Mem. S.A.It. Suppl. Vol. 13, 117

Memorie della
Supplementi

(Super)computing within integrated e-Infrastructures

F. Pasian

INAF - Information Systems Unit and INAF - O.A.Trieste, Via G.B.Tiepolo 11, I-34143 Trieste, Italy, e-mail: pasian@oats.inaf.it

Abstract. Also in the case of astrophysics, the capability of performing "Big Science" requires the availability of advanced computing: large HPC facilities and a Grid infrastructure. But computational resources alone are far from being enough for the community. A whole set of e-infrastructures (network, computing nodes, data repositories, applications) need to interoperate smoothly.
In this paper, after a survey of the situation in Europe and Italy in the fields of Grid and HPC, the data processing and simulations of the Planck mission are described as a practical example of the complexity of astrophysics applications. The need is evidenced for a complex e-Infrastructure for astrophysics, combining applications, data, computing, network within an integrated view, in which the Virtual Observatory is expected to bring in its own standards, tools and facilities.

Key words. HPC – grid computing – data grid – e-infrastructure – applications – Virtual Observatory

1. Status of advanced computing in Europe and in Italy

1.1. The Grid

The Grid is the field of advanced computing where the most massive progress has been made, and where Europe has displayed world leadership, providing through the EU-funded EGEE (Enabling Grids for E-sciencE) projects an infrastructure for research. Building upon this experience, the European Grid Initiative (EGI) Design Study was initiated, and represents an effort to establish a sustainable grid infrastructure in Europe. Driven by the needs and requirements of the research community, it is expected to enable the next leap in research infrastructures, thereby supporting collaborative scientific discoveries in the European Research Area (ERA).

The main foundations of EGI are the National Grid Initiatives (NGI), which operate the grid infrastructures in each country. EGI will link existing NGIs and will actively support the setup and initiation of new NGIs. The EGI Design Study is a project funded by the European Commission's 7th Framework Programme.

The Italian NGI, i.e. the national coordinator for the different pieces of the Italian e-Infrastructure present in EGEE and EGI will be IGI (the Association for the Italian Grid Infrastructure), which has been established to

Send offprint requests to: F. Pasian
Correspondence to: INAF-OATS, via G.B.Tiepolo 11, I-34143 - Trieste - Italy

provide long term sustainability for the Italian Grid. IGI will take over its formal predecessor, the Grid.it project, to extend its benefits.

The IGI partners are: the National Institute for Nuclear Physics (INFN), the National Research Council (CNR), the ELETTRA Consortium, the National Institute for Alternative Energies (ENEA), National Institute for Astrophysics (INAF), the National Institute for Geophysics and Vulcanology (INGV), the University of Calabria (UNICAL), the University of Naples 'Federico II' (UNINA), the University of Perugia (UNIPG), the University of Piemonte (UNIPNM), the Sicilian COMETA Consortium, the Sardinian COSMOLAB Consortium, the Consortium for the Italian Research Network (GARR) and the SPACI Consortium.

IGI focuses on setting up and operating a common e-Infrastructure for the Italian Sciences, and includes the main public Resource Providers and Computing Centers, in addition to various Regional Initiatives and other related projects. IGI will provide a consistent and coordinated Italian strategy as a step towards the European Grid Initiative (EGI) and as an interface to the EU Grid infrastructure projects e-IRG (e-Infrastructure Reflection Group) and ESFRI (European Strategy Forum on Research Infrastructures), and to other international activities as the need arises.

In spring 2008, MIUR has asked the research institutions to provide their roadmap for Italian research infrastructures. All institutions associated within IGI have jointly submitted a proposal for the Italian Grid Infrastructure. The proposal has been built so as to be coordinated with the section of the ESFRI roadmap dealing with e-infrastructures. At this stage, it is expected that appropriate funding will be provided by MIUR to support the IGI infrastructure.

