



L'Exascale, apports et défis

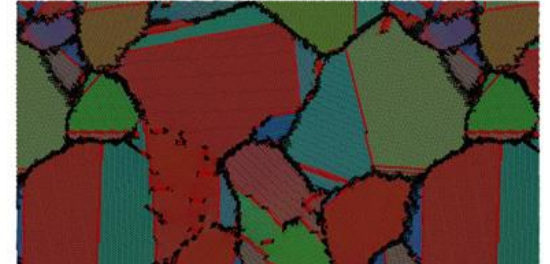
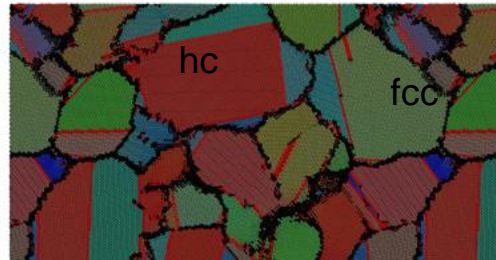
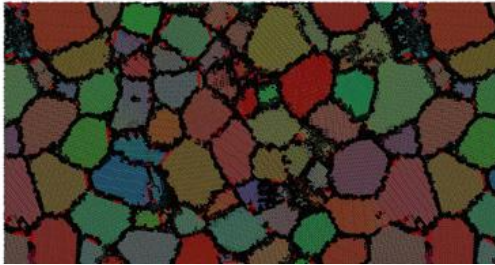
30 septembre 2014

Apports prévisibles

- Apports attendus de l'exascale décisifs dans divers domaines de la science, de l'industrie (retombées économiques), de la défense
- Recensés (notamment) par le DOE, le DARPA (USA), l' EESI, le CRESTA (Europe), le CSCI (France)
- Rapports à destination des décideurs pour motiver les investissements (importants...)
- Quelques exemples (parmi beaucoup de possibles)

Matériaux

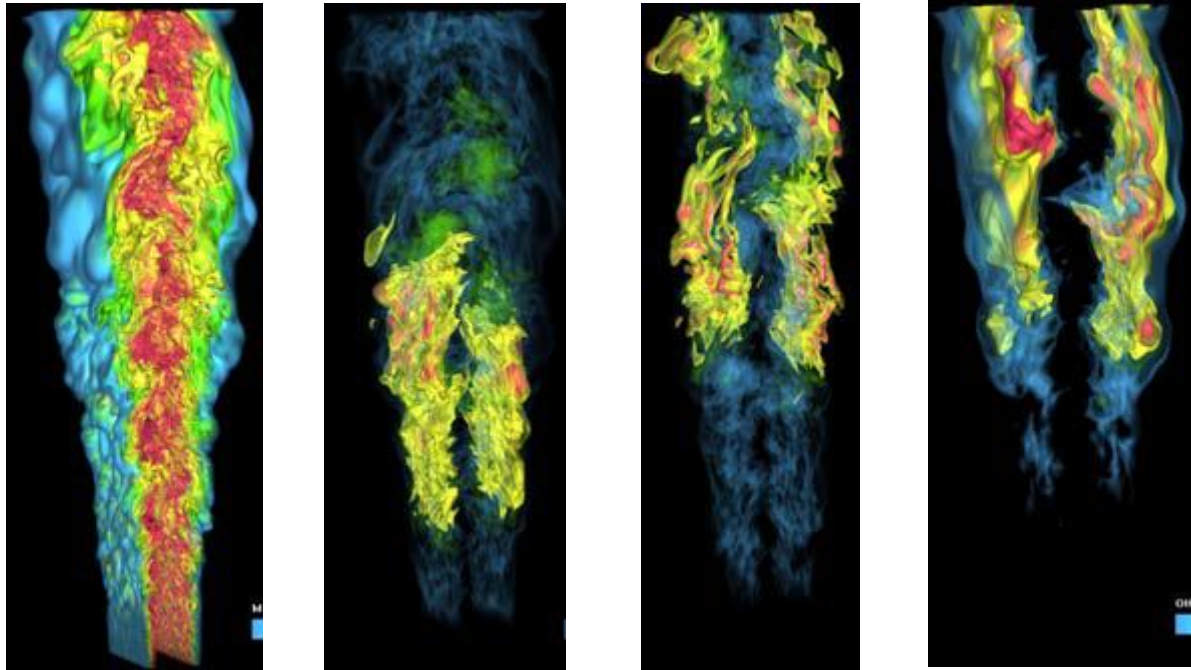
- Simuler au niveau atomique à des échelles de temps (10^{-6} s) et d'espace (10^{-6} m) industrielles
- Croissance de grains à $0.75 T_F$ (DOE 2009), simulations sur quelques ns



