



# ALICE February/March exercise summary

Latchezar Betev  
WLCG Collaboration Workshop  
22 April 2008

# ALICE in February/March CCRC

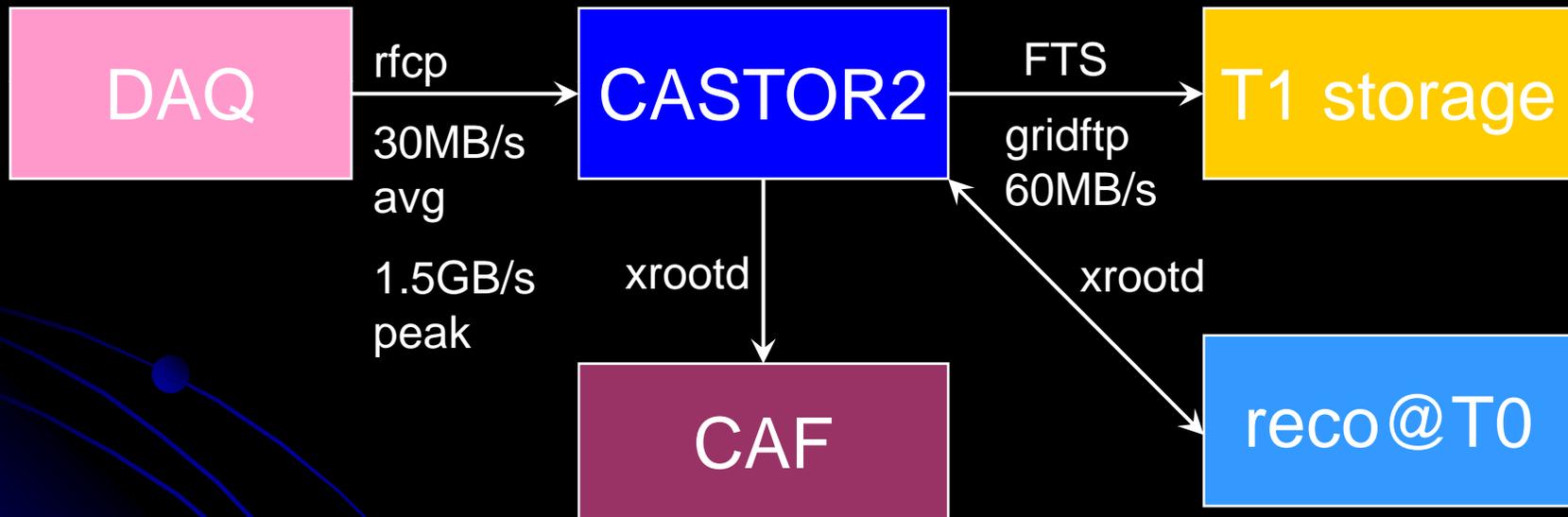
- February/March phase of CCRC'08 overlaps with ALICE second detector commissioning exercise (RUN II)
- Live detector data – cosmic trigger, calibration data
- Practically full complement of ALICE detectors installed and active
- Standard data taking regime - 'global runs' with all detectors participating

# Offline tasks

- Registration of data in CASTOR2 (T0) and on the GRID
- Replication T0->T1
- Conditions data gathering and publication on the GRID
- Quasi-online reconstruction
  - Pass 1 at T0
  - Pass 2 at T1s
  - Replication of ESDs to CAF/T2s
- Quality control
- MC production and user analysis at CAF/T2s

