



Enabling Grids for E-science

# AMGA - Official Metadata Service for EGEE

*Salvatore Scifo*

*INFN Catania*

*BioinfoGrid*

*Bari, March 8<sup>th</sup> - 10<sup>th</sup> 2006*

[www.eu-egee.org](http://www.eu-egee.org)



- **Background and Motivation for AMGA**
- **Interface, Architecture and Implementation**
- **Metadata Replication with AMGA**
- **Gilda use cases**

- **Grids often contain millions of files spread over several storage sites.**
- **Users and applications need an efficient mechanism**
  - to find the files of interest
  - to discover and query information about their contents
- **This is provided**
  - by associating descriptive attributes (metadata) to files
  - by exposing this information in catalogues, accessible and searchable by user and client application

- Metadata service must expose a **complete** but **simple interface**, in order to make all users able to use it easily.
- It should be **flexible** and support **dynamic schemas** in order to serve many (**all is wished**) application domains.
- The service must also allow **structured** and **hierarchical metadata** in order to implement any **logical collections**.
- **Collection** refers metadata grouped by any logical entity meaning. (for example, a collection can describe all file video in any encoded format).

- It must be designed with **scalability** in mind in order to deal with the **large number of entries** (several millions).
- **security** is required to provide **different access levels** to different users.
- **Quality of service has to ensure**
  - **Hide network latency** – Improved performance for WAN clients
  - **Disconnected computing** – Local replicas for off-line access (laptops)
  - **DB Independent replication** – GRID environment is heterogeneous
  - **Improve reliability and scalability** – No single point of failure

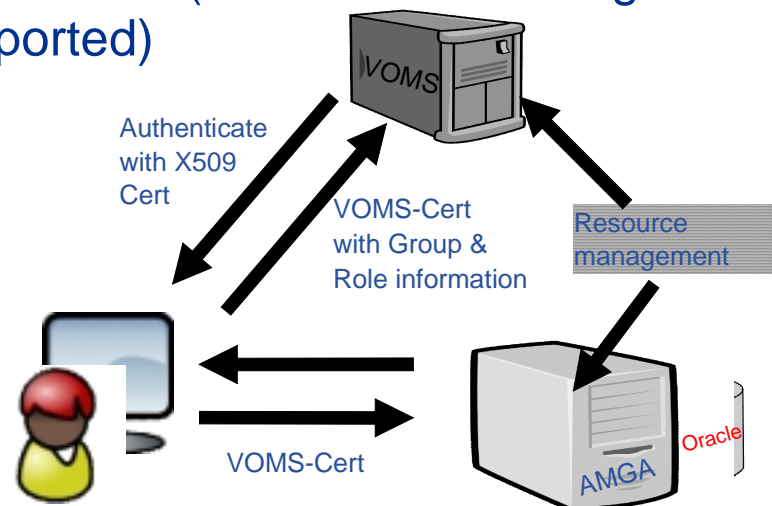
- **AMGA is a metadata service for the Grid**
  - It represents a database access service for Grid applications which allows user, and user jobs to discovery data describing their files in order to access them in the appropriate way.
- **AMGA is a service based on RDBMS.**
  - It allows to define metadata schemas according to users and applications needs
  - It provides a replication layer which makes databases locally available to user jobs and replicate the changes between the different participating databases.
- **AMGA has been designed to provide a best integration with the Grid environment**
  - Metadata Service is a **Grid component**
  - **Grid security** compliant
  - **Hide DB heterogeneity**

- **AMGA comes from the evaluation of existing Metadata Services from HEP experiments**
  - AMI (ATLAS), RefDB (CMS), Alien Metadata Catalogue (ALICE)
  - Similar **goals**, similar **concepts**
  - **Application domain dependent** designed, (not so good to reuse)
  - Several **technical limitations**: large answers, scalability, speed, lack of flexibility
  
- **ARDA proposed an interface for Metadata access on the GRID**
  - **AMGA – ARDA Metadata Grid Application**
    - **Based on requirements of LHC experiments**
    - **General purpose - not bound to a particular application domain**
    - **Designed jointly with the gLite/EGEE team**
    - **Incorporates feedback from GridPP (Grid Particle Physicists)**

- **Began a prototype to evaluate the Metadata Interface**
  - Evaluated by community since the beginning:
    - LHCb and Ganga
- **Adopted as the official EGEE Metadata Interface**
  - First release with upcoming gLite 1.5
  - Also available as standalone component
- **Expanding user community**
  - HEP, Biomed, UNOSAT, ...

- **Dynamic Schemas**
  - Schemas can be modified at runtime by client
    - Create, delete schemas
    - Add, remove attributes
- **Metadata organised as an hierarchy**
  - Schemas can contain sub-schemas
  - Analogy to file system:
    - Schema  $\leftrightarrow$  Directory; Entry  $\leftrightarrow$  File
- **Flexible Queries**
  - SQL-like query language
  - Joins between schemas are provided

- **Access control**
  - All entries in a directory sharing the same **ACL**
  - Groups of users are also supported (**Unix style permissions**)
- **Secure connections – SSL**
  - Provided by web services
- **Client Authentication is based on**
  - Username/password
  - General X509 certificates
  - Grid-proxy certificates (VOMS **Organization Management System** is supported)



