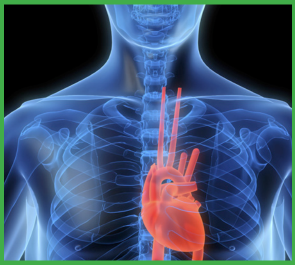
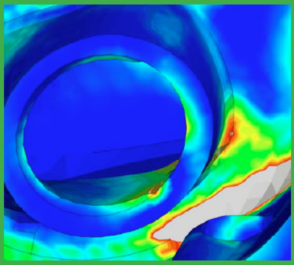
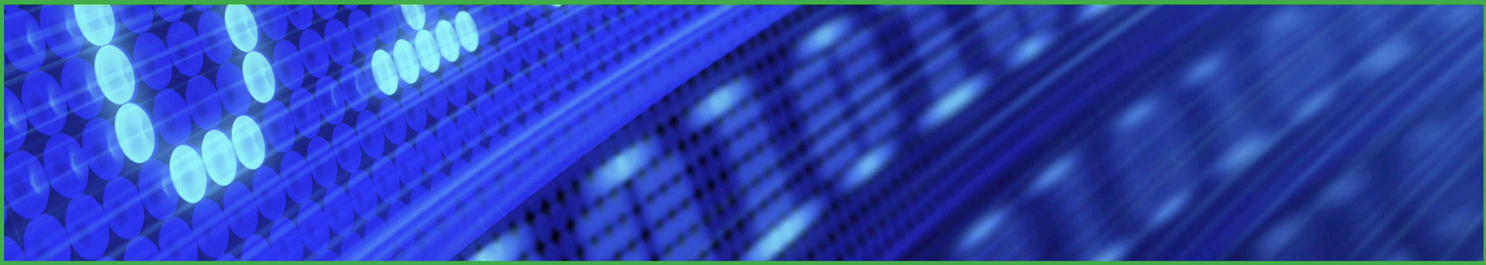
A Strategy for Research and Innovation through High Performance Computing





Editors

Mark Sawyer, Business Development and Project Manager, EPCC
Mark Parsons, Executive Director, EPCC; Associate Dean for e-Research, University of Edinburgh

PlanetHPC is supported under Objective “Computing Systems” of Challenge 3 “Components and Systems” of the ICT Programme of the European Commission.

Executive summary

HPC is without doubt a key enabling technology for many technologically advanced nations in the 21st century. Many countries world-wide are investing in HPC and some, most notably the USA, China and Japan are investing vast sums of public money on related infrastructure.

The importance of HPC is summarised well by the statement contained in a recent report by industry experts IDC:

"Today, to Out-Compute is to Out-Compete".

The rise in performance and capability of HPC is now threatened by technological factors, in particular energy requirements and the fact that microprocessor speeds are no longer expected to increase in the way they have in the past.

Europe has a strong tradition of research and commercialisation of HPC applications. On one hand, Europe is threatened by global competitors making huge investments. On the other hand Europe is well placed due to its world leading position in mobile and embedded computing, offering a strong competitive advantage when it comes to addressing some of the major challenges of HPC – energy-efficiency, dependability, and real-time responsiveness. Europe must view the challenges ahead as an opportunity to invest for its future by developing new technology, exploiting synergies with other domains, providing advanced facilities, educating its workforce and promoting innovative use of HPC. As the HPC landscape changes, the economies which adapt the fastest will be the ones that will gain the greatest benefits.

Unless action is taken, the societal and economic benefits of HPC and computing in general will stall. The leadership that Europe has shown in many domains of computing and its applications over the last 20 years must be maintained and expanded if the European economy is to remain innovative and competitive. Investment in research, development and training carried out at the European level together with measures to promote the early industrial and commercial uptake of new technology will be essential elements in a successful strategy, in which HPC is one of the key drivers.

A long term programme of R&D must be initiated to overcome the major technological hurdles that are identified in this report. Areas which should be prioritised include:

- Highly scalable methods for modelling and simulation that can exploit massive parallelism and data locality.
- New programming models and tools, targeted at massively parallel and heterogeneous environments.

- Decoupling application development from HPC.
- Technologies to support new and emerging applications which require robust HPC with real-time capability.
- Data-intensive HPC.
- Low-energy computing from both an architectural and application perspective.

The research programme must be focused ultimately on delivering results that are relevant to tackling the economic and societal challenges that lie ahead. It must be ambitious and aim for a transformation of the HPC market in Europe, rather than incremental change.

It must address the widest possible spectrum of applications and involve all actors in the HPC domain. To be successful in terms of industrial competitiveness, it is important that all relevant industrial actors collaborate: industrial end-users, HPC service providers, system vendors, Independent Software Vendors, research software providers, application experts and HPC consultancies.

The preparation for this programme of research needs to start as soon as possible in Framework 7 and continue through Horizon 2020. The following actions should be taken immediately to stimulate the HPC market and to prepare the foundation for future longer term R&D.

1. HPC pilot networks which stimulate the HPC solutions market place.
2. Research & development activities, which stimulate technology development in the HPC domain, and transfer of technology to the HPC domain from other computing domains and constituencies. This is essential if Europe is to maintain its leading position in this key technology.
3. Visioning, roadmapping and constituency-building activities, which prepare a long term strategy for R&D&I initiatives under Horizon 2020.

These actions will pave the way for longer term research aimed at ensuring that Europe fully exploits HPC to 2020 and beyond.

The recommendations in this report are endorsed by industry leaders and HPC experts across Europe, with whom the editors have carried out extensive consultation under the PlanetHPC project.

Contents

Contributors

Amedro	Brian	INRIA, France
Appa	Jamil	BAE Systems, UK
Aston	Martin	CFMS, UK
Barr	John	The 451 Group, UK
Björkman	Gunnar	Volvo Technology Corporation, Sweden
Bodin	François	CAPS Enterprise, France
De Bosschere	Koen	University of Ghent, Belgium
Butcher	Phil	Wellcome Trust Sanger Institute, UK
Falter	Hugo	ParTec Cluster Competence Center, Germany
Fischer	Rolf	INTES, Germany
Geiger	Alfred	T-Systems, Germany
Gaudry	Eloi	FFT, Belgium
Hempel	Rolf	DLR, Germany
Lanteri	Stéphane	INRIA, France
Lavignon	Jean-François	Bull, France
Lecina	Gerard	Dassault Systèmes, France
Lecomber	David	Allinea, UK
Linden	Johannes	Fraunhofer SCAI, Germany
Lonsdale	Guy	scapos AG, Germany
Lundin	Tomas	Saab Group, Sweden
Maccallum	Peter	Cancer Research UK, UK
Öster	Per	CSC, Finland
Özer	Emre	ARM, UK
Papadopoulo	Jean	Teratec, France
Parsons	Mark	EPCC, UK
Piatkowski	Piotr	COMARCH, Poland
de Pommery	Etienne	ESI Group, France
Ramirez	Alex	BSC, Spain
Reinemo	Sven-Arne	Simula Research Laboratory, Norway
Requena	Stéphane	GENCI, France
Resch	Michael	HLRS, Germany
Roucairol	Gerard	Teratec, France
Sawyer	Mark	EPCC, UK
Schelkle	Erich	Automotive Simulation Centre Stuttgart, Germany
Simmendinger	Christian	T –Systems, Germany
Taylor	Gary	Brunel University, UK
Vile	Adam	Excelian, UK
Wesner	Stefan	HLRS, Germany
Wray	Francis	EPCC, UK

Responsible officers at the European Commission: Panagiotis Tsarchopoulos, Max Lemke, Unit INFSO-G3 Embedded Systems and Control.

Other European Commission participants: Bernard Fabianek, Jorge Gasos, Leonardo Flores.

Introduction

In the 21st century, High Performance Computing (HPC) is without doubt a key enabling technology for technologically advanced nations. Many countries world-wide are investing in HPC and some, most notably the USA, China and Japan are investing vast sums of public money on related infrastructure.

In this report, HPC is considered in its broadest definition covering demanding applications that are compute and/or data-centric, which exploit parallelism on a large scale. HPC systems can range from a desktop computer with an accelerator, through clusters of servers and data centres up to high-end custom supercomputers.

HPC has had a major impact on industry and commerce and is an established and indispensable tool in many industrial and societal sectors. New applications are constantly emerging which are able to exploit HPC. The benefits made possible by HPC go beyond a positive return on investment. Improvements in healthcare, the development of efficient transportation systems, the quest for renewable and clean energy sources, and support to decision making through faster than real-time simulations on live data in emergency response are examples of how HPC can have a major societal impact. The importance of HPC is summarised well by a statement contained in a report by industry experts IDC[1]:

"Today, to Out-Compute is to Out-Compete".

Many achievements and benefits have been made possible by the remarkable advances in HPC over the last 20 years. The computational capability of the world's fastest computers has increased by a factor of over a million in that time. This significant increase in computing power has not just benefitted the users of the world's top supercomputers; systems which are routinely used by industry today have capabilities many times those of elite computers a decade ago. HPC is now well on its way to being an indispensable tool for SMEs and businesses of all kinds.

This continuing rise in performance and capability is now threatened by technological factors:

- Energy consumption is a major challenge. Continuing to build ever larger machines using today's technology will lead to prohibitive energy requirements.
- Processor speeds are no longer expected to increase in the way they have in the past. The only route to improved performance of applications is through massive parallelism.

