Distributed European Infrastructure for Supercomputing Applications

DEISA- The Distributed European Infrastructure for Supercomputing Applications

www.deisa.eu

Erwin Laure
PDC-HPC, erwinl@pdc.kth.se
(based on slides provided by Stefan Heinzel)

NSC09

October 13, 2009, Linköping

SEVENTH FRAMEWORK PROGRAMME
e-infrastructure

RI-222919

European HPC Strategy

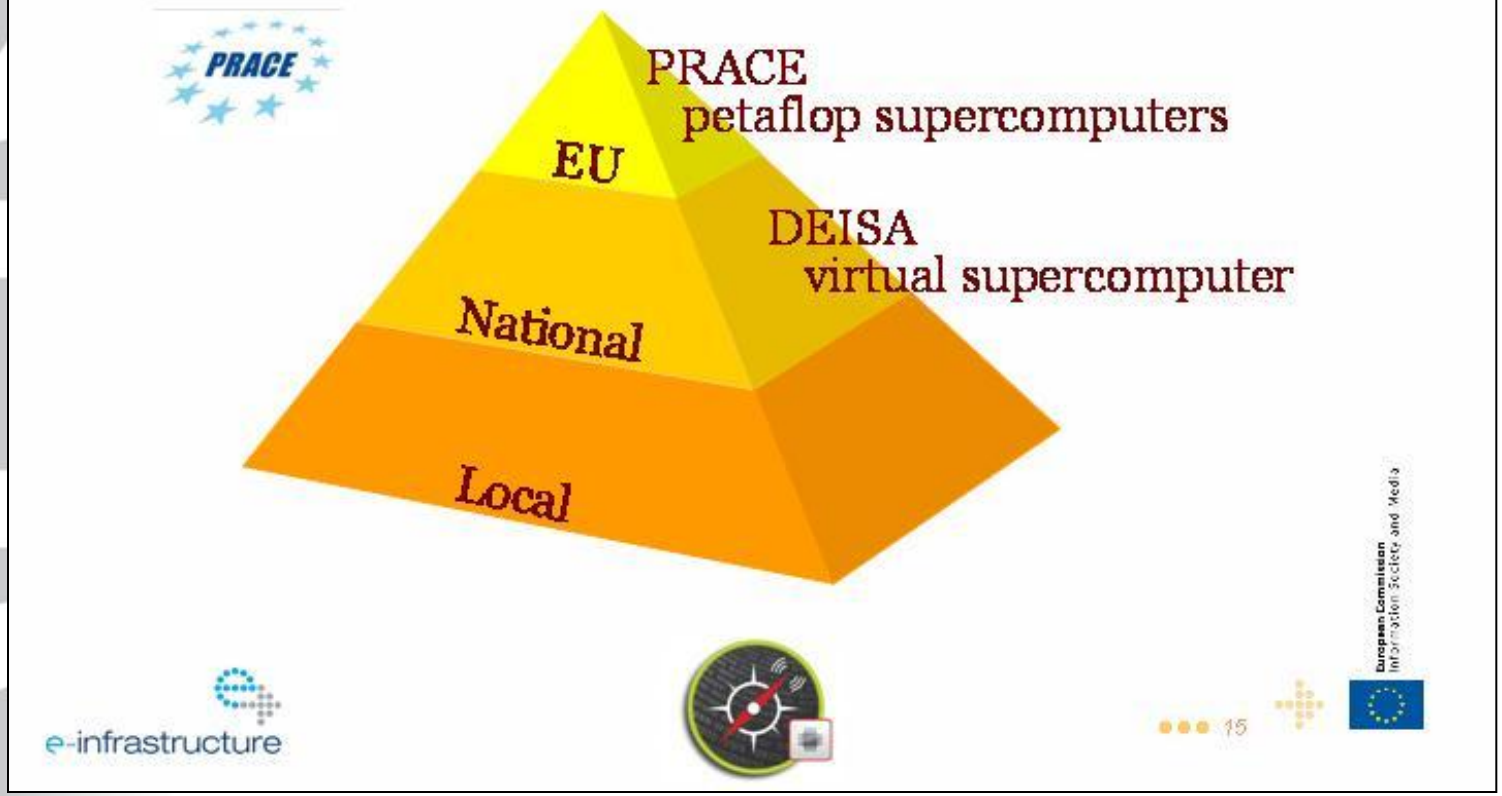
Distributed European Infrastructure for Supercomputing Applications

ESFRI Report 2006, p. 65

European High-Performance Computing Service

A European strategic approach to high-performance computing, concentrating the resources in a limited number of world top-tier centres in an overall infrastructure connected with associated national, regional and local centres, forming a scientific computing network to utilise the top-level machines.

