

*AFS & Kerberos Best Practices Workshop
University of Michigan, Ann Arbor June 12-16 2006*

Implementing GRID interoperability

G. Bracco, P. D'Angelo, L. Giammarino*,
S.Migliori, A. Quintiliani, C. Scio**, F. Simoni

**ENEA INFO, C.R. ENEA Frascati
V. E. Fermi 45 Frascati ROMA (Italy)
bracco@frascati.enea.it**

(*) CASPUR, Roma Italy, () Esse3Esse, Roma, Italy**

GRID interoperability

The concept of **Computational GRID** originated from the availability of powerful computing systems, distributed over WAN and connected by fast networks.

GRID functionalities “*unique authentication, authorization, resource access and resource discovery*” (Foster & Kesselman “*Anatomy of the GRID*” 2001) are implemented by the GRID **middleware** but a general accepted multi-platform standard does not exist at the moment.

The integration of different GRID infrastructures can nevertheless be attempted by implementing **gateways**.

Resource sharing can be thus obtained, **minimizing the invasiveness** of a specific GRID middleware inside the hosting infrastructure and **reducing the requirements** concerning **platform/OS** and **firewall security**.

Our gateway implementation takes advantage of having OpenAFS in our GRID infrastructure.

Outline

- ENEA-GRID
- EGEE / EGEE-II Grid Project
- Gateway implementation:
 - architecture
 - modified software components for compatibility with AFS
 - gssklog/gssklogd
 - lcmaps
 - Status and discussion

ENEA

[Italian National Agency for New Technologies, Energy and Environment]

12 Research sites and a **Central Computer and Network Service** (ENEA-INFO) with **6 computer centres** managing multi-platform resources for serial & parallel computation and graphical post processing.



ENEA GRID

INFO-ENEA computational resources:

- **Hardware**: ~100 hosts and ~650 cpu : IBM SP; SGI Altix & Onyx; Linux clusters 32/ia64/x86_64; Apple cluster; Windows servers. Most relevant resource: IBM SP5 192 nodes
- **software**: commercial codes (fluent, ansys, abaqus..); elaboration environments (Matlab, IDL, SAS..)

ENEA GRID mission [started 1999]:

- provide a **unified user environment** and an homogeneous access method for all ENEA researchers, irrespective of their location.
- implement tools to facilitate the **integration** of department and individual resources.

ENEA GRID architecture

GRID functionalities (unique authentication, authorization, resource access and resource discovery) are provided using “mature”, multi-platform components:

Distributed File System: **OpenAFS**

Resource Manager: **LSF Multicluster [www.platform.com]**

Unified user interface: **Java & Citrix Technologies**

These components constitute the ENEA-GRID Middleware.

OpenAFS

- user homes, software and data distribution
- integration with LSF
- user authentication/authorization [still kerberos 4]

6 db-servers in 4 sites

- OpenAFS 1.4.0 / Scientific Linux 4.2
- Frascati (3), Bologna, Casaccia, Trisaia

15 File-servers in 6 sites

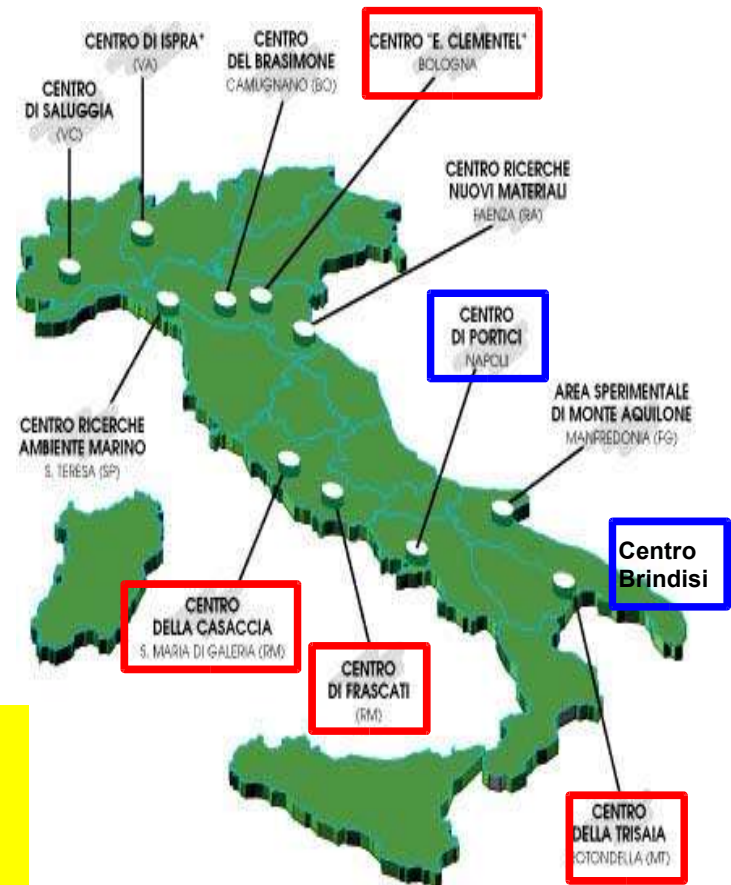
- OpenAFS: Scientific Linux 3.0.4, 4.3, CASPUR BigBox 3.0, Solaris 9, AIX 5.3.
- AFS Transarc 3.6.x / AIX 4.3.3/ Solaris 9

