Journal of Physics: Conference Series **219** (2010) 042007 doi:10.1088/1742-6596/219/4/042007

Event selection services in ATLAS

J Cranshaw[1,10], T Cuhadar-Donszelmann[2], E Gallas[3], J Hrivnac[4], M Kenyon[5], H McGlone[5], D Malon[1], M Mambelli[6], M Nowak[7], F Viegas[8], E Vinek[9] and Q Zhang[1]

[1]Argonne National Laboratory, HEP Division, Building 360, 9700 S. Cass Avenue, Argonne, IL 60439, USA.
[2]University of Sheffield, Department of Physics and Astronomy, Hounsfield Road, Sheffield S3 7RH, United Kingdom.
[3]Oxford University, Department of Physics, Denys Wilkinson Building, Keble Road, Oxford OX1 3RH, United Kingdom.
[4]LAL, Univ Paris-Sud, IN2P3/CNRS, Orsay, France.
[5]University of Glasgow, Department of Physics and Astronomy, Glasgow G12 8QQ, United Kingdom.
[6]Enrico Fermi Institute, University of Chicago, Chicago, IL 60637, USA.
[7]Brookhaven National Laboratory, P.O. Box 5000 Upton, NY 11973-5000, USA.
[8]CERN, CH-1211 Genève 23, Switzerland.
[9]University of Vienna, Dr.-Karl-Lueger-Ring 1,1010 Vienna, Austria.

[10]Email: Jack.Cranshaw@cern.ch

Abstract. ATLAS has developed and deployed event-level selection services based upon event metadata records ("TAGS") and supporting file and database technology. These services allow physicists to extract events that satisfy their selection predicates from any stage of data processing and use them as input to later analyses. One component of these services is a web-based Event-Level Selection Service Interface (ELSSI). ELSSI supports event selection by integrating run-level metadata, luminosity-block-level metadata (e.g., detector status and quality information), and event-by-event information (e.g., triggers passed and physics content). The list of events that survive after some selection criterion is returned in a form that can be used directly as input to local or distributed analysis; indeed, it is possible to submit a skimming job directly from the ELSSI interface using grid proxy credential delegation. ELSSI allows physicists to explore ATLAS event metadata as a means to understand, qualitatively and quantitatively, the distributional characteristics of ATLAS data. In fact, the ELSSI service provides an easy interface to see the highest missing ET events or the events with the most leptons, to count how many events passed a given set of triggers, or to find events that failed a given trigger but nonetheless look relevant to an analysis based upon the results of offline reconstruction, and more. This work provides an overview of ATLAS event-level selection services, with an emphasis upon the interactive Event-Level Selection Service Interface.

1. The ATLAS Metadata Challenge

The ATLAS detector is expected to take data at the rate of 10^9 events/year during regular LHC operation. A simulation event rate of roughly 20-50% of this size is also expected. This will produce petabytes of data at different levels of processing. ATLAS has already generated a petabyte of data before collisions have started at the LHC. Sorting through a distributed data store of this size is a challenge which must be met by the metadata system. For conditions, data file organization, etc., ATLAS was able to build on successful experience from previous experiments. ATLAS has tried to take this a step further and build an event-level selection capability into our system. This has been done at smaller scales in previous experiments[1].

2. ATLAS Data Products

ATLAS data are processed in stages which produce progressively more refined and filtered data products[2]. The data are divided into streams at the first stage, RAW, based on groups of triggers for different physics signatures, e.g.. Egamma, Jet, Muon, etc. These streams are allowed to overlap. Further data products called Event Summary Data (ESD) and Analysis Object Data (AOD) are also produced as managed productions. ESD are calibrated and clustered detector data (O(1 MB/event)). The AOD are reconstructed physics objects such as electrons, photons, jets, etc., (O(200 kB/event)). Beyond the online trigger selection none of these data products has any further filtering applied. The final stage of official, managed production, which produces Derived Physics Data (DPD), is still being defined and developed. It will have the option of both further filtering and further streaming.

3. ATLAS Event Level Metadata

The ATLAS event-level metadata system is called TAGS. It contains event metadata and sufficient information to navigate to the selected events. The basic structure of this system uses elements of the POOL Collections package[3] provided by the LCG. The package works through an interface which hides the details of the storage mechanism. Currently, two storage mechanisms are supported: file-based in ROOT[4] and relational database resident using CORAL[5]. ATLAS makes use of both storage mechanisms during the creation of the TAGS[6].

The metadata stored in the collection is event metadata, and the references point to event headers in the data store. The TAGS contain the following groups of metadata.

