RUPRECHT-KARLS-
UNIVERSITÄT
HEIDELBERG

Introduction to High Performance Computing

Lecture 01 - Introduction

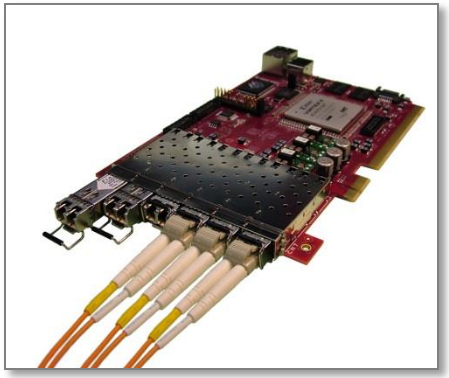
Holger Fröning
Institut für Technische Informatik
Universität Heidelberg

Persons

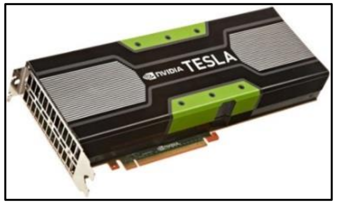
- Lecturer: JProf. Dr. Holger Fröning
  - Juniorprofessur, Institut für Technische Informatik, Universität Heidelberg
  - Email: froening@uni-hd.de, Web: http://ce.uni-hd.de
- Discussions/Questions
  - During/after the lecture
  - Or via email
- Consulting hours after appointment
- Exercises: Benjamin Baumann

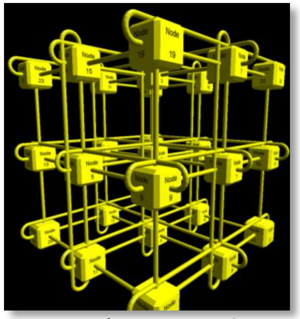
Computer Engineering Group

- "Currently sold on BSP styles of computing for data intensive problems"
- Research
  - Parallel computing, computer architecture, interconnection networks and hardware design with a recent focus on application-specific heterogeneous computing, data movement optimizations and associated power and energy aspects
  - "Most advanced group investigating data movements for accelerated systems"
- Teaching
  - Major „Application-Specific Computing (ASC)"
- Nvidia CUDA Research and Teaching Center
- Interactions
  - Technical University of Valencia, Spain
  - University of Castilla-La Mancha, Albacete, Spain
  - Georgia Institute of Technology, U.S.
  - Computer Architecture Group, UOH, Germany
  - CERN, Switzerland
  - AMD, NVidia, Xilinx, Intel, SAP, EXTOLL, ...

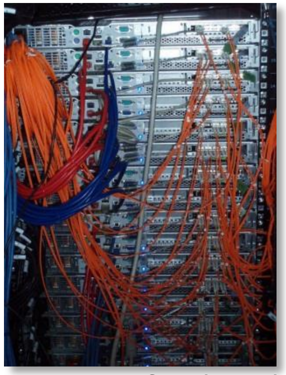


FPGA-based specialized interconnection network





Visualization of a network topology



Prototype of a shared-memory cluster

Get involved!

- In the case of interest plenty of opportunities exist
  - Hardware/Software design up to benchmark level
  - GPUs, CPUs, memory architecture, parallel programming, software environments, performance evaluation, etc

- MSc- and BSc-theses
- Project work

- Seminar work
  - Advanced Seminar „Computer Engineering“
  - 24.10., 14:00 ct, INF501, R102

RUPRECHT-KARLS-
UNIVERSITÄT
HEIDELBERG

Organisatorisches

Studiengänge

- MSc „Technische Informatik“
- Studierende mit Haupt- oder Nebenfach Informatik
  - MSc, BSc, LA
- Am Thema Interessierte
  - Physik, DKFZ, Mathematik, …

- Studien- bzw. Prüfungsleistung:
  - Voraussetzung Leistungsnachweis
    - Punktzahl Übungen >= 50%
  - Leistungsnachweis: **10.02.2015**
    - Mündliche Prüfung mit Schein

Aufbau

- VL „Introduction to HPC“
  - Vorlesung/Übung als 2+2
  - 13 Termine im WS 2014/2015
  - Vorlesung Di 14:00-16:00 (ct), Seminarraum A3.04
  - 1 Übung Di 16:00-18:00, Terminalraum B2.15
    - Arbeit in 2er Gruppen
- Anmeldung per Email
  - Zu Vorlesungsbeginn
  - Mit Name, Matrikelnummer, Studiengang
- Unterlagen (vorlesungsbegleitende Folien, Übungsblätter, Papers)
  - Auf Moodle, Schlüssel ist „mpi_send“

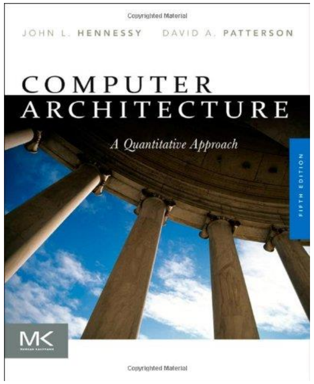
Outline

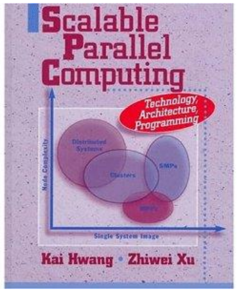
	Datum	Vorlesung	Übung
1	21.10.2014	Introduction	Reading
2	28.10.2014	Introduction to Message Passing	MPI Basics
3	04.11.2014	Basics	MPI latency & bandwidth, MMULT considerations
4	11.11.2014	Parallel Computing	MMULT MPI
5	***18.11.2014***	Practical Parallel Programming Example	Relaxation single-threaded
6	25.11.2014	Messaging	Relaxation MPI 1 (1D partitioning)
7	02.12.2014	Characteristics	N-Body Preparation
8	09.12.2014	Benchmarks	N-Body Implementation
9	16.12.2014	GPU Computing	CUDA Basics
10	13.01.2015	GPU Computing II	CUDA MMULT + MPI
11	20.01.2015	Interconnection Networks 1	CUDA MMULT with communication
12	27.01.2015	Interconnection Networks 2/3	-
13	03.02.2015	Systems / Future	Final discussion

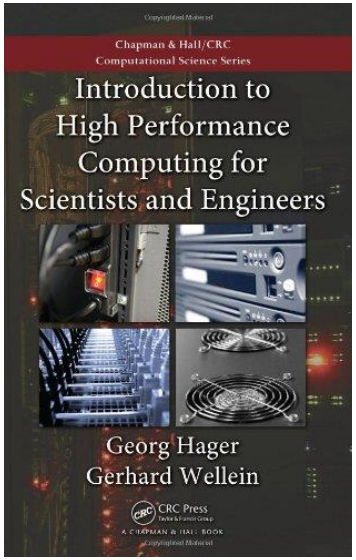
Literatur

- Vorlesungsbegleitende Folien werden laufend zur Verfügung gestellt.

