

The AMGA Metadata Service



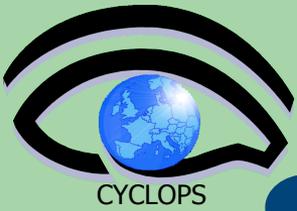
Antonio Calanducci
National Institute of Nuclear Physics
INFN Catania
EGEE NA3 Training & Dissemination
CYCLOPS Second Training Workshop
Chania (Crete), 05th-07th May 2008



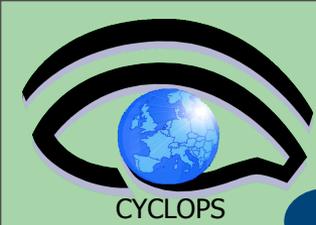
Contents

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

Why Grid needs Metadata?

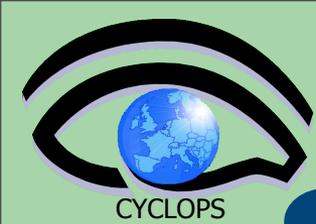


CYCLOPS



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



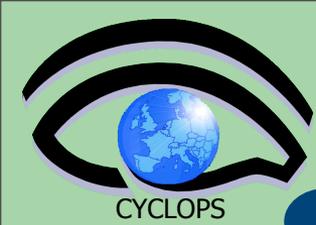
Why Grid needs Metadata?

- 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



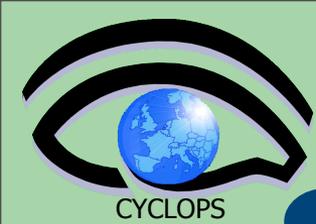
Basic Metadata Concept





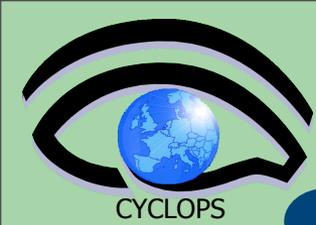
Basic Metadata Concept

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



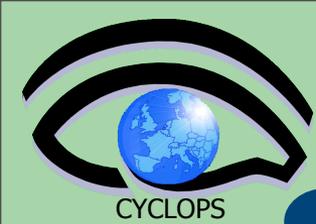
Basic Metadata Concept

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



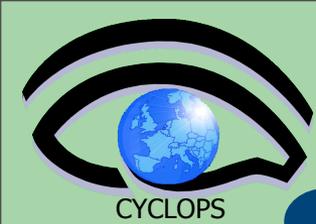
Basic Metadata Concept

- **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,...)



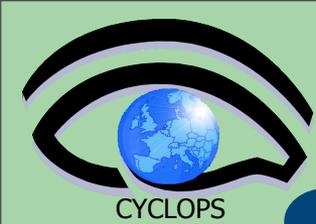
Basic Metadata Concept

- **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





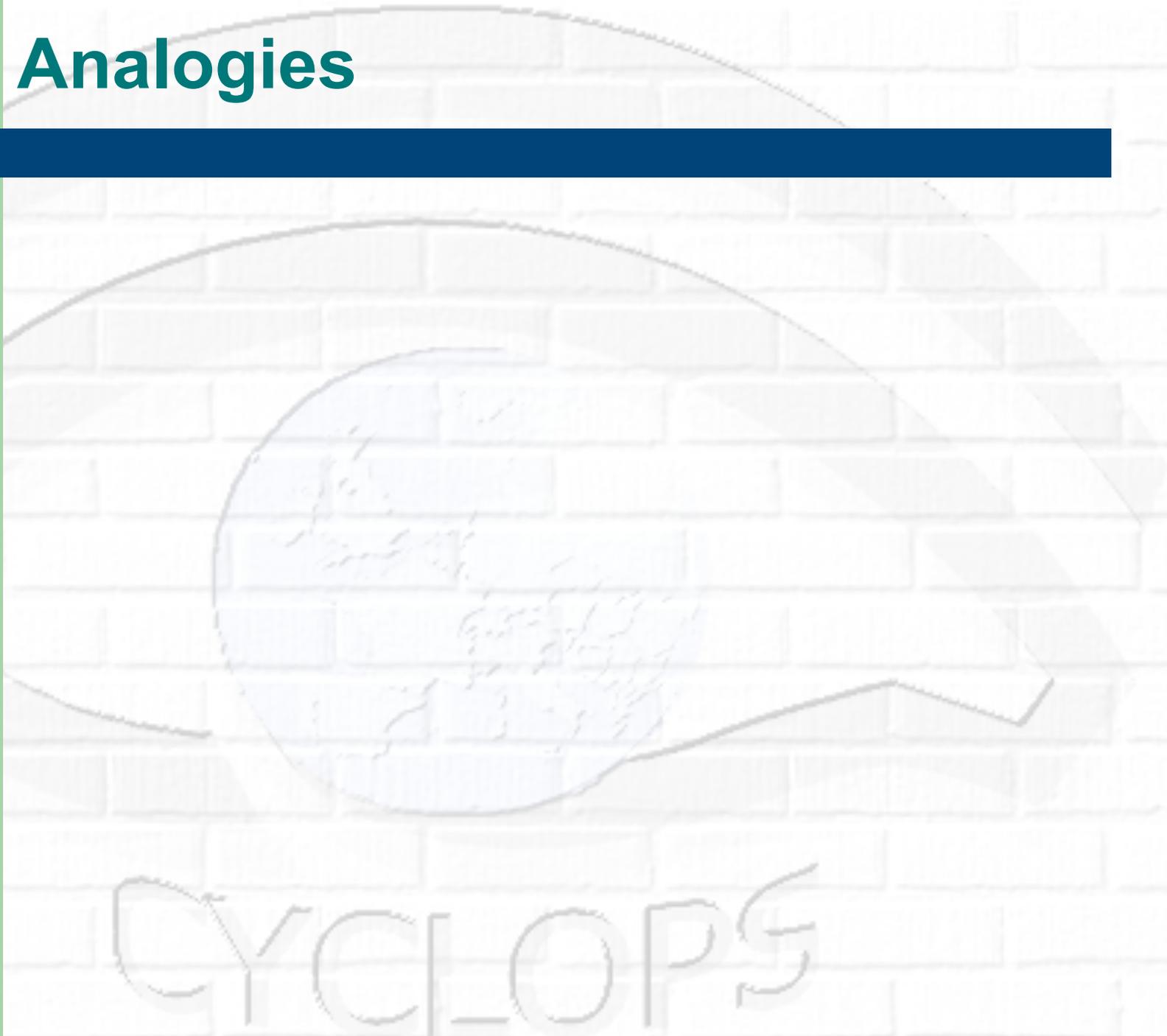
Basic Metadata Concept

- **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 \leftrightarrow table schema
 - collection \leftrightarrow db table
 - attribute \leftrightarrow schema column
 - entry \leftrightarrow table row/record





