ATLAS
Data Management Tests

Dario Barberis

CERN & Genoa University

ATLAS Data Management Activities

- The Distributed Data Management Operations team supports in 2007:
  - Data taking with cosmic rays and data distribution to Tier-1/2s
  - Functional tests of data distribution
  - High-throughput tests of data distribution Tier-0→Tier-1s→Tier-2s
  - Continuous simulation production (close to 6M events/week currently)
- During the last few months of 2007 and the first half of 2008, the FDR (Full Dress Rehearsal) will bring together and integrate the (so far) independent and complementary tests:
  - Data transfer from the online to Tier-0
  - Internal Tier-0 data processing and data movement
  - Data quality monitoring, calibration and reconstruction jobs running in real time on Tier-0
  - Data distribution to Tier-1s and Tier-2s
  - Data access by analysis jobs at Tier-2s
  - Reprocessing at Tier-1s
  - Continuous simulation production at Tier-2s
- Finally, the CCRC tests will run the ATLAS FDR at the same time as the other experiments' data transfer and processing tests

Earlier Cosmic Data Taking Tests

- **M3 Test back in June 2007**
  - Was mostly detector test
  - Last few days also wrote data into Castor 1
  - Only 1 SFO (Sub-Farm Output processor) could be used→ rate limitation
  - Data used for detector studies

- **TDAQ Technical Run in August 2007**
  - To test writing into Castor 2
  - Used all 5 SFO's → ~50 MB/s achieved
  - Expected 320 MB/s (64 MB/s per SFO)
  - Problem with writing/reading in SFO's
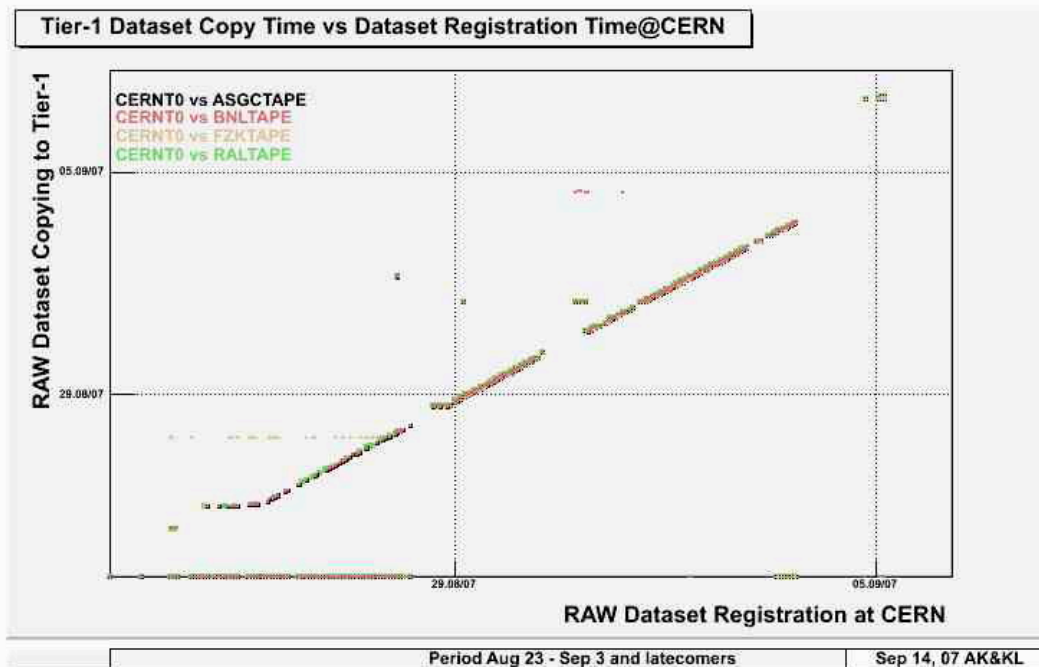
M4 Objectives

