CERN **IT** Department

ALICE Environment on the GRID

A. Abramyan, S. Bagansco, S. Banerjee,
L. Betev, F. Carminati, D. Goyal, A. Grigoras,
C. Grigoras, M. Litmaath, N. Manukyan,
M. Martinez, A. Montiel, J. Porter, P. Saiz,
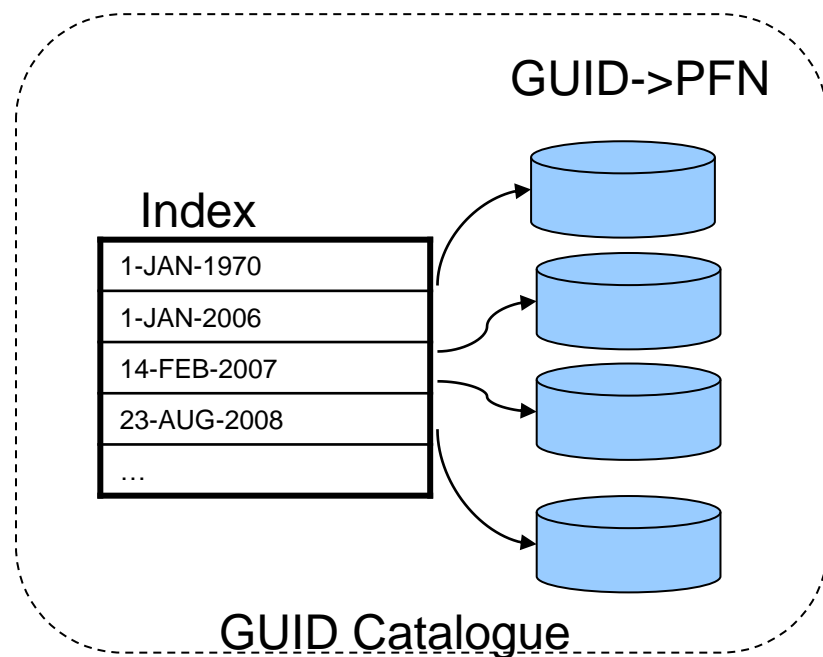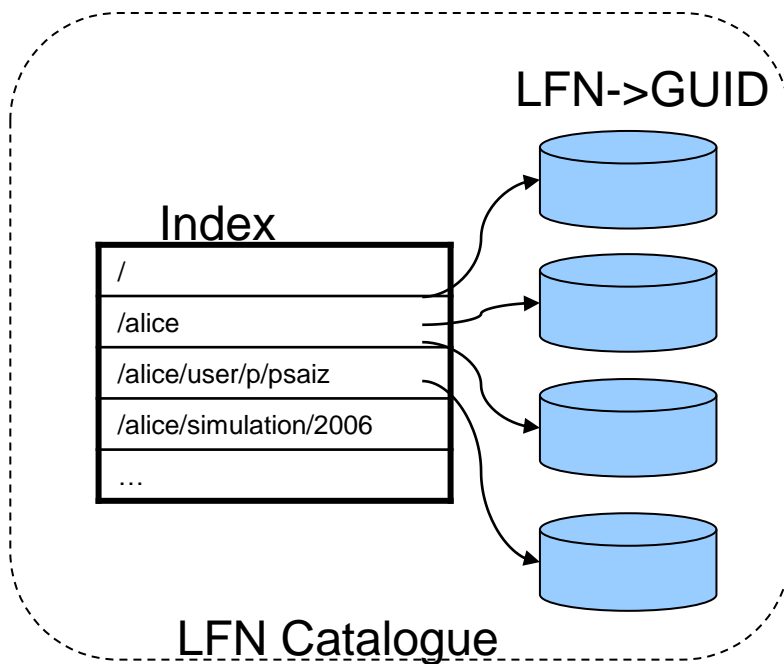S. Sankar, S. Schreiner, J. Zhu

Outline

- AliEn
  - File Catalogue
  - TaskQueue
  - Data access
- AliEn in ALICE
- New development
- Conclusions

