# *The ALICE Data Challenge*

F.Carminati

Grid Deployment Board

9 September 2004

# *Purpose and structure of PDC04*

- Test and validate the ALICE Offline computing model:
  - Produce and analyse ~10% of the data sample collected in a standard data-taking year
  - Use the entire ALICE off-line framework: AliEn, AliRoot, LCG, PROOF…
  - Experiment with Grid enabled distributed computing
  - Triple purpose: test of the middleware, the software and ***physics analysis*** of the produced data for the Alice PPR

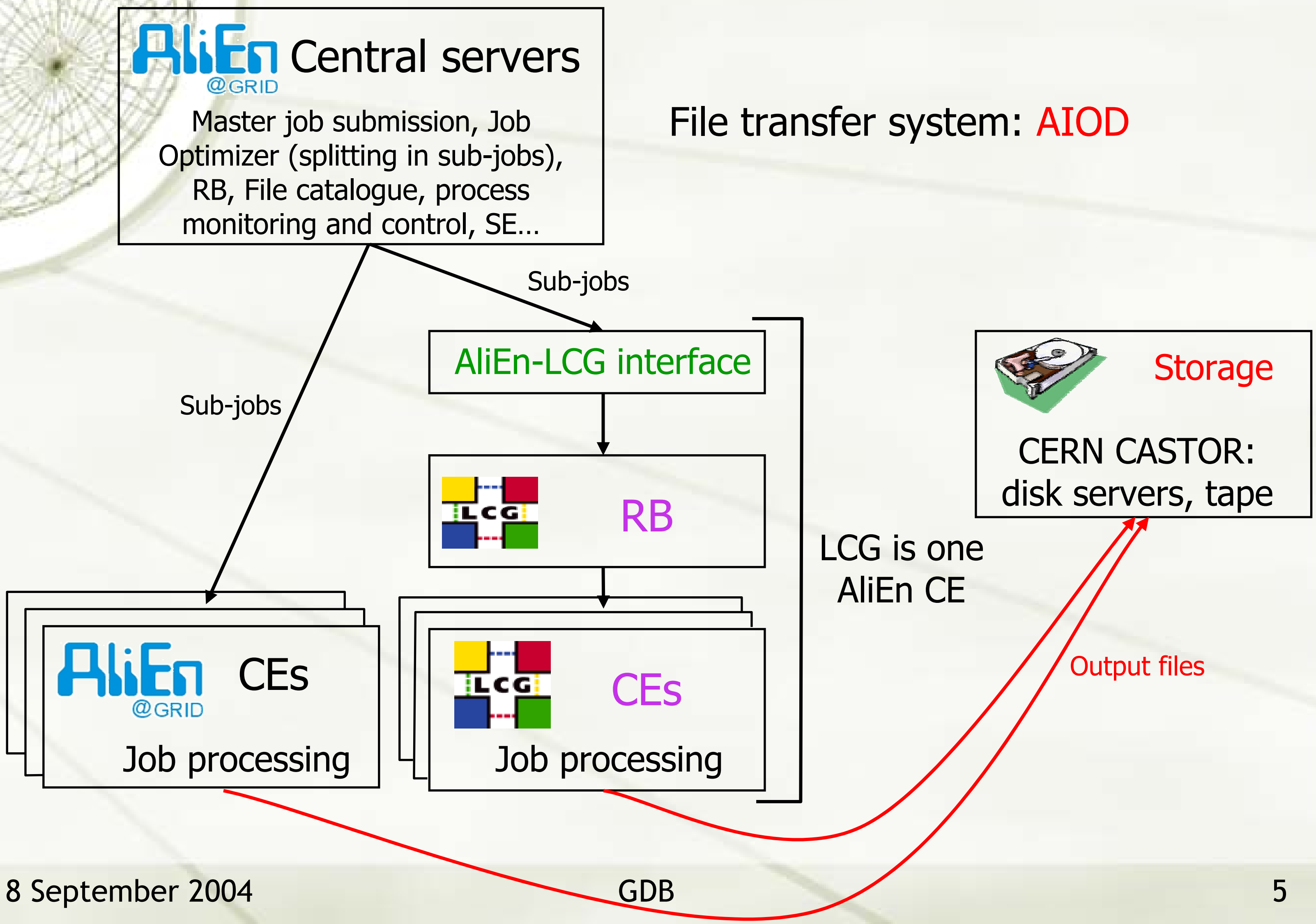
# *Purpose and structure of PDC04*

- Three phases
  - Phase I - Distributed production of underlying Pb+Pb events with different centralities (impact parameters) and of p+p events
  - Phase II - Distributed production mixing different signal events into the underlying Pb+Pb events (reused several times)
  - Phase III – Distributed analysis
- Principles:
  - True GRID data production and analysis: all jobs are run on the GRID, using only ***AliEn*** for access and control of native computing resources and, through an interface, the LCG resources
    - Use of *vanilla* LCG middleware (thanks to S.Bagnasco, P.Cerello INFN Torino)
  - In phase III ***GLite+ARDA***

