→ Aktuelle Technologien: immer schneller (und grüner)
Ranking in den TOP500
und trotzdem gibt es noch Probleme …

4.1 Technologiefortschritte
4.2 State of the Art
4.3 aktuelle Probleme

4.1 Technologiefortschritte (I)

- Zur Erinnerung (→ Historie):
  - 60er Jahre: CDC 6600 Control Data Corp.
    mit 10 Funktionseinheiten für verschiedene Operationen in der CPU
    + 10 E/A-Prozessoren
  - 70er Jahre: erster Vektorrechner: Cray-1 (1976)
    mehrere gleiche Ops über Vektordaten
    10 Pipelinestufen für Gleitkomma-Ops

    erste massiv parallele Systeme (SIMD)
    ILLIAC (1972, 64 Knoten), ICL DAP (1977)
    CM-2 (Thinking Machines, 1985, einige Tausend Knoten)

4.1 Technologiefortschritte (II)

- 80/90er Jahre: Verbreitung von RISC-CPUs, → MIMD
  gemeinsamer Speicher für alle CPUs
  (bei großer Anzahl problematisch…):
  C.mmp (16 PDP/11-40 Prozessoren)
  RP3 (IBM, 1985, bis 512 CPUs)
  KSR-1 (Kendall Square Research, 1992)

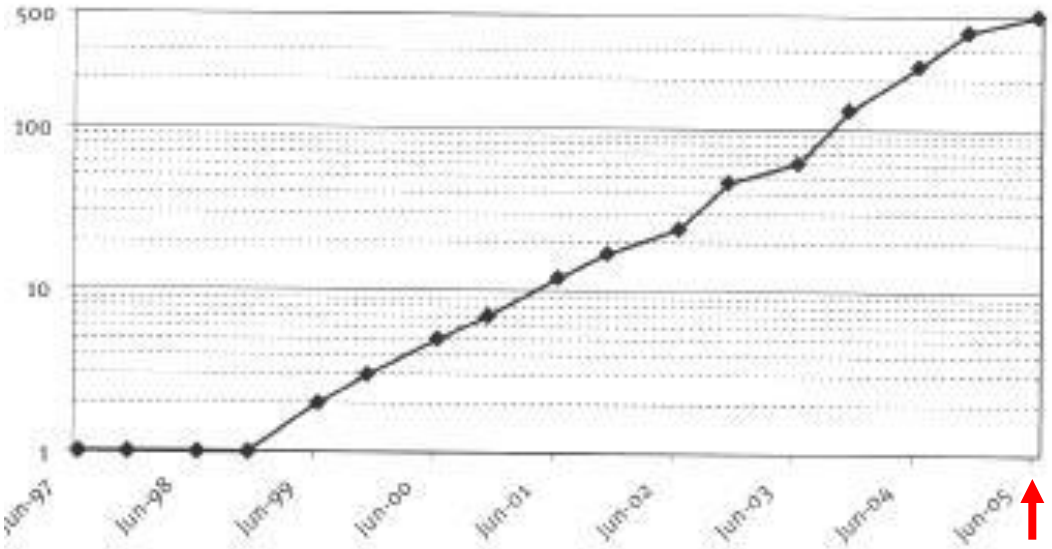
  oder ohne gemeins. Speicher: message passing
  Intel iPSC
  nCUBE Hypercube
  Suprenum (GMD Deutschland)

  Zwischenstellung: CRAY T3D (1994)

- 90/2000er Jahre: viele Cluster aus Standard-CPUs/PCs, GPUs,
  Grid Computing, Cloud,...

4.1 Technologiefortschritte (III)

- Erster „Megaflops-Rechner“: CDC 7600, 1971
  (Vorläufer der Vektorrechner)
- Erster „Gigaflops-Rechner“: Cray-2, 1986
- Erster „Terafops-Rechner“: ASCI Red, 1997
- Erster „Petaflops-Rechner“: IBM Blue Gene/P „Roadrunner“, 2008
- Erster „Exaflops-Rechner“? (akt. Prognose: etwa 2019)



Anzahl von „Teraflops-Rechnern“ in den TOP500

Quelle: Meuer:
The TOP500 Project.
in: Informatik Spektrum
Jg. 31, H. 3/2008, S. 203-222

Seit Juni 2005 sind in den TOP500 nur noch „Teraflops-Rechner“ enthalten!

4.1 Technologiefortschritte (IV)

Einige technologische Einflussfaktoren auf die Entwicklung im Bereich der Parallelverarbeitung sind z.B.:

1.) Leistungsexplosion (und Preisverfall) bei Hardware
- vgl. Moores Gesetz
- „MIPS-Preis“: 1991 Intel 486: 225 $/MIPS
  1997 Intel Pentium II: 4 $/MIPS
  2004 Intel Pentium 4: ca. 5 ct/MIPS
  2007: Intel Core Duo: 1,6 ct/MIPS
- Festplattenspeicher: 1991: 1 MB = 5 $
  1999: 1 MB = 2..5 ct
  2008: 1 MB < 0,1 ct
- Multiprozessoren, Multicore-Prozessoren, GPUs

4.1 Technologiefortschritte (V)

2.) Entwicklung von Netzwerken

- Anfänge ca. 1972, „Ethernet“ erstmals 1976
  - 1983 Ethernet-Standard IEEE 802.3
  - 1995 Fast Ethernet
  - 1998 Gigabit-Ethernet (u.a. für Cluster)
  - 2002 10Gigabit-Ethernet
  - 2010 100Gigabit-Ethernet
  - Ethernet-Technologie ist preiswert, robust und breit verfügbar
- und weitere Hochgeschwindigkeitsnetzwerktechnologien:
  - Myrinet (Glasfaser, Switches, proprietäres GM-Protokoll)
  - Scalable Coherent Interconnect SCI (Punkt-zu-Punkt)
  - InfiniBand (Industriestandard für direkten Hauptspeicherzugriff über Netz)
  - QsNet (für parallel Zugriffe in symmetrischen Multiprozessorsystemen)
  - Intel Omni-Path…
- sowie wachsende Bandbreite (Datenübertragungsrate) und geringe Latenz (Verzögerung durch Datentransport im Netz)

4.1 Technologiefortschritte (VI)

Vergleich von Netzwerktechnologien in Relation zum Bus

Technologie	Typ	Bandbreite in MByte/s	Latenz in µsec
Hauptspeicher	Bus	> 1000	< 0.01
Fast Ethernet	switch	11	70
GBit Ethernet	switched	110	30
Myrinet-2000	switched	248	6,3
SCI	point-to-point	326	2,7
InfiniBand	switched	805	7,5
QsNet, QsNet[11]	switched	340 bzw. 900	4

Quelle: Bengel u.a.: Masterkurs parallele und verteilte Systeme. Vieweg/Teubner, 2008

4.1 Technologiefortschritte (VII)

3.) Funkverbindungen und mobile Geräte

- Wireless Personal Area Networks (WPANs)
  zur Vernetzung kleinerer Geräte z.B. mittels
  - Infrarotverbindung (IrDA)
  - Bluetooth
  - DECT (Telefonie)
  - GSM und UMTS (Mobiltelefonie)
- Wireless Local Area Networks (WLAN)
  mit größerer Sendeleistung/Reichweite

4.1 Technologiefortschritte (VIII)

4.) Internet

- = Infrastruktur bzw. Netzwerktechnologie für alle
  - Client-Server-Netze
  - Cluster-Netze
  - Peer-to-Peer-Netzwerke
  - Grid-Netze
- vor allem WWW als Technologietreiber
  - Web 2.0:
    - Kollaboration/Partizipation,
    - Web-Services
    - Service-orientierte Architekturen (SOA)
  - Web 3.0:
    - Semantic Web, interagierende Software-Agenten *interpretieren* Daten auf Basis von Taxonomien, Ontologien (Beschreibung eines Wissensgebietes)
  - Web 4.0: „Web Intelligence“
  - E-Business, E-Applikationen,…

4.1 Technologiefortschritte (IX)

Zur Beherrschung und zum lfd. Betrieb von Clustern und Grids mit Tausenden von Prozessoren sind folg. Technologien relevant:

