



# 単精度190Tflops GPUクラスタ (長崎大)の紹介

長崎大学工学部  
超高速メニーコアコンピューティングセンター  
テニュアトラック助教  
濱田 剛

# 概要

- GPU (Graphics Processing Unit)について簡単に説明します.
- GPUクラスタが得意とする応用問題を議論し
- 長崎大学でのGPUクラスタによる
  - ◆ 取組方針
  - ◆  $N$ 体計算の高速化に関する研究内容を紹介します.
- まとめ

# GPU (Graphics Processing Unit)

- パソコンの3D描画アクセラレータ
- 演算性能が急速に向上
  - ◆  $\sim 1$  T flops/board
- オンボードメモリバンド幅
  - ◆  $\sim 160$  GB/s
- 低価格
  - ◆ 400 Gflops程度のボードが2万円弱



# GPGPU

- General-Purpose computation on GPUs
- GPUをN体計算や流体計算などに応用する試み
- 応用方法はメーカーごとに若干異なる. 互換性もなし
  - ◆ CUDA (Nvidia)
  - ◆ CAL(AMD)
  - ◆ Ct (Intel)
  - ◆ OpenCL(AMDはやる気満々. Intel/Nvidiaは...)

# CUDA (Compute Unified Device Architecture)

- NVIDIA GPU用プログラミング環境
  - ◆ 言語, ライブラリ, コンパイラ
- 他と比べGPGPUユーザーが多い
  - ◆ 早期に(2006年11月)公開, 無料
- 超並列プログラミングが基本
- ボードあたり5千~1万スレッド程度の並列性が必要で性能を出すのが大変
- GeForce 8800以降に対応

# PCクラスタと比べたGPUクラスタ

- PCクラスタにもいろいろあるが、大きな違いは1つ
  - ◆ ノード当たりの演算性能が高い
- 見方を変えると
  - ◆ ネットワーク性能が相対的に低い
- 1ソケット100万円とかのPCクラスタと比べると
  - ◆ 同程度の演算性能を1/10のコストで手入可能
- 似ているようで使い方は結構違う

# 超高速メニーコアコンピューティングセンター (Nagasaki Advanced Computing Center: NACC)

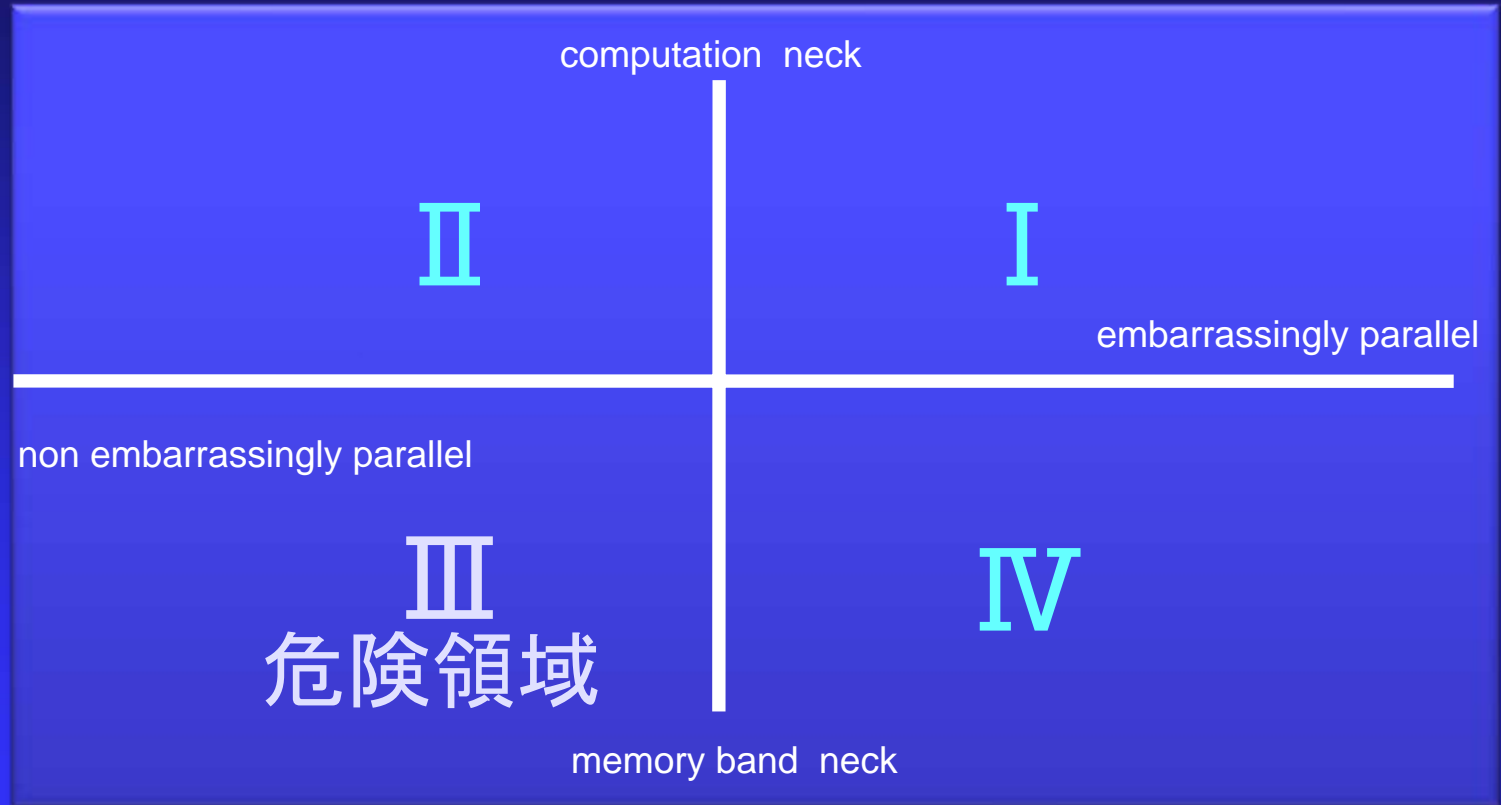
- 平成21年度から運営開始
  - ◆ センター長： 小栗 清
- 短期的な目標
  - ◆ GPUクラスタの応用方法に関する研究
  - ◆ 地方大学でも導入可能な低コスト・スーパーコンピュータの実現
- 中・長期的な目標
  - ◆ 学内における並列計算機応用研究の促進
  - ◆ 次世代LSIアーキテクチャ・コンパイラ技術の開発