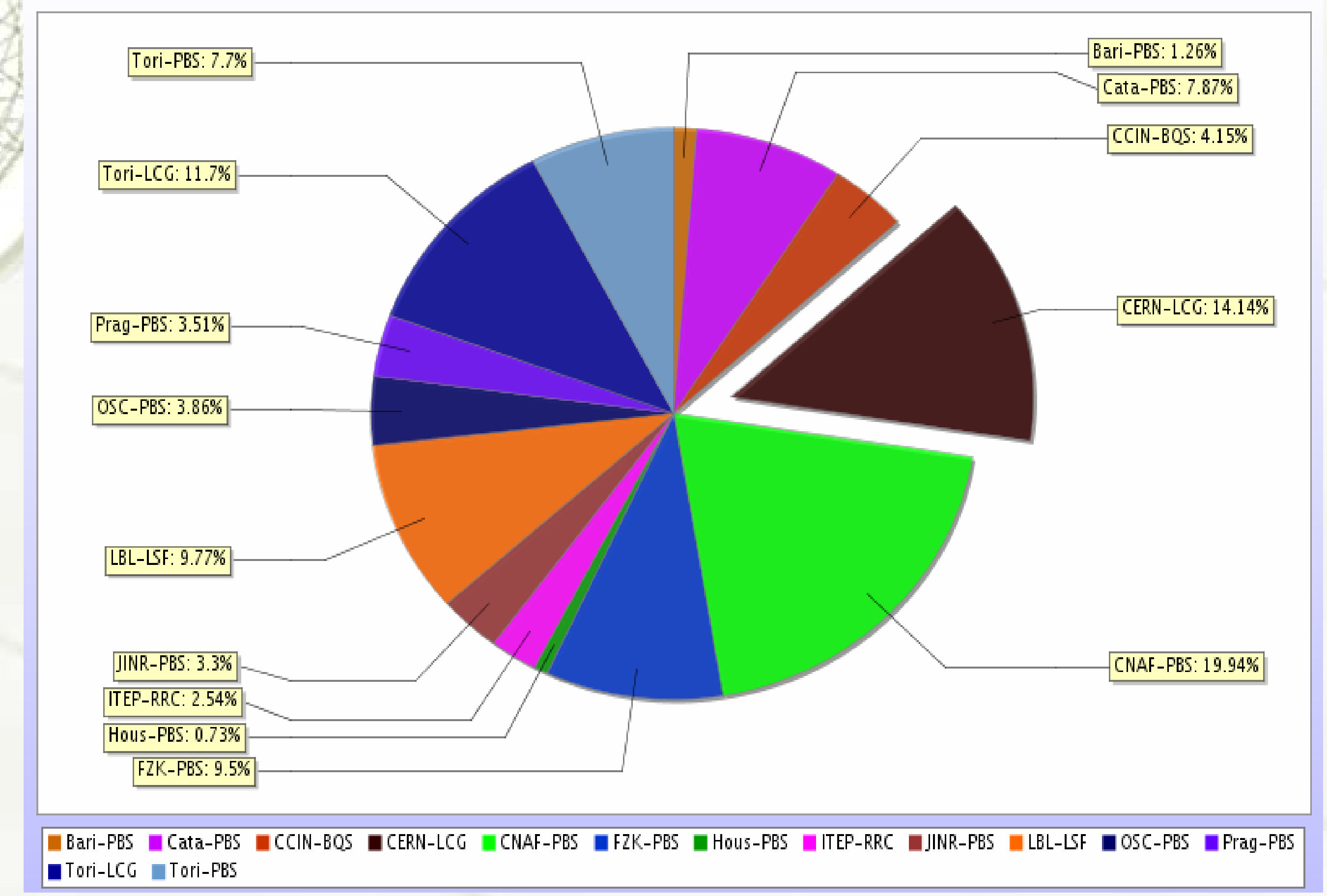
# *Phase I - statistics*

- Conditions and events
  - 5 centrality (impact parameter) conditions, 120k Pb+Pb events produced in 56k jobs
  - 8 hours (cent 1), 5 hours (per 1), 2.5 hours (per2-per5)
  - 3.8 million files in the AliEn file catalogue, 1.3 million in Castor@CERN (26 TB of data)
  - CPU: 285 MSI-2k hours (one 2.8 GHz PC working for 35 years)