- **C++ multiprocess server**

- **Backends**

- Oracle, MySQL, PostgreSQL, SQLite

- **Front Ends**

- **TCP Streaming**

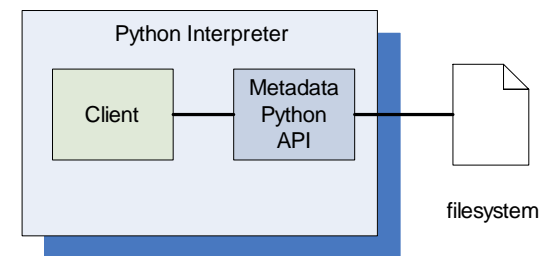
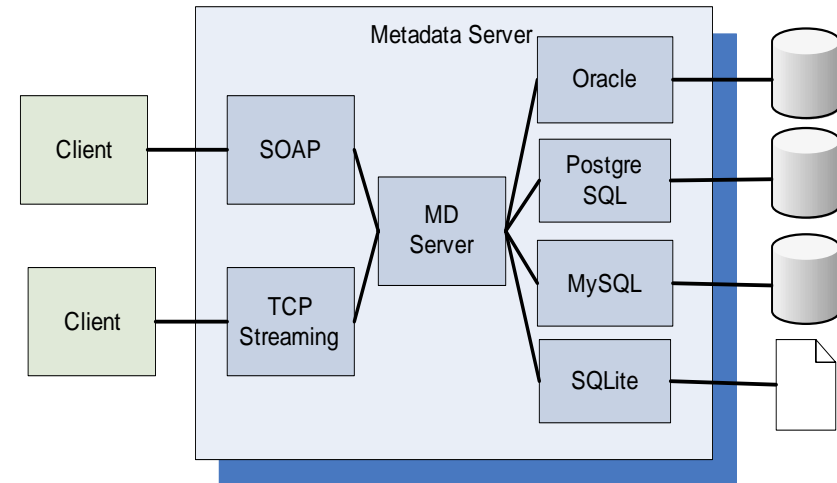
- *High performance*
      - *Client API for C++, Java, Python, Perl, Ruby*

- **SOAP (web services)**

- *Interoperability*
      - *Scalability*

- **Standalone Python Library implementation**

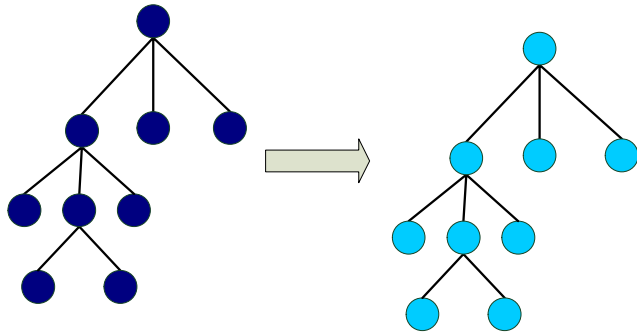
- Data stored on file system



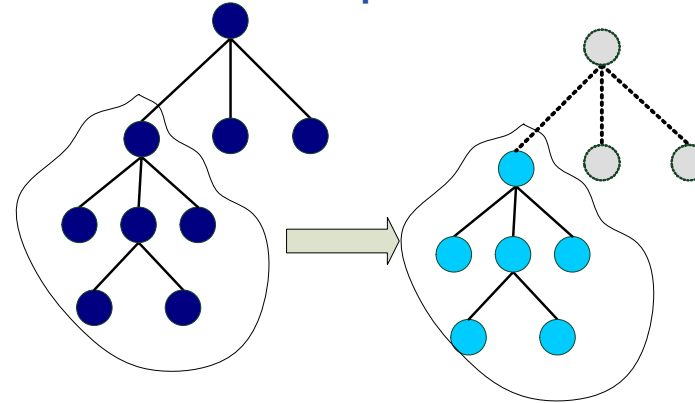
- **AMGA provides an replication/federation mechanisms**
- **Motivation**
  - **Scalability** – Support hundreds/thousands of concurrent users
  - **Geographical distribution** – Hide network latency
  - **Reliability** – No single point of failure
  - **DB Independent replication** – Heterogeneous DB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

## Some use cases

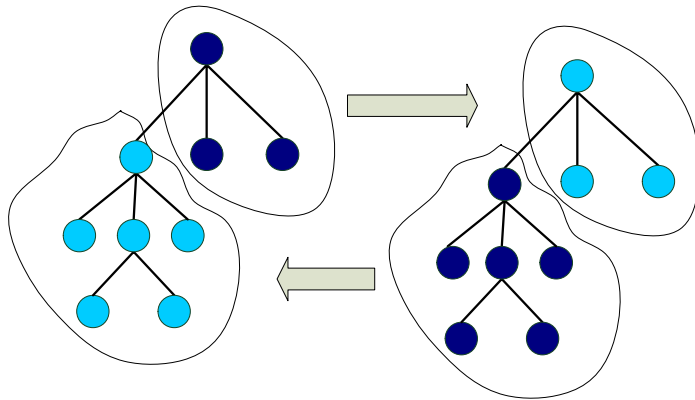
Full replication



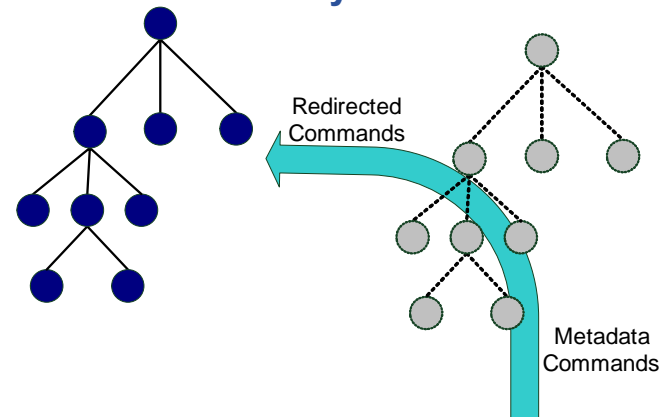
Partial replication



Federation



Proxy



- **AMGA – Metadata Service of gLite**
  - Part of **gLite 1.5**
  - Useful for simplified DB access
  - Integrated on the Grid environment (Security)
- **Replication/Federation under development**
- **Tests show good performance/scalability**
- **Already deployed by several Grid Applications**
  - LHCb, ATLAS, Biomed, ...
- **AMGA Web Site**

