

A sustainable High Performance Computing Ecosystem for Europe

P.J.C. Aerts, K. Koski, F. Lozano

Contents

A sustainable High Performance Computing Ecosystem for Europe.....	1
<u>Contents.....</u>	<u>2</u>
<u>Executive summary.....</u>	<u>3</u>
<u>1 Introduction.....</u>	<u>5</u>
<u>2 Role and scope of the HPC in Europe Taskforce (HET).....</u>	<u>6</u>
<u>3 An Ecosystem for High Performance Computing</u>	<u>7</u>
<u>3.1 Performance pyramid.....</u>	<u>7</u>
<u>3.2 The need for a European High Performance Computing policy.....</u>	<u>8</u>
<u>3.3 Human aspect has a key role.....</u>	<u>8</u>
<u>3.4 Why the top-of-the-pyramid resources are needed?.....</u>	<u>9</u>
<u>4 Supercomputing: Embedding in the general e-Infrastructure.....</u>	<u>10</u>
<u>4.1 Introduction.....</u>	<u>10</u>
<u>4.2 Supercomputing, Networking and the European Science Grid.....</u>	<u>10</u>
<u>4.2.1 Supercomputers must be part of the European Science Grid infrastructure.....</u>	<u>11</u>
<u>4.2.2 Grid developments support the integration of all e-resources and supercomputers are no exception.....</u>	<u>12</u>
<u>4.2.3 A supercomputer grid.....</u>	<u>12</u>
<u>4.2.4 One size does not fit all.....</u>	<u>12</u>
<u>4.3 Relation to other grid resources, data explosion.....</u>	<u>12</u>
<u>4.4 The pyramidal computing resources infrastructure</u>	
<u>-No top without a body-.....</u>	<u>13</u>
<u>5 The software environment.....</u>	<u>15</u>
<u>5.1 Introduction.....</u>	<u>15</u>
<u>5.2 Software development and maintenance.....</u>	<u>16</u>
<u>5.3 Selection of software development projects.....</u>	<u>16</u>
<u>5.4 Involvement of industry.....</u>	<u>16</u>
<u>6 Prioritisation and peer review, relation with the research councils.....</u>	<u>17</u>
<u>7 Organisational and operational matters.....</u>	<u>17</u>
<u>8 Funding</u>	<u>18</u>
<u>9 HPC Ecosystem and sustainability.....</u>	<u>18</u>
<u>10 Involving European industry.....</u>	<u>19</u>
<u>10.1 Promoting European HPC industry.....</u>	<u>19</u>
<u>10.2 Industrial usage of HPC resources.....</u>	<u>21</u>
<u>11 References and terminology.....</u>	<u>21</u>
<u>11.1 References to other processes.....</u>	<u>21</u>
<u>11.2 Terminology.....</u>	<u>22</u>

Executive summary

HET, High Performance Computing in Europe Taskforce, was established in June 2006 by eleven countries aiming at more intense collaboration in providing competitive resources for computational science in Europe. The target of HET is to propose a strategy and actions to develop a sustainable HPC Ecosystem in Europe which takes into account the main elements in solving the computing problems such as top-class supercomputers, existing national computing infrastructures, software development and need for competence development for computational science. The HET work includes the European scientific case for high-end computing developed during 2005 – 2006, which demonstrates a strong need for computational resources with a performance far beyond the currently available systems.

At the moment there is no major European resources which could match the capability computing performance of USA or Japan. That is an intolerable situation for the competitiveness of European computational science. The scope of HET work focuses both to the top of the computing performance, systems with a petaflops (10^{15} calculations in second) capability, and activities enabling utilizing such a system with extreme computing power efficiently. It is important to note that the computing hardware alone is inadequate in solving scientific problems – petaflop systems need also petaflop software, scalable algorithms, competent scientists and integration to the national computing centers for running that major part of the computing load which does not require the most capable and expensive systems. That is the main reason why instead of top resources only, HET discusses the whole HPC ecosystem.

The basic element of the HPC ecosystem is the performance pyramid. At the top we propose 1-3 supercomputer systems which would be funded through national sources, with additional funding from a European level. The middle layer consists of a number of national and regional supercomputers, which should have a high enough performance to be capable of running most of the computing load below petaflop and in addition serving as a development platform for the most scalable code directed to the petaflop systems. The bottom of the pyramid represents the local level of activity including the strong competence base of scientists in multiple European countries, renewal and new skills through local education and ability to raise visibility of scientific computing to attract new users. In addition to the well defined roles in each level of the pyramid, the transparent interaction between the layers is a major requirement.

The European Science Grid, the well connected infrastructure of all e-resources integrated at the hardware, software and communication level and the services utilizing them to benefit the research community at large (from purely scientific to applied or even commercial research) *is* the e-resources eco-system. The supercomputers, top or lower level, need to be integrated in the European Science Grid as any other resources. Through tight integration and optimal utilisation of distributed resources and skills in European countries it is possible to achieve high efficiency for the operations.

Due to the international competition the quality of infrastructure available for the research plays a major role. Computing performance is one important factor in achieving competitiveness, thus establishing a sufficient European infrastructure is a necessity. In order to close the gap currently existing from resource level available in Europe to those of USA and Japan, additional competitive tools should be utilized, such as investments to scientific software development and code optimisation, middleware development, data repositories and efficient networking. With the

right support, competence in designing petascale software could become competitive advantage for Europe.

Sustainable HPC ecosystem for Europe is a target which can be achieved gradually. EU FP7 provides one tool to develop HPC in Europe further. During the FP7 the European HPC ecosystem can be established based on the existing national infrastructures and active participation of other stakeholders for funding and using the HPC services. The stakeholders include – in addition to scientists using the services, nations funding computing infrastructure and computing centers providing services – European industry as a user and producer for parts of the computing environment, European Union and current EU supported Grid projects. Open questions, to which HET work targets to propose solutions, include funding models, peer review processes and organisation of the European level resources.

