

ACKNOWLEDGMENT: This work has been supported in part by the French National Research Agency (ANR) through COSINUS program (project MIDAS no. ANR-09-COSI-009).



Midapack : Développement d'une bibliothèque de calcul haute performance pour l'analyse des données cosmologiques

**Frédéric Dauvergne, Pierre Cargemel, Maude Le Jeune,
Radek Stompor**

Huitièmes journées informatiques de l'IN2P3-IRFU,
La Londe Les Maures,
22-25 Octobre 2012

Plan

- ▶ Evolution récente du calcul intensif ;
- ▶ La cosmologie comme cadre d'application ;
- ▶ Description de la bibliothèque de calcul Midapack ;
- ▶ Un exemple d'application avec la cartographie du fond diffus cosmologique.

La simulation numérique et la science

« La simulation numérique a révolutionné notre approche de la recherche scientifique ».

Théorie $\longrightarrow$ Expérience

Théorie $\longrightarrow$ Simulation $\longleftrightarrow$ Expérience

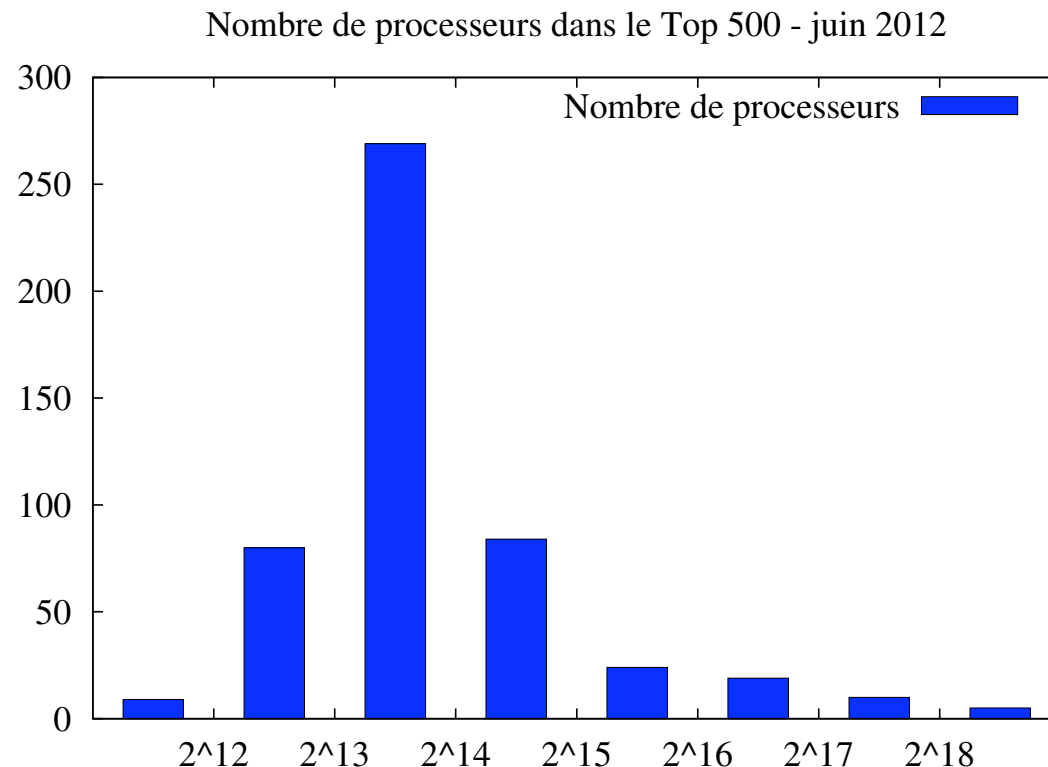
Le calcul haute performance :

- ▶ C'est un outil pour la recherche ;
- ▶ C'est un domaine de recherche scientifique en soit.

Nombre de processeurs dans le top 500 en juin 2012

► Augmentation significative en 5 ans :

41% des supercalculateurs avait entre 1024 et 2048 processeurs en juin 2007 ; Il y en a 54% entre 8192 et 16384 en Juin 2012.



(Source :
<http://www.top500.org/>)

Il y a en moyenne 26856 processeurs par supercalculateur du top 500 en juin 2012.

Puissance de calcul actuelle



(<http://www.top500.org/>)

Le plus puissant supercalculateur actuel :

