



L'ordinateur supraconducteur

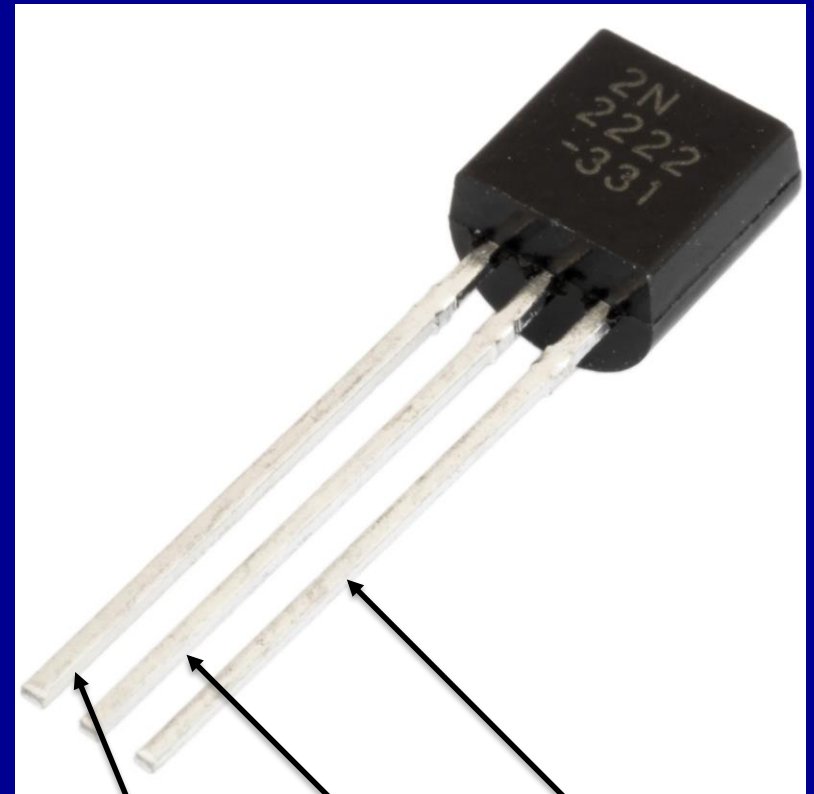
Pascal FEBVRE

Université Savoie Mont Blanc
IMEP-LAHC – CNRS UMR5130

Remerciements :

Akira Fujimaki, Juergen Kunert, Ziad Melhem, Alain Ravex, Georges Waysand

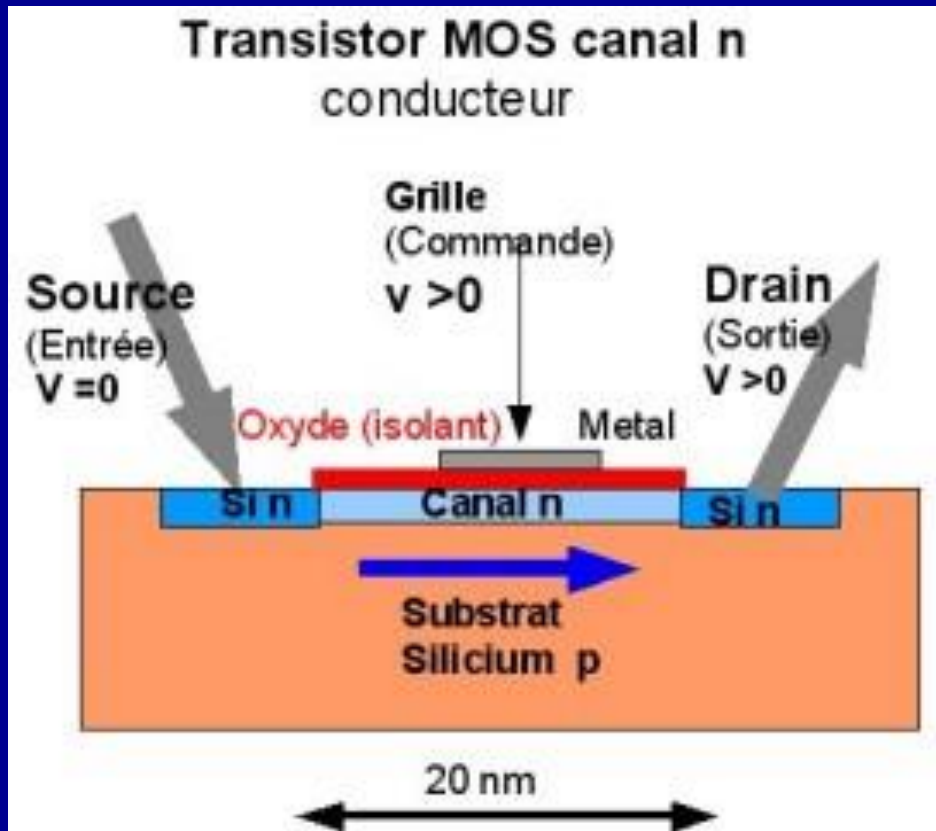
Le transistor



Inventé en 1947 par J. Bardeen, W. Shockley et W. Brattain, chercheurs aux Laboratoires Bell (Prix Nobel de physique 1956)

La microélectronique

source



grille

drain

Le pupitre de commande d'ordinateur IBM 360/67 -1970



le pupitre de commande de /0001 l'ordinateur IBM 360/67

1970

ACONIT 20203

Tableau de contrôle de l'ordinateur IBM 360 de l'IMAG exploité par le CICG (Centre Interuniversitaire de Calcul de Grenoble).
Les ordinateurs de l'époque n'étaient pas très fiables mais ils étaient réparables. Ils contenaient des milliers de cartes enfichables et remplaçables.
Ce tableau permettait de visualiser et éventuellement modifier pratiquement toutes les variables de l'unité centrale pour surveiller ou configurer l'ordinateur ou repérer les pannes.

fabriqué en 1965

volume : 8 m³

mémoire : 128 Ko (1965) à 4 096 Ko (1969)

Source : Exposition bâtiment IMAG - Campus universitaire de Grenoble

L'ordinateur massivement parallèle - 1992



l'ordinateur massivement parallèle DEC MPP 12000

1992

ACONIT 12169

Les machines parallèles sont nées d'une idée simple : accélérer l'exécution d'un programme informatique en le découpant en plusieurs tâches pouvant s'exécuter sur plusieurs processeurs.

La machine présentée contient 512 processeurs MasPar, répartis sur 16 cartes de 32 processeurs.

Cette machine était utilisée au CHU de Grenoble pour des études d'imagerie médicale. Elle était pilotée à partir d'un mini-ordinateur DEC Station 500 qui gérât les entrées/sorties et les unités de disque.

16 cartes de 32 processeurs 4 bits
fréquence d'horloge : 12,5 MHz
technologie CMOS 1 μ m

Source : Exposition bâtiment IMAG - Campus universitaire de Grenoble

Supercalculateurs pour l'astronomie

ALMA : 66 antennes au Chili (54 x 12m diam. ; 12 x 7m diam.)

SKA : environ 3000 antennes de 15 mètres de diamètre



Atacama Large Millimeter Array (ALMA) - Source : ESO

Spectromètre auto-corrélateur de ALMA



L'un des quatre quadrants de l'autocorrélateur ALMA – Source : ESO

Serveurs Facebook - Lulea - Suède



Serveurs Google - The Dalles - Oregon



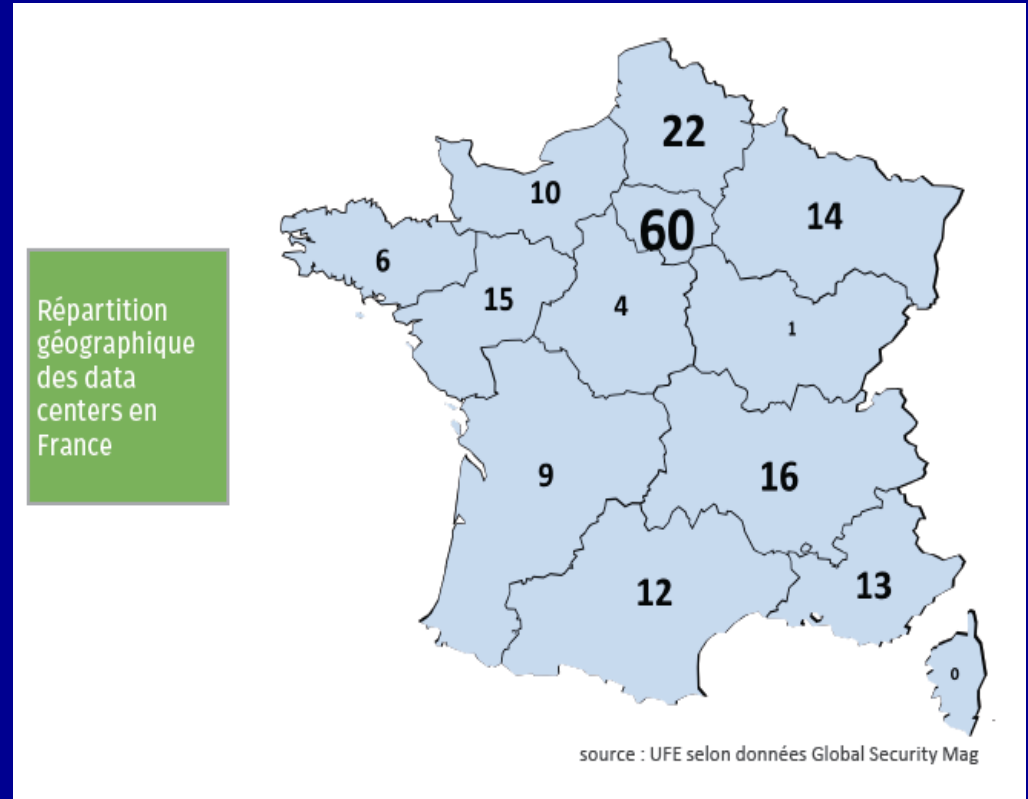
Copyright Google

Centres de données

Les serveurs de données (data centres) en France

La facture énergétique de ces sites (électricité pour le traitement des données mais aussi pour la climatisation) constitue en moyenne 40% de leurs coûts de fonctionnement.

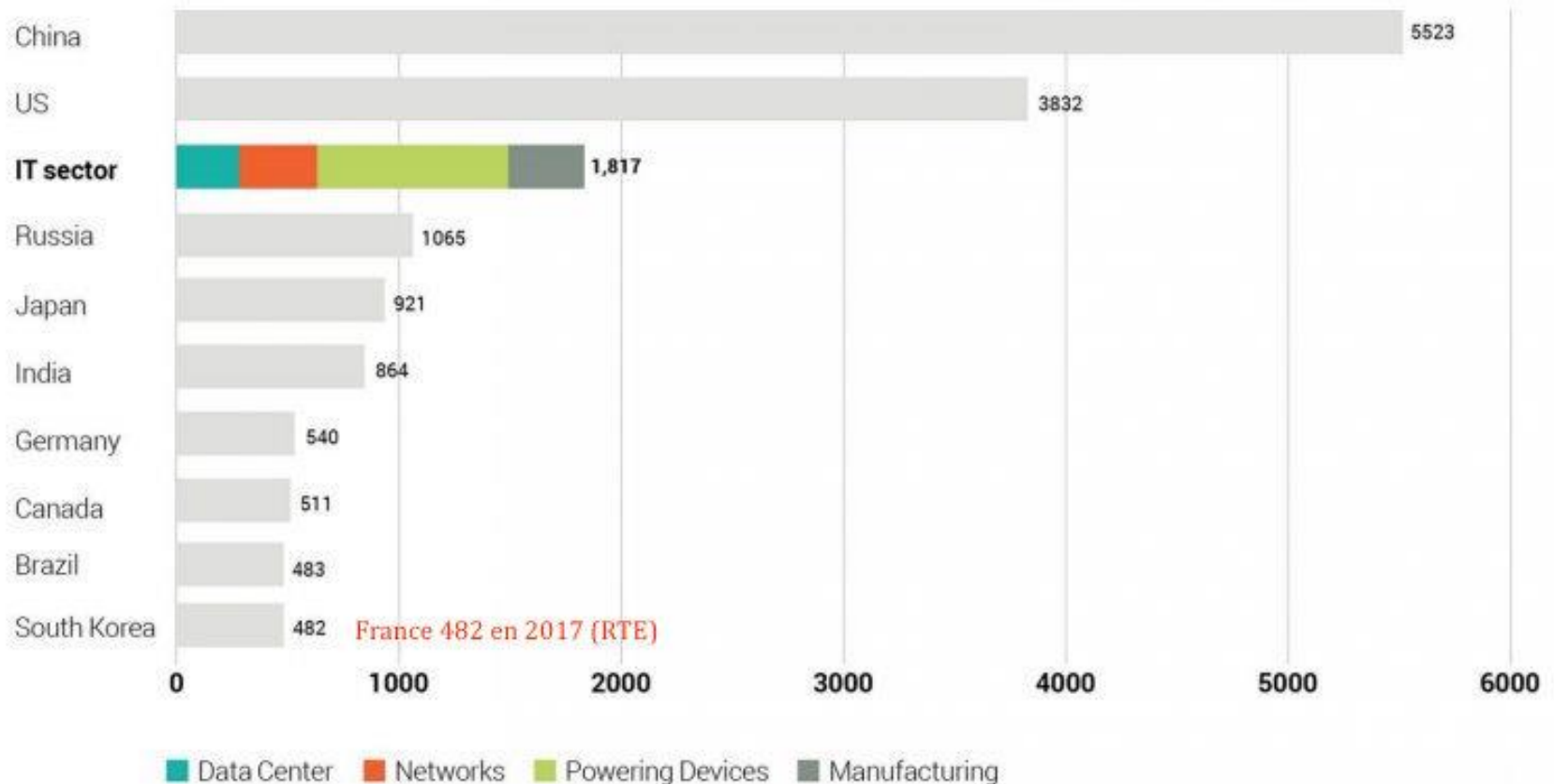
**9% de la
consommation
électrique en France
due aux serveurs de
données agés de
douze ans en
moyenne**



L'envoi d'un mail avec pièce jointe de 1 Mo transitant sur différents serveurs entraîne à lui seul une consommation électrique aussi importante qu' une ampoule basse consommation de 20 watts allumée pendant une heure.

Le problème de la consommation énergétique

2012 Electricity Consumption; Countries Compared to IT Sector in billion kWh



Source: *Emerging Trends in Electricity Consumption for Consumer ICT*, Peter Corcoran and Andres Andrae (2013) and *CIA World Factbook*. China/Russia/Canada figures are from 2014.

