ATLAS SC4 Service phase

Report on Tier0 exercise

ATLAS Tier0 + Distributed Data Management teams
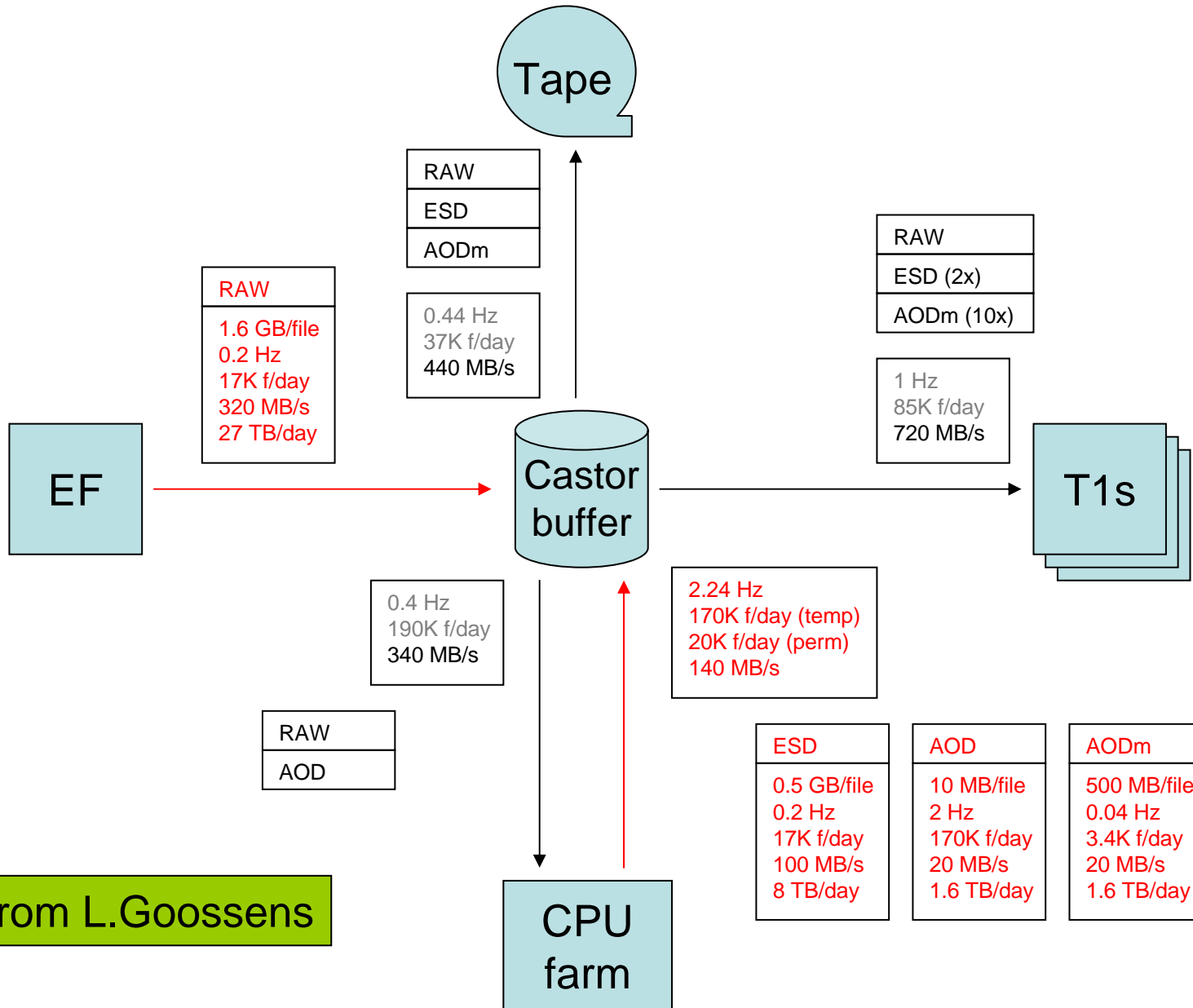
Outline

- Service phase
  - Tier0 exercise
  - DQ2
  - Status

Service Phase

- **Main exercise**: Tier0
  - Coordinators
    - Internal Tier0: Luc Goossens; Export to T1s: DDM group
  - Repeat of last year tests
  - Involves Tier0, all Tier1s and a subset of volunteer Tier2s
  - Only goal:
    - Run complete flow @ nominal rate for all T1s
  - Very similar to SC4 throughput phase
    - But going down to T2s, proper cataloguing.. this will be new!
  - Phase 1 (19th Juny for 3 weeks):
    - Repetition of last year tests, with DQ2 0.2.x
  - Phase 2 (> September):
    - **Consolidation**
    - Consider 'failing T1' scenarios, cleanup of pools @ Tier0, ..