- Literatur
  - Georg **Hager**, Gerhard **Wellein**: *Introduction to High Performance Computing for Scientists and Engineers*, Chapman & Hall, 2010
  - John L. **Hennessy** & David A. **Patterson**: *Computer Architecture – A quantitative approach*, 5th Ed., Morgan Kaufmann, 2011
  - David A. **Bader** (Ed.): *Petascale Computing: Algorithms and Applications,* Chapman & Hall, 2007
  - Kai **Hwang** und Zhiwei Xu: *Scalable Parallel Computing: Technology, Architecture, Programming,* Mcgraw Hill Book Co, 1997







Feedback

- Rückmeldungen und Fragen
  - Gebt mir Rückmeldungen bzgl. Änderungen oder Ergänzungen zum Stoff
  - Stellt Fragen, auch per Email!
  - Macht mich auf Fehler aufmerksam!
  - Nutzt (außerhalb der Vorlesung) die Möglichkeit, mit Mit-Studierenden zu diskutieren
- Selbstlernmöglichkeiten
  - Literatur
  - Bekanntgegebene Weblinks/Papers
  - Infrastruktur am ZITI

Vorkenntnisse

- Voraussetzungen:
  - C Kenntnisse, Linux
- Empfohlen:
  - Grundlagen von Rechnerarchitektur, Betriebssysteme
- Empfohlene Vorlesungen
  - VL “Parallele Rechnerarchitekturen von Prof. Brüning
    - Parallel in diesem Semester
  - VL “GPU-Computing”
    - Parallel in diesem Semester
  - VL „Advanced Parallel Computing“
    - Folgendes Semester
    - Vertiefung „Application-specific Computing“

Exercises: Scientific Review

- Reading & Feedback based on paper review
  - Ideal review here is 3 sentences for each of the following:
    1. Primary contribution
    2. Key insight of the contribution
    3. Your opinion/reaction to the content
  - Review: rating relative to all other papers (of this venue)
    - strong reject, weak reject, weak accept, accept
  - Old papers: optionally include some comments on how right this paper was
- Reading:
  - Michael J. Flynn and Patrick Hung. 2005. Microprocessor Design Issues: Thoughts on the Road Ahead. *IEEE Micro* 25, 3 (May 2005), 16-31.
  - Walker, 2008, benchmarking Amazon EC2 for high-performance scientific computing, *;login: The USENIX Magazine, 33(5).*
- Provide review (summary) using Moodle until next Tuesday
- Discussion round as part of the exercise

RUPRECHT-KARLS-
UNIVERSITÄT
HEIDELBERG

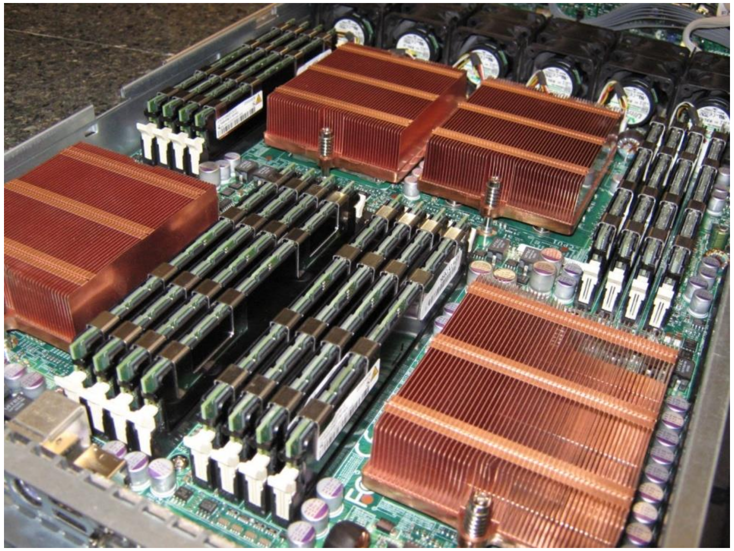
Introduction to High Performance Computing

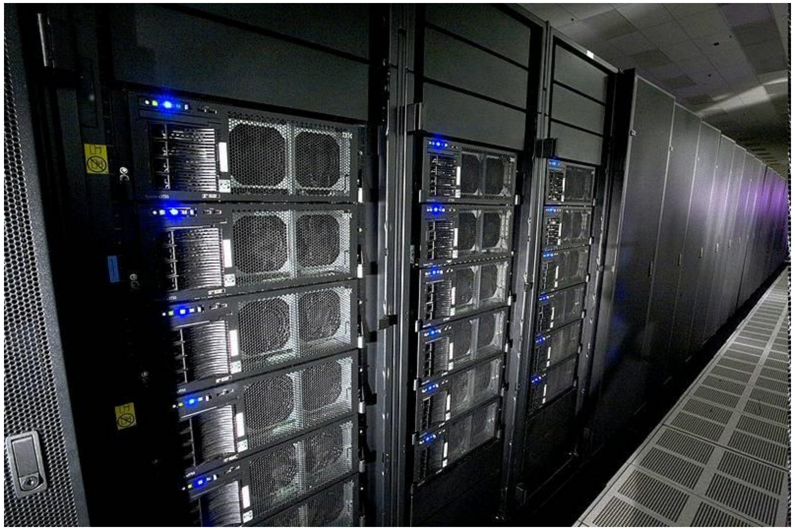
Questionnaire

Buzzword	
Multi-Core	
Infiniband	
TOP500	
Virtual Memory	
ISA	
Pthread/OpenMP	
Demand Paging	
Double Precision	
Virtual Machine	
Tianhe-2	
MapReduce	

