

CentraleSupélec

Mineure CalHauI

Evolution, marché et gestion des machines parallèles

Stéphane Vialle

universit  Paris-Saclay Sciences et technologies de l'information et de la communication (STIC)

RISEGRID Grand Est

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

CentraleSupélec

Evolution, march  et gestion des machines parall es

- Le « top500 » - march  des machines parall es
- Le « top500 green » -  volution du ratio Flop/Watt
- Cluster vs Supercalculateur : co t et co t total
- Gestion et utilisation d'une machine parall e

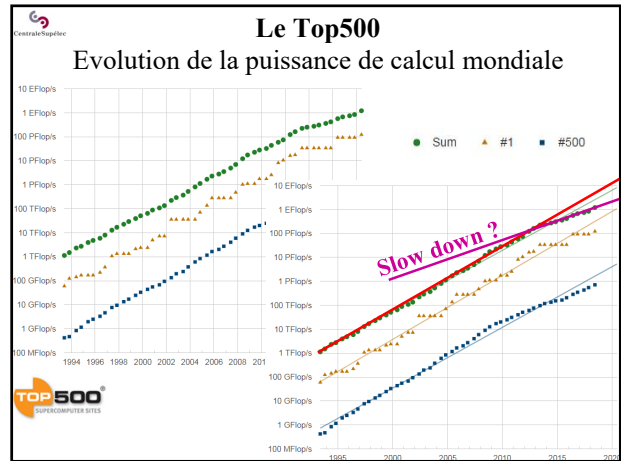
2

CentraleSupélec

Le Top500 et le march  des super-calculateurs

<http://www.top500.org>

- Mis   jour en novembre et en juin, depuis 1993
- Liste des 500 plus grosses machines
- Des informations sur leurs architectures
- Des statistiques sur l' volution des technologies et des march s



Rank	Country	System Name	Processor	Nodes	TFlops	KWatt
1	China	Tianhe-2 (MilkyWay-2) - TH-1TB-FEP Cluster, Xeon ES-2692 12C 2.200GHz, TH Express-2, Intel Xeon Phi 3151P, NUDT	3,120,000	33,862.7	54,702.4	17,895
2	United States	Titan - Cray XK7, Opteron 6200, Cray Gemini interconnect, NVIDIA K20x	540,640	17,590.0	27,112.5	8,209
3	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	Japan	K computer, SPARC64 VIIIfx 2.0GHz, Tofu interconnect, Fujitsu	705,024	10,510.0	11,280.0	12,660
5	United States	Mira - BlueGene/Q, Power BQC 16C 1.60GHz, Custom IBM	786,432	8,586.6	10,066.3	3,945
6	United States	Trinity - Cray XC40, Xeon ES-2698v3 16C 2.3GHz, Aries interconnect, Cray Inc.	301,056	8,100.9	11,078.9	4,233
7	Switzerland	Piz Daint - Cray XC30, Xeon ES-2692 12C 2.00GHz, Cray Inc., NVIDIA K20x	115,984	6,271.0	7,788.9	2,325
8	Germany	Hazel Hen - Cray XC40, Xeon ES-2680v3 12C 2.5GHz, Aries interconnect, Cray Inc.	185,088	5,440.2	7,403.5	2,610
9	Saudi Arabia	Shaheen II - Cray XC40, Xeon ES-2698v3 16C 2.3GHz, Aries interconnect, Cray Inc.	196,608	5,537.0	7,235.2	2,834
10	United States	Stampede - PowerEdge C9200, Xeon ES-2690 8C 2.70GHz, Infiniband FDR, Intel Xeon Phi 5450P, Dell	462,442	5,168.1	8,520.1	4,510