- Event identification: Run number, event number, luminosity block, …
- Information on reconstructed quantities: number of electrons, number of primary vertices, tranverse momentum of the leading particles, …
- Triggers passed stored as bitmasks.
- Quality information based on reconstruction, physics objects, …

The list of TAG attributes is extensible, and new metadata can be added within the limits of available storage and performance. TAG content is managed by the same procedures as AOD and ESD content.

The TAG content is extracted from physics data during AOD merging as illustrated in figure 1. It is done at this stage because most TAG metadata is derived from AOD quantities, but it can be run at any stage where these quantities are available. Certain quantities are also taken from the Conditions database[7]. The resulting TAG data contains navigational references to the merged AOD, the ESD, and the RAW.

AOD merging is done as a managed production at the Tier-0 during initial processing, and at Tier-1s during reprocessing. The TAGS produced in parallel to the AOD merging process are stored in POOL Collection files. These files are collected into file datasets and distributed using ATLAS distributed data management (DDM)[8]. As the files are made available at CERN, they are loaded into a relational database, which allows a global view of the ATLAS event store.

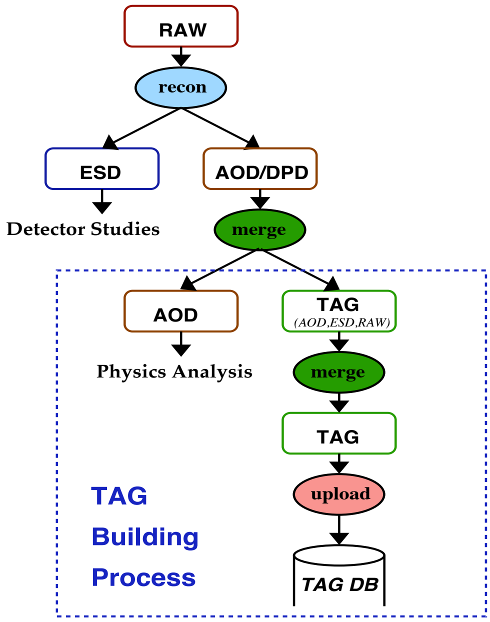

Figure 1: Extraction of TAG metadata and TAG database creation. TAGS contain pointers to AOD,ESD,RAW as noted in the figure.

4. ATLAS Non-event-level Metadata

In order to support event-level selections, metadata of other granularities is needed. The primary examples are data quality information, normally recorded for luminosity blocks or entire runs, and configuration information such as which triggers were active during a run. The latter is needed to decode the trigger decision bitmasks stored in the TAG metadata. Most of this non-event metadata is stored in the Conditions database in a format queryable by interval of validity[9]. For event-level selections, the necessary metadata is transferred into a relational schema alongside the TAGS data to efficiently support cross queries and queries on keys other than interval of validity. The schema is shown in figure 2. These data are distributed along with the TAG data to sites hosting a relational TAG database.

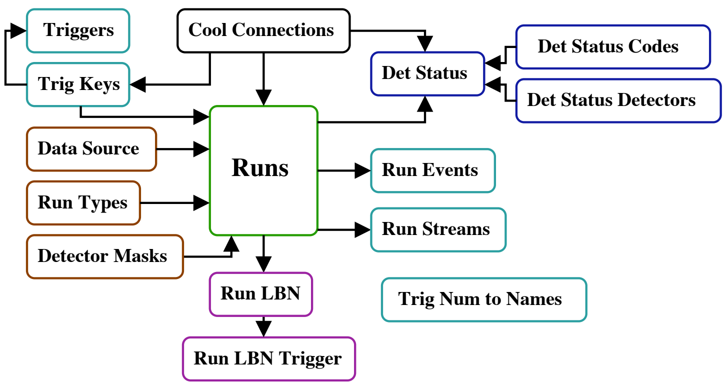

Figure 2. Non-event level metadata schema

5. An Integrated Event Selection Tool

The tool presented in this paper is called the Event Level Selection Service Interface (ELSSI), and it has been developed for ATLAS physics analysis. It brings together all of the metadata needed to do event-level selections in a manner which allows users to query metadata about the full ATLAS event store. This tool is web-based and the output of this tool is currently a POOL Collection file, which can be used as input to a distributed analysis.

The system is built to pick up user grid certificates, and consequently the capability to submit grid jobs directly from ELSSI is being developed and tested. The Acacia software[10] is used to provide the grid-aware capabilities of ELSSI. The following sections contain a snapshot of the current ELSSI capabilities.

