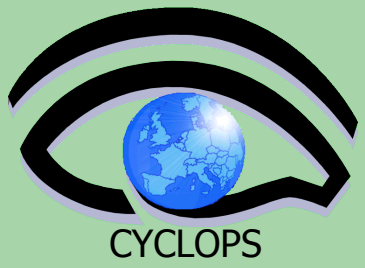
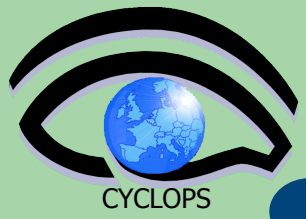
The AMGA Metadata Service

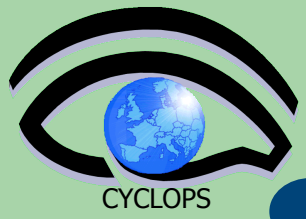


Antonio Calanducci
National Institute of Nuclear Physics
INFN Catania
EGEE NA3 Training & Dissemination
CYCLOPS Second Training Workshop
Chania (Crete), 05th-07th May 2008



Contents

- Metadata services background and motivation
- Architecture and features of AMGA
- Grid DB Access with AMGA
- Use cases



Why Grid needs Metadata?

Why Grid needs Metadata?

- Grids allow to save millions of files spread over several storage sites.

Why Grid needs Metadata?

- 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

Why Grid needs Metadata?

- 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

Basic Metadata Concept

Basic Metadata Concept

- **Entries** – Representation of real world entities which we are attaching metadata to for describing them

Basic Metadata Concept

- **Entries** – Representation of real world entities which we are attaching metadata to for describing them
- **Attribute** – key/value pair

Basic Metadata Concept

- **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,…)

Basic Metadata Concept

- **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

Basic Metadata Concept

- **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

Basic Metadata Concept

- **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

Basic Metadata Concept

- **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

Basic Metadata Concept

- **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

Analogies

Analogies

- Analogy to the RDBMS world:
  - **schema ←→ table schema**
  - **collection ←→ db table**
  - **attribute ←→ schema column**
  - **entry ←→ table row/record**

Analogies

- Analogy to the RDBMS world:
  - **schema ←→ table schema**
  - **collection ←→ db table**
  - **attribute ←→ schema column**
  - **entry ←→ table row/record**

Analogies

- Analogy to the RDBMS world:
  - **schema ←→ table schema**
  - **collection ←→ db table**
  - **attribute ←→ schema column**
  - **entry ←→ table row/record**

- Analogy to a file system:
  - **Collection ←→ Directory**
  - **Entry ←→ File**

CYCLOPS

Example: Movie Trailers

Information Society and Media

Example: Movie Trailers

- Movie trailers files (entries) saved on Grid Storage Elements and registered into File Catalogue

Example: Movie Trailers

- Movie trailers files (entries) saved on Grid Storage Elements and registered into File Catalogue
- We want to add metadata to describe movie content.

Example: Movie Trailers

- 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

Example: Movie Trailers

- 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

Trailer's example

Entry names	Title	Ru	Cast	LFN
8c3315c1-811f-4823-a778-60a203439689	My Best Friend's wedding	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

Trailer's example

Attribute

Entry names	Title	Ru	Cast	LFN
8c3315c1-811f-4823-a778-60a203439689	My Best Friend's wedding	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