<http://project-arda-dev.web.cern.ch/project-arda-dev/metadata/>



Enabling Grids for E-scienceE

# AMGA - Gilda Use Cases

**Salvatore Scifo**

**INFN Catania**

(Slides courteously provided by Tony Calanducci)

*BioinfoGrid*

*Bari, 08-10th March 2006*

[www.eu-egee.org](http://www.eu-egee.org)



University of  
Coimbra



Information Society  
and Media



- To better understand how AMGA works think of
  - schemas  $\leftrightarrow$  tables
  - attributes  $\leftrightarrow$  columns
  - entries  $\leftrightarrow$  rows
- **AMGA Metadata** is list of attributes associated with entries according to a user defined schema.
- **Schema** is a set of attributes
- **Entry** is the abstraction of directory/file mapped by the metadata server
- **Collection** is a set of entries associated with a schema

- **Attribute** – typed key/value pair associated with entries
  - **Type** – The type (int, float, string,...)
  - **Name/Key** – The name of the attribute
  - **Value** - Value of an entry's attribute
- **Analogy Examples**
  - >createdir /jobs (*create table jobs*)
  - >addattr /jobs jobStatus int (*alter table jobs add column jobStatus int*)
  - >addentry /jobs/job1 jobStatus 0 (*insert into jobs (jobstatus) values(1)*)
  - >updateattr /jobs jobStatus 1 jobID>100 (*update jobs set jobStatus=1 where JobID>100*)

- **LHCb-bookkeeping**
  - Migrated bookkeeping metadata to ARDA prototype
    - 20M entries, 15 GB
    - Large amount of static metadata
- **Ganga project**
  - Job management system
    - Developed jointly by Atlas and LHCb
  - Uses AMGA for storing information about job status
    - Small amount of highly dynamic metadata

- **AMGA Datatypes**

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

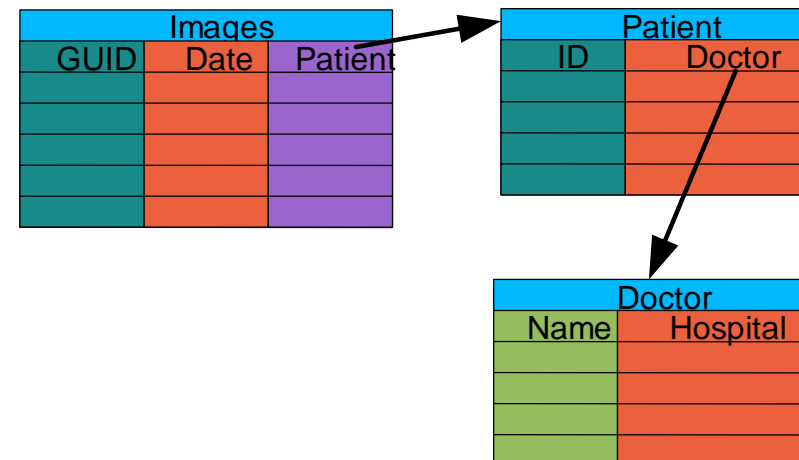
- Using the above datatypes you are sure that your metadata can be easily moved to all supported back-ends
- If you do not care about DB portability, you can use, in principle, as entry attribute type ALL the datatypes supported by the back-end, even the more esoteric ones (PostgreSQL Network Address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

- **Biomed**

- **Medical Data Manager – MDM**
  - Store and access medical images and associated metadata on the Grid
  - Built on top of gLite 1.5 data management system
  - Demonstrated at last EGEE conference (October 05, Pisa)
- **Strong security requirements**
  - Patient data is sensitive
  - Data must be encrypted
  - Metadata access must be restricted to authorized users