Almost all users of HPC either use their own systems or gain direct access to systems run by third parties. The cloud computing paradigm has been highly successful in the enterprise computing domain, but there is, as yet, no analogue of cloud computing that can deliver true high performance. Despite its lack of impact so far in the HPC domain, cloud computing is too important an issue to ignore. The business models on which cloud computing is based cannot currently be replicated in the HPC domain due to technological barriers. For clear technical reasons, attempts to provide HPC while maintaining the virtualised access model that underpins cloud have failed, except in a few niche examples. Some service providers allow HPC partitions to customers on pay-per-use via a web interface, and call this cloud

[1] http://www.hpcuserforum.com/EU/downloads/SR03S10.15.2010.pdf

computing, although in fact the resources are provided by direct access rather than as virtual machines.

HPC systems of the future will be greatly different from those of today. New skills, techniques and access models must be found to use them effectively:

- Applications must be made to scale so that they can use processors in numbers that are several orders of magnitude higher than today.
- HPC must be an easily accessible technology to businesses of all types supporting a huge range of applications.
- HPC must become more usable. It should not be necessary to be an expert in HPC to use it in a business context. It must be easy for developers to build and integrate applications based on HPC.
- Developers must not lose the ability to migrate applications to other HPC systems (current and future) with relative ease to avoid vendor lock-in.

Europe has a strong tradition of research and commercialisation of HPC applications. On one hand, Europe is threatened by global competitors making huge investments. On the other hand Europe is well placed due to its world leading position in mobile and embedded computing, offering a strong competitive advantage when it comes to addressing some of the major challenges of HPC – energy-efficiency, dependability, and real-time responsiveness. Europe must view the challenges ahead as an opportunity to invest for its future by developing new technology, exploiting synergies with other domains, providing advanced facilities, educating its workforce and promoting innovative use of HPC. As the HPC landscape changes, the economies which adapt the fastest will be the ones that will gain the greatest benefits.

Unless action is taken, the societal and economic benefits of HPC and computing in general will stall. The leadership that Europe has shown in many domains of computing and its applications over the last 20 years must be maintained and expanded if the European economy is to remain innovative and competitive. Investment in research, development and training carried out at the European level together with measures to promote the early industrial and commercial uptake of new technology will be essential elements in a successful strategy, in which HPC is one of the key drivers.

Meeting these challenges requires long-term, focused support. This document outlines a starting point for this support through immediate action within Framework Programme 7, together with the outline for a longer-term programme of industry-driven research and development under Horizon 2020.

The economic importance of HPC

What is HPC?

High Performance Computing (HPC) can be defined in many ways. For some it means the small collection of the world's most powerful supercomputers. For others it means any computer system larger than a regular desktop PC used for modelling and simulation. For others still, its application is broader than simply modelling and simulation and covers large-scale data processing, real-time HPC and other applications.

For the purpose of this report, the broad definitions of HPC applications and systems presented above are used: HPC applications are compute and/or data-centric applications that exploit parallelism on a large scale. HPC systems range from a desktop computing with an accelerator, through clusters of servers and data centres up to high-end custom supercomputers

Over the past fifty years modelling and simulation have come to complement theory and experiment as a key component of the scientific method. Used where the problem is too big, too small, too short or too long for physical experiments, many of our scientific breakthroughs and technological advances rely on models simulated on high-performance computers. At the same time, the digital revolution of the past two decades has led to unprecedented demands for high-performance data-processing systems. More recently, real-time HPC systems, sometimes embedded in large infrastructures, have appeared which can assist humans in instantaneous decision taking (for example a medical simulation that supports a surgeon performing an operation).

There can be no question that HPC will continue to be a vital technology, and that demand for greater computing power will continue. As described later in this report there are applications that will require dedicated systems that are hundreds of times more powerful than those of today. In other applications there will be a need to run multiple simulations at the same time to explore a design space. Both the capability and the capacity of HPC machines must increase to meet these demands.

The forthcoming PlanetHPC Roadmap is intended to encompass the whole spectrum of HPC from the largest to the smallest systems and from traditional modelling and simulation to data-processing applications. In choosing this broad ecosystem, we acknowledge that this roadmap must align with roadmaps and initiatives from other communities such as those from the supercomputing community (PRACE[2]), the Cloud/Grid community (EGI[3], EU strategy for Cloud Computing), and the embedded and mobile systems communities (ARTEMIS[4], HIPEAC[5]). It is our intention that it does so.

---

[2] http://www.prace-project.eu

[3] http://www.egi.eu

[4] http://www.artemis.eu/

[5] http://www.hipeac.net

Why is HPC important to industrialised economies?

Many countries world-wide are investing in HPC and some, most notably the USA, China and Japan are investing vast sums of public money on related infrastructure. In Europe, there has been an established HPC community for more than two decades. The first HPC centres appeared in Europe in the mid to late 1980s and these centres have cooperated in many ways over the years. Much of this cooperation has been focused around scientific and industrial collaboration (some financially supported by the European Commission and industry), and involvement in international standards (for example the development of MPI and OpenMP involved a number of European HPC experts). Europe has an acknowledged strength in HPC applications, and has many world leading Independent Software Vendors (ISVs)

HPC is a key enabling technology in which all major economies are investing.

Design and manufacturing are central drivers of the European economy. To maintain our key positions on the world stage in areas as diverse as the automotive and transport, pharmaceutical, financial, biological and renewable energy sectors requires that Europe invests in its ability to model, to simulate and to process data. This will drive the technological advancements needed to develop our future products and services. HPC is a key enabler of scientific and industrial innovation today. It is quite clear that the global economies that invest in HPC are those that will, over time, gain the greatest competitive advantage and reap the largest economic benefits. Hence the current scramble to invest in large-scale leading-edge HPC systems worldwide. However, the economic success or failure of this approach is not necessarily related to who has the fastest supercomputer in the world. Of much more importance than being number 1 in the top500[6] list are applications that can deliver the most valuable results. Europe has always had a key strength in such applications, which it must now exploit and refresh.

HPC is a key enabler for making Europe's products innovative and of high quality at competitive prices.

Europe's strengths in HPC

European engagement in HPC has been strong for the past 20 years. Although a number of European HPC vendors have come and gone in this period, many HPC centres, largely based in Universities or Research Centres and established in the late 1980s and early 1990s, have prospered and grown. Most of these centres today exist to serve the European scientific community although a few also engage on a regular basis with European industry. These centres have excellent expertise in the provision of HPC applications and services.

Europe has a history of leadership in HPC going back over 30 years.

Furthermore, Europe has a strong tradition of commercialising the modelling and simulation applications originating from our Universities. These are nowadays sold by ISVs. Many of our industries rely on applications owned and developed by ISVs. However, this process of commercialisation has also made it difficult for many ISVs to respond to the rapid changes we have seen in the scale and complexity of HPC systems.

There has been good take-up of HPC in some industrial sectors in Europe, and there are many commercial success stories. Opportunities exist for further exploitation in new application areas.

---

[6] http://www.top500.org

In terms of infrastructure, at the top level of HPC the PRACE and DEISA projects have, for the first time, led to a degree of harmonisation within Europe. They have established the concept of HPC as a European infrastructure which European scientists and industry can access for doing research, irrespective of what is provided by their company or their country at a national level.

Europe has a tradition of commercialising and using HPC.

The commercial use of HPC has been growing rapidly for the past two decades. This growth has come not only from increased use within the traditional sectors of aerospace and automotive, but also from new sectors such as finance and biotechnology. It is economically important therefore that Europe has a clear strategy for investment in all aspects of HPC within the timeframe of Horizon 2020.

The challenges facing European HPC

To reap the benefits that HPC brings in terms of new or enhanced products and services requires investment both in industrially relevant HPC applications and in models for their delivery to business and research.

A key challenge relates to the need to modernise ISV codes. Europe has many ISVs that have established themselves as world leaders in their fields. However the commercial pressures of running their businesses with limited user bases and revenue streams can make it difficult for ISVs to innovate, particularly if they are small companies. It must be made easier and less risky for ISVs to invest in long term R&D into new modelling and simulation techniques, while at the same time expanding their user base into new application areas. There is a major opportunity for both the software suppliers and users if the right conditions can be created.

We need to modernise existing software and exploit new applications.

A second key challenge is the need to exploit HPC in new applications and services and ensuring that a clear route to market exists for them. An innovative HPC applications sector will stimulate innovation by Europe's businesses. If Europe does not address this, then gaps in the market will either persist to the detriment of business and society, or they will be filled by our competitors, weakening Europe's global position.

A wider user base must be established (in particular SMEs) with access to reliable and affordable processor capability via reliable, high bandwidth connections. The costs of running and procuring HPC systems must be lowered, and training must be provided.

Using HPC must be made significantly simpler.

HPC use must become simpler. This third key challenge on a first glance appears technical, in reality however it has a strong economic dimension in the context of high labour costs. Programming models and tools must be improved, if not revolutionised, to enable application developers to use their skills more efficiently to exploit parallelism, or to port and optimise a code from one system to another.

We need to move beyond the current situation and bring together experts from industry (on the supply side as well as on the services and user side), from the mathematical and physical sciences, and from informatics and computational sciences to take a clear step forward in Europe. No single community can do this alone – it must be a joint effort. The forthcoming PlanetHPC Roadmap will to propose how we overcome these challenges.

The key industrial and societal applications: established and emerging

This section describes some application areas where HPC can have a major industrial and societal impact. Many of these application areas already have a history of successful use of HPC, and others are new or emerging. These examples show that HPC is an enabling technology which can yield huge benefits for Europe's economic competitiveness, the wellbeing of its citizens and its progress towards a sustainable future.

