ATLAS Quarterly Report and Plans

Dario Barberis

CERN & Genoa University/INFN

Outline

- Tier-0 and data-taking activities

- Data distribution

- Data reprocessing

- Simulation Production

- Distributed analysis

- Plans

Tier-0 and data-taking activities

- October 2009: global cosmics runs

- Mid-November: start of LHC data

  - Open trigger (low thresholds)

  - Full calorimeter read-out
    - 5 MB/event on average

  - Instantaneous rate limited to 800 MB/s, but average event rate very low
    - Large RAW, but small ESD etc.

- Cosmics runs are interleaved with LHC runs when the machine is off

  - They are needed (together with beam halo) for detector alignment to constrain the weak distorsion modes that cannot be constrained by tracks originating from the collision point

- Accumulated almost 1 PB including replicas

- All data processed in real time at Tier-0
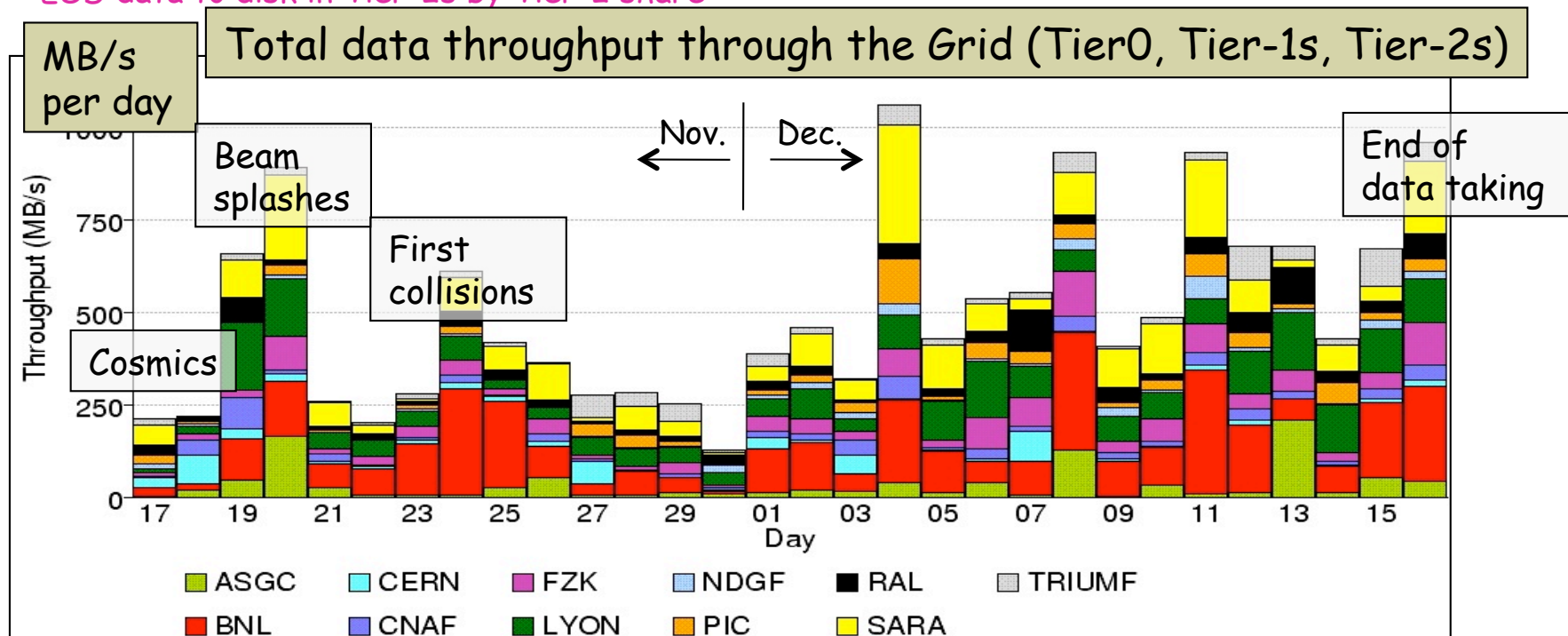
  - No surprises wrt MC events