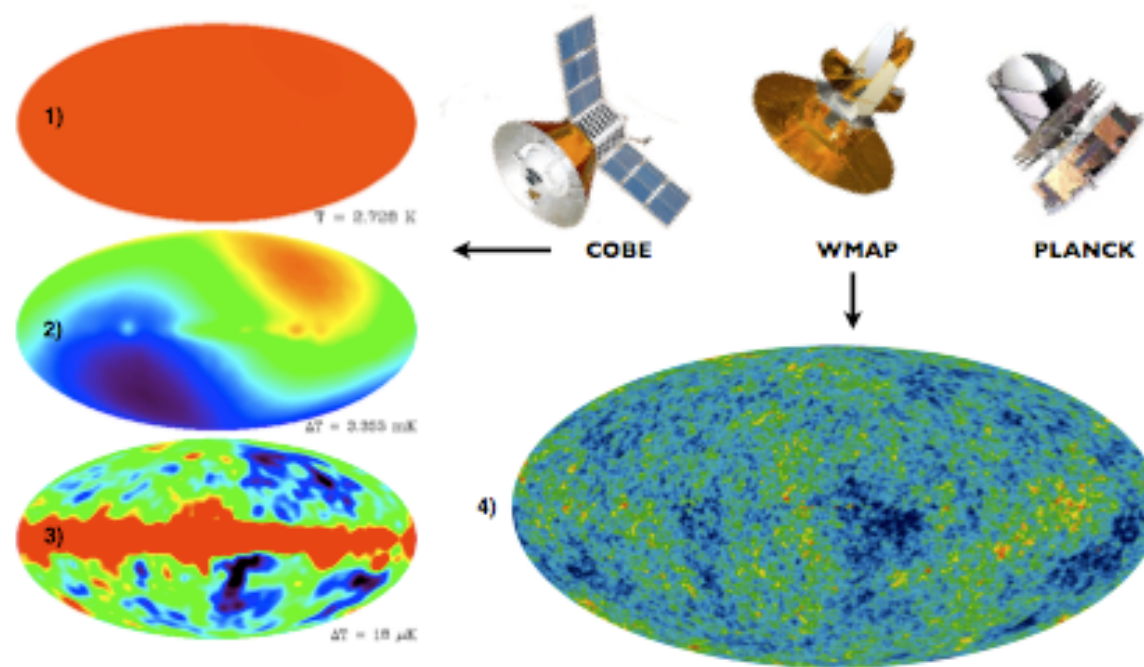
- ▶ Sequoia, IBM BlueGene/Q, Power PC, 1.60 GHz.
- ▶ Site : DOE/NNSA/LLNL, USA
- ▶ 98 304 nœuds de calcul, 1 572 864 cœurs
- ▶ 16.33 Pflop/s

Volume de données des expériences dans la cosmologie

COBE (1989) : 10 Gigabytes of data

PLANCK (2012) : 1 Terabyte of data

CMBPOL (2020) : 0.5 Petabyte of data (estimation)



(J. Errard, 2012)

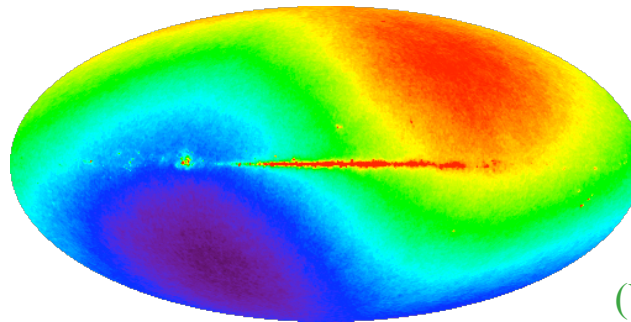
➡ Nécessite une importante puissance de calcul

Un peu de cosmologie

- ▶ Le CMB est extrêmement isotrope ;
- ▶ La température est globalement uniforme avec une température moyenne de 2.725 K (-270 °C) ;

Fluctuation de température pour le CMB :

$$\Delta T = 3.1 \text{ mK}$$

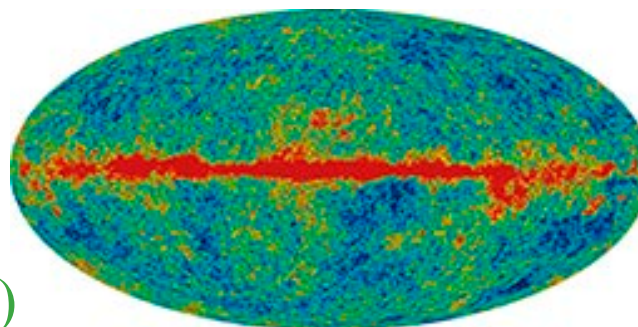


Variation dues au mouvement du Système Solaire

(WMAP, Nasa)

Après soustraction de ce dipôle :

$$\Delta T = 200 \text{ } \mu\text{K}$$



En rouge, c'est le plan galactique

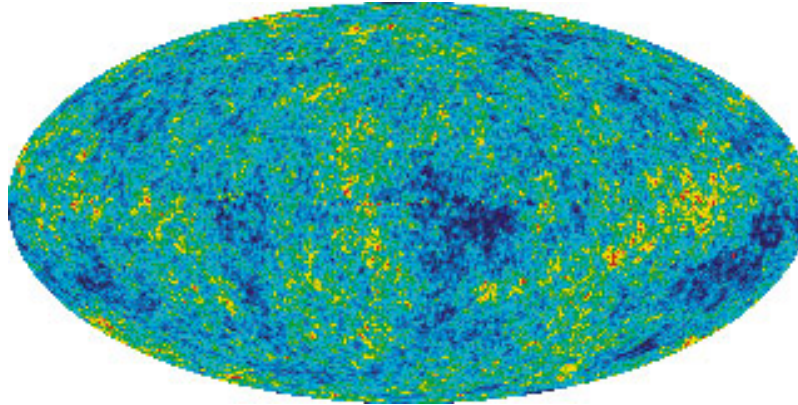
(61 GHz)