Si Internet était un pays



En 2015 les serveurs de données représentaient **4% de la production mondiale** d'énergie, en croissance de **5% par an** (source RTE)

Impact écologique



<https://www.fournisseur-energie.com/>

Monnaies virtuelles

L'énergie d'une transaction peut évaporer entre 700 et 1400 litres d'eau.



La monnaie virtuelle produit autant de CO2 par an qu'un million de vols transatlantiques



Photo: Stefan Chow

Harsh Conditions: Inner Mongolia has some of the cheapest electricity prices in the world (4 U.S. cents per kilowatt-hour, a government-reduced rate), which is the primary reason miners are setting up shop here. But it comes with a trade-off: The climate outside Bitmain's warehouses can be brutal, especially in the summer.

<https://www.theguardian.com/technology/2018/jan/17/bitcoin-electricity-usage-huge-climate-cryptocurrency>

Besoin de supercalculateurs avec performances accrues

La demande de machines à haute performance et de serveurs va continuer d'augmenter. Nous avons besoin :

- de meilleurs modèles prédictifs pour le changement climatique et la prévision météorologique ;
- de mieux comprendre la formation de l'Univers ;
- de comprendre la physique sub-atomique ;
- de modéliser des cellules, pour la génétique, les biotechnologies ;
- de simuler les fonctions du cerveau, ...



Mais la puissance requise ne peut pas augmenter au même rythme

Loi de Moore : le nombre de transistors par unité de surface double environ tous les deux ans



125mm² (11mm x 11 mm)

Technologie : 16 nm

Fréquence d'horloge : 2,34 GHz

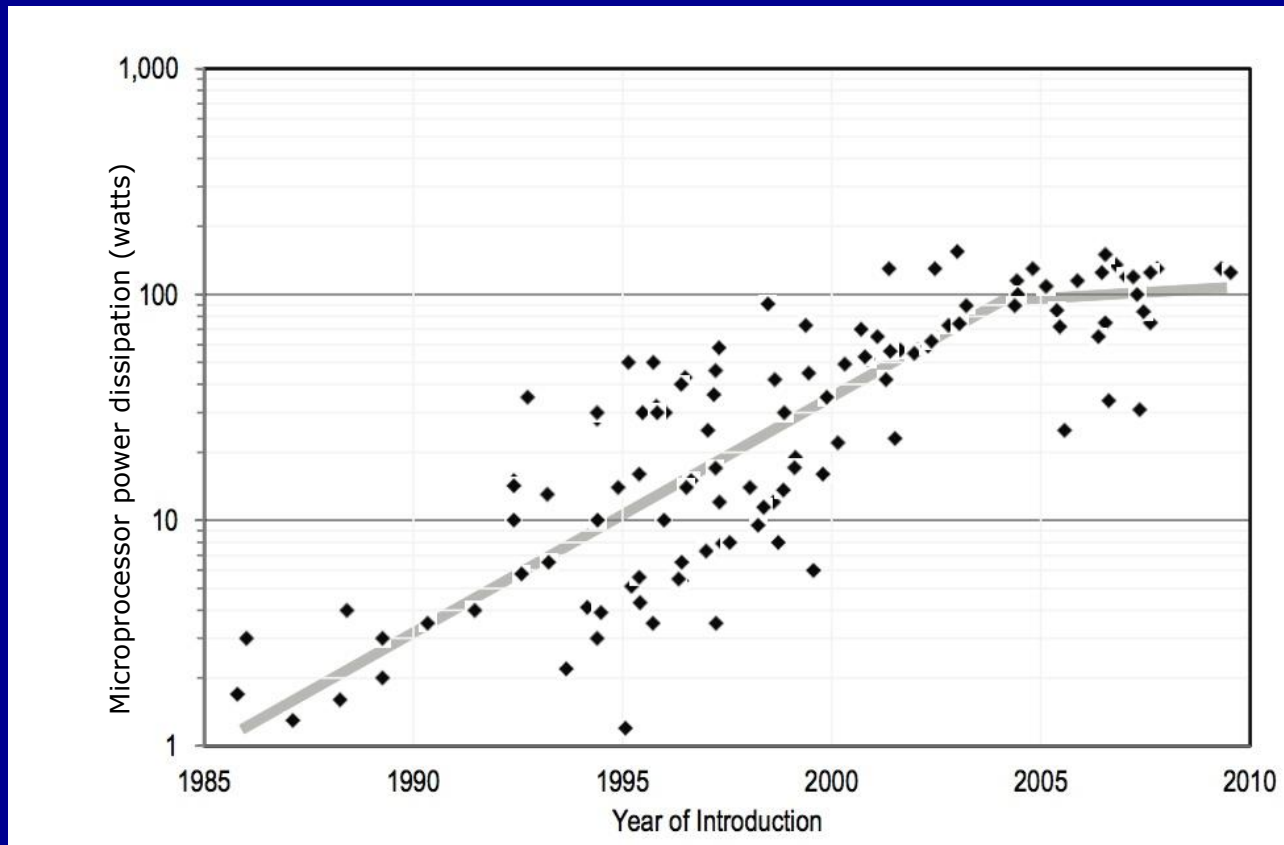
2,64 milliards de transistors/cm²



Loi d'échelle de Dennard

Une réduction comparable de la puissance de chaque composant est réalisée pour garder constante la puissance dissipée par la puce.

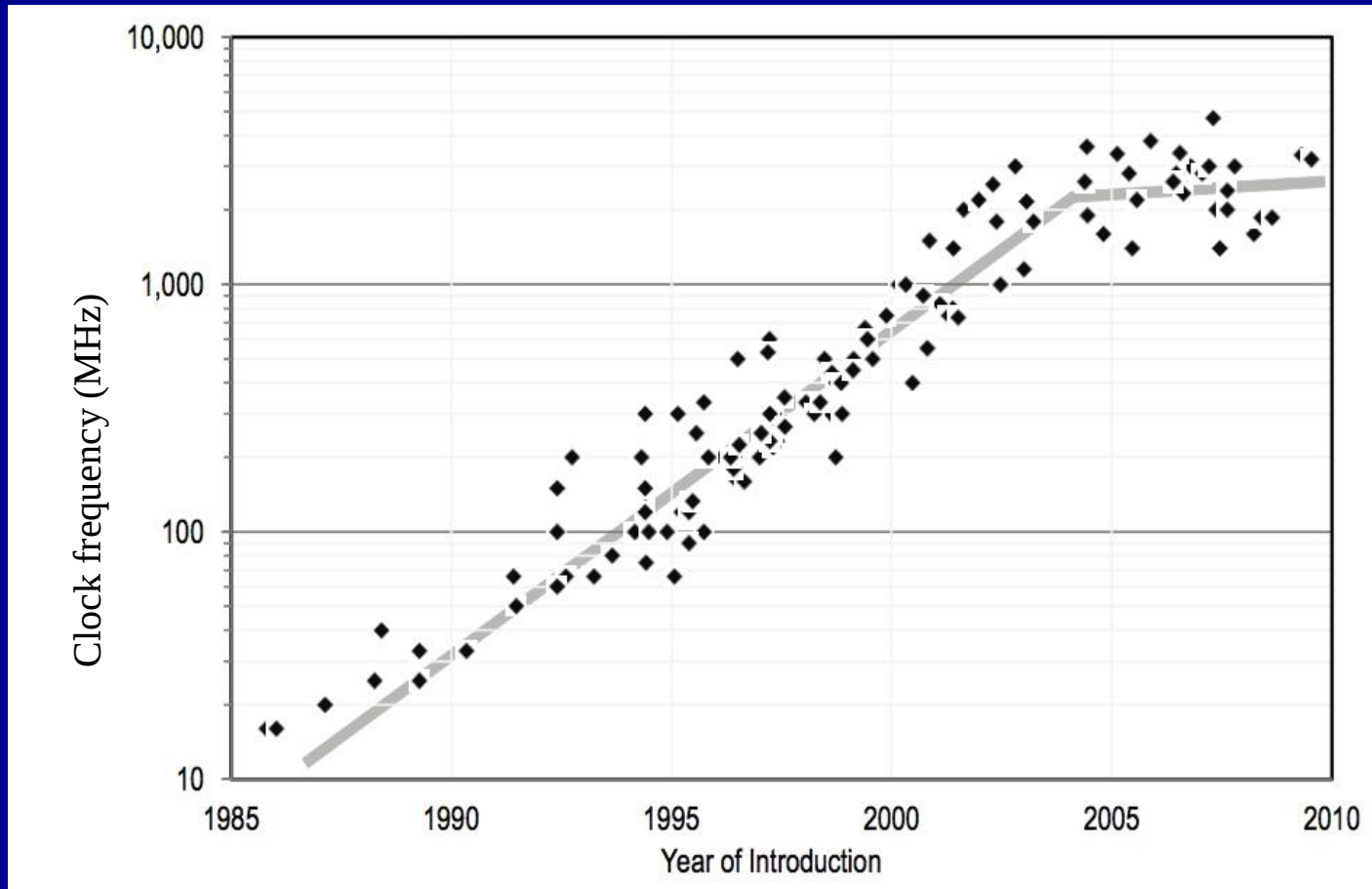
1985 : 1 watt/cm² 2016 : 145 watts/cm²



Source : THE FUTURE OF COMPUTING PERFORMANCE - Game Over or Next Level? Copyright 2011 by the National Academy of Sciences of the USA

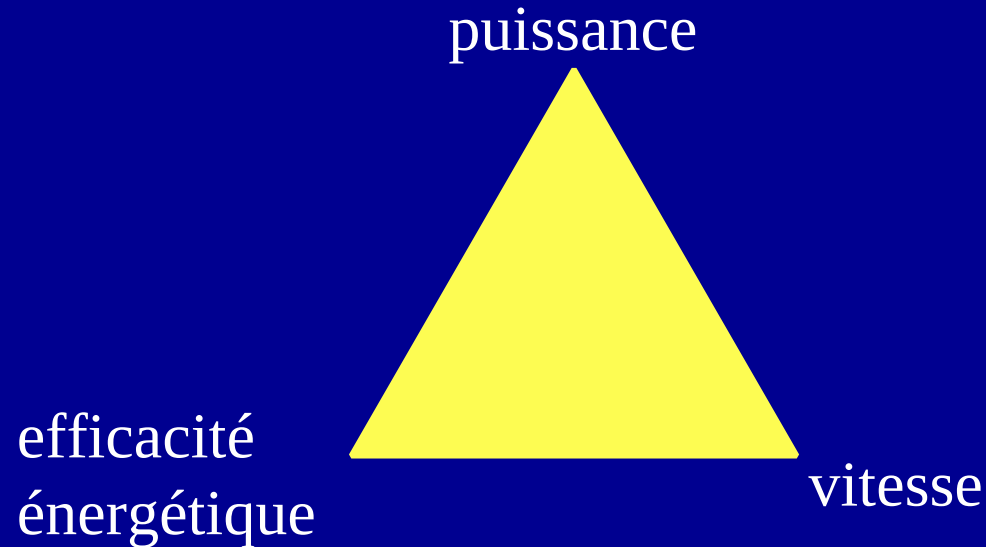
Fréquence d'horloge des microprocesseurs

Les fréquences des microprocessors sont passées d'environ 10 MHz en 1985 à 3 GHz en 2005 : augmentation de 40% chaque année pendant deux décennies.



Source : THE FUTURE OF COMPUTING PERFORMANCE - Game Over or Next Level? Copyright 2011 by the National Academy of Sciences of the USA

Métrique pour comparer les technologies



Définitions

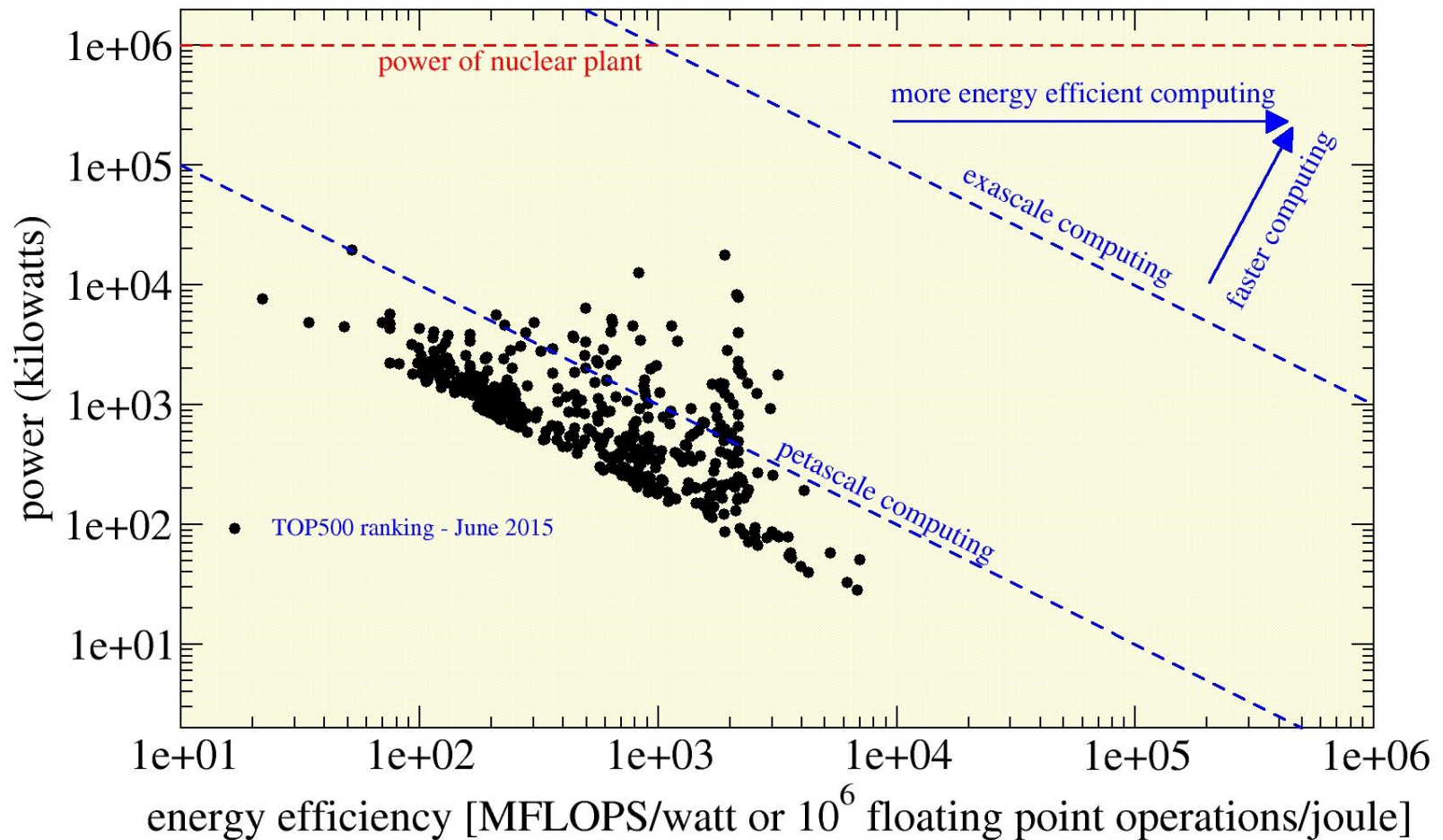
- FLOP : FLoating Point Operation (opérations en virgule flottante)
- efficacité énergétique = nombre de FLOPs par joule
- vitesse (fréquence d'horloge) = nombre de FLOPs par seconde