- Virtualisierung, Utility Computing (Bereitstellung und verbrauchsabhängige Abrechnung von IT-Leistungen in Form von Services)
- Selbstorganisierende Systeme
  - „Autonomic computing“
    - selbst konfigurierend
    - selbst heilend
    - selbst optimierend
    - selbst schützend
  - „Organic Computing“
    - Technisches System stellt sich dynamisch und selbstanpassend auf seine Umgebung ein, d.h. zusätzlich selbst-erklärend und kontextbewusst

4.2 State of the art (Ia): 49. TOP500 List (06/2017):

Nr	Name	Computer	Total Cores	Acceler.	Rmax [TFlop/s]	Power (kW)	OS Fam.
1	Sunway TaihuLight	Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway	10649600		93014,594	15371	Linux
2	Tianhe-2 (MilkyWay-2)	TH-IVB-FEP Cluster, Intel Xeon E5-2692 12C 2.200GHz, TH Express-2, Intel Xeon Phi	3120000	2736000	33862,7	17808	Linux
3	Piz Daint	Cray XC50, Xeon E5-2690v3 12C 2.6GHz, Aries interconnect , NVIDIA Tesla P100	361760	297920	19590	2271,99	Linux
4	Titan	Cray XK7, Opteron 6274 16C 2.200GHz, Cray Gemini interconnect, NVIDIA K20x	560640	261632	17590	8209	Linux
5	Sequoia	BlueGene/Q, Power BQC 16C 1.60 GHz, Custom	1572864		17173,224	7890	Linux
6	Cori	Cray XC40, Intel Xeon Phi 7250 68C 1.4GHz, Aries interconnect	622336		14014,7	3939	Linux
7	Oakforest-PACS	PRIMERGY CX1640 M1, Intel Xeon Phi 7250 68C 1.4GHz, Intel Omni-Path	556104		13554,6	2718,7	Linux
8		K computer, SPARC64 VIIIfx 2.0GHz, Tofu interconnect	705024		10510	12659,89	Linux
9	Mira	BlueGene/Q, Power BQC 16C 1.60GHz, Custom	786432		8586,612	3945	Linux
10	Trinity	Cray XC40, Xeon E5-2698v3 16C 2.3GHz, Aries interconnect	301056		8100,9	4232,63	Linux

Rmax - Maximal LINPACK performance achieved

Quelle: top500.org

4.2 State of the art (Ib): 49. TOP500 List (06/2017):

Rank	Site	System	Cores	Rmax (TFlop/s)	Rpeak (TFlop/s)	Power (kW)
1	National Supercomputing Center in Wuxi China	**Sunway TaihuLight** - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway NRCPC	10,649,600	93,014.6	125,435.9	15,371
2	National Super Computer Center in Guangzhou 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
3	Swiss National Supercomputing Centre (CSCS) Switzerland	**Piz Daint** - Cray XC50, Xeon E5-2690v3 12C 2.6GHz, Aries interconnect , NVIDIA Tesla P100 Cray Inc.	361,760	19,590.0	25,326.3	2,272
4	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
5	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
6	DOE/SC/LBNL/NERSC United States	**Cori** - Cray XC40, Intel Xeon Phi 7250 68C 1.4GHz, Aries interconnect Cray Inc.	622,336	14,014.7	27,880.7	3,939
7	Joint Center for Advanced High Performance Computing Japan	**Oakforest-PACS** - PRIMERGY CX1640 M1, Intel Xeon Phi 7250 68C 1.4GHz, Intel Omni-Path Fujitsu	556,104	13,554.6	24,913.5	2,719
8	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
9	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
10	DOE/NNSA/LANL/SNL United States	**Trinity** - Cray XC40, Xeon E5-2698v3 16C 2.3GHz, Aries interconnect Cray Inc.	301,056	8,100.9	11,078.9	4,233

Rmax - Maximal LINPACK performance achieved

Quelle: top500.org

4.2 State of the art (IIa)

http://top500.org

Aktuelle TOP500 (06/2017):

- In the latest rankings, the **Sunway TaihuLight**, a system developed by China's National Research (NRCPC) and installed at the National Supercomp. Center in Wuxi, maintains its top position. With a Linpack performance of 93 petaflops, TaihuLight is far and away the most powerful number-cruncher on the planet.
- **Tianhe-2, (Milky Way-2),** a system developed by China's National University of Defense Technology (NUDT) and deployed at the National Supercomp. Center in Guangzho, occupies the no. 2 position with a Linpack mark of 33.9 petaflops. Tianhe-2 was the number one system in the TOP500 list for three consecutive years, until TaihuLight eclipsed it in June 2016.
- The new no. 3 is the upgraded **Piz Daint**, a Cray XC50 system installed at the Swiss National Supercomp. Centre (CSCS). The upgrade was accomplished with additional NVIDIA Tesla P100 GPUs, doubling the Linpack performance of the system's previous mark of 9.8 petaflops in November 2016, which itself was the result of a significant upgrade. Piz Daint's current Linpack result of 19.6 petaflops enabled the system to climb five positions in the rankings.

4.2 State of the art (IIb)

http://top500.org

Aktuelle TOP500 (06/2017):

- As a result of the Piz Daint upgrade, **Titan**, a Cray XK7 system installed at the Department of Energy's (DOE) Oak Ridge National Lab., drops to no. 4 in the rankings. Its Linpack mark of 17.6 petaflops has remained constant since it was installed in 2012.
- **Sequoia** (17.2 petaflops), an IBM BlueGene/Q system installed at the DOE's Lawrence Livermore National Laboratory, at no. 5;
- **Cori** (14.0 petaflops), a Cray XC40 system housed at the National Energy Research Scientific Computing Center (NERSC), at no. 6;
- **Oakforest-PACS** (13.6 petaflops), a Fujitsu PRIMERGY system running at Japan's Joint Center for Advanced High Performance Computing, at no. 7;
- Fujitsu's **K computer** (10.5 petaflops), installed at the RIKEN Advanced Institute for Computational Science (AICS), at no. 8;
- **Mira** (8,6 petaflops), an IBM BlueGene/Q system installed at DOE's Argonne National Laboratory, at no. 9
- **Trinity** (8.1 petaflops), a Cray XC40 system running at Los Alamos National Laboratory, at no. 10.

4.2 State of the art (IIc)

http://.top500.org

Aktuelle TOP500 (06/2017, Forts.): **No.1 from since 06/2016:**

- **Sunway TaihuLight** is currently up and running at the National Supercomputing Center in the city of Wuxi/China.
- will be used for various research and engineering work, in areas such as climate, weather & earth systems modeling, life science research, advanced manufacturing, and data analytics.
- was developed by the National Research Center of Parallel Computer Engineering & Technology (NRCPC),
- Processor: SW26010, a 260-core chip that can crank out just over 3 teraflops.
- Power consumption: 15371 kW

4.2 State of the art (IId)

http://top500.org/blog/lists/2014/11/press-release/

Aktuelle TOP500 (06/2017, Forts.):

- **Aggregate performance** on the TOP500 rose to 749 petaflops, a 32 percent jump from a year ago. The slower growth in list performance is a trend that began in 2013
- Intel continues to be the dominant **supplier** of TOP500 chips. Either Xeon or Xeon Phi processors power 464 of the 500 systems. IBM Power processors are in 21 systems, while AMD Opteron CPUs are present in 6 systems.
- A total of 91 systems are now using **accelerator/coprocessor** technology. The most popular choices are NVIDIA GPUs, which are present in 74 systems, and Xeon Phi coprocessors, which are employed in 17 systems.
- For **system interconnects**, Ethernet and InfiniBand continue to be the most prevalent technologies. Ethernet is present in 207 systems; InfiniBand is present in 178. However, for the top 100 systems, their relative share changes dramatically, with Ethernet installed in just a single system, while InfiniBand is used in 42 of these elite machines. Intel Omni-Path interconnect technology, is now installed in 38 systems.

