ATLAS Computing: From Commissioning to 7TeV Data

Graeme Stewart
for the ATLAS collaboration

University of Glasgow | Department of Physics & Astronomy

ICHEP 2010, Paris

GridPP
UK Computing for Particle Physics

Outline

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

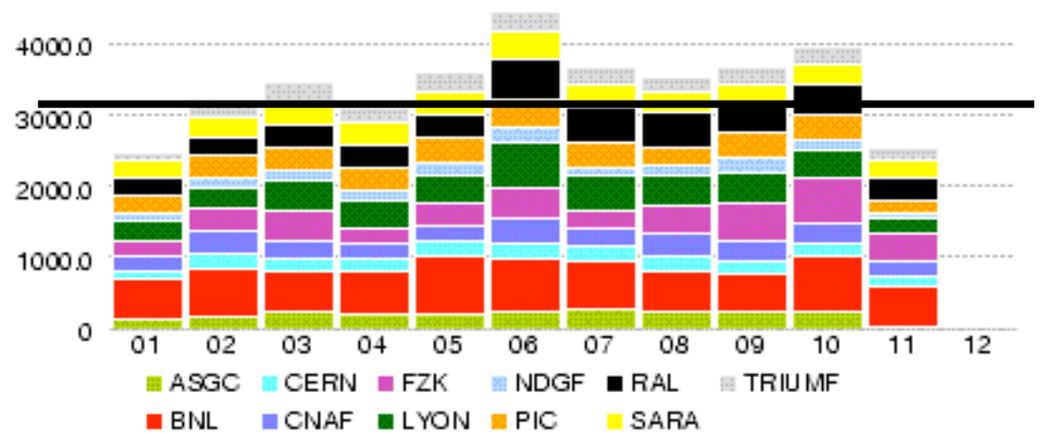
STEP09 Exercise

- 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 → Tier-1 → Tier-2
- Reprocessing at Tier-1s (from tape)
- Large scale analysis activity at Tier-2s
- Full scale simulation activity going on

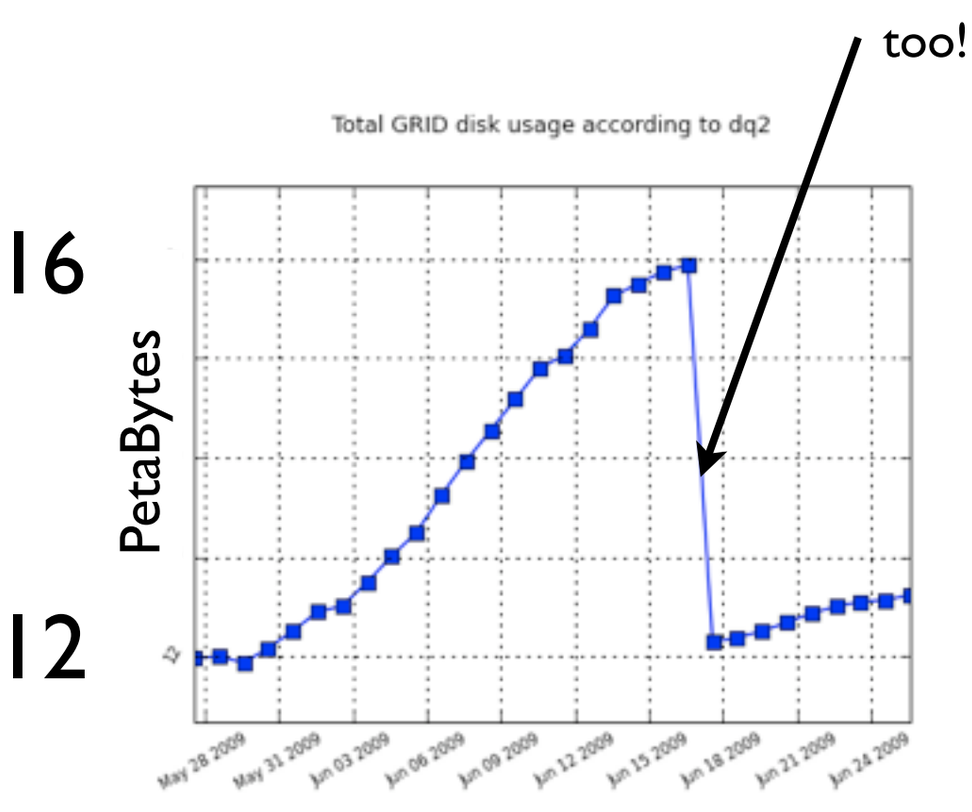
STEP 09 Data Distribution

- Data distribution to T1s and (most) T2s worked

3GB/s



We can delete too!



		Transfers
Cloud	Efficiency	Throughput
ASGC	99%	397 MB/s
BNL	84%	1128 MB/s
CERN	100%	334 MB/s
CNAF	98%	561 MB/s
FZK	85%	556 MB/s
LYON	96%	620 MB/s
NDGF	84%	137 MB/s
PIC	93%	429 MB/s
RAL	99%	838 MB/s
SARA	53%	262 MB/s
TRIUMF	100%	297 MB/s

