

Grape-DRを用いた 平面波基底第一原理計算

鳥取大学 工学部 吉本 芳英

アクセラレータの活用

稲葉班と連携



GRAPE-DR



NVIDIA Tesla

構造サンプリングに必要な物

- 計算の経済性：性能／¥、性能／W
- 計算の精度：高コスト近似の実用化

アクセラレータ活用の計画

計算科学と計算機科学の連携を実践

● Phase I

- 平面波基底の第一原理計算を対象
- 計算機科学側で開発された技術をとにかく活用
- 対話にはよくわかっているたたき台が必要

● Phase II

- 高度な計算(cRPA, GW, Hybrid functional...)を対象
- 構造の評価には高度な計算が必要

Grape-DRシステムの 作成

部品

部品

- Grape-DR 4 chip model (稲葉グループより借用)

部品

- Grape-DR 4 chip model (稲葉グループより借用)
- マザーボード:ASUS P6T6 WS Revolution (Intel X58)

部品

- Grape-DR 4 chip model (稲葉グループより借用)
- マザーボード:ASUS P6T6 WS Revolution (Intel X58)
- CPU:intel Core i7 960 (4 core 3.2 GHz)

部品

- Grape-DR 4 chip model (稲葉グループより借用)
- マザーボード:ASUS P6T6 WS Revolution (Intel X58)
- CPU:intel Core i7 960 (4 core 3.2 GHz)
- メモリ:TEAM製 PC3-10600 4GB*6

部品

- Grape-DR 4 chip model (稲葉グループより借用)
- マザーボード:ASUS P6T6 WS Revolution (Intel X58)
- CPU:intel Core i7 960 (4 core 3.2 GHz)
- メモリ:TEAM製 PC3-10600 4GB*6
- ハードディスク:Seagate Barracuda (500GB)

部品

- Grape-DR 4 chip model (稲葉グループより借用)
- マザーボード:ASUS P6T6 WS Revolution (Intel X58)
- CPU:intel Core i7 960 (4 core 3.2 GHz)
- メモリ:TEAM製 PC3-10600 4GB*6
- ハードディスク:Seagate Barracuda (500GB)
- ビデオカード:SAPPHIRE HD 4350 512MB DDR2

部品

- Grape-DR 4 chip model (稲葉グループより借用)
- マザーボード:ASUS P6T6 WS Revolution (Intel X58)
- CPU:intel Core i7 960 (4 core 3.2 GHz)
- メモリ:TEAM製 PC3-10600 4GB*6
- ハードディスク:Seagate Barracuda (500GB)
- ビデオカード:SAPPHIRE HD 4350 512MB DDR2
- 電源:AcBel R9 Power 1100W

