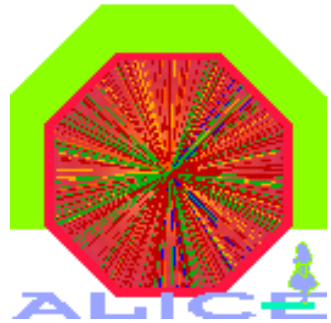
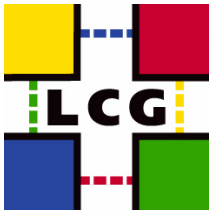
ARDA Workshop, CERN March, 8th, 2005



"ALICE @ ARDA" Distributed Analysis on LCG

P. Cerello on behalf of the ALICE Offline Team



Outline

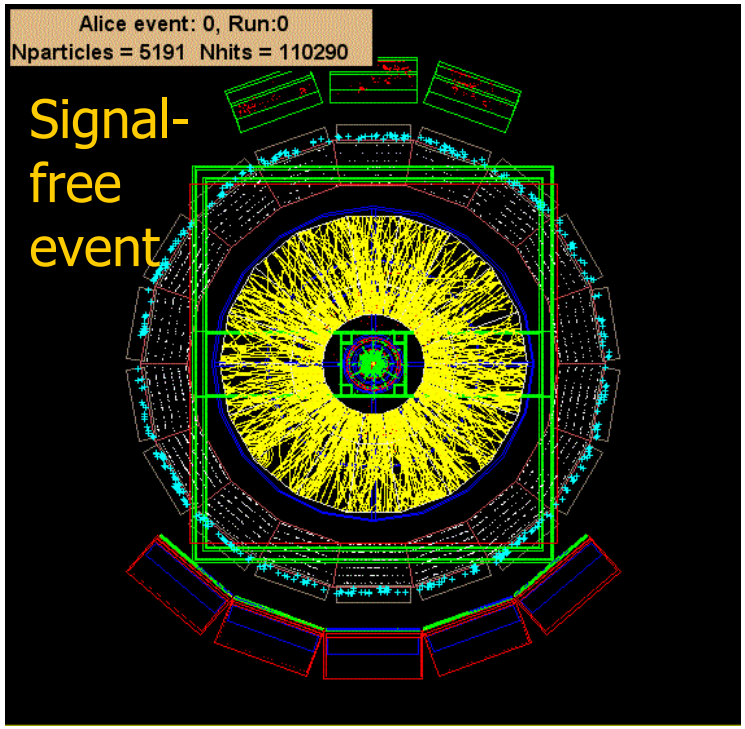
1. **Where do we stand?**
2. Where do we go?

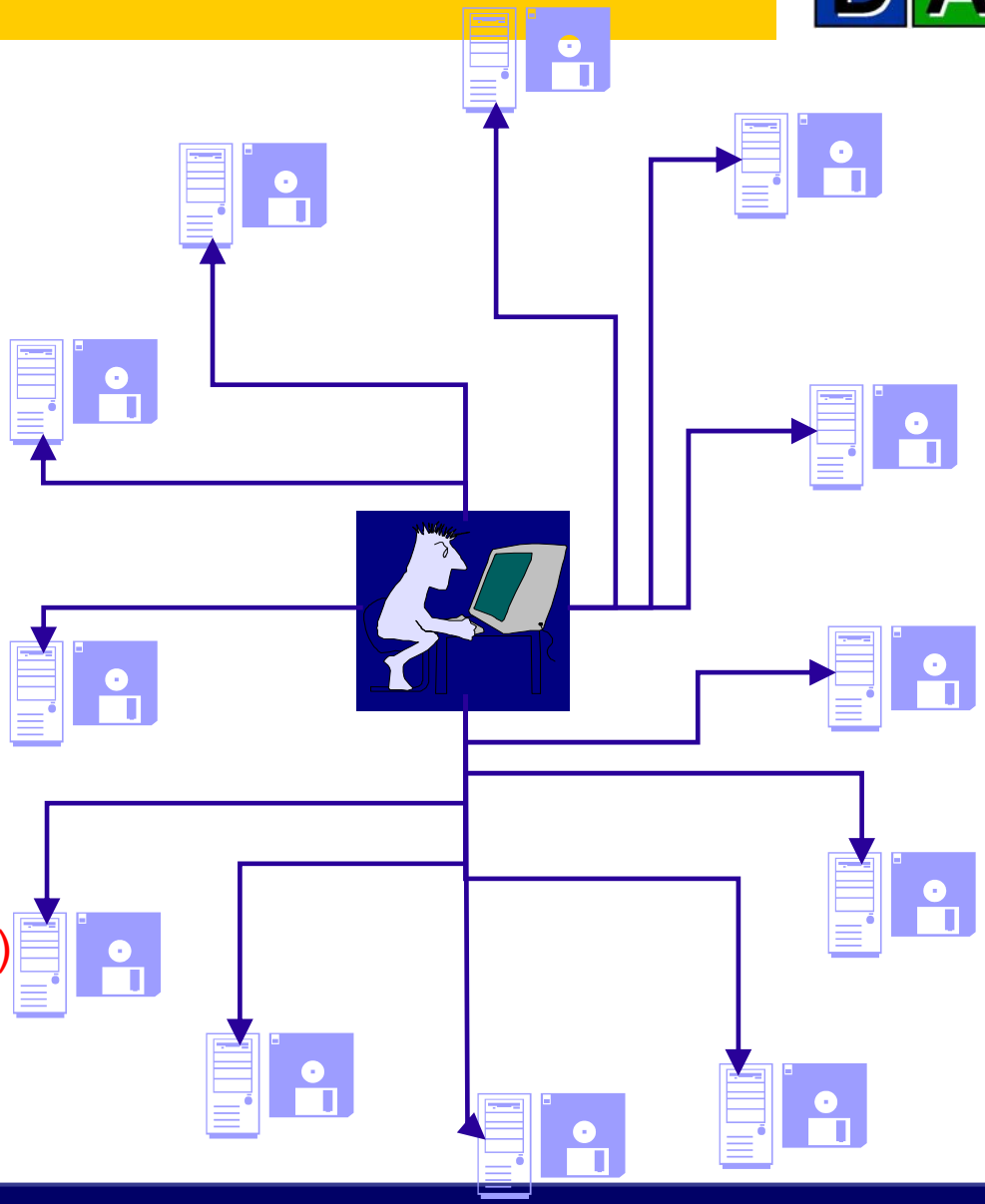
DC2004: Goals, structure, tasks

- Test and validate the ALICE computing model
  - Produce and analyse ~10% of the data of a standard data-taking year
  - Use the complete offline chain: AliEn, AliROOT, LCG and in Phase 3 – gLite+PROOF and the ALICE ARDA analysis prototype
  - ***Test*** of the software and ***physics analysis*** of the data for the PPR
- ***Do all of the above ENTIRELY on the GRID***
- Structure – divided in three phases:
  - Phase 1 - Production of underlying Pb+Pb and p+p events
    - Completed on time June 2004
  - Phase 2 - Mixing of different signal events with underlying Pb+Pb events (up to 50 times)
    - Completed on time September 2004
  - Phase 3 – Distributed analysis
    - Suspended

