社団法人　電子情報通信学会
THE INSTITUTE OF ELECTRONICS,
INFORMATION AND COMMUNICATION ENGINEERS

信学技報
TECHNICAL REPORT OF IEICE.

データ転送プロトコル GridFTP のための 並列 TCP コネクション数調整機構の性能評価

伊藤　建志† 　大崎　博之† 　今瀬　眞†

† 大阪大学 大学院情報科学研究科
〒 565-0871 大阪府吹田市山田丘 1-5
E-mail: †{t-itou,oosaki,imase}@ist.osaka-u.ac.jp

あらまし　グリッドコンピューティングでは、大容量のデータを効率的に転送するために GridFTP と呼ばれるデータ転送プロトコルが用いられている。GridFTP は、複数の TCP コネクションを並列に確立することによりスループットの向上を図る「並列データ転送」と呼ばれる機能をサポートしている。しかし、GridFTP が高いスループットを実現するためには、転送に利用する並列 TCP コネクション数を、ネットワーク環境に応じて適切に設定しなければならない。我々はこれまで、グリッドのミドルウエア層から計測できる情報のみを用いて、並列 TCP コネクション数を調整する機構 GridFTP-APT (GridFTP with Automatic Parallelism Tuning) を提案した。しかし、GridFTP-APT の制御パラメータが、GridFTP-APT の性能に与える影響や、バックグラウンドトラヒックが存在する環境のような一般的な環境における GridFTP-APT の性能はこれまで明らかにされていない。そこで本稿では、シミュレーション実験により GridFTP-APT の性能を明らかにする。その結果、GridFTP-APT の最適な制御パラメータを明らかにするとともに、GridFTP-APT はバックグラウンドトラヒックの存在する環境下でも良好に動作すること、また RED ルータを導入することにより GridFTP セッション間の公平性が改善されることを示す。
キーワード　グリッド、GridFTP、並列 TCP コネクション、GridFTP-APT、制御パラメータ、公平性

Performance Evaluation of Automatic Parallelism Tuning Mechanism for Data Transfer Protocol GridFTP

Takeshi ITO†, Hiroyuki OHSAKI†, and Makoto IMASE†

† Graduate School of Information Science and Technology, Osaka University
1-5 Yamadaoka, Suita, Osaka, 565-0871 Japan
E-mail: †{t-itou,oosaki,imase}@ist.osaka-u.ac.jp

Abstract　GridFTP has been used as a data transfer protocol to effectively transfer a large volume of data in Grid computing. GridFTP supports a feature called *parallel data transfer* that improves throughput by establishing multiple TCP connections in parallel. However, in order to achieve high GridFTP throughput, the number of TCP connections should be optimized based on the network status. We have proposed an automatic parallelism tuning mechanism called *GridFTP-APT (GridFTP with Automatic Parallelism Tuning)* that adjusts the number of parallel TCP connections only using information measurable at Grid middleware. However, the effect of control parameters on the performance of GridFTP-APT has not been investigated. Also, the performance of GridFTP-APT in a realistic network environment has not been investigated. In this paper, we therefore investigate the performance of GridFTP-APT through simulation experiments. We reveal the optimal setting of GridFTP-APT control parameters. Moreover, we show that GridFTP-APT operates effectively in a network with background traffic, and fairness among GridFTP sessions can be improved by introducing RED routers.
Key words　Grid, GridFTP, Parallel TCP Connections, GridFTP-APT (GridFTP with Automatic Parallelism Tuning), Control Parameter, Fairness

1 はじめに

グリッドコンピューティングにおいて大容量のデータを効率的に転送するためのプロトコルとして、GridFTP が使用されている [1, 2]。GridFTP は、TCP の問題点を解消することを目的として設計されており、既存の FTP にさまざまな機能が追加されている。例えば GridFTP は、複数の TCP コネクションによる並列データ転送や、TCP ソケットバッファサイズの自動交渉といった機能を持つ。

GridFTP の有効性は、並列 TCP コネクション数などの制御パラメータの設定に大きく依存する、また、GridFTP が良好に動作する制御パラメータの値は、ネットワークの状況に応じて変化することが知られている [3–5]。しかし、ネットワークの状況に応じて、GridFTP の制御パラメータを適切に設定することは容易ではない。GridFTP のプロトコルの仕様には、GridFTP のサーバ–クライアント間で、並列 TCP コネクション数を指定するコマンドが含まれている。しかし並列 TCP コネクション数をどのように決定すれば良いかは、GridFTP のプロトコルでは規定されていない。

我々はこれまで、GridFTP の並列 TCP コネクション数調整機構 GridFTP-APT (GridFTP with Automatic Parallelism Tuning) を提案した [6]。GridFTP-APT は、GridFTP クライアント上で動作し、GridFTP のグッドプットを最大化するように並列 TCP コネクション数を自動的に調整する。GridFTP-APT は、GridFTP のグッドプットが、並列 TCP コネクション数に関して上に凸の関数であるという性質を利用している。GridFTP-APT は、最大化問題の数値計算アルゴリズムを用いて、最適な並列 TCP コネクション数を探索する。

