Petaflop Computing in the European HPC Ecosystem

Cray Users' Group

May 7th, 2008

Kimmo Koski

CSC – The Finnish IT Center for Science

CSC

Topics

1. **Terminology and definitions**
2. **HPC in EU Frame Programs**
3. **ESFRI and IT services**
4. **Petaflop computing in Europe**
5. **New HPC Ecosystem**
6. **Case CSC**
7. **Some conclusions**

CSC

Terminology and pointers

- **HPC**
  - High Performance Computing
- **HET,** http://www.hpcineuropetaskforce.eu/
  - 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
- **e-IRG,** http://www.eirg.eu
  - 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,** http://cordis.europa.eu/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.
- **RI**
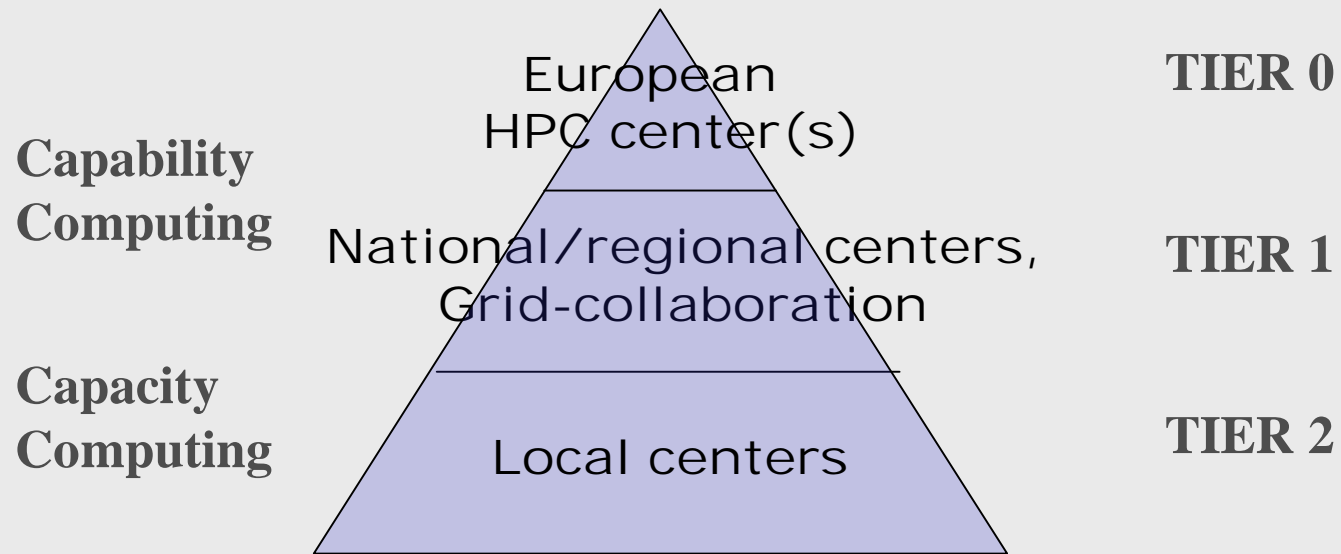 - Research Infrastructure

CSC

Terminology and pointers (cont.)

- **PRACE**
  - Partnership for Advanced Computing in Europe
  - EU FP7 project for preparatory phase in building the European petaflop computing centers, based on HET work
- **DEISA, https://www.deisa.org/**
  - 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, http://www.eu-egee.org/**
  - 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.
- **EGI, http://www.eu-egi.org/**
  - An effort to establish a sustainable grid infrastructure in Europe
- **GÉANT2, http://www.geant2.net/**
  - Seventh generation of pan-European research and education network

CSC

Performance Pyramid

Capability Computing

Capacity Computing

European HPC center(s)

National/regional centers, Grid-collaboration

Local centers

TIER 0

**TIER 1**

TIER 2

CSC

Need to remember about petaflop/s...

➢ **What do you mean with petaflop/s?**

1. Theoretical petaflop/s?
2. LINPACK petaflop/s?
3. Sustained petaflop/s for a single extremely parallel application?
4. Sustained petaflop/s for multiple parallel applications?

➢ **Note that between 1 and 4 there might be several years**

➢ **Petaflop/s hardware needs petaflop/s applications, which are not easy to program, or not even possible in many cases**

• Do we even know how to scale over 100000 processors ...