Dario Barberis:
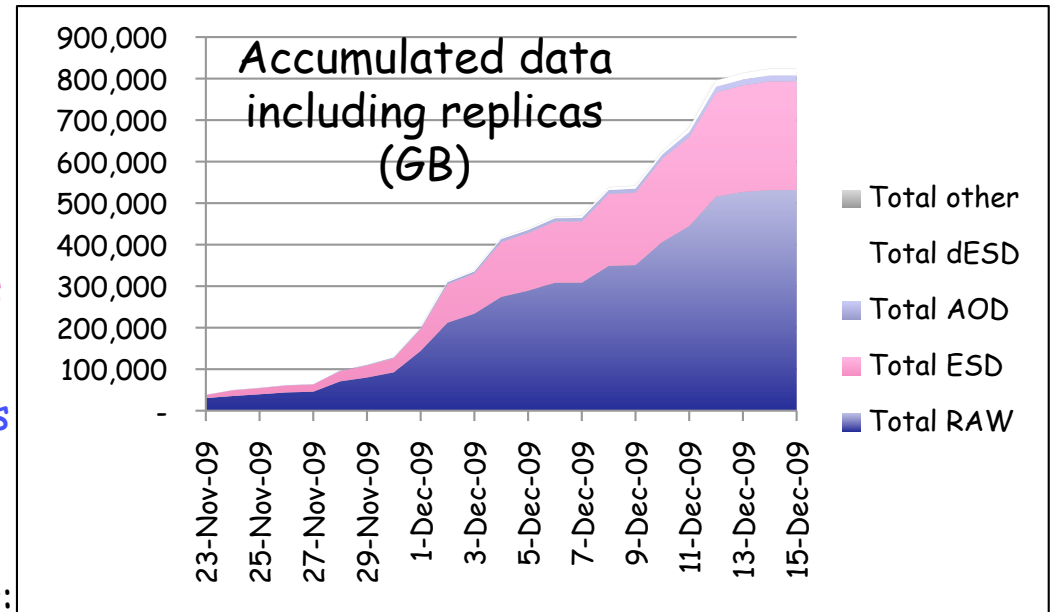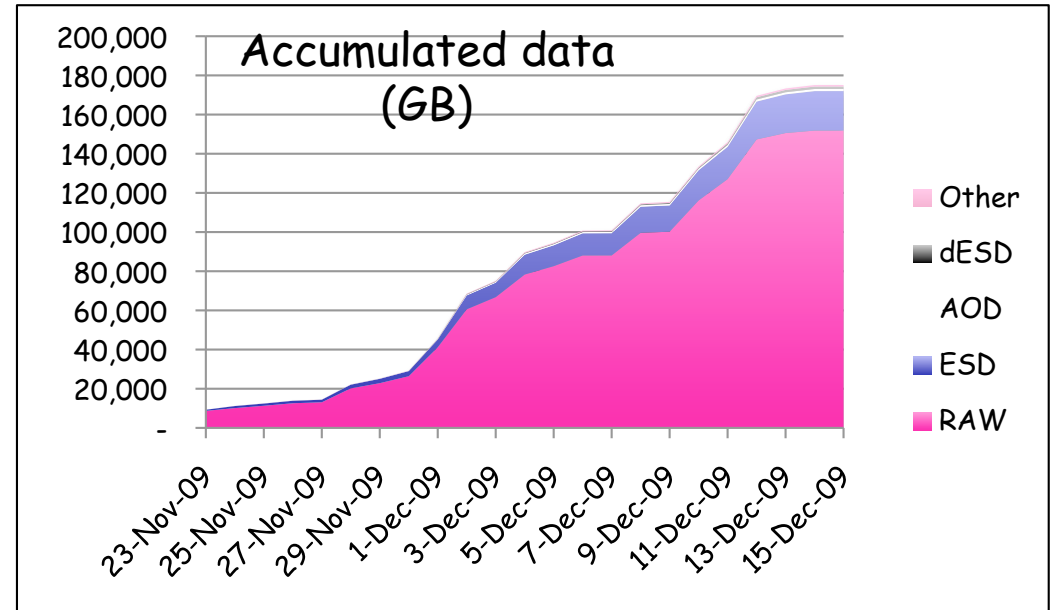
Data distribution pattern