- All components to create a GRID
- File Catalogue
  - UNIX-like file system
  - Mapping to physical files
  - Metadata information
  - SE discovery
- Transfer Model
  - With different plugins
- TaskQueue
  - Job Agent & pull model
  - Automatic installation of software packages
  - Simulation, reconstruction, analysis...
- Developed by ALICE
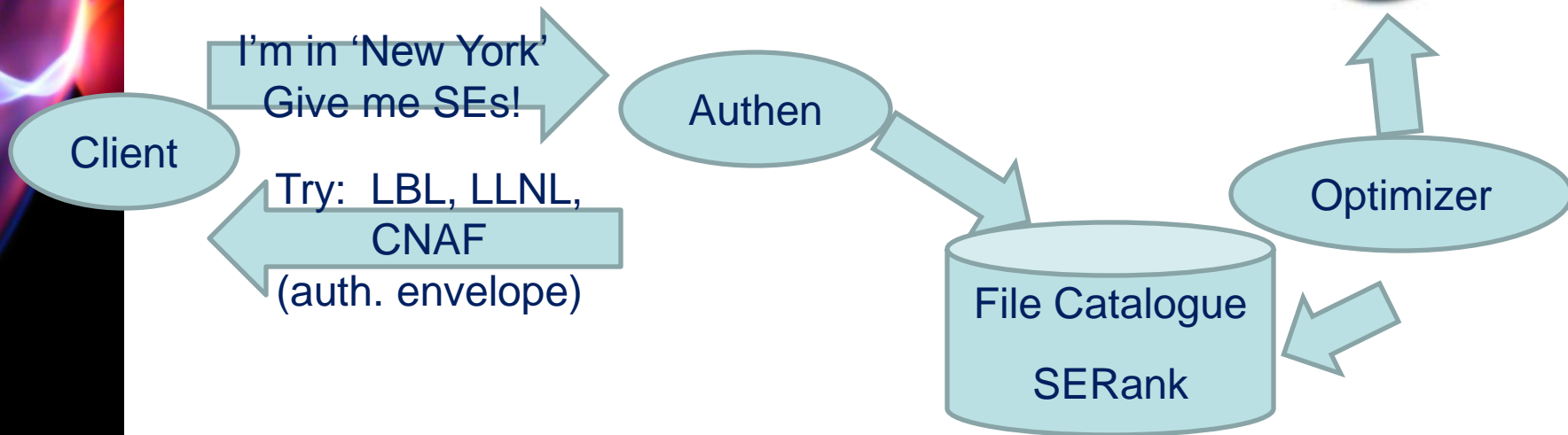 - Used by several communities
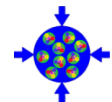
- Global Unique name space
  - Mapping from LFN to PFN
- UNIX-like file system interface
- Powerful metadata catalogue
- Automatic SE selection
- Integrated quota system
- Multiple storage protocols: xrootd, torrent, srm, file
- Collections of files
- Physical file archival
- Roles and users

