The GENIUS Grid portal with robot certificate: a success story from Bioinformatics

Giuseppe LA ROCCA

INFN - Sezione di Catania

giuseppe.larocca@ct.infn.it

Workshop CCR-INFN Grid 2010,

17-21 May 2010,

Acireale

Outline

- **Current state-of-the-art for the Grid Security**
  - **Introduction to Robot certificates in e-Science;**
  - **Installation and Configuration.**

- **The Genius Grid Portal & Robot certificates**
  - **Architecture;**
  - **The Users Tracking System.**

- **The use case: phylogenetic analysis on large scale**
  - **Main services, features;**
  - **Integration.**

- **Conclusions**

Grid Security: where are we now ?

- **Grid technology allows users to share a wide *pletora* of distributed computational resources regardless of their geographical location, but unfortunately…**

Virtual services are exposed to the users through rather complex Command Line Interfaces or API languages;

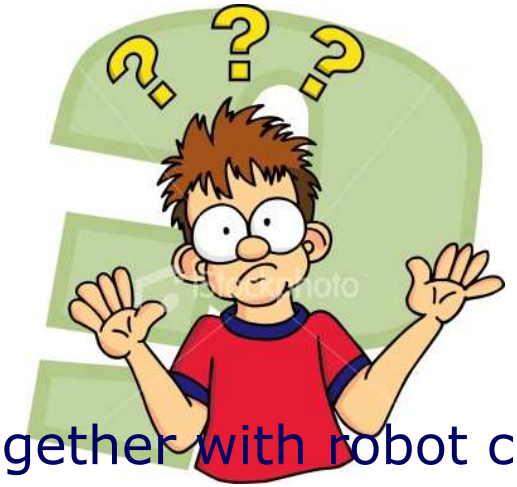
Grid security is indeed based on the Public Key Infrastructure (PKI) of X.509 certificates and the procedure to get and manage those certificates is unfortunately not straightforward;

Up to now, the high security policy requested to access distributed computing resources has been a rather *big limiting factor* when trying to broaden the usage of Grids into a wide community of users;

Grid Security: where are we now ?

User has to be a member of a Virtual Organization (VO) before to access Grid infrastructures;

User needs an account on one of the trusted User Interface (UI) for the experiment.



Grid portals together with robot certificates can provide an added value to make Grids more appealing for non-expert users.

Robot certificates in a nutshell

- **Robot certificates have been introduced to allow users, who are not familiar with deal personal certificates and don't belong to any VOs, to experience the Grid paradigm for research activity and reduce the initial barriers.**
  - They are extremely useful, for instance, to automate grid service monitoring, data processing production, distributed data collection systems;
  - Basically these certificates can be used to identify a person responsible for an unattended service or process acting as client and/or server.

Function

Subscriber Name

```
Your identity: /C=IT/O=GILDA/OU=Robots/L=INFN
  Catania/CN=Robot:MrBayes - Giuseppe La Rocca
Creating temporary proxy
  ............................... Done
Contacting  voms.ct.infn.it:15001
  [/C=IT/O=INFN/OU=Host/L=Catania/CN=voms.ct.infn.it]
  "gilda" Done
Creating proxy
  ...................................................
  ........................... Done
Your proxy is valid until Thu May  8 21:42:05 2008
```

Robot Certificates & tokens

- **In order to strong reduce the risks to have the portal certificate compromised, the INFN CA decided to issue this new certificate on board of the Aladdin eToken PRO smart cards.**

- **Each smart card can support several robot certificates: one for each application user wants to share with others.**
  - An user's PIN is prompted every time user try to read the certificate stored on the smart card to generate a proxy;
  - A first prototype of Grid Portal (https://glite-tutor1.ct.infn.it) using robot certificate to generate an user's proxy has been successfully designed.

