

分子動力学専用計算機 MDGRAPE

Yousuke Ohno ohno@riken.jp

Senior Technical Scientist
Laboratory for Computational Molecular Design
Center for Biosystems Dynamics Research (BDR)
生命機能科学研究センター計算分子設計研究チーム
RIKEN

- GRAPE-1Aのコピー
- DREAM-1/1A (disk base, for fluid dynamics) 大学院杉本研
- GRAPE-3 複製指導
- 観測 (近赤外、可視光広視野) 大学院杉本研・理研戎崎研
- GSE/MACE (大規模行列専用計算機) 理研戎崎研
- MDGRAPE-3 理研泰地研 (横浜)
- cppmd (汎用機用MDコード) 理研泰地研 (神戸)
- MDGRAPE-4/4A 理研泰地研 (神戸・大阪)
- MDGRAPE-5(開発中) 理研泰地研 (神戸・大阪)

■ GRAPE-1A(福重)の複製の作成

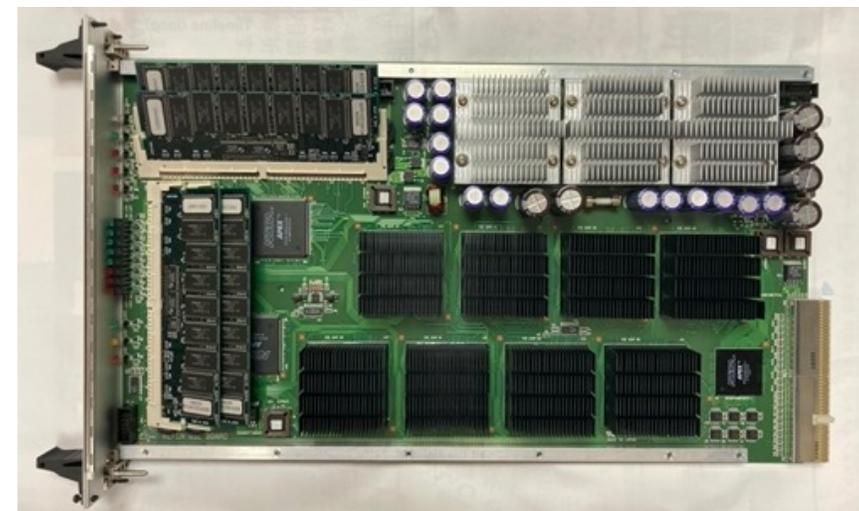
- ▷ 野本研の課題
- ▷ 牧野さんの指導

■ DREAM-1A

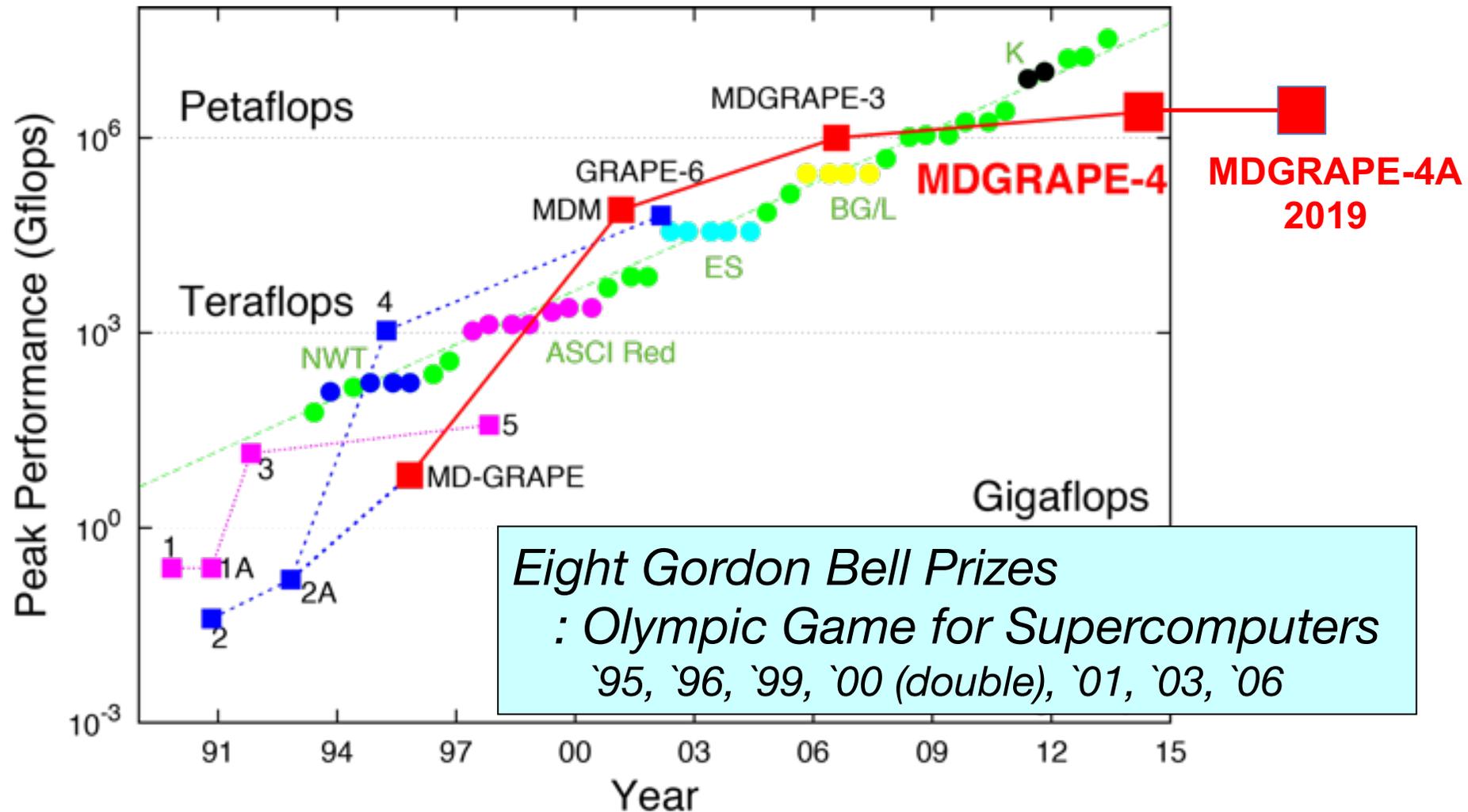
- ▷ Disk REsource Array Machine
- ▷ ハードディスクを主記憶として使う
- ▷ 牧野さんの発案
- ▷ 4並列