- Entries can be retrieved by LFN or GUID
- Flexible table structure, allowing for easy scalability
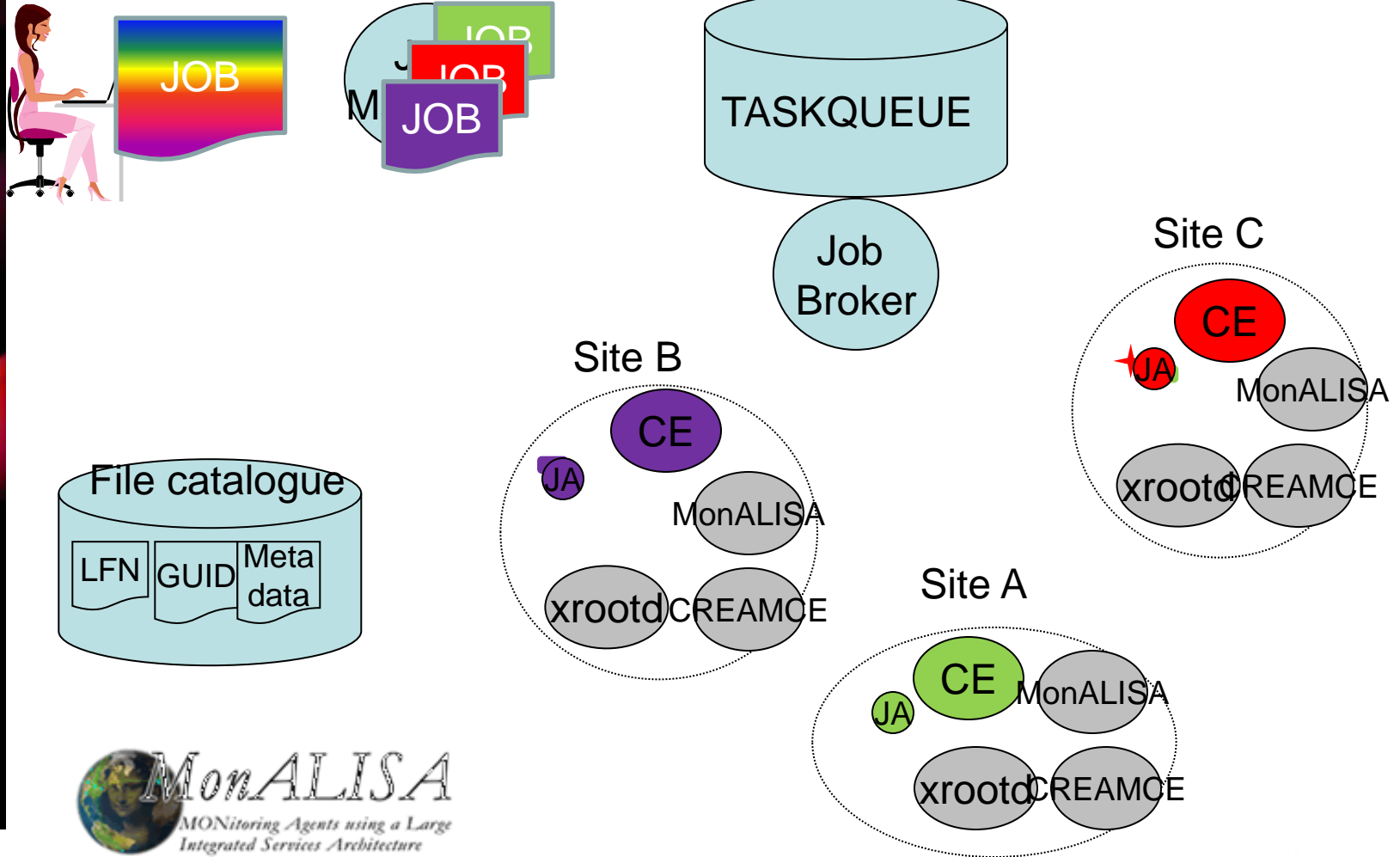- ALICE: 350M entries (270GB DB size)

LFN->GUID

Index

| / |
| /alice |
| /alice/user/p/psaiz |
| /alice/simulation/2006 |
| ... |

LFN Catalogue

GUID->PFN

Index

| 1-JAN-1970 |
| 1-JAN-2006 |
| 14-FEB-2007 |
| 23-AUG-2008 |
| ... |

GUID Catalogue

- Mostly xrootd protocol

- Jobs only download configuration files

- Data files are accessed remotely

  – The closest replica to the job, local replica first

- JobAgent uploads N replicas of the output

- Schedules data transfers only involve the raw data

  – Done via xrd3cp

- Use of torrent for AliEn and software package distribution

CERN**IT** Department

- Client automatically directed to the best SE to store/retrieve files
  - Find the closest working SE of a given QoS
- Built-in retrial mechanism

Client

I'm in 'New York' Give me SEs!

Try: LBL, LLNL, CNAF
(auth. envelope)

