

東京大学を中心とした研究チームがインターネット速度記録を更新

- IPv4, IPv6 の両速度記録を更新 -

東京大学大学院情報理工学系研究科

平木 敬 (090-6482-9169)

2006年3月9日

東京大学、WIDE プロジェクトと内外のネットワーク機関、企業による国際共同研究チームが11月に共同で実施した実験の結果が米国 Internet2 により IPv4 (11月10日実験)、IPv6 (11月12日実験) の Internet2 Land Speed Record (インターネット速度記録) として認定されました。本成果は、東京大学が代表機関として研究を実施している文部科学省科学技術振興調整費によるものです。

記録は IPv4 では 30,000 km (実際には 32,372 km) のネットワーク上で 7.99 Gbps のデータ転送速度、239,820 terabit-meters per second (Tb-m/s) の距離バンド幅積であり、IPv6 では 30,000 km (実際には 35,755 km) のネットワーク上で 6.18 Gbps のデータ転送速度 185,400 terabit-meters per second (Tb-m/s) の距離バンド幅積です。これは、同研究チームが持っていた IPv4 および IPv6 におけるインターネット速度記録を双方とも約 11% 更新したものです。

この結果、IPv4 および IPv6 の合計 4 種目全てのインターネット速度記録において日本を中心とした研究チームが引き続き保持しています。

記録に用いたインターネットは、

シアトル 東京 シカゴ アムステルダム シアトル (IPv4) および シアトル 東京 シカゴ (IPv6) であり、シアトルに設置した東京大学の 2 台の PC 間でデータ転送を行いました。

この記録に関して特筆すべきことは、以下の通りです：

[IPv4 インターネット速度記録に関して]

- (1) 東京大学とマイクロソフト株式会社の間で 2005 年 6 月に締結しました産学連携に関する覚書に基づく研究の成果としてこの記録が達成されました。
- (2) 記録達成のためのシステムのオペレーティングシステムに、Microsoft Windows Server 2003 を用いました。両端に設置したサーバとネットワークインタフェースカードも、特殊なものではなく一般的に入手可能なものであり、今回記録を実現した技術は広く利用可能となります。
- (3) 7.994Gbps のデータ転送速度は、TCP 転送を受信するサーバの CPU 能力で決まりました。今後、CPU 速度の向上により更に転送バンド幅が高くなることが期待されます。

[IPv6 インターネット速度記録に関して]

- (4) 記録達成のためのシステムのオペレーティングシステムに、Linux オペレーティングシステム、両端に設置したサーバとネットワークインタフェースカードには Chelsio 社製 N210 を用いました。これらは特殊なものではなく一般的に入手可能なものであり、今回記録を実現した技術は広く利用可能となります。
- (5) 6.18Gbps のデータ転送速度は、TCP 転送を受信するサーバの I/O バス能力で決まりました。今後、PC の I/O バス速度の向上により更に転送バンド幅が高くなることが期待されます。

[IPv4、IPv6 共通の特筆点]

- (6) ネットワーク経路中において主要なルータ・スイッチの過半数に国産スイッチを使用しました。日本のルータ・スイッチ技術が世界最高レベルに達していることを実証しました。
(日立製 GS4000 および富士通製 XG-800 スイッチ)
- (7) WIDE, JGN2 など日本の科学技術用学術ネットワークと世界の学術ネットワーク間の密な結合を、具体的な形で示しました。

実験の詳細は

<http://data-reservoir.adm.s.u-tokyo.ac.jp/lsr-20051110/sub.html> (IPv4 記録)

<http://data-reservoir.adm.s.u-tokyo.ac.jp/lsr-20051112/sub.html> (IPv6 記録)

を参照して下さい。

Internet2 からのプレス発表は

<https://mail.internet2.edu/wws/arc/i2-news/>

Internet2 Land Speed Record (インターネット2 ランドスピードレコード) は

<http://lsr.internet2.edu/> に過去の記録を見ることができます。

この結果は、日本の高速インターネット技術が世界の最高レベルであることを具体的な形として示しています。

1. IPv4 インターネット速度記録更新に関する詳細

実験日時：2005年11月10日

実験場所 米国シアトル市内 Pacific Northwest Gigapop 内の東京大学・WIDE の施設

シアトル・ネットワーク接続

シアトル 東京

東京ネットワーク接続

日本国内ネットワーク

東京 シカゴ

Pacific Northwest Gigapop

IEEAF

T-LEX(WIDE プロジェクトが運営)