# 経緯

- 平成19年度文部科学省科学技術振興調整費(若手研究者の自立的な研究環境整備促進)採択課題「地方総合大学における若手人材育成戦略」として開始
  - ◆ 128台のGPU クラスタを構築 (2008/3)
- 平成21年度政府補正予算-長崎大学概算要求「新世代GPU型スーパーコンピューティング環境の実現」でサポート
  - ◆ GPUの台数を256台に増強 (2008/11)
- 「超高速メニーコアコンピューティングセンター」を工学部内に設置



# 応用分野の分類



NACCでは領域 I, II, IVをターゲット

# NACC GPUクラスタの対象領域

- 領域 I, II, IVのみをターゲット
  - ◆ N体計算(領域I, II, IV)
  - ◆ 電磁波・流体計算(領域IV)
  - ◆ 画像処理・統計処理(領域I, IV)
- 領域IIIに分類される応用(並列FFT, 並列CFDなど)は視野に入れない
  - ◆ ネットワークがボトルネック
  - ◆ 高価な高速・低レイテンシネットワークが必要
  - ◆ 長崎大ではI, II, IVの領域を開拓することを優先

# 長崎大学GPUクラスタ仕様(2008/11)

- Core2Quad Q6600が128ノード
- 各ノードにNvidia GeForce GPUを
  - ◆ 8800GTS-OC x 128 ボード
  - ◆ 9800GTX+ x 128ボード
- 単精度190 Tflops
- 1 GbE 48ポートスイッチ(NetGear) x 4 台をスタック  
接続
- 開発コスト : 2千万円程度

# NGC190T (NACC GPU Cluster)



部屋面積 24 m<sup>2</sup>

部屋給電能力 640 A

1人で1週間もあれば組立可能

# GPUによるN体計算

- 長崎大学でのGPU研究発展の原点
- GPUによるN体計算を深く理解することが他の応用分野に展開する際に重要
- 以後少々詳しく説明します.

# N体計算

基礎方程式  
(重力N体)

