

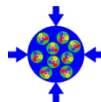
The AliEn File Catalogue

DDM/Rucio workshop

8-December-2011

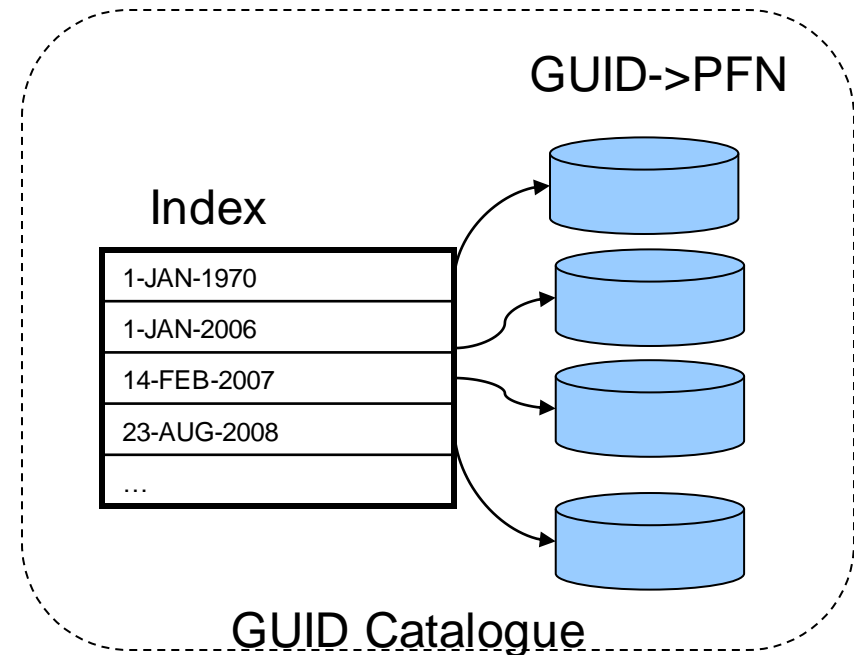
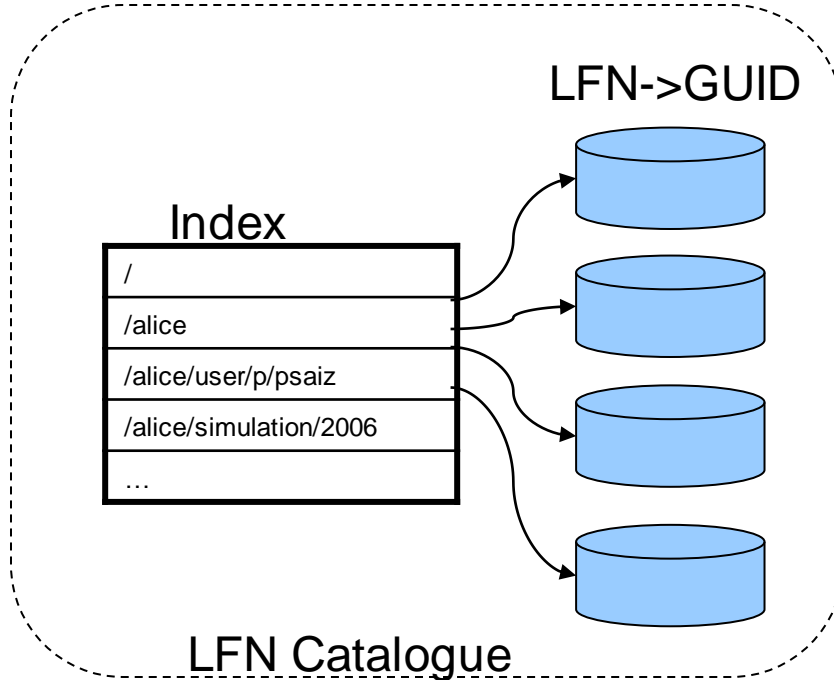
A. Abramyan, L. Betev, D. Goyal, A. Grigoras,
C. Grigoras, N. Manukyan, J. Porter, P. Saiz,
S. Sankar, S. Schreiner