■ GSE/MACE

- ▷ 大規模密行列専用計算機
- ▷ 密行列の直交対角化
 - 密度汎関数理論 (量子力学)
- ▷ ボード・システム・ソフトウェアを担当
 - ASIC 泰地 (統計数理研究所 当時)



History of MDGRAPE



■ 分子動力学 Molecular Dynamics

- ▷ 原子（分子）で構成された粒子系：水中タンパク質
- ▷ クーロン力、ファンデルワールスカ
 - 二体中心力なので、重力用パイプラインの拡張で対応できる
- ▷ 共有結合
 - 2-4体の相互作用だが、数が少ないので汎用CPUで

■ GRAPE-2A (1991)

- ▷ 高精度版

■ MD-GRAPE (1995)

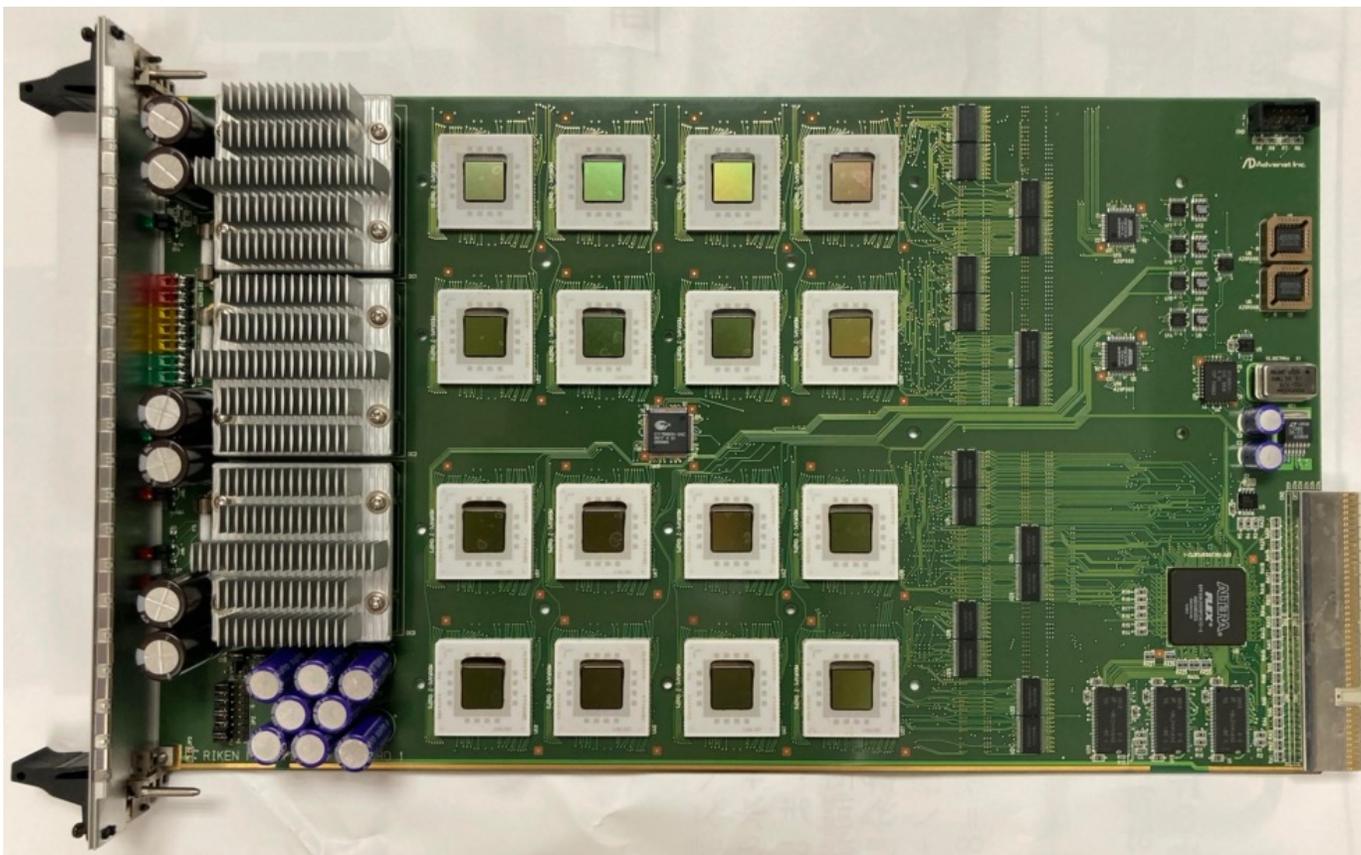
■ WINE

- ▷ Ewald法専用計算機

- 波数空間で周期境界条件下の遠距離ペア計算を高速化するアルゴリズム

