ATLAS Data Preparation for Physics Analysis

II PCI2010 & PCI2011 kick-off workshops

Farida Fassi
IFIC, Valencia, Spain

Valencia, Spain, January, 2012
OUTLINE

- Give an overview on data preparation in ATLAS
- how centrally produced data samples are organized to allow coherent data and MC analysis

- Main ingredients of ATLAS data analysis
  - Several data Formats
  - Applications to reduce data volume
  - Steps need to be understood beforehand:
    - Luminosity & LumiBlocks
    - Data quality & Good Runs Lists
    - trigger-aware analysis
Ongoing data-taking

- LHC machine has delivered in 2011 more than expected data
- pp runs finished with > 5 fb-1

A data set of 16 fb⁻¹ is expected to be collected by the end of 2012

Only “stable beam” data with entire ATLAS “ready” and “green data quality flags” is usable for Physics Analysis
Data Flow: from detector to analysis

Point 1
trigger

Tier-0
prompt reconstruction

Tier-1
reprocessings

Calibrations & Alignment

Data Quality Monitoring Histograms

Tier-2/3 analysis

Simulated datasets produced by Tier1/2 during reprocessings

Organized group ntuple production done at Tier-1 using PANDA/ProdSys
At Point 1: data streams definition

- High-level trigger output is organized into **streams**
  - Each HLT chain belongs to one:
    - **Express stream:**
    - **Physics streams:**
      - Egamma, JetTauEtmiss, Muon, MinBias
    - **Special-use streams:**
      - Cosmics, calibration, debug

Determine if the conditions need to be updated

- Once the streams are delivered \( \rightarrow \) Reconstruction starts at Tier-0
- Happens **very soon** after data are taken, even for a run which is still **ongoing**
- Experts have **48 hours** to update the **conditions database** for full reconstruction
Data Reconstruction and distribution

- Bulk reconstruction takes the majority of the Tier0 resources
- Uses an AMI tag to run athena reconstruction with a specific set of options:
  - Runs athena reconstruction, now with an AMI f-tag
  - You can do this too:
    - `Reco_trf.py AMI=f422 inputBSFile=[some raw data]`
- Goal of the bulk reconstruction
  - Produces the diferente data Format

- Data Distribution performed as follows:
  - Tier-1 sites get RAW data
  - Also get reconstruction output data format
  - Usually not needed for physics analysis
  - Group production runs at Tier 1-2 sites to create D*PDs
datasets REPORCESSING

- **Reconstruction software & data formats**
  - **RecExCommon**: single application to reconstruct all types of data (data/MC, cosmics/collisions), with data-driven auto-configuration.
  Interface the so-called: **Reco_trf**
  - Detailed reconstructed objects written in **Event Summary Data** (ESD) files, for detector & reconstruction studies
  - Smaller **Analysis Object Data** (AOD) files are derived from ESD, for physics analysis
  - For coherent data/MC, three sources are combined together:
    - simulated data from a reprocessing
    - real data from the same reprocessing
    - recent real data from calibrated Tier0 reco
    - as long as Tier0 software is “frozen”
Data Formats Flow

- Event Source
  - Monte Carlo Generation
  - Simulation
  - Digitization
  - Data Online
    - ByteStream
  - Conversion
  - "signals"

- Reconstruction
  - Pattern recognition
  - Calibrations
  - Jet/Trk/etc finding

- ESD
  - Ex: Track w/ digits, Jets w/ cells, Trigger Info, etc...

- ESD-->AOD Converter
  - Build "particles"
  - Select good particles

- TAG
  - Num of Jets, muons, etc
  - Most E Jet, photon, etc
  - "Bits" set by physics analysis groups

- Analysis Data
  - Composite particles and their daughters
  - Final selections
  - Visualization
  - Fitting
  - Results of Analysis algorithms

- Athena Analysis
  - Selections
  - Combinatorics
  - Calculating Observables
  - Making Ntuples

- AOD
  - Collections of Particles (INavigable4Momentum): Photons, Electrons, Jets, Muons, etc…
Huge amount of ATLAS data!
that’s needed to cover all aspects of the experiment,
but most of it can be filtered out for your specific analysis.

