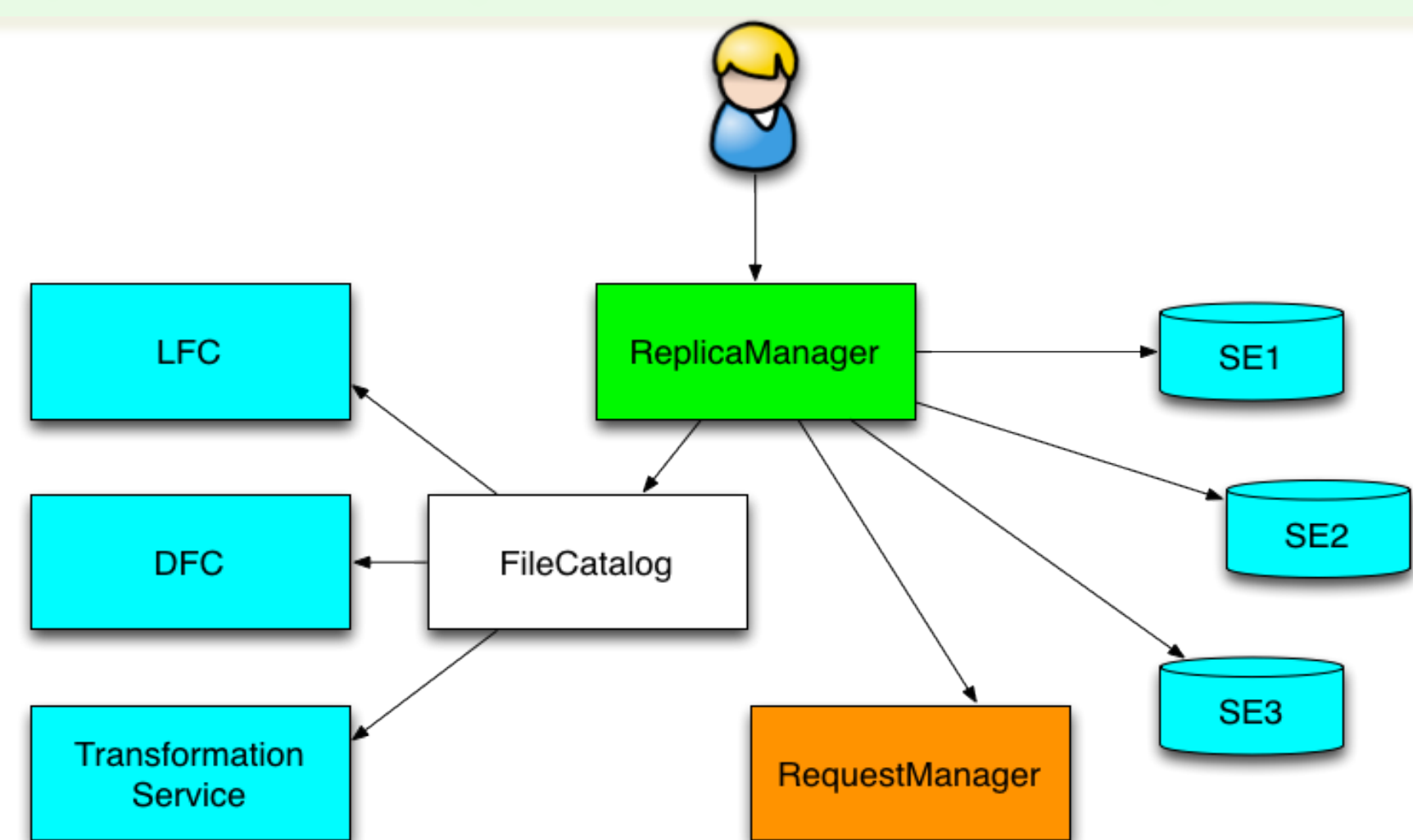


## File Catalog Framework

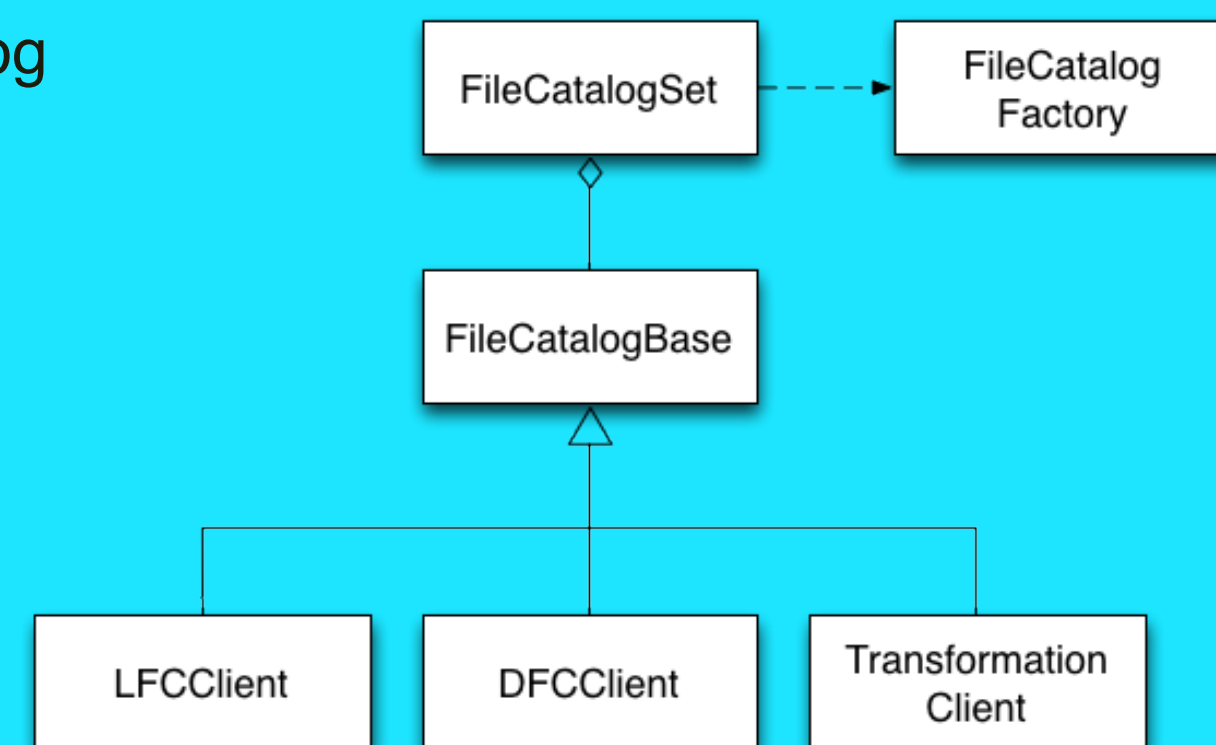
DIRAC is capable of working with several File Catalog services simultaneously. The FileCatalog user interface object can contain several Catalog Clients as defined by the configuration parameters. Each Catalog operation is executed by each Client. In case of data registration failures, the failover requests are sent to the RequestManager service for later retry once the service is available again.



DIRAC introduces a File Catalog Client abstraction layer with specific implementations for a number of catalog services:

