Norwegian Infrastructure for Computational Science

Jacko Koster,
Jørn Amundsen (speaker)

UNINETT Sigma

About UNINETT Sigma

- U.S. Is responsible for coordinating the Norwegian e-infrastructure currently operated by the Universities
- Subsidiary of UNINETT
- The e-infrastructure is financed by the Research Council and in-kind contributions from the Universities
- Provides an agreements framework for operating distributed systems as one logical infrastructure
  - The “metacenter” defines a common infrastructure
- Provides common tools for user administration and system monitoring
- Maintains international collaboration on related projects on behalf of Norway

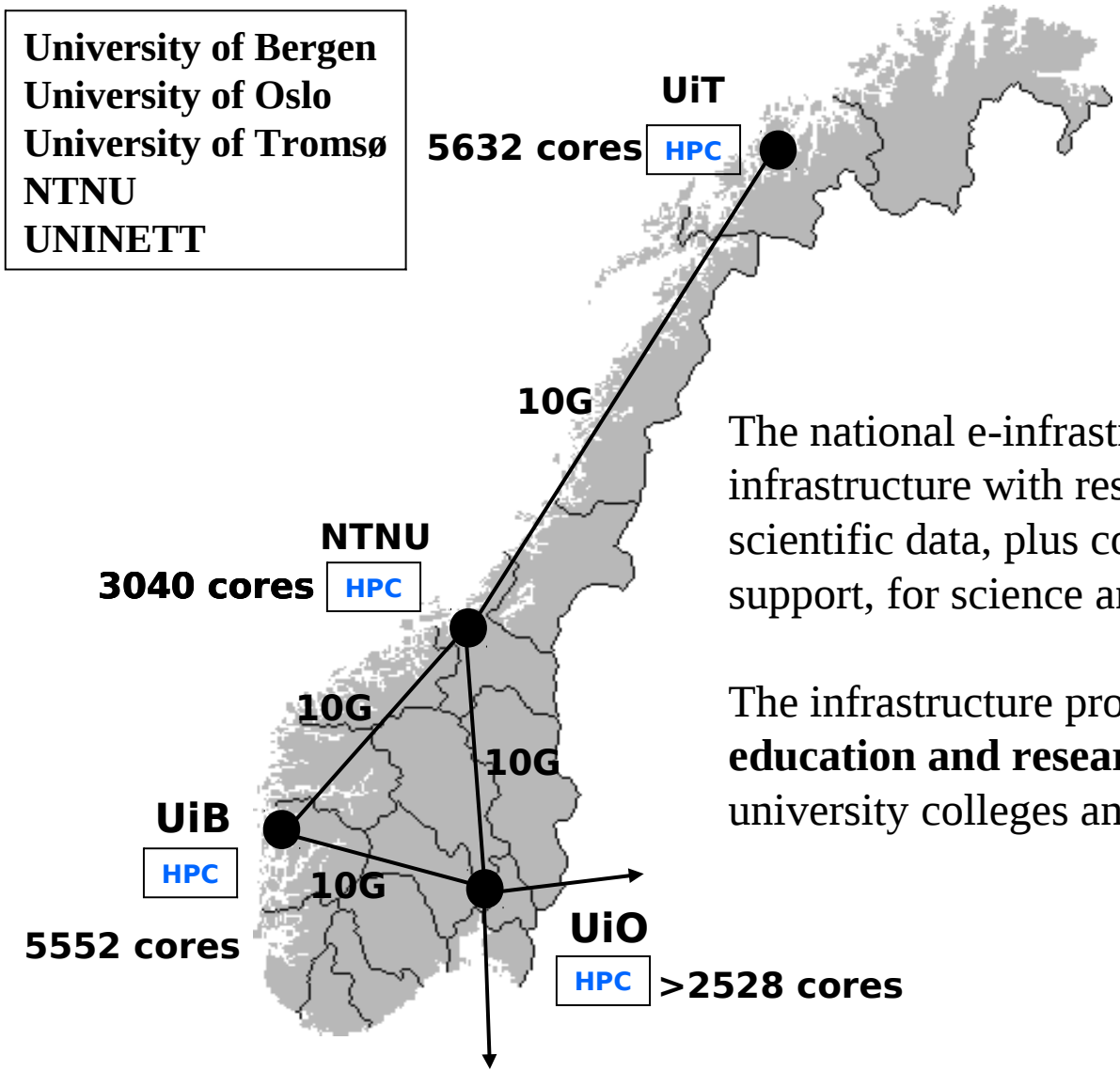
- Resources are for research and education in the U-H sector

Role at the Winter School

- Inform about current and near-future e-infrastructure
- Explain what are your options and what to expect on computational (e-infrastructure) needs in connection with PhD work or other research
- Although not mandatory, make it possible for those interested to do exercises on larger scale or on larger systems
- Assist on (larger scale) computational needs or problems during the winter school week
- Join skilled user communities to learn and to exchange opinions on e-infrastructure needs

National e-infrastructure

University of Bergen
University of Oslo
University of Tromsø
NTNU
UNINETT



HPC: Notur: 2005 → 2014
Data: NorStore: 2007 → 2013

The national e-infrastructure is a **distributed** infrastructure with resources for computation and scientific data, plus corresponding operations and support, for science and research in Norway

The infrastructure provides resources and services for **education and research** at **all** Norwegian universities, university colleges and research institutes

National e-infrastructure

National Infrastructure for High Performance Computing

Notur II - Procurement, operations and support of HPC resources. Heterogeneous resources.

