



Gemeinsames TUM-LMU Seminar



Hochleistungsrechner: Aktuelle Trends und Entwicklungen

Einführungsveranstaltung

18. Oktober 2017

Web:

<https://www.lrr.in.tum.de/lehre/wintersemester-1718/seminare/hochleistungsrechner-aktuelle-trends-und-entwicklungen/>

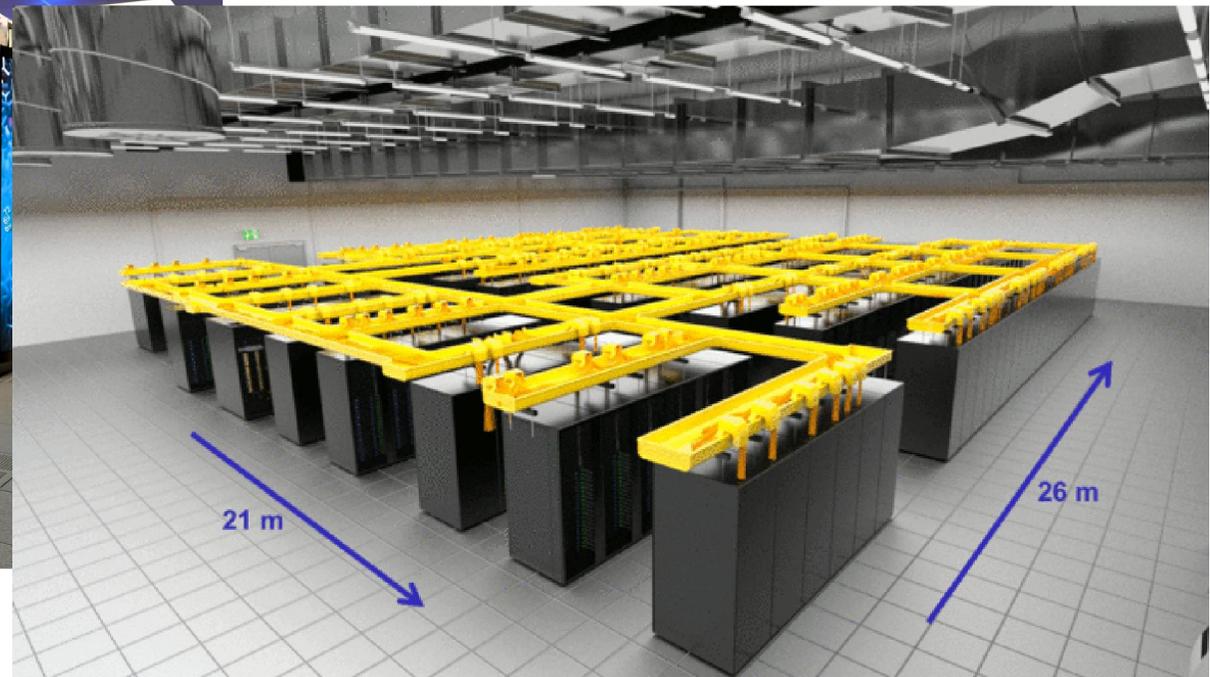
<http://www.mnm-team.org/teaching/Seminare/2017ws/HPCTrends/>