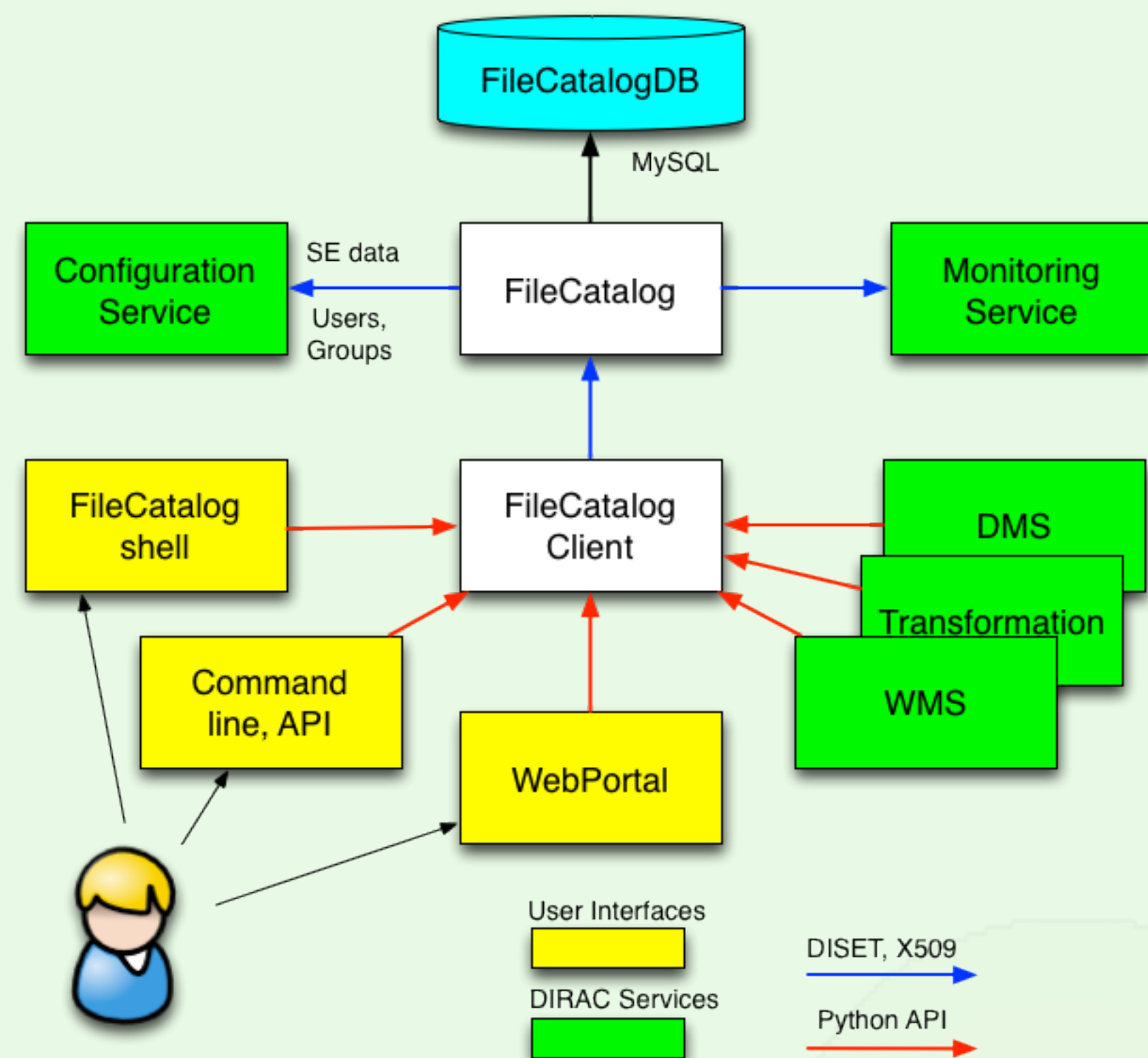
- LCG File Catalog Service (**LFC**)
- DIRAC File Catalog Service (**DFC**)
- DIRAC Transformation Service

More clients can be added for specific community services. For example, **LHCb** uses *Bookkeeping Client* for its custom provenance database service.



