

# 今後の「専用計算機」

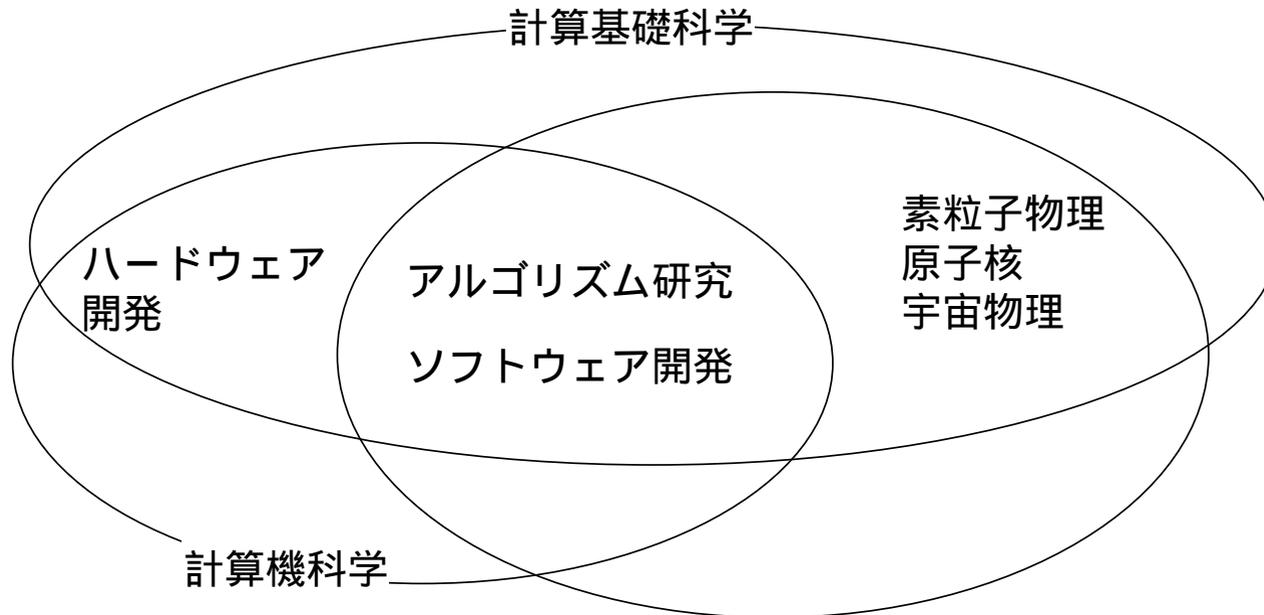
国立天文台  
理論研究部/天文シミュレーションプロジェクト (CfCA)  
牧野淳一郎



計算基礎科学シンポジウム「シミュレーションでせまる宇宙の成り立ち～素粒子・原子核・宇宙物理の展望」 2008/6/16

# まとめ

- 「普通のやり方」より 1-2桁価格性能比が良い計算機を作るのは、それほど難しいわけではない。
- それでサイエンスとして意味があるところまでもっていくには、計算機を作る、ソフトウェアを書く、計算して論文にするところまでが一体になった研究グループである必要あり



- 「計算基礎科学」はそれが可能

# 今日の話の構成

- 計算機アーキテクチャの進化
- QCD 専用計算機と GRAPE
- GRAPE-DR
- 今後の方向

# 計算機アーキテクチャの進化

## 歴史

- 1960年代 単に、高くて速い計算機
  - － 代表例: CDC 6(7)600 (Cray 設計)
- 1970年代 ベクトル プロセッサ
  - － 代表例: Cray-1, CDC-Star
- 1980年代 いろいろあった
  - － 日本のベクトル
  - － クレイの並列
  - － いろんな並列計算機



From Computer Desktop Encyclopedia  
Reproduced with permission.  
© 1996 Cray Research, Inc.