- All RAW to disk and tape in each Tier-1 by Tier-1 share
  - Moreover all RAW go to disk at BNL, Lyon and SARA
  - *Normal is tape in each Tier-1 by Tier-1 share*
  - No extra RAW data to disk at CERN except CAF
- All ESD to disk in each Tier-1
  - *Normal is 2 copies distributed over all Tier-1s*
  - Full ESD copy to disk at CERN
  - ESD data to disk in Tier-2s by Tier-2 share

- AOD and dESD to disk in all Tier-1s
  - *Normal is 2 copies kept in all Tier-1s only*
  - Copied to disk in Tier-2s by Tier-2 share (total ~18 copies)
  - *Normal is 10 copies in the Tier-2s only*
- Additional copies will be reduced dynamically to make room for 2010 data

MB/s per day

Total data throughput through the Grid (Tier0, Tier-1s, Tier-2s)

Beam splashes

First collisions

Cosmics

Nov. ← | → Dec.

End of data taking

Throughput (MB/s)

1000 — 750 — 500 — 250 — 0

Day: 17 19 21 23 25 27 29 01 03 05 07 09 11 13 15

ASGC | CERN | FZK | NDGF | RAL | TRIUMF
BNL | CNAF | LYON | PIC | SARA
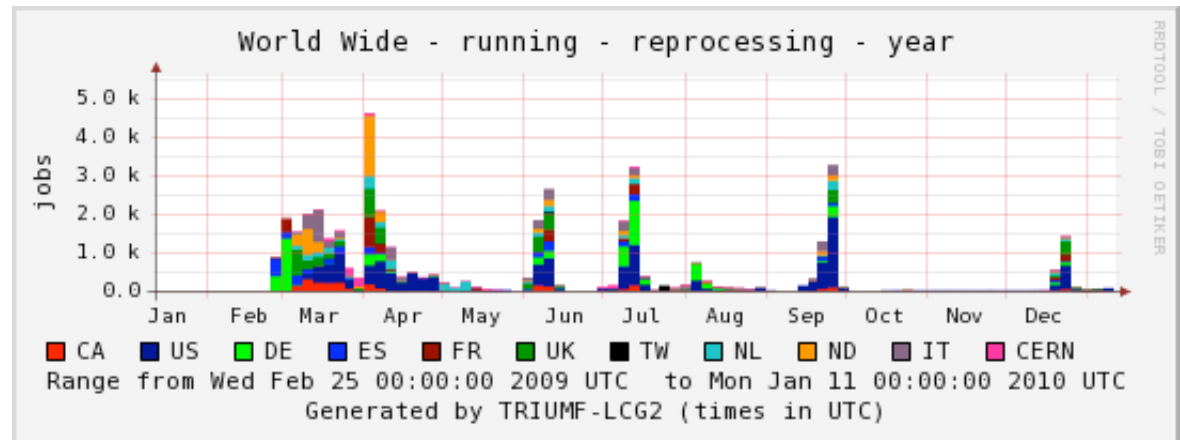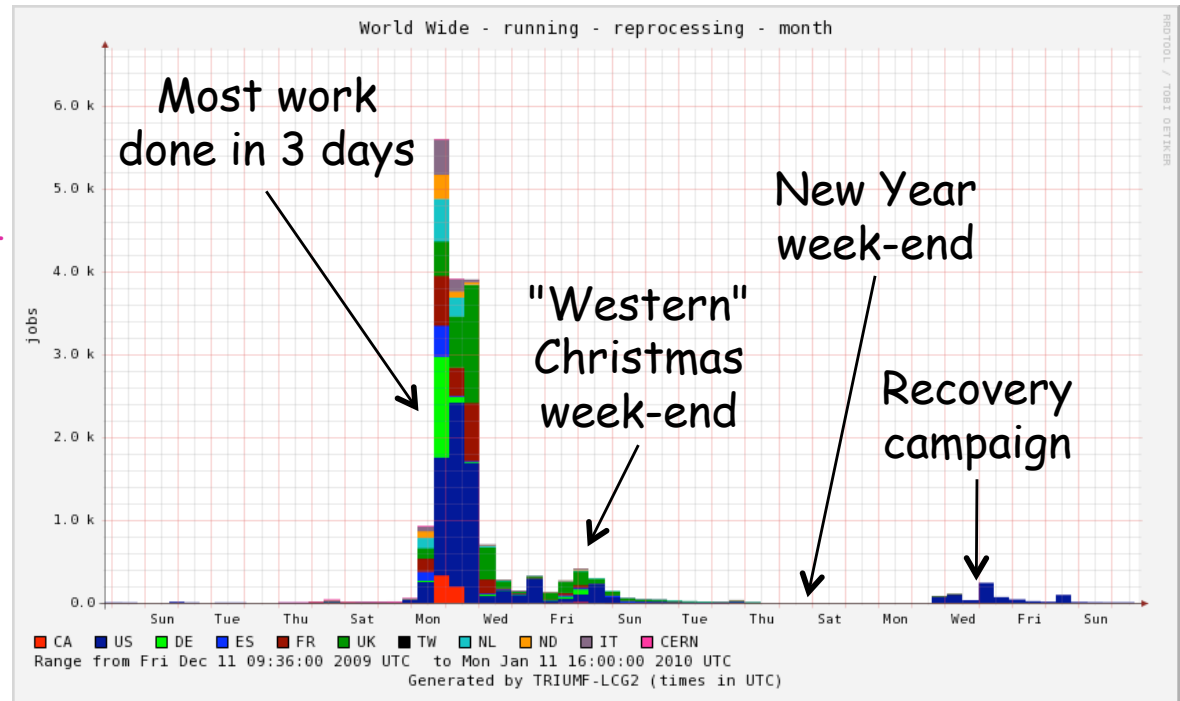
4

