



Plan for shutdown and data-taking preparation

Latchezar Betev
ALICE

Registration/replication of RAW

- Data volumes (2008)
 - Total 310 TB at CERN + 200 TB replicated at T1s
 - Includes cosmic data, detector calibration
- Tools and procedures
 - T0 registration -> CASTOR2 and AliEn file catalogue
 - Routine operation, high redundancy of servers and software
 - Problem-free operation
 - All updates carried out transparently by IT/FIO
 - Replication FTD/FTS
 - Tested extensively@CCRC08
 - All goals (rate/stability) met and exceeded
 - Routine operation during data taking

Registration/replication of RAW (2)

- Planned updates
 - Move to FTS on SCL4 – done
 - Validation of new SRM endpoint at CERN
 - Indirectly through FTS – December
 - Update of castoralice@cern to v.2.1.8
 - As soon as IT/FIO plans allow it

Registration/replication of RAW (3)

- Data cleanup
 - Some of the collected datasets have been identified as candidates for removal
 - Old MC productions are no longer needed
 - Ongoing discussion detector groups and physics coordination
 - Preliminary estimate – ~600GB of master copy+replicas
 - Discussed with CERN Castor experts
- Policy for tape replicas to be discussed with T1s
 - Possible complete scratch of RAW@T1 (tapes) and re-replication
 - Benefit – new scale test of FTS + re-spin of tapes

Registration/replication of RAW (4)

- No major issues
- Confidence in the storage and Grid tools is high
 - Middleware and computing centres storage (T0 and T1s) have been fully certified

Offline reconstruction

- Significant improvements of the processing configurations for all detectors
- Automatic configuration of detector reconstruction parameters implemented and tested
- Quasi-online reconstruction status
 - Reconstruction trigger system tested
 - Conditions data present, run registered at T0, Online QA completed
 - Pass 1 reconstruction at T0 steered automatically

Offline reconstruction (2)

- All reconstructible runs from 2008 cosmics data taking are processed
 - Emphasis on 'First physics' detectors
 - Selected runs already re-processed as 'Pass 2'
- Re-processing of all cosmics data – general 'Pass 2'
 - After completion of alignment and calibration studies by detectors
 - Detector code improvements in AliRoot
 - Processing machinery on standby

Offline reconstruction (3)

- Development of quasi-online processing framework
 - Further refinement of Online QA – requires detector development
 - Speed up the launch of reconstruction jobs to assure ‘hot copy’ of the RAW data
 - This is ongoing process, new pieces will be tested gradually as they become available with (monthly) AliRoot releases
- The middleware and fabric are fully tested for ‘pass 1’ (T0) RAW data processing

Conditions data - Shuttle

- Shuttle (subsystem DBs to Grid conditions data publisher) system is in operation since 2 years
 - In production regime for the whole 2008
- Detector algorithms (DAs) within Shuttle have evolved significantly since last period
- High stability of primary sources of conditions data: DCS, DAQ DBs and configuration servers
- Toward the end of last cosmics data taking period (August) – all pieces, including DAs 'Green'

