ALICE Physics Data Challenge '05 and LCG Service Challenge 3

Latchezar Betev / ALICE

Geneva, 6 April 2005

LCG Storage Management Workshop

Outline

➢ Goals and structure of PDC'05

➢ Basic parameters and participation plans

➢ Services, processing layout

➢ Timeline

Goals and structure

➢ The transfer framework can only be discussed as a part of the overall ALICE Grid strategy

➢ PDC'05 : Test and validation of the remaining parts of the ALICE Offline computing model:
  - ➢ Quasi-online reconstruction of RAW data at CERN (T0), without calibration
  - ➢ 'push' of the data from CERN to T1's
  - ➢ Second phase (delayed) reconstruction at T1's **with calibration** and remote storage

➢ ***Do all of the above <u>again</u> ENTIRELY on the GRID***

➢ Structure – logically divided in three phases:
  - ➢ Phase 1 – Slow production of events on the GRID, storage at CERN
  - ➢ Phase 2 (during SC3) – Pass 1 reconstruction at CERN, push data from CERN to T1's, Pass 2 reconstruction at T1s with calibration and storage:
    - ➢ Phase 2 (throughput phase of SC3) – how fast we can push data out
  - ➢ Phase 3 – Analysis of data (batch) and ***interactive analysis with PROOF***
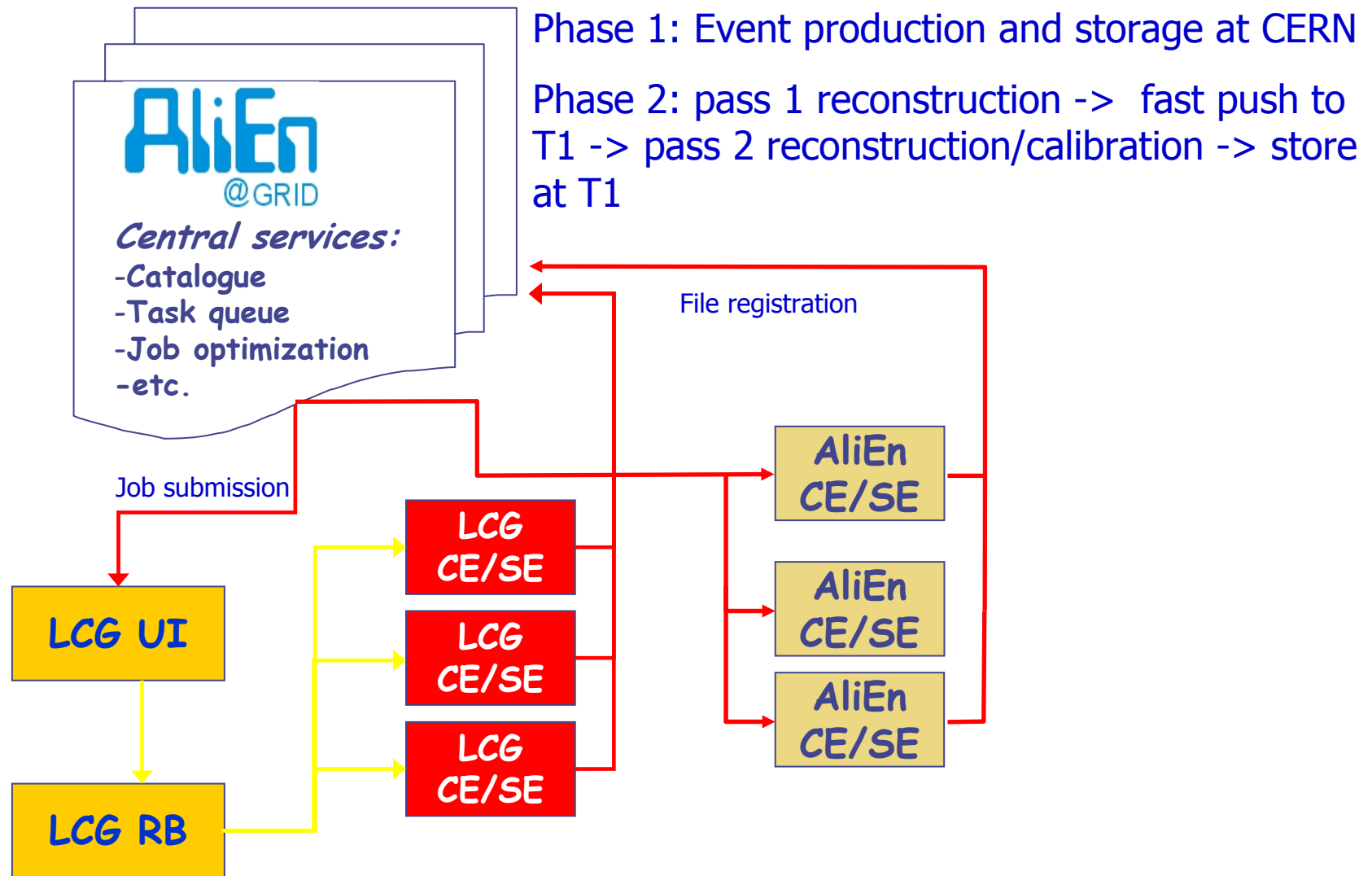
Parameters of the exercise

➢ Data production:

- ➢ Physics driven, list of physics signals to be produced is being collected now

- ➢ The processed data has to be preserved – capacity for its storage has been pledged by the participating computing centres for ALICE

- ➢ Potentially as large data volume as in PDC'04 (50-80TB) – ALICE detector and physics studies require new MC generated signals and improved statistics of the already produced ones