Phase I - Simulation

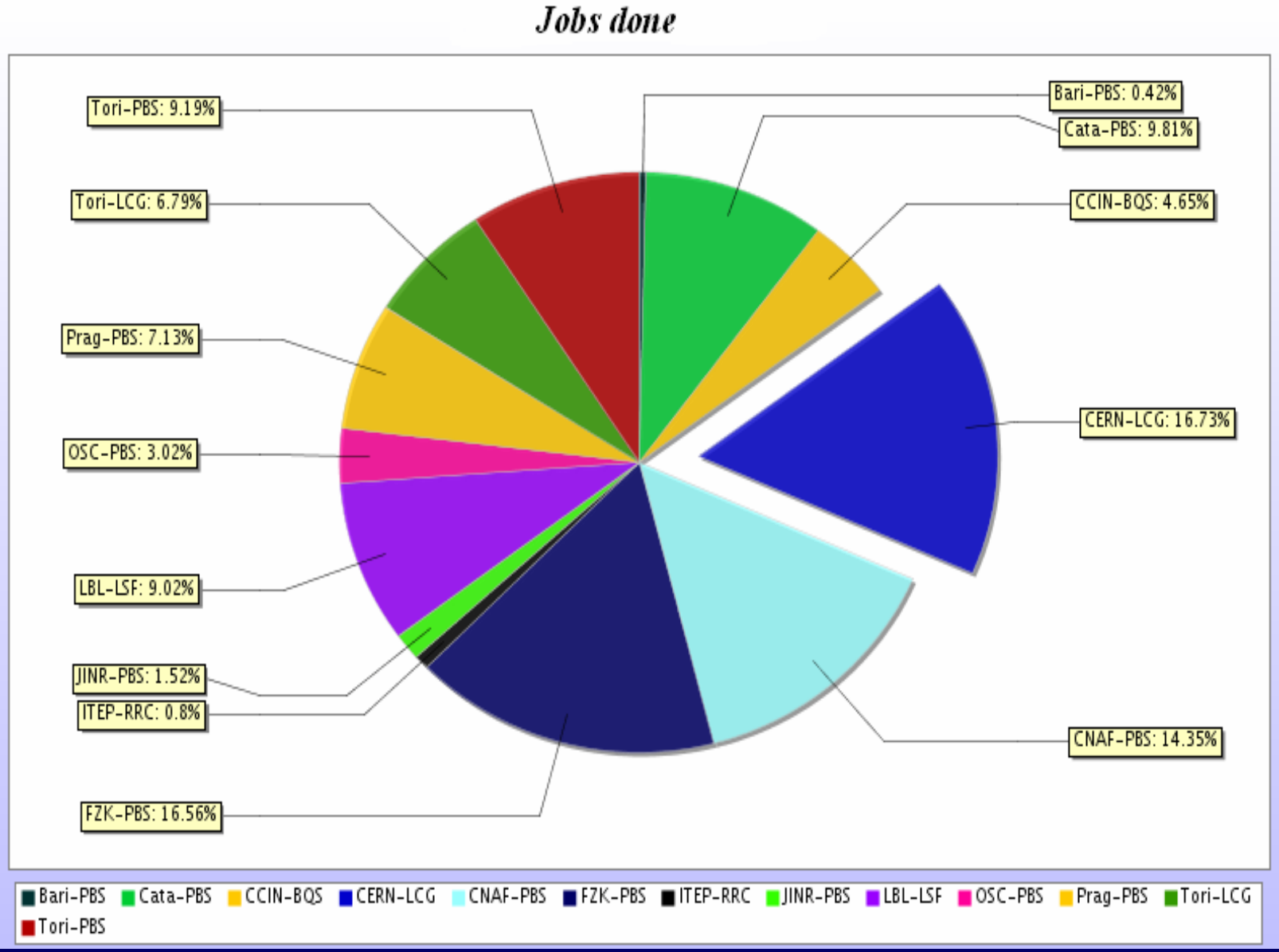
Small input with interaction conditions





Large **distributed** output (1 GB/event)
with simulated detector response
Long execution time (10 hours/event)

