

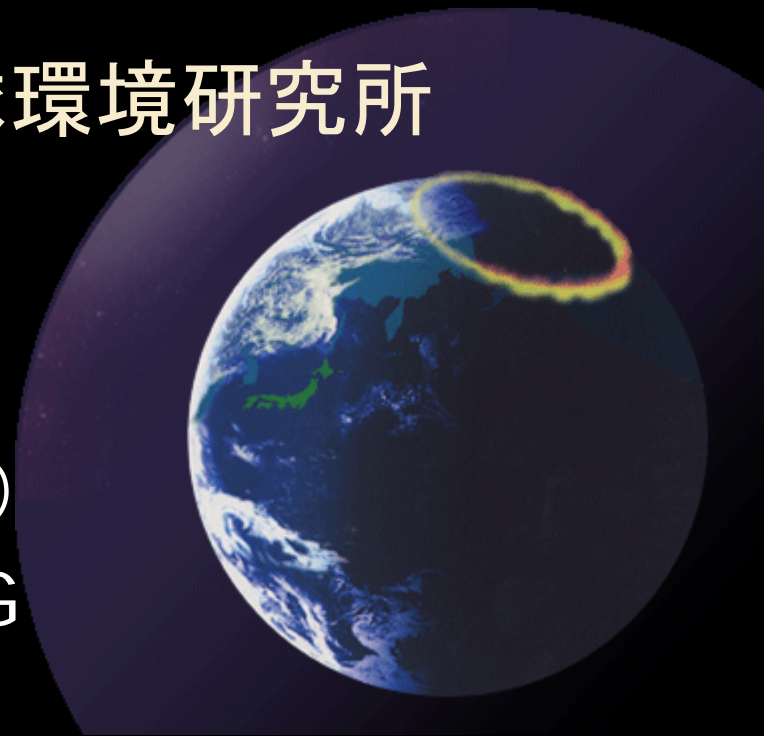
# スカラー型超並列計算機に 向けたプラズマ運動論コード の性能評価

梅田隆行  
名古屋大学太陽地球環境研究所

協力:

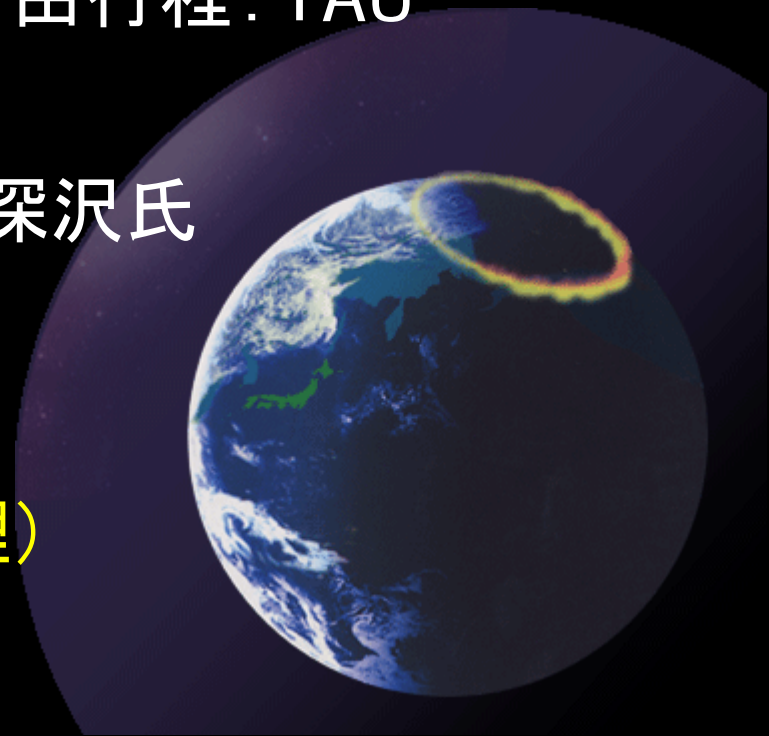
深沢圭一郎(九大基盤センター)

SS研マルチコアクラスタ性能WG



# プラズマシミュレーション

- 宇宙の99.9%以上の体積を占めているのはプラズマ(電離気体)
- その大部分は希薄であり、無衝突。
  - ex. 太陽風プラズマの平均自由行程: 1 AU
- 様々な近似階層が存在
  - 磁気流体力学(MHD) ⇒ 深沢氏
  - 運動論: 粒子 ⇒ 藤本氏
  - ブラソフ
  - (無衝突プラズマの第一原理)
  - ハイブリッド ⇒ 寺田氏



# プラズマ“運動論”コード

- 基本方程式: プラズマの運動

オイラー系 (分布関数)

$$\frac{\partial f}{\partial t} + \mathbf{v} \cdot \frac{\partial f}{\partial \mathbf{x}} + \frac{q}{m} (\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \frac{\partial f}{\partial \mathbf{v}} = 0$$

ブラソフ方程式

ラグランジュ系 (粒子)

$$\begin{aligned} \frac{d\mathbf{r}_n}{dt} &= \mathbf{v}_n \\ \frac{d\mathbf{v}_n}{dt} &= \frac{q_n}{m_n} (\mathbf{E} + \mathbf{v}_n \times \mathbf{B}) \end{aligned}$$