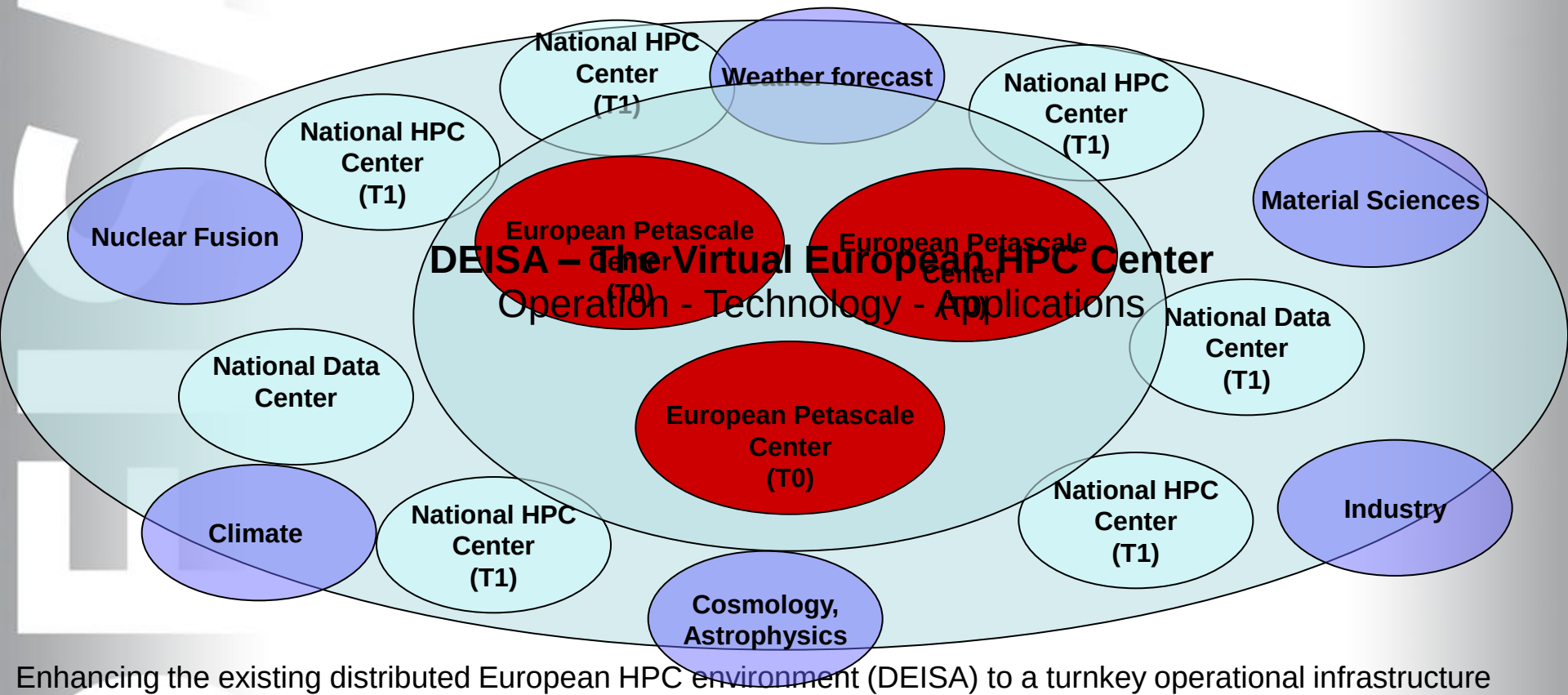
new "petaflop" supercomputers



Mario Campolargo, OGF23, June 2008

Vision and Strategy

Distributed European Infrastructure for Supercomputing Applications



Enhancing the existing distributed European HPC environment (DEISA) to a turnkey operational infrastructure

Advancing the computational sciences in Europe by supporting user communities and extreme computing projects
Enhancing the service provision by offering a complete variety of options of interaction with computational resources

Integration of T-1 and T-0 centres with a transparent access from and into the national data repositories

DEISA Partners and Associate Partners

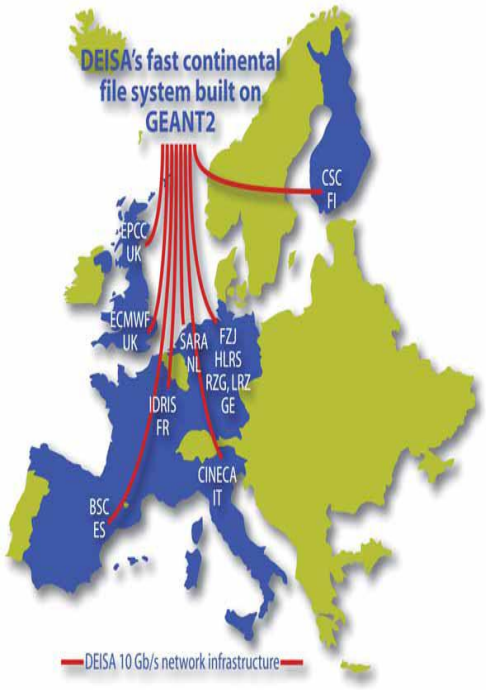
Distributed European Infrastructure for Supercomputing Applications

DEISA: May 1st, 2004 – April 30th, 2008

DEISA2: May 1st, 2008 – April 30th, 2011

DEISA 2008/2009
Operating the European HPC Infrastructure

Distributed European Infrastructure for Supercomputing Applications



>1 PetaFlop/s
Aggregated peak performance

Most powerful European supercomputers for most challenging projects

Top-level Europe-wide application enabling

Grand Challenge projects performed on a regular basis

Community Support

JÜLICH FORSCHUNGSZENTRUM

CINECA Consorzio Interuniversitario

SEVENTH FRAMEWORK PROGRAMME e-infrastructure

Supercomputing Resources

Distributed European Infrastructure for Supercomputing Applications

2009

DEISA partners resources :

