
CERN

Szymon Skorupinski

Prepare for the Worst:

Reliable Data Protection with Oracle RMAN and Oracle Data Guard

Outline

- Who am I
- CERN and Oracle
- Backup system
- Recovery system
- Backups from standby databases
- Our strategy and recommendations
- Conclusions

Outline

- Who am I
- CERN and Oracle
- Backup system
- Recovery system
- Backups from standby databases
- Our strategy and recommendations
- Conclusions

Who am I

- Work with Oracle technologies for 7+ years
  - DBA in CERN's IT Department Database Group since April 2011
- Honoured to be responsible for backup and recovery of CERN's Oracle databases
- The first international conference appearance...
  - ...now, so please be tolerant ;)
- Certificates
  - Oracle Certified Professional – 10g
  - IBM Certified System Administrator – AIX 6.1
- Contact
  - sskorupi@cern.ch

Outline

- Who am I
- CERN and Oracle
- Backup system
- Recovery system
- Backups from standby databases
- Our strategy and recommendations
- Conclusions

CERN

- European Organization for Nuclear Research
  - World's largest centre for scientific research, founded in 1954
  - Research: Seeking and finding answers to questions about the Universe
  - Technology, International collaboration, Education

Twenty Member States

Austria, Belgium, Bulgaria, Czech Republic, Denmark, Finland, France, Germany, Greece, Italy, Hungary, Netherlands, Norway, Poland, Portugal, Slovakia, Spain, Sweden, Switzerland, United Kingdom