One of the major issues for HET work is how to ensure sustainability of such an ecosystem, including a strong base but also a sufficient top of the pyramid. At least the following issues help in increasing the life-time of the collaborative infrastructure:

- Integration and compatibility with national infrastructures
- Building on the current HPC-related work, such as successful European research infrastructures projects
- Sufficient amount of communication, experience and skill sharing between the national centres
- Steering by an European group with strong mandate
- Financial support from European union
- Acknowledgement of the importance of computational resources in increasing number of sciences
- Creating persistent research groups in HPC based in current experienced ones, and promoting new ones
- Creating persistent application development and optimization groups based in current experienced ones, and promoting new ones

1 Introduction

Research Computing has reached a level of maturity where it has become an integral, yet growing part of day-to-day research. This component¹ in research, complementing theory and experiment, is to be considered of strategic importance to Europe and so is the infrastructure to support it. High Performance Computing is an essential element in that infrastructure. It is the leading edge of Research Computing that paves the way for future directions in a diversity of scientific disciplines. It is also extremely important for key areas in research on which Europe's future and position in world will depend. A *science case* has been developed describing the broadness of applicability of Advanced Research Computing and the tangible future research products that can be the result of an established European High Performance Computer infrastructure.

Europe has a number of national and regional supercomputing centres, which are equipped with systems of considerable performance. However, these resources are widely distributed and resources are mainly used by national scientists – no ‘European-level’ computing resource exists, which could be used for the most demanding computational tasks. European scientists requiring extreme computing performance beyond the national supercomputers have limited possibilities, since the available systems in Europe lack behind the performance that is available in Japan or USA. In addition, both USA and Japan have already announced projects for installing computers with performance exceeding petaflop/s (10^{15} floating point operations in a second), a kind of project that has no match in Europe.

Computing power alone does not guarantee the best of scientific results and Peta-scale² supercomputers require much more than just the computing hardware, such as peta-scale application software, scalable algorithms, new programming models, parallelisation skills and competence in computational science. Cross-boundary collaboration and scalable software development enables us to close a part of the gap through HPC competence. However, to be able to fully serve the growing requirements for HPC, European level peta-scale systems for a limited number of top researchers need to be provided and strong backing by national and regional computing centres. This requires focus on the whole European HPC Ecosystem – from different hardware through scalable software to development of scientific applications.

The target of HET is to propose actions to develop a sustainable HPC Ecosystem for Europe, which includes all required elements from competence and network to top class resources for the most demanding computational tasks. Such an ecosystem is described here in its broadest coherence, together with an onset to an organisational structure and links to research councils and industry.

¹ Also commonly referred to as the “Third Methodology”

² See footnote in chapter 2.

2 Role and scope of the HPC in Europe Taskforce (HET)

The HPC in Europe Taskforce (HET) is formed by representatives of a number of European countries running national HPC services. Although the Taskforce is limited in its direct participation due to practical reasons, the Taskforce is open to input from all countries and reports voluntarily to the e-IRG to keep in touch with the European research community at large.

HET is to provide recommendations for building a sustainable European HPC Ecosystem with a focus on the very top of the performance pyramid, in *petascale*³ computing. Such a European policy should encompass the following elements:

- 1) A rationale for a well embedded HPC policy at the European level;
- 2) Within such a policy for the leading edge of computing, define or describe the role of:
 - the network (chapter 5);
 - the grid (chapter 5);
 - the coherence of resources within the grid (The European Science Grid), embedding of supercomputing (chapter 5);
 - the architectures of HPC computer servers (chapter 5);
 - the data access, service, durability and storage (chapter 5);
 - the software environment (scientific codes, commercial codes, optimisation issues, compilers for new architectures, etc.) (chapter 6);
 - prioritisation and peer review, relation with research councils (chapter 7);
 - organisational issues, who does what, who reports to whom (chapter 8) and operational matters;
 - funding and sustainability (chapter 9);
 - relations with vendors and parties on the development side (chapter 11);
 - relations with industry on the client side (chapter 11);
- 3) Realisation and rollout.

The first question is discussed in chapter 4 and should become more apparent through the whole text. The second question will be discussed in the following chapters and the third question will be elaborated in the closing remarks.

³ Petascale refers to systems with a theoretical peak performance of 10^{15} Floating point operations per second (PetaFlop/s). For normal general purpose processing, the theoretical peak is less relevant than the balanced combination of memory size, internal network and processor performance. Also the very balanced systems are, however, commonly referred to as Petascale computers.

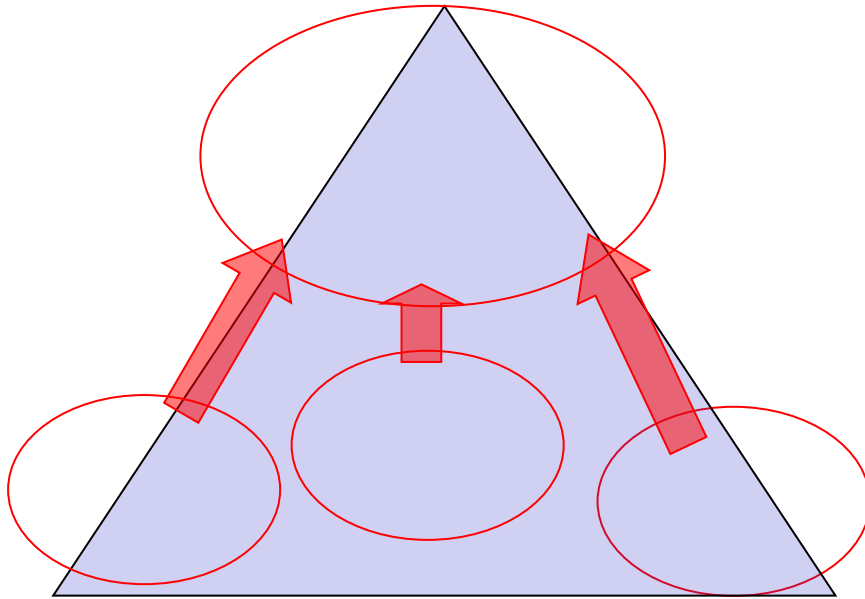


Figure 1: HET scope covers most of the performance pyramid with a focus to the top layers, but clearly indicating the areas that enable efficient usage of the petaflop computers.