- **AMGA used as metadata server**

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



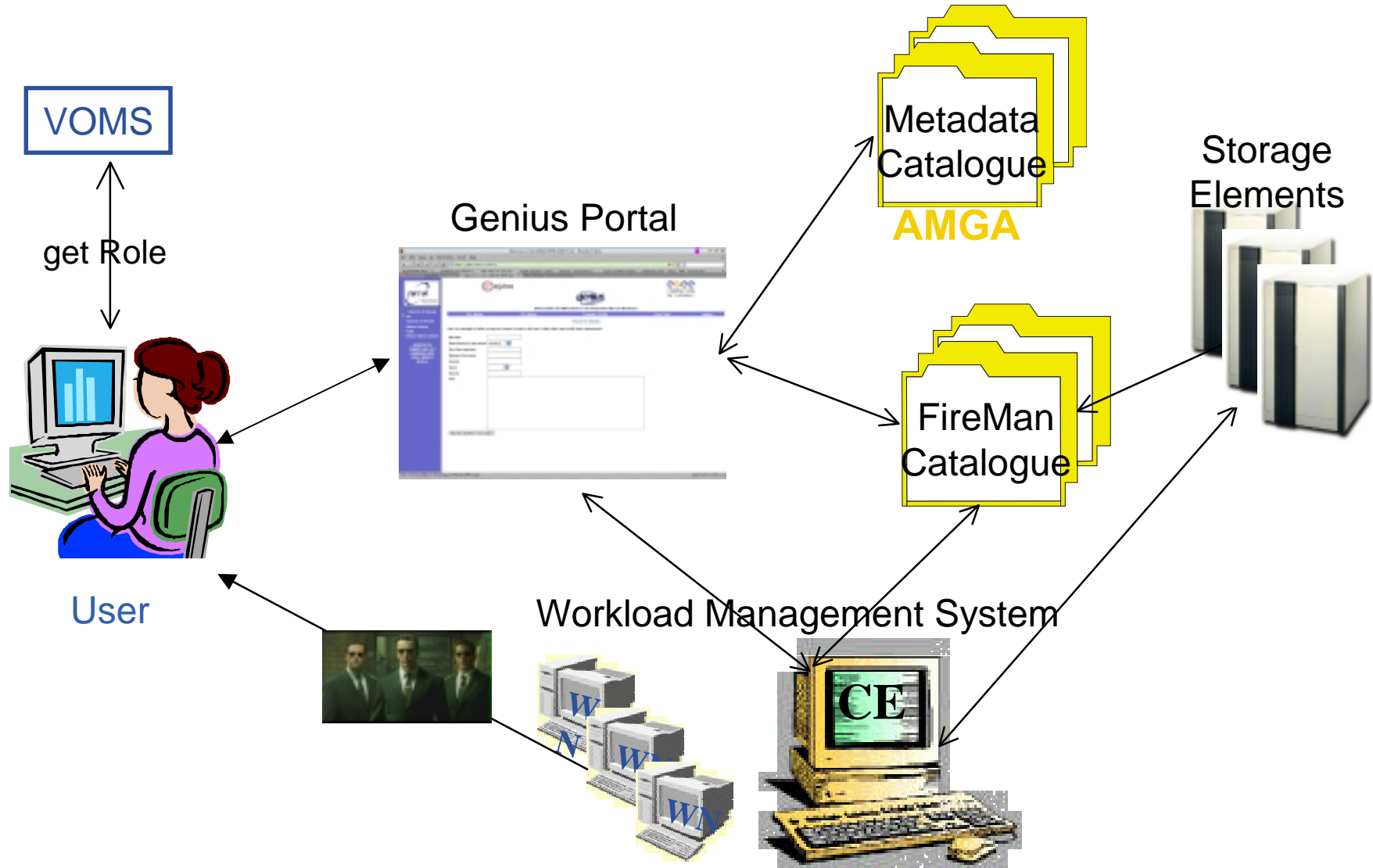
- **More details at**

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

- **gMOD**

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

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



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



Title	Shrek 2				
Run Time	92	Country	USA	Release Date	2004
Genre	Action	Director	Mikee Stelz		
Cast	Mike Myers ... Shrek (voice) Eddie Murphy ... Donkey (voice) Cameron Diaz ... F (voice) John Cleese ... King (voice) Rupert Everett ... Prince Charming (voice) Jo Page/Eli/Moblerus/Woblerus's Son (voice) Cody Cameron ... Pteocchio/Steve (voice) Christopher Knights ... Blind Mouse (voice) David P. Smith ... Harold/Mar				
Plot Outline	The film picks up right where the first movie ended... Shrek and Fiona return for The only problem is that they have no idea that their daughter is now an ogre.				

```

DNS: /trailers/Shrek2.mpg
Created: 2005-10-13 17:23:58.000
Modified: 2005-10-13 17:23:58.000
Size: 6100996
Type: File
Expires: Never
GUID: 0000000e-7e0d-134e-a731-010e00000000
Created: 2005-10-13 17:24:00.000
Modified: 2005-10-13 17:24:00.000
Size: 6100996
Checksum: 00000000
Status: 0
User: /C:/IT/0*00100A/00*Personal Certificate/5*DNAP
Group: egee-group
User rights: pdrvl-gs
Group rights: pdrvl-gs
  
```

- **gLibrary**

- **Motivations**

- Huge amounts of data can be saved on SEs, but how can we easily find later a file that we need?
  - (if you have good memory, its **GUID** could be a solution but it is **not so easy**)
  - **File Catalogues** just let us to arrange files in folders and subfolders, **no way to query on their contents**
  - **Metadata Catalogues** are a possible solution, but not always “affordable” especially for non expert users (**powerful but complex to use**)

- **Requirements**

- easy to use, fast, secure, extensible
- **Multimedia files**
  - Images
  - Movies
  - Audio Files
  - Office Documents (Powerpoint, Word, Excel, OpenOffice)
  - E-Mails, PDFs, HTMLs
  - Customized versions of well-know document type (ex. EGEE PPTs)

- **Example 1:**
  - Locate all theoretical PowerPoint presentations (Type) about FireMan (Keywords) written in 2005 (Date);
  - Find all the movies (Type) in which Julia Roberts (Cast) performed together with Hugh Grant (Cast) produced in USA (Country) in 2004 (ReleaseDate);
  - Find all the audio files (Type) in mp3 (Format) of Alanis Morissette (Singer) that last more than 3 minutes (Runtime).
- **Example 2:**
  - A doctor is looking for brain (keyword) DICOM (Type) images of male (Gender) patients older than 65 (Age).
- **Example 3:**
  - A job can work as a storage crawler: it scans pre-existing files in Storage Elements to extract relevant metadata that will be published on gLibrary for further data mining.

