

「日本の研究チームがインターネット速度記録全種目を獲得」

- 従来の IPv6 速度記録を大幅に更新 -

東京大学大学院情報理工学系研究科

平木 敬

2005 年 11 月 17 日

要 旨

東京大学、WIDE プロジェクト、(株)富士通コンピュータテクノロジーズ、チェルシオ・コミュニケーションズ(米国)、JGN2(日本)、パシフィック・ノースウェストギガポップ(米国)、NTT コミュニケーションズ(株)が 10 月 28 日、29 日に共同で実施した実験の結果が米国 Internet2 の Internet2 Land Speed Record (インターネット速度記録) に IPv6 における 2 種目で認定されました。これは、カリフォルニア工科大学が持っていた IPv6 におけるインターネット速度記録の 2.3 倍の記録です。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保持しています。

記録に用いたインターネットは、東京 シカゴ 東京 シアトル 東京 であり、東京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。データ転送速度は 1 秒間に 5.58 ギガビット (約 56 億ビット) 毎秒です。将来のインターネット標準規格として、日本を中心に推進している IPv6 インターネット規格を用いて、従来から用いられてきた IPv4 インターネットとほぼ同等の性能が達成できることを実証しました。

実験の詳細は <http://data-reservoir.adm.s.u-tokyo.ac.jp/lsr-20051029/sub.html> 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/2005-11/msg00004.html>

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。

<http://data-reservoir.adm.s.u-tokyo.ac.jp/press/lsr-20051115-j> に本発表文書があります。

詳 細

東京大学、WIDE プロジェクト、(株)富士通コンピュータテクノロジーズ、チェルシオ・コミュニケーションズ(米国)、JGN2(日本)、パシフィック・ノースウェストギガポップ(米国)、NTT コミュニケーションズ(株)が10月28日、29日に共同で実施した実験の結果が米国 Internet2 の Internet2 Land Speed Record (インターネット速度記録)に IPv6 における2種目で認定されました。これは、カリフォルニア工科大学が持っていた IPv6 におけるインターネット速度記録の2.3倍です。

今回の記録は、米国カリフォルニア工科大学が平成17年1月19日に達成した記録の2.3倍の記録を達成し、IPv6 を用いた単一ストリーム TCP クラスと複数ストリーム TCP クラスの双方での記録となったものです(表1)。この結果、12月24日に東京大学、WIDE プロジェクト他の国際共同チームが達成した IPv4 を用いた単一ストリーム TCP クラス、複数ストリーム TCP クラスと併せ、インターネット速度記録の全種目において日本を中心とした研究チームが保持しました。

この記録は、今後、インターネットの普及に伴う IP アドレスの不足を解消し、また将来開発されるであろう様々な技術の基盤として位置づけられている IPv6 においても、10ギガビットネットワークの高効率利用が特殊な方式なしで実現することを示しました。現在、多くの PC やサーバが扱っている大多数のアプリケーション、例えばウェブページのアクセス、メール通信、ファイル転送などが直接的に利用できることを意味し、今後数年間のネットワーク利用の高速化が、標準のイーサネット、標準の TCP 通信方式で実現することが可能であることを示しました。

