gLite Data Management Components

Presenter : Ákos Frohner, CERN, IT-GD
On behalf of the Grid Data Management Team

EGEE'07, 1-5 Oct 2007, Budapest Hungary

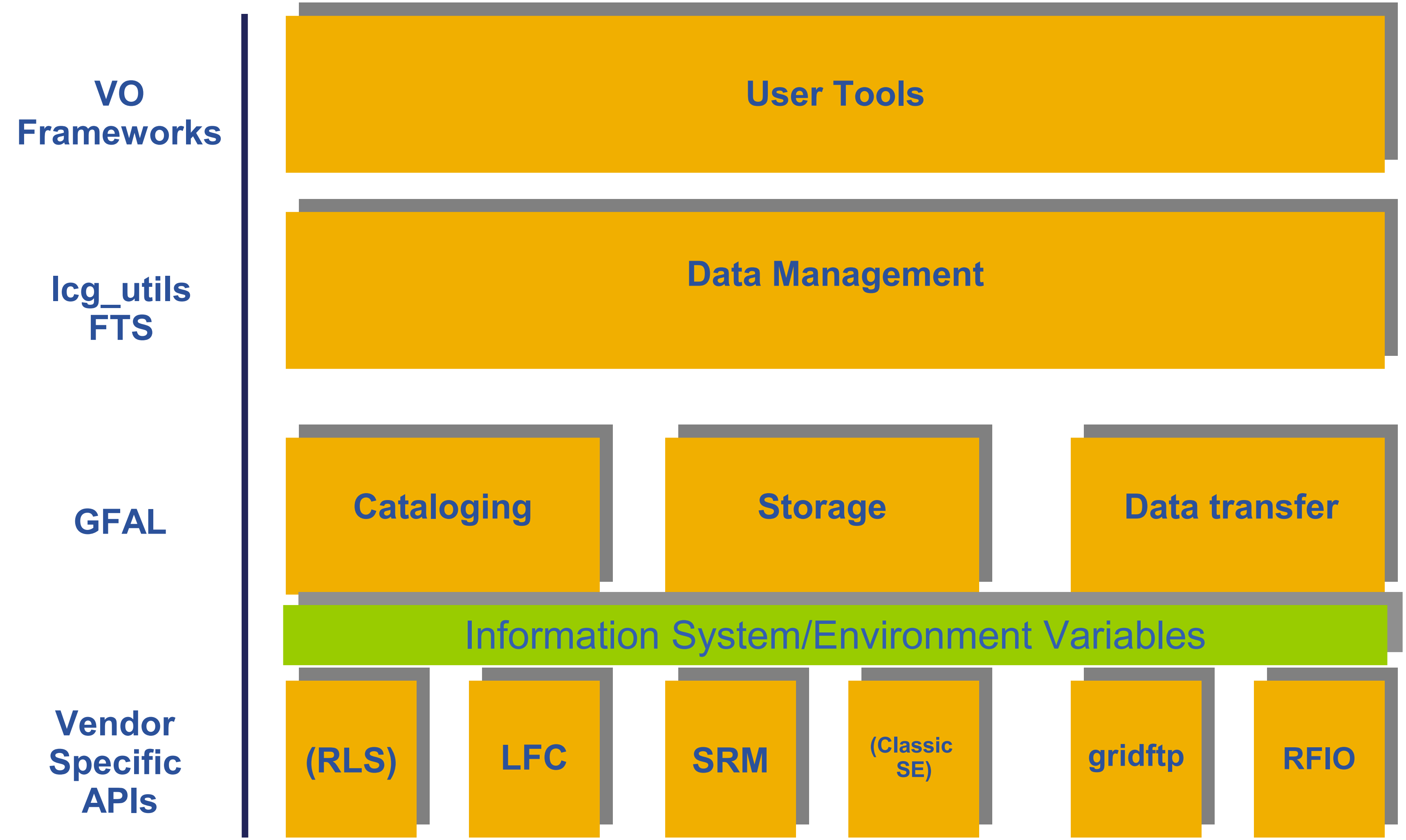
Akos.Frohner@cern.ch

Lightweight Middleware for Grid Computing

www.eu-egee.org

Information Society and Media

EGEE Data Management



VO Frameworks

User Tools

lcg_utils FTS

Data Management

GFAL

Cataloging

Storage

Data transfer

Information System/Environment Variables

Vendor Specific APIs

(RLS)