(WMAP, Nasa)

Un peu de cosmologie

- CMB déterminé par 5 ans de données WMAP :

$$\Delta T = 200 \mu\text{K}$$

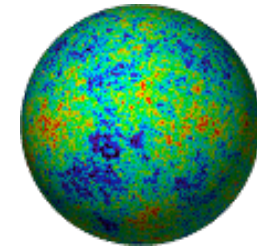


Le signal venant de notre galaxie a été soustrait en utilisant des données provenant de plusieurs fréquences.

Midapack : Motivation pour une bibliothèque de calcul

Fournir des outils de calcul pour l'analyse des données cosmologiques

- ▶ Haute performance ;
- ▶ Massivement parallèle ;
- ▶ Portable ;
- ▶ Incluant de nouvelles approches algorithmiques ;
- ▶ Capable de gérer un volume de données très important.



➡ Résoudre l'équation du « mapmaking » par le maximum de vraisemblance pour le fond diffus cosmologique (CMB)

Définition du problème « mapmaking » pour le CMB

Le problème « mapmaking » par le maximum de vraisemblance consiste à déterminer la meilleure estimation du « ciel » à partir d'un signal de données brutes temporelles, d'une stratégie de scan et d'une description du bruit.

Cela s'exprime sous la forme d'une équation par :

$$A_{pt}^T N_{tt}^{-1} A_{tp} \tilde{x}_p = A_{pt}^T N_{tt}^{-1} d_t$$

Avec:

- ▶ d_t le signal des données brutes temporelles ;
- ▶ A_{tp} la matrice de pointage qui relie chaque terme de d_t à un pixel du CMB par la relation :

$$d_t = A_{tp} x_p + n_t$$

- ▶ n_t le bruit instrumental, est supposé Gaussien et stationnaire par morceau.

Représentation du bruit instrumental

Les informations statistiques du bruit sont contenu dans la matrice de covariance $N_{tt} = \langle n_t n_t^T \rangle$

On assume généralement que N_{tt}^{-1} est

- ▶ une matrice symétrique et définie positive ;
- ▶ une matrice bande diagonale Toeplitz par morceau

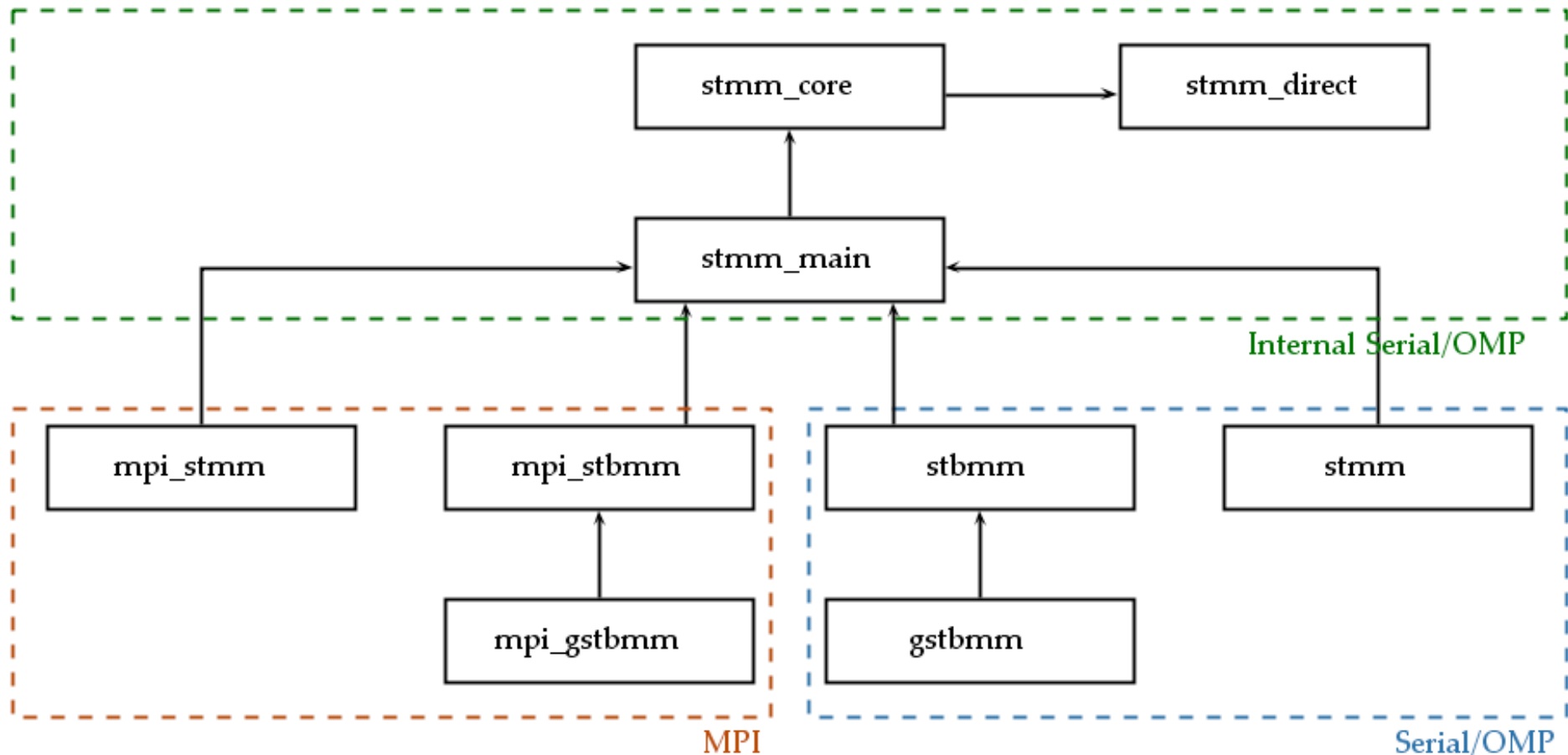
Cela nous permet d'exploiter des algorithmes performants :

