Virtual Computing Environments for Problem Solving on Grids

Jean-Pierre Antikidis°, Guillaume Alléon+ and Toan Nguyen*

°Centre National d'Etudes Spatiales (CNES), Toulouse (France)
+ EADS (Common Research Center), Toulouse (France)
*Institut National de Recherche en Informatique et Automatique (INRIA)
655, Av. de l'Europe, Montbonnot, F-38334 Saint-Ismier Cedex (France)
E-mail: Toan.Nguyen@inrialpes.fr - Web page: http://www.inrialpes.fr/opale

Key Words: Parallel and distributed computing, Grid computing

ABSTRACT

The HEAVEN project aims to create a European advanced services platform offering a multi-purpose hosting of operational applications in research, science, business and community services.

Usually, grids provide sophisticated interfaces to distributed resources management as well as application execution and monitoring in wide and local area networks. These networks may connect thousands of computers by high-speed of up to 40 Gigabits/sec links. The computing resources include nodes made of thousands of processors, and terabytes of storage media.

In order to use them, the state-of-the-art in grid computing requires the deployment of complex software tools and environments which are usually not accessible to end-users, i.e. application field experts. They must call for the support and expertise of highly skilled software engineers to deploy, execute and maintain the computing environments.

The ultimate goal of the HEAVEN *upperware* is to make these computing environments transparent to the application users by providing them a high-level interface to:

- define their required computing environment, in terms of computing and storage resources
- specify the applications deployment on these resources
- execute and monitor the applications

Along the line of the strategic objective "Grid-based systems for complex problem solving", it is an R&D project in the field of "Enabling Application Technologies" based on grid infrastructures, although it is not a grid by itself. It is intended to enable the creation of "*virtual processing systems*" of any configuration by all users that are able to manipulate a suitable high-level description language. *Thus it will be possible to command a grid infrastructure to mimic any type of processing system.* This could become the basis for a future generalised service allowing the integration of all kinds of services *without the need to deploy any specific infrastructure*: it is a "virtual platform" concept. The approach relies therefore on a seamless integration and cross-fertilization of the technology-push and application-pull paradigms.

HEAVEN is the missing *upperware* layer between the middleware grid services and the users' applications (Figure 3). It will enable:

- user communities to implement unlimited and extendable sets of services at marginal cost
- exploit the huge amounts of existing data to provide the users with sophisticated information and services without having to manipulate the original data

- the use of grid infrastructures by a large variety of communities able to create their applications hosted by HEAVEN

The HEAVEN effort is therefore an attempt to gain rapid advantage out of emerging grid technologies. It will support a new approach for data providers and users and offer developers an unprecedented flexibility in the design of services at minimal cost.

Thus, HEAVEN is based on a vision capitalising on the outstanding features of grids to create a *general purpose enabling facility hosting virtual computing systems.*

Basically, the modus operandi for HEAVEN is: if it is possible to command a grid to mimic or simulate a computer system, it is possible to use this simulation as a surrogate of a real system and use it directly to run the applications. Practically, if one can draw a computer system including PC-clusters, wide-area networks, etc, on a monitor like in any CAD software system, it will be possible to create that system with the help of a suitable high-level software layer, here called *upperwar*e, in such a way that it behaves like the real system being designed.

The important feature here is that HEAVEN is the simulated system and the running system itself, altogether. The background concept in HEAVEN is therefore originating from ideas applied in advanced aircraft manufacturing and electronic circuits design in which, where, instead of creating an object from scratch, it is first designed as a virtual object on a computer. This approach is called virtual design and HEAVEN goes one step further by proposing not only to design computerized environments also by virtual modelling but to make use of such a virtual (and simulated) environment as the actual running machine. The direct impact of computerised environments is that *there is no difference in terms of application results between the real computer system and the simulated virtual one*. Further, the power of existing grid environments makes performance penalties in such simulated environments negligible. Therefore, the raw processing power not being a limiting factor, it becomes possible to completely virtualize the design and implementation of a computing environment in a single step by commanding a grid to mimic the desired system. The flexibility and processing power thus introduced in operational systems development is nearly unlimited. In a matter of a few clicks, one can imagine the capability provided to the users to expand their processing environments without actual new hardware and software deployment.

