

# gLite Data Management System Architecture

**Antonio Calanducci**

**[antonio.calanducci@ct.infn.it](mailto:antonio.calanducci@ct.infn.it)**

National Institute of Nuclear Physics

INFN Catania

EGEE NA3 Training & Dissemination

**CyGrid - Training the trainers event**

**Nicosia (Cyprus), University of Cyprus, 06/05/2009**

- **Grid Data Management Challenge**
- **Storage Elements and SRM**
- **File Catalogs and DM tools**
- **Metadata Service**

- **Heterogeneity**

- Data are stored on different storage systems using different access technologies

- **Distribution**

- Data are stored in different locations – in most cases there is no shared file system or common namespace
- Data need to be moved between different locations

- **Data description**

- Data are stored as files: need a way to describe files and locate them according to their contents

- Need common interface to storage resources
  - Storage Resource Manager (SRM)
- Need to keep track where data is stored
  - File and Replica Catalogs
- Need scheduled, reliable file transfer
  - File transfer service
- Need a way to describe files' content and query them
  - Metadata service

- **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”
  - Locatable by the WMS (data requirements in JDL)
  
- **Also...**
  - WMS can send (small amounts of) data to/from jobs: Input and Output Sandbox
  - Files may be copied from/to local filesystems (WNs, UIs) to the Grid (SEs)

- The **Storage Element** is the service which allow a user or an application to store data for future retrieval
- **Manage local storage (disks) and interface to Mass Storage Systems(tapes) like**
  - HPSS, CASTOR, DiskeXtender (UNITREE), ...
- **Be able to manage different storage systems uniformly and transparently for the user (providing an SRM interface)**
- **Support basic file transfer protocols**
  - GridFTP mandatory
  - Others if available (https, ftp, etc)
- **Support a native I/O (remote file) access protocol**
  - POSIX (like) I/O client library for direct access of data (GFAL)

She is running a job which needs:  
Data for physics event reconstruction  
Simulated Data  
Some data analysis files  
She will write files remotely too

They are at CERN  
In dCache

They are at Fermilab  
In a disk array

They are at Nikhef  
in a classic SE



## dCache

Own system, own protocols and parameters

## gLite DPM

Independent system from dCache or Castor

## Castor

No connection with dCache or DPM

You as a user need to know all the systems!!!

## dCache

Own system, own protocols and parameters

## gLite DPM

Independent system from dCache or Castor

## Castor

No connection with dCache or DPM

SRM

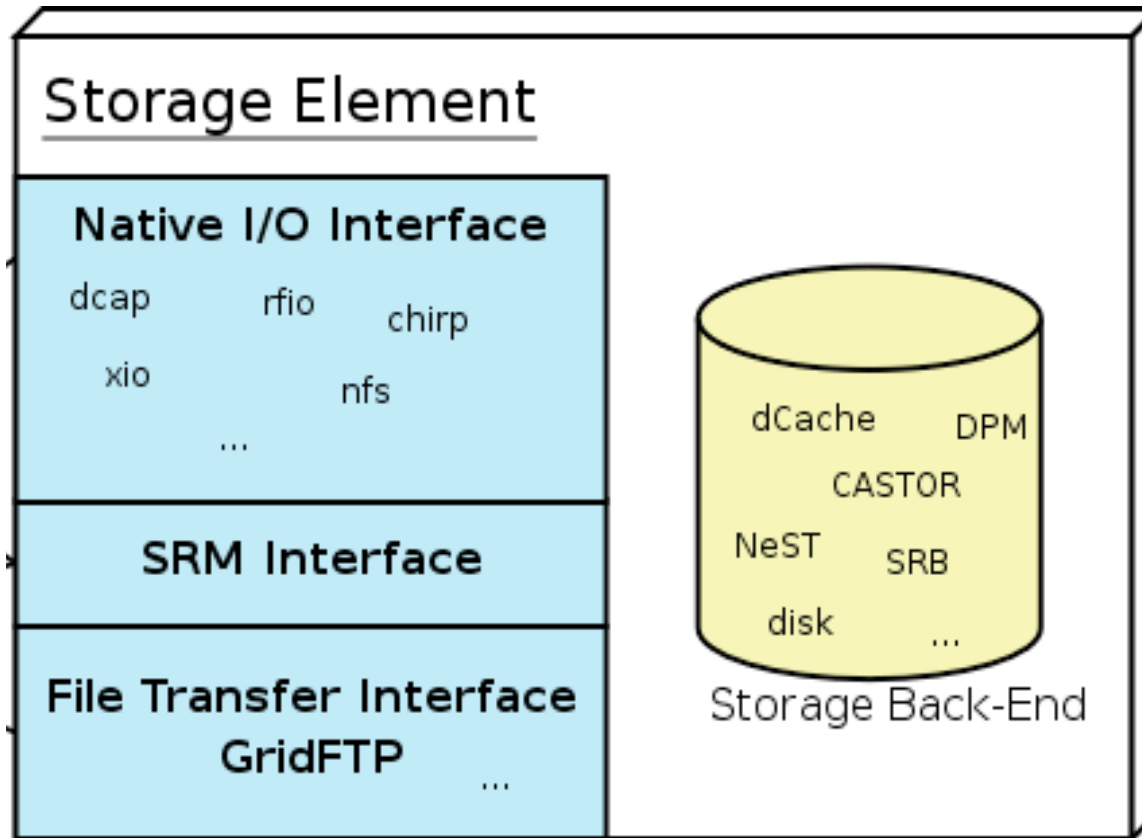
I talk to them on your behalf

I will even allocate space for your files

And I will use transfer protocols to send your files there



- **Data are stored on disk pool servers or Mass Storage Systems**
- storage resource management needs to take into account
  - Transparent access to files (migration to/from disk pool)
  - File pinning
  - Space reservation
  - File status notification
  - Life time management
- **The SRM (Storage Resource Manager) takes care of all these details**
  - The SRM is a single interface that takes care of local storage interaction and provides a Grid
- In gLite, interactions with the SRM is hidden by higher level services (DM tools and APIs)

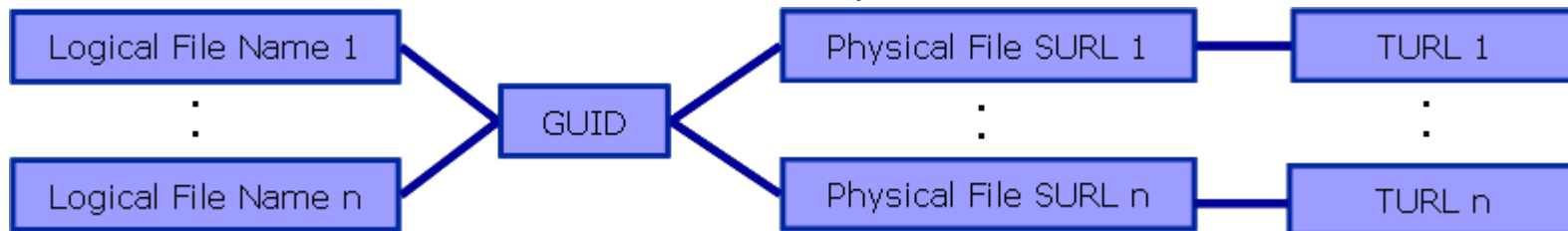


- **Logical File Name (LFN)**
  - An alias created by a user to refer to some item of data, e.g. “lfn:/grid/gilda/20030203/run2/track1”
- **Globally Unique Identifier (GUID)**
  - A non-human-readable unique identifier for an item of data, e.g. “guid:f81d4fae-7dec-11d0-a765-00a0c91e6bf6”
- **Site URL (SURL) (or Physical File Name (PFN) or Site FN)**
  - The location of an actual piece of data on a storage system

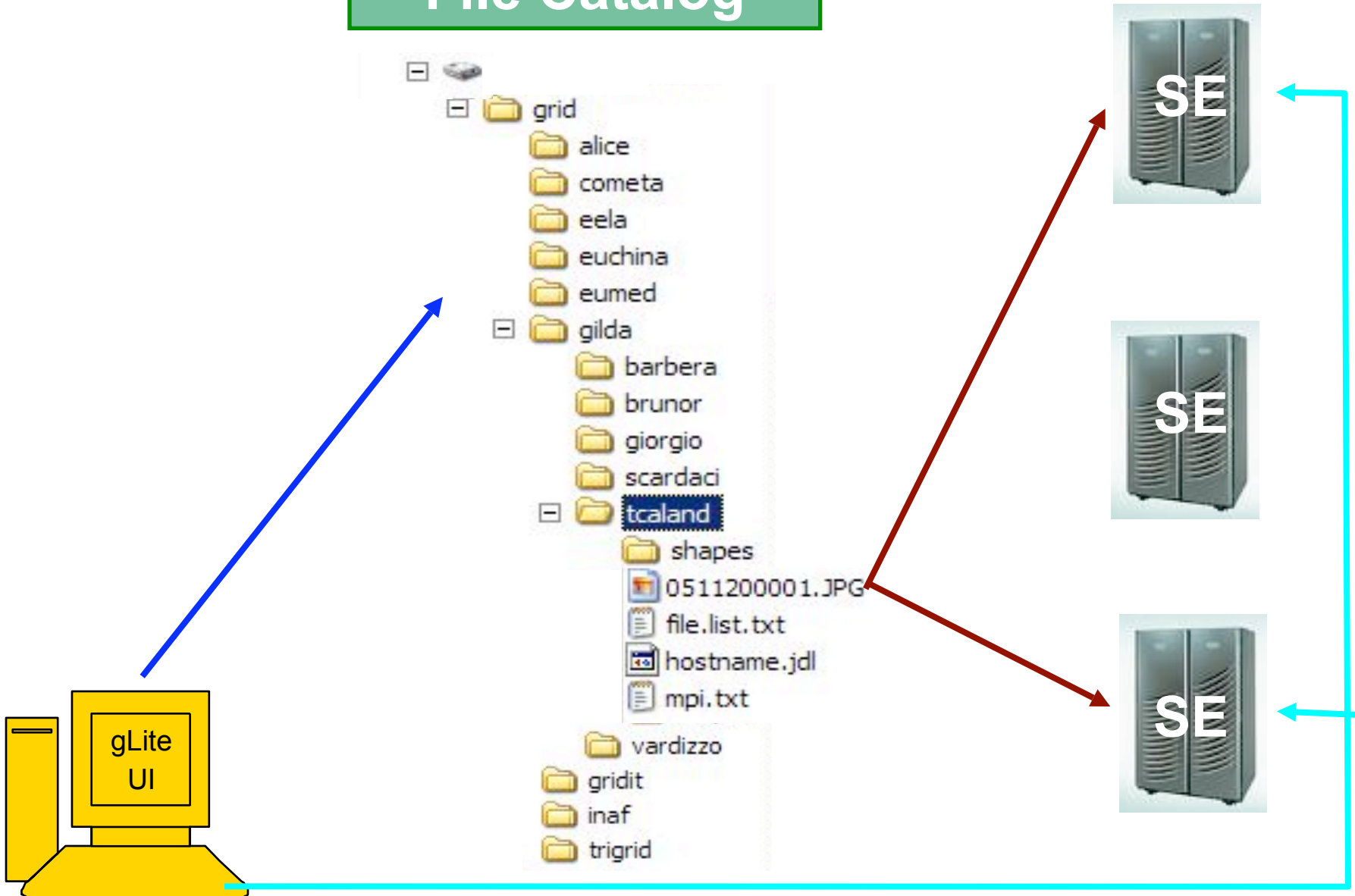
e.g. “srm://grid009.ct.infn.it/dpm/ct.infn.it/gilda/output10\_1” (SRM) “sfn://lxshare0209.cern.ch/data/alice/ntuples.dat” (Classic SE)

- **Transport URL (TURL)**
  - Temporary locator of a replica + access protocol: understood by a SE, e.g.