WIDE, JGN2

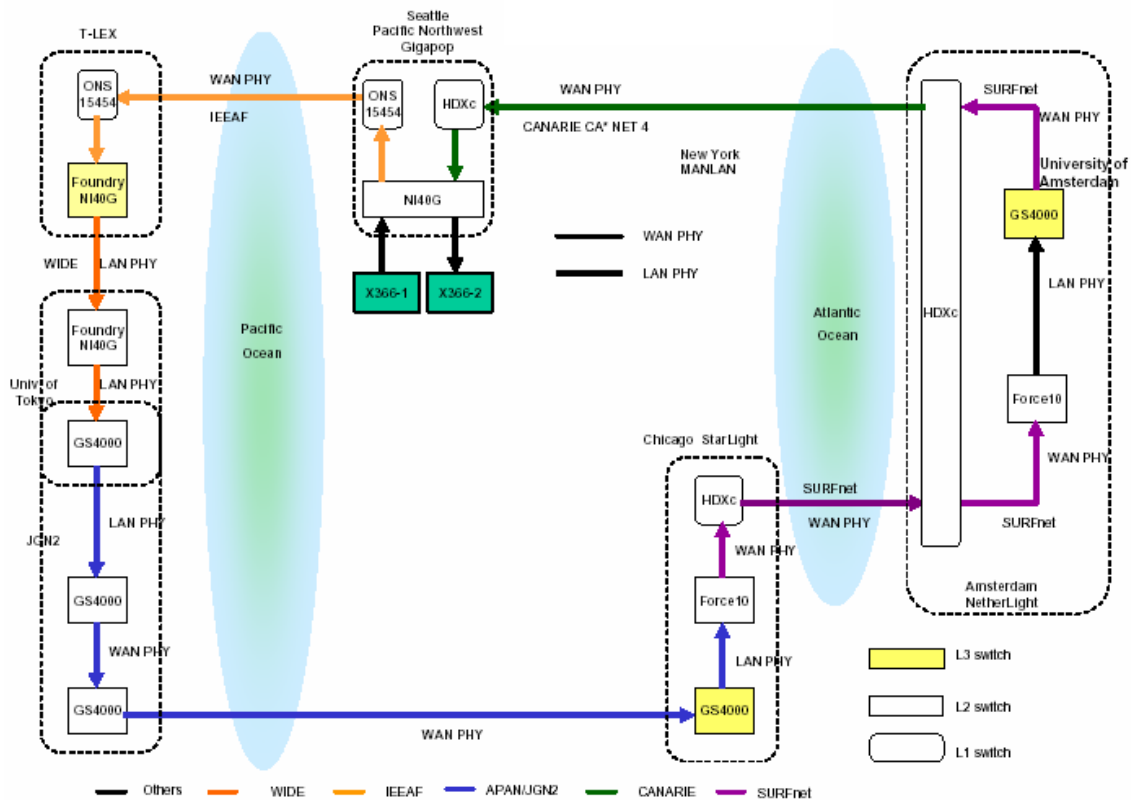
JGN2

シカゴネットワーク接続	StarLight
シカゴ アムステルダム	SURFnet
アムステルダム・ネットワーク接続	NetherLight, SARA
アムステルダム シアトル	CANARIE

2 台のサーバ間の距離は、中間にあるルータ間の最短距離の和で測ると 32,372 km となりますが、Internet2 のランドスピードレコードのルールにより、30,000 km として記録は計算されました。このネットワーク上のデータ転送速度が 7.994Gbps であったことから、今回認定された記録は **239,820** テラビット・メートル/秒となりました。

この記録は単一 TCP ストリームを用いたもので、IPv4 通信での従来の記録の 1.11 倍 であるため、LSR に認定されました。また、従来の複数ストリーム TCP クラスの記録の 1.11 倍であることから、複数ストリーム TCP クラスの記録としても認定された次第です。なお、ネットワークの往復遅延時間 (RTT) は約 500 ミリ秒です。

ネットワーク経路、ルータ、スイッチの詳細 (IPv4)