National Infrastructure for Scientific Data

NorStore - Infrastructure for scientific data collections, open for all research environments. Preservation of data.

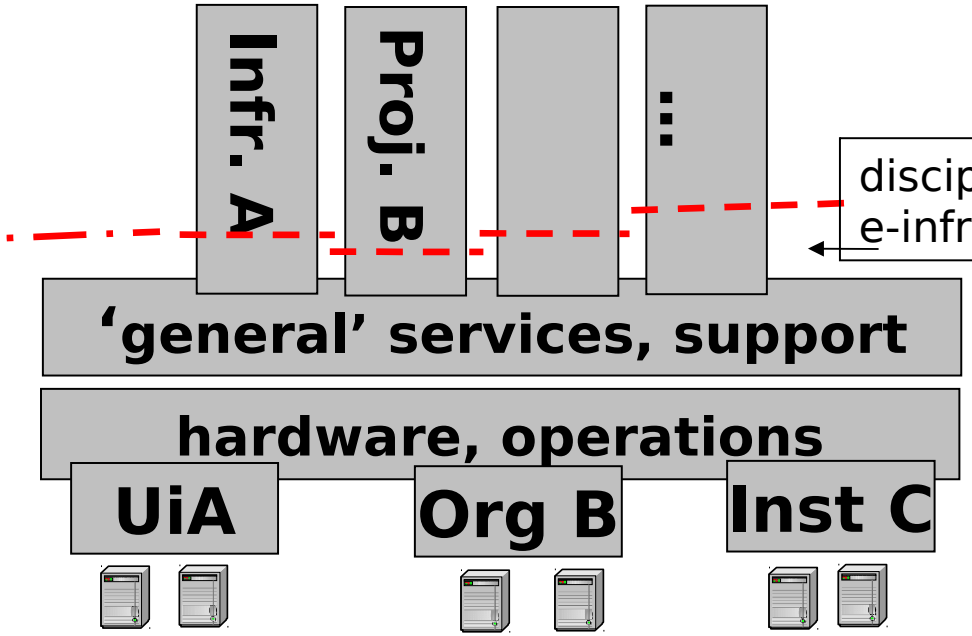
National Grid Initiative

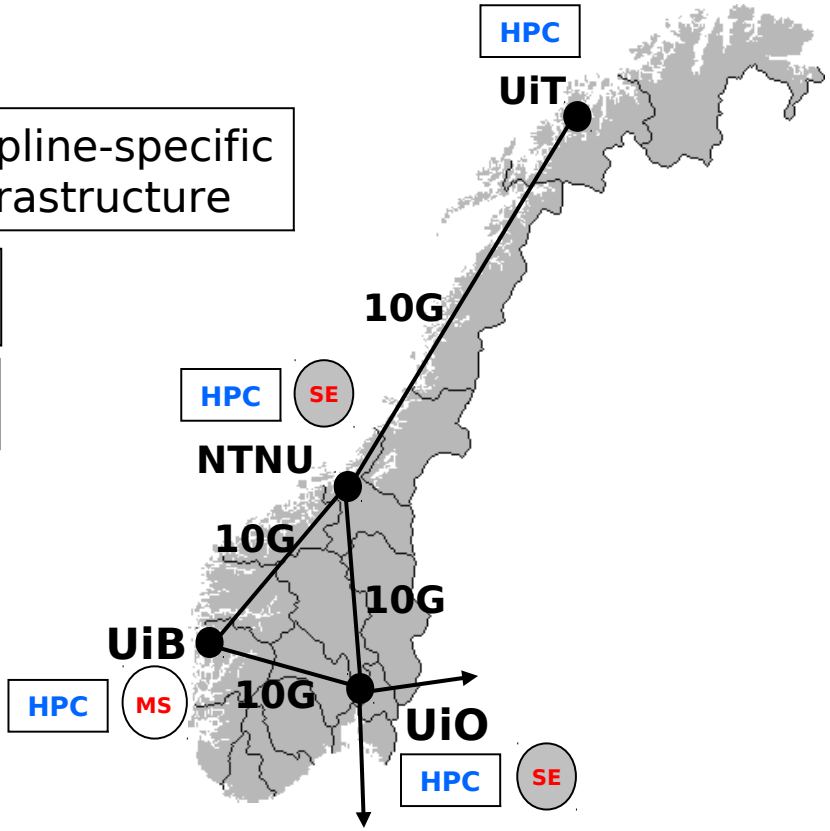
NorGrid - Environments for distributed (virtualized) computing. Coupling of resources in Notur and NorStore, improve utilization of resources, services for distributed data management. Norwegian contribution to WLCG: Nordic Tier-1 and national Tier-2.