ニュートン-ローレンツ式

+ マックスウェル方程式 (電磁場) と電荷保存の式

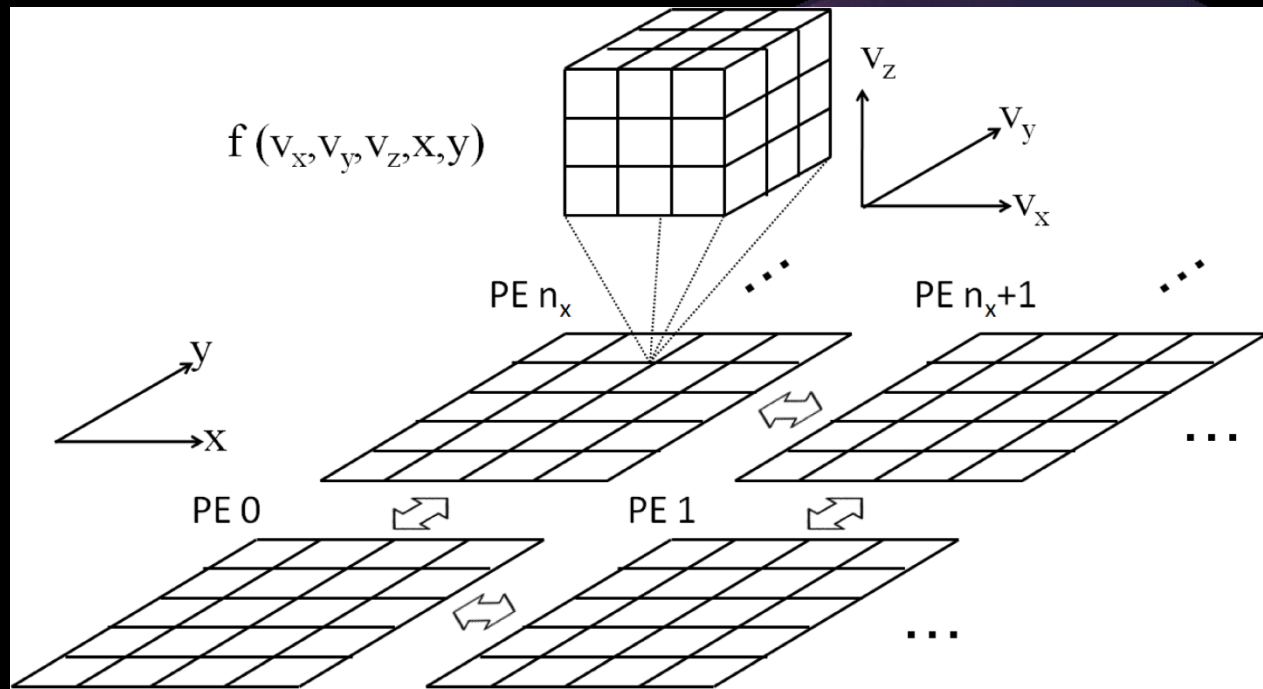
$$\begin{aligned} \nabla \times \mathbf{B} &= \mu_0 \mathbf{J} + \frac{1}{c^2} \frac{\partial \mathbf{E}}{\partial t} & \nabla \cdot \mathbf{E} &= \frac{\rho}{\epsilon_0} \\ \nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t} & \nabla \cdot \mathbf{B} &= 0 \end{aligned}$$

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} = 0$$

# ブラソフコード

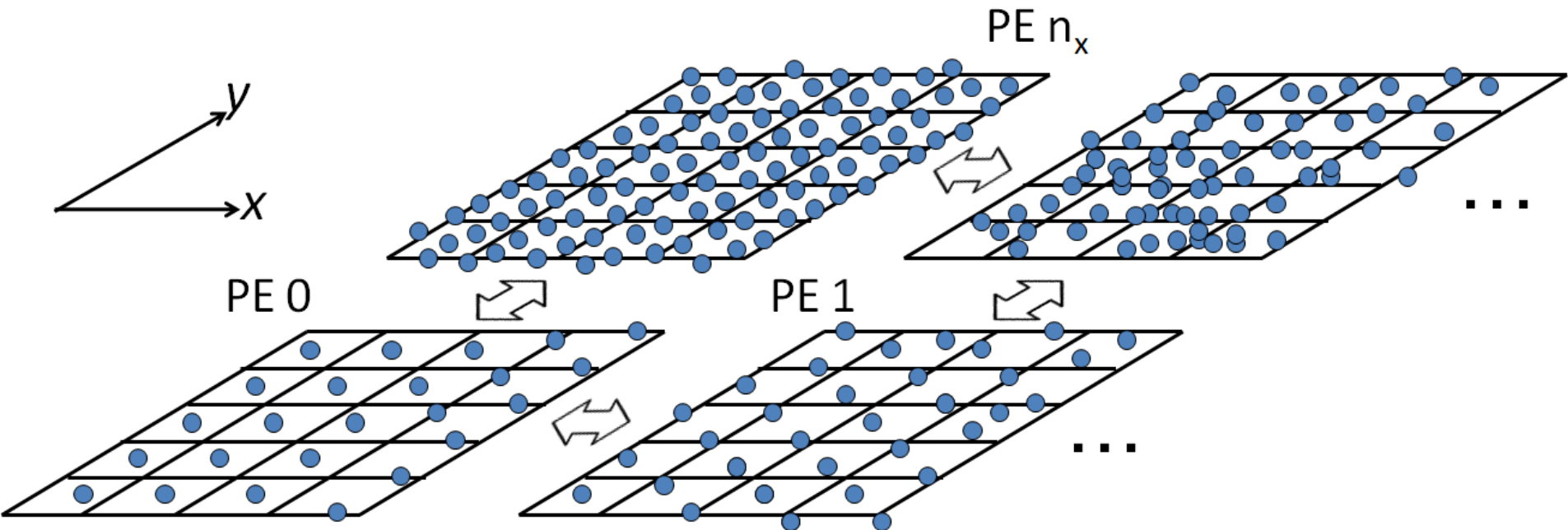
- 格子系、全てのオイラー変数、連続アクセス
- ただし、超多重ループ(4-6次元)  
個々のループ長が短い
  - ex.  $40 \times 40 \times 40 \times 40 \times 40 \sim 4\text{GB}$