- Voir exposé G Zerah

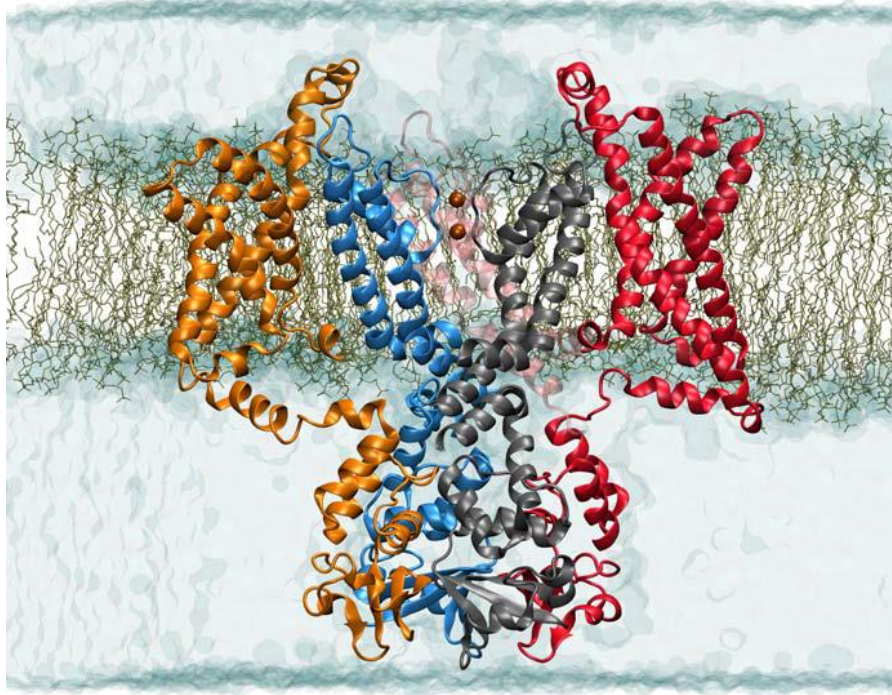
Combustion

- Simulation fine de la combustion pour améliorer le rendement et diminuer les émissions des moteurs
- Flamme d'éthylène (DOE 2010), concentrations espèces



Biologie

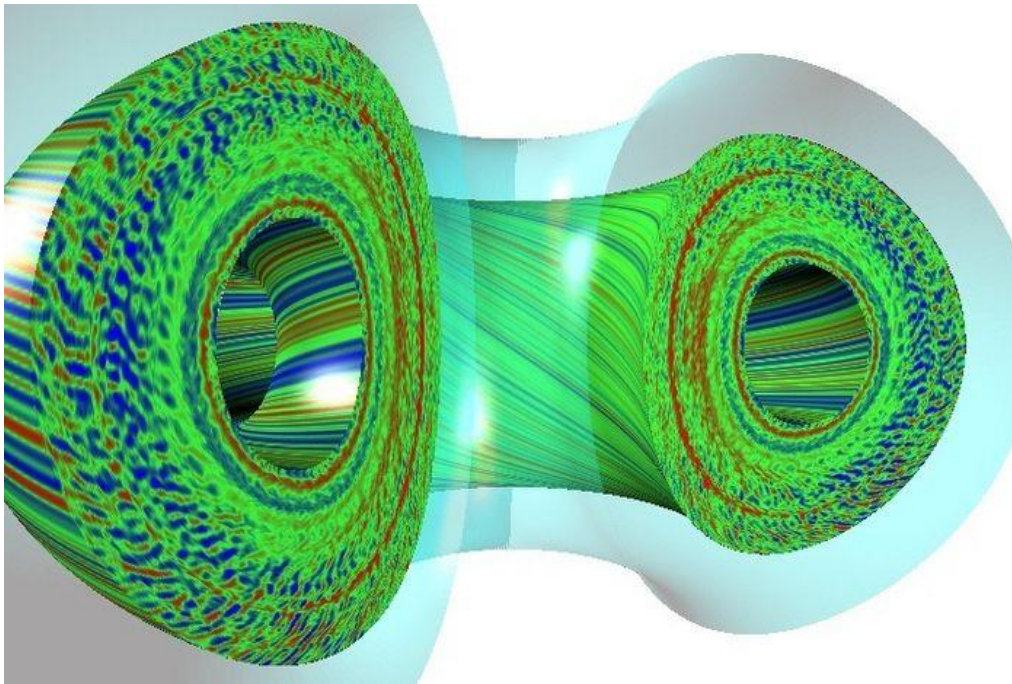
- Simulation de molécules biologiques en interaction
- Voir exposé M Masella