Seven Observer States

European Commission, USA, Russian Federation, India, Japan, Turkey, UNESCO

Associate Member States

Israel, Serbia

Candidate State

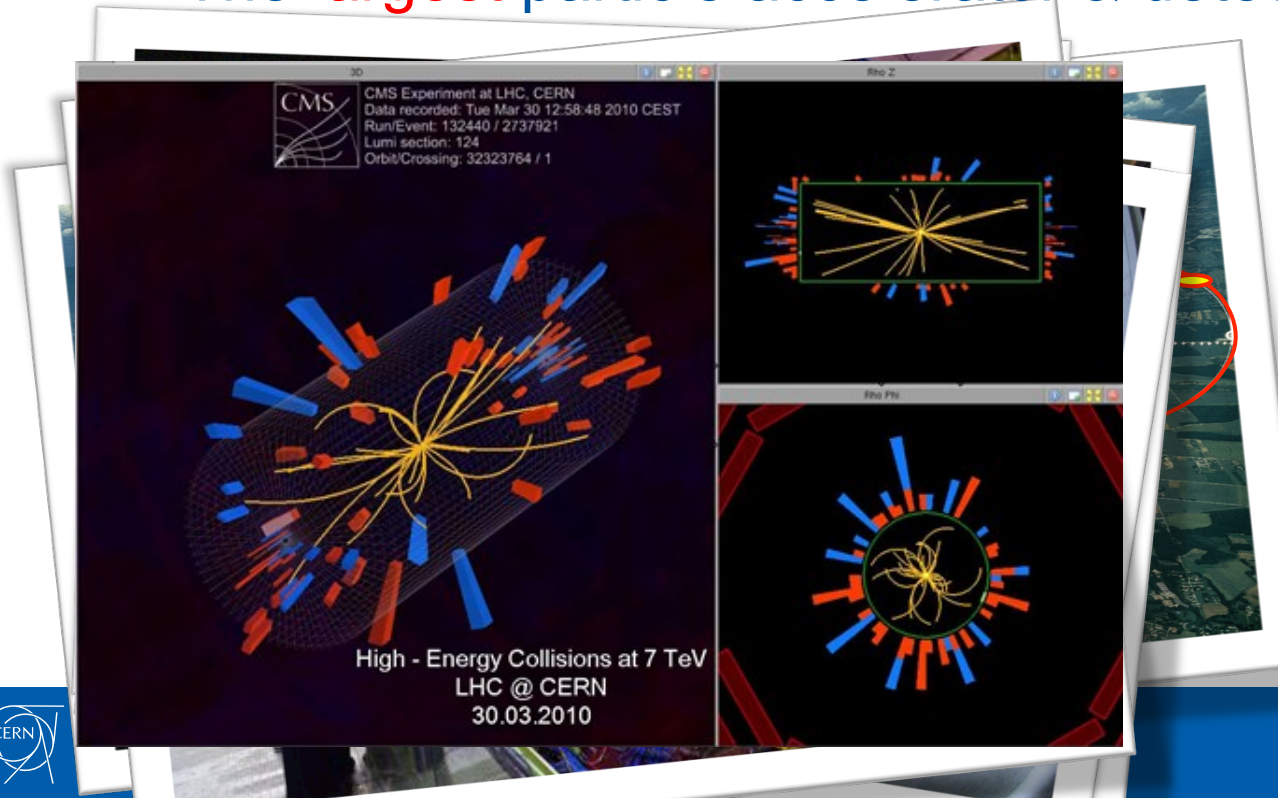
Romania

People

~2400 Staff, ~900 Students, post-docs and undergraduates, ~9000 Users, ~2000 Contractors

LHC

- The largest particle accelerator & detectors



17 miles (27km) long underground tunnel
Thousands of superconducting magnets
Coldest place in the Universe
-271.3 °C (1.9 K)
but also...
...hottest place in the Universe :)
100k times hotter than the heart of the Sun
600 million collisions per second detected by about 150 millions of sensors

WLCG

- World's largest computing grid

More than 20 Petabytes
of data stored and analysed
every year

Over 68 000 physical CPUs
Over 305 000 logical CPUs

157 computer centres in 36
countries

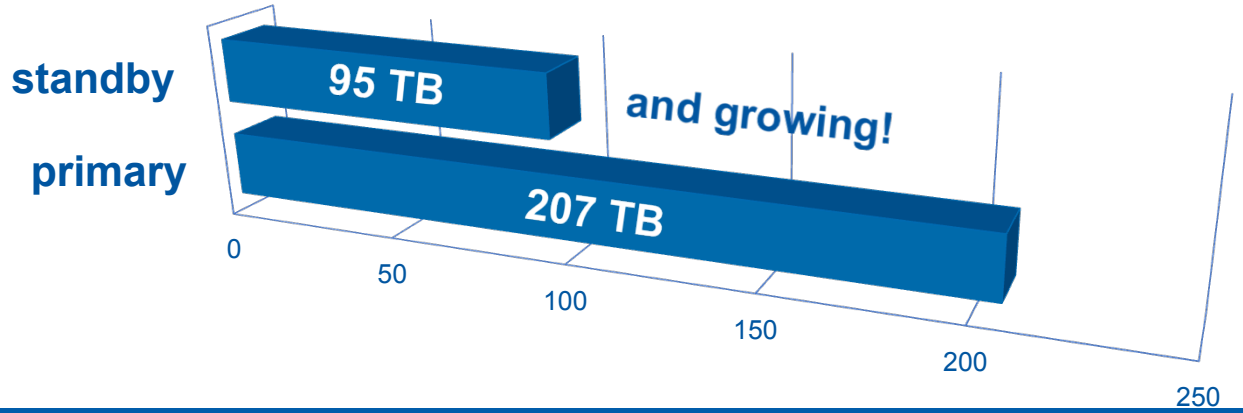
More than 8000 physicists with
real-time access to LHC data

CERN's databases

- CERN IT-DB Group manages Oracle databases used by LHC and its experiments, as well as financial and administrative ones
  - Also MySQL On Demand service now in production
- >100 Oracle databases, most of them RAC
  - Mostly NAS storage plus some SAN with ASM
- >70 Oracle databases backed up to tapes
  - On average ~5.1 TB of redo daily, ~302 TB of datafiles in total
- The biggest Oracle databases at CERN
  - LHC logging database ~145 TB, expected growth up to 70 TB / year
  - 13 production experiments' databases ~122 TB in total

Backups source

- 10 (~95 TB) out of the biggest production experiments' databases are already backed up to tapes from standby with Active Data Guard option
  - Still ongoing process of changing backups source for the rest

