

Mineure CalHau1

Evolution, marché et gestion des machines parallèles

Stéphane Vialle



Stephane.Vialle@centralesupelec.fr
<http://www.metz.supelec.fr/~vialle>

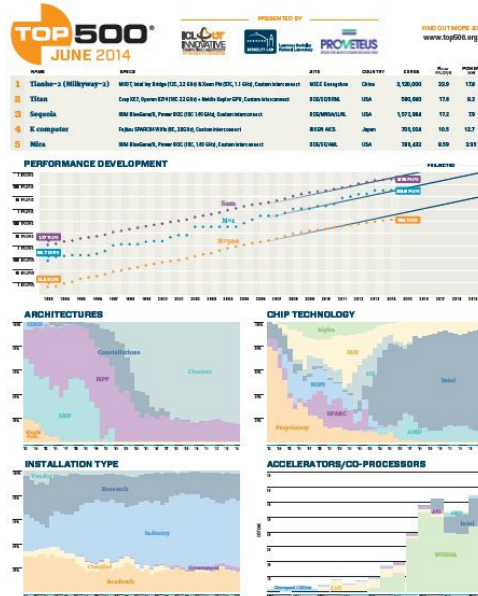
Evolution, marché et gestion des machines parallèles

- **Le « top500 » - marché des machines parallèles**
- Le « top500 green » - évolution du ratio Flop/Watt
- Cluster vs Supercalculateur : coût et coût total
- Gestion et utilisation d'une machine parallèle

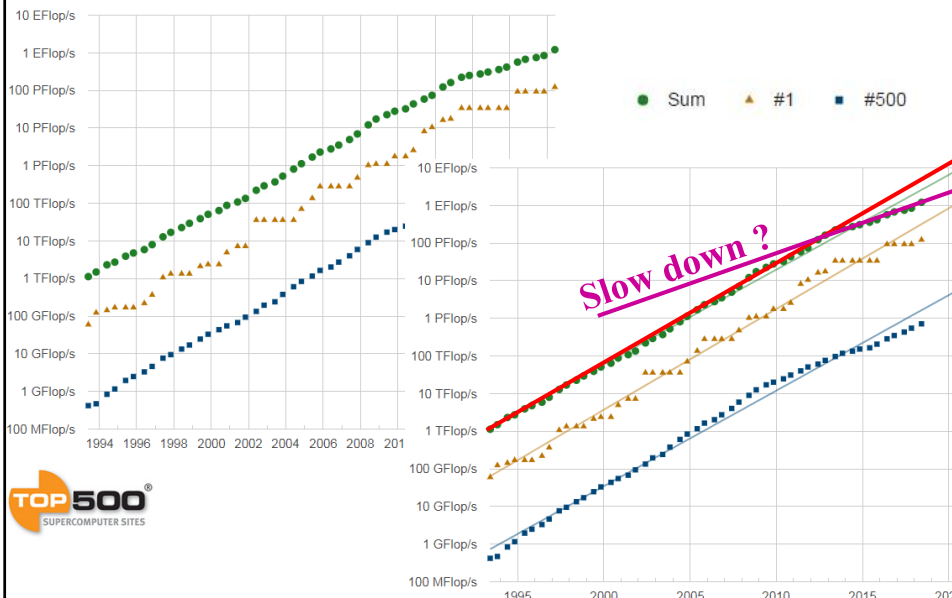
Le Top500 et le marché des super-calculateurs

<http://www.top500.org>

- Mis à jour en novembre et en juin, depuis 1993
- Liste des 500 plus grosses machines
- Des informations sur leurs architectures
- Des statistiques sur l'évolution des technologies et des marchés



Le Top500 Evolution de la puissance de calcul mondiale



TOP 500 [®] SUPERCOMPUTER SITES					TFlops	KWatt
T O P 1 0 n o v 2 0 1 5	1	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 3151P NUDT	3,120,000	33,862.7	54,902.4 17,808
	2	DOE/SC/Oak Ridge National Laboratory United States	Titan - Cray XK7, Opteron 6272 2.200GHz, Cray Gemini interconnect, NVIDIA K20x Cray Inc.	560,640	17,590.0	27,112.5 8,209
	3	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
	4	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
	5	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
	6	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
	7	Swiss National Supercomputing Centre (CSGS) Switzerland	Piz Daint - Cray XC30, Xeon E5-2670 8C 2.600GHz, Aries interconnect, NVIDIA K20x Cray Inc.	115,984	6,271.0	7,788.9 2,325
	8	HLRS - Höchstleistungsrechenzentrum Stuttgart Germany	Hazel Hen - Cray XC40, Xeon E5-2680v3 12C 2.5GHz, Aries interconnect Cray Inc.	185,088	5,640.2	7,403.5
	9	King Abdullah University of Science and Technology Saudi Arabia	Shaheen II - Cray XC40, Xeon E5-2698v3 16C 2.3GHz, Aries interconnect Cray Inc.	196,608	5,537.0	7,235.2 2,834
	10	Texas Advanced Computing Center/Univ. of Texas United States	Stampede - PowerEdge C8220, Xeon E5-2680 8C 2.700GHz, Infiniband FDR, Intel Xeon Phi SF10P Dell	462,462	5,168.1	8,520.1 4,510

TOP 500 [®] SUPERCOMPUTER SITES					TFlops	KWatt
T O P 1 0 n o v 2 0 1 6	1	National Supercomputing Center in Wuxi China	Sunway TaihuLight - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway NRCC	10,649,600	93,014.6	25,435.9 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 3151P NUDT	3,120,000	33,862.7	54,902.4 17,808
	3	DOE/SC/Oak Ridge National Laboratory United States	Titan - Cray XK7, Opteron 6272 2.200GHz, Cray Gemini interconnect, NVIDIA K20x Cray Inc.	560,640	17,590.0	27,112.5 8,209
	4	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
	5	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
	6	Joint Center for Advanced High Performance Computing Japan	Daiichi-3, PRIMERGY CX1640 M1, Intel Xeon Phi 7250 68C 1.4GHz, Intel Omni-Path Fujitsu	556,104	13,554.6	24,913.5 2,719
	7	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
	8	Swiss National Supercomputing Centre (CSGS) Switzerland	Piz Daint - Cray XC50, Xeon E5-2698v3 16C 2.6GHz, Aries interconnect, NVIDIA Tesla P100 Cray Inc.	206,720	9,779.0	15,988.0 1,312
	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

TOP 500 [®] SUPERCOMPUTER SITES				TFlops	KWatt	
T O P 1 0 0 v 2 0 1 7	1	Sunway TaihuLight - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway , NRCPC National Supercomputing Center in Wuxi China	10,649,600	93,014.6	25,435.9	15,371
	2	Tianhe-2 (MilkyWay-2) - TH-1B 12C 2.200GHz, TH Express-2000, Intel Xeon E5-2692 NUDT National Super Computer Center in Guangzhou China	3,120,000	33,862.7	54,902.4	17,808
	3	Piz Daint - Cray XC50, Xeon E5-2690v3 12C 2.6GHz, Aries interconnect, NVIDIA Tesla P100, Cray Inc. Swiss National Supercomputing Centre (CSCS) Switzerland	361,760	19,590.0	25,326.3	2,272
	4	Gyokou - ZettaScaler-2.2 HPC system, Xeon D-1571 16C 1.3GHz, Infiniband EDR, PEZY-SC2 700Mhz , ExaScaler Japan Agency for Marine-Earth Science and Technology Japan	19,860,000	19,135.8	28,192.0	1,350
	5	Titan - Cray XK7, Opteron 6274 16C 2.200GHz, Cray Gemini interconnect, NVIDIA K20x, Cray Inc. DOE/SC/ERDC - National Laboratory United States	560,640	17,590.0	27,112.5	8,209
	6	Sequoia - BlueGene/Q, Power BQC 16C 1.60 GHz, Custom , IBM DOE/NNSA/LNL United States	1,572,864	17,173.2	20,132.7	7,890
	7	Trinity - Cray XC40, Intel Xeon Phi 7250 68C 1.4GHz, Aries interconnect , Cray Inc. DOE/NNSA/ANL/SNL United States	979,968	14,137.3	43,902.6	3,844
	8	Cori - Cray XC40, Intel Xeon Phi 7250 68C 1.4GHz, Aries interconnect , Cray Inc. DOE/SC/ERDC/NERSC United States	622,336	14,014.7	27,880.7	3,939
	9	Oakforest-PACS - PRIMERGY CX1640 M4, Intel Xeon Phi 7250 68C 1.4GHz, Intel Omni-Path , Fujitsu Intel Center for Advanced High Performance Computing Japan	556,104	13,554.6	24,913.5	2,719
	10	K computer, SPARC64 VIIIfx 2.0GHz, Tofu interconnect , Fujitsu RIKEN Advanced Institute for Computational Science [AICS] Japan	705,024	10,510.0	11,280.4	12,660

