

Einführungsveranstaltung zum Hauptseminar

Akzeleratorarchitekturen

Prof. Dr. Dieter Kranzlmüller

Dr. Karl Furlinger - fuerling@nm.ifi.lmu.de

MNM-Team

LMU München

Prof. Dr. Carsten Trinitis - trinitic@in.tum.de

Dr. Josef Weidendorfer - weidendo@in.tum.de

Dr. Jens Breitbart - j.breitbart@tum.de

Lehrstuhl für Rechnertechnik und Rechnerorganisation

TU München

Web:

<http://www.lrr.in.tum.de/public/HauptseminarAkzeleratorarchitekturenWS14>

<http://www.nm.ifi.lmu.de/teaching/Seminare/2014ws/Hauptseminar/>





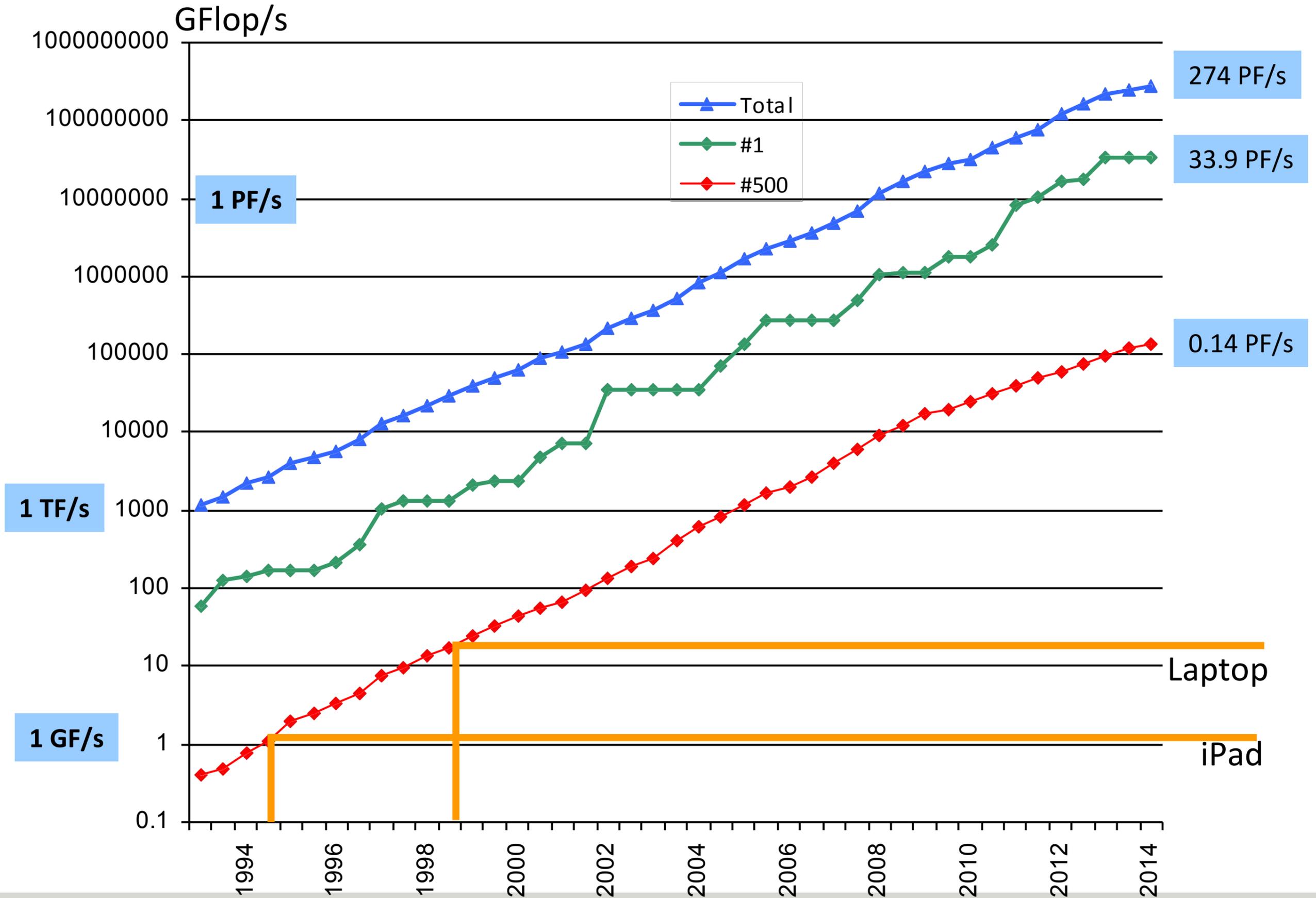
- Akzeleratoren := Beschleunigerhardware die insbes. Im Hochleistungsrechnen (HPC) in letzter Zeit verstärkt Einsatz findet
- Warum Akzeleratoren?
 - Um schneller zum Ergebnis zu gelangen (time to solution)
 - Um Effizienz zu steigern (zB Platz- oder Energiebedarf)
- Fragestellungen
 - Wann werden Akzeleratoren eingesetzt?
 - Wie sind sie aufgebaut?
 - Wie verwendet man sie?
 - Einschränkungen?