“rfio://lxshare0209.cern.ch//data/alice/ntuples.dat”



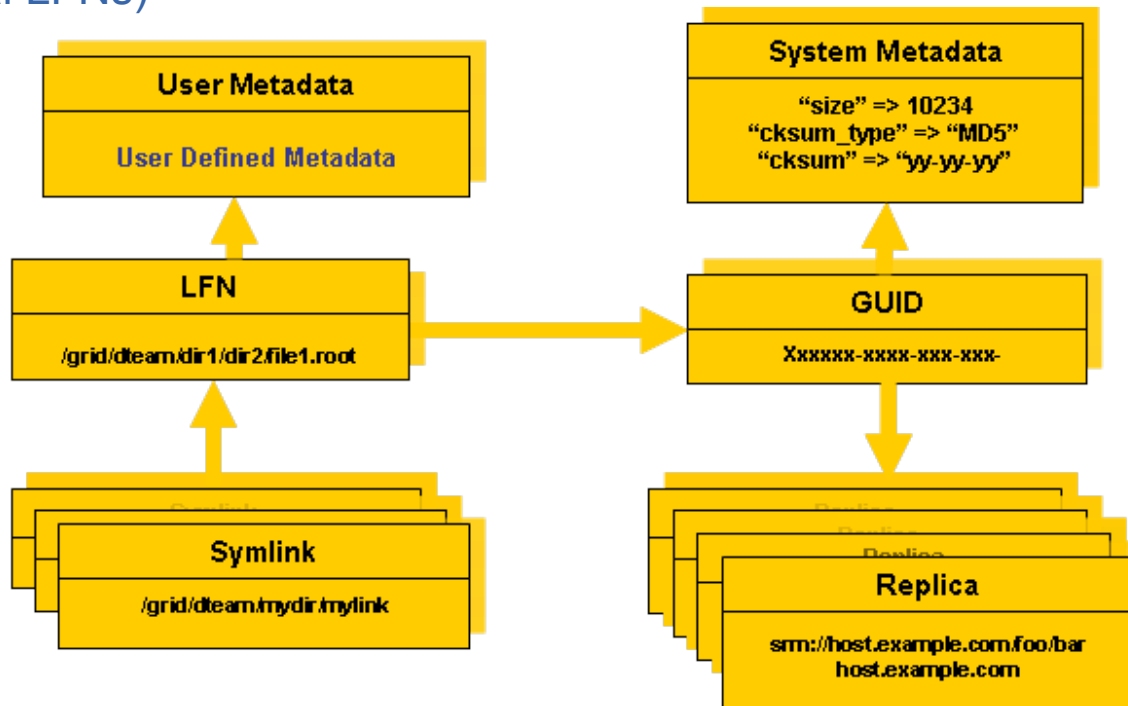
## File Catalog



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



- Cursors for large queries
- Timeouts and retries from the client
- User exposed transactional API (+ auto rollback on failure)
- **Hierarchical namespace** and namespace operations (for LFNs)
- Integrated GSI Authentication + Authorization
- Access Control Lists (Unix Permissions and POSIX ACLs)
- Checksums
- Integration with VOMS (VirtualID and VirtualGID)

## Listing the entries of a LFC directory

```
lfc-ls [-cdiLIRTu] [--class] [--comment] [--deleted] [--display_side] [--ds]
  path...
```

where *path* specifies the LFN pathname (mandatory)

- Remember that **LFC has a directory tree structure**
- **/grid/<VO\_name>/<you create it>**



- All members of a VO have read-write permissions under their directory
- You can set LFC\_HOME to use relative paths
  - > *lfc-ls /grid/gilda/tony*
  - > *export LFC\_HOME=/grid/gilda*
  - > *lfc-ls -l tony*
  - > *lfc-ls -l -R /grid*

-l : long listing  
 -R : list the contents of directories recursively: **Don't use it!**

## Creating directories in the LFC

***lfc-mkdir [-m mode] [-p] path...***

- Where *path* specifies the LFC pathname
- Remember that while registering a new file (using lcg-cr, for example) the corresponding destination directory must be created in the catalog beforehand.
- Examples:
  - > ***lfc-mkdir /grid/gilda/tony/demo***

You can just check the directory with:

> **lfc-ls -l /grid/gilda/tony**

```
drwxr-xrwx  0 19122  1077
```

```
0 Jun 14 11:36 demo
```



## Creating a symbolic link

***lfc-ln -s file linkname***

***lfc-ln -s directory linkname***

Create a link to the specified *file* or *directory* with *linkname*

– *Examples:*

**> *lfc-ln -s /grid/gilda/tony/demo/test /grid/gilda/tony/aLink***



Let's check the link using *lfc-ls* with long listing (-l):

**> *lfc-ls -l***

```
lrwxrwxrwx 1 19122 1077 0 Jun 14 11:58 aLink ->/grid/gilda/tony/demo/test
```

```
drwxr-xrwx 1 19122 1077 0 Jun 14 11:39 demo
```

## Summary of the LFC Catalog commands

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

## Low level methods (many POSIX-like):

lfc_access	lfc_deleteclass	lfc_listreplica	lfc_setacl
lfc_aborttrans	lfc_delreplica	lfc_lstat	lfc_setatime
lfc_addreplica	lfc_endtrans	lfc_mkdir	lfc_setcomment
lfc_apiinit	lfc_enterclass	lfc_modifyclass	lfc_seterrbuf
lfc_chclass	lfc_errmsg	s	lfc_setfsz
lfc_chdir	lfc_getacl	lfc_opendir	lfc_starttrans
lfc_chmod	lfc_getcomment	lfc_queryclass	lfc_stat
lfc_chown	t	lfc_readdir	lfc_symlink
lfc_closedir	lfc_getcwd	lfc_readlink	lfc_umask
lfc_creat	lfc_getpath	lfc_rename	lfc_undelete
lfc_delcomment	lfc_lchown	lfc_rewind	lfc_unlink
lfc_delete	lfc_listclass	lfc_rmdir	lfc_utime
	lfc_listlinks	lfc_selectsrvr	send2lfc

## Interactions with SE require some components:

- File catalog services to locate replicas
- SRM
- File access mechanism to access files from the SE on the WN

## GFAL does all this tasks for you:

- Hides all these operations
- Presents a POSIX interface for the I/O operations

- Single shared library in threaded and unthreaded versions

libgfal.so, libgfal\_pthr.so

- Single header file: **gfal\_api.h**

- User can create all commands needed for storage management
- It offers as well an interface to SRM

## Supported protocols:

- file (local or nfs-like access)
- dcap, gsidcap and kdcap (dCache access)
- rfio (castor access) and gsirfio (dpm)

```
int gfal_access (const char *path, int amode);
int gfal_chmod (const char *path, mode_t mode);
int gfal_close (int fd);
int gfal_creat (const char *filename, mode_t mode);
off_t gfal_lseek (int fd, off_t offset, int whence);
int gfal_open (const char * filename, int flags, mode_t mode);
ssize_t gfal_read (int fd, void *buf, size_t size);
int gfal_rename (const char *old_name, const char *new_name);
ssize_t gfal_setfilchg (int, const void *, size_t);
int gfal_stat (const char *filename, struct stat *statbuf);
int gfal_unlink (const char *filename);
ssize_t gfal_write (int fd, const void *buf, size_t size);
```

```
int gfal_closedir (DIR *dirp);
```

```
int gfal_mkdir (const char *dirname, mode_t mode);
```

```
DIR *gfal_opendir (const char *dirname);
```

```
struct dirent *gfal_readdir (DIR *dirp);
```

```
int gfal_rmdir (const char *dirname);
```

int **create\_alias** (const char \*guid, const char \*lfn, long long size)

int **guid\_exists** (const char \*guid)

char \***guidforpfn** (const char \*surl)

char \***guidfromlfn** (const char \*lfn)

char \*\***lfnforguid** (const char \*guid)

int **register\_alias** (const char \*guid, const char \*lfn)

int **register\_pfn** (const char \*guid, const char \*surl)

int **setfilesize** (const char \*surl, long long size)

char \***surlfromguid** (const char \*guid)

char \*\***surlsfromguid** (const char \*guid)

int **unregister\_alias** (const char \*guid, const char \*lfn)

int **unregister\_pfn** (const char \*guid, const char \*surl)

int **deletesurl** (const char \*surl)

int **getfilemd** (const char \*surl, struct stat64 \*statbuf)

int **set\_xfer\_done** (const char \*surl, int reqid, int fileid, char \*token, int oflag)

int **set\_xfer\_running** (const char \*surl, int reqid, int fileid, char \*token)

char \***turlfromsurl** (const char \*surl, char \*\*protocols, int oflag, int \*reqid, int \*fileid, char \*\*token)

int **srm\_get** (int nbfiles, char \*\*surls, int nbprotocols, char \*\*protocols, int \*reqid, char \*\*token, struct srm\_filestatus \*\*filestatuses)

int **srm\_getstatus** (int nbfiles, char \*\*surls, int reqid, char \*token, struct srm\_filestatus \*\*filestatuses)



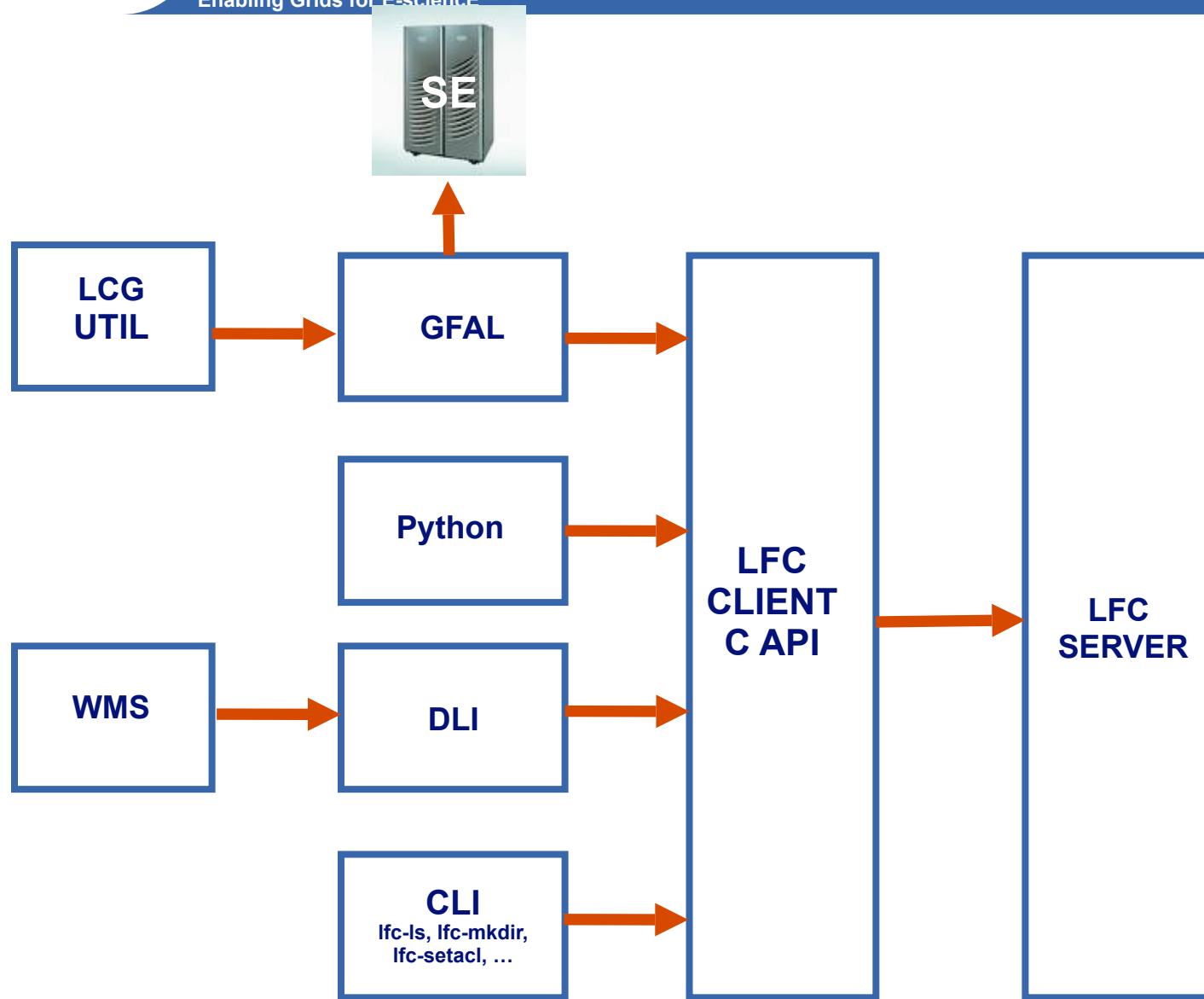
- **High level interface (CL tools and APIs) to**
  - Upload/download files to/from the Grid (UI,CE and WN <----> SEs)
  - Replicate data between SEs and locate the best replica available
  - Interact with the file catalog
- ***Definition:* A file is considered to be a **Grid File** if it is both physically present in a SE and registered in the File Catalog**
- **lcg-utils ensure the consistency between files in the Storage Elements and entries in the File Catalog**