3 An Ecosystem for High Performance Computing

3.1 Performance pyramid

The Performance Pyramid consists of three layers:

- The European level capability computing centers (Tier0), which represent the highest available computing power and provide computing cycles to top research groups over the country borders
- The national and regional computing centers (Tier1) with sufficient computing capacity to serve HPC users and to facilitate the access ramp to the resources of the European level centers
- The local computing centers (Tier2) in universities, in research labs or in other organisations

The different Tier-layers contain HPC systems of different capabilities, but the whole of the pyramid includes all relevant building blocks and a grid supported interconnection between them. This kind of service requires, in addition to the relevant computing power, for example efficient storage systems, networks, middleware and scalable software. HPC services provided by different levels of centers require also expertise to run the systems and develop/optimize/enable the applications in each of the levels.

In addition to providing sufficient resources in each of the Tier-layers, it is important to enable flexible and efficient interoperability, in some cases even strong integration, between different layers of the pyramid.

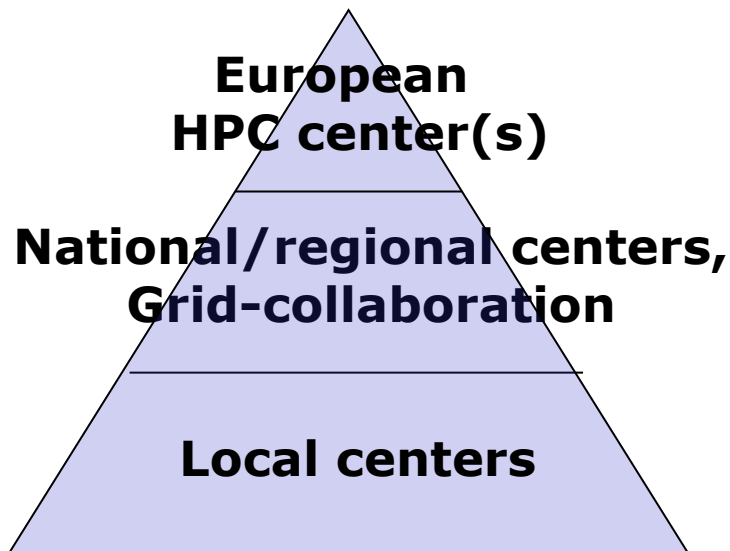


Figure 2: The Performance Pyramid

3.2 The need for a European High Performance Computing policy

Much attention is being paid recently to the internet and grid developments. Their development is of high importance and the activities are highly profiled. The views behind grids and communication are attractive and appealing due to their general nature. But a grid is only as strong as the resources it can couple together and add value to. Supercomputers are the leading edge of resources for Research Computing that pave the way for future directions in an abundance of scientific disciplines. It is also extremely important for key areas in research on which Europe's future and position in world will depend. A *science case* has been developed describing the broadness of applicability of Advanced Research Computing and the tangible future research products that can be the result of an established European High Performance Computer infrastructure. Only by including supercomputers in a European Grid environment the very goals set by our governments for our competitiveness in research can be reached.

3.3 Human aspect has a key role

To support the HPC ecosystem it is necessary to focus on competence development and other human aspects in various levels. Expertise and its availability in all levels of the computational pyramid have a key role in widening the HPC user base to new disciplines and strengthening the work in the already mature areas.

The growing investment is required for example in:

- Code development and optimisation skills in all levels of the pyramid – for users of small local systems as well as national or European level systems.
- Research oriented training activities
- Access to expertise and scientific/technical support
- Access to best practises, such as code libraries
- Collaborative tools promoting sharing of information between different research groups, both within the same discipline and multi-scientific

Establishing a collaborative and stimulating HPC atmosphere makes it possible to create different kind of networks of expertise, which will enable more efficient technology and competence transfer aiming at active interaction between different stakeholders, such as computational scientists, HPC service providers, hardware and software vendors and policy makers.

3.4 Why the top-of-the-pyramid resources are needed?

In competitive science and research only one thing counts: the time to solution of any particular problem. Therefore sufficient resources for qualified research should be readily available. It also means that High Performance Computing (HPC) cannot be considered without the broader context of all resources for Research Computing in Europe. Research Computing requires high end facilities to fit or match with the European and national research profiles and with the nature of the very application codes that are being used. This applies to the dimensions of each resource and its architecture. The European and national research profiles are determined by general or social interest, emphases placed by research councils and the sheer number of researchers in a particular field.

So where does HPC distinguish itself from any other type of research computing? The distinction is defined by the means problems can be solved that are intrinsically *latency bound*.

Supercomputers, as they were called at the time of single and multiprocessing, have always distinguished themselves by their ability *to solve latency bound problems*. The first latency hurdle is the time to access memory from the CPU. This still is a crucial factor which has not been overcome yet. Much, however, has been done to hide this latency by the introduction of a layer of memory caches, combined with compiler intelligence. It is expected that new technological advances in chip architecture will further reduce the latency problem on the processor level.

Present day computing facilities, which all belong to a category of massively parallel computers (whether the processors are scalar or vector type and 32- or 64 bits wide), still face a formidable latency problem, namely in the network that connects the processors. Dedicated internal networks are still factors faster than general purpose networks, but still are hardly capable of matching the ever increasing clock frequencies of the processors. And even if they ever would, the latency still would be critically dependent on the distance that the signals will have to travel: no more than 30 cm per nanosecond, due to the absolute limitation by the speed of light in vacuum.

Communication system	Latency (nanoseconds)
Supercomputer internal networks HW latency	3 – 5
Supercomputer internal networks SW latency	800 – 2000
Commercial internal network SW latency	2000 – 50000
Speed-of-light in glass fiber (HW) latency per km	10000
Long distance protocol (SW) latencies	> 1000000

Table 1: Overview of indicative system and network latencies in communication processes (HW= hardware, SW=software). Figures are best estimates, the lower the figure the better.

