Architecture of LHC File Catalog

Valeria Ardizzone

INFN Catania – EGEE-II NA3/NA4

Health e-Child Application Developer Tutorial
Cern(Geneve), September 25-26, 2006

Introduction

- **Assumptions:**
  - Users and programs produce and require data
  - the lowest granularity of the data is on the file level (we deal with files rather than data objects or tables)
    - Data = files

- **Files:**
  - Mostly, write once, read many
  - Located in Storage Elements (SEs)
  - Several replicas of one file in different sites
  - Accessible by Grid users and applications from “anywhere”

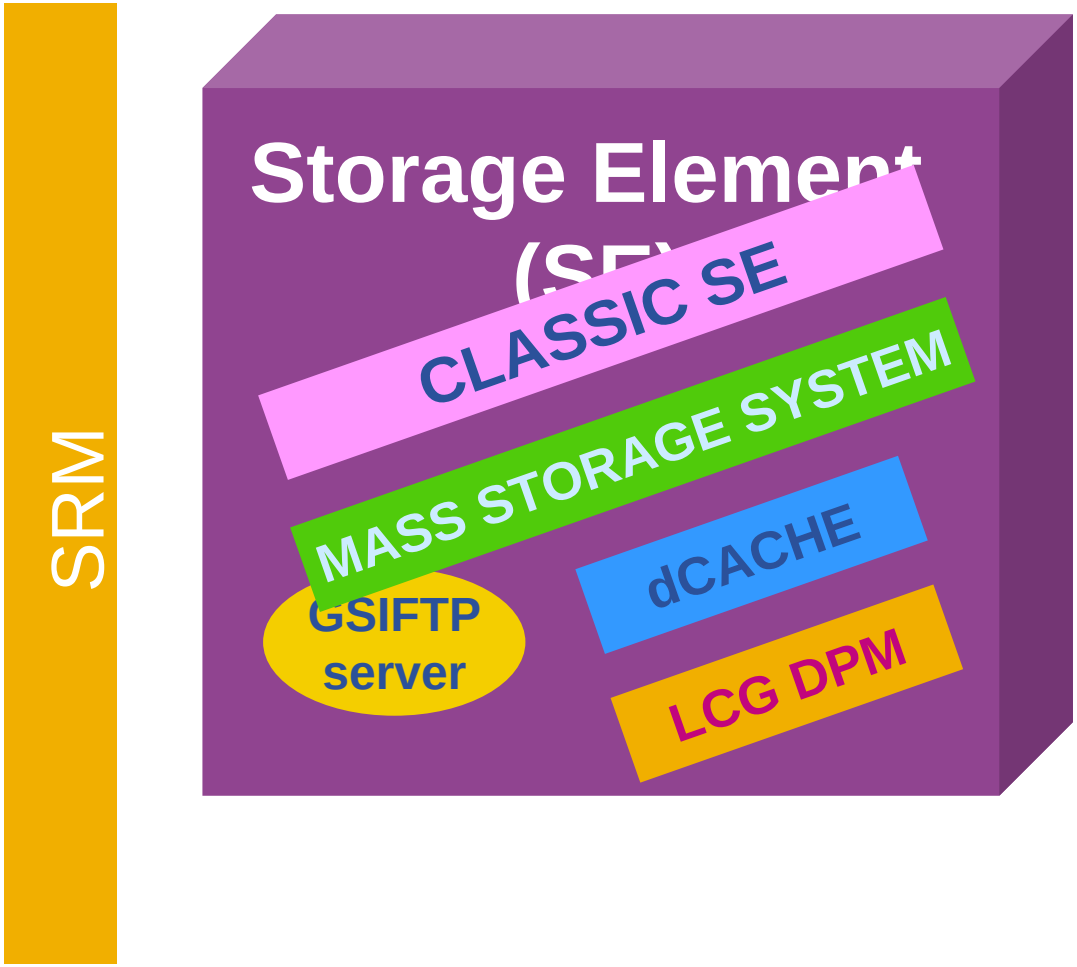
- **Also…**
  - Files may be copied from/to local filesystems (WNs, UIs) to the Grid (SEs)

Storage Element (SE)

- **Def: The Storage Element is the service which allows a user or an application to store data for future retrieval.**
- **User is responsible to manage the available space in the SE.**
- **gLite3.0 support basic file transfer protocols**
  - GridFTP mandatory (GSI enabled FTP)
  - Others if available (https, ftp, etc)

