



Enabling Grids for
E-science in Europe

www.eu-egee.org

GENIUS/GILDA Tutorial, Catania, 24-25.05.2004

Data Services

Valeria Ardizzone
EGEE NA4 Generic Applications



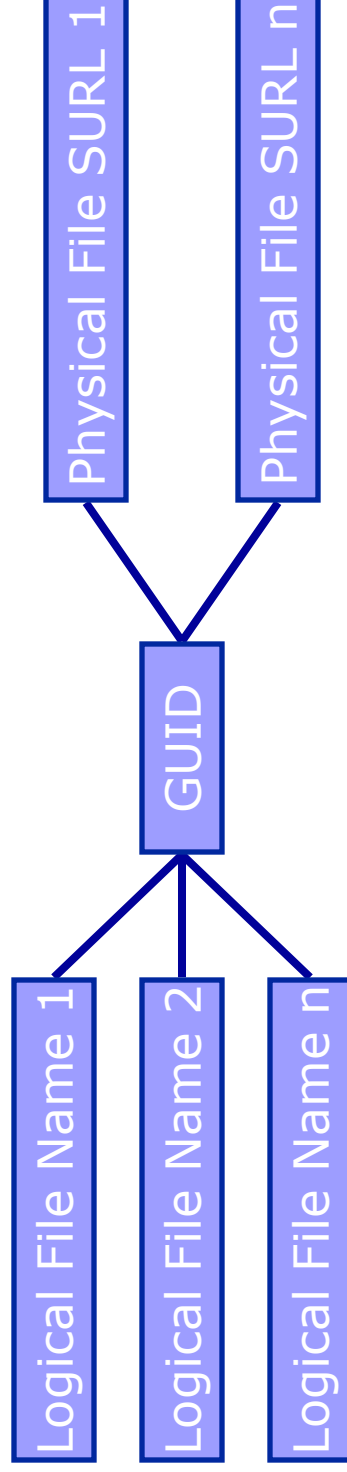
Contents

- **Data Management services**
- **Naming Conventions**
- **Management commands**
- **Catalog commands**
- **Optimization commands**
- **File Transfer commands**
- **Exercises**



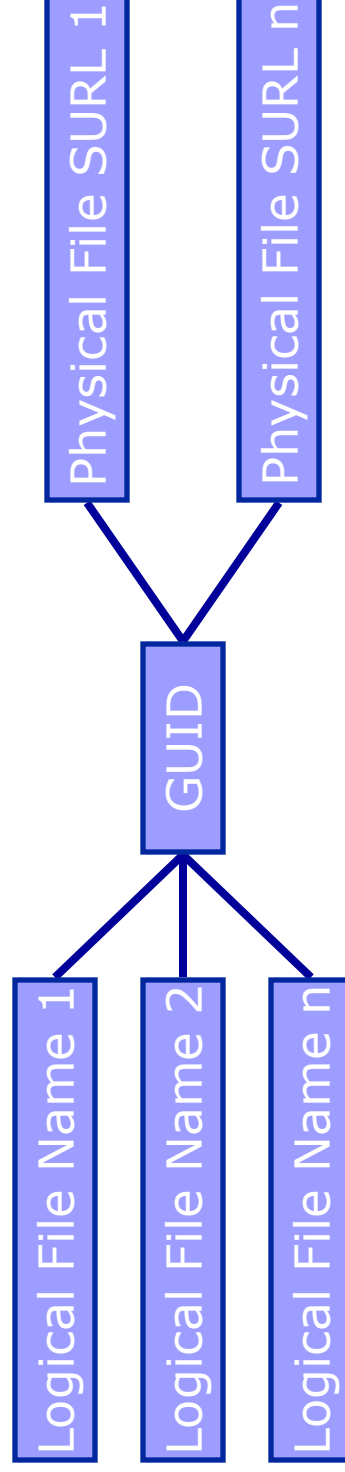
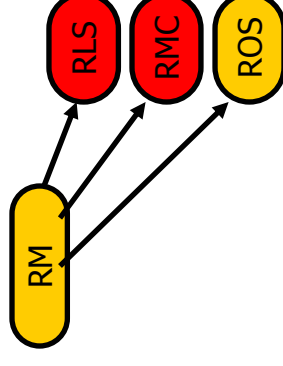
Naming Conventions