Design and manufacturing

Manufacturing is still the driving force of the European economy, contributing over €6,500 billion in GDP and providing more than 30 million jobs. It covers more than 25 different industrial sectors, largely dominated by SMEs, and generates annually €1,500 billion of added value.[7]

HPC can contribute to the vision of a high value, high quality manufacturing industry in Europe.

HPC has revolutionised the design and manufacturing of a wide range of products. It can not only help produce better products, but can also dramatically reduce the time and cost of design and production. These benefits have been achieved by HPC simulations of the finished product with a reduced need for building prototypes. A CAD model is created and the physical conditions for the finished product (for example the fluid flow in a turbine or pump) are simulated numerically. High performance computing is required to produce a result with the required accuracy. Virtual reality applications (sometime called "life experience") are being used to allow designers to explore and interact with virtual realisations of their products and systems. Real-time HPC is essential if these applications are to give the required fidelity and responsiveness. These applications allow huge cost savings to be achieved by removing the need to build prototypes for testing and by optimising for ease of manufacture.

HPC therefore enables higher-value products to be made and brought more rapidly to market whilst avoiding prohibitive development costs. It allows better product definition and reduced risk through better up-front analysis and decision-making. The continued ability to improve manufacturing designs and processes is essential for Europe's competitiveness to be maintained. There will be increasing demand for higher-fidelity simulations and faster throughput of simulation results in order to stay ahead of global competition. The following sections outline the importance of HPC in key industries and in the factories of the future.

Investment in HPC will help the EU maintain a strong market position.

Automotive

HPC is widely used in the automotive industry for designing efficient and safe vehicles, and many opportunities lie ahead as the industry moves towards alternatives to fossil fuels.

Aerodynamics has become an established part of vehicle design. Reducing the drag coefficient can greatly increase the fuel efficiency of vehicles, making them cheaper and more attractive for consumers and at the same time reducing

[7] See http://ec.europa.eu/research/industrial_technologies/fof-facts-and-figures_en.html

CO_2 emissions. This has been achieved by using simulation software, mainly from commercial third parties, running on HPC machines.

There is an international move towards electrically powered and hybrid cars, with efforts such as the European Green Car Initiative carrying out related R&D. HPC will be an essential technology in the development of new batteries and motors that will enable electric cars to become a reality.

HPC has also been used to design cars which offer greater protection for drivers and passengers. The costs of designing for safety have been greatly reduced by the use of HPC, since the use of computer simulation reduces the number of expensive physical tests that need to be carried out.

The automotive industry in Europe will continue to be driven by market forces which demand reliable, efficient and safe vehicles at competitive prices. Competition from producers elsewhere is intense, and hence Europe must use and invest in technologies such as HPC which will help it maintain a strong market position.

Automotive applications use the full spectrum of HPC systems: embedded, clusters and supercomputers.

The automotive industry foresees a need for greater computing power on all levels, for applications such as simulating combustion processes in engines, improving the crash worthiness of vehicles and their fuel economy. Key technical challenges will include generating meshes to model geometries in sufficient detail, and the need for coupling different simulations to understand the interactions between them. Modelling the full process chain in engineering consists of different levels of abstraction with different characteristics in terms of granularity, inherent parallelism and locality of data as well as connectivity requirements, e.g. business processes, multiphysics and multiscale coupling of disciplines, parameter studies, stochastic models and mesh partitioning.

In addition to market forces there are pressures from policy makers, such as the need to develop transport that does not rely on fossil fuels and is low in CO_2 emissions. HPC has an essential part to play in the research required in these areas.

Aerospace

The aerospace industry in Europe has been a major user of HPC where it is a key enabling technology. Major companies such as Airbus, Rolls-Royce and BAE systems, as well as smaller companies in the supply chain are experienced users of HPC.

The success of HPC can be seen by the great improvements in fuel efficiency of aircraft in the last few decades. Use of simulation tools has improved aircraft fuel efficiency dramatically; the fuel used per passenger mile today is around 30% of what it was 40 years ago. The industry aims to reduce fuel consumption by a further 25%. The Airbus A380 uses around 3 litres of fuel per 100 passenger km, which makes it comparable with a family car carrying two occupants. The A380 has been designed as a low-noise aircraft, which has been achieved through acoustic simulations using HPC.

The Advisory Council for Aeronautic Research in Europe has a vision for 2020[8]

[8] http://www.acare4europe.org/docs/Vision%202020.pdf

which sets ambitious goals that deal with a considerable reduction of exhaust gas (NO_x 80%, CO_2 50%) and noise (10-20 dB). Air traffic is expected to increase by a factor of 3 and accidents should go down by 80%. Passenger costs should drop by as much as 50% and flights must become available in all but the most extreme weather conditions.

In aerospace the key challenges for HPC are related to real-time requirements, complex multiphysics models, and optimisation requiring several orders more of computing power.

To achieve these goals the industry needs the real-time simulation of a full aircraft in flight. The computational power needed to cope with this is estimated to be a factor of 10^7–10^8 higher than today's capability. Only a major increase in capability will enable the European aircraft industry to develop the necessary technology and optimise their products from an economical and sustainable point of view. The demand for compute power will grow further with the increased use of numerical optimisation in the design process, which requires many simulations using different parameter sets. On the positive side, those "ensemble simulations" lend themselves perfectly to the massively parallel systems that are expected to be available by the year 2020.

To be able to design aircraft engines that fulfil the envisaged reductions in air pollution and noise, high-fidelity, multi-physics simulations will be required. The main challenges here will be the development of accurate turbulence models and efficient numerical methods for the coupled simulation of fluid, structural, thermal and chemical effects. Because of the moving geometry, many aspects can only be studied with computationally intensive unsteady simulations. Achieving this will require access to systems with the necessary computational power together with software capable of scaling to utilize that power. Today's software is not able to do this; major investment is required into algorithms and modelling techniques for the goals to be realized.

Many space research applications greatly benefit from HPC technology. In "classical" areas, such as the simulation of atmospheric re-entry of space vehicles or the gas flow in rocket motors, the development issues are very similar to those of aeronautics as outlined above.

Another area where HPC acts as an enabling technology is the processing of large-scale data streams received from satellite payloads. The astronomy mission Gaia, for example, scheduled for launch in 2012, will measure the positions, distances and proper motions of one billion stars in our galaxy, totalling one trillion single observations. Key elements of the data processing are a large object-oriented database and HPC processing. Although Gaia will observe 500 times more stars with 100 times better accuracy than its predecessor mission Hipparcos in the early 1990s, it will still only cover one per cent of the stars in our galaxy.

A topic of strategic relevance is the cataloguing and removal of space debris in Earth orbit. At ESA and in several national European space programmes, related initiatives are in their early phases. Observations of debris objects as small as one centimetre will be made, both with radar and laser technology. About 600,000 such objects are estimated currently to be in low Earth orbit. HPC technology will be used to determine orbits and process databases. In terms of HPC, this is not a high-end application, as most computations can be performed on moderately-sized clusters, perhaps including GPGPU accelerator cards. Somewhat more demanding in terms of compute power will be the removal of space debris using ground-based laser devices, which is expected to be feasible in about ten years.

Manufacturing in the future

The Factories of the Future (FoF) Public Private Partnership (PPP) recognises the challenges to manufacturing in Europe and the need for more innovation, describing the situation as follows:

> *"The European manufacturing sector faces an intense and growing competitive pressure in global markets. European companies are faced with a continuous competition in the high-tech sectors from other developed economies, such as the US, Japan and Korea.*
>
> *Manufacturing has to address the challenge of producing more products with less material, less energy and less waste. Our living standards are on the rise; global manufacturing today has to meet a constantly increasing demand for consumer goods. Manufacturing has to improve its innovation activity. New ideas have to be transformed into new products and processes. Many of the manufacturing companies are SMEs and only a few of them have research capacity and the financial potential to implement high-risk innovative manufacturing technologies[9]."*

European manufacturing will rely heavily on investment in HPC.

The FoF PPP strategic multi-annual roadmap[10] identifies modelling and simulation as key technologies for designing products and processes if these challenges are to be met. Such simulations will require HPC to achieve the required degree of fidelity and speed. Technological challenges that must be met include adapting existing applications or developing new ones for various product types and integrating HPC into the workflow and production process.

As many manufacturers are SMEs they may be unable to risk investing in application software and HPC systems until the technology is proven. Therefore programmes that encourage small companies to evaluate and adopt HPC-enabled modelling and simulation by reducing the barriers will be required, for example by providing low-cost or pay-per-use access to applications, HPC systems and services.

Services and utilities

HPC has the ability to transform the service and utility sector by enabling applications and providing the necessary computing power to manage and control complex systems. Many of the applications in this sector require HPC to be accessible as a robust technology that is constantly available without downtime and able to deliver results in real-time. Meeting these requirements will be critical if these applications are to keep pace with the rest of the world.

[9] See: http://ec.europa.eu/research/industrial_technologies/fof-challenges-ahead_en.html

[10] See http://ec.europa.eu/research/industrial_technologies/pdf/ppp-factories-of-the-future-strategic-multiannual-roadmap-info-day_en.pdf

Energy

HPC will help make the most of existing energy sources, and will be vital for developing clean and renewable energy.

Efficient use of existing energy sources such as oil and gas and the development of alternative sources which are sustainable and low in CO_2 emissions are high priorities globally. HPC is a key technology which can contribute to this goal. Europe can show leadership by investing in HPC to help secure clean future energy sources.