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



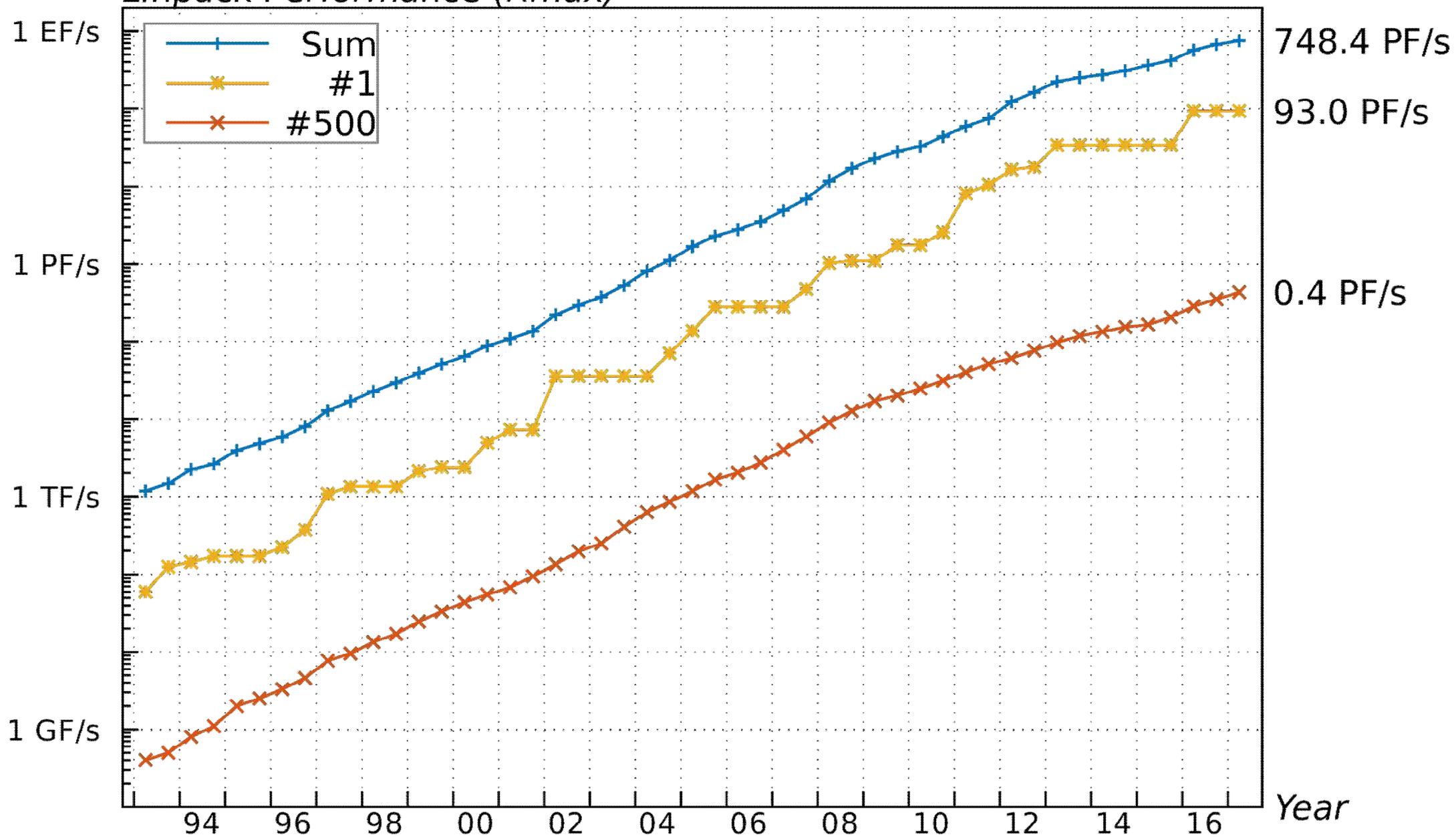
(Titan, Oak Ridge Lab, USA)



(SuperMUC, Garching, Deutschland)