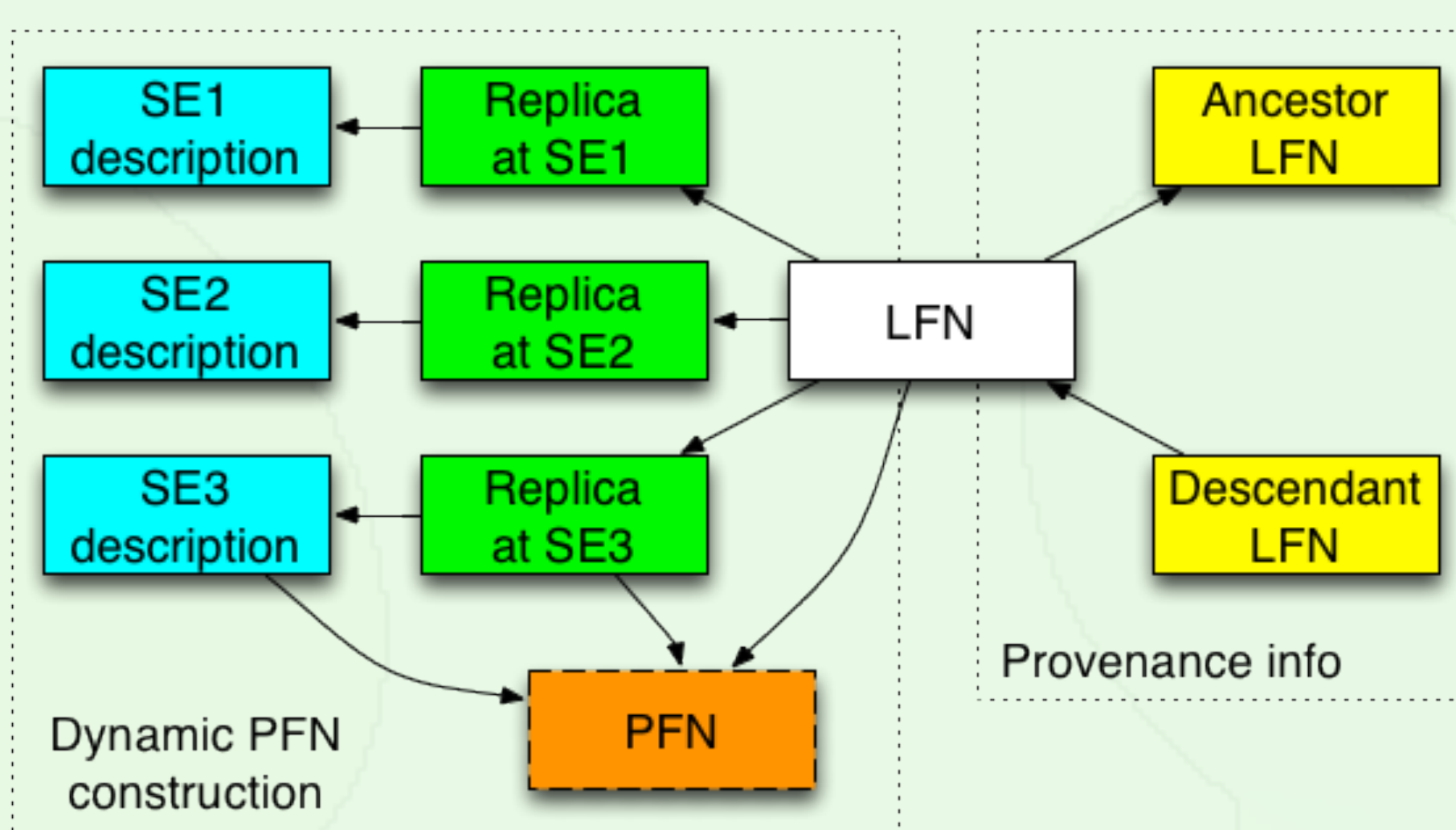
## File Catalog Service

The File Catalog Service is built entirely in the DISET Framework of the DIRAC Project. It is tightly coupled with multiple other DIRAC services and provides several User Interfaces



## Replica Catalog

The Logical File Names ( LFN ) are stored in a hierarchical directory structure. The Replica data consists of references to Storage Elements containing Physical Files. The Physical File Names ( PFN ) are constructed on the fly on client requests.