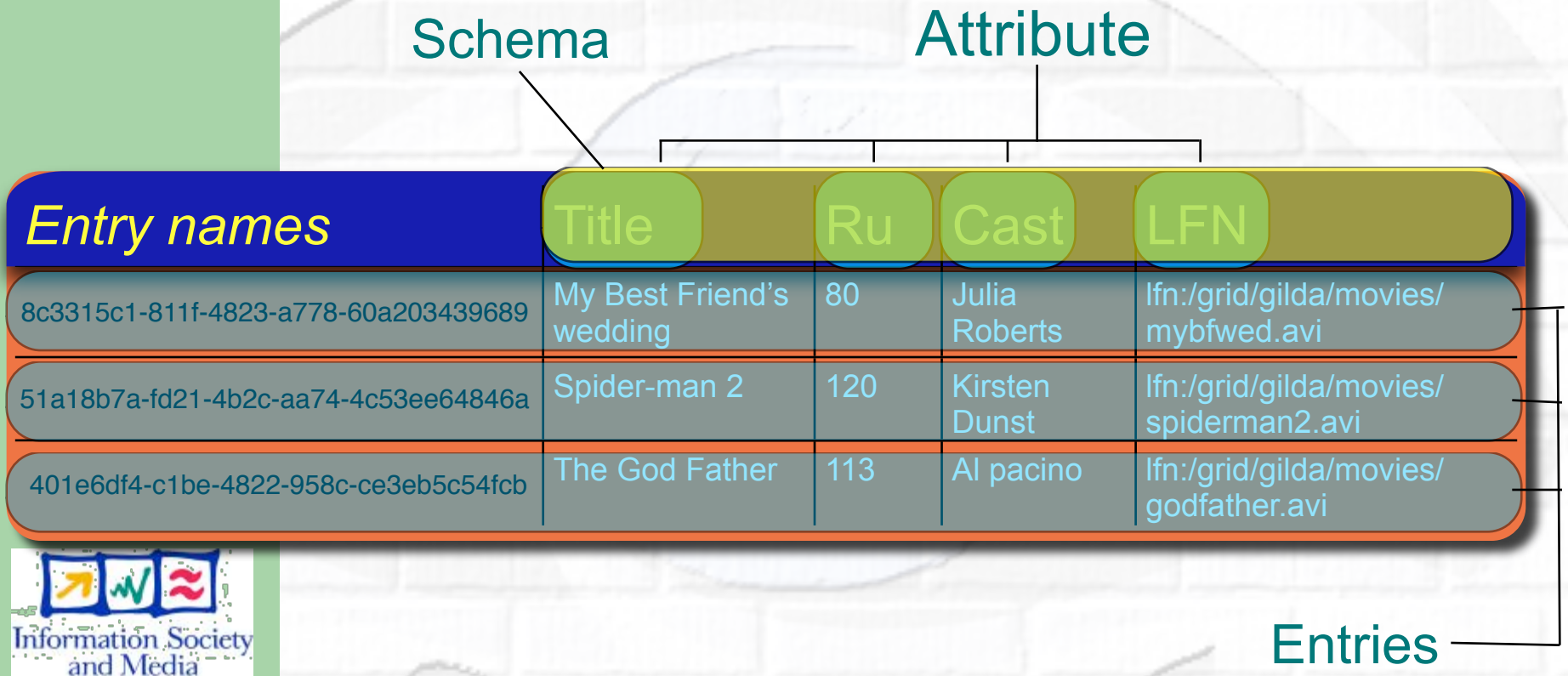
Trailer's example

Schema

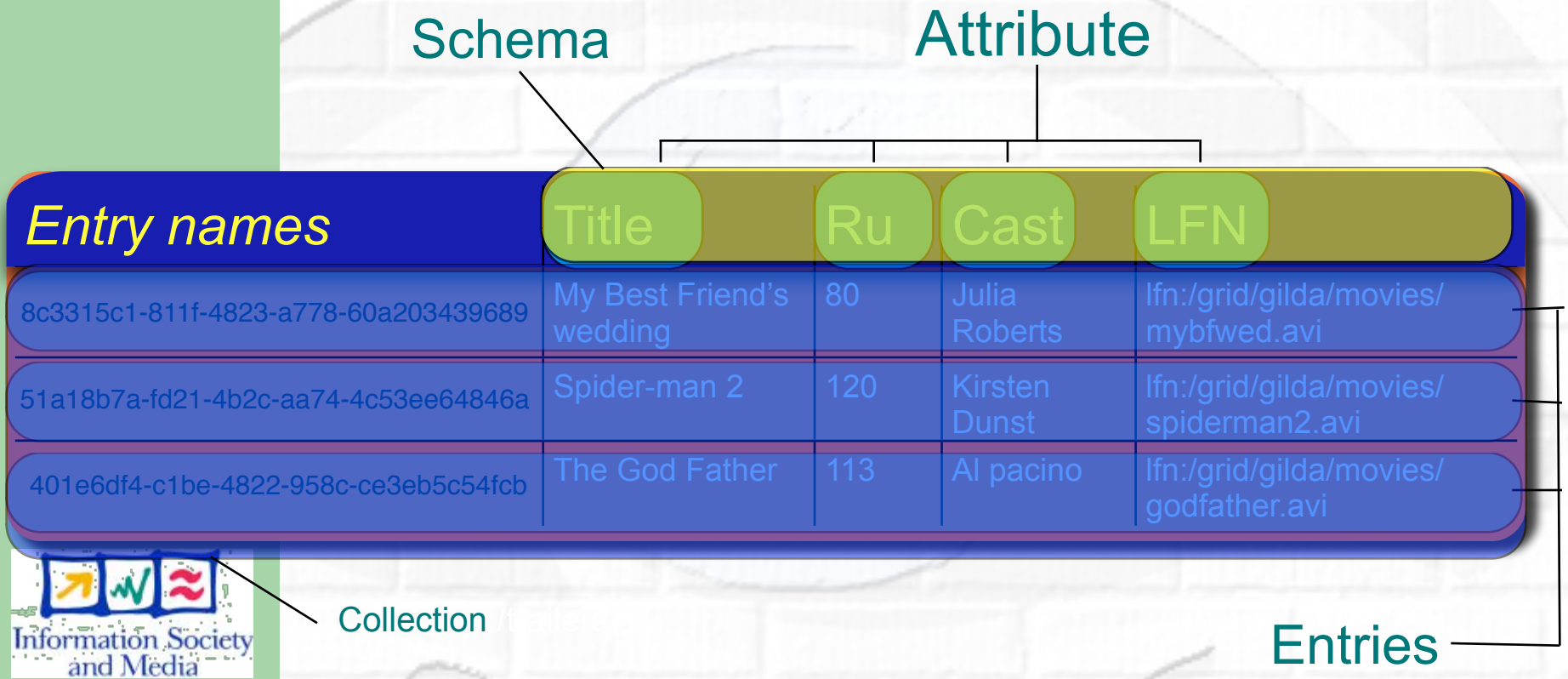
Attribute

Entry names	Title	Ru	Cast	LFN
8c3315c1-811f-4823-a778-60a203439689	My Best Friend's wedding	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

Trailer's example



Trailer's example

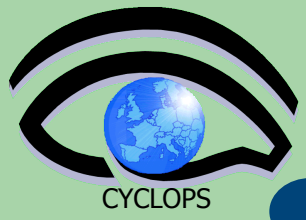


Example

```
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
```



Metadata on the Grid

Metadata on the Grid

- Information about files -- but not only!

Metadata on the Grid

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

Metadata on the Grid

- 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

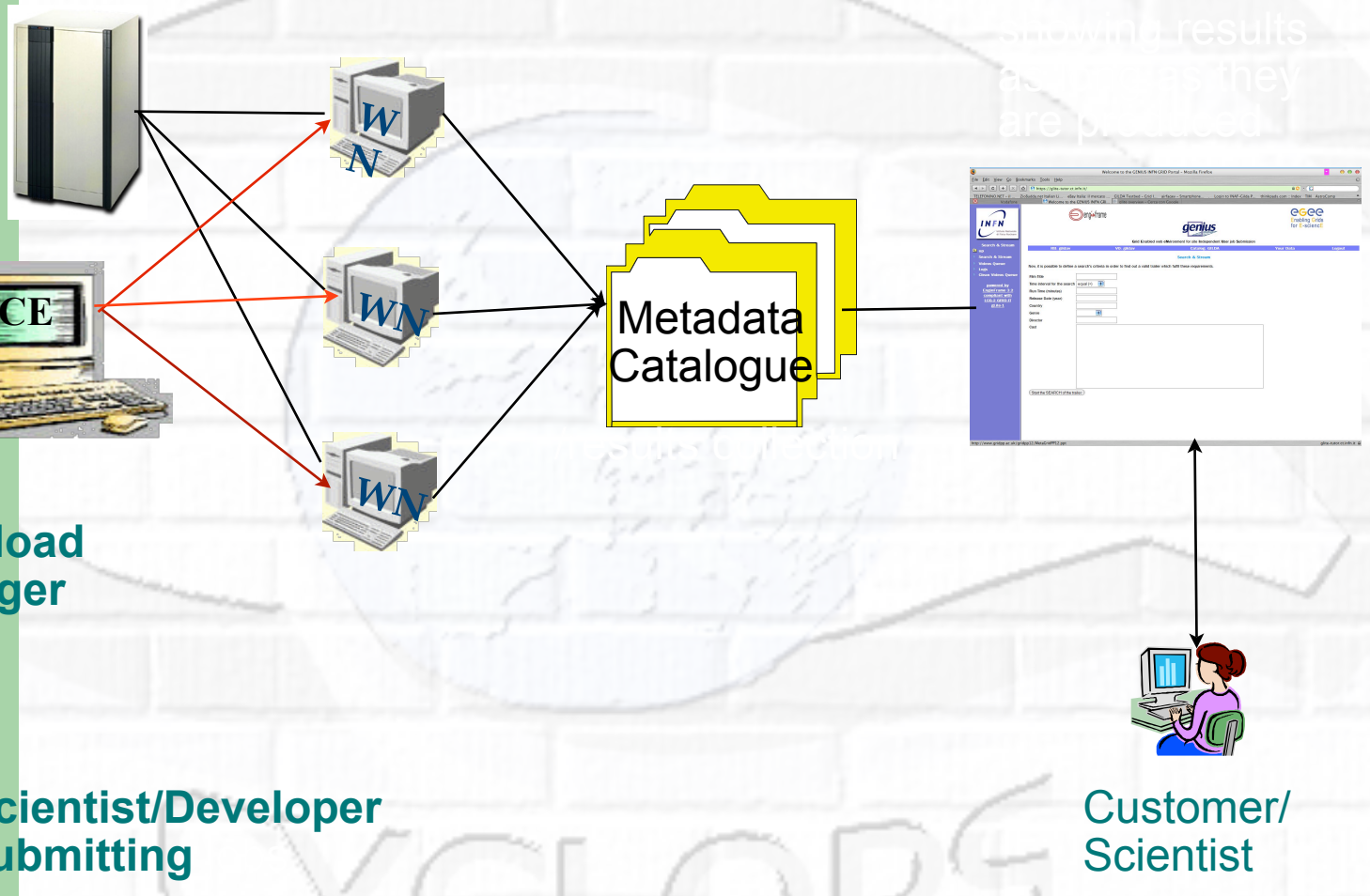
Metadata on the Grid

- 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 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”

Metadata on the Grid

- 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 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