Tier0

Tape

| RAW |
| ESD |
| AODm |

| 0.44 Hz |
| 37K f/day |
| 440 MB/s |

| RAW |
| ESD (2x) |
| AODm (10x) |

| 1 Hz |
| 85K f/day |
| 720 MB/s |

RAW
1.6 GB/file
0.2 Hz
17K f/day
320 MB/s
27 TB/day

EF → Castor buffer → T1s

| 0.4 Hz |
| 190K f/day |
| 340 MB/s |

| 2.24 Hz |
| 170K f/day (temp) |
| 20K f/day (perm) |
| 140 MB/s |

| RAW |
| AOD |

CPU farm

ESD	AOD	AODm
0.5 GB/file	10 MB/file	500 MB/file
0.2 Hz	2 Hz	0.04 Hz
17K f/day	170K f/day	3.4K f/day
100 MB/s	20 MB/s	20 MB/s
8 TB/day	1.6 TB/day	1.6 TB/day

Slide from L.Goossens

4

Service Phase

- Remaining exercises:
  - Distributed Production
  - Distributed Analysis
  - Reprocessing
- More details to be sent out soon!!…
  - (ATLAS CSC/SC4 coordination to sent out details for remaining exercises)
  - **But** important milestone is success of **Tier0 exercise**:
    - **Validation of DQ2 and WLCG data management infrastructure,** which is now being done as part of Tier0 exercise;
   - Remaining exercises depend on this outcome..
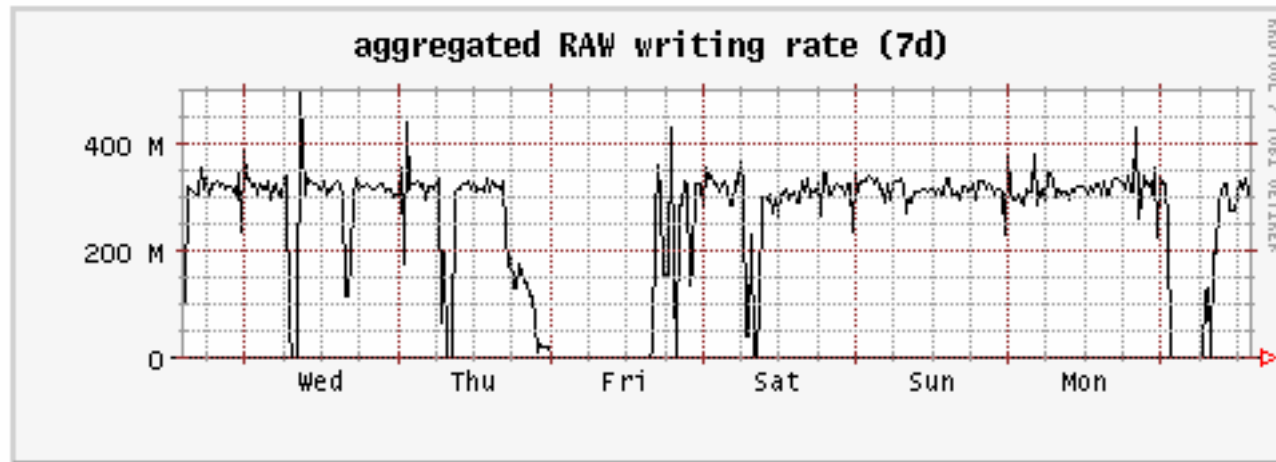
T0 internal

- web log at
  https://uimon.cern.ch/twiki/bin/view/Atlas/AtlasTierZero
  - all our problems with LSF and Castor