700 registered users

2 TB data

Migration from AFS Transarc to OpenAFS almost completed in the last year.

Kerberos 5 migration for this year.



EGEE/EGEE-II



<http://www.eu-egee.org>: “Expanding from originally two scientific fields, high energy physics and life sciences, EGEE now integrates applications from many other scientific fields, ranging from geology to computational chemistry”

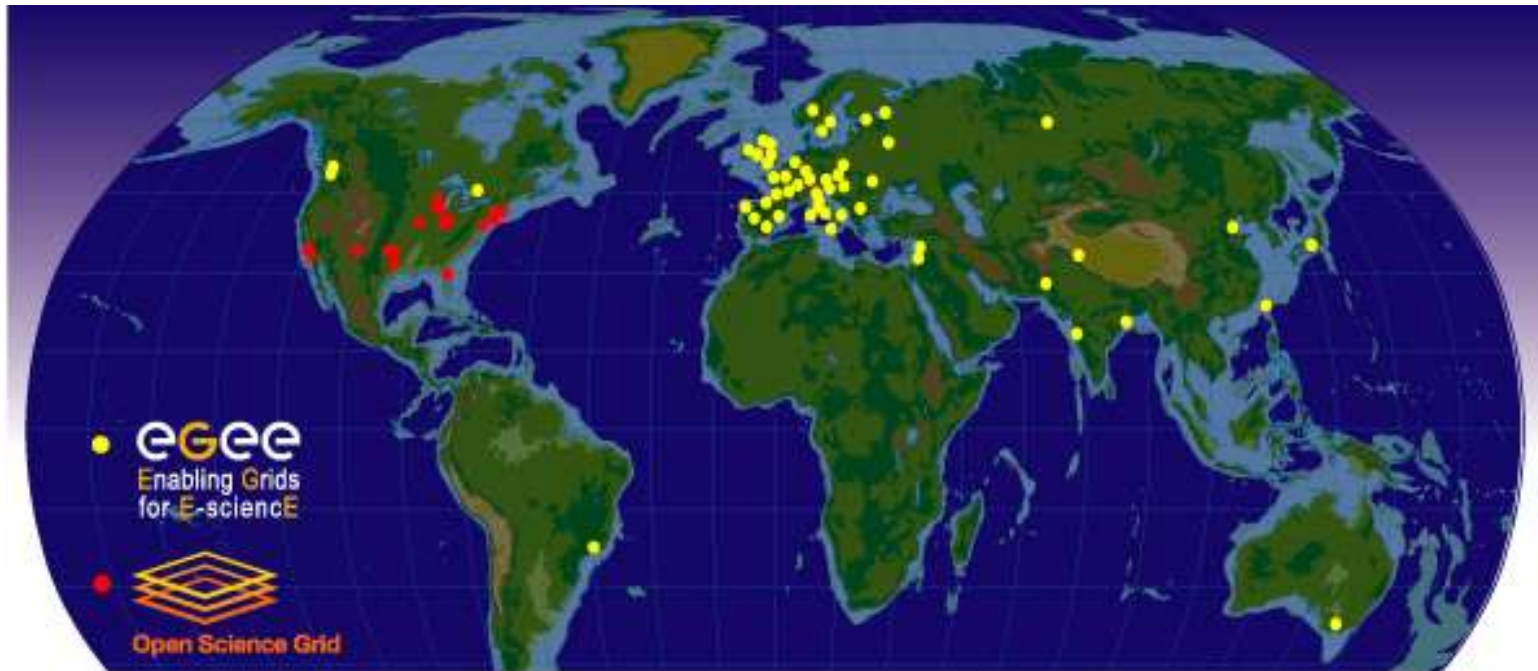
EGEE is one of the two biggest EU GRID projects; [2+2 years] second phase EGEE-II started April 2006

- 90 institutions in over 30 countries, ~30000 cpu
- **ENEA is a funded partner, [~100 cpu 20% of time, AIX]**

EGEE architecture is based both on components (globus, condor,..) coming from previous GRID projects [VDT Virtual Data Toolkit, Datagrid, Datatag,.....] and on new systems developed by the project itself and integrated in the **EGEE middleware: Glite**

EGEE, OSG, LCG

EGEE is integrated with OSG, the US “Open Science GRID”, into LCG [LHC Computing GRID] to provide the resources required by LHC, the “Large Hadron Collider Experiment” [starting at CERN 2007]



A map of the worldwide LCG infrastructure operated by EGEE and OSG.

