DEISA- The Distributed European Infrastructure for Supercomputing Applications

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

NSC09
October 13, 2009, Linköping
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

- PRACE
  - petaflop supercomputers
- DEISA
  - virtual supercomputer
- EU
- National
- Local

Mario Campolargo, OGF23, June 2008
Vision and Strategy

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

Erwin Laure, PDC
DEISA Partners and Associate Partners

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

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

Erwin Laure, PDC
DEISA 2008/2009
Operating the European HPC Infrastructure

>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
Supercomputing Resources

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
DEISA dedicated high speed network

- 1 Gb/s GRE tunnel
- 10 Gb/s wavelength
- 10 Gb/s routed
- 10 Gb/s switched

Networks involved:
- sara
- CSC
- RZG
- LRZ
- UKERNA
- SURFnet
- FUNET
- DFN
- GARR
- SUNET
- NEST
- RENATER
- Rediris
- CINECA

DEISA - Distributed European Infrastructure for Supercomputing Applications
DEISA Global File System
(based on MC-GPFS)
DEISA Software Layers

- Multiple ways to access
- Workflow management
- Common production environment
- Presentation layer
- Single monitor system
- Job rerouting
- Co-reservation and co-allocation
- Job management layer and monitor
- Data staging tools
- Data transfer tools
- WAN shared filesystem
- Data management layer
- Unified AAA
- DEISA Sites
- Network connectivity
- Network and AAA layers
Real needs of HPC users

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
DEISA operational and system Services

Pan-European operational Services
- INCA monitoring
- Help desk
- Operator on Duty
- Maintenance, installation and configuration management

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"

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

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 Europe are assigned to the most challenging projects, the most appropriate supercomputer architecture selected for each project and the most appropriate experts for the application support


- Involvement of ~ 160 research institutes and universities 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

2 Astro Science projects
Institute for Solar Physics
Achievements and Scientific Impact

Brochures can be downloaded from http://www.deisa.eu/publications/results
Evolution of user categories in DEISA

Support of Virtual Communities and EU projects

Single project support DEISA Extreme Computing Initiative

Early adopters (Joint Research Activities)

Preparatory phase FP6 DEISA FP7 DEISA2

2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
Virtual Community Support

Fusion energy research:

European Fusion Development Agreement

EU Fusion fOR Iter Applications (EUFORIA)

Life Sciences:

Virtual laboratory for infectious diseases
Virtual Community Support

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

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?
Moore’s law is holding – needs reinterpretation

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

From Herb Sutter<hsutter@microsoft.com>
Amdahl’s law exists and implies dramatic problems in the range of 100K - 1M cores

Million core systems on the horizon
Current Status (10k-200K cores)
BGL@ LLNL 200K, BGP@ANL 160K, 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

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

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