4.2 State of the art (IIe)

http://top500.org/blog/lists/2014/11/press-release/

Aktuelle TOP500 (06/2017, Forts.):

- In the **system vendor** arena, Hewlett Packard Enterprise (HPE) claims the most TOP500 systems, with 144. These include 25 systems originally installed by SGI, which HPE purchased in 2016. Lenovo is the second most popular vendor, with 88 systems, and Cray is in third place, with 57.
- Cray systems, however, continue to lead in **overall performance**, claiming 21.4 percent of the list's total performance. HPE is well back in second place, with an overall performance share of 16.7 percent. Thanks to its number one Sunway TaihuLight system, NRCPC retains the third spot with 12.5 percent of the total performance.
- **Energy efficiency** on the list continues to rise, as reflected in the latest Green500 results. The top four positions are all occupied by newly installed systems in Japan, with the upgraded Piz Daint supercomputer capturing the number five spot. All of these use NVIDIA's latest P100 GPUs. In fact, the top 13 systems on the latest Green500 are all equipped with the P100 hardware.

4.2 State of the art (IIf)

http://top500.org/blog/lists/2014/11/press-release/

Aktuelle TOP500 (06/2017, Forts.):

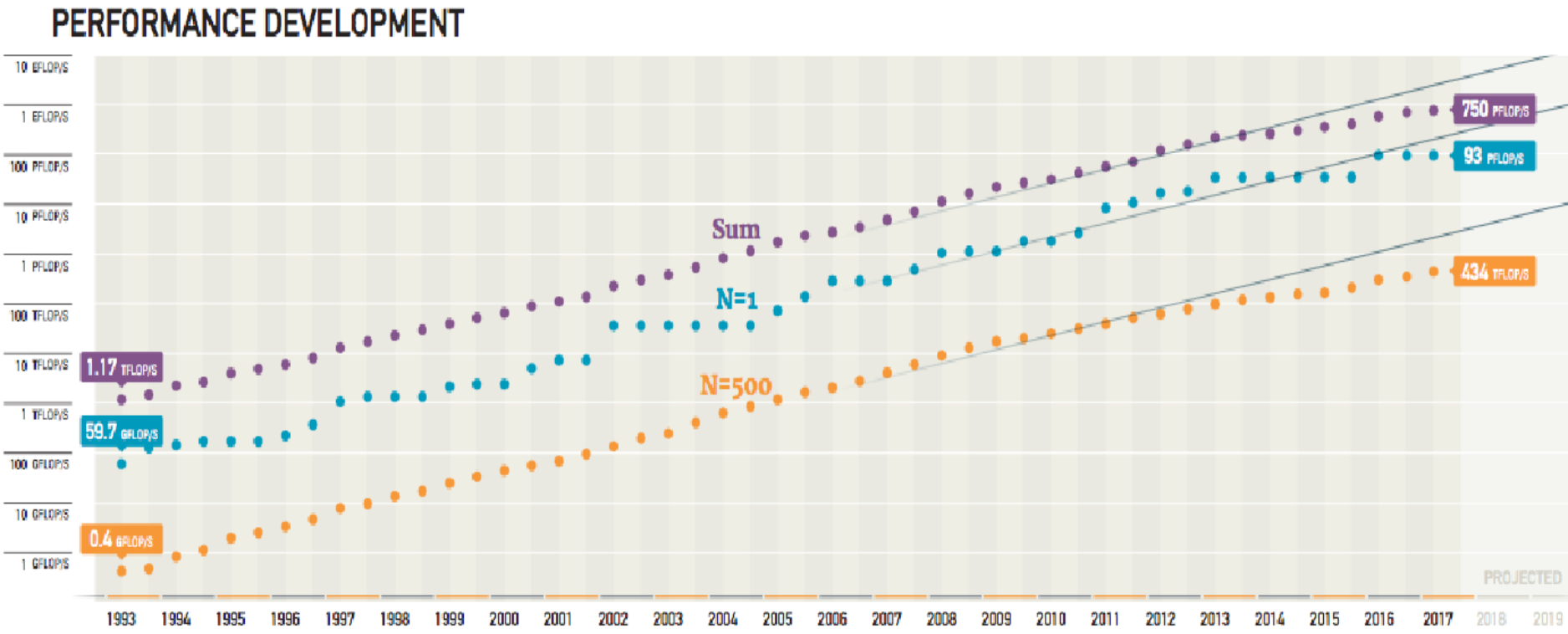
- China and the USA are neck-and-neck in the performance category with the USA holding 33.5% of the overall installed performance while China is second with 31.2% of the overall installed performance.
- There are 138 systems with performance greater than a Pflop/s on the list, up from 117 six months ago.
- In the Top 10, the No. 2 system, Tianhe-2, the No. 6 Cori, and the No. 7 Oakforest-PACS uses Intel Xeon Phi processors to speed up their computational rate. The No. 3 system Piz Daint and the No. 4 system Titan are using NVIDIA GPUs to accelerate computation
- Intel continues to provide the processors for the largest share (92.8 percent) of TOP500 systems.
- Ninety-three (93.0) percent of the systems use processors with eight or more cores, sixty-eight (68.6) percent use twelve or more cores, and twenty-seven (27.2) percent eighteen or more cores.
- In Europe, Germany is the clear leader with 28 systems followed by France with 18 and the UK with 17 systems.

4.2 State of the art (IIIa)

Entwicklung der Performance

Quelle: www.top500.org

Aktuelle TOP500 (06/2017):

