Building the European High-Performance Computing Ecosystem

Linköping

October 17th, 2007

Kimmo Koski

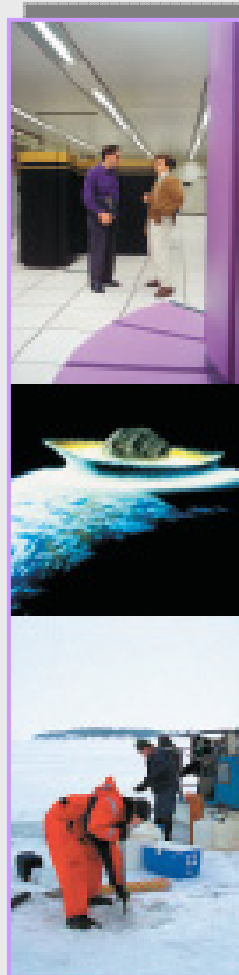**CSC – The Finnish IT Center for Science**

CSC

CSC Fact Sheet

Operated on a non-profit principle

All shares to the Ministry of Education of Finland in 1997

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

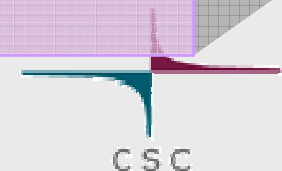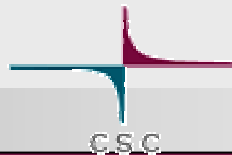Founded in 1971 as a technical support unit for Univac 1108

CSC Turnover in 2006 15.6 M€, 150 employees

Since March 2005, facilities in Keilaniemi, Espoo

First supercomputer Cray X-MP/EA 416 in 1989

Funet started in 1984

CSC

Topics

1. **Terminology and definitions**
2. **Towards European petaflop computing**
   - ➢ Scientific case for petaflop computing
   - ➢ HPC in Europe Taskforce (HET)
   - ➢ Partnership for Advanced Computing in Europe (PRACE)
3. **Middle layer of performance pyramid**
   - ➢ DEISA
   - ➢ EGI/EGEE
4. **New HPC Ecosystem in Europe**
5. **Nordic opportunities**

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.
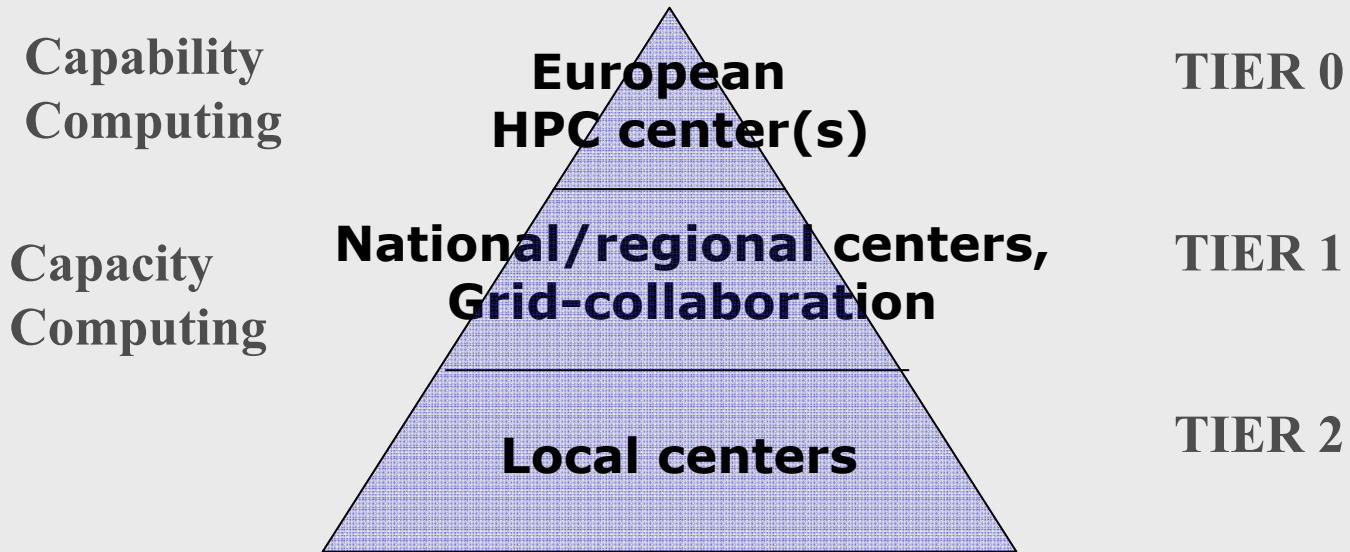
CSC

