

黒川 原佳 (理化学研究所)

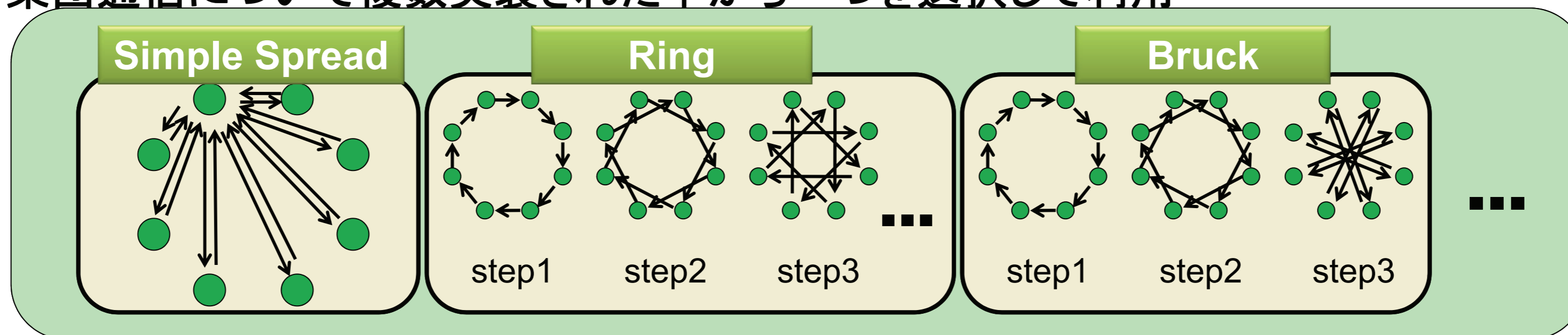
負荷バランスや通信性能が予測困難な状況を想定した 集団通信アルゴリズムの動的選択技術に関する研究



研究の目的: 実行時の状況に応じたアルゴリズム選択による集団通信の高速化

■ 集団通信アルゴリズム:

- 各集団通信について複数実装された中から一つを選択して利用



- 従来: メッセージサイズ、プロセス数に応じて選択。閾値は静的に設定。

⇒ **実行時にアルゴリズムの優劣が変動する環境やプログラムに対応できない**

■ 実行時にアルゴリズムの優劣が変動する要因:

■ 通信性能の変動

例) RICC(Riken Integrated Cluster of Clusters)

