

DIE TOP500-LISTE

- ▶ Vergleich von Rechnersystemen
- ▶ Die TOP500-Liste
- ▶ Details vom November 2009
- ▶ Historische Entwicklung wichtiger Aspekte
- ▶ Leistungsentwicklung
- ▶ Ausgewählte Systeme
- ▶ Ein Blick zurück

Die TOP500-Liste

Die zehn wichtigsten Fragen

- ▶ In welcher Einheit wird die Leistung angegeben?
- ▶ Wie wird die Leistung evaluiert?
- ▶ Welche Zielsetzung verfolgt das TOP500-Projekt?
- ▶ In welchen Größenordnungen liegen die stärksten Rechner?
- ▶ Welche Hersteller dominieren wie den Markt?
- ▶ Welches sind aktuelle Anwendergebiete?
- ▶ Welche Prozessorenarchitekturen und –familien dominieren?
- ▶ Welche Verbindungstechnologien dominieren in welchem Bereich?
- ▶ Wie verhält sich die Leistungssteigerung zu Moore's Law?
- ▶ Welche Leistung brachten Systeme 1993?

▶ 47

Hochleistungsrechnen - © Thomas Ludwig

15.04.2010

Vergleich von Rechnersystemen

Komplexe Fragestellung

- ▶ Erster Ansatz: FLOPS (floating point operations per second)
- ▶ Theoretisches Maximum ergibt sich aus der Anzahl der Zyklen pro Gleitkommaoperation

Bewertung durch sogenannte Benchmark-Programme

- ▶ Synthetische Benchmarks (meist Assembler)
- ▶ CPU-Benchmark (meist numerische Programme)

▶ 48

Hochleistungsrechnen - © Thomas Ludwig

15.04.2010

Siehe: <http://en.wikipedia.org/wiki/Flops>,
http://en.wikipedia.org/wiki/Benchmark_%28computing%29

Vergleich von Rechnersystemen...

Der parallele LINPACK-Benchmark

- ▶ Entwickelt von Jack Dongarra (Knoxville, TN)
- ▶ Ist gleichzeitig eine vollwertige Bibliothek für lineare Algebra
- ▶ Benchmark: dicht besetztes Gleichungssystem
- ▶ R_{\max} ist maximale Leistung bei Problemgröße N_{\max}
- ▶ $N_{1/2}$ ist Problemgröße bei Leistung $R_{1/2}$
- ▶ R_{peak} ist die theoretische Maximalleistung

► 49

Hochleistungsrechnen - © Thomas Ludwig

15.04.2010

Siehe: <http://en.wikipedia.org/wiki/LINPACK>

Die TOP500-Liste

Website www.top500.org

- ▶ Hans Meuer (Universität Mannheim)
- ▶ Jack Dongarra (Univ. Tennessee, Knoxville)
- ▶ Erich Strohmeier (NERSC/LBNL)
- ▶ Horst Simon (NERSC/LBNL)

Zwei Aktualisierungen pro Jahr

- ▶ Juni: International Supercomputing Conference Deutschland
- ▶ November: Supercomputing Conference USA

Basiert auf dem LINPACK-Benchmark

► 50

Hochleistungsrechnen - © Thomas Ludwig

15.04.2010

Siehe: <http://www.top500.org/>

Rank	Site	Computer/Year Vendor	Cores	R _{max}	R _{peak}	Power
1	Oak Ridge National Laboratory United States	Jaguar - Cray XT5-HE Opteron Six Core 2.6 GHz / 2009 Cray Inc.	224162	1759.00	2331.00	6950.60
2	DOE/NNSA/LANL United States	Roadrunner - BladeCenter QS22/LS21 Cluster, PowerXCell 8i 3.2 Ghz / Opteron DC 1.8 GHz, Voltaire Infiniband / 2009 IBM	122400	1042.00	1375.78	2345.50
3	National Institute for Computational Sciences/University of Tennessee United States	Kraken XT5 - Cray XT5-HE Opteron Six Core 2.6 GHz / 2009 Cray Inc.	98928	831.70	1028.85	
4	Forschungszentrum Juelich (FZJ) Germany	JUGENE - Blue Gene/P Solution / 2009 IBM	294912	825.50	1002.70	2268.00
5	National SuperComputer Center in Tianjin/NUDT China	Tianhe-1 - NUDT TH-1 Cluster, Xeon E5540/E5450, ATI Radeon HD 4870 2, Infiniband / 2009 NUDT	71680	563.10	1206.19	
6	NASA/Ames Research Center/NAS United States	Pleiades - SGI Altix ICE 8200EX, Xeon QC 3.0 GHz/Nehalem EP 2.93 Ghz / 2009 SGI	56320	544.30	673.26	2348.00
7	DOE/NNSA/LLNL United States	BlueGene/L - eServer Blue Gene Solution / 2007 IBM	212992	478.20	596.38	2329.60
8	Argonne National Laboratory United States	Blue Gene/P Solution / 2007 IBM	163840	458.61	557.06	1260.00

TOP500 Nov 2009 1-8 / Ein paar wichtige Daten sollten Sie kennen.

Nr 1: Mit über 220.000 Prozessorkernen verwendet Jaguar einen 6-Kern-AMD-Prozessoren. Braucht auch entsprechend viel Energie.

Nr 2: Roadrunner verwendet zwei Prozessortypen: Opteron und Cell-Prozessor

Nr 4: Das in Jülich stehende System JUGENE hat fast 300.000 Prozessoren vom Typ PowerPC 450 850 MHz.

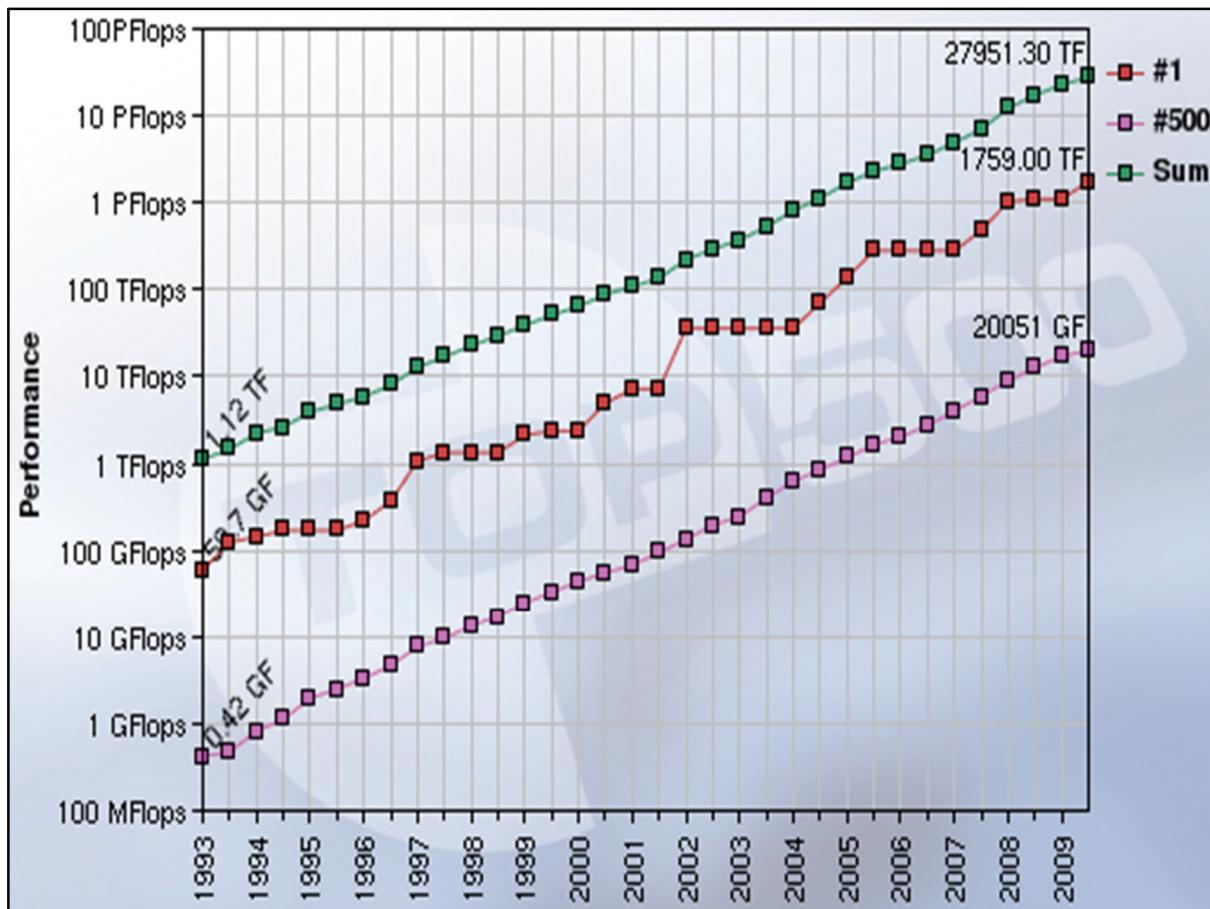
9	Texas Advanced Computing Center/Univ. of Texas United States	Ranger - SunBlade x6420, Opteron QC 2.3 Ghz, Infiniband / 2008 Sun Microsystems	62976	433.20	579.38	2000.00
10	Sandia National Laboratories / National Renewable Energy Laboratory United States	Red Sky - Sun Blade x6275, Xeon X55xx 2.93 Ghz, Infiniband / 2009 Sun Microsystems	41616	423.90	487.74	
11	DOE/NNSA/LLNL United States	Dawn - Blue Gene/P Solution / 2009 IBM	147456	415.70	501.35	1134.00
12	Moscow State University - Research Computing Center Russia	Lomonosov - T-Platforms T-Blade2, Xeon 5570 2.93 GHz, Infiniband QDR / 2009 T-Platforms	35360	350.10	414.42	
13	Forschungszentrum Juelich (FZJ) Germany	JUROPA - Sun Constellation, NovaScale R422-E2, Intel Xeon X5570, 2.93 GHz, Sun M9/Mellanox QDR Infiniband/Partec Parastation / 2009 Bull SA	26304	274.80	308.28	1549.00
14	KISTI Supercomputing Center Korea, South	TachyonII - Sun Blade x6048, X6275, IB QDR M9 switch, Sun HPC stack Linux edition / 2009 Sun Microsystems	26232	274.80	307.44	1275.96
15	NERSC/LBNL United States	Franklin - Cray XT4 QuadCore 2.3 GHz / 2008 Cray Inc.	38642	266.30	355.51	1150.00
16	Oak Ridge National Laboratory United States	Jaguar - Cray XT4 QuadCore 2.1 GHz / 2008 Cray Inc.	30976	205.00	260.20	1580.71

TOP500 Nov 2009 9-16 / Ein paar wichtige Daten sollten Sie kennen.

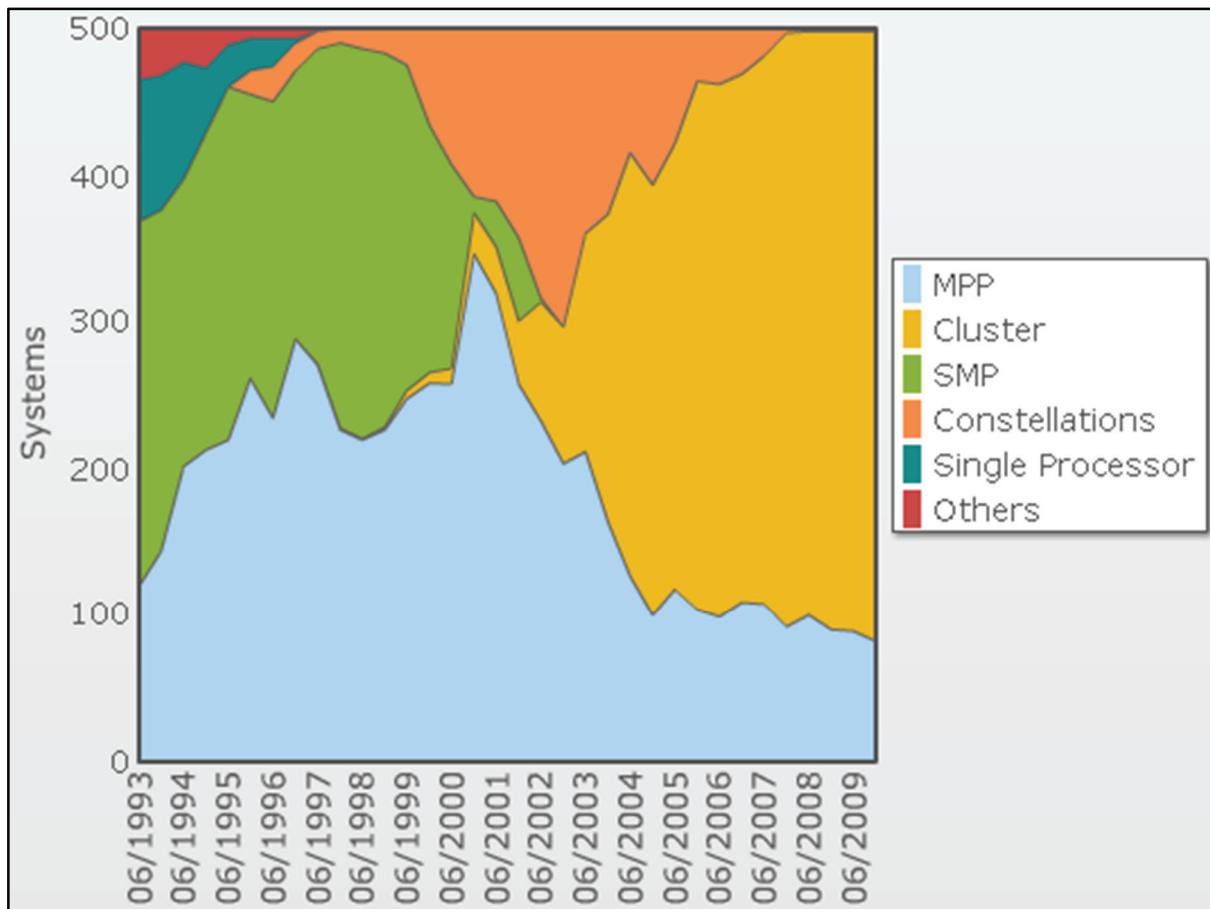
Rank	Site	System	Cores	R _{max}	R _{peak}
4	Forschungszentrum Juelich (FZJ) Germany	Blue Gene/P Solution IBM	294912	825.5	1002.7
13	Forschungszentrum Juelich (FZJ) Germany	Sun Constellation, NovaScale R422-E2, Intel Xeon X5570, 2.93 GHz, Sun M9/Mellanox QDR Infiniband/Partec Parastation Bull SA	26304	274.8	308.28
35	DKRZ - Deutsches Klimarechenzentrum Germany	Power 575, p6 4.7 GHz, Infiniband IBM	8064	115.9	151.6
39	HLRN at Universitaet Hannover / RRZN Germany	SGI Altix ICE 8200EX, Xeon QC E5472 3.0 GHz/X5570 2.93 GHz SGI	10240	107.1	120.73
40	HLRN at ZIB/Konrad Zuse-Zentrum fuer Informationstechnik Germany	SGI Altix ICE 8200EX, Xeon QC E5472 3.0 GHz/X5570 2.93 GHz SGI	10240	107.1	120.73
46	Max-Planck-Gesellschaft MPI/IPP Germany	Power 575, p6 4.7 GHz, Infiniband IBM	6848	98.42	128.74
50	IT Service Provider Germany	Cluster Platform 3000 BL2x220, E54xx 3.0 Ghz, Infiniband Hewlett-Packard	10240	94.74	122.88
82	Leibniz Rechenzentrum Germany	Altix 4700 1.6 GHz SGI	9728	56.52	62.26
94	HWW/Universitaet Stuttgart Germany	NEC HPC 140Rb-1 Cluster, Xeon X5560 2.8Ghz, Infiniband NEC	5376	50.79	60.21
102	Max-Planck-Gesellschaft MPI/IPP Germany	Blue Gene/P Solution IBM	16384	47.73	55.71
110	Forschungszentrum Juelich (FZJ) Germany	QPACE SFB TR Cluster, PowerXCell 8i, 3.2 GHz, 3D-Torus IBM	4608	43.01	55.71

TOP500 Nov 2009 Deutschland: 27 Systeme auf den Plätzen von 4 bis 481.