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
- **NDGF,** www.ndgf.org
  - Nordic Data Grid Facility

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

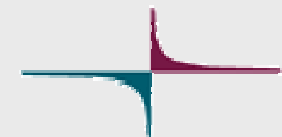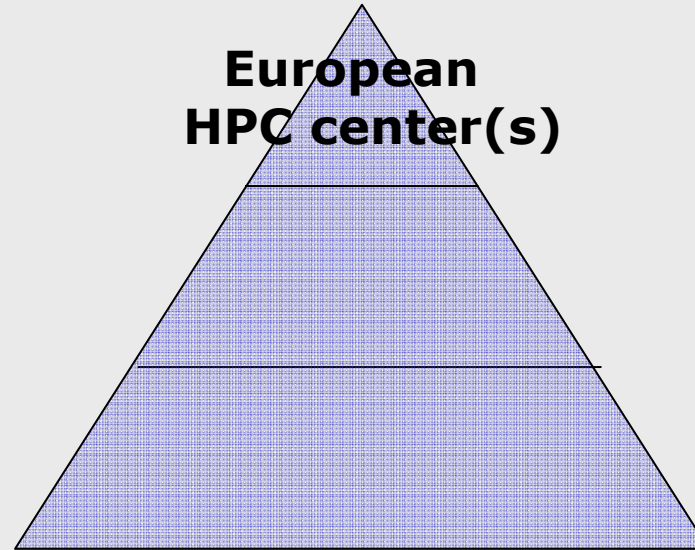# Computational science infrastructure

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

Petaflop computing

**European
HPC center(s)**

Scientific Case prepared cooperatively by panels of European scientists

November 2005: Barcelona Workshop

    Identify key challenge applications

February 2006: Cadarache Workshop

    Refine challenges which need a European Tier 0 infrastructure

CSC

The report through scientific areas

- **Weather, Climatology, Earth Sciences**

- **Astrophysics, HEP and Plasma Physics**

- **Material Science, Chemistry and Nanoscience**

- **Life Sciences**

- **Engineering**

CSC

The Need for European commitment in Capability Computing

In summary, key panel recommendations:

➢ **« Providing scientists and engineers with access to capability computers of leadership-class must be recognized as an essential strategic priority in Europe. »**

➢ **« The panel stresses that these resources should be reserved for the most exigent computational tasks of high potential value. »**