Data distribution performance

- All data were delivered to Tier-1s and Tier-2s using open datasets
  - RAW during data-taking (run in progress)
  - ESD etc during Tier-0 processing, as soon as outputs were available
- Data were available for analysis at Tier-2s on average 4 hours after data-taking
  - Including the time for Tier-0 processing
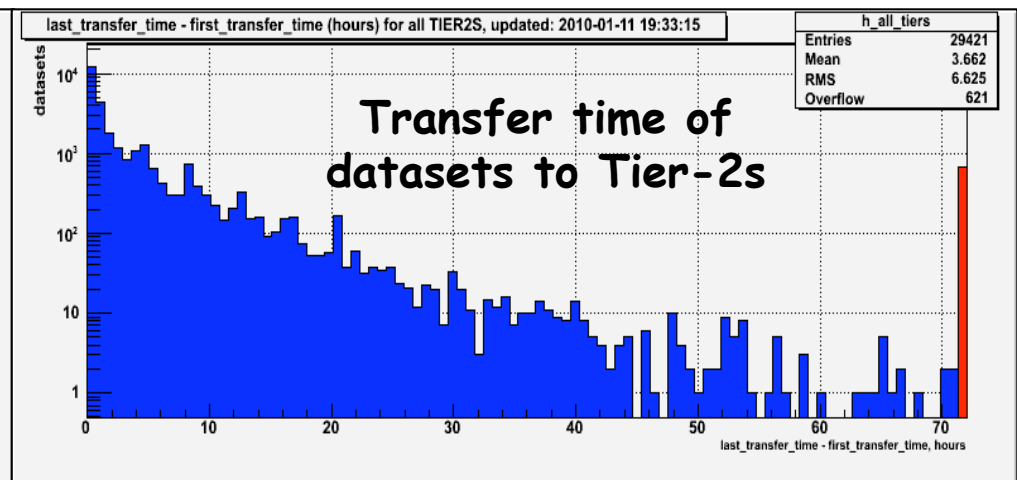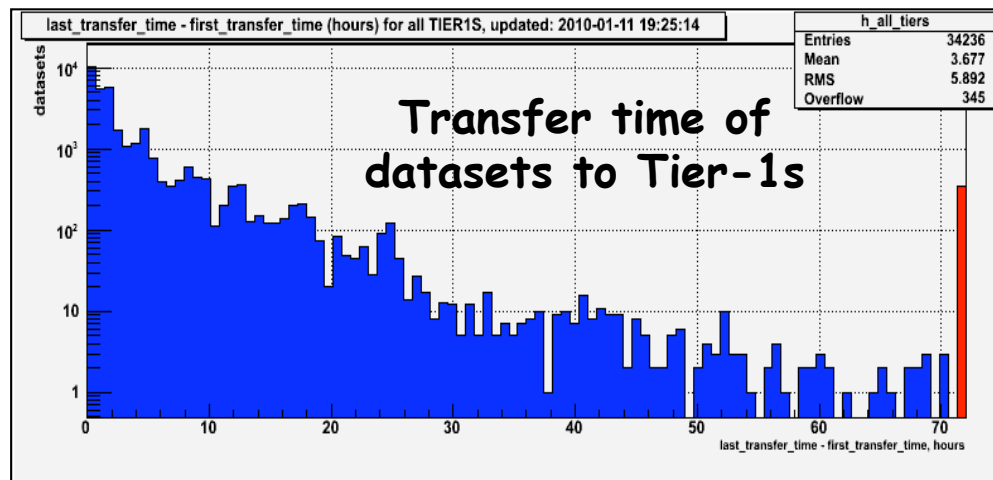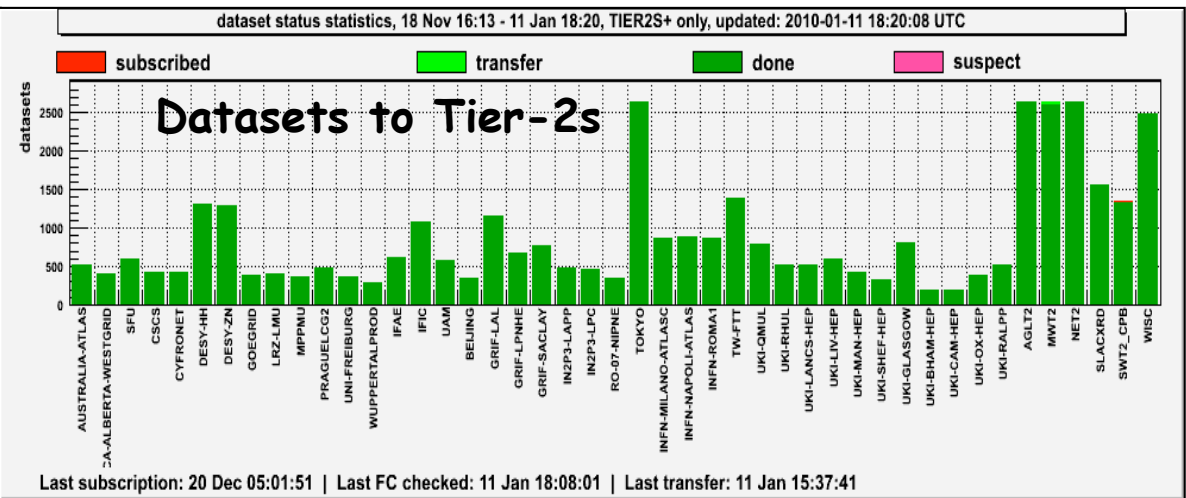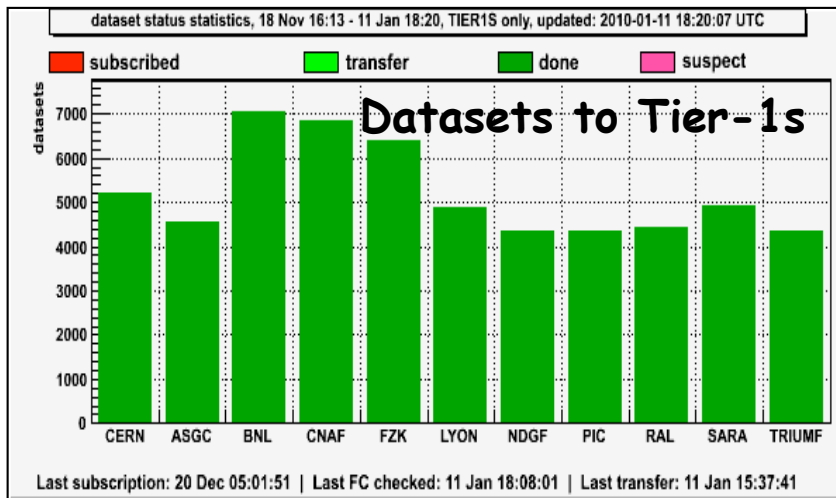
Data reprocessing