TOP 500 [®] SUPERCOMPUTER SITES				TFlops	KWatt	
T O P 1 0 0 v 2 0 1 8	1	Summit - IBM Power System AC922, IBM POWER9 22C 3.07GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband , IBM ORNL - Oak Ridge National Laboratory United States	2,397,824	143,500.0	200,794.4	9,783
	2	Sinerg - IBM Power System S922LC, IBM POWER9 22C 3.1GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband , IBM / NVIDIA / Mellanox ORNL - Oak Ridge National Laboratory United States	1,572,400	94,640.0	125,712.3	7,438
	3	Sunway TaihuLight - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway , NRCPC National Supercomputing Center in Wuxi China	10,649,600	93,014.6	125,435.9	15,371
	4	Tianhe-2A - TH-1B, SSE Cluster, Intel Xeon E5-2692v2 12C 2.2GHz, TH Matrix-2000, NUDT National Super Computer Center in Guangzhou China	4,981,760	61,444.5	100,678.7	18,482
	5	Piz Daint - Cray XC50, Xeon E5-2690v3 12C 2.6GHz, Aries interconnect , NVIDIA Tesla P100, Cray Inc. Swiss National Supercomputing Centre (CSCS) Switzerland	361,760	21,230.0	27,154.3	2,384
	6	Trinity - Cray XC40, Xeon E5-2698v3 16C 2.3GHz, intel Xeon Phi 7250 68C 1.4GHz, Aries interconnect , Cray Inc. DOE/NNSA/ANL/SNL United States	979,968	20,158.7	41,461.2	7,578
	7	AI Bridging Cloud Infrastructure - PRIMERGY CX2570 M4, Xeon Gold 6148 20C 2.4GHz, NVIDIA Tesla V100 SXM2, Infiniband EDR , Fujitsu National Institute of Advanced Industrial Science and Technology [AIST] Japan	391,680	19,880.0	32,576.6	1,649
	8	SuperMUC-NG - ThinkSystem SD530, Xeon Platinum 8174 24C 3.1GHz, Intel Omni-Path , Lenovo Leibniz Rechenzentrum Germany	305,856	19,476.6	26,873.9	
	9	Titan - Cray XK7, Opteron 6274 16C 2.200GHz, Cray Gemini interconnect, NVIDIA K20x, Cray Inc. ORNL - Oak Ridge National Laboratory United States	560,640	17,590.0	27,112.5	8,209
	10	Sequoia - BlueGene/Q, Power BQC 16C 1.60 GHz, Custom , IBM DOE/NNSA/LNL United States	1,572,864	17,173.2	20,132.7	7,890

Accélérateur conçu en Chine spécialement pour upgrader le Tianhe-2!

Suite à l'embargo sur les Xeon-phi du gouvernement Américain...

Le Top500

N°1 en novembre 2018

143.5 Pflops ($\times 1.54$ vs previous n°1)

- 9 216 processors IBM POWER9 22C 3.07GHz
- 27 648 GPU Volta GV100
→ 2 282 544 « cores »
- interconnect: Dual-rail Mellanox EDR Infiniband

9.8 MWatt ($\times 0.64$ vs previous n°1)

Flops/Watt : $\times 2.4$

Summit Overview

OpenPOWER

Compute Node
2 x POWER9
6 x NVIDIA GV100
NVMe-compatible PCIe 1600 GB SSD

Components
IBM POWER9
• 22 Cores
• 4 Thread/core
• NVLink

NVIDIA GV100
• 7 TF
• 16 GB @ 0.9 TB/s
• NVLink

Compute Rack
18 Compute Servers
Warm water (70°F) direct-cooled components
RDHX for air-cooled components

Compute System
10.2 PB Total Memory
256 compute racks
4,668 compute nodes
Mellanox EDR Infiniband
200 PFLOPS
~13 MW

GPFS File System
250 PB storage
2.5 TB/s read, 2.5 TB/s write

25 GB/s EDR (8-2 ports)
512 GB DRAM (CC84)
96 GB HBM (32 Banks)
Coherent Shared Memory

39.7 TB Memory/rack
55 kW max power/rack

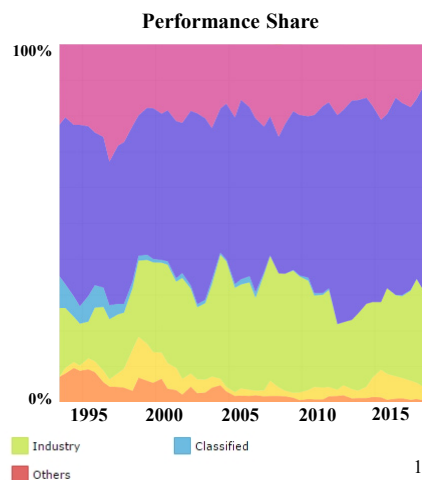
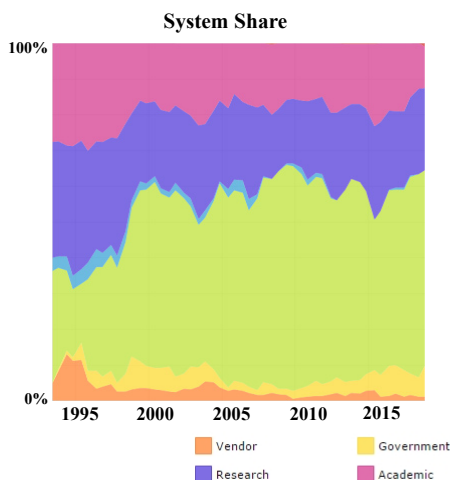
Le Top500

Evolution du marché



De + en + de supercalculateurs dans l'industrie

Mais des systèmes plus puissants dans la recherche



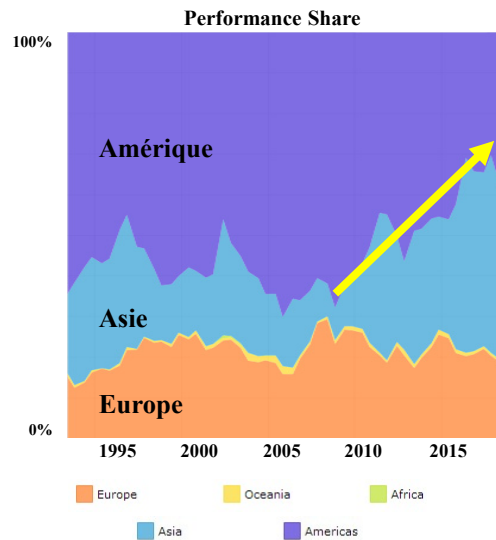
Le Top500

Evolution de la répartition géographique

L'Amérique du nord possède la plus grosse partie de la capacité de calcul mondiale... Mais a nettement diminué !

L'Europe se maintient depuis 2000

L'Asie se développe clairement (Chine + Japon) ...



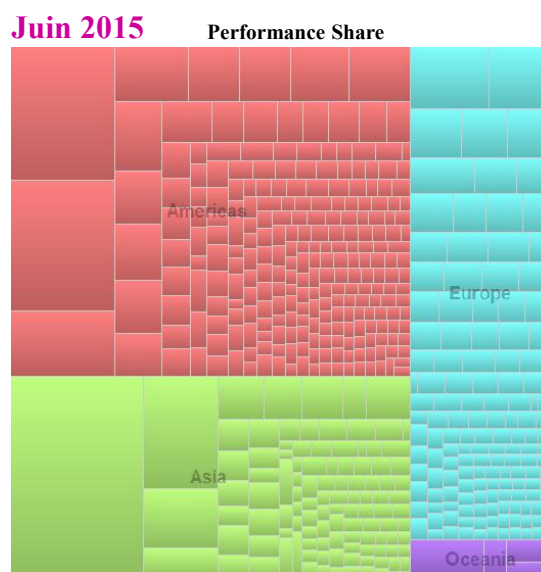
Le Top500

Evolution de la répartition géographique

L'Amérique du nord possède la plus grosse partie de la capacité de calcul mondiale, **et de très grosses machines**

L'Europe se maintient depuis 2000

L'Asie se développe (Chine + Japon), **et possède la plus grosse machine en juin 2015**