- It is built on top of many gLite grid services: a **Metadata Catalogue + File Catalogue + Storage Elements**
- The **SEs** to contain Files
- The **File Catalogues** (LFC and/or FiReMan) to map files locations
- The **Metadata Catalogue (AMGA)** to store and organize metadata in order to provide information about their type and contents.
- **gLibrary** defines the following collections:
  - /gLibrary contains generic metadata for each entry (main collection)
  - /gLAudio, /gLImage, /gLVideo, /gLPPT, /EGEEPPT, /gLDoc, ... (derived collection for “additional features”)
  - /gLTypes
    - It keeps the associations between document types and the names of the collection that contains the “additional features”
    - It is used by gLibrary to find out where it has to look when new document types are added into the system (extensibility)
  - /gLKeys is used to store Decryption Keys

<b>Collection</b>		<b>/gLibrary</b>		
<b>Entry Names</b>	<b>Attributes</b>			
	<b>FileName</b>	<b>PathName</b>	<b>Type</b>	<b>Submitter</b>
4ffaafc8-26e7-4826-b460-3d5bf08081a4	DedicatoAte.mp3	/grid/gilda/calanducci	Audio	Tony Calanducci
00454dca-a269-4b93-8a45-c4012af05600	ardizzonelarocca_is_231005.ppt.gpg	/grid/gilda/calanducci/EGEE	EGEEDOC	Tony Calanducci

## /gLibrary (continuum)

<b>Attributes</b>				
<b>SubmissionDate</b>	<b>Encryption</b>	<b>Description</b>	<b>Keywords</b>	<b>CreationDate</b>
2006-01-05 00:00:00	false	Canzone delle vibrazioni che ha ricevuto un enorme successo tra i teenagers nel 2003	Vibrazioni	2004-02-05 00:00:00
2005-01-05 16:44:22	true	gLite Information System	R-GMA, RGMA, BDII, IS	2005-10-05 23:40

<b>Collection</b>	<b>/gLTypes</b>
<b>Entry names</b>	<b>Attributes</b>
	<b>Path</b> ( <i>refers to a collection</i> )
Audio	/gLAudio
Image	/gLImage
Video	/gLVideo
Documents	/gLDOC
PowerPoint	/gLppt
EGEEDOC	/EGEEPPT

<b>Collection</b>	<b>/gLKeys</b>
<b>Entry names</b>	<b>Attributes</b>
	<b>Passphrase</b>
00454dca-a269-4b93-8a45-c4012af05600	ardizzo

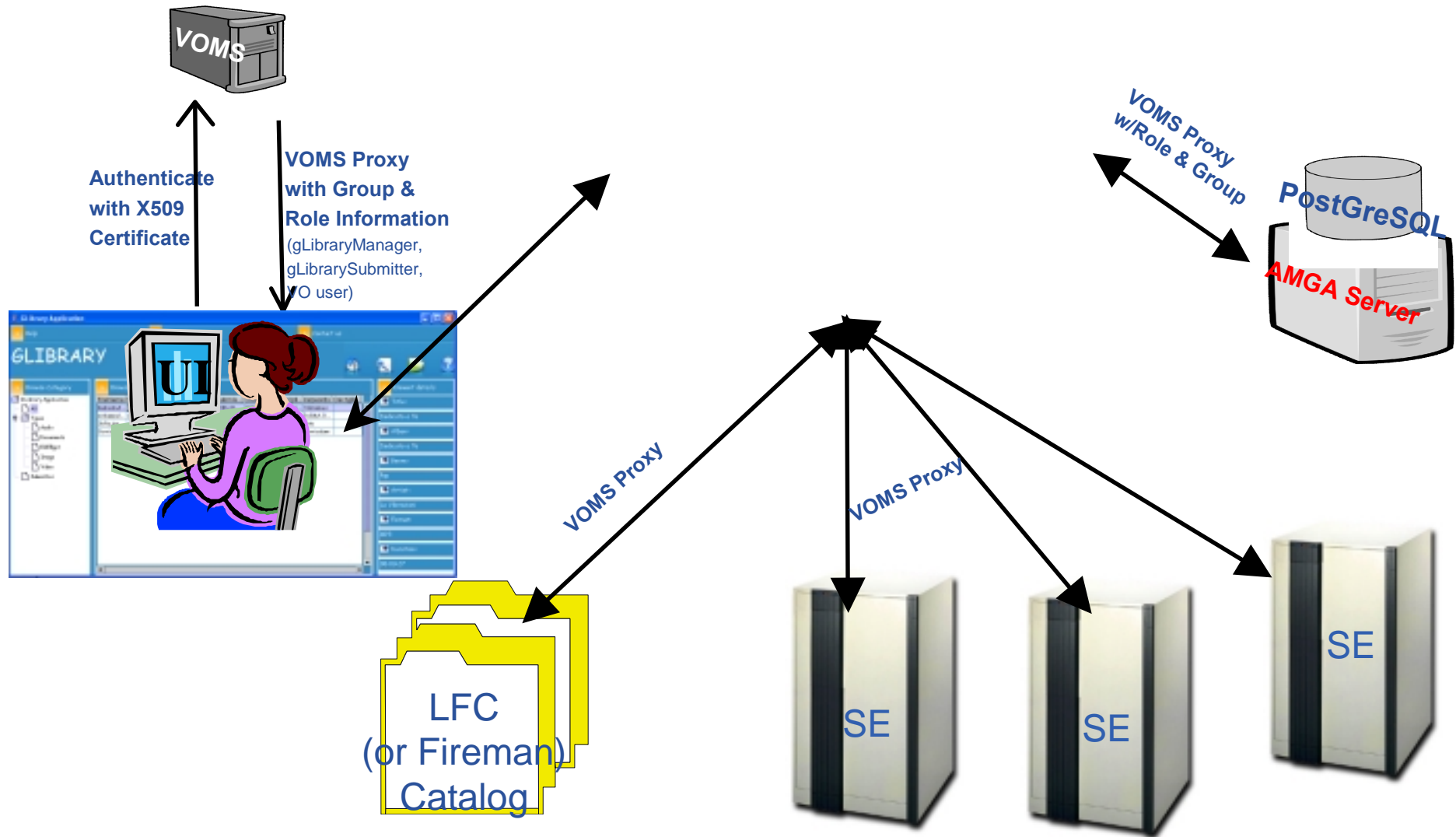
“additional features”

