

分子動力学専用計算機 MDGRAPE-3の開発と 生命科学・工学への応用

泰地 真弘人

理化学研究所 ゲノム科学総合研究センター
システム情報生物学研究グループ
高速分子シミュレーション研究チーム
taiji@gsc.riken.jp

大規模シミュレーションによる 分子生物学

- 定量的な実験データの蓄積
 - タンパク質の構造
 - 転写ネットワーク、プロテオーム、メタボローム
 - イメージング
 - 表現型情報
 - 個人の配列情報
- シミュレーションの適用範囲の拡大
 - 分子動力学
 - 量子化学計算
 - ネットワークのモデリング・シミュレーション
 - 遺伝統計解析、表現形解析 - biometrics

MDGRAPE-3(Protein Explorer)計画

- 1Petaflopsの分子動力学専用計算機
- 世界初のPetaflops計算機を目指す
- タンパク3000プロジェクトの一環
- 2002年度開発開始、2006年度完成
- 専用化により、100倍以上優れたコスト・性能比を実現
- 省スペース・省電力
- 現在くみ上げ中

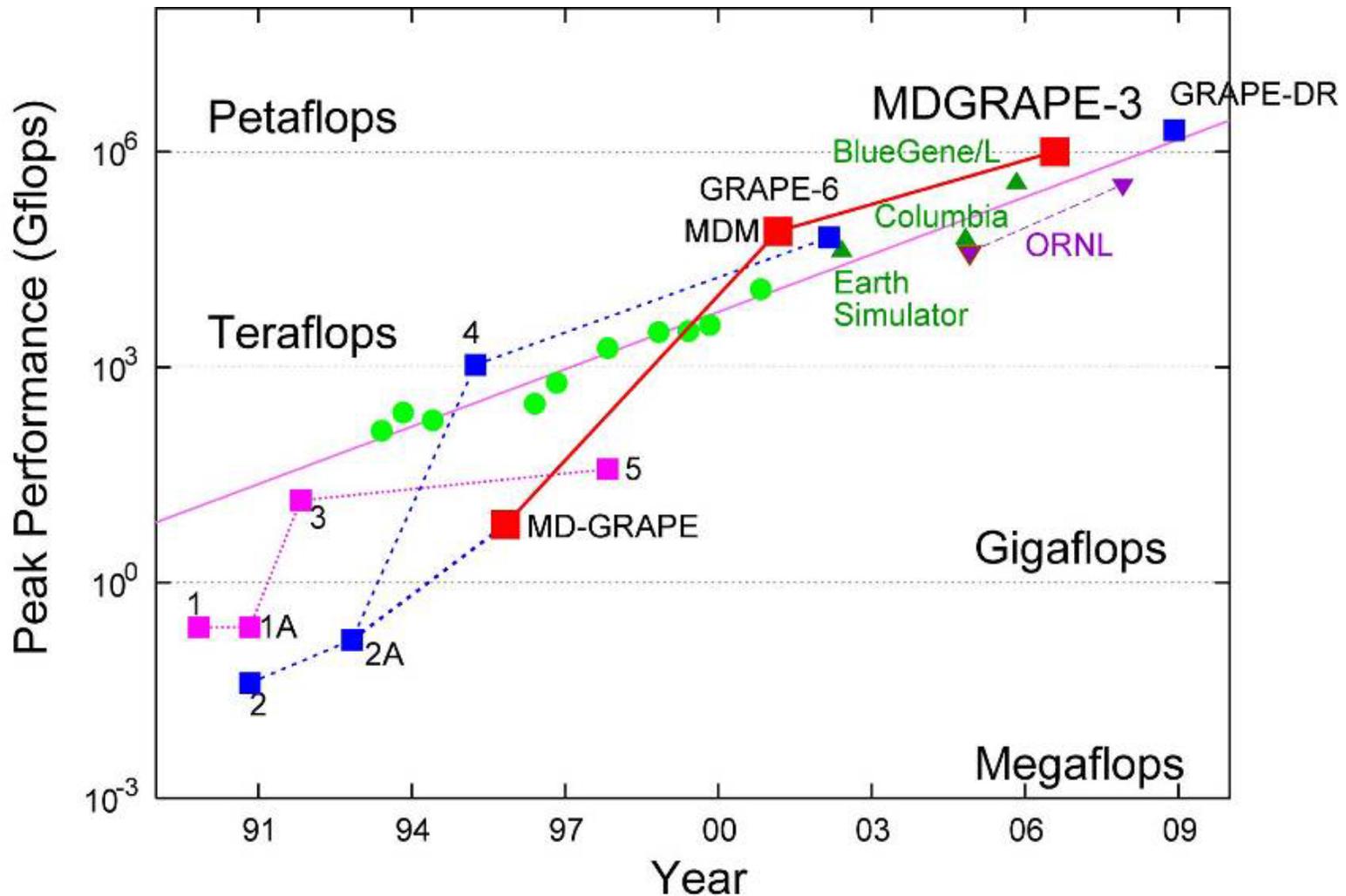