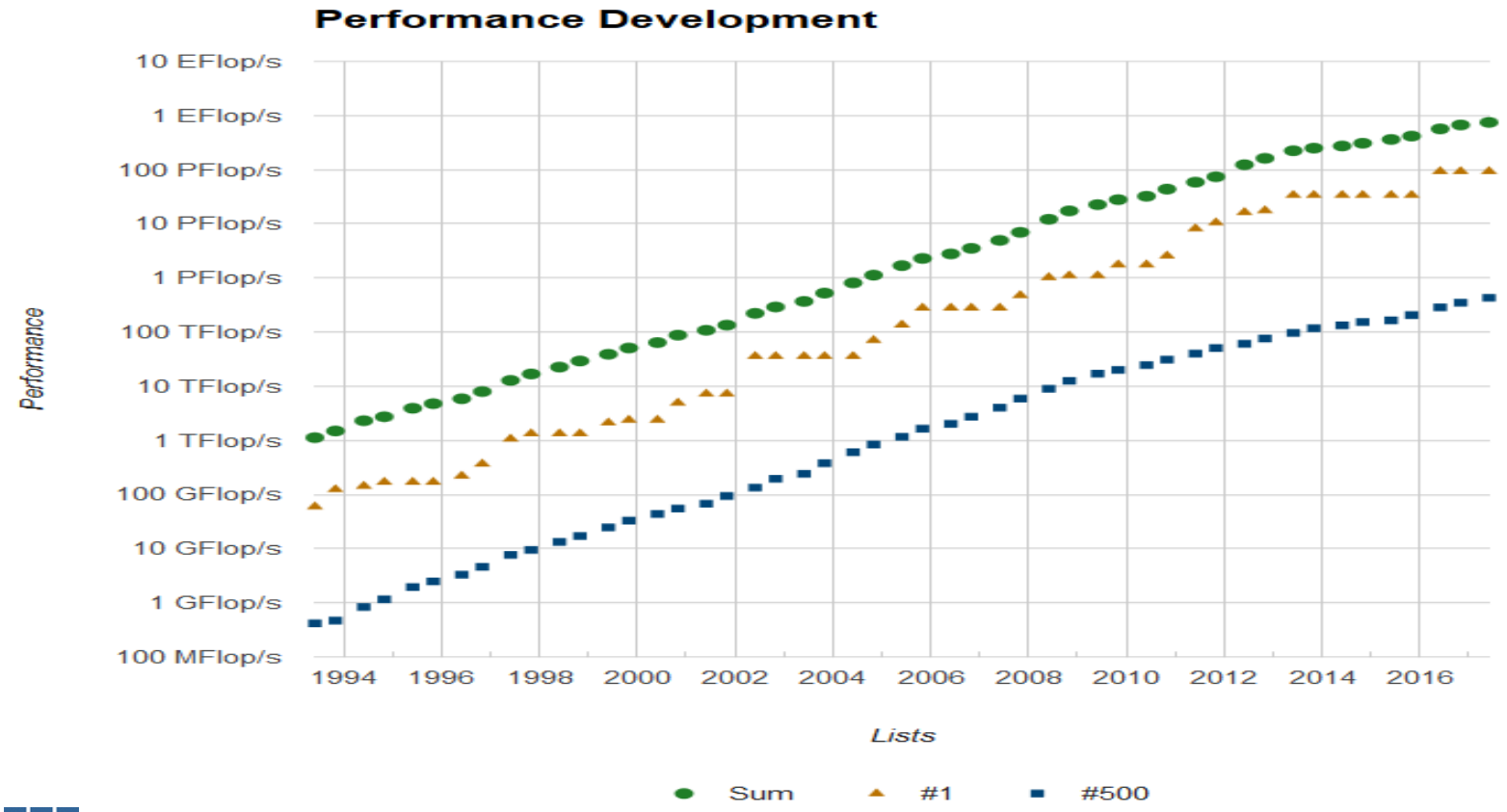


4.2 State of the art (IIIb)

Entwicklung der Performance

Quelle: www.top500.org

Aktuelle TOP500 (06/2017):



4.2 State of the art (IV)

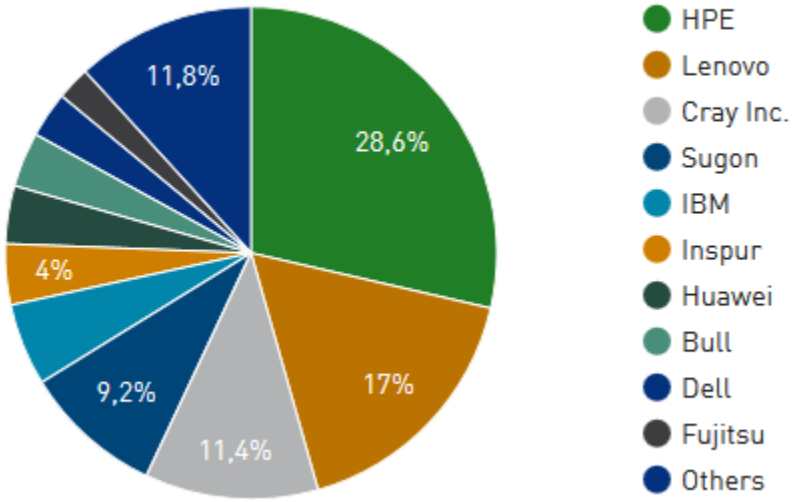
- Parallele Architekturen (und deren Programmierung) werden bald keine Nische mehr sein, sondern Alltag
- SIMD und Vektor-Prinzipien sind in akt. CPUs schon „drin“
- starker Trend zu Integration von Acceleratoren (GPUs!)
- HPC-Systeme sind inzwischen fast „schlüsselfertig“
  → können in fast allen Bereichen eingesetzt werden, nicht nur in Forschung/Lehre, sondern auch breiter Industrieeinsatz!

- in TOP500-06/2017
  - vorherrschende Systemarchitektur: Cluster (86,4%)
  - vorherrschende Prozessorfamilien: Intel (ca. 90%)
  - vorherrschende Verbindungstechnologie: 10G Ethernet + (39%) und Infinniband (35%)
  - vorherrschendes Betriebssystem: Linux (99,6%) !

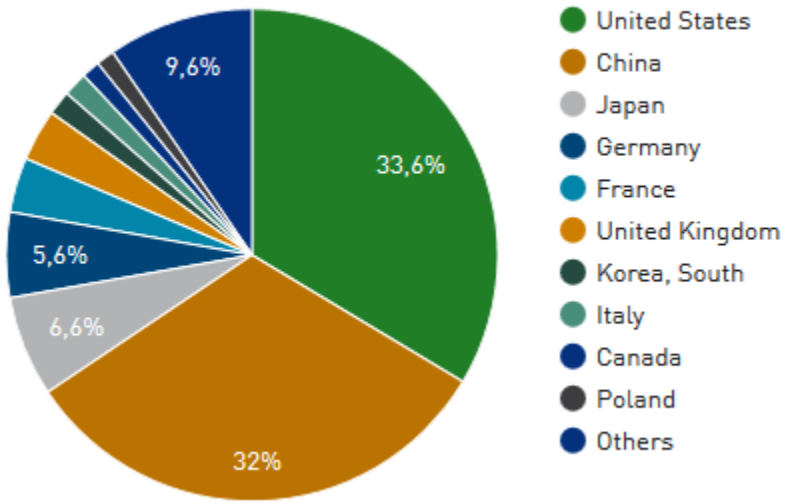
4.2 State of the art (Va)