Simulation
de canaux
ioniques
(DOE 2010)

Fusion

- Simulations gyrocinétiques dans les tokamaks

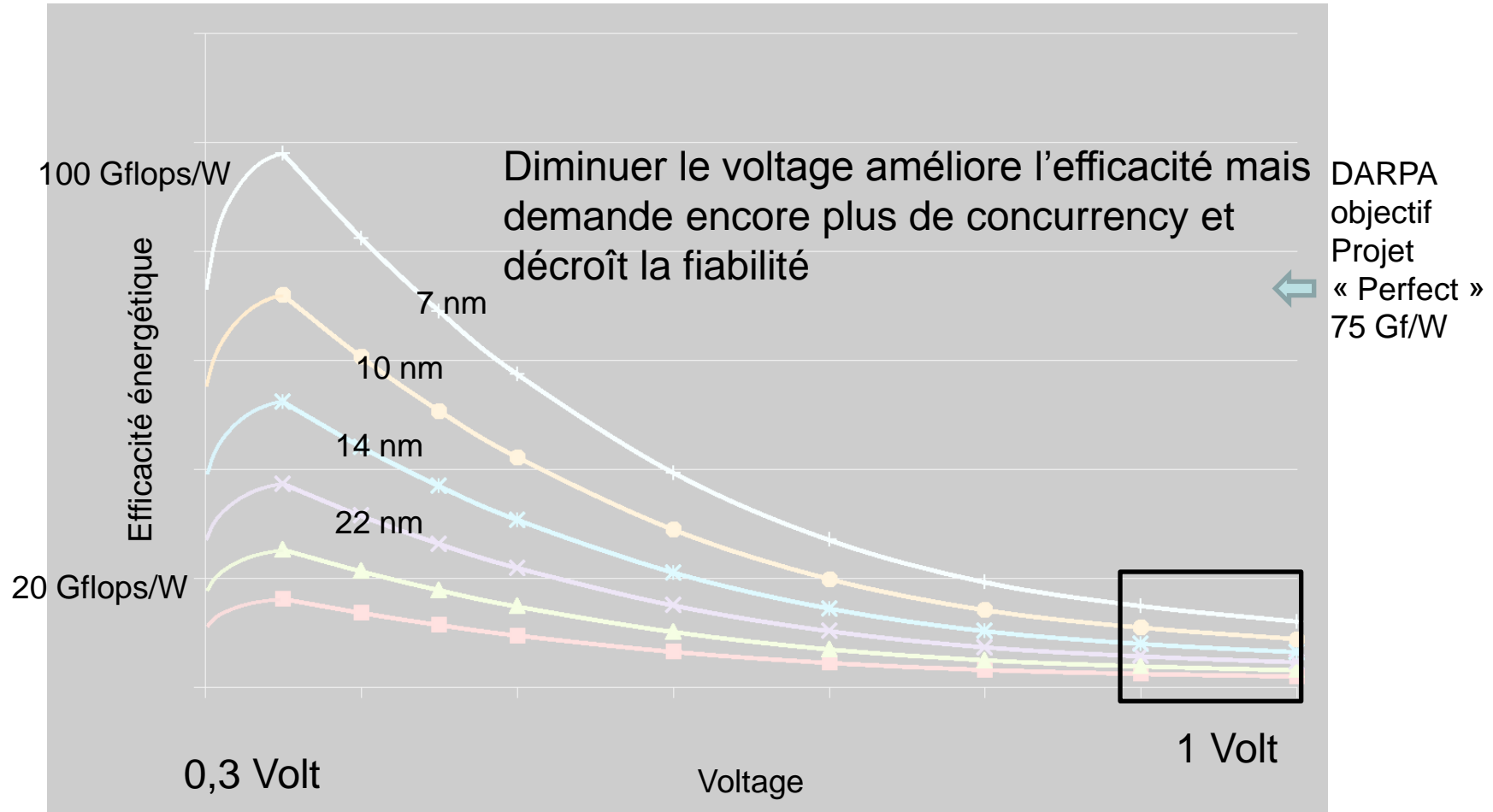


Calculs 5D
très coûteux
Déjà des
simulations
à qq 10^{11} ddl
(V Grandgirard,
CEA)