L. Robertson / CERN <http://hepix.caspar.it/spring2006/>

EGEE middleware

EGEE middleware systems:

- ◆ Authentication and Authorization System [VO]
- ◆ Information System
- ◆ The Workload Management System
- ◆ The Data Management System
- ◆ Monitoring System
- ◆ Accounting System

LCG middleware has been adopted by EGEE at the beginning of the project and the first production version of Glite [3.0] has been very recently released.

This presentation refers to LCG 2.7 middleware version.

EGEE supported platforms:

- ◆ Scientific Linux 3, i386, IA64
- ◆ Porting to others still in progress
<http://www.grid.ie/porting>

[AFS compatibility](#)

gssklog package (<ftp://achilles.ctd.anl.gov/pub/DEE>) permits to get an AFS token from a Globus X509 certificate.

EGEE VO based approach is not supported in gssklog-0.11.

ENEA EGEE site implementation (1)

Basic components of an EGEE site installation:

- ◆ User Interface [UI]
- ◆ Storage element [SE]
- ◆ Computing Element [CE] (with support to LSF)
- ◆ Worker Nodes [WN]

ENEA implementation requirements:

- ◆ EGEE users => mapped to some standard ENEA-GRID AFS users
- ◆ Access to AIX platform => **NO Middleware on Worker Nodes**
- ◆ LSF resources must be used whenever possible
 - **Batch submission**: bsub
 - **Information**: lshosts, bhosts, lshosts, bqueues,...
 - **Prompt job execution**: lsruncmd
 - Batch queues dedicated to EGEE GRID users

ENEA EGEE site implementation (2)

Implementation architecture:

- ◆ UI & SE: Linux, Standard EGEE configuration
- ◆ CE: Linux, modified to become a “gateway”
- ◆ WN: any platform/OS with support for AFS/LSF
 - share user homes with CE using AFS
 - delegate all grid commands concerning file transfer to/from GRID to the CE by means of the LSF lsruncmd command.
 - the CE has the correct security context
 - AFS security and quota management guarantee reliability

ENEA EGEE site implementation (3)

How:

- ◆ Any WN middleware command concerning file transfer is wrapped with a lsrunc script.
- ◆ The wrapped WN middleware is located in AFS.
- ◆ The relevant PATHs in the environment of the EGEE job on the WN are modified so that the wrapped middleware is used
 - LSF wrapper configuration for the GRID dedicated queues
- ◆ The CE is installed with all the packages required by a WN
- ◆ EGEE GRID users require an AFS token:
 - gssklogd modified to be compatible with EGEE middleware
 - EGEE middleware on the CE must properly call gssklog.