DPD: read all events and filter some of them with athena algorithm(s)
simplest and most flexible
four levels of data reduction
- Filter complete events which is known as skimming
- Reduce event content by filtering containers, objects and/or properties (operations are: trim, thin, slim)
- some “primary” DPDs centrally produced
  - example: large MET, high p_T electron/muon, group “secondary” DPDs

TAG: only read pre-selected events
- Fastest method, can pick individual events
- Filtering criteria must be pre-defined in TAG
Steps needed beforehand
Luminosity

- Raw input recorded online in units of LumiBlocks (~2 minutes interval)
- Actual luminosity derived offline and stored in COOL database

Key ingredients to define luminosity that you’re running over, you need:
- Either Time range or machine conditions
  - 7 TeV collisions from stable beams
  - data approved for summer conferences (GoodRunList)

- Detector/offline data quality (DQ flags)
  - full ATLAS detector ON
  - no “hot channel” in calorimeter, good MissingET

- Trigger Configuration
  - Your primary physics trigger (like EF_mu18)
  - Both active and unprescaled
Few word on Data Periods

- Runs are grouped into Data Periods that correspond to stable detector conditions.
- If no significant changes, they usually correspond to about a month of data-taking.
- Period changes when something important changes for data quality.

<table>
<thead>
<tr>
<th>E</th>
<th>/+E1</th>
<th>COMA; runQuery; Containers</th>
<th>LAr problem in EMBA: 6 FEB's not operational</th>
<th>2011-Apr-30: 2011-May-03</th>
<th>180614: 180776</th>
</tr>
</thead>
<tbody>
<tr>
<td>F</td>
<td>/-F1:F3</td>
<td>COMA; runQuery; Containers</td>
<td>50ns data after the technical stop. The MDT's switched to &quot;MROD trailer suppression mode&quot; but this is not expected to impact physics. For Tier0 reconstruction improved conditions were applied to correct for the EMEC sagging and for the tilt of the ID w.r.t. the solenoid field.</td>
<td>2011-May-15: 2011-May-25</td>
<td>182013: 182519</td>
</tr>
<tr>
<td>G</td>
<td>/-G1:G6</td>
<td>COMA; runQuery; Containers</td>
<td>new trigger cache which uses pileup noise suppression EMEC sagging correction. In particular jet and tau trigger threshold should be sharper w.r.t. the offline. SCT is now also in X1X readout mode (but this should not impact physics as the detector is well timed in).</td>
<td>2011-May-27: 2011-Jun-14</td>
<td>182726: 183462</td>
</tr>
<tr>
<td>H</td>
<td>/+H1:H4</td>
<td>COMA; runQuery; Containers</td>
<td>update to L1 muon firmware to fix problem with misconfiguration affecting A-side only. The muon trigger efficiency is expected to be improved at eta=0.8-1.0 with respect to previous data periods. Starting at H3 a b-jet trigger was fixed. See comment in H3.</td>
<td>2011-Jun-16: 2011-Jun-28</td>
<td>183544: 184169</td>
</tr>
<tr>
<td>I</td>
<td>/-I1:I4</td>
<td>COMA; runQuery</td>
<td>data after the July technical stop: the 4 LAr FEBs in Layer2 are recovered!</td>
<td>2011-Jul-13: 2011-Jul-29</td>
<td>185353: 186493</td>
</tr>
</tbody>
</table>

https://atlas-tagservices.cern.ch/tagservices/RunBrowser/index.html
Data Quality Flags and GoodRunLists

- Data Quality (DQ) Flags are **simple indicators** of data quality, obtained from monitoring histograms output of the full reconstruction.

