Event Selection Services in ATLAS

Jack Cranshaw
Argonne National Laboratory
for the ATLAS TAG Development Team

The ATLAS Data Model

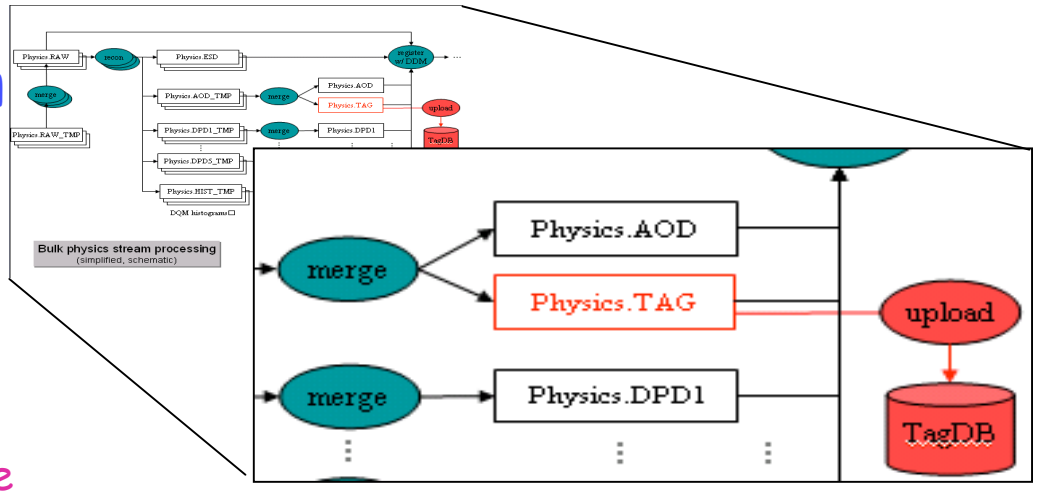
- The expected data taking for a standard LHC year is 10^7 seconds.
- The detector will take data at roughly 200 Hz.
  - Limited by offline data storage, not detector trigger rate. ***Could go up!***
  - Does not take into account simulation data which will have comparable size (20-50% of RAW data). We already have roughly 1 PB of data and the LHC has yet to have collisions.
- Multiple stages of processing (RAW 2 MB/ev, ESD 500 MB/ev, AOD 250 kB/ev, DPD 10-100 kB/ev) with multiple versions.
- Extensive calibration systems and data needed.
  - Refinement based on iterative processing and physics results over the course of 1 month to 1 year.
- Metadata key to finding data.
  - Provided by multiple systems at multiple sites in multiple formats.
  - Modulo some scaling and distributed computing problems we can build on previous experimental experience for much on geometry and detector conditions.
  - But providing event-level metadata has rarely been done and never at these scales.

Components Adopted From LCG

- TAG software is built on top of the POOL Collections package (see Andrea's talk).
- ATLAS is the primary user of this software package and contributes much of the development effort.
- Two persistency mechanisms
  - File based: ROOT
    - Schema object, TTree with TAG data, TTree with metadata (for collection/file)
  - Relational Database: CORAL
    - Collection definition tables, TAG data tables, File navigation tables, metadata table for collection metadata.
- This provides a flexible system for
  - TAG creation using grid processes (files).
  - Global TAG selections (relational).
  - Analysis of selected events (file).
- Commissioning of this system is ongoing. ATLAS software challenges have led to the development and implementation of new features in POOL. In general, changes are required at each order of magnitude increase in data volume either because of performance needs or increasing data complexity.

TAG Data Storage and Distribution

- Two things are done with the produced TAG data.
  - It is grouped into file datasets and distributed by the ATLAS Distributed Data Management system.
  - After transfer to CERN, the files are uploaded to an Oracle RAC system at CERN and at Oracle installations at participating Tier 1's (BNL, TRIUMF, …).
- This is done centrally from CERN by processes writing locally to the CERN database or over the network to the remote databases.
  - In general this requires running several writing processes in parallel, and the system can scale by adding processes



- More details on the poster ***"ATLAS TAGS Database distribution and management"***.