Vorlesungsinhalte

- Einführung in HPC
  - Gebiet: „Rechnerarchitektur“
  - Großgebiet: Technische Informatik
    - „Innenleben von Rechnern“

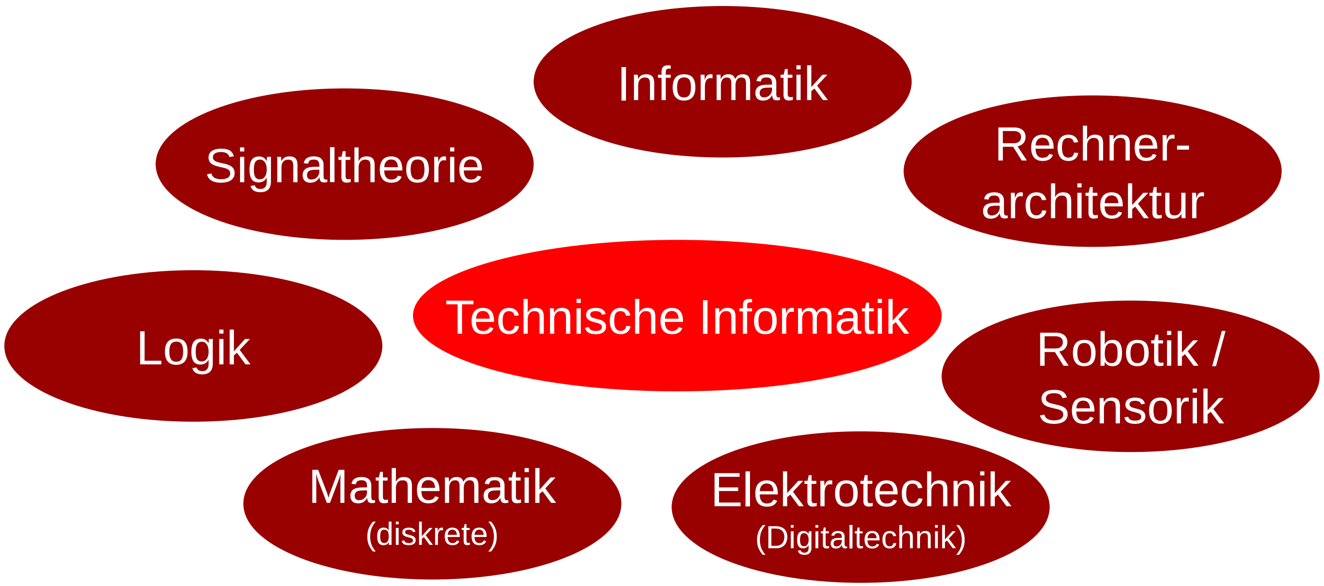




Courtesy: wikipedia.org

Vorlesungsinhalte

- Großgebiet: Technische Informatik

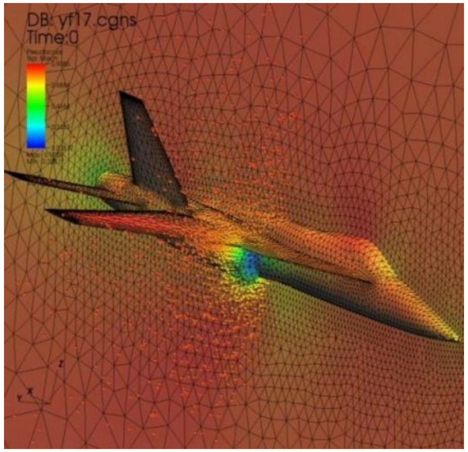


High Performance Computing - Definition

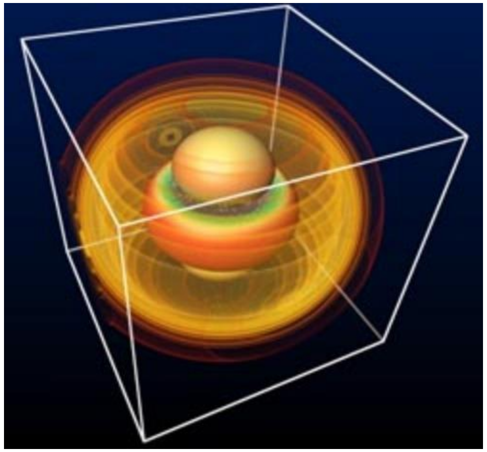
- Versuche:
  1. *“By HPC we mean all clusters of servers and associated infrastructure used to solve technical/scientific or analytic problems that are computationally intensive or data intensive, and use such techniques as simulation, modeling, etc. This includes technical servers but EXCLUDES desktop computers or workstations used for technical computing.”*
  2. *“Verwendung von spezialisierten (evtl. hochparallelen) Systemen zur beschleunigten Berechnung von rechen- bzw. datenintensiven Problemen, d.h. zur Reduzierung der Laufzeit im Vergleich zu Standardrechensystemen durch Erhöhung der Rechenleistung.”*
- Verwandte Themen
  - Supercomputing - Exascale

Vorlesungsinhalte

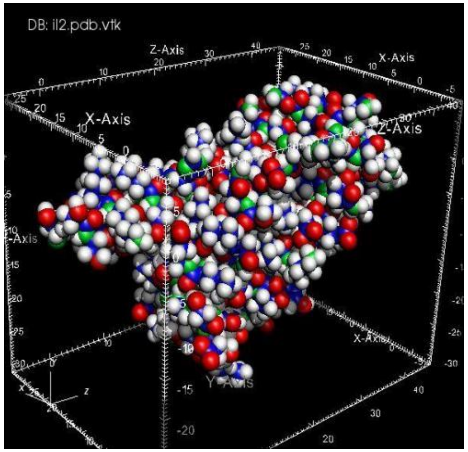
1. Basics
2. Parallel computing
3. MPI & messaging
4. Benchmarks & application characteristics
5. Data-parallel processors
6. Interconnection networks
7. Future



