Mem. S.A.It. Vol. 80, 504

Memorie della

The DEISA HPC Grid for astrophysical applications

C. Gheller

CINECA – Via Magnanelli 6/3 I-40033 Casalecchio di Reno (Bo), Italy
e-mail: `c.gheller@cineca.it`

Abstract. DEISA is a consortium of leading national supercomputing centres that currently deploys and operates a persistent, production quality, distributed supercomputing environment with continental scope. Scientific applications are supported by the DEISA Extreme Computing Initiative (DECI), which has the goal of enhancing the impact of the DEISA research infrastructure on leading European science and technology. This initiative consists of the identification, enabling, deploying and operating "flagship" applications which must deal with complex, demanding, innovative simulations that would not be possible without the DEISA infrastructure, and which would benefit from the exceptional resources of the Consortium.

1. Introduction

DEISA (Distributed European Infrastructure for Supercomputer Applications - `http://www.deisa.eu`) is a consortium of eleven major national supercomputing centres in Europe. The involved sites are FZJ-Julich, LRZ, HLRS and RZG Garching (Germany), ECMWF and EPCC (UK), IDRIS-CNRS (France), CSC (Finland), CINECA (Italy), BSC (Spain) and SARA (The Netherlands). The partners in this project consider their mission and their responsibility to provide visionary leadership in the area of high-performance computing in Europe. Therefore, the consortium decided to create and operate a distributed petascale supercomputing facility, whose integrated power has reached several hundreds of teraflops in 2008. The principal objective of this Integrated Infrastructure Initiative (I3 - `http://cordis.europa.eu/infrastructures/home.html`) is to advance computational science in leading scientific and industrial disciplines by deploying an innovative Grid-empowered infrastructure to enhance and reinforce High-performance Computing in Europe. This super-cluster with appropriate software and with dedicated high-speed networks, represents a way open, in the immediate future, for innovative and creative thinking that enhances the impact of existing infrastructures. It is assumed that this approach will provide answers to a number very demanding requirements arising from the scientific community. The DEISA infrastructure is based on the tight coupling of homogeneous (with respect to architecture and system software) national supercomputers, to provide a distributed supercomputing platform operating in multi-cluster mode. The production capability of the distributed platform provides a substantial European added-value to the existing national infrastructures.

Send offprint requests to: C. Gheller

Leading scientists across Europe are exploiting the bundled supercomputing power and the related global data management infrastructures in a coherent and comfortable way. A special focus is set on grand-challenge applications from key scientific areas like material sciences, climate research, astrophysics, life sciences and fusion-oriented energy research. Since 2005, the focus has been enhanced by defining the DEISA Extreme Computing initiative (DECI, `http://www.deisa.eu/science/deci`, see also section 3). The DEISA research infrastructure is also open, under certain conditions, to users of non-member organisations.

DEISA plans to collaborate with a large number of projects or institutions in Europe. The first priority has been co-operation with other FP6 infrastructure projects in HPC (HPC-Europa, `http://www.hpc-europa.org/`) or Grid computing (EGEE, `http://www.eu-egee.org/`). DEISA also deploys very close collaborations with other projects emerging in the R&D area.

The United States is also deploying a technology-driven, grid-empowered infrastructure, TeraGrid (`http://www.teragrid.org/`). In this project, five computing systems across the country are integrated to form a unique, single system platform using very highbandwidth, dedicated networks to guarantee highest performance. DEISA can be seen as the European counterpart to TeraGrid. Though implementation strategies are not identical, very close collaboration with TeraGrid has been initiated.

2. DEISA Architecture Overview

DEISA is technology neutral, and no technology commitments of any kind have been made. The systems integrated into the DEISA infrastructure have been chosen mainly because of site specific strategic issues. Technology choices follow from the system capability to adapt to a preestablished operational model, and to provide real services to end users. The DEISA supercomputing environment has been structured into two layers. The core project deploys an inner research infrastructure, which is the distributed European supercomputing facility. This part of the project is strongly dependent on the deployment of an internal, dedicated high-bandwidth network interconnect between computing platforms, and is strongly focused in HPC. It uses those Grid and multi-cluster technologies that are relevant to achieve large-scale, tight-system integration of a super-cluster at the continental level. In this core infrastructure, state-of-the-art Grid technologies are working transparently in the background and are not directly seen by the end users.

The extended project deploys an outer research infrastructure, whose purpose is to efficiently interface the European distributed supercomputing facility to other supercomputing infrastructures in Europe.

These Grid services are supported by the standard high-bandwidth connectivity for general scientific users provided by the National Research and Education Networks (NRENs) and GANT.

