

## 論文の内容の要旨

論文題目 A STUDY OF HIGH-QUALITY NETWORK TRANSMISSION ON  
COMPUTER CLUSTERS

(計算機クラスタにおける高品質ネットワーク送信に関する研究)

氏名 松葉 浩也

近年の計算機技術は単純化、大規模化、集中化の傾向にあり、コモディティハードウェアを用いた汎用計算機を大量に導入することで多様なニーズに応えるデータセンター型の計算機資源提供方法が主流になりつつある。商業用のデータセンターのみならず大学や研究所などの大型計算機においても汎用部品から成るスーパーコンピュータの導入が進んでおり、その汎用性を活用すべく、各種サーバのホスティングやデスクトップコンピュータの代替機能などの科学技術計算用途以外の新たな利用方法が模索されている。

大勢の利用者が多様なアプリケーションを同時に実行する計算機センターにおいては、資源の平等かつ正確な分配は必須の機能である。特にネットワーク資源は多くの利用者が、ファイル転送、Web サービス、遠隔操作、グリッドコンピューティングなどの多様なデータを一本のアップリンクを共有して転送するため、それらを高速かつ正確な帯域で扱うことはサービス品質上重要である。

本論文では計算機センターのネットワーク資源について、特に計算機センターとセンター外部の計算機が通信を行う状況において、高速かつ正確な帯域での通信を実現する手法を提案する。ユーザーごとの正確な帯域割り当てはバックボーンへの接続ルーターで行う方法や、各サーバにおいて出力帯域制限を設ける手法が考えられるが、本論文ではその中

間的な方法として「I/O サーバ」と呼ばれるネットワーク出力専用のサーバの設置を提案し、そのサーバにすべてのデータを集めて帯域制御を行う方法を提案する。この方式の利点は次の3点である。一点目は各サーバと I/O サーバの間をクラスタ用高速インターコネクタで接続することにより、データ混雑によるパケットロスのない安定した通信が可能であること、二点目は各サーバと I/O サーバ間にユーザーごとに独立したフローコントロールプロトコルを定義することにより、パケットロスのない帯域制御を他のユーザーの通信の影響を受けずに実現可能であること、三点目は複数の I/O サーバにデータを分散して送信することにより、各サーバが持つイーサネットインタフェースの性能を越えるバンド幅での通信が可能になることである。

提案手法の実現手法として、ユーザーレベルライブラリによる方式とカーネルレベルの方式の双方を設計、実装する。前者の中心は ICBC と呼ばれる通信ライブラリであり、複数の I/O サーバを経由した通信を単純な送受信 API で提供する。後者の中心は分散ネットワークインタフェースと呼ばれるネットワークデバイスドライバであり、各サーバに導入されたソフトウェアデバイスドライバが送信データを I/O サーバに転送し、I/O サーバは流量制限を適用しつつ、それらをインターネットに送信する。また、SFCP と呼ばれる各サーバと I/O サーバ間の通信プロトコルも提案し、送信過多によるパケット廃棄を防止することで安定した通信を実現する。

提案手法の評価として、並列アプリケーションを用いた複数並列計算機間のグリッドコンピューティング性能、ファイル転送を想定した大量データ転送性能、および Web サーバとしての計算機センター利用を想定した HTTP トラフィック転送性能を、最大使用可能帯域を設定した状態で測定する。測定の結果、通信の両端の計算機間で単純に TCP 接続を確立する既存の手法と比較して、並列アプリケーション性能で最大 25%、ファイル転送性能で 5.4 倍、HTTP トラフィックで 5.3% の性能向上が確認された。また、ベンチマーク中のネットワークトラフィックの解析により、バックボーンネットワークは与えられた帯域を越えることなく、かつ上限に近い帯域で使用されていることが確認された。

提案した手法は、高速インターコネクションが使用できる計算機センターにおいて外部との通信に求められる機能、性能として必要十分であり、汎用ハードウェアと新規ソフトウェアのみによる汎用的な技術である。