■ MDM : MDGRAPE-2/WINE-2 (1999 理研戎崎研究室 成見)

- ▷ クーロンカ・ファンデルワールス力専用ASIC
- ▷ Ewald法専用ASIC



MDGRAPE-2
4 chip

MDGRAPE-2
16 chip

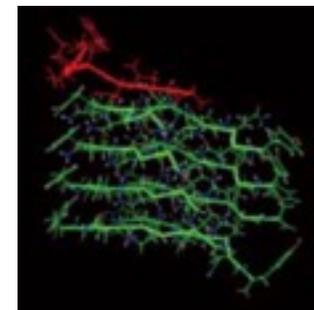
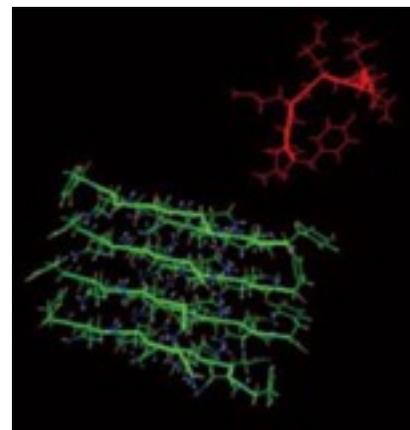
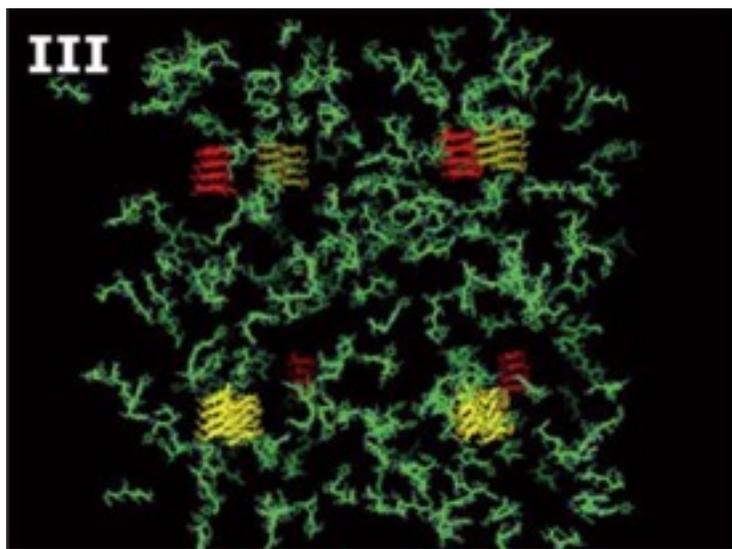
■ 2006 (理研泰地研究室 泰地、成見、大野)

- ▷ 1 PFlops
- ▷ クーロンカ・ファンデルワールスカ専用パイプライン
- ▷ 三角関数対応：フーリエ変換（WINE機能）
- ▷ 20 pipeline/chip
 - 40 particles by virtual parallelization
- ▷ 4800 chip
 - 12 chip/board
 - 400 boards