■ Supercomputing Systeme am LRZ:

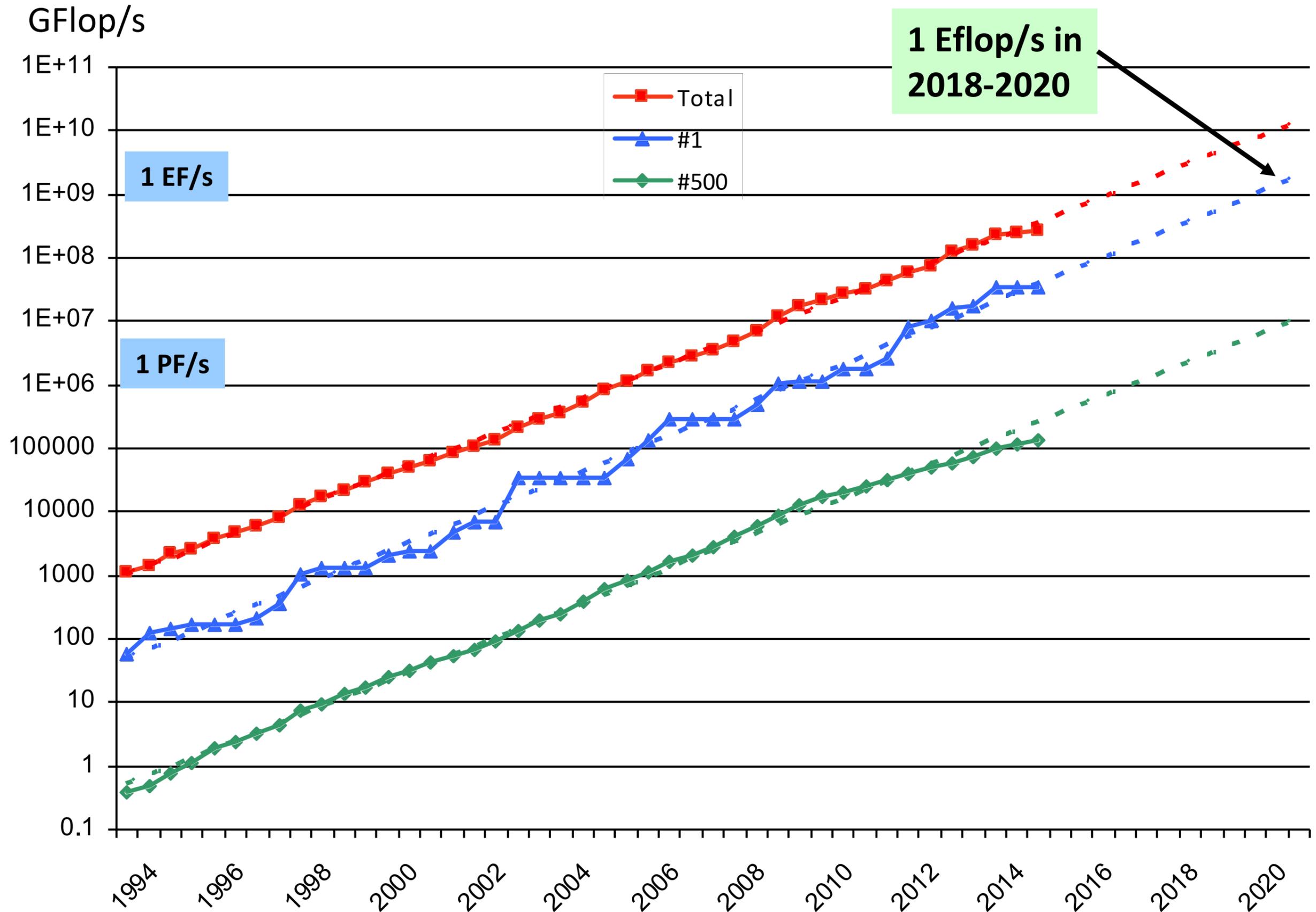
System	Era	Peak Performance	Power Consumption	Investment Costs	Total Operating Costs (incl. Power)	Power Bill
HLRB I: Hitachi SR8000	2000 - 2006	1.3 TFLOP/s	0.5 MW	29 M€	13 M€	3 M€
HLRB II: sgi ALTIX 4700	2006 - 2011	62 TFLOP/s	1 MW	35 M€	16 M€	7 M€
SuperMUC: IBM iDataPlex	2012 - 2016	3000 TFLOP/s	3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