Alle Angaben: TOP500-06/2017

Vendors System Share

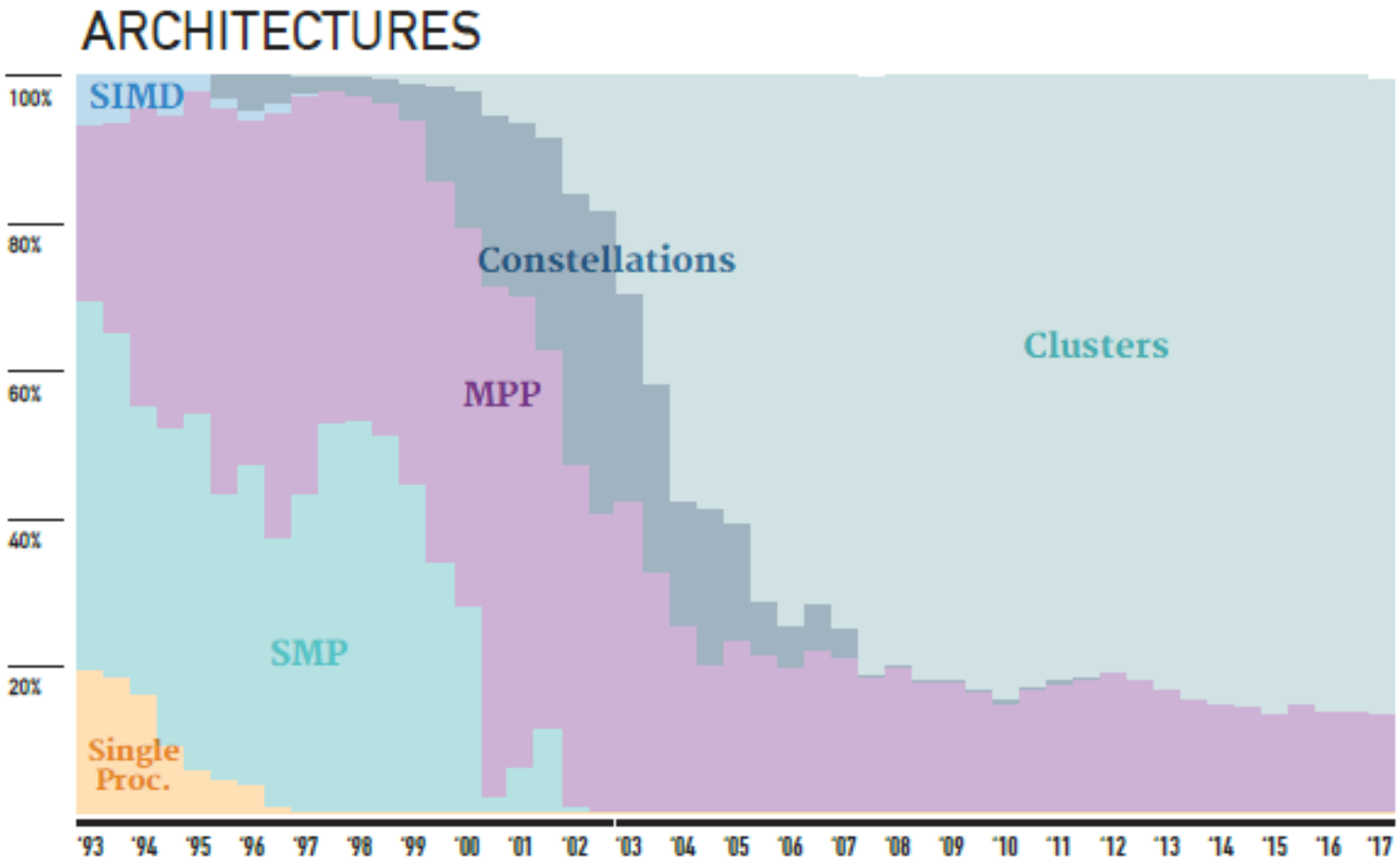


Countries System Share



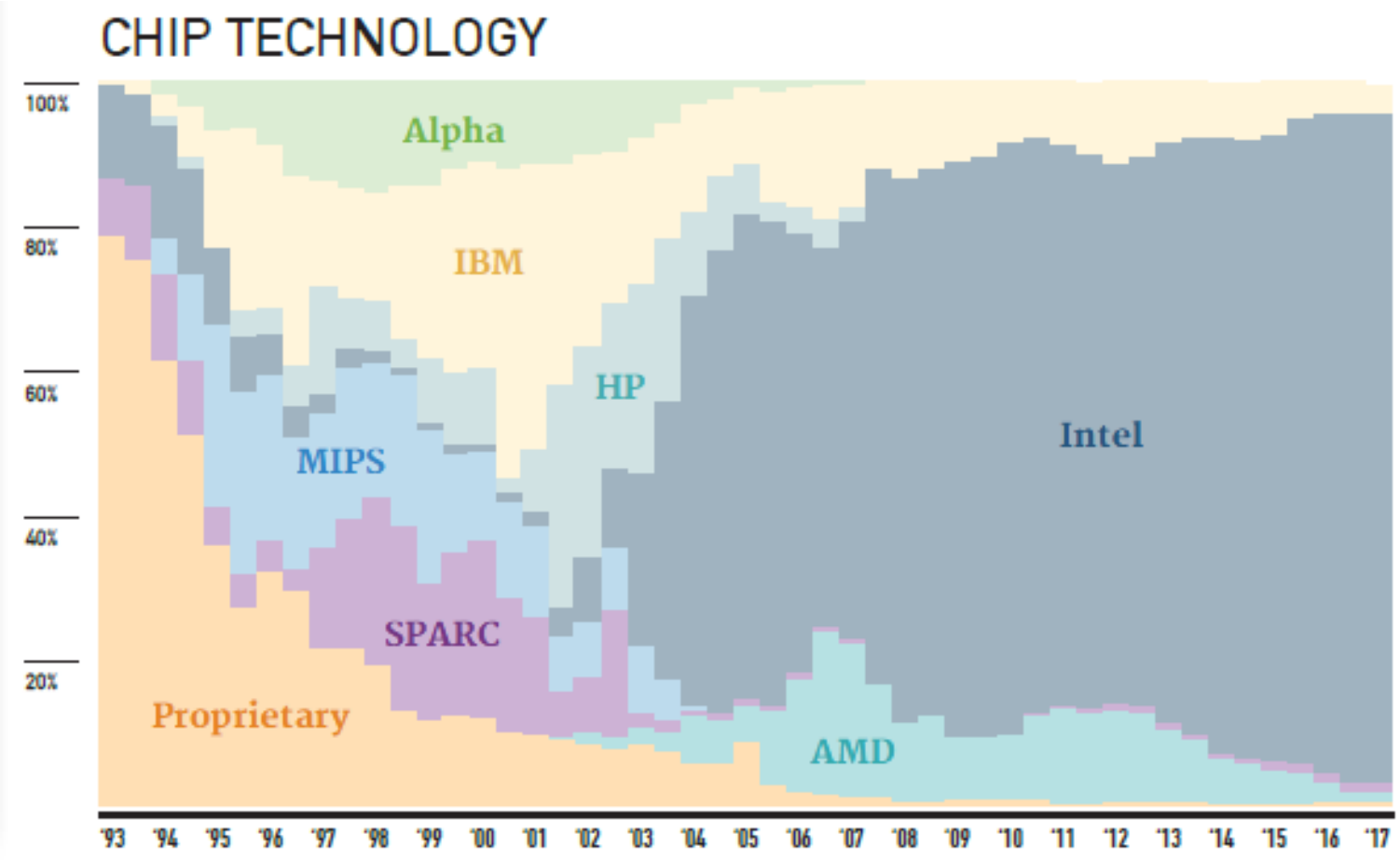
4.2 State of the art (Vb)

Alle Angaben: TOP500-06/2017



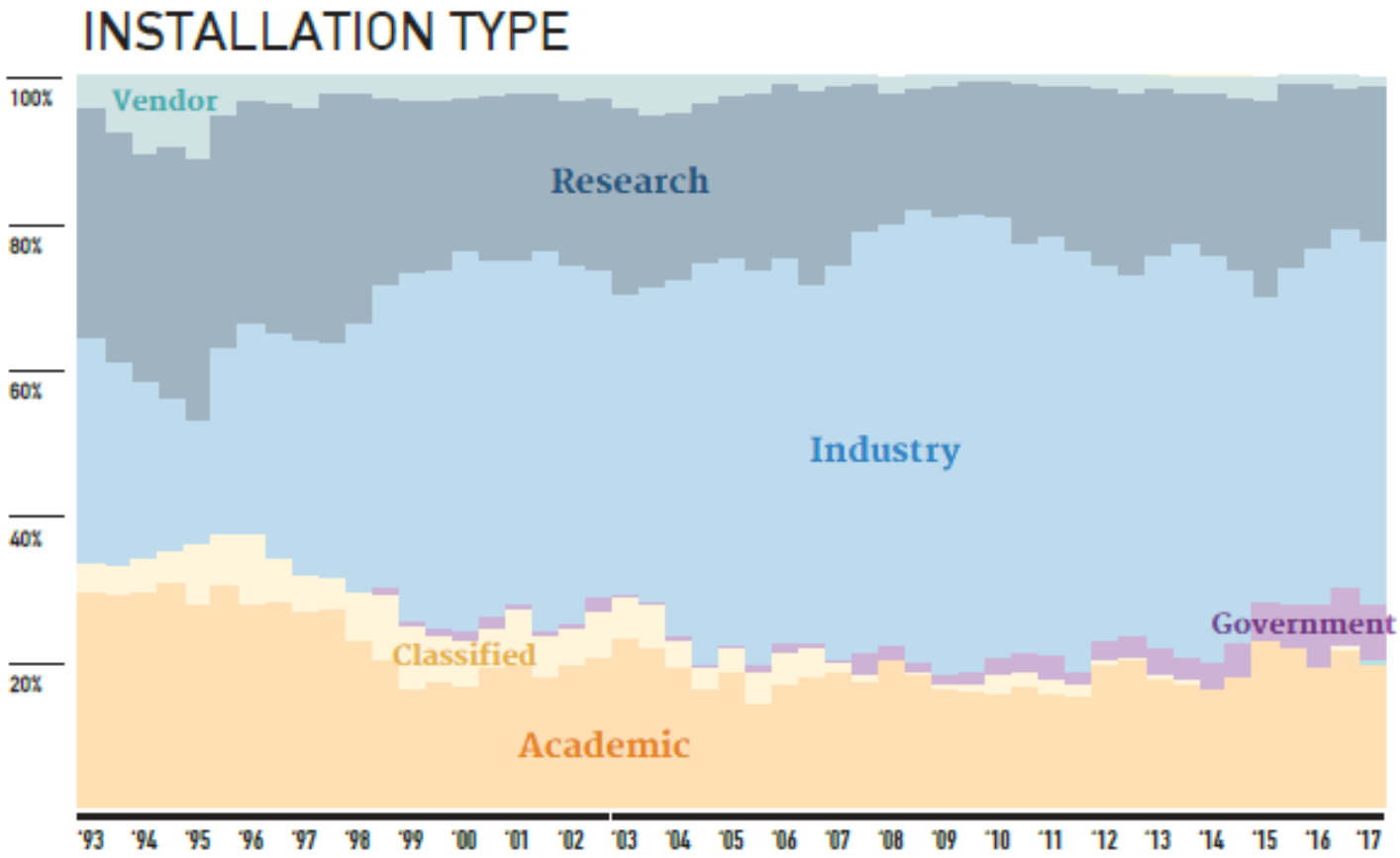
4.2 State of the art (Vc)

