

# アプリケーション作業部会 スペック検討サブWG 中間報告

牧野淳一郎

東京工業大学理工学研究科

理学研究流動機構

合同作業部会 2011/10/15

# スペック検討サブWG 中間報告

## 概要

- 各分野のアプリケーションの特性について予備的な検討をおこなった
- それに基づいて、アーキテクチャのイメージとして4通り想定した
- アーキテクチャのイメージをベースに、アプリケーションの特性を再検討した
- アプリケーションとアーキテクチャの対応について今日は報告する
- 但し、アプリケーションの特性の評価はかなり予備的なものである

# 予備検討の方針(1)

- アーキテクチャ部会での検討では、B/F、メモリ量、ネットワークバンド幅等について非常に狭い範囲しか想定していないように思われた
- 消費電力当り性能は、エクサスケール実現にとって大きな壁である。
- B/F はアプリケーションの効率に大きく影響する一方、消費電力当り性能にも大きな影響をもつ
- メモリ量、ネットワークバンド幅も、大きく変えれば電力に影響する

アプリケーション側で、B/F、メモリ量、ネットワークバンド幅の必要量をだしておこう

# What I learned from Steve Jobs

— Guy Kawasaki

1. **Experts are clueless**
2. **Customers cannot tell you what they need**
3. Jump to the next curve
4. The biggest challenges beget best work
5. Design counts
6. Changing your mind is a sign of intelligence
7. "Value" is different from "price"
8. A players hire A+ players

(いくつか省略)

## 予備検討の方針(2)

- アプリケーション、アルゴリズムにより B/F 等への要求は変わる
- 特に、同じアプリケーションでもアルゴリズムが変われば、またアルゴリズムが同じでも系のサイズ等だけによっても要求は変わる

といった問題があるので、

- 各分野に、重要なアプリケーション(計算法、系のサイズ等含めて)を選定してもらい、それぞれについて要求を見積もってもらう
- それらをいくつかのタイプに分類できるかどうか検討する

という方針を考えた。

# テンプレート

- タイトル
- 担当者名、メールアドレス
- アプリケーション、アルゴリズムに関する説明 (自由形式)
- 想定する単一ノード演算性能 (デフォルト 100Tflops)
- ネットワークバンド幅、レイテンシ
- メモリ容量、バンド幅
- オンチップメモリ容量
- ストレージ容量、速度
- 補足説明 (これはスライドでも文章でも)

# 記入例

- タイトル宇宙粒子系シミュレーション
- 担当者名、メールアドレス牧野淳一郎 makino@geo.titech.ac.jp
- アプリケーション、アルゴリズムに関する説明: ダークハローから星形成にいたる様々な階層での重力による構造形成シミュレーション。粒子法で、ガス・ダークマター・星等の重力相互作用の計算がメインである。計算方法はバーンズハットツリーないし高速多重極法。粒子数は  $10^{10}$  から  $10^{15}$ 、粒子数\*ステップ数で  $10^{19}$  程度

# 記入例 (続き)

- 想定する単一ノード演算性能 100Tflops
- ネットワークバンド幅、レイテンシ 30GB/s, 1  $\mu$  s
- メモリ容量、バンド幅 100GB, 60GB/s
- オンチップメモリ容量 100MB
- ストレージ容量、速度 1モデルあたり 10 – 100EB 程度使いたい。書き込み速度 1-10PB/s
- 補足説明 (省略)

# アプリケーション

## 38 アプリケーション

分野	数
1	7
2	13
3	4
4	8
5	7

# 分野 1

---

分子動力学計算

1 分子粒度生化学反応拡散系シミュレーション

細胞、細胞間生化学反応・拡散系シミュレーション

循環器系連続体力学・生化学反応計算

次世代シーケンサ解析プログラム

人間の全脳規模に相当する  $10^{11}$  の神経細胞からなる

リアリスティックな神経回路モデル

---

# 分野2

---

第一原理分子動力学法 (実空間基底)

第一原理分子動力学法 (平面波基底)

phase による電子状態計算 (+ MD 計算)

実時間密度汎関数法

超高精度電子状態計算

短距離古典分子動力学シミュレーション

長距離古典分子動力学シミュレーション

ナノ構造体電子・電磁場ダイナミクス法

クラスターアルゴリズム量子モンテカルロ法

feram によるリラクサー強誘電体の誘電率の

周波数依存性の分子動力学計算 (アレイジョブ)

