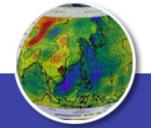




The gLite Data Management System

Giuseppe LA ROCCA
INFN Catania
giuseppe.larocca@ct.infn.it







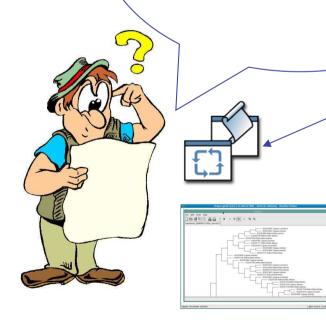


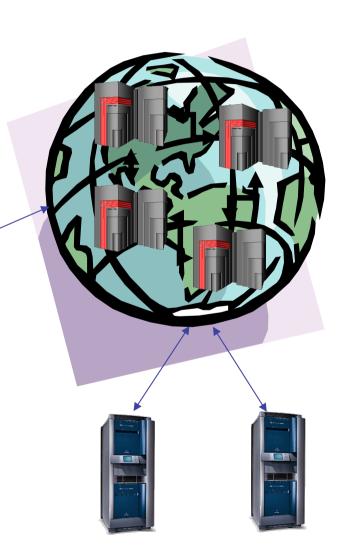
- Storage & Protocols
 - Types of Storage
- The Storage Resource Manager (SRM)
- Grid file referencing schemes
- LFC File Catalogue
 - Architecture
 - LFC commands
- File & Replica Management Client Tools
- File Transfer Service (FTS)
- References
- Hands on

www.euasiagrid.or



I need to collect data from the GRID Storage Elements to run my application. How can I do?





- www.euasiagrid.org
- The Storage Element is the service which allows a user or an application to store data for future retrieval.
- All data in a SE must be considered read-only and therefore can not be changed unless physically removed and replaced.

Protocol	Type	GSI secure	Description	Optional
GSIFTP	File Transfer	Yes	FTP-like	No
gsidcap	File I/O	Yes	Remote file access	Yes
insecure RFIO	File I/O	No	Remote file access	Yes
secure RFIO (gsirfio)	File I/O	Yes	Remote file access	Yes

- The GSIFTP protocol offers the functionalities of FTP, but with support for GSI. It is responsible for secure, fast and efficient file transfers to/from Storage Elements.
- RFIO was developed to access tape archiving systems, such as CASTOR (CERN Advanced STORage manager) and it comes in a secure and an insecure version.
- The gsidcap protocol is the GSI enabled version of the dCache native access protocol, dcap.

- In WLCG/EGEE, different types of Storage Elements are available:
- CASTOR. It consists in a disk buffer frontend to a tape mass storage system. A virtual file system (namespace) shields the user from the complexities of the disk and tape underlying setup. File migration between disk and tape is managed by a process called "stager". The native storage protocol, the insecure RFIO, allows access of files in the SE. Since the protocol is not GSIenabled, only RFIO access from a location in the same LAN of the SE is allowed. With the proper modifications, the CASTOR disk buffer can be used also as disk-only storage system.

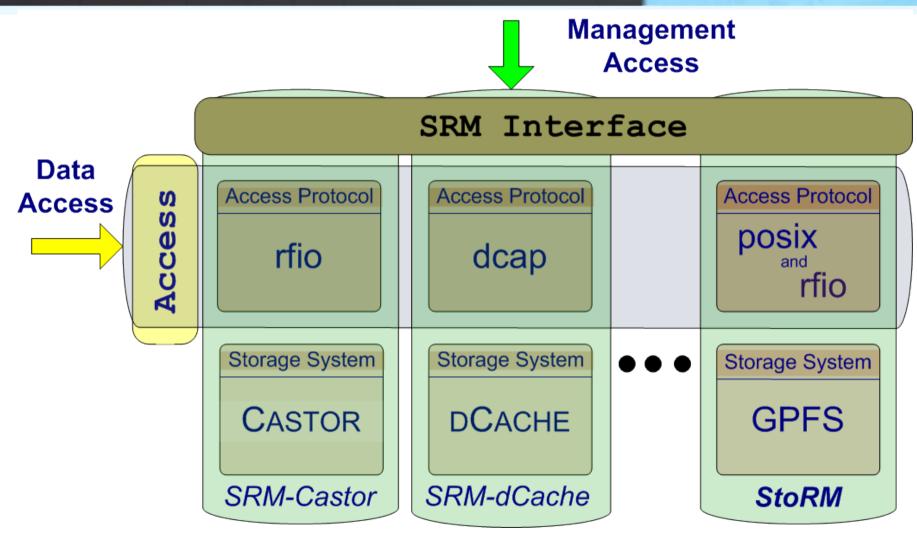
- StoRM. It has been designed to support space reservation and direct access (native POSIX I/O call), as well as other standard libraries (like RFIO).
- StoRM takes advantage from high performance parallel file systems like GPFS (from IBM).
 - In addition, standard POSIX file systems are supported (XFS from SGI and ext3).
- StoRM takes advantage of ACL support provided by the underlying file systems to implement the security models