MDGRAPE-3 チップ
165Gflops



MDGRAPE-3テスト基板

History of GRAPE computers



なぜ専用計算機が重要か

- 汎用計算機の問題点
 - 半導体集積度の向上を活かしきれない
 - 発熱の問題
 - これらの問題は今後ますます深刻になる
 - 専用計算機では
 - 集積度の問題 **根本的に解決可能**
 - 発熱の問題 かなり緩和可能
- 価格性能比、電力性能比が100倍以上向上**

Highly-Parallel Operations in a LSI for MD simulations

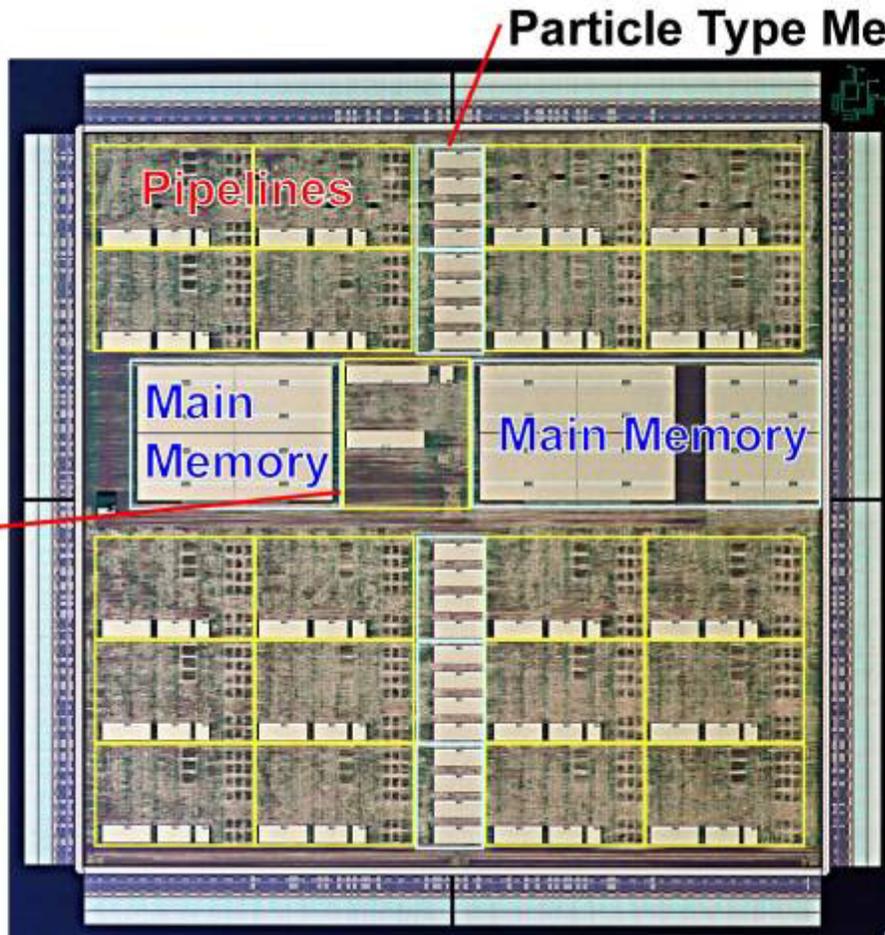
- For special-purpose computers
 - Broadcast Memory Architecture
 - Efficient : 660 operations/cycle/chip in MDGRAPE-3 chip
 - possible to increase according to Moore's law
- In case of MD:

MDGRAPE	600 nm	1 pipeline	1Gflops
MDGRAPE-2	250 nm	4 pipelines	16Gflops
MDGRAPE-3	130 nm	20 pipelines	165Gflops

Power Efficiency of Special-Purpose Computers

- 同等技術での比較
- General-Purpose Processors
 - Pentium 4 (0.13 μm , 3GHz, FSB800) ... 82W
 - 14W/Gflops
- Molecular Dynamics Processors
 - MDGRAPE-2 chip (0.25 μm) ... 1W/Gflops
 - MDGRAPE-3 chip (0.13 μm) ... 0.1W/Gflops

MDGRAPE-3 chip

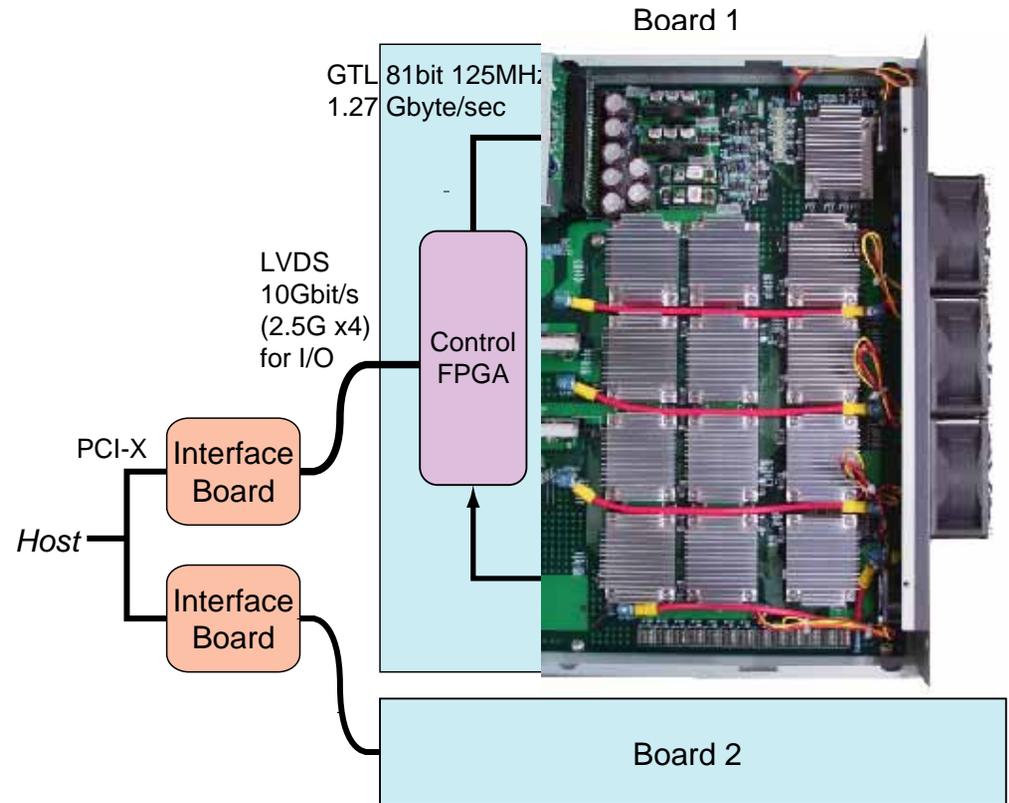


130 nm
 Vcore=+1.2V
 250MHz (worst)
 15.7 mm X 15.7 mm
 6.1 M random gates
 + 9 Mbit memory
 17W at 300 MHz
 1444 pin FCBGA