■ amyloid (プリオンの断片)の凝集

- ▷ Gordon Bell Prize Honorable Mention 2006
- ▷ 1400万原子 ns シミュレーション 55 TFlops



■ フーリエ変換機能のX線構造解析への応用

- ▷ Gordon Bell Prize Finalist 2007

■ アクセラレータ型

- ▷ チップあたりの粒子数が少ないと効率が悪い

■ 大規模並列

- ▷ 4800個のASIC、100台の汎用計算機

- ▷ 計算対象全体の粒子数が少ないと並列化効率も悪い

- 1400万原子 ns シミュレーション 55 TFlops

- ▷ ストロングスケールリングに弱い

- ▷ 大規模タンパク質MDで主流のPMEは効率が悪い

- Particle Mesh Ewald, メッシュに割り当てた電荷分布を使いFFTで遠距離のクーロン相互作用を計算

- FFT は $O(N\log N)$

- 直接計算はカットオフ距離内の近距離ペアだけなので $O(N)$

■ アクセラレータ型

▷ 計算量の多い部分だけ専用回路、あとは汎用計算機

– 開発要素が少ない

▶ 汎用計算機（OSやソフトウェア環境も含めて）は既製品でよい

▷ 効率は問題の性質に依存

重力多体問題の場合

– 重力相互作用の計算量が支配的 $O(N^2)$

▶ データ量は $O(N)$

▶ 積分等 $O(N)$ の部分は加速しなくてもよい

▶ 専用回路と汎用計算機の通信も $O(N)$ で通信時間も問題になりにくい

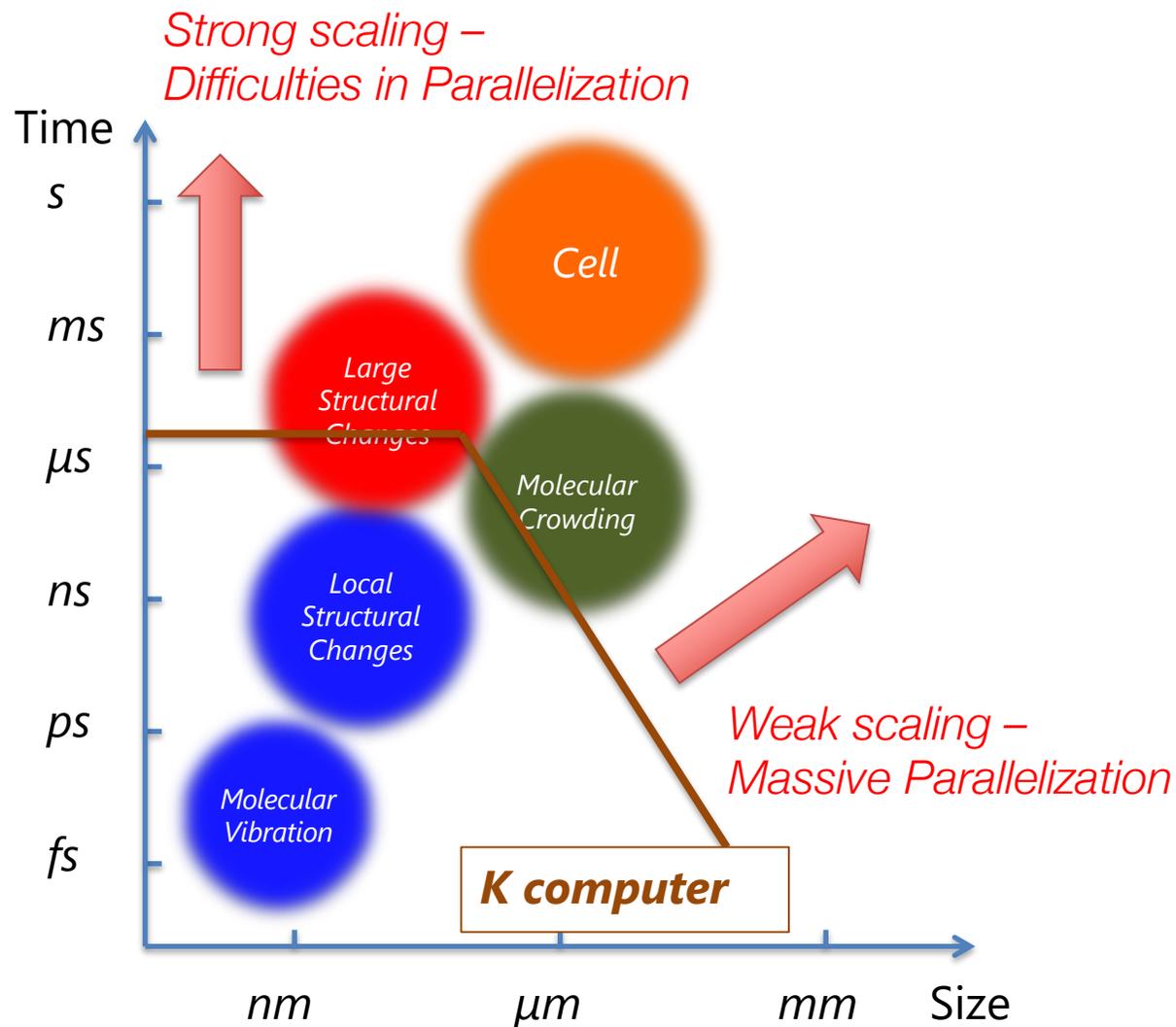
– 粒子数が多いほど効率が良い

▷ 粒子数が少ないと効率が悪い

– 大規模並列化する場合ストロングスケールリングは苦手