LFC

SRM

(Classic SE)

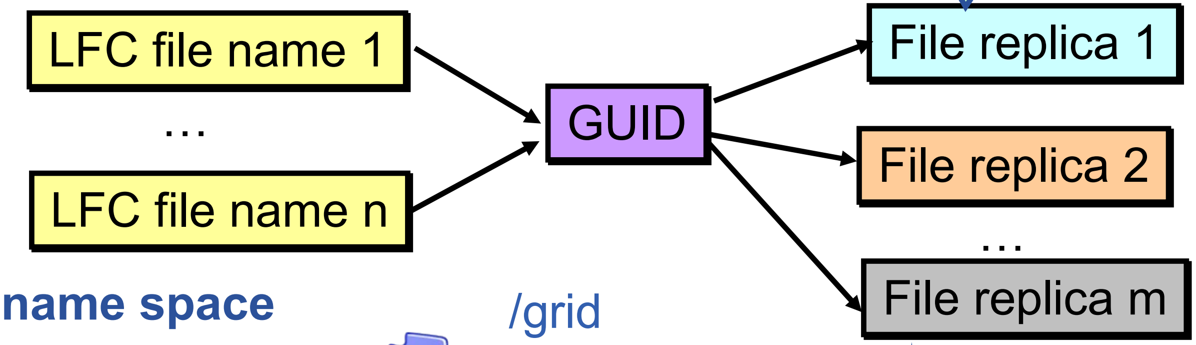
gridftp

RFIO

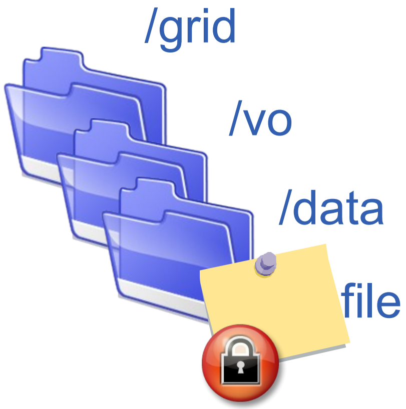
LCG “File” Catalog

- **The LFC stores mappings between**
  - Users’ file names and file locations on the Grid
  - Stores Permissions and
  - Ownership
  - Simple metadata

“Replicas” are “Copies”



- **Provides a hierarchical name space**
- **Supports GSI security model**
  - Including VOMS based ACLs
    - Very fine grained control
    - Implementation based on virtual IDs
  - Soon: encrypted channels
- **Simple DLI interface**
  - Data Location Interface
  - GUID <---> Location
  - Integration with WMS&RBs



All files are “Write Once”