Rank	Country	System Name	Processor	Nodes	TFlops	KWatt
1	China	National Supercomputing Center in Wuzi, Sunway Taihulight - Sunway MPP, Sunway SW26010 2600 1.45GHz, Sunway NRCC	10,647,600	33,862.7	54,702.4	17,895
2	China	National Super Computer Center in Guangzhou, Tianhe-2 (MilkyWay-2) - TH-1TB-FEP Cluster, Intel Xeon ES-2692 12C 2.200GHz, TH Express-2, NUDT, Xeon Phi 3151P	3,120,000	33,862.7	54,702.4	17,895
3	United States	DOE/SC/Oak Ridge National Laboratory, Titan - Cray XK7, Opteron 6200, Cray Gemini interconnect, NVIDIA K20x	540,640	17,590.0	27,112.5	8,209
4	United States	DOE/NNSA/LNL, Sequoia - BlueGene/Q, Power BQC 16C 1.60 GHz, Custom IBM	1,572,864	17,173.2	20,132.7	7,890
5	United States	DOE/SC/Oak Ridge National Laboratory, Mira - BlueGene/Q, Power BQC 16C 1.60GHz, Custom IBM	786,432	8,586.6	10,066.3	3,945
6	Japan	Joint Center for Advanced High Performance Computing, K computer, SPARC64 VIIIfx 2.0GHz, Tofu interconnect, Fujitsu, Xeon Phi 7250 68C, PRIMERBY CX1640 M1, Intel	556,104	13,554.6	24,913.5	2,719
7	Japan	RIKEN Advanced Institute for Computational Science (AICS), K computer, SPARC64 VIIIfx 2.0GHz, Tofu interconnect, Fujitsu	705,024	10,510.0	11,280.0	12,660
8	Switzerland	Swiss National Supercomputing Centre (SNLC), Piz Daint - Cray XC50, Xeon ES-2692 12C 2.00GHz, Aries interconnect, NVIDIA Tesla P100, Cray Inc.	206,720	9,779.0	15,988.0	1,312
9	United States	DOE/SC/Argonne National Laboratory, Mira - BlueGene/Q, Power BQC 16C 1.60GHz, Custom IBM	786,432	8,586.6	10,066.3	3,945
10	United States	DOE/NNSA/LNL/SNL, Trinity - Cray XC40, Xeon ES-2698v3 16C 2.3GHz, Aries interconnect, Cray Inc.	301,056	8,100.9	11,078.9	4,233

Rank	System Name	Country	Processors	TFlops	KWatt
1	Summit@Light - Summit MPP, Summit SW20110 2007	China	10,649,600	93,014.6	25,435.9
2	Tianhe-2 (MilkyWay-2) - TH-2012, Tianhe-2 Super Computer Co	China	3,120,000	33,862.7	17,308
3	Piz Daint - Cray XE6, NVIDIA Tesla P100, Cray Inc.	Switzerland	361,760	19,590.0	25,326.3
4	Oyukou - ZettaScale-2 HPC system, Xeon D-1571 140 1.30GHz, Intel Xeon Phi 7205	Japan	19,860,000	19,135.8	28,192.0
5	Titan - Cray XE6, NVIDIA K80, Cray Inc.	United States	560,640	17,590.0	27,112.5
6	Sequoia - BlueGene/Q, Power BQC 140 1.60 GHz, Custom, IBM	United States	1,572,864	17,173.2	20,132.7
7	Trinity - Cray XE6, Intel Xeon Phi 7250 680 1.40GHz, Arries	United States	979,968	14,137.3	43,902.6
8	Cori - Cray XE6, Intel Xeon Phi 7250 680 1.40GHz, Arries	United States	622,336	14,014.7	27,880.7
9	Oakforest-PACS - PRIMERY DX1640 1.40GHz, Intel Omni-Path, Fujitsu	Japan	556,104	13,554.6	24,913.5
10	K computer, SPARC64 VIIfx, 2.0GHz, Tofu interconnect, Fujitsu	Japan	705,024	10,510.0	11,280.4

Rank	System Name	Country	Processors	TFlops	KWatt
1	Summit@Light - Summit MPP, Summit SW20110 2007	China	10,649,600	93,014.6	25,435.9
2	Tianhe-2 (MilkyWay-2) - TH-2012, Tianhe-2 Super Computer Co	China	3,120,000	33,862.7	17,308
3	Piz Daint - Cray XE6, NVIDIA Tesla P100, Cray Inc.	Switzerland	361,760	19,590.0	25,326.3
4	Oyukou - ZettaScale-2 HPC system, Xeon D-1571 140 1.30GHz, Intel Xeon Phi 7205	Japan	19,860,000	19,135.8	28,192.0
5	Titan - Cray XE6, NVIDIA K80, Cray Inc.	United States	560,640	17,590.0	27,112.5
6	Sequoia - BlueGene/Q, Power BQC 140 1.60 GHz, Custom, IBM	United States	1,572,864	17,173.2	20,132.7
7	Trinity - Cray XE6, Intel Xeon Phi 7250 680 1.40GHz, Arries	United States	979,968	14,137.3	43,902.6
8	Cori - Cray XE6, Intel Xeon Phi 7250 680 1.40GHz, Arries	United States	622,336	14,014.7	27,880.7
9	Oakforest-PACS - PRIMERY DX1640 1.40GHz, Intel Omni-Path, Fujitsu	Japan	556,104	13,554.6	24,913.5
10	K computer, SPARC64 VIIfx, 2.0GHz, Tofu interconnect, Fujitsu	Japan	705,024	10,510.0	11,280.4

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 (×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

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 possédait 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) ...

