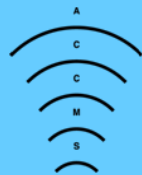


戦略的高性能計算システム開発に 関するワークショップ

京都大学 学術情報メディアセンター
中島 浩

<http://www.open-supercomputer.org/workshop/>



Facts (1)

■ 趣旨

100P-1EFlops のシステム仕様と、そのための（主にシステムS/Wに関する）研究開発RMを、（主に若手の）研究者により議論する

■ 開催

- #1@**金沢** ('10.8.2) : **所信表明**
- #2@**秋葉原** ('10.11.27) : **所信表明+シニアパネル**
- #3@**京都** ('11.2.5) : **テーマ別パネル**
- #4@**秋葉原** ('11.5.28) : RM kickoff
- #5@**鹿児島** ('11.7.26) : RM **の議論**

■ 主催

- **スパコンセンター@筑波大・東大・東工大・京大**
- **産総研 情報研究部門** (#2, #4~)
- **理研 情報基盤センター** (#3~)

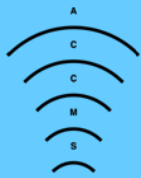


Facts (2)

■ 参加者

	発表者				その他参加者				合計			
	ac	or	co	計	ac	or	co	計	ac	or	co	計
#1	19	4	6	29	11	1	4	16	30	5	10	45
#2	17	3	2	22	16	3	4	23	33	6	6	45
#3	16	3	1	20	22	5	2	29	38	8	3	49
#4	17	2	3	22	22	10	3	35	39	12	6	57

- #4の発表者はRM作成ボランティアの若手研究者



所信表明(#1 & #2)のポイント (1)

アプリ & ライブラリ (8件)

■ アプリ屋の問題意識

- システム革新に伴う性能ポータビリティ
→ アプリケーションへの自動チューニング組み込み
- I/O & post proc. の共通基盤技術

■ ライブラリ屋の問題意識

- アプリへの適合技術 = ライブラリ利用によるメンテ最小化
- システムへの適合技術 (性能, 精度, 電力, 耐故障, ...)
→ ライブラリのミドルウェア / システムソフトウェア化
- 問題規模 & 計算規模の爆発的増大に伴う計算精度問題



所信表明(#1 & #2)のポイント (2)

プログラミング言語 (9件)

■ HPC的言語屋の問題意識

■ システム革新に伴う記述複雑化

(MPI/PGAS + OpenMP + CUDA + ...)

■ 非HPCプログラミング(Java, C++, Ruby, Haskell, Go, ...) との乖離 (Fortran/C + MPI + ...)

■ 非HPC的言語屋の問題意識

■ 高抽象・高水準言語の復権／HPC適用のチャンス

← HPCアプリ記述の複雑化

← 非HPCアプリ／非HPC的処理 in HPCアプリの増加

■ 本当に移植性・生産性は上がるのか？

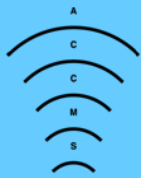
■ 理論／処理系／最適化はシステム革新に追隨できるか？



所信表明(#1 & #2)のポイント (3)

ミドルウェア (7件) + システムソフトウェア (10件) [1]

- **(広域) Data Intensive 関連の問題意識**
 - 計算集約だけでなくデータ集約への対応の必要性
 - データ集約に適合した技術開発 (データ / 計算配置、ネットワーク帯域確保、スケジューリング、...)
- **ストレージ関連の問題意識**
 - ストレージシステム多様化 (SSD, 広域分散ファイル)
 - 大規模化に伴う一貫性維持のボトルネック化
 - アプリに応じた構成 / アクセス最適化 / 管理の必要性
 - 並列 I/O の機能強化 & 普及

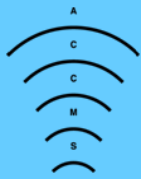


所信表明(#1 & #2)のポイント (4)

ミドルウェア (7件) + システムソフトウェア (10件) [2]