- August 23 – September 3

- Using 4 SFOs: rate < ~250 MB/s

- Data written into Castor 2

- Full Tier-0 operation

- ESD (Event Summary Data, output of reconstruction) data subscribed from Tier-1s to Tier-2s
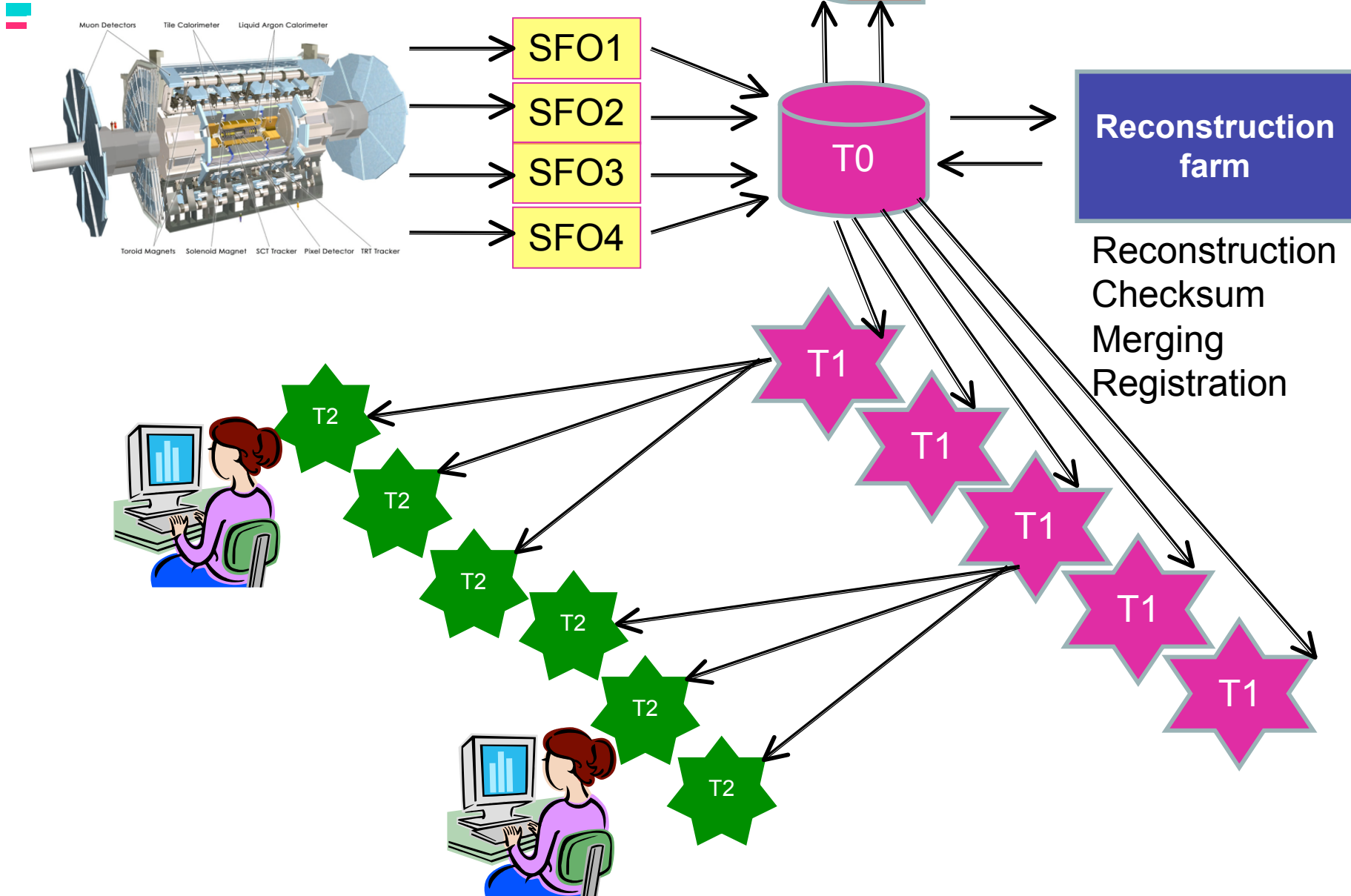
- Analyse M4 data at Tier-2s

M4 Objectives

✓ August 23 – September 3

✓ Using 4 SFOs: rate < ~250 MB/s

✓ Data written into Castor 2 (~40 TB)