The Top500 List (June 2014)



Die Top 500 Liste - Extrapolation



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

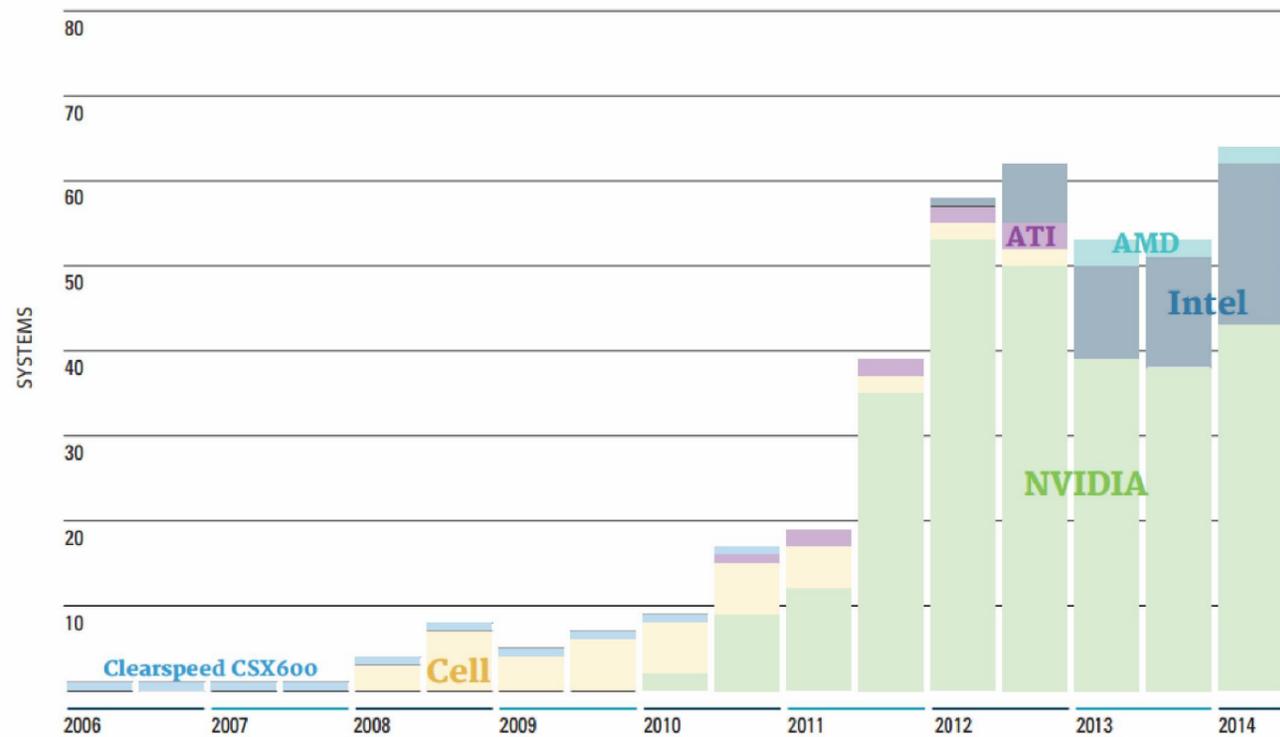
Source: Green500 List

Current 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 4x 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