- 共有メモリが小さいと不利  
(6次元は「京」では難しい)



# 粒子コード

- 格子系(連続体、オイラー変数)の電磁場と、ラグランジュ変数の粒子が混在
- 格子データ(メモリ)へのアクセスがランダム
- プロセス内の粒子数が不均一



# 目的

- 名大基盤センターの次期スパコンに向けた、プラズマ運動論コードの性能評価
- 特に、ノード内(マルチコア環境)の性能を重視  
⇒ノード内性能の向上が全体性能の向上につながる

## 講演内容(成果)

1. ブラソフ・粒子、各コードのノード内性能
2. 3次元FDTD(電磁場ソルバ)のチューニング
3. ソーティングのOpenMPスレッド並列化

# 計測環境

※8Bulldozer modules

システム:

- FX1(名大)
- 京(理研)
- FX10(富士通)
- HA8000(東大)
- HX600(名大)
- DELL R815-1(STEL)
- DELL R815-2(STEL)
- DELL R610(STEL)
- RX200S6(九大)

CPU:

SPARC64 VII

SPARC64 VIIIfx

SPARX64 IXfx

Opteron (Barcelona)

Opteron (Shanghai)

Opteron (Magny-Cours)

Opteron (Interlagos)

Xeon (Nehalem)

Xeon (Westmere)

コア数:

4

8

16

4 x 4

4 x 4

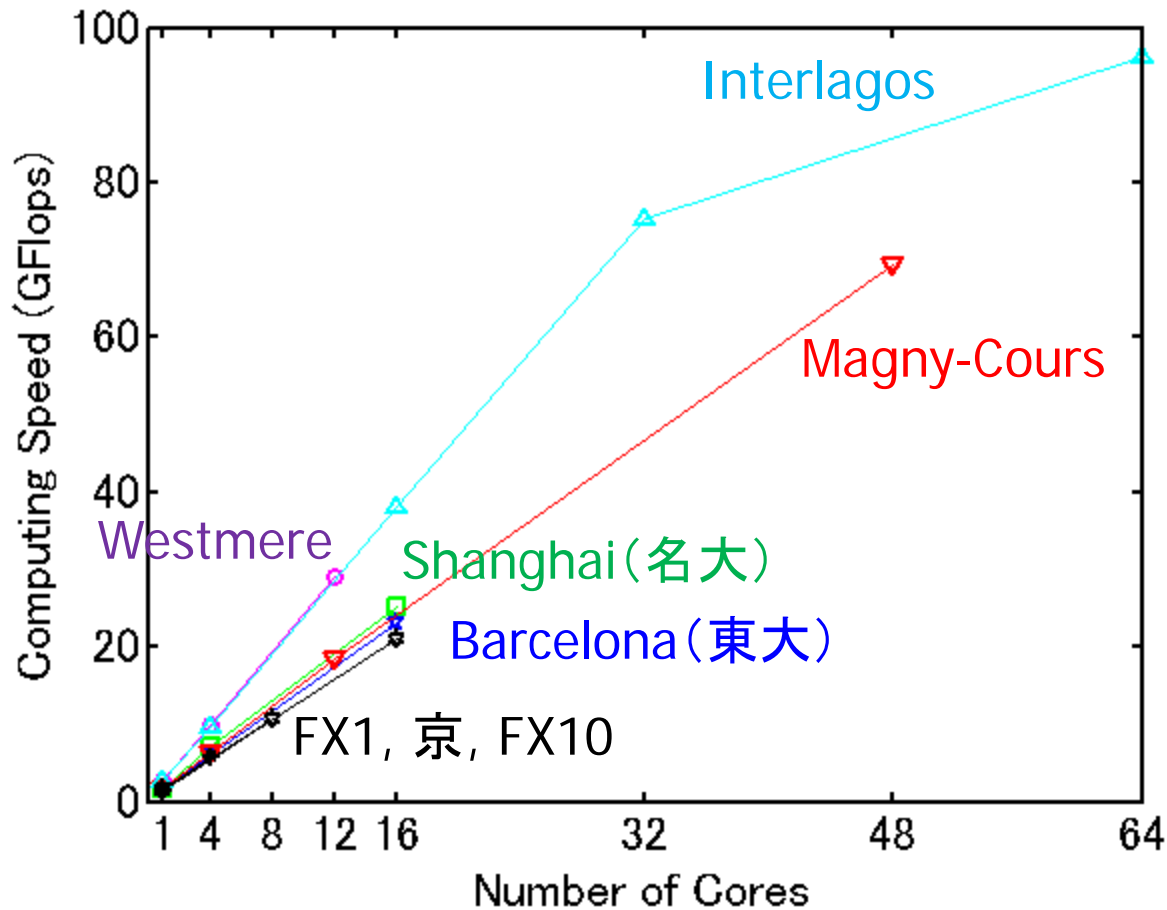
12 x 4

16※ x 4

4 x 2

6 x 2

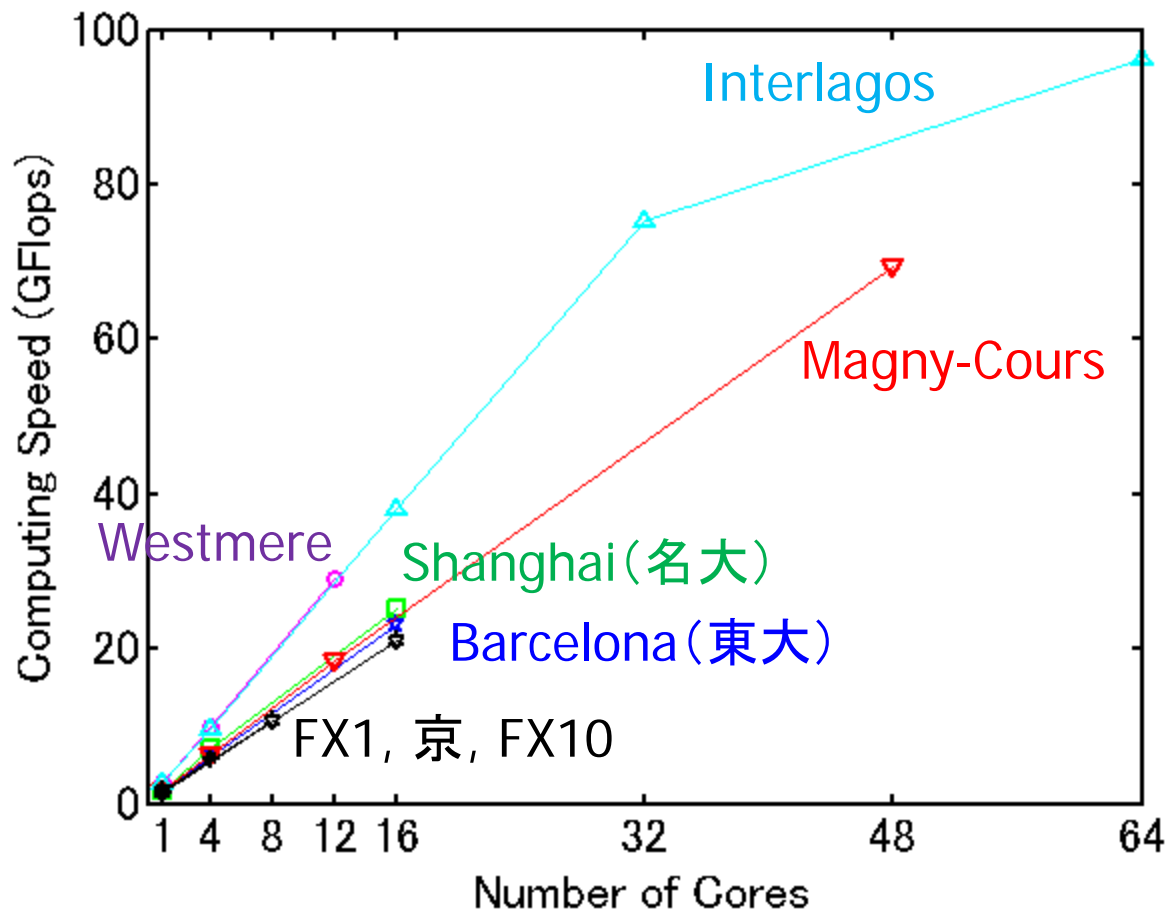
# ブラソフコード実行性能①



- 1GB/coreの弱いスケーリング  
40x20 x 30<sup>3</sup> /core
- FX1はコンパイラのバージョンアップで性能が向上
- 京/FX10は8命令/サイクル
- FX1/京/FX10はほぼ同じ性能