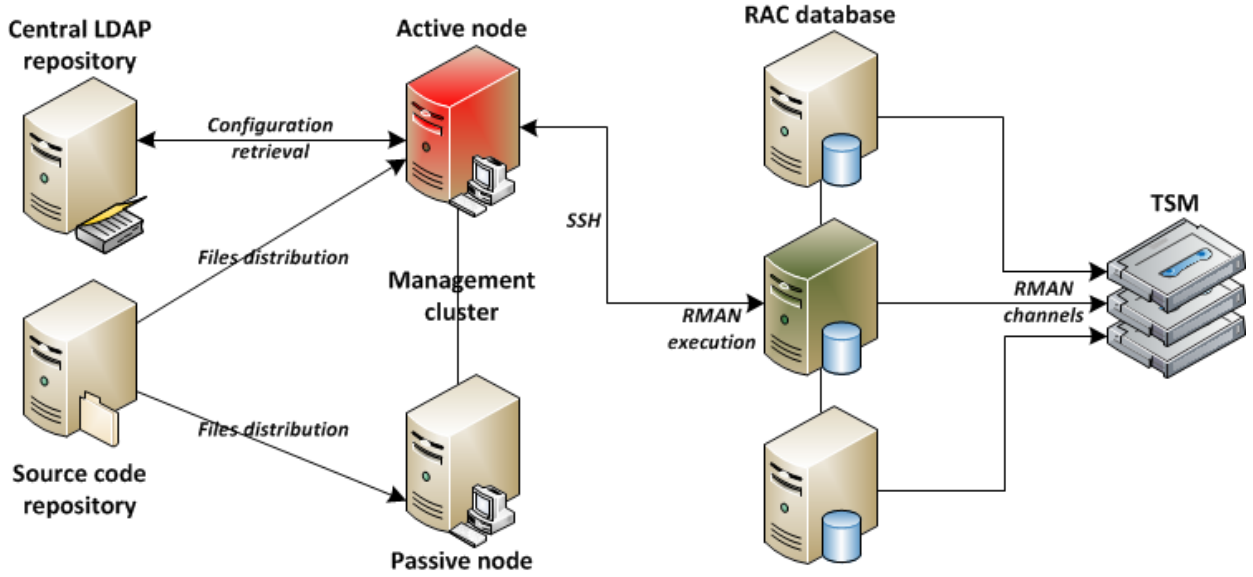


Outline

- Who am I
- CERN and Oracle
- Backup system
- Recovery system
- Backups from standby databases
- Our strategy and recommendations
- Conclusions

Our backup system

- Centrally managed and clustered
  - Scheduling on all nodes, but real execution only on active one



Our backup system

- Highly configurable
  - Target database, e.g.
    - Backup retention
    - Channel definitions
    - Archivelog deletion policy
  - Central LDAP repository, e.g.
    - TSM information – server and TDPO node
    - Hostname where RMAN is executed
    - Recovery catalog connection details
    - Delay for recovery of copy or archivelog deletions

Our backup system

- Simply extendable
  - Plugins
    - NAS snapshots
    - Read-only tablespaces backup
    - Full or selected schemas export
  - Templates with RMAN commands definition
    - Dynamically modified by backup system

```
$ cat backup_level_arch0.tpl
backup tag 'BR_TAG' archivelog all
format '%d_%T_%U_arch' delete all input;
```

Our backup system

- Easily manageable
  - Suspend backups
    - Specified or all databases
    - All databases using specified TSM server
  - Stop all ongoing backups or just for specific database
  - Suspend backup pre-checks
    - Target database up and running
    - Recovery catalog up and running
  - Flag `-simulate` to safely check effects of your command

Outline

- Who am I
- CERN and Oracle
- Backup system
- Recovery system
- Backups from standby databases
- Our strategy and recommendations
- Conclusions

Our recovery system

- Available as open source on Sourceforge
  - Developed by Ruben Gaspar Aparicio
- Used for
  - Validation of backup strategy
  - Sanity tests of backups sent to tape
- (Almost) full isolation from production database
  - Recovery catalog not touched
    - Restored controlfile is sufficient in almost all cases
  - User jobs dropped after recovery
  - Trimmed **`tnsnames.ora`** file

Our recovery system

- Easy to install on any machine
  - Recoveries scheduled using Cron
  - Three machines used at CERN
- Many features
  - Restore to filesystem or ASM
  - Configurable cleanup after recovery
  - Time interval for recovery
    - Hours since current time or specific date
  - Partial restore
    - For PITR – e.g. single schema
    - For VLDB – e.g. only RW part with some percentage of RO tablespaces

