phases is to move the data as close to the outer tracks of the disks as possible. Note that at this stage of the rebalance, the EST_MINUTES will keep showing 0. This is a 'feature' that will hopefully be addressed in the future. The time to complete this phase will again depend on the number of disks, reason for rebalance, etc. Overall time should be a fraction of the second phase.

Notes about rebalance operations:

- Rebalance is *per file* operation.
- An ongoing rebalance is restarted if the storage configuration changes either when we alter the configuration, or if the configuration changes due to a failure or an outage. If the new rebalance fails because of a user error a manual rebalance may be required.
- There can be one rebalance operation per disk group per ASM instance in a cluster.
- Rebalancing continues across a failure of the ASM instance performing the rebalance.
- The REBALANCE clause (with its associated POWER and WAIT/NOWAIT keywords) can also be used in ALTER DISKGROUP commands for ADD, DROP or RESIZE disks.

Tuning rebalance operations:

If the POWER clause is not specified in an ALTER DISKGROUP statement, or when rebalance is implicitly run by ADD/DROP/RESIZE disk, then the rebalance power defaults to the value of the ASM_POWER_LIMIT initialization parameter. We can adjust the value of this parameter dynamically. Higher power limit should result in a shorter time to complete the rebalance, but this is by no means linear and it will depend on the (storage system) load, available throughput and underlying disk response times.

The power can be changed for a rebalance that is in progress. We just need to issue another ALTER DISKGROUP ... REBALANCE command with different value for POWER. This interrupts the current rebalance and restarts it with modified POWER.

Relevant initialization parameters and disk group attributes:

ASM_POWER_LIMIT

The ASM_POWER_LIMIT initialization parameter specifies the default power for disk rebalancing in a disk group. The range of values is 0 to 11 in versions prior to 11.2.0.2. Since version 11.2.0.2 the range of values is 0 to 1024, but that still depends on the disk group compatibility (see the notes below). The default value is 1. A value of 0 disables rebalancing.

For disk groups with COMPATIBLE, ASM set to 11.2.0.2 or greater, the operational range of values is 0 to 1024 for the rebalance power.

For disk groups that have COMPATIBLE.ASM set to less than 11.2.0.2, the operational range of values is 0 to 11 inclusive.

Specifying 0 for the POWER in the ALTER DISKGROUP REBALANCE command will stop the current rebalance operation (unless you hit bug 7257618).

_DISABLE_REBALANCE_COMPACT

Setting initialization parameter _DISABLE_REBALANCE_COMPACT=TRUE will disable the compacting phase of the disk group rebalance - for all disk groups.

_REBALANCE_COMPACT

this is a hidden disk group attribute. Setting _REBALANCE_COMPACT=FALSE will disable the compacting phase of the disk group rebalance - for that disk group only.

_ASM_IMBALANCE_TOLERANCE

this initialization parameter controls the percentage of imbalance between disks. Default value is 3%.

Processes

The following table has a brief summary of the background processes involved in the rebalance operation.

Process	Description
ARBn	ASM Rebalance Process. Rebalances data extents within an ASM disk group. Possible processes are ARB0-ARB9 and ARBA.
RBAL	ASM Rebalance Master Process. Coordinates rebalance activity. In an ASM instance, it coordinates rebalance activity for disk groups. In a database instances, it manages ASM disk groups.
Xnnn	Exadata only - ASM Disk Expel Slave Process. Performs ASM post-rebalance activities. This process expels dropped disks at the end of an ASM rebalance.

When a rebalance operation is in progress, the ARBn processes will generate trace files in the background dump destination directory, showing the rebalance progress.

Views

In an ASM instance, V\$ASM_OPERATION displays one row for every active long running ASM operation executing in the current ASM instance. GV\$ASM_OPERATION will show cluster wide operations.

During the rebalance, the OPERATION will show REBAL, STATE will shows the state of the rebalance operation, POWER will show the rebalance power and EST_MINUTES will show an estimated time the operation should take.

In an ASM instance, V\$ASM_DISK displays information about ASM disks. During the rebalance, the STATE will show the current state of the disks involved in the rebalance operation.

Is your disk group balanced:

run the following query in your ASM instance to get the report on the disk group imbalance.

```
SQL> column "Diskgroup" format A30
```

```
SQL> column "Imbalance" format 99.9 Heading "Percent|Imbalance"
```

```
SQL> column "Variance" format 99.9 Heading "Percent|Disk Size|Variance"
```

```
SQL> column "MinFree" format 99.9 Heading "Minimum|Percent|Free"
```

```
SQL> column "DiskCnt" format 9999 Heading "Disk|Count"
```

```
SQL> column "Type" format A10 Heading "Diskgroup|Redundancy"
```

```
SQL> SELECT g.name "Diskgroup",
```

```
100*((max((d.total_mb-d.free_mb)/d.total_mb)-min((d.total_mb-d.free_mb)/d.total_mb))/max((d.total_mb-d.free_mb)/d.total_mb)) "Imbalance",
```

```

100*(max(d.total_mb)-min(d.total_mb))/max(d.total_mb) "Variance",
100*(min(d.free_mb/d.total_mb)) "MinFree",
count(*) "DiskCnt",
g.type "Type"
FROM v$asm_disk d, v$asm_diskgroup g
WHERE d.group_number = g.group_number and
d.group_number <> 0 and
d.state = 'NORMAL' and
d.mount_status = 'CACHED'
GROUP BY g.name, g.type;
Diskgroup          Imbalance  Variance  Free Count Redundancy
-----

```

```

ACFS                .0      .0  12.5  2 NORMAL
DATA                .0      .0  48.4  2 EXTERN
PLAY               3.3      .0  98.1  3 NORMAL
RECO                .0      .0  82.9  2 EXTERN

```

Explanation-2:

Dynamic Storage Configuration:

ASM

enables you to change the storage configuration without having to take the database offline. It automatically rebalances—redistributes file data evenly across all the disks of the disk group—after you add disks to or drop disks from a disk group.

Should a disk failure occur, ASM automatically rebalances to restore full redundancy for files that had extents on the failed disk. When you replace the failed disk with a new disk, ASM rebalances the disk group to spread data evenly across all disks, including the replacement disk.

Tuning Rebalance Operations:

The V\$ASM_OPERATION view provides information that can be used for adjusting ASM_POWER_LIMIT and the resulting power of rebalance operations. The V\$ASM_OPERATION view also gives an estimate in the EST_MINUTES column of the amount of time remaining for the rebalance operation to complete. You can see the effect of changing the rebalance power by observing the change in the time estimate.

Effects of Adding and Dropping Disks from a Disk Group:

ASM automatically rebalances whenever disks are added or dropped. For a normal drop operation (without the FORCE option), a disk is not released from a disk group until data is moved off of the disk through rebalancing. Likewise, a newly added disk cannot support its share of the I/O workload until rebalancing completes. It is more efficient to add or drop multiple disks at the same time so that they are rebalanced as a single operation. This avoids unnecessary movement of data.

For a drop operation, when rebalance is complete, ASM takes the disk offline momentarily, and then drops it, setting disk header status to FORMER.

You can add or drop disks without shutting down the database. However, a performance impact on I/O activity may result.

Explanation-3:

ASM Rebalance:

The Rebalance operation provides an even distribution of file extents across all disks in the diskgroup. The rebalance is done on each file to ensure balanced I/O load.

The RBAL background process manages the rebalance activity. It examines the extent map for each file and redistributes the extents to new storage configuration. The RBAL process will calculate estimation time and the work required to perform the rebalance activity and then message the ARBx processes to actually perform the task. The number of ARBx process starts is determined by the parameter ASM_POWER_LIMIT.

There will be one I/O for each ARBx process at a time. Hence the impact of physical movement of file extents will be low. The asm_power_limit parameter determines the speed of the rebalance activity. It can have values between 0 and 11. If the value is 0 no rebalance occurs. If the value is 11 the rebalance takes place at full speed. The power value can also be set for specific rebalance activity using Alter Diskgroup statement.

The rebalance operation has various states, they are

WAIT: No operations are running for the group.

RUN: A rebalance operation is running for the group.

HALT: The [DBA](#) has halted the operation.

ERROR: The operation has halted due to errors.

You can query the V\$ASM_OPERATION to view the status of rebalance activity.

The rebalance activity is an asynchronous operation, i.e., the operation runs in the background while the users can perform other tasks. In certain situation you need the rebalance activity to finish successfully before performing the other tasks. To make the operation synchronous you add a keyword WAIT while performing the rebalance as shown below.

```
SQL> Alter diskgroup ASMDB Add Disk '/dev/sdc4' Rebalance power 4 WAIT;
```

The above statement will not return the control to the user unless the rebalance operation ends.

Explanation-4:

Manually Rebalancing Disk Groups

You can manually rebalance the files in a disk group using the REBALANCE clause of the [ALTER DISKGROUP](#) statement. This would normally not be required, because ASM automatically rebalances disk groups when their configuration changes. You might want to do a manual rebalance operation if you want to control the speed of what would otherwise be an automatic rebalance operation.

The POWER clause of the ALTER DISKGROUP...REBALANCE statement specifies the degree of parallelism, and thus the speed of the rebalance operation. It can be set to a value from 0 to 11. A value of 0 halts a rebalancing operation until the statement is either implicitly or explicitly re-run. The default rebalance power is set by the ASM_POWER_LIMIT initialization parameter. See ["Tuning Rebalance Operations"](#) for more information.

The power level of an ongoing rebalance operation can be changed by entering the rebalance statement with a new level.

The ALTER DISKGROUP...REBALANCE command by default returns immediately so that you can issue other commands while the rebalance operation takes place asynchronously in the background. You can query the V\$ASM_OPERATION view for the status of the rebalance operation.

If you want the ALTER DISKGROUP...REBALANCE command to wait until the rebalance operation is complete before returning, you can add the WAIT keyword to the REBALANCE clause. This is especially useful in scripts. The command also accepts a NOWAIT keyword, which invokes the default behavior of conducting the rebalance operation asynchronously. You can interrupt a rebalance running in wait mode by typing CTRL-C on most platforms. This causes the command to return immediately with the message [ORA-01013](#): user requested cancel of current operation, and then to continue the rebalance operation asynchronously.

Additional rules for the rebalance operation include the following:

- An ongoing rebalance command will be restarted if the storage configuration changes either when you alter the configuration, or if the configuration changes due to a failure or an outage. Furthermore, if the new rebalance fails because of a user error, then a manual rebalance may be required.
- The ALTER DISKGROUP...REBALANCE statement runs on a single node even if you are using Oracle Real Application Clusters (Oracle RAC).
- ASM can perform one disk group rebalance at a time on a given instance. Therefore, if you have initiated multiple rebalances on different disk groups, then Oracle processes this operation serially. However, you can initiate rebalances on different disk groups on different nodes in parallel.
- Rebalancing continues across a failure of the ASM instance performing the rebalance.
- The REBALANCE clause (with its associated POWER and WAIT/NOWAIT keywords) can also be used in ALTER DISKGROUP commands that add, drop, or resize disks.

Note:

Oracle will restart the processing of an ongoing rebalance operation if the storage configuration changes.

Furthermore, if the next rebalance operation fails because of a user error, then you may need to perform a manual rebalance.

Example: Manually Rebalancing a Disk Group

The following example manually rebalances the disk group dgroup2. The command does not return until the rebalance operation is complete.

```
ALTER DISKGROUP dgroup2 REBALANCE POWER 5 WAIT;
```

Tuning Rebalance Operations

If the POWER clause is not specified in an ALTER DISKGROUP statement, or when rebalance is implicitly run by adding or dropping a disk, then the rebalance power defaults to the value of the ASM_POWER_LIMIT initialization parameter. You can adjust the value of this parameter dynamically.

The higher the power limit, the more quickly a rebalance operation can complete. Rebalancing takes longer with lower power values, but consumes fewer processing and I/O resources which are shared by other applications, such as the database.

The default value of 1 minimizes disruption to other applications. The appropriate value is dependent on your hardware configuration, performance requirements, and availability requirements

If a rebalance is in progress because a disk is manually or automatically dropped, then increasing the power of the rebalance shortens the time frame during which redundant copies of that data on the dropped disk are reconstructed on other disks.

The V\$ASM_OPERATION view provides information for adjusting ASM_POWER_LIMIT and the resulting power of rebalance operations. The V\$ASM_OPERATION view also gives an estimate in the EST_MINUTES column of the amount of time remaining for the rebalance operation to complete. You can see the effect of changing the rebalance power by observing the change in the time estimate.

6. What is significance of re-balance power?

7. In what situations we need to re-balance the disk?

8. In what situation asm instance will automatically re-balance the disk?

9. Explain about disk group managements?

Refer the Meta link for Create, Delete, and Add

10. What are different types of redundancies in ASM & explain?

<http://blog.trivadis.com/b/mathiaszarick/archive/2009/05/04/asm-disk-groups-redundancy-at-diskgroup-level-vs-redundancy-at-template-level-are-there-differences.aspx>

External redundancy,

Normal redundancy,

High redundancy.

The availability of normal redundancy configuration option for Automatic Storage Management (ASM) on Oracle Database Appliance starting with OAK version 2.4 allows for additional usable space on Oracle Database Appliance (about 6 TB with Normal Redundancy versus about 4 TB with High Redundancy). This is great news for many customers. Some environments, such as test and development systems, may benefit significantly as a result of this new option. However, the availability of Normal Redundancy option obviously should not be taken to mean that choosing Normal Redundancy may be the best approach for all database environments. High redundancy would still provide a better and more resilient option (and may be a preferred choice) for mission critical production systems. It is therefore an option and not the default configuration choice. Many customers may choose to use Normal Redundancy for test, development, and other non-critical environments and High Redundancy for production and other important systems.

In general, ASM supports three types of redundancy (mirroring*) options.

High Redundancy - In this configuration, for each primary extent, there are two mirrored extents. For Oracle Database Appliance this means, during normal operations there would be three extents (one primary and two secondary) containing the same data, thus providing “high” level of protection. Since ASM distributes the partnering extents in a way that prevents all extents to be unable due to a component failure in the IO path, this configuration can sustain at least two simultaneous disk failures on Oracle Database Appliance (which should be rare but is possible).

Normal Redundancy - In this configuration, for each primary extent, there is one mirrored (secondary) extent. This configuration protects against at least one disk failure. Note that in the event a disk fails in this configuration, although there is typically no outage or data loss, the system operates in a vulnerable state, should a second disk fail while the old failed disk replacement has not completed. Many Oracle Database Appliance customers thus prefer the High Redundancy configuration to mitigate the lack of additional protection during this time.

External Redundancy - In this configuration there are only primary extents and no mirrored extents. This option is typically used in traditional non-appliance environments when the storage sub-system may have existing redundancy such as hardware mirroring or other types of third-party mirroring in place. Oracle Database Appliance does not support External Redundancy.

*ASM redundancy is different from traditional disk mirroring in that ASM mirroring is a logical-physical approach than a pure physical approach. ASM does not mirror entire disks. It mirrors logical storage entities called ‘extents’ that are allocated on physical disks. Thus, all “mirrored” extents of a set of primary extents on a given disk do not need to be on a single mirrored disk but they could be distributed across multiple disks. This approach to mirroring provides significant benefits and flexibility. ASM uses intelligent, Oracle Database Appliance architecture aware, extent placement algorithms to maximize system availability in the event of disk failure(s).

11. How to copy file to/from ASM from/to file system?

By using [ASMCMD](#) CP command

You can use RMAN or DBMS_FILE_TRANSFER.COPY_FILE procedure to copy the files to/from ASM from/to Filesystem. Starting from Oracle 11g, you can use cp command in asmcmd to perform the same between ASM Diskgroups and also to OS Filesystem.

12. How to find out the databases, which are using the ASM instance?

[ASMCMD](#)> lsct

```
SQL> select DB_NAME from V$ASM_CLIENT;
```

13. What is Striping and Mirroring? What are different types of striping and Mirroring in ASM & their differences?

ASM provides striping by dividing files into equal-sized extents. Fine-grained striping extents are 128KB in size. For Oracle 10g, coarse-grained striping extents are 1MB in size. For Oracle 11g, coarse-grained striping extents can be 1, 2, 4, 8, 16, 32, or 64MB in size. Striping spreads each file extent evenly across all disks in the assigned disk group. ASM also provides automatic mirroring of ASM files and allows the mirroring level to be specified by group. This mirroring occurs at the extent level. If a disk group is mirrored, each extent has one or more mirrored copies, and mirrored copies are always kept on different disks in the disk group.

There are three ASM mirroring options:

- Two-way mirroring – Each extent has one mirrored copy in this option.
- Three-way mirroring – Each extent has two mirrored copies in this option.
- Unprotected mirroring – ASM provides no mirroring in this option, which is used when mirroring is provided by the disk subsystem.

ASM Striping

Explanation-1:

ASM can use variable size data extents to support larger files, reduce memory requirements, and improve performance.

Each data extent resides on an individual disk.

Data extents consist of one or more allocation units.

The data extent size is:

- Equal to AU for the first 20,000 extents (0–19999)
- Equal to 4 × AU for the next 20,000 extents (20000–39999)
- Equal to 16 × AU for extents above 40,000

ASM stripes files using extents with a coarse method for load balancing or a fine method to reduce latency.

- Coarse-grained striping is always equal to the effective AU size.
- Fine-grained striping is always equal to 128 KB.

Explanation-2:

ASM stripes files across all the disks within the disk group thus increasing performance, each stripe is called an 'allocation unit'. ASM offers two types of striping which is dependent on the type of database file.

Striping is a technique where [data](#) is stored on multiple disk drives by splitting up the data and accessing all of the disk drives in parallel. Striping significantly speeds up disk drive performance.

Example: [RAID](#) - RAID 0 is data striping

ASM stripes its files across all the disks that belong to a disk group. It remains unclear if it follows a strict [RAID 3](#) fashion of striping or a variant of RAID 3 that facilitates easy addition and removal of disks to and from the disk group. Oracle Corporation recommends that all the disks that belong to a disk group have the same size, in which case each disk gets the same number of extents. However, if a DBA configures disks of different sizes, each disk might get a different number of extents — based upon the size of the disk. An allocation unit typically has a size of 1MB. ASM stripes help make data more reliably available and more secure than in other Oracle storage implementations.

Types of Striping:

Coarse Striping	used for datafile, archive logs (1MB stripes)
Fine Striping	used for online redo logs, controlfile, flashback files(128KB stripes)

ASM Mirroring

Disk mirroring provides data redundancy, this means that if a disk were to fail Oracle will use the other mirrored disk and would continue as normal. Oracle mirrors at the extent level, so you have a primary extent and a mirrored

extent. When a disk fails, ASM rebuilds the failed disk using mirrored extents from the other disks within the group, this may have a slight impact on performance as the rebuild takes place.

All disks that share a common controller are in what is called a *failure group*, you can ensure redundancy by mirroring disks on separate failure groups which in turn are on different controllers, ASM will ensure that the primary extent and the mirrored extent are not in the same failure group. When mirroring you must define failure groups otherwise the mirroring will not take place.

Types of Mirroring:

- **External redundancy** - doesn't have failure groups and thus is effectively a no-mirroring strategy
- **Normal redundancy** - provides two-way mirroring of all extents in a disk group, which result in two failure groups
- **High redundancy** - provides three-way mirroring of all extents in a disk group, which result in three failure groups

Note: After creating a diskgroup you cannot change the redundancy level. If you want to change it then create a separate diskgroup and move the files to that diskgroup (using RMAN restore or DBMS_FILE_TRANSFER).

14. What are Diskgroup's and Failuregroups?

A disk group consists of multiple disks and is the fundamental object that ASM manages. Each disk group contains the metadata that is required for the management of space in the disk group. The ASM instance manages the metadata about the files in a Disk Group in the same way that a file system manages metadata about its files. However, the vast majority of I/O operations do not pass through the ASM instance. In a moment we will look at how file I/O works with respect to the ASM instance.

Diskgroup is a terminology used for logical structure which holds the database files. Each Diskgroup consists of Disks/Raw devices where the files are actually stored. Any ASM file is completely contained within a single disk group. However, a disk group might contain files belonging to several databases and a single database can use files from multiple disk groups.

The primary component of ASM is the **disk group**. A disk group consists of a grouping of disks that are managed together as a unit. You configure ASM by creating disk groups to store database files. Oracle provides SQL statements that create and manage disk groups, their contents, and their metadata.

The disk group type determines the levels of mirroring that files in the disk group can be created with. You specify disk group type when you create the disk group.

If you do not specify a disk group type (redundancy level) when you create a disk group, the disk group defaults to normal redundancy.

The files in a high redundancy disk group are always 3-way mirrored, files in an external redundancy disk group have no ASM mirroring, and files in a normal redundancy disk group can be 2-way or 3-way mirrored or unprotected, with 2-way mirroring as the default. Mirroring level for each file is set by templates, which are described later in this section.

Disks

The disks in a disk group are referred to as ASM disks. On Windows operating systems, an ASM disk is always a partition. On all other platforms, an ASM disk can be:

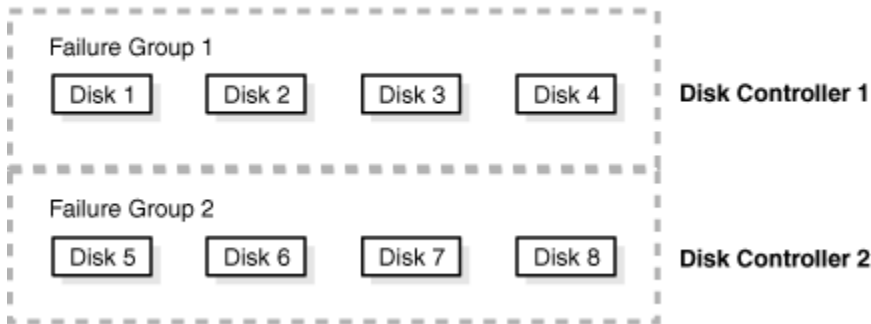
- A partition of a logical unit number (LUN)
- A network-attached file
- When an ASM instance starts, it automatically discovers all available ASM disks. Discovery is the process of determining every disk device to which the ASM instance has been given I/O permissions (by some operating system mechanism), and of examining the contents of the first block of such disks to see if they are recognized as belonging to a disk group. ASM discovers disks in the paths that are listed in an initialization parameter, or if the parameter is NULL, in an operating system-dependent default path.

Failure Groups

Failuregroups are used when using Normal/High Redundancy. They contain the mirrored ASM extents and must be containing different disks and preferably on separate disk controller.

Failure groups define ASM disks that share a common potential failure mechanism. An example of a failure group is a set of SCSI disks sharing the same SCSI controller. Failure groups are used to determine which ASM disks to use for storing redundant copies of data. For example, if two-way mirroring is specified for a file, ASM automatically stores redundant copies of file extents in separate failure groups. Failure groups apply only to normal and high redundancy disk groups. You define the failure groups in a disk group when you create or alter the disk group.

Disk Group 1



Files

Files written on ASM disks are **ASM files**, whose names are automatically generated by ASM. You can specify user-friendly **alias names** (or just **aliases**) for ASM files, and you can create a hierarchical directory structure for these aliases. Each ASM file is completely contained within a single disk group, and is evenly spaced over all of the ASM disks in the disk group.

Templates

Templates are collections of file attribute values, and are used to set mirroring and striping attributes of each type of database file (datafile, control file, redo log file, and so on) created in ASM disk groups. Each disk group has a default template associated with each file type. See ["Managing Disk Group Templates"](#) for more information.

You can also create your own templates to meet unique requirements. You can then include a template name when creating a file, thereby assigning desired attributes on an individual file basis rather than on the basis of file type. See ["About ASM Filenames"](#) for more information.

15. Can ASM be used as replacement for RAID?

ASM is supposed to stripe the data and also mirror the data (if Using Normal, High Redundancy). So this can be used as an alternative for RAID 0+1 solutions

16. What are the background processes in ASM?

RABL- Rebalancer: It opens all the device files as part of disk discovery and coordinates the ARB processes for rebalance activity.

ARBx - Actual Rebalancer: They perform the actual rebalancing activities. The number of ARBx processes depends on the ASM_POWER_LIMIT init parameter.

ASMB - ASM Bridge: This process is used to provide information to and from the Cluster Synchronization Service (CSS) used by ASM to manage the disk resources. It is also used to update statistics and provide a heartbeat mechanism.

Process Description

RBAL	Opens all device files as part of discovery and coordinates the rebalance activity
ARBn	One or more slave processes that do the rebalance activity
GMON	Responsible for managing the disk-level activities such as drop or offline and advancing the ASM disk group compatibility
MARK	Marks ASM allocation units as stale when needed
Onnn	One or more ASM slave processes forming a pool of connections to the ASM instance for exchanging messages
PZ9n	One or more parallel slave processes used in fetching data on clustered ASM installation from GV\$ views

17. What are the file types that ASM support and keep in disk groups?

Control files
Flashback logs
Data Pump dump sets

Data files
DB SPFILE
Data Guard configuration
Temporary data files
RMAN backup sets
Change tracking bitmaps
Online redo logs
RMAN data file copies
OCR files
Archive logs
Transport data files
ASM SPFILE

Note: Oracle executable and ASCII files, such as alert logs and trace files, cannot be stored in ASM disk groups.

18. How many ASM Diskgroups can be created under one ASM Instance?

ASM imposes the following limits:

63 disk groups in a storage system

10,000 ASM disks in a storage system

Two-terabyte maximum storage for each ASM disk (non-Exadata)

Four-petabyte maximum storage for each ASM disk (Exadata)

40-exabyte maximum storage for each storage system

1 million files for each disk group

ASM file size limits (database limit is 128 TB):

External redundancy maximum file size is 140 PB.

Normal redundancy maximum file size is 42 PB.

High redundancy maximum file size is 15 PB.

19. What process does the rebalancing?

20. How does ASM provides Redundancy?

When you create a disk group, you specify an ASM disk group type based on one of the following three redundancy levels:

- * Normal for 2-way mirroring – When ASM allocates an extent for a normal redundancy file; ASM allocates a primary copy and a secondary copy. ASM chooses the disk on which to store the secondary copy in a different failure group other than the primary copy.

- * High for 3-way mirroring. In this case the extent is mirrored across 3 disks.

- * External to not use ASM mirroring. This is used if you are using Third party Redundancy mechanism like RAID, Storage arrays.

21. Can we change the Redundancy for Diskgroup after its creation?

No, we cannot modify the redundancy for Diskgroup once it has been created. To alter it we will be required to create a new Diskgroup and move the files to it. This can also be done by restoring full backup on the new Diskgroup. Following metalink note describes the steps

Note.438580.1 – How To Move The Database To Different Diskgroup

22. Unable to open the ASM instance. What is the reason?

ASM instance does not have open stage. It has got only two options

- * Nomount- This starts the ASM instance

- * Mount- At this stage, Diskgroup defined in ASM_DISKGROUPS parameter are mounted

When you try to open the ASM instance, you get following error

```
SQL> alter database open;
```

```
alter database open
```

```
*
```

```
ERROR at line 1:
```

```
ORA-15000: command disallowed by current instance type
```

23. Can ASM instance and database (rdbms) be on different servers?

ASM instance and Database (rdbms) have to be present on same server. Otherwise it will not work.

24. Can we see the files stored in the ASM instance using standard UNIX commands?

No, you cannot see the files using standard unix commands like ls. You need to use utility called asmcmd to do this.

This is present in 10.2 and above.

25. Can we use ASM for storing Voting Disk/OCR in a RAC instance?

In oracle 11gR1 and below, you cannot use ASM for storing the voting disk and OCR. It is due to the fact that Clusterware starts before ASM instance and it should be able to access these files which are not possible if you are storing it on ASM. You will have to use raw devices or OCFS or any other shared storage.

In Oracle 11gR2 we can store them in ASM.

26. Does ASM instance automatically rebalances and takes care of hot spots?

No. This is a myth and ASM does not do it. It will initiate automatic rebalance only when a new disk is added to Diskgroup or we drop a disk from existing Diskgroup.

27. What is ASMLIB?

ASMLIB is the support library for the ASM. ASMLIB allows an Oracle database using ASM more efficient and capable access to diskgroups. The purpose of ASMLIB, is to provide an alternative interface to identify and access block devices. Additionally, the ASMLIB API enables storage and operating system vendors to supply extended storage-related features.

28. What is SYSASM role?

Starting from Oracle 11g, SYSASM role can be used to administer the ASM instances. You can continue using SYSDBA role to connect to ASM but it will generate following warning messages at time of startup/shutdown, create Diskgroup/add disk, etc

Alert entry

WARNING: Deprecated privilege SYSDBA for command 'STARTUP'

29. Can we use BCV to clone the ASM Diskgroup on same host?

Diskgroup Cloning is not supported on the same host using BCV. You have no other option to use except RMAN DUPLICATE. You can find more detail on BCV and ASM in below whitepaper

30. Can we edit the ASM Disk header to change the Diskgroup Name?

No. This cannot be done. Please find details in

31. What is Kfed?

kfed is a utility which can be used to view the ASM Disk information. Syntax for using it is

kfed read devicename

32. Can we use block devices for ASM Disks?

Yes. Starting from Oracle Database 10.2 block devices can be used directly for ASM Disks in Linux. This is not true for other Unix based systems where block devices are not supported yet.

Along with this it is recommended to use a Device mapping functionality so that disk mapping is preserved after disk failure. This is important when you have devices as /dev/sda,/dev/sdb,/dev/sdc and due to some reason the devices are not detected at next reboot (say /dev/sdb), the system will map the incorrect device (i.e /dev/sdc will be marked as /dev/sdb). You can use following methods for preserving disk names

-udev – the role of udev is to provide device persistency and naming consistency. This is especially important for the Oracle Cluster Registry (OCR) and Voting disks required by Oracle Clusterware.

-ASMLIB – ASMLIB will provide device management specifically for ASM disk devices.

33. Is it mandatory to use disks of same size and characteristics for Diskgroups?

No, it is not mandatory to use the disks of same size and characteristics for Diskgroups though it is a Recommended Practice.

Same size disk for Failuregroups in Normal/High redundancy will prevent issues like ORA-15041 as the file extents needs to be mirrored across the disks. Also as Oracle distributes data based on capacity, so larger disk will have more data stored in it and which will result in higher I/O to disk and eventually can lead to sub-optimal performance.

Moreover having disks of different characteristic like varying disk speed can impact the performance.

When managing disks with different size and performance capabilities, best practice is to group them into disk groups according to their characteristics. So you can use higher speed disks for your database files while other disks can be part of Diskgroup used for Flash Recovery Area.

34. Do we need to install ASM and Oracle Database Software in different ORACLE_HOME?

No. Again installing ASM and Oracle Database Software in different ORACLE_HOME is not mandatory but a best practice. This is useful in cases when we need to have multiple databases using same ASM instance and you need to patch only one of them. E.g You need to apply a CBO patch to one of 10.2 database while your other 10.1 database using different installation does not require it. In this case having a ASM_HOME separate from 10.2 ORACLE_HOME will allow your 10.1 database to keep running. Thus this approach is useful for High Availability.

35. What is the maximum size of Disk supported by ASM?

ASM supports disks up to 2Tb, so you need to ensure that lun size should be less than 2Tb. 10.2.0.4 and 11g database will give error if you try to create a diskgroup with ASM disks having disk size >2Tb.

36. I have created Oracle database using DBCA and having a different home for ASM and Oracle Database. I see that listener is running from ASM_HOME. Is it correct?

This is fine. When using different home for ASM, you need to run the listener from ASM_HOME instead of ORACLE_HOME.

37. How does the database interact with the ASM instance and how do I make ASM go faster?

ASM is not in the I/O path so ASM does not impede the database file access. Since the RDBMS instance is performing raw I/O, the I/O is as fast as possible.

38. Do I need to define the RDBMS FILESYSTEMIO_OPTIONS parameter when I use ASM?

No. The RDBMS does I/O directly to the raw disk devices, the FILESYSTEMIO_OPTIONS parameter is only for filesystems.

39. Why Oracle recommends two diskgroups?

Oracle recommends two diskgroups to provide a balance of manageability, utilization, and performance.

40. We have a 16 TB database. I'm curious about the number of disk groups we should use; e.g. 1 large disk group, a couple of disk groups, or otherwise?

For VLDBs you will probably end up with different storage tiers; e.g. with some of our large customers they have Tier1 (RAID10 FC), Tier2 (RAID5 FC), Tier3(SATA), etc. Each one of these is mapped to a diskgroup.

These custs mapped certain tablespaces to specific tiers; eg, system/rollback/syaux and latency sensitive tablespaces in Tier1, and not as IO critical on Tier2, etc.

For 10g VLDBs its best to set an AU size of 16MB, this is more for metadata space efficiency than for performance. The 16MB recommendation is only necessary if the diskgroup is going to be used by 10g databases. In 11g we introduced variable size extents to solve the metadata problem. This requires compatible.rdbms & compatible.asm to be set to 11.1.0.0. With 11g you should set your AU size to the largest I/O that you wish to issue for sequential access (other parameters need to be set to increase the I/O size issued by Oracle). For random small I/Os the AU size does not matter very much as long as every file is broken into many more extents than there are disks.

41. We have a new app and don't know our access pattern, but assuming mostly sequential access, what size would be a good AU fit?

For 11g ASM/RDBMS it is recommended to use 4MB ASM AU for disk groups. See Metalink Note 810484.1

42. Would it be better to use BIGFILE tablespaces, or standard tablespaces for ASM?

The use of Bigfile tablespaces has no bearing on ASM (or vice versa). In fact most database object related decisions are transparent to ASM.

Nevertheless, Bigfile tablespaces benefits:

Fewer datafiles - which means faster database open (fewer files to open), Faster checkpoints, as well fewer files to manage. But you'll need careful consideration for backup/recovery of these large datafiles.

43. What is the best LUN size for ASM?

There is no best size! In most cases the storage team will dictate to you based on their standardized LUN size. The ASM administrator merely has to communicate the ASM Best Practices and application characteristics to storage folks :

- Need equally sized / performance LUNs
- Minimum of 4 LUNs
- The capacity requirement
- The workload characteristic (random r/w, sequential r/w) & any response time SLA

Using this info , and their standards, the storage folks should build a nice LUN group set for you.

In most cases the storage team will dictate to you what the standardized LUN size is. This is based on several factors, including RAID LUN set builds (concatenated, striped, hypers, etc..). Having too

many LUNs elongates boot time and is it very hard to manage On the flip side, having too few LUNs makes array cache management difficult to control and creates un-manageable large LUNs (which are difficult to expand).

The ASM administrator merely has to communicate to SA/storage folks that you need equally sized/performance LUNs and what the capacity requirement is, say 10TB. Using this info, the workload characteristic (random r/w, sequential r/w), and their standards, the storage folks should build a nice LUN group set for you Having too many LUNs elongates boot time and is it very hard to manage (zoning, provisioning, masking, etc...)...there's a \$/LUN barometer!

44. In 11g RAC we want to separate ASM admins from DBAs and create different users and groups. How do we set this up?

A. For clarification

- Separate Oracle Home for ASM and RDBMS.
- RDBMS instance connects to ASM using OSDBA group of the ASM instance.

Thus, software owner for each RDBMS instance connecting to ASM must be a member of ASM's OSDBA group.

- Choose a different OSDBA group for ASM instance (asmdba) than for RDBMS instance (dba).
- In 11g, ASM administrator has to be member of a separate SYSASM group to separate ASM Admin and DBAs.

Operating system authentication using membership in the group or groups designated as OSDBA, OSOPER, and OSASM is valid on all Oracle platforms.

A typical deployment could be as follows:

ASM administrator:

User : asm

Group: oinstall, asmdba(OSDBA), asmadmin(OSASM)

Database administrator:

User : oracle

Group: oinstall, asmdba(OSDBA of ASM), dba(OSDBA)

ASM disk ownership : asm:oinstall

Remember that Database instance connects to ASM instance as sysdba. The user id the database instance runs as needs to be the OSDBA group of the ASM instance.

A typical deployment could be as follows:

ASM administrator:

User : asm

Group: oinstall, asmdba(OSDBA), asmadmin(OSASM)

Database administrator:

User : oracle

Group: oinstall, asmdba(OSDBA of ASM), dba(OSDBA)

A typical deployment could be as follows:

ASM administrator:

User : asm

Group: oinstall, asmdba(OSDBA), asmadmin(OSASM)

Database administrator:

User : oracle

Group: oinstall, asmdba(OSDBA of ASM), dba(OSDBA)

45. Can my RDBMS and ASM instances run different versions?

Yes. ASM can be at a higher version or at lower version than its client databases. There's two components of compatibility:

Software compatibility

Diskgroup compatibility attributes:

compatible.asm

compatible.rdbms

46. Where do I run my database listener from; i.e., ASM HOME or DB HOME?

It is recommended to run the listener from the ASM HOME. This is particularly important for RAC env, since the listener is a node-level resource. In this config, you can create additional [user] listeners from the database homes as needed.

47. How do I backup my ASM instance?

Not applicable! ASM has no files to backup, as it does not contain controlfile, redo logs etc.

48. When should I use RMAN and when should I use ASMCMD copy?

RMAN is the recommended and most complete and flexible method to backup and transport database files in ASM. ASMCMD copy is good for copying single files

- Supports all Oracle file types
- Can be used to instantiate a Data Guard environment
- Does not update the controlfile
- Does not create OMF files

49. I'm going to do add disks to my ASM diskgroup, how long will this rebalance take?

Rebalance time is heavily driven by the three items:

- 1) Amount of data currently in the diskgroup
- 2) IO bandwidth available on the server
- 3) ASM_POWER_LIMIT or Rebalance Power Level

50. We are migrating to a new storage array. How do I move my ASM database from storage A to storage B?

Given that the new and old storage are both visible to ASM, simply add the new disks to the ASM disk group and drop the old disks. ASM rebalance will migrate data online.

Note 428681.1 covers how to move OCR/Voting disks to the new storage array

51. Is it possible to unplug an ASM disk group from one platform and plug into a server on another platform (for example, from Solaris to Linux)?

No. Cross-platform disk group migration not supported. To move datafiles between endian-ness platforms, you need to use XTTS, Datapump or Streams.

52. How does ASM work with multipathing software?

It works great! Multipathing software is at a layer lower than ASM, and thus is transparent.

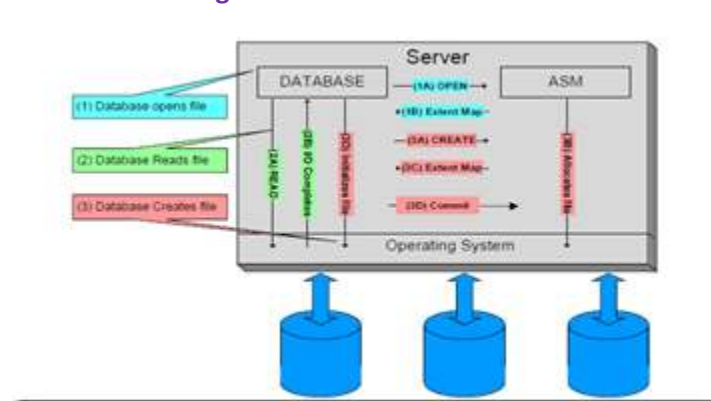
You may need to adjust ASM_DISKSTRING to specify only the path to the multipathing pseudo devices.

53. Is ASM constantly rebalancing to manage "hot spots"?

No...No...Nope!! ASM provides even distribution of extents across all disks in a disk group. Since each disk will equal number of extents, no single disk will be hotter than another. Thus the answer NO, ASM does not dynamically move hot spots, because hot spots simply do not

occur in ASM configurations. Rebalance only occurs on storage configuration changes (e.g. add, drop, or resize disks).

54. Draw the Diagram that how database interacts with ASM when a request is to read or open a datafile.



1A. Database issues open of a database file

1B. ASM sends the extent map for the file to database instance. Starting with 11g, the RDBMS only receives first 60 extents the remaining extents in the extent map are paged in on demand, providing a faster open

2A/2B. Database now reads directly from disk

3A.RDBMS foreground initiates a create tablespace for example

3B. ASM does the allocation for its essentially reserving the allocation units for the file creation

3C. Once allocation phase is done, the extent map is sent to the RDBMS

3D. The RDBMS initialization phase kicks in. In this phase the initializes all the reserved AUs

3E. If file creation is successful, then the RDBMS commits the file creation

Going forward all I/Os are done by the RDBMS directly.

55. Can my disks in a diskgroup can be varied size? For example one disk is of 100GB and another disk is of 50GB. If so how does ASM manage the extents?

Yes, disk sizes can be varied, Oracle ASM will manage data efficiently and intelligent by placing the extents proportional to the size of the disk in the disk group, bigger diskgroups have more extents than lesser ones.

56. What is Intelligent Data Placement?

57. What is ASM preferred Mirror read? How does it useful?

58. What is ACFS?

59. What is ADVM?

60. What is ASM Template?

Collections of attributes used by ASM during file creation are known as templates. Templates are used to simplify ASM file creation by mapping complex file attribute specifications into a single named object (template). Each Oracle file type has its own default template. Each disk group contains its own set of definition templates. Template names only have to be unique within a single ASM disk group, a template of the same name can exist in different disk groups with each separate template having their own unique properties.

Administrators can change the attributes of the default templates or add their own templates. This lets an administrator specify the appropriate file creation attributes as a template. However, if a DBA needs to change an ASM file attribute after a file has been created, then the file must be copied using RMAN into a new file created with a different template that contains the new attributes.

61. Why Oracle recommends two diskgroups. Why?

A. Oracle recommends two diskgroups to provide a balance of manageability, utilization, and performance. To reduce the complexity of managing ASM and its diskgroups, Oracle recommends that generally no more than two diskgroups be maintained and managed per RAC cluster or single ASM instance.

Database work area: This is where active database files such as datafiles, control files, online redo logs, and change tracking files used in incremental backups are stored. This location is indicated by DB_CREATE_FILE_DEST.

Flash recovery area: Where recovery-related files are created, such as multiplexed copies of the current control file and online redo logs, archived redo logs, backup sets, and flashback log files. This location is indicated by DB_RECOVERY_FILE_DEST.

Having one DATA container means only place to store all your database files, and obviates the need to juggle around datafiles or having to decide where to place a new tablespace. By having one container for all your files also means better storage utilization. Making the IT director very happy. If more storage capacity or IO capacity is needed, just add an ASM disk....all online activities. You have to ensure that this storage pool container houses enough spindles to accommodate the IO rate of all the database objects Bottom line, one container == one pool manage, monitor, and track Note however, that additional diskgroups may be added to support tiered storage classes in Information Lifecycle Management (ILM) or Hierarchical Storage Management (HSM) deployments

Oracle RAC FAQ

1. What is the use of RAC? What is RAC? How is it different from standalone database?
2. What are the prerequisites for RAC setup?
3. What is oracle clusterware? Benefits of Clusterware? What are the Cluster ware components?
4. What is OCR file?
5. What is voting file/disk and how many files should be there? Why do we have to create odd number of voting disk?
6. How to backup of OCR file?
7. How to recover OCR file?
8. What is local OCR?
9. How to check backup of OCR files?
10. How to backup of voting file?
11. How do I identify the voting disk location?
12. How do I identify the OCR file location?
13. If voting disk/OCR file got corrupted and don't have backups, how to get them?
14. Who will manage OCR files?
15. Who will backup of OCR files?
16. What is VIP? Advantages and disadvantages of VIP? Need and Importance of VIP?
17. What are Oracle Cluster ware/Daemon processes and what they do?
18. What are the special background processes for RAC?
19. What are structural changes or new features in 11g R2 RAC?
20. What is the cache fusion?
21. What is the purpose of Private Interconnect?
22. What is split brain syndrome?
23. What are various IPs used in RAC? Or How may IPs we need in RAC?
24. What is the use of SCAN IP (SCAN name) and will it provide load balancing?
25. How many SCAN listeners will be running?
26. What is FAN?
27. What is FCF?
28. What is TAF and TAF policies?
29. How will you upgrade RAC database?
30. What are rolling patches and how to apply in RAC?
31. How to add/remove a node?
32. What are node apps?
33. What is gsd (Global Service Daemon)?
34. How to do load balancing in RAC? What is client balancing and server side balancing?
35. What are the uses of services? How to find out the services in cluster?
36. How to find out the nodes in cluster (or) how to find out the master node?
37. How to know the public IPs, private IPs, VIPs in RAC?
38. What utility is used to start DB/instance?
39. How can you shutdown single instance in the cluster environment?
40. What is HAS (High Availability Service) and the commands?
41. How many nodes are supported in a RAC Database?
42. What is fencing?
43. Why Cluster ware installed in root (why not oracle)?

44. What are the wait events in RAC?
45. What is the difference between cr block and cur (current) block?
46. Why Node Eviction happens on Oracle RAC?
47. What are the initialization parameters that must have same value for every instance in an Oracle RAC database?
48. What is Miscount (MC) in Oracle RAC?
49. What is the use of CSS Heartbeat Mechanism in Oracle RAC?
50. What happens if latencies to voting disks are longer?
51. What is CSS miscount?
52. How to change the CSS miscount default value?
53. How to start and stop CRS?
54. How to move regular DB to an ASM disk group?
55. What is a NIC card and HBA card?
56. What is a TPS?
57. What is the use of crs_getperm command?
58. What is the use of crs_profile?
59. Where will you check for RAC log files?
60. What is OCFS?
61. What is Oracle Cluster Ware?
62. What is a resource?
63. How to register a resource?
64. What does crs_start / crs_stop does?
65. What is the difference between Oracle Cluster ware and CRS?
66. What is Oracle recommendation for interconnect?
67. List the commands used to manage RAC?
68. What are the log file locations for RAC?
69. How to restore OCR file if corrupted?
70. How to compare all nodes with cluvfy?
71. How to manage ASM in RAC?
72. Where are the Cluster ware files stored on a RAC environment?
73. Where are the database software files stored on a RAC environment?
74. What kind of storage we can use for the shared Cluster ware files?
75. What kind of storage we can use for the RAC database storage?
76. What is a CFS?
77. What is an OCFS2?
78. Which files can be placed on an Oracle Cluster File System?
79. Do you know another Cluster Vendor?
80. How is possible to install a RAC if we don't have a CFS?
81. What is a raw device?
82. What is a raw partition?
83. When to use CFS over raw?
84. When to use raw over CFS?
85. What CRS is?
86. Why we need to have configured SSH or RSH on the RAC nodes?
87. Is the SSH, RSH needed for normal RAC operations?
88. Do we have to have Oracle RDBMS on all nodes?
89. What are the restrictions on the SID with a RAC database? Is it limited to 5 characters?
90. Does Real Application Clusters support heterogeneous platforms?

91. What is the Load Balancing Advisory?
92. What is the Cluster Verification Utility (cluvfy)?
93. Are there any issues for interconnect when sharing the same switch as the public network by using VLAN to separate the network?
94. What versions of the database can I use the cluster verification utility (cluvfy) with?
95. If I am using Vendor Clusterware such as Veritas, IBM, Sun or HP, do I still need Oracle Clusterware to run Oracle RAC 10g?
96. Is RAC on VM Ware supported?
97. What is hangcheck timer used for?
98. Is the hangcheck timer still needed with Oracle RAC 10g?
99. What files can I put on Linux OCFS2?
100. Is it possible to use ASM for the OCR and voting disk?
101. Can I change the name of my cluster after I have created it when I am using Oracle Clusterware?
102. What the O2CB is?
103. What is the recommended method to make backups of a RAC environment?
104. What command would you use to check the availability of the RAC system?
105. What is the minimum number of instances you need to have in order to create a RAC?
106. Name two specific RAC background processes
107. Can you have many database versions in the same RAC?
108. What was RAC previous name before it was called RAC?
109. What RAC component is used for communication between instances?
110. What is the difference between normal views and RAC views?
111. Which command will we use to manage (stop, start) RAC services in command-line mode?
112. How many alert logs exist in a RAC environment?
113. How do you know you lost the voting disk?
114. What format is the OCR file?
115. What will happen if we lost the voting disk?
116. What is the network protocol you used in configuring RAC?
117. How you check the health of Your RAC Database?
118. If there is some issue with virtual IP how will you troubleshoot it? How will you change virtual ip?
119. What kind of backup strategy you follow for your Databases?
120. What will you backup in your RAC Database?
121. How to recover your RAC Database?
122. What kind of backup strategy you are following for application server?
123. How your Add node to your RAC Database?
124. For a Database created with ASM on RAC How you would add one more ASM configuration?
125. How you add node for a RAC cluster? Step by step?
126. Which CRS process starts first?
127. What are the ways to configure TAF and Load Balancing?
128. When to use -repair parameter of ocrconfig command?
129. What is crs_stat? What is the meaning of TARGET and STATUS column in crs_stat command output?
130. What is service? How to use services to gain maximum use of RAC?
131. What is split Brain Syndrome? How Oracle Cluster ware handles it?
132. What is STONIH algorithm?
133. What is cache fusion? Which Database background process facilitates it?
134. What is GRD? Where does it reside?
135. Diff between cluster file system, RAW device and ASM?
136. Architecture of RAC?

137. Explain how Instance Recovery takes place in RAC?
138. How does your client connect to VIP or public IP? Or is it your choice?
139. Can private IP be changed?
140. What does root.sh do when you install 10g RAC? What is the importance of executing oraInstRoot.sh and root.sh scripts in Oracle Standalone and RAC environment?
141. How listener does handle requests in RAC?
142. How can cache fusion improve or degrade performance?
143. Will you increase parallelism if you have RAC, to gain inter instance parallelism? What are the considerations to decide?
144. What is single point of failure in RAC?
145. A query running fast on one node is very slow on other node. All the nodes have same configurations. What could be the reasons in RAC environment?
146. Does RMAN behave differently in RAC?
147. Can archive logs be placed on ASM disk? What about on RAW?
148. Write a RMAN script for taking backup of the database including arch log files in RAC?
149. Write a sample script for RMAN for the recovery if all the instance are down.(First explain the procedure how you will restore?
150. Clients are performing some operation and suddenly one of the datafile is experiencing problem what do you do in RAC?
151. What is the difference between a OS cluster and a RAC cluster?
152. What happens when a DML is issued in a RAC environment, how are requests for common buffers handled in a RAC environment?
153. Explain about checkpoint and local & Remote listener in RAC?
154. Explain LOCK Monitoring in RAC?
155. Describe a scenario in which a vendor cluster ware is required, in addition to the Oracle 10g Cluster ware?
156. How new connection establish in Oracle RAC?
157. What are the characteristics of VIP in oracle RAC?
158. What information being written in the vote disk when Split brain syndrome occurs?
159. Who does RAC do incase node becomes inactive?
160. When can I use TAF or FCF?
161. I have a java application that uses JDBC and a RAC database - should I use FCF?
162. Will I have to change my application?
163. What happens if the GSD does not run in 10g RAC, Will there be any impact to the 10g RAC when GSD is not running? Or else GSD should run in 10g RAC even?
164. Is it possible in a RAC environment to force the database node that you want to connect too?
165. If my OCR and Voting Disks are in ASM, can I shut down the ASM instance?
166. I have changed my spfile with alter system set parameter_name with scope=spfile. The spfile is on ASM storage and the database will not start.
167. How do I use DBCA in silent mode to set up RAC and ASM?
168. How does OCR mirror work? What happens if my OCR is lost / corrupt?
169. How do troubleshoot node reboot?
170. What do you do if you see GC CR BLOCK LOST in top 5 Timed Events in AWR Report?
171. SRVCTL cannot start instance, I get the following error PRKP-1001 CRS-0215, however SQLPLUS can start it on both nodes? How do you identify the problem?
172. What are the major RAC wait events?
173. What is usage of CRS_RELOCATE command?
174. What is the use of CRS_GETPERM and CRS_SETPERM?
175. What is the use of CRS_REGISTER and CRS_UNREGISTER?

176. What is the use of CRS_PROFILE?
177. What components in RAC must reside in shared storage?
178. What is the significance of using cluster-aware shared storage in an Oracle RAC environment?
179. Give few examples for solutions that support cluster storage?
180. How can we configure the cluster interconnect?
181. How do users connect to database in an Oracle RAC environment?
182. What are the characteristics controlled by Oracle services feature?
183. Which enables the load balancing of applications in RAC?
184. Give situations under which VIP address failover happens?
185. What is the significance of VIP address failover?
186. What are the administrative tools used for Oracle RAC environments?
187. How do we verify that RAC instances are running?
188. Where can we apply FAN UP and DOWN events?
189. State the use of FAN events in case of a cluster configuration change
190. Why should we have separate homes for ASM instance?
191. What is rolling upgrade? Can rolling upgrade be used to upgrade from 10g to 11g database?
192. Can the DML_LOCKS and RESULT_CACHE_MAX_SIZE be identical on all instances?
193. What two parameters must be set at the time of starting up an ASM instance in a RAC environment?
194. How does an Oracle Clusterware manage CRS resources?
195. Name some Oracle Clusterware tools and their uses?
196. What are the modes of deleting instances from Oracle Real Application cluster Databases?
197. How do we remove ASM from an Oracle RAC environment?
198. How do we verify that an instance has been removed from OCR after deleting an instance?
199. What are the performance views in an Oracle RAC environment?
200. What is the difference between server-side and client-side connection load balancing?
201. Give the usage of srvctl?
202. What is the purpose of the ONS daemon?
203. What is is Dynamic Remastering?
204. What are RAC based services? What are difference between normal database service and RAC services?
205. What happens if one of the node is not able to access the voting disk?
206. What happens if all of the nodes not able to access the voting disk?
207. What happens if one of the node is not able to communicate via private interconnect?
208. What happens if all of the nodes not able to communicate via private interconnect?
209. What is split brain syndrome?
210. Which background daemons initiates node eviction?
211. Which background daemon starts clusterware or the resources?
212. What are my options for load balancing with Oracle RAC? Why do I get an uneven number of connections on my instances?
213. How can a customer mask the change in their clustered database configuration from their client or application? (I.E. So I do not have to change the connection string when I add a node to the Oracle RAC database)?
214. What is the Load Balancing Advisory?
215. How do I enable the load balancing advisory?
216. Why do we have a Virtual IP (VIP) in Oracle RAC 10g or 11g? Why does it just return a dead connection when its primary node fails?
217. What are my options for setting the Load Balancing Advisory GOAL on a Service?
218. While executing root.sh power loss or pressed the CTLRT+C key, so what is next step to do?
219. How can you connect to a specific node in a RAC?

- 220. What is the Oracle Recommendation for backing up voting disk?
- 221. How can we add and remove multiple voting disks?
- 222. When can we use -force option?

1. What is the use of RAC? What is RAC? How is it different from standalone database?

1. 100% uptime
2. Performance
3. Scalability

2. What are the prerequisites for RAC setup?

(1) Checking the Hardware Requirements

- **Physical memory:** At least 1GB RAM.
- **Swap space:** If RAM is between 1 GB and 2 GB then make swap space to 1.5 times of the size of RAM. If RAM is more than 2GB then make swap space to the equal of the size of the RAM.
- **Temporary space:** At least 400 MB typically in /tmp directory.
- **Processor type (CPU):** Need to be certified with the version of the Oracle software being installed.
- **Hard Disk Space:** 1.5 GB for oracle database home directory + 1.5GB for the ASM home directory + 120 oracle clusterware software installation + Two Oracle Clusterware components OCR 256 MB each, or 512 MB total disk space + Three Oracle Clusterware components Voting Disk 256 MB each, or 768 MB total disk space.
- All the nodes in the cluster must have same hardware architecture, however can have machines of different speeds and size in the same cluster.

On UNIX system you can check hardware components as follows.

- To determine physical RAM size, **# grep MemTotal /proc/meminfo**
- To determine the configured swap space, **# grep SwapTotal /proc/meminfo**
- To determine the amount of disk space available in the /tmp directory, **# df -k /tmp**
- To determine free disk space on the system, **#df -h** or **#df -k**
- To determine whether the system architecture, **# grep "model name" /proc/cpuinfo**

(2) Checking the Network Requirements

Network Hardware Requirements:

- One private interconnect is needed for Oracle Clusterware for the use to synchronize each instance's use of the shared resources and for Oracle RAC to interconnect to transmit data blocks that are shared between the instances. Thus each node needs at least two network interface cards, or network adapters. One adapter is for the public network and the other adapter is for the private network
- The **public interface names associated with the network adapters for each network must be the same on all nodes**, and the private interface names associated with the network adapters should be the same on all nodes. For example if in server1/node1, eth0 is the public interface then on server2/node2 eth0 must be the public interface.
- For the public network, each network adapter must support TCP/IP.
- For the private network, the interconnect must support the user datagram protocol (UDP) using high-speed network adapters and switches that support TCP/IP.
- Note UDP is the default interconnect protocol for Oracle RAC, and TCP is the interconnect protocol for Oracle Clusterware.
- For the private network, the endpoints of all designated interconnect interfaces must be completely reachable on the network. All nodes must be accessible between through private network. You can check the reachability by ping command.

Network Parameter Requirements:

If NFS is used for the shared storage, then you must set the values for the NFS buffer size parameters rsize and wsize to at least 16384. Oracle recommends that you use the value 32768.

You can set the value by updating the /etc/fstab file on each node with an entry similar to the following,

clusternode:/vol/DATA/oradata /home/oradata/app nfs

rw,bg,vers=3,tcp,hard,nointr,timeo=600,rsize=32768,wsize=32768,actimeo=0 1 2

IP Address Requirements:

- You must have at least three IP addresses available for each node
 1. An IP address for the public interface, Interface name should be the name of the node name.
 2. An IP address for the private interface. Interface name should be hostname-priv.
 3. One virtual IP address with an associated network name, Interface name should be hostname-VIP.
- The VIP is on the same subnet as your public interface and its address will be not used currently in the network.

- For public and virtual addresses register with an associated network name in DNS. If you do not have an available DNS, then record the all network names and interface names in the system hosts file, /etc/hosts.
- Identify the interface names and associated IP addresses for all network adapters by running the following command on each node:

/sbin/ifconfig

(3)Node Time Requirements

Ensure that each member node of the cluster is set as closely as possible to the same date and time. Oracle strongly recommends using the Network Time Protocol (NTP) feature of most operating systems for this purpose.

(4)Verifying the Installed Operating System and Software Requirements

- To determine which distribution and version of Linux is installed, run the following command as the root user:

#cat /etc/issue

- The Linux kernel is updated to fix bugs. These kernel updates are referred to as erratum kernels or errata levels. To determine if the required errata level is installed, use the following procedure as the root user:

#uname -r

3. What is oracle clusterware? Benefits of Clusterware? What are the Cluster ware components?

Clusterware:

Oracle Clusterware is the software that enables servers to operate together as if they are one server. Each server looks like any standalone server. However, each server has additional processes that communicate with each other so the separate servers appear as if they are one server to applications and end users.

The benefits of using a cluster ware:

- Scalability for applications
- Using lower-cost hardware
- Ability to fail over
- Ability to grow the capacity over time by adding servers, when needed

Voting Disk - Oracle RAC uses the voting disk to manage cluster membership by way of a health check and decides cluster ownership among the instances in case of network failures. The voting disk must reside on shared disk.

Oracle Cluster Registry (OCR) - Maintains cluster configuration information as well as configuration information about any cluster database within the cluster. The OCR must reside on shared disk that is accessible by all of the nodes in your cluster. The daemon OCSSd manages the configuration info in OCR and maintains the changes to cluster in the registry.

- Oracle Cluster Registry or OCR is a component of Oracle Cluster ware Framework.
- It stores profile attribute information.
- Oracle RAC consists of series of resources.
- Other applications can also be treated as a resource.
- OCR contains information pertaining to instance-to-node mapping

Note: you can't have more than two OCRs.

Virtual IP (VIP) - When a node fails, the VIP associated with it is automatically failed over to some other node and new node re-arps the world indicating a new MAC address for the IP. Subsequent packets sent to the VIP go to the new node, which will send error RST packets back to the clients. This results in the clients getting errors immediately.

- It is used for failover and RAC management

Using virtual IP we can save our TCP/IP timeout problem because Oracle notification service (ONS) maintains communication between each nodes and listeners. Once ONS found any listener down or node down, it will notify another nodes and listeners. While new connection is trying to establish connection to failure node or listener, virtual IP of failure node automatically divert to surviving node and session will be establishing in another surviving node. This process doesn't wait for TCP/IP timeout event. Due to this new connection gets faster session establishment to another surviving nodes/listener.

Virtual IP (VIP) is for fast connection establishment in failover dictation. Still we can use physical IP address in Oracle 10g in listener if we have no worry for failover timing. We can change default TCP/IP timeout using operating system utilities/commands and kept smaller. But taking advantage of VIP (Virtual IP address) in Oracle 10g RAC database is advisable.

4. What is OCR file?

RAC configuration information repository that manages information about the cluster node list and instance-to-node mapping information. The OCR also manages information about Oracle Clusterware resource profiles for customized applications. Maintains cluster configuration information as well as configuration information about any cluster

database within the cluster. The OCR must reside on shared disk that is accessible by all of the nodes in your cluster. The daemon OCSSd manages the configuration info in OCR and maintains the changes to cluster in the registry. Oracle Cluster Registry (OCR) and recovering it. Oracle Clusterware automatically creates OCR backups every four hours and it always retains the last three backup copies of the OCR. The CRSD process that creates the backups also creates and retains an OCR backup for each full day and then at the end of a week a complete backup for the week. So there is a robust backup taking place in the background. And you guessed it right; you cannot alter the backup frequencies. This is meant to protect you, the DBA, so that you can copy these generated backup files at least once daily to a different device from where the primary OCR resides. These files are located at %CRS_home/cdata/my_cluster

5. What is voting file/disk and how many files should be there? Why do we have to create odd number of voting disk? (<http://oracleinaction.blogspot.in/2012/12/votedisk.html>)

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.

Oracle Clusterware uses the voting disk to determine which instances are members of a cluster. The voting disk must reside on a shared disk. Basically all nodes in the RAC cluster register their heart-beat information on these voting disks. The number decides the number of active nodes in the RAC cluster. These are also used for checking the availability of instances in RAC and remove the unavailable nodes out of the cluster. It helps in preventing split-brain condition and keeps database information intact. The split brain syndrome and its affects and how it has been managed in oracle is mentioned below.

For high availability, Oracle recommends that you have a minimum of three voting disks. If you configure a single voting disk, then you should use external mirroring to provide redundancy. You can have up to 32 voting disks in your cluster.

How many files:

Oracle recommends that you do not use more than five voting disks. The maximum number of voting disks that is supported is 15.

Why do we have to create odd number of voting disk?

Explanation-1:

The odd number of voting disks configured provides a method to determine who in the cluster should survive. A node must be able to access more than half of the voting disks at any time. For example, let's have a two node cluster with an even number of let's say 2 voting disks. Let Node1 is able to access voting disk1 and Node2 is able to access voting disk2. This means that there is no common file where clusterware can check the heartbeat of both the nodes. If we have 3 voting disks and both the nodes are able to access more than half i.e. 2 voting disks, there will be at least on disk which will be accessible by both the nodes. The clusterware can use that disk to check the heartbeat of both the nodes. Hence, each node should be able to access more than half the number of voting disks. A node not able to do so will have to be evicted from the cluster by another node that has more than half the voting disks, to maintain the integrity of the cluster. After the cause of the failure has been corrected and access to the voting disks has been restored, you can instruct Oracle Clusterware to recover the failed node and restore it to the cluster.

Note: Loss of more than half your voting disks will cause the entire cluster to fail

Explanation-2:

As far as voting disks are concerned, a node must be able to access strictly more than half of the voting disks at any time. So if you want to be able to tolerate a failure of n voting disks, you must have at least 2n+1 configured. (n=1 means 3 voting disks). You can configure up to 32 voting disks, providing protection against 15 simultaneous disk failures.

Oracle recommends that customers use 3 or more voting disks in Oracle RAC 10g Release 2. Note: For best availability, the 3 voting files should be physically separate disks. It is recommended to use an odd number as 4 disks will not be any more highly available than 3 disks, 1/2 of 3 is 1.5...rounded to 2, 1/2 of 4 is 2, once we lose 2 disks, our cluster will fail with both 4 voting disks or 3 voting disks.

- Does the cluster actually check for the vote count before node eviction? If yes, could you explain this process briefly?

Yes. If you lose half or more of all of your voting disks, then nodes get evicted from the cluster, or nodes kick themselves out of the cluster.

- What's the logic behind the documentation note that says that "each node should be able to see more than half of the voting disks at any time"?

The answer to first question itself will provide answer to this question too.

6. How to backup of OCR file?

```
#ocrconfig -manualbackup
```

```
#ocrconfig -export file_name.dmp
```

```
#ocrdump -backupfile my_file
```

```
$cp -p -R /u01/app/crs/cdata /u02/crs_backup/ocrbackup/RAC1
```

7. How to recover OCR file?

```
#ocrconfig -restore backup_file.ocr
```

```
#ocrconfig -import file_name.dmp
```

8. What is local OCR?

```
/etc/oracle/local.ocr
```

```
/var/opt/oracle/local.ocr
```

10. How to check backup of OCR files?

```
#ocrconfig -showbackup
```

11. How to backup of voting file?

```
dd if=/u02/ocfs2/vote/VDFile_0 of=$ORACLE_BASE/bkp/vd/VDFile_0
```

```
crsctl backup css votedisk -- from 11g R2
```

12. How do I identify the voting disk location?

```
# crsctl query css votedisk
```

18. How do I identify the OCR file location?

check /var/opt/oracle/ocr.loc or /etc/ocr.loc

ocrcheck

13. If voting disk/OCR file got corrupted and don't have backups, how to get them?

We have to install Cluster ware.

14. Who will manage OCR files?

cssd will manage OCR.

15. Who will backup of OCR files?

crsd will take backup.

16. What is VIP? Advantages and disadvantages of VIP? Need? Importance of VIP?

Public Interface: Used for normal network communications to the node

Private Interface: Used as the cluster interconnect

Virtual (Public) Interface: Used for failover and RAC management

Need of VIP:

To assure that oracle clients quickly failover when a node fails

From Oracle 10g, virtual IP considers to configure listener. Using virtual IP we can save our TCP/IP timeout problem because Oracle notification service maintains communication between each nodes and listeners. Once ONS found any listener down or node down, it will notify another nodes and listeners with same situation. While new connection is trying to establish connection to failure node or listener, virtual IP of failure node automatically divert to surviving node and session will be establishing in another surviving node. This process doesn't wait for TCP/IP timeout event. Due to this new connection gets faster session establishment to another surviving nodes/listener.

