Medical Data Management

Ákos Frohner on behalf of the Grid DM Team
CERN – JRA1 All Hands meeting, 2007-10-25

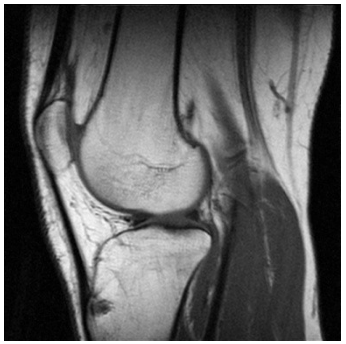
Biomed Requirements

Problem : Medical institutes request data storage encryption

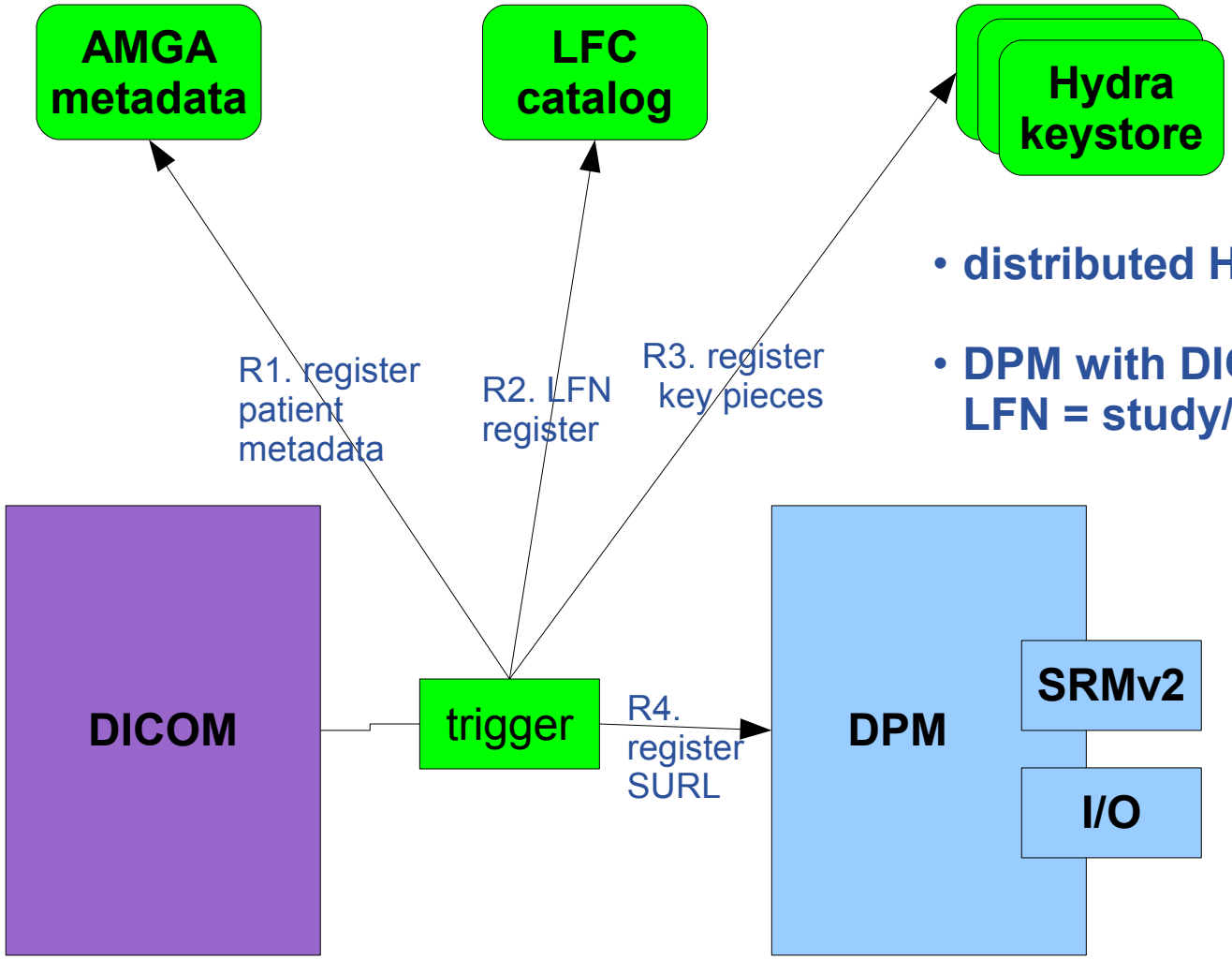
- Use of the DICOM standard for medical image handling
- Image retrieval and storage from/in DICOM servers : security issues