Authen

File Catalogue

SERank

Optimizer

ES

- Central TaskQueue with all jobs
  - Priorities & quotas on user/role level

CERN IT Department

JOB

JOB

JOB

TASKQUEUE

Job Broker

Site C

CE

JA

MonALISA

xrootd CREAMCE

Site B

CE

JA

MonALISA

xrootd CREAMCE

File catalogue

LFN GUID Meta data

Site A

CE

JA

MonALISA

xrootd CREAMCE

MonALISA
MONitoring Agents using a Large
Integrated Services Architecture

AliEn user interface			
		ARC/	
CONDOR	AliEn		LCG/EGEE/gLITE
		NORDUGRID	

Nice! I STILL do not have to worry about ever changing GRID environment…

More than 80 centres
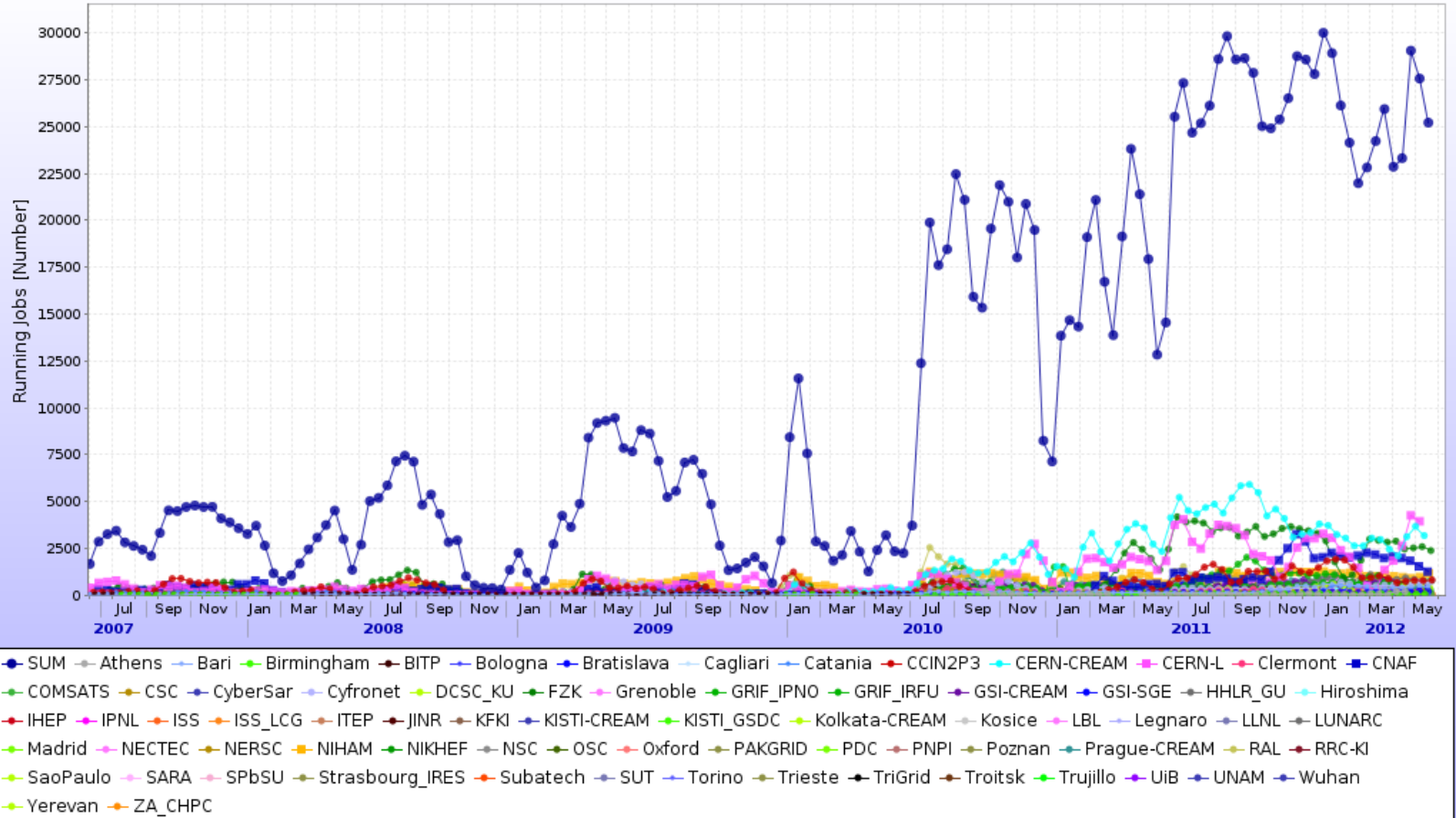
CERN **IT** Department