When a node fails, the VIP associated with it is automatically failed over to some other node and new node re-arms the world indicating a new MAC address for the IP. Subsequent packets sent to the VIP go to the new node, which will send error RST packets back to the clients. This results in the clients getting errors immediately.

Without using VIPs or FAN, clients connected to a node that died will often wait for a TCP timeout period (which can be up to 10 min) before getting an error. As a result, you don't really have a good HA solution without using VIPs.

It returns a dead connection IMMEDIATELY, when its primary node fails. Without using VIP IP, the clients have to wait around 10 minutes to receive ORA-3113: "end of file on communications channel". However, using Transparent Application Failover (TAF) could avoid ORA-3113.

Advantage of Virtual IP deployment in Oracle RAC:

Using VIP configuration, client can be able to get connection fast even fail over of connection request to node. Because vip automatically assign to another surviving node faster and it can't wait for TNS timeout old fashion.

Disadvantage of Virtual IP deployment in Oracle RAC:

Some more configurations is needed in system for assign virtual IP address to nodes like in /etc/hosts and others. Some misunderstanding or confusion may occur due to multiple IP assigns in same node.

Importance for VIP configuration:

The VIPs should be registered in the DNS. The VIP addresses must be on the same subnet as the public host network addresses. Each Virtual IP (VIP) configured requires an unused and resolvable IP address.

17. What are Oracle Cluster ware/Daemon processes and what they do?

OCSSD, CRSd, EVMD

OCSSD (cluster synchronization Services):

1. CSS provides basic Group Services Support; it is a distributed group membership system that allows applications to coordinate activities to achieve a common result.
2. Group services use vendor cluster ware group services when it is available.
3. Lock services provide the basic cluster-wide serialization locking functions; it uses the First In, First out (FIFO) mechanism to manage locking
4. Node services uses OCR to store data and updates the information during reconfiguration, it also manages the OCR data which is static otherwise.

Functionality: enables basic cluster services ==> new/ lost node information is written to OCR

-The node membership functionality

- Basic locking

Failure of the process: Node restart.

OCRSD (cluster Registry Services): A daemon process that is used for synchronization between different RAC nodes and synchronization between ASM and Oracle database instances.

1. Manages the RAC resources (a database, an instance, a service, a listener, a virtual IP (VIP) address, an Application process)
2. Manages the Oracle Cluster Registry and stores the current known state in the Oracle Cluster Registry
3. Runs as 'root' on UNIX and automatically restarts in case of failure.
4. CRSd manages the resources like starting and stopping the services and failing-over the application Resources. It spawns separate processes to manage application resources.
5. Failure of this daemon causes the node to be rebooted to avoid split-brain situations.

Failure of the process: the crsd restarts automatically, without restarting the node.

CRSd can run in 2 modes:

reboot mode -> when crsd starts all the resources are restarted.

restart mode -> when crsd starts the resources are started as these were before the shutdown.

When CRS is installed on the cluster where a 3rd-party clusterware is integrated (there are 2 clusterware on the cluster)

-> CRSd manages:

When CRS is the ONLY ONE clusterware on the cluster

-> CRSd manages:

- Oracle RAC services and resources
- the node membership functionality (by CSSd, but CSS is managed by CRSd)

COMMENT:

In order to start the crsd we need:

- The public interface, the private interface and the virtual IP (VIP) should be up and running !
- These IPs must be pingable to each other.

Note: CRS requires the public interface, private interface and the [Virtual IP](#) (VIP) for the operation. All these interfaces should be up and running should be able to ping each other before starting CRS Installation. Without the above network infrastructure CRS cannot be installed

Functionality:

1. It will update all the changes in the OCR

2. Start, stop of the resources
3. Failover of the application resources
4. Node recovery
5. Automatically restarts the RAC resources when a failure occurs

Event Management (EVM): A background process that publishes events that crs creates. which runs the EVMD process. The daemon spawns processes called evmlogger and generates the events when things happen. The evmlogger spawns new children processes on demand and scans the callout directory to invoke callouts. Death of the EVMD daemon will not halt the instance and will be restarted. Run as a oracle user.

Functionality: - a background process that publishes events that crs creates.

Failure of the process: the evmd restarts automatically, without restarting the node.

CRS Process	Functionality	Failure of the Process	Run AS
OPROCD - Process Monitor	provides basic cluster integrity services	Node Restart	root
EVMD - Event Management	spawns a child process event logger and generates callouts	Daemon automatically restarted, no node restart	oracle
OCSSD - Cluster Synchronization Services	basic node membership, group services, basic locking	Node Restart	oracle
CRSD - Cluster Ready Services	resource monitoring, failover and node recovery	Daemon restarted automatically, no node restart	root

18. What are the special background processes for RAC?

DIAG, LCKn, LMD, LMSn, LMON

LMS: The Global Cache Service background processes (LMSn) manage requests for data access between the nodes of the cluster. Each block is assigned to a specific instance using the same hash algorithm that is used for global resources. The instance managing the block is known as the resource master. When an instance requires access to a specific block, a request is sent to an LMS process on the resource master requesting access to the block. The LMS process can build a read-consistent image of the block and return it to the requesting instance, or it can forward the request to the instance currently holding the block.

The LMS processes coordinate block updates, allowing only one instance at a time to make changes to a block and ensuring that those changes are made to the most recent version of the block. The LMS process on the resource master is responsible for maintaining a record of the current status of the block, including whether it has been updated.

In Oracle 9.0.1 and Oracle 9.2 there can be up to 10 LMSn background processes (LMS0 to LMS9) per instance; in Oracle 10.1 there can be up to 20 LMSn background processes (LMS0 to LMS9, LMSa to LMSj) per instance; in Oracle 10.2 there can be up to 36 LMSn background processes (LMS0 to LMS9, LMSa to LMSz). The number of required LMSn processes varies depending on the amount of messaging between the nodes in the cluster.

LMON: In a single-instance database, access to database resources is controlled using enqueues that ensure that only one session has access to a resource at a time and that other sessions wait on a first in, first out (FIFO) queue until the resource becomes free. In a single-instance database, all locks are local to the instance. In a RAC database there are global resources, including locks and enqueues that need to be visible to all instances. For example, the database mount lock that is used to control which instances can concurrently mount the database is a global enqueue, as are library cache locks, which are used to signal changes in object definitions that might invalidate objects currently in the library cache.

The Global Enqueue Service Monitor (LMON) background process is responsible for managing global enqueues and resources. It also manages the Global Enqueue Service Daemon (LMD) processes and their associated memory areas. LMON is similar to PMON in that it also manages instance and process expirations and performs recovery processing on global enqueues. In Oracle 10.1 and below there is only one lock monitor background process.

LMDn: The current status of each global enqueue is maintained in a memory structure in the SGA of one of the instances. For each global resource, three lists of locks are held, indicating which instances are granted, converting, and waiting for the lock.

The LMD background process is responsible for managing requests for global enqueues and updating the status of the enqueues as requests are granted. Each global resource is assigned to a specific instance using a hash algorithm. When an instance requests a lock, the LMD process of the local instance sends a request to the LMD process of the

remote instance managing the resource. If the resource is available, then the remote LMD process updates the enqueue status and notifies the local LMD process. If the enqueue is currently in use by another instance, the remote LMD process will queue the request until the resource becomes available. It will then update the enqueue status and inform the local LMD process that the lock is available.

The LMD processes also detect and resolve deadlocks that may occur if two or more instances attempt to access the two or more enqueues concurrently.

In Oracle 10.1 and below there is only one lock monitor daemon background process named LMD0.

LCK0: The instance enqueue background process (LCK0) is part of GES. It manages requests for resources other than data blocks—for example, library and row cache objects. LCK processes handle all resource transfers not requiring Cache Fusion. It also handles cross-instance call operations.

In Oracle 9.0.1 there could be up to ten LCK processes (LCK0 to LCK9). In Oracle 9.2 and Oracle 10.1 and 10.2 there is only one LCK process (LCK0).

DIAG: The DIAG background process captures diagnostic information when either a process or the entire Instance fails. This information is written to a subdirectory within the directory specified by the BACKGROUND_DUMP_DEST initialization parameter. The files generated by this process can be forwarded to Oracle Support for further analysis.

There is one DIAG background process per instance. It should not be disabled or removed. In the event that the DIAG background process itself fails, it can be automatically restarted by other background processes.

19. What are structural changes or new features in 11g R2 RAC?

- Grid & ASM are on one home,
- Voting disk & ocrfile can be on the ASM
- SCAN
- By using srvctl, we can manage diskgroups, home, ons, eons, filesystem, srvpool, server, scan, - scan_listener, gns, vip, oc4j,
- GSD

We can store everything on the [ASM](#). We can store OCR & voting files also on the ASM.

- ASMCA
- [Single Client Access Name \(SCAN\)](#) - eliminates the need to change tns entry when nodes are added to or removed from the Cluster. RAC instances register to SCAN listeners as remote listeners. SCAN is fully qualified name. Oracle recommends assigning 3 addresses to SCAN, which create three SCAN listeners.
- Clusterware components: crfmond, crflogd, GIPCD.
- AWR is consolidated for the database.
- 11g Release 2 Real Application Cluster (RAC) has server pooling technologies so it's easier to provision and manage database grids. This update is geared toward dynamically adjusting servers as corporations manage the ebb and flow between data requirements for datawarehousing and applications.
- By default, LOAD_BALANCE is ON.
- GSD (Global Service Daemon), [gsdctl](#) introduced.
- GPnP profile.
- Cluster information in an XML profile.
- Oracle [RAC OneNode](#) is a new option that makes it easier to consolidate databases that aren't mission critical, but need redundancy.
- [raconeinit](#) - to convert database to RacOneNode.
- raconefix - to fix RacOneNode database in case of failure.
- racone2rac - to convert RacOneNode back to RAC.
- Oracle Restart - the feature of Oracle Grid Infrastructure's [High Availability Services](#) (HAS) to manage associated listeners, ASM instances and Oracle instances.
- [Oracle Omotion](#) - Oracle 11g release2 RAC introduces new feature called Oracle Omotion, an online migration utility. This Omotion utility will relocate the instance from one node to another, whenever instance failure happens.
- Omotion utility uses Database Area Network (DAN) to move Oracle instances. Database Area Network (DAN) technology helps seamless database relocation without losing transactions.
- Cluster Time Synchronization Service (CTSS) is a new feature in Oracle 11g R2 RAC, which is used to synchronize time across the nodes of the cluster. CTSS will be replacement of NTP protocol.

- [Grid Naming Service \(GNS\)](#) is a new service introduced in Oracle RAC 11g R2. With GNS, Oracle Clusterware (CRS) can manage Dynamic Host Configuration Protocol (DHCP) and DNS services for the dynamic node registration and configuration.
- Cluster interconnect: Used for data blocks, locks, messages, and SCN numbers.
- Oracle Local Registry (OLR) - From Oracle 11gR2 "Oracle Local Registry (OLR)" something new as part of Oracle Clusterware. OLR is node's local repository, similar to OCR (but local) and is managed by OHASD. It pertains data of local node only and is not shared among other nodes.
- Multicasting is introduced in 11gR2 for private interconnect traffic.
- I/O fencing prevents updates by failed instances, and detecting failure and preventing split brain in cluster. When a cluster node fails, the failed node needs to be fenced off from all the shared disk devices or diskgroups. This methodology is called I/O Fencing, sometimes called Disk Fencing or failure fencing.
- Re-bootless node fencing (restart) - instead of fast re-booting the node, a graceful shutdown of the stack is attempted.
- Clusterware log directories: acfs*
- HAIP (IC VIP).
- Redundant interconnects: NIC bonding, HAIP.
- RAC background processes: DBRM – Database Resource Manager, PING – Response time agent.
- Virtual Oracle 11g RAC cluster - Oracle 11g RAC supports virtualization.

20. What is the cache fusion?

Transferring of data between RAC instances by using private network. Cache Fusion is the remote memory mapping of Oracle buffers, shared between the caches of participating nodes in the cluster. When a block of data is read from datafile by an instance within the cluster and another instance is in need of the same block, it is easy to get the block image from the instance which has the block in its SGA rather than reading from the disk.

Whoever knows the basics of RAC should very well be aware of the fact that CACHE FUSION is one of the most important and interesting concepts in a RAC setup. As the name suggests, CACHE FUSION is the amalgamation of cache from each node/instance participating in the RAC, but it is not any physically secured memory component which can be configured unlike the usual buffer cache (or other SGA components) which is local to each node/instance.

We know that every instance of the RAC database has its own local buffer cache which performs the usual cache functionality for that instance. Now there could be occasions when a transaction/user on instance A needs to access a data block which is being owned/locked by the other instance B. In such cases, the instance A will request instance B for that data block and hence accesses the block through the interconnect mechanism. This concept is known as CACHE FUSION where one instance can work on or access a data block in other instance's cache via the high speed interconnect.

Cache Fusion architecture helps resolve each possible type of contentions that could be thought of in a multi-node RAC setup. We will look at them in detail in coming sections but first let us understand few very important terms/concepts which will be useful in understanding the contentions which we are going to discuss in later sections.

Global Cache Service

Global Cache Service (GCS) is the heart of Cache Fusion concept. It is through GCS that data integrity in RAC is maintained when more than one instance need a particular data block. Instances look up to the GCS for fulfilling their data block needs.

GCS is responsible for:

1. Tracking the data block
2. Accepting the data block requests from instances
3. Informing the holding instance to release the lock on the data block or ship a CR image
4. Coordinating the shipping of data blocks as needed between the instance through the interconnect
5. Informing the instances to keep or discard PIs

More about the above functions will be clear from the following discussion on contention. Please note that GCS is available in the form of the background process called LMS.

Past Image: The concept of Past Image is very specific to RAC setup. Consider an instance holding 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. Basically, PI is the copy of the data block before the block is written to the disk.

- There can be more than one PI of the block at a time across the instances. In case there is some instance crash/failure in the RAC and a recovery is required, Oracle is able to re-construct the block using these Past Images from all the instances.

When a block is written to the disk, all Past Images of that block across the instances are discarded. GCS informs all the instances to do this. At this time, the redo logs containing the redo for that data block can also be overwritten because they are no longer needed for recovery.

Consistent Read

A consistent read is needed when a particular block is being accessed/modified by transaction T1 and at the same time another transaction T2 tries to access/read the block. If T1 has not been committed, T2 needs a consistent read (consistent to the non-modified state of the database) copy of the block to move ahead. A CR copy is created using the UNDO data for that block. A sample series of steps for a CR in a normal setup would be:

1. Process tries to read a data block
2. Finds an active transaction in the block
3. Then checks the UNDO segment to see if the transaction has been committed or not
4. If the transaction has been committed, it creates the REDO records and reads the block
5. If the transaction has not been committed, it creates a CR block for itself using the UNDO/ROLLBACK information.
6. Creating a CR image in RAC is a bit different and can come with some I/O overheads. This is because the UNDO could be spread across instances and hence to build a CR copy of the block, the instance might have to visit UNDO segments on other instances and hence perform certain extra I/O. Possible contentions in a RAC setup and How CACHE FUSION helps resolve them

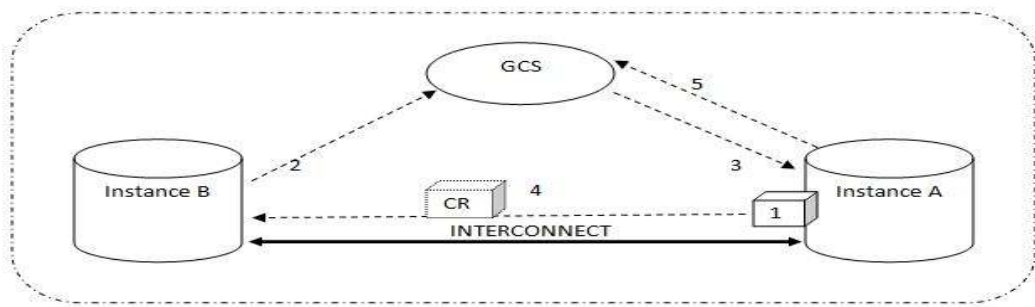
As mentioned above, CACHE FUSION helps resolve all the possible contentions that could happen between instances in a RAC setup. There are 3 possible contentions in a RAC setup which we are going to discuss in detail here with a mention of cache fusion where ever applicable.

Our discussion thus far should help understand the following discussion on contentions and their resolutions better.

1. **Read/Read contention:** Read-Read contention might not be a problem at all because the table/row will be in a shared lock mode for both transactions and none of them is trying an exclusive lock anyways.
2. **Read/Write contention:** This one is interesting.

Here is more about this contention and how the concept of cache fusion helps resolve this contention

1. A data block is in the buffer cache of instance A and is being updated. An exclusive lock has been acquired on it.
 - b. After some time instance B is interested in reading that same data block and hence sends a request to GCS. So far so good – Read/Write contention has been induced
 - c. GCS checks the availability of that data block and finds that instance A has acquired an exclusive lock. Hence, GCS asks instance A to release the block for instance B.
 - d. Now there are two options – either instance A releases the lock on that block (if it no longer needs it) and lets instance B read the block from the disk OR instance A creates a CR image of the block in its own buffer cache and ships it to the requesting instance via interconnect
 - e. The holding instance notifies the GCS accordingly (if the lock has been released or the CR image has been shipped)
 - f. Creation of CR image, shipping it to the requesting instance and involvement of GCS is where CACHE FUSION comes into play



3. Write/Write contention:

This is the case where both instance A as well as B are trying to acquire an exclusive lock on the data block. A data block is in the buffer cache of instance A and is being updated. An exclusive lock has been acquired on it

- Instance B send the data block request to the GCS
- GCS checks the availability of that data block and finds that instance A has acquired an exclusive lock. Hence, GCS asks instance A to release the block for instance B
- There are 2 options - either instance A releases the lock on that block (if it no longer needs it) and lets instance B read the block from the disk OR instance A creates a PI image of the block in its own buffer cache, makes the redo entries and ships the block to the requesting instance via interconnect
- Holding instance also notifies the GCS that lock has been released and a PI has been preserved
- Instance B now acquires the exclusive lock on that block and continues with its normal processing. At this point GCS records that data block is now with instance B
- The whole mechanism of resolving this contention with the due involvement of GCS is attributed to the CACHE FUSION.

PI image VS CR image

Let us just halt and understand some basic stuff - Wondering why CR image used in Read-Write contention and PI image used in Write-Write contention? What is the difference?

- CR image was shipped to avoid Read-Write type of contention because the requesting instance doesn't want to perform a write operation and hence won't need an exclusive lock on the block. Thus for a read operation, the CR image of the block would suffice. Whereas for Write-Write contention, the requesting instance also needs to acquire an exclusive lock on the data block. So to acquire the lock for write operations, it would need the actual block and not the CR image. The holding instance hence sends the actual block but is liable to keep the PI of the block until the block has been written to the disk. So if there is any instance failure or crash, Oracle is able to build the block using the PI from across the RAC instances (there could be more than one PI of a data block before the block has actually been written to the disk). Once the block is written to the disk, it won't need a recovery in case of a crash and hence associated PIs can be discarded.
- Another difference of course is that the CR image is to be shipped to the requesting instance where as the PI has to be kept by the holding instance after shipping the actual block.

UNDO?

This discussion is not about UNDO management in RAC but here is a brief about UNDO in a RAC scenario. UNDO is generated separately on each instance just similar to a standalone database. Each instance has its own UNDO tablespace. The UNDO data of all instances is used by holding instance to build CR image in case of contention

What is Cache Fusion and how does this affect applications:

Cache Fusion is new parallel database architecture for exploiting clustered computers to achieve scalability of all types of applications. Cache Fusion is a shared cache architecture that uses high speed low latency interconnects available today on clustered systems to maintain database cache coherency. Database blocks are shipped across the interconnect to the node where access to the data is needed. This is accomplished transparently to the application and users of the system. As Cache Fusion uses at most a 3 point protocol, this means that it easily scales to clusters with a large numbers of nodes. For more information about cache fusion see the following links: Additional

Information can be found at:

Note: 139436.1 Understanding 9i Real Application Clusters Cache Fusion

21. What is the purpose of Private Interconnect?

Clusterware uses the private interconnect for cluster synchronization (network heartbeat) and daemon communication between the clustered nodes. This communication is based on the TCP protocol. RAC uses interconnect for cache fusion (UDP) and inter-process communication (TCP).

The private interconnect is the physical construct that allows inter-node communication. It can be a simple crossover cable with UDP or it can be a proprietary interconnect with specialized proprietary communications protocol. When setting up more than 2- nodes, a switch is usually needed. This provides the maximum performance for RAC, which relies on inter-process communication between the instances for cache-fusion implementation.

22. What is split brain syndrome?

Explanation-1: Oracle ID 1425586.1

It will arise when two or more instances attempt to control a cluster database. In a two-node environment, one instance attempts to manage updates simultaneously while the other instance attempts to manage updates. In Oracle RAC environment all the instances/servers communicate with each other using high-speed interconnects on the private network. This private network interface or interconnect are redundant and are only used for inter-instance oracle data block transfers. Now talking about split-brain concept with respect to oracle RAC systems, it occurs when the instance members in a RAC fail to ping/connect to each other via this private interconnect, but the servers are all physically up and running and the database instance on each of these servers is also running. These individual nodes are running fine and can conceptually accept user connections and work independently. So basically due to lack of communication the instance thinks that the other instance that it is not able to connect is down and it needs to do something about the situation. The problem is if we leave these instances running, the same block might get read, updated in these individual instances and there would be data integrity issue, as the blocks changed in one instance, will not be locked and could be over-written by another instance. Oracle has efficiently implemented check for the split brain syndrome.

When there is no network heart bit then Split brain is done using Disk (vote disk) information

Explanation-2:

It occurs when the instance members in a RAC fail to ping/connect to each other via this private interconnect, but the servers are all physically up and running and the database instance on each of these servers is also running. These individual nodes are running fine and can conceptually accept user connections and work independently. So basically due to lack of communication the instance thinks that the other instance that it is not able to connect is down and it needs to do something about the situation. The problem is if we leave these instance running, the same block might get read, updated in these individual instances and there would be data integrity issue, as the blocks changed in one instance, will not be locked and could be over-written by another instance. Oracle has efficiently implemented check for the split brain syndrome.

In RAC if any node becomes inactive, or if other nodes are unable to ping/connect to a node in the RAC, then the node which first detects that one of the node is not accessible, it will evict that node from the RAC group. e.g. there are 4 nodes in a rac instance, and node 3 becomes unavailable, and node 1 tries to connect to node 3 and finds it not responding, then node 1 will evict node 3 out of the RAC groups and will leave only Node1, Node2 & Node4 in the RAC group to continue functioning.

The split brain concepts can become more complicated in large RAC setups. For example there are 10 RAC nodes in a cluster. And say 4 nodes are not able to communicate with the other 6. So there are 2 groups formed in this 10 node RAC cluster (one group of 4 nodes and other of 6 nodes). Now the nodes will quickly try to affirm their membership by locking controlfile, then the node that lock the controlfile will try to check the votes of the other nodes. The group with the most number of active nodes gets the preference and the others are evicted. Moreover, I have seen this node eviction issue with only 1 node getting evicted and the rest function fine, so I cannot really testify that if thats how it work by experience, but this is the theory behind it.

When we see that the node is evicted, usually oracle rac will reboot that node and try to do a cluster reconfiguration to include back the evicted node.

You will see oracle error: ORA-29740, when there is a node eviction in RAC. There are many reasons for a node eviction like heart beat not received by the controlfile, unable to communicate with the clusterware etc.

Explanation-3:

Voting Disk is a file that resides in the shared storage area and must be accessible by all nodes in the cluster. All nodes in the cluster register their heart-beat information in the voting disk, so as to confirm that they are all operational. If heart-beat information of any node in the voting disk is not available that node will be evicted from

the cluster. The CSS (Cluster Synchronization Service) daemon in the clusterware maintains the heart beat of all nodes to the voting disk. When any node is not able to send heartbeat to voting disk, then it will reboot itself, thus help avoiding the split-brain syndrome.

For high availability, Oracle recommends that you have a minimum of three or odd number (3 or greater) of voting disks.

According to Oracle – “An absolute majority of voting disks configured (more than half) must be available and responsive at all times for Oracle Clusterware to operate.” which means to survive from loss of ‘N’ voting disks, you must configure atleast ‘2N+1’ voting disks.

Suppose you have 5 voting disks configured for your 2 Node environment, then you can survive even after loss of 2 voting disks.

Keep in mind that, having multiple voting disks is reasonable if you keep them on different disks/volumes/san arrays so that your cluster can survive even during the loss of one disk/volume/array. So, there is no point in configuring multiple voting disks on a single disk/lun/array.

But there is a special scenario, where all the nodes in the cluster can see the all voting disks but the cluster-interconnect between the nodes failed, to avoid split-brain syndrome in this scenario, node eviction must happen.

But the question here is which one?

According to Oracle – “The node with the lower node number will survive the eviction (The first node to join the cluster)”. So, the very first one that joined in the cluster will survive from eviction.

Note: Well in a 2 node setup, the split brain concept doesn't really work. In 2 node rac I have seen that if there is any real problem with the instance/server, node eviction happen. As there are no node groups formed, as there are only 2 nodes. If one node is unable to contact the other it will evict the node, which ever node find out first, will evict the other node. **(Refer Note ID: 219361.1)**

In order to avoid split brain syndrome, any node should be able to access at least 1/2 +1 of the total vote disks created

23. What are various IPs used in RAC? Or How many IPs we need in RAC?

Public IP, Private IP, Virtual IP, SCAN IP

24. What is the use of SCAN IP (SCAN name) and will it provide load balancing?

Single Client Access Name (SCAN) is a new Oracle Real Application Clusters (RAC) 11g Release 2, feature that provides a single name for clients to access an Oracle Database running in a cluster. The benefit is clients using SCAN do not need to change if you add or remove nodes in the cluster.

25. How many SCAN listeners will be running?

Three SCAN listeners only.

26. What is FAN?

Applications can use Fast Application Notification (FAN) to enable rapid failure detection, balancing of connection pools after failures, and re-balancing of connection pools when failed components are repaired. The FAN process uses system events that Oracle publishes when cluster servers become unreachable or if network interfaces fail. FAN is a notification mechanism that RAC uses to notify other processes about configuration and service level information such as includes service status changes, such as UP or DOWN events. Applications can respond to FAN events and take immediate action. FAN UP and DOWN events can apply to instances, services, and nodes. RAC publishes the FAN events the minute any changes are made to the cluster. So, instead of waiting for the application to check on individual nodes to detect an anomaly, the applications are notified by FAN events and are able to react immediately.

This feature will allow application to publish events related to node, instance or any service up/down events and event are published using ONS (RAC component) and Advanced Queues. ONS will send notification to application and Load balance advisory framework to change goodness and delta value for particular instance. FAN cleans up any connections when the failure occurs. It keeps track of service and instance for each connection.

When down event is published by ONS in RAC environment:

- a. Routes new requests to remaining instances
- b. Throws exceptions if applications are in middle of transactions. As mentioned earlier, DML is not supported for failover so applications should have capability to deal with DML transaction and throw appropriate errors or exceptions to the users.

When up event is published by ONS in RAC environment:

- a. Creates new connections to remaining instances
- b. Distributes new requests evenly to all instances based on Node load and instance load

c. Updates Load Balance advisory

27. What is FCF?

Fast Connection Failover provides high availability to FAN integrated clients, such as clients that use JDBC, OCI, or ODP.NET. If you configure the client to use fast connection failover, then the client automatically subscribes to FAN events and can react to database UP and DOWN events. In response, Oracle gives the client a connection to an active instance that provides the requested database service.

Fast Connection Failover is a feature of Oracle clients that have integrated with FAN HA Events.

Oracle JDBC Implicit Connection Cache, Oracle Call Interface (OCI), and Oracle Data Provider for .Net (ODP.Net) include fast connection failover.

With fast connection failover, when a down event is received, cached connections affected by the down event are immediately marked invalid and cleaned up.

FCF is an application-level failover mechanism: the database-tier notifies the application-tier by means of a FAN (fast application notification) message, distributed via the ONS (Oracle Notification Service) daemons. Where database connections are pooled, as offered by Oracle JDBC, the driver then has the opportunity to clean up stale connections if a node fails. One very nice feature is that when a node comes back online (i.e. a FAN message is sent out) the JDBC driver will automatically make new connections to the node to re-balance the pool.

28. What is TAF and TAF policies?

Explanation-1:

TAF is a database session-level connection failover mechanism

It works via the OCI interface only so applies to applications using thick JDBC drivers and those using the OCI libraries (e.g. written in C, like OID). During failover a new session will be started on an alternative node (though it can be pre-connected ahead of the failure) and this can optionally re-open a cursor and advance it to the point it was at prior to failure (providing underlying data hasn't changed).

Transparent Application Failover (TAF) - A runtime failover for high availability environments, such as Real Application Clusters and Oracle Real Application Clusters Guard, TAF refers to the failover and re-establishment of application-to-service connections. It enables client applications to automatically reconnect to the database if the connection fails, and optionally resume a SELECT statement that was in progress. This reconnect happens automatically from within the Oracle Call Interface (OCI) library.

Transparent Application Failover (TAF) is a feature of the Oracle Call Interface (OCI) driver at client side. It enables the application to automatically reconnect to a database, if the database instance to which the connection is made fails. In this case, the active transactions roll back.

Tnsnames Parameter: FAILOVER_MODE

e.g (failover_mode=(type=select)(method=basic))

Failover Mode Type can be Either SESSION or SELECT.

Session failover will have just the session to failed over to the next available node. With SELECT, the select query will be resumed.

TAF can be configured with just server side service settings by using dbms_service package.

TAF Policy:

1-None: Don't use TAF

2-Basic: Establish connection at failover time only

3-Pre-connect: Establish connection to both preferred instance and backup instance.

TAF Types:

1- Select: Oracle Net keeps track of all SELECT statements. Tracking how many rows have been fetched back to the client for each cursor associated with a SELECT statement. If the connection to the instance is lost, Oracle Net establishes a connection to another Oracle RAC node and re-executes the SELECT statements, repositioning the cursors so the client can continue fetching rows as if nothing has happened. The SELECT failover approach is best for data warehouse systems that perform complex and time-consuming transactions.

2- Session: When the connection to an instance is lost, SESSION failover results only in the establishment of a new connection to another Oracle RAC node; any work in progress is lost. SESSION failover is ideal for online transaction processing (OLTP) systems, where transactions are small.

Explanation-2:

Transparent Application Failover (TAF) - A runtime failover for high availability environments, such as Real Application Clusters and Oracle Real Application Clusters Guard, TAF refers to the failover and re-establishment of application-to-

service connections. It enables client applications to automatically reconnect to the database if the connection fails, and optionally resume a SELECT statement that was in progress. This reconnect happens automatically from within the Oracle Call Interface (OCI) library.

```
SUM1111DB =  
(DESCRIPTION =  
  (ADDRESS_LIST =  
    (LOAD_BALANCE=ON)  
    (FAILOVER=ON)  
    (ADDRESS=(PROTOCOL=TCP)(HOST=ORADB3)(PORT=1521))  
    (ADDRESS=(PROTOCOL=TCP)(HOST=ORADB4)(PORT=1521))  
  )  
(CONNECT_DATA =  
  (SERVICE_NAME = SUMSKYDB.SUMMERSKYUS.COM)  
  (FAILOVER_MODE = (TYPE=SELECT)(METHOD=BASIC)  
  )  
)  
)
```

29. How will you upgrade RAC database?

Refer Document

30. What are rolling patches and how to apply in RAC?

Types or Rolling Patches:

Rolling patch with a shared home: We had to apply yet another opatch, but we are only allowed 4 hours of downtime per system per month and we already used our monthly budget, so we need a way to apply an opatch without any downtime.

On some of our systems, it is not an issue. They are RAC systems where each node has its own \$ORACLE_HOME on its own server. We take one node down, apply the patch, start the node, stop the other node, apply the patch, and start other node. Patch installed on both nodes, no downtime for our customers. Win-Win.

But what do we do about our other systems?

The ones which share a single \$ORACLE_HOME on a filer? Where we need to take both nodes down for applying the patch?

A co-worker came up with a brilliant idea:

Stop one node. Use the filer power to duplicate \$ORACLE_HOME. Connect node to new home, just make the change in /etc/fstab, the database will never notice the difference.

Apply patch in new home. Start database in new home. Now stop the second node and connect it to the new home as well. Start the node in the new home. We have a patched DB with no downtime in a shared home system! We even have a built in rollback – connect one node after the other back to the old home, where we didn't apply the patch. In my experience rollback of opatches don't always work, so having a sure rollback plan is a great bonus.

We tested it today in a staging environment and it seems to work well. Now we just need to convince management that we should do it in production. It looks like a great solution, but in my experience management hates approving any plan that does not appear in Oracle manuals. For all their talk of innovation and "thinking outside the box" they are a very conservative bunch. I can understand the extreme risk aversion of IT management, but if you never do anything new, you can never improve, and thats also risky.

31. How to add/remove a node?

Refer Document

32. What are node apps?

VIP, listener, ONS, GSD

33. What is gsd (Global Service Daemon)?

A component that receives requests from SRVCTL to execute administrative job tasks, such as startup or shutdown. The command is executed locally on each node, and the results are returned to SRVCTL. GSD is installed on the nodes by default.

1. The GSD daemon has been replaced with the CRS in Oracle Database 10g
2. GSD is not mandatory for 10G RAC to function properly.
3. It is present only for backward compatibility with 9i RAC.
4. Also going forward with 11G RAC , GSD is disabled by default and it is not started.

34. How to do load balancing in RAC? What is Client balancing and server side balancing?

Explanation-1:

Client Side Connect-Time Load Balance: The client load balancing feature enables clients to randomize connection requests among the listeners.

This is done by client TNSNAMES Parameter: **LOAD_BALANCE**.

The (load_balance=yes) instructs SQLNet to progress through the list of listener addresses in the address_list section of the net service name in a random sequence. When set to OFF, instructs SQLNet to try the addresses sequentially until one succeeds.

Client Side Connect-Time failover:

This is done by client TNSNAMES Parameter: **FAILOVER**

The (failover=on) enables clients to connect to another listener if the initial connection to the first listener fails.

Without connect-time failover, Oracle Net attempts a connection with only one listener.

Server Side Listener Connection Load Balancing:

With server-side load balancing, the listener directs a connection request to the best instance currently providing the service.

INIT parameter **REMOTE_LISTENER** should be set. When set, each instance registers with the TNS listeners running on all nodes within the cluster.

There are two types of server-side load balancing:

Load Based — Server side load balancing redirects connections by default depending on node load. This is default.

Session Based — Session based load balancing takes into account the number of sessions connected to each node and then distributes the connections to balance the number of sessions across the different nodes.

From 10g release 2 the service can be setup to use load balancing advisory. This means connections can be routed using **SERVICE TIME** and **THROUGHPUT**. Connection load balancing means the goal of a service can be changed, to reflect the type of connections using the service.

Explanation-2:

Client-Side Connection Load Balancing: This load balancing method has been available since Oracle 8i. When a user session attempts to connect to the database, the database's listener will assign the session randomly to one of the listed multiple listening endpoints. [Listing 1](#) shows an example of the **TNSNAMES.ORA** network configuration file entries for an alias named **CSLB** that uses the **LOAD_BALANCING=ON** directive to implement this load balancing feature.

While this load balancing method is certainly simplest to implement, it also has an obvious limitation: *the listener has no idea if the session has been assigned to an endpoint whose corresponding database server is already overloaded.*

Moreover, since the listener is essentially picking the connection completely at random, there is no guarantee that the connection chosen will even be available at that time. This may force the session to wait for a considerable length of time – perhaps even minutes, a relative eternity in computing timeframes! – until the operating system indicates to the listener that the connection is unavailable, which causes the user session to fail eventually with an **ORA-01034**

ORACLE not available error.

Client-Side Connection Failover: Obviously, balancing the overall number of connections is a desirable goal, but what happens if the chosen connection point is unavailable because the database server's listener is no longer active? To forestall this, Oracle 8i provided the capability to determine if the connection that has been chosen at random is still "alive" and, if not, continue to try the other listed connection points until a live connection is found. This simple method tends to limit the chance of a lost connection, but unfortunately it must rely on TCP/IP timeouts to determine if a connection is alive or dead, and this means that an application may wait several seconds (or even longer!) before it receives a notification that the connection has been terminated.

I've laid out the **TNSNAMES.ORA** entries to activate client-side connection failover in [Listing 2](#). They are almost identical to Listing 1 with the notable exception of one more directive: **FAILOVER=ON**.

Server-Side Load Balancing:

The two previous methods will adequately handle the distribution of user sessions across available resources while helping to guarantee that no session will wait an excessive time to find a currently active address on which to connect. Clearly, a better solution was needed, and Oracle 9i offered one: *server-side load balancing*. This method divides the connection load evenly between all available listeners by determining the total number of connections on each listener, and then distributing new user session connection requests to the least loaded listener(s) based on the total number of sessions already connected. While a bit more complex to implement because it requires

configuration of multiple listeners, it most definitely helps to even out connections across all available listeners in a database system.

To implement server-side load balancing, at least two listeners must be configured. Also, the **REMOTE_LISTENERS** initialization parameter must be added to the database's **PFILE** or **SPFILE** so that the database knows to search out the value provided in that parameter in the database server's **TNSNAMES.ORA** configuration file. When server-side load balancing is activated, each listener that contributes a listening endpoint communicates with the other listener(s) via each database instance's PMON process. Oracle then determines how many user connections each listener is servicing, and it will distribute any new connection requests so that the load is balanced evenly across all servers. The entries in **TNSNAMES.ORA** direct the listeners to share information about the relative load connectivity. As shown in [Listing 3](#), I've gathered these required changes to **TNSNAMES.ORA**. I've also created a new PL/SQL package, **HR.PKG_LOAD_GENERATOR**, that incorporates three different methods to generate user sessions in an attempt to "overload" an Oracle 10gR2 database listener. A simple shell script, **RandomLoadGenerator.sh**, calls a few SQL command files that in turn make calls to the package's procedures and thus create a sample workload of approximately 40 connections against an Oracle 10gR2 database that contains the standard sample schemas.

Load Balancing In Oracle 10g Real Application Clusters Environments:

These three methods are actually quite effective for distribution of incoming connections evenly across multiple listeners in any single-instance database configuration. An Oracle 10gR2 Real Application Clusters (RAC) clustered database, on the other hand, needs more robust load balancing capabilities because of the nature of that environment.

A RAC clustered database comprises at least two (and usually many more) *nodes*, each running a separate *instance* of the clustered database. In addition, a RAC database usually needs to supply a *minimum amount* of connections and resources to *several applications*, each with dramatically different *resource needs* depending on the current business processing cycle(s), so the application load that's placed on each instance in the clustered database therefore can be dramatically different at different times of the day, week, month, and year. Finally, it's likely that a RAC clustered database will need to guarantee a minimum *cardinality* (i.e. a specific number of nodes on which the application needs to run at all times) to one or more mission-critical applications.

RAC Services: Starting in Oracle 8i, an Oracle database could dynamically register a database directly with its corresponding listener(s) based on the settings for the **SERVICE_NAMES** initialization parameter through the database's Process Monitor (**PMON**) background process. To completely support this feature, Oracle strongly suggested that the **SERVICE_NAME** parameter should be used instead of the original **SID** parameter in the **TNSNAMES.ORA** configuration file so that an incoming user session could immediately identify the database instance to which a session intended to connect.

Oracle 10g RAC leverages this service naming feature to distribute application connections efficiently across a RAC clustered database. For example, a clustered database may need to support three different applications, **OLTP**, **DSS**, and **ADHOC**. The **OLTP** application is the main order entry application for this enterprise computing environment, and therefore it needs a minimum *cardinality* of two cluster database instances at all times. The **DSS** application, on the other hand, supports extraction, transformation and load (ETL) operations for the enterprise's data warehouse, and thus it requires a minimum cardinality of just one instance. Likewise, the **ADHOC** application supports OLAP and general user query execution against the data warehouse, but it too only requires a minimum cardinality of a single instance.

Oracle 10gR2 RAC: Server-Side Connect-Time Load Balancing:

To demonstrate the implementation of load balancing features in a RAC environment, I'll use a relatively straightforward testing platform: a simple two-node RAC clustered database, **RACDB**, with two instances, **RACDB1** and **RACDB2**, configured on two nodes (**RACLinux1** and **RACLinux2**, respectively). I've set up this configuration using two VMWare Virtual Machines, each running CentOS Linux Enterprise Server 3 Release 8 (kernel 2.4.21-40) as the guest configuration.

[Listing 4](#) shows the **SRVCTL** commands I've issued to create and start three new application services, **OLTP**, **DSS**, and **ADHOC**, on the **RACDB** clustered database. Note that I've specified both the **RACDB1** and **RACDB2** instances as the *preferred* instances for all three applications. When I execute **SRVCTL** commands to create these services, Oracle 10g automatically adds these service name values to the **SERVICE_NAMES** parameter for each instance in the cluster database.

I've also set up three application aliases for these applications in the client **TNSNAMES.ORA** entries configuration files. Note that I've specified the **SERVICE_NAME** parameter as **RACDB** so that all nodes in the cluster can participate in distributing the load of these applications across the cluster. I've also used the two nodes' *virtual IP addresses* (**raclinux1-vip** and **raclinux2-vip**) as the connection points for these services. This guarantees that if any one of the

listeners or instances servicing these applications should fail, ONS will automatically relocate any new connection requests to a new listener alias on another surviving node.

To complete the configuration of *server-side connection load balancing* for this RAC clustered database, note that I've set the ***.REMOTE_LISTENERS=RACDB_LISTENERS** initialization parameter in the database's shared SPFILE. I've also added a corresponding **RACDB_LISTENERS** entry to the TNSNAMES.ORA file in the Oracle Home of each node in the cluster. Each database's PMON process will now automatically register the database with the database's local listener as well as *cross-register* the database with the listeners on all other nodes in the cluster. In this mode, the nodes *themselves* decide which node is least busy, and then will connect the client to that node.

It's also important to realize that in a RAC environment, the server-side load balancing methodology differs slightly from the methodology used in a single-instance environment because Oracle 10gR2 discriminates whether the incoming connection has been requested as either a *dedicated* or a *shared server* connection:

- If a *dedicated* session is requested, then the listener will select the instance first on the basis of the *node that is least loaded*; if all nodes are *equally* loaded, it will then select the instance that has the least load.
- For a *shared server* connection, however, the listener goes one step further. It will also check to see if all of the available *instances* are equally loaded; if this is true, the listener will place the connection on the *least-loaded dispatcher* on the selected instance.

Example:

<http://www.databasejournal.com/features/oracle/article.php/3666396/Oracle-10gR2-RAC-Load-Balancing-Features.htm>

<http://www.orafaq.com/node/1840>

```
/*
|| Oracle 10gR2 RAC LBA Features Listing
||
|| Demonstrates Oracle 10gR2 Load Balancing Advisory (LBA) features for
|| Real Application Clusters, including:
|| - How to set up client-side load balancing and failover
|| - How to set up server-side load balancing
|| - How to set up Load Balancing Advisory features
|| - How to monitor the efficiency and outcomes of the Load Balancing Advisory
||
|| Author: Jim Czuprynski
||
|| Usage Notes:
|| This script is provided to demonstrate various features of Oracle 10gR2
|| Load Balancing Advisor, and it should be carefully proofread before
|| executing it against any existing Oracle database to insure that no
|| potential damage can occur.
*/
/*
|| Listing 1: Setting up client-side connection load balancing
*/
#####
# Add these entries to each client's TNSNAMES.ORA configuration file
# to enable Client-Side Load Balancing ONLY (i.e., no failover)
#####
CSLB_ONLY =
  (DESCRIPTION =
    (LOAD_BALANCE = ON)
    (ADDRESS = (PROTOCOL = TCP)(HOST = raclinux1-vip)(PORT = 1521))
    (ADDRESS = (PROTOCOL = TCP)(HOST = raclinux2-vip)(PORT = 1521))
```

```

(CONNECT_DATA =
  (SERVER = DEDICATED)
  (SERVICE_NAME = racdb)
)
)
/*
|| Listing 2: Setting up client-side connection load balancing plus failover
*/
#####
# Add these entries to each client's TNSNAMES.ORA configuration file
# to enable Client-Side Load Balancing PLUS Failover
#####
CSLB_FAILOVER =
  (DESCRIPTION =
    (ADDRESS = (PROTOCOL = TCP)(HOST = raclinux1-vip)(PORT = 1521))
    (ADDRESS = (PROTOCOL = TCP)(HOST = raclinux2-vip)(PORT = 1521))
    (LOAD_BALANCE = ON) # Activates load balancing
    (FAILOVER = ON) # Activates failover
    (CONNECT_DATA =
      (SERVER = DEDICATED)
      (SERVICE_NAME = racdb)
    )
  )
)

/*
|| Listing 3: Setting up server-side connection load balancing features.
|| Note that server-side load balancing requires:
|| 1.) New entries in every client's TNSNAMES.ORA file for the new alias
|| 2.) New entries in the TNSNAMES.ORA file of every node in the cluster
||    to include the REMOTE_LISTENER setting
|| 3.) The addition of *.REMOTE_LISTENER parameter to all nodes in cluster
||    to force each node's Listener to register with each other
*/
#####
# Add these entries to each server's TNSNAMES.ORA file to enable Server-Side
# Load Balancing:
#####
LISTENERS_RACDB =
  (ADDRESS_LIST =
    (ADDRESS = (PROTOCOL = TCP)(HOST = raclinux1-vip)(PORT = 1521))
    (ADDRESS = (PROTOCOL = TCP)(HOST = raclinux2-vip)(PORT = 1521))
  )
)
SSLB =
  (DESCRIPTION =
    (ADDRESS = (PROTOCOL = TCP)(HOST = raclinux1-vip)(PORT = 1521))
    (ADDRESS = (PROTOCOL = TCP)(HOST = raclinux2-vip)(PORT = 1521))
    (LOAD_BALANCE = ON)
    (FAILOVER = ON)
    (CONNECT_DATA =
      (SERVER = DEDICATED)
      (SERVICE_NAME = racdb)
    )
  )
)
)
-----
-- Run this command to add the REMOTE_LISTENERS initialization parameter to

```

```

-- the common SPFILE for all nodes in the RAC clustered database:
-----
ALTER SYSTEM SET REMOTE_LISTENER = LISTENERS_RACDB SID='*' SCOPE=BOTH;
/*
|| Listing 4: Setting up Load Balancing Advisory features in an Oracle 10g
||      Real Applications Cluster (RAC) clustered database environment
*/
#####
# Create, register, and start three new services with
# Cluster-Ready Services
#####
srvctl add service -d racdb -s ADHOC -r racdb1,racdb2
srvctl start service -d racdb -s ADHOC
srvctl add service -d racdb -s DSS -r racdb1,racdb2
srvctl start service -d racdb -s DSS
srvctl add service -d racdb -s OLTP -r racdb1,racdb2
srvctl start service -d racdb -s OLTP
/*
|| Listing 5: Using DBMS_SERVICE.MODIFY_SERVICE to configure RAC services
||      to use Load Balancing Advisory features in an Oracle 10g
||      Real Applications Cluster (RAC) clustered database environment
*/
-----
-- Configuring existing RAC services to use the Load Balancing Advisory:
-- 1.) ADHOC: No Load Balancing Advisory
-- 2.) DSS: Load Balancing Advisory with Service Time goal
-- 3.) OLTP: Load Balancing Advisory with Throughput goal
-- Note that Advanced Queueing (AQ) tracking is also activated.
-----
BEGIN
  DBMS_SERVICE.MODIFY_SERVICE(
    service_name => 'ADHOC'
    ,aq_ha_notifications => TRUE
    ,goal => DBMS_SERVICE.GOAL_NONE
    ,clb_goal => DBMS_SERVICE.CLB_GOAL_LONG
  );
  DBMS_SERVICE.MODIFY_SERVICE(
    service_name => 'DSS'
    ,aq_ha_notifications => TRUE
    ,goal => DBMS_SERVICE.GOAL_SERVICE_TIME
    ,clb_goal => DBMS_SERVICE.CLB_GOAL_SHORT
  );
  DBMS_SERVICE.MODIFY_SERVICE(
    service_name => 'OLTP'
    ,aq_ha_notifications => TRUE
    ,goal => DBMS_SERVICE.GOAL_THROUGHPUT
    ,clb_goal => DBMS_SERVICE.CLB_GOAL_SHORT
  );
END;
/
-----
-- Confirm these services' configuration by querying DBA_SERVICES:
-----
SET PAGESIZE 50
SET LINESIZE 110

```



```

TTITLE 'Services Configured to Use Load Balancing Advisory (LBA) Features|
(From DBA_SERVICES)'
COL name          FORMAT A16   HEADING 'Service Name' WRAP
COL created_on    FORMAT A20   HEADING 'Created On' WRAP
COL goal          FORMAT A12   HEADING 'Service|Workload|Management|Goal'
COL clb_goal      FORMAT A12   HEADING 'Connection|Load|Balancing|Goal'
COL aq_ha_notifications FORMAT A16 HEADING 'Advanced|Queueing|High-|Availability|Notification'
SELECT
    name
    ,TO_CHAR(creation_date, 'mm-dd-yyyy hh24:mi:ss') created_on
    ,goal
    ,clb_goal
    ,aq_ha_notifications
FROM dba_services
WHERE goal IS NOT NULL
    AND name NOT LIKE 'SYS%'
ORDER BY name
;
TTITLE OFF
/*
|| Listing 6: Using the GV$SERVICEMETRIC global view to track how RAC
||          services are responding to the Load Balancing Advisor
*/
TTITLE 'Current Service-Level Metrics|(From GV$SERVICEMETRIC)'
BREAK ON service_name NODUPPLICATES
COL service_name  FORMAT A08   HEADING 'Service|Name' WRAP
COL inst_id       FORMAT 9999   HEADING 'Inst|ID'
COL beg_hist      FORMAT A10    HEADING 'Start Time' WRAP
COL end_hist      FORMAT A10    HEADING 'End Time' WRAP
COL intsize_csec  FORMAT 9999   HEADING 'Intvl|Size|(cs)'
COL goodness      FORMAT 999999 HEADING 'Good|ness'
COL delta         FORMAT 999999 HEADING 'Pred-|icted|Good-|ness|Incr'
COL cpuperccall   FORMAT 9999999 HEADING 'CPU|Time|Per|Call|(mus)'
COL dbtimepercall FORMAT 9999999 HEADING 'Elpsd|Time|Per|Call|(mus)'
COL callspersec   FORMAT 9999999 HEADING '# Of|User|Calls|Per|Second'
COL dbtimepersec  FORMAT 9999999 HEADING 'DBTime|Per|Second'
COL flags         FORMAT 999999 HEADING 'Flags'
SELECT
    service_name
    ,TO_CHAR(begin_time, 'hh24:mi:ss') beg_hist
    ,TO_CHAR(end_time, 'hh24:mi:ss') end_hist
    ,inst_id
    ,goodness
    ,delta
    ,flags
    ,cpuperccall
    ,dbtimepercall
    ,callspersec
    ,dbtimepersec
FROM gv$servicemetric
WHERE service_name IN ('OLTP', 'DSS', 'ADHOC')
ORDER BY service_name, begin_time DESC, inst_id
;
CLEAR BREAKS
TTITLE OFF

```

35. What are the uses of services? How to find out the services in cluster?

Applications should use the services to connect to the Oracle database. Services define rules and characteristics (unique name, workload balancing, failover options, and high availability) to control how users and applications connect to database instances.

36. How to find out the nodes in cluster (or) how to find out the master node?

`olsnodes` -- Which ever displayed first, is the master node of the cluster.

Select MASTER_NODE from V\$GES_RESOURCE;

To find out which is the master node, you can see `ocssd.log` file and search for "master node number".

37. How to know the public IPs, private IPs, VIPs in RAC?

`olsnodes -n -p -i`

```
node1-pub  1  node1-prv  node1-vip
node2-pub  2  node2-prv  node2-vip
```

38. What utility is used to start DB/instance?

`srvctl start database -d database_name`

`srvctl start instance -d database_name -i instance_name`

39. How can you shutdown single instance?

Change `cluster_database=false`

`srvctl stop instance -d database_name -i instance_name`

40. What is HAS (High Availability Service) and the commands?

HAS includes ASM & database instance and listeners.

`crsctl check has`

`crsctl config has`

`crsctl disable has`

`crsctl enable has`

`crsctl query has releaseversion`

`crsctl query has softwareversion`

`crsctl start has`

`crsctl stop has [-f]`

41. How many nodes are supported in a RAC Database?

10g Release 2, support 100 nodes in a cluster using Oracle Clusterware, and 100 instances in a RAC database.

With 10g Release 2, we support 100 nodes in a cluster using Oracle Clusterware, and 100 instances in a RAC database. Currently DBCA has a bug where it will not go beyond 63 instances. There is also a documentation bug for the `max-instances` parameter. With 10g Release 1 the Maximum is 63.

42. What is fencing?

I/O fencing prevents updates by failed instances, and detecting failure and preventing split brain in cluster. When a cluster node fails, the failed node needs to be fenced off from all the shared disk devices or disk groups. This methodology is called I/O Fencing, sometimes called Disk Fencing or failure fencing.

Nodes in a RAC cluster can fall victim to conditions called Split Brain and Amnesia. These conditions usually occur from a temporary network disconnect. Because of disconnect, the "sick" node thinks it is the only node in the cluster, and forms its own "sub cluster" consisting only of itself.

In this case, the cluster needs to correct the issue. Traditional clusters use a process called STONITH (Shoot the Other Node in the Head) in order to correct the issue; this simply means the healthy nodes kill the sick node. Oracle's Clusterware does not do this; instead, it simply gives the message "Please Reboot" to the sick node. The node bounces itself and rejoins the cluster.

There are other methods of fencing that are utilized by different hardware/software vendors. When using Veritas Storage Foundation for RAC (VxSF RAC), you can implement I/O fencing instead of node fencing. This means that instead of asking a server to reboot, you simply close it off from shared storage.

43. Why Clusterware installed in root (why not oracle)?

Need document

44. What are the wait events in RAC? Differences in Oracle RAC wait events?

`gc current block 2-way`

`gc current block 3-way`

`gc current block busy`

`gc current buffer busy`

`gc current block congested`

gc current block 2-way:

An instance requests authorization for a block to be accessed in current mode to modify a block, the instance mastering the resource receives the request. The master has the current version of the block and sends the current copy of the block to the requestor via Cache Fusion and keeps a Past Image (.PI)

If you get this then do the following

- Analyze the contention, segments in the "current blocks received" section of AWR
- Use application partitioning scheme
- Make sure the system has enough CPU power
- Make sure the interconnect is as fast as possible
- Ensure that socket send and receive buffers are configured correctly

gc current block 3-way:

An instance requests authorization for a block to be accessed in current mode to modify a block, the instance mastering the resource receives the request and forwards it to the current holder of the block, asking it to relinquish ownership. The holding instance sends a copy of the current version of the block to the requestor via Cache Fusion and transfers the exclusive lock to the requesting instance. It also keeps a past Image (PI).

Use the above actions to increase the performance

gc current block busy:

The requestor will eventually get the block via cache fusion but it is delayed due to one of the following

- The block was being used by another session on another session
- was delayed as the holding instance could not write the corresponding redo record immediately

If you get this then do the following

- Ensure the log writer is tuned

gc current buffer busy:

This is the same as above (gc current block busy), the difference is that another session on the same instance also has requested the block (hence local contention)

gc current block congested:

This is caused if heavy congestion on the GCS, thus CPU resources is stretched

Differences in Oracle RAC wait events?

Oracle RAC is somewhat of a unique case of an Oracle environment, but everything learned about wait events in the single instance database also applies to clustered databases. However, the special use of a global buffer cache in RAC makes it imperative to monitor inter-instance communication via the cluster-specific wait events such as gc cr request and gc buffer busy. Understanding these wait events will help in the diagnosis of problems and pinpointing solutions in a RAC database.

Focus on the buffer cache and its operations:

The main difference to keep in mind when monitoring a RAC database versus a single-instance database is the buffer cache and its operation. In a RAC environment, the buffer cache is global across all instances in the cluster and hence the processing differs. When a process in a RAC database needs to modify or read data, Oracle will first check to see if it already exists in the local buffer cache. If the data is not in the local buffer cache the global buffer cache will be reviewed to see if another instance already has it in their buffer cache. In this case the remote instance will send the data to the local instance via the high-speed interconnect, thus avoiding a disk read.

Monitoring an Oracle RAC database often means monitoring this situation and the amount of requests going back and forth over the RAC interconnect. The most common wait events related to this are gc cr request and gc buffer busy (note that in Oracle RAC 9i and earlier these wait events were known as "global cache cr request " and "global cache buffer busy" wait events).

gc cr request

The gc cr request wait event specifies the time it takes to retrieve the data from the remote cache. In Oracle 9i and prior, gc cr request was known as global cache cr request. High wait times for this wait event often are because of: RAC Traffic Using Slow Connection - typically RAC traffic should use a high-speed interconnect to transfer data between instances, however, sometimes Oracle may not pick the correct connection and instead route traffic over the slower public network. This will significantly increase the amount of wait time for the gc cr request event. The oradebug command can be used to verify which network is being used for RAC traffic:

```
SQL> oradebug setmypid
SQL> oradebug ipc
```

This will dump a trace file to the location specified by the user_dump_dest Oracle parameter containing information about the network and protocols being used for the RAC interconnect.

Inefficient Queries - poorly tuned queries will increase the amount of data blocks requested by an Oracle session. The more blocks requested typically means the more often a block will need to be read from a remote instance via the interconnect.

gc buffer busy acquire and gc buffer busy release

The gc buffer busy acquire and gc buffer busy release wait events specify the time the remote instance locally spends accessing the requested data block. In Oracle 11g you will see gc buffer busy acquire wait event when the global cache open request originated from the local instance and gc buffer busy release when the open request originated from a remote instance. In Oracle 10g these two wait events were represented in a single gc buffer busy wait, and in Oracle 9i and prior the "gc" was spelled out as "global cache" in the global cache buffer busy wait event. These wait events are all very similar to the buffer busy wait events in a single-instance database and are often the result of:

Hot Blocks - multiple sessions may be requesting a block that is either not in buffer cache or is in an incompatible mode. Deleting some of the hot rows and re-inserting them back into the table may alleviate the problem. Most of the time the rows will be placed into a different block and reduce contention on the block. The DBA may also need to adjust the pctfree and/or pctused parameters for the table to ensure the rows are placed into a different block.

Inefficient Queries - as with the gc cr request wait event, the more blocks requested from the buffer cache the more likelihood of a session having to wait for other sessions. Tuning queries to access fewer blocks will often result in less contention for the same block.

Buffer busy global cache:

This wait event falls under the umbrella of 'global buffer busy events'. This wait event occurs when a user is waiting for a block that is currently held by another session on the same instance and the blocking session is itself waiting on a global cache transfer.

Buffer busy global CR:

This wait event falls under the umbrella of 'global buffer busy events'. This wait event occurs when multiple CR requests for the same block are submitted from the same instance before the first request completes, users may queue up behind it

Global cache busy:

This wait event falls under the umbrella of 'global buffer busy events'. This wait event means that a user on the local instance attempts to acquire a block globally and a pending acquisition or release is already in progress.

Global cache cr request:

this wait event falls under the umbrella of 'global cache events'. This wait event determines that an instance has requested a consistent read version of a block from another instance and is waiting for the block to arrive.

Global cache null to s and global cache null to x:

This wait event falls under the umbrella of 'global cache events'. These events are waited for when a block was used by an instance, transferred to another instance, and then requested back again.

Global cache open s and global cache open x:

This wait event falls under the umbrella of 'global cache events'. These events are used when an instance has to read a block from disk into cache as the block does not exist in any instances cache. High values on these waits may be indicative of a small buffer cache, therefore you may see a low cache hit ratio for your buffer cache at the same time as seeing these wait events.

Global cache s to x:

This wait event falls under the umbrella of 'global cache events'. This event occurs when a session converts a block

from shared to exclusive mode.

45. What is the difference between cr block and cur (current) block?

The current block contains changes for all the committed and yet-to-be-committed transactions. A consistent read (CR) block represents a consistent snapshot of the data from a previous point in time. Applying undo/rollback segment information produces consistent read versions. Thus, a single data block can reside in many buffer caches under shared resources with different versions.

Multi-version data blocks help to achieve read consistency. The read consistency model guarantees that the data block seen by a statement is consistent with respect to a single point in time and does not change during the statement execution. Readers of data do not wait for other writer's data or for other readers of the same data. At the same time, writers do not wait for other readers for the same data. Only writers wait for other writers if they attempt to write. As mentioned earlier, the undo (rollback) segment provides the required information to construct the read-consistent data blocks. In case of a multi-instance system, like the RAC database, the requirement for the same data block may arise from another instance. To support this type of requirement, past images of the data blocks are created within the buffer cache.

46. Why Node Eviction happens on Oracle RAC?

Oracle Clusterware evicts the node when following condition occur:

- Node is not pinging via the network heartbeat
- Node is not pinging the Voting Disk
- Node is hung or busy and is unable to perform the above two tasks

Most cases the error cause is written to disk. If no error following the Metalink note: ID 559365.1 to use Diagwait option which will give 10 seconds for the node to write logs to error log file.

```
#crsctl set css diagwait 13 -force
```

```
#crsctl get css diagwait
```

```
#crsctl check crs
```

```
#crsctl unset css diagwait -f
```

47. What are the initialization parameters that must have same value for every instance in an Oracle RAC database?

- ACTIVE_INSTANCE_COUNT
- ARCHIVE_LAG_TARGET
- CLUSTER_DATABASE
- CLUSTER_DATABASE_INSTANCES
- CONTROL_FILES
- DB_BLOCK_SIZE
- DB_DOMAIN
- DB_FILES
- DB_NAME
- DB_RECOVERY_FILE_DEST
- DB_RECOVERY_FILE_DEST_SIZE
- MAX_COMMIT_PROPAGATION_DELAY
- TRACE_ENABLED
- UNDO_MANAGEMENT

48. What is Miscount (MC) in Oracle RAC?

The Cluster Synchronization Service (CSS) on RAC has Miscount parameter. This value represents maximum time, in seconds, that a network heartbeat can be missed before entering into a cluster reconfiguration to evict the node. The default value is 30 seconds (Linux 60 seconds in 10g, 30 sec in 11g).

49. What is the use of CSS Heartbeat Mechanism in Oracle RAC? (Metalink Note: 294430.1)

The CSS of the Oracle Clusterware maintains two heartbeat mechanisms

1. The disk heartbeat to the voting device and
2. The network heartbeat across the interconnect (This establishes and confirms valid node membership in the cluster).

Both of these heartbeat mechanisms have an associated timeout value. The disk heartbeat has an internal i/o timeout interval (DTO Disk TimeOut), in seconds, where an i/o to the voting disk must complete. The miscount parameter (MC), as stated above, is the maximum time, in seconds, that a network heartbeat can be missed. The disk heartbeat i/o timeout interval is directly related to the miscount parameter setting. The Disk TimeOut(DTO) = Miscount(MC) - 15 seconds (some versions are different).

50. What happens if latencies to voting disks are longer?

If I/O latencies to the voting disk are greater than the default Disk Time Out (DTO), then the cluster may experience CSS node evictions.

51. What is CSS misscount?

The CSS misscount represents the maximum seconds the network heartbeat can be missed before entering into cluster reconfiguration and evict the node. The default CSS misscount is 30 seconds. (only for 10g Linux it is 60 seconds).

52. How to change the CSS misscount default value? (Metalink Note: 284752.1)

1) Shut down CRS on all but one node. For exact steps use Note 309542.1

2) Execute crsctl as root to modify the misscount:

```
$ORA_CRS_HOME/bin/crsctl set css misscount
```

where is the maximum i/o latency to the voting disk +1 second

3) Reboot the node where adjustment was made

4) Start all other nodes shutdown in step 1

53. How to start and stop CRS?

Note: Typically the Oracle clusterware starts up automatically during startup.

```
cd /etc/init.d
```

```
init.crs stop
```

```
init.crs start
```

To disable crs to start during next reboot. It will not bring down running crs.

```
init.crs enable
```

```
init.crs disable
```

Start Oracle Clusterware

```
crsctl start crs
```

Stop Oracle Clusterware

```
crsctl stop crs
```

54. How to move regular DB to an ASM disk group?

The following are the steps involved in moving regular db files to ASM disk group.

Assume:

a. Oracle RAC instance is up already

b. DB name to be moved PROD

c. RAC db and normal DB both are in same instance.

1. Install and bring up Oracle RAC instance and ASM disk group.

2. Comment control file location in the DB you want to move and add ASM disk name for control_file.

```
ex. control_file="+DATA_GRP"
```

3. SQL> startup nomount

```
SQL> Show parameter => Control_files will show new disk grp
```

4. Use RMAN to move control file from regular disk to ASM using restores command.

```
rman> connect target
```

```
rman> restore controlfile from '/u01/oracle/PROD/cntrl01.ctl';
```

5. Verify using asmcmd

```
asmcmd> cd DSK_GRP/DATA_GRP/PROD
```

```
asmcmd> ls => you can see new controlfile under PROD directory.
```

6. Now mount the DB

```
sqlplus "/as sysdba"
```

```
sql> alter database mount;
```

7. Now use RMAN to move the data files.

```
rman
```

```
connect target (connected to PROD)
```

```
rman> backup as copy database format '+DATA_GRP';
```

Note: you can use asmcmd to monitor the data file movements to ASM.

8. rman> swith database to copy;

9. sqlplus "/as sysdba" ; alter database open;

```
select * from v$datafile;
```

```
select * from v$tempfile;
```

```
select * from v$controlfile;
```

select * from v\$logfile;

11. sql> alter database drop logfile '/u01/.../redo01.log';

alter database add logfile '+DATA_GRP';

Note: Repeat same step for all log files except current used logfile.

select * from v\$log to find which one is current

12. alter system switch logfile;

drop the first one which was being used.

