

ALICE Data Challenge 04



Federico Carminati
GDB 15 June, 2004



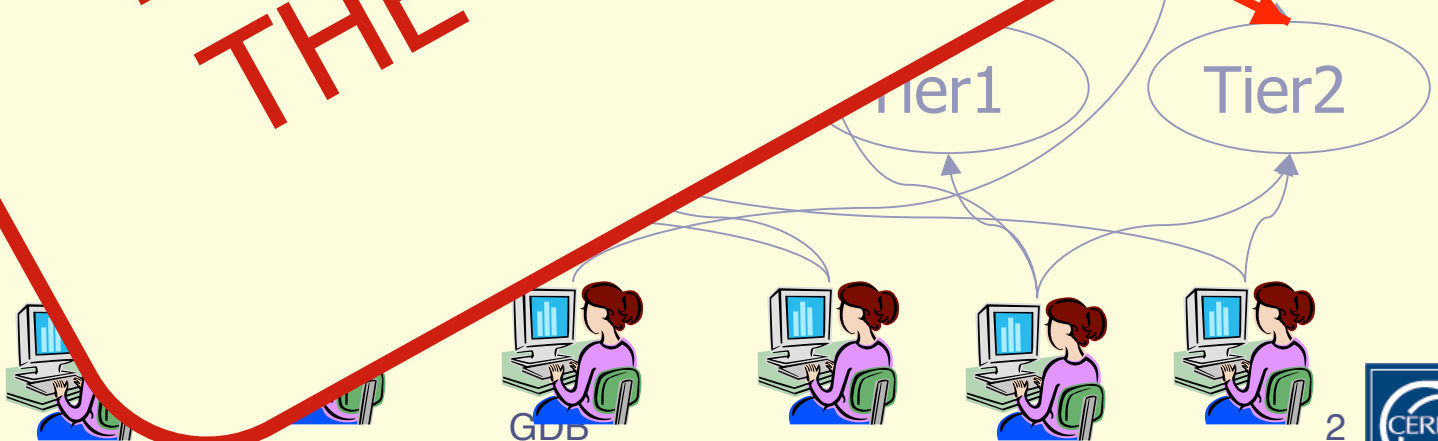


PDC 3 s

En job control
a transfer

Production of RAW
Shipment of RAW to
Reconstruction
Analysis