ただし、GridFTP-APT は 3 種類の制御パラメータを持つが、これらのパラメータ設定が、GridFTP-APT の性能にどの程度影響するかは十分明らかにされていない。また、バックグラウンドトラヒックが存在する環境のような、一般的な環境における GridFTP-APT の性能も十分に評価されていない。GridFTP-APT を現実のネットワーク環境で利用するためには、GridFTP-APT が良好に動作する制御パラメータ設定を明らかにする必要がある [7]。例えば、リンクの帯域や伝搬遅延などのシステムパラメータによって、最適な制御パラメータ設定がどの程度変化するかを明らかにする必要がある。

そこで本稿では、シミュレーション実験により、GridFTP-APT の性能評価を行う。まず、GridFTP-APT の制御パラメータ設定が、GridFTP-APT の性能にどのような影響を与えるかを明らかにする。具体的には、並列 TCP コネクション数の乗算増加量など、GridFTP-APT の制御パラメータが、GridFTP の平均ファイル転送レートに与える影響を調査する。その結果、最適な GridFTP-APT の制御パラメータの設定方法を明らかにする。

次に、バックグラウンドトラヒックが存在する環境における GridFTP-APT の性能を調査する。バックグラウンドトラヒックとして、一定レートで送信される UDP トラヒックを用いる。その結果、バックグラウンドトラヒックの帯域に応じて、GridFTP-APT は利用可能帯域を有効に利用できることを示す。

さらに、複数の GridFTP セッションが存在する環境における、GridFTP セッション間の公平性を調査する。GridFTP はスループット向上のために並列 TCP コネクションを用いるが、これによって GridFTP セッション間の公平性が劣化することも予想される。そこで本稿では、GridFTP-APT を用いる時に、GridFTP セッション間の公平性がどの程度実現されるかを明らかにする。その結果、RED ルータを用いることにより、GridFTP セッション間の公平性が改善されることを示す。

本稿の構成は以下の通りである。2 章では、GridFTP の並列 TCP コネクション数調整機構 GridFTP-APT の概要を説明する。3 章では、シミュレーション実験により GridFTP-APT の有効性を定量的に評価する。最後に 4 章において、本稿のまとめと今後の課題を述べる。

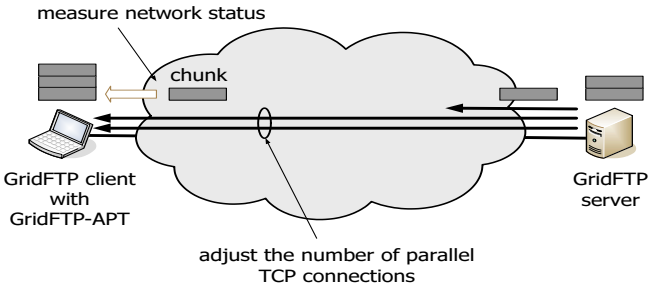

図 1 GridFTP の並列 TCP コネクション数調整機構 GridFTP-APT の概要

Fig. 1 GridFTP-APT (GridFTP with Automatic Parallelism Tuning) overview

2 並列 TCP コネクション数調整機構 GridFTP-APT

本章では、GridFTP-APT の概要を説明する。GridFTP-APT の詳細については、文献 [6] を参照されたい。

GridFTP-APT は、GridFTP クライアント上で動作する機構である (図 1)。GridFTP-APT は、転送したいファイルを「チャンク」と呼ばれるブロックに分割して転送する。チャンク転送ごとに、GridFTP-APT は並列 TCP コネクション数を自動的に調整する。

GridFTP-APT では、最大化問題の数値計算アルゴリズムを用いて、チャンク転送ごとに GridFTP のグッドプットを計測する。その結果に応じて、GridFTP のグッドプットが最大化されるように、並列 TCP コネクション数を調整する。GridFTP-APT では、最大化問題の数値計算アルゴリズムとして、黄金分割探索法 (Golden Section Search Method) を用いる [8]。GridFTP-APT では、黄金分割探索法を用いて、GridFTP のグッドプットが最大となる並列 TCP コネクション数を数値的に探索する。

2.1 並列 TCP コネクション数調整アルゴリズム

まず、黄金分割探索法を適用するために、GridFTP のグッドプットが上に凸となる、並列 TCP コネクション数の範囲 (以下、ブラケットと呼ぶ) を探索する。

並列 TCP コネクション数を小さな値から開始し、GridFTP のグッドプットが減少し始めるまで、各チャンク転送ごとに並列 TCP コネクション数を乗算的に増加させる。これにより、GridFTP のグッドプットが最大となる点を含む、並列 TCP コネクション数の範囲 (ブラケット) を決定する。

以下では、チャンク転送に用いる並列 TCP コネクション数を N、その時に計測された GridFTP のグッドプットを $G(N)$ と表記する。また、k 回前のチャンク転送時に用いた並列 TCP コネクション数を N_{-k} と表記する。