Data Available

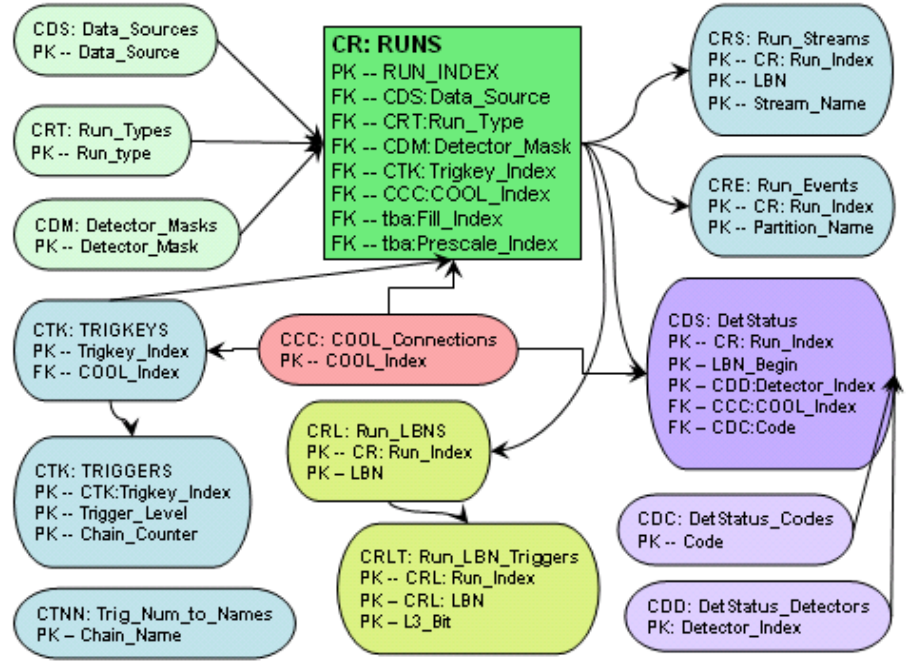
- TAG production is done during merging of AOD files during Tier0 reconstruction at CERN or on the grid at Tier1 facilities during reprocessing using a transform provided and maintained by the ATLAS Physics Analysis Tools group.
- Thus far TAG production and processing has been irregular
  - Produced regularly with non-standard content for detector commissioning exercises (10^7 events). Files merged into larger files to cover a full run, but not imported into the database.
    - TAGS from reprocessing will be loaded.
  - Produced for dedicated exercises designed to mimic detector data such as the Full Dress Rehearsal (10^6 events).
  - Produced for some Monte Carlo data, but the power of TAGS are larger when extracting data from heterogeneous samples than from homogeneous samples.

An Integrated Approach to Event Selection

- We have developed an integrated system for doing event selections called the Event Level Selection Service Interface (ELSSI). This is deployed as a web interface which will be described in the following slides.
- As an integrated system it brings together not just the information in the TAG database, but also information on
  - quality conditions,
  - detector status,
  - trigger configuration information,
  - …
    - Much of this information sits in other database systems or files.
- ELSSI is also grid-aware using the Acacia software (http://acacia.wiki.sourceforge.net/) and requires a grid certificate for access
  - Allows a level of security on access.
  - Allows the possibility to submit grid jobs using the user's grid certificate directly from ELSSI.
- While the TAG data is replicated to remote sites, the services are being prototyped at CERN before deploying to other sites.