CSC

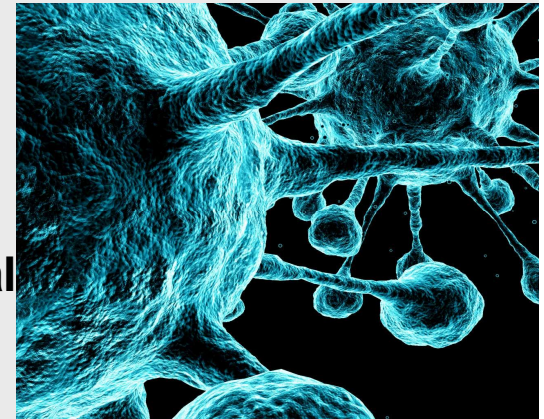# Computational science infrastructure
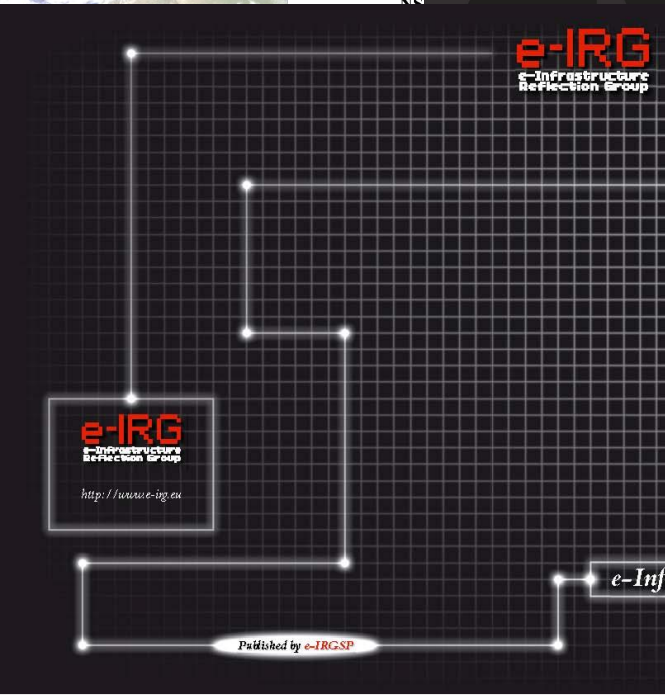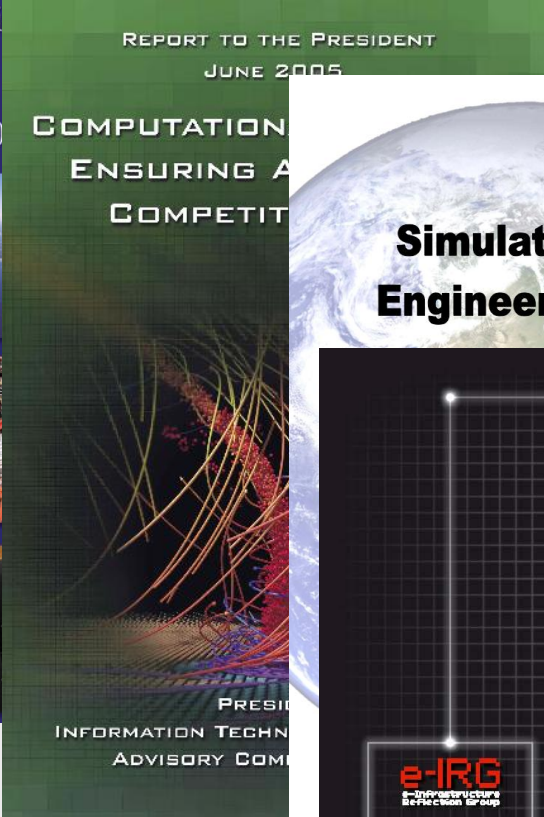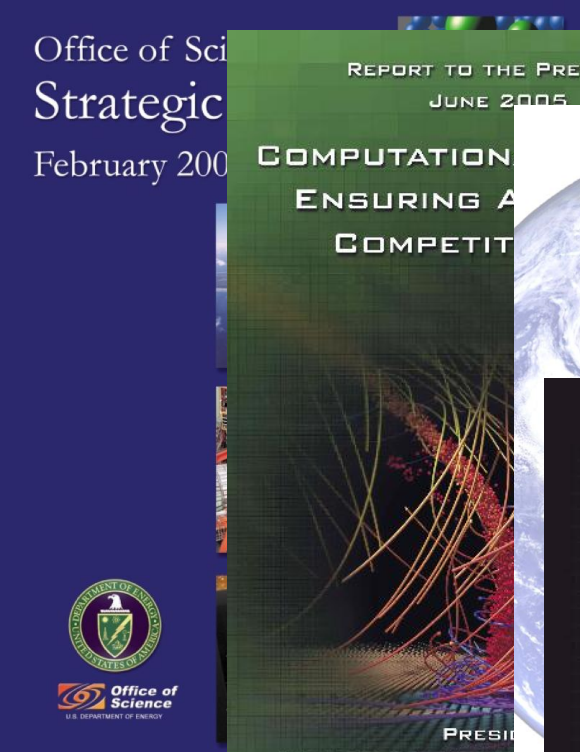
HPC in Europe: from FP6 to FP7

The Era of EU Frame Program 6,

Moving towards FP7

CSC

European HPC after FP6

- ➢ **Multiple Grid projects with varying results – learning for collaboration**
- ➢ **Early experiences about interoperability between national HPC centers**
- ➢ **Communities start to form, in various levels**
- ➢ **Research community more active in computational science domain**
- ➢ **European Union targets in creating sustainable infrastructures**

- ➢ **Petaflop computing raised to European agenda, scientific case for high-end computing available**

CSC

DEISA –Distributed European Infrastructure for Supercomputing Applications

- A consortium of leading national supercomputing centres deploying and operating a persistent, production quality, distributed supercomputing environment with continental scope

- Grid-enabled FP6 funded Research Infrastructure

- A 4-year-project started on May 2004

- Total budget is *37,1 M€* (incl. DEISA and eDEISA contracts), EU funding - *20.9 M€*

- **>200 VOs from several scientific domains**
  - Astronomy & Astrophysics
  - Civil Protection
  - Computational Chemistry
  - Comp. Fluid Dynamics
  - Computer Science/Tools
  - Condensed Matter Physics
  - Earth Sciences
  - Fusion
  - High Energy Physics
  - Life Sciences
- **Further applications under evaluation**