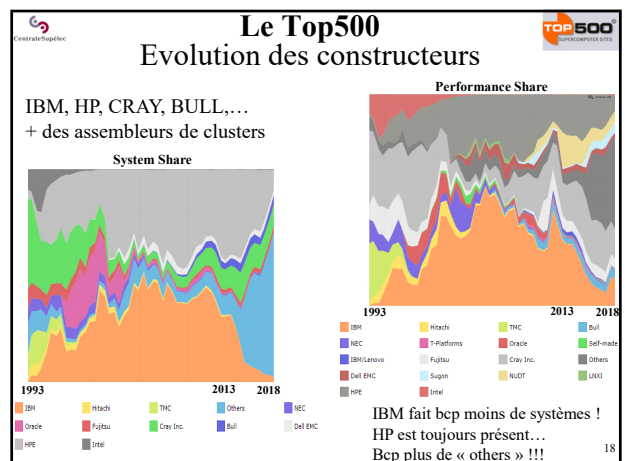
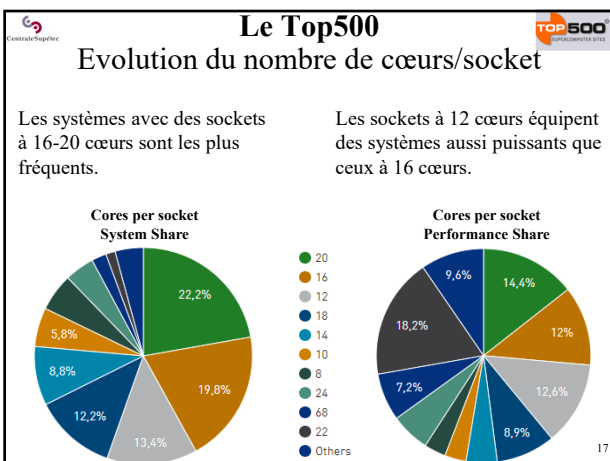
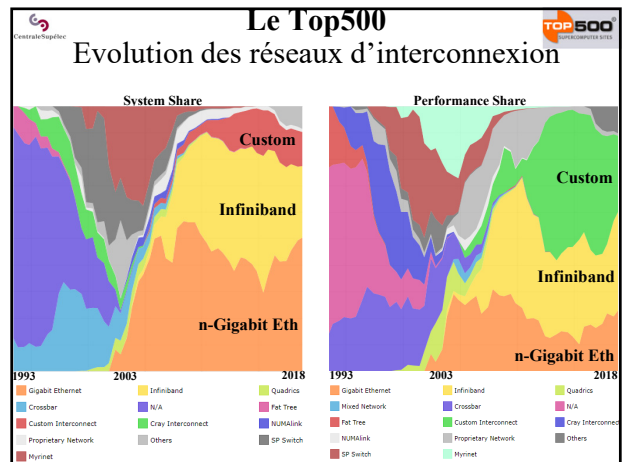
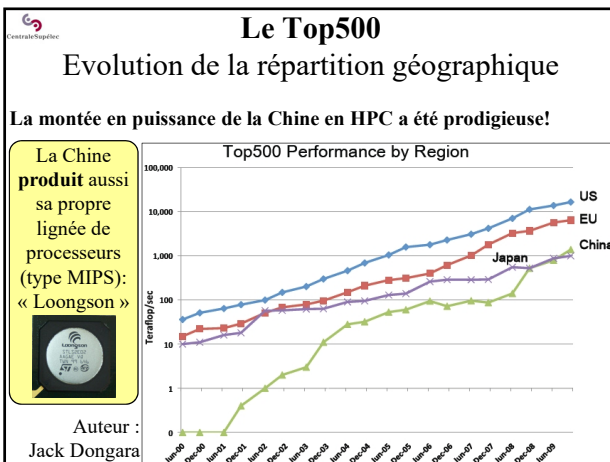
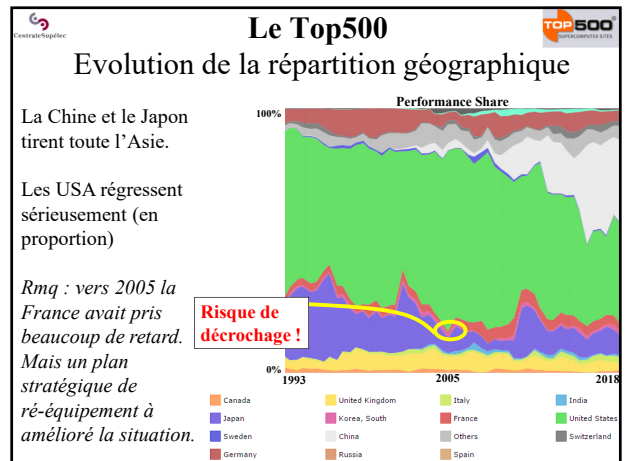
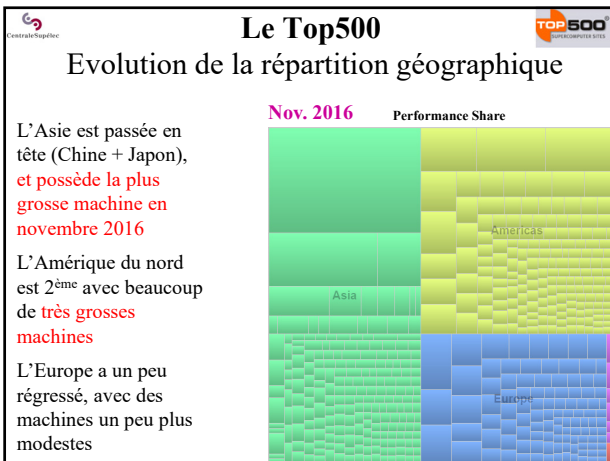
Le Top500

