



Retour d'expériences de convolutions 3D sur GPU

*pour un code de calcul de structure électronique à base
d'ondelettes*

Matthieu Ospici (stagiaire), ***Luigi Genovese*** (CEA)
T.Deutsch, ***Jean-François Méhaut*** (UJF, INRIA)

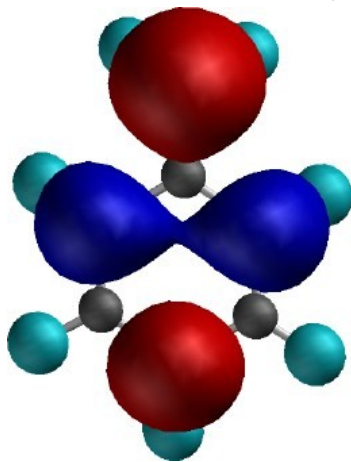
Les méthodes *ab initio* (pas de paramètres ajustables)

La théorie de la fonctionnelle de la densité (DFT) :

Calcule **l'énergie totale d'un solide ou molécule dans son état *fondamental*** (Hohenberg & Kohn 1964, Kohn & Sham 1965).

$$-\frac{1}{2m}\nabla^2\psi(\mathbf{r}) + V_{\text{ions}}(\mathbf{r})\psi(\mathbf{r}) + V_{\text{hxc}}[\psi](\mathbf{r})\psi(\mathbf{r}) = \varepsilon\psi(\mathbf{r})$$

Pour chaque électron, correspond une fonction d'onde ψ à représenter dans tout l'espace

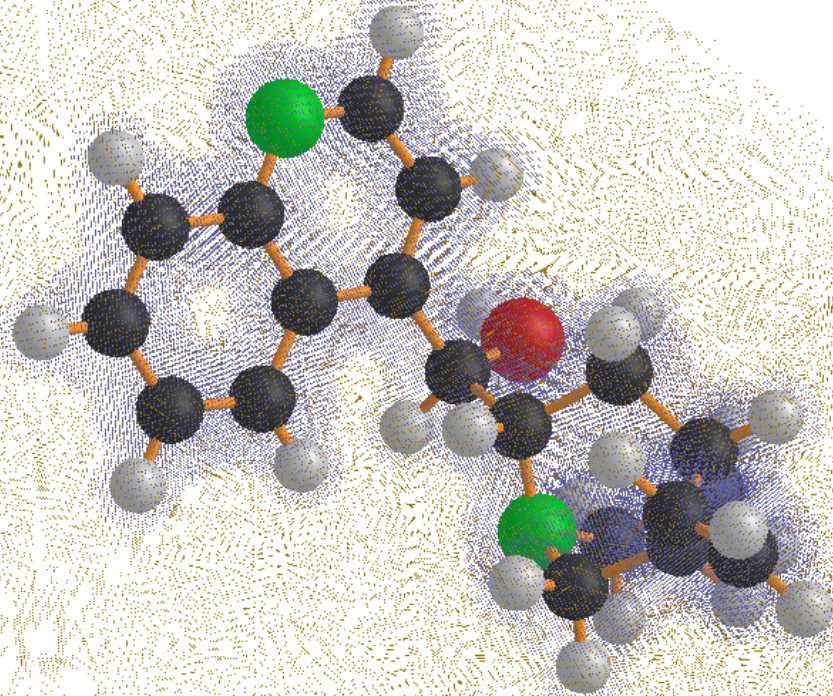
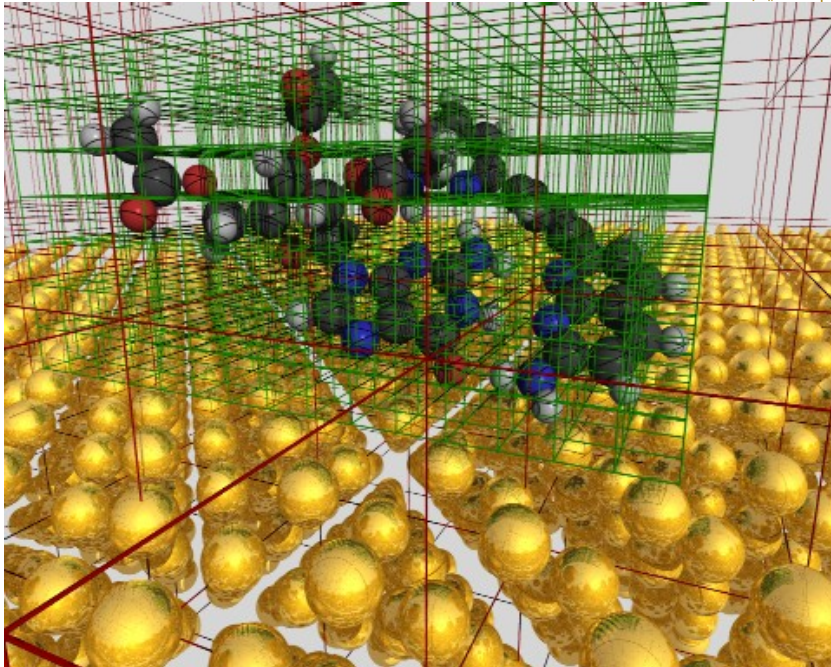