部品

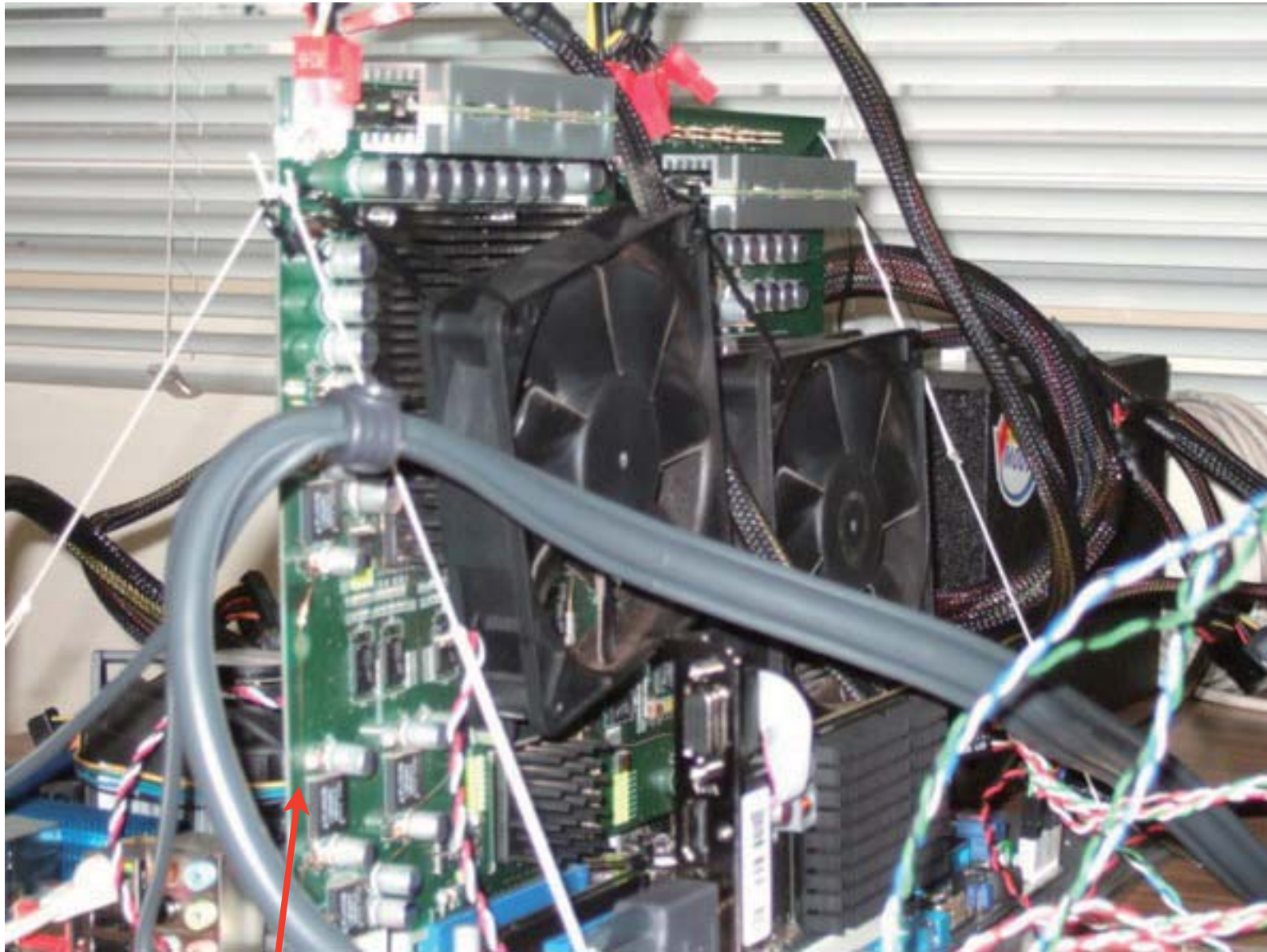
- Grape-DR 4 chip model (稲葉グループより借用)
- マザーボード:ASUS P6T6 WS Revolution (Intel X58)
- CPU:intel Core i7 960 (4 core 3.2 GHz)
- メモリ:TEAM製 PC3-10600 4GB*6
- ハードディスク:Seagate Barracuda (500GB)
- ビデオカード:SAPPHIRE HD 4350 512MB DDR2
- 電源:AcBel R9 Power 1100W
- ケース:なし