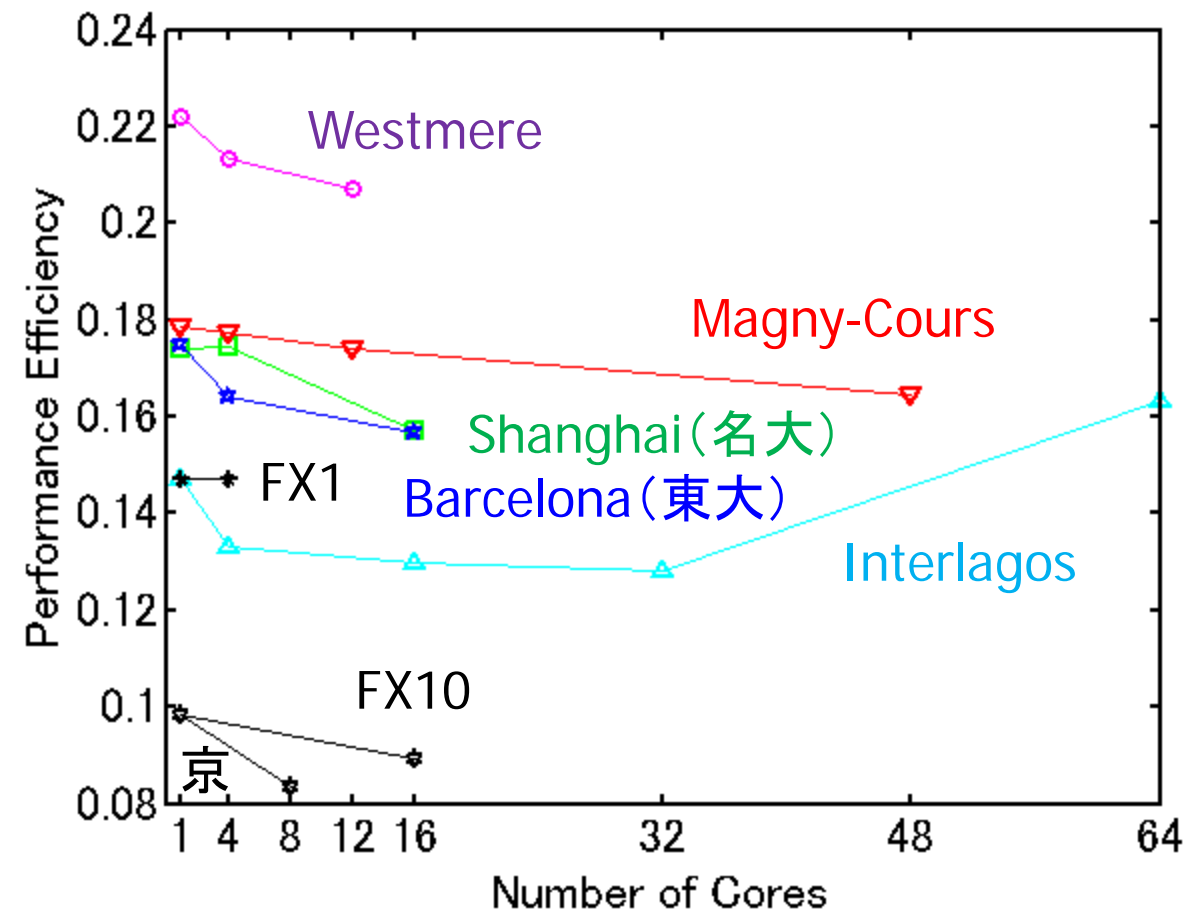
# ブラソフコード実行性能②



- FX1/京/FX10は東大T2Kに劣る
- 旧世代Opteronはほぼ同じ性能
- Bulldozerは、8命令/サイクルの32モジュールを64コアで共有
- 32コアまではWestmereと同等

Interlagosの32コアと64コアは同じ理論性能

# ブラソフコード実行効率



- 8命令/サイクルのCPU (京, FX1, Interlagos)の実行効率が低い
- Westmereは20%超
- 旧世代Opteronはほぼ同じ性能  
⇒新世代CPU用に再チューニングする必要あり

※京とFX10は強いスケーリング

# 粒子コード性能評価

- 粒子の速度を計算:

- 各粒子の位置の電磁界データにアクセス  
6つ配列データ(ex,ey,ez,bx,by,bz)のロード

$$\frac{d\mathbf{v}}{dt} = \frac{q}{m} (\mathbf{E} + \mathbf{v} \times \mathbf{B})$$

- 電流密度を計算:

- 各粒子の電流と電荷を格子に足しこむ  
4つの配列データ(jx,jy,jz,rho)のロード・ストア

$$\mathbf{J} = \sum q\mathbf{v}$$

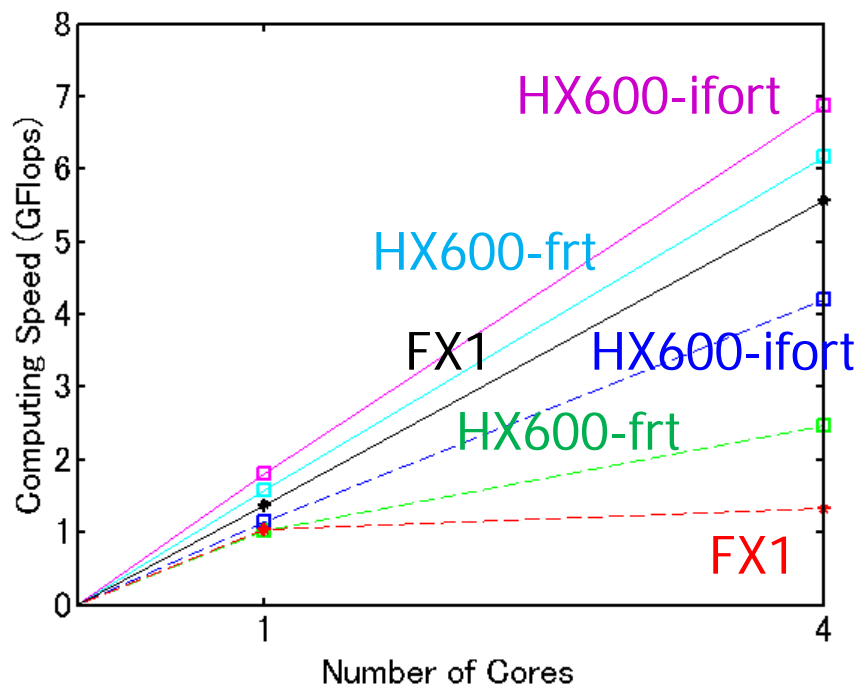
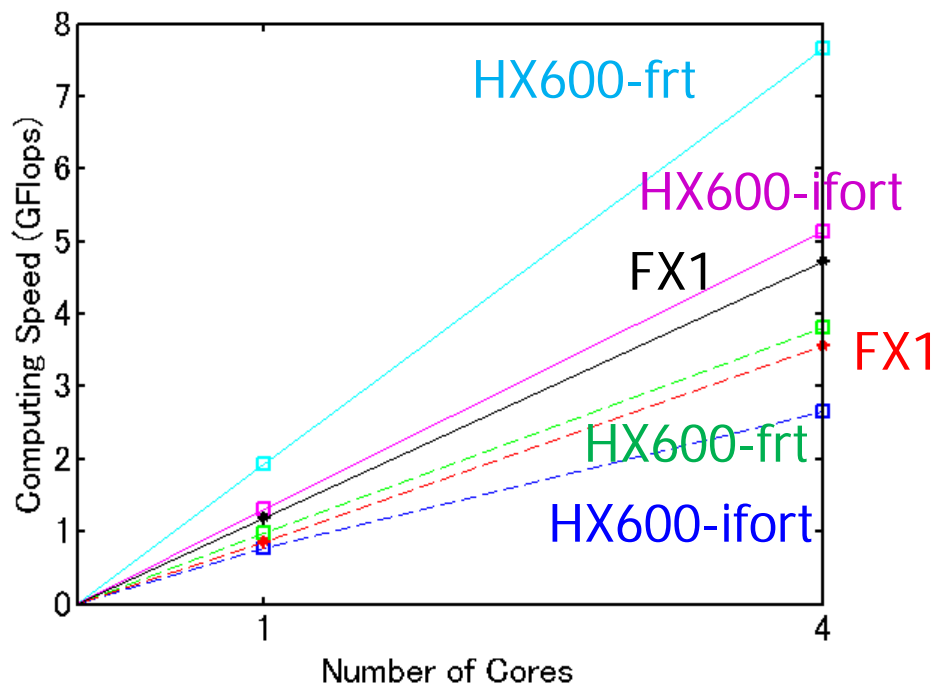
- ソーティング:

- 速度・電流のメモリアクセスが連続的になるように、  
粒子の配列を整頓  
基本的に整数演算

# 粒子コード実行性能 (ソートあり・なし)

速度更新

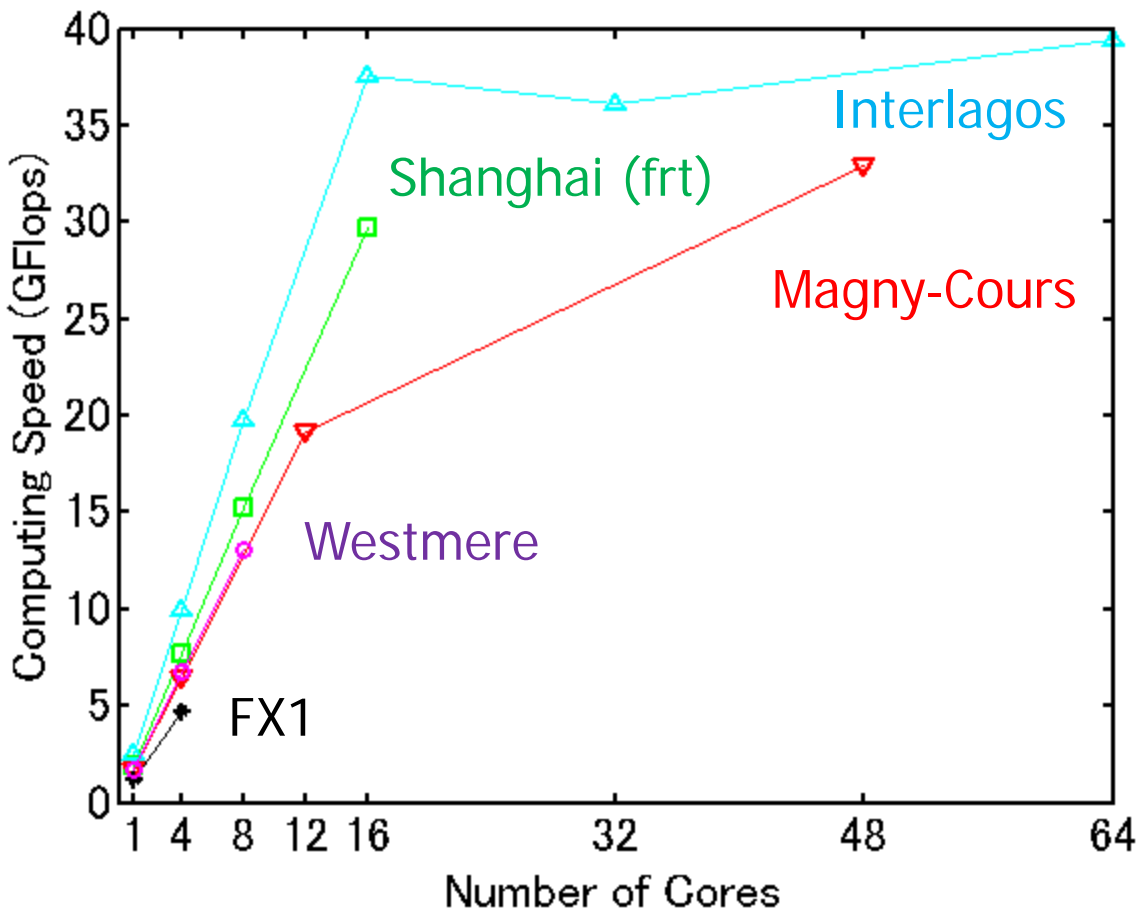
電流密度



- シリアル:ソートによって1.5-2倍性能が向上
- 並列:ソートによるキャッシュ競合の軽減、スケーラビリティ向上
- frt はロードが速く、ifort はストアが速い？

