

超長距離・超高速インターネットを用いたファイル転送速度の世界記録を更新
ー 24,000km のインターネット上で、7.56Gbps のディスク間データ転送を実現 ー
ー BWC2003 において最高バンド幅・距離積賞を受賞 ー

連絡先：東京大学大学院情報理工学系研究科

平木 敬

電話：03-5841-4039

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

キーポイント バンド幅・距離積で世界最高値を記録 (181,440 terabit-meters/second)
過去最長の 24000 km 離れた 2 点間での 7.56Gbps の超高速データ転送
シングルファイル転送の実現

概要

東京大学、富士通研究所、富士通コンピュータテクノロジーによるデータレゼボワール・プロジェクトチーム（リーダー：東京大学情報理工学系研究科、平木 敬教授）は、米国アリゾナ州フェニックスにて開催した SC2003 国際会議において実施されたインターネットの高速利用に関するコンテストであるバンド幅チャレンジ 2003 に参加した。同チームは、11 月 19 日にバンド幅チャレンジの公式測定を行い、日米間（フェニックス（図 1）＝東京＝ポートランド＝東京（図 2）を経路とする）24000 km (15000 マイル) の超高速通信（最大容量 8.2Gbps）において、TCP/IP 通信により、ネットワーク上のバンド幅 7.56 Gbps のディスク間データ転送速度（図 4、5）でデータを転送した。長距離データ転送の指標であるバンド幅・距離積は 181,440 terabit-meters/second であり、同プロジェクトチームが 11 月 7 日に実現した世界記録を約 1.44 倍更新した。また、ネットワークの平均使用効率も約 92% の高効率であった。

24,000 km におよぶ長距離のギガビットレベルのデータ通信の実現は従来の日米通信距離と比較して、倍以上長距離であり、24,000 km が世界の半周より長いことから、世界中の何れの 2 箇所の間において、超高速のデータ転送が可能であることを実証した。

この記録が、SC2003 バンド幅チャレンジ参加者において最大のバンド幅・距離積であることから、「最高バンド幅・距離積賞・ネットワークテクノロジー賞」を受賞した。データレゼボワールプロジェクトチームが、現時点で最も高速かつ遠距離にデータを転送するテクノロジーであることが実証された。また、データレゼボワールプロジェクトは、バンド幅チャレンジ 2002 における「最高効率賞」に引き続き、2 年連続の受賞である。

7.56Gbps のディスク間ファイル転送速度は、DVD のディスク 1 枚を 5 秒以内に転送する速度であり、長距離インターネットの使用形態に対し、新たな領域を切り開くものである。

また、本成果は今回開発した技術を用いることにより日米間、日欧間における超高速ネットワークを用いた科学技術研究、特に、現在は日常的に DLT テープを用いて観測施設から大学へのデータの転送を行っている、高エネルギー実験、天体観測実験等、巨大な実験観測データを共有する研究に対しネットワークを用いたデータ転送の基本技術が確立していることを示した。

実験形態

実験は、米国アリゾナ州フェニックスから東京まで、NTT コミュニケーションズが運用するネットワーク、APAN が運用する APAN ネットワーク、国立情報学研究所が運営する SUPER Sinet、東京から米国オレゴン州ポートランド間の IEEAF が運用する OC-192(9.6Gbps)ネットワーク接続を折り返し、東京に両端のサーバおよびディスクを配置して実施した（図3）。実験で用いたネットワークのボトルネックは、フェニックスから東京への太平洋を横断する部分で、8.2Gbps の最大ネットワーク容量である。ネットワークの遅延時間は、RTT（ラウンドトリップタイム）約 350 ミリ秒である（図6）。

使用機器は片側に、32 台のサーバ、128 台のディスクが用いられ、Foundry 社製 BigIron8000 スイッチを介してアリゾナ州フェニックスから東京へデータ転送し、東京において Juniper 社製 T320 ルータで再び米国オレゴン州ポートランドへデータ転送し、米国側で折り返して東京までデータ転送する形態で実施した。

