

広域分散ファイルシステム Gfarm と連動した高速ファイル転送ツール High-speed File Transfer Tool with the Gfarm File System

渡邊 英伸^{1*}; 黒澤 隆²; 村田 健史¹

WATANABE, Hidenobu^{1*}; KUROSAWA, Takashi²; MURATA, Ken T.¹

¹ 独立行政法人 情報通信研究機構, ² 株式会社 日立ソリューションズ東日本

¹National Institute of Information and Communications Technology, ²Hitachi Solutions East Japan, Ltd.

巨大なデータ量を扱うハイパフォーマンス・コンピューティング (HPC) は、非常に大きなデータ量を扱う数値シミュレーション等に利用されている。近年、計算クラスタ内に数百から何千ものサーバを含む大規模な処理環境になるまでに至り、大量のストレージリソースが消費されている。さらに、ストレージリソースはエクサバイトオーダ以上のサイズが要求されるまでになり、スケールアウトが可能な広域分散型のストレージシステムが注目を集めている。情報通信研究機構 (NICT) は、観測データやシミュレーションデータなど、あらゆる科学データを収集・蓄積すると同時に解析環境も提供する科学研究向けのクラウドシステム (NICT サイエンスクラウド) を構築している。NICT サイエンスクラウドは、国内 5 地区 (東京, 名古屋, 京都, 大阪, 沖縄) にあるデータセンターに分散配置した計算機を 10Gbps の L2 高速バックボーンネットワーク網である JGN-X で接続し、オープンソースの広域分散ファイルシステムの Gfarm を用いて約 3PB の広域分散型ストレージシステムを運用している。

HPC 等を想定した広域分散型ストレージシステムは、大容量データに対して高速なデータ I/O とデータ転送が重要となる。Gfarm は、ハイパフォーマンス・コンピューティング・インフラ (HPCI) の共有ストレージに採用されており、高速なデータ I/O を実現することが可能である。一方、データ転送にはインターネットで利用される標準の通信規約である TCP を採用している。TCP は長距離・高遅延のネットワークにおいて伝送遅延の問題が知られており、Gfarm では、TCP マルチストリーミングによってデータ転送の高速化を図っている。しかしながら、ネットワークが長距離・高遅延になればなるほど、高速化の効率が低くなっているのが実情である。

我々は Gfarm と連動する高速ファイル転送ツールを開発した。データ転送の通信プロトコルにオープンソースの通信ライブラリである UDT (UDP-based Data Transfer) プロトコルを採用し、簡易な並列データ転送制御機構を有する。UDT プロトコルは、UDP によるデータのバルク転送と RTT (Round Trip Time) に依存しない独自のフロー制御や輻輳制御を提供し、長距離・高遅延のネットワークにおいては TCP よりも高速なデータ転送が可能である。本発表では、開発した並列ファイル転送ツールを紹介するとともに、基本的な性能について報告する。

キーワード: 広域分散型ストレージ, Gfarm, 高速ファイル転送, UDT