研究テーマ：分散共有型研究データ利用基盤の整備（1/2）

（プロジェクト番号　JGN2-A16058）

研究機関：　東京大学情報理工学系研究科、東京大学理学系研究科

WIDEプロジェクト

研究の概要：

東京大学、WIDEプロジェクト、および多くのネットワーク関連機関が連携して超遠距離・10Gbpsのネットワークを構築し、IPv6 およびIPv4 を用いたTCPによる超高効率（最大バンド幅に対して98%以上の実質バンド幅）通信を実現した。更に、同ネットワークを用いてディスク間でデータを共有するシステム：Data Reservoirの実証実験を実施し、iSCSIプロトコルを用いたディスク間データ転送で7Gbpsのシングルストリームのバンド幅を得た。これらの結果により、米国Internet2から、10回にわたり、Internet2 Land Speed Recordが認定された。実験参加者は合計約40名である。

研究の目的：

Internetの利用にとり、TCP通信の高速化と、ディスク間データ転送の高速化は、アプリケーションに直結する重要課題である。我々は、インターネットを用いる国際的な研究データの交換の高速化を目的として、その基礎となる、シングルストリームTCP通信の高速化、超遠距離化と、iSCSIプロトコルを用いたディスク間シングルストリームデータ転送の高効率化を研究の目的とした。我々のアプローチは、データ転送の高速化、高効率化を、通信プロトコルのレイヤ間を協調させること、ネットワークハードウェアレベルに手を入れることを特徴としている。一方、これまで研究開発された高速TCP通信方式は、コンジェスチョン・ウィンドウ・サイズとパケットロスや遅延時間の変化を関連付ける方法として最適化することにより通信の高速化を実現したが、長距離10Gbps通信では、無力であった。また、並列通信ストリーム間の負荷分散の実現は、ディスク間転送の高速化には不可欠である。他の通信との公平性を保ちつつ、並列通信ストリーム間の適正な負荷分散の実現も、本研究の目的に含まれる。