# ✦ Job structure and production (phase I)

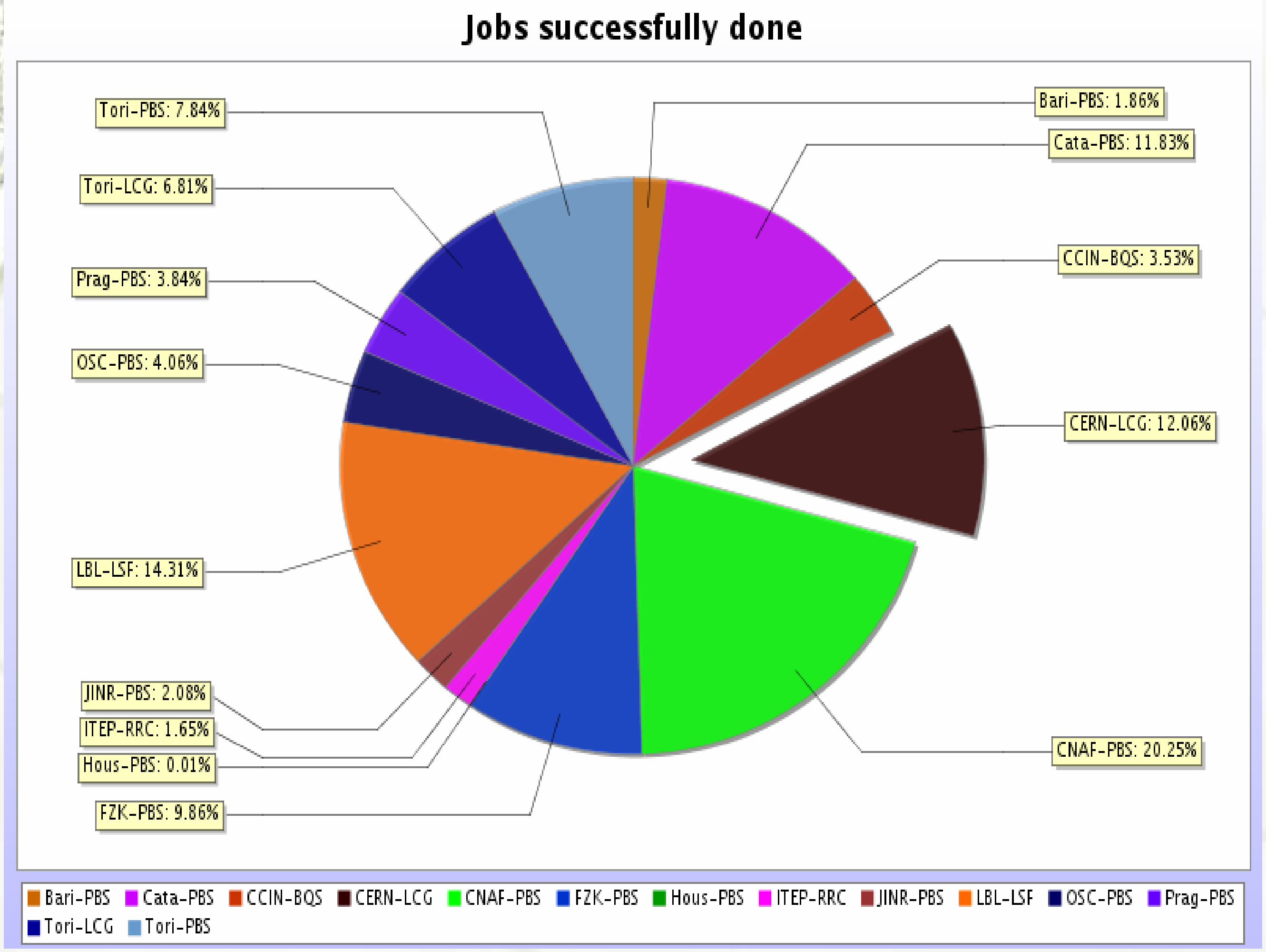


## ✦ Phase I CPU contributions




CEs: 15 directly controlled through AliEn + CERN-LCG and Torino-LCG (Grid.it)