Non-Event Metadata

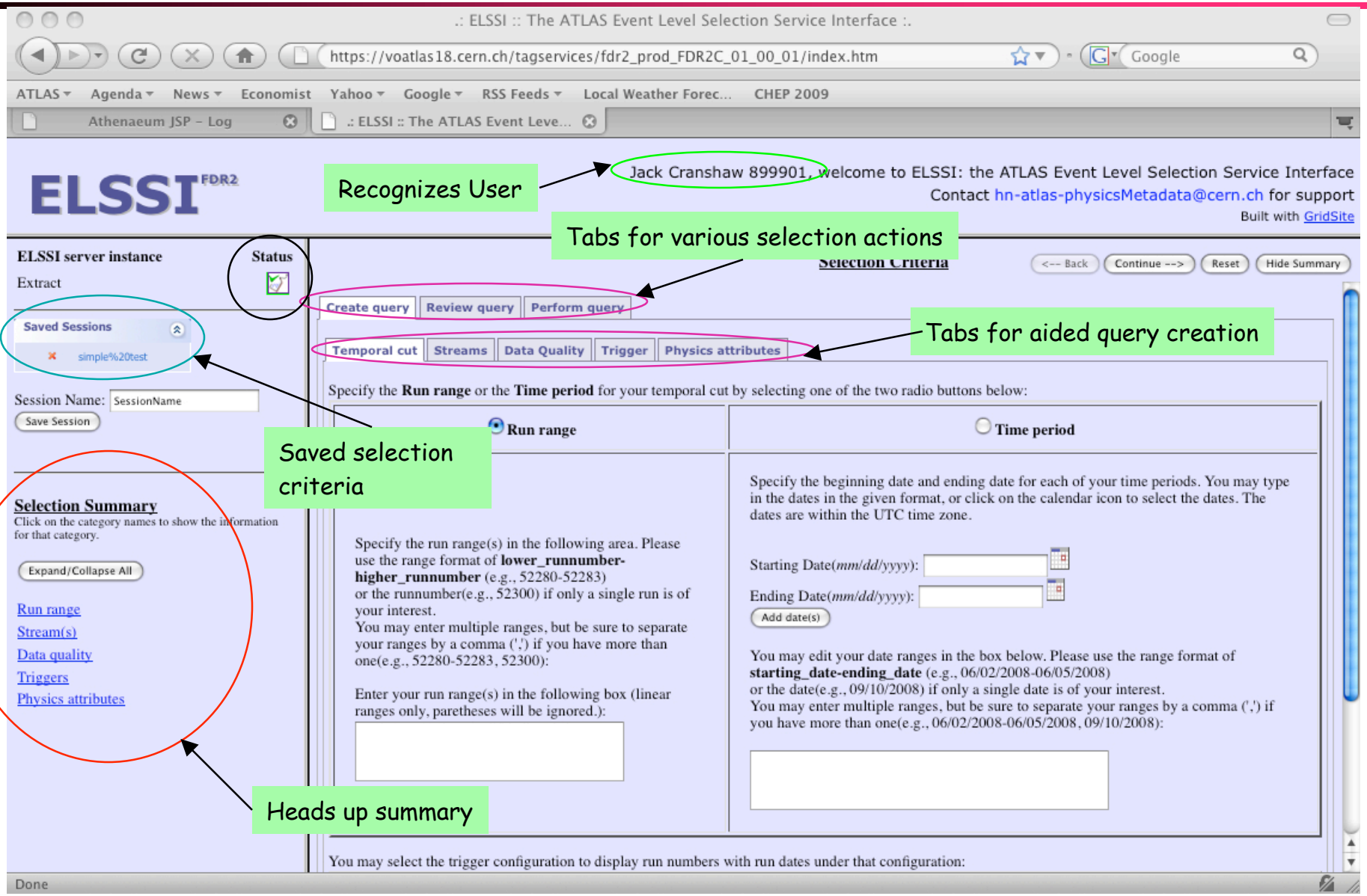
- A meaningful event selection in ELSSI requires information on triggers and other "conditions" which change at intervals coarser than event-by-event. For performance reasons, this information is extracted from the sources that it is originally written into, e.g. COOL; and stored in a dedicated schema for use by ELSSI.
- These data are distributed along with the TAG data to sites hosting a relational TAG database.



Supporting Services

- ELSSI requires several supporting services
  - A stable and clonable hardware system with Apache and php installed.
  - Grid authentication for access to database resources and for eventual submission of grid jobs.
  - Extraction of the selected events and their metadata into a format which can be used as input to an athena job. This service is essentially a wrapper around one of the POOL Utilities: CollAppend.
- Along with ELSSI these services are deployed on two machines at CERN, one for development and one for production users. The code is managed in cvs, but is not part of ATLAS software releases.
- An initial test installation of ELSSI + services was deployed to the Midwest Tier2 at UChicago.

