DATA MANAGEMENT SYSTEM

Christophe Haen

5th DIRAC Workshop, 28.05.15, Ferrara

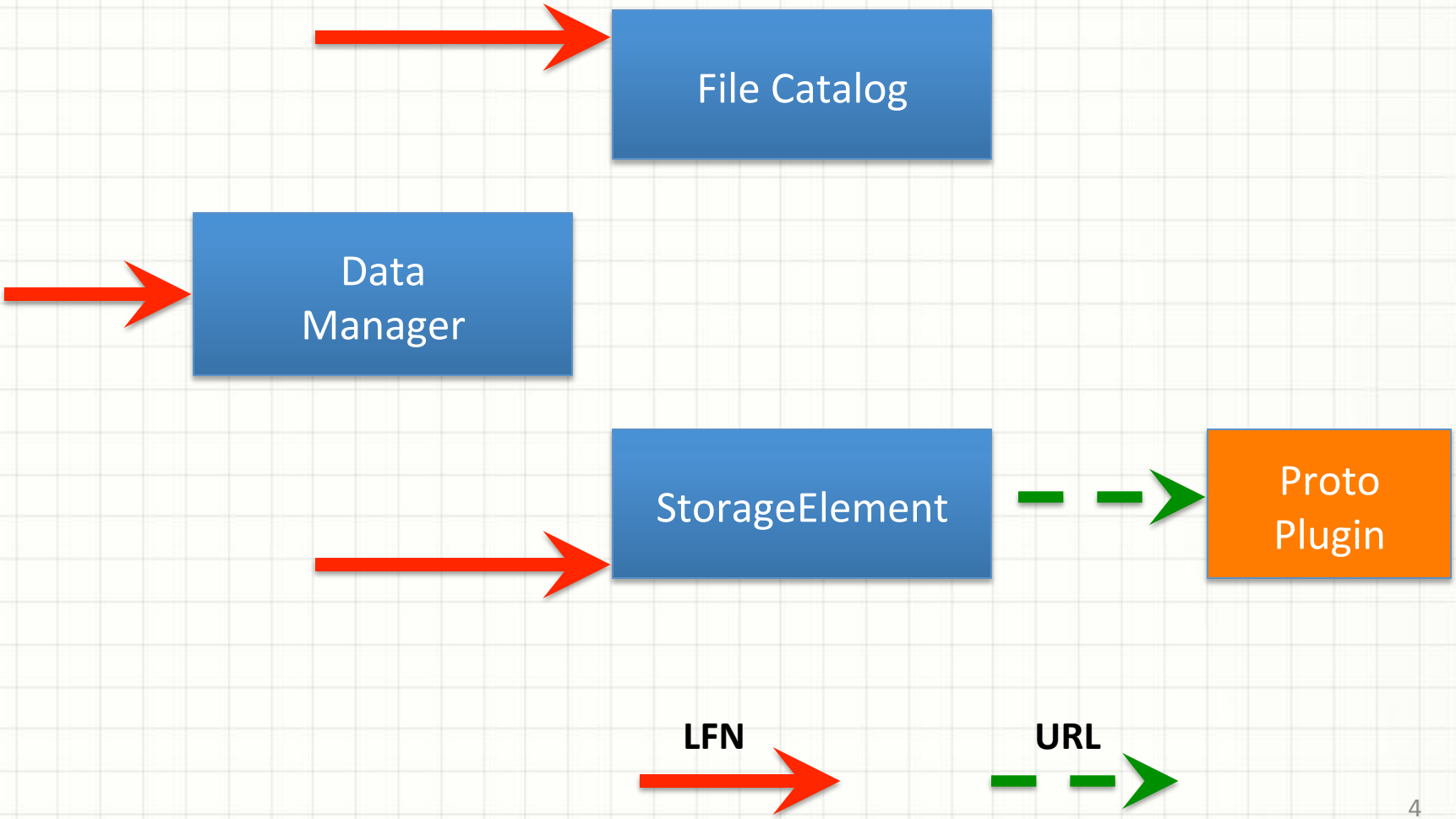
Major change: Good bye PFN

- Continuity of RFC #17
- PFN is meaningless
- Lot of confusion (getPFNForLFn, getPfnForProtocol, getPfnPath, getAccessUrl, …)
- No need for it until the very last moment when physically accessing the file