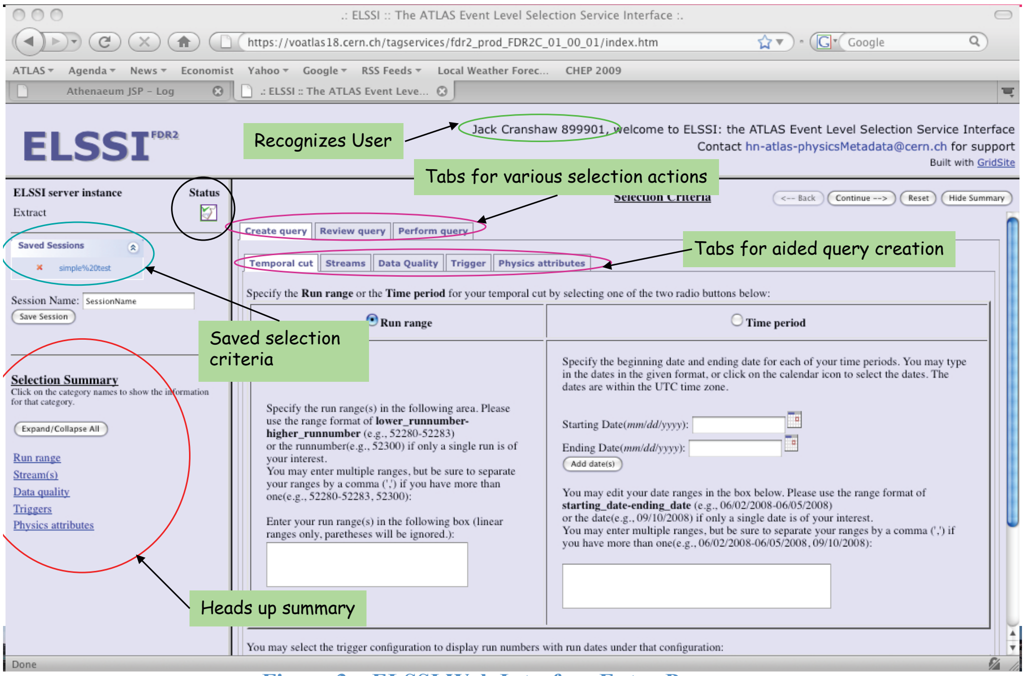

Figure 3a. ELSSI Web Interface Entry Page

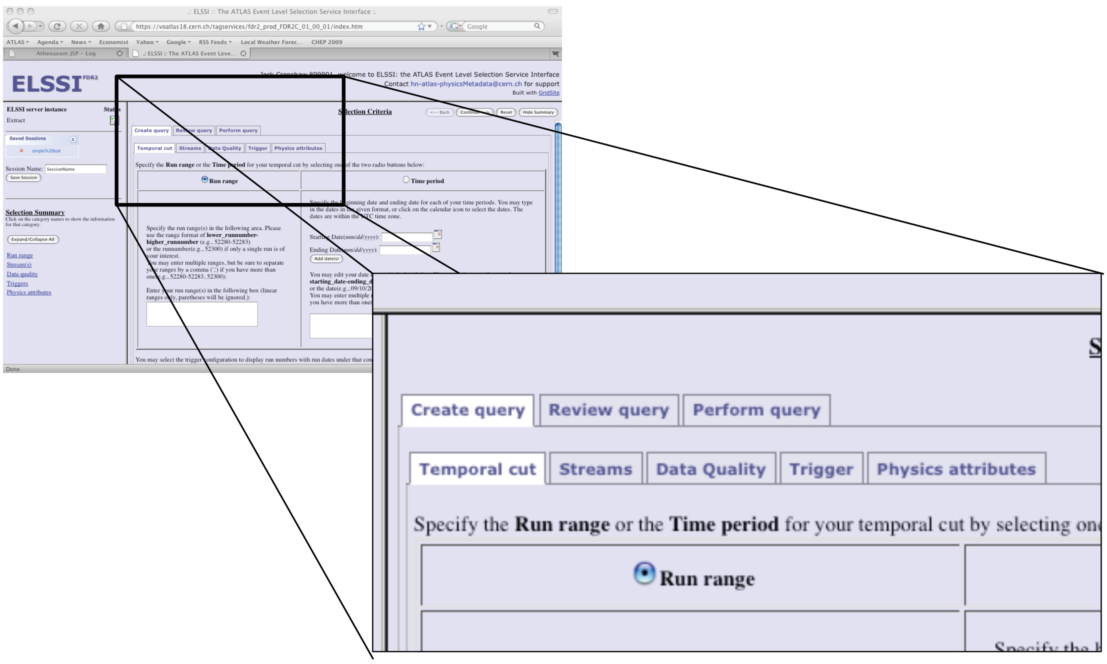

Figure 3b Close-up of the master tabs used to “Create”, “Review”, and “Perform” Selections

5.1. The Web Interface

As shown in Figure 4a, the ELSSI web interface is set in a multi-framed environment.

