

CRESCO HPC SYSTEM INTEGRATED INTO ENEA GRID ENVIRONMENT

G. Bracco, S. Podda, S. Migliori, P. D'Angelo, A. Quintiliani, D. Giammattei, M. De Rosa, S. Pierattini, G. Furini, R. Guadagni, F. Simoni, A. Perozziello, A. De Gaetano, S. Pecoraro, A. Santoro, C. Sciò*, A. Rocchi, A. Funel, S. Raia, G. Aprea, U. Ferrara, D. Novi**, G. Guarnieri

ENEA-FIM, L.Tevere Thaon di Revel (Roma) Italy, (*) Esse3Esse (**) CRIAI



Summary

This poster describes the architecture of the high performance computing (HPC) system that has been installed to provide the required computing power to the CRESCO project applications and the dedicated activity required to integrate CRESCO HPC system into ENEA-GRID infrastructure.

ENEA, the Italian agency for the energy, environment and new technologies, has a substantial experience in GRID technologies and its multi-platform HPC resources are integrated in the ENEA-GRID infrastructure.

CRESCO (Computational Research Center for Complex Systems) is an ENEA Project, co-funded by the Italian Ministry of University and Research (MUR). The project is functionally built around a HPC platform and 3 scientific thematic laboratories:

- the Computing Science Laboratory, hosting activities on HW and SW design, GRID technology and HPC platform management
- The activities in the Life Science domain, ranging from the "post-omic" sciences (genomics, interactomics, metabolomics) to Systems Biology;
- the Complex Networks Systems Laboratory, hosting activities on complex technological infrastructures, for the analysis of Large National Critical Infrastructures

CRESCO HPC system consists of more than 2700 computing cores, divided into three main sections. A section is dedicated to applications with high memory requirements (42 nodes with 16 cores and 32 or 64 GB memory for a total of 672 cores), a section dedicated to high scalable applications (256 nodes with 8 cores and 16 GB memory, for a total of 2048 cores) and a third experimental section providing systems with Cell processors (4 blades), FPGA (6 VIRTEX systems) and high performance video adapters dedicated to computational applications. High bandwidth and low latency connections are provided by an InfiniBand 4xDDR network. The main storage consists of an IBM/DDN 9550 system with 160 TB raw data, organized in a GPFS file system.