EGEE Standard site layout

User Interface

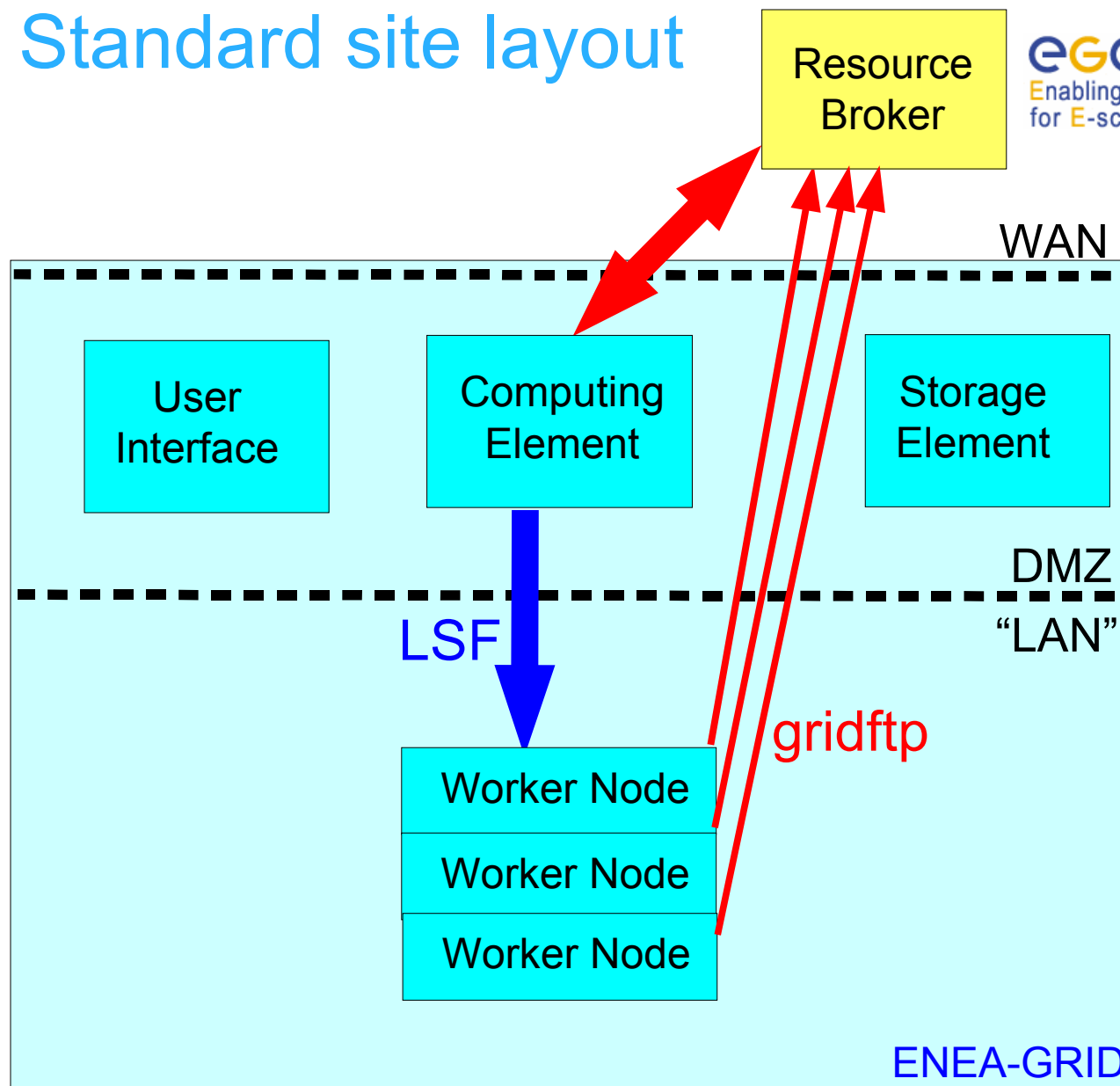
Storage Element

Computing

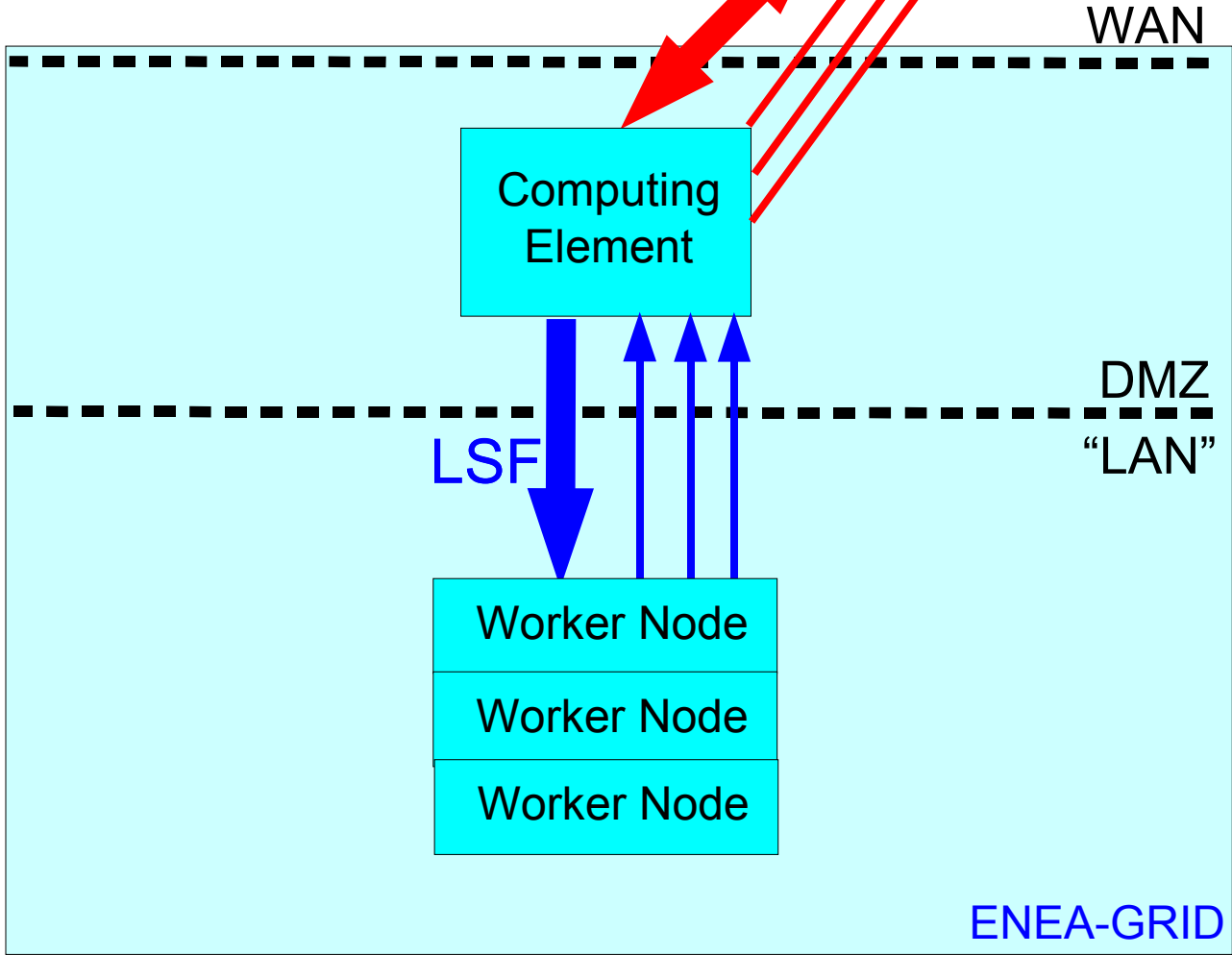
Element: accepts jobs sent by the RB and submits them to the local batch system (LSF)

