ATLAS Computing: From Commissioning to 7TeV Data

Graeme Stewart
for the ATLAS collaboration
Outline

• STEP09 Summary and Lessons
• Data Preparation and Calibration
• Tier-0 Operations and Workflow
• Reprocessing
• Data Distribution
• 7TeV Analysis
STEP09 Exercise

• Attempt to do a full chain exercise of ATLAS distributed computing in June 2009
  • Done in concert with other LHC experiments
    • Important for multi-experiment sites
  • Data distribution from Tier-0 $\rightarrow$ Tier-1 $\rightarrow$ Tier-2
  • Reprocessing at Tier-1s (from tape)
  • Large scale analysis activity at Tier-2s
  • Full scale simulation activity going on
STEP 09 Data Distribution

- Data distribution to T1s and (most) T2s worked

3GB/s

Peaks of 5.5GB/s

We can delete too!
STEP09 Tier-1 Reprocessing

- 6/10 Tier-1s validated
- 3/10 Close
- Problems were generally understood
- But system shown to be complex and somewhat fragile

<table>
<thead>
<tr>
<th>TI</th>
<th>Base Target</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASGC</td>
<td>10 000</td>
<td>4 782</td>
</tr>
<tr>
<td>BNL</td>
<td>50 000</td>
<td>99 276</td>
</tr>
<tr>
<td>CNAF</td>
<td>10 000</td>
<td>29 997</td>
</tr>
<tr>
<td>FZK</td>
<td>20 000</td>
<td>17 954</td>
</tr>
<tr>
<td>LYON</td>
<td>30 000</td>
<td>29 187</td>
</tr>
<tr>
<td>NDGF</td>
<td>10 000</td>
<td>28 571</td>
</tr>
<tr>
<td>PIC</td>
<td>10 000</td>
<td>47 262</td>
</tr>
<tr>
<td>RAL</td>
<td>20 000</td>
<td>77 017</td>
</tr>
<tr>
<td>SARA</td>
<td>30 000</td>
<td>28 729</td>
</tr>
<tr>
<td>TRIUMF</td>
<td>10 000</td>
<td>32 481</td>
</tr>
</tbody>
</table>
STEP09 Analysis

- Massive analysis did work
- Hammercloud infrastructure a great success
- But site performance very variable
- Learned how to optimise performance
- Weaknesses in ATLAS root file layouts were identified (affected remote i/o drastically)
The Data Cometh

- 20 Nov 2009: First collisions in Atlas $\sqrt{s} = 900$ GeV
- 6 Dec 2009: LHC “stable beams”: Inner detector at nominal voltage.
- 8 Dec 2009: LHC world record $\sqrt{s} = 2.36$ TeV collisions
- 30 March 2010: $\sqrt{s} = 7$ TeV collisions
- Steep rises in LHC delivered luminosity
Visual Overview

Simon George: “ATLAS High Level Trigger...”. This track, today 1120

Peter Onyisi: “Operation of ATLAS detector with first collisions at 7TeV...” This track, today 1500
Data Quality

- Online data quality per subdetector is loaded into COOL, along with LHC status
- DQ resolution is per-lumi block (2 minutes)
- Tier-0 prompt reconstruction populates histograms every 10 minutes for further DQ assessment
- This means DQ can be assessed efficiently during long runs
- More than 20 000 histograms are generated on demand per run per stream and are cached for future use
# Data Quality Display

- Automated and manual checks used

## LBs

<table>
<thead>
<tr>
<th>LBs</th>
<th>PIX0</th>
<th>PIXB</th>
<th>PIXEA</th>
<th>PIXEC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Run 155669</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 → 253</td>
<td>0.0 % 0, 286</td>
<td>0.0 % 0, 1170</td>
<td>0.0 % 0, 144</td>
<td>0.0 % 0, 144</td>
</tr>
<tr>
<td>253 → 254</td>
<td>5.7 % 67, 1170</td>
<td>1.4 % 2, 144</td>
<td>66.7 % 96, 144</td>
<td></td>
</tr>
<tr>
<td>254 → 255</td>
<td>85.5 % 1000, 1170</td>
<td>96.5 % 139, 144</td>
<td></td>
<td></td>
</tr>
<tr>
<td>255 → 256</td>
<td>97.1 % 1136, 1170</td>
<td>97.9 % 141, 144</td>
<td></td>
<td></td>
</tr>
<tr>
<td>256 → 257</td>
<td>1.0 % 3, 286</td>
<td>97.9 % 280, 286</td>
<td></td>
<td></td>
</tr>
<tr>
<td>257 → 312</td>
<td>97.9 % 280, 286</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>312 → 313</td>
<td>14.3 % 41, 286</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>313 → 314</td>
<td>67.3 % 787, 1170</td>
<td>72.9 % 105, 144</td>
<td></td>
<td></td>
</tr>
<tr>
<td>314 → 317</td>
<td>0.0 % 0, 286</td>
<td>0.0 % 0, 1170</td>
<td>0.0 % 0, 144</td>
<td>0.0 % 0, 144</td>
</tr>
</tbody>
</table>
**Final Data Quality**

- Inefficiencies dominated by ‘warm starts’ after LHC declares stable beams
- Overall efficiency is 95%