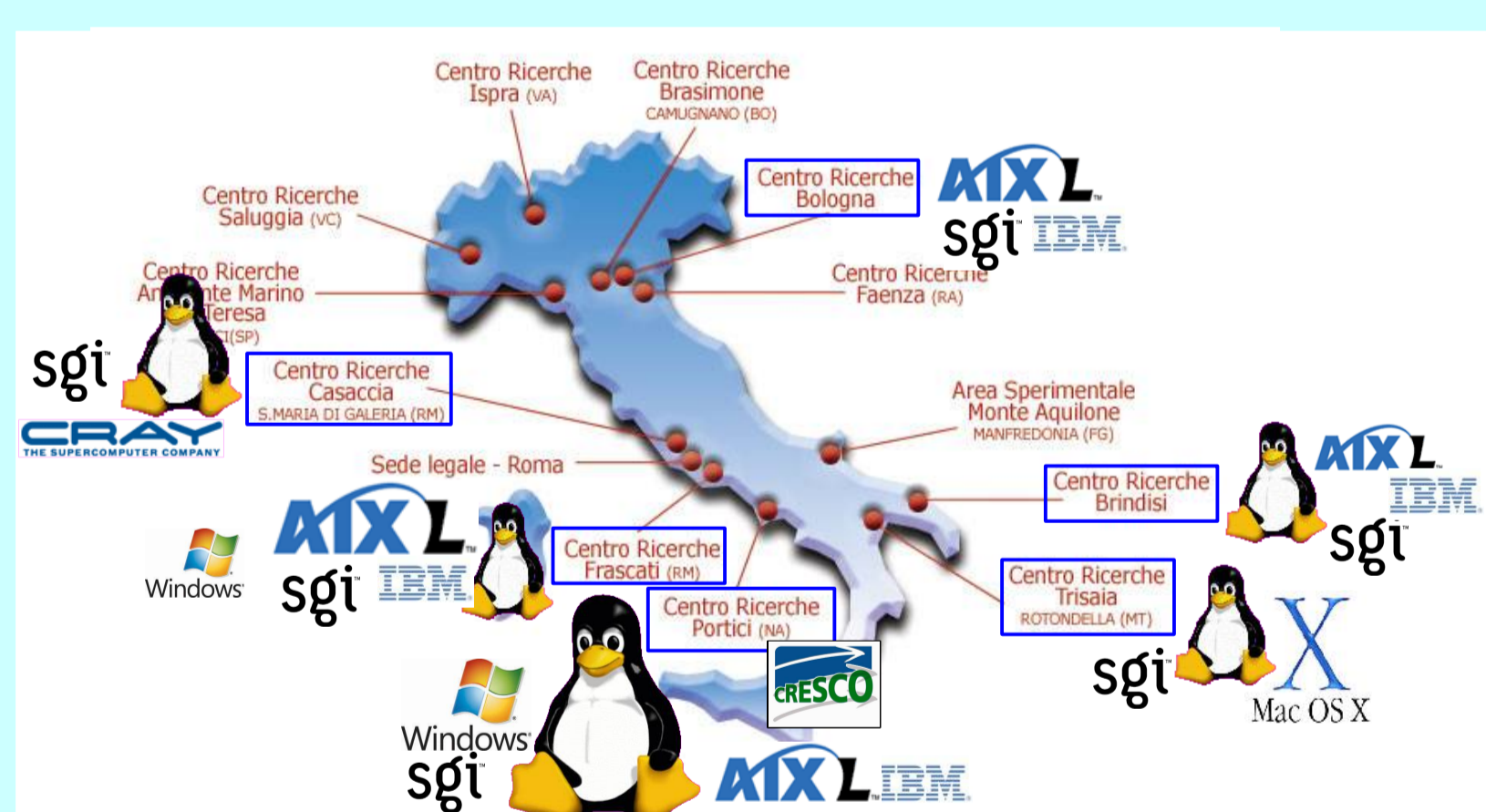
CRESCO HPC system has been integrated into ENEA-GRID infrastructure which has been developed to provide a unified environment for all the main ENEA HPC resources.

The main software components of ENEA-GRID are the multi-site resource manager LSF Multicluster, the OpenAFS distributed file system, the integrated Kerberos 5 authentication and a Java and Web based Graphical User interface making use of CITRIX technologies.

The choice of mature, reliable and multi-platform software components has permitted along the years to integrate in a GRID oriented infrastructure HPC resources at the state of the art performances, with minimal changes in the user environment

ENEA GRID

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



ENEA GRID mission [started 1999]:

- provide an **unified user environment** and an **homogeneous access method** for all ENEA researchers and their collaborators, irrespective of their location.
- optimize the utilization of the available 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; FreeNX**

These components constitute the ENEA-GRID Middleware.

OpenAFS

- user homes, software and data distribution
- user authentication/authorization, Kerberos V

Integration with other GRIDS: EGEE, PON Projects/GRISU

The **SPAGO** [Shared proxy Approach for GRID Objects] technology has been developed by ENEA to implement gateways to other GRID Middleware as gLite.

CRESCO HPC system

CRESCO HPC system has been designed with the aim of offering a general purpose facility based on the leading multi-core x86_64 technology.

The performance for the CRESCO HPC plant set-up has ranked #125 in the June 2008 top500 list with Rmax=17.1 TeraFlops (rank #2 between the Italian HPC systems in the list)