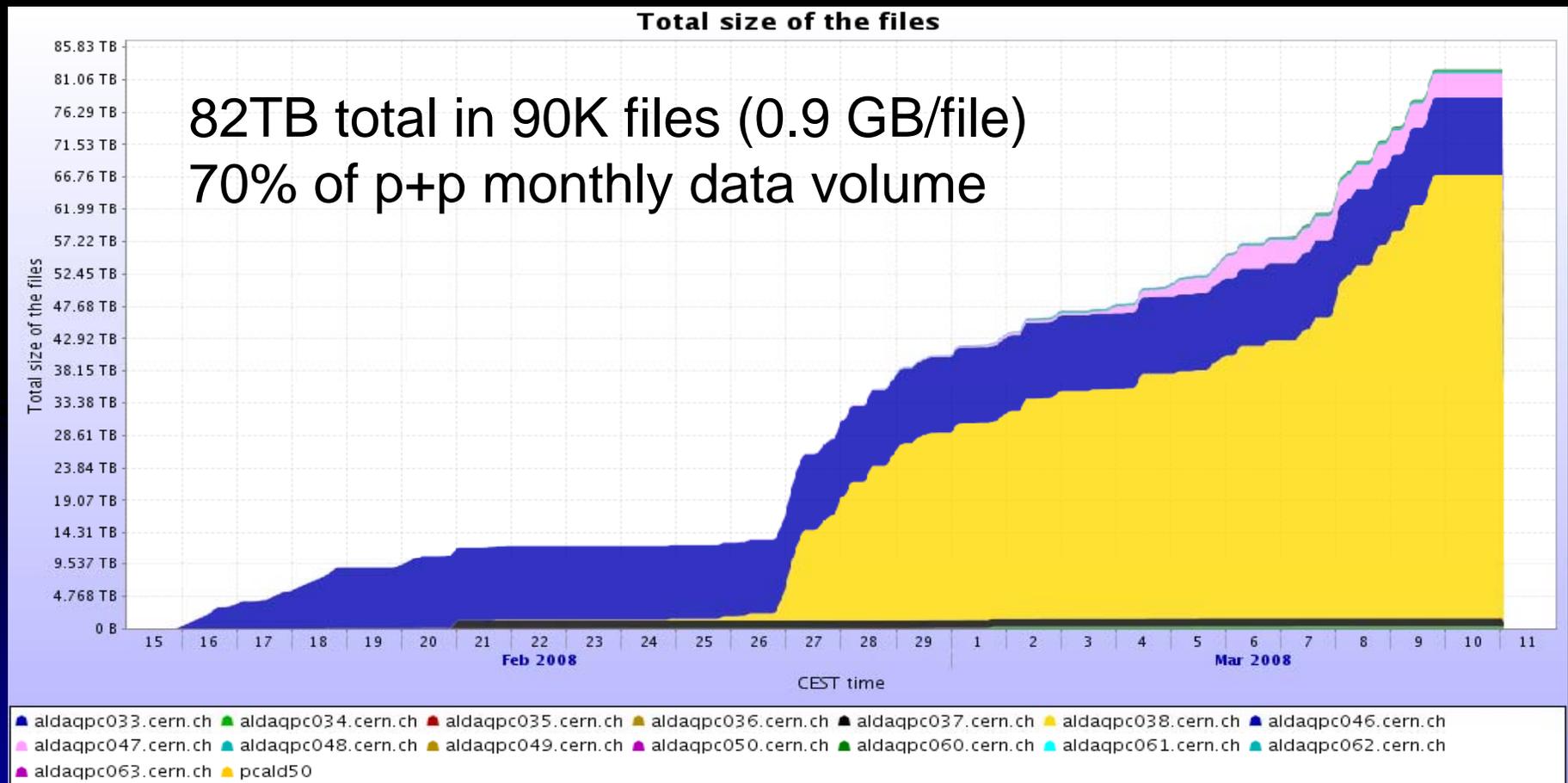
# Data path and rates

Rates correspond to  $\frac{1}{2}$  standard p+p ALICE data taking rate (from DAQ) + full p+p data taking rate (replication)

