

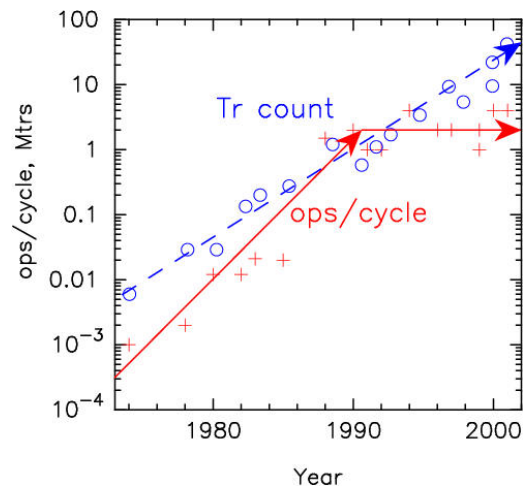
## VPM (Vector Processor Matrix) 次世代HPCアーキテクチャ

戎崎俊一、泰地真弘人(理化学研究所)  
牧野淳一郎(東大理学系研究科)

## 科学技術用計算機(既存1)

- ベクトル計算機
  - 計算/通信~2
  - 通信幅に膨大な費用
    - 地球シミュレータ 40Tflops/400億円
- PCクラスタ
  - 計算/通信~20
  - まだ割高
    - 理研クラスタ 12.5Tflops/38億円
- SMPクラスタ
  - 計算/通信~10
  - 欠点がなくなんでもそつなくなす
- GRAPE型専用計算機
  - 計算/通信~300
  - 割安
    - MDM 78Tflops/5億円
    - GRAPE-6 64Tflops/5億円
  - 融通が利かない 実パイプライン

## CPUの性能変化



## 科学技術用計算機(続き)

- 計算通信比が価格性能比を決める
- 実効/理論性能比はベクトルとGRAPEで高く、PC クラスタで低い
- ベクトルで実効性能が高いのは、計算通信比と蓄積されたノウハウのため
- GRAPEで高いのは応用に適合したメモリアーキテクチャのノウハウのため
- 単なる GRID では性能がでない

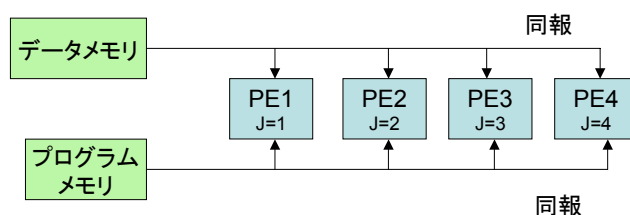
## VPM (Vector Processor Matrix)

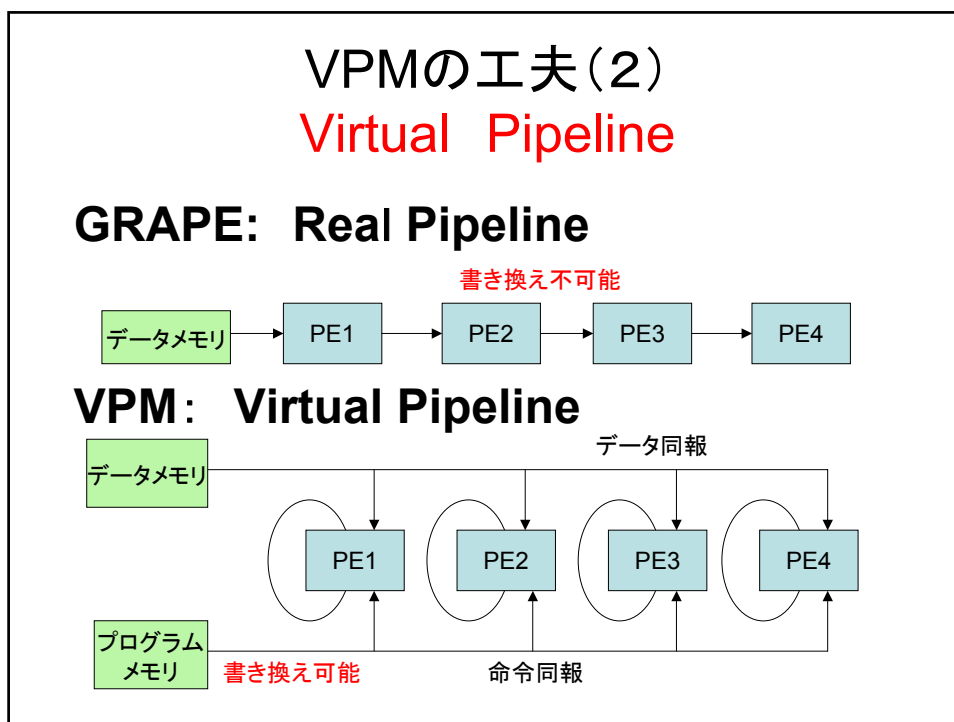
- GRAPEの価格性能比を維持、柔軟性を確保
- 計算／データ比が高いHPCアプリケーションに特化
  - 重力多体問題、分子動力学、境界要素法
  - 第一原理電子状態計算(二電子積分、行列対角化)
  - ルジャンドル変換など
  - 密行列ソルバ