Monitoring of running application



CYCLOPS

Use a Metadata services to exchange data among running jobs

Information Society and Media

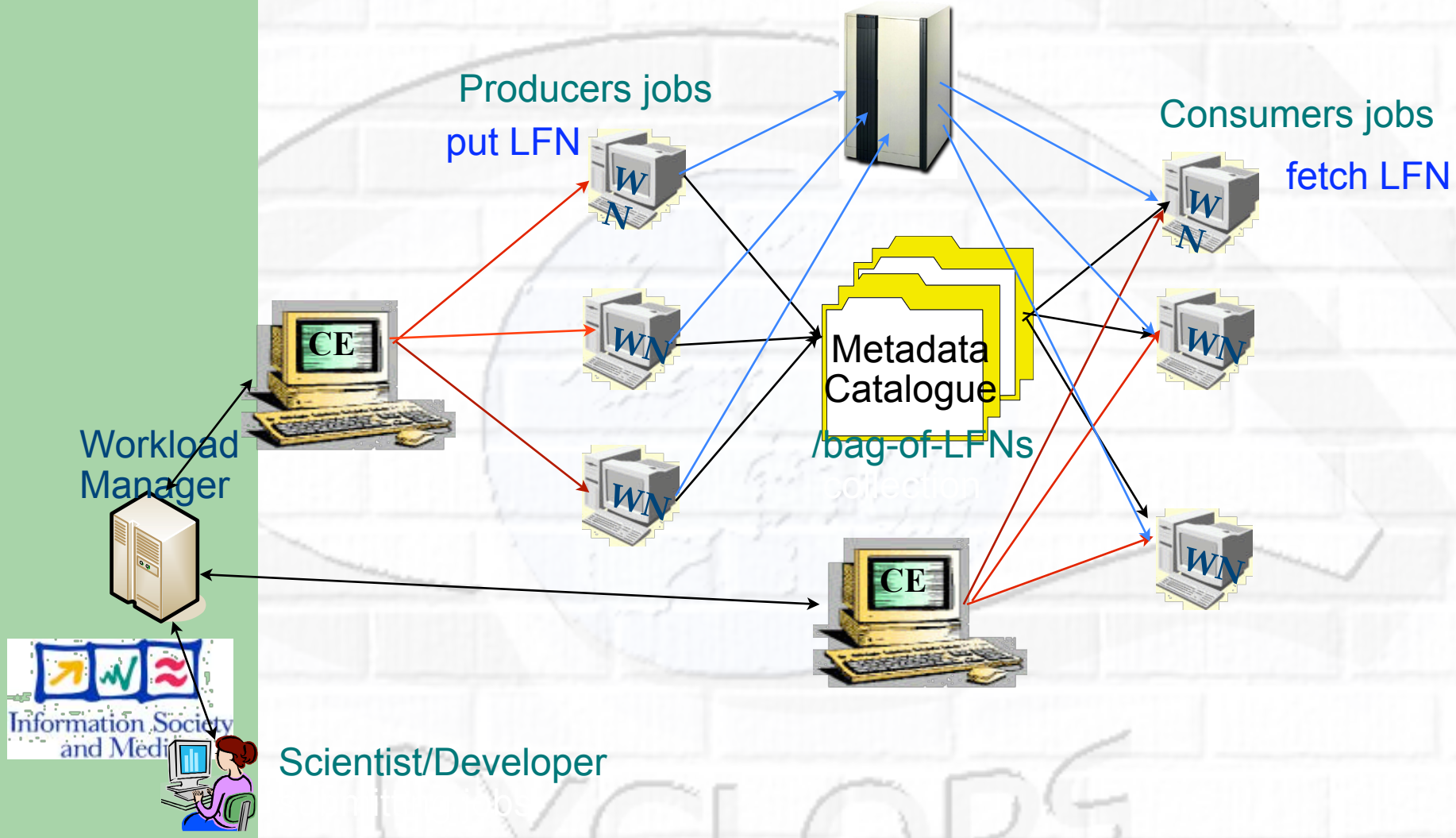
Use a Metadata services to exchange data among running jobs

- 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

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
- 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

Information exchanging among grid peers



The AMGA Metadata Catalogue

The AMGA Metadata Catalogue

- Official metadata service for the gLite middleware

The AMGA Metadata Catalogue

- Official metadata service for the gLite middleware
- AMGA: Arda Metadata Grid Application

The AMGA Metadata Catalogue

- Official metadata service for the gLite middleware
- AMGA: Arda Metadata Grid Application
- Provide a complete but simple interface, in order to make all users able to use it easily.

The AMGA Metadata Catalogue

- Official metadata service for the gLite middleware
- 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

The AMGA Metadata Catalogue

- Official metadata service for the gLite middleware
- 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
- Grid security is provided to grant different access levels to different users.

The AMGA Metadata Catalogue

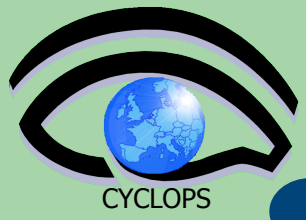
- Official metadata service for the gLite middleware
- AMGA: Arda Metadata Grid Application
- Provide a complete but simple interface, in order to make all users able to use it easily.
- Designed with scalability in mind in order to deal with large number of entries
- Grid security is provided to grant different access levels to different users.
- Flexible with support to dynamic schemas in order to serve several application domains

The AMGA Metadata Catalogue

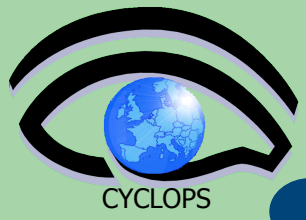
- Official metadata service for the gLite middleware
- AMGA: Arda Metadata Grid Application
- Provide a complete but simple interface, in order to make all users able to use it easily.
- Designed with scalability in mind in order to deal with large number of entries
- Grid security is provided to grant different access levels to different users.
- Flexible with support to dynamic schemas in order to serve several application domains
- Allow hierarchical metadata schemas

The AMGA Metadata Catalogue

- Official metadata service for the gLite middleware
- 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
- 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
- Allow hierarchical metadata schemas



AMGA Features



AMGA Features

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

AMGA Features

- **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

AMGA Features

- **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

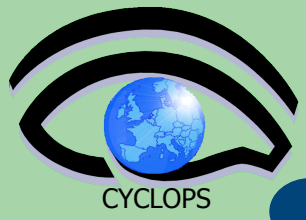
```
selectattr /gLibrary:FileName /gLAudio:Author /gLAudio:Album
'/gLibrary:FILE=/gLAudio:FILE and like(/gLibrary:FileName, "%.mp3")'
```

AMGA Features

- **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**



AMGA Security

AMGA Security

- Unix style permissions - users and groups

AMGA Security

- Unix style permissions - users and groups
- ACLs – Per-collection or per-entry.

AMGA Security

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

AMGA Security

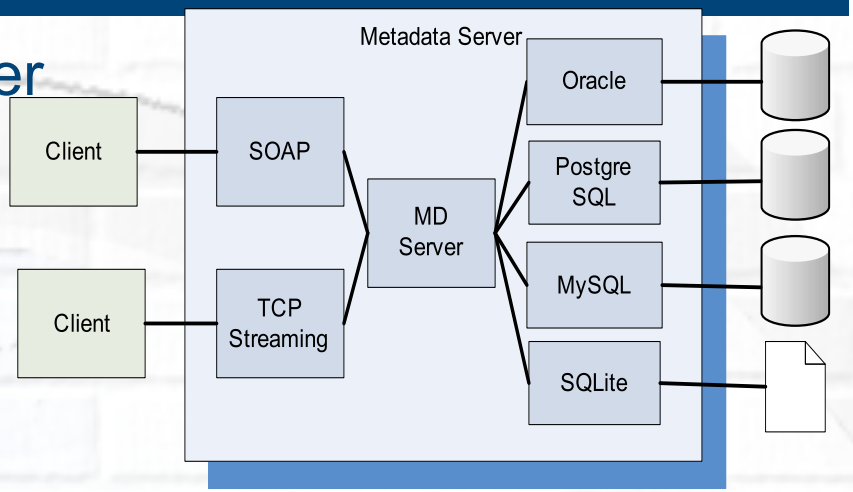
- Unix style permissions - users and groups
- ACLs – Per-collection or per-entry.
- Secure client/server connections – SSL
- Client Authentication based on
  - Username/password
  - General X509 certificates (DN based)
  - Grid-proxy certificates (DN based)

AMGA Security

- Unix style permissions - users and groups
- ACLs – Per-collection or per-entry.
- 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