Analogies

- Analogy to the RDBMS world:
 - schema \leftrightarrow table schema
 - collection \leftrightarrow db table
 - attribute \leftrightarrow schema column
 - entry \leftrightarrow table row/record





Analogies

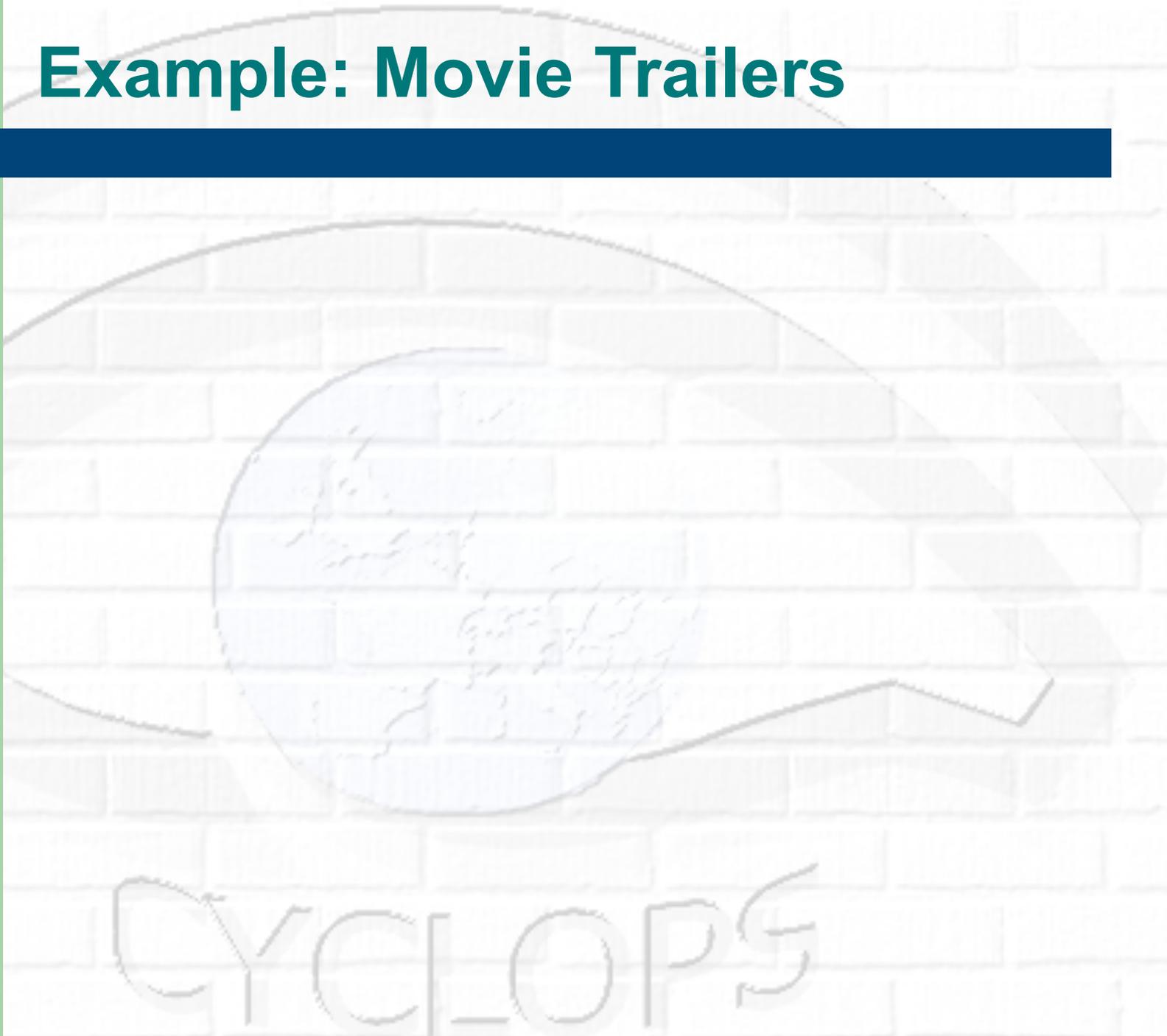
- Analogy to the RDBMS world:
 - schema \leftrightarrow table schema
 - collection \leftrightarrow db table
 - attribute \leftrightarrow schema column
 - entry \leftrightarrow table row/record

- Analogy to a file system:
 - **Collection** \leftrightarrow **Directory**
 - **Entry** \leftrightarrow **File**





Example: Movie Trailers





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

<i>Entry names</i>	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



CYCLOPS



Trailer's example

Attribute

<i>Entry names</i>	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



CYCLOPS



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



CYCLOPS



Trailer's example

Schema

Attribute

Entry names

Title

Ru

Cast

LFN

	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

Entries



CYCLOPS



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



Collection

Entries





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> 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
```