Major change: Good bye PFN

- A file is identified by its LFN
- A replica is identified by (LFN, SE)
- ***No PFN/URL can be given to a StorageElement***
- Long and painful development, but a lot of simplification

Major change: Good bye PFN



Major change: Good bye PFN

- Paves the way for protocol negotiation (e.g. xroot -> xroot, …)
- ***The LFN must start with your VO name***

Adapt your extension code

Minor change: StorageElement cache

- Instantiating a SE is not so cheap…
- Lots of “tricks” implemented on the client side

  → Change the SE implementation
- Done ! (v6r12)
- Cache system with life time
- Totally transparent to the clients ☺

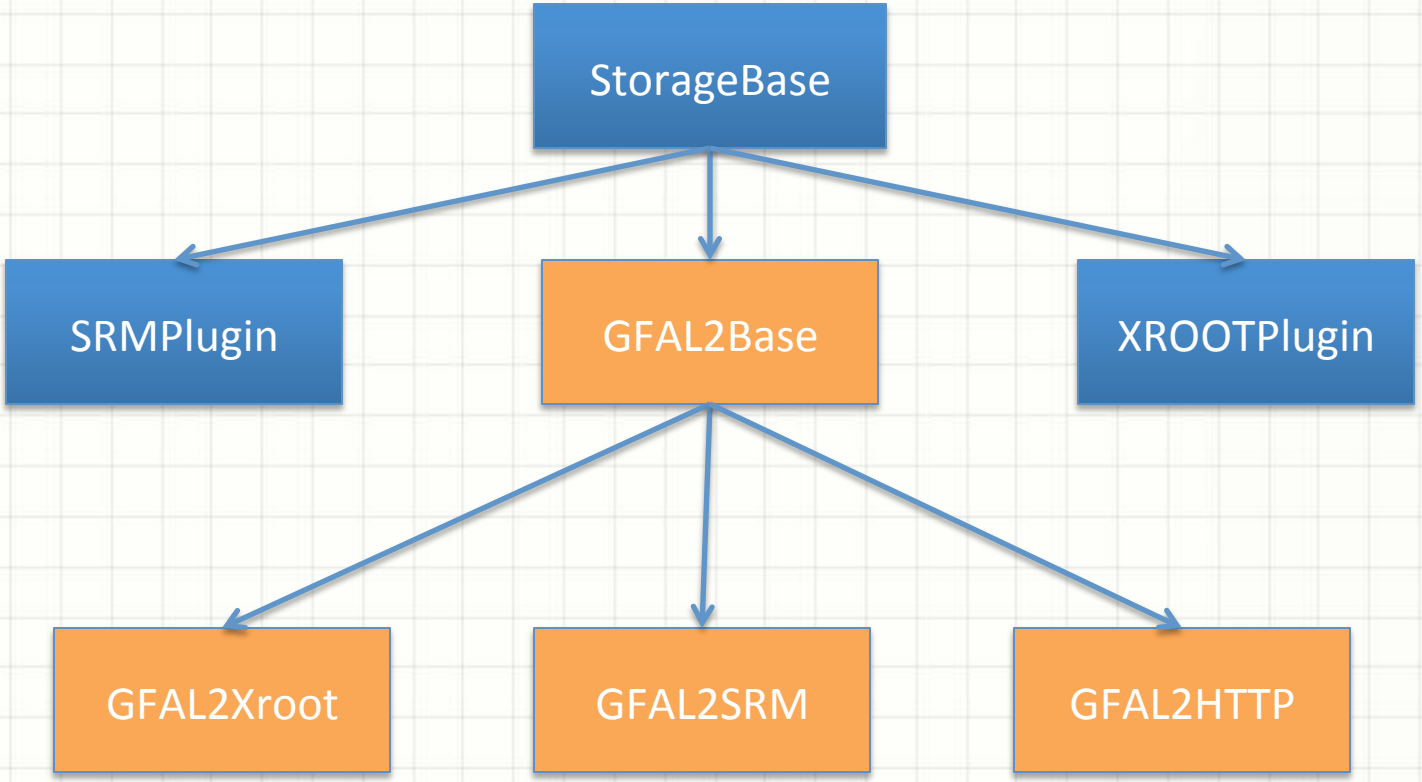
Minor change: Base SE

- Many SE with a very similar configuration
- Cumbersome to maintain the configuration
- Solution: base SE as a “template”
- Different than Alias !
- Alias -> reference (programing)
- Base SE -> inheritance