- An "ultra-fast" reprocessing campaign was run on 21-31 December 2009
  - Using the last Tier-0 software cache plus a few last-minute bug fixes (release 15.5.4.10) and most up-to-date calibrations and alignments for the whole period
- Thanks to site people for their support!
- Only 22 RAW->ESD jobs failed out of 130148 and 27 ESD->AOD jobs out of 10001
  - A few software bugs being followed up, affecting beam splash events
- Next reprocessing round will take place in February
  - Using release 15.6.3.X built now
  - Will also be a test of releases 15.6.X.Y to be used at Tier-0 next month
  - SLC5/gcc4.3 only!!!

Most work done in 3 days

"Western" Christmas week-end

New Year week-end

Recovery campaign
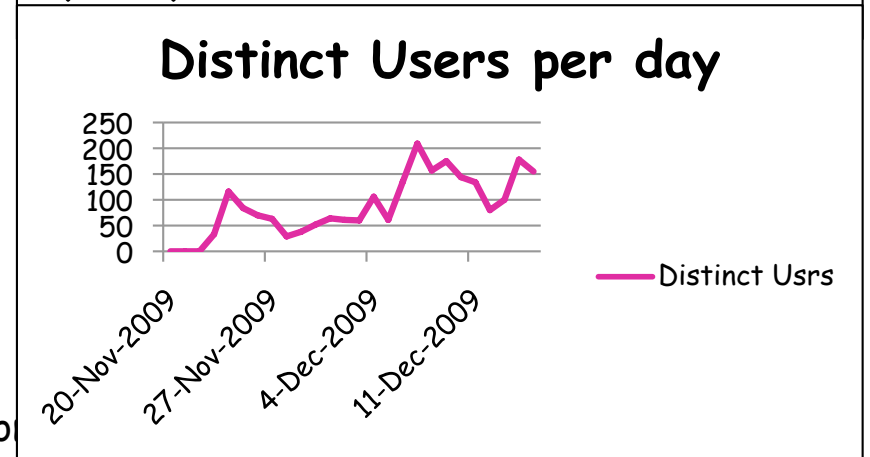
Dario Barberis: ATLAS Computing
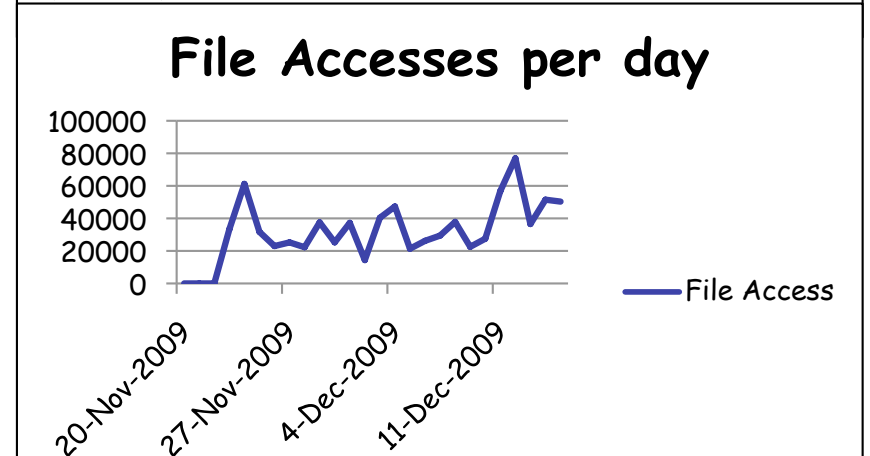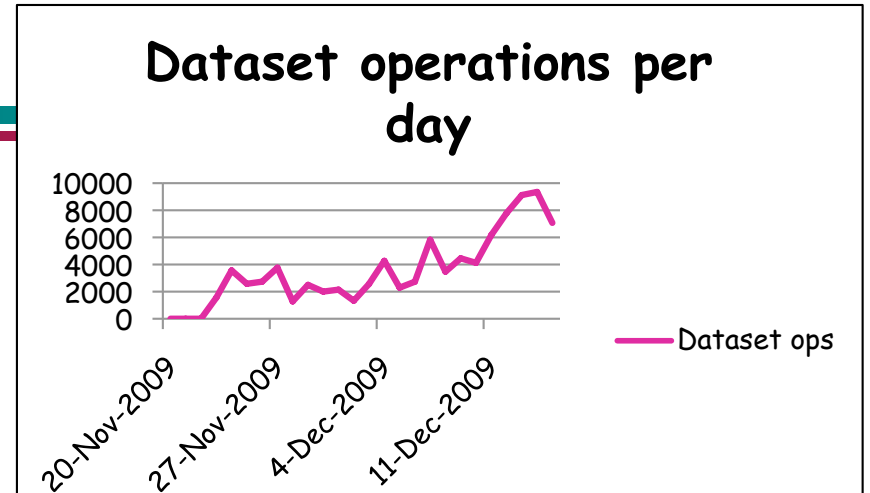
6