- dCache. It consists of a server and one or more pool nodes. The server represents the single point of access to the SE and presents files in the pool disks under a single virtual file system tree. Nodes can be dynamically added to the pool. The native gsidcap protocol allows POSIX-like data access. dCache is widely employed as disk buffer frontend to many mass storage systems, like HPSS and Enstore, as well as a disk-only storage system.
- LCG Disk pool manager. It's a lightweight disk pool manager, suitable for relatively small sites (max 10 TB of total space). Disks can be added dynamically to the pool at any time. Like in dCache and CASTOR, a virtual file system hides the complexity of the disk pool architecture. The secure RFIO protocol allows file access from the WAN.



www.euasiagrid.org







The Storage Resource Manager

The Storage Resource Manager (SRM) has been designed to be the single interface for the management of disk and tape storage resources.

Any type of Storage Element in WLCG/EGEE offers an SRM interface except for the Classic SE, which is being phased out.

- SRM hides the complexity of the resources setup behind it and allows the user to request files, keep them on a disk buffer for a specified lifetime, reserve space for new entries, and so on.
 - In gLite, interactions with the SRM is hidden by high level services (DM tools and APIs)

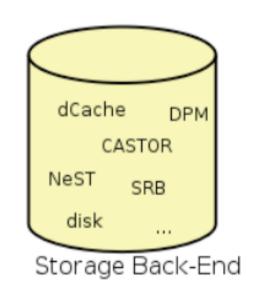


Native I/O Interface

dcap rfio chirp xio nfs

SRM Interface

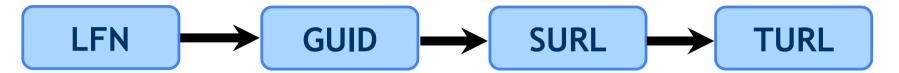
File Transfer Interface
GridFTP
http/https



