Partnership for Advanced Computing in Europe

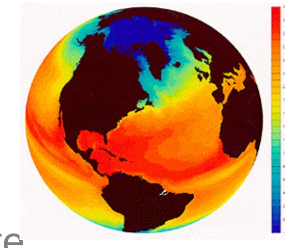
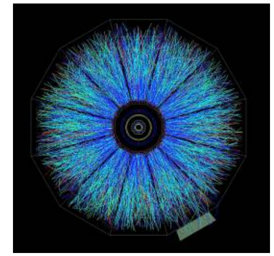
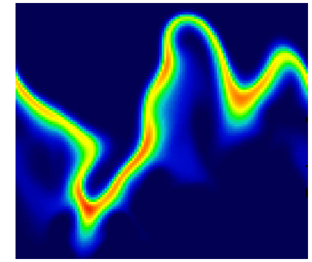
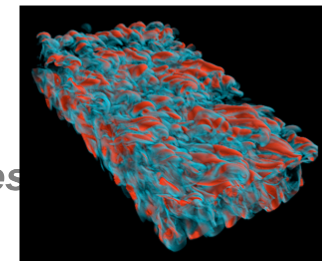
Introduction to HPC, Leon Kos, UL

PRACE Autumn School 2013 - Industry oriented HPC simulations, University of Ljubljana, Slovenia

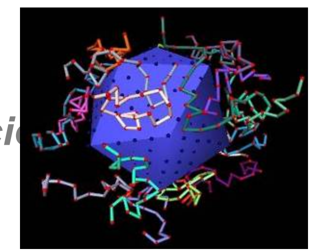
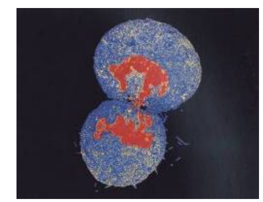
25 members of PRACE

- Germany: GCS - GAUSS Centre for Supercomputing e.V
- Austria: JKU - Johannes Kepler University of Linz
- Belgium: DGO6-SPW – Service Public de Wallonie
- Bulgaria: NCSA - Executive agency
- Cyprus: CaSToRC –The Cyprus Institute
- Czech Republic: VŠB - Technical University of Ostrava
- Denmark: DCSC - Danish Center for Scientific Computing
- Finland: CSC - IT Center for Science Ltd.
- France: GENCI - Grand Equipement National de Calcul Intensif
- Greece: GRNET - Greek Research and Technology Network S.A.
- Hungary: NIIFI - National Information Infrastructure Development Institute
- Ireland: ICHEC - Irish Centre for High-End Computing
- Israel: IUCC - Inter-University Computation Center
- Italy: CINECA - Consorzio Interuniversitario
- Norway: SIGMA – UNINETT Sigma AS –
- The Netherlands: SURFSARA: SARA Computing and Networking Services
- Poland: PSNC – Instytut Chemii Bioorganicznej Pan
- Portugal: FCTUC – Faculdade Ciencias e Tecnologia da Universidade de Coimbra
- Serbia: IPB - Institute of Physics Belgrade
- ***Slovenia: ULFME - University of Ljubljana, Faculty of Mechanical Engineering***
- Spain: BSC – Barcelona Supercomputing Center – Centro Nacional de Supercomputación
- Sweden: SNIC – Vetenskapsrådet – Swedish Research Council
- Switzerland: ETH – Eidgenössische Technische Hochschule Zürich
- Turkey: UYBHM – Ulusal Yuksek Basarimli Hesaplama Merkezi,
- UK: EPSRC – The Engineering and Physical Sciences Research Council

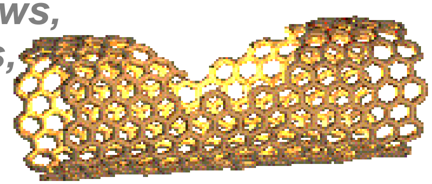
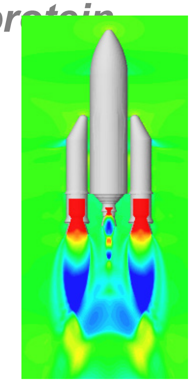
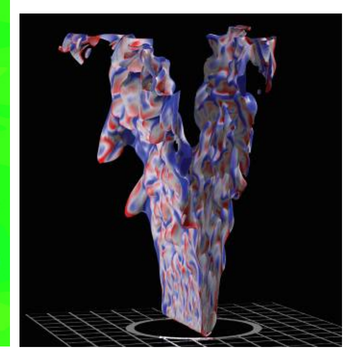
Why supercomputing?

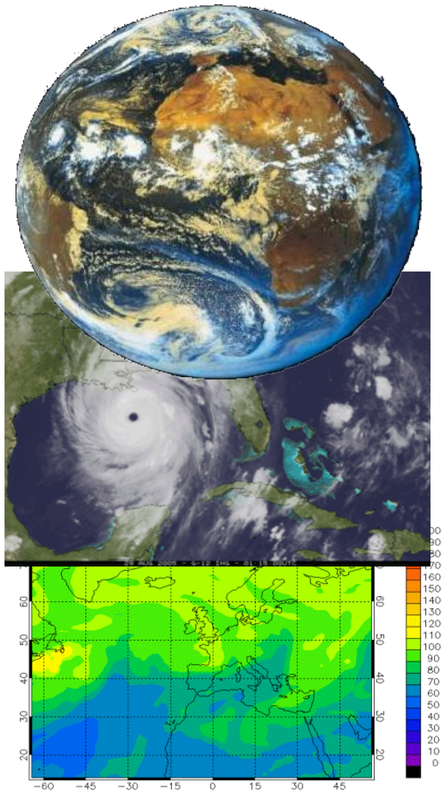
- Weather, Climatology, Earth Science
  - degree of warming, scenarios for our future climate.
  - understand and predict ocean properties and variations
  - weather and flood events
- ***Astrophysics, Elementary particle physics, Plasma physics***
  - ***systems, structures which span a large range of different length and time scales***
  - ***quantum field theories like QCD, ITER***