**DO IT ALL ON
THE GRID!!!!**



June 15, 2004

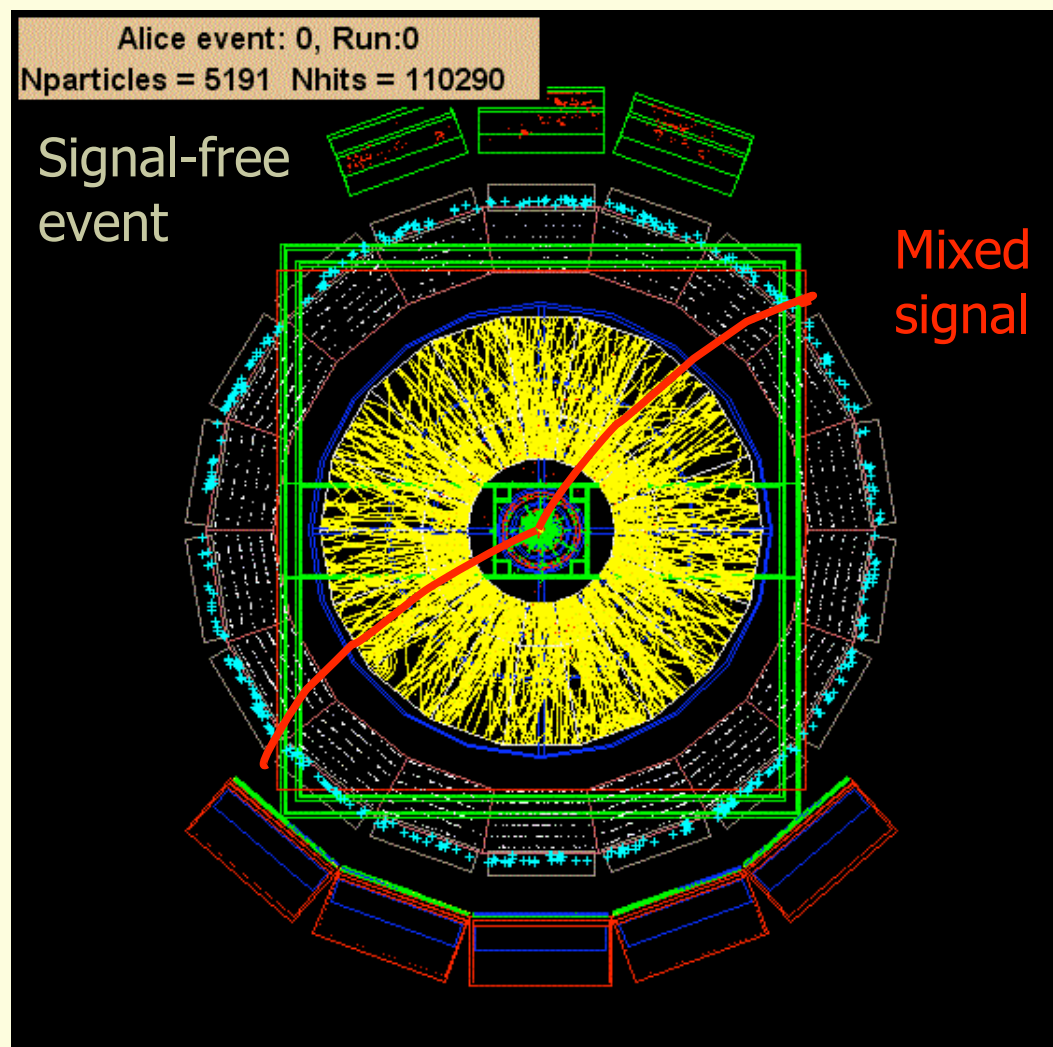
GDB

2





Merging



GDB



June 15, 2004

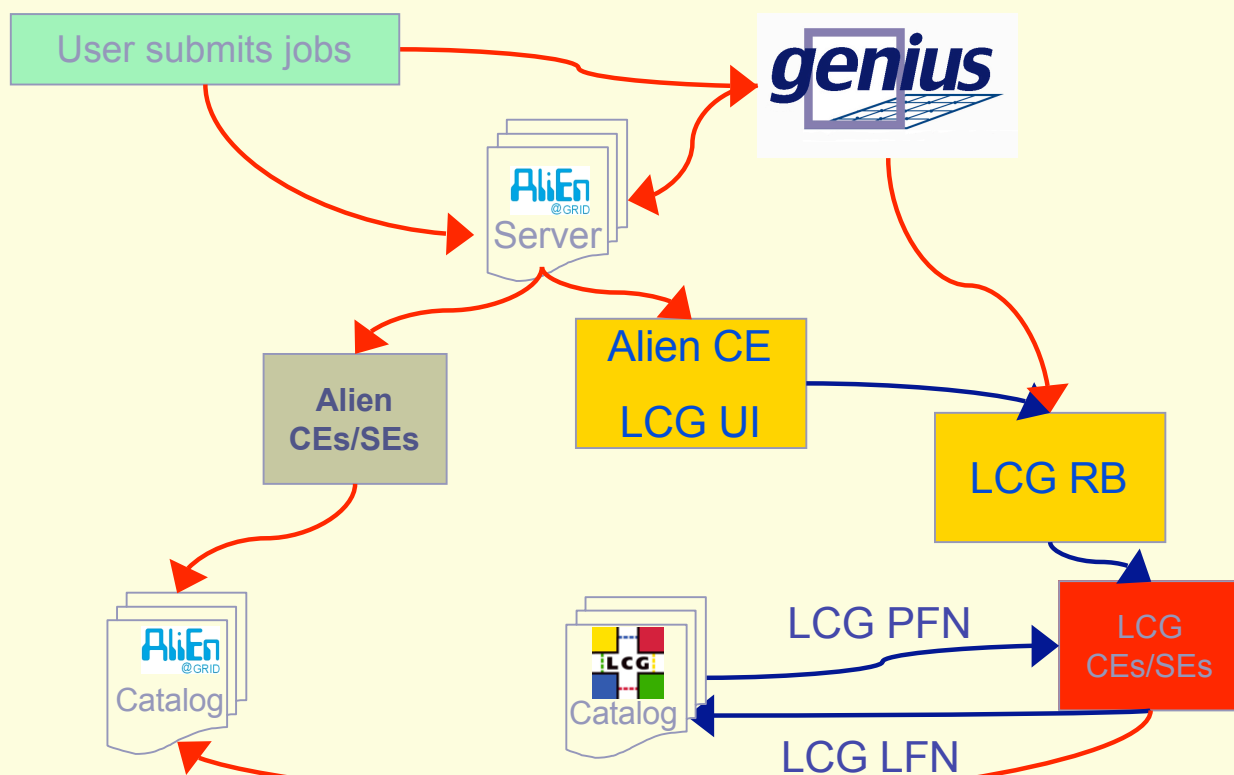
3





AliEn, Genius & EDG/LCG

- LCG-2 is one CE of AliEn, which integrates LCG and non LCG resources
 - If LCG-2 can run a large number of jobs, it will be used heavily
 - If LCG-2 cannot do that, AliEn selects other resources, and it will be less used