Peaks of 5.5GB/s

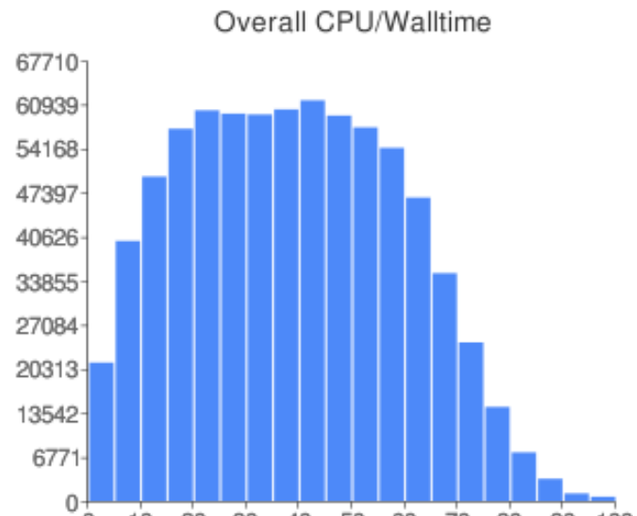
STEP09 Tier-1 Reprocessing

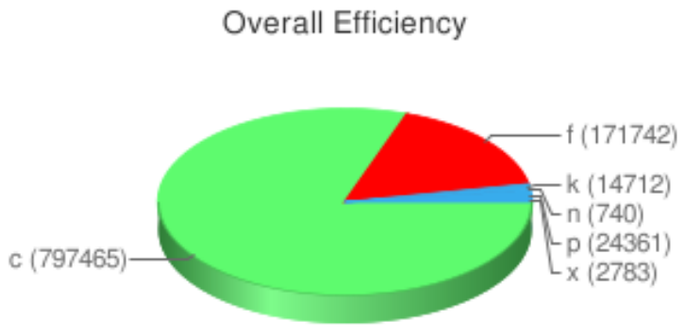
- 6/10 Tier-1s validated
- 3/10 Close
- Problems were generally understood
- But system shown to be complex and somewhat fragile

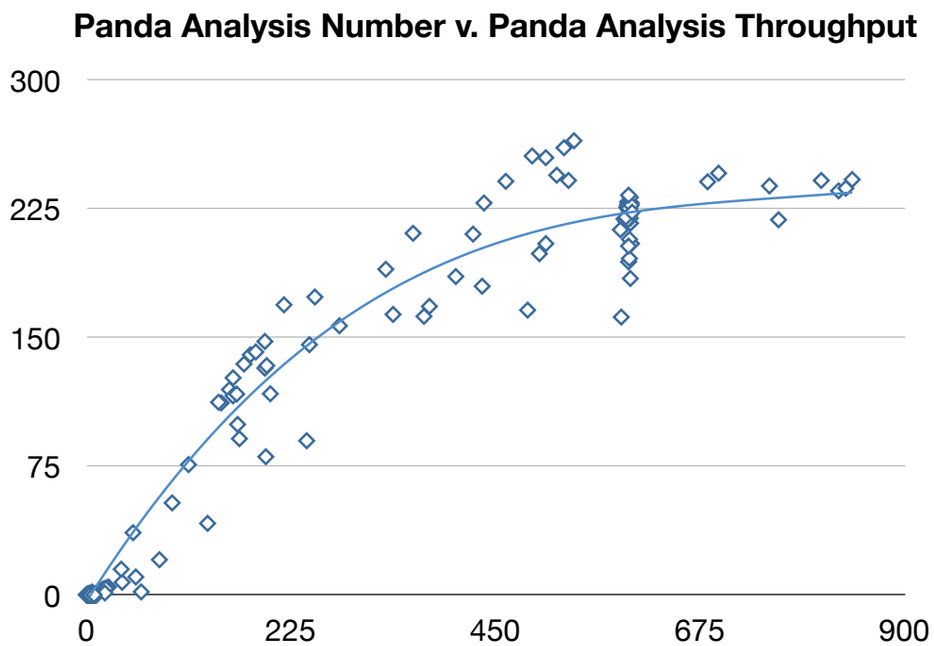
T1	*Base Target*	*Result*
ASGC	10 000	4 782
BNL	50 000	99 276
CNAF	10 000	29 997 ☆
FZK	20 000	17 954
LYON	30 000	29 187
NDGF	10 000	28 571 ☆
PIC	10 000	47 262 ☆
RAL	20 000	77 017 ☆
SARA	30 000	28 729
TRIUMF	10 000	32 481 ☆

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)