- monitoring at
  http://atlas.web.cern.ch/Atlas/GROUPS/SOFTWARE/DC/Tier0/monitoring/

T0 internal

- already reached nominal rates last Jan for periods of 12h (see CHEP paper)

- additional goals

  - run for 3 weeks continuous at nominal rate

    - see slide, we are not so far off

  - add conditions DB access and tag DB uploading

    - done

  - put in place extensive monitoring

    - we are monitoring ~45 variables at 5 min interval

      - IO rates, running jobs, pending jobs, backlog, tag upload times, LSF operation times, DQ2 times, LFC times, job error rates, C2 pool usage, migration queue, ...

    - we monitor every rfcp we do

      - > 1e6 so far

      - start time, end time, server used, transfer time, file size

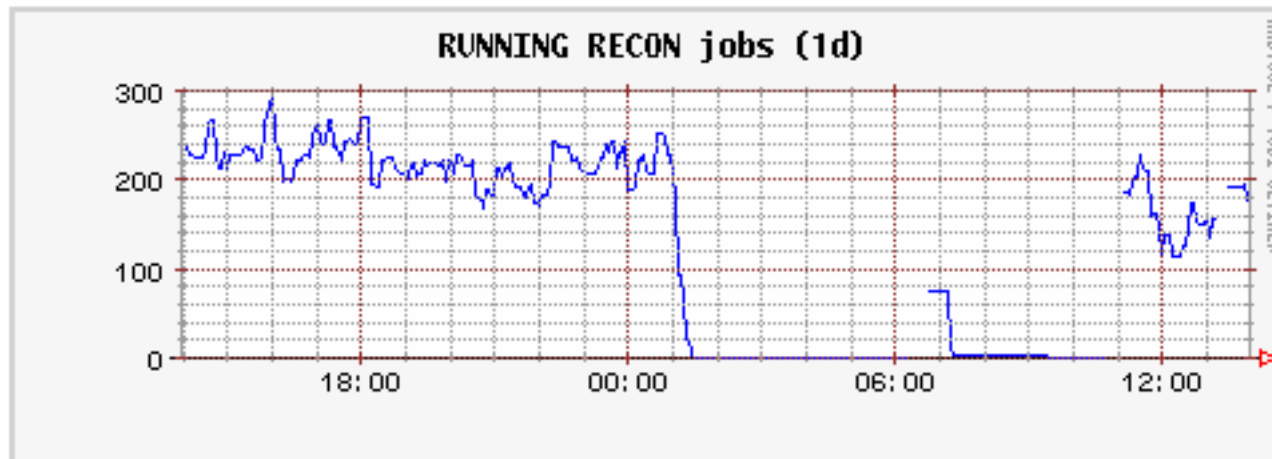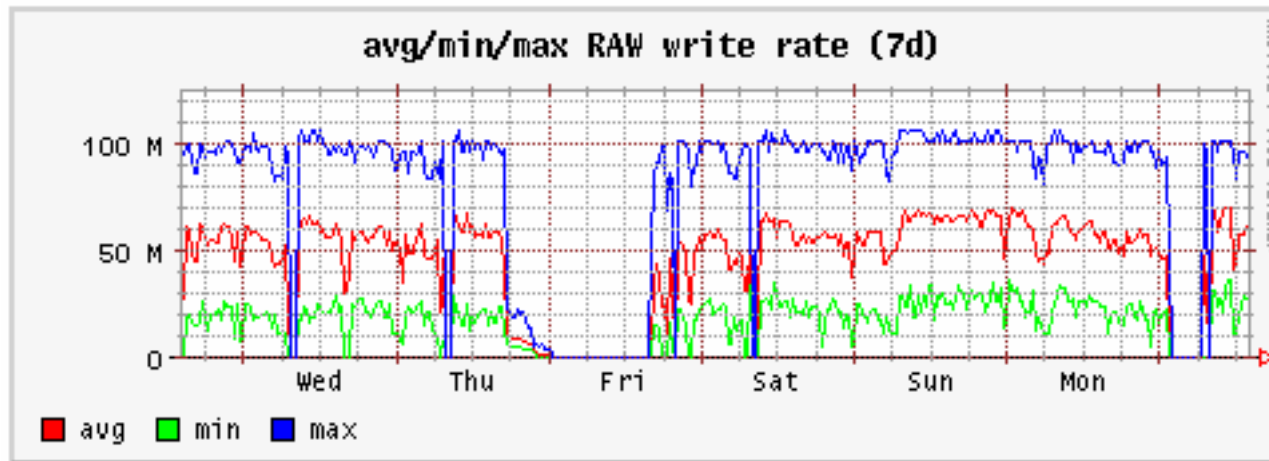   - we probably have better Castor monitoring now than Castor team