■ Anzahl der Systeme mit Akzeleratoren in der Top500 Liste

ACCELERATORS/CO-PROCESSORS

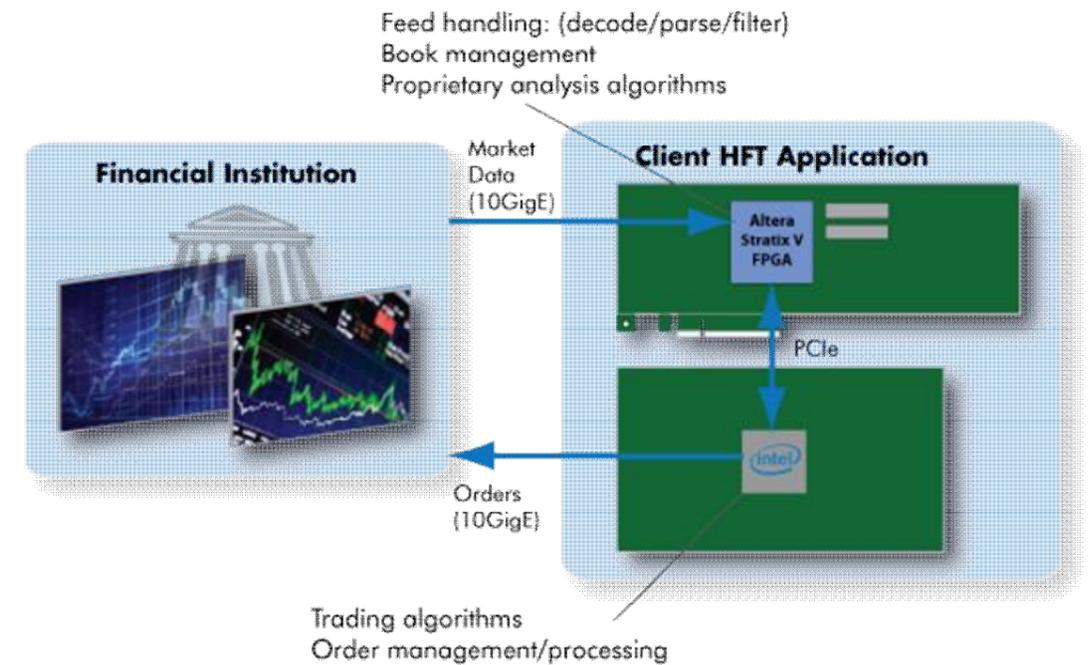


■ Top 5 (Juni 2014)

	NAME	SPECS	SITE	COUNTRY	CORES	R _{MAX} PFLDP/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