The overview in Table1 indicates the broadness of latencies of different components in a grid. It explains the very reason why supercomputer performance will never be attainable through grid computing, where by definition systems are involved that are spatially (far) apart. High end HPC (sometime referred as “capability computing”) can only be reached through systems that are densely packed, contain dedicated high performance low latency networks and have amended operating systems and HPC-compilers to optimally benefit from the architecture and enable any remaining latencies to be hid as much as possible. This combination of requirements is still much under development.

4 Supercomputing: Embedding in the general e-Infrastructure

4.1 Introduction

This chapter deals with the interrelation between network, grids and other resources including data, interoperability with national infrastructures and current European HPC activities (for example DEISA and other grid projects, Geant2 and optical fibre), but only in as far as these components have a direct relation with the focus of HET, namely the availability and efficient utilisation of Tier0 or top-of-the-line resources.

Enabling advanced research computing requires development work in lower levels, fast communication between components of the grid and fast and transparent access. Other issues, such as scalable software development, code parallelisation and optimisation, competence development in computational science and middleware development are discussed in chapter 6.

Two discussions can be distinguished on embedding supercomputing in a larger environment. One is supercomputing as a resource within the larger European Grid Environment, the European Science Grid. The other is supercomputing as the top of some hierarchical pyramidal resources concept. The resources are all nodes in the grid. Both points of view and their connection will be discussed next.

4.2 Supercomputing, Networking and the European Science Grid

Suppose one sees a future with one coherent e-Infrastructure for Europe (to start with) and, due to the connective nature of all resources involved, refer to it the European Science Grid⁴. Then the question arises what relation there is or is to be between HPC and such a European Science Grid. All the more while in the introduction it is argued that supercomputing performance cannot be obtained from grid-connectivity.

Without the existence of our European back-bone network (GÉANT) supported by the national research networks through the NREN’s, the notion of a grid that builds on top of that would be empty. As supercomputing was already defined as a means to solve latency bound problems, a supercomputing supporting network needs to be key on both delivering at the lowest latencies and the highest bandwidth available. Recent developments in switched optical networking are very helpful to say the least.

⁴ See various e-IRG White Paper documentations

For supercomputing the network is needed for bulk data-exchange, storage and retrieval, between systems and between user and system, real-time processing, visualisation and operational matters, such as cross scheduling and multi-system management.

There are at least four elements, ranked not in any particular order of relevance, but for reference only, connecting supercomputing with the grid developments:

1. In order for all relevant Research Computing to be conducted at the highest competitive level in Europe, supercomputers must be part of any European Science Grid infrastructure;
2. Grid developments support the *integration* of all e-resources and supercomputers are no exception in that respect. A scientific research activity is only rarely dependent on just a single IT component, including supercomputing: there may be data pre-processing, a real time component (with a sensor grid for example), visualisation both interactively or as post-processing. This can only be achieved by making supercomputing part of the whole European Science Grid infrastructure. Also the links between different layers of the computing resources pyramid are to be established through grid middleware;
3. A grid from supercomputers can (and already is, to a certain extent) also be formed, to add value to each individual component for the purpose of sharing a storage address space, cross scheduling or parallel processing of relatively individual computing components (just as in the standard grid);
4. Even at the supercomputing level there is no “one size fits all” for all applications, which means that there should be a transparent and flexible environment from which to choose a best resource, which is suitably done within a grid working environment;

The European Science Grid, the well connected infrastructure of all e-resources integrated at the hardware, software and communication level and the services utilizing them to benefit the research community at large (from purely scientific to applied or even commercial research) *is* the e-resources eco-system.

Its elements will be elaborated on in the next subsections.

4.2.1 Supercomputers must be part of the European Science Grid infrastructure

Whereas much of the *resources in* European Science Grid (*not the grid itself*) may be established in many ways and too much extent will grow naturally, supercomputing requires special efforts to be available. This is already true at a national level (many countries do not presently have a significant facility) but certainly at the European level. A science case has been developed separately to go with this document of HET, where the need for capability computing in sufficient capacity is clearly funded from a scientific and research viewpoint. An e-infrastructure bypassing the need for supercomputers would therefore be utterly incomplete.

On the other hand a supercomputer resources infrastructure without a proper grid embedding would be usable on the short term, but inefficient from three points of view: (i) from a working environment point of view as (ii) from a financial point of view, -i.e. all advocated added values of the European grid infrastructure would be lacking-, and (iii) most importantly because without the grid the access ramps to the high supercomputers are difficult to establish, resulting into an progressive isolation of the top system.

4.2.2 Grid developments support the integration of all e-resources and supercomputers are no exception

Supercomputers are well distinguishable elements in a computer infrastructure. They are able to support solving latency bound problems in science but in order to do so, they are to be embedded in an overall infrastructure. The processes that supercomputers are supposed to support may require vast amounts of input of various natures, batch mode, interactive or both types of processing or even real time processing, they may generate large amounts of data for intermediate purposes or as a final result that may in turn require post-processing, visualisation, data (re-)distributions, etc. Rather than inventing separate solutions for bulk storage, transparent data access, long term preservation or curation of data just for the sake of supercomputing all such requirements are to be part of the already ongoing grid-software environment developments that can as well be fully beneficial to supercomputing. All benefits that the grid development may offer researchers apply to the HPC community. The only thing that a grid is *not able* to yield is being an alternative to latency bound computing *per se*.

4.2.3A supercomputer grid

In particular, within the broader European Science Grid concept, a grid “densification” can be formed through special gridification of the existing and future supercomputers in the grid. This special gridification goes well beyond the features a grid normally offers and is especially meant to serve the HPC community to efficiently use the HPC resources. The present endeavours are embodied by the DEISA⁵ project. Such a supercomputer grid is not focussing at running a single application across multiple systems, but on overall efficiency including but not limited to cross scheduling, transparent high-speed and low latency data access, etc.

4.2.4 One size does not fit all

The only really relevant criterion for successful science, provided that facilities are there in the first place, is the time to solution. This requires a best fit to project the user code on the machine architecture. Therefore there is no such thing as a one-size-fits-all supercomputer for all codes, even though also the computing architectures are constantly evolving and are able to address increasing number of different applications efficiently.