For example, in Figure 5, the application systems can then be seen by the users as standalone operational system for optimisation in aerodynamics. The system is transparently implemented on top of a grid including several PC-clusters remotely located that are connected by a high-speed wide-area network. Further, the boxes are iconic representation of several application components that are parallel implementations of CFD solvers, genetic algorithm optimisers and meshing algorithms. This is transparent to the end-users that can define the application solely in terms of software components linked by transitional links.

The advent of high-speed gigabits/sec networks and the simultaneous use of large PC-clusters, including hundreds or thousands of processors, have given rise to tremendous expectations on complex problems solving in the scientific community, including engineering, aeronautics, automotive, nuclear and environmental areas. The expected size of the problems being addressed is supposed to rise simultaneously by several orders of magnitude, and the complexity and the accuracy of the solutions found to these problems are expected to lead the way to unprecedented applications, e.g., aircraft in flight dynamics simulation, long range weather forecast and accurate climate modelling.

As problem complexity grows, the solutions rely on efficient numerical methods and flexible tools. The more complex are the problems to be solved, the more sophisticated are the solutions likely to be [6]. At the same time, the complexity of the computing infrastructures grows rapidly [5]. Hardware infrastructures become highly complex layers of components, both hardware and software. Bringing the powerful computing environments to the casual users becomes nowadays a challenging task. Indeed, engineers and researchers in application fields need sophisticated and flexible interfaces to benefit from the computing power provided by superscalar computers, PC-clusters and grid-computing environments. However, these environments seldom lend themselves to easy use by the practioners.

The goal of the HEAVEN project is thus to provide a powerful and yet simple to use framework by which the users:

- define their needs in terms of computing resources
- specify their requirements in terms of application software, data access and storage
- execute and control their application processes

To some extent, HEAVEN provides a software layer that manages resources in a grid computing environment in much the same way that today's operating systems on PCs manage the desktop computer resources transparently through icons and mouse interfaces – to the benefit of every user in a simple, affordable and easy way, even if he (she) doesn't know anything about computer technology (Figure 5). The target user communities are engineering and research project members that require powerful computing environments, e.g., simulation in the aerospace industry, climate modelling, nuclear simulation programs, etc.

The core of the HEAVEN framework is an *upperware* (Figure 1) which is built on top of existing middleware likely to be encountered in current grid-computing environments, e.g., GRASP, UNICORE, GLOBUS and EGEE. In fact, HEAVEN goes one step further: instead of implementing computerized environments by the direct interfacing with grids, the *upperware* is developed with two goals:

- simulate the computer system required by the users
- make use of this simulator as the actual running environment for their applications

By coupling these two features, we create true "virtual computer factories" enabling the users to build transparent systems that are hidden (and simulated) by the grids, which ultimately become themselves transparent. The potential outcomes of this approach are enormous, but it is only realistic now using the computing power provided by grid environments. By pushing further the frontiers of computing power, the advent of grids make it beneficial to trade raw computing power with system design flexibility.

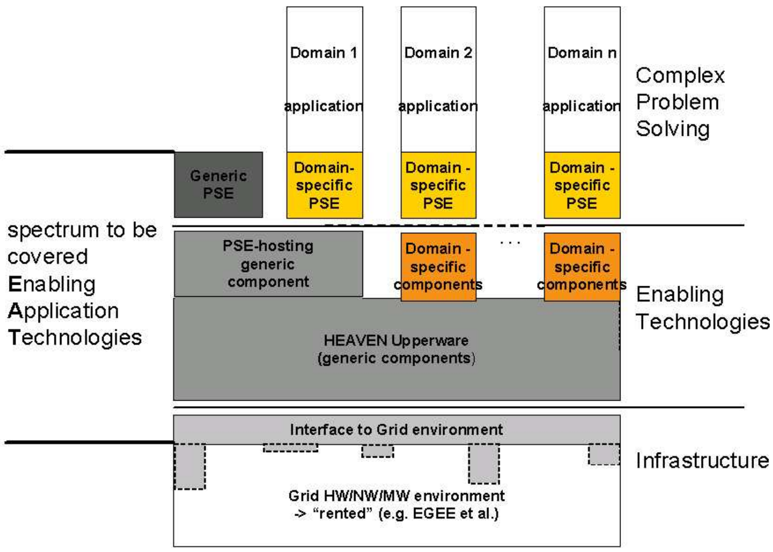

Figure 1. The HEAVEN architecture.