11 DEISA partner sites, including 12 of the Top 100 most powerful supercomputers in the world

Higher than 1 PF aggregated Peak performance on state-of-the art supercomputers

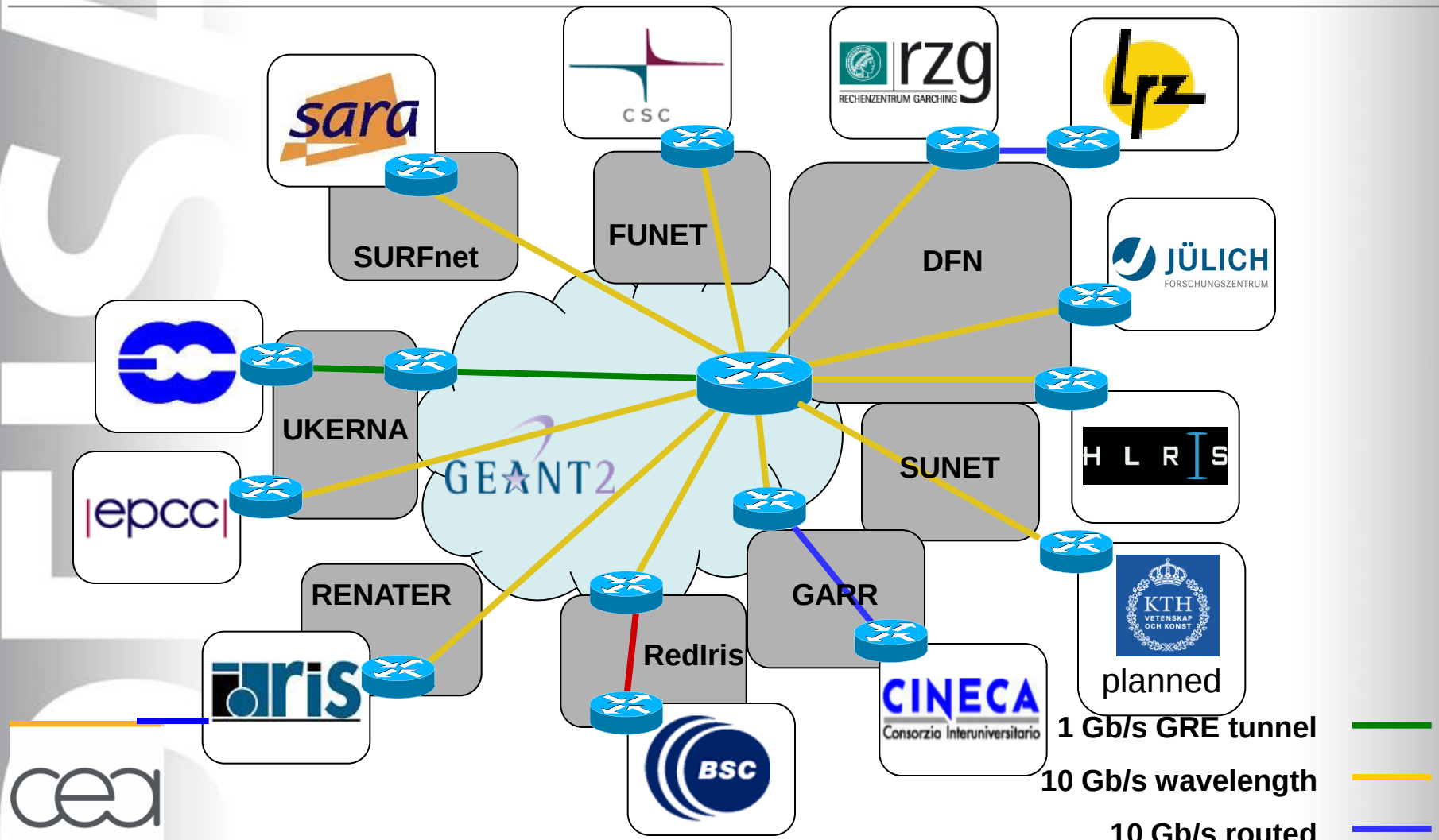
Cray XT4/XT5 Linux
Power5, Power6, AIX / Linux
IBM BlueGene/P, Linux
IBM PowerPC, Linux (MareNostrum)
SGI ALTIX 4700 (Itanium2 Montecito), Linux
NEC SX8 vector system, Super UX