Calculs intensifs :

- 10 à 1 000 atomes selon les calculateurs
- 40 à 10000 fonctions d'onde

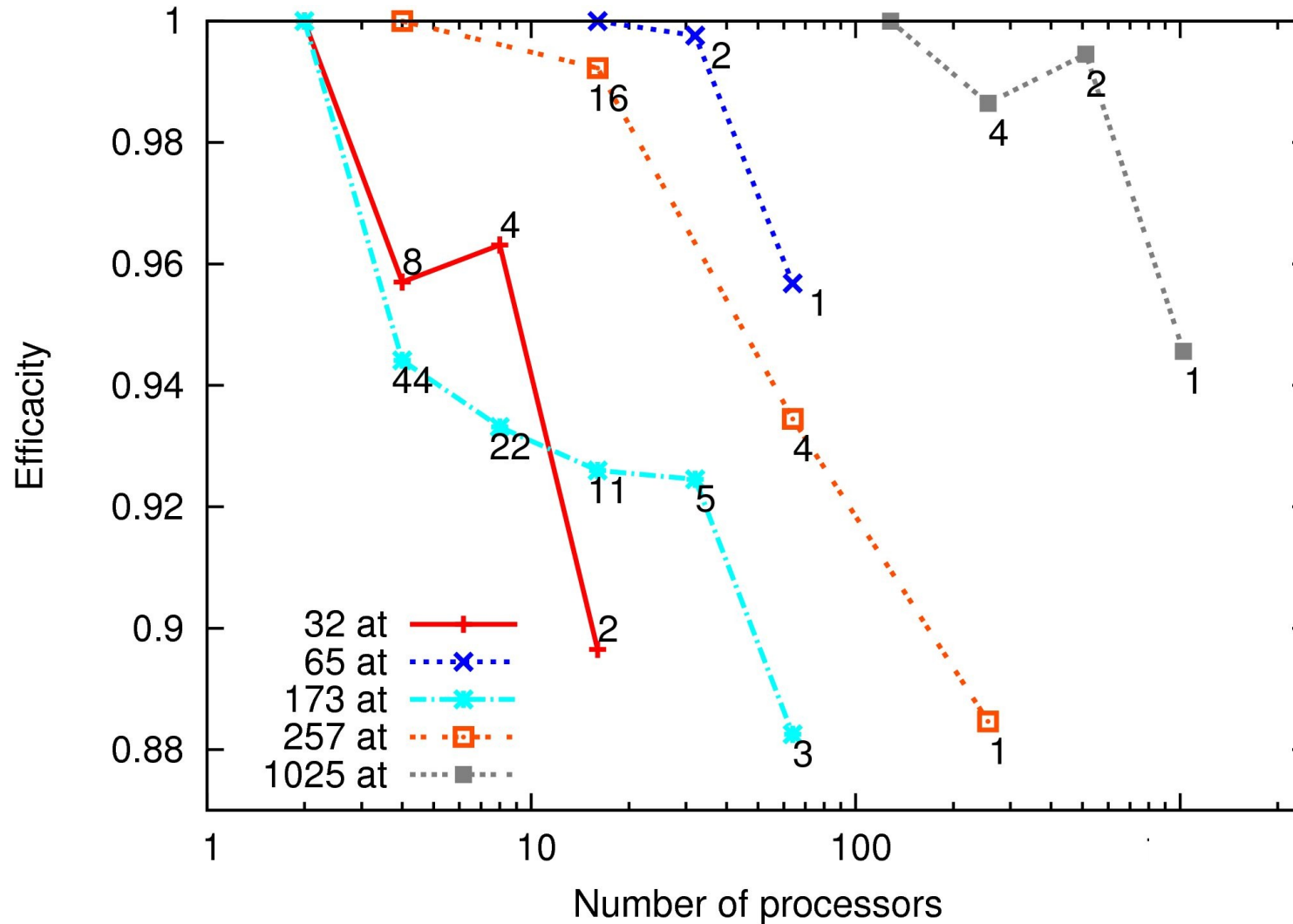
Maillage adaptatif: raffiner où c'est nécessaire

Molécule de 44 atomes et
le maillage adaptatif associé



Ajuster automatiquement la base en fonction des propriétés à étudier (ex: photo-émission)
ou des conditions aux bords (solide, surface, fil, solutions,...)