ディスクに格納されたデータは、IP 上のストレージプロトコルである iSCSI プロトコルを用い、TCP/IP により転送された。長距離通信における遅延時間問題を克服するために、Comet TCP 技術を使用した。また、ソフトウェアを用いた最適化方式として、Transmission Rate Controlled TCP 方式（転送レート制御を用いた TCP）および Balanced Parallel TCP 方式（平衡並列 TCP 方式）を用いた。

本実験は文部科学省科学技術振興調整費、「科学技術研究向け超高速ネットワーク利用基盤整備」に基づき実施された。実験は エヌ・ティ・ティ・コミュニケーションズ株式会社、IEEAF、APAN、WIDE プロジェクト、Tyco Telecom、国立情報学研究所、ジュニパーネットワークス株式会社、シスコシステムズ株式会社、物産ネットワークス株式会社、ネットワンシステムズ株式会社、デジタルテクノロジー株式会社との協力により実現した。

なお、本稿は

<http://data-reservoir.adm.s.u-tokyo.ac.jp/Phoenix-press.html> に Web ページとして公開されています。

参考情報

<http://data-reservoir.adm.s.u-tokyo.ac.jp/Phoenix-press.html>

データレゼボワールプロジェクトチームによる、現在の記録に関するプレスリリース

<http://data-reservoir.adm.s.u-tokyo.ac.jp/Portland-press.html>

データレゼボワールプロジェクトチームが11月7日に実現した記録に関するプレスリリース。

<http://www.guinnessworldrecords.com/index.asp?id=58445>

ギネスに登録されているネットワーク通信の世界記録。今回の記録は、この記録の約 7.6 倍のバンド幅・距離積である。

<http://info.web.cern.ch/info/Press/PressReleases/Releases2003/PR15.03ESpeedrecord.html>

<http://archives.internet2.edu/guest/archives/I2-NEWS/log200306/msg00003.html>

現在公表されているネットワーク通信の世界記録（2つ目の URL はその前の記録）。今回の記録は、この記録の約 4.7 倍のバンド幅距離積である。ただし、このデータはディスク間ファイル転送ではなく、メモリ間データ転送で測定された。

http://www.aist.go.jp/aist_j/press_release/pr2002/pr20021121/pr20021121.html

これまでの日米間データ転送速度記録に関するプレスリリース。今回の記録は、この記録をバンド幅・距離積で約 25 倍、単純バンド幅で約 10.7 倍上回っている。

データレゼボワール (Data Reservoir : データの貯水池)

科学技術振興調整費・先導的研究基盤の整備「科学技術計算向け超高速ネットワーク利用基盤の整備」により東京大学、(株)富士通研究所、(株)富士通コンピュータテクノロジーズが共同で研究開発を進めている、科学技術研究用の遠隔地データ共有システム。従来からのグリッド技術が通常のファイルシステムを用いてデータ共有することにより、ネットワークバンド幅の効率的利用が困難であることに対し、記憶層で直接データ共有を実現することにより、データレゼボワールでは95%を越すネットワークバンド幅利用効率を実現し、2002年に実施された SC2002 Bandwidth Challenge において最高効率賞を受賞した。2003年度には国内外の10組の科学技術研究拠点に配備される。

研究開発担当者	東京大学情報理工学系研究科	平木、稲葉、亀澤
	東京大学理学系研究科	玉造
	東京大学情報基盤センター	中村
	(株)富士通研究所	陣崎、下見、河合、下國、 長沼、的場、都築
	(株)富士通コンピュータテクノロジーズ	来栖、坂元、古川、水口、 中野、柳沢、鳥居、生田

回線関係協力：WIDE Project

NTT コミュニケーションズ

IEEAF

APAN SuperSINET

東京大学情報基盤センター	加藤、関谷
東京大学情報理工学系研究科	山本
NTT コミュニケーションズ	村上、福田、長谷部
IEEAF	Don Riley
KDDI	小西、北辻
国立情報学研究所	浅野、松方、藤野

詳細は <http://www.adm.s.u-tokyo.ac.jp/data-reservoir/> を参照のこと。

TCP/IP

インターネットで最も一般的に利用されているプロトコルで、エラー・リカバリー機能があり信頼性のある通信である反面、遠距離通信において高転送速度を出すことは困難であることが知られている。なお、これまでのネットワーク転送速度世界記録は、通常の TCP/IP ではなく、長距離・高速転送用にプロトコルを改造した FAST TCP が用いられている。