Running Jobs

ES — Distribution of successful jobs

Done jobs statistics
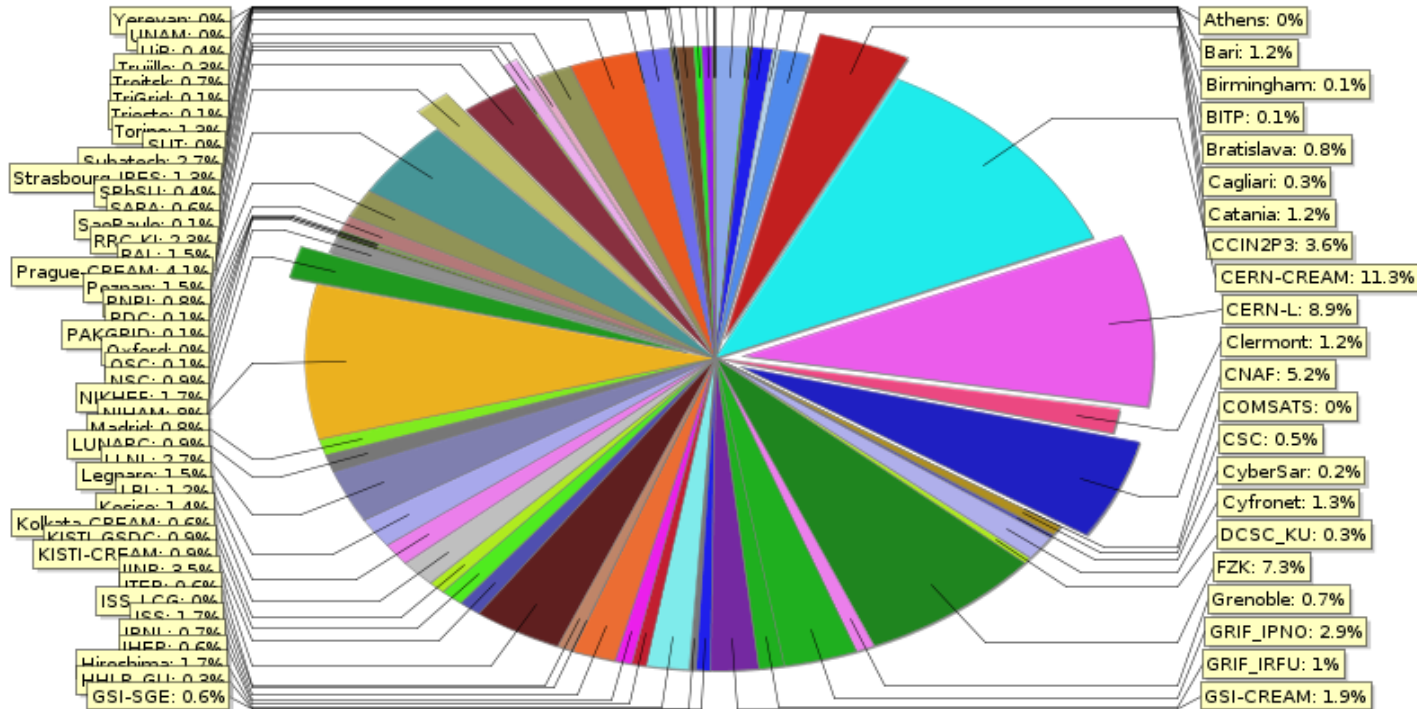
CH-1211 Geneva 23 Switzerland
www.cern.ch/it

21 May 2012 — *Pablo Saiz*

13

- ## Several major AliEn iterations
  - Starting in 2001
  - Scale testing
  - Evaluating different approaches
    - And learning from the exercises
- ## Keeping (almost) the same CLI
  - Even if backward incompatible under the hood

Scotland,
VO to VO

Germany,
ORACLE

Armenia,
XML model
PackMan

China,
Trust Model

USA,
JA memory

Switzerland,
Main dev.