Adapter au calcul massivement parallèle

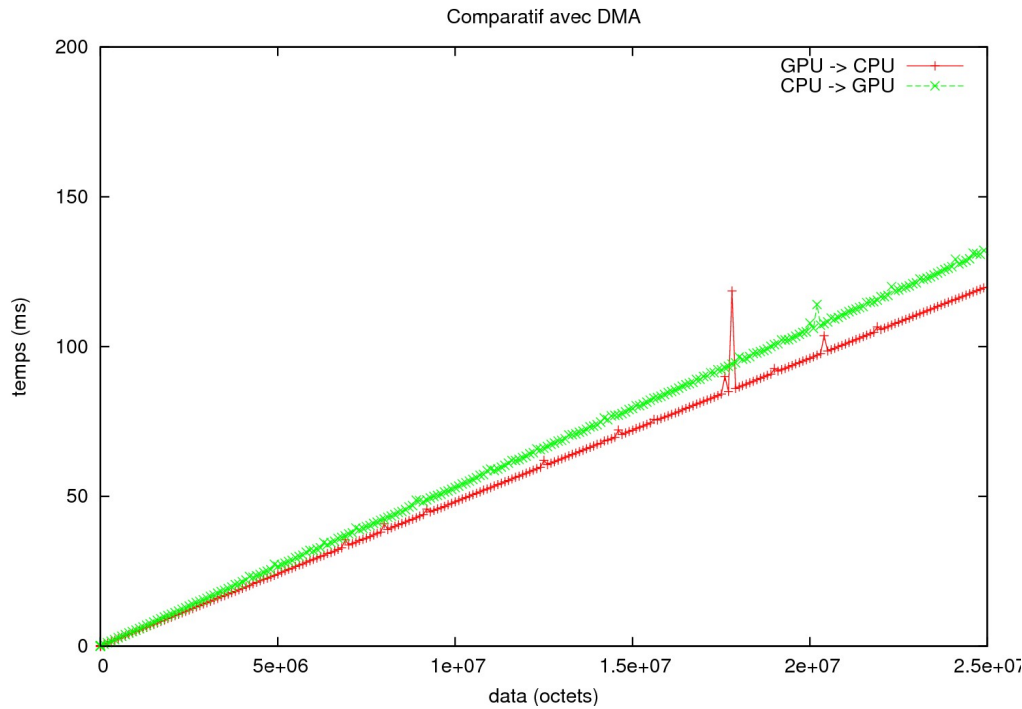


Initiative d'utilisation des cartes graphiques (INRIA) :

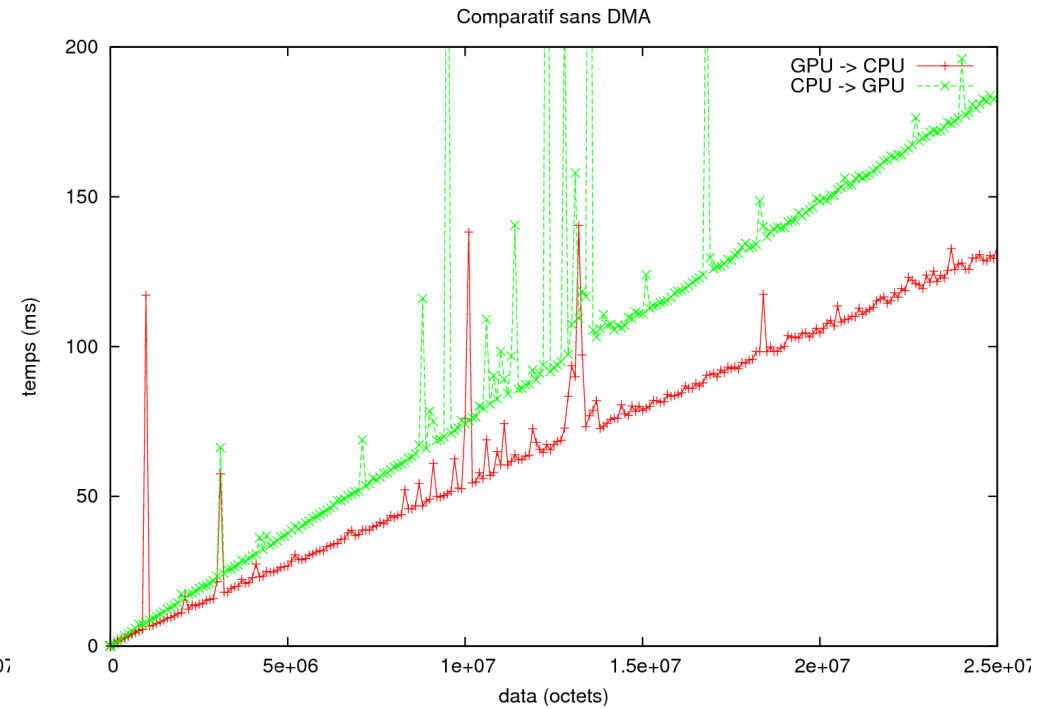
- convolutions 3D sur 8600GTS, 8800 GTX et Tesla
- multi-GPU



