

**Einführungsveranstaltung zum Hauptseminar**

## **Hochleistungsrechner: Aktuelle Trends und Entwicklungen**

**Web:**

<http://www.lrr.in.tum.de/lehre/wintersemester-1516/seminare/hochleistungsrechner-aktuelle-trends-und-entwicklungen/>  
<http://www.mnm-team.org/teaching/Seminare/2015ws/Hauptseminar/>

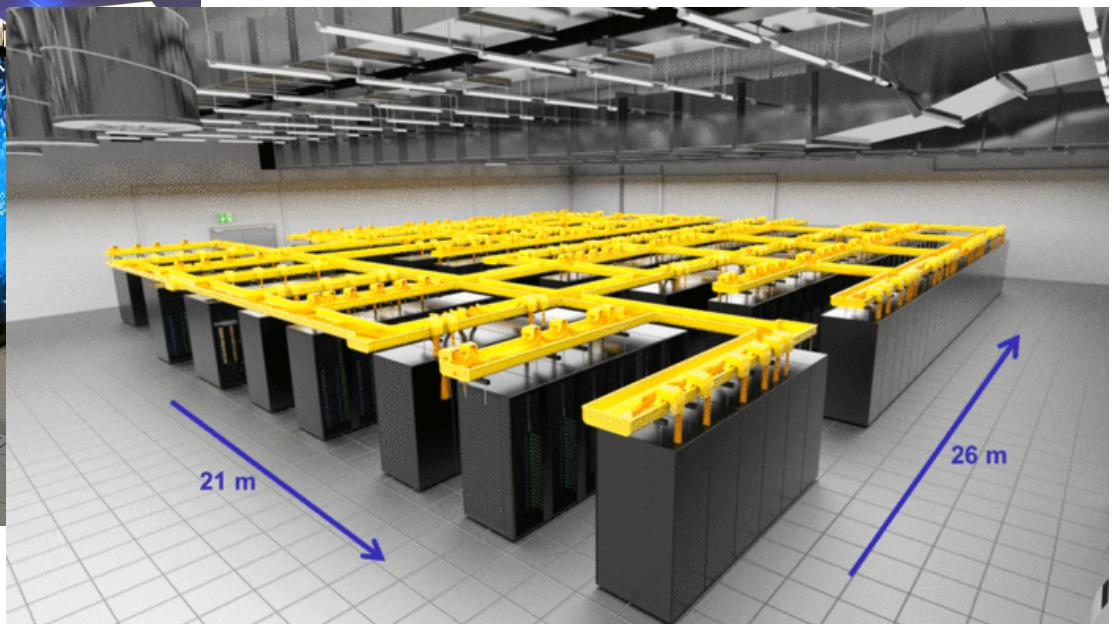


# Hochleistungsrechnen

- Engl. HPC = **High Performance Computing**
- Die „Formel 1“ der Informatik
- Große, schnelle, laute, teure Rechner
- A.k.a. Supercomputer