Installation & Configuration /1

eToken PRO Smartcard Specifications	
• Operating systems	Windows 98/98SE/Me/2000/XP/NT4.0 SP6 and later/Vista; Mac OS X; Linux
• API & standards support	PKCS#11 v2.01, CAPI (Microsoft Crypto API), Siemens/Infineon APDU commands
	PC/SC, X.509 v3 certificate storage, SSL v3, IPSec/IKE
• Models (by memory size)	32K, 64K
• On board security algorithms	RSA 1024-bit / 2048-bit, DES, 3DES (Triple DES), SHA1
• Security certifications	Common Criteria EAL5/EAL5+ (smart card chip) / EAL4+ (smart card OS)
• ISO specification support	Support for ISO 7816 1 to 4 specifications
• Memory data retention	At least 10 years
• Memory cell rewrites	At least 500,000

Before installing PKI Client 4.55, `PCSC-lite`, `PCSC-lite-lib` and `CCID` packages must be installed in your system

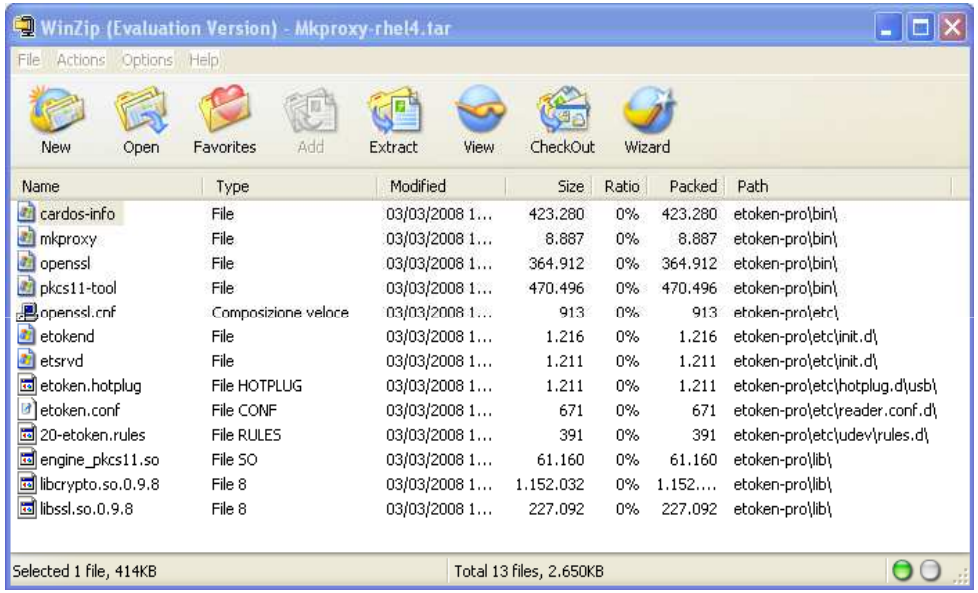
- Maybe you can find these packages in your repo.
  - These packages have dependencies between each other.
- Start the daemon: `/etc/init.d/pcscd start`

The eToken PKI Client includes all the necessary files and drivers to support eToken integration.

- It also includes the eToken Properties configuration tool, which enables easy user management of the eToken password and name.
- Install: `rpm –ivh pkiclient-full-4.55-34.i386.rpm`

Installation & Configuration /2

- **The `Mkproxy-rhel4.tar.gz` tarball contains all the required binaries for RHEL4 compatible platforms.**



Name	Type	Modified	Size	Ratio	Packed	Path
cardos-info	File	03/03/2008 1...	423.280	0%	423.280	etoken-pro\bin\
mkproxy	File	03/03/2008 1...	8.887	0%	8.887	etoken-pro\bin\
openssl	File	03/03/2008 1...	364.912	0%	364.912	etoken-pro\bin\
pkcs11-tool	File	03/03/2008 1...	470.496	0%	470.496	etoken-pro\bin\
openssl.cnf	Composizione veloce	03/03/2008 1...	913	0%	913	etoken-pro\etc\
etokend	File	03/03/2008 1...	1.216	0%	1.216	etoken-pro\etc\init.d\
etsrvd	File	03/03/2008 1...	1.211	0%	1.211	etoken-pro\etc\init.d\
etoken.hotplug	File HOTPLUG	03/03/2008 1...	1.211	0%	1.211	etoken-pro\etc\hotplug.d\usb\
etoken.conf	File CONF	03/03/2008 1...	671	0%	671	etoken-pro\etc\reader.conf.d\
20-etoken.rules	File RULES	03/03/2008 1...	391	0%	391	etoken-pro\etc\udev\rules.d\
engine_pkcs11.so	File SO	03/03/2008 1...	61.160	0%	61.160	etoken-pro\lib\
libcrypto.so.0.9.8	File 8	03/03/2008 1...	1.152.032	0%	1.152....	etoken-pro\lib\
libssl.so.0.9.8	File 8	03/03/2008 1...	227.092	0%	227.092	etoken-pro\lib\