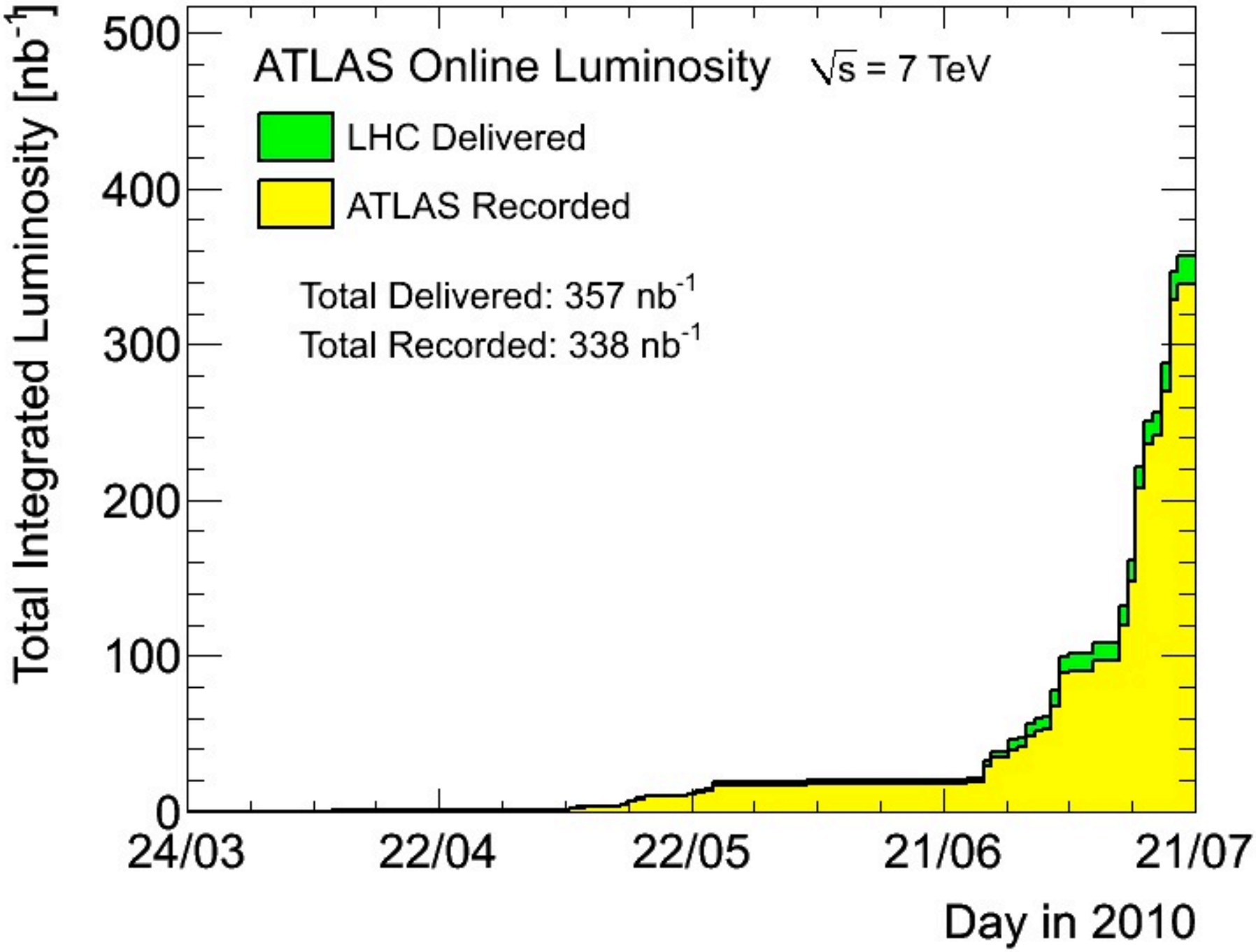






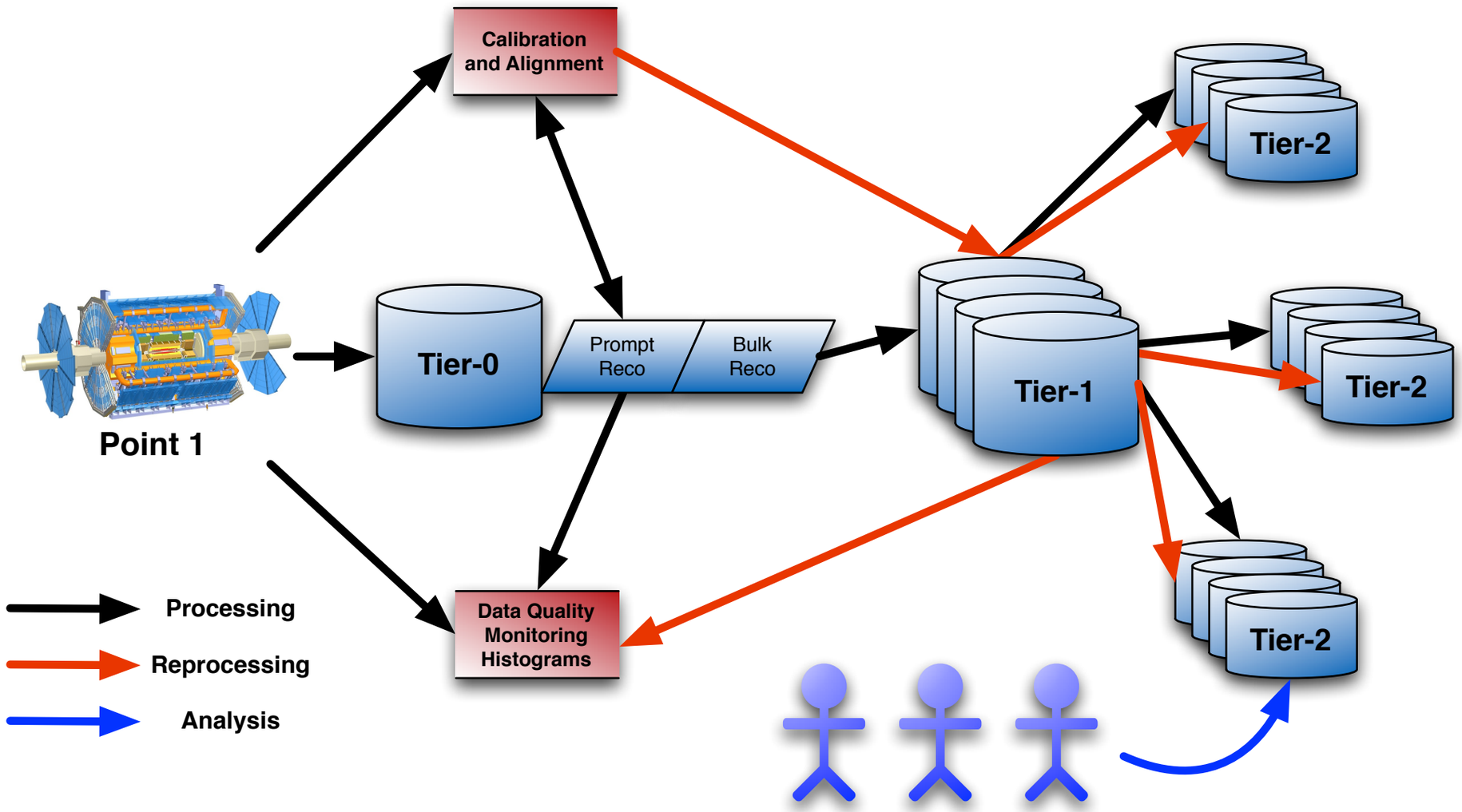
The Data Cometh

- 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

Unrecoverable

Maybe Correctable

Good for analysis

- Automated and manual checks used

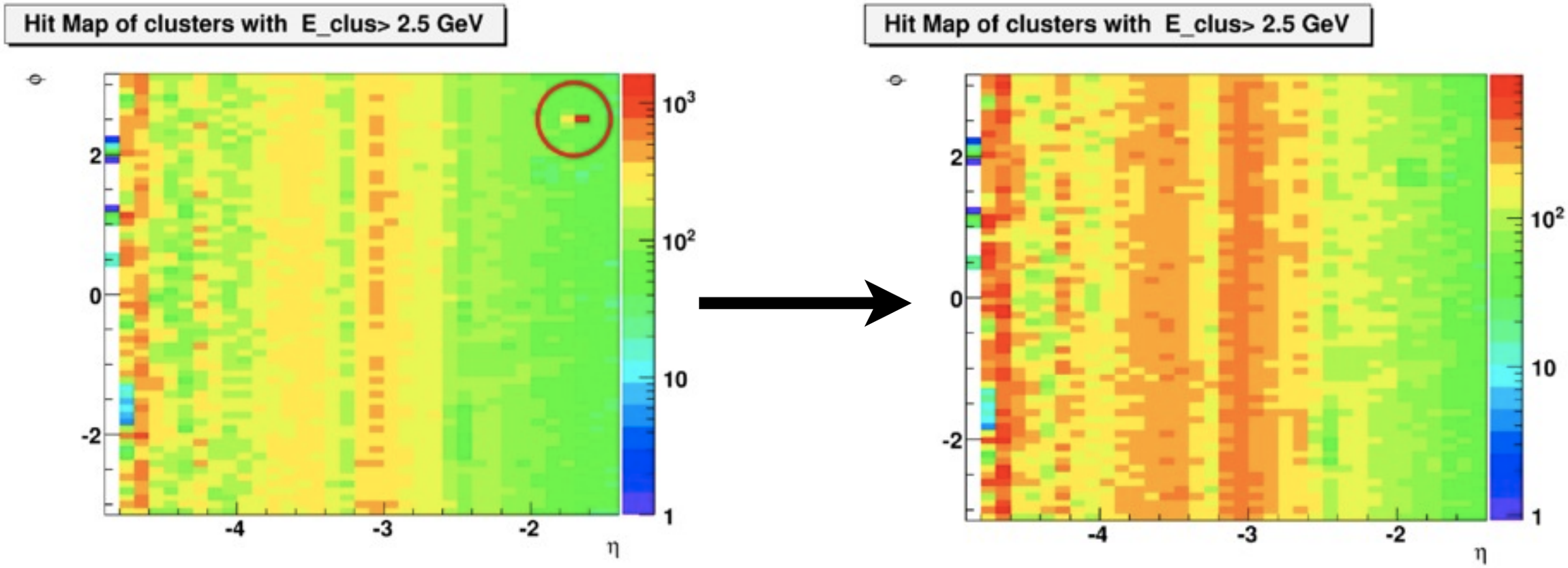
Final Data Quality

Inner Tracking Detectors			Calorimeters				Muon Detectors			