ELSSI Web Page Overview



ELSSI Tour: Stream Selection

- Things to note:
  - Atlas data is divided into streams based on related triggers.
  - The same data is reprocessed multiple times.

Create query | Review query | Perform query

Temporal cut | Streams | Data Quality | Trigger | Physics attributes

Indicate the stream(s) from which you would like to select by clicking in a check-box. Collections of data sets will be displayed according to your selection.
You may select all collections in each single stream, or you can select individual collections of your interest.

Streams	Collections under each stream of "fdr08_run2" runs
☑ Bphys	All Bphys fdr08_run2_physics_Bphys_o3_f47_m26 **fdr08_run2_physics_Bphys_o3_f48_m27** fdr08_run2_physics_Bphys_o3_f8_m10
☑ Egamma	**All Egamma** fdr08_run2_physics_Egamma_o3_f47_m26 fdr08_run2_physics_Egamma_o3_f48_m27 fdr08_run2_physics_Egamma_o3_f8_m10
☑ Jet	All Jet fdr08_run2_physics_Jet_o3_f47_m26 fdr08_run2_physics_Jet_o3_f48_m27 **fdr08_run2_physics_Jet_o3_f8_m10**
☑ Minbias	All Minbias fdr08_run2_physics_Minbias_o3_f47_m26 **fdr08_run2_physics_Minbias_o3_f48_m27** fdr08_run2_physics_Minbias_o3_f8_m10
☑ Muon	All Muon **fdr08_run2_physics_Muon_o3_f47_m26** fdr08_run2_physics_Muon_o3_f48_m27 fdr08_run2_physics_Muon_o3_f8_m10

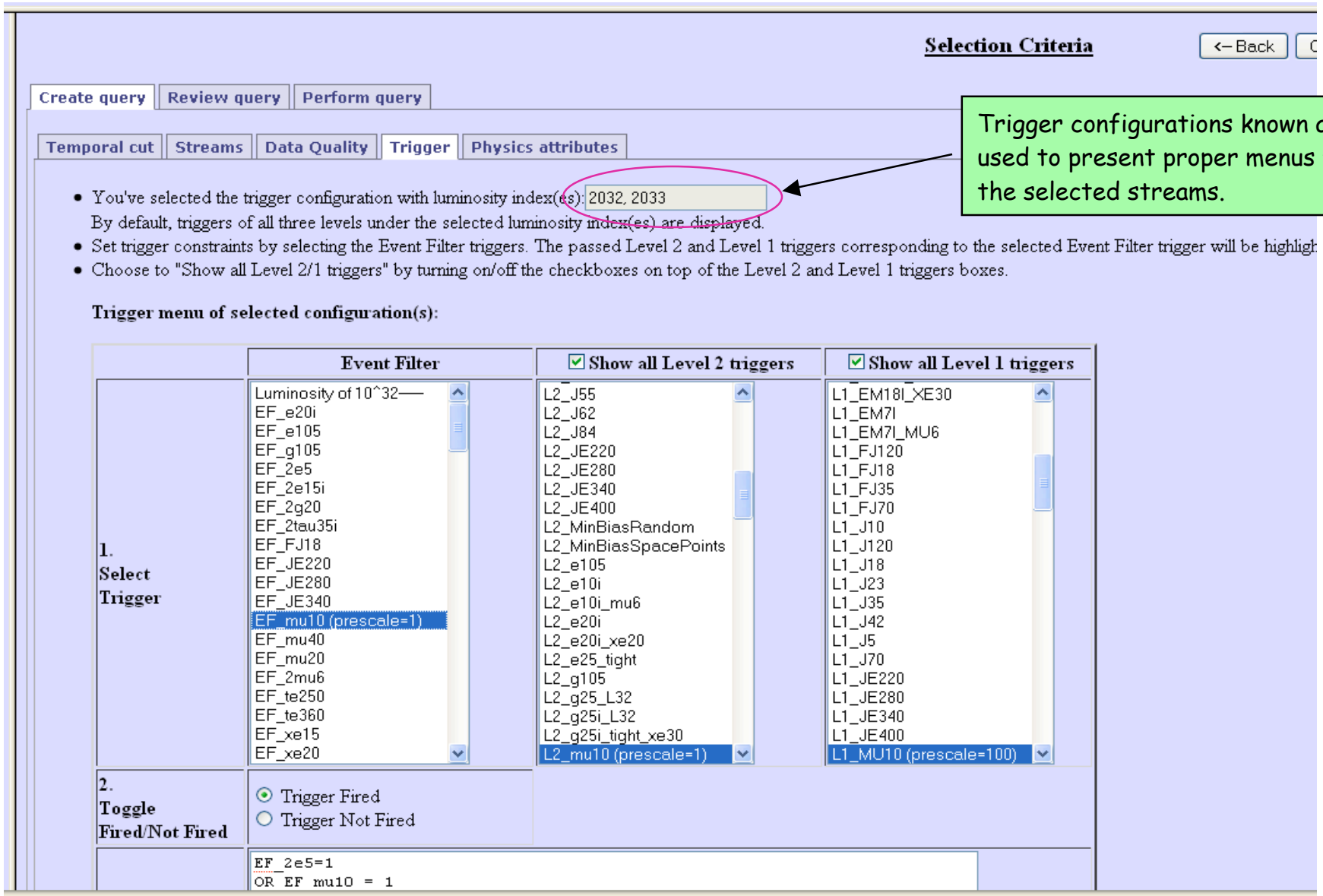
Selected streams:

```
fdr08_run2_physics_Bphys_o3_f48_m27
fdr08_run2_physics_Egamma_o3_f47_m26
fdr08_run2_physics_Egamma_o3_f48_m27
fdr08_run2_physics_Egamma_o3_f8_m10
fdr08_run2_physics_Jet_o3_f8_m10
fdr08_run2_physics_Minbias_o3_f48_m27
fdr08_run2_physics_Muon_o3_f47_m26
```