✓ Full Tier-0 operation

✓ ESD (Event Summary Data, output of reconstruction) data subscribed from Tier-1s to Tier-2s
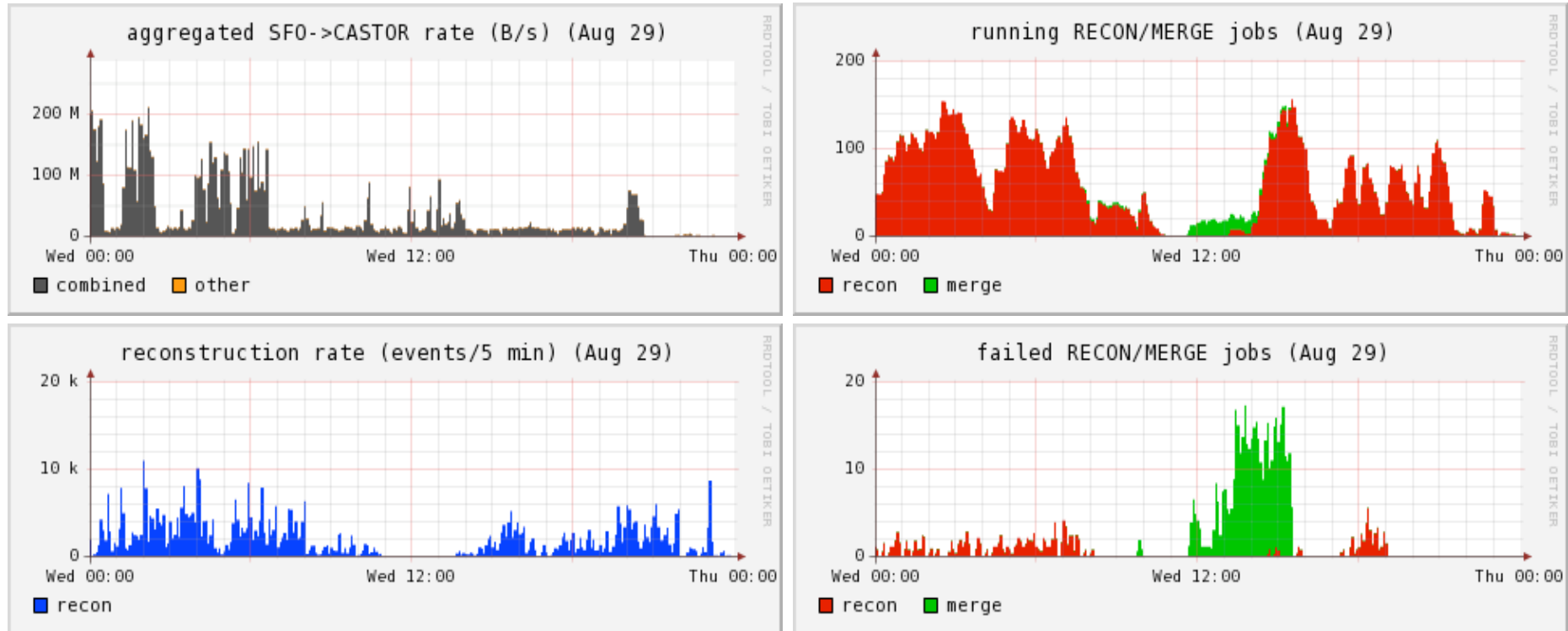
✓ Analyse M4 data at Tier-2s

M4 Cosmic Run (Aug 23 -Sep 3 2007)