## ✦ Phase I Completed jobs contributions




CERN-LCG and Torino-LCG (Grid.it): 19%

# *Issues*

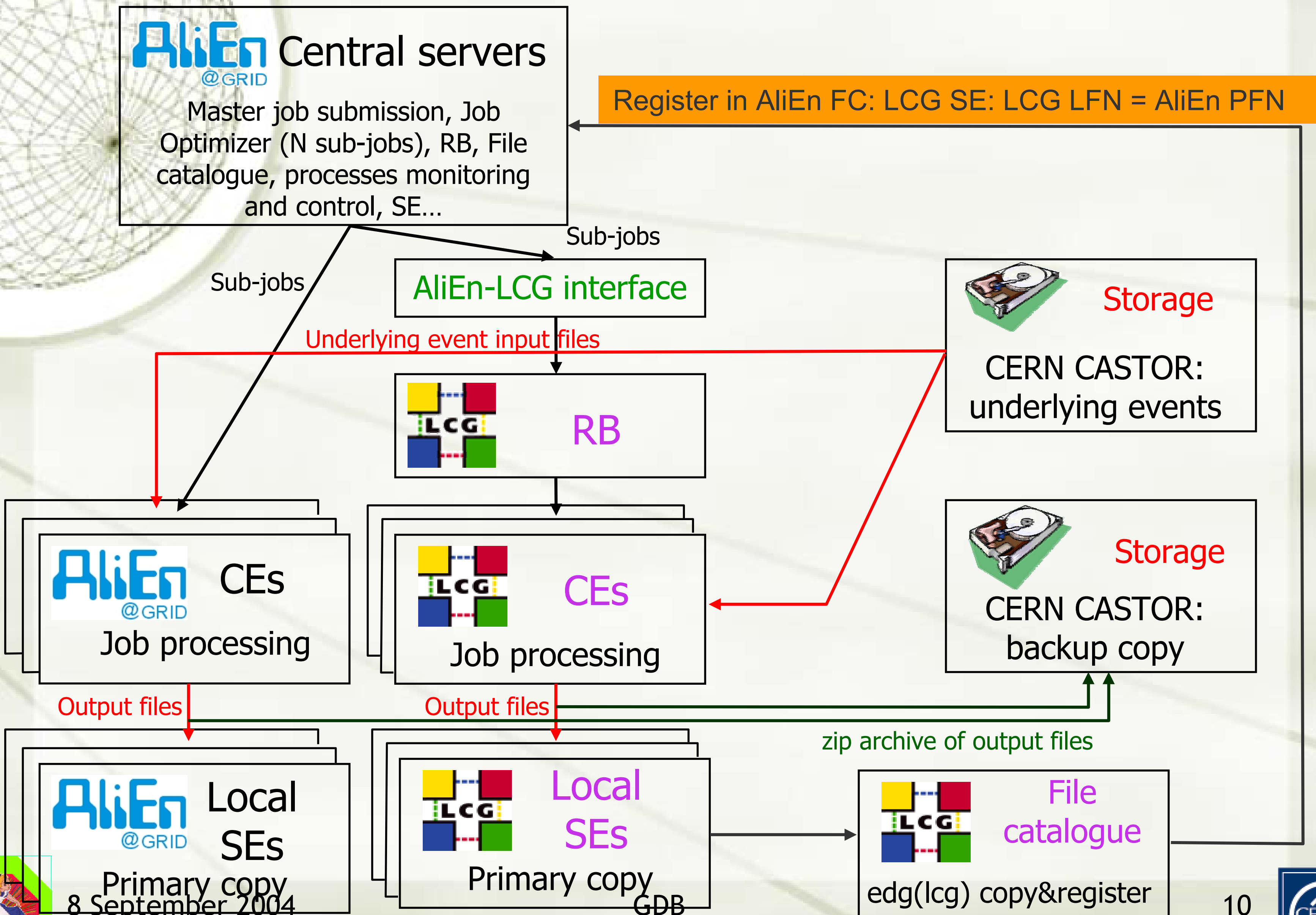
- CASTOR stager number of files limit (effective 250k)
  - Solved by splitting the data on two stagers
- Persistent problems with local configurations reducing the availability of LCG sites
  - Frequent *black holes*
  - Problems often come back (e.g. nfs mounts!)
  - Local disk space on WN
- Quality of the information in the II
- Workload Management System does not ensure an even distribution of jobs in the different centres
- Lack of support for *bulk operations* makes the WMS response time critical
- *KeyHole* approach and lack of appropriate monitoring and reporting tools make debugging difficult

# Phase II (started 1/07) – statistics

- In addition to phase I
  - Distributed production of signal events and merging with phase I events
  - Network and file transfer tools stress
  - Storage at remote SEs and stability (crucial for phase III)
- Conditions, jobs …:
  - 110 conditions total
  - 1 million jobs
  - 10 TB produced data
  - 200 TB transferred from CERN
  - 500 MSI2k hours CPU
- End by 30 September