Information on the role of INAF in Grid usage and developments can be found in Benacchio and Pasian eds. (2008) and in Vuerli et al. (2008).

1.2. HPC

Complementary to EGEE's characteristics, the EU-funded DEISA projects have enhanced Europe's capability of supporting "Big Science" through the distributed use of HPC resources. In the DEISA paradigm, a computationally intensive task is assigned to one of the HPC nodes, and the relevant data are stored in a distributed storage system. DECI (the DEISA Extreme Computing Initiative) was created to demonstrate through scientific cases the soundness of the concept.

As the next step, PRACE (PartneRship for Advanced Computing in Europe) will create a pan-European HPC service comprising all the major supercomputing centres. This infrastructure will be managed as a single European entity. PRACE prepares the creation of a persistent pan-European HPC service, consisting in one or more "tier-0" centres providing European researchers with access to capability Petaflop-scale facilities and forming the top level of the European HPC ecosystem. Such an ecosystem is expected to include "tier-1" (national) facilities in the 100 Gflops range, and "tier-2" (regional) centres in the 10 Gflops range.

PRACE is a project funded in part by the European Commission's 7th Framework Programme. It is the natural extension of DEISA, and is aimed at sustainability. CINECA is the Italian representative in PRACE. More information on European HPC activities can be found in (Bassini (2008)).

Another proposal called HPC-ISI has been submitted for the MIUR roadmap, designed to be complementary with IGI, and dealing with the need of an appropriate national HPC facility. Partners in this endeavour are: CINECA, the condensed matter Institute of the National Research Council (CNR-INFM), INAF, the National Institute for Oceanography, Geophysics and Seismology (OGS), CNISM, INSTM, the International Centre for Theoretical Physics (ICTP), the International School for Advanced Studies (SISSA), the CMCC, the Protezione Civile, the Istituto Ortopedico Rizzoli. The purpose of the proposal is the enhancement of an existing facility so as to reach capabilities having pan-European relevance, as referenced in the ESFRI 2008 Roadmap and as specified within the PRACE framework. The goal is to achieve the possibility of creating a Tier-0 centre, while the absolute need is to implement a Tier-1 facility having a computing power in the order of hundreds of Gflops. The interoperability with the Tier-2 centres implemented by the PON Consortia is also an important achievement to be pursued.

For the time being, while waiting for these developments to become reality, the astrophysics community may take advantage of the new INAF-CINECA memorandum of understanding. The title of the agreement is “HPC for Astrophysics on new-generation systems; archiving and exploitation of results” and its duration is 3 years, from January 2008 to December 2010. The processing power allocated to INAF is 200 equivalent-CPUs on the CINECA BCX system, corresponding to 1752000 CPU hours. The agreement contains the definition of Service Level Agreement (SLA) policies. In order to follow-up the evolution of HPC systems in terms of service, R&D, benchmarks, etc., CINECA and INAF co-finance a support position for INAF activities at CINECA. As in past, the agreement is controlled by two committees: the Management Committee is composed of two CINECA and two INAF members and, besides defining the SLA, is expected to solve issues in the provision of services; the Scientific Committee acts as computing time allocation Committee (TAC) and identifies new scientific and/or technical topics for collaborative work.

2. Planck processing and simulations as a practical example

In the following, the data processing and simulations for the Planck mission are described, as a practical example of the complexity of astrophysics applications. Massive computing is an important aspect, but is only part of the overall Planck problem.

2.1. The magnitude of the problem

Planck is an ESA mission (with NASA contribution) aimed at exploring the Cosmic Microwave Background (CMB) and is due to be launched in fall 2008. It represents the new generation of CMB experiments, significantly improving the results of the NASA WMAP mission. It will perform an all-sky survey at 9 microwave frequencies from 30 to 857 GHz (most of them include polarisation information) for a period of more than 18 months. The size of the mission is in the order of $O(10^{12})$ observations, $O(10^8)$ sky pixels, $O(10^4)$ spectral multipoles. For details on the Planck mission, see e.g. Clavel and Tauber (2005).