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



CentralSuperSite

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

CentralSuperSite

Top 500 « vert » des supercalculateurs

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

THE GREEN 500

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

CentralSuperSite

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	352.28	Max-Planck-Gesellschaft MPI/IPP	Blue Gene/P Solution	62.20	40
3	345.95	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

CentralSuperSite

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 @ 4.0 Ghz, Infiniband	34.63	220
2	530.33	Repsol YPF	BladeCenter QS22 Cluster @ 3.2 Ghz, Infiniband	26.38	429
2	530.33	Repsol YPF	BladeCenter QS22 Cluster @ 3.2 Ghz, Infiniband	26.38	430
2	530.33	Repsol YPF	BladeCenter QS22 Cluster @ 3.2 Ghz, Infiniband	26.38	431
5	458.33	DOE/NNSA/LANL	BladeCenter QS22/LS21 Cluster, PowerXCell @ 3.2 Ghz / Opteron DC 1.8 Ghz, Infiniband	138	41
5	458.33	IBM Poughkeepsie Benchmarking Center	BladeCenter QS22/LS21 Cluster, PowerXCell @ 3.2 Ghz / Opteron DC 1.8 Ghz, Infiniband	138	42
7	444.94	DOE/NNSA/LANL	BladeCenter QS22/LS21 Cluster, PowerXCell @ 3.2 Ghz / Opteron DC 1.8 Ghz, 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

CentralSuperSite

Top 500 « vert » des supercalculateurs

Les « Cell » en tête, et les performances progressent

Nov. 2009

Green500 Rank	MFLOPS/W	Site*	Computer*	Total Power (kW)	TOP500 Rank*
1	722.98	Forschungszentrum Juelich (FZJ)	QPACE SFB TR Cluster, PowerXCell @ 3.2 Ghz, 3D-Torus	59.49	110
1	722.98	Universitaet Regensburg	QPACE SFB TR Cluster, PowerXCell @ 3.2 Ghz, 3D-Torus	59.49	111
1	722.98	Universitaet Wuppertal	QPACE SFB TR Cluster, PowerXCell @ 3.2 Ghz, 3D-Torus	59.49	112
4	458.33	DOE/NNSA/LANL	BladeCenter QS22/LS21 Cluster, PowerXCell @ 3.2 Ghz / Opteron DC 1.8 Ghz, Infiniband	276	29
4	458.33	IBM Poughkeepsie Benchmarking Center	BladeCenter QS22/LS21 Cluster, PowerXCell @ 3.2 Ghz / Opteron DC 1.8 Ghz, Infiniband	138	78
6	444.25	DOE/NNSA/LANL	BladeCenter QS22/LS21 Cluster, PowerXCell @ 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/E5440, ATI Radeon HD 4870 2, Infiniband	1484.8	5
9	378.77	King Abdullah University of Science and Technology	Blue Gene/P Solution	504	16
9	378.77	EDF R&D	Blue Gene/P Solution	252	49