$$\text{vitesse} = \text{puissance} * \text{efficacité énergétique}$$

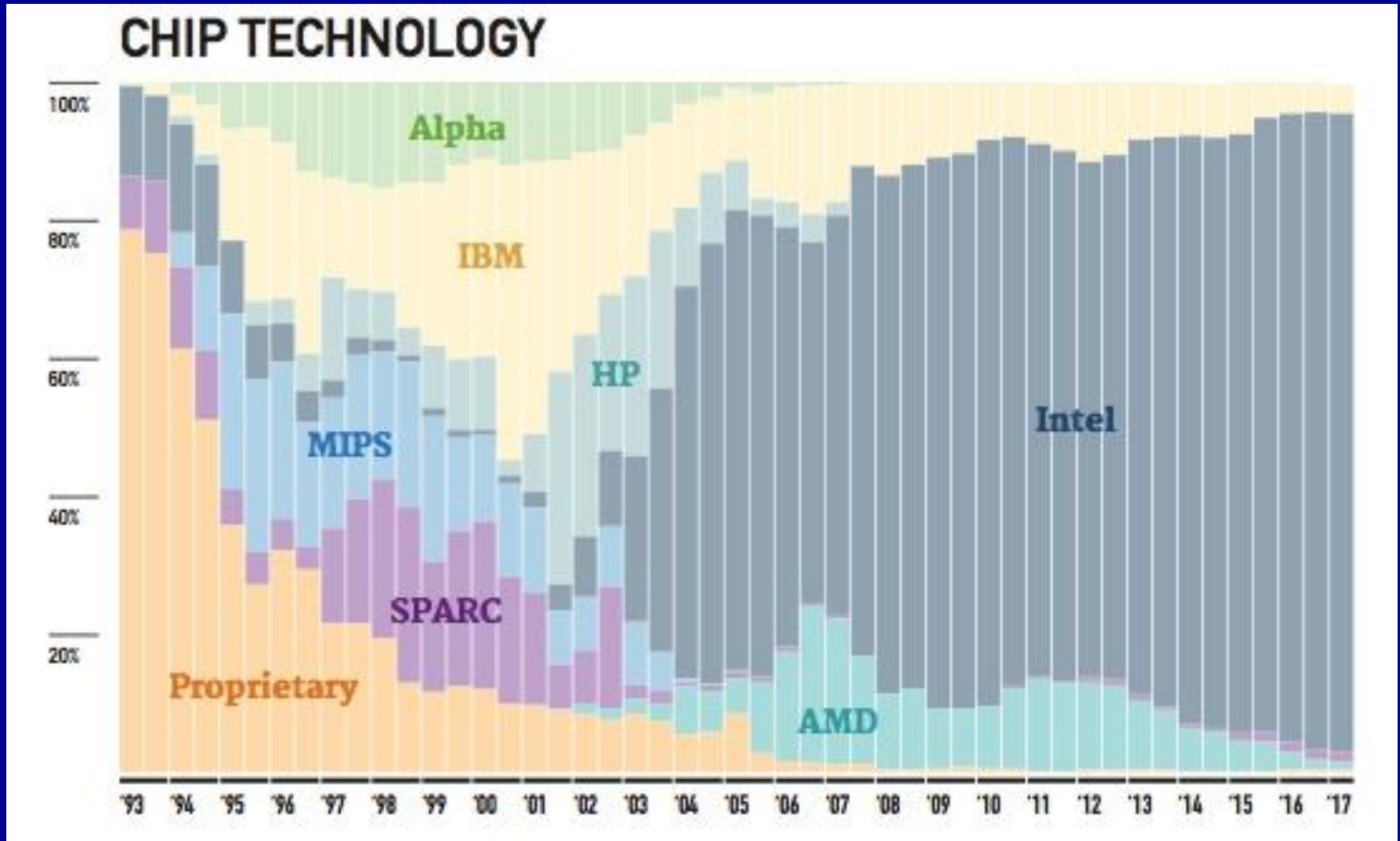
Etat de l'art des supercalculateurs semi-conducteurs

Performance and power of high-end computing

September 2015 - total power consumption : 0.61 GW



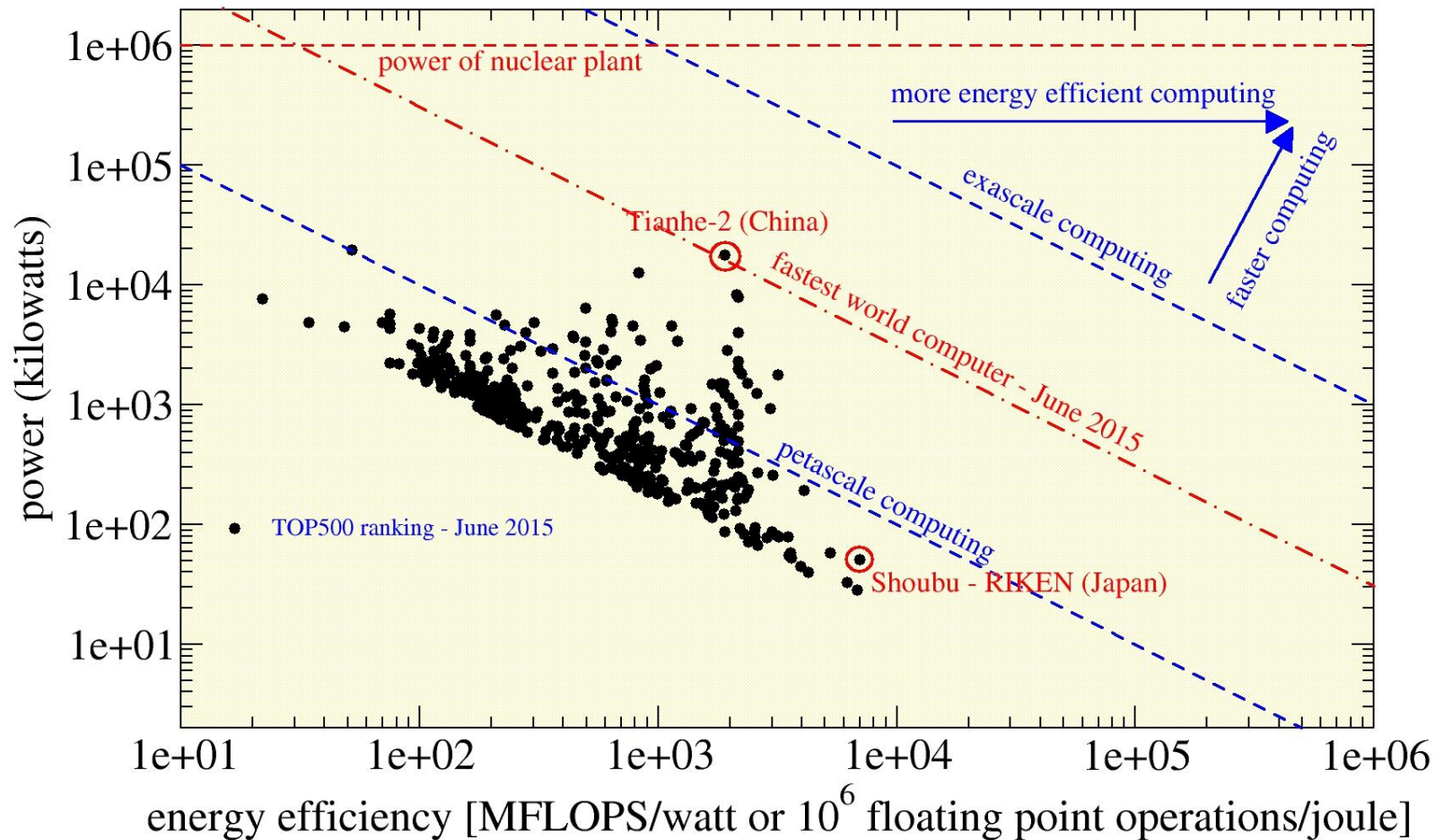
Fabricants de puces semi-conductrices



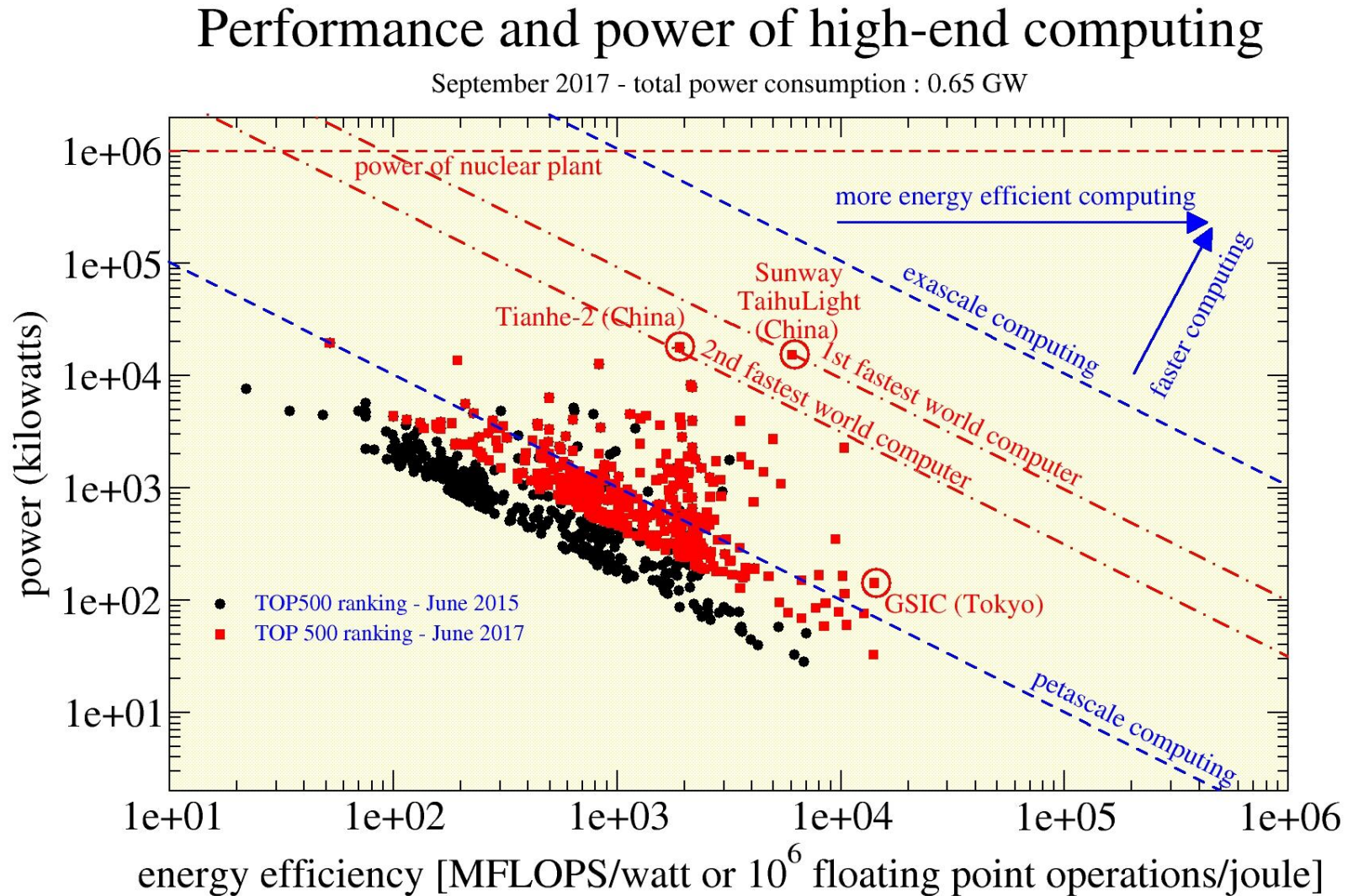
Etat de l'art des supercalculateurs semi-conducteurs