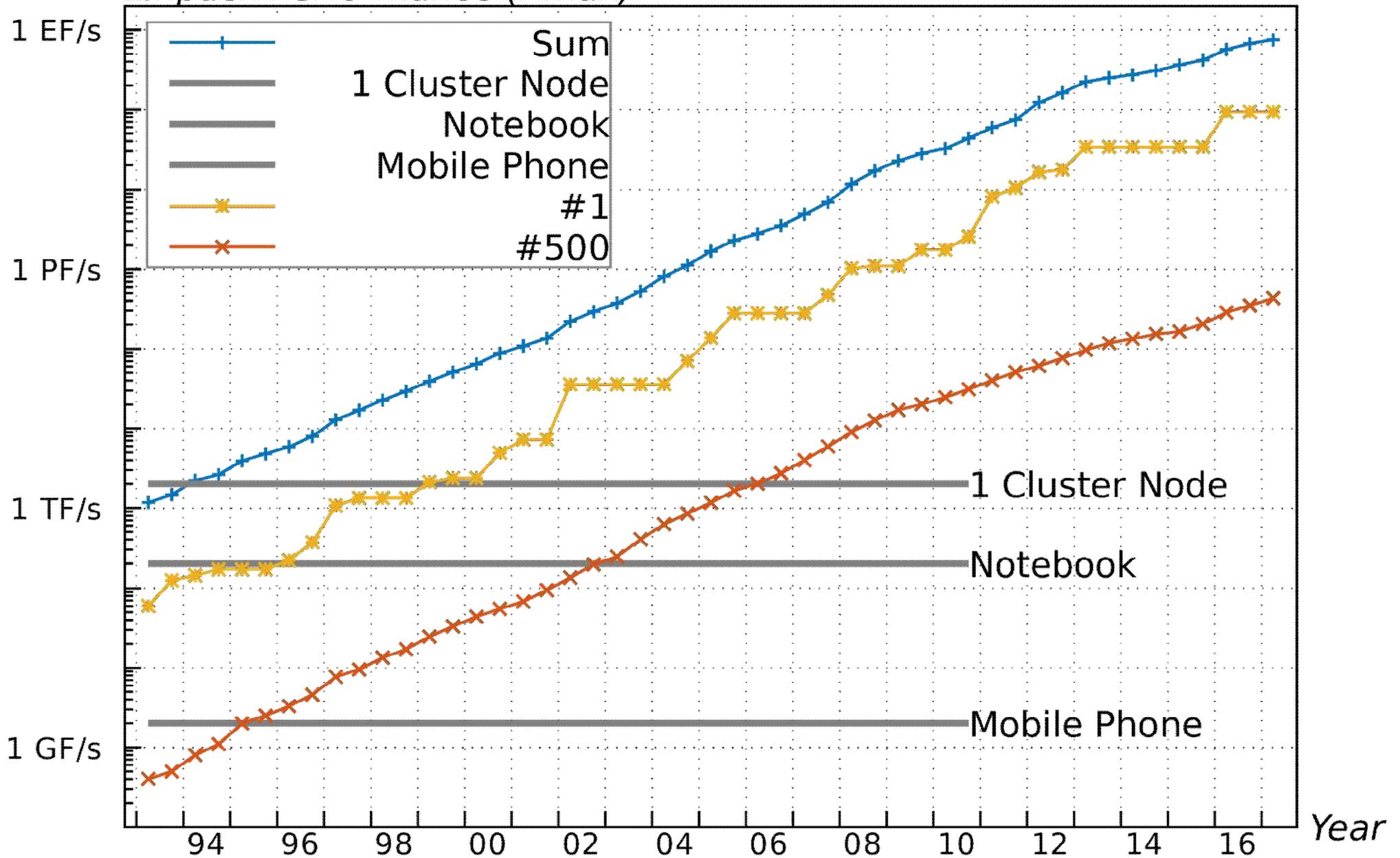
The Top 500 List (2)

Linpack Performance (Rmax)

