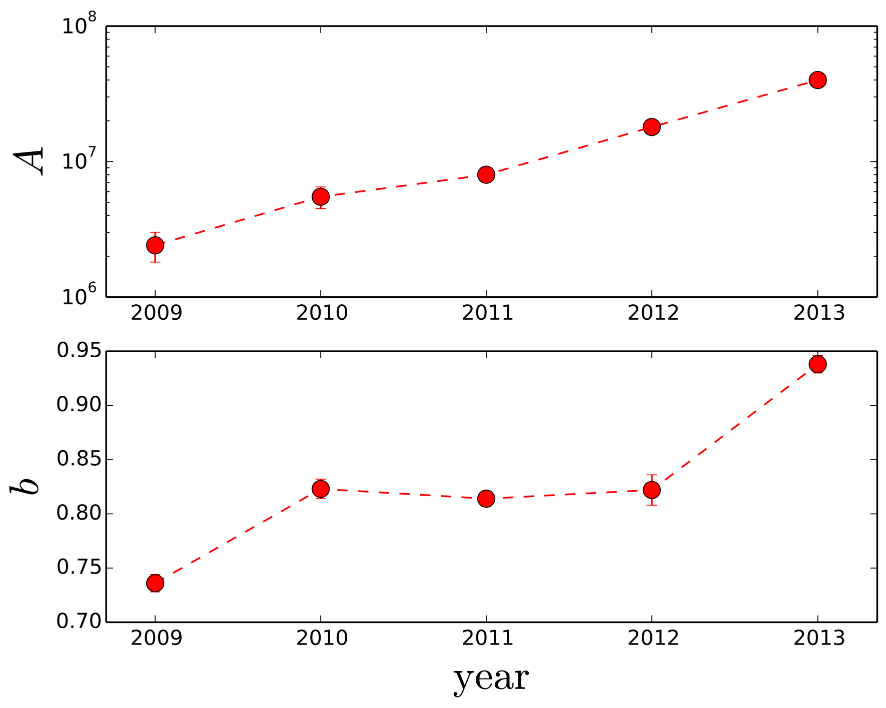
The role of medium size facilities in the HPC ecosystem: the case of the new CRESCO4 cluster integrated in the ENEAGRID infrastructure

Role of medium size HPC Clusters

Since a long time it has been established that performance data populating the TOP500 list tend to follow Zipf distributions, e.g. $R_{\max} = A/k^b$ with k denoting the rank in the list and A and b representing time depending constants. In particular, it is known that A tends to increase exponentially in time as a reflection of the Moore's law. More subtle is the trend of the exponent b: while displaying a slight decrease over the years 1999–2007, it shows a slow growth since 2009 up to date, thus signaling an increase in time of the unbalance of computing power between higher and lower ranks in the list. More precisely, Fig. 1 shows the trend of the power–law parameters of $R_{\max}$ as obtained from numerical fits to the yearly data, with interpolations neglecting the lowest 100 items in the lists, as they show a slightly steeper decrease than the rest.

On the technical side, high scalable applications must be tailored to the specific hardware solutions adopted by the leading edge machines, such as special communication networks and/or specific co–processor architectures. Moreover, the access rules to large size machines usually require a heavy amount of bureaucracy which, though well justified by the cost and value of the resources made available, result in hampering fast exploration of new ideas and methodologies. As a consequence, medium size clusters and flexible access can represent an answer for scalability training and developing.

Medium size clusters are optimal for numerical problems which do not intrinsically require top class machines (partitioning large machines in a plethora of small subsections often results in a waste of precious resources). For this reason, it is not rare that scientific and technological communities prefer to invest R&D funds to purchase their own machine, even though this ends up in occupying the lowest ranks of the TOP500 list (if ever at all...). In this context the Italian HPC ecosystem is highly concentrated and rather different from other European countries, such as Germany, France and United Kingdom, as most of the Italian TOP500 resources are polarized in a limited number of large machines and institutions, totalling five TOP500 entries compared to the 20–30 of the other cited countries.



Hardware Architecture of CRESCO4

CRESCO4 is made of six racks. Five of them are occupied by computing nodes, whereas the last one gathers all network devices. The peak computing power of the machine amounts to 101 TFLOPs and the measured HPL efficiency is 85%.

- The system is composed by 38 Supermicro F617R3–FT chassis, each hosting 8 dual CPU nodes. Each CPU, specifically an Intel E5–2670, hosts in its turn 8 cores, for a total number of 4864 cores. These operate at a clock frequency of 2.6 GHz;
- The system is provided with a RAM memory of 4 GB per core;
- Computing nodes access a DDN S2A9900 storage system, for a total storage amount of 480 TB;
- Computing nodes are interconnected via an Infiniband 4xQDR QLogic/Intel 12800-180 switch (432 ports, 40Gbps).

