

CentraleSupélec

Mineure CalHau1

## Evolution, marché et gestion des machines parallèles

**Stéphane Vialle**  
 Sciences et Technologies de l'Information et de la Communication (STIC)

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

---

---

---

---

---

---

---

---

---

---

CentraleSupélec

## 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

2

---

---

---

---

---

---

---

---

---

---

CentraleSupélec

## 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

---

---

---

---

---

---

---

---

---

---



**TOP 500**

**T  
O  
P  
1  
0  
0  
n  
o  
v  
e  
m  
b  
r  
e  
2  
0  
1  
7**

Rank	System	Country	Processors	TFlops	KWatt
1	Sumway TaihuLight - Sumway MPP, Sumway SW26010 2007	China	1,435Hz, TH Equipes, Xeon ES-2692	10,649,600	93,014.6
2	Tianhe-2 (MilkyWay-2) - TH Equipes, Xeon ES-2692	China	120 2.20GHz, TH Equipes, Xeon Phi 7205	3,120,000	33,862.7
3	Piz Daint - Cray XC260, Intel Xeon E5-2690 v3 120 2.6GHz, Aries	Switzerland	interconnect: NVIDIA Tesla P100	361,760	19,590.0
4	Gyokkou - ZettaScale-2 HPC system, Xeon D-1571 140 1.30GHz	Japan	interconnect: NVIDIA K80	19,860,000	19,135.8
5	Titan - Cray XC260, NVIDIA K80 120 2.20GHz, Cray Gemini	United States	interconnect: NVIDIA K80	560,640	17,590.0
6	Sequoia - BlueGene/Q, Power BQC 140 1.60 GHz, Custom	United States	interconnect: IBM	1,572,864	17,173.2
7	Trinity - Cray XC260, Intel Xeon Phi 7250 680 1.4GHz, Aries	United States	interconnect: Cray Inc.	979,968	14,137.3
8	Cori - Cray XC260, Intel Xeon Phi 7250 120 1.4GHz, Aries	United States	interconnect: Cray Inc.	622,336	14,014.7
9	Oakforest-PA03 - PRIMERGY CX1640 M4, Intel Xeon Phi 7250 120	Japan	1.4GHz, Intel Omni-Path, Fujitsu	556,104	13,554.6
10	K computer, SPARC64 VIIIfx, 2.0GHz, Tofu interconnect, Fujitsu	Japan	SPARC64 VIIIfx, Advanced Institute for Computational Science (AICS)	705,024	10,510.0



