Evolution of the ATLAS Metadata Architecture and Databases

Borut Paul Kersevan, Davide Costanzo, Dario Barberis

Ljubljana/Sheffield/Genova

Introduction

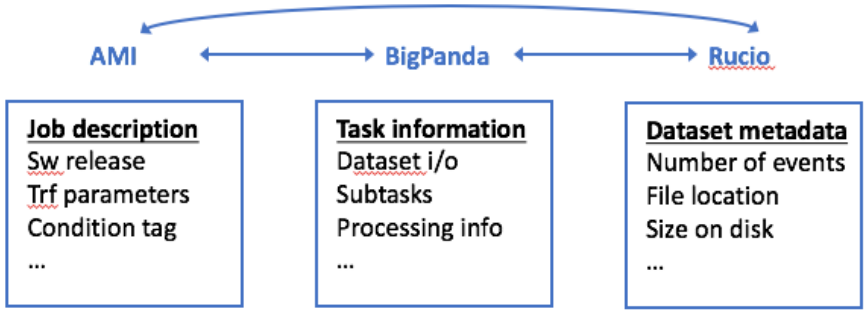
- The current ATLAS metadata architecture and related database system **is mature and working well** in Run 2.
  - **Meaning now is the time to explore its future evolution and introduce adiabatic (gradual, parallel) changes towards Runs 3 and 4.**
  - There are several aspects we need to explore:
    - **The metadata we store,**
    - **the storage/DB technologies for this metadata,**
    - **the way we mine and access this metadata and**
    - **the user interface to this metadata.**
  - These aspects need to be checked on how they match the computing model and technologies future evolution and developed accordingly.

Metadata in ATLAS

- Metadata used and stored in ATLAS has very different origins, types and granularity:
  - **Detector information** (running conditions) and corresponding MC conditions.
  - **Data and MC processing configurations**:
    - SW release, job configuration on various processing steps (RAW -> AOD -> DAOD…) (ProdSys/PanDA)
  - **Data storage information:**
    - Dataset location, volume, formats, replicas (Rucio)
    - Location and basic metadata of recorded data events (EventIndex)
  - **Analysis-level physics information**: cross-sections of MC physics processes, derived pile-up conditions, various types of post-processing optimizations of reconstructed objects and systematics uncertainties from theoretical and experimental sources.
  - **Management-level physics information:** Which data and MC are being actively used, in what formats, for which analysis/papers…

Storage of Metadata

- **Storage of metadata is very varied:**
  - Some metadata stored in-file (requires open/read file)
  - Some metadata stored in databases of various types: Oracle, Hadoop… (requires DB access)
  - A lot of useful physics analysis-level information in TWiki (required read-write access and control)
- **We have a largish set of tools that produce metadata and aggregate it (separately):**
  - Metadata from Tier-0, MC production, production system, …
  - Aggregation through specialized tools with DB back-ends and separate UI: AMI (Main ATLAS Metadata Catalogue), COMA, Run Query, Data Summary, BigPanda monitoring, Rucio data management tools …
  - Often another level of parsing these sources with clever scripting generating TWiki pages and adding futher metadata information manually:
    - From NLO corrections to what MC samples correspond to what MC data



