

GRAPE-DR と GRAPE-8

牧野淳一郎
理学研究流動機構

の前に

**の前に
規定発表の真似事的何か**

研究室構成、論文数等

	人数	論文数
教授	1	1
特任准教授	1	1
特任助教	1	0
D3	1	0
その他		7

- 本人が好きそうなことを
- 成果がでそうな(人がまだやってない)方向に

Kominami, J. D., Makino, J., Daisaka, H. 2011. Binary Formation in Planetesimal Disks. I. Equal Mass Planetesimals. PASJ 63, 1331-1344.

Oshino, S., Funato, Y., Makino, J. 2011. Particle-Particle Particle-Tree: A Direct-Tree Hybrid Scheme for Collisional N-Body Simulations. PASJ 63, 881.

Fujii, M. S., Baba, J., Saitoh, T. R., Makino, J., Kokubo, E., Wada, K. 2011. The Dynamics of Spiral Arms in Pure Stellar Disks. The Astrophysical Journal 730, 109.

外部資金

- HPCI 戦略プログラム 2011-15 3000万くらい
- 基盤 A(代表) 2009-11 900万くらい
- 基盤 S 2009-11 500万くらい

来年度は今のところ HPCI 戦略プログラムのみ (特別推進に出してたけど落ちた)

共同研究グループ

- 鹿児島大学 和田、藤井: 銀河円盤、銀河内の星団進化
- 一橋大学 台坂: GRAPE ハードウェア
- 会津大学 中里: GRAPE ハードウェア、ソフトウェア
- IAS Piet Hut, Leiden U Simon Portegies Zwart, Northwestern U Will Farr, 筑波大学 谷川 ... : 星団の力学進化
- 筑波大学 似鳥、石山、: N体コード開発、宇宙論、
- UC London, 井上: 銀河形成
- Kyung Hee U S. S. Kim: 銀河中心、星形成

講義他

- 講義: なし
- 学科内委員: なし
- 学内委員: GSIC 運営委員
- 学外委員: 良く憶えてないけど一番面倒だったのは研究振興局情報課の下にできた「今後のHPC技術の研究開発を検討する作業部会」
- 社会貢献
 - 福島原発事故関係でシンポジウム等、座談会、雑誌原稿等
 - 貢献かどうか不明な気も