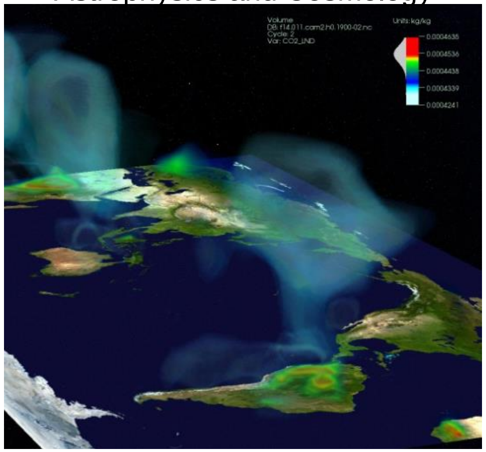
CFD



Astrophysics and Cosmology



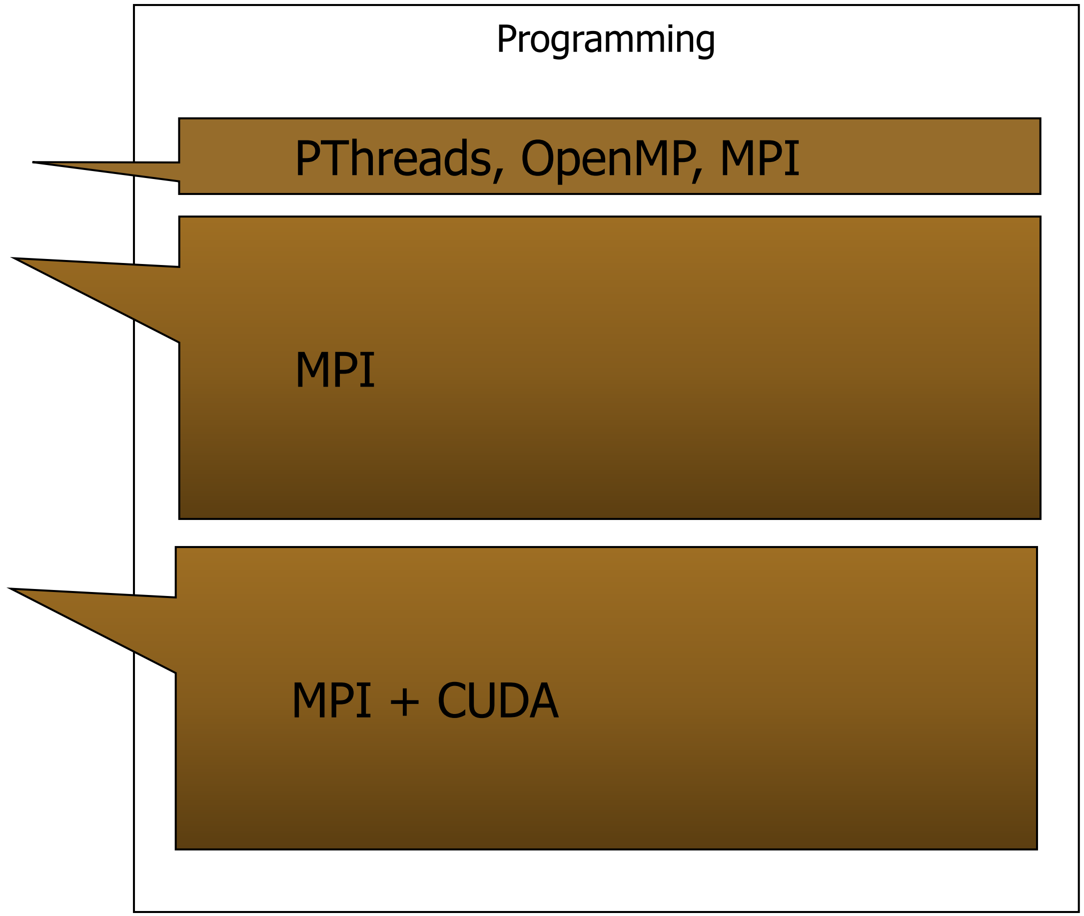
Molecular Dynamics



Weather and Climate Research

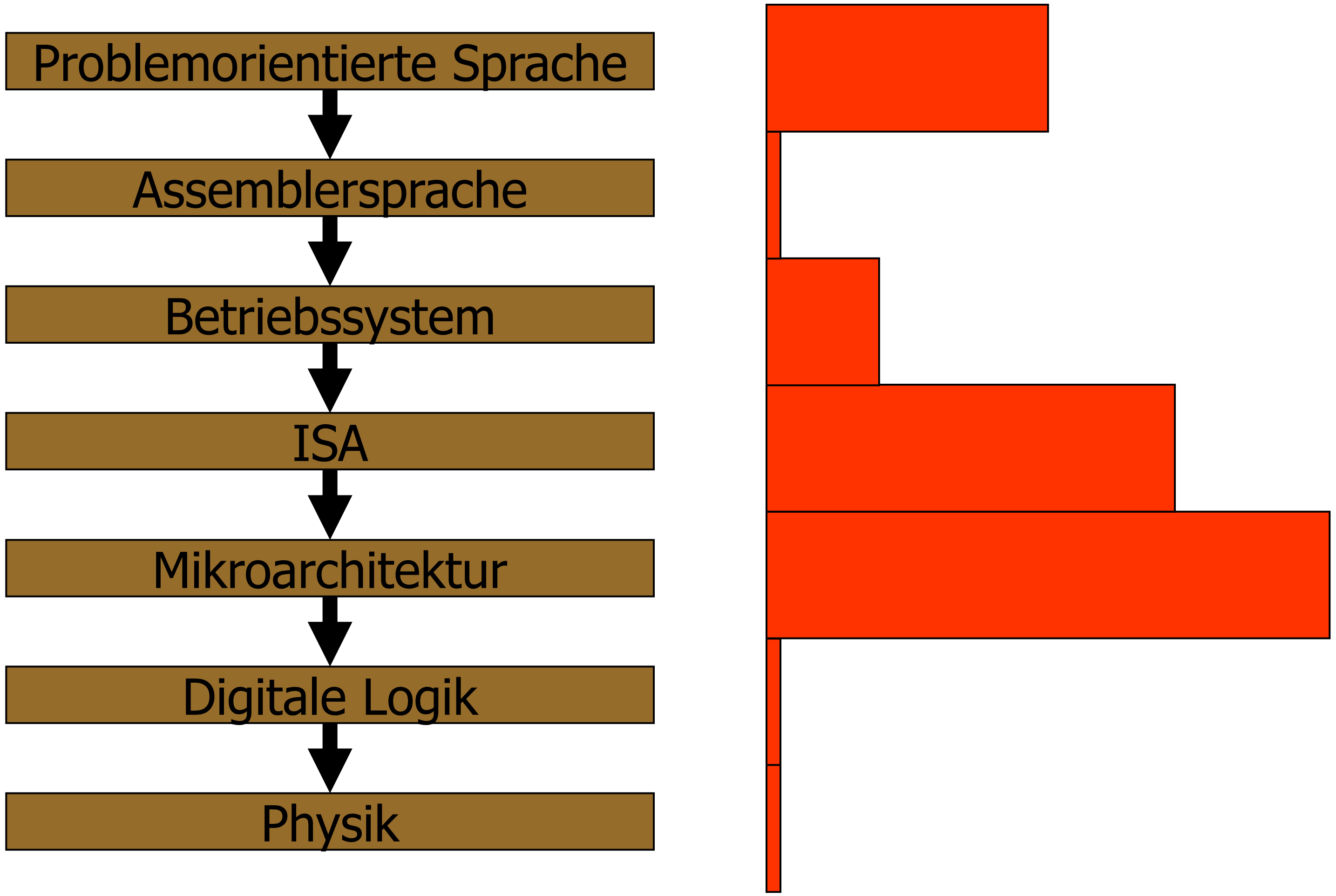
Vorlesungsinhalte

1. Basics
2. Parallel computing
3. MPI & messaging
4. Benchmarks & application characteristics
5. Data-parallel processors
6. Interconnection networks
7. Future