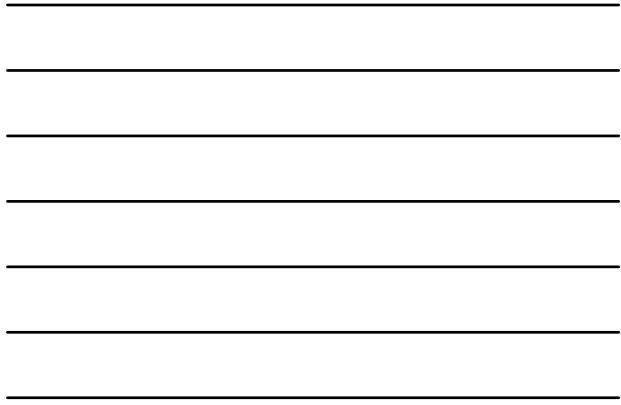
**TOP 500**

**T  
O  
P  
1  
0  
0  
n  
o  
v  
e  
m  
b  
r  
e  
2  
0  
1  
8**

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

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

Rank	System	Country	Processors	TFlops	KWatt
1	Sumway TaihuLight - Sumway MPP, Sumway SW26010 2007	China	1,435Hz, TH Equipes, Xeon ES-2692	10,649,600	93,014.6
2	Tianhe-2 (MilkyWay-2) - TH Equipes, Xeon ES-2692	China	120 2.20GHz, TH Equipes, Xeon Phi 7205	3,120,000	33,862.7
3	Piz Daint - Cray XC260, Intel Xeon E5-2690 v3 120 2.6GHz, Aries	Switzerland	interconnect: NVIDIA Tesla P100	361,760	19,590.0
4	Gyokkou - ZettaScale-2 HPC system, Xeon D-1571 140 1.30GHz	Japan	interconnect: NVIDIA K80	19,860,000	19,135.8
5	Titan - Cray XC260, NVIDIA K80 120 2.20GHz, Cray Gemini	United States	interconnect: NVIDIA K80	560,640	17,590.0
6	Sequoia - BlueGene/Q, Power BQC 140 1.60 GHz, Custom	United States	interconnect: IBM	1,572,864	17,173.2
7	Trinity - Cray XC260, Intel Xeon Phi 7250 680 1.4GHz, Aries	United States	interconnect: Cray Inc.	979,968	14,137.3
8	Cori - Cray XC260, Intel Xeon Phi 7250 120 1.4GHz, Aries	United States	interconnect: Cray Inc.	622,336	14,014.7
9	Oakforest-PA03 - PRIMERGY CX2570 M4, Xeon Gold 6148 200 2.4GHz, NVIDIA Tesla V100 120, Infiniband EDR, Fujitsu	Japan	interconnect: NVIDIA Tesla V100 120, Infiniband EDR, Fujitsu	556,104	13,554.6
10	K computer, SPARC64 VIIIfx, 2.0GHz, Tofu interconnect, Fujitsu	Japan	SPARC64 VIIIfx, Advanced Institute for Computational Science (AICS)	705,024	10,510.0



**Le Top500**  
N°1 en novembre 2018

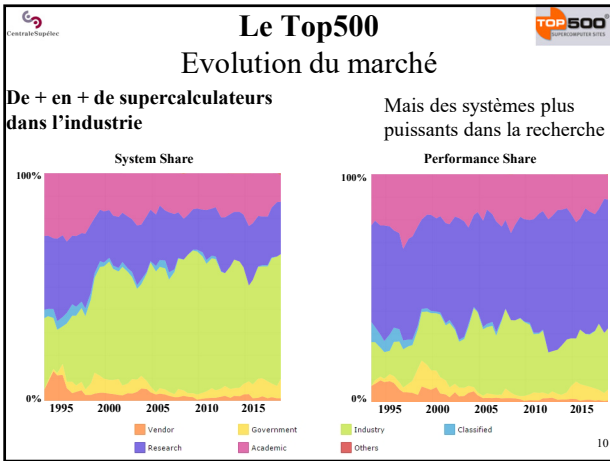
**143.5 Pflops** (×1.54 vs previous n°1)

- 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** (×0.64 vs previous n°1)