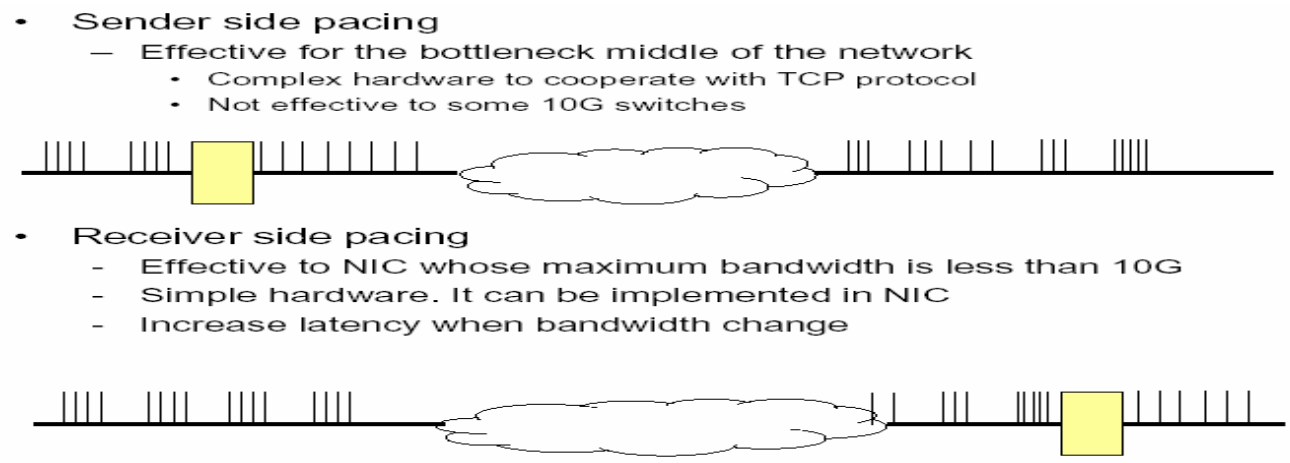

図1．ネットワーク効率を低下させるクラスタリングと、Pacingによる解決

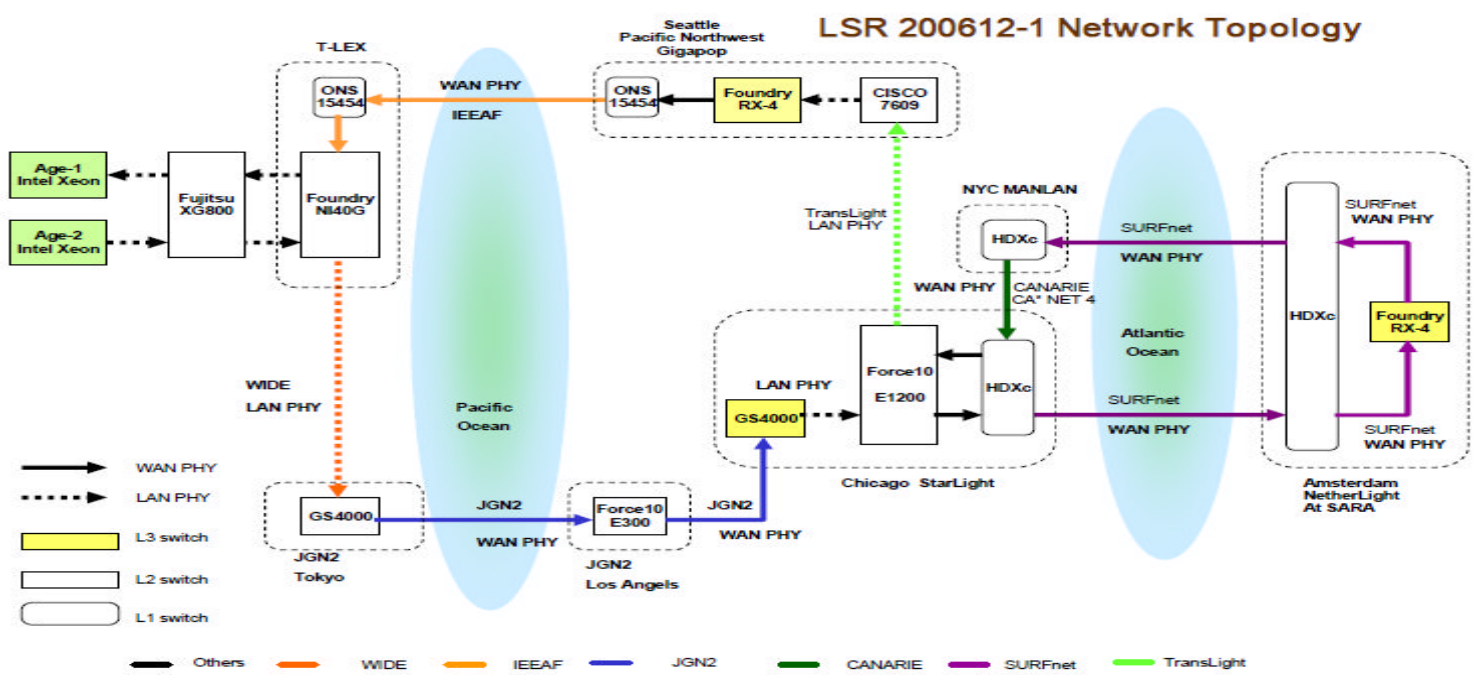

図2　実験で用いたネットワーク・機器構成

研究テーマ:分散共有型研究データ利用基盤の整備(2/2)

(プロジェクト番号　JGN2-A16058)

研究機関:　東京大学情報理工学系研究科、東京大学理学系研究科

WIDEプロジェクト

実験機器構成:

実験で用いたネットワーク構成、使用機器の例(IPv6TCP通信実験の場合)を図2に示す。

研究開発成果:

(1-1) データレゼボワールIPv6実証実験展示では、米国ペンシルバニア州ピッツバーグにおいて開催されたSC05展示実験場と、東京、アムステルダムを結ぶ国際ネットワークを用い、TCPシングルストリーム通信を、Pacingを用いて安定化させる実験を実施し、引き続きH17年度には米国ワシントン州シアトルと東京、アムステルダムを結ぶ国際ネットワークを用い、IPv6の超高速データ転送実験を実施した。その結果、適切なPacingを用いると、RTTが500msの超長距離通信においても、安定してTCPシングルストリーム通信が実現することが示された。

(1-2) データレゼボワールIPv6超高速TCP実験では、米国フロリダ州タンパにおいて開催されたSC06の展示実験場と、東京、アムステルダムを結ぶ国際ネットワークを用い、データ転送における実転送バンド幅の揺らぎと、それによる性能低下の測定と、受信側およびACK信号のPacingによる通信安定化実験を実施した。その結果、IPv4通信において、擬似遅延装置を通した場合と同等の安定した転送性能を得ることに成功した。

(1-3) データレゼボワール超高速ディスク転送実験では、東京・米国間でIPv4、TCP,iSCSIプロトコルを用いて、シングルストリームおよび2本のストリームによるディスク間通信実験を実施した。通信実験では、ディスク管理ソフトウェア、通信管理ソフトウェア、ユーザプログラムを適切にプロセッサ間負荷分散することにより、シングルストリームで7Gbps,　2本のストリームで8.6Gbpsの実データ転送速度を得た。この結果と、前年度までの結果を総合することにより、実用的かつ安定したファイル共有が、大陸間の10Gbpsネットワーク上で実現することが実証できた。

これらの結果は、レイヤ間協調技術を用いることにより、既存のネットワーク層プロトコルを用いても、超遠距離TCP通信を効率化が可能であることを示している。次のステップとして、ユーザが直接的に用いるアプリケーションから、このレベルの性能を引き出し、多数のエンドユーザが性能を享受できるシステムを構築することである。また、本研究で使用したFPGAを用いたネットワークハードウェアは、非常に有用であることがJGN2を用いた実験で実証された。これの商用化、実用化は、起業を含め、重要な将来課題である。

プロジェクトのアピールポイント

本成果は、広く使われている TCP プロトコルを変更することなく、シングルストリームおよび並列ストリームの TCP 通信、ディスク・ディスク間データ転送を、超遠距離・高速ネットワークを介して実現することを特徴としている。したがって、ソフトウェア処理の効率化、プロセッサや I/O バス速度の向上により、これまで困難といわれていたエンドユーザによる直接的な 10Gbps ネットワークが可能となることが、新規性・独自性をもつ。この特質から、本技術を適用できる範囲は、データ共有だけに限定されるものでなく、Web システム、P2P システム、マルチメディアシステム、GRID システムなど広範な適用分野を持ち、大きな社会的効果をもたらすものである。

プロジェクトの自己評価

本プロジェクトは、当初想定した、国際間10Gbpsネットワークの、TCPにより効率的利用を達成するという全体目標にもとづき、研究計画を立案した。目標達成を実現するための障害は、数多くあらわれたが、それらも、新規に研究開発した方法により克服された。したがって、本プロジェクトは、研究計画に基づき適切に進められたと判断する。