## Replica Management

<b>lcg-cp</b>	<b>Copies a grid file to a local destination</b>
<b>lcg-cr</b>	<b>Copies a file to a SE and registers the file in the catalog</b>
<b>lcg-del</b>	<b>Delete one file</b>
<b>lcg-rep</b>	<b>Replication between SEs and registration of the replica</b>
<b>lcg-gt</b>	<b>Gets the TURL for a given SURL and transfer protocol</b>
<b>lcg-sd</b>	<b>Sets file status to “Done” for a given SURL in a SRM request</b>

## File Catalog Interaction

<b>lcg-aa</b>	<b>Add an alias in LFC for a given GUID</b>
<b>lcg-ra</b>	<b>Remove an alias in LFC for a given GUID</b>
<b>lcg-rf</b>	<b>Registers in LFC a file placed in a SE</b>
<b>lcg-uf</b>	<b>Unregisters in LFC a file placed in a SE</b>
<b>lcg-la</b>	<b>Lists the alias for a given SURL, GUID or LFN</b>
<b>lcg-lg</b>	<b>Get the GUID for a given LFN or SURL</b>
<b>lcg-lr</b>	<b>Lists the replicas for a given GUID, SURL or LFN</b>



- Grids allow to save **millions of files** spread over several storage sites.
- Users and applications need an efficient mechanism
  - to **describe** files
  - to **locate** files based on their contents
- **This is achieved by**
  - associating descriptive attributes to files
    - Metadata is **data about data**
  - answering user queries against the associated information

- **Entries** – Representation of real world entities which we are attaching metadata to for describing them
- **Attribute** – key/value pair
  - **Type** – The type (int, float, string,...)
  - **Name/Key** – The name of the attribute
  - **Value** - Value of an entry's attribute
- **Schema** – A set of attributes
- **Collection** – A set of entries associated with a schema
- **Metadata** - List of attributes (including their values) associated with entries

- **Movie trailers files (entries) saved on Grid Storage Elements and registered into File Catalogue**
- **We want to add metadata to describe movie content.**
- **A possible schema:**
  - Title -- varchar
  - Runtime -- int
  - Cast -- varchar
  - LFN -- varchar
- **A metadata catalogue will be the repository of the movies' metadata and will allow to find movies satisfying users' queries**

<b><i>Entry names</i></b>	<b>Title</b>	<b>Ru</b>	<b>Cast</b>	<b>LFN</b>
8c3315c1-811f-4823-a778-60a203439689	My Best Friend's	80	Julia Roberts	lfn:/grid/gilda/movies/mybfwed.avi
51a18b7a-fd21-4b2c-aa74-4c53ee64846a	Spider-man 2	120	Kirsten Dunst	lfn:/grid/gilda/movies/spiderman2.avi
401e6df4-c1be-4822-958c-ce3eb5c54fcb	The God Father	113	Al pacino	lfn:/grid/gilda/movies/godfather.avi

## Attribute



### Entry names

	Title	Ru	Cast	LFN
8c3315c1-811f-4823-a778-60a203439689	My Best Friend's	80	Julia Roberts	lfn:/grid/gilda/movies/mybfwed.avi
51a18b7a-fd21-4b2c-aa74-4c53ee64846a	Spider-man 2	120	Kirsten Dunst	lfn:/grid/gilda/movies/spiderman2.avi
401e6df4-c1be-4822-958c-ce3eb5c54fcb	The God Father	113	Al pacino	lfn:/grid/gilda/movies/godfather.avi



Schema

Attribute

**Entry names**

Title

Ru

Cast

LFN

Entry names	Title	Ru	Cast	LFN
8c3315c1-811f-4823-a778-60a203439689	My Best Friend's	80	Julia Roberts	lfn:/grid/gilda/movies/mybfwed.avi
51a18b7a-fd21-4b2c-aa74-4c53ee64846a	Spider-man 2	120	Kirsten Dunst	lfn:/grid/gilda/movies/spiderman2.avi
401e6df4-c1be-4822-958c-ce3eb5c54fcb	The God Father	113	Al pacino	lfn:/grid/gilda/movies/godfather.avi

Schema

Attribute

**Entry names**

**Title**

**Ru**

**Cast**

**LFN**

8c3315c1-811f-4823-a778-60a203439689

My Best Friend's

80

Julia Roberts

lfn:/grid/gilda/movies/mybfwed.avi

51a18b7a-fd21-4b2c-aa74-4c53ee64846a

Spider-man 2

120

Kirsten Dunst

lfn:/grid/gilda/movies/spiderman2.avi

401e6df4-c1be-4822-958c-ce3eb5c54fcb

The God Father

113

Al pacino

lfn:/grid/gilda/movies/godfather.avi

Entries

Schema

Attribute

**Entry names**

Title

Ru

Cast

LFN

8c3315c1-811f-4823-a778-60a203439689

My Best Friend's

80

Julia Roberts

lfn:/grid/gilda/movies/mybfwed.avi

51a18b7a-fd21-4b2c-aa74-4c53ee64846a

Spider-man 2

120

Kirsten Dunst

lfn:/grid/gilda/movies/spiderman2.avi

401e6df4-c1be-4822-958c-ce3eb5c54fcb

The God Father

113

Al pacino

lfn:/grid/gilda/movies/godfather.avi

Collection /trailers

Entries



- **Information about files -- but not only!**

- Information about files -- but not only!
- metadata can **describe** any grid entity/object
  - ex: JobIDs - add logging information to your jobs

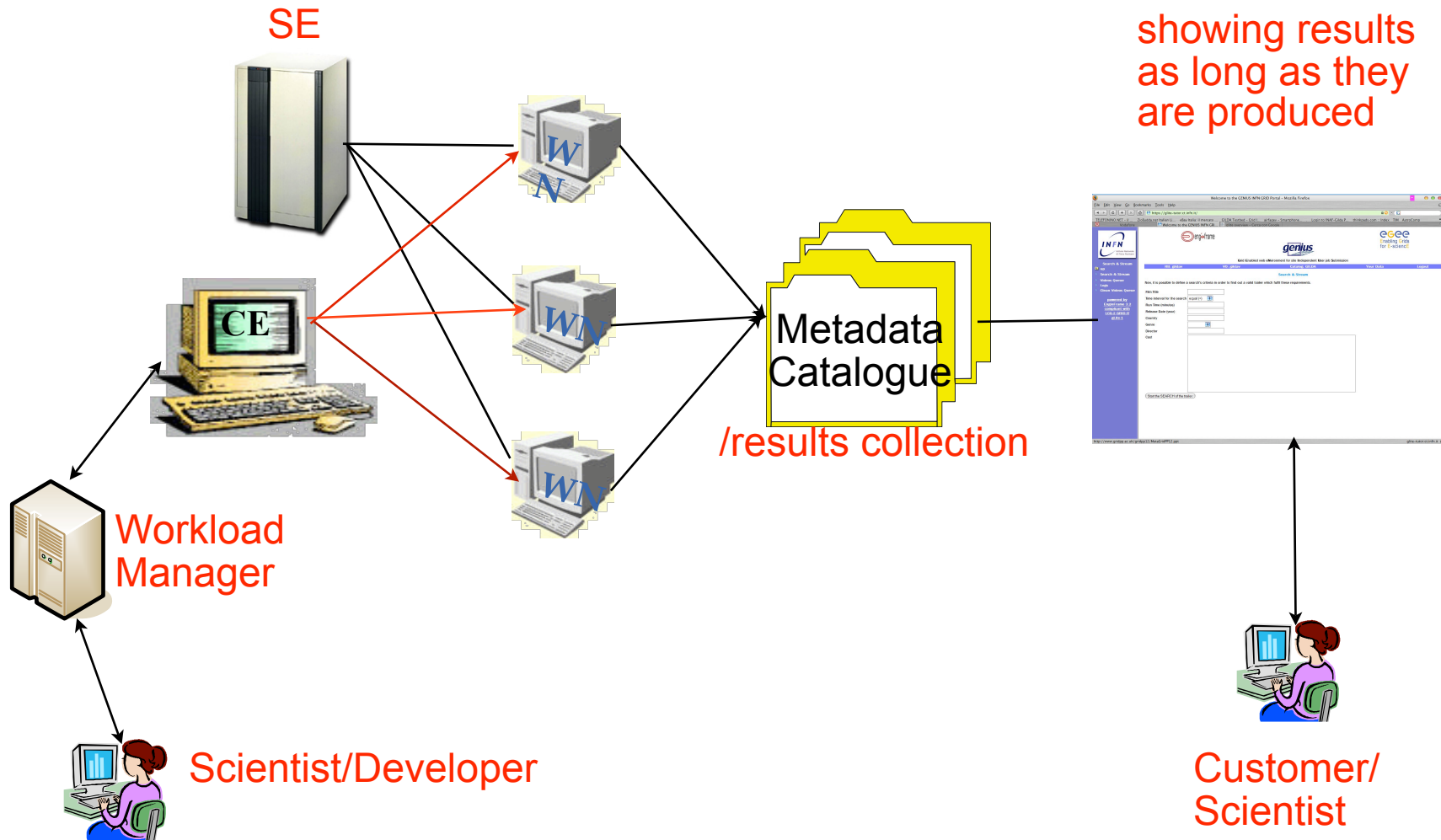
- Information about files -- but not only!
- metadata can **describe** any grid entity/object
  - ex: JobIDs - add logging information to your jobs
- **monitoring** of running applications:
  - ex: ongoing results from running jobs can be published on the metadata server

- Information about files -- but not only!
- metadata can **describe** any grid entity/object
  - ex: JobIDs - add logging information to your jobs
- **monitoring** of running applications:
  - ex: ongoing results from running jobs can be published on the metadata server
- **Inputset** for a storm of **parametric jobs**



- **Information about files -- but not only!**
- **metadata can describe any grid entity/object**
  - ex: JobIDs - add logging information to your jobs
- **monitoring of running applications:**
  - ex: ongoing results from running jobs can be published on the metadata server
- **Inputset for a storm of parametric jobs**
- **information exchanging among grid peers**
  - ex: producers/consumers job collections: master jobs produce data to be analyzed; slave jobs query the metadata server to retrieve input to “consume”

- Information about files -- but not only!
- metadata can **describe** any grid entity/object
  - ex: JobIDs - add logging information to your jobs
- **monitoring** of running applications:
  - ex: ongoing results from running jobs can be published on the metadata server
- **Inputset** for a storm of **parametric jobs**
- **information exchanging** among grid peers
  - ex: producers/consumers job collections: master jobs produce data to be analyzed; slave jobs query the metadata server to retrieve input to “consume”
- **Simplified DB access** on the grid
  - Grid applications that needs structured data can model their data schemas as metadata



showing results as long as they are produced

/results collection

Customer/Scientist

- **/grid/my\_simulation/input**

entry	x1	x2	y1	y2	step	isTaken	found	output
1	9453.1	9453.32	-439.93	-439.91	0.0006	JobID1234	No pillars	
2	9342.13	3435	3423	2343.2	0.003	No		
3	34254.3	342342	432.43	132	0.002	No		
..... and so on								

- **This collection lists all the parameter set to be run on the Grid**
- **On the WN, one of the inputset is selected and “isTaken” is set = JOB\_ID of the job that has fetched it**
- **Results is also written in the “found” column to monitor the simulation**
  - so users can check the simulation from a UI, querying the metadata server, or from a WebPage (using APIs for ex)
- **StdOutput can be copied also into the “output” text column**