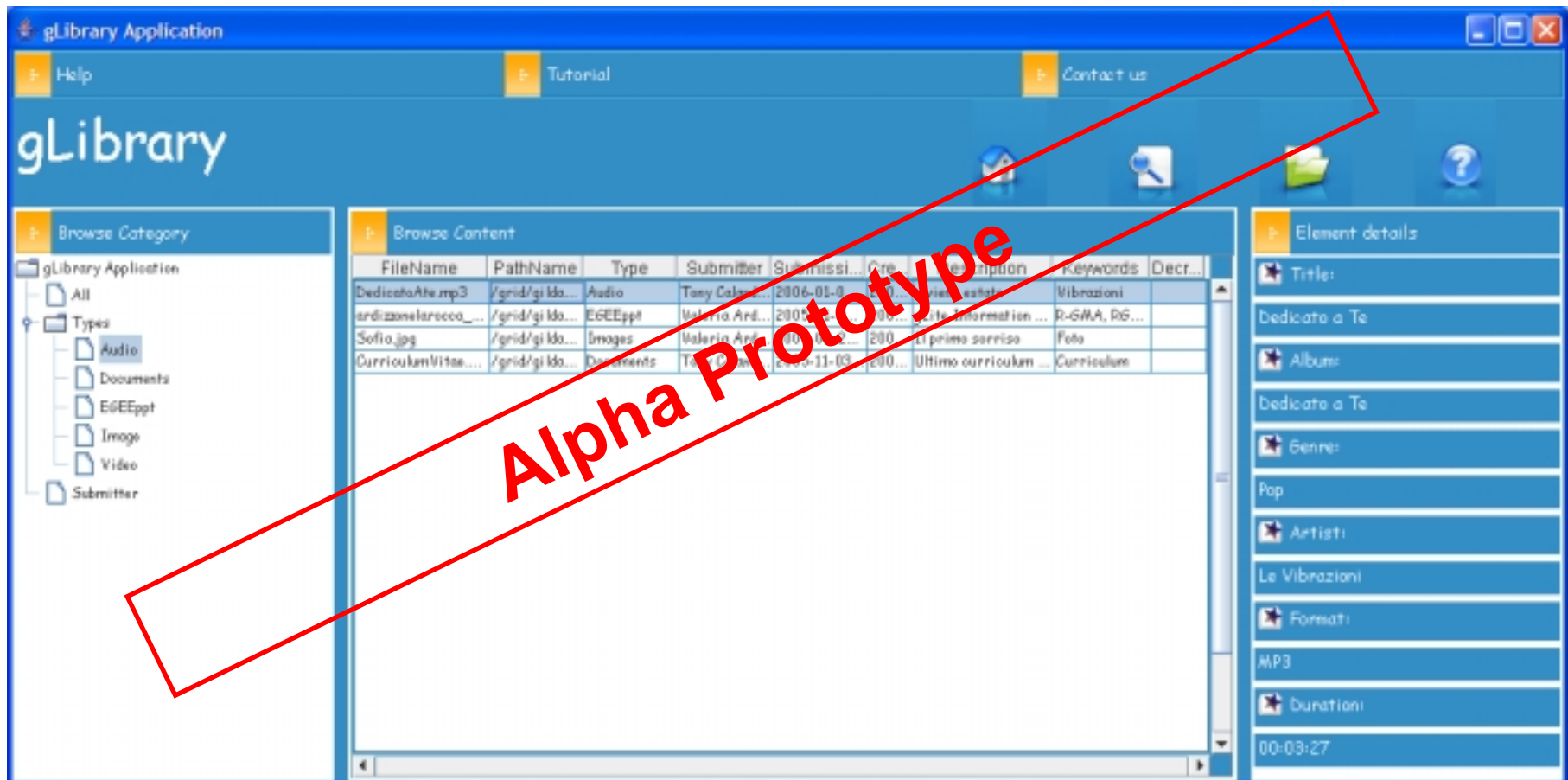
<b>Collection</b>	<b>/EGEEPPT</b>							
<b>Entry names</b>	<b>Attributes</b>							
	<b>Title</b>	<b>Runtime</b>	<b>Author</b>	<b>Type</b>	<b>Date</b>	<b>Event</b>	<b>Speaker</b>	<b>Topic</b>
00454dca-a269-4b93-8a45-c4012af05600	Information Systems	00:30:00	Valeria Ardizzone, Giuseppe La Rocca	Theoretical	2005-10-23	4 <sup>th</sup> EGEE Conference	Giuseppe La Rocca, Valeria Ardizzone	R-GMA, BDII