$$\frac{d^2 r_i}{dt^2} = \sum_{j \neq i} -\frac{G m_j r_{ij}}{r_{ij}^3}$$



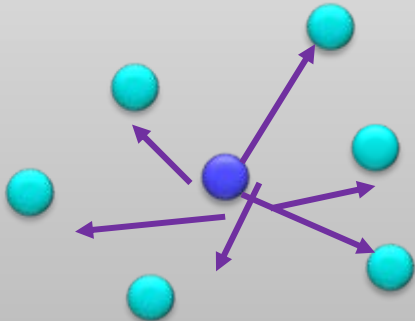
逐次処理部分  
(積分計算など)  
を担当



並列処理部分  
(相互作用の計算)  
を担当

# N体計算のアルゴリズム

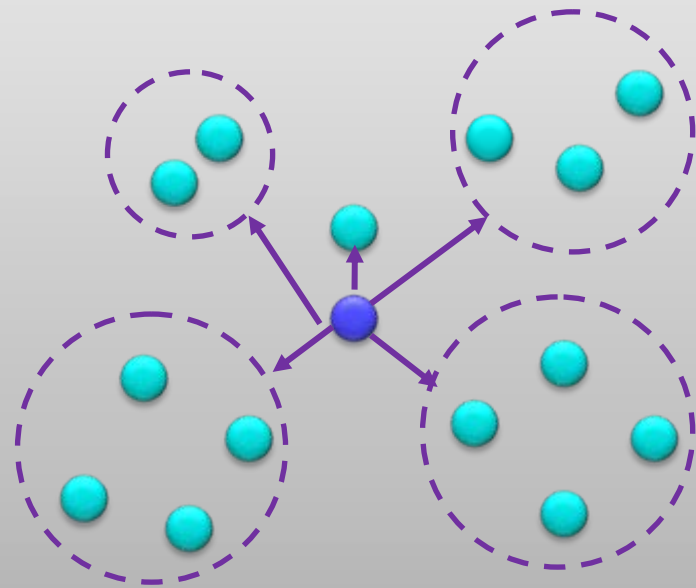
Direct Summation  
Algorithm



$O(N^2)$ : Small

領域Iに近い問題

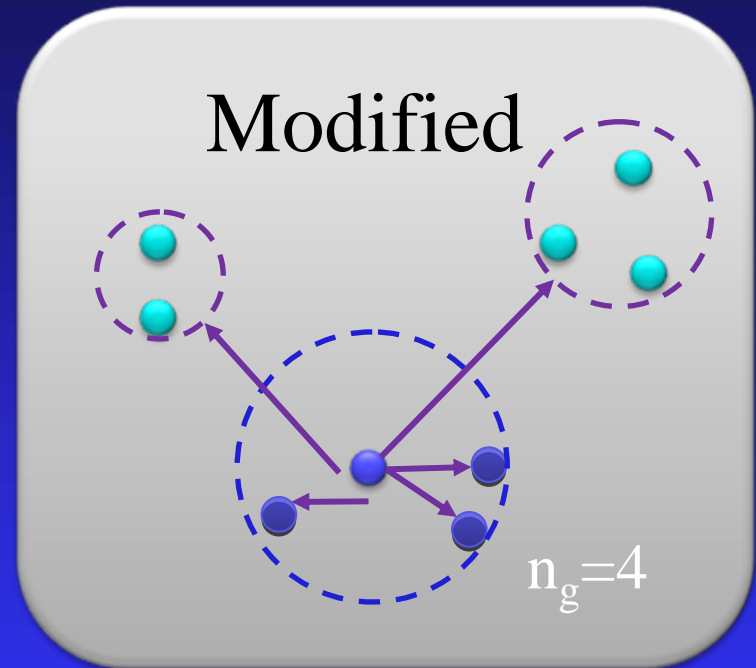
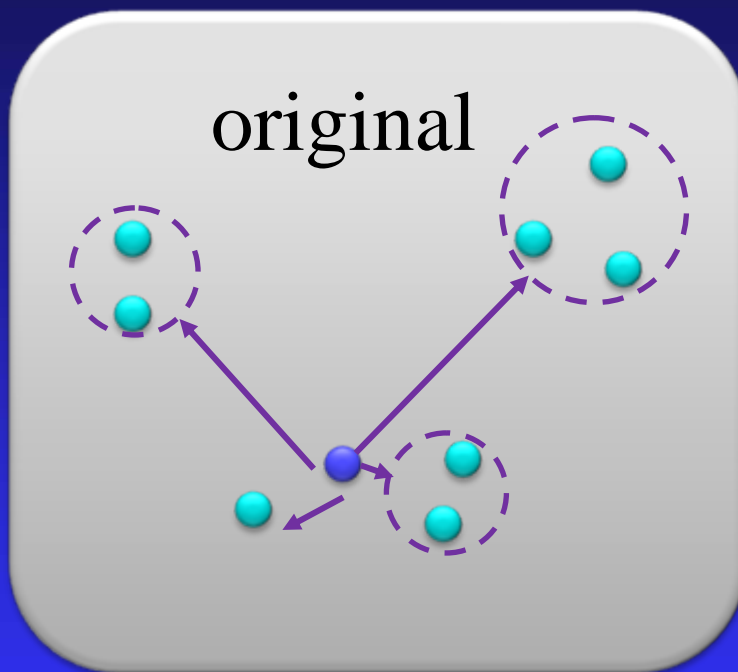
Tree code