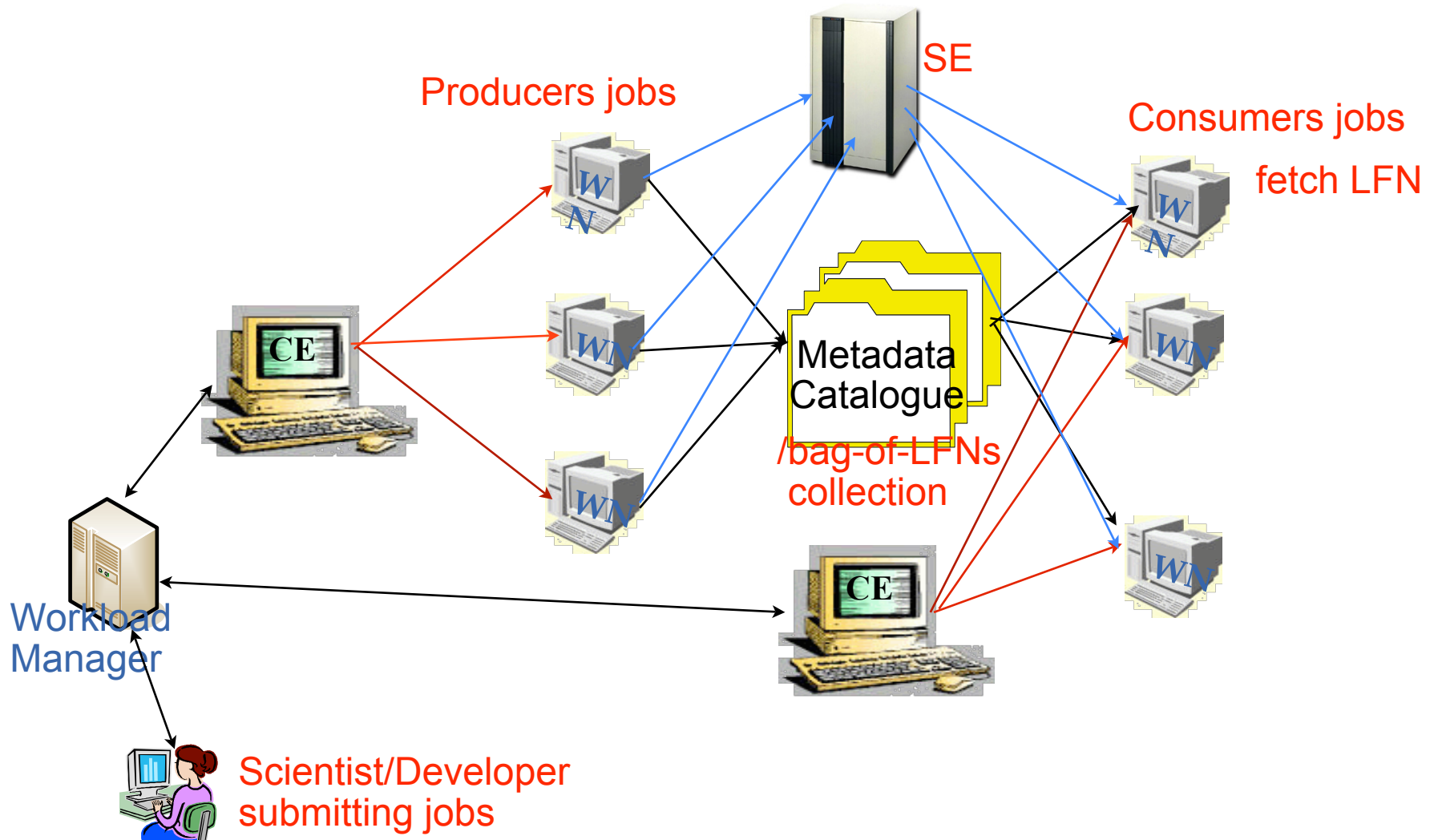


Enabling Grids for E-science

# Use a Metadata services to exchange data among running jobs

- **Suppose we have two sets of jobs:**
  - **Producers:** they generate a file, store on a SE, register it onto the LFC File Catalogue assigning a LFN
  - **Consumers:** they will take a LFN, download the file and elaborate it

- **Suppose we have two sets of jobs:**
  - **Producers:** they generate a file, store on a SE, register it onto the LFC File Catalogue assigning a LFN
  - **Consumers:** they will take a LFN, download the file and elaborate it
- **A Metadata collection can be used to share the information generated by the **Producers**; it could act as a “bag-of-LFNs” (bag-of-task model) from which **Consumers** can fetch file for further elaboration**





- **Official metadata service for the gLite middleware**
  - but no dependencies from gLite software
  - **it can be used with other grid technologies/other environments**
- **AMGA: Arda Metadata Grid Application**
- Provide a **complete but simple interface**, in order to make all users able to use it easily.
- Designed with **scalability** in mind in order to deal with **large number of entries**
  - based on a lightweight and streamed text-based protocol, like HTTP/SMTP
- **Grid security** is provided to grant **different access levels** to different users.
- **Flexible** with support to dynamic schemas in order to serve several application domains
- Simple installation by tar source, RPMs or Yum/YAIM



- **Dynamic Schemas**

- Schemas can be modified at runtime by client
  - Create, delete schemas
  - Add, remove attributes

- **Dynamic Schemas**
  - Schemas can be modified at runtime by client
    - Create, delete schemas
    - Add, remove attributes
- **AMGA collections are hierarchical organized**
  - Collections can contain sub-collections
  - Sub-collections can inherit/extend parent collection' schema

- **Dynamic Schemas**
  - Schemas can be modified at runtime by client
    - Create, delete schemas
    - Add, remove attributes
- **AMGA collections are hierarchical organized**
  - Collections can contain sub-collections
  - Sub-collections can inherit/extend parent collection' schema
- **Flexible Queries**
  - SQL-like query language
  - Different join type (inner, outer, left, right) between schemas are provided

```
selectattr /gLibrary:FileName /gLAudio:Author /gLAudio:Album  
'/gLibrary:FILE=/gLAudio:FILE and like(/gLibrary:FileName, "%.mp3")`
```

- **Dynamic Schemas**
  - Schemas can be modified at runtime by client
    - Create, delete schemas
    - Add, remove attributes
- **AMGA collections are hierarchical organized**
  - Collections can contain sub-collections
  - Sub-collections can inherit/extend parent collection' schema
- **Flexible Queries**
  - SQL-like query language
  - Different join type (inner, outer, left, right) between schemas are provided

```
selectattr /gLibrary:FileName /gLAudio:Author /gLAudio:Album
'/gLibrary:FILE=/gLAudio:FILE and like(/gLibrary:FileName, "%.mp3")`
```

- ▶ Support for **Views, Constraints, Indexes**

```
Query> selectattr /trailers:Title Runtime FILE 'Runtime > 80'  
>> Amelie of Montmartre  
>> 122  
>> 004405ac-da9a-1417-92db-c1ced08dbeef  
>> American Pie 2  
>> 108  
>> 006d56b4-d7d1-1417-8417-c1ced08dbeef  
>> Batman Begins  
>> 141  
>> 0072f510-db33-1417-b12e-c1ced08dbeef  
>> The Fast and The Furious  
>> 106  
>> 00737e72-d8cb-1417-871f-c1ced08dbeef  
>> Madagascar  
>> 86  
>> 0069b608-d95c-1417-9fd1-c1ced08dbeef  
>> The Matrix
```

```
Query> ls  
>> 004405ac-da9a-1417-92db-c1ced08dbeef  
>> 006d56b4-d7d1-1417-8417-c1ced08dbeef  
>> 0072f510-db33-1417-b12e-c1ced08dbeef  
>> 00737e72-d8cb-1417-871f-c1ced08dbeef  
>> 0069b608-d95c-1417-9fd1-c1ced08dbeef  
>> 0010bf6c-d9cc-1417-a38c-c1ced08dbeef  
>> 002e3966-d877-1417-8b9c-c1ced08dbeef
```

```
Query> listattr /trailers  
>> Title  
>> varchar(200)  
>> Runtime  
>> int  
>> Country  
>> varchar(25)  
>> ReleaseDate  
>> int  
>> Director  
>> varchar(80)  
>> PlotOutline  
>> text  
>> Cast  
>> varchar(2048)  
>> Genre  
>> varchar(100)  
>> Image  
>> text
```





- **Unix style permissions - users and groups**

- **Unix style permissions - users and groups**
- **ACLs – Per-collection or per-entry (table row).**

- **Unix style permissions - users and groups**
- **ACLs – Per-collection or per-entry (table row).**
- **Secure client/server connections – SSL**

- **Unix style permissions - users and groups**
- **ACLs – Per-collection or per-entry (table row).**
- **Secure client/server connections – SSL**
- **Client Authentication based on**
  - Username/password
  - General X509 certificates (DN based)
  - Grid-proxy certificates (DN based)

- **Unix style permissions - users and groups**
- **ACLs – Per-collection or per-entry (table row).**
- **Secure client/server connections – SSL**
- **Client Authentication based on**
  - Username/password
  - General X509 certificates (DN based)
  - Grid-proxy certificates (DN based)
- **VOMS support:**
  - VO attribute maps to defined AMGA user
  - VOMS Role maps to defined AMGA user
  - VOMS Group maps to defined AMGA group

- **C++ multiprocess server**

- **Backends**

- Oracle, MySQL 4/5, PostgreSQL, SQLite

- **Front Ends**

- **TCP text streaming**

- High performance
      - Client API for C++, Java, Python, Perl, PHP

- **SOAP (deprecated)**

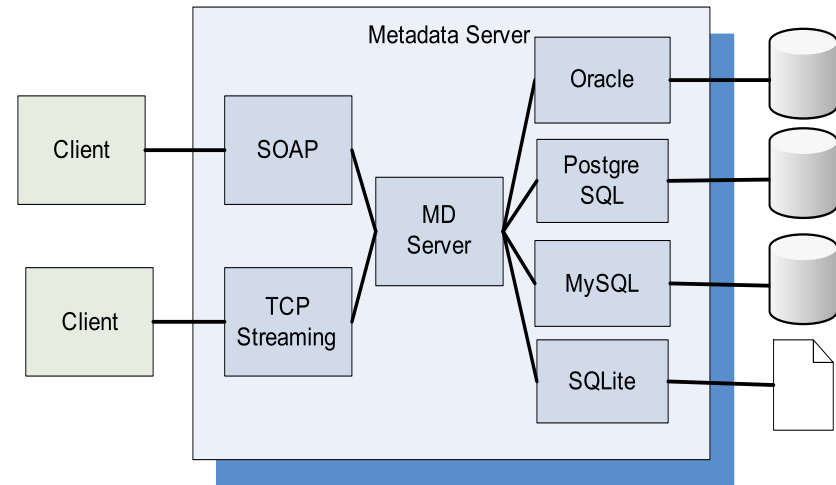
- Interoperability
      - Scalability

- **WS-DAIR Interface (new in AMGA 2.0)**

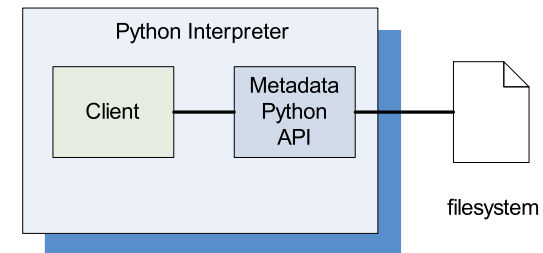
- **WS-enable environment**

- **Standalone Python Library implementation**

- Data stored on file system



- AMGA server runs on SLC3/4, Fedora Core, Gentoo, Debian



	PostgreSQL	MySQL	Oracle	SQLite	Python
<b>int</b>	integer	int	number(38)	int	int
<b>float</b>	double precision	double precision	float	float	float
<b>varchar(n)</b>	character varying(n)	character varying(n)	varchar2(n)	varchar(n)	string
<b>timestamp</b>	timestamp w/o TZ	datetime	timestamp(6)	unsupported	time (unsupp.)
<b>text</b>	text	text	long	text	string
<b>numeric(p,s)</b>	numeric(p,s)	numeric(p,s)	numeric(p,s)	numeric(p,s)	float

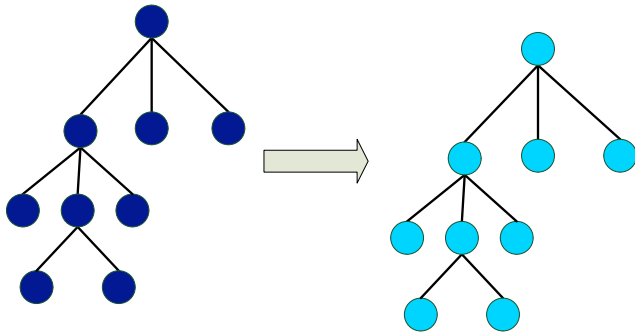
- ▶ Using the above datatypes you are sure that your metadata can be easily moved to all supported back-ends
- ▶ If you do not care about DB portability, you can use, in principle, as entry attribute type ALL the datatypes supported by the back-end, even the more esoteric ones (PostgreSQL Network Address type or Geometric ones)