Tier-1 Dataset Copy Time vs Dataset Registration Time@CERN

CERNT0 vs ASGCTAPE
CERNT0 vs BNLTAPE
CERNT0 vs FZKTAPE
CERNT0 vs RALTAPE

RAW Dataset Copying to Tier-1

RAW Dataset Registration at CERN

Period Aug 23 - Sep 3 and latecomers | Sep 14, 07 AK&KL

Dario Barberis: AT

Tape

SFO1

SFO2

SFO3

SFO4

T0

Reconstruction farm

Reconstruction
Checksum
Merging
Registration

T1

T1

T1

T1

T1

T2

T2

T2

T2

T2

T2

Tier-0 Monitoring
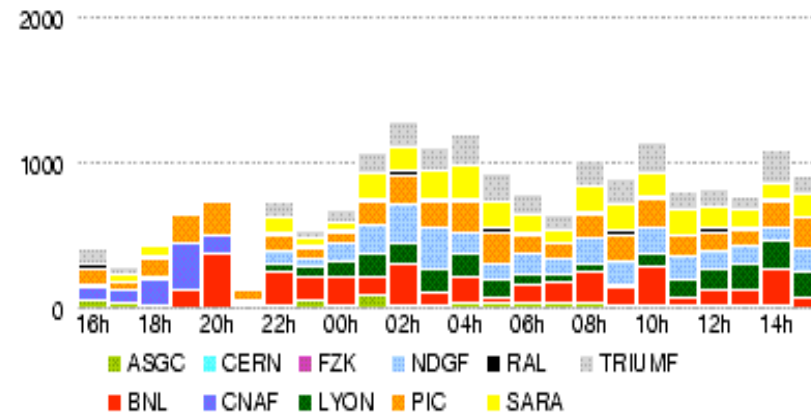

Snapshot from August 29

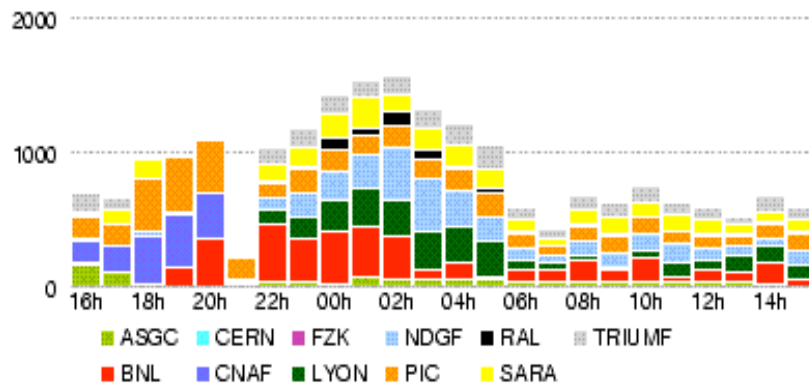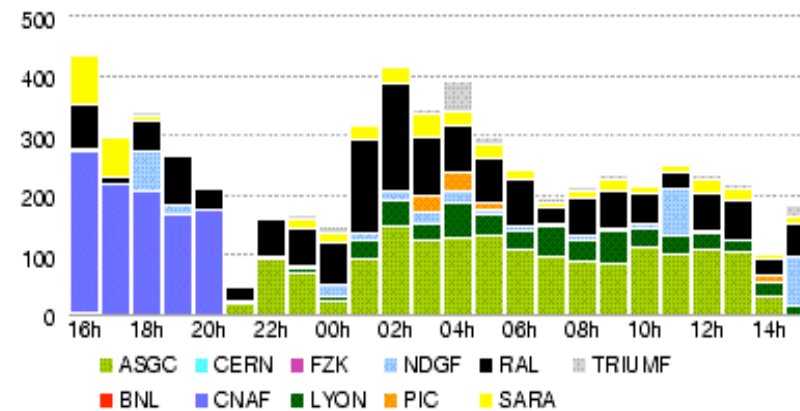# Tier-0→Tier-1 August 31st