- **After unpacking the tarball, copy over the files to their respective locations:**

```
cp -rp etoken/bin/*  /usr/local/bin
cp -rp etoken/lib/*  /usr/local/lib
cp –rp etoken/etc/openssl.cnf /usr/local/etc
```

Edit /usr/local/bin/mkproxy

larocca

File Edit View Terminal Tabs Help

```
info "Starting Aladdin eToken PRO proxy generation"

# Apply defaults
SLOT=${SLOT:-0}
VALID=${VALID:-12:00}
PROXY_SUGGEST=/tmp/x509up_u`id -u`
PROXY="${X509_USERPROXY:-$PROXY_SUGGEST}"
# the next 3 variables are referenced from openssl.cnnf
export PROXY_PATHLENGTH=${PROXY_PATHLENGTH:-2}
export PROXY_POLICY=${PROXY_POLICY:-normal_policy}
export PROXY_STYLE=${PROXY_STYLE:-legacy_proxy}
BITS=${BITS:-512}

DATE_CMD="date -d"

debug "Output File: $PROXY"

MYDIR=${0%/*}
if [ "$MYDIR" = "${MYDIR#/}" ]
then
    MYDIR=$PWD/$MYDIR
fi
MYDIR=${MYDIR%/*}

export LD_LIBRARY_PATH="$MYDIR/lib:$LD_LIBRARY_PATH"
export PKCS11_ENG="$MYDIR/lib/engine_pkcs11.so"
#export PKCS11_MOD="$MYDIR/lib/libetpkcs11.so"
export PKCS11_MOD="/usr/lib/libetPkcs11.so"
if [ ! -r "$PKCS11_MOD" ]
then
    export PKCS11_MOD="/usr/local/lib/libetpkcs11.so"
fi
export OPENSSL="$MYDIR/bin/openssl"
export OPENSSL_CONF="$MYDIR/etc/openssl.cnf"

if [ `uname -s` = "Darwin" ]
then
    export DYLD_LIBRARY_PATH="$MYDIR/lib"
    DATE_CMD="echo"
elif [ `uname -o` = "Cygwin" ]
then
    export PKCS11_ENG="$MYDIR/lib/engine_pkcs11.dll"
    export PKCS11_MOD="$WINDIR\system32\\\etpkcs11.dll"
    export PATH=$MYDIR/bin:$MYDIR/lib:$PATH
fi

if [ ! -r "${PKCS11_ENG}" ]
then
    echo "Error: cannot find PKCS11 engine (engine_pkcs11) to use." >&2
        exit 2
fi
if [ ! -r "${PKCS11_MOD}" ]
```

Basic Requirements

- **The mkproxy script has been tested on:**
  - **Windows XP (using cygwin)**
  - **Linux Fedora Core 5 and 8**
  - **Linux CentOS 4**
  - **Scientific Linux 4 and 5**
  - **Linux OpenSuse 10 (suse10)**
  - **In the near future we hope to test it on MacOS X.**

Administrating your eToken

- **Before to start initialize your token, set the administrator password and upload your certificate**
- **To access the graphics Quick Function Menu right-click the eToken icon in the system tray or from Start -> Programs -> eToken -> eToken Properties**