- **Granularity of these information is a luminosity block**, from many sources, set for each LumiBlocks:
  - Detectors (divided by barrel, 2 endcaps)
  - Trigger (by slice)
  - Offline combined perf. (e/mu/tau/jet/MET/etc)

- Every working group should decide which set of **DQ flags** is needed for each analysis, which can then be used to generate a **GoodRunList (GRL)**:
  - method is available to make your analysis aware of GRLs
  - Standard lists centrally produced by DQ group

- More details:
  - [https://twiki.cern.ch/twiki/bin/viewauth/Atlas/DataQuality](https://twiki.cern.ch/twiki/bin/viewauth/Atlas/DataQuality)
Coherent datasets in data/MC

- MC11compain is aimed to provide MC data for comparison purpose with real data.
- Currently we have two versions:
  - MC11a (preliminary) and MC11b (final)
- Intended to simulate the major detector changes in proportion to the affected luminosity.
- Emulate the pileup conditions seen in data.
- **Pileup Reweighting**
  - The newest MC11b is simulated with a pileup distribution corresponding to the 2011 data.
  - However, it’s only a good approximation, thus you should still reweight your MC with the official tool.

More details:
https://twiki.cern.ch/twiki/bin/view/AtlasProtected/ExtendedPileupReweighting
Brief overview on trigger-awareness

ATLAS data comes in trigger streams that group several trigger items together.
Physics analysis must be done for specific items only.

- Luminosity can only be calculated for trigger items.
- Trigger composition of streams change over time.
- Streams are not reproduced with simulated data.

Trigger-aware framework provide the main functionality.
Stages in Analysis

- Use TAG to quickly select subset of events which are interesting for analysis. (Skim)
- Starting from the AOD
  - **Stage 0**: Re-reconstruction, re-calibration, selection (AOD)
    - Redo some clustering/track fitting, calculate shower shapes, apply corrections, etc...
  - **Stage 1**: Selection/Overlap removal/complicated analysis (AOD/DPD)
    - Select electrons/photons → find jets on remaining clusters → btag → calculate MET
    - Perform observable calculation, combinatorics + kinematic fitting, ...
  - **Stage 2**: Interactive analysis (AOD/DPD)
    - Final selections, plots, studies.
  - **Stage 3**: Statistical Analysis

Physics group

Analysis group

Personal
**ATLAS Distributed Data Analysis System**

Basic model:
- data is pre-distributed to sites, jobs are brokered to site which has data

Grid details mostly hidden for users - automatic job splitting according dataset size

Output registered in DQ2 and automatically collected to “home-SE”

**New data distribution mechanism – PD2P:** data is distributed based on user demand

Various ways how to submit and choice of Grid gateway
Locate your data: AMI

ATLAS datasets are collections of files
You are probably heard about “Dataset Discovery” which means finding the names of valid datasets to use in your analysis

https://indico.cern.ch/getFile.py/access?contribId=21&sessionId=6&resId=1&materialId=slides&confId=115155
## ATLAS Releases status: production and installation

[Read more](http://atlas-computing.web.cern.ch/atlas-computing/projects/releases/status/)

[Read more](http://panda.cern.ch:25980/server/pandamon/query?overview=releaseinfo)

### Latest Production Releases

<table>
<thead>
<tr>
<th>Release</th>
<th>Patch Release</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>17.5.1</td>
<td>ATlasProduction 17.5.1.1</td>
<td>Upgrade studies on the grid</td>
</tr>
<tr>
<td>17.0.4</td>
<td>ATlasProduction 17.0.4.2</td>
<td>center --&gt; eco migration and ROOT bug fix</td>
</tr>
<tr>
<td>16.6.7</td>
<td>ATlasProduction 16.6.7.17</td>
<td>Update for generation, simulation and analysis-related packages</td>
</tr>
<tr>
<td>16.6.5</td>
<td>ATlasProduction 16.6.5.5</td>
<td>TAg reprocessing</td>
</tr>
<tr>
<td>16.6.3</td>
<td>ATlasProduction 16.6.3.7</td>
<td>Past physics monitoring</td>
</tr>
<tr>
<td>16.6.2</td>
<td>ATlasProduction 16.6.2.6</td>
<td>Physics validation processing MB10b.</td>
</tr>
<tr>
<td>16.4.2</td>
<td>ATlasProduction 16.4.2.2</td>
<td>Upgrade studies on the grid</td>
</tr>
</tbody>
</table>