Alle Angaben: TOP500-06/2017



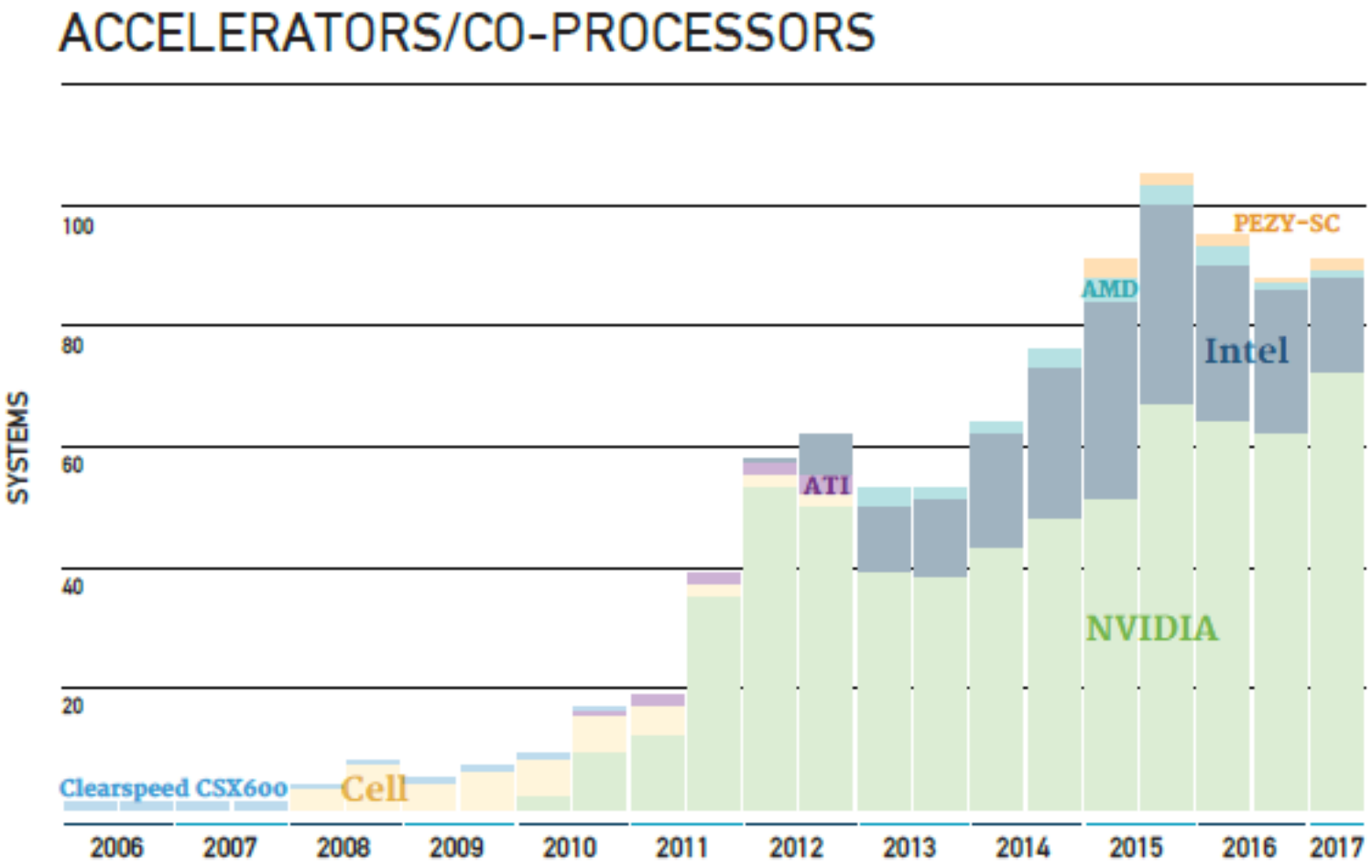
4.2 State of the art (Vd)

Alle Angaben: TOP500-06/2017



4.2 State of the art (Ve)

Alle Angaben: TOP500-06/2017



4.2 State of the art (VI)

- **Bell's Law** (1972):

Technologiefortschritte bei Halbleitern, Speichern, Benutzerinterfaces und Netzwerktechnik führen etwa alle 10 Jahre zu einer neuen, preisgünstigen „Computing-Plattform".

Eine solche Plattform wird ca. 10 Jahre lang als Basis für eine industrielle Infrastruktur fungieren.

Für solche Computing-Plattformen gelten etwa 10-Jahreszyklen für:

- Forschung
- Einführung und Reife
- Hauptanwendung („prime use")
- Ausklingen der Hauptanwendung („past prime usage")

4.2 State of the art (VII)

- **Bell's Law** (1972): Anwendung auf High Performance Computing:

HPC Computer Classes

Class	Early Adoption starts:	Prime Use starts:	Past Prime Usage starts:
Data Parallel Systems	Mid 70's	Mid 80's	Mid 90's
Custom Scalar Systems	Mid 80's	Mid 90's	Mid 2000's
Commodity Cluster	Mid 90's	Mid 2000's	Mid 2010's ???
Accelerators or Embedded Proc	Mid 2000's	Mid 2010's ?!	Mid 2020's ???

Quelle: top500.org (2015)

Vectorrechner, SIMD

Massiv parallele Syst. Scalar SMPs

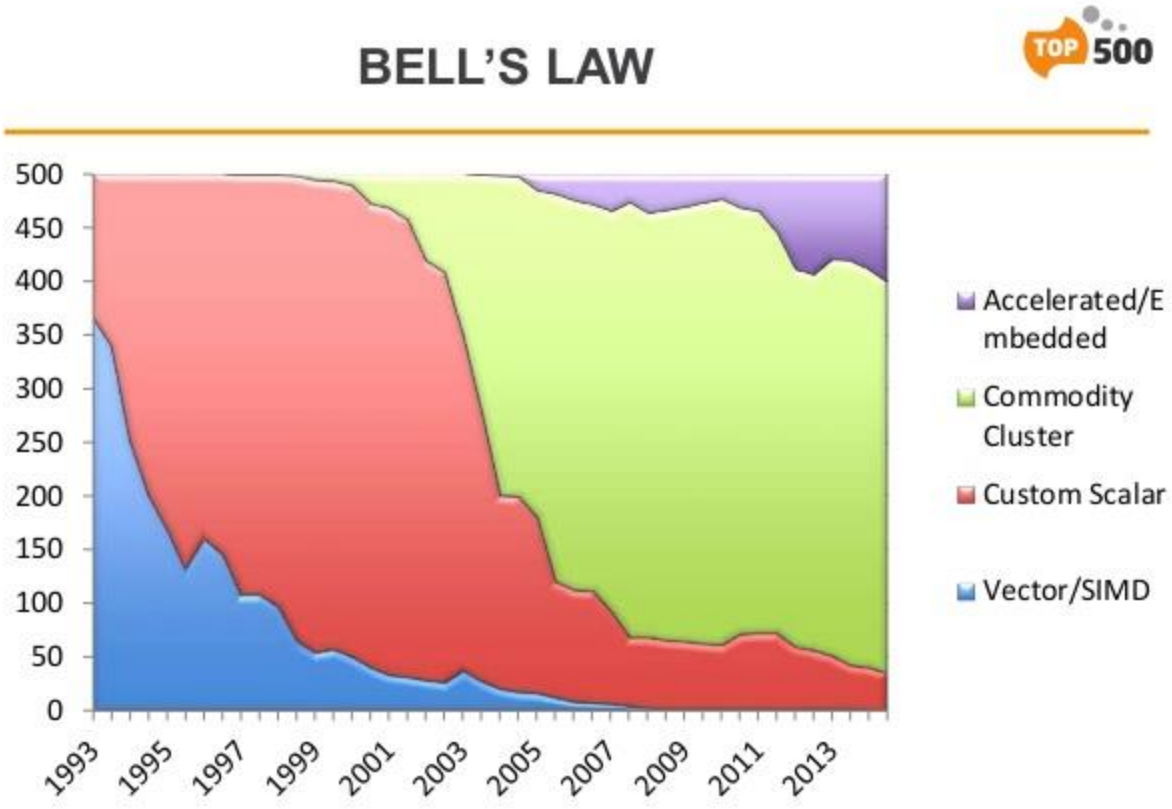
Commodity clusters: PC-Cluster, Blades…

Acceleratoren, energieeffiziente Syst. wie BlueGene?