Major to come: gfal2

- Currently, only SRM really working. Xroot existing but not used
- One library per protocol
- Gfal is not supported anymore
- Gfal2 support multiple protocols and is heavily used
- Implemented, not certified yet

Major to come: gfal2



DataLogging

- Keep track of *everything* that happens to a file
- Currently being worked on
- A priori transparent to the client, everything done in the DataManager, StorageElement and FileCatalog
- Answer the question “what happened to that file?”

DIRACFS Prototype

- Contribution of Xiaobo Li
- Fuse FS for DFC browsing
- Any idea what to do with it ?

Error reporting

- Not good….
- Will be the first place to use the new Error system presented by Adry
- Will be good ☺

Catalog instantiation

- Problem: no granularity when running multiple catalogs
- LHCb example: Bookkeeping only for production files
- Proposal: conditional usage of the catalogs
- CAUTION: different from permissions

Catalog instantiation

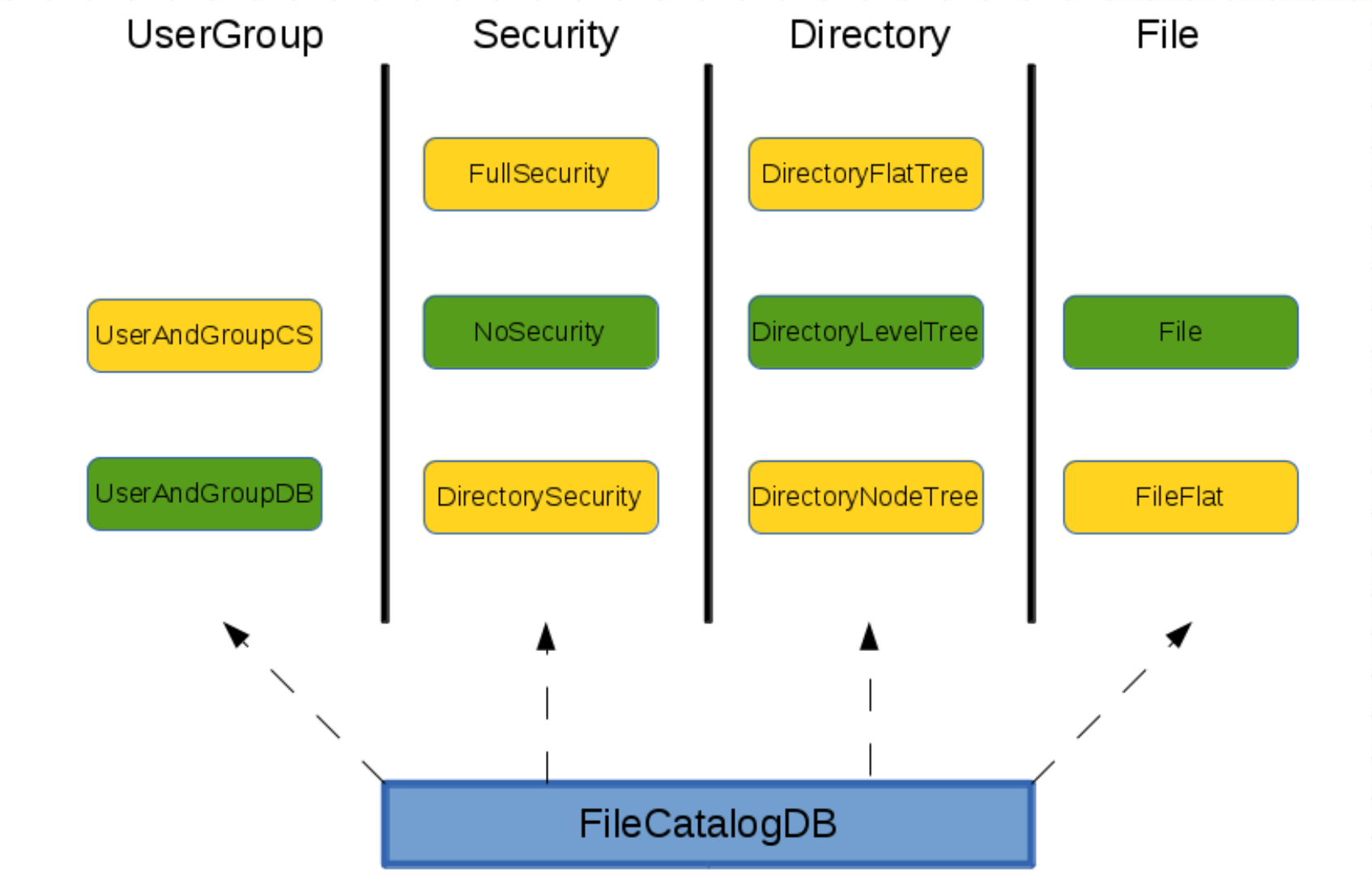
- New section in the CS for each catalog
  - Option name: catalog method name
  - Magic keys: ALL, READ, WRITE
  - Value: boolean expression
    - [plugin1 = arg1] & [[plugin2 = arg2] | [plugin2 = arg3]]