➢ **« This would require putting in place an appropriate process to screen proposals, and to run the resources as permanent research infrastructure. »**
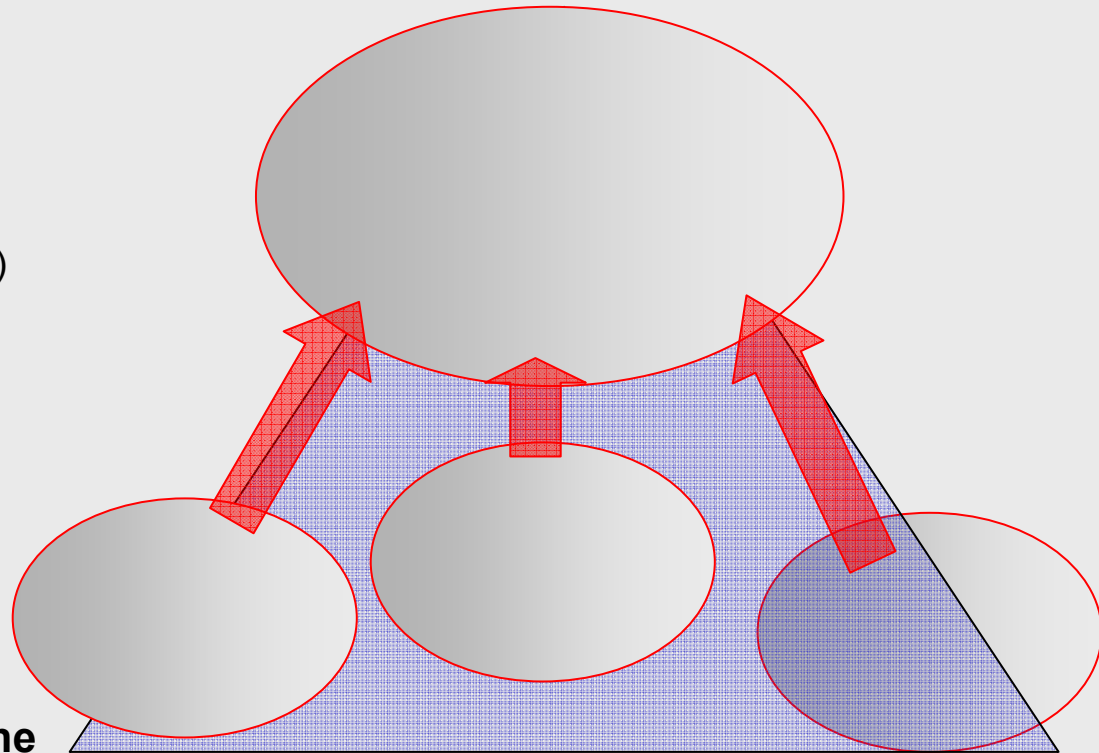
CSC

HPC in Europe Taskforce (HET)

➢ **Temporary taskforce shaping the European strategy for petaflop computing**

➢ **Founded in June 2006**

➢ **Strategy work delivered in January 2007**

➢ **Members from 11 European countries**

➢ **Chaired by CSC, Finland**

➢ **Complete documentation available:**
  http://www.hpcineuropetaskforce.eu/

HET Scope: HPC Ecosystem

- ➢ **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**

HET outcome

➢ **Entry in ESFRI list for petaflop computing**

➢ **Papers:**

- Scientific case for European HPC (most work done by previous HPCEUR project)
- Proposal for funding models
- Proposal for peer review process
- Views for HPC Ecosystem
- Summary paper with recommendations

➢ **Good team spirit with a common approach**

➢ **Basis for practical implementation**

- Consortium for ESFRI Preparatory phase
- Memorandum of Understanding for European Tier 0 HPC service

CSC

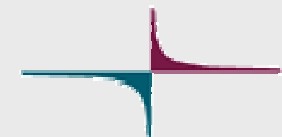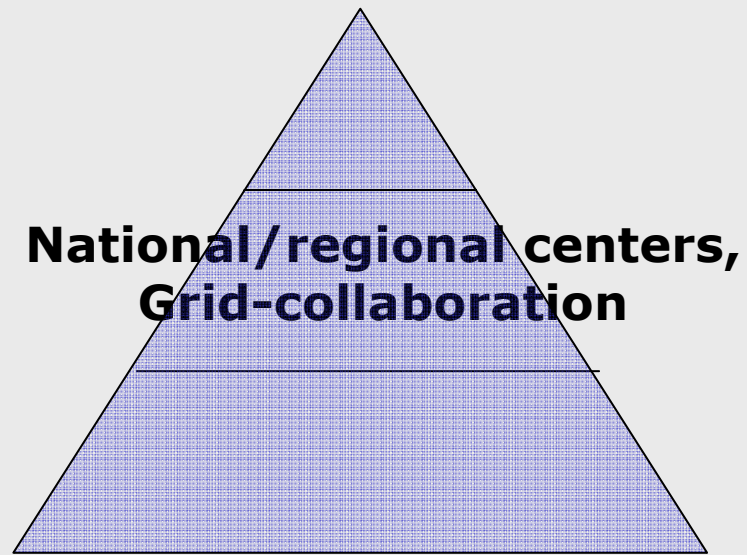Signing of the MoU in Berlin 17.4.2007

What is going to happen with PRACE?