### Release availability at sites, cache None

### Caches found (and number of sites):

<table>
<thead>
<tr>
<th>Caches found</th>
<th>(and number of sites)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conditions</td>
<td>193</td>
</tr>
<tr>
<td>5.28.00</td>
<td>4(4)</td>
</tr>
<tr>
<td>5.28.00c</td>
<td>7(6)</td>
</tr>
<tr>
<td>2.5.4</td>
<td>17.3.1</td>
</tr>
<tr>
<td>17.3.0</td>
<td>255</td>
</tr>
<tr>
<td>17.0.4</td>
<td>237</td>
</tr>
<tr>
<td>17.0.3</td>
<td>258</td>
</tr>
<tr>
<td>17.0.2</td>
<td>255</td>
</tr>
<tr>
<td>17.0.1</td>
<td>262</td>
</tr>
<tr>
<td>17.0.0</td>
<td>267</td>
</tr>
<tr>
<td>16.7.0</td>
<td>106</td>
</tr>
<tr>
<td>16.6.7</td>
<td>270</td>
</tr>
<tr>
<td>16.6.6</td>
<td>269</td>
</tr>
<tr>
<td>16.6.5</td>
<td>269</td>
</tr>
<tr>
<td>16.6.4</td>
<td>269</td>
</tr>
<tr>
<td>16.6.3</td>
<td>270</td>
</tr>
<tr>
<td>16.6.2</td>
<td>270</td>
</tr>
<tr>
<td>16.4.2</td>
<td>272</td>
</tr>
<tr>
<td>16.0.1</td>
<td>273</td>
</tr>
<tr>
<td>16.0.0</td>
<td>269</td>
</tr>
<tr>
<td>15.6.0</td>
<td>286</td>
</tr>
<tr>
<td>15.5.0</td>
<td>280</td>
</tr>
<tr>
<td>15.4.0</td>
<td>280</td>
</tr>
<tr>
<td>15.3.0</td>
<td>280</td>
</tr>
<tr>
<td>15.2.0</td>
<td>280</td>
</tr>
<tr>
<td>15.1.0</td>
<td>280</td>
</tr>
<tr>
<td>15.0.0</td>
<td>280</td>
</tr>
<tr>
<td>14.5.0</td>
<td>280</td>
</tr>
<tr>
<td>14.4.0</td>
<td>280</td>
</tr>
<tr>
<td>14.3.0</td>
<td>280</td>
</tr>
<tr>
<td>14.2.0</td>
<td>280</td>
</tr>
<tr>
<td>14.1.0</td>
<td>280</td>
</tr>
<tr>
<td>14.0.0</td>
<td>280</td>
</tr>
<tr>
<td>13.9.0</td>
<td>280</td>
</tr>
<tr>
<td>13.8.0</td>
<td>280</td>
</tr>
<tr>
<td>12.9.0</td>
<td>280</td>
</tr>
<tr>
<td>12.8.0</td>
<td>280</td>
</tr>
<tr>
<td>12.7.0</td>
<td>280</td>
</tr>
<tr>
<td>12.6.0</td>
<td>280</td>
</tr>
<tr>
<td>12.5.0</td>
<td>280</td>
</tr>
<tr>
<td>12.4.0</td>
<td>280</td>
</tr>
<tr>
<td>12.3.1</td>
<td>280</td>
</tr>
<tr>
<td>12.0.2</td>
<td>280</td>
</tr>
<tr>
<td>12.0.1</td>
<td>280</td>
</tr>
<tr>
<td>12.0.0</td>
<td>280</td>
</tr>
<tr>
<td>11.5.0</td>
<td>280</td>
</tr>
<tr>
<td>11.4.0</td>
<td>280</td>
</tr>
<tr>
<td>11.3.0</td>
<td>280</td>
</tr>
<tr>
<td>11.2.0</td>
<td>280</td>
</tr>
<tr>
<td>10.5.0</td>
<td>280</td>
</tr>
<tr>
<td>10.4.0</td>
<td>280</td>
</tr>
<tr>
<td>10.3.0</td>
<td>280</td>
</tr>
<tr>
<td>10.2.0</td>
<td>280</td>
</tr>
<tr>
<td>10.1.0</td>
<td>280</td>
</tr>
<tr>
<td>10.0.2</td>
<td>280</td>
</tr>
<tr>
<td>9.9.0</td>
<td>280</td>
</tr>
<tr>
<td>9.8.0</td>
<td>280</td>
</tr>
<tr>
<td>9.7.0</td>
<td>280</td>
</tr>
<tr>
<td>9.6.0</td>
<td>280</td>
</tr>
<tr>
<td>9.5.0</td>
<td>280</td>
</tr>
<tr>
<td>9.4.0</td>
<td>280</td>
</tr>
<tr>
<td>9.3.0</td>
<td>280</td>
</tr>
<tr>
<td>9.2.0</td>
<td>280</td>
</tr>
<tr>
<td>9.1.0</td>
<td>280</td>
</tr>
<tr>
<td>9.0.0</td>
<td>280</td>
</tr>
</tbody>
</table>