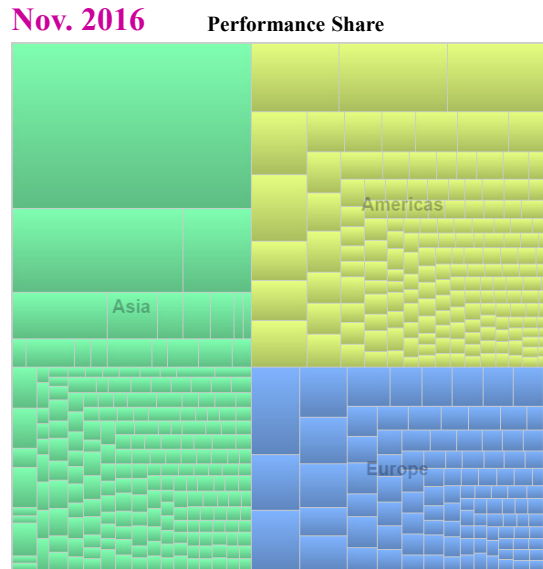
Le Top500

Evolution de la répartition géographique

L'Asie est passée en tête (Chine + Japon),
et possède la plus grosse machine en novembre 2016

L'Amérique du nord est 2^{ème} avec beaucoup de très grosses machines

L'Europe a un peu régressé, avec des machines un peu plus modestes



Le Top500

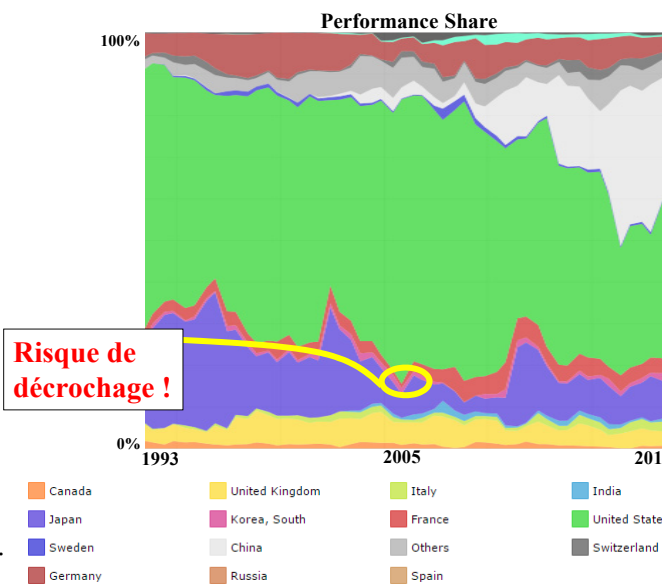
Evolution de la répartition géographique

La Chine et le Japon tirent toute l'Asie.

Les USA régressent sérieusement (en proportion)

Rmq : vers 2005 la France avait pris beaucoup de retard. Mais un plan stratégique de ré-équipement à amélioré la situation.

Risque de décrochage !



Le Top500

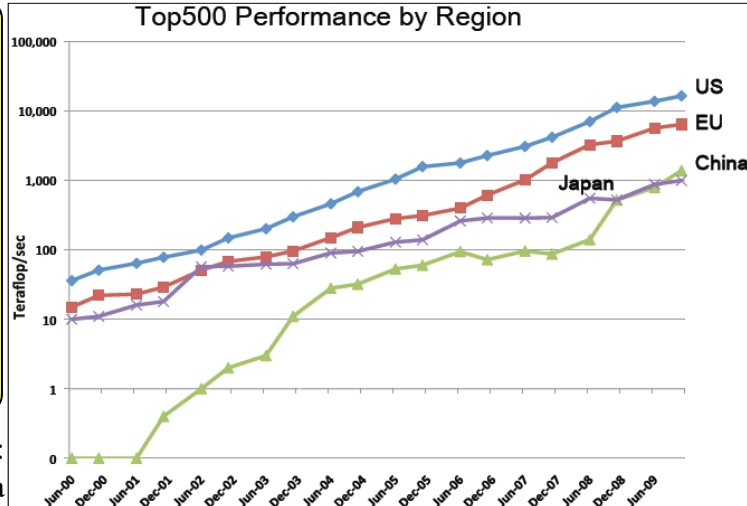
Evolution de la répartition géographique

La montée en puissance de la Chine en HPC a été prodigieuse!

La Chine produit aussi sa propre lignée de processeurs (type MIPS): « Loongson »

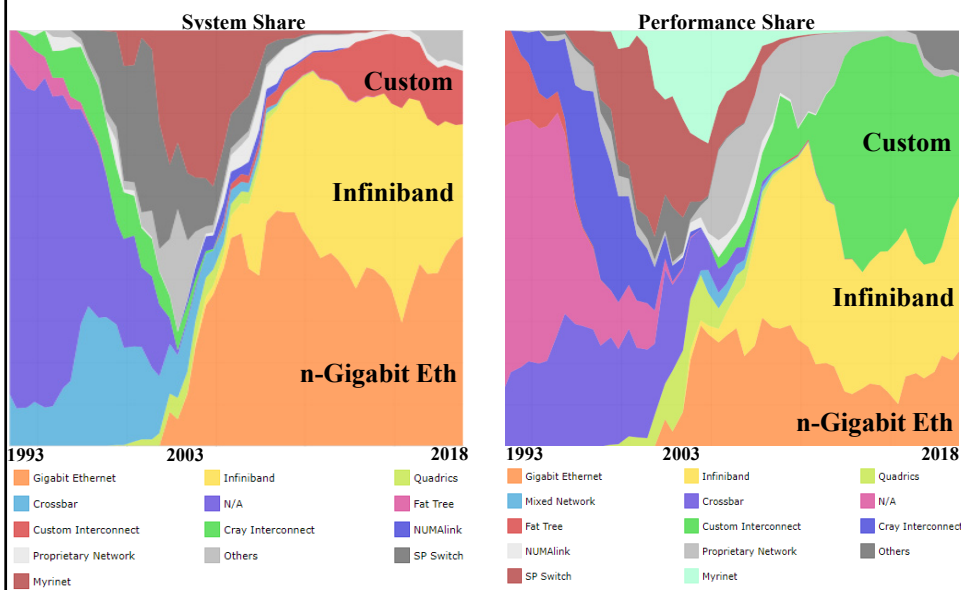


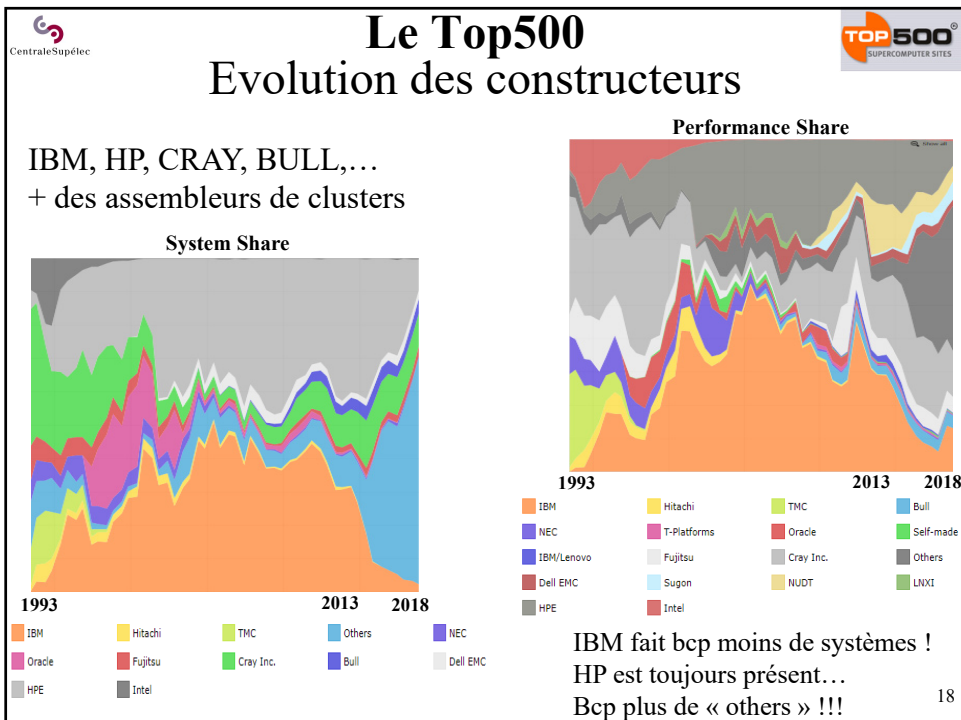
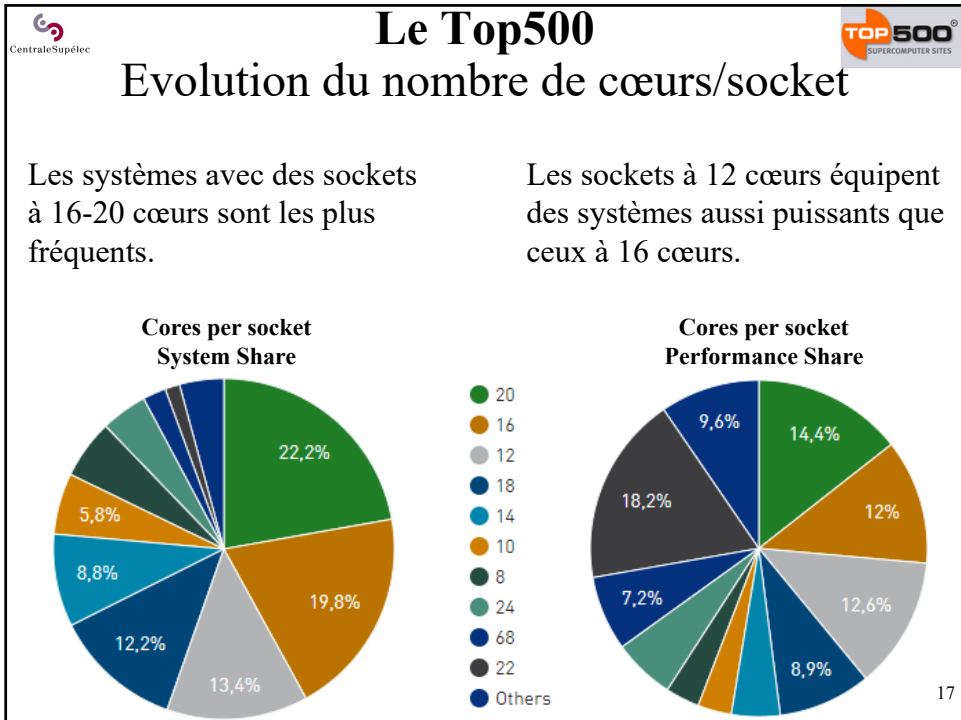
Auteur : Jack Dongara