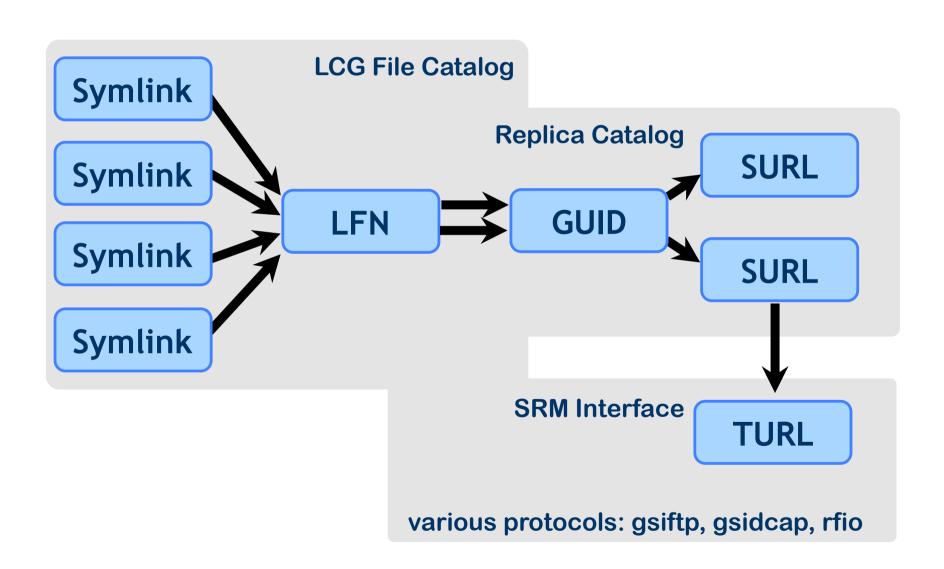


www.euasiagrid.or





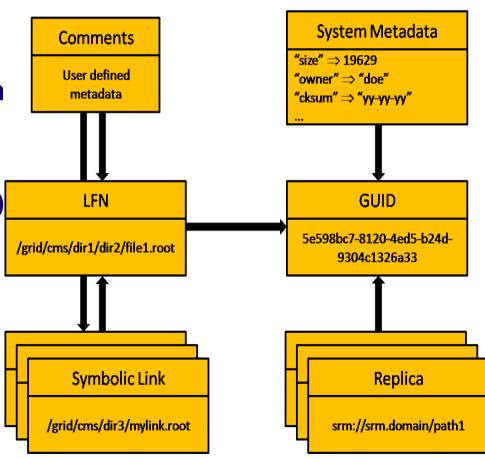
- Logical File Name (LFN)
 - Ifn:/grid/gilda/input-file
- Grid Unique IDentifier (GUID)
 - guid:4d57edef-fa5c-4512-a345-1c838916b357
- Storage URL (for a specific replica, on a specific Storage Element)
 - srm://aliserv6.ct.infn.it/gilda/generated/2007-11-13/ fileb366f371-b2c0-485d-b12c-c114edaf4db4
 - sfn://se01.athena.hellasgrid.gr/data/dteam/doe/file1
- Transport URL (for a specific replica, on an SE, with a specific protocol)
 - gsiftp://aliserv6.ct.infn.it/gilda/generated/2007-11-13/ fileb366f371-b2c0-485d-b12c-c114edaf4db4



- Users and applications need to locate files (or replicas) on the Grid. The LCG File Catalogue is the service which maintains mappings between LFN(s), GUID and SURL(s).
- The catalogue publishes its endpoint in the Information Service so that it can be discovered by Data Management tools and other services (the WMS for example).
- It consists of a unique catalogue, where the LFN is the main key. Further LFNs can be added as symlinks to the main LFN.
 - System metadata are supported, while for user metadata only a single string entry is available



- LFN acts as main key in the database.
 It has:
 - Symbolic links to it (additional LFNs)
 - System metadata
 - Information on replicas
 - One field of user metadata
 - Access Control Lists
 - Integration with VOMS (VirtualID and VirtualGID)
 - C API language





- User can interact with the file catalogue through CLIs and C APIs.
 - The environment variable LFC_HOST
 (e.g.: LFC_HOST=gilda-lfc.ct.infn.it)
 must contains the host name of the LFC server
 to be used.
- The directory structure of the LFC namespace has the form: /grid/<VO>/<subpaths>
 - Users of a given VO will have read and write permissions only under the corresponding
 <VO> subdirectory.



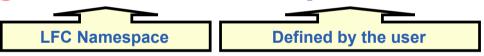
LFC Commands

www.euasiagrid.or

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



- Listing the entries of a LFC directory
 - IfC-IS [-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/tutorials/taipei02
export LFC_HOME=/grid/gilda/tutorials
lfc-ls -l taipei02
lfc-ls -l -R /grid
```

- Creating directories in the LFC
 - Ifc-mkdir [-m mode] [-p] path...
- Where path specifies the LFC pathname
- Remember that while registering a new file (using lcgcr, for example) the corresponding destination directory must be created in the catalog beforehand.
- Examples:

lfc-mkdir /grid/gilda/<YOUR DIRECTORY>

Created by the user



- Creating a symbolic link
 - Ifc-In -s file linkname
 - Ifc-In -s directory linkname
 - Create a link to the specified file or directory with linkname

Examples:

- lfc-ln -s /grid/gilda/test /grid/gilda/aLink

Original File Symbolic Link

Let's check the link using Ifc-Is with long listing

- lfc-ls -l aLink
lrwxrwxrwx 1 19122 1077 0 Jun 14 11:58 aLink > /grid/gilda/test

- LFC allows to attach to a file or directory an access control list (ACL), a list of permissions which specify who is allowed to access or modify it. The permissions are very much like those of a UNIX file system: read (r), write (w) and execute (x).
- In LFC, users and groups are internally identified as numerical virtual uids and virtual gids, which are virtual in the sense that they exist only in the LFC namespace.
 - A user can be specified as a name, as a virtual uid or as a DN.
 - A group can be specified as name, as a virtual gid or as a VOMS FQAN.
- A directory in LFC has also a default ACL (which is the ACL associated to any file or directory being created under that directory). After creation, the ACLs can be freely changed.
 - When creating a sub-directory, its default ACL is inherited from the parent directory

\$ lfc-getacl /grid/gilda/tutorials/test-acl