The architecture of the HEAVEN framework is described in Figure 1. It relies on a set of Problem Solving Environments (PSE) tuned and dedicated to specific application areas: aerodynamics, thermonuclear fission, climate modelling, etc.

These are implemented using domain-specific software components, e.g., 3D flow solvers in aerodynamics. These components lie on top of the generic components forming the *upperware*: this includes modules for the management of the user-specific environments, management of the applications and interface with the underlying grid computing environments (Figure 3).

The interface with the underlying grids is based on generic tools to interface resources management tools, user logging and accounting, application monitoring tools provided by the grid software infrastructures, e.g., GRAM for resource management in the GLOBUS toolkit [5].

The HEAVEN environment is proposed by a consortium including research, academic and industry partners from various countries in Europe. The implementation plan is described in Figure 2. The various tasks are self-explanatory and their logical links are given. The overall duration of the project is 4.5 years, which are divided in three successive phases. The first phase lasting 18 months is a specification and proof of concept development step. The second phase, lasting two years, is a software development and testcases validation step. The third phase, lasting one year, is a dissemination and in-depth test phase. The testbeds include a variety of prototypes and testcases submitted by academic, research and industry partners in various application fields. Potential customers are associated to the project from the beginning of the first phase. Indeed, the project leader is by itself a large Earth Observation agency. This will guarantee the expected usability and impact of the project.

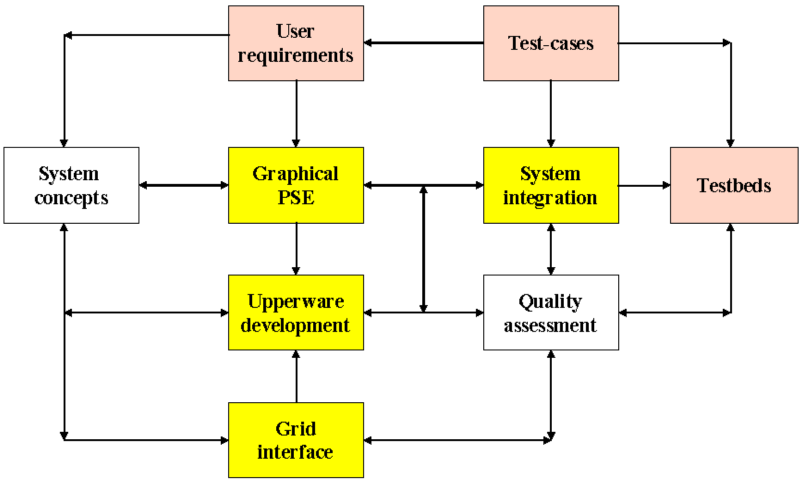

Figure 2. The HEAVEN implementation plan.

The validation testcases include optimisation in aerodynamics, fisheries resource management in the Mediterranean sea, particle physics data management and analysis, geophysical satellite image processing and storage for environmental issues, operational oceanography, vascular surgery, and noise pollution in urban areas.

The HEAVEN framework is the result of working groups including a variety of participants from various European countries, including research institutions as well as industry from Greece, Cyprus, France, Italy and Great-Britain. The authors wish to thank Stefano Beco, Iain Le Duc, Guillaume Alléon, Nektarios Chrysoulakis and George Papadopoulos for their contributions to the HEAVEN project.

REFERENCES

[1] T. Nguyen. Grids : opportunities and drawbacks. MACSInet Workshop "Modeling, Analysis of coupled PDE systems" and AFFRST "Innovative Methods in Advanced Technologies", Invited lecture, Sophia-Antipolis, October 2003.

[2] T. Nguyen. Virtual organisations and multidisciplinary applications. European Workshop on "Grid-based Virtual Organisations & collaborative e-Entreprise applications", European Commission., DG INFSO-F2, BAE Systems, London, May 2003.

[3] T. Nguyen. Grid-computing in mutidisciplinary CFD optimization problems. Parallel CFD 2003 Conference. Invited lecture, Moscow (Russia), May 2003.

[4] *Next Generation Grid(s): European Grid Research 2005-2010*. European Commission. Expert Group Report. June 2003.