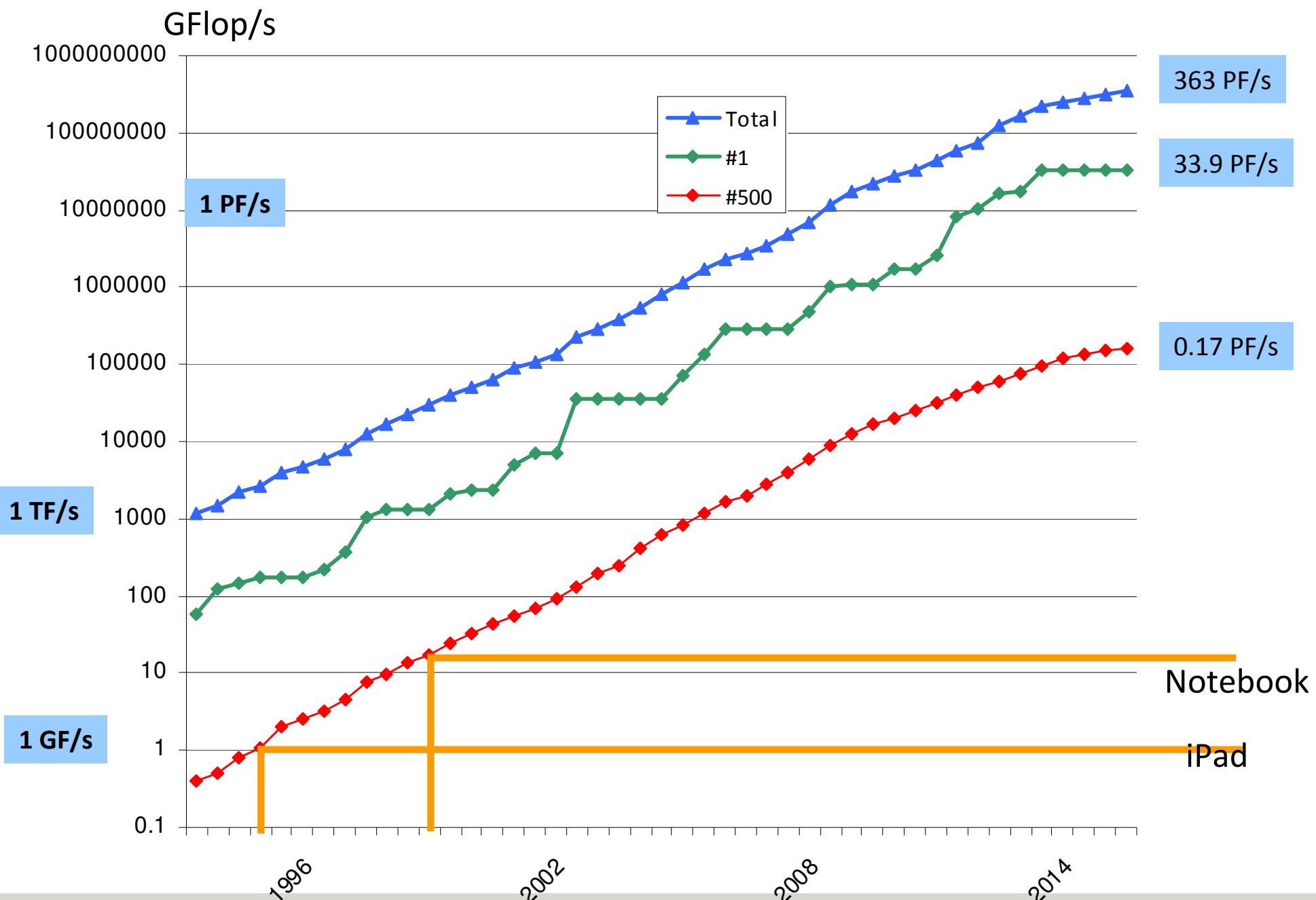


(Titan, Oak Ridge Lab, USA)