Das DKRZ steht auf Platz 3 in Deutschland.

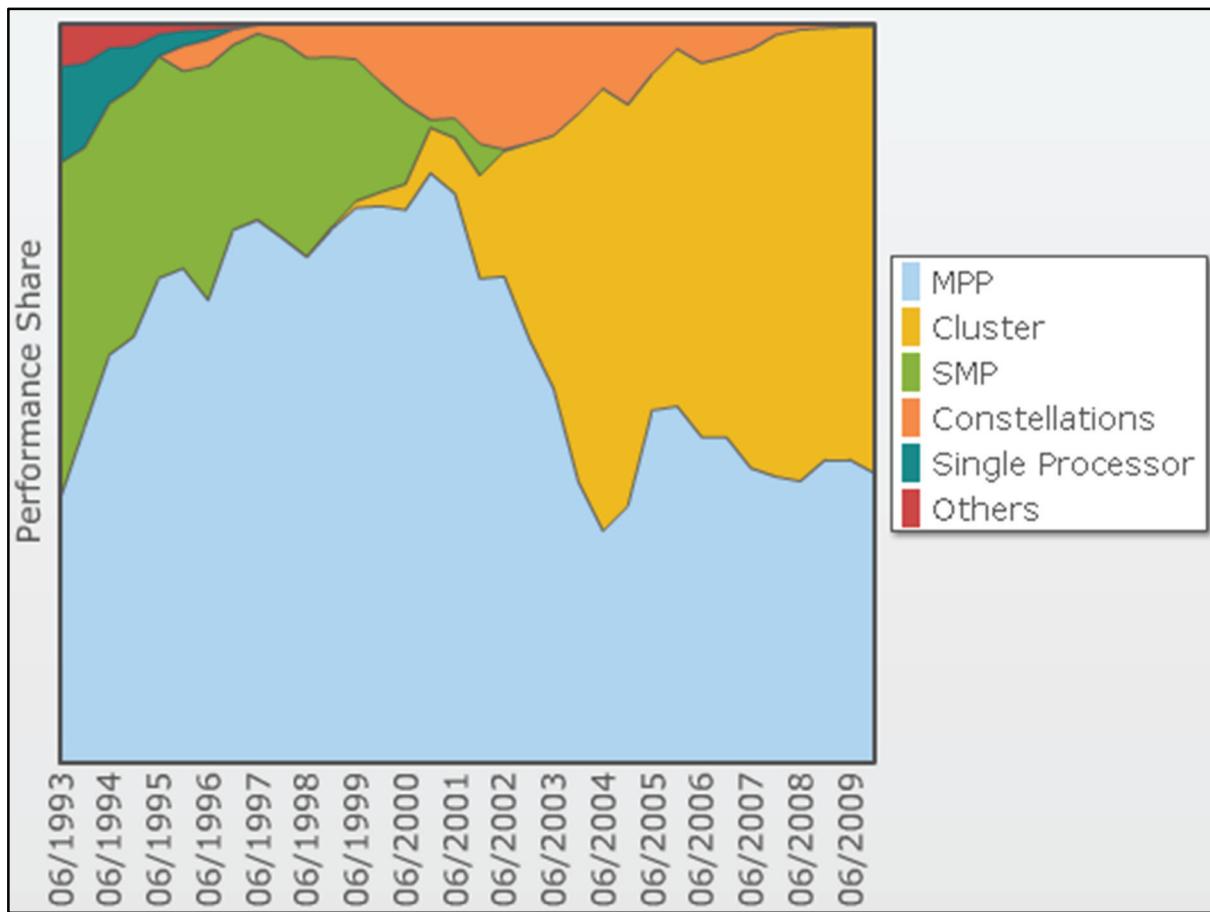


TOP500 Leistung: Die aktuellen Nr. 1-Rechner erzeugen immer ein Plateau über mehrere Listen hinweg (Z.B. ab 06/2002 5x der japanische Earth Simulator von NEC auf Platz 1). Die Leistung des Nr. 500-Rechners und die aggregierte Gesamtleistung aller 500 Rechner steigen in dieser Grafik linear, wobei die y-Achse logarithmisch aufgetragen ist.



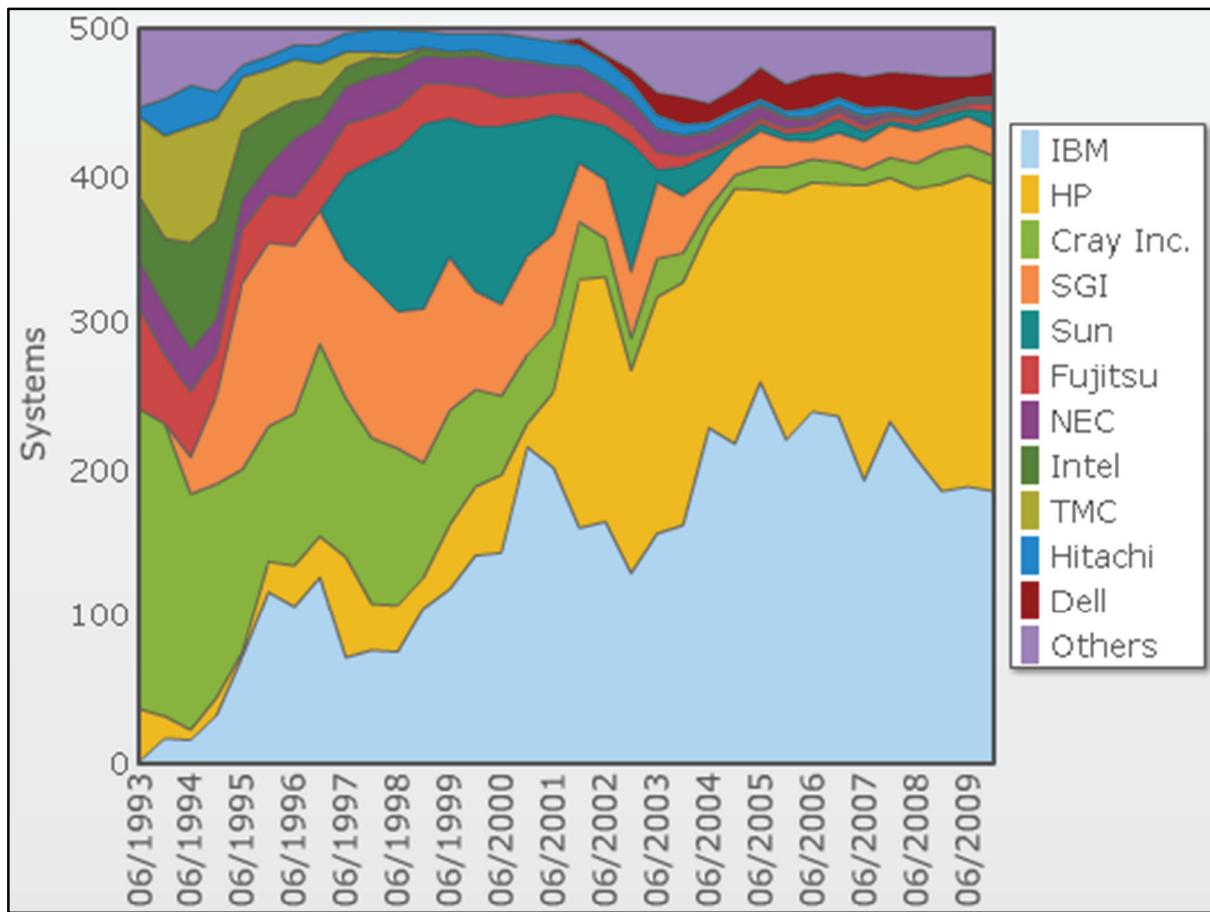
TOP500 Architekturen (# Systeme)

Die überwiegende Zahl der Installationen sind vom Typ Cluster. Beachten Sie die ausgestorbenen Architekturen: Einzelprozessor, SMP und Constellation.



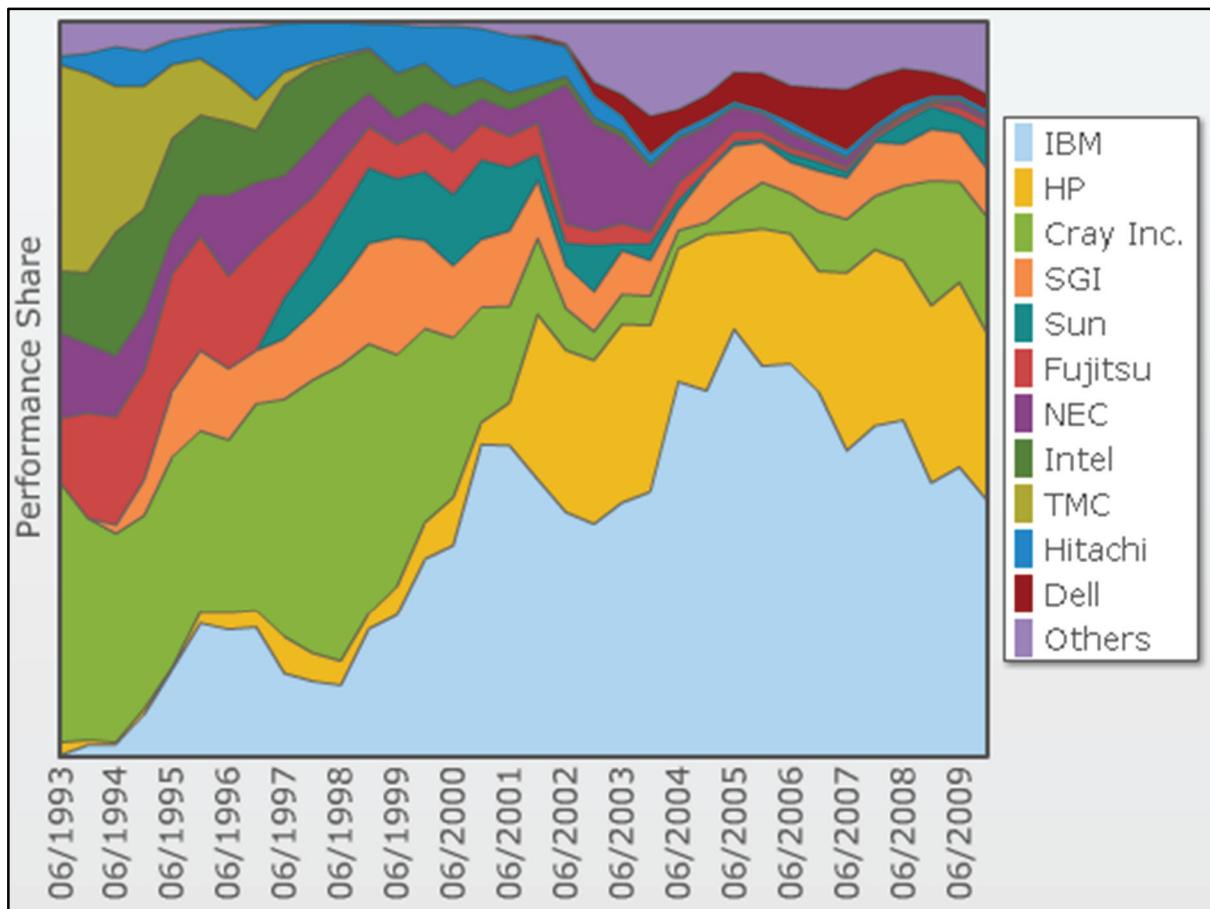
TOP500 Architekturen (% Leistung)

Der Leistungsanteil der Cluster ist geringer, weil die Menge der BlueGene-Systeme nicht vom Typ Cluster ist und insgesamt viel Leistung liefert.



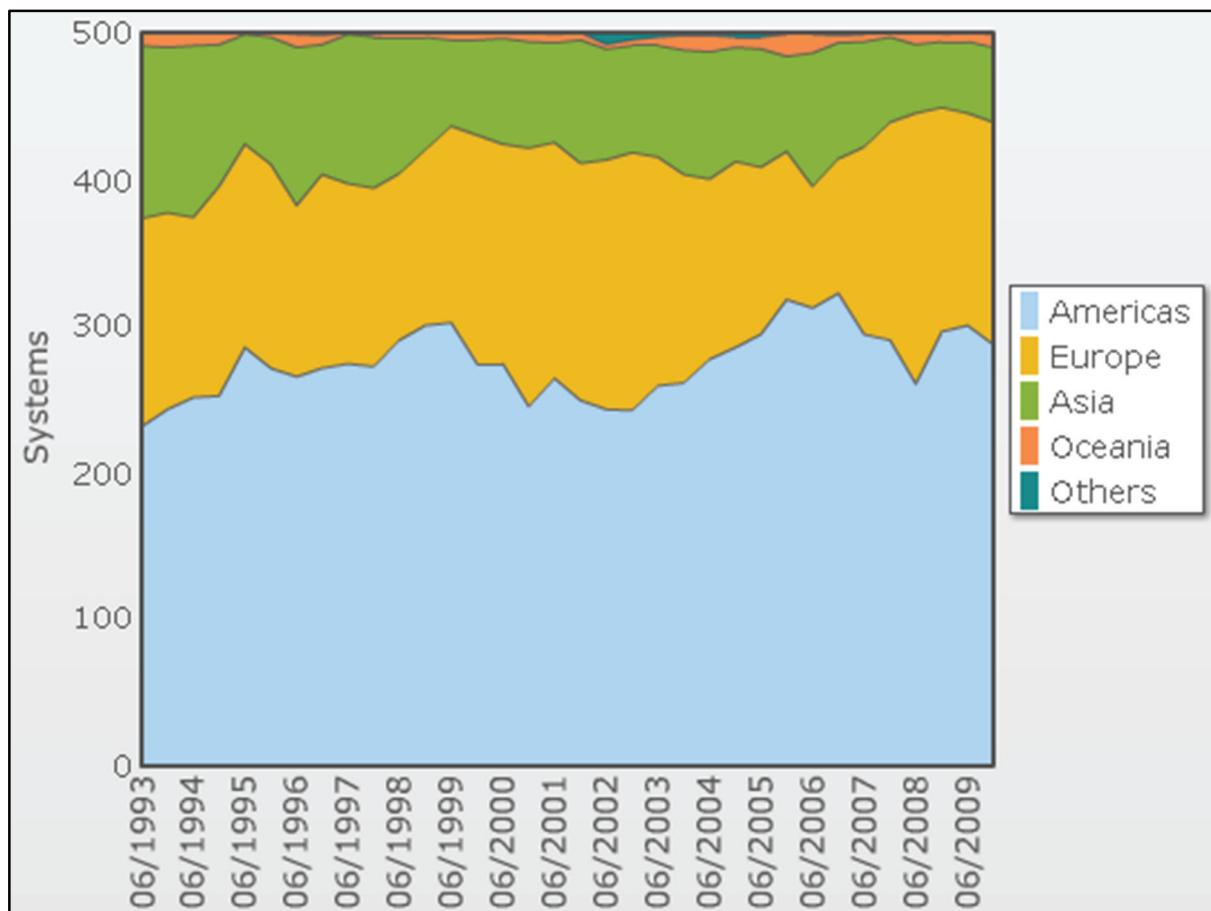
TOP500 Hersteller (# Systeme)

IBM und HP sind die Führer mit etwa gleich vielen Systemen am Markt.