The Planck data set consists of several items: first, the raw data, consisting in detector readouts, pointing information and housekeeping data, all amounting to roughly 2-5 Tbytes of data. The derived products (including mission deliverables) are: systematic templates, noise filters, ..., sky maps at all Planck frequencies, foregrounds maps, power spectra and cosmological parameters, plus additional auxiliary products.

All of this amounts to a total of about 200-500 Tbytes of data stored. All of these products need to be simulated to build and test the various levels of the data processing pipeline.

For Planck, the number of measured samples $N_t \sim 5.10^{11}$ (corresponding to 72 detectors, 12 months); the number of pixels $N_p \sim 6.10^8$ (corresponding to the intensity and polarisation maps (I, q, u); the number of spectra is 6; the maximum multipole is $\sim 3.10^3$; the number of multipoles is $N_l = 6.3.10^3$ (corresponding to the TT, EE, BB, TE, TB, EB spectra).

The computation on these data is pretty complex. The approximate analysis kernel algorithms are FFT in the time domain ($N_t lnN$ operations required), SHT in the pixel domain ($N_p^{3/2}t$ operations), implicit multiplication and/or inversion of a rank $O(10^8)$ matrix, and various other methods such as sparse matrix algebra, Monte Carlo simulations, PGC solvers, etc.

Planck data analysis will require $O(10^{19})$ operations corresponding to 10 million TFlop i.e. 10 ExaFlop or more per year (!), thus marking a transition in the CMB data analysis from serial to parallel on basically all levels of data processing.

For all of this data processing system to be ready before launch, the whole mission has been simulated hundreds of times at different levels of accuracy. In this case the whole simulations loop is basically the following: cosmological parameters are given as an input to the whole process; an ideal CMB sky is created; foregrounds are added generating reference sky maps; the sky is then 'observed' with instrument parameters as an input; Time-Ordered Data (TOD) are thus created mimicking the output from the mission. It is only at this stage that the data are processed as if it was the Ground Segment's pipeline: first data reduction, generating sky maps in all frequencies; then the merging of frequencies and the separation of astrophysical components which create component maps; then the power spectrum coefficients C_l are computed, and finally the cosmological parameters are evaluated. The comparison between input and output cosmological parameters is the figure of merit of the whole simulation+processing+analysis process.

Of course, many different sets of cosmological and instrumental parameters have been given as an input to the simulation process, including e.g. undesired instrumental effects or systematics to be removed during the processing. Instrumental knowledge and detail of simulations have increased over time, therefore the whole computational chain has been iterated many times. Even after launch, simulations will continue, to try understanding unexpected effects one may find in the real data.

2.2. Porting to EGEE and DEISA

To be able to perform meaningful simulations, HPC centres have been used, both in Italy and the US. However, also the use of the Grid and the EU HPC network have been tested.

On the Grid, first on Grid.it then on the EGEE infrastructure, the user defines the simulation which needs to be performed: the cosmological and instrumental parameters are sent to Grid nodes; the ideal sky is generated and 'observed', generating the TOD. Some analysis and simplified mapmaking is then performed, on the data corresponding to 1 instrument channel per Node and the resulting frequency maps (ideal, observed skies, mapmaking residuals) are sent back to the user (Taffoni et al. (2005)).

Similarly a Planck simulation environment has been set up on DEISA to produce large quantities of simulated time ordered data. The objective was to test and improve the core algorithms (optimal mapmaking, power spectrum estimation) using realistic volumes of data (full-sky, full-resolution). A second aim was to study the impact of certain instrument features, in particular the systematic effects due to the main beam asymmetries. Comparing ideal (symmetrical) beams with realistic ones requires a 4π full sky convolution with each beam.

A discussion of these experiences can be found in Gheller et al. (2007).