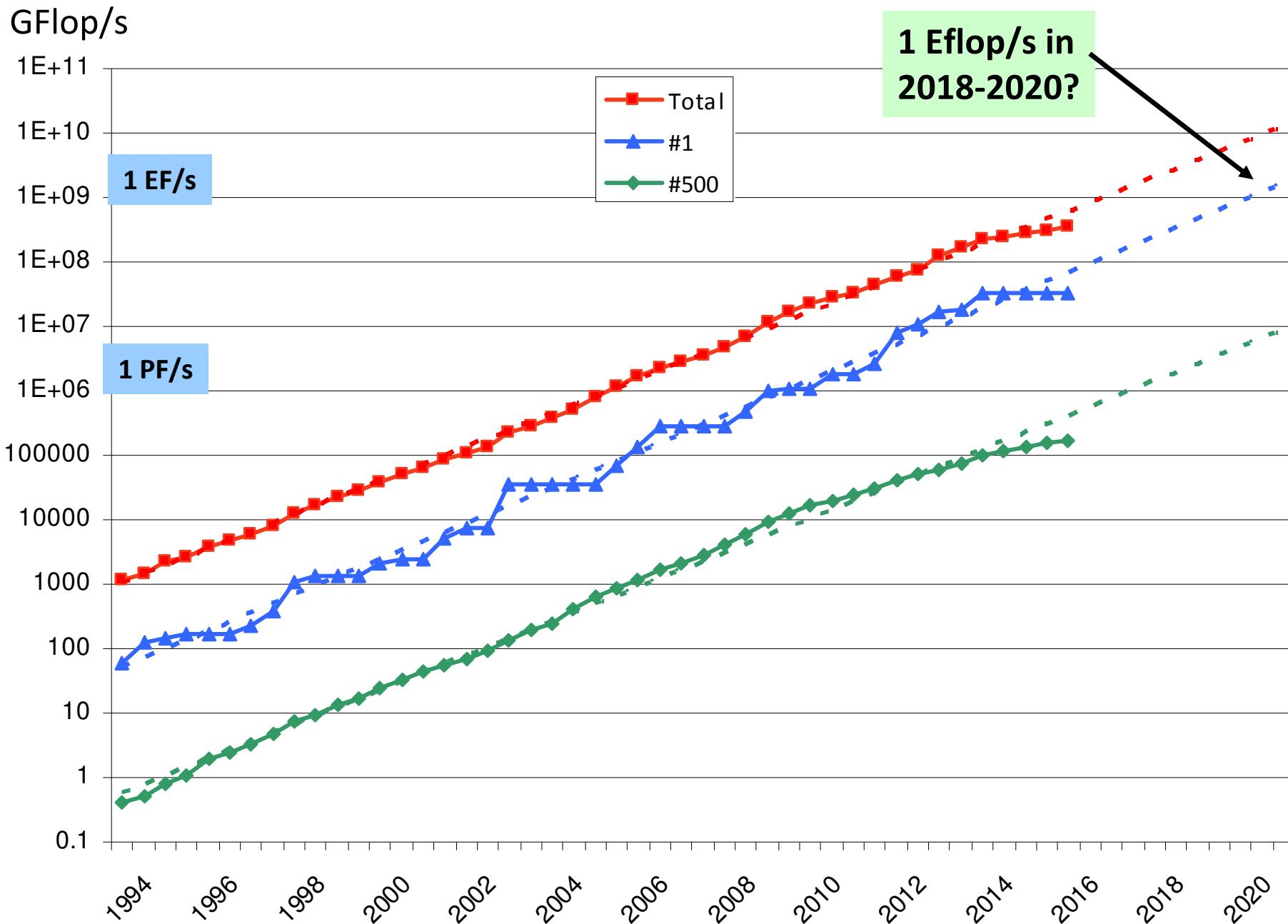


(SuperMUC, Garching, Deutschland)

# Die Top500 Liste (Juni 2015)



# Die Top500 Liste - Extrapolation



# Herausforderungen: Energieeffizienz

## ■ Supercomputing Systeme am LRZ:

System	Era	Peak Performance	Power Consumption	Investment Costs	Total Operating Costs (incl. Power)	Power Bill
HLRB I: Hitachi SR8000	2000 - 2006	1.3 TFLOP/s	0.5 MW	29 M€	13 M€	3 M€
HLRB II: sgi ALTIx 4700	2006 - 2011	62 TFLOP/s	1 MW	35 M€	16 M€	7 M€
SuperMUC: IBM iDataPlex	2012 - 2016	3000 TFLOP/s	3 MW	48 M€	35 M€	22 M€

■ Leistungsfähigkeit von HPC Systemen erhöht sich sogar schneller als Moore's Law

■ Aber: Energieverbrauch wird immer mehr zum entscheidenden Faktor, insbes. Für Exascale

- Ziel: 20 MW für 1 ExaFlop = 50000 MFlops/Watt

# Entwicklung der Energieeffizienz – Green 500

Cray XT line of systems

System	MFLOPS/Watt
Cray XT3 (2004)	60
Cray XT4 (2006)	130
Cray XT5 (2007)	150
Cray XT6 (2009)	260
Cray XE6 (2010)	360
Cray XC30 (2012)	620

IBM BlueGene line of systems

System	MFLOPS/Watt
IBM BlueGene/L (2005)	204
IBM BlueGene/P (2007)	370
IBM BlueGene/Q (2011)	2100

Accelerator-Based

System	MFLOPS/Watt
„Beacon“ IB/Xeon/Xeon Phi (2012)	2500
Cray XK7 (Opteron/NVIDIA Kepler) (2012)	2100