2. IPv6 インターネット速度記録更新に関する詳細

実験日時：2005年11月13日

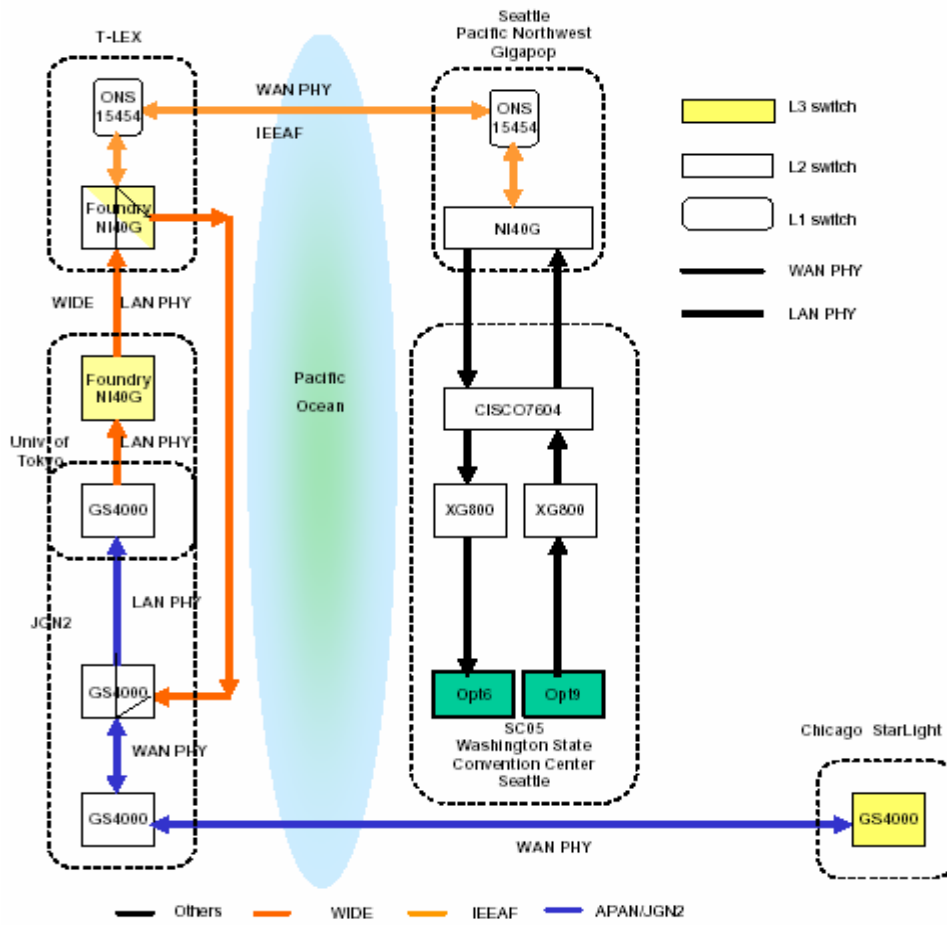
実験場所 米国シアトル市内 スーパーコンピューティング国際会議（SC05）展示場内の東京大学のブース

SC05 会場内ネットワーク接続		SCinet
シアトル・ネットワーク接続		Pacific Northwest Gigapop
シアトル 東京		IEEAF
東京ネットワーク接続		T-LEX(WIDE プロジェクトが運営)
日本国内ネットワーク		WIDE, JGN2
東京 シカゴ	JGN2	
シカゴネットワーク接続		StarLight
シカゴ 東京	JGN2	
東京 シアトル		IEEAF

2台のサーバ間の距離は、中間にあるルータ間の最短距離の和で測ると35,755 kmとなりますが、Internet2 のランドスピードレコードのルールにより、30,000 km として記録は計算されました。このネットワーク上のデータ転送速度が6.18Gbpsであったことから、今回認定された記録は **185,400** テラビット・メートル/秒となりました。

この記録は単一 TCP ストリームを用いたもので、IPv6 通信での従来の記録の1.11倍であるため、LSR に認定されました。また、従来の複数ストリーム TCP クラスの記録の1.11倍であることから、複数ストリーム TCP クラスの記録としても認定された次第です。なお、ネットワークの往復遅延時間（RTT）は約500ミリ秒です。

ネットワーク経路、ルータ、スイッチの詳細 (IPv6)



実験で用いたネットワーク経路図



この実験は、次の企業による機材の提供やサポートを得て実現されました：Cisco Systems、Anue Systems、東陽テクニカ、Foundry Networks、アラクサラネットワークス(株)、(株)日立製作所、(株)PFU、物産ネットワークス、およびネットワークワンシステムズ。