No. jobs / month - all

98k jobs/day

- OPS
- Non-LHC
- LHC
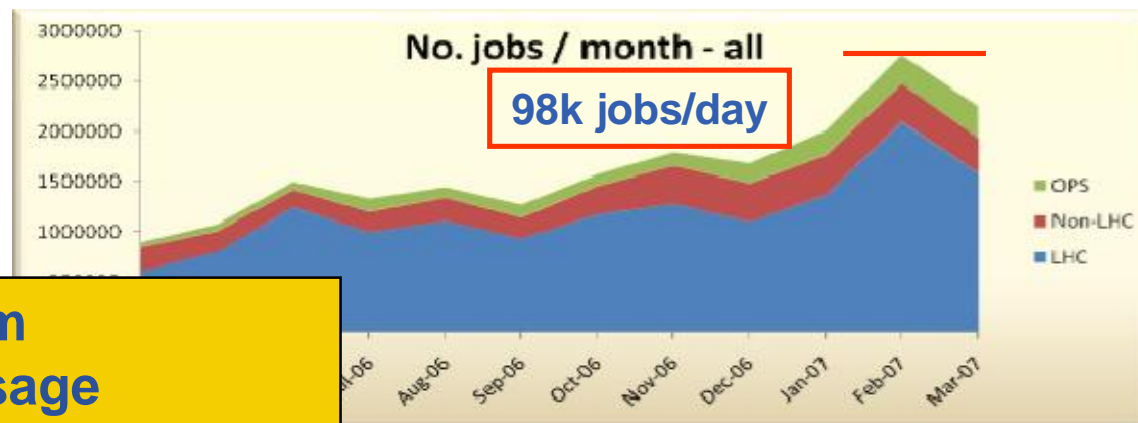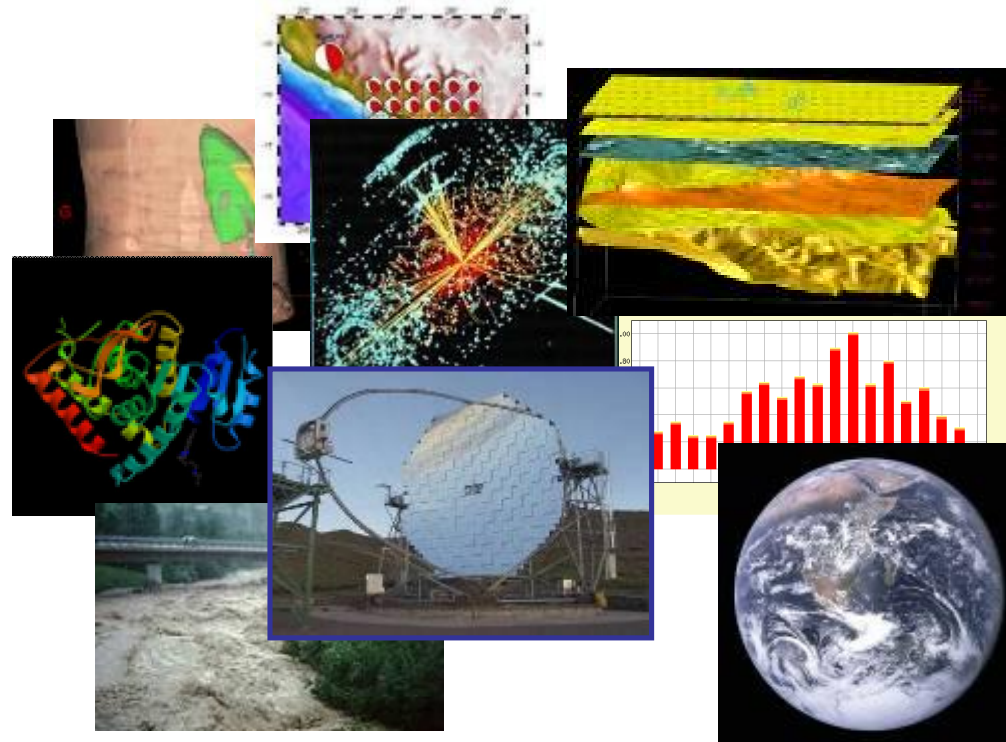
Applications have moved from testing to routine and daily usage

~80-90% efficiency

European Grid Initiative

Goal:
- Long-term sustainability of grid infrastructures in Europe

Approach:
- Establishment of a new federated model bringing together NGIs to build the EGI Organisation

EGI Organisation:
- Coordination and operation of a common multi-national, multi-disciplinary Grid infrastructure
    - To enable and support international Grid-based collaboration
    - To provide support and added value to NGIs
    - To liaise with corresponding infrastructures outside Europe

EGI Objectives:
- Ensure the long-term sustainability of the European e-infrastructure
- Coordinate the integration and interaction between National Grid Infrastructures
- Operate the European level of the production Grid infrastructure for a wide range of scientific disciplines to link National Grid Infrastructures

EGI Vision:

http://www.eu-egi.org/vision.pdf

Policy and strategy work

- ➢ **HET: HPC in Europe Taskforce**
  http://www.hpcineuropetaskforce.eu/

- ➢ **e-IRG: e-Infrastructure**
  Reflection Group
  http://www.e-irg.org/

- ➢ **ESFRI: European Strategy Forum**
  on Research Infrastructures
  http://www.cordis.lu/esfri/

HPC Ecosystem to support the top

➢ **The upper layers of the pyramid**
- HPC centers / services
- European projects (HPC/Grid, networking, …)

➢ **Activities which enable efficient usage of upper layers**
- Inclusion of national HPC infrastructures
- Software development and scalability issues
- Competence development