Worker Nodes:

performe the computation and the WN middleware sends back job results to RB



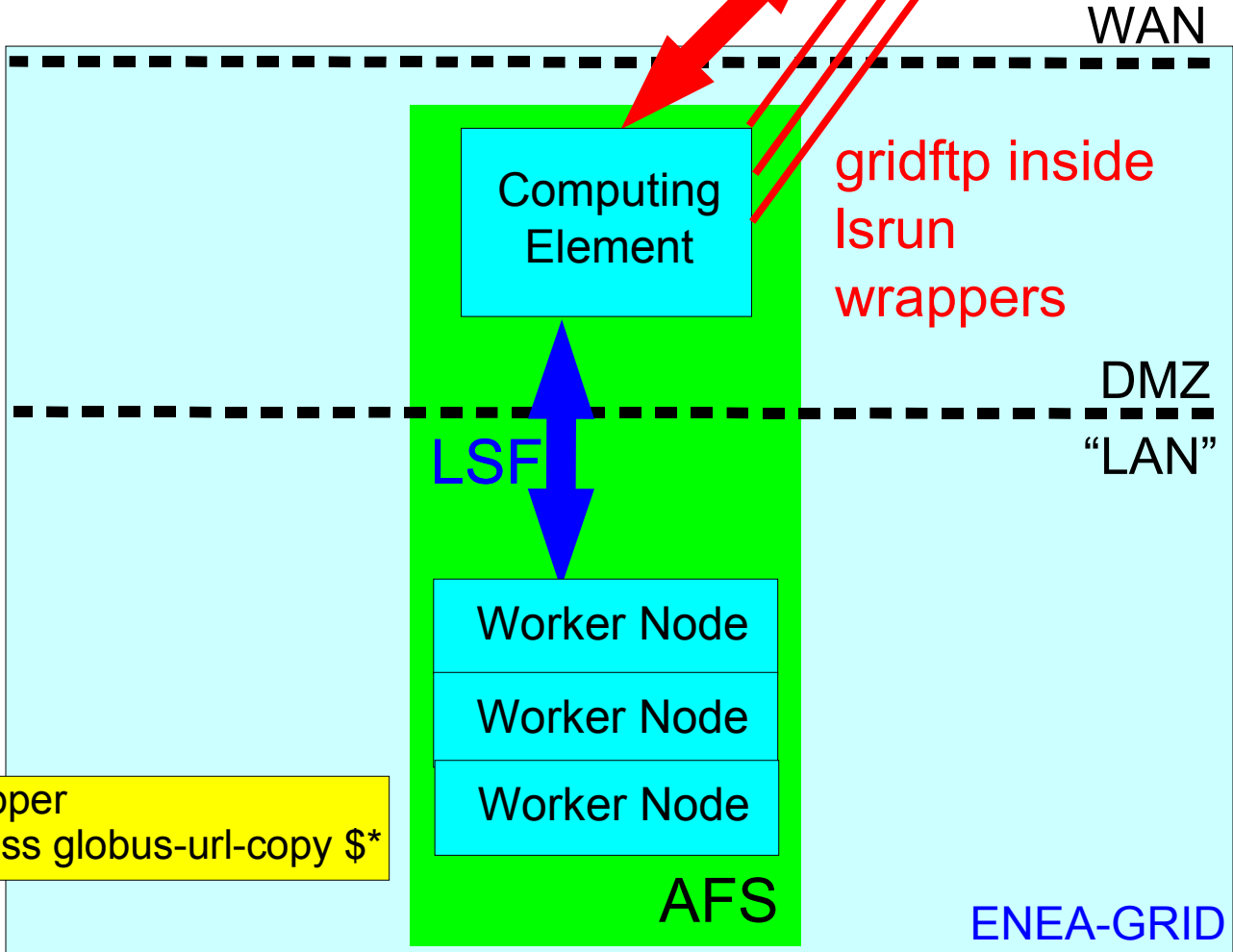
EGEE ENEA site implementation (1)