# 粒子コード(速度更新部) 実行性能

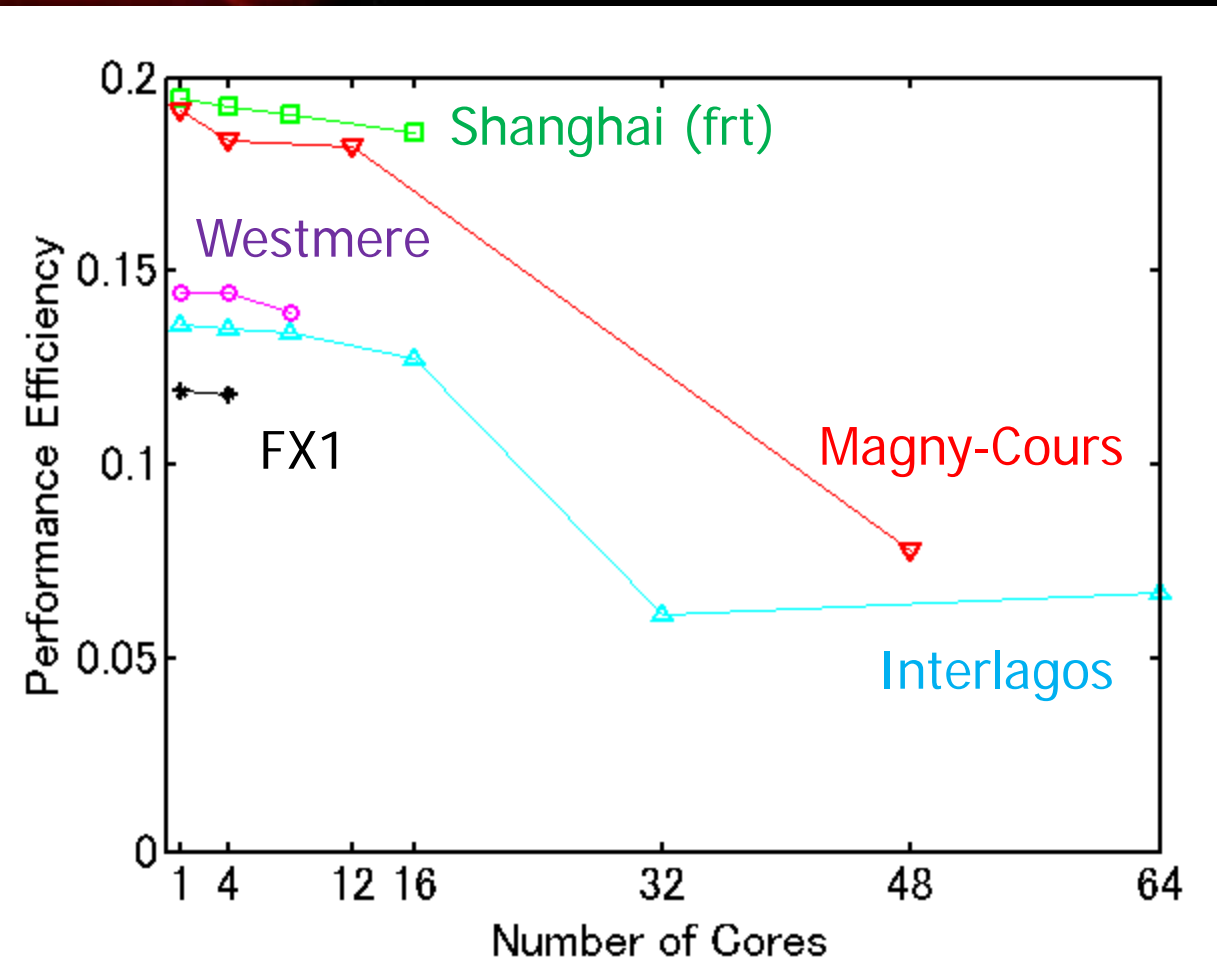
$$\frac{dv}{dt} = \frac{q}{m} (E + v \times B)$$



- 256粒子/cellの強いスケーリング
- 16コアまではBulldozerが高速
- 旧世代CPUはどれも似たような性能  
⇒ Xeonが遅い...
- 16コアまではスケール

# 粒子コード(速度更新部) 実行効率

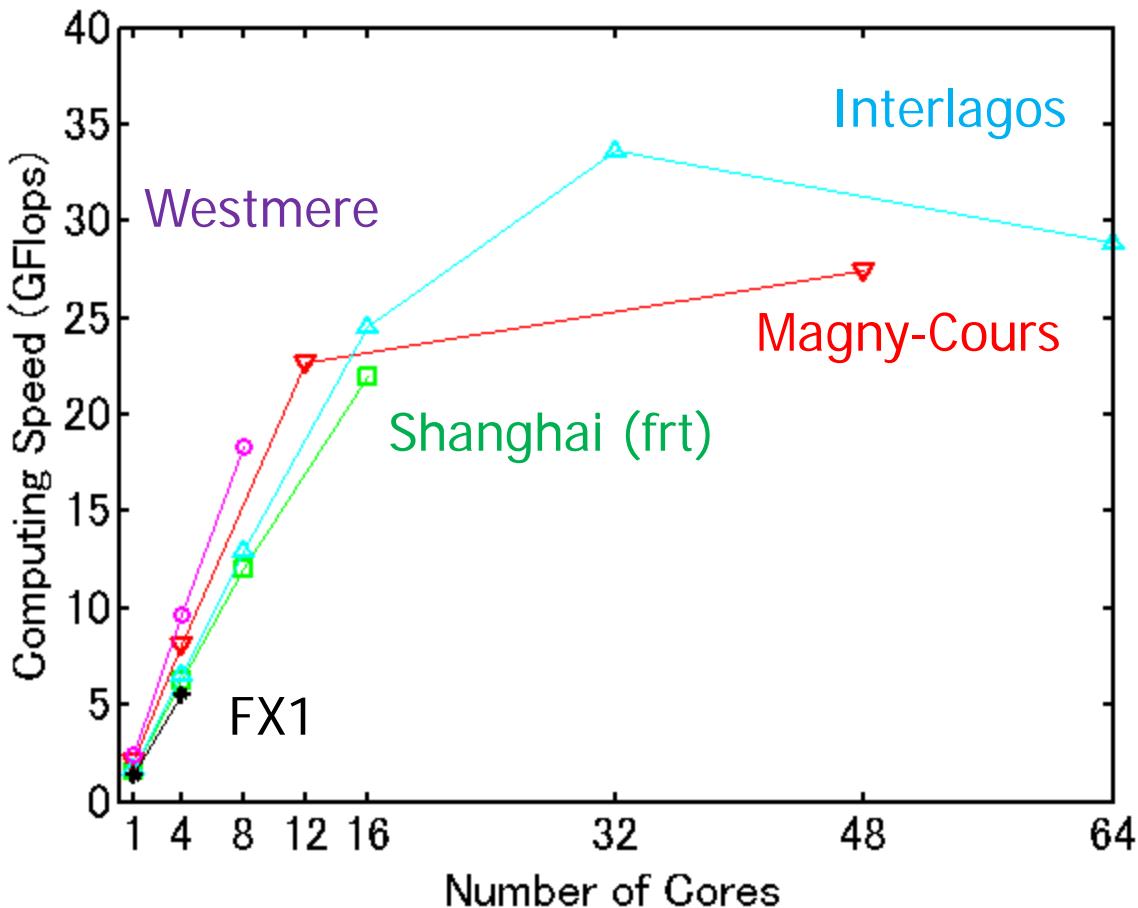
$$\frac{dv}{dt} = \frac{q}{m} (E + v \times B)$$