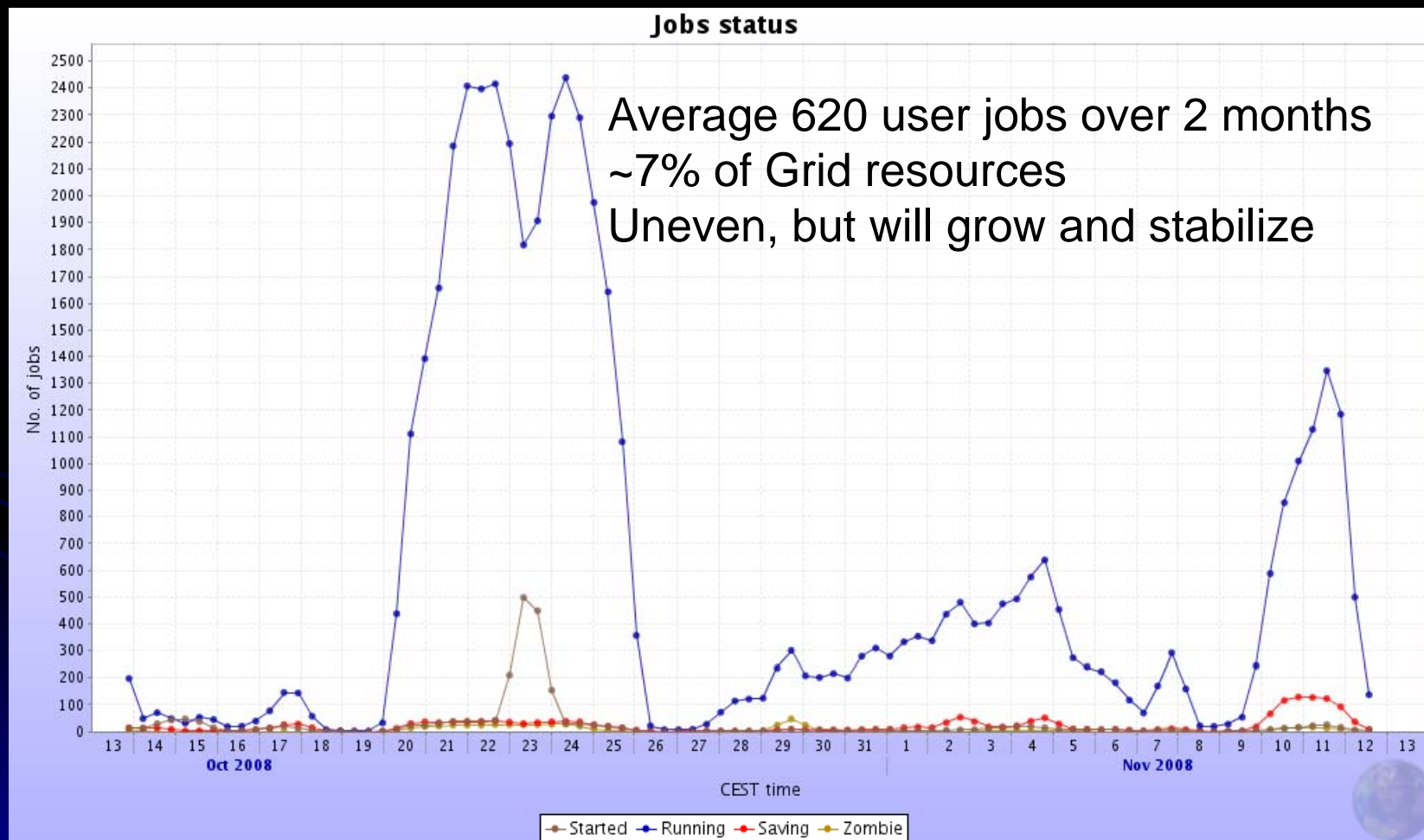
Conditions data – Shuttle (2)

- Major efforts concentrated on filling in the missing conditions data
 - Critical LHC parameters
 - Newly installed detectors and control hardware
- Some improvements on replication of conditions data on the Grid (file replicas)
 - So far, ALICE maintains 2 full replicas
- Conditions data access is the area with least problems on the Grid
 - Both in terms of publication and client access