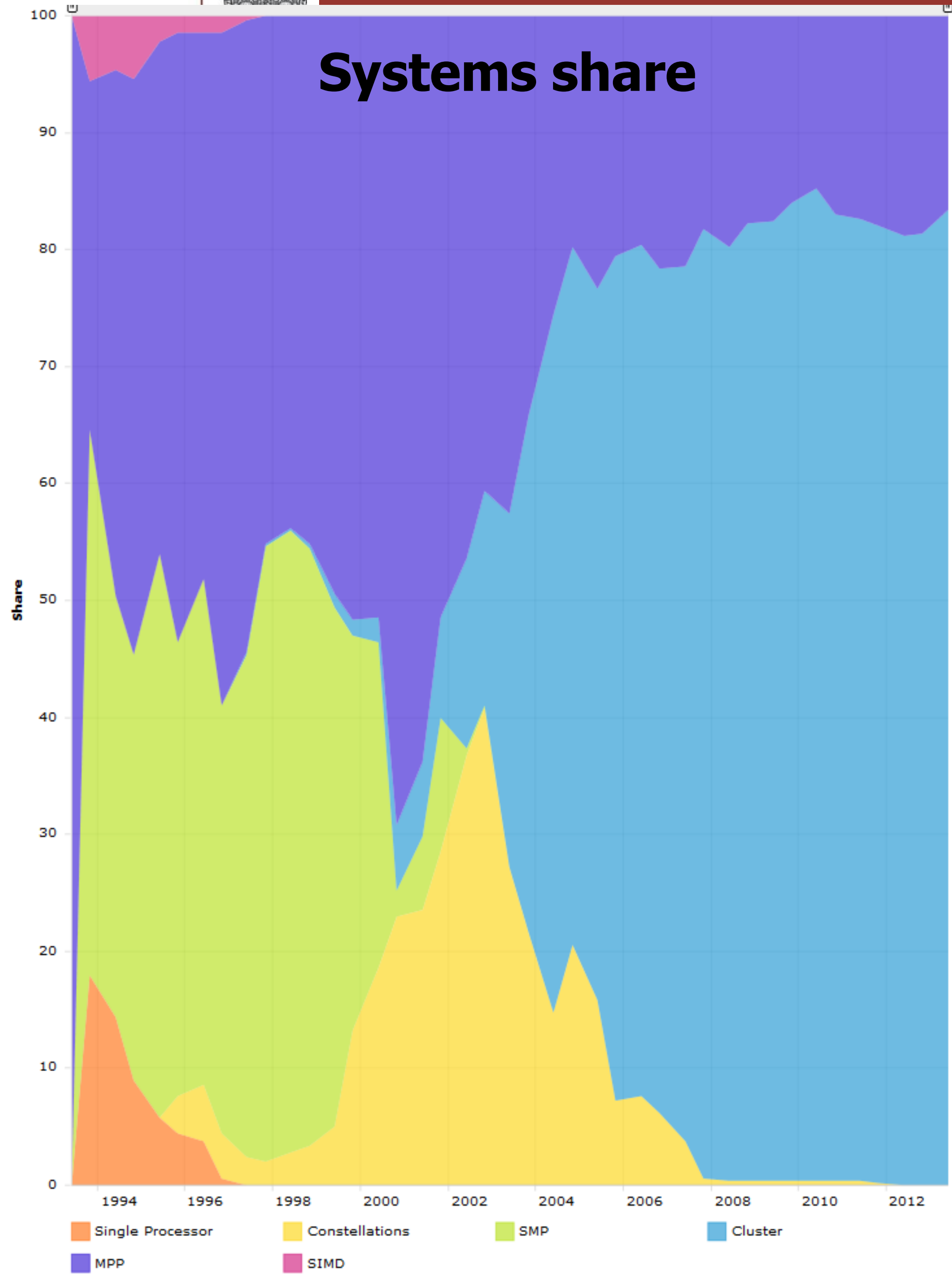


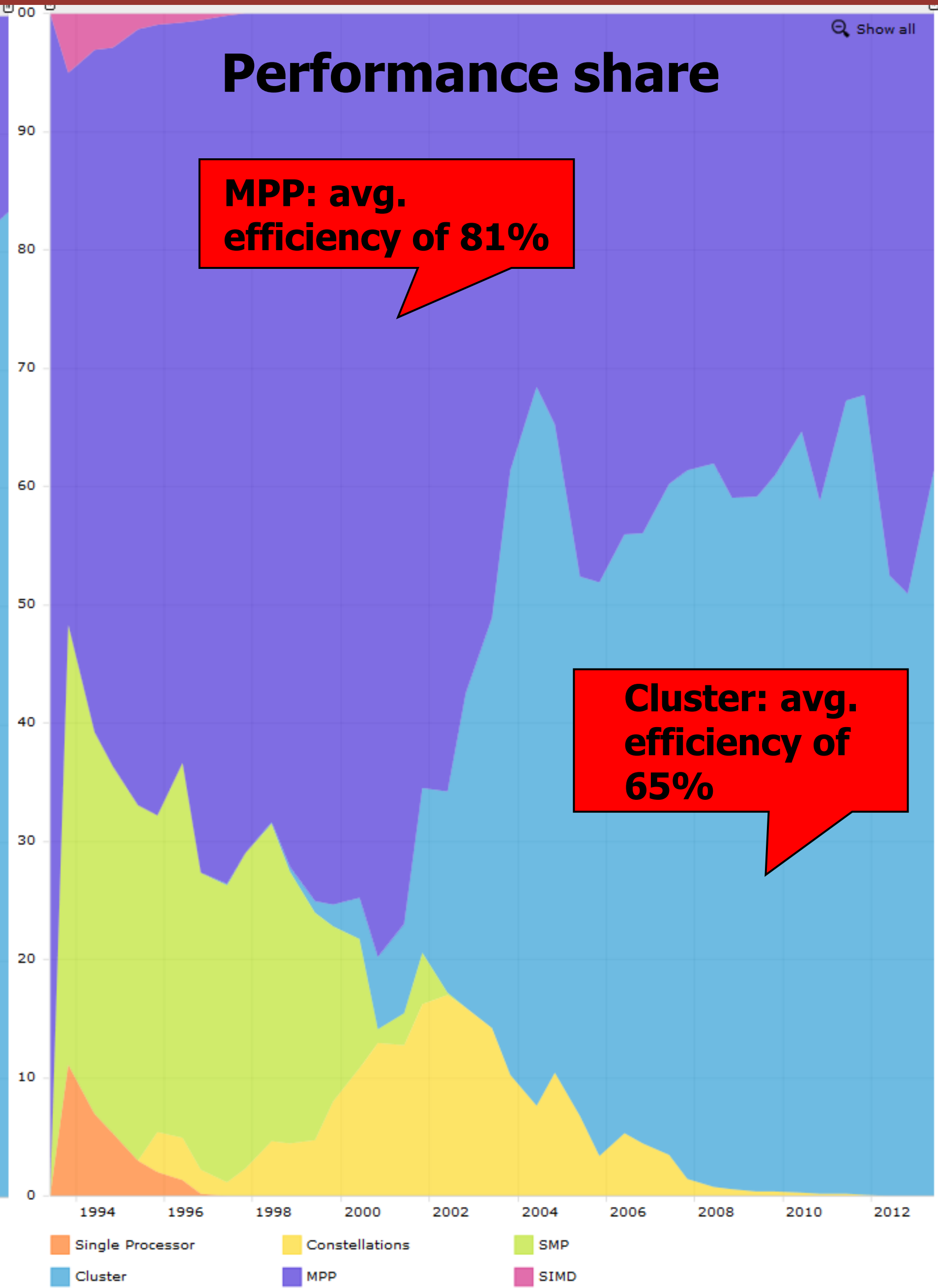
Vorlesungsinhalte

Schichten eines Rechensystems



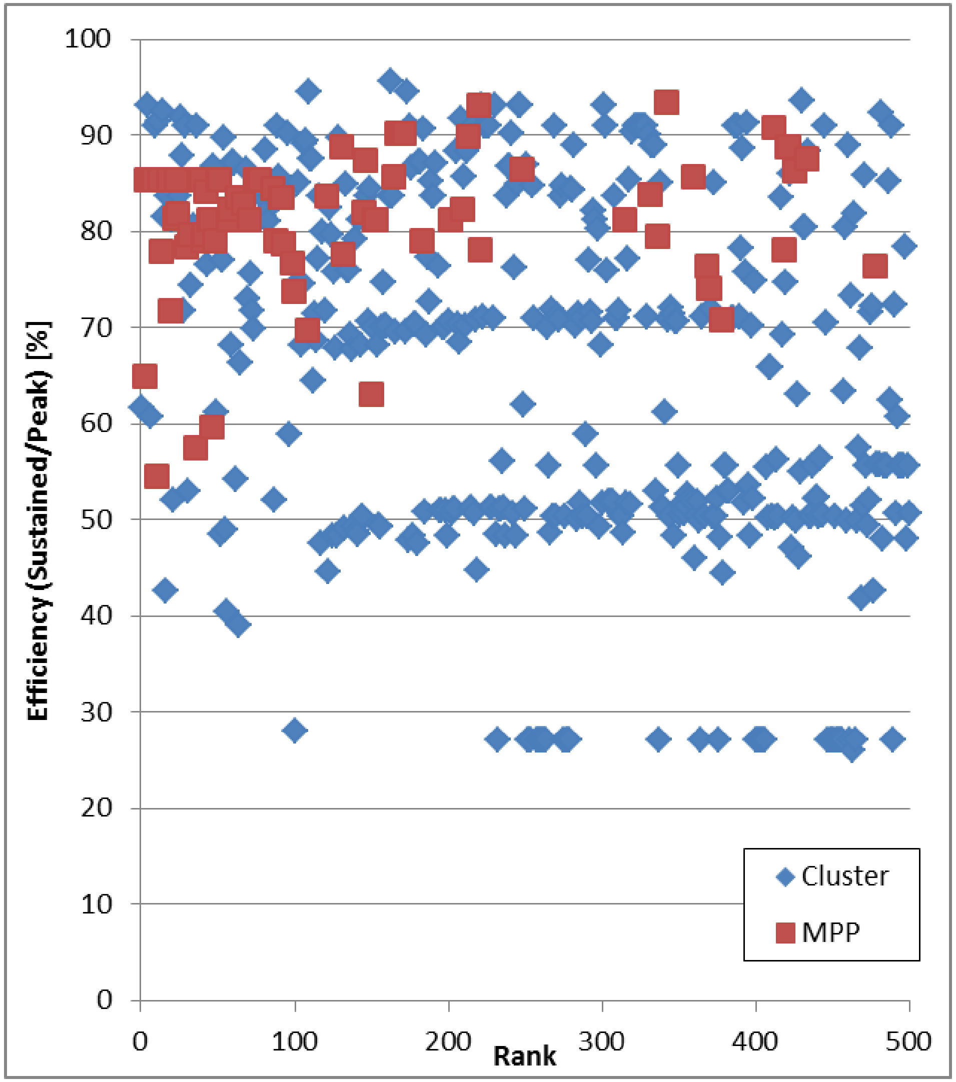
TOP500


Systems share
Share
100
90
80
70
60
50
40
30
20
10
0
1994
1996
1998
2000
2002
2004
2006
2008
2010
2012
Single Processor
Constellations
SMP
Cluster
MPP
SIMD


Performance share
Show all
MPP: avg. efficiency of 81%
Cluster: avg. efficiency of 65%
90
80
70
60
50
40
30
20
10
0
1994
1996
1998
2000
2002
2004
2006
2008
2010
2012
Single Processor
Constellations
SMP
Cluster
MPP
SIMD

RUPRECHT-KARLS-
UNIVERSITÄT
HEIDELBERG

TOP500: Efficiency (Sustained/Peak)



- TOP500 (06/2013)
  - Efficiency: 26%/69%/96% (min/avg/max)
  - Rmax/core: 2.8/12/95 GFLOPs (min/avg/max)
- ~60 accelerated systems
  - GPUs/MICs
  - Efficiency: 27%/57%/77% (min/avg/max)
  - 7 accelerated systems back in 2008

Power Consumption

- Constraints of Energy/Power Consumption
  - Economical: Costs