| Signal | Signals / Underlying event | Underlying events | MB per signal event | kSI2Ks per signal event | TB | [MSI2K x h] | | |
|---|---|---|---|---|---|---|---|---|
| Jets | | cent1 | | | | | cycles: | 2 |
| Jets PT 20-24 GeV/c | 5 | 1666 | 5.2 | 940 | 0.09 | 4.35 | | |
| Jets PT 24-29 GeV/c | 5 | 1666 | 5.2 | 946 | 0.09 | 4.38 | | |
| Jets PT 29-35 GeV/c | 5 | 1666 | 5.3 | 952 | 0.09 | 4.41 | | |
| Jets PT 35-42 GeV/c | 5 | 1666 | 5.3 | 958 | 0.09 | 4.43 | | |
| Jets PT 42-50 GeV/c | 5 | 1666 | 5.4 | 964 | 0.09 | 4.46 | | |
| Jets PT 50-60 GeV/c | 5 | 1666 | 5.4 | 970 | 0.09 | 4.49 | | |
| Jets PT 60-72 GeV/c | 5 | 1666 | 5.5 | 976 | 0.09 | 4.52 | | |
| Jets PT 72-86 GeV/c | 5 | 1666 | 5.5 | 982 | 0.09 | 4.54 | | |
| Jets PT 86-104 Gev/c | 5 | 1666 | 5.6 | 988 | 0.09 | 4.57 | | |
| Jets PT 104-125 GeV/c | 5 | 1666 | 5.6 | 994 | 0.09 | 4.6 | | |
| Jets PT 125-150 GeV/c | 5 | 1666 | 5.7 | 1000 | 0.09 | 4.63 | | |
| Jets PT 150-180 GeV/c | 5 | 1666 | 5.7 | 1006 | 0.09 | 4.66 | | |
| Total signal | 199920 | | | | 1.08 | 54.04 | | |
| Jets with quenching | | cent1 | | | | | cycles: | 2 |
| Total signal | 199920 | | | | 1.08 | 54.04 | | |
| Jets | | per1 | | | | | cycles: | 2 |
| Jets PT 20-24 GeV/c | 5 | 1666 | 2.6 | 940 | 0.04 | 2.18 | | |
| Jets PT 24-29 GeV/c | 5 | 1666 | 2.6 | 946 | 0.04 | 2.19 | | |
| Jets PT 29-35 GeV/c | 5 | 1666 | 2.65 | 952 | 0.04 | 2.2 | | |
| Jets PT 35-42 GeV/c | 5 | 1666 | 2.65 | 958 | 0.04 | 2.22 | | |
| Jets PT 42-50 GeV/c | 5 | 1666 | 2.7 | 964 | 0.04 | 2.23 | | |
| Jets PT 50-60 GeV/c | 5 | 1666 | 2.7 | 970 | 0.04 | 2.24 | | |
| Jets PT 60-72 GeV/c | 5 | 1666 | 2.75 | 976 | 0.05 | 2.26 | | |
| Jets PT 72-86 GeV/c | 5 | 1666 | 2.75 | 982 | 0.05 | 2.27 | | |
| Jets PT 86-104 Gev/c | 5 | 1666 | 2.8 | 988 | 0.05 | 2.29 | | |
| Jets PT 104-125 GeV/c | 5 | 1666 | 2.8 | 994 | 0.05 | 2.3 | | |
| Jets PT 125-150 GeV/c | 5 | 1666 | 2.85 | 1000 | 0.05 | 2.31 | | |
| Jets PT 150-180 GeV/c | 5 | 1666 | 2.85 | 1006 | 0.05 | 2.33 | | |
| Total signal | 199920 | | | | 0.54 | 27.02 | | |
| Jets with quenching | | per1 | | | | | cycles: | 2 |
| Total signal | 199920 | | | | 0.54 | 27.02 | | |
| PHOS | | cent1 | | | | | cycles: | 1 |
| Jet-Jet PHOS | 1 | 20000 | 8.6 | 3130 | 0.17 | 17.39 | | |
| Gamma-jet PHOS | 1 | 20000 | 8.6 | 3130 | 0.17 | 17.39 | | |
| Total signal | 40000 | | | | 0.34 | 34.78 | | |
| D0 | | cent1 | | | | | cycles: | 1 |
| D0 | 5 | 20000 | 2.3 | 820 | 0.23 | 22.77 | | |
| Total signal | 100000 | | | | 0.23 | 22.77 | | |
| Charm & Beauty | | cent1 | | | | | cycles: | 1 |
| Charm (semi-e) + J/psi | 5 | 20000 | 2.3 | 820 | 0.23 | 22.78 | | |
| Beauty (semi-e) + Y | 5 | 20000 | 2.3 | 820 | 0.23 | 22.78 | | |
| Total signal | 200000 | | | | 0.46 | 45.56 | | |
| MUON | | cent1 | | | | | cycles: | 1 |
| Muon coctail cent1 | 100 | 20000 | 0.04 | 67 | 0.08 | 37.22 | | |
| Muon coctail HighPT | 100 | 20000 | 0.04 | 67 | 0.08 | 37.22 | | |
| Muon coctail single | 100 | 20000 | 0.04 | 67 | 0.08 | 37.22 | | |
| Total signal | 6000000 | | | | 0.24 | 111.66 | | |
| MUON | | per1 | | | | | cycles: | 1 |
| Muon coctail per1 | 100 | 20000 | 0.04 | 67 | 0.08 | 37.22 | | |
| Muon coctail HighPT | 100 | 20000 | 0.04 | 67 | 0.08 | 37.22 | | |
| Muon coctail single | 100 | 20000 | 0.04 | 67 | 0.08 | 37.22 | | |
| Total signal | 6000000 | | | | 0.24 | 111.66 | | |
| All signals | | | | | 4.75 | 488.55 | | |
| MUON | | per4 | | | | | cycles: | 1 |
| Muon coctail per4 | 5 | 20000 | | | | | | |
| Muon coctail single | 100 | 20000 | | | | | | |
| proton-proton | | no merging | | | | | cycles: | 1 |
| proton-proton | 100000 | | | | | | | |