# 歴史のつづき

- 1990年代
  - ベクトル 衰退
  - 並列計算機 会社倒産
  - PC クラスタ 生き残る

何故そうなったか？

- ベクトルは高くなった
- 並列計算機 やっぱり高くなった

何に比べて？

- PC クラスタ



# どれくらい高いか？

1975:	Cray-1	100MF	10M\$	Cray	<b>50倍お得</b>
	PDP-11/70	10kF?	50K\$?		
1985:	Cray XMP	1GF	10M\$	XMP	<b>20倍お得</b>
	PC-AT	30kF?	5K\$		
1995:	VPP-500	100GF	30M\$?	VPP	<b>3倍損</b>
	Dec Alpha	300MF	30K\$		
2005:	SX-8	10TF	50M\$?	SX-8	<b>60倍損</b>
	Intel PD	12 GF	1K\$		

**30年間で 3000倍スパコンは割高になった**

何故か？

# ベクトルプロセッサとマイクロプロセッサ

- Cray-1、その後の普通のベクトルプロセッサ: 主記憶のバンド幅と演算速度が大体つりあう。つまり、メモリが 100GB/s なら 数十 Gflops。
- マイクロプロセッサ: そういうことはあまり考えないで演算速度あげた。10GB/s で 50Gflops とか。
- それでも演算器は少ない
- システムの価格は演算速度ではなくチップ間転送バンド幅で決まっている。

# マイクロプロセッサが HPC の将来か？

x86 マイクロプロセッサの進歩はスーパーコンピュータの進歩を 10-20 年遅れで追いかけている

フルデコード乗算器	CDC 7600 (1971)	i80860(1989)
密結合並列	Cray XMP (1982)	Pen D (2005)
分散メモリ並列	数値風洞 (1993)	???
衰退	1995-	2015-?

# QCD 専用計算機と GRAPE

## マイクロプロセッサベースの QCD 専用計算機

- マイクロプロセッサが良い解になる前後にスタート (1980 年頃)
- 分散メモリ並列計算機の起源
- 今頃マイクロプロセッサの話？

## カスタムパイプラインの GRAPE

- たまたまカスタム LSI がよかった時にスタート (1990 年頃)
- GRAPE-6、MDGRAPE-3 まで、価格性能比では同時期最高

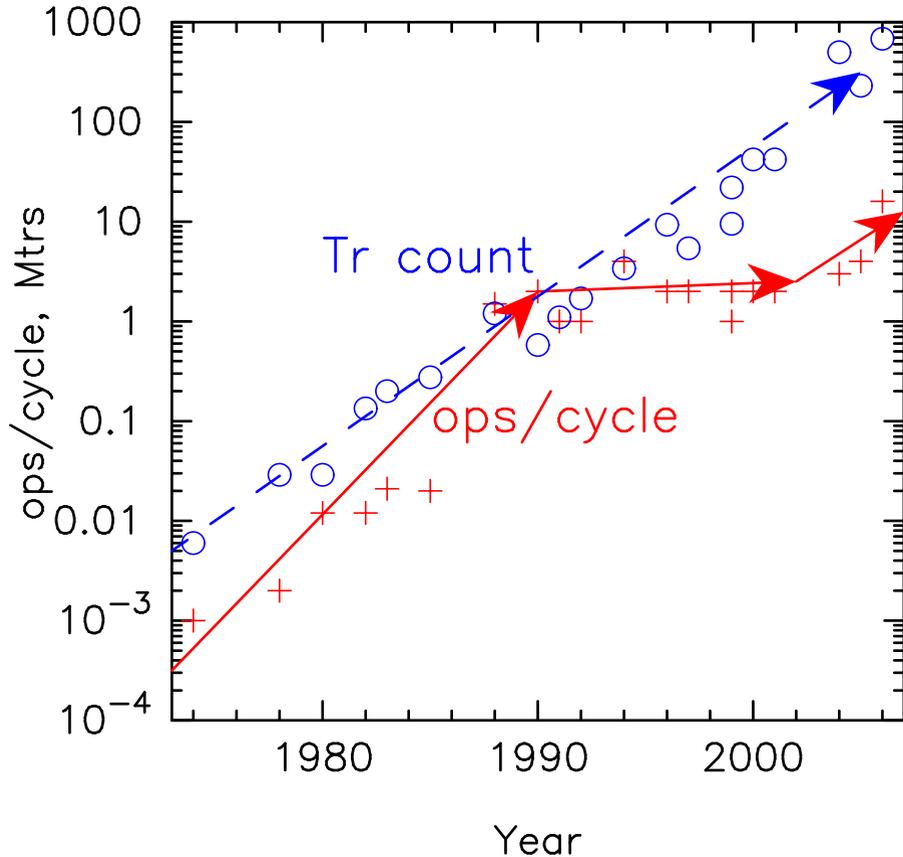
現在の共有の問題: LSI 開発コスト

# GRAPE-DR

- カスタム LSI による専用パイプライン:開発コストが高騰
  - 応用範囲を広げる必要あり
- 汎用マイクロプロセッサは効率が低下中
  - 専用パイプラインでなくても効率上げられる
- 2008年度で、 1Pflops/15億円