Performance and power of high-end computing

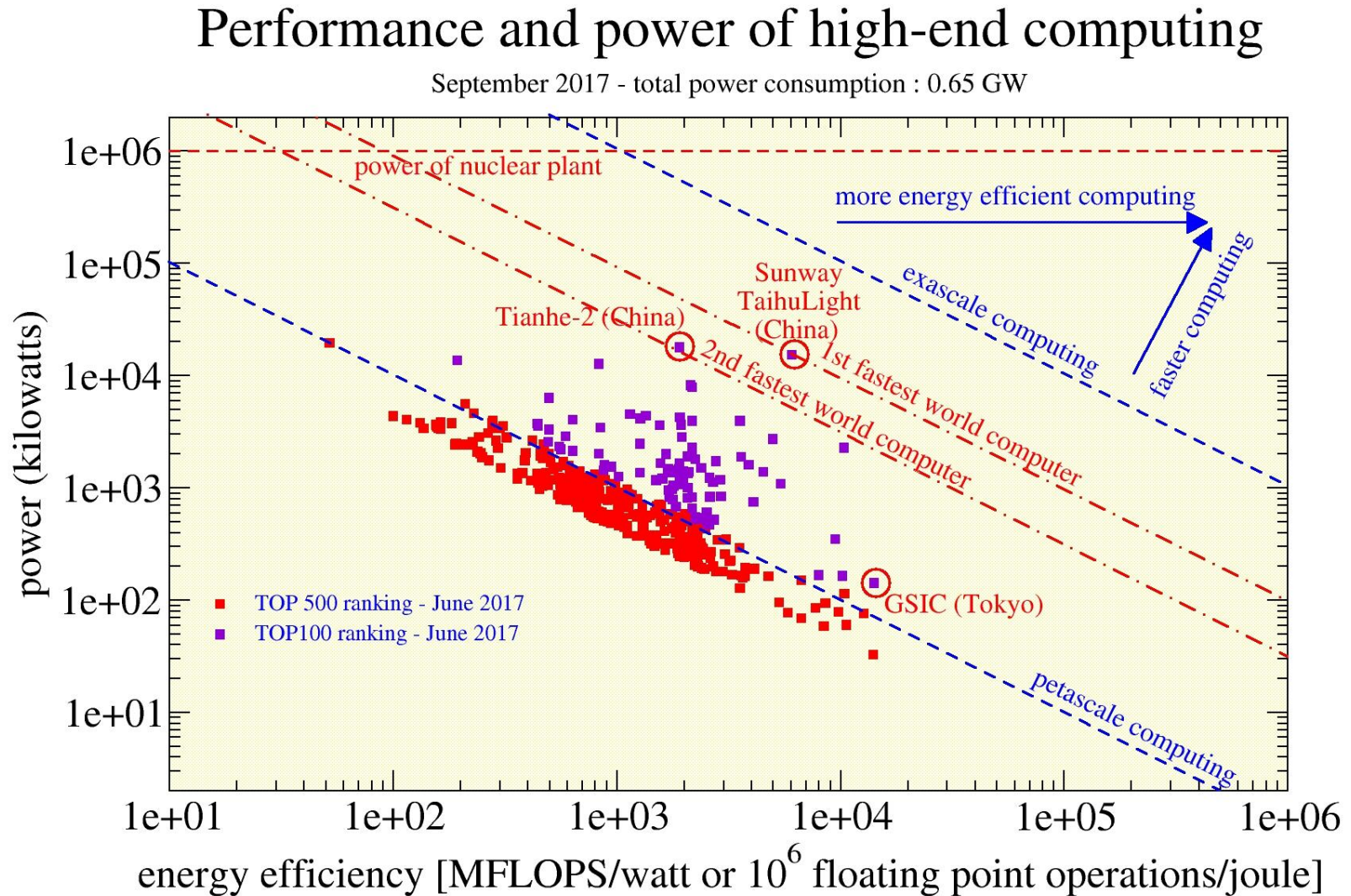
September 2015 - total power consumption : 0.61 GW



Etat de l'art des supercalculateurs semi-conducteurs



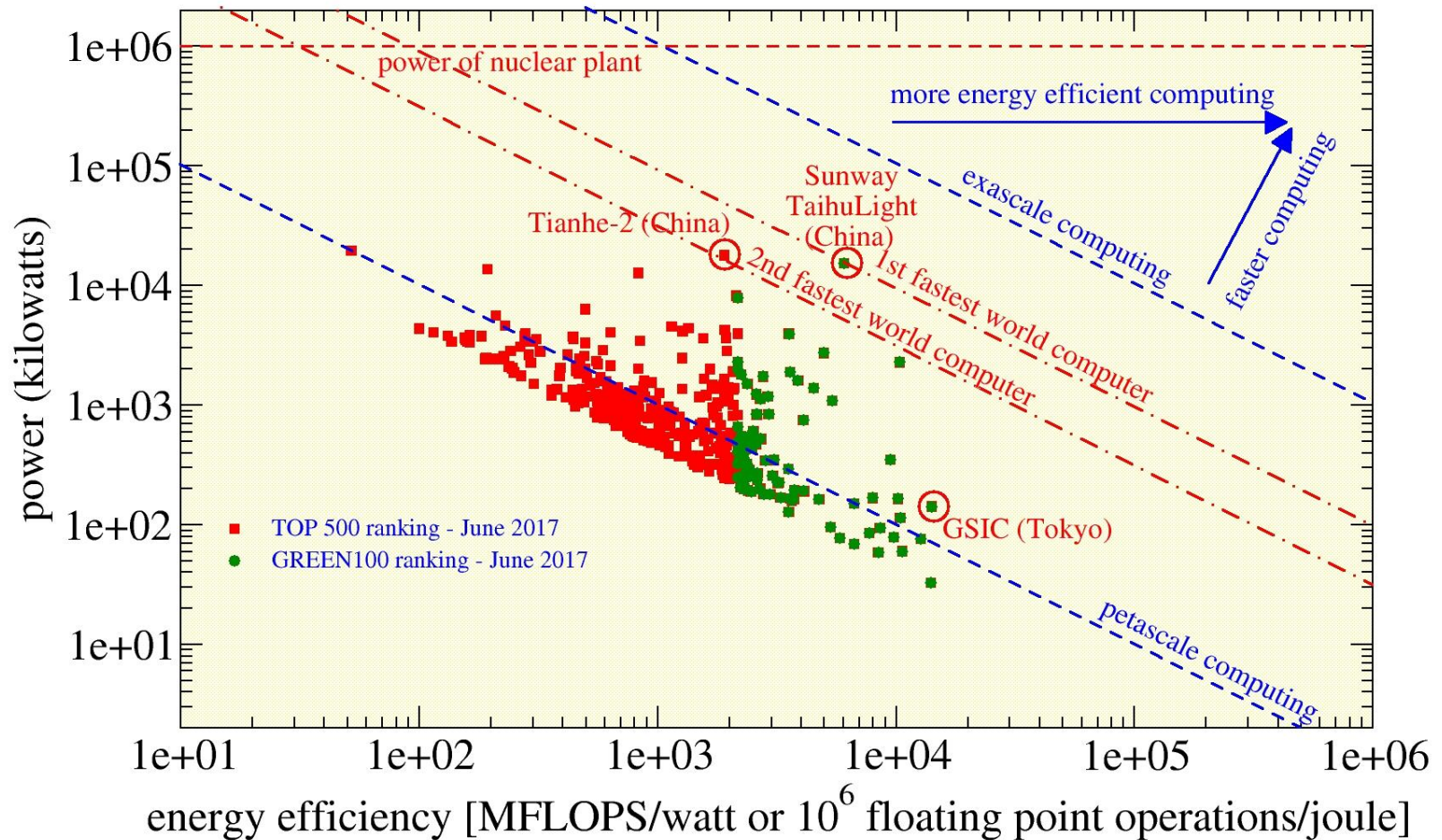
Etat de l'art des supercalculateurs semi-conducteurs



Etat de l'art des supercalculateurs semi-conducteurs

Performance and power of high-end computing

September 2017 - total power consumption : 0.65 GW



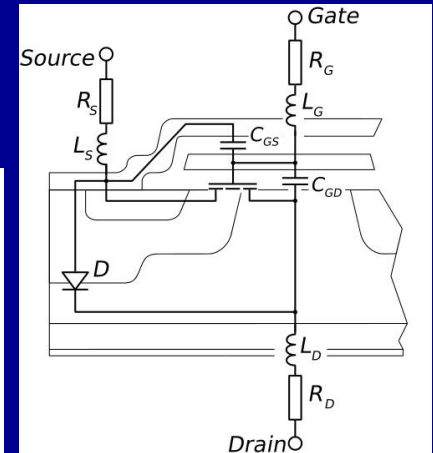
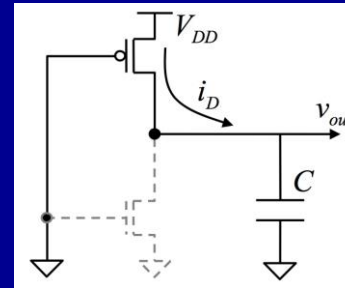
Semiconducteurs : le produit énergie-délai

Semiconducteurs : la puissance dynamique est la quantité limitante :

- P_{dd} est la puissance dynamique
- V_{dd} est la tension de polarisation
- C est la capacité intrinsèque de la porte

$$P_{dd} = CV_{dd}^2 f$$

$$\tau = \frac{CV_{dd}}{I_d}$$



Délai intrinsèque de la porte :

- I_d est le courant de saturation du drain

réduire la taille du transistor
réduire la tension

Le produit énergie-délai (EDP) est :

$$EDP = \frac{\tau P_{dd}}{f} = \frac{CV_{dd}}{I_d} CV_{dd}^2 = C V_{dd}^3 / I_d$$

$$EDP = 1/(\text{efficacité énergétique}^2 * \text{puissance}) = \text{puissance}/\text{vitesse}^2$$

Semiconducteurs : projections pour la prochaine décennie

Summary Table of ITRS Technology Trend Targets

Year of Production	2013	2015	2017	2019	2021	2023	2025	2028
Logic Industry "Node Name" Label	"16/14"	"10"	"7"	"5"	"3.5"	"2.5"	"1.8"	
Logic ½ Pitch (nm)	40	32	25	20	16	13	10	7
Flash ½ Pitch [2D] (nm)	18	15	13	11	9	8	8	8
DRAM ½ Pitch (nm)	28	24	20	17	14	12	10	7.7
FinFET Fin Half-pitch (new) (nm)	30	24	19	15	12	9.5	7.5	5.3
FinFET Fin Width (new) (nm)	7.6	7.2	6.8	6.4	6.1	5.7	5.4	5.0
6-t SRAM Cell Size(um ²) [@60f2]	0.096	0.061	0.038	0.024	0.015	0.010	0.0060	0.0030
MPU/ASIC HighPerf 4t NAND Gate Size(um ²)	0.248	0.157	0.099	0.062	0.039	0.025	0.018	0.009
4-input NAND Gate Density (Kgates/mm) [@155f2]	4.03E+03	6.37E+03	1.01E+04	1.61E+04	2.55E+04	4.05E+04	6.42E+04	1.28E+05
Flash Generations Label (bits per chip) (SLC/MLC)	64G /128G	128G /256G	256G / 512G	512G / 1T	512G / 1T	1T / 2T	2T / 4T	4T / 8T
Flash 3D Number of Layer targets (at relaxed Poly half pitch)	16-32	16-32	16-32	32-64	48-96	64-128	96-192	192-384
Flash 3D Layer half-pitch targets (nm)	64nm	54nm	45nm	30nm	28nm	27nm	25nm	22nm
DRAM Generations Label (bits per chip)	4G	8G	8G	16G	32G	32G	32G	32G
450mm Production High Volume Manufacturing Begins (100Kwspm)				2018				
Vdd (High Performance, high Vdd transistors)[**]	0.86	0.83	0.80	0.77	0.74	0.71	0.68	0.64
1/(CV/I) (1/psec) [**]	1.13	1.53	1.75	1.97	2.10	2.29	2.52	3.17
On-chip local clock MPU HP [at 4% CAGR]	5.50	5.95	6.44	6.96	7.53	8.14	8.8	9.9
Maximum number wiring levels [unchanged]	13	13	14	14	15	15	16	17
MPU High-Performance (HP) Printed Gate Length (GLpr) (nm) [**]	28	22	18	14	11	9	7	5
MPU High-Performance Physical Gate Length (GLph) (nm) [**]	20	17	14	12	10	8	7	5
ASIC/Low Standby Power (LP) Physical Gate Length (nm) (GLph)[**]	23	19	16	13	11	9	8	6

**** Note:** from the PIDS working group data; however, the calibration of Vdd, GLph, and I/CV is ongoing for improved targets in 2014 ITRS work

Source : International Technology Roadmap for Semiconductors – 2013 edition – Executive summary

Produit énergie-délai : projection

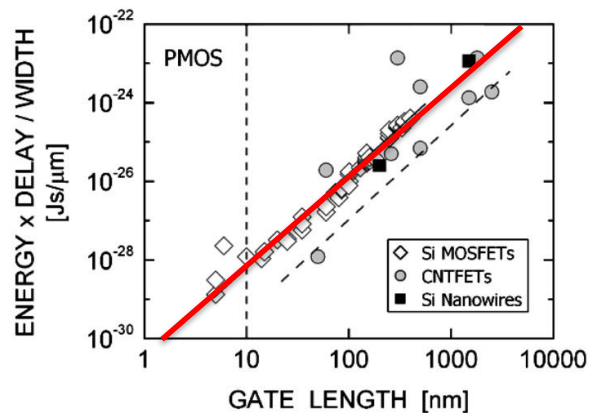


Fig. 6. Energy-delay product per device width versus transistor physical gate length of PMOS transistors.

Source : Robert Chau et al, IEEE Trans. Nanotechnology, Vol.. 4, No. 2, March 2005