The performance variation can be dramatic. This is why different architectures are required for optimal time-to-solutions for a wide selection of user codes. Differences can be found in the memory architecture or substructures, the IO to CPU-balance, the memory per node/CPU/cluster sizes, programming model etc.

4.3 Relation to other grid resources, data explosion

Today computing is an integral part of science. Most of the scientific disciplines are dependent of information technology. This is true also for a number of scientific instruments, especially when they produce vast amounts of data. Typically the data derived from the scientific instruments needs to be processed further, often requiring thousands of processors and high amounts of storage capacity. Grid infrastructures can be used to gather the data, but in many cases supercomputer systems are required for processing the data.

⁵ DEISA= Distributed European Infrastructure for Supercomputing, a FP6 supported project

Examples of scientific instruments include:

- Particle accelerators, such as CERN Large Hadron Collider (LHC)
- Radio telescopes
- Satellites, such as Planck-satellite
- Synchrotrons
- Microscopes
- Fusion reactors, such as ITER

A number of potential future scientific instruments, which in most cases require HPC resources and have substantial challenges related to data management, have been listed in the ESFRI⁶ list of opportunities.

Utilisation of HPC systems can possibly replace part of the need for expensive instruments, for example by realistic simulations requiring often huge amount of computing capability, apart, of course, of producing results that can never be achieved by experimentation, due to impracticalities or ethical reasons (such as operative experiments on humans). On the other hand, the ability to simulate even more complex phenomena can also accelerate the need of instruments in order to verify or further develop the simulation results. In any case there is a connection and potential synergy between experiments and modelling, which, however, can only work if the time scales for modelling are similar to those of the experiment.

4.4 The pyramidal computing resources infrastructure -No top without a body-

Supercomputers are often referred to as forming the top of a pyramid of resources, which then in turn are referred to as mid-level systems and basic resources at the bottom. Although the various pyramids may yet differ in height or broadness of base, this picture makes much sense. The picture usually reflects at the same time different layers of responsibility, such as – going from bottom to top – to individuals and research groups, universities or institutes, national and/or European responsibilities. Again at the same time these responsibilities reflect usually also the cost in terms of real money of investment in any of these layers. The higher the direct investment level, the more cooperation and the more strategic visions are required to get the funding and the broader context in which the funding takes place. But the pyramid is not just a series of resources ranked after performance and stacked on top of each other. The pyramid is essential to a solid research computing environment and should be kept in tact at all times⁷.

For the foreseeable future there seems no limit in wanted computer resources/cycles for science and research. This insatiable need can only be satisfied if at all levels of the pyramid the required resources are available and sustainable: desktop facilities, local clusters (or a gridified cluster environment) for direct and on demand access, higher level systems for (very) demanding codes at the national level (or an equivalent gridified environment) and national/European level supercomputers for “qualified” science/research. The pyramidal structure is essential for this scheme, because building on the top of the pyramid only makes sense if the base is well founded. If the middle layer, that is due to be formed by national HPC systems, is absent, the work permeates one way or another to either other layers. This leads to a cost ineffective use of those resources if not directly to a shortfall in scientific production and quality. The disappearance of the middle layer may readily result from a funding scheme for the top where effectively the

⁶ ftp://ftp.cordis.europa.eu/pub/esfri/docs/esfri-roadmap-report-26092006_en.pdf

⁷ At all times: anyway as long as there are “significant developments at the very top of the computing market”.

funding for the middle layer is used to pay for the top. Such a scheme would not yield a continuum of resources and certainly not more cycles for scientists but merely form another bureaucratic complication for the access to the still urgently needed cycles by scientists and researchers.

The grid infrastructure is the vehicle that glues all resources in the pyramid to one transparent environment and that has to guarantee that all resources in the pyramid will be optimally used for their very purposes. This should not be underutilised nor be infinitely queued up.

The layers in the pyramid are admittedly not very sharp. A system based on a given technology may be in the basic layer if the number of processors is small, in the middle layer if the number is significant and at the top if their number is abundant. But in addition to these layers based on performance of “size” it is also the access possibilities that co-determine the layers: at the basic level the access is instantaneous, at the middle level there may be queues or appointments with colleagues or other user groups about sharing policies. At the top an additional hurdle must be taken, which usually is *peer review*. But without local and institutional/national and easily accessible resources it will be difficult to make one’s case for access to the peer reviewed environment.

The form of the European-wide pyramid may, in addition to the elements already mentioned, become more dependent on the total cost of ownership/cost per cycle than has been the case. This is due to a few parameters that were around already but have become differently valued more recently. Such parameters include energy cost, heat production, support cost.

For European ecosystem it is crucial to be able to collaborate and share resources among centers both within the same level of the performance pyramid and between different levels. As the national infrastructures will most likely continue to dominate in terms of funding, the one possible scenario presented in Figure 6.1 is relevant and demonstrates the potentially wide distribution of resources creating opportunities in utilizing optimal timing in installing systems.

It is important to provide the suitable level of resources in the pyramid for the research groups in order to guarantee cost efficient usage and still reserve the high-end resources only to those applications that can really benefit from such a system. An example of the cost optimisation attempts is provided for example in an article by Gordon Bell, Jim Gray and Alex Szalay⁸.