23

CentralSuperSite

Top 500 « vert » des supercalculateurs

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

Nov. 2010

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
24	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 RC X5570, Nvidia GPU, Linux/Windows	1243.80
3	933.06	NCSA	Hybrid Cluster Core i3 2.93Ghz Dual Core, Nvidia	36.00
4	828.67	RNEN Advanced Institute for Computational Science	K computer, SPARC64 VIIIx2, 2.0Ghz, Tofu Interconnect	57.96
5	773.38	Universitaet Wuppertal	QPACE SFB TR Cluster, PowerXCell @ 3.2 Ghz, 3D-Torus	57.54
6	773.38	Universitaet Regensburg	QPACE SFB TR Cluster, PowerXCell @ 3.2 Ghz, 3D-Torus	57.54
7	773.38	Forschungszentrum Juelich (FZJ)	QPACE SFB TR Cluster, PowerXCell @ 3.2 Ghz, 3D-Torus	57.54
8	740.78	Universitaet Frankfurt	Supermicro Cluster, OC Opteron 2.1 Ghz, ATI Radeon GPU, Infiniband	385.00
9	677.12	Georgia Institute of Technology	HP ProLiant SL390s G7 Xeon RC X5560 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

Juin 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	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 E5-2650, 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 2.8, Nvidia GPU	160.00
6	824.56	RIKEN Advanced Institute for Computational Science (AICS)	K computer, SPARC64 VIII, 2.0GHz, Tofu Interconnect	9898.56
7	773.38	Forschungszentrum Juelich (FZJ)	QPACE SFB TR Cluster, QH, 3D-Torus, PowerCell, sl, 3.2	57.54
8	773.38	Universitaet Regensburg	QPACE SFB TR Cluster, QH, 3D-Torus, PowerCell, sl, 3.2	57.54
9	773.38	Universitaet Wuppertal	QPACE SFB TR Cluster, QH, 3D-Torus, PowerCell, sl, 3.2	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	DOENNSALLNL	BlueGene/Q, Power BQC 16C 1.600GHz, Custom	41.10
2	2,100.88	IBM Thomas J. Watson Research Center	BlueGene/Q, Power BQC 16C 1.600GHz, Custom	41.10
3	2,100.86	DOENNSALLNL	BlueGene/Q, Power BQC 16C 1.600GHz, Custom	82.20
4	2,100.86	DOENNSALLNL	BlueGene/Q, Power BQC 16C 1.600GHz, Custom	82.20
5	2,100.86	DOENNSALLNL	BlueGene/Q, Power BQC 16C 1.600GHz, Custom	82.20
6	2,100.86	DOENNSALLNL	BlueGene/Q, Power BQC 16C 1.600GHz, 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 1.600GHz, Custom	493.10
9	2,099.50	Science and Technology Facilities Council - Daresbury Laboratory	BlueGene/Q, Power BQC 16C 1.600GHz, Custom	575.30
10	2,099.46	Forschungszentrum Juelich (FZJ)	BlueGene/Q, Power BQC 16C 1.600GHz, Custom	657.50