PDC is working to integrate Dell PowerEdge (Ekman) and the SNIC PRACE Prototype

RI-222919

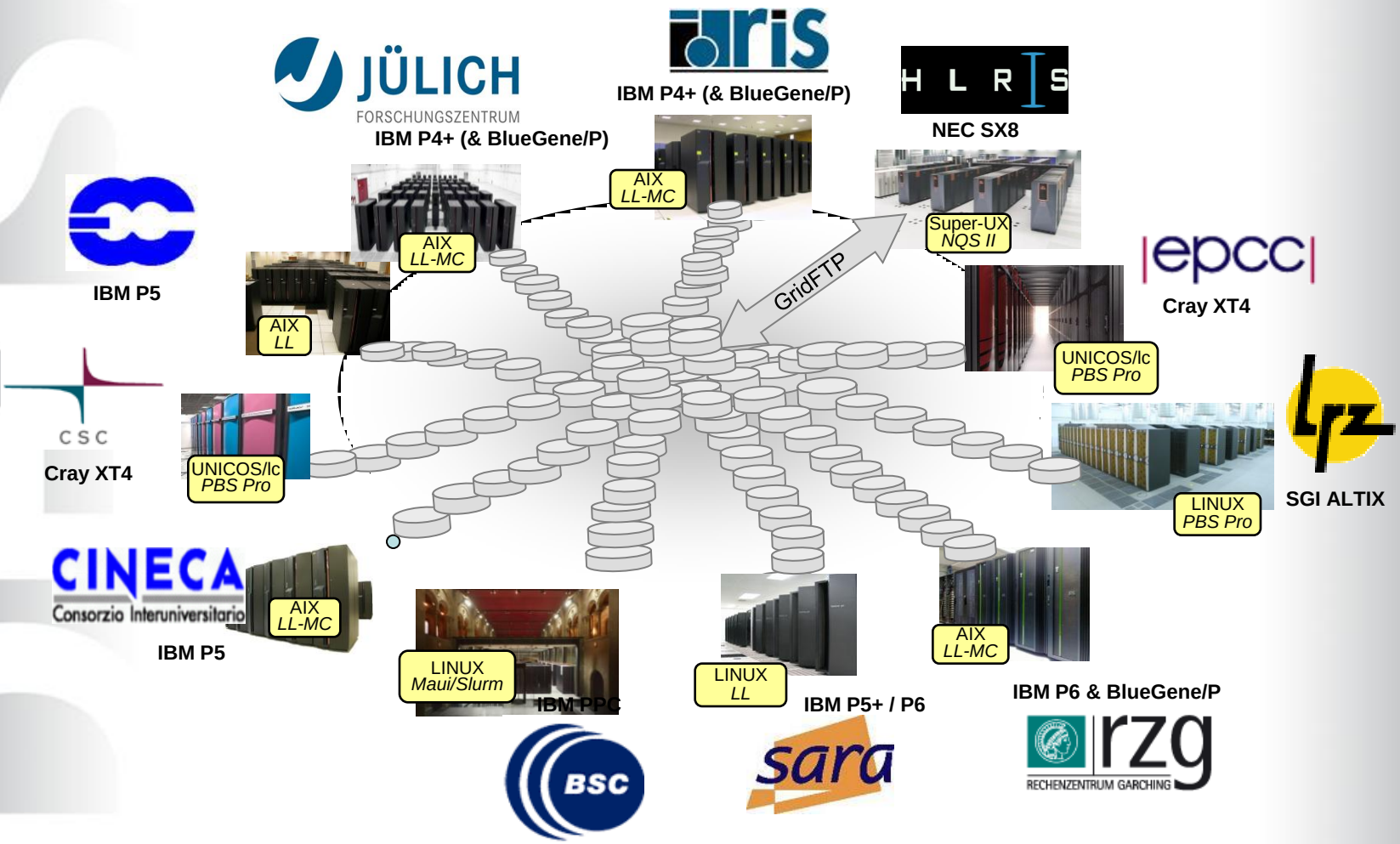
DEISA dedicated high speed network

Distributed European Infrastructure for Supercomputing Applications



DEISA Global File System
(based on MC-GPFS)

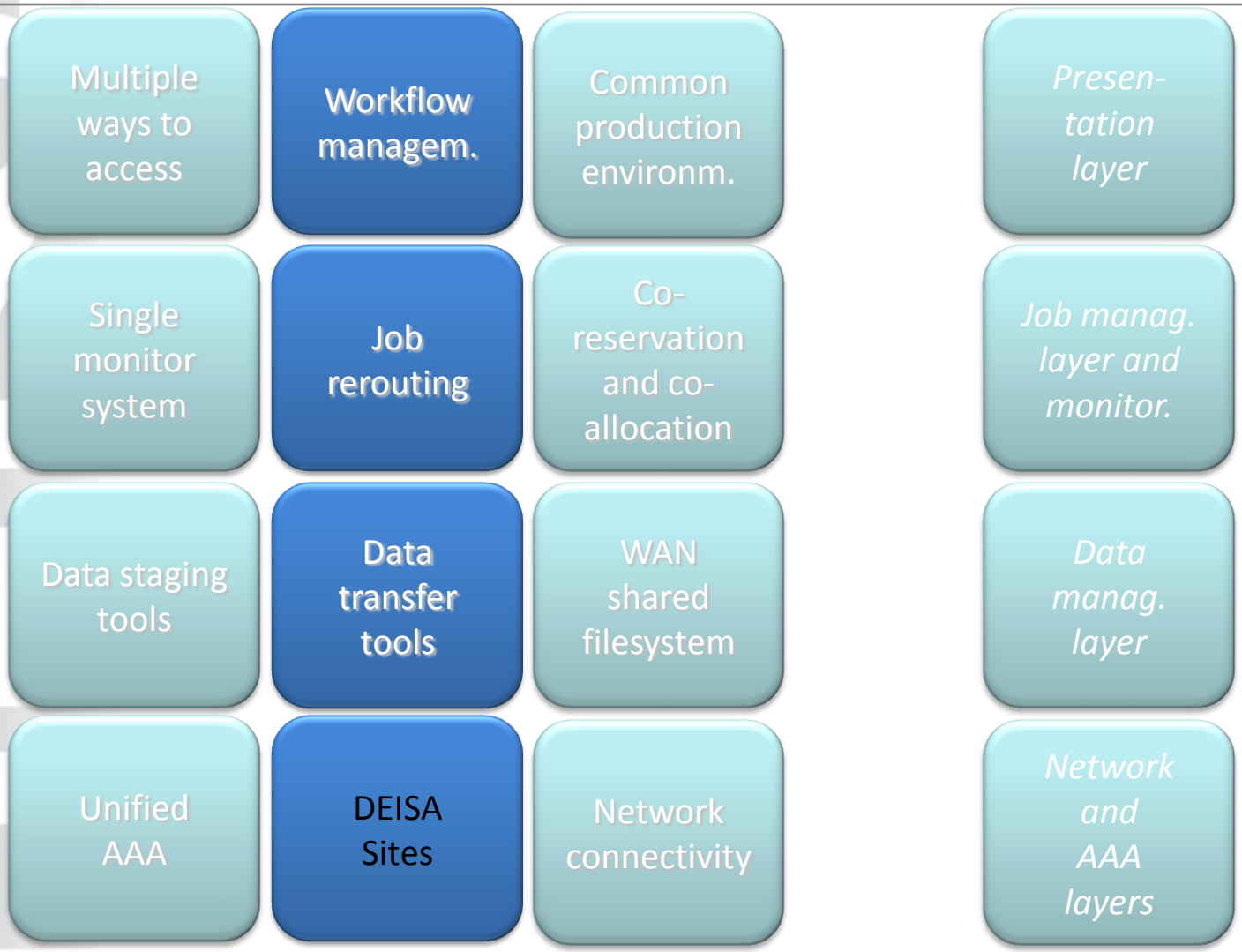
Distributed European Infrastructure for Supercomputing Applications