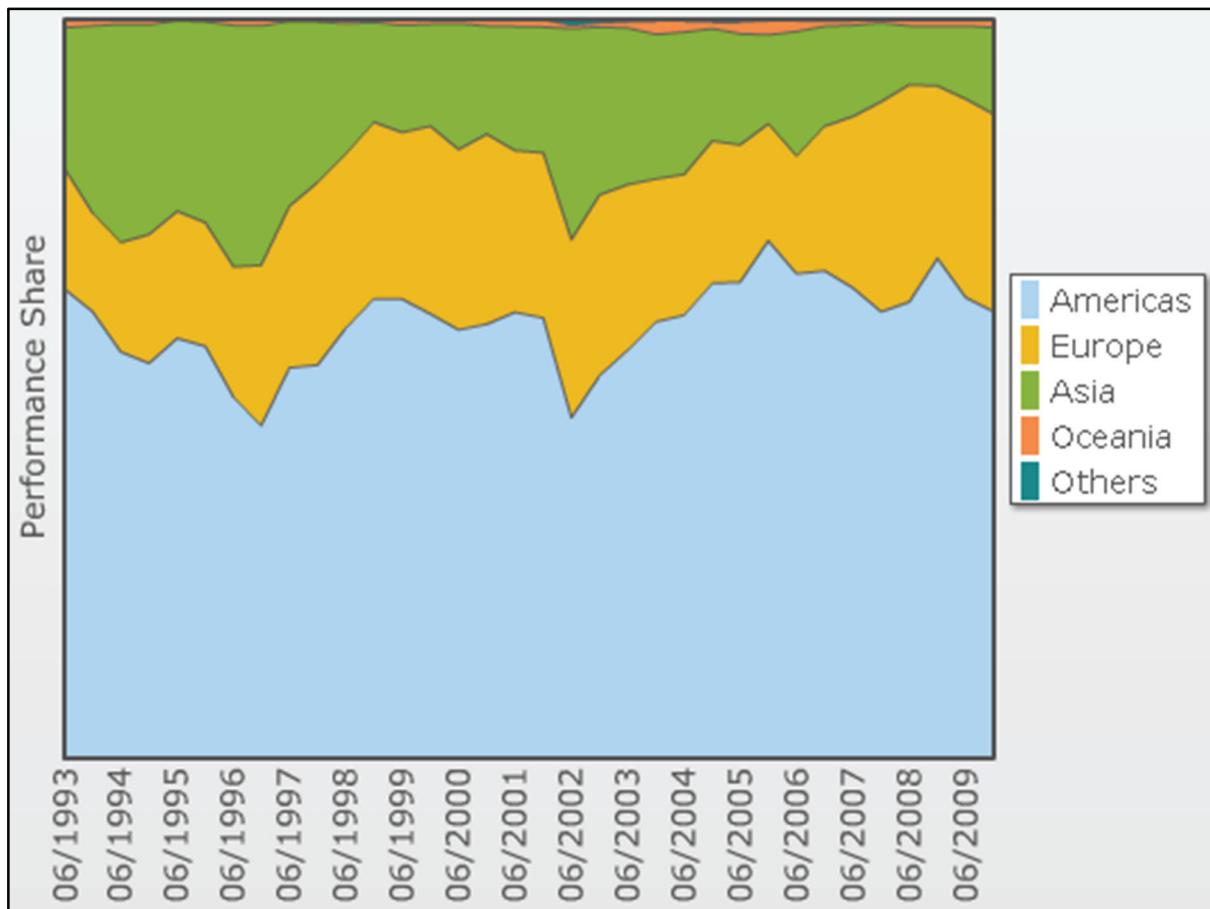


TOP500 Hersteller (% Leistung)

An der Leistung hat IBM aber wegen der BlueGene-Systeme einen viel größeren Gesamtanteil

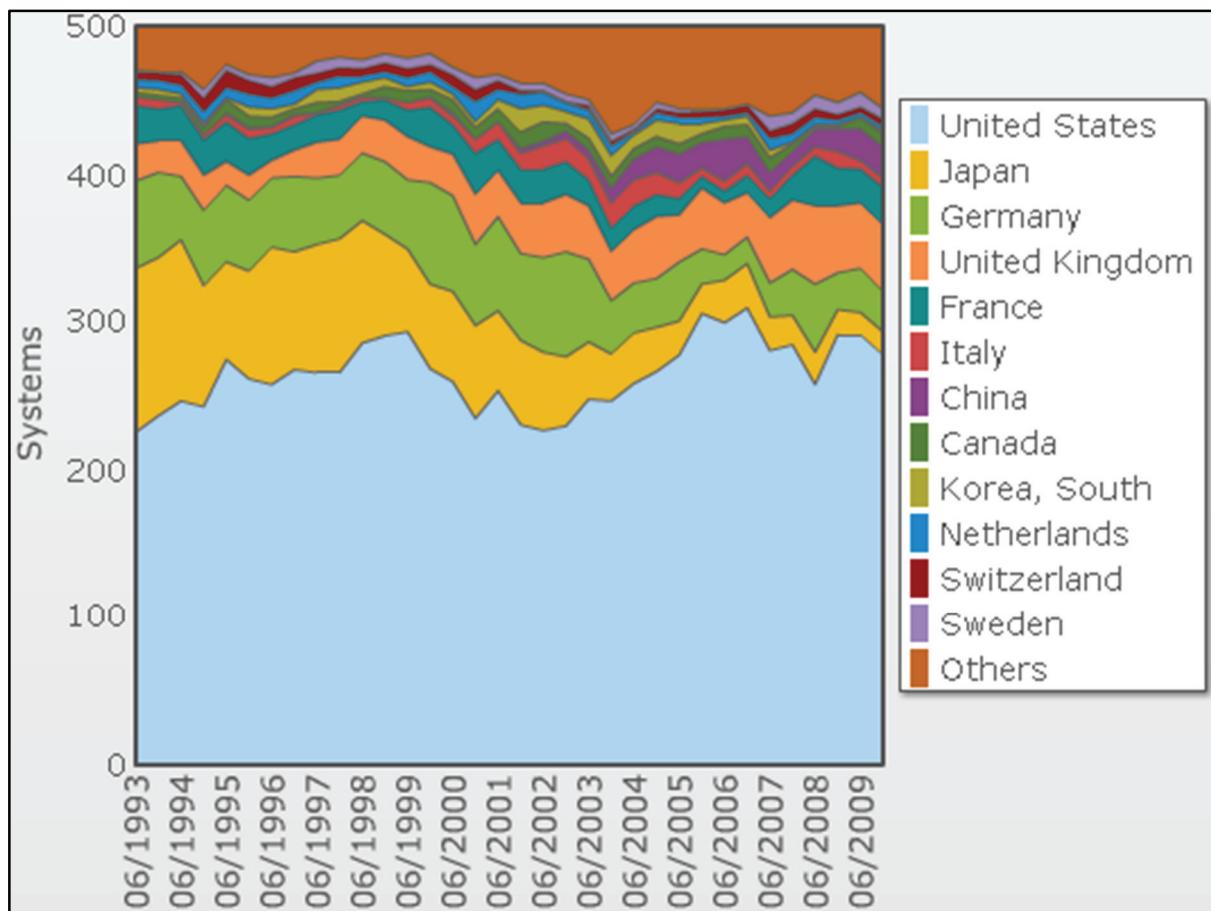


TOP500 Kontinente (# Systeme)

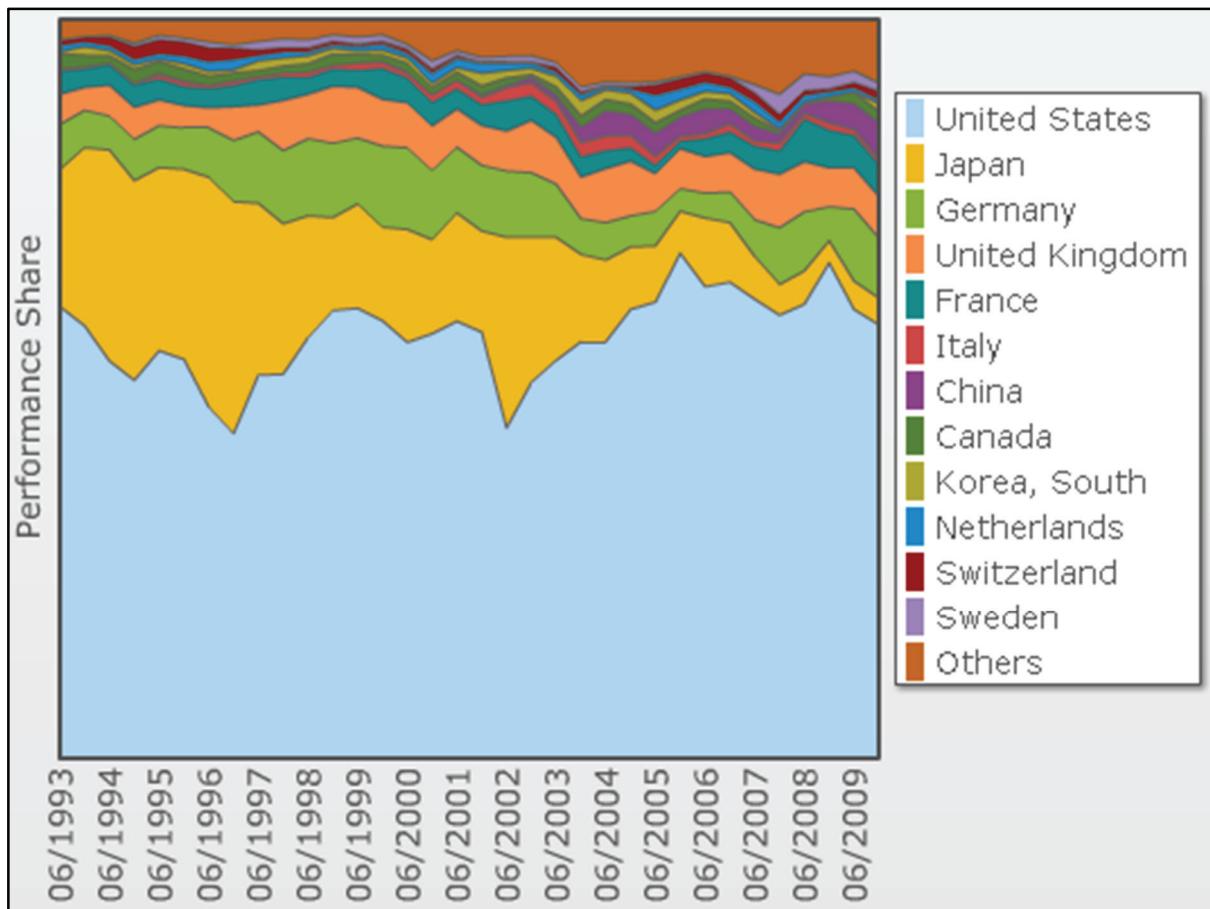


TOP500 Kontinente (% Leistung)

Man sieht Juni 2002 schön den Einfluß des japanischen Earth-Simulator von NEC: ein System mit sehr hoher Leistung, das Japan nach vorne bringt (grün!). (Ebenso Juni 1996 eine starke japanische Maschine von Hitachi).

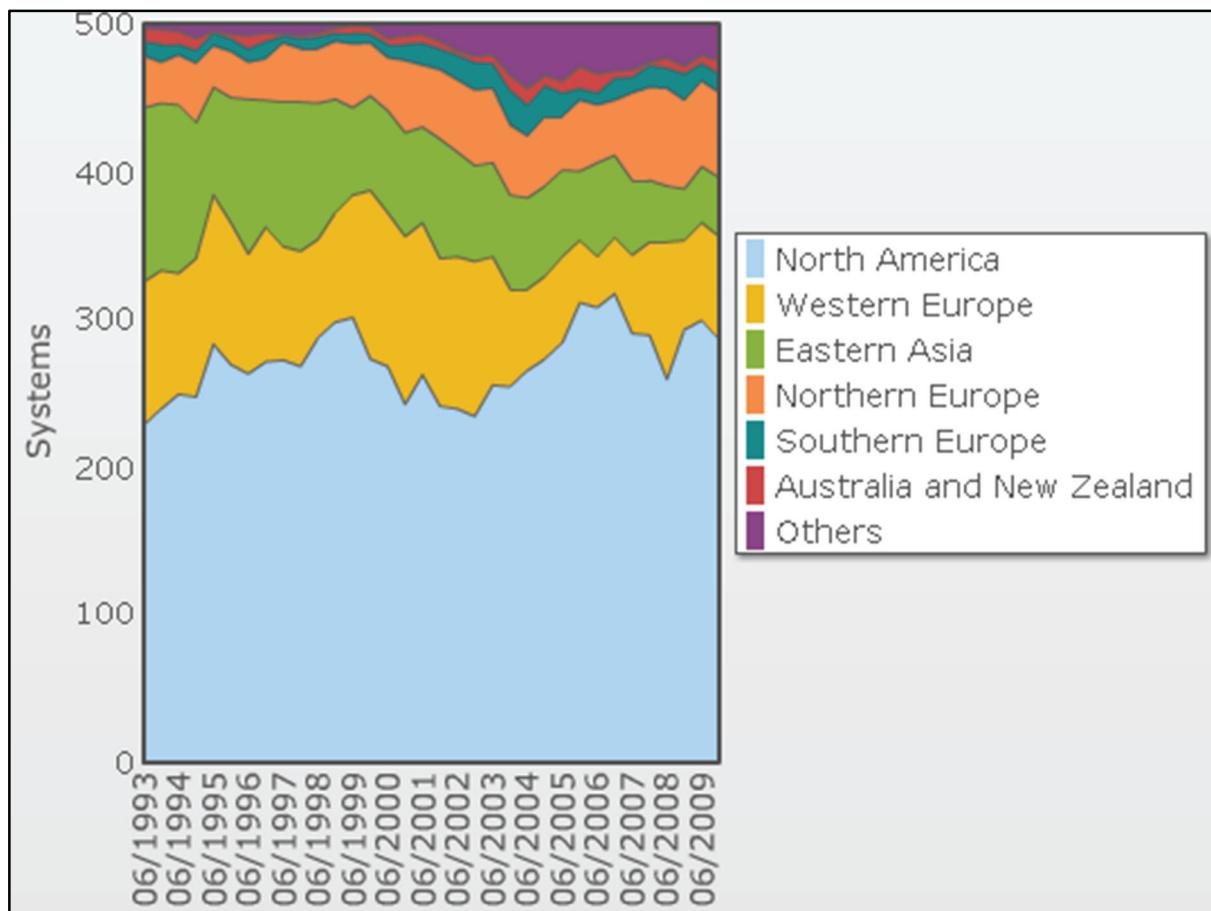


TOP500 Länder (# Systeme)