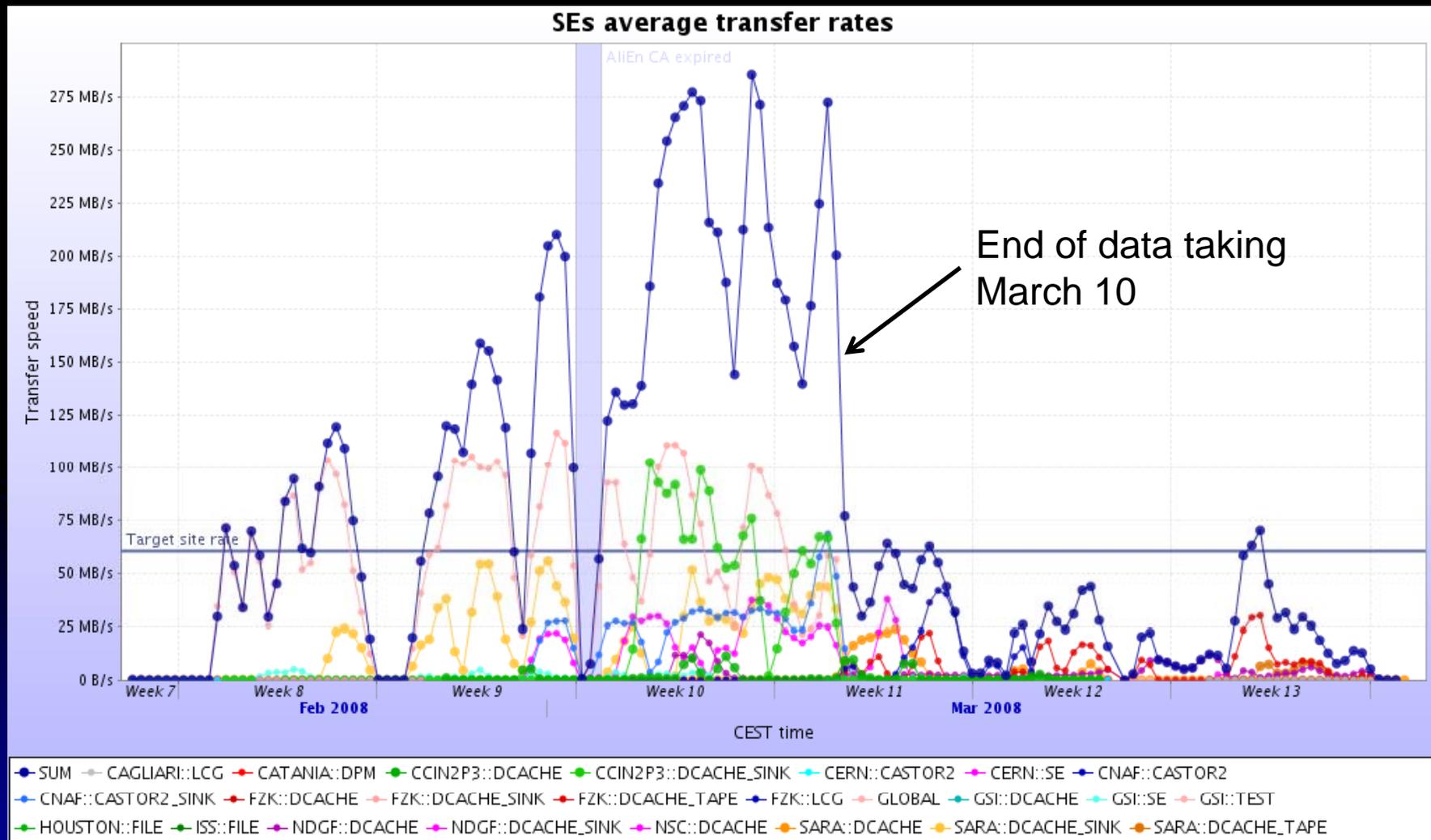


# Time line

- The ALICE exercise ran from 15 February to 10 March

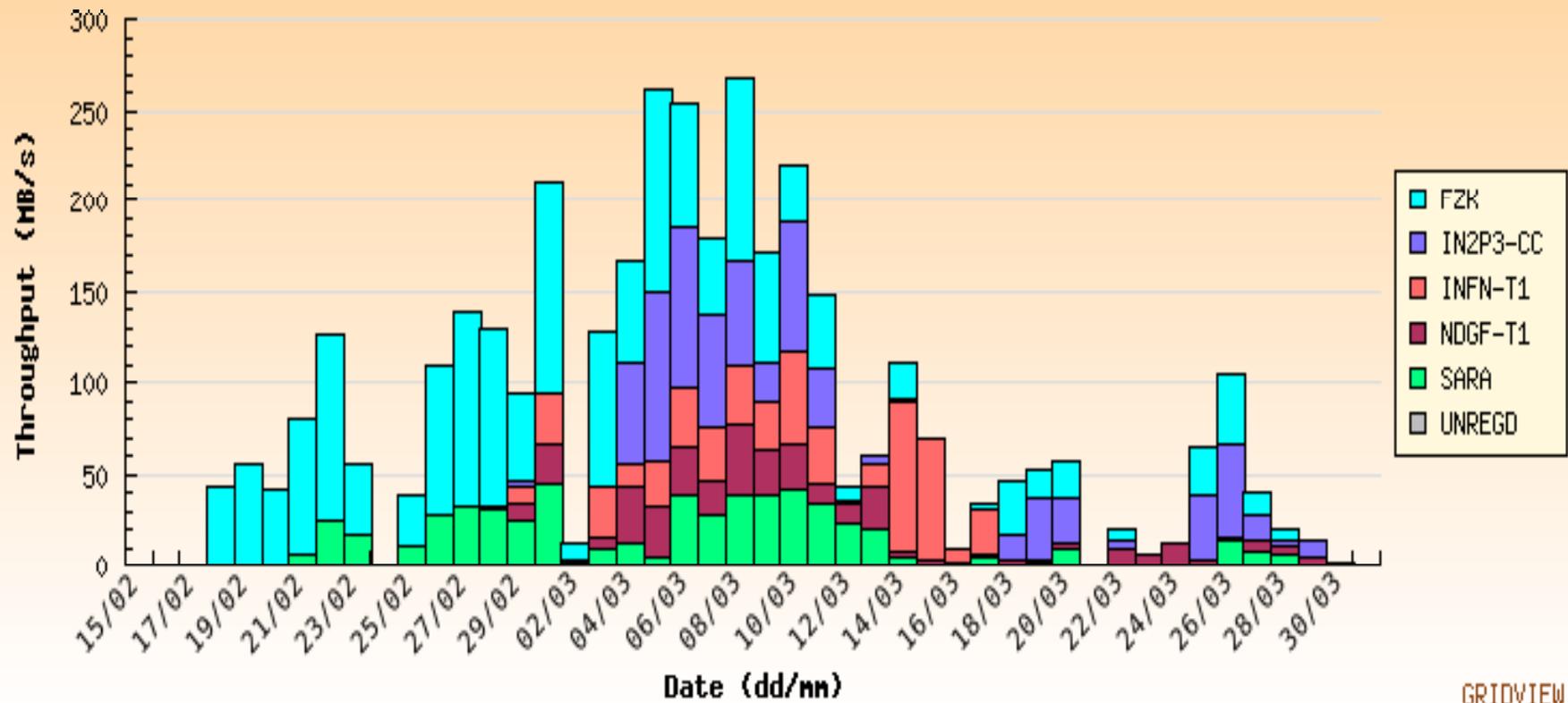


# Replication T0->T1



# Replication T0->T1

Averaged Throughput From 15/02/08 To 31/03/08  
Data Transfer For 'Alice' From All Sites To All Sites