```
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
```



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





CYCLOPS

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 Society
and Media



CYCLOPS

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”





CYCLOPS

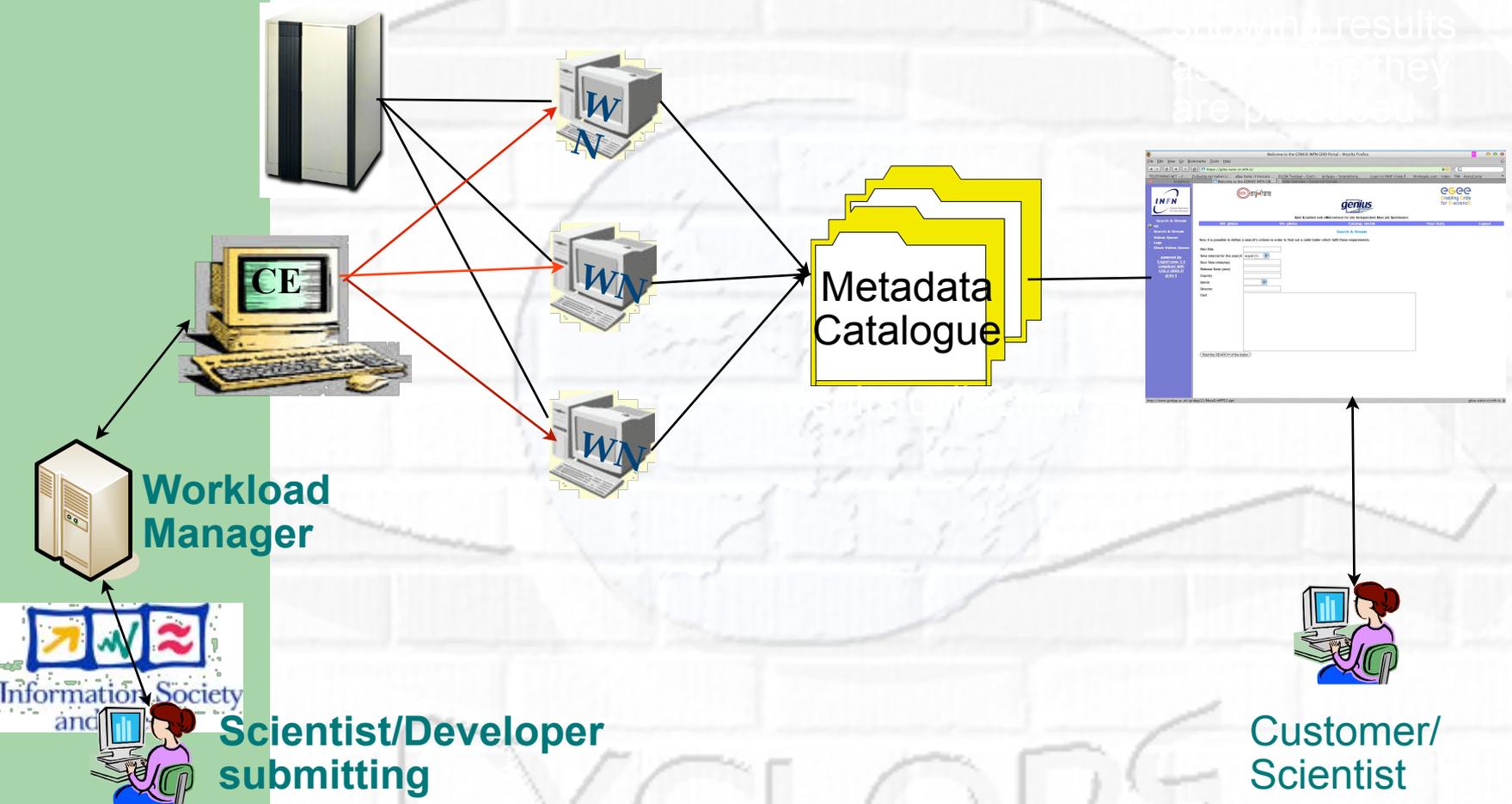
Metadata on the Grid

- 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



Monitoring results
as they are produced



**Scientist/Developer
submitting**

**Customer/
Scientist**



Use a Metadata services to exchange data among running jobs

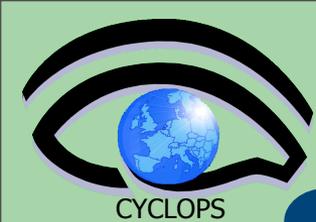


CYCLOPS



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

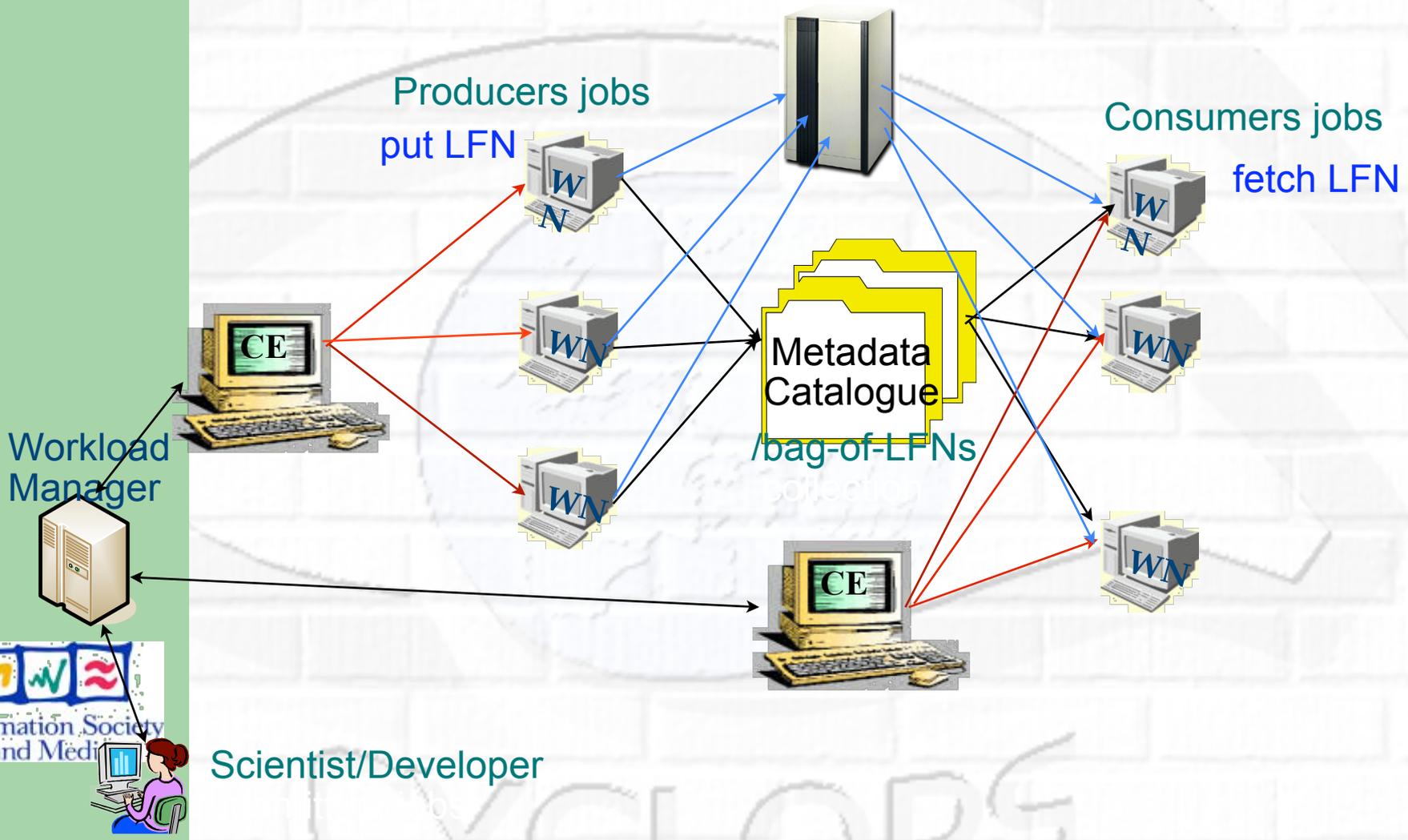


Use a Metadata services to exchange data among running jobs

- 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



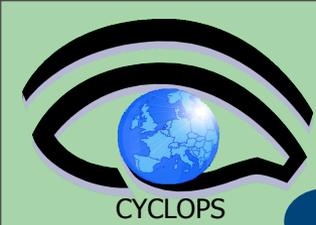
Information exchanging among grid peers





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**

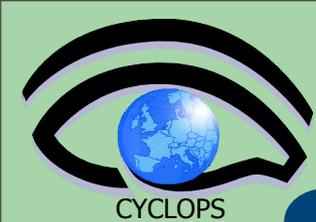




The AMGA Metadata Catalogue

- 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.

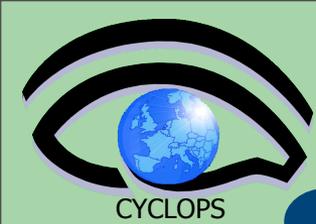




The AMGA Metadata Catalogue

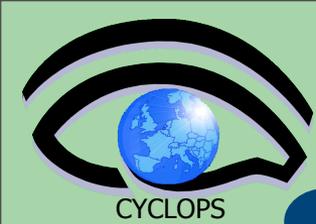
- Official metadata service for the gLite middleware
- AMGA: **Arda Metadata Grid Application**
- Provide a **complete** but **simple interface**, in order to make all users able to use it easily.
- Designed with **scalability** in mind in order to deal with **large number of entries**
- **Grid security** 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



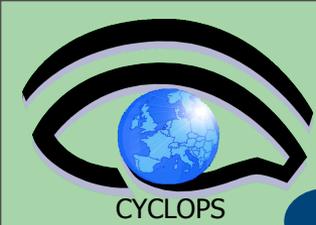
The AMGA Metadata Catalogue

- 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's schema



CYCLOPS

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's 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") '
```