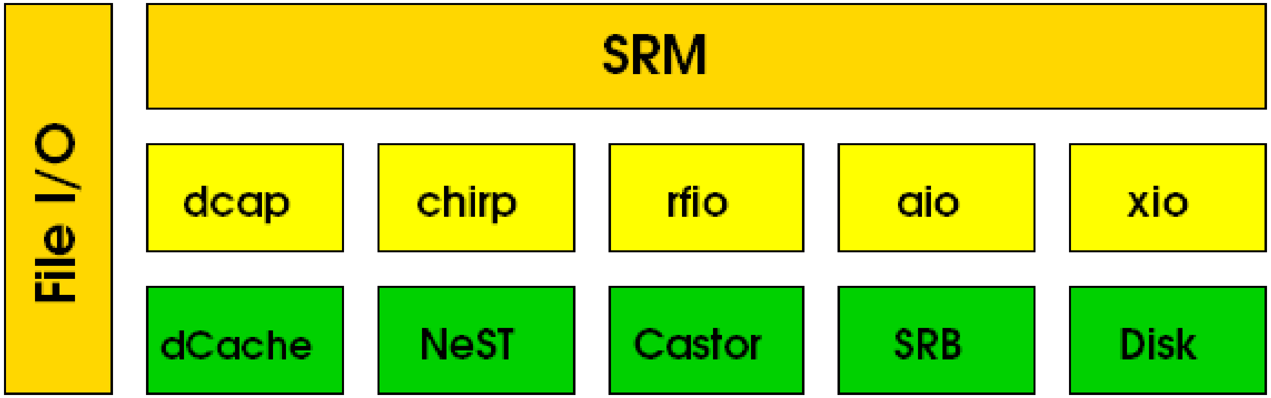
LFC

- **MySQL and ORACLE back-ends**
  - Ensures scalability and allows small scale deployment
  - Read only replication of catalogue (awaiting wider deployment)
- **Multi-threaded C server**
  - Supports multiple instances for load balancing
- **Thread-safe C clients**
  - Python & Perl bindings
  - Command line interface
- **Supports sessions to avoid authentication costs**
  - GSI is very expensive!
- **Bulk methods to reduce the number of round trips**
  - Under test by ATLAS --> 20 times faster
- **Widely used in EGEE:**
  - **largest LFC instance contains 8 millions entries**

DPM - Storage Element

- **Storage Resource Manager (SRM)**
  - Standard that hides the storage system implementation (disk or active tape)
  - handles authorization
  - Web service based on HTTPG
  - translates SURLs (Storage URL) to TURLs (Transfer URLs)
  - disk-based: DPM, dCache, Storm; tape-based: Castor, dCache
  - SRM-2.2
    - Space tokens (manage space by VO/USER), advanced authorization,
    - Better handling of directories, lifetime

- ***File I/O*: *posix-like* access from local nodes or the grid**
  - ➔ GFAL (Grid File Access Layer)