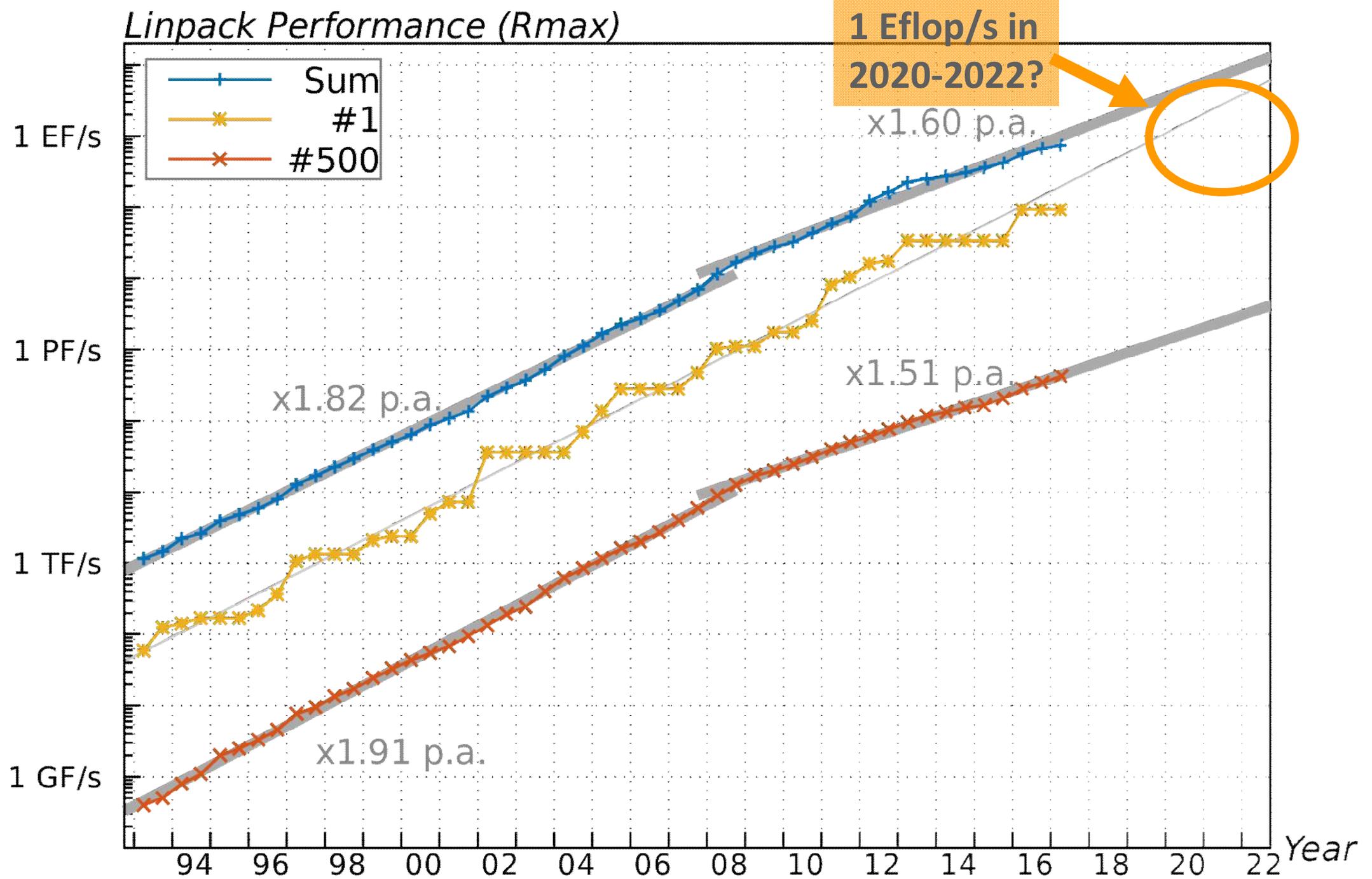


The Top 500 List (3)

Linpack Performance (Rmax)



The Top 500 List (Extrapolation)