**Flops/Watt : ×2.4**






---

---

---

---

---

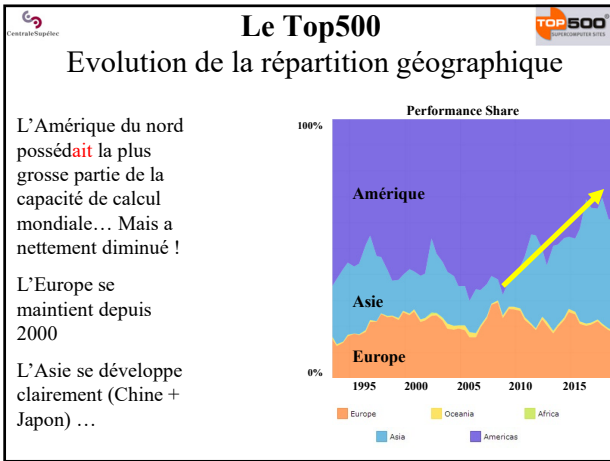
---

---

---

---

---




---

---

---

---

---

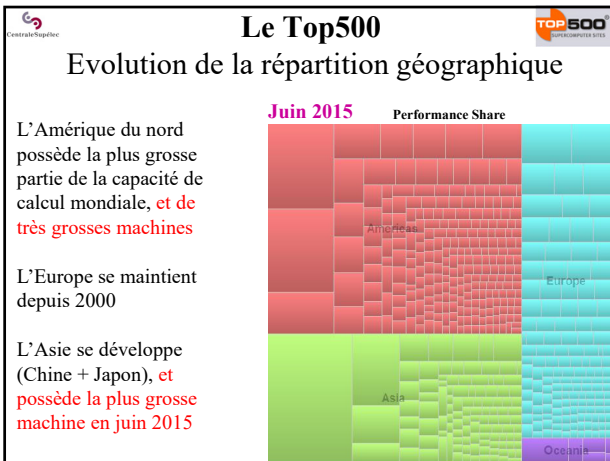
---

---

---

---

---




---

---

---

---

---

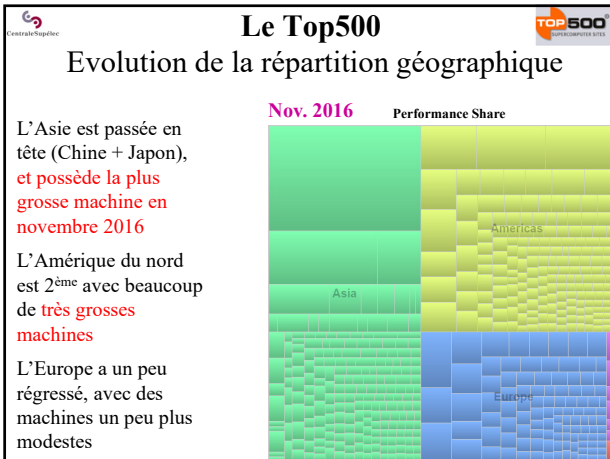
---

---

---

---

---




---

---

---

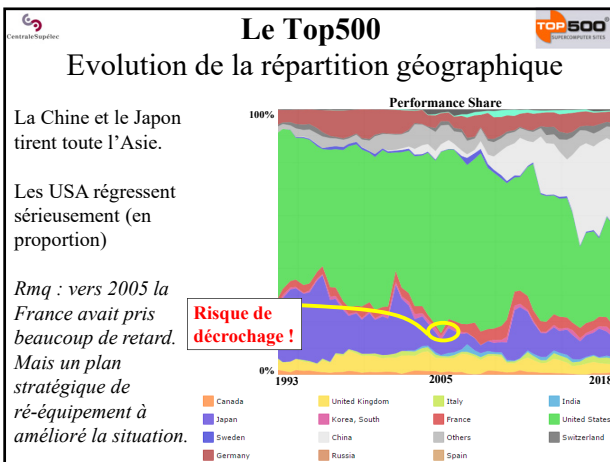
---

---

---

---

---




---

---

---

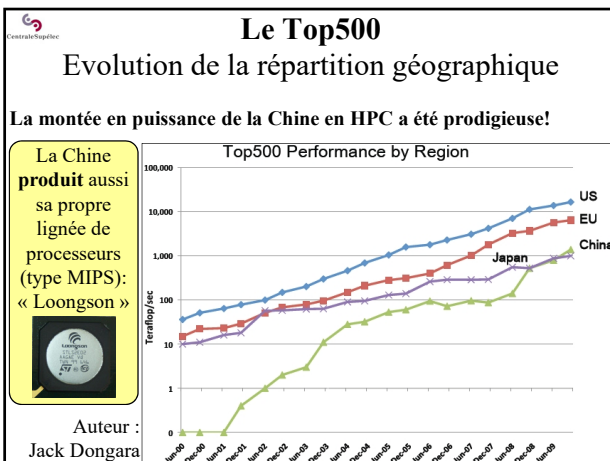
---

---

---

---

---




---

---

---

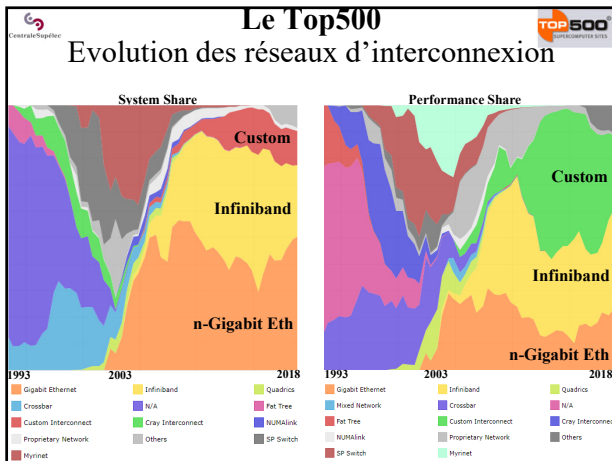
---

---

---

---

---




---

---

---

---

---

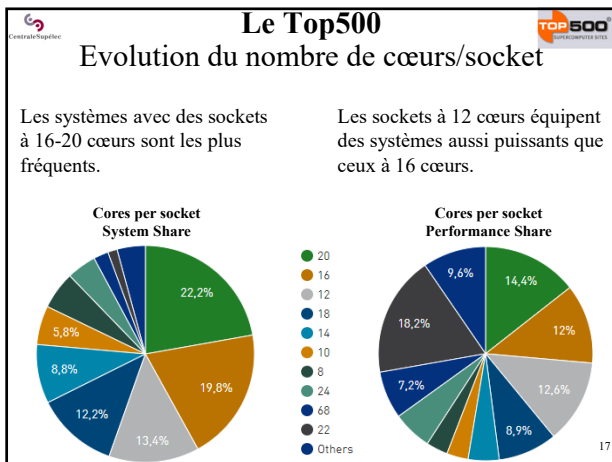
---

---

---

---

---




---

---

---

---

---

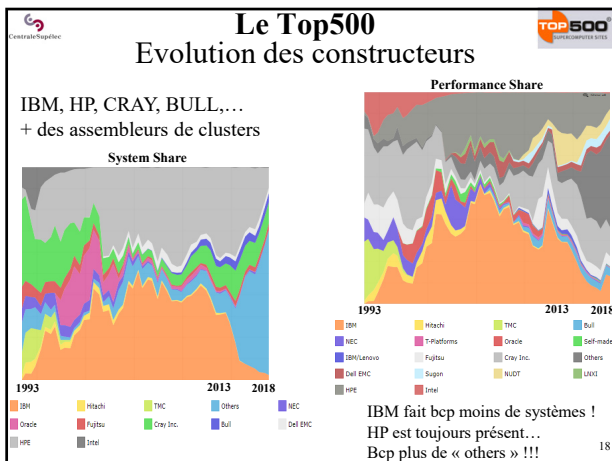
---

---

---

---

---




---

---

---

---

---

---

---

---

---

---

CentralesSupélec

## 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

19

---

---

---

---

---

---

---

---

CentralesSupélec

## Top 500 « vert » des supercalculateurs

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

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.

SUPERMICRO  
Proud Sponsor of The Green500

Recent Green500 News

---

---

---

---

---

---

---

---

CentralesSupélec

## 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	BM - 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	210.58	Harvard University	eServer Blue Gene Solution	44.80	170
7	210.58	High Energy Accelerator Research Organization /KEK	eServer Blue Gene Solution	44.80	171