7

T0 internal

- run for 3 weeks continuous at nominal rate

8
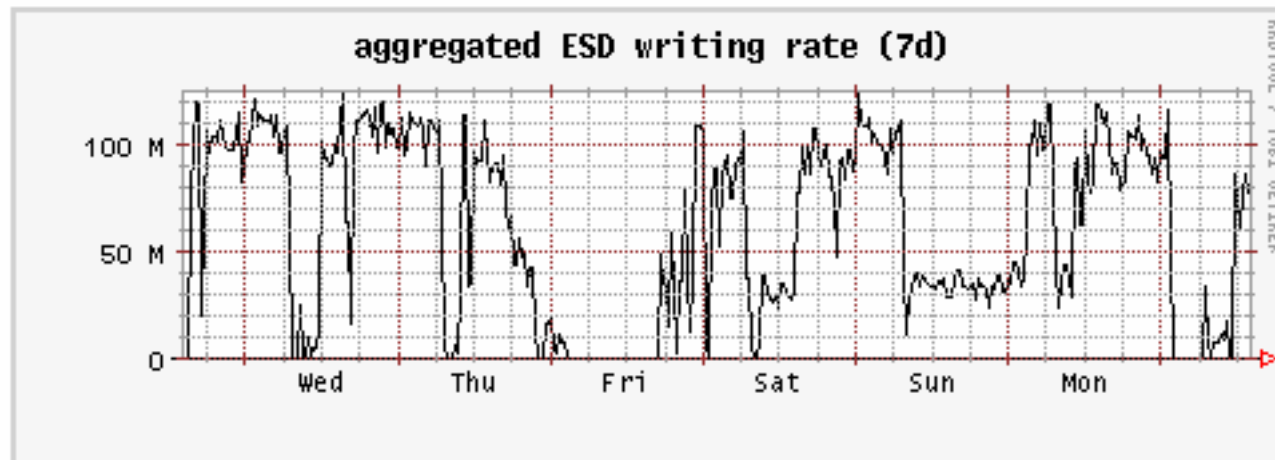
T0 internal

- extensive monitoring

DQ2

- DQ2, our Distributed Data Management system which builds on top of Grid data transfer tools, is based on:
  - A hierarchical definition of datasets
  - Central dataset catalogues
  - Data blocks as units of file storage and replication
  - Distributed file catalogues
  - Automatic data transfer mechanisms using distributed services (dataset subscription system)
- DQ2 allows the implementation of the basic ATLAS Computing Model concepts, as described in the Computing TDR (June 2005):
  - Distribution of raw and reconstructed data from CERN to the Tier-1s
  - Distribution of AODs (Analysis Object Data) to Tier-2 centres for analysis
  - Storage of simulated data (produced by Tier-2s) at Tier-1 centres for further distribution and/or processing

Status

- Status:
  - Web log:
    - https://uimon.cern.ch/twiki/bin/view/Atlas/DDMSc4
- Goal:
  - Run complete flow (T0 + T1s export) at nominal rate
  - Aim is not to provide a constant stream of data to Tier1s (dteam does it) but rather to provide a realistic stream of data from the Tier0 to Tier1s
    - Even if there are failures, ramp up/downs, etc
  - Much more than throughput rate:
    - LFCs, FTS 'job logic', bookkeeping, monitoring, …