$O(N \log N)$ : Large

領域IIの問題

# Modified Tree code

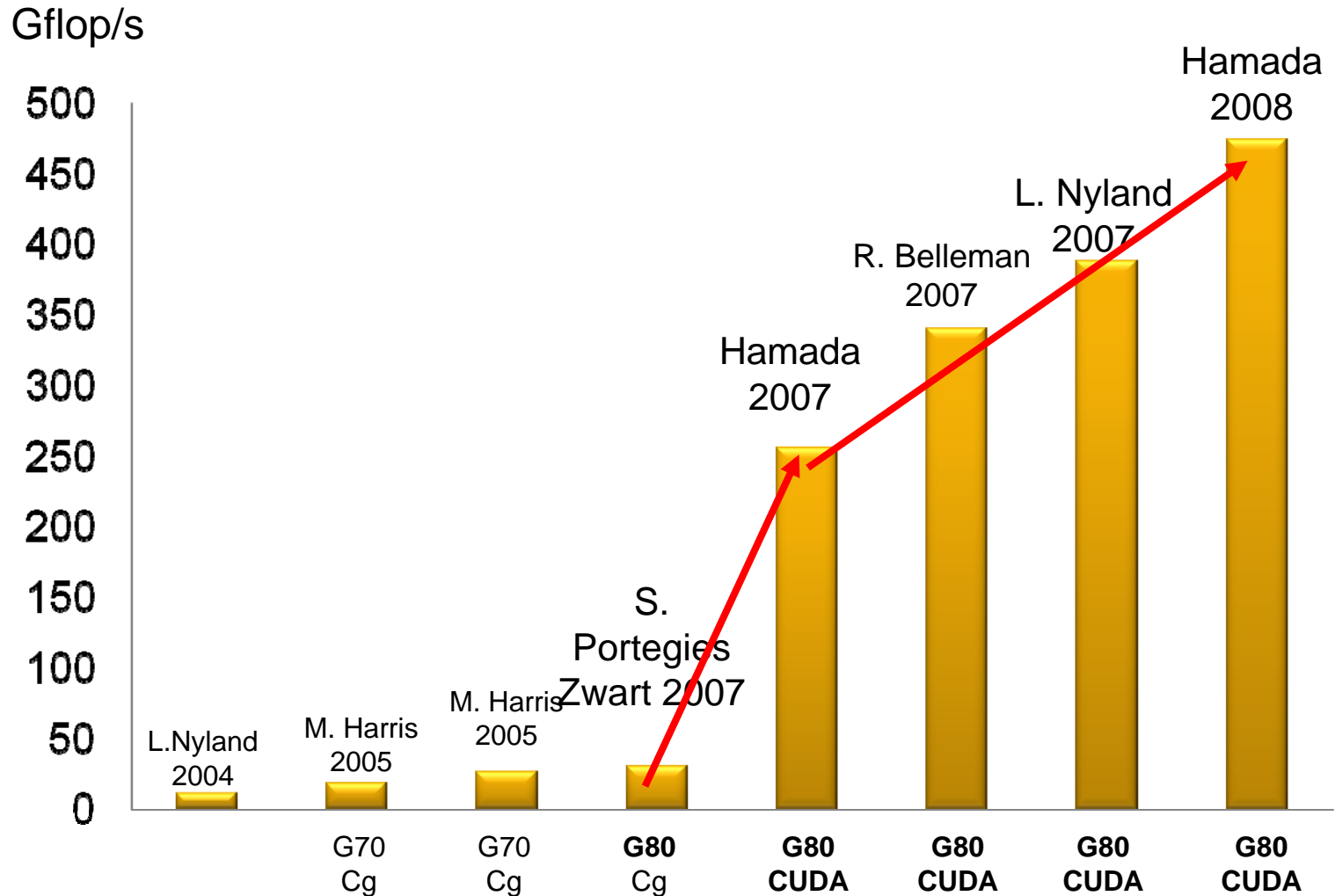


$n_g$  個の粒子が相互作用リストを共有

	Original	Modified
List creation cost on the host	$N$	$N/n_g$
Interaction list length	shorter	longer



# GPUでのN体計算(Direct Sum.)



単体GPUのN-bodyの性能推移(直接和, 1相互作用38演算換算)

# 最適化

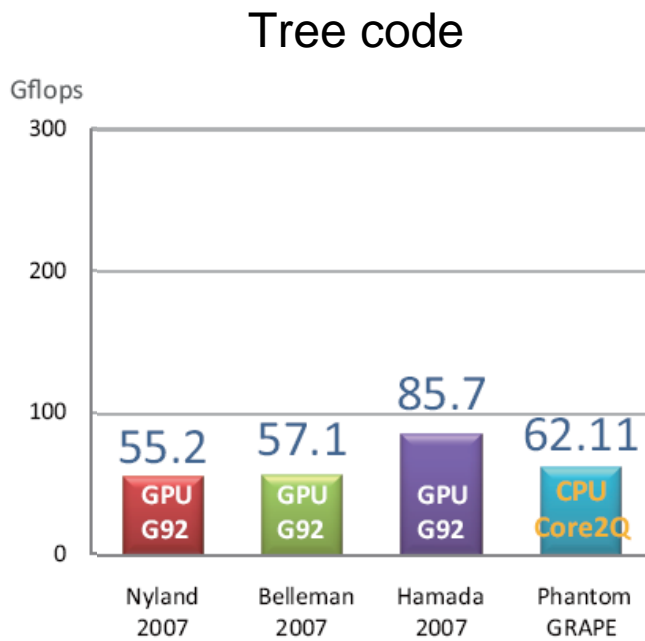
- Shared Memoryの利用
- ループアンロール
- メモリアクセスの最適化

- 詳細は

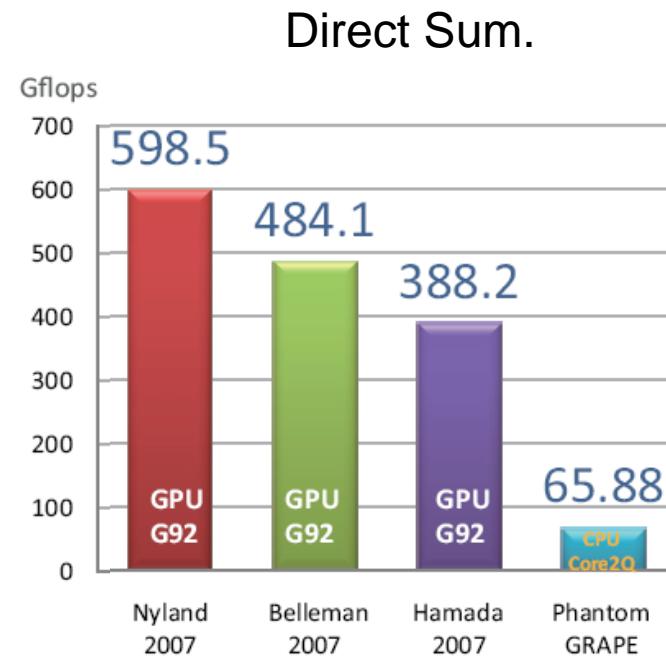
情報処理学会会誌「情報処理」  
2009年2月号参照

(成見, 濱田, 小西, 情報処理学会会誌「情報処理」2009年2月号,  
IPSIJ-MGN500207)

