ATLAS : File and Dataset Metadata Collection and Use

S Albrand[1], J Fulachier[1], E J Gallas[2], F Lambert[1]

on behalf of the **ATLAS Collaboration**

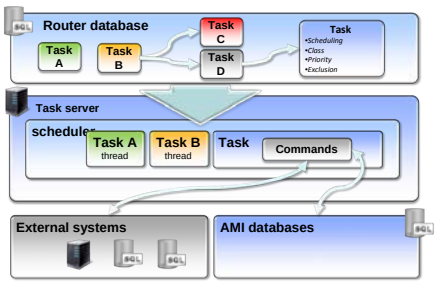
1. Introduction

- The ATLAS dataset search catalogs (AMI) combines information about ATLAS datasets and files from many sources and derives quantities.
  - nFiles
  - nEvents
  - Production Status
  - Averages of cross-section over EVGEN datasets, propagated along the production chain.

The data sources are:

- **ATLAS Tier 0 database.** Tier 0 is the application which forms the files of data coming from the Detector Acquisition into datasets of "real data".
- **ATLAS Production system and Production Task Request. (ProdSys).** Physicists make requests to the production system for simulation tasks or for reconstruction of real data
- **ATLAS Distributed Data Management. (DDM).** All datasets are registered in the DDM system, which manages the physical location and existence of datasets.
- **ATLAS Conditions Metadata (COMA**). A relational view of the ATLAS Conditions Database which stores all detector conditions indexed by run and luminosity block, or by interval of validity.

2. The AMI Framework Task Server



- Data is entered by a set of specialized tasks controlled by the task server.
- Information is available from different sources in a chaotic way.
- The task server imposes time-sharing.
  - One cannot allow a peak in production tasks finishing to allow a backlog of input from Tier0 to develop.
  - Little and very often is best.
- Some tasks also must store a "stop point" which is usually the data source time stamp of the time when the last AMI read was successful.

The next time the task runs it will start from this time stamp.

3. Tier 0 : Semaphore Mechanism

- T0M (Tier 0 Management) determines which datasets go to AMI.
- AMI looks every 60 seconds.
- T0M raises a "READY" flag on a dataset .
- AMI reads file and dataset information.
- AMI flips "READY" to "DONE" (write privileges on flag column in Tier0 DB)
- Approx. 9Hz. Insertion rate; much faster for datasets with lots of files.

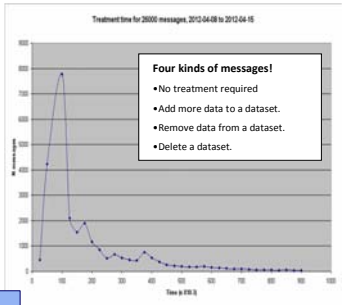
4. Production System : Timestamp mechanism

- Read everything greater than lastUpdateTime and TaskNumber.
  - Reader (AMI) must decide what is relevant.
  - "Secure programming" = be ready for surprises.
- Read in 20 task bites – gear change if backlog > 1000 tasks.
- Looks every 60 seconds. Copies dataset, file, and metadata output from jobs
- Treats about 7 jobs a second on average. Each job has several files.
- Multithreaded. (8 tasks treated in parallel)
- Profiling has revealed we can still gain some time.

5. DDM : Publish/Subscribe

- Registration & Deletion of data using Active MQ/Stomp "Publish/subscribe" protocol.
- Very reliable and fast.

Speed of treating Active MQ messages



6. COMA : Symbiosis

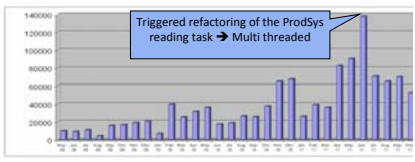
- AMI and COMA "think" they are part of the same application. Parts of COMA were rendered "AMI Compliant"
- COMA has benefited from the AMI infrastructure, in particular pyAMI, the web service client. AMI writes some aggregated quantities in the COMA DB.
- AMI has benefited from the access to Conditions Data.

Run and period info from **COMA**.
Other info from **AMI**.

7. Is AMI loading scalable?

- Insertion in AMI is longer than pure SQL insert operations on a database.
  - Many coherence checks,
  - derivation of quantities etc.
- Although almost all the time we have spare capacity we have observed backlogs from time to time – usually with massive numbers of finished production jobs arriving within a short period.
- Some obvious optimisations still not attempted.
- Scalable in medium term.

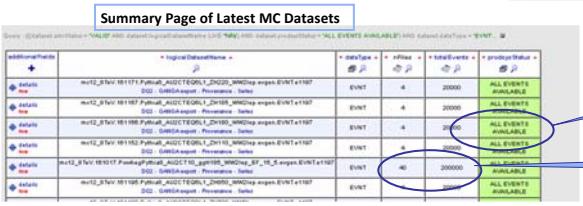
Datasets registered in AMI per month since May 2009



8. Some Examples of Derived Quantities

- Complete dataset provenance. (T0 & ProdDB)
- Number of files and events available in the dataset is updated every time fresh information arrives. (T0, ProdDB, DDM)
- Production Status. (T0, ProdDB, DDM)
- Average, min and max cross section recorded for the simulated, transported down the production chain. (ProdDB)
- Lost luminosity blocks in reprocessed data. (T0, ProdDB, DDM)
- Run period reprocessing errors (ProdDB & COMA)
- Datasets in run periods. (COMA)

Summary Page of Latest MC Datasets



Status combines info from ProdDB & DDM

nEvents & nFiles recalculated as necessary

Affiliations

(1) Laboratoire de Physique Subatomique et Corpusculaire, Université Joseph Fourier Grenoble 1, CNRS/IN2P3, INPG, 53 avenue des Martyrs, 38026, Grenoble, FRANCE

(2) Department of Physics, Oxford University, Denys Wilkinson Building, Keble Road, Oxford OX1 3RH, UNITED KINGDOM