GridFTP-APT は、以下のようにブラケットを探索する。

(1) 並列 TCP コネクション数 N を初期化

$$N \leftarrow N_0$$

ここで、N_0 は並列 TCP コネクション数の初期値である。

(2) チャンクを転送し、その時の GridFTP のグッドプット $G(N)$ を計測する。

(3) 以下の不等式が成立すれば、ブラケットを (N_{-2}, N_{-1}, N) と決定し、アルゴリズムを終了する。成立しなければステップ (4) へ進む。

$$G(N) < G(N_{-1})$$

(4) 並列 TCP コネクション数 N を次式のように増加させ、ステップ (2) に戻る。

$$N \leftarrow \alpha \times N$$

ここで $\alpha(> 1)$ は制御パラメータである。なお、並列 TCP コネ

クション数は 1 以上の整数であるため、実際には、上記によって得られる N に最も近い整数を用いる。

次に、上記のように決定されたブラケット (l, m, r) に対して、チャンク転送中に黄金分割探索法を適用し、GridFTP のグッドプットが最大となる、並列 TCP コネクション数を探索する。

GridFTP-APT は、以下のように、最適な並列 TCP コネクション数 N を探索する。

(1) 並列 TCP コネクション数 N を以下のように決定する。

$$N \leftarrow \begin{cases} l + (m-l)\nu & \text{if } m-l > r-m \\ m + (r-m)\nu & \text{otherwise} \end{cases} \quad (1)$$

ここで ν は黄金分割比 $(= (3-\sqrt{5})/2)$ である。なお、並列 TCP コネクション数は 1 以上の整数であるため、実際には、上記によって得られる N に最も近い整数の値を用いる。

(2) チャンクを転送し、その時の GridFTP のグッドプット $G(N)$ を計測する。

(3) 以下の不等式が成立すれば、ステップ (4) へ進む。

$$G(N) > G(m)$$

成立しなければ、ブラケットを以下のように変更し、ステップ (1) に戻る。

$$(l, m, r) \leftarrow \begin{cases} (l, m, N) & \text{if } m < N \\ (N, m, r) & \text{otherwise} \end{cases}$$

(4) ブラケットを以下のように変更し、ステップ (1) に戻る。

$$(l, m, r) \leftarrow \begin{cases} (m, N, r) & \text{if } m < N \\ (l, N, m) & \text{otherwise} \end{cases} \quad (2)$$

2.2 チャンクサイズ決定アルゴリズム

GridFTP-APT では、次回のチャンク転送時の GridFTP のグッドプットを予測することにより、チャンク転送に要する時間ができるだけ一定となるよう、チャンクサイズを動的に変更する。具体的には、GridFTP-APT は、以下のようにチャンクサイズ X を決定する。

ブラケットを探索する時は、次回のチャンク転送時の GridFTP のグッドプットを、前回および前々回の GridFTP グッドプットの比から $G(N_{-1}) \times G(N_{-1})/G(N_{-2})$ と予測し、以下のようにチャンクサイズを決定する。

$$X \leftarrow G(N_{-1}) \frac{G(N_{-1})}{G(N_{-2})} \Delta$$

ここで、Δ は制御パラメータであり、チャンク転送時間の目標値を意味している。

なお、最初のチャンク転送時には、前回および前々回の GridFTP のグッドプット $G(N_{-1})$、$G(N_{-2})$ が不明であるため、以下のようにチャンクサイズ X を決定する。

$$X \leftarrow \frac{N_0 W}{R} \Delta$$

ここで、W は、TCP ソケットバッファサイズであり、R はラウンドトリップ時間である。ラウンドトリップ時間は、制御チャンネルにおけるコマンドの応答時間から計測する。また、2 回目のチャンク転送時には、前々回の GridFTP のグッドプット $G(N_{-2})$ が不明であるため、以下のようにチャンクサイズ X を決定する。

$$X \leftarrow \alpha\, G(N_{-1}) \Delta$$

黄金分割法を適用し、最適な並列 TCP コネクション数を探索する時は、次回のチャンク転送時の GridFTP のグッドプットを、ブラケット (l, m, r) 内の 2 点の GridFTP グッドプットの内挿によって予測し、以下のようにチャンクサイズを決定する。

$$X \leftarrow \begin{cases} ((1-\xi)\, G(l) + \xi\, G(m))\, \Delta & \text{if } N < m \\ ((1-\xi)\, G(m) + \xi\, G(r))\, \Delta & \text{otherwise} \end{cases}$$

$$\xi \leftarrow \begin{cases} \frac{N-l}{m-l} & \text{if } N < m \\ \frac{N-m}{r-m} & \text{otherwise} \end{cases}$$

表 1 シミュレーションで用いたパラメータ設定
Table 1 Parameter configuration used in simulation

ボトルネックリンク帯域 B	100	[Mbit/s]
ボトルネックリンクの伝搬遅延 τ	10, 50	[ms]
ルータのバッファサイズ	100	[packet]
TCP ソケットバッファサイズ	64	[Kbyte]
TCP のパケット長	1,500	[byte]
並列 TCP コネクション数の初期値 N_0	4	
並列 TCP コネクション数の乗算増加量 α	2	
チャンク転送時間の目標値 Δ	3	[s]
GridFTP で転送するファイルサイズ F	1, 5, 10	[Gbyte]