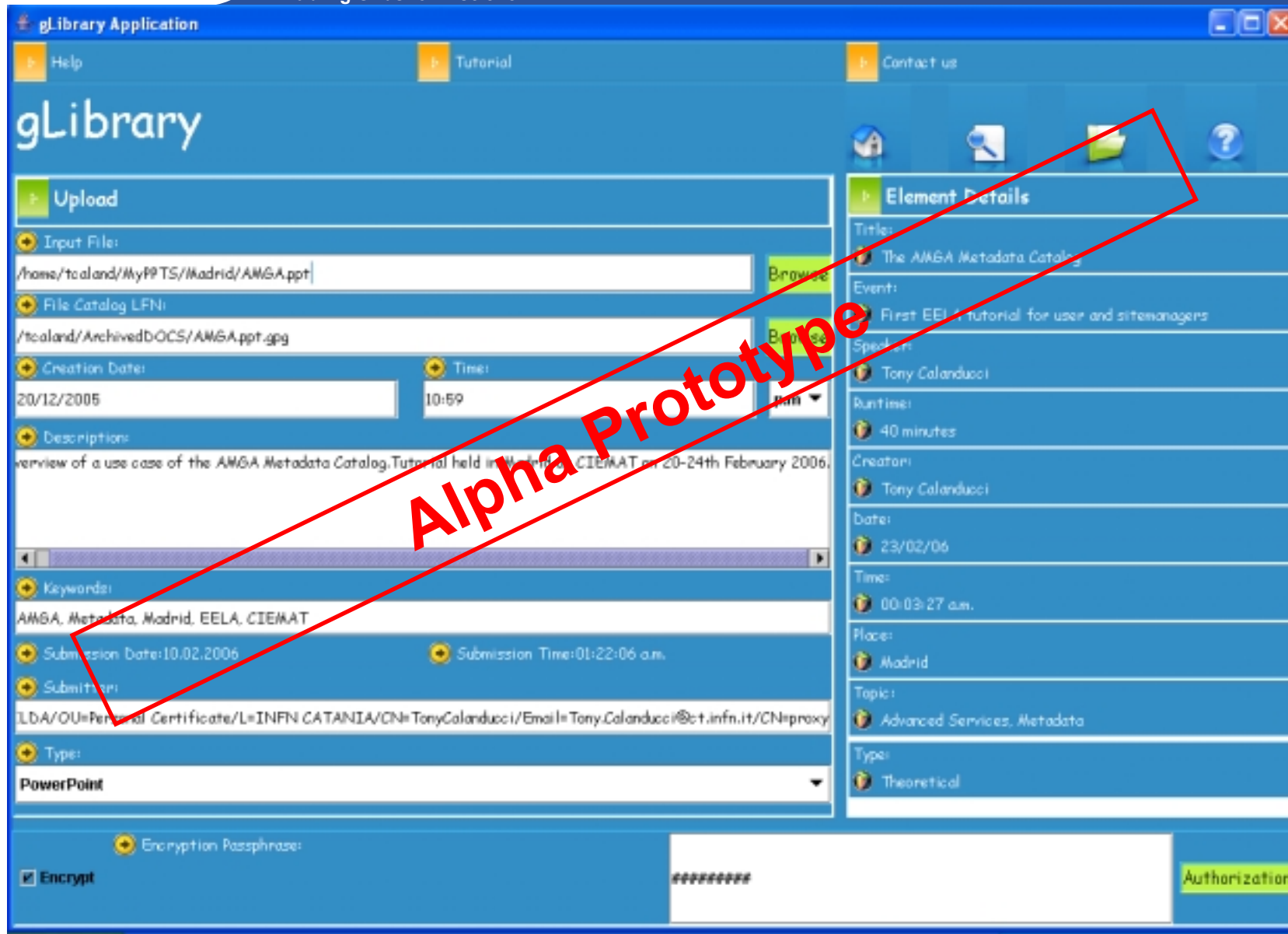
<b>Collection</b>	<b>/gLAudio</b>					
<b>Entry names</b>	<b>Attributes</b>					
	<b>SongTitle</b>	<b>Duration</b>	<b>Album</b>	<b>Genre</b>	<b>Singer</b>	<b>Format</b>
4ffaafc8-26e7-4826-b460-3d5bf08081a4	Dedicato A Te	00:03:27	Dedicato A Te	Pop	Le Vibrazioni	MP3

- **User Requirements:**
  - a valid proxy with VOMS extensions
  - VOMS Role and Group needed to be recognized by gLibrary as a contents manager.
- **3 kinds of users:**
  - **gLibraryManager:** (s)he can create new content type and allows a generic VO user to become gLibrarySubmitter
  - **gLibrarySubmitters:** they can add new entries and define access rights on the entries they create.
    - Fine-grained permission (reading, writing, listing, decrypting) settings on each entry: whole VO members, VO groups, list of DNs
  - **generic VO users:** browse and make queries (on entries they have access to)
- **Basic level of cryptography:**
  - New files saved on SEs can be encrypted beforehand with a symmetric passphrase that will be saved in /gLKeys. Only selected users (that have a specific DN in the subject of their VOMS proxy) can access the passphrase and decrypt the file.

- **Heavy exploitation of AMGA features**
  - support for VOMS proxy authentication
  - fine-grained authorization capabilities to set ACLs per entry basis to restrict access to the decryption keys.
    - Allow gLibrarySubmitters to control which users (based on DNs, VOMS Roles and Groups) can list and get the attributes' value for the submitted entries
- **GUI Front-ends (to achieve the “easy of use” promise):**
  - Java SWING GUI to be run on a Grid UserInterface (JVM required) -- prototype is under way
  - Portlet based front-end will be deployed in GENIUSPHERE and made available for any other JSR168 compliant portlets container
    - Both use AMGA Java APIs