Italy,
CREAMCE

India,
File deletion

South Korea,
Quota system

Chile,
Trust model

CERN
IT
Department
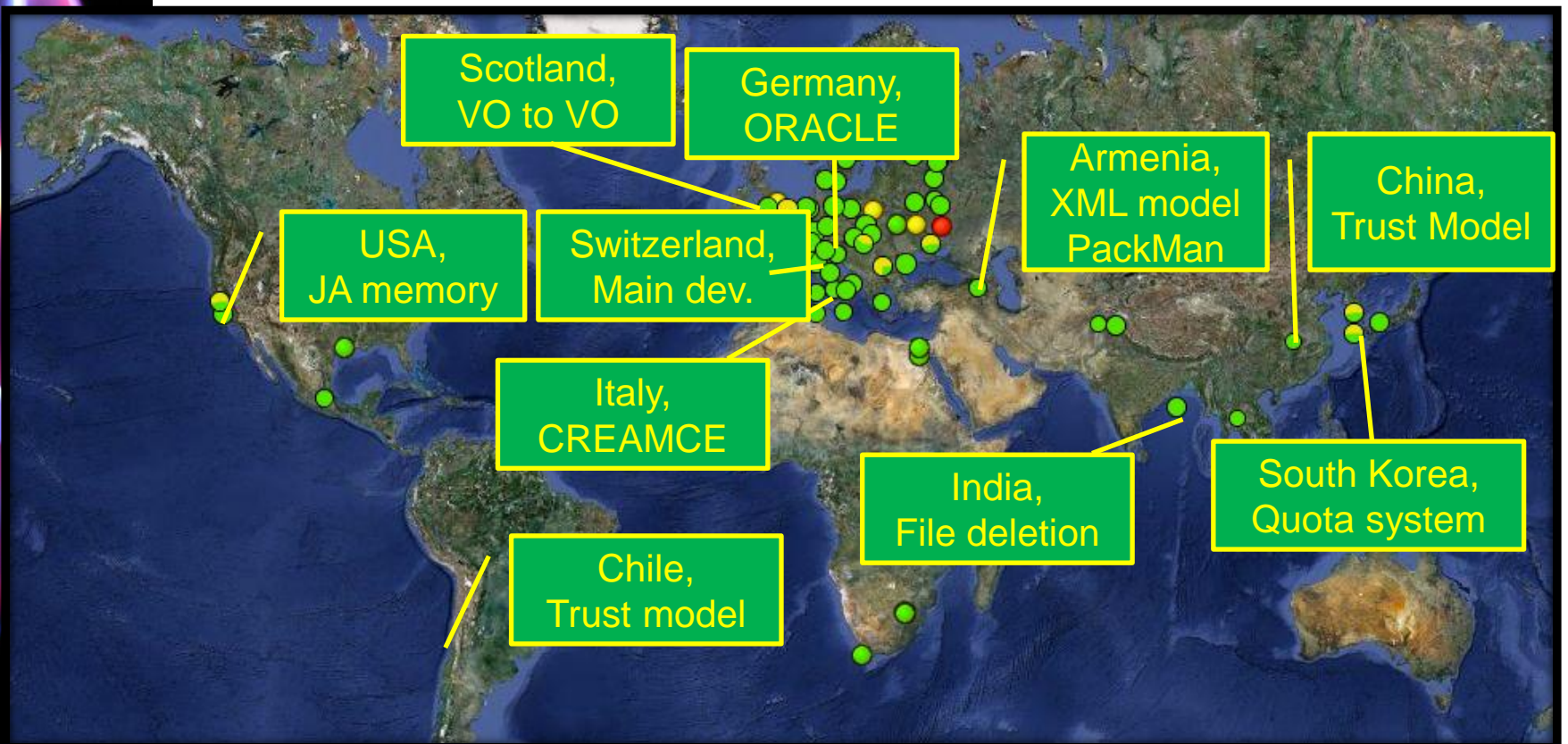
- Plenty of new improvements
  - Catalogue simplification
  - Extreme Job Brokering
  - Removal of PackMan
  - New JDL fields
  - Proxy renewal
 - Job Memory checkup
  - …
- And baseline for new development

- Postpone splitting of job until last moment
- Decide data to be analyzed based on current location of JA & files not analyzed yet
- Can define Max/Min number of files to be analyzed
  - Even if the files are not local
- Less subjobs:
  - Easier merging

CERN**IT**
Department

- File popularity

- Interactive jobs

- Correlate Monitoring

- Multi core jobagents

- Catalogue crawler

- Error classification

- Distributed brokering

- Trust model

- …

The Future

NEXT EXIT →

- Multicore jobagent
  - One agent per core (overkill)

    or

  - One agent per machine (needs development)
- Combining jobs with similar input
  - And dispatch them together
- Distribute (part of) brokering to the vobox
  - Sent multiple jobs to Vobox
    - Let the vobox distribute them among the JA
  - Reduce load on central services
  - Possibility to reuse files from previous jobs