Source: Green500 List

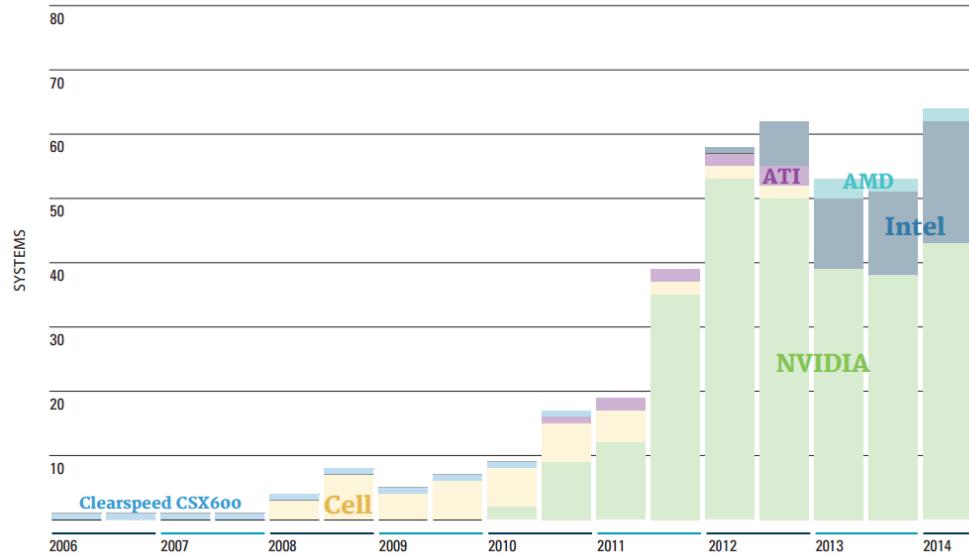
# 2014 Green 500 List

Green500 Rank	MFLOPS/W	Site*	Computer*	Total Power (kW)
1	4,389.82	GSIC Center, Tokyo Institute of Technology	TSUBAME-KFC - LX 1U-4GPU/104Re-1G Cluster, Intel Xeon E5-2620v2 6C 2.100GHz, Infiniband FDR, NVIDIA K20x	34.58
2	3,631.70	Cambridge University	Wilkes - Dell T620 Cluster, Intel Xeon E5-2630v2 6C 2.600GHz, Infiniband FDR, NVIDIA K20	52.62
3	3,517.84	Center for Computational Sciences, University of Tsukuba	HA-PACS TCA - Cray 3623G4-SM Cluster, Intel Xeon E5-2680v2 10C 2.800GHz, Infiniband QDR, NVIDIA K20x	78.77
4	3,459.46	SURFsara	Cartesius Accelerator Island - Bullx B515 cluster, Intel Xeon E5-2450v2 8C 2.5GHz, InfiniBand 4× FDR, Nvidia K40m	44.40
5	3,185.91	Swiss National Supercomputing Centre (CSCS)	Piz Daint - Cray XC30, Xeon E5-2670 8C 2.600GHz, Aries interconnect , NVIDIA K20x Level 3 measurement data available	1,753.66
6	3,131.06	ROMEO HPC Center - Champagne-Ardenne	romeo - Bull R421-E3 Cluster, Intel Xeon E5-2650v2 8C 2.600GHz, Infiniband FDR, NVIDIA K20x	81.41
7	3,019.72	CSIRO	CSIRO GPU Cluster - Nitro G16 3GPU, Xeon E5-2650 8C 2GHz, Infiniband FDR, Nvidia K20m	86.20
8	2,951.95	GSIC Center, Tokyo Institute of Technology	TSUBAME 2.5 - Cluster Platform SL390s G7, Xeon X5670 6C 2.93GHz, Infiniband QDR, NVIDIA K20x	927.86
9	2,813.14	Exploration & Production - Eni S.p.A.	HPC2 - iDataPlex DX360M4, Intel Xeon E5-2680v2 10C 2.8GHz, Infiniband FDR, NVIDIA K20x	1,067.49
10	2,678.41	Financial Institution	iDataPlex DX360M4, Intel Xeon E5-2680v2 10C 2.800GHz, Infiniband, NVIDIA K20x	54.60