Numerical projects on CRESCO4

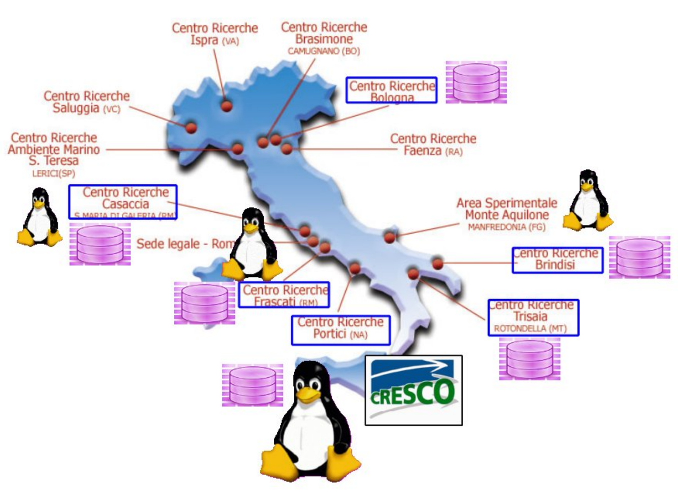
The numerical projects on CRESCO4 in 2014 cover the following research themes:

Material Science	Technologies for Energy & Industry
Storage of Hydrogen European flagship on graphene Materials for nuclear technologies Organic-inorganic adhesion	Thermo–acoustic instabilities in combustion Turbolence/combustion interaction Subgrid LES models Development of numerical algorithms Extension to supercritical fluxes
Energy & Environmental Modelling	**Nuclear Fusion**
Climate models for the Mediterranean sea Air pollution dispersion	Simulation of Alfven waves Physics and design technologies for nuclear fusion devices

Integration within ENEAGRID

As well as its predecessors CRESCO1/2/3, CRESCO4 is also integrated within ENEAGRID, a large infrastructure for cloud computing, which includes all the ENEA computing resources installed at the various ENEA research centres in Italy. ENEAGRID is characterized by solid structural components, offering reliability and easy management, and by web interfaces which have been developed and customized so as to ensure a friendly user environment:

- Kerberos 5 authentication
- File systems:
  - AFS/OpenAFS geographic file system
  - GPFS: parallel file system and WAN
- Resource Management: LSF Multi-cluster
- Web graphic user interfaces:
  - NX/FARO
  - Jobrama: a tool to monitor running jobs & administrate accounts
- System monitoring: Zabbix
- Web administration of users and projects: WARC



The computing resources currently provided to the user are x86–64 Linux systems (the CRESCO HPC Linux Clusters gather ~ 10.000 cores) and dedicated systems (e.g. GPU systems). The storage systems are both AFS and GPFS based. The computing resources spread across six ENEA research centres in Italy (as in the map shown above).

These projects are well suited for a medium size machine for various reasons. The exploratory nature of some of them takes advantage of the flexibility in the access to the computational resources while for others the limited scalability of the software could hardly benefit from running on a very large machine. In some important cases the limit in scalability is due to the number of available licenses of proprietary codes which, in some contexts such as Computational Fluid Dynamics, are an essential tool for technological development activities. In the latter case, the proprietary codes are often available for standard computing platforms only, as the market driving forces discourage software companies from developing code versions specific to the very specialized architectures of some of the top–class HPC systems. For all these reasons the choice of a medium size cluster such as CRESCO4 fits optimally with a large fraction of our community needs.

G. Ponti, F. Palombi, D. Abate, F. Ambrosino, G. Aprea, T. Bastianelli, F. Beone, R. Bertini, G. Bracco, M. Caporicci, B. Calosso, M. Chinnici, R. Ciavarella, A. Colavincenzo, A. Cucurullo, P. D'Angelo, M. De Rosa, P. De Michele, A. Funel, G. Furini, D. Giammattei, S. Giusepponi, R. Guadagni, G. Guarnieri, A. Italiano, S. Magagnino, A. Mariano, B. Mastroianni, G. Mencuccini, C. Mercuri, S. Migliori, P. Ornelli, S. Pecoraro, A. Perozziello, S. Pierattini, S. Podda, F. Poggi, A. Quintiliani, A. Rocchi, C. Sciò, F. Simoni, A. Vita

Technical Unit for the Development of IT and ICT Systems
ENEA — Italian Agency for New Technologies, Energy and Sustainable Economic Development
Lungotevere Thaon di Revel 76 – 00196 Rome, Italy
e–mail: giovanni.ponti@enea.it