Grid batch user analysis

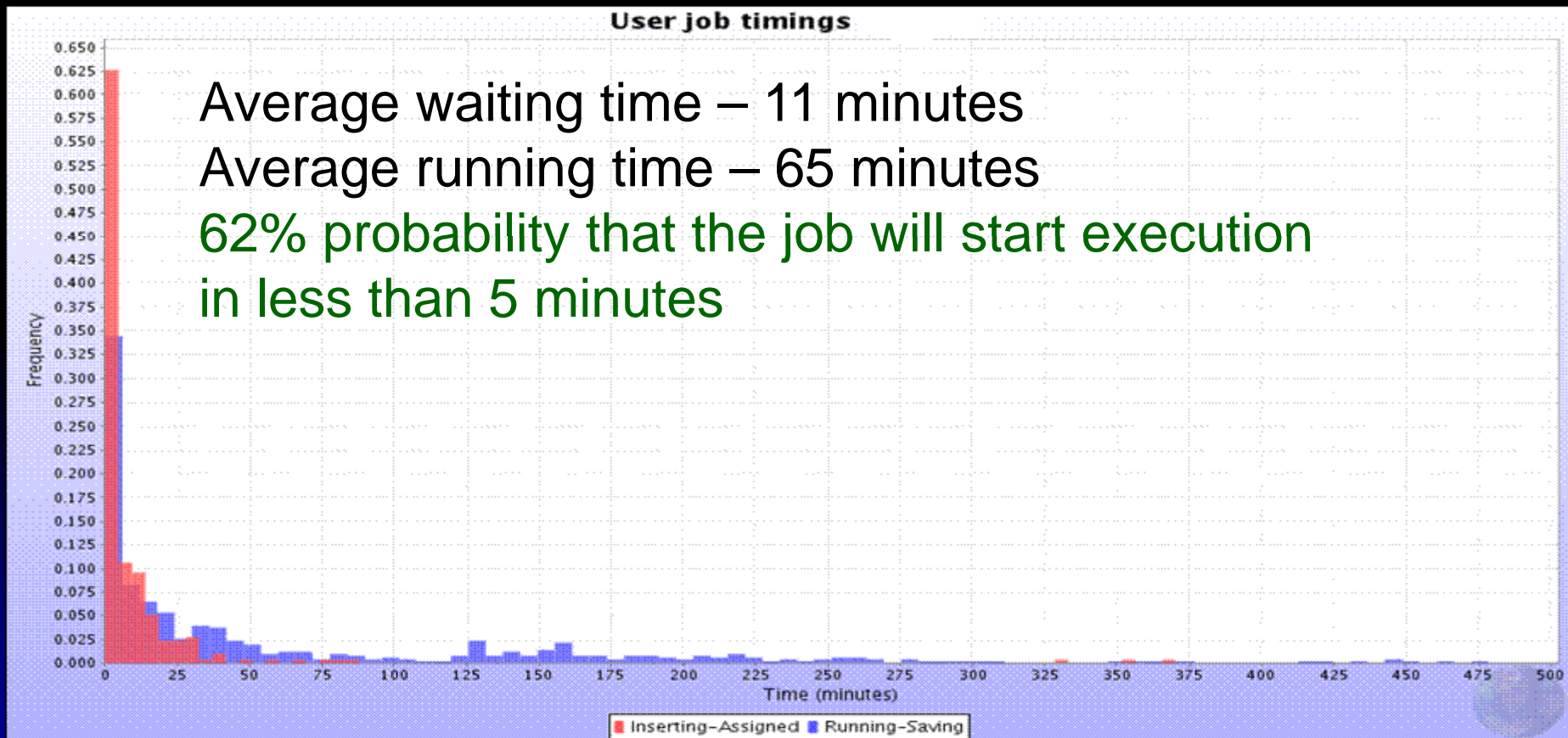
- **High importance task**
- Predominantly MC data analysis
 - New productions will accelerate the trend and user presence
- RAW (cosmics) data is still to be re-processed to qualify for analysis
 - First pass has been extensively analyzed already
- Ever increasing number of users on the Grid
 - 426 registered, ~120 active