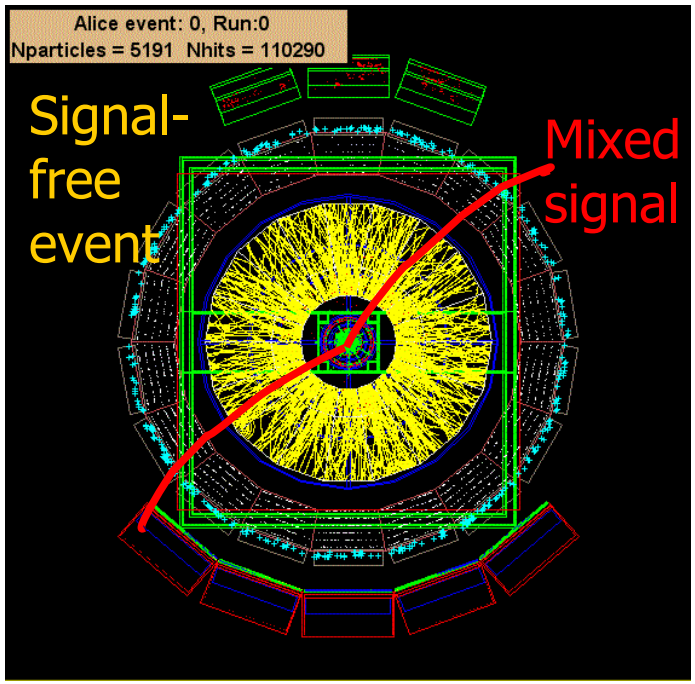
Phase I results



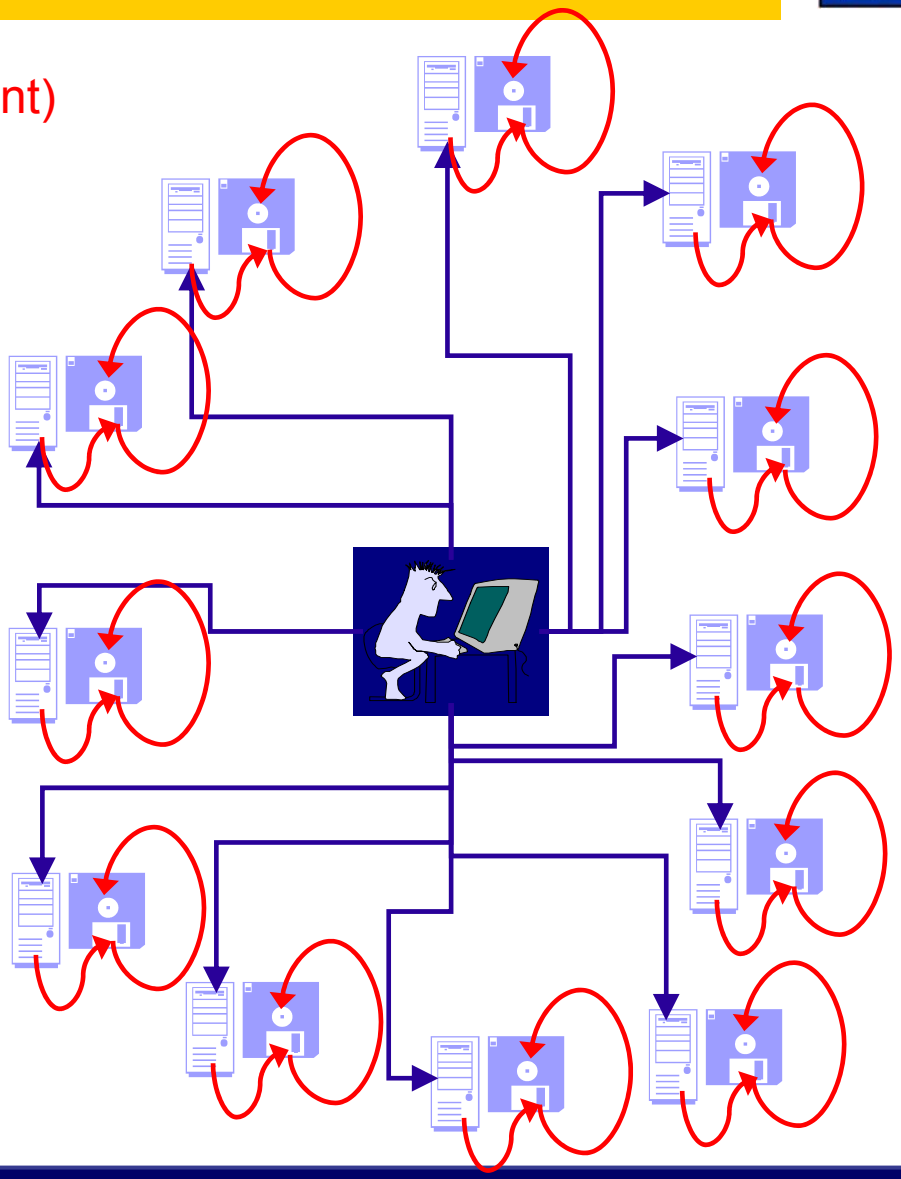
not

Phase II - Merging & Reconstruction

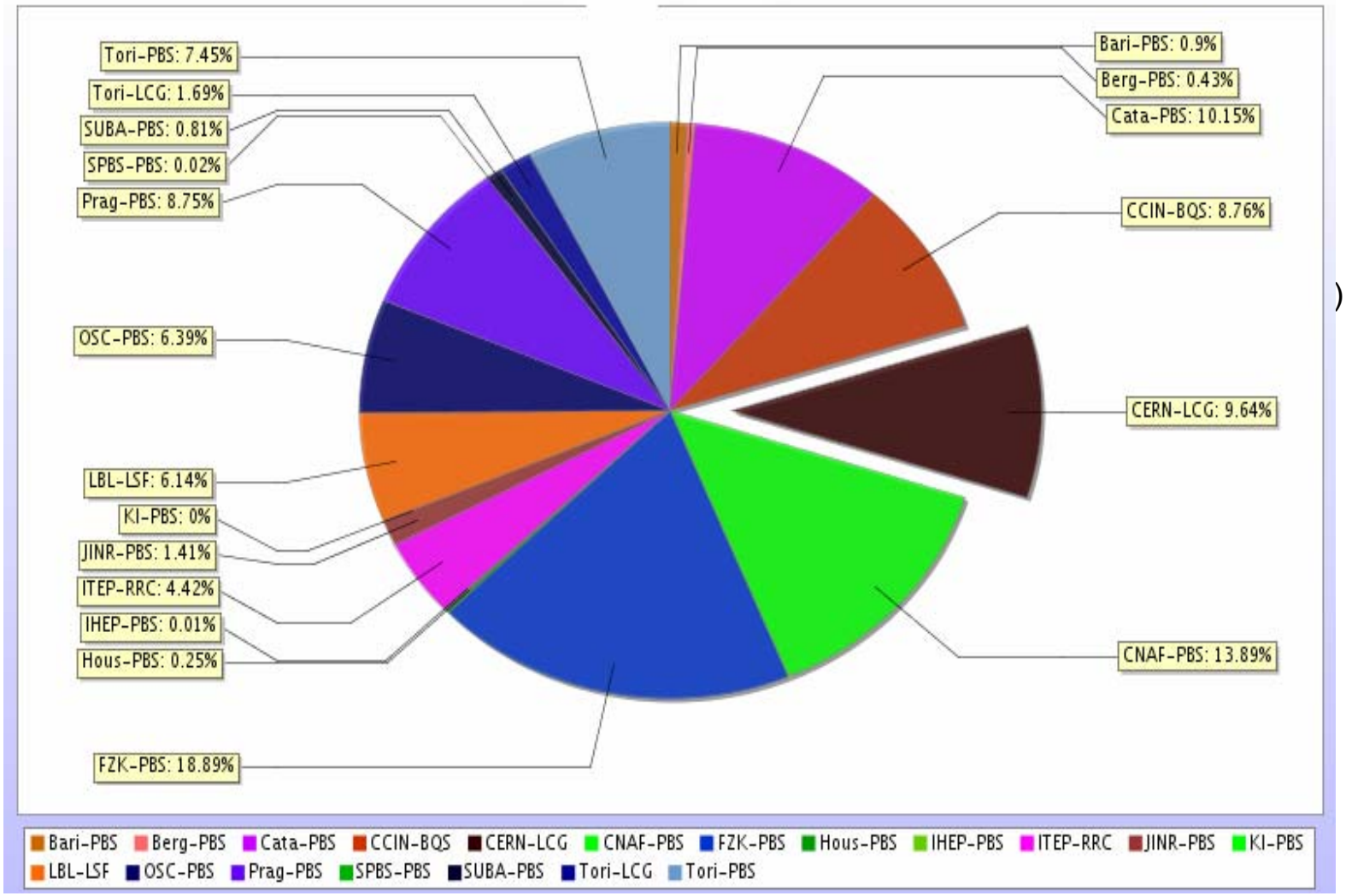
Large **distributed** input (1 GB/event)