Throughput MB/s
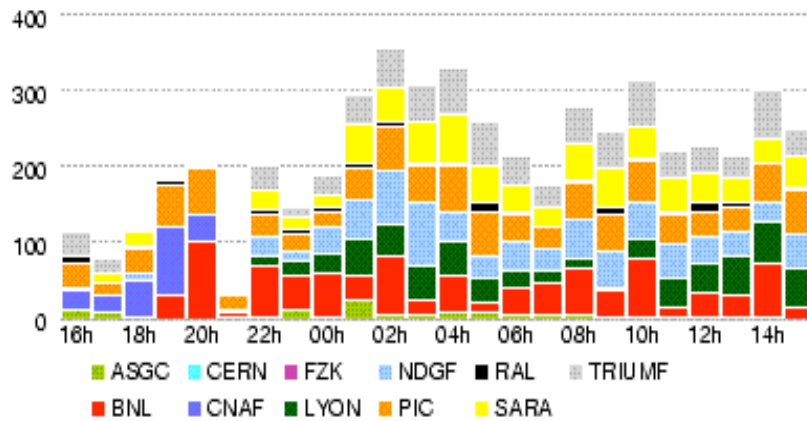
Data transfered GB/h

Completed file transfers

Total number of errors

Dario Barberis: ATLAS Data Management Tests
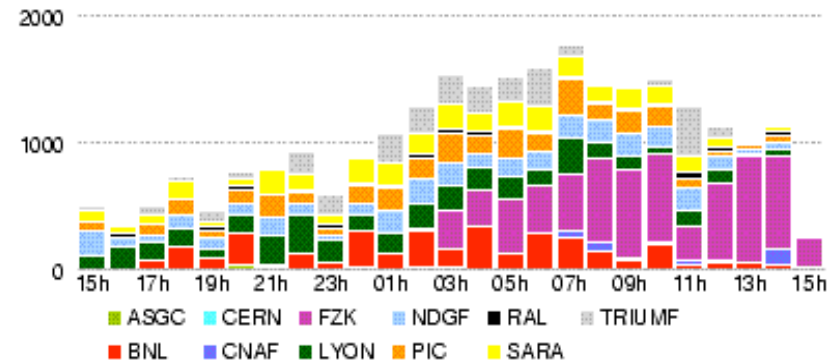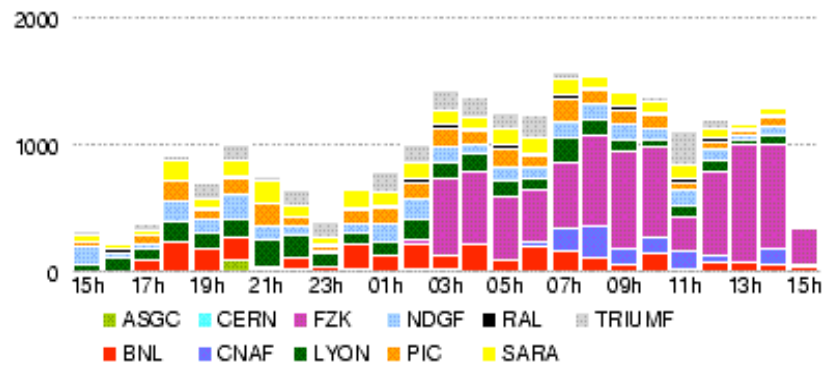
Recovery works
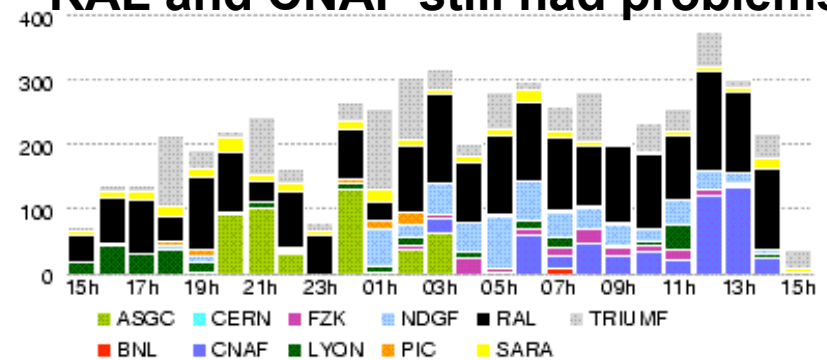
Throughput MB/s
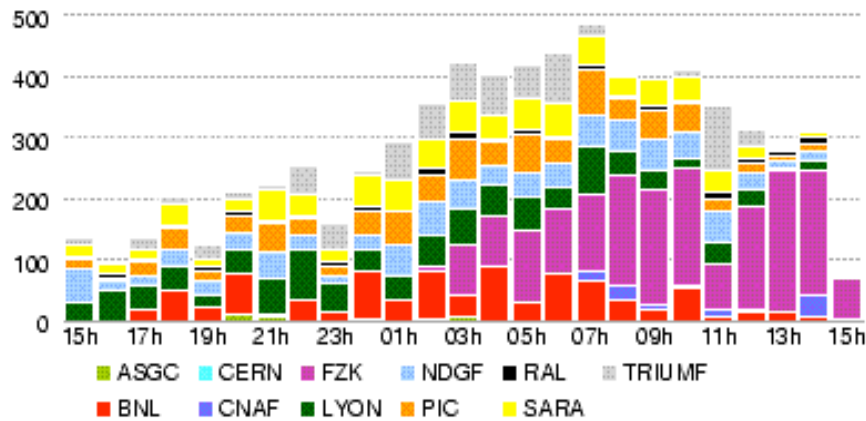
Data transfered GB/h