### Cloud Site Releases Caches

<table>
<thead>
<tr>
<th>Cloud</th>
<th>Site</th>
<th>Release</th>
<th>Conditions</th>
<th>None</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA</td>
<td>ANALYSIS_ALBERTA-WG1</td>
<td>17.3.1</td>
<td>17.3.1.1</td>
<td>17.3.1.2</td>
</tr>
<tr>
<td>CA</td>
<td>ANALYSIS_ALBERTA-WG1</td>
<td>17.3.0</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td>CA</td>
<td>ANALYSIS_ALBERTA-WG1</td>
<td>17.0.3</td>
<td>AtlasPhysics-17.0.3.1</td>
<td>AtlasPhysics-17.0.3.3</td>
</tr>
<tr>
<td>CA</td>
<td>ANALYSIS_ALBERTA-WG1</td>
<td>17.0.2</td>
<td>AtlasPhysics-17.0.2.10</td>
<td>AtlasPhysics-17.0.2.31</td>
</tr>
</tbody>
</table>
Setup the Analysis Job-Panda (1)

- Using PanDA (Production and Distributed Analysis System for ATLAS)

Client Tools:

Panda Analysis Dashboard

The current pathena version is 0.3.62. The current ganga version is 5.6.10. It is recommended you use the current versions.

Documentation on user analysis with Panda:
- Distributed analysis on Panda - overview page
- Client tools for Panda analysis jobs
- pathena: how to submit athena analysis jobs
- prun: how to submit ROOT and general jobs
- pbook: bookkeeping for Panda analysis jobs
- psequencer: how to perform sequential jobs/operations

Frequently asked questions:
- Full FAQ
- How is job priority calculated?
- Online/Offline status of sites

- PanDA client consists of tools to submit or manage analysis jobs on PanDA
  - **pathena**
    - How to submit Athena jobs
  - **prun**
    - How to submit general jobs
    - (ROOT, python, sh, exe, ...)
  - **pbook**
    - Bookkeeping (retry, kill) of analysis jobs
  - **psequencer**
    - How to perform sequential jobs/operations
    - (e.g. submit job + download output)
Output Datasets and DATRI

Files that are stored on the grid will be placed within Containers and Datasets.