RI-222919

DEISA Software Layers

Distributed European Infrastructure for Supercomputing Applications



RI-222919

Real needs of HPC users

Distributed European Infrastructure for Supercomputing Applications

HPC users are conservative, standard access methods are preferred, no interest in complicated middleware stacks.

Global Login

- *"HPC users prefer a personal Login in each system"*
  - *Unicore (installed at PDC) and gsi-ssh (pending at PDC)*
  - *LDAP for global user management (installed at PDC)*

Comfortable Data Access

- *"HPC users need a global, fast and comfortable access to their data"*
  - *GPFS (pending at PDC)*

Common Production Environment

- *"HPC users do not need an identical but an equivalent HPC software stack"*
  - *Pending at PDC*

Global Help Desk

- *"HPC users wish one central point of contact and as local as possible"*

Application Support

- "HPC users need help in scalability and adaptation to different architectures"
  - Local support at PDC planned

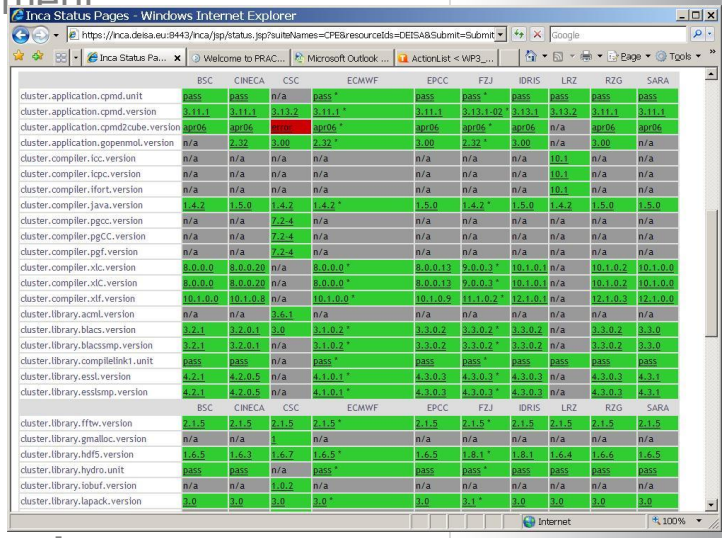
RI-222919

DEISA operational and system Services

Distributed European Infrastructure for Supercomputing Applications

Pan-European operational Services

- INCA monitoring
- Help desk
- Operator on Duty
- Maintenance, installation and configuration management



	BSC	CINECA	CSC	ECMWF	EPCC	FZJ	IDRIS	LRZ	RZG	SARA
cluster.application.cpmd.unit	pass	pass	n/a	pass *	pass	pass *	pass	pass	pass	pass
cluster.application.cpmd.version	3.11.1	3.11.1	3.13.2	3.11.1 *	3.11.1	3.13.1-02 *	3.13.1	3.13.2	3.11.1	3.11.1
cluster.application.cpmd2cube.version	apr06	apr06	error	apr06 *	apr06	apr06 *	apr06	n/a	apr06	apr06
cluster.application.gopenmol.version	n/a	2.32	3.00	2.32 *	3.00	2.32 *	3.00	n/a	3.00	n/a
cluster.compiler.icc.version	n/a	n/a	n/a	n/a	n/a	n/a	n/a	10.1	n/a	n/a
cluster.compiler.icpc.version	n/a	n/a	n/a	n/a	n/a	n/a	n/a	10.1	n/a	n/a
cluster.compiler.ifort.version	n/a	n/a	n/a	n/a	n/a	n/a	n/a	10.1	n/a	n/a
cluster.compiler.java.version	1.4.2	1.5.0	1.4.2	1.4.2 *	1.5.0	1.4.2 *	1.5.0	1.4.2	1.5.0	1.5.0
cluster.compiler.pgcc.version	n/a	n/a	7.2-4	n/a	n/a	n/a	n/a	n/a	n/a	n/a
cluster.compiler.pgCC.version	n/a	n/a	7.2-4	n/a	n/a	n/a	n/a	n/a	n/a	n/a
cluster.compiler.pgf.version	n/a	n/a	7.2-4	n/a	n/a	n/a	n/a	n/a	n/a	n/a
cluster.compiler.xlc.version	8.0.0.0	8.0.0.20	n/a	8.0.0.0 *	8.0.0.13	9.0.0.3 *	10.1.0.1	n/a	10.1.0.2	10.1.0.0
cluster.compiler.xlC.version	8.0.0.0	8.0.0.20	n/a	8.0.0.0 *	8.0.0.13	9.0.0.3 *	10.1.0.1	n/a	10.1.0.2	10.1.0.0
cluster.compiler.xlf.version	10.1.0.0	10.1.0.8	n/a	10.1.0.0 *	10.1.0.9	11.1.0.2 *	12.1.0.1	n/a	12.1.0.3	12.1.0.0
cluster.library.acml.version	n/a	n/a	3.6.1	n/a	n/a	n/a	n/a	n/a	n/a	n/a
cluster.library.blacs.version	3.2.1	3.2.0.1	3.0	3.1.0.2 *	3.3.0.2	3.3.0.2 *	3.3.0.2	n/a	3.3.0.2	3.3.0
cluster.library.blacssmp.version	3.2.1	3.2.0.1	n/a	3.1.0.2 *	3.3.0.2	3.3.0.2 *	3.3.0.2	n/a	3.3.0.2	3.3.0
cluster.library.compilelink1.unit	pass	pass	n/a	pass *	pass	pass *	pass	pass	pass	pass
cluster.library.essl.version	4.2.1	4.2.0.5	n/a	4.1.0.1 *	4.3.0.3	4.3.0.3 *	4.3.0.3	n/a	4.3.0.3	4.3.1
cluster.library.esslsmp.version	4.2.1	4.2.0.5	n/a	4.1.0.1 *	4.3.0.3	4.3.0.3 *	4.3.0.3	n/a	4.3.0.3	4.3.1
	BSC	CINECA	CSC	ECMWF	EPCC	FZJ	IDRIS	LRZ	RZG	SARA
cluster.library.fftw.version	2.1.5	2.1.5	2.1.5	2.1.5 *	2.1.5	2.1.5 *	2.1.5	2.1.5	2.1.5	2.1.5
cluster.library.gmalloc.version	n/a	n/a	1	n/a	n/a	n/a	n/a	n/a	n/a	n/a
cluster.library.hdf5.version	1.6.5	1.6.3	1.6.7	1.6.5 *	1.6.5	1.8.1 *	1.8.1	1.6.4	1.6.6	1.6.5
cluster.library.hydro.unit	pass	pass	n/a	pass *	pass	pass *	pass	pass	pass	pass
cluster.library.iobuf.version	n/a	n/a	1.0.2	n/a	n/a	n/a	n/a	n/a	n/a	n/a
cluster.library.lapack.version	3.0	3.0	3.0	3.0 *	3.0	3.1 *	3.0	3.0	3.0	3.0