CLUPAN による相図の計算 (アレイジョブ)

phonopy による熱伝導率などの第一原理計算

厳密対角化

---

# 分野3

---

地震動計算

3D地球モデルに対する理論地震波形計算

地球環境モデルMIROC

全球雲解像大気大循環モデルNICAM

---

# 分野 4

---

核融合プラズマ流体解析 GT5D

経路積分分子動力学 PIMD

FMO 計算

流体解析 VCAD

可視化

構造解析 ADVENTURE

流体解析 FFB

流体解析 UPACS

---

# 分野5

---

格子 QCD 計算 オーバーラップ型クォーク

格子 QCD 計算 クローバー型クォーク

核分裂現象シミュレーション

軽い原子核の第一原理計算

粒子系シミュレーション (大粒子数)

粒子系シミュレーション (小粒子数)

宇宙流体シミュレーション

---

# 検討結果

(詳しい資料は別に)

- B/F とネットワークバンド幅は関係あり。
- B/F 要求高いがネットワークは弱くていいものはある。逆はない
- ランダムアクセス、非数値計算等、この軸ではよく表現できない要求もある

とりあえず、 B/F とメモリ量にしばって結果を紹介する

# 分野1

(各アプリケーションの回答は公開用資料では削除)

# タイプわけの観点

## 観察:

- B/F は 0.1 以上の高いものと、桁で小さいものに分かれる
- メモリ要求にも非常に幅がある

## 注意事項:

- 分野によってはまだ十分な検討が進んでいない
- 分野によっては、そもそもアプリケーション・アルゴリズムの進化が速いために定量的な要求を明確にしにくいところもある

# タイプわけの観点

「アプリケーションタイプ」でなくて「アーキテクチャタイプ」としてみた。

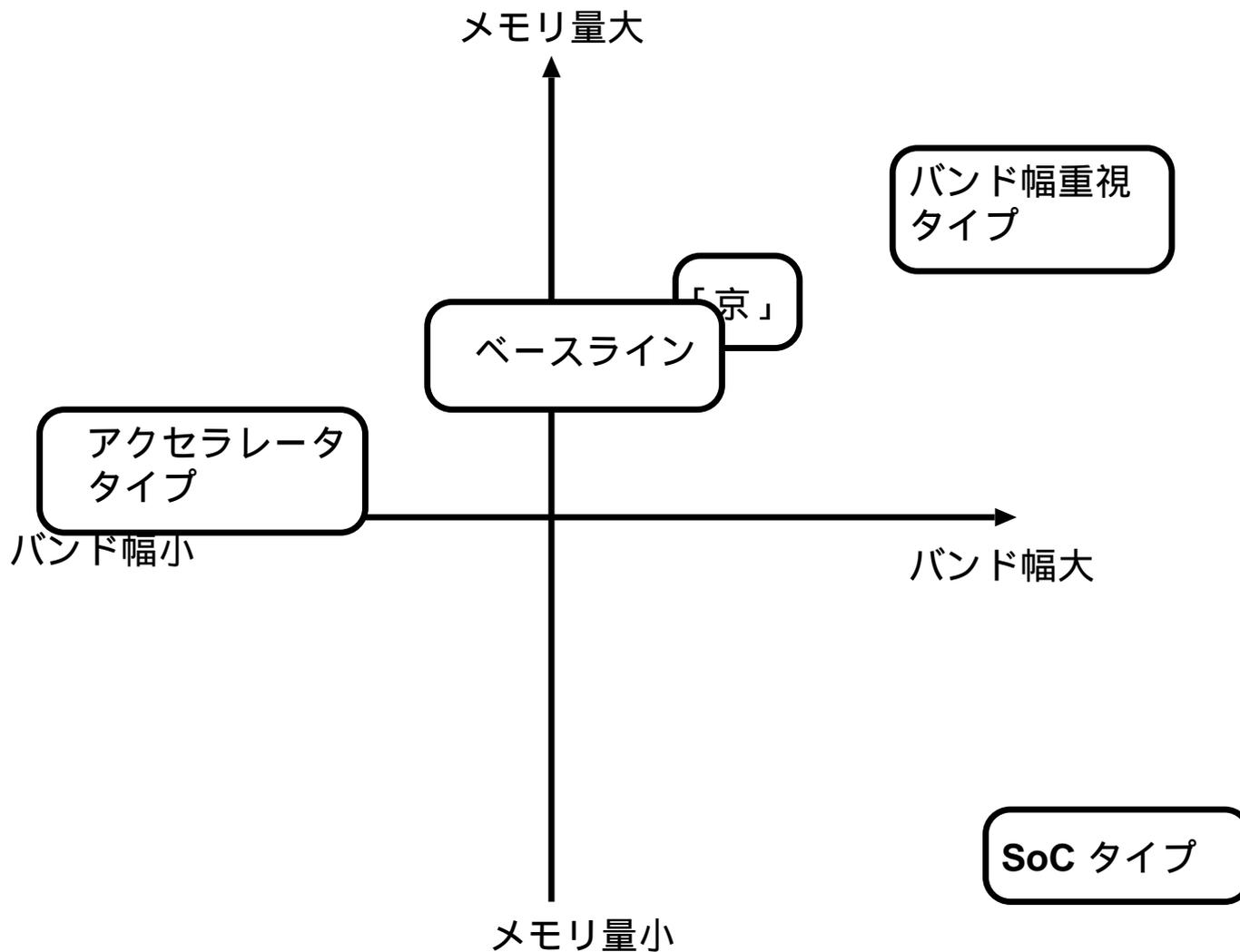
- そのほうが物理的制約をイメージしやすい
- アーキテクチャ側との議論もしやすい？