L'année de l'architecture BlueGene de l'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 2.800GHz, Infiniband FDR, Intel Xeon Phi 5110P	44.89
2	2,351.10	King Abdulaziz City for Science and Technology	SANAM - Adtech ESC 4000, Xeon E5-2650 8C 2.800GHz, 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 (SCSC)	Todi - Cray XK7, Opteron 6274 16C 2.200GHz, Cray Gemini interconnect, NVIDIA Tesla K20	129.00
5	2,102.12	Forschungszentrum Juelich (FZJ)	JJUQUEEN - 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	BQ-Dev - BlueGene/Q, Power BQC 16C 1.600GHz, Custom Interconnect	41.09
7	2,101.39	DOENNSALLNL	rzuseq - BlueGene/Q, Power BQC 16C 1.600GHz, Custom	41.09
8	2,101.39	IBM Thomas J. Watson Research Center	BlueGene/Q, Power BQC 16C 1.600GHz, Custom	41.09
9	2,101.32	IBM Thomas J. Watson Research Center	BlueGene/Q, Power BQC 16C 1.60 GHz, Custom	82.19
10	2,101.12	Ecole Polytechnique Federale de Lausanne	CADMO5 BQ10 - 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/104R4-1G Cluster, Intel Xeon E5-2620v2 6C 2.100GHz, Infiniband FDR, NVIDIA K20x	27.78
2	3,631.84	Wilkes - Dell T140 Cluster, Intel Xeon E5-2630v2 6C 2.600GHz, Infiniband FDR, NVIDIA K20	52.62	
3	3,631.84	HA-PACS TCA - Cray 3625G4-SM Cluster, Intel Xeon E5-2680v2 10C 2.800GHz, Infiniband QDR, NVIDIA K20x	78.77	
4	3,631.84	Swiss National Supercomputing Centre (SCSC)	romeo - Bull R421-E3 Cluster, Intel Xeon E5-2650v2 8C 2.600GHz, Infiniband FDR, NVIDIA K20x	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 3004-G7, Xeon X5670 6C 2.930GHz, Infiniband QDR, NVIDIA K20x	922.54
7	2,702.16	University of Arizona	iDataFlex DX360M4, Intel Xeon E5-2650v2 8C 2.600GHz, Infiniband FDR, NVIDIA K20x	53.62
8	2,629.10	Max-Planck-Gesellschaft MPVPP	CraneFly F436M4, Intel Xeon E5-2680v2 10C 2.800GHz, Infiniband, NVIDIA K20x	269.94
9	2,629.10	Financial Institution	CraneFly F436M4, Intel Xeon E5-2680v2 10C 2.800GHz, Infiniband, NVIDIA K20x	55.62
10	2,358.69	CSIRO	CSIRO GPU Cluster - Xeon Phi 3100, Xeon E5-2650 8C 2.800GHz, Infiniband FDR, NVIDIA K20m	71.01

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

Top 500 « vert » des supercalculateurs

Juin 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/104R4-1G Cluster, Intel Xeon E5-2620v2 6C 2.100GHz, Infiniband FDR, NVIDIA K20x	34.58
2	3,631.70	Cambridge University	Wilkes - Dell T140 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 3625G4-SM Cluster, Intel Xeon E5-2680v2 10C 2.800GHz, Infiniband QDR, NVIDIA K20x	78.77
4	3,459.46	SURFsara	Canterius Accelerator Island - Xeon E5-2450v2 6C 2.5GHz, Infiniband 4x FDR, Nvidia K40m	44.40
5	3,185.91	Swiss National Supercomputing Centre (SCSC)	romeo - Bull R421-E3 Cluster, Intel Xeon E5-2650v2 8C 2.600GHz, Infiniband FDR, NVIDIA K20x	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 - Xeon Phi 3100, 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 SL 3004-G7, Xeon X5670 6C 2.930GHz, Infiniband QDR, NVIDIA K20x	927.86
9	2,813.14	Exploration & Production - Eni S.p.A.	HPCC - Catalyst 2000M4, Intel Xeon E5-2680v2 10C 2.80GHz, Infiniband FDR, NVIDIA K20x	1,067.49
10	2,678.41	Financial Institution	CraneFly F436M4, Intel Xeon E5-2680v2 10C 2.800GHz, Infiniband, NVIDIA K20x	54.60

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 ESC 4000 P2R2926, 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	Suiken - ExaScaler 32, Intel Xeon E5-2660v2 10C 2.20GHz, Infiniband FDR, AMD FirePro S9150, CPU: 110 cores/socket	37.83
3	4,447.58	GSIC Center, Tokyo Institute of Technology	TSUBAME-KFC - LX 1U-4GPU/104R4-1G Cluster, Intel Xeon E5-2620v2 6C 2.100GHz, Infiniband FDR, NVIDIA K20x	35.39
4	3,962.73	Cray Inc.	Storax - Cray C600m, Intel Xeon E5-2690v2 10C 2.2GHz, Infiniband FDR, Nvidia K40m	44.54
5	3,631.70	Cambridge University	Wilkes - Dell T140 Cluster, Intel Xeon E5-2630v2 6C 2.600GHz, Infiniband FDR, NVIDIA K20	52.62
6	3,543.32	Financial Institution	iDataFlex 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 CS360 Cluster, Intel Xeon E5-2680v2 10C 2.800GHz, Infiniband QDR, NVIDIA K20x	78.77
8	3,459.46	SURFsara	Canterius Accelerator Island - Xeon E5-2450v2 6C 2.5GHz, Infiniband 4x FDR, Nvidia K40m	44.40
9	3,185.91	Swiss National Supercomputing Centre (SCSC)	romeo - Bull R421-E3 Cluster, Intel Xeon E5-2650v2 8C 2.600GHz, Infiniband FDR, NVIDIA K20x	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