UNINETT operates the **national hybrid network** for research and education (10 Gbit/s backbone between 5 universities)

National e-infrastructure

Research projects;
Research Infrastructure





njord.hpc.ntnu.no

IBM p575+, 8-cpu dual-core nodes
2976 cores, power5+ 1.9 GHz
120 TB storage
IBM HPS interconnect
AIX, LoadLeveler
Installed 11/2006, upgrade 11/2009

Electricity usage > 1 MWh (excl. cooling)

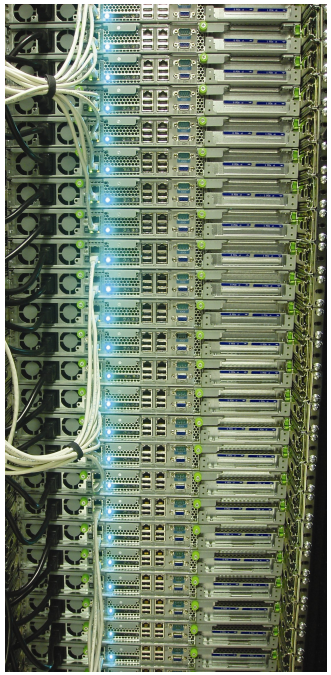
titan.uio.no

Sun x2200, 2-cpu quad-core nodes
2528 cores, Opteron 2.3 GHz
10 TB storage
InfiniBand
Linux CentOS
Upgraded 10/2007

hexagon.bccs.uib.no

Cray XT4, quad-core nodes
5552 cores, Opteron 2.3 GHz
288 TB storage
SeaStar2 interconnect
Linux (Cray variant)
Installed 01/2008

Total 16700 cores, giving ca. 150 million CPU-hours per year.
New resources in 2012.



stallo.uit.no

HP, 2-cpu quad-core nodes
5632 cores, Xeon 2,66 GHz
128 TB storage
InfiniBand (55%)
Linux CentOS
Installed 11/2007

HPC Usage

	hexagon	njord	stallo	titan	ALL
Chemistry	6.5%	2.4%	67.6%	11.5%	**26.6%**
Geosciences	31.4%	5.3%	3.4%	7.5%	**14.3%**
Physics	30.0%	2.9%	12.6%	33.8%	**21.9%**
CFD	1.9%	31.5%	0.9%	0.0%	**4.8%**
Math & Informatics	2.8%	0.0%	4.4%	2.6%	**2.9%**
Biosciences	7.9%	0.0%	1.1%	29.4%	**9.2%**
Structural	0.0%	0.0%	0.0%	0.0%	**0.0%**
Medical	0.4%	1.9%	0.4%	0.3%	**0.6%**
Economics	0.0%	0.0%	0.0%	0.4%	**0.1%**
Materials	5.2%	5.2%	6.7%	2.6%	**5.2%**
Marine Technology	0.0%	16.5%	0.5%	0.0%	**2.2%**
Linguistics	0.0%	0.0%	0.0%	0.8%	**0.2%**
Petroleum	0.0%	0.0%	0.9%	0.0%	**0.3%**
Forecast	0.0%	19.5%	0.0%	0.0%	**2.4%**
Other	13.8%	14.6%	1.4%	11.1%	**9.4%**
SUM (hours)	**39476375**	**13782396**	**35933194**	**23206072**	**112398037**
SUM	100.00%	100.00%	100.00%	100.00%	100.00%

2012 Upgrades

New investments:

All compute resources will be replaced or upgraded in 2012.

NTNU: Vilje, SGI Altix ICE X; installation February 2012; 1440 compute nodes, 23040 cores; 479 TFLOPS; 8-core 2.6 GHz Intel Sandy Bridge, Infiniband FDR14 interconnect

UiB: Hexagon, Cray XE6; upgrade March 2012; 696 compute nodes, 22 272 cores; 205 TFLOPS; 16-core 2.3 GHz AMD Opteron Interlagos, 16-core, Cray Gemini interconnect

2012 Upgrades (2)

UiO: tender published, installation June-August 2012; compute cluster with partitions for parallel computation (InfiniBand interconnect), large memory nodes, large I/O nodes, GPU nodes

UiT: tender published, first installation June-August 2012; compute cluster for parallel computation (InfiniBand)

All systems will run a RPM based Linux distro, either SLES or CentOS.

Access

Access to the national e-infrastructure is by application:

- Two calls for proposals per year. Once can apply for up to four periods (2 years) in a single application.
- Proposals are evaluated by Resource Allocation Committee, appointed by Research Council of Norway.

One can apply for the following resources:

- Time on computing resources (Notur)
- Storage resources and related services (NorStore)
- Application support: porting, application enabling, etc.

The e-infrastructure does not develop application software, this remains the reponsibility of the researcher.

Miscellaneous

Nordic High Performance Computing system

Procurement, operations and support of an HPC system, financed by Denmark, Iceland, Norway, Sweden; Installation in Reykjavik (gardar.nhpc.hi.is); Operational period 2012-2014.

Cloud Computing

Integrate cloud services in a transparent manner in national infrastructure; Both computing (EC2) and storage (S3); Creation of shared user/community spaces with metadata capabilities. Collaboration Norway and Sweden.

Prototype cloud backed storage (WebDAV interface): “plug in the cloud as additional disk in your computer”.

Researchers can apply for access to cloud similar to applying for access to national resources. Pilot usage in 2011.

Miscellaneous

GPU competence project

Duration April 2012 – December 2014.

Includes:
- porting of selected applications to GPU environment
- training of system administrators and end-users
- procurement of GPU system for piloting/production
- ca. 2 FTE's

Prepare for future large procurements including GPU

NorStore

A national infrastructure for **scientific data**

Objective: Develop and maintain national infrastructure for data services storage, publishing, archive, sharing and preservation of scientific datasets over a wide area of disciplines.

Funded in part by Research Council of Norway (2010-2013).

NorStore is a consortium between UiT, NTNU, UiB, UiO and UNINETT Sigma (coordinator).

A major upgrade to the storage infrastructure is currently in progress.