## VPMにおける工夫(1)

### 同報並列計算

- HPCアプリの多くは、計算/データ $\gg$ 1
  - 重力・静電力計算: 行列対角化: 密行列直接解法
  - 二電子積分: 境界要素法
  - 計算/通信 $\gg$ 1でもよい
    - 一つのデータを何度も使いまわす

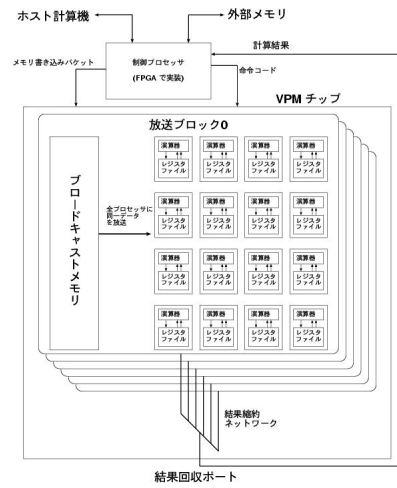




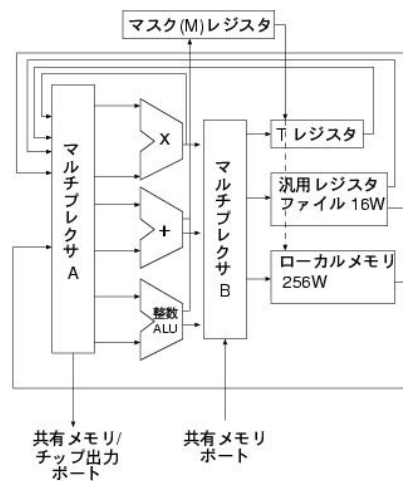
## GRAPE-DR VPM型の最初のプロトタイプ

- 振興調整費
  - 2004年度開始 5年 × 3億円
  - 2.0~2.8Pflops(単精度) 2008年度
  - GRAPE-DRチップ
    - 1.4Gflops/0.35Gflops(単/倍精度)
    - メモリバンド幅5.6GB/s 計算/通信=200
  - システム
    - 1024 PE/チップ 2048チップ
    - 256ノードのPCクラスタ
- アプリケーション
  - 重力多体、分子動力学、境界要素法

# VPMシステム



# PEの構成



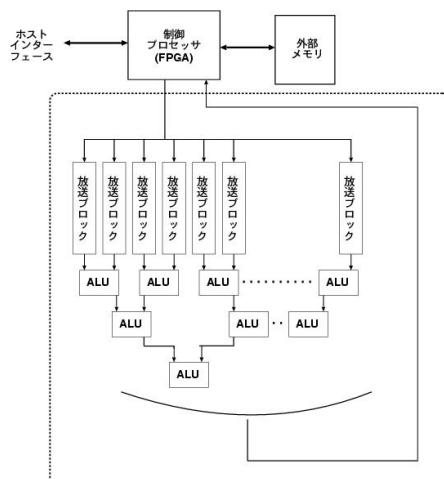
## GRAPE-DRチップ

- 90nmテクノロジー
- 2億トランジスタ
- 1.4Tflops／350Gflops(単精度／倍精度)
- 700MHz動作
- 1024個のPE内蔵
- 20W程度の消費電力

## GRAPE-DRチップ機能

- ブロードキャストユニット32個(PE 1024個)
- 外部メモリからブロードキャストメモリへの転送バンド幅 5.6GB/s
- ブロードキャストメモリから(リダクションツリーを通過しての)
- 外部メモリ/ホスト計算機へのデータ転送速度 1.4GB/s

## VPMチップ構造



## プロセッサエレメントの仕様

- 浮動小数点単精度乗算:1演算／1サイクル
- 浮動小数点倍精度加減算:1演算／サイクル
- 浮動小数点倍精度乗算:1演算／4サイクル
  
- 内積演算:乗算・加算並列実行可能
- ローカルメモリ:倍精度 256 語

## ブロードキャストユニットの仕様

- プロセッサエレメント 32 個にデータを供給
- ブロードキャストメモリ 1K語(デュアルポート)
- ブロードキャストメモリからPEへのデータ放送機能
- 個別PEによる読み書きを可能とする

