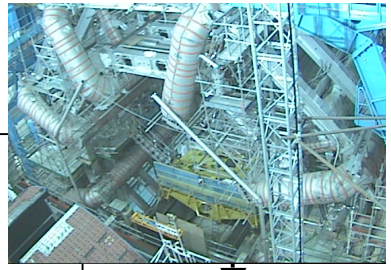
ATLAS & 3D

Dario Barberis

(CERN & Genoa University)

Calib/Align in the Data Flow



L1

L2

Event Filter

RAW

GeomDB

CondDB

File Catalogue

Reconstruction

ESD

AOD Production

AOD

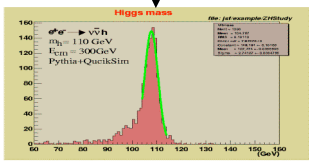
TAG

Future

DPD

Physics Analysis

Event Selection



Event Generation

SIM

Detector Simulation

Pileup

Digitization

Trigger Simulation

SIM

Artwork by Stefan Stonjek

WLCG 3D Replication Tests

Once Oracle streams are established between Tier-0 and Tier-1:

- ATLAS can replicate the 'small' databases such as the Geometry DB
- Start to look at COOL replication, which will involve more data
  - a lot depends on the subdetectors:
    - how they implement their conditions data

Geometry DB Replication

- As COOL is not deployed yet, the Geometry DB database application creates the main database server load in the ATLAS distributed production on the Grids
- Recent doubling of the ATLAS Oracle RAC capacities at Tier-0 and the POOL SQLite database replication technique resolved the Geometry DB workflow bottleneck in distributed production
  - The Geometry DB connect failures are now reduced to the level of 0.2%
- Currently we do have enough database server capacities to service the anticipated increase in distributed production levels (by a factor of 2-3) through the end of this year
- Large scale production experience with SQLite replication provided important feedback to the POOL SQLite developers
  - Support for the SQLite connection time-out is required (for nfs)

COOL Status in Offline

- COOL software fully integrated into Athena from Summer 2005
  - Integration consistent with old Lisbon MySQL implementation,
    - Subdetector clients can migrate adiabatically

Data model:

- small data payloads directly 'inline' in COOL
- for large payloads, COOL stores:
  - POOL-token refs to data in POOL files or
  - POOL object-relational (only in prototype form now)
- Functionality and performance testing ongoing in both online and offline environments
- Currently small amounts of conditions data in COOL
- Production: uses Lisbon ConditionsDB (MySQL)
  - Primarily conditions data from 2004 Combined Test Beam (CTB)
  - Full migration to COOL is underway:
    - CTB data being moved to phase out Lisbon by Summer 06
- Commissioning: conditions data from subdetectors
  - Already started
  - Going to COOL not Lisbon MySQL DB
- COOL usage grows in simulation and reconstruction now :
  - COOL data volumes and use will increase significantly

Deployment Next Steps

- Need in production:
  - 'Static' replication of COOL conditions data from February 2006
  - Dynamic replication from June 2006
- Customers:
  - Ongoing support for commissioning data
  - Major simulation production from April 2006
  - Major reconstruction production from July 2006
    - Including calibration
- Closed loop cycles:
  - reconstruct, improve calibration, re-reconstruct

Replication Strategies

- COOL API-level copy from Oracle->SQLite
  - for static replicas - works
- Explore Oracle streams for COOL Tier-0 -> Tier-1
  - as soon as 3D ready
- Need dynamic COOL API-level copy (updating replicas)
  - from June 2006
- Evaluate FroNTier-based replication
  - ATLAS will start with using the FroNTier setup of the 3D project
  - We will start with tests of the Geometry DB replication
  - The performance evaluation milestone for COOL replication is scheduled before June
  - To do significant scalability tests beyond the capabilities of the shared service we plan to deploy our own service
- Experience with different solutions in first half of 2006 will guide replication choices at various Tiers for calibration-closed loop cycles in second half of 2006

Cache Consistency

- We appreciate CERN studies of the cache consistency
- We plan to do practical testing with Frontier to determine what our FroNTier caching methods will be with COOL
- Payloads are cached and the IOV lookups are not:
  - The payload lookups will carry an identifier for the particular version/object being retrieved
  - There are use cases when the IOV lookups are not worth caching because they will rarely be repeated
  - Or the IOV lookups could only be cached in certain circumstances
    - e.g. in HLT caching IOV lookups could be very useful
- In that way we would avoid the cache consistency problems
  - as long as we do not reuse these identifiers
    - which we do not in the COOL model

Ideas for Cache Managing

- To solve the FroNTier cache invalidation issue we will manage cache consistency in a 'controlled' production environment with well-defined 'database content' releases:
  - For example, we wouldn't launch the distributed reconstruction of a particular dataset (e.g. a particular period of data, or a set of MC samples) until the appropriate conditions DB data are available
  - Before launching, we could invalidate ALL of our caches, and let them repopulate themselves as data is requested
- By construction, we would not then be expecting this data to change 'under our feet' during a reconstruction pass
- The same applies in the HLT environment, where we could invalidate caches before each run start
- But this obviously does not apply in chaotic end user environments, and for individuals developing new calibrations

These are just ideas at this stage - they will need to be tested

Conclusions

2006 in ATLAS is a year of the Computing System Commissioning (CSC)

- COOL commissioning is a part of CSC
- ATLAS Tier-0 tests (part of WLCG SC4) include
  - Growing ATLAS database applications: (1) Realistic Conditions DB: Geometry DB, Magnetic Field, etc; (2) Calibration and alignment; (3) TAG database
  - A production quality system using 3D tools is scheduled for September
    - By then, we need to have made choices regarding: Frontier vs replicas, the proper use of Oracle streams, caching strategies etc.
    - We are interested in testing 3D tools as soon as they are ready
- Database project milestones after completion of Tier-0 tests in September:
  - Database infrastructure for TDAQ scalability test
    - scheduled for October-November
  - Deployment of database infrastructure to support large scale distributed production (following FroNTier evaluation in ATLAS)