13. Now vi init.ora and put full path for controlfile for DB to start properly.

*.control_file="+DATA_GRP/PROD/controlfile/current.333.433.3333"

13. vi init.ora => Change location of arc to ASM

*.log_archive_dest_1='LOCATION='+DATA_GRP/PROD' => if you omit PROD it will not work properly.

14. alter system switch logfile; => now the new arc will go to ASM.

15. END

55. What is a NIC card and HBA card?

Oracle RAC requires a NIC or HBA card which enables the computer to talk to network or to a storage subsystem.

There are different speeds of HBA card: 1Gbit/s, 2Gbit/s, 4, 8, 10, 20 Gbits/s

HBA has a unique World Wide Name (WWN),

which is similar to an Ethernet MAC address in that it uses an Organizationally

Unique Identifier (OUI) assigned by the IEEE.

56. What is a TPS?

http://www.dba-oracle.com/m_transactions_per_second.htm

57. What is the use of crs_getperm command?

Used to get permission information.

crs_getperm

Usage: crs_getperm resource_name [-u user|-g group] [-q]

crs_getperm ora.dudb.dudb1.inst

Name: ora.dudb.dudb1.inst

owner:oracle:rw,pgroup:oinstall:rw,other::r--,

58. What is the use of crs_profile?

Used to create, validate, delete and update a profile for RAC.

59. Where will you check for RAC log files?

- \$ORA_CRS_HOME/crs/log Contains trace files for the CRS resources.
- \$ORA_CRS_HOME/crs/init Contains trace files of the CRS daemon during startup. Good place to start with any CRS login problems.
- \$ORA_CRS_HOME/css/log The Cluster Synchronization (CSS) logs indicate all actions such as reconfigurations, missed check-ins, connects, and disconnects from the client CSS listener. In some cases, the logger logs messages with the category of auth.crit for the reboots done by Oracle. This could be used for checking the exact time when the reboot occurred.
- \$ORA_CRS_HOME/css/init Contains core dumps from the Oracle Cluster Synchronization Service daemon (OCSSD) and the process ID (PID) for the CSS daemon whose death is treated as fatal. If abnormal restarts for CSS exist, the core files will have the format of core..
- \$ORA_CRS_HOME/evm/log Log files for the Event Volume Manager (EVM) and evmlogger daemons. Not used as often for debugging as the CRS and CSS directories.
- \$ORA_CRS_HOME/evm/init PID and lock files for EVM. Core files for EVM should also be written here.
- \$ORA_CRS_HOME/srvm/log Log files for Oracle Cluster Registry (OCR), which contains the details at the Oracle cluster level.
- \$ORA_CRS_HOME//log Log files for Oracle Clusterware (known as the cluster alert log), which contains diagnostic messages at the Oracle cluster level. This is available from Oracle database 10g R2.

60. What is OCFS?

Explanation-1:

<http://decipherinfosys.wordpress.com/2007/11/12/ocfs-asm-raw-devices-and-regular-filesystem/>

Oracle Clustered File System, is a file system format that is proprietary to Oracle. This file system is designed with clustering in mind, hence the name. OCFS was designed by Oracle to make it possible for DBA's to run a RAC on a shared file system, without having to use RAW devices. There are other clustered file systems on the market, however Oracle offers OCFS at no cost. The initial version of OCFS (OCFS 1) targeted making clustered data storage easier to manage. Only database files, such as data files, control files, and redo log files (archived also) can be stored in these file system (in addition to the files the file system keeps to maintain the shared cluster storage). Oracle has now released OCFS2, which has been expanded to include the storage of Oracle binaries and scripts. OCFS is not included with every operating system. Certain OS's are now including OCFS as an option for file system formatting, but for others the OCFS software will need to be downloaded from Oracle.

Pros:

1. The file system was designed with Oracle clustering in mind and it is free.
2. Eliminates the need to use RAW devices or other expensive clustered file systems.
3. With the advent of OCFS2, binaries, scripts, and configuration files (shared Oracle home) can be stored in the file system. Making the management of RAC easier.

Cons:

1. With OCFS version 1, regular files cannot be stored in the file system, however this issue is eliminated with OCFS2.

Explanation-2:

Oracle Cluster File System (OCFS) presents a consistent file system image across the servers in a cluster. OCFS allows administrators to take advantage of a file system for the Oracle database files (data files, control files, and archive logs) and configuration files. This eases administration of the Oracle Real Application Clusters.

61. What is Oracle Cluster Ware?

a. It is a framework which contains application modeling logic.

Invokes application aware agents.

Performs resource recovery. When a node goes down, Clusterware framework recovers the application by relocating the resources to a live node.

This can be done for non-Oracle applications as well, for example. xclock.

b. Clusterware also hosts OCR cache.

The Oracle Clusterware requires two clusterware components:

a voting disk to record node membership information and the

Oracle Cluster Registry/Repository (OCR) to record cluster configuration information.

The voting disk and the OCR must reside on shared storage.

62. What is a resource?

A resource is a Oracle Clusterware manager application.

'Profile attributes' for a resource is stored in Oracle Cluster Registry.

63. How to register a resource?

a. Use crs_profile to create .CAP file with configuration details.

b. use crs_register to read .CAP file and update the OCR.

c. Resources can have dependencies. It will start in order and failover as a single unit.

64. What does crs_start / crs_stop do?

Reads config info from OCR and calls agent with command 'start'.

The agents (can be user written) actually stop the resource.

crs_start => read OCR config info => calls 'Control Agent' with command start. => Control agent stops the resource.

crs_stop => read OCR config info => call 'Control agent' with 'stop' => control agent stops app.

65. What is the difference between Oracle Clusterware and CRS?

Oracle Clusterware is formerly known as Cluster Ready Services (CRS). It is an integrated cluster management solution that enables you to link multiple servers so that they function as a single system or cluster. The Oracle Clusterware simplifies the infrastructure required for RAC because it is integrated with the Oracle Database. In addition, Oracle Clusterware is also available for use with single-instance databases and applications that you deploy

on clusters

Note: The commands stating with crs_ are still valid and same.

66. What is Oracle recommendation for interconnect?

Oracle recommends that you configure a redundant interconnect to prevent interconnect from being a single point of failure.

Oracle also recommends that you use User Datagram Protocol (UDP) on a Gigabit Ethernet for your cluster interconnects.

Crossover cables are not supported for use with Oracle Cluster ware or RAC databases.

67. List the commands used to manage RAC?

crs_profile, crs_register, crs_relocate, crs_getperm, crs_setperm, crs_stat, srvctl , crsctl

crsctl check crs

crsctl check cssd

crsctl check evmd

crsctl add css votedisk - adds a new voting disk

crsctl delete css votedisk - removes a voting disk

crsctl enable crs - enables startup for all CRS daemons

crsctl disable crs - disables startup for all CRS daemons

crsctl start crs - starts all CRS daemons.

crsctl stop crs - stops all CRS daemons. Stops CRS resources

crsctl query crs activeversion

CRS active version on the cluster is [10.2.0.1.0]

ocrdump

ocrconfig

ocrconfig -showbackup

ocrconfig -repair ocr

ocrconfig -replace

ocrconfig -export/-import

ocrconfig -upgrade

ocrcheck - no param needed.

ocrcheck

68. What are the log file locations for RAC?

cd \$ORACLE_HOME/log//client

-- when you execute command like OIFCFG, OCRCONFIG and etc

-- a log file will be created here.

cd \$ORACLE_HOME/log//crsd

cd \$ORACLE_HOME/log//racg

69. How to restore OCR file if corrupted?

Do the following to restore our OCR on Unix/Linux Systems.

To show the backups, type the commands ocrconfig

showbackup

Check the contents by doing ocrdump -backupfile my_file

Go to bin and stop the CRS. crs stop on all nodes.

Perform the restore ocrconfig

restore my_file

Restart the nodes crs start

We have spoken and seen the CVU (Cluster Verification Utility) play a crucial role during installation in our RAC on VMware Series. Check the OCRs integrity. Get a verbose output of all of the nodes by doing this: cluvfy comp ocr -n all -verbose

70. How to compare all nodes with cluvfy?

cluvfy comp ocr -n all

71. How to manage ASM in RAC?

Administering ASM Instances with SRVCTL in RAC:

Use the following command to add configuration information to an existing ASM instance:

srvctl add asm -n mynode_name -i myasm_instance_name -o myoracle_home

If, however, you choose not to add the -l option, then the changes are propagated throughout the entire ASM

instance pool.

To remove an ASM instance, use the following syntax:

```
srvctl remove asm -n mynode_name [-i myasm_instance_name]
```

In order to enable an ASM instance, use the following syntax:

```
srvctl enable asm -n mynode_name [-i myasm_instance_name]
```

In order to disable an ASM instance use the following syntax:

```
srvctl disable asm -n mynode_name [-i myasm_instance_name]
```

Note that you can also use the SRVCTL utility to start, stop, and get the status of an ASM instance. See the examples below.

To start an ASM instance, do the following

```
srvctl start asm -n mynode_name [-i myasm_instance_name] [-o start_options] [-c | -q]
```

To stop an ASM instance, type the following syntax:

```
srvctl stop asm -n mynode_name [-i myasm_instance_name] [-o stop_options] [-c | -q]
```

To list the configuration of an ASM instance do the following:

```
srvctl config asm -n mynode_name
```

To get the status of an ASM instance, see the following syntax:

```
srvctl status asm -n mynode_name
```

72. Where are the Cluster ware files stored on a RAC environment?

The Cluster ware is installed on each node (on an Oracle Home) and on the shared disks (the voting disks and the CSR file)

73. Where are the database software files stored on a RAC environment?

The base software is installed on each node of the cluster and the database storage on the shared disks.

74. What kind of storage we can use for the shared Cluster ware files?

- OCFS (Release 1 or 2)
- raw devices
- third party cluster file system such as GPFS or Veritas

75. What kind of storage we can use for the RAC database storage?

- OCFS (Release 1 or 2)
- ASM
- raw devices
- third party cluster file system such as GPFS or Veritas

76. What is a CFS?

A cluster File System (CFS) is a file system that may be accessed (read and write) by all members in a cluster at the same time. This implies that all members of a cluster have the same view.

77. What is an OCFS2?

The OCFS2 is the Oracle (version 2) Cluster File System which can be used for the Oracle Real Application Cluster.

78. Which files can be placed on an Oracle Cluster File System?

- Oracle Software installation (Windows only)
- Oracle files (controlfiles, datafiles, redologs, files described by the bfile datatype)
- Shared configuration files (spfile)
- OCR and voting disk
- Files created by Oracle during runtime

Note: There are some platform specific limitations.

79. Do you know another Cluster Vendor?

HP Tru64 UNIX, VERITAS, Microsoft

80. How is possible to install a RAC if we don't have a CFS?

This is possible by using a raw device.

81. 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.

82. What is a raw partition?

A raw partition is a portion of a physical disk that is accessed at the lowest possible level. A raw partition is created when an extended partition is created and logical partitions are assigned to it without any formatting. Once formatting is complete, it is called cooked partition.

83. When to use CFS 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.

84. 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.

85. What CRS is?

Oracle RAC 10g Release 1 introduced Oracle Cluster Ready Services (CRS), a platform-independent set of system services for cluster environments. In Release 2, Oracle has renamed this product to Oracle Cluster ware.

86. 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.

87. 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.

88. Do we have to have Oracle RDBMS on all nodes?

Each node of a cluster that is being used for a clustered database will typically have the RDBMS and RAC software loaded on it, but not actual data files (these need to be available via shared disk).

89. What are the restrictions on the SID with a RAC database? Is it limited to 5 characters?

The SID prefix in 10g Release 1 and prior versions was restricted to five characters by install/ config tools so that an ORACLE_SID of up to max of 5+3=8 characters can be supported in a RAC environment. The SID prefix is relaxed up to 8 characters in 10g Release 2, see bug 4024251 for more information.

90. Does Real Application Clusters support heterogeneous platforms?

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

91. What is the Load Balancing Advisory?

To assist in the balancing of application workload across designated resources, Oracle Database 10g Release 2 provides the Load Balancing Advisory. This Advisory monitors the current workload activity across the cluster and for each instance where a service is active; it provides a percentage value of how much of the total workload should be sent to this instance as well as service quality flag.

92. What is the Cluster Verification Utility (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.

93. Are there any issues for interconnect when sharing the same switch as the public network by using VLAN to separate the network?

RAC and Cluster ware deployment best practices suggests that the interconnect (private connection) be deployed on a stand-alone, physically separate, dedicated switch. On big network the connections could be instable.

94. 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.

95. If I am using Vendor Clusterware such as Veritas, IBM, Sun or HP, do I still need Oracle Clusterware to run Oracle RAC 10g?

Yes. When certified, you can use Vendor Clusterware however you must still install and use Oracle Clusterware for RAC. Best Practice is to leave Oracle Clusterware to manage RAC. For details see Metalink Note 332257.1 and for Veritas SFRAC see 397460.1.

96. Is RAC on VMWare supported?

Yes.

97. What is hangcheck timer used for?

The hangcheck timer checks regularly the health of the system. If the system hangs or stop the node will be restarted automatically.

There are 2 key parameters for this module:

-> hangcheck-tick: this parameter defines the period of time between checks of system health. The default value is 60 seconds; Oracle recommends setting it to 30seconds.

-> hangcheck-margin: this defines the maximum hang delay that should be tolerated before hangcheck-timer resets the RAC node.

98. Is the hangcheck timer still needed with Oracle RAC 10g?

Yes.

99. What files can I put on Linux OCFS2?

For optimal performance, you should only put the following files on Linux OCFS2:

- Datafiles
- Control Files
- Redo Logs
- Archive Logs
- Shared Configuration File (OCR)
- Voting File
- SPFILE

100. Is it possible to use ASM for the OCR and voting disk?

No, the OCR and voting disk must be on raw or CFS (cluster file system).

101. Can I change the name of my cluster after I have created it when I am using Oracle Clusterware?

No, you must properly uninstall Oracle Clusterware and then re-install.

102. What the O2CB is?

The O2CB is the OCFS2 cluster stack. OCFS2 includes some services. These services must be started before using OCFS2 (mount/ format the file systems).

103. What is the recommended method to make backups of a RAC environment?

RMAN to make backups of the database, dd to backup your voting disk and hard copies of the OCR file.

104. What command would you use to check the availability of the RAC system?

crs_stat -t -v (-t -v are optional)

105. What is the minimum number of instances you need to have in order to create a RAC?

You can create a RAC with just one server.

106. Name two specific RAC background processes?

RAC processes are: LMON, LMDx, LMSn, LKCx and DIAG.

107. Can you have many database versions in the same RAC?

Yes, but Clusterware version must be greater than the greater database version.

108. What was RAC previous name before it was called RAC?

OPS: Oracle Parallel Server

109. What RAC component is used for communication between instances?

Private Interconnect.

110. What is the difference between normal views and RAC views?

A RAC view has the prefix 'G'. For example, GV\$SESSION instead of V\$SESSION

111. Which command will we use to manage (stop, start) RAC services in command-line mode?

srvctl

112. How many alert logs exist in a RAC environment?

A- One for each instance.

113. How do you know you lost the voting disk?

Need document

If your RAC setup using single voting disk and if it is lost. The CRS will crash and your db server will be shut down. If it contain redundant voting disk it works until all other voting disks fail.

114. What format is the OCR file?

115. What will happen if we lost the voting disk?

If you have only 3 votedisk, So yes, If you lose 1/2 or more of all of your voting disks, then nodes get evicted from the cluster.

If you lose 1/2 or more of all of your voting disks, then nodes get evicted from the cluster, or nodes kick themselves out of the cluster. It doesn't threaten database corruption. Alternatively you can use external redundancy which means you are providing redundancy at the storage level using RAID.

For this reason when using Oracle for the redundancy of your voting disks, Oracle recommends that customers use 3 or more voting disks in Oracle RAC 10g Release 2. Note: For best availability, the 3 voting files should be physically separate disks. It is recommended to use an odd number as 4 disks will not be any more highly available than 3 disks, 1/2 of 3 is 1.5...rounded to 2, 1/2 of 4 is 2, once we lose 2 disks, our cluster will fail with both 4 voting disks or 3 voting disks.

116. What is the network protocol you used in configuring RAC?

- Public interface names must be the same for all nodes. If the public interface on one node uses the network adapter eth0, then you must configure eth0 as the public interface on *all* nodes. Network interface names are case-sensitive.
- You should configure the same private interface names for all nodes as well. If eth1 is the private interface name for the first node, then eth1 should be the private interface name for your second node. Network interface names are case-sensitive.
- The network adapter for the public interface must support TCP/IP.
- The network adapter for the private interface must support the user datagram protocol (UDP) using high-speed network adapters and a **network switch** that supports TCP/IP (Gigabit Ethernet or better).

117. How you check the health of Your RAC Database?

'crsctl' command from root or oracle user can be used to check the clusterware health But for starting or stopping we have to use root user or any privilege user.

```
$ crsctl check crs
```

118. If there is some issue with virtual IP how will you troubleshoot it? How will you change virtual ip?

To change the VIP (virtual IP) on a RAC node, use the command

```
$ srvctl modify nodeapps -A new_address
```

119. What kind of backup strategy you follow for your Databases?

We follow different backup strategy for our Databases depend on type of Database. We use different kind of Backup strategy for Production, Test, Performance, Demo, Development Databases. But the main aim is to recover the Database with minimal or no Data loss:

Production Databases:

Backup strategy for Production Database is as follows:

RMAN BACKUP:

incremental level 0 =>Weekly Basis at 6am -- Full backup of Database with archive logs and copy of Current control file

incremental level 1 =>Mon, Tues, Thurs, Friday at 6am --Changes from recent back to a particular day

cumulative backup =>Wed, Saturday at 6am -- changes from the lowest level i.e mon-wed and Thur-Sat.

While deciding the backup strategy for our production system of 300GB we had in our mind the following points:

- 1)Backup should be schedule at less peak hours.
- 2) With no loss we should recover Database in case of any Disaster.

EXPDP Backup:

Export Datapump backup on daily basis at 9pm.

We should have one datapump backup which should be most recent to recover the lost of Table or any Data. Below points are same:

- 1)Backup should be schedule at less peak hours.
- 2) With no loss we should recover Database in case of any Disaster.

Test Databases:

Usually Test Database is almost same as production in terms of Data. However whenever we want to test some patch or any script before applying to production we can apply in test and then apply in production. I usually prefer to have same backup strategy as production for Test Databases.

Development Database:

In a development Database. We can go for below backup strategy, However if you have space and enough infrastructure you can repeat the same backup strategy as above.

Expdp full backup:

In a Development environment, We should have full Database logical Backup up to date and should be schedule on daily basis, so that whenever there is some table drop or table backup is requested by developer you can restore that table from your Logical backup.

COLD RMAN BACKUP:

We can schedule cold RMAN backup on every Sundays at 9am (any time which you feel is convenient without affecting much to developers and end users). Below is the link for implementing RMAN cold

120. What will you backup in your RAC Database?

Backup strategy of RAC Database:

An RAC Database consists of

- 1) OCR
- 2) Voting disk &
- 3) Database files, control files, redo log files & Archive log files

121. How to recover your RAC Database?

122. What kind of backup strategy you are following for application server?

Complete Oracle Application Server Environment Backup can be done using the below techniques,

A complete Oracle Application Server environment backup includes:

- * A full backup of all files in the middle-tier Oracle homes (this includes Oracle software files and configuration files)
- * A full backup of all files in the Infrastructure Oracle home (this includes Oracle software files and configuration files)
- * A complete cold backup of the Metadata Repository
- * A full backup of the Oracle system files on each host in your environment.

Oracle AS Backup and Recovery Tool can be used for taking Oracle Application server backup.

123. How you Add node to your RAC Database?

To add a new node (server) to your RAC Database use the command 'srvctl' with the option given below.

Method-1:

```
$ srvctl add nodeapps -n newserver_name -o $ORACLE_HOME -A 149.181.220.1/255.255.255.0/eth1
```

Note: The -A flag precedes an address specification.

Method-2:

Run the addNode.sh script

On an existing node from the Oracle_home/oui/bin directory, run the addNode.sh script

124. For a Database created with ASM on RAC How you would add one more ASM configuration?

We can use DBCA in Silent Mode to Add ASM and Database Instances to Target Nodes

We can use the DBCA in silent mode to add instances to nodes on which you have extended an Oracle Clusterware home and an Oracle Database home. Use the following syntax where password is the password as given below:

```
$dbca -silent -addInstance -nodeList node -gdbName gdbname [-instanceName instname]
-sysDBAUserName sysdba -sysDBAPassword password
```

Note: We can use Oracle Enterprise Manager grid control also to do the same task.

125. How you add node for a RAC cluster? Step by step?

Below are the steps for adding non-rac node to an RAC Database:

I) Prerequisite Steps for Extending Oracle RAC to Target Nodes:

The following steps describe how to set up target nodes to be part of your cluster:

Step 1, "Make physical connections"

Step 2, "Install the operating system"

Step 3, "Create Oracle users"

Step 4, "Verify the installation" => use cluvfy utility for verification of clusterware installation

Eg: cluvfy stage -post hws -n node_list | all [-verbose]

II) Extend Oracle Clusterware to Target Nodes

In the above step you have to stop clusterware services with 'crsctl' and create a clone environment by copying file and making identical copy of clusterware home.

III) Configure Shared Storage on Target Nodes

Depending on the environment existing whether it is having asm, ocfs2, raw or any vendor shared storage make the environment same as the source. If the ASM HOME and ORACLE RAC database HOME exists in oracle then you don't need to do anything as ASM home to a node will happen implicitly if it is not so the case, you must first extend the Oracle Clusterware home (CRS_home), ASM home, and then the Oracle home (in that order), in order to add the new node to the cluster.

IV) Add the Oracle Real Application Clusters Database Homes to Target Nodes

We can add the Oracle RAC database home to target nodes using either of the following methods:

1) Extending the Database Home to Target Nodes Using Oracle Universal Installer in Interactive Mode (OR)

2) Extending the Database Home to Target Nodes Using Oracle Universal Installer in Silent Mode

Let us see the 2nd method which doesn't involve user interaction:

We can optionally run addNode.sh in silent mode,

replacing steps 1 through 6, as follows where node1, node1+1,

and so on are the target nodes to which you are adding the Oracle RAC database home.

* Ensure that you have successfully installed the Oracle Database with the Oracle RAC software on at least one node in your cluster environment.

* Ensure that the \$ORACLE_HOME environment variable identifies the successfully installed Oracle home.

Go to Oracle_home/oui/bin and run the addNode.sh script.

In the following example, node1, node1+1 (and so on) are the nodes that you are adding:

```
addNode.sh -silent "CLUSTER_NEW_NODES={node1, node1+1, ... node1+n}"
```

You can also specify the variable=value entries in a response file, known as filename, and you can run the addNode script as follows:

```
addNode.sh -silent -responseFile filename
```

Command-line values always override response file values.

v) Add ASM and Oracle RAC Database Instances to Target Nodes

We can add ASM and RAC Database Instances with the help of DBCA.

After you terminate your DBCA session, run the following command to verify the administrative privileges on the target node and obtain detailed information about these privileges where nodelist consists of the target nodes:

```
cluvfy comp admprv -o db_config -d oracle_home -n nodelist [-verbose]
```

Above the steps in brief so that we can crack the interview. The actual steps might be in detailed which we have to plan and do to avoid issues.

126. Which CRS process starts first?

127. What are the ways to configure TAF and Load Balancing?

128. When to use -repair parameter of ocrconfig command?

Use the ocrconfig -repair command to repair an OCR configuration on the node from which you run this command.

Use this command to add, delete, or replace an OCR location on a node that may have been stopped while you made changes to the OCR configuration in the cluster. OCR locations that you add must exist, have sufficient permissions, and, in the case of Oracle ASM disk groups, must be mounted before you can add them.

Syntax

```
ocrconfig -repair -add file_name | -delete file_name | -replace  
current_file_name -replacement new_file_name
```

Usage Notes

You must run this command as root.

Oracle High Availability Services must be started to successfully complete the repair.

The Cluster Ready Services daemon must be stopped before running ocrconfig -repair.

The file_name variable can be a valid OCR and either a device name, an absolute path name of an existing file, or the name of an Oracle ASM disk group. For example:

```
/dev/raw/raw1  
/oradbocfs/crs/data.ocr  
d:\oracle\mirror.ocr  
+newdg
```

If you specify an Oracle ASM disk group, the name of the disk group must be preceded by a plus sign (+).

You can only use one option with ocrconfig -repair at a time.

Running this command only modifies the local configuration and it and only affects the current node.

Example

To repair an OCR configuration:

```
# ocrconfig -repair -delete +olddg
```

129. What is crs_stat? What is the meaning of TARGET and STATUS column in crs_stat command output?

The crs_stat command provides status information for resources on the cluster nodes. To query resources with the crs_stat command, resource files must have read and execute permissions (r and x permissions on UNIX-based systems). An exception is the -g option, that anyone can use to verify whether a resource exists.

Resources are either ONLINE or OFFLINE as shown in the STATE attribute. An application resource in the ONLINE state is running successfully on a cluster node. This cluster node is shown indicating its state.

The TARGET value shows the state to which Oracle Clusterware attempts to set the resource. If the TARGET value is ONLINE and a cluster node fails, then Oracle Clusterware attempts to restart the application on another node if possible. If there is a condition forcing a resource STATE to be OFFLINE, such as a required resource that is OFFLINE, then the TARGET value remains ONLINE and Oracle Clusterware attempts to start the application or application resource once the condition is corrected.

A TARGET value for all non-application resources should be ONLINE unless the resource has a failure count higher than the failure threshold, in which case the TARGET is changed to OFFLINE. The Oracle Clusterware then treats the resource as if its STATE were OFFLINE. If the STATE is ONLINE and the TARGET is OFFLINE, then you can reset the target value to ONLINE using the crs_start command.

The verbose status -v gives additional information that may be useful, especially for troubleshooting. The RESTART_COUNT value shows how many times an application resource has been restarted on a single cluster node. The maximum number of restarts before Oracle Clusterware stops restarting the application is equal to RESTART_ATTEMPTS. FAILURE_COUNT shows the number of times that a resource has failed within the FAILURE_INTERVAL defined in the application profile. The maximum number of failures before Oracle Clusterware stops restarting an application is equal to the value set for the FAILURE_THRESHOLD parameter. If a cluster node fails and applications are waiting to be relocated due to the profile FAILOVER_DELAY attribute, then the verbose status also shows the FAILOVER_STATUS field. The FAILOVER_STATUS field is not shown at any other time. The FAILOVER_STATUS field shows the node the application failed on and how much time is left waiting for that node to restart before restarting on another node.

130. What is service? How to use services to gain maximum use of RAC?

Note: Explain about the TAF services

Applications should use the services feature to connect to the Oracle database. Services enable us to define rules and characteristics to control how users and applications connect to database instances.

Services are used to manage the workload in Oracle RAC, the important features of services are

- used to distribute the workload
- can be configured to provide high availability
- provide a transparent way to direct workload

The view v\$services contains information about services that have been started on that instance, here is a list from a fresh RAC installation

```
SQL> select name, network_name, creation_date, goal, dtp, aq_ha_notification, clb_goal from v$services;
```

NAME	NETWORK_NAME	CREATION_DATE	GOAL	DTP	AQ_HA_NOTIFICATION	CLB_GOAL
racdb_taf	racdb_taf	06-JAN-10	NONE	N	NO	LONG
racdbXDB	racdbXDB	06-JAN-10	NONE	N	NO	LONG
racdb.idevelopment.info	racdb.idevelopment.info	06-JAN-10	NONE	N	NO	LONG
SYS\$BACKGROUND		06-JAN-10	NONE	N	NO	SHORT
SYS\$USERS		06-JAN-10	NONE	N	NO	SHORT

The table above is described below

- **Goal** - allows you to define a service goal using service time, throughput or none
- **Connect Time Load Balancing Goal** - listeners and mid-tier servers contain current information about service performance
- **Distributed Transaction Processing** - used for distributed transactions
- **AQ_HA_Notifications** - information about nodes being up or down will be sent to mid-tier servers via the advance queuing mechanism
- **Preferred and Available Instances** - the preferred instances for a service, available ones are the backup instances

You can administer services using the following tools

- DBCA
- EM (Enterprise Manager)
- DBMS_SERVICES
- Server Control (srvctl)

Two services are created when the database is first installed, these services are running all the time and cannot be disabled.

- **sys\$background** - used by an instance's background processes only
- **sys\$users** - when users connect to the database without specifying a service they use this service

131. What is split Brain Syndrome? How Oracle Clusterware handles it?

132. What is STONIH algorithm?

133. What is cache fusion? Which Database background process facilitate it?

134. What is GRD? Where does it reside?

Explanation-1: RAC uses GRD (Global Resources Directory) to record information about how resources are used within a clustered database.

Global Resource directory , a common memory structures across SGA's, in other words this is the combination of GCS/GES memory structures (infact synchronizing all the times through cluster interconnect messages). All the resources in the cluster group form a central repository called GRD. which is integrated and distributed across the nodes memory structures. Each instance masters some of the resources (buffer) based on their weightage and accessibility) and together all formed called GRD. Basically a combination of GES and GCS.

Explanation-2: The Global Resource Directory (GRD) contains information about the current status of all shared resources. It is maintained by the GCS and GES to record information about resources and enqueues held on these resources. The GRD resides in memory and is used by the GCS and GES to manage the global resource activity. It is distributed throughout the cluster to all nodes. Each node participates in managing global resources and manages a portion of the GRD.

When an instance reads data blocks for the first time, its existence is local; that is, no other instance in the cluster has a copy of that block. The block in this state is called a current state (XI). The behavior of this block in memory is similar to any single-instance configuration, with the exception that GCS keeps track of the block even in a local mode. Multiple transactions within the instance have access to these data blocks. Once another instance has requested the same block, then the GCS process will update the GRD, changing the role of the data block from local to global.

Explanation-3: The RAC environment includes many resources such as multiple versions of data block buffers in buffer caches in different nodes, Oracle uses locking and queuing mechanisms to coordinate lock resources, data and interinstance data requests. Resources such as data blocks and locks must be synchronized between nodes as nodes within a cluster acquire and release ownership of them. The synchronization provided by the Global Resource Directory (GRD) maintains a cluster wide concurrency of the resources and in turn ensures the integrity of the shared data. Synchronization is also required for buffer cache management as it is divided into multiple caches, and each instance is responsible for managing its own local version of the buffer cache. Copies of data are exchanged between nodes, this sometimes is referred to as the *global cache* but in reality each nodes buffer cache is separate and copies of blocks are exchanged through traditional distributed locking mechanism.

Global Cache Services (GCS) maintain the cache coherency across buffer cache resources and Global Enqueue Services (GES) controls the resource management across the clusters non-buffer cache resources.

Cache Coherency : Cache coherency identifies the most up-to-date copy of a resource, also called the **master** copy, it uses a mechanism by which multiple copies of an object are keep consistent between Oracle instances. Parallel Cache Management (PCM) ensures that the master copy of a data block is stored in one buffer cache and consistent copies of the data block are stored in other buffer caches, the process **LCKx** is responsible for this task.

The lock and resource structures for instance locks reside in the GRD (also called the DLM), its a dedicated area within the shared pool. Details about the data blocks resources and cached versions are maintained by GCS. Additional details such as the location of the most current version, state of the buffer, role of the data block (local or global) and ownership are maintained by GES. Global cache together with GES form the GRD. Each instance maintains a part of

the GRD in its SGA. The GCS and GES nominate one instance, this will become the **resource master**, to manage all information about a particular resource. Each instance knows which instance master is with which resource.

135. Diff between cluster file system, RAW device and ASM?

Platforms and file types with each storage option

Storage option	Platforms	File types supported	File types not supported
Raw	All platforms	Database, CRS	Software/Dump files, Recovery
ASM	All platforms	Database, Recovery	CRS, Software/Dump
Certified Vendor CFS	AIX, HP Tru64 UNIX, SPARC Solaris	All	None
	HP-UX, HP Tru64 UNIX, SPARC Solaris	All	None
OCFS	Windows, Linux	Database, CRS, Recovery	Software/Dump files
NFS	Linux, SPARC Solaris	All	None

Raw devices: Raw devices need little explanation. As with single-instance Oracle, each tablespace requires a partition. You will also need to store your software and dump files elsewhere.

Pros: You won't need to install any vendor or Oracle-supplied clusterware or additional drivers.

Cons: You won't be able to have a shared oracle home, and if you want to configure a flash recovery area, you'll need to choose another option for it. Manageability is an issue. Further, raw devices are a terrible choice if you expect to resize or add tablespaces frequently, as this involves resizing or adding a partition.

NFS: NFS also requires little explanation. It must be used with a certified NAS device; Oracle has certified a number of NAS filers with its products, including products from EMC, HP, NetApp and others. NFS on NAS can be a cost-effective alternative to a SAN for Linux and Solaris, especially if no SAN hardware is already installed.

Pros: Ease of use and relatively low cost.

Cons: Not suitable for all deployments. Analysts recommend SANs over NAS for large-scale transaction-intensive applications, although there's disagreement on how big is too big for NAS.

Vendor CFS and LVMs: If you're considering a vendor CFS or LVM, you'll need to check the 10g Real Application Clusters Installation Guide for your platform and the Certify pages on MetaLink. A discussion of all the certified cluster file systems is beyond the scope of this article. Pros and cons depend on the specific solution, but some general observations can be made:

Pros: You can store all types of files associated with the instance on the CFS / logical volumes.

Cons: Depends on CFS / LVM. And you won't be enjoying the manageability advantage of ASM.

OCFS: OCFS is the Oracle-supplied CFS for Linux and Windows. This is the only CFS that can be used with these platforms. The current version of OCFS was designed specifically to store RAC files, and is not a full-featured CFS. You can store database, CRS and recovery files on it, but it doesn't fully support generic filesystem operations. Thus, for example, you cannot install a shared ORACLE_HOME on an OCFS device. The next version of OCFS, OCFS2, is currently out in beta version and will support generic filesystem operations, including a shared ORACLE_HOME.

Pros: Provides a CFS option for Linux and Windows.

Cons: Cannot store regular filesystem files such as Oracle software.

Easier to manage than raw devices, but not as manageable as NFS or ASM

ASM: Oracle recommends ASM for 10g RAC deployments, although CRS files cannot be stored on ASM. In fact, RAC installations using Oracle Database Standard Edition must use ASM.

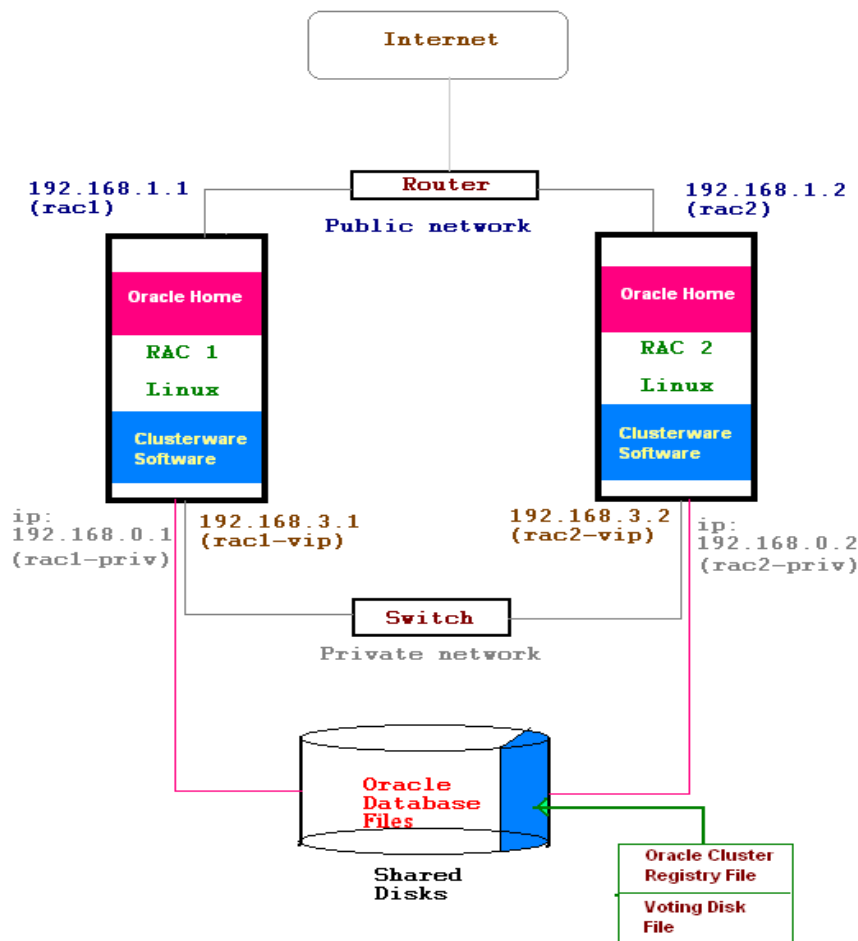
ASM is a little bit like a logical volume manager and provides many of the benefits of LVMs. But it also provides benefits LVMs don't: file-level striping/mirroring, and ease of manageability. Instead of running LVM software, you run an ASM instance, a new type of "instance" that largely consists of processes and memory and stores its information in the ASM disks it's managing.

Pros: File-level striping and mirroring; ease of manageability through Oracle syntax and OEM.

Cons: ASM files can only be managed through an Oracle application such as RMAN. This can be a weakness if you prefer third-party backup software or simple backup scripts. Cannot store CRS files or database software.

Note: We've seen that there's an array of storage options for the shared storage device in your RAC. These options depend on your platform, and many options don't store all types of database files, meaning they have to be used in conjunction with another option. For example, a DBA wanting to use ASM to store database files might take a 12-disk SAN, create 11 ASM disks for the database files and flash recovery area, leave the 12th disk raw and store CRS files on it, and maintain separate ORACLE_HOMES on the non-shared disks on each node.

136. Architecture of RAC



137. Explain how Instance Recovery takes place in RAC?

Explanation-1:

RAC relies on the cluster services for failure detection. The cluster services are a distributed kernel component that monitors whether cluster members can communicate with each other and through this process enforces the rule of cluster membership. This is taken care of by the Cluster Synchronization Service (CSS) with the CSSD process. The functions performed by CSS can be listed below.

1. Forms a cluster, add/remove members to/from a cluster.
2. Tracks which members in a cluster are active.
3. Maintains a cluster membership list, which is consistent on all member nodes.
4. Provides timely notification of membership changes.

When a node polls another node (target) in the cluster, and the target has not responded successfully after repeated attempts, a timeout occurs after approximately 60 seconds.

Among the responding nodes, the node that was started first and that is alive declares that the other node is not responding and has failed. This node becomes the new MASTER and starts evicting the non-responding node from the cluster. Once eviction is complete, cluster reformation begins. The reorganization process regroups accessible nodes and removes the failed ones.

LMON is a background process that monitors the entire cluster to manage the global resource. By constantly probing the other instances, it checks and manages instance death and associated recovery for the Global Cache Service (GCS). When a node joins or leaves the cluster, it handles reconfiguration of locks and associated resources. LMON handles the part of recovery associated with global resources. Failover of a service is also triggered by the EVMD process by firing a down event.

Once the reconfiguration of the nodes is complete, Oracle, in coordination with the EVMD and CRSD, performs several tasks.

1. Database/Instance recovery.
2. Failover of VIP system service.
3. Failover of the user/database services to another instance.

Database/Instance Recovery.

After a node in the cluster fails, it goes through several steps of recovery to complete changes at both the instance (cache) level and database level:

1. During the first phase of recovery, Global Enqueue Services (GES) remasters the enqueues, and Global Cache Services (GCS) remasters its resources from the failed instance among the surviving instances.
 2. The first step in the GCS remastering process is for Oracle to assign a new incarnation number.
 3. Oracle determines how many more nodes are remaining in the cluster. (Nodes are identified by a numeric starting with zero and incremented by one for every additional node in the cluster).
 4. IN AN ATTEMPT TO RECREATE THE RESOURCE MASTER OF THE FAILED INSTANCE, ALL GCS RESOURCE REQUESTS AND WRITE REQUESTS ARE TEMPORARILY SUSPENDED (GRD IS FROZEN).
 5. All the dead shadow process related to the GCS are cleaned from the failed instance.
 6. After enqueues are reconfigured, one of the surviving instances can grab the instance recovery enqueue.
 7. At the same time as GCS resources are remastered, SMON determines the set of blocks that need recovery. This set is called the Recovery set. With Cache Fusion an instance ships the contents of its block to the requesting instance without writing that dirty block to the disk (i.e. the on-disk version of the blocks may not contain the changes that are made by either instance). Because of this behavior, SMON needs to merge the content of all the online redo logs of each failed instance to determine the recovery set and the order of recovery.
 8. At this stage, buffer space for recovery is allocated, and the resources that were identified in the previous reading of the redo logs are claimed as recovery resources. THIS IS DONE TO AVOID OTHER INSTANCES FROM ACCESSING THOSE RESOURCES.
 9. A new master node for the cluster is created (A NEW MASTER NODE IS ONLY ASSIGNED IF THE FAILED NODE WAS THE PREVIOUS MASTER NODE IN THE CLUSTER). All GCS shadow processes are now traversed from a frozen state, and this completes the reconfiguration process.
 10. During the remastering of GCS from the failed instance (during cache recovery), MOST WORK ON THE INSTANCE PERFORMING RECOVERY IS PAUSED, AND WHILE TRANSACTION RECOVERY TAKES PLACE, WORKS OCCUR AT A SLOWER PACE.
- Subsequently, Oracle starts the database recovery process and begins the cache recovery process (i.e., rolling forward committed transactions). This is made possible by reading the redo log files of the failed instance. Because of the shared storage subsystem, redo log files of all instances participating in the cluster are visible to other instances. This makes any one instance that detected the failure read the redo log files of the failed instance and start the recover process.
11. After completion of the cache recovery, Oracle starts the transaction recovery operation i.e. roll forward the committed transaction and rollback the uncommitted transactions.

Explanation-2:

There are basically two types of failure in a RAC environment: instance and media. Instance failure involves the loss of one or more RAC instances, whether due to node failure or connectivity failure. Media failure involves the loss of one or more of the disk assets used to store the database files themselves.

If a RAC database undergoes instance failure, the first node still available that detects the failed instance or instances will perform instance recovery on all failed instances using the failed instances redo logs and the SMON process of the surviving instance. The redo logs for all RAC instances are located either on an OCFS shared disk asset or on a RAW file system that is visible to all the other RAC instances. This allows any other node to recover for a failed RAC node in the event of instance failure.

Recovery using redo logs allows committed transactions to be completed. Non-committed transactions are rolled back and their resources released.

There are experts with over a dozen years of working with Oracle databases that have yet to see an instance failure result in a non-recoverable situation with an Oracle database. Generally speaking, an instance failure in RAC or in normal Oracle requires no active participation from the DBA other than to restart the failed instance when the node becomes available once again.

If, for some reason, the recovering instance cannot see all of the datafiles accessed by the failed instance, an error will be written to the alert log. To verify that all datafiles are available, the ALTER SYSTEM CHECK DATAFILES command can be used to validate proper access.

Instance recovery involves nine distinct steps. The Oracle manual only lists eight, but in this case, the actual instance failure has been included:

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.

Instance recovery is automatic, and other than the performance hit to surviving instances and the disconnection of users who were using the failed instance, recovery is invisible to the other instances. If RAC failover and transparent application failover (TAF) technologies are properly utilized, the only users that should see a problem are those with in-flight transactions.

Note: One word of caution, during testing for this listing, an instance could not be brought back up after failure, a rare occurrence. A kill -9 was done on the SMON process on AULTLINUX1, within the Linux/RAC/RAW environment. AULTLINUX2 continued to operate and recovered the failed instance; however, an attempted restart of the instance on AULTLINUX1 yielded a Linux Error: 24: Too Many Files Open error. This was actually caused by something blocking the SPFILE link. Once the instance was pointed towards the proper SPFILE location during startup, it restarted with no problems.

138. How does your client connect to VIP or public IP? Or is it your choice?

139. Can private IP be changed?

Yes,

Using "oifcfg" command to change ip public sub and ip interconnect sub.

```
$oifcfg getif
bond0 6.9.23.32 global public
bond1 192.168.2.0 global cluster_interconnect

#oifcfg delif -global bond0
#oifcfg setif -global bond0/6.9.17.224:Public

#oifcfg delif -global bond1
#oifcfg setif -global bond1/192.168.20.0:cluster_interconnect
```

if you need to change the hostname of a server, you would have to re-install Clusterware - see Oracle Clusterware documentation for more information. But it does not need to be a complete re-installation. Just delete the node whose hostname you want to change from the cluster, change the hostname, and add the node as a new node. The only disadvantage this approach would have is that your node number would increase. E.g. you have a 2 node cluster and you need to change the hostname of what is node 1, then this node would have the node number 3 after the procedure and node number 1 would not exist in this cluster anymore.

If you change your VIP to a different subnet, then you have to change your Public IP to the same subnet as well (vice versa)..

You may not have to change your Private IPs (let's say 10.10.10.x) because these are your interconnect addresses and are "private" to the RAC nodes...

How to Change Interconnect/Public Interface IP or Subnet in Oracle Clusterware

Doc ID: 283684.1

Modifying the VIP or VIP Hostname of a 10g or 11g Oracle Clusterware Node

Doc ID: 276434.1

Considerations when Changing the Database Server Name or IP

Doc ID: 734559.1

Preparing For Changing the IP Addresses Of Oracle Database Servers

Doc ID: 363609.1

To Change hostname:

<http://www.pythian.com/blog/changing-hostnames-in-oracle-rac/>

140. What does root.sh do when you install 10g RAC? What is the importance of executing oraInstRoot.sh and root.sh scripts in Oracle Standalone and RAC environment?

As a part of post installation steps, we execute two scripts 'oraInstRoot.sh' and 'root.sh' scripts and Oracle also suggest to backup the 'oraInstRoot.sh' and 'root.sh' scripts. These two scripts we should execute as 'root' user as it displays after the Oracle software installation completes.

Importance of running 'oraInstRoot.sh' script:

The first Script that we run is 'oraInstRoot.sh' which is located in \$ORACLE_BASE/orainventory(/u01/app/oracle/orainventory) path. We execute 'oraInstRoot.sh' script for the following purposes:

- 1) It creates the inventory pointer file (/etc/orainst.loc), This file shows the inventory location and group it is linked to.
- 2) It Changes groupname of the orainventory directory to oinstall group.

Importance of running 'root.sh' script:

The Second script that we run is 'root.sh' script which is located in \$ORACLE_HOME (/u01/app/oracle/product/11.2.0/db_1) path. We execute 'root.sh' for the following purposes:

- 1) It will Create a /etc/oratab file. This is the file which we use to make automatic Database shutdown and startup. It is very important file.
- 2) It Sets the Oracle base and home environments.
- 3) It sets an appropriate permission to the OCR base directory
- 4) it creates the OCR backup and Network Socket directories.
- 5) It modifies the ownership to 'root' user on the Oracle base and Cluster home filesystem.
- 6) It Configures the OCR and voting disks (only on the first node)
- 7) It Starts the Clusterware daemons.
- 8) It adds Clusterware daemons to the inittab file.
- 9) It verifies whether the Clusterware is up on all nodes.
- 10) On the last node, initiates ./vipca in silent mode to configure nodeapps, such as, GSD, VIP, and ONS for all the nodes.
- 11) It verifies the super user privileges.
- 12) It Creates a trace directory. The 'trace' directory is again very vital for generating trace file to keep a track on user sessions in case of any error, troubleshooting and diagnosis purpose.
- 13) It Generates OCR keys for the 'root' user.
- 14) It Adds daemon information to the inittab file
- 15) starts up the Oracle High Availability Service Daemon (OHASD) process.
- 16) Creates and configures an ASM instance and starts up the instance.
- 17) Creates required ASM disk groups, if ASM is being used to put OCR and voting files.
- 18) Starts up the Cluster Ready Service Daemon (CRSD) process
- 19) Creates the voting disk file.
- 20) It Puts the voting disk on the Voting disk, if ASM type is selected.
- 21) It Displays voting disk details
- 22) Stops and restarts a cluster stack and other cluster resources on the local node
- 23) Backs up the OCR to a default location
- 24) It Installs the cvuqdisk-1.0.7-1 package
- 25) It Updates the Oracle inventory file.
- 26) Completes with the UpdateNodeList success operation.

When 'root.sh' is executed on the last node of the cluster, the following set of actions are likely to be performed by the script:

- 1) It Sets Oracle base and home environmental variables.
- 2) The /etc/oratab file will be created
- 3) It Performs the super user privileges verification.
- 4) Adds trace directories
- 5) It Generates OCR keys for the 'root' user.
- 6) Adds a daemon to inittab

- 7) Starts the Oracle High Availability Service Daemon (OHASD) process.
- 8) It Stops/starts a cluster stack and other cluster resources on the local node
- 9) Performs a backup of the OCR file
- 10) Installs the cvuqdisk-1.0.7-1 package
- 11) Updates the Oracle inventory file.
- 12) Completes with UpdateNodeList success operation.
141. How listener does handle requests in RAC?

Connections in RAC:

For failover configuration we should need to configure our physical ip of host name in listener configuration. Listener process is accepting new connection request and handover user process to server process or dispatcher process in Oracle.

Means using listener new connection is being established by Oracle. Once connection get established there is no need of listener process. If new connection is trying to get session in database and listener is down then what will be happening. User process gets error message and connection fails. Because listener is down in same host or something else problem. But in Oracle RAC database environment database is in sharing mode. Oracle RAC database is shared by all connected nodes. Means more than 1 listeners are running in various nodes.

In Oracle RAC database if user process is trying to get connection with some listener and found listener is down or node is down then Oracle RAC automatically transfer this request to another listener on another node. Up to Oracle 9i we use physical IP address in listener configuration. Means if requested connection gets failed then it will be diverting to another node using physical IP address of another surviving node. But during this automatically transfer, connection should need to wait up to get error message of node down or listener down using TCP/IP connection timeout. Means session should need to wait up to getting TCP/IP timeout error dictation. Once error message is received oracle RAC automatically divert this new connection request to another surviving node.

Using physical IP address there is biggest gap to get TCP/IP timeout for failover suggestion. Session should need to wait for same timeout. High availability of Oracle RAC depends on this time wasting error message.

Example:

In order to do this you need to connect to your database and set the LOCAL_LISTENER to a tns entry that points to the local server and REMOTE_LISTENER to the other servers.

Let's say you have 2 servers in the cluster: node1 and node2 and the db name is orcl. Add in your tnsnames entries for orcl1 (node1 only), orcl2 (node2 only) and rac (both node1 and node2).

Now connect to your database and set the following:

```
alter system set local_listener=orcl1 instance='orcl1';
```

```
alter system set local_listener=orcl2 instance='orcl2';
```

```
alter system set remote_listener=rac instance='*';
```

That way each instance will know who is the listener on its computer and all listeners in the cluster. If you do that correctly, Oracle will update all listeners with the instance information.

142. How can cache fusion improve or degrade performance?

143. Will you increase parallelism if you have RAC, to gain inter instance parallelism? What are the considerations to decide?

144. What is single point of failure in RAC?

The term SPOF (single point of failure) is used when one component fails causing the unavailability from all database(s) on your cluster

Note: OCR MIRROR is not a failover of the OCR so if you have no redundant storage for this file, you have a SPOF

Example: OCR and Voting Disks non-mirrored are SPOF

The single points of failure are:

- Firewall
- Application Server
- Fabric Switch
- SAN array

A failure of any one of these single points will result in unscheduled downtime, no matter how well the RAC cluster is designed and tuned.

It is critical to ensure that there is no single point of failure in a high availability configuration. Figure 4.2 illustrates exactly what eliminating single points of failure means.

145. A query running fast on one node is very slow on other node. All the nodes have same configurations. What could be the reasons in RAC environment?

146. Does RMAN behave differently in RAC?

No, it behaves the same way in both single instance and clustered RAC instance.

147. Can archive logs be placed on ASM disk? What about on RAW?

148. Write a RMAN script for taking backup of the database including arch log files in RAC?

```
run
{
allocate channel sbt_table .... ;
backup database ;
backup archive log all delete all input;
}
```

149. Write a sample script for RMAN for the recovery if all the instance are down.(First explain the procedure how you will restore)

Bring all nodes down.

Start one Node

Restore all datafiles and archive logs.

Recover 1 Node.

Open the database.

bring other nodes up.

Confirm that all nodes are operational.

Configure your client TNS so that it connects to PROD3 first then PROD2 and both fails then to prod1

150. Clients are performing some operation and suddenly one of the datafile is experiencing problem what do you do in RAC?

Bring the datafile offline recover the datafile.

151. What is the difference between a OS cluster and a RAC cluster?

152. what happens when a DML is issued in a RAC environment, how are requests for common buffers handled in a RAC environment?

A DML behaves on RAC in a similar manner as on a single node instance except a small change.

Each node on RAC has its own buffer cache. In case of DML, it looks for the data block in a local cache. If the current copy of block is not present, cache fusion is applied to get the latest copy from the local cache of other blocks. If this is not possible then the block is read into the local cache from disk and updated similar to a single node instance. The buffer cache of other blocks is fused with the current block later as required.

153. Explain about checkpoint and local & remote listener in RAC?

Checkpoint: If you want to checkpoint all instances in an Oracle RAC cluster then you would use the alter system checkpoint global command. A global checkpoint is the default checkpoint in Oracle RAC. The alter system checkpoint local command will cause a checkpoint to occur on the local node only.

Example:

1. ALTER SYSTEM SET CHECKPOINT LOCAL

Affects only the instance to which you are currently connected

2. ALTER SYSTEM CHECKPOINT or ALTER SYSTEM CHECKPOINT GLOBAL

Affects all instances in the cluster database

LOCAL_LISTENER AND REMOTE_LISTENER:

LOCAL_LISTENER on each node should point to the listener on that node. REMOTE_LISTENER should point to all listeners on all nodes if you want server side load balancing, otherwise don't set REMOTE_LISTENER.

Example configuration:

node1:

LOCAL_LISTENER_NODE1 =

(ADDRESS_LIST =

(ADDRESS =

(PROTOCOL = TCP)

(HOST = node1-vip)

(PORT = 1521)

)

)

```

REMOTE_LISTENERS =
  (ADDRESS_LIST =
    (ADDRESS =
      (PROTOCOL = TCP)
      (HOST = node1-vip)
      (PORT = 1521)
    )
    (ADDRESS =
      (PROTOCOL = TCP)
      (HOST = node2-vip)
      (PORT = 1521)
    )
  )
LOCAL_LISTENER=LOCAL_LISTENER_NODE1
REMOTE_LISTENER=REMOTE_LISTENERS node2:
LOCAL_LISTENER_NODE2 =
  (ADDRESS_LIST =
    (ADDRESS =
      (PROTOCOL = TCP)
      (HOST = node2-vip)
      (PORT = 1521)
    )
  )
)
REMOTE_LISTENERS =
  (ADDRESS_LIST =
    (ADDRESS =
      (PROTOCOL = TCP)
      (HOST = node1-vip)
      (PORT = 1521)
    )
    (ADDRESS =
      (PROTOCOL = TCP)
      (HOST = node2-vip)
      (PORT = 1521)
    )
  )
)
LOCAL_LISTENER=LOCAL_LISTENER_NODE2

```

REMOTE_LISTENER=REMOTE_LISTENERS on a 2-node cluster your REMOTE_LISTENER can point to a single listener but i find it easier to keep REMOTE_LISTENER identical on all nodes.

Note: The purpose of REMOTE_LISTENER is to connect all instances with all listeners so the instances can propagate their load balance advisories to all listeners. if you connect to a listener, this listener uses the advisories to decide who should service your connect. if the listener decides its local instance(s) are least loaded and should service your connect it passes your connect to the local instance. if the node you connected to is overloaded, the listener can use TNS redirect to redirect your connect a less loaded instance.

Explanation with Example:

suppose we have 2-node cluster: host1 and host2, with VIP address host1-vip and host2-vip respectively.

and one RAC database (orcl) running on this cluster; instance 1 (orcl1) on host1, and instance 2 (orcl2) on host2

we have listener_host1 running on host1, and listener_host2 running on host2.

listener_host1 is considered local listener for orcl1 instance, while listener_host2 is considered remote listener for that same orcl1 instance (because the listener is not running on the same machine as the database instance).

similarly, listener_host2 is considered local listener for orcl2 instance, and considered as remote listener for orcl1.

to make this consideration a real configuration, we configure the 2 parameters local_listener and remote_listener for both instances as below:

orcl1.local_listener=(address of listener_host1)

orcl1.remote_listener=(addresses of both listener_host1 and listener_host2)

orcl2.local_listener=(address of listener_host2)

orcl2.remote_listener=(addresses of both listener_host1 and listener_host2)

(as you see, we can simply use both listeners for the remote listener, as a simple configuration. But of course you could have configured orcl1.remote_listener=(address of listener_host2) only.)

with such configuration, both listeners in the cluster know about both instances, and about both hosts (statistics about host load, and instance load). and can make decision about forwarding a client connection request to the other node if it's less loaded. Which is the mechanism behind server-side load balancing.

clients are generally configured with tnsnames with both VIP addresses of the 2 hosts (i.e. can connect to either listener). so if a client attempts the connection to the database with the first IP (which is listener_host1), and suppose host1 is a bit loaded that host2, in such case the listener_host1 knows there is another instance orcl2 running on host2 that's less loaded. in such case, listener_host1 sends a redirect packet to the client asking him to transparently reconnect to listener_host2 to establish the database connection.

without such configuration of remote listener, each listener knows only about the local instance, and have nothing to do but connecting the client to the instance running on the same host as the listener. in such case you have only what's called client-side load balancing.

154. Explain LOCK Monitoring in RAC?

Oracle RAC is extremely complex and special scripts are required to identify locks within any RAC cluster. The monitoring of the RAC global enqueue services (GES) process is performed using the GV\$ENQUEUE_STAT view. The RAC resources managed by the GES include the following lock areas:

- Transaction locks – It is acquired in the exclusive mode when a transaction initiates its first row level change. The lock is held until the transaction is committed or rolled back.
- Library Cache locks - When a RAC database object (such as a table, view, procedure, function, package, package body, trigger, index, cluster, or synonym) is referenced during parsing or compiling of a SQL, DML or DDL, PL/SQL, or Java statement, the process parsing or compiling the statement acquires the library cache lock in the correct mode.
- Dictionary Cache Locks - Global enqueues are used in the cluster database mode. The data dictionary structure is the same for all Oracle instances in a cluster database, as it is for instances in a single-instance database.

155. Describe a scenario in which a vendor clusterware is required, in addition to the Oracle 10g Clusterware?

If you choose external redundancy for the OCR and voting disk, then to enable redundancy, the disk subsystem must be configurable for RAID mirroring/vendor clusterware. Otherwise, your system may be vulnerable because the OCR and voting disk are single points of failure.

156. How new connection establish in Oracle RAC?

For failover configuration we should need to configure our physical ip of host name in listener configuration. Listener process is accepting new connection request and handover user process to server process or dispatcher process in Oracle.

Means using listener new connection is being established by Oracle. Once connection get established there is no need of listener process. If new connection is trying to get session in database and listener is down then what will be happening. User process gets error message and connection fails. Because listener is down in same host or something else problem. But in Oracle RAC database environment database is in sharing mode. Oracle RAC database is shared by all connected nodes. Means more than 1 listeners are running in various nodes.

157. What are the characteristics of VIP in oracle RAC?

In Oracle RAC database if user process is trying to get connection with some listener and found listener is down or node is down then Oracle RAC automatically transfer this request to another listener on another node. Up to Oracle 9i we use physical IP address in listener configuration. Means if requested connection gets failed then it will be diverting to another node using physical IP address of another surviving node. But during this automatically transfer, connection should need to wait up to get error message of node down or listener down using TCP/IP connection timeout. Means session should need to wait up to getting TCP/IP timeout error dictation. Once error message is received oracle RAC automatically divert this new connection request to another surviving node.

Using physical IP address there is biggest gap to get TCP/IP timeout for failover suggestion. Session should need to wait for same timeout. High availability of Oracle RAC depends on this time wasting error message.

Virtual IP (VIP) is for fast connection establishment in failover dictation. Still we can use physical IP address in Oracle 10g in listener if we have no worry for failover timing. We can change default TCP/IP timeout using operating system

utilities or commands and kept smaller. But taking advantage of VIP (Virtual IP address) in Oracle 10g RAC database is advisable. There is utility also provided to configure virtual IP (vip) with RAC environment called VIPCA. Default path is \$ORA_CRS_HOME/bin. During installation of Oracle RAC, it is executed.

158. What information being written in the vote disk when Split brain syndrome occurs?

159. How does RAC do incase node becomes inactive?

In RAC if any node becomes inactive, or if other nodes are unable to ping/connect to a node in the RAC, then the node which first detects that one of the node is not accessible, it will evict that node from the RAC group. e.g. there are 4 nodes in a rac instance, and node 3 becomes unavailable, and node 1 tries to connect to node 3 and finds it not responding, then node 1 will evict node 3 out of the RAC groups and will leave only Node1, Node2 & Node4 in the RAC group to continue functioning.

The split brain concepts can become more complicated in large RAC setups. For example there are 10 RAC nodes in a cluster. And say 4 nodes are not able to communicate with the other 6. So there are 2 groups formed in this 10 node RAC cluster (one group of 4 nodes and other of 6 nodes). Now the nodes will quickly try to affirm their membership by locking controlfile, then the node that lock the controlfile will try to check the votes of the other nodes. The group with the most number of active nodes gets the preference and the others are evicted. Moreover, I have seen this node eviction issue with only 1 node getting evicted and the rest function fine, so I cannot really testify that if thats how it work by experience, but this is the theory behind it.

When we see that the node is evicted, usually oracle rac will reboot that node and try to do a cluster reconfiguration to include back the evicted node.

You will see oracle error: ORA-29740, when there is a node eviction in RAC. There are many reasons for a node eviction like heart beat not received by the controlfile, unable to communicate with the clusterware etc.

A good metalink note on understanding node eviction and how to address is Note ID: 219361.1

The CSS (Cluster Synchronization Service) daemon in the clusterware maintains the heart beat to the voting disk.

160. When can I use TAF or FCF?

Stating the obvious, but only when you have a database that can failover: this typically is RAC, but it could also be a Data Guard Primary-Standby site or even a single instance cold failover cluster. For both mechanisms you need to connect using database services (rather than traditional SIDs) since it's important that the service is not tied to an instance. TAF has been around since 8.1.5 whereas FCF is newer (10.1 onwards).

161. I have a java application that uses JDBC and a RAC database - should I use FCF?

Well, the answer is probably yes, providing you're on a recent Oracle JDBC driver (e.g. 10.2.0.4.0). There's no performance penalty to using it - you just need to set it up. Even if your application doesn't take account of failures flagged by the FCF mechanism at least you will always ensure that the connection pool has live connections.

162. Will I have to change my application?

To get the highest levels of reliability, yes, you will need to add additional error handling to re-apply your database work. This is true for both TAF (assuming you're not just doing read-only transactions) and FCF. If you have short lived transactions and don't want (or can't) change your application, then FCF can quickly clean up stale connections in the pool and your application will only suffer failures where it has a connection checked out from the pool (e.g. is in the middle of database activity).

163. What happens if the GSD does not run in 10g RAC, Will there be any impact to the 10g RAC when GSD is not running? Or else GSD should run in 10g RAC even?

GSD is not mandatory for 10G RAC to function properly.

It is present only for backward compatibility with 9i RAC.

Also going forward with 11G RAC , GSD is disabled by default and it is not started.

The GSD daemon has been replaced with the CRS in Oracle Database 10g

164. Is it possible in a RAC environment to force the database node that you want to connect too?

Yes and intelligent load balancing is an important tuning technique. You use the *TNSNAMES.ORA* file to direct "like-minded" end-users to specific nodes, and you can use TAF to specify failover nodes.

Again, in my years as a RAC DBA, I recommend functional load balancing, whereby like-minded transactions are grouped together on each node (e.g. one node for order queries, one node for customer queries). Using such a "functional" load balancing scheme, cache fusion pinging is greatly reduced and performance is improved.

One major RAC tuning issue is minimizing pinging access the [cache fusion layer](#), and smart DBA's will segregate users by node (i.e. customer queries on node 1, product queries on node 2, etc).

I don't recommend that "automatic load balancing" for all RAC databases and it's smart to group your end-users to nodes based on their query types.

Load-balancing for RAC involves extensive manual configuration to use a round-robin configuration to distribute the load among the instances. Starting in Oracle 10g release 2, we have a brand-new [load-balancing advisory](#) that promises to cut-down the manual effort for RAC load balancing between instances, but this does not take a functional approach to RAC load balancing, load balancing by the "type" of data being requested.

Again, a "functional" load balancing scheme is best, especially since pinging can become a major RAC bottleneck.

165. If my OCR and Voting Disks are in ASM, can I shut down the ASM instance?

No. You will have to stop the Oracle Clusterware stack on the node on which you need to stop the Oracle ASM instance. Either use "crsctl stop cluster -n *node_name*" or "crsctl stop crs" for this purpose.

166. I have changed my spfile with alter system set parameter_name with scope=spfile. The spfile is on ASM storage and the database will not start.

How to recover: </p>

In \$ORACLE_HOME/dbs

. oraenv <instance_name>

sqlplus "/" as sysdba

startup nomount

create pfile='recoversp' from spfile

/

shutdown immediate

quit

Now edit the newly created pfile to change the parameter to something sensible.