User job distribution



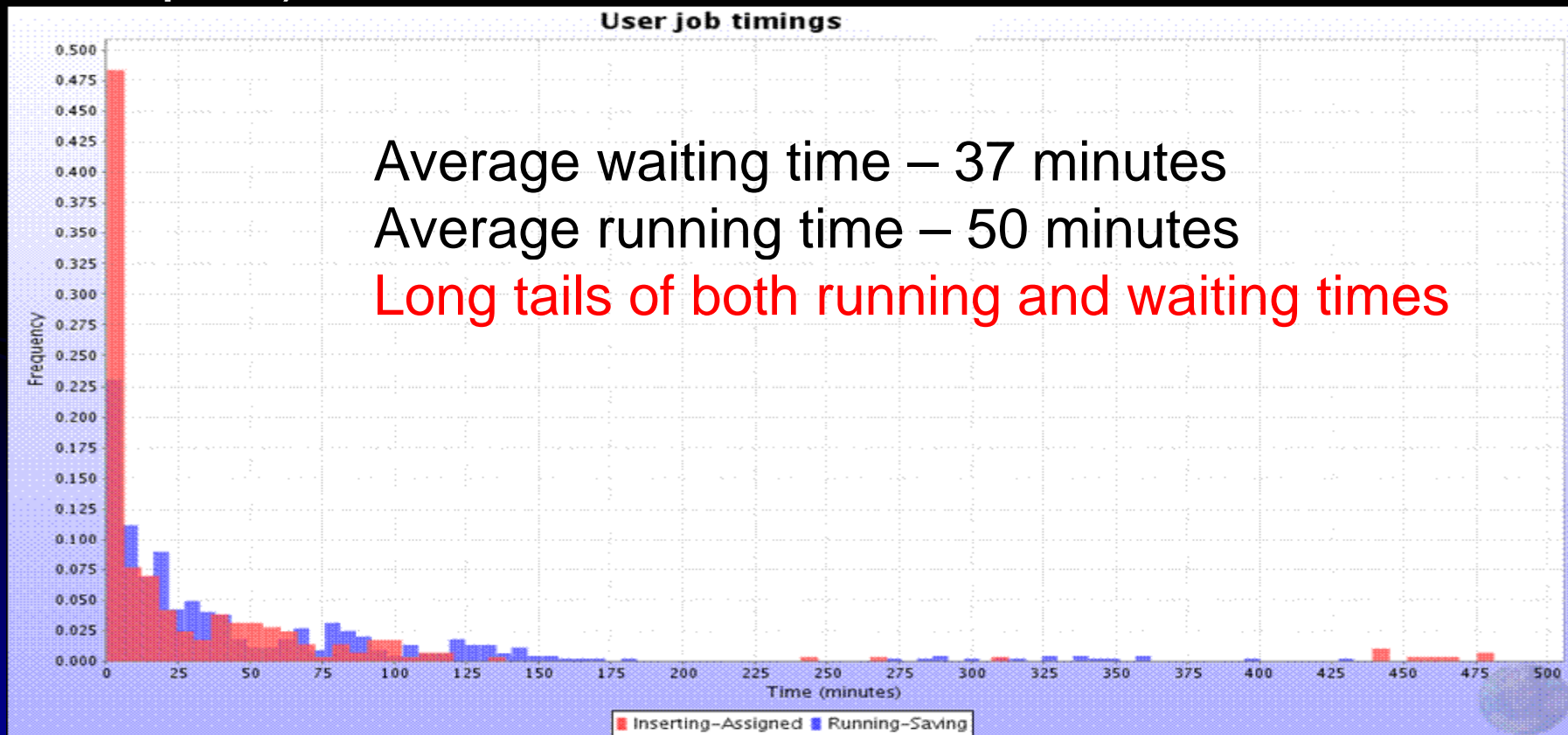
Grid responsiveness – type 1

Jobs not requiring large input data (typically MC)



Grid responsiveness – type 2

Jobs exclusively running on ESDs (large input data samples)



Grid batch analysis

- Internal ALICE prioritization within the common task queue works well
 - Production user is 'demoted' in favor of normal users in the queue
 - Generic pilot jobs assure fast execution of user jobs at the sites
- Case 1 (MC) jobs can be regarded as the 'maximum Grid efficiency'
 - Until there is contention for resources between users (*)
 - (*) is being addressed by the 'banking system', which allows users to 'pay' for higher task queue priority

Grid batch analysis (2)