Restore saved stream selection

Reset

ELSSI Tour: Trigger Selection



Trigger configurations known and used to present proper menus for the selected streams.

ELSSI Tour: Review/Refine Query

Selection Criteria ← Back | Continue → | Reset

Create query | Review query | Perform query

Your streams are :

```
fdr08_run2_physics_Bphys_o3_f48_m27
fdr08_run2_physics_Egamma_o3_f47_m26
fdr08_run2_physics_Egamma_o3_f48_m27
fdr08_run2_physics_Egamma_o3_f8_m10
fdr08_run2_physics_Jet_o3_f8_m10
fdr08_run2_physics_Minbias_o3_f48_m27
```

Your constraints and the resulting query

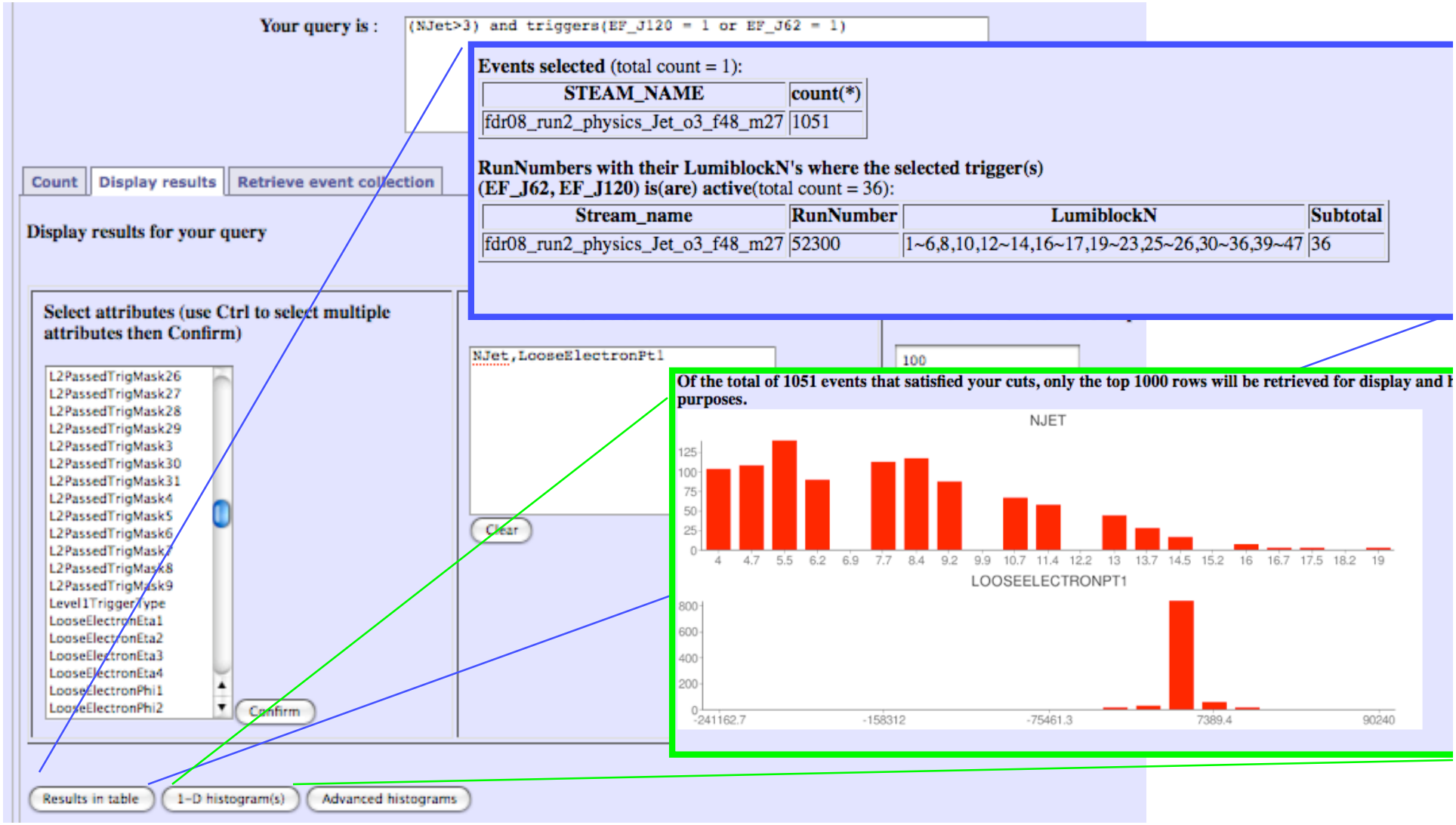
You may in the following editable areas modify your query *if necessary*, using AND, OR and parenthesis, to form your SQL constraints(e.g., (a=1 OR b=2) AND c=3). Do NOT insert newline carriage return) into each constraint expression.
(Note: The units in TAG database are in MeV instead of GeV.)

Run ranges or time periods (Enter linear ranges only, paretheses will be ignored.)	Physics cuts (Remember to add abs() to your η and P_T cuts as required!)	Trigger constraints
52280-52304	NLooseElectron>2 AND abs(LooseElectronPt2)>100000	EF_2e5=1 OR EF_mu10 = 1