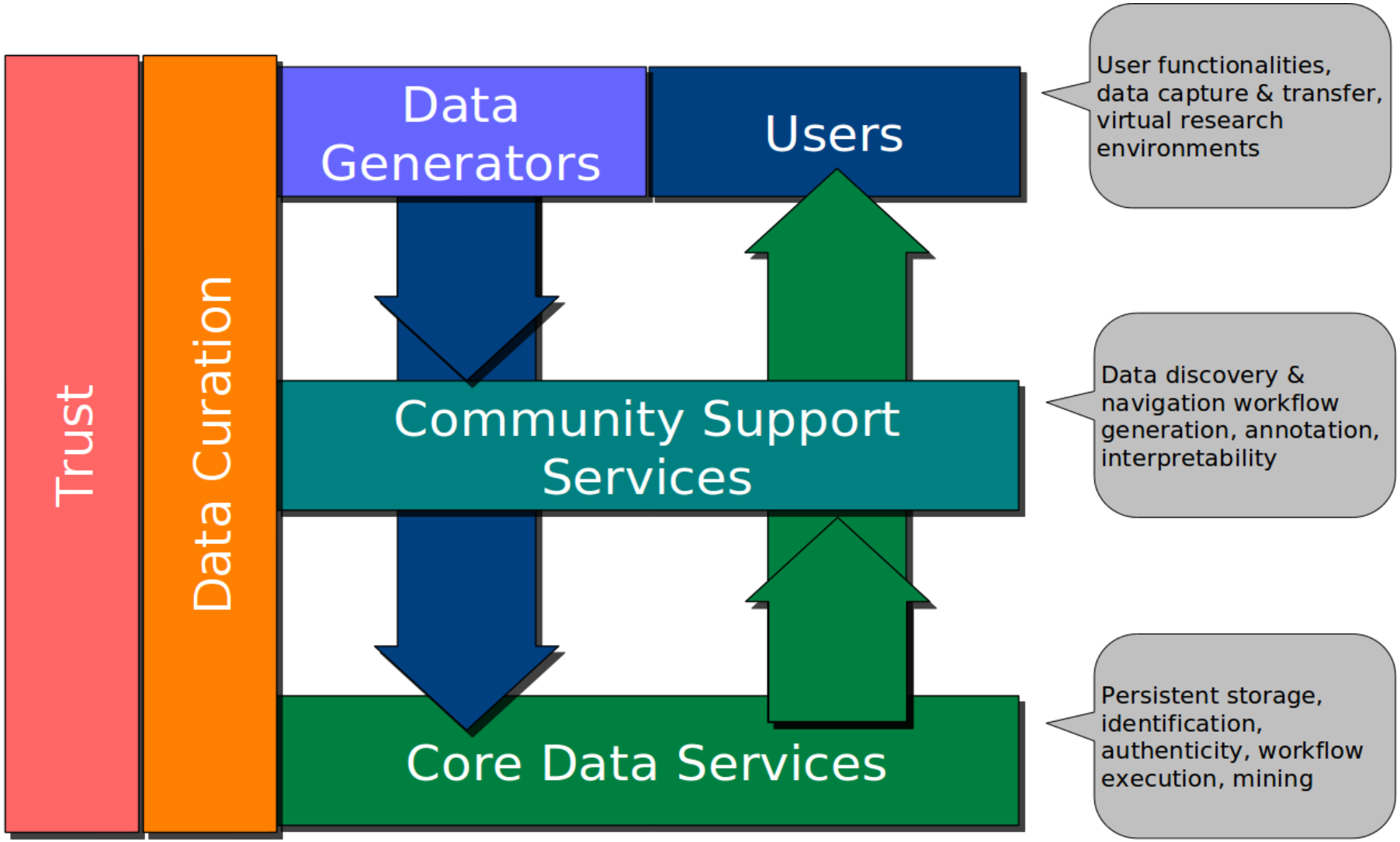


Fig. from “Riding the wave”, EC High Level Expert Group on Scientific Data

NorStore: aims

- develop and operate a national sustainable infrastructure for scientific data

- facilitate services for storing, publishing, sharing and curating digital data across scientific domains, including:

    - long-term preservation of scientific data

    - improving the reusability of scientific data

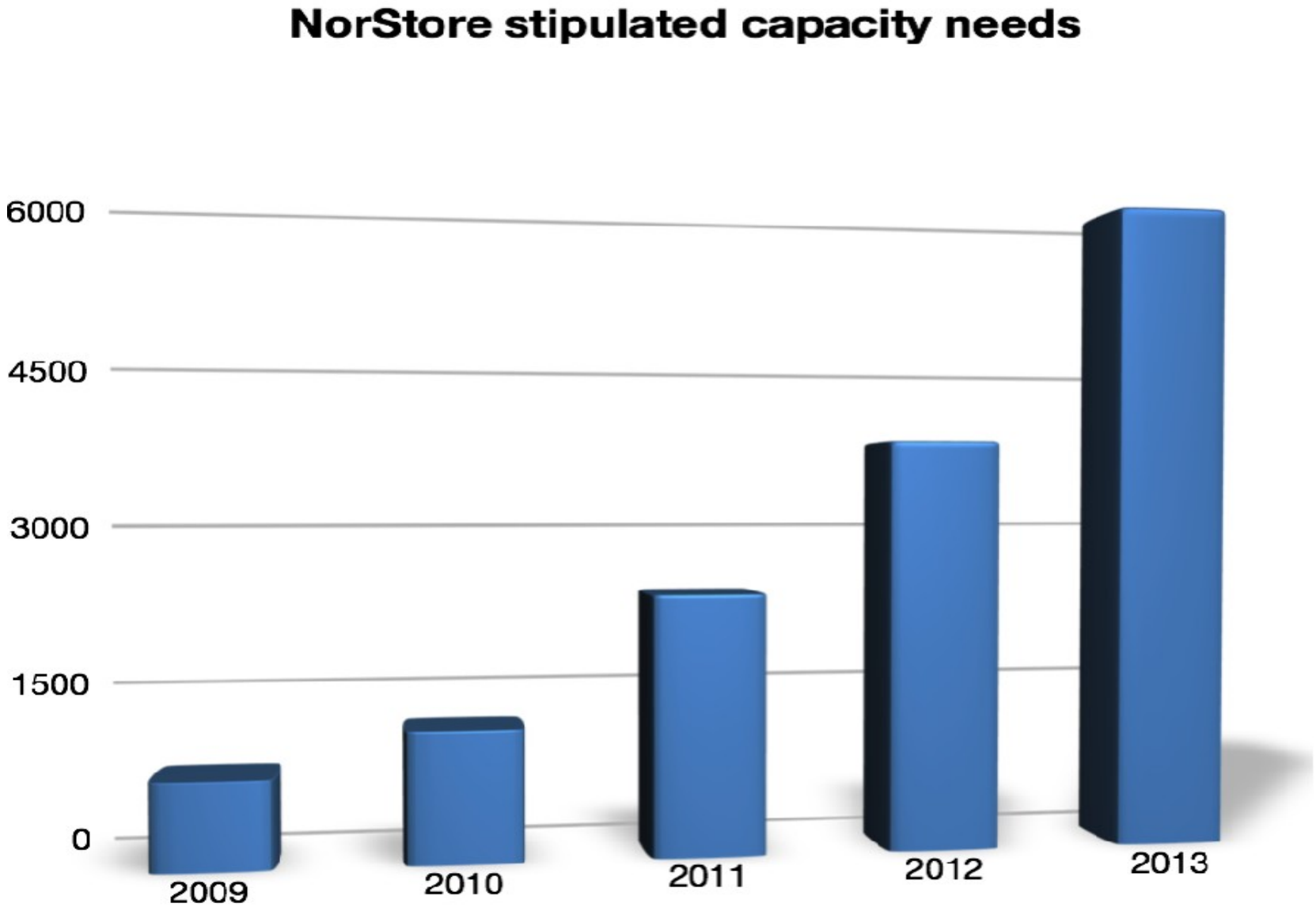
    - facilitate and promote the establishment of digital scientific repositories

- contribute to unification of interfaces to infrastructure for scientific data within Norway and abroad

- ensure cost-efficiency of the infrastructure by national coordination of large-scale investments and operation

- become a leading and highly successful permanent e-infrastructure for a broad range of scientific disciplines

NorStore: aims



Community building

Increased attention on identifying/supporting **community** needs

Establish interfaces with **national communities**:

- national contact group that represents the community
- longer-term (multi-year) advanced user support

Requirements specification from / reporting to the contact group

Commitment from Notur/NorStore and community to activity plan

Implemented for bioinformatics; in progress for climate, linguistics, chemistry; application to be submitted for a national CFD community.

Communities need to cover issues like applications, licensing and storage infrastructure.

