

GRAPE-DR

牧野淳一郎
(東京大学理学系研究科天文学専攻)

概要

- GRAPE を振り返る
- 何故 GRAPE とは違うことを始めたか
- 今後の HPC の方向

GRAPE を振り返る

- 基本的には大成功であったと思う (自画自讃)
 - GRAPE-4, GRAPE-6, MDM: 実アプリケーションでの計算速度世界一を実現。同時期、同等 (以下) の速度の汎用計算機の 1/100 のコスト
 - 世界中で広く使われている
 - * 東大、国立天文台、筑波大、東工大、....
 - * ケンブリッジ、ハイデルベルグ、ロチェスター、....
 - 全世界で 30 以上の研究機関、100 Tflops 以上の計算速度
 - 科学的な成果も結構でている (特に海外の機関が導入したのはその現われ)

では、何故そのまま続けないか？

- LSI の初期コストが高くなりすぎた

1990 1 μ m 1500万円

1997 0.25 μ m 1億円

2004 90nm 3億円以上 (某海外メーカー)

2006 65nm 10億円以上 (某国内メーカー)

理論天文にしか使えないものに出せる限度を超えた？

ではどうするか？

1. 頑張っって沢山お金をとる。「巨大科学」を目指す
2. やめて普通の計算機を使う
3. FPGA で色々やる
4. なんか違うものを作る

1-3 は他にやってる人がいくらでもいるので、
人がやらないことを = 4

どう違うものを？

目標:

- GRAPE くらい速くて安い
- 天文以外に色々できる

そんなことは可能か？

どう違うものを？

目標:

- GRAPE くらい速くて安い
- 天文以外に色々できる

そんなことは可能か？

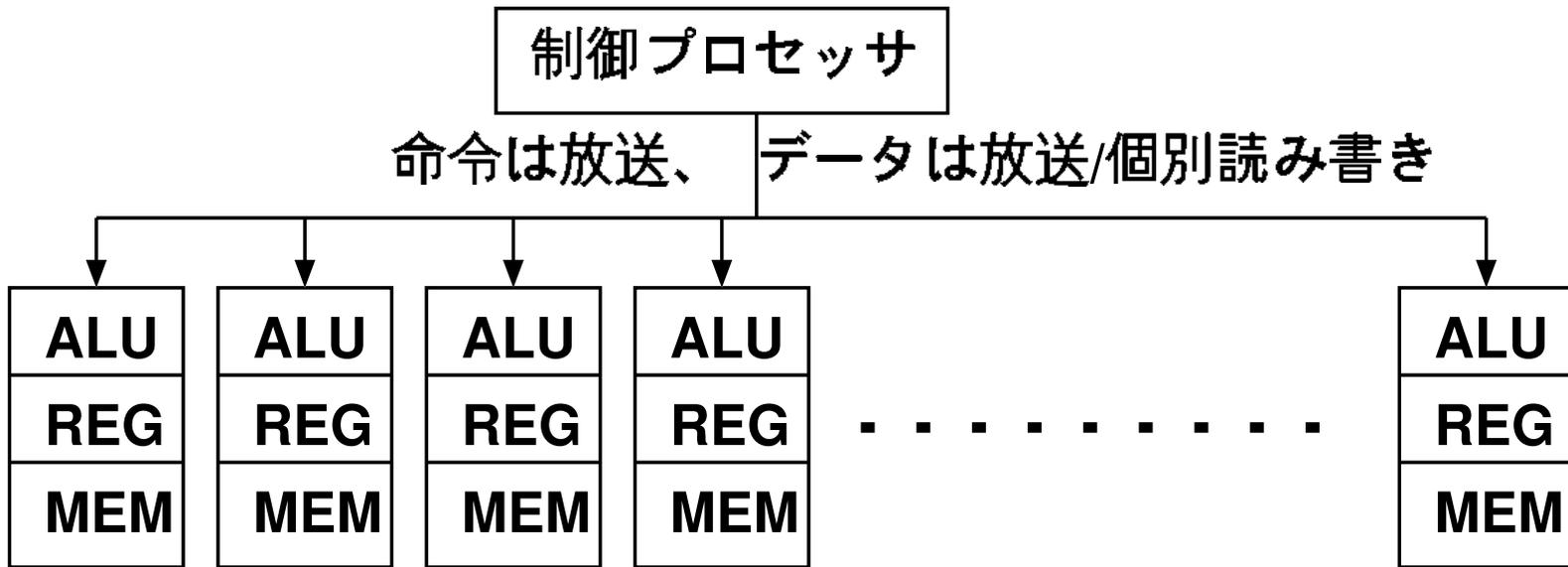
もちろん、全く文字通りには無理。

GRAPE が速い2つの理由

1. 相互作用計算に特化したパイプラインプロセッサでトランジスタ利用効率を上げる
 - こっちは捨てる
2. 多数の(論理)パイプラインが計算中は同じデータを受け取ることでメモリバンド幅を節約
 - こっちは残す

ブロードキャスト並列モデル

ブロードキャスト並列モデル



- 実は古典的な SIMD とこの図は同じ。
- 実際の GRAPE-DR はちょっと違う。
- 違いは想定する使い方からくる。

実際のGRAPE-DR チップ

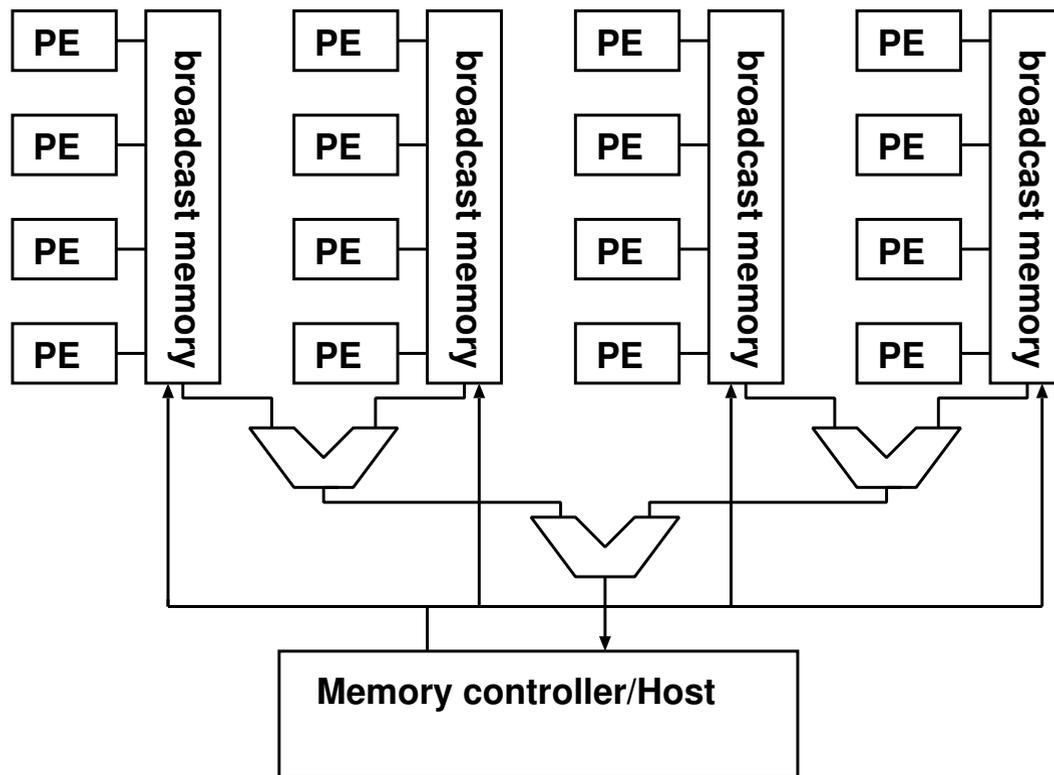
PE(要素プロセッサ) は放送メモリとだけ接続

放送メモリからそれにつながった PE には

- 放送
- ランダム読み書き

チップ外ポートから放送メモリには

- 放送
- 縮約 (総和など)
しながら読み出し
- ランダム読み書き



アプリケーションに対する考え方

- Memory Wall が問題にならないようなアプリケーションのみを対象にする
- 3つの型に特化
 - 散乱実験型
 - 粒子間相互作用型
 - 密行列型
- 可能ならばアプリケーションを書き換える