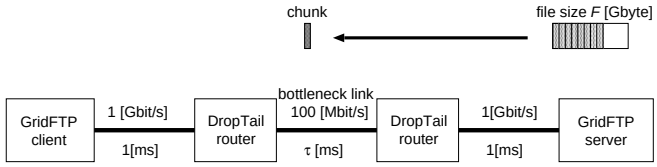

図 2 シミュレーションに用いたネットワークのトポロジ (制御パラメータの影響の調査)
Fig. 2 Network topology used in simulation for investigating the effect of GridFTP-APT control parameters

3 シミュレーション

本章では、シミュレーション実験により、GridFTP-APT の性能評を定量的に評価する。シミュレーションには、ns2 シミュレータ [9] を一部修正して使用した。シミュレーションで用いたパラメータ設定を表 1 に示す。以降のシミュレーションでは、特に断りのない限り表 1 のパラメータを用いている。

3.1 GridFTP-APT 制御パラメータの影響

まず、GridFTP-APT の制御パラメータの設定が、GridFTP-APT の性能に与える影響を調査する。具体的には、GridFTP-APT の制御パラメータである、並列 TCP コネクション数の乗算増加量 α や、チャンク転送時間の目標値 Δ、並列 TCP コネクション数の初期値 N_0 の設定が GridFTP の性能にどの程度影響するかを明らかにする。

シミュレーションに用いたネットワークのトポロジを図 2 に示す。GridFTP サーバとクライアントが 2 台のルータを介して接続されている。GridFTP サーバから GridFTP クライアントに対して単一のファイルを転送し、その時の平均ファイル転送レートを計測した。平均ファイル転送レートとは、ファイル転送における転送レートの平均であり、ファイルサイズをファイル転送に要した時間転送で除することによって計算した。

まず、GridFTP-APT がブラケット探索時に用いる制御パラメータ α が、GridFTP-APT の性能にどのような影響を与えるかを調査する。制御パラメータ α は、並列 TCP コネクション数の乗算増加量であるため、GridFTP-APT のブラケット探索の速さおよび精度に影響を与えると予想される。

以下では、紙面の都合により、システムパラメータとしてボトルネックリンクの伝搬遅延 τ および転送するファイルサイズ F を変化させた時の結果のみを示す。一般に、ボトルネックリンクの帯域 B も重要なシステムパラメータの一つである。しかし、シミュレーション実験の結果、GridFTP-APT の制御パラメータ設定はボトルネックリンクの帯域 B の影響をそれほど受けないことが分かったため、本稿ではそれらのシミュレーショ

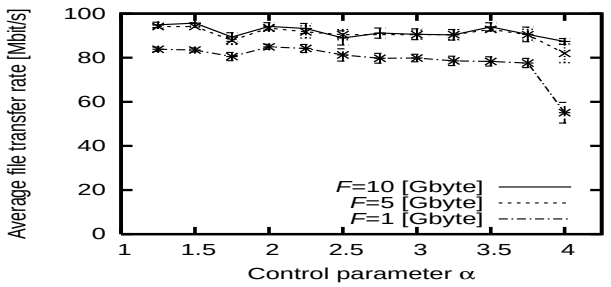

図 3　GridFTP-APT の制御パラメータ α と 平均ファイル転送レートの関係 ($\tau = 10$ [ms])

Fig. 3　GridFTP-APT control parameter α vs. average file transfer rate ($\tau = 10$ [ms])

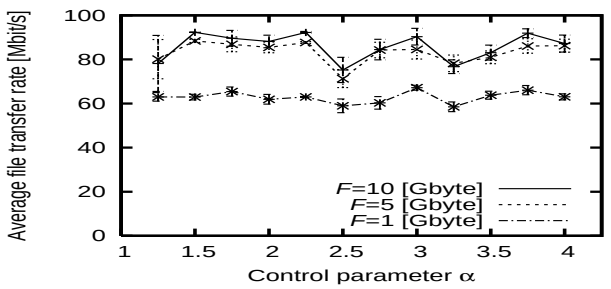

図 4　GridFTP-APT の制御パラメータ α と平均ファイル転送レートの関係 ($\tau = 50$ [ms])

Fig. 4　GridFTP-APT control parameter α vs. average file transfer rate ($\tau = 50$ [ms])

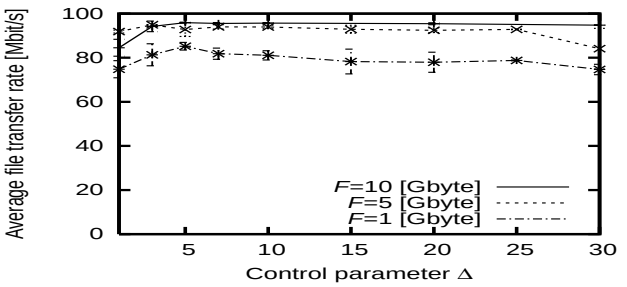

図 5　GridFTP-APT の制御パラメータ Δ と 平均ファイル転送レートの関係 ($\tau = 10$ [ms])