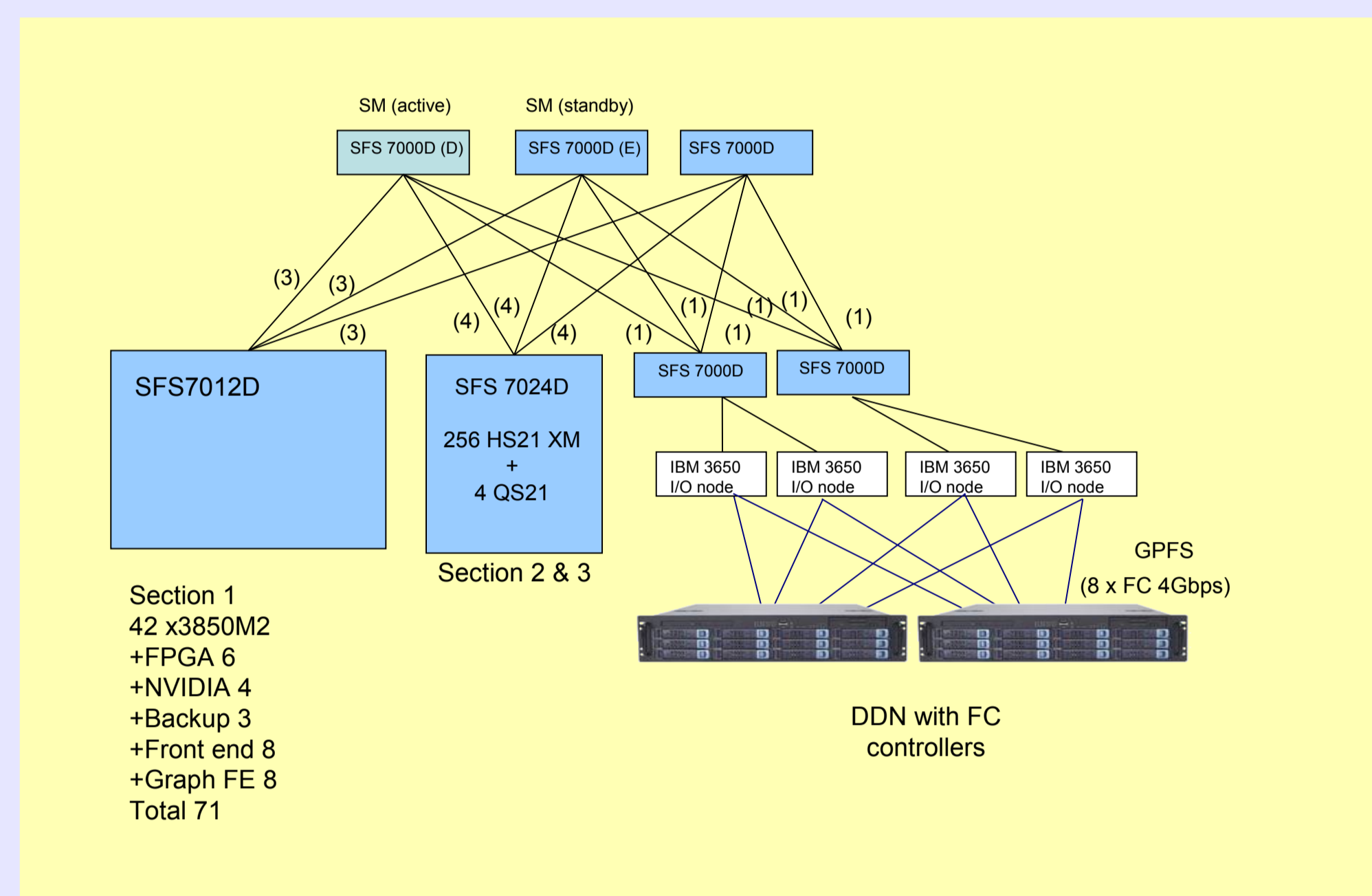
In order to provide the best environment for different types of applications the system consists of two main sections respectively oriented (1) for **high memory request** and moderate parallel scalability and (2) for limited memory and **high scalability** cases. Both sections are interconnected by a common Infiniband 4X DDR network (IB) and can operate as a single large integrated system.

The first main section is composed by 42 fat nodes IBM x3850-M2 with 4 Xeon Quad-Core Tigerton E7330 processors (2.4GHz/1066MHz/6MB L2), 32 GB RAM (12 extra-fat nodes with 64 GB RAM). The total number of cores in the first section is then equal to 672.

The second main section is composed by 256 blades IBM HS21 each supporting dual Xeon Quad-Core Clovertown E5345 processors (2.33GHz/1333MHz/8MB L2), 16 GB RAM for total of 2048 cores. The blades are hosted by the 14 slots blades chassis for a total of 19 chassis and each blade has a dedicated IB connection. The larger system created by joining the two main sections has 2720 cores.

