ALICE Plans and requirements for May run and beyond

Latchezar Betev

WLCG Collaboration Workshop

23 April 2008

Offline tasks in May CCRC'08

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

Expected data volumes

- Standard p+p data taking:
  - 60MB/sec RAW
  - 120 TB total RAW volume /month
  - ~24TB total ESD volume / month (without re-processing)
- February/March – 70% of monthly volume achieved with real RAW data
- We expect to reach above 80% in May and 100% p+p rate in the subsequent months, up to the start of LHC

Storage requirement – T1s

- Additional resources needed for May exercise (80% p+p scenario)
- Disk will store ESDs from RAW
  - Assuming ESD+other files 20% of RAW

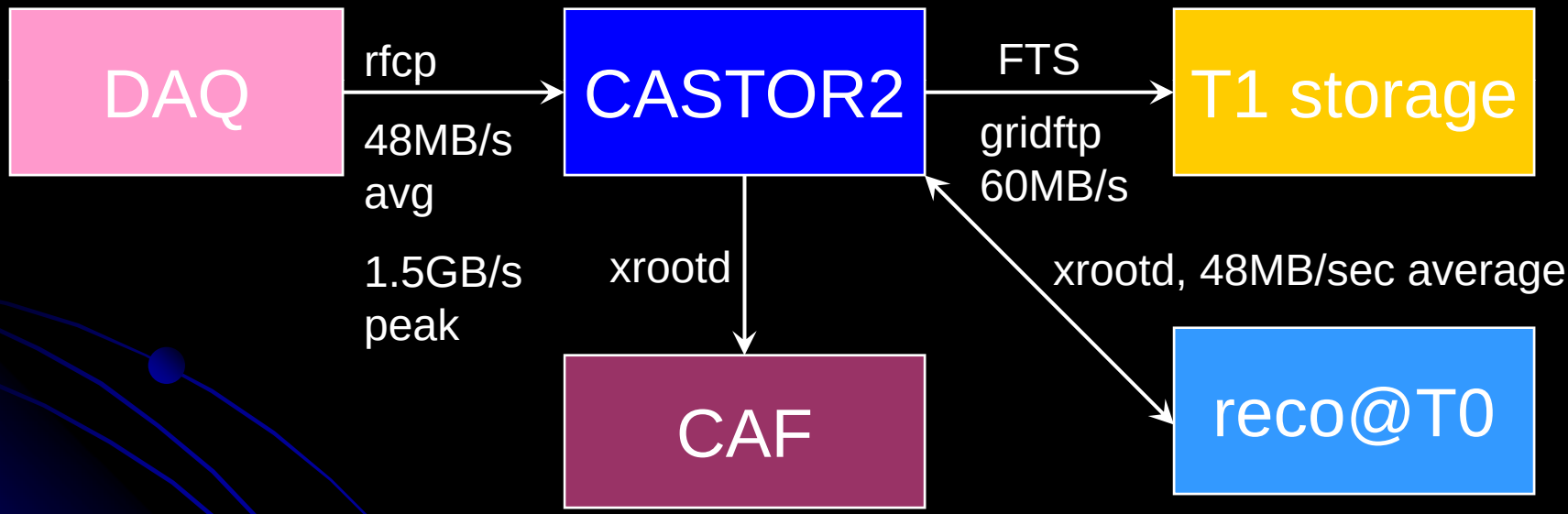
Tier 1 site	Disk space (TB)	Tape space (TB)
CCIN2P3 (15%)	3	14
CNAF (15%)	3	14
GridKA (45%)	9	44 (already deployed)
NDGF (15%)	3	14
RAL (5%)	1	5
T1-NL (5%)	1	5
TOTAL	20	96

Data management – T1s

- Three virtual SEs per T1 site, different mountpoints
  - T1D0 for RAW
    - NEW - T1D0ESDs for ESD production
    - The above is implemented at CERN and GridKA to optimize tape dirves access patterns
  - T0D1 for ESDs
  - T1D0R for complementary data (keeping 60MB/sec constant rate out of T0, limited use in May and beyond)