# タイプわけ

以下の4タイプ

タイプ	B/F	メモリ量 (1TF)	消費電力 (1EF)	演算性能) (20MW)	バンド幅 (20MW)
ベースライン	0.1	10-100GB	20MW	1EF	0.1EB/s
SoC	4	5-10MB	2-5MW	4-10EF	16-40EB/s
アクセラレータ	0.001	1-10GB	4-10MW	2-5EF	2-5PB/s
バンド幅重視	1	1TB	120MW	0.15EF	0.15EB/s

# こんな感じ



# ベースライン

- 「京」の延長、スカラ並列+SIMD
- アーキテクチャ部会ででていた絵を想定
- B/F 0.1
- メモリは 1TF あたり 100GB 程度以下

# SoC

- 外付けメモリがないか、極端に低バンド幅
- 軽量コアを非常に多数集積して、電力当り性能を上げる
- オンチップメモリに対しては 高B/F
- メモリはチップ当り 1GB 以下程度
- ネットワークは 100GB/s 程度

# 電力効率10倍は本当？

- 書いてある値: 2-5pJ/演算
- GRAPE-DR: TSMC 90nm プロセス。static leak かなり大きい
- チップだけだと 200Gflops, 50W
- 演算器は大体半分 (レジスタファイルの電力消費が異様に大きいため)
- 演算器部分だけだと大体 100pJ/演算
- 電圧スケールリングが全くないとすると 11nm で 10pJ
- なにかしら電圧スケールリング的なものがあと数倍あればいい