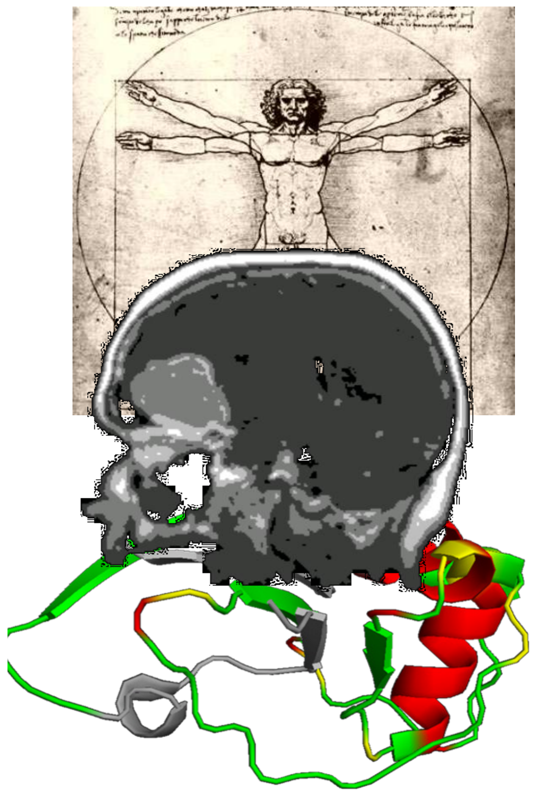
- ***Material Science, Chemistry, Nanoscience***
  - ***understanding complex materials, complex chemistry, nanosci***
  - ***the determination of electronic and transport properties***
- ***Life Science***
  - ***system biology, chromatin dynamics, large scale protein dynamics, pr association and aggregation, supramolecular systems, medicine***
- ***Engineering***
  - ***complex helicopter simulation, biomedical flows, gas turbines and internal combustion engines, forest fires, green aircraft,***
  - ***virtual power plant***

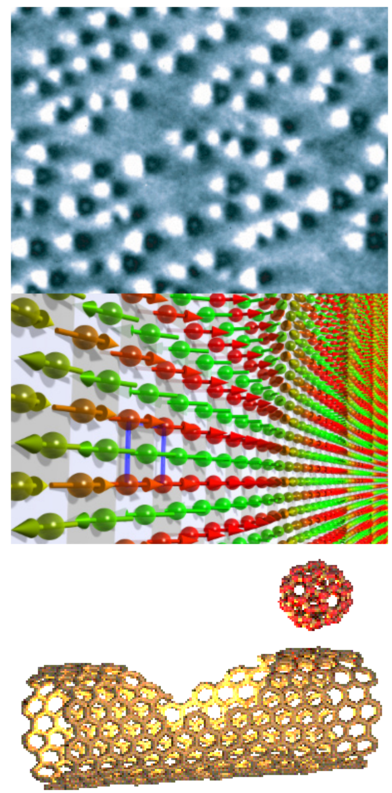
Supercomputing drives science with simulations



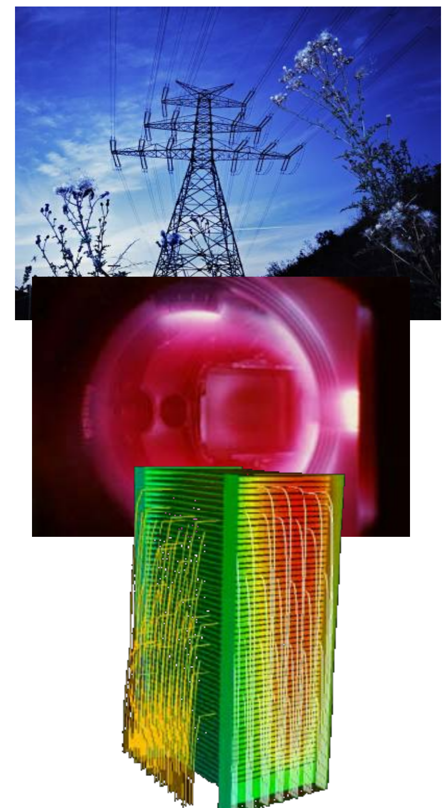
Environment
Weather/ Climatology
Pollution / Ozone Hole



