Supercomputing and Grid Joint Use Case

Peter Kunszt
ETH Zürich / Swiss National Supercomputing Centre CSCS

CAPACITIES

e-infrastructure

CSCS
Swiss National Supercomputing Centre

Careful with The Terminology

- High Performance Computing HPC
  - Layman's definition: LINPACK top500 ranking of HW
  - Supercomputing community definition: Very few jobs completely max out a very large resource – resource hungry applications
  - Cluster computing community definition: Usage of powerful modern clusters in general – large number of applications
- Grid Computing
  - Sold as: Ubiquitous computing like electric power from the outlet
  - Advantages praised as: Standard interfaces, apps are not resource-bound
  - Common misconception: It is for free (seti@home)
  - Includes also High Throughput Computing, Collaborative Computing, Cloud Computing..

Terminology suggestion

- Avoid the terms HPC and Grid as they have too many conflicting connotations
- Use
  - **Supercomputing** : Very resource-hungry, highly machine-specific optimized simulations, on **cutting edge systems**
  - **Cluster computing**: Applications able to run on general purpose **mainstream systems** with ever-increasing capacitiy and capability
  - **Collaborative computing**: User domains in need of several **heterogeneous geographically distributed systems**
  - **On-demand computing**: Usage of **third-party resources** on a pay-per-use basis (mostly to deal with spike overload of local resources)
  - **Desktop computing**: Exploitation of **spare desktop cycles**

Terminology suggestion

- Avoid the terms HPC and Grid as they have too many conflicting connotations
- Use
  - **Supercomputing** : Very resource-hungry, highly machine-optimized simulations, on **cutting edge** systems
  - **Cluster computing**
  - **Grid computing**: Usage of **third-party resources** on a pay-per-use basis (mostly to deal with spike overload of local resources)
  - **Desktop computing**: Exploitation of **spare desktop cycles**

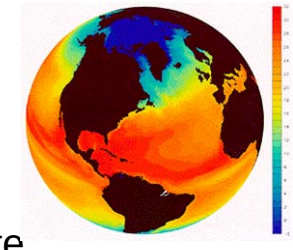
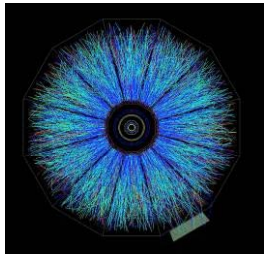
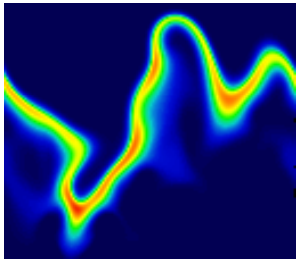
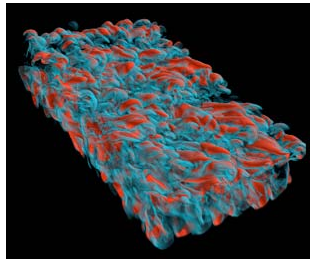
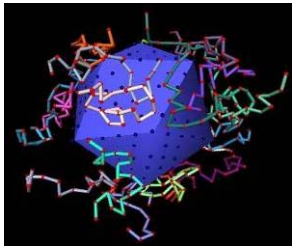
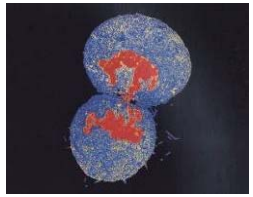
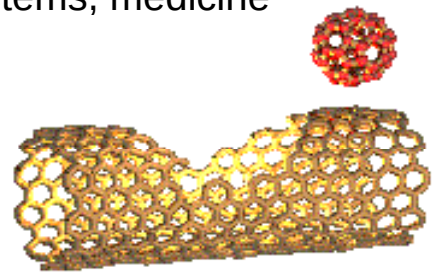
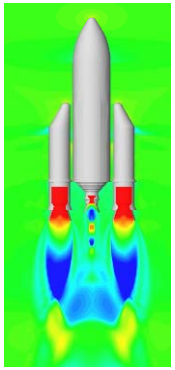
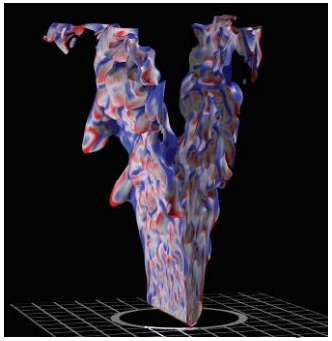
EGI has to consider all of these COMPLEMENTARY resources