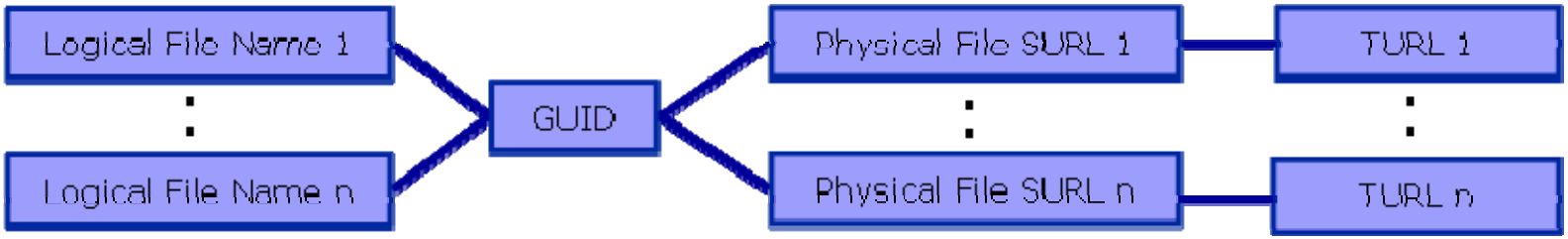
Storage Resource Management (SRM)

- SRM has been designed to be a single interface for the management of disk and tape storage resource.
- SRM is a storage management protocol, no file access or file transport one.

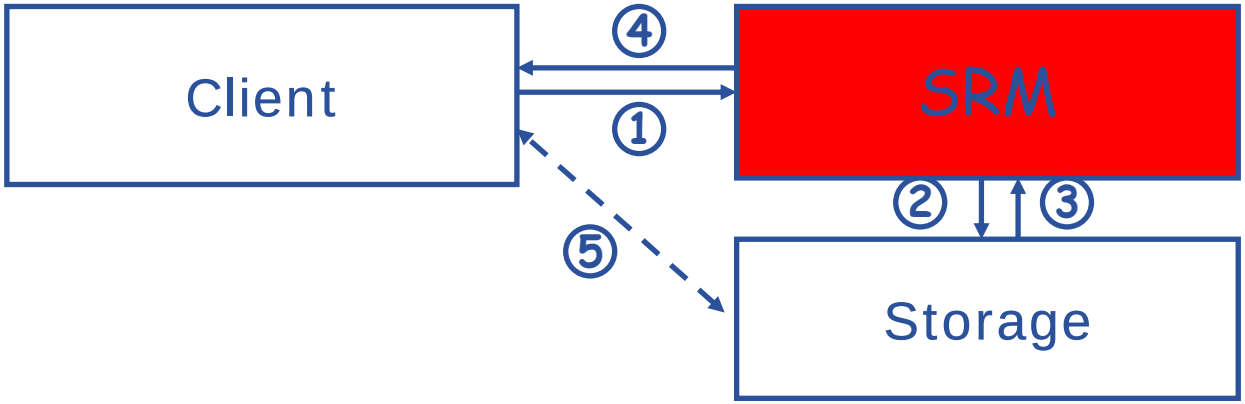


Files Naming conventions

- **Logical File Name (LFN)**
  - An alias created by a user to refer to some item of data, e.g. "lfn:/grid/gilda/tony/simple2.dat"
- **Globally Unique Identifier (GUID)**
  - A non-human-readable unique identifier for an item of data, e.g. "guid:3a69a819-2023-4400-a2a1-f581ab942044"
- **Site URL (SURL)**
  - Gives indication on which place (Storage Element) the file is actually found.
  - Understood by the SRM interface
  - "srm://aliserv6.ct.infn.it/dpm/ct.infn.it/home/gilda/generated/2006-07-10/filef7a916f7-159b-48df-9159-877f2d3c6f58"
- **Transport URL (TURL)**
  - Temporary locator of a replica+access protocol: understood by the backend MSS "gsiftp://aliserv6.ct.infn.it/aliserv6.ct.infn.it:/gpfs/dpm/gilda/2006-07-10/filef7a916f7-159b-48df-9159-877f2d3c6f58.46193.0"