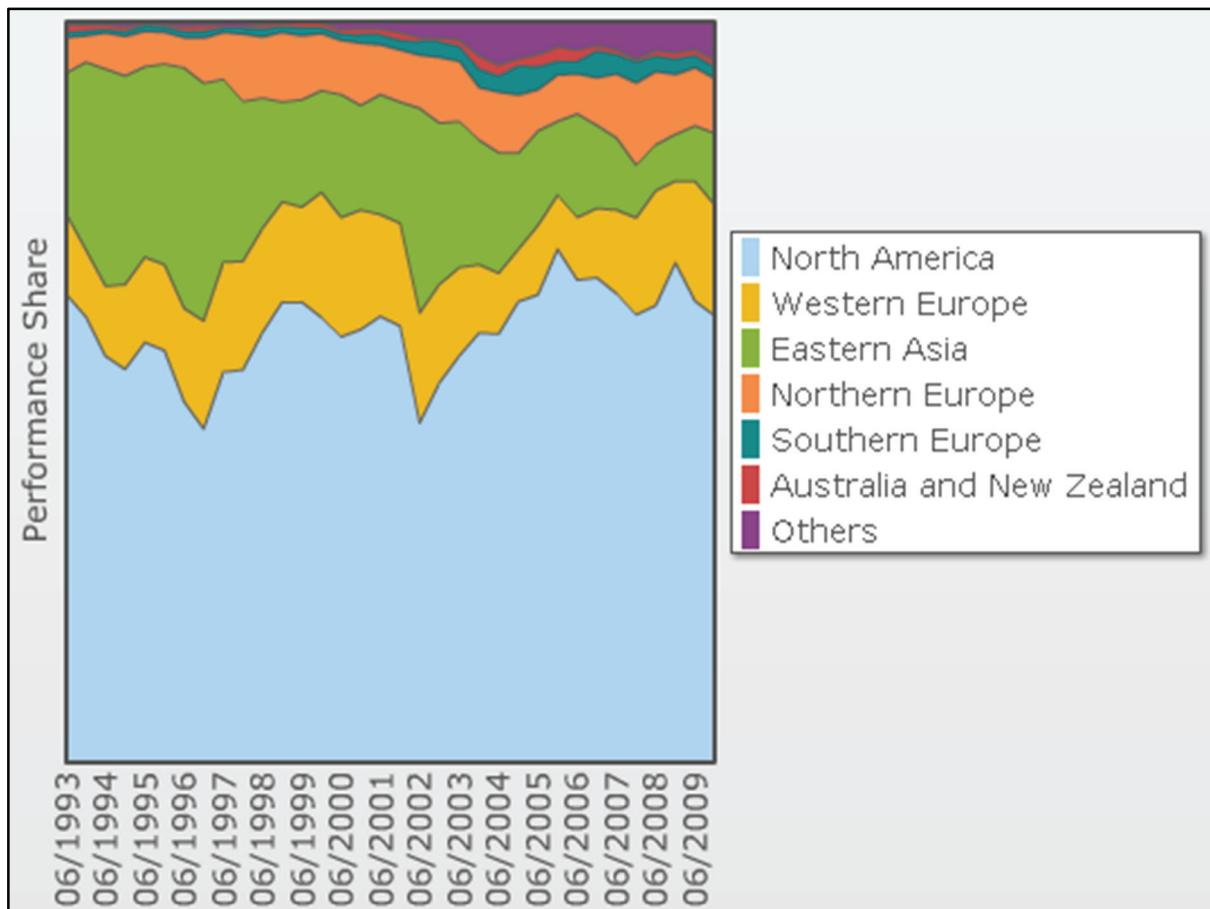


TOP500 Länder (% Leistung)

Wieder sieht man die Spitzen der Hitachi und des NEC Earth Simulator (gelb!).

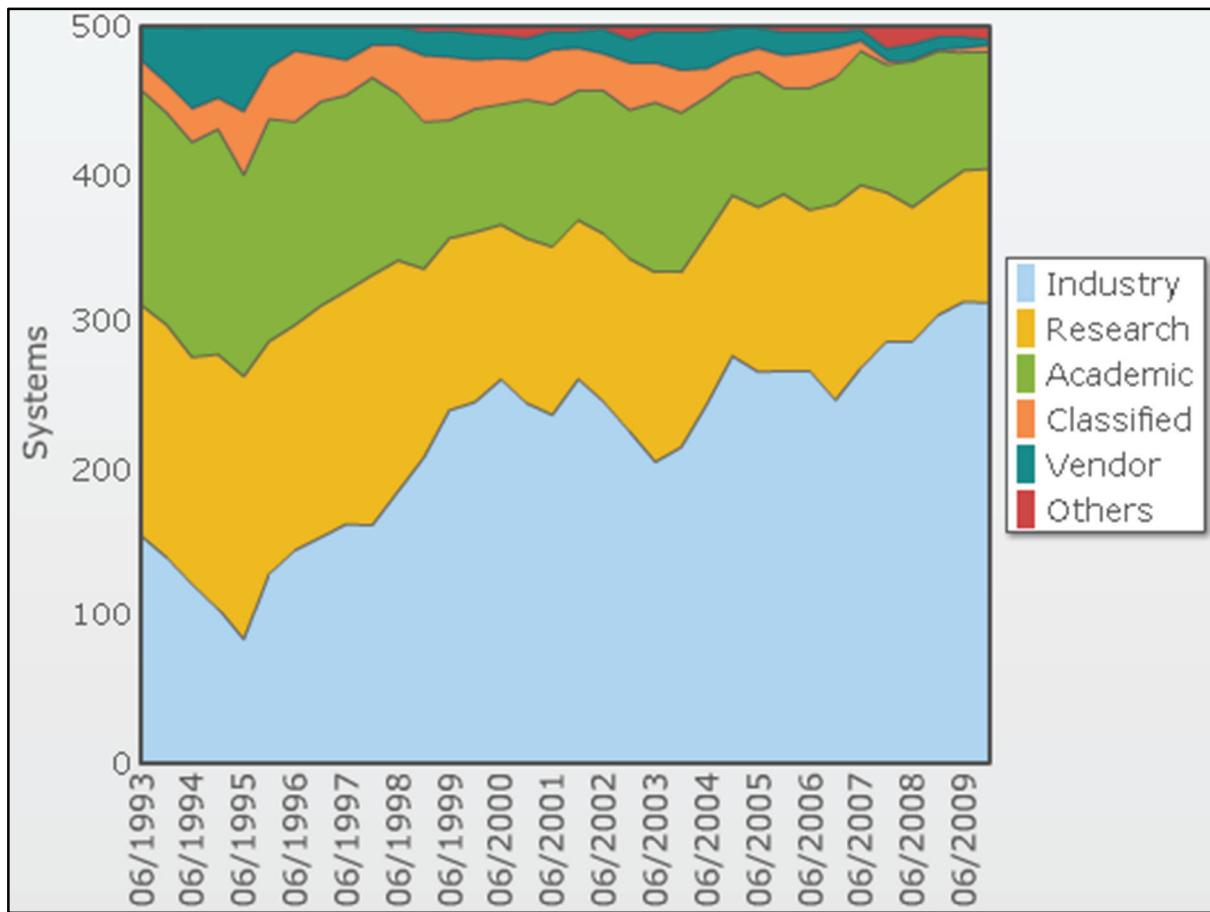


TOP 500 Geographische Regionen (# Systeme)



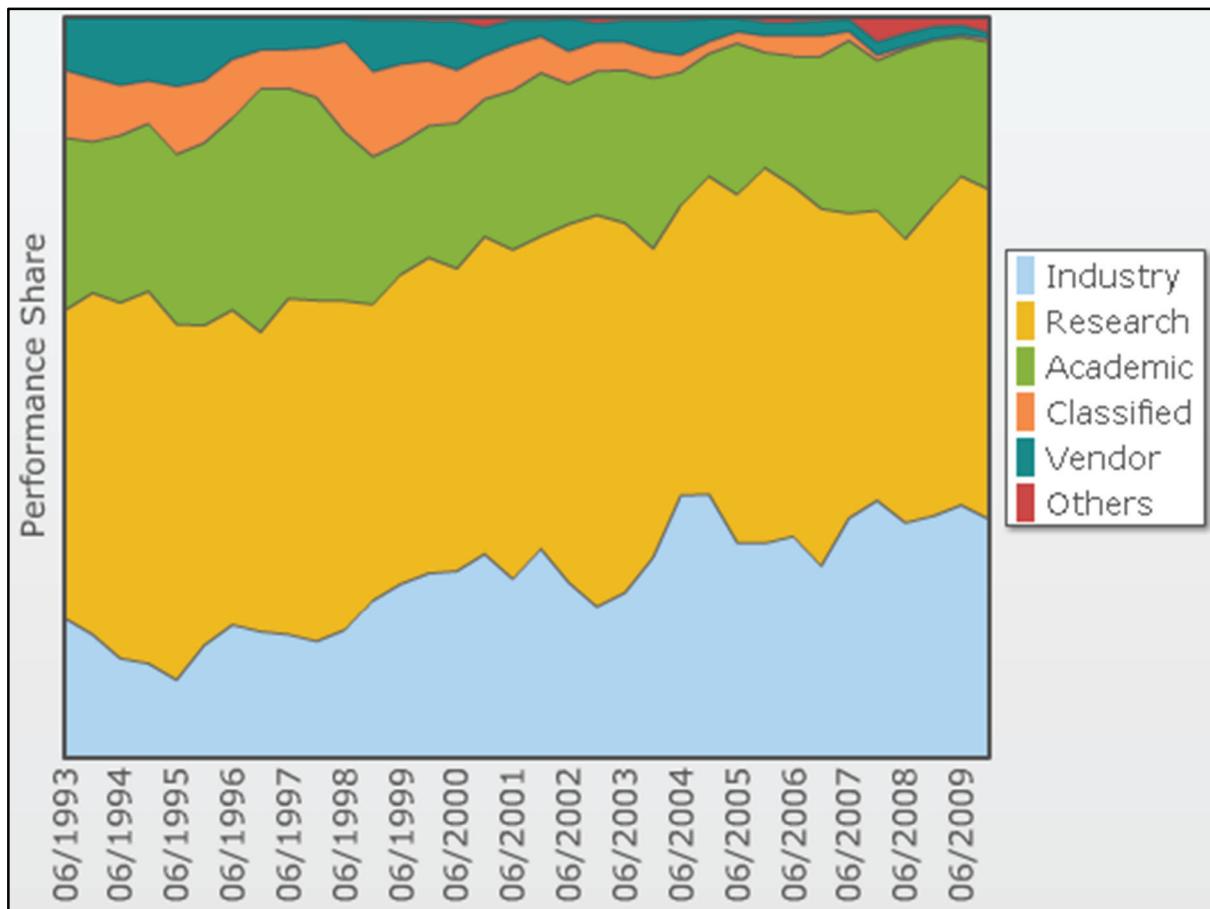
TOP500 Geographische Regionen (% Leistung)

Japan deutlich zu sehen (grün!).



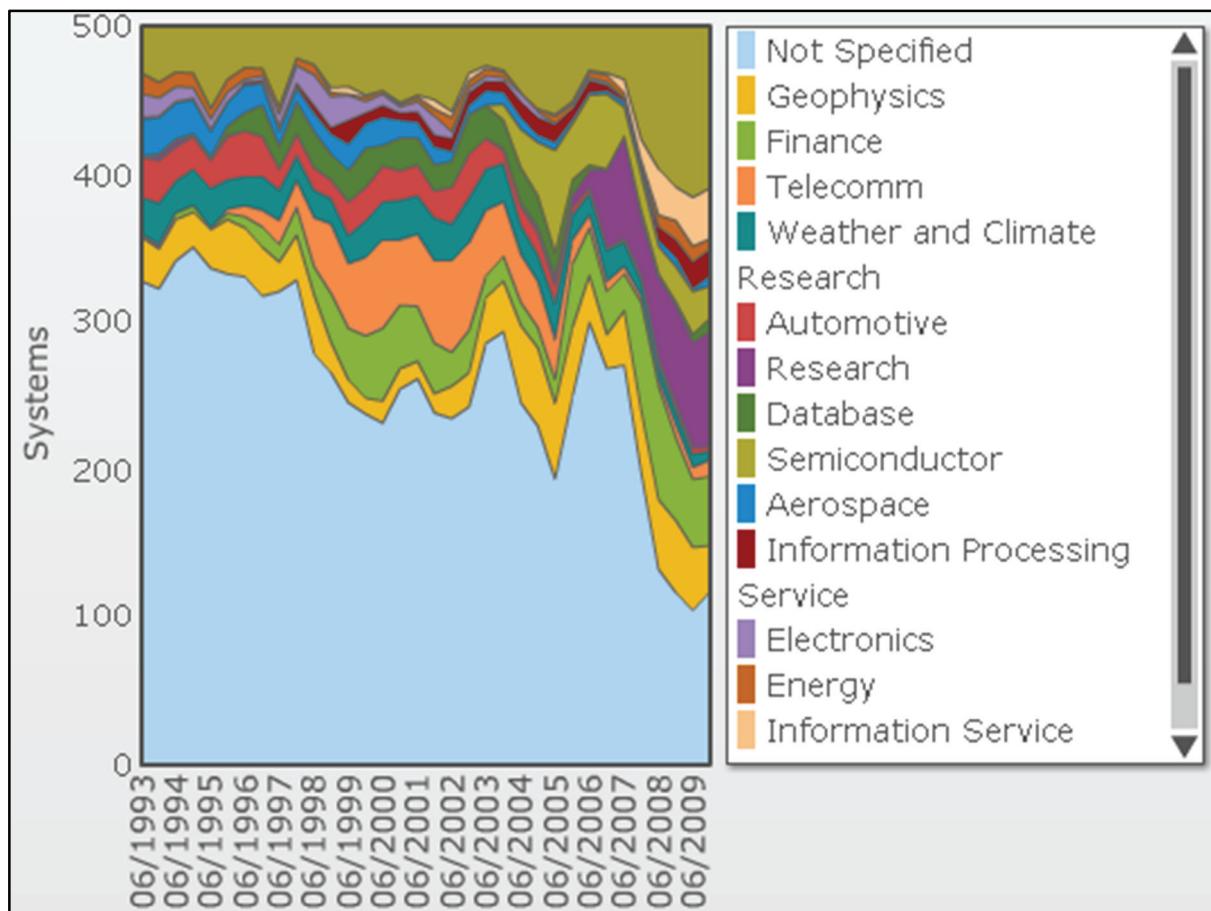
TOP500 Anwendergebiet (# Systeme)

Es werden zunehmend mehr Systeme in der Industrie installiert, d.h. das Hochleistungsrechnen findet Eingang in die Praxis. Insbesondere Ingenieurwissenschaften und Finanzwirtschaft.