- ▶ Algorithmes basés sur la FFT;
- ▶ Besoins en mémoire réduits.

➡ C'est généralement considéré comme un bon compromis pour la représentation de la corrélation du bruit du CMB.

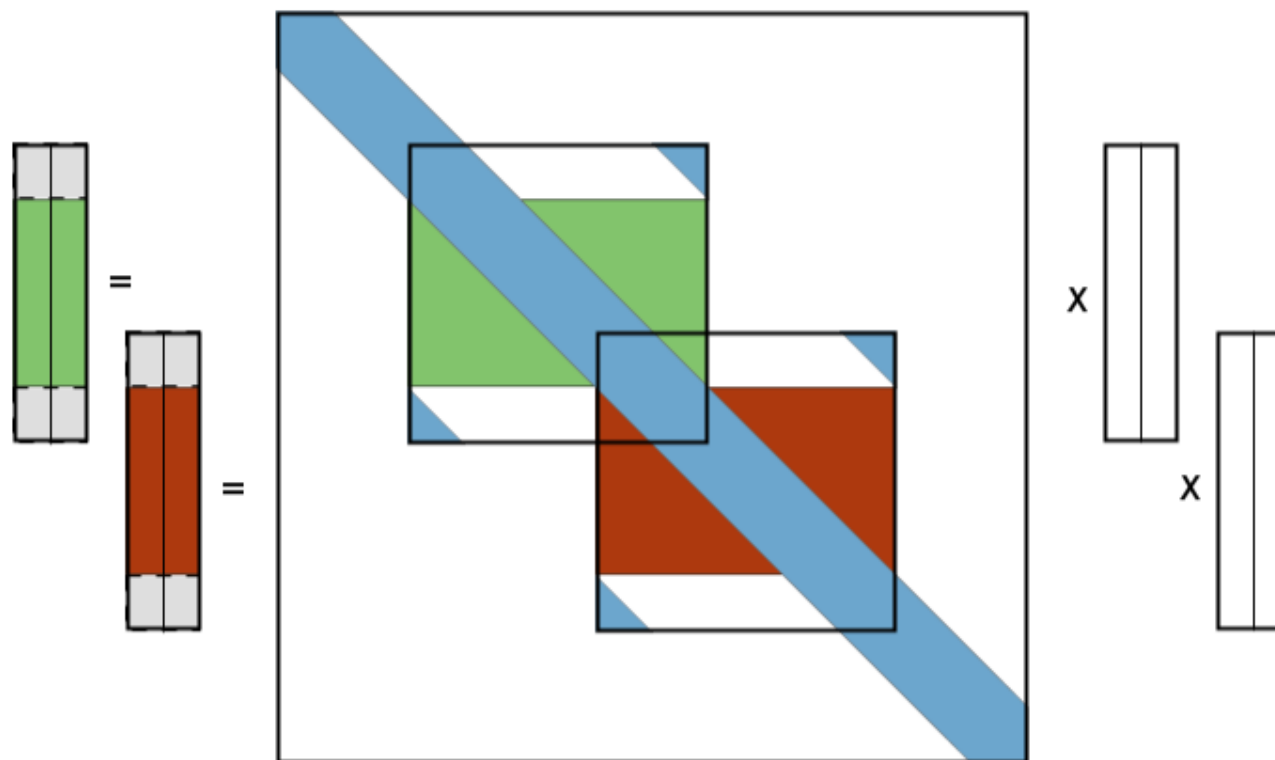
Module « algèbre de Toeplitz »