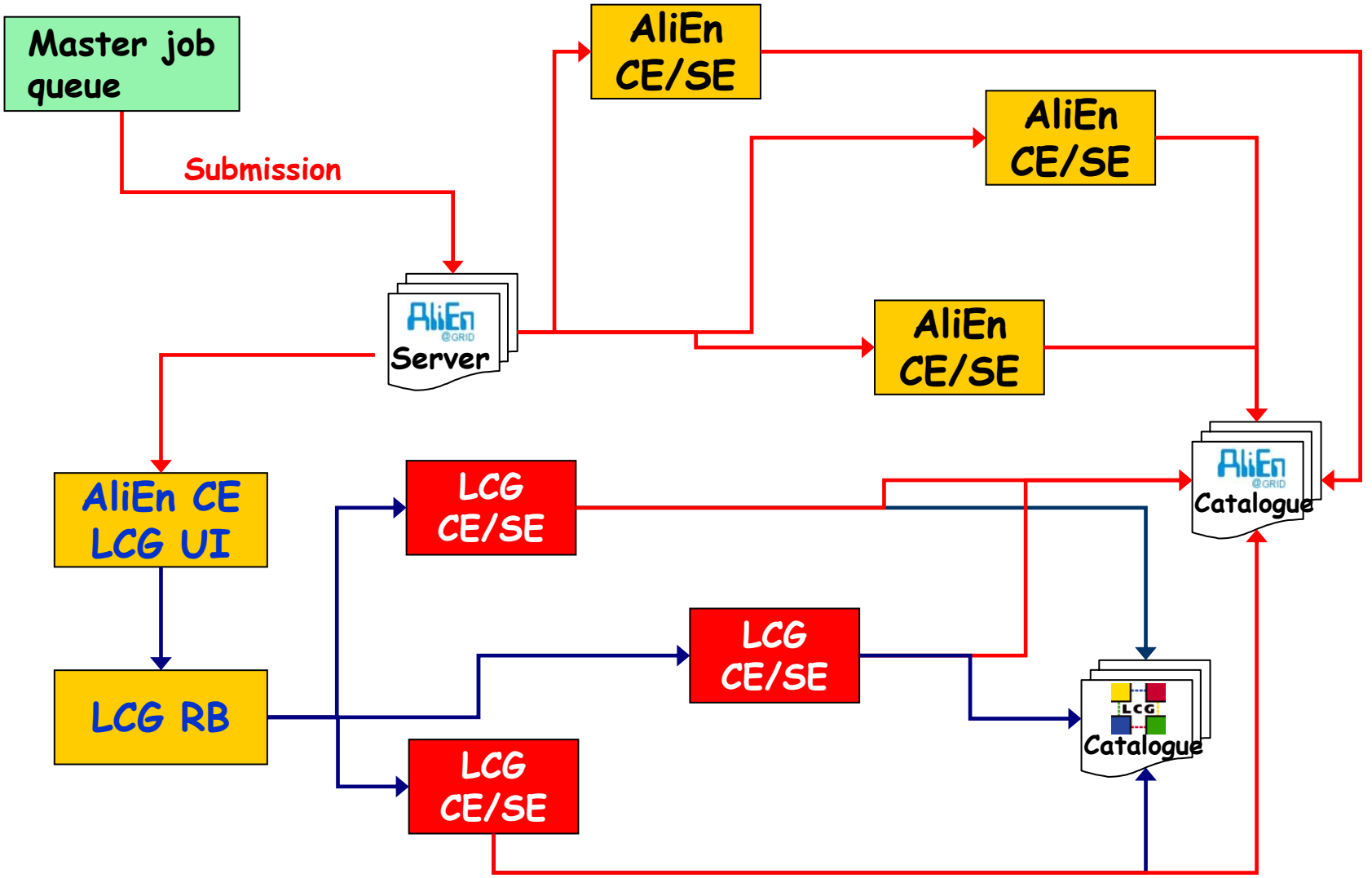
Fairly Large **distributed** output (100 MB/event, 7MB files) with reconstructed events



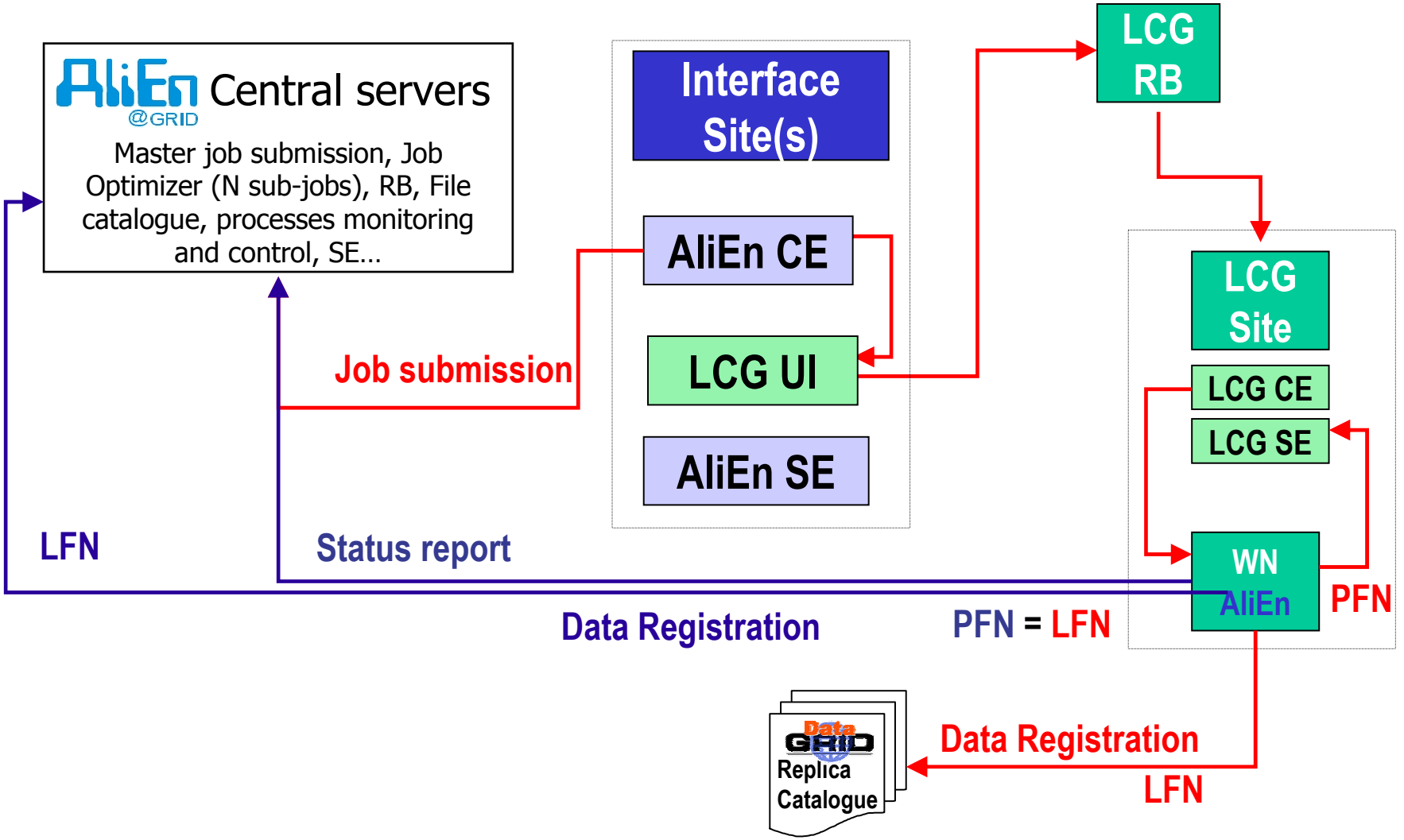
Phase II results



Phase I & II layout: a "Meta-Grid"