Our recovery system

- Even more features
  - Export after recovery
    - Full or list of schemas
    - Could be stored on tape or offsite file server
  - Dry run with script generation only
    - Target database connection used to gather needed data
  - On-the-fly generation of TDPO and recovery scripts based on
    - Central LDAP repository
    - Database metadata
  - Copy of successful recovery scripts kept

Outline

- Who am I
- CERN and Oracle
- Backup system
- Recovery system
- Backups from standby databases
- Our strategy and recommendations
- Conclusions

Backups from standby databases

- Only taken on physical standby are interchangeable
  - No performance impact on primary
    - BCT overhead could also be removed
- Possible even in 10g, but now
  - One recovery catalog could (must) be used
    - **`DB_UNIQUE_NAME`** different for each database
  - Fast incremental backups possible with Active Data Guard
  - Controlfile backups are also interchangeable
    - Only SPFILE backups are not allowed to be used on another database

New terms

- Association of backups
  - **SITE_KEY** set to target **DB_UNIQUE_NAME**
  - Unassociated files (null **SITE_KEY**) are treated like they are associated with current target database
- Changing association

```
RMAN> change ... reset db_unique_name;
RMAN> change ... reset db_unique_name to ...;
RMAN> catalog ...;
```

New terms

- Accessibility of backups
  - On disk – accessible only for associated database
  - On tape – accessible for all databases
- **`BACKUP`**, **`RESTORE`**, **`CROSSCHECK`** work on any accessible backups
- Default behaviour could be changed for specific session

```
RMAN> set backup files for device type disk to accessible;
```

  - Undocumented!

Database registration

- Only primary must be explicitly registered

```
RMAN> register database;
```

  - Standby will be automatically registered during first connection
  - It could also be done even before standby creation

```
RMAN> configure db_unique_name dgtest connect identifier
      'dgtest';
RMAN> configure db_unique_name dgtest_adg connect
      identifier 'dgtest_adg';
RMAN> list db_unique_name of database;
```

Database unregistration

- Unregistering all or specific standby database

```
RMAN> unregister database;
RMAN> unregister db_unique_name ...;
```

  - Backups still usable by other databases
  - According to docs, **INCLUDING BACKUPS** could be added to remove backups metadata, but...

```
RMAN-06014: command not implemented yet: unregister db &
bck
```

RMAN configuration

- Parameters to be set only on primary (inherited by standbys)
  - Retention policy
  - Database unique names

```
RMAN-05021: this configuration cannot be changed for a
BACKUP or STANDBY control file
```

- The others could (should) be set differently on each database, e.g.
  - Archivelog deletion policy
  - Snapshot controlfile name
  - Controlfile autobackup format

RMAN configuration

- The same controlfile autobackup format on primary and standby could lead to

```
RMAN-03009: failure of Control File and SPFILE Autobackup
command on ORA_SBT_TAPE_1 channel at 09/20/2012 11:47:17
ORA-19506: failed to create sequential file,
name="DGTEST_c-3733054949-20120920-ff", parms=""
ORA-27027: sbtremove2 returned error
ORA-19511: Error received from media manager layer, error
text:
ANU2614E Invalid sequence of function calls to Data
Protection for Oracle
```

RMAN configuration

- Moreover, sequence is set to max (-ff) – subsequent backups from the same day overwrite previous one (even if format is changed)

```
Handle: DGTEST_ADG_RAC7_c-3733054949-20120920-ff
Completion Time: 20-SEP-2012 11:50:06
Ckp SCN: 6303359618630, Media: 1877
```

- After new backup

```
Handle: DGTEST_ADG_RAC7_c-3733054949-20120920-ff
Completion Time: 20-SEP-2012 11:50:49
Ckp SCN: 6303359619993, Media: 1877
```

RMAN configuration

- Possible to set configuration for all databases while connected only to one

```
RMAN> configure ... for db_unique_name ...;
```

- But look out, at least with archivelog deletion policy
  - After setting for standby on primary, visible on standby, but not applied
  - Resetting on standby makes it effective
  - Workaround – configure all target databases one by one

Resynchronization

- Needed to ensure metadata consistency between recovery catalog and controlfile
  - Partial of full triggered automatically by most of RMAN commands
- Normal resynchronization
  - From controlfile to recovery catalog
- Reverse resynchronization
  - From recovery catalog to primary or standby controlfile