Then:

sqlplus "/" as sysdba

startup pfile='recoversp' (or whatever you called it in step one).

create spfile='+DATA/GASM/spfileGASM.ora' from pfile='recoversp'

/

N.B.The name of the spfile is in your original init(instance_name).ora so adjust to suit

shutdown immediate

startup

167. How do I use DBCA in silent mode to set up RAC and ASM?

If you already have an ASM instance/diskgroup then the following creates a RAC database on that diskgroup (run as the Oracle user):

```
$ORACLE_HOME/bin/dbca -silent -createDatabase -templateName General_Purpose.dbc -gdbName $SID -sid $SID -sysPassword $PASSWORD -systemPassword $PASSWORD -sysmanPassword $PASSWORD -dbsnmpPassword $PASSWORD -emConfiguration LOCAL -storageType ASM -diskGroupName $ASMGROUPNAME -datafileJarLocation $ORACLE_HOME/assistants/dbca/templates -nodeinfo $NODE1,$NODE2 -characterset WE8ISO8859P1 -obfuscatedPasswords false -sampleSchema false -oratabLocation /etc/oratab
```

The following will create a ASM instance & 1 diskgroup (run as the ASM/Oracle user)

```
$ORA_ASM_HOME/bin/dbca -silent -configureASM -gdbName NO -sid NO -emConfiguration NONE -diskList $ASM_DISKS -diskGroupName $ASMGROUPNAME -nodeinfo $NODE1,$NODE2 -obfuscatedPasswords false -oratabLocation /etc/oratab -asmSysPassword $PASSWORD -redundancy $ASMREDUNDANCY where ASM_DISKS = '/dev/sda1,/dev/sdb1' and ASMREDUNDANCY='NORMAL'
```

168. How does OCR mirror work? What happens if my OCR is lost / corrupt?

OCR is the Oracle Cluster Registry, it holds all the cluster related information such as instances, services. The OCR file format is binary and starting with 10.2 it is possible to mirror it. Location of file(s) is located in: /etc/oracle/ocr.loc in ocrconfig_loc and ocrmirrorconfig_loc variables. Obviously if you only have one copy of the OCR and it is lost or corrupt then you must restore a recent backup, see ocrconfig utility for details, specifically -showbackup and -restore flags. Until a valid backup is restored the Oracle Clusterware will not startup due to the corrupt/missing OCR file.

The interesting discussion is what happens if you have the OCR mirrored and one of the copies gets corrupt? You would expect that everything will continue to work seamlessly. Well.. Almost.. The real answer depends on when the corruption takes place.

If the corruption happens while the Oracle Clusterware stack is up and running, then the corruption will be tolerated and the Oracle Clusterware will continue to function without interruptions. Despite the corrupt copy. DBA is advised to repair this hardware/software problem that prevent OCR from accessing the device as soon as possible; alternatively, DBA can replace the failed device with another healthy device using the ocrconfig utility with -replace flag.

If however the corruption happens while the Oracle Clusterware stack is down, then it will not be possible to start it up until the failed device becomes online again or some administrative action using ocrconfig utility with -overwrite flag is taken. When the Clusterware attempts to start you will see messages similar to:

total id sets (1), 1st set (1669906634,1958222370), 2nd set (0,0) my votes (1), total votes (2)

2006-07-12 10:53:54.301: [OCRRAW][1210108256]proprioini:disk 0 (/dev/raw/raw1) doesn't have enough votes (1,2)

2006-07-12 10:53:54.301: [OCRRAW][1210108256]proprseterror: Error in accessing physical storage [26]

This is because the software can't determine which OCR copy is the valid one. In the above example one of the OCR mirrors was lost while the Oracle Clusterware was down. There are 3 ways to fix this failure:

a) Fix whatever problem (hardware/software?) that prevent OCR from accessing the device.

b) Issue "ocrconfig -overwrite" on any one of the nodes in the cluster. This command will overwrite the vote check built into OCR when it starts up. Basically, if OCR device is configured with mirror, OCR assign each device with one vote. The rule is to have more than 50% of total vote (quorum) in order to safely make sure the available devices contain the latest data. In 2-way mirroring, the total vote count is 2 so it requires 2 votes to achieve the quorum. In the example above there isn't enough vote to start if only one device with one vote is available. (In the earlier example, while OCR is running when the device is down, OCR assign 2 vote to the surviving device and that is why this surviving device now with two votes can start after the cluster is down). See warning below

c) This method is not recommended to be performed by customers. It is possible to manually modify ocr.loc to delete the failed device and restart the cluster. OCR won't do the vote check if the mirror is not configured. See warning below

EXTREME CAUTION should be exercised if choosing option b or c above since data loss can occur if the wrong file is manipulated; please contact Oracle Support for assistance before proceeding.

169. How do troubleshoot node reboot?

Note 265769.1 Troubleshooting CRS Reboots

Note.559365.1 Using Diagwait as a diagnostic to get more information for diagnosing Oracle Clusterware Node evictions.

170. What do you do if you see GC CR BLOCK LOST in top 5 Timed Events in AWR Report?

This is most likely due to a fault in interconnect network.

Check netstat -s

if you see "fragments dropped" or "packet reassemblies failed", Work with your system administrator find the fault with network.

171. SRVCTL cannot start instance, I get the following error PRKP-1001 CRS-0215, however SQLPLUS can start it on both nodes? How do you identify the problem?

Set the environmental variable SRVM_TRACE to true.. And start the instance with srvctl. Now you will get detailed error stack.

172. What are the major RAC wait events?

In a RAC environment the buffer cache is global across all instances in the cluster and hence the processing differs. The most common wait events related to this are gc cr request and gc buffer busy

GC CR request: the time it takes to retrieve the data from the remote cache

Reason: RAC Traffic Using Slow Connection or Inefficient queries (poorly tuned queries will increase the amount of data blocks requested by an Oracle session. The more blocks requested typically means the more often a block will need to be read from a remote instance via the interconnect.)

173. What is usage of CRS_RELOCATE command?

The crs_relocate command relocates applications and application resources as specified by the command options that you use and the entries in your application profile. The specified application or application resource must be registered and running under Oracle Clusterware in the cluster environment before you can relocate it. The command displays a message if you specify a cluster node that is unavailable or if the attempt to relocate failed. You must have full administrative privileges to use this command. When you perform a crs_relocate command, Oracle Clusterware first runs the stop entry point of the action program on the node on which it is currently running. The Oracle Clusterware then performs the start entry point of the action program to start it on a new node.

If Oracle Clusterware fails to stop the application or application resource on the current node due to an action program error, then it marks it as UNKNOWN. You cannot run crs_relocate on a resource in this state. Instead, run a crs_stop -f command on the resource and restart it with crs_start to return it to the ONLINE state before you attempt to relocate it again. If Oracle Clusterware fails to restart an application resource, then you may need to check the resource action program.

If the action program start entry point fails to run successfully, then the stop entry point is run. If the stop entry point fails to run successfully, then the state is marked as UNKNOWN and relocation attempts are stopped. If the stop entry point succeeds, then the state is set to OFFLINE. The target state remains ONLINE however, so subsequent cluster node failures or restarts can cause Oracle Clusterware to attempt to restart the application. If you have not specified the node to which to relocate it and if there are available cluster nodes that satisfy the placement criteria, then Oracle Clusterware attempts to start the application on one of these available nodes.

If one or more user-defined attributes have been defined for application resources, then you can specify values for these attributes when relocating an application with the `crs_relocate` command. The specified value is passed to the action program as an environment variable with the attribute name.

The actions that Oracle Clusterware takes while relocating an application resource are echoed on the command line. You can also monitor them using the Event Manager (EVM). Standard error and standard output from a resource action program that is started by the `crs_relocate` command are redirected to the standard error and standard output for `crs_relocate`. Note that if the Oracle Clusterware daemon starts an application, then standard error and standard output of the action program is lost. Within an action program, you can check for user invocation of the action program using reason codes.

Syntax:

```
crs_relocate resource_name [-c cluster_node] [-q]
```

Example:

The following example relocates an application resource to the node known as rac1:

```
crs_relocate postman -c rac1
Attempting to stop `postman` on node `rac2`
Stop of `postman` on node `rac1` succeeded
Attempting to start `postman` on node `rac1`
Start of `postman` on node `rac1` succeeded
```

The following example attempts to relocate all application resources from node rac2 to node rac1:

```
crs_relocate -s rac2 -c rac1
Attempting to stop `postman` on node `rac2`
Stop of `postman` on node `rac2` succeeded.
Attempting to start `postman` on node `rac1`
Start of `postman` on node `rac1` succeeded.
Attempting to stop `calc` on node `rac2`
Stop of `calc` on node `rac2` succeeded.
Attempting to start `calc` on node `rac1`
Start of `calc` on node `rac1` succeeded.
```

If a user-defined attribute `USR_DEBUG` has been defined, then the following example runs the stop and start entry point of the action program with the `USR_DEBUG` environment variable set to FALSE. This overrides any value set in the application profile. In the corresponding action program, if you add the following line to the appropriate section of the action program, then you can view the value:

```
echo $USR_DEBUG
```

Then run the following command:

```
# crs_relocate USR_DEBUG=false database
```

174. What is the use of `CRS_GETPERM` and `CRS_SETPERM`?

CRS_GETPERM:

Inspects the permissions associated with a resource.

Syntax:

```
crs_getperm resource_name [-u user] [-g group]
```

To obtain the permissions associated with a resource, use the following syntax:

```
crs_getperm resource_name
```

Example:

To list the permissions associated with the postman application, use the following command:

```
crs_getperm postman
```

CRS_SETPERM:

Modifies the permissions associated with a resource. This command is similar to the `chmod` command in UNIX-based systems or the Windows desktop options, in this order: File, Properties, Security, and Permissions.

Syntax:

```
crs_setperm resource_name -u aclstring [-q]
crs_setperm resource_name -x aclstring [-q]
crs_setperm resource_name -o user_name [-q]
crs_setperm resource_name -g group_name [-q]
```

In the previous example syntax, -u updates the acl string, -x deletes the acl string, -o changes the owner of the resource, and -g changes the primary group of the resource, and aclstring is one of the following:

```
user:username:rwX
group:groupname:r-x
other::r--
```

Example:

The following example modifies the permissions on the admin1 user for the postman application:

```
crs_setperm postman -u user:admin1:r-x
```

175. What is the use of CRS_REGISTER and CRS_UNREGISTER?

CRS_REGISTER:

The `crs_register` command registers one or more applications specified with the *resource_name* parameter for each application. This command requires write access to the target application. This command only succeeds if the profile is found either in the default location or in the directory specified by the `-dir` option. An application must be registered in order for Oracle Clusterware to monitor the resource or to start, restart, or relocate a highly available application associated with an application resource. An application registration must be updated for any changes to an application profile to take effect. This command requires full administrative privileges.

An application can be registered or updated only if the Oracle Clusterware daemon is active and an Oracle Clusterware application profile exists in the profile directory for this application. If fields are missing from an application profile, then the profile is merged with the default profile template and Oracle uses the values in the default profile template.

The ownership and default permissions for the application are set during the registration process. You can register any .cap file as long as the permissions for the file permit read and write, for example, `crs_profile` and `crs_register` must be able to read the file. By default, the user who registers the application is the owner of the application. If the profile cannot be registered, then Oracle Clusterware displays messages explaining why. Use the `crs_stat` command to verify that the application is registered.

Syntax:

You can use the `crs_register` command to register and update applications. Use the following `crs_register` syntax to register an application:

```
crs_register resource_name [-dir directory_path] [...] [-u] [-f] [-q]
```

The *resource_name* [...] parameter can be the name of one or more application resources as specified in an application profile. If you do not specify any options, then the `crs_relocate` command relocates each specified application resource according to its placement policy and required resource lists. The Oracle Clusterware does not relocate a resource if there are interdependent application resources unless you specify the `-f` option. A profile must exist for the application that you are registering.

The `-dir` option specifies where the .cap file is if it is not in the default directory.

Use `crs_register -u` immediately after a `crs_profile -update` or a manual edit of an application profile to ensure that the changes take effect immediately.

Use the following `crs_register` command syntax to update a registered application:

```
crs_register resource_name -update [option ...] [-o option,...] [-q]
```

Example:

The following example registers an application named postman for which an application profile exists in the *CRS home/crs/profile* directory:

```
CRS home/bin/crs_register postman
```

Note: The profile will be in the *crs/profile* directory if the profile is created by the root user. The profile will be in the *crs/public* directory if the profile is created by any other user.

CRS_UNREGISTER: The `crs_unregister` command removes the registration information of Oracle Clusterware resources from the binary Oracle Clusterware registry database. The Oracle Clusterware will no longer acknowledge this resource. An application associated with a resource that is unregistered is no longer highly available. You must have full administrative privileges to use this command.

Upon successful completion of the `crs_unregister` command, the resource is removed from the online Oracle Clusterware environment. You cannot unregister a resource that is a required resource for another resource. You must stop the resource by using the `crs_stop` command before unregistering it.

Syntax:

Use the `crs_unregister` command with the following syntax:

```
crs_unregister resource_name [...] [-q]
```

The only option available for this command is `-q`, that runs the `crs_unregister` command in quiet mode, which means no messages are displayed.

Example:

The following example unregisters a highly available application called `postman`:

```
crs_unregister postman
```

176. What is the use of CRS_PROFILE?

Creates, validates, deletes, and updates an Oracle Clusterware application profile. It works on the user's copy of a profile. This command does not operate against the database that Oracle Clusterware is using.

You can also use `crs_profile` to generate a template script. The `crs_profile` command creates new application profiles and validates, updates, deletes, or lists existing profiles. An application profile assigns values to attributes that define how a resource should be managed or monitored in a [cluster](#). For the root user, profiles are written to the `CRS home/crs/profile` directory. For non-privileged users, the profile is written to the `CRS home/crs/public` directory. Values in the profile that are left blank are ignored and may be omitted if not required. Omitted profile variables that are required for a resource type can cause validation or registration to fail.

After you have created an application profile and registered the application with Oracle Clusterware using the `crs_register` command, you can use other Oracle Clusterware commands, such as `crs_stat`, `crs_start`, `crs_stop`, `crs_relocate`, and `crs_unregister`, on the application. You must register applications using the `crs_register` command *before* the other Oracle Clusterware commands are available to manage the application. The `crs_profile` command may have other options for defining user-defined attributes in a profile.

Syntax:

Use the `crs_profile` command with the following syntax to create an application profile template:

```
crs_profile -create resource_name -t application [-a action_script] [-B executable_pathname] [-dir directory] [-d description] [-p placement_policy] [-h hosting_nodes] [-r required_resources] [-l optional_resources] [-o option,...] [attribute_flag attribute_value] [...] [-f] [-q]
```

To create an application profile from an application profile template:

```
crs_profile -create resource_name -l template_file [-f] [-q]
```

To validate the application profile syntax of a profile, enter the following command:

```
crs_profile -validate resource_name [-q]
```

To list one or more application profiles:

```
crs_profile -print [resource_name [...]] [-q]
```

To create an application profile template from an existing application profile:

```
crs_profile -template resource_name [-O template_file] [-q]
```

To update an application profile:

```
crs_profile -update resource_name [option [...]] [-q]
```

To delete an application profile and its associated action program:

```
crs_profile -delete resource_name [-q]
```

Note: The `crs_profile -delete` command deletes the resource profile file but does not delete the action script file.

Example:

The following is an example of using the `crs_profile` command to create the application profile for an application named `dtcalc`:

```
crs_profile -create dtcalc
```

The following is an example of using the `crs_profile` command to validate the application profile named `dtcalc`:

```
crs_profile -validate dtcalc
```

If you do not specify either the `-a` or `-B` options, then the profile is created without an action program value. You should create a readable, executable script and update the profile with this script's value before attempting to start the application. If you specify the `-a` option alone, then the action program specified must exist at the specified location or in the default directory if no path is specified. Otherwise, the command fails.

177. What components in RAC must reside in shared storage?

All datafiles, controlfiles, SPFiles, redo log files must reside on cluster-aware shared storage.

178. What is the significance of using cluster-aware shared storage in an Oracle RAC environment?

All instances of an Oracle RAC can access all the datafiles, controlfiles, SPFILE's, redo log files when these files are hosted out of cluster-aware shared storage which are group of shared disks.

179. Give few examples for solutions that support cluster storage?

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

180. How can we configure the cluster interconnect?

- Configure User Datagram Protocol (UDP) on Gigabit Ethernet for cluster interconnects.
- On UNIX and Linux systems we use UDP and RDS (Reliable data socket) protocols to be used by Oracle Clusterware.
- Windows clusters use the TCP protocol.

181. How do users connect to database in an Oracle RAC environment?

Users can access a RAC database using a client/server configuration or through one or more middle tiers, with or without connection pooling. Users can use oracle services feature to connect to database.

182. What are the characteristics controlled by Oracle services feature?

The characteristics include a unique name, workload balancing, failover options, and high availability.

183. Which enables the load balancing of applications in RAC?

Oracle Net Services enable the load balancing of application connections across all of the instances in an Oracle RAC database.

184. Give situations under which VIP address failover happens?

VIP addresses failover happens when the node on which the VIP address runs fails; all interfaces for the VIP address fails, all interfaces for the VIP address are disconnected from the network.

185. What is the significance of VIP address failover?

When a VIP address failover happens, Clients that attempt to connect to the VIP address receive a rapid connection refused error. They don't have to wait for TCP connection timeout messages.

186. What are the administrative tools used for Oracle RAC environments?

Oracle RAC cluster can be administered as a single image using the below

- OEM (Enterprise Manager),
- SQL*PLUS,
- Server control (SRVCTL),
- Cluster Verification Utility (CLUVFY),
- DBCA and NETCA

187. How do we verify that RAC instances are running?

Issue the following query from any one node connecting through SQL*PLUS.

```
$connect sys/sys as sysdba
```

```
SQL>select * from V$ACTIVE_INSTANCES;
```

The query gives the instance number under INST_NUMBER column, host instance name under INST_NAME column.

188. Where can we apply FAN UP and DOWN events?

FAN UP and FAN DOWN events can be applied to instances, services and nodes.

189. State the use of FAN events in case of a cluster configuration change?

During times of cluster configuration changes, Oracle RAC high availability framework publishes a FAN event immediately when a state change occurs in the cluster. So applications can receive FAN events and react immediately. This prevents applications from polling database and detecting a problem after such a state change.

190. Why should we have separate homes for ASM instance?

It is a good practice to have ASM home separate from the database home (ORACLE_HOME). This helps in upgrading and patching ASM and the Oracle database software independent of each other. Also, we can deinstall the Oracle database software independent of the ASM instance.

191. What is rolling upgrade? Can rolling upgrade be used to upgrade from 10g to 11g database?

It is a new ASM feature from Database 11g. ASM instances in Oracle database 11g release (from 11.1) can be upgraded or patched using rolling upgrade feature. This enables us to patch or upgrade ASM nodes in a clustered

environment without affecting database availability. During a rolling upgrade we can maintain a functional cluster while one or more of the nodes in the cluster are running in different software versions.

No, it can be used only for Oracle database 11g releases (from 11.1).

192. Can the DML_LOCKS and RESULT_CACHE_MAX_SIZE be identical on all instances?

These parameters can be identical on all instances only if these parameter values are set to zero.

193. What two parameters must be set at the time of starting up an ASM instance in a RAC environment?

The parameters CLUSTER_DATABASE and INSTANCE_TYPE must be set.

194. How does an Oracle Clusterware manage CRS resources?

Oracle Clusterware manages CRS resources based on the configuration information of CRS resources stored in OCR(Oracle Cluster Registry).

195. Name some Oracle Clusterware tools and their uses?

- OIFCFG - allocating and deallocating network interfaces.
- OCRCONFIG - Command-line tool for managing Oracle Cluster Registry.
- OCRDUMP - Identify the interconnect being used.
- CVU - Cluster verification utility to get status of CRS resources

196. What are the modes of deleting instances from Oracle Real Application cluster Databases? We can delete instances using silent mode or interactive mode using DBCA (Database Configuration Assistant).

197. How do we remove ASM from an Oracle RAC environment?

We need to stop and delete the instance in the node first in interactive or silent mode. After that ASM can be removed using srvctl tool as follows:

```
srvctl stop asm -n node_name
```

```
srvctl remove asm -n node_name
```

We can verify if ASM has been removed by issuing the following command:

```
srvctl config asm -n node_name
```

198. How do we verify that an instance has been removed from OCR after deleting an instance?

Issue the following srvctl command:

```
srvctl config database -d database_name
```

```
cd CRS_HOME/bin
```

```
./crs_stat
```

199. What are the performance views in an Oracle RAC environment?

We have v\$ views that are instance specific. In addition we have GV\$ views called as global views that has an INST_ID column of numeric data type. GV\$ views obtain information from individual V\$ views.

200. What is the difference between server-side and client-side connection load balancing?

Client-side balancing happens at client side where load balancing is done using listener. In case of server-side load balancing listener uses a load-balancing advisory to redirect connections to the instance providing best service.

201. Give the usage of srvctl?

- srvctl start instance -d db_name -i "inst_name_list" [-o start_options]
- srvctl stop instance -d name -i "inst_name_list" [-o stop_options]
- srvctl stop instance -d orcl -i "orcl3,orcl4" -o immediate
- srvctl start database -d name [-o start_options]
- srvctl stop database -d name [-o stop_options]
- srvctl start database -d orcl -o mount

202. What is the purpose of the ONS daemon?

The Oracle Notification Service (ONS) daemon is an daemon started by the CRS clusterware as part of the nodeapps. There is one ons daemon started per clustered node.

The Oracle Notification Service daemon receives a subset of published clusterware events via the local evmd and racgimom Clusterware daemons and forward those events to application subscribers and to the local listeners.

This in order to facilitate:

- a. the FAN or Fast Application Notification feature or allowing applications to respond to database state changes.
- b. the 10gR2 Load Balancing Advisory, the feature that permit load balancing across different RAC nodes dependent of the load on the different nodes. The rdbms MMON is creating an advisory for distribution of work every 30seconds and forward it via racgimom and ONS to listeners and applications.

203. What is Dynamic Remastering?

Every instance (node) owns the block(buffers) called master node for that blocks when accessed first time. If the other node requires the same, it has to send a request to the owning node to get the blocks (CR/CUR modes). If the requests are more than 50 in an hour the mastership will be transferred to the other node.

204. What are RAC based services? What are difference between normal database service and RAC services?

- Is a means of grouping sessions that are doing the same kind of work
- Provides single-system image instead of multiple instances image
- Is a part of the regular administration tasks that provide dynamic service-to-instance allocation
- Is the base for high availability of connections
- Provides a new performance-tuning dimension
- Normal services not maintained in data dictionaries where the Rac services maintained in data dictionary.

205. What happens if one of the node is not able to access the voting disk?

The master node OCSSD in the instance verifies the votes in the voting disk periodically and ensure quorum is matched, if quorum is not matched then it posts the failure nodes OCSSD to evict the cluster.

Or Node OCSSD recognises it and evicts the node

206. What happens if all of the nodes not able to access the voting disk?

This can be lead to split brain syndrome, each node acts as master node, with the disktimeout setting clusterware waits until that period can be delayed using disktimeout setting to reasonable value using crsctl. all nodes reboot.

207. What happens if one of the node is not able to communicate via private interconnect?

Node OCSSD recognises it and evicts the node

208. What happens if all of the nodes not able to communicate via private interconnect?

This can be lead to split brain syndrome, each node acts as master node, with the csstimeout setting clusterware waits until that period , can be delayed using csstimeout setting to reasonable value using crsctl. all nodes reboot.

209. What is split brain syndrome?

Each node acts as master since there is communication or common storage access break down.

210. Which background daemons initiate node eviction?

Ocssd

211. Which background daemon starts clusterware or the resources?

crsd

212. What are my options for load balancing with Oracle RAC? Why do I get an uneven number of connections on my instances?

All the types of load balancing available currently (9i-10g) occur at connect time.

This means that it is very important how one balances connections and what these connections do on a long term basis.

Since establishing connections can be very expensive for your application, it is good programming practice to connect once and stay connected. This means one needs to be careful as to what option one uses. Oracle Net Services provides load balancing or you can use external methods such as hardware based or clusterware solutions.

The following options exist prior to Oracle RAC 10g Release 2 (for 10g Release 2 see Load Balancing Advisory):

Random

Either client side load balancing or hardware based methods will randomize the connections to the instances.

On the negative side this method is unaware of load on the connections or even if they are up meaning they might cause waits on TCP/IP timeouts.

Load Based

Server side load balancing (by the listener) redirects connections by default depending on the RunQ length of each of the instances. This is great for short lived connections. Terrible for persistent connections or login storms. Do not use this method for connections from connection pools or applicaton servers

Session Based

Server side load balancing can also be used to balance the number of connections to each instance. Session count balancing is method used when you set a listener parameter, `prefer_least_loaded_node_listener-name=off`. Note listener name is the actual name of the listener which is different on each node in your cluster and by default is `listener_nodename`.

Session based load balancing takes into account the number of sessions connected to each node and then distributes the connections to balance the number of sessions across the different nodes.

213. How can a customer mask the change in their clustered database configuration from their client or application? (I.E. So I do not have to change the connection string when I add a node to the Oracle RAC database)

The combination of Server Side load balancing and Services allows you to easily mask cluster database configuration changes. As long as all instances register with all listeners (use the LOCAL_LISTENER and REMOTE_LISTENER parameters), server side load balancing will allow clients to connect to the service on currently available instances at connect time.

The load balancing advisory (setting a goal on the service) will give advice as to how many connections to send to each instance currently providing a service. When a service is enabled on an instance, as long as the instance registers with the listeners, the clients can start getting connections to the service and the load balancing advisory will include that instance in its advice.

With Oracle RAC 11g Release 2, the Single Client Access Name (SCAN) provides a single name to be put in the client connection string (as the address). Clients using SCAN never have to change even if the cluster configuration changes such as adding nodes.

214. What is the Load Balancing Advisory?

To assist in the balancing of application workload across designated resources, [Oracle Database](#) 10g Release 2 provides the Load Balancing Advisory. This Advisory monitors the current workload activity across the cluster and for each instance where a service is active; it provides a percentage value of how much of the total workload should be sent to this instance as well as service quality flag. The feedback is provided as an entry in the Automatic Workload Repository and a FAN event is published. The easiest way for an application to take advantage of the load balancing advisory, is to enable Runtime Connection Load Balancing with an integrated client.

215. How do I enable the load balancing advisory?

The load

balancing advisory requires the use of services and Oracle Net connection load balancing.

To enable it, on the server: set a goal (service_time or throughput, and set CLB_GOAL=SHORT) on your service.

For client, you must be using the connection pool.

For JDBC, enable the datasource parameter FastConnectionFailoverEnabled.

For ODP.NET enable the datasource parameter Load Balancing=true.

216. Why do we have a Virtual IP (VIP) in Oracle RAC 10g or 11g? Why does it just return a dead connection when its primary node fails?

The goal is application availability.

When a node fails, the VIP associated with it is automatically failed over to some other node. When this occurs, the following things happen.

- (1) VIP detects public network failure which generates a FAN event.
- (2) the new node re-arms the world indicating a new MAC address for the IP.
- (3) connected clients subscribing to FAN immediately receive ORA-3113 error or equivalent. Those not subscribing to FAN will eventually time out.
- (4) New connection requests rapidly traverse the tnsnames.ora address list skipping over the dead nodes, instead of having to wait on TCP-IP timeouts

Without using VIPs or FAN, clients connected to a node that died will often wait for a TCP timeout period (which can be up to 10 min) before getting an error.

As a result, you don't really have a good HA solution without using VIPs and FAN. The easiest way to use FAN is to use an integrated client with Fast Connection Failover (FCF) such as JDBC, OCI, or ODP.NET.

217. What are my options for setting the Load Balancing Advisory GOAL on a Service?

The load balancing advisory is enabled by setting the GOAL on your service either through PL/SQL DBMS_SERVICE package or EM DBControl Clustered Database Services [page](#). There are 3 options for GOAL:

None – Default setting, turn off advisory

THROUGHPUT – Work requests are directed based on throughput. This should be used when the work in a service completes at homogenous rates. An example is a trading system where work requests are similar lengths.

SERVICE_TIME – Work requests are directed based on response time. This should be used when the work in a service completes at various rates. An example is an internet shopping system where work requests are various lengths

Note: If using GOAL, you should set CLB_GOAL=SHORT

218. While executing root.sh power loss or pressed the CTLRT+C key, so what is next step to do?

Refer [Meta link](#): 220970.1

"Is it supported to rerun root.sh from the Oracle Clusterware installation ?

Rerunning root.sh after the initial successful install of the Oracle Clusterware is expressly discouraged and unsupported. We strongly recommend not doing it.

In case where root.sh is failing to execute for the on an initial install (or a new node joining an existing cluster), it is OK to re-run root.sh after the cause of the failure is corrected (permissions, paths, etc.). In this case, please run rootdelete.sh to undo the local effects of root.sh before re-running root.sh. "

If You want to deinstall Oracle Clusterware, You can follow this DOC to do it [Deinstalling Oracle Clusterware from a UNIX

Environment|http://download.oracle.com/docs/cd/E11882_01/em.112/e12255/oui5_cluster_environment.htm#BABBDDJBF]

Note: Running the rootcrs.pl command flags -deconfig -force enables you to deconfigure Oracle Clusterware on one or more nodes without removing installed binaries. This feature is useful if you encounter an error on one or more cluster nodes during installation when running the root.sh command, such as a missing operating system package on one node. By running rootcrs.pl -deconfig -force on nodes where you encounter an installation error, you can deconfigure Oracle Clusterware on those nodes, correct the cause of the error, and then run root.sh again.

219. How can you connect to a specific node in a RAC environment?

tnsnames.ora ensure that you have INSTANCE_NAME specified in it.

220. What is the Oracle Recommendation for backing up voting disk?

Oracle recommends us to use the dd command to backup the voting disk with a minimum block size of 4KB.

221. How can we add and remove multiple voting disks?

If we have multiple voting disks, then we can remove the voting disks and add them back into our environment using the following commands, where path is the complete path of the location where the voting disk resides:

```
crsctl delete css votedisk path
```

```
crsctl add css votedisk path
```

222. When can we use -force option?

If our cluster is down, then we can include the -force option to modify the voting disk configuration, without interacting with active Oracle Clusterware daemons. However, using the -force option while any cluster node is active may corrupt our configuration.

Preferred Sites:

<http://appsdbaera.blogspot.in/2013/03/rac-interview-questions.html#!/2013/03/rac-interview-questions.html>

http://lazyappsdba.blogspot.in/2010/08/rac-interview-q_24.html

<http://orajourn.blogspot.in/2007/06/rac-class-day-4.html>

<http://e-university.wisdomjobs.com/oracle-dba-interview-questions/oracle-dba-interview-questions/question-1.html>

Oracle RMAN FAQ

1. What is RMAN? Benefits of RMAN?
2. Which Files must be backed up?
3. When you take a hot backup putting Tablespace in begin backup mode, Oracle records SCN # from header of a database file. What happens when you issue hot backup database in RMAN at block level backup? How does RMAN mark the record that the block has been backed up? How does RMAN know what blocks were backed up so that it doesn't have to scan them again?
4. What are the Architectural components of RMAN?
5. What are Channels?
6. Why is the catalog optional?
7. What does complete RMAN backup consist of?
8. What is a Backup set?
9. What is a Backup piece?
10. What is the use of [RMAN Restore Preview](#)?
11. Where should the catalog be created?
12. How many times does oracle ask before dropping a catalog?
13. How to view the current defaults for the database?
14. How to resolve the ora-19804 error?
15. What are the various reports available with RMAN?
16. What does backup incremental level=0 database do?
17. What is the difference between DELETE INPUT and DELETE ALL command in backup?
18. How do I backup archive log?
19. How do I do an incremental backup after a base backup?
20. In catalog database, if some of the blocks are corrupted do the system crash, How will you recover?
21. You have taken a manual backup of a datafile using o/s. How RMAN will know about it?
22. Where RMAN keeps information of backups if you are using RMAN without Catalog?
23. How do you see information about backups in RMAN?
24. How RMAN improves backup time?
25. List the encryption options available with RMAN?
26. What are the steps required to perform in \$ORACLE_HOME for enabling the RMAN backups with netbackup or TSM tape library software?
27. What is the significance of incarnation and DBID in the RMAN backups?
28. List at least 6 advantages of RMAN backups compare to traditional hot backups?
29. How do you enable the autobackup for the controlfile using RMAN?
30. How do you identify what are the all the target databases that are being backed-up with RMAN database?
31. What is the difference between cumulative incremental and differential incremental backups?
32. How do you identify the block corruption in RMAN database? How do you fix it?
33. How do you clone the database using RMAN software? Give brief steps? When do you use crosscheck command?
34. What is the difference between obsolete RMAN backups and expired RMAN backups?
35. List some of the RMAN catalog view names which contain the catalog information?
36. What is db_recovery_file_dest ? When do you need to set this value?
37. How do you setup the RMAN tape backups?
38. How do you install the RMAN recovery catalog?
39. When do you recommend hot backup? What are the pre-reqs?
40. What is the difference between physical and logical backups?
41. What is RAID? What is RAID0? What is RAID1? What is RAID 10?
42. What are things which play major role in designing the backup strategy?
43. What is hot backup and what is cold backup?
44. How do you test that your recovery was successful?
45. How do you backup the Flash Recovery Area?
46. How to enable Fast Incremental Backup to backup only those data blocks that have changed?

47. How do you set the flash recovery area?
48. How can you use the CURRENT_SCN column in the V\$DATABASE view to obtain the currentSCN?
49. How do you identify the expired, active, obsolete backups? Which RMAN command you use?
50. Explain how to setup the physical standby database with RMAN?
51. What is auxiliary channel in RMAN? When do you need this?
52. What is RMAN and how does one use it?
53. What kind of backup are supported by RMAN?
54. What is the Flash Recovery Area?
55. How do you use the V\$RECOVERY_FILE_DEST view to display information regarding the flash recovery area?
56. How can you display warning messages?
57. How to use the best practice to use Oracle Managed File (OMF) to let Oracle database to create and manage the underlying operating system files of a database?
58. How do you monitor block change tracking?
59. How do you use the V\$BACKUP_DATAFILE view to display how effective the block change tracking is in minimizing the incremental backup I/O?
60. How do you backup an individual tablespaces?
61. How do you backup datafiles and control files?
62. Use a fast recovery without restoring all backups from their backup location to the location specified in the controlfile?
63. What are the Oracle Enhancements for RMAN in 10g?
64. What are the benefits of Global Scripting in RMAN?
65. What is FRA? When do we use?
66. What is Channel? How do you enable the parallel backups with RMAN?
67. What are RTO, MTBF, and MTTR?
68. How do you enable the encryption for RMAN backups?
69. What is the difference between restoring and recovering?
70. What are the various tape backup solutions available in the market?
71. Outline the steps for recovering the full database from cold backup?
72. Explain the steps to perform the point in time recovery with a backup which is taken before the resetlogs of the db?
73. Outline the steps involved in TIME based recovery from the full database from hot backup?
74. Is it possible to take Catalog Database Backup using RMAN? If Yes How?
75. Can a schema be restored in Oracle 9i RMAN when the schema having numerous table spaces?
76. Outline the steps for changing the DBID in a cloned environment?
77. How do you identify the expired, active, obsolete backups? Which RMAN command you use?
78. Explain how to setup the physical standby database with RMAN?
79. List the steps required to enable the RMAN backup for a target database?
80. How do you verify the integrity of the image copy in RMAN environment?
81. Outline the steps involved in SCN based recovery from the full database from hot backup?
82. Outline the steps involved in CANCEL based recovery from the full database from hot backup?
83. Is it possible to store specific tables when using RMAN DUPLICATE feature? If yes how?
84. Difference between catalog and nologcat?
85. Outline the steps for recovery of missing data file?
86. Outline the steps for recovery with missing online redo logs?
87. Outline steps for recovery with missing archived redo logs?
88. Difference between using recovery catalog and control file?
89. Can we use same target database as catalog?
90. How do you know how much RMAN task has been completed?
91. From where list & report commands will get input?
92. Command to delete archive logs older than 7 days?
93. How many days backup, by default RMAN stores?
94. What is the use of crosscheck command in RMAN?
95. What are the differences between crosscheck and validate commands?
96. Which is one is good, differential (incremental) backup or cumulative (incremental) backup?

97. What is Level 0, Level 1 backup?
98. Can we perform level 1 backup without level 0 backup?
99. Will RMAN put the database/tablespace/datafile in backup mode?
100. What is snapshot control file?
101. What is the difference between backup set and backup piece?
102. How to do cloning by using RMAN?
103. You loss one datafile and DB is running in ARCHIVELOG mode. You have full database backup of week/day old and don't have backup of this (newly created) datafile. How do you restore/recover file?
104. What is obsolete backup & expired backup?
105. What is the difference between hot backup & RMAN backup?
106. How to put manual/user-managed backup in RMAN (recovery catalog)?
107. What are new features in Oracle 11g RMAN?
108. What is the difference between auxiliary channel and maintenance channel?
109. Can we take rman backup if database is in no archivelog mode?
110. What is incarnation?
111. How do I go about backing up my online redo logs?
112. Outline the steps for recovery with missing online redo logs?
113. Outline steps for recovery with missing archived redo logs?
114. What is RMAN and How to configure it?

1. What is RMAN? Benefits of RMAN?

Recovery Manager (RMAN) is a utility that can manage your entire Oracle backup and recovery activities.

Benefits-1:

- Incremental backups that only copy data blocks that have changed since the last backup
- Tablespaces are not put in backup mode, thus there is no extra redo log generation during online backups.
- Detection of corrupt blocks during backups
- Parallelization of I/O operations
- Automatic logging of all backup and recovery operations
- Built-in reporting and listing commands

Benefits-2:

Central Repository

Incremental Backup

Corruption Detection

Advantage over tradition backup system:

Copies only the filled blocks i.e. even if 1000 blocks is allocated to datafile but 500 are filled with data then RMAN will only create a backup for that 500 filled blocks.

Incremental and accumulative backup

Catalog and no catalog option

Detection of corrupted blocks during backup;

Can create and store the backup and recover scripts.

Increase performance through automatic parallelization (allocating channels) less redo generation.

Benefits-3:

No Extra Costs, It is available free.

RMAN introduced in Oracle 8 it has become simpler with new version and easier than user managed backups.

Proper Security

You are 100% sure your database has been backed up.

It contains details of backup taken in the central repository

Facility of testing validity of backups also commands like cross check to check the status of backup.

Oracle 10g has got further optimized incremental backups which has resulted in improvement of performance during backup and recovery time

Parallel operation are supported

Better Querying facility for knowing different details of backup

No Extra redo generated when backup is taken compared to online backup

Without rman which results in saving of space in hard disk

RMAN is an intelligent tool

Maintains repository of backup metadata

Remembers backup locations

Knows what needs backup set locations

Knows what needs to be backed up

Knows what is required for recovery

Know what backups are redundant

It handles database corruptions

2. Which Files must be backed up?

Database Files (with RMAN)

Control Files (with RMAN)

Offline Redolog Files (with RMAN)

INIT.ORA (manually)

Password Files (manually)

3. When you take a hot backup putting Tablespace in begin backup mode, Oracle records SCN # from header of a database file. What happens when you issue hot backup database in RMAN at block level backup? How does RMAN mark the record that the block has been backed up? How does RMAN know what blocks were backed up so that it doesn't have to scan them again?

In 11g, there is Oracle Block Change Tracking feature. Once enabled; this new 10g feature records the modified since last backup and stores the log of it in a block change tracking file. During backups RMAN uses the log file to identify the specific blocks that must be backed up. This improves RMAN's performance as it does not have to scan whole datafiles to detect changed blocks.

Logging of changed blocks is performed by the CTRW process which is also responsible for writing data to the block change tracking file. RMAN uses SCNs on the block level and the archived redo logs to resolve any inconsistencies in the datafiles from a hot backup. What RMAN does not require is to put the tablespace in BACKUP mode, thus freezing the SCN in the header. Rather, RMAN keeps this information in either your control files or in the RMAN repository (i.e., Recovery Catalog).

4. What are the Architectural components of RMAN?

- RMAN executable
- Server processes
- Channels
- Target database
- Recovery catalog database (optional)
- Media management layer (optional)
- Backups, backup sets, and backup pieces

5. What are Channels?

A channel is an RMAN server process started when there is a need to communicate with an I/O device, such as a disk or a tape. A channel is what reads and writes RMAN backup files. It is through the allocation of channels that you govern I/O characteristics such as:

Type of I/O device being read or written to, either a disk or a sbt_tape

Number of processes simultaneously accessing an I/O device

Maximum size of files created on I/O devices

Maximum rate at which database files are read

Maximum number of files opens at a time

6. Why is the catalog optional?

Because RMAN manages backup and recovery operations, it requires a place to store necessary information about the database. RMAN always stores this information in the target database control file. You can also store RMAN metadata in a recovery catalog schema contained in a separate database. The recovery catalog schema must be stored in a database other than the target database.

7. What does complete RMAN backup consist of?

A backup of all or part of your database

A backup consists of one or more backup sets

8. What is a Backup set?

A logical grouping of backup files -- the backup pieces -- that are created when you issue an RMAN backup command.

A backup set is RMAN's name for a collection of files associated with a backup. A backup set is composed of one or more backup pieces.

9. What is a Backup piece?

A physical binary file created by RMAN during a backup. Backup pieces are written to your backup medium, whether to disk or tape. They contain blocks from the target database's datafiles, archived redo log files, and control files.

When RMAN constructs a backup piece from datafiles, there are a several rules that it follows:

- A datafile cannot span backup sets
 - A datafile can span backup pieces as long as it stays within one backup set
 - Datafiles and control files can coexist in the same backup sets
 - Archived redo log files are never in the same backup set as datafiles or control files
- RMAN is the only tool that can operate on backup pieces. If you need to restore a file from an RMAN backup, you must use RMAN to do it. There's no way for you to manually reconstruct database files from the backup pieces. You must use RMAN to restore files from a backup piece.

10. What is the use of RMAN Restore Preview?

The PREVIEW option of the RESTORE command allows you to identify the backups required to complete a specific restore operation. The output generated by the command is in the same format as the LIST command. In addition the PREVIEW SUMMARY command can be used to produce a summary report with the same format as the LIST SUMMARY command. The following examples show how these commands are used:

```
# Spool output to a log file
SPOOL LOG TO c:\oracle\rmancmd\restorepreview.lst;
# Show what files will be used to restore the SYSTEM tablespace's datafile
RESTORE DATAFILE 2 PREVIEW;
# Show what files will be used to restore a specific tablespace
RESTORE TABLESPACE users PREVIEW;
# Show a summary for a full database restores
RESTORE DATABASE PREVIEW SUMMARY;
# Close the log file
SPOOL LOG OFF;
```

11. Where should the catalog be created?

The recovery catalog to be used by rman should be created in a separate database other than the target database. The reason was that the target database will be shutdown while datafiles are restored.

12. How many times does oracle ask before dropping a catalog?

The default is two times one for the actual command, the other for confirmation.

13. How to view the current defaults for the database?

```
RMAN> show all;
```

14. How to resolve the ora-19804 error?

Basically this error is because of flash recovery area been full. One way to solve is to increase the space available for flashback database.

```
SQL>ALTER SYSTEM SET DB_RECOVERY_FILE_DEST_SIZE=5G; --It can be set to K, M or G
```

15. What are the various reports available with RMAN?

```
RMAN> list backup;
rman> list archive;
```

16. What does backup incremental level=0 database do?

Backup database level=0 is a full backup of the database.

```
RMAN> backup incremental level=0 database;
```

you can also use backup full database; which means the same thing as level=0;

17. What is the difference between DELETE INPUT and DELETE ALL command in backup?

Generally speaking LOG_ARCHIVE_DEST_n points to two disk drive locations where we archive the files, when a command is issued through rman to backup archivelogs it uses one of the location to backup the data. When we specify delete input the location which was backedup will get deleted, if we specify delete all all log_archive_dest_n will get deleted.

DELETE all applies only to archived logs.

```
RMAN> Delete expired archivelog all;
```

18. How do I backup archive log?

```
RMAN> backup archivelog all;
RMAN> backup archivelog all delete input;
RMAN> backup archivelog all delete all;
```

19. How do I do an incremental backup after a base backup?

```
RMAN> backup incremental level 1 database;
```

20. In catalog database, if some of the blocks are corrupted do the system crash, how will you recover?

Using RMAN BLOCK RECOVER command

21. You have taken a manual backup of a datafile using o/s. How RMAN will know about it?

You have to catalog that manual backup in RMAN's repository by command

```
RMAN> catalog datafilecopy '/DB01/BACKUP/users01.dbf';
```

Restrictions:

- > Accessible on disk
- > A complete image copy of a single file

22. Where RMAN keeps information of backups if you are using RMAN without Catalog?

RMAN keeps information of backups in the control file.

CATALOG vs NOCATALOG

The difference is only who maintains the backup records like when is the last successful backup incremental differential etc.

In CATALOG mode another database (TARGET database) stores all the information.

In NOCATALOG mode controlfile of Target database is responsible.

23. How do you see information about backups in RMAN?

RMAN> List Backup;

Use this SQL to check

SQL> SELECT Sid totalwork sofar FROM v\$session_longops WHERE Sid =153;

24. How RMAN improves backup time?

RMAN backup time consumption is very less than compared to regular online backup as RMAN copies only modified blocks

25. List the encryption options available with RMAN?

RMAN offers three encryption modes: transparent mode, password mode and dual mode.

26. What are the steps required to perform in \$ORACLE_HOME for enabling the RMAN backups with netbackup or TSM tape library software?

All the steps to take an rman backup with TSM tape library as follows

1. Install TDPO (default path /usr/Tivoli/tsm/client/oracle/)

2. Once u installed the TDPO automatically one link is created from TDPO directory to /usr/lib

Now we need to create soft link between OS to ORACLE_HOME

\$ ln -s /usr/lib/libiobk64.a \$ORACLE_HOME/lib/libobk.a(very important)

3. Uncomment and Modify tdpo.opt file which in

/usr/tivoli/tsm/client/oracle/bin/tdpo.opt as follows

DSMI_ORC_CONFIG /usr/Tivoli/tsm/client/oracle/bin64/dsm.opt

DSMI_LOG /home/tmp/oracle

TDPO_NODE backup

TDPO_PSWDPATH /usr/tivoli/tsm/client/oracle/bin64

4. Create dsm.sys file in same path and add the entries

SErvername <Server name >

TCPPort 1500

Passwordaccess prompt

Nodename backup

Enablelanfree yes

TCPSERVERADDRESS <Server Address>

5. Create dsm.opt file add an entry

SErvername <Server name >

6. Then take backup

RMAN>run

{

allocate channel t1 type 'sbt_tape' parms

'ENV (TDPO_OPTFILE /usr/tivoli/tsm/client/oracle/bin64/tdpo.opt)';

backup database include current controlfile;

release channel t1;

}

27. What is the significance of incarnation and DBID in the RMAN backups?

When you have multiple databases you have to set your DBID (Database Id) which is unique to each database. You have to set this before you do any restore operation from RMAN.

There is possibility that incarnation may be different of your database. So it is advised to reset to match with the current incarnation. If you run the RMAN command ALTER DATABASE OPEN RESETLOGS then RMAN resets the target database automatically so that you do not have to run RESET DATABASE. By resetting the database RMAN considers the new incarnation as the current incarnation of the database.

28. List at least 6 advantages of RMAN backups compare to traditional hot backups?

RMAN has the following advantages over Traditional backups:

1. Ability to perform INCREMENTAL backups

2. Ability to recover one block of datafile
3. Ability to automatically backup CONTROLFILE and SPFILE
4. Ability to delete the older ARCHIVE REDOLOG files, with the new one's automatically.
5. Ability to perform backup and restore with parallelism.
6. Ability to report the files needed for the backup.
7. Ability to RESTART the failed backup, without starting from beginning.
8. Much faster when compared to other TRADITIONAL backup strategies.

29. How do you enable the autobackup for the controlfile using RMAN?

Issue command at rman prompt.....

RMAN> configure controlfile autobackup on;

RMAN> configure controlfile autobackup format for device type disk to '\$HOME/BACKUP/RMAN/F.bkp';

30. How do you identify what are all the target databases that are being backed-up with RMAN database?

You don't have any view to identify whether it is backed up or not. The only option is connect to the target database and give list backup this will give you the backup information with date and timing.

31. What is the difference between cumulative incremental and differential incremental backups?

Differential backup: This is the default type of incremental backup which backs up all blocks changed after the most recent backup at level n or lower.

Cumulative backup: Backup all blocks changed after the most recent backup at level n-1 or lower.

32. How do you identify the block corruption in RMAN database? How do you fix it?

Using v\$sql_block_corruption view u can find which blocks corrupted.

Rman>> block recover datafile <fileid> block <blockid>;

Using the above statement u recover the corrupted blocks.

First check whether the block is corrupted or not by using this command

SQL> select file# block# from v\$database_block_corruption;

File# block

2 507

The above block is corrupted...

Connect to Rman to recover the block use this command...

Rman>blockrecover datafile 2 block 507;

The above command recovers the block 507

Verify: Rman>blockrecover corruption list;

33. How do you clone the database using RMAN software? Give brief steps? When do you use crosscheck command?

Check whether backup pieces proxy copies or disk copies still exist.

Two commands available in RMAN to clone database:

1) Duplicate

2) Restore.

34. What is the difference between obsolete RMAN backups and expired RMAN backups?

The term obsolete does not mean the same as expired. In short obsolete means "not needed" whereas expired means "not found."

35. List some of the RMAN catalog view names which contain the catalog information?

RC_DATABASE_INCARNATION

RC_BACKUP_COPY_DETAILS

RC_BACKUP_CORRUPTION

RC_BACKUP-DATAFILE_SUMMARY to name a few

36. What is db_recovery_file_dest? When do you need to set this value?

If Database Flashback option is on then use this option.

37. How do you setup the RMAN tape backups?

RMAN Target /

run

{

Allocate channel ch1 device type sbt_tape maxpiecesize 4g

Format 'D_U_T_t';

sql 'alter system switch logfile';

Backup database;

```

backup archivelog from time 'sysdate-7';
Backup Format ' D_CTLFILE_P_ U_ T_ t' Current controlfile;
release channel ch1;
}

```

This is backup script for Tivoli Backup Server

38. How do you install the RMAN recovery catalog?

Steps to be followed:

- 1) Create connection string at catalog database.
- 2) At catalog database create one new user or use existing user and give that user a recovery_catalog_owner privilege.
- 3) Login into RMAN with connection string
 - a) Export ORACLE_SID
 - b) Rman target catalog @connection string
- 4) Rman> create catalog;
- 5) Register database;

39. When do you recommend hot backup? What are the pre-reqs?

Database must be Archivelog Mode

Archive Destination must be set and LOG_ARCHIVE_START TRUE (EARLIER VERSION BEFORE 10G)

40. What is the difference between physical and logical backups?

In Oracle Logical Backup is "which is taken using either Traditional Export/Import or Latest Data Pump". Where as Physical backup is known "when you take Physical O/s Database related Files as Backup".

41. What is RAID? What is RAID0? What is RAID1? What is RAID 10?

RAID: It is a redundant array of independent disk

RAID0: Concatenation and stripping

RAID1: Mirroring

42. What are things which play major role in designing the backup strategy?

Designing a good backup strategy it will not only be simply backup but also a contingency plan. In this case you should consider the following:

1. How long is the allowable down time during recovery? - If short you could consider using dataguard.
2. How long is the backup period? - If short I would advise to use RMAN instead of user managed backup.
3. If limited disk space for backup never use user managed backup.
4. If the database is large you could consider doing full rman backups on a weekend and do a incremental backup on a weekday.
5. Schedule your backup on the time where there is least database activity this is to avoid resource hugging.
6. Backup script should always be automized via scheduled jobs. This way operator would never miss a backup period.
7. Retention period should also be considered. Try keeping atleast 2 full backups (current and previous backup).

Cold backup: shutdown the database and copy the datafiles with the help of

O.S. Command: This is simply copying of datafiles just like any other text file copy.

Hot backup: backup process starts even though database in running. The process to take a hot backup is

- 1) sql> alter database begin backup;
- 2) copy the datafiles.
- 3) after copying

```
sql> alter database end backup;
```

Begin backup clause will generate the timestamp. it'll be used in backup consistency i.e. when begin backup pressed it'll generate the timestamp. During restore database will restore the data from backup till that timestamp and remaining backup will be recovered from archive log.

43. What is hot backup and what is cold backup?

Hot backup when the database is online cold backup is taken during shut down period

44. How do you test that your recovery was successful?

```
SQL> SELECT count(*) FROM flashback_table;
```

45. How do you backup the Flash Recovery Area?

```
RMAN> BACKUP RECOVERY FILES;
```

The files on disk that have not previously been backed up will be backed up. They are full and incremental backup sets, control file auto-backups, archive logs and datafile copies.

46. How to enable Fast Incremental Backup to backup only those data blocks that have changed?

```
SQL> ALTER DATABASE enable BLOCK CHANGE TRACKING;
```

47. How do you set the flash recovery area?

```
SQL> ALTER SYSTEM SET db_recovery_file_dest_size = 100G;
```

```
SQL> ALTER SYSTEM SET db_recovery_file_dest = '/u10';
```

48. How can you use the CURRENT_SCN column in the V\$DATABASE view to obtain the currentSCN?

```
SQL> SELECT current_scn FROM v$database;
```

49. How do you identify the expired, active, obsolete backups? Which RMAN command you use?

Use command:

```
Rman > crosscheck backup;
```

```
Rman > crosscheck archivelog all;
```

```
Rman > listbackup;
```

```
Rman > list archive logall
```

50. Explain how to setup the physical stand by database with RMAN?

```
$ Export ORACLE_SID=TEST $ rman target /
```

```
RMAN> show all;
```

Using target database controlfile instead of recovery catalog RMAN configuration parameters are:

```
CONFIGURE RETENTIONPOLICY TO RECOVERY WINDOW OF 1 DAYS;
```

```
CONFIGURE BACKUP OPTIMIZATION
```

51. What is auxiliary channel in RMAN? When do you need this?

An auxiliary channel is a link to auxiliary instance. If you do not have automatic channels configured, then before issuing the DUPLICATE command, manually allocate at least one auxiliary channel with in the same RUN command.

52. What is RMAN and how does one use it?

Recovery Manager (or RMAN) is an Oracle provided utility for backing-up, restoring and recovering Oracle Databases. RMAN ships with the database server and doesn't require a separate installation. The RMAN executable is located in your ORACLE_HOME/bin directory.

53. What kind of backup are supported by RMAN?

Backup Sets
Datafiles
Copies
OS Backup

54. What is the Flash Recovery Area?

It is a unified storage location for all recovery-related files and activities in an Oracle Database. It includes Control File, Archived Log Files, Flashback Logs, Control File Autobackups, Data Files, and RMAN files.

55. How do you use the V\$RECOVERY_FILE_DEST view to display information regarding the flash recovery area?

```
SQL> SELECT name, space_limit, space_used, space_reclaimable, number_of_files FROM v$recovery_file_dest;
```

56. How can you display warning messages?

```
SQL> SELECT object_type, message_type, message_level, reason, suggested_action FROM dba_outstanding_alerts;
```

57. How to use the best practice to use Oracle Managed File (OMF) to let Oracle database to create and manage the underlying operating system files of a database?

```
SQL> ALTER SYSTEM SET db_create_file_dest = '/u01/oradata/feroz';
```

```
SQL> ALTER SYSTEM SET db_create_online_dest_1 = '/u02/oradata/feroz';
```

58. How do you monitor block change tracking?

```
SQL> SELECT filename, status, bytes FROM v$block_change_tracking;
```

It shows where the block change-tracking file is located, the status of it and the size.

59. How do you use the V\$BACKUP_DATAFILE view to display how effective the block change tracking is in minimizing the incremental backup I/O?

```
SQL> SELECT file#, AVG(datafile_blocks), AVG(blocks_read), AVG (blocks_read/datafile_blocks), AVG(blocks) FROM v$backup_datafile WHERE used_change_tracking = 'YES' AND incremental_level > 0 GROUP BY file#; If the AVG (blocks_read/datafile_blocks) column is high then you may have to decrease the time between the incremental backups.
```

60. How do you backup an individual tablespaces?

```
RMAN> CONFIGURE DEFAULT DEVICE TYPE TO DISK;
```

```
RMAN> BACKUP TABLESPACE system;
```

61. How do you backup datafiles and control files?

```
RMAN> BACKUP DATAFILE 3;
```

RMAN> BACKUP CURRENT CONTROLFILE;

62. Use a fast recovery without restoring all backups from their backup location to the location specified in the controlfile?

RMAN> SWITCH DATABASE TO COPY;

63. What are the Oracle Enhancement for RMAN in 10g?

Flash Recovery Area

Incrementally Updated Backups

Faster Incremental Backups

SWITCH DATABASE COMMAND.

Binary Compression

Global Scripting

Duration Clause

Configure This

Oracle Enhancement for Rman in 10g

Automatic Channel Failover

Compress Backup Sets

Recovery through Reset Logs

Cross Backup Sets

64. What are the benefits of Global Scripting in RMAN?

Oracle Database 10g provides a new concept of global scripts, which you can execute against any database registered in the recovery catalog, as long as your RMAN client is connected to the recovery catalog and a target database simultaneously. CPISOLUTION.COM

Example of local and global scripts:

RMAN> print script full_backup to file 'my_script_file.txt'

RMAN> create global script global_full_backup

65. What is FRA? When do we use?

Flash recovery area where you can store not only the traditional components found in a backup strategy such as control files, archived log files, and Recovery Manager (RMAN) datafile copies but also a number of other file components such as flashback logs. The flash recovery area simplifies backup operations, and it increases the availability of the database because many backup and recovery operations using the flash recovery area can be performed when the database is open and available to users.

Because the space in the flash recovery area is limited by the initialization parameter DB_RECOVERY_FILE_DEST_SIZE the Oracle database keeps track of which files are no longer needed on disk so that they can be deleted when there is not enough free space for new files. Each time a file is deleted from the flash recovery area, a message is written to the alert log.

A message is written to the alert log in other circumstances. If no files can be deleted, and the recovery area used space is at 85 percent, a warning message is issued. When the space used is at 97 percent, a critical warning is issued. These warnings are recorded in the alert log file, are viewable in the data dictionary view

DBA_OUTSTANDING_ALERTS, and are available to you on the main page of the EM Database Control

66. What is Channel? How do you enable the parallel backups with RMAN?

Channel is a link that RMAN requires to link to target database. This link is required when backup and recovery operations are performed and recorded. This channel can be allocated manually or can be preconfigured by using automatic channel allocation.

The number of allocated channels determines the maximum degree of parallelism that is used during backup, restore or recovery. For example, if you allocate 4 channels for a backup operation, 4 background processes for the operation can run concurrently.

Parallelization of backup sets allocates multiple channels and assigns files to specific channels. You can configure parallel backups by setting a PARALLELISM option of the CONFIGURE command to a value greater than 1 or by manually allocating multiple channels.

RMAN> CONFIGURE DEVICE TYPE PARALLELISM 2 BACKUP TYPE TO COMPRESSED BACKUPSET;

67. What are RTO, MTBF, and MTTR?

RTO: Recovery Time objective-is the maximum amount of time that the database can be unavailable and still satisfy SLA's

MTBF (Mean Time Between Failure)-

MTTR (Mean Time to recover)- fast recovery solutions

68. How do you enable the encryption for RMAN backups?

If you wish to modify your existing backup environment so that all RMAN backups are encrypted, perform the following steps:

- Set up the Oracle Encryption Wallet
- Issue the following RMAN command:

```
RMAN> CONFIGURE ENCRYPTION ALGORITHM 'AES256'; -- use 256 bit encryption
```

```
RMAN> CONFIGURE ENCRYPTION FOR DATABASE ON; -- encrypt backups
```

69. What is the difference between restoring and recovering?

Restoring involves copying backup files from secondary storage (backup media) to disk. This can be done to replace damaged files or to copy/move a database to a new location.

Recovery is the process of applying redo logs to the database to roll it forward. One can roll-forward until a specific point-in-time (before the disaster occurred), or roll-forward until the last transaction recorded in the log files.

```
SQL> connect SYS as SYSDBA
```

```
SQL> RECOVER DATABASE UNTIL TIME '2001-03-06:16:00:00' USING BACKUP CONTROLFILE;
```

```
RMAN> run {
```

```
    set until time to _date('04-Aug-2004 00:00:00', 'DD-MON-YYYY HH24:MI:SS');
```

```
    restore database;
```

```
    recover database;
```

```
}
```

70. What are the various tape backup solutions available in the market?

71. Outline the steps for recovering the full database from cold backup?

72. Explain the steps to perform the point in time recovery with a backup which is taken before the resetlogs of the db?

73. Outline the steps involved in TIME based recovery from the full database from hot backup?

74. Is it possible to take Catalog Database Backup using RMAN? If Yes How?

75. Can a schema be restored in oracle 9i RMAN when the schema having numerous table spaces?

76. Outline the steps for changing the DBID in a cloned environment?

77. How do you identify the expired, active, obsolete backups? Which RMAN command you use?

78. Explain how to setup the physical stand by database with RMAN?

79. List the steps required to enable the RMAN backup for a target database?

80. How do you verify the integrity of the image copy in RMAN environment?

81. Outline the steps involved in SCN based recovery from the full database from hot backup?

82. Outline the steps involved in CANCEL based recovery from the full database from hot backup?

83. Is it possible to store specific tables when using RMAN DUPLICATE feature? If yes how?

84. Difference between catalog and nologcat?

85. Outline the steps for recovery of missing data file?

86. Outline the steps for recovery with missing online redo logs?

87. Outline steps for recovery with missing archived redo logs?

88. Difference between using recovery catalog and control file?

When new incarnation happens, the old backup information in control file will be lost. It will be preserved in recovery catalog.

In recovery catalog, we can store scripts.

Recovery catalog is central and can have information of many databases.

89. Can we use same target database as catalog?

No. The recovery catalog should not reside in the target database (database to be backed up), because the database can't be recovered in the mounted state.

90. How do u know how much RMAN task has been completed?

By querying v\$rman_status or v\$session_longops

91. from where list & report commands will get input?

92. Command to delete archive logs older than 7days?

```
RMAN> delete archivelog all completed before sysdate-7;
```

93. How many days backup, by default RMAN stores?

94. What is the use of crosscheck command in RMAN?

Crosscheck will be useful to check whether the catalog information is intact with OS level information.

95. What are the differences between crosscheck and validate commands?

96. Which is one is good, differential (incremental) backup or cumulative (incremental) backup?

A differential backup, which backs up all blocks changed after the most recent incremental backup at level 1 or 0

A cumulative backup, which backs up all blocks changed after the most recent incremental backup at level 0

97. What is Level 0, Level 1 backup?

A level 0 incremental backup, which is the base for subsequent incremental backups, copies all blocks containing data, backing the datafile up into a backup set just as a full backup would. A level 1 incremental backup can be either of the following types:

A differential backup, which backs up all blocks changed after the most recent incremental backup at level 1 or 0

A cumulative backup, which backs up all blocks changed after the most recent incremental backup at level 0

98. Can we perform level 1 backup without level 0 backup?

If no level 0 backup is available, then the behavior depends upon the compatibility mode setting. If compatibility < 10.0.0, RMAN generates a level 0 backup of the file contents at the time of the backup. If compatibility is >= 10.0.0, RMAN copies all blocks changed since the file was created, and stores the results as a level 1 backup. In other words, the SCN at the time the incremental backup is taken is the file creation SCN.

99. Will RMAN put the database/tablespace/datafile in backup mode?

No

100. What is snapshot control file?

101. What is the difference between backup set and backup piece?

Backup set is logical and backup piece is physical.

102. How to do cloning by using RMAN?

RMAN> duplicate target database

103. You loss one datafile and DB is running in ARCHIVELOG mode. You have full database backup of 1 week/day old and don't have backup of this (newly created) datafile. How do you restore/recover file?

Create the datafile and recover that datafile.

SQL> alter database create datafile '...path..' size n;

RMAN> recover datafile file_id;

104. What is obsolete backup & expired backup?

A status of "expired" means that the backup piece or backup set is not found in the backup destination.

A status of "obsolete" means the backup piece is still available, but it is no longer needed. The backup piece is no longer needed since RMAN has been configured to no longer need this piece after so many days have elapsed, or so many backups have been performed.

105. What is the difference between hot backup & RMAN backup?

For hot backup, we have to put database in begin backup mode, then take backup.

RMAN won't put database in backup mode.

106. How to put manual/user-managed backup in RMAN (recovery catalog)?

By using catalog command.

RMAN> CATALOG START WITH '/tmp/backup.ctl';

107. What are new features in Oracle 11g RMAN?

108. What is the difference between auxiliary channel and maintenance channel?

109. Can we take rman backup if database is in no archive log mode?

110. What is incarnation?

111. How do I go about backing up my online redo logs?

You don't. Online redo logs should never, ever be included in a backup, regardless of whether that backup is performed hot or cold. The reasons for this are two-fold. First, you physically cannot backup a hot online redo log, and second there is precisely zero need to do so in the first place because an archive redo log is, by definition, a backup copy of a formerly on-line log. There is, however, a more practical reason: backing up the online logs yourself increases the risk that you will lose committed data.

In short, it's not necessary, it's not possible, and it's dangerous to even try.

Can you explain why it's not possible a bit more?

I can, of course, issue 'copy *.log' commands at the command prompt whenever it suits me, so it would seem to be very possible for me.

The golden rule in Oracle backups is: you cannot ever copy anything hot without the resulting copy being, internally, complete garbage. That's because it takes a finite amount of time to copy a file, and during that time the contents of the original may change. The copy will therefore end up with bits of itself at one time and bits at other times.

Now, Oracle provides a mechanism to patch up that sort of mess when it encounters it inside a copied-then-restored

data file. The mechanism is called “recovery”, and it works by applying redo to the internal bits of the data file so that the oldest bits get rolled forward till they ‘catch up’ to the youngest bits. Eventually, the entire file gets to one, consistent, point of time and can be rolled forward from there. In other words, redo makes internally inconsistent data files internally consistent and usable.

But you only apply redo to data files. A recovery does not apply redo to control files. Instead, Oracle provides a different mechanism to permit hot backups of the control file: alter database backup controlfile to ‘c:\somewhere’. This is a SQL command that generates a read-consistent image of the control file. It’s guaranteed by Oracle to be internally consistent in the first place, so it needs no redo applied to it to make it usable.

In both of these cases, therefore, Oracle has provided a mechanism to make it possible to take hot copies of data files and control files -one to prevent internal inconsistencies in the first place, and one to sort them out after they are encountered. But here’s the punch line: neither of those mechanisms applies to online redo logs. Since there is no ‘repair inconsistencies’ or ‘avoid inconsistencies’ mechanism for online redo logs, it follows that they cannot be copied hot without the contents of the copies being immediately rendered 100% useless.