# Top500 Systeme mit Akzeleratoren

## ■ Anzahl der Systeme mit Akzeleratoren in der Top500 Liste

### ACCELERATORS/CO-PROCESSORS



## ■ Top 5 (Juni 2015)

	NAME	SPECS	SITE	COUNTRY	CORES	R <sub>MAX</sub> PFLOP/S	POWER MW
1	Tianhe-2 (Milkyway-2)	NUDT, Intel Ivy Bridge (12C, 2.2 GHz) & Xeon Phi (57C, 1.1 GHz), Custom interconnect	NSCC Guangzhou	China	3,120,000	33.9	17.8
2	Titan	Cray XK7, Operon 6274 (16C 2.2 GHz) + Nvidia Kepler GPU, Custom interconnect	DOE/SC/ORNL	USA	560,640	17.6	8.2
3	Sequoia	IBM BlueGene/Q, Power BQC (16C 1.60 GHz), Custom interconnect	DOE/NNSA/LLNL	USA	1,572,864	17.2	7.9
4	K computer	Fujitsu SPARC64 VIIIfx (8C, 2.0GHz), Custom interconnect	RIKEN AICS	Japan	705,024	10.5	12.7
5	Mira	IBM BlueGene/Q, Power BQC (16C, 1.60 GHz), Custom interconnect	DOE/SC/ANL	USA	786,432	8.59	3.95

# HPC Systeme der nahen Zukunft

	Cori <sup>1</sup> (NERSC)	Sierra, Summit <sup>2</sup> (LLNL, ORNL)	Aurora <sup>3</sup> (ANL)
Inst. Year	2016	2017-2018	2018-2019
Peak Perf.	>30 PF	Complex node architecture	180-450 PF
Manycore nodes (70+ cores)	7 MW 3 TF	10 MW >40 TF	Heterogeneous nodes (accelerator based)
No. of Nodes	9,300	3,400	>50,000
Node Architecture	Intel Xeon Phi (Knights Landing) Cray Aries	IBM Power 9 + Nvidia Volta Dual Rail EDR Infiniband	Intel Xe (Knight) 2 <sup>nd</sup> gen Intel Omni Path
Memory	DDR4 + On-package high BW memory	DRAM + stacked DRAM	7 PB DRAM + Persistent Memory
	Integrated High-speed interconnect	www.olcf.ornl.gov/summit/	3: http://aurora.alcf.anl.gov/

- Aktuelle Trends und Entwicklungen bei HPC Systemen
  - Welche neuen Entwicklungen gibt es in bezug auf HW, Programmierung, Laufzeit- und Betriebssystemen, ...
  - Es tut sich einiges...
- “Evolutionär” und kurz vor der Einführung, zB.
  - Persistent Memory (NVRAM)
  - High bandwidth RAM (zB. Xeon Phi MCDRAM)
- “Revolutionär” und radikal, im Prototypen-Stadium, zB
  - Neuromorphic chips
  - Adiabatic Quantum Computer

- Ziel des Seminars: Darstellung eines abgegrenzten wissenschaftlichen Themas
  - Grundlage sind aktuelle publizierte Arbeiten
  - Selbständige Erarbeitung von wissenschaftlichen Texten
  - Kritische Auseinandersetzung und Bewertung des Materials
  - Erstellen einer Arbeit, in dem das Thema umfassend besprochen wird
    - Korrekte Verwendung von Literatur
    - Richtiges Zitieren
  - Präsentation der Ergebnisse im Rahmen eines Vortrags am Semesterende
  - Unterstützt durch Betreuer während des Semesters

# Organisatorisches

## ■ Bewertungsgrundlagen für Leistung

- Schriftliche Ausarbeitung: Verständlichkeit, Inhalt, Gliederung, Literatur, Schreibstil
- Präsentation: Verständlichkeit, Vortragsstil, Beherrschung des Themas
- Mitarbeit: Präsenz bei Veranstaltungen, Einhalten von Abgabeterminen, Fortschritt über das Semester, Zusammenarbeit mit Betreuer

## Themenübersicht (1)

### ■ Neuromorphic Chips / IBM TrueNorth / SpiNNaker

*Neurons and synapses in hardware to model the brain*

- Merolla, Paul A., et al. "A million spiking-neuron integrated circuit with a scalable communication network and interface." *Science* 345.6197 (2014): 668-673.
- Painkras, Eustace, et al. "Spinnaker: a multi-core system-on-chip for massively-parallel neural net simulation." *Custom Integrated Circuits Conference (CICC), 2012 IEEE*. IEEE, 2012.