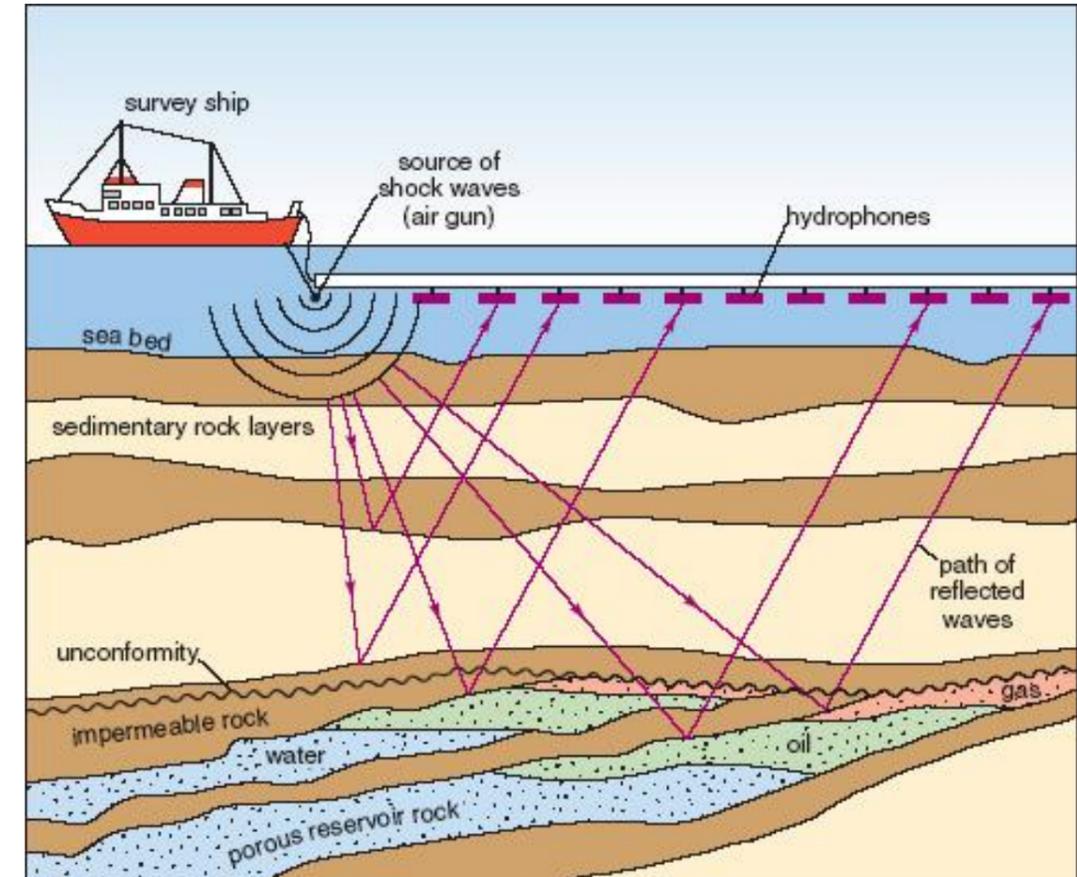
■ Algorithmic / low-latency trading

- Beste Leistung bringt unmittelbare Marktvorteile



■ Seismic Exploration

- Große Datenmengen, großer Berechnungsaufwand



- Das Seminar wird angeboten in Zusammenarbeit zw. LMU und TU München
 - Hauptseminar für Master Studenten entsprechend Modulkatalog

- 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

■ Bewertungsgrundlagen für Leistung

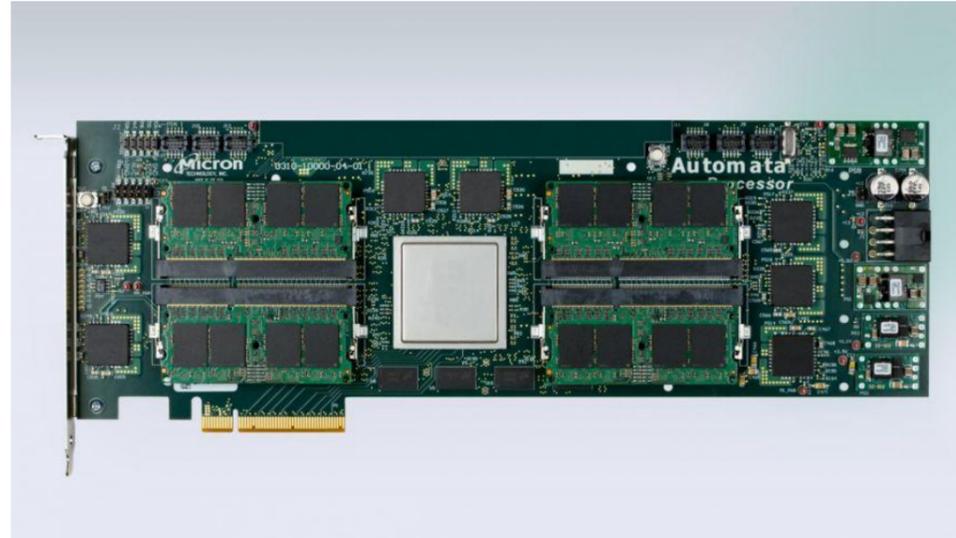
- Schriftliche Ausarbeitung: Verständlichkeit, Inhalt, Gliederung, Literatur, Schreibstil
- Präsentation: Verständlichkeit, Vortragsstil, Beherrschung des Themas
- Mitarbeit: Präsenz bei Veranstaltungen, Einhalten von Abgabeterminen, Fortschritt über das Semester, Zusammenarbeit mit Betreuer