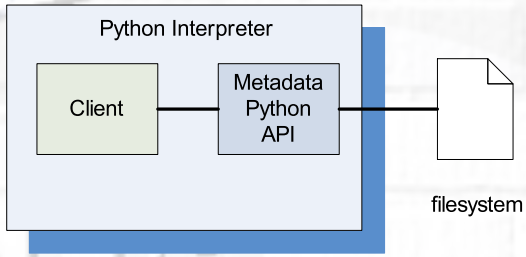
AMGA Implementation

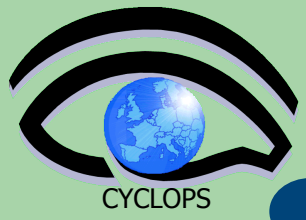
- C++ multiprocess server
  - **Backends**
    - Oracle, MySQL 4/5, PostgreSQL, SQLite
  - **Front Ends**
    - **TCP text streaming**
      - High performance
      - Client API for C++, Java, Python, Perl, PHP
    - **SOAP (web services)**
      - Interoperability
      - Scalability



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

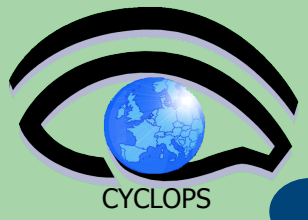
- Standalone Python Library implementation
  - Data stored on file system





AMGA Datatypes

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



AMGA Datatypes

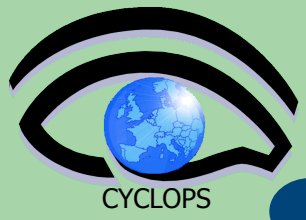
	PostgreSQL	MySQL	Oracle	SQLite	Python
int	integer	int	number(38)	int	int
float	double precision	double precision	float	float	float
varchar(n)	character varying(n)	character varying(n)	varchar2(n)	varchar(n)	string
timestamp	timestamp w/o TZ	datetime	timestamp(6)	unsupported	time (unsupp.)
text	text	text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

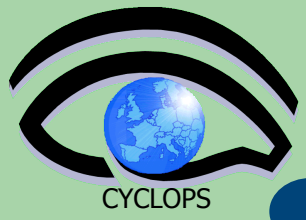
AMGA Datatypes

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



Accessing AMGA from UI/WNs



Accessing AMGA from UI/WNs

- 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 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

Accessing AMGA from UI/WNs

- 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 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 Web Interface

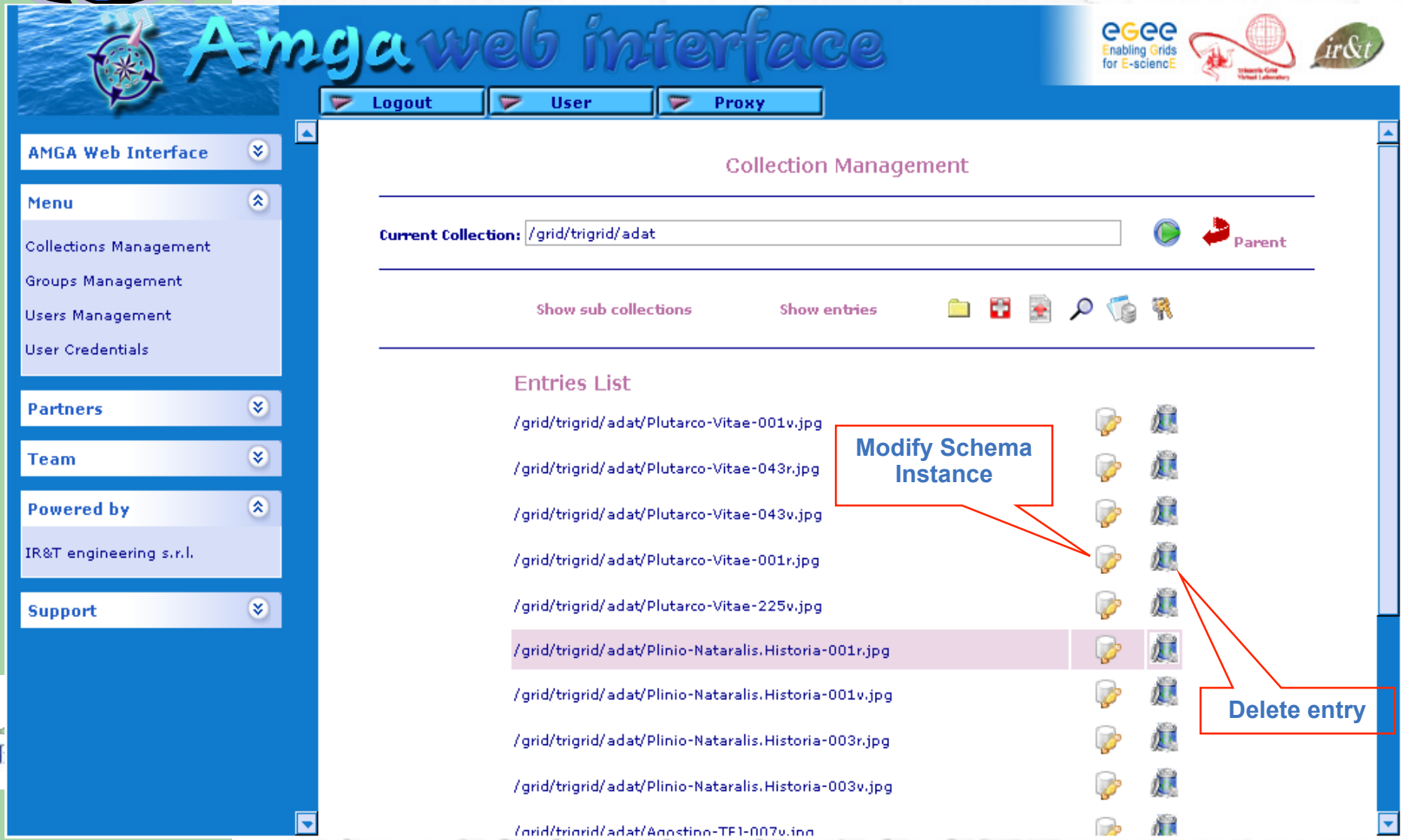
Provide your user ID to log on AMGA Server

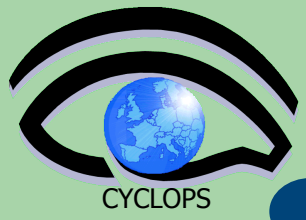
User ID: SAMMY Enter

MESSAGE: Proxy file has not been found, upload it.

op\PROXY\x509up_u502_gilda Sfoglia... Upload

Collection Management

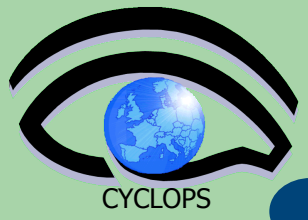




Existing DB access with AMGA

Existing DB access with AMGA

- Since AMGA 1.2.10, a new import feature allow to access existing DB table

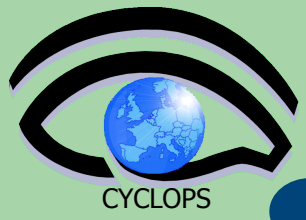


Existing DB access with AMGA

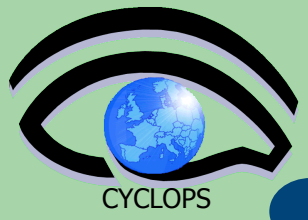
- Since AMGA 1.2.10, a new import feature allow to access existing DB table
- Once imported into AMGA the tables from or more DBs you want to access through AMGA, you can exploit many of the features brought to you by AMGA for your existing tables

Existing DB access with AMGA

- Since AMGA 1.2.10, a new import feature allow to access existing DB table
- Once imported into AMGA the tables from 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

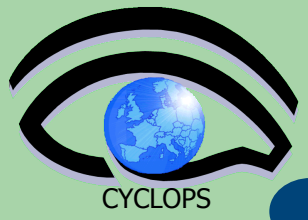


Early adopters of AMGA



Early adopters of AMGA

- LHCb-bookkeeping
  - Migrated bookkeeping metadata to ARDA prototype
    - 20M entries, 15 GB
    - Large amount of static metadata
  - Feedback valuable in improving interface and fixing bugs
  - AMGA showing good scalability