Using an Aladdin eToken PRO to generate Grid Proxies

- **Once your grid certificate and private key are safely stored on your eToken, you can generate proxies directly from it.**

```
$ mkproxy
Starting Aladdin eToken PRO proxy generation
Found X.509 certificate on eToken:
label: (eTCAPI) MrBayes's GILDA ID
id: 39453945373335312d333545442d343031612d384637302d32384636363930
Your identity: /C=IT/O=GILDA/OU=Robots/L=INFN Catania/CN=MrBayes
Generating a 512 bit RSA private key
writing new private key to 'proxykey.FM6588'
engine "pkcs11" set.
Signature ok
subject=/C=IT/O=GILDA/OU=Robots/L=INFN Catania/CN=MrBayes/CN=proxy
Getting CA Private Key
PKCS#11 token PIN: ************
Your proxy is valid until: Thu Feb 18 01:22:01 CET 2010
```

Testing the smart card

```
$ pkcs11-tool --module=/usr/lib/libeTPkcs11.so -L

Available slots:
Slot 0          AKS ifdh 00 00
  token label:  eToken
  token manuf:   Aladdin Ltd.
  token model:   eToken
  token flags:   rng, login required, PIN initialized, token initialized, other
flags=0x200
  serial num  :  001c33f9
Slot 1          (empty)
Slot 2          (empty)
Slot 3          (empty)
 [..]
Slot 13         (empty)
Slot 14         (empty)
Slot 15         (empty)
Slot 16         (empty)
```