# RAW replication status

- The target rate of 60MB/sec was achieved, average over 3 weeks was **125MB/sec**
- Period beyond 10 March – leftover files, specific datasets replication to T2s (calibration data)
- 90% of the data was replicated quasi-online
- Simultaneous replication together with ATLAS/ CMS and LHCb not a problem
- Short interruptions of transfers (up to one day) is manageable – the data is still ‘hot, on disk’ at the source
- (see later for exceptions)
- **Many thanks to the CERN IT/FIO group for the extensive help in the debugging/optimization of the ALICE FTS transfers**

# MSS operation

- Optimization of file sizes and access patterns
- Extensive study at CERN during the Feb/March exercise (applicable to MSS at T1s?)
  - Feedback and advice from CERN IT/FIO
- Issues
  - Average file sizes too low (inefficient tape writing)
  - Too many (repeat) mounts of same tapes (inefficient robot usage)

# MSS operation (2)

- Solutions (partial) applicable for May and beyond
  - RAW data chunks size 1GB->10GB
    - Subsequently ESDs 0.1 GB->1GB (without any further consolidation)
  - Pre-staging of datasets targeted for FTS replication
    - For 'late' transfers – files no longer in disk buffer
  - Optimization of MSS mount points
    - Separate for RAW data and ESDs
  - Use of file archives
    - All job output in a single file stored to MSS
  - Regular seminars with ALICE users on good Grid practices
    - In addition - moving entirely to disk-based storage (with replication) for all user files

# xrootd-enabled storage summary

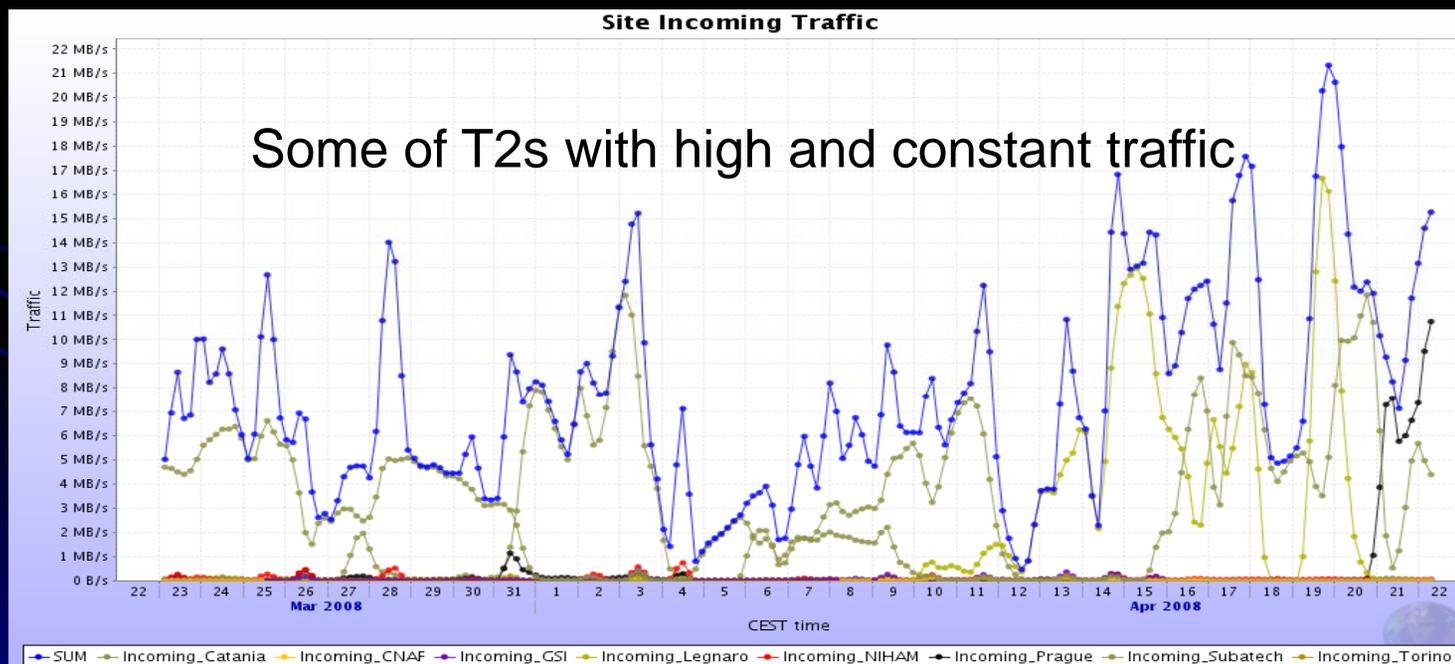
- **dCache** – stable, new development for support of advanced xrootd features is ongoing
  - CCIN2P3, GridKA, NDGF, NL-T1
- **CASTOR2** – stable (see later), most heavily used (current version 2.1.6)
  - CERN, CNAF
  - RAL – pending installation
- **dCache and DPM** at T2s (15 active SEs) – stable, used for MC production and user analysis