- The output of your Grid jobs will end up in a dataset.
- Datasets and filenames must be uniquely named

User data must follow:  **user.<nickname>.xxx**

<table>
<thead>
<tr>
<th>User:jobID</th>
<th>Created</th>
<th>Latest</th>
<th>Jobs</th>
<th>Pre-run</th>
<th>Running</th>
<th>Holding</th>
<th>Finished</th>
<th>Failed</th>
<th>Cancel</th>
</tr>
</thead>
</table>

In: `mc10_7TeV.105009.J0_pythia_jetjet.merge.AOD.e574_s934_s946_r2299_r2300_tid358712_00`
Out: `user.farida.mc10_7TeV.105009.J0_pythia_jetjet.merge.AOD.e574_s934_s946_r2299_r2300_v1/`

**DaTri interface**

Allows large data transfers to be scheduled to avoid disruption.

- For user datasets, much faster to download from local T2 SCRATCHSPACE than job site.
Running Jobs on the grid: Practical Advice

1) Always test your job locally before submitting to the Grid
2) Use the latest version of pathena/Ganga
3) Submit your jobs on container datasets (merged)
4) If your dataset is on tape, you need to request for data replication to a disk storage first
5) Do not specify any site name in your job submission, pathena/Ganga will choose the best site available
6) Check release and installation status

Where and how you will store your output datasets?
- Request a data replication: your output files will stay as a dataset on Grid.
  → You need to freeze your dataset before requesting for replication
- Download onto your local disk using dq2-get

Note that: by default, user datasets are created on “SCRATCHDISK” at the site where the jobs run.
This disk space is not a permanent place, users have ~15 days to retrieve the data
Replication is needed if user want to see their data on the grid.
Summary

- All data are processed at the CERN Tier0 in two stages
- Express to quickly evaluate the data, and perform calibrations which are used for
- Bulk reconstruction, and then distribution to worldwide ATLAS Tier-centers

- Only reprocessings produce coherent data samples for data & MC combined with Tier0 reco for recent data
- Physics analysis must use specific trigger item(s) and a GoodRunList to...
  - assure good data quality
  - calculate luminosity
  - be coherent with MC
Backup
Data Formats

- **RAW:**
  - Stored in Bytestream format
  - One copy is kept at CERN (tape)
  - One distributed over Tier1’s (disk)
  - Small quantities can be copied to Tier2/group space for special study

- **ESD**
  - Produced from RAW at Tier0 (first pass reco) and Tier1 (reprocessing)
  - One ESD data copy at CERN (tape), two distributed over Tier1’s (disk)
  - Small quantities can be copied to Tier2

- **AOD**
  - Produced from ESD at Tier0 (first pass reco) and Tier1 (reprocessing)
  - At most 2 versions on disk at any given time. 2+1 copies at Tier1’s, 10+10 copies at Tier2’s

- **dESD**
  - Derived from ESD, for detector and performance studies, 10 copies across Tier2’s

- **dAOD:**
  - Derived from AOD, targeted toward physics analysis, definition driven by needs of group analyses, to be stored on Tier2 group space

- **ntuple (D3PD):**
  - Produced by running over d(ESD)/d(AOD), ntuple dumper tools must be validated (official ATLAS status), produced under group/individual control, stored in group space, or locally
ATLAS Computing changes

Dynamic data placement:
- Pre-defined, static data distribution according to fixed shares of certain data types found to be inefficient
  - lots of data is centrally distributed to many sites but hardly used
  - lots of data gets used but not covered by central distribution
- DatRI: User/group dataset replication triggered by users
- System in DQ2, large volumes must be confirmed by local managers
- – users copy only data they really need (in general …)
- PD2P: Panda dynamic data placement replication based on analysis request
  - triggers replication of popular datasets to different site
  - most T2 data to be replicated this way