



# Evolution of ATLAS Databases

Dario Barberis

Genoa University/INFN



# Database applications

- Online (ATONR):
  - Detector Control System (DCS/PVSS)
  - Trigger and detector configuration
  - Conditions for data-taking (COOL)
  - SFO-Tier0 handshake
- ADC (ADCR):
  - DDM/DQ2
  - ProdSys/Panda
- Grid services (LCGR):
  - LFC, FTS
- Offline (ATLR):
  - Detector Control System (DCS/PVSS)
  - Trigger and detector configuration
  - Conditions for offline processing (COOL)
  - Detector geometry
  - Tier-0 and CAF job control
  - Metadata:
    - AMI
    - COMA
    - TAG catalogue
- "Archive" (but not quite) (ATLARC):
  - TAG database
  - Real archive of some old data



# Online databases

- Current detector data-taking needs are well satisfied by the current setup (ATONR + CORAL server)
- Requirements are in terms of data volume rather than functionality
  - Mainly DCS/PVSS data
    - Sliding window algorithm already implemented keeping 12-18 months
      - All data are in ATR anyway
  - Talk later this morning (Stefan Schlenker)
- Further optimisation may be needed for Oracle 11g
  - Talk later this morning (Gancho Dimitrov)



# Offline databases

- Geometry and Trigger DB now also distributed with Frontier
  - Remove the need for the hybrid SQLite/Frontier DB access for analysis
  - Keep SQLite only for production tasks
- Continue evolving COOL in functionality and performance
- Optimisation of existing database applications using Oracle 11g
  - Transition planned for next Winter
  - Talk later this morning (Gancho Dimitrov)



# ADC databases

- Had to split them off the ATR cluster to isolate sources of high load and avoid interferences with COOL and other applications (Tier-0)
- ADC launched a number of R&D projects earlier this year, including:  
**Study of the usability and performance of NoSQL databases for ADC applications**
- Two talks later today:
  - ProdSys/Panda (Maxim Potekhin)
  - DDM/DQ2 (Vincent Garonne)



# TAG database

- TAGs — event-level metadata records — are by volume one of the most demanding applications
- ~2 kB/event (including indices) x 400 Hz x multiple reprocessing passes
  - And matching amounts of simulation
  - Scale is in the 10s of terabytes
  - Scalability is an increasing challenge
- Databases hosted at multiple sites (Switzerland, UK, Germany, Spain, Canada)
  - No site hosts everything
- ATLAS TAGs are hosted both in POOL/ROOT files and in Oracle
- Underlying technology — LCG POOL Collections — is designed to be DB technology independent
  - Original implementations included MySQL
    - Dropped because this was not a CERN-IT supported technology
- Obvious candidate for technology alternatives
  - Especially those that support efficient parallel read-only shared-nothing (or shared-little — nothing is ever quite so simple) access patterns
  - Column-wise databases also of potential interest
    - Demo project (Petaminer) used TAGs with ROOT as column-wise storage backend to MySQL
  - No real effort has been available for such explorations, but the situation may be changing
- ATLAS is interested in collaboratively exploring both:
  - NoSQL implementations to support event selection services (Luc Goossens)
  - Optimizations that may be possible within Oracle 11g (Gancho Dimitrov)



# Outlook

- We have a working system
  - Sure, it can be improved
  - Also the amount of data is steadily increasing
- We work towards two parallel aims:
  - Improvements and optimisations of the current applications
  - R&D projects following the DB technology evolution
- We won't disrupt data taking and data analysis
  - All major component replacements can only take place during the LHC downtimes and after adequate load testing