



フロー制御および再送制御機能を実装したUDP 高速ファイル転送システム

NTTアクセスサービスシステム研究所

ひらし もりみち おくむら やすゆき
平橋 盛通 / 奥村 康行

2時間映画(1GBに圧縮)を約10秒で転送可能な高速ファイル転送技術を開発しました。この技術は、高速回線におけるファイル転送の問題点を回避し、大容量ファイルを短時間で転送可能としました。

高速ファイル転送のニーズ

近年、ADSLやFTTHの普及によりアクセスネットワークの高速化が急速に進んでいます。また、ネットワークの高速化と歩調を合わせて、コンテンツも従来のテキストや静止画に加え、音声や動画も増加しています。現在のダウンロード(ファイル転送)サービス用のコンテンツは、短時間で転送可能な小容量のものが大部分を占めており、転送に時間を要する動画などの大容量のものはあまり見受けられません。大容量コンテンツを数秒で転送できるようになれば、ダウンロードサービスの市場が爆発的に広がると予想されており、高速ネットワークの性能を最大限に引き出して大容量コンテンツを短時間に端末間で交換する機能の実現が求められています。

ファイル転送における問題点

信頼性を必要とするファイル転送をIPネットワークを介して行う場合は、通常、再送機能を持つTCP(Transmission Control Protocol)を用います。しかし、ギガビットクラス回線でTCPファイル転送を行うと、2つの要因により十分な転送速度が得られませ

ん。1つはTCP転送方式、もう1つがハードディスクの読み書き速度に関する要因です。詳細は次のとおりです。

TCP転送方式の問題点

- (1) 端末内部処理ならびに確認応答(ACK: acknowledgement)のオーバヘッド

TCPでは2つのデータパケットの受信に対して、1つのACKパケットが返信されます。しかし、高速転送時には次のような問題があります。通常の端末では、パケットを受信するたびにプロセッサに割り込みが発生し、2パケットに1回は送受信処理が中断されることになり、処理効率は高くありません。また、頻繁なACKパケットの送受信はI/Oバスの競合によるスループットの低下を招きます。さらにACKによるネットワーク帯域の消費も無視できません。例えば、イーサネットを利用して1Gbit/sの速度でデータ送信を行えば、ACKパケットの帯域は数十Mbit/sになります(図1(a))。

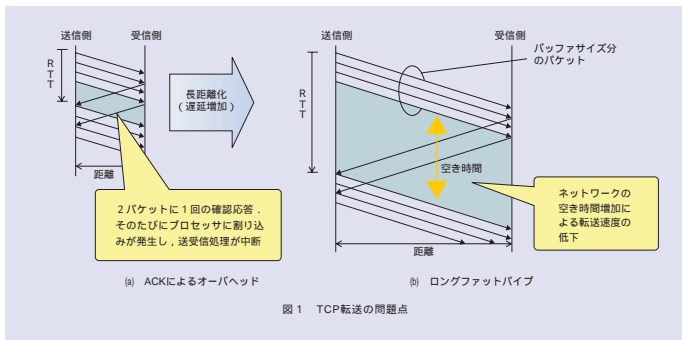
- (2) ロングファットパイプ

TCPでは、ACKが戻ってきてから次のパケットを送信します。このため広帯域であっても、長距離の回線の場合にはACKが届くまでに時間がかかり、結果として回線の空き時間が増加し、

スループットが低下します(図1(b))。

- (3) ウィンドウサイズによる帯域制御の限界

TCPの転送速度を向上する手法として、ウィンドウサイズを操作する方法がよく利用されますが、ネットワークの帯域は端末の環境や接続先によって数kbit/s~数百Mbit/sまで大きく変化するため、あらゆる場合に適用する最適なウィンドウサイズは一意には決められません。例えば、遅延の大きな高速リンクに最適化されたウィンドウサイズを低速遅延の低速リンクで使用すると、バッファに蓄積しきれないパケットが廃棄され、極端にスループットが劣化する場合があります。また、TCPではウィンドウサイズ内のパケットに対する送信速度調節機構がなく、パースティックにトラヒックが送信されるため、過大なウィンドウサイズの使用は受信側での取りこぼしや、ネットワークの不安定性を招く可能性があります。さらに、輻輳回避モードでは1RTT(往復遅延時間)当り1セグメント(1パケット)の割合でウィンドウサイズが増加するため、高速リンクにおいて十分な転送速度が得られるウィンドウサイズに到達するためには非常に長い時間が必要となります。