– 演算量自体を減らすアルゴリズムとの相性

Scaling Challenges in MD



- $\sim 30,000$ FLOP/particle/step
- System size : $N=10^5$
- 3 GFLOP/step
- 3TFLOPS effective performance
1msec/step = 200nsec/day
Possible
- 3PFLOPS effective performance
1 μ sec/step = 200 μ sec/day???
Difficult, but important

並列効率の限界により汎用計算機ではタンパク質の長時間シミュレーションは困難
アクセラレーターも加速効率に問題

■ SoC(System on Chip)

- ▷ 専用回路だけでなくCPU・メモリやネットワークも一つのLSIに。
1chipで1ノードが完結する。
- ▷ 汎用CPUやネットワークと専用回路がLSI内部で接続されるので専用回路のデータ入出力が高速になる。
 - ノード間（チップ間）通信も独自に最適化・高速化可能
 - OS・ドライバ等のオーバーヘッドを減らすことも可能。
- ▷ 設計・開発要素が多い
 - CPU、ネットワーク（チップ間、チップ内）
 - ソフトウェア
 - 独自アーキテクチャ用に移植
 - ハードウェアの制御
- ▷ ANTON(D. E. Show Research)が先行

■ MDGRAPE-4/4A

▷ 汎用CPU

- 4: Xtensa (Tensilica社), 4A: RISC-Vを参考にASIP (SynopsysのCPU開発環境)で独自設計
- コンパイラも開発環境で生成
- OSなし、MDコードを直接実行

▷ ネットワーク (チップ間)