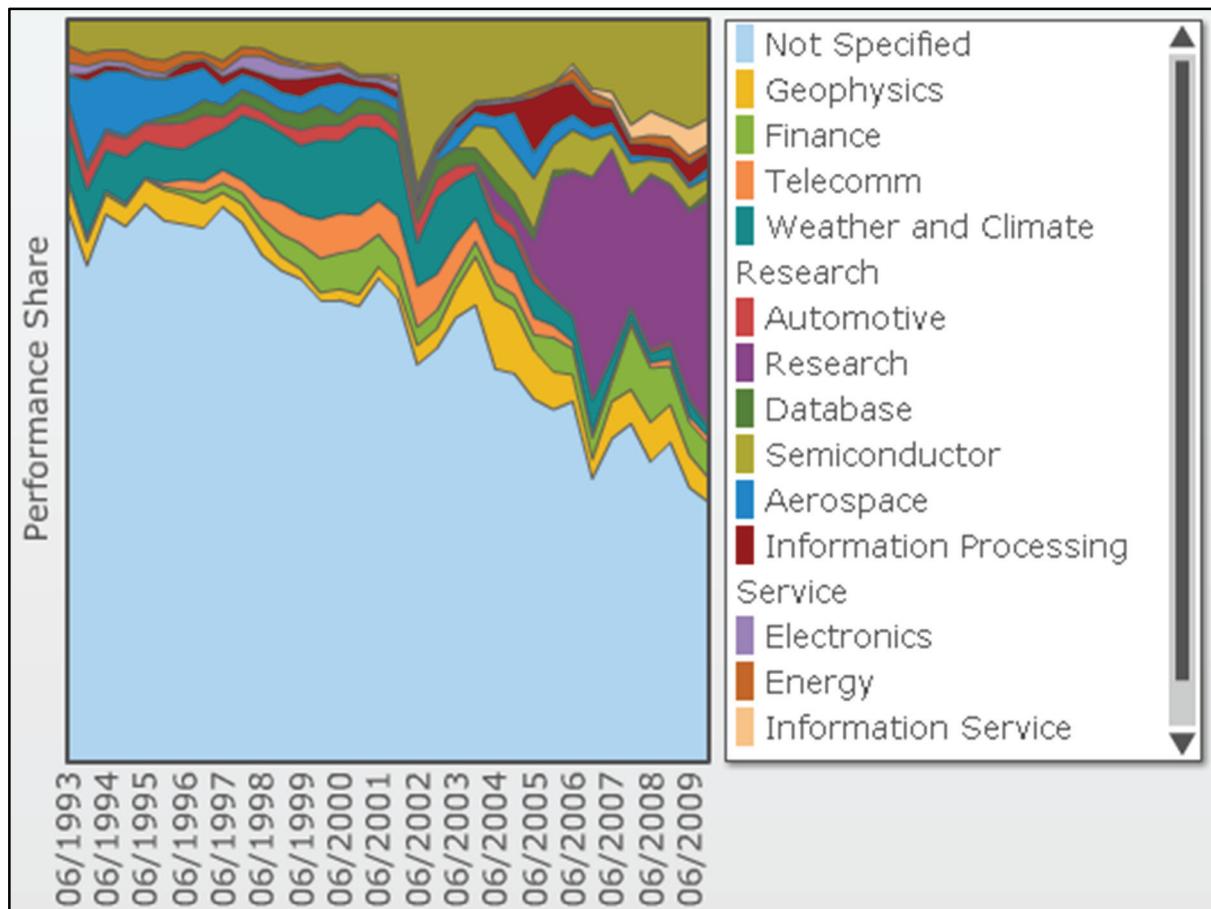


TOP500 Anwendergebiet (% Leistung)

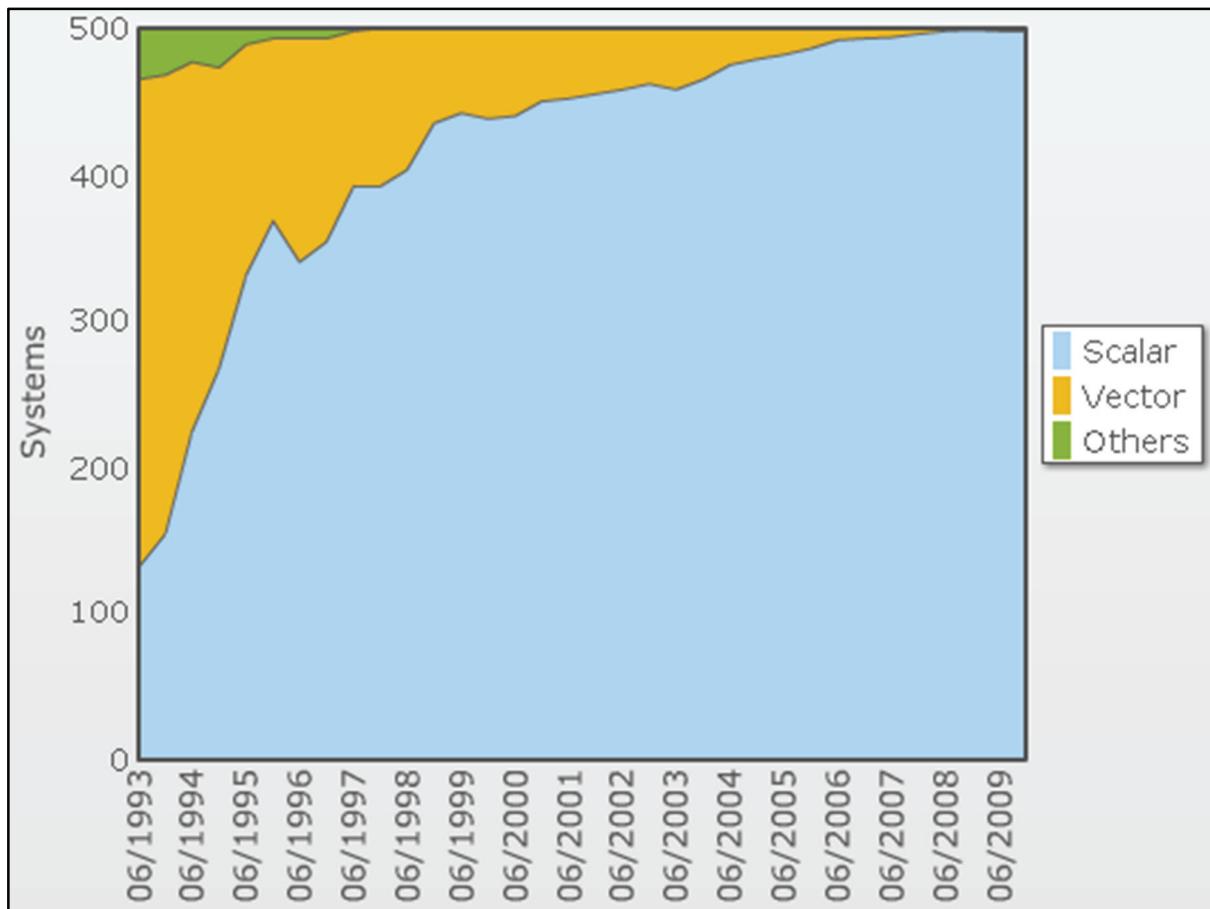
Das leistungsstärkste Gebiet bleibt aber die Forschung. D.h. die Industrie installiert eher kleine bis mittlere Systeme.



TOP500 Anwendungsgebiet (# Systeme)

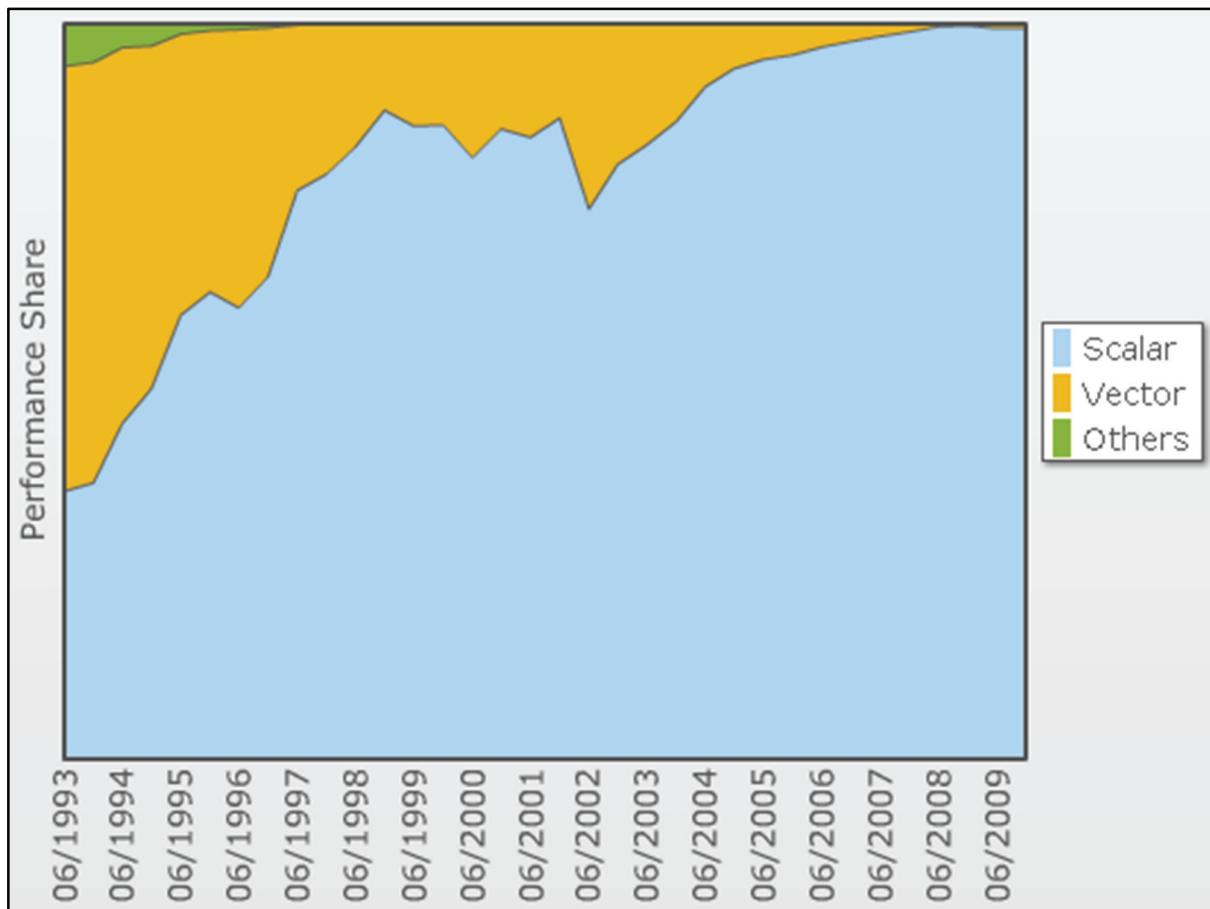


TOP500 Anwendungsgebiet (% Leistung)



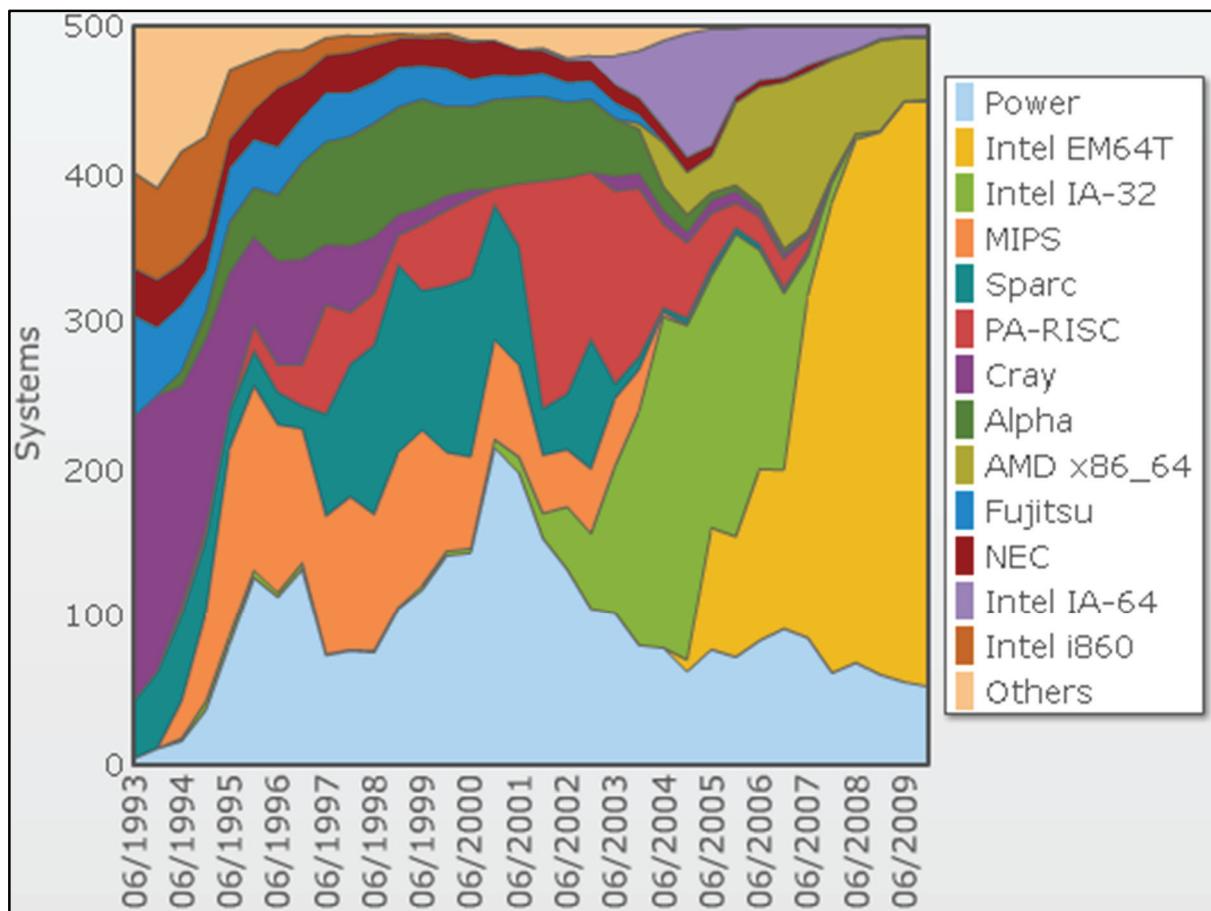
TOP500 Prozessorarchitektur (# Systeme)

Nur noch Skalarprozessoren.



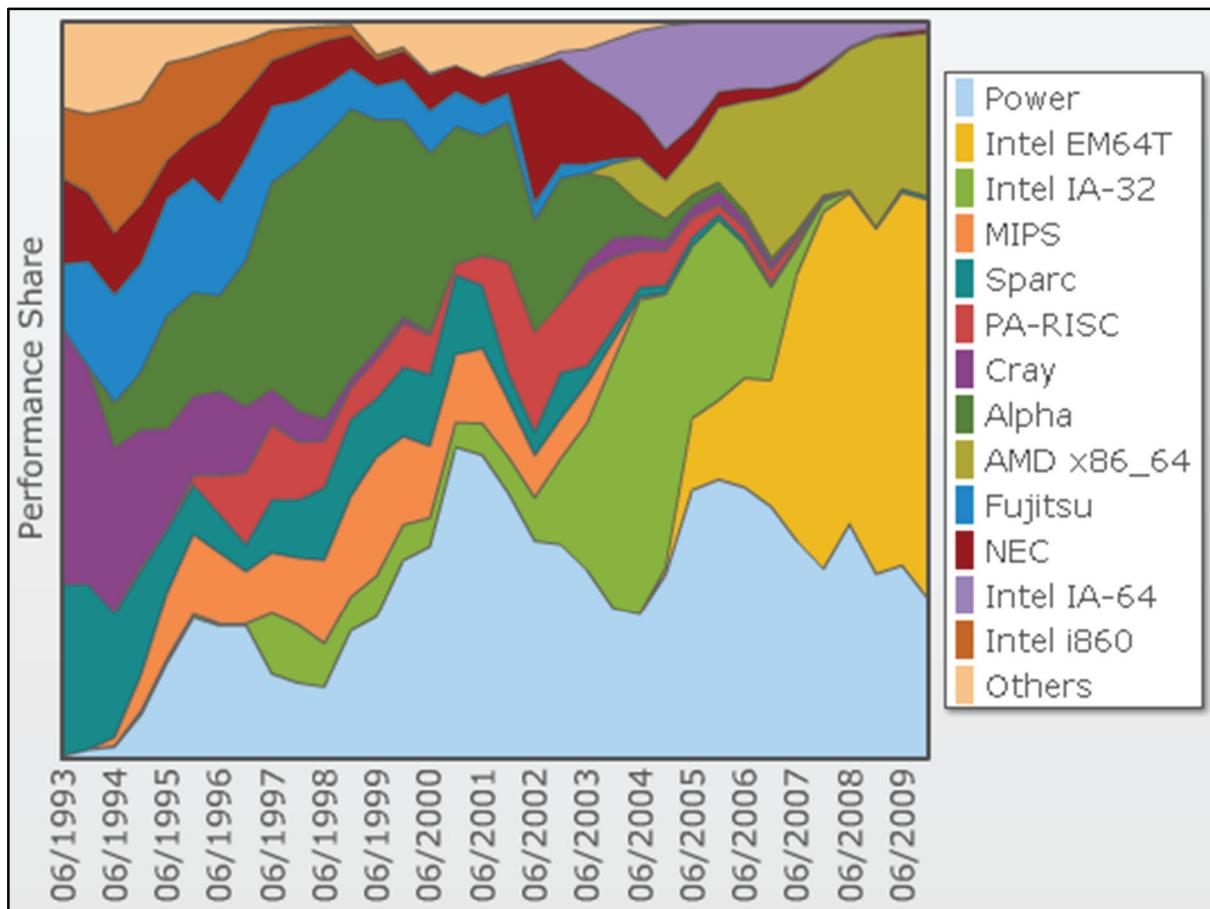
TOP500 Prozessorarchitektur (% Leistung)

06/2002 ein letztes Aufzucken der Vektorprozessoren im NEC Earth-Simulator.