<http://www-iepm.slac.stanford.edu/lsr2/>

今回の記録はネットワークアダプタレベルで遠距離における高速高信頼通信を実現する Comet TCP 技術によって実現した。通信アプリケーションおよびオペレーティングシステムで動作するプロトコルは通常の TCP を用いた。

Transmission Rate Controlled TCP 方式

ネットワークに対する流量制御に、TCP における輻輳ウィンドウの大きさの制御だけでなく、パケットの送信間隔を精密に制御することにより、パケットの損失率を低く抑え、長距離転送において高速データ転送を実現する技術。Transmission Rate Controlled TCP 方式は既存の TCP のウィンドウ制御アルゴリズムに一切の変更を加えていないところに大きな特徴がある。

Balanced Parallel TCP 方式

データ通信の高速化を目的として、複数の TCP 通信を同時並行的に使用する場合に発生する、各々の TCP 通信間のアンバランスを修正して、最

適通信を実現する技術。

バンド幅・距離積 長距離・超高速ネットワーク上におけるデータ通信能力を比較する指標。TCP 通信の困難度がバンド幅の増大とネットワーク遅延時間の積が増大すると増すことから用いられる。なお、バンド幅・距離積は、ネットワーク上（光ファイバ）に存在する通信中のデータの量と解釈される。

同プロジェクトチーム以外で達成されたこれまでの世界記録は 38,421 terabit-meters/second（CERN のプレスリリースによる）であったことに
対し、今回 181,440 terabit-meters/second を達成した。これは、従来の
記録の約 4.7 倍である。

ディスク間ファイル転送速度

ネットワーク速度は、両端に接続されたサーバのメモリ：メモリ間転送
速度およびディスク：ディスク間転送速度で測ることができる。

実際のファイル転送にとり、ディスク間ファイル転送速度が実際に利用
可能な転送速度である。また、多くのファイルを同時に多数転送して転送
速度を増加させる場合があるが、今回の実験では、1 個のファイルを最大
速度で転送することが可能なことが特色である。

SC2003

スーパーコンピューティングおよびハイパフォーマンスコンピューティ
ング分野における世界で最も高く評価されている IEEE コンピュータソ
サエティ主催の国際会議で、毎年 11 月に米国内で開催される。SC2003 は
米国アリゾナ州フェニックスで 11 月 15 日から 21 日の間開催された。
SC2003 では、論文発表の他に、大規模な展示会、最高速の計算システムに
対して与えられるゴードンベル賞、バンド幅チャレンジ賞などハイパーフォ
ーマンスコンピューティングに関わる賞のセッション、多数の招待講演など
が同時に開催された。なお、2003 年度のゴードンベル賞は東京大学理学系
研究科天文学専攻の牧野助教授が受賞した。

バンド幅チャレンジ(High-performance bandwidth challenge)

SC2003 の会場と、外部研究機関などとの間でネットワークを用いた実験を
行い、各実験チームが達成したネットワーク使用状況を実験内容とあわせて
判断することにより、6 個の部門に対して最も優れた実験チームを各々 1 チ
ームが表彰された。最高バンド幅・距離積賞は、ネットワークの最も基本的
な性能であるデータ転送性能を比較したものであり、データレゼボワールが
データ転送能力に関して最も優れたテクノロジーであることを示している。な
お、データレゼボワールプロジェクトチーム以外の受賞チームは

- ・ アプリケーション賞 (Application Award) Multi-continent Telescience
米国カリフォルニア大学サンディエゴ校 (UCSD) ,大阪大学サイバーメディアセンター
他

- アプリケーション基盤賞 (Application Foundation Award) Project DataSpace
米国国立データマイニングセンター(National Center for Data Mining)
- 双方向データ転送賞 (Bi-directional Award) Distributed Lustre File System
Demonstration
米国ローレンス・リバモア研究所 (Lawrence Livermore Lab.)
- 分散インフラ賞 (Distributed Infrastructure Award) Trans-Pacific Grid Datafarm
産業技術研究所グリッド技術センター
- ツール賞 (Tools Award) High Performance Grid Enabled Data Movement with GridFTP
米国サンディエゴ・スーパーコンピュータセンター (San Diego Supercomputer Center)