# アクセラレータ

- SoC タイプのチップをベースラインシステムのアクセラレータにするイメージ
- アクセラレータとホストCPUの間の転送速度は 100GB/s 程度

# バンド幅重視

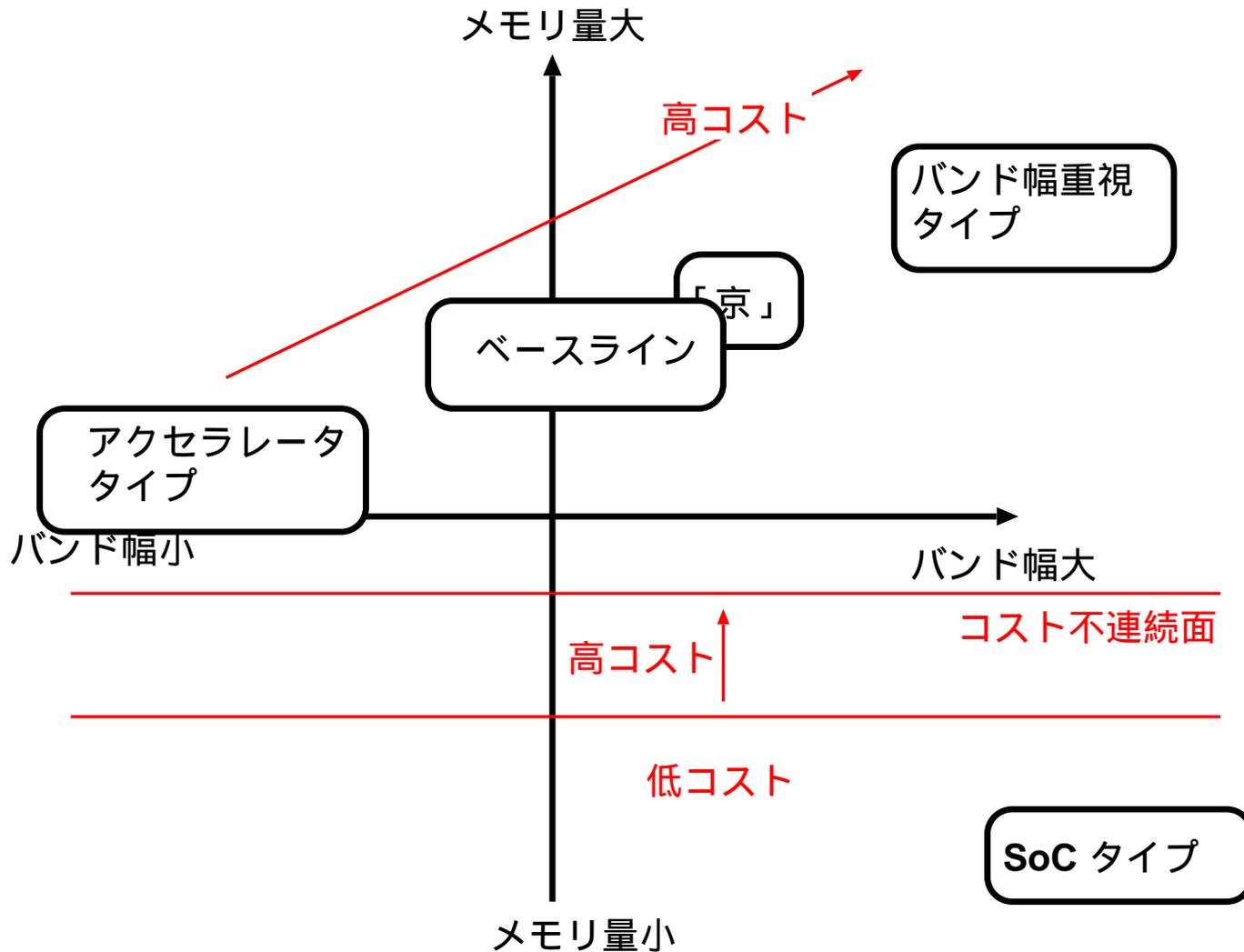
- とにかくがんばって B/F を1程度まで上げる
- 必然的に電力当り性能は下がる
- メモリバンド幅当り電力は Memory Cube の値を外挿した。

# タイプわけ

以下の4タイプ

タイプ	B/F	メモリ量 (1TF)	消費電力 (1EF)	演算性能) (20MW)	バンド幅 (20MW)
ベースライン	0.1	10-100GB	20MW	1EF	0.1EB/s
SoC	4	5-10MB	2-5MW	4-10EF	16-40EB/s
アクセラレータ	0.001	1-10GB	4-10MW	2-5EF	2-5PB/s
バンド幅重視	1	1TB	120MW	0.15EF	0.15EB/s

# 電力その他も考えると



# アプリケーション分類結果

各分野からの提出の通り

タイプ	数
ベースライン	12
SoCタイプ	4
アクセラレータタイプ	5
バンド幅重視タイプ	17

# アプリケーション分類結果

いくつか牧野が勝手に修正したもの(但し、かなり確実にこれでよい、というものののみ変更した)

タイプ	数
ベースライン	8
SoC タイプ	5
アクセラレータタイプ	10
バンド幅重視タイプ	15

# バンド幅必要なアプリケーション

- 圧縮性流体、地震波動、構造解析等の有限要素法ないし差分法
- 分野2の多くのアプリケーション。FFTを使うため？別の理由？

「別の理由」の一部: 「プログラミングの容易さ」

- サイエンスを進める上では重要
- 国際競争等を考えた時にどう考えるべきか？

# まとめ

- ベースライン、SoC、アクセラレータ、バンド幅重視の4アーキテクチャを想定し、ある程度定量的にアプリケーションを分類した
- この4タイプがほぼ 3:1:2:4 の比率程度になった
- 但し、検討が不十分の時にはバンド幅要求は高くなる傾向があるであろう
- 電力当り性能では2桁の幅があることも考慮し、サイエンスターゲットもかんがみながらアーキテクチャの「ベストミックス」を考えるべきではないか？

# 今後の検討課題等

- これらの数字の精度。もっと高い性能が実現できる、あるいは低い性能しか実現できないということはないか。
- 4種類必要かどうか
- アプリケーションの実行性能評価

# もっと色々な検討課題

- 単精度で2倍、可能なら4倍速いようなアーキテクチャに対する要求は大きい。また、さらに低い精度についても要求がある。
- 4倍精度についても、極端に遅くならないようなハードウェアサポートがあることが望ましい。これは必ずしもIEEE標準の112ビット仮数にこだわるものではない。
- コアの機能、特に整数演算性能、分岐等の性能
- ネットワークの機能。同期・縮約等
- I/O
- プログラミングモデル
- Capacity computing への対応
- その他