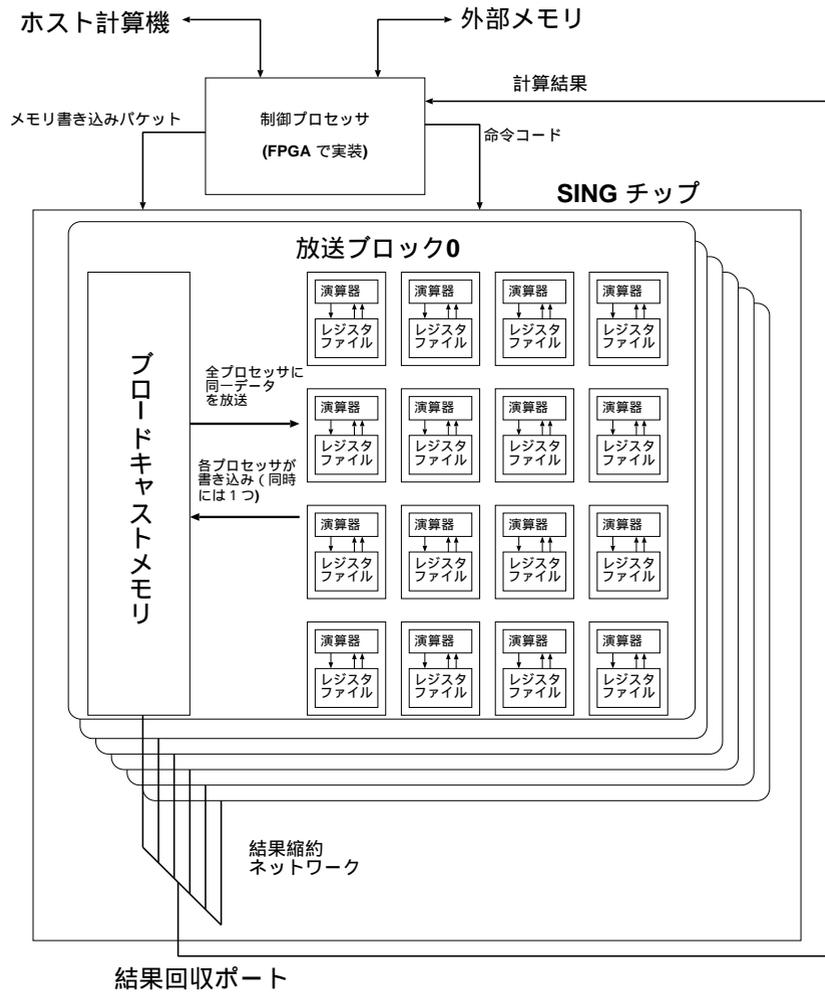
# トランジスタは沢山ある

## マイクロプロセッサの「進歩」



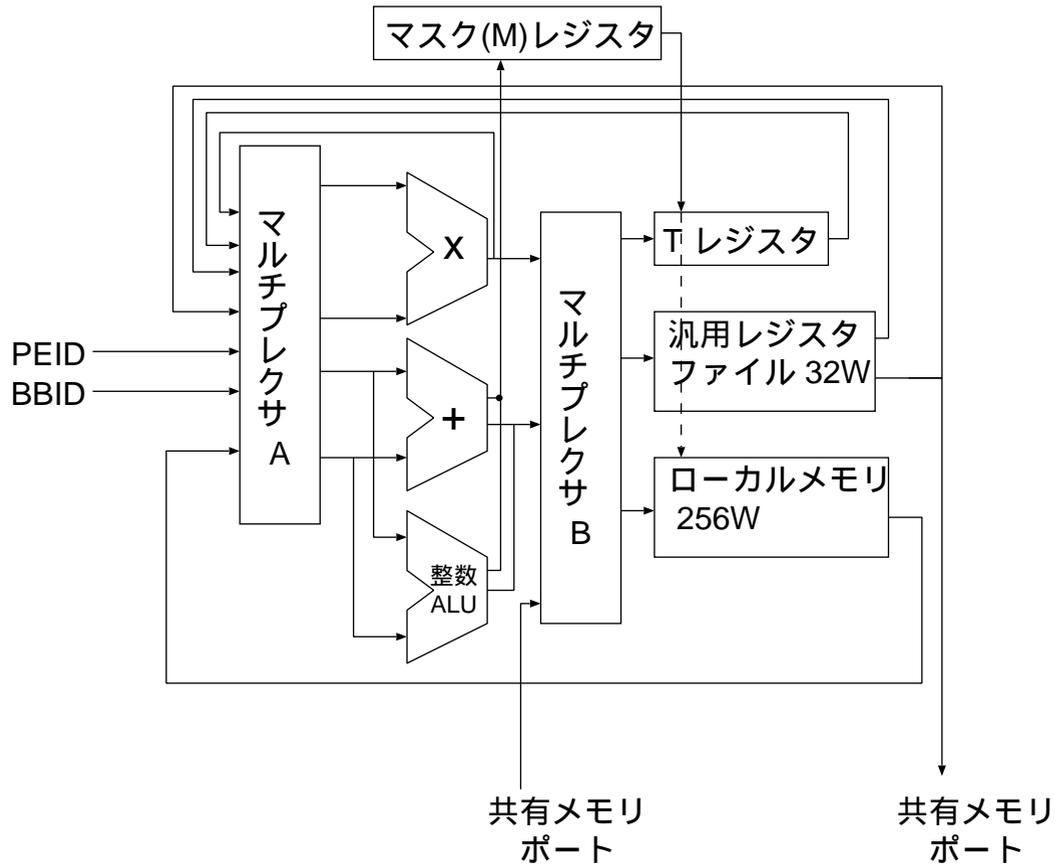
- トランジスタ: 15年で1000倍
- 演算器の数: 同じ期間に8倍
- 100倍分がどこかに消えた
- 最も良かった時でもチップ上の演算器の割合は5%くらい