Catalog instantiation

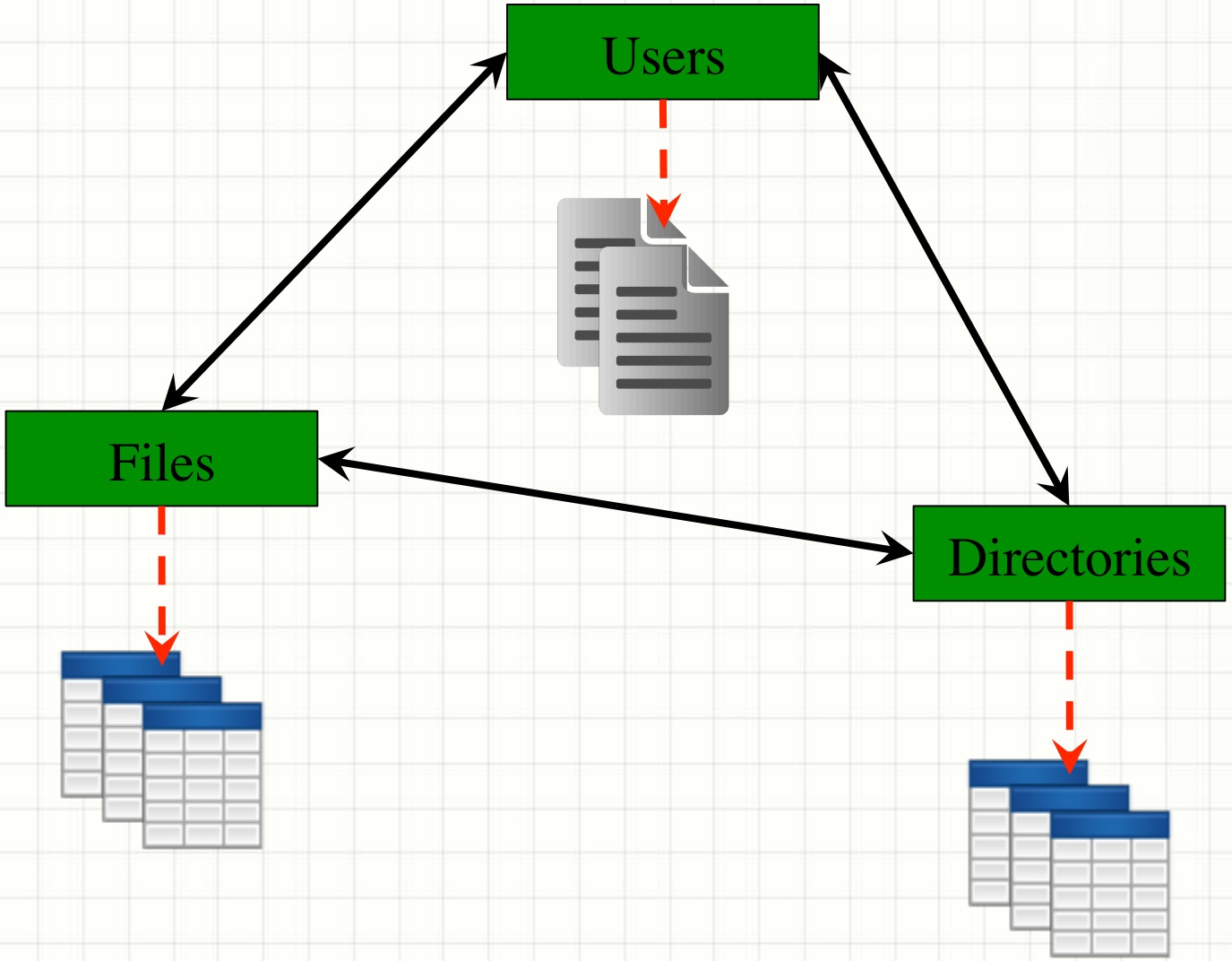
- [filenamePlugin=startswith('/lhcb/MC')] & ! [lfnLengthPlugin= >8]
- filenamePlugin=endswith('.root')