8	210.58	BM - Almaden Research Center	eServer Blue Gene Solution	44.80	172
9	210.58	BM Research	eServer Blue Gene Solution	44.80	173
10	210.58	BM 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 PowerXCell 8i 3.2 Ghz, Infiniband	34.63	220
2	530.33	Repsol YPF	BladeCenter QS22 Cluster 8i 3.2 Ghz, Infiniband	26.38	429
2	530.33	Repsol YPF	BladeCenter QS22 Cluster 8i 3.2 Ghz, Infiniband	26.38	430
2	530.33	Repsol YPF	BladeCenter QS22 Cluster 8i 3.2 Ghz, Infiniband	26.38	431
5	458.33	DOE/NNSA/LANL	BladeCenter QS22/LS21 Cluster PowerXCell 8i 3.2 Ghz / Opleron DC 1.8 Ghz, Infiniband	138	41
5	458.33	IBM Poughkeepsie Benchmarking Center	BladeCenter QS22/LS21 Cluster PowerXCell 8i 3.2 Ghz / Opleron DC 1.8 Ghz, Infiniband	138	42
7	444.94	DOE/NNSA/LANL	BladeCenter QS22/LS21 Cluster PowerXCell 8i 3.2 Ghz / Opleron DC 1.8 Ghz, Voltare 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 / Opleron DC 1.8 Ghz, Infiniband	276	29
4	458.33	IBM Poughkeepsie Benchmarking Center	BladeCenter QS22/LS21 Cluster PowerXCell 8i 3.2 Ghz / Opleron DC 1.8 Ghz, Infiniband	138	76
6	444.25	DOE/NNSA/LANL	BladeCenter QS22/LS21 Cluster PowerXCell 8i 3.2 Ghz / Opleron DC 1.8 Ghz, Voltare Infiniband	2345.5	2
7	428.91	National Astronomical Observatory of Japan	GRAPE-DR accelerator Cluster, Infiniband	51.2	445
8	378.24	National Supercomputer Center in Tianjin/NUDT	NUDT TH-1 Cluster, Xeon E5540E5408, ATI Radeon HD 4870 2, Infiniband	1484.8	5
9	378.77	King Abdulaziz University of Science and Technology	Blue Gene/P Solution	504	18
9	378.77	EDF R&D	Blue Gene/P Solution	252	49