- **Global Unique** name space
 - Mapping from LFN to PFN
- **UNIX-like** file system interface
- Powerful **metadata** catalogue Wildcard queries
- **Automatic** SE selection Rules
- Integrated **quota** system
- Multiple storage **protocols**: xrootd, torrent, srm, file
- **Collections** of files Datasets
- Physical file **archival**
- **Roles and users** Accounts
- Used by **several experiments**



CBM

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

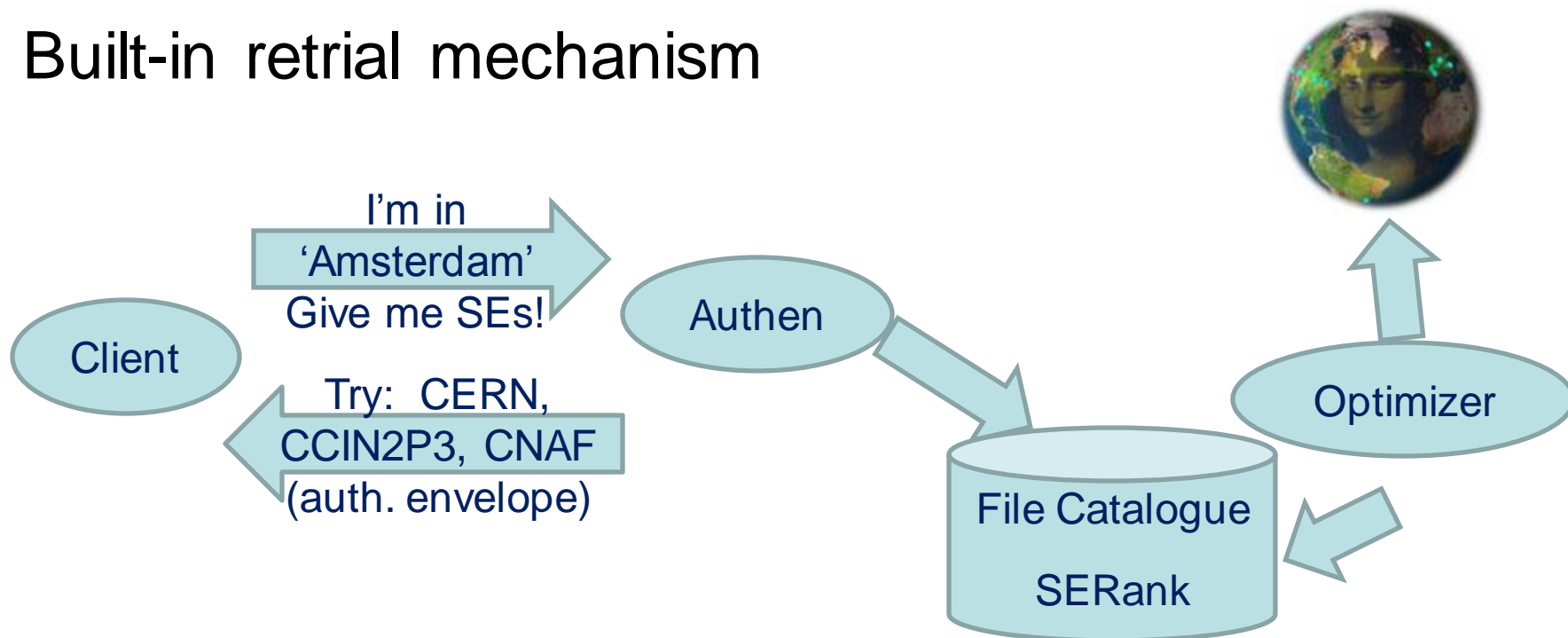


- A collection is a list of LFNs
- Appears as a single entry in the catalogue
 - And can be split according to the location
- Facilitate actions on multiple files: cp, mirror, analyze...

```
[aliendb06c.cern.ch:3307] /alice/data/2010/ > ls -al LHC10c/000120677/collection
crwxr-xr-x  alidaq  alidaq  9564798201  May 24 00:05  collection
[aliendb06c.cern.ch:3307] /alice/data/2010/ > listFilesFromCollection LHC10c/000120677/collection
3C67BA10-66A9-11DF-8001-5523A10ABEEF (from the file /alice/data/2010/LHC10c/000120677/raw/10000120677048.10.root)
61AEF5FE-66A9-11DF-8001-5523A10ABEEF (from the file /alice/data/2010/LHC10c/000120677/raw/10000120677048.100.root)
6538478E-66A9-11DF-8001-5523A10ABEEF (from the file /alice/data/2010/LHC10c/000120677/raw/10000120677048.110.root)