Le Top500

Evolution des réseaux d'interconnexion





Evolution, marché et gestion des machines parallèles

- Le « top500 » - marché des machines parallèles
- **Le « top500 green » - évolution du ratio Flop/Watt**
- Cluster vs Supercalculateur : coût et coût total
- Gestion et utilisation d'une machine parallèle

Top 500 « vert » des supercalculateurs

<http://www.green500.org/>

THE
GREEN
500

STATISTICS SEARCH

LIST GROUPING DISPLAY

Ranking the World's Most
ENERGY-EFFICIENT SUPERCOMPUTERS



HOME ABOUT GREEN LISTS NEWS RESOURCES FAQ CONTACT

Environmentally Responsible Supercomputing

The Green500 provides rankings of the most energy-efficient supercomputers in the world. We raise awareness about power consumption, promote alternative total cost of ownership performance metrics, and ensure that supercomputers only simulate climate change and not create it.

SUPERMICR

Proud Sponsor of The Green500

Recent Green500 News

Top 500 « vert » des supercalculateurs

Le classement est très différent de celui du Top500 !

Janvier 2008

Green500 Rank	MFLOPS/W	Site*	Computer*	Total Power (kW)	TOP500 Rank*
1	357.23	Science and Technology Facilities Council - Daresbury Laboratory	Blue Gene/P Solution	31.10	121
2	352.25	Max-Planck-Gesellschaft MPI/IPP	Blue Gene/P Solution	62.20	40
3	346.95	IBM - Rochester	Blue Gene/P Solution	124.40	24
4	336.21	Forschungszentrum Juelich (FZJ)	Blue Gene/P Solution	497.60	2
5	310.93	Oak Ridge National Laboratory	Blue Gene/P Solution	70.47	41
6	210.56	Harvard University	eServer Blue Gene Solution	44.80	170
7	210.56	High Energy Accelerator Research Organization /KEK	eServer Blue Gene Solution	44.80	171
8	210.56	IBM - Almaden Research Center	eServer Blue Gene Solution	44.80	172
9	210.56	IBM Research	eServer Blue Gene Solution	44.80	173
10	210.56	IBM Thomas J. Watson Research Center	eServer Blue Gene Solution	44.80	174

Top 500 « vert » des supercalculateurs

Les architectures hétérogènes à base d'IBM Cell passent en tête.

Nov. 2008

Green500 Rank	MFLOPS/W	Site*	Computer*	Total Power (kW)	TOP500 Rank*
1	536.24	Interdisciplinary Centre for Mathematical and Computational Modelling, University of Warsaw	BladeCenter QS22 Cluster, 8i 4.0 Ghz, Infiniband, PowerXCell	34.63	220
2	530.33	Repsol YPF	BladeCenter QS22 Cluster, 8i 3.2 Ghz, Infiniband, PowerXCell	26.38	429
2	530.33	Repsol YPF	BladeCenter QS22 Cluster, 8i 3.2 Ghz, Infiniband, PowerXCell	26.38	430
2	530.33	Repsol YPF	BladeCenter QS22 Cluster, 8i 3.2 Ghz, Infiniband, PowerXCell	26.38	431
5	458.33	DOE/NNSA/LANL	BladeCenter QS22/LS21 Cluster, PowerXCell 8i 3.2 Ghz / Opteron DC 1.8 Ghz, Infiniband	138	41
5	458.33	IBM Poughkeepsie Benchmarking Center	BladeCenter QS22/LS21 Cluster, PowerXCell 8i 3.2 Ghz / Opteron DC 1.8 Ghz, Infiniband	138	42
7	444.94	DOE/NNSA/LANL	BladeCenter QS22/LS21 Cluster, PowerXCell 8i 3.2 Ghz / Opteron DC 1.8 Ghz, Voltaire Infiniband	2483.47	1
8	371.67	ASTRON/University Groningen	Blue Gene/P Solution	94.5	75
9	371.67	IBM - Rochester	Blue Gene/P Solution	126	56
9	371.67	RZG/Max-Planck-Gesellschaft MPI/IPP	Blue Gene/P Solution	126	57

Top 500 « vert » des supercalculateurs

Nov. 2009

Les « Cell » en tête, et les performances progressent

Green500 Rank	MFLOPS/W	Site*	Computer*	Total Power (kW)	TOP500 Rank*
1	722.98	Forschungszentrum Juelich (FZJ)	QPACE SFB TR Cluster, PowerXCell 8i, 3.2 GHz, 3D-Torus	59.49	110
1	722.98	Universitaet Regensburg	QPACE SFB TR Cluster, PowerXCell 8i, 3.2 GHz, 3D-Torus	59.49	111
1	722.98	Universitaet Wuppertal	QPACE SFB TR Cluster, PowerXCell 8i, 3.2 GHz, 3D-Torus	59.49	112
4	458.33	DOE/NNSA/LANL	BladeCenter QS22/LS21 Cluster, PowerXCell 8i, 3.2 Ghz / Opteron DC 1.8 GHz, Infiniband	276	29
4	458.33	IBM Poughkeepsie Benchmarking Center	BladeCenter QS22/LS21 Cluster, PowerXCell 8i, 3.2 Ghz / Opteron DC 1.8 GHz, Infiniband	138	78
6	444.25	DOE/NNSA/LANL	BladeCenter QS22/LS21 Cluster, PowerXCell 8i, 3.2 Ghz / Opteron DC 1.8 GHz, Infiniband	2345.5	2
7	428.91	National Astronomical Observatory of Japan	GRAPE-DR accelerator Cluster, Infiniband	51.2	445
8	379.24	National SuperComputer Center in Tianjin/NUDT	NUDT TH-1 Cluster, Xeon E5540/E5450, ATI Radeon HD 4870 2, Infiniband	1484.8	5
9	378.77	King Abdullah University of Science and Technology	Blue Gene/P Solution	504	18
9	378.77	EDF R&D	Blue Gene/P Solution	252	49

23

Top 500 « vert » des supercalculateurs

Nov. 2010

Le 1^{er} prototype BG/Q en tête, suivi de clusters de GPUs

Green500 Rank	MFLOPS/W	Site*	Computer*	Total Power (kW)
1	1684.20	IBM Thomas J. Watson Research Center	NNSA/SC Blue Gene/Q Prototype	38.80
2+	1448.03	National Astronomical Observatory of Japan	GRAPE-DR accelerator Cluster, Infiniband	24.59
2	958.35	GSIC Center, Tokyo Institute of Technology	HP ProLiant SL390s G7 Xeon 6C X5670, Nvidia GPU, Linux/Windows	1243.80
3	933.06	NCSA	Hybrid Cluster Core i3 2.93Ghz Dual Core, NVIDIA C2050, Infiniband	36.00
4	828.67	RIKEN Advanced Institute for Computational Science	K computer, SPARC64 VIIIx 2.0GHz, Tofu interconnect	57.96
5	773.38	Universitaet Wuppertal	QPACE SFB TR Cluster, PowerXCell 8i, 3.2 GHz, 3D-Torus	57.54
5	773.38	Universitaet Regensburg	QPACE SFB TR Cluster, PowerXCell 8i, 3.2 GHz, 3D-Torus	57.54
5	773.38	Forschungszentrum Juelich (FZJ)	QPACE SFB TR Cluster, PowerXCell 8i, 3.2 GHz, 3D-Torus	57.54
8	740.78	Universitaet Frankfurt	Supermicro Cluster, QC Opteron 2.1 GHz, ATI Radeon GPU, Infiniband	385.00
9	677.12	Georgia Institute of Technology	HP ProLiant SL390s G7 Xeon 6C X5660 2.8Ghz, Nvidia Fermi, Infiniband QDR	94.40
10	636.36	National Institute for Environmental Studies	GOSAT Research Computation Facility, nvidia	117.15