- ➢ **Project will start 1.1.2008**
- ➢ **Consortium partners (14 countries)**
    - • Austria, Finland, France, Germany, Greece, Italy, Norway, Poland, Portugal, Spain, Sweden, Switzerland, The Netherlands, United Kingdom
- ➢ **Two years, 10+10 MEUR volume**
- ➢ **Prototypes for petaflop computing during 2008-2009**
- ➢ **Target to have the first center operational in 2009-2010**
- ➢ **Open issues to be solved during the preparatory phase:**
    - • Which companies to prototype and where to place them?
    - • Who will host the petaflop centers?
    - • Who will pay for construction?
    - • Who can use the resources and under which conditions?
    - • How to link with other projects, for example DEISA?

CSC

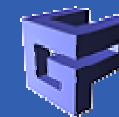# Middle layer

National/regional centers, Grid-collaboration

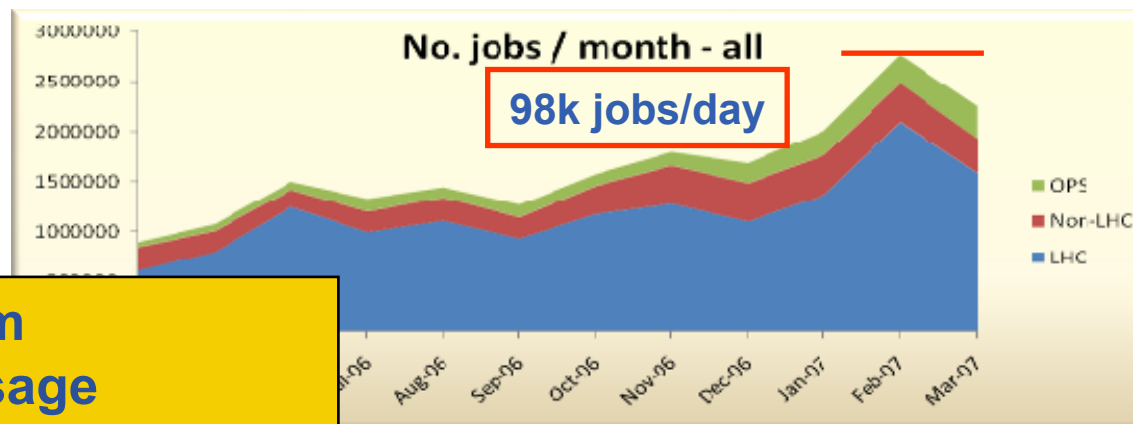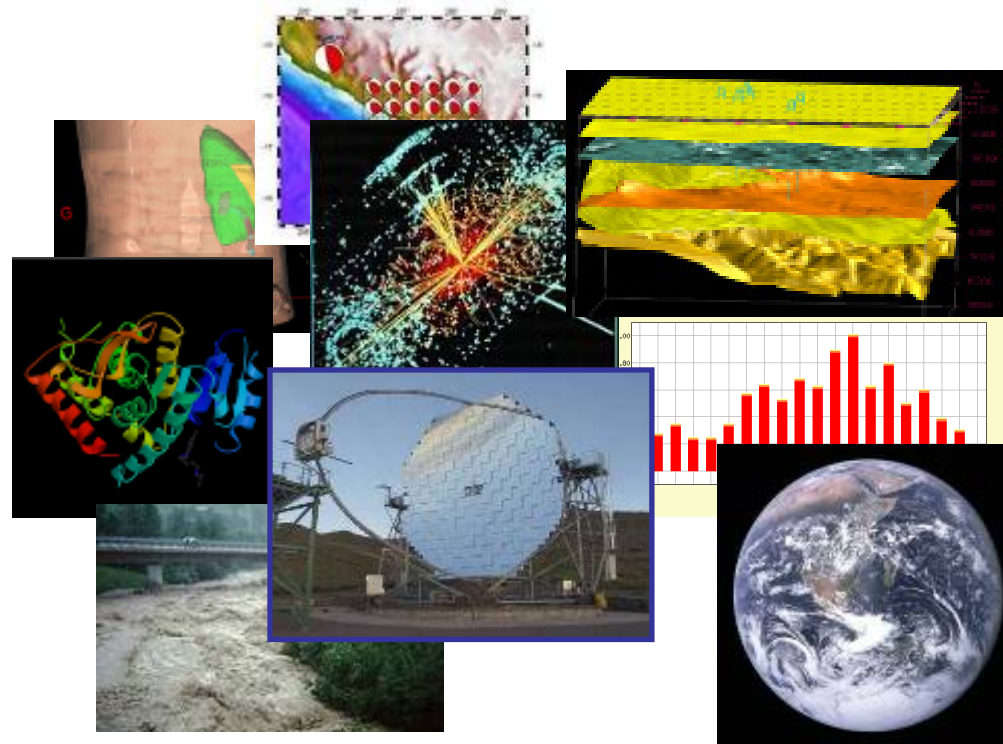# DEISA –Distributed European Infrastructure for Supercomputing Applications