### ■ Adiabatic Quantum Computers (D-Wave)

- McGeoch, Catherine C., and Cong Wang. "Experimental evaluation of an adiabatic quantum system for combinatorial optimization." *Proceedings of the ACM International Conference on Computing Frontiers*. ACM, 2013.
- Rønnow, Troels F., et al. "Defining and detecting quantum speedup." *Science* 345.6195 (2014): 420-424.

## ■ Micron Automata

Noch frei!

*Special purpose accelerator for computing with finite state automata / regular expressions, ...*

- Sabotta, Christopher. "Advantages and challenges of programming the Micron Automata Processor." (2013).

## ■ Anton / Anton 2 (D.E. Shaw)

Noch frei!

*Massively Parallel Special Purpose System for MD calculations*

- Shaw, David E., et al. "Anton 2: raising the bar for performance and programmability in a special-purpose molecular dynamics supercomputer." *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*. IEEE Press, 2014.

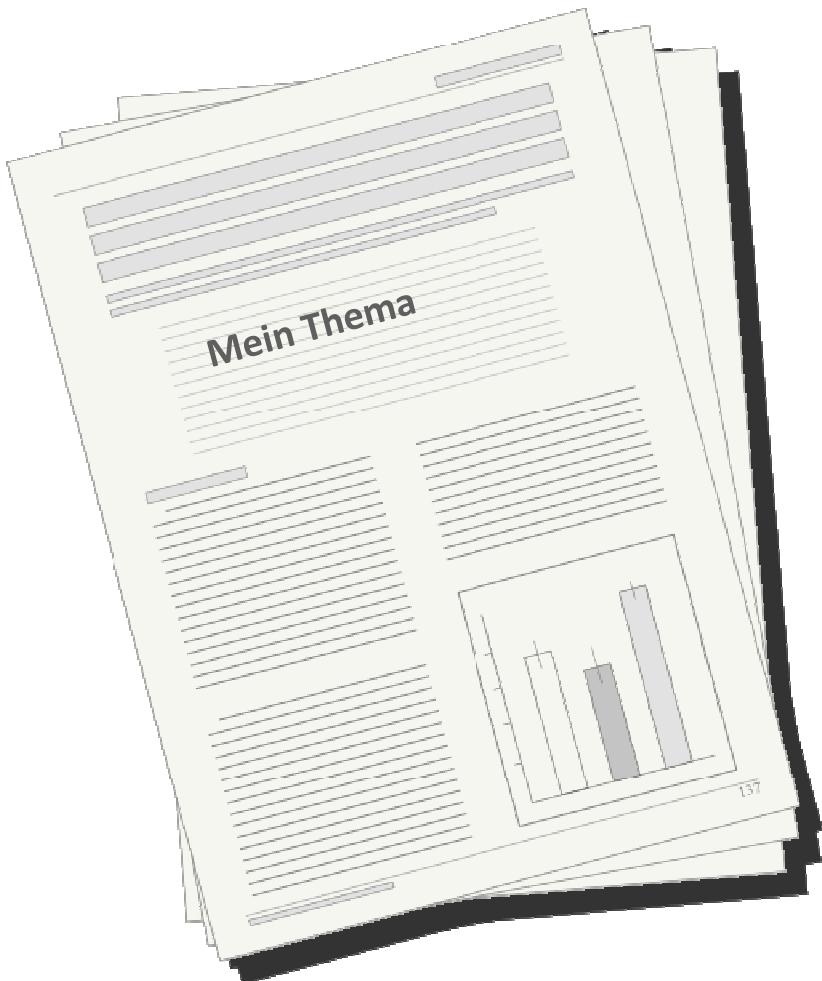
## Themenübersicht (3)

- Hochgeschwindigkeitsnetze: Cray Aries / Dragonfly Topologie
  - Faanes, Greg, et al. "Cray cascade: a scalable HPC system based on a Dragonfly network." *Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis*. IEEE Computer Society Press, 2012.
  - Besta, Maciej, and Torsten Hoefer. "Slim fly: a cost effective low-diameter network topology." *High Performance Computing, Networking, Storage and Analysis, SC14: International Conference for*. IEEE, 2014.
- Non-Volatile RAM
  - Li, Dong, et al. "Identifying opportunities for byte-addressable non-volatile memory in extreme-scale scientific applications." *Parallel & Distributed Processing Symposium (IPDPS), 2012 IEEE 26th International*. IEEE, 2012.

