

次世代メニーコア型スパコンのための システムソフトウェア

堀 敦史

システムソフトウェア技術部会 部会長（理研 AICS）

清水 正明

システムソフトウェア技術部会 副部会長（日立）

2019/12/13



- IHK/McKernel
 - 最新情報
- Process-in-Process (PiP)
 - 新しい並列実行モデル

McKernel

- **マルチ OS カーネル**

- 軽量カーネルと（重量）Linux カーネルのコンビネーション
- （OSサービスをあまり必要としない）並列アプリは McKernel 上で動作
- 頻度の低いシステムコールは Linux カーネルに委託

- **特長**

- ノイズが非常に少ない
- Linux に比べ機能追加が容易

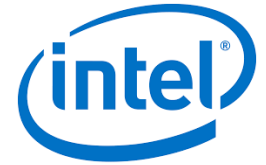
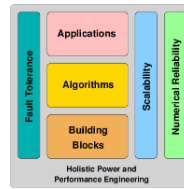
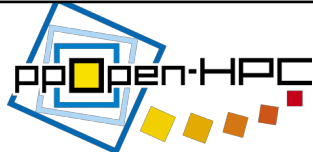


Process-in-Process (PiP)

- マルチプロセスモデル (MPI) とマルチスレッドモデル (OpenMP) の
いいとこ取り
- **特長**
 - 容易なプログラミング (共有メモリとは似て非)
 - 高性能はノード内通信
 - In-situ などへの応用の可能性

McKernel と PiP の開発状況

- 富岳への導入を目指し
 - Aarch64 へのポーティング（ほぼ終了）
 - 品質向上（x86_64、aarch64）



Parallel Multigrid Methods on Manycore Clusters with IHK/McKernel

Kengo Nakajima^{*1,*2}, Balazs Gerofi^{*2}, Yutaka Ishikawa^{*2}, Masashi Horikoshi^{*3}

***1: Information Technology Center, University of Tokyo, *2: RIKEN R-CCS, *3: Intel**

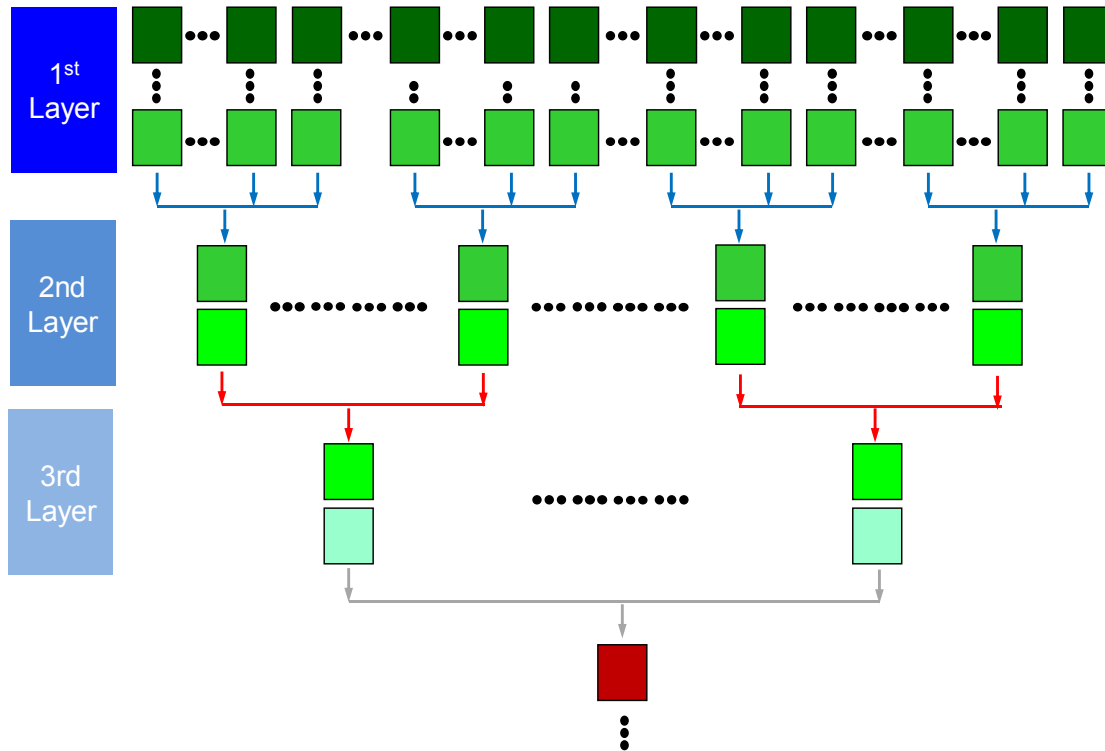
185th R-CCS Café, December 2, 2019

based on the presentation at:

ScalA19: 10th Workshop on Latest Advances in Scalable Algorithms for Large-Scale Systems in conjunction with SC19, November 18, 2019, Denver, CO



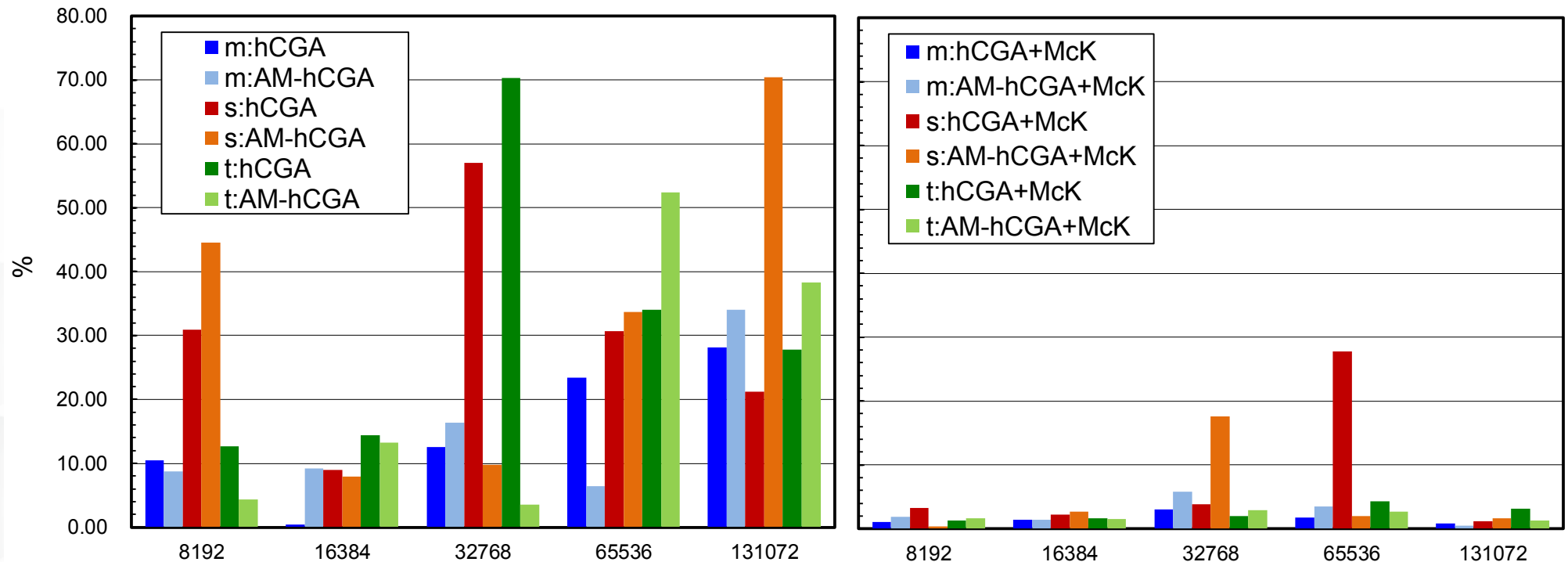
Proposed Method: AM-*h*CGA Adaptive Multilevel *h*CGA



- If the number of MPI processes is $O(10^4)$, *h*CGA is effective
- If the number of MPI processes is $O(10^6-10^7)$, number of processes at the 2nd level of *h*CGA could be $O(10^4)$.
 - 2-Layers might not be enough for more processes
 - More levels are needed ?
- AM-*h*CGA
 - 3-Layers in this work

McKernel の応用例 3/4

- もともと実行時間のバラツキが大きい傾向があった（問題サイズがtiny）
- Linux 上では提案方式の有効性が確認できず
- McKernel上でははっきりと有効性が確認できた（OFP上の評価）



AM-hCGA: Effects of IHK/McKernel

Computation Time for hCGA= 100%, ■ Communication, ■ Others

More significant reduction of comm. time by IHK/McKernel in “Tiny”