Simulation production

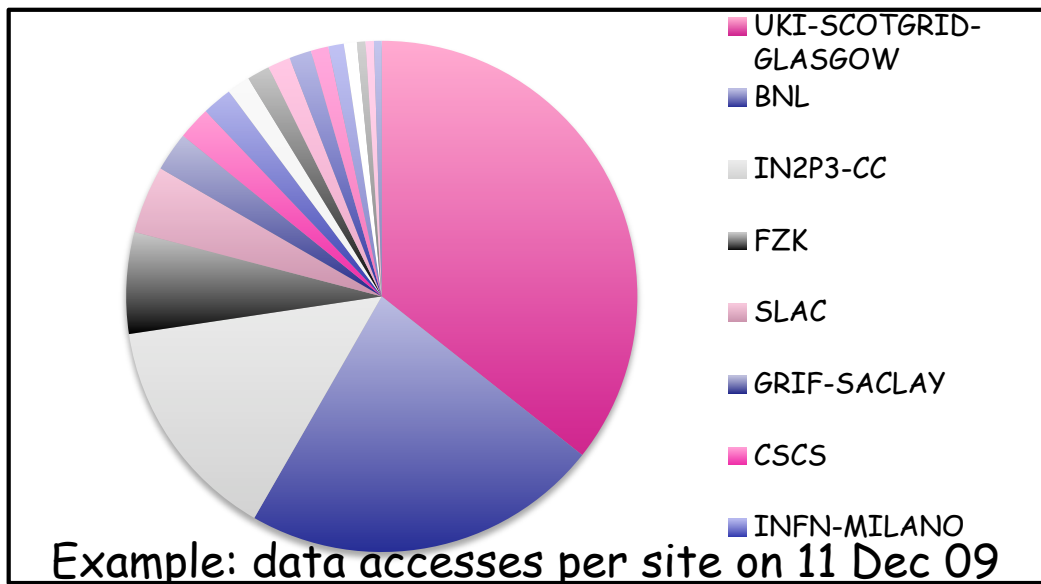- Simulation production continues in the background all the time
  - Only limited by physics requests and the availability of disk space for the output
- Parallel effort underway for MC reconstruction and reprocessing
  - Including reprocessing of MC09 900 GeV and 2.36 TeV samples with AtlasTier0 15.5.4.10 reconstruction, same release as for data reprocessing