- Case 2 (ESDs) can be improved
 - Trivial - more data replication
 - Not an option – no storage resources
 - Analysis train – grouping many analysis tasks in a common data set – is in active development/test phase
 - Allows for better CPU/Wall and even load on the storage servers

Prompt analysis

- PROOF enabled facilities currently available at two sites
 - CAF@CERN
 - GSIAF@GSI Darmstad
- Extremely popular with users
- Fast processing of
 - MC ESDs – critical for first physics
 - Calibration and alignment iterative tasks
- CAF@CERN was recently upgraded with higher CPU and local disk capacity
- ALICE encourages centres to install analysis facilities for the local physics community
- Good integration with the Grid on the level of data exchange

Job submission to WMS

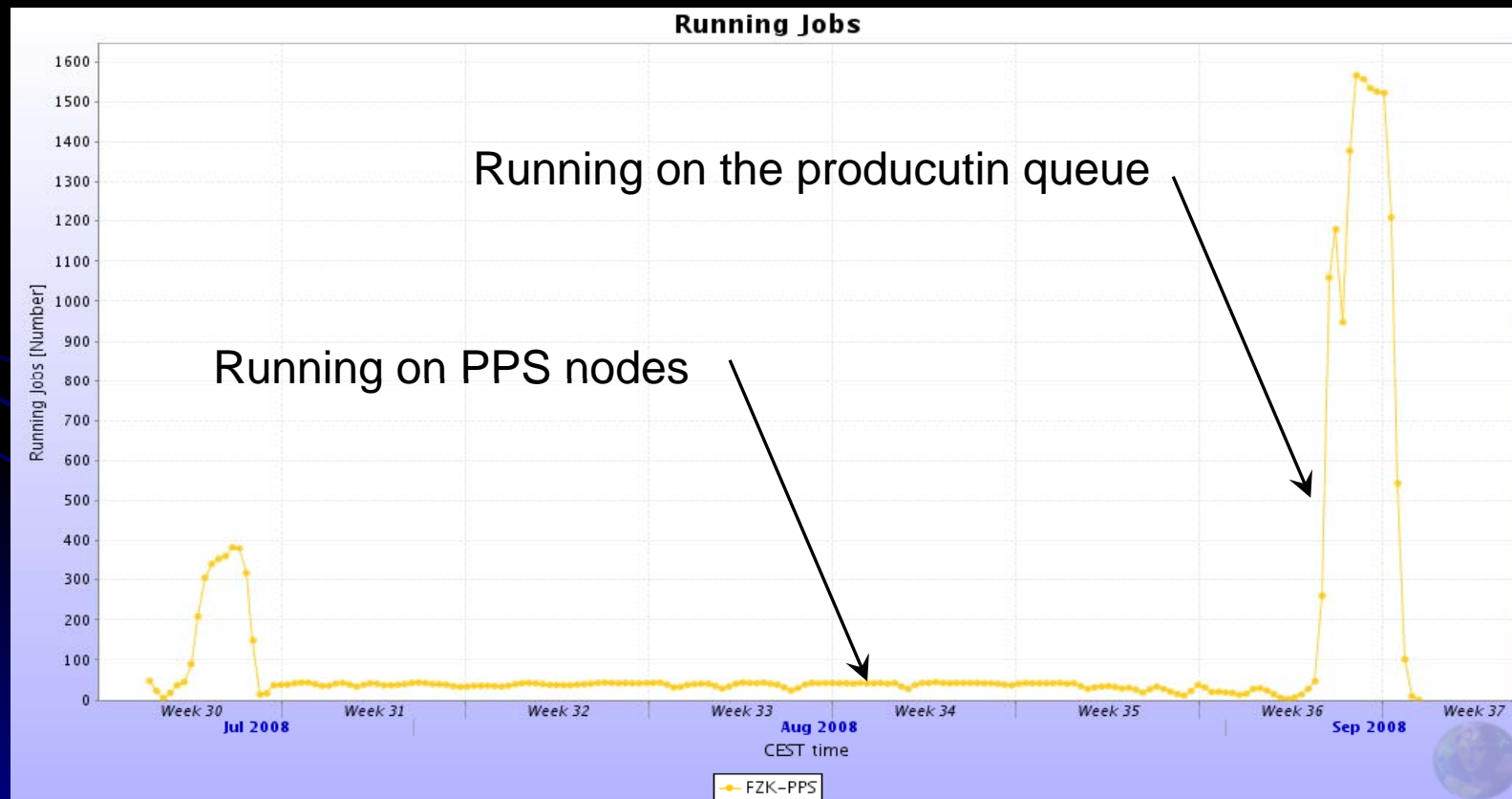
- ALICE sites migration to WMS submission
 - Completed last week
 - ALICE is **RB free**
- Configuration of WMS submission
 - WMS-specific parameters to be implemented in the ALICE job submission system
 - Regular discussion with WMS experts at CERN – very efficient support of IT/GS