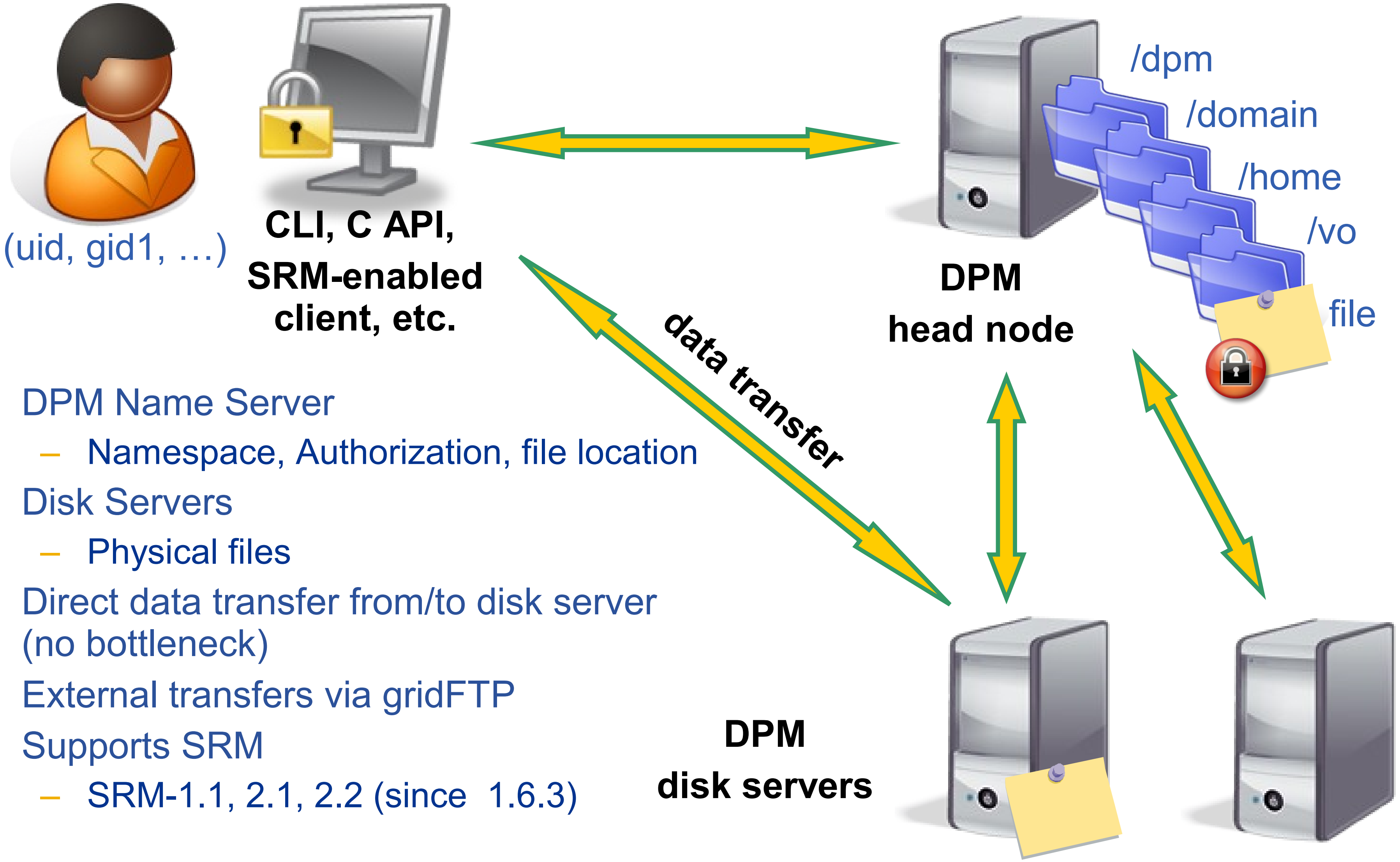
File I/O

SRM

dcap | chirp | rfio | aio | xio

dCache | NeST | Castor | SRB | Disk

DPM: user's point of view



- DPM Name Server
  - Namespace, Authorization, file location
- Disk Servers
  - Physical files
- Direct data transfer from/to disk server (no bottleneck)
- External transfers via gridFTP
- Supports SRM
  - SRM-1.1, 2.1, 2.2 (since 1.6.3)

DPM

Addresses the storage needs of Tier-2 and smaller sites

- Focus on easy setup and maintenance
- Multi-threaded C implementation
- Name server DB
  - Keeps track of the status of files and their physical locations
  - MySQL and ORACLE back ends
    - Simplifies integration in existing local DB infrastructure
    - Ensures scalability
  - Shares code with LFC --> fix once run twice!
- Thread-safe C client and command line interface
  - http/https DPM browser (implemented, very soon to be released)
  - users and site managers interact with DPM at different levels
- GSI and VOMS based authorization and fine grained ACLs
  - Implemented via virtual IDs -> no excessive use of pool accounts
  - Pool access control on VO basis

DPM

Different file access and transfer protocols

- Secure Remote File Input/Output (RFIO)
  - Secure file transfer and manipulation.
  - Implementation of thread-safe C client and a command line interface
  - Support of streaming mode.
- GSIFTP allows remote file transfer
  - New gridftp plugin is implemented to support gridFTP-2
- Xrootd: usable but still limited
  - no support of grid/voms certificates yet
- https/http: web access based on Apache.
  - Protocol http or https can be specified at transfer time