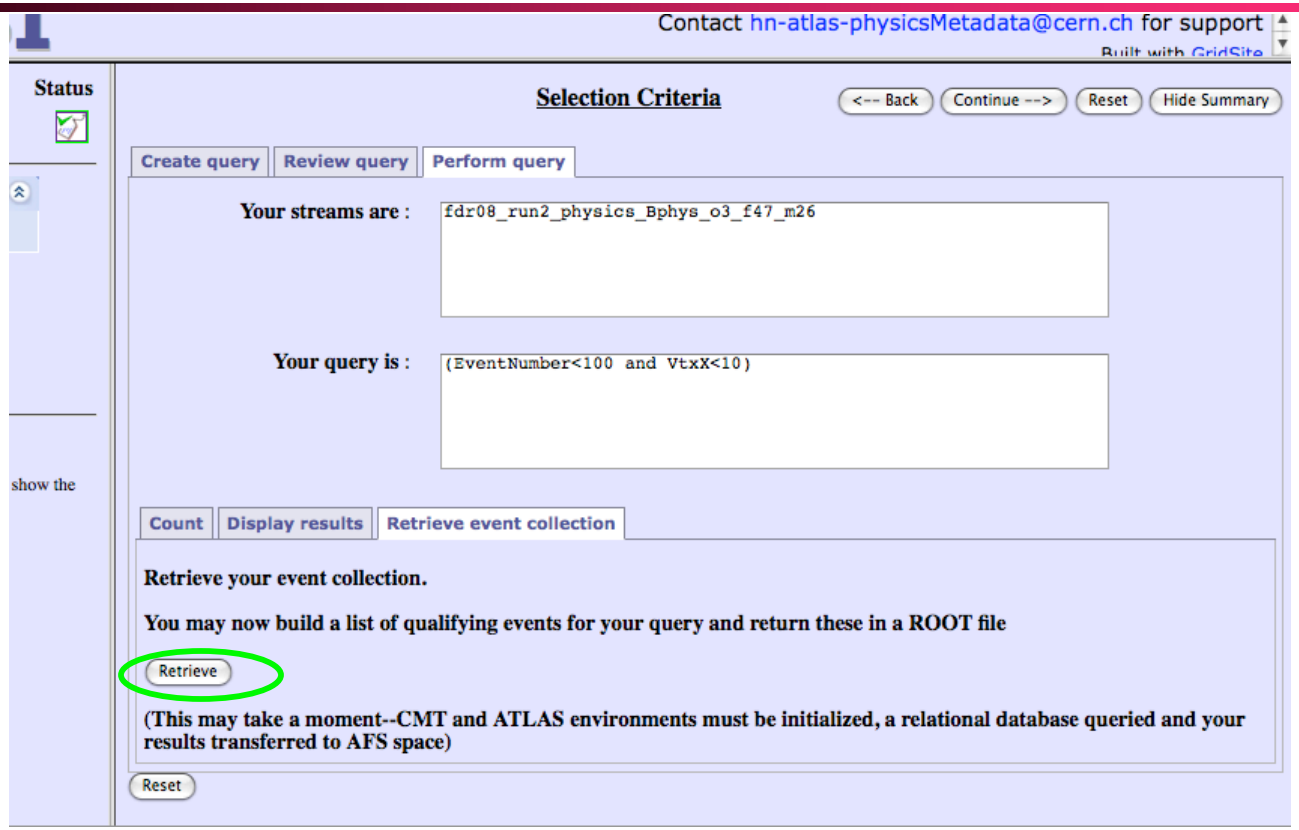
The resulting query:

```
((RunNumber >= 52280 and RunNumber <= 52304)) and (NLooseElectron>2 AND
abs(LooseElectronPt2)>100000) and triggers(EF_2e5=1 OR EF_mu10 = 1)
```