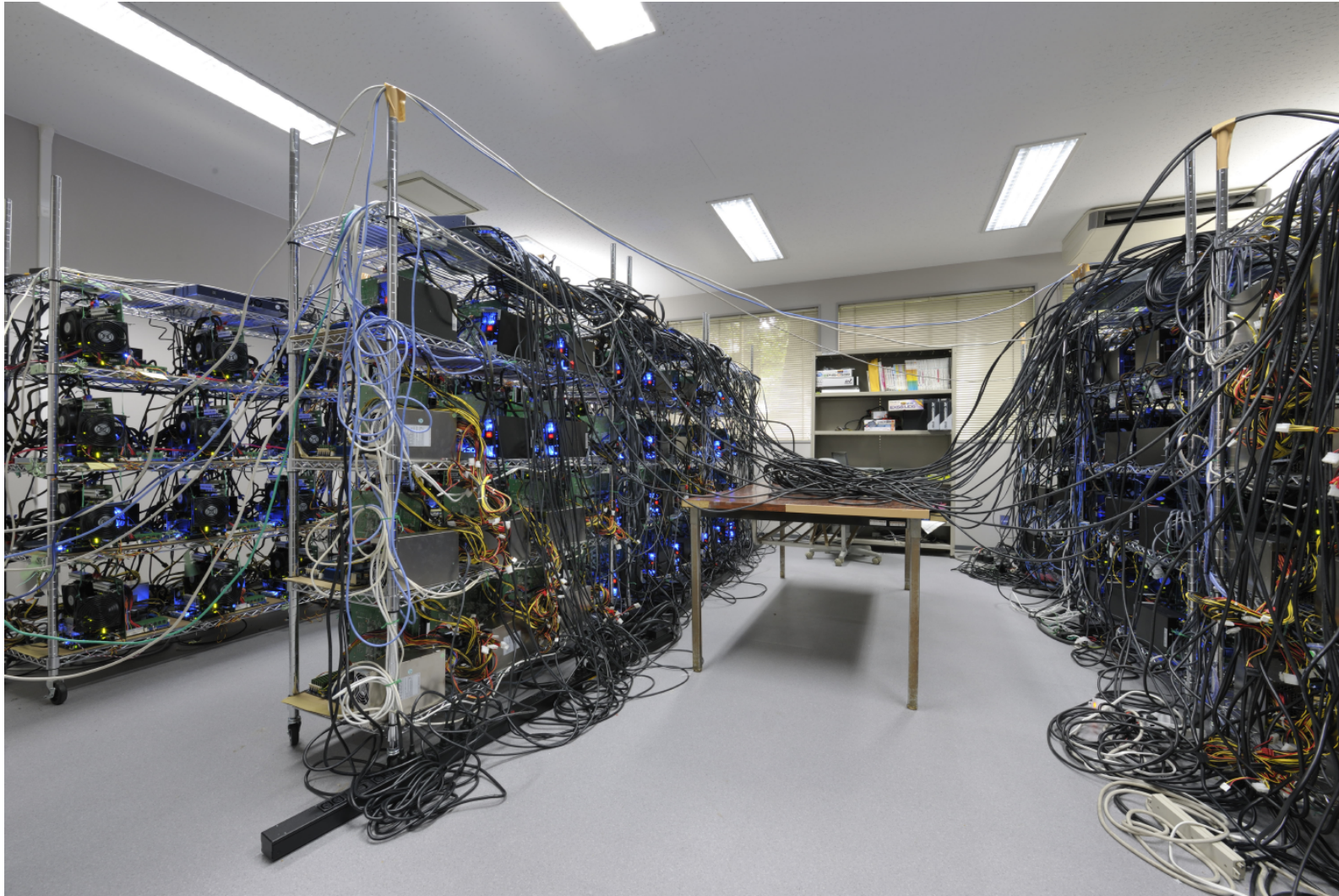
GRAPE-DR と GRAPE-8

- GRAPE-DR
- GRAPE-8

GRAPE-DR

- 従来の GRAPE(重力計算の専用ハードウェア) とは違いアーキテクチャ
- 「汎用」ということで予算とる
- 非常に単純で小さなメモリだけをもつ演算プロセッサを多数並べる SIMD 構成
- 何故か最近の GPU と良く似た方向
- 開発予算: 振興調整費 (FY2004-2008)

GRAPE-DR クラスタシステム



GRAPE-DR クラスタシステム

前のスライドの写真とった時の構成:

- 128-ノード, 128-ボード (105Tflops peak)
- ホスト計算機: Intel Core i7+X58 12-24 GB メモリ
- ネットワーク: x4 DDR インフィニバンド
- 現在部屋整備中のはず (年度内に工事が終わるはず)

性能

行列乗算

$M=N$, $K=2048$,
720 Gflops

$N=K=2048$, 500 Gflops

現在カード 1 枚の性能・電力当り性能としては世界一

LU 分解

行列サイズに対する性能 (Gflops)