■ 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

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

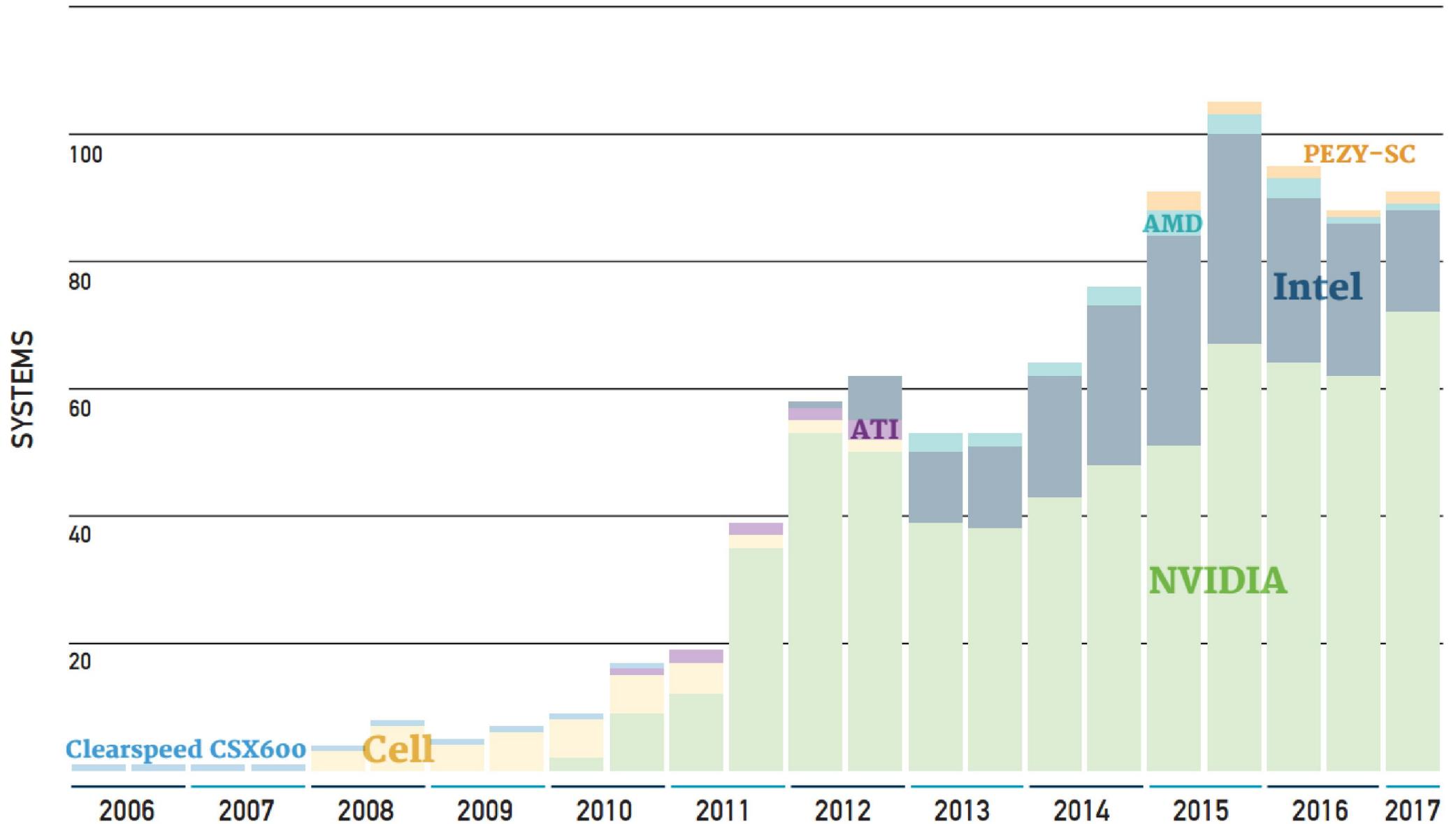
System	MFLOPS/Watt
„Beacon“ IB/Xeon/Xeon Phi (2012)	2500
Cray XK7 (Opteron/NVIDIA Kepler) (2012)	2100

Source: Green500 List

Green 500 List (June 2017)

Rank	TOP500 Rank	System	Cores	Rmax (TFlop/s)	Power (kW)	Power Efficiency (GFlops/watts)
1	61	TSUBAME3.0 - SGI ICE XA, IP139-SXM2, Xeon E5-2680v4 14C 2.4GHz, Intel Omni-Path, NVIDIA Tesla P100 SXM2 , HPE GSIC Center, Tokyo Institute of Technology Japan	36,288	1,998.0	142	14.110
2	465	kukai - ZettaScaler-1.6 GPGPU system, Xeon E5-2650Lv4 14C 1.7GHz, Infiniband FDR, NVIDIA Tesla P100 , ExaScalar Yahoo Japan Corporation Japan	10,080	460.7	33	14.046
3	148	AIST AI Cloud - NEC 4U-8GPU Server, Xeon E5-2630Lv4 10C 1.8GHz, Infiniband EDR, NVIDIA Tesla P100 SXM2 , NEC National Institute of Advanced Industrial Science and Technology Japan	23,400	961.0	76	12.681
4	305	RAIDEN GPU subsystem - NVIDIA DGX-1, Xeon E5-2698v4 20C 2.2GHz, Infiniband EDR, NVIDIA Tesla P100 , Fujitsu Center for Advanced Intelligence Project, RIKEN Japan	11,712	635.1	60	10.603
5	100	Wilkes-2 - Dell C4130, Xeon E5-2650v4 12C 2.2GHz, Infiniband EDR, NVIDIA Tesla P100 , Dell University of Cambridge United Kingdom	21,240	1,193.0	114	10.428

