



GridPP

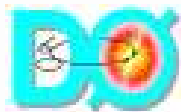
UK Computing for Particle Physics



Metadata



Paul Millar



UNIVERSITY
of
GLASGOW



- Introduction,
- Group news,
- Metadata Workshop,
- Documentation,
- Metadata interface,
- AMI interface,
- AMI-GANGA,
- PTF/web-services,
- The future.





The Metadata group exists to **examine commonalities** across all the High Energy Physics experiments' metadata handling, at the technology, interface and schema level.

The aim of the group is to ensure metadata services are deployed as much as possible using **web services** and other grid standards.





Facilitate good communication:

- Monthly Video/Phone meetings,
- Have mailing list
- Write documentation of current situation
- Organise, or help organise workshops

Provide resources:

- **wiki** - <http://www.physics.gla.ac.uk/metadata/>
- CVS server
- Mailing list
- Development cluster (WAN testing)

Work on specific areas:

- ATLAS-gLite, ATLAS-Ganga, web-services, ...





GridPP

UK Computing for Particle Physics

News

- **Mòrag Burgon-Lyon left in June**
 - Taking on the reigns of the group
- **Steven Hanlon will be leaving in August,**
- **We hope to have the two posts filled by August.**



UNIVERSITY
of
GLASGOW



GridPP
UK Computing for Particle Physics

Grenoble Workshop



- Experiments were invited to a workshop dedicated to metadata, from 25 to 28 April 2005.

- Programme organised by **Mòrag Burgon-Lyon**.
- Everything else organised by **Solvieg Albrand** (with help from **Jocelyne Riffault** and **Anne Wolfers**). **Many Thanks!**



UNIVERSITY
of
GLASGOW



Workshop outcomes



- 20 people attended.
- Representatives from ATLAS, LCG, CMS, BaBar, LHCb, CDF.
- 24 talks presented
 - available on-line
- 6 discussion sessions
 - minutes available on-line
- Discussion sessions resulted in spin-off collaborations
- Improved cross-experiment communication





- Documents the different metadata schemata
- Drew information from:
 - **ATLAS** (AMI, DC2) schema
 - **BaBar** schema
 - **CDF & D0**
 - **CMS** TMDB and RefDB
 - **LHCb** schema (Bookkeeping)
- Release 1.01 now available
- Feedback welcome!





Metadata interface

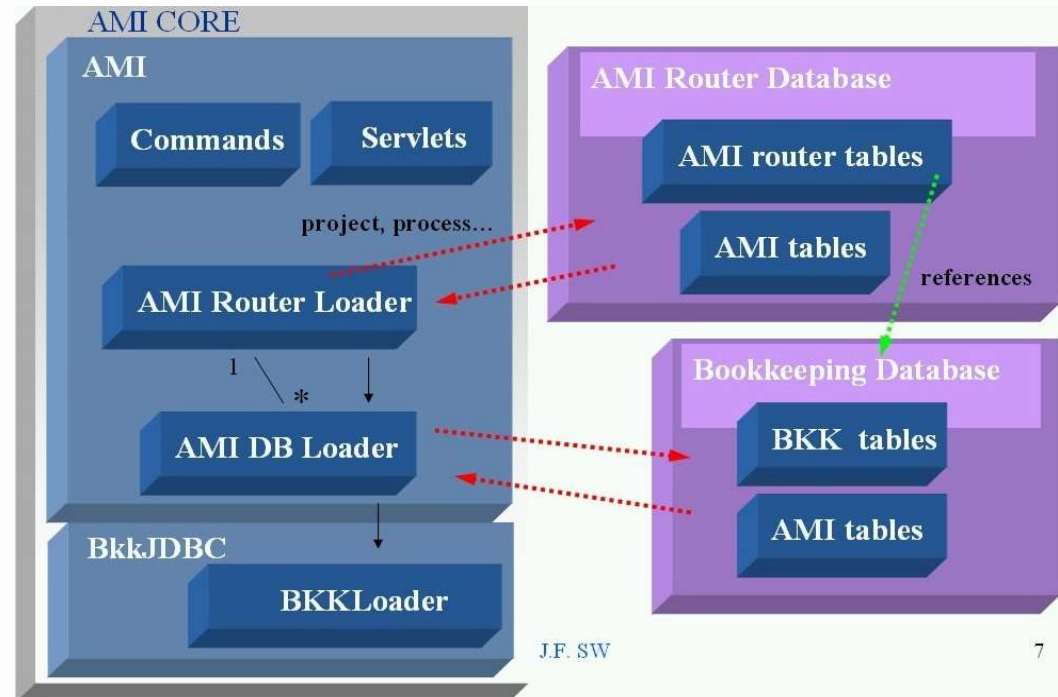
- Two generic metadata interfaces: **gLite** and **LCG**
 - Experiments had their own, or no well-defined one.
- In February, a face-to-face meeting of **PTF** was dedicated to metadata and interface.
- The new interface was designed in collaboration between gLite and LCG, with involvement from the metadata group:
 - presence, use-case documentation, ...
- EGEE and LCG are implementing interface
 - other issues have higher priority, so not clear when this will be released.
- ATLAS are implementing this interface within **AMI**.