図 1. アリゾナ州フェニックス、
SC2003 会場におけるデータレゼボワ
ール



図2 東京側に設置した
データレゼボワール

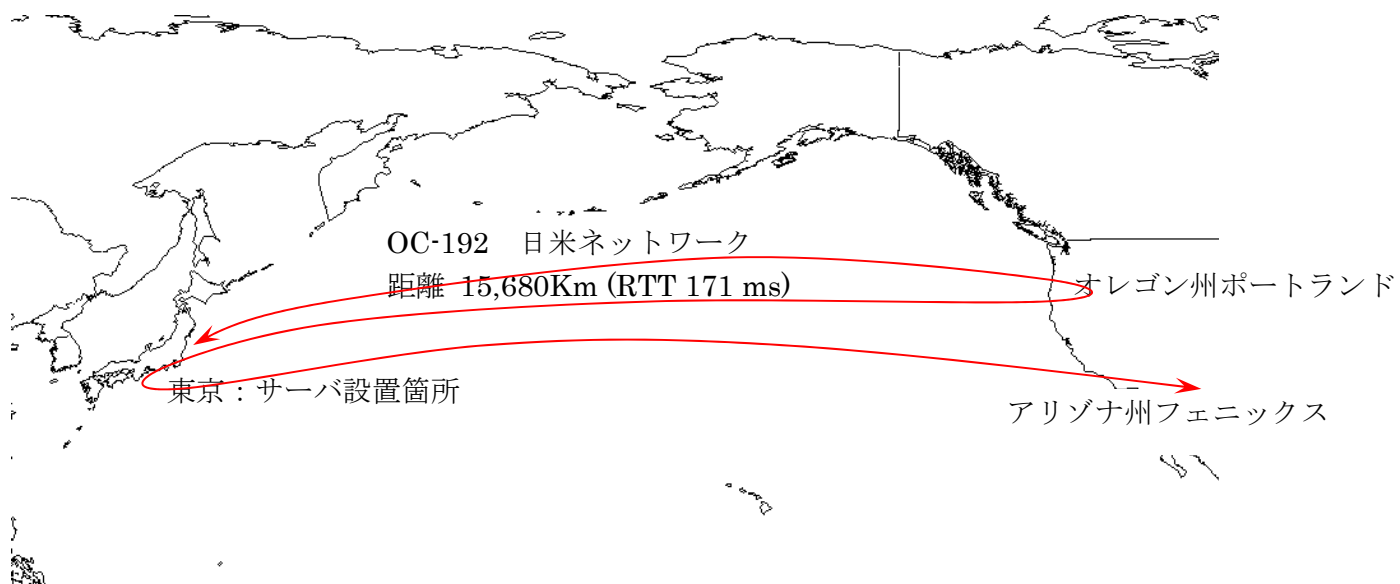


図3. 日米接続ネットワーク

Bandwidth Over Time (Current Max Datapoint: 7.56 Gb/sec)