Right, I can understand not copying them if I’m doing hot backups. I can see that now. But actually, I’m doing cold backups. Presumably these considerations don’t apply, and I could take a copy of the files if I wanted to?

You could, it’s true. A cold file can be copied at your leisure: by definition, its contents aren’t going to change whilst you’re copying it, so there is no risk of internal corruption or inconsistency in the copy. But you still have to face the “it’s pointless” argument.

For a start, you’re presumably in archive log mode, and ARCn has been busy taking copies of every one of your online logs for you. You shutting down the database and taking a fresh copy of one or two of them doesn’t exactly bring much that’s new or beneficial to the party. What’s more, the second you open up your database after taking the new backup, the contents of your backup are out of date.

Third, and most important, you never need to restore online redo logs to perform any database recovery -so what is the point of having a backup of something you will never actually need to use?

But it’s not actually true, is it, that ARCn has copied every log already -because it doesn’t copy the CURRENT online log, until after it ceases to be the current one, does it?

Quite right. There is always one log (and only one log, incidentally) which definitely hasn’t been archived by ARCn yet, and that’s the one which is currently being written to by LGWR (and hence has a status of CURRENT in the V\$LOG dynamic performance view). But try and think logically: if that log is truly ‘CURRENT’, it must be in use and is therefore hot... and you can’t copy it because the copy will be internally inconsistent and unusable. If the database is shut down, then the current log isn’t truly ‘current’ (because there’s no LGWR to write to it!). It’s cold, and could be copied -but you never need to copy it, because the copy would be instantly out of date once you restart the instance, and in any case you never need to restore online redo logs under any circumstances. Either way you look at it, you either can’t copy the current log, or there’s no point in doing so.

What you are really saying, of course, is that you are worried that ARCn has not yet copied the CURRENT log, and are feeling a bit nervous about that. I totally agree that you should be nervous about this: the current redo log is definitely the weakest point in the entire Oracle architecture, and its loss would indeed result in committed data being lost.

But those are not grounds to copy the current log. Those are grounds to make sure you never lose the current log in the first place -and to do that, you should be employing hardware mirroring and Oracle multiplexing (making each log group consist of multiple members).

OK, I accept that it is pointless and unnecessary. But it won’t exactly do any harm if I do back them up (cold!) will it? Yes it will, actually. Or, rather, yes it could.

Imagine a production environment in which you have, despite all advice to the contrary, taken backups of your online redo logs (hot or cold, it makes no difference). Suppose you lose a small and fairly unimportant data file from that system. Recovery should be a piece of cake: restore the damaged data file, recover datafile X, alter tablespace X online. All committed data back, no sweat. Now imagine that in the heat of the moment, you were to restore *.dbf, instead of just x.dbf... well, recovery takes a lot longer, and the database is down for the duration, completely unnecessarily, because SYSTEM has to be recovered. But you still get all your committed data back, no worries. A totally successful recovery that just wasn’t performed as efficiently as it might have been.

Now take that scenario one stage further: if the restore had been of *.* , rather than *.dbf. At that point, you have just overwritten your existing online redo logs with old versions that were backed up last night. Now the loss of some of those up-to-date logs are not a problem, because you’ve got archives of them. But the loss of the previously current log, and its replacement by an out-of-date backup is terminal: there’s no way to replace the redo that log contained, and hence you have just lost committed data.

Now you might laugh at that, and say you would never be so stupid. But it's happened. I've seen it happen. Oh, and confession-time: I've done it myself. In short, the mere presence of the online logs in a backup set is a risk. One over-eager restore operation later, and committed data will have been lost that shouldn't have been.

I don't mind doing unnecessary things when 'there's no harm' in doing it. But in the case of backing up your online redo logs, it is not only unnecessary but it can definitely do harm.

I have a note here in Oracle's own documentation: 'In cases where the entire database needs to be restored, the process is simplified if the online redo logs have been backed up'. So Oracle itself says it's OK to do it.

That is because Oracle documentation describes what is technically possible, not what is pragmatically the safest or best thing to do. If you have closed your database down cleanly, then you can make a copy of the 'online' logs which will not be internally inconsistent. I said as much earlier. But there is no **need** to do so.

It is again true that the recovery process in the event of complete loss of the database with backups of redo logs available would be something along the lines of:

```
copy c:\backup\*. * c:\oracle\ora92
```

```
startup
```

But the equivalent without the redo logs present would be:

```
copy c:\backup\*.dbf c:\oracle\ora92
```

```
copy c:\backup\*.ctl c:\oracle\ora92
```

```
startup mount;
```

```
recover database until cancel using backup controlfile;
```

```
cancel;
```

```
alter database open resetlogs;
```

Which is definitely a bit more complex and a fair bit more typing. Is it so outrageously more complex, however, that it is worth risking the loss of committed data to avoid having to do it this way? I certainly don't think so. If care and concern for your organisation's data doesn't move you to a similar viewpoint, just consider: a DBA that manages to lose committed data usually loses his or her job shortly thereafter. Self-interest and self-preservation alone should cause you to avoid taking unnecessary risks when the risk-minimised approach is not exactly rocket science.

So yes, the Oracle documentation is correct: cleanly shutdown databases technically can have their redo logs backed up. But I'm right too: common sense, safety, caution and pragmatism dictates that you never, ever back them up.

And if you follow my advice, it will always be correct, whether your database is hot, cold, in archivelog mode or noarchivelog mode. If you follow Oracle's advice, however, you have to remember that it only applies to noarchivelog databases which have been shutdown cleanly. And, personally, I like laws of nature which are always true, not working shortcuts whose precise applicability depends on a variety of factors.

Finally, take a look at RMAN, the backup utility written and supplied by Oracle. It is syntactically impossible to get that tool to backup online redo logs. That ought to tell you something: Oracle's own best practice is not to back them up. So follow their lead: leave your online logs out of your backups!

[Troubleshooting Techniques](#)

Outline the steps for recovery of missing data file?

Losing Datafiles Whenever you are in NoArchivelog Mode:

```
#####
```

If you are in noarchivelog mode and you loss any datafile then whether it is temporary or permanent media failure, the database will automatically shut down. If failure is temporary then correct the underline hardware and start the database. Usually crash recovery will perform recovery of the committed transaction of the database from online redo log files. If you have permanent media failure then restore a whole database from a good backup. How to restore a database is as follows:

If a media failure damages datafiles in a NOARCHIVELOG database, then the only option for recovery is usually to restore a consistent whole database backup. As you are in noarchivelog mode so you have to understand that changes after taken backup is lost.

If you logical backup that is export file you can import that also.

In order to recover database in noarchivelog mode you have to follow the following procedure.

1)If the database is open shutdown it.

```
SQL>SHUTDOWN IMMEDIATE;
```

2)If possible, correct the media problem so that the backup database files can be restored to their original locations.

3)Copy all of the backup control files, datafiles to the default location if you corrected media failure. However you can restore to another location. Remember that all of the files not only the damaged files.

4)Because online redo logs are not backed up, you cannot restore them with the datafiles and control files. In order

to allow the database to reset the online redo logs, you must have to do incomplete recovery:

```
RECOVER DATABASE UNTIL CANCEL
```

```
CANCEL
```

5)Open the database in RESETLOGS mode:

```
ALTER DATABASE OPEN RESETLOGS;
```

In order to rename your control files or in case of media damage you can copy it to another location and then by setting (if spfile)

```
STARTUP NOMOUNT
```

```
alter system set control_files='+ORQ/orq1/controlfile/control01.ctl','+ORQ/orq1/controlfile/control02.ctl'
```

```
scope=spfile;
```

```
STARTUP FORCE MOUNT;
```

In order to rename data files or online redo log files first copy it to new location and then point control file to new location by,

```
ALTER DATABASE RENAME FILE '+ORQ/orq1/datafile/system01.dbf';
```

```
TO '+ORQ/orq1/datafile/system02.dbf';
```

Losing Datafiles Whenever you are in Archivelog Mode:

```
#####
```

If the datafile that is lost is under SYSTEM tablespace or if it is a datafile contain active undo segments then database shuts down. If the failure is temporary then correct the underline hardware and start the database. Usually crash recovery will perform recovery of the committed transaction of the database from online redo log files.

If the datafile that is lost is not under SYSTEM tablespace and not contain active undo segments then the affected datafile is gone to offline. The database remains open. In order to fix the problem take the affected tablespace offline and then recover the tablespace.

112. Outline the steps for recovery with missing online redo logs?

Redo log is CURRENT (DB was shut down cleanly)

If the CURRENT redo log is lost and if the DB is closed consistently, OPEN RESETLOGS can be issued directly without any transaction loss. It is advisable to take a full backup of DB immediately after the STARTUP.

Redo log is CURRENT (DB was not shut down cleanly)

When a current redo log is lost, the transactions in the log file are also lost before making to archived logs. Since a DB startup can no more perform a crash recovery (since all the now-available online log files are not sufficient to startup the DB in consistent state), an incomplete media recovery is the only option. We will need to restore the DB from a previous backup and restore to the point just before the lost redo log file. The DB will need to be opened in RESETLOGS mode. There is some transaction loss in this scenario.

```
RMAN> RESTORE CONTROLFILE FROM '<backup tag location>';
```

```
RMAN> ALTER DATABASE MOUNT;
```

```
RMAN> RESTORE DATABASE;
```

```
RMAN> RECOVER DATABASE UNTIL TIME "to_date('MAR 05 2009 19:00:00','MON DD YYYY HH24:MI:SS')";
```

```
RMAN> ALTER DATABASE OPEN RESETLOGS;
```

113. Outline steps for recovery with missing archived redo logs?

If a redo log file is already archived, its loss can safely be ignored. Since all the changes in the DB are now archived and the online log file is only waiting for its turn to be re-written by LGWR (redo log files are written circularly) the loss of the redo log file doesn't matter much. It may be re-created using the command

```
SQL> STARTUP MOUNT;
```

```
SQL> ALTER DATABASE CLEAR LOGFILE GROUP <group#>;
```

This will re-create all groups and no transactions are lost. The database can be opened normally after this.

114. What is RMAN and How to configure it?

A. RMAN is an Oracle Database client that performs backup and recovery tasks on your databases and automates administration of your backup strategies. It greatly simplifies the dba jobs by managing the production database's backing up, restoring, and recovering database files.

This tool integrates with sessions running on an Oracle database to perform a range of backup and recovery activities, including maintaining an RMAN repository of historical data about backups. There is no additional installation required for this tool. Its by default get installed with the oracle database installation. The RMAN environment consists of the utilities and databases that play a role in backing up your data. You can access RMAN through the command line or through Oracle Enterprise Manager.

115. Why to use RMAN?

A. RMAN gives you access to several backup and recovery techniques and features not available with user-managed backup and recovery. The most noteworthy are the following:

-- **Automatic specification of files to include in a backup** : Establishes the name and locations of all files to be backed up.

-- **Maintain backup repository** : Backups are recorded in the control file, which is the main repository of RMAN metadata. Additionally, you can store this metadata in a recovery catalog,

-- **Incremental backups** : An incremental backup stores only blocks changed since a previous backup. Thus, they provide more compact backups and faster recovery, thereby reducing the need to apply redo during datafile media recovery.

-- **Unused block compression** : In unused block compression, RMAN can skip data blocks that have never been used

-- **Block media recovery** : You can repair a datafile with only a small number of corrupt data blocks without taking it offline or restoring it from backup.

-- **Binary compression** : A binary compression mechanism integrated into Oracle Database reduces the size of backups.

-- **Encrypted backups** : RMAN uses backup encryption capabilities integrated into Oracle Database to store backup sets in an encrypted format.

-- **Corrupt block detection** : RMAN checks for the block corruption before taking its backup.

116. How RMAN works?

A. RMAN backup and recovery operation for a target database are managed by *RMAN client*. RMAN uses the target database control file to gather metadata about the target database and to store information about its own operations. The RMAN client itself does not perform backup, restore, or recovery operations. When you connect the RMAN client to a target database, RMAN allocates *server sessions* on the target instance and directs them to perform the operations. The work of backup and recovery is performed by server sessions running on the target database. A *channel* establishes a connection from the *RMAN client* to a target or auxiliary database instance by starting a *server session* on the instance. The channel reads data into memory, processes it, and writes it to the output device. When you take a database backup using RMAN, you need to connect to the target database using RMAN Client. The RMAN client can use Oracle Net to connect to a target database, so it can be located on any host that is connected to the target host through Oracle Net. For backup you need to allocate explicit or implicit channel to the target database. An RMAN channel represents one stream of data to a device, and corresponds to one database server session. This session dynamically collect information of the files from the target database control file before taking the backup or while restoring.

For example If you give ' Backup database ' from RMAN, it will first get all the datafiles information from the controlfile. Then it will divide all the datafiles among the allocated channels. (roughly equal size of work as per the datafile size). Then it takes the backup in 2 steps. In the first step the channel will read all the Blocks of the entire datafile to find out all the formatted blocks to backup. Note : RMAN do not take backup of the un formatted blocks. In the second step it take backup of the formatted blocks. This is the best advantage of using RMAN as it only take backup of the required blocks. Lets say in a datafile of 100 MB size, there may be only 10 MB of use full data and rest 90 MB is free then RMAN will only take backup of those 10 MB.

117. What O/S and oracle user privilege required to use RMAN?

A. RMAN always connect to the target or auxiliary database using the SYSDBA privilege. In fact the SYSDBA keywords are implied and cannot be explicitly specified. Its connections to a database are specified and authenticated in the same way as SQL*Plus connections to a database.

The O/S user should be part of the DBA group . For remote connection it needs the password file

Authentication.Target database should have the initialization parameter REMOTE_LOGIN_PASSWORDFILE set to EXCLUSIVE or SHARED.

118. RMAN terminology:

A target database: An Oracle database to which RMAN is connected with the TARGET keyword. A target database is a database on which RMAN is performing backup and recovery operations. RMAN always maintains metadata about its operations on a database in the control file of the database.

A recovery Catalog: A separate database schema used to record RMAN activity against one or more target databases. A recovery catalog preserves RMAN repository metadata if the control file is lost, making it much easier to restore and recover following the loss of the control file. The database may overwrite older records in the control file, but RMAN maintains records forever in the catalog unless deleted by the user.

Backup sets: RMAN can store backup data in a logical structure called a backup set, which is the smallest unit of an

RMAN backup. One backup set contains one or more datafiles a section of datafile or archive logs.

Backup Piece: A backup set contains one or more binary files in an RMAN-specific format. This file is known as a backup piece. Each backup piece is a single output file. The size of a backup piece can be restricted; if the size is not restricted, the backup set will comprise one backup piece. Backup piece size should be restricted to no larger than the maximum file size that your filesystem will support.

Image copies: An image copy is a copy of a single file (datafile, archive log, or controlfile). It is very similar to an O/S copy of the file. It is not a backupset or a backup piece. No compression is performed.

Snapshot Controlfile: When RMAN needs to resynchronize from a read-consistent version of the control file, it creates a temporary snapshot control file. The default name for the snapshot control file is port-specific.

Database Incarnation: Whenever you perform incomplete recovery or perform recovery using a backup control file, you must reset the online redo logs when you open the database. The new version of the reset database is called a new incarnation. The reset database command directs RMAN to create a new database incarnation record in the recovery catalog. This new incarnation record indicates the current incarnation.

119. What is RMAN Configuration and how to configure it?

A. The RMAN backup and recovery environment is preconfigured for each target database. The configuration is persistent and applies to all subsequent operations on this target database, even if you exit and restart RMAN. RMAN configured settings can specify backup devices, configure a connection to a backup device, policies affecting backup strategy, encryption algorithm, snapshot controlfile location and others.

By default there are few default configuration are set when you login to RMAN. You can customize them as per your requirement. Any time you can check the current setting by using the "Show all" command. CONFIGURE command is used to create persistent settings in the RMAN environment, which apply to all subsequent operations, even if you exit and restart RMAN. For details of the Configuration kindly refer Note <>

120. How Many catalog database I can have?

A. You can have multiple catalog databases for the same target database. But at a time you can connect to only 1 catalog database via RMAN. It's not recommended to have multiple catalog databases.

121. What is the advantage of catalog database?

Catalog database is a secondary storage of backup metadata. Its very useful in case you lost the current controlfile, as all the backup information are there in the catalog schema. Secondly from controlfile the older backup information are aged out depending upon the control_file_record_keep_time. RMAN catalog database maintain the history of data. Kindly refer the note <> for more details on relation between retention policy and control_file_record_keep_time.

122. What is the difference between catalog database & catalog schema?

A. Catalog database is like any other database which contains the RMAN catalog user's schema

122. Catalog database compatibility matrix with the target database?

A. refer [Note 73431.1](#) : RMAN Compatibility Matrix

123. What happen if catalog database lost?

A. Since catalog database is a option one there is no direct effect of loss of catalog database. Create a new catalog database and register the target database with the newly created catalog one. All the backup information from the target database current controlfile will be updated to the catalog schema. If any backup information which is aged out from the target database then you need to manually catalog those backup pieces.

124. Can I regulate the size of backup piece and backupset?

A. Yes ! You can set max size of the backupset as well as the backup piece. By default one RMAN channel creates a single backupset with one backup piece in it. You can use the MAXPIECESIZE channel parameter to set limits on the size of backup pieces. You can also use the MAXSETSIZE parameter on the BACKUP and CONFIGURE commands to set a limit for the size of backup sets.

125. What is the difference between backup set and Image copy backup?

A : A backup set is an RMAN-specific proprietary format, whereas an image copy is a bit-for-bit copy of a file. By default, RMAN creates backup sets

126. What is RMAN consistent backup and inconsistent backup?

A. A consistent backup occurs when the database is in a consistent state. That means backup of the database taken after a shutdown immediate, shutdown normal or shutdown transactional. If the database is shutdown with abort option then it's not a consistent backup.

A backup when the database is Up and running is called an inconsistent backup. When a database is restored from an inconsistent backup, Oracle must perform media recovery before the database can be opened, applying any pending changes from the redo logs. You cannot take inconsistent backup when the database is in Non-Archive log mode.

127. Can I encrypt RMAN backup?

A. RMAN supports backup encryption for backup sets. You can use wallet-based transparent encryption, password-based encryption, or both. You can use the CONFIGURE ENCRYPTION command to configure persistent transparent encryption. Use the SET ENCRYPTION, command at the RMAN session level to specify password-based encryption.

128. Can RMAN take backup to Tape?

Yes! You can use RMAN for the tape backup. But RMAN cannot able to write directly to tape. You need to have third party Media Management Software installed. Oracle has published an API specification which Media Management Vendor's who are members of Oracle's Backup Solutions Partner program have access to. Media Management Vendors (MMVs) then write an interface library which the Oracle server uses to write and read to and from tape.

129. Where can I get the list of supported Third party Media Management Software for tape backup?

RMAN should not be used with that Media Manager until the MMV has certified that their software works with RMAN. Either contact your Media Manager, or check the RMAN home page for updates on which MMVs have certified their products on which platforms:

<http://www.oracle.com/technology/deploy/availability/htdocs/bsp.htm>

starting from oracle 10g R2 oracle has its Own Media management software for the database backup to tape called OSB.

130. How RMAN Interact with Media manager?

Before performing backup or restore to a media manager, you must allocate one or more channels or configure default channels for use with the media manager to handle the communication with the media manager. RMAN does not issue specific commands to load, label, or unload tapes. When backing up, RMAN gives the media manager a stream of bytes and associates a unique name with this stream. When RMAN needs to restore the backup, it asks the media manager to retrieve the byte stream. All details of how and where that stream is stored are handled entirely by the media manager.

131. What is Proxy copy backup to tape?

Proxy copy is functionality, supported by few media manager in which they handle the entire data movement between datafiles and the backup devices. Such products may use technologies such as high-speed connections between storage and media subsystems to reduce load on the primary database server. RMAN provides a list of files requiring backup or restore to the media manager, which in turn makes all decisions regarding how and when to move the data.

132. What is Oracle Secure backup?

Oracle Secure Backup is a media manager provided by oracle that provides reliable and secure data protection through file system backup to tape. All major tape drives and tape libraries in SAN, Gigabit Ethernet, and SCSI environments are supported.

133. Why Recovery catalog?

It is always recommended to have the recovery catalog. If the target database controlfiles are lost recovery can become difficult if not impossible. Having recovery catalog makes the DBA life easier at the time of critical scenario. Even for larger system the use of a recovery catalog can increase the backup performance.

Recovery Catalog Schema can be created in the Target database or in any test or development database. It is not at all recommended to create the recovery catalog in the target database itself. Make sure to have a separate database for the recovery catalog always. Also its recommended to create the recovery catalog database in a different machine. If creating the Recovery catalog database in different machine is not possible then ensure that the recovery catalog and target databases do not reside on the same disk. If both your recovery catalog and your target database suffer hard disk failure, your recovery process is much more difficult. If possible, take other measures as well to eliminate common points of failure between your recovery catalog database and the databases you are backing up.

The recovery catalog contains information about RMAN operations, including:

- + Datafile and archived redo log backup sets and backup pieces
- + Datafile copies
- + Archived redo logs and their copies
- + Tablespace and datafiles on the target database
- + Stored scripts, which are named user-created sequences of RMAN commands
- + Persistent RMAN configuration settings

133. How to create recovery catalog?

Creating recovery catalog is a 3 step process. The recovery catalog is stored in the default tablespace of the recovery catalog schema. SYS cannot be the owner of the recovery catalog.

1. Creating the Recovery Catalog Owner
2. Creating the Recovery Catalog
3. Registering the target database

1. Creating the Recovery Catalog Owner

1.1 Size of recovery catalog schema :

Size of recovery catalog schema depends on

- a) The number of databases monitored by the catalog.
- b) The rate at which archived redo log generates in the target database
- c) The number of backups for each target database
- d) RMAN stored scripts stored in the catalog

1.2 Creating the Recovery Catalog Owner

Start by creating a database schema (usually called rman). Assign an appropriate tablespace to it and grant it the recovery_catalog_owner role. Look at this example:

```
% sqlplus '/ as sysdba'
```

```
SQL> CREATE USER rman IDENTIFIED BY rman
```

```
      DEFAULT TABLESPACE tools
```

```
      TEMPORARY TABLESPACE temp
```

```
      QUOTA UNLIMITED ON tools;
```

```
SQL> GRANT CONNECT, RECOVERY_CATALOG_OWNER TO rman;
```

2. Creating the Recovery Catalog

log in to rman and create the catalog schema. Look at this example:

In the below example "catdb " is the catalog database connection string. Before creating the recovery catalog make sure to have the tnsnames.ora entry for the catalog database in the target server and the listener must be up and running in the catalog database server. You must be able to connect to the catalog database from sqlplus from the target server.

```
% rman catalog rman/rman @ catadb
```

```
RMAN> CREATE CATALOG;
```

3. Registering the target database

After making sure the recovery catalog database is open, connect RMAN to the target database and recovery catalog database and register the database . Make sure that your target database is either open or in Mount stage. Look at this example:

```
% rman TARGET / CATALOG rman/rman @ catdb
```

```
RMAN> REGISTER DATABASE;
```

RMAN creates rows in the catalog tables to contain information about the catalog database .Copy all the pertinent data from the controlfile into the catalog, synchronizing the catalog with the control file. You can register multiple target databases in a single recovery catalog, if they do not have duplicate DBIDs. RMAN uses the DBID to distinguish one database from another.

How to upgrade recovery catalog Schema?

When you upgrade target database to the latest version you need to upgrade the RMAN catalog schema. Connect to RMAN from the target database so that you can use, the target database's RMAN executable. Look at the **example**:

```
% rman target / catalog rman/rman @ catdb
```

```
RMAN> UPGRADE CATALOG;
```

```
RMAN-06435: recovery catalog owner is rman
```

```
RMAN-06442: enter UPGRADE CATALOG command again to confirm catalog upgrade
```

```
RMAN> UPGRADE CATALOG;
```

Issuing 'upgrade catalog' will only upgrade the catalog schema to be compatible with the higher release of RMAN; it will not upgrade the catalog database in any way. You have to connect to recovery catalog database catdb and run "upgrade catalog" twice.

134. How to upgrade recovery catalog database?

Upgrading the recovery catalog database is same as the any other database upgrade steps. Upgrading the catalog database do not upgrade the catalog database schema.

135. How to remove catalog?

The "drop catalog;" command to remove an RMAN catalog. These commands need to be entered twice to confirm the operation. Look at the example :

```
RMAN> DROP CATALOG;
```

136. How to unregister the target database from the recovery catalog?

From 10G onwards, the process is simplified by introducing a new RMAN command to unregister the target database from the recovery catalog. Look at the example :

```
RMAN> UNREGISTER DATABASE <database_name> ;
```

The command "unregister database " should be executed only at the RMAN prompt. This is a restriction to use this command. Also RMAN must be connected to the recovery catalog in which the target database is registered. (Ref. [Note 252800.1](#))

Prior to release 10G, in order to unregister the target database you need to execute the following statement in the recovery catalog database connected as recovery catalog schema owner. Look at the example:

```
% sqlplus rman/rman @catdb
```

```
SQL > DBMS_RCVCAT.UNREGISTERDATABASE (db_key, db_id);
```

To unregister a database from the recovery catalog prior to Oracle 10g (Ref. [Note 1058332.6](#)).

137. How to backup of the Recovery Catalog?

Recovery catalog database is just like any other database. This database backup need to be taken every time after the target database backup. You can take Physical backup or Logical backup of the catalog database. You can use RMAN for the backup of the recovery catalog database.

Few guidelines for recovery catalog database

- + Run the recovery catalog database in ARCHIVELOG mode so that you can do point-in-time recovery if needed.
- + Set the retention policy to a REDUNDANCY value greater than 1.
- + Do not use another recovery catalog as the repository for the backups.
- + Configure the control file autobackup feature to ON.

138. How to restore and Recover recovery catalog from Backup?

Restoring and recovering the recovery catalog is much like restoring and recovering any other database

Compatibility of the Recovery Catalog

When you use RMAN with a recovery catalog in an environment where you have run past versions of the database, you can wind up with versions of the RMAN client, recovery catalog database, recovery catalog schema, and target database that all originated in different releases of the database.

Here is a note which gives detailed information about the compatibility matrix

Ref. [Note.73431.1](#) RMAN Compatibility Matrix

139. How to identify recovery catalog schema version?

The schema version of the recovery catalog is stored in the recovery catalog itself. The information is important in case you maintain multiple databases of different versions in your production system, and need to determine whether the catalog schema version is usable with a specific target database version.

To determine schema version of recovery catalog connect to catalog database from the recover catalog user and then query RCVER table. Look at the example:

```
% sqlplus rman/rman @catdb
```

```
SQL > SELECT * FROM rcver;
```

```
VERSION
```

```
-----
```

```
11.01.00
```

If the table displays multiple rows, then the highest version in the RCVER table is the current catalog schema version.

For example, assume that the rcver table displays the following rows:

```
VERSION
```

```
-----
```

```
08.01.07
```

```
09.02.00
```

```
10.02.00
```

Oracle Data Guard FAQ

1. What is Data Guard and Benefits of Data Guard?
2. What are the types of Standby databases and their benefits?
3. How to setup Data Guard?
4. What are different types of modes in Data Guard and which is default?
5. How many standby databases we can create (in 10g/11g)?
6. What are the parameters we've to set in primary/standby for Data Guard?
7. What is the use of FAL_SERVER & FAL_CLIENT is it mandatory to set these?
8. How to find out backlog of standby?
9. If you didn't have access to the standby database and you wanted to find out what error has occurred in a data guard configuration, what view would you check in the primary database to check the error message?
10. How can u recover standby which far behind from primary (or) without archive logs how can we make standby sync?
11. What is snapshot standby (or) how can we give a physical standby to user in READ WRITE mode and let him do updates and revert back to standby?
12. What are new features in 11g Data Guard?
13. What are the uses of standby redo log files? In what scenario standby redo logs are used?
14. What is DG_CONFIG?
15. What is RTA (real-time apply) mode MRP?
16. What is the difference between normal MRP (managed apply) and RTA MRP (real time apply)?
17. What are various parameters in LOG_ARCHIVE_DEST and its use?
18. What is the difference between SYNC/ASync, LGWR/ARCH, and AFFIRM/NOAFFIRM?
19. What is Data Guard broker (or) Can we add/delete/create/drop the datafile at standby database? What is the use of DGMGRL?
20. What is STATICCONNECTIDENTIFIER property used for?
21. What is failover/switchover (or) what is the difference between failover & switchover?
22. What are the background processes involved in Data Guard?
23. What happens if standby out of sync with primary? How will you resolve it?
24. How will you sync if archive is got deleted in primary?
25. Can we change protection mode online?
26. How will add a datafile in standby environment?
27. Can we add/delete/create/drop the datafile at standby database?
28. If Standby database does not receive the redo data from the primary database, how will you diagnose?
29. You can't mount the standby database what is the reason?
30. How do you do network tuning for redo transmission in data guard?
31. How to troubleshoot the slow disk performance on standby database?
32. Does log files size should be same as primary server? If sizes are not same what will happen?
33. What is RFS process on Standby Database?
34. How to identify which transport mode (Archiver or Log Writer) you are using to ship?
35. How to check if you are using Real-Time Apply?
36. How to identify standby redo logs?
37. How to see members of standby redo log file?
38. How to add Standby Redo Log File Group to a Specific Group Number?
39. What are the different services available in Oracle Data Guard?
40. What are the different Protection modes available in Oracle Data Guard?
41. How to check what protection mode of primary database in your Oracle Data Guard?
42. How to change protection mode in Oracle Data Guard setup?
43. What are the advantages of using Physical standby database in Oracle Data Guard?
44. What is physical standby database in Oracle Data Guard?
45. What is Logical standby database in Oracle Data Guard?
46. What are the advantages of Logical standby database in Oracle Data Guard?
47. What is the usage of DB_FILE_NAME_CONVERT parameter in Oracle Data Guard setup?

48. What is the usage of LOG_FILE_NAME_CONVERT parameter in Oracle Data Guard setup?
49. Explain the parameter which is used for standby database?

1. What is Data Guard and Benefits of Data Guard?

Oracle Data Guard ensures high availability, data protection, and disaster recovery for enterprise data. Data Guard provides a comprehensive set of services that create, maintain, manage, and monitor one or more standby databases to enable production Oracle databases to survive disasters and data corruptions. Data Guard maintains these standby databases as transactionally consistent copies of the production database. Then, if the production database becomes unavailable because of a planned or an unplanned outage, Data Guard can switch any standby database to the production role, minimizing the downtime associated with the outage. Data Guard can be used with traditional backup, restoration, and cluster techniques to provide a high level of data protection and data availability. With Data Guard, administrators can optionally improve production database performance by offloading resource-intensive backup and reporting operations to standby systems.

A **Data Guard configuration** consists of one production database and one or more standby databases. The databases in a Data Guard configuration are connected by Oracle Net and may be dispersed geographically. There are no restrictions on where the databases are located, provided they can communicate with each other. For example, you can have a standby database on the same system as the production database, along with two standby databases on other systems at remote locations.

You can manage primary and standby databases using the SQL command-line interfaces or the Data Guard broker interfaces, including a command-line interface (DGMGRL) and a graphical user interface that is integrated in Oracle Enterprise Manager.

A standby database is a transactionally consistent copy of the primary database. Using a backup copy of the primary database, you can create up to nine standby databases and incorporate them in a Data Guard configuration. Once created, Data Guard automatically maintains each standby database by transmitting redo data from the primary database and then applying the redo to the standby database.

Similar to a primary database, a standby database can be either a single-instance Oracle database or an Oracle Real Application Clusters database.

Following are the different benefits in using Oracle Data Guard feature in your environment.

- High Availability.
- Data Protection.
- Offloading Backup operation to standby database.
- Automatic Gap detection and Resolution in standby database.
- Automatic Role Transition using Data Guard Broker.

Benefits of Data Guard:

Disaster recovery, data protection, and high availability: Data Guard provides an efficient and comprehensive disaster recovery and high availability solution. Easy-to-manage switchover and failover capabilities allow role reversals between primary and standby databases, minimizing the downtime of the primary database for planned and unplanned outages.

Complete data protection: Data Guard can ensure no data loss, even in the face of unforeseen disasters. A standby database provides a safeguard against data corruption and user errors. Storage level physical corruptions on the primary database do not propagate to the standby database. Similarly, logical corruptions or user errors that cause the primary database to be permanently damaged can be resolved. Finally, the redo data is validated when it is applied to the standby database.

Efficient use of system resources: The standby database tables that are updated with redo data received from the primary database can be used for other tasks such as backups, reporting, summations, and queries, thereby reducing the primary database workload necessary to perform these tasks, saving valuable CPU and I/O cycles. With a logical standby database, users can perform normal data manipulation on tables in schemas that are not updated from the primary database. A logical standby database can remain open while the tables are updated from the primary database, and the tables are simultaneously available for read-only access. Finally, additional indexes and materialized views can be created on the maintained tables for better query performance and to suit specific business requirements.

Flexibility in data protection to balance availability against performance requirements: Oracle Data Guard offers maximum protection, maximum availability, and maximum performance modes to help enterprises balance data availability against system performance requirements.

Automatic gap detection and resolution: If connectivity is lost between the primary and one or more standby databases (for example, due to network problems), redo data being generated on the primary database cannot be sent to those standby databases. Once a connection is reestablished, the missing archived redo log files (referred to as a gap) are automatically detected by Data Guard, which then automatically transmits the missing archived redo log files to the standby databases. The standby databases are synchronized with the primary database, without manual intervention by the DBA.

Centralized and simple management: The Data Guard broker provides a graphical user interface and a command-line interface to automate management and operational tasks across multiple databases in a Data Guard configuration. The broker also monitors all of the systems within a single Data Guard configuration.

Integration with Oracle Database: Data Guard is a feature of Oracle Database Enterprise Edition and does not require separate installation.

Automatic role transitions: When fast-start failover is enabled, the Data Guard broker automatically fails over to a synchronized standby site in the event of a disaster at the primary site, requiring no intervention by the DBA. In addition, applications are automatically notified of the role transition.

2. What are the types of Standby databases and their benefits?

Physical standby database

Provides a physically identical copy of the primary database, with on disk database structures that are identical to the primary database on a block-for-block basis. The database schema, including indexes, are the same. A physical standby database is kept synchronized with the primary database, though Redo Apply, which recovers the redo data received from the primary database and applies the redo to the physical standby database.

A physical standby database can be used for business purposes other than disaster recovery on a limited basis. Data Guard maintains a physical standby database by performing Redo Apply. When it is not performing recovery, a physical standby database can be open in read-only mode, or it can be opened temporarily in read/write mode if Flashback Database is enabled.

Redo Apply: The physical standby database is maintained by applying redo data from the archived redo log files or directly from standby redo log files on the standby system using the Oracle recovery mechanism. The recovery operation applies changes in redo blocks to data block using the data-block address. The database cannot be opened while redo is being applied.

Open read-only: A physical standby database can be open in read-only mode so that you can execute queries on the database. While opened in read-only mode, the standby database can continue to receive redo data, but application of the redo data from the log files is deferred until the database resumes Redo Apply.

Although the physical standby database cannot perform both Redo Apply and be opened in read-only mode at the same time, you can switch between them. For example, you can perform Redo Apply, then open it in read-only mode for applications to run reports, and then change it back to perform Redo Apply to apply any outstanding archived redo log files. You can repeat this cycle, alternating between Redo Apply and read-only, as necessary.

The physical standby database is available to perform backups. Furthermore, the physical standby database will continue to receive redo data even if archived redo log files or standby redo log files are not being applied at that moment.

Open read/write: A physical standby database can also be opened for read/write access for purposes such as creating a clone database or for read/write reporting. While opened in read/write mode, the standby database does not receive redo data from the primary database and cannot provide disaster protection.

The physical standby database can be opened temporarily in read/write mode for development, reporting, or testing purposes, and then flashed back to a point in the past to be reverted back to a physical standby database. When the database is flashed back, Data Guard automatically synchronizes the standby database with the primary database, without the need to re-create the physical standby database from a backup copy of the primary database.

Benefits of a Physical Standby Database

A physical standby database provides the following benefits:

Disaster recovery and high availability: A physical standby database enables a robust and efficient disaster recovery and high availability solution. Easy-to-manage switchover and failover capabilities allow easy role reversals between primary and physical standby databases, minimizing the downtime of the primary database for planned and unplanned outages.

Data protection: Using a physical standby database, Data Guard can ensure no data loss, even in the face of unforeseen disasters. A physical standby database supports all datatypes, and all DDL and DML operations that the primary database can support. It also provides a safeguard against data corruptions and user errors. Storage level physical corruptions on the primary database do not propagate to the standby database. Similarly, logical corruptions

or user errors that cause the primary database to be permanently damaged can be resolved. Finally, the redo data is validated when it is applied to the standby database.

Reduction in primary database workload: Oracle Recovery Manager (RMAN) can use physical standby databases to off-load backups from the primary database saving valuable CPU and I/O cycles. The physical standby database can also be opened in read-only mode for reporting and queries.

Performance: The Redo Apply technology used by the physical standby database applies changes using low-level recovery mechanisms, which bypass all SQL level code layers; therefore, it is the most efficient mechanism for applying high volumes of redo data.

Logical standby database

Contains the same logical information as the production database, although the physical organization and structure of the data can be different. The logical standby database is kept synchronized with the primary database through SQL Apply, which transforms the data in the redo received from the primary database into SQL statements and then executing the SQL statements on the standby database.

A logical standby database can be used for other business purposes in addition to disaster recovery requirements. This allows users to access a logical standby database for queries and reporting purposes at any time. Also, using a logical standby database, you can upgrade Oracle Database software and patch sets with almost no downtime. Thus, a logical standby database can be used concurrently for data protection, reporting, and database upgrades.

A logical standby database is initially created as an identical copy of the primary database, but it later can be altered to have a different structure. The logical standby database is updated by executing SQL statements. This allows users to access the standby database for queries and reporting at any time. Thus, the logical standby database can be used concurrently for data protection and reporting operations.

Data Guard automatically applies information from the archived redo log file or standby redo log file to the logical standby database by transforming the data in the log files into SQL statements and then executing the SQL statements on the logical standby database. Because the logical standby database is updated using SQL statements, it must remain open. Although the logical standby database is opened in read/write mode, its target tables for the regenerated SQL are available only for read-only operations. While those tables are being updated, they can be used simultaneously for other tasks such as reporting, summations, and queries. Moreover, these tasks can be optimized by creating additional indexes and materialized views on the maintained tables.

A logical standby database has some restrictions on datatypes, types of tables, and types of DDL and DML operations.

Benefits of a Logical Standby Database

A logical standby database provides similar disaster recovery, high availability, and data protection benefits as a physical standby database. It also provides the following specialized benefits:

Efficient use of standby hardware resources: A logical standby database can be used for other business purposes in addition to disaster recovery requirements. It can host additional database schemas beyond the ones that are protected in a Data Guard configuration, and users can perform normal DDL or DML operations on those schemas any time. Because the logical standby tables that are protected by Data Guard can be stored in a different physical layout than on the primary database, additional indexes and materialized views can be created to improve query performance and suit specific business requirements

Reduction in primary database workload: A logical standby database can remain open at the same time its tables are updated from the primary database, and those tables are simultaneously available for read access. This makes a logical standby database an excellent choice to do queries, summations, and reporting activities, thereby off-loading the primary database from those tasks and saving valuable CPU and I/O cycles.

3. How to setup Data Guard?

Step for Physical Standby

- Enable forced logging
- Create a password file
- Configure a standby redo log
- Enable archiving
- Set up the primary database initialization parameters
- Configure the listener and tnsnames to support the database on both nodes

4. What are different types of modes in Data Guard and which is default?

Three Modes are there in Data Guard

Maximum performance: This is the default protection mode. It provides the highest level of data protection that is possible without affecting the performance of a primary database. This is accomplished by allowing transactions to

commit as soon as all redo data generated by those transactions has been written to the online log.

Maximum protection: This protection mode ensures that no data loss will occur if the primary database fails. To provide this level of protection, the redo data needed to recover a transaction must be written to both the online redo log and to at least one standby database before the transaction commits. To ensure that data loss cannot occur, the primary database will shut down, rather than continue processing transactions.

Maximum availability: This protection mode provides the highest level of data protection that is possible without compromising the availability of a primary database. Transactions do not commit until all redo data needed to recover those transactions has been written to the online redo log and to at least one standby database.

5. How many standby databases we can create (in 10g/11g)?

Till [Oracle 10g](#), 9 standby databases are supported.

From [Oracle 11g R2](#), we can create 30 standby databases.

6. What are the parameters we've to set in primary/standby for Data Guard?

DB_NAME

DB_UNIQUE_NAME

LOG_ARCHIVE_CONFIG

LOG_ARCHIVE_DEST_2

FAL_SERVER

FAL_CLIENT

DB_FILE_NAME_CONVERT

LOG_FILE_NAME_CONVERT

STANDBY_FILE_MANAGEMENT

7. What is the use of FAL_SERVER & FAL_CLIENT is it mandatory to set these?

FAL = fetch archive log, it points to the ORACLE TNS service.

FAL_SERVER: need to mention the service name of the remote database, If from Primary database, FAL_SERVER value should point to the standby. If from Standby database, then FAL_SERVER value should point to the primary database.

FAL_CLIENT: points to its own TNS service name, i.e. from primary you need to mention primary TNS service, if from Standby mention the local TNS service.

FAL_CLIENT and FAL_SERVER are initialization parameters used to configure log gap detection and resolution at the standby database side of a physical database configuration. This functionality is provided by log apply services and is used by the physical standby database to manage the detection and resolution of archived redo logs.

FAL_CLIENT and FAL_SERVER only need to be defined in the initialization parameter file for the standby database(s). It is possible; however, to define these two parameters in the initialization parameter for the primary database server to ease the amount of work that would need to be performed if the primary database were required to transition its role.

FAL_CLIENT specifies the TNS network services name for the standby database (which is sent to the FAL server process by log apply services) that should be used by the FAL server process to connect to the standby database. The syntax would be:

FAL_CLIENT=<net_service_name_of_standby_database>

FAL_SERVER specifies the TNS network service name that the standby database should use to connect to the FAL server process. The syntax would be:

FAL_SERVER=<net_service_name_of_primary_database>

8. How to find out backlog of standby?

```
SELECT ROUND ((SYSDATE - A.NEXT_TIME)*24*60) AS "BACKLOG", M.SEQUENCE#-1 "SEQ APPLIED", M.PROCESS, M.STATUS
```

```
FROM V$ARCHIVED_LOG A, (SELECT PROCESS, SEQUENCE#, STATUS FROM V$MANAGED_STANDBY WHERE PROCESS LIKE '%MRP %') M WHERE A.SEQUENCE# = (M.SEQUENCE#-1);
```

9. If you didn't have access to the standby database and you wanted to find out what error has occurred in a data guard configuration, what view would you check in the primary database to check the error message?

Check the V\$DATAGUARD_STATUS view.

```
SELECT MESSAGE FROM V$DATAGUARD_STATUS;
```

10. How can u recover standby which far behind from primary (or) without archive logs how can we make standby sync?

By using RMAN incremental backup

11. What is snapshot standby (or) how can we give a physical standby to user in READ WRITE mode and let him do updates and revert back to standby?

Till Oracle 10g, create guaranteed restore point, open in read write, let him do updates, flashback to restore point, start MRP.

From Oracle 11g, convert physical standby to snapshot standby, let him do updates, convert to physical standby, start MRP.

12. What are new features in 11g Data Guard?

13. What are the uses of standby redo log files? In what scenario standby redo logs are used?

A standby redo log is required for the maximum protection and maximum availability modes and the LGWR ASYNC transport mode is recommended for all databases. Data Guard can recover and apply more redo data from a standby redo log than from archived redo log files alone.

You should plan the standby redo log configuration and create all required log groups and group members when you create the standby database. For increased availability, consider multiplexing the standby redo log files, similar to the way that online redo log files are multiplexed.

If the real-time apply feature is enabled, log apply services can apply redo data as it is received, without waiting for the current standby redo log file to be archived. This results in faster switchover and failover times because the standby redo log files have been applied already to the standby database by the time the failover or switchover begins.

A standby redo log resides on the standby database site. The standby redo log file is similar to an online redo log, except that a standby redo log is used to store redo data that has been received from a primary database.

Oracle Data Guard used to have the onerous problem of losing the last redo log. If the primary instance crashed, the "current" redo log (as written by the LGWR process) would need to be flushed (with a log switch) before the most recent changes could be applied to the standby database. If you could not flush the current redo, data could be lost forever.

Note: The standby redo logs are populated with redo information as fast as the primary redo logs, rather than waiting for the redo log to be archived and shipped to the standby database. This means that the standby redo log has more current information than the log apply mechanism because it took a "shortcut" and was written to the standby, bypassing the traditional archiving and FTP to the standby database.

Standby Redo Logs (SRL) : is similar to Online Redo Log (**ORL**) and only difference between two is that Standby Redo Log is used to store redo data received from **another database** (primary database).

Standby Redo Logs are only used if you have the **LGWR** as transport mode to Remote Standby Database.

Scenarios Standby Redo Logs are required:

Standby Redo Log is required if

1) Your standby database is in maximum protection or maximum availability modes. (Physical Standby Database can run in one of three modes – Maximum Protection, Maximum Availability and Maximum Performance)

or

2) If you are using Real-Time Apply on Standby Database.

or

3) If you are using Cascaded Destinations

Things good to know about SRL

i) Standby Redo Logs should be same size as of Online Redo Logs. (The RFS process will attach Standby Redo Logs only if they are of same size as of Online Redo Log)

ii) Although the standby redo log is only used when the database is running in the standby role, Oracle recommends that you create a standby redo log on the primary database so that the primary database can switch over quickly to the standby role without the need for additional DBA intervention.

iii) Standby redo logs can be created even after the standby has been created. In this case create the SRL's on the primary before the creation of SRL on the standby database. (Standby Redo Log is not mandatory for Primary Database but its good practice and useful in role conversion from Primary to Standby Database)

iv) It is a best practice/recommendation to maintain Standby Redo Logs (SRLs) on both the standby AND primary database when using LGWR transport mode regardless of protection mode (Maximum Protection/Performance/Availability).

14. What is DG_CONFIG?

LOG_ARCHIVE_CONFIG enables or disables the sending of redo logs to remote destinations and the receipt of remote redo logs, and specifies the unique database names (DB_UNIQUE_NAME) for each database in the Data Guard configuration.

Specifies a list of up to 9 unique database names (defined with the DB_UNIQUE_NAME initialization parameter) for all of the databases in the Data Guard configuration

15. What is RTA (real-time apply) mode MRP?

If the real-time apply feature is enabled, log apply services can apply redo data as it is received, without waiting for the current standby redo log file to be archived. This results in faster switchover and failover times because the standby redo log files have been applied already to the standby database by the time the failover or switchover begins.

16. What is the difference between normal MRP (managed apply) and RTA MRP (real time apply)?

By default, log apply services wait for the *full* archived redo log file to arrive on the standby database before applying it to the standby database. Redo data transmitted from the primary database is received by the remote file server process (RFS) on the standby system where the RFS process writes the redo data to either archived redo log files or standby redo log files. However, if you use standby redo log files, you can enable **real-time apply**, which allows Data Guard to recover redo data from the current standby redo log file as it is being filled up by the RFS process

17. What are various parameters in LOG_ARCHIVE_DEST and its use?

18. What is the difference between SYNC/ASYNC, LGWR/ARCH, and AFFIRM/NOAFFIRM?

19. What is Data Guard broker (or) what is the use of DGMGRL?

The Oracle Data Guard broker is a distributed management framework that automates and centralizes the creation, maintenance, and monitoring of Data Guard configurations. The following list describes some of the operations the broker automates and simplifies:

- Creating Data Guard configurations that incorporate a primary database, a new or existing (physical, logical, or snapshot) standby database, redo transport services, and log apply services, where any of the databases could be Oracle Real Application Clusters (RAC) databases.
- Adding additional new or existing (physical, snapshot, logical, RAC or non-RAC) standby databases to an existing Data Guard configuration, for a total of one primary database, and from 1 to 9 standby databases in the same configuration
- Managing an entire Data Guard configuration, including all databases, redo transport services, and log apply services, through a client connection to any database in the configuration.
- Managing the protection mode for the broker configuration
- Invoking switchover or failover with a single command to initiate and control complex role changes across all databases in the configuration.
- Configuring failover to occur automatically upon loss of the primary database, increasing availability without manual intervention
- Monitoring the status of the entire configuration, capturing diagnostic information, reporting statistics such as the redo apply rate and the redo generation rate, and detecting problems quickly with centralized monitoring, testing, and performance tools.

You can perform all management operations locally or remotely through the broker's easy-to-use interfaces: the Data Guard management pages in Oracle Enterprise Manager, which is the broker's graphical user interface (GUI), and the Data Guard command-line interface called DGMGRL.

Benefits:

The broker's interfaces improve usability and centralize management and monitoring of the Data Guard configuration. Available as a feature of the Enterprise Edition and Personal Edition of the Oracle database, the broker is also integrated with the Oracle database and Oracle Enterprise Manager. This broker attributes result in the following benefits:

Disaster protection: By automating many of the manual tasks required to configure and monitor a Data Guard configuration, the broker enhances the high availability, data protection, and disaster protection capabilities that are inherent in Oracle Data Guard. Access is possible through a client to any system in the Data Guard configuration, eliminating any single point of failure. If the primary database fails, the broker automates the process for any one of the standby databases to replace the primary database and take over production processing. The database availability that Data Guard provides makes it easier to protect your data.

Higher availability and scalability with Oracle Real Application Clusters (RAC) Databases: While Oracle Data Guard broker enhances disaster protection by maintaining transactionally consistent copies of the primary database, Data Guard, configured with Oracle high availability solutions such as Oracle Real Application Clusters (RAC) databases, further enhances the availability and scalability of any given copy of that database. The *intrasite* high availability of an Oracle RAC database complements the *intersite* protection that is provided by Data Guard broker.

Consider that you have a cluster system hosting a primary Oracle RAC database comprised of multiple instances sharing access to that database. Further consider that an unplanned failure has occurred. From a Data Guard broker perspective, the primary database remains available as long as at least one instance of the clustered database continues to be available for transporting redo data to the standby databases. Oracle Clusterware manages the availability of instances of an Oracle RAC database. Cluster Ready Services (CRS), a subset of Oracle Clusterware, works to rapidly recover failed instances to keep the primary database available. If CRS is unable to recover a failed instance, the broker continues to run automatically with one less instance. If the last instance of the primary database fails, the broker provides a way to fail over to a specified standby database. If the last instance of the primary database fails, and fast-start failover is enabled, the broker can continue to provide high availability by automatically failing over to a pre-determined standby database.

The broker is integrated with CRS so that database role changes occur smoothly and seamlessly. This is especially apparent in the case of a planned role switchover (for example, when a physical standby database is directed to take over the primary role while the former primary database assumes the role of standby). The broker and CRS work together to temporarily suspend service availability on the primary database, accomplish the actual role change for both databases during which CRS works with the broker to properly restart the instances as necessary, and then start services defined on the new primary database. The broker manages the underlying Data Guard configuration and its database roles while CRS manages service availability that depends upon those roles. Applications that rely on CRS for managing service availability will see only a temporary suspension of service as the role change occurs in the Data Guard configuration.

Note that while CRS helps to maintain the availability of the individual instances of an Oracle RAC database, the broker coordinates actions that maintain one or more physical or logical copies of the database across multiple geographically dispersed locations to provide disaster protection. Together, the broker and Oracle Clusterware provide a strong foundation for Oracle's high-availability architecture.

Automated creation of a Data Guard configuration: The broker helps you to logically define and create a Data Guard configuration consisting of a primary database and (physical or logical, snapshot, RAC or non-RAC) standby databases. The broker automatically communicates between the databases in a Data Guard configuration using Oracle Net Services. The databases can be local or remote, connected by a LAN or geographically dispersed over a WAN. Oracle Enterprise Manager provides a wizard that automates the complex tasks involved in creating a broker configuration, including:

Adding an existing standby database or a new standby database created from existing backups taken through Enterprise Manager

Configuring the standby control file, server parameter file, and datafiles

Initializing communication with the standby databases

Creating standby redo log files

Enabling Flashback Database if you plan to use fast-start failover

Although DGMGRL cannot automatically create a new standby database, you can use DGMGRL commands to configure and monitor an existing standby database, including those created using Enterprise Manager.

Easy configuration of additional standby databases: After you create a Data Guard configuration consisting of a primary and a standby database, you can add up to eight new or existing, physical, snapshot, or logical standby databases to each Data Guard configuration. Oracle Enterprise Manager provides an Add Standby Database wizard to guide you through the process of adding more databases. It also makes all Oracle Net Services configuration changes necessary to support redo transport services and log apply services across the configuration.

Simplified, centralized, and extended management: You can issue commands to manage many aspects of the broker configuration. These include:

- Simplify the management of all components of the configuration, including the primary and standby databases, redo transport services, and log apply services.
- Coordinate database state transitions and update database properties dynamically with the broker recording the changes in a broker configuration file that includes profiles of all the databases in the configuration. The broker propagates the changes to all databases in the configuration and their server parameter files.
- Simplify the control of the configuration protection modes (to maximize protection, to maximize availability, or to maximize performance).
- Invoke the Enterprise Manager verify operation to ensure that redo transport services and log apply services are configured and functioning properly.

Simplified switchover and failover operations: The broker simplifies switchovers and failovers by allowing you to invoke them using a single key click in Oracle Enterprise Manager or a single command at the DGMGRL command-line interface (referred to in this documentation as *manual failover*). For *lights-out administration*, you can enable fast-start failover to allow the broker to determine if a failover is necessary and to initiate the failover to a pre-specified target standby database automatically, with no need for DBA intervention. Fast-start failover can be configured to occur with no data loss or with a configurable amount of data loss.

Fast-start failover allows you to increase availability with less need for manual intervention, thereby reducing management costs. Manual failover gives you control over exactly when a failover occurs and to which target standby database. Regardless of the method you choose, the broker coordinates the role transition on all databases in the configuration. Once failover is complete, the broker posts the DB_DOWN event to notify applications that the new primary is available.

Note that you can use the DBMS_DG PL/SQL package to enable an application to initiate a fast-start failover when it encounters specific conditions

Only one command is required to initiate complex role changes for switchover or failover operations across all databases in the configuration. The broker automates switchover and failover to a specified standby database in the broker configuration. Enterprise Manager enables you to select a new primary database from a set of viable standby databases (enabled and running, with normal status). The DGMGRL SWITCHOVER and FAILOVER commands only require you to specify the target standby database before automatically initiating and completing the many steps in switchover or failover operations across the multiple databases in the configuration.

Built-in monitoring and alert and control mechanisms: The broker provides built-in validation that monitors the *health* of all of the databases in the configuration. From any system in the configuration connected to any database, you can capture diagnostic information and detect obvious and subtle problems quickly with centralized monitoring, testing, and performance tools. Both Enterprise Manager and DGMGRL retrieve a complete configuration view of the progress of redo transport services on the primary database and the progress of Redo Apply or SQL Apply on the standby database.

The ability to monitor local and remote databases and respond to events is significantly enhanced by the broker's health check mechanism and tight integration with the Oracle Enterprise Manager event management system.

Transparent to application: Use of the broker is possible for any database because the broker works transparently with applications; no application code changes are required to accommodate a configuration that you manage with the broker.

20. What is STATICCONNECTIDENTIFIER property used for?

To set the Data Guard property *StaticConnectIdentifier* to use a SID instead of service name to configure fast start failover method

21. What is failover/switchover (or) what is the difference between failover & switchover?

Switchover – This is done when both primary and standby databases are available. It is pre-planned.

Failover – This is done when the primary database is NO longer available (ie in a Disaster). It is not pre-planned.

A switchover (or graceful switchover) is a planned role reversal between the primary and the standby databases. This is used when there is a planned outage on the primary database or primary server and you do not want to have extended downtime on the primary database. The switchover allows you to switch the roles of the databases so that the standby databases now becomes a primary databases and all your users and applications can continue operations on the “new” primary database (on the standby server). During the switchover operation there is a small outage. How long the outage lasts, depends on a number of factors including the network, the number and sizes of the redo logs. The switchover operation happens on both the primary and standby database.

A failover operation is what happens when the primary database is no longer available. The failover operation only happens on the standby database. The failover operation activates the standby database and turns this into a primary database. This process cannot be reversed so the decision to failover should be carefully made. The failover process is initiated during a real disaster or severe outage.

Automatic Failover

Automatic failover is where the software determines when the standby database should be activated to become the new primary database. There are numerous conditions that can occur (ie: network glitches/outages) in any system which theoretically could disrupt communications between the primary and standby sites. Because of the importance of this decision and the number of variances, we believe it is best not to automate this process but to leave it in the hands of a human.

Switchover

Switchover is the act of change the standby database into the primary but in a controlled manor, the planned event means that it is safe from data loss because the primary database must complete all redo generation on the production data before allowing the switchover to commence. The switchback does not exists as it is a switchover but in the reserve order, which would restore the database back on its original server. This planned event normally happens during a quiet period, the reason for the switchover might be DR testing, patch, hardware changes, implementing RAC, etc.

Once the switchover is complete the redo from the new primary will send it to the remaining standby servers, including the old primary, if using either grid control or the broker this should be all automatically do for you, but if you are using SQLPlus you have to performance some manual work.

You always start the switchover on the primary database, the actual switchover command is below whether you are using Grid Control, Broker or SQLPlus.

start the switchover (primary)	alter database commit to switchover to standby;
-----------------------------------	---

When the switchover command is executed the redo generation is stopped, all DML related cursors are invalidated and users are either prevented from executing transactions or terminated and the current redo log is archived for each thread. A special switchover marker called the *EOR* (end of redo) is then placed in the header of the next sequence for each thread, and the online redo files are archived a second time, sending the final sequences to the standby databases. At this point the physical standby database is closed and the final log switch is done without allowing the primary database to advance the sequence numbers for each thread.

After the *EOR* redo is sent to the standby databases, the original primary database is finalized as a standby and its control file backed up to the trace file and converted to the correct type of standby control file. In the case of a physical standby switchover the managed recovery process (MRP) is automatically started on the original primary to apply the final archive logs that contain the *EOR* so that all the redo ever generated is processed. The primary is then dismounted and must then be restarted as a standby database in at least the *mount* state.

The standby database must received this *EOR* redo otherwise the switchover cannot occur, once this redo has been received and applied to complete the switchover you run the following command, this will be automatic if you are using the Grid Control or the Broker

complete the switchover (new primary)	alter database commit to switchover to primary;
---------------------------------------	---

The physical standby switchover will wait for the MRP process to exit after processing the *EOR* redo and then convert the standby control file into a normal production control file. The final thing to do is to open the database for general production use

complete the switchover (new primary)	alter database open;
---------------------------------------	----------------------

A logical standby also has to wait for the *EOR* redo from the primary to be applied and SQL apply to shut down before the switchover command can complete, once the *EOR* has been processed, the GUARD can be turned off and production processing can begin.

Failover

A failover is a unplanned event when something has happened to hardware, networking, etc. This is when you invoke you DR procedures (hopefully documented), and you will have full confidence in getting the new primary up and running as quickly as possible. Unlike the switchover which begins on the primary, no primary is involved which means you will not be able to get the redo from the primary. Depending on what protection mode you have chosen there may be data loss (less you have a Maximum Protection mode enabled), you start be telling Data Guard to apply the remaining redo that it can. Once the redo has been applied you run the same command that you do with a physical standby to switchover the standby to a primary

complete the switchover (new primary)	alter database commit to switchover to primary;
---------------------------------------	---

Once difference is when the switchover has completed the protection mode will be *maximum performance* regardless what it was before, to get it back to your original [protection mode](#) you must get a standby database back up and running, then manually execute the steps to get it into the [protection mode](#) you want.

change the protection mode

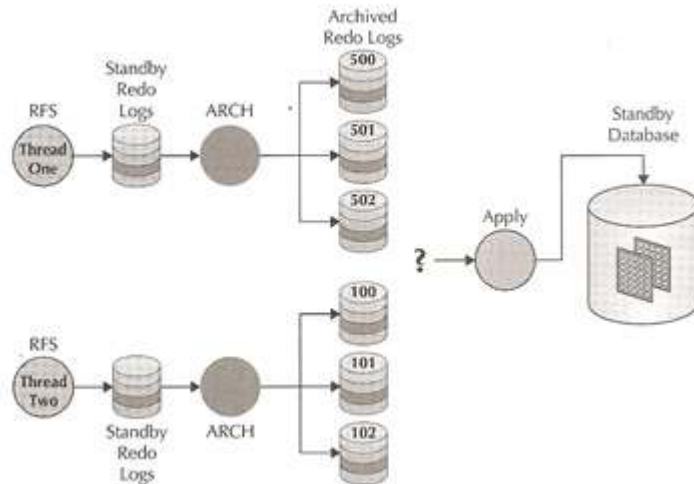
Choose what level of protection you require

sql> alter database set standby to maximize performance;

sql> alter database set standby to maximize availability;

sql> alter database set standby to maximize protection;

If you are using a protection mode that may result in data, the received archive redo logs are merged into a single thread and the sequence is sorted on the dependant transaction, this merged thread is then applied to the standby database up until the last redo. This may take sometime if using a RAC environment as the redo data has to be transfers from each instance.



Since the redo heartbeat is sent every 6 seconds or so, the general rule is that you may lose 6 seconds of redo during a failover but this is a best guess. At failover the merging thread will look at the last log of the disconnected thread and use the last heartbeat in it to define the consistent point, throwing away all the redo that the surviving nodes had been sending all along.

22. What are the background processes involved in Data Guard?

MRP, LSP,

23. What happens if standby out of sync with primary? How will you resolve it?

24. How will you sync if archive is got deleted in primary?

25. Can we change protection mode online?

26. How will add a datafile in standby environment?

27. Can we add/delete/create/drop the datafile at standby database?

You cannot rename the datafile on the standby site when the STANDBY_FILE_MANAGEMENT initialization parameter is set to AUTO. When you set the STANDBY_FILE_MANAGEMENT initialization parameter to AUTO, use of the following SQL statements is not allowed:

ALTER DATABASE RENAME

ALTER DATABASE ADD/DROP LOGFILE

ALTER DATABASE ADD/DROP STANDBY LOGFILE MEMBER

ALTER DATABASE CREATE DATAFILE AS

If you attempt to use any of these statements on the standby database, an error is returned. For example:

SQL> ALTER DATABASE RENAME FILE '/disk1/oracle/oradata/payroll/t_db2.log' to 'dummy';

alter database rename file '/disk1/oracle/oradata/payroll/t_db2.log' to 'dummy'

*

ERROR at line 1:

ORA-01511: error in renaming log/datafiles

ORA-01270: RENAME operation is not allowed if STANDBY_FILE_MANAGEMENT is auto

28. If Standby database does not receive the redo data from the primary database, how will you diagnose?

If the standby site is not receiving redo data, query the V\$ARCHIVE_DEST view and check for error messages. For example, enter the following query:

SQL> SELECT DEST_ID "ID",

2> STATUS "DB_status",

3> DESTINATION "Archive_dest",

4> ERROR "Error"


```

5> FROM V$ARCHIVE_DEST WHERE DEST_ID <=5;
ID DB_status Archive_dest          Error
-----

```

```

1 VALID   /vobs/oracle/work/arc_dest/arc
2 ERROR   standby1                ORA-16012: Archivelog standby database identifier mismatch
3 INACTIVE
4 INACTIVE
5 INACTIVE

```

5 rows selected.

If the output of the query does not help you, check the following list of possible issues. If any of the following conditions exist, redo transport services will fail to transmit redo data to the standby database:

The service name for the standby instance is not configured correctly in the tnsnames.ora file for the primary database.

- The Oracle Net service name specified by the LOG_ARCHIVE_DEST_ *n* parameter for the primary database is incorrect.
- The LOG_ARCHIVE_DEST_STATE_ *n* parameter for the standby database is not set to the value ENABLE.
- The listener.ora file has not been configured correctly for the standby database.
- The listener is not started at the standby site.
- The standby instance is not started.
- You have added a standby archiving destination to the primary SPFILE or text initialization parameter file, but have not yet enabled the change.
- The databases in the Data Guard configuration are not all using a password file, or the SYS password contained in the password file is not identical on all systems.
- You used an invalid backup as the basis for the standby database (for example, you used a backup from the wrong database, or did not create the standby control file using the correct method).

29. You can't mount the standby database what is the reason?

You cannot mount the standby database if the standby control file was not created with the ALTER DATABASE CREATE [LOGICAL] STANDBY CONTROLFILE ... statement or RMAN command. You cannot use the following types of control file backups:

- An operating system-created backup
- A backup created using an ALTER DATABASE statement *without* the PHYSICAL STANDBY or LOGICAL STANDBY option

30. How do you do network tuning for redo transmission in data guard?

For optimal performance, set the Oracle Net SDU parameter to 32 kilobytes in each Oracle Net connect descriptor used by redo transport services.

The following example shows a database initialization parameter file segment that defines a remote destination netserv:

```
LOG_ARCHIVE_DEST_3='SERVICE=netserv'
```

The following example shows the definition of that service name in the tnsnames.ora file:

```

netserv=(DESCRIPTION=(SDU=32768)(ADDRESS=(PROTOCOL=tcp)(HOST=host) (PORT=1521))
(CONNECT_DATA=(SERVICE_NAME=srvc)))

```

The following example shows the definition in the listener.ora file:

```

LISTENER=(DESCRIPTION=(ADDRESS_LIST=(ADDRESS=(PROTOCOL=tcp)
(HOST=host)(PORT=1521))))
SID_LIST_LISTENER=(SID_LIST=(SID_DESC=(SDU=32768)(SID_NAME=sid)
(GLOBALDBNAME=srvc)(ORACLE_HOME=/oracle)))

```

If you archive to a remote site using a high-latency or high-bandwidth network link, you can improve performance by using the SQLNET.SEND_BUF_SIZE and SQLNET.RECV_BUF_SIZE Oracle Net profile parameters to increase the size of the network send and receive I/O buffers.

31. How to troubleshoot the slow disk performance on standby database?

If asynchronous I/O on the file system itself is showing performance problems, try mounting the file system using the Direct I/O option or setting the FILESYSTEMIO_OPTIONS=SETALL initialization parameter. The maximum I/O size setting is 1 MB.