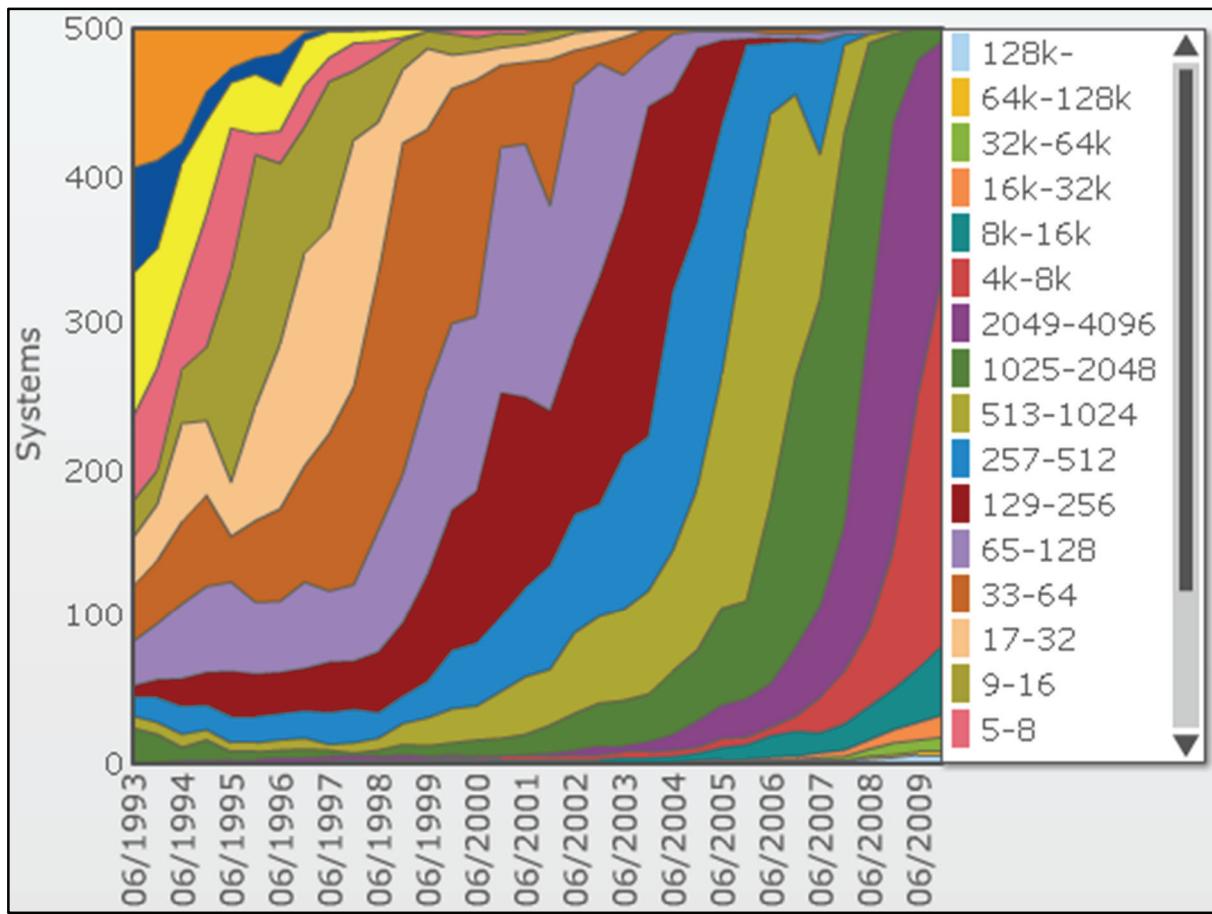
TOP500 Prozessorfamilie (# Systeme)

Intel dominiert die Anzahl der installierten Systeme vor Power und AMD.



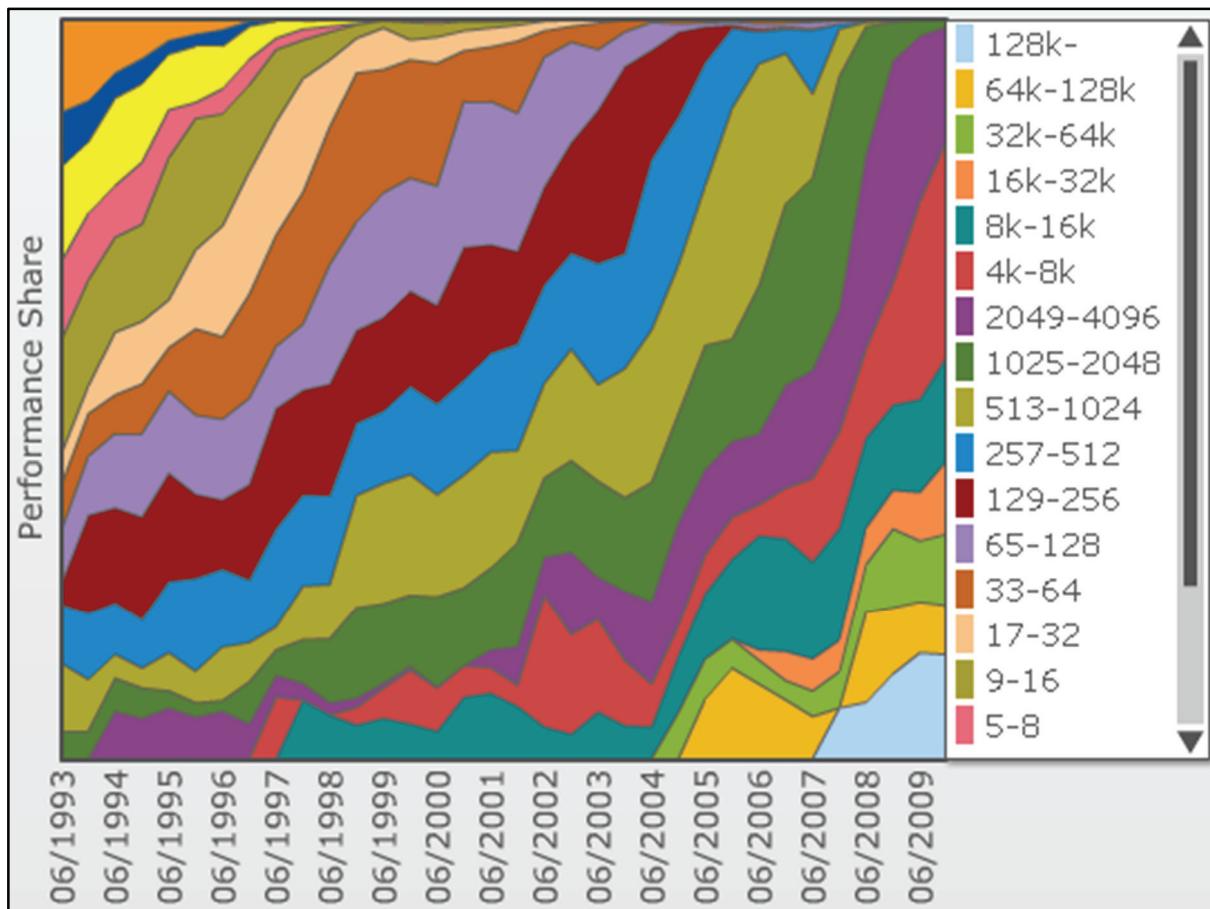
TOP500 Prozessorfamilie (% Leistung)

Im Leistungsanteil ist Power allerdings deutlich höher wegen der BlueGene-Systeme.



TOP500 Prozessoranzahl (# Systeme)

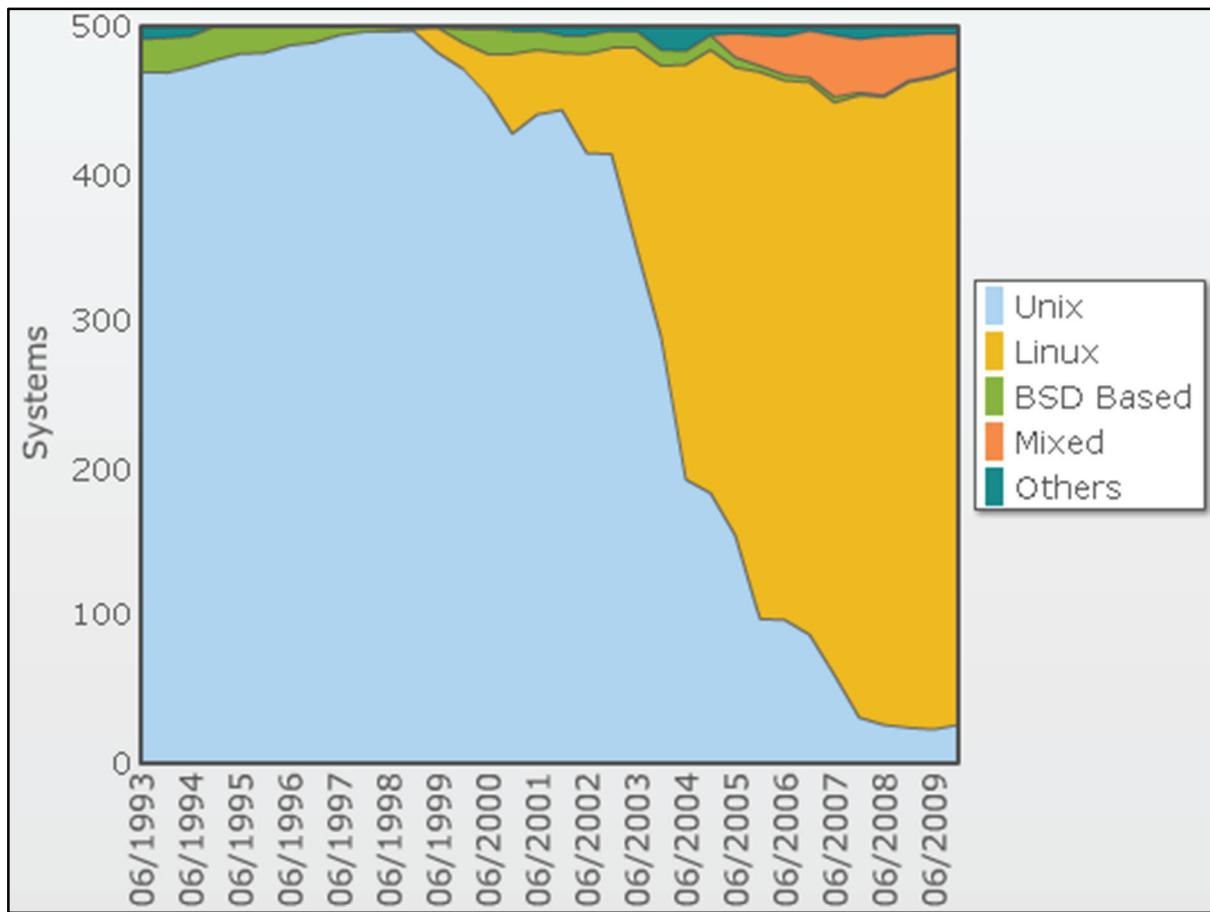
Nur wenige Systeme mit über 128K Prozessorkernen ...



TOP500 Prozessoranzahl (% Leistung)

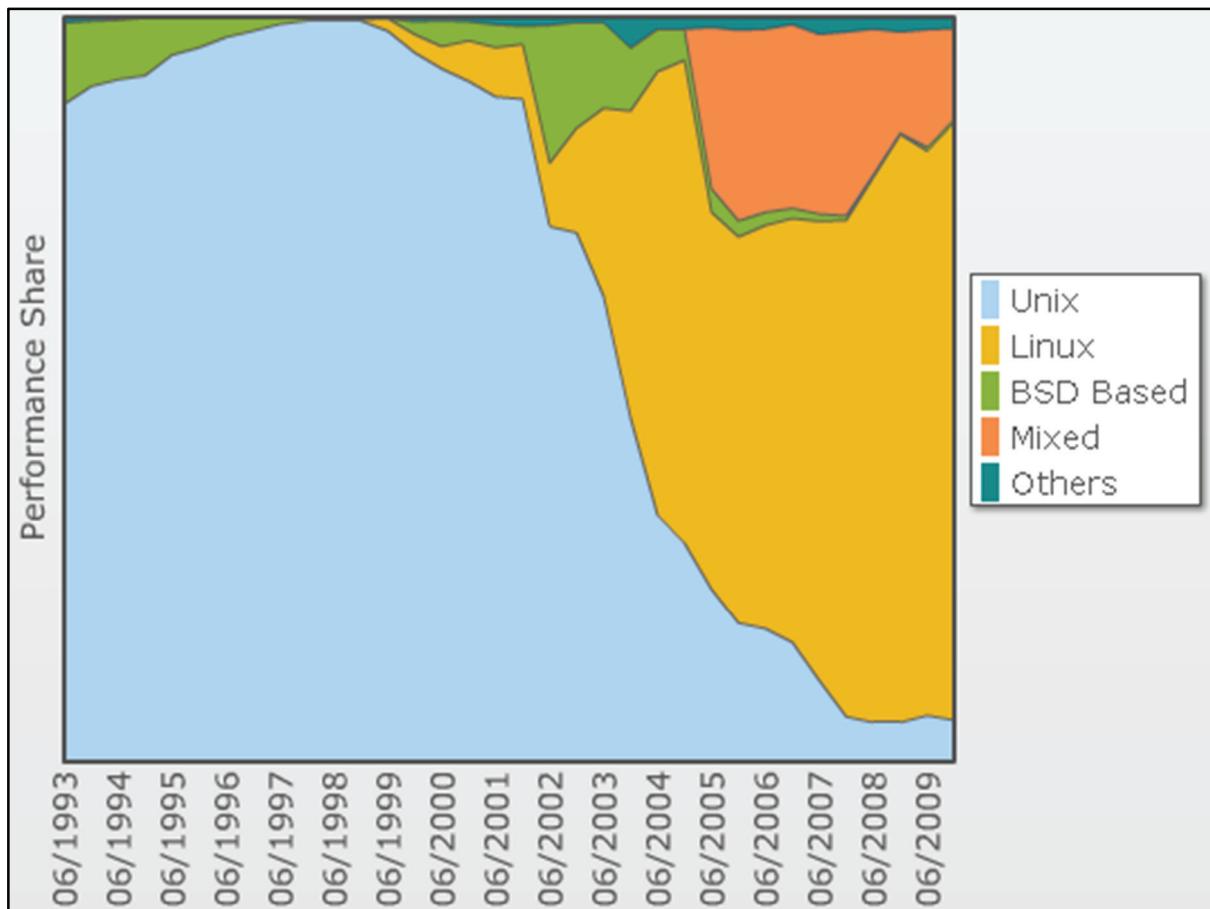
... aber die wenigen bringen einen deutlich sichtbaren Leistungsanteil (fast alles BlueGene-Systeme).