Diagramme de dépendances des routines API



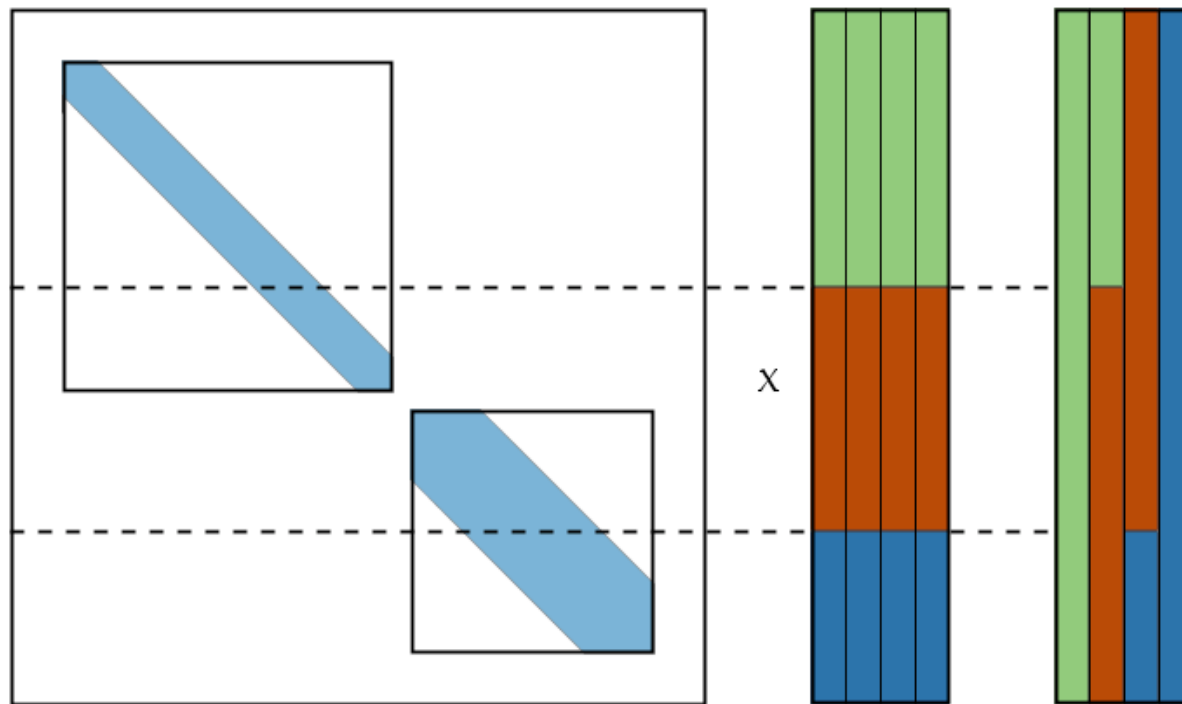
L'algorithme en fenêtre glissante - séquentiel

- ▶ Limite la taille des données à copier ;
- ▶ Les performances sont très dépendantes de la taille de ce bloc.



Toeplitz block diagonale – Parallèle MPI

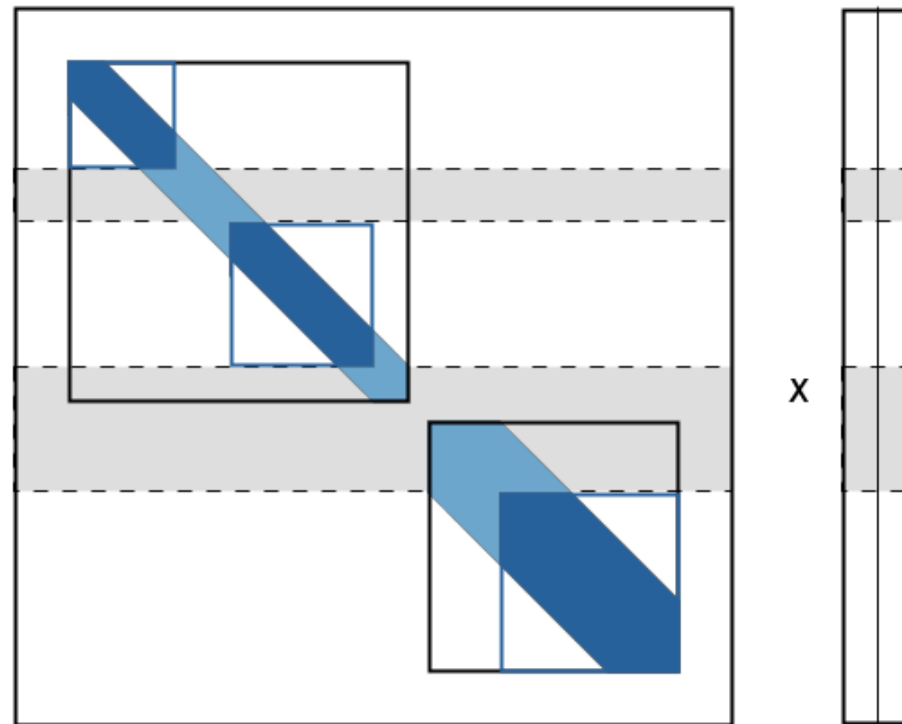
- ▶ Utilise des blocs flottants ;
- ▶ Les communications entre voisins sont faites seulement si besoin ;
- ▶ Il est possible de définir plus de blocs que nécessaire.



➡ Le nombre de blocs à définir localement dépend fortement de
La distribution des données.