32. Does log files size should be same as primary server? If sizes are not same what will happen?

If you have configured a standby redo log on one or more standby databases in the configuration, ensure the size of the standby redo log files on each standby database exactly matches the size of the online redo log files on the primary database.

At log switch time, if there are no available standby redo log files that match the size of the new current online redo log file on the primary database:

- The primary database will shut down if it is operating in maximum protection mode,
or
- The RFS process on the standby database will create an archived redo log file on the standby database and write the following message in the alert log:
- No standby log files of size <#> blocks available.

For example, if the primary database uses two online redo log groups whose log files are 100K, then the standby database should have 3 standby redo log groups with log file sizes of 100K.

Also, whenever you add a redo log group to the primary database, you must add a corresponding standby redo log group to the standby database. This reduces the probability that the primary database will be adversely affected because a standby redo log file of the required size is not available at log switch time.

33. What is RFS process on Standby Database?

RFS – Remote File System on standby database receives data from Primary Database and writes it to Disk.

34. How to identify which transport mode (Archiver or Log Writer) you are using to ship?

```
SQL> show parameter log_archive_dest_  
log_archive_dest_  
log_archive_dest_  
SERVICE=visr12_standby [ARCH | LGWR]
```

If neither the ARCH or LGWR attribute is specified, the default is ARCH.

35. How to check if you are using Real-Time Apply?

```
SQL> SELECT DEST_ID, RECOVERY_MODE FROM V$ARCHIVE_DEST_STATUS;
```

36. How to identify standby redo logs?

```
SQL> select * from v$standby_log;
```

37. How to see members of standby redo log file?

```
SQL> select * from v$logfile where type='STANDBY';
```

38. How to add Standby Redo Log File Group to a Specific Group Number?

```
SQL> alter database add standby logfile group 4 (  
'<full_path_for_srl>/log04a.dbf',  
'<full_path_for_srl>/log04b.dbf'  
) size 50m;
```

39. What are the different services available in Oracle Data Guard?

Following are the different Services available in Oracle Data Guard of Oracle database.

- Redo Transport Services.
- Log Apply Services.
- Role Transitions

40. What are the different Protection modes available in Oracle Data Guard?

Following are the different protection modes available in Data Guard of Oracle database you can use any one based on your application requirement.

- Maximum Protection
- Maximum Availability
- Maximum Performance

41. How to check what protection mode of primary database in your Oracle Data Guard?

By using following query you can check protection mode of primary database in your Oracle Data Guard setup.

```
SELECT PROTECTION_MODE FROM V$DATABASE;
```

For Example:

```
SQL> select protection_mode from v$database;
```

```
PROTECTION_MODE
```

```
-----
```

```
MAXIMUM PERFORMANCE
```

42. How to change protection mode in Oracle Data Guard setup?

By using following query you can change the protection mode in your primary database after setting up required value in corresponding LOG_ARCHIVE_DEST_n parameter in primary database for corresponding standby database.
ALTER DATABASE SET STANDBY DATABASE TO MAXIMUM [PROTECTION | PERFORMANCE | AVAILABILITY];

Example: alter database set standby database to MAXIMUM PROTECTION;

43. What are the advantages of using Physical standby database in Oracle Data Guard?

Advantages of using Physical standby database in Oracle Data Guard are as follows.

- High Availability.
- Load balancing (Backup and Reporting).
- Data Protection.
- Disaster Recovery.

44. What is physical standby database in Oracle Data Guard?

Oracle Standby database are divided into physical standby database or logical standby database based on standby database creation and redo log apply method. Physical standby database are created as exact copy i.e block by block copy of primary database. In physical standby database transactions happen in primary database are synchronized in standby database by using Redo Apply method by continuously applying redo data on standby database received from primary database. Physical standby database can offload the backup activity and reporting activity from Primary database. Physical standby database can be opened for read-only transactions but redo apply won't happen during that time. But from 11g onwards using Active Data Guard option (extra purchase) you can simultaneously open the physical standby database for read-only access and apply redo logs received from primary database.

45. What is Logical standby database in Oracle Data Guard?

Oracle Standby database are divided into physical standby database or logical standby database based on standby database creation and redo log apply method. Logical standby database can be created similar to Physical standby database and later you can alter the structure of logical standby database. Logical standby database uses SQL Apply method to synchronize logical standby database with primary database. This SQL apply technology converts the received redo logs to SQL statements and continuously apply those SQL statements on logical standby database to make standby database consistent with primary database. Main advantage of Logical standby database compare to physical standby database is you can use Logical standby database for reporting purpose during SQL apply i.e Logical standby database must be open during SQL apply. Even though Logical standby database are opened for read/write mode, tables which are in synchronize with primary database are available for read-only operations like reporting, select queries and adding index on those tables and creating materialized views on those tables. Though Logical standby database has advantage on Physical standby database it has some restriction on data-types, types of DDL, types of DML and types of tables.

46. What are the advantages of Logical standby database in Oracle Data Guard?

- Better usage of resource
- Data Protection
- High Availability
- Disaster Recovery

47. What is the usage of DB_FILE_NAME_CONVERT parameter in Oracle Data Guard setup?

DB_FILE_NAME_CONVERT parameter is used in Oracle Data Guard setup that to in standby databases.

DB_FILE_NAME_CONVERT parameter are used to update the location of data files in standby database. These parameter are used when you are using different directory structure in standby database compare to primary database data files location.

48. What is the usage of LOG_FILE_NAME_CONVERT parameter in Oracle Data Guard setup?

LOG_FILE_NAME_CONVERT parameter is used in Oracle Data Guard setup that to in standby databases.

LOG_FILE_NAME_CONVERT parameter are used to update the location of redo log files in standby database. These parameter are used when you are using different directory structure in standby database compare to primary database redo log file location.

49. Explain the parameter which is used for standby database?

The LOG_ARCHIVE_CONFIG parameter enables or disables the sending of redo streams to the standby sites. The DB_UNIQUE_NAME of the primary database is dg1 and the DB_UNIQUE_NAME of the standby database is dg2. The primary database is configured to ship redo log stream to the standby database. In this example, the standby database service is dg2.

Next, `STANDBY_FILE_MANAGEMENT` is set to `AUTO` so that when Oracle files are added or dropped from the primary database, these changes are made to the standby databases automatically. The `STANDBY_FILE_MANAGEMENT` is only applicable to the physical standby databases.

Setting the `STANDBY_FILE_MANAGEMENT` parameter to `AUTO` is recommended when using Oracle Managed Files (OMF) on the primary database. Next, the primary database must be running in `ARCHIVELOG` mode.

Oracle Data Pump FAQ

1. What is Data Pump? Explain?
2. What are the new features of Data Pump?
3. What are the Init.ora parameters that affect the performance of Data Pump?
4. How Data pump accesses loading and unloading of Data?
5. What is use of CONSISTENT option in exp?
6. What is use of DIRECT=Y option in exp?
7. What is use of COMPRESS option in exp?
8. How to improve exp performance?
9. How to improve imp performance?
10. What is use of INDEXFILE option in imp?
11. What is use of IGNORE option in imp?
12. What are the differences between expdp and exp (Data Pump or normal exp/imp)?
13. Why expdp is faster than exp (or) why Data Pump is faster than conventional export/import?
14. How to improve expdp performance?
15. How to improve impdp performance?
16. In Data Pump, where the jobs info will be stored (or) if you restart a job in Data Pump, how it will know from where to resume?
17. What is the order of importing objects in impdp?
18. How to import only metadata?
19. How to import into different user/tablespace/datafile/table?
20. How to export/import without using external directory?
21. . Using Data Pump, how to export in higher version (11g) and import into lower version (10g), can we import to 9i?
22. Using normal exp/imp, how to export in higher version (11g) and import into lower version (10g/9i)?
23. How to do transport tablespaces (and across platforms) using exp/imp or expdp/impdp?
24. Explain about Remapping in oracle using data pump?

Answers

1. What is Data Pump? Explain?

Oracle 10g offers several new features, one of which is Data Pump technology for fast data movement between databases. Most Oracle shops still use their traditional export and import utility scripts rather than this new technology. Data Pump technology is entirely different from the export/import utility, although they have a similar look and feel. Data Pump runs inside the database as a job, which means jobs are somewhat independent of the process that started the import or export. Another advantage is that other DBAs can login to the database and check the status of the job. The advantages of Data Pump, along with Oracle's plan to deprecate the traditional import/export utilities down the road, make Data Pump a worthwhile topic for discussion.

Oracle claims Data Pump offers a transfer of data and metadata at twice the speed of export and twenty to thirty times the speed of the import utility that DBAs have been using for years. Data Pump manages this speed with multiple parallel streams of data to achieve maximum throughput. Please note that Data Pump does not work with utilities older than the 10g release 1 utility.

Data Pump consists of two components: the Data Pump export utility called "expdp," to Export the objects from a database, and the Data Pump Import utility called "impdp," to load the objects into database. Just like traditional export and import utilities, the DBA can control these jobs with several parameters.

For example:

```
$expdp username/password (other parameters here)
```

```
$impdp username/password (other parameters here)
```

We can get a quick summary of all parameters and commands by simply issuing

```
$expdp help=y
```

```
$impdp help=y
```

Similar to the export and import utility, Data Pump export and import utilities are extremely useful for migrating especially large databases from an operating system and importing them into a database running on a different platform and operating system in a short amount of time.

The Oracle supplied package, DBMS_DATAPUMP, can be used to implement the API, through which you can access the Data Pump export and import utilities programmatically. In other words, we can create a much powerful, custom Data Pump utility using Data Pump technology, if you have hundreds of databases to manage.

One of the interesting points is how Data Pump initiates the export session. In the traditional export utility, the user process writes the exported data to the disk that is requested from the server process, as a part of regular session. The Data Pump expdp - user process launches a server-side process or job that writes data to disks on the server node, and this process runs independently of the session established by expdp client. However, similar to the traditional export utility, Data Pump writes the data into dump files in an Oracle proprietary format that only the Data Pump import utility can understand.

As stated earlier, Data Pump is a server-based utility, rather than client-based; dump files, log files, and SQL files are accessed relative to server-based directory paths. Data Pump requires you to specify directory paths as directory objects. A directory object maps a name to a directory path on the file system.

a. The following SQL statements create a user, a directory object named dpump_dir1 and grants the permissions to the user.

```
SQLPLUS system/manager@TDB10G as sysdba
```

```
SQL> create user dpuser identified by dpuser;
```

```
SQL> grant connect, resource to dpuser;
```

```
SQL> CREATE DIRECTORY dpump_dir1 AS '/opt/app/oracle';
```

```
SQL> grant read, write on directory dpump_dir1 to dpuser
```

b. Let us see how the INCLUDE and EXCLUDE parameters can be used to limit the load and unload of particular objects. When the INCLUDE parameter is used, only the objects specified by it will be included in the export. When the EXCLUDE parameter is used, all objects except those specified by it will be included in the export: Assume we have EMP, EMP_DETAILS and DEPT tables owned by dpuser.

```
$ expdp dpuser/dpuser@TDB10G schemas=dpuser  
include= TABLE:\"IN (\\EMP\\', \\DEPT\\')\"  
directory=dpump_dir1 dumpfile=dpuser.dmp log=dpuser.log
```

```
$expdp dpuser/dpuser@TDB10G schemas=dpuser
```

```
exclude=TABLE:"= \EMP_DETAILS\"
```

```
directory=dpump_dir1 dumpfile=dpuser2.dmp logfile=dpuser.log
```

As stated earlier, Data pump performance can be significantly improved by using the PARALLEL parameter. This should be used in conjunction with the "%U" wildcard in the DUMPFILE parameter to allow multiple dumpfiles to be created or read:

```
$expdp dpuser/dpuser@TDB10G schemas=dpuser
```

```
directory=dpump_dir1 parallel=4 dumpfile=dpuser_%U.dmp logfile=dpuser.log
```

Data Pump API:

The Data Pump API, DBMS_DATAPUMP, provides a high-speed mechanism to move the data from one database to another. Infact, the Data Pump Export and Data Pump Import utilities are based on the Data Pump API. The structure used in the client interface of this API is a job handle. Job handle can be created using the OPEN or ATTACH function of the DBMS_DATAPUMP package. Other DBA sessions can attach to a job to monitor and control its progress so that remote DBA can monitor the job that was scheduled by an on-site DBA.

The following steps list the basic activities involved in using Data Pump API.

1. Execute DBMS_DATAPUMP.OPEN procedure to create job.
2. Define parameters for the job like adding file and filters etc.
3. Start the job.
4. Optionally monitor the job until it completes.
5. Optionally detach from job and attach at later time.
6. Optionally, stop the job
7. Restart the job that was stopped.

Example of the above steps:

Declare

```
P_handle number; --- -- Data Pump job handle
```

```
P_last_job_state varchar2(45); ---- -- To keep track of job state
```

```
P_job_state varchar2(45);
```

```
P_status ku$_Status ----- -- The status object returned by get_status
```

```
BEGIN
```

```
P_handle:=DBMS_DATAPUMP.OPEN ('EXPORT','SCHEMA', NULL,'EXAMPLE', and 'LATEST');
```

```
-- Specify a single dump file for the job (using the handle just returned)
```

```
-- and a directory object, which must already be defined and accessible
```

```
-- to the user running this procedure
```

```
DBMS_DATAPUMP.ADD_FILE (p_handle,'example.dmp','DMPDIR');
```

```
-- A metadata filter is used to specify the schema that will be exported.
```

```
DBMS_DATAPUMP.METADATA_FILTER (p_handle,'SCHEMA_EXPR','IN ('dpuser')');
```

```
-- Start the job. An exception will be generated if something is not set up
```

```
-- Properly.
```

```
DBMS_DATAPUMP.start_job (p_handle);
```

```
----The export job should now be running.
```

The status of the job can be checked by writing a separate procedure and capturing the errors and status until it is completed. Overall job status can also be obtained by querying "SELECT * from dba_datapump_jobs".

Conclusion:

Oracle Data Pump is a great tool for the fast movement of data between the databases and much of this performance improvement is derived from the use of parameter "parallelism." Even when the Transportable Tablespace feature of Oracle is used to move self-contained data between the databases, Data Pump is still required for handling the extraction and recreation of the metadata for that tablespace. Whenever possible, Data Pump performance is further maximized by using Direct-Path driver. Otherwise, Data Pump accesses the data using an External Table access driver. Data Pump provides flexibility, with the implementation of parameters such as INCLUDE, EXCLUDE, QUERY, and TRANSFORM that gives the DBA more control of data and objects being loaded and unloaded. With all of these features, Data Pump is a welcome addition to DBA tools in a world that constantly redefines the size of the "large database".

2. What are the new features of Data Pump?

New Features of Data Dump that improve the performance of Data movement:

Below are some of the features that differentiate the traditional export and import utility from Data Pump. These features not only enhance the speed of the data transfer but also are handy for the DBA to assess how the job would run before actually running Data Dump.

Parallel Threads: The parallel parameter specifies the maximum number of threads of active execution operating on behalf of the export job. This execution set consists of a combination of worker processes and parallel I/O server processes. Because each active worker processes or I/O server process works on one active file at a time, the DBA must specify a sufficient number of files. Therefore, the value the DBA specifies for this parameter should be less than or equal to the number of files in the dump file set. This important parameter helps the DBA to make a trade-off between resource consumption and the elapsed time.

Ability to restart the job: The ability to restart a job is an extremely useful feature if DBA is involved in moving large amounts of data. The Data Pump job can be restarted without any data loss or corruption after unexpected failure or if the DBA stopped the job with stop_job parameter.

Ability to detach from and reattach the job: This allows other DBAs to monitor jobs from multiple locations. We can attach the Data Pump export and import utilities to one job at a time but we can have multiple clients attached to the same job.

Support for export and import operations over the network: The NETWORK_LINK parameter initiates an export using a database link. It means that the system, to which expdp is connected, contacts the source database referenced by the source_database_link, retrieves data from it and writes the data to a dump file set back on the connected system.

Ability to change the name of source datafile to a different name: The DBA can change the name of the source datafile to a different name in all DDL statements where the source datafile is referenced.

Support for filtering the metadata: The DBA can filter metadata using the “EXCLUDE” and “INCLUDE” options. If the object is excluded, all of its dependent objects are also excluded. For example, EXCLUDE=CONSTRAINT will exclude all constraints except “NOT NULL” and constraints needed for table creation, which includes:

INCLUDE=TABLE:"IN('EMPLOYEES','DEPARTMENTS')"

- **Space Estimate:** The DBA can estimate how much space an export job will consume, without actually performing the export.

- **Query Parameter:** The DBA can filter data during the export by specifying a clause for a “SELECT” statement.

- **Content Parameter:** The DBA can specify what is exported or imported, for example, Meta data only or data only or both.

3. What are the Init.ora parameters that affect the performance of Data Pump?

Oracle recommends the following settings to improve performance.

Disk_Asynch_io= true

Db_block_checking=false

Db_block_checksum=false

Additionally, the number of processes and sessions allowed to the database must be set to high, to allow for maximum parallelism.

4. How Data pump accesses loading and unloading of Data?

Oracle has provided direct path to unload or export operations since Oracle 7.3. This method has been very useful for DBAs that want a quick export of the database and this process has been further enhanced in the Data Pump technology. Oracle uses the direct path method for loading (impdp) and unloading (expdp) when the structure of the tables allows it. If the table is part of a cluster, or it has a global index on a partitioned table, then Data Pump accesses the data in a different method called External Table. Both the direct path load and external table method support the same external data representation, so we can load the data that was unloaded with External Table method and vice versa.

5. What is use of CONSISTENT option in exp?

Cross-table consistency. Implements SET TRANSACTION READ ONLY. Default value N.

6. What is use of DIRECT=Y option in exp?

Setting direct=yes, to extract data by reading the data directly, bypasses the SGA, bypassing the SQL command-processing layer (evaluating buffer), so it should be faster. Default value N.

7. What is use of COMPRESS option in exp?

Imports into one extent. Specifies how export will manage the initial extent for the table data. This parameter is helpful during database re-organization. Export the objects (especially tables and indexes) with COMPRESS=Y. If table was spanning 20 Extents of 1M each (which is not desirable, taking into account performance), if you export the table with COMPRESS=Y, the DDL generated will have initial of 20M. Later on when importing the extents will be

coalesced. Sometime it is found desirable to export with COMPRESS=N, in situations where you do not have contiguous space on disk (tablespace), and do not want imports to fail.

8. How to improve exp performance?

- a. Set the BUFFER parameter to a high value. Default is 256KB.
- b. Stop unnecessary applications to free the resources.
- c. If you are running multiple sessions, make sure they write to different disks.
- d. Do not export to NFS (Network File Share). Exporting to disk is faster.
- e. Set the RECORDLENGTH parameter to a high value.
- f. Use DIRECT=yes (direct mode export).

9. How to improve imp performance?

- a. Place the file to be imported in separate disk from datafiles.
- b. Increase the DB_CACHE_SIZE.
- c. Set LOG_BUFFER to big size.
- d. Stop redolog archiving, if possible.
- e. Use COMMIT=n, if possible.
- f. Set the BUFFER parameter to a high value. Default is 256KB.
- g. It's advisable to drop indexes before importing to speed up the import process or set INDEXES=N and building indexes later on after the import. Indexes can easily be recreated after the data was successfully imported.
- h. Use STATISTICS=NONE
- i. Disable the INSERT triggers, as they fire during import.
- j. Set Parameter COMMIT_WRITE=NOWAIT (in [Oracle 10g](#)) or COMMIT_WAIT=NOWAIT (in [Oracle 11g](#)) during import.

10. What is use of INDEXFILE option in imp?

Will write DDLs of the objects in the dumpfile into the specified file.

11. What is use of IGNORE option in imp?

Will ignore the errors during import and will continue the import.

12. What are the differences between expdp and exp (Data Pump or normal exp/imp)?

Explanation-1:

[Data Pump](#) is server centric (files will be at server).

Data Pump has APIs, from procedures we can run Data Pump jobs.

In Data Pump, we can stop and restart the jobs.

Data Pump will do parallel execution.

Tapes & pipes are not supported in Data Pump.

Data Pump consumes more undo tablespace.

Data Pump import will create the user, if user doesn't exist.

Explanation-2:

- Impdp/Expdp has self-tuning unities. Tuning parameters that were used in original Export and Import, such as BUFFER and RECORDLENGTH, are neither required nor supported by Data Pump Export and Import.
- Data Pump represent metadata in the dump file set as XML documents rather than as DDL commands.
- Impdp/Expdp use parallel execution rather than a single stream of execution, for improved performance.
- In Data Pump expdp full=y and then impdp schemas=prod is same as of expdp schemas=prod and then impdp full=y where in original export/import does not always exhibit this behavior.
- Expdp/Impdp access files on the server rather than on the client.
- Expdp/Impdp operate on a group of files called a dump file set rather than on a single sequential dump file.
- Sequential media, such as tapes and pipes, are not supported in oracle data pump. But in original export/import we could directly compress the dump by using pipes.
- The Data Pump method for moving data between different database versions is different than the method used by original Export/Import.
- When you are importing data into an existing table using either APPEND or TRUNCATE, if any row violates an active constraint, the load is discontinued and no data is loaded. This is different from original Import, which logs any rows that are in violation and continues with the load.
- Expdp/Impdp consume more undo tablespace than original Export and Import.
- If a table has compression enabled, Data Pump Import attempts to compress the data being loaded. Whereas, the original Import utility loaded data in such a way that if a even table had compression enabled, the data was not compressed upon import.

- Data Pump supports character set conversion for both direct path and external tables. Most of the restrictions that exist for character set conversions in the original Import utility do not apply to Data Pump. The one case in which character set conversions are not supported under the Data Pump is when using transportable tablespaces.
- There is no option to merge extents when you re-create tables. In original Import, this was provided by the COMPRESS parameter. Instead, extents are reallocated according to storage parameters for the target table.
- Differences between Data Pump impdp and import utility
- The original import utility dates back to the earliest releases of Oracle, and it's quite slow and primitive compared to Data Pump. While the old import (imp) and Data Pump import (impdp) do the same thing, they are completely different utilities, with different syntax and characteristics.
Here are the major syntax differences between import and Data Pump impdp:
 - Data Pump does not use the BUFFERS parameter
 - Data Pump export represents the data in XML format
 - A Data Pump schema import will recreate the user and execute all of the associated security privileges (grants, user password history).
 - Data Pump's parallel processing feature is dynamic. You can connect to a Data Pump job that is currently running and dynamically alter the number of parallel processes.
 - Data Pump will recreate the user, whereas the old imp utility required the DBA to create the user ID before importing.

13. Why expdp is faster than exp (or) why Data Pump is faster than conventional export/import?

Data Pump is block mode, exp is byte mode.

Data Pump will do parallel execution.

Data Pump uses direct path API.

14. How to improve expdp performance?

Using parallel option which increases worker threads. This should be set based on the number of cpus.

15. How to improve impdp performance?

Using parallel option which increases worker threads. This should be set based on the number of cpus.

16. In Data Pump, where the jobs info will be stored (or) if you restart a job in Data Pump, how it will know from where to resume?

Whenever Data Pump export or import is running, Oracle will create a table with the JOB_NAME and will be deleted once the job is done. From this table, Oracle will find out how much job has completed and from where to continue etc.

Default export job name will be SYS_EXPORT_XXXX_01, where XXXX can be FULL or SCHEMA or TABLE.

Default import job name will be SYS_IMPORT_XXXX_01, where XXXX can be FULL or SCHEMA or TABLE.

17. What is the order of importing objects in impdp?

Tablespaces

Users

Roles

Database links

Sequences

Directories

Synonyms

Types

Tables/[Partitions](#)

Views

Comments

Packages/Procedures/Functions

[Materialized views](#)

18. How to import only metadata?

CONTENT= METADATA_ONLY

19. How to import into different user/tablespace/datafile/table?

REMAP_SCHEMA

REMAP_TABLESPACE

REMAP_DATAFILE

REMAP_TABLE

REMAP_DATA

20. How to export/import without using external directory?

21. Using **Data Pump**, how to export in higher version (**11g**) and import into lower version (**10g**), can we import to **9i**?

22. Using normal exp/imp, how to export in higher version (11g) and import into lower version (10g/**9i**)?

23. How to do **transport tablespaces** (and across platforms) using exp/imp or expdp/impdp?

24. Explain about Remapping in oracle using data pump?

Schema, Table and Data Remapping in Oracle Datapump

As you know Datapump is the Oracle preferred tool for moving data and is soon will be the only option because traditional exp/imp utilities will be deprecated.

In following sections we will look at how you can remap schema, table and data remapping.

Data remapping allows you to manipulate sensitive data before actually placing the data inside the dump file. This can happen on done at different stages including during schema remap, table remap and remapping of individual rows inside tables i.e. data remapping. We will look at them one by one in this section.

Schema Remapping

When you export a schema or some objects of a schema from one database and import it to the other then import utility expects the same schema to be present in second database. For example if you export EMP table of SCOTT schema and import it to another then import utility will try to locate the SCOTT schema in second database and if not present, it may create it for you depending on the options you specified.

But if you want to create the EMP table in SH schema instead. The remap_schema option of impdp utility will allow you to accomplish that. For example

```
$ impdp userid=rman/rman@orcl dumpfile=data_pump:SCOTT.dmp remap_schema=SCOTT:sh
```

This functionality was available in before 11g versions of database but had a different syntax and naming which has been changed in favor of more broader concept that goes beyond just schema mapping.

Table Remapping

On a similar ground you can also import data from one table into a table with a different name by using the REMAP_TABLE option. If you want to import data of EMP table to EMPTEST table then you just have to provide the REMAP_TABLE option with the new table name.

This option can be used for both partitioned and nonpartitioned tables.

On the other side however table remapping has the following restrictions.

- If partitioned tables were exported in a transportable mode then each partition or subpartition will be moved to a separate table of its own.
- Tables will not be remapped if they already exist even if you specify the TABLE_EXISTS_ACTION to truncate or append.
- The export must be performed in non transportable mode.

The syntax of REMAP_TABLE is as follows:

```
REMAP_TABLE=[old_schema_name.old_table_name]:[new_schema_name.new_table_name]
```

Data Remapping

The remapping option is used to remap data rows which is an extremely powerful feature. You can modify the rows while exporting or importing them. It is worth mentioning that as opposed to the schema and table level remapping which is only a logical mapping to happen at import time, data remapping can be done while creating the dump file i.e. expdp or it can be done while importing a dump file i.e. impdp.

To use this all you have to do is to create a function to perform the actual manipulation and wrap it inside a package and then pass the name when exporting data or importing data. The decision on when to use this depends on your requirement. If you want to store manipulated data inside the dump file then you can use it while exporting and if you don't then you may use it while importing.

We will now look at an example to test this functionality. Let's take the example of the EMP table inside the SCOTT schema. During the export process we will set the salary column value to a fix value of 5000. Firstly we create the package to actually perform this operation.

Oracle Performance Tuning FAQ

1. What are the major focuses of Performance tuning?
2. How does Oracle aid performance tuning?
3. Why is performance tuning a menacing area for DBA's?
4. What are the approaches towards performance tuning?
5. What is a systematic approach to performance tuning?
6. What are the Oracle's suggestions towards systematic tuning?
7. What are the effects of poor database design?
8. What is reactive performance tuning?
9. Which is useful – systematic or reactive tuning?
10. We have an application whose code can't be changed. Can we improve its performance?
11. What is the use of SQL over procedural languages?
12. What is query processing?
13. What is query optimization?
14. What are the techniques used for query optimization?
15. What are the phases of a SQL statement processing?
16. What is Parsing?
17. Mention the steps in the creation of a parse tree?
18. Where does the parse tree generation take place?
19. What is Optimization/what happens during optimization phase?
20. How does a CBO generate an optimal execution plan for the SQL statement?
21. What are the parts of an optimizer phase?
22. What is query rewrite phase?
23. What is execution plan generation phase/physical execution plan execution plan generation phase?
24. What are the factors considered by a physical query/execution plan?
25. Which generates the query plan/what is generated by optimizer?
26. How does the optimizer choose the query plan/what is cost-based query optimization?
27. What are the factors affecting the cost of a execution plan?
28. What happens after choosing the low-cost physical query plan?
29. What is a heuristic strategy?
30. What are unary and binary operations?
31. What is an optimal operation processing strategy?
32. What are the heuristic-processing strategies?
33. What is query execution?
34. What is the crucial step in SQL statement processing?
35. What is the job of an optimizer?
36. What is an index?
37. Why is an index efficient?
38. When do we need to index tables?
39. Why an index does traverse a table's row faster?
40. How do you set up tablespaces during an Oracle installation?
41. You see multiple fragments in the SYSTEM tablespace, what should you check first?
42. What are some indications that you need to increase the SHARED_POOL_SIZE parameter?
43. What is the general guideline for sizing DB_BLOCK_SIZE and DB_MULTI_BLOCK_READ for an application that does many full table scans?

44. What is the fastest query method for a table?
45. Explain the use of TKPROF? What initialization parameter should be turned on to get full TKPROF output?
46. When looking at v\$sysstat you see that sorts (disk) is high. Is this bad or good? If bad, how do you correct it?
47. When should you increase copy latches? What parameters control copy latches?
48. Where can you get a list of all initialization parameters for your instance? How about an indication if they are default settings or have been changed?
49. Describe hit ratio as it pertains to the database buffers. What is the difference between instantaneous and total hit ratio; which should be used for tuning?
50. Discuss row chaining, how does it happen? How can you reduce it? How do you correct it?
51. When looking at the estat events report you see that you are getting busy buffer waits. Is this bad? How can you find what is causing it?
52. If you see contention for library caches how you can fix it?
53. If you see statistics that deal with "undo" what are they really talking about?
54. If a tablespace has a default pct increase of zero what will this cause (in relationship to the smon process)?
55. If a tablespace shows excessive fragmentation what are some methods to defragment the tablespace? (7.1,7.2 and 7.3 only)
56. How can you tell if a tablespace has excessive fragmentation?
57. You see the following on a status report:
redo log space requests 23
redo log space wait time 0
Is this something to worry about? What if redo log space wait time is high? How can you fix this?
58. What can cause a high value for recursive calls? How can this be fixed?
59. If you see a pin hit ratio of less than 0.8 in the estat library cache report is this a problem? If so, how do you fix it?
60. If you see the value for reloads is high in the estat library cache report is this a matter for concern?
61. You look at the dba_rollback_segs view and see that there is a large number of shrinks and they are of relatively small size, is this a problem? How can it be fixed if it is a problem?
62. You look at the dba_rollback_segs view and see that you have a large number of wraps is this a problem?
63. In a system with an average of 40 concurrent users you get the following from a query on rollback extents:
ROLLBACK CUR EXTENTS are

R01 11
R02 8
R03 12
R04 9
SYSTEM 4
You have room for each to grow by 20 more extents each. Is there a problem? Should you take any action?
64. You see multiple extents in the temporary tablespace. Is this a problem?
65. What operation query optimizer performs?
66. What do you mean by "Throughput" or "Best response Time"?
67. How optimizer mode can be change?
68. What is optimizer mode SQL Hint and what it does?
69. What are optimizer statistics how it can be collected?
70. What do you mean by Histograms?
71. What is an access path for query optimizer?
72. What are the different types of access path that can be followed by optimizer?
73. What is the explain plan? And what type of information explain plan contains?
74. What is TKPROF?

75. What is SQL Trace?
76. What is the Explain plan statement disadvantage?
77. What is the Explain plan statement advantage?
78. What is the plan table? Describe its purpose?
79. How you can create plan table if plan table already not exists?
80. What are the important fields of plan table?
81. How you run explain plan statement?
82. What are the methods can be used to display plan table output (Execution Plan)?
83. Why and when should one tune?
84. What database aspects should be monitored?
85. Where should the tuning effort be directed?
86. What tuning indicators can one use?
87. What tools/utilities does Oracle provide to assist with performance tuning?
88. What is STATSPACK and how does one use it?
89. When is cost based optimization triggered?
90. How can one optimize %XYZ% queries?
91. Where can one find I/O statistics per table?
92. My query was fine last week and now it is slow. Why?
93. Why is Oracle not using the damn index?
94. When should one rebuild an index?
95. How does one tune Oracle Wait events?
96. What is the difference between DBFile Sequential and Scattered Reads?
97. What is the use of statistics?
98. How to generate explain plan?
99. How to check explain plan of already ran SQLs?
100. How to find out whether the query has ran with RBO or CBO?
101. What are top 5 wait events (in AWR report) and how you will resolve them?
102. What are the init parameters related to performance/optimizer?
103. What are the values of optimizer_mode init parameters and their meaning?
104. What is the use of AWR, ADDM, ASH?
105. How to generate AWR report and what are the things you will check in the report?
106. How to generate ADDM report and what are the things you will check in the report?
107. How to generate ASH report and what are the things you will check in the report?
108. How to generate STATSPACK report and what are the things you will check in the report?
109. How to generate TKPROF report and what are the things you will check in the report?
110. What is Performance Tuning?
111. Types of Tunings?
112. What mainly Database Tuning contains?
113. What is an optimizer?
114. Types of Optimizers?
115. Which init parameter is used to make use of Optimizer?
116. Which optimizer is the best one?
117. What are the pre requisite to make use of Optimizer?
118. How do you collect statistics of a table?
119. What is the diff between compute and estimate?
120. What will happen if you set the optimizer_mode=choose? Ans: If the statistics of an object is available then CBO used. if not RBO will be used
121. Data Dictionary follows which optimizer mode?

122. How do you delete statistics of an object?
123. How do you collect statistics of a user/schema?
124. How do you see the statistics of a table?
125. What are chained rows?
126. How do you collect statistics of a user in Oracle Apps?
127. How do you create a execution plan and how do you see?
128. How do you know what sql is currently being used by the session?
129. What is a execution plan?
130. How do you get the index of a table and on which column the index is?
131. Which init paramter you have to set to by pass parsing?
132. How do you know which session is running long jobs?
133. How do you flush the shared pool?
134. How do you get the info about FTS?
135. How do you increase the db cache?
136. Where do you get the info of library cache?
137. How do you get the information of specific session?
138. What you'll check whenever user complains that his session/database is slow?
139. Customer reports a application slowness issue, and you need to evaluate database performance. What do you look at for 9i and for 11g?
140. You have found a long running sql in your evaluation of system health of database, what do you look for to determine why sql is slow?
141. You have a windows service is crashing, how can you determine the sqls related to the windows service?
142. What is proactive tuning and reactive tuning?
143. Describe the level of tuning in oracle?
144. What is Database design level tuning?
145. Explain rule-based optimizer and cost-based optimizer?
146. What are object datatypes? Explain the use of object datatypes?
147. What is translate and decode in oracle?
148. What is oracle correlated sub-queries? Explain with an example?
149. Explain union and intersect with examples?
150. What is difference between open_form and call_form? What is new_form built-in in oracle form?
151. What is advantage of having disk shadowing/ Mirroring in oracle?

Answers

1. What are the major focuses of Performance tuning?

Performance tuning focuses primarily on writing efficient SQL, allocating appropriate computing resources, and analyzing wait events and contention in a system.

2. How does Oracle aid performance tuning?

Oracle provides several options to aid performance tuning, such as partitioning large tables, using materialized views, storing plan outlines, using tools like Automatic Optimizer statistics collection feature, packages like DBMS_STATS, SQL Tuning Advisor to tune SQL statements, etc.

3. Why is performance tuning a menacing area for DBA's?

Like many other features of Oracle like exp/imp, backup recovery this field can't be automated. This is one area that requires a lot of detective work on the part of application programmers and DBA's to see some process is running slower than expected, why we can't scale applications to a larger number of users without problems like performance degradation etc. This is an area where our technical knowledge must be used along with constant experimentation and observation.

4. What are the approaches towards performance tuning?

We can follow either a systematic approach or a reactive approach for performance tuning.

5. What is a systematic approach to performance tuning?

It is mandatory to design the database properly at initial stages to avoid potential problems. It is mandatory to know the nature of application that a database is going to support. With a clear idea on the application's nature database can be optimally created by allocating appropriate resources to avoid problems when the application is moved to production. Most production moves cause problem because of the scalability problems with the applications. So, oracle recommends tuning database at inception stage. This is systematic approach to performance tuning.

6. What are the Oracle's suggestions towards systematic tuning?

Oracle suggests a specific design approach with the following steps. This is a top down approach:

- 1) Design the application correctly
- 2) Tune the application SQL code
- 3) Tune memory
- 4) Tune I/O
- 5) Tune contention and other issues

7. What are the effects of poor database design?

A poor database design results in poor application performance. We have to tune the application code and some database resources such as memory, CPU, I/O owing to performance degradation. An application performs well in development and testing. Will there be any performance problem when it is moved to production?

Production moves may cause problems due to scalability. We can't simulate the original load in test and development. So problems may crop up at times as the application may be performing poor due to scalability problems.

8. What is reactive performance tuning?

Performance tuning is an iterative process. We as a DBA may have to tune applications which is designed and implemented in production. The performance tuning at hits stage is referred to as reactive performance tuning.

9. Which is useful – systematic or reactive tuning?

The performance tuning steps to improve the performance of a database depends on the stage at which we get the input and on the nature of the application. DBA's can assist the developers to write optimal code that is scalable based on systematic approach. Mostly the real life problems that are encountered after production moves have to be solved by reactive performance tuning.

10. We have an application whose code can't be changed. Can we improve its performance?

We can improve the application performance without changing base SQL code by optimizing the SQL performance. Oracle has come up with SQL Advisor tool that helps SQL performance. We can make use of SQL Advisor tools' SQL Profiles to improve performance, though we can't touch the underlying SQL.

11. What is the use of SQL over procedural languages?

SQL isn't a procedural language in which we have to specify the steps to be followed to achieve the statement goal. We don't have to specify how to accomplish a task(say data retrieval) using SQL, rather we can specify as to what needs to be done.

12. What is query processing?

When a user starts a data retrieval operation, the user's SQL statement goes through several sequential steps that together constitute query processing. Query processing is the transformation of the SQL statement into efficient execution plan to return the requested data from the database.

13. What is query optimization?

Query optimization is the process of choosing the most efficient execution plan. The goal is to achieve the result with least cost in terms of resource usage. Resources include I/O and CPU usage on the server where the database is running. This is a means to reduce the execution times of the query, which is the sum of the execution times of the all component operations of the query.

14. What are the techniques used for query optimization?

Cost-based optimization, heuristic strategy is used for query optimization.

15. What are the phases of a SQL statement processing?

An user's SQL statement goes through the parsing, optimizing, and execution stages. If the SQL statement is a query(SELECT), data has to be retrieved so there's an additional fetch stage before the SQL processing is complete.

16. What is Parsing?

Parsing primarily consists of checking the syntax and semantics of the SQL statements. The end product of the parse stage of query compilation is the creation of a parse tree, which represents the query structure. The parse tree is then sent to the logical query plan generation stage.

17. Mention the steps in the creation of a parse tree?

- 1) The SQL statement is decomposed into relational algebra query that's analyzed to see whether it's syntactically correct.
- 2) The query then undergoes semantic checking.
- 3) The data dictionary is consulted to ensure that the tables and the individual columns that are referenced in the query do exist, as well as all the object privileges.
- 4) The column types are checked to ensure that the data matches the column definition.
- 5) The statement is normalized so that it can be processed more efficiently
- 6) The query is rejected if it is incorrectly formulated
- 7) Once the parse tree passes all the syntactic and semantic checks, it is considered a valid parse tree, and it's sent to the logical query plan generation stage.

18. Where does the parse tree generation take place?

The parse tree generation takes place in the library cache portion of the SGA(system global Area)./

19. What is Optimization/what happens during optimization phase?

During the optimization phase, Oracle uses its optimizer (CBO (cost-based optimizer)) to choose the best access method for retrieving data for the tables and indexes referred to in the query.

20. How does a CBO generate an optimal execution plan for the SQL statement?

Using the statistics we provide and the hints specified in the SQL queries, the CBO produces an optimal execution plan for the SQL statement.

21. What are the parts of an optimizer phase?

An optimizer phase can be divided into two distinct parts: the query rewrite phase and the physical execution plan generation phase.

22. What is query rewrite phase?

In this phase, the parse tree is converted into an abstract logical query plan. This is an initial pass at an actual query plan, and it contains only a general algebraic reformulation of the initial query. The various nodes and branches of the parse tree are replaced by operators of relational algebra.

23. What is execution plan generation phase/physical execution plan execution plan generation phase?

During this phase, Oracle transforms the logical query plan into a physical query plan.

The optimizer may be faced with a choice of several algorithms to solve a query. It needs to choose the most efficient algorithm to answer a query, and it needs to determine the most efficient way to implement the operations. The optimizer determines the order in which it will perform the steps.

24. What are the factors considered by a physical query/execution plan?

Following factors are considered by a physical query or an execution plan:

- 1) The various operations (eg: joins) to be performed during the query
- 2) The order in which the operations are performed
- 3) The algorithm to be used for performing each operation

- 4) The best way to retrieve data from disk or memory
- 5) The best way to pass data from one operation to another during the query

25. Which generates the query plan/what is generated by optimizer?

The optimizer generates several valid physical query plans. All the physical query plans are potential execution plans.

26. How does the optimizer choose the query plan/what is cost-based query optimization?

The optimizer generates several physical query plans that are potential execution plans. The optimizer then chooses among them by estimating the cost of each possible physical plan based on the table and index statistics available to it, and selecting the plan with the lowest estimated cost. This evaluation of the possible physical query plans is called cost-based query optimization.

27. What are the factors affecting the cost of an execution plan?

The cost of executing a plan is directly proportional to the amount of resources such as I/O, memory and CPU necessary to execute the proposed plan.

28. What happens after choosing the low-cost physical query plan?

The optimizer passes the low-cost physical query plan to the Oracle's query execution engine.

29. What is a heuristic strategy?

The database uses a less systematic query optimization technique known as the heuristic strategy.

30. What are unary and binary operations?

A join operation is called a binary operation; an operation like selection is called a unary operation.

31. What is an optimal operation processing strategy?

In general an optimal strategy is to perform unary operations first so the more complex and time-consuming binary operations use smaller operands. Performing as many of the possible unary operations first reduces the row sources of the join operations.

32. What are the heuristic-processing strategies?

- 1) Perform selection operation early so that we can eliminate a majority of the candidate rows early in the operation. If we leave most rows in until the end, we're going to do needless comparisons with the rows we're going to get rid of later
- 2) Perform projection operations early so that we limit the number of columns we have to deal with
- 3) If we need to perform consecutive join operation, perform the operations that produce the smaller join first
- 4) Compute common expressions once and save the results

33. What is query execution?

During the final stage of a query processing, the optimized query (the physical query plan that has been selected) is executed. If it's a SELECT statement the rows are returned to the user. If it's an INSERT, UPDATE or DELETE statement, the rows are modified. The SQL execution engine takes the execution plan provided by the optimization phase and executes it.

34. What is the crucial step in SQL statement processing?

Of the three steps involved in the SQL statement processing, the optimization process is the crucial one because it determines the all important question of how fast our data will be retrieved.

35. What is the job of an optimizer?

The job of an optimizer is to find the optimal/best plan to execute our DML statements such as SELECT, INSERT, UPDATE and DELETE. Oracle uses CBO to help determine efficient methods to execute queries.

36. What is an index?

An index is a data structure that takes the value of one or more columns of a table (the key) and returns all rows/requested-columns in a row quickly.

37. Why is an index efficient?

The efficiency of an index comes from the fact that it lets us find necessary rows without having to scan all the rows of a table. They need a fewer disk I/O's than if we had to scan the table and hence are efficient.

38. When do we need to index tables?

We need to index tables only when the queries will be selecting a small portion of the table. If our query is retrieving rows that are greater than 10 or 15 percent of the total rows in the table, we may not need an index.

39. Why an index does traverse a table's row faster?

Indexes prevent a full table scan, so it is inherently a faster means to traverse a table's row

A tablespace has a table with 30 extents in it. Is this bad? Why or why not?

Multiple extents in and of themselves aren't bad. However if you also have chained rows this can hurt performance.

40. How do you set up tablespaces during an Oracle installation?

You should always attempt to use the Oracle Flexible Architecture standard or another partitioning scheme to ensure proper separation of SYSTEM, ROLLBACK, REDO LOG, DATA, TEMPORARY and INDEX segments.

41. You see multiple fragments in the SYSTEM tablespace, what should you check first?

Ensure that users don't have the SYSTEM tablespace as their TEMPORARY or DEFAULT tablespace assignment by checking the DBA_USERS view.

42. What are some indications that you need to increase the SHARED_POOL_SIZE parameter?

Poor data dictionary or library cache hit ratios, getting error ORA-04031. Another indication is steadily decreasing performance with all other tuning parameters the same.

43. What is the general guideline for sizing DB_BLOCK_SIZE and DB_MULTI_BLOCK_READ for an application that does many full table scans?

Oracle almost always reads in 64k chunks. The two should have a product equal to 64 or a multiple of 64.

44. What is the fastest query method for a table?

Fetch by rowid

45. Explain the use of TKPROF? What initialization parameter should be turned on to get full TKPROF output?

The tkprof tool is a tuning tool used to determine CPU and execution times for SQL statements. You use it by first setting timed_statistics to true in the initialization file and then turning on tracing for either the entire database via the sql_trace parameter or for the session using the ALTER SESSION command. Once the trace file is generated you run the tkprof tool against the trace file and then look at the output from the tkprof tool. This can also be used to generate explain plan output.

46. When looking at v\$sysstat you see that sorts (disk) is high. Is this bad or good? If bad, how do you correct it?

If you get excessive disk sorts this is bad. This indicates you need to tune the sort area parameters in the initialization files. The major sort parameter is the SORT_AREA_SIZE parameter.

47. When should you increase copy latches? What parameters control copy latches?

When you get excessive contention for the copy latches as shown by the "redo copy" latch hit ratio. You can increase copy latches via the initialization parameter LOG_SIMULTANEOUS_COPIES to twice the number of CPUs on your system.

48. Where can you get a list of all initialization parameters for your instance? How about an indication if they are default settings or have been changed?

You can look in the init.ora file for an indication of manually set parameters. For all parameters, their value and whether or not the current value is the default value, look in the v\$parameter view.

49. Describe hit ratio as it pertains to the database buffers. What is the difference between instantaneous and total hit ratio; which should be used for tuning?

Hit ratio is a measure of how many times the database was able to read a value from the buffers versus how many times it had to re-read a data value from the disks. A value greater than 80-90% is good, less could indicate problems. If you take the ratio of existing parameters this will be a cumulative value since the database started. If you do a comparison between pairs of readings based on some arbitrary time span, this is the instantaneous ratio for that time span. Generally speaking an instantaneous reading gives more valuable data since it will tell you what your instance is doing for the time it was generated over.

50. Discuss row chaining, how does it happen? How can you reduce it? How do you correct it?

Row chaining occurs when a VARCHAR2 value is updated and the length of the new value is longer than the old value and won't fit in the remaining block space. This results in the row chaining to another block. It can be reduced by setting the storage parameters on the table to appropriate values. It can be corrected by export and import of the effected table.

51. When looking at the estat events report you see that you are getting busy buffer waits. Is this bad? How can you find what is causing it?

Buffer busy waits may indicate contention in redo, rollback or data blocks. You need to check the v\$waitstat view to see what areas are causing the problem. The value of the "count" column tells where the problem is, the "class" column tells you with what. UNDO is rollback segments, DATA is data base buffers.

52. If you see contention for library caches how can you fix it?

Increase the size of the shared pool.

53. If you see statistics that deal with "undo" what are they really talking about?

Rollback segments and associated structures

54. If a tablespace has a default pct increase of zero what will this cause (in relationship to the smon process)?

The SMON process won't automatically coalesce its free space fragments.

55. If a tablespace shows excessive fragmentation what are some methods to defragment the tablespace? (7.1,7.2 and 7.3 only)

In Oracle 7.0 to 7.2 The use of the 'alter session set events 'immediate trace name coalesce level ts#';' command is the easiest way to defragment contiguous free space fragmentation. The ts# parameter corresponds to the ts# value found in the ts\$ SYS table. In version 7.3 the 'alter tablespace coalesce;' is best. If free space isn't contiguous then export, drop and import of the tablespace contents may be the only way to reclaim non-contiguous free space.

56. How can you tell if a tablespace has excessive fragmentation?

If a select against the dba_free_space table shows that the count of tablespaces extents is greater than the count of its data files, then it is fragmented.

57. You see the following on a status report:

redo log space requests 23

redo log space wait time 0

Is this something to worry about? What if redo log space wait time is high? How can you fix this?

Since wait time is zero, no. If wait time was high it might indicate a need for more or larger redo logs.

58. What can cause a high value for recursive calls? How can this be fixed?

A high value for recursive calls is caused by improper cursor usage, excessive dynamic space management actions, and or excessive statement re-parses. You need to determine the cause and correct it. By either delinking applications to hold cursors, use proper space management techniques (proper storage and sizing) or ensure repeat queries are placed in packages for proper reuse.

59. If you see a pin hit ratio of less than 0.8 in the estat library cache report is this a problem? If so, how do you fix it?

This indicates that the shared pool may be too small. Increase the shared pool size.

60. If you see the value for reloads is high in the estat library cache report is this a matter for concern?

Yes, you should strive for zero reloads if possible. If you see excessive reloads then increase the size of the shared pool.

61. You look at the dba_rollback_segs view and see that there is a large number of shrinks and they are of relatively small size, is this a problem? How can it be fixed if it is a problem?

A large number of small shrinks indicates a need to increase the size of the rollback segment extents. Ideally you should have no shrinks or a small number of large shrinks. To fix this just increase the size of the extents and adjust optimal accordingly.

62. You look at the dba_rollback_segs view and see that you have a large number of wraps is this a problem?

A large number of wraps indicates that your extent size for your rollback segments are probably too small. Increase the size of your extents to reduce the number of wraps. You can look at the average transaction size in the same view to get the information on transaction size.

63. In a system with an average of 40 concurrent users you get the following from a query on rollback extents:

ROLLBACK CUR EXTENTS are

R01 11

R02 8

R03 12

R04 9

SYSTEM 4

You have room for each to grow by 20 more extents each. Is there a problem? Should you take any action?

No there is not a problem. You have 40 extents showing and an average of 40 concurrent users. Since there is plenty of room to grow no action is needed.

64. You see multiple extents in the temporary tablespace. Is this a problem?

As long as they are all the same size this isn't a problem. In fact, it can even improve performance since Oracle won't have to create a new extent when a user needs one.

65. What operation query optimizer performs?

Optimizer performs following operations.

Evaluation of expressions & checks: Syntactic (Syntax of query), Semantic (objects exist and accessible)

Statement transformation: Transform view and sub-queries into a equivalent join / base tables

Choice of optimizer goals: Chooses throughput or best response time.

Choice of access paths: Chooses one or more available access paths

Choice of join orders: Chooses which pair of tables is joined first

66. What do you mean by "Throughput" or "Best response Time"?

Throughput-Default: (amount of work in a particular period of time). Best for Batch applications (Report application) because the user concerned with the time necessary for the application to complete.

Best response time: Least amount of resources necessary to process the first row accessed by a SQL statement. Best for interactive application (Forms application, OLTP) because user is waiting to see the few accessed by the statement.

67. How optimizer mode can be change?

The optimizer mode can be set by using initialization Parameter (OPTIMIZER_MODE) the following mode can be set.

RULE	The rule-based optimizer is used
CHOOSE	Optimizer use CBO with available statistics. RBO will be use instead if no statistics are available
ALL_ROWS	Optimizer uses CBO approach regardless of statistics with a goal of best throughput. Appropriate for data warehousing this is the default value of the parameter
FIRST_ROWS_n	Optimizer uses CBO approach, regardless statistics, with a goal of best response time to return the first n number of rows; (n = 1, 10, 100, or 1000). Appropriate for OLTP type queries
FIRST_ROWS	The optimizer uses a mix of cost and heuristics(learning themselves) to find a best plan for fast delivery of the first few rows. (for backward compatibility use FIRST_ROWS_n instead)

- If optimizer uses CBO approach with no statistics, then optimizer uses internal information (No of DB blocks allocated for these tables)
- It is recommended that always use the cost-based optimizer because it can recognize materialized views while RBO does not recognize them.

```
ALTER SESSION SET optimizer_mode = first_rows_1; --current session
```

```
ALTER SYSTEM SET optimizer_mode = first_rows_1; --instance level
```

68. What is optimizer mode SQL Hint and what it does?

Hints can be used to direct query optimizer to use a specific optimization technique for a query and can override the OPTIMIZER_MODE initialization parameter for that SQL statement. For example

```
SELECT /*+ ALL_ROWS */ empno, ename  
FROM emp;
```

If a hint is incorrect or invalid, Oracle ignores the hint without causing an error.

69. What are optimizer statistics how it can be collected?

Statistics are used by query optimizer to choose best execution plan. Statistics are stored in data dictionary.

DBMS_STATS Use to store/collect statistics in data dictionary for the use of query optimizer

ANALYZE Prior to Oracle8i used to collect statistics

You can collect statistics for the complete database or for particular objects (manually or automatically).

70. What do you mean by Histograms?

Histograms give the optimizer a more detailed view of the distribution of data values in the column. DBMS_STATS package can be used to create a histogram.

For table columns that contain values with large variations in number of duplicates, called **skewed data**.

71. What is an access path for query optimizer?

Access path are depend on the statement WHERE clause and it's FROM clause. Optimizer generates possible execution plans and estimates the cost of each plan using statistics. Finally optimizer chooses execution plan with the lowest estimated cost.

72. What are the different types of access path that can be followed by optimizer?

1- Full table scan

this type of scan reads all rows from a table up to the high water mark (HWM). The HWM marks the last block in the table that has ever had data written to it. Optimizer can use full table scan in case of Lack of Index, Large Amount of Data, and Small Table

2- ROW ID scan

the rowid of a row specifies the datafile and data block containing the row and the location of the row in that block. Oracle first obtains the rowids of the selected rows and then locates each selected row in the table based on its rowid. Optimizer uses when any columns in the statement not present in the index.

3- Index Scan

In this method, a row is retrieved by traversing the index. The index contains not only the indexed value, but also the rowids of rows.

Index Unique	This scan returns a single rowid if statement contains a UNIQUE or PRIMARY KEY constraint.
---------------------	--

Scans	
Index Range Scans	This scan can return more than one row. Data can be ascending or descending. Optimizer uses when it finds range operations (>, <, <>, >=, <=, between)
Full Scans	Optimizer chooses when the statistics indicate that it is going to be more efficient than a Full table scan. Optimizer chooses if statement is without where clause and all the select columns included in the index and At least one of the index columns is not null. (performs single block I/O)
Fast Full Index Scans	Alternative to a full table scan the difference is, it performs multi-block reads and can not be used against bitmap indexes
Index Joins	An index join is a hash join of several indexes that together contain all the table columns that are referenced in the query
Bitmap Indexes	A bitmap join uses a bitmap for key values and a mapping function that converts each bit position to rowid.

4- Cluster Access A cluster scan is used to retrieve, from a table stored in an indexed cluster

5- Hash Access A hash scan is used to locate rows in a hash cluster, based on a hash value

6- Sample Table Scan This access path is used when a statement's FROM clause includes the SAMPLE clause or the SAMPLE BLOCK clause.

--scan to access 1% of the employees table

```
SELECT * FROM employees SAMPLE BLOCK (1);
```

73. What is the explain plan? And what type of information explain plan contains?

The EXPLAIN PLAN statement displays execution plans chosen by the Oracle optimizer for SELECT, UPDATE, INSERT, and DELETE statements. A statement's execution plan is the sequence of operations Oracle performs to run the statement.

It shows the following information in a statement:

Ordering of the tables

Access method

Join method for tables

Data operations (filter, sort, or aggregation)

Optimization (cost and cardinality of each operation)

Partitioning (set of accessed partitions)

Parallel execution (distribution method of join inputs)

74. What is TKPROF?

Formats a trace file into a more readable format for performance analysis, before you can use TKPROF, you need to generate a trace file and locate it.

75. What is SQL Trace?

SQL trace files are text files and used to debug performance problems, execution plan and other statistics.

76. What is the Explain plan statement disadvantage?

Explain Plan is not as useful when used in conjunction with tkprof since the trace file contains the actual execution path of the SQL statement. Use Explain Plan when anticipated execution statistics are desired without actually executing the statement.

77. What is the Explain plan statement advantage?

Main advantage is that it does not actually run the query - just parses the SQL. In the early stages of tuning explain plan gives you an idea of the potential performance of your query without actually running it.

78. What is the plan table? Describe its purpose?

A global temporary table (automatically created) that Oracle fills when you issue "Explain plan" command for an SQL statement for all users.

79. How you can create plan table if plan table already not exists?

UTLXPLAN.SQL script if plan table not already exists and creates table Named PLAN_TABLE

```
SQL> CONN sys/password AS SYSDBA
```

```
SQL> @$ORACLE_HOME/rdbms/admin/utlxplan.sql
```

```
SQL> GRANT ALL ON sys.plan_table TO public;
```

```
SQL> CREATE PUBLIC SYNONYM plan_table FOR sys.plan_table;
```

80. What are the important fields of plan table?

Most important fields within the plan table are (operation, option, object_name,id, parent_id.)

81. How you run explain plan statement?

EXPLAIN PLAN FOR SELECT last_name FROM employees;

--Using EXPLAIN PLAN with the STATEMENT ID Clause

EXPLAIN PLAN SET STATEMENT_ID = 'st1'

FOR SELECT last_name FROM employees;

--Using EXPLAIN PLAN with the INTO Clause

EXPLAIN PLAN INTO my_plan_table

FOR SELECT last_name FROM employees;

82. What are the methods can be used to display plan table output (Execution Plan)?

The execution plan can be display by using following methods

1- Using simple query:(base is PLAN_TABLE)

Display execution plan for the the last "EXPLAIN PLAN" command. You need to format result yourself.

2- Utlxpls.sql or utlxplp.sql scripts (for serial or parallel queries) (base is PLAN_TABLE)

Displays the contents of a PLAN_TABLE. Makes it much easier to format and display execution plans.

@ORACLE_HOME\RDBMS\ADMIN\Utlxpls.sql --FOR serial queries

@ORACLE_HOME\RDBMS\ADMIN\utlxplp.sql --FOR parallel queries

Note: Executing individual scripts or using DBMS_XPLAN is same.

3- Using DBMS_XPLAN (As of 9i) (base is PLAN_TABLE)

DBMS_XPLAN.DISPLAY function that displays the contents of a PLAN_TABLE. Makes it much easier to format and display execution plans.

SELECT * FROM TABLE(DBMS_XPLAN.DISPLAY);

DBMS_XPLAN.DISPLAY_AWR Function look up an historical SQL statement captured in Oracle 10g's Automatic Workload Repository (AWR), and display its execution plan. This gives you a seven-day rolling window of history that you can access.

4- Using V\$SQL_PLAN Views (base is SQL Statement)

After the statement has executed V\$SQL_PLAN views can be used to display the execution plan of a SQL statement. Its definition is similar to the PLAN_TABLE. It is the actual execution plan and not the predicted one – just like tkprof and even better than Explain Plan.

V\$SQL_PLAN_STATISTICS provides actual execution statistics (output rows and time) for every operation

V\$SQL_PLAN_STATISTICS_ALL combines V\$SQL_PLAN and V\$SQL_PLAN_STATISTICS information

Both v\$sql_plan_statistics and v\$sql_plan_statistics_all are not populated by default. The option statistics_level=all must be set.

5- Using Toad (base is SQL Statement) TOOLS > SGA Trace / Optimization

83. Why and when should one tune?

One of the biggest responsibilities of a DBA is to ensure that the Oracle database is tuned properly. The Oracle RDBMS is highly tunable and allows the database to be monitored and adjusted to increase its performance.

One should do performance tuning for the following reasons:

The speed of computing might be wasting valuable human time (users waiting for response);

Enable your system to keep-up with the speed business is conducted; and

Optimize hardware usage to save money (companies are spending millions on hardware).

Although this FAQ is not overly concerned with hardware issues, one needs to remember than you cannot tune a Buick into a Ferrari.

84. What database aspects should be monitored?

One should implement a monitoring system to constantly monitor the following aspects of a database. Writing custom scripts, implementing Oracle's Enterprise Manager, or buying a third-party monitoring product can achieve this. If an alarm is triggered, the system should automatically notify the DBA (e-mail, page, etc.) to take appropriate action.

Infrastructure availability:

- Is the database up and responding to requests
- Are the listeners up and responding to requests
- Are the Oracle Names and LDAP Servers up and responding to requests
- Are the Web Listeners up and responding to requests

Things that can cause service outages:

- Is the archive log destination filling up?
- Objects getting close to their max extents
- Tablespace running low on free space/ Objects what would not be able to extend
- User and process limits reached

Things that can cause bad performance:

See question “What tuning indicators can one use?”.

85. Where should the tuning effort be directed?

Consider the following areas for tuning. The order in which steps are listed needs to be maintained to prevent tuning side effects. For example, it is no good increasing the buffer cache if you can reduce I/O by rewriting a SQL statement.

Database Design (if it's not too late):

Poor system performance usually results from a poor database design. One should generally normalize to the 3NF. Selective denormalization can provide valuable performance improvements. When designing, always keep the “data access path” in mind. Also look at proper data partitioning, data replication, aggregation tables for decision support systems, etc.

Application Tuning:

Experience showed that approximately 80% of all Oracle system performance problems are resolved by coding optimal SQL. Also consider proper scheduling of batch tasks after peak working hours.

Memory Tuning:

Properly size your database buffers (shared pool, buffer cache, log buffer, etc) by looking at your buffer hit ratios. Pin large objects into memory to prevent frequent reloads.

Disk I/O Tuning:

Database files need to be properly sized and placed to provide maximum disk subsystem throughput. Also look for frequent disk sorts, full table scans, missing indexes, row chaining, data fragmentation, etc.

Eliminate Database Contention:

Study database locks, latches and wait events carefully and eliminate where possible.

Tune the Operating System:

Monitor and tune operating system CPU, I/O and memory utilization. For more information, read the related Oracle FAQ dealing with your specific operating system.

86. What tuning indicators can one use?

The following high-level tuning indicators can be used to establish if a database is performing optimally or not:

- Buffer Cache Hit Ratio

Formula: Hit Ratio = (Logical Reads – Physical Reads) / Logical Reads

Action: Increase DB_CACHE_SIZE (DB_BLOCK_BUFFERS prior to 9i) to increase hit ratio

- Library Cache Hit Ratio

Action: Increase the SHARED_POOL_SIZE to increase hit ratio

87. What tools/utilities does Oracle provide to assist with performance tuning?

Oracle provides the following tools/ utilities to assist with performance monitoring and tuning:

- TKProf
- UTLBSTAT.SQL and UTLESTAT.SQL – Begin and end stats monitoring
- Statspack
- Oracle Enterprise Manager – Tuning Pack

88. What is STATSPACK and how does one use it?

Statspack is a set of performance monitoring and reporting utilities provided by Oracle from Oracle8i and above. Statspack provides improved BSTAT/ESTAT functionality, though the old BSTAT/ESTAT scripts are still available. For more information about STATSPACK, read the documentation in file \$ORACLE_HOME/rdbms/admin/spdoc.txt.

Install Statspack:

```
cd $ORACLE_HOME/rdbms/admin
```

```
sqlplus “/ as sysdba” @spdrop.sql — Install Statspack -
```

```
sqlplus “/ as sysdba” @spcreate.sql— Enter tablespace names when prompted
```

Use Statspack:

```
sqlplus perfstat/perfstat
```

```
exec statspack.snap; — Take a performance snapshots
```

```
exec statspack.snap;
```

- Get a list of snapshots

```
select SNAP_ID, SNAP_TIME from STATS$SNAPSHOT;
```

@spreport.sql — Enter two snapshot id's for difference report

Other Statspack Scripts:

- sppurge.sql – Purge a range of Snapshot Id's between the specified begin and end Snap Id's
- spauto.sql – Schedule a dbms_job to automate the collection of STATPACK statistics
- spcreate.sql – Installs the STATSPACK user, tables and package on a database (Run as SYS).
- spdrops.sql – Deinstall STATSPACK from database (Run as SYS)
- sppurge.sql – Delete a range of Snapshot Id's from the database
- spreport.sql – Report on differences between values recorded in two snapshots
- sprunc.sql – Truncates all data in Statspack tables.

89. When cost based optimization triggered?

It's important to have statistics on all tables for the CBO (Cost Based Optimizer) to work correctly. If one table involved in a statement does not have statistics, Oracle has to revert to rule-based optimization for that statement. So you really want for all tables to have statistics right away; it won't help much to just have the larger tables analyzed.

Generally, the CBO can change the execution plan when you:

1. Change statistics of objects by doing an ANALYZE;
2. Change some initialization parameters (for example: hash_join_enabled, sort_area_size, db_file_multiblock_read_count).

90. How can one optimize %XYZ% queries?

It is possible to improve %XYZ% queries by forcing the optimizer to scan all the entries from the index instead of the table. This can be done by specifying hints.

If the index is physically smaller than the table (which is usually the case) it will take less time to scan the entire index than to scan the entire table.

91. Where can one find I/O statistics per table?

The UTLESTAT report shows I/O per tablespace but one cannot see what tables in the tablespace has the most I/O. The \$ORACLE_HOME/rdbms/admin/catio.sql script creates a sample_io procedure and table to gather the required information. After executing the procedure, one can do a simple SELECT * FROM io_per_object; to extract the required information.

For more details, look at the header comments in the \$ORACLE_HOME/rdbms/admin/catio.sql script.

92. My query was fine last week and now it is slow. Why?

The likely cause of this is because the execution plan has changed. Generate a current explain plan of the offending query and compare it to a previous one that was taken when the query was performing well. Usually the previous plan is not available.

Some factors that can cause a plan to change are:

- which tables are currently analyzed? Were they previously analyzed? (ie. Was the query using RBO and now CBO?)
- Has OPTIMIZER_MODE been changed in INIT.ORA?
- Has the DEGREE of parallelism been defined/changed on any table?
- Have the tables been re-analyzed? Were the tables analyzed using estimate or compute? If estimate, what percentage was used?
- Have the statistics changed?
- Has the INIT.ORA parameter DB_FILE_MULTIBLOCK_READ_COUNT been changed?
- Has the INIT.ORA parameter SORT_AREA_SIZE been changed?
- Have any other INIT.ORA parameters been changed?
- What do you think the plan should be? Run the query with hints to see if this produces the required performance.