- Logical File Name (**LFN**)
 - An alias created by a user to refer to some item of data
e.g. “lfn:cms/20030203/run2/track1”
- Storage URL (**SURL**) (or Physical File Name (**PFN**))
 - The location of an actual piece of data on a storage system e.g. “srm://pcrd24.cern.ch/flatfiles/cms/output10_1”
- Globally Unique Identifier (**GUID**)
 - A non-human readable unique identifier for an item of data
e.g. “guid:f81d4fae-7dec-11d0-a765-00a0c91e6bf6”



Replica Metadata Catalog (RMC) Replica Location Service (RLS)

- **RMC:**
 - Stores LFN-GUID mappings
- **RLS:**
 - Stores GUID-SURL mappings



Management Commands(1/3)

copyAndRegisterFile

args: source

option: dest, lfn, protocol, streams

- Copy a file into grid-aware storage and register the copy in the Replica Catalog as an atomic operation. Upon successful completion this call returns the GUID of this file, that can be used to retrieve it again.

```
>> edg-rm --vo gilda copyAndRegisterFile  
file:///home/vardizzo/empty-localfile -d grid009.ct.infn.it -l  
lfn:myfirst
```

guid:2fdce224-af32-11d8-bca4-a3c587356e55

Management Commands(2/3)

replicateFile

args: source (LFN | GUID | SURL)

option: dest, protocol, streams

- Replicate a file between grid-aware stores and register the replica in the Replica Catalog as an atomic operation. It is also prints the SURL of the new replica.

```
>>edg-rm --vo gilda replicateFile guid:2fdce224-af32-11d8-bca4-  
a3c587356e55 -d grid015.pd.infn.it  
sfn://grid015.pd.infn.it/shared/gilda/generated/2004-05-  
26/filee27fab1c-af33-11d8-8585-b431461c4a4d
```

or

```
>>edg-rm --vo gilda replicateFile lfn:myfirst -d grid015.pd.infn.it  
sfn://grid015.pd.infn.it/shared/gilda/generated/2004-05-  
26/filee27fab1c-af33-11d8-8585-b431461c4a4d
```

Management Commands(3/3)

deleteFile

args: source(LFN | GUID | SURL).

option: SEHost, all available.

- Delete a file from storage and unregister it. With the all available tag, all the replicas may be deleted by specifying the GUID.

```
>>edg-rm --vo gilda deleteFile  
      guid:2fdce224-af32-11d8-bca4-a3c587356e55  
      --all-available
```

or

```
>>edg-rm --vo gilda deleteFile  
      guid:2fdce224-af32-11d8-bca4-a3c587356e55  
      -s grid015.pd.infn.it
```

Catalog Commands (1/2)

- **listGUID**
args: LFN | SURL
 - Print the GUID associated with an LFN or SURL.

```
>> edg-rm --vo gilda listGUID lfn:myfirst  
guid:2fdce224-af32-11d8-bca4-a3c587356e55
```

- **printInfo**
args: file (optional)
 - Print the information needed by the Replica Manager to screen or to a file.
- ```
>>edg-rm --vo gilda printInfo
```



# Catalog Command(2/2)

- **listReplicas**  
args: LFN | GUID | SURL
    - List all replicas of a file.
- ```
>> edg-rm --vo gilda listReplicas guid:2fdce224-af32-11d8-bca4-  
a3c587356e55  
sfn://grid015.pd.infn.it/shared/gilda/generated/2004-05-26/fileebdac483-  
af35-11d8-ae9c-8ac5558ff824  
sfn://grid009.ct.infn.it/flatfiles/gilda/generated/2004-05-26/file26b940d3-  
af32-11d8-bca4-a3c587356e55
```
- or
- ```
>> edg-rm --vo gilda listReplicas lfn:myfirst
sfn://grid015.pd.infn.it/shared/gilda/generated/2004-05-26/fileebdac483-
af35-11d8-ae9c-8ac5558ff824
sfn://grid009.ct.infn.it/flatfiles/gilda/generated/2004-05-26/file26b940d3-
af32-11d8-bca4-a3c587356e55
```