The Computing Element: a gateway system where all the grid commands are executed.

Worker Nodes: perform the computation and ask the CE to send back job results to RB.

EGEE ENEA site implementation (2)



CE and WN share user homes using AFS.

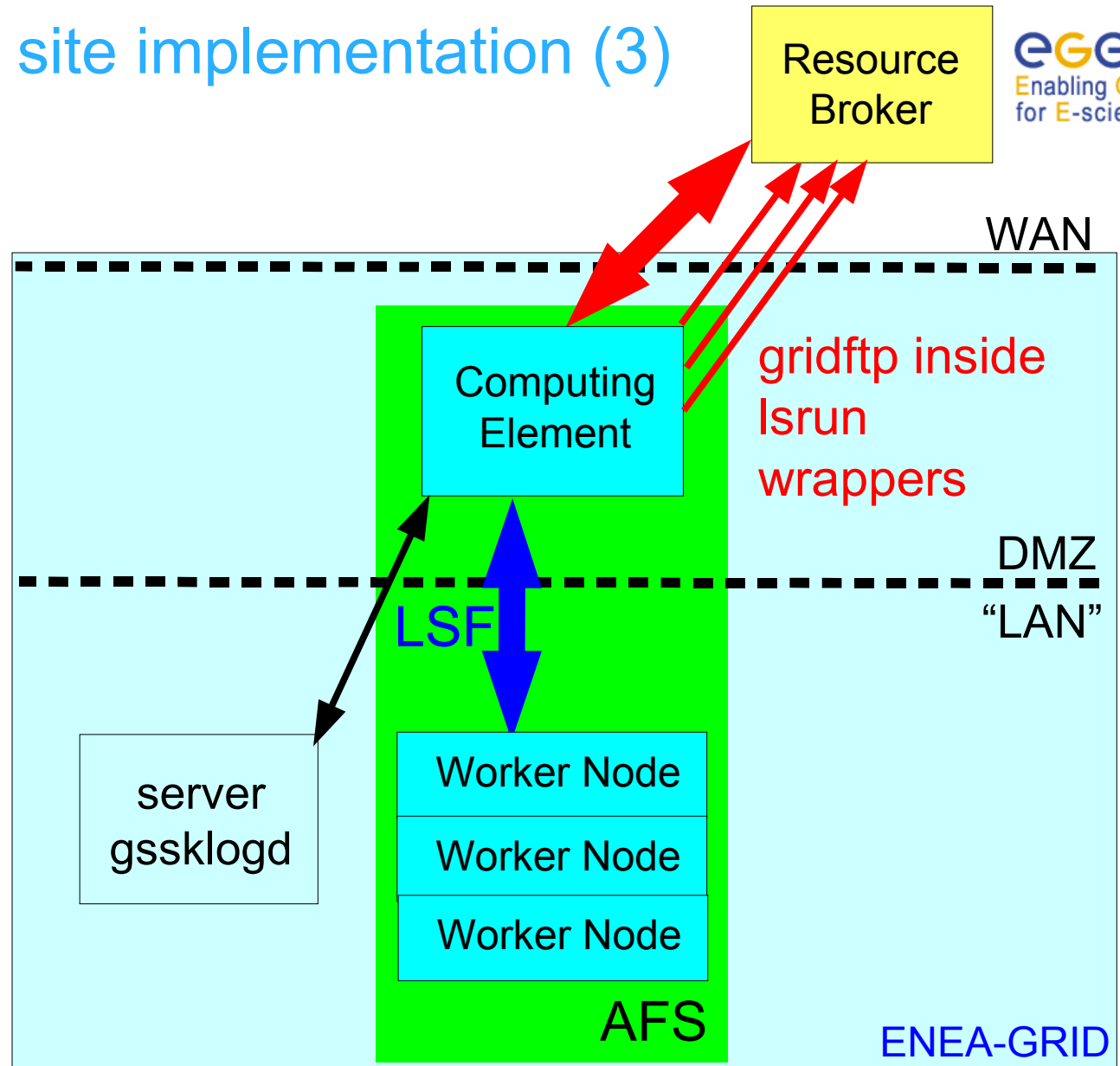
In the environment of the job on the WN the PATHs are defined so that grid data transfer commands are:

- 1) taken from AFS
- 2) wrapped with Irun

e.g. globus-url-copy-> wrapper
wrapper: Irun -m ce.address globus-url-copy \$*

EGEE ENEA site implementation (3)

GRID users are AFS users so a modified gssklogd server is required to obtain an AFS token from the X509 certificate.



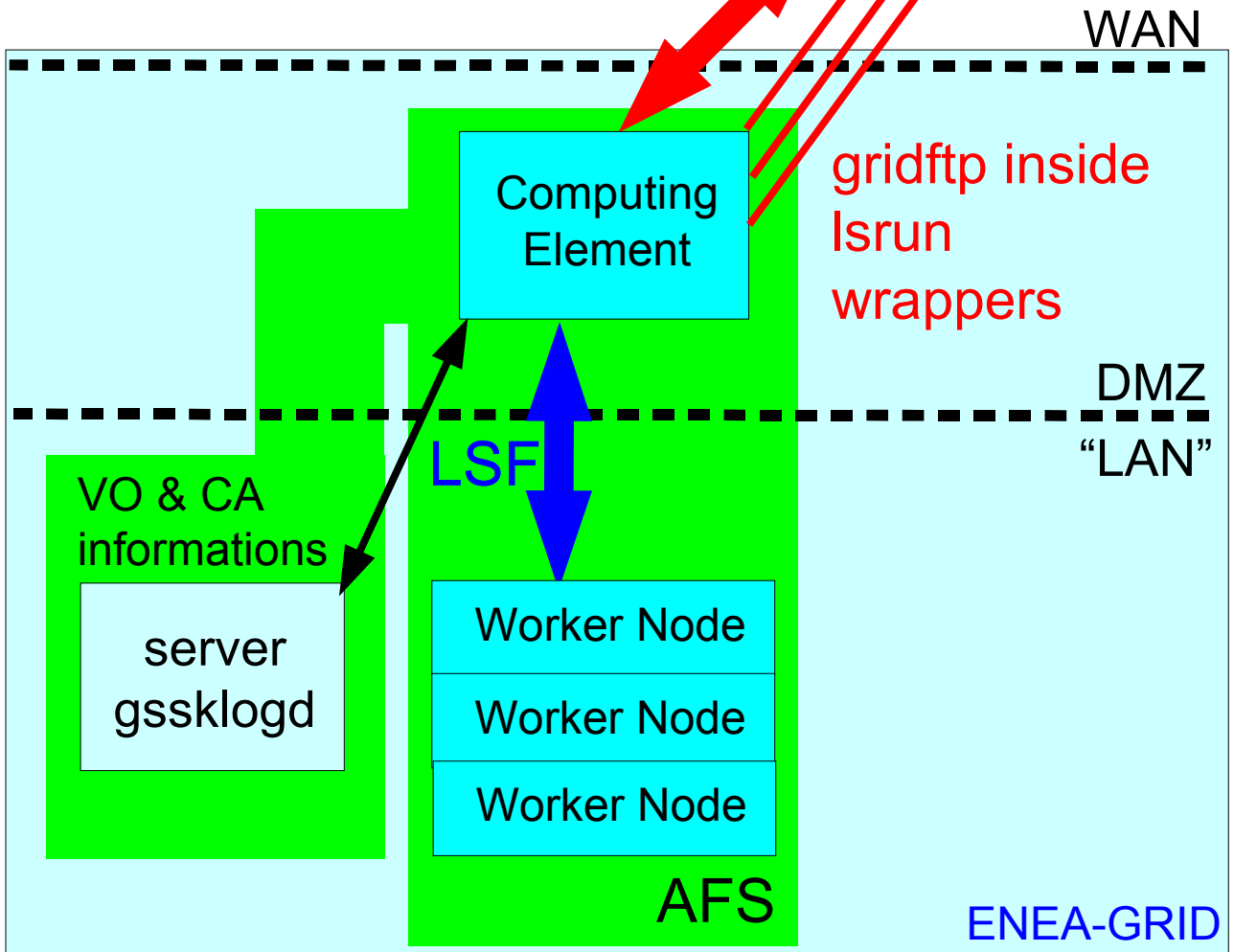
EGEE ENEA site implementation (4)



GRID users are AFS users so a modified gssklogd server is required to obtain an AFS token from the X509 certificate.