Toeplitz block diagonale « gappy » – Parallèle MPI

- ▶ Utilise des blocs flottants ;
- ▶ Les communications entre voisins sont faites seulement si besoin ;
- ▶ Il est possible de définir plus de blocs que nécessaire.

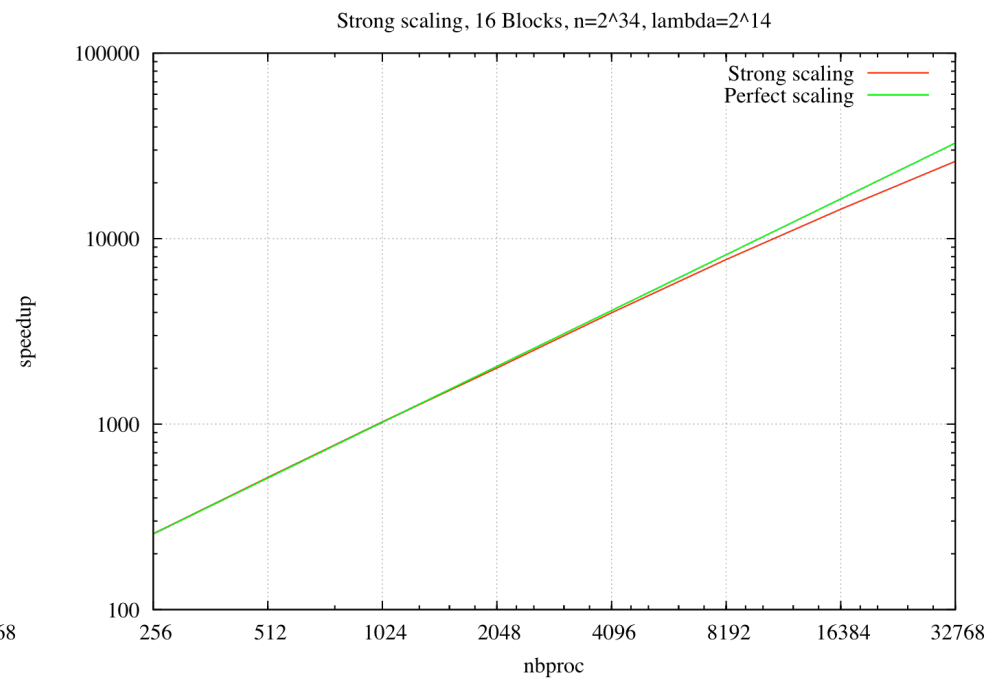
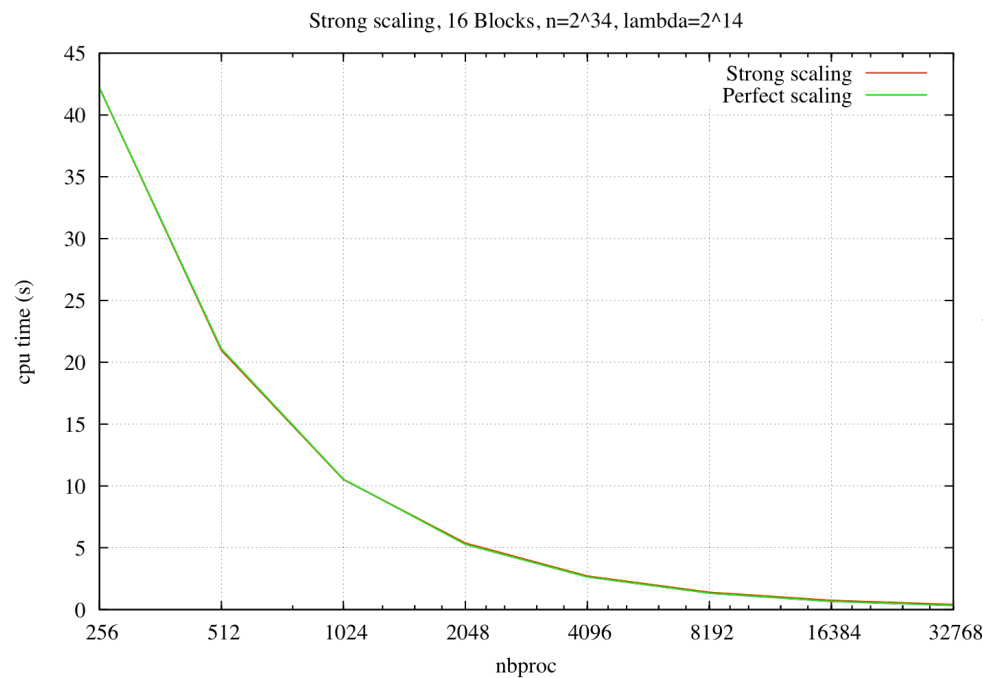


➡ Le nombre de blocs à définir localement dépend fortement de
La distribution des données.

Résultat d'extensibilité pour le « produit Toeplitz »

Extensibilité forte :

- Le volume de données reste constant, seul le nombre de processus change.