M. Taiji, Proc. Hot Chips 14, in CD-ROM (2004)

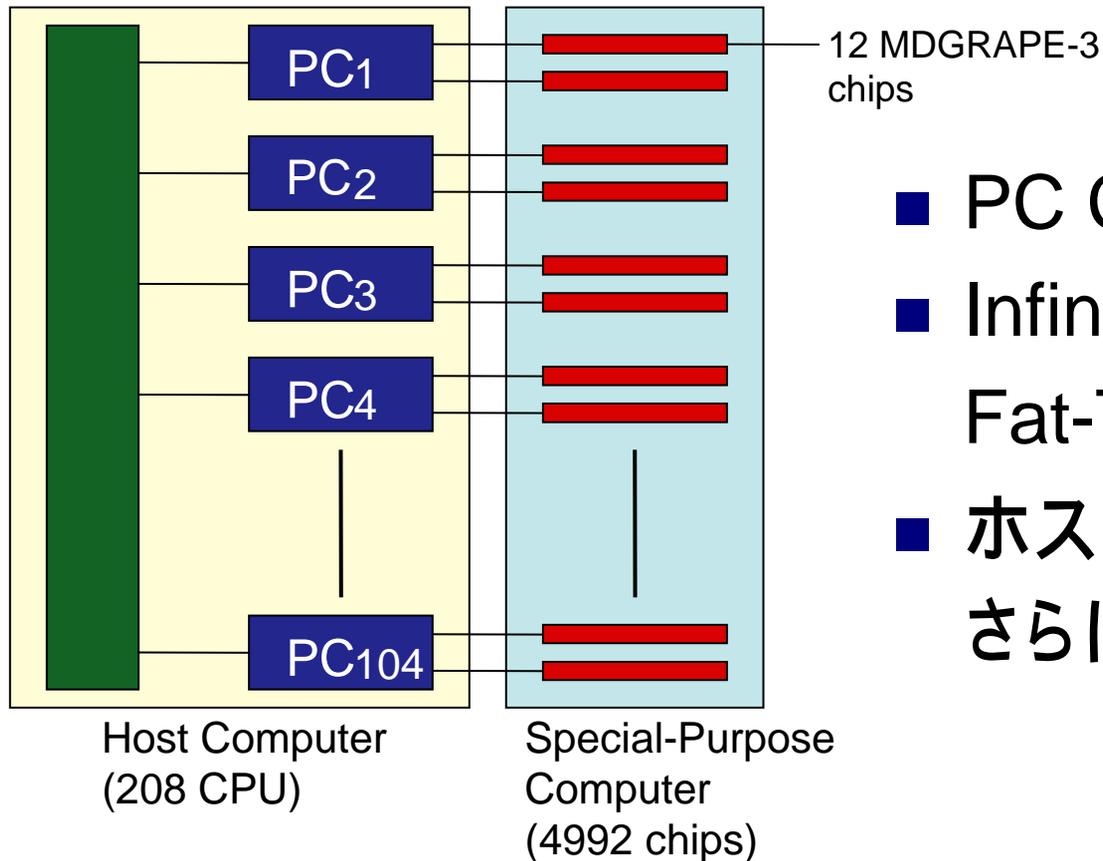
MDGRAPE-3システム

- 12チップ/ボード
- 2ボード/2Uサブラック
= 4 Tflops
- PCI-X接続
MDM-1aに比べて倍以上の転送性能
(演算あたり)
- 2チップのPCI-X
ボードも製作



PCI-X board:
330 Gflops

システム構成



- PC Cluster (Intel Xeon)
- Infiniband Switch:
Fat-Tree
- ホスト計算機は今年度
さらに増強予定

MDGRAPE-3 cluster



大規模分子動力学シミュレーションで できること

■ 分子間相互作用の予測

- Protein-Protein interactions
- Protein-Ligand interactions (screening)
- Protein-DNA/RNA interactions
- SNPsによる違い
- 化学修飾による違い

