ATLAS calibration/alignment at Tier-2 centres

Roger Jones, Richard Hawkings

LCG workshop, 13/6/06

- Calibration/alignment plans and the role of Tier-2 centres
  - One role of Tier-2 in ATLAS
  - Calibration alignment model
  - Conditions database
    - Replication and implications for Tier-2 centres
  - Calibration/alignment challenge
  - Calibration centres for muon calibration - a special case
  - Concluding remarks
- NB: At present, many things are not yet clear, and require much more real-world experience …

Role of Tier-2s

- Role of Tier-2s according to the ATLAS computing model
  - Simulation production
  - Physics group and end-user analysis
  - Code development
  - Calibration/alignment for 'local-interest' subdetectors
    - Institutes with responsibility for calibration of a particular subdetector expect to do their calibration processing at nearby Tier-2 centres
- Data requirements to support this
  - Data samples to host: TAG and AOD, some samples of ESD and RAW data for development, ESD and (possibly) RAW data for calibration samples
    - Access to distributed data management system, to manage local storage elements
  - Conditions data for simulation production (small), and for analysis
    - Analysis will likely require access to a limited subset of the full conditions data
  - Conditions data for calibration tasks
    - Larger amounts of conditions data, but for particular subdetectors / data periods
- So far, Tier-2 concentrated on simulation - very limited conditions data needs

Calibration / alignment model

- First pass calibration done at CERN (except muon stream, see later)
  - In 24 hours after end of fill, process and analyse calibration streams, produce and verify first pass alignment constants…
    - Processing resources are part of CERN Tier-0/CAF
    - Calibration will also depend on previous calibration - amount of ‘per run’ recalibration will not be known until experience with real data is gained
  - … Prompt reconstruction of physics data, distribution to Tier-1s, Tier-2s, etc.
- Then, study pass 1 data, prepare new calibrations ready for reprocessing
  - ATLAS expects to reprocess whole data sample 1-2 times per year, at Tier-1s
  - Calibration will be based on detailed analysis of AOD, ESD and some RAW data
    - Processing done primarily at Tier-2 and Tier-1 centres
  - Calibrations will be uploaded from originating sites to CERN central databases
    - Probably file-based uploading - see later
  - New calibrations distributed to Tier-1 centres for subsequent raw data reprocessing
- Once raw data is reprocessed and distributed, process can be repeated

Conditions data model

- ATLAS conditions database contains all non-event data needed for simulation, reconstruction and analysis
  - Calibration/alignment data, also DCS (slow controls) data, subdetector and trigger configuration, monitoring, …
    - Key concept is data stored by ‘interval of validity’ (IOV) - run/event or timestamp
    - Some meta-data may be stored elsewhere (luminosity blocks, run level information)
- Several technologies employed:
  - Relational databases: COOL for IOVs and some payload data, other relational database tables referenced by COOL
    - COOL databases can be stored in Oracle, MySQL DBs, or SQLite file-based DBs
    - Accessed by ‘CORAL’ software (common database backend-independent software layer) - CORAL applications are independent of underlying database
    - Mixing technologies an important part of database distribution strategy
  - File based data (persistified calibration objects) - stored in files, indexed / referenced by COOL
    - File based data will be organised into datasets and handled using DDM (same system as used for event data)

Relational database data

- Replication of relational-database based conditions data (COOL and others):
  - Tier-0 hosts master copy of all data in Oracle (O(1 TB/year))
  - Oracle Streams technology used to replicate data to Oracle servers at Tier-1
    - Native Oracle technology, for keeping a replica in sync - 'duplicates' all database writes in slave servers by extracting data from master server's change logs
    - Works equally well for COOL and other relational database data (application-neutral)
- All Tier-1 sites should have local access to conditions data from Oracle
  - Performant-enough access for reconstruction of full RAW data samples
- Options for Tier-2s:
  - Access Oracle server of nearest Tier-1
    - OK for small scale access, limited by network latencies and load on Tier-1 server
  - Extract needed COOL data into an SQLite file (tools exist)
    - A 'one shot' replication, only practical for a subset of data (e.g. for simulation use case)
  - Maintain a 'live' database copy in MySQL - run a local MySQL condDB server
    - Tool being developed to synchronise two COOL databases and copy recent updates
    - Will probably be needed for sites doing significant calibration work
    - Again, only practical for subsets of the full conditions database