Comparatif débit avec et sans DMA (8600 GTS)



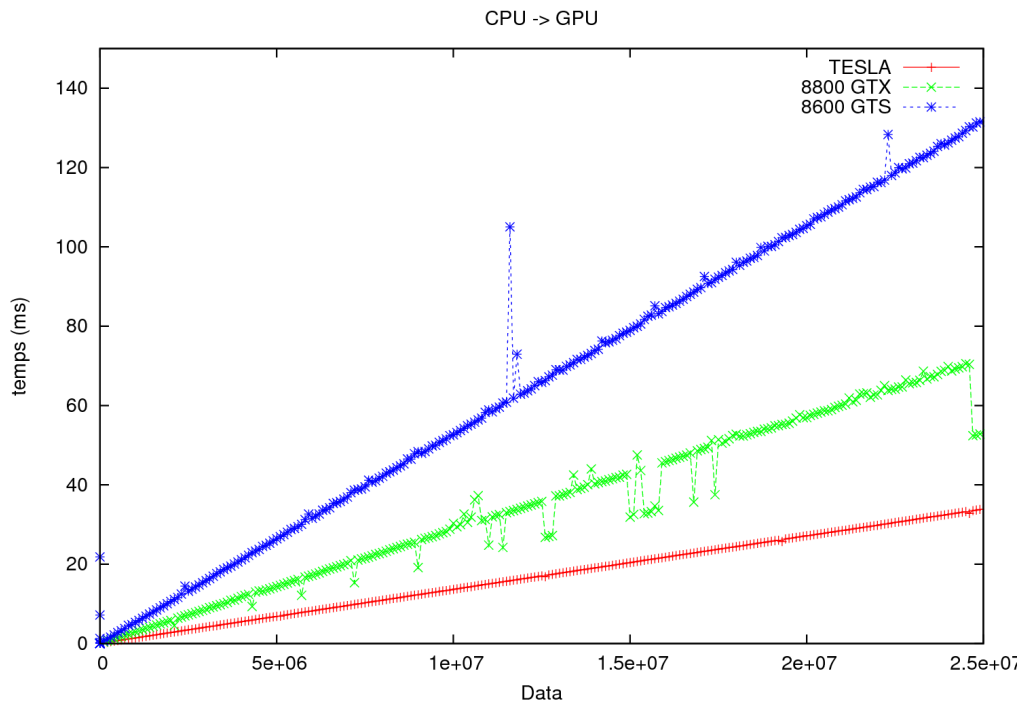
Débit : 532Mo/s (CPU -> GPU)
716Mo/s (GPU -> CPU)