93. Why is Oracle not using the damn index?

This problem normally only arises when the query plan is being generated by the Cost Based Optimizer. The usual cause is because the CBO calculates that executing a Full Table Scan would be faster than accessing the table via the index. Fundamental things that can be checked are:

- USER_TAB_COLUMNS.NUM_DISTINCT – This column defines the number of distinct values the column holds.
- USER_TABLES.NUM_ROWS – If NUM_DISTINCT = NUM_ROWS then using an index would be preferable to doing a FULL TABLE SCAN. As the NUM_DISTINCT decreases, the cost of using an index increase thereby is making the index less desirable.
- USER_INDEXES.CLUSTERING_FACTOR – This defines how ordered the rows are in the index. If CLUSTERING_FACTOR approaches the number of blocks in the table, the rows are ordered. If it approaches the number of rows in the table, the rows are randomly ordered. In such a case, it is unlikely that index entries in the

same leaf block will point to rows in the same data blocks.

- Decrease the INIT.ORA parameter DB_FILE_MULTIBLOCK_READ_COUNT – A higher value will make the cost of a FULL TABLE SCAN cheaper.
- Remember that you MUST supply the leading column of an index, for the index to be used (unless you use a FAST FULL SCAN or SKIP SCANNING).
- There are many other factors that affect the cost, but sometimes the above can help to show why an index is not being used by the CBO. If from checking the above you still feel that the query should be using an index, try specifying an index hint. Obtain an explain plan of the query either using TKPROF with TIMED_STATISTICS, so that one can see the CPU utilization, or with AUTOTRACE to see the statistics. Compare this to the explain plan when not using an index.

94. When should one rebuild an index?

You can run the 'ANALYZE INDEX VALIDATE STRUCTURE' command on the affected indexes – each invocation of this command creates a single row in the INDEX_STATS view. This row is overwritten by the next ANALYZE INDEX command, so copy the contents of the view into a local table after each ANALYZE. The 'badness' of the index can then be judged by the ratio of 'DEL_LF_ROWS' to 'LF_ROWS'.

95. How does one tune Oracle Wait events?

Some wait events from V\$SESSION_WAIT and V\$SYSTEM_EVENT views:

Event Name: Tuning Recommendation:

db file sequential read Tune SQL to do less I/O. Make sure all objects are analyzed. Redistribute I/O across disks.
buffer busy waits Increase DB_CACHE_SIZE (DB_BLOCK_BUFFERS prior to 9i)/ Analyze contention from SYS.V\$BH
log buffer space Increase LOG_BUFFER parameter or move log files to faster disks

96. What is the difference between DBFile Sequential and Scattered Reads?

Both "db file sequential read" and "db file scattered read" events signify time waited for I/O read requests to complete. Time is reported in 100's of a second for Oracle 8i releases and below, and 1000's of a second for Oracle 9i and above. Most people confuse these events with each other as they think of how data is read from disk. Instead they should think of how data is read into the SGA buffer cache.

db file sequential read:

A sequential read operation reads data into contiguous memory (usually a single-block read with p3=1, but can be multiple blocks). Single block I/Os are usually the result of using indexes. This event is also used for rebuilding the controlfile and reading datafile headers (P2=1). In general, this event is indicative of disk contention on index reads.

db file scattered read:

Similar to db file sequential reads, except that the session is reading multiple data blocks and scatters them into different discontinuous buffers in the SGA. This statistic is NORMALLY indicating disk contention on full table scans. Rarely, data from full table scans could be fitted into a contiguous buffer area, these waits would then show up as sequential reads instead of scattered reads.

The following query shows average wait time for sequential versus scattered reads:

prompt "AVERAGE WAIT TIME FOR READ REQUESTS"

select a.average_wait "SEQ READ", b.average_wait "SCAT READ"

from sys.v_\$system_event a, sys.v_\$system_event b

where a.event = 'db file sequential read'

and b.event = 'db file scattered read';

97. What is the use of statistics?

98. How to generate explain plan?

99. How to check explain plan of already ran SQLs?

100. How to find out whether the query has ran with RBO or CBO?

101. What are top 5 wait events (in AWR report) and how you will resolve them?

<http://satya-dba.blogspot.in/2012/10/wait-events-in-oracle-wait-events.html>

db file sequential read => tune indexing, tune SQL (to do less I/O), tune disks, increase buffer cache. This event is indicative of disk contention on index reads. Make sure all objects are analyzed. Redistribute I/O across disks. The wait that comes from the physical side of the database. It related to memory starvation and non selective index use. Sequential read is an index read followed by table read because it is doing index lookups which tells exactly which block to go to.

db file scattered read => disk contention on full table scans. Add indexes, tune SQL, tune disks, refresh statistics, and create materialized view. Caused due to full table scans may be because of insufficient indexes or unavailability of updated statistics.

db file parallel read => tune SQL, tune indexing, tune disk I/O, increase buffer cache. If you are doing a lot of partition activity then expect to see that wait even. It could be a table or index partition.

db file parallel write => if you are doing a lot of partition activity then expect to see that wait even. It could be a table or index partition.

db file single write => if you see this event than probably you have a lot of data files in your database.

control file sequential read

control file parallel write

log file sync => committing too often, archive log generation is more. Tune applications to commit less, tune disks where redo logs exist, try using nologging/unrecoverable options, log buffer could be too large.

log file switch completion => May need more log files per group.

log file parallel write => Deals with flushing out the redo log buffer to disk. Disks may be too slow or have an I/O bottleneck. Look for log file contention.

log buffer space => Increase LOG_BUFFER parameter or move log files to faster disks. Tune application, use NOLOGGING, and look for poor behavior that updates an entire row when only a few columns change.

log file switch (checkpoint incomplete) => May indicate excessive db files or slow IO subsystem.

log file switch (archiving needed) => Indicates archive files are written too slowly.

redo buffer allocation retries => shows the number of times a user process waited for space in the redo log buffer.

redo log space wait time => shows cumulative time (in 10s of milliseconds) waited by all processes waiting for space in the log buffer.

buffer busy waits/ read by other session => Increase DB_CACHE_SIZE. Tune SQL, tune indexing, we often see this event along with full table scans, if the SQL is inserting data, consider increasing FREELISTS and/or INITRANS, if the waits are on segment header blocks, consider increasing extent sizes.

free buffer waits => insufficient buffers, process holding buffers too long or i/o subsystem is over loaded. Also check you db writes may be getting clogged up.

cache buffers lru chain => Freelist issues, hot blocks.

no free buffers => Insufficient buffers, dbwr contention.

latch free

latch: session allocation

latch: in memory undo latch => If excessive could be bug, check for your version, may have to turn off in memory undo.

latch: cache buffer chains => check hot objects.

latch: cache buffer handles => Freelist issues, hot blocks.

direct path write => You wont see them unless you are doing some appends or data loads.

direct Path reads => could happen if you are doing a lot of parallel query activity.

direct path read temp or direct path write temp => this wait event shows Temp file activity (sort,hashes,temp tables, bitmap) check pga parameter or sort area or hash area parameters. You might want to increase them.

library cache load lock

library cache pin => if many sessions are waiting, tune shared pool, if few sessions are waiting, lock is session specific.

library cache lock => need to find the session holding the lock, look for DML manipulating an object being accessed, if the session is trying to recompile PL/SQL, look for other sessions executing the code.

undo segment extension => If excessive, tune undo.

wait for a undo record => Usually only during recovery of large transactions, look at turning off parallel undo recovery.

enqueue wait events => Look at V\$ENQUEUE_STAT

SQL*Net message from client

SQL*Net message from dblink

SQL*Net more data from client

SQL*Net message to client

SQL*Net break/reset to client

102. What are the init parameters related to performance/optimizer?

optimizer_mode = choose

optimizer_index_caching = 90

optimizer_index_cost_adj = 25

optimizer_max_permutations = 100

```
optimizer_use_sql_plan_baselines=true  
optimizer_capture_sql_plan_baselines=true  
optimizer_use_pending_statistics = true;  
optimizer_use_invisible_indexes=true  
_optimizer_connect_by_cost_based=false  
_optimizer_compute_index_stats= true;
```

103. What are the values of optimizer_mode init parameters and their meaning?

optimizer_mode = choose

104. What is the use of AWR, ADDM, and ASH?

105. How to generate AWR report and what are the things you will check in the report?

106. How to generate ADDM report and what are the things you will check in the report?

107. How to generate ASH report and what are the things you will check in the report?

108. How to generate STATSPACK report and what are the things you will check in the report?

109. How to generate TKPROF report and what are the things you will check in the report?

The tkprof tool is a tuning tool used to determine cpu and execution times for SQL statements. Use it by first setting timed_statistics to true in the initialization file and then turning on tracing for either the entire database via the sql_trace parameter or for the session using the ALTER SESSION command. Once the trace file is generated you run the tkprof tool against the trace file and then look at the output from the tkprof tool. This can also be used to generate explain plan output.

110. What is Performance Tuning?

Making optimal use of system using existing resources called performance tuning.

111. Types of Tunings?

1. CPU Tuning 2. Memory Tuning 3. IO Tuning 4. Application Tuning 5. Database Tuning

112. What mainly Database Tuning contains?

1. Hit Ratios 2. Wait Events

113. What is an optimizer?

Optimizer is a mechanism which will make the execution plan of an sql statement

114. Types of Optimizers?

1. RBO(Rule Based Optimizer) 2. CBO(Cost Based Optimizer)

115. Which init parameter is used to make use of Optimizer?

optimizer_mode= rule----RBO cost---CBO choose-----First CBO otherwise RBO

116. Which optimizer is the best one?

CBO

117. What are the prerequisites to make use of Optimizer?

1. Set the optimizer mode 2. Collect the statistics of an object

118. How do you collect statistics of a table?

analyze table emp compute statistics or analyze table emp estimate statistics

119. What is the difference between compute and estimate?

If you use compute, The FTS will happen, if you use estimate just 10% of the table will be read

120. What will happen if you set the optimizer_mode=choose?

If the statistics of an object is available then CBO used, if not RBO will be used.

121. Data Dictionary follows which optimizer mode?

RBO

122. How do you delete statistics of an object?

analyze table emp delete statistics

123. How do you collect statistics of a user/schema?

EXEC DBMS_STATS.GATHER_SCHEMA_STATS (SCOTT)

124. How do you see the statistics of a table?

select num_rows, blocks, empty_blocks from dba_tables where tab_name='emp'

125. What are chained rows?

These are rows, it spans in multiple blocks

126. How do you collect statistics of a user in Oracle Apps?

fnl_stats package

127. How do you create an execution plan and how do you see?

1. @?/rdbms/admin/utlxplan.sql ----- it creates a plan_table
2. explain set statement_id='1' for select * from emp;
3. @?/rdbms/admin/utlxpls.sql -----it display the plan

128. How do you know what sql is currently being used by the session?

by going v\$sql and v\$sql_area

129. What is a execution plan?

Its a road map how sql is being executed by oracle db?

130. How do you get the index of a table and on which column the index is?

dba_indexes and dba_ind_columns

131. Which init paramter you have to set to by pass parsing?

cursor_sharing=force

132. How do you know which session is running long jobs?

by going v\$session_longops

133. How do you flush the shared pool?

alter system flush shared_pool

134. How do you get the info about FTS?

using v\$sysstat

135. How do you increase the db cache?

alter table emp cache

136. Where do you get the info of library cache?

v\$librarycache

137. How do you get the information of specific session?

v\$mystat

138. What you'll check whenever user complains that his session/database is slow?

139. Customer reports a application slowness issue, and you need to evaluate database performance. What do you look at for 9i and for 11g.

On oracle 9i, look at the statspack report , on oracle 11g look at the AWR report. In both reports look top sqls listed in elapsed time or cpu time. At sqlplus look sql_text from v\$sql where disk_reads is high.

140. You have found a long running sql in your evaluation of system health of database, what do you look for to determine why sql is slow?

Use explain plan to determine the execution plan of the sql. When looking at execution plan look for indexes being used, full table scans on large tables.

141. You have a windows service is crashing, how can you determine the sqls related to the windows service?

Use sql trace to trace the username and program associated with the trace file. Use tkprof to analyze the sql trace and determine the long running sqls.

142. What is proactive tuning and reactive tuning?

In Proactive Tuning, the application designers can then determine which combination of system resources and available Oracle features best meet the needs during design and development. In reactive tuning the bottom up approach is used to find and fix the bottlenecks. The goal is to make Oracle run faster.

143. Describe the level of tuning in oracle?

A.System-level tuning involves the following steps:

1. Monitoring the operating system counters using a tool such as top, gtop, and GKrellM or the VTune analyzer's counter monitor data collector for applications running on Windows.
2. Interpreting the counter data to locate system-level performance bottlenecks and opportunities for improving the way your application interacts with the system.
- 3.SQL-level tuning:Tuning disk and network I/O subsystem to optimize the I/O time, network packet size and dispatching frequency is called the server kernel optimization.

Distribution of data can be studied by the optimizer by collecting and storing optimizer statistics. This enables intelligent execution plans. Choice of db_block_size, db_cache_size, and OS parameters

(db_file_multiblock_read_count, cpu_count, &c), can influence SQL performance. Tuning SQL Access workload with physical indexes and materialized views.

144. What is Database design level tuning?

The steps involved in database design level tuning are:

1. Determination of the data needed by an application (what relations are important, their attributes and structuring the data to best meet the performance goals)

2. Analysis of data followed by normalization to eliminate data redundancy.
3. Avoiding data contention.
4. Localizing access to the data to the partition, process and instance levels.
5. Using synchronization points in Oracle Parallel Server.
6. Implementation of 8i enhancements that can help avoid contention are:
 - a. Consideration on partitioning the data
 - b. Consideration over using local or global indexes.

145. Explain rule-based optimizer and cost-based optimizer?

A. Oracle decides how to retrieve the necessary data whenever a valid SQL statement is processed. This decision can be made using one of two methods:

1. Rule Based Optimizer

if the server has no internal statistics relating to the objects referenced by the statement then the RBO method is used. This method will be deprecated in the future releases of oracle.

2. Cost Based Optimizer

The CBO method is used if internal statistics are present. The CBO checks several possible execution plans and selects the one with the lowest cost based on the system resources.

146. What are object datatypes? Explain the use of object datatypes?

Object data types are user defined data types. Both column and row can represent an object type. Object types instance can be stored in the database. Object datatypes make it easier to work with complex data, such as images, audio, and video. Object types provide higher-level ways to organize and access data in the database. The SQL attributes of Select into clause, i.e. SQL % Not found, SQL % found, SQL % Isopen, SQL % Rowcount.

1. % Not found: True if no rows returned

E.g. If SQL%NOTFOUND then return some_value

2. % found: True if at least one or more rows returned

E.g. If SQL%FOUND then return some_value

3. %Isopen: True if the SQL cursor is open. Will always be false, because the database opens and closes the implicit cursor used to retrieve the data

4. %Rowcount: Number of rows returned. Equals 0 if no rows were found (but the exception is raised) and a 1, if one or more rows are found (if more than one an exception is raised).

147. What is translate and decode in oracle?

1. Translate: translate function replaces a sequence of characters in a string with another set of characters. The replacement is done single character at a time. Syntax:

`translate(string1, string_to_replace, replacement_string)`

Example:

`translate ('1tech23', '123', '456');`

2. Decode: The DECODE function compares one expression to one or more other expressions and, when the base expression is equal to a search expression, it returns the corresponding result expression; or, when no match is found, returns the default expression when it is specified, or NA when it is not.

Syntax:

`DECODE (expr , search, result [, search , result]... [, default])`

Example:

`SELECT employee_name, decode(employee_id, 10000, 'tom', 10001, 'peter', 10002, 'jack' 'Gateway') result FROM employee;`

148. What is oracle correlated sub-queries? Explain with an example?

A query which uses values from the outer query is called as a correlated sub query. The subquery is executed once and uses the results for all the evaluations in the outer query. Example:

Here, the sub query references the employee_id in outer query. The value of the employee_id changes by row of the outer query, so the database must rerun the subquery for each row comparison. The outer query knows nothing about the inner query except its results.

```
select employee_id, appraisal_id, appraisal_amount From employee
where
appraisal_amount < (select max(appraisal_amount)
from employee e
where employee_id = e. employee_id);
```

149. Explain union and intersect with examples?

1. UNION: The UNION operator is used to combine the result-set of two or more SELECT statements. Tables of both the select statement must have the same number of columns with similar data types. It eliminates duplicates. Syntax:

```
SELECT column_name(s) FROM table_name1
```

```
UNION
```

```
SELECT column_name(s) FROM table_name2
```

Example:

```
SELECT emp_Name FROM Employees_india
```

```
UNION
```

```
SELECT emp_Name FROM Employees_USA
```

2. INTERSECT allows combining results of two or more select queries. If a record exists in one query and not in the other, it will be omitted from the INTERSECT results.

150. What is difference between open_form and call_form? What is new_form built-in in oracle form?

A. Open_form opens the indicated form. Call_form not just opens the indicated form, but also keeps the parent form alive. When new_form is called, the new indicated form is opened and the old one is exited by releasing the memory. The new form is run using the same Run form options as the parent form.

151. What is advantage of having disk shadowing/ Mirroring in oracle?

A. Fast recovery of data in case of Disk failure. Improved performance since most OS supports volume shadowing that can direct file I/O request to use the shadow set of files instead of the main set of files.

AWR vs ADDM vs ASH

AWR : automatic workload repository

The AWR is used to collect performance statistics including:

- Wait events used to identify performance problems.
- Time model statistics indicating the amount of DB time associated with a process from the V\$SESS_TIME_MODEL and V\$SYS_TIME_MODEL views.
- Active Session History (ASH) statistics from the V\$ACTIVE_SESSION_HISTORY view.
- Some system and session statistics from the V\$SYSSTAT and V\$SESSTAT views.
- Object usage statistics.
- Resource intensive SQL statements.

I will not get into Details how to generate AWR since i mention it before on my Blog .

ADDM : automatic database diagnostic monitor

analyzes data in the Automatic Workload Repository (AWR) to identify potential performance bottlenecks. and we use it for the following cases :

- CPU bottlenecks
- Undersized memory structures
- I/O capacity issues
- High load SQL statements
- RAC specific issues
- Database configuration issues
- Also provides recommendations on hardware changes, database configuration & schema changes.

Generate ADDM :

- Login to SQL
- @\$ORACLE_HOME/rdbms/admin/addmrpt.sql
- enter system password when you asked for .
- Specify a begin_snap from the list and press Enter.
- Specify the end_snap from the list and press Enter.
- Report Name

ASH : Active Session History

statistics from the in-memory performance monitoring tables also used to track session activity and simplify performance tuning.

ASH reports Give the following information :

- Top User Events (frequent wait events)
- Details to the wait events
- Top Queries
- Top Sessions
- Top Blocking Sessions
- Top DB Object.
- Activity Over Time

Generate ASH reports :

The Best way to do that using OEM. (Enterprise manager).

Oracle Installaion FAQ

1. What are the pre-requisites of Oracle installation?
2. What are the scripts to run to successful of software installation? Explain them?
3. What are post installation Tasks?
4. Explain the steps for manual/silent installation of oracle software?
5. Explain the steps for manual creation of oracle database?
6. What is use of catalog.sql, catproc.sql and pupbld.sql? Explain?
7. What is ORAINVENTORY and the default location?
8. What is ORAINST.LOC file? Explain?
9. Explain the manual un-installation of oracle software without using uninstall Wizard?
10. Explain the manual deletion of oracle database?
11. If Oracle inventory is corrupted or missing? How to recover?

Answers

1. What are the pre-requisites of Oracle installation? (http://docs.oracle.com/html/B15521_01/toc.htm) (<http://naveenkumarsr.wordpress.com/2010/05/11/oracle10g-silent-installation/>)

- ❏ [Log In to the System as root](#)
- ❏ [Check the Hardware Requirements](#)
- ❏ [Check the Software Requirements](#)
- ❏ [Create Required UNIX Groups and User](#)
- ❏ [Create Required Directories](#)
- ❏ [Configure Kernel Parameters](#)
- ❏ [Mount the Product Disc](#)
- ❏ [Log In as the oracle User and Configure the oracle User's Environment](#)
- ❏ [Install Oracle Database 10g](#)
- ❏ [Install Products from the Oracle Database 10g Companion CD](#)

Log In to the System as root

Before you install the Oracle software, you must complete several tasks as the root user. To log in as the root user, complete one of the following procedures:

Note:

You must install the software from an X window workstation, an X terminal, or a PC or other system with X server software installed.

If you are installing the software from an X Window System workstation or X terminal:

Start a local terminal session, for example, an X terminal (xterm).

If you are not installing the software on the local system, enter the following command to enable remote hosts to display X applications on the local X server:

```
$ xhost +
```

If you are not installing the software on the local system, use the ssh, rlogin, or telnet command to connect to the system where you want to install the software:

```
$ telnet remote_host
```

If you are not logged in as the root user, enter the following command to switch user to root:

```
$ su - root
```

```
password:
```

```
#
```

If you are installing the software from a PC or other system with X server software installed:

Note:

If necessary, see your X server documentation for more information about completing this procedure. Depending on the X server software that you are using, you may need to complete the tasks in a different order.

Start the X server software.

Configure the security settings of the X server software to permit remote hosts to display X applications on the local system.

Connect to the remote system where you want to install the software and start a terminal session on that system, for example, an X terminal (xterm).

If you are not logged in as the root user on the remote system, enter the following command to switch user to root:

```
$ su - root
```

```
password:
```

Check the Hardware Requirements

The system must meet the following minimum hardware requirements:

Requirement	Minimum Value
Physical memory (RAM)	512 MB (524288 KB)

Requirement	Minimum Value
Swap space	1 GB (1048576 KB) or twice the size of RAM On systems with 2 GB or more of RAM, the swap space can be between one and two times the size of RAM
Disk space in /tmp	400 MB (409600 KB)
Disk space for software files	2.5 GB (2621440 KB) This value includes 1 GB (1048576 KB) of disk space required to install the Oracle Database 10g Products from the Companion CD (optional, but recommended).
Disk space for database files	1.2 GB (1258290 KB)

To ensure that the system meets these requirements, follow these steps:

To determine the physical RAM size, enter the following command:

```
# grep MemTotal /proc/meminfo
```

If the size of the physical RAM installed in the system is less than 512 MB, you must install more memory before continuing.

To determine the size of the configured swap space, enter the following command:

```
# grep SwapTotal /proc/meminfo
```

If necessary, see your operating system documentation for information about how to configure additional swap space.

To determine the amount of free disk space available in the /tmp directory, enter the following command:

```
# df -h /tmp
```

If there is less than 400 MB of disk space available in the /tmp directory, complete one of the following steps:

Delete unnecessary files from the /tmp directory to achieve the required disk space.

Set the TEMP and TMPDIR environment variables when setting the oracle user's environment (described later).

Extend the file system that contains the /tmp directory. If necessary, contact your system administrator for information about extending file systems.

To determine the amount of free disk space available on the system, enter the following command:

```
# df -h
```

This command displays the disk space usage on all mounted file systems. To complete the installation, the system must satisfy either of the following conditions:

3.7 GB (3879731 KB) of free disk space is available on two file systems: one with at least 2.5 GB (2621440 KB) free for the Oracle software and another with at least 1.2 GB free for the preconfigured database

3.7 GB of free disk space is available for the Oracle software and database on a single file system

Note:

While installing the Oracle database on a disk drive separate from the software does provide a performance improvement, for best performance, the Oracle database files should be distributed across three or more disks. The Oracle Database Installation Guide for UNIX Systems describes this more complex and time-consuming type of installation. However, this type of installation is recommended only for experienced users.

Check the Software Requirements

The system must meet the following minimum software requirements, depending on your Linux distribution and version.

Red Hat Enterprise Linux ES/AS 2.1 (Update 3 or higher)

Kernel version 2.4.9 errata 34 (e.34) or higher must be installed

The following packages (or later versions) must be installed:

make-3.79

openmotif-2.1.30

gcc-2.96-128

gcc-c++-2.96-128

libstdc++-2.96-128

glibc-2.2.4-32

Red Hat Enterprise Linux ES/AS 3 (Update 2 or higher)

Kernel version 2.4.21-15 or higher must be installed
The following packages (or later versions) must be installed:

gcc-3.2.3-34
gcc-c++-3.2.3-34
glibc-2.3.2-95.20
make-3.79.1
openmotif21-2.1.30-8
setarch-1.3-1
compat-db-4.0.14-5
compat-gcc-7.3-2.96.128
compat-gcc-c++-7.3-2.96.128
compat-libstdc++-7.3-2.96.128
compat-libstdc++-devel-7.3-2.96.128
SUSE Linux Enterprise Server 8 (Service Pack 3 or higher)

Kernel version 2.4.21-138 or higher must be installed
The following packages (or higher versions) must also be installed:

gcc-3.2.2-38
gcc-c++-3.2.2-38
glibc-2.2.2-124
make-3.79.1
openmotif-2.2.2-124
SUSE Linux Enterprise Server 9

Kernel version 2.6.5-7.5 or higher must be installed
The following packages (or higher versions) must also be installed:

gcc-3.3.3-43
gcc-c++-3.3.3-43
glibc-2.3.3-98
libaio-0.3.98-18
libaio-devel-0.3.98-18
make-3.80
openmotif-libs-2.2.2-519.1

To ensure that the system meets these requirements, follow these steps:

To determine which distribution and version of Linux is installed, enter the following command:

```
# cat /etc/issue
```

Note:

Only the listed distributions and versions are currently certified and supported.

To determine whether the required packages are installed, enter commands similar to the following:

```
$ rpm -q package_name
```

If a required package is not installed, or if the version is lower than the required version, install the package from your operating system distribution media or download the required package version from your Linux vendor's Web site.

To determine whether the required kernel version is installed, enter the following command:

```
# uname -r
```

If the kernel version is lower than the required version, download and install the required version or a higher version from your Linux vendor's Web site.

Create Required UNIX Groups and User

The following local UNIX groups and user must exist on the system:

The oinstall group (the Oracle Inventory group)

The dba group (the OSDBA group)

The oracle user (the Oracle software owner)

The oinstall and dba groups and the oracle user may already exist on your system. To determine whether they exist already, and if necessary, to create them, follow these steps:

To determine whether the oinstall and dba groups exist, enter the following commands:

```
# grep oinstall /etc/group
```

```
# grep dba /etc/group
```

If the output from these commands shows the specified group name, that group already exists.

If necessary, enter the following commands to create the oinstall and dba groups:

```
# /usr/sbin/groupadd oinstall
```

```
# /usr/sbin/groupadd dba
```

To determine whether the oracle user exists and belongs to the correct groups, enter the following command:

```
# id oracle
```

If the oracle user exists, this command displays information about the groups to which the user belongs. The output should be similar to the following, indicating that oinstall is the primary group and dba is a secondary group:

```
uid=502(oracle) gid=502(oinstall) groups=502(oinstall),503(dba)
```

If necessary, complete one of the following actions:

If the oracle user exists, but its primary group is not oinstall or it is not a member of the dba group, enter the following command:

```
# /usr/sbin/usermod -g oinstall -G dba oracle
```

If the oracle user does not exist, enter the following command to create it:

```
# /usr/sbin/useradd -g oinstall -G dba oracle
```

This command creates the oracle user and specifies oinstall as the primary group and dba as the secondary group.

Enter the following command to set the password of the oracle user:

```
# passwd oracle
```

Create Required Directories

Create directories with names similar to the following and specify the correct owner, group, and permissions for them:

```
/u01/app/oracle (the Oracle base directory)
```

```
/u02/oradata (an optional Oracle datafile directory)
```

The Oracle base directory must have 2.5 GB (2621440 KB) of free disk space, or 3.7 GB (3879731 KB) of free disk space if you choose not to create a separate Oracle datafile directory. The Oracle datafile directory must have 1.2 GB of free disk space.

Note:

If you do not want to create a separate Oracle datafile directory, you can install the datafiles in a subdirectory of the Oracle base directory (not recommended for production databases).

To determine where to create these directories, follow these steps:

Enter the following command to display information about all mounted file systems:

```
# df -h
```

This command displays information about all of the file systems mounted on the system, including:

The physical device name

The total amount, used amount, and available amount of disk space

The mount point directory for that file system

From the display, identify either one or two file systems that meet the following requirements:

Two file systems:

Identify one file system with 2.5 GB of free disk space, for the Oracle base directory, and another file system with 1.2 GB of free disk space for the Oracle datafile directory.

One file system:

Identify one file system with 3.7 GB of free disk space, for both the Oracle base directory and the Oracle datafile directory.

Note the name of the mount point directory for each file system that you identified.

In the following examples, /u01 is the mount point directory used for the software and /u02 is the mount point directory used for the Oracle datafile directory. You must specify the appropriate mount point directories for the file systems on your system.

To create the required directories and specify the correct owner, group, and permissions for them, follow these steps:

Note:

In the following procedure, replace /u01 and /u02 with the appropriate mount point directories that you identified in Step 3 previously.

Enter the following command to create subdirectories in the mount point directory that you identified for the Oracle base directory:

```
# mkdir -p /u01/app/oracle
```

If you intend to use a second file system for the Oracle database files, create an oradata subdirectory in the mount point directory that you identified for the Oracle datafile directory (shown as /u02 in the examples):

```
# mkdir /u02/oradata
```

Change the owner and group of the directories that you created to the oracle user and the oinstall group:

```
# chown -R oracle:oinstall /u01/app/oracle
```

```
# chown -R oracle:oinstall /u02/oradata
```

Change the permissions on the directories that you created to 775:

```
# chmod -R 775 /u01/app/oracle
```

```
# chmod -R 775 /u02/oradata
```

Configure Kernel Parameters

Verify that the kernel parameters shown in the following table are set to values greater than or equal to the recommended value shown. The procedure following the table describes how to verify and set the values.

Parameter	Value	File
semmsl semmns semopm semmni	250 32000 100 128	/proc/sys/kernel/sem
shmall	2097152	/proc/sys/kernel/shmall
shmmax	Half the size of physical memory (in bytes)	/proc/sys/kernel/shmmax
shmmni	4096	/proc/sys/kernel/shmmni
file-max	65536	/proc/sys/fs/file-max
ip_local_port_range	1024 65000	/proc/sys/net/ipv4/ip_local_port_range

Note:

If the current value for any parameter is higher than the value listed in this table, do not change the value of that parameter.

To view the current value specified for these kernel parameters, and to change them if necessary, follow these steps: Enter commands similar to the following to view the current values of the kernel parameters:

Note:

Make a note of the current values and identify any values that you must change.

Parameter	Command
semmsl, semmns, semopm, and semmni	# /sbin/sysctl -a grep sem This command displays the value of the semaphore parameters in the or listed.
shmall, shmmax, and shmmni	# /sbin/sysctl -a grep shm
file-max	# /sbin/sysctl -a grep file-max
ip_local_port_range	# /sbin/sysctl -a grep ip_local_port_range This command displays a range of port numbers.

If the value of any kernel parameter is different to the recommended value, complete the following steps: Using any text editor, create or edit the /etc/sysctl.conf file and add or edit lines similar to the following:

Note:

Include lines only for the kernel parameter values that you want to change. For the semaphore parameters (kernel.sem), you must specify all four values. However, if any of the current values are larger than the recommended value, specify the larger value.

```
kernel.shmall = 2097152
kernel.shmmax = 2147483648
kernel.shmmni = 4096
kernel.sem = 250 32000 100 128
fs.file-max = 65536
net.ipv4.ip_local_port_range = 1024 65000
```

By specifying the values in the `/etc/sysctl.conf` file, they persist when you reboot the system.
Enter the following command to change the current values of the kernel parameters:

```
# /sbin/sysctl -p
```

Review the output from this command to verify that the values are correct. If the values are incorrect, edit the `/etc/sysctl.conf` file, then enter this command again.

On SUSE systems only, enter the following command to cause the system to read the `/etc/sysctl.conf` file when it reboots:

```
# /sbin/chkconfig boot.sysctl on
```

Set Shell Limits for the oracle User

To improve the performance of the software on Linux systems, you must increase the following shell limits for the oracle user:

Shell Limit	Item in limits.conf	Hard Limit
Maximum number of open file descriptors	nofile	65536
Maximum number of processes available to a single user	nproc	16384

To increase the shell limits:

Add the following lines to `/etc/security/limits.conf` file:

```
oracle    soft  nproc  2047
oracle    hard  nproc  16384
oracle    soft  nofile 1024
oracle    hard  nofile 65536
```

Add the following line to the `/etc/pam.d/login` file, if it does not already exist:

```
session required /lib/security/pam_limits.so
```

Depending on the oracle user's default shell, make the following changes to the default shell start-up file:

For the Bourne, Bash, or Korn shell, add the following lines to the `/etc/profile` file (or the `/etc/profile.local` file on SUSE systems):

```
if [ $USER = "oracle" ]; then
    if [ $SHELL = "/bin/ksh" ]; then
        ulimit -p 16384
        ulimit -n 65536
    else
        ulimit -u 16384 -n 65536
    fi
fi
```

For the C or tcsh shell, add the following lines to the `/etc/csh.login` file (or the `/etc/csh.login.local` file on SUSE systems):

```
if ( $USER == "oracle" ) then
    limit maxproc 16384
    limit descriptors 65536
endif
```

Mount the Product Disc

The Oracle Database 10g software is available on both CD-ROM and DVD-ROM. These discs are in ISO 9660 format with Rockridge extensions.

On most Linux systems, the product disc mounts automatically when you insert it into the drive. To verify that the disc is mounted correctly, follow these steps:

If necessary, enter a command similar to following to eject the currently mounted disc, then remove it from the drive:

Red Hat:

```
# eject /mnt/cdrom
```

SUSE:

```
# eject /media/cdrom
```

In this example, /mnt/cdrom or /media/cdrom is the mount point directory for the CD-ROM drive, depending on your distribution.

Insert the disc into the CD-ROM or DVD-ROM drive.

To verify that the disc mounted automatically, enter a command similar to the following:

Red Hat:

```
$ ls /mnt/cdrom
```

SUSE:

```
$ ls /media/cdrom
```

If this command fails to display the contents of the disc, enter a command similar to the following, depending on your distribution:

Red Hat:

```
# mount /mnt/cdrom
```

SUSE:

```
# mount /media/cdrom
```

Log In as the oracle User and Configure the oracle User's Environment

You run the Installer from the oracle account. However, before you start the Installer you must configure the environment of the oracle user. To configure the environment, you must:

Set the default file mode creation mask (umask) to 022 in the shell startup file.

Set the DISPLAY, ORACLE_BASE, and ORACLE_SID environment variables.

To set the oracle user's environment, follow these steps:

Start another terminal session.

Enter the following command to ensure that X Window applications can display on this system:

```
$ xhost +
```

Complete one of the following steps:

If the terminal session is not connected to the system where you want to install the software, log in to that system as the oracle user.

If the terminal session is connected to the system where you want to install the software, switch user to oracle:

```
$ su - oracle
```

To determine the default shell for the oracle user, enter the following command:

```
$ echo $SHELL
```

Open the oracle user's shell startup file in any text editor:

Bash shell (bash) on Red Hat:

```
$ vi .bash_profile
```

Bourne shell (sh), Bash shell on SUSE, or Korn shell (ksh):

```
$ vi .profile
C shell (csh or tcsh):
% vi .login
```

Enter or edit the following line in the shell startup file, specifying a value of 022 for the default file mode creation mask:

```
umask 022
```

Save the file and exit from the editor.

To run the shell startup script, enter the following command:

Bash shell on Red Hat:

```
$ . ./bash_profile
```

Bourne shell, Bash shell on SUSE, or Korn shell:

```
$ . ./profile
```

C shell:

```
% source ./login
```

If you determined that the /tmp directory had insufficient free disk space when checking the hardware requirements, enter the following commands to set the TEMP and TMPDIR environment variables. Specify a directory on a file system with sufficient free disk space.

Bourne, Bash, or Korn shell:

```
$ TEMP=/directory
```

```
$ TMPDIR=/directory
```

```
$ export TEMP TMPDIR
```

C shell:

```
% setenv TEMP /directory
```

```
% setenv TMPDIR /directory
```

If you are not installing the software on the local system, enter the following command to direct X applications to display on the local system:

Bourne, Bash, or Korn shell:

```
$ DISPLAY=local_host:0.0 ; export DISPLAY
```

C shell:

```
% setenv DISPLAY local_host:0.0
```

In this example, local_host is the host name or IP address of the system that you want to use to display the Installer (your workstation or PC).

Enter commands similar to the following to set the ORACLE_BASE and ORACLE_SID environment variables:

Bourne, Bash, or Korn shell:

```
$ ORACLE_BASE=/u01/app/oracle
```

```
$ ORACLE_SID=sales
```

```
$ export ORACLE_BASE ORACLE_SID
```

C shell:

```
% setenv ORACLE_BASE /u01/app/oracle
```

```
% setenv ORACLE_SID sales
```

In these examples, /u01/app/oracle is the Oracle base directory that you created earlier and sales is the name that you want to call the database (typically no more than five characters).

Enter the following commands to ensure that the ORACLE_HOME and TNS_ADMIN environment variables are not set:

```
Bourne, Bash, or Korn shell:
$ unset ORACLE_HOME
$ unset TNS_ADMIN
```

```
C shell:
% unsetenv ORACLE_HOME
% unsetenv TNS_ADMIN
```

To verify that the environment has been set correctly, enter the following commands:

```
$ umask
$ env | more
```

Verify that the umask command displays a value of 0022, 022, or 22 and that the environment variables you set in this section have the correct values.

Install Oracle Database 10g

After configuring the oracle user's environment, start the Installer and install the Oracle software, as follows:

Note:
The following examples show paths to the runInstaller script on a CD-ROM. If you are installing the software from DVD-ROM, use a command similar to the following:
\$ /mount_point/db/runInstaller

To start the Installer, enter the following commands:

```
Red Hat:
$ cd /tmp
$ /mnt/cdrom/runInstaller
```

```
SUSE:
$ cd /tmp
$ /media/cdrom/runInstaller
```

If the Installer does not appear, see the Oracle Database Installation Guide for UNIX Systems for information about how to troubleshoot X display problems.

Use the following guidelines to complete the installation:

The following table describes the recommended action for each Installer screen.

Note:
If you have completed the tasks listed previously, you can complete the installation by choosing the default values on most screens.

If you need more assistance, or if you want to choose an option that is not a default, click Help for additional information.

If you encounter errors while installing or linking the software, see the Oracle Database Installation Guide for UNIX Systems for information about troubleshooting.

Screen	Recommended Action
Welcome to the Oracle Database 10g Installation	Specify the following information, then click Next. Oracle Home Location Verify that the path shown is similar to the following: oracle_base/product/10.1.0/db_1 Installation Type Select Enterprise Edition or Standard Edition. UNIX DBA Group Select the name of the OSDBA group that you created earlier, for example dba. Global Database Name Specify a name for the database, followed by the domain name of the system:

Screen	Recommended Action
	<p>sales.your_domain.com</p> <p>Database Password/Confirm Password</p> <p>Specify and confirm the password that you want to use for the following administrative database accounts:</p> <p>SYS, SYSTEM, SYSMAN, and DBSNMP</p>
Specify Inventory Directory and Credentials	<p>Note: This screen appears only during the first installation of Oracle products on a system. Specify the following information, then click Next.</p> <p>Enter the full path of the inventory directory:</p> <p>Verify that the path is similar to the following, where oracle_base is the value you specified for the ORACLE_BASE environment variable:</p> <p>oracle_base/oraInventory</p> <p>Specify operating system group name:</p> <p>Verify that the group specified is the Oracle Inventory group that you created earlier:</p> <p>oinstall</p>
Run oraInstRoot.sh	<p>If prompted, run the following script in a separate terminal window as the root user:</p> <p>oracle_base/oraInventory/oraInstRoot.sh</p>
Summary	Review the information displayed, then click Install.
Install	The Install screen displays status information while the product is being installed.
Configuration Assistants	<p>The Configuration Assistants screen displays status information for the configuration assistants that configure the software and create a database.</p> <p>After the Database Configuration Assistant finishes, click OK to continue.</p>
Setup Privileges	<p>When prompted, run the following script in a separate terminal window as the root user:</p> <p>oracle_home/root.sh</p> <p>In this example, oracle_home is the directory where you installed the software. The correct path is displayed on the screen.</p> <p>Press Return to accept the default values for each prompt displayed by the script. When the script finishes, click OK.</p>
End of Installation	<p>The configuration assistants configure several Web-based applications, including Oracle Enterprise Manager Database Control. This screen displays the URLs configured for these applications. Make a note of the URLs used.</p> <p>The port numbers used in these URLs are also recorded in the following file:</p> <p>oracle_home/install/portlist.ini</p> <p>To exit from the Installer, click Exit, then click Yes.</p>

Install Products from the Oracle Database 10g Companion CD

The Oracle Database 10g Companion CD contains products that improve the performance of or complement Oracle Database 10g. For most installations, Oracle recommends that you install Oracle Database 10g Products from the Companion CD.

Note:

If you intend to use Oracle JVM or Oracle interMedia, you must install Oracle Database 10g Products from the Companion CD. This installation optimizes the performance of those products on your system.

Products Included on the Companion CD

The Companion CD includes two sets of products:

Oracle Database 10g Products

Includes Oracle Database Examples, natively compiled Java libraries for Oracle JVM and Oracle interMedia, Oracle Text supplied knowledge bases, and Legato Single Server Version (LSSV)

Note:

You must install these products into the same Oracle home directory as Oracle Database 10g Release 1 (10.1.0).

Oracle Database 10g Companion Products

Includes Oracle HTTP Server and Oracle HTML DB

Note:

You must install Oracle HTTP Server into its own Oracle home directory. You must install Oracle HTML DB either with Oracle HTTP Server, or into an Oracle home directory that contains Oracle HTTP Server.

The following subsection describes how to install Oracle Database 10g Products. For more information about the products on the Companion CD, and for more detailed information about installing them, see the Oracle Database Companion CD Installation Guide which is located on the Companion CD.

Installing Oracle Database 10g Products

To install Oracle Database 10g Products, follow these steps:

As the root user, mount the Oracle Database 10g Companion CD CD-ROM or the Oracle Database 10g DVD-ROM.

For more information about mounting discs, see [Section 9, "Mount the Product Disc"](#).

If necessary, log in as the Oracle software owner user that you used to install Oracle Database 10g (typically oracle).

Enter a command similar to the following to start the Installer:

CD-ROM installation:

```
$ /mount_point/runInstaller
```

DVD-ROM installation:

```
$ /mount_point/companion/runInstaller
```

The following table describes the recommended action for each Installer screen:

Screen	Recommended Action
Welcome	Click Next.
Specify File Locations	In the Destination section, select the Name or Path value that specifies the Oracle home directory where you installed Oracle Database 10g, then click Next. The default Oracle home path is similar to the following: oracle_base/product/10.1.0/db_1
Select a Product to Install	Select Oracle Database 10g Products, then click Next.
Summary	Review the information displayed, then click Install.
Install	The Install screen displays status information while the product is being installed.
Setup Privileges	When prompted, run the following script in a separate terminal window as the root user: oracle_home/root.sh In this example, oracle_home is the directory where you installed the software. The correct path is displayed on the screen. Note: Unless you want to install Legato Single Server Version, enter 3 to quit the installation of LSSV. When the script finishes, click OK.
End of Installation	To exit from the Installer, click Exit, then click Yes.

2. What are the scripts to run to successful of software installation? Explain them?

a. Orainstroot.sh

b. root.sh

Note: Both the script should be run as root user

oraInstRoot.sh:

It is located in \$ORACLE_BASE/oraInventory

Usage:

- It creates the inventory pointer file (/etc/oraInst.loc), The file shows the inventory location and group it is linked to.
- Changing groupname of the oraInventory directory to oinstall group

root.sh:

It is located in \$ORACLE_HOME directory

Usage:

root.sh script performs many things, namely

- It changes or correctly sets the environment variables
- copying of few files into /usr/local/bin , the files are dbhome, oraenv, coraenv etc.
- creation of /etc/oratab file or adding database home and SID's entry into /etc/oratab file.

3. What are post installation Tasks?

[Required Post-installation Tasks](#)

[Recommended Post-installation Tasks](#)

[Required Product-Specific Post-installation Tasks](#)

[Installing Oracle Database 10g Products from the Companion CD](#)

Required Post-installation Tasks

You must perform the tasks described in the following sections after completing an installation:

[Downloading and Installing Patches](#)

[Running Oracle Enterprise Manager Java Console](#)

[Connecting with Instant Client](#)

[Configuring Oracle Products](#)

Downloading and Installing Patches

Check the OracleMetalink Web site for required patches for your installation. To download required patches:

Use a Web browser to view the OracleMetalink Web site:

<http://metalink.oracle.com>

Log in to OracleMetalink.

Note:

If you are not an OracleMetalink registered user, click Register for MetaLink! and follow the registration instructions.

On the main OracleMetalink page, click Patches.

Select Simple Search.

Specify the following information, then click Go:

In the Search By field, choose Product or Family, then specify RDBMS Server.

In the Release field, specify the current release number.

In the Patch Type field, specify Patchset/Minipack.

In the Platform or Language field, select your platform.

Running Oracle Enterprise Manager Java Console

In addition to using Oracle Enterprise Manager Database Control or Grid Control to manage an Oracle Database 10g database, you can also use the Oracle Enterprise Manager Java Console to manage databases from this release or previous releases. The Java Console is installed by the Administrator installation type.

Note:Oracle recommends that you use Grid Control or Database Control in preference to the Java Console when possible.

To start the Java Console, follow these steps:

Set the ORACLE_HOME environment variable to specify the Oracle home directory where you installed Oracle Client.

Set the shared library path environment variable of the system to include the following directories:

Platform	Environment Variable	Required Setting
Linux x86-64	LD_LIBRARY_PATH	\$ORACLE_HOME/lib32:\$ORACLE_HOME/lib:\$LD_LIBRARY_PATH

Enter the following command to start the Java Console:

\$ \$ORACLE_HOME/bin/oemapp

Connecting with Instant Client

If you installed the Instant Client installation type, you can configure users' environments to enable dynamically linked client applications to connect to a database as follows:

Set the appropriate shared library path environment variable for your platform to specify the directory that contains the Instant Client libraries. For the Instant Client installation type, this directory is the Oracle home directory that you specified during the installation, for example:

/u01/app/oracle/product/10.1.0/client_1

The following table shows the appropriate environment variable for the platform:

Platform	Environment Variable
Linux x86-64	LD_LIBRARY_PATH

Use one of the following methods to specify database connection information for the client application:

Specify a SQL connect URL string using the following format:

//host:port/service_name

Set the TNS_ADMIN environment variable to specify the location of the tnsnames.ora file and specify a service name from that file.

Set the TNS_ADMIN environment variable and set the TWO_TASK environment variable to specify a service name from the tnsnames.ora file.

Note:

You do not have to specify the ORACLE_HOME environment variable.

Configuring Oracle Products

Many Oracle products and options must be configured before you use them for the first time. Before using individual Oracle Database products or options, see the appropriate manual in the product documentation library, available on the Oracle Documentation Library CD-ROM, the DVD-ROM, or on the OTN Web site.

Recommended Post-installation Tasks

Oracle recommends that you perform the tasks described in the following section after completing an installation:

[Backing Up the root.sh Script](#)

[Configuring New or Upgraded Databases](#)

[Setting Up User Accounts](#)

[Generating the Client Static Library](#)

Backing Up the root.sh Script

Oracle recommends that you back up the root.sh script after you complete an installation. If you install other products in the same Oracle home directory, then the Oracle Universal Installer updates the contents of the existing root.sh script during the installation. If you require information contained in the original root.sh script, then you can recover it from the backed up root.sh file.

Configuring New or Upgraded Databases

Oracle recommends that you run the utlrp.sql script after creating or upgrading a database. This script recompiles all PL/SQL modules that might be in an invalid state, including packages, procedures, and types. This is an optional step but Oracle recommends that you do it during installation and not at a later date.

To run the utlrp.sql script, follow these steps:

Switch user to oracle.

Use the oraenv or coraenv script to set the environment for the database where you want to run the utlrp.sql script:

For the Bourne, Bash or Korn shell:

```
$ . /usr/local/bin/oraenv
```

For the C shell:

```
% source /usr/local/bin/coraenv
```

When prompted, specify the SID for the database.

Start SQL*Plus, as follows:

```
$ sqlplus "/ AS SYSDBA"
```

If necessary, start the database:

```
SQL> STARTUP
```

Run the utlrp.sql script:

```
SQL> @?/rdbms/admin/utlrp.sql
```

Setting Up User Accounts

For information about setting up additional user accounts, see the Oracle Database Administrator's Reference for UNIX Systems.

Generating the Client Static Library

The client static library (libclntst10.a) is not generated during installation. If you want to link your applications to the client static library, you must first generate it as follows:

Switch user to oracle.

Set the ORACLE_HOME environment variable to specify the Oracle home directory used by the Oracle Database installation. For example:

Bourne shell (sh), Bash shell (bash), or Korn shell (ksh):

```
$ ORACLE_HOME=/u01/app/oracle/product/10.1.0/db_1
```

```
$ export ORACLE_HOME
```

C shell (csh or tcsh):

```
% setenv ORACLE_HOME /u01/app/oracle/product/10.1.0/db_1
```

Enter the following command:

```
$ $ORACLE_HOME/bin/genclntst
```

Required Product-Specific Post-installation Tasks

The following sections describe platform-specific post-installation tasks that you must perform if you installed and intend to use the products mentioned:

[Configuring Oracle Net Services](#)

[Configuring Oracle Label Security](#)

[Installing Natively Compiled Java Libraries for Oracle JVM and Oracle interMedia](#)

[Installing Oracle Text Supplied Knowledge Bases](#)

[Configuring Oracle Messaging Gateway](#)

[Configuring Oracle Precompilers](#)

Note:

You need only perform post-installation tasks for products that you intend to use.

Configuring Oracle Net Services

If you have a previous release of Oracle software installed on this system, you might want to copy information from the Oracle Net tnsnames.ora and listener.ora configuration files from the previous release to the corresponding files for the new release.

Note:

The default location for the tnsnames.ora and listener.ora files is the \$ORACLE_HOME/network/admin/ directory. However, you can also use a central location for these files, for example /var/opt/oracle/etc.

Modifying the listener.ora File

If you are upgrading from a previous release of Oracle Database, Oracle recommends that you use the current release of Oracle Net listener instead of the listener from the previous release.

To use the listener from the current release, you may need to copy static service information from the listener.ora file from the previous release to the version of that file used by the new release.

For any database instances earlier than release 8.0.3, add static service information to the listener.ora file. Oracle Database releases later than release 8.0.3 do not require static service information.

Modifying the tnsnames.ora File

Unless you are using a central tnsnames.ora file, copy Oracle Net service names and connect descriptors from the previous release tnsnames.ora file to the version of that file used by the new release.

If necessary, you can also add connection information for additional database instances to the new file.

Configuring Oracle Label Security

If you installed Oracle Label Security, you must configure it in a database before you use it. You can configure Oracle Label Security in two ways; with Oracle Internet Directory integration and without Oracle Internet Directory integration. If you configure Oracle Label Security without Oracle Internet Directory integration, you cannot configure it to use Oracle Internet Directory at a later stage.

Note:

To configure Oracle Label Security with Oracle Internet Directory integration, Oracle Internet Directory must be installed in your environment and the Oracle database must be registered in the directory.

See Also:

For more information about Oracle Label Security enabled with Oracle Internet Directory, see the Oracle Label Security Administrator's Guide.

Installing Natively Compiled Java Libraries for Oracle JVM and Oracle interMedia

If you plan to use Oracle JVM or Oracle interMedia, Oracle strongly recommends that you install the natively compiled Java libraries (NCOMPs) used by those products from the Oracle Database 10g Companion CD. These libraries are required to improve the performance of the products on your platform.

For information about how to install products from the Companion CD, see the ["Installing Oracle Database 10g Products from the Companion CD"](#) section.

Installing Oracle Text Supplied Knowledge Bases

An Oracle Text knowledge base is a hierarchical tree of concepts used for theme indexing, ABOUT queries, and deriving themes for document services. If you plan to use any of these Oracle Text features, you can install two supplied knowledge bases (English and French) from the Oracle Database 10g Companion CD.

Note:

You can extend the supplied knowledge bases depending on your requirements. Alternatively, you can create your own knowledge bases, possibly in languages other than English and French. For more information about creating and extending knowledge bases, see the Oracle Text Reference.

For information about how to install products from the Companion CD, see the ["Installing Oracle Database 10g Products from the Companion CD"](#) section.

Configuring Oracle Messaging Gateway

To configure Oracle Messaging Gateway, see the section about Messaging Gateway in the Oracle Streams Advanced Queuing User's Guide and Reference manual. When following the instructions listed in that manual, refer to this section for additional platform-specific instructions about configuring the listener.ora, tnsnames.ora, and mgw.ora files.

Modifying the listener.ora File for External Procedures

To modify the \$ORACLE_HOME/network/admin/listener.ora file for external procedures:

Back up the listener.ora file.

Ensure that the default IPC protocol address for external procedures is set as follows:

(ADDRESS = (PROTOCOL=IPC)(KEY=EXTPROC))

Add static service information for a service called mgwextproc by adding lines similar to the following to the SID_LIST parameter for the listener in the listener.ora file:

```
(SID_DESC =  
  (SID_NAME = mgwextproc)  
  (ENVS = platform-specific_env_vars)  
  (ORACLE_HOME = oracle_home)  
  (PROGRAM = extproc_agent)  
)
```

In this example:

The ENVS parameter defines the shared library path environment variable and any other required environment variables.

The following table lists the environment variables and required values that you must specify for each platform. In the shared library path environment variable, you must also add any additional library paths required for non-Oracle messaging systems, for example WebSphere MQ or TIBCO Rendezvous.

Platform	ENVS Parameter Setting
Linux x86-64	EXTPROC_DLLS=/oracle_home/lib32/libmgwagent.so, LD_LIBRARY_PATH=/oracle_home/jdk/jre/lib/i386:/oracle_home/jdk/jre/lib/i386/server server:/oracle_home/lib32

oracle_home is the path of the Oracle home directory.

extproc_agent is the external procedure agent executable file. The following table lists the correct executable file for each platform:

Platform	Agent Executable File
Linux x86-64	extproc32

The following examples show sample listener.ora files on Linux x86-64:

Note:
In the following examples, the PLSExtProc service is the default service for PL/SQL external procedures.

```
SID_LIST_LISTENER =
(SID_LIST =
(SID_DESC =
(SID_NAME = PLSExtProc)
(ORACLE_HOME = /u01/app/oracle/product/10.1.0/db_1)
(PROGRAM = extproc)
)
(SID_DESC =
(SID_NAME = mgwextproc)
(ENVS = EXTPROC_DLLS=/u01/app/oracle/product/10.1.0/db_1/lib32/
libmgwagent.sl,LD_PRELOAD=/u01/app/oracle/product/10.1.0/db_1/jdk/jre/
lib/PA-
RISC/server/libjvm.sl,SHLIB_PATH=/u01/app/oracle/product/10.1.0/db_1/jdk/jre/lib/PA_RISC:/u01/app/oracle/prod
uct/10.1.0/db_1/jdk/jre/lib/PA_RISC/server:/u01/app/oracle/product/10.1.0/db_1/lib32)
(ORACLE_HOME = /u01/app/oracle/product/10.1.0/db_1)
(PROGRAM = extproc32)
)
)
)
```

Modifying the tnsnames.ora File for External Procedures

To modify the \$ORACLE_HOME/network/admin/tnsnames.ora file for external procedures:

Back up the tnsnames.ora file.

In the tnsnames.ora file, add a connect descriptor with the net service name MGW_AGENT, as follows:

```
MGW_AGENT =
(DESCRIPTION=
(ADDRESS_LIST= (ADDRESS= (PROTOCOL=IPC)(KEY=EXTPROC)))
(CONNECT_DATA= (SID=mgwextproc) (PRESENTATION=RO)))
```

In this example:

The value specified for the KEY parameter must match the value specified for that parameter in the IPC protocol address in the listener.ora file.

The value of the SID parameter must match the service name in the listener.ora file that you specified for the Oracle Messaging Gateway external procedure agent in the previous section (mgwextproc).

Setting up the mgw.ora Initialization File

To modify the \$ORACLE_HOME/mgw/admin/mgw.ora file for external procedures, set the CLASSPATH environment variable to include the classes in the following table and any additional classes required for Oracle Messaging Gateway to access non-Oracle messaging systems, for example WebSphere MQ or TIBCO Rendezvous classes:

Classes	Path
Oracle Messaging Gateway	ORACLE_HOME/mgw/classes/mgw.jar
JRE internationalization	ORACLE_HOME/JRE/lib/i18n.jar
JRE runtime	ORACLE_HOME/JRE/lib/rt.jar
Oracle JDBC	ORACLE_HOME/jdbc/lib/ojdbc14.jar
Oracle internationalization	ORACLE_HOME/jdbc/lib/orai18n.jar
SQLJ	ORACLE_HOME/sqlj/lib/translator.zip ORACLE_HOME/sqlj/lib/runtime12.zip
JMS Interface	ORACLE_HOME/rdbms/jlib/jmscommon.jar
Oracle JMS implementation	ORACLE_HOME/rdbms/jlib/aqapi3.jar
Java Transaction API	ORACLE_HOME/jlib/jta.jar

Note:

All the lines in the mgw.ora file should be less than 1024 characters.

Configuring Oracle Precompilers

The following section describes post-installation tasks for Oracle precompilers.

[Configuring Pro*C/C++](#)

Note:

All precompiler configuration files are located in the \$ORACLE_HOME/precomp/admin directory.

Configuring Pro*C/C++

Verify that the PATH environment variable setting includes the directory that contains the C compiler executable.

[Table 4-1](#) shows the default directories and the appropriate commands to verify the path setting, depending on your platform and compiler.

Table 4-1 C/C++ Compiler Directory

Platform	Path	Command
Linux x86-64	/usr/bin	\$ which gcc

Installing Oracle Database 10g Products from the Companion CD

The Oracle Database 10g Companion CD contains additional products that you can install. Whether you need to install these products depends on which Oracle Database products or features you plan to use. If you plan to use the following products or features, Oracle strongly recommends that you complete the Oracle Database 10g Products installation from the Companion CD:

Oracle JVM

Oracle interMedia

Oracle Text

To install Oracle Database 10g Products from the Companion CD, follow these steps:

Note:

For more detailed installation information, see the Oracle Database Companion CD Installation Guide, which is available on the Companion CD.

Insert the Oracle Database 10g Companion CD or the Oracle Database 10g DVD-ROM into the disc drive.

If necessary, log into the system as the user who installed Oracle Database (typically the oracle user).

To start the Installer, enter the following commands where directory_path is the CD-ROM mount point directory or the path of the companion directory on the DVD-ROM:

```
$ cd /tmp
```

\$ /directory_path/runInstaller

If the Installer does not appear, see the ["X Windows Display Errors"](#) section for information about troubleshooting.

Use the following guidelines to complete the installation:

On the Specify File Locations screen, select the Oracle home name and path for the Oracle Database 10g installation where you want to install the products.

On the Select a Product to Install screen, select Oracle Database 10g Products.

Unless you want to install Legato Single Server Version, eEnter 3 at the prompt displayed by the root.sh script.

4. Explain the steps for manual/silent installation of oracle software?

a. Create the oraInst.loc file.

b. Prepare a response file.

c. Run Oracle Universal Installer in silent or response file mode.

d. If you completed a software-only installation, then run Net Configuration Assistant and Database Configuration Assistant in silent or response file mode if required.

Explanation-1:

1) Customize value in response file

important note: copy from Oracle sample file: <oracle_installation_dir>/database/response/db_install.rsp

oracle.install.responseFileVersion=/oracle/install/rspfmt_dbinstall_response_schema_v11_2_0

oracle.install.option=INSTALL_DB_SWONLY

UNIX_GROUP_NAME=oinstall

INVENTORY_LOCATION=/oracle/OraInventory

ORACLE_HOME=/oracle/product/11.2.0/template

ORACLE_BASE=/oracle

oracle.install.db.InstallEdition=EE

oracle.install.db.DBA_GROUP=dba

oracle.install.db.OPER_GROUP=dba

DECLINE_SECURITY_UPDATES=true

2) Silent installation

`./runInstaller -silent -noconfig -responseFile /u01/download/db11ginstall.rsp`

Starting Oracle Universal Installer...

Checking Temp space: must be greater than 120 MB. Actual 22753 MB Passed

Checking swap space: must be greater than 150 MB. Actual 3817 MB Passed

Preparing to launch Oracle Universal Installer from /tmp/OraInstall2011-06-17_09-02-07PM.

Please wait..

\$ You can find the log of this install session at:

/oracle/OraInventory/logs/installActions2011-06-17_09-02-07PM.log

The following configuration scripts need to be executed as the "root" user.

#!/bin/sh

#Root scripts to run

/oracle/product/11.2.0/template/root.sh

To execute the configuration scripts:

1. Open a terminal window

2. Log in as "root"

3. Run the scripts

4. Return to this [window](#) and hit "Enter" key to continue

Successfully Setup Software

*/ * tricks and tips */*

Trick 1 # If oracle home is not empty and we didn't use *runInstaller *** -force* option, it will die with following messages

CAUSE: The chosen installation conflicted with software already installed in the given Oracle home.

ACTION: Install into a different Oracle home.

Trick 2 # If system are not satisfied with Oracle installation requirement, it will show following messages.

CAUSE: Some of the optional prerequisites are not met. See logs for details. /oracle/OraInventory/logs/...log

ACTION: Identify the list of failed prerequisite checks from the log: /oracle/OraInventory/logs/....log. Then either from the log file or from installation manual find the appropriate configuration to meet the prerequisites and fix it

manually.

[WARNING] [INS-13014] Target environment do not meet some optional requirements.

Trick 3 # If DECLINE_SECURITY_UPDATES is not set to FALSE, Oracle will try to set up OCM (Oracle configuration management with metalink credential), it will die with following messages.

[SEVERE] – Email Address Not Specified

Trick 4 # If DBA and OS group is not specified properly, it may die with following messages

CAUSE: User is not a member of one or more of the chosen OS groups.

ACTION: Please choose OS groups of which user are a member.

Explanation-2:

a. Download Oracle 11g from the website in the link below

b. Unzip both zip files in the same directory

c. Create users and groups

Create the groups oinstall and dba as follows

```
# groupadd oinstall
```

```
# groupadd dba
```

Create user oracle and set the dba group as the user's primary group and oinstall as the secondary

```
# useradd oracle
```

```
# usermod -g dba -G oinstall oracle
```

d. Create the file systems paths

Regarding this step this is just a guide according Oracle OFA (Optimal Flexible Architecture)

Create the directories as follows

```
ORACLE_BASE: /u01/app/oracle/
```

```
ORACLE_HOME: /u01/app/oracle/product/11.2.0/db1
```

e. Prerequisites for installation

You should read carefully and follow the Oracle guide for all the prerequisites otherwise the installation might fail.

Pay attention specifically in the following sections

- Checking the Software Requirements
- Checking Resource Limits for the Oracle Software Installation Users
- Configuring Kernel Parameters for Linux

f. Create the response file

Within the "database/response" directory, edit the response file db_install.rsp for my installation I have used the following parameters.

Make sure the directories mentioned in the file are created before the installation.

I have highlighted the parameters I actually set, the others I used the default or left blank.

```
oracle.install.option=INSTALL_DB_SWONLY
```

```
ORACLE_HOSTNAME=
```

```
UNIX_GROUP_NAME=oinstall
```

```
INVENTORY_LOCATION=/etc/oraInv
```

```
SELECTED_LANGUAGES=en
```

```
oracle.install.db.InstallEdition=EE
```

```
oracle.install.db.isCustomInstall=false
```

```
oracle.install.db.customComponents=oracle.server:11.2.0.1.0,oracle.sysman.ccr:10.2.7.0.0,oracle.xdk:11.2.0.1.0,oracle.rdbms.oci:11.2.0.1.0,oracle.network:11.2.0.1.0,oracle.network.listener:11.2.0.1
```

```
0,oracle.rdbms:11.2.0.1.0,oracle.options:11.2.0.1.0,oracle.rdbms.partitioning:11.2.0.1.0,oracle.oraolap:11.2.0.1.0,oracle.rdbms.dm:11.2.0.1.0,oracle.rdbms.dv:11.2.0.1.0,oracle.rdbms.lbac:11.2.0.1.0
```

```
,oracle.rdbms.rat:11.2.0.1.0
```

```
oracle.install.db.DBA_GROUP=dba
```

```
oracle.install.db.OPER_GROUP=dba
```

```
oracle.install.db.CLUSTER_NODES=
```

```
oracle.install.db.config.starterdb.type=GENERAL_PURPOSE
```

```
oracle.install.db.config.starterdb.globalDBName=db1
```

```
oracle.install.db.config.starterdb.SID=db1
```

```
oracle.install.db.config.starterdb.characterSet=AL32UTF8
```

```
oracle.install.db.config.starterdb.memoryOption=true
```

```
oracle.install.db.config.starterdb.memoryLimit=
```

```

oracle.install.db.config.starterdb.installExampleSchemas=false
oracle.install.db.config.starterdb.enableSecuritySettings=true
oracle.install.db.config.starterdb.password.ALL=manager
oracle.install.db.config.starterdb.password.SYS=
oracle.install.db.config.starterdb.password.SYSTEM=
oracle.install.db.config.starterdb.password.SYSMAN=
oracle.install.db.config.starterdb.password.DBSNMP=
oracle.install.db.config.starterdb.control=DB_CONTROL
oracle.install.db.config.starterdb.gridcontrol.gridControlServiceURL=
oracle.install.db.config.starterdb.dbcontrol.enableEmailNotification=false
oracle.install.db.config.starterdb.dbcontrol.emailAddress=email@something.com
oracle.install.db.config.starterdb.dbcontrol.SMTPServer=email@something.com
oracle.install.db.config.starterdb.automatedBackup.enable=false
oracle.install.db.config.starterdb.automatedBackup.osuid=
oracle.install.db.config.starterdb.automatedBackup.ospwd=
oracle.install.db.config.starterdb.storageType=
oracle.install.db.config.starterdb.fileSystemStorage.dataLocation=
oracle.install.db.config.starterdb.fileSystemStorage.recoveryLocation=
oracle.install.db.config.asm.diskGroup=
oracle.install.db.config.asm.ASMNSMPPassword=
MYORACLESUPPORT_USERNAME=
MYORACLESUPPORT_PASSWORD=
SECURITY_UPDATES_VIA_MYORACLESUPPORT=false
DECLINE_SECURITY_UPDATES=true
PROXY_HOST=
PROXY_PORT=
PROXY_USER=
PROXY_PWD=

```

Ensure the parameters SECURITY_UPDATES_VIA_MYORACLESUPPORT and DECLINE_SECURITY are correctly otherwise you may get the “[SEVERE] – Email Address Not specified” error.

g. Run the root.sh file

After the installation runs the root.sh file as root (you will be prompted by the installation).