A third **experimental section** consists of 3 subsections dedicated to special processor architectures:

- 4 blades IBM QS21 with 2 Cell BE Processors 3.2 Ghz each.
- 6 nodes IBM x3755, 4 sockets AMD Dualcore 8222 equipped with a FPGA VIRTEX5 LX330 card
- 4 node IBM x 3755, 4 sockets AMD Dualcore 8222 with a NVIDIA Quadro FX 4500X2 [3] and 4700X2 [1] video card; the latter support CUDA while all of them support the GP-GPU development environment.



Architecture of the InfiniBand network including the IBM/DDN 9550 storage system.

The 4 I/O Nodes, directly FC attached to the storage, are the GPFS NSD servers.

The IB network is based on a CISCO SFS 7024 (288 ports), a CISCO SFS 7012 (144 ports) and 5 CISCO SFS 7000 (120 ports).

The Ethernet network consists of one CISCO 4506 (240 ports), 3 CISCO 4948 (144 ports) and 3 CISCO 3750G (144 ports).

The storage of CRESCO HPC system is provided by an IBM DCS9550 system, 160 TB raw space based on 500 GB SATA Hard Disk. An IBM Tape Library IBM TS3500 provides the backup facility.

The power required to run the system has been estimated to 150 kw and proper cooling systems have been provided.

The operating system is RedHat EL 5.1 and the usual set of Portland and Intel Compilers are available.

A GPFS parallel file system is shared via Infiniband between all the computing nodes of all the main section of the system. User homes are located in an OpenAFS file system, one of the base elements of the ENEA-GRID infrastructure.

The three sections together with other 35 service machines (front-end, controls, file-servers, installation servers) and storage and network components make use of a total of 18 standard racks (19", 42 U).

ENEA-GRID computational resources

• **Hardware:**

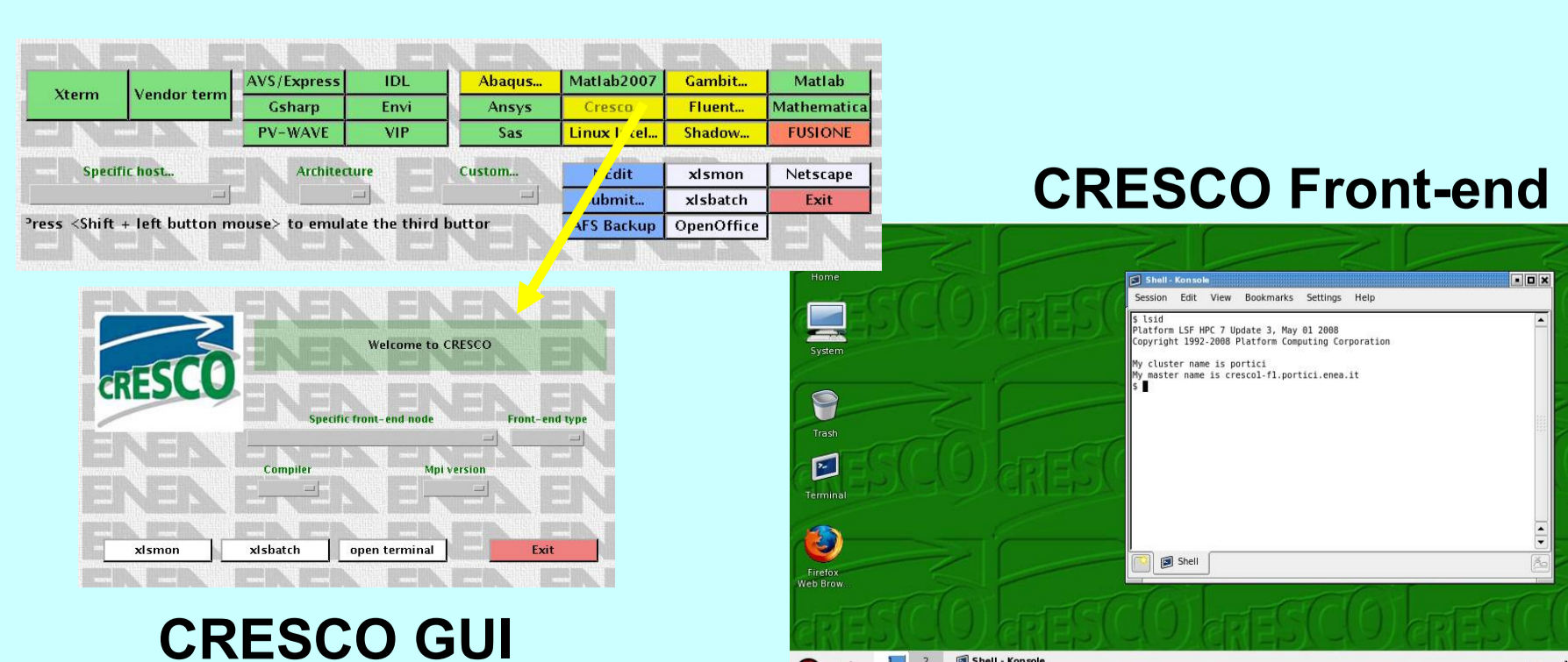
- CRESCO ~300 hosts & 3000 cores mainly x86_64
- Elsewhere ~110 hosts and ~730 cores : IBM SP; SGI Altix & Onyx; Linux clusters 32/ia64/x86_64; Apple cluster; Windows servers. Most relevant AIX resources: IBM SP5 258 cpu; 3 frames of IBM SP4 96 cpu