⁸ Petascale Computational Systems: Balanced CyberInfrastructure in a Data-Centric World (September 2005)

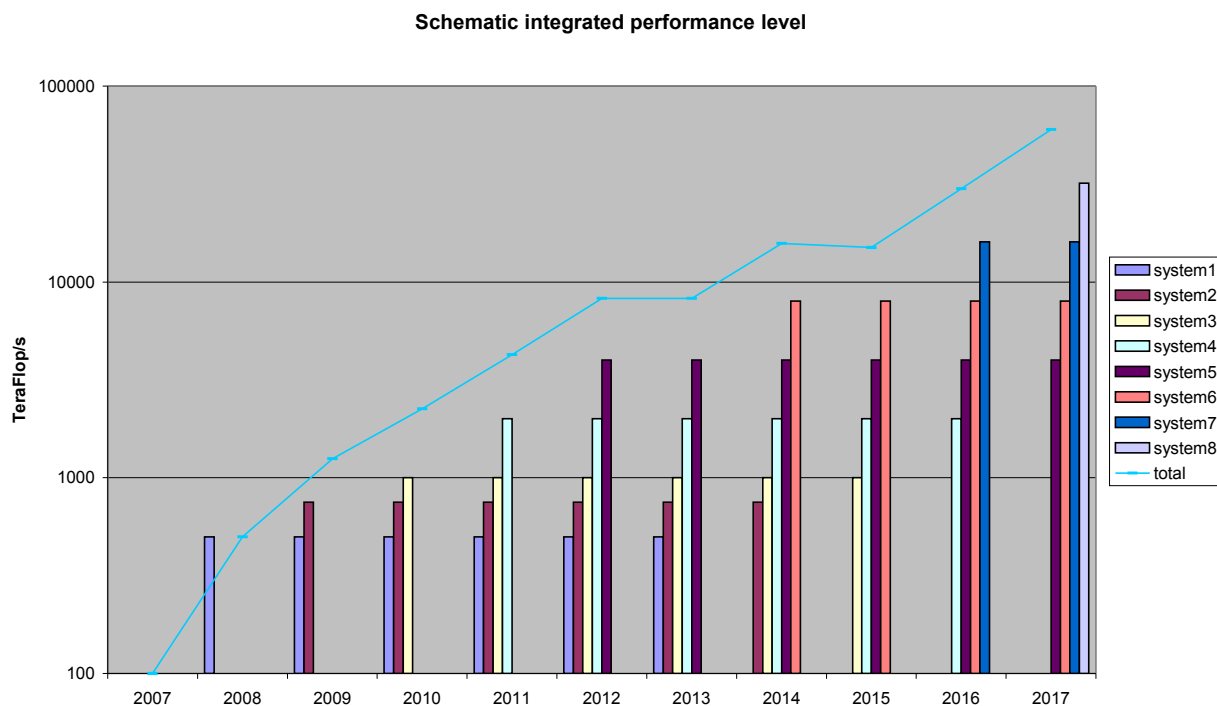


Figure 3: Schematic view on the integrated performance with an 18-months investment cycle, a doubling of the performance of each new system and an 6-year depreciation period (y-axis is logarithmic)

5 The software environment

5.1 Introduction

Software is crucial to any efficient computer use. The more expensive the computer system is, the more it pays off to invest in optimised software. Most likely a from scratch development of new software, based on algorithms selected in direct conjunction with a machines architecture contributes more to science than a system upgrade to yield the same results. Software is also a domain where Europe could really make a difference.

There are several aspects to the topic software relevant to HPC:

- how to organise software development, maintenance and optimization within the HPC Ecosystem
- how to select software development projects
- how to engage the software development industry into developing SW for HPC
- how to create a role for European SW industry for HPC systems

Note that everything needs not to be developed from scratch. Many elements of this topic may find their place automatically, once the word is spread that Europe's engagement for HPC is really a long term commitment and not a one time effort.

5.2 Software development and maintenance

New computer architectures require new software development or at least (re-)optimisation of existing codes. However, the architecture of the future most likely bring new programming paradigms to support such new architectures. To benefit from leading edge technology to produce leading edge science new software is indispensable. The DARPA initiative⁹ for the development of new computer architectures for Petaflop computing involved the design of new programming languages and operating systems to go along with the new hardware designs. Europe has much experience in software development, but from scratch development of codes for specific research is rare. Also the cooperation of computer scientists with disciplinary oriented application fields can be improved upon.

The software development for HPC may be rather specific to this domain. But the organisation to get this off the ground best links to the more general elements of software development as described in the e-IRG Roadmap¹⁰ (Software life cycle management, open standards, software repositories etc.)

Since computer performance in all the levels of the pyramid is highly dependant on scalability and quality of the software, the HPC Ecosystem can not work properly without sufficient effort in software development work.

5.3 Selection of software development projects

In as far as scientific software development is concerned, there are two tracks:

- from scratch development of new codes;
- optimisation and adaptation of existing scientific user codes to new architectures.

HET sees a role for both the tracks, but through different channels of organisation. The prioritisation of the development of new codes should be left to the very disciplines within which the code is to serve research. Only the means to do the code development should be part of the HPC Ecosystem project and budget. Such projects typically take several man years to finish.

The prioritisation of code optimisation could be left to the research councils as well, but not to disciplinary committees. Rather a more general committee could conduct a light weight procedure to allocate typically a few man weeks to man months to do the optimisations. After all it is beneficial to the whole user community if an expensive system is used in the most efficient way.

5.4 Involvement of industry

As already stated, Europe does have a record in code development in many fields. Notably Europe plays an important role in the gaming industry. But Europe contributed to the BLAS-routines¹¹, available on any general purpose computer for scientific computing. The NAG-library is European and several CFD-codes, Finite Element codes, computational chemistry packages and statistical code packages have European roots. Many community codes used in a wide area of scientific computing are developed and supported in Europe. Also in compiler design Europe has an impact. But Europe could do much better. Such a spin off may certainly be expected from a long term European commitment to HPC. The mere expectations that will arise from such a long term commitment will naturally raise the interest of industry. Discussions with industry are to be organised to maximally exploit any industrial interest in this field from the very start. Such

⁹ DARPA initiative 2003

¹⁰ e-IRG e-Infrastructures Roadmap, available through www.e-irg.eu

¹¹ BLAS for Basic Linear Algebra Subroutines

discussions should yield direct input for future pre-competitive procurements for leading edge supercomputers.

6 Prioritisation and peer review, relation with the research councils

The HET has been setup with the objective to contribute towards the improvement of the European Research Infrastructure in the area of scientific computing and data handling. It has clarified the requirement to make competitive world class supercomputers available for researchers involved in large size simulation and modeling, and has put forward a scientific case for such.