17 Milliards de données (doubles)
= 136 Milliards d'octets

$$\text{speedup} = \frac{t_1}{t_n}$$

Module « Matrice de pointage »

Cette matrice très creuse est de taille (m, n) avec :

- ▶ m la taille de la série temporelle locale au processeur ;
- ▶ n le nombre de pixels parcouru localement ;
- ▶ Le nombre de valeurs non nulles par lignes est très faible (de l'ordre de 1 à 5) pour les applications en cosmologie.

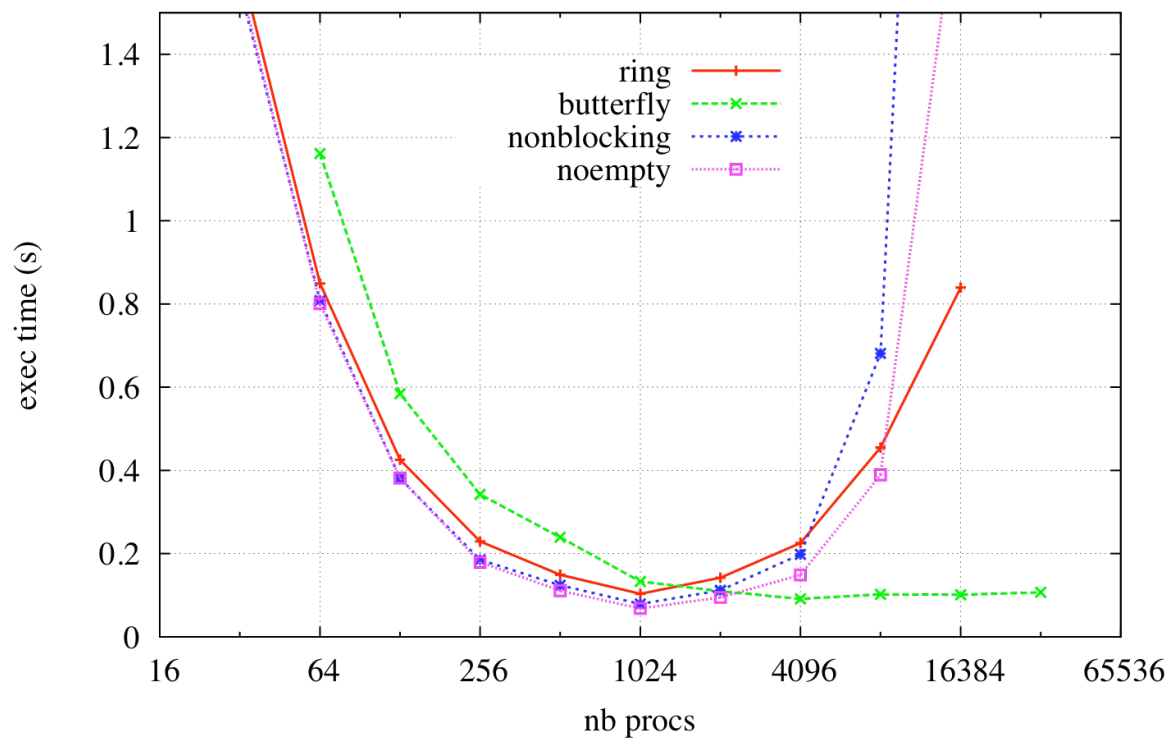
Les pixels peuvent être très éparpillés sur les processeurs :

- ▶ Les communications entre processeurs peuvent donc être très complexes pour le produit avec la matrice A_t ;
- ▶ Le temps de communication domine généralement le temps de calcul pour cette opération.

➡ Utilisation d'algorithmes spécifiques pour optimiser les communications.

Résultat d'extensibilité pour le « produit dépointage »

strong scaling hopper : $V := A^t U$, $nt=1638.4$ millions, $nnz=1$, $npix$ 12 millions



Un premier résultat de « mapmaking »

Résolution de l'équation du « mapmaking » avec la méthode du gradient conjugué préconditionné ;

Stratégie de scan :

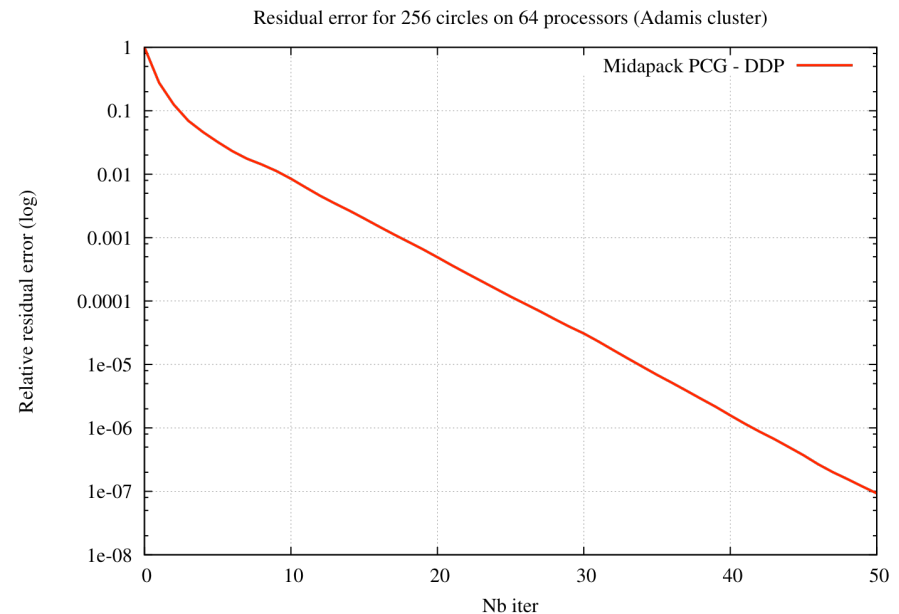
- ▶ 256 intervalles stationnaires de taille homogène 2^{20} ;
- ▶ Effectués par translation horizontale ;
- ▶ 1 valeur non nulle par ligne.

