

高性能並列計算システム はどうあるべきか

鴨志田良和 (東京大学)

kamo@cc.u-tokyo.ac.jp

5年後にセンター運用可能な高性能並列計算機システムは、どういうシステム仕様が考えられ、そのために、今後どういう研究開発をしていくべきなのか

日々やるべきこと、将来をみること

- 現在のスパコン運用において改善すべき点を地道に拾って
いって、真面目に解決していく
 - 不便なことに慣らされすぎてはいないか
 - 初心忘るべからず
- 「10年で1000倍」にとらわれすぎるべきでない
 - 目標は必要
 - 本当に使えるものはなにか考える
 - 最善の選択は常に可能ではない
 - 性能上の問題を顕在化し、原因を調べることができる仕組みが必要

日々直面する問題

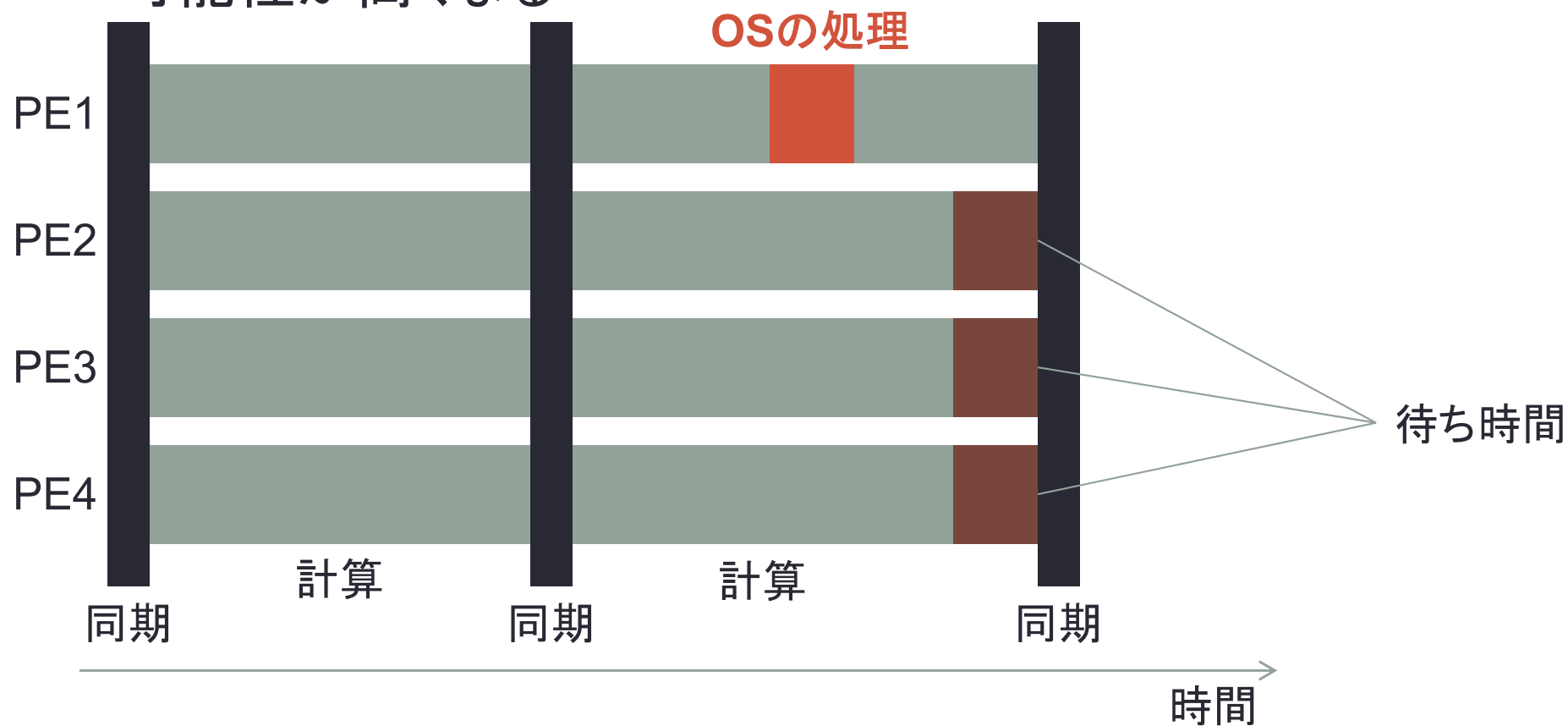
- バッチ？インタラクティブ？
- 遅いファイルシステム
 - システムが大規模化するとより深刻
- 閉じた世界
 - 外の環境と同時に使うことが想定されていない
 - 無人アカウントと利用資格
- 外国の人が使いにくい
 - 国際的な共同研究の妨げになっているのでは？
 - 日本居住の外国人、海外居住の日本人
- コンパイラ
 - ハードウェア提供メーカー製コンパイラでは、Mpich, OpenMPI(動的リンク版)のコンパイルすらできないことがある
 - たとえば、コマンドラインオプションをGCC等に合わせるべきでは？

行うべき研究開発

- 性能上の問題の顕在化と、原因調査のしくみ
 - 「ノード数が増えるとアプリケーションがスケールしない」問題を解決しやすくしたい
 - 大規模にならないと見えてこない問題
 - OS jitter問題(次ページ)など
 - 計算機的设计時点では見えていない問題もあるかもしれない
 - システムの問題、アプリケーションの問題
- **モニタリング→データ解析→可視化をスケーラブルに行うシステムが必要**

OS Jitter問題

- 同期しながら進んでいく処理は、ひとりの仕事に邪魔が入ると他の人も待たされる
- システム大規模化するほど、だれか一人の仕事に邪魔が入る可能性が高くなる



研究のテーマ

- プロセスごとの統計情報、CPUの性能カウンタや、ネットワークデバイスのカウンタ等の情報を時々刻々収集・集約
 - システムへの負荷をなるべく小さくして収集したい
 - ノード数が100万なら、およそ数GB/秒のデータ処理
- 大量の情報から重要なものを抽出し、提示する
- オンライン・リアルタイムで扱うことができる性能を実現する

まとめ

- 高性能並列計算システムはどうあるべきか？
 - 性能が予測とどのように異なっているかが、見えるようなシステム
- 本当に使える高性能並列計算システムを作るためには
 - システムのハード・ソフトからアプリケーションまで、幅広い知識が必要
 - 運用やユーザビリティからの視点も必要
 - 専門家同士の協力必要
- システムモニタリング
 - データの羅列から問題解決に至るまでにはやはり横断的な知識と分析能力が必要
 - 自分も、横断的な知識をより厚みのあるものにして、高性能並列計算の分野でリーダーシップを発揮していけるようになりたい