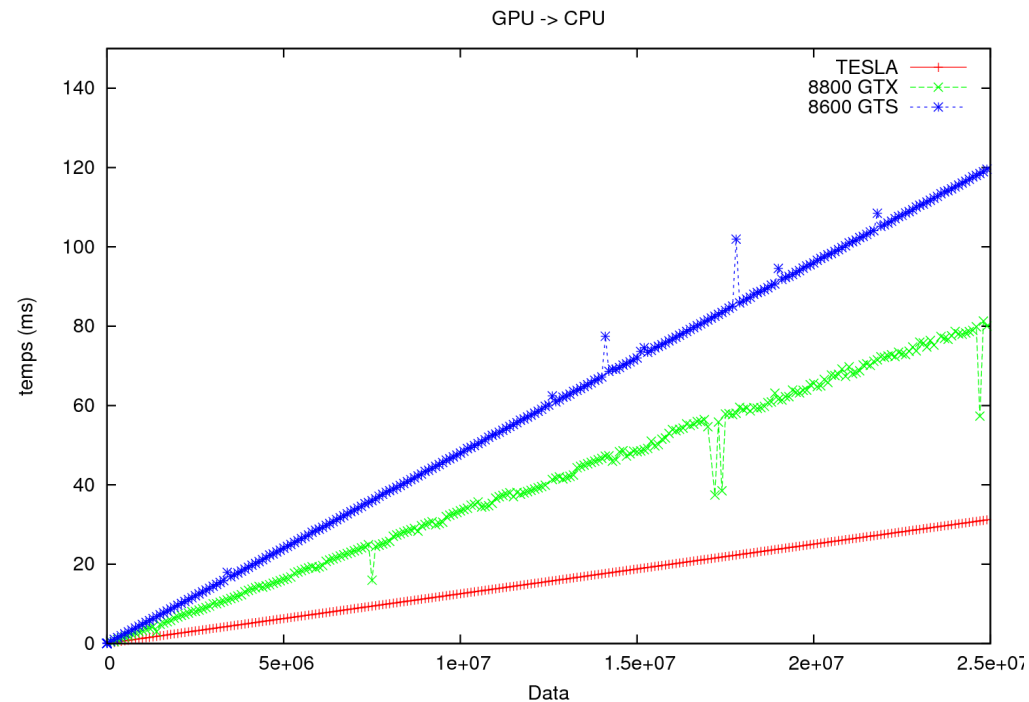
Débit : 756Mo/s (CPU -> GPU)
840Mo/s (GPU -> CPU)

Latence identique : 20,6 μ s

Comparaison débit : 3 cartes graphiques



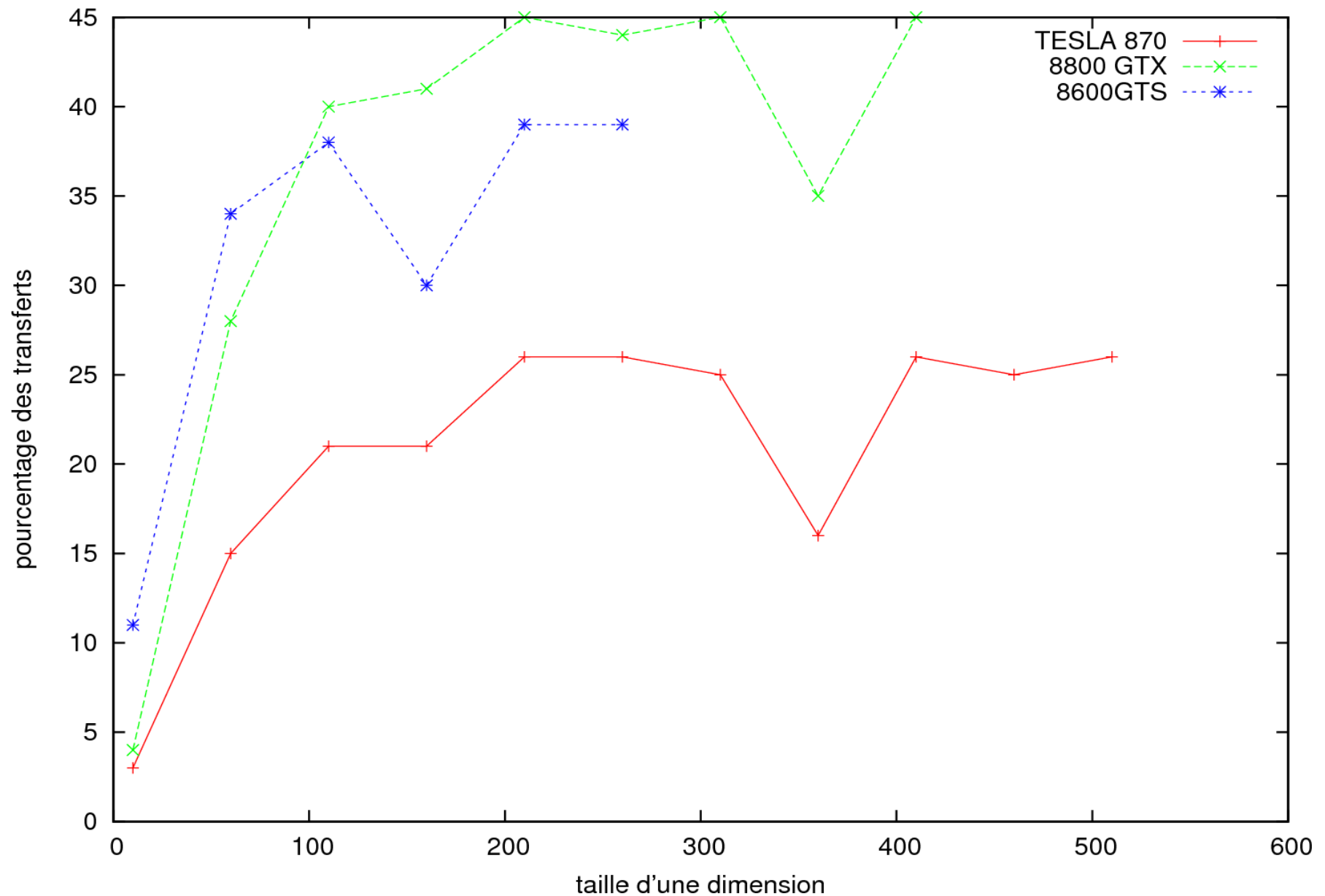
Tesla : 2,97 Go/s
8800 GTX : 1,4 Go/s
8600 GTS : 756 Mo/s