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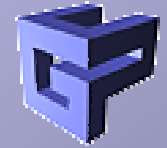
- **>200 VOs from several scientific domains**
  - Astronomy & Astrophysics
  - Civil Protection
  - Computational Chemistry
  - Comp. Fluid Dynamics
  - Computer Science/Tools
  - Condensed Matter Physics
  - Earth Sciences
  - Fusion
  - High Energy Physics
  - Life Sciences
- **Further applications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

No. jobs / month - all

98k jobs/day

OPS
Nor-LHC
LHC

Applications have moved from testing to routine and daily usage
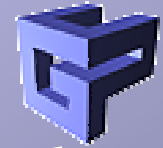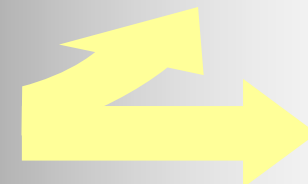
~80-90% efficiency

European Grid Initiative

Goals:

- 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 Design Study (EGI_DS)

- Project Proposal, submitted to the European Commission for funding within **FP7-INFRASTRUCTURES-2007-1**, 1.2.1 Design Studies (May 2, 2007)

Participant no.	Participant organisation name	Short name	Country
1 (Coordinator)	Institut für Graphische und Parallele Datenverarbeitung der Johannes Kepler Universität Linz	GUP	A
2	Greek Research and Technology Network – GRNET S.A.	GRNET	GR
3	Istituto Nazionale di Fisica Nucleare	INFN	I
4	CSC – Scientific Computing Ltd.	CSC	FI
5	CESNET, z.s.p.o.	CESNET	CZ
6	European Organization for Nuclear Research	CERN	CH
7	Verein zur Förderung eines Deutschen Foschungsnet-zen – DFN-Verein	DFN	D
8	Science & Technology Facilities Council	STFC	UK
9	Centre National de la Recherche Scientifique	CNRS	F

New HPC Ecosystem is being built…

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

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**

➢ **Target to establish an active European community for HPC: infrastructure, resource sharing, communication and collaboration over country borders**

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

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, although the prototypes for tomorrow's petaflop systems are usually today's production systems
    - **Innovation is more in scalable software than hardware**

CSC

What is there for Nordic Countries?

Opportunities…

➢ **… can be used only with a strong national infrastructure and competence**

➢ **Skills to utilize top resources**

➢ **High level national services to make us appealing**

➢ **Resources which can be changed for petaflop cycles**

➢ **Nordic people have a very good reputation in EU-projects**

➢ **Focus on where the competence will be concentrated – domestic or outside of home country?**

CSC

More opportunities

➢ **Focus on workload sharing where applicable: HPC, Grid, Network, …**

- Virtual teams based on existing national centers or other organizations

➢ **Synergy for ESFRI Roadmap projects**

➢ **Nordic Grand Challenge Syrvey – continuous process?**

➢ **Nordforsk calls for proposals**

➢ **Who is willing to fund infrastructure in addition to the research? Some research can not be done without high-level infrastructure.**

CSC

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 issue**

➢ **Nordic countries have a great opportunity to benefit, in case we can adapt the game to the different Nordic governance systems**

CSC