The GENIUS Grid Portal architecture

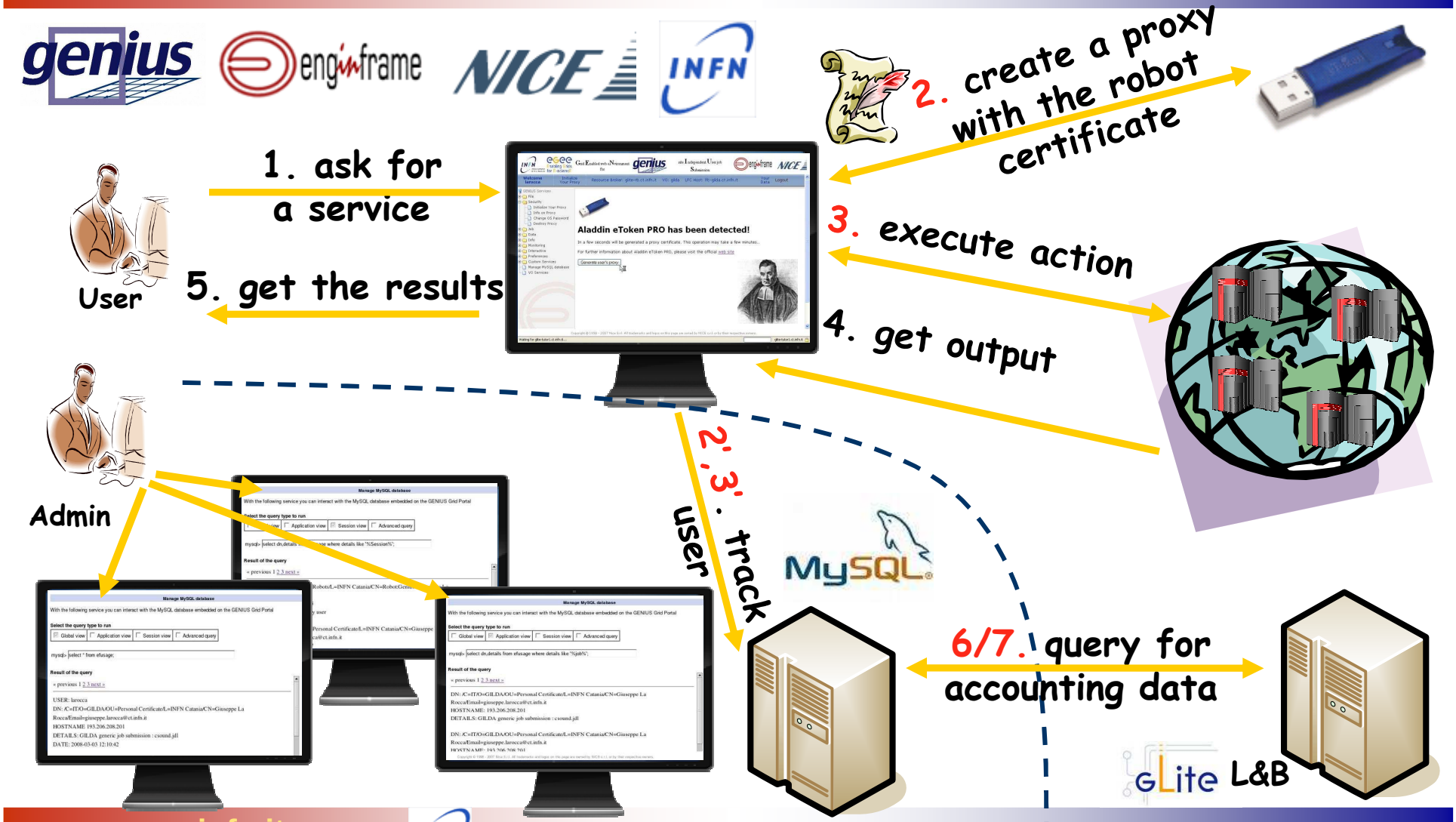
www.enginframe.com

www.nice-italy.com

www.infn.it

- **The GENIUS Grid portal (ver 4.2 is free for educational) is built on top of the EnginFrame Java/XML framework;**
- **It's a gateway to European EGEE Project middleware (it's easily customizable for other middleware);**
- **It allows to expose gLite-enabled applications via web browser as well as Web Services.**

The extended XML/Java EnginFrame framework



The Users Tracking System /1

With the following service you can interact with the User(s) Tracking System embedded on the GENIUS Grid Portal

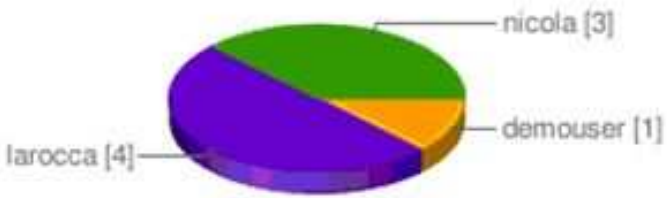
Select the view type

☑ Global view | ☐ Application view | ☐ Session view | ☐ Advanced query

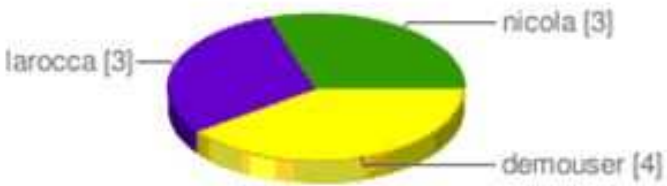
« previous 1 2 next »

#	USER	HOSTNAME	JOBID	TIME STAMP	DETAILS
1	larocca	193.206.208.201	2QHoxZx-4cbtTn8UeGc_6g	2010-01-29 08:57:11	MrBayes+JST job submission : run_job.jdl
2	larocca	193.206.208.201	N/A	2010-01-29 08:51:45	Session started by user
3	demouser	193.206.208.201	UIBlOWFs_XTucSJDt2flRA	2010-01-28 18:05:37	MrBayes+JST job submission : run_job.jdl
4	nicola	193.206.208.201	N/A	2010-01-28 15:30:14	Session closed by user

Session(s) statistics



Application(s) statistics



The Users Tracking System /2

Querying the L&B server **grid-test-53.trigrid.it**

larocca

#	JobID	Running (Time stamp)	Done (Time stamp)	CPU Time (*)
1	HVTeoADSub0OnZoZ4fql_g	2010-03-17 11:52:19	2010-03-17 11:53:26	67
2	CV__GZ9AcHroJTdxrlqxAg	2010-03-24 15:09:00	2010-03-24 15:19:43	643
3	hRz7rNmNsra2IekFYz-F5g	2010-03-24 15:14:01	2010-03-24 15:34:43	1242

demouser

#	JobID	Running (Time stamp)	Done (Time stamp)	CPU Time (*)
1	8J3uOM4nxDZcpGcK0vZDGQ	2010-03-17 11:51:37	2010-03-17 11:52:07	30
2	9WUpa4oUm4gJazSHK1UKeA	2010-03-17 17:32:36	2010-03-17 17:48:18	942
3	9bBvamtDXXTWXn0_H_s6OQ	2010-03-16 10:47:11	2010-03-16 11:02:49	938