2.3. Lessons learned

Planck processing is only very partially serial (the checking of telemetry to verify inflight the health of instruments); most of the rest is heavily parallel. Simulations on the Grid (EGEE) have proved very useful to check (pretty fast) the impact of a specific systematic effect on the scientific result of the mission. HPC (DEISA) allowed to produce relatively quickly massive sets of simulated data and to perform and test data processing steps having a memory footprint requiring a large centralised RAM.

However, it appeared pretty obvious that computing power alone is not enough to cope with the big data reduction, processing and analysis challenge Planck is facing. In other words, there is the absolute need to:

- transparently access the data structure (for efficiency);
- work in a documented data processing environment (for integrity)
- give external access to data (for distribution of results).

A dedicated system is therefore required during the mission to support secure data handling and processing operations. That is now available at INAF OATS for the Data Processing Centre (DPC) of the LFI instrument, and integrates a certain number of subsystems. To cope with the need for computing power the CINECA CLX system has been lent to INAF; the data processing applications are handled and integrated in the pipeline by a workflow management system, the Process Coordinator (ProC), tightly integrated with a Planck-dedicated Data Management subsystem (DMC) which in turn accesses a storage system managing several dozens of TeraBytes. In June a network connection (initially at 100 Mbps, then at 1 Gbps) will allow external users to access the public section of the storage area.

3. An integrated e-Infrastructure

The above-mentioned experience has evidenced the need of integration among the various infrastructural components of the Planck DPC(s).

3.1. Concepts and requirements

Some general concepts can be derived. All applications (dealing with theory, numerical simulations, data reduction, data analysis, ...) have different requirements, but in principle share the same infrastructure.

To offer scientists a useful service, all of these components need to be thought as integrated, or at least fully interoperable. In other words, the various infrastructure components (applications, computing, data) should interact seamlessly exchanging information, and be based on a strong underlying network component. This is depicted in Figure 1.

Considering first the computing infrastructure, the facilities available to the scientists can be very different: from the laptop or desktop PC to the HPC centre, passing through local clusters and “the Grid”. From the user perspective, ideally, all facilities should be seen homogeneously; in reality, they all tend to have different access modes. As a result, users find obstacles to their ideal exploitation.

As for the data infrastructure, again from the user perspective there is the need to transparently and homogeneously access a wide variety of data (multi-frequency and multi-instrument observations and numerical simulations). The World-wide Astronomical Virtual Observatory is progressively fulfilling this requirement providing seamless access to data centres and facilities through its standards, but this is only the beginning of the story. User queries could be expressed in natural language, or a user query may imply, besides a data infrastructure, the implicit use of applications and computing resources. This, again, requires tight integration of the various infrastructure components.

3.2. The path forward

The first thing to do is to enhance the basic infrastructure: network and computing resources. The GARR Network needs to be improved by increasing the bandwidth and eliminating the bottlenecks; in particular, last mile connections to the individual local sites need to be upgraded. At the same, resources must be found to allow Italy to harmonise with the European plan for HPC by allowing to build at very minimum a national Tier-1 facility, aiming at a PetaFlop Tier-0 system.

Not in contrast with this approach, the Grid concept should become far more widespread than currently. The old perception of the Grid re-using the idle cycles of a set of PC’s (as was in the SETI@home experiment that originated the Grid concept) is nowadays completely misleading – from the computing and storage viewpoints, the Grid has now the possibility of linking powerful clusters and even HPC hardware, provided it uses the proper middleware.