## Top 500 « vert » des supercalculateurs

Nov. 2010 Le 1<sup>er</sup> 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	1448.63	National Astronomical Observatory of Japan	GRAPE-DR accelerator Cluster, Infiniband	24.59
3	958.35	GSIC Center, Tokyo Institute of Technology	HP ProLiant SL390s G7 Xeon 6C X5670 Nvidia	1243.80
4	933.06	NCSA	Hybrid Cluster Core i3 2.93Ghz Dual Core NVIDIA	38.00
5	828.67	RWEN Advanced Institute for Computational Science	K computer SPARC64 VIII, 2.0GHz, Tolu Interconnect	57.96
6	773.38	Universitaet Wuppertal	QPACE SFB TR Cluster PowerXCell 8i 3.2 Ghz, 3D-Torus	57.54
6	773.38	Universitaet Regensburg	QPACE SFB TR Cluster PowerXCell 8i 3.2 Ghz, 3D-Torus	57.54
6	773.38	Forschungszentrum Juelich (FZJ)	QPACE SFB TR Cluster PowerXCell 8i 3.2 Ghz, 3D-Torus	57.54
9	740.78	Universitaet Frankfurt	Supermicro Cluster, OC Opleron 2.1 GHz ATI	385.00
10	677.42	Georgia Institute of Technology	HP ProLiant SL390s G7 Xeon 6C X5660 2.8GHz Nvidia Fermi Infiniband QDR	94.40
11	636.36	National Institute for Environmental Studies	GOSAT Research Computation Facility, nvidia	117.15



**Top 500 « vert » des supercalculateurs**

**Juin 2011** Le 2<sup>nd</sup> 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	1664.20	IBM Thomas J. Watson Research Center	NNSA/SC Blue Gene/Q Prototype 1	38.80
3	1375.88	Nagasaki University	DEGIMA Cluster, Intel E5, ATI Radeon GPU	34.24
4	958.35	GSIC Center, Tokyo Institute of Technology	HP ProLiant SL390s G7 Xeon E5 X5670, Nvidia GPU, Linux/Windows	1243.80
5	891.88	CINECA / SCS - SuperComputing Solution	iDataFlex DX360M3, Xeon 2, nvidia GPU	160.00
6	824.56	RIKEN Advanced Institute for Computational Science (AICS)	K computer, SPARC64 VIIIx, 2.0GHz, Tofu Interconnect	9898.56
7	773.38	Forschungszentrum Juelich (FZJ)	QPACE SFB TR Cluster, PowerCell, 3.2 GHz, 3D-Torus	57.54
8	773.38	Universitaet Regensburg	QPACE SFB TR Cluster, PowerCell, 3.2 GHz, 3D-Torus	57.54
9	773.38	Universitaet Wuppertal	QPACE SFB TR Cluster, PowerCell, 3.2 GHz, 3D-Torus	57.54
10	718.13	Universitaet Frankfurt	Supermicro Cluster, OC Opteron 2.1 GHz, ATI, Radeon GPU, Infiniband	416.78

---

---

---

---

---

---

---

---

---

---

**Top 500 « vert » des supercalculateurs**

**Juin 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	DOE/NNSA/LLNL	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 Q6824M, Xeon E5-2670 8C 2.800GHz, Infiniband FDR, Intel Xeon Phi 5110P	44.89
2	2,351.10	King Abdulaziz City for Science and Technology	SANAM - Adtech ESC 4000 GDR G2, Xeon E5-2650 8C 2.000GHz, Infiniband FDR, AMD FirePro S10000	179.15
3	2,142.77	DOERSC/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 6274 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	BQDev - 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 1.600GHz, Custom	41.09
9	2,101.12	IBM Thomas J. Watson Research Center	BlueGene/Q, Power BQC 16C 1.60 GHz, Custom	82.19
10	2,101.12	Ecole Polytechnique Fédérale de Lausanne	CADMOS BQ10 - BlueGene/Q, Power BQC 16C 1.600GHz, Custom Interconnect	82.19