DC2004 Phase I & II

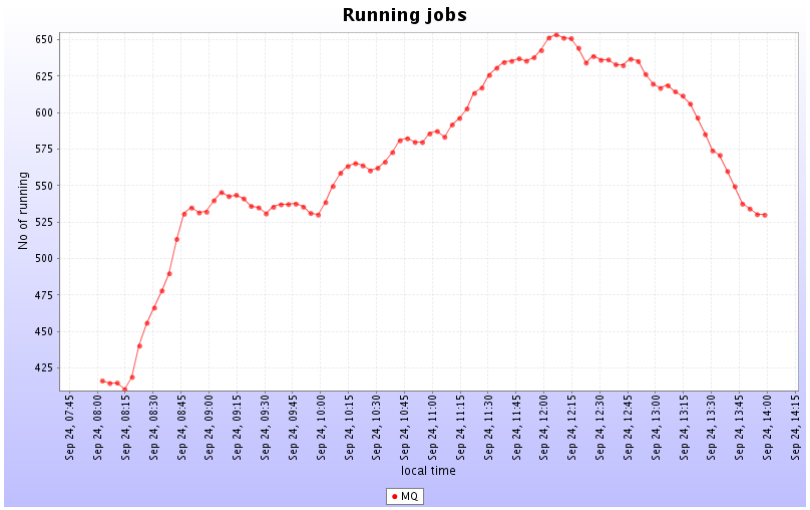


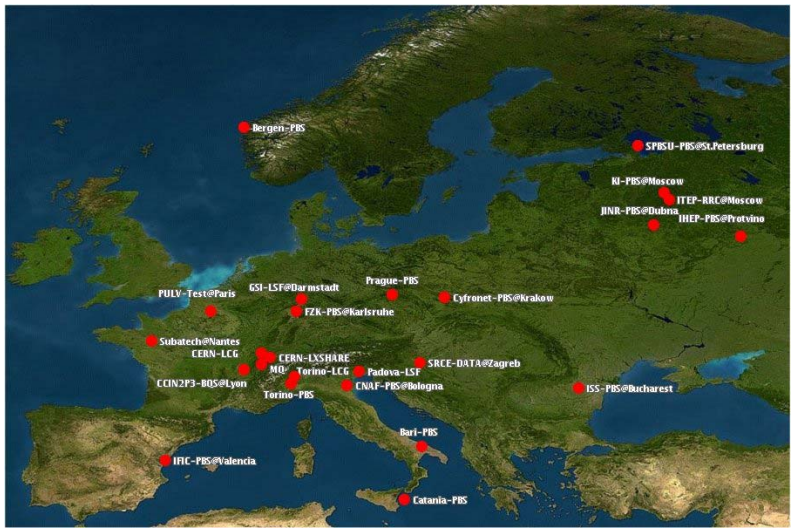
Production history

- ALICE repository – history of the entire DC
- ~ 1,000 monitored parameters:
  - Running, completed processes
  - Job status and error conditions
  - Network traffic
  - Site status, central services monitoring
  - ….
- 7 GB data
- 24 M records with 1 minute granularity – analysed to improve GRID performance

MonALISA
MONitoring Agents using a Large Integrated Services Architecture

- Statistics
  - 400 000 jobs, 6 hours/job, 750 MSi2K hours
  - 9M entries in the AliEn file catalogue
  - 4M physical files at 20 AliEn SEs in centres world-wide
  - 30 TB stored at CERN CASTOR
  - 10 TB stored at remote AliEn SEs
  - 200 TB network transfer CERN –> remote computing centres
  - AliEn efficiency observed >90%
  - LCG observed efficiency 60% (see GAG document)