SRM Interactions

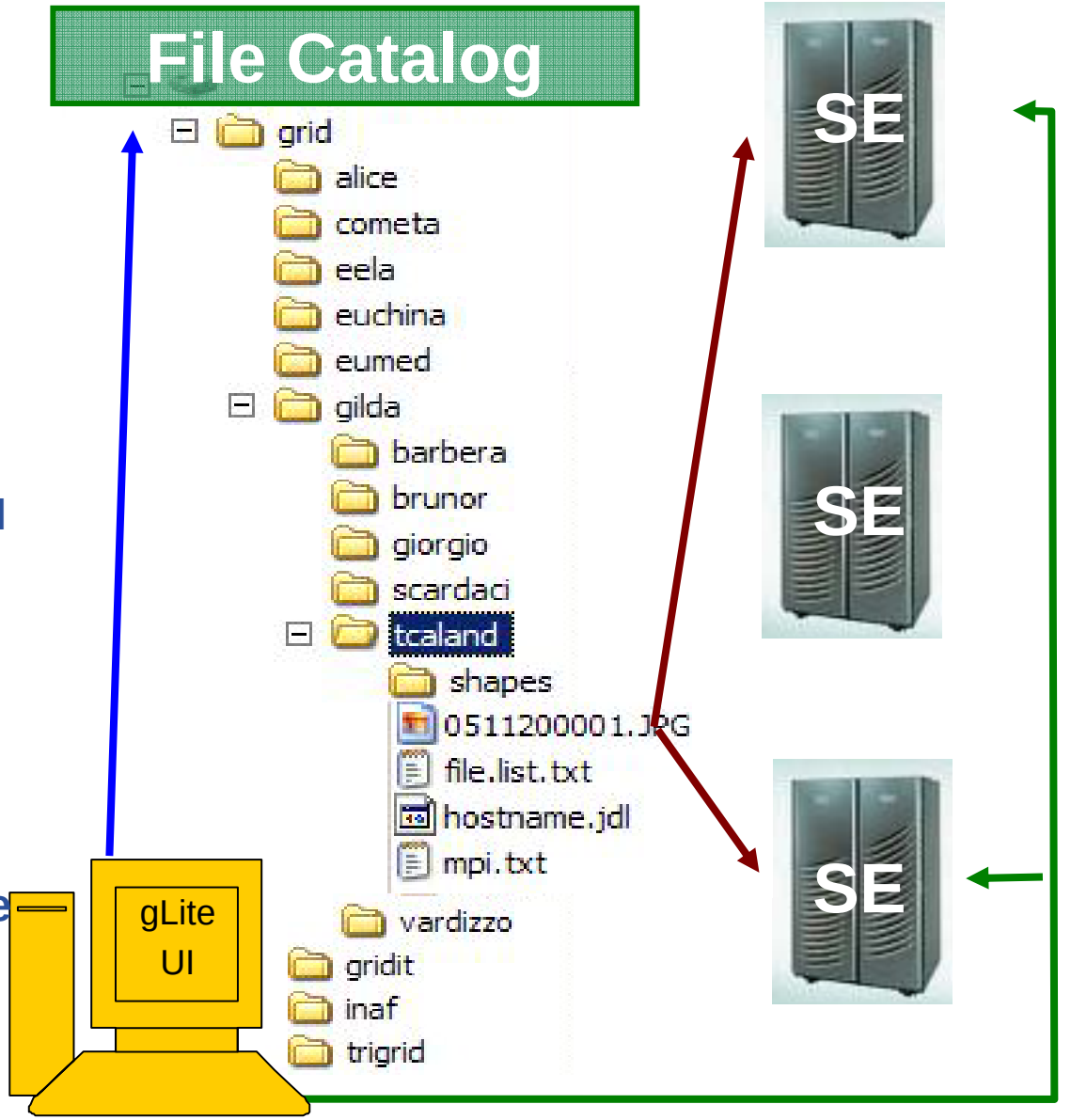


1. The client asks the SRM for a file providing an SURL (Site URL)
2. The SRM asks the storage system to provide the file
3. The storage system notifies the availability of the file and its location
4. The SRM returns a TURL (Transfer URL), i.e. the location from where the file can be accessed
5. The client interacts with the storage using the protocol specified in the TURL