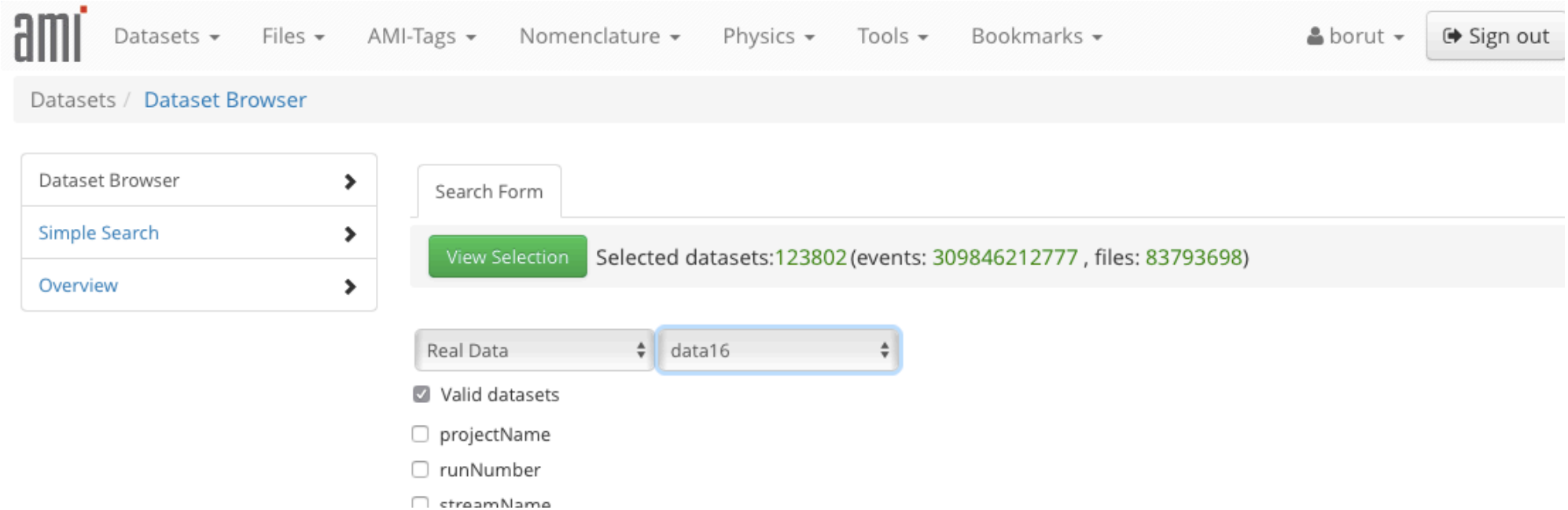
Mining and Access to Metadata

- **The specialized tools are quite good at mining the metadata they aggregate.**
  - Using 'traditional' relational DB searches and are optimized to do it or 'progressive' tools in NoSQL DBs like Hadoop, both working very well…
  - **The challenge is:**
    - adding dynamic information (this dataset is being used by that analysis, recommended by experts for certain types of analyses, obsolete − do not use, etc..) and
    - Maintaining cross-information between the separate metadata collections (use this MC to analyze the data from that detector running period …).
    - Such information then added through TWikis, where it is harder to keep current and validate.

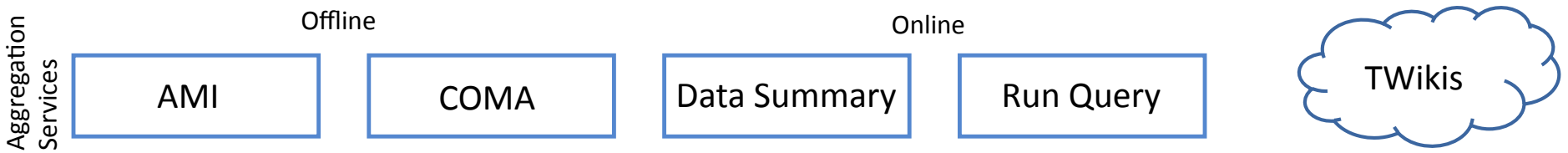
Name	Datasets	Release	AMI tag for data	Physics Container	AMI tag for MC
May2010 reprocessing ("repro04")	As April2010, and adding 2010 7 TeV data up to run 155160	15.6.9.8	r1297 (runs 141748-154817) and r1299 (runs 154822-155160), change of DB release only between the two 'r' tags	repro04_v01	**ESD** (if requested): r1302 (default setup) r1303 (particle ID and cal. hits in ESD for Jet/Etmiss samples) r1304 (GEO-10-01-00 - extra material - sample reconstruction) r1305 (samples without the EM constant term for e/gamma group)