- **TCP Streaming Front-end**
  - **mdcli** & **mdclient** CLI and C++ API (md\_cli.h, MD\_Client.h)
  - **Java** Client API and command line **mdjavaclient.sh** & **mdjavacli.sh** (also under Windows !!)
  - **Python** and **Perl** Client API
  - **PHP** Client API – NEW
    - developed totally by the GILDA team – INFN CT
  - AMGA **Web Interface** (AMGA WI) ---NEW
    - Developed totally by the GILDA team – INFN CT
    - Based on JAVA AMGA Standard APIs
    - Web Application using standard as JSP Custom Tags, Servlet
- **SOAP Frontend (WSDL)**
  - C++ gSOAP
  - AXIS (Java)
  - ZSI (Python)

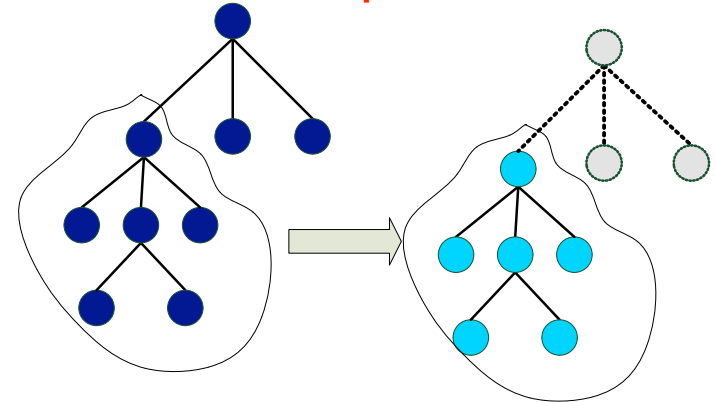


- AMGA provides a **replication/federation** mechanisms
- **Motivation**
  - **Scalability** – Support hundreds/thousands of concurrent users
  - **Geographical distribution** – Hide network latency
  - **Reliability** – No single point of failure
  - **DB Independent replication** – Heterogeneous DB systems
  - **Disconnected computing** – Off-line access (laptops)
- **Architecture**
  - **Asynchronous** replication
  - **Master-slave** – writes only allowed on the master
  - **Application level** replication
    - Replicate Metadata commands
  - **Partial replication** – supports replication of only sub-trees of the metadata hierarchy

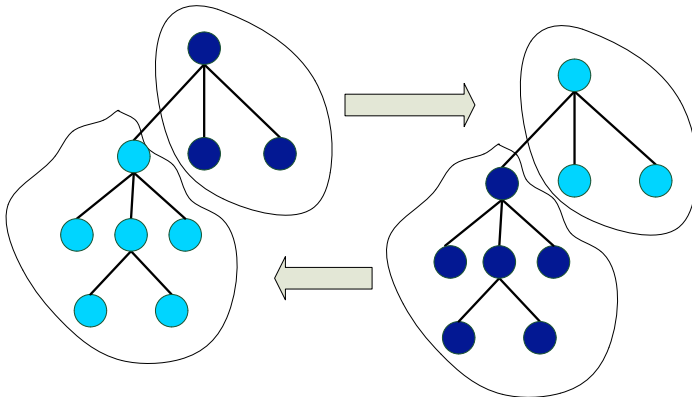
**Full replication**



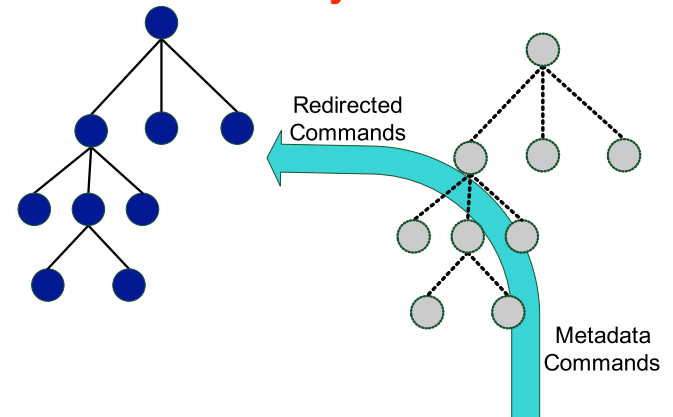
**Partial replication**



**Federation**



**Proxy**



- Since AMGA 1.2.10, a new **import** feature allow to access existing DB table
- Once imported into AMGA the tables from one or more DBs you want to access through AMGA, you can exploit many of the features brought to you by AMGA for your existing tables
- **Advantages:**
  - your db tables can be accessed by grid users/applications, using grid authentication (VOMS proxies)/authorization with ACLs
  - exploiting AMGA federation features you can access several databases together from the Grid

- To remember: AMGA stores its own tables in its DB backend
- To access and existing DB you have 2 option:
  - import the tables of the DB you want to access to into AMGA DB backend
  - viceversa, add AMGA DB backed tables to the DB you want to access to
- Use the **import** command by root to “mount” you table into the AMGA collection hierarchy

```
Query> whoami
>> root
Query> createdir /world
Query> cd /world/
Query> import world.City /world/City
Query> import world.Country /world/Country
Query> import world.CountryLanguage /world/CountryLanguage
```

- **Properly set up authorization on the imported tables:**

```
Query> acl_remove /world/City/ system:anyuser
Query> acl_remove /world/Country system:anyuser
Query> acl_add /world/ gilda:users rx
Query> acl_show /world
>> root rwx
>> gilda:users rx
>> system:anyuser rx
Query> selectattr City:CountryCode City:Name 'like(City:Name, "Am%")
limit 5'
>> NLD
>> Amsterdam
>> NLD
>> Amersfoort
>> BRA
>> Americana
>> ECU
>> Ambato
>> IDN
```

- ▶ **More information on existing DB access @:**

- ▶ <http://amga.web.cern.ch/amga/importing.html>
- ▶ <https://grid.ct.infn.it/twiki/bin/view/GILDA/AMGADBAccess>

- **Objective:**
  - implement native SQL query processing functionality in AMGA
- **Current Status:**
  - direct SQL data statement in SQL92 Entry Level has been implemented in the 1.9 release
    - Including 4 statements: SELECT, DELETE, UPDATE and INSERT
    - ALL SQL commands should be issued in UPPERCASE
- **Entry name:**
  - when a new entry is created with addentry/addentries, a name has to be assigned (filling the “file” column in the AMGA db backend)
    - in the INSERT implementation, it’s filled automatically with a random guid

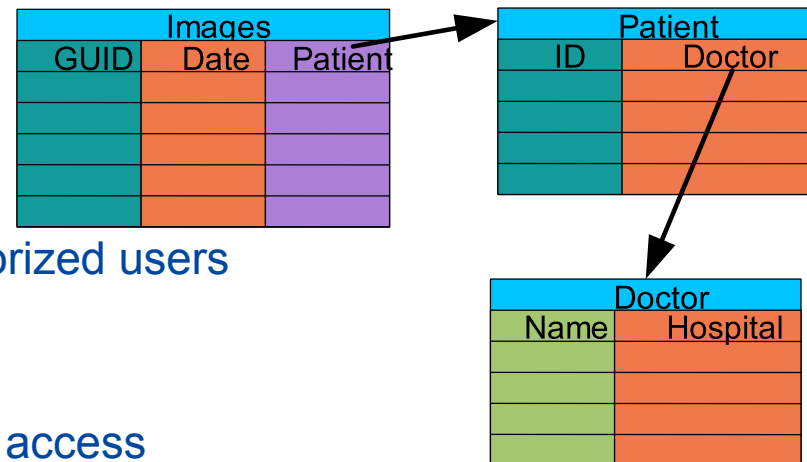
```
Query> INSERT INTO `City` VALUES (1,'Kabul','AFG','Kabol',1780000)
>> Operation Success
Query> dir /world/City/
>> /world/City/80b4fe646ed11dda02100304873049
>> entry
Query> SELECT COUNT (*) FROM /world/City
>> 3429
Query> SELECT * FROM /world/City WHERE Name LIKE '%Catani%'
>> 1472
>> Catania
>> ITA
>> Sisilia
>> 337862
Query> SELECT /world/City:Name, /world/City:District, /world/Country:Name, /
world/Country:Region, /world/Country:Continent FROM /world/City, /world/Country
WHERE /world/City:Name LIKE '%Catani%' AND Code = 'ITA'
>> Catania
>> Sisilia
>> Italy
>> Southern Europe
>> Europe
```

## Medical Data Manager – MDM

- Store and access medical images and associated metadata on the Grid
- Built on top of gLite 1.5 data management system
- Demonstrated at last EGEE conference (October 05, Pisa)

## Strong security requirements

- Patient data is sensitive
- Data must be encrypted
- Metadata access must be restricted to authorized users



## AMGA used as metadata server

- Demonstrates authentication and encrypted access
- Used as a simplified DB

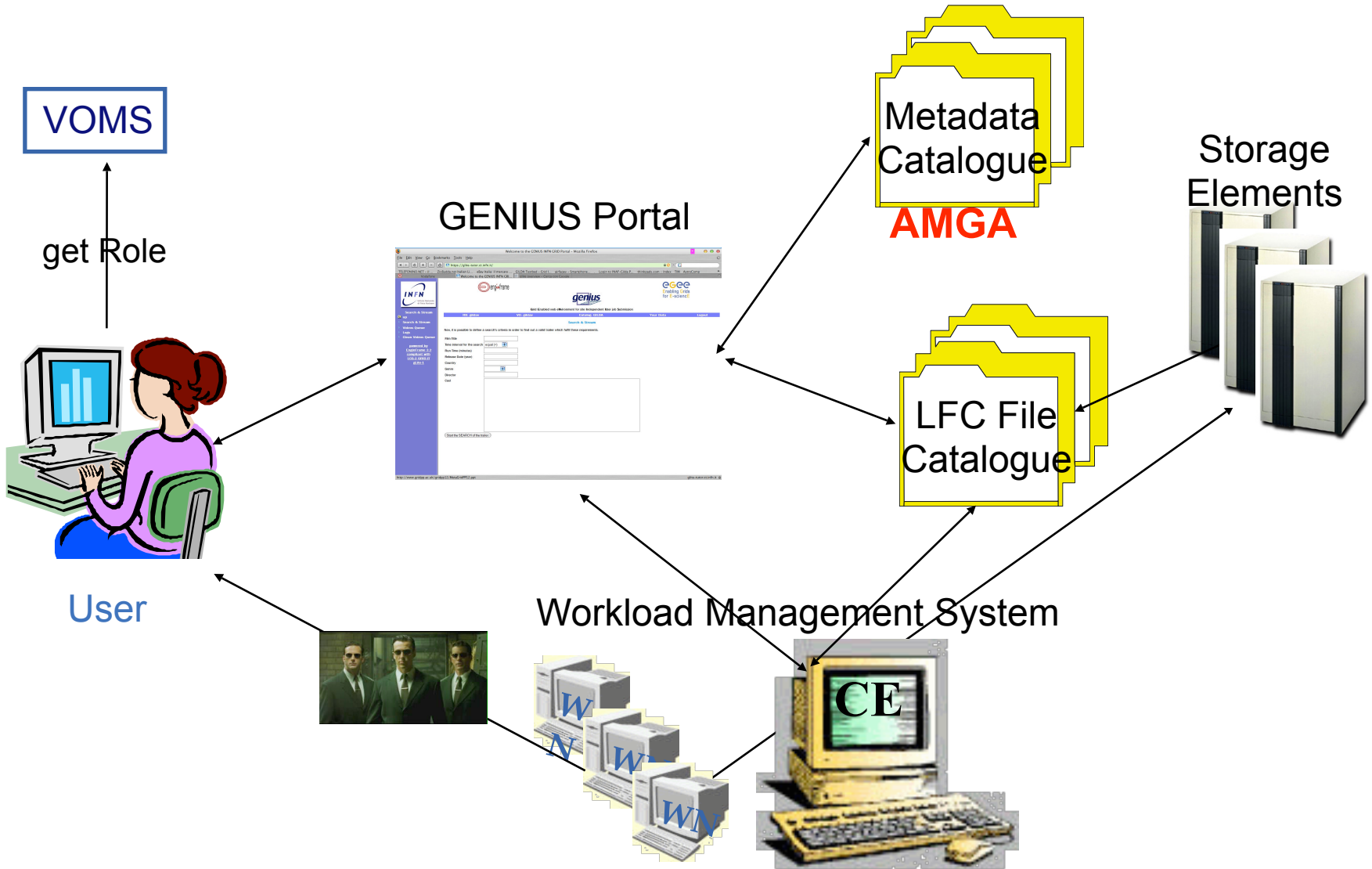
## More details at

- <https://uimon.cern.ch/twiki/bin/view/EGEE/DMEncryptedStorage>



- gMOD provides a **Video-On-Demand** service
- User chooses among a list of video and the chosen one is streamed in real time to the video client of the user's workstation
- For each movie a lot of details (Title, Runtime, Country, Release Date, Genre, Director, Case, Plot Outline) are stored and users can search a particular movie querying on one or more attributes
- Two kind of users can interact with gMOD: **TrailersManagers** that can administer the db of movies (uploading new ones and attaching metadata to them); **GILDA VO users (guest)** can browse, search and choose a movie to be streamed.