Information Society
and Media



CYCLOPS

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's schema
- **Flexible Queries**
 - SQL-like query language
 - Different join type (inner, outer, left, right) between schemas are provided



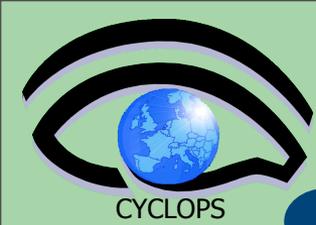
```
selectattr /gLibrary:FileName /gLAudio:Author /gLAudio:Album  
'/gLibrary:FILE=/gLAudio:FILE and like(/gLibrary:FileName, "%.mp3")`
```

▶ 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



CYCLOPS

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)



Information Society
and Media



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

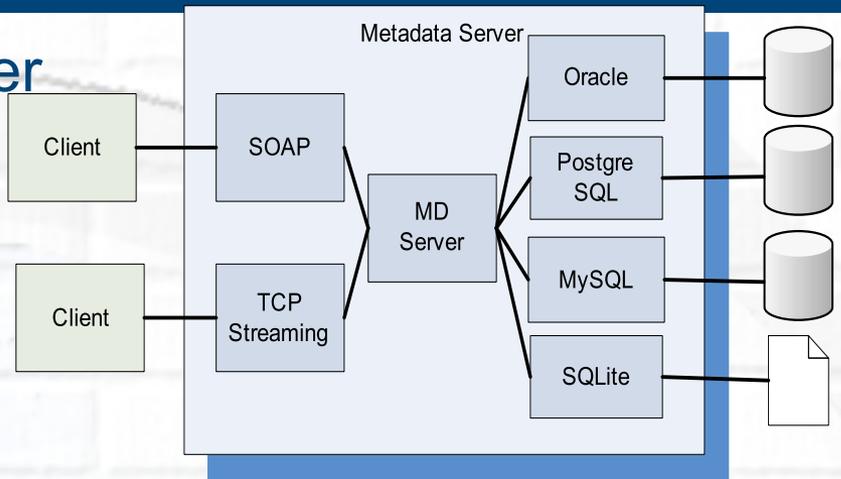
* C++ multiprocess server

- Backends

- Oracle, MySQL 4/5, PostgreSQL, SQLite

- Front Ends

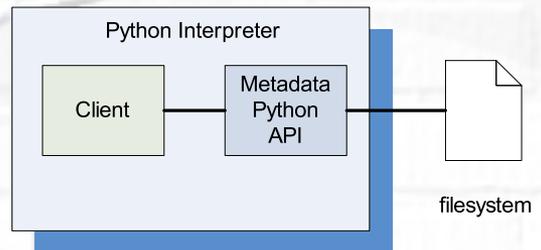
- **TCP text streaming**
 - High performance
 - Client API for C++, Java, Python, Perl, PHP
- **SOAP (web services)**
 - Interoperability
 - Scalability



- 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



CYCLOPS



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 (un supp.)
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



CYCLOPS



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



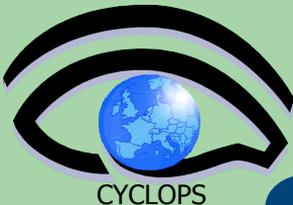
CYCLOPS



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)





CYCLOPS

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

Sfoggia...