The long-term goal is to allow complete interoperation of HPC centres: initially, Tier-2 (regional) facilities shall be integrated with the Grid, followed by full integration of local clusters (‘Tier 3’) and, as final goal, achieving inclusion of Tier-0/1 within a common scheme. In any case, a necessary step is to achieve as soon as possible the full compatibility between

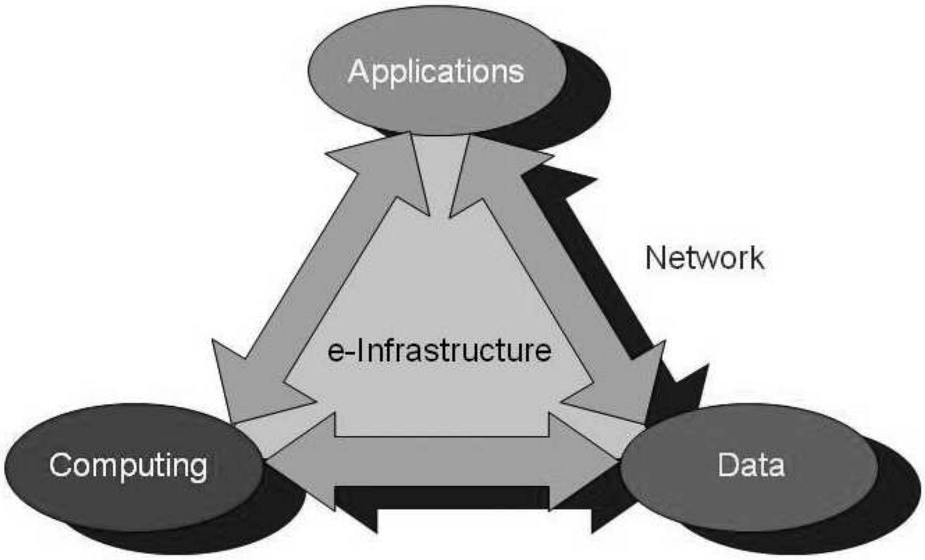

Fig. 1. The components of the e-infrastructure (applications, computing, data) interact seamlessly exchanging information, and are based on a strong underlying network component.

HPC and Grid at least at the User Interface level.

There is a further point. Up to now, the Grid has mainly delivered computing power, the main issue that its implementers, mostly involved in large High-Energy Physics experiments (e.g. at CERN), needed to solve. Accessing data, and databases using the Grid paradigm is a step still to be improved. Some activities are being carried out, also with the participation of INAF, within EGEE (Taffoni et al. (2006)). This is of course a step in the direction of interoperability.

The requirement this interoperability needs to fulfill is shown in Figure 2: applications, computing power, data repositories and databases holding metadata or catalogues should be accessed as a single utility. We may call it Grid, or Cloud (as is now fashionable), or Virtual Observatory (which may be more familiar to our community), ... It does not really matter, provided the goal is the same.

The far-end goal, in any case, is achieving the full interoperation of the relevant applications (data processing, analysis, mining, bibliography, semantic web) to start building a Knowledge Infrastructure for Astronomy.

4. A "political" conclusion

As a conclusion, the need for interoperable e-infrastructure(s) to improve scientific results implies a number of steps to be taken and the development and integration of tools and facilities: common (or at least compatible) user interfaces to computing resources; transparent access to observations and numerical simulations through the Virtual Observatory; integrated data processing pipelines; data mining; semantic web applications.

More in general, it is to be noted that **e-Infrastructures are absolutely necessary**