… scheint sich zu bestätigen …

4.2 State of the art (VIII)

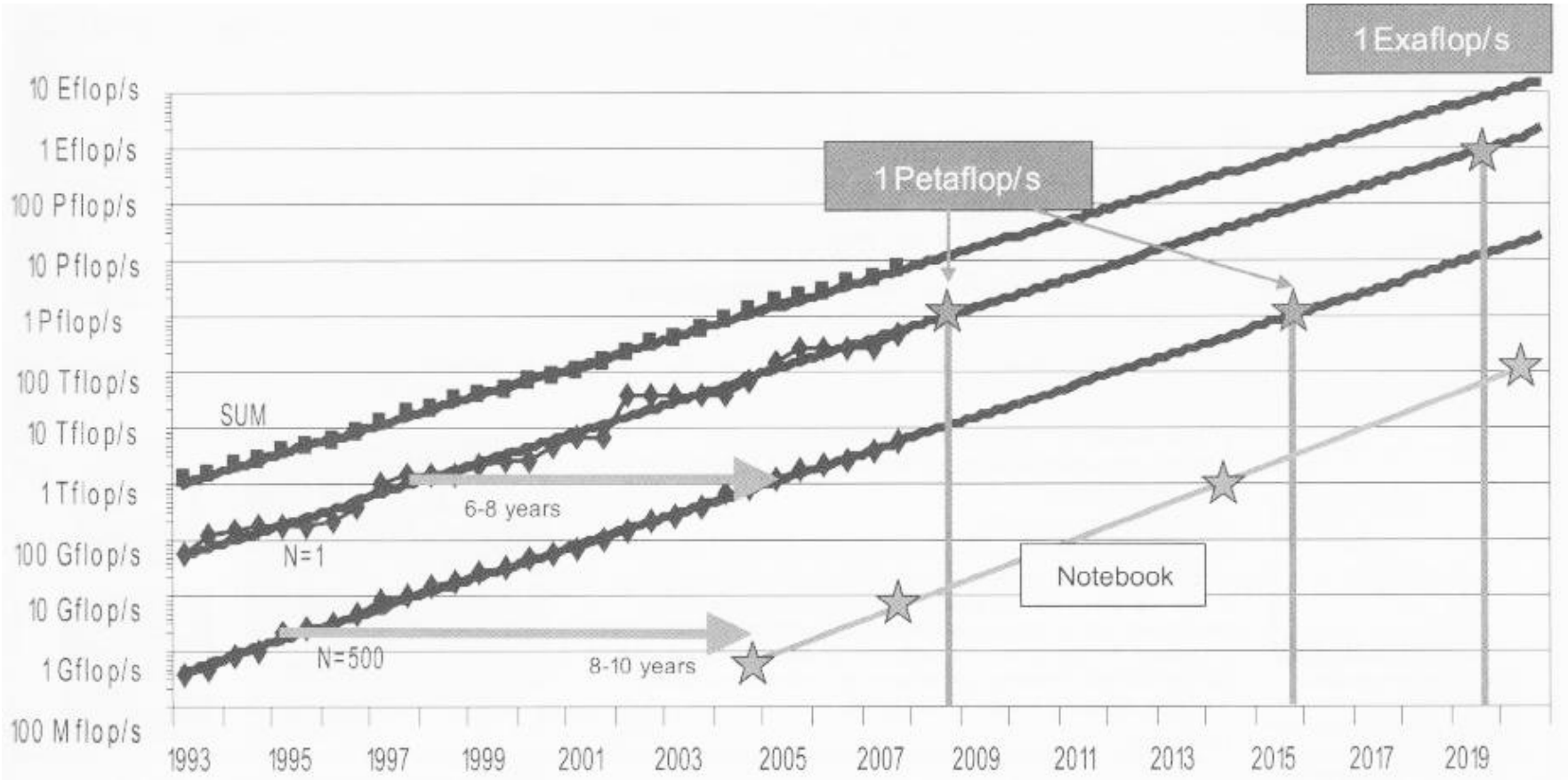
- „Beweis“ (aus TOP500-Historie):



Quelle: www.top500.org (2015)

4.2 State of the art (IX)

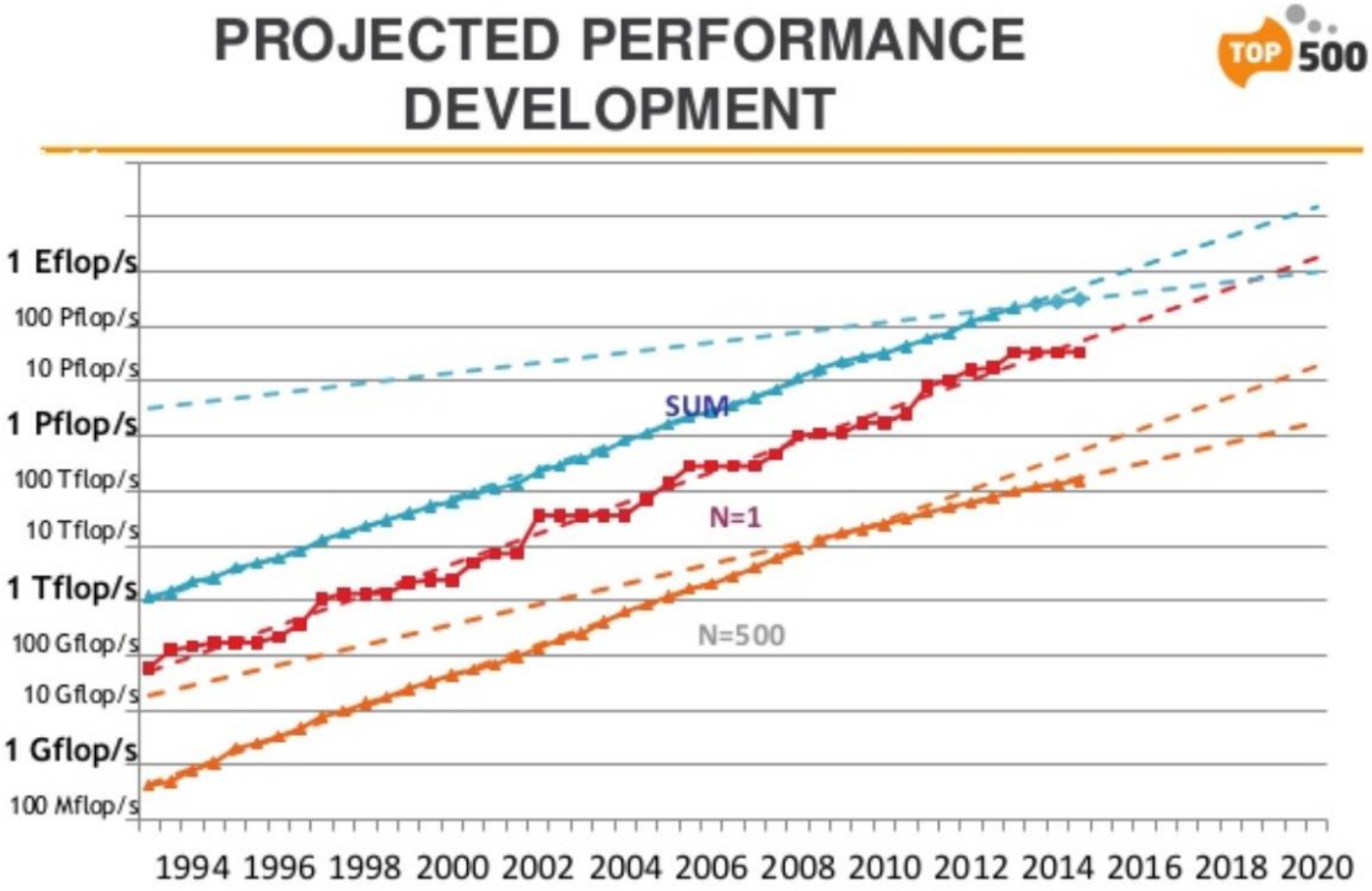
„Prognose“ (2008, auf Basis der TOP500-Historie):



Quelle: Meuer: The TOP500 Project. in: Informatik Spektrum Jg. 31, H. 3/2008, S. 203-222

4.2 State of the art (X)

Aktuelle „Prognose“ (auf Basis der TOP500-Historie):



Quelle: www.top500.org (2015)

4.2 State of the art (XIIa)

Trend „Green IT“ auch (erst recht?) im HPC-Bereich!