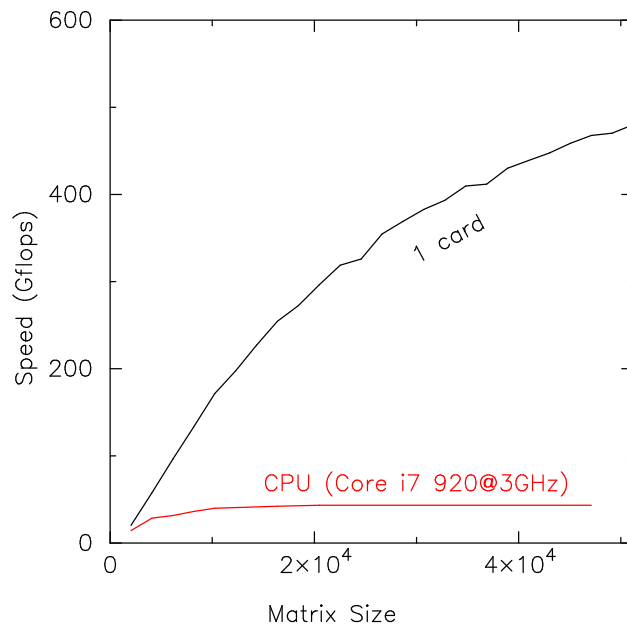
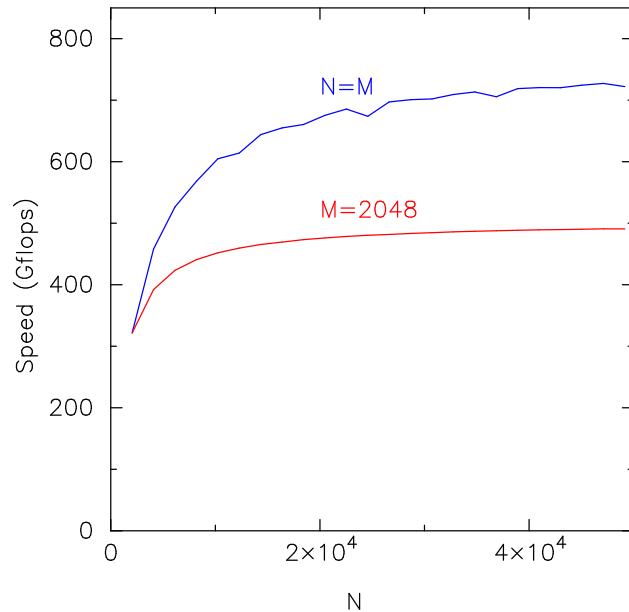
上:GRAPE-DR カード 1 枚

下:Core i7 920(3GHz 動作)

行列格納順序の変更・処理途中で転置

その他色々新しいアルゴリズム開発

1 カードの性能として世界一



Little Green 500, June 2010

Green500 Rank	MFLOPS/W	Site*	Computer*	Total Power (kW)
1	815.43	National Astronomical Observatory of Japan	GRAPE-DR accelerator Cluster, Infiniband	28.67
2	773.38	Forschungszentrum Juelich (FZJ)	QPACE SFB TR Cluster, PowerXCell 8i, 3.2 GHz, 3D-Torus	57.54
2	773.38	Universitaet Regensburg	QPACE SFB TR Cluster, PowerXCell 8i, 3.2 GHz, 3D-Torus	57.54
2	773.38	Universitaet Wuppertal	QPACE SFB TR Cluster, PowerXCell 8i, 3.2 GHz, 3D-Torus	57.54
5	536.24	Interdisciplinary Centre for Mathematical and Computational Modelling, University of Warsaw	BladeCenter QS22 Cluster, PowerXCell 8i 4.0 Ghz, Infiniband	34.63

電力当りの性能で、しつこいですが、世界一を達成

#2: IBM PowerXCell, #9: NVIDIA Fermi

GRAPE-8

GRAPE-DR の次をどうするか？

- GRAPE-DR 的な、プログラマブルなものの次世代機
- 従来の GRAPE 的専用機
 - － カスタムチップを作るのは無理
 - － Structured ASIC (後述) を想定

2008 年にに前者を特別推進、後者を基盤 (A) で申請して、前者はヒアリングでこけたので後者になった。

Structured ASIC

カスタム LSI と FPGA (論理回路変更可能な LSI) の中間

	ASIC	FPGA	Str. A.
--	------	------	---------

初期費用 (10 億円)	1	0.001	0.02
--------------	---	-------	------

回路規模 (1 千万ゲート)	20	1	1
----------------	----	---	---

ゲート単価 (円/1000 ゲート)	0.3	10	2
--------------------	-----	----	---

良くいうと、ASIC より初期費用安くて FPGA より量産費用安い

GRAPE-8 チップ

- eASIC 社 Nextreme2 N2X740 チップ使用、74万素子
(1素子は10ゲート程度に相当)
- 重力計算パイプライン 48本集積
- 250MHz 動作
- 480Gflops 相当の速度、消費電力 13W、37Gflops/W
- チップ単体の電力当り性能で IBM BG/Q の 10倍程度