➢ **Interoperability between the layers**

CSC
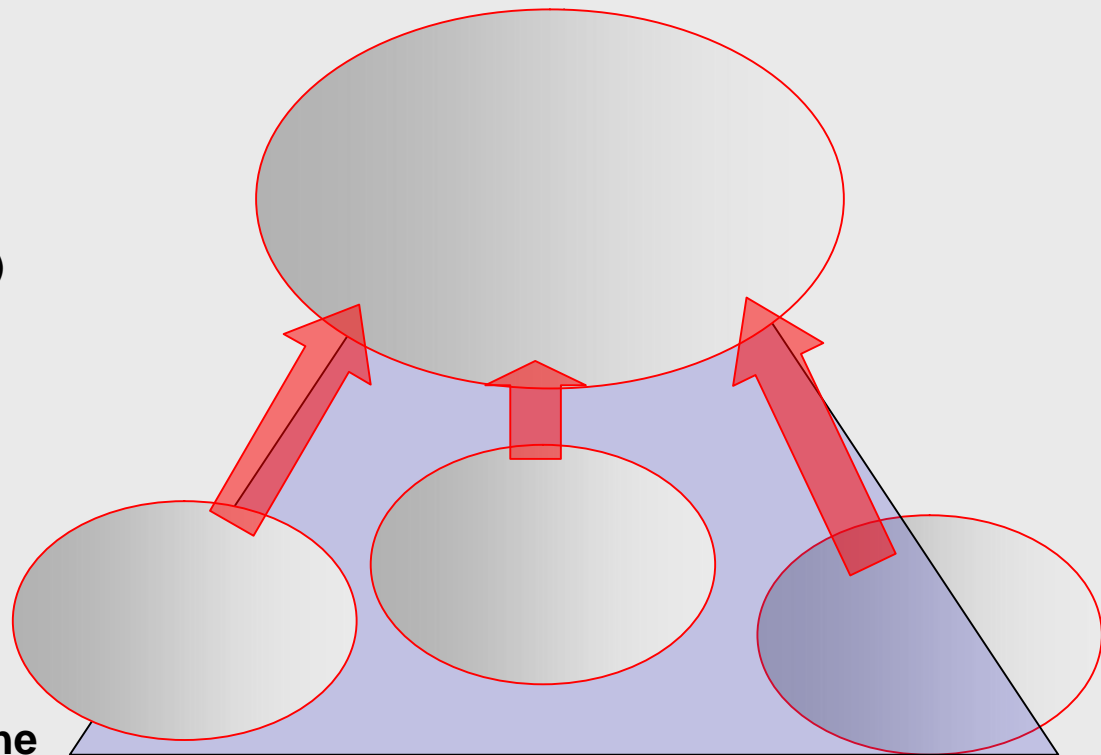
Targets for European HPC collaboration 2007 onwards

- ➤ **Continuation of existing grid projects (DEISA, EGEE …) and development in GEANT2 network infrastructure**
- ➤ **Building European petaflop computing services integrated in the full HPC ecosystem according to the performance pyramid model (PRACE)**
- ➤ **Maximal synergy between PRACE and DEISA (integration after some time?)**
- ➤ **Interoperability between PRACE and EGI/EGEE**
- ➤ **Building up research infrastructure services for ESFRI roadmap**
- ➤ **Linking the policy works to support optimally each other: ESFRI and e-IRG**
- ➤ **Target to establish an active European community for HPC: infrastructure, resource sharing, communication and collaboration over country borders**

CSC

Next steps

- ➢ **From project base organization to sustainable infrastructures**
- ➢ **From disciplinary IT silos to horizontal services and synergy**
- ➢ **From hardware orientation to full HPC ecosystem model, including software, data and competence development etc.**
- ➢ **From sub-optimization through "I'll do all by myself" model to collaboration**

ESFRI

ESFRI

- ➤ **Strategy Forum with a consulting role to EU**
- ➤ **Wide representation of scientists in various disciplines**
- ➤ **Roadmap process for major new European research infrastructures (range of 10-1000 MEUR for an infrastructure)**
- ➤ **Roadmap published in 2006**
  - 35 projects labeled mature
  - One of the projects European HPC Service
- ➤ **Preparatory projects for each project**
  - 1-4 years
  - Deadline for project call was May 2nd, 2007
- ➤ **ESFRI-list update in process**
  - Renewed list targeted for autumn 2008

European Strategy Forum
on Research Infrastructures **ESFRI**

EUROPEAN ROADMAP
FOR RESEARCH
INFRASTRUCTURES

Report 2006

CSC

Impact of ESFRI

- **Rising a lot of interest**
  - Scientific communities
  - EU
  - National priorities
- **Preparatory phase call by EU**
- **National funding**
- **Political and non-political discussions for hosting of ESFRI infrastructures**
- **Obvious need for prioritising**

- **NOTE: ESFRI list includes only the new infrastructures. The existing ones have development plans, too**

CSC

ICT infrastructure and ESFRI

- **Only one of the projects is from ICT sector**
  - PRACE for petaflop computing
- **All of the projects need ICT infrastructure at some level**
  - Computing, data handling, software development, networks, …
  - Will this be properly understood is a good question
- **Need for a strong horizontal ICT infrastructure to avoid overlapping work**

- **And the ESFRI-list is being updated just now**
  - Should there be more ICT entries in the updated list?
  - Data handling and software development would be good candidates…