Terminology suggestion

- Avoid the terms HPC and Grid as they have too many conflicting connotations
- Use
  - **Supercomputing** : Very resource-hungry, highly machine-specific optimized simulations, on **cutting edge systems**
  - **Cluster computing**: Applications able to run on general purpose **mainstream systems** with ever-increasing capacitiy and capability
  - **Collaborative computing**: User domains in need of several **heterogeneous geographically distributed systems**
  - **On-demand computing**: Usage of **third-party resources** on a pay-per-use basis (mostly to deal with spike overload of local resources)
  - **Desktop computing**: Exploitation of **spare desktop cycles**

This Talk

SC: The Scientific Case

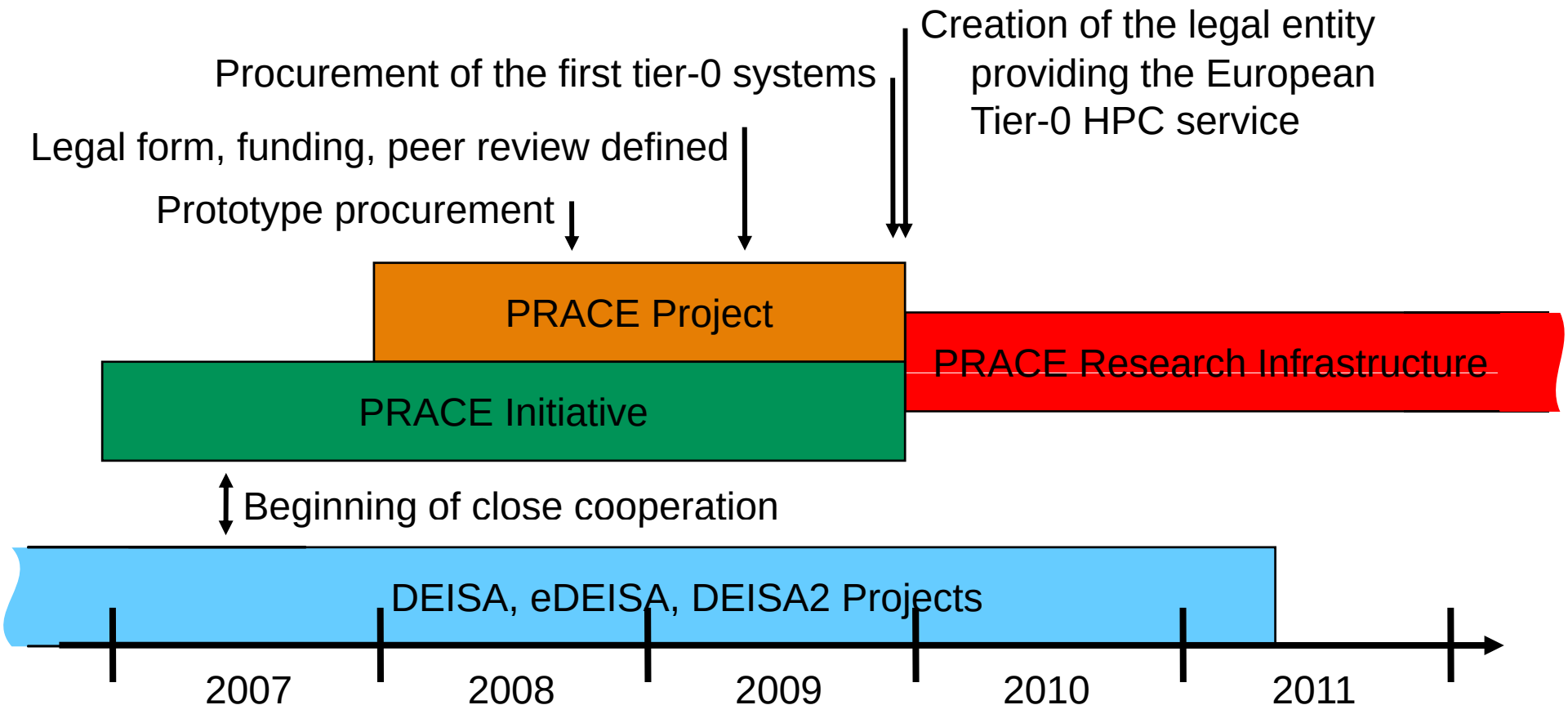
- Weather, Climatology, Earth Science
  - degree of warming, scenarios for our future climate.
  - understand and predict ocean properties and variations
  - weather and flood events
- Astrophysics, Elementary particle physics, Plasma physics
  - systems, structures which span a large range of different length and time scales
  - quantum field theories like QCD, ITER
- Material Science, Chemistry, Nanoscience
  - understanding complex materials, complex chemistry, nanoscience
  - the determination of electronic and transport properties
- Life Science
  - system biology, chromatin dynamics, large scale protein dynamics, protein association and aggregation, supramolecular systems, medicine
- Engineering
  - complex helicopter simulation, biomedical flows, gas turbines and internal combustion engines, forest fires, green aircraft,
  - virtual power plant