  - Technological: supply, distribution, cooling
  - Ecological

→ „Heat Wall“

„You can put more cores on a die than you can afford to turn on“

Cray 3 from 1993: 90kW

- Supercomputer today: MWatts
- Google‘s Power Budget
- NSA Number Cruncher

Leistungsaufnahme - 2008

TOP500 List - November 2008 (1-100)

R_{max} and R_{peak} values are in TFlops. For more details about other fields, check the TOP500 description.

Power data in KW for entire system

next

Rank	Site	Computer/Year Vendor	Cores	R_{max}	R_{peak}	Power
1	DOE/NNSA/LANL United States	Roadrunner - BladeCenter QS22/LS21 Cluster, PowerXCell 8i 3.2 Ghz / Opteron DC 1.8 GHz , Voltaire Infiniband / 2008 IBM	129600	1105.00	1456.70	2483.47
2	Oak Ridge National Laboratory United States	Jaguar - Cray XT5 QC 2.3 GHz / 2008 Cray Inc.	150152	1059.00	1381.40	6950.60
3	NASA/Ames Research Center/NAS United States	Pleiades - SGI Altix ICE 8200EX, Xeon QC 3.0/2.66 GHz / 2008 SGI	51200	487.01	608.83	2090.00
4	DOE/NNSA/LLNL United States	BlueGene/L - eServer Blue Gene Solution / 2007 IBM	212992	478.20	596.38	2329.60
5	Argonne National Laboratory United States	Blue Gene/P Solution / 2007 IBM	163840	450.30	557.06	1260.00
6	Texas Advanced Computing Center/Univ. of Texas United States	Ranger - SunBlade x6420, Opteron QC 2.3 Ghz, Infiniband / 2008 Sun Microsystems	62976	433.20	579.38	2000.00
7	NERSC/LBNL United States	Franklin - Cray XT4 QuadCore 2.3 GHz / 2008 Cray Inc.	38642	266.30	355.51	1150.00
8	Oak Ridge National Laboratory United States	Jaguar - Cray XT4 QuadCore 2.1 GHz / 2008 Cray Inc.	30976	205.00	260.20	1580.71
9	NNSA/Sandia National Laboratories United States	Red Storm - Sandia/ Cray Red Storm, XT3/4, 2.4/2.2 GHz dual/quad core / 2008 Cray Inc.	38208	204.20	284.00	2506.00
10	Shanghai Supercomputer Center China	Dawning 5000A - Dawning 5000A, QC Opteron 1.9 Ghz, Infiniband, Windows HPC 2008 / 2008 Dawning	30720	180.60	233.47	