散乱実験型

- 多数の PE が、独立にイベントを計算
 - イベント間の相互作用はない、または非常に少ない
 - * 光線追跡：レンズ自動設計
 - * アナログ回路シミュレーション：回路自動設計
 - * 放射線伝播のモンテカルロ計算：検出器設計
- “Embarassingly Parallel” とほぼ対応
- 古典的 SIMD 機と同様の振る舞い：
 - Goodyear MPP, ICL DAP, TMC CM-1/2, Maspar MP-1/2
 - 極端に少ないメモリ
 - PE 間通信が遅い
- 計算速度と通信速度の比：
 - 散乱実験の計算がどれだけ複雑かで決まる

粒子間相互作用型

$$f_i = \sum_j f(x_i, x_j)$$

- 他の「粒子」との「相互作用」を縮約
 - 全ての相互作用を並列に計算可能
 - 同じ「粒子」のための計算結果を高速に縮約する必要
- 計算手順
 - PE に相互作用を受ける粒子をロード
 - 相互作用を及ぼす粒子をロード
 - 計算機終了したら結果を縮約しながら回収
 - 計算速度とチップ外への通信速度の比:
相互作用を及ぼす粒子数に比例

密行列型

$$c_{ij} = \sum_k a_{ik} b_{kj}$$

- 計算手順

- 行列が PE に収まるところまで分割。それから
- 行列 A の部分行列を PE にロード
- B の1列を分解して各グループにロード
- 各 PE で B の部分列と A の部分行列の積を計算
- 計算が終わったものから順次回収。グループ間で合計

- 計算速度・通信速度の比はチップ全体にロードできる行列のサイズに依存

- メモリサイズの平方根に比例して通信速度を落とせる

これでどれくらい速くできるか？

GRAPE-DR チップの場合:

2006/1 テープアウト

2006/4/X ES

90 nm, 18mm角チップに

- 512 プロセッサ
- ピーク性能 1 Tflops (単精度)
0.5Tflops (倍精度)

参考: AMD Athlon 64 FX-60: 10.4Gflops
(0.01 Tflops)

最終システムのイメージ

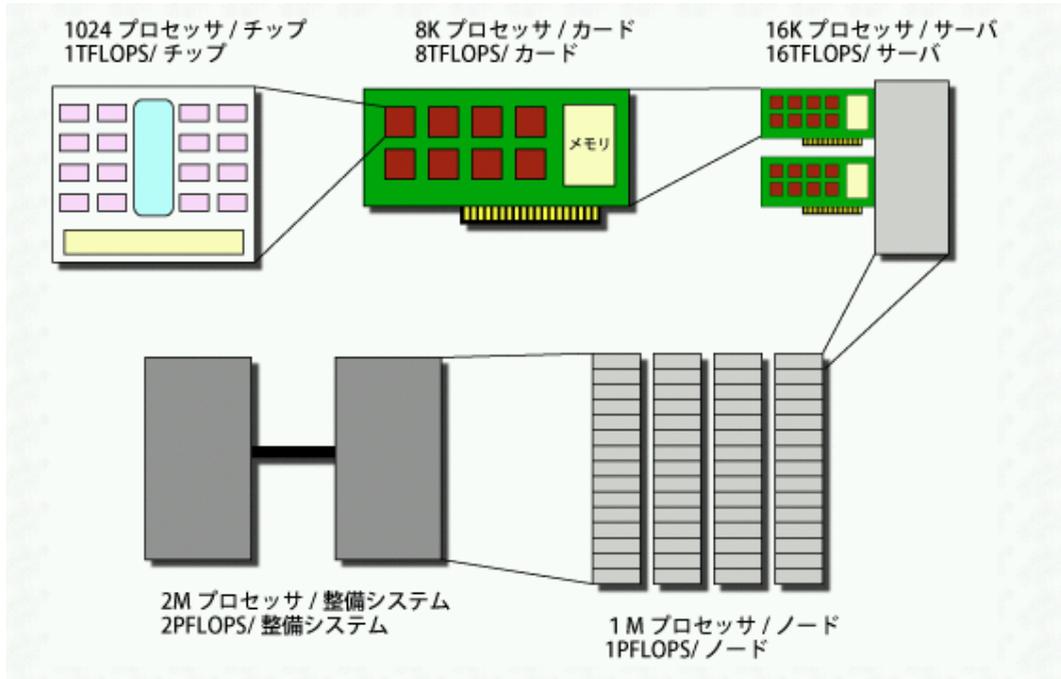
2007年度に基本的には完成 (2008年度に増強)