⇒ Energie-Effizienz aktueller HPC-Systeme
⇒ Energieeffizienz = MFLOPS / W_{att}

RankingListe: www.green500.org

aktuell: Liste 06/2017

4.2 State of the art (XIIb)

Energieeffizienz = MFLOPS / Watt

Nr	TOP500 Rank	Name	Computer	Site	Country	Power kW	Power Effeciency GFlops/Watts
1	61	TSUBAME3.0	SGI ICE XA, IP139-SXM2, Xeon E5-2680v4 14C 2.4GHz, Intel Omni-Path, NVIDIA Tesla P100 SXM2	GSIC Center,	Japan	141,6	14,11
2	465	kukai	ZettaScaler-1.6 GPGPU system, Xeon E5-2650Lv4 14C 1.7GHz, Infiniband FDR, NVIDIA Tesla P100	Yahoo Japan	Japan	32,8	14,046
3	148	AIST AI Cloud	NEC 4U-8GPU Server, Xeon E5-2630Lv4 10C 1.8GHz, Infiniband EDR, NVIDIA Tesla P100 SXM2	National Instit	Japan	75,78	12,681
4	305	RAIDEN GPU sub	NVIDIA DGX-1, Xeon E5-2698v4 20C 2.2GHz, Infiniband EDR, NVIDIA Tesla P100	Center for Ad	Japan	59,9	10,603
5	100	Wilkes-2	Dell C4130, Xeon E5-2650v4 12C 2.2GHz, Infiniband EDR, NVIDIA Tesla P100	University of (	United Kingdo	114,4	10,428
6	3	Piz Daint	Cray XC50, Xeon E5-2690v3 12C 2.6GHz, Aries interconnect , NVIDIA Tesla P100	Swiss Nation	Switzerland	2272	10,398
7	69	Gyoukou	ZettaScaler-2.0 HPC system, Xeon D-1571 16C 1.3GHz, Infiniband EDR, PEZY-SC2	Japan Agency	Japan	164	10,226
8	220	Research Comput	SGI Rackable C1104-GP1, Xeon E5-2650v4 12C 2.2GHz, Infiniband EDR, NVIDIA Tesla P100	National Instit	Japan	78,64	9,797
9	31		NVIDIA DGX-1/Penguin Relion 2904GT, Xeon E5-2698v4 20C 2.2GHz/ E5-2650v4, Mellanox Infiniband	Facebook	United States	349,5	9,462
10	32	DGX Saturn V	NVIDIA DGX-1, Xeon E5-2698v4 20C 2.2GHz, Infiniband EDR, NVIDIA Tesla P100	NVIDIA Corpo	United States	349,5	9,462

Quelle: www.green500.org (Juni2017)

4.2 State of the art (XIIc)

Quelle: www.green500.org (Juni2017)

- Energy efficiency on the list continues to rise
- The top four positions are all occupied by newly installed systems in Japan, with the upgraded Piz Daint supercomputer capturing the number five spot. All of these use NVIDIA's latest P100 GPUs. In fact, the top 13 systems on the latest Green500 are all equipped with the P100 hardware.
- The most energy-efficient system on the Green500 list is the new TSUBAME 3.0, a modified HPE ICE XA system installed at the Tokyo Institute of Technology. It achieved 14.110 gigaflops/watt during its 1.998-petaflop Linpack performance run. It is ranked number 61 on the TOP500.
- The no. 2 Green500 entry is the kukai system at the Yahoo Japan Corporation. Built by Exascaler, this system achieves 14.045 gigaflops/watt, just 0.3 percent behind TSUBAME 3.0.
- Piz Daint, the fifth-ranked supercomputer on the Green500, conducted a power-optimized run of the Linpack benchmark, achieving 10.4 gigaflops/watt. At number three on the TOP500, it represents the most energy-efficient supercomputer in the top 50 of that list.

4.3 aktuelle Probleme (I)

- Hardwareentwicklung (vor allem CPUs) gibt die „Geschwindigkeit“ vor – die Software hinkt hinter her …

  „It appears to be easier to build parallel machines then to use them.“ [Papadopoulus, G. 1987]

- geeignete Aufteilung der Aufgaben/Problem ist schwierig
- z.T. hoher Overhead für Koordinierung der Teilaufgaben/-abläufe
- wachsende Zahl von Systemen mit sehr vielen (> 1000) CPUs aber Leistungsfähigkeit der Tools oft begrenzt:
  - wie visualisiert man den Zustand von 1000 CPUs?
  - wie modelliert man das Verhalten von 1000 CPUs?
  - wie optimiert man den Quellcode? Wie testet man? usw.
- Komplexe HPC-Projekte basieren oft auf verschiedenen Modellen (vgl. Wettervorhersage: versch. Modelle für Wasser, Atmosphäre, Erde…)
  → Softwareentwicklung besonders anspruchsvoll!

4.3 aktuelle Probleme (II)

- es entstehen z.T. inhomogene Systeme
  (Roadrunner: Power-CPUs + Cell-CPUs mit unterschiedl. Taktraten!)
  → Portabilität schwierig
  Alternativen evtl. bei Multi-Core und Many-Core-Architekturen
- Speicher und Netzwerk sind der Engpass:
  - Verdopplung CPU-Leistung ca. alle 1,5-2 Jahre (Moore´s Law!)
  - Verdopplung DRAM-Geschwindigkeit ca. alle 6 Jahre

  → speichereffiziente Algorithmen nötig!
  (möglichst hoher Anteil lokaler Speicherzugriffe erforderlich)
- profitiert der Mainstream (die gängigen Anwendungen) von den Fortschritten der HW-Entwicklung (und dem „Wettlauf“ für TOP500)?
  Vgl. auch Kritik am TOP500-Benchmarking (vordergründige Orientierung an Peak-Performance und CPU-Zahl, kaum Beachtung des Einflusses der Netzwerkbelastung, zeigt nicht die Anwendungsleistung, was sagt eine einzige Zahl aus?)

4.3 aktuelle Probleme (III)

- z.Zt. noch Probleme mit modernen Multicore-Architekturen:
  - insbesondere Programmierung/Programmier-Tools von Multicore-CPUs noch nicht ausgereift,
  - die Anwendungseigenschaften entscheiden stark über den „Gewinn“ (Parallelisierbarkeit, insbes. auf Thread-Level, Lokalitätsverhalten, Datenabhängigkeiten, …),
  - es gibt viele alte Anwendungen, die von Multicores nicht profitieren
  - technische Komplexität der Kerne selbst ist (noch zu) hoch
  - BS-Unterstützung für Multicores noch ausbaufähig (z.B. bzgl. nutzerdefinierter Zuordnung von Threads zu Cores)
  - OpenMP wird aber zunehmend von Compilern unterstützt

- Zuverlässigkeit komplexer Systeme!?
  - hohe Zahl von Knoten → Ausfallwahrscheinlichkeit?
  - genügend Redundanz?
  - Multicore-Prozessoren könnten dieses Problem etwas mildern