The requirements for the top-of-the-pyramid resources have been defined taken into account the principles of interoperability with existing national and thematic resources, and ensuring the most relevant scientific contribution. It is expected that scientists using the top-of-the-pyramid resource are expected to contribute in advancing science through numerical simulation. This calls for an evaluation process based on peer review to allocate computer resources.

Appropriate steering committees are required for launching the calls for proposals requesting computing time from European resources and evaluating and ranking the proposals. This should be organized through a European level steering body, but still taking in account the national research councils and possible national prioritization work as much as possible. In the evaluation both the scientific value of the project and implementation of the computational algorithms – such as their scalability and efficiency – should be considered.

The peer review process is fully described in an associated document.

7 Organisational and operational matters

The organisation of the European level HPC services is most probably heavily influenced by the division of funding for the Tier0 resources. Since most of the funding will be required from national sources, the control over the organisation is accordingly mostly national and needs to interoperate with the existing organisational processes.

A collaborative organisation for taking care of the part funded by European Union, in case such funding is granted, is necessary. Depending on the funding volume and resource coverage various stakeholders should be included in decision making, such as the sites or countries hosting the European systems, national HPC policy groups, European Union and scientific community using the high-end computing services. In addition, the organisation should coordinate or combine the activities with existing European grid projects for HPC, for example with the DEISA project. The organisational form of the European HPC services could be, for example, a long-term project structure spanning over the whole 7th EU frame program. Based on the experience obtained during the construction of the HPC service a more sustainable model, for example a legal entity formed by major stakeholders, could be developed in a later stage.

Organisation of European petascale HPC services need to consider at least the following points:

- Level of distribution: How many petascale centers will be established and how are they distributed geographically? What kind of interaction is required with other centers with regional or national resources? Decoupling of hardware and technical/scientific support?
- Utilisation of existing premises suitable for high performance computing systems
- Synergy and workload distribution with major EU-funded grid projects, such as DEISA and EGEE-II

Operations for petaflop systems require special arrangements in terms of premises, electricity and cooling. It is important to consider the cost and functional requirements for the supporting structure including the above mentioned technical requirements and need for personnel to efficiently run the services.

Additionally new technologies induce new computing paradigms and future technologies will be no different. This requires that much support be offered to users of advanced equipment at an expert level that goes well beyond technical support. Field expertise in combination with technology insight is required to further science through HPC. This again changes the balance between the initial investment cost and total cost of ownership over the lifetime of a machine.

8 Funding

Since the budgetary requirements are far beyond that is expected to be invested in this area through EU Research Infrastructures funding, the national funding has the key role in building a sustainable infrastructure. In order to take the maximal benefit of the technological refresh delivered by technological advances, software support must be part of this infrastructure

For funding scenarios in establishing the center(s) with an extreme computing power there are a number of alternatives. One of the most popular scenarios is complementing national funding of some of the major European countries with sufficient funding from European Union, in order to open part of the resource for European usage through peer review process. This is naturally dependant on the terms of EU funding as also willingness of the hosting country to include European activities into the national systems.

9 HPC Ecosystem and sustainability

The full HPC Ecosystem considering the whole performance pyramid is a vast area including in practise the whole European scientific community. Thus it is necessary to focus HET work to the parts of HPC Ecosystem which have a key impact to the competitiveness of European computational science: the upper layers of performance pyramid, collaboration among the European nations and their supercomputing centres and supporting activities, such as scalable software development and competence development.

The key issues for building European level HPC Ecosystem include:

- Efficient use of existing national infrastructures and HPC expertise, since they will be in any case the most dominant funding source for HPC
- Competence and skills utilization over the country borders, allowing HPC centers to specialize increasingly to provide services utilizing their strengths

- Activities supporting top of the pyramid (peta-scale software etc.) in order to use the extreme performance and the most costly resources as efficiently as possible, and for the suitable kind of scientific problems
- Strong scientific case for high end computing requirements
- Close relationship with scientific community and industry main actors
- Utilization of best practices from the existing infra projects to speed up the production usage
- Rules of the game for resource exchange to agree on the cost division
- The interaction of the different layers in the performance pyramid for optimal cost-efficient load sharing purposes
- The access rights and restrictions for European scientists including decisions on priorities and methods to wide coverage in Europe
- Continuous feedback and interaction between the European scientific community, HPC policy groups and HPC centers

Sustainability is the key issue for building a European HPC Ecosystem with infrastructure and services targeted to be usable for decades through continuous renewal. Supercomputers, as any computers, get old in a few years, and can also be purchased and installed in less than a year of calendar time. However, building a supercomputer centre with sufficient personnel with technical and scientific competence can take 10 years.

How can European HPC Ecosystem become sustainable? At least the following issues help in increasing the life-time of the collaborative infrastructure:

- Integration and compatibility with national infrastructures, since they will be maintained anyway
- Building on the current HPC-related work, such as successful European research infrastructures projects, which have already developed working practices and kind of ‘informal standards’
- Sufficient amount of communication between the national centres
- Steering by an European group with strong mandate
- Financial support from European union
- Acknowledgement of the importance of computational resources in increasing number of sciences
- Creating persistent research groups in HPC based in current experienced ones, and promoting new ones
- Creating persistent application development and optimization groups based in current experienced ones, and promoting new ones

10 Involving European industry

10.1 Promoting European HPC industry

There are a limited number of European HPC vendors compared to those of USA and Japan. This makes it even more difficult to compete for computing resources since installations of petascale computers in USA and Japan, for example, support at the same time the domestic computing industry. In order to provide competitive HPC resources to European researchers, we need to balance non European systems purchases with European ones when available.

It is important to ensure that European centers running supercomputer procurements do have a possibility to choose from a wide variety of solutions based on their competitiveness in terms of price/performance and quality. At the same time it is equally important to consider possibilities to advance European HPC industry through the activities combined with development in European HPC centers.

There are a number of methods which can increase European share of the computing industry. Even if there are not many European HPC hardware vendors or processor manufacturers, the more we can shift the R&D activities of non-European vendors to Europe, the more we increase the European impact. Already now some of the main players do have manufacturing in Europe and even more have established Europe-based research centers or have research agreements with European centers. By increasing a trend to transfer activities to Europe, for example by giving it a suitable weight factor in evaluation of HPC procurements, we can not only promote European HPC industry but also increase the quality of our HPC services with proximity of R&D from European customers.