- Built on top of gLite services:
- **Storage Elements**, sited in different place, physically contain the movie files
- **LFC**, the File Catalogue, keeps track in which Storage Element a particular movie is located
- **AMGA** is the repository of the detailed information for each movie, and makes possible queries on them
- The **Virtual Organization Membership Service (VOMS)** is used to assign the right role to the different users
- The **Workload Management System (WMS)** is responsible to retrieve the chosen movie from the right Storage Element and stream it over the network down to the user's desktop or laptop



gMOD is accessible through the Genius Portal (<https://glite-demo.ct.infn.it>)

Welcome to the GENIUS INFN GRID Portal - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

https://glite-tutor.ct.infn.it/

TELEFONINO.NET - il ... ZioBudda.net Italian Li... e8av Italia: il mercato ... GILDA Testbed - Grid I... airfacev - Smartphone... Login to INAF-Gilda P... thinkpads.com :: Index TIM AstroComp

Vodafone Welcome to the GENIUS INFN GRI...

Istituto Nazionale di Fisica Nucleare  
 EnginFrame  
 genius  
 EGEE Enabling Grids for E-science

Grid Enabled web eNvironment for site Independent User job Submission

RB: gildav VO: gildav Catalog: GILDA Your Data Logout

You have selected the trailer **/trailers/Shrek2.mpg**.  
These are the attributes of the trailer you have chosen.

Title	<b>Shrek 2</b>				
Run Time	92	Country	USA	Release Date	2004
Genre	Action	Director	William Steig		
Cast	Mike Myers .... Shrek (voice) Eddie Murphy .... Donkey (voice) Cameron Diaz .... Fiona (voice) John Cleese .... King (voice) Rupert Everett .... Prince Charming (voice) James Earl Ray .... Nobleman/Nobleman's Son (voice) Cody Cameron .... Pinocchio/Three (voice) Christopher Knights .... Blind Mouse (voice) David P. Smith .... Herald/Marshmallow				
Plot Outline	The film picks up right where the first movie ended... Shrek and Fiona return from their honeymoon. The only problem is that they have no idea that their daughter is now an ogre.				

**powered by**  
 EnginFrame 3.2  
 compliant with  
 LCG-2 GRID-IT  
 glite-1

```

LFN: /trailers/Shrek2.mpg
Created: 2005-10-13 17:23:58.000
Modified: 2005-10-13 17:23:58.000
Size: 6100996
Type: File
Expires: Never
GUID: 000f0e2e-7c0d-134e-a731-c1ced08dbeef
Created: 2005-10-13 17:24:03.000
Modified: 2005-10-13 17:24:03.000
Size: 6100996
Checksum: 00000000
Status: 0
User: /C=IT/O=GILDA/OU=Personal Certificate/L=INAF
Group: eege-group
User rights: pdrwl-gs
Group rights: pdrwl-gs
    
```

Done

glite-tutor.ct.infn.it

- INFN-developed tool totally gLite based
- It allows to **store, organize, search** and **retrieve** digital assets on a Grid environment with an intuitive front-end
- What we mean by **Digital Assets**:



**FAIR INVOICE**

Fair: CareerFair Live  
 Bill To: Sample Company  
 Recruiting  
 Joe Apple  
 1237 Steel Road  
 Suite SD 47564-8373

Invoice Date	Invoice Number
10/07/2006	41
Registration ID	Site ID
43	3

**Invoice Summary**  
 Billing Date: 04/03/2006 Amount Due: \$201.00 Due By: 05/03/2006 Payment Method: CK  
 Rent check payments to: Strategy Ten, LLC, 9834 Dawn Trail, San Antonio, TX, 78254  
 Make checks payable to: "CareerFair Live"

**Invoice Details**

Item	Price	Qty	Est. Price
Wednesday Base Package	\$1.00	1	\$1.00

Genre	Artist	Album
All (9 Genres)	All (4 Artists)	All (3 Albums)
Acoustic	Elton John	In Between Dreams
Dance	Jack Johnson	Love Songs
Disco	Regina Spektor	Top Of The Spot 2008
General Jazz	Tears For Fears	
House		
Other		
Pop		
Rock		
Vocal		

Name	Time	Artist	Album
<input checked="" type="checkbox"/> Better Together	3:28	Jack Johnson	In Between Dreams
<input checked="" type="checkbox"/> Never Know	3:33	Jack Johnson	In Between Dreams
<input checked="" type="checkbox"/> Banana Pancakes	3:12	Jack Johnson	In Between Dreams
<input checked="" type="checkbox"/> Good People	3:29	Jack Johnson	In Between Dreams
<input checked="" type="checkbox"/> No Other Way	3:10	Jack Johnson	In Between Dreams
<input checked="" type="checkbox"/> Sitting, Waiting, Wishing	3:04	Jack Johnson	In Between Dreams
<input checked="" type="checkbox"/> Staple It Together	3:16	Jack Johnson	In Between Dreams
<input checked="" type="checkbox"/> Situations	1:17	Jack Johnson	In Between Dreams
<input checked="" type="checkbox"/> Crying Shame	3:06	Jack Johnson	In Between Dreams
<input checked="" type="checkbox"/> If I Could	2:25	Jack Johnson	In Between Dreams
<input checked="" type="checkbox"/> Breakdown	3:33	Jack Johnson	In Between Dreams
<input checked="" type="checkbox"/> Belle	1:43	Jack Johnson	In Between Dreams
<input checked="" type="checkbox"/> Do You Remember	2:25	Jack Johnson	In Between Dreams
<input checked="" type="checkbox"/> Constellations	3:22	Jack Johnson	In Between Dreams
<input checked="" type="checkbox"/> Daniel	3:56	Elton John	Love Songs
<input checked="" type="checkbox"/> Sowing the Seeds of Love (Spot Vodafone)	4:05	Tears For Fears	Top Of The Spot 2008
<input checked="" type="checkbox"/> Fidelity (Spot Kinder Maxi)	3:47	Regina Spektor	Top Of The Spot 2008



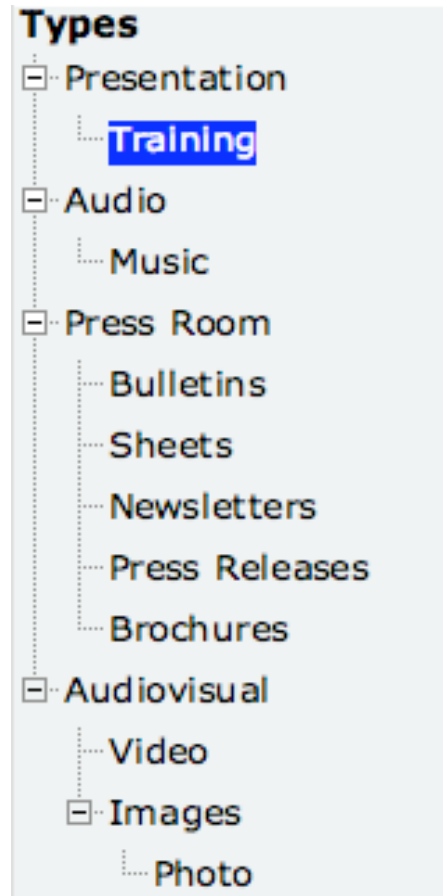
- “Types” and “Categories” definition by repository providers/admins:

- **Assets are organized by type:**

- a list of specific attributes to describe each kind of asset to be managed by the system
- hierarchical (a child type shares and extend parent’s attributes)
- queried during searches

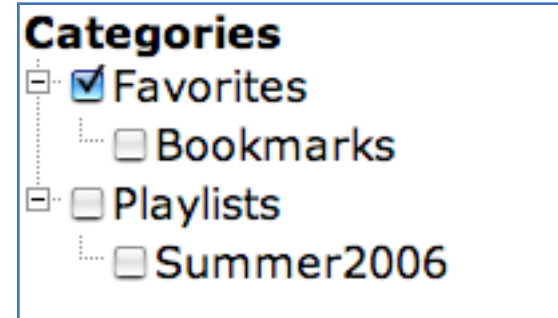
EXAMPLE OF TYPES AND ATTRIBUTES’ LIST

Type	Attributes’ list
Audio	Format, Bitrate, Samplerate, Time
Music	(Format, Bitrate, Samplerate, Time), Name, Artist, Album, Genre, Tracknumber, Year, Artwork, Lyric, Rating
Presentation	Format, NumOfPages
Training	(Format, NumOfPages), Title, Runtime, Speaker, Author, Subject, Event, Date, Type
(Root)	FileName, SubmissionDate, Description, Keywords, LastModificationDate, Size

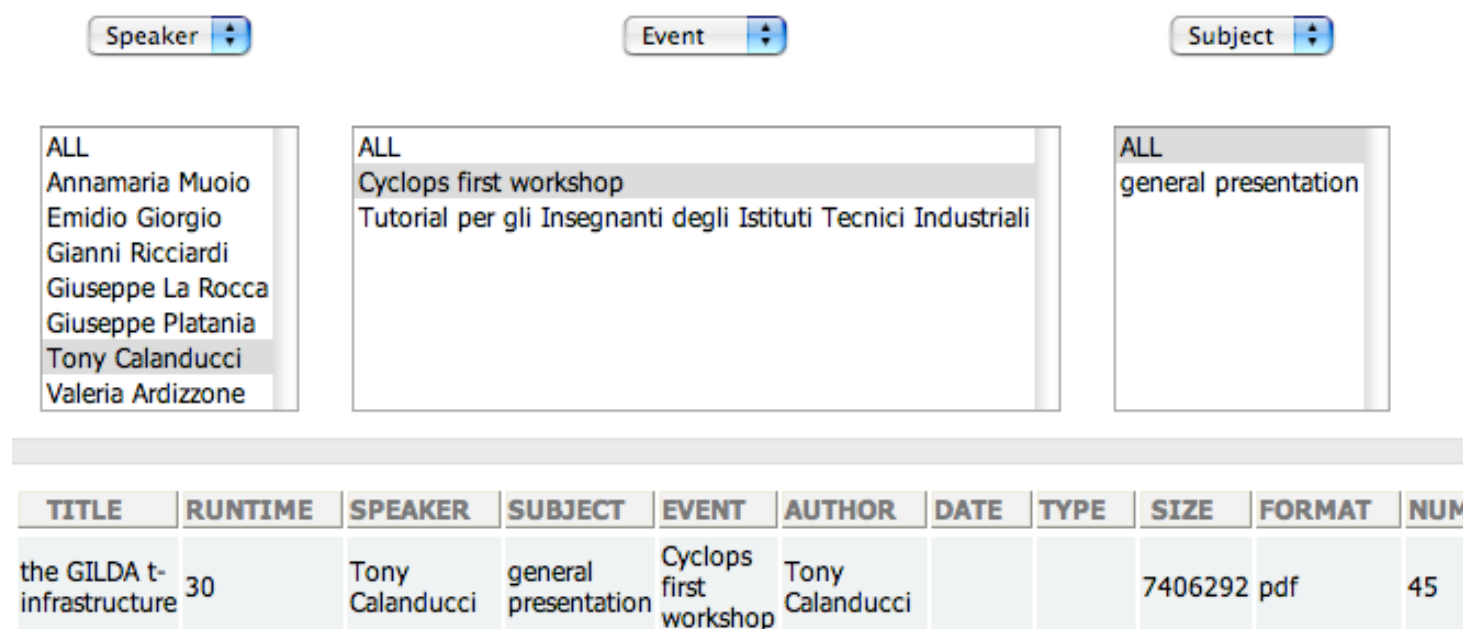


- and/or organized by **category/ collection:**

- Group together related assets of different type;
- Useful also to define subsets of assets belonging to the same type
- Multiple category assignment per asset (tagging)



- **Assets can be browsed selecting a type (or category) and selecting one or more filters:**
  - attributes of the selected types, chosen from a defined list, used to narrow the result set
- **Filter application is cascading and context-sensitive: the selection of a filter value dynamically influences subsequent filter values (“à la iTunes” browsing)**
  - Classical search by description and keywords available too



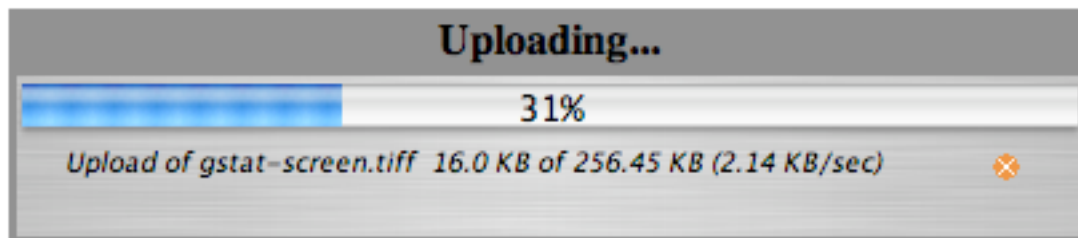
The screenshot shows a search interface with three filter dropdowns: Speaker, Event, and Subject. The Speaker dropdown is open, showing a list of names including Tony Calanducci. The Event dropdown is open, showing 'Cyclops first workshop' and 'Tutorial per gli Insegnanti degli Istituti Tecnici Industriali'. The Subject dropdown is open, showing 'general presentation'. Below the filters is a table of search results.