CSC

ESFRI infrastructures and other infrastructures require professional services for computing, data handling, networks, software development, parallel computing, computational methods etc.
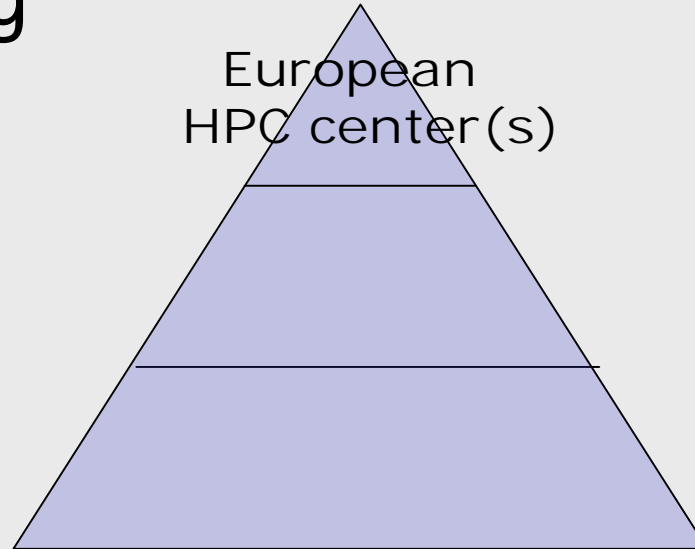
These tasks do not necessarily vary much between sciences and should not be done separately to each research infrastructure.
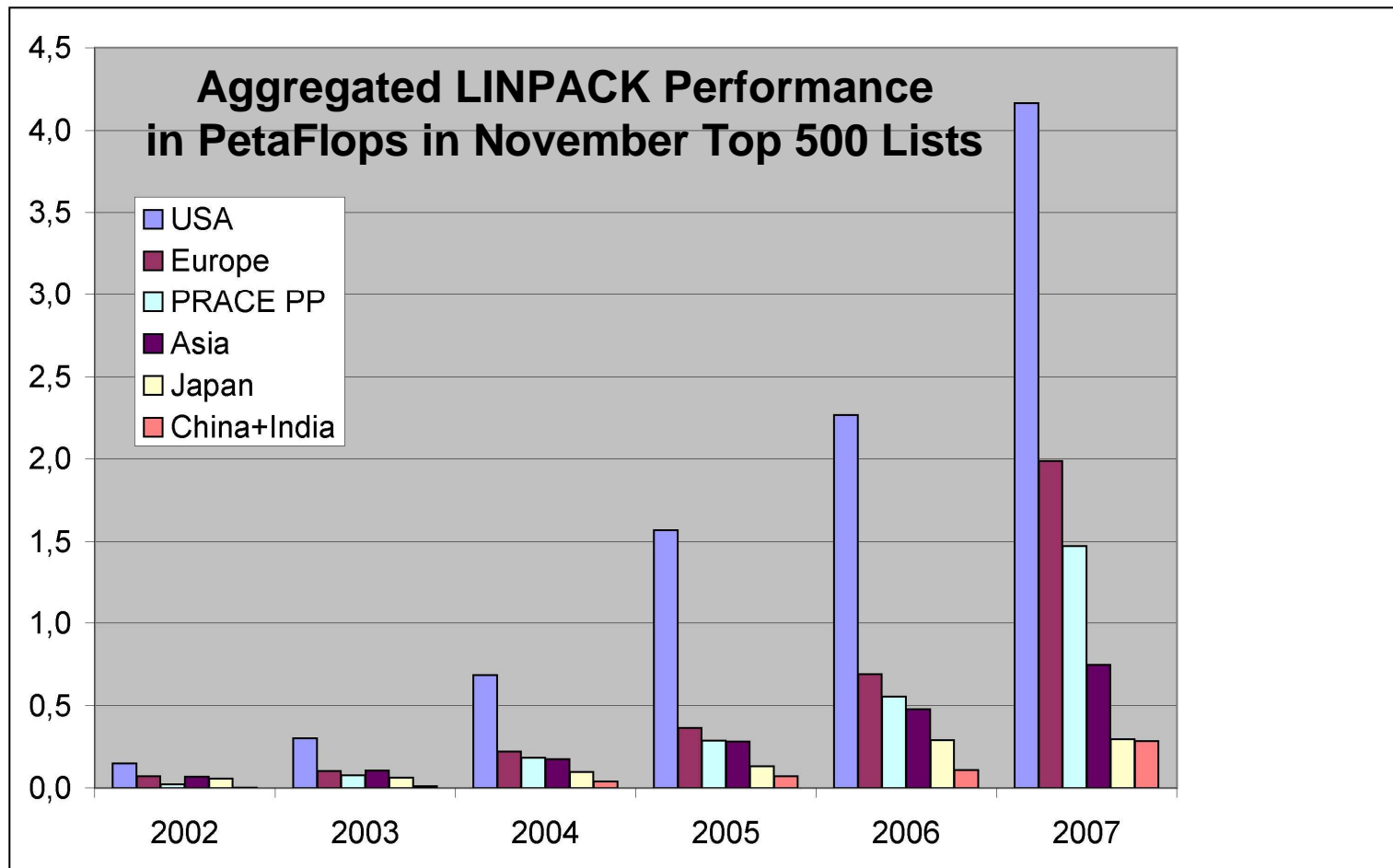
CSC

# Petaflop/s computing

European
HPC center(s)

CSC

Europe's current position in HPC

Aggregated LINPACK Performance
in PetaFlops in November Top 500 Lists

Legend: USA, Europe, PRACE PP, Asia, Japan, China+India

The ESFRI Vision for a European HPC service

European HPC-facilities at the top of an HPC provisioning pyramid

- Tier-0: 3-5 European facilities
- Tier-1: National facilities
- Tier-2: Regional/University Centres
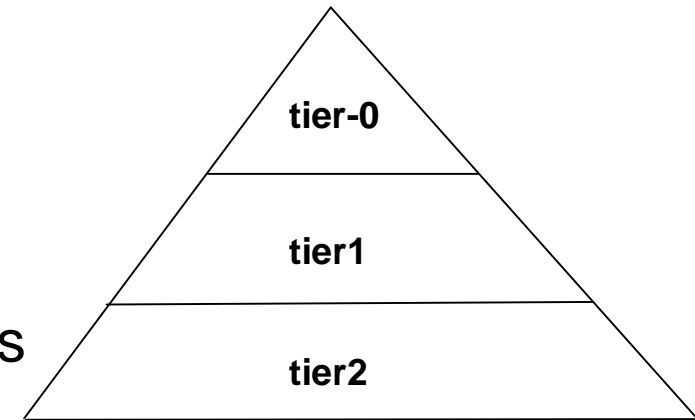
Creation of a European HPC ecosystem involving all stakeholders