Status

- Started June 19th with all Tier1s
  - (Tier1s are active and running all the time except for scheduled downtimes)
- Took a few days to get all DQ2 installations validated
- Then first ramp up was quite successful, rapidly overtaking last year's "record"
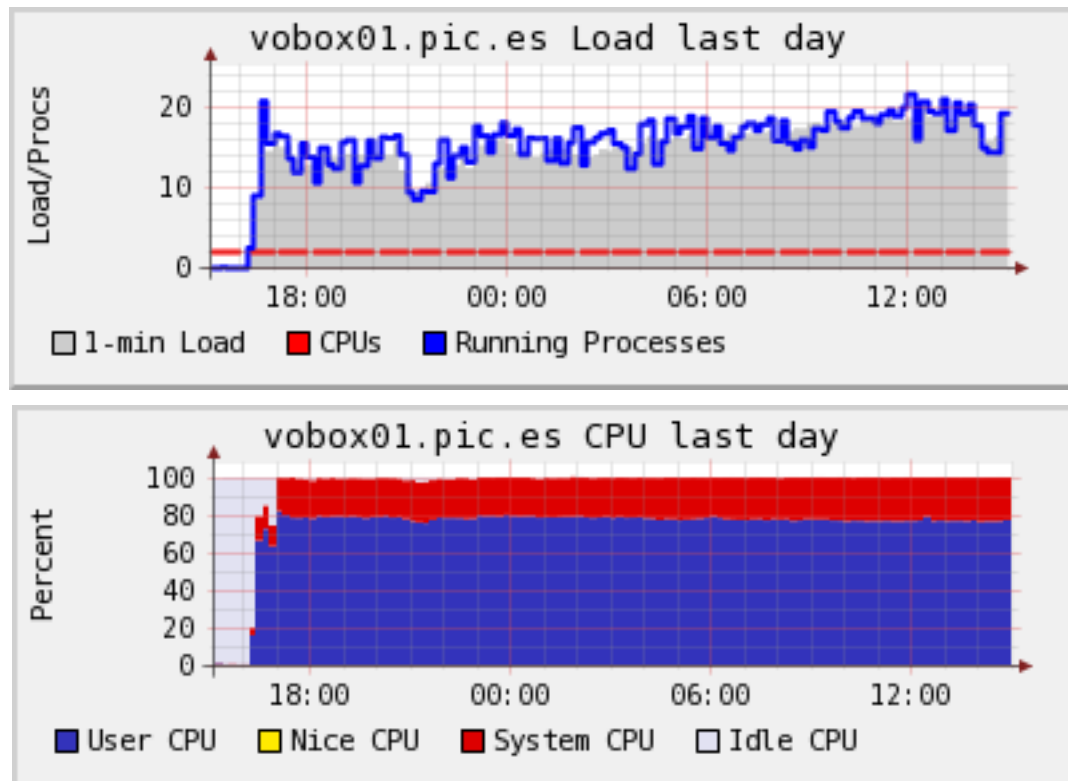
20th June

First problems

- DQ2 heavier on VO BOX h/w than initially thought
- Stable sites:
  - BNL, LYON, PIC, SARA, CNAF, TRIUMF
- Unstable sites:
  - RAL, ASGC, FZK
- Overloading LFCs with too many parallel requests
- DQ2 monitoring did not scale