24

Top 500 « vert » des supercalculateurs

Jun 2011 Le 2nd prototype BG/Q en tête, suivi de clusters de GPUs

Green500 Rank	MFLOPS/W	Site*	Computer*	Total Power (kW)
1	2097.19	IBM Thomas J. Watson Research Center	NNSA/SC Blue Gene/Q Prototype 2	40.95
2	1684.20	IBM Thomas J. Watson Research Center	NNSA/SC Blue Gene/Q Prototype 1	38.80
3	1375.88	Nagasaki University	DEGIMA Cluster, Intel i5 Infiniband QDR ATI Radeon GPU	34.24
4	958.35	GSIC Center, Tokyo Institute of Technology	HP ProLiant SL390s G7 Xeon 6C X5670, Nvidia GPU, Linux/Windows	1243.80
5	891.88	CINECA / SCS - SuperComputing Solution	iDataPlex DX360M3, Xeon 2.4 Infiniband nVidia GPU	160.00
6	824.56	RIKEN Advanced Institute for Computational Science (AICS)	K computer, SPARC64 VIIIfx 2.0GHz, Tofu interconnect	9898.56
7	773.38	Forschungszentrum Juelich (FZJ)	QPACE SFB TR Cluster PowerXCell 8i, 3.2 GHz, 3D-Torus	57.54
8	773.38	Universitaet Regensburg	QPACE SFB TR Cluster PowerXCell 8i, 3.2 GHz, 3D-Torus	57.54
9	773.38	Universitaet Wuppertal	QPACE SFB TR Cluster PowerXCell 8i, 3.2 GHz, 3D-Torus	57.54
10	718.13	Universitaet Frankfurt	Supermicro Cluster, QC Opteron 2.1 GHz, ATI Radeon GPU, Infiniband	416.78

Top 500 « vert » des supercalculateurs

Jun 2012 Que des BG/Q dans le Top10 !!!

Green500 Rank	MFLOPS/W	Site*	Computer*	Total Power (kW)
1	2,100.88	DOE/NNSA/LLNL	BlueGene/Q, Power BQC 16C 1.60GHz, Custom	41.10
2	2,100.88	IBM Thomas J. Watson Research Center	BlueGene/Q, Power BQC 16C 1.60GHz, Custom	41.10
3	2,100.86	DOE/NNSA/LLNL	BlueGene/Q, Power BQC 16C 1.60GHz, Custom	82.20
4	2,100.86	DOE/NNSA/LLNL	BlueGene/Q, Power BQC 16C 1.60GHz, Custom	82.20
5	2,100.86	RIKEN Advanced Institute for Computational Science (AICS)	BlueGene/Q, Power BQC 16C 1.60GHz, Custom	82.20
6	2,100.86	University of Rochester	BlueGene/Q, Power BQC 16C 1.60GHz, Custom	82.20
7	2,100.86	IBM Thomas J. Watson Research Center	BlueGene/Q, Power BQC 16C 1.60 GHz, Custom	82.20
8	2,099.56	University of Edinburgh	BlueGene/Q, Power BQC 16C 1.60GHz, Custom	493.10
9	2,099.50	Science and Technology Facilities Council - Daresbury Laboratory	BlueGene/Q, Power BQC 16C 1.60GHz, Custom	575.30
10	2,099.46	Forschungszentrum Juelich (FZJ)	BlueGene/Q, Power BQC 16C 1.60GHz, Custom	657.50

L'année de l'architecture BlueGene d'IBM... mais ça ne dure pas.

Top 500 « vert » des supercalculateurs

Nov. 2012 Des machines avec accélérateurs devant les BG/Q

Green500 Rank	MFLOPS/W	Site*	Computer*	Total Power (kW)
1	2,499.44	National Institute for Computational Sciences/University of Tennessee	Beacon - Appro GreenBlade GB824M, Xeon E5-2670 8C 2.600GHz, Infiniband FDR, Intel Xeon Phi 5110P	44.89
2	2,351.10	King Abdulaziz City for Science and Technology	SANAM - Adtech ESC4000/FDR G2, Xeon E5-2650 8C 2.000GHz, Infiniband FDR, AMD FirePro S10000	179.15
3	2,142.77	DOE/SC/Oak Ridge National Laboratory	Titan - Cray XK7, Opteron 6274 16C 2.200GHz, Cray Gemini interconnect, NVIDIA K20x	8,209.00
4	2,121.71	Swiss Scientific Computing Center (CSCS)	Todi - Cray XK7, Opteron 6272 16C 2.100GHz, Cray Gemini interconnect, NVIDIA Tesla K20 Kepler	129.00
5	2,102.12	Forschungszentrum Juelich (FZJ)	JUQUEEN - BlueGene/Q, Power BQC 16C 1.600GHz, Custom Interconnect	1,970.00
6	2,101.39	Southern Ontario Smart Computing Innovation Consortium/University of Toronto	BQGdev - BlueGene/Q, Power BQC 16C 1.600GHz, Custom Interconnect	41.09
7	2,101.39	DOE/NNSA/LLNL	rzuseq - BlueGene/Q, Power BQC 16C 1.60GHz, Custom	41.09
8	2,101.39	IBM Thomas J. Watson Research Center	BlueGene/Q, Power BQC 16C 1.80GHz, Custom	41.09
9	2,101.12	IBM Thomas J. Watson Research Center	BlueGene/Q, Power BQC 16C 1.80 GHz, Custom	82.19
10	2,101.12	Ecole Polytechnique Federale de Lausanne	CADMOS BG/Q - BlueGene/Q, Power BQC 16C 1.600GHz, Custom Interconnect	82.19

Top 500 « vert » des supercalculateurs

Nov. 2013 Que des machines avec accélérateurs GPU, et flop/W ↗ ↗

Green500 Rank	MFLOPS/W	Site*	Computer*	Total Power (kW)
1	4,503.17	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	27.78
2	3,631.00	Wilkes - Dell T800 Cluster	Intel Xeon E5-2630v2 6C 2.600GHz, Infiniband FDR, NVIDIA K20	52.62
3	3,130.95	HA-PACS TCA - Cray 3623G4-SM Cluster	Intel Xeon E5-2680v2 10C 2.800GHz, Infiniband QDR, NVIDIA K20x	78.77
4	2,785.91	Swiss National Supercomputing Centre (CSCS)	BizData - Cray XC30, Xeon E5-2670 8C 2.600GHz, Aries interconnect, NVIDIA K20x Level 3 measurement data available	1,753.66
5	3,130.95	ROMEO HPC Center - Champagne-Ardenne	romeo - Bull R421-E3 Cluster, Intel Xeon E5-2650v2 8C 2.600GHz, Infiniband FDR, NVIDIA K20x	81.41
6	3,068.71	GSIC Center, Tokyo Institute of Technology	TSUBAME 2.5 - Cluster Platform SL 300s-C7, Xeon X5670 6C 2.930GHz, Infiniband QDR, NVIDIA K20x	922.54
7	2,702.16	University of Arizona	iDataPlay DX360M4, Intel Xeon E5-2650v2 8C 2.600GHz, Infiniband FDR1, NVIDIA K20x	53.62
8	2,629.10	Max-Planck-Gesellschaft MPI/IPP	iDataPlay DY360M4, Intel Xeon E5-2680v2 10C 2.800GHz, Infiniband, NVIDIA K20x	269.94
9	2,629.10	Financial Institution	iDataPlay DY360M4, Intel Xeon E5-2680v2 10C 2.800GHz, Infiniband, NVIDIA K20x	55.62
10	2,358.69	CSIRO	CSIRO GPU Cluster - Nils G15 3GPU, Xeon E5-2650 8C 2.000GHz, Infiniband FDR, Nvidia K20m	71.01

Arrivée massive du GPU...
...et ça va durer !