Upload





Collection Management



Amga web interface

[Logout](#) [User](#) [Proxy](#)

AMGA Web Interface

Menu

- Collections Management
- Groups Management
- Users Management
- User Credentials

Partners

Team

Powered by

IR&T engineering s.r.l.

Support

Collection Management

Current Collection: Parent

[Show sub collections](#) [Show entries](#)

Entries List

/grid/trigrig/adat/Plutarco-Vitae-001v.jpg		
/grid/trigrig/adat/Plutarco-Vitae-043r.jpg		
/grid/trigrig/adat/Plutarco-Vitae-043v.jpg		
/grid/trigrig/adat/Plutarco-Vitae-001r.jpg		
/grid/trigrig/adat/Plutarco-Vitae-225v.jpg		
/grid/trigrig/adat/Plinio-Nataralis.Historia-001r.jpg		
/grid/trigrig/adat/Plinio-Nataralis.Historia-001v.jpg		
/grid/trigrig/adat/Plinio-Nataralis.Historia-003r.jpg		
/grid/trigrig/adat/Plinio-Nataralis.Historia-003v.jpg		
/grid/trigrig/adat/Agostino-TF1-007v.inn		

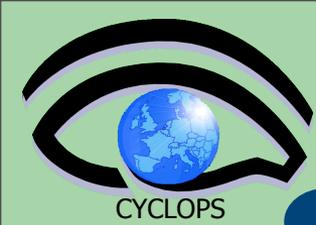
Modify Schema Instance

Delete entry



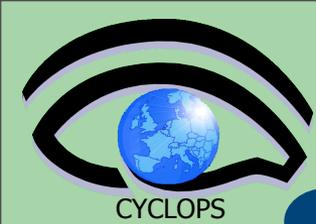
Existing DB access with AMGA





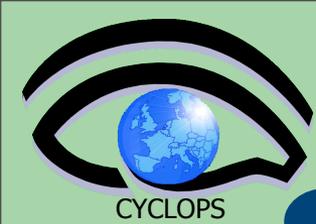
Existing DB access with AMGA

- 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



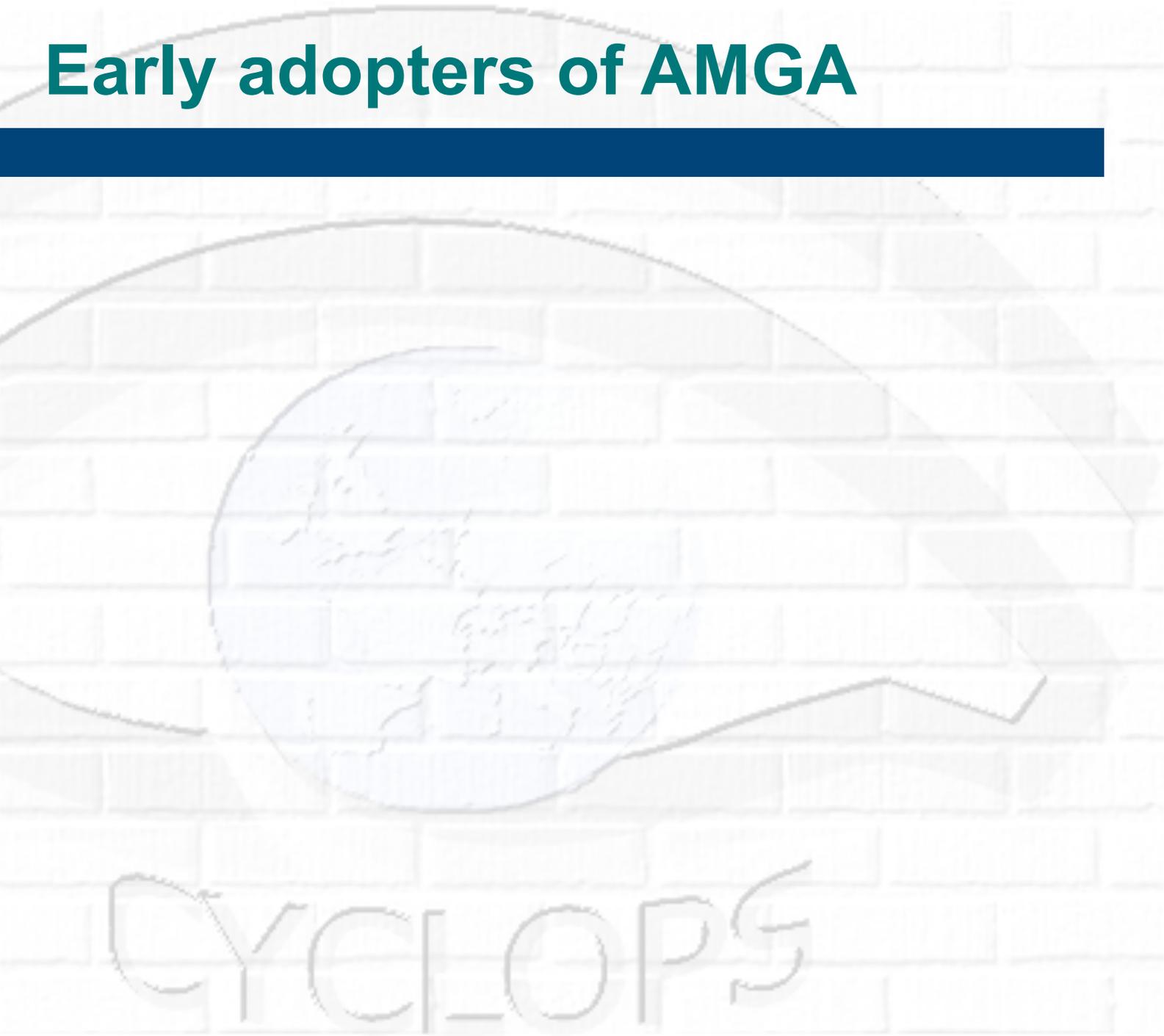
Existing DB access with AMGA

- 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



CYCLOPS

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

Images		
GUID	Date	Patient

Patient	
ID	Doctor

Doctor	
Name	Hospital

- **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: grid Movie On Demand



CYCLOPS



gMOD: grid Movie On Demand

- gMOD provides a **Video-On-Demand** service



CYCLOPS



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



CYCLOPS



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





CYCLOPS

gMOD interactions

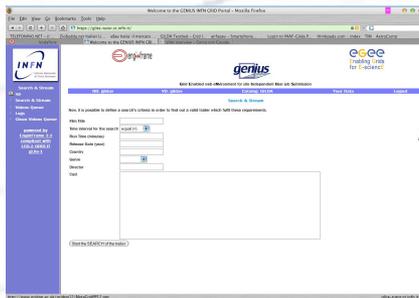
VOMS

get Role