- HPC service providers on all tiers
- Grid Infrastructures
- Scientific and industrial user communities
- The European HPC hard- and software industry

First Steps and Achievements

Production of the HPC part of
the ESFRI Roadmap;
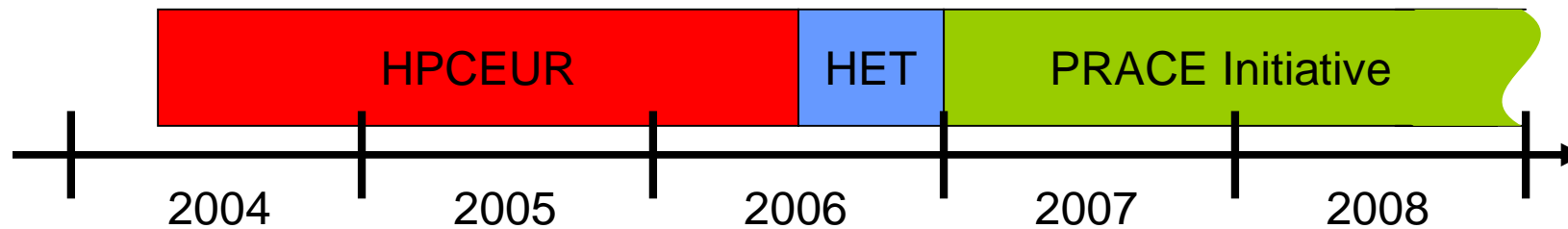Creation of a vision,
involving 15 European countries

Signature of the MoU

Submission of
the project proposal

Bringing scientists together
Creation of the Scientific Case

Approval of the project

Kick-off

| HPCEUR | HET | PRACE Initiative |

2004   2005   2006   2007   2008

PRACE – Project Facts

Objectives of the PRACE Project:

- Prepare the contracts to establish the PRACE permanent Research Infrastructure as a single Legal Entity in 2010 including governance, funding, procurement, and usage strategies.
- Perform the technical work to prepare operation of the Tier-0 systems in 2009/2010 including deployment and benchmarking of prototypes for Petaflop/s systems and porting, optimising, peta-scaling of applications

Project facts:

- Partners: 16 Legal Entities from 14 countries
- Project duration: January 2008 – December 2009
- Project budget: 20 M € ,   EC funding: 10 M €

PRACE – Project Consortium

Principal Partners

General Partners

tier 0

tier 1

The next tasks:

… growing into a persistent Research Infrastructure

Define the legal form and governance

Secure initial and continuous funding

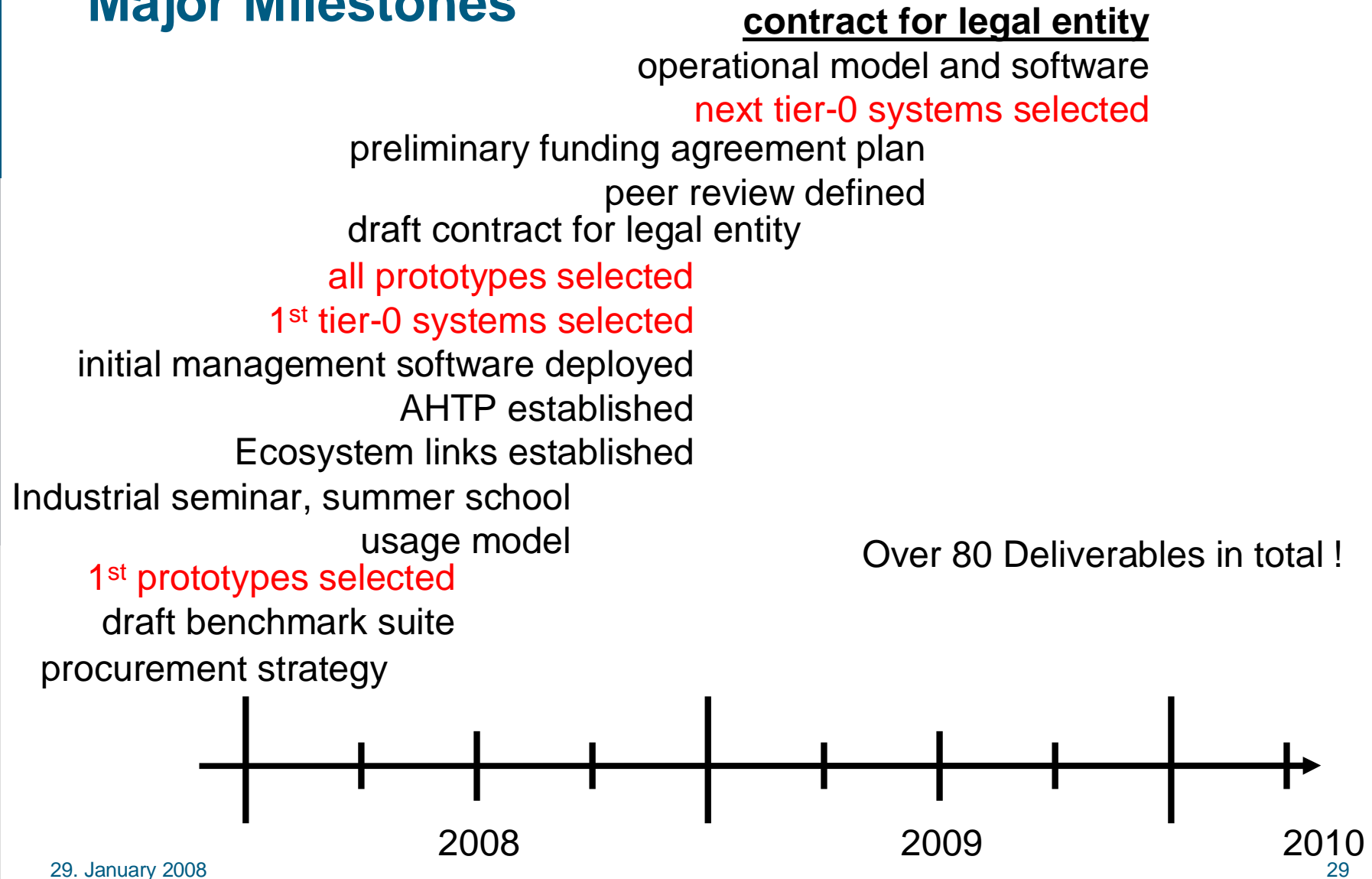Prepare procurement and installation of the first petaflop/s
systems