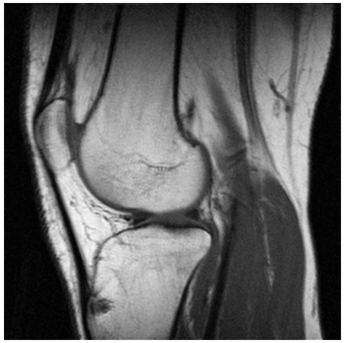
Medical Data Management

Problem : Medical institutes request data storage encryption

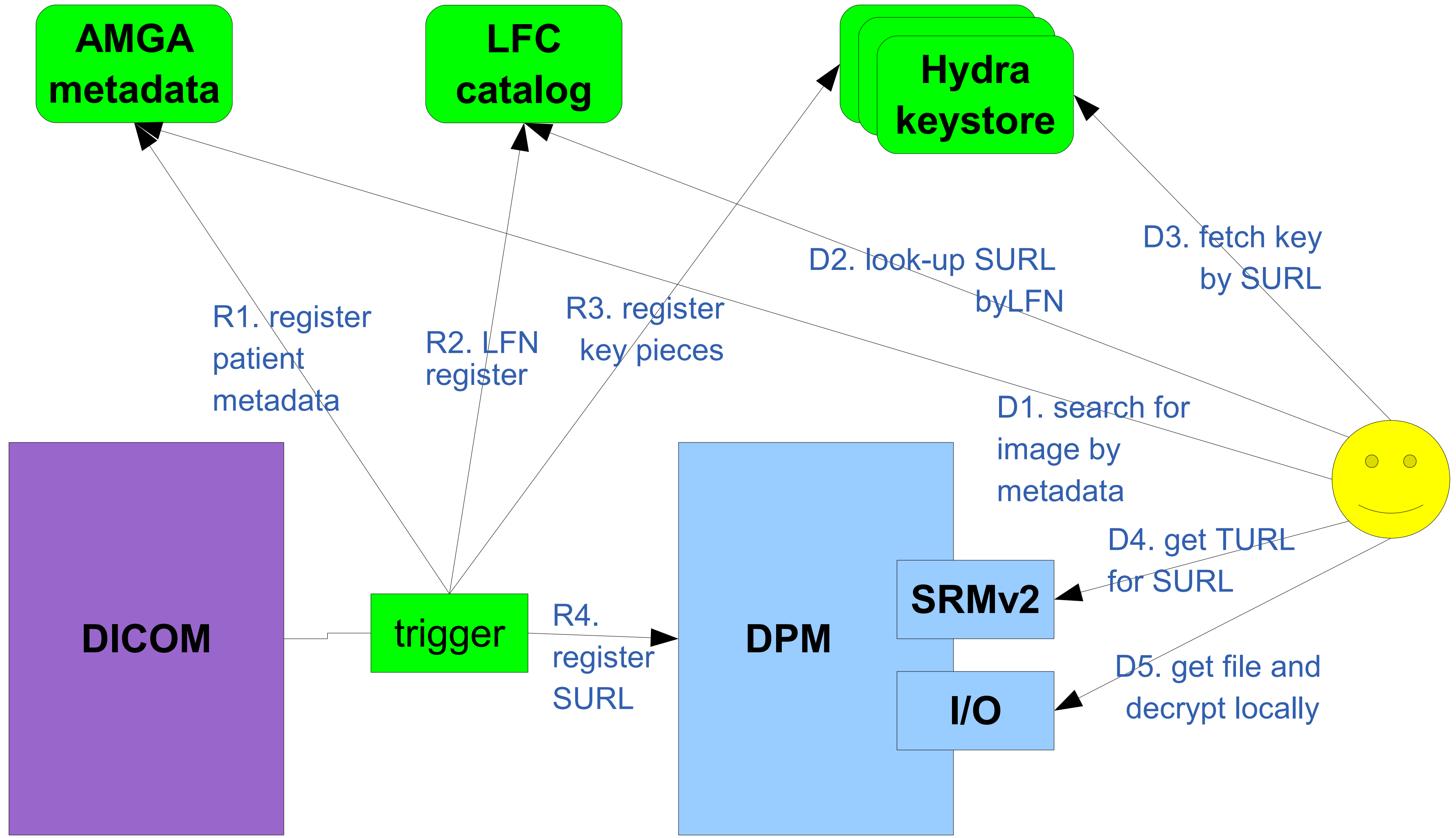
- Use of the DICOM standard for medical image handling
- Image retrieval and storage from/in DICOM servers : security issues

Solution : Extension of the data management tools (under way)

- File encryption on the fly, local decryption
- Use of HYDRA for split key management
- Use of the LFC to register/retrieve system data
  - Replicas location, filesize, …
- Use of srmv2 to get the turls
- Use of I/O protocols, gridftp to load medical images
- Access control based on VOMS