Europe

EUROPE

NORWAY

EUDAT **NORSTORE**

EGI **NORGRID**

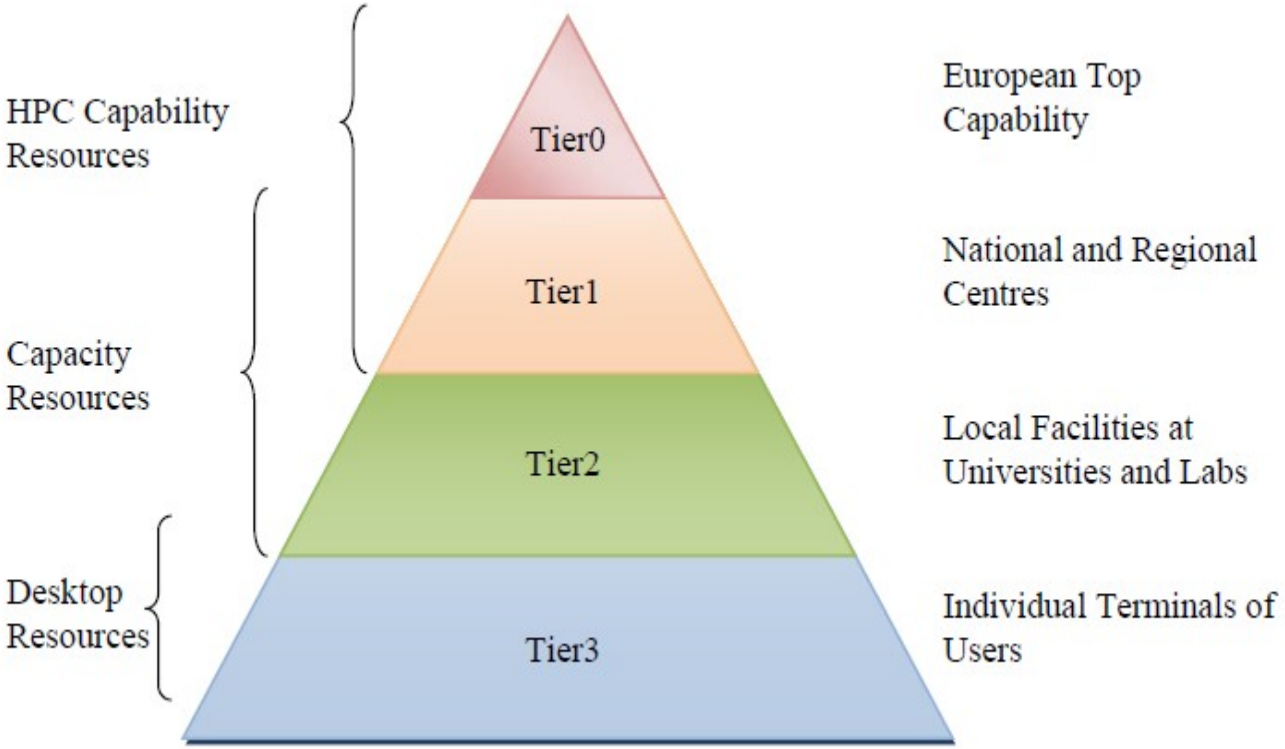
PRACE **NOTUR**

GEANT **UNINETT**

e-infrastructure: ubiquitous research environments for accessing and sharing resources and tools

European HPC Infrastructure

Performance pyramid:



PRACE

DEISA

eGee
Enabling Gri
for E-scienc

eGI

PRACE

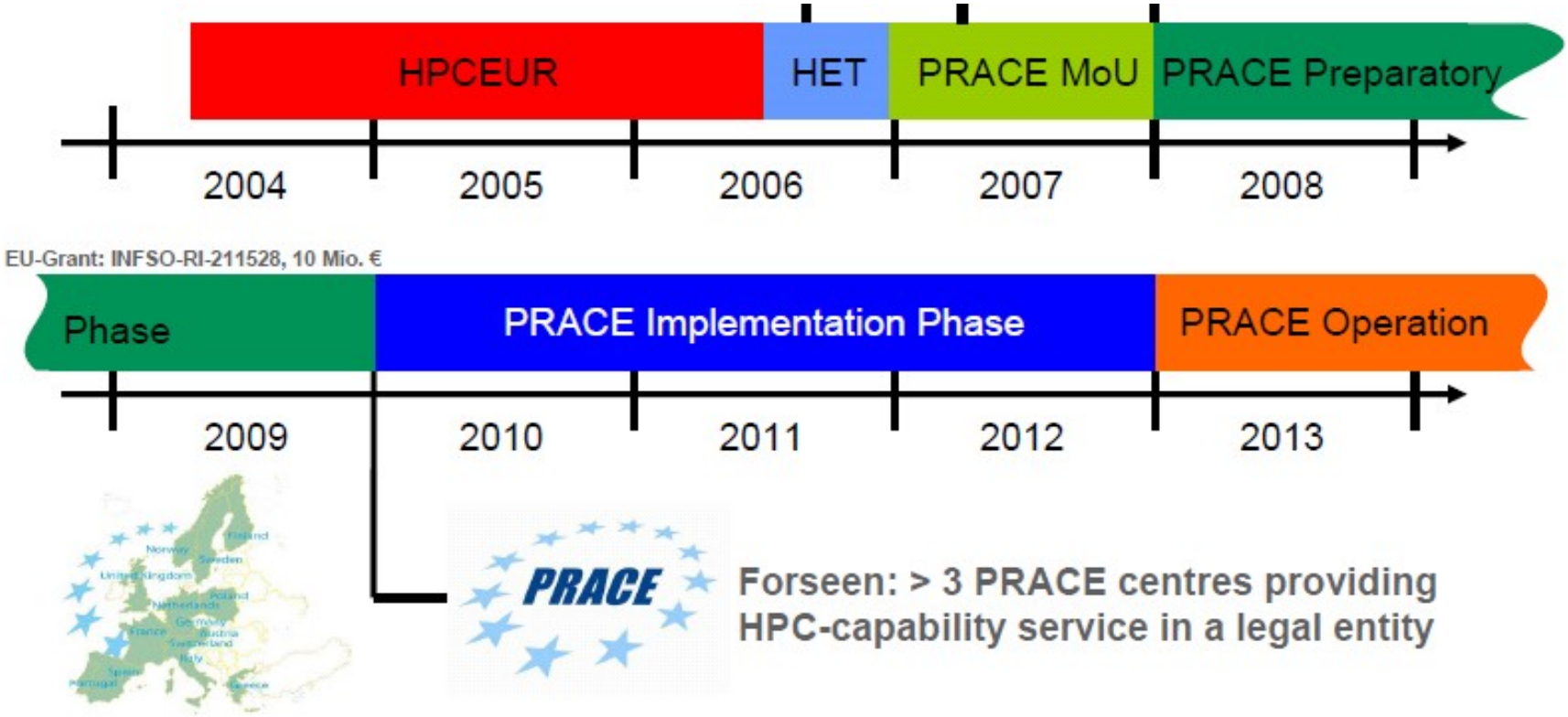
Partnership for Advanced Computing in Europe

PRACE is an ESFRI initiative on large scale HPC resources

Ambition:

- **Mid 2009: 1 Petaflop system in the top 5**
- **Late 2010: 2 Petaflop systems in the top 5; 1 Petaflop system in top 10**
- **2011: over 10 Petaflop in the top 5**
- **2020: the Exaflop in the top 5**

PRACE roadmap



Preparatory Phase: 10 M€ / yr (50% from EU)
Implementation Phase: 100-150 M€ / yr (20 M€ / yr from EU)