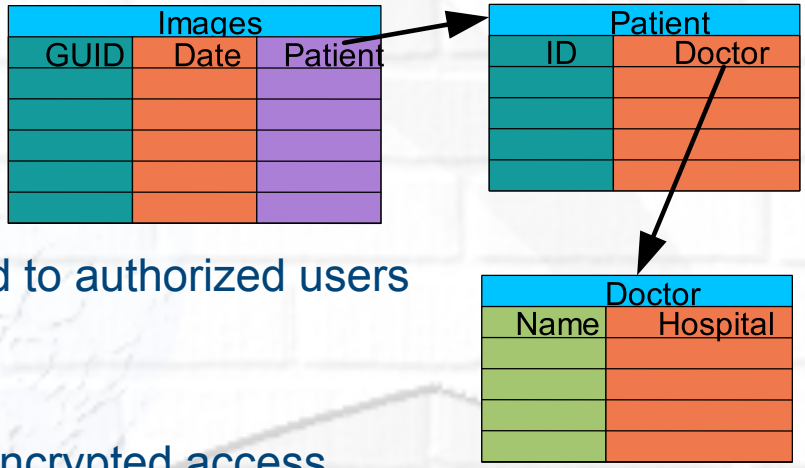


Early adopters of AMGA

- LHCb-bookkeeping
  - Migrated bookkeeping metadata to ARDA prototype
    - 20M entries, 15 GB
    - Large amount of static metadata
  - Feedback valuable in improving interface and fixing bugs
  - AMGA showing good scalability
- Ganga
  - Job management system
    - Developed jointly by Atlas and LHCb
  - Uses AMGA for storing information about job status
    - Small amount of highly dynamic metadata

Biomed - MDM

- 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

CYCLOPS

gMOD: grid Movie On Demand

Information Society and Media

gMOD: grid Movie On Demand

- gMOD provides a Video-On-Demand service

gMOD: grid Movie On Demand

- 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

gMOD: grid Movie On Demand

- 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

gMOD: grid Movie On Demand

- 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.

gMOD under the hood

gMOD under the hood

- Built on top of gLite services:

gMOD under the hood

- Built on top of gLite services:
- Storage Elements, sited in different place, physically contain the movie files

gMOD under the hood

- 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

gMOD under the hood

- 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

gMOD under the hood

- 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

gMOD under the hood

- 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 interactions

CYCLOPS

VOMS

get Role

User

Metadata Catalogue

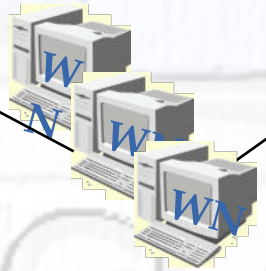
AMGA

Storage Elements

GENIUS Portal

LFC File Catalogue

Workload Management System



CE

Information Society and Media

gMOD screenshot

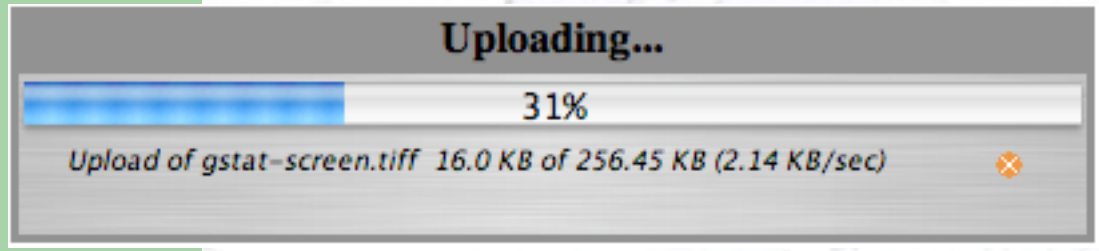
gMOD is accesible through the Genius Portal (https://glite-demo.ct.infn.it)

What is gLibrary

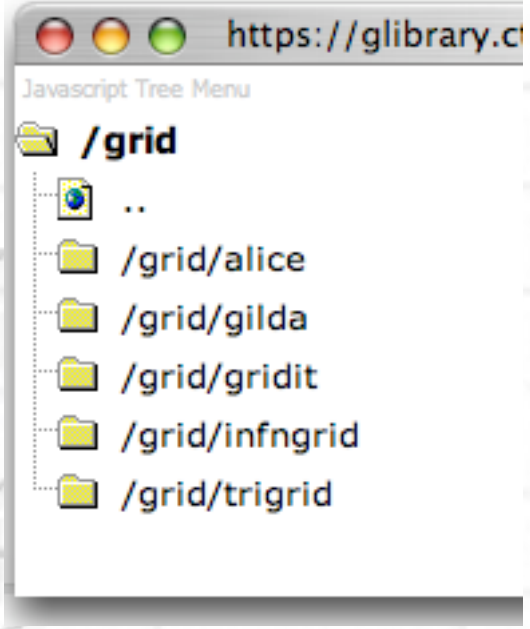
- gLibrary challenge is to offer a multiplatform, flexible, secure and intuitive system to handle digital assets on a Grid Infrastructure.
- By Digital Asset, we mean any kind of content and/or media represented as a computer file. Examples:
  - Images
  - Videos
  - Presentations
  - Office documents
  - E-mails, web pages
  - Newsletters, brochures, bulletins, sheets, templates
  - Receipts, e-books
  - ... (only the imagination can make a limit)
- It allows to store, organize, search and retrieve those assets on a Grid environment.

Store assets on the Grid

- User's local assets are uploaded to one or more (as replicas) Storage Systems the user is authorized on
  - Uploads are managed through Java Applets: a direct GSIFTP copy is done from the local file to the chosen Storage Element

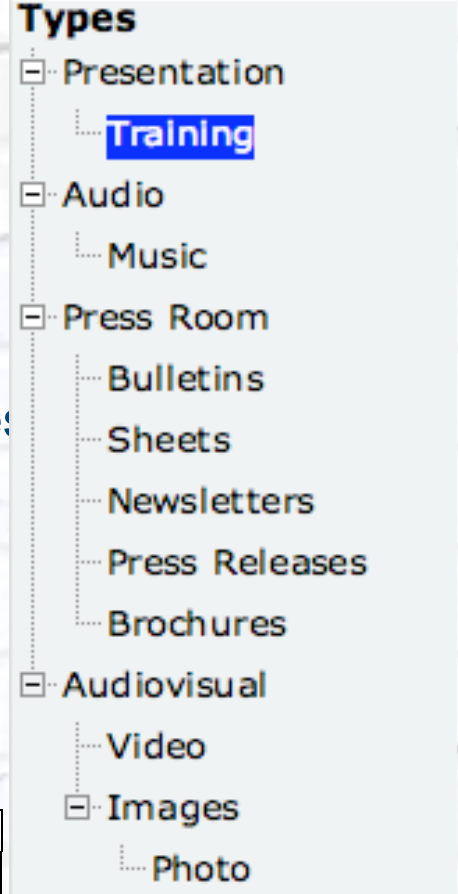


- **File already on the Grid can be managed by gLibrary too**
  - a File Catalogue browser is integrated to select existing grid files.