- The Banner Frame (across the top of the page) includes:
    - The browser version (and targeted datasets).
    - The recognition of the user based on his/her grid certificate.
    - The contact link for comments/questions.
- The Session Frame (down the left hand side) includes:
    - The status of the Extract Server (described later).
    - A text box allowing the user to save the current selections.
    - A list of previously saved selections, which can be reloaded.
    - A summary of the current selection, which the user can expand or collapse.
- The Browser Frame (shown in figure 4b) includes all the main functionality for the TAGS application metadata selection, display, and extraction functions.
    - Buttons:
        - The user can collapse the left (Session) frame with the "Hide Summary" button.
        - The user can reset all criteria with the "Reset" button.
        - The user can navigate through the recommended selection path using the "Continue" and "Back" buttons.
    - Under the "Create Query" tab a multi-tabbed environment guides the user through the selection criteria. Use of the "Continue" button ensures that each criterion is used to refine the selections presented in the tabs that follow. For example, a Run selection in the "Temporal" tab will narrow the list of allowed choices in trigger conditions to those that are configured in those Runs.
    - To review the selections of the current session, the user can:
        - Check the full selection in the summary in the Session frame.
        - Click on an earlier or later tab to review and revise selections.
        - Review and edit the selection criteria to modify the logic of the final criteria in the text boxes in the "Review Query" tab.
    - The "Perform Query" tab contains sub-tabs for presenting the results of the current selection (executing the criteria against the TAGS database)
        - Count the events that meet the selection criteria.
        - Display any event-wise metadata for the events meeting the selection criteria in tabular or histogram (1-D plot) formats.
        - Retrieve the selected events into a file.

The back end database is an Oracle system, where the POOL Collection data table is indexed on all attributes. This makes the count of events fast on a large set of data and keeps the system scalable and flexible to any combination of physics based cuts. This roughly triples the storage requirements versus the non-indexed version. Current plans and resources allow us to keep two versions of the TAGS available online: the two most recent processings.

5.2. Extraction and Analysis of Selected Events

After going through the query creation step, the user can then ask to see the events in the Display tab or proceed to obtain the ROOT files using the Extract functionality. The extraction functionality is provided by a web service that produces a ROOT file with the complete TAG records for the selected events using POOL utilities wrapped in python. The user can retrieve the ROOT file on the AFS file system or via a web link.

The extraction output can be used as input to an ATLAS reconstruction framework job or to a distributed analysis job using GANGA[11] or pathena[12]. At this point the results are handled using tools developed for those activities, and no further TAG-specific services are deployed. In addition to the event references and their metadata, the extracted POOL file also contains information about the luminosity domain and other relevant non-event metadata needed to support physics analyses such as cross section calculations.

5.3. Supporting Services

Like any software deployment ELSSI requires a supporting hardware infrastructure and supporting software services. The architecture is shown in figure 3. The services needed by ELSSI are the following

- A standard web server with Apache and php installed. CERN provides images which are easy to clone and replace.
- 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.

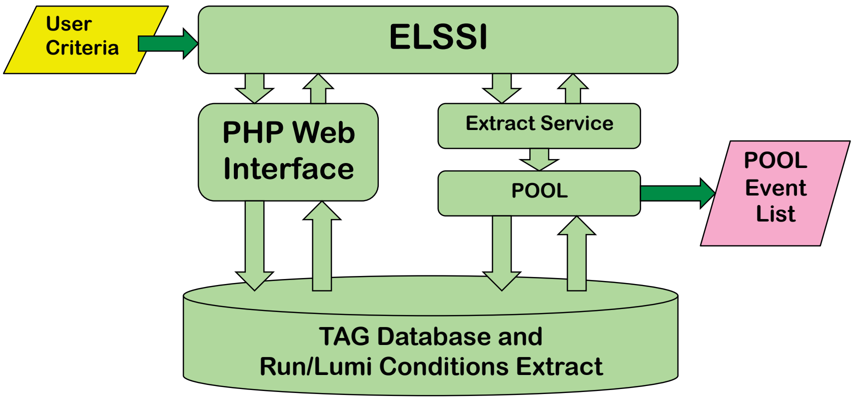

Figure 4. ELSSI Services Architecture.