Pan European User Environment

- Common Production Environment
  - Adaptations for new architectures
  - **Integration of 7 new HPC systems and technologies in the last 10 months**

DEISA's PAN-European Application Group

"Europe-wide top-level support for applications enabling"

Distributed European Infrastructure for Supercomputing Applications

Current Scope

- DEISA Extreme Computing Initiative (DECI)
  - Grand Challenge projects performed on a regular basis
- Virtual Community Support
  - Dedicated community support for Fusion, Climate, Astrophysics and Material Science

Scientific Areas - From basic research to applied science

- Astro, Fusion, Climate, Life Sciences, Material Sciences and Engineering

Tasks - Many experts from different sites for various needs

- Identification, enabling, deploying and operation of “flagship” applications in many areas of science and technology
- Workflows and coupled applications
- Hyperscaling of huge parallel applications
- Provision European Benchmark Suite for HPC systems

DEISA Extreme Computing Initiative

Distributed European Infrastructure for Supercomputing Applications

Yearly DECI calls launched since 2005 enhancing science and research

Multi-national proposals strongly encouraged to foster European collaboration

Applications are selected on the basis of scientific excellence innovation potential and relevance criteria

Once approved, the most powerful HPC architectures in Euro assigned to the most challenging projects, the most appropriate supercomp architecture selected for each project and the most appropriate experts fo e application support

2 Astro Science projects
Institute for Solar Physics

Projects from DECI calls 2005, 2006, 2007, 2008, 20 .

Involvement of ~ 160 research institutes and uni ersities from 15 European countries

Austria	Finland	France	Germany	Hungary
Italy	Netherlands	Poland	Portugal	Romania
Russia	Spain	Sweden	Switzerland	UK

with collaborators from

four other continents

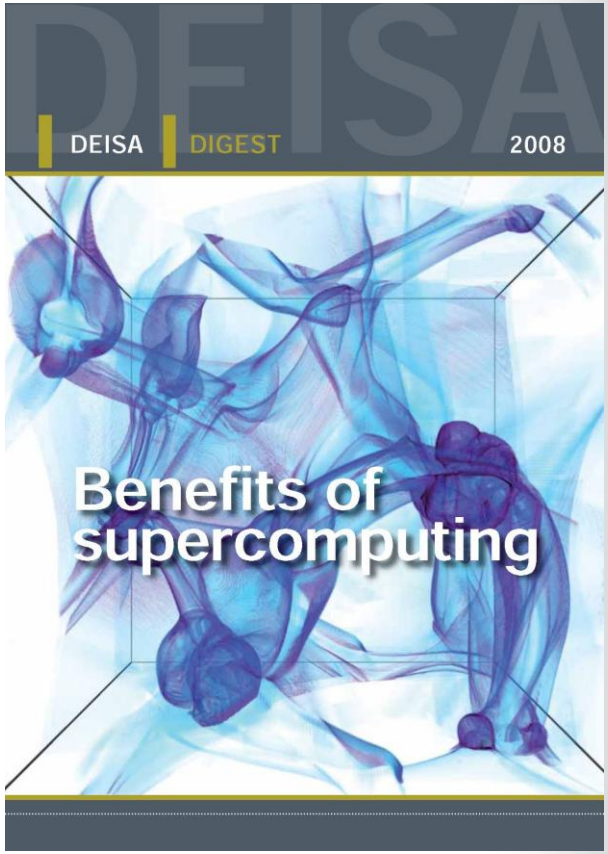
North America, South America, Asia, Australia