Top 500 « vert » des supercalculateurs

Jun 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 16Brick, 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 32U/256SC Cluster, Intel Xeon E5-2660v2 10C 2.2GHz, Infiniband FDR PEZY-SC CPU, 10 cores/socket, 2.2GHz	32.59
4	5,271.81	GSJ Helmholtz Center	ASUS ESC4000 FDR/Q2S, 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/104Brick 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.	Storm - Cray CS-Storm, Intel Xeon E5-2680v2 10C 2.2GHz, Infiniband FDR Nvidia K40m	44.54
8	3,631.70	Cambridge University	Wilkes - Dell T700 Cluster, Intel Xeon E5-2630v2 6C 2.600GHz, Infiniband FDR NVIDIA K20	52.82
9	3,614.71	TU Dresden, ZIH	Taura - IBM PowerPC S816, Xeon E5-2680v3 12C 2.5GHz, Infiniband FDR Nvidia K80	58.01
10	3,543.32	Financial Institution	Intel Xeon E5-2680v2 10C 2.800GHz, 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 E5-2618L3 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-KFODL - LX 1U-4GPU/104Brick Cluster, Intel Xeon E5-2620v2 6C 2.1GHz, Infiniband FDR NVIDIA Tesla K80	51.13
3	5,271.81	GSJ Helmholtz Center	ASUS ESC4000 FDR/Q2S, 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	Super Cluster W701, 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	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	Total Power(kW)
1	9462.1	NVIDIA Corporation	NVIDIA DGX-1, Xeon E5-2690v4 20C 2.2GHz, Infiniband EDR NVIDIA Tesla P100	349.5
2	7453.5	Swiss National Supercomputing Centre (CSCS)	Cray XC50, Xeon E5-2690v3 12C 2.4GHz, Infiniband FDR NVIDIA Tesla P100 CPU: 8 cores/socket, 2.3GHz	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	6666.6	Computing Center in Wuxi	Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway	15371
5	6000.0	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/S/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 64C 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 !

Top 500 « vert » des supercalculateurs

Nov. 2017

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

Rank	System	Cores	Power (kW)	Power Efficiency (GFlops/Watt)
1	Shoubu system B - ZettaScaler-1.6, Xeon E5-2618L3 8C 2.3GHz, Infiniband EDR PEZY-SC CPU, 8 cores/socket, 2.3GHz	794,400	642.0	17,000
2	Shoubu system B - ZettaScaler-1.6, Xeon E5-2618L3 8C 2.3GHz, Infiniband EDR PEZY-SC CPU, 8 cores/socket, 2.3GHz	742,424	788.2	47 16,759
3	Shoubu system B - ZettaScaler-1.6, Xeon E5-2618L3 8C 2.3GHz, Infiniband EDR PEZY-SC CPU, 8 cores/socket, 2.3GHz	794,400	824.7	58 16,657
4	Shoubu system B - ZettaScaler-1.6, Xeon E5-2618L3 8C 2.3GHz, Infiniband EDR PEZY-SC CPU, 8 cores/socket, 2.3GHz	22,440	1,070.9	97 15,113
5	Shoubu system B - ZettaScaler-1.6, Xeon E5-2618L3 8C 2.3GHz, Infiniband EDR PEZY-SC CPU, 8 cores/socket, 2.3GHz	19,868,000	19,135.8	1,350 14,773
6	Shoubu system B - ZettaScaler-1.6, Xeon E5-2618L3 8C 2.3GHz, Infiniband EDR PEZY-SC CPU, 8 cores/socket, 2.3GHz	135,828	8,125.8	792 13,704
7	Shoubu system B - ZettaScaler-1.6, Xeon E5-2618L3 8C 2.3GHz, Infiniband EDR PEZY-SC CPU, 8 cores/socket, 2.3GHz	23,400	941.0	76 12,681
8	Shoubu system B - ZettaScaler-1.6, Xeon E5-2618L3 8C 2.3GHz, Infiniband EDR PEZY-SC CPU, 8 cores/socket, 2.3GHz	11,712	635.1	68 10,603
9	Shoubu system B - ZettaScaler-1.6, Xeon E5-2618L3 8C 2.3GHz, Infiniband EDR PEZY-SC CPU, 8 cores/socket, 2.3GHz	21,240	1,193.0	114 10,428
10	Shoubu system B - ZettaScaler-1.6, Xeon E5-2618L3 8C 2.3GHz, Infiniband EDR PEZY-SC CPU, 8 cores/socket, 2.3GHz	341,740	19,950.0	2,272 10,398