If we have a very limited view on what HPC services require, e.g. consider only computing hardware, the backward European position may easily seem insurmountable. Fortunately this is not the whole truth. By focusing on the whole HPC Ecosystem, all levels of the performance pyramid, we do find a number of competitive factors which help us to close the gap to USA and Japan:

- Balanced performance pyramid with efficient collaboration across the different levels
- Code optimization and scientific code development: sustained performance is the main focus
- Strong multi-scientific competence across the EU countries with effective skills development through training
- Strengths of the European industry in different HPC sectors
- Experience of the European industry in building top size computers
- Excellence in building centralized and distributed cluster systems with impact from optimization of scientific code (integrator role)
- Effective collaboration between EU countries
- Low bureaucracy in accessing computing resources
- Europe is leading in middleware development: ARC, gLite, Unicore, AAA, security, tools for performance analysis, grid, etc.

Therefore, HET proposes establishing an interactive forum with major vendors, discussing for example suitable criteria deemed necessary for vendor partnerships in order to qualify for pre-commercial procurements. Such criteria could include for example:

- Significant R&D activities in Europe about processor architecture, computer science (OS, compilers, tools, etc.), file system, file and data management, visualization or communication systems
- Significant R&D partnerships with universities or research labs on any of the topics above
- Significant hardware manufacturing or integration, even through a major partner

Significant may be qualified through a minimum level of contribution in any of the forms listed above and expected from any vendor to participate in the pre-commercial procurement.

In addition the charter may include some additional criteria such as adhering, for example, to open source software.

10.2 Industrial usage of HPC resources

High Performance Computing has traditionally been mostly used by academic users. Industrial usage in Europe is not very wide, although encouraging examples fortunately exist. Computational science based technology transfer from academic environments to industry is slow and in many cases realized through experienced people changing their job from academic to industry, not by companies starting intensive training programs in applying HPC and computational methods to their existing environments. One of the obvious bottlenecks is the lack of trained experts in computational science with a strong competence in parallelisation and optimisation techniques, for example. Another factor is the difficult porting process from old serial codes to parallel applications demanded by the modern computers with 1000+ processing elements.

However, there are also a number of industrial companies that have been using HPC resources already from the beginning. Some of the most typical industries are automotive and aerospace industry, even though the scope of industrial applications has been widened during the recent years. Drug design, development of new materials and petroleum industry are examples of areas in which HPC is not anything new. Bioinformatics, medical imaging and telecommunications are areas in which the computational methods have been taken to use more recently.

The business model however with such partners has to be flexible and innovative. Experience shows that, while not being particularly interested in buying cycles at HPC centres (a low added value for both partners with practical difficulty ranging from security issues to code porting), industrial partners can establish excellent collaboration projects when they want to explore a new technology or test innovative ideas. Performing such activities at HPC centres give them the possibility to interact, to learn and to share with the newest level of technology and with the highest level of scientific and technological innovations by working with research groups: the outcome is new knowledge generated, which is then re-injected in the innovation pipe and in fine, in the economic development.

Increasing the HPC utilisation of European industry is one of the main priorities since a working ecosystem requires commercial utilisation and funding mechanisms to be sustainable. To transfer the competence and increase involvement of the industry is not necessarily easy, but is a necessity in the near future. And vice versa, an available first-class HPC infrastructure will help European industry to become more competitive.

11 References and terminology

11.1 References to other processes

The vision for a European High Performance Computing policy as here developed by the HPC in Europe Taskforce is not a stand alone vision on supercomputing and also not a vision by just this Taskforce. The description of the sustainable ecosystem for HPC is in line with the recommendations of the e-IRG Task Force on Sustainable e-Infrastructures (SeI)¹², a further development based on the official e-IRG Roadmap¹³ and in line with the outcome of various EU-meetings, in Bruxelles¹⁴ and elsewhere.

HET utilizes the work from HPCEUR¹⁵ project.

11.2 Terminology

HPC

High Performance Computing

HET

High Performance Computing in Europe Taskforce, established in June 2006 with a mandate to draft a strategy for European HPC ecosystem

Petaflop/s

Performance figure: 10^{15} floating point operations (calculations) in second

Petascale

Systems or applications capable of petaflop/s computing performance

Tier0 – Tier2

Different performance levels of computing infrastructures. Tier0 stands for the most efficient European services, Tier1 for national or regional services and Tier2 for local services

Latency

Delay required for the communication

e-IRG

e-Infrastructure Reflection Group. e-IRG is supporting the creation of a framework (political, technological and administrative) for the easy and cost-effective shared use of distributed electronic resources across Europe - particularly for grid computing, storage and networking.

ESFRI

European Strategy Forum on Research Infrastructures. The role of ESFRI is to support a coherent approach to policy-making on research infrastructures in Europe, and to act as an incubator for international negotiations about concrete initiatives. In particular, ESFRI is preparing a European Roadmap for new research infrastructures of pan-European interest.

HPCEUR

Project aiming at establishing European high-end supercomputer installations was run until June 2006, after which the project was formally ended and the remaining work transferred to HET

¹² <http://www.eirg.eu/publ/>

¹³ <http://www.eirg.eu/roadmap/eIRG-roadmap.pdf>

¹⁴ <ftp://ftp.cordis.europa.eu/pub/ist/docs/rn/changes-060511-DD-event-supercomp-report.pdf>

¹⁵ HPCEUR project aiming at establishing European high-end supercomputer installations was run until June 2006, after which the project was formally ended and the remaining work transferred to HET

DEISA

Distributed European Infrastructure for Supercomputing Applications. DEISA is a consortium of leading national supercomputing centers that currently deploys and operates a persistent, production quality, distributed supercomputing environment with continental scope.

EGEE-II

Enabling Grid for E-science. The project provides researchers in academia and industry with access to a production level Grid infrastructure, independent of their geographic location.

GÉANT2

Seventh generation of pan-European research and education network