連絡先

東京大学大学院情報理工学系研究科コンピュータ科学専攻 教授

平木 敬 e-mail: hiraki@is.s.u-tokyo.ac.jp

電話：090-6482-9169

実験参加者

東京大学 平木 敬、稲葉真理、玉造潤史、菅原 豊、稲上克史、

吉野剛史、青嶋奈緒

WIDE プロジェクト 加藤 朗、山本成一、関谷勇司

Microsoft(米国) (IPv4 実験のみ)

チェルシオ・コミュニケーションズ(米国) (IPv6 実験のみ)

Felix Marti、Wael Nouredine

JGN2 豊田麻子、中村一彦、小林克志、月岡陽一、田中 仁、

服部佳隆、堀真也

Pacific Northwest Gigapop

Jan Eveleth、Bill Mar、Beth Garrison

NTT コミュニケーションズ 長谷部克幸、村上満雄、根尾美由紀、稲田 友

(株)富士通コンピュータテクノロジーズ

来栖竜太郎、坂元真和、古川裕希、生田祐吉

ネットワーク担当機関(西から東)

WIDE

JGN2

IEEAF(Internet Educational Equal Access Foundation)

CANARIE (IPv4 のみ)

SURFnet (IPv4 のみ)

ネットワーク接続ポイント機関

T-LEX, JGN2, Pacific Northwest Gigapop, StarLight,

(IPv4 のみ) University of Amsterdam, NetherLight, SARA

東京大学データレゼボワール/GRAPE-DR プロジェクトは文部科学省科学技術振興調整費に基づいた研究プロジェクトで、科学技術データの国際的な共有システムの構築および天文学や物理学、生

命科学のシミュレーションを実施する高速計算エンジンの実現を目指しています。GRAPE-DR プロジェクトでは 2008 年に 2PFLOPS の計算エンジンと数 10Gbps のネットワークを利用する国際的な研究基盤を構築します。詳細は

<http://data-reservoir.adm.s.u-tokyo.ac.jp/> および

<http://grape-dr.adm.s.u-tokyo.ac.jp/> をご覧下さい。連絡先は平木敬教授

hiraki@is.s.u-tokyo.ac.jp (Tel:03-5841-4039 または 090-6482-9169) です。

WIDE Project は、インターネットに関連する各種技術の実践的な研究開発を行う研究プロジェクトであり、1988 年から活動を行っています。133 の企業と 11 の大学を含む多くの組織と連携して、インターネットの様々な発展に寄与してきました。Root DNS サーバの一台である M.ROOT-SERVERS.NET の運用を 1997 年から行っており、また IEEAF 太平洋リンクのホストとして T-LEX (<http://www.t-lex.net/>) の運用も行っております。詳細は

<http://www.wide.ad.jp/> をご覧下さい。連絡先は、press@wide.ad.jp (Tel:0466-49-3618, Fax:0466-49-3622) です

マイクロソフトは、1975 年に創立され、人々とビジネスの持つ可能性を最大限に引き出す支援をするソフトウェア、サービス、ソリューションを提供するグローバルカンパニーです。詳しくは <http://www.microsoft.com/japan> をご覧下さい。

チェルシオ・コミュニケーションズ(本社：米国カリフォルニア州)は 10 ギガビット・イーサネットのネットワークインタフェースカードと、プロトコルプロセッサ技術を開発し、実験で用いた T110 10 ギガビットイーサネットプロトコルエンジンを製造・発売しています。詳細は <http://www.chelsio.com/> をご覧下さい。

JGN2 は、1999 年 4 月から 2004 年 3 月まで運用された JGN(Japan Gigabit Network: 研究開発用ギガビットネットワーク)を発展させた新たな超高速・高機能研究開発テストベッド・ネットワークとして、独立行政法人情報通信研究機構が 2004 年 4 月から運用を開始したオープンなテストベッド・ネットワーク環境です。産・学・官・地域などと連携し、次世代のネットワーク関連技術の一層の高度化や多彩なアプリケーションの開発など、基礎的・基盤的な研究開発から実証実験まで推進することを目指しています。

JGN2 は、全国規模の IP ネットワーク、光波長ネットワーク、光テストベッドの研究開発環境を提供しています。また、2004 年 8 月から日米回線を整備し、国内外の研究機関とも連携して研究開発を推進しています。詳細は <http://www.jgn.nict.go.jp/> をご覧ください。