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



Registration



- distributed Hydra keystore
- DPM with DICOM back-end
  LFN = study/series/SOP ID
- back-end is part of a standard DPM service
- encrypted channel to LFC

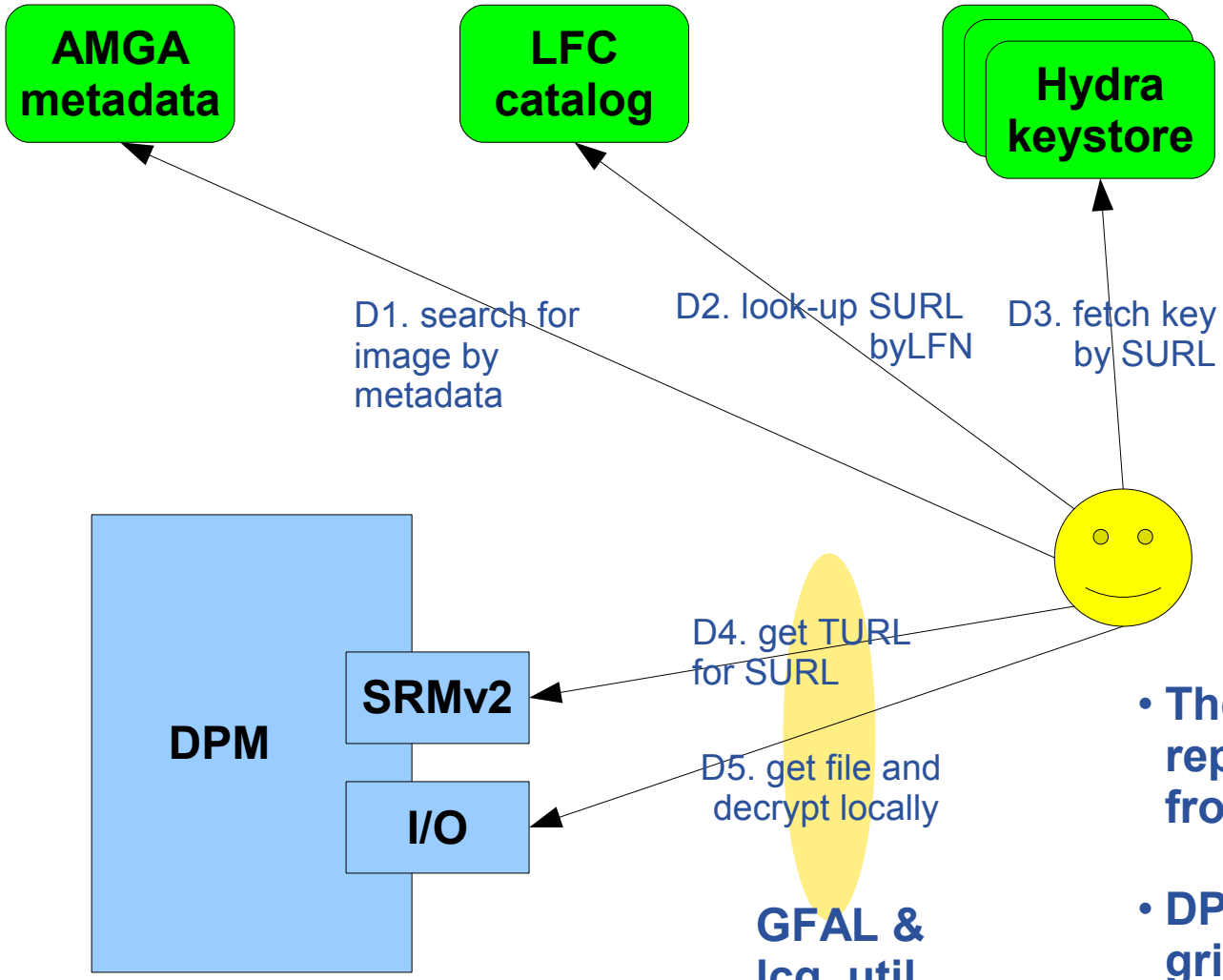
Key splitting

Keys are split for security and reliability reasons using the Shamir's Secrect Sharing Scheme (org.glite.security.ssss)

- **standalone library and CLI**
- **modified Hydra service and Hydra client library/CLI**
- **the client contacts all services for key registration, retrieval and to change permissions**
  - there is no synchronization or transaction coordinator service

```
$ glite-ssss-split-passwd -q 5 3 secret
137c9547aba101ef 6ee7adbbaacac1ef 1256bcc160eda592 fdabc259cdfbacc9
    3113be83f203d794
$ glite-ssss-join-passwd -q 137c9547aba101ef NULL 1256bcc160eda592 NULL
    3113be83f203d794
secret
```