Over 100 MCPUh awarded

RI-222919

Achievements and Scientific Impact

Distributed European Infrastructure for Supercomputing Applications

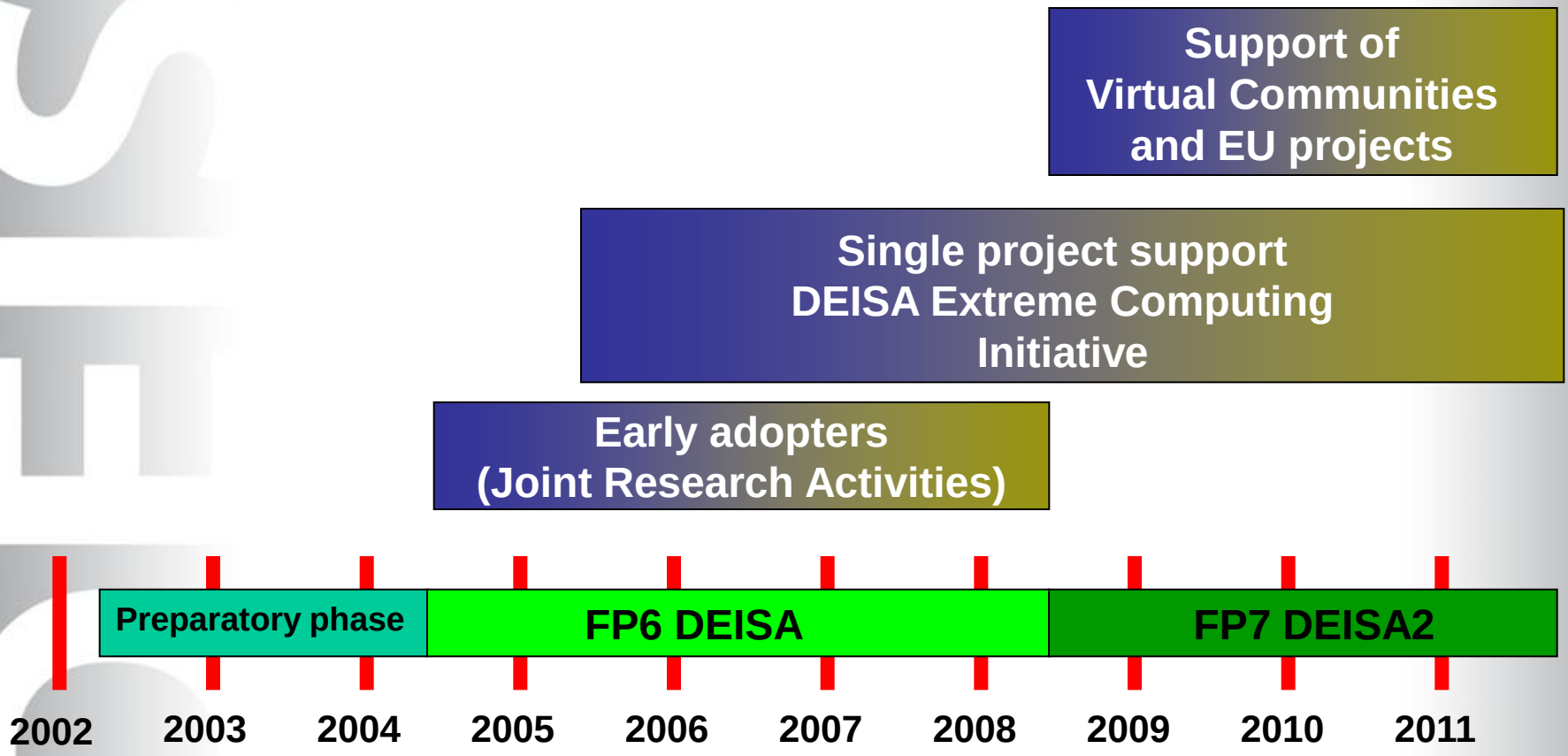


Brochures can be downloaded from http://www.deisa.eu/publications/results

RI-222919

Evolution of user categories in DEISA

Distributed European Infrastructure for Supercomputing Applications



Virtual Community Support

Distributed European Infrastructure for Supercomputing Applications

Fusion energy research:

www.efda.org

European Fusion Development Agreement

www.euforia-project.eu

EU Fusion fOR Iter Applications (EUFORIA)

Life Sciences:

www.virolab.org

Virtual laboratory for infectious diseases

Virtual Community Support

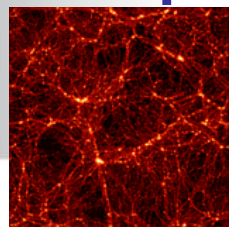
Distributed European Infrastructure for Supercomputing Applications

Earth Sciences:

www.enes.org

European Network for Earth System Modelling

Astrophysics/Cosmology:



http://www.mpa-garching.mpg.de/Virgo/

Virgo Consortium is an international grouping of scientists carrying out supercomputer simulations of the formation of galaxies, ...

http://www.sciops.esa.int/project=PLANCK

Planck is a Mission (M3) of ESA's Horizon 2000 Scientific Programme

Challenges to support Petaflop Applications

Distributed European Infrastructure for Supercomputing Applications

Sustained Petaflop Systems as of 2012

Already designed by 3-6 companies and partly announced for delivering (Blue Waters - IBM Power7, NCSA - 2011; Sequoia - IBM BlueGeneQ, LLNL- 2011; Riken Project - 2012)

Budget, Centres, Local Infrastructure

National budget commitment of different member states
Centres in Europe identified
Local infrastructure being established