- Very flexible
- Easy to write new plugins
- Cons: extra calls

DFC


UserGroup
Security
Directory
File
FullSecurity
DirectoryFlatTree
UserAndGroupCS
NoSecurity
DirectoryLevelTree
File
UserAndGroupDB
DirectorySecurity
DirectoryNodeTree
FileFlat
FileCatalogDB

DFC



DFC for LHCb

- Focus on the Replica catalog functionalities
- At least end of run2 in mind
  - 20M directories
  - 100M files
  - 200M replicas
- Enforce consistency of entities

- Hosted on the CERN MySQL Database On Demand instances

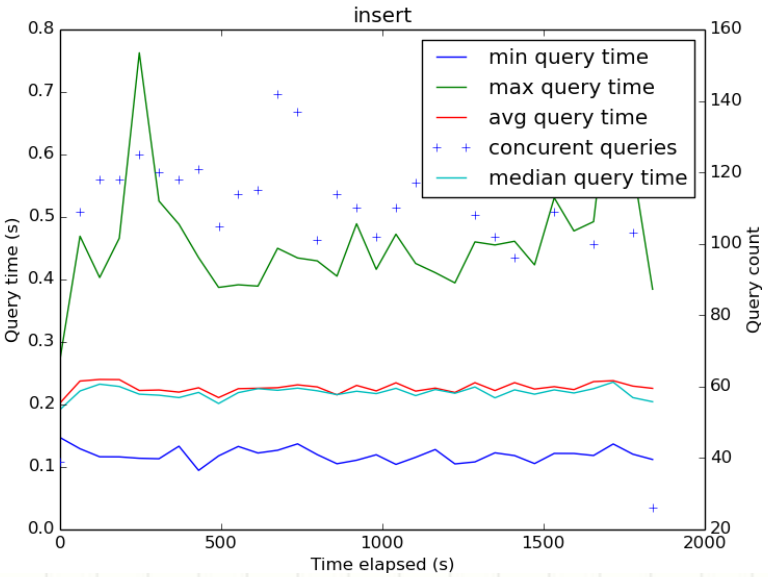