The discovery of oil and gas relies on HPC to investigate the structure of rock formations by analysing reflected seismic waves. Processing software is used to analyse seismic data. Many companies have invested in developing such software for in-house use or licensing to other users. Greater HPC performance is needed to improve the accuracy of the analysis and hence produce a better understanding of geological structures and of the likely location of oil reserves.

Modelling and simulation using HPC is also vital to the oil production process. The ability to simulate the behaviour of oil reservoirs allows different strategies for producing oil to be evaluated so that the best scheme can be selected and implemented. The effects of injecting water into the reservoir can be simulated, and the positioning of wells can be planned for optimal extraction of oil. Software for oil-reservoir simulation tends to scale modestly with perhaps tens of processors being the limit. Small percentage improvements in the way that oil is extracted translate into huge revenues. More performance is needed to improve the accuracy of the simulations so that better use is made of existing reserves. This will require more powerful HPC systems and research into simulation techniques which scale better.

HPC offers great potential in the emerging application of real-time management of electricity networks[11]. Many European countries are planning or beginning major asset replacement and renewal programmes with regard to electricity distribution networks. For example UK Power Networks is committed to major UK distribution network asset replacement and renewal over the next five years. Such an initiative represents an ideal opportunity for the massively distributed deployment of state-of-the-art network metering, monitoring and instrumentation.

The HPC architectural requirements that must be met to support this type of application are different from the off-line, data-centre approach of many engineering applications. Robust, real-time HPC systems will be required to perform data analysis and data mining to support electricity-distribution management systems. Major research and development for a high-performance ICT infrastructure will also be required to exploit fully the return on investment of such new assets and to defer further asset deployment. Future potential financial benefits and cost savings to large-scale distribution network owners, operators and customers can be measured in hundreds of millions of Euros.

In order to provide suitable infrastructures that will support electric vehicles, the generation, distribution and use of electricity must be modelled and optimised, with all actors – from generation plants and energy distributors to households and vehicles – integrated in real-time models. To address this challenge HPC resources are indispensable.

[11] See: http://www.hiperdno.eu

Digital media

The advances in computer generated images (CGI) which have been possible in the last two decades have been enabled by, among other things, availability of affordable processing power. The first feature-length film to be created entirely using CGI was made in 1995.

A single frame of a CGI movie takes around 6 hours to render on a single processor, with more complex scenes over ten times longer. As computing power has increased, the time to render a frame has not reduced; instead the image quality has become much greater. The use of specialised Graphical Processing Units (GPUs) is widespread. In fact the development of GPUs has been driven by the rendering needed for computer games. In addition, frames can be rendered independently, making this an embarrassingly parallel problem. Rendering is carried out by 'farming' the individual frames to a cluster of processors. Render farms used by film companies are often run in-house or with trusted partners rather than in public clouds (confidentiality of data is of extreme importance in this industry). A consequence of this approach is that a reduction in time to render an individual frame can only come about as a result of faster processors, not more processors. Rendering on a multi-core processor will not be faster than on a single core running at the same processor speed.

Although the major centres for cinema film production are outside Europe (USA and India) there are global opportunities for production companies and a thriving European industry already exists, for example in Soho in London. Development in Europe of HPC systems and expertise suited to this kind of workload and addressing the non-functional aspects (such as security and robustness) could place Europe in a strong position to attract business and become a world leader.

Financial services

HPC has become an indispensable tool in the financial services industry where it is used to carry out complex calculations to estimate the value of assets, pricing of products and assessment of risks.

Due to the fast-moving nature of financial trading, these applications require real-time response to streams of input data together with high reliability and security. Many applications are developed in-house integrating with third-party software such as databases.

Robust, reliable and secure HPC systems are the key requirements for HPC in this sector.

Quality of life, wellbeing and sustainability

HPC has the potential to improve radically the quality of life of EU citizens, for example by enabling better healthcare, greener transport and safer, more energy-efficient buildings. Applying HPC to these areas will result in economic and social benefits.

HPC can dramatically improve healthcare and the diagnosis and treatment of diseases.

Healthcare and medicine

Medical science is increasingly turning to computational models to study the possible effects of drugs and surgical interventions, before moving on to patient trials. In Europe, the Virtual Physiological Human[12] (VPH) project is carrying out pioneering research into biomedical modelling and simulation of the human body. This will improve our ability to predict, diagnose and treat disease, and have a dramatic impact on the future of healthcare and the pharmaceutical and medical device industries. HPC is essential in order to simulate the many complex interactions on multiple scales. Research is needed into how to integrate computer models and extract knowledge from the results. The medical sector promises both greater wellbeing for citizens and economic benefits for the companies supplying the industry. HPC has the potential to enhance the lives of millions of European citizens by allowing better drugs, therapies and diagnostic techniques to be developed. Prevention, diagnosis and treatment of disease not only enhance the quality of life of individuals, but can also reduce the costs of long-term care and may prolong the active and working life of individuals. When applied to the population at large, these economic benefits can be substantial.

In biomedical applications, similar abstraction levels as in engineering with different characteristics in terms of granularity, inherent parallelism and locality of data as well as connectivity requirements can be identified. Examples include disease and ecosystem models, system biology and coupled metabolic-physiological models, Bayesian or stochastic network models, analysis of large data sets and optimisation of individual bioinformatics applications.

Efficient transport

HPC has been used in Intelligent Transport Systems (ITS) for the simulation of traffic flows to enable the better design of road networks. Modelling the individual behaviour of vehicles on a large scale is an application well suited to HPC systems due to its inherently parallel nature. This is a key element in the design of road systems to reduce congestion, which has significant economic and environmental impact.

More efficient and greener transport can be achieved using HPC.

Road transport is becoming much more complex in terms of in-vehicle intelligence and roadside sensors. The convergence of technologies for location awareness and mobile communications presents an opportunity to manage traffic on a regional or even national scale to reduce congestion, journey times and fuel consumption, and hence make more efficient use of the existing road networks. Building new motorways can cost up to €100m per mile and has a major impact on the environment. Therefore making more efficient use of existing transport assets is imperative across Europe. Doing this at a regional or national level requires massive computing resources.

HPC, together with other technologies, can enable this, but only if it can be better integrated with the sensor-rich and inherently distributed reality of traffic systems. The systems which will manage road traffic in the future will require HPC that is able to produce real-time results and be robust and highly available. Down-time will not be an option in managing road networks. Investment will be needed to ensure that HPC has the required characteristics and can be integrated with other technologies.

[12] http://www.vph-noe.eu/

As described earlier, the vision for the future of personal transport is electric and hybrid vehicles. In addition to developing the vehicles, an infrastructure to support them must be designed, built and managed. This is a massive undertaking that will require HPC. Simulation and design of the electrical systems required for recharging batteries will require new applications to be developed, and managing the resulting network will require robust and reliable real-time systems.

Safe and efficient buildings

HPC is used in the construction industry for the design of buildings and systems. Modelling and simulation have helped to improve the energy efficiency and safety of buildings.

To make a building safer in the event of fire, designers want to implement systems which can aggregate data from a number of sources (for example smoke detectors, CCTV, air conditioning systems) and combine them with the results of simulations (for example to predict how the fire might spread, and using simulations of human behaviour as an evacuation proceeds) to produce and evaluate scenarios for tackling the fire and mitigating risks to people and property. Advanced structural mechanics is required to analyse how buildings are affected by the fire, while the spread of the fire can be modelled using computational fluid dynamics.

Emergency response

Numerous emergency response scenarios can be envisaged where rescue and relief operations require decision support. Examples include major road incidents, floods and fires. Faster than real-time simulation of the systems that need to be controlled is needed so that response strategies can be rapidly evaluated and the best plan of action put into place. Emergency teams can be trained by creating virtual scenarios and learning how to tackle them, thereby ensuring that the teams are better equipped to react to a real situation.

These are applications where on-demand HPC will be necessary, as they will be characterised by infrequent, but very intense use of computing power. Having systems on stand-by waiting to be used is not likely to be cost effective.

Why Europe needs to act now

Many of the achievements and benefits described in the previous section have been made possible by the remarkable advances in HPC over the last 20 years. The computational capability of the world’s fastest computers has increased by a factor of a million in that time, and machines which are routinely used in industry and academia today are equivalent to many times the power of the fastest system of a decade ago. Few, if any, other technologies can claim this kind of progress, which has been underpinned by the ever increasing performance of the basic commodity components (microprocessors, memory, interconnect and storage) coupled with the adoption of standards at all levels including programming models. Year-on-year, users of HPC have been able to re-use their parallel codes on machines with ever larger numbers of faster processors, achieving ever greater performance.

This continued progression is threatened by disruptive technologies such as multi-core and GPGPUs, which require increased effort and system-specific skills for porting and optimising application codes. However, Europe is well placed, if it acts now in a concerted fashion, to emerge from this period of disruption as a leader in the application of HPC technology.

Technology disruption

Technology disruption is a threat to HPC progress.

Industry has come to rely on a continual increase in computing performance accompanied by a lowering of cost. However, this scenario of continued exponential growth is facing disruption due to a number of factors. Firstly, the effect of Moore's Law must come to an end at some point which will limit the performance of individual components. Secondly, the electrical power consumption of a Petaflop/s system today is several megawatts; this implies that an Exaflop/s system built using today's technology would consume an unfeasibly large amount of energy. Thirdly, many applications are reaching the limit of scalability on today's large systems which have tens or hundreds or thousands of processor cores. An Exascale system may have tens of millions of processor cores. Although Petaflop/s systems are at the leading edge today, all previous experience suggests they will be commonplace departmental-level systems by the end of this decade. However there are few academic and no industrial applications that operate today in the Petaflop/s regime.