具体的には、CD 一枚のデータを0.7秒、DVD 一枚を5秒、100ギガバイトのディスク一個分のデータを約110秒で地球の反対側に送ることに相当するものです。

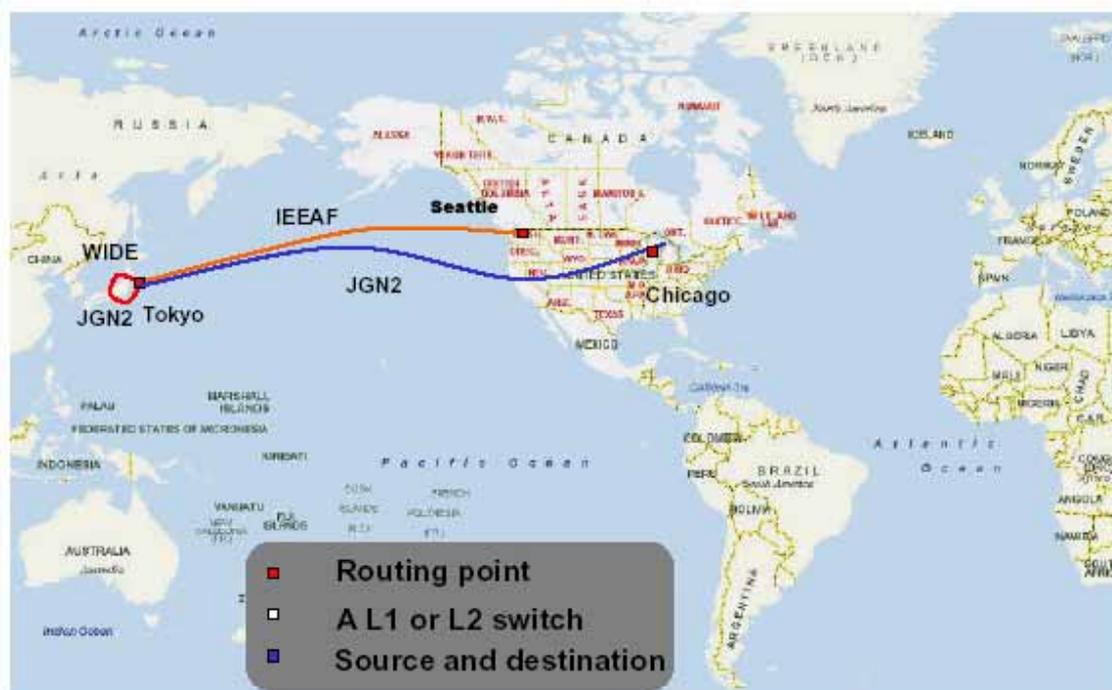
インターネット速度記録を実現したネットワークは、東京に設置された2台の東京大学のサーバ間を、下記ネットワークで結合しました。

東京ネットワーク接続	T-LEX(WIDE プロジェクトが運営)
東京 シカゴ	JGN2
シカゴ 東京	JGN2
東京 東京大学	WIDE
東京大学 東京	JGN2
東京 シアトル	IEEAF
シアトル・ネットワーク接続	Pacific Northwest Gigapop
シアトル 東京	IEEAF

2 台のサーバ間の距離は、中間にある機器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 5.58Gbps であったことから、今回認定された記録は 167,400 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 2 . 3 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 2 . 3 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。

実験で用いたネットワーク経路図



年月日	バンド幅距離積 テラビット・メートル/秒	研究グループ	クラス
2000年3月21日	4278	Microsoft Qwest Communications University of Washington USC Information Sciences Institute	IPv4 Single Stream
2000年3月29日	5384	Microsoft Qwest Communications University of Washington USC Information Sciences Institute	IPv4 Multiple Stream
2002年4月9日	4933	University of Alaska at Fairbanks Faculty of Science of the University of Amsterdam SURFnet	IPv4 Single Stream
2002年8月22日	40	Oregon Gigapop NYSERNet University of Oregon	IPv6 Multiple Stream
2002年9月27日	2517	ARNES DANTE RedIRIS	IPv6 Single Stream
2002年10月9日	5154	ARNES DANTE RedIRIS	IPv6 Single Stream
2002年11月19日	10136	Nationaal Instituut voor Kernfysica en Hoge-Energiefysica (NIKHEF) Universiteit van Amsterdam (UvA) Stanford Linear Accelerator Center (SLAC) California Institute of Technology (Caltech)	IPv4 Single and Multiple stream
2003年2月23日	23888	California Institute of Technology (Caltech) CERN Los Alamos National Laboratory (LANL) Stanford Linear Accelerator Center (SLAC)	IPv4 Single and Multiple stream
2003年5月6日	6947	California Institute of Technology (Caltech) CERN	IPv6 Single and Multiple stream
2003年10月10日	38420	California Institute of Technology (Caltech) CERN	IPv4 Single and Multiple stream
2003年11月11日	61752	California Institute of Technology (Caltech) CERN	IPv4 Single and Multiple stream
2004年2月22日	68431	California Institute of Technology (Caltech) CERN	IPv4 Multiple stream
2004年4月14日	69073	SUNET Sprint	IPv4 Single stream
2004年5月6日	77699	California Institute of Technology (Caltech) CERN	IPv4 Multiple stream
2004年6月25日	104529	California Institute of Technology (Caltech) CERN	IPv4 Multiple stream
2004年6月28日	103583	California Institute of Technology (Caltech) CERN	IPv4 Single stream
2004年9月12日	124935	SUNET Sprint	IPv4 Single and Multiple stream