File-based conditions data

- Some conditions data stored in files:
  - Large calibration data objects, stored using POOL technology (as event data)
  - Other types of data, e.g. files of monitoring histograms
- Organise into conditions datasets using standard ATLAS DDM tools
  - Expect O(100 GB/year) of calibration data - small compared to event data
    - Perhaps more for histograms/monitoring data
  - Reconstruction/analysis jobs will require local access to specified datasets
    - Stored on DDM-managed local storage, as for event data being processed, or even downloaded to worker node
    - DDM / DQ2 instance to manage the storage and maintain catalogues could be at Tier-2, or at Tier-1
    - … but Tier-2 sites must be ‘DDM-aware’
  - End users will want to download specific datasets, e.g. histogram sets for their subdetector, download locally to Tier-2 or even to their laptops
    - Again using DDM end-user tools - retrieve datasets from local Tier-2 or nearest Tier-1

Frontier

- Frontier is an interesting alternative to traditional database replication
  - A fourth (read-only) technology for CORAL - database access requests are translated to http page requests
    - These are served by a Tomcat web server sitting in front of a relational database - server translates page request back to SQL and queries real relational database
    - Server returns result as web page (can be gzipped to avoid XML space overheads)
    - Frontier client (CORAL) translates web page request back to SQL result for client program (e.g COOL)
  - Putting a web proxy cache (squid) between client and server allows queries to be cached
    - When many clients make the same query (= request same web page), only the first one will go all the way to the database, rest will be satisfied from squid cache
    - Reduces queries on the server, and network traffic
  - In a distributed environment, could have e.g. squid caches at Tier-1s or even at local Tier-2s, to satisfy most requests as locally and as quickly as possible
- First steps in trying this out for ATLAS conditions data (CMS more advanced)
  - Many questions (e.g. stale caches), but could be an attractive alternative for Tier-2s - deploy a squid cache instead of a MySQL replica

Calibration data challenge

- So far in ATLAS, Tier-2s have only really done simulation/reconstruction
  - With static replicas of conditions data in SQLite files, or preloaded MySQL replicas - required conditions data already known in advance
- ATLAS calibration data challenge (late 2006) will change this
  - Reconstruct misaligned/miscalibrated data, derive calibrations, rereconstruct and iterate - as close as possible to real data
  - Will require 'live' replication of new data out to Tier-1/2 centres
- Technologies to be used @ Tier-2
  - Will need COOL replication either by local MySQL replicas, or via Frontier
    - Currently just starting on ATLAS tests of Frontier - need to get experience
    - Decision in a few months on what to use for calibration data challenge
    - Frontier is also of interest in online environment (database replication for trigger farm)
  - Will definitely need DDM replication of new conditions datasets (sites subscribe to evolving datasets)
  - External sites will submit updates as COOL SQLite files to be merged into central CERN Oracle databases

Muon calibration use case

- A few Tier-2 sites designated as muon ‘calibration centres’
  - Receive special stream of muon data extracted from level 2 trigger: ~ 100 GB/day
    - Probably transferred via Tier-1 for tape backup
  - Process this locally at Tier-2 on a farm of O(100 machines)
  - Store intermediate results in a local Oracle-based calibration database, which is replicated to CERN using Oracle streams replication
  - Calibration results (to be used in prompt reconstruction) will be derived from this data and entered into COOL in the usual way
    - Time critical operation - prompt reconstruction needs these results in < 24 hours
- Goes beyond the calibration requirements of a standard Tier-2 site
  - Need for dedicated local Oracle database expertise and higher ‘quality of service’ and response time for problems

Concluding remarks

- Little experience of calibration/alignment activities so far, especially in an organised production environment
  - Tier-2 have concentrated on simulation/reconstruction of simulated data
- Some requirements on Tier-2s are clear:
  - Need for CPU resources for calibration/alignment
  - Access to event and conditions datasets using ATLAS DDM tools
  - Access to local SQLite-based database replicas of parts of conditions database
- Others are not so clear:
  - Need for dedicated MySQL service for live conditions data ?
  - Need for Froniter squid caches ?
  - … will become clearer in next few months and from experience with calibration data challenge
- Probably will not have ‘standard’ requirements for a Tier-2
  - A lot will depend on what the users at that Tier-2 want to do (simulation,analysis, calibration,..)