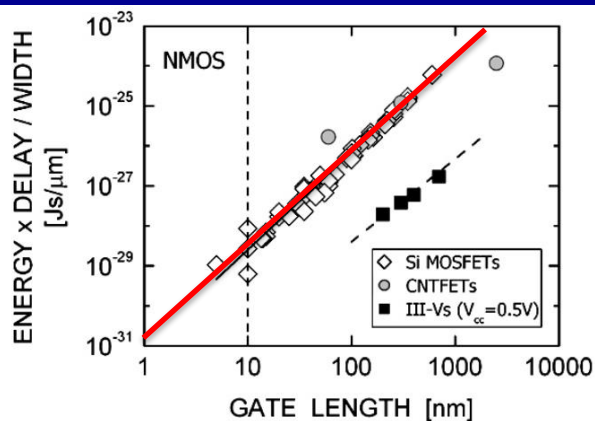
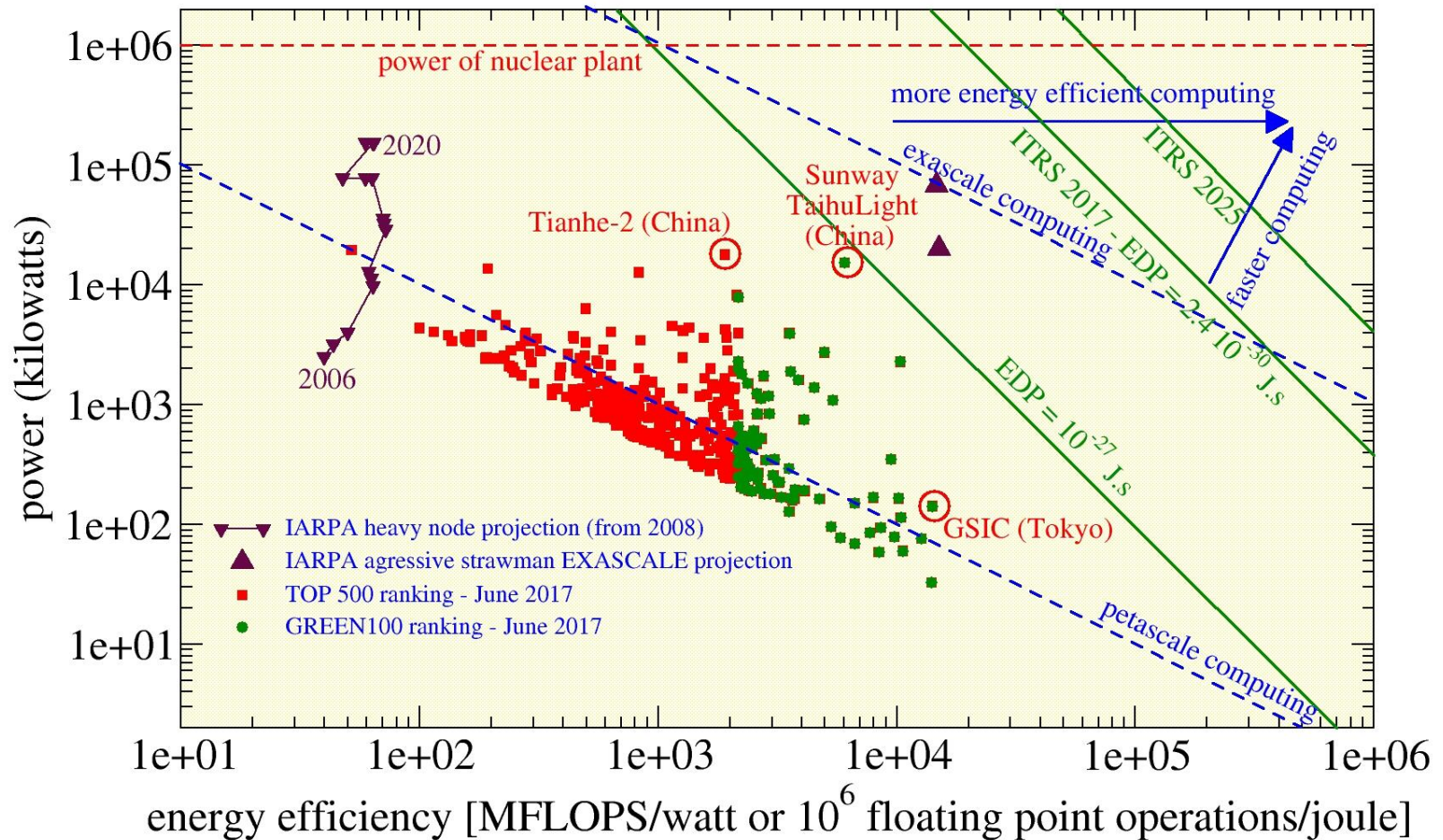


Fig. 7. Energy-delay product per device width versus transistor physical gate length of NMOS transistors.

Year	gate length	EDP (J.s/μm)	EDP (J.s)
2013	20	3,9E-28	7,9E-30
2015	17	2,7E-28	4,6E-30
2017	14	1,7E-28	2,4E-30
2019	12	1,2E-28	1,5E-30
2021	10	8,0E-29	8,0E-31
2023	8	4,8E-29	3,8E-31
2025	7	3,5E-29	2,5E-31
2028	5	1,6E-29	8,1E-32
2040	1	4,0E-31	4,0E-34

Performance and power of high-end computing

September 2017 - total power consumption : 0.65 GW



La jonction Josephson



La jonction Josephson est l'équivalent du transistor pour l'électronique supraconductrice

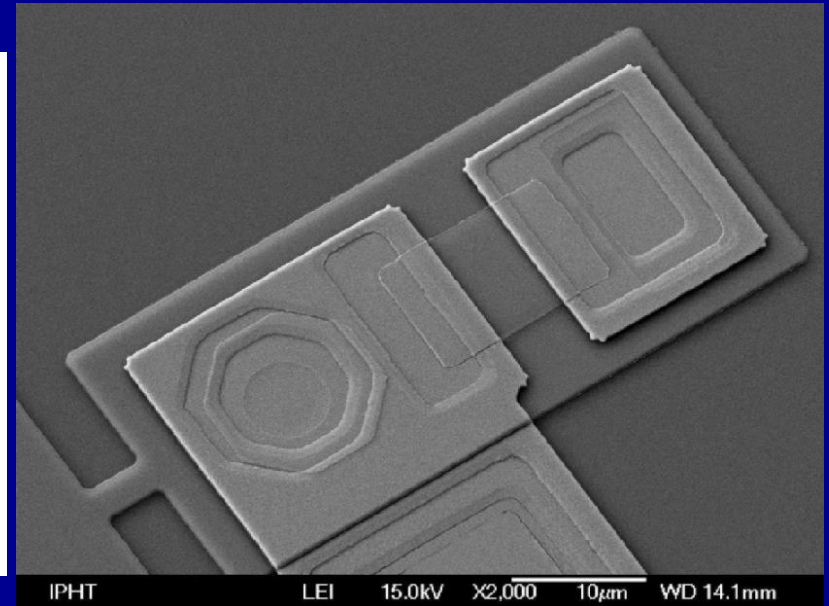
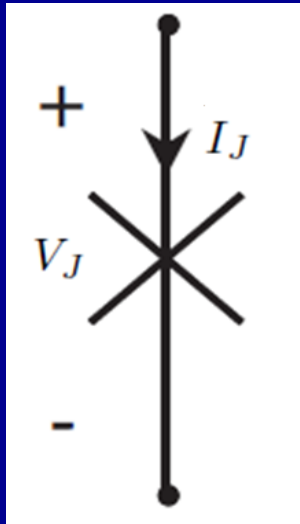
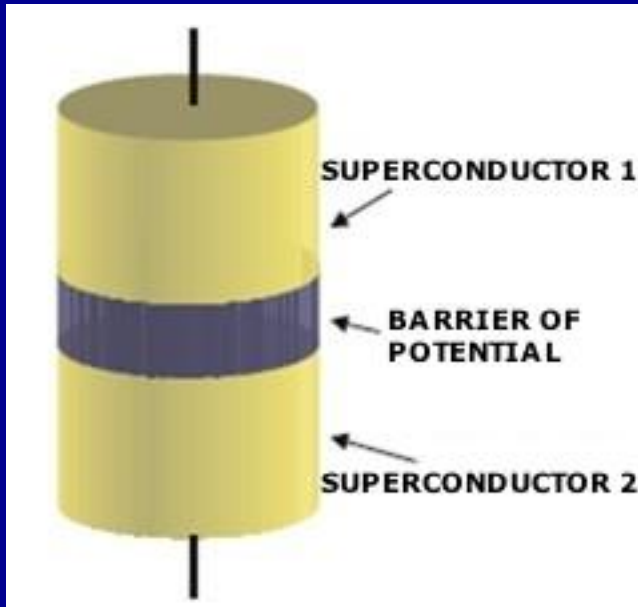
Brian Josephson

En 1962, Brian Josephson prédit qu'un super-courant peut circuler par effet tunnel à travers une barrière isolante séparant deux matériaux supraconducteurs.

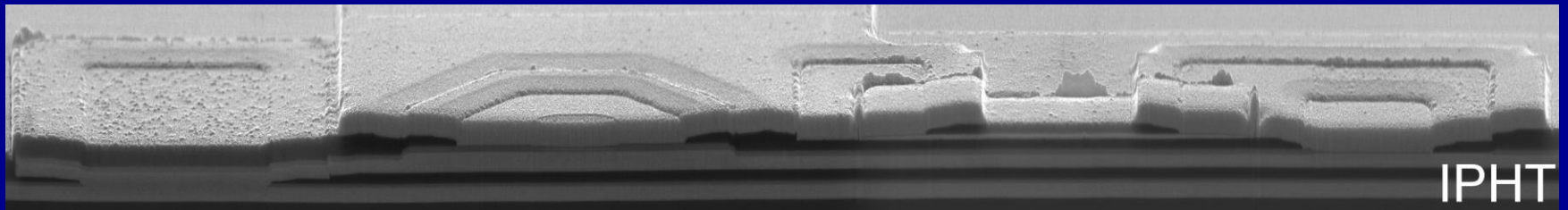
C'est la naissance du composant de base de l'électronique supraconductrice: la jonction Josephson. Il démontre également deux équations régissant le fonctionnement de celle-ci. Il obtiendra le prix Nobel de physique en 1973.

La jonction Josephson

La jonction Josephson : élément actif en électronique supraconductrice

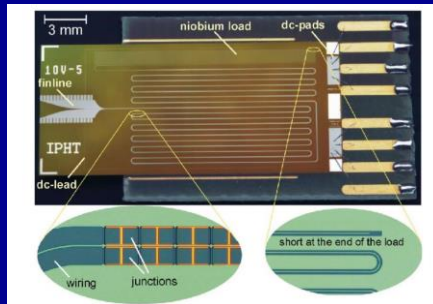


Matériaux les plus utilisés : Nb/Al-AlOx/Nb @ 4,2 K

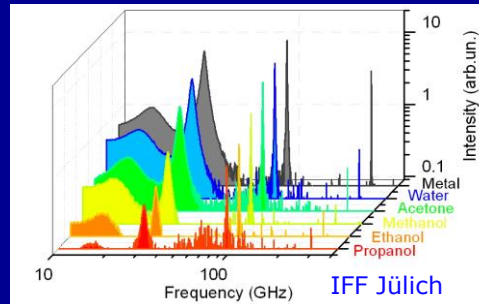


3 μm Photographies de la FLUXONICS Foundry - IPHT Iéna - Allemagne

Les applications de la jonction Josephson



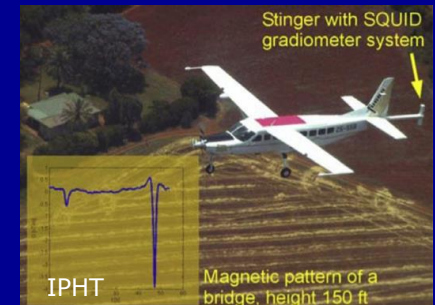
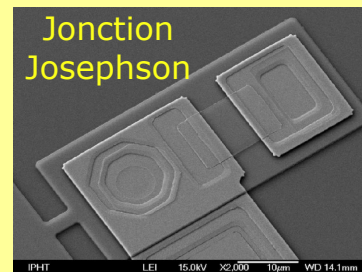
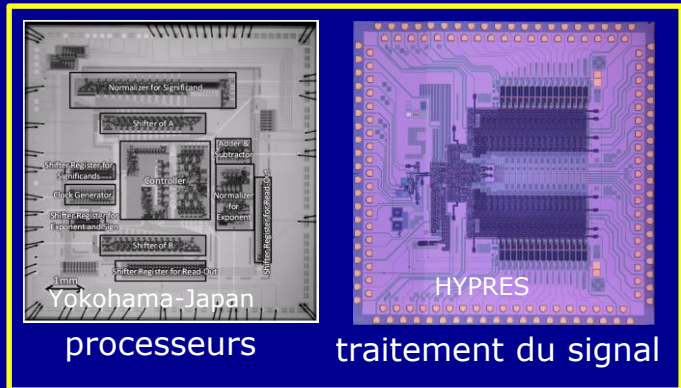
métrologie



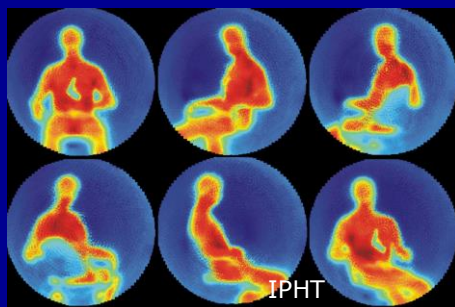
spectroscopie



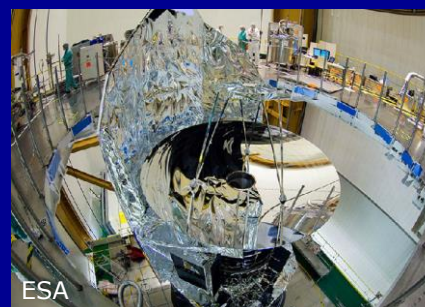
magnéto-encéphalographie



géophysique



sécurité



radio-astronomie

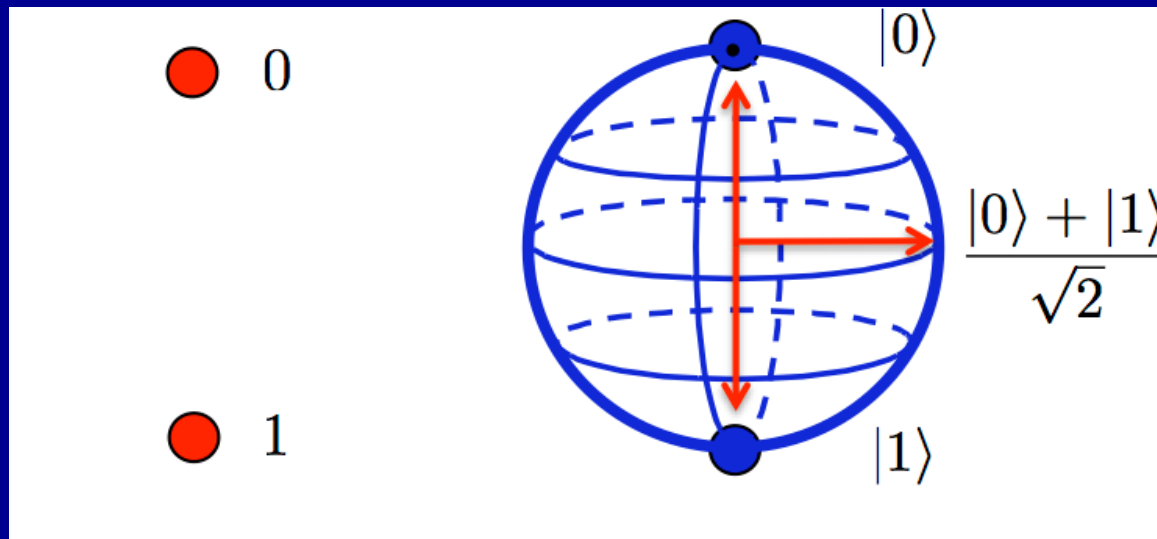


imagerie par champ magnétique

Electronique supraconductrice et calcul quantique

Deux technologies reposent sur la même technologie à base de jonctions Josephson utilisant des matériaux supraconducteurs :