User

GENIUS Portal



Metadata Catalogue

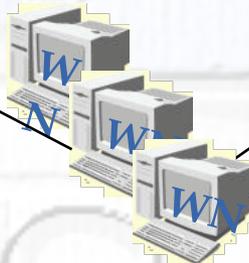
AMGA

LFC File Catalogue

Storage Elements



Workload Management System



Information Society and Media



gMOD screenshot

gMOD is accesible 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... eBay 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...

INFN
Istituto Nazionale di Fisica Nucleare

enginframe

genius

egee
Enabling Grids for E-scienceE

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 choosen.

Title	Shrek 2				
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) Jeffrey Tambor Puss in Boots (voice) Vladek Plazek 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-II
glite-1

Information Science and Media

VLC media player

File Visualizza Impostazioni Audio Video Navigazione Aiuto

0:01:21 / 0:02:30 x1.00 C:\Documents and Settings\Tony\I

glite-tutor.ct.infn.it



What is gLibrary

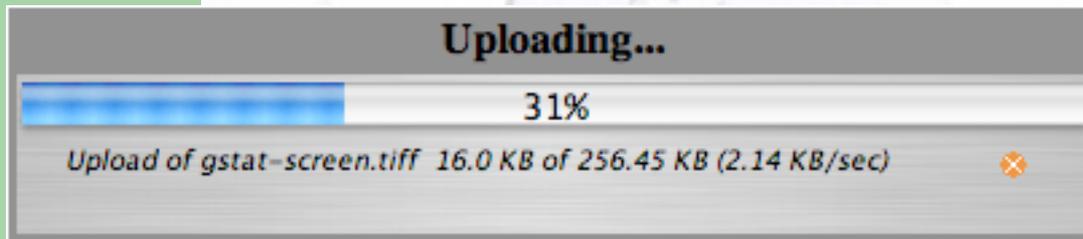
- gLibrary challenge is to offer a **multiplatform**, **flexible**, **secure** and **intuitive** system to handle digital assets on a Grid Infrastructure.
- By Digital Asset, we mean any kind of content and/or media represented as a computer file. Examples:
 - Images
 - Videos
 - Presentations
 - Office 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.



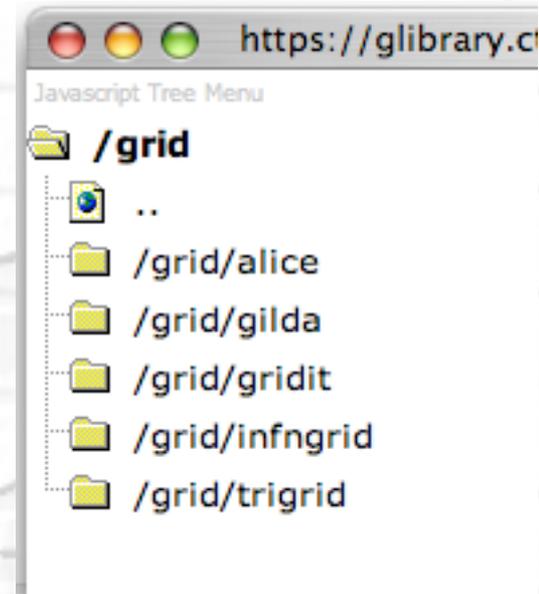


Store assets on the Grid

- 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.

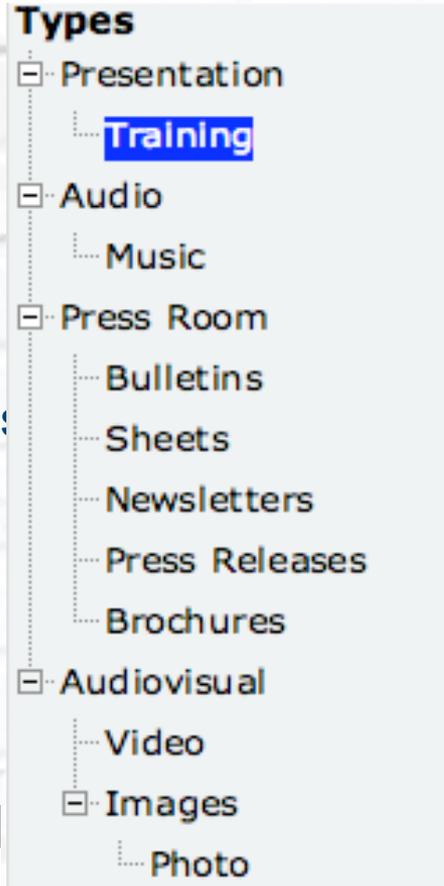




CYCLOPS

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

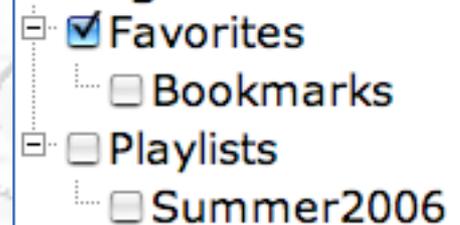


- **Assets can be organized also by category:**
 - Group together related assets of different types;
 - Useful also to define subset of assets belonging to the same type
 - Multiple category assignment per asset