Establish the peer review process for academic usage

Promote Europe wide collaboration between scientific
communities using leading edge scientific simulation

Encourage new projects to increase software and simulation
competence

Major Milestones

<u>contract for legal entity</u>

operational model and software

next tier-0 systems selected

preliminary funding agreement plan

peer review defined

draft contract for legal entity

all prototypes selected

1st tier-0 systems selected

initial management software deployed

AHTP established

Ecosystem links established

Industrial seminar, summer school

usage model

1st prototypes selected

draft benchmark suite

procurement strategy

Over 80 Deliverables in total !

2008   2009   2010

Opportunities ahead

We build upon

- the HPC expertise of 14 European countries, created through HPC service provisioning and projects like DEISA
- the expressed support of our national governments, the European Commission and many scientific communities
- an excellent team-spirit grown during the past years of HPCEUR, HET, PRACE and other joint endeavors

The time is right

- to boost European competitiveness
- to position ourselves as a leader in HPC and its applications
- to create and shape the European HPC ecosystem

New HPC Ecosystem is being built…

Main players in European HPC Ecosystem

➢ **ESFRI Roadmap and 35 new research infrastructures**

➢ **PRACE – Petaflop computing centers**

➢ **EU-supported infrastructure projects, such as EGEE, DEISA, GEANT2 and OMII-Europe**

➢ **European Grid Initiative, EGI**

➢ **Policy groups, such as ESFRI and e-IRG**

➢ **Regional activities, such as NDGF**

➢ **National Infrastructures**

➢ **International centers, such as CERN, EBI and ECMWF**

➢ **User communities with HPC requirements, such as fusion or climate**

CSC

Some Key Issues

- ➢ **Sustainability**
  - EGEE and DEISA are projects with an end
  - PRACE and EGI are targeted to be sustainable with no definitive end
- ➢ **ESFRI and e-IRG**
  - How do the research side and infrastructure side work together?
  - Two-directional input requested
- ➢ **Requirement for horizontal services**
  - Let's not create disciplinary IT silos
  - Synergy required for cost efficiency and excellence
- ➢ **ICT infrastructure is essential for research**
  - The role of computational science is growing
- ➢ **Renewal and competence**
  - Will Europe run out of competent people?
  - Will training and education programs react fast enough?

CSC

New market for European HPC

- ➢ **35 ESFRI list new research infrastructure projects, most of them starting a preparatory phase project late 2007**
  - 1-4 years
  - 1-7 MEUR * 2 (petaflop computing 10 MEUR * 2)
- ➢ **Successful new research infrastructures start construction 2009-2011**
  - 10-1000 MEUR per infrastructure
- ➢ **Existing infrastructures are also growing**
- ➢ **Results:**
  - Growing RI market, considerably rising funding volume
  - Need for horizontal activities (computing, data, networks, computational methods and scalability, application development,…)
  - Real danger to build disciplinary silos instead of searching IT synergy

> 1 BEUR for ICT ?

CSC

Advice for HPC vendors: Europe wants to develop HPC Ecosystem <u>in Europe</u>

➢ What can you do in Europe?

- **Manufacturing**
- **Research**
- **Software development**
- **Integration work**
- **Training**
- **Other**

➢ Pre-commercial procurement used increasingly

➢ Prototyping is part of the petaflop/s project

- **The prototypes for tomorrow's petaflop systems are sometimes today's production systems**
- **Innovation is both in scalable software and hardware**

CSC

Case CSC

CSC at a glance

- Founded in 1970 as a technical support unit for Univac 1108

- Reorganized as a company, CSC - Scientific Computing Ltd. in 1993

- All shares to the Ministry of Education of Finland in 1997

- Operates on a non-profit principle

- Facilities in Espoo, close to Otaniemi community (of 15,000 students and 16,000 technology professionals)

- Staff 160 and growing
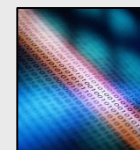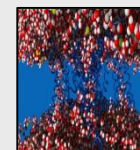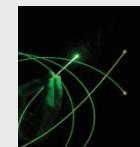
- Turnover 2007 16.7 million euros

CSC's services

FUNET SERVICES

COMPUTING SERVICES

APPLICATION SERVICES

STORAGE AND CURATION SERVICES

INFORMATION MANAGEMENT SERVICES

CSC

MISSION:
CSC, as a part of the Finnish national research
structure, develops and offers
high quality information technology
services