Pixel	SCT	TRT	LAr EM	LAr HAD	LAr FWD	Tile	MDT	RPC	TGC	CSC
97.1	98.2	100	93.8	98.8	99.1	100	97.9	96.1	98.1	97.4

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

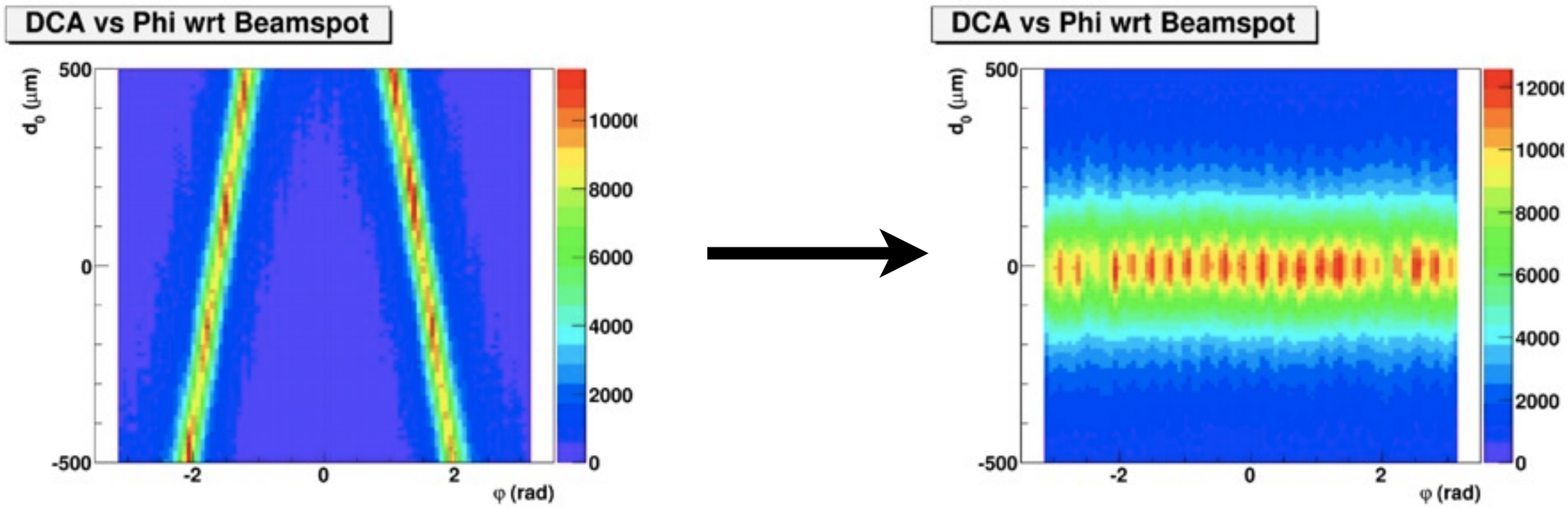
- Inefficiencies dominated by 'warm starts' after LHC declares stable beams
- Overall efficiency is 95%