If Europe acts now, it is well placed to emerge from this period of disruption as a leader in HPC technology and its application.

These factors are already forcing changes in the ICT industry which will impact on HPC. Multi-core processors are now ubiquitous. Processors with 2 or 4 cores became commonplace in the mid-2000s, and higher core counts are now available. This trend is set to continue, with the expectation among many industry analysts that processors with hundreds or even thousands of cores will be the norm by 2020. Graphics accelerators are becoming increasingly popular for performing vector-type floating point operations at low cost and energy. Programming such devices efficiently presents challenges in terms of the scalability of applications and the restrictions on memory and interprocessor bandwidth which make locality of data an important consideration.

New business and access models

New access and business models for compute resources and application services are evolving in the enterprise computing field as a result of cloud computing. This has raised expectation among some observers that the same will happen for HPC. As noted earlier in this report, there has so far been little impact on HPC, due to significant technological barriers. This may change and may have an impact on the market if the technological barriers can be overcome.

We therefore stand at a crossroads in HPC. The future of high performance computing will not simply be a continuation of the past two decades. Components, architectures, programming models and applications will need to undergo changes, potentially revolutionary ones, if systems based on millions of processing elements are to be effectively used.

Europe must invest in HPC to continue to reap the proven benefits.

Unless action is taken, the societal and economic benefits of HPC and computing in general will stall. The leadership that Europe has shown in many domains of computing and its applications over the last 20 years

must be maintained and expanded if the European economy is to remain innovative and competitive. Investment in research, development and training carried out at the European level, together with measures to promote the early industrial and commercial uptake of new technology will be essential elements in a successful strategy, in which HPC is one of the key drivers.

A European opportunity

Investment is needed now to ensure that Europe is able to meet the challenge to its competitiveness from emerging and established economic powers that will exploit HPC. Among these established economic powers, the USA has always made major investments in HPC with at least half of the current Top500 systems located in the country. Japan once again has the world's most powerful computer, the 8 Petaflop/s K computer, following an investment program of over 100bn yen (around €1bn). Among the new economic powers, China and India have ambitious HPC programmes.

The disruption in technology must be viewed as an opportunity for Europe to invest for its future in technology development, facilities provision, education and promoting innovative use of HPC. With the changing landscape of HPC, the economies which adapt the fastest will be the ones that will gain the greatest benefits.

Key European strengths

Europe has key strengths in HPC which must be reinvigorated and exploited.

Europe has key strengths that make it well positioned to implement an industrially relevant HPC strategy:

- Europe has expertise in low energy devices and in embedded computing technology.
- Many world-class ISVs are European.
- There is established use of HPC in many industries and proven success stories.
- A rich set of HPC centres able to work in collaboration has been established.
- Europe has many trained experts and education programmes.
- Initiatives such as DEISA and PRACE give Europe a head start in supporting the research communities.

Weaknesses

Europe has fallen behind in terms of HPC investment.

- According to the report by IDC[13] Europe's investment in HPC has fallen behind that of many competitors in recent years.
- There is a skills shortage in HPC, and a reliance on those who have skills in both HPC and an application domain.
- Many of the companies which can benefit the most from HPC are small

[13] See http://www.hpcuserforum.com/EU/downloads/SR03S10.15.2010.pdf

and cannot risk making the investments needed. This includes, in particular, Independent Software Vendors.

- Europe has few hardware providers and European system vendors only have small market shares.

Opportunities

Technology disruption will demand changes to the way HPC is used in the future. Europe has an opportunity to lead the way.

- The technology disruption underway will require new programming methods and tools, and must lead to a modernisation of applications. Europe has strengths in programming methods and tools and can lead the way in this area.

- HPC and embedded technologies are converging. There are opportunities for technologies for cross fertilisation between these distinct application domains. For example, the HPC community can profit from the expertise of the embedded-systems community in making systems more energy-efficient, dependable and ready for time-critical applications. The embedded computing community can learn from the HPC community how to better exploit parallelism.

- With more flexible access, simpler use, and affordability, more companies will be able to use HPC to support business decisions in a way that hasn't been possible before. In particular, there are a great many SMEs across many application domains, for which HPC and HPC-based applications remain out of reach. There is huge potential for Europe to exploit this in terms of reinforcing the competitiveness of our SMEs. Their use of HPC must be indirect through their application codes and services.

- Mobile computing is widespread as we move to a situation where people and devices are almost always connected. Ten years ago we were sometimes on-line; now we are occasionally off-line. By 2020 it is likely that almost everything and everyone will be permanently on-line. This opens opportunities for the ICT market in general and HPC in particular.

- There is a growing acceptance of cloud computing as a model for service delivery. If this could be replicated in the HPC domain, it would open up the possibilities of using HPC, in particular for small companies, by lowering the risks associated with adopting HPC. The business model and rationale underpinning cloud computing (i.e. access to a reliable service without high capital and operational costs) are still some way from being supported by current technology with respect to HPC. Finding the correct balance between ease of use through virtualisation, and attaining good performance must be found.

Threats

- Established and emerging competitors are increasing their investment in HPC and developing HPC expertise, which will enable them to get the benefits of HPC sooner than Europe and so take a competitive edge.

- Increased spending on infrastructure by Europe's competitors will enable their economies to take up HPC much faster. They will be able to out-compute and out-compete.

- Countries outside Europe are providing users and developers with low cost access to HPC capability. Europe must ensure it is able to compete globally, but must do so without distorting the market.

- The skills shortage will be exacerbated if the best HPC career opportunities exist outside Europe – there will be a 'brain drain'.

- Investment must be targeted at the areas where it will have maximum market impact. Focussing only at the highest or lowest level of HPC carries the risk of market failure. Due to the convergence of computing disciplines, HPC investment must be seen in the context of overall investment in R&D&I in computing and its applications in general, including synergies across disciplines.

Investment by our global competitors is a threat to Europe. We must avoid a brain drain.

Key challenges and research issues

The performance of high-end HPC systems has, to a close approximation, increased by a factor of 1000 in each of the last two decades. With the world's most powerful systems already past the Petaflop/s level, a projection of this trend leads to the prediction that by 2020 Exascale computing will be possible, and perhaps even commonplace.

Energy consumption and limits to processor speeds are major issues that will lead to new computer architectures.

However progress towards this is threatened by two related factors:

- Based on current technology, systems with Exascale performance will use excessive amounts of energy.
- The performance of processing elements can no longer be assumed to follow Moore's Law and double every 18 months to 2 years due to physical constraints.

There are market drivers which provide an incentive for developing lower-power devices. Energy consumption is a recognised problem applicable from data-centres down to hand-held devices. Progress in lower-power devices may make it feasible to build future systems with Exascale performance and acceptable power usage by using millions of processing cores. However this provides a challenge of its own: how do we exploit the new computer architectures that will emerge from this disruptive technology?

There is a strong push by providers of generic cloud computing resources to enter the HPC market and offer HPC resources on demand. However the current services are not truly HPC according to the definitions used in this report and widely accepted in industry. It is possible that some of these barriers may be overcome in the coming years and a pro-active approach will help European actors to direct and make early use of any such advances. It remains a widely held belief across the HPC community that it will remain impossible to deliver true HPC through cloud computing except in some special cases.

Many HPC applications are so-called legacy codes that have been developed over many years following a model of HPC which has been stable during that period, but which may not persist for much longer. Most of these legacy applications cannot be changed quickly, easily or cheaply. The users of these applications have enjoyed year-on-year improvements in performance as that of microprocessors has improved and HPC architectures have remained relatively unchanged. Massive investment will be needed to migrate or re-implement these applications to exploit future HPC architectures. In particular, research and development will be needed to exploit the high levels of parallelism (millions of threads) that are anticipated in future systems.

New applications generate new requirements such as the need for robust and reliable HPC.

Applications are emerging which will present new requirements for HPC. Robust, real-time or even faster than real-time and reliable HPC is required for applications which deal with constant operation and monitoring of infrastructure such as electricity generation and distribution, transport control, industrial processes and decision support in crisis situations. Furthermore, the ability to deal with enormous amounts of data from numerous diverse sources such as always-connected sensor networks is a clear requirement for HPC in the future. In complex forecasting applications, which build on modelling complete process chains from designing and manufacturing to maintenance and business processes, HPC applications have to be integrated with management and services applications, which traditionally run on server farms.

The following sections outline challenges not only at the technological and applications levels, but also those challenges that must be overcome to ensure the rapid and effective take-up of HPC solutions, which will be a vital component in Europe's HPC strategy.

Technologies

Computing must become more energy efficient.

A key challenge for the future is to develop low-energy systems and components, and use them effectively within HPC systems. Much work is going on in the embedded systems field to develop low-energy processors, driven by the requirements of the hand-held market. HPC systems must leverage these developments by designing architectures which can use such components. Tools such as energy-aware compilers and run-time environments are envisaged.

Multi-core, many-core and heterogeneous systems must be mastered.

The trend towards systems with millions of processing cores presents a major challenge. Alongside this, the likelihood seems to be that systems with heterogeneous processing cores (such as GPUs today, and perhaps something else in the future) will also be prevalent. Programming models, languages and software engineering methodologies must be found to manage and exploit this massive parallelism and heterogeneity. This must go hand-in-hand with research into new algorithms and numerical methods.