---

---

---

---

---

---

---

---

---

---



**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 E5-2618L3 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-1.4.168Brick, Xeon E5-2618L3 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 32U755SC Cluster, Intel Xeon E5-2690v2 10C 2.2GHz, Infiniband FDR PEZY-SC CPU, 10 cores/socket, 2.2GHz	32.59
4	5,271.81	GSI Helmholtz Center	ASUS ESC4000 EDR JCS, 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 2.100GHz, Infiniband FDR NVIDIA K20c	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.	Storm - Cray CS-Storm, Intel Xeon E5-2680v2 10C 2.8GHz, Infiniband FDR Nvidia K40m	44.54
8	3,631.70	Cambridge University	Wilkes - Dell T670 Cluster, Intel Xeon E5-2630v3 6C 2.600GHz, Infiniband FDR NVIDIA K20	52.82
9	3,614.71	TU Dresden, ZIH	Taur - 2GBE - Bullitbulx R400, Xeon E5-2680v3 12C 2.5GHz, Infiniband FDR Nvidia K80	58.01
10	3,543.32	Financial Institution	Intel E5-2690v2 10C 2.8GHz, Infiniband, NVIDIA K20c	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 E5-2618L3 8C 2.3GHz, Infiniband FDR PEZY-SC CPU, 8 cores/socket, 2.3GHz	50.32
2	5,331.70	GSIC Center, Tokyo Institute of Technology	TSUBAME-KFCIDL - 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 EDR JCS, 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	Suiren Cluster W780, Xeon E5-2640v3 8C 2.8GHz, 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 TS1000 HPC Server, Xeon E5-2620v3 6C 2.4GHz, 10G Ethernet, NVIDIA Tesla K40	58.00
7	3,775.45	Internet Service	Inspur TS1000 HPC Server, Intel Xeon E5-2620v2 6C 2.1GHz, 10G Ethernet, NVIDIA Tesla K40	110.00
8	3,775.45	Internet Service	Inspur TS1000 HPC Server, Intel Xeon E5-2620v2 6C 2.1GHz, 10G Ethernet, NVIDIA Tesla K40	110.00
9	3,775.45	Internet Service	Inspur TS1000 HPC Server, Intel Xeon E5-2620v2 6C 2.1GHz, 10G Ethernet, NVIDIA Tesla K40	110.00
10	3,775.45	Internet Service	Inspur TS1000 HPC Server, Intel Xeon E5-2620v2 6C 2.1GHz, 10G Ethernet, NVIDIA Tesla K40	32 110.00

---

---

---

---

---

---

---

---

---

---

---

---

**Top 500 « vert » des supercalculateurs**

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

Green500 Rank	MFLOPS/W	Site	System	Power(kW)
1	9462.1	NVIDIA Corporation	NVIDIA DGX-1, Xeon E5-2699v4 20C 2.2GHz, Infiniband EDR NVIDIA Tesla P100	349.5
2	7453.5	Swiss National Supercomputing Centre (CSCS)	Cray XC30, Xeon E5-2695v4 12C 2.4GHz, Aries interconnect, NVIDIA Tesla P100	1312
3	6673.8	Advanced Center for Computing and Research, RIKEN	ZettaScaler-1.6, Xeon E5-2618L3 8C 2.3GHz, Infiniband FDR PEZY-SC CPU, 8 cores/socket, 2.3GHz	150.0
4	6140.0	Swiss National 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	77
6	4985.7	Joint Center for Advanced High Performance Computing	PRIMERGY CX1640 M1, Intel Xeon Phi 7250	2718.7
7	4688.0	DDE/SC/Argonne National Laboratory	Cray XC40, Intel Xeon Phi 7230 64C	1087
8	4112.1	Stanford Research Computing Center	Cray CS-Storm, Intel Xeon E5-2690v2 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	748.1
10	3836.6	Thomas Jefferson National Accelerator Facility	KOI Cluster, Intel Xeon Phi 7230 64C	111 33

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

---

---

---

---

---

---

---

---

---

---

---

---



## 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

---

---

---

---

---

---

---

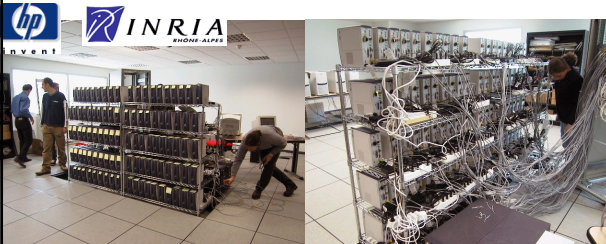
---

## 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 :

**Evolution contexte grappe**

40

---

---

---

---

---

---

---

---

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

**CRAY**

**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 ? »**

42

---

---

---

---

---

---

---

---

CentralesSupélec

## 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

---

---

---

---

---

---

---

---

CentralesSupélec

## 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-habituels !

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

---

---

---

---

---

---

---

---

CentralesSupélec

## 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 ☺

Musée Ferrari - Modène

---

---

---

---

---

---

---

---

---

---

Evolution, marché et gestion des machines parallèles

**FIN**

---

---

---

---

---

---

---

---

---

---