ケースなし？



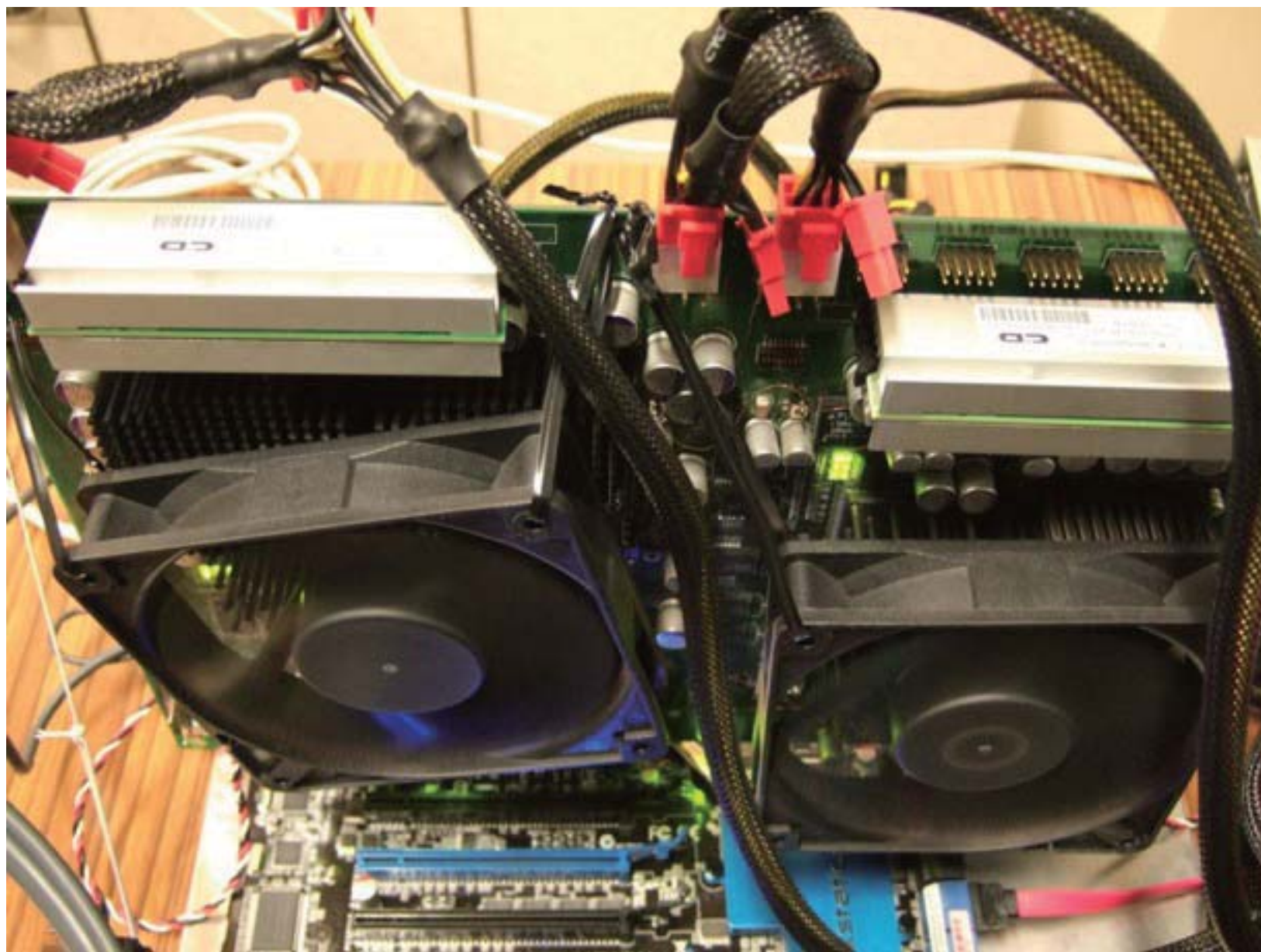
マザーボードを取り付けている金属板の正体



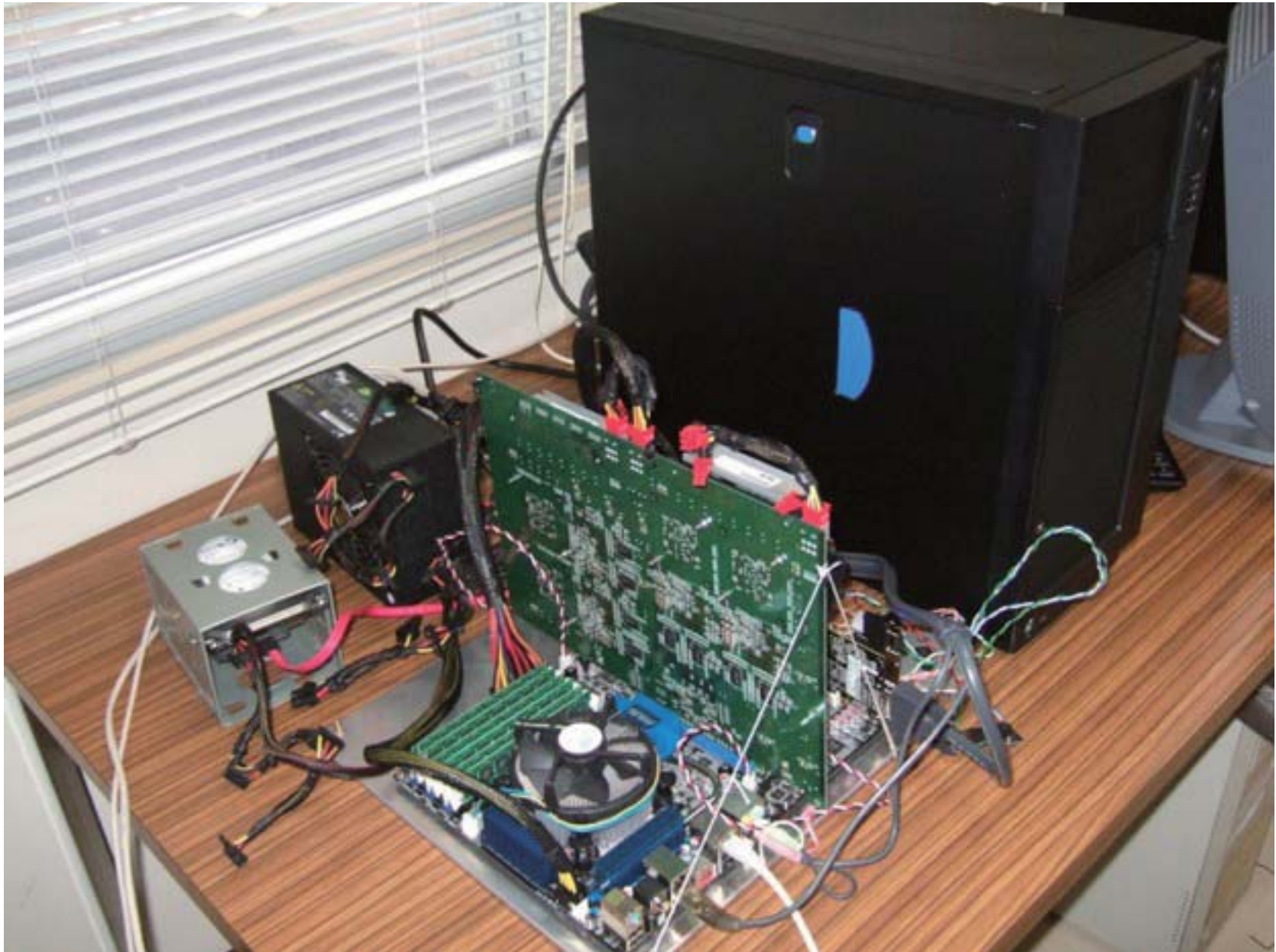


ボード上に銅線の空中配線がある

ファンはねじりっ子でつってある



システム全景



システムの体積は小さくない

右はwestmere * 2





ハードディスクと電源



感想

- このパッケージングはやり過ぎではないか？
 - ノイズの放射が気になる
 - ボードをひもで固定しないと、曲がるし、ぐらぐらする
- 汎用化することで経済性を確保するために
 - 計算機システムを構成する他の部品との調和を守るべき
 - ATXのフォームファクタは守るべき
 - ノイズの放射はさけるべき
 - 正常に動作しないシステムは速度0。リスク回避も必要ではないか。

素人が考えることは
当然検討されたはず

あえてこうなっている理由？

- 多くのchipを単一のボードに集積したい
 - ボード性能／ボード製造コストの向上
 - ボード上でのchip間通信をしたい
 - ボードのスロットの有効活用
 - 1 chipにつきPCIe x 4が限度／で十分と判断
 - 1スロットは最大PCIe x 16、束ねて使いたい
 - 1システム中のchip数を増やしたい
- GEMMは低優先度

他の（保守的な）可能性？

- 1ボード1 chip, PCIe x 8 or x 16
- ボードを薄く作る（ファン？）寸法は守る
- 1システムに2か3ボード
- ピーキーな性能ではなく粘り強い性能
 - 幅広いアプリケーションで性能を出して
支持を広げる

借用後、正常に起動するまで
2ヶ月ほどかかった

正常に起動するまでに問題だったこと

- 正しいドライバを入手すること
 - K&F computing版と国立天文台版がある
- 正しいfirmwareをインストールすること
 - 制御チップのプログラムのこと
 - チップ自体の不揮発メモリに書き込んである
 - ボードによって入っているfirmwareが違っていた
 - 入っているfirmwareの種類を区別する方法はない
 - 稲葉班側では問題なく動いたのはこのため