Job submission to WMS (2)

- In general – ALICE has no concerns with regards to job submission through WMS
- ALICE **does not** require WMS at every site
- ALICE **requires** sufficient number of WMS instances per region (typically country) to support the workload

CREAM CE

- Test CREAM-CE instance provided at GridKA
- Intensive functionality and stability tests from July to September 2008



CREAM CE (2)

- Impressive service stability – maintenance free for the entire period
- Mode of operation
 - Second VO-box for simultaneous submission to gLite-CE and CREAM-CE (T1s)
 - Change of configuration parameter on the primary VO-box
- Deployment (following recommendations of GD)
 - CREAM-CE instance in parallel with gLite CE
 - Negotiations ongoing with sites

SLC5

- Presently tested at CERN PPS
 - ALICE code – compatibility mode (SLC4 compiled)
 - SLC5 32-bit WNs – completed
 - SLC5 64-bit WNs – ongoing
- ALICE wishes to move to SL(C)5 at all sites as soon as possible
- The ALICE code is already ported and is being validated

High volume MC production

- New round of production in preparation
- Modifications of AliRoot code reflecting
 - Inclusion of new detectors in 2008
 - Change in existing detector configuration
- Following the requirements of ALICE Physics Coordination/Physics working groups
- Expected start – December 2008
- Presently running short test production cycles

Storage

- New storage (at T2s)
 - Gradually increasing the number of sites with xrootd-enabled SEs
- Every SE is validated and tuned independently
- Emphasis on disk-based SEs for analysis
 - Including a capacity at T0/T1s
 - Storage types remain unchanged
 - T1D0 for RAW, production, custodial ESD/AOD storage
 - T0D1 for analysis: ESD/AOD replicas

Storage – MSS use optimization

- File size increase

- RAW 1GB->10GB - presently at 3GB due to event size/processing time needed
- Secondary files 100MB->1GB
- Aggressive use of archives, containing multiple files

- File access patterns

- Emphasis on prompt replication and reconstruction – files are still in the MSS disk buffer
- Analysis and chaotic file access moved away from MSS SEs

Summary of shutdown period activities

- RAW and old MC data – partial cleanup of storages (November 2008 – January 2009)
- Replication tools – moved to new FTS/SRM, testing through re-replication of portions of RAW (December 2008 – January 2009)
- RAW data processing (better code and detector calibration) – second pass reconstruction of cosmics data (January 2009 – February 2009)

Summary of shutdown period activities (2)

- Grid batch analysis – introduction of analysis train for common analysis tasks (January 2009)
- New middleware
 - Fine tuning of WMS use (December 2008)
 - Migration to CREAM-CE (December 2008 -?)
 - Move to SL(C)5 – as soon as possible
- High volume p+p MC production (December 2008 – March 2009)

Summary of shutdown period activities (3)

- Storage
 - Validation of new SEs (Ongoing)
 - Analysis SEs at T0/T1s (Ongoing)
- ALICE detector end of upgrade period (March 2009)
- New round of cosmics and calibration data taking (April-Start of LHC)
 - The standard data-taking ready regime of Offline will start in April as well