What is a file catalog

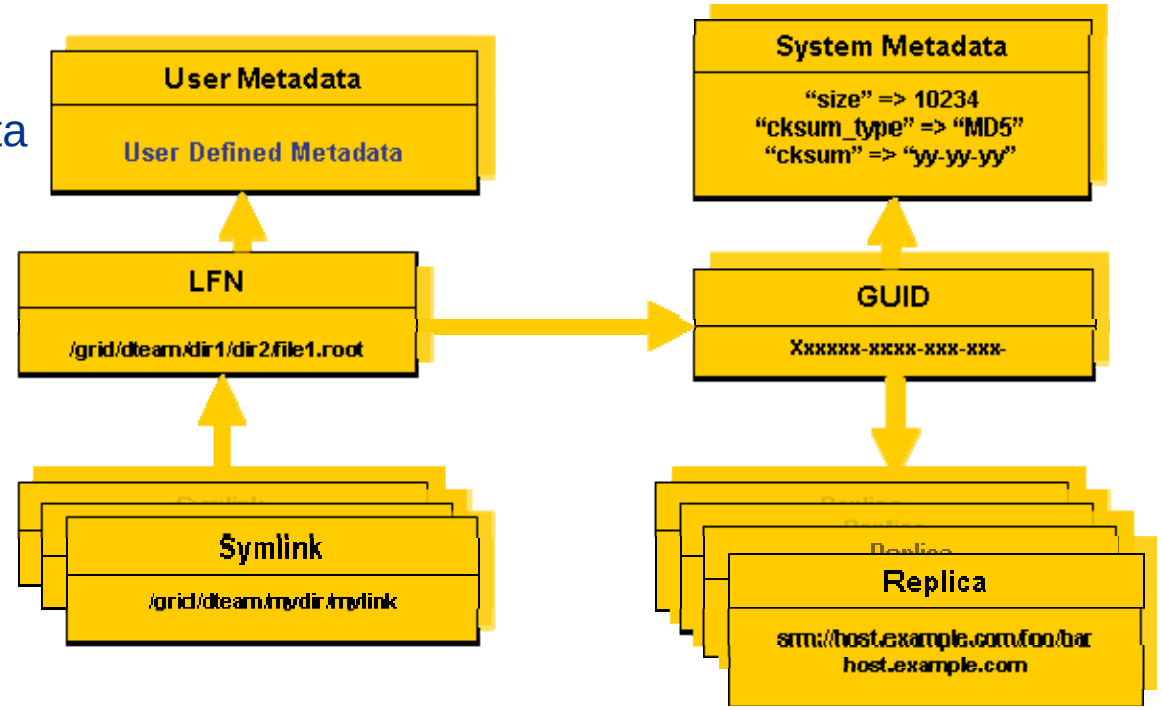
LFC Features:

- **Mantaining mappings between LFN(s),GUID and SURL(s).**
- **Best performance and less security problem than old RLS.**
- **It supports transactions, roll-back, sessionsm bulk queries.**
- **It is a unique catalog where LFN is the main key.**
- **Hierarchiacal name-space for LFNs.**
- **Symlink to main LFN.**
- **System metadata.**
- **User metadata (but only a single string entry).**



The LFC (LCG File Catalog)

- **It keeps track of the location of copies (replicas) of Grid files**
- **LFN acts as main key in the database. It has:**
  - Symbolic links to it (additional LFNs)
  - Unique Identifier (GUID)
  - System metadata
  - Information on replicas
  - One field of user metadata



LFC commands

Summary of the LFC Catalog commands

lfc-chmod	Change access mode of the LFC file/directory
lfc-chown	Change owner and group of the LFC file-directory
lfc-delcomment	Delete the comment associated with the file/directory
lfc-getacl	Get file/directory access control lists
lfc-ln	Make a symbolic link to a file/directory
lfc-ls	List file/directory entries in a directory
lfc-mkdir	Create a directory
lfc-rename	Rename a file/directory
lfc-rm	Remove a file/directory
lfc-setacl	Set file/directory access control lists
lfc-setcomment	Add/replace a comment

References

- VOMS and ACLs in Data Management
  https://twiki.cern.ch/twiki/bin/view/LCG/VomsAndAcls
- GILDA -- Data Management quickstart
  https://grid.ct.infn.it/twiki/bin/view/GILDA/DataManagement
- SEE-GRID2 LFC JAVA API
  http://wiki.egee-see.org/index.php/SEE-GRID_File_Management_Java_API
- LFC Documentation
  https://twiki.cern.ch/twiki/bin/view/LCG/DataManagementDocumentation
- EGEE Middleware Support
  https://twiki.cern.ch/twiki/bin/view/EGEE/EGEEMiddlewareSupport

Introduction to the LFC Java API practicals

- **Browse to the agenda page:**
  http://agenda.cern.ch/fullAgenda.php?ida=a063217

- **Follow the link "more information" on the topic "Practicals on LFC API usage"**