Top 500 « vert » des supercalculateurs

Jun 2014

Que des machines à base d'accélérateurs GPU

Green500 Rank	MFLOPS/W	Site*	Computer*	Total Power (kW)
1	4,389.82	GSIC Center, Tokyo Institute of Technology	TSUBAME-KFC - LX 1U-4GPU/104R... 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 3622G4-SM Cluster, Intel Xeon E5-2680v2 10C 2.800GHz, Infiniband QDR, NVIDIA K20x	78.77
4	3,459.46	SURFsara	Cartesius Accelerator Island - Bullx PE15 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 - iDataPlay-DX360M4, Intel Xeon E5-2680v2 10C 2.8GHz, Infiniband FDR, NVIDIA K20x	1,067.49
10	2,678.41	Financial Institution	iDataPlay-DX360M4, Intel Xeon E5-2680v2 10C 2.800GHz, Infiniband, NVIDIA K20x	54.60

29

Top 500 « vert » des supercalculateurs

Nov. 2014

« Presque que » des machines à base d'accélérateurs GPU

Green500 Rank	MFLOPS/W	Site*	Computer*	Total Power (kW)
1	5,271.81	GSI Helmholtz Center	L-CSC - ASUS ESC4000 FDR/C2S, Intel Xeon E5-2690v2 10C 3GHz, Infiniband FDR, AMD FirePro S9150 Level 1 measurement data available	57.15
2	4,945.63	High Energy Accelerator Research Organization /KEK	Suilren - ExaScaler 32U P56SC Cluster, Intel Xeon E5-2660v2 10C 2.2GHz, Infiniband FDR, PEZY-SC CPU, 10 cores/socket	37.83
3	4,447.58	GSIC Center, Tokyo Institute of Technology	TSUBAME-KFC - LX 1U-4GPU/104R... Cluster, Intel Xeon E5-2620v2 6C 2.100GHz, Infiniband FDR, NVIDIA K20x	35.39
4	3,962.73	Cray Inc.	Storm1 - Cray CS Storm, Intel Xeon E5-2660v2 10C 2.2GHz, Infiniband FDR, Nvidia K40m Level 3 measurement data available GPU Nvidia	44.54
5	3,631.70	Cambridge University	Wilkes - Dell T620 Cluster, Intel Xeon E5-2630v2 6C 2.600GHz, Infiniband FDR, NVIDIA K20	52.62
6	3,543.32	Financial Institution	iDataPlay-DX360M4, Intel Xeon E5-2680v2 10C 2.800GHz, Infiniband, NVIDIA K20x	54.60
7	3,517.84	Center for Computational Sciences, University of Tsukuba	HA-PACS TCA - Cray CS300 Cluster, Intel Xeon E5-2680v2 10C 2.800GHz, Infiniband QDR, NVIDIA K20x	78.77
8	3,459.46	SURFsara	Cartesius Accelerator Island - Bullx PE15 cluster, Intel Xeon E5-2450v2 8C 2.5GHz, InfiniBand 4x FDR, Nvidia K40m	44.40
9	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
10	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

30

Top 500 « vert » des supercalculateurs

Juin 2015

3 machines sans accélérateurs en tête !

Green500 Rank	MFLOPS/W	Site*	Computer*	Total Power (kW)
1	7,031.58	RIKEN	Shoubu - ExaScaler-1.4 80Brick, Xeon E5-2618Lv3 8C 2.3GHz, Infiniband FDR PEZY-SC CPU, 8 cores/socket, 2.3GHz	50.32
2	6,842.31	High Energy Accelerator Research Organization /KEK	Suiren Blue - ExaScaler-1.4 16Brick, Xeon E5-2618Lv3 8C 2.3GHz, Infiniband PEZY-SC CPU, 8 cores/socket, 2.3GHz	28.25
3	6,217.04	High Energy Accelerator Research Organization /KEK	Suiren - ExaScaler 32U256SC Cluster, Intel Xeon E5-2660v2 10C 2.2GHz, Infiniband FDR PEZY-SC CPU, 10 cores/socket, 2.2GHz	32.59
4	5,271.81	GSI Helmholtz Center	ASUS ESC4000 FDR/G2S, Intel Xeon E5-2690v2 10C 3GHz, Infiniband FDR AMD FirePro S9150	57.15
5	4,257.88	GSIC Center, Tokyo Institute of Technology	TSUBAME-KFC - LX 1U-4GPU/104Re-1G Cluster, Intel Xeon E5-2620v2 6C 2.1GHz, Infiniband FDR NVIDIA K20x	39.83
6	4,112.11	Stanford Research Computing Center	XStream - Cray CS-Storm, Intel Xeon E5-2680v2 10C 2.8GHz, Infiniband FDR Nvidia K80	190.00
7	3,962.73	Cray Inc.	Storm1 - Cray CS-Storm, Intel Xeon E5-2660v2 10C 2.2GHz, Infiniband FDR Nvidia K40m	44.54
8	3,631.70	Cambridge University	Wilkes - Dell T620 Cluster, Intel Xeon E5-2630v2 6C 2.6GHz, Infiniband FDR NVIDIA K20	52.62
9	3,614.71	TU Dresden, ZIH	Taurus GPUs - Bull-bulx R400, Xeon E5-2680v3 12C 2.5GHz, Infiniband FDR Nvidia K80	58.01
10	3,543.32	Financial Institution	DataPlex-DX360M4, Intel Xeon E5-2680v2 10C 2.8GHz, Infiniband, NVIDIA K20x	54.60

Top 500 « vert » des supercalculateurs

Nov. 2015

1 seule machine sans accélérateurs en tête

Green500 Rank	MFLOPS/W	Site*	Computer*	Total Power (kW)
1	7,031.58	Institute of Physical and Chemical Research (RIKEN)	Shoubu - ExaScaler-1.4 80Brick, Xeon E5-2618Lv3 8C 2.3GHz, Infiniband FDR PEZY-SC CPU, 8 cores/socket, 2.3GHz	50.32
2	5,331.79	GSIC Center, Tokyo Institute of Technology	TSUBAME-KFC/DL - LX 1U-4GPU/104Re-1G Cluster, Intel Xeon E5-2620v2 6C 2.1GHz, Infiniband FDR NVIDIA Tesla K80	51.13
3	5,271.81	GSI Helmholtz Center	ASUS ESC4000 FDR/G2S, Intel Xeon E5-2690v2 10C 3GHz, Infiniband FDR AMD FirePro S9150	57.15
4	4,778.46	Institute of Modern Physics (IMP), Chinese Academy of Sciences	Sugon Cluster W7801, Xeon E5-2640v3 8C 2.6GHz, Infiniband QDR, NVIDIA Tesla K80 GPU Nvidia	65.00
5	4,112.11	Stanford Research Computing Center	XStream - Cray CS-Storm, Intel Xeon E5-2680v2 10C 2.8GHz, Infiniband FDR Nvidia K80	190.00
6	3,856.90	IT Company	Inspur TS10000 HPC Server, Xeon E5-2620v3 6C 2.4GHz, 10G Ethernet NVIDIA Tesla K40	58.00
7	3,775.45	Internet Service	Inspur TS10000 HPC Server, Intel Xeon E5-2620v2 6C 2.1GHz, 10G Ethernet NVIDIA Tesla K40	110.00
8	3,775.45	Internet Service	Inspur TS10000 HPC Server, Intel Xeon E5-2620v2 6C 2.1GHz, 10G Ethernet NVIDIA Tesla K40	110.00
9	3,775.45	Internet Service	Inspur TS10000 HPC Server, Intel Xeon E5-2620v2 6C 2.1GHz, 10G Ethernet NVIDIA Tesla K40	110.00
10	3,775.45	Internet Service	Inspur TS10000 HPC Server, Intel Xeon E5-2620v2 6C 2.1GHz, 10G Ethernet NVIDIA Tesla K40	32 110.00

CentraleSupélec **Top 500 « vert » des supercalculateurs**

Nov. 2016 Nouveaux GPU NVIDIA en tête, puis des Xeon-phi KnL

Green500 Rank	MFLOPS/W	Site	System	Total Power(kW)
1	9462.1	NVIDIA Corporation	NVIDIA DGX-1, Xeon E5-2698v4 20C 2.2GHz, Infiniband EDR, NVIDIA Tesla P100	349.5
2	7453.5	Swiss National Supercomputing Centre (CSCS)	Cray XC50, Xeon E5-2690v3 12C 2.6GHz Aries interconnect, NVIDIA Tesla P100	1312
3	6673.8	Advanced Center for Computing and Communication, RIKEN	ZettaScaler-1.6, Xeon E5-2618Lv3 8C 2.3GHz, Infiniband FDR, PEZY-SC2	150.0
4	6055.0	Supercomputing Center in Wuxi	Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway	15371
5	4985.7	Technology Solutions GmbH	PRIMERGY CX1640 M1, Intel Xeon Phi 7210 64C 1.3GHz, Intel Omni-Path	77
6	4985.7	Joint Center for Advanced High Performance Computing	PRIMERGY CX1640 M1, Intel Xeon Phi 7250 68C 1.4GHz, Intel Omni-Path	2718.7
7	4688.0	DOE/SC/Argonne National Laboratory	Cray XC40, Intel Xeon Phi 7230 64C 1.3GHz, Aries interconnect	1087
8	4112.1	Stanford Research Computing Center	Cray CS-Storm, Intel Xeon E5-2680v2 10C 2.8GHz, Infiniband FDR, Nvidia K80	190
9	4086.8	Academic Center for Computing and Media Studies (ACCMS), Kyoto University	Cray XC40, Intel Xeon Phi 7250 68C 1.4GHz, Aries interconnect	748.1
10	3836.6	Thomas Jefferson National Accelerator Facility	KOI Cluster, Intel Xeon Phi 7230 64C 1.3GHz, Intel Omni-Path	111