2004年11月8日	184877	California Institute of Technology (Caltech) CERN CENEC	IPv4 Multiple stream
2004年11月9日	148850	University of Tokyo Fujitsu Computer Technologies WIDE Chelsio Communications	IPv4 Single stream
2004年12月25日	216300	University of Tokyo Fujitsu Computer Technologies WIDE Chelsio Communications APAN JGN2 CANARIE SURFnet Universiteit van Amsterdam	IPv4 Single and Multiple stream
2005年1月19日	72225	California Institute of Technology(Caltech) CERN	IPv6 Single and Multiple stream
2005年10月28日	91838	University of Tokyo Fujitsu Computer Technologies WIDE Chelsio Communications JGN2 Pacific Northwest Gigapop	IPv6 Single and Multiple stream
2005年10月29日	167400	University of Tokyo Fujitsu Computer Technologies WIDE Chelsio Communications JGN2 Pacific Northwest Gigapop	IPv6 Single and Multiple stream

表 1 . 過去のインターネット速度記録

この記録に関して特筆すべきことは、以下の通りです：

(1) データ通信が、インターネット規格として重要である IPv6 で実現したことは、この技術が将来、携帯通信、ユビキタス情報技術を含む広い分野へ適用可能であることを意味していること。Web アクセス、ファイル転送、メールシステム、GRID などは想定される応用分野です。

(2) 両端に設置したサーバは、特殊なものではなく一般的に入手可能な標準的部品を組み合わせた PC であること (約 30 万円の PC および普通に入手可能な 10 ギガビットのネットワークアダプタ)。

(3) 5.58Gbps のデータ転送速度は、TCP を実現するソフトウェアとネットワークの安定性で決まりました。東京大学が研究開発した「レイヤ間協調による最適化」技術が遠距離・高バンド幅 TCP 通信の持つ問題点をここでも解決したことを意味していること。

(4) 日本のネットワーク技術が世界最高レベルに達していること、日本と世界の学術ネットワーク間の密な結合を、具体的な形で示したこと。

なお、本成果は、科学技術振興調整費「重要課題解決型研究等の推進」プログラムの実施課題「分散共有型研究データ利用基盤の整備」によるものです。

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、東陽テクニカ、Foundry Networks、物産ネットワークス、およびネットワンシステムズ。

連絡先

東京大学大学院情報理工学研究科コンピュータ科学専攻 教授 平木 敬

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、
青嶋奈緒、中村 誠

(株)富士通コンピュータテクノロジーズ

来栖竜太郎、坂元眞和、古川裕希、生田祐吉

WIDE プロジェクト 加藤 朗、山本成一、関谷勇司

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

チェルシオ・コミュニケーションズ(米国) Felix Marti、Wael Nouredine

JGN2 中村一彦、豊田麻子、小林克志、月岡陽一、田中 仁、服部佳隆

Pacific Northwest Gigapop

Jan Eveleth、Bill Mar、Beth Garrison

ネットワーク担当機関(西から東)

WIDE

JGN2

IEEAF(Internet Educational Equal Access Foundation)

ネットワーク接続ポイント機関

T-LEX、Pacific Northwest Gigapop

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。

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/>) の運用も行ってあります。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

㈱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、データレゼボワールはこの技術を背景に開発を行っております。本システムは科学研究では最も標準的なLinuxプラットフォーム上で、ストレージ、ネットワークを並列に動作させることにより、おのおのの能力を最大限に活かし、高いスケーラビリティにより世界最速の転送性能を実現しています。詳細は <http://www.ctec.fujitsu.com> 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/> をご覧下さい。

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。

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つことになる。詳細は

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

Internet2 Land Speed Record では、11 個のルールを定め、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別個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>