68E76D24-66A9-11DF-8001-5523A10ABEEF (from the file /alice/data/2010/LHC10c/000120677/raw/10000120677048.120.root)
6C99E104-66A9-11DF-8001-5523A10ABEEF (from the file /alice/data/2010/LHC10c/000120677/raw/10000120677048.130.root)
70478C8E-66A9-11DF-8001-5523A10ABEEF (from the file /alice/data/2010/LHC10c/000120677/raw/10000120677048.140.root)
73F4D1E8-66A9-11DF-8001-5523A10ABEEF (from the file /alice/data/2010/LHC10c/000120677/raw/10000120677048.150.root)
77A60474-66A9-11DF-8001-5523A10ABEEF (from the file /alice/data/2010/LHC10c/000120677/raw/10000120677048.160.root)
7BBA4386-66A9-11DF-8001-5523A10ABEEF (from the file /alice/data/2010/LHC10c/000120677/raw/10000120677048.170.root)
7EC525F0-66A9-11DF-8001-5523A10ABEEF (from the file /alice/data/2010/LHC10c/000120677/raw/10000120677048.180.root)
821E1BC6-66A9-11DF-8001-5523A10ABEEF (from the file /alice/data/2010/LHC10c/000120677/raw/10000120677048.190.root)
43DCC43E-66A9-11DF-8001-5523A10ABEEF (from the file /alice/data/2010/LHC10c/000120677/raw/10000120677048.20.root)
857E2892-66A9-11DF-8001-5523A10ABEEF (from the file /alice/data/2010/LHC10c/000120677/raw/10000120677048.200.root)
89464BBC-66A9-11DF-8001-5523A10ABEEF (from the file /alice/data/2010/LHC10c/000120677/raw/10000120677048.210.root)
8CBF8E84-66A9-11DF-8001-5523A10ABEEF (from the file /alice/data/2010/LHC10c/000120677/raw/10000120677048.220.root)
```