Fig. 5　GridFTP-APT control parameter Δ vs. average file transfer rate ($\tau = 10$ [ms])

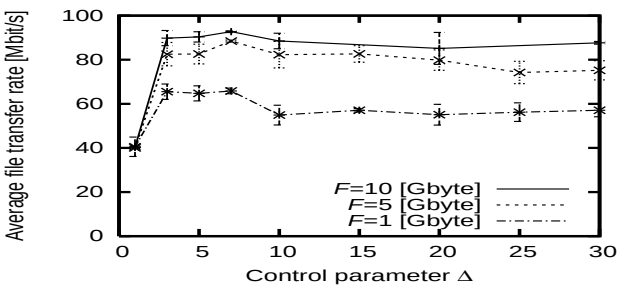

図 6　GridFTP-APT の制御パラメータ Δ と平均ファイル転送レートの関係 ($\tau = 50$ [ms])

Fig. 6　GridFTP-APT control parameter Δ vs. average file transfer rate ($\tau = 50$ [ms])

ン結果は省略する。

図 3 および図 4 に、ボトルネックリンクの伝搬遅延がそれぞれ $\tau = 10$ および $\tau = 50$ [ms] の時に、GridFTP-APT の制御パラメータ α を 1 ~ 4 と変化させた場合の GridFTP の平均ファイル転送レートを示す。20 回のシミュレーションを繰り返し、これらの値の平均値を計算した。図中には、ファイルサイズが $F = 1, 5, 10$ [Gbyte] の時の結果を 95% 信頼区間とともに示している。

これらの図より、GridFTP-APT の制御パラメータ α の設定によって、GridFTP の平均ファイル転送レートが変化することがわかる。特に、ボトルネックリンクの伝搬遅延が小さく、かつファイルサイズが小さい時には、α の値が大きくなるにつれ、GridFTP の平均ファイル転送レートが低下している。

この理由は以下のように説明できる。制御パラメータ α の値が大きすぎる場合、GridFTP-APT は、ブラケット探索時に並列 TCP コネクション数を急激に増加させる。このためブラケットが大きくなり、黄金分割探索法を用いて最適な並列 TCP コネクション数を調整するまでに要する時間が長くなると考えられる。その結果、GridFTP-APT の転送開始直後のスループットが低下し、平均ファイル転送レートが減少すると考えられる。

これらの図より、ネットワークの伝搬遅延やファイルサイズによらず、制御パラメータ α は 2 ~ 3 程度に設定すれば良いことがわかる。

次に、GridFTP-APT がチャンクサイズ決定に用いる制御パラメータ Δ が、GridFTP-APT の性能にどのような影響を与えるかを調査する。制御パラメータ Δ の値により、各チャンク転送に要する時間が変化する。このため、制御パラメータ Δ は、GridFTP-APT の並列 TCP コネクション数が収束するまでの時間に特に影響を与えると予想される [6]。

図 5 および図 6 にボトルネックリンクの伝搬遅延がそれぞれ $\tau = 10$ および $\tau = 50$ [ms] の時に、GridFTP-APT の制御パラメータ Δ を 1 ~ 30 [s] と変化させた場合の、GridFTP の平均ファイル転送レートを示す。10 回のシミュレーションを実行し、これらの値の平均値を計算した。図中には、ファイルサイズが $F = 1, 5, 10$ [Gbyte] の時の結果を 95% 信頼区間とともに示している。

これらの図より、制御パラメータ Δ の設定により、GridFTP の平均ファイル転送レートが変化することがわかる。特に、ボトルネックリンクの伝搬遅延が大きい時 (図 6) には、Δ の値が大きくなるにつれ、GridFTP の平均ファイル転送レートが減少している。

この理由は以下のように説明できる。制御パラメータ Δ の値が大きすぎる場合、1 回のチャンク転送に要する時間が長くなるために、並列 TCP コネクション数が最適化されるまでの時間が長くなる。このため、GridFTP のファイル転送レートが低下していると考えられる。

このことを確認するため、GridFTP-APT の制御パラメータ Δ が 3 または 10 [s] の時の、GridFTP の平均ファイル転送レートの時間的変動を示す (図 7)。図中には、チャンク転送ごとに計測した、GridFTP-APT のスループットの時間的変動を示している。この図からも、制御パラメータ Δ の設定により、GridFTP-APT が並列 TCP コネクション数を最適化するまでの時間が大きく異なることが分かる。

また、図 6 では、逆に Δ の値が小さすぎる場合にも、GridFTP の平均ファイル転送レートが減少している。制御パラメータ Δ の値が小さすぎる場合には、GridFTP-APT は、並列 TCP コネクションのスループットを正確に計測できない。このため、並列 TCP コネクション数が最適化されず、平均ファイル転送レートが減少すると考えられる。

なお、ボトルネックリンクの伝搬遅延およびファイルサイズにかかわらず、制御パラメータ Δ の値が 3 [s] 程度の時に、GridFTP-APT は最も高い平均ファイル転送レートを実現しているという点は注目に値する。