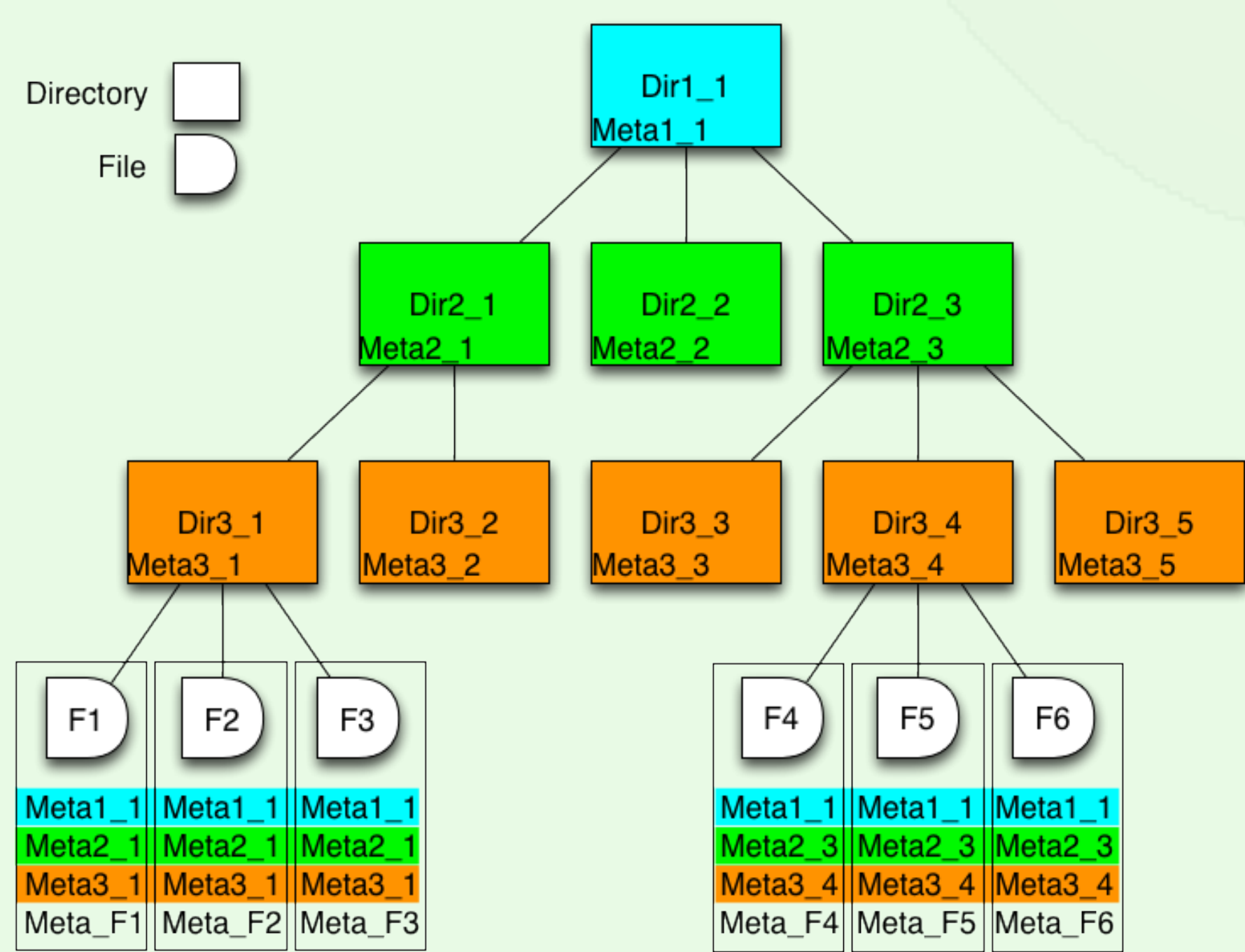


Features:

- Support for ancestor-descendant relations between LFNs
- Pluggable Security Modules (NoSecurity, GlobalRead, POSIX like)
- Support for efficient “disk usage” reports
- Emphasis on efficiency of bulk queries

## User Metadata Catalog

The User Metadata can be defined for each directory and/or file in the same hierarchy as the Replica Catalog. Subdirectories and files are inheriting metadata of their parent directories.