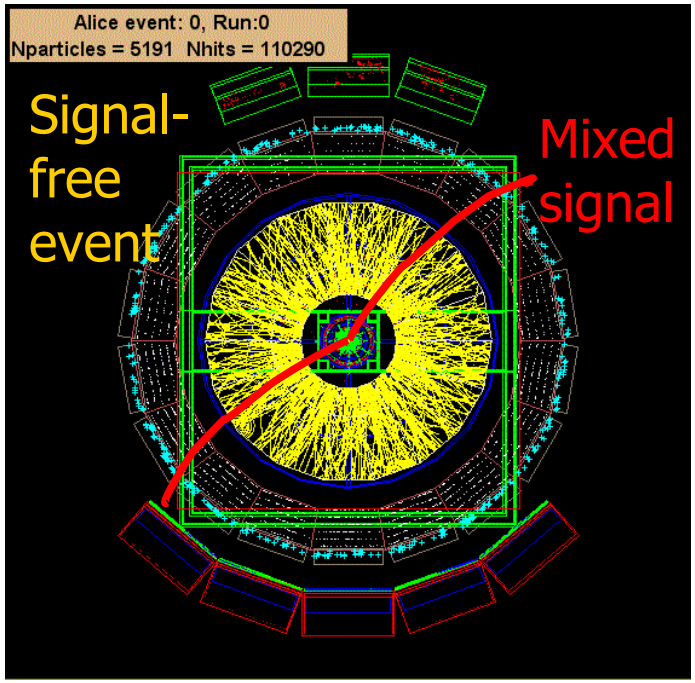


Summary of PDC'04

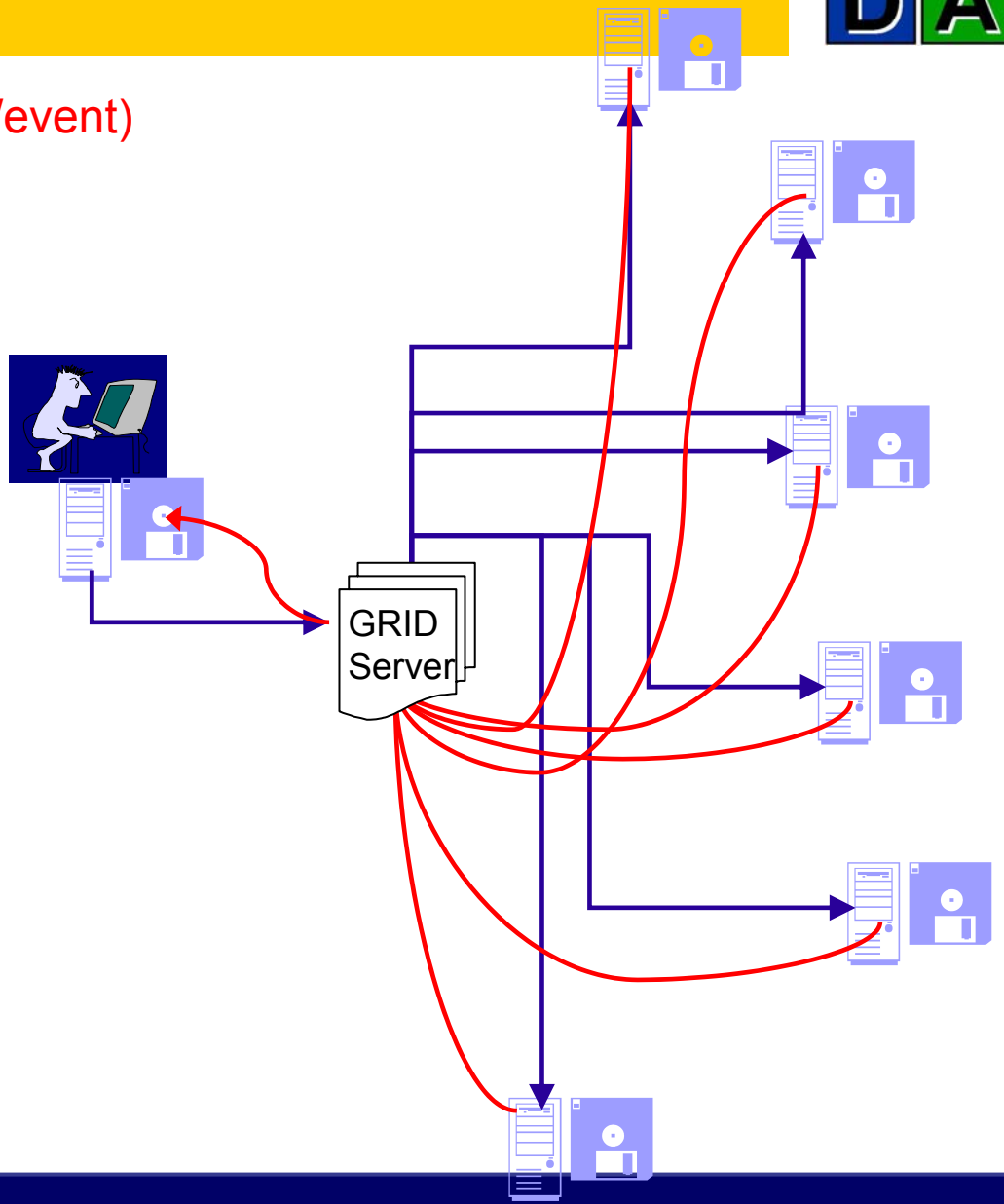
- Computing resources
  - It took some effort to 'tune' the resources at the remote computing centres
  - The centres' response was very positive – more CPU and storage capacity was made available during the PDC
- Middleware
  - AliEn proved to be fully capable of executing high-complexity jobs and controlling large amounts of resources
  - Functionality for Phase 3 has been demonstrated, but cannot be used
  - LCG MW proved adequate for Phase 1, but not for Phase 2 and in a competitive environment
- ALICE computing model validation:
  - AliRoot – all parts of the code successfully tested
  - Computing elements configuration
    - Need for a high-functionality MSS shown
    - Phase 2 distributed data storage schema proved robust and fast
  - Data Analysis could not be tested

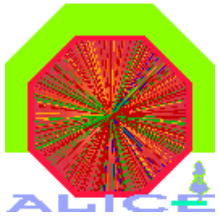
Phase III – (Interactive) Analysis

Large **distributed** input (100 MB/event)



Fairly small **merged** output

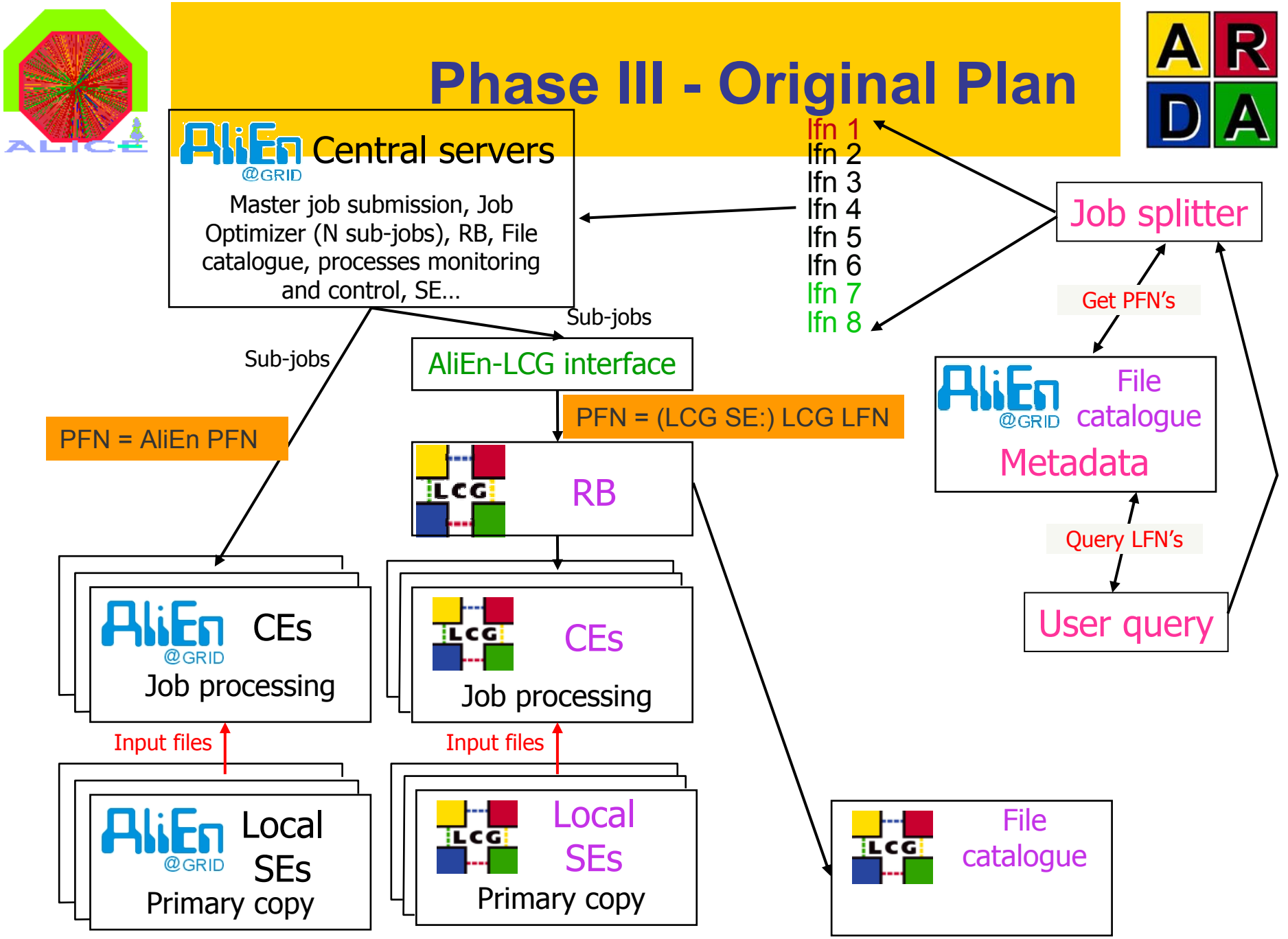


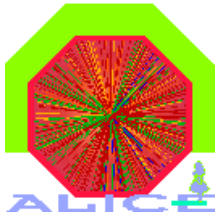


Phase III - Design strategy

- Distributed Input
  - 5.4 M files, on about 30 different (AliEn & LCG) SEs
- Do not move the input
  - Algorithm definition by a user on site S
  - On site S, Input Query to the Data Catalogue, based on selected MetaData
  - Input files typically from many SEs, so
    - Split them in N subgroups defined by files stored on a given SE
    - Split the task into N sub-tasks, to be run in parallel on the CEs associated to SEs containing a fraction of the input files
    - Run the N sub-tasks in parallel
    - Merge the output on the user's site S
- How?
  - From a ROOT shell on the user's site
  - As a set of (AliEn & LCG) subjobs
  - As a set of interactive dynamical PROOF slaves driven by a PROOF master