TITLE	RUNTIME	SPEAKER	SUBJECT	EVENT	AUTHOR	DATE	TYPE	SIZE	FORMAT	NUM
the GILDA t-infrastructure	30	Tony Calanducci	general presentation	Cyclops first workshop	Tony Calanducci			7406292	pdf	45

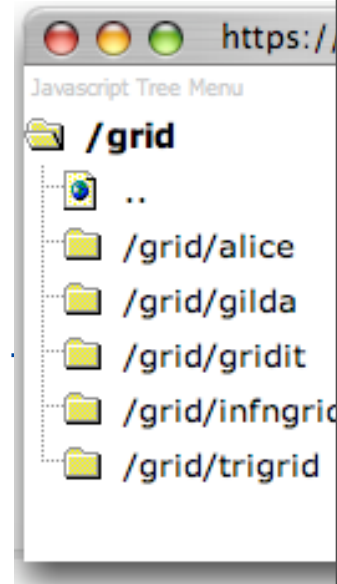


## Users can upload their local assets on one or more (creating replicas) Storage Elements of the Grid

- Uploads managed through Java Applets



- Files already on SE can be included in a digital library by File Catalogue browser



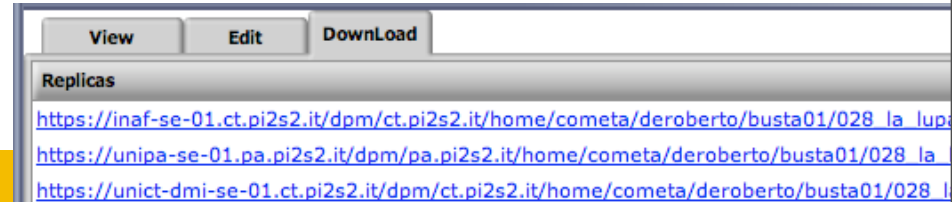
## Download from SEs to the users' laptop/desktop:

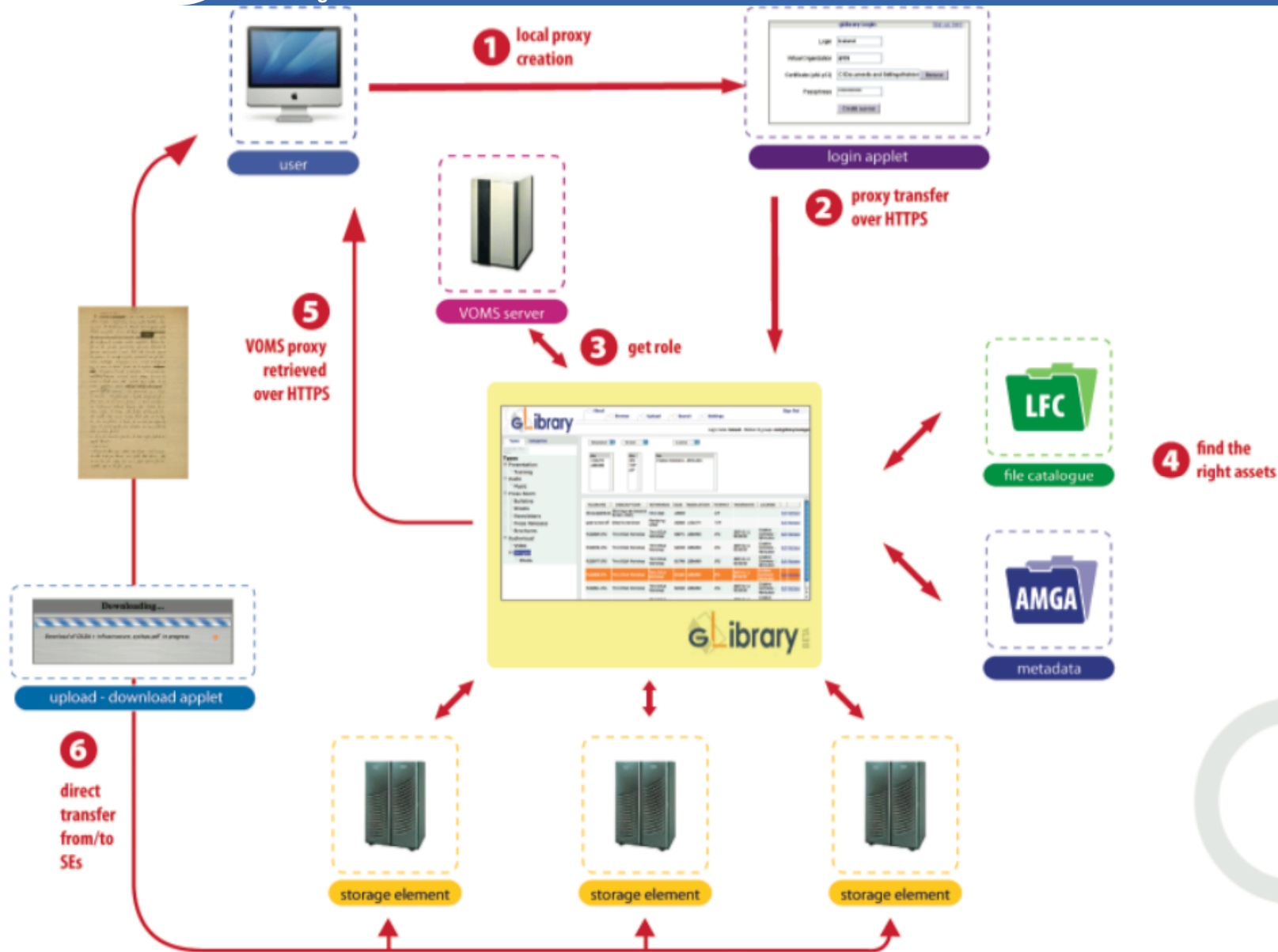
- selection of a replica link from a list
- download java applet

### List of replicas:

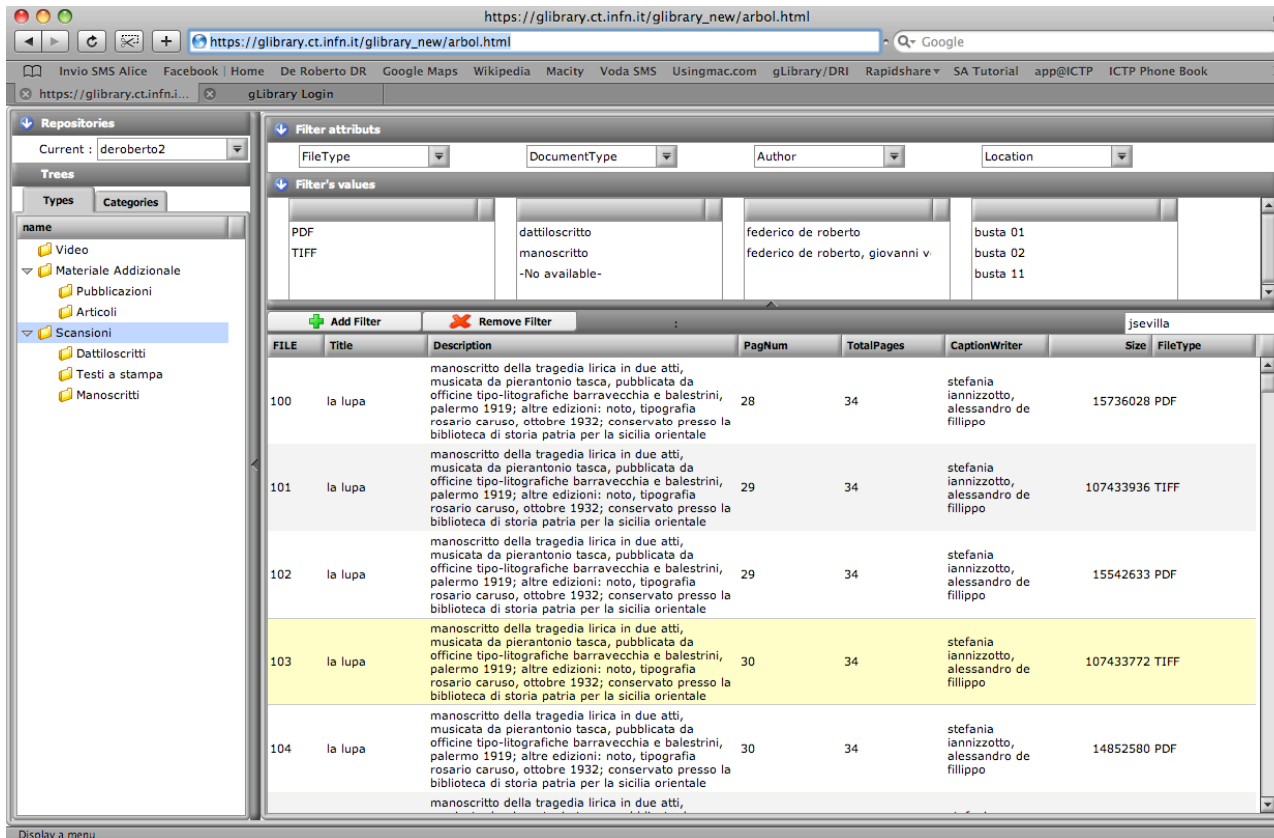
<https://infn-se-01.ct.pi2s2.it/dpm/ct.pi2s2.it/home/cometa/generated/2008-06-14/filea203fe0e-08ed-495b-8d16-73a6fa275e8e>  
<https://unime-se-01.me.pi2s2.it/dpm/me.pi2s2.it/home/cometa/generated/2008-06-14/file59c6ebad-6833-4e6d-bc49-f7241e7f45b8>  
<https://unipa-se-01.pa.pi2s2.it/dpm/pa.pi2s2.it/home/cometa/generated/2008-06-14/file680f1e17-530e-4a9e-a9d0-1df0c9cf6d86>

## Latest NEWS: transfers can be handled over HTTPS with X.509 Grid Cert AuthN/AuthZ





- **Implemented as Web 2.0 application**
  - AJAX and Javascript are strongly used to offer a desktop like user experience
  - Business logic implemented using PHP 5 OOP support



Modal Window

View Edit Download

Replicas

[https://inaf-se-01.ct.pi2s2.it/dpm/ct.pi2s2.it/home/cometa/deroberto/busta01/029\\_la\\_lupa.tif](https://inaf-se-01.ct.pi2s2.it/dpm/ct.pi2s2.it/home/cometa/deroberto/busta01/029_la_lupa.tif)

[https://unipa-se-01.pa.pi2s2.it/dpm/pa.pi2s2.it/home/cometa/deroberto/busta01/029\\_la\\_lupa.tif](https://unipa-se-01.pa.pi2s2.it/dpm/pa.pi2s2.it/home/cometa/deroberto/busta01/029_la_lupa.tif)