Caractéristiques probables d'une machine de classe exaflopique

- Une telle puissance ne pourra être obtenue qu'avec un parallélisme à plusieurs niveaux
 - Cœurs : SIMD (le retour de la vectorisation...)
 - Nœuds à grand nombre de cœurs
 - Parallélisme inter- nœud
- La forme précise de la machine est encore très floue
- La date à laquelle la première machine sera construite dépendra des avancées (efficacité énergétique notamment) sur les processeurs (exposé G Colin de Verdière) et du montant de l'investissement (initial + consommation électrique) consenti

Efficacité énergétique vs gravure et voltage



Un design (très...) hypothétique

- Cœurs à registres vectoriels : 64 opérations/cycle d'horloge, fréquence 1,25 GHz, 100 GFlops
- Nœuds de quelque 10^2 cœurs, 10 TFlops/ nœud,
- 10^5 nœuds

- Possible (à budget tendant vers l'infini) en 2015 :
 - Nœuds Xeon Phi 72 cœurs 3 Tflops crête,
 - Environ $3 \cdot 10^5$ nœuds

Mais puissance d'environ 100 MWatts... Une puissance « raisonnable » est plutôt de 20 MW : solutions alternatives (exposé B de Dinechin) et/ou progrès de l'électronique et des processeurs
la puissance de calcul n'est pas tout : problématique des flux de données

Flux de données

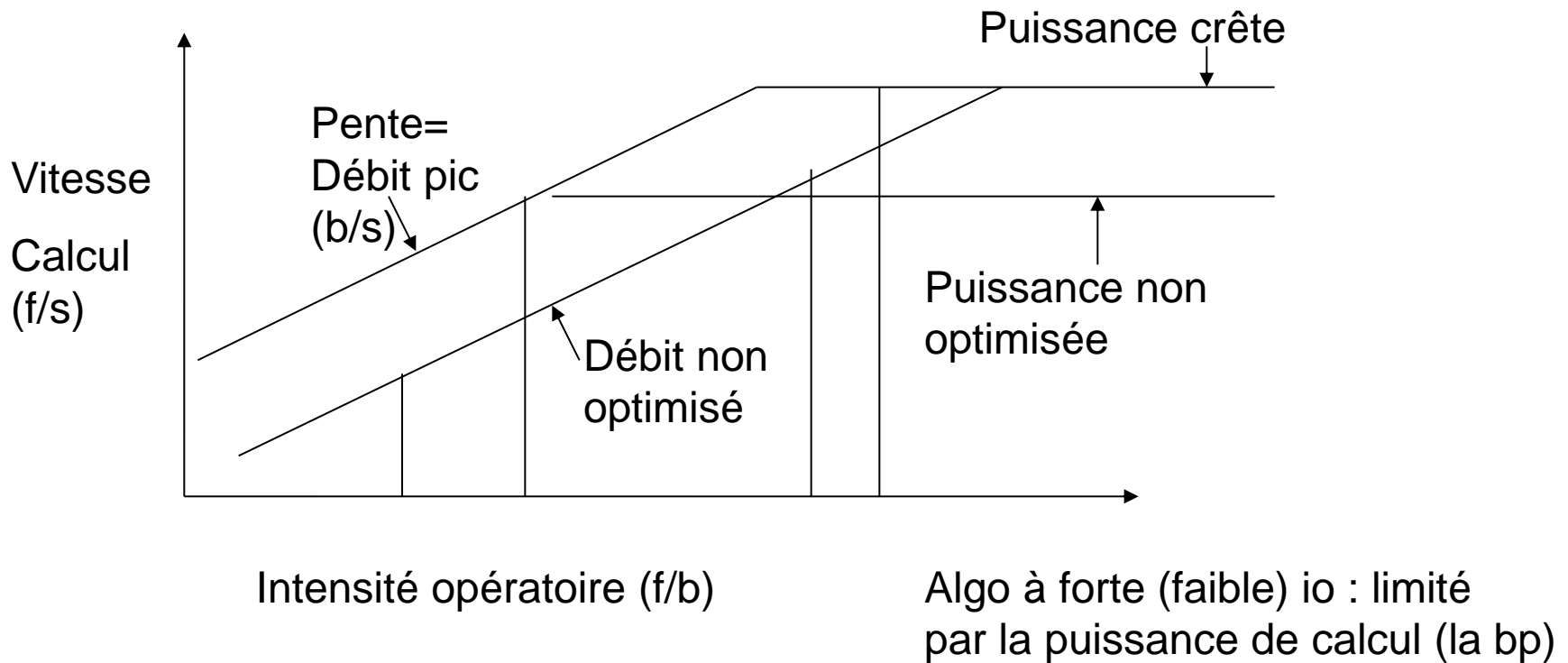
- Progrès attendus calculs : qq 10^{-11} J/Flops, qq10 MW/ExFlops à l'horizon 2020, mais
- Progrès attendus sur la latence/bande passante moindres que sur la puissance de calcul
 - Hiérarchie de mémoires de capacités et temps d'accès variables
 - Mouvements de données coûteux en temps et en énergie
- Limiter les mouvements de données, quitte à faire plus de calculs