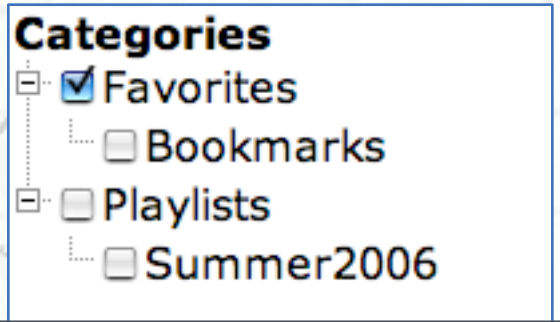


Organize assets

- All entries are organized according to their *type*:
  - a list of specific attributes to describe each kind of assets to be managed by the system;
  - hierarchical (child type shares parent's attributes)
  - defined by the gLibrary administrators
  - queried by users



- **Assets can be organized also by category:**
  - Group together related assets of different types;
  - Useful also to define subset of assets belonging to the same type
  - Multiple category assignment per asset

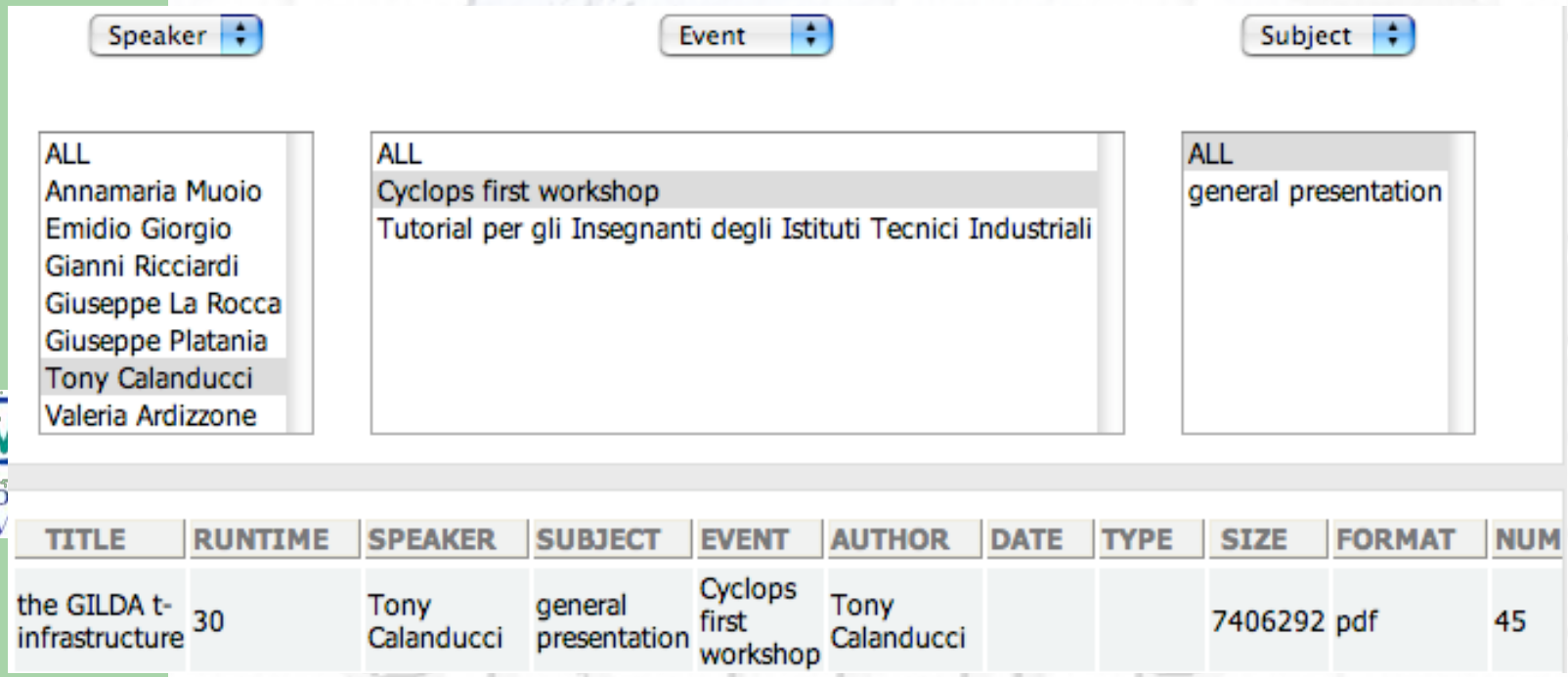


EXAMPLE OF TYPES AND ATTRIBUTES' LIST

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*

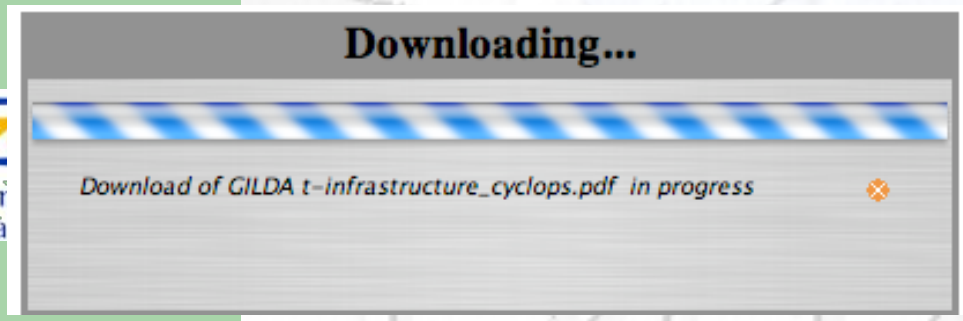
Search assets

- Assets are browsed selecting a type (or category) and selecting one or more filters:
  - type attributes 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"* browser)
- Classic search available too



Retrieve assets from the Grid

- User is presented with a list of asset replicas
- Download from the chosen storage element is matter of a mouse click
- Transfer handled over GridFTP with a Java Applet

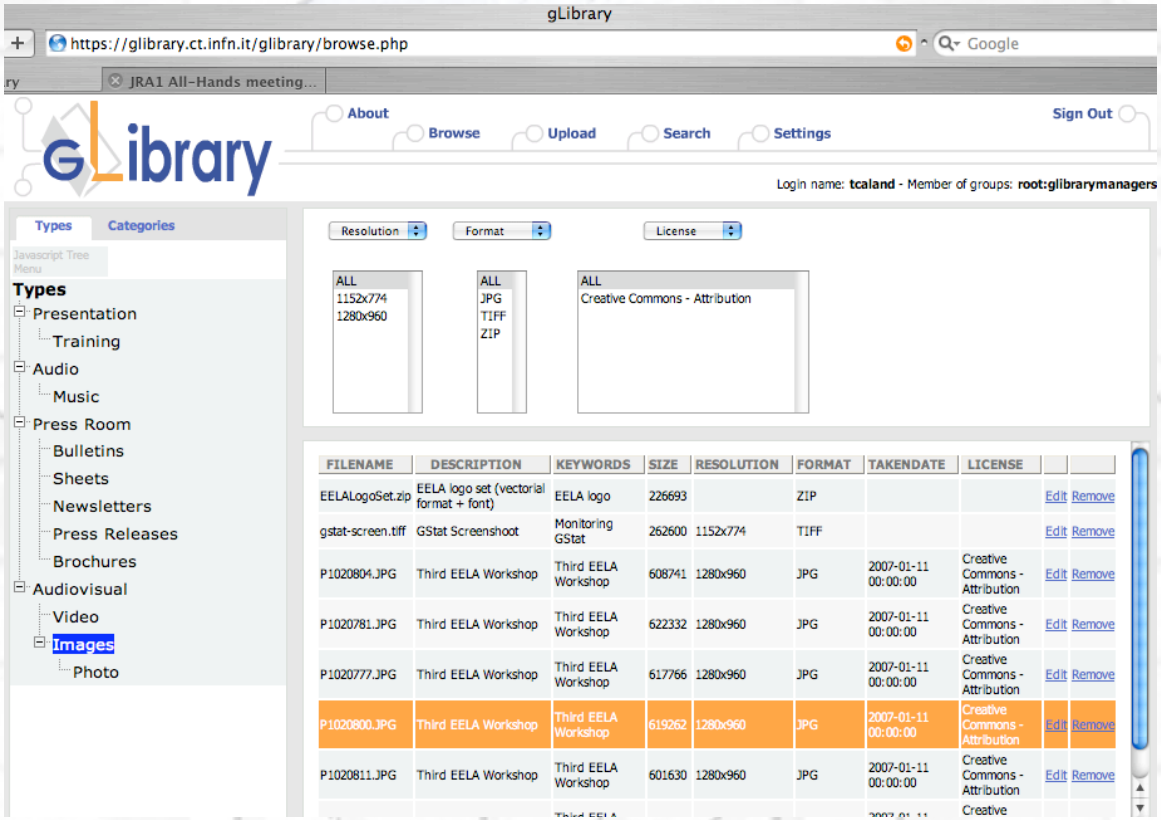


FileName	GILDA t-infrastructure_cyclops.pdf
TypeID	Training
CategoryIDs	
SubmissionDate	2007-05-03 11:46:00
Description	GILDA t-infrastructure
Keywords	t-infrastructure
LastModificationDate	2007-04-10 11:14:02
Size	7406292
Encrypted	
Format	pdf
NumOfPages	45
Title	the GILDA t-infrastructure
Runtime	30
Speaker	Tony Calanducci
Author	Tony Calanducci
Subject	general presentation
Event	Cyclops first workshop
Date	
Type	
Thumb	
OWNER	tcaland
FILE	107
PERMISSIONS	rwx
GROUP_RIGHTS	r-x