- Übersicht über das Thema “Akzeleratoren”
 - Hintergrund, Einsatzgebiete, Varianten, Motivation, Beschränkungen
- Architektur von modernen NVIDIA GPUs
- Architektur von AMD GPUs/APUs
- Intel Xeon Phi (Many Integrated Cores Architektur)
- FPGA Maxeler: Architektur
- FPGA Maxeler: Beispiele
- Beschleuniger in mobilen Chips (SoCs: Tegra3, OMAP, Apple A6/7)

- Programmierung mit CUDA
- Programmierung mit OpenCL
- Programmierung mit Pragmas: OpenMP/OpenACC
- Programmierung mit OmpSS
- Programmierung mit StarPU
- Simulatoren für Beschleunigerhardware
- Hybride Programmierung von Rechnerverbänden mit Akzeleratoren
- Akzeleratoren und Virtualisierung

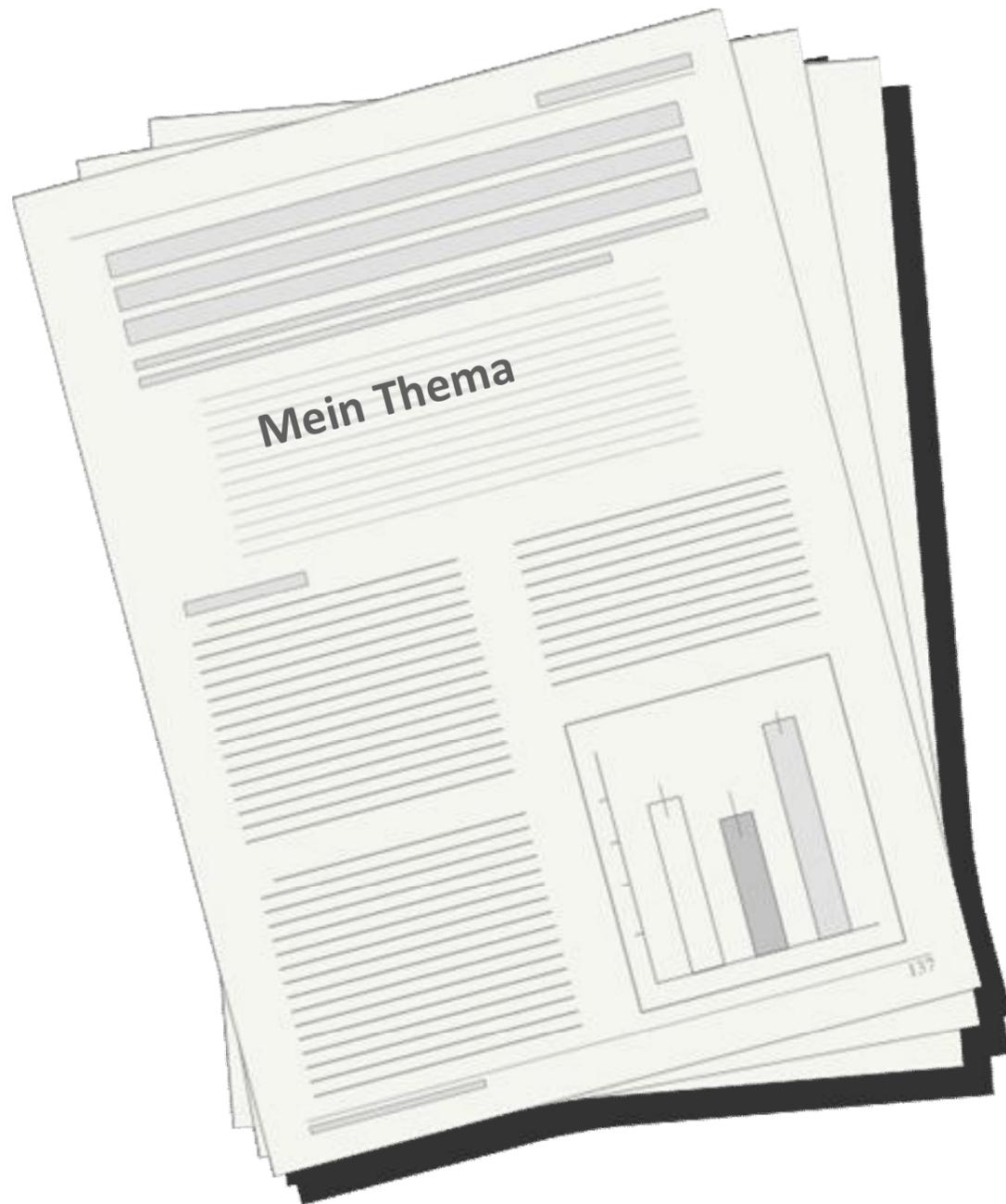
■ Micron Automata

- System zur effizienten Bearbeitung von endl. Automaten, reg. Ausdrücken,... (Micron)



■ Anton / Anton 2

- Massiv-paralleles System zur MD Berechnung (D.E. Shaw)



Schriftliche Ausarbeitung

Erstellt mit LaTeX

Hauptseminar: 15-20 Seiten



Symbolbild!

Präsentation im Blockseminar am Ende des Semesters

Hauptseminar: ca. 30 Minuten Vortrag + Diskussion

■ TUM

- Oliver Barta - **Simulatoren für Beschleuniger - Josef**
- Nikolaus Dafinger - **Maxeler Data Flow Engine: Architektur - Carsten**
- Nico Hartmann - **Beschleuniger für Suche in Zeichenketten: "Micron Automata Processor" - Karl**