Interessant wäre eine Projektion für die kommenden zehn Jahre.



TOP500 Betriebssystemfamilie (# Systeme)

Linux dominiert die Betriebssysteme.



TOP500 Betriebssystemfamilie (% Leistung)

Gemischte Systeme anteilig stark wegen des Einsatzes auf BlueGene.



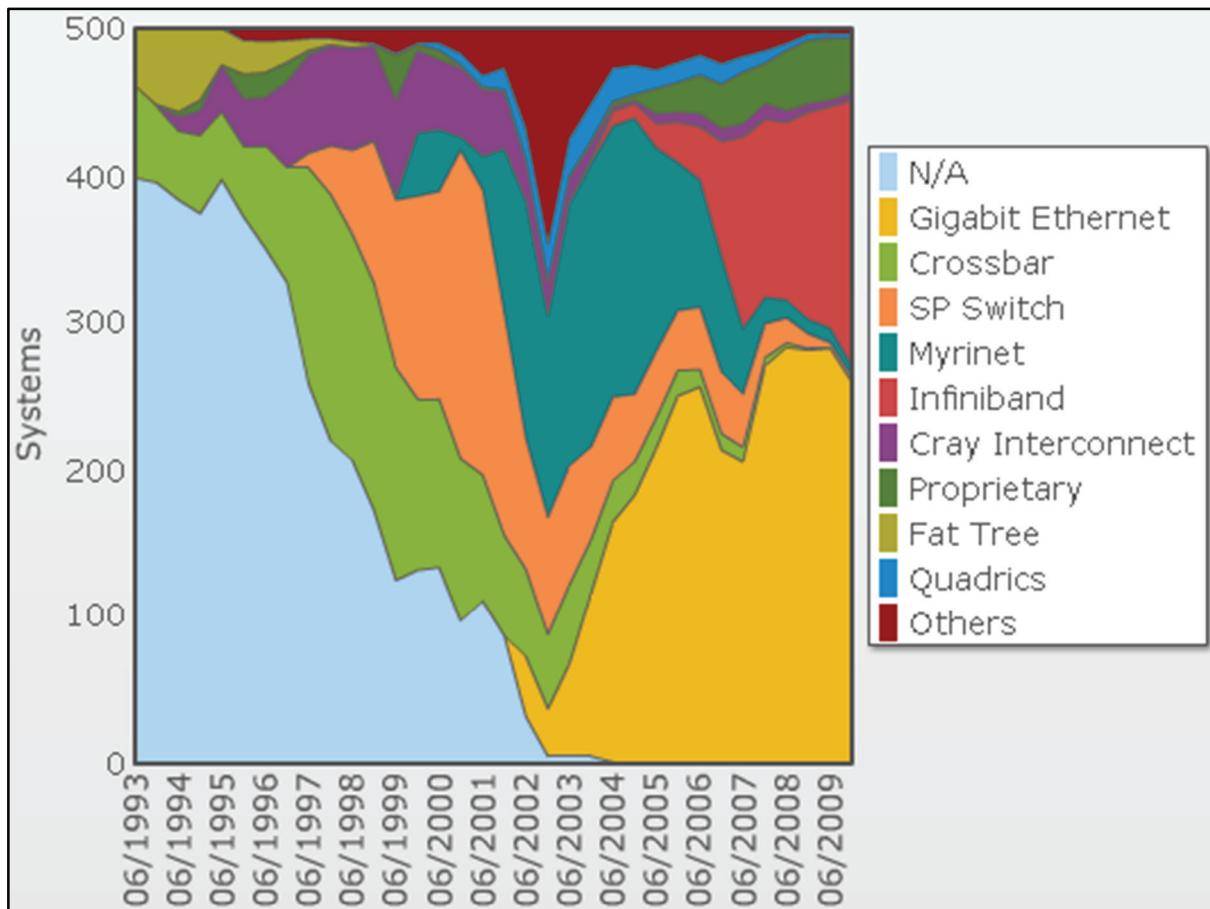
Verbindungstechnologien

In dedizierten Parallelrechnern

- ▶ Spezielle Hochleistungsnetze mit verschiedenen Topologien

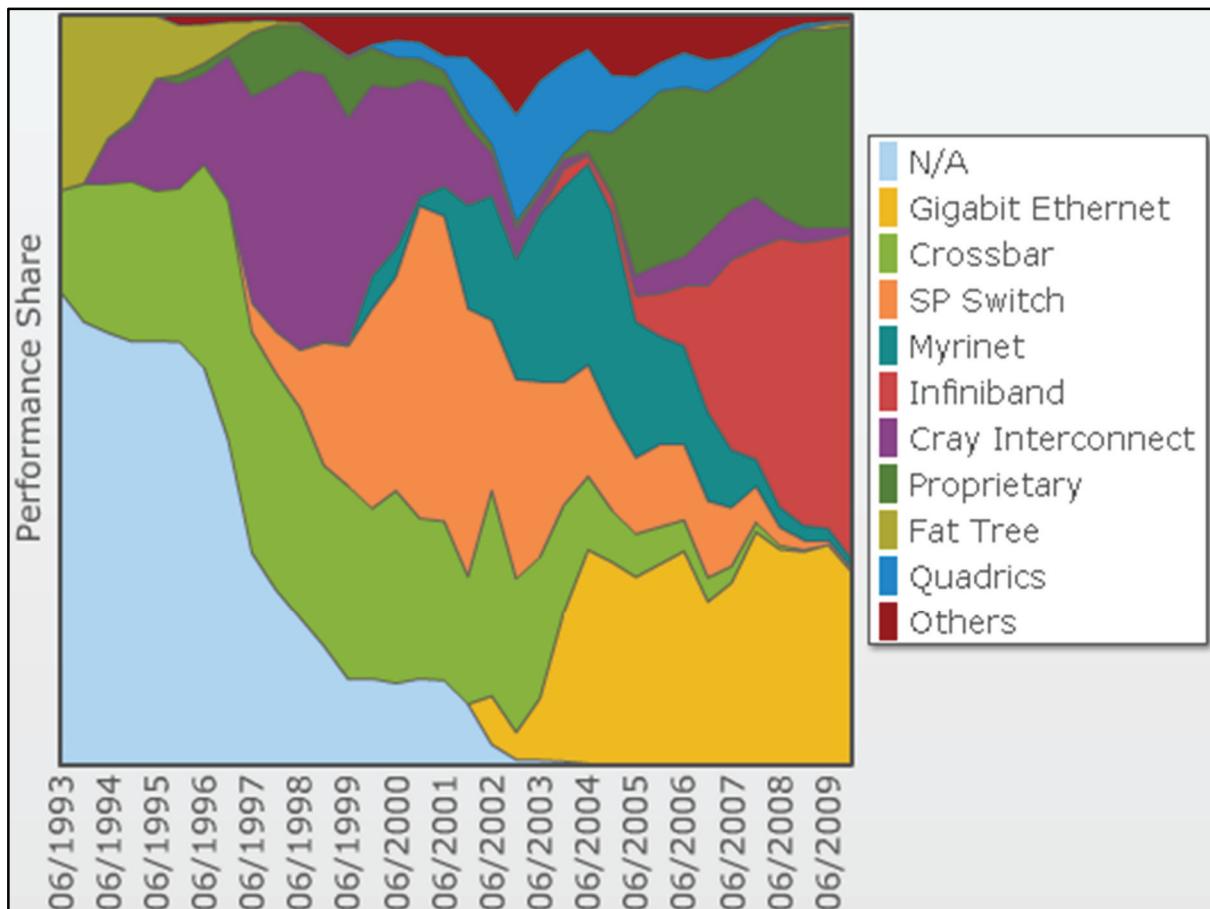
In Cluster-Architekturen

- ▶ Preiswert: Gigabit-Ethernet
- ▶ Teuer: Myrinet
- ▶ Sehr teuer: InfiniBand, Quadrics



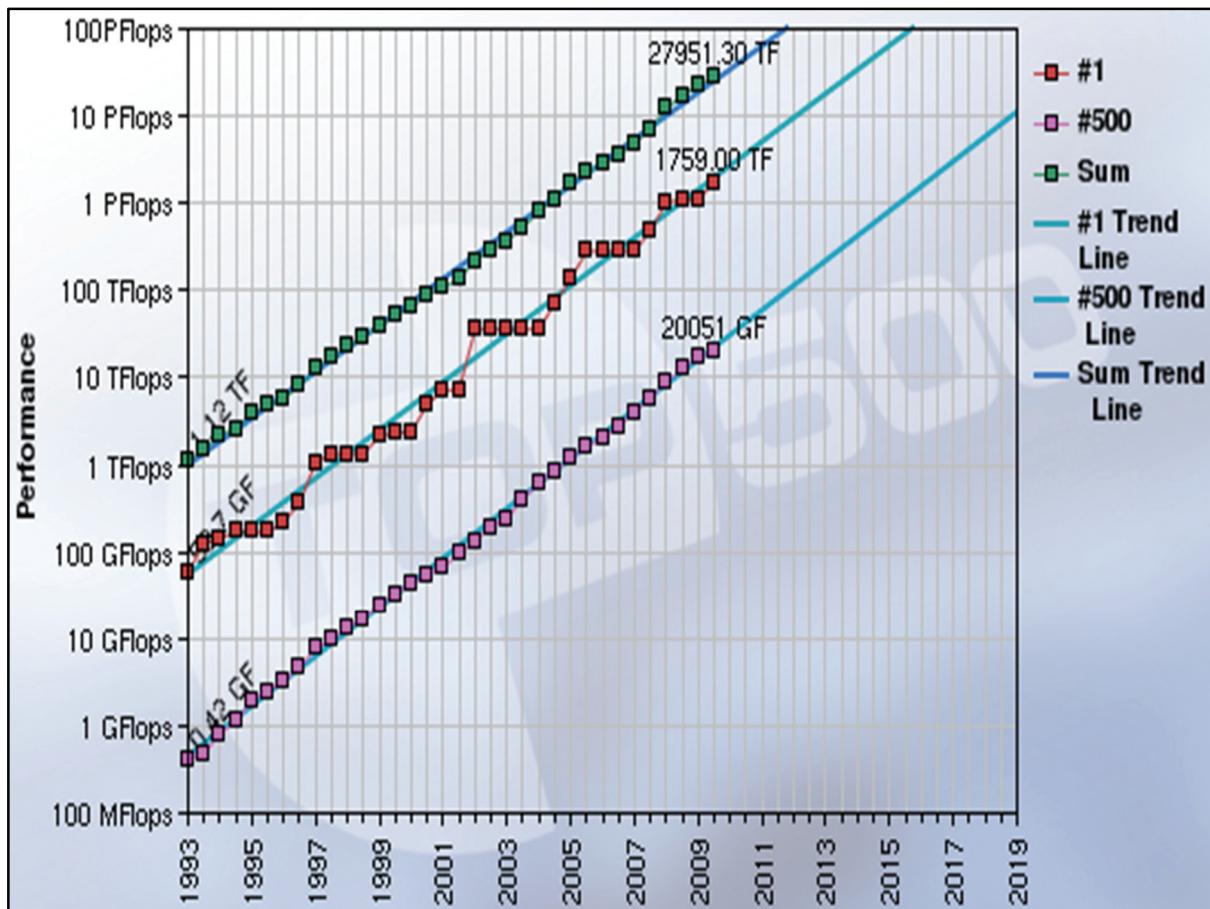
Verbindungstechnologie / Systeme

Gigabit-Ethernet dominiert die Installationen.

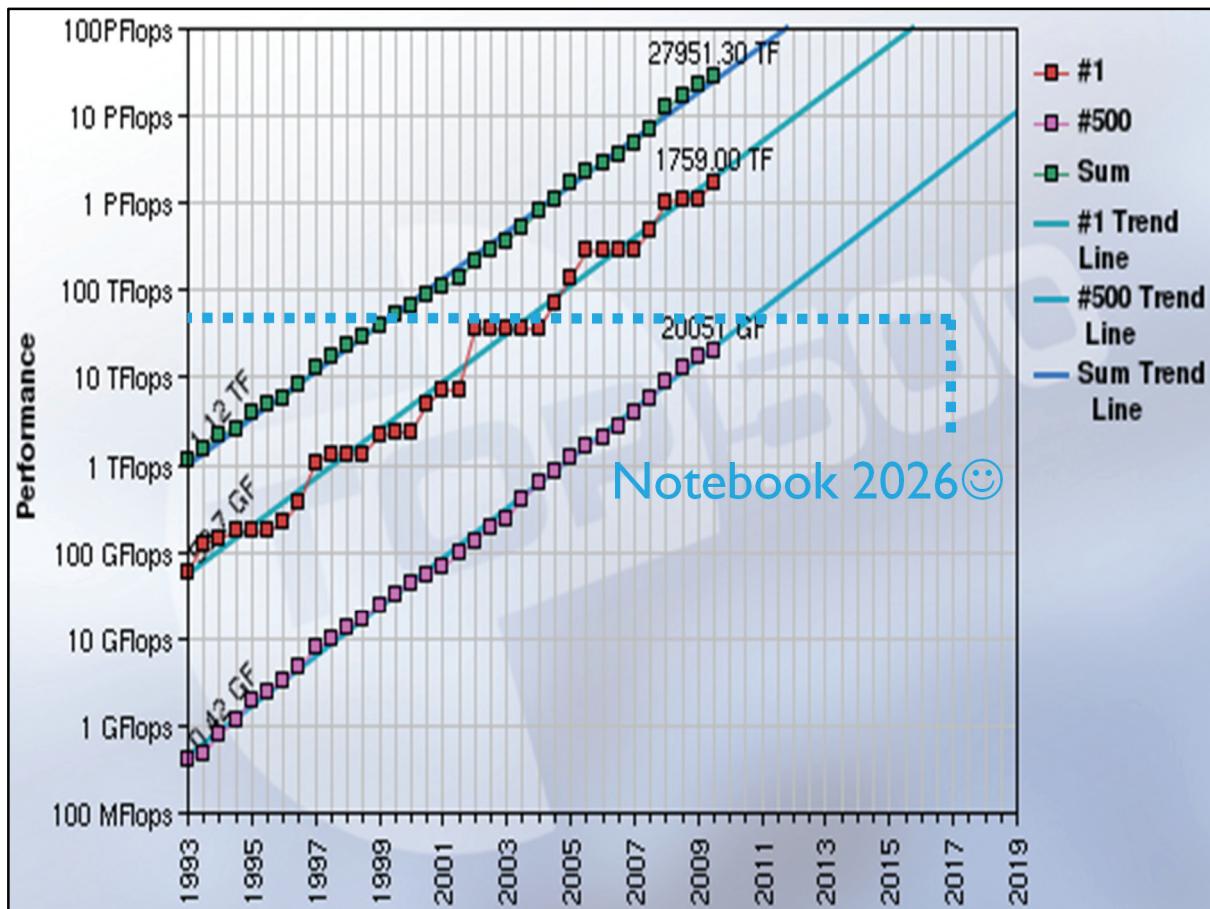


Verbindungstechnologie / Leistung

Allerdings hält Gigabit einen kleinen Leistungsanteil, weil es meist in kleineren und mittleren Systemen verwendet wird. Infiniband ist das wichtigste Netz für leistungsfähige Systeme. BlueGene hat ein proprietäres Netz, wie deutlich erkennbar ist.