ELSSI Tour: Perform Query

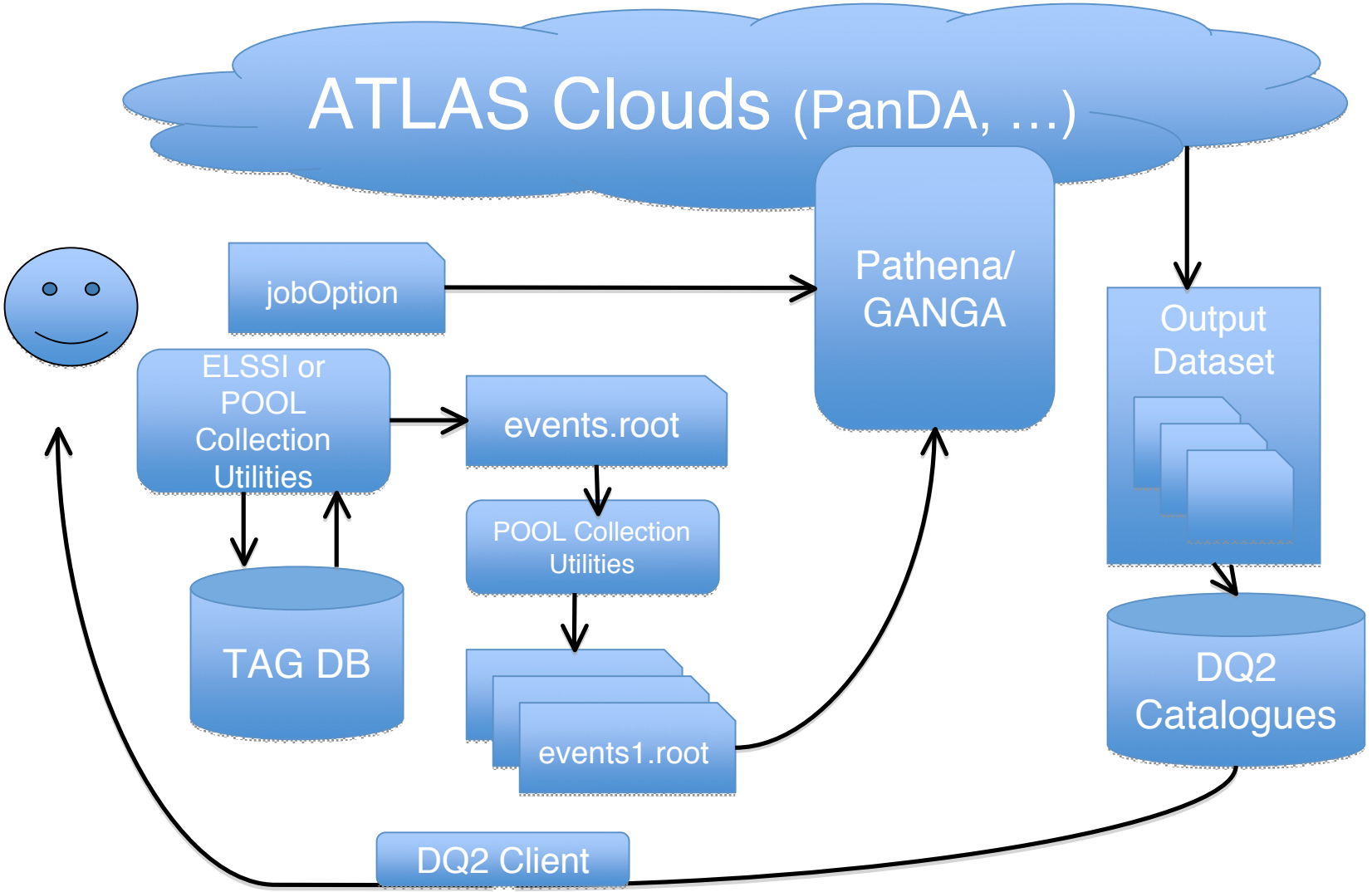


ELSSI Tour: Selection Extraction



- The final page contains a button which calls the Extract service which returns a hyperlink to an POOL Collection ROOT file (also available on afs).
- With an appropriate catalog (provided by DA/DDM tools), this file can be used as input to an athena job locally or on the grid.

Example: TAG's with Distributed Analysis



Summary

- ATLAS has developed an event-level metadata system called TAG which is being deployed through a web interface called ELSSI to provide physicists with a tool for doing analysis.
- This tool
  - Uses the POOL Collections package as the basis for storing the data.
  - Data is produced and distributed in ROOT files.
  - It is also stored in relational databases at CERN and participating ATLAS Tier 1 sites for use by ELSSI.
  - In order to support ELSSI event selections, coarser grain (run, lumiblock) data is extracted from other sources (COOL,…) and stored in a set of tables with the TAG data.
  - ELSSI can extract the selected events references and metadata and write them to a file which can then be used as input to an athena job run on the grid.

Future

- Submission of a skimming job to the grid as part of ELSSI
  - Fixed skim of certain data type (AOD,ESD,…). Prototype using delegated grid credentials already demonstrated.
  - User defined job
- More links to outside metadata sources
  - Standalone Run/Lumiblock level browser whose output can be used by ELSSI or athena.
  - Dataset level data in the AMI database.
  - More information on target data storage using DDM tools.
- Expanded collection-level and file-level metadata
  - Can an analysis which uses ELSSI calculate a cross section?
    - Piece-wise infrastructure tested, but need end-to-end validation.
- More flexibility on data presented and awareness of 'schema' evolution or 'schema' compatibilities.
- Deployment of services to other sites
  - Load balancing, cloud hierarchies, failback options, …
- Further automation and streamlining of the process.