What is AMI?

- Flexible system for building database applications.
- Based on self-describing databases.
- Many heterogeneous databases are accessed via a router database.
- Access via web servlets or SOAP web service.





- Implementation of generic Metadata Catalogue interface as a wrapper for AMI.
 - Provides a simple, interoperable interface.
 - Currently implementing Metadata Query Language
 - A simplified sub-set of SQL defined by EGEE.
- Security using Grid Certificates.
 - AMI can recognise a user by his certificate.
 - Certificates are registered with AMI via a servlet.
 - Users are mapped to roles which provide privileges.
 - Hope to use VOMS to provide roles in future...





New AMI features

- The 'Request a Task' interface allows jobs to be submitted to the Atlas Grid Production system through AMI.
 - Simple way to submit a job.
 - All datasets produced through the RAT are properly logged in AMI.
 - Keeps the metadata up to date with all data.
- Database Connection Pooling.
 - Adds to the scalability of the application.
- AMI Administration interface.





- A new Physics Metadata interface.
 - AMI has a generic servlet interface.
 - Users find this somewhat opaque!
 - Want to build a web interface dedicated to physicists' needs.
 - Requirements/ideas being gathered.
- Deployment at Lyon Tier 1 / CERN.
- Use of delegation tools to contact other services on users' behalf.
 - In current Atlas model AMI will handle only datasets.
 - Contact another service to find out about files.





- GANGA: a component-based UI for job definition and management.
- AMI-GANGA interface plug-in
 - Allow user to access AMI tables from within GANGA
 - UI designed to be familiar to AMI web UI users.
 - Uses positive features of AMI web and (in general) hide complexity to become more user friendly.
- Allows users to query without knowledge of SQL or AMI schema
 - but still allows people to get their hands dirty
- Currently creates ATHENA job-options file.





Simple search

Search AMI tables - construct a general search (all users)

Please select your Project and Process
DC1 simulation

Select the entity you want to search for
dataset

Define your search criteria
physicsWorkingGroup

Attribute = Higgs add

dataset.physicsWorkingGroup = 'Higgs'

Search

results

Select all

Property to add LFN

Select the Options file to add to /afs/cern.ch/user/

Add

- Currently in development
- Search suitable non-expert user.
- Current design is similar to AMI web search.
- Puts LFNs in ATHENA Job-options file.
- Visual design is in prototype and may change.





SQL search

Search AMI tables - construct an sql query (Advanced users)

SQL query
Please select your project and processing step

Project Process

SELECT .