➢ SC3 - ALICE participation currently being discussed:

- ➢ Computing and storage capacities

- ➢ Network capacity: reserved/shared and elements (PDC data transfer volume ~200 TB)

- ➢ Availability of services (more later)

- ➢ Monitoring: experiment specific and interaction with the SC3 tools

Services availability

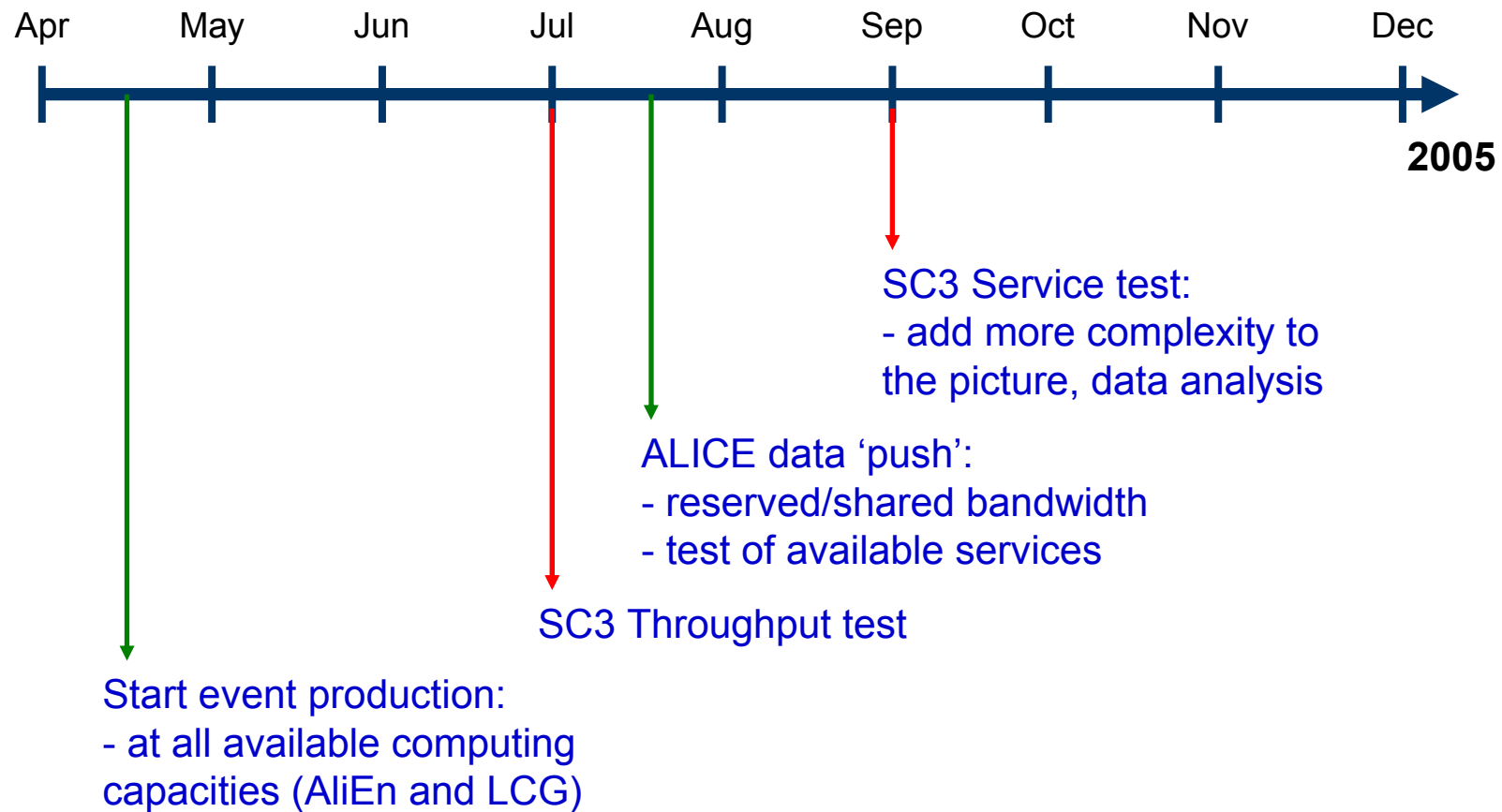- Transfer and storage services (part 1 of SC3):
  - GridFTP
  - gLite FTS
  - SRM
  - Experimental solution
  - Other
    - How will these become 'stable services' for the experiments and with what functionality
    - Be deployed on sufficiently large computing and storage capacities on the sites (more than on test-beds)

- Other Grid services (part 2 of SC3):
  - We assume that EGEE middleware available on LCG-SC3

ALICE PDC'05 layout

Phase 1: Event production and storage at CERN

Phase 2: pass 1 reconstruction -> fast push to T1 -> pass 2 reconstruction/calibration -> store at T1

AliEn @GRID

Central services:
- Catalogue
- Task queue
- Job optimization
- etc.

File registration

Job submission

LCG UI

LCG RB

LCG CE/SE

LCG CE/SE

LCG CE/SE

AliEn CE/SE

AliEn CE/SE

AliEn CE/SE

PDC'05 timeline

Apr   May   Jun   Jul   Aug   Sep   Oct   Nov   Dec

2005

SC3 Service test:
- add more complexity to
  the picture, data analysis

ALICE data 'push':
- reserved/shared bandwidth
- test of available services

SC3 Throughput test

Start event production:
- at all available computing
capacities (AliEn and LCG)

PDC'05 continuation

- Last step:
  - Analysis of the reconstructed data
- That is at the end of SC3:
  - We have some more time to think about it