## 行列積

1. 行列A、Bの要素を外部メモリに格納
2. Bの部分小行列を各PEのレジスタに格納
3. Aの部分列を全PEに順に放送、PEでは積和命令実行( $c_{ij}=a_{i1} * b_{1j} + a_{i2} * b_{2j} \dots + a_{ik} * b_{kj}$ )
4. グループにまたがったりダクションを行いつつ、結果cを回収



## 重力計算

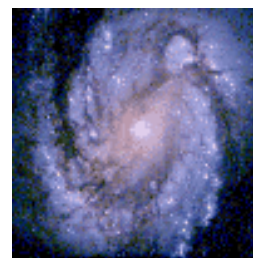
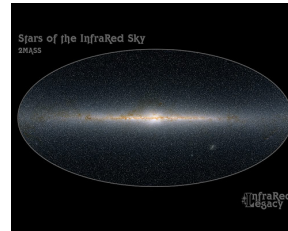
- 全粒子の位置座標を外部メモリに格納
- 力を計算する場所の位置座標群を各PEのレジスタに格納
- 力を与える位置情報を全PEに放送、PEでその粒子からの力の寄与を計算して積算
- グループにまたがったリダクションを行いながら、積算結果を回収

## GRAPE-8 実用マシン

- 目標
  - 10Pflops    2010年度    50億円
- GRAPE-8チップ
  - 1. 0Tflops/チップ (倍精度)
  - ~10000チップ
  - 計算/通信            50~100 (開発要素)
  - 1024ノードのPCクラスタ
- アプリケーション
  - 重力多体、分子動力学、境界要素法
  - 第一原理計算 (2電子積分、密行列の対角化)

## 銀河を作る

- 銀河の大きさ
  - 3万光年
  - 1千億太陽質量
- 分解能
  - 空間~30光年
  - 質量~太陽質量
- 粒子数
  - 空間分解能の要請  
 $1000^3 = \sim 10$ 億個
  - 質量分解能の要請~  
1000億個



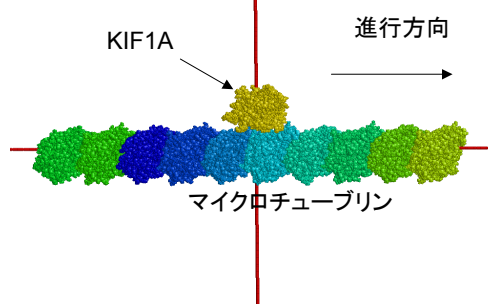
## 惑星を作る

- 惑星形成
  - 原始太陽系星雲の中で微惑星の形成
  - 微惑星が公転しながら、重力散乱と合体・成長を繰り返す
  - 数十億年で惑星になる
- 1000万個の粒子でこれをシミュレーション
  - 10Pflops



## マックスウェルの悪魔

- 分子モーター
  - エネルギー効率60%超
  - もっとも簡単なものはブラウン・モーター
- なぜ生体分子モーターはこんなに効率が良いのか？



## 心を作る

- 脳は1000億個の神経細胞で構成
- 神経細胞=1万入力、1出力、1kHz動作のゲート=2kflopsのプロセッサ
- 全体としては2000Pflopsの計算機でエミュレート可能
- 20Pflopsマシンはその100分の一の性能を持つ。
- そろそろ心を作れてもいいかもしれない。

## これからの課題

- FFT
  - ポアソン方程式の解法
    - 他的高速アルゴリズムの開発
  - データ解析
    - 低精度→FPGAベースマシン
- 地球環境シミュレーション
  - 全球コード: 大部分はルジャンドル変換の計算
    - VPMIに向く

## FPGA

### (Field Programmable gate Array)

- 低精度のHPC計算には威力を発揮
  - 銀河シミュレーション(低精度型GRAPE)
  - 類似度検索(ゲノムや文章の解析など)
  - 画像・データ処理(FFT、位相回復など)
- Bioler-3の開発(千葉大と共同開発)
- DAP-DNA (IP-Flex社)のソフトウェア開発

## まとめ

- VPM:高(計算/通信)でも高い実行性能
  - フォン・ノイマン・ボトルネックの回避
- 多くの科学技術計算に適用可能
  - 重力多体、分子動力学
  - 行列演算(LINPACKなど)
  - 第一原理電子状態計算など
    - 二電子積分、密行列対角化
- GRAPE-DR:プロトタイプ
  - 2008年に2Pflops 15億円
- GRAPE-8:実用マシン(応用範囲を拡大)
  - 2010年に10Pflops 50億円
- ベクトルマシン、PCクラスタと補完的