[5] *IST Grid projects Inventory and Roadmap*. GRIDSTART Project. Document IR-D2.2.1.2-V1.0. June 2003.

[6] M. Andreoli, A. Janka, J.A. Desideri. Free-form deformation parameterization for multilevel 3D shape optimization in aerodynamics. *Research Report n° 5019. INRIA*. November 2003.

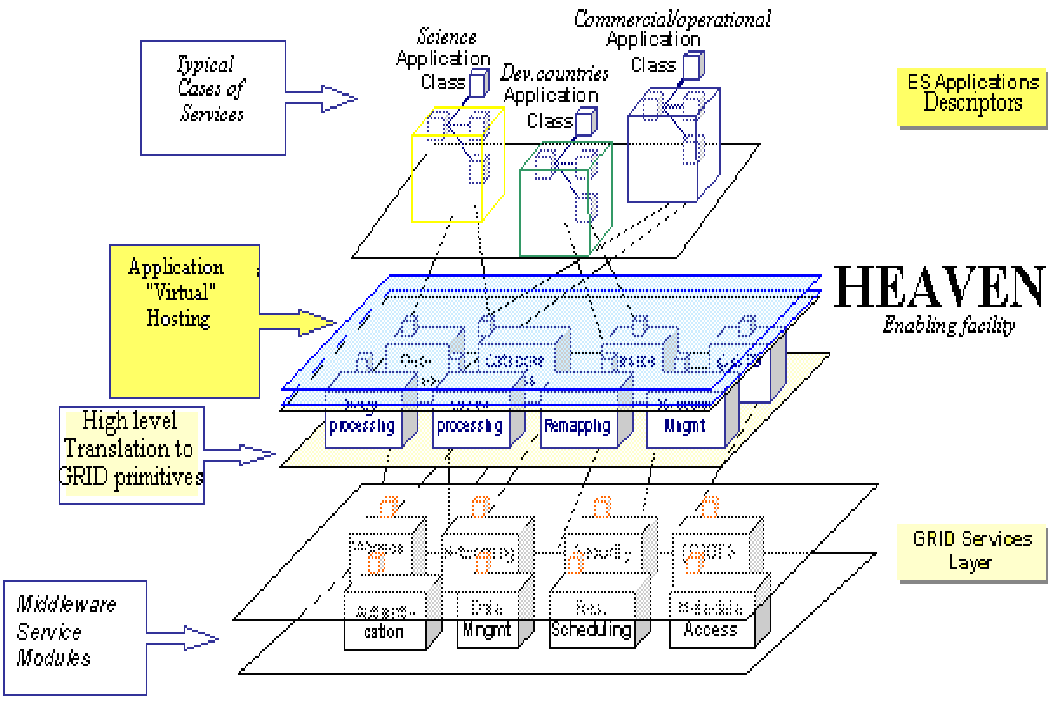

Figure 3. The HEAVEN deployment.