Resynchronization

- Resynchronizing all environment from one place

```
RMAN> resync catalog from db_unique_name all;
```

  - Even changes done without catalog will be refreshed
  - The same SYSDBA password in password files required
  - Target database connection must use TNS

```
RMAN-03002: failure of resync from db_unique_name command
at 08/30/2012 15:09:59
ORA-17629: Cannot connect to the remote database server
```

Resynchronization

- Sensitive for any misconfigurations

```
RMAN-01005: ORA-20079: full resync from primary database
is not done
```

- RMAN debug showed that error appeared while resynchronizing tempfiles information
  - Different number of tempfiles on both databases
    - Tempfiles not created on standby after adding on primary (`standby_file_management` set to auto) – no redo is generated
    - Manual creation on standby solved the problem
    - But cannot be reproduced on test

Outline

- Who am I
- CERN and Oracle
- Backup system
- Recovery system
- Backups from standby databases
- Our strategy and recommendations
- Conclusions

Our backup strategy

- Datafile backups on standby
- Archivelog backups on primary
  - To avoid delays if communication with standby is lost
  - To have SPFILE and controlfile backups
    - Controlfile backups interchangeable, but additional work needed after restore (e.g. drop of standby logs)
- We could easily switch backups back to primary in case of problems

Our backup strategy

- Things to be kept in mind
  - Checking apply lag on standby where backups are taken
    - Views: **`V$ARCHIVED_LOG, V$DATAGUARD_STATS, V$RECOVERY_PROGRESS`**
    - **`STANDBY_MAX_DATA_DELAY`** session parameter
  - Resolving gaps on standby using incremental backups from primary
    - No backups ready to use
    - **`BACKUP ... FROM SCN`** longer without BCT

Archivelog deletion policies

- Primary

```
RMAN> configure archivelog deletion policy to applied on
      all standby backed up 1 times to device type sbt;
```

- Standby

```
RMAN> configure archivelog deletion policy to applied on
      all standby;
```

- For databases with downstream capture
  - **SHIPPED** instead of **APPLIED**

Automatic recoveries

- Recovery catalog connection useful
  - Restored controlfile not aware of all backups
- Without recovery catalog connection
  - Catalog backups after controlfile restore
    - Possible even for backup pieces on tape

```
RMAN> catalog device type 'sbt_tape'
      backuppiece '<handle>';
```

  - Too many changes required and error-prone
  - Better to test recovery as close as possible to real exercise

Other recommendations

- Fast incremental backups
  - BCT not used without database restart after enabling
    - Even if enabled in mount state
    - Check **`USED_CHANGED_TRACKING`** column in **`V$BACKUP_DATAFILE`**
  - Apply fix for bug 12312133 – standby crashes during RMAN incremental backup with BCT enabled
    - **`ORA-600 [krcccb_busy]/[krccckp_scn]`**
  - Reminder
    - Active Data Guard option needed!

Other recommendations

- Minor problem during backup from standby
  - **`TOTALWORK`** value doubled in **`V$SESSION_LONGOPS`**
- Good to be aware of
  - Current log not archived on standby during backup
    - **`ARCHIVE_LAG_TARGET`** could be used to force log switch on primary for low redo generation periods
- Be careful with plugged in tablespaces
  - Bug 13000553 – null datafile name on standby after adding new datafile to plugged in tablespace on primary
    - Causes resync to crash

Other recommendations

- Backups on standby in RAC
  - All instances used for channel allocation in the same state

```
RMAN-03009: failure of backup command on ORA_SBT_TAPE_1
channel at 06/27/2012 09:00:32
ORA-01138: database must either be open in this instance
or not at all continuing other job steps, job failed will
not be re-run
```

- Consistent naming convention is important, e.g.
  - TNS names used in channels definition on each database

Outline

- Who am I
- CERN and Oracle
- Backup system
- Recovery system
- Backups from standby databases
- Our strategy and recommendations
- Conclusions

Conclusions

- Backups from standby work well
  - If properly configured
- Impact on whole environment and procedures
  - Management a little bit more complicated
  - **`DB_UNIQUE_NAME`** change could entail
    - CRS reconfiguration
    - Data and other files location change when using OMF
- Reconfiguration needed after switchover or failover
  - Archivelog deletion policy for example

Questions?

Thank you!

www.cern.ch