ハードディスクの読み書き速度の問題点
端末がデータを受信してファイルとして認識するには、通常メインメモリを経由してハードディスクに書き込む処理が必要です。しかし、一般のハードディスクへの読み書き速度は数十Mbit/sであり、大容量コンテンツをギガビットクラスで転送しても、ファイルとして認識できるまでには時間がかかります。

ギガビットクラス回線でのファイル転送では、以上のいずれかがボトルネックになり、高スループット値が得られないのが現状です。

UDP高速ファイル転送システム

このような高速時におけるファイル転送の課題を解決し、ギガビットクラスのファイル転送速度を得るために、新たに考案したUDP (User Datagram Protocol) 高速ファイル転送システムでは、高速転送に適したUDPをデータ転送に使用し(制御情報はTCPを使用)、さらにPC内のメインメモリ上でファイルを認識することで高速化を図りました。

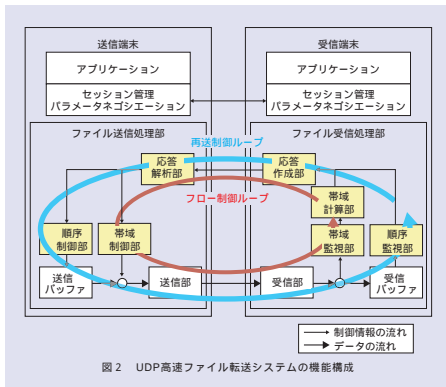
UDP転送方式の改良

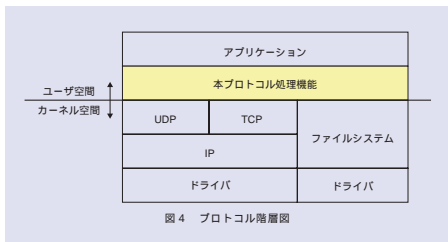
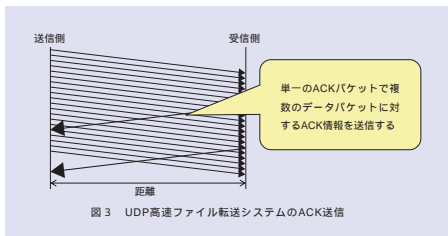
UDPはACK機能を持たずデータを一方的に送信する特徴を持っており、通常ストリーミング配信などで使われています。

そのUDPに改良を加えました。

(1) 再送制御機能の実装

UDPでデータを転送する際に問題になるのは、パケットロスやエラーによ





て損失したパケットの扱いです。受信したパケットから、元のコンテンツを正しく再構成するためには、受信データの並べ替えやデータの再送が必須です。そこで、UDPパケットのデータ部分にシーケンス番号を付与してパケットを管理し、パケット損失が起きた場合でも損失したパケットを再送することで、元のコンテンツを再構成することができるようにしました(図2)。

(2) フロー制御機能の実装

TCPのウィンドウサイズによる速度調整機構で大きな問題であったパスト的なトラフィック送出を防ぎ、ネットワークの状況に応じた帯域が即座に得られるように、パケットの転送速度とパケット損失数から送信速度を自動的に調整する