LCG LFN = AliEn PFN
GDB



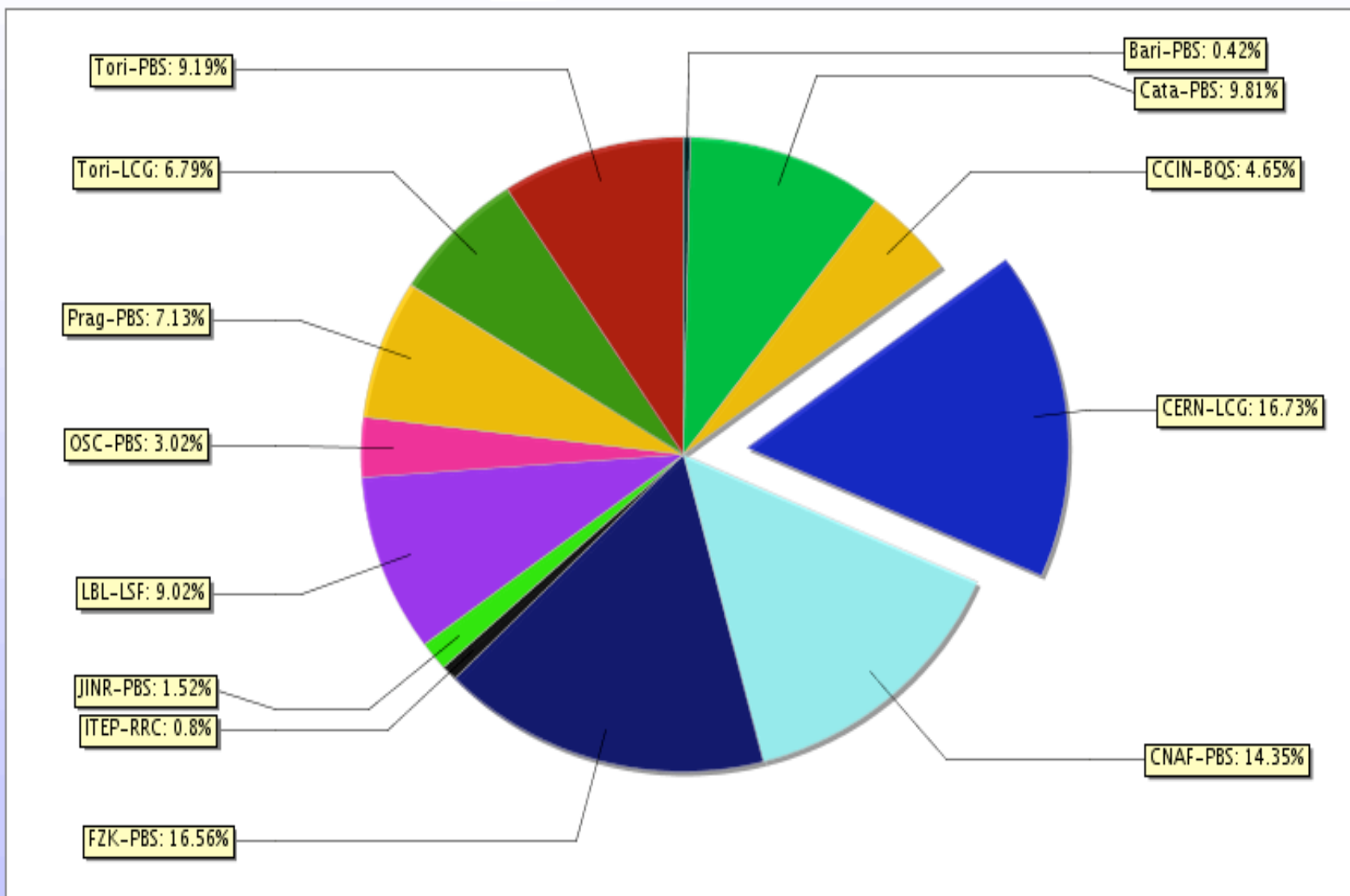
June 15, 2004

4





Jobs done



Legend for the pie chart:

- Bari-PBS
- Cata-PBS
- CCIN-BQS
- CERN-LCG
- CNAF-PBS
- FZK-PBS
- ITEP-RRC
- JINR-PBS
- LBL-LSF
- OSC-PBS
- Prag-PBS
- Tori-LCG
- Tori-PBS