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

Peter Waller: "ATLAS Data Quality Monitoring...". Poster

Michael Böhler: "Processing, Calibration and Reprocessing of ATLAS Data...". Poster

David Miller: "Luminosity and Beam Spot Determination...". Poster

Tier-0

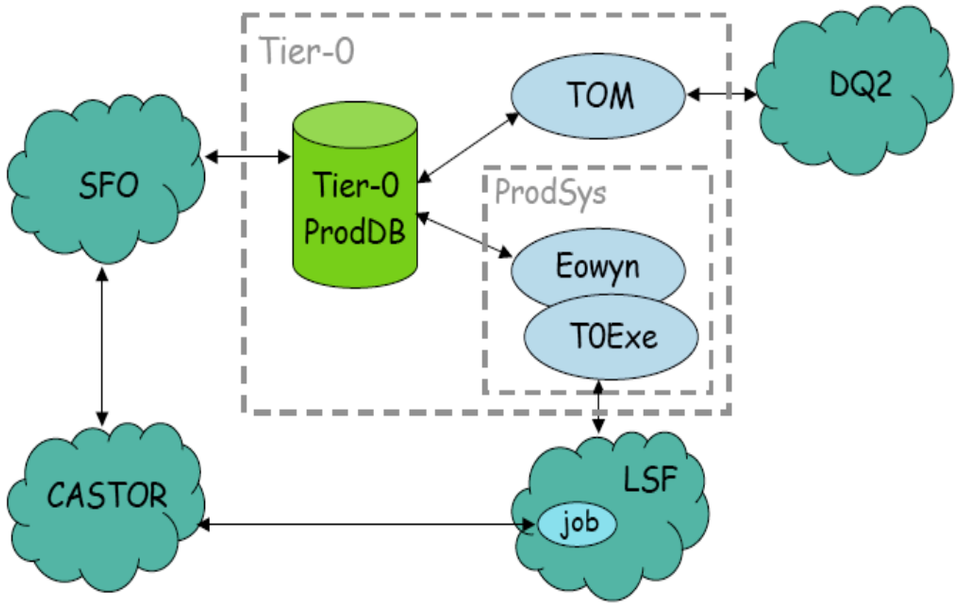
- ATLAS Tier-0 plays a pivotal role:
  - Accept data from online and ensure it's archived to tape
  - Process express, calibration and physics streams
  - Export data to Tier-1 and calibration Tier-2s, as well as CAF
    - Data has to be registered in ATLAS Distributed Data Management system