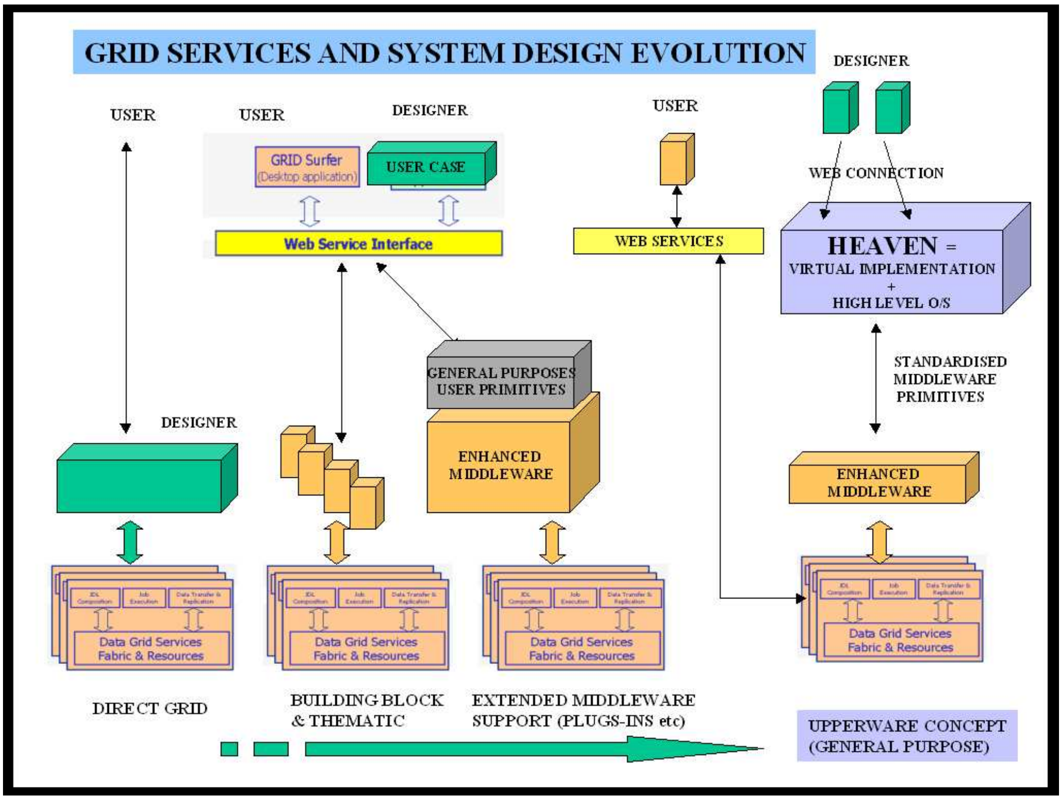

Figure 4. The HEAVEN design approach

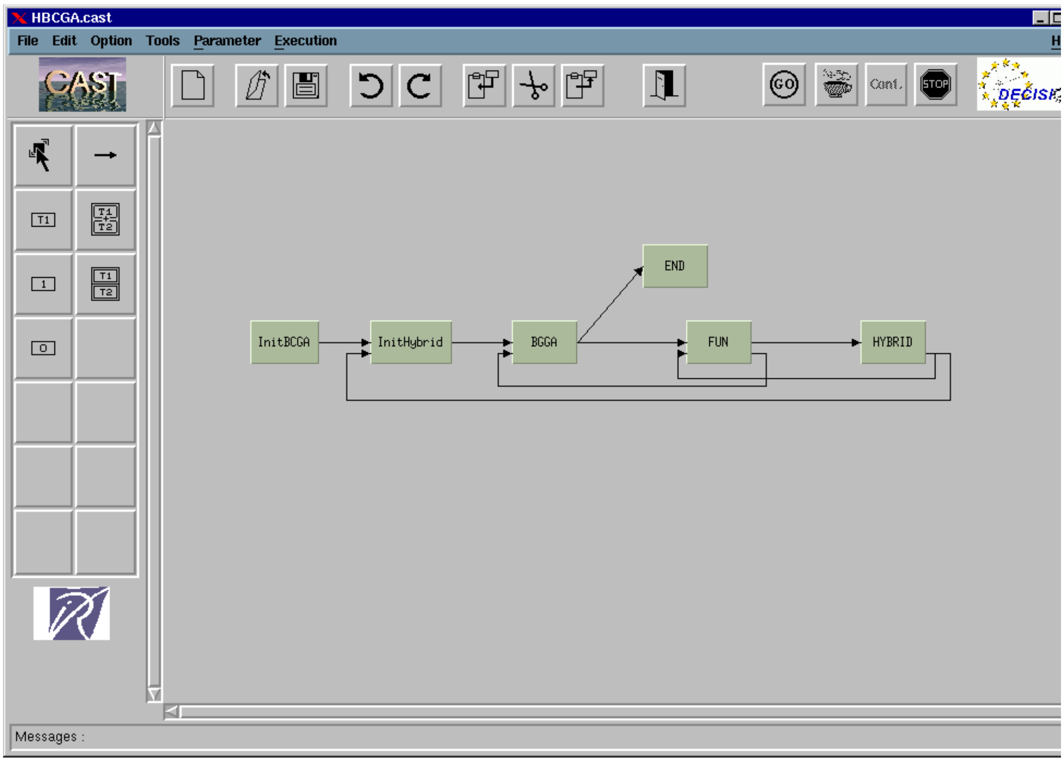

Figure 5. An example aerodynamics application.