FZK recovered at 02:00 on Sept. 2nd

Completed file transfers

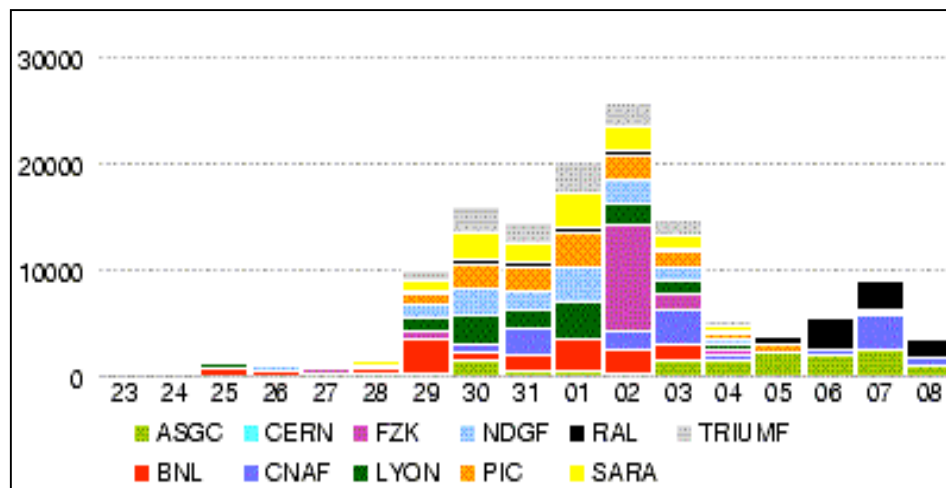## Total number of errors

RAL and CNAF still had problems

Dario Barberis: ATLAS Data Management Tests
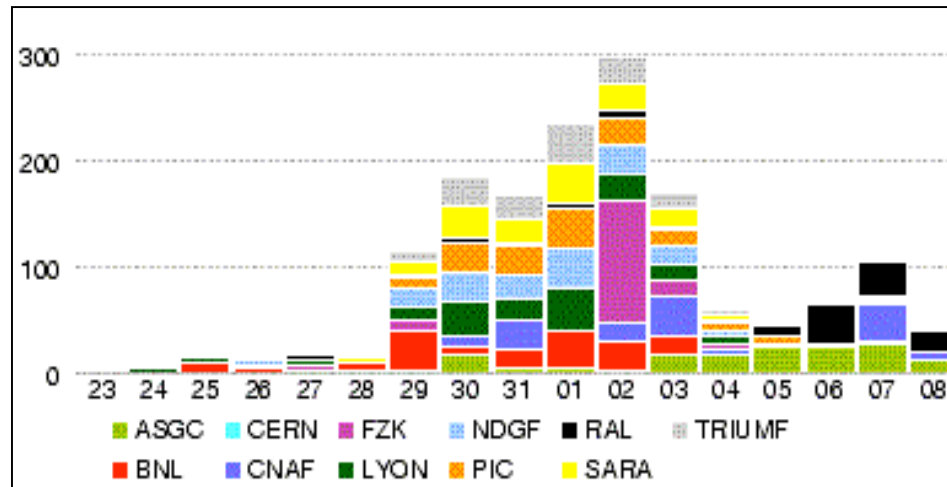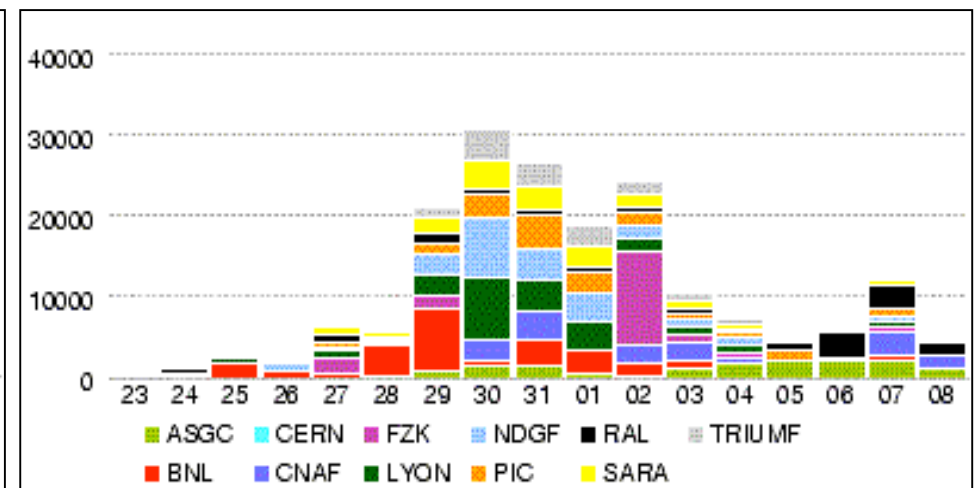
The whole M4 run

Total throughput (MB/s)
Aug 23 – Sep 8

Data transferred (GB/day)
Aug 23 – Sep 8

Completed file transfers
Aug 23 – Sep 8

Tier-0 → Tier-1 Export

- Tier-0/DQ2 upload started Aug 24 (evening), worked smoothly
  - Low latencies (usually <1h) between arrival of data in CASTOR and registration with DQ2

- Tier-0 → Tier-1 data export could be established to all 10 Tier-1s
  - Worked well from the beginning for 6/10 Tier-1s
  - Problems with remaining 4 were resolved in the course (CNAF, FZK) or towards the end (RAL, ASGC) of the M4 exercise
  - Further Tier-1 → Tier-2 export succeeded for some Tier-2s

- O(70-80%) of the data subscriptions were processed successfully
  - Problems with remaining 20-30% still have to be understood and are being investigated (e.g. cancelled subscriptions)

- Real-time data analysis could be done simultaneously at some European and US Tier-1/2 sites

- Summary and more detailed overview statistics will be produced

Plans & Outlook for M5+

- **Checksums strategy**
  - Have to be provided by the online BEFORE transfering data to Castor