以上の考察から、ネットワークの伝搬遅延やファイルサイズによらず、制御パラメータ Δ は 3 [s] 程度に設定すれば良いことがわかる。

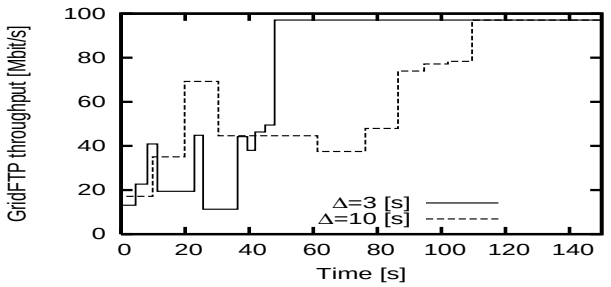

図 7　GridFTP セッションのスループットの時間的変動 ($\tau = 50$ [ms])
Fig. 7　Evolution of GridFTP throughput ($\tau = 50$ [ms])

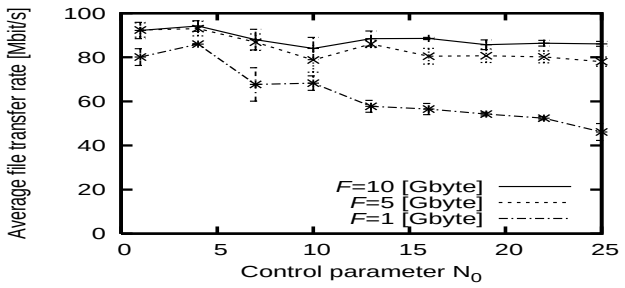

図 8　GridFTP-APT の制御パラメータ N_0 と平均ファイル転送レートの関係 ($\tau = 10$ [ms])
Fig. 8　GridFTP-APT control parameter Δ vs. average file transfer rate ($\tau = 10$ [ms])

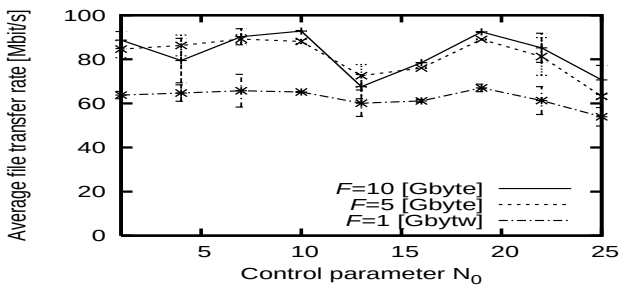

図 9　GridFTP-APT の制御パラメータ N_0 と平均ファイル転送レートの関係 ($\tau = 50$ [ms])
Fig. 9　GridFTP-APT control parameter Δ vs. average file transfer rate ($\tau = 50$ [ms])

次に、GridFTP-APT の制御パラメータ N_0 が、GridFTP-APT の性能にどのような影響を与えるかを調査する。制御パラメータ N_0 は、並列 TCP コネクション数の初期値であるため、並列 TCP コネクション数の最適化の速さに影響を与えると予想される。

図 8 および図 9 に、ボトルネックリンクの伝搬遅延がそれぞれ $\tau = 10$ および $\tau = 50$ [ms] の時に、GridFTP-APT の制御パラメータ N_0 を 1 ~ 25 と変化させた場合の、GridFTP の平均ファイル転送レートを示す。図中には、ファイルサイズが $F = 1, 5, 10$ [Gbyte] の時の結果を 95% 信頼区間とともに示している。

この図より、GridFTP-APT の制御パラメータ N_0 の設定により、GridFTP の平均ファイル転送レートが変化することがわかる。ただしファイルサイズが大きい時には、N_0 の設定による影響は小さいことも分かる。最適な制御パラメータ N_0 の値はボトルネックリンクの伝搬遅延には依存するが、ファイルサイズとはほぼ無関係であることも分かる。ボトルネックリンクの伝搬遅延が大きい場合 (図 9) には、十分な平均ファイル転送レートを実現できる制御パラメータ N_0 の範囲は広いことも分かる。

以上の考察から、最適な制御パラメータ N_0 は、ネットワークの伝搬遅延に依存はするが、ファイルサイズによらず 4 ~ 8 程度に設定すれば良いことがわかる。

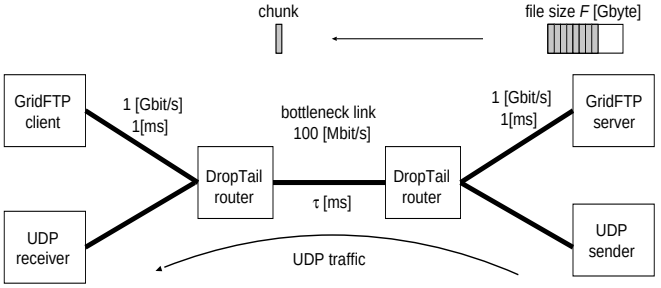

図 10　シミュレーションに用いたネットワークのトポロジ (バックグラウンドトラヒックの影響の調査)
Fig. 10　Network topology used in simulation for investigating the effect of background traffic