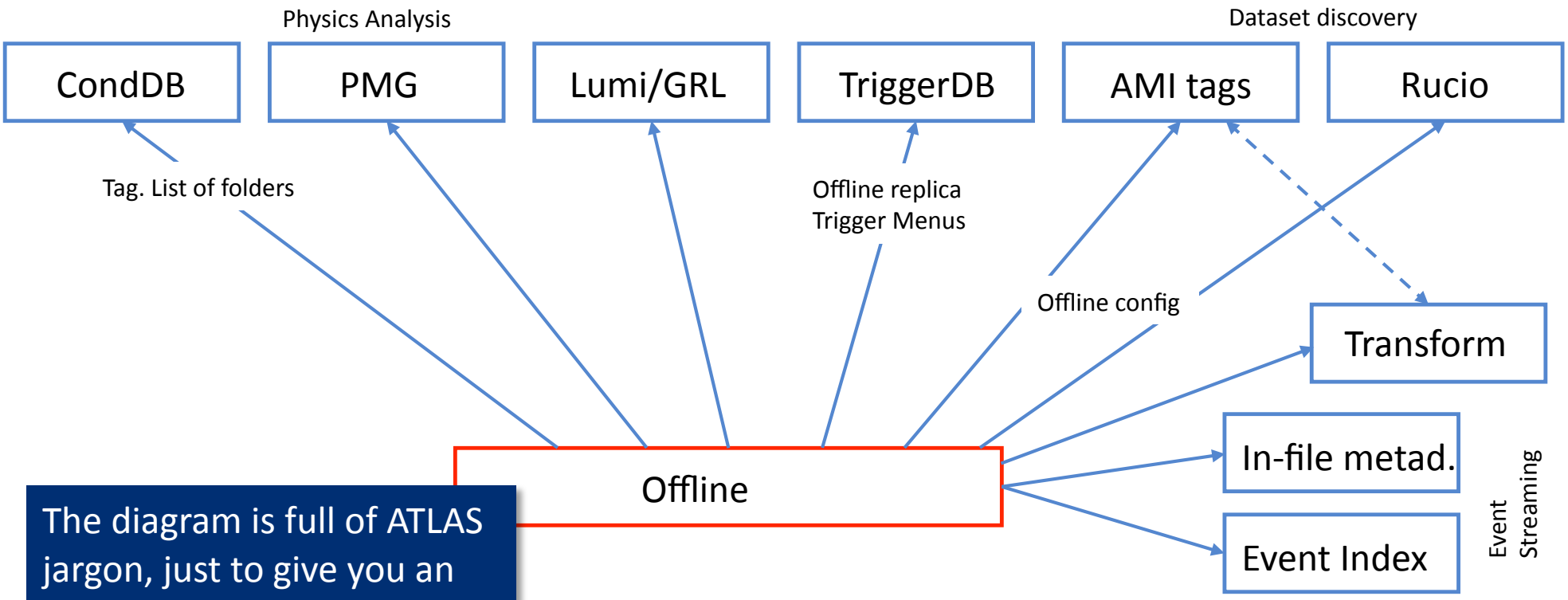
User Interface to Metadata

- **There is not just one type of user, interfaces ideally designed accordingly:**
  - Web UI with user-friendly access and metadata discovery for a student doing physics analysis.
  - Low-level 'command line' or API access for the experts to script and process further.
- **In practice both are often difficult to achieve to everyone's satisfaction and there is always room for improvement.**
  - UI design by physicists for physicists sometimes more successful than others, we are all tough customers… Example of a nice design below.



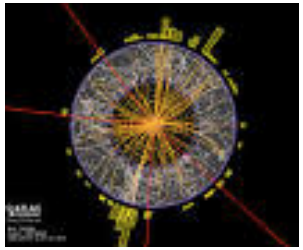
Offline-centric Diagram of the ATLAS Metadata Flow





The diagram is full of ATLAS jargon, just to give you an impression of scale and complexity

New approach: Event-level ATLAS Metadata



- In the future of ATLAS the **event should be the atomic information unit for metadata**.
- **The motivation comes from the anticipated future data processing and storage technologies:**
  - Dynamic data aggregation in data storage to best accommodate the specifics of the storage location and data transfer, i.e. the files are no longer static quantities.
  - Event-level granularity of event processing in heavily parallel computing environments ('Event service' solutions on HPC, already being tried out in ATLAS).
- It also simplifies the handling of e.g. loss of data/MC and or extension of data/MC statistics.
- **Events can be either from data or from Monte Carlo:**
  - A different "representation" of the event can be available to reflect the processing stage under consideration:
    - RAW, Analysis Object Data (AOD), or derived DAOD.
  - The event metadata should carry the provenance information of the event, as well as the (link to the) logical location of the event itself.

Dynamic Aggregation of Event-level Metadata

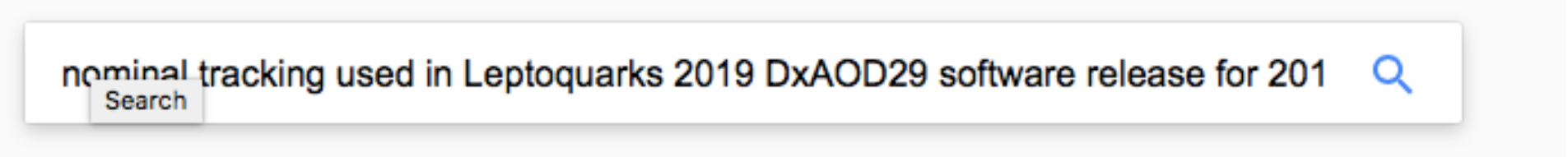
- **Collections** are built from events with the same characteristics:
  - e.g. from the same luminosity block ('Data Sample')
  - generated with the same event generator configuration ('MC Sample').
- **Collections will carry all the metadata with physics content and production system configurations**.
  - The metadata at the collection level will be dynamical, and stored into a "**whiteboard**".
  - The information of the whiteboard will change over time as collections are extended, physics information is improved/added, or the collection is declared obsolete.
  - A versioning mechanism will be needed to manage this evolution.

Whiteboard, Data mining and User Interface

- The UI to the whiteboard should offer both expert (API) and analysis-level graphical Web interface.
  - **A powerful free-form search involving:**
    - MC conditions ('nominal tracking geometry'),
    - detector configurations ('nominal mag. field')
    - processing metadata (software version used)

  should be made available to the users, with as little limitations for cross-referencing as possible.
  - **A simple way to:**
    - add expert or individual collections,
    - annotate metadata to these collections
    - add labels to these ('the ATLAS official ttbar background MC sample' , 'the VdM scan data', 'used for Leptoquark 2018 analysis')

  is required for the whiteboard to perform.
- Such labels should be easy to pass around (URL link) and discoverable by other users in a straightforward 'Google-like' free-form search.



Ongoing activities and Plans

- Currently these plans are taking shape beyond pure wish-lists.
- We are starting the first R&D on potential whiteboard solutions, coordinating the effort with the development of next-generation existing toolsets (e.g. AMI 2.0)
  - The next long shutdown is an occasion to put this new infrastructure in place, hence **the R&D needs to start now!**
- There is a long way to go but we are optimistic:
  - We are driven by demand and future requirements.
  - The new improvements will be added adiabatically whenever they become ready for production-level use.
  - Of course LS2 is a good time to phase in major new developments