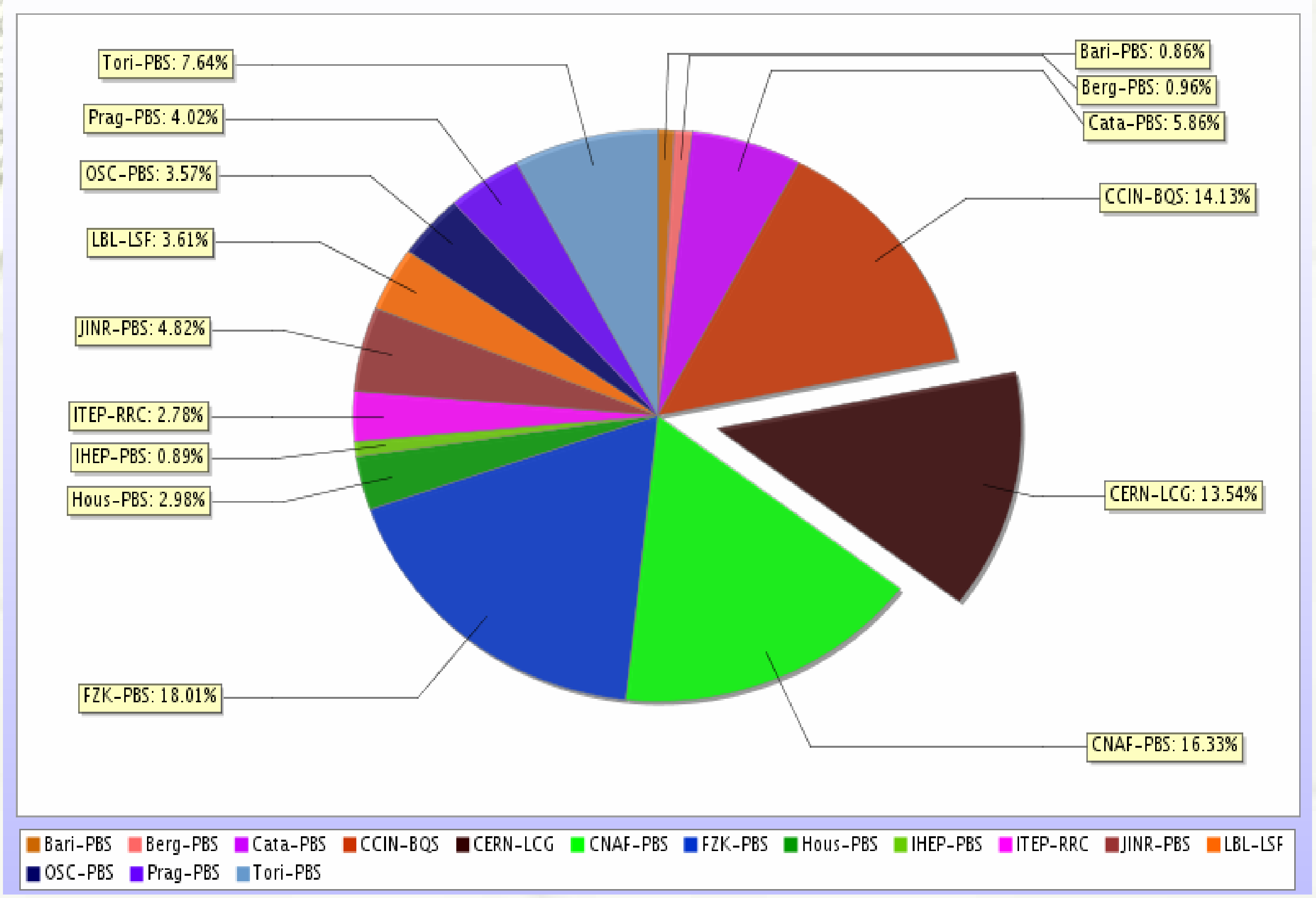
# Structure of event production in phase II