```
# file: /grid/gilda/tutorials/test-acl
# owner: /C=IT/O=INFN/OU=Personal Certificate/L=Catania/
CN=Giuseppe La Rocca/Email=giuseppe.larocca@ct.infn.it
# group: gilda
user::rwx
group::rwx #effective:rwx
other::r-x
default:user::rwx
default:group::rwx
default:other::r-x
```

In this example, the owner and all users in the gilda group have full privileges to the directory, while other users cannot write into it.



lfc-setacl [-d] [-m] [-s] acl_entries path

The -m option means that we are modifying the existing ACL. Other options of lfc-setacl are -d to remove ACL entries, and -s to replace the complete set of ACL entries.

acl_entries is a coma separated list of entries. Each entry has colon separated fields: ACL type, id (uid or gid), permission. Only directories can have default ACL entries!

The entries look like:

user::perm

user:uid:perm

group:perm

group:gid:perm

mask:perm

other:perm

defaul::user:perm

defaul::user:uid:perm

defaul::group:perm

defaul::group:gid:perm

default::mask:perm

deafult::other:perm



Modify the ACL of a directory

Lets's change default ACL, with read/write permission for user and group, and no privileges for others.

The syntax we apply here is modify (-m) default (d:) for user (u:), and the same of course for group and others.



Adding metadata information

The lfc-setcomment and lfc-delcomment commands allow the user to associate a comment with a catalogue entry and delete such comment. This is the only user-defined metadata that can be associated with catalogue entries.

The comments for the files may be listed using the --comment option of the Ifc-Is command. This is shown in the following example:

- \$ lfc-setcomment /grid/gilda/file1 "My metadata"
- \$ lfc-ls --comment /grid/gilda/file1
 /grid/gilda/file1 My metadata

 The LCG Data Management tools allow users to copy files between UI, WN and a SE, to register entries in the file catalogue and replicate files between SEs.

lcg-cp	Copies a Grid file to a local destination
lcg-cr	Copies a file to a SE and registers it in the catalogue
lcg-del	Deletes one file (either one replica or all the replicas)
lcg-rep	Copies a file from one SE to another SE and registers it in the catalogue
lcg-gt	Gets the TURL for a given SURL and transfer protocol
lcg-aa	Adds an alias in the catalogue for a given GUID
lcg-ra	Removes an alias in the catalogue for a given GUID
lcg-rf	Registers in the catalogue a file residing on a SE
lcg-uf	Unregisters in the catalogue a file residing on a SE
lcg-la	Lists the aliases for a given LFN, GUID or SURL
lcg-lg	Gets the GUID for a given LFN or SURL
lcg-lr	Lists the replicas for a given LFN, GUID or SURL

The --vo <vo name> option, to specify the virtual organisation of the user, is present in all commands, except for lcg-gt. Its usage is mandatory unless the variable LCG_GFAL_VO is set (e.g.: export LCG_GFAL_VO=gilda)

Timeouts

The commands lcg-cr, lcg-del, lcg-gt, lcg-rf, lcg-sd and lcg-rep all have timeouts implemented.

By using the option -t, the user can specify a number of seconds for the timeout.

The default is 0 seconds, that is no timeout.

If we got a times out during the performing of an operation, all actions performed till that moment are rolled back, so no broken files are left on a SE and no existing files are not registered in the catalogues.

 For all Icg-* commands to work, the environment variable LCG_GFAL_INFOSYS must be set to point to a top BDII in the format <hostname>:<port>, so that the commands can retrieve the necessary information

```
export LCG_GFAL_INFOSYS=gilda-bdii.ct.infn.it:2170
```

 The VO_<VO>_DEFAULT_SE variable specifies the default SE for the VO. \$ lcg-cr --vo gilda -d aliserv6.ct.infn.it \
file:/home/larocca/file1

guid: 6ac491ea-684c-11d8-8f12-9c97cebf582a

where the only argument is the local file to be uploaded and the -d <destination> option indicates the SE used as the destination for the file. The command returns the file GUID.

If no destination is given, the SE specified by the VO_<VO>_DEFAULT_SE environmental variable is taken.

The -P option allows the user to specify a relative path name for the file in the SE. If no -P option is given, the relative path is automatically generated.





Uploading a file to the Grid /2

www.euasiagrid.org

The following are examples of the different ways to specify a destination:

```
-d aliserv6.ct.infn.it
-d srm://aliserv6.ct.infn.it/data/gilda/my_file
-d aliserv6.ct.infn.it -P my_dir/my_file
```

The -1 < 1fn > option can be used to specify a LFN:

guid:db7ddbc5-613e-423f-9501-3c0c00a0ae24

The -g option allows to specify a GUID (otherwise automatically created):

```
$ lcg-cr --vo gilda -d aliserv6.ct.infn.it \
-g guid:baddb707-0cb5-4d9a-8141-a046659d243b \
file: 'pwd'/file2
```

guid:baddb707-0cb5-4d9a-8141-a046659d243b



Attention!

This option should not be used except for expert users and in very particular cases. Because the specification of an existing GUID is also allowed, a misuse of the tool may end up in a corrupted GRID file in which replicas of the same file are in fact different from each other.



www.euasiagrid.org



```
$ lcg-rep -v --vo gilda -d <SECOND SE> \
guid:db7ddbc5-613e-423f-9501-3c0c00a0ae24
Source URL:
sfn://aliserv6.ct.infn.it/data/gilda/larocca/file1
File size: 30
Destination specified: <SECOND SE>
Source URL for copy:
gsiftp://aliserv6.ct.infn.it/data/gilda/larocca/file1
Destination URL for copy:
gsiftp://<SECOND SE>/data/gilda/generated/2004-07-09/
  file50c0752c-f61f-4bc3-b48e-af3f22924b57
# streams: 1
Transfer took 2040 ms
Destination URL registered in LRC: srm://<SECOND SE>/
  data/gilda/generated/2004-07-09/file50c0752c-
  f61f-4bc3-b48e-af3f22924b57
```

www.euasiagrid.org EUAsiaGr

```
$ lcg-lr --vo gilda \
lfn:/grid/gilda/tutorials/larocca/my_alias1

srm://aliserv6.ct.infn.it/data/gilda/generated/
2004-07-09/file79aee616-6cd7-4b75-8848-f091
srm://<SECOND_SE>/data/gilda/generated/2004-07-08/
file0dcabb46-2214-4db8-9ee8-2930
```

Again, a LFN, the GUID or a SURL can be used to specify the file.

The lcg-lg command returns the GUID associated with a specified LFN or SURL.

```
$ lcg-lg --vo gilda lfn:/grid/gilda/test.txt
guid:cf93526e-807a-43a6-9262-f55a3989623c
```



Listing replicas, GUIDs and aliases /2

The lcg-la command can be used to list the LFNs associated with a particular file, which can be identified by its GUID, any of its LFNs, or the SURL of one of its replicas:

```
$ lcg-la --vo gilda \
guid:cf93526e-807a-43a6-f55a3989623c
```

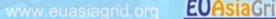
lfn:/grid/gilda/test.txt

The lcg-aa (add alias) command allows the user to add a new LFN to an existing GUID:

```
$ lcg-aa --vo gilda \
  guid:baddb707-0cb5-4d9a-8141-a046659d243b lfn:/
  grid/gilda/new_alias
```

The lcg-ra command (remove alias) allows a user to remove an LFN from an existing GUID:

```
$ lcg-ra --vo gilda \
  guid:baddb707-0cb5-4d9a-8141-a046659d243b lfn:/
  grid/gilda/my alias1
```





```
$ lcg-cp --vo gilda -t 100 -v lfn:/grid/gilda/
  tutorials/pippo.txt file:/tmp/pippo.txt
Source URL: lfn:/grid/gilda/pippo.txt
File size: 104857600
Source URL for copy:
gsiftp://aliserv6.ct.infn.it:/storage/gilda/2007-07-06/
  input2.dat.10.0
Destination URL: file:///tmp/myfile
# streams: 1
# set timeout to 100 (seconds)
  85983232 bytes 8396.77 KB/sec avg 9216.11
Transfer took 12040 ms
```



A file stored on a SE and registered in LFC can be deleted using the lcg-del command.

- If a SURL is provided as argument, then that particular replica will be deleted.
- If a LFN or GUID is given instead then the -s <SE>
 option must be used to indicate which one of the
 replicas must be erased
- \$ lcg-del --vo gilda -s aliserv6.ct.infn.it \
 quid:91b89dfe-ff95-4614-bad2-c538bfa28fac

• If the -a option is used, all the replicas of the given file will be deleted and unregistered from the catalog.