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



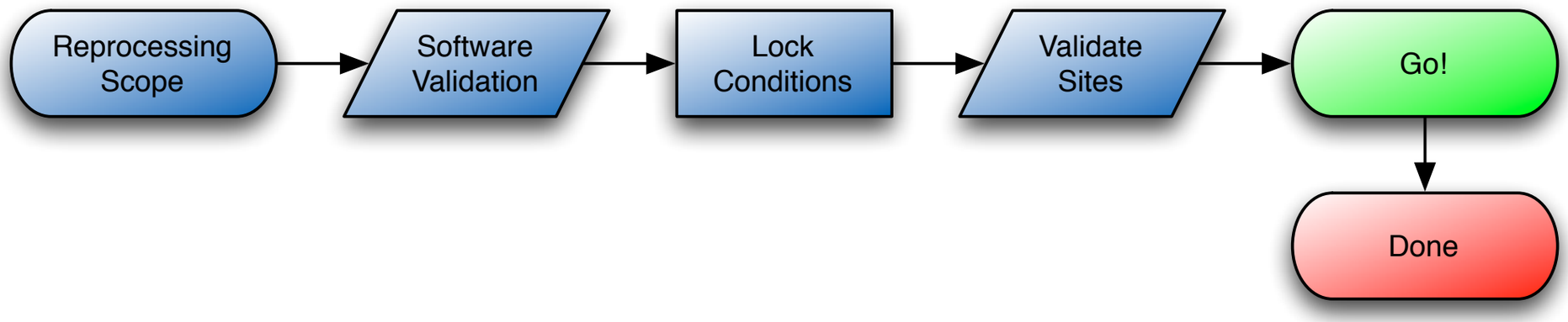
Tier-0 7TeV Statistics

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

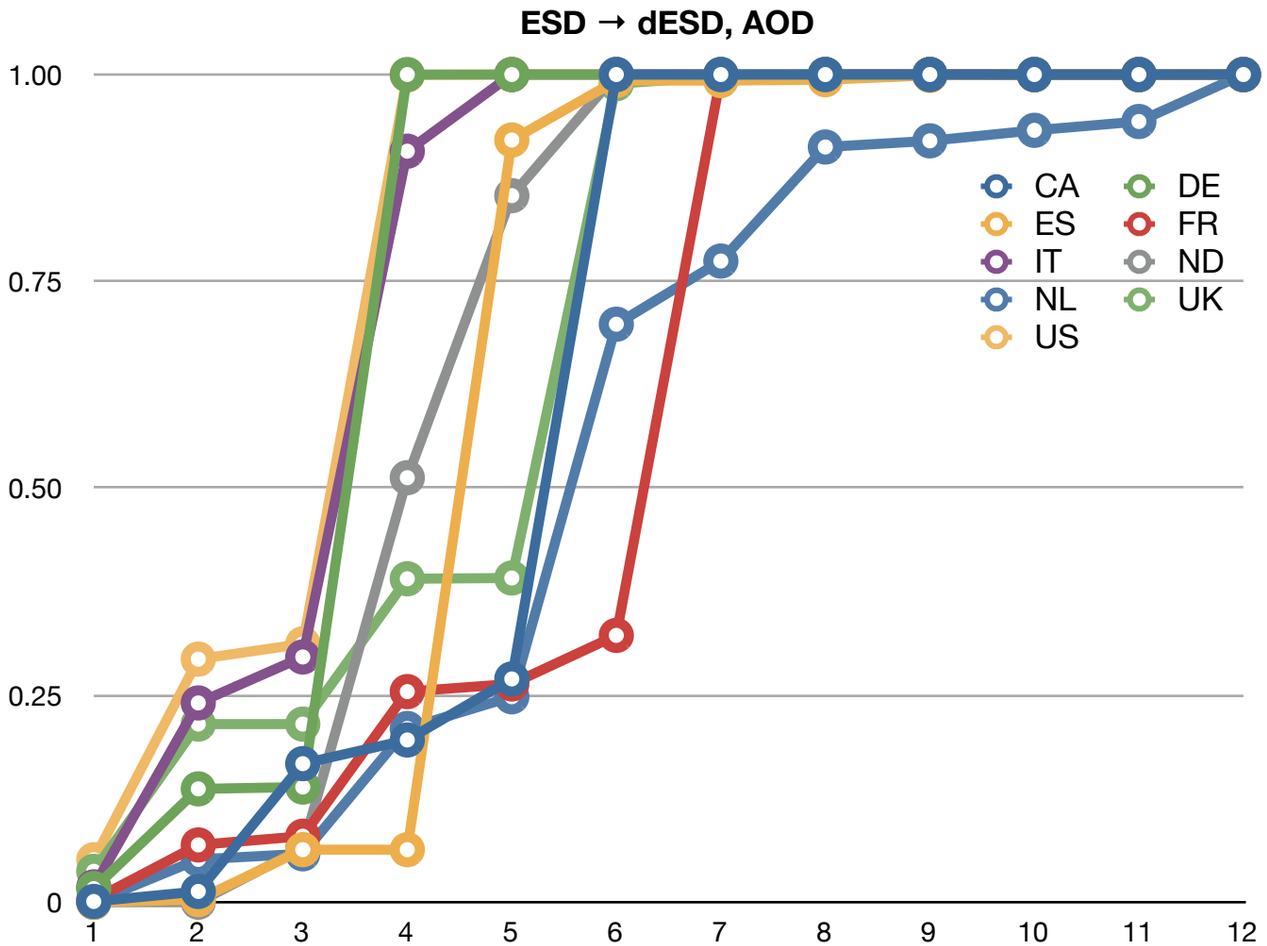
- March 30 - July 21: Tier-0 has run 1.86M jobs consuming 243 years of CPU time

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

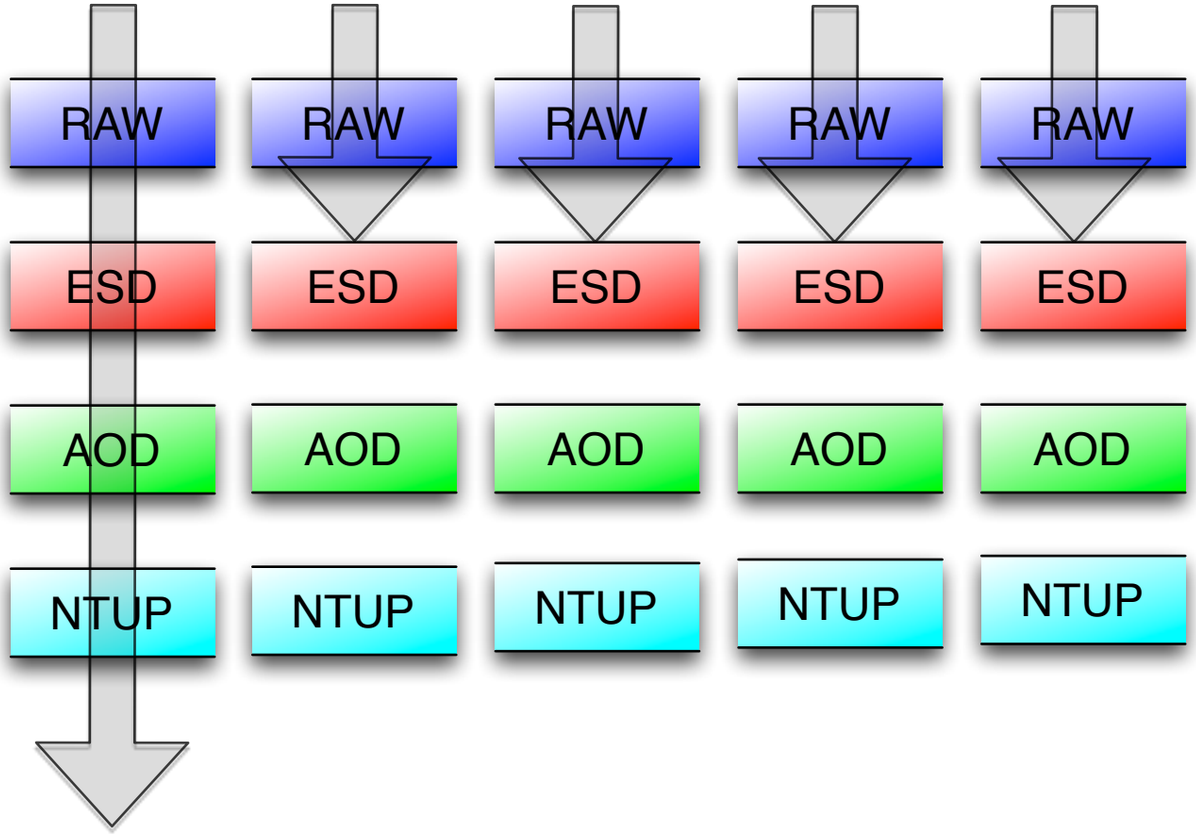


- With ten Tier-1s involved there's lots of scope for problems
- Operationally heavy
  - But sites do respond
- ATLAS Distributed Computing team successful in achieving 100% of events processed in April and May

total jobs	9577	9540	7233	13375	1964	6886	26676	19252	25197	119700
total done	9577	9540	7233	13375	1964	6886	26676	19252	25197	119700
% %	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