Ageing Society
Medicine
Biology

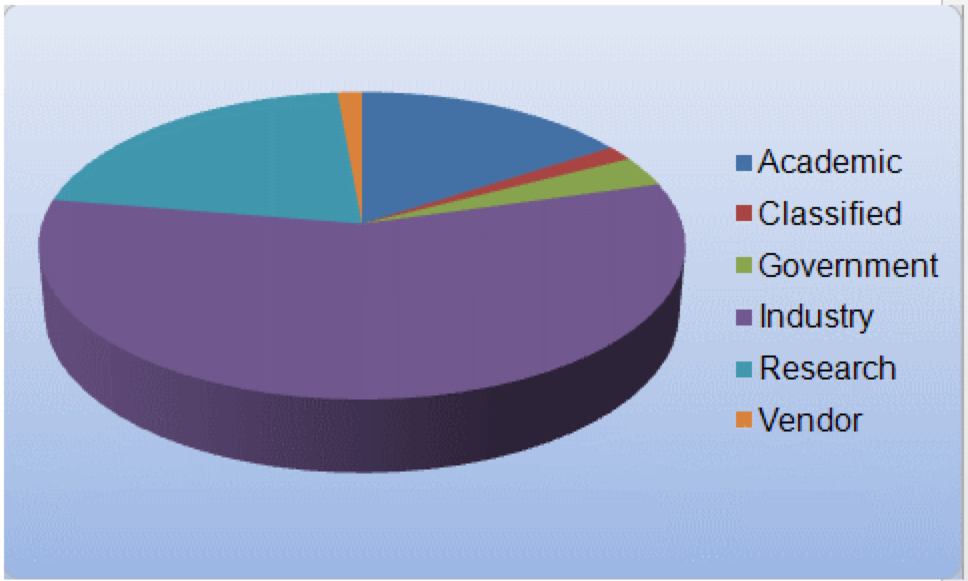


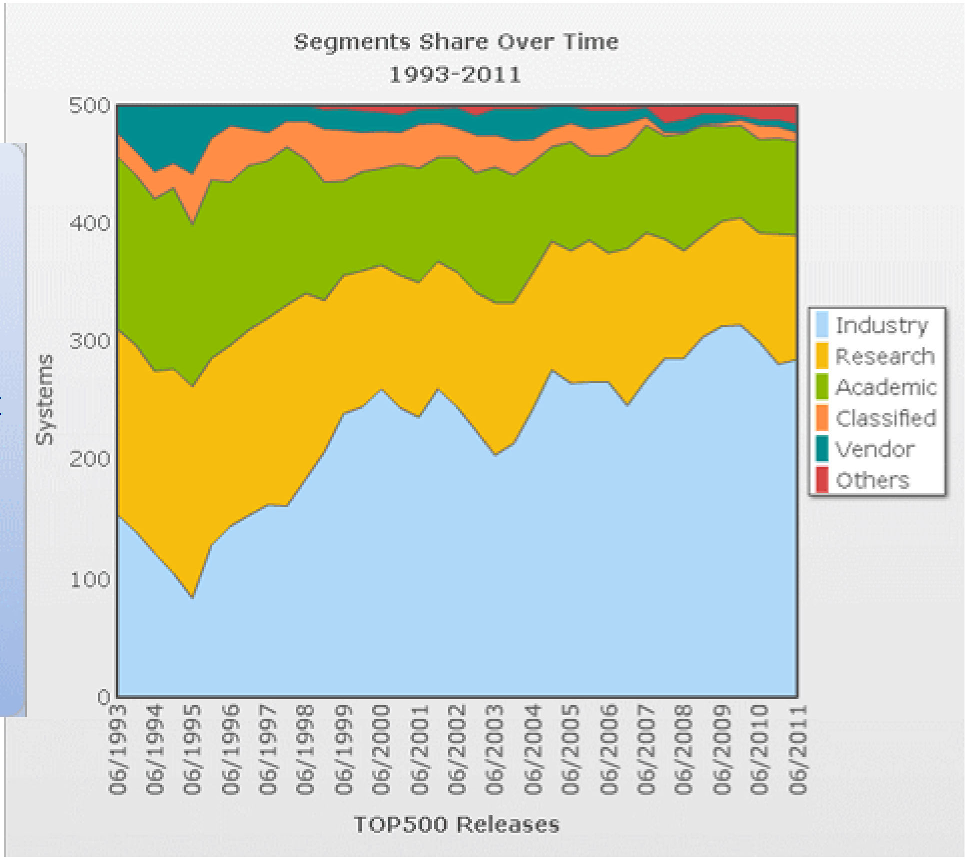
Materials/ Inf. Tech
Spintronics
Nano-science



Energy
Plasma Physics
Fuel Cells

Computing tshares in the TOP500 list





Large HPC systems around the world

FZJ

2010 1st PRACE System - JUGENE

- BG/P by Gauss Center for Supercom at Juelich
    - 294,912 CPU cores, 144 TB memory
    - 1 PFlop/s peak performance
    - 825.5 TFlop/s Linpack
    - 600 I/O nodes (10GigE) > 60 GB/s I/O
    - 2.2 MW power consumption
    - 35% for PRACE

GENCI

2011 2nd PRACE system – CURIE

- ***Bull, 1.6PF, 92160 cores, 4GB/core***
- ***Phase 1, December 2010, 105 TF***
  - ***360 four Intel Nehalem-EX 8-core nodes, 2.26 GHz CPUs (11,520 cores), QDR Infiniband fat-tree***
  - ***800 TB, >30GB/sec, local Lustre file system***
- ***Phase 1.5 Q2 2011***
  - ***Conversion to 90 16-socket, 128 core, 512 GB nodes***
- ***Phase 2, Q4 2011, 1.5 TF***
  - ***Intel Sandy-Bridge***
  - ***10PB, 230GB/sec file system***

HLRS

2011 3rd PRACE System – HERMIT

- ***Cray XE6 (Multi-year contract for $60+M)***
  - ***Phase 0 – 2010***
    10TF, 84 dual socket 8-core AMD Magny-Cours CPUs, 1344 cores in total, 2 GHz, 2GB/core, Gemini interconnect
  - ***Phase 1 Step 1 – Q3 2011***
    AMD Interlagos, 16 cores,1 PF
    2 – 4 GB/core
    2.7 PB file system, 150 GB/s I/O
  - ***Phase 2 – 2013***
    Cascade, first order for Cray, 4- 5 PF