33

*6/10 sont des Xeon-Phi...
...mais ca ne dure pas !*

CPU: 8 cores/socket, 2.3GHz

CentraleSupélec **Top 500 « vert » des supercalculateurs**

Nov. 2017

TOP500 Rank	Rank	System	Cores	Rmax (TFlop/s)	Power (kW)	Power Efficiency (GFlops/watt)
1	259	Shoubu system B - Zettai, Xeon D-1571 16C 1.3GHz, Infiniband EDR, PEZY-SC2 PEZY Computing / Exascaler Inc. Advanced Center for Computing and Communication, RIKEN Japan	794,400	842.0	50	17.009
2	307	Suiren2 - ZettaScaler, Xeon D-1571 16C 1.3GHz, Infiniband EDR, PEZY-SC2 PEZY Computing / Exascaler Inc. High Energy Accelerator Research Organization / KEK Japan	762,624	788.2	47	16.759
3	276	Sakura - ZettaScaler, Xeon E5-2618Lv3 8C 2.3GHz, Infiniband EDR, PEZY-SC2 PEZY Computing / Exascaler Inc. PEZY Computing K.K. Japan	794,400	824.7	50	16.657
4	149	DGX SaturnV Volta - NVIDIA DGX-1, 2698v4 20C 2.2GHz, Infiniband EDR, NVIDIA Tesla V100 , Nvidia Nvidia Corporation United States	22,440	1,070.0	97	15.113
5	4	Gyokoku - ZettaScaler, Xeon D-1571 16C 1.3GHz, Infiniband EDR, PEZY-SC2 PEZY Computing / Exascaler Inc. Japan Agency for Marine Science and Technology Japan	19,860,000	19,135.8	1,350	14.173
6	13	TSUBAME3.0 - SGI ICE XA, Xeon E5-2698v4 14C 2.4GHz, Intel Omni-Path, NVIDIA Tesla P100 , HPE GSEC Center, Tokyo Institute of Technology Japan	135,828	8,125.0	792	13.704
7	195	AIST AI Cloud - NEC 4U-8B, Xeon E5-2698v4 10C 1.8GHz, Infiniband EDR, NVIDIA Tesla P100 , NEC National Institute of Advanced Industrial Science and Technology Japan	23,400	961.0	76	12.681
8	419	RAIDEN GPU subsystem - NVIDIA DGX-1, Xeon E5-2698v4 20C 2.2GHz, Infiniband EDR, NVIDIA Tesla P100 , Center for Advanced Intelligence Project, AIST Japan	11,712	635.1	60	10.603
9	115	WRites-2 - Dell EMC, Xeon E5-2698v4 12C 2.2GHz, Infiniband EDR, NVIDIA Tesla P100 , University of Cambridge United Kingdom	21,240	1,193.0	114	10.428
10	3	Plz Daint - Dell EMC, Xeon E5-2698v4 12C 2.6GHz, Aries interconnect, NVIDIA Tesla P100 , Swiss National Supercomputing Centre (CSCS) Switzerland	361,760	19,590.0	2,272	10.398

34

Pb d'énergie au Japon (suite de Fukushima)
→ 7 du Top10 vert au Japon !
→ 4 du Top10 vert en techno japonaise
ZettaScaler-PEZY-SC2

Et des machines à GPU P100

Top 500 « vert » des supercalculateurs

Nov. 2018

Pb d'énergie au Japon (suite de Fukushima)

→ 4 du Top10 vert au Japon, mais 1 en ZettaScaler-PEZY-SC2

8 du Top10 en GPU NVIDIA (V100 et P100)

Rank	TOP500 Rank	System	Cores	Rmax (TFlop/s)	Power (kW)	Power Efficiency (GFlops/watts)
1	375	Shoubu system B - ZettaScaler-PEZY-SC2, Xeon D-1571 16C 1.3GHz, Infiniband EDR Exascaler Inc. Advanced Center for Computing and Communication, RIKEN Japan	953,280	1,063.3	60	17.604
2	374	96X SaturnV Volta - NVIDIA Tesla V100, ES-2698v4 20C 2.2GHz, Infiniband EDR NVIDIA Corporation United States	22,440	1,070.0	97	15.113
3	1	Summit - IBM Power System AC922, IBM POWER9 22C 3.07GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband DOE/SC/Oak Ridge National Laboratory United States	2,397,824	143,500.0	9,783	14.668
4	7	AI Bridging Cloud Infrastructure (ABC) - PRIMERGY CX2570 M4, Xeon Gold 6148 20C 2.4GHz, NVIDIA Tesla V100 5XM2, Infiniband EDR, Fujitsu National Institute of Advanced Industrial Science and Technology (AIST) Japan	391,680	19,880.0	1,649	14.423
5	22	TSUBAME3.0 - SGI ICE XA, 13V-5XM2, Xeon E5-2680v4 14C 2.4GHz, Intel Omni-Path, NVIDIA Tesla P100-8XM2, HPE GSIC Center, Tokyo Institute of Technology Japan	135,828	8,125.0	792	13.704
6	2	Sierra - NVIDIA Volta GV100, P22LC, IBM POWER9 22C 3.15GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband, IBM Lambda 13 DOE/NNSA/LLNL United States	1,572,480	94,640.0	7,438	12.723
7	444	AIST AI Cloud - NEC A-1800 server, Xeon ES-2630Lv4 10C 1.8GHz, Infiniband EDR, NVIDIA Tesla P100-8XM2, NEC National Institute of Advanced Industrial Science and Technology Japan	23,400	961.0	76	12.681
8	411	MareNostrum P9 CTE - IBM Power System AC922, IBM POWER9 22C, Dual-rail Mellanox EDR Infiniband, NVIDIA Tesla V100, IBM National Institute of Advanced Industrial Science and Technology Spain	19,440	1,018.0	86	11.865
9	38	Advanced Computing System (PreE) - Sugon TCR600, Hygon Dhyana 32C 2GHz, Deep Computing Processor, 200Gb 6D-Torus, Sugon Sugon China	163,840	4,325.0	380	11.382
10	20	Taiwania 2 - QCT QuantaGrid D520, Xeon Gold 6148 18C 3GHz, Mellanox InfiniBand EDR, NVIDIA Tesla V100 5XM2, Quanta Computer / Taiwan National Center for High Performance Computing Taiwan	170,352	9,000.0	798	11.285

35

Top 500 « vert » des supercalculateurs

Dans le Top10 vert :

- Les Technologies des accélérateurs matériels « Cell » et « Xeon-phi » ont tenu quelques années...
- Les architecture IBM Blue Gene ont tenu aussi quelques années...
- **Les GPU sont bien installés depuis 10 ans !**
(dominé par NVIDIA)
- La technologie ZettaScaler – PEZY-SC (Japan) est à surveiller aussi.
- Le Japon a des soucis critique de rendement Flop/Watt !

36

Evolution, marché et gestion des machines parallèles

- Le « top500 » - marché des machines parallèles
- Le « top500 green » - évolution du ratio Flop/Watt
- **Cluster vs Supercalculateur : coût et coût total**
- Gestion et utilisation d'une machine parallèle

37

Cluster vs Supercalculateur : coût et coût total Apparition d'un cluster de PCs...

1994 : 1er cluster de PCs, appelé "Beowulf"

- En 1994, T. Sterling et D. Becker (CESDIS) assemblent un "cluster" avec leurs machines :
 - 16 mono-processeurs DX4 (intel-486)
 - réseau Ethernet (10Mb/s)
 - configuration Linux + envoi de messages (PVM/MPI)



- **C'est un succès : très bonnes performances et pas cher !**

- Les "clusters" vont commencer à se répandre ...