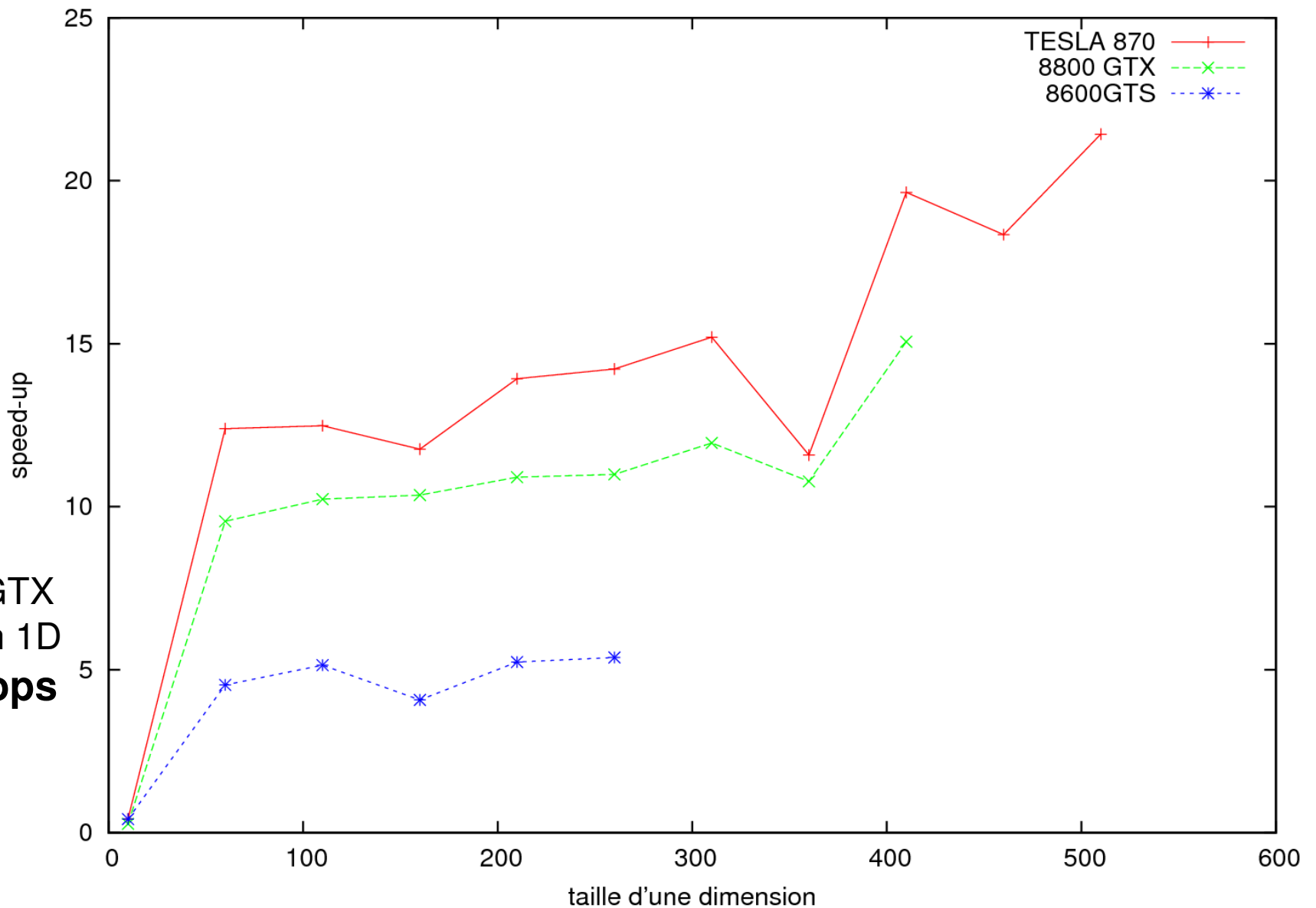
Tesla : 3,12 Go/s
8800 GTX : 1,2 Go/s
8600 GTS : 840 Mo/s