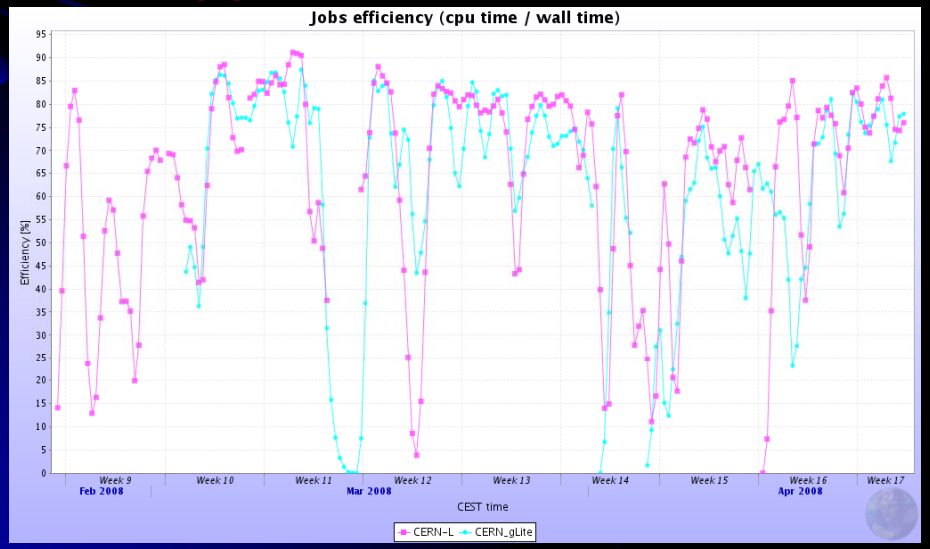
Data path and rates

Rates correspond to 80% standard p+p ALICE data taking rate (from DAQ) + full p+p data taking rate (replication)



Job duration/efficiency

- Difficult to predict – the job duration is determined by the quality of the RAW data
  - Amount of noise in the detector – big events, but little content
  - Trigger selectivity
- Substantial improvements expected over Feb/March status (jobs were mostly I/O bound)
  - Upper limit – 10hours with 10GB input file/job



Average efficiency for reconstruction jobs (last 2 months)

Jobs efficiency (cpu time / wall time)

Farm	Last value	Min	Avg	Max
CERN-L	76.03	0	63.93	100
CERN_gLite	77.91	0	62.07	97.83
Total	**76.97**		**63**	

MSS operation

- Issues
  - Average file sizes too low (inefficient tape writing)
  - Too many (repeat) mounts of same tapes (inefficient robot usage)
- From May – RAW data chunks of 10GB (from 1GB)
  - Presently being tested by ALICE DAQ group
- Pre-staging of data sets prior to processing and replication
  - Already in production at CERN

Services and criticality

ALICE Wiki: http://twiki.cern.ch/twiki//bin/view/ALICE/CCRC08

Critical service list

Rank	Definition	Max downtime (hrs)
10	Critical	2
7	Serious disruption	8
5	Reduced efficiency	12

Service	Rank	Comment
Site VO boxes	10	If restored within 2 hours - no loss of jobs
CASTOR2+xrootd@T0	10	Same as above
MSS@T1s	7	
FTS T0->T1	5	Will affect primarily second pass reconstruction at T1s
gLite WMS or RB	5	Failover mechanism 1/3 used, efficiency will be affected if all 3 fail
PROOF@CAF	5	Especially relevant during daytime

OS and software updates

- Ongoing – migration of all remaining sites/VO-boxes from SLC3 to SLC4
  - From May on, ALICE will no longer support sites still with SLC3
- Deployment of ALICE VO-box services (AliEn v.2-15 SLC4) starting next week
  - In parallel the new reconstruction/simulation applications (SLC4) will be deployed
- We do prefer to have the VO-box installation (i686, x86_64) corresponding to the site WNs installation
  - Avoid working in compatibility mode

OS and software updates (2)

- gLite 3.1 VO-box suite on x86_64
  - OK with compatibility libraries installed
  - Many thanks to CERN IT/GS (Patricia Mendez, Maarten Litmaath) for their help with native x86_64 installation
  - Instructions are in preparation
- The migration of VO-boxes is ongoing, working directly with the sites

T2s – MC and analysis

- Two new rounds of MC production planned
  - Different LHC collision energy startup scenarios
- Continue (aggressive) deployment of xrootd-enabled storage (dCache and DPM)
- Analysis of MC and ESDs from RAW data

Calendar of activities

- ALICE commissioning exercise (RUN III) – data taking starts 18 May
  - Registration/replication/quasi-online reco
- Prior to that, we will reconstruct February/March RAW @T1s with new version of AliRoot
- MC production
  - Round 1 – from 5 May to 15 May
  - Round 2 – from 16 May
- If required by WLCG, ALICE can participate to the common data transfers T0->T1 from 5 may onward

Summary

- The May phase of CCRC'08 will be a continuation of February/March exercise
- The operations are governed by the ALICE detector commissioning exercise (RUN III)
  - The experiment plan is to be in continuous data-taking from 18 of May onward
- Focus is on
  - Data management (deployment of storage at T2s, optimization of storage at T0/T1)
  - Quasi-online processing of RAW data