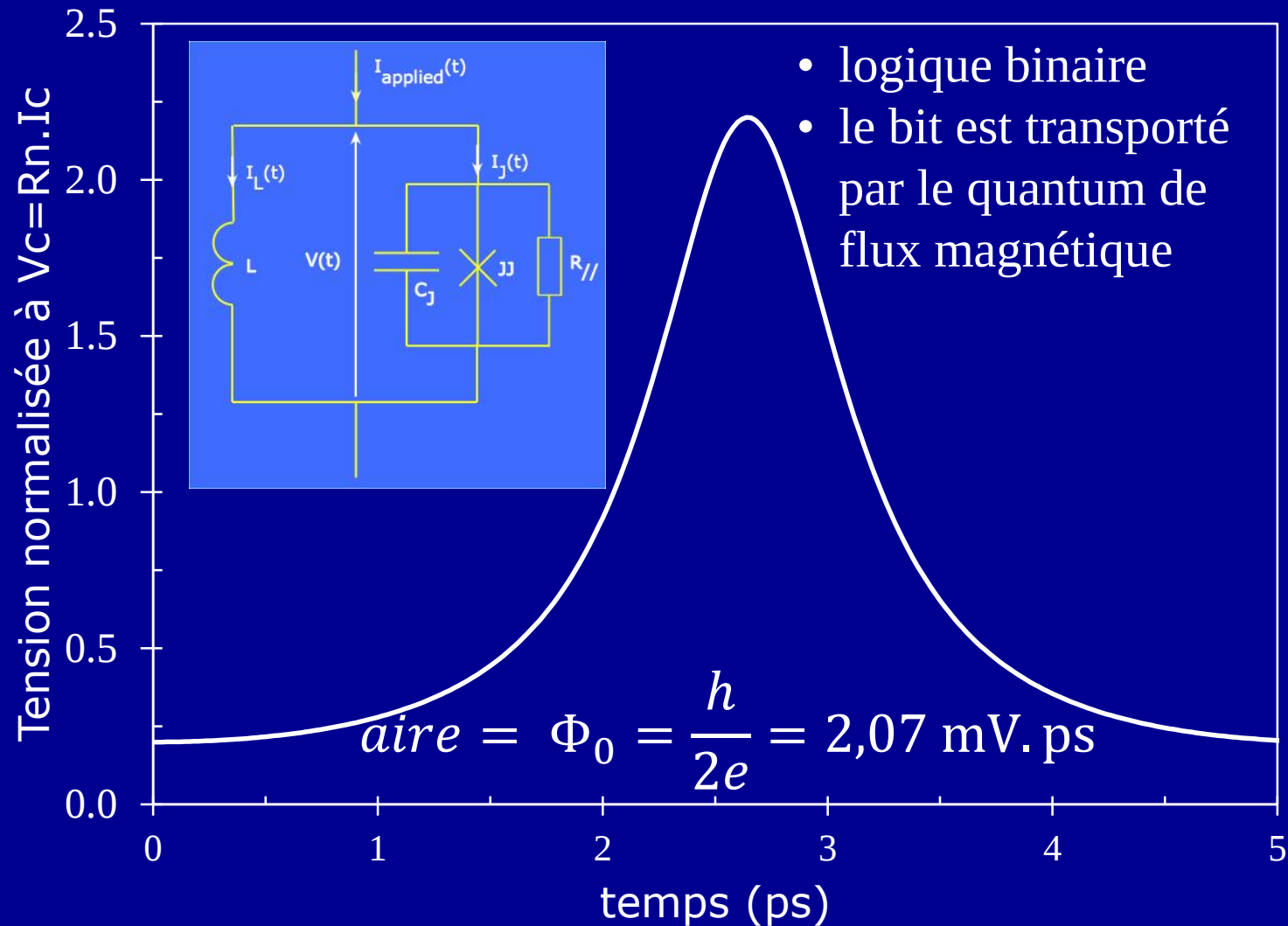
- électronique supraconductrice binaire
- le calcul quantique utilisant la superposition quantique des états "0" et "1"



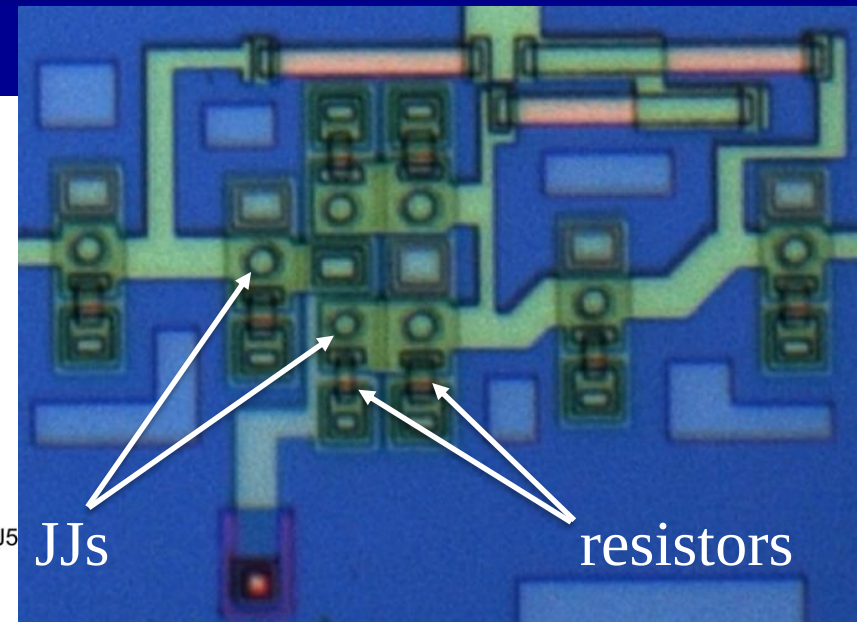
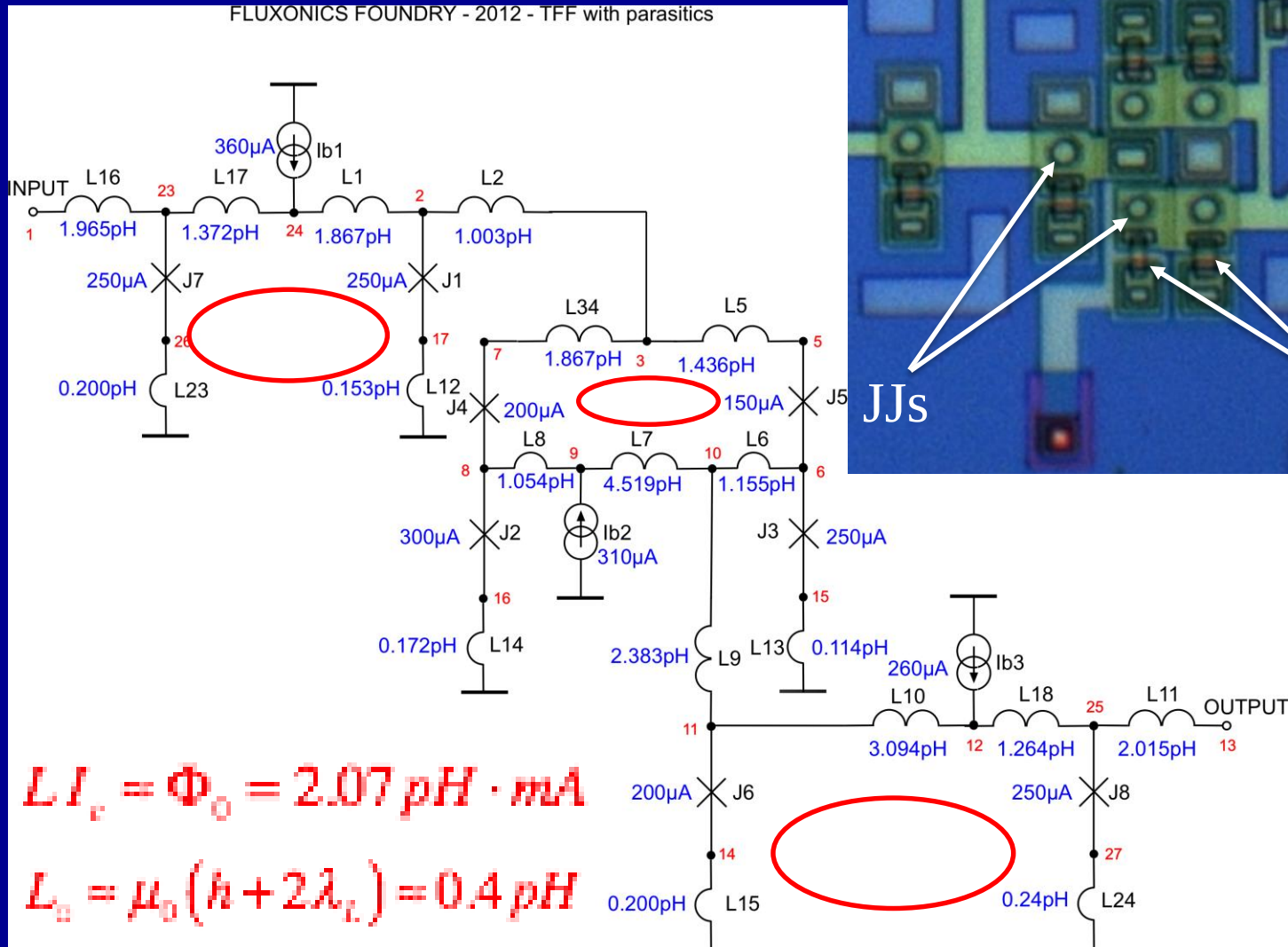
bit classique

bit quantique ou qu-bit

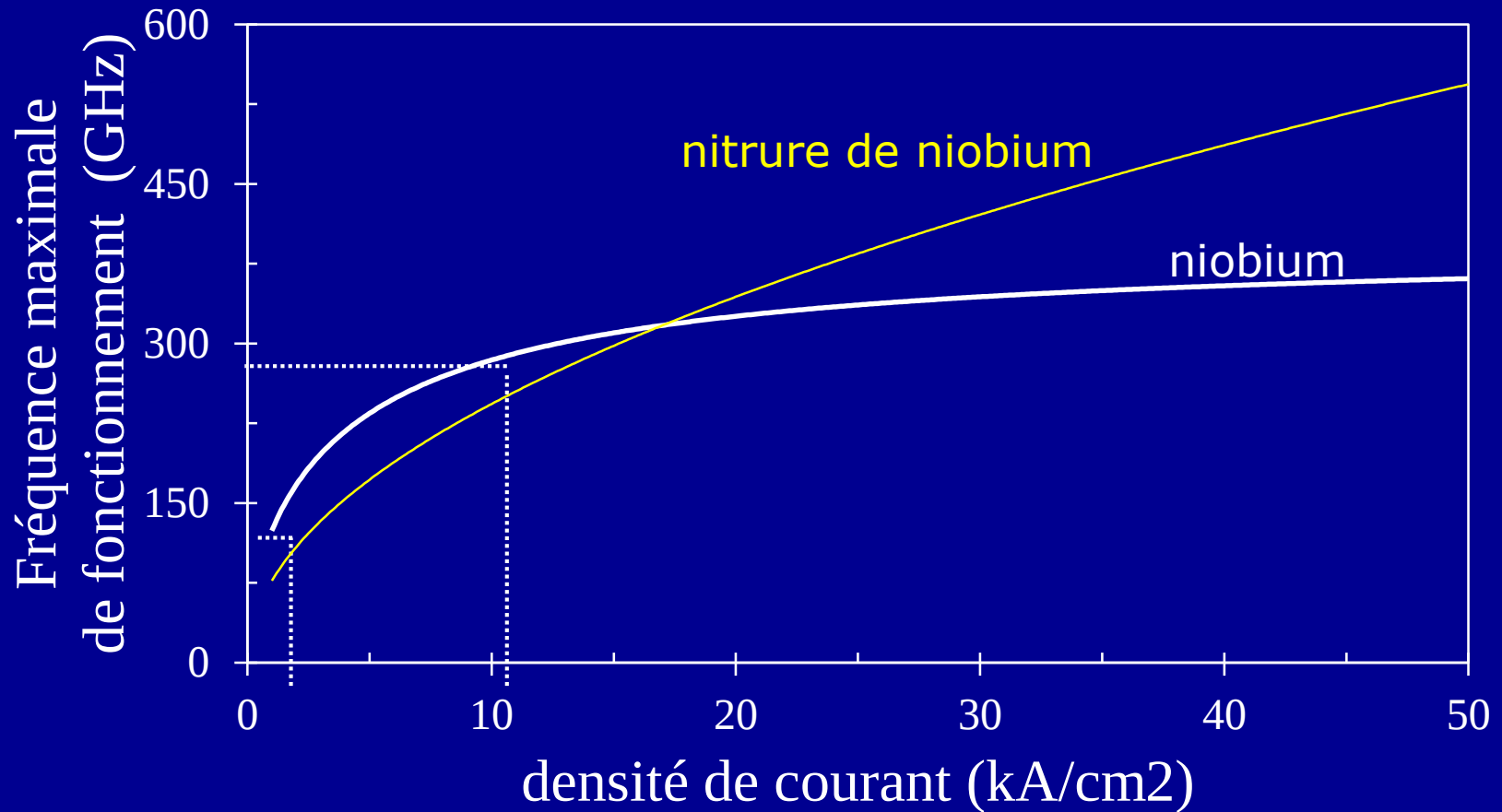
Logique à quanta de flux magnétique (logique RSFQ)