The PRACE Mission

- Create a persistent pan-European HPC service
  - Provide European researchers with world-class computing resources
  - Establish the top-level of the European HPC ecosystem involving national, regional and topical HPC centres
  - Deploy several leadership computing systems at selected tier-0 centres
- 14 European Countries participate

  - Several are willing to fund and operate a tier-0 centre

PRACE Roadmap



Supercomputing‘ Use Case Categories

- **Instrument Data Analysis**. There is one or several sources of experimental data to be analyzed.
  - Astrophysics, Geophysics, Satellites, Nanosciences
- **Simulation**. Simulation of large systems
  - Plasma physics, Fluid dynamics, Solid state physics, QCD, Comp.chem, Life Sciences, Climate modeling
- **Time Critical Applications**. Urgent computing
  - Earthquake, tsunami predictions, Weather models
- **Complex Workflows**. Correlation of complex data
  - Life sciences, nanosciences, molecular dynamics

Additional Needs

- **Instrument Data Analysis**. There is c
  sources of experimental data to be an
  - Astrophysics, Geophysics, Satellites, Nanos
- **Simulation**. Simulation of large syste
  - Plasma physics, Fluid dynamics, Solid state
    Comp.chem, Life Sciences, Climate modelin
- **Time Critical Applications**. Urgent c
  - Earthquake, tsunami predictions, Weather m
- **Complex Workflows**. Correlation of c
  - Life sciences, nanosciences, molecular dyna

Collaborative Services

Clusters

Storage

Network

Transfer

Security

Optimization

How people work today

- Projects are formed and scientific proposals are assembled
- Requesting CPU time and storage space – equivalent of a grant
- There is a peer review process that allocates time – based on technical feasability, merit and scientific value and merit
- Project duration 1-3 years
- Many have had long histories, communities that have been working together for years and have established processes
- Very close collaboration with resource providers
- SSH access to machines, direct queue submission

How people work today: DEISA

- 5-10% of site‘s resources dedicated to DEISA open call
- Submission through UNICORE possible but not mandatory
- Possibility to optimize resources, matching the application with the resource
- ‚Local‘ resources are usually dedicated to the groups
- People have dedicated people working on computing issues, scientific computing experts

Differences to current vision

- Project, not VO oriented
  - No continuity beyond that
  - Very clear process how to apply and enter
  - Peer review allocated resource, not ‚taking with you through NGI‘
- Inherently NOT matching the Grid ‚ubiquitousness‘
  - not at all like ‚computing power from the outlet‘
  - Specialized optimization and knowledge necessary to build SC applications
- Supercomputing resources are nonstandard
  - Restricted usage of custom and specialized resources
  - Will rarely adhere to established standards

Questions to address

- SSC
  - What role does it have concerning Supercomputing resources?
- Virtual Organisations
  - What is the motivation to build one? To join one?
  - How is one established and dismantled?
- Project – VO – SSC - NGI – EGI relations
  - Who is responsible for what? Who provides what resource?
- Peer review vs. My own resource vs. NGI resource