New programming models must be developed and mastered, supported by tools to improve developer productivity.

Robust and reliable real-time HPC will be required.

The ever increasing number of components in HPC systems makes them more prone to failure. Most HPC systems of today have little or no tolerance to the failure of single components. Future systems will have to exhibit much greater degrees of fault tolerance, and be able to reconfigure themselves to compensate for failing components. It may be necessary for applications to be developed with fault tolerance in mind.

Application developer productivity must match progress in enterprise computing.

The productivity of developers in enterprise computing has increased dramatically in recent years thanks to the adoption of standards and the emergence of effective development tools. This has also been the case in embedded systems; toolsets from companies such as Codeplay and Critical Blue have enjoyed widespread take-up. The HPC community must learn from this, and achieve better productivity. It must be possible for application experts and software experts to work together to generate applications rapidly and efficiently. This is not a challenge that can be solved by tools alone; very few in the HPC community consider the development of fully automatic parallelisation tools to be feasible.

Applications

New methods for performing simulation and modelling need to be developed which can exploit parallelism and heterogeneous computing on a massive scale.

Major investment in applications is needed.

Innovation must not stop at the level of porting existing software to new hardware. There must be research into new methods of modelling and simulation, and new numerical methods and algorithms which must take account of massive parallelism, potentially unreliable systems and heterogeneous architectures.

For industry to gain the most benefit, HPC applications need to be coupled with each other, and with other parts of the workflow.

Research is needed into how HPC simulations can be coupled effectively to simulate multiple phenomena. This is well illustrated by the automotive and aerospace industries where there is a clear need for multi-physics simulations. HPC must also be integrated better with workflows and across the supply chain. Standards will play an important role in this.

Data-intensive applications will emerge as an important domain as we seek to deal with the data-deluge.

Data holds the key to knowledge discovery.

With the volume of data from measurements, observations and the results of simulations rising at an exponential rate, extracting knowledge from data will become an ever more important issue. Research efforts must be directed towards collecting, storing and curating data and extracting knowledge from databases.

Accelerating take-up

SMEs must be encouraged to take up technology.

Many of the potential beneficiaries of HPC (particularly new users) are small end-user companies that cannot afford to take on the risks of adopting new and unproven technology. Programmes must be initiated which will give these companies access not only to HPC infrastructure but also to HPC expertise. It is important to understand here that for user SMEs, HPC is just an enabling technology for using new applications. In that sense HPC needs to be marketed indirectly through applications and services that use HPC.

New access and business models must be explored.

Cloud computing is changing the landscape in many commercial and societal applications. The HPC community, in particularly software and service providers must exploit the opportunities this period of change is bringing. For example, new licensing schemes as part of pay-per-use business models for HPC-based application services will open new revenue streams for software owners and service providers, while at the same time bringing HPC applications to large groups of actors, for whom HPC is currently beyond their reach.

Migration pathways for legacy applications must be found.

Many companies are reliant on legacy applications which represent huge investments and are difficult to change. Migration pathways, either for porting existing applications or for operating them side-by-side with new applications must be provided. Businesses must be able to operate effectively through the technological changes that lie ahead.

SMEs must be given incentives to take up HPC.
Software vendors must see benefits in innovating for HPC.
The shortage of skills must be addressed.

Encouraging innovation by Independent Software Vendors.

Many ISVs are small companies which cannot afford to invest in major restructuring or rewriting of their software. They must be encouraged to innovate and participate in co-design with system developers and end-users to ensure knowledge transfer which will lead to systems and applications which together deliver high-quality results and excellent performance.

Raising awareness of the changing technology landscape.

Many end-users are simply unaware of the potential technology disruption that is pending, and believe that it will be business as usual for HPC in the future.

HPC needs to break the dependence on dual expertise (HPC and domain specific).

Furthermore there is an inflexible dependence on experts who need to be skilled both in HPC and the target domain application. Models of collaboration must be found to break this dependence.

A vision for HPC-driven innovation in Europe

The preceding sections of this report have identified the domain, applications, opportunities, key challenges and research issues relevant to HPC driven innovation in Europe today. This section draws these threads together and proposes a concrete plan that will lead to increased innovation in Europe through the application of HPC.

The broad vision

The European economy relies on continuous innovation in industrial and commercial products and services for growth. As this document has described, in a wide range of domains this innovation can be supported through the use of HPC-enabled applications. By investing in accelerating the uptake and use of HPC, the European Commission can have a direct and positive impact on the products and services offered by European companies and hence create economic growth.

Meeting the challenges requires long-term, focused support that should start immediately.

In order to realise this vision we must tackle the interlinked challenges of:

- Developing new programming models, languages and methodologies for future HPC architectures and tools to increase developer productivity.
- Finding ways to build and use HPC systems which are energy efficient.
- Adapting existing modelling and simulation techniques, and developing new ones, so that they scale to massive degrees of parallelism.
- Creating an environment which encourages take-up of technology by involving end-users, providing access to infrastructure and thereby lowering risks.
- Exploiting the convergence of computing domains (for example mobile, embedded, HPC, distributed computing and sensors).

Meeting these challenges requires long-term, focused support. In the next sections we outline a starting point for this support through immediate action within Framework Programme 7, together with the outline for a longer-term programme of research and development.

Synergy and convergence of ICT technologies

Investment in HPC must take account of synergy between various usage models of HPC, and the convergence with other technologies such as mobile communications, embedded processing and sensors. Some of the challenges facing HPC are common across different ICT domains. The need for greater energy efficiency applies across all computing fields; supercomputing centres and data centres face huge electricity costs, while at the embedded systems level low power devices are needed to extend battery life. Innovation in this area is needed at the large installation level, where designing efficient cooling can reduce costs, and at the component level where devices are being developed that can manage their power consumption intelligently.

Applications in all areas of ICT demand increasing levels of dependability and robustness. In this context the HPC community can learn lessons from both management and services systems and embedded systems.

The widespread availability of mobile computing opens up great opportunities for HPC. Users are able to interact with HPC from anywhere with many types of device, and there is the potential to integrate data from millions of sensors. The scope for new applications is immense, particularly in managing and controlling infrastructures such as road networks and electricity grids. Systems used in this way must have exceptional robustness and reliability.

HPC actors

The effective use of HPC is often complex because of the diverse skills and expertise that are required to solve specific problems. This complexity has led to slow adoption of the technology by many companies, particularly SMEs, over the past 20 years. Many of the barriers to adoption we face today were apparent 20 years ago. The actors involved in the successful use of HPC include many, but not necessarily all, of the following organisations:

The use of HPC involves many different actors, each of which can benefit through collaboration.

- **End users:** the problem holders. These companies range from some of Europe's largest companies to some of the smallest SMEs. They are characterised by their need to solve or understand a problem through use of HPC. They may own and develop their own software applications or use applications developed by third parties. Many large companies have a mixture of in-house and third-party applications.
- **HPC service providers**: public or private providers of HPC system resources. The services provided go well beyond access simply to computing hardware, but include access to storage, development tools, management and accounting systems and take into account security, networking and data transfer considerations. This category includes cloud computing providers who offer HPC-like services.
- **HPC systems vendors:** companies that build and sell HPC systems.
- **Independent Software Vendors (ISVs):** companies that provide HPC software applications to solve specific problems or tools to support HPC development. Access to such applications is controlled through specific licensing terms linking the application to individual systems and/or users. Europe has considerable strength in this area.
- **Research software providers:** academic and research organisations that provide HPC software applications to solve specific problems or tools to support HPC development. Often such provision is a by-product of the scientific research being undertaken by the organisation. Access to such applications is often through Open Source licensing. Again, Europe has considerable strength in this area.

- **Application experts:** companies and academic organisations that can provide modelling and simulation consultancy support to assist the end users with their HPC use, in particular with regard to instantiating models and interpreting the results. This is often a highly skilled activity where very specific problems are well understood by the organisation's staff.
- **HPC consultancies:** companies and academic organisations that can provide parallel computing and supercomputing expertise to application owners (end users, ISVs and research software providers) to optimise and port applications between different HPC systems.

Each of these actors may have a specific role to play in any successful HPC application scenario. Many market players act in more than one of these segments.

Collaborations and interactions

As has already been noted, the use of HPC is complex and, particularly when a user is new to the domain, may require significant help and support in order to realise its benefits. It is also very unusual, despite claims to the contrary, for any single service provider to be able to provide all of the necessary components of a successful solution. Collaboration between actors is therefore vital for the success of individual projects.

Many scenarios for collaboration between the actors can be envisaged. Some are listed below, although this list is by no means comprehensive:

Scenario 1: Porting in-house code to a new HPC system.

An end-user which is also a code-owner wishes to port its code and run it on an HPC machine. It works with an HPC expert to develop the code, and an HPC-service provider to do production runs. The code owner retains IP and the HPC expert receives a consultancy fee.

Scenario 2: Porting ISV code to a new HPC system.

An end-user needs a commercial code to run on a new HPC system. The ISV works with an HPC expert to develop the code. The ISV retains the IP and the HPC expert and application expert receive consultancy fee or royalties.

Scenario 3: New end-user using an application.

An inexperienced end-user wishes to trial HPC for their business. The other stakeholders act to create a possible solution which is evaluated by the end-user.

Scenario 4: New functionality for ISV code.

In response to end-user demand and to reach new markets, an ISV adds functionality to its code. It works with application experts and HPC providers to develop the new code which it tests and tunes using HPC. The ISV retains the IP.

Scenario 5: New application development.