```
$ lcg-del --vo gilda -a \
guid:91b89dfe-ff95-4614-bad2-c538bfa28fac
```



The lcg-rf command allows to register a file physically present in a SE, creating a GUID-SURL mapping in the catalogue.

The -g <GUID> option allows to specify a GUID (otherwise automatically created).

```
$ lcg-rf --vo gilda \
  -g guid:baddb707-0cb5-4d9a-8141-a046659d243b \ srm://
  aliserv6.ct.infn.it/data/gilda/generated/2004-07 08/
  file0dcabb46-2214-4db8-9ee8-2930de1
```

guid:baddb707-0cb5-4d9a-8141-a046659d243b

Icg-uf allows to delete a GUID-SURL mapping (respectively the first and second argument of the command) from the catalogue:

```
$ lcg-uf --vo gilda \
  guid:baddb707-0cb5-4d9a-8141-a046659d243b \
  srm://aliserv6.ct.infn.it/data/gilda/generated/
  2004-07 08/file0dcabb46-2214-4db8-9ee8-2930de1
```

If the last replica of a file is unregistered, the corresponding GUID-LFN mapping is also removed.



Attention!

It <u>does not remove</u> any physical replica from the SE.

www.euasiagrid.org

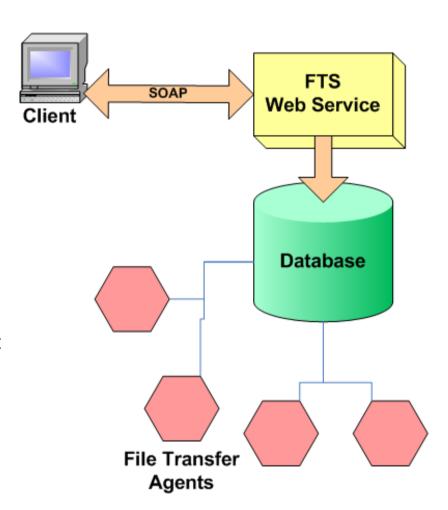
- The File Transfer Service (FTS) is the lowest-level data movement service defined in gLite.
 - It is responsible for moving sets of files from one site to another.
 - It is designed for point to point movement of physical files (no file routing via intermediate storage).
 - The FTS has dedicated interfaces for managing the network resource and to display statistics of ongoing transfers.
 - The FTS handles internally the SRM negotiation between the source and destination SEs and the management of the underlying GridFTP transfers.





 The clients. The FTS client libraries or command line tools are used by the applications to communicate with the FTS.

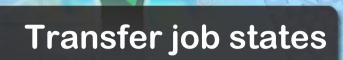
- •
- The FTS Web Service. The web service component is implemented as a Tomcat web application. This is a proper web service that implements the WSDL as defined by the FTS interface. Based on the Web Service Description Language document, anyone can build their own clients in their own preferred language. The web service connects to the database through JDBC using a Tomcat database connection pool.
- The FTS Database. There is a MySQL and an Oracle implementation of the FTS schema. This is the only persistency point in the system and it scales only as well as the corresponding backend allows (Oracle scales better than MySQL, obviously). There can be only a single instance of this database for a given FTS instance.
- The File Transfer Agents. It is an agent that actually performs the file transfers on the <u>channels</u> that the FTS manages.

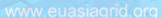


- **Transfer Job:** a set of files to be transferred in a source/ destination pair format;
- File: a source/destination SURL pair to be transferred;
- Job State: a function of the individual file states constituting the Job;
- File State: the state of an individual file transfer;
- Channel: a specific network pipe used for file transfers.
 - Production channels are high bandwidth, dedicated network pipe between Tier-0, Tier-1's and other major Tier-2's centers.
 - Non-production channels are assigned typically to open networks and do not guarantee a minimum throughput as production channels do.

Basic Concepts /2

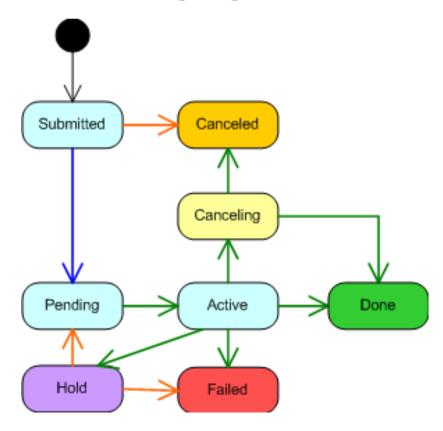
- The transfer jobs are processed asynchronously (batch mode). Upon submission, a job identifier is returned to the user.
- This identifier can be used to query the status of the job as it progresses through the system or cancel the job.
- Once a job has been submitted to the system it is assigned to a transfer channel based on the SEs containing the source and the destination.







- Submitted: the job has been submitted to FTS but not yet assigned to a channel
- Pending: the job has been assigned to a channel and files are waiting for being transferred
- Active: the transfer for some of the job's files is ongoing
- Canceling: the job is being canceled
- Canceled: the job has been canceled
- Done: all files in a job were successfully transferred
- Failed: some file transfers in a job have failed
- Hold: the job has aborted and requires manual interventions (moving it to Pending or Failed)





www.euasiagrid.org



 Before submitting a job, the user is expected to upload an appropriate password-protected long-term proxy to the MyProxy server used by FTS.

The following user-level commands for submitting, querying and canceling jobs are described here:

glite-transfer-submit
glite-transfer-status

Submits a transfer job
Displays the status of an ongoing
transfer job

glite-transfer-list

Lists all submitted transfer jobs owned by the user

glite-transfer-cancel

Cancels a transfer job

Some additional administrative commands are described here

glite-transfer-channel-add

Create a new channel with defined parameters on

FTS

glite-transfer-channel-list

Displays details of the given channel defined on FTS

glite-transfer-channel-set

Allows administrator to set a channel 'Active' or 'Inactive'



 Once a user has successfully registered a long-term proxy to a MyProxy server, he can submit a transfer job. He can do it either by specifying the source-destination pair in the command line:

```
$ myproxy-init -s myproxy-fts.cern.ch -d

$ glite-transfer-submit -m myproxy-fts.cern.ch \
-s https://w-fts.grid.sinica.edu.tw:8443/sc3/glite-data-
transfer-fts/services/FileTransfer \

srm://srm.sara.nl/pnfs/srm.sara.nl/data/source_file \
srm://srm.cnaf.infn.it/castor/cnaf.infn.it/grid/destination

Enter MyProxy password:
Enter MyProxy password again:

c2e2cdb1-a145-11da-954d-944f2354a08b
```

Querying the job status

 The following example shows a query to FTS to infer information about the state of a transfer job:

```
$ glite-transfer-status \
-s https://w-fts.grid.sinica.edu.tw:8443/sc3/glite-data-
transfer-fts/services/FileTransfer \
-l c2e2cdb1-a145-11da-954d-944f2354a08b

Source: srm://srm.grid.sara.nl/pnfs/grid.sara.nl/data/
source_file
Destination: srm://sc.cr.cnaf.infn.it/castor/cnaf.infn.it/
grid/destination

State: Pending
Retries: 0
Reason: (null)
Duration: 0
```



Listing and Canceling data transfer

• Listing...

```
$ glite-transfer-list \
-s https://w-fts.grid.sinica.edu.tw:8443/sc3/glite-data-
transfer-fts/services/FileTransfer
...
c2e2cdb1-a145-11da-954d-944f2354a08b Pending
```

• Cancelling..

```
$ glite-transfer-cancel \
-s https://w-fts.grid.sinica.edu.tw:8443/sc3/glite-data-
transfer-fts/services/FileTransfer \
c2e2cdb1-a145-11da-954d-944f2354a08b
```

gLite 3 User Guide – Manual Series

https://edms.cern.ch/file/722398/1.2/gLite-3-UserGuide.pdf

- 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
- File Transfer Services
 - http://cern.ch/egee-jra1-dm/FTS/default.htm
- LFC and DPM documentation
 - https://uimon.cern.ch/twiki/bin/view/LCG/ DataManagementDocumentation

- Connect to the training infrastructure using the information reported in the tutorial sheet
- Run the hands-on available in this web link:
- http://www.euasiagrid.org/wiki/index.php/
 Data_Management
- Enjoy!