GRAPE-8 ボード



GRAPE-8 ボード

- ボード論理回路設計は台坂
- PCI-Express インターフェースは FPGA を使う。GRAPE-DR の設計流用
- GRAPE-8 チップ 2 個搭載
- 理論ピーク性能 960 Gflops 相当、消費電力 47W。20Gflops/W

計算精度は十分か？

ツリーコードと直接計算のハイブリッドなら「多分」大丈夫

PPPT (Particle-Particle, Particle-Tree), Oshino et al. 2012

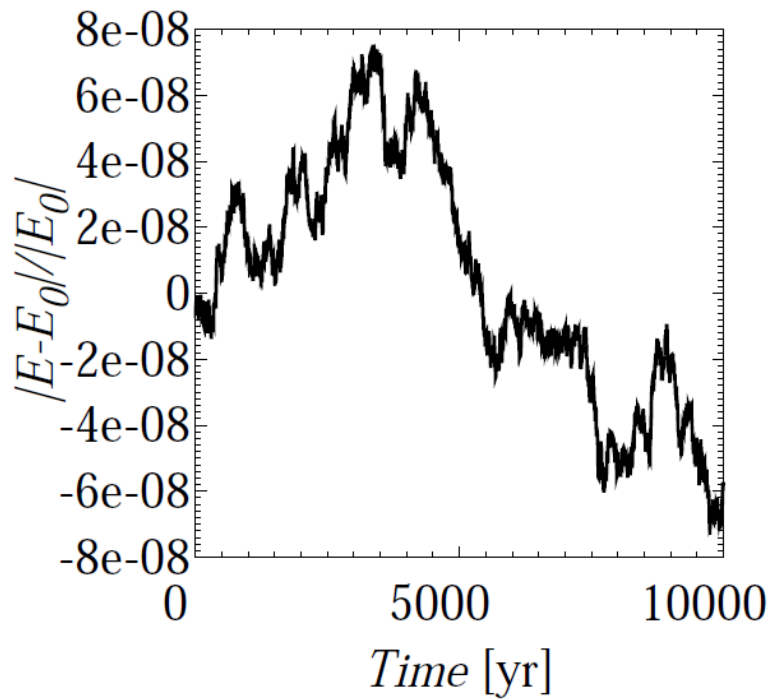
$$H = H_{Hard} + H_{Soft},$$

$$H_{Hard} = \sum_{i=1}^N \left(\frac{p_i^2}{2m_i} - \frac{Gm_i m_{\odot}}{r_i} \right) - \sum_{i < j}^N \frac{Gm_i m_j}{r_{ij}} W(r_{ij}),$$