# T2s in CCRC'08

- Normal operation – MC production and user analysis
- Special focus on deployment of xrootd-enabled storage (dCache and DPM)
- Last 2 months - 15 new T2 storage elements in Germany, Italy, France, Czech Republic, Romania, Russia
  - Total capacity ~200TB
- Very good set of HowTo's and operation manuals developed with the active participation of T2 experts

# T2s in CCRC'08 (2)

- Storage is used for
  - MC and RAW data processing ESDs
  - Specific calibration runs
  - User storage

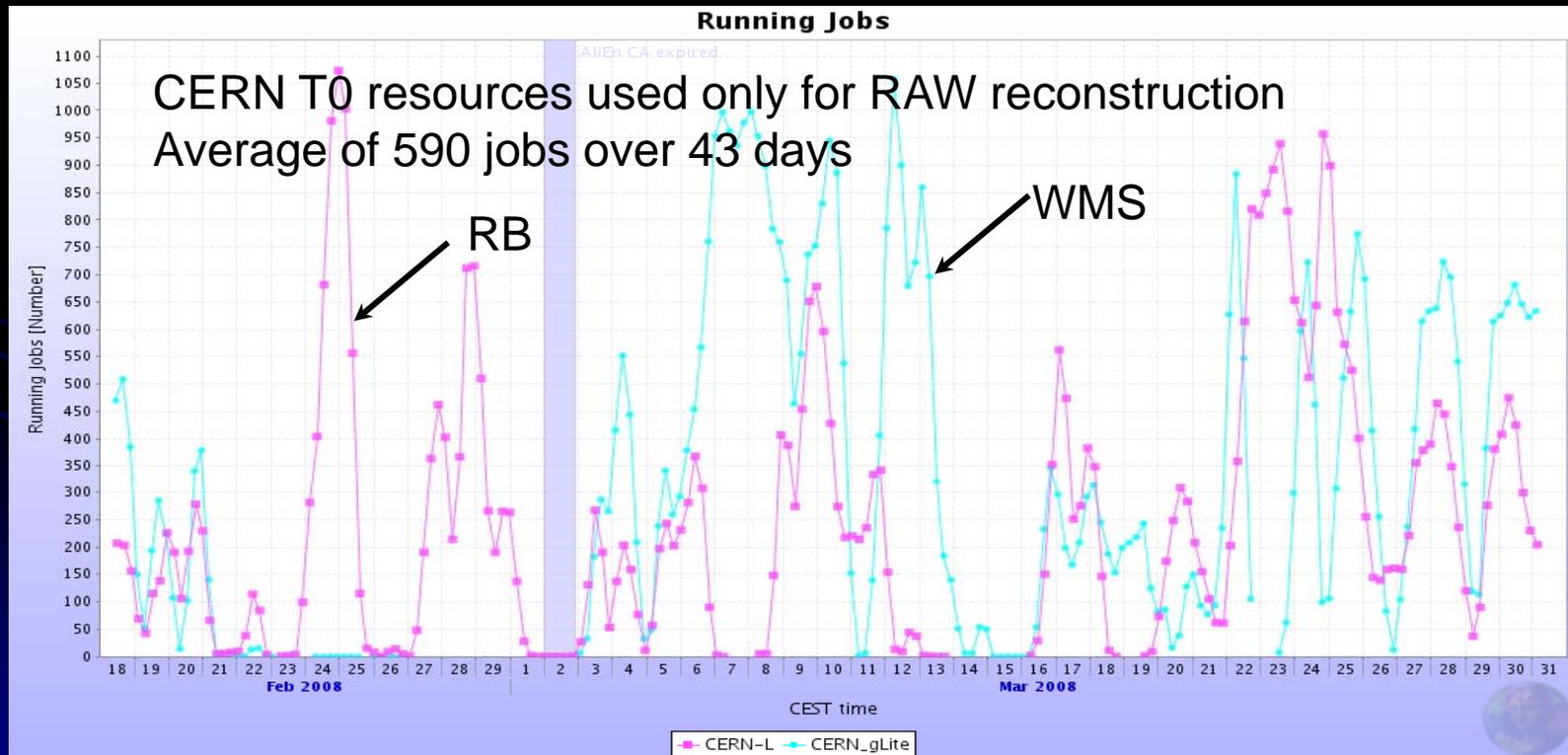


## T2s in CCRC'08 (3)

- The T2s (at present 62 centres) are providing ~50% of all ALICE resources
- Independently of CCRC T0/T1 activities, the T2 centres load and function remains constant (MC, analysis)
- **Very enthusiastic** site participation in storage deployment and debugging

# RAW data production

- Feb/March - Pass 1 reconstruction@T0
- Special data taking regime – several updates of the reconstruction software needed



## RAW data production (2)

- Mixed submission – LCG RB and gLite WMS
  - Very good stability of WMS
- Issues – blocked jobs due to data access
  - RAW data (cosmics) contains few track, very quick reconstruction (high I/O component)
  - As a consequence – concurrent access to the same storage server by many jobs
  - Problem understood, fix deployed
  - In addition: improvements in the ALICE job wrapper
    - stop the blocked process quickly (in development)

## RAW data production (3)

- Quasi-online production was difficult to achieve from the very start
  - First time the real detector data is 'seen' by the Offline code (especially as 'global' detector)