Cellules en électronique supraconductrice



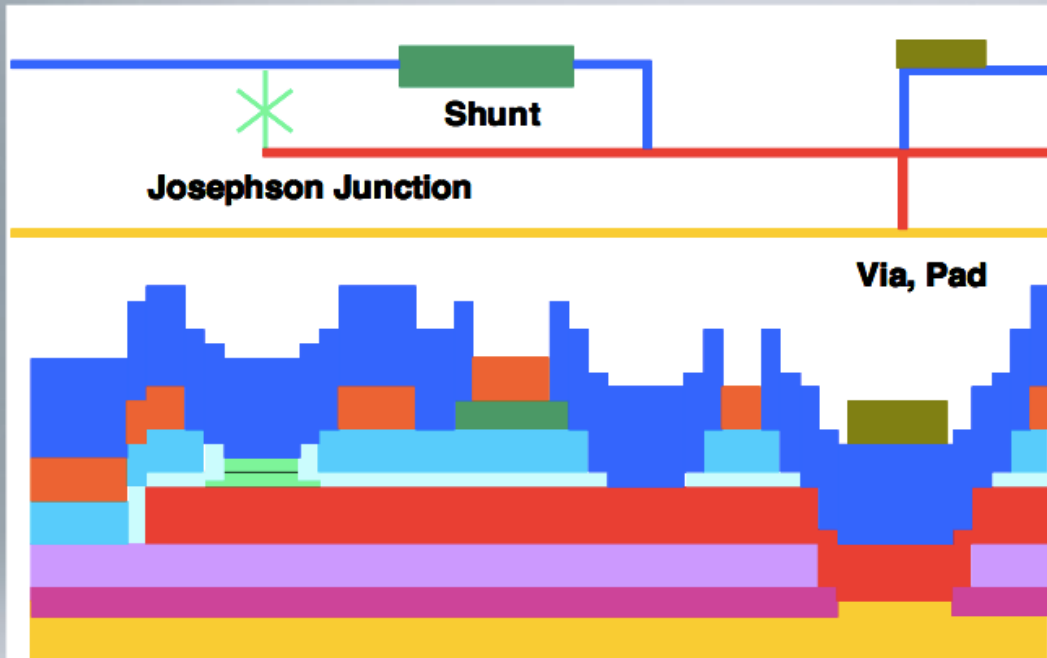
Fréquence maximale de fonctionnement en électronique supraconductrice



Valide pour des jonctions SIS shuntées de manière externe

RSFQ process – cross section

iphtjena



Layer	Thickness	Material
R2	50 nm	Au
M2	350 nm	Nb
I2	150 nm	SiO
R1	80 nm	Mo
I1B	150 nm	SiO
I1A	70 nm	Nb ₂ O ₅
T1	60 nm 12 nm 30 nm	Nb Al ₂ O ₃ Nb
M1	250 nm	Nb
I0B	200 nm	SiO
I0A	50 nm	Nb ₂ O ₅
M0	200 nm	Nb



Procédés des fonderies d'électronique supraconductrice

PROCEDE	densité de courant [kA/cm ²]	aire minimale [μm ²]	Intégration maximale	Fréquence maximale [GHz]
Hypres #03-10-45	0.03 1.0 4.5	~ 3.14	15 000	80 GHz RnIc=1,3mV @ 4,5 kA/cm ²
Hypres #S45/100/200	0.1 1 4.5 10 20 30	~ 0.4	10 000	200 GHz @ 30 kA/cm ²
MIT Lincoln Lab SFQx	10 20 50	~ 0.06	~ 800 000	240 GHz ^[SEP] RnIc=2,17 mV @50 kA/cm ²
ADP2	10	1.0	1100 JJ/mm ²	80 GHz
STP2	2.5 - 20	0.25 - 4.0	100 JJ/mm ² - > 2,000 JJ/mm ²	30 GHz - 150 GHz
HSTP	10	1.0	70 000	80 GHz
Fluxonics standard	1	12.5	100 JJ/mm ²	40 GHz RnIc =0,256 mV
INRiM SNIS	up to 100	25	1,000 JJ/mm ²	300 GHz RnIc =0,1mV - 0,7mV
NIST Nb/Nbx Si1-x/Nb	up to 110	?	70 000	300 GHz
INRiM SNIS 3D FIB	up to 100	0.1	10 000 JJ/mm ²	300 GHz ^[SEP] RnIc=0,1mV - 0,7mV

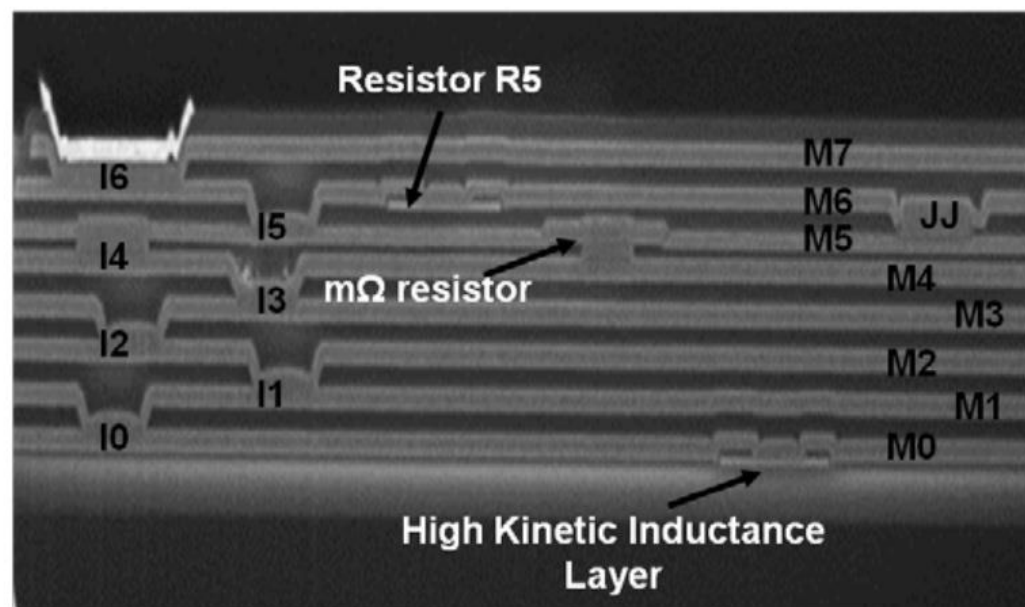
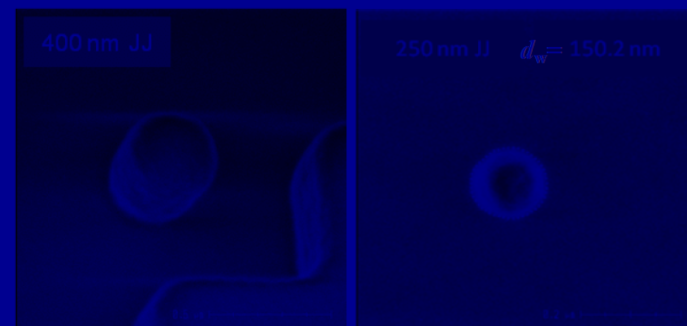
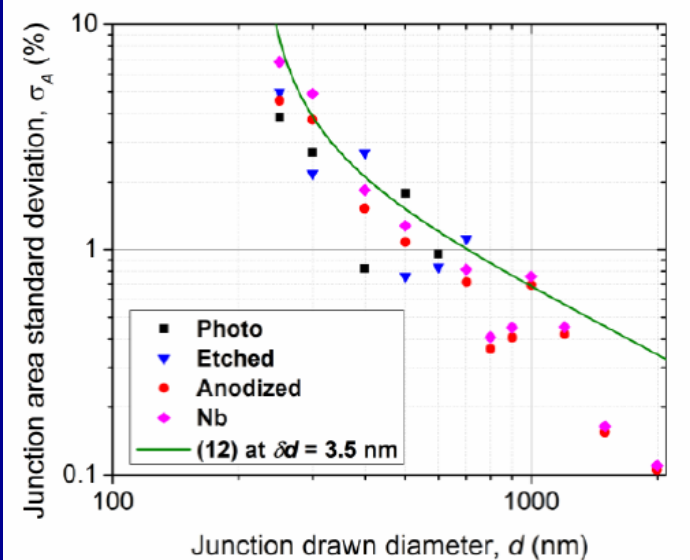


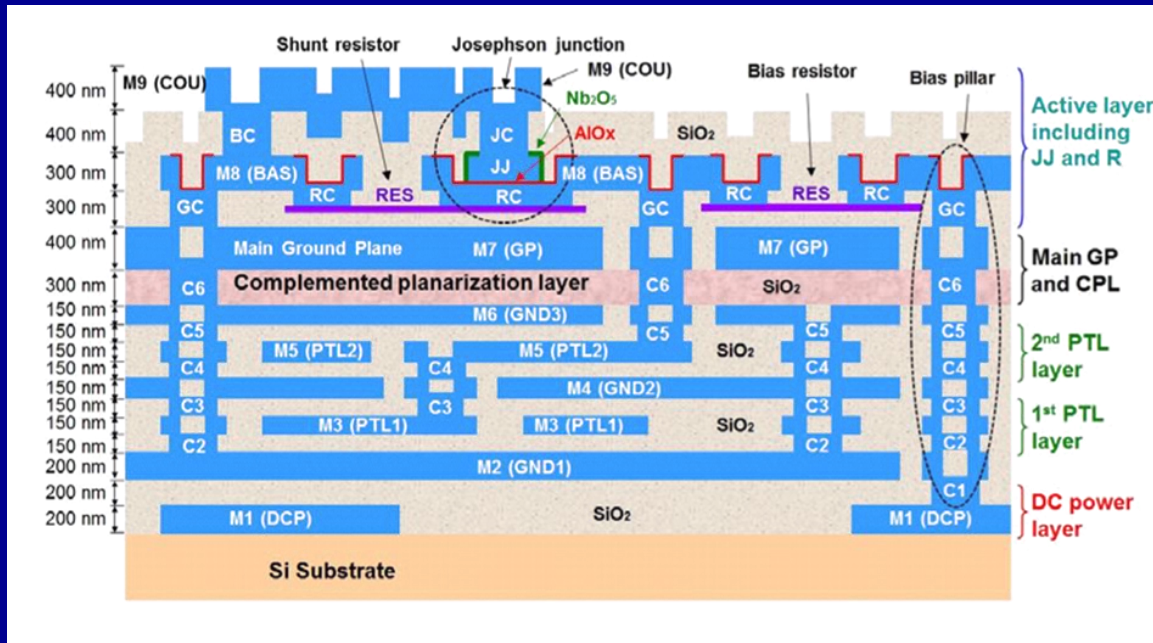
Fig. 10. Cross section of a wafer fabricated by the SFQ5ee process. The labels of metal layers and vias are the same as in Table I. New features of the SFQ5ee are shown: a high kinetic inductance layer under M0 and a layer of mΩ-range resistors between M4 and M5 layers.



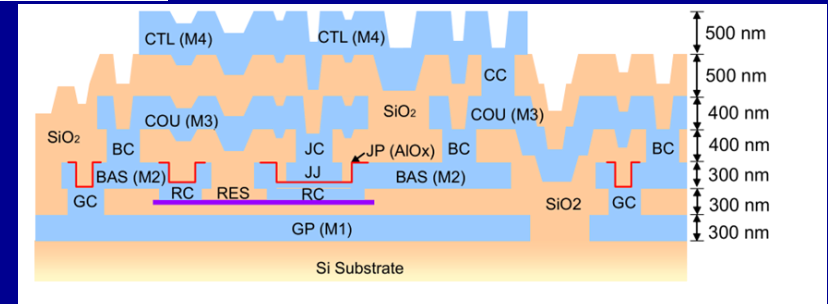
nombre de JJ $\approx 800,000$ JJ/puce

S. K. Tolpygo et al, "Fabrication Process and Properties of Fully- Planarized Deep-Submicron Nb/Al-AlOx/Nb Josephson Junctions for VLSI Circuits," *IEEE TAS 2015*

S.K. Tolpygo et al., "Developments towards a 250-nm, fully planarized fabrication process with ten superconducting layers and self-shunted Josephson junctions," arxiv_1704.07683 (20017); *IEEE TAS*.



Procédé ADP 2 (9 niveaux de métal - 10kA/cm²)



Procédé STP2 (4 niveaux de métal - 2,5 or 20kA/cm²)

S. Nagasawa et al., "Nb 9-Layer Fabrication Process for Superconducting Large-Scale SFQ Circuits and Its Process Evaluation," IEICE 2014

S. Nagasawa, T. Satoh, and M. Hidaka, "Uniformity and Reproducibility of Submicron 20kA/cm² Nb/AIOx/Nb Josephson Junction Process," ISEC 2015

Microprocesseurs supraconducteurs

Ils existent !