# Tree codeではもうひと工夫が必要



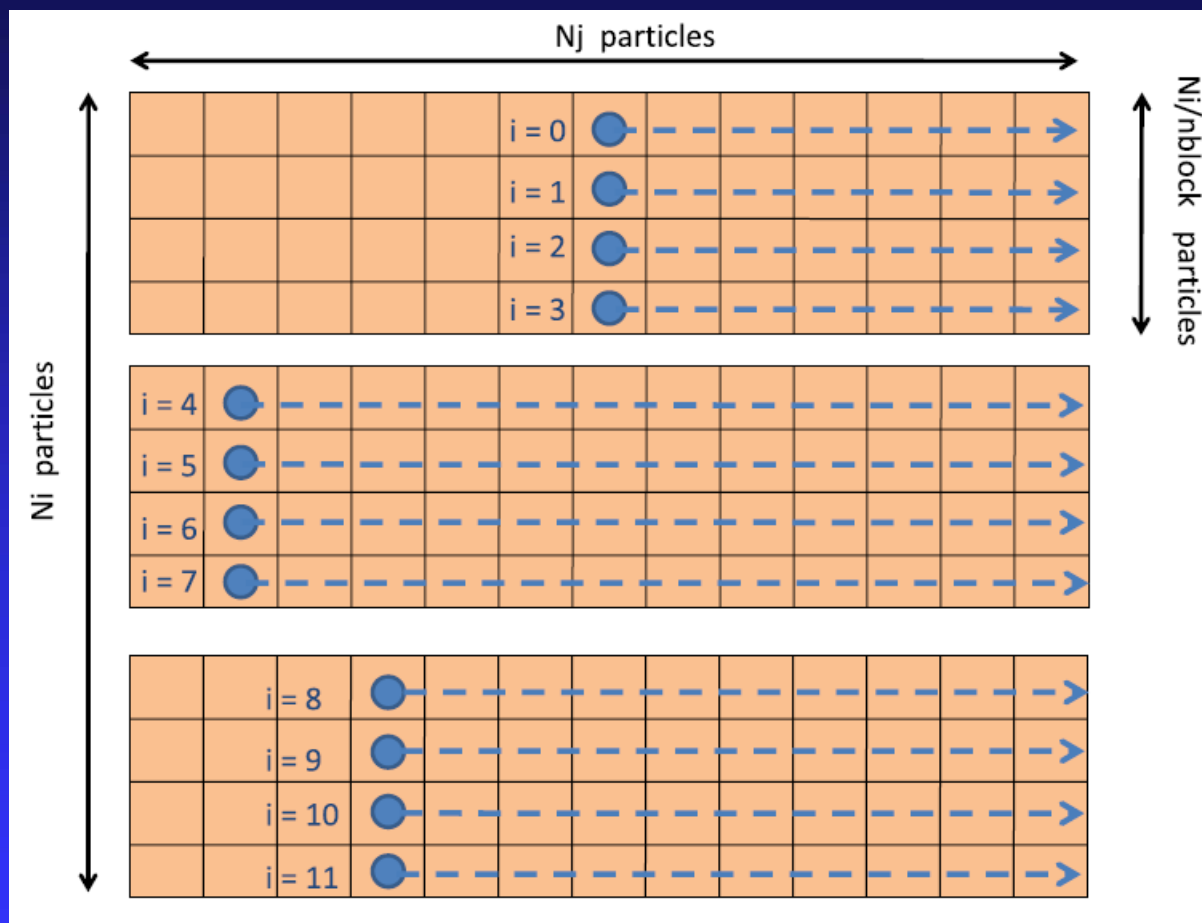
N= 4 M  
plummer



N=64 K

# 性能の出し方(Direct Sum.の場合)

全てのスレッドが別々の*i*粒子への力を計算 (*i*並列)