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 with supporting services was also deployed to the USATLAS Midwest Tier2 facilities at the University of Chicago.

6. Outlook and future developments

ATLAS is deploying an event-level selection capability based on an event-level metadata system called TAGS. These data are manipulated using the POOL Collection package provided by the LCG. An interface (ELSSI) has been built which uses TAGS and non-event level metadata to allow users to select events based on the known relevant criteria: stream, trigger, data quality, physics content, etc.
The current ELSSI is a mature prototype that is being tested and already used by several physics groups. Further developments are underway to increase its robustness and functionality. These include the following:

- Allowing submission of a skimming job directly to the grid from ELSSI. This can take several forms, not all of which will necessarily be implemented within ELSSI.
    1. A skim based on job options provided by ELSSI.
    2. A skim based on job options provided by the user.
    3. A skim based on job options and custom code provided by the user.
- Provide more links to outside metadata sources
    1. A Run/Lumi-block level browser is being developed for users who may not need event-level capabilities or have other reasons for browsing run metadata. ELSSI is being extended to accept the output of this tool to configure the event-level selection.
    2. Provide links to dataset level data in the AMI database for users to understand stream content and the processes that created the TAG metadata.

3. Provide more information on target data storage for use of the ELSSI output in an analysis using DDM tools. I.e. where does the user need to run the job to access the selected events, or which data need to be pulled to run the job.

- Validate that the collection-level and file-level metadata is sufficient for physics analyses. The piece-wise infrastructure has been tested, but still needs an end-to-end validation.
- Different productions of TAGS have different characteristics: TAGS from physics simulations, for example, are very different from TAGS from cosmic ray comissioning. Currently ELSSI normally has to be customized when these characteristics change. There is a need for increased flexibility with respect to the data presented, and for awareness of schema evolution and issues of schema compatibility.
- Deploy ELSSI and supporting services to other sites. This requires packaging to conform to existing ATLAS deployment procedures. The deployed copies would allow for load balancing, cloud hierarchies, fallback options, limited capacity on a single site, and more.
- Further automation and streamlining of the process to improve performance and usability.

7. Acknowledgements

This work supported in part by the U.S. Department of Energy, Division of High Energy Physics, under Contract DE-AC02-06CH11357. We also gratefully acknowledge the support of the UK Science and Technology Facilities Council.

References

[1] L.A.T Bauerdick, et. al., “Event indexing systems for efficient selection and analysis of HERA data”, Comp. Phys. Comm. 137, pp 236-246, 2001.
L. Lueking, et. al., “The Sequential access model for Run II data management”, FERMILAB-CONF-98-382-E, 1998,
[2] Atlas Computing Group, “Computing Technical Design Report”, CERN LHCC-2005-022 ISBN 92-9083-250-9, 2005.
[3] A. Valassi, et. al., “LCG Persistency Framework (POOL, CORAL, COOL) - Status and Outlook”, these proceedings.
http://pool.cern.ch
[4] R. Brun and F. Rademachers, “ROOT: an object oriented data analysis framework”, Nucl. Instrum. Meth. A 389, pp 81-86, .
http://root.cern.ch.
[5] A. Valassi, et. al., “LCG Persistency Framework (POOL, CORAL, COOL) - Status and Outlook”, these proceedings.
[6] F. Viegas, et. al., “Relational databases for conditions data and event selection in ATLAS”, J. Phys. Conf. Ser. 119 042032, 2008.
[7] A. Vaniachine, et. al., "Advanced Technologies for Scalable ATLAS Conditions Database Access on the Grid", these proceedings.
[8] M. Lassnig, et. al., “Managing ATLAS data on a petabyte-scale with DQ2”, J. Phys. Conf. Ser. 119 062017, 2008.
[9] F. Viegas, et. al., “The ATLAS TAGS Database Distribution and Management: Operational Challenges of a Multi-terabyte Distributed Database System”, these proceedings.
[10] http://acacia.wiki.sourceforge.net
[11] J. Elmsheuser, et. al., “Distributed Analysis using GANGA on the EGEE/LCG Infrastructure”, J. Phys. Conf. Ser. 119 072014, 2008.
[12] T. Maeno for the ATLAS collaboration, “PanDA: Distributed production and distributed analysis system for ATLAS”, J. Phys. Conf. Ser. 119 062036, 2008.
M. Mambelli, et. al., “Job Optimization in ATLAS TAG-based Distributed Analysis”, these proceedings.