CORE1 β – 2006

10955 jonctions Josephson

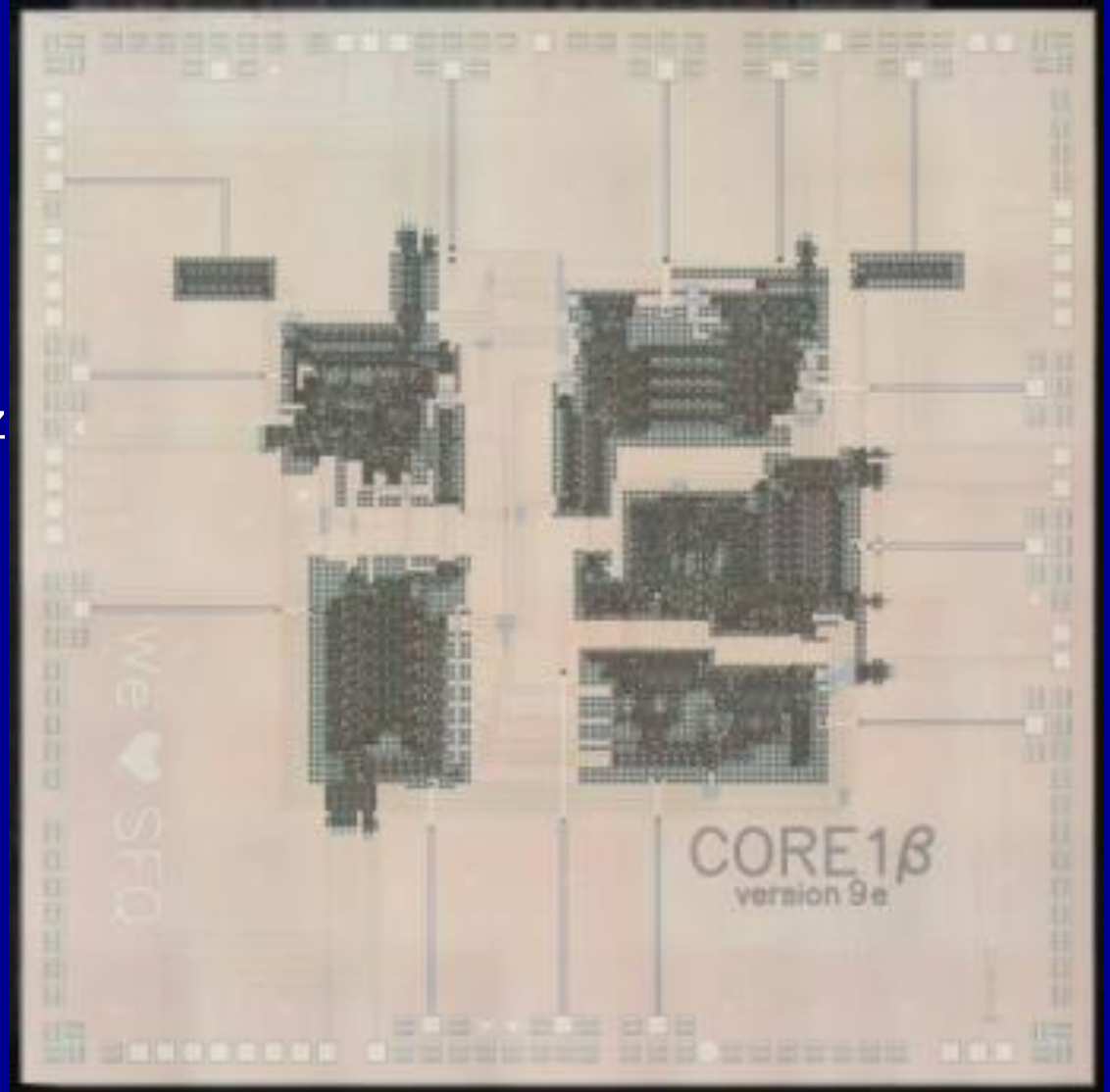
Fréquence d'horloge : 25 GHz

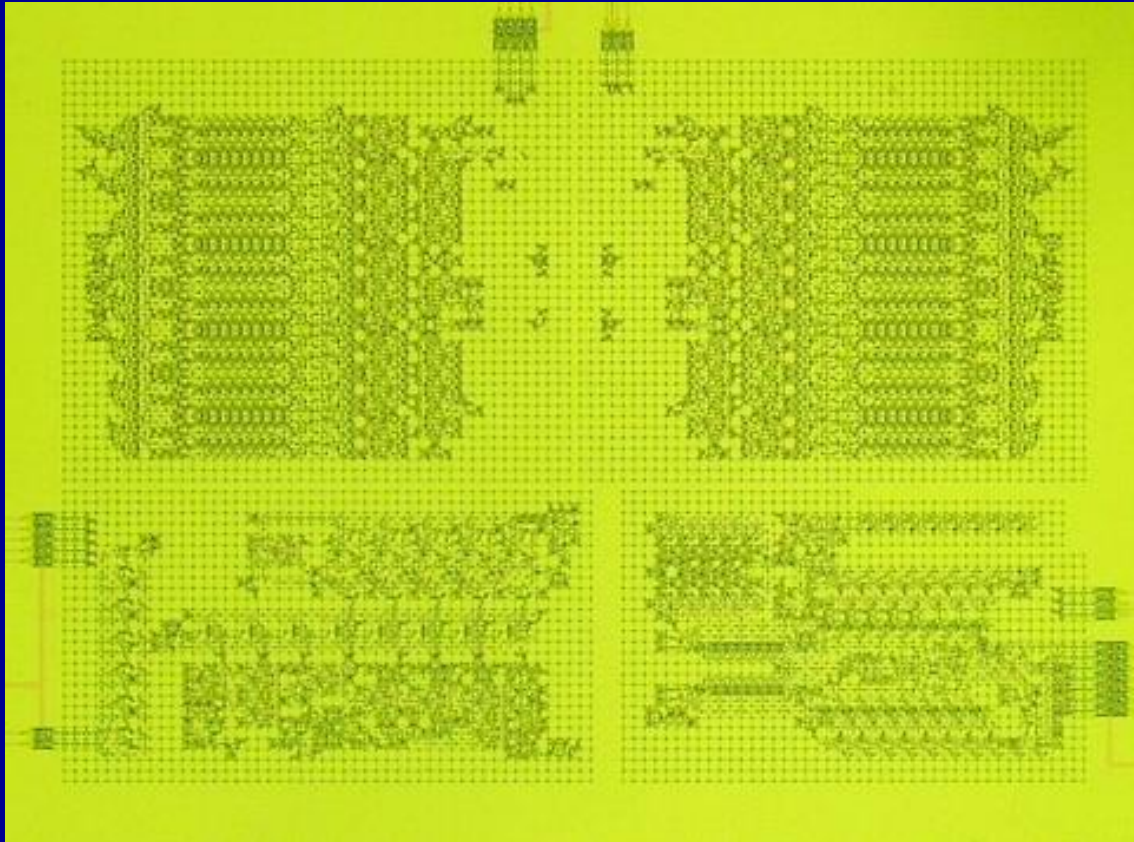
1,4 milliards d'opérations/sec

3,3 mW



Source: Akira FUJIMAKI





COREe2 (2017)

10655 JJs

500 MIPS

2,4 mW

210 GIPS/W

Exécution de programmes

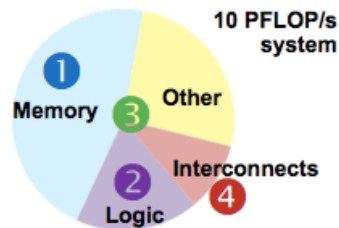
50 GHz

Mémoire intégrée

Source : Prof. Akira FUJIMAKI - Université de Nagoya

Supercalculateurs supraconducteurs

Performance (PFLOP/s):	1	10	100	1,000
Power budget (@ 4 K)	1.5 W	10 W	100 W	1,000 W
Logic (RQL, $I_c = 25 \mu A$, 8.3 GHz) • processor cores	0.18 W • 40,200	1.8 W • 402,000	18 W • 4,020,000	180 W • 40,200,000
Memory (1 B/FLOP, JMRRAM) • quantity (1 B/FLOPS)	0.46 W 1 PB	4.6 W 10 PB	46 W 100 PB	460 W 1,000 PB
Interconnects (VCSELs @ 40 K)	0.1 W	1 W	10 W	100 W
Other (structure, radiation heat leaks)	0.76 W	2.6 W	26 W	260 W
Total	1.5 W	10 W	100 W	1,000 W
• Computation efficiency (goal: $\geq 5 \times 10^{11}$ FLOPS/W)	0.7×10^{11} FLOPS/W	2.5×10^{11} FLOPS/W	5×10^{11} FLOPS/W	5×10^{11} FLOPS/W



Conclusions:

- Energy-efficient superconducting computers are possible
- Priorities:

Memory → **Logic** → **System** → **Interconnects**

Un système à base d'électronique supraconductrice à 4K



Récepteur numérique :
hauteur : 1,8 m
largeur : 50 cm

Processeurs à base de logique binaire

L'information est transportée par des quanta de flux magnétiques (des vortex) et non par des tensions comme avec les technologies à base de semi-conducteurs

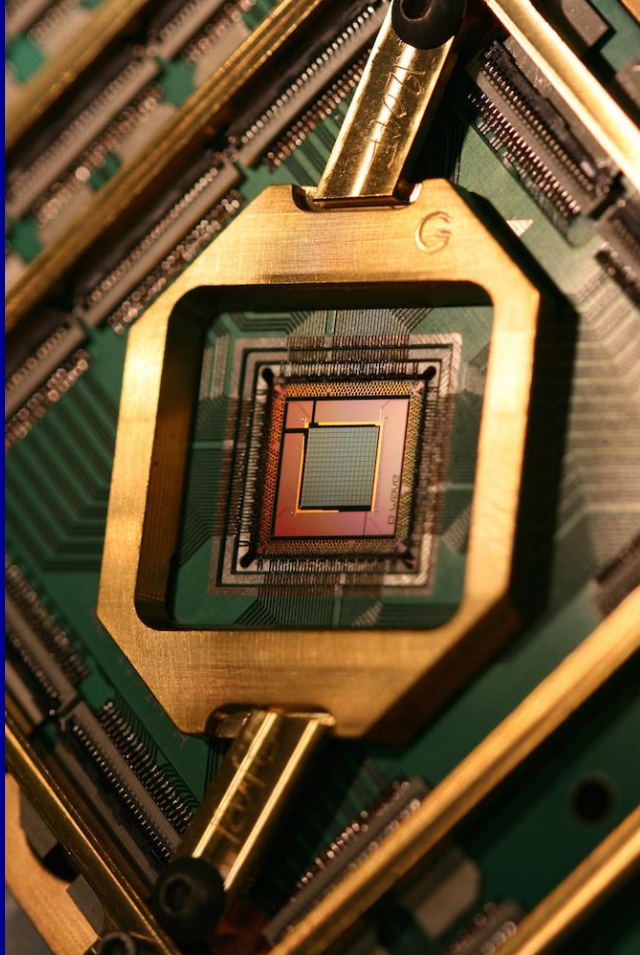
Source : Deep Gupta – HYPRES

Réfrigérateurs 4K disponibles commercialement



Source: Alain Ravex - Absolut Systems

L'ordinateur quantique



Source : D-Wave



D-Wave's latest processor has 2,000 qubits — far surpassing the capacity of previous models.

Source : E. Gibney, "Quantum computer gets design upgrade," Nature, January 2017

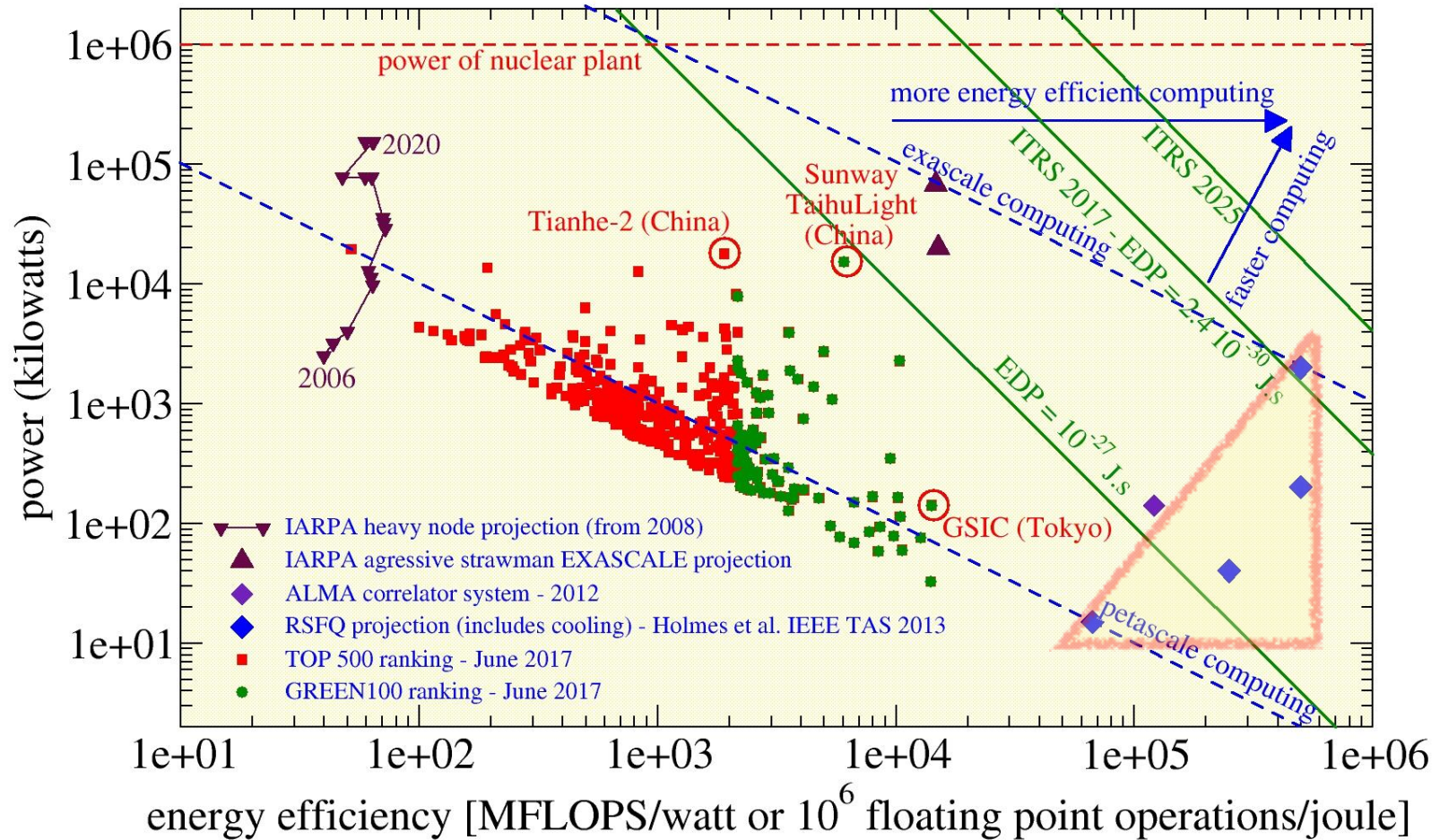
The IBM vision : 50 Qu-bits



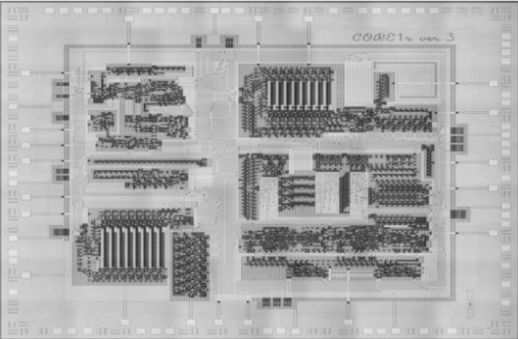
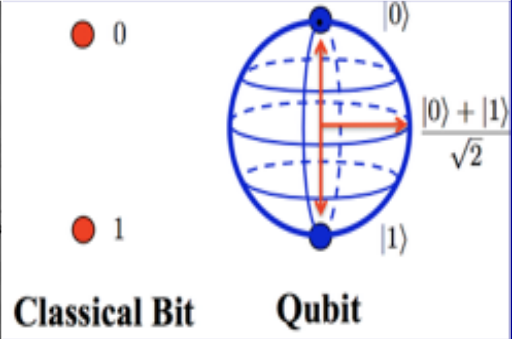
Comparaison des technologies

Performance and power of high-end computing

September 2017 - total power consumption : 0.65 GW



De la technologie CMOS aux ordinateurs quantiques

	<i>CMOS</i>	<i>calcul classique avec supraconducteurs</i>	<i>calcul quantique</i>
<i>technologie</i>			
<i>efficacité énergétique</i>	10^4 MFLOPS/W	5.10^5 MFLOPS/W (inclut le refroidissement)	??
<i>vitesse</i>	2,4 GHz	100 GHz	dépend du problème
<i>portes par cm²</i>	$2,64 \cdot 10^9$	$0,8 \cdot 10^6$	$2 \cdot 10^3$
<i>température de travail</i>	300 K	4 K	0,01 K