ATLAS Databases and Related Tools

Dario Barberis

Genoa University/INFN

Outline

- This talk:
  - ATLAS database group organization
  - Recap of main database applications
  - Database distribution and access strategy
  - Split of "ADC" databases
- Next talk (Gancho Dimitrov & Florbela Viegas):
  - Size of databases
  - Load on servers
  - TAG database
- This afternoon (John DeStefano):
  - Frontier servers: ATLAS experience and setup optimization

Outline

- This talk:
  - ATLAS database group organization
  - Recap of main database applications
  - Database distribution and access strategy
  - Split of "ADC" databases
- Next talk (Gancho Dimitrov & Florbela Viegas):
  - Size of databases
  - Load on servers
  - TAG database
- This afternoon (John DeStefano):
  - Frontier servers: ATLAS experience and setup optimization

Offline Database Coordination: Dario Barberis

Database Software: Elizabeth Gallas

- COOL related tools:
R. Hawkings ➔ M. Borodin
- Liaison to LCG-AA:
E. Gallas, D. Barberis
- Conditions data tagging:
P. Laycock ➔ M. Plamondon
- Geometry DB: V. Tsulaia
- Trigger DB: J. Stelzer

Liaison to Online Databases:

Rainer Bartoldus

Database Operations: Dario Barberis

- DB administration, monitoring and optimisation:
G. Dimitrov, F. Viegas
- Liaison to WLCG Service Coordination:
  - Coordination: D. Barberis
  - Technical (including 3D):
G. Dimitrov, F. Viegas
- DB releases: V. Tsulaia
- ADC servers liaisons:
F. Donno, S. Baranov
- Frontier/Squid support:
J. DeStefano, R. Walker
- Conditions data files:
A. De Salvo

Metadata and TAGs: David Malon

- TAG infrastructure and contents:
D. Malon, J. Cranshaw, T. Donszelmann
- TAG uploading tools and operations:
E. Vinek, F. Viegas
- Conditions metadata:
E. Gallas
- AMI: S.Albrand

June 2010 – February 2011

Overview – Oracle usage in ATLAS

Oracle is used extensively: every stage of data taking & analysis

- Configuration
  - Detector Description – Geometry
  - OKS – Configuration databases for the TDAQ
  - Trigger – Trigger Configuration (online and simulation)
  - PVSS – Detector Control System (DCS) Configuration & Monitoring
- File and Job management
  - T0 – Tier-0 processing
  - DQ2/DDM – distributed file and dataset management
  - Dashboard – monitor jobs and data movement on the ATLAS grid
  - PanDA – workload management: production & distributed analysis

  } "ADC" Databases
- Dataset selection catalogue
  - AMI (dataset selection catalogue)
- Conditions data (non-event data for offline analysis)
  - Conditions Database in Oracle
  - [POOL files in DDM (referenced from the Conditions DB) – not Oracle data!]
- Event summary - event-level metadata
  - TAGs – ease selection of and navigation to events of interest

Database distribution and access (1)

- Only Geometry, Trigger and Conditions data are accessed by ATLAS jobs running on the Grid
  - The other databases are used
    - Only by jobs running at CERN
    - Through special web interfaces (AMI, TAGs)
- Trigger and Conditions DB are distributed from CERN to the Tier-1s with Oracle Streams for analysis access
- The Geometry database (including only the base geometry, not alignments and calibrations that are in the Conditions DB) is distributed to each Grid site in the DB release
- Conditions data files are distributed to each Grid site using DDM
- Special Conditions Database Releases are created for reprocessing, containing:
  - An SQLite extraction of the COOL data for the run range to reprocess
  - The conditions data files for that run range
- DB and CDB releases are distributed on the Grid, downloaded on the worker node by each Grid job and accessed from the local disk

Database distribution and access (2)

- We want to prevent any interference between scheduled production activities and chaotic user analysis
  - For the benefit of both sides
- Simulation jobs and analysis jobs running on simulated data only need the DB release with which the simulation was produced (or any more recent one)
- Reprocessing jobs should need only the Conditions DB release
  - This statement was found to be false in the current reprocessing campaign
    - Found several direct accesses to Oracle from Python code during job configuration (magnetic field and some detector status parameters)
    - Discussions in progress on how to solve this problem "forever"
- Analysis jobs access conditions data through local Squids and Frontier servers placed at Tier-1s and CERN
  - Analysis runs mostly at Tier-2s and Tier-3s
    - Local caching of data is useful to reduce access times
  - Load on the Oracle servers is reduced also for jobs running at Tier-1s

Database distribution and access (3)

- ATLAS has 10 Tier-1s, therefore Oracle streams were set up initially for all of them
- Early in 2009 we asked the sites that wanted to build experience with Frontier servers to go ahead and install it
  - At that time we had no idea of the load on Oracle and Frontier servers due to analysis jobs but we wanted to be on the safe side
- We now have 7 Frontier servers (CERN + 6 Tier-1s) and measure light loads
  - As a side-effect of the Frontier deployment, it turned out that Oracle databases without Frontier server should not be accessed by any job
    - Apart from the problem explained in the previous slide
- Discussion active in ATLAS Computing Model and Resource groups (CREM+ICB) on the consolidation of distributed database services:
  - How many Oracle/Frontier servers do we need, and which size, for analysis tasks?
  - Do we want to use Oracle/Frontier instead of CDB releases also for reprocessing?
    - And in this case how much more load shall we have?

Statement from ATLAS ICB

- The ICB took note of the two options presented for Frontier deployment in ATLAS :
  - a) Continue with adiabatic approach used so far. Let sites willing to support Oracle/Frontier servers continue to do so but do not force them to have a very fail-proof service, as the global ATLAS system is robust through redundancy.
  - b) Have 2 large and well-maintained sites (CERN and BNL would be the natural choices for ATLAS) and discontinue the other sites.
- Moving to Squid to all Tier-1 while they have been told years ago that ORACLE deployment is crucial for ATLAS conditions data is quite a change in ATLAS DB policy. Many sites have invested a lot both in hardware and manpower.
- Many questions were raised and some quantitative numbers and scenarios will be presented by Dario at the next ICB (early December).
  - What is the expected load on 2 Frontier servers?
  - What is the viability of a system based on 2 servers only?
- Option b) implies that all Tier-1s would have to deploy a Squid server which is currently not the case.
- While option b) would greatly simplify the operation of DBs within ATLAS, it has some implications that need to be evaluated with care. A statement cannot be issued on such a short time scale.
- Meanwhile all Tier-1 representatives should inform sites about this possible change of ATLAS DB strategy as it may have some financial (and human) consequences.

Split of ADC databases

- DDM and PanDA are heavy users of central databases
  - They can interfere with other DB activities (COOL etc) and with each other
- We wish to separate ADC databases (DDM, PanDA and a few other bits) from the rest of central database services at the end of this year
- Testing of a separate instance is starting now, aiming at a production service mid-January
  - Needs 2-day interruption of all ADC services

Outline

- This talk:
  - ATLAS database group organization
  - Recap of main database applications
  - Database distribution and access strategy
  - Split of "ADC" databases
- Next talk (Gancho Dimitrov & Florbela Viegas):
  - Size of databases
  - Load on servers
  - TAG database
- This afternoon (John DeStefano):
  - Frontier servers: ATLAS experience and setup optimization