FROM ,

WHERE . =

. =

. =

. like

results

Property to add (LFN must be in pool file catalog to run)

Options file to add to

- Prototype of an “advanced” search.
- Requires knowledge of both SQL and AMI schema.
- Greater flexibility in what can be queried.
- With flexibility, inherent danger, for example creating Cartesian products.





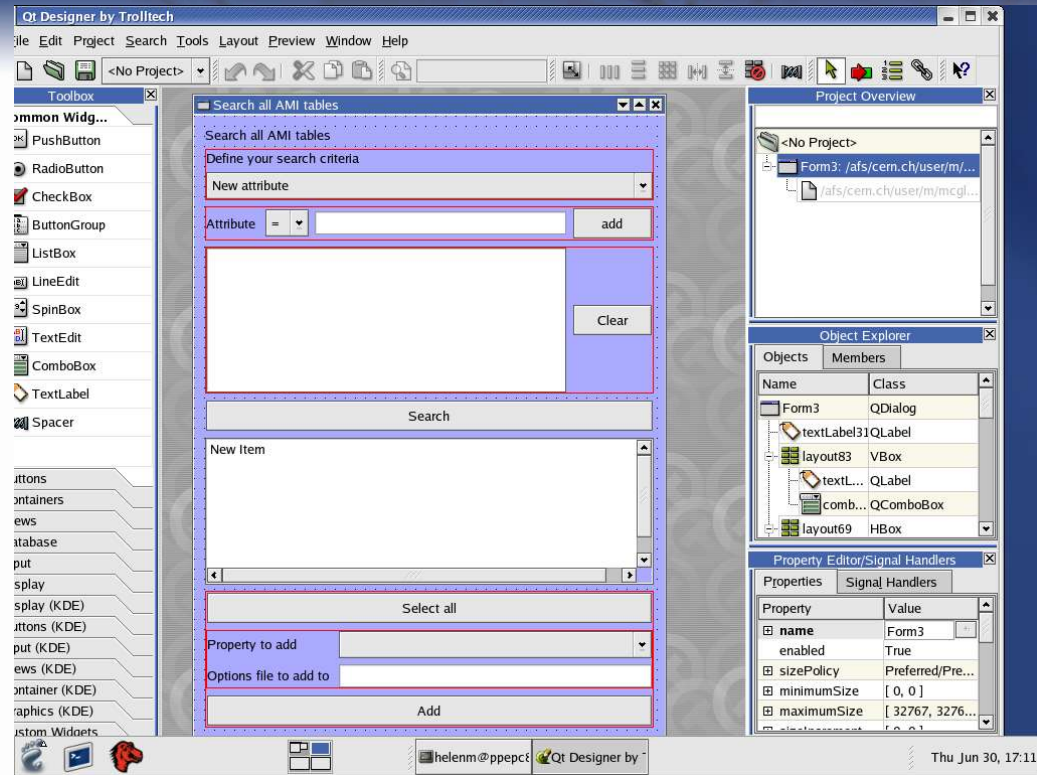
Design process

How it works:

- Written in PyQt and Python
- WXPYthon not used due to excessive dependencies

Design process:

- Qt designer (IDE)
- pyuic (XML to Python)
- python



Communicating with AMI:

- python-ized AMI cmd
- pyAMI (ZSI based webservice)
- AMI (server)

Future:

Develop full functionality
Consider further improvements ...





- Subgroup of the PTF
- Run by Abdeslem Djaoui
- Has mailing list, wiki, and CVS area.
- Current task is to build a webservices cookbook.
- Imports work from EGEE, but is not limited to EGEE
- Future is unclear (since PTF dissolved, 20th June)





GridPP

UK Computing for Particle Physics

The future

- Continue with facilitating cross-experiment communication
- More experiment focused work
- Continue development of common interface: supported by AMI/ATLAS, LHCb, (others ?)



UNIVERSITY
of
GLASGOW



GridPP

UK Computing for Particle Physics



Questions?



UNIVERSITY
of
GLASGOW