

# 隣接通信に対する通信衝突を考慮した 通信性能向上のためのタスク配置最適化の評価

森江 善之†‡, 南里 豪志†‡

†九州大学情報基盤研究開発センター

‡独立行政法人科学技術振興機構, CREST

{morie.yoshiyuki.404@m., nanri@cc}.kyushu-u.ac.jp

## 1 はじめに

近年の並列計算機の大規模化は著しく、スーパーコンピュータの性能評価を行う TOP500 において 1 位となった「京」コンピュータは 8 万個もの計算ノードを持つ。このように大規模化した並列計算機においては、そのネットワークをクロスバ網で構成することはハードウェアコストの問題で困難である。このため、メッシュ/トーラスやファットツリーと言ったネットワークポロジが用いられる。しかし、これらのネットワークポロジでは、各計算ノード間において経路を共有するリンクがあるため、通信衝突が発生し、通信性能が悪化する。現在、計算機センター等においてこのような大規模並列計算機の導入が進められている。このような計算機を利用する場合にはその一部分の利用であってもそのネットワークポロジの影響を受け通信衝突が起こりうるため、対処が必要となる。著者らはネットワークポロジに起因する通信衝突を考慮し、通信性能を向上させる研究を行っている [1]。本稿では、隣接通信の通信パターンに対する通信衝突を考慮した通信性能向上のためのタスク配置最適化の効果の調査を行った。

## 2 通信衝突を考慮した通信性能向上のためのタスク配置最適化

通信衝突の通信性能への影響を考慮するためには、通信衝突の発生を検出する必要がある。通信衝突は、複数のメッセージが同時に同一リンクの同一方向へ要求がなされた場合に発生する。このため、本タスク最適化では、同時にメッセージの転送が開始される通信の集合を与える。この集合を Concurrent Communication Set (CCS) と呼ぶ。そして、どのリンクで同一 CCS における通信が競合するかを調べる。これらにより通信衝突による通信性能への影響を加味できるようにした。以下、通信衝突を考慮した通信性能向上のためのタスク配置最適化を Task

Allocation with Checking Concurrent Communication (TACCC) と呼ぶこととする。

ここで、式 1 に TACCC の目的関数を示す。

$$\sum_{t=0}^{N-1} \max_{i=0 \dots n-1} M(t, i, \text{tork}(i)) \times \text{coll}(t, \pi(i), \pi(\text{tork}(i))) / B \quad (1)$$

$t$  は CCS の ID,  $N$  は CCS の総数である。 $i$  はタスクの ID,  $n$  はタスクの総数である。 $\pi(i)$  はタスク  $i$  が割り付けられている計算ノードの ID を返す関数である。 $\text{tork}(i)$  はタスク  $i$  の宛先のタスクの ID を返す関数である。 $B$  は各リンクの通信帯域幅である。 $M(t, i, j)$  は  $t$  番目の CCS におけるタスク  $i$  からタスク  $j$  への 1 つ通信のデータ量を返す関数である。また、 $\text{coll}(t, p, q)$  は  $t$  番目の CCS における計算ノード  $p, q$  間の経路の各リンクで要求される通信数の最大値を返す関数である。この目的関数の値を最小とするタスク配置を simulated annealing を用いて求め、適用することで通信性能向上を図る。

## 3 隣接通信の通信パターン

本稿では、数値計算の分野において適用範囲の広い隣接通信の通信パターンに対し TACCC の評価を行った。隣接通信において論理的なトポロジと物理的なトポロジが一致した場合は、全ての通信の通信距離が 1 ホップとなり、通信衝突による通信性能の悪化は発生しない。しかし、論理的なトポロジと物理的なトポロジが一致しない場合は複数ホップの通信が発生するため、通信衝突が発生する可能性がある。この点を考慮して評価を行った。

## 参考文献

- [1] 森江 善之, 末安 直樹, 松本 透, 南里 豪志, 石畑 宏明, 井上 弘士, 村上 和彰, “通信タイミングを考慮した衝突削減のための MPI ランク配置最適化技術,” 情報処理学会論文誌コンピューティングシステム, Vol.48, No.13, pp.192-202, Aug. 2007.