ACCELERATORS/CO-PROCESSORS



HPC Systeme der (nahen) Zukunft (USA)

	Cori ¹ (NERSC)	Sierra, Summit ² (LLNL, ORNL)	Aurora ³ (ANL)
Inst. Year	2016	2017-2018	2018-2019
Peak Perf.	>30 PF	Complex node architecture	180-450 PF
Power	7 MW	10 MW	
Performance/Power	3 TF	>40 TF	
No. of Nodes	9,300	3,400	>50,000
Node	Intel Xeon Phi (Knights Landing)	IBM Power 9 + Nvidia Volta	Intel Xeon (Knights Landing)
Interconnect	Cray Aries	Dual Rail EDR Infiniband	2 nd gen Intel Omni Path
Memory	DDR4 + On-package high BW memory	DRAM + stacked DRAM	7 PB DRAM + Persistent Memory

Manycore nodes (70+ cores)

Complex node architecture

Heterogeneous nodes (accelerator based)

Heterogeneous memory (regular DRAM+stacked RAM)

Persistent memory (NVRAM)

Integrated High-speed interconnect

1: <https://www.nersc.gov/users>

2: www.olcf.ornl.gov/summit/

3: <http://aurora.alcf.anl.gov/>

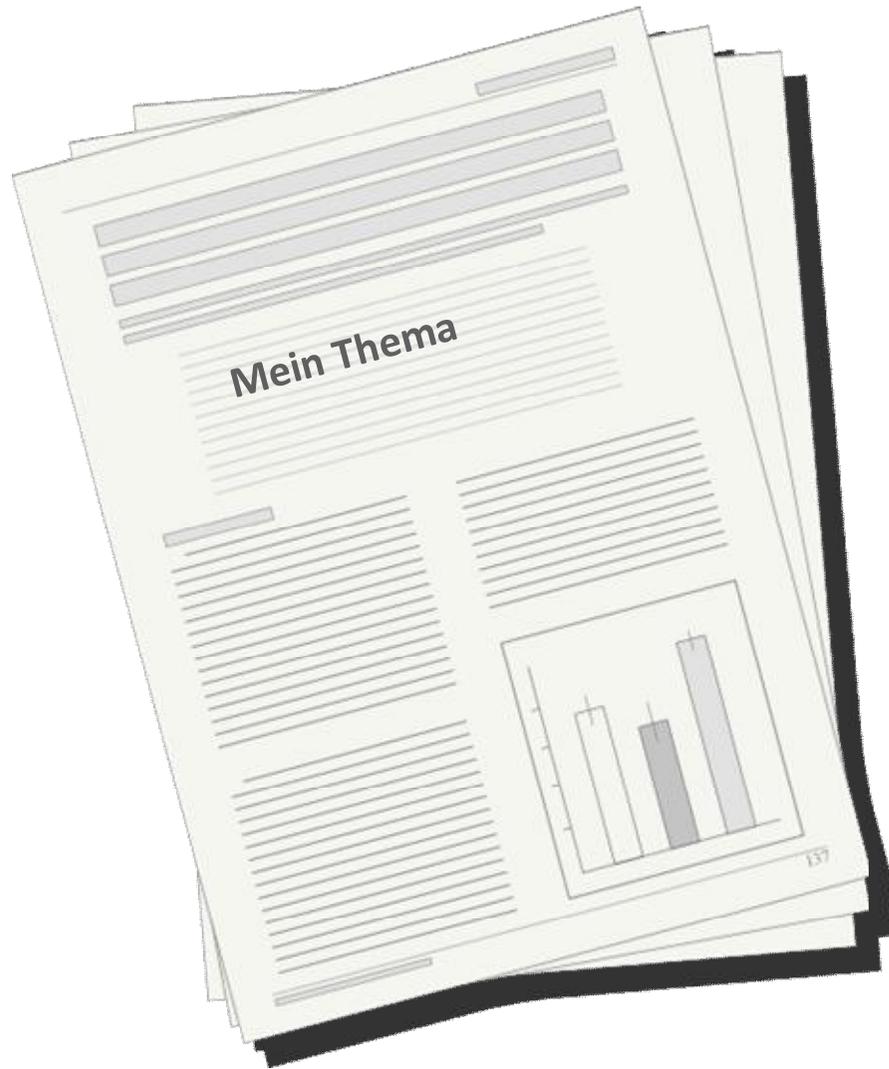
- Aktuelle Trends und Entwicklungen bei HPC Systemen
 - Welche neuen Entwicklungen gibt es in bezug auf HW, Programmierung, Laufzeit- und Betriebssystemen, ...