# Optimization Commands(1/2)

- **listBestFile**  
args: LFN | GUID
    - Return the 'best' replica for a given logical file identifier.
- ```
>> edg-rm --vo gilda listBestFile lfn:myfirst  
sfn://grid009.ct.infn.it/flatfiles/gilda/generated/2004-05-  
26/file26b940d3-af32-11d8-bca4-a3c587356e55  
or  
>> edg-rm --vo gilda listBestFile guid:2fdce224-af32-11d8-bca4-  
a3c587356e55  
sfn://grid009.ct.infn.it/flatfiles/gilda/generated/2004-05-  
26/file26b940d3-af32-11d8-bca4-a3c587356e55
```

Optimization Commands(2/2)

- **getBestFile**
args: LFN | GUID
 - Return the SURL that has the smallest access cost from the local store (or from the store specified by `-d` option).

```
>> edg-rm --vo gilda getBestFile lfn:myfirst  
sfn://grid009.ct.infn.it/flatfiles/gilda/generated/2004-05-  
26/file26b940d3-af32-11d8-bca4-a3c587356e55
```

or

```
>> edg-rm --vo gilda getBestFile guid:2fdce224-af32-11d8-bca4-  
a3c587356e55  
sfn://grid009.ct.infn.it/flatfiles/gilda/generated/2004-05-  
26/file26b940d3-af32-11d8-bca4-a3c587356e55
```

File Transfer Commands

- **copyFile**
args: source(LFN | GUID | SURL), dest(non-grid file)
option: stream, protocol, force
 - Copy a file to a non-grid destination.>>edg-rm -vo gilda **copyFile**
 lfn:myfirst file:///home/bob/analysis/data5.dat
- **list**
args: dir
 - List the directory contents on an SRM or a GridFTP server.>>edg-rm -vo gilda **list**
 sfn://tbed0101.cern.ch/flatfile/SE00/dteam

Data Management - Exercise A (1/2)

1. Create an empty local file to work with:
>>touch empty-local-file
2. Perform a copyAndRegisterFile command with:
**>>edg-rm --vo gilda copyAndRegisterFile
file:///home/miouser/empty-local-file -d grid009.ct.infn.it -l
lfn:myfirstDM
guid:b793f080-f417-11d7-b584-857330072702**
3. To check that this file exists, use the listReplicas command:
**>> edg-rm --vo gilda listReplicas guid:b793f080-f417-11d7-b584-
857330072702
sfn://grid015.pd.infn.it/shared/gilda/generated/2004-05-
27/file1b4d82b6-afb7-11d8-9264-97b731718da7**

Data Management - Exercise A (2/2)

4. Delete this file specifying the SURL to be deleted.
**>> edg-rm --vo gilda deleteFile
sfn://grid015.pd.infn.it/shared/gilda/generated/2004-05-
27/file1b4d82b6-afb7-11d8-9264-97b731718da7**
5. List all the Replicas of this file.
**>> edg-rm --vo gilda listReplicas lfn:mysecondDM
sfn://grid009.ct.infn.it/flatfiles/gilda/generated/2004-05-
27/file85b21a82-afb6-11d8-a830-ef8ecb195f0f**

Data Management Exercise B (1/2)

1. Execute the exercise 1 until step 3, for another file with
lfn:mySecondDM;
2. Verify this exists with:
>> **edg-rm --vo gilda listGUID lfn:mySecondDM**
guid:8b929ec3-afb6-11d8-a830-ef8ecb195f0f
3. To replicate this to another storage element
>> **edg-rm --vo gilda replicateFile guid:8b929ec3-afb6-11d8-a830-ef8ecb195f0f -d grid015.pd.infn.it**
sfn://grid015.pd.infn.it/shared/gilda/generated/2004-05-27/file1b4d82b6-afb7-11d8-9264-97b731718da7

Data Management - Exercise B (2/2)

4. Using `listReplicas` again shows the two distinct replicas:

```
>> edg-rm --vo gilda listReplicas lfn:mySecondDM
```

```
sfn://grid015.pd.infn.it/shared/gilda/generated/2004-05-  
27/file1b4d82b6-afb7-11d8-9264-97b731718da7  
sfn://grid009.ct.infn.it/flatfiles/gilda/generated/2004-05-  
27/file85b21a82-afb6-11d8-a830-ef8ecb195f0f
```