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	System	Cores	Power (kW)	Power Efficiency (GFlops/Watt)
1	Shoubu system B - ZettaScaler-1.6, Xeon E5-2618L3 8C 2.3GHz, Infiniband EDR PEZY-SC CPU, 8 cores/socket, 2.3GHz	955,280	1,043.0	60 17,454
2	Shoubu system B - ZettaScaler-1.6, Xeon E5-2618L3 8C 2.3GHz, Infiniband EDR PEZY-SC CPU, 8 cores/socket, 2.3GHz	22,440	1,070.9	97 15,113
3	Shoubu system B - ZettaScaler-1.6, Xeon E5-2618L3 8C 2.3GHz, Infiniband EDR PEZY-SC CPU, 8 cores/socket, 2.3GHz	2,977,824	143,500.0	9,783 14,468
4	Shoubu system B - ZettaScaler-1.6, Xeon E5-2618L3 8C 2.3GHz, Infiniband EDR PEZY-SC CPU, 8 cores/socket, 2.3GHz	791,480	19,880.0	1,649 14,423
5	Shoubu system B - ZettaScaler-1.6, Xeon E5-2618L3 8C 2.3GHz, Infiniband EDR PEZY-SC CPU, 8 cores/socket, 2.3GHz	135,828	8,125.8	792 13,704
6	Shoubu system B - ZettaScaler-1.6, Xeon E5-2618L3 8C 2.3GHz, Infiniband EDR PEZY-SC CPU, 8 cores/socket, 2.3GHz	1,572,480	94,450.0	7,438 12,723
7	Shoubu system B - ZettaScaler-1.6, Xeon E5-2618L3 8C 2.3GHz, Infiniband EDR PEZY-SC CPU, 8 cores/socket, 2.3GHz	23,400	941.0	76 12,681
8	Shoubu system B - ZettaScaler-1.6, Xeon E5-2618L3 8C 2.3GHz, Infiniband EDR PEZY-SC CPU, 8 cores/socket, 2.3GHz	19,440	1,018.0	86 11,865
9	Shoubu system B - ZettaScaler-1.6, Xeon E5-2618L3 8C 2.3GHz, Infiniband EDR PEZY-SC CPU, 8 cores/socket, 2.3GHz	143,840	4,325.0	380 11,382
10	Shoubu system B - ZettaScaler-1.6, Xeon E5-2618L3 8C 2.3GHz, Infiniband EDR PEZY-SC CPU, 8 cores/socket, 2.3GHz	178,352	9,000.0	798 11,385

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 !

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**

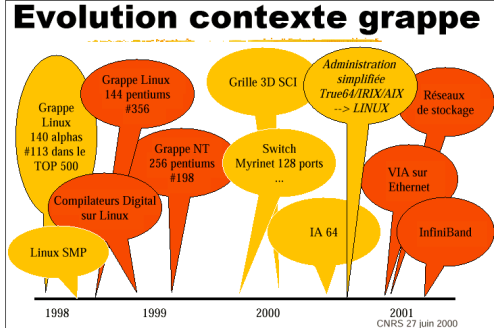
- Configuration d'intranet standard dans une entreprise.
- Travail : surtout du développement de couches systèmes (ingénieurs)
 - ex : reboot complet et rapide, update config rapide, ...
- Puis de la recherche algorithmique (chercheurs)




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

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


Evolution contexte grappe



1998 1999 2000 2001
CNRS 27 juin 2000

40

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 ? »

42

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-habituels !

! 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 ☺

Musée Ferrari - Modène

47

Evolution, marché et gestion des machines parallèles

FIN

48