nicola

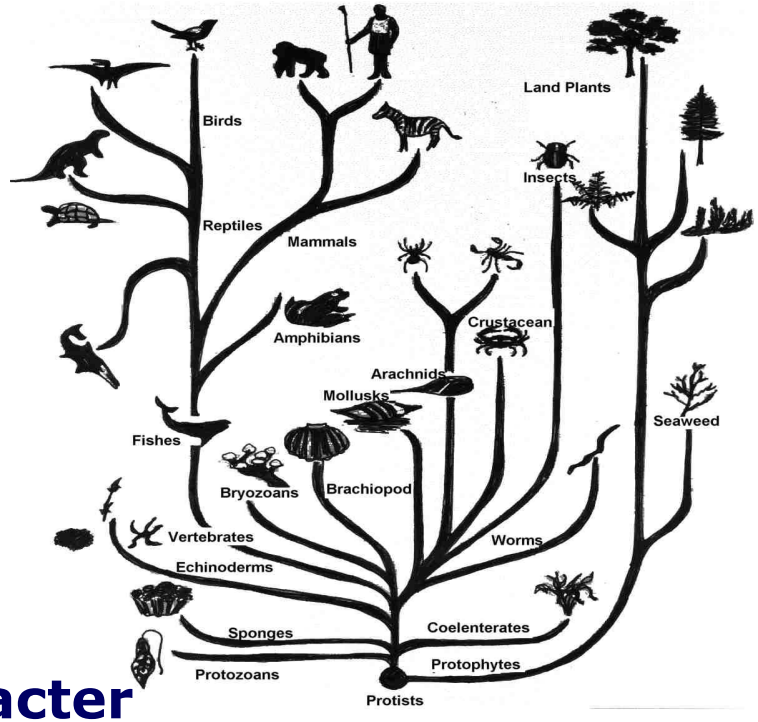
#	JobID	Running (Time stamp)	Done (Time stamp)	CPU Time (*)
1	UHpzlbQoblFZS30-WlpfeQ	2010-03-16 11:32:45	2010-03-16 11:47:57	912