Latence à 20 μ s pour toutes les cartes

Temps de transferts par rapport au temps total (convolutions 3D)



Facteur d'accélération (convolutions 3D)

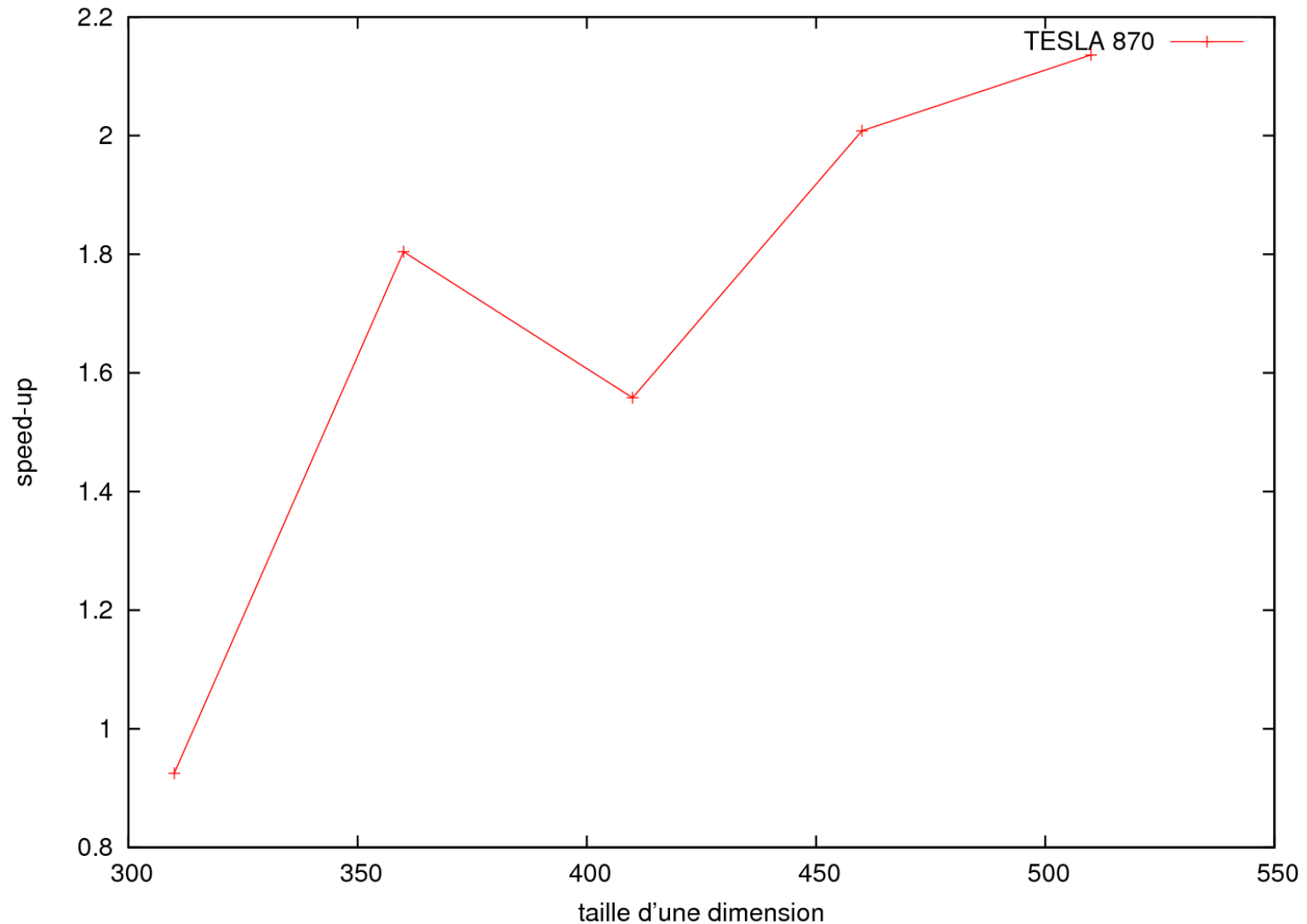


Par rapport à un Intel Xeon E5472, compilation avec icc 10.1

Sur 8800 GTX
convolution 1D
99,8 Gflops



- Création des threads et cudaMalloc avant le transfert mémoire et le calcul
- Speed-up à partir de 310x310x310
- Speed-up de 2,13 sur 4 Tesla C870 et 4 convolutions 510x510x510