ピーク性能 2Tflops (単精度)

プロセッサチップ 4096 個。

プロセッサボード当り 4 チップ (PCI-X/Express)

ホストは 512 ノードの PC クラスタ



純粋 GRAPE に比べたらどうか？

GRAPE-DR と同じ技術、同じチップサイズで
GRAPE-6 後継を作ると:

- トランジスタ数 20 倍
- クロック 5 倍
- 性能は 3 Tflops

GRAPE-DR の GRAPE-6 のエミュレーション:

性能差は 6-8 倍程度

まあ、しょうがないかな、、、 くらい

- 汎用マイクロプロセッサよりはるかによい
- 色々使えるので面白い

どんなことができそうか

- 分子動力学その他あらゆる粒子法
- 量子化学計算 (2電子積分とか)
- 密行列計算 (LINPACK、固有値、対角化等)
- その他直交変換 (FFT は性能低い)
- 差分法でもメッシュ当りの計算量が大ければ結構
なんとか

割合色々できそう。

HPC の方向

- 70 年代: ベクトル
- 80 年代: ベクトル並列、その他並列
- 90 年代: 価格性能比ではPC クラスタ

スーパーコンピューターの意味

年	マシン/速度/価格	マシン/速度/価格	比
1975	Cray-1 100Mflops/10M\$	PDP-11/70 10kflops?/50K\$?	50:1
1985	Cray XMP 1Gflops/10M\$	PC-AT 30kflops?/5K\$	17:1
1995	VPP-500 100Gflops/30M\$?	Dec Alpha 300Mflops/30K\$	1:3
2005	SX-8 10TF/50M\$?	Intel P4(D) 12 Gflops/1K\$	1:60

昔はスーパーコンピューターとはミニコンより価格性能比が良かった。今は逆。

何故スパコンは高くなったか？

答えは簡単：並列計算機になったから

並列である → 並列にするために余計な回路がある
→ 高くなる

なるべく高くならないようにしないとお金をかけた
わりに見返りが少ない

GRAPE-DR は一つの方向：並列にするための回
路を極限まで減らす

まとめ

- GRAPE-DR では、従来の GRAPE と違ってプログラム可能にした
- その分数倍価格性能比が落ちるが、応用範囲ははるかに広がる
- 今後の HPC の、少なくとも1つの方向ではあるのではないか？

梅村さんからの宿題

- 次の10年を考えた各分野での発展性
- 基礎分野と応用分野の関係をどのように考えていくか
- 世界情勢と日本の位置

http://jun.artcompsci.org/articles/future_sc/face.html

次の10年を考えた各分野での発展性

- x86系マイクロプロセッサ: 性能発展のブレーキとなっていた、プロセッサコアの複雑化の方向が断念された: この先5年程度は順調に性能が上がると思われる。
- つまり: それ以外のアーキテクチャはとても大変
- 汎用ネットワークの性能向上は十分ではないが、他のものが勝てるほどではない

計算科学としては、、、このタイプの速い計算機に上手くなる分野は発展、そうでないところは停滞。

基礎分野と応用分野の関係をどのように考えていくか

計算科学の分野において「基礎科学」に国家予算を投入してきたことの意義は？

もっとも重要な例: QCD 用並列計算機

70年代、80年代の分散メモリで単純なネットワークをもった並列計算機の開発の主要な推進力 = QCD 計算

初めから「QCD 計算以外の広い応用」とか考えてたら作れない。

つまり、過去の歴史は明らかに、基礎科学は独自の方向を目指すべき、ということを示している。

現実の問題: 開発コストの高騰 (加速器、望遠鏡の状況に類似)
開発コスト: 相当程度まで技術的な問題「検証、レイアウトの人件費」。技術的に解決すべき

世界情勢と日本の位置

世界といってもプロセッサから作ろうというのは多分アメリカ、中国だけ。

中国は基本的にはアメリカの後追い

アメリカは、

- 物量で多様な方向を追求
- 大量生産を前提とするマイクロプロセッサ開発

物量で対抗するのは消耗戦より「上手く」やる必要がある。