# *Improvements*

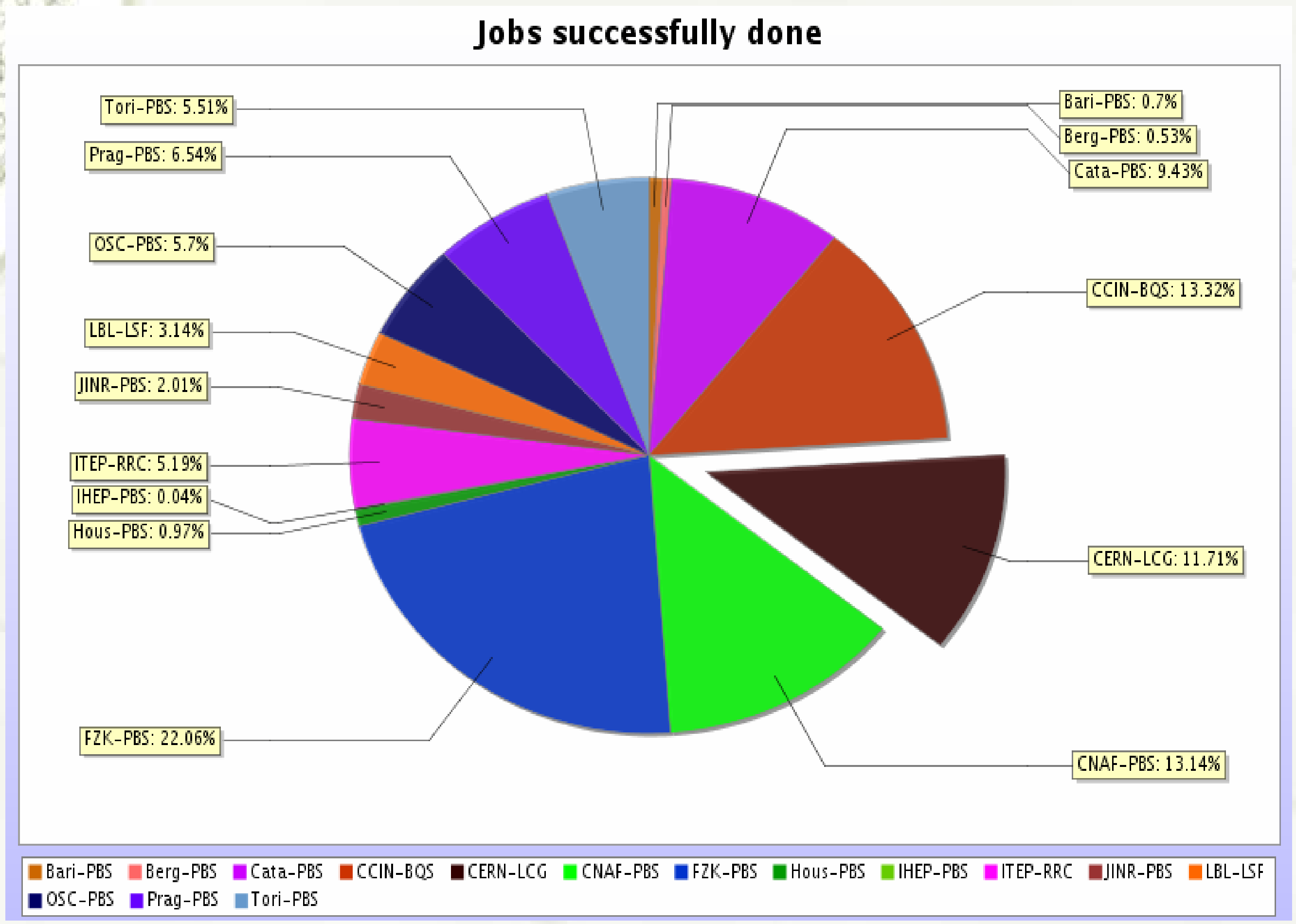
- Production system improvements
  - Implemented tar archiving, mainly to accommodate the limitations of the taping systems (less files, large size)
  - Fast resubmission of failed jobs – in this phase all jobs ***must*** finish
  - More sophisticated job monitoring tools, including singe job monitoring from start to finish (fast debugging)
    - Very fruitful collaboration with CMS on MonaLisa