- Interface Fortran
- Utilisation des CudaBlas
- Optimisation du multi-GPU et massivement parallèle
- Projet ANR et européen d'utilisation d'architectures hybrides

MULTI-GPU	speed-up	Temps total des 4 convolutions	Temps de calcul (une convolution)	CPU->GPU (1)	GPU -> CPU (1)	Temps total d'une convolution	Taille
	0.925	1190.695	157.731	58.180	69.454	315.286	310x310x310
	1.804	1506.953	465.128	80.337	109.033	665.033	360x360x360
	1.558	1581.567	359.862	110.405	188.224	702.216	410x410x410
	2.008	1833.701	556.673	188.777	262.148	1036.385	460x460x460
	2.136	2239.591	732.397	223.383	345.702	1334.323	510x510x510

	Temps total	Calcul	CPU->GPU	GPU -> CPU	Cuda Malloc	Host Malloc	Calcul CPU	Acc	Acc2	Acc3	Taille
TESLA	1,683	0,228	0,030	0,021	0,020	1,357	0,129	0,077	0,566	0,422	10x10x10
	4,051	1,178	0,339	0,306	0,351	1,505	27,201	6,715	23,091	12,392	60x60x60
	16,388	9,278	1,858	1,701	0,701	2,232	167,992	10,251	18,106	12,485	110x110x110
	50,268	32,995	5,621	5,173	1,564	3,696	529,438	10,532	16,046	11,763	160x160x160
	93,756	57,214	12,649	12,226	3,183	6,145	1175,680	12,540	20,549	13,925	210x210x210
	174,056	107,532	23,970	22,492	5,792	10,144	2249,326	12,923	20,918	14,225	260x260x260
	301,116	190,241	40,593	37,540	9,617	16,190	4185,462	13,900	22,001	15,203	310x310x310
	718,382	547,013	63,534	58,478	14,992	23,786	7873,475	10,960	14,394	11,585	360x360x360
	687,521	420,690	93,867	86,207	22,030	49,493	12100,140	17,600	28,763	19,643	410x410x410
	1015,047	643,756	133,323	121,704	30,873	63,837	16881,639	16,631	26,224	18,343	460x460x460
	1317,369	820,617	180,337	165,845	41,838	79,715	25626,629	19,453	31,228	21,430	510x510x510
8800 GTX	2,464	0,301	0,085	0,026	0,039	1,942		0,052	0,429	0,267	10x10x10
	4,879	1,142	0,678	0,731	0,291	1,740		5,575	23,819	9,551	60x60x60
	20,351	7,589	3,852	4,344	0,726	3,205		8,255	22,136	10,231	110x110x110
	59,961	24,555	11,794	13,310	1,881	6,953		8,830	21,561	10,355	160x160x160
	125,024	48,214	26,490	30,061	3,853	13,382		9,404	24,385	10,907	210x210x210
	239,380	90,933	50,257	56,908	8,159	26,567		9,396	24,736	10,991	260x260x260
	401,792	159,139	85,088	96,631	11,878	39,788		10,417	26,301	11,954	310x310x310
	811,000	446,810	133,523	151,236	21,211	59,330		9,708	17,622	10,779	360x360x360
	918,352	361,441	196,888	223,912	27,120	87,930		13,176	33,477	15,063	410x410x410
8600 GTS	0.44800	0.23200	0.02800	0.02500	0.00800	0.13000		0.28795	0.55603	0.41613	10x10x10
	6.78100	3.25900	1.24700	1.06200	0.33200	0.44700		4.01136	8.34643	4.53199	60x60x60
	34.89900	18.39800	7.10900	6.40800	0.84200	1.37400		4.81366	9.13099	5.14004	110x110x110
	135.97301	87.33500	21.61400	19.66900	2.12100	3.74900		3.89370	6.06215	4.06938	160x160x160
	236.89400	128.51199	48.82800	44.44100	4.19800	8.01500		4.96290	9.14841	5.23266	210x210x210
	452.29099	235.98700	92.85900	84.53000	7.82600	26.02800		4.97318	9.53157	5.37554	260x260x260