Your installation is done. To create a database, prepare the init.ora file (located in the ORACLE_HOME/dbs and set the ORACLE_SID environment variable.

Then run sqlplus as sysdba and start the database in no mount mode.

5. Explain the steps for manual creation of oracle database?

- Create directories for datafiles, redo log files and archive log files
- Create init.ora file with minimum parameters
- Set the SID for the current session
- Connect to SQLPLUS
- Create spfile from existing pfile
- Startup the instance in nomount
- Run the create database script
- Create the catalog by running catalog.sql and catproc.sql
- Run the Script pupbld.sql with SYSTEM user
- Create the user tablespace, local, auto allocate
- Create a user and assign the tablespace
- Add the database entries into the /etc/oratab file
- Setup the listener

Explanation:

Create the directories

First, create the directories you need for the datafiles, eg:

Don't need to create the admin directories in 11g since introduction

of diag_dest

```

mkdir -p /mnt/raid/dborafiles/11gr2/admin
mkdir -p /mnt/raid/dborafiles/11gr2/admin/bdump
mkdir -p /mnt/raid/dborafiles/11gr2/admin/cdump
mkdir -p /mnt/raid/dborafiles/11gr2/admin/udump
mkdir -p /mnt/raid/dborafiles/11gr2/datafiles
mkdir -p /mnt/raid/dborafiles/11gr2/redo

```

For a production setup, each of these areas is probably a separate mount point on different disks etc.

Create a minimal init.ora

This file should go into \$ORACLE_HOME/dbs and be called initSID.ora:

```

control_files = (/mnt/raid/dborafiles/ora11gr2/datafiles/control01.ora,
                /mnt/raid/dborafiles/ora11gr2/datafiles/control02.ora,
                /mnt/raid/dborafiles/ora11gr2/datafiles/control03.ora)

```

```

undo_management = auto

```

```

db_name = ora11gr2

```

```

db_block_size = 8192

```

```

# 11G (oracle will create subdir diag and all the required subdirs)

```

```

diagnostic_dest = /mnt/raid/dborafiles/ora11gr2

```

```

# Pre 11G specify these parameters

```

```

# background_dump_dest = /mnt/raid/dborafiles/ora11gr2/admin/bdump

```

```

# core_dump_dest = /mnt/raid/dborafiles/ora11gr2/admin/cdump

```

```

# user_dump_dest = /mnt/raid/dborafiles/ora11gr2/admin/udump

```

Set the SID for your session

```

export ORACLE_SID=ora11gr2

```

Connect to SQLPLUS

```

$ sqlplus /nolog

```

```

SQL11g> connect / as sysdba

```

Create the SPFILE

```

$ create SPFILE from PFILE='/dboracle/product/11.2.0/dbhome_1/dbs/init11gr2.ora'

```

Startup the instance

```

SQL11g> startup nomount

```

Create the database

```

create database ora11gr2

```

```

  logfile group 1 ('/mnt/raid/dborafiles/ora11gr2/redo/redo1.log') size 10M,

```

```

    group 2 ('/mnt/raid/dborafiles/ora11gr2/redo/redo2.log') size 10M,

```

```

    group 3 ('/mnt/raid/dborafiles/ora11gr2/redo/redo3.log') size 10M

```

```

  character set WE8ISO8859P1

```

```

  national character set utf8

```

```

  datafile '/mnt/raid/dborafiles/ora11gr2/datafiles/system.dbf'

```

```

    size 50M

```

```

    autoextend on

```

```

    next 10M

```

```

    extent management local

```

```

  sysaux datafile '/mnt/raid/dborafiles/ora11gr2/datafiles/sysaux.dbf'

```

```

    size 10M

```

```

    autoextend on

```

```

    next 10M

```

```

  undo tablespace undo

```

```

    datafile '/mnt/raid/dborafiles/ora11gr2/datafiles/undo.dbf'

```

```

    size 10M

```

```

    autoextend on

```

```

  default temporary tablespace temp

```

```

    tempfile '/mnt/raid/dborafiles/ora11gr2/datafiles/temp.dbf'

```

```

    size 10M

```

```

    autoextend on

```

(TODO - unsure about setting max files sizes on these files)

Create the catalogue etc:

```
SQL11G> @$ORACLE_HOME/rdbms/admin/catalog.sql
```

```
SQL11G> @$ORACLE_HOME/rdbms/admin/catproc.sql
```

As SYSTEM (not SYS) run the following:

```
SQL11G> @$ORACLE_HOME/sqlplus/admin/pupbld.sql
```

(not doing this doesn't cause any harm, but a warning is displayed when logging into SQLPLUS if it is not run)

The database is now basically ready to use, but there no users and no users tablespace. Note it is also NOT in archive log mode, so is certainly not production ready, but may be good enough for a non-backed up dev instance.

Create the users tablespace, local, auto allocate:

```
SQL>CREATE TABLESPACE users DATAFILE '/mnt/raid/dborafiles/ora11gr2/datafiles/users_01.dbf'
```

```
SIZE 50M
```

```
autoextend on
```

```
maxsize 2048M
```

```
EXTENT MANAGEMENT LOCAL AUTOALLOCATE;
```

Create a user:

```
SQL11G> create user sodonnel
```

```
identified by sodonnel
```

```
default tablespace users
```

```
temporary tablespace temp;
```

```
SQL11G> alter user sodonnel quota unlimited on users;
```

```
SQL11G> grant connect, create procedure, create table, alter session to sodonnel;
```

Ensure the database comes up at startup time:

Add a line to /etc/oratab to tell Oracle about the instance. This is used by the dbstart command, which will start all the database specified in this file:

```
ora11gr2:/dboracle/product/11.2.0/dbhome_1:Y
```

To start all instances use dbstart and to stop use dbshut.

TODO - control script to autostart databases when the machine boots.

Setup the listener

At this point, only people on the local machine can connect to the database, so the last step is to setup the listener.

All you need to do here is add a file called listener.ora in \$ORACLE_HOME/network/admin, and have it contain something like the following:

```
LISTENER =
```

```
(DESCRIPTION_LIST =
```

```
(DESCRIPTION =
```

```
(ADDRESS = (PROTOCOL = TCP)(HOST = hostname)(PORT = 1521))
```

```
(ADDRESS = (PROTOCOL = IPC)(KEY = EXTPROC))
```

```
)
```

```
)
```

Creating a tnsnames.ora file at this point would be a good idea too. It also goes into

\$ORACLE_HOME/network/admin:

```
ora11gr2 =
```

```
(DESCRIPTION =
```

```
(ADDRESS_LIST =
```

```
(ADDRESS = (PROTOCOL = TCP)(Host = localhost)(Port = 1521))
```

```
)
```

```
(CONNECT_DATA =
```

```
(SERVICE_NAME = ora11gr2)
```

```
)
```

```
)
```

6. What is use of catalog.sql, catproc.sql and pupbld.sql? Explain?

CATALOG.SQL: creates the views of the data dictionary tables, the dynamic performance views, and public synonyms for many of the views. Grants PUBLIC access to the synonyms. It is located in \$ORACLE_HOME/rdbms/admin/catalog.sql.

CATPROC.SQL: Runs all scripts required for or used with PL/SQL. It is located in \$ORACLE_HOME/rdbms/admin/catproc.sql.

PUPBLD.SQL: To create PRODUCT_USER_PROFILE table, a Database Administrator (DBA) must run the PUPBLD.SQL script located in the Oracle home directory

This script must be run as user SYSTEM. If you do not run, you may be getting USER PROFILE warnings when connecting via SQL*Plus to an Oracle database and what scripts to run to get rid of the warning and/or error messages due to the fact that the USER PROFILE environment has not been setup.

7. What is ORAINVENTORY and the default location?

Explanation-1:

The oraInventory is the location for the OUI (Oracle Universal Installer)'s bookkeeping. The inventory stores information about:

- * All Oracle software products installed in all ORACLE_HOMES on a machine
- * Other non-Oracle products, such as the Java Runtime Environment (JRE)

Explanation-2:

In the oraInventory directory, the Installer keeps track of what is installed, and information on how to deinstall products. Many of the libraries needed to run the Installer are also in this directory. The Oracle Installer uses the "oraInst.loc" file to determine the location of the oraInventory directory. If this file is blank or contains an invalid entry the Installer will give an error about unable to find oraInventory or may provide an incomplete listing of installed products.

Explanation-3:

What is oraInventory

What is oraInventory

The OraInventory is the location for the OUI's Book keeping

The inventory stores information about.

All the Oracle software products installed on all ORACLE_HOMES on a machine

Other non-oracle products such as Java Runtime env's (JRE)

Binary OraInventory

Before OUI 2.X (or 11.5.7 or earlier) the inventory was binary, the binary oraInventory maintains inventory in binary format

XML Inventory

Starting from OUI 2.X and 11.5.8 information in the inventory is stored in the Extensible Markup Language (XML) format.

The XML format allows easier diagnostic of the problem and faster loading of data.

XML inventory is divided into 2 components.

Global Inventory

Global Inventory holds information about Oracle Products on a Machine, The inventory contains the high level list of all oracle products installed on a machine such as ORACLE_HOMES or JRE.

It doesn't have any information about the details of patches applied on each ORACLE_HOMES.

There should be only one per machine. Its location is defined in the oraInst.loc in /etc (on Linux) or /var/opt/oracle (solaris).

Local Inventory

There is one Local inventory per ORACLE_HOME.

Inventory inside each Oracle Home is called as local Inventory or ORACLE_HOME Inventory. This Inventory holds information to that ORACLE_HOME only.

Can I have multiple Global Inventories on a machine?

Can you have multiple global Inventory and answer is YES you can have multiple global Inventory but if your upgrading or applying patch then change Inventory Pointer oraInst.loc to respective location.

If you are following single global Inventory and if you wish to uninstall any software then remove it from Global Inventory as well.

What to do if my Global Inventory is corrupted?

If your global Inventory is corrupted, you can recreate global Inventory on machine using Universal Installer and attach already installed oracle home by option

-attachHome

./runInstaller -silent -attachHome -invPtrLoc \$location_to_oraInst.loc ORACLE_HOME=Oracle_Home_Location

ORACLE_HOME_NAME=Oracle_Home_Name CLUSTER_NODES={}

Do I need to worry about oraInventory during oracle Apps 11i cloning ?

No, Rapid Clone will update both Global & Local Inventory with required information, you don't have to worry about Inventory during Oracle Apps 11i cloning.

How to Move oraInventory from one location to other?

Find the current location of the central inventory (Normally \$ORACLE_BASE/oraInventory):

Open the oraInst.loc file in /etc and check the value of inventory_loc

```
cat /etc/oraInst.loc
```

```
inventory_loc=/u01/app/oracle/oraInventory
```

```
inst_group=oinstall
```

Remark: The oraInst.loc file is simply a pointer to the location of the central inventory (oraInventory)

Copy the oraInventory directory to the destination directory

```
cp -Rp /u01/app/oracle/oraInventory /u02/app/oracle/oraInventory
```

Edit the oraInst.loc file to point to the new location

```
vi /etc/oraInst.loc
```

```
inventory_loc=/u02/app/oracle/oraInventory
```

```
inst_group=dba
```

8. What is ORAINST.LOC file? Explain?

In the oraInventory directory, the Installer keeps track of what is installed, and information on how to deinstall products. Many of the libraries needed to run the Installer are also in this directory. The Oracle Installer uses the "oraInst.loc" file to determine the location of the oraInventory directory. If this file is blank or contains an invalid entry the Installer will give an error about unable to find oraInventory or may provide an incomplete listing of installed products.

Usage:

If you are planning to apply one-off Patch or you are planning to install the latest Patchset and the oraInst.loc (Oracle Inventory Pointer) requires. If it is not located in /var/opt/oracle or /etc, you need to manually tell OPatch or the OUI (Oracle Universal Installer) where to find the oraInst.loc file?

To determine where oraInventory is created, opatch needs to read

/var/opt/oracle/oraInst.loc or /etc/oraInst.loc depending upon the Platform.

By default, opatch searches /var/opt/oracle/oraInst.loc or /etc/oraInst.loc

If oraInst.loc does not exist or was not created in /var/opt/oracle or /etc, you need to tell opatch where to find oraInventory.

Like \$ opatch apply -invPtrLoc <path>/oraInst.loc.

Don't delete oraInst.loc from /var/opt/oracle.

If you plan to install Oracle products using Oracle Universal Installer in silent or suppressed mode, you must manually create the oraInst.loc file if it does not already exist. This file specifies the location of the Oracle Inventory directory where Oracle Universal Installer creates the inventory of Oracle products installed on the system. If Oracle software has been installed previously on the system, the oraInst.loc file might already exist. If the file does exist, you do not need to create a file.

AIX: /etc and on HP-UX, Solaris, and Tru64 UNIX: /var/opt/oracle Location for oraInst.loc.

Still if you're very much decided to delete you can delete /oracle/<SID>/102_64/oraInst.loc. But let it be !!!!! a small file with some imp info?

Note: A pure oracle 10g installation creates 3 oraInst.loc.

1) \$ORACLE_HOME/oraInst.loc

2) /etc/oraInst.loc or /var/opt/oracle/oraInst.loc depending on the OS

3) on the inventory directory (typically /oracle/oraInventory/oraInst.loc)

9. Explain the manual un-installation of oracle software without using uninstall Wizard?

Note: In essence, uninstalling Oracle involves two major steps, removing the executable code from the server and downgrading the Oracle data dictionary. **Note:** An uninstall cannot be recovered, so make sure to take a full backup of all server-side components before uninstalling Oracle.

Manually uninstalling Oracle

For experienced DBA's, the manual uninstall is the preferred method of uninstalling Oracle

Step 1: Nuke all Oracle processes

The uninstall requires stopping all Oracle instances, listeners and daemon processes that are associated with the release:

```
$ORACLE_HOME/bin/shutdown abort
```

```
$ORACLE_HOME/bin/emctl stop dbconsole
```

```
$ORACLE_HOME/bin/lsnrctl stop
$ORACLE_HOME/bin/isqlplusctl stop
```

Step 2: Take a full backup of \$ORACLE_HOME

In UNIX/Linux you can make a backup copy in case you need anything later:

```
root> tar -xvf $ORACLE_HOME > /dev/rmt0/bkup
```

Step 3: Remove \$ORACLE_HOME and all subdirectories

You can use the DBCA uninstaller function if you are on 10g or beyond. In UNIX/Linux you can remove the software footprint as follows:

```
root> cd $ORACLE_HOME
root> rm -RF *
```

Step 4: Remove entries in shared parm files

The database may participate in server-side parameter files and you want to complete the uninstallation process by removing references to the database in these locations specified by the "dest" parameters in the init.ora file (e.g.

log_archive_dest_n):

tnsnames.ora

protocol.ora

listener.ora

oratab (/etc or /var/opt/oracle)

Assisted Oracle uninstall

For beginners and those who are unfamiliar with the components of Oracle, the Oracle DBCA has [Assisted uninstall procedures](#):

Step 1: Nuke all Oracle processes

The uninstall requires stopping all Oracle instances, listeners and daemon processes that are associated with the release:

```
$ORACLE_HOME/bin/shutdown abort
$ORACLE_HOME/bin/emctl stop dbconsole
$ORACLE_HOME/bin/lsnrctl stop
$ORACLE_HOME/bin/isqlplusctl stop
```

Step 2: Start the Database Creation Assistant (DBCA) and choose the remove database option

Step 3: Use the assisted deinstall wizard. Execute the Oracle Universal Installer (OUI)

```
$ORACLE_HOME/oui/bin/runInstaller
```

Step 4. In the OUI Welcome window, click the "Deinstall Products" button.

Step 5. In the OUI software Inventory screen, select the Oracle home and the products that you want to remove, cross your fingers and mash the "Remove" button.

Uninstall Oracle on Windows:

Its easy to uninstall Oracle on Windows :

Uninstall all Oracle components using the Oracle Universal Installer (OUI).

Run regedit.exe and delete the HKEY_LOCAL_MACHINE/SOFTWARE/ORACLE key. This contains registry entires for all Oracle products.

Delete any references to Oracle services left behind in the following part of the registry

(HKEY_LOCAL_MACHINE/SYSTEM/CurrentControlSet/Services/Ora*). It should be pretty obvious which ones relate to Oracle.

Reboot your Windows box.

Delete the "C:\Oracle" directory, or whatever directory is your ORACLE_BASE.

Delete the "C:\Program Files\Oracle" directory.

Empty the contents of your "C:\temp" directory.

Empty your recycle bin.

10. Explain the manual deletion of oracle database?

Identify all instances associated with the Oracle home:

To identify all instances associated with the Oracle home that you want to remove, issue the following command

```
1$ cat /etc/oratab |grep "/"
```

The output will be similar to the following:

```
1db10g1:/u01/app/oracle/product/10.2.0/db10g1:N
```

There is one instance associated with the /u01/app/oracle/product/10.2.0/db10g1 Oracle Home directory

Remove database:

To completely remove Oracle Database software, you must remove any installed databases. To remove an Oracle database:

a. Log in as oracle user:

```
1$ su - oracle
```

b. Set the environment for the database that you want to remove:

```
1$ . /usr/local/bin/oraenv      # for Bourne, Bash, or Korn shell:
```

```
2$ source /usr/local/bin/coraenv # for C shell
```

c. At the prompt, specify the SID for the database that you want to remove.

d. Start the Database Configuration Assistant by invoking dbca

e. Click Next on Welcome Screen, In Operations Window Select Delete a Database

Deleting the Oracle CSS Daemon Configuration:

Log in as root user and set environment variables

```
1# ORACLE_HOME=/u01/app/oracle/product/10.2.0/db10g1
```

```
2# export ORACLE_HOME
```

Enter the below command to delete the CSS daemon configuration from this Oracle home:

```
1# $ORACLE_HOME/bin/localconfig delete
```

Removing Oracle Software:

The following steps describe how to use Oracle Universal Installer to remove Oracle software from an Oracle home. Always use Oracle Universal Installer to remove Oracle software. Do not delete any Oracle home directories without first using Oracle Universal Installer to remove the software.

a. Login as oracle user:

```
1$ su - oracle
```

b. Set the environment for the database that you want to remove:

c. Stop any Oracle background processes running in this Oracle home:

- | | |
|---|--|
| 1. Grid Control for Database Management | \$ORACLE_HOME/bin/emctl stop agent |
| 2. Database Control for Database Management | \$ORACLE_HOME/bin/emctl stop dbconsole |
| 3. Oracle Net listener | \$ORACLE_HOME/bin/lsnrctl stop |
| 4. iSQL*Plus | \$ORACLE_HOME/bin/isqlplusctl stop |
| 5. Ultra Search | \$ORACLE_HOME/bin/searchctl stop |

d. Start Oracle Universal Installer

```
1$ $ORACLE_HOME/oui/bin/runInstaller
```

e. In the Welcome window, click Deinstall Products. In the Inventory screen, select the Oracle home and the products that you want to remove, then click Remove. Oracle Universal Installer displays a confirmation window asking you to confirm that you want to deinstall the products and their dependent components. When the products have been deleted, click Cancel to exit from Oracle Universal Installer, and then click Yes.

Removing Oracle 10g Software by Hand

Always use Oracle Universal Installer to remove Oracle software. Do not delete any Oracle home directories without first using Oracle Universal Installer to remove the software. I do not recommend the manual approach, exercise caution while deleting directories with root privilege.

If an attempt to remove the Oracle Software has failed for some reason, use below information to remove oracle manually.

Login as root

```
1# ORACLE_BASE=/u01/app; export ORACLE_BASE
```

```
2# ORACLE_HOME=$ORACLE_BASE/oracle/product/10.2.0/db10g1; export ORACLE_HOME
```

```
3# $ORACLE_HOME/bin/localconfig delete
```

This will stop Oracle CSS daemon and deletes the configuration

```
1# rm -f /etc/inittab.cssd
```

```
2# rm -rf /etc/oracle
```

This removes CSS daemon script and CSS configuration directory.

Issue below commands to remove

- | | |
|--------------------------------|---|
| 1# rm -rf \$ORACLE_BASE/* | #<--- Removes Entire Oracle Software Directory |
| 2# rm -f /etc/orainst.loc | #<--- Removes Install Loc |
| 3# rm -f /etc/oratab | #<--- Removes OraTab used by dbstart and dbshut scripts |
| 4# rm -f /usr/local/bin/dbhome | #<--- Removes Database Home Identifier Script |
| 5# rm -f /usr/local/bin/oraenv | #<--- Removes Env Script used by Bourne, Bash, or Korn shell |

```
6# rm -f /usr/local/bin/coraenv    #&lt;--- Removes Env Script used by C shell
```

Login as oracle and comment out below environment variables from oracle users' .bash_profile

```
1#export ORACLE_HOME=$ORACLE_BASE/oracle/product/10.2.0/db10g1
```

```
2#export PATH=$PATH:$ORACLE_HOME/bin
```

```
3#export LD_LIBRARY_PATH=$ORACLE_HOME/lib
```

11. If Oracle inventory is corrupted or missing? How to recover?

From 10g onwards, you can reverse engineer and recreate your Oracle inventory if it gets corrupted or accidentally deleted, thereby avoiding time consuming re-installation of Oracle S/W or any other unsupported tricks

You get the below error when opatch command is issued

```
oracle@myhost:/app/oracle$ opatch lsinventory
```

```
Invoking OPatch 11.2.0.1.6
```

```
Oracle Interim Patch Installer version 11.2.0.1.6
```

```
Copyright (c) 2011, Oracle Corporation. All rights reserved.
```

```
Oracle Home      : /app/oracle/product/10.2/db
```

```
Central Inventory : /app/oracleai/oralInventory
```

```
from             : /etc/oralnst.loc
```

```
OPatch version   : 11.2.0.1.6
```

```
OUI version      : 10.2.0.3.0
```

```
Log file location : /app/oracle/product/10.2/db/cfgtoollogs/patch/patch2011-12-27_13-19-08PM.log
```

```
OPatch failed to locate Central Inventory.
```

Possible causes are:

The Central Inventory is corrupted

The oralnst.loc file specified is not valid.

LsInventorySession failed: OPatch failed to locate Central Inventory.

OPatch failed with error code 73

```
oracle@myhost:/app/oracle$
```

You may also get this error because of incorrect inventory location. So it is a good idea to make sure the location of inventory is specified correctly in one of the following files depending upon your OS.

/etc/oralnst.loc

Contents of oralnst.loc

```
Bash-3.2$ cat /etc/oralnst.loc
```

```
inventory_loc=/app/oralInventory
```

```
inst_group=dba
```

Resolution:

If the error occurred due to missing or corrupt inventory, then you can recreate the inventory following the steps below.

Backup your existing oracle corrupted inventory if it exists.

Run the following OUI command from the Oracle home whose inventory is corrupt or missing.

```
Cd $ORACLE_HOME/oui/bin
```

```
./runInstaller -silent -attachHome ORACLE_HOME="/app/oracle/product/10.2/db"
```

```
ORACLE_HOME_NAME="Ora10202Home"
```

Note: Even though -attachHome was introduced with OUI version 10.1, it is documented with OUI 10.2 and higher.

UNIX for Oracle DBA

1. What's the difference between soft link and hard link?
2. How you will read a file from shell script?
3. What's the use of umask?
4. What is crontab and what are the arguments?
5. How to find operating system (OS) version?
6. How to find out the run level of the user?
7. How to delete 7 days old trace files?
8. How to get 10th line of a file (by using grep)?
9. (In Solaris) how to find out whether it's 32bit or 64bit?
10. What is paging?
11. What is top command?
12. How to find out the status of last command executed?
13. How to find out number of arguments passed to a shell script?
14. What is the default value of umask?
15. How to add user in Solaris/Linux?
16. How can you check the files permission in Unix OS?
17. How can you schedule job in Unix OS?
18. Which tool you use to copy files from windows to linux or linux to windows?
19. How can you search file in Unix OS?
20. Which command is used to see the disk space details in Unix OS?
21. Which comand is used to see files usage details in Unix OS?
22. Which command is used to schedule job without user intervention and backend?
23. What are frequently used OS command for DBA's?
24. Which command is used to copy or synchronizing two Directories in a secure way on any Unix environment?
25. Which command is useful to see line by line display in Unix environment?
26. Which command can be used to view file in readonly mode?
27. How we can delete the files which are few days(N days) old?
28. How to FIND whether OS (kernel) is 32 or 64 bit in Unix Operating systems?
29. Which command can be use to convert Windows file to Unix file?
30. How can you determine the space left in a file system?
31. How can you determine the number of SQLNET users logged in to the UNIX system?
32. What command is used to type files to the screen?
33. What command is used to remove a file?
34. Can you remove an open file under UNIX?
35. How do you create a decision tree in a shell script?
36. What is the purpose of the grep command?
37. The system has a program that always includes the word nocomp in its name, how can you determine the number of processes that are using this program?
38. What is an inode?
39. The system administrator tells you that the system hasn't been rebooted in 6 months, should he be proud of this?
40. What is redirection and how is it used?
41. How can you find dead processes?
42. How can you find all the processes on your system?

43. How can you find your id on a system?
44. What is the finger command?
45. What is the easiest method to create a file on UNIX?
46. What does >> do?
47. If you aren't sure what command does a particular UNIX function what is the best way to determine the command?
48. How to know whether server is 32 or 64 bit?
49. How do you automate starting and shutting down of databases in Unix/Linux?
50. How do you see how many oracle database instances are running?
51. You have written a script my_backup.sh to take backups. How do you make it run automatically every week?
52. What is OERR utility?
53. How do you see Virtual Memory Statistics in Linux?
54. How do you see how much hard disk space is free in Linux?
55. What is SAR?
56. What is SHMMAX?
57. Swap partition must be how much the size of RAM?
58. What is DISM in Solaris?
59. How do you see how many memory segments are acquired by Oracle Instances?
60. How do you see which segment belongs to which database instances?
61. What is VMSTAT?
62. How do you set Kernel Parameters in Red Hat Linux, AIX and Solaris?
63. How do you remove Memory segments?
64. What is the difference between Soft Link and Hard Link?
65. What is stored in oratab file?
66. How do you see how many processes are running in Unix/Linux?
67. How do you kill a process in Unix?
68. Can you change priority of a Process in Unix?

Answers

1. What's the difference between soft link and hard link?

A symbolic (soft) linked file and the targeted file can be located on the same or different file system while for a hard link they must be located on the same file system, because they share same inode number and an inode table is unique to a file system, both must be on the same file system.

2. How you will read a file from shell script?

```
while read line
do
```

```
  echo $line
done < file_name
```

3. What's the use of umask?

Will decide the default permissions for files.

4. What is crontab and what are the arguments?

The entries have the following elements:

field	allowed values
-------	----------------

minute	0-59
--------	------

hour	0-23
------	------

day of month	1-31
--------------	------

month	1-12
-------	------

day of week	0-7 (both 0 and 7 are Sunday)
-------------	-------------------------------

user	Valid OS user
------	---------------

command	Valid command or script
---------	-------------------------

? ? ? ? ? command

| | | | _____ day of the week (0-6, 0=Sunday)

| | | | _____ month (1-12)

| | | _____ day of the month (1-31)

| | _____ hour (0-23)

| _____ minute (0-59)

5. How to find operating system (OS) version?

```
uname -a
```

6. How to find out the run level of the user?

```
uname -r
```

7. How to delete 7 days old trace files?

```
find ./trace -name *.trc -mtime +7 -exec rm {} \;
```

8. How to get 10th line of a file (by using grep)?

9. (In Solaris) how to find out whether it's 32bit or 64bit?

10. What is paging?

11. What is top command?

top is a operating system command, it will display top processes which are taking high cpu and memory.

12. How to find out the status of last command executed?

```
$?
```

13. How to find out number of arguments passed to a shell script?

```
$#
```

14. What is the default value of umask?

```
022
```

15. How to add user in Solaris/Linux?

```
useradd command
```

16. How can you check the files permission in Unix OS?

'ls' command with the below option is used for checking files permission.

```
$ ls -altr
```

```
total 198
```

```
drwxr-xr-x 2 root root 4096 Aug 8 2008 srv
drwxr-xr-x 2 root root 4096 Aug 8 2008 mnt
drwxr-xr-x 3 root root 4096 Mar 7 05:10 home
drwxr-xr-x 7 root root 4096 Mar 7 05:47 lib64
```

17. How can you schedule job in Unix OS?

'crontab' command is used for scheduling job in Unix OS. Crontab can be installed and removed as followed: For commands that need to be executed repeatedly (e.g. hourly, daily or weekly), use crontab, which has the following options:

crontab filename Install filename as your crontab file.

crontab -e Edit your crontab file.

crontab -l Show your crontab file.

crontab -r Remove your crontab file.

18. Which tool you use to copy files from windows to linux or linux to windows?

Winscp can be used for copying files from windows to linux. Winscp is easy to install GUI based utility for copying files in binary mode by default. However 'ftp' command line utility can also be used.

19. How can you search file in Unix OS?

Using 'find' command we can search file in unix OS.

```
$ find /home/oracle -name "*.vim*" -print
```

```
/home/oracle/.viminfo
```

where: /home/oracle: path of the directory where you are searching file.

20. Which command is used to see the disk space details in Unix OS?

We use 'df -h' =>Linux OS, 'bdf' =>HP-UX OS, 'df -g' =>AIX OS

```
$ df -h
```

Note: In IBM-AIX OS we have to use 'df -g' ; In HP-UX OS we have to use 'bdf' for checking the Disk Space details .

21. Which comand is used to see files usage details in Unix OS?

'du' command is used to see directories and file usage details in Unix OS.

```
$ cd /u01
```

```
[oracle@node1.u01]$ du -csh *
```

22. Which command is used to schedule job without user intervention and backend?

'nohup' command is used to run a command without user intervention. If we use '&' symbol then the command will run in backend. 'nohup' command is my favourite command for running export and import jobs. I use this command very often.

Eg: nohup sh file_name.sh >file_output.out &

23. What are frequently used OS command for DBA's?

Ans: mkdir =>for creating directories in Unix OS; cd =>For changing the directory;

rm =>For removing files; rmdir =>For removing directories; grep =>For searching characters in a file; man =>User manual for all commands; useradd=>for creating OS user; chmod=>For granting permission to files and

Directories; chown=>For changing ownership of Directories.

Eg: mkdir test =>Create test directory

cd /u01/test =>Will change the path of Directory

man grep =>user manual for all command man has it own advantage.

chmod -R 775 /u01/test =>Grants permissions as read,write,execute to owner,groups

useradd -d oracle =>For Creating Operating system user as Oracle.

24. Which command is used to copy or synchronizing two Directories in a secure way on any Unix environment?

Ans : 'rsync' command can be used to copy or synchronize two directories. It is very important command and very handy tool for copying files fast for a DBA, Ofcourse 'scp' can also be use. But I really like using 'rsync' .Below is the example:

Eg:

```
[oracle@node1 HEALTH_CHECK_DB]$ rsync -av /u04/HEALTH_CHECK_DB/ /u05/SCRIPTS_DBA_TASKS/
```

Note: If you want to copy files from one linux(unix server) to other again it can be very handy, below is the syntax:

```
rsync -avc my_stuff2/ user@remotehost:~/mystuff3/
```

Important: rsync copies only those files which are not copied and this is definitely useful in synchronizing source and destination directories. Hence it is very fast.

25. Which command is useful to see line by line display in Unix environment?

Ans: 'less' command is used to see line-by-line display('more' for page-by-page display).I find less more useful specially for seeing log files and to find the errors or Warnings in the log files.

Eg: less test1.log =>Will display test1.log file

26. Which command can be used to view file in readonly mode?

Ans:'view' command can be used to view file in readonly mode.A very good option to see cronjob file specially because at any this file should not get modified by mistake as all your daily jobs will be scheduled using cronjob.

Eg: view crontab.oracle

27. How we can delete the files which are few days(N days) old?

Ans: To save the disk space you might be deleting the old files or backup which are 1 week(2 weeks) old or depending on your Disk Space and other requirement. We should automate these tasks as a DBA. We can do this as follows:

For Unix environment:

Eg: If I want to delete files from a path which are 7 Days old:

Write one shell script as given below:

```
#remove_files_back.sh
```

```
#Removing 7 days old dump files
```

```
find /u03/DB_BACKUP_TESTDB/expdp_fulldb_backup -mtime +6 -exec rm {} \;
```

Where: find =>Finding the file; /u03/DB_BACKUP_TESTDB/expdp_fulldb_backup =>path;

-mtime=>Modified time,Here I'm giving 7 days(>6);-exec rm =>execute removing for files. Now as per your convenience schedule cronjob for doing this task,For example every sunday at 9 pm than:

```
00 21 * * 0 /u03/DB_BACKUP_TESTDB/expdp_fulldb_backup/remove_files_back.sh 2>&1
```

```
>/u05/DB_BACKUP_TESTDB/logs/CRONJOBS_LOGS/TESTDB_BACK_cron.log
```

Note: For Windows Environment you can create a .bat file as follows:

```
--remove_file_back.bat
```

```
forfiles /p "D:\dbbackup\testdb\daily" /s /d -7 /c "cmd /c del @file : date >=7 days >NUL"
```

Where:All files which are 7 days old are removed;D:\dbbackup\testdb\daily =>path.

However Please make a note Please don't use the above command for your /(root) directory or any software specified and confirm and test in test environment before using on the actual system.

28. How to FIND whether OS (kernel) is 32 or 64 bit in Unix Operating systems?

```
file /usr/bin/initdb
```

29. Which command can be use to convert Windows file to Unix file?

'dos2unix' command can be use for this purpose.

After copying files with winscp use dos2unix utility to convert.

Unix uses a convention that a line is ended by a line feed character (Ascii 10). Windows/DOS uses a convention that a line is ended by a two character sequence, carriage-return line-feed (Ascii 13 then ascii 10). The dos2unix command converts for Windows/DOS format to Unix format. It writes the result to standard output. To write to a file, just redirect the standard output to the file. For example, use

```
$dos2unix myfile >mynewfile
```

30. How can you determine the space left in a file system?

There are several commands to do this: du, df, or bdf

31. How can you determine the number of SQLNET users logged in to the UNIX system?

SQLNET users will show up with a process unique name that begins with oracle, if you... do a ps -ef|grep oracle|wc -l you can get a count of the number of users.

32. What command is used to type files to the screen?

```
cat, more, pg
```

33. What command is used to remove a file?

```
rm
```

34. Can you remove an open file under UNIX?

```
yes
```

35. How do you create a decision tree in a shell script?

depending on shell, usually a case-esac or an if-endif or fi structure

36. What is the purpose of the grep command?

grep is a string search command that parses the specified string from the specified file or files

37. The system has a program that always includes the word nocomp in its name, how can you determine the number of processes that are using this program?

```
ps -ef|grep *nocomp*|wc -l
```

38. What is an inode?

An inode is a file status indicator. It is stored in both disk and memory and tracks file status. There is one inode for each file on the system.

39. The system administrator tells you that the system hasn't been rebooted in 6 months, should he be proud of this?

Maybe. Some UNIX systems don't clean up well after themselves. Inode problems and dead user processes can accumulate causing possible performance and corruption problems. Most UNIX systems should have a scheduled periodic reboot so file systems can be checked and cleaned and dead or zombie processes cleared out.

40. What is redirection and how is it used?

Redirection is the process by which input or output to or from a process is redirected to another process. This can be done using the pipe symbol "|", the greater than symbol ">" or the "tee" command. This is one of the strengths of UNIX allowing the output from one command to be redirected directly into the input of another command.

41. How can you find dead processes?

```
ps -ef|grep zombie — or — who -d depending on the system.
```

42. How can you find all the processes on your system?

Use the ps command

43. How can you find your id on a system?

Use the "who am i" command.

44. What is the finger command?

The finger command uses data in the passwd file to give information on system users.

45. What is the easiest method to create a file on UNIX?

Use the touch command

46. What does >> do?

The ">>" redirection symbol appends the output from the command specified into the file specified. The file must already have been created.

47. If you aren't sure what command does a particular UNIX function what is the best way to determine the command?

The UNIX man -k command will search the man pages for the value specified. Review the results from the command to find the command of interest.

48. How to know whether server is 32 or 64 bit?

```
getconf KERNEL_BITS
```

49. How do you automate starting and shutting down of databases in Unix/Linux?

One of the approaches is to use dbstart and dbshut scripts by init.d

Another way is to create your own script. To do that, create your own script "dbora" in /etc/init.d/ directory

```
# touch /etc/init.d/dbora
```

```
#!/bin/sh
```

```
# chkconfig: 345 99 10
```

```
# description: Oracle auto start-stop script.
```

```
# Applies to Oracle 10/11g
```

```
#
```

```
# Set ORA_HOME
```

```
# Set ORA_OWNER
```

```
ORA_HOME=/u01/app/oracle/product/10.2.0/db_1
```

```
ORA_OWNER=oracle
```

```
if [ ! -f $ORA_HOME/bin/dbstart ]
```

```
then
```

```
echo "Oracle startup: Error $ORA_HOME/bin/dbstart doesn't exist, cannot start "
```

```
exit
```

```
fi
```

```
case "$1" in
```

```
'start')
```

```
# Start the Oracle databases:
```

```
# The following command assumes that the oracle login
```

```
# will not prompt the user for any values
```

```

su - $ORA_OWNER -c "$ORA_HOME/bin/dbstart $ORA_HOME"
touch /var/lock/subsys/dbora
;;
'stop')
# Stop the Oracle databases:
su - $ORA_OWNER -c "$ORA_HOME/bin/dbshut $ORA_HOME"
rm -f /var/lock/subsys/dbora
;;
esac
Edit the "/etc/oratab" file and set the start flag of desired instance to 'Y'
MYDB1:/u01/app/oracle/product/10.2.0:Y
#Add dbora to init.d
[root@host ~]# chkconfig --add dbora
#Set the right permissions
chmod 750 /etc/init.d/dbora

```

50. How do you see how many oracle database instances are running?

Issue the following command "ps -ef | grep pmon"

```

[oracle@host ~]$ ps -ef | grep pmon | grep -v grep
oracle 7200 1 0 21:16 ? 00:00:00 ora_pmon_my_db_SID
oracle 9297 9181 0 21:42 pts/0 00:00:00 grep pmon

```

This will show within the paths returned the names of all instances (if you are OFA compliant - Oracle Flexible Architecture).

#Count them all:

```

[oracle@host ~]$ ps -ef | grep pmon | grep -v grep | wc -l

```

51. You have written a script my_backup.sh to take backups. How do you make it run automatically every week?

The Crontab will do this work.

Crontab commands:

```

crontab -e    (edit user's crontab)
crontab -l    (list user's crontab)
crontab -r    (delete user's crontab)
crontab -i    (prompt before deleting user's crontab)

```

Crontab syntax :

crontab entry consists of five fields: day date and time followed by the user (optional) and command to be executed at the desired time

```

* * * * * user command to be executed

```

```

- - - - -
| | | | |
| | | | +----- day of week(0-6)the day of the week (Sunday=0)
| | | +----- month (1-12)    the month of the year
| | +----- day of month (1-31) the day of the month
| +----- hour (0-23)          the hour of the day
+----- min (0-59)            the exact minute
#Run automatically every week - every 6th day of a week (Saturday=6)
* * * * 6 root /home/root/scripts/my_backup.sh
#TIP: Crontab script generator:
http://generateit.net/cron-job/

```

52. What is OERR utility?

Oerr is an Oracle utility that extracts error messages with suggested actions from the standard Oracle message files. Oerr is installed with the Oracle Database software and is located in the ORACLE_HOME/bin directory.

Usage: oerr facility error

Facility is identified by the prefix string in the error message.

For example, if you get ORA-7300, "ora" is the facility and "7300" is the error. So you should type "oerr ora 7300".

If you get LCD-111, type "oerr lcd 111", and so on. These include ORA, PLS, EXP, etc.

The error is the actual error number returned by Oracle.

Example:

\$ oerr ora 600

ora-00600: internal error code, arguments: [%s], [%s], [%s], [%s], [%s], [%s], [%s], [%s]

*Cause: This is the generic internal error number for Oracle program exceptions. This indicates that a process has encountered an exceptional condition.

*Action: Report as a bug - the first argument is the internal error number

53. How do you see Virtual Memory Statistics in Linux?

There is several ways to check mem stats: cat /proc/meminfo, top, free, vmstat...

"cat /proc/meminfo"

[user@host ~]\$ cat /proc/meminfo

[user@host ~]\$ top (sort by memory consumption by pressing SHIFT+O and then press "n")

[user@host ~]\$ free

TIP: free -m shows results in MB

[user@host ~]\$ vmstat

54. How do you see how much hard disk space is free in Linux?

"df" - reports filesystem disk space usage

TIP: "df -h" shows results in human readable format M/G/T

[user@host ~]\$ df -h

55. What is SAR?

SAR stands for Specific Absorption Rate, which is the unit of measurement for the amount of RF energy absorbed by the body when using a mobile phone.

SAR is an active remote sensing system; SAR antenna on a satellite that is orbiting the Earth and so on...

The question should be rather like this: What does sar command do in UNIX/LINUX like systems?

sar - Collect, report, or save system activity information.

[user@host ~]\$ sar

TIP: ls -la /var/log/sa/sar* ; man sar

More info: <http://computerhope.com/unix/usar.htm>

56. What is SHMMAX?

A: shmmx — maximum size (in bytes) for a UNIX/Linux shared memory segment

DESCRIPTION (docs.hp.com)

Shared memory is an efficient InterProcess Communications (IPC) mechanism.

One process creates a shared memory segment and attaches it to its address space.

Any processes looking to communicate with this process through the shared memory segment, then attach the shared memory segment to their corresponding address spaces as well.

Once attached, a process can read from or write to the segment depending on the permissions specified while attaching it.

How to display info about shmmx:

[user@host ~]\$ cat /etc/sysctl.conf | grep shmmx

kernel.shmmx=3058759680

57. Swap partition must be how much the size of RAM?

A tricky question coz opinions about this are always a good topic for any discussions.

In the past; once upon the time, when systems used to have 32/64 or max 128Mb of RAM memory it was recommended to allocate as twice as your RAM for swap partition.

Nowadays we do have a bit more memory for our systems.

An advice: always take software's vendor recommended settings into account and then decide.

For example, Oracle recommends always minimum 2GB for swap and more for their products - depends of the product and systems size.

Example of ORACLE Database 11g recommendations:

Amount of RAM	Swap Space
---------------	------------

Between 1 GB and 2 GB	1.5 times the size of RAM
-----------------------	---------------------------

Between 2 GB and 16 GB	Equal to the size of RAM
------------------------	--------------------------

More than 16 GB	16 GB
-----------------	-------

Another common example:

Equal the size of RAM, if amount of RAM is less than 1G

Half the size of RAM for RAM sizes from 2G to 4G.

More than 4G of RAM, you need 2G of Swap.

TIP: To determine the size of the configured swap space in Linux, enter the following command:

```
[user@host]~ grep SwapTotal /proc/meminfo
```

```
SwapTotal: 2096472 kB
```

To determine the available RAM and swap space use "top" or "free".

58. What is DISM in Solaris?

DISM = Dynamic Intimate Shared memory, which is used to support Oracle in Solaris Environment.

DISM is only supported from Solaris 9 and above version (not recommended to use in older version).

Until Solaris 8, only ISM (Intimate Shared Memory) which is of 8kb page size, from Solaris 9 Sun has introduced a new added feature which is DISM, which supports up to 4mb of page size.

Intimate Shared Memory

On Solaris systems, Oracle Database uses Intimate Shared Memory (ISM) for shared memory segments because it shares virtual memory resources between Oracle processes. ISM causes the physical memory for the entire shared memory segment to be locked automatically.

On Solaris 8 and Solaris 9 systems, dynamic/pageable ISM (DISM) is available. This enables Oracle Database to share virtual memory resources between processes sharing the segment, and at the same time, enables memory paging. The operating system does not have to lock down physical memory for the entire shared memory segment.

Oracle Database automatically selects ISM or DISM based on the following criteria:

- Oracle Database uses DISM if it is available on the system, and if the value of the SGA_MAX_SIZE initialization parameter is larger than the size required for all SGA components combined. This enables Oracle Database to lock only the amount of physical memory that is used.

- Oracle Database uses ISM if the entire shared memory segment is in use at start-up or if the value of the SGA_MAX_SIZE parameter is equal to or smaller than the size required for all SGA components combined.

Regardless of whether Oracle Database uses ISM or DISM, it can always exchange the memory between dynamically sizable components such as the buffer cache, the shared pool, and the large pool after it starts an instance. Oracle Database can relinquish memory from one dynamic SGA component and allocate it to another component.

Because shared memory segments are not implicitly locked in memory, when using DISM, Oracle Database explicitly locks shared memory that is currently in use at start-up. When a dynamic SGA operation uses more shared memory, Oracle Database explicitly performs a lock operation on the memory that is put to use. When a dynamic SGA operation releases shared memory, Oracle Database explicitly performs an unlock operation on the memory that is freed, so that it becomes available to other applications.

Oracle Database uses the oradism utility to lock and unlock shared memory. The oradism utility is automatically set up during installation. It is not required to perform any configuration tasks to use dynamic SGA.

59. How do you see how many memory segments are acquired by Oracle Instances?

ipcs - provides information on the ipc facilities for which the calling process has read access

#UNIX: SEGSZ

```
root> ipcs -pmb
```

```
IPC status from <running system> as of Mon Sep 10 13:56:17 EDT 2001
```

```
T   ID   KEY      MODE    OWNER   GROUP  SEGSZ  CPID
```

```
Shared Memory:
```

```
m   2400 0xeb595560 --rw-r----- oracle dba 281051136 15130
```

```
m    601 0x65421b9c --rw-r----- oracle dba 142311424 15161
```

#Linux: bytes

```
[user@host ~]$ ipcs
```

```
----- Shared Memory Segments -----
```

```
key   shmid  owner   perms  bytes  nattch status
```

```
0x00000000 32769  oracle  644   122880  2      dest
```

60. How do you see which segment belongs to which database instances?

This can be achieved with help of ipcs tool and sqlplus; oradebug ipc

#Linux

```
[user@host ~]$ ipcs
```

```
----- Shared Memory Segments -----
```

```
key   shmid  owner   perms  bytes  nattch status
```

```
0x00000000 32769  oracle  644   122880  2      dest
```

#UNIX:

```
root> ipcs -pmb
```

IPC status from <running system> as of Mon Sep 10 13:56:17 EDT 2001

T	ID	KEY	MODE	OWNER	GROUP	SEGSZ	CPID
---	----	-----	------	-------	-------	-------	------

Shared Memory:

m	32769	0xeb595560	--rw-r----	oracle	dba	281051136	15130
---	-------	------------	------------	--------	-----	-----------	-------

m	601	0x65421b9c	--rw-r----	oracle	dba	142311424	15161
---	-----	------------	------------	--------	-----	-----------	-------

m	702	0xe2fb1874	--rw-r----	oracle	dba	460357632	15185
---	-----	------------	------------	--------	-----	-----------	-------

m	703	0x77601328	--rw-r----	oracle	dba	255885312	15231
---	-----	------------	------------	--------	-----	-----------	-------

#record value of shmid "32769" (ID in UNIX)

```
[user@host ~]$ sqlplus /nologin
```

```
SQL> connect system/manager as sysdba;
```

```
SQL> oradebug ipc
```

#Information have been written to the trace file. Review it.

In case of having multiple instances, grep all trace files for shmid 32769 to identify the database instance corresponding to memory segments.

#scrap of trace file MY_SID_ora_17727.trc:

Area	Subarea	Shmid	Stable Addr	Actual Addr
------	---------	-------	-------------	-------------

1	1	32769	000000038001a000	000000038001a000
---	---	-------	------------------	------------------

61. What is VMSTAT?

vmstat - Reports virtual memory statistics in Linux environments.

It reports information about processes, memory, paging, block IO, traps, and cpu activity.

```
[user@host ~]$ vmstat
```

```
procs -----memory----- ---swap-- ----io---- --system-- ----cpu----
```

```
r b swpd free buff cache si so bi bo in cs us sy id wa
```

```
3 0 170224 121156 247288 1238460 0 0 18 16 1 0 3 2 95 0
```

62. How do you set Kernel Parameters in Red Hat Linux, AIX and Solaris?

sysctl - configure kernel parameters at runtime

EXAMPLES

```
/sbin/sysctl -a (Display all values currently available)
```

```
/sbin/sysctl -w kernel.shmmax = 3058759680 ( -w this option changes a sysctl setting)
```

To modify settings permanetly edit /etc/sysctl - kernel sysctl configuration file and then issue the following command:

```
/sbin/sysctl -p /etc/sysctl.conf ( Load in sysctl settings from the file specified or /etc/sysctl.conf if none given)
```

63. How do you remove Memory segments?

We can use ipcs and ipcrm command

ipcs - provides information on ipc facilities

ipcrm - removes a message queue, semaphore set or shared memory id

First kill all Oracle database processes; A shared memory object is only removed after all currently attached processes have detached.

#UNIX

```
root> ps -ef | grep $ORACLE_SID | grep -v grep | awk '{print $2}' | xargs -i kill -9 {}
```

```
root> ipcs -pmb #displays held memory
```

IPC status from /dev/kmem as of Tue Sep 30 11:11:11 2011

T	ID	KEY	MODE	OWNER	GROUP	SEGSZ	CPID	LPID
---	----	-----	------	-------	-------	-------	------	------

Shared Memory:

m	25069	0x4e00e002	--rw-r----	oracle	dba	35562418	2869	23869
---	-------	------------	------------	--------	-----	----------	------	-------

m	10x4bc0eb18	--rw-rw-rw-	root	root	31008	669	669	
---	-------------	-------------	------	------	-------	-----	-----	--

RAM memory segment owned by Oracle is ID=25069.

```
root> ipcrm -m 25069 #this command will release that memory segment
```

64. What is the difference between Soft Link and Hard Link?

"Symbolic links" (symlinks/soft link) are a special file type in which the link file actually refers to a different file or directory by name.

When most operations (opening, reading, writing, and so on) are passed the symbolic link file, the kernel automatically "dereferences" the link and operates on the target

of the link. But remove operation works on the link file itself, rather than on its target!

A "hard link" is another name for an existing file; the link and the original are indistinguishable.

They share the same inode, and the inode contains all the information about a file.

It will be correct to say that the inode is the file. Hard link is not allowed for directory (This can be done by using mount with --bind option)!

ln - makes links between files. By default, it makes hard links; with the "-s" option, it makes symbolic (soft) links.

Synopses:

ln [OPTION]... TARGET [LINKNAME]

ln [OPTION]... TARGET... DIRECTORY

EXAMPLE:

#hard links

```
[user@host ~]$ touch file1
```

```
[user@host ~]$ ln file1 file2
```

```
[user@host ~]$ ls -li
```

```
total 0
```

```
459322 -rw-r--r-- 2 userxxx users 0 May 6 16:19 file1
```

```
459322 -rw-r--r-- 2 userxxx users 0 May 6 16:19 file2 (the same inode, rights, size, time and so on!)
```

```
[user@host ~]$ mkdir dir1
```

```
[user@host ~]$ ln dir1 dir2
```

```
ln: `dir1': hard link not allowed for directory
```

#symbolic links

```
[user@host ~]$ rm file2 #hard link removed
```

```
[user@host ~]$ ln -s file1 file2 #symlink to file
```

```
[user@host ~]$ ln -s dir1 dir2 #symlink to directory
```

```
[user@host ~]$ ls -li
```

```
total 12
```

```
459326 drwxr-xr-x 2 userxxx users 4096 May 6 16:38 dir1
```

```
459327 lrwxrwxrwx 1 userxxx users 4 May 6 16:39 dir2 -> dir1 (dir2 refers to dir1)
```

```
459322 -rw-r--r-- 1 userxxx users 0 May 6 16:19 file1
```

```
459325 lrwxrwxrwx 1 userxxx users 5 May 6 16:20 file2 -> file1 (different inode, rights, size )
```

```
[user@host ~]$ rm file2 #will remove a symlink NOT a target file; file1
```

```
[user@host ~]$ rm dir2
```

```
[user@host ~]$ ls -li
```

```
[user@host ~]$ ls -li
```

```
total 4
```

```
459326 drwxr-xr-x 2 userxxx users 4096 May 6 16:38 dir1
```

```
459322 -rw-r--r-- 1 userxxx users 0 May 6 16:19 file1
```

```
[user@host ~]$ info coreutils ln #(should give you access to the complete manual)
```

65. What is stored in oratab file?

This file is being read by ORACLE software, created by root.sh script which is being executed manually during the software installation and updated by the Database Configuration Assistant (dbca) during the database creation.

File location: /etc/oratab

ENTRY SYNTAX:

\$ORACLE_SID:\$ORACLE_HOME:<N|Y>:

\$ORACLE_SID - Oracle System Identifier (SID environment variable)

\$ORACLE_HOME - Database home directory

<N|Y> Start or not resources at system boot time by the start/stop scripts if configured.

Multiple entries with the same \$ORACLE_SID are not allowed.

EXAMPLES:

```
[user@host ~]$ cat /etc/oratab
```

```
MYDB1:/u01/app/oracle/product/10.2.0/db:Y
```

```
emagent:/u01/app/oracle/product/oem/agent10g:N
```

```
client:/u01/app/oracle/product/10.2.0/client_1:N
```

```
emcli:/u01/app/oracle/product/oem/emcli:N
```

66. How do you see how many processes are running in Unix/Linux?

ps" with "wc" or "top" does the job.

ps - report a snapshot of the current processes

wc - print the number of newlines, words, and bytes in files

top - display Linux tasks (better solution)

In other words ps will display all running tasks and wc will count them displaying results:

```
[user@host ~]$ ps -ef | wc -l
```

149

```
[user@host ~]$ ps -ef | grep -v "ps -ef" | wc -l #this will not count ps process executed by you
```

148

#using top

```
[user@host ~]$ top -n 1 | grep Tasks
```

Tasks: 148 total, 1 running, 147 sleeping, 0 stopped, 0 zombie

67. How do you kill a process in Unix?

A: kill - terminate a process

killall - kill processes by name

kill -9 <PID> #kill process with <PID> by sending SIGKILL 9 Term Kill signal

killall <process name>

EXAMPLES:

```
[user@host ~]$ ps -ef | grep mc
```

```
user 31660 31246 0 10:08 pts/2 00:00:00 /usr/bin/mc -P /tmp/mc-user/mc.pwd.31246
```

```
[user@host ~]$ kill -9 31660
```

Killed

#killall

```
[user@host ~]$ killall mc
```

Terminated

```
[user@host ~]$ killall -9 mc
```

Killed

68. Can you change priority of a Process in Unix?

YES. nice & renice does the job.

nice - Runs COMMAND with an adjusted scheduling priority. When no COMMAND specified prints the current scheduling priority.

ADJUST is 10 by default. Range goes from -20 (highest priority) to 19 (lowest).

renice - alters priority of running processes

EXAMPLES: (NI in top indicates nice prio)

```
[user@host ~]$ mc
```

```
[user@host ~]$ top
```

PID	USER	PR	NI	VIRT	RES	SHR	S	%CPU	%MEM	TIME+	COMMAND
1896	userxxx	17	0	7568	1724	1372	S	0	0.0	0:00.03	mc

```
[user@host ~]$ nice -n 12 mc #runs the mc command with prio of 12.
```

```
[user@host ~]$ top
```

PID	USER	PR	NI	VIRT	RES	SHR	S	%CPU	%MEM	TIME+	COMMAND
1763	userxxx	26	12	6832	1724	1372	S	0	0.0	0:00.03	mc

```
[root@host ~]# renice +16 1763 #must be a root user
```

1763: old priority 12, new priority 16

FAQ Asked by Companies

NDS

1. Tell me something about yourself ?
2. what is your daily activities
3. what are your roles and responsibilities ?
4. what is meant by alert log ? what do you observe means what type of errors (monitoring)
5. what is an instance ? what components does it contain ?
6. what is SGA ? what it contains ?
7. what is background process ? what it contains ?
8. when oracle process is going which b/g process run first ?
9. how can oracle process authenticate is a valid user user or not ?
10. what happens at startup Nomount ?
11. what happens at startup Mount stage ?
12. when oracle process is reading a SP FILE which component it is likely to check ?
13. what is purpose of control file ?
14. what is purpose of redolog file ?
15. why do you mirror redo logs ?
16. what ,s your BACKUP STRATEGY ?
17. What is export and import strategy ?
18. what does controlfile contain ?
19. what are components effects the controlfile size ?

SIEMENS

1. How do you alter table space?
2. FINGER command
3. TOUCH command
4. IMPORT/EXPORT what happens if we keep DIRECT = Y?
5. Day started and supposes 90 % data file filled? What do you do?
6. LOCAL/Dictionary tablespaces differences?
7. STATS PACK
8. MTS (Multi Thread Services)
9. COLD/HOT Differences?
10. Control files BACKUP?
11. Trace files?
12. Listener.ora?

Q SOFT

1. How many instances in your company?
2. Suppose your db is 90 GB how do u assign SGA?
3. What are the components of SGA?
4. Database Buffer cache is useful for what?
5. What is DML Statement?
6. Suppose a redo log file fails what would be happen what is going on?
7. Difference between Shutdown Transactional and NORMAL?
8. What is SP FILE? Suppose u have sp1, sp2, pfile which will run first?
9. Suppose 90 % of data file filled? What do u do?
10. HOT BACKUP ADVANTAGES?
11. Write command on user assign Temp, Default TS? What contains SYSTEM TS?
12. RBS and UNDO?
13. Installation on Linux, parameter in KERNEL?
14. How many tables?
15. What is Hash Partioning ?

16. Export utility ? SHOW =Y & DIRECT =Y explain it ?
17. How can use quota ?
18. PGA_AGGREGATE_TARGET?
19. Explain about Alert log file ?
20. Optimiser RBO,CBO how can we say RBO AND CBO ?
21. Logical Backup ?
22. When you create indexes ? After inserting data you created index on same?

ORACLE

1. Suppose SID= 100 how can you check in OS level?
2. Altering a data file? Any pre requisite?
3. What is your setup i.e. Environment?
4. How can you verify log verification?
5. Auto Extend on?
6. U added a data file how can u see that?
7. How many processes are running? how can we see that ?
8. How can I increase my performance?
9. Difference between ls -ltr and ls -l?

CHENNAI COMPANY

1. OS level disk free space? (df -h)
2. Hot backup steps?
3. controlfile command
4. BACKUP TIME?
5. Compress =y (IMPORT /EXPORT) ?
6. Does export possible in diff OS 's ?
7. Version of OS ?
8. How can you identify port No ?
9. Default Port No ?
10. While Installing Oracle on Linux what is the last step ?
11. Can user create a db ?
12. What is DB links ?
13. Which user can Install Oracle ? (root user or \$ user)
14. How many memory layers are in the shared pool ?
15. How do you find out in the RMAN catalog a particular archive log has been backed up ?
16. What is ORA 01113 & 01110 ?
17. What is Dual Table ?
18. What is for Reco ?
19. Explain the setup in your company ?
20. How many redologs are generated in your company and tell the size ?

FINCH

1. what is check point ?
2. How to manage Idle Users ? (idle_time in profile)
3. BACKUP Strategy of your company ?
4. When u take cold backup and RMAN Backup ?
5. Monitoring of Dynamic Views & Data Dictionary Views ?
6. SQL Tuning ?
7. what is semaphore while installing ORACLE ?
8. Datafile size ?
9. How do u find the overall size of database ?
10. How do u find USED and FREE space in DATABASE ?
11. Shared Realm memory does not exist ?
12. Status of Redologs in ARCHIVE and NOARCHIVE mode of DB ?
13. What happens internally when we keep in BEGIN BACKUP MODE ?

14. What are the advantages of CATALOG DB ?
15. What is RMAN Repository ?

IBM

1. Tell me about yourself ?
2. Backup Strategy ?
3. what happens internally when we keep in BEGIN BACKUP MODE ?
4. How do u include user managed backups in RMAN ?
5. What happens internally while taking backup using RMAN ?
6. If a block is found corrupt while taking backup using RMAN does RMAN take backup of that datafile or will it terminate processing ?
7. what is DBWR,LGWR ?
8. What is SMON and PMON ?
9. How does ckpt process helpful during instance recovery ?
10. How can we image copy of datafile using RMAN ?
11. How can we compress backup set ?
12. what are new features in 10g release 2 ?
13. How do you Configure RMAN ?

WIPRO

I ROUND

1. Tell about yourself ?
2. What are your daily activities ?
3. What type of DB u have i.e. OLTP/DEVELOPMENT/OLAP ?
4. Undo tablespace OPTIONS ?
5. How many types one can create db ?
6. Suppose u have deleted datafile ? can we recover or not provided no backup is available ?
7. What is DBWR ?
8. What are features of oracle 9i ?
9. What is core level sql statement ?
10. Suppose waste data inserted into a table ? I want to remove that ? how ?
11. 4 pm meeting going on ? db crash ? what is scenario ?

II ROUND

1. What is your company profile ?
2. Which type of backups are using ?
3. Difference between RMAN and User Managed ?
4. Difference between Export /Import ? what Features ?
5. OLTP backup manager ?
6. SQL statement Tuning ?
7. Row buffer cache ?
8. STATS PACK ?

MIND TREE

1. Definition of Function, Package , Procedure ?
2. What do you mean by dual table why it is required ?
3. Snapshot too old error 01555 ?
4. Difference between HOT & COLD backup ? Why it is required to be archive in online backup ?
5. What is meant by Trigger ?

TCS (MUMBAI)

1. Difference between 8i ,9i and 10g ? Explain grid ?
2. How many components in SGA and DATABASE ?
3. What is checkpoint ?
4. What is PMON ?
5. I want to change NOARCHIVE TO ARCHIVE . What are the steps ?
6. COLD & HOT backup steps ?
7. Difference between RMAN and LOGICAL backups ?
8. While starting the Linux , what will happen . First GUI or any command to enter to GUI ?
9. SQL Loader , suppose while loading data error will occur ? will I load or not ? If not what is the solution ?
10. Session ? what we can able to see in that ?
11. Suppose in 24/7 Env 5 datafiles are lost ? how can you recover ? Steps ?
12. What is difference between Windows and Unix Environment ?
13. Suppose one user opened a session ? How can we see that session is opened or not at O/S level ?
14. How many memory layers are there in Shared Pool ?
15. How do you find out in the RMAN catalog if a particular archive log has been backup or not ?
16. How can you tell how much space is left on a given file system ?

WIBEN TECHNOLOGIES

1. Daily activities ?
2. Database version ?
3. DB Size ?
4. HOW many instances you have and how many development boxes u have ?
5. Team size ?
6. Tell about physical backups ?
7. How to take physical backups can you explain ?
8. What is logical backup ?
9. What is your backup strategy ?
10. Difference between logical and physical backup ?

AXA

1. Tell me about yourself ?
2. Team size ?
3. How many DB u have production , development & QA ?
4. SIZE of your Production DB ?
5. Size of your DEVELOPMENT DB ?
6. What is the size of your Redo log ?
7. Size of SGA ?
8. Control file contains what parameters ?
9. what does Pfile Contains ?
10. What is PCT Free ?
11. what is db block size ?
12. What is SQL Loader ?
13. How can you load data in Sql Loader ?
14. I have one updated row how can you insert already updated table by using SQL Loader ?
15. How can you increase TS Size and Datafile Size ?
16. What is Dirty buffer ?
17. what is alert log ? what is it contain ?
18. ANY issues did u face ?
19. Generally what errors you got ?
20. Can you tell me what is diff between 9i and 10g ?
21. Explain 10g features ?
22. Tell me Oracle installation steps ?
23. What is your backup strategy ?
24. When u take hot backup and how much time it take to complete ?

25. what is cold backup ?
26. What is db refresh ? How you do db refresh ?
27. What is Stats Pack Analysis ?
28. What is PT ? Have you ever done it ?
29. Can you tell me how many controlfiles you have ?
30. If I have created 9 controlfiles what happens ? why ?
31. Control file contains what ?
32. If listener is failed what is the error ?
33. One of the user not able to connect to db he is getting listener failed error ? what is the reason ? what error ?
34. what is recovery catalog
35. What is TKPROF ?
36. One of my user is complaining that system is so slow ? what do u do ?
37. If the user does not have the privilege to enable session , As a DBA what u do?
38. Suppose If have 1000 statements in my report and I want to see the top 5 resource consuming statements ? how to see those statements ?
39. How to enable session at Instance level ?

IBM

1. how do you apply patch in rac
2. what is cache fusion.
3. how do you add disk to asm disk group
4. what is incarnation
5. how do you sync DG when there is some archive logs missing
6. difference between incremental and cumulative backup's
7. what is rebalancing
8. do you know about OS watcher
9. how to add node in RAC
10. what happens if archive dest is full
11. how will perform clone
12. how to upgrade a rac database.
13. what do you mean rolling upgrade.
14. what are diff protection mode available is DG
15. diff between exp and datapump
16. what is inventory location
17. what happens if inventory is corrupted
18. what is fractured block.
19. what is partial check point
20. what are background processes in asm
21. how will find out the number of clients running under one ASM ins
22. how do you recover undo with no down time
23. how recover a lost datafile.
24. what is the background process that write into alert log file
25. sga_target,sga_max_size
26. root.sh,orainstroot.sh
27. how to find nodes in RAC