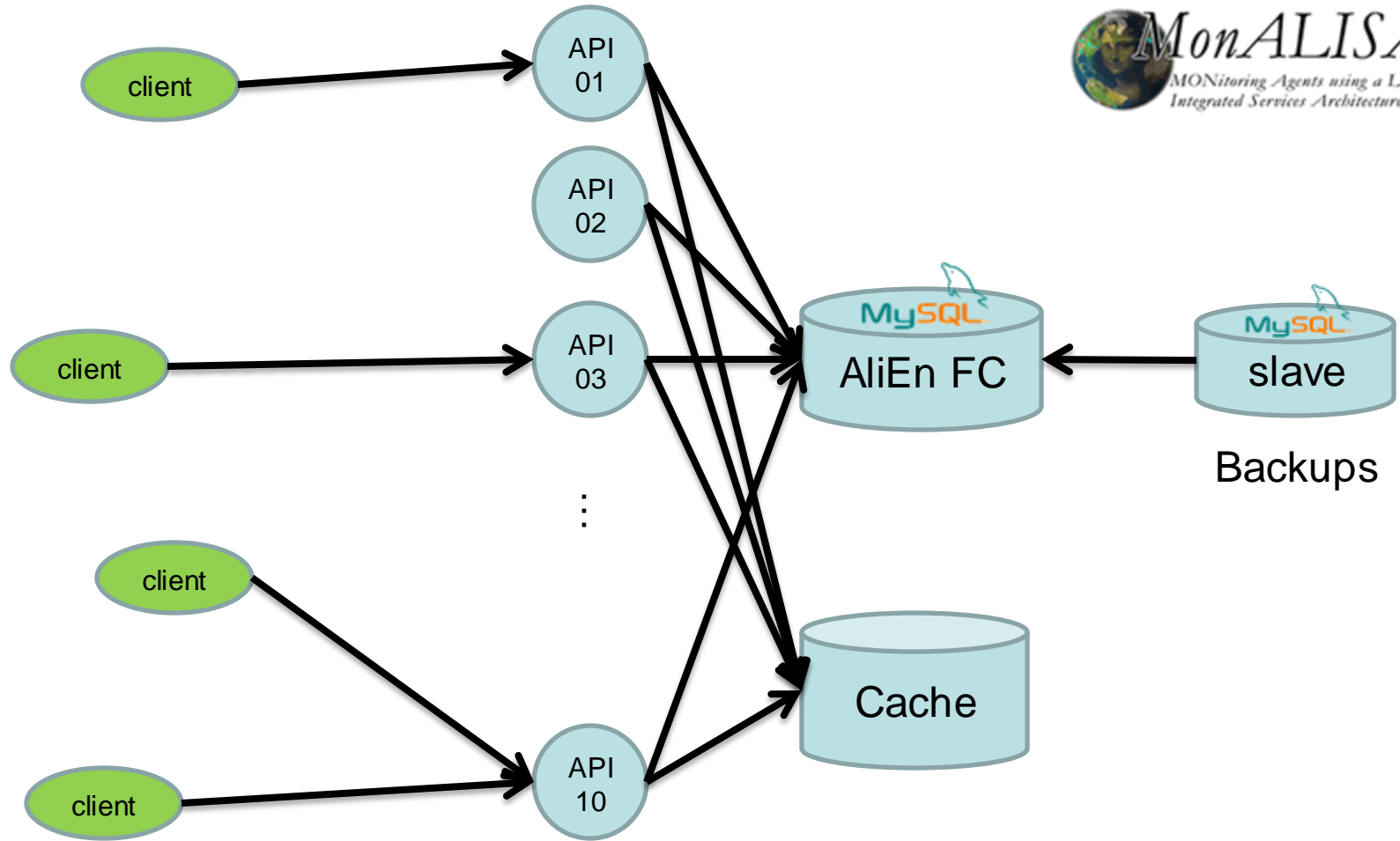
- Mostly **xrootd** protocol
- Jobs only download configuration files
- Data files are accessed **remotely**
 - The **closest replica** to the job, local replica first
- JobAgent uploads **N replicas** of the output
- Schedules data transfers only involve the raw data
 - Done via **xrd3cp**
- Use of **torrent** for AliEn and software package distribution