LRZ

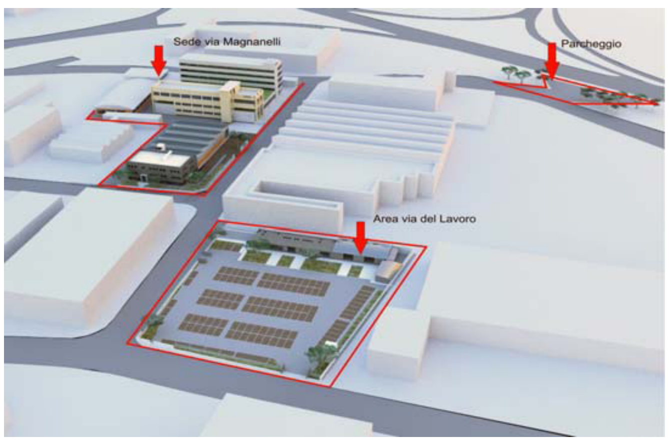
2011/12 4th PRACE system

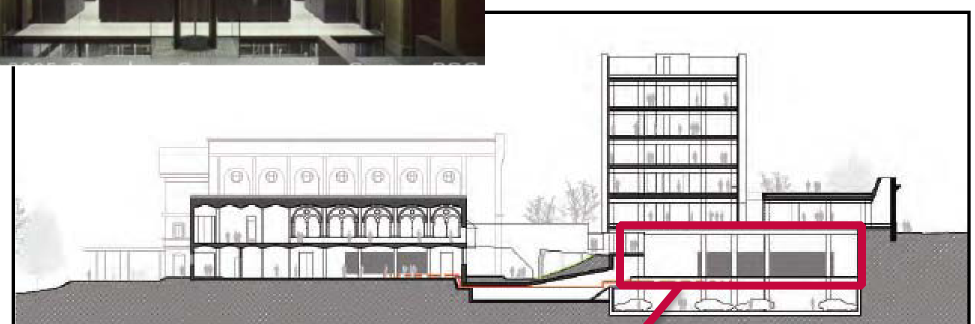
- ***IBM iDataPlex (€83M including operational costs)***
  - ***>14,000 Intel Sandy-Bridge CPUs, 3 PF (~110,000 cores), 384 TB of memory***
  - ***10PB GPFS file system with 200GB/sec I/O,
2PB 10GB/sec NAS
LRZ <13MW***
  - ***Innovative hot water cooling (60C inlet, 65C outlet) leading to 40 percent less energy consumption compared to air-cooled machine.***

BSC and CINECA

- 2012/2013 5th and 6th PRACE Systems

CINECA
2.5 PF





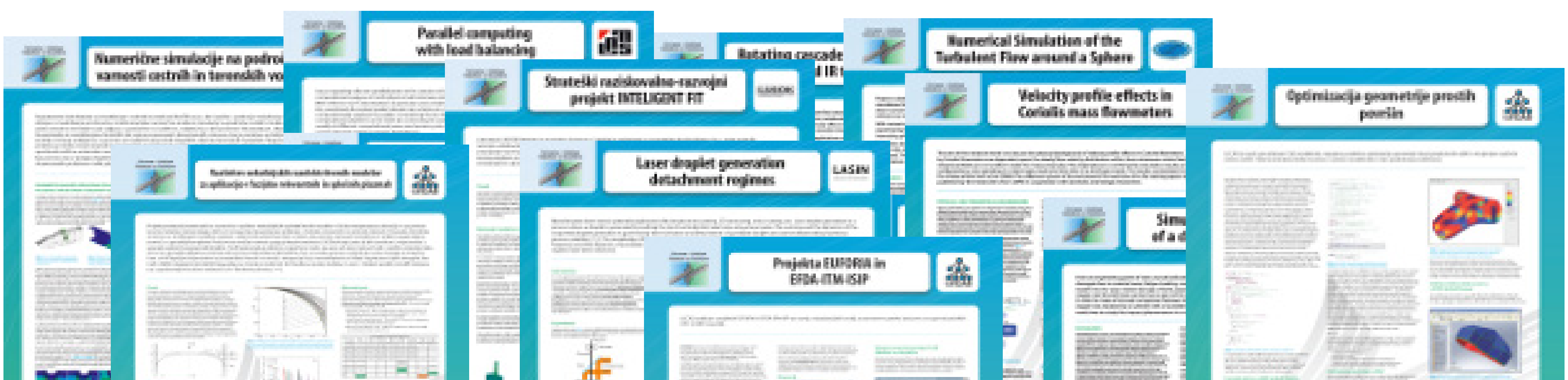
Computing Facility
10 MW 2013

Supercomputing at UL FME --HPCFS for ?

- ***Some examples of previous projects***

What HPCFS is used for?

- ***Complex enginering research problems demands parallel processing***
- ***Education of new generation of students on II cycle ob Bologna process***
- ***Cooperation with other GRID and HPC centres***