Source: top500.org

Leistungsaufnahme - 2013

Rank	Site	System	Cores	Rmax (TFlop/s)	Rpeak (TFlop/s)	Power (kW)
1	National University of Defense Technology China	**Tianhe-2 (MilkyWay-2)** - TH-IVB-FEP Cluster, Intel Xeon E5-2692 12C 2.200GHz, TH Express-2, Intel Xeon Phi 31S1P NUDT	3,120,000	33,862.7	54,902.4	17,808
2	DOE/SC/Oak Ridge National Laboratory United States	**Titan** - Cray XK7 , Opteron 6274 16C 2.200GHz, Cray Gemini interconnect, NVIDIA K20x Cray Inc.	560,640	17,590.0	27,112.5	8,209
3	DOE/NNSA/LLNL United States	**Sequoia** - BlueGene/Q, Power BQC 16C 1.60 GHz, Custom IBM	1,572,864	17,173.2	20,132.7	7,890
4	RIKEN Advanced Institute for Computational Science (AICS) Japan	K computer, SPARC64 VIIIfx 2.0GHz, Tofu interconnect Fujitsu	705,024	10,510.0	11,280.4	12,660
5	DOE/SC/Argonne National Laboratory United States	**Mira** - BlueGene/Q, Power BQC 16C 1.60GHz, Custom IBM	786,432	8,586.6	10,066.3	3,945
6	Texas Advanced Computing Center/Univ. of Texas United States	**Stampede** - PowerEdge C8220, Xeon E5-2680 8C 2.700GHz, Infiniband FDR, Intel Xeon Phi SE10P Dell	462,462	5,168.1	8,520.1	4,510
7	Forschungszentrum Juelich (FZJ) Germany	**JUQUEEN** - BlueGene/Q, Power BQC 16C 1.600GHz, Custom Interconnect IBM	458,752	5,008.9	5,872.0	2,301
8	DOE/NNSA/LLNL United States	**Vulcan** - BlueGene/Q, Power BQC 16C 1.600GHz, Custom Interconnect IBM	393,216	4,293.3	5,033.2	1,972
9	Leibniz Rechenzentrum Germany	**SuperMUC** - iDataPlex DX360M4, Xeon E5-2680 8C 2.70GHz, Infiniband FDR IBM	147,456	2,897.0	3,185.1	3,423
10	National Supercomputing Center in Tianjin China	**Tianhe-1A** - NUDT YH MPP, Xeon X5670 6C 2.93 GHz, NVIDIA 2050 NUDT	186,368	2,566.0	4,701.0	4,040

Source: top500.org

Datacenter vs. Cloud vs. HPC

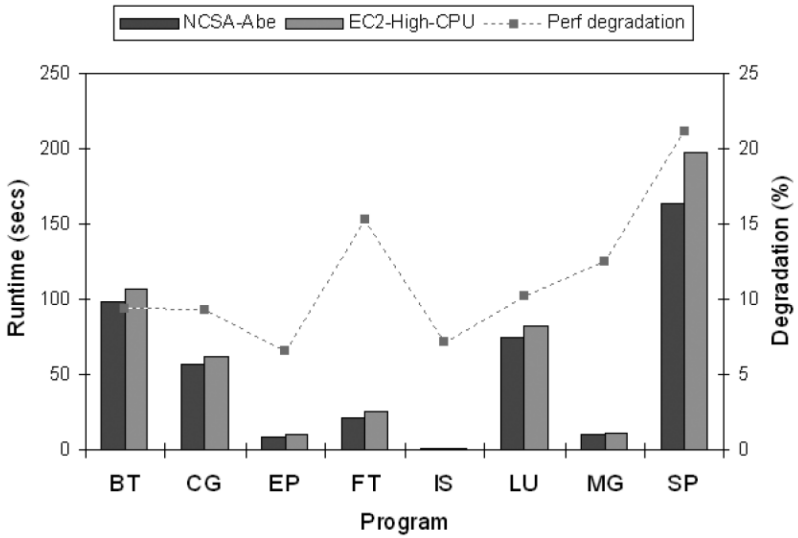
Characteristic	Cloud/Datacenter	HPC
Parallelism	Massively parallel	Massively parallel
Nodes	Commodity x86	Commodity x86
Local storage	Huge (~1-2TB)	Medium (<1TB)
Network	Ethernet	Typically **not** Ethernet
Coupling	*Loose*	*Tight*
Technical details	2x Xeon-SMP 4 cores per socket at 2.33GHz 7GB memory 1690GB storage Network: „High I/O performance“	2x Xeon-SMP 4 cores per socket at 2.33GHz 8GB memory 73GB storage Network: Infiniband

Walker, 2008, benchmarking Amazon EC2 for high-performance scientific computing
;login: The USENIX Magazine, 33(5)
http://www.usenix.org/publications/login/2008-10/openpdfs/walker.pdf

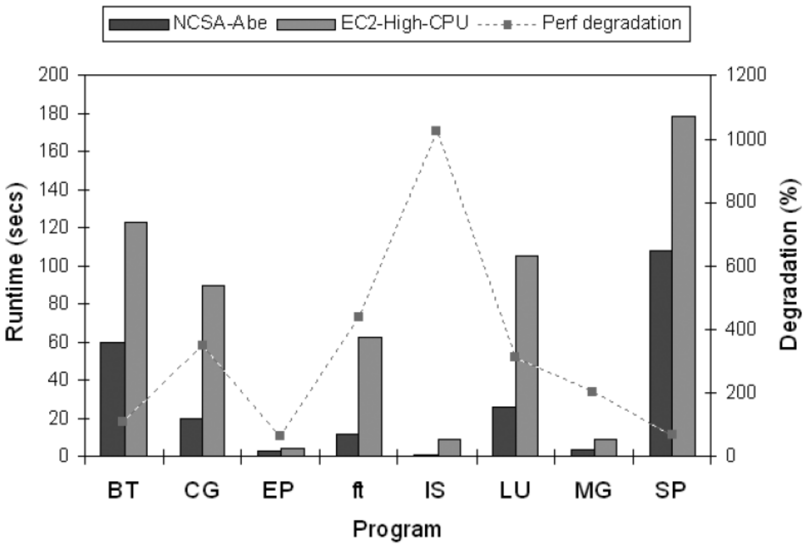
Cloud vs. HPC

- Cloud (Amazon EC2) vs. Cluster (Infiniband, NCSA-Abe)
  - NAS Parallel Benchmarks (NPB)

- Node-level execution
  - OpenMP
  - 8 cores, 1 node



- Cluster-level execution
  - MPI
  - 32 cores (resp. 16)



Summary

- Demand for computing power still not satisfied and won't be satisfied in the near future
  - During the design phase almost every product gets in touch with HPC
  - Scientific questions
  - Technical problems
- Primary tool: high parallelization
  - Interconnection network moves into the spotlight
- Power consumption is a pervasive problem, not only for huge installations
- Browse http://top500.org for plenty of information about facts and trends of HPC

Questions?

Next lecture:

Introduction to Message Passing