IEEAF (Internet Educational Equal Access Foundation) は通信帯域や機材の寄贈を受け、国際的な研究教育コミュニティに提供することを目的とした非営利組織です。IEEAF 太平洋回線は Tyco Telecom によって 5 年間の IRU として寄贈された二番目の 10Gbps の海底ケーブルによる回線です。最初の回線である IEEAF 大西洋リンクはニューヨークとオランダのグローニガンを接続しており、2002 年から運用されています。IEEAF の寄贈は現在 17 のタイムゾーンに跨っています。詳細は <http://www.ieeaf.org/> を参照して下さい。

Pacific Northwest Gigapop はアメリカ北西部に於ける次世代インターネットの基盤で、アプリケーションの協調やテストベッドであり、国際的な peering 交換地点である Pacific Wave を運用し、また WIDE と協調して IEEAF 太平洋回線の運用を行っています。PNWGP と Pacific Wave は共に広帯域な国際および国内のネットワークを、ワシントン、アラスカ、アイダホ、モンタナ、オレゴンの各州や太平洋地域の大学や研究機関、先端的な R&D やニューメディア企業へと接続しています。詳細は <http://www.pnw-gigapop.net/> をご覧ください。

(株)富士通コンピュータテクノロジーズ(略称“CTEC”)は電子機器関係の組み込みハードウェア・ソフトウェアの開発を行う会社です。特にコンピュータの基盤であるサーバ、ストレージ、ネットワーク関連の開発については長い経験と最新の技術をもって開発を行っており、データレゼポワールはこの技術を背景に開発を行っております。詳細は <http://www.ctec.fujitsu.com/> をご覧ください。

NTT コミュニケーションズ株式会社は、日本の電気通信事業者として、日本国内及び国際の電気通信サービスを提供しており、「グローバル IP ソリューションカンパニー」を事業ビジョンとして、インターネットをはじめとするグローバルネットワークサービスの構築、システムインテグレーション、アウトソーシングビジネスの展開を行っています。

StarLight は先進的光ネットワーク基盤として、高性能アプリケーションのためのネットワークサービスを実現しています。StarLight は、2001 年夏から運用を開始し、多くのネットワークと接続するとともに光多重ネットワークに対する光スイッチングをおこなっており、APAN/JGN2, Abilene, SURFnet, CANARIE, NLR など多くのネットワークを接続しています。詳しくは <http://www.startap.net/starlight/> を参照してください。

SURFnet はオランダの 150 を越える高等教育機関や研究所を接続しているネットワークの運用と開発を行っており、国際的な先端的な研究ネットワークの一つです。SURFnet は、オランダ通商産業省、教育機関、研究機関のプロジェクトである次世代ネットワーク Gigaport の実現に責任を持っており、国内の知識基盤の強化を図ります。光および IP ネットワークやグリッドはプロジェクトの主要な要素です。詳細は <http://www.surfnet.nl/> をご覧ください。

CANARIE はカナダの先端インターネットの非営利組織で、次世代研究ネットワークやそのアプリケーション、サービスを促進しています。CANARIE はカナダを先進ネットワークで世界のリーダーと位置づけ、メンバーやプロジェクトパートナー、カナダ政府によってサポートされています。CANARIE はカナダの研究教育ネットワークである CA*net4 の開発および運用を行っています。詳しくは <http://www.canarie.ca/> をご覧ください。

アムステルダム大学アドバンスト・インターネット研究センターはアムステルダム大学理学部の施設で、インターネットの高度利用に関するアーキテクチャとプロトコルの研究を実施しています。特にセキュリティ、認証、Grid 環境と仮想研究室や e-サイエンス応用に関する最先端の研究を行っ

ています。詳しくは、<http://www.science.uva.nl/research/air/>
をご覧ください。

用語解説

Internet2 Land Speed Record インターネットを用いたデータ転送能力を、バンド幅 距離積で評価し、その最高記録である。米国 Internet2 で認定する。カテゴリーは、IPv4, IPv6 について、TCP の単一ストリームと複数ストリームの 4 クラスに分けられる。今回は、IPv6 の TCP 単一ストリームおよび複数ストリームのカテゴリーに対応する認定である。前回の認定とあわせ、全カテゴリーを東京大学、WIDE プロジェクトを中心とする共同研究チームが持っている。詳細は

<http://lsr.internet2.edu/>

<http://lsr.internet2.edu/history.html>

Internet2 Land Speed Record では、11 のルールを定め、記録を認定しています。以下にルールの要約を示します：

ルール 1) データ転送は、100Km 以上距離が離れ、少なくとも 2 個のルータが中間にあるネットワーク上を、10 分以上の期間実施することが必要である。また、距離の最大値は 30000km である。