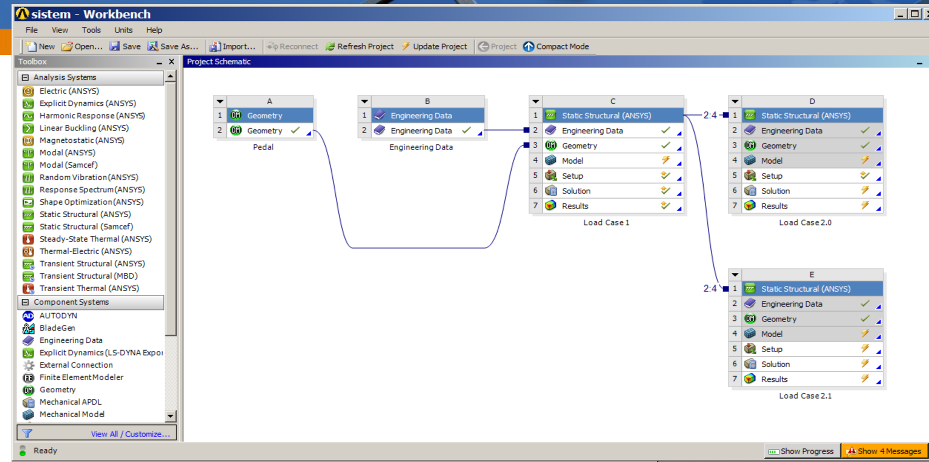


Long term goals

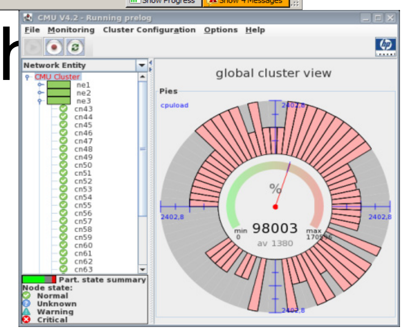
- Extension of computing capabilities

- In-house development of custom codes
- ***Installation of commercial and open-source code***
- ***ANSYS Multiphysics, OpenFOAM,..***
- ***Cooperation in EU projects***
- ***Advantage is if having HPC and knowledge about it***
- ***Introducing (young) researchers***
  - ***Center for modelling, simulations and optimization in cooperation on severale levels at university and intra universities***
- ***Promotion of FS/UL, science, research and increased awareness***

Software at HPCFS



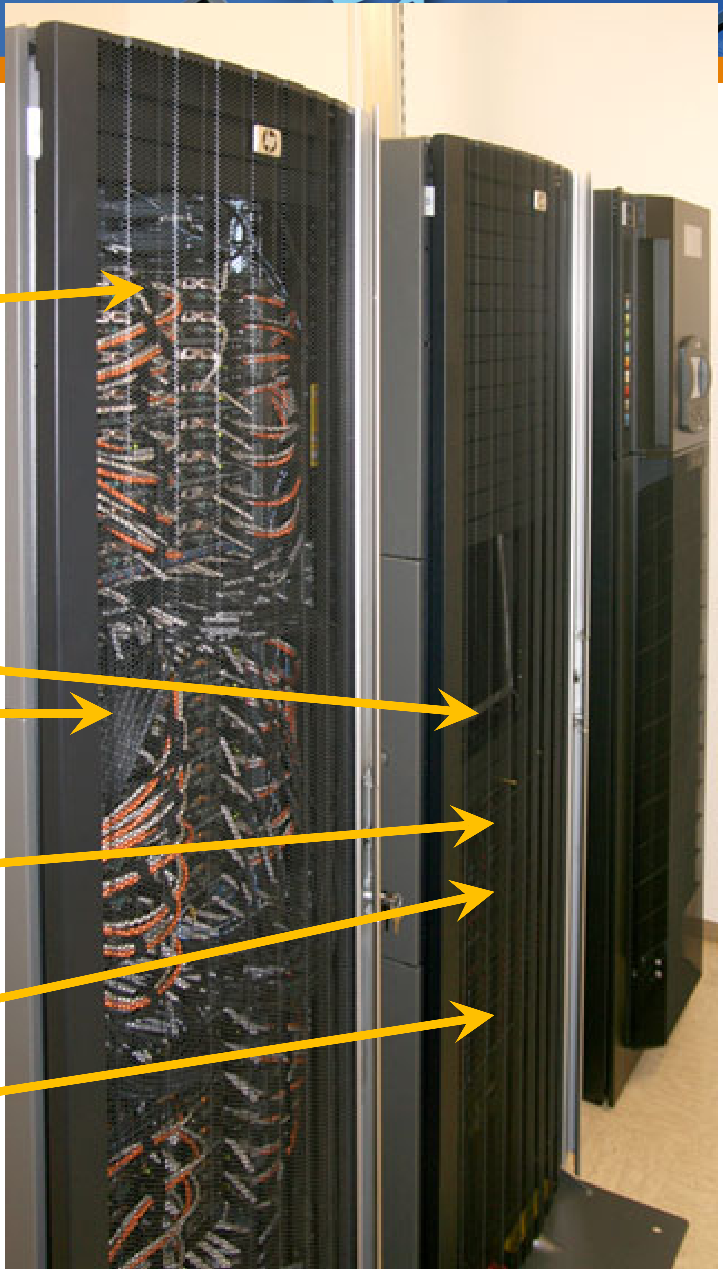
- Linux (CentOS 6.4)
- Remote desktop NX
- Development environment and LSF batch sch



- Compilers C++, Fortran (Python, R, ...)
- Parallel programming with MPI, OpenMP
- Open-source and commercial packages for simulations (ANSYS)
- Servers for support of the researsch and development