The core supercomputing platform provided a single system image of an initially homogeneous super-cluster, constituted of several IBM P690, P690+ and P655+ national systems, as well as global data management at a continental scale. This super-cluster was homogeneous in the sense that the basic processing units are IBM processors and the operating system is AIX, but the national configurations integrated in this super-cluster are different. They mixed 32 processor P690 and P690+ nodes, and 4-8 processor P655+ nodes. The Phase 1 initial core platform was composed of 3660 processors, each with a peak performance of at least 5 Gigaflops (FZJ, 1312 processors; IDRIS-CNRS, 1024 processors; RZG, 812 processors; CINECA, 512 processors). The second generation Phase 2 core platform benefits from several improvements. Along with additional IBM systems a heterogeneous extension has been started by the integration of huge Linux systems, like those BSC, LRZ and SARA, and vector machines, like that of HLRS.

The DEISA core infrastructure relies on software that operates underneath standard Grid middleware like Load Leveler Multicluster, Unicore and/or the Globus

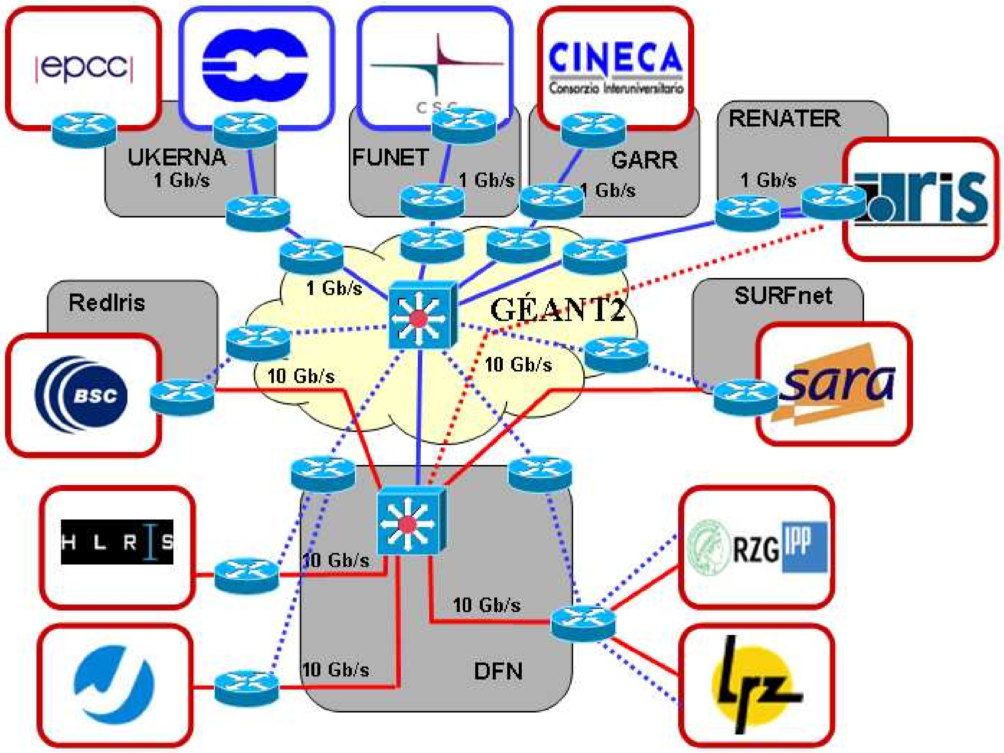

Fig. 1. The DEISA GPC Grid with the continental filesystem.

Toolkit. The distributed file system used provides global data management and high-performance distributed supercomputing at a continental scale. Global distributed file systems are sophisticated software environments necessary to provide a single system image in a clustered platform. They avoid data replication at each computing node, because they allow data to be shared by all nodes. The DEISA facility adopts an extension of IBMs GPFS (Global Parallel File System). A supercomputing cluster operating in a DEISA site is composed of a huge amount of autonomous computing nodes each with their own operating system, a shared large disk space and an internal, often proprietary, network connecting all these elements.

To deploy a European supercomputing cluster, DEISA just extends this idea to a continental scale by adding one extra layer of dedicated network connectivity across computing nodes residing in different national sites, and by deploying a global file system with continental scope. Users do not see cluster internals. They have a unified view of the whole environment. Data sets can be accessed on all computing nodes independent of their location and transparent to the user. Therefore, a user does not need to know on which nodes its application is executed and where the data resides.