[https://unict-dmi-se-01.ct.pi2s2.it/dpm/ct.pi2s2.it/home/cometa/deroberto/busta01/029\\_la\\_lupa.tif](https://unict-dmi-se-01.ct.pi2s2.it/dpm/ct.pi2s2.it/home/cometa/deroberto/busta01/029_la_lupa.tif)

View	Edit	Download
FileName:	029_la_lupa.tif	
TypeID:	2	
CategoryIDs:		
SubmissionDate:	2009-03-02 19:54:00	
Description:	manoscritto della tragedia lirica in due atti, musicata da pierantonio tasca, pubblicata da officine tipografiche barravecchia e balestrini, palermo 1919; altre edizioni: noto, tipografia rosario caruso, ottobre 1932; conservato presso la biblioteca di storia patria per la sicilia orientale	
Keywords:	verismo, federico de roberto, la lupa, giovanni verga, officine tipo-litografiche barravecchia e balestrini palermo, tipografia rosario caruso noto, società di storia patria per la sicilia orientale catania, manoscritti letterari moderni, la.mu.s.a., facoltà di lettere e filosofia, università degli studi di catania, società di storia patria per la sicilia orientale	
LastModificationDate:	2007-08-11 20:59:00	
Size:	107433936	
Thumb:		
FileType:	TIFF	
ImageWidth:	5100	
ImageHeight:	7020	
CreatorTool:	Adobe Photoshop CS2 Windows	
XResolution:	600	
YResolution:	600	
Title:	la lupa	
Author:	federico de roberto	
CaptionWriter:	stefania iannizzotto, alessandro de fillippo	
CopyrightStatus:	Copyrighted	

Add Filter Remove Filter

Modal Window

View Edit Download

FileName:  TypeID:

CategoryIDs:  SubmissionDate:

Description:  Keywords:

LastModificationDate:  Size:

Thumb:  FileType:

ImageWidth:  ImageHeight:

CreatorTool:  XResolution:

YResolution:  Title:

Author:  CaptionWriter:

CopyrightStatus:  CopyrightNotice:

PagNum:  TotalPages:

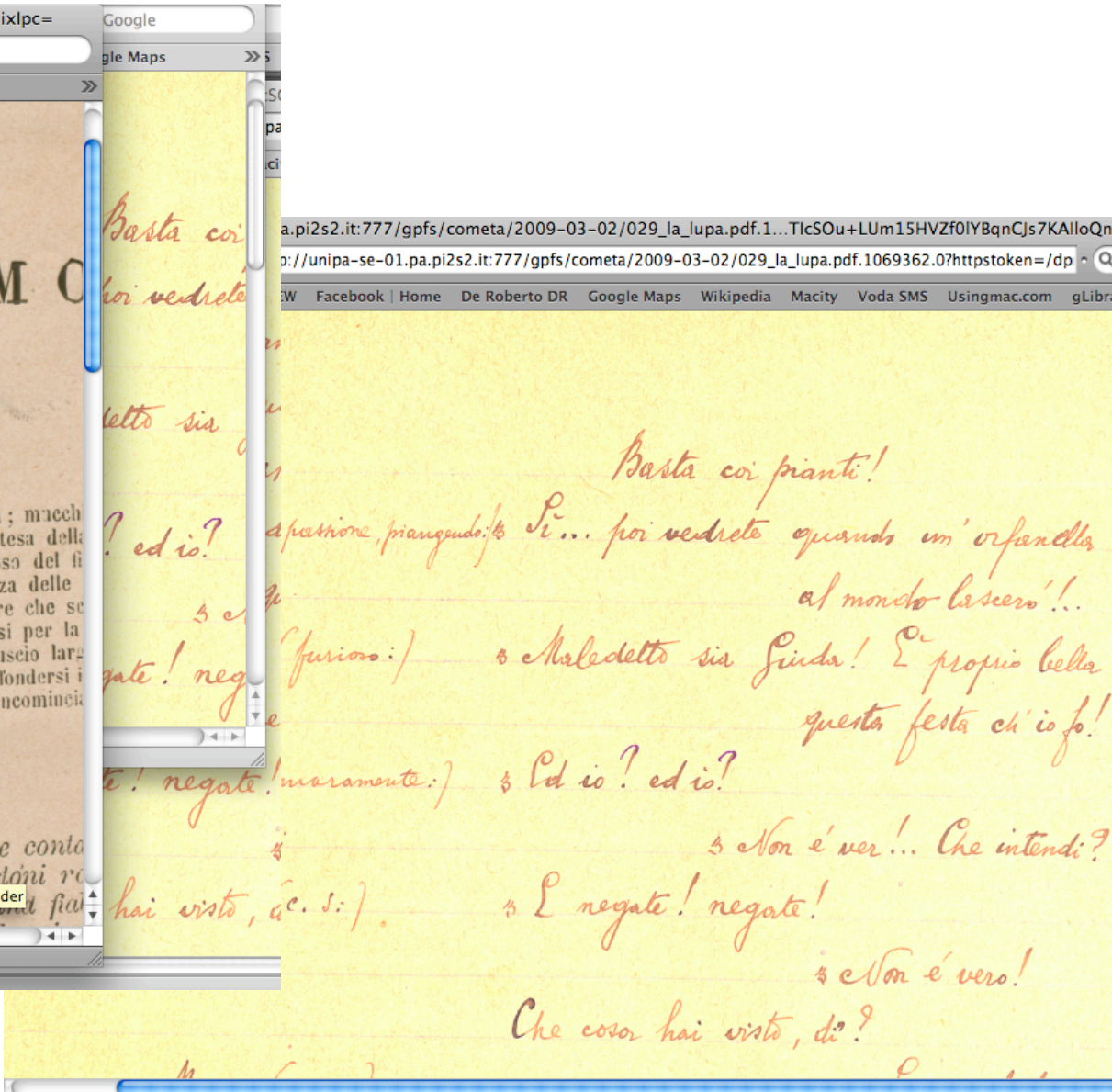
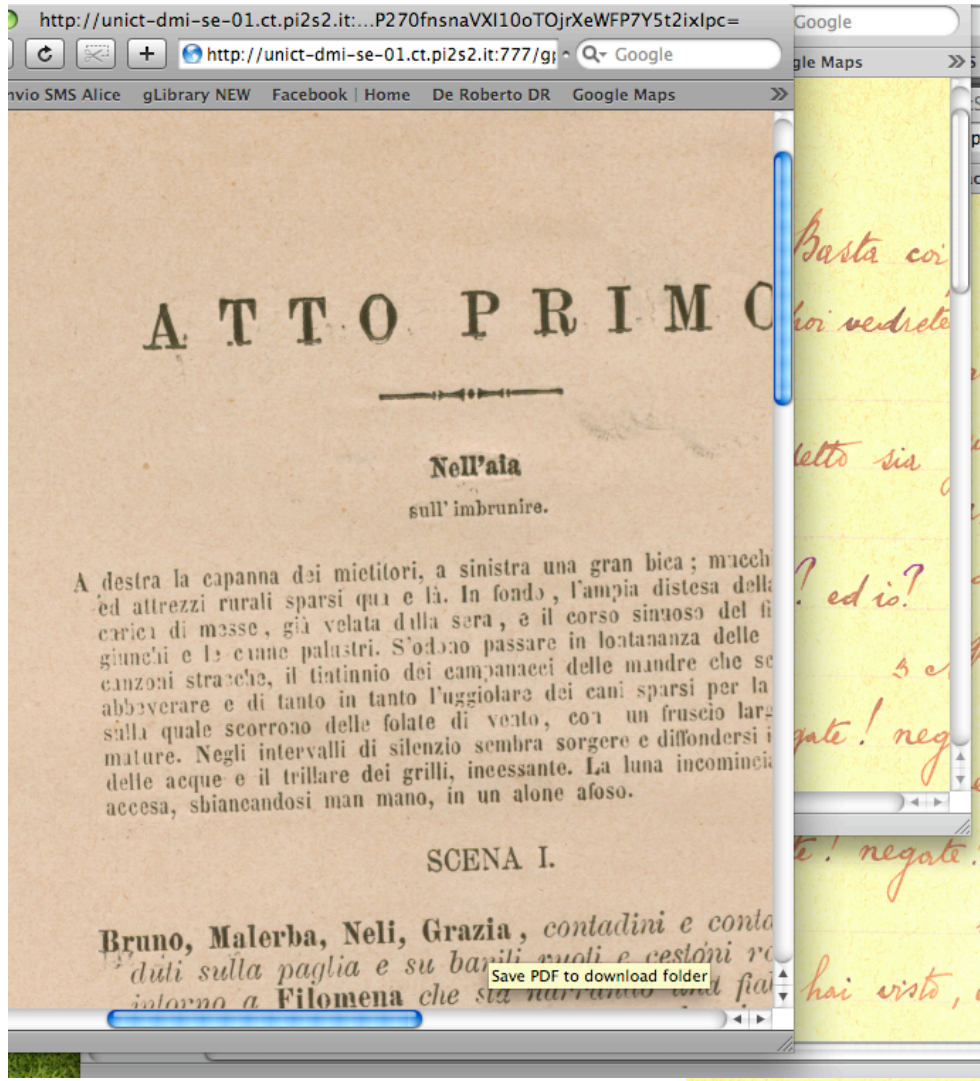
PhysicalFormat:  ScanDate:

ScanQuality:  DocumentGenre:

DocumentType:  PublicationYear:

officine tipo-		
chia e		
ipografia		
società di		
licilia orientale	107433936	TIFF
letterari		
facoltà di		
iversità degli		
età di storia		
rientale		
roberto, la		
, officine tipo-		
chia e		
ipografia		
società di		
licilia orientale	15542633	PDF
letterari		
facoltà di		
iversità degli		
età di storia		
rientale		
roberto, la		
, officine tipo-		
chia e		
ipografia		
società di		
storia patria per la sicilia orientale	107433772	TIFF

103	la lupa	pierantonio tasca, pubblicata da officine tipo-litografiche barravecchia e balestrini, palermo 1919; altre edizioni:	30	34	rosario caruso noto, società di storia patria per la sicilia orientale 107433772	TIFF
-----	---------	--	----	----	--	------



- **Storage Element** – save data and provide a common interface
  - Storage Resource Manager (SRM) Castor, dCache, DPM, ...
  - Native Access protocols rfiio, dcap, nfs, ...
  - Transfer protocols gsiftp, ftp, ...
- **Catalogs** – keep track where data are stored
  - File Catalog
  - Replica Catalog
  - Metadata Catalog

LCG File Catalog (LFC)

AMGA Metadata Catalogue
- **Data Movement** – schedules reliable file transfer
  - File Transfer Service gLite FTS  
(manages physical transfers)

- **gLite documentation homepage**
  - <http://glite.web.cern.ch/glite/documentation/default.asp>
- **DM subsystem documentation**
  - <http://egee-jra1-dm.web.cern.ch/egee-jra1-dm/doc.htm>
- **LFC and DPM documentation**
  - <https://uimon.cern.ch/twiki/bin/view/LCG/DataManagementDocumentation>

- **AMGA Web Site**

<http://cern.ch/amga>

- **AMGA Manual**

[http://amga.web.cern.ch/amga/downloads/amga-manual\\_1\\_3\\_0.pdf](http://amga.web.cern.ch/amga/downloads/amga-manual_1_3_0.pdf)

- **AMGA API Javadoc**

<http://amga.web.cern.ch/amga/javadoc/index.html>

- **AMGA Web Frontend**

<http://gilda-forge.ct.infn.it/projects/amgawi/>

- **AMGA Basic Tutorial**

<https://grid.ct.infn.it/twiki/bin/view/GILDA/AMGAHandsOn>

- **More information on existing DB access @:**

- <http://amga.web.cern.ch/amga/importing.html>

- <https://grid.ct.infn.it/twiki/bin/view/GILDA/AMGADBaccess>



- **gLibrary BETA homepage:**
  - [https://glibrary.ct.infn.it/glibrary\\_new/](https://glibrary.ct.infn.it/glibrary_new/)
- **gLibrary paper:**
  - [https://glibrary.ct.infn.it/glibrary/downloads/gLibrary\\_paper\\_v2.pdf](https://glibrary.ct.infn.it/glibrary/downloads/gLibrary_paper_v2.pdf)