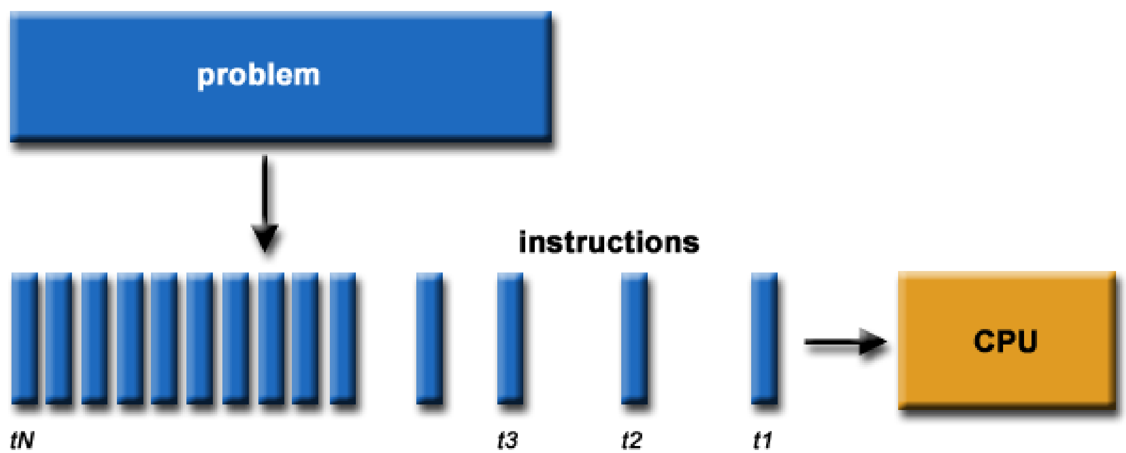
Hardware of the cluster PRELOG at ULFME

- 64 computing nodes
  - *768 cores X5670*
  - *1536 threads*
- ***3 TB RAM***
- ***Login node***
- ***Infiniband network***
- ***QDR x4 „fat tree“***
- ***File servers***
  - ***NFS 25TB***
  - ***LUSTRE 12TB+22TB***
- ***Virtualization servers***
- ***1Gbit Connection to ARNES***



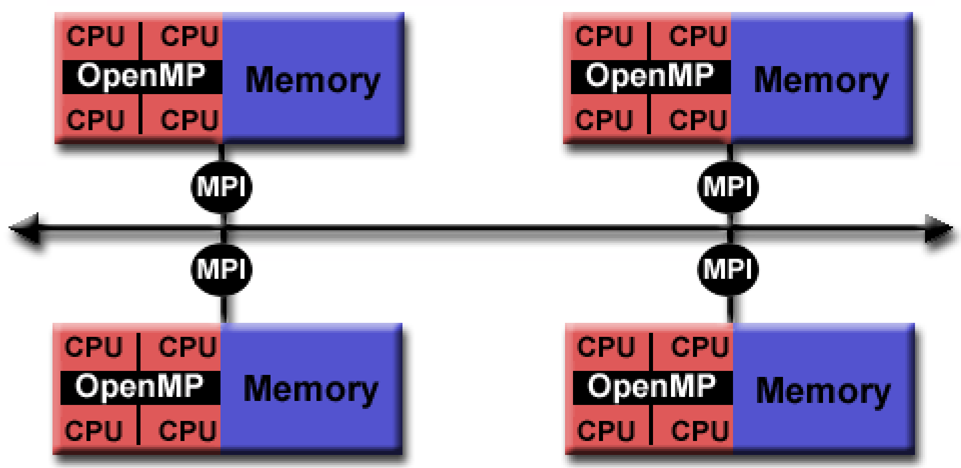
Introduction to parallel computing

- Usually is the program written for serial execution on one processor
- We divide the problem into series of commands that can be executed in paralllel
- Only one command at a time can be executed on one CPU



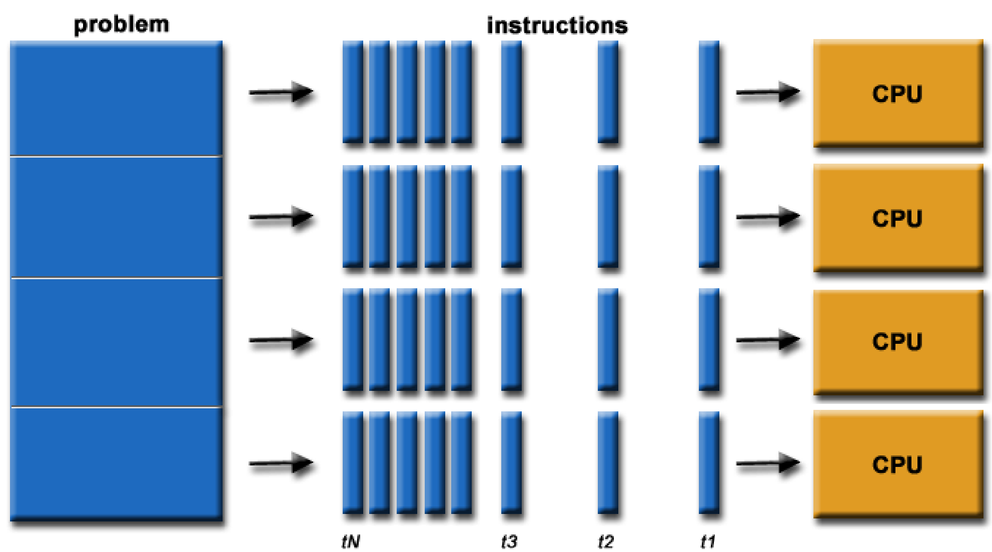
Parallel programming models

- Threading
- ***OpenMP* – *automatic parallelization***
- Distributed memory model = ***Message Passing Interface (MPI)* – *manual parallelization needed***
- ***Hybrid model OpenMP/MPI***