- Dynamic Task Execution Systems / Runtime Systems: Stanford Legion, Open Communitiy Runtime
  - <http://legion.stanford.edu/>
  - Bauer, Michael, et al. "Legion: expressing locality and independence with logical regions." Proceedings of the international conference on high performance computing, networking, storage and analysis. IEEE Computer Society Press, 2012.
- Fehlertoleranz: Containment domains Noch frei!
  - Chung, Jinsuk, et al. "Containment domains: A scalable, efficient and flexible resilience scheme for exascale systems." *Scientific Programming* 21.3-4 (2013): 197-212.

- SOCs (Systems on a Chip) for HPC
  - Rajovic, Nikola, et al. "Supercomputing with commodity CPUs: are mobile SoCs ready for HPC?." *High Performance Computing, Networking, Storage and Analysis (SC), 2013 International Conference for.* IEEE, 2013.
- Programmieren von FPGAs mit OpenCL
- Architektur von AMD GPUs/APUs und ihre Programmierung
- Intel Xeon Phi (Many Integrated Cores Architektur) und Programmierung

# Ziele



**Schriftliche Ausarbeitung**  
Erstellt mit LaTeX  
Hauptseminar: 8-10 Seiten



**Symbolbild!**

**Präsentation im Blockseminar am Ende des Semesters**  
Hauptseminar: ca. 25-30 Minuten Vortrag + Diskussion

# Themen- und Betreuer

	<b>Teilnehmer</b>	<b>Thema</b>	<b>Betreuer</b>
TUM	Thorsten Fuchs	Programmierung von FPGAs für HPC (z.B. mit OpenCL)	Carsten Trinitis
TUM	Fabio Gratl	Manycore-Architekturen (z.B. Xeon Phi)	Carsten Trinitis
TUM	Johannes Klicpera	Verbindung von CPUs und GPUs: AMD APUs	Jens Breitbart
TUM	Markus Loipfinger	Non-Volatile Memory (NVM): Typen und Nutzung für HPC	Josef Weidendorfer
TUM	Paul Nieleck	Mobile Architekturen für HPC (z.B. AMD ARM Opterons, EU Montblanc)	Matthias Maiterth
TUM	Dmitry Tegunov	"Processing in Memory" (IBM Active Memory Cube)	Josef Weidendorfer
TUM	Andreas Seibold	Aktuelle Speichertechnologien: High-Bandwidth-Memory/Hybrid-Memory-Cube	Jens Breitbart
TUM	Sophie Sepp	<del>Heterogene Systeme und passende Laufzeitsysteme, (StarPU, OMPSS)</del>	<del>Tobias Fuchs</del>
LMU	Daniyal Kazempour	Neuromorphic Chips (IBM TrueNorth / Spinnacker)	Tobias Fuchs
LMU	Stefan Effenberger	Hochleistungsnetze (Cray Aries, Extoll)	Matthias Maiterth
LMU	Jon-Magnus Maier	Adiabatic Quantum Computers	Karl Fürlinger
LMU	Matthias Höschel	ANTON / ANTON2?	Karl Fürlinger

# Ablauf des Seminars und Termine

Okt.

Nov.

Dez.-Jan.



Einführungsveranstaltung  
Di 20.10.2015 / LRZ, Garching

Präsentations- und Arbeitstechnik  
Do. 5. Nov. 2015, Oettingenstr. 67, Raum 061, 16:15 Uhr

*Abgabe: Gliederung, 8.11.2015*

Alle Abgaben per Email  
an den Betreuer

Gastvortrag zum Thema: TBA

*Abgabe: Entwurf Seminararbeit, 6.12.2015*

*Abgabe: Finale Seminararbeit, 20.12.2015*

*Abgabe: Präsentationsfolien, 10.01.2016*

Blockseminar Frauenchiemsee  
3.-4.02.2016 (Mittwoch-Donnerstag)