■ その他諸々の問題意識

- **耐故障:** アプリに適合した耐故障戦略
故障の予兆検出 & 回避
- **通信:** アプリに適合した networking
- **OS・RT:** 軽量化・jitterless 化
ヘテロシステムへの対応
非HPC的(データ集約的)アプリ対応
- **モニタ:** 膨大な稼動・性能情報からの有益情報の抽出
アプリに適合したシステムパラメータチューニング



所信表明(#1 & #2)のポイント (5)

ハードウェア&アーキテクチャ (12件)

■ メモリに関する問題意識

- 一定のバンド幅確保にはプロセッサ／メモリ積層しかない
(あるいは容量／バンド幅トレードオフ)
- 狭バンド幅／小メモリ容量の許容へのパラダイムシフト？
- 新メモリ階層？ メモリをバイパスする結合網？

■ 結合網に関する問題意識

- 光技術 (&プロセッサ／スイッチ積層) ？
- メモリをバイパスするレジスタ／キャッシュ間通信？

■ 消費電力に関する問題意識

- ハードウェア (e.g. FPGA) への問題マッピング？
- 細粒度消費電力制御？



シニアパネル@#2のポイント

- 朴(M), 泰地, 中島, 中村, 平木, 松岡
- テーマ
 - 100P/1E システムの big picture
 - 技術的挑戦と解決への道筋
 - 日本の役割 & 舞台
- 主な論点
 - システム&アプリの定量化・モデリングを現段階から
 - メモリ・結合網のバンド幅低下は所与としての挑戦的工夫
 - システムソフト(&アプリ)との co-design は必然
 - many coreのチップ内低遅延によるパラダイムシフト?



アプリ&ライブラリのポイント

- **中島(M), 片桐, 高橋, 丸山, 村主, 矢作**
- **テーマ**
 - ExaScale へのマイルストーン
 - Co-Design と自分の役割
- **主な論点**
 - **垂直方向の研究開発のためのインタフェース定義**
 - domain specific な並列計算記述・実装インタフェース
 - 100P/1E specific な機能(省電力, 耐故障, 動的最適化, ...)のためのAPI
 - **ライブラリの課題:ヘテロ, 省電力, 精度, 耐故障, ...**
 - **(過度な) 汎用性からの脱却 in ハード & システムソフト**



言語 & ミドルウェアのポイント

- 佐藤(M), 近藤, 滝沢, 千葉, 姫野, 平木, 八杉
- テーマ
 - 言語 & Prog. Model はアーキテクチャを牽引できるか
 - 日本で実施すべき研究開発は何か
- 主な論点
 - ハードの隠蔽 (incl. 省電力, 耐故障, ...) は必然 & 好機
 - What & How 記述の分離も必然
 - domain specific は必然(やむを得ない)
 - DSL/Framework 構築(半)自動化はCSでは常識化
 - 挑戦的アプリなら言語・モデルの革新を許容



システムソフト&アーキテクチャのポイント

- 松岡(M), 清水, 須田, 泰地, 建部, 中村, 堀
- テーマ
 - 課題 & 研究開発アプローチ
 - マイルストーン/ロードマップ
- 主な論点
 - scale-out のための革新的(通信)技術
 - ソフトウェア可視の新メモリ階層導入
 - メモリバイパス通信機構の導入
 - これら (& others) の継続的・統一的な提示が重要
 - 低レベル(ハードウェア可視)~高レベル(for app.)のプログラミング階層化
 - 諸々の新機構を含む性能モニタリング&チューニング



まとめ (所感): 期待以上の熱気に加えて...

- **挑戦的な課題の共有**
 - **並列性の爆発的増大への対応**
 - **negative issues への対応 & その隠蔽**
 - **メモリ・通信バンド幅低下(&遅延増加)による複雑化(階層深化)**
 - **非均質性, 省電力, 耐故障, ...**
- **application awareness & co-design**
 - **頻出キーワード=「アプリに適合した〇〇〇」**
 - **app. aware の意識(or 必要性)は確実に向上**
 - **co-design のための liaison CS屋の必要性**
- **general vs specific**
 - **blindly general purpose という意見は聞こえず**
 - **app. specific ではなく (algorithm) domain specific であることが重要**
 - **ハードよりむしろソフトウェアスタックの specificity**