3. The DECI program and the Planck Mission

The DEISA Extreme Computing Initiative has been launched in May 2005 by the DEISA Consortium, as a way to enhance its impact on science and technology. The main purpose of this initiative is to enable a number of grand challenge applications in all ar-

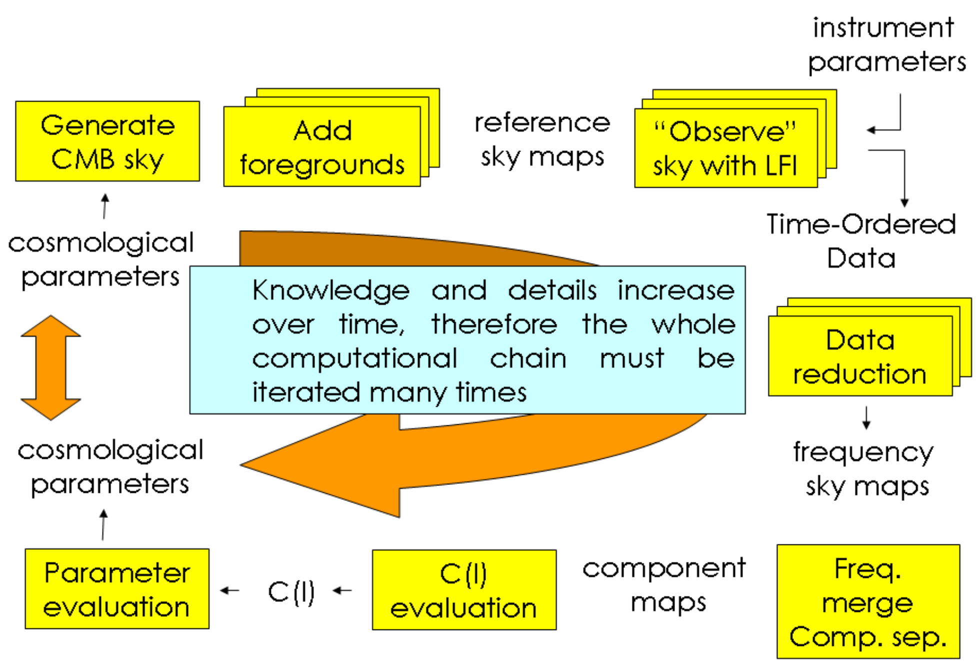

Fig. 2. The Planck software test pipeline.

eas of science and technology. These leading, ground breaking applications must deal with complex, demanding and innovative simulations that would not be possible without the DEISA infrastructure, and which benefit from the exceptional resources provided by the Consortium.

The DECI applications are expected to have requirements that cannot be fulfilled by the national services alone. During the time frame 2007-2008 45 DECI projects expressed by a number of different scientific areas, from Computational Fluid Dynamics to astrophysics, from Earth sciences to Computational Biophysics, run. The Planck mission (see `http://www.rssd.esa.int/Planck`), with the LFI-SIM project could exploit the DEISA infrastructure, being approved as a DECI project in 2006.

Planck is a mission of the European Space Agency whose ultimate goal is to determine the geometry and contents of the Universe, and to check the correctness of the theories on the birth and evolution of the Universe. To achieve this goal, it will measure with unprecedented accuracy and level of detail the fluctuations of the cosmic microwave background radiation (CMB) that contain fundamental information on the properties of the primordial universe. ESA manages the project, develops and procures the spacecraft, integrates the instruments into the spacecraft, and will launch and operate it. Planck is currently planned to be launched on an Ariane 5 rocket during 2008, together with the Herschel Space Observatory (see `http://www.rssd.esa.int/Herschel`).

Data detected by Planck, will be handled by a complex software pipepline (see figure 2) which encompasses all the steps needed to go from the raw measures to the reconstruction of the CMB map and, finally, to the estimate of the spectrum and of the cosmological parameters. All these steps must be checked iteratively to verify the correctness of the algorithms and optimized in order the get the most accurate results with the maximum computational efficiency. This progressive refinement process requires to simulate the whole mission many times.

The various pipline steps require large computational and storage resources, which can be provided by Grid infrastructures. And, in fact, Planck could exploit the two major computational grid environments, EGEE and DEISA, representing a unique example of interopearability between the two different large scale infrastractures.

In particular, the DEISA infrastructure has been used for those steps of the analysis pipeline which require huge computational resources. This means not just powerful CPUs and large memories, but also massive and reliable storage and fast networks. The mapmaking operation and TODs generations required high performance dedicated systems as those provided by DEISA.

4. Conclusions

DEISA intends to contribute to the significant enhancement of capabilities and capacities of high-performance computing in Europe by the integration of leading national supercomputing infrastructures. To realise this mission, it deploys and operates the distributed multi-terascale European computing platform based on the strong coupling of existing national supercomputers and operates as a virtual European supercomputing centre. The main target of the DEISA project is the provision of effective HPC services to science and research. The DECI program was conceived to identify, deploy and operate a number of flagship applications in selected areas of science and technology, chosen on the basis of scientific excellence, innovation potential and relevance criteria. The LFI-SIM Planck project, represents an example of successful DECI project and it is also an example how a research project can use the resources provided by two different grid infrastructures like EGEE and DEISA in an effective way, exploiting the complementary characteristics of the two environments.