DFC for LHCb

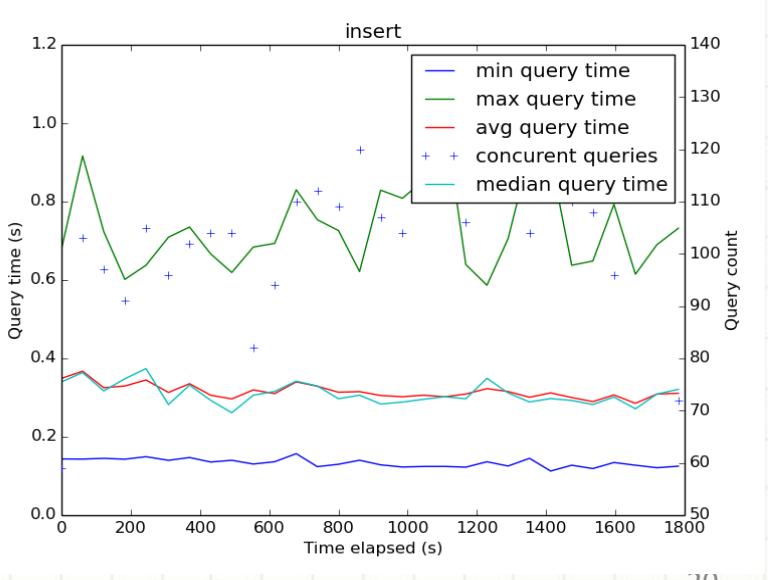
- Only for files/replicas & directories
  - Beauty of AOP
- Introduced the use of Foreign keys
  - Consistency at the cost of coupling in the DB
- Use of stored procedures
  - Compiled “once”
  - Ease manual interventions if necessary
- Use of closure
  table to store the hierarchy

DFC perf with SP

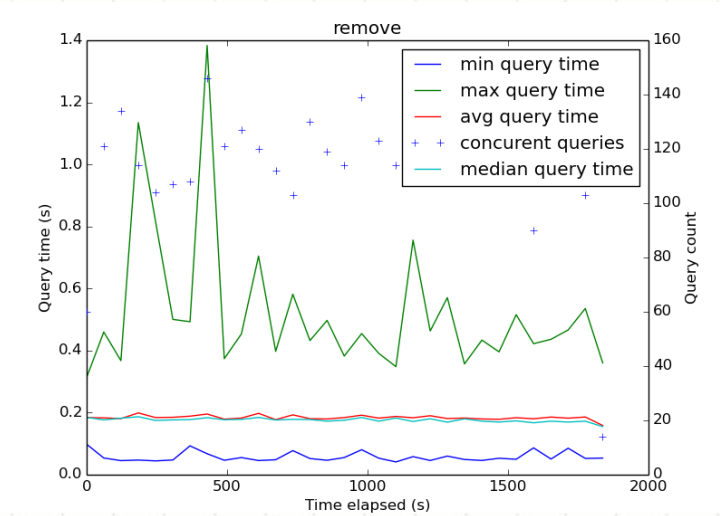
- Access pattern
  - Read: 90%
  - Write: 5%
  - Delete: 5%

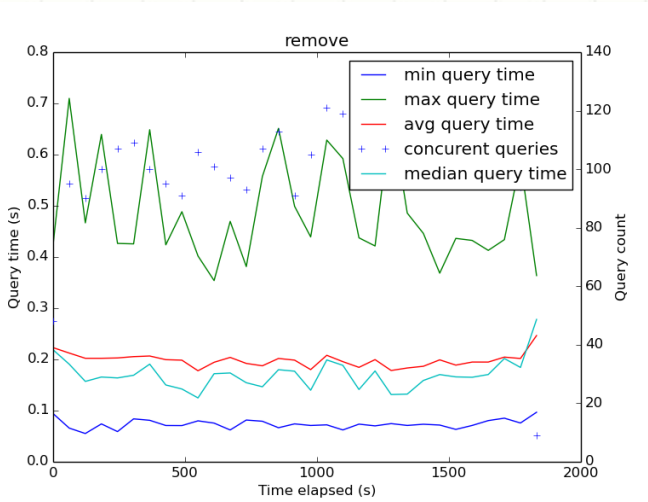
- Bulk operations (1-1k files)