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



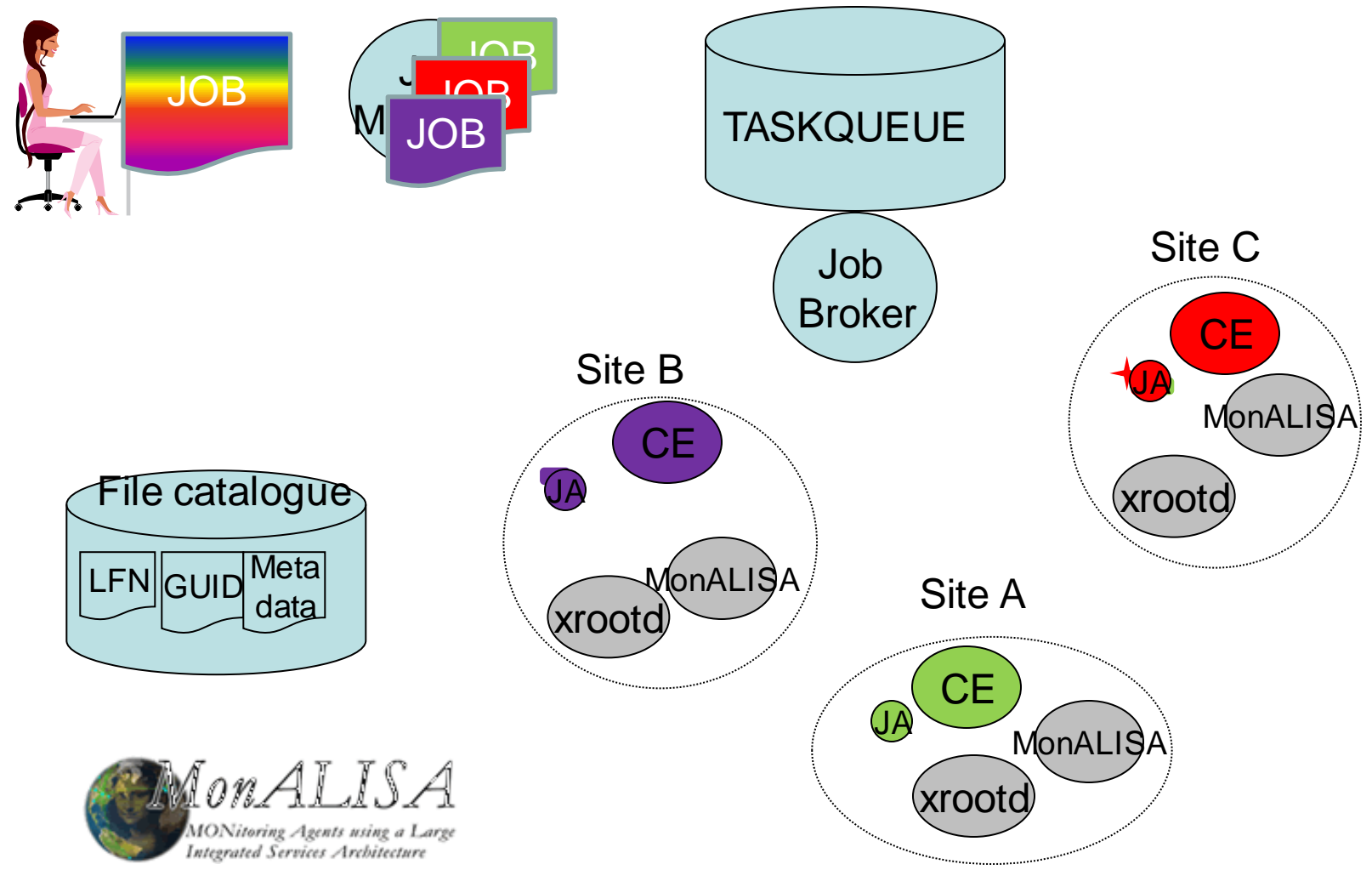
- Based on a **dynamic “distance”** metric from an IP address to a SE
 - Starting from the **network topology**
 - Same site, same AS, same country, continent...
 - Last **functional test** excludes non-working SEs
 - Altered by
 - Reliability
 - Remaining **free space**
 - A small **random factor** to assure 'democratic' data distribution

- Storage-independent
- Handled centrally by the Authen AliEn service
- Checks client credentials and catalogue permissions and issues **access tickets**
 - XML block signed and encrypted by Authen
- The client hands these tickets to the respective storage and (for writes) notifies the catalogue of the successful operation
- Implemented as **xrootd plugin**
 - By Andreas Peters



- Catalogue:
 - LFN: 350 M Size of DB: 270 GB
- Storage:
 - 60 SE
 - Files: 200 M Size on storage: 20PB
- Average values over one day:

Operation	Frequency	Speed	Speed of 90%
Read	350Hz	0.27 s	0.06 s
Find (metadata)	40 Hz	0.08 s	0.04 s
Write	8 Hz	2.5 s	0.2 s