AFS can be used to provide to the gssklogd daemon the information about VO and CA

These information are automatically updated by the CE



AFS related software

The implementation has required to modify two components related with AFS as EGEE Grid users need a token.

- User mapping: `lcmaps_afs` in EGEE LCMAPS package
- Token granting for users belonging to a VO: `gssklogd`

User mapping in a EGEE site (1)

LCAS [Local Centre Authorization Service] and **LCMAPS** [Local Credential Mapping Service] are the **EGEE services** that make the connection between GRID users certificates and local UNIX userids on the local site. The systems support:

- The standard globus X509 certificate
 - **gridmap file**: [*distinguished_name local_userid*]
"/O=dutchgrid/O=users/O=sara/CN=Dutch User" griduser
- The extension of X509 certificate holding VO information [VOMS service]
 - **gridmap file**: [*VO_data voms_pool_account_prefix*]
"/VO=enea/GROUP=/enea" .enea

“voms pool account”: predefined local users which are dynamically assigned to the members of the VO when running job on the site.
[e.g. for VO **enea** pool users are **enea001**, **enea002**, ..., **enea020**]

pool accounts can be used also with standard X509 certificates:
"/O=dutchgrid/O=users/O=sara/CN=Dutch User" .dteam

User mapping in a EGEE site (2)

LCMAPS [Local Credential Mapping Service] is a modular service.

A **lcmaps_afs** module exists with the purpose to obtain an AFS tokens for the local mapped user using gssklog but **the standard module is not compatible with VOMS** “pool account”

A modified module has been prepared to achieve compatibility:
gssklog is invoked always with -principal parameter

Note: due to the software architecture of the CE the relevant user processes on the CE can not be started in the same pag shell; in fact they are managed by two independent services:

- **authorization and user mapping** are managed by lcas/lcmaps
- **job submission** is managed by the edg-gatekeeper

The reliability of a **“user based” token** instead that a “pag based” token is delegated to the CE middleware architecture.

gssklogd/gssklog (1)

Gssklog package has been developed by Doug Engert to obtain an AFS token using a GSS [Generic Security Service] implementation.

It has been used together with Globus Toolkit GSSAPI to obtain an AFS token from a X509 certificate.

The package include a daemon: `[gssklogd]` and a client `[gssklog]`.

`gssklogd` runs on a server where the AFS cell keyfile is available and requires a service certificate and key. A map file is also required `/etc/grid-security/afsgrid-mapfile` containing the `X509 distinguished name` and the `AFS userid(s)`:

```
"/O=Grid/O=Globus/OU=anl.gov/CN=John Doe" jdoe,altuser
```

The client `gssklog` requires the proxy file generated by the globus command `grid-proxy-init`; an AFS token is obtained by the command:

```
gssklog -server server_address -principal jdoe
```

gssklogd/gssklog (2)

EGEE uses of an extension of X509 certificate to hold VO information [VOMS] and the proxy is generated using command:

```
voms-proxy-init -voms VO_name
```

gssklogd has been modified to provide tokens for **pool account users** on the basis of the content of the VO information contained in the user certificate. The daemon requires now also the VO certificate.

The **afsgrid-mapfile** syntax has been extended, e.g. for ENEA VO:

```
"/VO=enea" .enea/020/
```

The client **gssklog** provides always the principal name:

```
gssklog -server server_address -principal enea004
```

Some users requires pool account without having VOMS extension

```
"/O=dutchgrid/O=users/O=sara/CN=Dutch User" .dteam
```

Status and discussion

ENEA Linux site

CE: egce.frascati.enea.it

SE: egse.frascati.enea.it

WN: 16 Linux P4, 1.8 Ghz, 1 GB

ENEA AIX site

CE: egceaix.frascati.enea.it

SE: egseaix.frascati.enea.it

WN: 32 cpu SP4, 1.1 Ghz, 32 GB

LIMITS of the gateway approach

- **Not a completely standard site** [but EGEE Certification job runs well]
 - GRID API are not available
 - Some WN monitoring components are unavailable
- **Scalability** : the number of gateway machines can be increased

What next and acknowledgements

What next

- Other platforms: ALTIX, Opteron Cluster, MacOSX..
- Application demonstration

ENEA GRID and enea.it AFS cell are operated with the support of many people in various ENEA sites:

S. Taglienti, R. Guadagni, A. Perozziello, A. De Gaetano, S. Pecoraro, D. Giammattei, G. Mencuccini., M. De Rosa, M. Caiazzo, A. Palumbo, G. Elmo, S. Pierattini, M. Impara, G. Furini, C. Zini...

Many thanks to Vincenzo Ciaschini [INFN, CNAF, Bologna] for helping us with VOMS libraries utilization.