  - Frequent releases of reconstruction software is essential
  - In general, we have processed all RAW data taken in February/March
  - Pass 1 at T0 – **successful test of CASTOR2 with xrootd**
  - Pass 2 at T1s – **ongoing, with new and improved version of the reconstruction software**

# Access to storage

- RAW data production and user analysis – direct access to storage through xrootd
  - No copying of data to WN
  - **Better ratio CPU/Wall time**, especially for analysis jobs (I/O bound)
- Limits of storage and optimal configuration is being established through stress-tests
  - At CERN (CASTOR2) for RAW data reconstruction
  - Same at T1s (April/May)
  - At T2s (dCache, DPM) for analysis jobs
    - Up to 200-300 concurrent tasks

# Summary

- For ALICE, the February/March phase of CCRC'08 exercise was very useful
  - Focus on data management
  - Brings all experiments together
  - Controlled tests, organization
- The focus of the May phase and beyond will continue to be data management
  - At the T1s - mostly deployment of additional capacity
  - At the T2s – more complex, involves a period of tuning for every new T2

## Summary (2)

- In general, the data registration, transfer and processing tools worked well
  - Many thanks to the experts at CERN, T1s and T2s for the close support, feedback and suggestions
- ALICE is continuing to re-process the RAW data taken in Feb/March
  - In addition to the standard MC production and user analysis
- Plans and resources requirements for the May exercise will be presented tomorrow