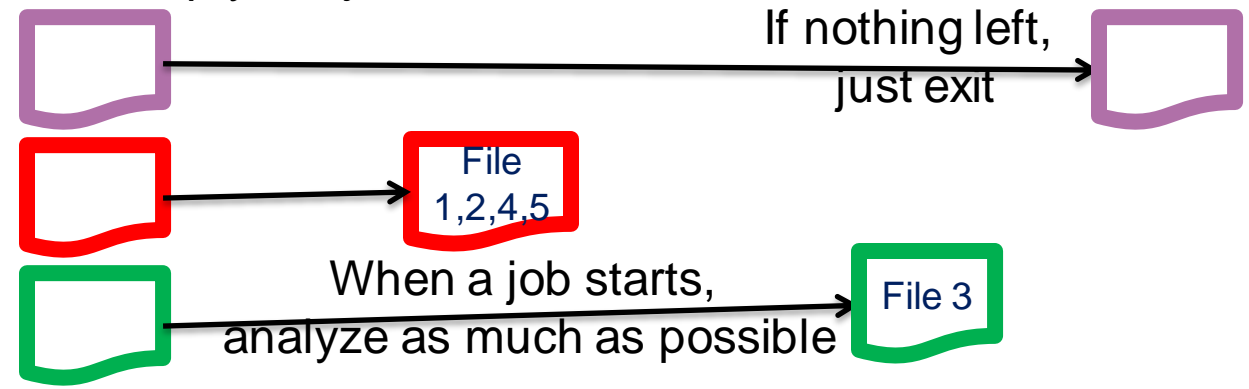
	Site A	Site B	Site C
File 1	✓	✓	
File 2		✓	✓
File 3	✓		✓
File 4	✓	✓	
File 5	✓	✓	✓



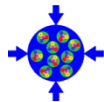
Current schema
Submit 4 jobs:



Broker per file

Submit 3 empty subjobs

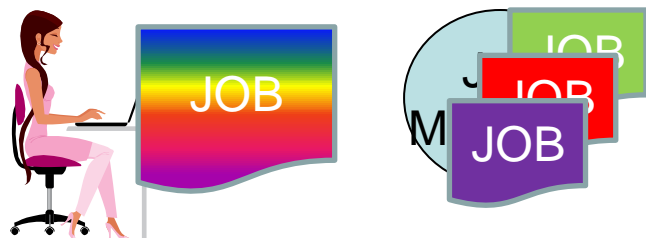


- Global Unique catalogue
 - UNIX-like file system
- Roles, collections, metadata, quotas, archives
- Automatic SE selection
 - With Authentication & Authorization
- Can be taken as a standalone component
 - Or combined with AliEn TaskQueue
- Used by several experiments   
 - And happy to get more customers 😊

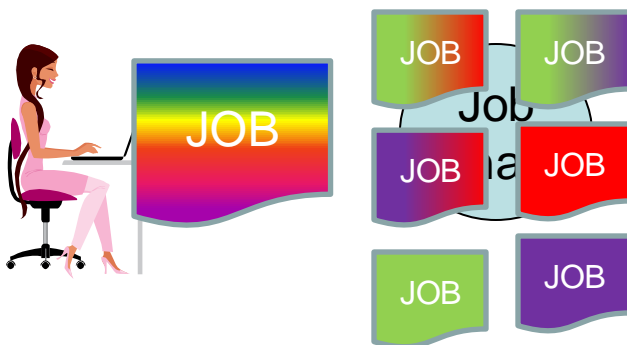
<http://alien.cern.ch>

- Backup

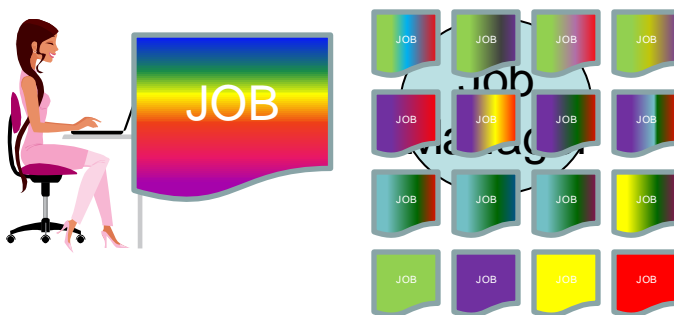
- Basic **commands**:
 - Identical to unix:
 - ls, cp, mv, rm, mkdir, cd, whoami, cat, vi, df, du,
 - With more options:
 - find (searching through name space and metadata)
- Advanced catalogue commands:
 - Quota based: fquota, jquota
 - File location: whereis (physical location of files), mirror (replication of files), listTransfers
 - File collections: createCollection
- Other AliEn commands:
 - Software packages: packman
 - Job interaction: submit, ps, top, kill, spy, trace