VISION 2012:
CSC – a leading center of excellence in
information technology for
science in the European research area

The Finnish model: Impact through centralization at CSC

➢ **One center: critical mass, fast decision making, services and infrastructure close to each other**

➢ **FUNET network services part of the national IT center – exceptional concept internationally**

➢ **Wide range of services from research networks to supercomputing and data management, from scientific applications to information management etc.**

➢ **Strong player in international research infrastructures**

➢ **Influence in European strategy working groups**

➢ **New demand for high capacity data transfer through new experiments and increased international collaboration**

➢ **Aiming at a key role as one major IT center for science in the future Ecosystem for European e-infrastructure**

CSC

Supercomputer as a part of Finnish Science Policy

➢ **Finland has a strong base for computational science: first centralized resources available already in mid 80's**

➢ **The impact of computational science is constantly growing: supercomputers are used increasingly often, and in new scientific disciplines**

➢ **The preconditions for using HPC capacity are strong in Finland: even if it is possible to purchase a supercomputer in less than one year, building up a competence center such as CSC takes easily 10 years**

➢ **Efficient computing environment attracts international top quality research and eases the collaboration possibilities of Finnish scientists**
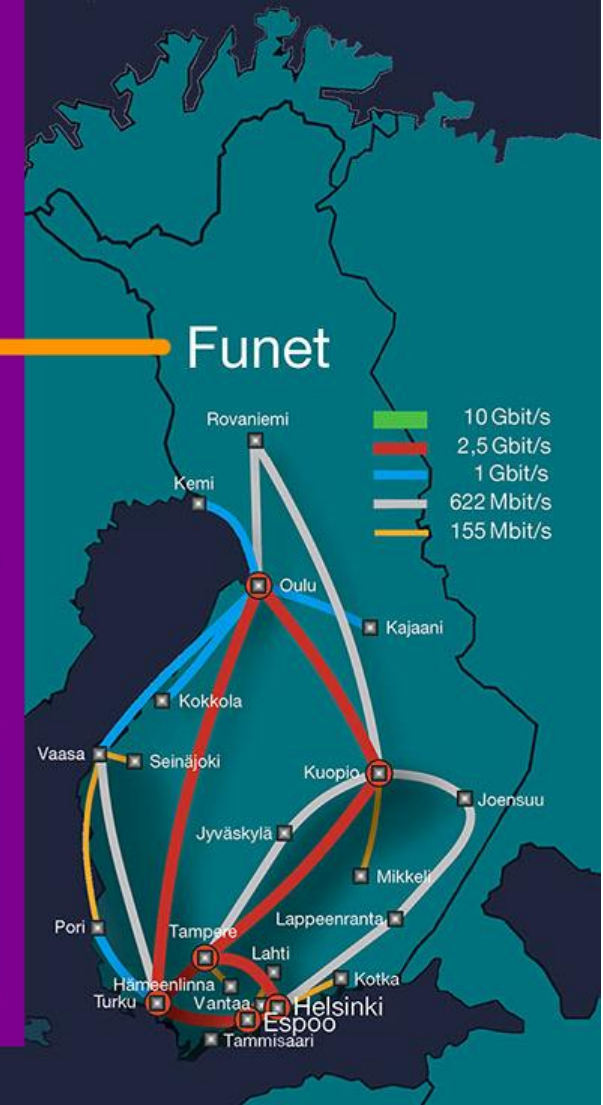
CSC
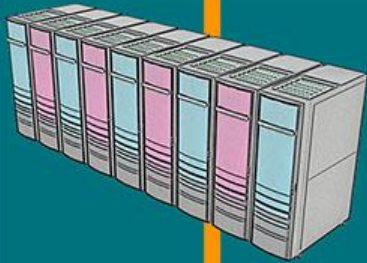
Examples of computational challenges in Finland

1. **Global model for seas: The future of the Golf stream, vital condition for Scandinavia**
2. **Connected models: Forests and nanoscale aerosols, factors for our future climate**
3. **How do cell membranes function**
4. **Development of more efficient drugs against cancer**
5. **New environmentally friendly pulp bleaching chemicals**
6. **Better, faster, cheaper with the aid of computational fluid dynamics**
7. **Accurate quantification of the age and composition of the universe using satellite observations**
8. **New type of solar cells**
9. **Quantum dots and wells as nano-electronics solutions**
10. **Computational modelling of fusion reactors**

CSC

CSC contribution in Europe

➢ **Partner in most European grid infrastructures**
  - DEISA
  - EGEE
  - EGI
➢ **Partner in selected ESFRI list items, focus on "horizontal" ICT services**
  - PRACE
  - ELIXIR
  - CLARIN
➢ **Collaboration with major user communities**
  - Bio, Climate, Fusion, …
➢ **Active participation in future European development**
  - Proposals for ESFRI-list update: data and software development

CSC

Conclusions

Some conclusions

➢ **HPC Ecosystem has a lot of acronyms** ☺
➢ **ESFRI Roadmap projects will have a major impact requiring horizontal (IT-infrastructure) services**
  • 'Research Infrastructure Market' will burst in two years
➢ **Interoperability between different European projects or infrastructures is said to be crucial, but this is not a technical challenge**
➢ **Collaboration has to develop in all levels, for example between different disciplines, between research communities and centers, between European projects and infrastructures, and between countries**
➢ **Since applying HPC is increasing and widening to new areas, funding volume has to grow at some stage**
➢ **The rules for the new HPC Ecosystem are not yet available – best practices can be adapted**

CSC