

RMAN history and internals

Tuomas Pystynen
Deepbase Consulting Oy

Who am I ?

- Independent consultant since 1997
 - Database and application server expertise
- Oracle Server Technologies
 - RMAN development
 - Oracle7 testing
- Oracle experience since version 4
 - Nokia, Oracle Finland, ...
- Member of the Oaktable Network

Online backup

- Introduced in Oracle V6
- Must put datafiles in hot backup mode
 - alter tablespace ... begin backup
 - freezes checkpoint SCN in datafile headers
 - generates extra redo to recover fractured blocks
- Backup datafiles using external utility
- Must put datafiles back in normal mode
 - alter tablespace ... end backup
- Online backups are inconsistent
 - must be recovered using archived/online logs

Checkpoint SCN

- Every datafile header has checkpoint SCN
 - datafile contains all changes to checkpoint SCN
 - updated when checkpoint completes
 - recovery starts from checkpoint SCN
- Checkpoint SCN is frozen in backup mode
 - external backup utility might read header last
 - other writes continue normally

Fractured database block

- External backup utility may use multiple reads to read a database block
 - it is possible DBWR writes the block in between
 - the database block in the backup is fractured
- Extra redo is generated in hot backup mode
 - when first change is made to a database block, whole block is written to redo log
 - to guarantee that fractured blocks can be repaired during recovery

Enterprise Backup Utility (EBU)

- Developed by Unix porting group for Oracle7
- Made backing up databases easier
 - especially when using raw devices
- Supported tape backups
 - defined the SBT interface used by RMAN

Recovery Manager (RMAN)

- Utility developed by database development for Oracle8
- Uses database server to backup database
- Keeps track of backups in control file and recovery catalog
- Simplifies backup and restore operations using information in the recovery catalog

DBMS_BACKUP_RESTORE

- A fixed PL/SQL package used by RMAN to backup and restore files
 - available also when database is not open
- No need to put datafile in hot backup mode
 - header block is always read first
 - header block is written out last
 - to detect when backup is not complete
- No problem with fractured blocks
 - always read database blocks in one IO
 - unless IO drivers break on IO to multiple read
 - always perform fractured block check
 - re-read database block if fractured
 - also verifies checksums if present

Backup sets

- Backup multiple datafiles or archived logs to one backup set
 - can write backup set to disk or tape
 - need a MML to write to tape
 - read multiple files in parallel to keep tape drive streaming
 - backup set can be divided in to multiple pieces
 - backup set format is proprietary to Oracle
 - may need to read whole backup set to restore a single file

Backup sets

- Database blocks which have never been used are excluded from backup sets
 - in Oracle10g R2 also blocks currently empty are excluded
 - restore recreates empty blocks
- Backup sets written to disk can be compressed in Oracle10g
 - takes only about 1/5 of the disk space
 - restore automatically uncompresses

Incremental backups

- Incremental backups contain only database blocks modified since previous backup
 - every database block contains the SCN of the last change
 - Oracle10g can also maintain a bitmap of changed blocks
 - alter database enable block change tracking using file
- Incremental backups are applied to restored datafiles in recovery phase
 - in Oracle10g incremental backups can be applied to datafile copies

Datafile copies

- An image copy identical to the datafile
 - except some information in the file header
- Datafile copies are always written to disk
 - can switch database to use the copy
 - no restore required

Proxy copies

- A copy created by an external utility
 - usually directly from disk to tape
- RMAN controls proxy backups
 - puts the datafile in backup mode
 - invokes external utility via SBT interface
 - passes filenames to MML
 - records the proxy copy in control file
 - takes data file out of backup mode
- Proxy copy is not seen by database server
 - no checking for corruption
 - incremental backups not possible

Controlfile

- DBMS_BACKUP_RESTORE records successful backups in controlfile
- Fixed views show controlfile information
 - v\$backup_set
 - v\$backup_piece
 - v\$backup_datafile
 - v\$datafile_copy ...
- Circular reuse records are kept for CONTROLFILE_KEEP_TIME days
 - if controlfile is large enough or can expand
 - otherwise records may be reused earlier

Controlfile

- Controlfile contains non-reusable records
 - database
 - tablespace
 - datafile
 - online redo log ...
- and circular reuse records
 - archived log
 - backup set and piece
 - backup datafile and datafile copy ...
- 34 different record types in 10g R2
 - almost a database itself

Controlfile transaction

- Session must obtain CF enqueue first
- Controlfile contains 2 versions of each block
 - bitmap indicates which version is current
 - transaction modifies non-current version
- At the end of transaction
 - controlfile sequence number is incremented
 - bitmap is updated and written out atomically
 - CF enqueue is released

RMAN language

- RMAN has its own language
 - can type commands interactively
 - can write scripts common operations
- RMAN parses and executes the language
 - parser is generated, not written by hand
 - RMAN generates PL/SQL blocks using templates defined in `recover.bsq`
 - PL/SQL blocks are executed using non-blocking OCI calls
 - one database connection per channel

Recovery Catalog

- Recovery catalog consist of tables stored in an Oracle database
 - not in the database to be backed up
 - in its own database schema
- Recovery catalog data is replicated from controlfile
 - using RESYNC CATALOG command
 - implicitly by other RMAN commands

Recovery catalog resync

- Full resync synchronizes all information
 - creates a snapshot of the controlfile
 - to avoid concurrency problems and
 - to get consistent results
 - full resync is used to ensure that information about database physical structure is up-to-date
 - for example backup database needs up-to-date list of all datafiles
- Partial resync only circular reuse records
 - reads the controlfile directly
 - retries if it gets inconsistent_read error
 - used when full resync is not needed
 - for example when backup completes

Recovery Catalog

- RMAN uses DBMS_RCVMAN package to query the recovery catalog
- RMAN can operate without recovery catalog
 - with some limitations using control file instead
 - a fixed version of DBMS_RCVMAN package accesses controlfile
- Recovery catalog contains more history
 - dropped tablespaces and datafiles
 - all backups since creation of the catalog

Backup

- Specify what to backup
 - database
 - tablespace(s)
 - datafiles(s)
 - archivelog range
- RMAN resolves what files to backup
- Easy to parallelize backup
 - define multiple channels

Restore

- Specify what needs to be restored
 - database
 - tablespace(s)
 - datafile(s)
 - archivelog range
- Optionally specify where to restore to
- RMAN will determine what files to restore and from what backups
 - usually the latest backup unless doing point-in-time recovery
 - can influence backup selection using tags

Recover

- Recover restored datafiles by applying
 - incremental backups
 - offline ranges
 - archived logs
- Applies incremental backups and offline ranges first
 - also NOLOGGING operations are recovered
 - using DBMS_BACKUP_RESTORE calls
- Applies archived logs last
 - using SQL RECOVER statement