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

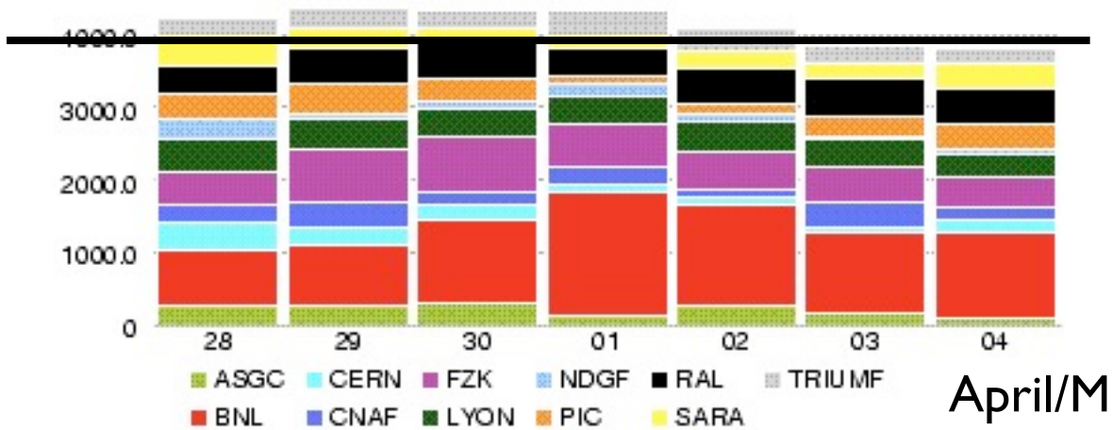


Early Running Processing and Reprocessing

- 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

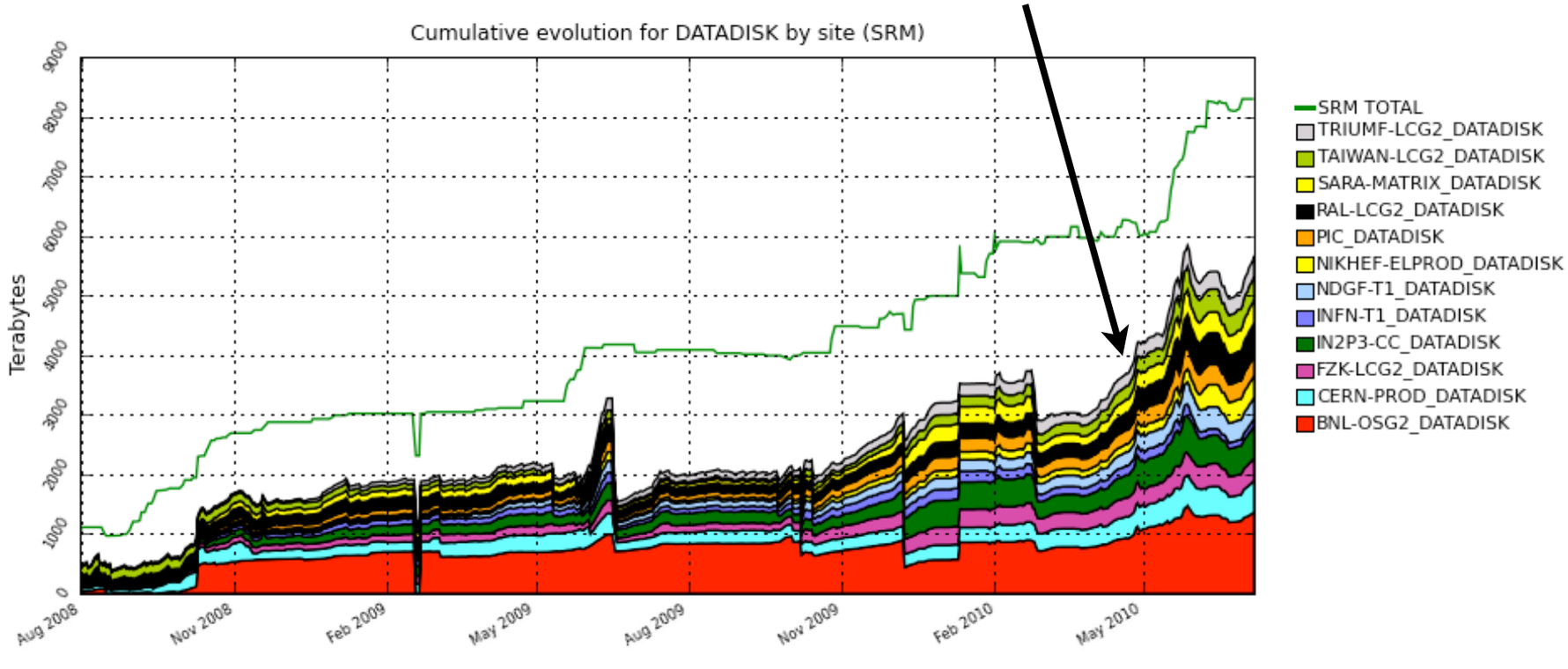
- In concert with data reprocessing we reprocess MC to assure consistency
- This leads to large volumes of data which need to be distributed after reprocessing campaigns
  - This takes a long time!
- Can lead to delays in 'interesting' data arriving