- 物理層は既製品、プロトコルは独自

▷ メモリ

▷ 専用パイプライン

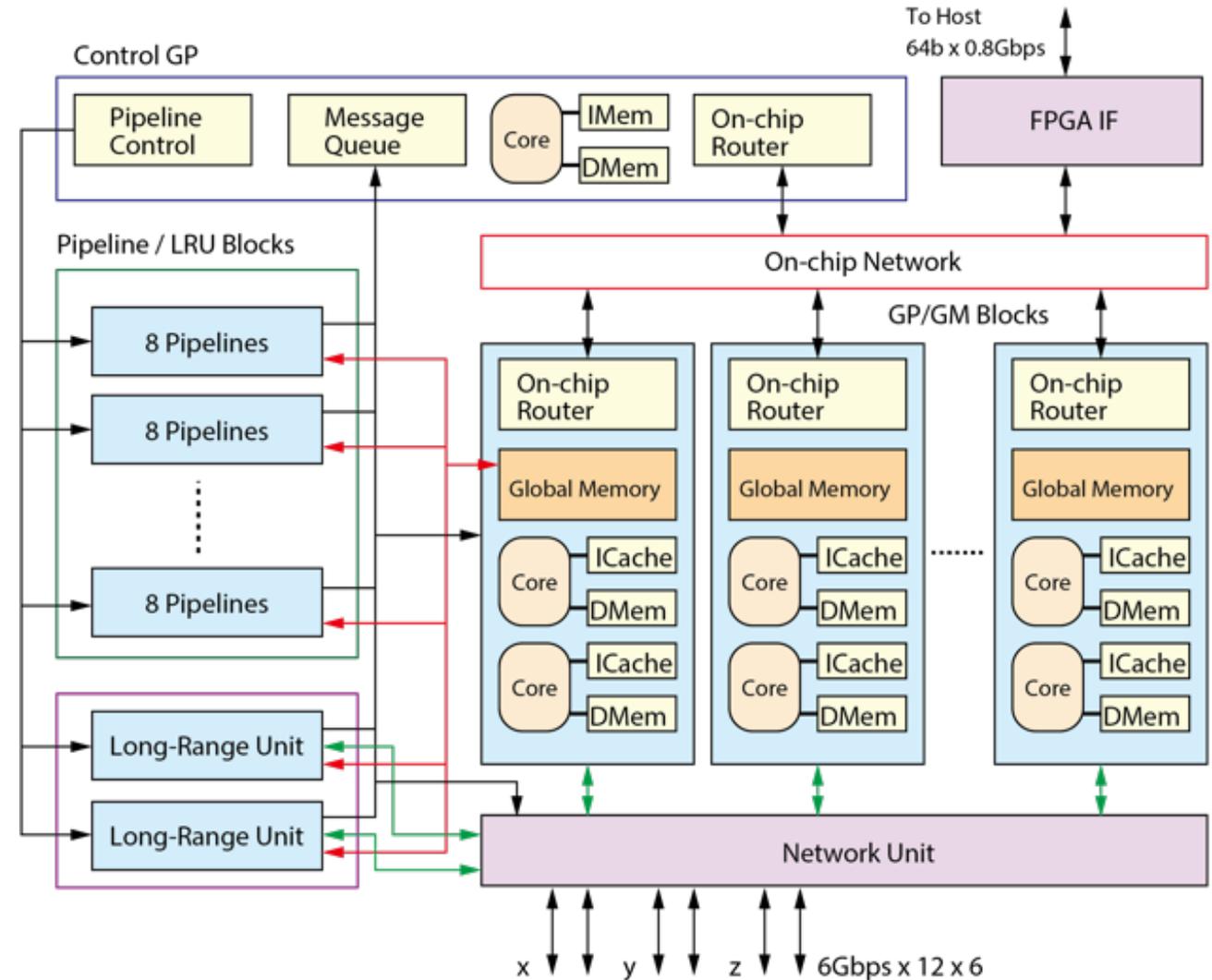
▷ 専用ASIC以外に遠距離相互作用を支援するFPGA

▷ ハードウェア 泰地、大村、大野、森本、Zhang

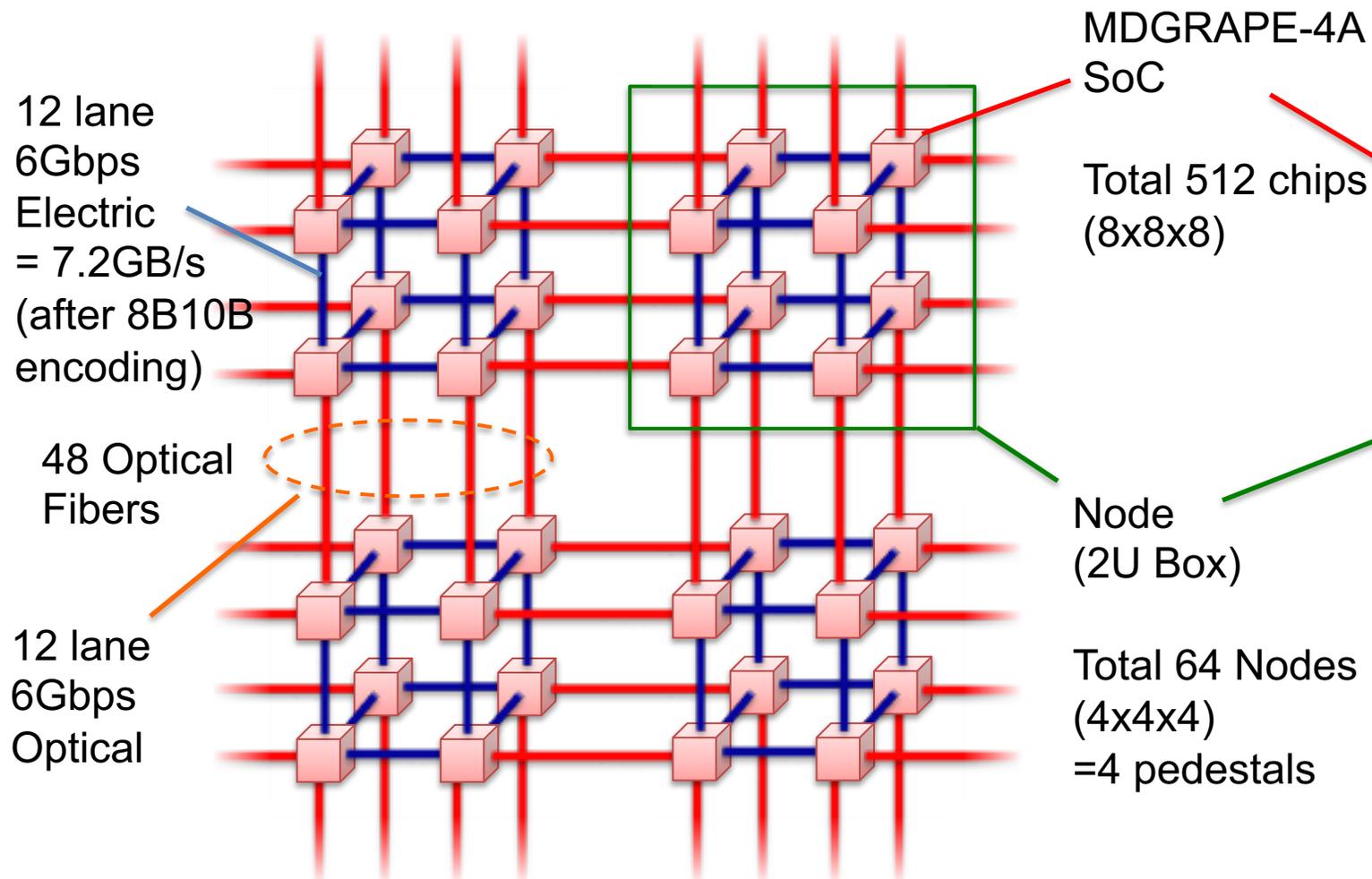
▷ ソフトウェア 泰地、大野、森元、小山、小松

MDGRAPE-4A System-on-Chip

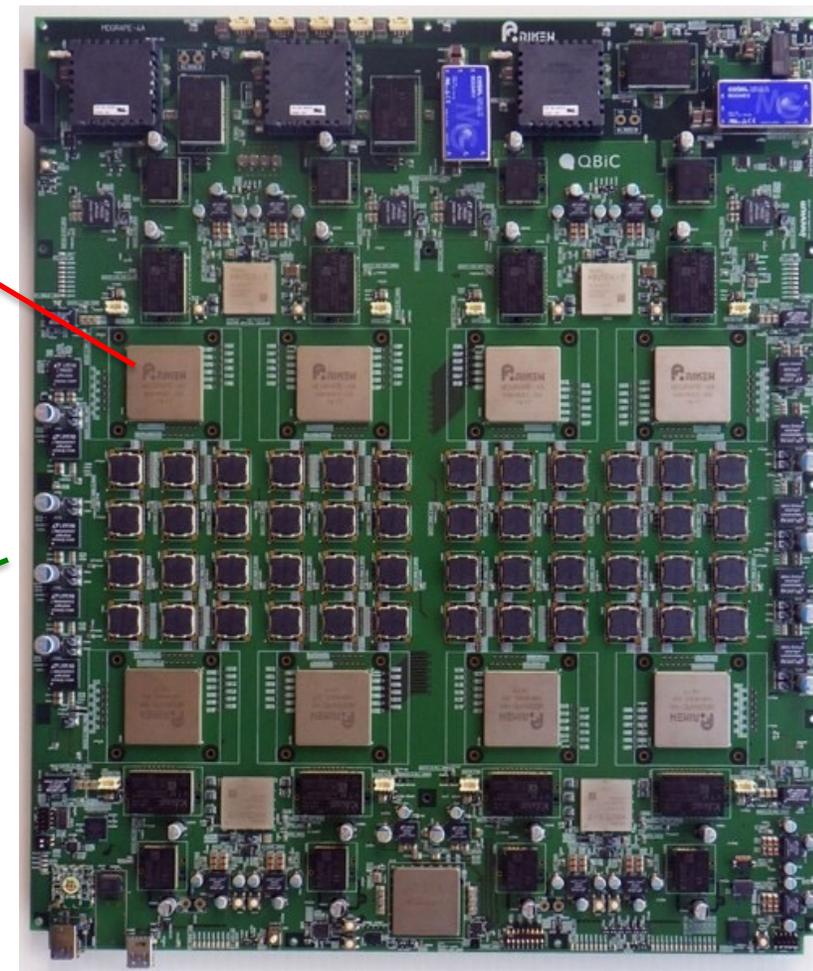
- ▷ Production : Alchip
- ▷ TSMC 40G (40 nm: 10 years old technology), 16.2mm X 16.2mm
- ▷ Logic ~ 50M Gates, SRAM ~ 72M bits
- ▷ 85W (worst case, actual : 40~50W)
- ▷ 0.8 GHz (pipeline)
0.6 GHz(the other)
- ▷ Most of RTL design has been done by us.



MDGRAPE-4A System



+ FPGA based oct-tree network (5GB/s duplex)