<table>
<thead>
<tr>
<th>Inner Tracking Detectors</th>
<th>Calorimeters</th>
<th>Muon Detectors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pixel</td>
<td>LAr EM</td>
<td>MDT</td>
</tr>
<tr>
<td>SCT</td>
<td>LAr HAD</td>
<td>97.9</td>
</tr>
<tr>
<td>TRT</td>
<td>LAr FWD</td>
<td>96.1</td>
</tr>
<tr>
<td></td>
<td>Tile</td>
<td>98.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>97.4</td>
</tr>
</tbody>
</table>

Luminosity weighted relative detector uptime and good quality data delivery during 2010 stable beams at \( \sqrt{s} = 7 \) TeV between March 30\textsuperscript{th} and July 16\textsuperscript{th} (in %)
Calibration Loops

- Calibration runs on the express stream and calibration streams
  - Express stream is ~10% of data, including high $P_T$ lepton and jet triggers
  - Calibration stream contains partially built events from calibration triggers
- Suppression of noisy channels for physics reconstruction
Calibration of Beam Spot

- After prompt reconstruction updated calibration constants are used for physics streams
- Nominal time for whole calibration loop is 36 hours with a manual signoff
Tier-0

- ATLAS Tier-0 plays a pivotal role:
  - Accept data from online and ensure it’s archived to tape
  - Process express, calibration and physics streams
  - Export data to Tier-1 and calibration Tier-2s, as well as CAF
  - Data has to be registered in ATLAS Distributed Data Management system
Tier-0 Workflow

- This display of workflows is also the shifter interface
- Boxes turn amber or red when there is a problem
T0 Design Highlights

• Robust handshake with online systems
• RAW merging and archive to tape
• Solid framework for running reconstruction
• T0 is 99.997% efficient
• Pool size 65 servers, 650TB, but main design criterion is i/o capacity of 6GB/s
• Additional merge pool for high availability of pre-merged data
# Tier-0 7TeV Statistics

<table>
<thead>
<tr>
<th>Data Type</th>
<th># Datasets</th>
<th># Files</th>
<th># Events</th>
<th>Total Size [TB]</th>
</tr>
</thead>
<tbody>
<tr>
<td>RAW (physics)</td>
<td>2466</td>
<td>532993</td>
<td>492 M</td>
<td>797</td>
</tr>
<tr>
<td>RAW (express)</td>
<td>403</td>
<td>62575</td>
<td>36 M</td>
<td>63</td>
</tr>
<tr>
<td>RAW (calibration)</td>
<td>3697</td>
<td>66643</td>
<td>292 M</td>
<td>28</td>
</tr>
<tr>
<td>ESD</td>
<td>3941</td>
<td>979967</td>
<td>631 M</td>
<td>600</td>
</tr>
<tr>
<td>AOD</td>
<td>3900</td>
<td>52322</td>
<td>625 M</td>
<td>41</td>
</tr>
<tr>
<td>DPD</td>
<td>3627</td>
<td>85805</td>
<td>141 M</td>
<td>117</td>
</tr>
<tr>
<td>NTUP</td>
<td>7736</td>
<td>87900</td>
<td>1283 M</td>
<td>62</td>
</tr>
<tr>
<td>HIST</td>
<td>3815</td>
<td>3778</td>
<td>591 M</td>
<td>0.6</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>29585</strong></td>
<td><strong>1871983</strong></td>
<td></td>
<td><strong>1708.6</strong></td>
</tr>
</tbody>
</table>

- March 30 - July 21: Tier-0 has run 1.86M jobs consuming 243 years of CPU time
Reprocessing

• Reprocessing occurs at Tier-1 sites, instead of Tier-0
• Two types: ‘fast’ and ‘full’
  • Fast uses software already known to be good
  • Full uses new versions of athena
• The aim here is for ‘best’ calibration constants, best software and 100% reprocessing success
Reprocessing in Practice

• With ten Tier-1s involved there’s lots of scope for problems
• Operationally heavy
  • But sites do respond
• ATLAS Distributed Computing team successful in achieving 100% of events processed in April and May

---

ICHEP 2010, Paris
Recent Improvements

- Take a vertical slice through the data processing to pickup any unexpected problems in later stages
- Setup ‘hospital queues’ at Tier-1s to deal with tricky events
Early Running Processing and Reprocessing

- RAW from the detector was put on disk at Tier-1s to help reprocessing.
- After May reprocessing Tier-0 went into a software freeze.
  - No changes to physics content of outputs.
  - Allows Tier-0 processed data to be merged into existing plots.
- Next reprocessing foreseen ~September with Athena 16.0.0.
Data Distribution

- In concert with data reprocessing we reprocess MC to assure consistency.
- This leads to large volumes of data which need to be distributed after reprocessing campaigns.
- This takes a long time!
- Can lead to delays in ‘interesting’ data arriving.

Disk Usage Ramp up an T1s

ICHEP 2010, Paris
Analysis

- ATLAS has been able to sustain continued high rates of analysis across the grid since LHC running began
- The system continues to scale up well
- Site tuning is a continuous process
Future Improvements

- Data Quality: Further automation of DQ signoff
- Tier-0: Better coupling to external components which can destabilise system
- Data Distribution: Distribute ‘interesting’ data more widely in an automatic way
- Analysis: Better (re)brokering of users jobs and automation of masking problematic sites
Conclusions

• After a long preparation ATLAS data preparation and computing were in a good state when LHC delivered data

• End to end systems from Tier-0, through data quality and calibration, to physics analysis are working well

• Systems will continue to evolve and improve

• Looking forwards to the challenges of more LHC data