DQ2 load average on VO BOXes

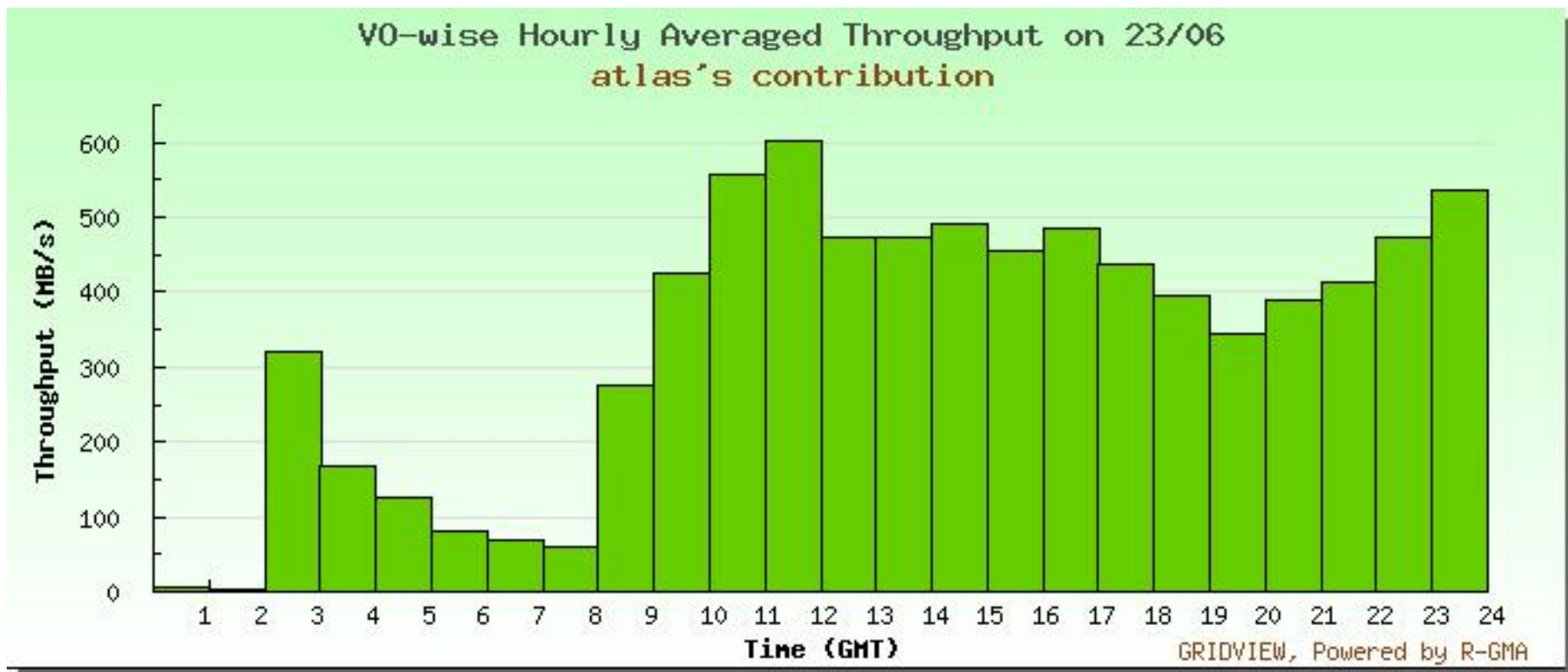- (thanks to Xavier Espinal for plots from PIC's VO BOX)

22th June

- Power cut at CERN:
  - took a long time to recover
    - not realistic 'exercise' hopefully!

23th June

- Aim: fast ramp up as Tier0 had been running overnight
- **Successfully done!!**

VO-wise Hourly Averaged Throughput on 23/06
atlas's contribution

Missing

- Not nominal rate at this phase:
  - ASGC not working
  - FZK with slow VO BOX (dual processor Pentium III!)
  - No AODs! (200MB/s aggregate)
  - RAL: occasional problems: reason not understood from our side
  - Assigning ~ MoU rates (but without AOD)
  - + occasional file transfer problems (GGUS tickets submitted)