Un cluster "Beowulf" était constitué de machines standards et d'un réseau standard, simplement configurés pour fonctionner ensemble

Cluster vs Supercalculateur : coût et coût total Les premiers « gros » clusters de PCs

1998/99 : cluster expérimental de 225 PC à Grenoble

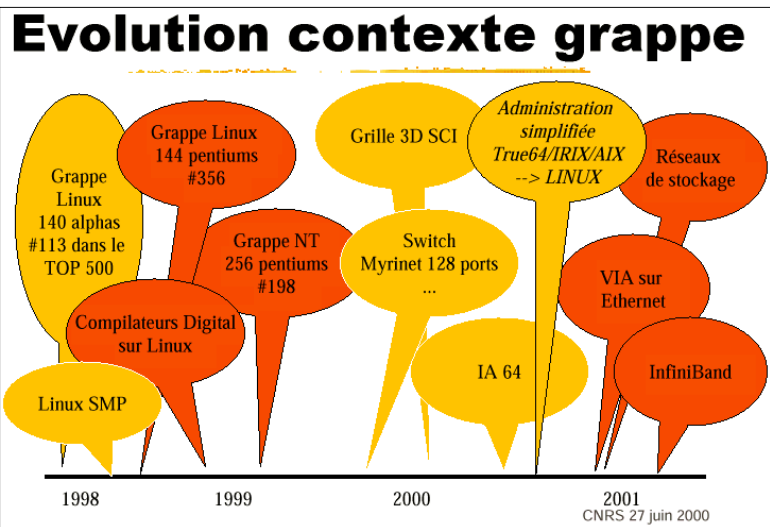
Rank 385 in
Top500

- Configuration d'intranet standard dans une entreprise.
- Travail : surtout du développement de couches systèmes (ingénieurs)
ex : reboot complet et rapide, update config rapide, ...
- Puis de la recherche algorithmique (chercheurs)



Cluster vs Supercalculateur : coût et coût total Les clusters viennent à maturité

Les clusters « percent » définitivement autour de l'an 2000 :



Cluster vs Supercalculateur : coût et coût total Coût réel des clusters ... ?



Cluster Systems Are Not Cheaper

CRAY critique les clusters, et défend les super-calculateurs

"A lot of people have argued that cluster systems are cheaper than semi-custom systems such as the X1. The fact of the matter is, they're not cheaper. When we did our total life-cycle cost analysis, we included a lot of factors, such as the facility costs, the power consumption, the support costs in terms of system staff, operations staff, and when all of those elements are included, it turns out that clusters are fairly expensive -- a lot more expensive than people would think, especially when you look at the relatively low level of sustained performance that you get out of the processors on a cluster system, and compare that with the higher levels of sustained performance on the Cray X1."

Paul Muzio, Support Infrastructure Director, Army High Performance Computing Research Center

Cluster vs Supercalculateur : coût et coût total Où mettre l'argent ?

Bilan des Clusters (ou « grappes ») :

- Très répandus, et les plus grosses machines sont des clusters.
- Les grands constructeurs proposent tous des clusters (ex: IBM).
- Configurations types :
 - Clusters de PCs en Gigabit-Ethernet (« Beowulf »)
 - Clusters de PCs à réseaux rapides (*InfiniBand*)
 - Clusters de (big) nœuds à réseau Cray (SuperCalculateur *Cray*)

• LA question récurrente : « que faire à prix constant ? »

- plus de nœuds ?
- plus de processeurs par nœuds ?
- meilleur réseau d'interconnexion ?

Pas de réponse
standard !

• LA question tabou : « coût total d'un cluster de PC ? »

Evolution, marché et gestion des machines parallèles

- Le « top500 » - marché des machines parallèles
- Le « top500 green » - évolution du ratio Flop/Watt
- Cluster vs Supercalculateur : coût et coût total
- **Gestion et utilisation d'une machine parallèle**

43

Gestion et utilisation d'une machine parallèle

Configuration classique :

- Cluster derrière des machines frontales
- Comptes spécifiques, home directories spécifiques, transfert de fichiers explicite
- Compilation sur une frontale
- Exécutions en mode 'batch', sans interactivité
- Exécutions interactives très limitées
- Files de jobs de priorités différentes, fonctions de :
 - la durée du job
 - le nombre de nœuds demandés
 - la priorité de l'utilisateur et de ses projets, son quota d'heures
 - parfois : les exécutions récentes du même utilisateur

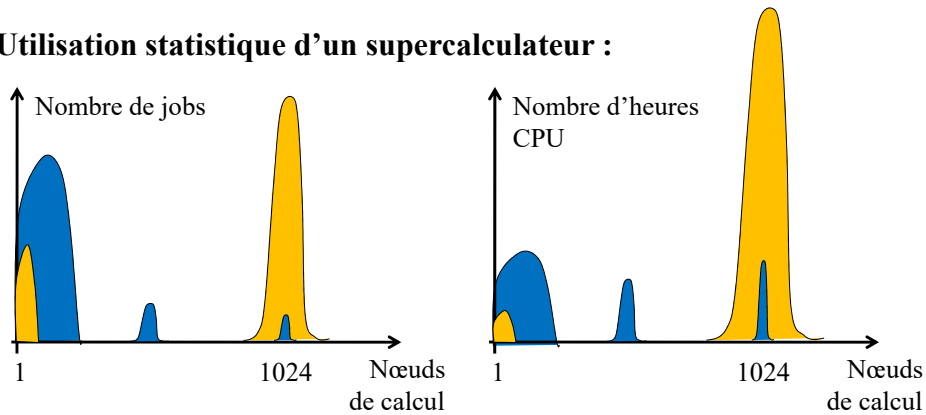
Attention :
peut dissuader
beaucoup
d'utilisateurs
non-habitués !



Trop de sécurité ou de contraintes d'équité peut tuer une communauté d'utilisateurs!

Gestion et utilisation d'une machine parallèle

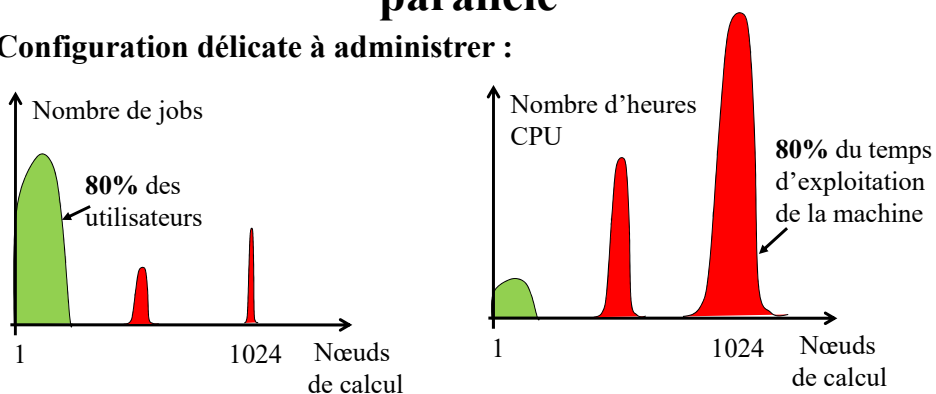
Utilisation statistique d'un supercalculateur :



- Un site où les utilisateurs sont familiers avec le calcul (massivement) parallèle
- Un site où les utilisateurs ne sont pas familiers avec le calcul parallèle, sauf quelques uns

Gestion et utilisation d'une machine parallèle

Configuration délicate à administrer :

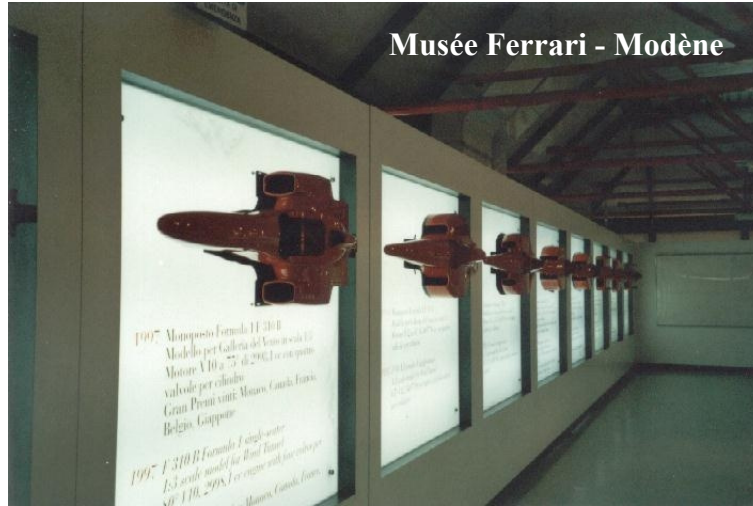


Deux communautés avec des besoins de configuration différents/opposés

- Majorité d'utilisateurs 'quasi-séquentiels' qui sont très minoritaire en % de temps d'exploitation
- Minorité d'utilisateurs 'massivement parallèles' qui sont très majoritaire en % du temps d'exploitation

Autres impact du HPC...

On distingue les années où la simulation numérique intensive permet de créer des carrosseries plus complexes ☺



47

Evolution, marché et gestion
des machines parallèles

FIN

48