Extreme Brokering

- Postpone splitting of job until last moment
- Decide data to be analyzed based on current location of JA & files not analyzed yet
- Can define Max/Min number of files to be analyzed
  - Even if the files are not local
- Less subjobs:
  - Easier merging
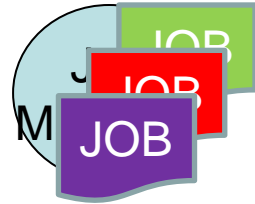
Works nicely if one replica per file
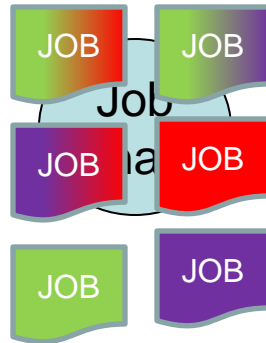
A bit more complex with 3 SE and 2 replicas
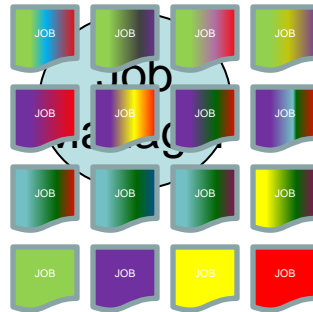
And a lot more with 50 SE and 3 replicas

21 May 2012   *Pablo Saiz*   *21*

Example

	Site A	Site B	Site C
File 1	✓	✓	
File 2		✓	✓
File 3	✓		✓
File 4	✓	✓	
File 5	✓	✓	✓

Current schema
Submit 4 jobs:

File1
File 4

File2

File3

File 5

Broker per file
Submit 3 empty subjobs

If nothing left, just exit

File1, 2,4,5

When a job starts, analyze as much as possible

File 3
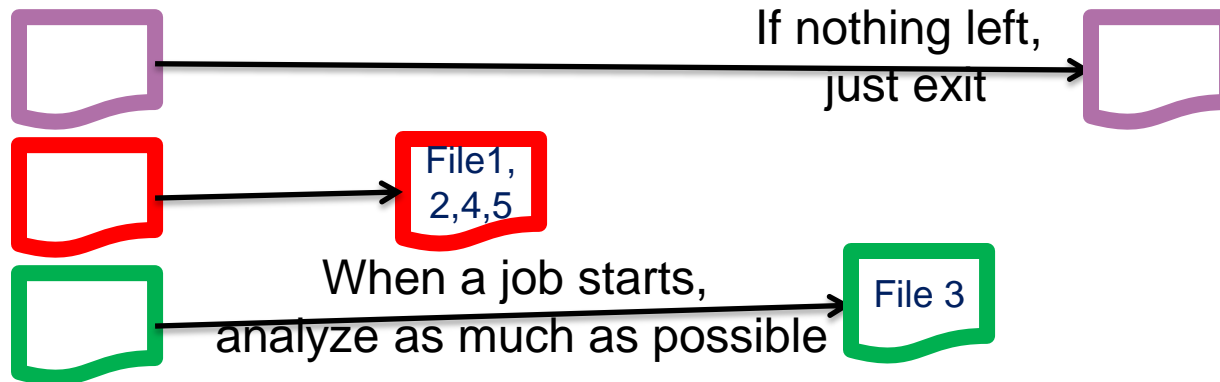
- ## Build system:
  - Multiple platforms
  - Integration & basic functionality tests
    - No API/access from ROOT tests ☹
  - Similar to the AliROOT, ROOT build systems
  - Running the whole system on a single machine
  - http://alienbuild.cern.ch:8888

- AliEn v2.20 ready for deployment
  - With plenty of new features and bug fixes

- Minimize upgrade downtime
  - Create testing setup with several sites, and with all the SE
  - More effort on testing (also from site admins)

- Deploy Test V0 with ALICE sites

- And say goodbye to v2-19 in two months

ES

- ALICE has been using the GRID since 2001
- AliEn provides an interface to the GRID
  - Access to multiple resources
  - Transparent to the user
- AliEn Components
  - File & metadata catalogue
  - TaskQueue
  - Transfer model
- Can be used by other communities
- Plenty of areas for research

http://alien.cern.ch