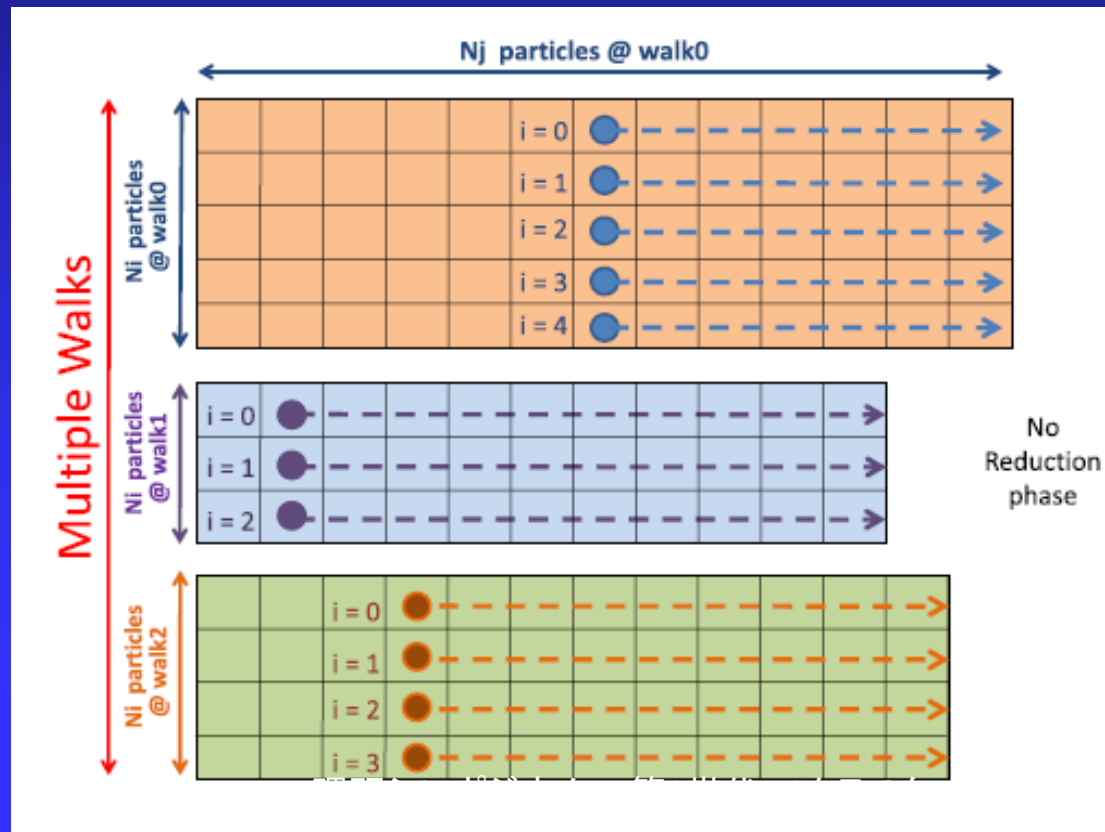


*i*粒子が数千～1万程度の場合に効率が高い

# 性能の出し方(Tree codeの場合)

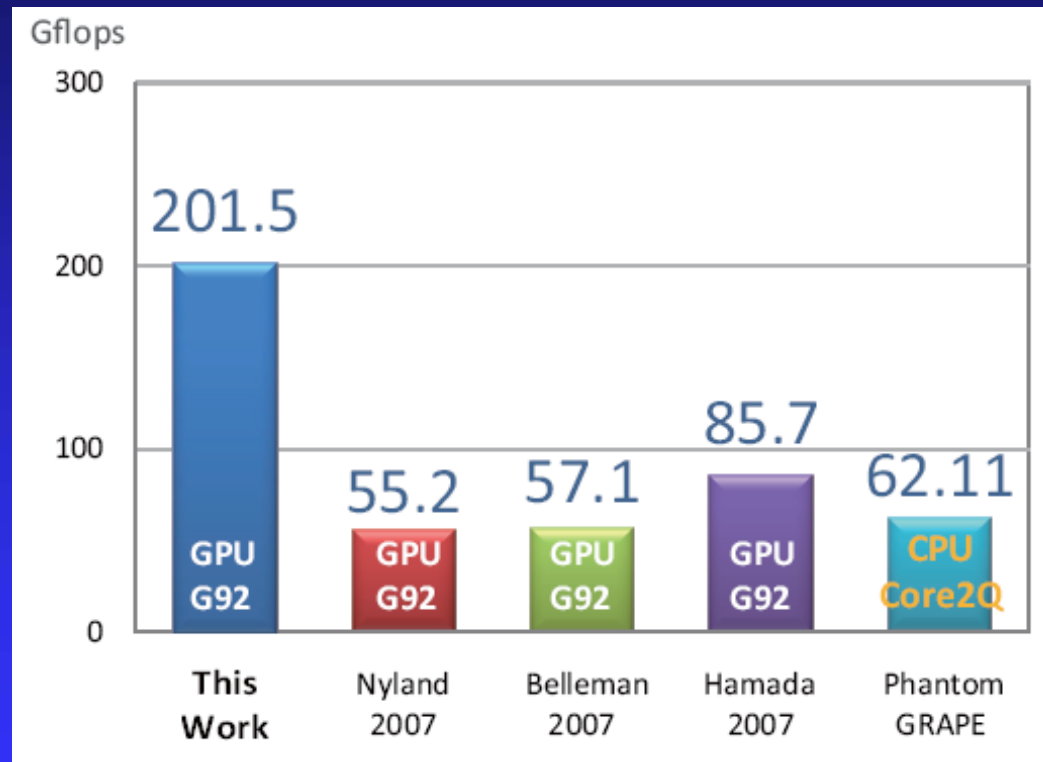
## ■ Multiple Walks ( Hamada et al 2008 )