Phase III - Original Plan





Phase III – Steps to get started

- Input Data
  - They are already registered in the AliEn & LCG Data Catalogues and stored on AliEn & LCG Storage Elements
  - Access the AliEn & LCG Catalogues from gLite
  - Input Data is scattered on about 30 sites
  - … we will likely have to do some data movement on LCG

- Output Data
  - Must be made available at the User's site for merging and optional registration to the AliEn Data Catalogue

The ALICE Grid Proposal for Using Computational Grids

F.Carminati

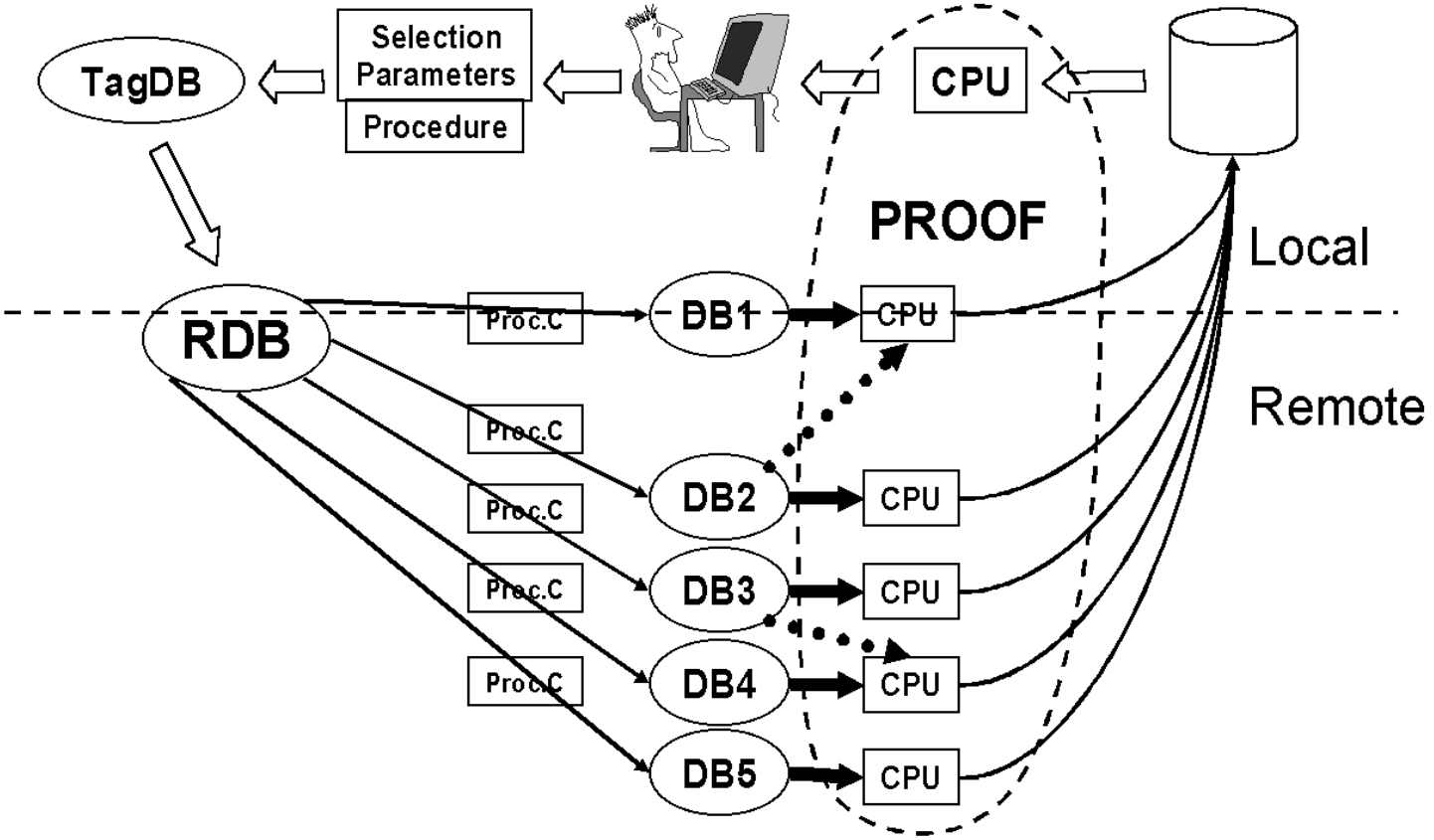
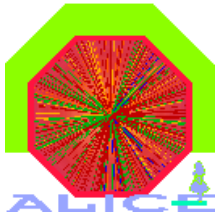