機能を備えました。

(3) ACKの削減

再送機能とフロー制御機能の実装により、本システムでも受信側からACKを送信する必要があります。TCPでは頻繁なACK送出が速度低下の原因の1つとなっているため、ACKの送出をデータパケットの受信とは独立に行うこととし、単一のACKパケットで複数のデータパケットに対するACK情報を送信します。同様の機能はTCPのSACK(Selective ACK)でも実現されていますが、本システムはオクテッド単位のACKではなく、パケット単位の確認応答を行います。これらにより、ACK送受信は効率化され、特に高速転送時の無駄な割り込みの発生やI/Oバスの競

合を最小限に抑えることができます。また、ロングファットパイプ問題も回避可能となります(図3)。

ファイル認識の高速化

端末の内部処理では、ハードディスクの書き込みと比較して、メインメモリでの処理は高速であるという特徴があります。そこで、ハードディスクにデータを転送せず、メインメモリ上のファイルシステム間でデータ転送を行うことで、ファイル認識の時間短縮を実現しました。

本システムは、LinuxPC上で動作するアプリケーションソフトウェアとして実現しており、専用ハードウェアやデバイスドライバは不要です(図4)。

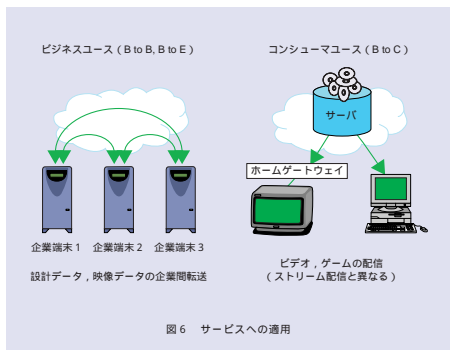
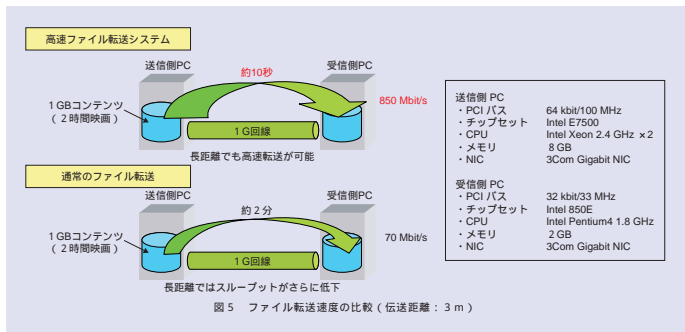
ファイル転送速度の比較

1000BASE-Tで直結した2台のLinux端末間で1GBのコンテンツを転送した場合、UDP高速ファイル転送システムでは約850 Mbit/s、TCPを使った従来のFTPでは約70 Mbit/sという結果が得られました(図5)。つまり、UDP高速ファイル転送システムを使うと高速ネットワークの性能を十分に活かした転送が可能となることを確認しました。

今後の展開

今回は、高速ファイル転送システムの機能と実験レベルでの効果について紹介しました。今後は、長距離回線におけるTCP転送方式のロングファットパイプ問題に対する解決策としての本方式の有効性と、TCPとUDPが混在する実ネットワークでの性能を検証していきます。

本システムは、現在バイク便を使って設計データや動画などの大容量データをやり取りしている企業などが、高速ネットワークを使って短時間でファイル転送を行うときに有効です。また、数百メガ



ビット・ギガビットクラスのアクセス回線が普及したときは、ゲームや映画などのコンシューマ向けダウンロードサービスの転送方法として展開することが可能です (図6)。

参考文献
(1) W・リチャード・スティヴンス: "詳解TCP/IP Vol.1プロトコル", ビアソン・エデュケーション, 2000.



(左から) 奥村 康行 / 平橋 盛通

大容量コンテンツを高速に転送するニーズは、近い将来ますます高まるものと思えます。そのため、さらに高速・高信頼で、使い勝手の良いファイル転送技術の確立を目指して開発を進めています。

問い合わせ先
NTTアクセスサービスシステム研究所
アクセスサービスネットワークアーキテクチャプロジェクト
TEL 043-211-3209
FAX 043-211-4577
E-mail okumura@ansl.ntt.co.jp