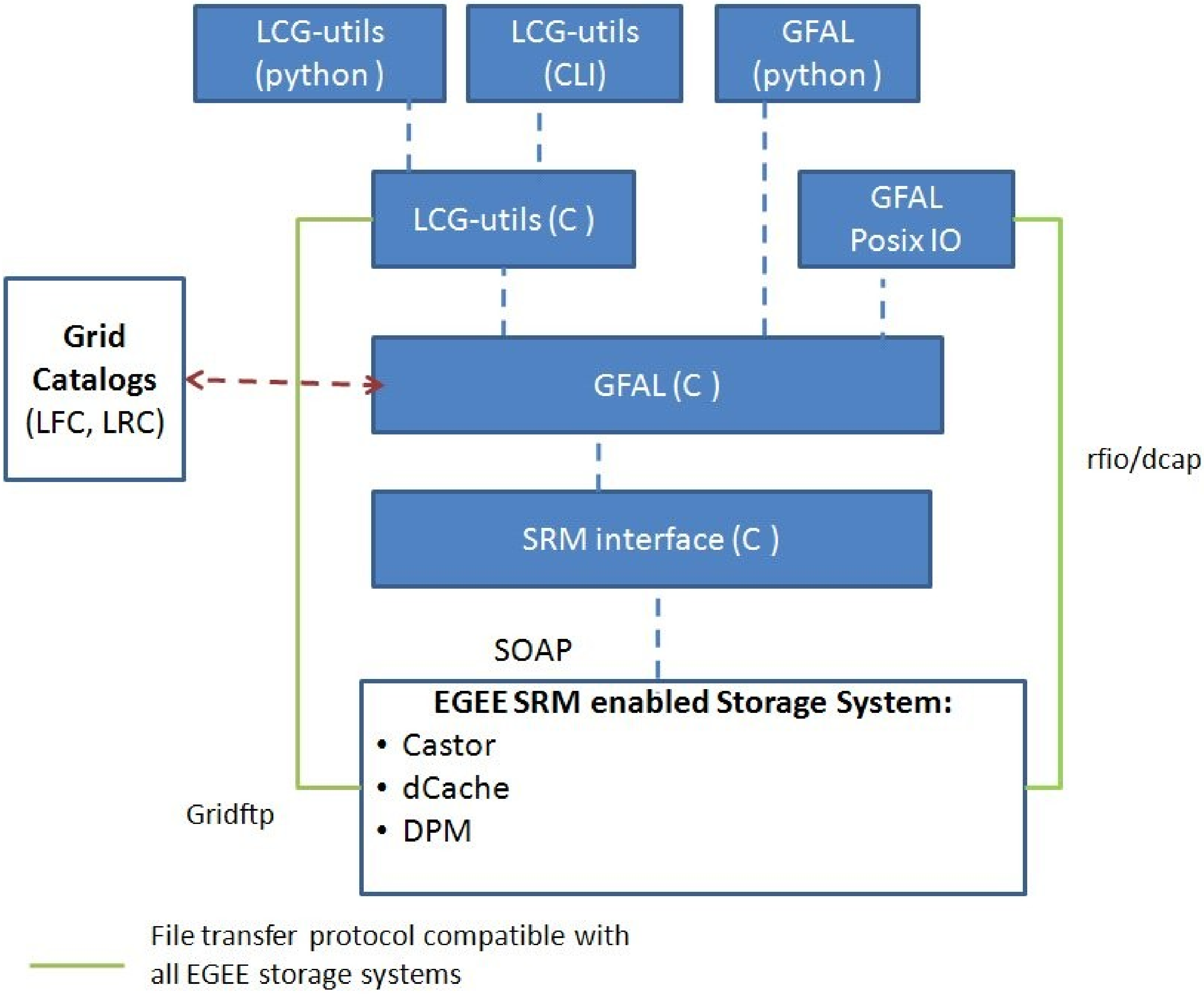
EDS Overview



Lcg-util/gfal

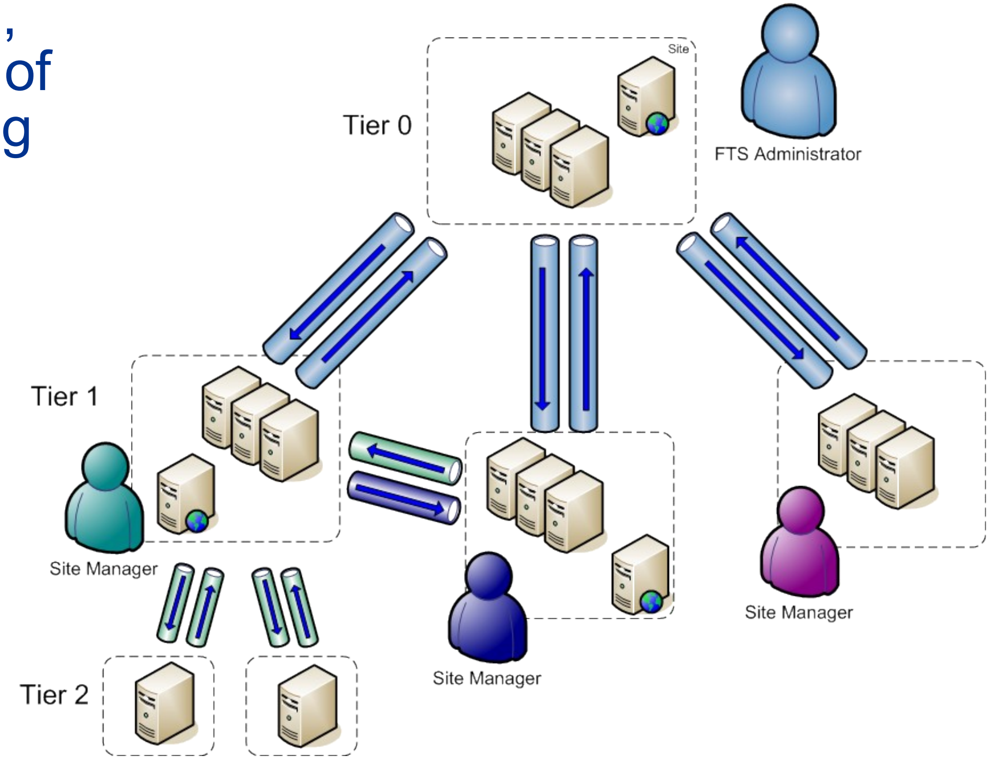
- **Purpose: Create the illusion of POSIX I/O**
  - Shield users from complexity
  - Interact with the information system, catalogue, SRMs
    - Can be used with/without information system/ catalogue
- **LCG-util :**
  - Command line and C-Api
  - Covers most common use cases
    - Replication, catalogue interaction etc,
  - high level tool box
- **Gfal:**
  - Posix like C API for file access
  - SRMv2.2 support
    - user space tokens for retention policy (custodial/replica) & access latency (online/nearline)

Lcg-util/gfal architecture



FTS overview

- **gLite File Transfer Service is a reliable data movement service (batch for file transfers)**
  - FTS performs bulk file transfers between multiple sites
  - Transfers are made between any SRM-compliant storage elements (both SRM 1.1 and 2.2 supported)
  - It is a **multi-VO** service, used to balance usage of site resources according to the SLAs agreed between a site and the VOs it supports
  - VOMS aware



fts
Grid File Transfer Service

FTS

- **Why is it needed ?**
  - For the user, the service it provides is the reliable point to point movement of Storage URLs (SURLs) and ensures you get your share of the sites' resources
  - For the site manager, it provides a reliable and manageable way of serving file movement requests from their VOs and an easy way to discover problems with the overall service delivered to the users
  - For the VO production manager, it provides ability to control requests coming from his users
    - Re-ordering, prioritization,…
  - The focus is on the "**service**" delivered to the user
    - It makes it easy to do these things well with minimal manpower

Is this all?

Definitely NO!!!

- The AMGA meta data catalogue (by Birger Koblitz)
  - Widely used by experiments
  - Is in the process to be integrated in the gLite distribution
- Many data management tools and services developed by Vos

- Lessons learned
  - A DM stack can only be developed with production feedback
  - The right balance between exposing details and hiding is hard to find
  - There will be more to do

Summary

Current status

- Data Management framework is usable
- LFC, FTS, DPM and lcg-util/gfal are used in production on a large scale

Outlook & Future

- ACL synchronization between LFC and SEs
- Improvements to lcg-util/gfal
  - e.g. flexibility to work independently of the LFC
- Better tools to check consistency in DPM
- Extension of Xrootd to support grid/voms certificates
- Finish medical data management implementation
- DPM : quota on pools and accounting
- Operational improvements to the FTS

- Continue the dialog with the user communities to focus effort