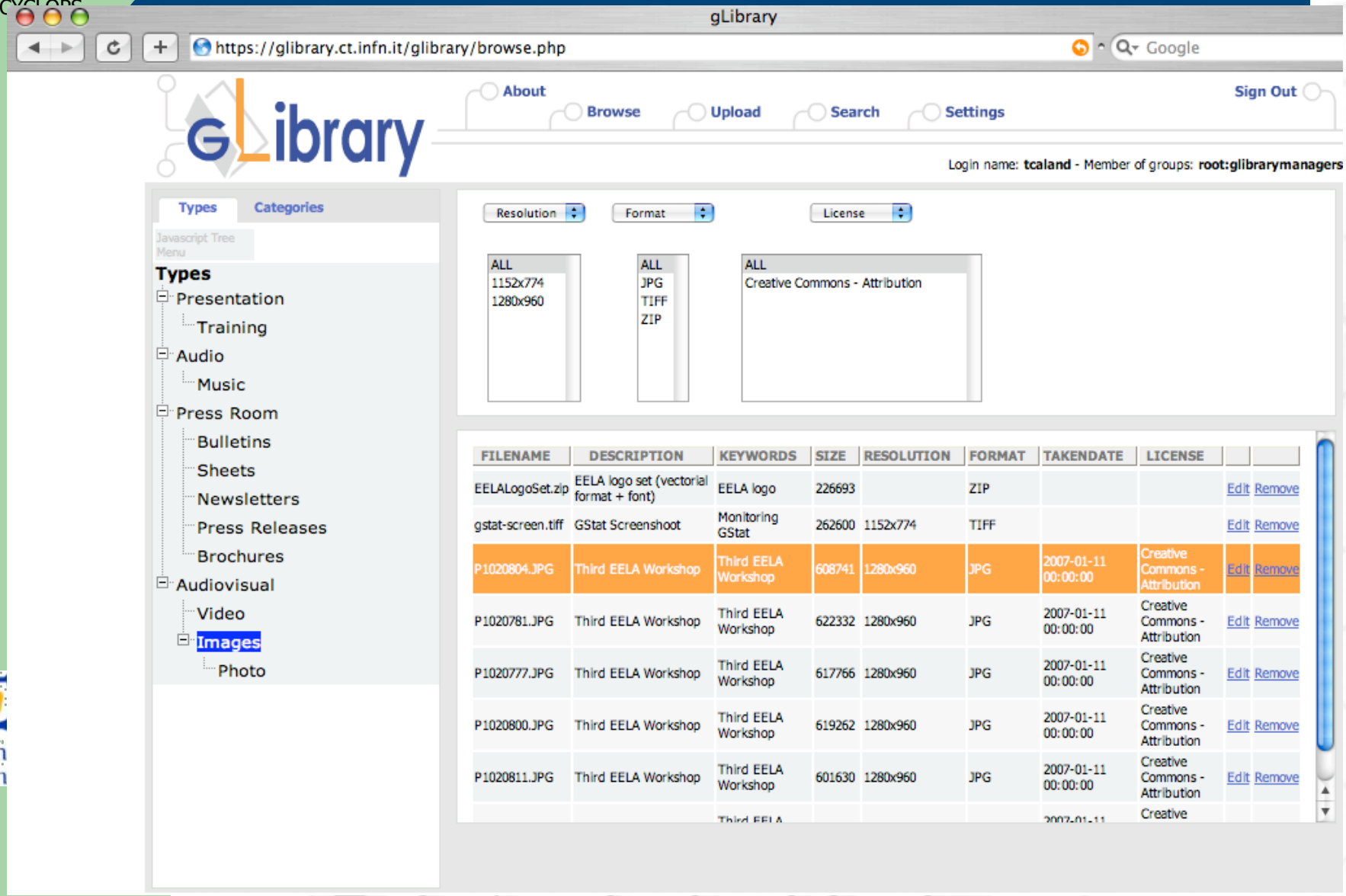
List of replicas:
srm://aliserv6.ct.infn.it/dpm/ct.infn.it/home/gilda/generated/2007-05-03/file2d8cc915-49e9-4040-ab1c-4a1329a2a8d6

Features

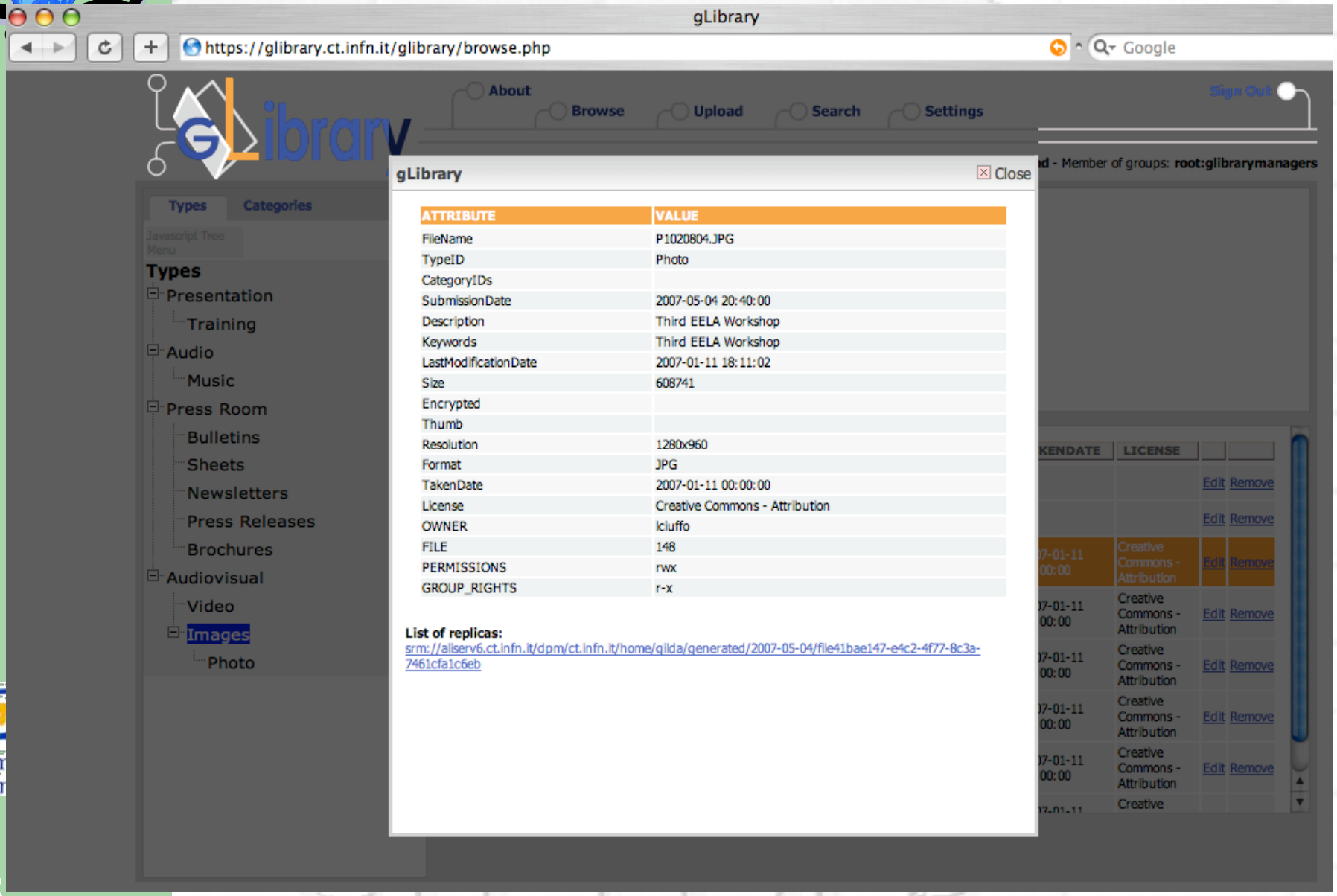
- Implemented as Web 2.0 application
  - AJAX and Javascript are strongly used to offer a desktop like user experience
  - Business logic implemented using PHP 5 OOP support