Wir sind auf dem Weg zum Exaflops-Computing.



Notebook 2026: Faktor 30.000 über dem von 2010

Leistungsentwicklung bis November 2009

Moore's Law

„Verdopplung der Transistorzahl alle 18 Monate“
(entspricht evtl. Leistungsverdopplung)

Jun93–Nov09: 16,5 Jahre = 11×18 Monate
Faktor $2^{11} = 2048$

Leistung Summe: 1 TFlop/s – 27.951 TFlop/s ($\times 27.951$)

Leistung #1: 60 GFlop/s – 1.759 TFlop/s ($\times 29.300$)

Leistung #500: 0,4GFlop/s – 20.041 GFlop/s ($\times 50.100$)

NECs Earth Simulator (6/2002-11/2008)

- ▶ 640 Knoten
- ▶ Zu je 8 Vektorprozess.
- ▶ 5120 Prozessoren
- ▶ 0,15 mikron Kupfer

- ▶ 200 MioUSD Rechner
- ▶ 200 MioUSD Gebäude und Kraftwerk
- ▶ Zum Zwecke der Klimaforschung etc.

- ▶ 10 TByte Hauptspeicher
- ▶ 700 TByte Festplatten
- ▶ 1,6 PByte Bandspeicher
- ▶ 83.000 Kupferkabel
- ▶ 2.800 km/220 t Kabel
- ▶ 3250qm
- ▶ Erdbebensicher

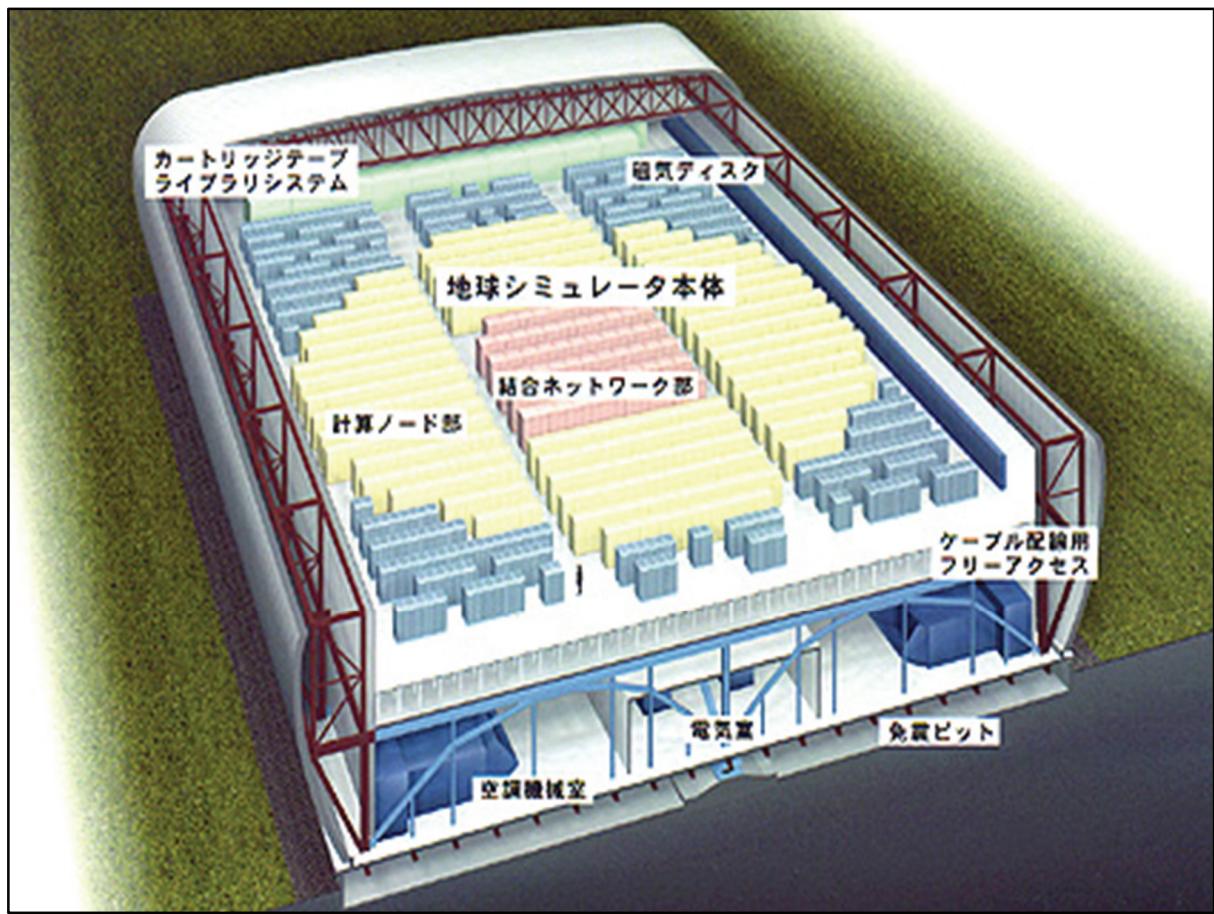
Computenik-Shock

► 84

Hochleistungsrechnen - © Thomas Ludwig

15.04.2010

Die Installation des Rechners traf die Amerikaner gleichermaßen schwer wie damals, als die Russen den Sputnik iins All schossen. Daher die Bezeichnung Computenik-Shock.



IBMs BlueGene-Programm

- ▶ Ursprünglich Hauptanwendung: Proteinfaltung
- ▶ Spezialprozessoren: leistungsschwach aber stromsparend – wenig Hauptspeicher pro Core
- ▶ Betriebssystem: mehrere Varianten von Linux
- ▶ Verbindungsnetz: proprietärer 3dim-Torus mit Zusatznetzen für globale Kommunikation, E/A und Verwaltung
- ▶ Systeme:
 - ▶ Blue Gene/L: 180TFlops/s, Ende 2004
 - ▶ Bis Ende 2005 verdoppelt auf 370TFlops/s
 - ▶ Blue Gene/P: 1PFlop/s, 2006/2007
 - ▶ Blue Gene/Q: 3PFlop/s, 2007/2008

► 87

Hochleistungsrechnen - © Thomas Ludwig

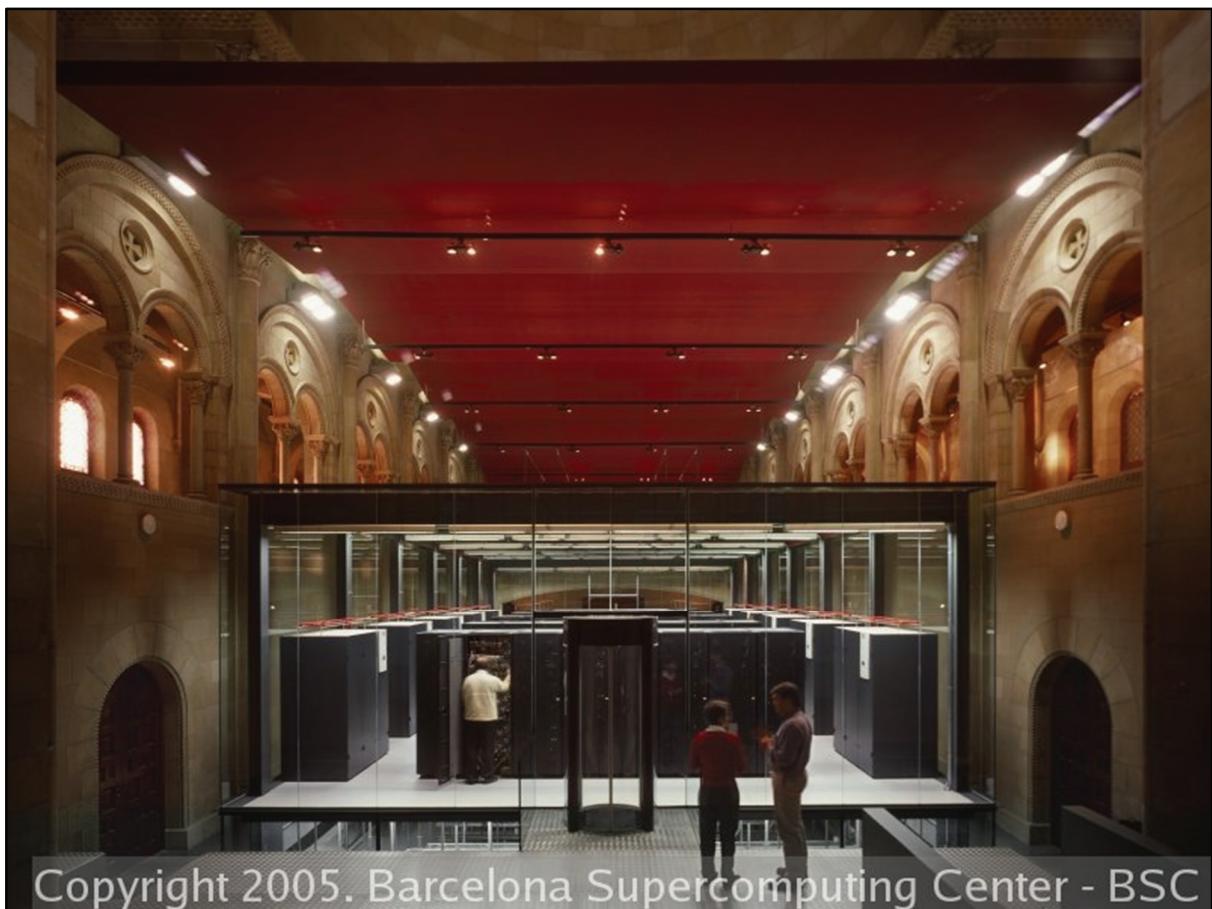
15.04.2010

Siehe: <http://en.wikipedia.org/wiki/Bluegene>

Gewinner
des
Schönheitswettbewerbs



Copyright 2006, Barcelona Supercomputing Center - BSC



Copyright 2005. Barcelona Supercomputing Center - BSC



Historische Sicht

- ▶ Die TOP500 im Juni 1993
- ▶ Deutschland in der TOP500 im Juni 1993

▶ 95

Hochleistungsrechnen - © Thomas Ludwig

15.04.2010

Rank	Manufacturer Computer/Procs	R _{max} R _{peak}	Installation Site Country/Year	Inst. type Installation Area	Nmax Nhalf	Computer Family Computer Type
1	TMC CM-5/1024/ 1024	59.70 131.00	Los Alamos National Laboratory USA/	Research Energy	52224 24064	TMC CM5 CM5
2	TMC CM-5/1024/ 1024	59.70 131.00	National Security Agency USA/	Classified	52224 24064	TMC CM5 CM5
3	TMC CM-5/544/ 544	30.40 70.00	Minnesota Supercomputer Center USA/	Industry	36864 16384	TMC CM5 CM5
4	TMC CM-5/512/ 512	30.40 66.00	NCSA USA/	Academic	36864 16384	TMC CM5 CM5
5	NEC SX-3/44R/ 4	23.20 26.00	NEC Fuchu Plant Japan/1990	Vendor	6400 830	NEC Vector SX3
6	NEC SX-3/44/ 4	20.00 22.00	Atmospheric Environment Service (AES) Canada/1991	Research Weather	6144 832	NEC Vector SX3
7	TMC CM-5/256/ 256	15.10 33.00	Naval Research Laboratory (NRL) USA/1992	Research	26112 12032	TMC CM5 CM5
8	Intel Delta/ 512	13.90 20.48	Caltech USA/	Academic	25000 7500	intel Paragon Paragon
9	Cray/SGI Y-MP C916/16256/ 16	13.70 15.24	Cray Research USA/	Vendor	10000 650	Cray Vector C90
10	Cray/SGI Y-MP C916/16256/ 16	13.70 15.24	DOE/Bettis Atomic Power Laboratory USA/1993	Research	10000 650	Cray Vector C90
11	Cray/SGI Y-MP C916/16256/ 16	13.70 15.24	DOE/Knolls Atomic Power Laboratory USA/1993	Research	10000 650	Cray Vector C90
12	Cray/SGI Y-MP C916/16128/ 16	13.70 15.24	ECMWF UK/1993	Research Weather	10000 650	Cray Vector C90
13	Cray/SGI Y-MP C916/161024/ 16	13.70 15.24	Government USA/1992	Classified	10000 650	Cray Vector C90
14	Cray/SGI Y-MP C916/161024/ 16	13.70 15.24	Government USA/1992	Classified	10000 650	Cray Vector C90

TOP500 im Juni 1993

Rank	Manufacturer Computer/Procs	R _{max} R _{peak}	Installation Site Country/Year	Inst. type Installation Area	Nmax Nhalf	Computer Family Computer Type
56	Fujitsu S600/20/ 1	4.01 5.00	Universitaet Aachen Germany/1991	Academic		Fujitsu VP VP2000
57	Fujitsu S600/20/ 1	4.01 5.00	Universitaet Karlsruhe Germany/1990	Academic		Fujitsu VP VP2000
60	TMC CM-5/64/ 64	3.80 8.19	GMD Germany/1993	Research	13056 6016	TMC CM5 CM5
65	Fujitsu S400/40/ 2	3.62 5.00	Universitaet Darmstadt Germany/1991	Academic		Fujitsu VP VP2000
66	Fujitsu S400/40/ 2	3.62 5.00	Universitaet Hannover / RRZN Germany/1991	Academic		Fujitsu VP VP2000
76	TMC CM-2/32k/ 1024	2.60 7.00	AMK Germany/1990	Classified		TMC CM2 CM2
98	Cray/SGI Y-MP8/832/ 8	2.14 2.67	Forschungszentrum Juelich (FZJ) Germany/1989	Research		Cray Vector YMP
102	Cray/SGI Y-MP8/864/ 8	2.14 2.67	Leibniz Rechenzentrum Germany/1992	Academic		Cray Vector YMP
142	TMC CM-5/32/ 32	1.90 4.10	Universitaet Wuppertal Germany/1992	Academic	9216 4096	TMC CM5 CM5
149	Intel XP/S5/ 66	1.90 3.30	Forschungszentrum Juelich (FZJ) Germany/1992	Research		intel Paragon Paragon
155	Intel XP/S5-32/ 66	1.90 3.30	Universitaet Stuttgart Germany/1992	Academic		intel Paragon Paragon
190	Cray/SGI CRAY-2s/4-128/ 4	1.41 1.95	DKRZ Germany/1988	Research Weather		Cray2/3 Cray 2
206	Cray/SGI CRAY-2/4-256/ 4	1.41 1.95	Universitaet Stuttgart Germany/1986	Academic		Cray2/3 Cray 2
218	TMC CM-2/16k/ 512	1.30 3.50	GMD Germany/1990	Research		TMC CM2 CM2
223	NEC SX-3/11/ 1	1.30 1.37	Universitaet Koeln Germany/1990	Academic	2816 192	NEC Vector SX3

TOP500 Deutschland im Juni 1993

Die TOP500-Liste Zusammenfassung

- ▶ Die Rechnerleistung wird mit einem numerischen Benchmark-Programm (LINPACK) evaluiert
- ▶ Die TOP500-Liste verzeichnet halbjährig die schnellsten Rechner weltweit
- ▶ Die schnellsten Rechner haben die Petaflops-Grenze durchbrochen
- ▶ IBM ist der bedeutendste Leistungslieferant
- ▶ Die USA dominieren den Markt
- ▶ Wichtigste Architektur: Cluster
- ▶ Wichtigste Vernetzung: Gigabit-Ethernet, Infiniband
- ▶ Fast nur noch CMOS off-the-shelf-Prozessoren
- ▶ Aktuelles Problem: Energiebedarf

▶ 98

Hochleistungsrechnen - © Thomas Ludwig

15.04.2010