- Bulldozerの効率が悪い
- 旧世代Opteronが得意な演算？
- 富士通コンパイラが得意な演算

# 粒子コード(電流密度部) 実行性能

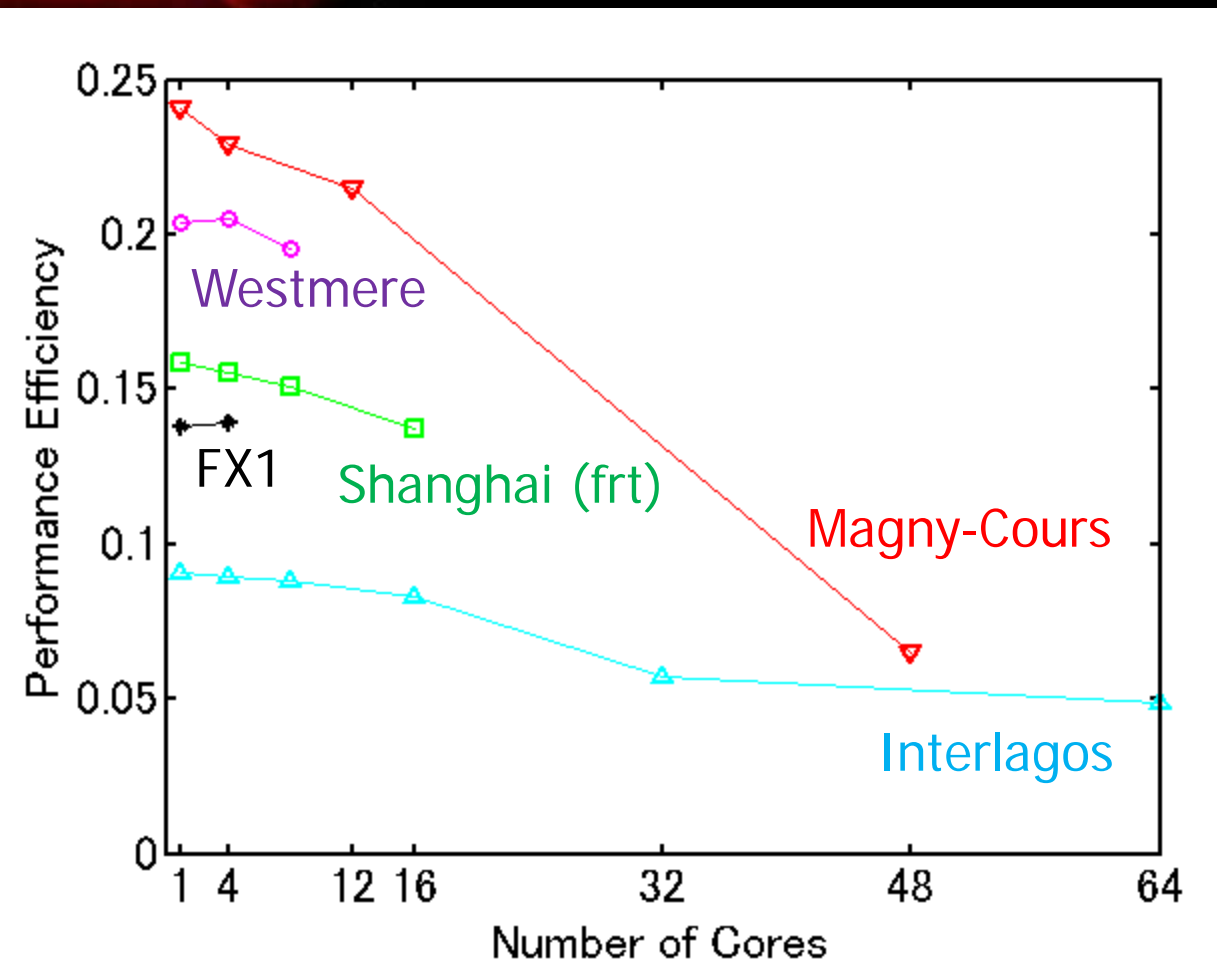
$$J = \sum qv$$



- 256粒子/cellの強いスケーリング
- Xeonが高速、Bulldozerは遅い
- 富士通コンパイラは苦手な演算
- 16コアまではスケール

# 粒子コード(電流密度部) 実行効率

$$J = \sum qv$$



- 256粒子/cellの強いスケーリング
- 旧Opteron+ifortが非常に高速
- 富士通コンパイラは苦手な演算
- Xeonも高効率、Bulldozerは遅い



# まとめ＋雑感

- 2011/11のバージョンアップによって性能が多少改善。
- コンパイラが得意・苦手な演算がある。
- 8演算/サイクルのコア用にチューニングする必要あり。
  
- 1024並列はほぼ動かず。256並列もHX600ではほぼ動かなかった。
- FX1の256並列は2日以内に動作。約20日×3ヶ月でCPU時間を使いきった。⇒使用料が高い...
  
- 10,000コア以上が定常的に使える環境を希望
- できれば、ノード単体の性能も高いほうが良い