4GB/s

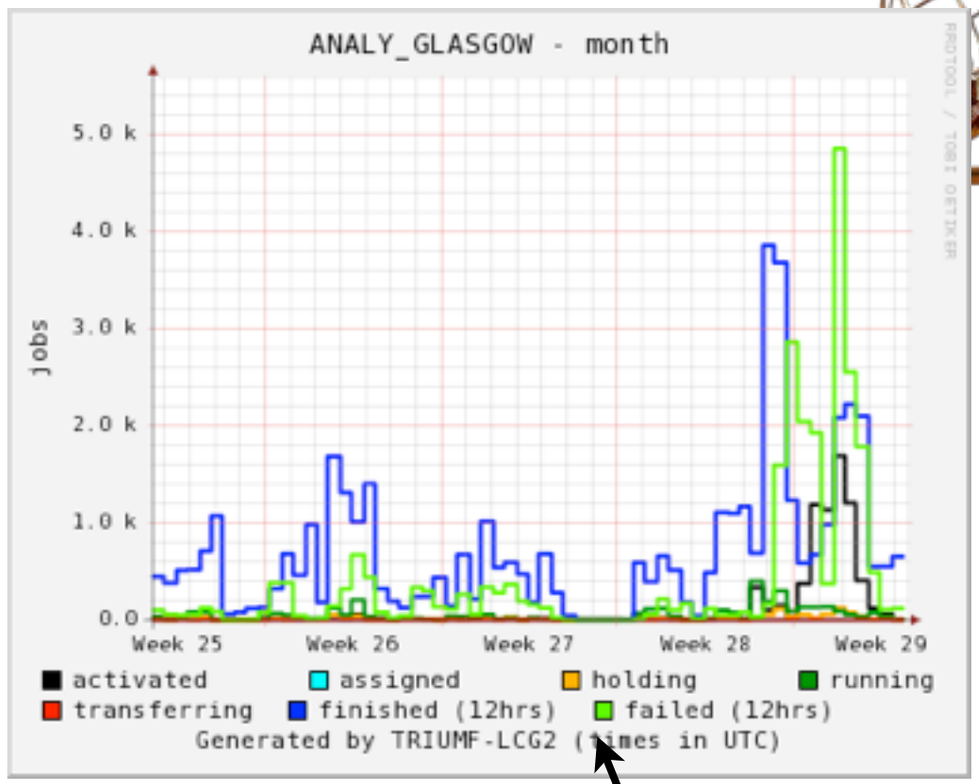
April/May 2010

Disk Usage Ramp up an T1s

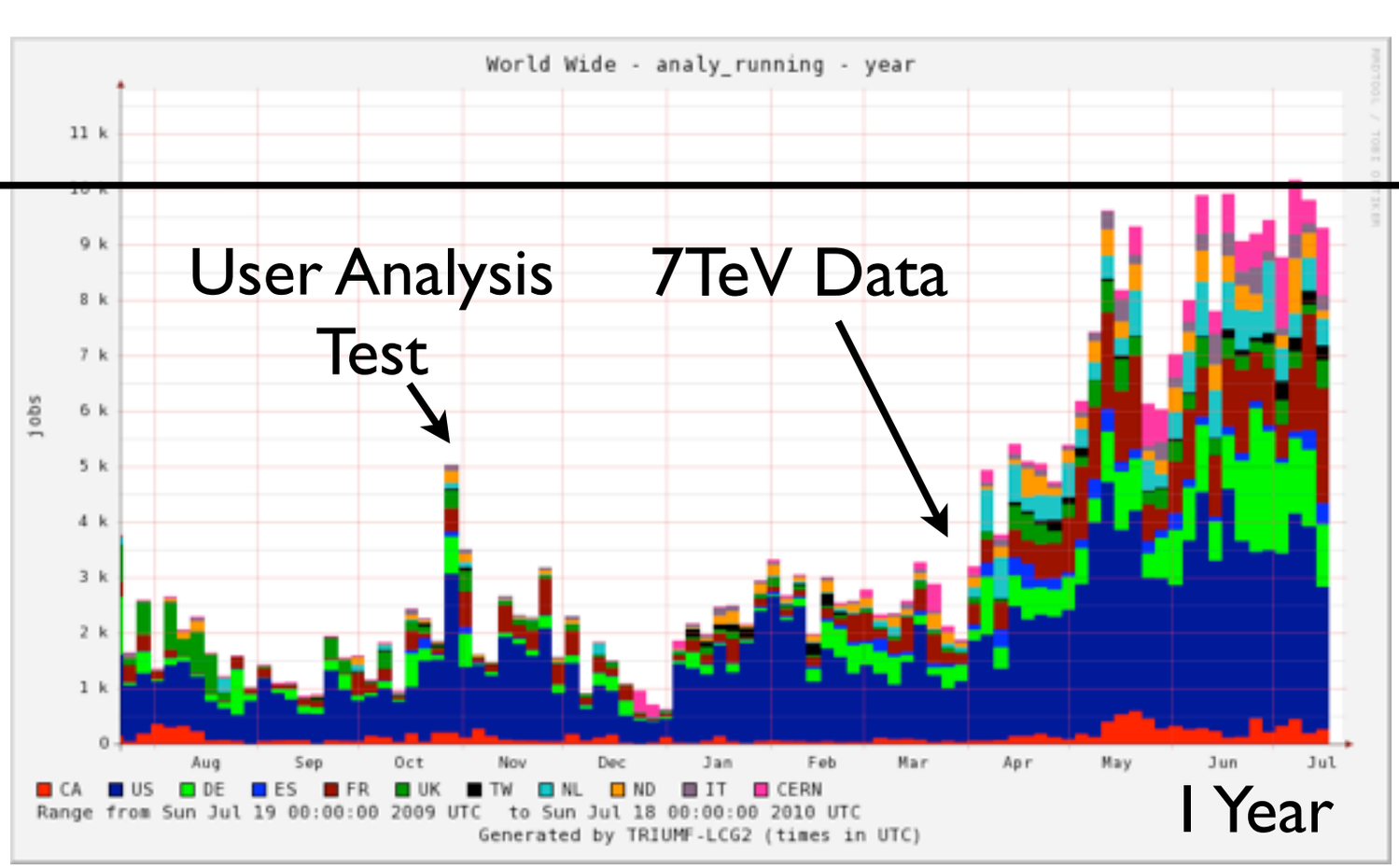


Analysis

- 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



10k running jobs



Actual per-site profile is spiky

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

Andrey Loginov: “ATLAS Upgrade for the sLHC...”. This track, 1715

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