# ✦ Phase II CPU contributions




CEs: 17 directly controlled through AliEn + CERN-LCG

# ✦ Phase II completed jobs contributions

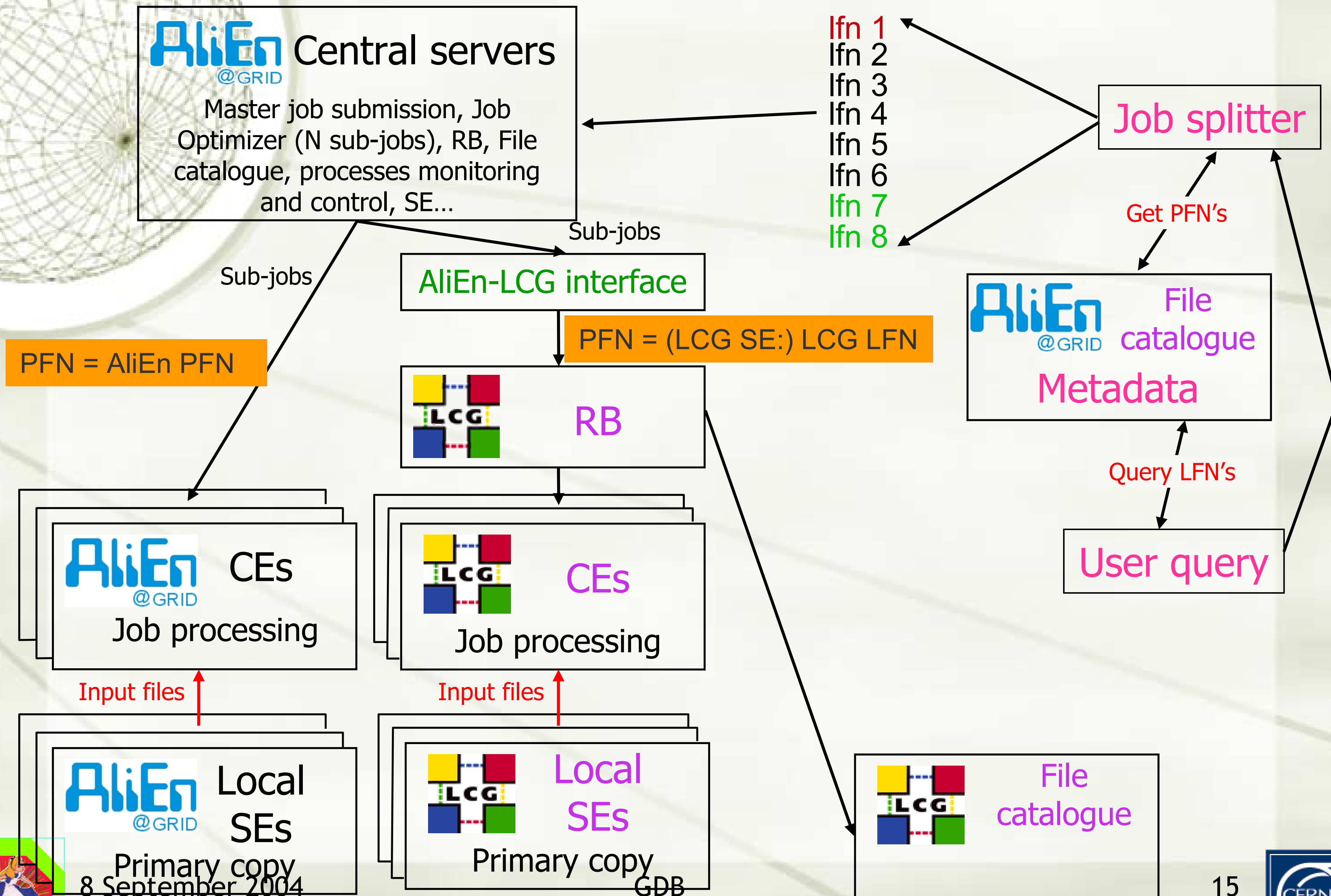



CEs: 17 directly controlled through AliEn + CERN-LCG

# *Issues*

- Same problems as before
  - Although actively followed by LCG support
- We *lived* through the transition from edg-* to lcg-* data management commands
  - It took us some time to learn how to use them
  - From time to time we still have batches of jobs failing to save due to local configuration and communication problems
    - Which have not been completely understood

## ✦ Structure analysis in phase 3



# *Phase III plans*

- We have a prototype working in AliEn
- We plan to use also LCG resources for interactive analysis
  - Which we understand is quite daring
- However we would like to do as much as possible of Phase III with gLite
- Phase III will go from October to the end of the year

# *Considerations*

- LCG is providing a (more and more) coherent infrastructure with a large quantity of resources
- Relations with LCG are good
  - With the inevitable strains of time-constrained exercise
  - Special thanks to P.Mendez for playing the difficult and ungrateful role of interface between the unruly ALICE mob and LCG support and site admins
  - Thanks to LCG for providing additional storage resources for a local SE at CERN exclusively for ALICE (3.5 TB).
- The middleware is improving within the limits of its current design and architecture
  - GAG is preparing a detailed feedback based on the experience of the Data Challenges
  - ALICE observations are similar to those of the other experiments

# *Conclusions*

- The ALICE DC04 started out with (almost unrealistically) ambitious objectives
- We are coming very close to reach this objectives and LCG has played an important role in this
- We are ready and willing to move to gLite as soon as possible and contribute to its evolution with our feedback