An application developer works with an HPC expert and a research software provider to develop the new application. They transfer the commercial rights to an ISV, or create a start-up.

Scenario 6: Reimplementation of applications for future systems.

An ISV realises that it must completely re-implement an application for future HPC systems. It works with the application experts, HPC experts and research software providers to develop the code, driven by requirements from end-users. All parties have an appropriate stake in the newly generated IP.

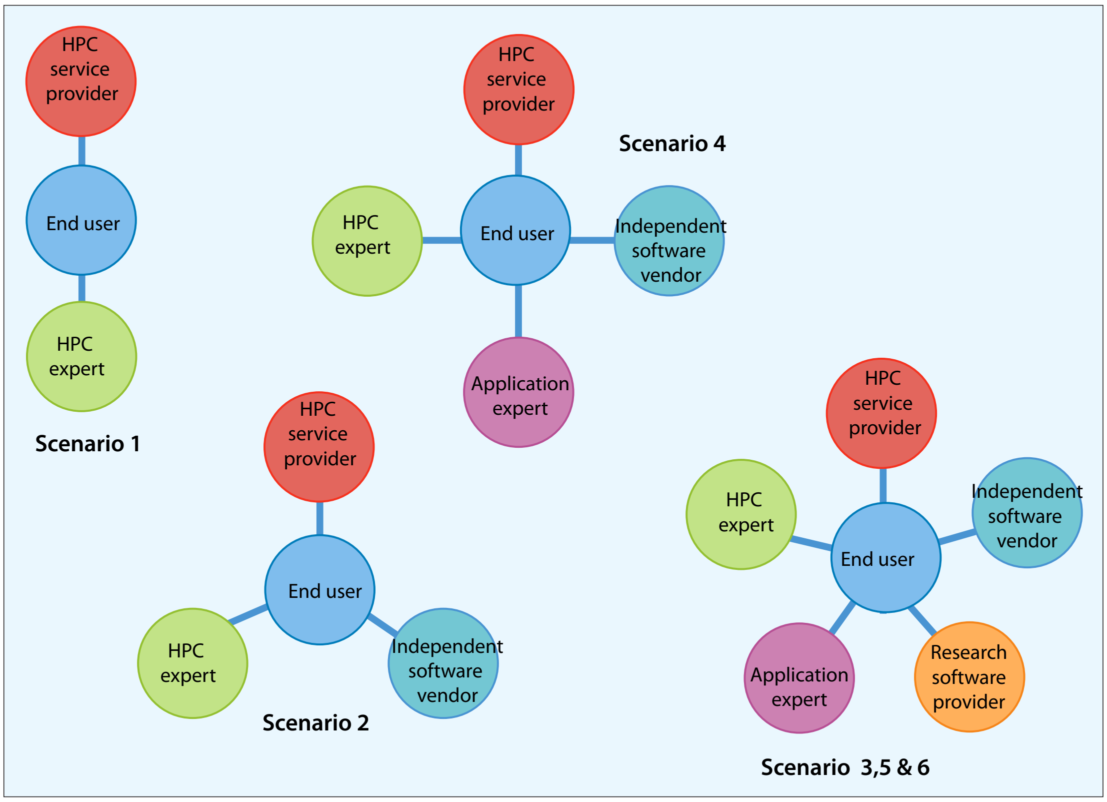


Figure 1: Interactions between HPC actors.

The following table shows how each actor benefits.

Scenario	End user	HPC Service Provider	Independent Software Vendor	Research Software Provider	Application Expert	HPC expert
1	Enhanced code and results from new HPC system	Potential new market	N/A	N/A	N/A	Fees, know-how
2	Results from new HPC system	Potential new market	Enhanced product	N/A	N/A	Fees know-how
3	Assessment of whether HPC is suitable for its business	Potential new market	Potential new customer	Know-how. Possible revenue stream	Know-how. Possible revenue stream	Know-how. Possible revenue stream
4	Results previously unobtainable	Potential new market	Enhanced product	N/A	Know-how. Possible revenue stream	Know-how. Possible revenue stream
5	Results previously unobtainable	Potential new market	Possible new product	Know-how. Possible revenue stream	Know-how. Possible revenue stream	Know-how. Possible revenue stream
6	Application available on future HPC systems: business continuity	Potential new market	Future-proofed product	Know-how. Possible revenue stream	Know-how. Possible revenue stream	Know-how. Possible revenue stream

Table 1: How different actors benefit in HPC applications.

Long term R&D strategy

A long term programme of R&D must be initiated to overcome the major technological hurdles that have been identified in this report.

A long term programme of R&D is needed, together with immediate actions which will prepare its foundation.

This report has identified some of the areas that should be prioritised for long term R&D:

- Highly scalable methods for modelling and simulation that can exploit massive parallelism and data locality.
- New programming models and tools, targeted at massively parallel and heterogeneous environments.
- Decoupling application development from HPC.
- Technologies to support new and emerging applications which require robust HPC with real-time capability.
- Data-intensive HPC.
- Low-energy computing from both an architectural and application perspective.

The research programme must be focused ultimately on delivering results that are relevant to tackling the economic and societal challenges that lie ahead. It must be ambitious and aim for a transformation of the HPC market in Europe, rather than incremental change.

It must involve all actors in the HPC domain and address the widest possible spectrum of applications.

The preparation for this programme of research needs to start as soon as possible in Framework 7 and continue through Horizon 2020; the following section outlines actions that should be taken immediately in support of it.

Immediate actions

Actions along three lines are proposed to stimulate the HPC market and to prepare the foundation for future longer term R&D described above:

- 1. HPC pilot networks which stimulate the HPC solutions marketplace.
- 2. Research & development activities, which stimulate technology development in the HPC domain, and transfer of technology to the HPC domain from other computing domains and constituencies. This is essential if Europe is to maintain its leading position in HPC.
- 3. Visioning, roadmapping and constituency-building activities, which prepare a long term strategy for R&D&I initiatives under Horizon 2020.

A programme to achieve the objectives listed above must combine a mixture of activities which bring the relevant actors together, coupled with a professional effective marketing and publicity campaign, and delivered in such a way as to stimulate the market for industrial and commercial HPC.

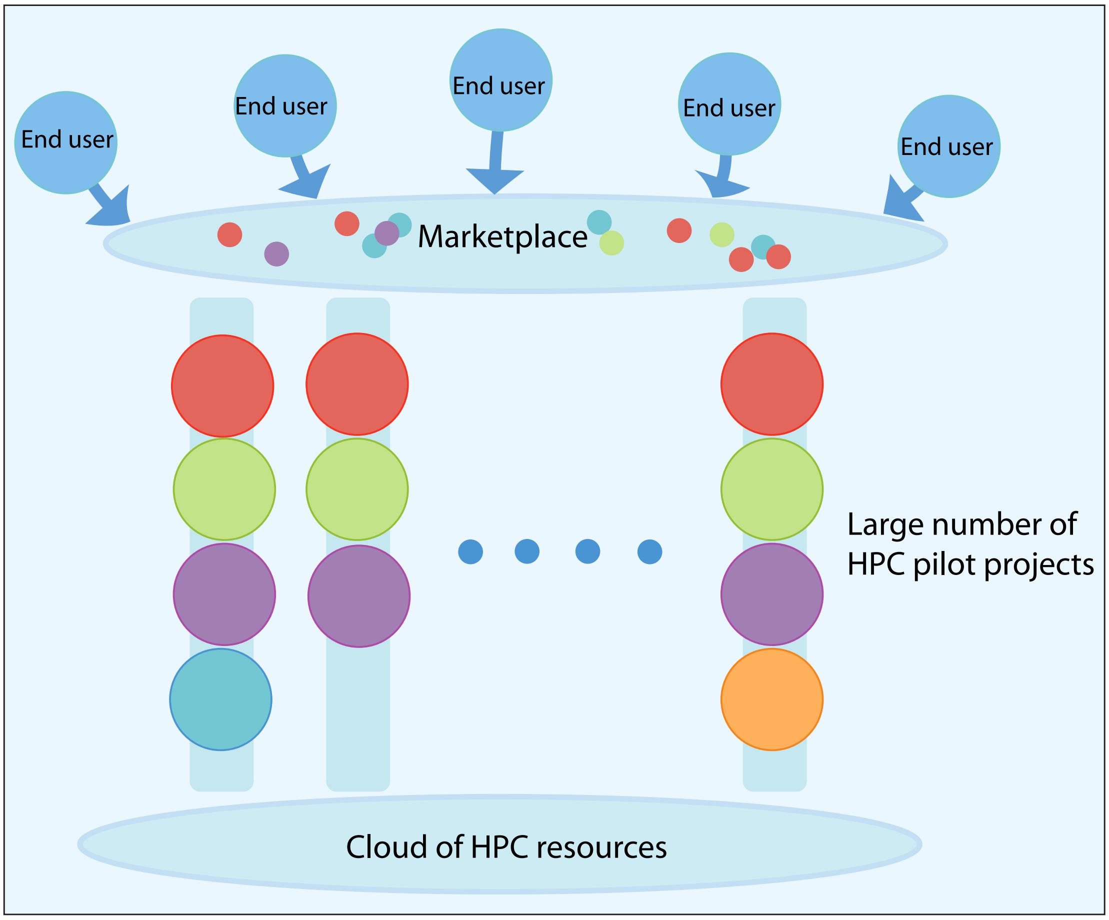
We propose that these immediate actions are carried out within the context of Framework 7 such that lessons can be learned for future investment in Horizon 2020.