- 任意の計算ノードにジョブを割り当てることにより、ジョブの充填率を最大化
- ジョブのランク配置によって通信衝突の頻度が変化

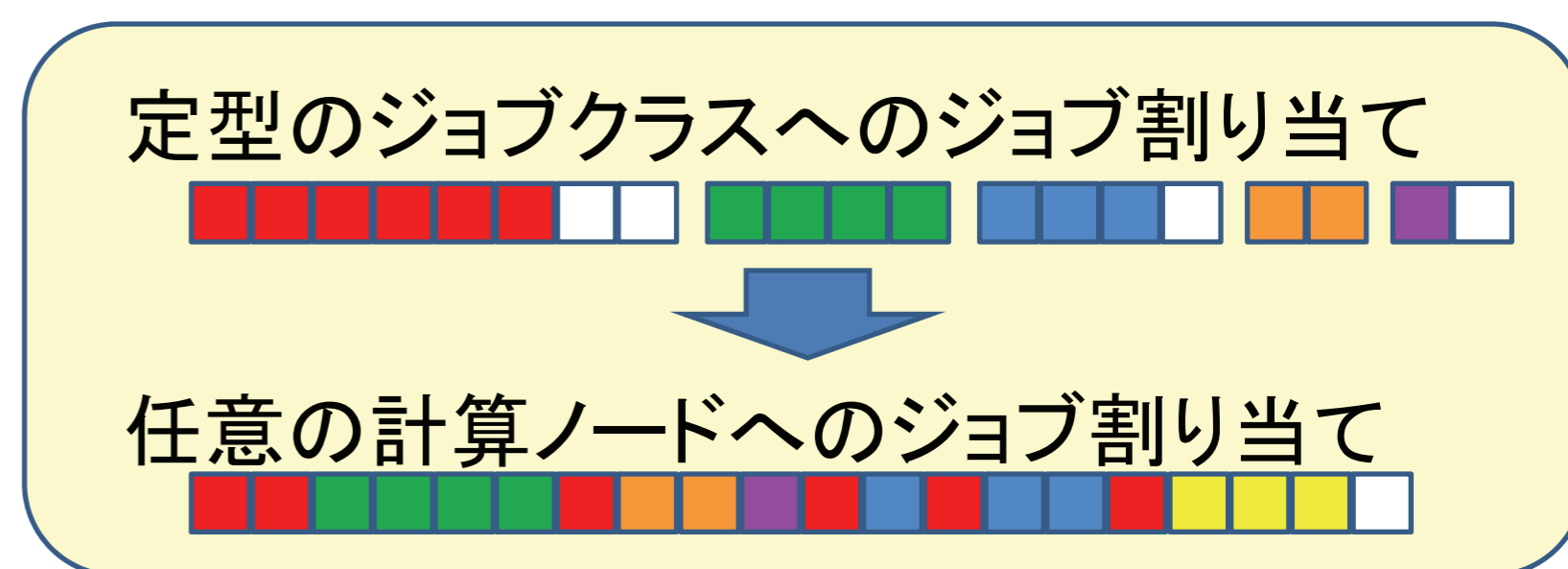
⇒ **最適アルゴリズムがジョブ毎に変動**

■ プログラムの負荷バランス

例) 疎行列による連立一次方程式の反復解法

- 入力データ(非零要素の分布)によって疎行列ベクトル積の負荷バランスが変化
- プロセス毎に集団通信の開始時間が変化

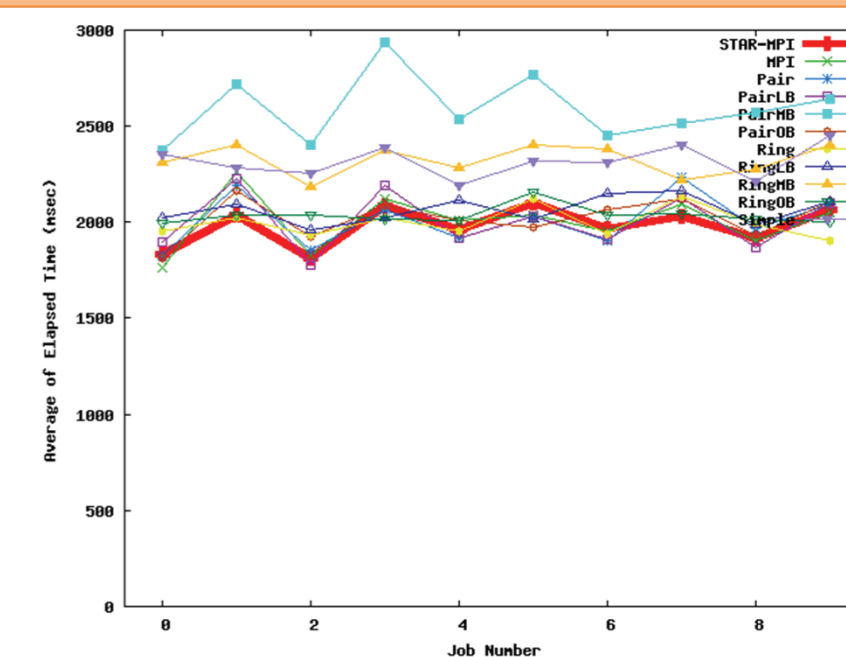
⇒ **最適アルゴリズムが入力データによって変動**



昨年度の成果: ジョブ毎に通信性能が変動する環境での動的アルゴリズム選択の効果検証

- 毎回、アルゴリズムの優劣関係が変動する中で、動的アルゴリズム選択技術によって常に最速のアルゴリズムと同程度の速度を維持

- 性能予測モデルを用いて選択対象のアルゴリズムを事前に絞り込むことによるオーバヘッド削減手法を提案



今年度の計画: 通信性能の変動, 負荷バランスそれぞれに応じたアルゴリズム選択

■ 通信性能に応じたアルゴリズム選択

- ノード配置による通信性能の変化の解析 ⇒ **より精度の高いアルゴリズム絞り込み**

■ 負荷バランスへの対応

■ 負荷不均衡の検知

- 最初の呼び出し時に隣接する番号のプロセスとping-pong-ping通信 ⇒ 待ち時間が異常に長い場合、「負荷不均衡情報」としてマスタープロセスに通知

■ アルゴリズムの絞り込み

- 「負荷不均衡情報」の数と待ち時間から負荷不均衡の度合いを判断し、**必要に応じて負荷不均衡に"弱い"アルゴリズムを除外**

■ 通信性能の変動と負荷バランスの双方を加味したアルゴリズム選択技術

■ 実アプリケーションによる効果の検証

- 量子化学計算プログラムOpenFMO等