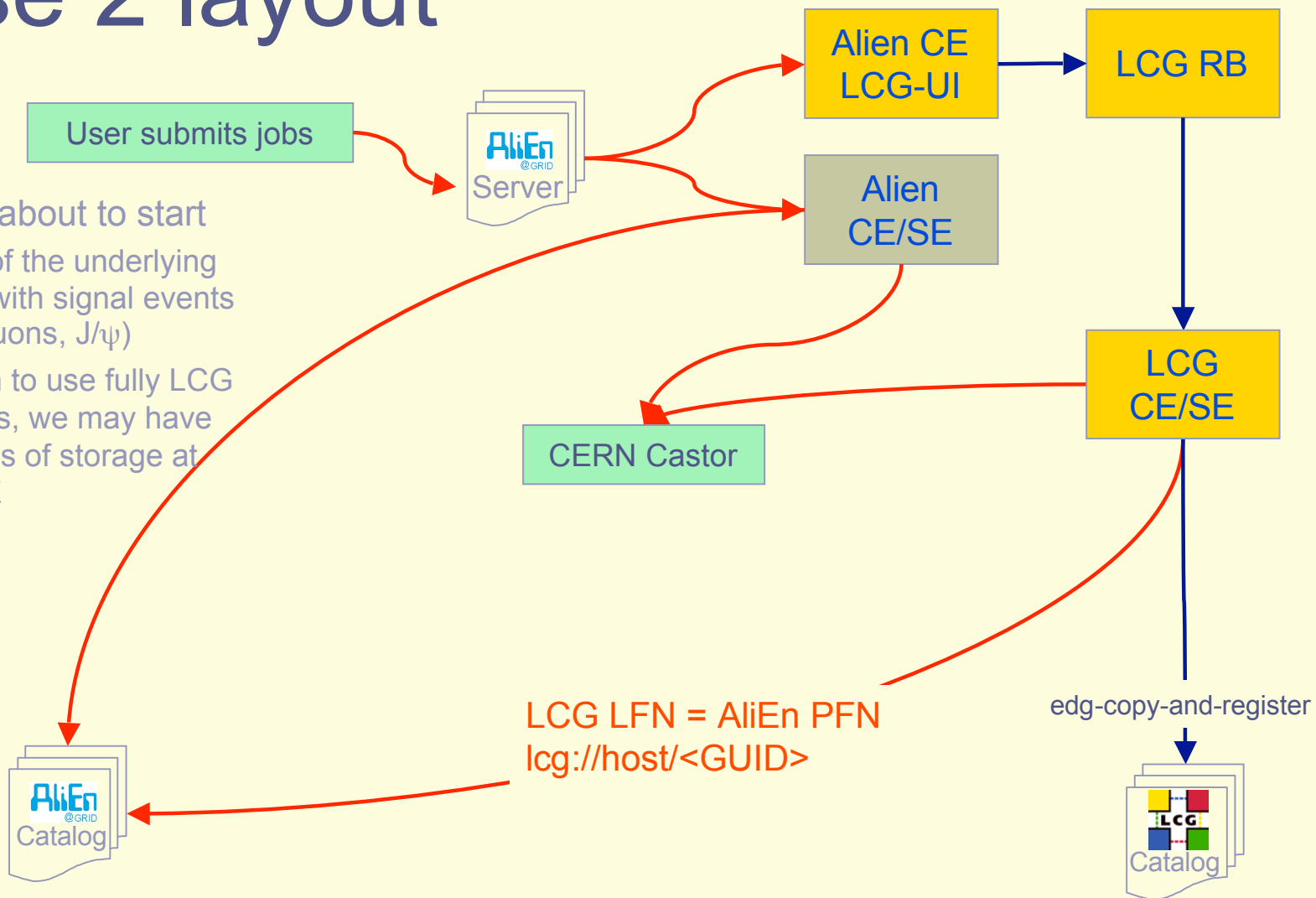




Phase 2 layout

Phase 2 -- about to start

- Mixing of the underlying events with signal events (jets, muons, J/ψ)
- We plan to use fully LCG DM tools, we may have problems of storage at local SE



June 15, 2004

GDB

6





Problems with Phase II

- Phase II will generate lots (1M) of (rather small ~7MB) files
- We would need an extra stager at CERN, but this is not available at the moment
- We could use some TB of disk space, but this too is not available
- We are testing a plug-in to AliEn using tar to bunch small files
- The space available on the local LCG storage elements seems very low... we will see
- Preparation of the LCG-2 JDL is more complicated, due to the use of the data management features
- This has introduced a two weeks delay -- we hope to start soon!