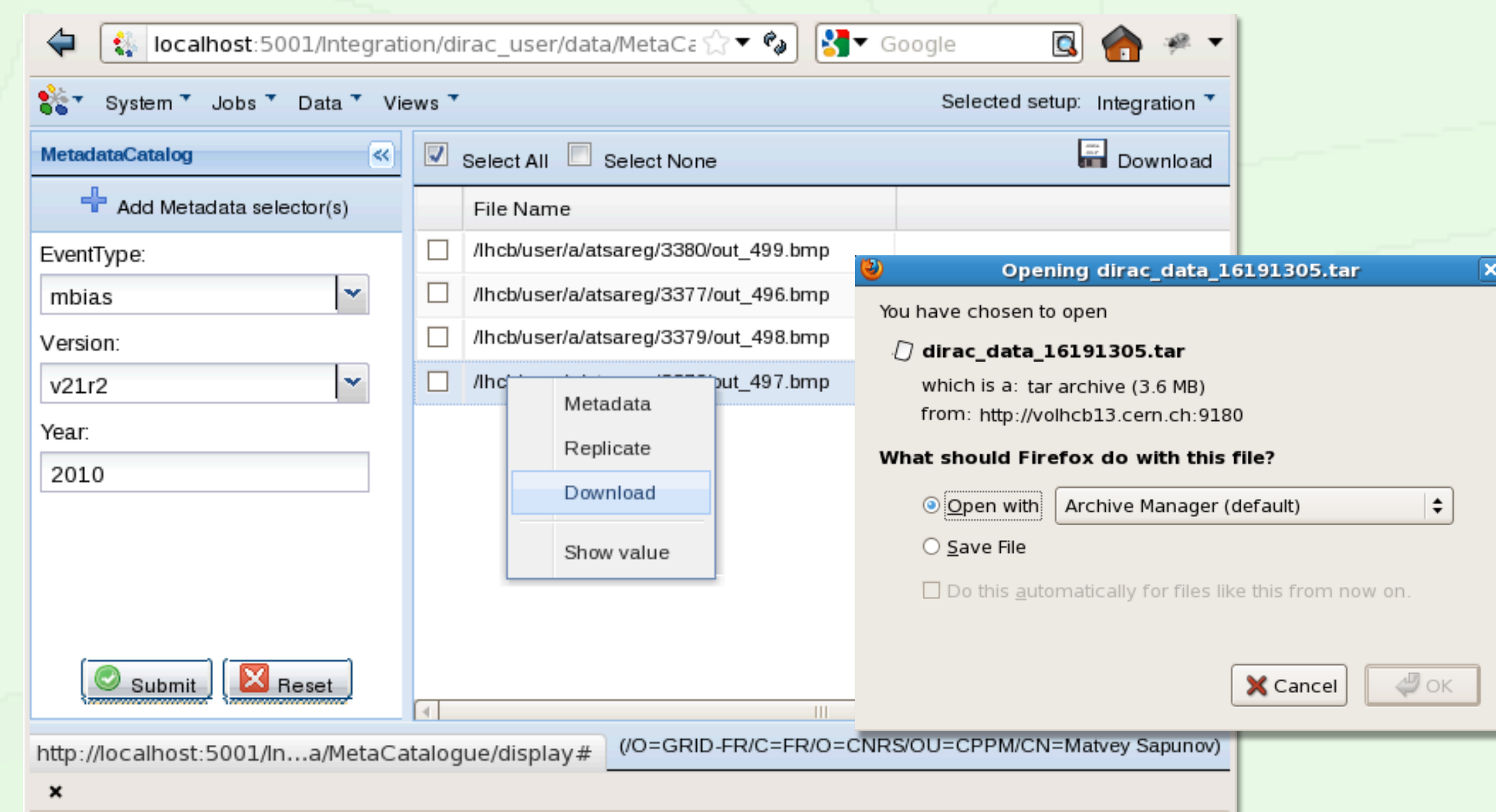


Example search for files by metadata:

**FC > find /lhcb/data Year>2010 EventType=mbias Version=v20r1,v21r2**

## Web User Interface

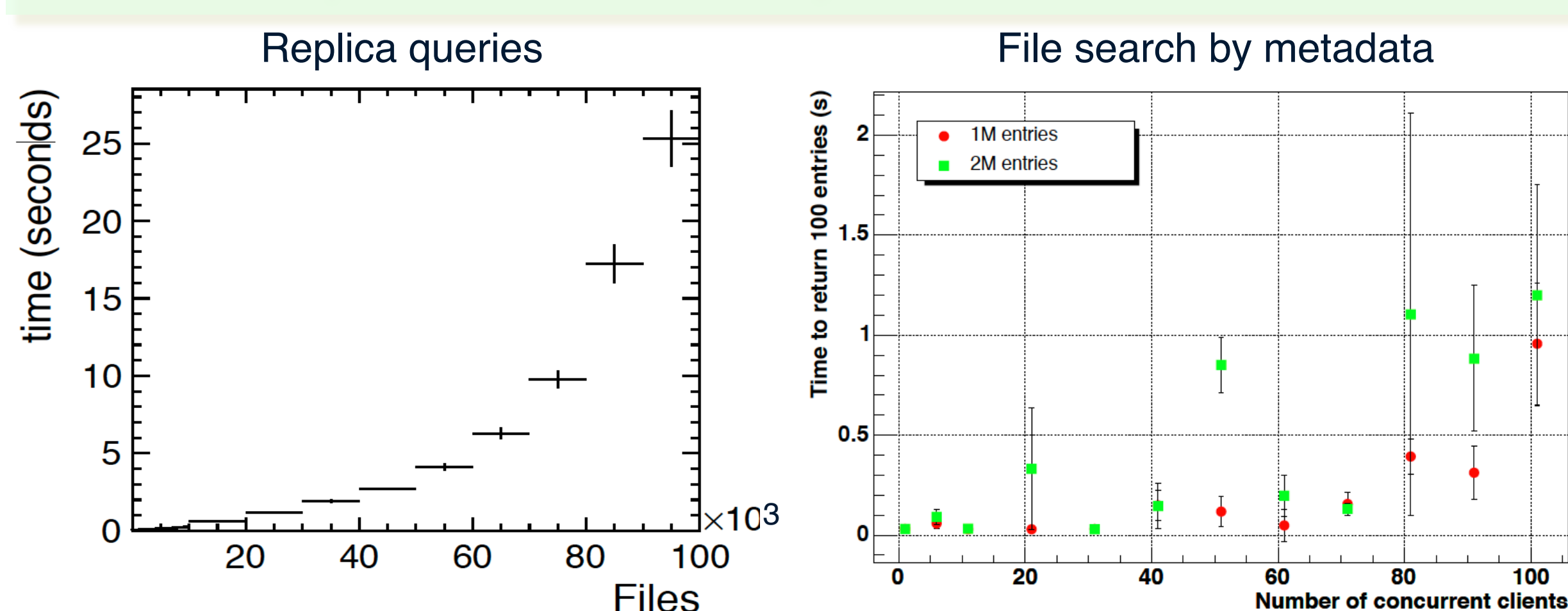
An intuitive Web Interface is making part of the DIRAC Web Portal. Support for data search by metadata as well as directory browsing is provided.



Common data operations like looking up file annotations, replication, data downloads and others are supported using user grid certificates. Access to other functionalities, e.g. bulk replication service is developed.

## Performance

Replica queries for random large collection of files are optimized due to a small footprint of the database. Storing metadata in a hierarchical structure helps to increase efficiency in the search of the data.



## Usage

The DIRAC File Catalog is used by several Collaborations for their production activities:

- ILC
- CLIC
- BES III



The DIRAC File Catalog with the complete set of the *LHCb* data ( 9M files, 1.2M directories ) shows performance superior to the one of LFC. Tests done by the BES Collaboration demonstrated the performance of the Metadata Catalog similar to that of the AMGA service. The DIRAC File Catalog is provided for several grid projects ( *GISELA*, *France-Grilles*, *IberGrid* ). Combining both Replica and Metadata Catalog in a single service offers clear advantages to the users.