- “Evolutionär” und kurz vor der Martreife, bzw bereits verfügbar, 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: Erarbeitung und Präsentation 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



Schriftliche Ausarbeitung

Erstellt mit LaTeX

Hauptseminar: 8-10 Seiten



Symbolbild!

Präsentation im Blockseminar am Ende des Semesters

Hauptseminar: ca. 25-30 Minuten Vortrag, danach Diskussion

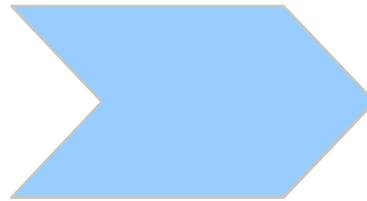
Metapher des Seminars: “Wiss. Konferenz”

Recherche,
Erstellung der
Arbeit



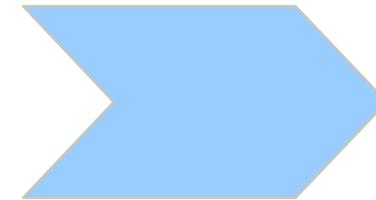
*Recherche,
Erstellung der
Seminararbeit*

Begutachtung
durch “Peers”



*Jede Seminararbeit
wird von 2 anderen
Teilnehmern gelesen
und begutachtet.*

Vortrag auf der
Konferenz



*Seminar ist in Abschnitte
(“Sessions”) gegliedert
die jeweils von einem
Teilnehmer geleitet werden.*

- (1) Neue Hochgeschwindigkeitsnetze (z.B. EXTOLL, OmniPath)
- (2) Aktuelle knoteninterne Verbindungsstrukturen (Intel UPI/NVLink)
- (3) Beschleuniger für "Deep Learning" (Intel Knights Mill, NVidia Volta, Google TPU, IBM TrueNorth, ...)
- (4) Beschleuniger für Spezialaufgaben (Zustandsautomaten (Micron Automata), Proteinstrukturanalyse (Anton), ...)
- (5) FPGA für HPC (Intel Xeon mit integriertem FPGA, ...)
- (6) Speichertechnologien (NVRAM, 3D Stacking, 3DXPoint) und ihre Nutzung

- (7) Neuerungen in aktuellen Mikroarchitekturen (AVX512, AMD Ryzen, ...)
- (8) Aktuelle GPU-Generationen (Nvidia Volta, AMD Polaris) (**vergeben**)
- (9) Fehlertoleranz im HPC (FTMPI, LAIK, ...) (**vergeben**)
- (10) Task-basierte Programmiermodelle (HALadapt, Charm++)
- (11) Neue OS-Konzepte (Unikernel, LibraryOS, HermitCore, ...)
- (12) Maschinelles Lernen im HPC (Ausfallvorhersage, Energieoptimierung, ...) (**vergeben**)

Ablauf des Seminars und Termine

Okt.



Einführungsveranstaltung
18. Oktober 2017

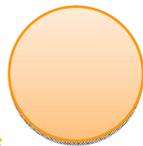


Präsentations- und Arbeitstechnik
TBA

Nov.

Abgabe: Gliederung, 10. Nov. 2017

Alle Abgaben per
Email an den Betreuer



Gastvortrag zum Thema: TBA

Dez.-Jan.

Abgabe: Erste vollständige Version der Arbeit, 15. Dezember 2017

Abgabe: Gutachten, 22. Dezember 2018

Abgabe: Finale Seminararbeit, 12. Januar 2018

Abgabe: Präsentationsfolien, Mitte Januar 2018



Blockseminar Frauenchiemsee
Wahrscheinlich 30./31. Januar 2018