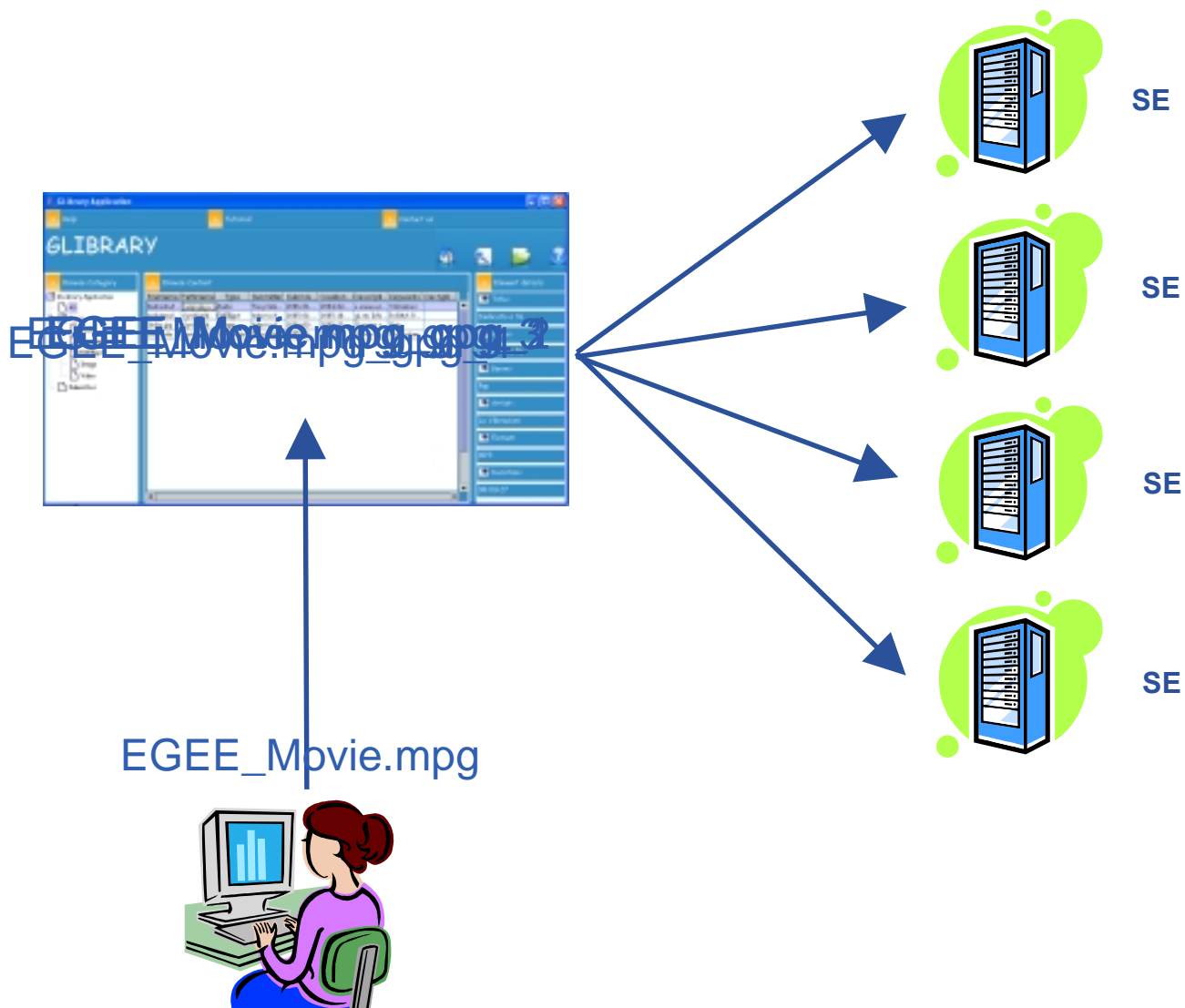


The screenshot displays the gLibrary application interface. The main window is titled "gLibrary Application" and features a navigation bar with "Help", "Tutorial", and "Contact us" buttons. The interface is divided into two main sections:

- Upload Section:**
  - Input File:** /home/tealand/MyPPTS/Madrid/AMGA.ppt
  - File Catalog LFN:** /tealand/ArchivedbOCS/AMGA.ppt.gpg
  - Creation Date:** 20/12/2005
  - Time:** 10:59
  - Descriptions:** overview of a use case of the AMGA Metadata Catalog. Tutorial held in Madrid, CIEMAT on 20-24th February 2006.
  - Keywords:** AMGA, Metadata, Madrid, EELA, CIEMAT
  - Submission Date:** 10.02.2006
  - Submission Time:** 01:22:06 a.m.
  - Submitter:** /LDAP/OU=Personal Certificate/L=INFN CATANIA/CN=TonyCalanducci/Email=Tony.Calanducci@ct.infn.it/CN=proxy
  - Type:** PowerPoint
  - Encryption Passphrase:**  Encrypt
- Element Details Section:**
  - Title:** The AMGA Metadata Catalog
  - Event:** First EEL tutorial for user and sitemanagers
  - Speaker:** Tony Calanducci
  - Runtime:** 40 minutes
  - Creator:** Tony Calanducci
  - Date:** 23/02/06
  - Time:** 00:03:27 a.m.
  - Place:** Madrid
  - Topic:** Advanced Services, Metadata
  - Type:** Theoretical

A red diagonal watermark reading "Alpha Prototype" is overlaid across the center of the interface. At the bottom right, there is an "Authorization" button.

- **Splitting of big files among several SEs (different chunks stored in different SEs):**
  - Enforce security of data: even if a chunk is intercepted it has no meaning alone.
  - Increase upload/download bandwidth
  - Possible implementation:
    - one more NumberOfChunks attribute in /gLibrary collection.
    - /gLChunks collection keeps track of FirstChunkGUID-Chunk#-ChunkGUID
- **Automatic extraction and population of metadata for well known document types**
  - use of GNU libextractor to extract metadata from HTML, PDF, PS, OLE2 (DOC, XLS, PPT), OpenOffice (sxw), StarOffice (sdw), DVI, MAN, MP3 (ID3v1 and ID3v2), OGG, WAV, EXIV2, JPEG, GIF, PNG, TIFF, DEB, RPM, TAR(.GZ), ZIP, ELF, REAL, RIFF (AVI), MPEG, QT and ASF
  - use of Lucene algorithm for indexing document types containing text
- **Evaluation of gLite Hydra Key Store to save decryptions keys**



- Born as an use case to demonstrate AMGA features
- Built on top of many gLite services
- Considering collaboration and integration with NA3 Document Digital Library System
- **Fast** → thanks to AMGA
- **Secure** → ACLs, encryption, and splitting
- **Easy to use** → User friendly Java GUI and portal soon available
- Easily **extensible** to support any document types (Medical Images and files, Invoices, Proceedings, Scientific Publications, Newspapers clips, ...)