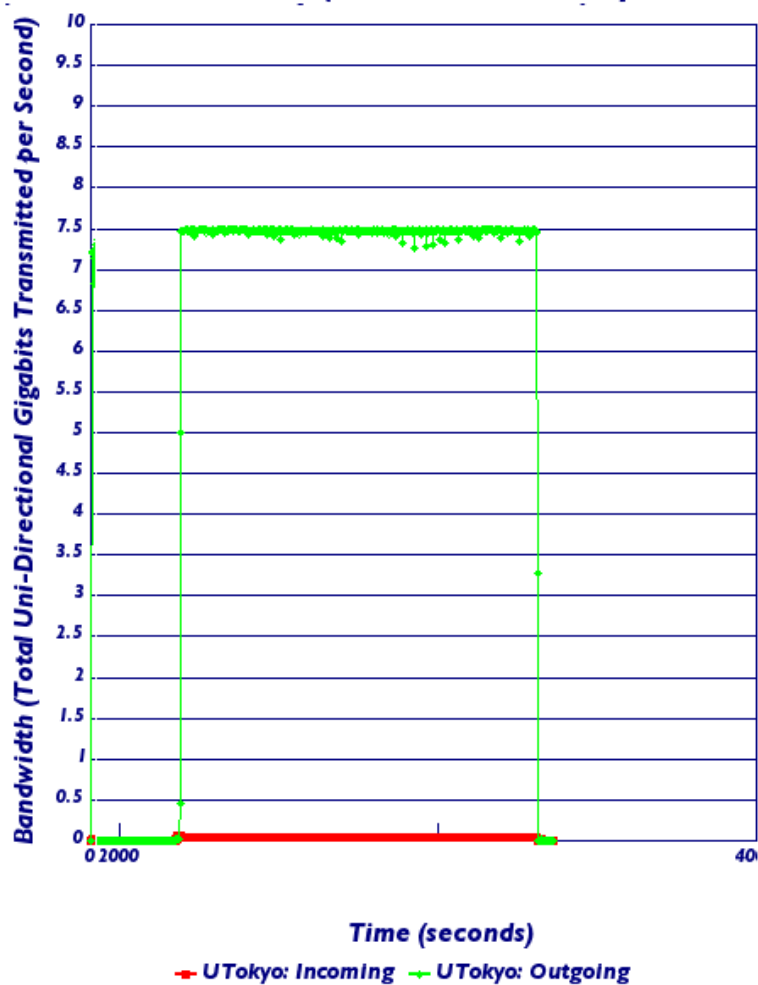


図4 送信バンド幅記録 (1)

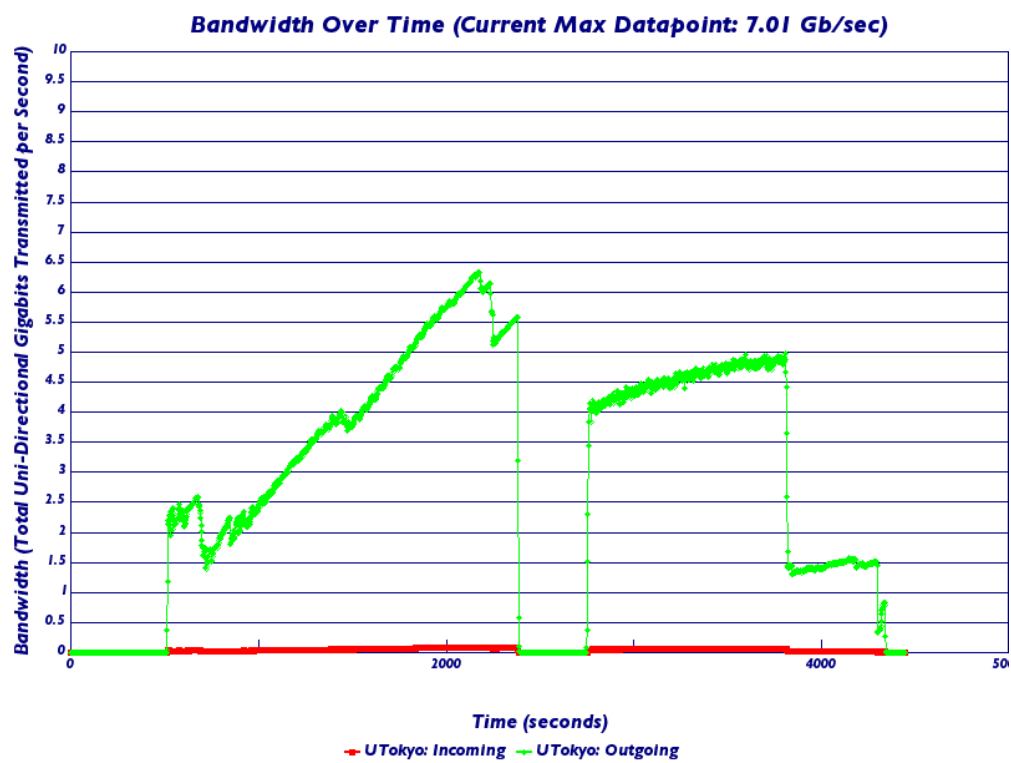


図5 送信バンド幅記録 (2)