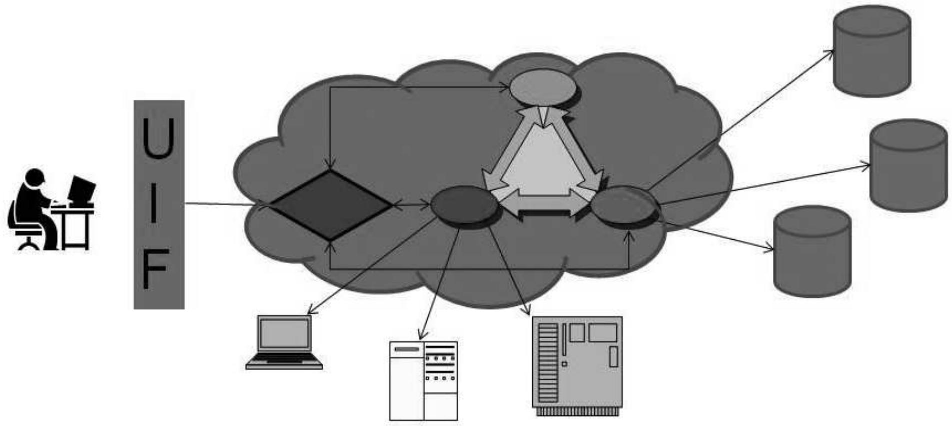

Fig. 2. The components of the e-infrastructure (applications, computing, data) accessed as a single utility. Access to data, applications and computing may be implicit.

for the development of science. But they **do not come for free, and cannot be given for granted**. This is well understood at the European level, and **e-Infrastructures are an integral part of EUs policies**.

ESFRI, which is a multi-disciplinary initiative for the implementation of research infrastructures, has focussed on the need for networking, capability and throughput computing, grid architectures, software, data management and curation as the main priorities.

ASTRONET is another initiative, dedicated to Astronomy and Astrophysics, aimed at defining the scientific priorities in our discipline for the next two decades, and the corresponding priorities in the allocation of resources. The ASTRONET working groups have recognised e-infrastructures as must-haves to tackle the challenges of the future: computing (both capacity and capability), theory and simulations and the virtual observatory, together with laboratories, must have priority over the rest, because without them science cannot be made.

And in Italy? Is this understood by everyone? Do we have a roadmap for the future?

Acknowledgements. The INAF-CINECA agreement is funded by INAF and managed by the Information Systems Unit of the INAF Projects Department. The support of its Director to INAF's participation in the IGI and HPC-ISI proposals to the MIUR roadmap is acknowledged.

The author wishes to thank Antonia Ghiselli and Mirco Mazzucato of INFN for information on Grid initiatives, Giovanni Erbacci and Claudio Gheller of CINECA for information on the status of HPC, and Francesca Perrotta of INAF-OATS for input on Planck data processing. Planck is the third M-sized mission of the ESA Horizon 2000+ programme. Instruments are built by international consortia (over 500 persons in 22 countries), funded by the national space agencies. Italian participation partially funded by ASI.

References

Bassini S. (2008), Perspectives in the development of HPC, these proceedings

Benacchio L., Pasian F. eds. (2007), Grid-enabled Astronomy, Polimetrica International Scientific Publisher (202 pages)

Clavel J., Tauber J.A. (2005) The Planck Mission, EAS Publications Series, Volume 15, pp.395-403

Gheller C., Leach S., Taffoni G., Vuerli G., Pasian F. (2007) Simulating the Planck

Mission on a Distributed Computing Architecture: the DEISA and EGEE experiences, in: Astronomical Data Analysis Software and Systems 16, R. Shaw, F. Hill & D. Bell, eds., ASP Conference Series,in press

Taffoni G., Castelli G., Smareglia R., Vuerli C., Zacchei A., Pasian F., Maino D., de Gasperis G., Ansari S.G., Tauber J., Ensslin T., Barbera R. (2005) Prototypes of a Computational Grid for the Planck Satellite, in: Astronomical Data Analysis Software and Systems 14, P.Shopbell, M.Britton & R.Ebert eds., ASP Conference Series, Vol. 347

Taffoni G., Vuerli C., Barisani A., Smareglia R., Volpato A., Pastore S., Baruffolo A., Pasian F., Benacchio L., Ambrosi E., Ghiselli A. (2006), Grid Data Source Engine: Grid gateway to the Virtual Observatory, in: Astronomical Data Analysis Software and Systems 15, C.Gabriel, C.Arviset, D.Ponz, & E.Solano eds., Publications of the Astronomical Society of the Pacific, Vol. 351

Vuerli C., Taffoni G., Sponza M., Pasian F. (2008), The Grid in INAF, these proceedings