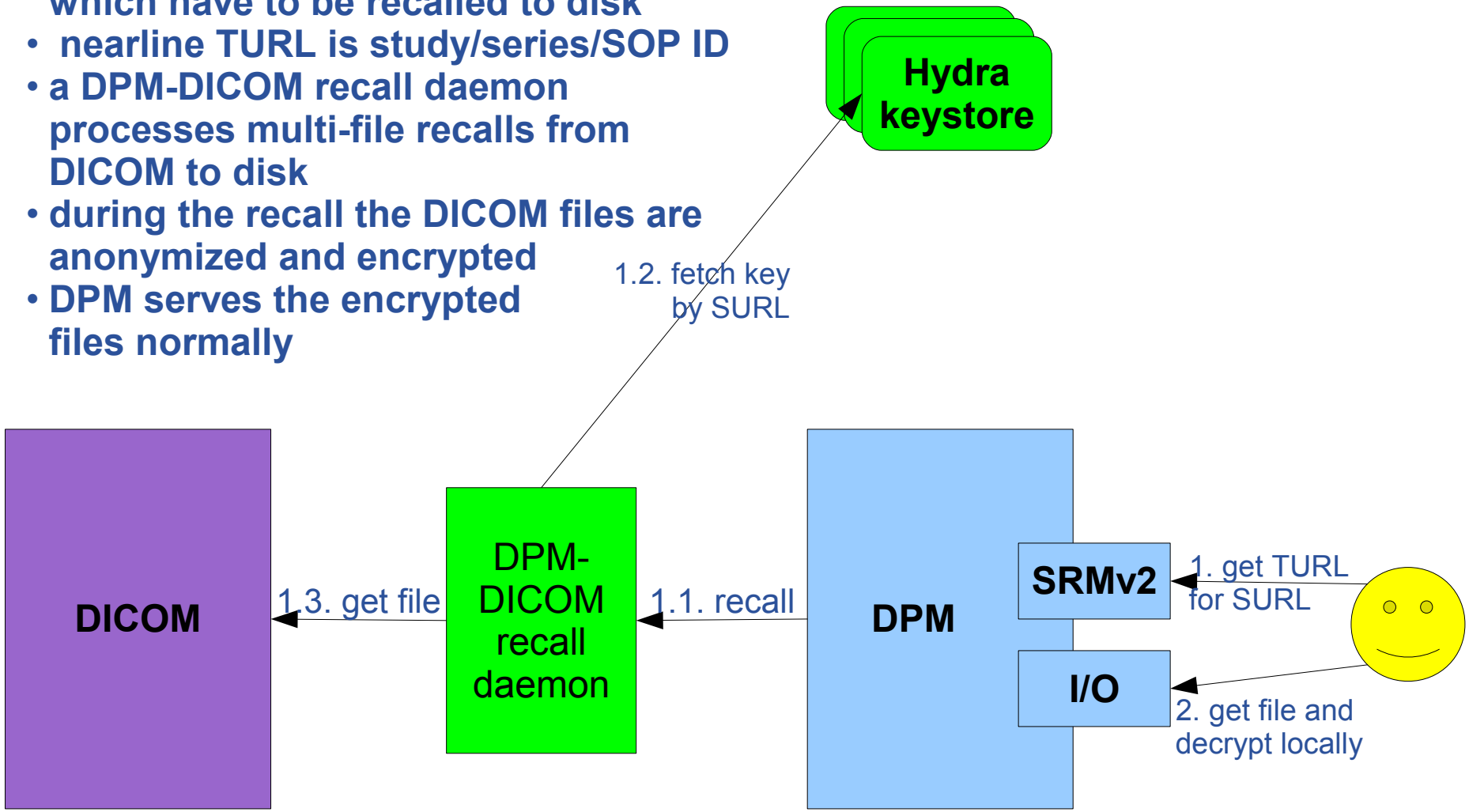
Retrieval



- **The encrypted file can be replicated and retrieved from any other SE.**
- **DPM I/O access via: gridftp, rfio(s), http(s)**

DPM-DICOM interaction

- DICOM files are marked as 'nearline', which have to be recalled to disk
- nearline TURL is study/series/SOP ID
- a DPM-DICOM recall daemon processes multi-file recalls from DICOM to disk
- during the recall the DICOM files are anonymized and encrypted
- DPM serves the encrypted files normally



Status and Plans

- DPM-DICOM recall daemon is coded, being tested
- DICOM image anonymization and encryption ready
- Hydra key splitting is ready
- Hydra/GFAL CLI is coded, being tested
- LFC encryption is coded, needs integration and tests
- Still need to integrate and test all these together