Works nicely if one
replica per file

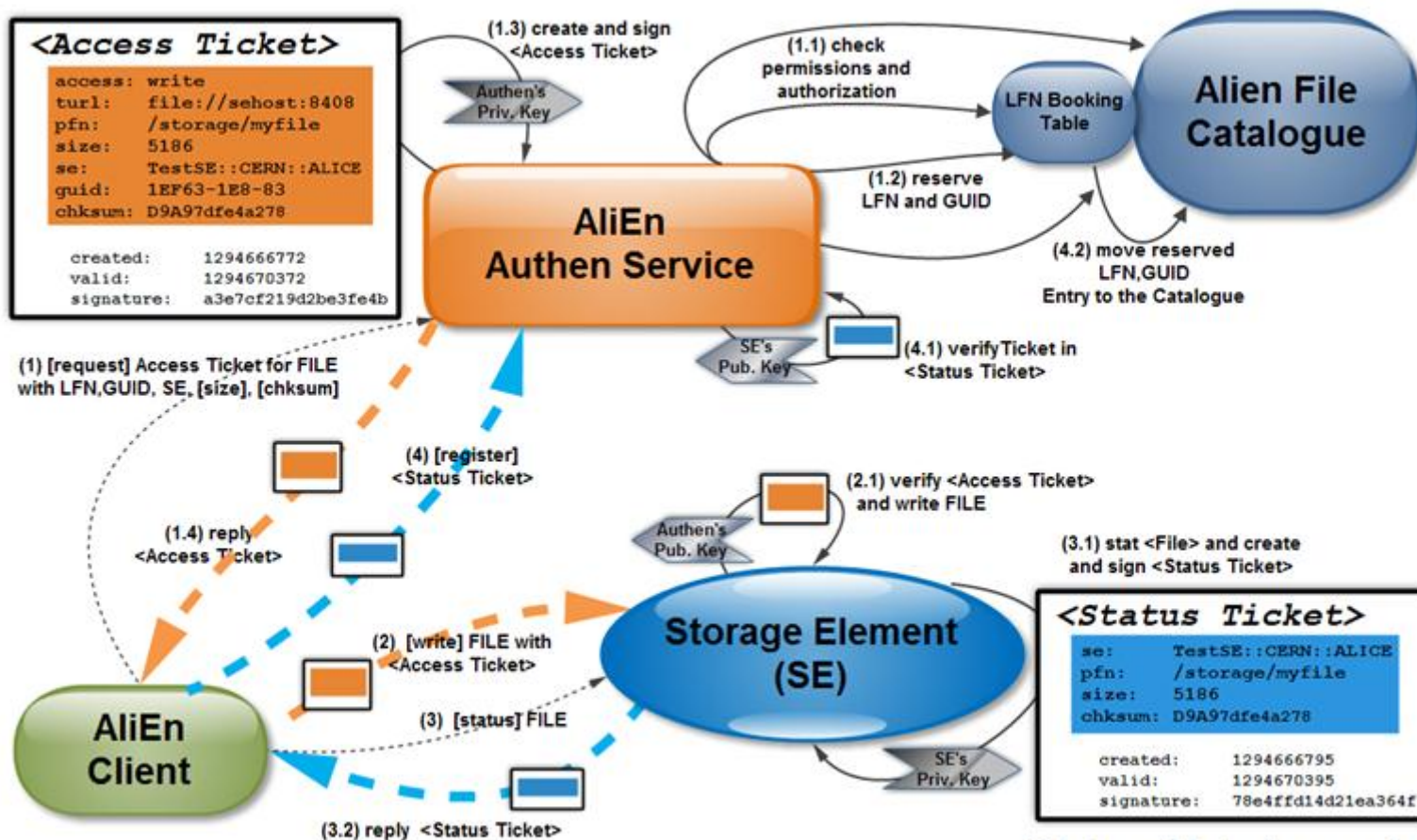


A bit more complex with
3 SE and 2 replicas



And a lot more with
50 SE and 3 replicas

Access Ticket proofs AuthN+AuthZ to the SE



Status Ticket proofs file's existence, size, and checksum to Authen