- Florian Janßen - **Maxeler Data Flow Engine: Beispiele - Carsten**
- Josef Stark - **Programmierung mit OpenCL - Jens**

■ LMU

- Fußenegger Markus - **Programmieren mit CUDA - Jens**
- Kirsch Michael - **AMD Architekturen - Jens**
- Kißlinger Michael - **Übersicht über das Thema Akzeleratoren - Carsten**
- Lohrer Andreas - **Eigenes Thema: IBM DB2 Analytics Accelerator for z/OS - Karl**
- Picker Stephan - **Eigenes Thema: Radix Sort auf Akzeleratorarchitekturen - Karl**
- Saumweber Bernhard – **Programmieren mit Pragmas (OpenACC/OpenMP) - Josef**
- Seifert Michael - **Intel Xeon Phi - Karl**
- Tänzel Tobias - **Virtualisierung und Akzeleratoren - Josef**
- Wallwitz Andrej – **FPGAs - Carsten**
- Wohlschlögl Markus - **Nvidia Architekturen - Jens**

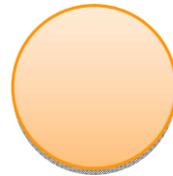
Ablauf des Seminars und Termine

Okt.



Einführungsveranstaltung
Do 16.10.2014 / TUM in Garching

Nov



Präsentations- und Arbeitstechnik
Do. 30. Oktober 2014, Oettingenstr. 67, Raum 061, 16:15 Uhr

Abgabe: Gliederung, 2.11.2014

Alle Abgaben per Email
an den Betreuer



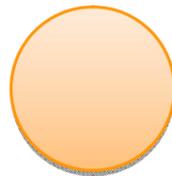
Gastvortrag zum Thema: TBA

Dez.-Jan.

Abgabe: Entwurf Seminararbeit, 7.12.2014

Abgabe: Finale Seminararbeit, 21.12.2014

Abgabe: Präsentationsfolien, 6.01.2015



Blockseminar Frauenchiemsee
14.-15.01.2015 (Mittwoch-Donnerstag)