Production jobs running in 2009 in each cloud

7
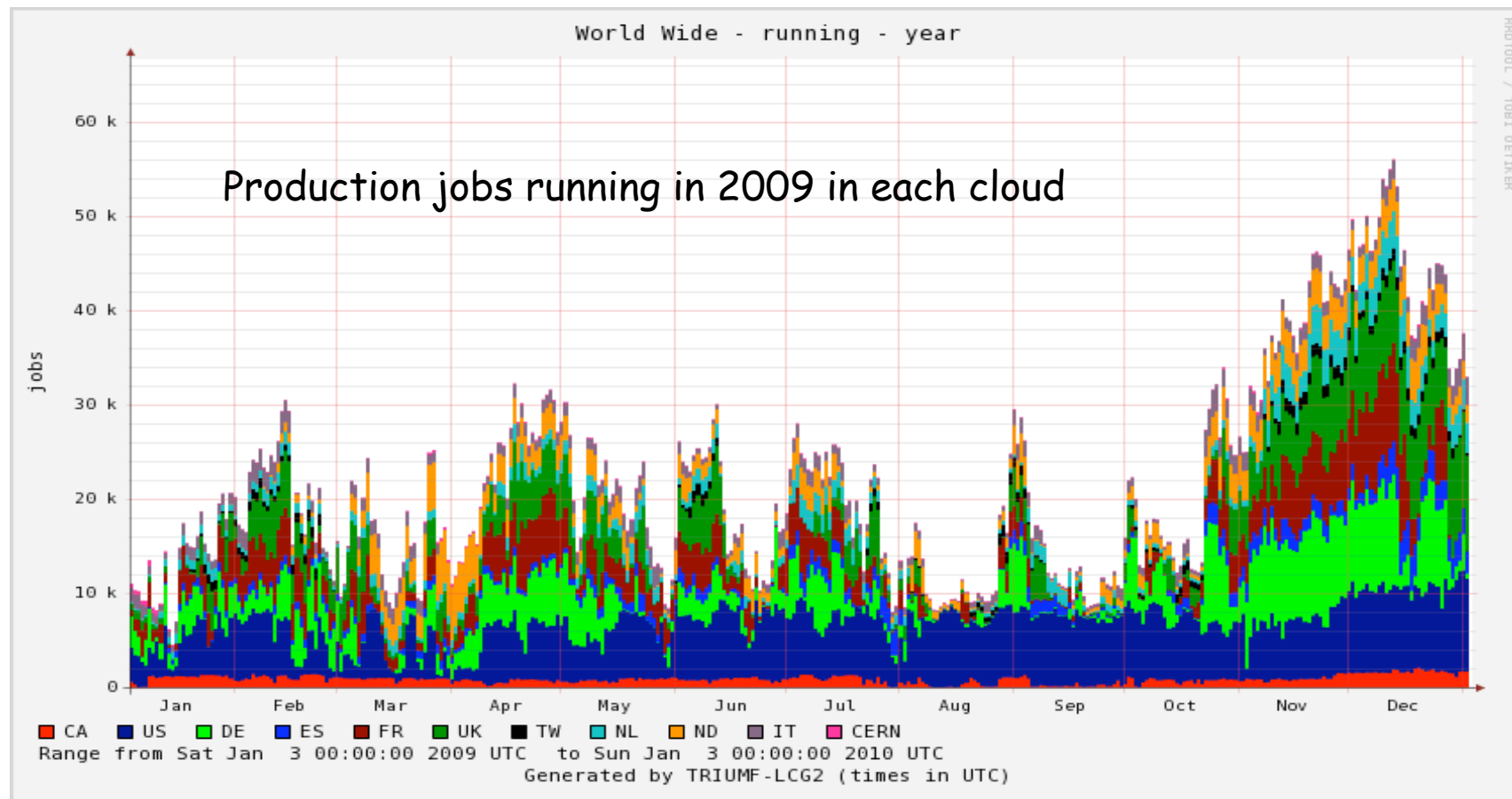
Analysis data access

- Data were analysed on the Grid already from the first days of data-taking
    - See plots
- Several "single" users submitted event selection and/or analysis jobs on behalf of their performance or physics working group
    - Outputs are ntuples that are copied to group space and then downloaded by end users
    - In this work model the number of real Grid users is somewhat underestimated

Dataset operations per day

— Dataset ops

File Accesses per day

— File Access

Distinct Users per day

— Distinct Usrs

Pie chart legend:
- UKI-SCOTGRID-GLASGOW
- BNL
- IN2P3-CC
- FZK
- SLAC
- GRIF-SACLAY
- CSCS
- INFN-MILANO

Example: data accesses per site on 11 Dec 09

Plans

- Restart data-taking with separate detector runs during January
  - No data export
- Start global cosmics run first week of February
  - Start of Tier-0 data processing and export
- Ready for LHC beams mid-February