Tight European Collaboration

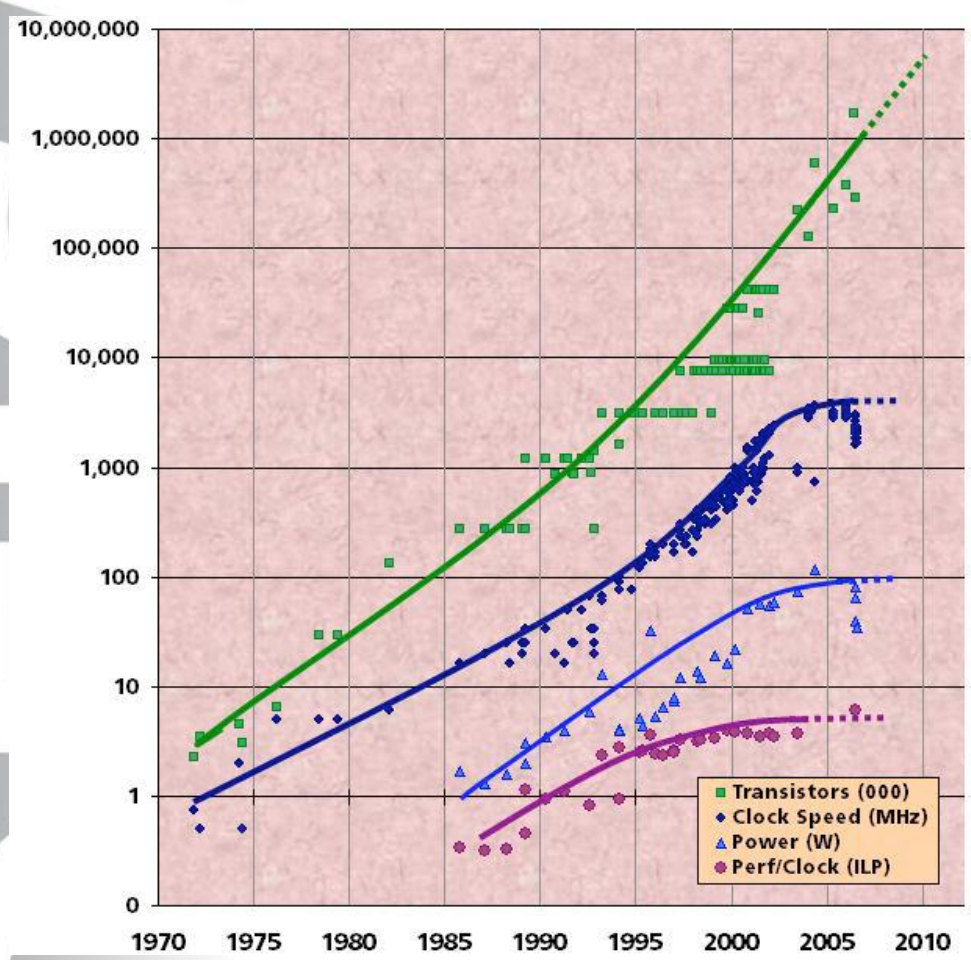
European Wide HPC Infrastructure addressed by DEISA
European Wide HPC Governance Structure and Petaflop Systems addressed by PRACE
European Wide network infrastructure, addressed by GEANT

All these issues seem well addressed!

But how about the progress with Petaflop Applications in Europe?

RI-222919

Moore's law is holding – needs reinterpretation



From Herb Sutter<hsutter@microsoft.com>

Moore's law is holding, in the number of transistors

- Transistors on an ASIC still doubling every 18 months at constant cost
- 15 years of *exponential* clock rate growth has ended

Moore's Law reinterpreted

- Performance improvements are now coming from the increase in the number of cores on a processor (ASIC)
- #cores per chip doubles every 18 months *instead* of clock
- 64-512 threads per node will become visible soon

Amdahl's law exists and implies dramatic problems in the range of 100K - 1M cores

Distributed European Infrastructure for Supercomputing Applications

Million core systems on the horizon

Current Status (10k-200K cores)

BGL@ LLNL 200K, BGP@ANL160K, XT5@ORNL 150K, BGP@Juelich 295K

Status 2011- 2013

200K-1.6 M core range will be achieved

Challenges for the applications on Petaflop systems

Improvement of existing codes will become complex and partly impossible

The use of O(100K) cores implies dramatic optimization effort

New paradigm as the support of a hundred threads in one node implies new parallelization strategies

and

Implementation of new parallel programming methods in existing large applications has not always a promising perspective

---> There is the need for new community codes

Challenges for new Community Codes

Distributed European Infrastructure for Supercomputing Applications

Development of a complex community code takes typically several years by several developers

- Are the structures in the scientific organizations appropriate?
- Do new important community codes mainly come from large research institutions?
- Will we have a monopole of a few community codes?
- How many Petaflop codes are available today?

Application efforts are in the responsibility of science departments

- *Top-down funding* by the EU and national governments for application development?
  - Scientific competition
  - Scientific fundamentals and independence

Overall support/funding for Petaflop Application should address

- A strong synergy between theory and computational science ab initio
- Provision of parallel numerical algorithms/libraries and tools

Summary

Distributed European Infrastructure for Supercomputing Applications

DEISA2 as the vector for the integration of Tier-0 and Tier-1 systems in Europe

To provide a lean and reliable turnkey operational solution for a robust and persistent European HPC ecosystem

Overall support for the enabling of Petaflop Applications via DECI or Community support

Bridging worldwide HPC projects: To facilitate the support of international science communities with computational needs traversing existing political boundaries

Increase usage by Swedish scientists