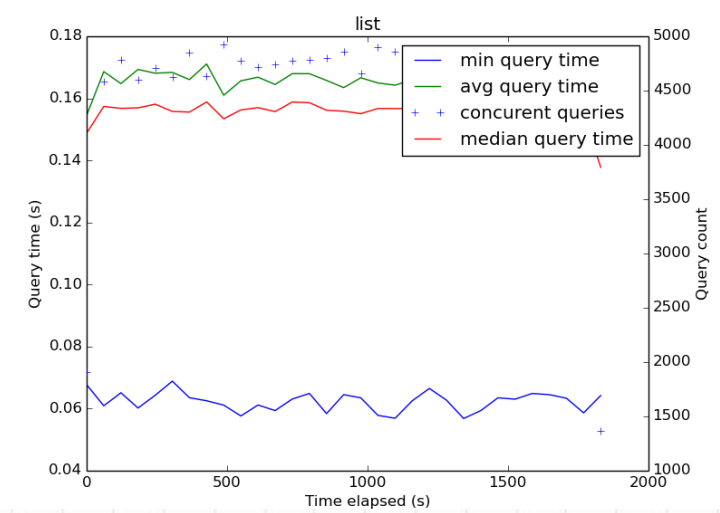


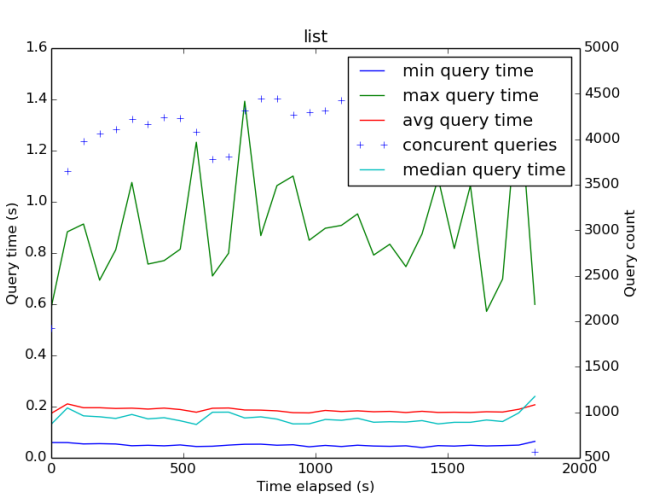


DFC perf with SP







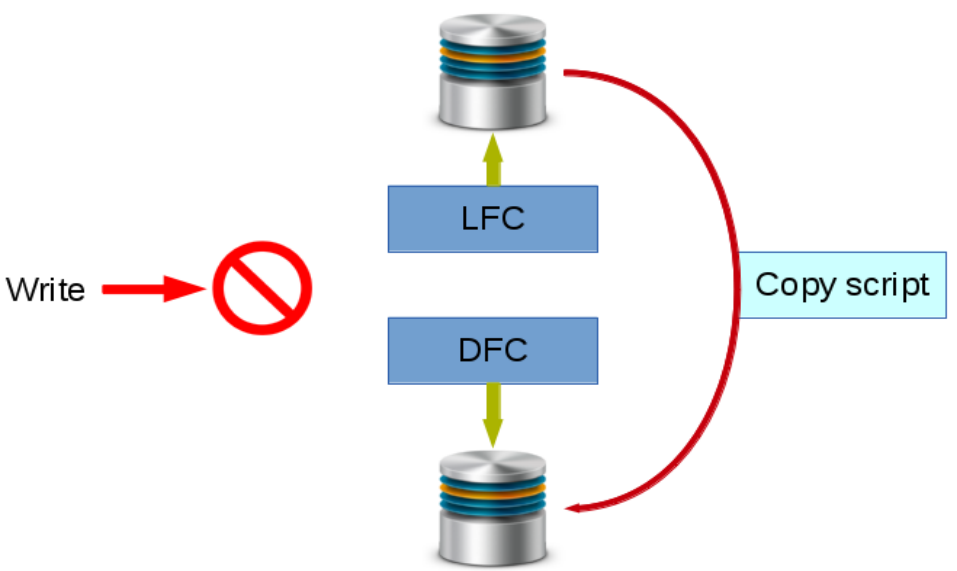


From LFC to DFC

- Before : cleanup of the LFC
- Actual migration (next slide)
- Keep synchronization for a few months
- Farewell party for the LFC