June 15, 2004

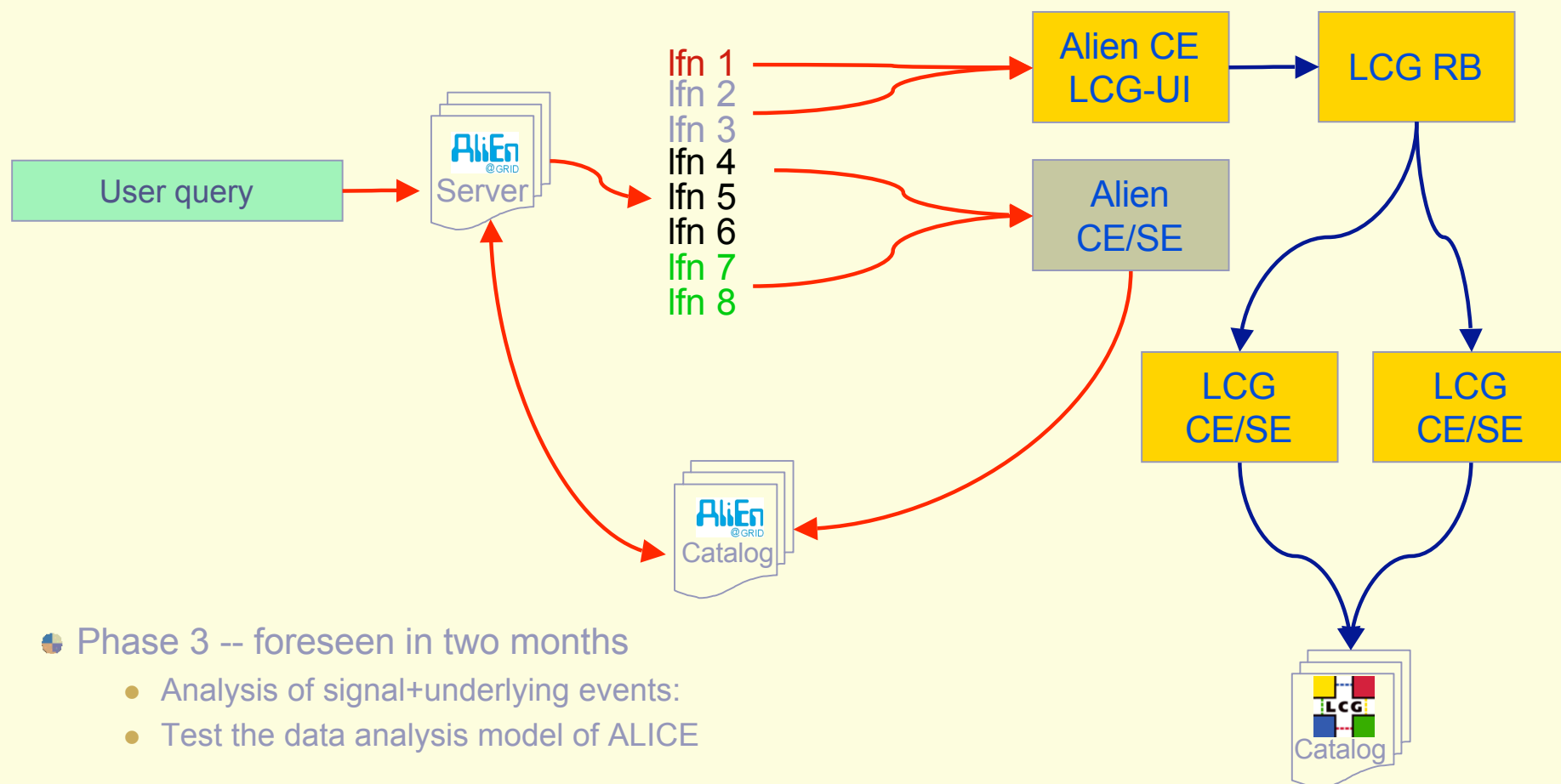
GDB

7





Phase 3 layout



- Phase 3 -- foreseen in two months
 - Analysis of signal+underlying events:
 - Test the data analysis model of ALICE