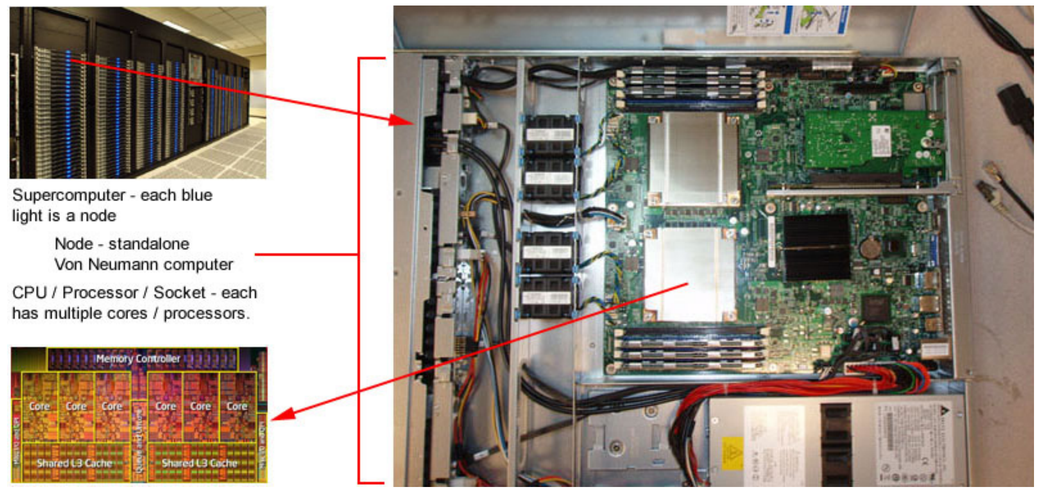
Embarrasingly simple parallel processing

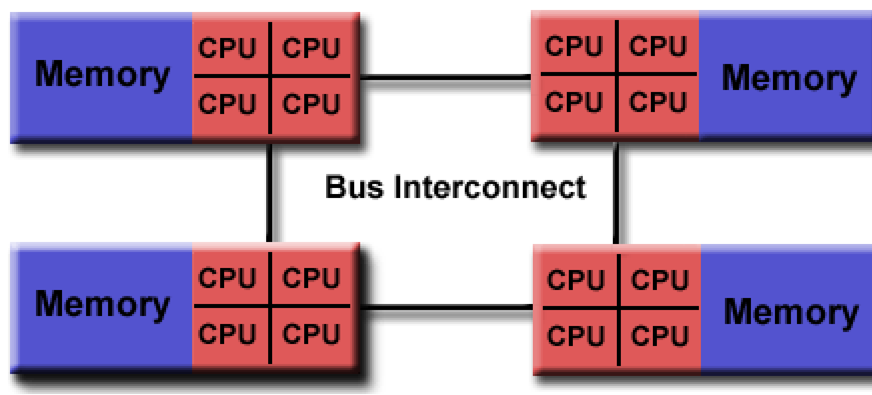
- Parallel processing of the same subproblems on multiple prooocessors
- No communication is needed between processes



Logical view of a computing node

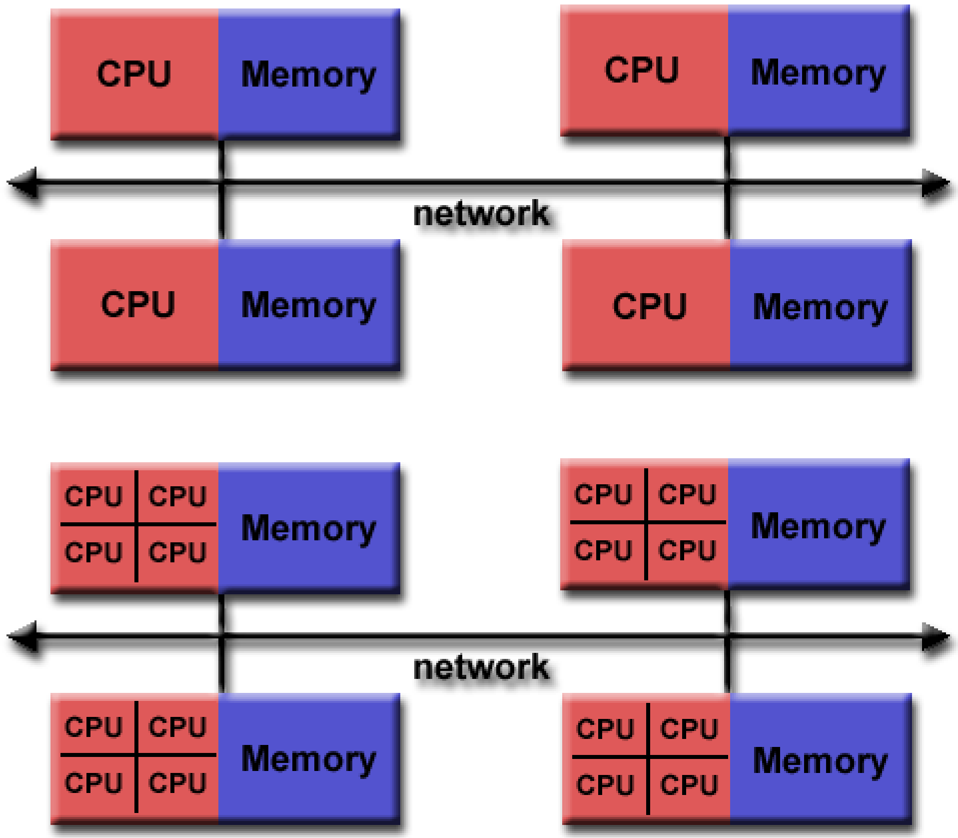
- Need to know computer architecture
- ***Interconnect bus for sharing memory between processors (NUMA interconnect)***





Nodes interconnect

- Distributed computing
- Many nodes exchange messages on
  - high speed,
  - low latency interconnect such as **Infiniband**