HPC pilot networks

As an initial step to meeting the first action line above, we propose the creation of one or more large-scale pilot networks, each targeted at a specific industry sector. These will bring together a critical mass of actors required to deliver a coherent set of clustered projects which tackle the innovation and research challenges outlined earlier in this document.

In Europe we already have experience of delivering such programmes of support through, for example, Europort, the HPCN Technology Transfer Node network and the EUTIST-Integrated Machine Vision network. Each of these activities delivered a significant number of highly innovative, end-user driven projects that tackled a broad set of technical problems faced by European industry at that time. This model of interaction has been demonstrated to work in the past and will work again. The diagram below summarises our vision:

The proposed large-scale pilot network (or networks) will consist of a marketplace of actors, offering the different components necessary for successful projects. An end-user will approach the marketplace to discover available providers in their problem-space and discuss their requirements with



the various providers. From these discussions, the various parties concerned will make a proposal for funding. The success of this approach has already been demonstrated in previous projects such as BEinGRID. The administrative processes must be lightweight in order to encourage participation by industry.

Proposals should be encouraged from ISVs wishing to modernise their codes driven by the needs of the end-users. Projects should be clustered around strategically important applications, and the results should be focused on the long-term use of the software, rather than short-term incremental fixes.

A large-scale network of projects should be set up to tackle innovation and research challenges and stimulate the market.

Proposals should be encouraged from organisations which develop tools. Prototype versions of new tools should be made available under preferential terms to other users in the network.

Service providers would make their HPC systems available as a cloud of HPC resources where the correct resources (in terms of scale, available storage, software licensing etc.) could be matched to each project and paid for appropriately. This approach would offer end-users the benefits of the conventional cloud computing model (i.e. pay-per-use and no capital costs) while retaining their flexibility to access the most appropriate architecture for each case, whether that be a tier-0 supercomputer, a cluster optimised for data-intensive processing or some other type of HPC system.

We envisage an initial set of projects being submitted with the proposals that will be augmented over time through Open Calls. Projects selected for funding will be chosen according to pre-agreed criteria based on the need to tackle specific technical and administrative challenges apparent in the HPC domain today.

Role and profile of participants

The pilot projects must involve all of the actors in mentioned on page 28, although not all may appear in every project and some participants may act in multiple roles. There must be a strong emphasis on SME involvement and application areas where HPC has yet to make a strong impact.

Benefits to participants

The following benefits can be foreseen for participants. All of the following arise through collaboration with other partners, enabled by a network approach.

- End User

  Ability to evaluate new technology at low risk.

  Opportunity to build business case for investment in new technology.

  Buy application services rather than licences and computing systems.

- HPC service providers

  New customers and revenue streams.

  Opportunity to build case for investment in new systems.

  New business models, e.g. pay-per-use based, IaaS.

- HPC systems vendors

  Better understanding of application requirements.

  Opportunity for co-design.

- ISVs

  New customers and revenue streams.

  Better understanding of users' requirements, leading to new features or new applications.

  Better performance of applications.

  Routes to migrating software to new computer architectures.

  Opportunity for co-design.

  New business models, e.g. pay-per-use based, SaaS.

- Research software providers

  Commercialisation opportunities for research software.

  Opportunity for co-design.

- Application experts

  Commercial opportunities to form one-stop-shop for end-users.

  Opportunity for co-design.

- HPC consultancies

  Opportunities to commercialise know-how through licensing with ISVs and end-users.

Business models and sustainability

Support from the European Commission must be targeted at ensuring market success and ensuring that European users and suppliers are given a competitive edge in the global marketplace. The participants in the pilot network must have the objectives of expanding their businesses. Those participants which are commercial organisations must do this either by using their current business models, or by adapting them to take advantage of the new opportunities. HPC centres that derive part of their income from commercial activities and partly from government sources need to find sustainable ways to continue the provision of services. The key to sustainability will be ensuring that the organisation that faces the customer – the business owner – has the best understanding of how to solve the customer's problem. There is not necessarily a single type of actor that can fulfil this role in every case; the pilot network will provide an ideal environment for the right collaboration and business models to be found.

The need for action at a European and local level

Addressing the issue at the European level is essential. It opens new markets for suppliers and customers, and exploits the European network of HPC experts. This will allow specialisation across member states rather than duplication of effort within them and enable actions targeted at European level challenges (e.g. green transport).

A Europe-wide publicity and marketing activity will ensure a pipeline of new end-users is brought to the marketplace while a complementary activity will measure the economic impact of the programme.

In terms of support and contact with SMEs, a regional dimension is essential: SMEs must be able to work with organisations that they trust, which are often local. This approach will also allow participants to focus on a specific sector of regional importance.

We envisage a model whereby companies can establish trust in new methods of working by following a regional approach, and then gain trust in the marketplace of suppliers through a European approach. In this way we see the local and European aspects of collaboration being essential and complementary.

Research and development activities

Synergies between HPC and other computing disciplines have been identified in this report. A programme of small research and development projects is proposed that should stimulate technology development within the HPC domain, and facilitate a two-way transfer of technology between the HPC and other domains. This should be related to such issues as energy efficiency, reliability, time criticality, hardware and software co-design, reconfigurability, customisation and tools.

These short term R&D activities will feed into the longer-term research programme.

In addition to building on European strengths in areas such as embedded and mobile computing, R&D actions must also aim at a structuring effect, bringing constituencies together who are rather distinct today but whose collaboration would have a high economic, technical and competitive potential.

Visioning, roadmapping and constituency-building activities

The immediate actions described above must be coupled to a longer-term programme of research and development. R&D themes which cut across the scenarios presented in earlier sections should be established with a view to their continuation through Horizon 2020.

The longer-term programme must be multidisciplinary, involving HPC experts, scientists (including computer scientists), mathematicians and engineers, together with application experts from all appropriate fields.

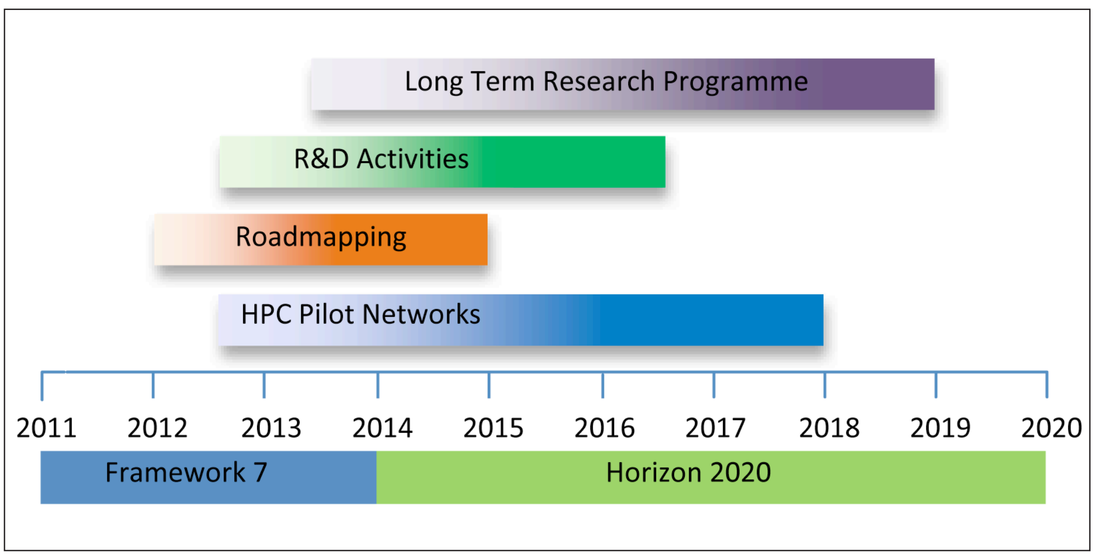
The themes that should form the basis of the long term programme have been mentioned earlier, although more may emerge as time progresses.

These themes should form the basis for a set of support actions which will identify promising research directions within their field, build constituencies and sketch potential research roadmaps. These should be analysed based on their market relevance and potential benefit to define in detail a longer-term research strategy for Horizon 2020.

These actions must be carried out in a way that allows cross-over of ideas to ensure that gap analysis can be done and a holistic approach can be taken. There must also be cross-over with strategic programmes such as Factories of the Future PPP, European Green Car Initiative, the Virtual Physiological Human project and other initiatives.

Timeline of activities

The preparation for this programme of research needs to start as soon as possible in Framework 7 and continue through Horizon 2020. The following timeline is proposed for actions:



Conclusion

This document promotes HPC as an important technology for Europe, one in which further investment is essential if its economic competitiveness is to be maintained.

This report describes the technological challenges that must be overcome for the full potential of HPC to be realised, and has put forward the case for Europe to take a leading role in tackling these issues.

This report outlines the numerous opportunities for using HPC in new applications, and the possibilities opened up by using HPC with other computing domains such as embedded systems and mobile computing.

A plan of research, development and innovation actions is proposed, aimed at stimulating the HPC market and addressing the long term challenges associated with HPC. These recommendations carry the endorsement of leading industrial and academic experts in the HPC field.

Acknowledgements

This report was prepared by the PlanetHPC project in consultation with leading industry and academic experts in the field of HPC. The contributors are listed at the start of the document. PlanetHPC is supported under Objective "Computing Systems" of Challenge 3 "Components and Systems" of the ICT Programme under contract number 248749. This document represents the views of the PlanetHPC experts. Neither PlanetHPC nor the European Commission is liable for any use that may be made of the information contained in this report.