Flux de données et performances

- Ratio byte/flop machine = bande passante mémoire (byte/s)/puissance de calcul (flop/s), voué à décroître...de $\sim 1\text{b/f}$ (actuel) à $\sim 0.1\text{b/f}$
- Ratio byte/flop algorithme = quantité de données transférées de la mémoire vers processeur/ nombre d'opérations réalisées sur ces données
- Exemple différences finies $C(i,j)=a*a(i,j)+b*[a(i,j-1)+a(i,j+1)+ a(i-1,j)+ a(i+1,j)]$ $8 N^2$ byte/ $6 N^2$ flop= $4/3$ b/f au dessus du ratio des futures machines
- Le modèle Roofline permet de visualiser les facteurs limitant les performances

Facteurs affectant les performances

- Modèle « roofline » (log-log)



Facteurs affectant les performances

- Variation de la puissance de calcul efficace suivant
 - L'efficacité d'utilisation du CPU (utilisation +/- efficace des accélérateurs vectoriels)
 - L'intensité opératoire de l'algorithme : une faible i.o. ($b/f \text{ algo} \ll b/f \text{ machine}$) conduit à des performances limitées par la bande passante
 - La bande passante effective : accès séquentiels ou aléatoires, « prefetching », cache, mémoire, ou réseau peut être bien moindre que la bp théorique
- Les pénalités en performances d'algos ou de programmations inadaptés seront importants

Une machine délicate à programmer...

- Parallélisme hétérogène (3 niveaux) et (très) massif
- Temps d'accès aux données et bande passante dépendant fortement de leur position (NUMA)
- Performances très variables suivant les algorithmes et la manière de programmer
- Faciliter le travail des informaticiens (organisation, génie logiciel, formation) sera (encore plus que maintenant) essentiel
- Voir exposé W Jalby

Bilan d'énergie et pannes

- La consommation d'énergie sera un des facteurs limitants à la construction et à l'utilisation des machines exaflopiques
- Calcul au niveau de précision juste suffisant , adaptation des voltages et des fréquences sont quelques une des pistes du calcul « energy aware »
- L'augmentation du nombre de processeurs va réduire le temps moyen entre pannes (1 jour, voire moins) : la défaillance d'un processeur au cours d'un calcul va devenir la règle
- Problème de résilience à traiter au niveau du matériel et des algorithmes

Conclusion

- L'exascale : une belle opportunité pour simuler plus finement et avec une physique plus précise dans de nombreux domaines de la science et de l'industrie, mais trois murs à franchir
- Le mur de l'énergie : gagner un facteur 10 sur la consommation des processeurs
- Le mur de la fiabilité/résilience : gérer les pannes, erreurs, inévitables vu le nombre et la nature des processeurs
- Le mur de la programmation : vectorisation, parallélisme « hybride », choix des algorithmes adaptés
- Un dialogue entre numériciens, informaticiens, constructeurs, et une R et D conséquente s'impose...