■ タンパク質の分子機構を原子レベルで知る

■ 構造：フォールディング、複合体、凝集体

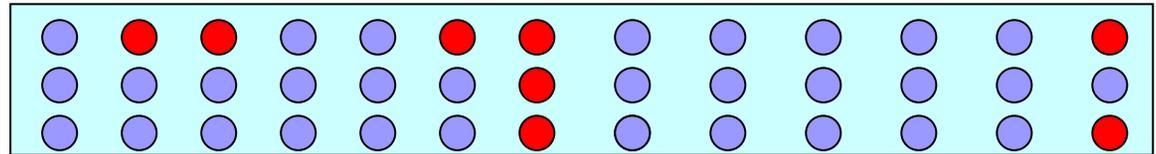
MDGRAPE-3による分子設計

- 生体分子間の結合エネルギーをMMPBSAや Thermodynamic IntegrationなどのMDベースの方法で精度よく見積もる
- 薬剤設計、タンパク質設計、機能推定
- 通常のドッキングプログラムの予測精度は低い

- 薬剤のスクリーニング
ペタフロップスあれば
500 ~ 数千対サンプル/日の能力
- 特に薬剤の最適化に有効
ー合成の前に結合能を評価できる

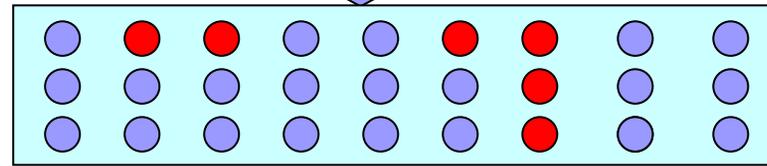
薬剤スクリーニング用 自動化パイプライン

データベース
(1,000,000化合物)



1次スクリーニング:
Docking Program

5000~10000 化合物: 7日



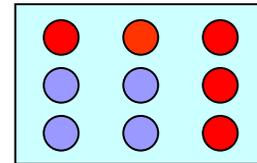
←
de novo

2次スクリーニング:
MDシミュレーション

MDパラメタの準備: 10日

MDシミュレーション: 10日.

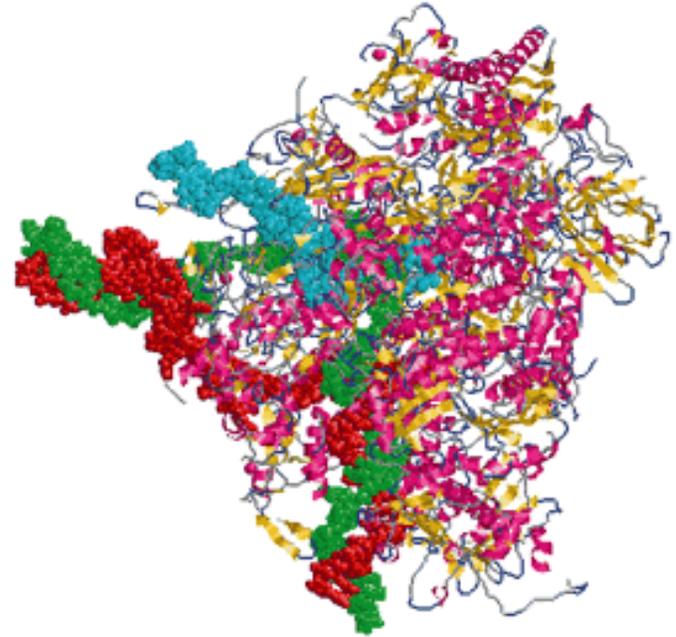
解析: 10日



1,000,000 化合物のスクリーニング
40日<=2ヶ月

大規模分子シミュレーションによる 機能解析

- RNA Polymerase II
~50万原子
- DNAをほどこいた状態の
モデリング
- Targetted MDによる
モーター機構の解析
- Petaflops級: 1M原子、マイクロ秒



タンパク質のフォールディング、 凝集体の形成・構造

- Pflopsの専用計算機
累積でミリ秒の計算が可能
- フォールディングの
動力学・統計力学
- タンパク質凝集体の形成過程・構造の解明

今後の課題 — 京速に向けて

- 実証研究
- 量子系
- 力場の改良
- 時間軸方向に延ばすには？
- ペタフロップスまでスケールするコード開発
- ペタフロップスを前提としたソフトウェア開発

MDGRAPE-4

- 20Pflops @ 2010
- 2Tflops/chip, 65 nm
- 薬剤スクリーニング等への応用では、
20Pflopsぐらいが一つの区切り
- 実験より
低いコスト + 短い期間
を実現