• **Software:**

- Commercial codes (fluent, ansys, abaqus...)
- Research codes. (mcpn/x, eranos, fluka...)
- Elaboration environments (Matlab, IDL, Scilab...)
- Windows Applications

ENEAGRID WEB GUI

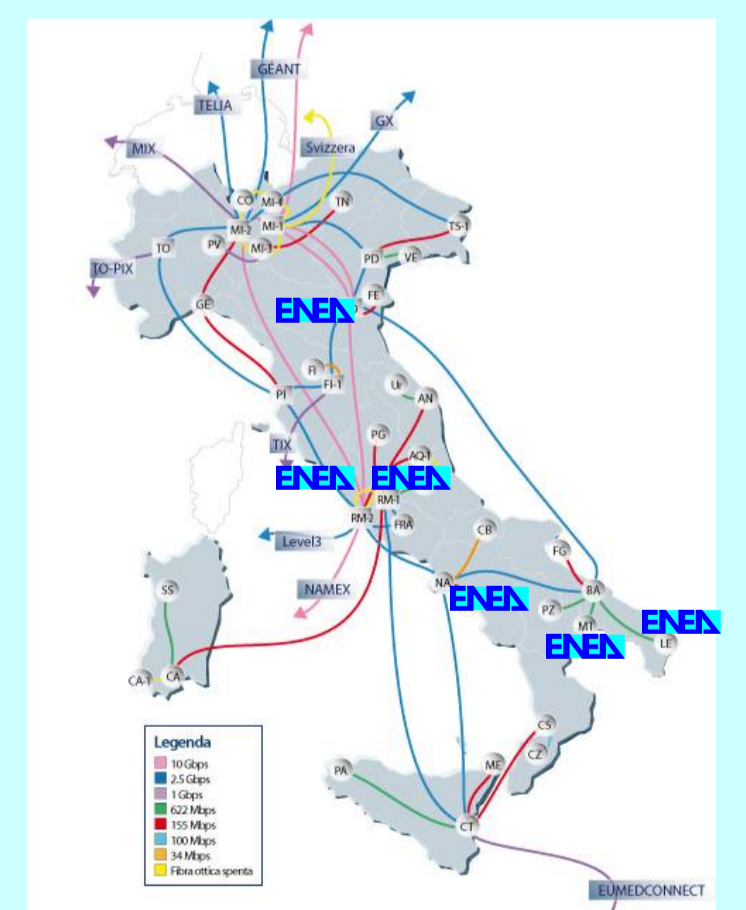
ENEA GRID unified GUI interface has been implemented using Citrix Technologies and recently by **FreeNX**, used also for access to CRESCO front-end desktop, via Java Applet



ENEA Network connection

ENEA computational resources are distributed over WAN, connected by GARR, the Italian Academic & Research Network

- ENEA-GARR**
- 9 PoP, 18-2000 Mb/s**
- Brindisi 150 Mb/s
- Bologna 30 Mb/s
- Casaccia 200 Mb/s
- Frascati 155 Mb/s
- Portici 2000 Mb/s
- Trisaia 18 Mb/s
- Palermo
- Pisa
- Roma Sede



CRESCO Project Computer Room