26th, 27th June

- Updates to DQ2
  - To activate monitoring
  - Solve problem with overloading FTS with too many requests
    - Requests were staying pending (particularly in sites with problems) and DQ2 kept sending requests and asynchronously canceling old requests: not terribly effective
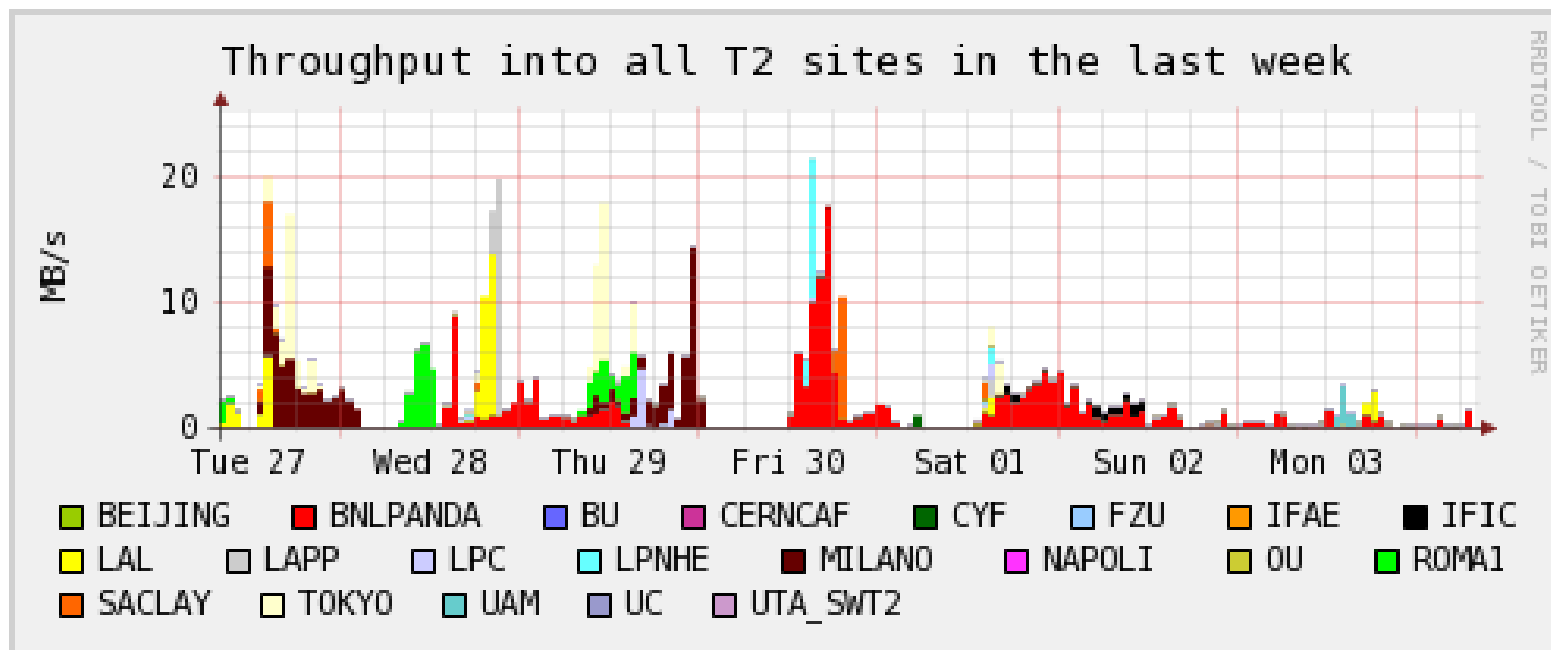
19

27th, 28th June

- Last day without AOD
- Monitoring enabled: useful for understanding site errors and submitting GGUS tickets