June 15, 2004

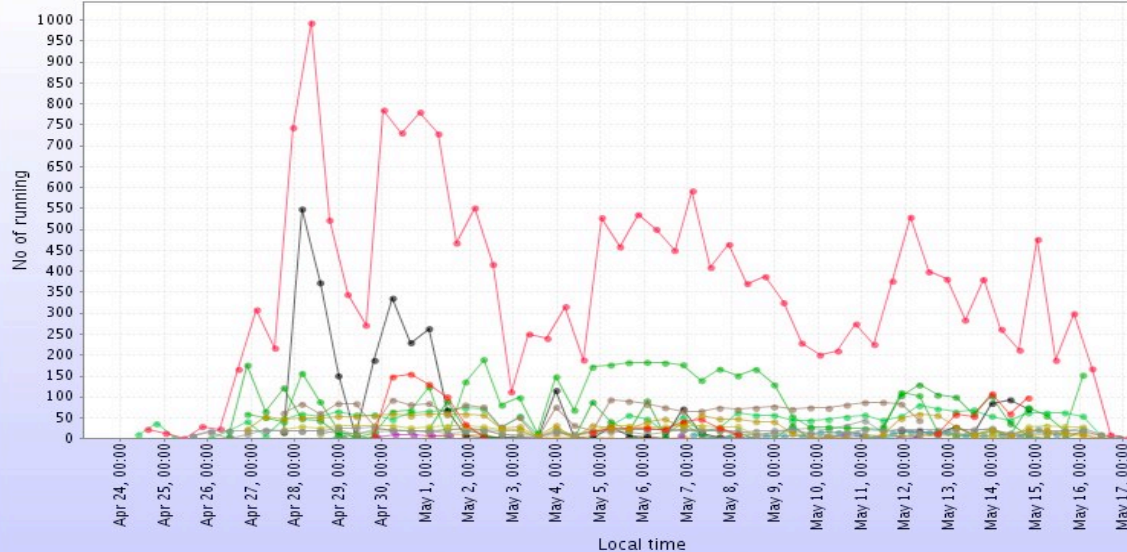
GDB

8





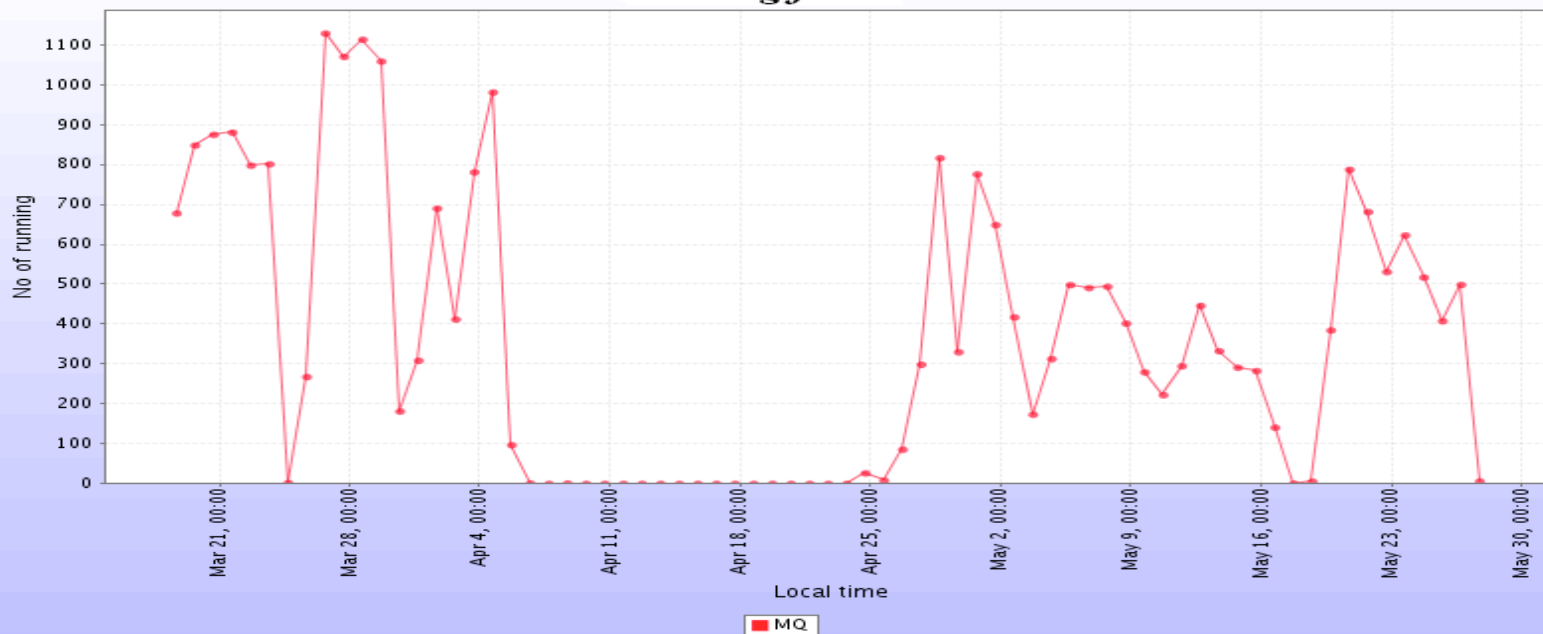
Running Processes



Farm	Last value	Min	Avg	Max
Bari-PBS	1	1	5.209	10
Cata-PBS	1	1	40.78	91
CCIN-BQS	2	1	17.06	52
CERN-LCG	72.1	1	84.59	693
CNAF-PBS	152.5	1	85.35	202
FZK-PBS	20.16	1	41.2	216
Hous-PBS	1	1	5.958	11
IFIC-PBS	1	1	1.429	2
IHEP-PBS	3.566	1	5.15	20
ITEP-RRC	10.54	1	9.879	34
JINR-PBS	8.861	2	14.9	25
LBL-LSF	43.85	1	68.55	101
MQ	6.243	4	265.8	1.107 K
OSC-PBS	28.43	1	24.28	34
Prag-PBS	1	1	16.71	36
PULV-SGE	1	1	1	1
SUBA-PBS	1	1	1	1
Tori-LCG	97.81	1	54.89	180
Tori-PBS	2	1	32.58	63
Total	455.1	23	776.3	2.828 K

■ Bari-PBS ■ Cata-PBS ■ CCIN-BQS ■ CERN-LCG ■ CNAF-PBS ■ FZK-PBS ■ Hous-PBS ■ IFIC-PBS ■ IHEP-PBS ■ ITEP-RRC ■ JINR-PBS ■ LBL-LSF
■ MQ ■ OSC-PBS ■ Prag-PBS ■ PULV-SGE ■ SUBA-PBS ■ Tori-LCG ■ Tori-PBS

Running jobs



■ MQ





ALICE Physics Data Challenges

Period (<u>milestone</u>)	Fraction of the final capacity (%)	Physics Objective
<u>06/01-12/01</u>	1%	pp studies, reconstruction of TPC and ITS
<u>06/02-12/02</u>	5%	<ul style="list-style-type: none">• First test of the complete chain from simulation to reconstruction for the PPR• Simple analysis tools• Digits in ROOT format
<u>01/04-06/04</u>	10%	<ul style="list-style-type: none">• Complete chain used for trigger studies• Prototype of the analysis tools• Comparison with parameterised MonteCarlo• Simulated raw data
NEW <u>05/05-07/05</u>	TBD	<ul style="list-style-type: none">• Refinement of jet studies• Test of new infrastructure and MW• TBD NEW
<u>01/06-06/06</u>	20%	<ul style="list-style-type: none">• Test of the final system for reconstruction and analysis



June 15, 2004

GDB

10





Considerations

- AliEn
 - Tools OK for DC running and resources control
 - Feedback from the CE and WN proved to be essential for early spotting of problems
 - Centralized and compact master services allow for fast upgrades
 - DM was working just fine (providing that underlying MSS systems work well)
 - File catalogue works great, 4M entries and no noticeable performance degradation
- LCG-2 provides many cycles, despite being the first to use it for production
 - But required continuous efforts and interventions (ALICE and LCG)
 - Some instabilities came from the LCG-RB and/or its local configurations
 - The LCG-SE is still very “fluid”, so we may expect instabilities
 - Analysis will be run with ARDA
 - LCG needs to be strongly “prompted” for resources
- MonALISA is valuable for monitoring, GridICE is more opaque
- There is difference between pledged and available resources!
- AliEn as meta-grid works well, across three grids, and this is a success in itself
 - “keyhole” approach, some things become awkward (e.g. monitoring!)