思い込みで混乱したこと

- 吉本はウェブ上の情報だけからK&F Computingのドライバを使った
 - 正しいドライバは国立天文台の物
- 動かないのは電源のためという説が稲葉班側が強く主張した
 - 稲葉班側でのテスト機はすぐに動いた
 - 問題は電源ではなくfirmwareだった

解決に必要なだったこと

- 正しい情報の入手
 - 市販製品とは異なり、webは当てにならない。
- リバースエンジニアリング
 - ドライバの古いコードをつかうと部分的に正常動作した。→firmwareが古いのでは？
- 現物での確認
 - firmwareが問題であることを確認するため、一度東京にボードを持ち込んだ

平面波基底プログラムの移植

- パッケージにはxTAPPを使う
- 行列積相当の部分だけ変更
 - 元々行列積相当の部分を分離してある
 - Gamma点計算に対応する部分だけ
- 原子間力（Hellman-Feynman力）は時間切れで対象外

ベンチマークの条件

- Pyrope(鉍物結晶の一種)160*4原子
- $N_{PW} = 85543, N_B = 1600, N_K = 1, N_{XYZ} = 128^3$
 - $E_c = 25 \text{ Ry}$: メモリが足りないため
- ランダムな初期波動関数
- 原子の電荷分布から初期電荷分布
- SCFが収束するまで、力の計算は除外

速度比較

	Grape-DR dgemm	Core i7 MKL dgemm	Core i7 hand opt	Core i7 naive
行列積 [s]	3346	9989	33872	
比	1	2.99	10.12	
FFT [s]	1335	1333	1339	
固有系 [s]	338	335	337	
その他 [s]	1770	1774	1764	
合計 [s]	6788	13431	37313	>100000
比	1	1.98	5.50	>14.7

分析

- MKLのdgemmは40GFlops程度と思われる

➡ Grape-DRのdgemmは120GFlops程度

✓ バンド数1600でまだ小さい

Grape-DRの最大性能*は400~600GFlops

300GFlops出すのに行列サイズ5000

- その他の部分は $O(N^2)$ のはず
 - その他が妙に大きい。分析の必要がある

* 中里@ 会津大学 行列乗算カーネルの性能評価の「牧野 秋期天文学会 (2010)」

予想

- 原子数が大きいと性能が出るはず
 - 演算量は $O(N^3)$ vs $O(N^2)$ ：原子数に比例して $O(N^3)$ が優越
 - 通信量も $O(N^2)$
- $640*8 = 5120$ 原子でMKLから4倍の速度？
 - MPI並列になるGrape-DRのクラスタが必要
 - メモリが64倍必要なので64ノード以上
 - 行列の対角化をScaLapackで行って効率33%を前提

予想速度 640*8原子

640原子

640*8原子, 64 MPI proc

	Grape-DR dgemm	Core i7 MKL dgemm	倍率	Grape-DR dgemm	Core i7 MKL dgemm
行列積 [s]	3346	9989	$8^3=512$	10707	79912
比	1	2.99		1	7.46
FFT [s]	1335	1333	$8^2=64$	1335	1333
固有系 [s]	338	335	$8^3=512$	8112	8040
その他 [s]	1770	1774	$8^2=64$	1770	1774
合計 [s]	6788	13431		21924	91059
比	1	1.98		1	4.15

Grape-DR 120GFlops

Grape-DR 300GFlops

ScaLapack 33% eff.

課題

- 力の計算をあからさまに行列積化
- 反転対称性がなく gamma点でもない場合の計算にはzgemmが必要
- $O(N^2)$ の部分の非効率を改善

電力の観点からの調査？

- 行列積, FFT, その他
 - 消費電力的な割合？
 - 電力効率？ (GFlops/W)
- 現在、可能性
 - メモリ律速の計算：CPUクロック下げる
 - 演算律速の単純計算：アクセラレータで実行
 - 演算律速の複雑計算：CPUが通常クロック

結論

- Grape-DRの本来の性能ができる
→原子数が5000以上？
 - 研究対象物で研究する場合
 - 連絡を密に
 - 実物を共有、できれば動くところまで指導
- ➡そういう人間関係を建設する必要