- **Improved communication with TDAQ/online**
  - Via online database, not only CASTOR name server look-ups as in M3/M4
  - Would need info about
    - file metadata (names, sizes, events/file, checksums if possible)
    - number of files/run, active SFOs/run
    - end-of-run

- **Offline Data Quality Monitoring working**
  - Low latencies (O(1-2h)) in Tier-0 processing as in M4 in principle would allow quasi-real-time DQ checking
  - Possibly useful/attractive even for the online and detector shifters

- **Tier-0 → Tier-1 export has to be further improved and streamlined**
  - In particular RAW data export has to become 100% reliable
  - Role of RAW vs. ESD vs. AOD not clear

Functional Tests of Data Distribution

- The Functional Tests (FT) are meant to check the data transfers between Tier-0/1/2 centres, simulating at low rate the complete data flow according to the computing model:

  - Tier-0→Tier-1s: RAW and 1st-pass ESD and AOD data
  - Tier-1→Tier-2s: RAW and ESD samples, and AOD data
  - Tier-1→Tier-1s: reprocessed ESD and AOD data

- Data are organised in a small number of datasets, each one consisting of ~30 files, of representative sizes, and are subscribed to sites using DDM tools

- We measure time and efficiency of data transfer, number of retries, etc.

13

Functional Test Examples

10 Tier-1s and 46 Tier-2 sites (1st week Oct'07):

- **ASGC**
  - ■ AU-ATLAS, TW-FTT,AU-UNIMELB
- **BNL**
  - ■ AGLT2, BU, MWT2, OU, SLAC, UTA, WISC
- **CNAF**
  - ■ LNF, MILANO, NAPOLI, ROMA1
- **FZK**
  - ■ CSCS,CYF,DESY-HH,DESY-ZN,FZU,LRZ, FREIBURG, WUP
- **LYON**
  - ■ BEIJING, CPPM, LAL, LAPP, LPC, LPNHE, NIPNE_02,NIPNE_07, SACLAY ,TOKYO
- **NDGF**
- **PIC**
  - ■ IFAE, IFIC, UAM,LIP
- **RAL**
  - ■ GLASGOW, LANCS, MANC, QMUL, DUR, EDINBURGH, OXF,CAM,LIV,BRUN,RHUL
- **SARA**
  - ■ IHEP, ITEP, JINR, PNPI, SINP
- **TRIUMF**
  - ■ ALBERTA, MONTREAL, SFU, TORONTO,UVIC

Test completed, not 100%, failed, not part.

Dario Barberis: ATLAS Data

Data Transfer from Tier-1s to Tier-1s

	ASGC	BNL	CNAF	FZK	LYON	NGDF	PIC	RAL	SARA	TRIUMF
ASGC										
BNL										
CNAF										
FZK										
LYON										
NDGF										
PIC										
RAL										
SARA										
TRIUMF										

Data Transfer from CERN to Tier-1s

	ASGC	BNL	CNAF	FZK	LYON	NDGF	PIC	RAL	SARA	TRIUMF
CERN										

100% , 90+% , >50%, less than 50%, of data transferred within 24h

DDM Functional Test Oct 2007. Dataset transfer time

Dataset transfer time (hours)

14
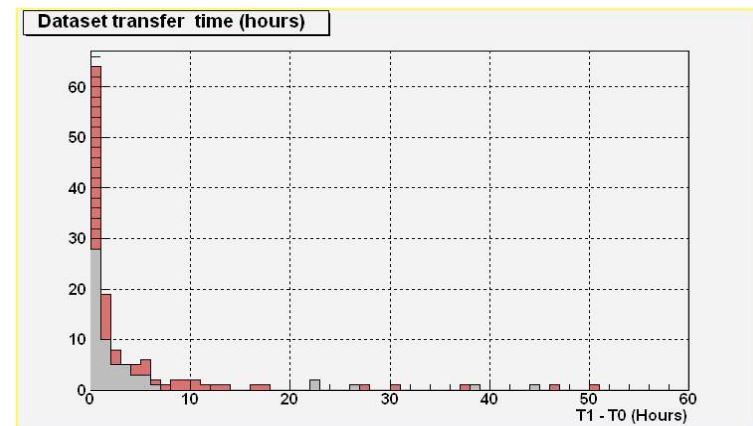
CSC, FDR and Data Taking (1)

- CSC (Computing System Commissioning) tests started in early 2006

    - Set of independent component tests, followed (in 2007) by integration tests

    - All test suites are kept active and used for each major software release

    - Tier-0 internal tests and Data Distribution tests exercised periodically with additional software components

    - The final integration test (Full Dress Rehearsal) has been in the meantime split into a number of phases of increasing complexity

- Cosmic Ray data taking runs:

    - M4 run: 2 weeks, August-September 2007

    - M5 run: 2 weeks, late October 2007

    - M6 run: 2 weeks, early February 2008

    - After that: continuous operation with detector commissioning activities

CSC, FDR and Data Taking (2)

- FDR (Full Dress Rehearsal) tests started in Spring 2007 and will run through late Spring 2008:

  - Sept/Oct 07: Data preparation for round 1 (round 0 was the CSC streaming test). Data from release 12 simulation, data preparation using release 13 (currently in validation phase).