$$H_{Soft} = \sum_{i < j}^N \frac{Gm_i m_j}{r_{ij}} (1 - W(r_{ij})),$$

- P³M スキームと同様に近距離力と遠距離力を分割
- 遠距離力はツリー、近距離力は独立時間刻み直接計算

PPPT テストラン



地球領域、微惑星 10^4 粒子

ツリーコードの力の誤差が主要因

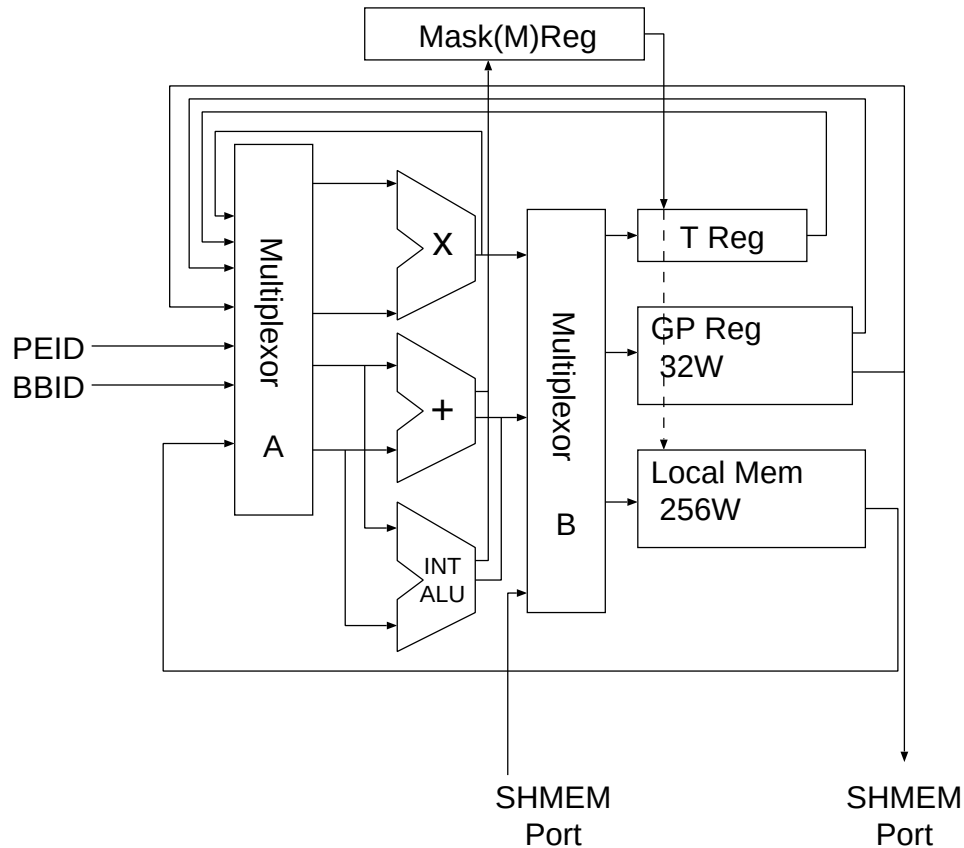
エラーがランダムウォークする (累積しない)

惑星形成では十分な精度

まとめ

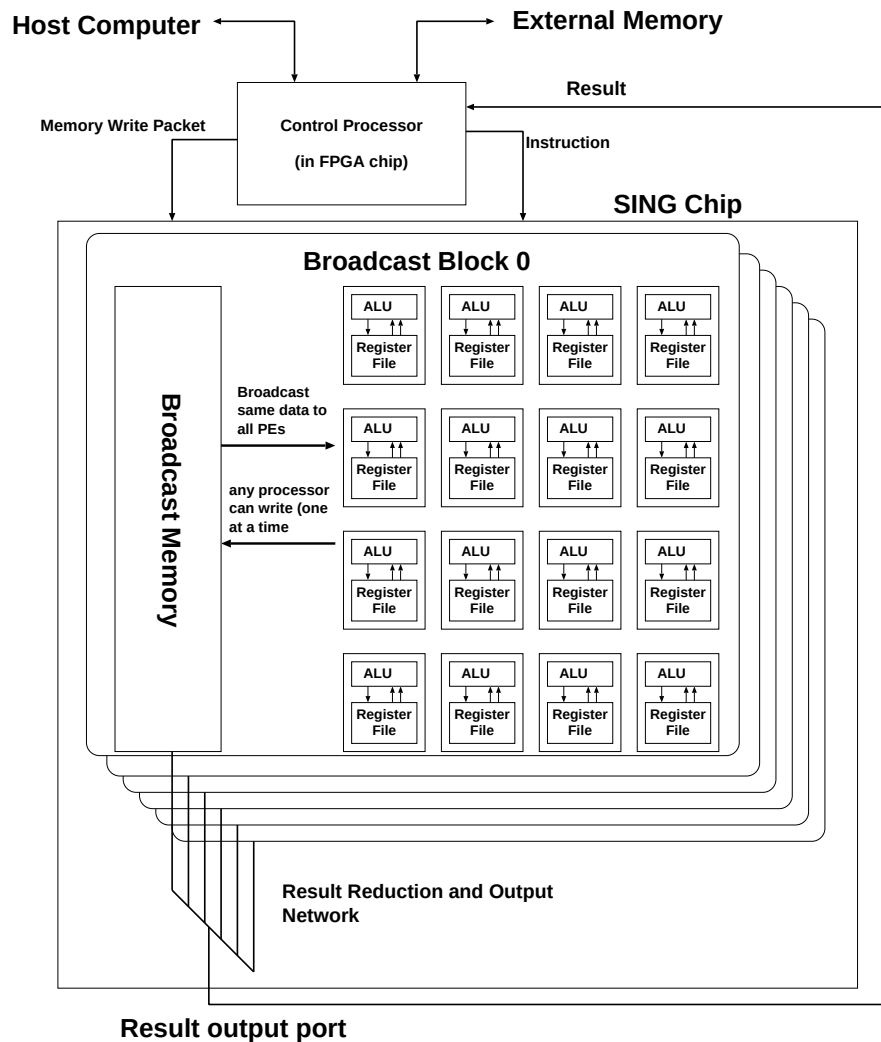
- GRAPE-DR は大体出来て、性能もでるようになった。
- カード 1 枚の性能は 4 コア CPU の 15(DGEMM)-20(重力) 倍
- GRAPE-8 もできた。PPPT スキームと組み合わせて星団、惑星形成に使う。
- 電力当り性能では世界一達成。