Figure 1: Schematic view of the analysis via PROOF facility



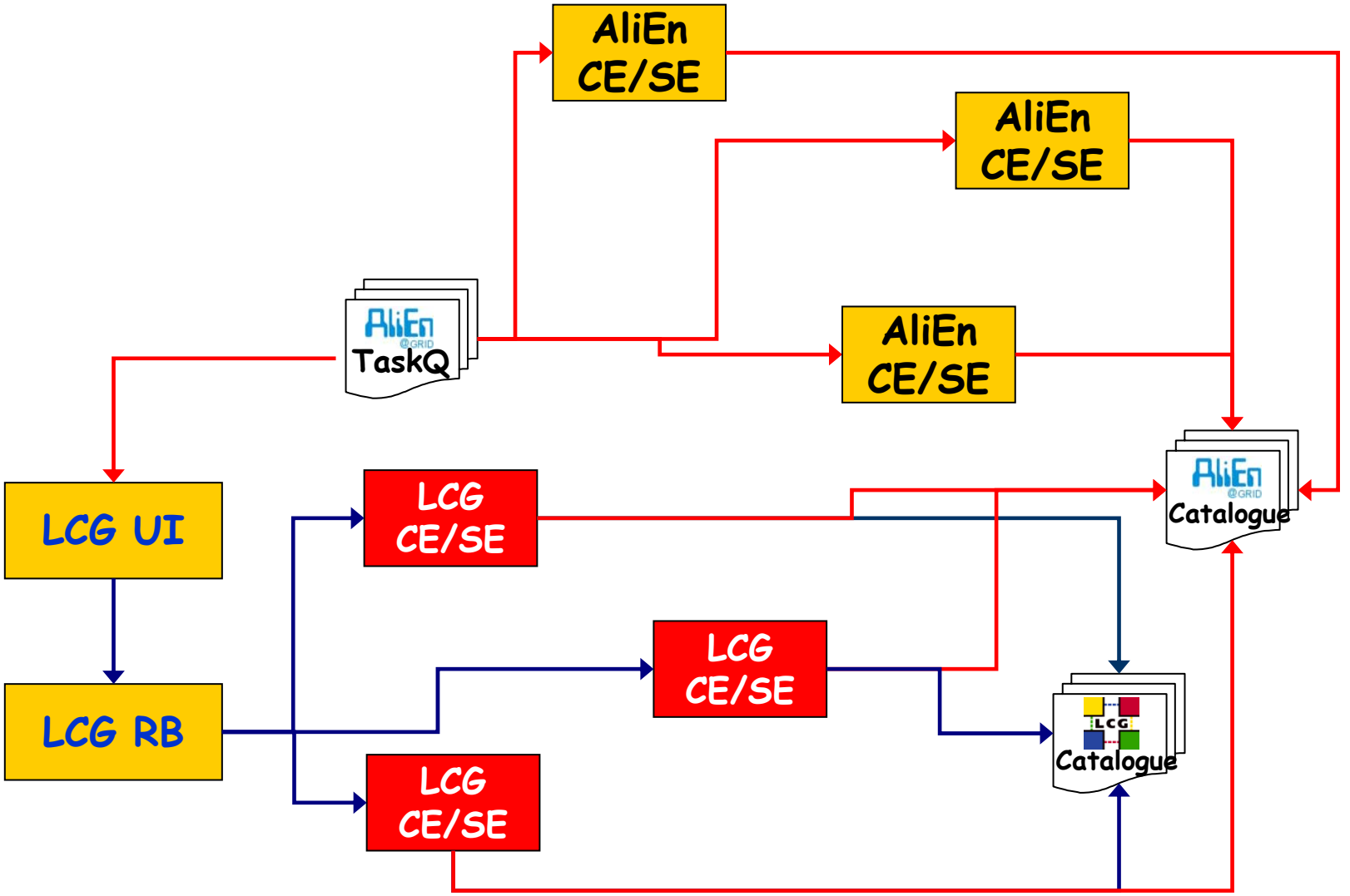
ALICE @ ARDA - Summary

- We view ARDA as a success
  - A working prototype of end-2-end interactive analysis based on gLite/AliEn components and PROOF was
    - **developed**
    - **deployed**
    - **demonstrated (SC2004)**
- But
  - Boundary conditions prevent its deployment as it is on the LCG production system
- So
  - We are developing an alternative strategy which aims at an equivalent functionality

Outline

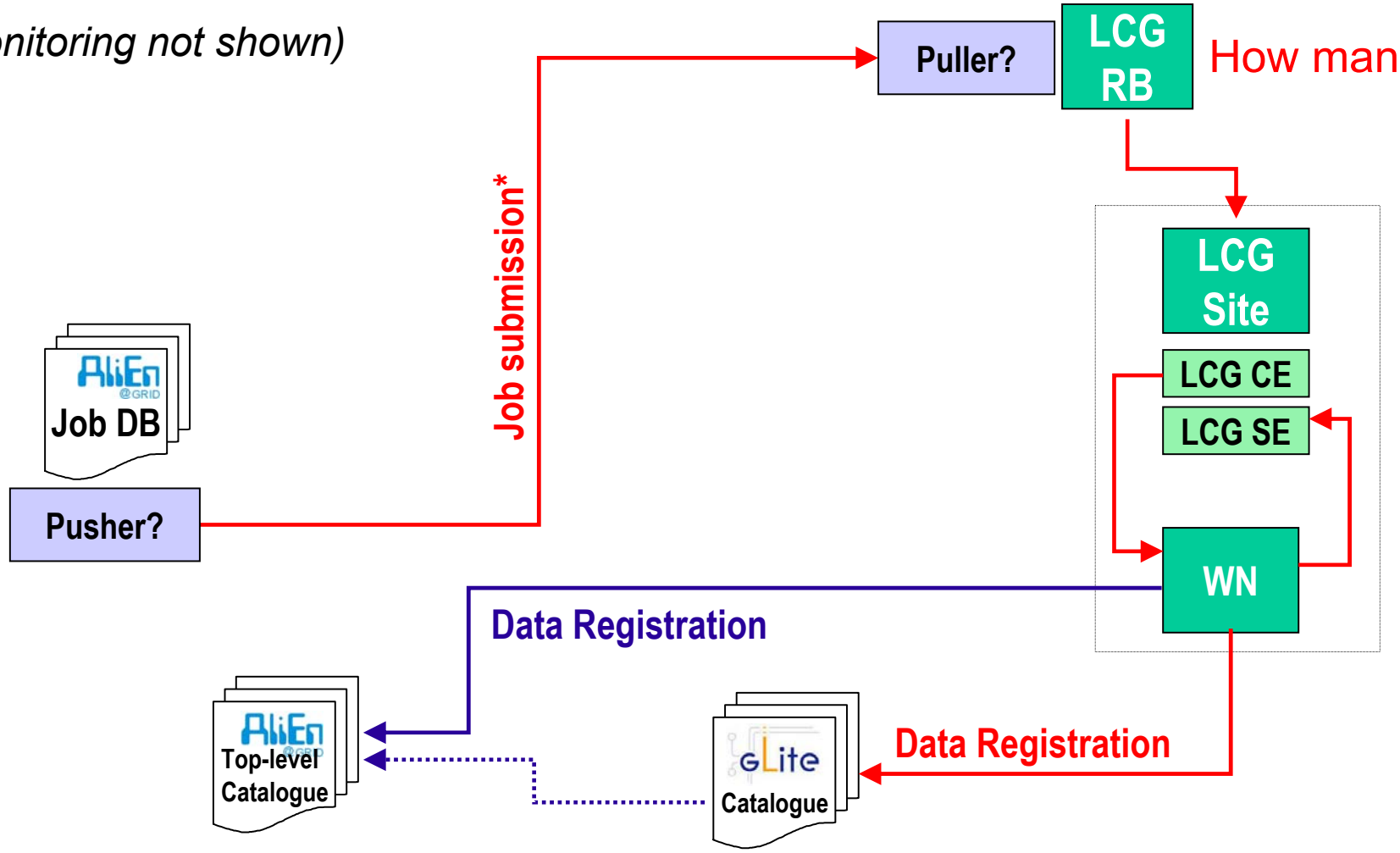
1. Where do we stand?
2. **Where do we go?**

Phase III layout: a "Meta-Grid"

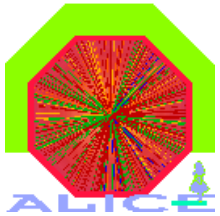


New layout

(Monitoring not shown)



* Single jobs can be part of a higher level job split at the AliEn level



Interactivity

It is still our goal

We will try to reach the same functionality of our present prototype using LCG

We are evaluating how to do that

We hope LCG evolution will be “user-driven”

more than the definition of its services has been until now