  - End of October 2007 onwards: Tier-0 operations test (shifts) and reconstruction runs using release 13 (including support for the M5 run).

  - Nov 07 - Feb 08: Reprocess at Tier-1s, make group DPDs (Derived Physics Data).

  - Dec 07 - Jan 08: Data production for final round using release 13 (assumes simulation has been validated).

  - Feb 08: Data preparation for final round using release 13.

  - Mar 08: Reconstruction final round using release 14. Tier-0/Tier-1 shifts.

  - Apr 08: DPD production at Tier-1s, user testing using release 14. Ideally same release as for first data.

  - April/May 08: More simulated data production in preparation for first data.

Timescales of Data Management Tests

- Early October 2007 (and repeat about once/month): Functional Test
  - Add/check Tier-2 sites, also improve Tier-1↔Tier-1 transfers
- Mid-October (and repeat as needed): data distribution throughput tests
- End Oct. — early Nov.: M5 cosmic run
- Mid-Nov. — end Dec.: FDR Phase 1
- January-February 2008: SRM 2.2 installed at Tier-1s and under test
- Early February 2008: M6 cosmic run
- Late February (compatibly with the M6 run): 1st phase of CCRC'08
  - Common data transfer and processing tests with the other LHC experiments
- March-April: FDR phase 2, cosmic runs, more throughput tests
- May 2008 (exact time depends on cosmic runs): 2nd phase of CCRC'08, including from the ATLAS side:
  - Data transfer from online (compatibly with online and detector constraints)
  - Full Tier-0 operation including calibration loop
  - Data export to Tier-1s (all) and Tier-2s (possibly all)
  - Retrieval from tape and reprocessing at Tier-1s of pre-placed data
  - Distribution of reprocessed data
  - Simulation production at Tier-2s and data distribution
  - Submission of "group analysis" jobs at Tier-1s and "user analysis" jobs at Tier-2s
- Target rate would be the ATLAS nominal rate (200 Hz), which we know is doable already now for most of the (independent) activities above
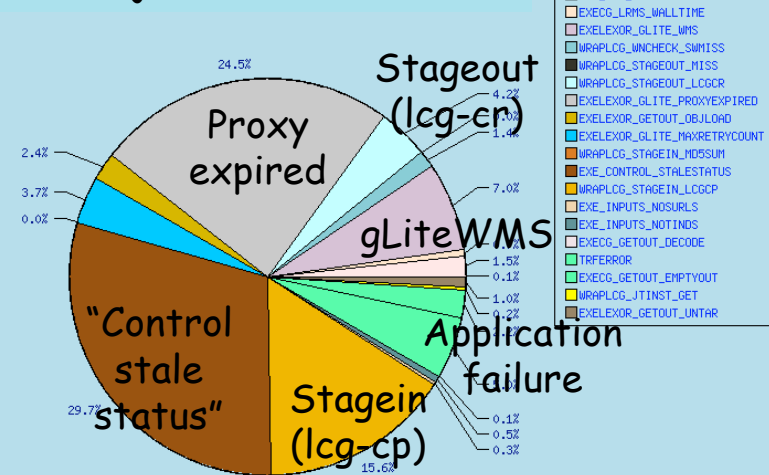
Outstanding Grid Issues

- Data Management: instability of SRMv1 servers

  - This is the major source of errors and wasted wall time for our production and analysis jobs

  - It could be also a source of poor CPU/WCT efficiency but we need to check

  - We were told a long time ago that SRMv2 will solve all these problems

- Resource management tools from the VO side are still missing and the end is not in sight

  - Job priorities, shares etc:

    - We can settle for a very coarse granularity (s/w installation, validation, production, analysis) and do the actual prioritisation ourselves (pilot jobs…)

  - ACLs and quotas for disk space management are needed soon

    - At least at the group/role level, eventually per user

    - This is not a new request but was postponed as SRM 2.2 has higher priority

Dario Barberis: ATLAS Data M

EGEE job errors 9/10/07

EGEE WCT errors 9/10/07