- ◆ 複数の $i$ 粒子のグループ(walk)を一度に計算



# Multiple Walksによる効果

Tree code

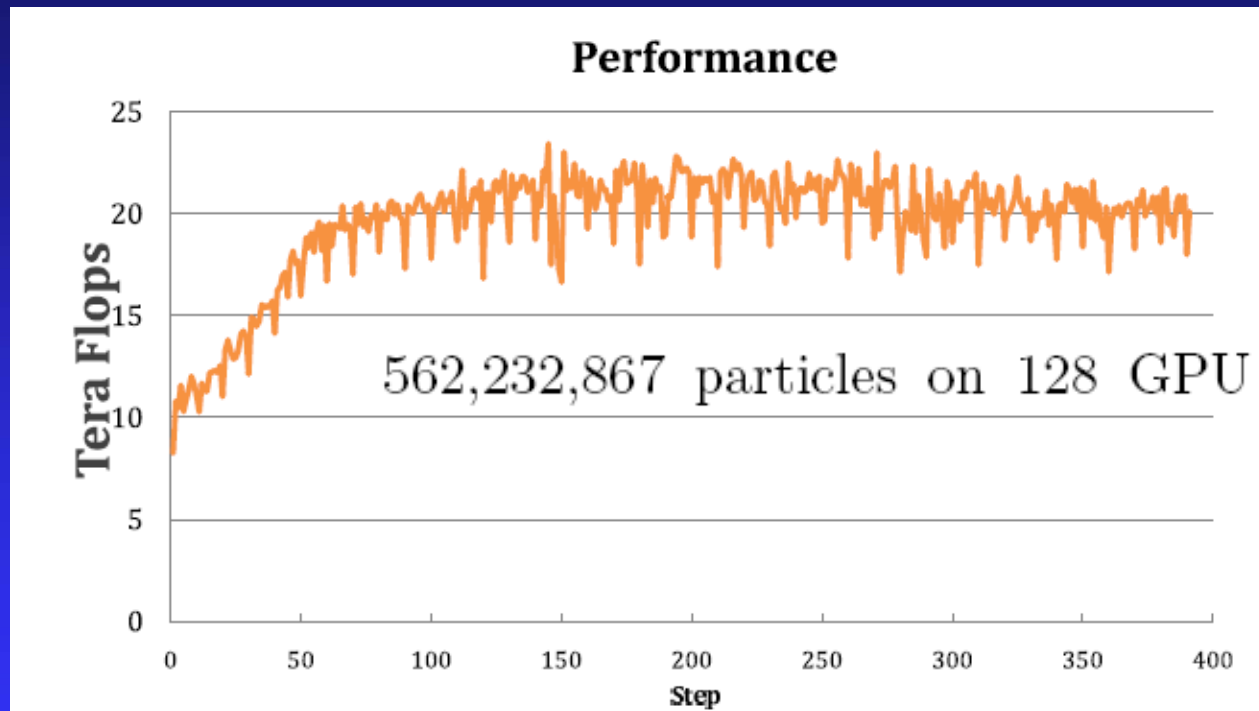


N= 4 M, plummer, Ncrit = 2000

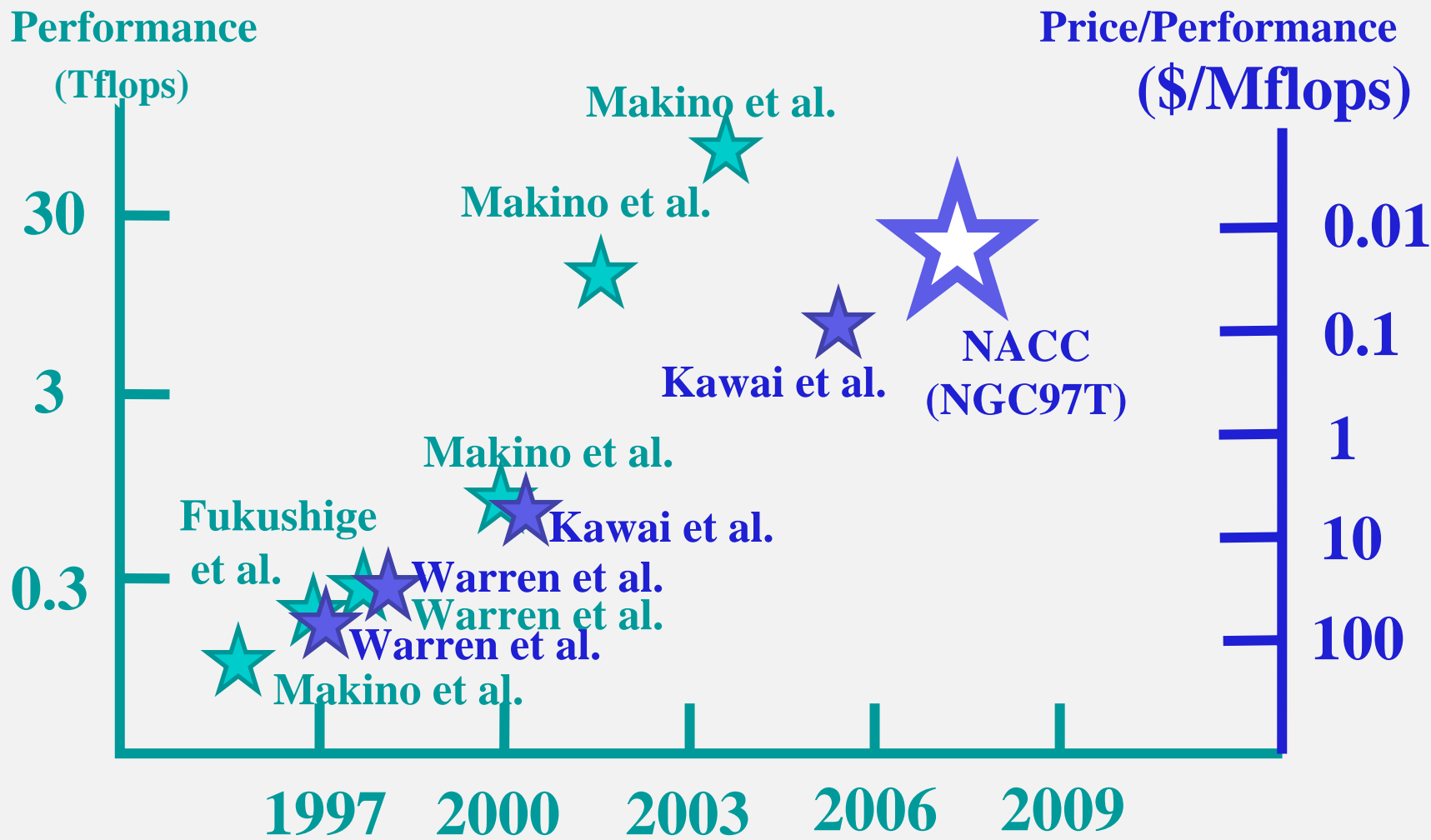
(Hamada et al. SC08)