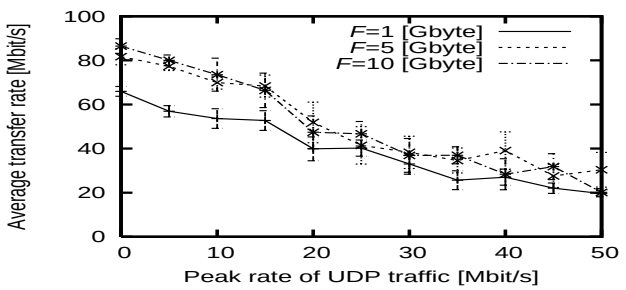

図 11　UDP トラヒックのピークレートと GridFTP の平均ファイル転送レートの関係 ($\tau = 50$ [ms])
Fig. 11　Peak rate of UDP traffic vs. average transfer rate ($\tau = 50$ [ms])

3.2　バックグラウンドトラヒックの影響

次に、バックグラウンドトラヒックが GridFTP-APT の性能に与える影響を調査する。シミュレーションに用いたネットワークのトポロジを図 10 に示す。GridFTP サーバとクライアントが 2 台のルータを介して接続されている。また、一定レートで UDP パケットを送受信するホストが、同じ 2 台のルータを介して接続されている。バックグラウンドトラヒックとして、パケット発生間隔が指数分布に従う UDP トラヒックを生成した。UDP トラヒックのパケット長は 1,500 [byte] とした。GridFTP サーバから GridFTP クライアントに対して単一のファイルを転送し、その時の平均ファイル転送レートを計測した。

図 11 に、UDP トラヒックのピークレートを 0 ~ 50 [Mbit/s] と変化させた場合の、GridFTP の平均ファイル転送レートを示す。この図より、バックグラウンドトラヒックの帯域に応じて、GridFTP-APT は利用可能帯域を有効に利用できていることが分かる。

3.3　複数の GridFTP セッション間の公平性

最後に、複数の GridFTP セッションが存在する環境における GridFTP セッション間の公平性を調査する。具体的には、各 GridFTP セッションのスループットの時間的変動を計測し、GridFTP セッション間でどの程度の公平性が実現できるかを明らかにする。

シミュレーションに用いたネットワークのトポロジを図 12 に示す。4 台の GridFTP サーバが 4 台の GridFTP クライアントに、2 台の DropTail ルータもしくは RED ルータを介して接続されている。

まず、ルータが DropTail ルータの場合の結果を示す。図 13 に、100 [s] ごとに各 GridFTP セッションがファイル転送を開始した時の、各 GridFTP セッションのスループットの時間的変動を示す。この図より、ボトルネックリンクを共有する GridFTP

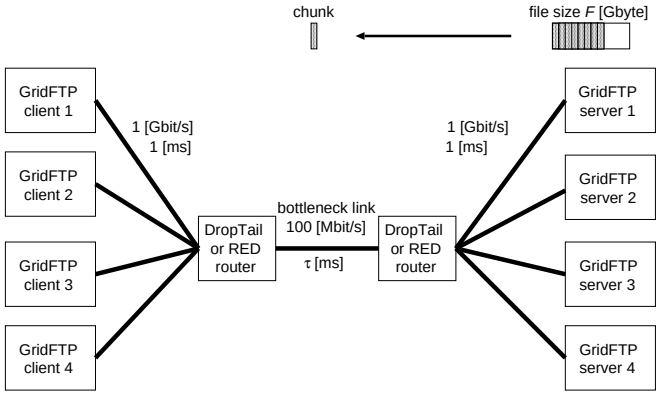

図 12　シミュレーションに用いたネットワークのトポロジ (複数の GridFTP セッション間の公平性の調査)

Fig. 12　Network topology used in simulation for investigating the fairness among multiple GridFTP sessions

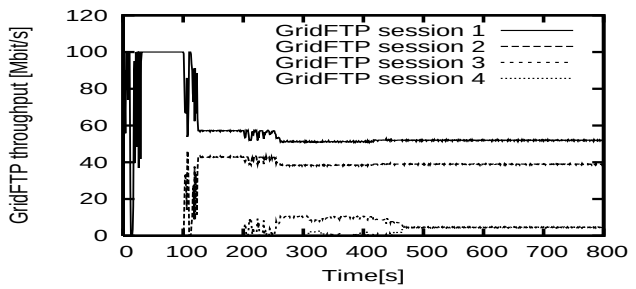

図 13　GridFTP セッションのスループットの時間的変動 (DropTail ルータの場合) ($\tau = 10$ [ms])

Fig. 13　Evolution of GridFTP throughput with DropTail routers ($\tau = 10$ [ms])

セッション数が増加するにつれ、GridFTP セッション間の公平性が低下していることがわかる。これは、GridFTP-APT 自体が、ネットワーク帯域の変動に追随するための機構を備えてないことが原因の一つと考えられる。しかし、不公平性が生じている要因の一つに、各 GridFTP セッションを構成している TCP コネクションと DropTail ルータそのものの問題がある。特に、ネットワークの帯域遅延積と比較してルータのバッファサイズが小さい時に、DropTail ルータに収容されている TCP コネクション間の公平性が劣化することが知られている [10]。