From LFC to DFC

- Stop all writes
- Keep the read in LFC
- Copy from Oracle LFC to MySQL DFC
  - Custom SQL and Python script
- DFC & LFC in write
  - Keep in sync
- Only DFC in read
  - Spot problem if any
- Failover system for running jobs



From LFC to DFC

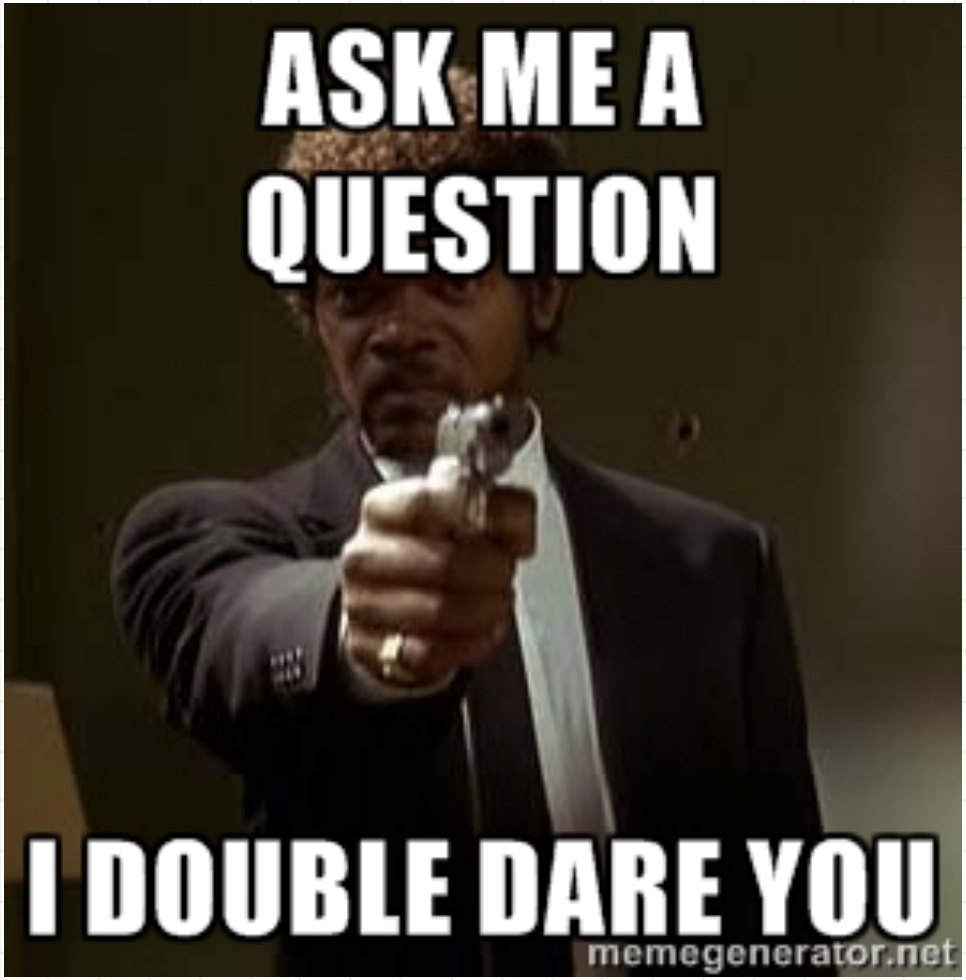
- Smooth migration
- Few surprises/overlooked aspects
  - Slight API difference LFC/DFC
  - VOMS roles -> DIRAC users/groups

DFC: Policy management

- Granularity of permissions too coarse (“can I write ?”)
- Move to more specific tests (“Can I add a replica ?”)
- New SecurityManager
  - “Groups of groups”: Files/Directories “belong” to several groups
  - VOMS aware
  - POSIX inspired
    - 3 levels: directory/file/replica
    - Not transversal

From LFC to DFC

- LHCb has done the hard work
- It should be much easier for you
- We can share everything with you
  - Benchmark scripts
  - Migration plans
  - Migration scripts
  - Celebration Champagne you will open once done


ASK ME A QUESTION
I DOUBLE DARE YOU
memegenerator.net