PRACE

PRACE-1IP: (07/2010 – 06/2012): establish Tier-0 service
- UiO: scaling of applications for Tier-0, efficient data center design, prototype system (numascale)
- NTNU: scaling of applications for Tier-0

PRACE-2IP: (09/2011 – 08/2013): establish Tier-1 service
- UiO: establish Norwegian Tier-1 service
- NTNU: scaling of applications for Tier-1, new technologies

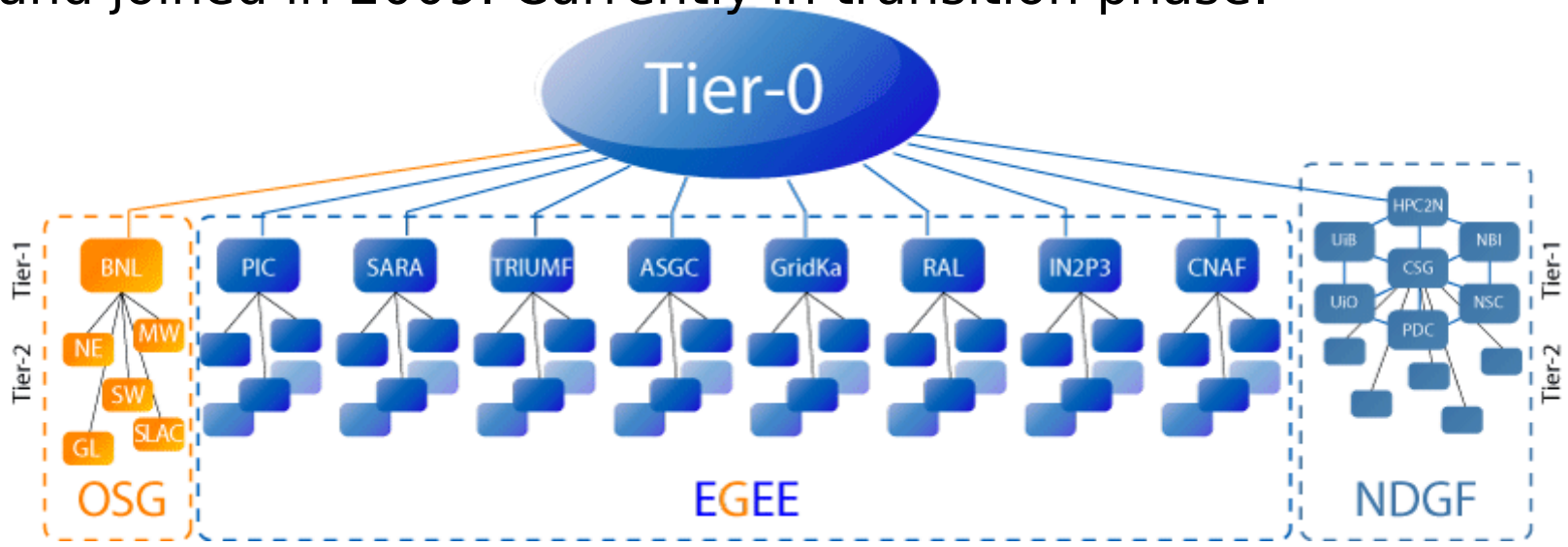
PRACE-3IP: (09/2012 – 08/2014):
- Pre-Commercial Procurement, joint procurement
- Establish relations with industry
- Continuation of PRACE-1IP

Nordic e-infrastructure

Nordic Data Grid Facility (2006-2011, 2012-2015)

Collaboration between Nordic countries, primarily to implement the Nordic contribution to World-wide Large Hadron Collider Grid (WLCG), but also starting to look at other initiatives, e.g., ELIXIR.

Builds on national projects: SweGrid/SNIC, CSC, DCSC, NorGrid. Iceland joined in 2009. Currently in transition phase.



European DCI

EGI is the Distributed Computing Infrastructure for Europe
EGI = egi.eu + NGI's + Europe's largest research organizations
- CERN
- EMBL