VO-wise Hourly Averaged Throughput on 28/06
atlas's contribution
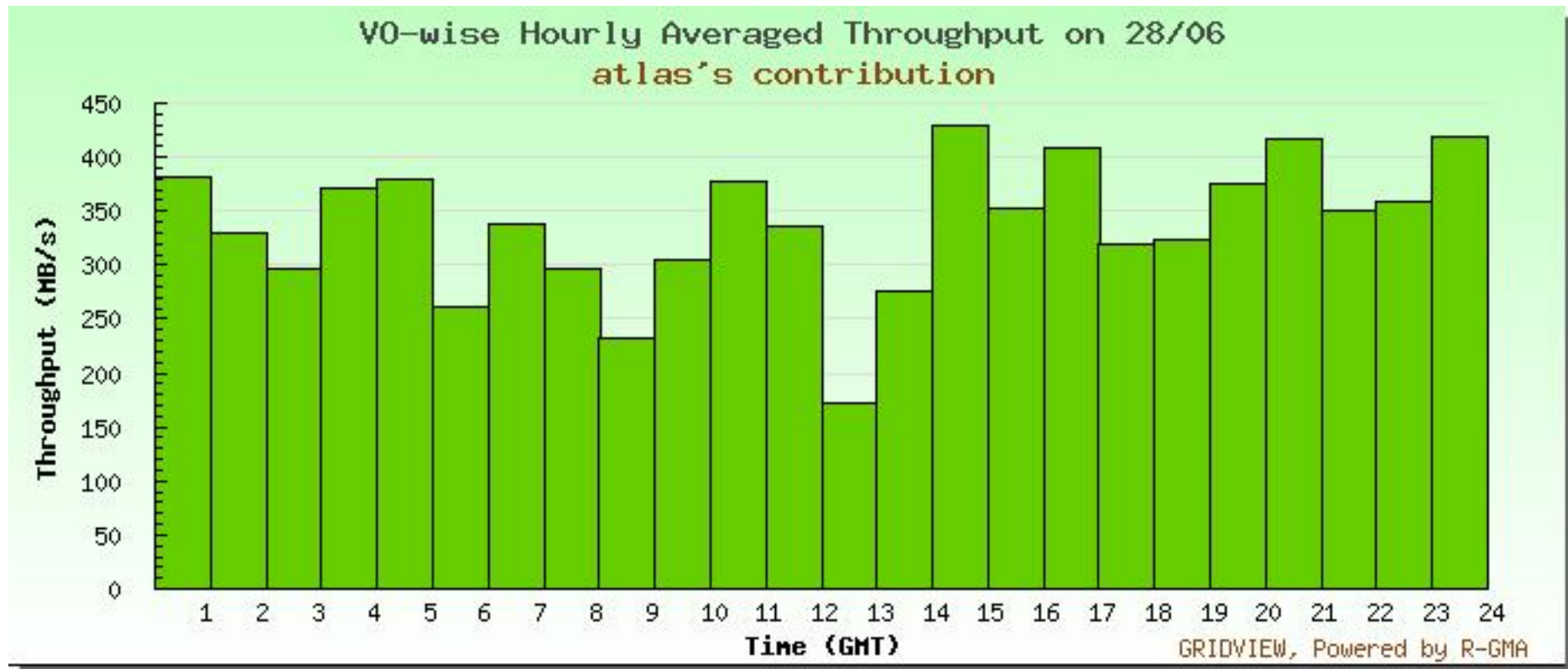
30th June

- Tier0 meltdown: due to LFC problems
- Report from Tier2 tests by ATLAS regional contacts

Tier2 tests

- Manually subscribing to AOD datasets using DQ2
- Triggered by ATLAS regional contacts
  - Not a constant exercise
    - Only done when regional contacts are available
    - LYON and CNAF: **successfully supporting their Tier2s** (have sustained 20MB/s for short periods: typically "few hours AOD production")
- ATLAS model:
  - VO BOXes at Tier1 also serve Tier2s
  - VO BOXes are **even more overloaded**!

Now…

- **Another rollout of DQ2 fixes**
  - Monitoring and logging improvements
  - Automatic recovery in case of crashed VO BOXes
    - Good indication the exercise is becoming **more mature** with '*unlikely*' events starting to occur

- **All Tier1s are active all the time**
  - DQ2 'accumulates' failures or pending transfers and cleans old data ~ once per day
    - because Tier0 disk pool cleans files older than 30hours

Not yet nominal rate?

- Tonight: problems with DQ2<->Tier0 plug-in
- Site problems:
  - ASGC (storage problems: unclear)
  - FZK (limited by insufficient VO BOX h/w)
  - CNAF (slow writing to tape; filling up disk)
- DQ2 problems:
  - Automatic recovery for 'harsh' failure conditions
- ***Would like to continue running for a few more weeks***
  - *Gain experience with > hours running (as we did this last weekend)*

24

"General" Problems

- Familiarity with DQ2 from ATLAS contacts
- Communication with:
  - ATLAS-site contacts
  - Sites' technical contacts
  - Service Challenges team
- GGUS
  - Often ineffective forwarding of tickets
    - Problems often require a coordinated 'team' to understand the origin, not a single group looking at 'their' own logs
  - (personal opinion: I'd definitely prefer chat or IRC at this stage)

Summary

- Have sustained ~ 500 MB/s for several hours on a few occasions:
  - Means all components are working together (from Tier0 to Tier1 export)
  - Close to achieving goal of nominal rate
  - Speculation:
    - could aim at >> throughput rate by making the exercise less realistic but not interesting for us: much rather have a realistic (even if more problematic) flow
- Considerably better than SC3!!
  - But still quite a bit to go to become stable…
- Limitation understood:
  - Running integrated data flow (our choice) makes ramping up 'slower'
  - Not having all Tier1s available