

## 審査の結果の要旨

氏名 松葉 浩也

近年の計算機環境は、並列数値計算からオフィスアプリケーションに至る様々な種類のアプリケーションが同一アーキテクチャをもつ計算機上で実行される傾向にある。これに伴い、同種の計算機を多数配置した計算機クラスターを、様々な用途に応じてクラスター内部を切り分けて利用する形態が一般化しつつある。このような計算機クラスターでは、同時のクラスターを使う利用者間の公平な計算資源の配分が重要視されている。ところがクラスターとクラスター外部を接続するネットワーク資源に関しては、その最適利用のために、各利用者の利用可能帯域を正確に設定しつつ、設定された帯域内で可能な限り高速で安定した通信を実現するような重要であるが、その有効な方式が確立されていない。

以上のような背景のもので、本論文では安定かつ正確な帯域でのクラスターからのネットワーク送信を目指して二つの方法を提案している。ひとつは **I/O Server Aggregation** と呼ぶ方法であり、本方式では外部との通信を中継する **I/O 専用ノード** の数を調整することで帯域制御を実現する。もうひとつは **Distributed Network Interface** と呼ばれる方法であり、**I/O サーバ** に帯域を絞った仮想ネットワークインタフェースを準備し、それを各計算専用ノードから利用するために利用者毎の独立したフロー制御を実現することのできる新たな通信プロトコルを用いる。本研究ではこれらの提案手法を代表的な通信パターンを模したベンチマークを用いて評価し、いずれの通信方法でも設定帯域を上回ることはない正確な帯域制御を実現し、既存手法の最大 5.4 倍となる高い性能を得ている。

本論文の第 1 章では、クラスターからクラスター外部への通信において一般的に高い性能を実現することだけでなく、各利用者が期待または設定した帯域が確実に得られることの重要性を指摘し、その実現のために提案する二つの方法、**I/O Server Aggregation** と **Distributed Network Interface** についての概要を紹介している。

第 2 章では、現在のクラスターアーキテクチャの紹介を行った後、クラスターから外部への通信について、その代表的な通信パターンとして、ファイル転送、クラスター間並列計算のための通信、および **HTTP** の応答通信を挙げている。これらの通信において安定かつ帯域保証された通信を実現するための問題として、**TCP** の性質による性能ロスとクラスターの各ノードが持つ外部通信用機器の性能限界を指摘している。

第 3 章では、本研究で提案する手法の土台として、外部との通信を中継する「**I/O サーバ**」と呼ばれるノードを導入する手法を紹介している。その上で **I/O サーバ** の数によって、クラスターからの総送信帯域の上限を規定する基本アイデアを提案している。また、提案手法の実現方法としてユーザーレベルの通信ライブラリを導入する場合とオペレーティングシステムカーネルに新たなドライバを組み込む場合の利害得失を議論し、双方のアプローチをもちいることを結論としている。

第 4 章では、ユーザーレベル通信ライブラリとして効率的かつ帯域保証された通信を実現する方法として **I/O Server Aggregation** を提案している。提案方法では **I/O サーバ** の数に上限を与えることで送信バンド幅の上限を定める。一方、各ノードが持つ外部通信用機器では性能が不足するため、クラスター内部向け高速インターコネクタを用いて単一メッセージを分割して複数の **I/O サーバ** に転送し、並列に **I/O サーバ** を用いることで広帯域通信を行う。

第 5 章では、安定かつ正確な帯域でのクラスターからのネットワーク送信を目指して二つ目の方法として、カーネルレベルでの目標実現手段として、**Distributed Network Interface** と呼ばれる方法を提案している。本方法では物理ネットワークインタフェースを時分割により帯域を絞った上で各ノードと接続し、帯域調整された送信を実現

する。特定の利用者が I/O サーバに大量にデータを送信して他の利用者の通信を阻害することがないように、新たなフロー制御プロトコルを提案し利用者毎の独立したフロー制御を提供している。

第 6 章では、本論文で提案した二方法を評価するために、バンド幅、MPI 並列計算、HTTP サーバの各種ベンチマークで性能評価を行っている。既存手法と比較して、バンド幅測定では最大 5.4 倍、並列計算では最大 125%、HTTP では 109%の性能を記録している。またいずれの場合においても指定帯域を超えないことを確認しており、提案方法が高い性能と帯域保証を両立させていることを示している。

第 7 章では、データセンター用の最新機器、性能向上のための通信プロキシ手法、Ethernet 上の特殊用途プロトコル、並列計算機用オペレーティングシステムおよび通信ライブラリなどの関連研究として紹介し、本研究で提案した手法との差異を議論している。

最終章（第 8 章）では、提案した二手法のまとめと残された課題について触れている。

以上まとめると、本論文では、クラスタ型計算機から安定かつ正確な帯域を保ちながらネットワーク送信するための二つ方法が提案されている。どちらの手法も初めて提案されたもので、その有効性が評価結果で確認されており実用性と学術性が高い。よって本論文は博士（情報理工学）の学位請求論文として合格と認められる。