# GRAPE-DR の構成



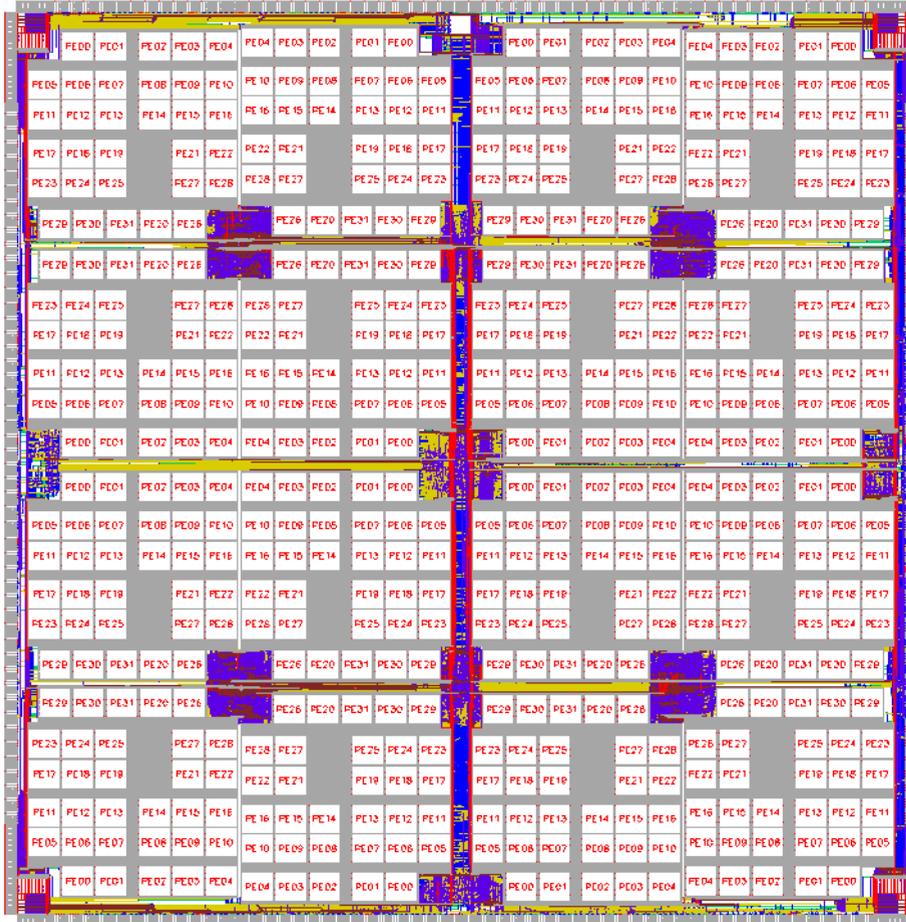
- 非常に多数のプロセッサエレメント (PE) を 1 チップに集積
- PE = 演算器 + レジスタファイル (メモリをもたない)
- チップ内に小規模な共有メモリ (PE にデータをブロードキャスト)。これを共有する PE をブロードキャストブロック (BB) と呼ぶ。
- 制御プロセッサ、外部メモリへのインターフェースを持つ