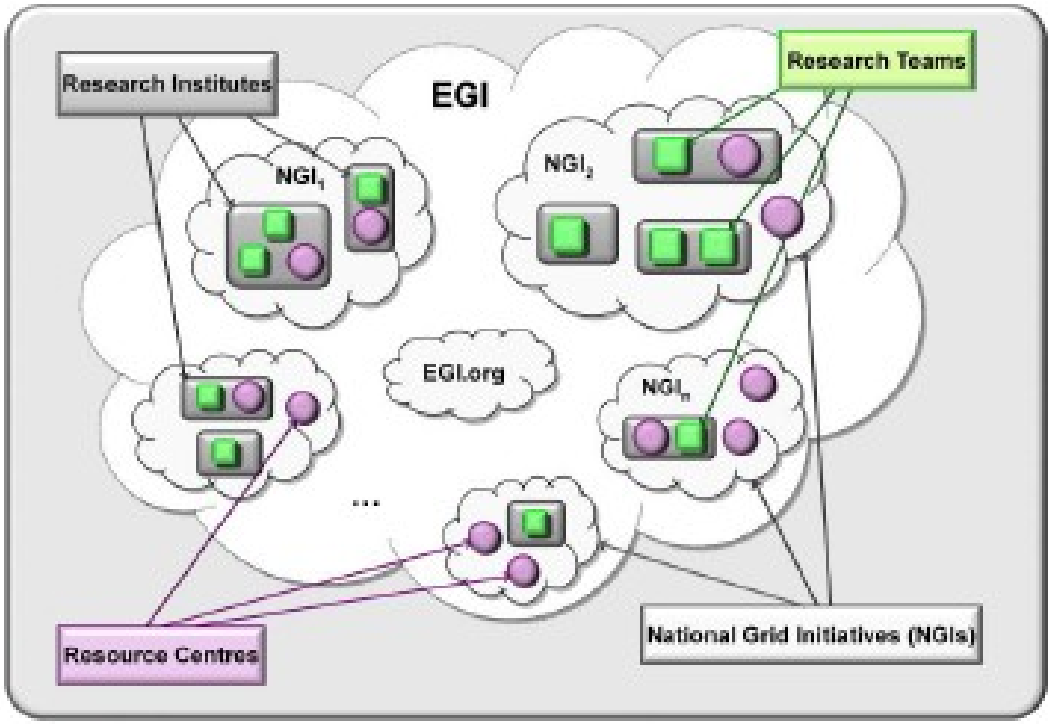


Figure 1: EGI and the NGIs

Specialized support functions are defined on top of EGI, e.g., for communities: HEP, chemistry, material science, fusion, ...

EGI-InSPIRE

Integrated Sustainable Pan-European Infrastructure for Researchers in Europe – egi.eu/projects/egi-inspire

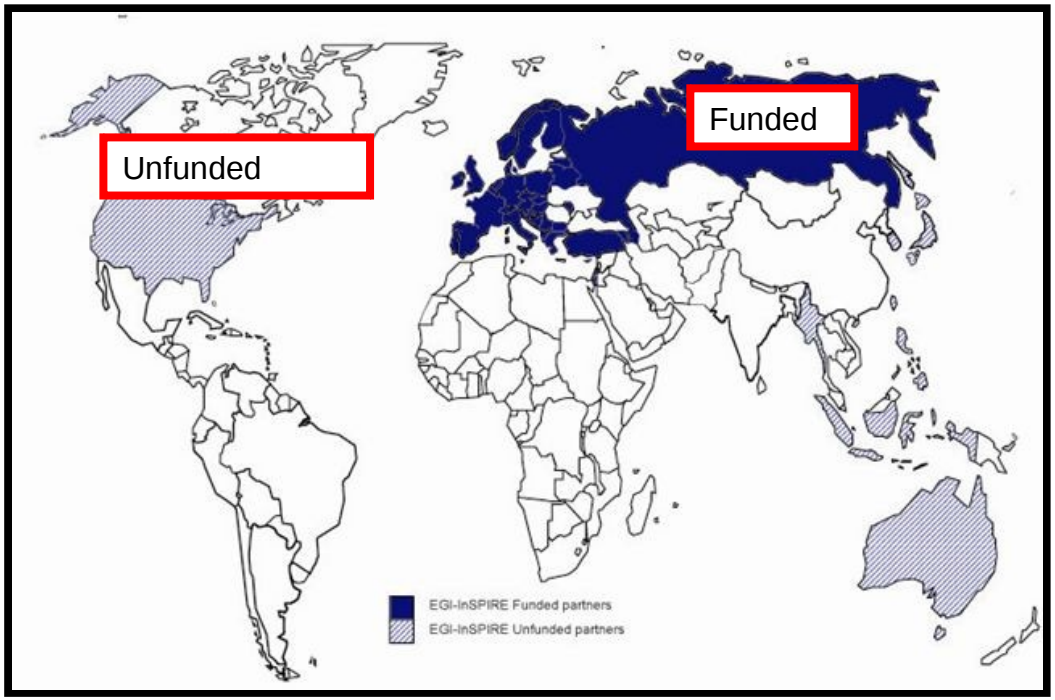
A 4 year project with €25M EC contribution

Project cost: €69M

Total Effort: ~€330M

Effort: 9261 PM

Start: 5/2010



Project Partners (51)

- egi.eu, 40 NGI's, 2 EIRO's
- Asia Pacific (8 partners)

More information

- For more information on Notur and NorStore, visit www.notur.no or www.norstore.no
  - events within the computing community/ies
  - information on how to apply for computing time
  - how to apply for storage resources
- Documentation is on http://docs.notur.no
- Dynamic load is on http://www.notur.no/hardware/status/
- User support is provided by email, to support@notur.no

Computing on Ve

- The SGI Altix ve.hpc.ntnu.no system is a temporary system in the advent of vilje.hpc.ntnu.no
- Runs forecasting production suite in parallel with njord
  => not useful for long jobs, only short test jobs!
- Compute nodes available ~ 7-9, 13-15, 19-21, 23-02
- System running Linux SLES11 with 2 login and 256 compute nodes (8-core 16-way HT, Intel Westmere)
- PBS Pro batch system
  => need to submit a UNIX shell script performing the computation
- Log into the system with ssh ve.hpc.ntnu.no
- User guide on **https://www.hpc.ntnu.no/display/hpc/Ve** (also as pdf file from author)

Computing on Ve (2)

- A compute job is a shell script with initial comment lines for job specification (duration, core count, etc.), see section “Job Execution” and subsection “Sample Batch Scripts” in the UG
- Job control with qsub, qdel and qstat (see UG or man pages, by typing ``man qsub´´ online)
- Notice the modules system is used to enable compilers and software packages on Ve as well as other systems in the e-infrastructure
- ``module load python´´ loads latest Python 2.x and a number of packages, see the UG

- **Sample job scripts to run IFEM will be provided by me / Arne Morten (~arnemort/pack/)**
- (temporary) login accounts on Ve is available by contacting me