ルール 2) データ転送を行うネットワークは、一つ以上の平常運用されている研究教育ネットワーク上を通過することが必要である。[Abilene](#), [ESnet](#), [CA*net3](#), [NREN](#) and [GEANT](#)はその例である。

ルール 3) データ転送は、一对の IP アドレス間を、RFC 793 と RFC 791 を満たす TCP/IP プロトコルで実施する必要がある。

ルール 4) 仕様する機器、ソフトウェアは Internet2 の米国メンバーが商品として入手可能か、またはオープンソースソフトウェアであることが必要である。

ルール 5) 申請には、認定に必要な全ての情報を添付することが必要である。

ルール 6) Internet Land Speed Record には、単一ストリームの TCP/IPv4、複数ストリームの TCP/IPv4、単一ストリームの TCP/IPv6、複数ストリームの TCP/IPv6 のクラスがある。単一ストリームの記録が複数ストリームの従来記録より大きい場合には、単一ストリームは複数ストリームと解釈する。

ルール 7) データは、ユーザ空間上のバッファ間で行われ、バッファごとにデータの内容が異なることが必要である。サムチェックなどの方法でデータが正しく伝達されたことを確認することが必要である。

ルール 8) 記録は、バンド幅と距離の積で評価する。距離は、端点と、途中のルータ間を、最短距離で結んだ距離の和とする。

ルール 9) 省略

ルール 10) 新たな記録は、それまでの記録を少なくとも 10 % 凌駕することが必要である。

ルール 11) 省略

Internet2 米国の学術ネットワーク組織。米国の 200 以上の大学、政府と産業界が連携して、次世代インターネット創造を目的とした研究と高等教育のための先進的ネットワーク技術・ネットワーク応用技術の開発を行っています。Internet2 は、インターネットの草創期に在ったアカデミア・産業界・政府の連携を再び現出させます。詳しい活動内容は、

<http://www.internet2.edu/> を参照してください。

インターネット 広義では、多数のユーザにより共有して使われ、国内・海外のネットワーク通信を実現しているネットワークを意味するが、通常は、ウェブページアクセス、メール、ファイル転送などに広く用いられている TCP を用いたネットワーク体系のことである。

IPv6 現在主に用いられている 32 ビットアドレスの IPv4 では今後のインターネットの急成長に対応できないため設定された次世代インターネット標準規格。128 ビットアドレスを持ち、世界中全てのインターネット機器に個別のアドレスを付けることが可能であるとともに、暗号化などを含む豊富な機能を持っている。

TCP Transmission Control Protocol の略。ネットワーク上では、エラーの発生、複数の通信の衝突による利用可能バンド幅の減少と不安定性が問題である。これを解決するため、受信側が送信側に正常に受理した確認（アクノレッジ）を返すことにより、送信側が正常に通信が行われたことを確認するとともに、次に送る流量を決定する通信プロトコル。正常なアクノレッジが返らなかった場合には、データの再送により信頼ある通信を実現する。

ウェブページのアクセス、メール通信、ファイルアクセス、データベースシステムなど私たちが日常的に使う通信の大部分は TCP を用いている。

単一ストリーム TCP 上記のように、TCP 通信では、受信側からの正常に通信が行われたことを通知するアクノレッジを待ちながらデータ通信を行う。従って、遅延時間の長いネットワークでは、ともすると低いバンド幅でしかデータ通信が行えず、一個だけの TCP コネクションだけで高いバンド幅を実現することは困難と考えられてきた。複数ストリーム TCP は、この問題を回避する目的で、一対のサーバ間で多数の TCP コネクションをはる通信方式である。今回の実験では、超遠距離通信における TCP の根本問題を解決したため、単一ストリームで困難なく高い性能を得ることができたため、複数ストリームで更に性能を上げる必要性はない。

レイヤ間協調最適化技術 東京大学が研究開発した高速通信最適化方式。従来、通信に対する最適化は、各通信レイヤで行われてきた。例えば、FAST TCP, High-Speed TCP、Scalable TCPはネットワークレイヤにおける通信最適化である。東京大学のデータレゼボワールドプロジェクトは、実際の通信線上でのパケットの最適配置や複数ストリーム間協調による新しいタイプの通信最適化方式を提案してきた。今回用いられた最適化は、TCPの流量制御と協調するパケット位置の最適化である。詳細は、同プロジェクトが SC2004 で発表した論文に示されている。論文は、<http://www.sc-conference.org/sc2004/schedule/pdfs/pap254.pdf>