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

Categories





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

Speaker ▾

Event ▾

Subject ▾

- ALL
- Annamaria Muoio
- Emidio Giorgio
- Gianni Ricciardi
- Giuseppe La Rocca
- Giuseppe Platania
- Tony Calanducci
- Valeria Ardizzone

- ALL
- Cyclops first workshop
- Tutorial per gli Insegnanti degli Istituti Tecnici Industriali

- ALL
- general presentation

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





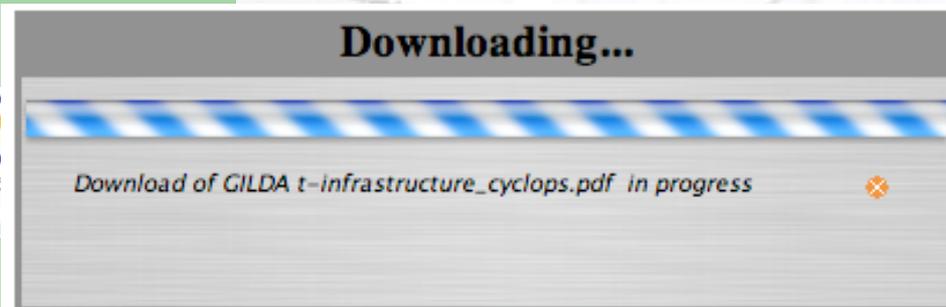
Retrieve assets from the Grid

- 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>





CYCLOPS

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

The screenshot shows the gLibrary web application interface. The browser address bar displays `https://glibrary.ct.infn.it/glibrary/browse.php`. The page header includes the gLibrary logo, navigation links (About, Browse, Upload, Search, Settings), and a Sign Out button. The user is logged in as `tcaland` with group `root:glibrarymanagers`.

The interface features a left sidebar with a 'Types' menu and a 'Categories' section. The 'Types' menu includes: Presentation, Training, Audio, Music, Press Room, Bulletins, Sheets, Newsletters, Press Releases, Brochures, Audiovisual, Video, and Images. The 'Categories' section includes: Photo.

The main content area has three filter boxes: Resolution (set to ALL), Format (set to ALL), and License (set to ALL). Below these filters is a table of files with columns: FILENAME, DESCRIPTION, KEYWORDS, SIZE, RESOLUTION, FORMAT, TAKENDATE, LICENSE, and actions (Edit, Remove).

FILENAME	DESCRIPTION	KEYWORDS	SIZE	RESOLUTION	FORMAT	TAKENDATE	LICENSE		
EELALogoSet.zip	EELA logo set (vectorial format + font)	EELA logo	22693		ZIP			Edit	Remove
gstat-screen.tiff	GStat Screenshot	Monitoring GStat	262600	1152x774	TIFF			Edit	Remove
P1020804.JPG	Third EELA Workshop	Third EELA Workshop	608741	1280x960	JPG	2007-01-11 00:00:00	Creative Commons - Attribution	Edit	Remove
P1020781.JPG	Third EELA Workshop	Third EELA Workshop	622332	1280x960	JPG	2007-01-11 00:00:00	Creative Commons - Attribution	Edit	Remove
P1020777.JPG	Third EELA Workshop	Third EELA Workshop	617766	1280x960	JPG	2007-01-11 00:00:00	Creative Commons - Attribution	Edit	Remove
P1020800.JPG	Third EELA Workshop	Third EELA Workshop	619262	1280x960	JPG	2007-01-11 00:00:00	Creative Commons - Attribution	Edit	Remove
P1020811.JPG	Third EELA Workshop	Third EELA Workshop	601630	1280x960	JPG	2007-01-11 00:00:00	Creative Commons - Attribution	Edit	Remove



Browsing screenshot



gLibrary

https://glibrary.ct.infn.it/glibrary/browse.php

Google

About Browse Upload Search Settings Sign Out

GLibrary

Login name: **tcaland** - Member of groups: **root:glibrarymanagers**

Types Categories

Javascript Tree Menu

Types

- Presentation
 - Training
- Audio
 - Music
- Press Room
 - Bulletins
 - Sheets
 - Newsletters
 - Press Releases
 - Brochures
- Audiovisual
 - Video
 - Images**
 - Photo

Resolution: ALL 1152x774 1280x960

Format: ALL JPG TIFF ZIP

License: ALL Creative Commons - Attribution

FILENAME	DESCRIPTION	KEYWORDS	SIZE	RESOLUTION	FORMAT	TAKENDATE	LICENSE		
EELALogoSet.zip	EELA logo set (vectorial format + font)	EELA logo	226693		ZIP			Edit	Remove
gstat-screen.tiff	GStat Screenshoot	Monitoring GStat	262600	1152x774	TIFF			Edit	Remove
P1020804.JPG	Third EELA Workshop	Third EELA Workshop	608741	1280x960	JPG	2007-01-11 00:00:00	Creative Commons - Attribution	Edit	Remove
P1020781.JPG	Third EELA Workshop	Third EELA Workshop	622332	1280x960	JPG	2007-01-11 00:00:00	Creative Commons - Attribution	Edit	Remove
P1020777.JPG	Third EELA Workshop	Third EELA Workshop	617766	1280x960	JPG	2007-01-11 00:00:00	Creative Commons - Attribution	Edit	Remove
P1020800.JPG	Third EELA Workshop	Third EELA Workshop	619262	1280x960	JPG	2007-01-11 00:00:00	Creative Commons - Attribution	Edit	Remove
P1020811.JPG	Third EELA Workshop	Third EELA Workshop	601630	1280x960	JPG	2007-01-11 00:00:00	Creative Commons - Attribution	Edit	Remove
		Third EELA				2007-01-11	Creative		