この問題は、ルータを DropTail ルータから RED ルータに変更することで、ある程度の改善が可能である。図 13 に、ルータを DropTail ルータから RED ルータに変更した時の、各 GridFTP セッションのスループットの時間的変動を示す。この図では、ルータの種類のみを変更しており、他のすべてのパラメータは図 13 と同じ値を用いている。図 14 より、ルータを DropTail ルータから RED ルータに変更することにより、GridFTP セッション間の公平性が大幅に改善されていることがわかる。これは、GridFTP セッションが用いている TCP コネクションが安定して動作しているためと考えられる。ただし、RED ルータを用いた場合でも、最初に転送を開始した GridFTP セッションが大きな帯域を獲得してしまっている。この問題を解決するためには、GridFTP-APT のアルゴリズム自体の修正が必要であろう。

4　まとめと今後の課題

本稿では、データ転送プロトコル GridFTP の並列 TCP コネクション数調整機構 GridFTP-APT の性能評価を行った。さまざまなシミュレーション実験により、最適な GridFTP-APT の制御パラメータの設定方法を明らかにするとともに、GridFTP はバックグラウンドトラヒックの存在する環境下でも良好に動作すること、また RED ルータを導入することにより GridFTP セッション間の公平性が改善されることを示した。

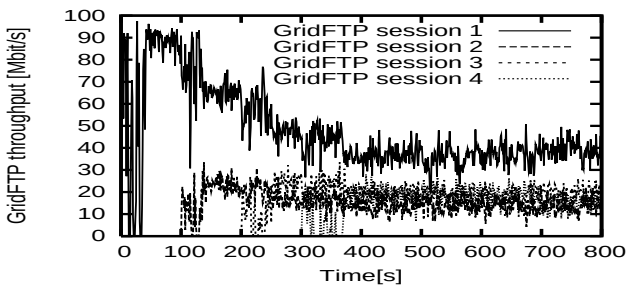

図 14　GridFTP セッションのスループットの時間的変動 (RED ルータの場合) ($\tau = 10$ [ms])

Fig. 14　Evolution of GridFTP throughput with RED routers ($\tau = 10$ [ms])

今後の課題として、3. 3 節でも議論したように、GridFTP セッション間の公平性の改善が挙げられる。また、並列 TCP コネクションの収束速度を向上するための、GridFTP-APT のアルゴリズム改善などが挙げられる。

謝　　辞

本研究を実施するにあたり、有意義な議論をしていただいた、大阪大学大学院情報科学研究科の村田正幸氏に感謝する。また、本研究の一部は、文部科学省における超高速コンピューター網形成プロジェクト (NAREGI) および、文部科学省における科学研究費補助金特定領域研究 (課題番号 16016261) の支援を受けている。ここに記して感謝の意を表する。

文　　献

[1] W. Allcock *et al.*, "GridFTP: Protocol extensions to FTP for the Grid," *GGF Document Series GFD.20*, Apr. 2003. Also available as `http://www.ggf.org/document/GFD.20.pdf`.

[2] I. Mandrichenko, W. Allcock, and T. Perelmutov, "GridFTP v2 protocol description," *GGF Document Series GFD.47*, May 2005. Also available as `http://www.ggf.org/document/GFD.47.pdf`.

[3] T. Ito, H. Ohsaki, and M. Imase, "On parameter tuning of data transfer protocol GridFTP in wide-area Grid computing," in *Proceedings of Second International Workshop on Networks for Grid Applications (GridNets 2005)*, pp. 415–421, Oct. 2005.

[4] T. J. Hacker, B. D. Athey, and B. Noble, "The end-to-end performance effects of parallel TCP sockets on a lossy wide-area network," in *Proceedings of the 16th IEEE-CS/ACM International Parallel and Distributed Processing Symposium (IPDPS)*, pp. 434–443, Apr. 2002.

[5] H. Sivakumar, S. Bailey, and R. L. Grossman, "PSockets: The case for application-level network striping for data intensive applications using high speed wide area networks," in *Proceedings of the 2000 ACM/IEEE Conference on Supercomputing*, pp. 4–10, Nov. 2000.

[6] T. Ito, H. Ohsaki, and M. Imase, "GridFTP-APT: Automatic parallelism tuning mechanism for data transfer protocol GridFTP," to be presented at *6th IEEE International Symposium on Cluster Computing and the Grid (CCGrid2006)*, May 2006.

[7] S. Floyd and V. Paxson, "Difficulties in simulating the Internet," *IEEE/ACM Transactions on Networking*, vol. 9, pp. 392–403, Feb. 2001.

[8] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press, 1992.

[9] "The network simulator – ns2." available at `http://www.isi.edu/nsnam/ns/`.

[10] S. Floyd and V. Jacobson, "Random early detection gateways for congestion avoidance," *IEEE/ACM Transactions on Networking*, vol. 1, pp. 397–413, Aug. 1993.