October 9th 2006, Maria Girone

ORACLE database backup and recovery policy at Tier 0

Introduction

The physics database services at CERN Tier 0 implement a backup strategy based on ORACLE RMAN 10g. This strategy allows for backups to tape but also to disk (flash backup), thus reducing substantially the recovery time for many common recovery scenarios.

This note explains the types of backup, retention policies and recovery strategies which are adopted by the IT-PSS group, with the aim of reviewing them with the experiments and the grid community prior to the LHC start-up.

Types of backup

The **backup to tape** uses the IBM Tivoli Storage Manager (TSM), supported by the IT-FIO group. Oracle RMAN has dedicated drivers to connect to TSM and manage backups stored in tapes in a special backup format. There are different kinds of backups done to tape.

- Full backup (level 0) - a complete copy of the database and control files
- Differential backup (level 1) - copy of blocks changed since latest level 0 or level 1 backup
- Cumulative backup (level 1) - copy of blocks changed since latest level 0 backup
- Archive logs - copy of the archive logs (files containing the log of the operations done in the database)

For the **backup on disk** first ORACLE RMAN copies all data files into the flash recovery area (typically on a different storage array than the data files), then all the subsequent backups are differential. We use a technique called "Incrementally Updated Backups" to maintain this type of backup.

The ORACLE's block change tracking feature is used to significantly reduce the latency and weight of the incremental DB backups (only changed DB blocks are read during a backup with this optimization).

Backup retention policy proposal

The **backup on tape retention policy** is set to 31 days. This guarantees that a copy of the database can be recovered in a time window of 31 days. In addition, we propose that a full backup is systematically performed and kept before any ORACLE software upgrade.

The schedule for the **backup to tape** is as follows

- **Full** - every 2 weeks
- **Incremental** (differential or cumulative) – daily
- **Archive logs** - every 30 minutes

The backup on disk retention is set to 2 days and allows for database recoveries in that time frame.

The schedule for the **backup on disk** is as follows

- **Full** - at database creation
- **Incremental** – daily

Recovery strategies

A dedicated system for recoveries is available. The system allows for periodic test recoveries, point-in-time recoveries (from disk within a two days latency) and tape-backup based disaster recoveries.

At present, a typical recovery takes 30MB/sec, assuming that a channel for tape reading is available. In addition, a overhead of a couple of hours is estimated for the DBA to understand and analyse the problem.

To give an idea, the recovery of a 300GB database, would take a total of 5 hours (3 for reading from tape plus 2 hours overhead).