Development of parallel codes

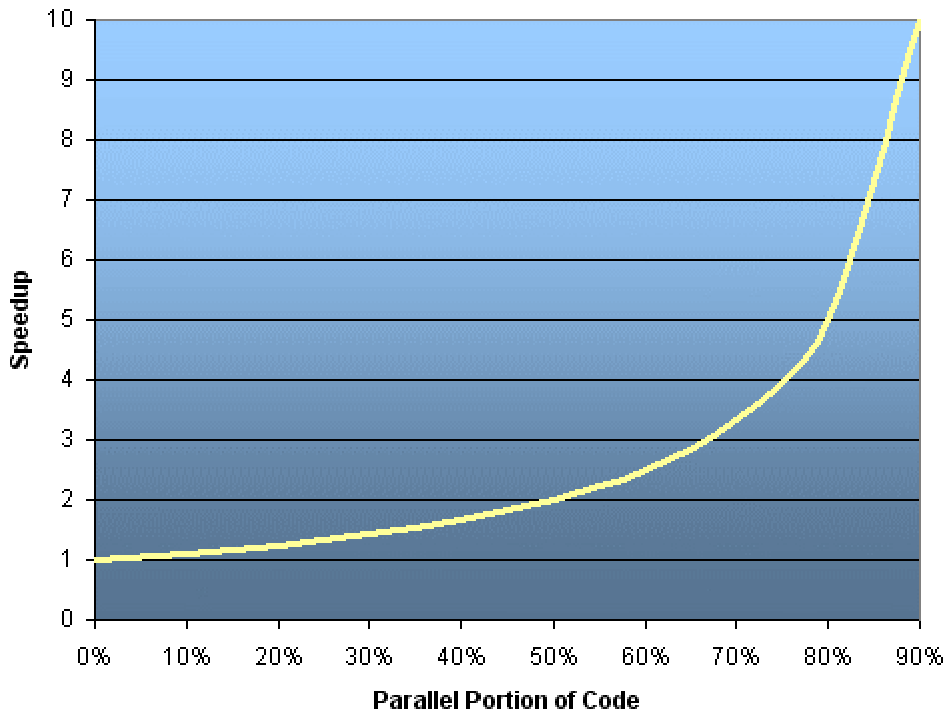
- Good understanding of the problem being solved in parallel
- How much of the problem can be run in parallel
- Bottleneck analysys and profiling gives good picture on scalability of the problem
- We optimize and parallelize parts that consume most of the computing time
- Problem needs to be disected into parts functionally and logically

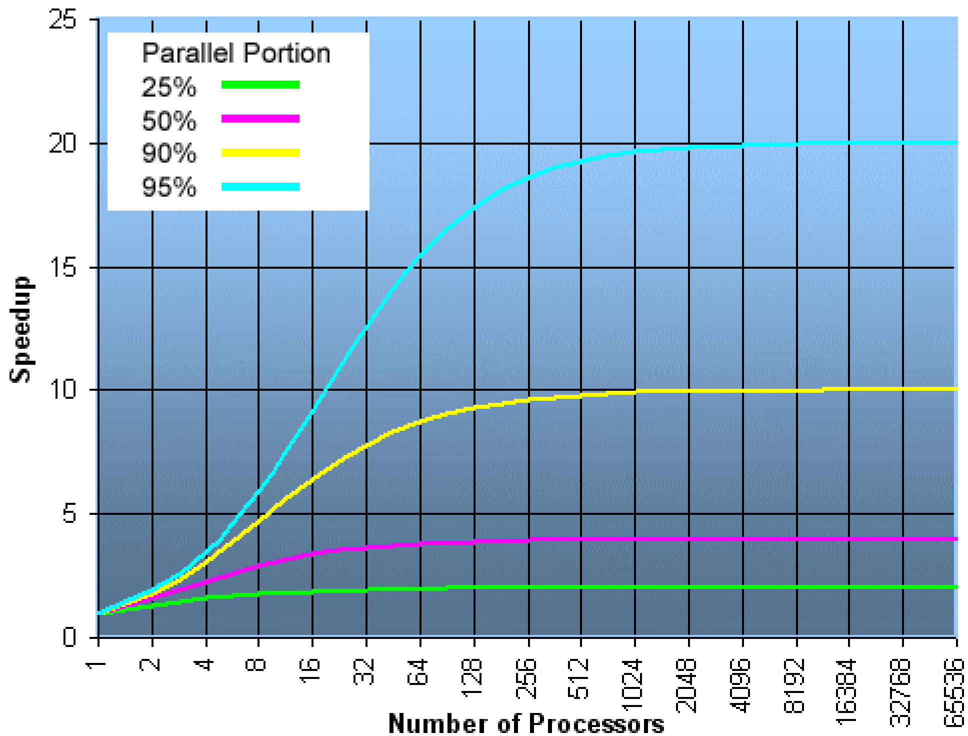
Interprocess communications

- Having little an infrequent communication between processes is the best
- Determining the largest block of code that can run in parallel and still provides scalability
- Basic properties
  - ***response time***
  - ***transfer speed - bandwidth***
  - ***interconnect capabilities***

Parallel portion of the code determines code scalability

- Amdahlov law ***Speedup = 1/(1-p)***





Questions and practicals on the HPCFS cluster

- Demonstration of the work on the cluster by repeating
- Access with NX client
- Learning basic Linux commands
- LSF scheduler commands
- Modules
- Development with OpenMP and OpenMPI parallel paradigms
- Excercises and extensions of basic ideas
- Instructions available at http://hpc.fs.uni-lj.si/