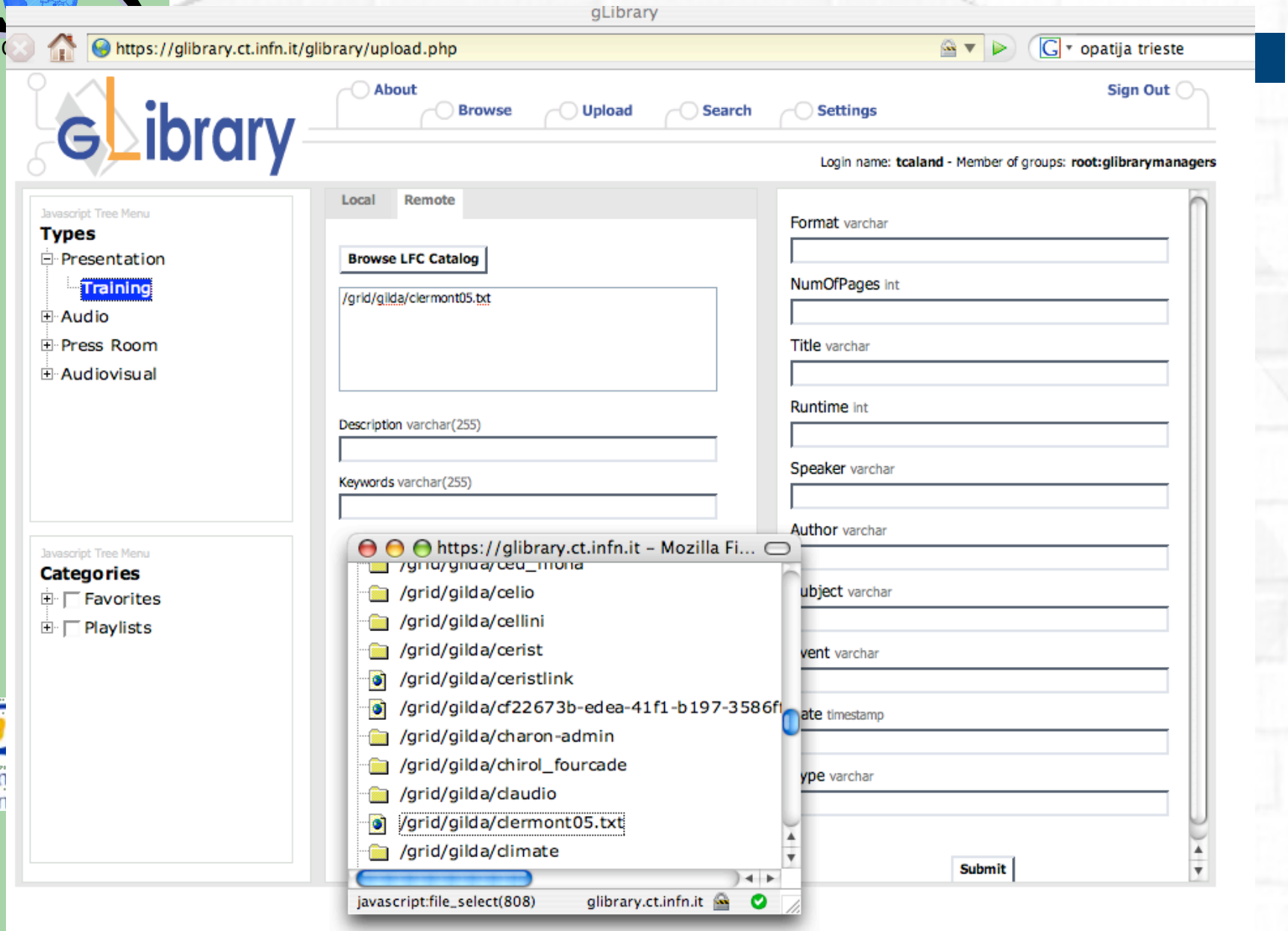
Browsing screenshot



Entry detail screenshot



Upload screenshot



Architecture overview

CYCLOPS

VOMS Server

LFC File Catalogue

AMGA Metadata Catalogue

SE

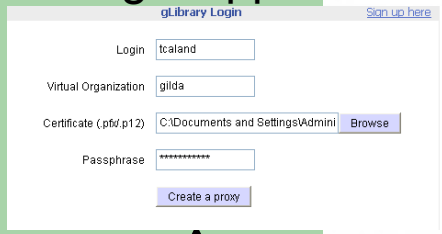
SE

SE

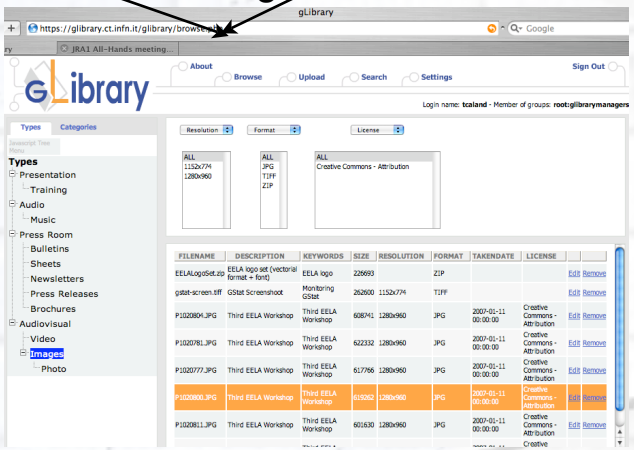
3. get role

4. find the right asset

Login applet

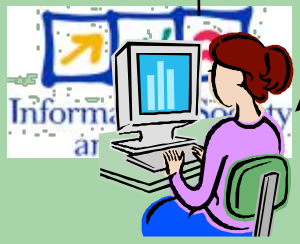


2. proxy transfer over HTTPS

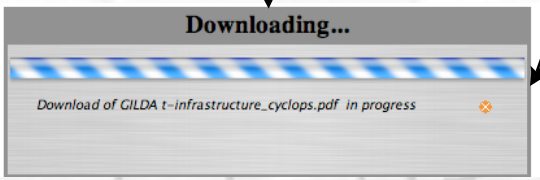


1. local proxy creation

5. proxy retrieved over HTTPS

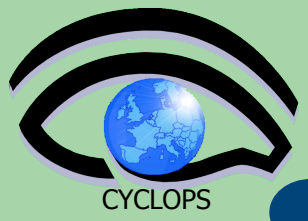


6. direct transfer from SE



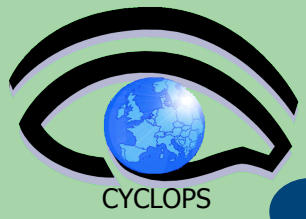
User

Upload/Download applet



Conclusion

- AMGA – Metadata Service of gLite
  - Part of gLite 3.1
  - Useful to realize simple Relational Schemas
  - Integrated on the Grid Environment (Security)
- Replication/Federation features
- Importing existing databases
- Tests show good performance/scalability
- Already deployed by several Grid Applications
  - LHCb, ATLAS, Biomed, gMOD, gLibrary, ADAT



References

- AMGA Web Site
  http://cern.ch/amga
- AMGA Manual
  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