# PE の構造



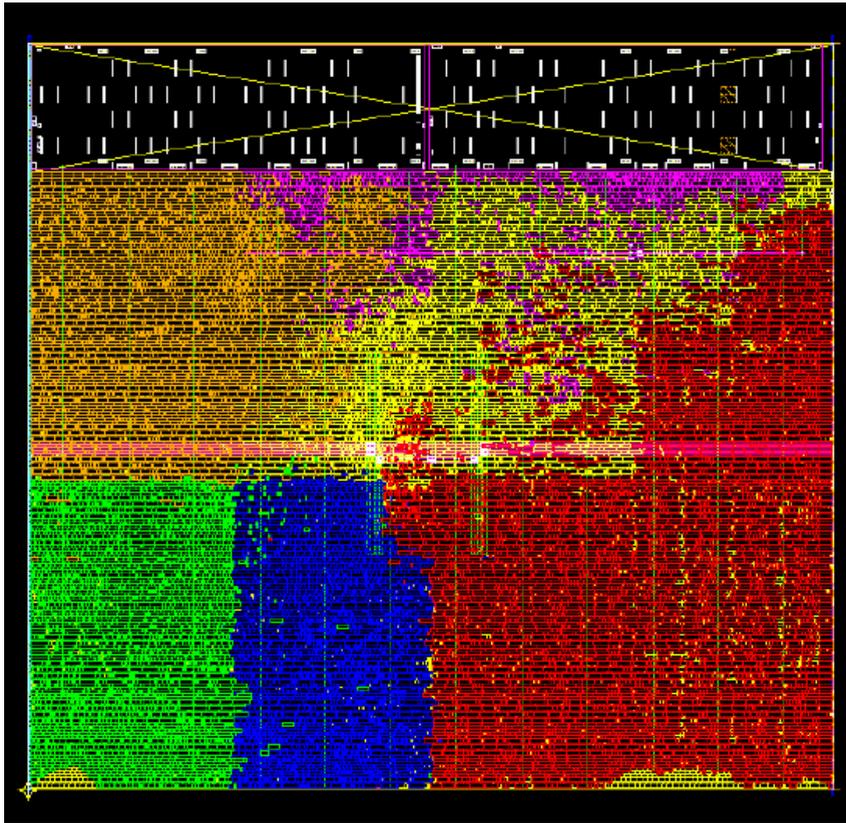
- 浮動小数点演算器
- 整数演算器
- レジスタ
- メモリ 256 語  
(K とか M では  
ない。)