# 128GPUでのtree code 演算性能

(Hamada et al. SC08)



# コスト性能



歴代Gordon Bell賞(ピーク性能部門)



歴代Gordon Bell賞(コスト性能部門)



# 電力性能

#Direct Sum. (Hamada et. al. submitted to FCCM'09)

	GRAPE-5	Bioler-3	Cray XD1	PROGRAPE-4	ASUS EN8800GTX	MSI N9800GTX+	Core2Quad Q6600	Atom 230
Device Chip	ASIC 300k gates	FPGA XC2VP70-5	FPGA XC2VP50-7	FPGA XC3S5000-5	GPU G80	GPU G92	CPU SSE	CPU SSE
Development Year	1997	2004	2004	2006	2007	2008	2007	2008
Chip technology	500 nm	130 nm	130 nm	90 nm	90 nm	65 nm	65 nm	45 nm
Chips/board	8	4	1	4	1	1	1	1
Pipelines/chip	2	16	10	16	N/A	N/A	N/A	N/A
Frequency (MHz)	80	133.3	120	100	1350	1890	2400	1600
Gflops/chip	24.3	81	45.6	60.8	470.8	687.1	70.3	6.35
Gflops/board	48.6	324.2	45.6	243.2	470.8	687.1	70.3	6.35
ratio of performance (against GRAPE-5)	1.0	6.7	0.9	5.0	9.7	14.1	1.45	0.13
Power Consumption per board (without host)	80 W	30 W	N/A	5 W	148 W	122 W	49 W	3.1 W
Power Consumption per chip	8 W	7.5 W	N/A	1.3 W	148 W	122 W	49 W	3.1 W
Gflops/Watt	0.61	11	N/A	49	3.2	5.6	1.43	2.05

- GPUはCPUよりも良いがFPGAには及ばない
- 今後は電力性能も優れたアーキテクチャの開発の余地 → NACC長期目標

# NACCでのその他の応用

- 電磁場解析 (領域I, IV)

- ✓ GPS用アンテナの設計

- 画像レジストレーション (領域IV)

- ✓ 災害時映像を用いたPIV解析

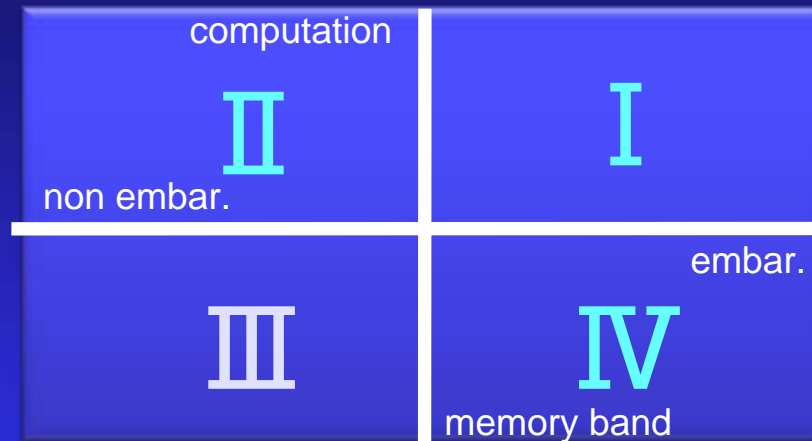
- 流体解析 (領域IV)

- ✓ 東シナ海藻類の生理生態学的課程と流体環境の相互作用のモデル構築

等々

# まとめ

## ■ GPUクラスタが得意とする領域 ... I, II, IV



- ◆ 領域IIIで無理に頑張らないのが長崎大の特徴
- ◆ 応用問題がどの領域に属するのかを理解するセンスがGPUクラスタでは特に重要
- ◆ N体計算における経験はそれらを理解する上でとても有用