Entry detail screenshot



gLibrary

https://glibrary.ct.infn.it/glibrary/browse.php

About Browse Upload Search Settings Sign Out

gLibrary

Types Categories

Types

- Presentation
 - Training
- Audio
 - Music
- Press Room
 - Bulletins
 - Sheets
 - Newsletters
 - Press Releases
 - Brochures
- Audiovisual
 - Video
 - Images**
 - Photo

gLibrary

ATTRIBUTE	VALUE
FileName	P1020804.JPG
TypeID	Photo
CategoryIDs	
SubmissionDate	2007-05-04 20:40:00
Description	Third EELA Workshop
Keywords	Third EELA Workshop
LastModificationDate	2007-01-11 18:11:02
Size	608741
Encrypted	
Thumb	
Resolution	1280x960
Format	JPG
TakenDate	2007-01-11 00:00:00
License	Creative Commons - Attribution
OWNER	lcluffo
FILE	148
PERMISSIONS	rwX
GROUP_RIGHTS	r-x

List of replicas:
<srm://allserv6.ct.infn.it/dpm/ct.infn.it/home/qilda/generated/2007-05-04/file41bae147-e4c2-4f77-8c3a-7461cfa1c6eb>

Member of groups: root:glibrarymanagers

KENDATE	LICENSE		
		Edit	Remove
		Edit	Remove
07-01-11 00:00	Creative Commons - Attribution	Edit	Remove
07-01-11 00:00	Creative Commons - Attribution	Edit	Remove
07-01-11 00:00	Creative Commons - Attribution	Edit	Remove
07-01-11 00:00	Creative Commons - Attribution	Edit	Remove
07-01-11 00:00	Creative Commons - Attribution	Edit	Remove
07-01-11	Creative		



Upload screenshot

The screenshot displays the gLibrary web application interface. At the top, the browser address bar shows the URL `https://glibrary.ct.infn.it/glibrary/upload.php`. The page header includes navigation links for **About**, **Browse**, **Upload**, **Search**, and **Settings**, along with a **Sign Out** button. The user is logged in as **tcaland**, a member of the **root:glibrarymanagers** group.

The main content area is divided into three sections:

- Types:** A tree menu with expandable categories: **Presentation**, **Training** (highlighted), **Audio**, **Press Room**, and **Audiovisual**.
- Categories:** A tree menu with expandable categories: **Favorites** and **Playlists**.
- Local/Remote:** A file selection interface. The **Remote** tab is active, showing a **Browse LFC Catalog** button and a text input field containing `/grid/gilda/clermont05.txt`. Below this are input fields for **Description** and **Keywords**.

On the right side, there is a form for metadata with the following fields:

- Format** (varchar)
- NumOfPages** (int)
- Title** (varchar)
- Runtime** (int)
- Speaker** (varchar)
- Author** (varchar)
- Subject** (varchar)
- Event** (varchar)
- Date** (timestamp)
- Type** (varchar)

A **Submit** button is located at the bottom right of the form.

An overlaid file explorer window shows the directory structure of the LFC Catalog, with `/grid/gilda/clermont05.txt` selected. The directory list includes:

- /grid/gilda/ced_mona
- /grid/gilda/celio
- /grid/gilda/cellini
- /grid/gilda/cerist
- /grid/gilda/ceristlink
- /grid/gilda/cf22673b-ede4-41f1-b197-3586f...
- /grid/gilda/charon-admin
- /grid/gilda/chinol_fourcade
- /grid/gilda/claudio
- /grid/gilda/clermont05.txt**
- /grid/gilda/climate

The browser's status bar at the bottom indicates the page title as `javascript:file_select(808)` and the URL as `glibrary.ct.infn.it`.



CYCLOPS

VOMS Server



Architecture overview



Login applet

gLibrary Login [Sign Up here](#)

Login:

Virtual Organization:

Certificate (.pfx/.p12):

Passphrase:

glibrary

https://glibrary.ct.inf.nu/glibrary/browse

gLibrary

Types Categories

Resolution: ALL | Format: ALL | License: ALL

FILENAME	DESCRIPTION	KEYWORDS	SIZE	RESOLUTION	FORMAT	TAKENDATE	LICENSE
EELAlogoSet.zip	EELA logo set (vector format + font)		22693		ZIP		
glib-screen.tif	Glib Screenshots	Monitoring Glib	26200	1152x774	TIFF		
F102080.JPG	Third EELA Workshop		60874	1280x960	JPG	2007-05-11 00:00:00	Creative Commons - Attribution
F102078.JPG	Third EELA Workshop		62232	1280x960	JPG	2007-05-11 00:00:00	Creative Commons - Attribution
F102077.JPG	Third EELA Workshop		61796	1280x960	JPG	2007-05-11 00:00:00	Creative Commons - Attribution
F102080.JPG	Third EELA Workshop		61202	1280x960	JPG	2007-05-11 00:00:00	Creative Commons - Attribution
F102081.JPG	Third EELA Workshop		60530	1280x960	JPG	2007-05-11 00:00:00	Creative Commons - Attribution

3. get role

4. find the right asset

2. proxy transfer over HTTPS

5. proxy retrieved over HTTPS

6. direct transfer from SE

Downloading...

Download of GILDA t-infrastructure_cyclops.pdf in progress

Upload/Download applet

1. local proxy creation



37 User



CYCLOPS

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



Information Society
and Media



CYCLOPS

References

- AMGA Web Site

<http://cern.ch/amqa>

- AMGA Manual

http://amqa.web.cern.ch/amqa/downloads/amqa-manual_1_3_0.pdf

- AMGA API Javadoc

<http://amqa.web.cern.ch/amqa/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>



Information Society
and Media