# レイアウト



- 32PE(要素プロセッサ)のブロックが16個
- 空いているところは共有メモリ
- チップサイズは18mm 角

# PE レイアウト



0.7mm by 0.7mm

Black: Local Memory

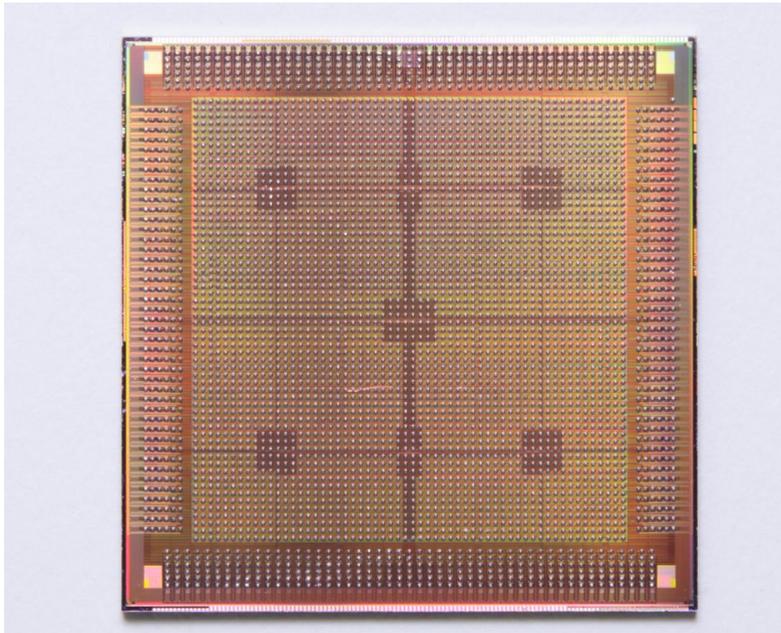
Red: Reg. File

Orange: FMUL

Green: FADD

Blue: IALU

# プロセッサチップとボード



- PCI-Express カード (16 レーン、通信速度 2GB/s)
- 4 GRAPE-DR チップ
- 理論ピークスピード 1TP(DP), 2TF(SP)

# 今後の方向

- 多演算プロセッサは今後 10-15 年間の方向
  - GRAPE-DR
  - GPGPU
  - マルチコアプロセッサ
- 計算基礎科学の使命の1つは、他の分野より早く多演算プロセッサを有効に使う方法を示すこと

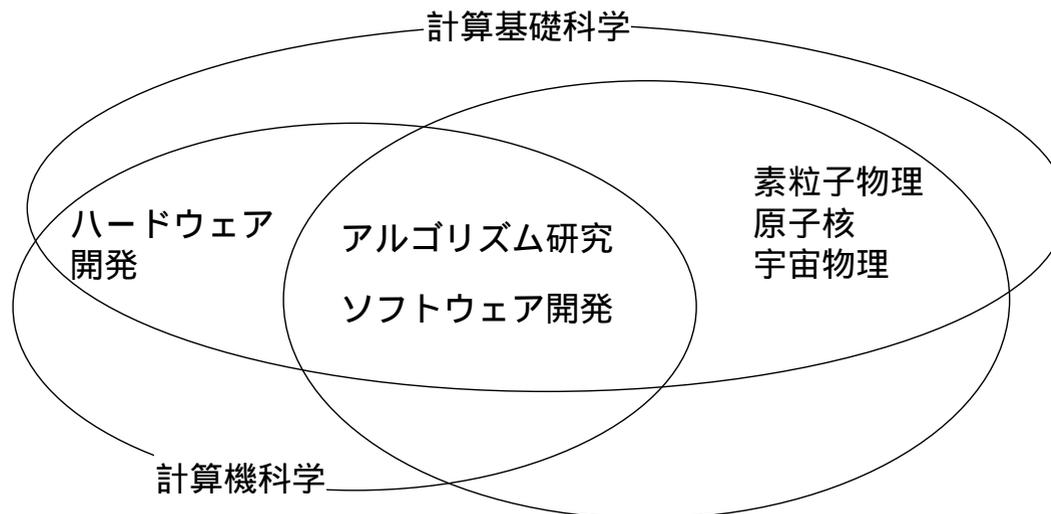
# 計算基礎科学の使命とは？

- 基礎科学の中での計算科学

- 自然法則の理解、法則による自然現象の理解

- 計算科学の中での基礎科学

- 比較的単純な自然現象を対象にすることで、計算科学の可能性、方向性を示す
- 計算機の使いかたから作り方まで



# おまけ: GRAPE-DR で QCD 計算はできるか

- 今のままの設計では効率低い
- 不足しているもの: チップ間ネットワークバンド幅
- BG/L: 5Gflops で 10Gbps
- GRAPE-DR: 250Gflops。今の構成だとノード間はチップ当たり 1Gbps くらい。200Gbps くらいは欲しい
  - 1GHz LVDS, 200 本。
  - 技術的には難しくない
  - コストもそんなににかかるわけではない