Matrice de Toeplitz :

- ▶ Largeur de bande de taille 2^{14} .

Distribution des données :

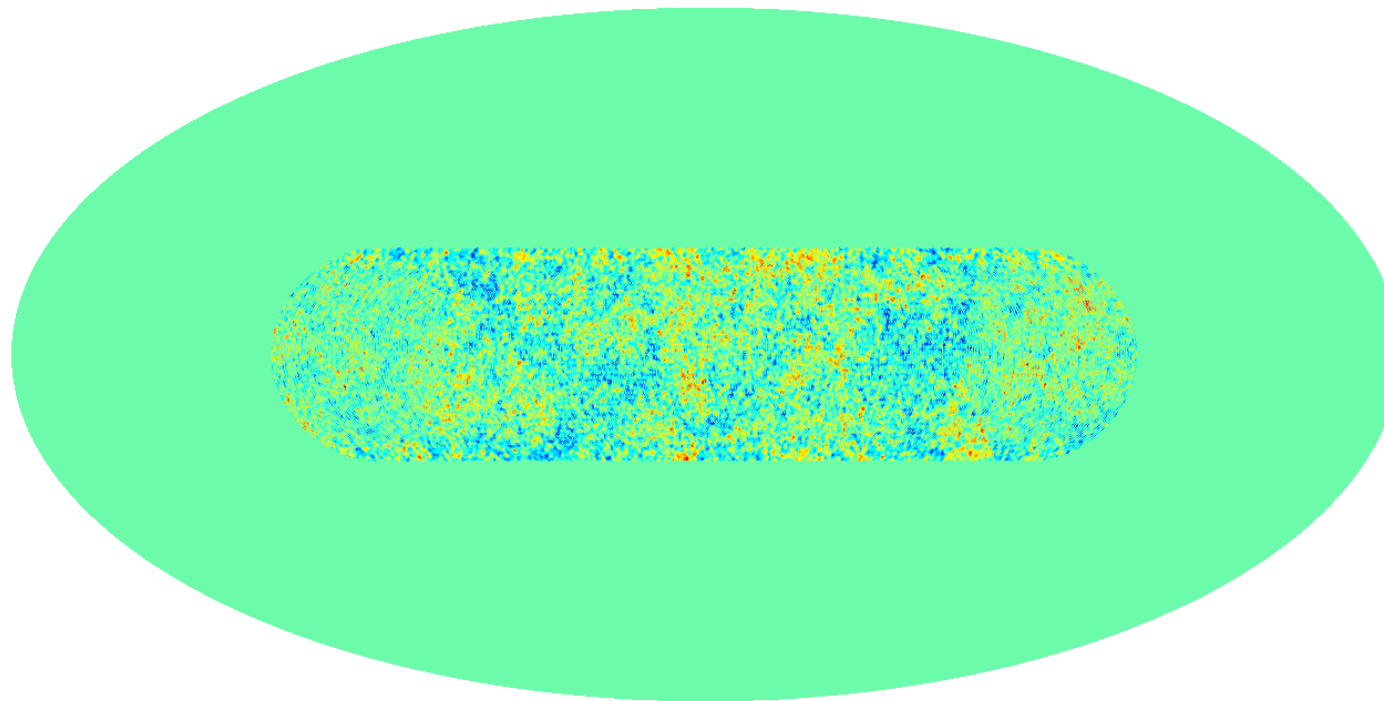
- ▶ Homogène par processus avec des intervalles stationnaires entiers.



Un premier résultat numérique

786432 pixels

Projection Mollweide



-391.  +400.

Conclusion

- ▶ Les éléments principaux de la librairie Midapack sont en place ;
- ▶ Cela fournit un ensemble de routines efficaces et flexibles ;
- ▶ Un paramétrage fin est aussi possible pour un utilisateur expert ;
- ▶ Il est déjà possible de résoudre l'équation du « mapmaking » ;
- ▶ **Midapack** est gratuit et téléchargeable sur le site web officiel (projet MIDAS no. ANR-09-COSI-009) :
http://www.apc.univ-paris7.fr/APC_CS/Recherche/Adamis/MIDAS09/software.html

Merci de votre attention.