*MDGRAPE-4A
System board*

MDGRAPE-4A system



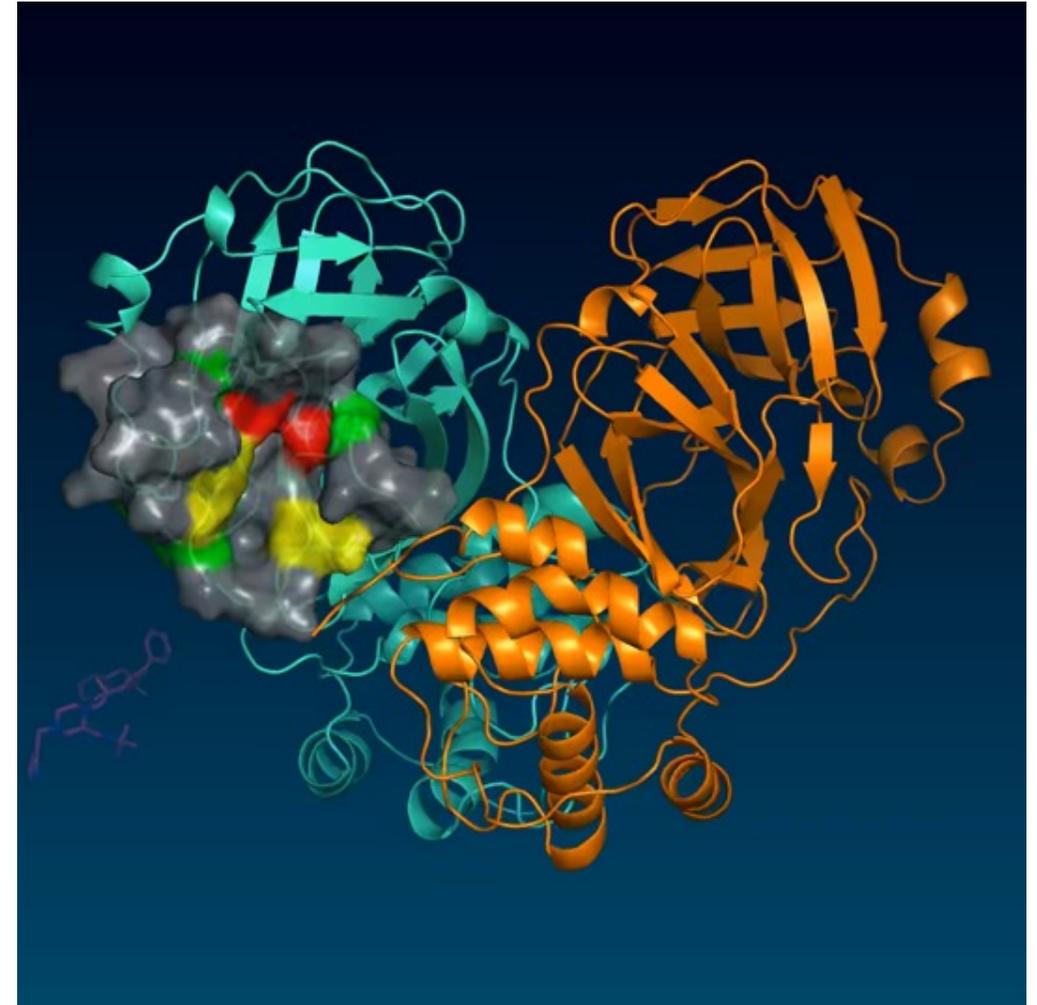
■ MD simulations of SARS-CoV-2 Main Protease (M^{pro} , 3CL protease).

- ▷ We performed ligand-binding simulations for seven HIV-1 protease inhibitors and SARS-CoV-2 M^{pro} .
- ▷ Revealed soft pocket structure upon ligand binding.

Komatsu T.S et al., *Sci. Rep.* **10**, 16986 (2020).
<https://doi.org/10.1038/s41598-020-74099-5>

Dr. Komatsu
(Physics)

Dr. Okimoto
(Pharmaceutical
Sciences)



SARS-CoV-2 Main Protease (6LU7)
+ indinavir, 6 μs simulation

■ハードウェア

▷ 2019年10月完成

■ソフトウェア

▷ Gromacs をもとにしたMDコード (NVT/NPT)、Gromacs の入力ファイルから計算可能

Original code has been provided by Prof. Lindahl

▷ Performance : $\sim 1 \mu\text{sec/day}$ for 100K atom systems (NVT)

5 times faster than that of conventional supercomputer

Slower than our target performance $\sim 5 \mu\text{sec/day}$

reason: some bugs in hardware

slow general-purpose cores

■設備・受託サービスで提供中

■ 二体相互作用以外的高速化

▷ イベント駆動型

- 汎用CPU+ソフトウェアによるハードウェア制御のオーバーヘッドを減らす
- 二体相互作用パイプライン、CPUによる結合力計算、通信等の管理・制御をハードウェア化

■ チップ内ネットワークの見直し

▷ NoC (Network on Chip)

■ 粗視化モデルへの対応

■ 高レベル言語での開発

▷ Chisel : Scalaの回路設計フレームワーク

- RISC-V 実装 Rocket Chip の設計言語

■ 回路設計

- ▷ 全体・イベント駆動 大野
- ▷ 汎用CPU 新稲
- ▷ 二体相互作用パイプライン 加藤・泰地・森本
- ▷ NoC 大村・新稲
- ▷ 遠距離相互作用 森本・泰地

■ FPGA 実装での検証

- ▷ 富士通 委託・協業

■ ASIC

- ▷ 合成 泰地
- ▷ RaaSへの参加 <https://raas-cip.org>
 - 黒田 忠広 東京大学教授が主催するチッププロトタイプ開発事業

- GRAPE-1Aの複製から専用計算機の道に
 - ▷ 牧野さんにはいろいろお世話になりました。
- GRAPE以外の専用計算機を経てMDGRAPEに

- MDGRAPE-4でアクセラレータ型からSoCへ
- MDGRAPE-4Aで実用計算
- MDGRAPE-5開発中