Porting the parallel version of "MrBayes" application to Grid

Case study from

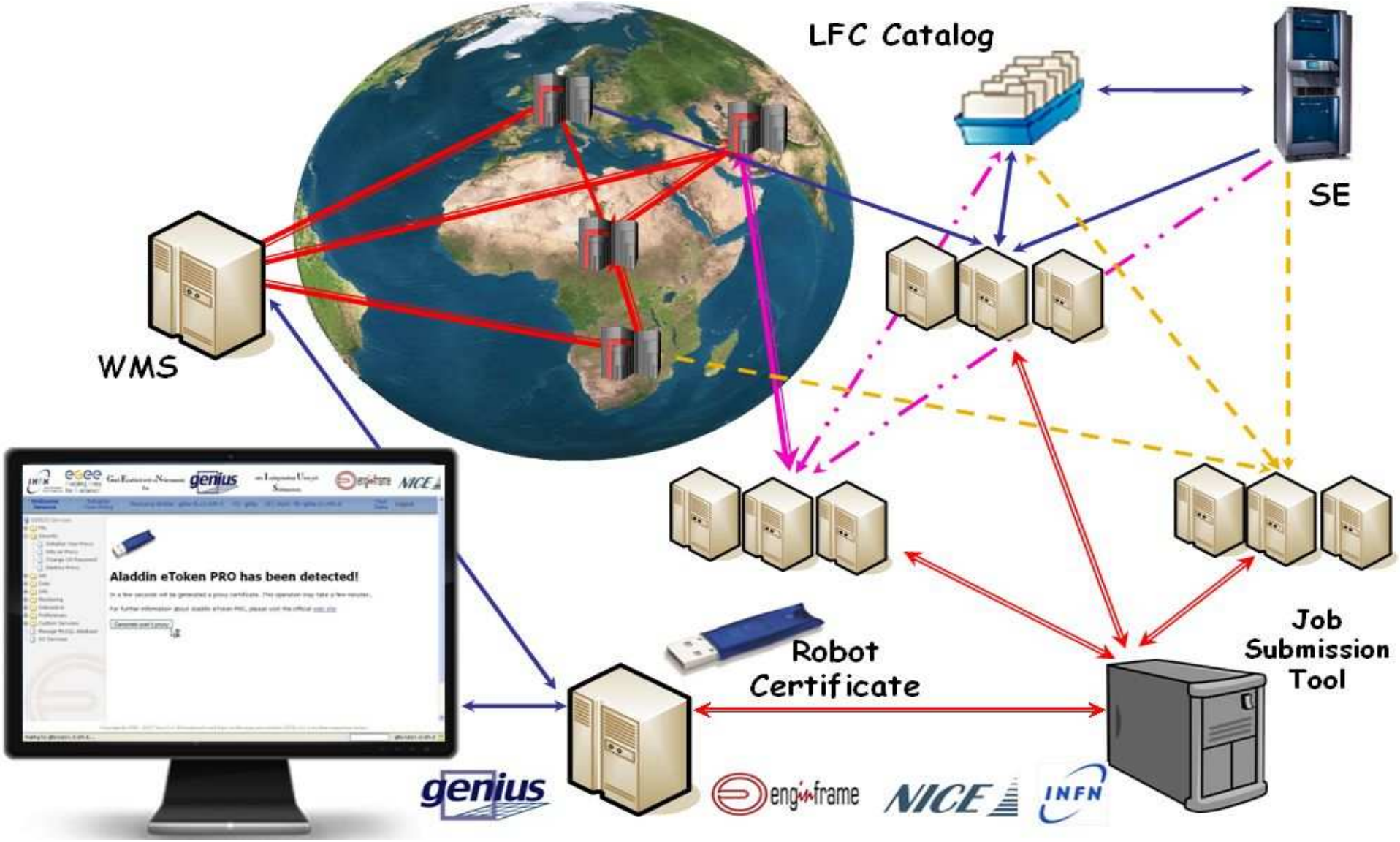
What is Phylogeny ?

- **The Science of estimating the evolutionary past**
  - Fossil data
  - Morphological data
  - Protein sequence data
  - DNA sequence data
  - etc..



- **MrBayes is a program for the Bayesian estimation of phylogeny.**
- **The program takes as input a character matrix in a NEXUS file format and produce some ASCII files in output.**
- **The application is CPU demanding, especially if the MPI version of the software is used.**

Phylogenetic analysis on large scale



JST characteristics

- **Job Submission Tool: is driven by the concept of "Task" as the applications are**
  - Each task could be independent or could be described as depended from another "Task"
  - Each task is described by a **"status"**
  - The task is executed by a wrapper that takes care of monitoring the task:
    - If the task is correctly executed the wrapper can change the status of the task from **"Free"** to **"Done"**
    - If a single step on the job execution fails, the whole task is considered failed and automatically rescheduled
- **JST tool takes care of *submitting* jobs, *retrieving* the output and *monitoring* the status of each task**
- **It is able to deal with accidental failure of grid services**
- **It is possible to change at run time the priority of each task/application**

Conclusions

- **This work is particularly relevant for all users who are not familiar with personal digital certificates.**

- **The valuable benefits introduced by robot certificates in e-Science can be extended to users belonging to different scientific domains, providing an asset in raising Grid awareness to a wide number of potential users.**

Links and References

- **Job Submission Tool (JST) [link]**
- **GENIUS + robot certificate [link]**
- **Using a smart card to generate grid proxies [link]**
- **The Aladdin eToken [link]**

For any information or enquiry:

Roberto BARBERA [roberto.barbera@ct.infn.it]
Alberto FALZONE [alberto.falzone@nice-software.com]
Giacinto DONVITO [giacinto.donvito@ba.infn.it]
Giorgio P. MAGGI [giorgio.maggi@ba.infn.it]
Giuseppe LA ROCCA [giuseppe.larocca@ct.infn.it]
Saverio VICARIO [saverio.vicario@gmail.com]
Luciano MILANESI [luciano.milanesi@itb.cnr.it]