プロセッサアーキテクチャ



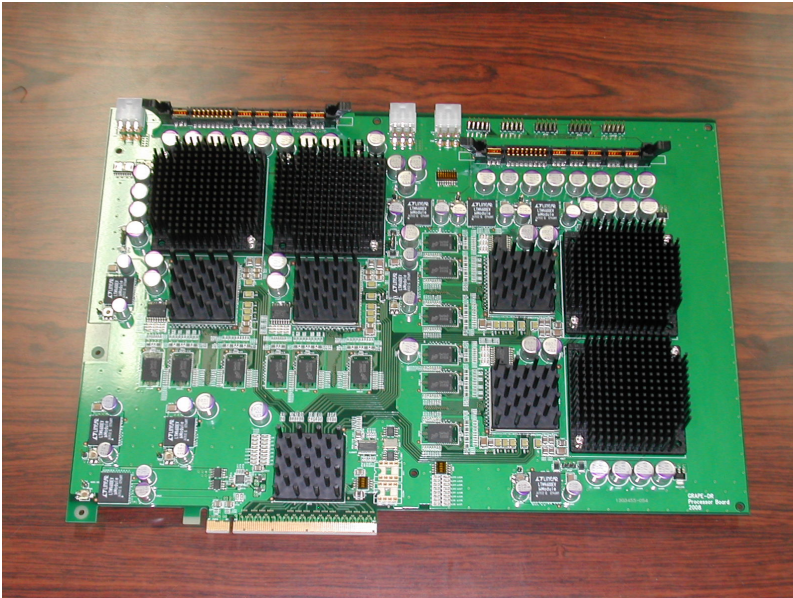
- 倍精度乗算器 (の半分)
- 倍精度加減算器
- 整数・論理演算器
- 32語のレジスタ
- 256語のメモリ
- 通信ポート

プロセッサチップ構成



- 32 個 (図では 16) のプロセッサで「放送ブロック」を構成。放送ブロックは 16 個。
- 放送ブロック毎に共有メモリ
- 外からの入力データは全ブロックに放送
- 各ブロックからの出力データを合計等しながら出力可能

プロセッサボード



PCIe x16 (Gen 1) インターフェース

Altera Arria GX を DRAM 制御、通信インターフェースに利用。

- 最大電力消費 200W くらい
- 433MHz くらいまでは問題なく動作
- 400 MHz での理論ピーク性能は 819Gflops
- KFCR で販売中