VPN 複数経路接続を用いた iSCSI 並列ストレージアクセス性能評価

千島　望† 　山口　実靖†† 　小口　正人†

† お茶の水女子大学　〒 112–8610 東京都文京区大塚 2–1–1
†† 工学院大学　〒 163–8677 東京都新宿区西新宿 1–24–2
E-mail: †nozomi@ogl.is.ocha.ac.jp, ††sane@cc.kogakuin.ac.jp, †††oguchi@computer.org

あらまし　近年，ストレージの管理コスト低減などの目的で SAN の導入が進んでおり，IP ネットワークを利用した IP-SAN として iSCSI が期待されている．しかし現状では SAN は主にサイト内のローカル環境のみで用いられている．そこで本研究では VPN を利用することにより，iSCSI を広域ネットワークに適用することを検討した．ネットワークストレージの信頼性と性能を向上させることを目的として，VPN マルチルーティング機能を用いて VPN 広域ネットワーク内を複数経路で接続し，さらにネットワークストレージの Target に並列ストレージシステムを利用した環境において，iSCSI ストレージアクセスの特性を解析，評価した．

キーワード　iSCSI，VPN，輻輳ウィンドウ，スループット，複数経路，RAID

Analysis of iSCSI Parallel Storage Access using Multi-routing VPN Connections

Nozomi CHISHIMA†, Saneyasu YAMAGUCHI††, and Masato OGUCHI†

† Ochanomizu University　Otsuka 2–1–1, Bunkyo-ku, Tokyo, 112–8610 Japan
†† Kogakuin University　1–24–2 Nishi-shinjuku, Shinjuku-ku, Tokyo, 163–8677 Japan
E-mail: †nozomi@ogl.is.ocha.ac.jp, ††sane@cc.kogakuin.ac.jp, †††oguchi@computer.org

Abstract　As the introduction of SAN progresses for the purpose of the storage management cost reduction, iSCSI is expected as a representative of IP-SAN that uses IP network. However, SAN is mostly used only in the server site currently. Thus we have claimed iSCSI can be employed in a WAN environment using VPN, as well as in a local environment. In this paper, we have observed the feature of iSCSI parallel storage access and evaluated its performance of the system when iSCSI access is performed using a multi-routing function of a VPN router.

Key words　iSCSI, VPN, Congestion Window, Throughput, multi-routing, RAID

1. はじめに

近年，インターネット技術の進展やマルチメディアアプリケーションの普及などにより，ユーザが蓄積し利用するデータ容量が爆発的に増加している．これに伴いストレージの増設，管理コストの増大が問題となっている．そこで SAN(Storage Area Network) が登場し，広く用いられるようになった．SAN はサーバとストレージを物理的に切り離し，各ストレージとサーバ間を相互接続してネットワーク化したもので，これにより各サーバにばらばらに分散していたデータの集中管理が実現された．

一般に SAN としてはファイバチャネルを用いる FC-SAN(Fibre Channel - SAN) が広く利用されている．しかし FC-SAN はファイバチャネルを用いているため高価となり，また接続距離に制約がある．一方，SAN に IP ネットワークを利用した IP-SAN として iSCSI が期待されている [1] [2]．iSCSI は，これまで DAS(Direct Attached Storage) で使われてきた SCSI コマンドを TCP/IP パケット内にカプセル化することにより，サーバ (Initiator) とストレージ (Target) 間でデータの転送を行う．今後インターネット技術の発展により，ギガビットクラスの回線の実現が期待され，iSCSI の有用性もさらに高まると考えられる．

現状において，SAN は主にサーバサイト内のみでしか使用されていない．しかし遠隔バックアップ等を目的として，離れたサイトのサーバとストレージを SAN で接続することが強く期待されている．そこで本稿では，VPN(Virtual Private Network) を利用することにより，ローカル環境で使用されている iSCSI を用いて広域ネットワーク上でリモートアクセスを行うことを検討した．

iSCSI は複雑な階層構成のプロトコルスタックで処理されており，バースト的なデータ転送も多いことから，通常のソケット通信と比較して，特に高遅延環境においては性能の劣化が著

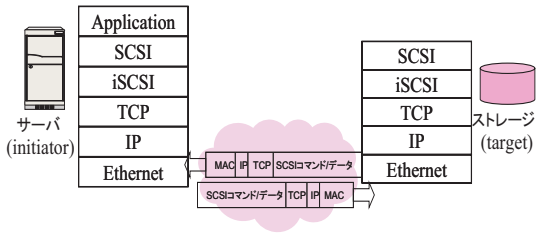

図 1　iSCSI

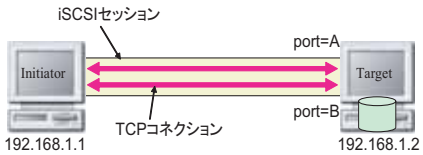

図 2　iSCSI 複数コネクション

しく，下位基盤の TCP/IP 層が提供できる限界性能を超えることはできない [3]．さらに広域環境で iSCSI 複数経路アクセスを行う場合，経路により遅延時間やネットワーク性能が異なるため，iSCSI で適応的なパケット処理を行うことが望ましい．

また，近年大規模なストレージシステムは多数のストレージノードをネットワーク結合した並列ストレージシステムとして構成されている．システムはストレージ仮想化機構により仮想化された単一の巨大ボリュームとして提供される．このような高機能並列ストレージシステムを利用する場合，システムへのアクセスネットワークがボトルネックとなってしまうと考えられる．そこで本研究ではネットワークの性能と信頼性を高めるために，VPN 広域ネットワーク内を複数経路で接続し，iSCSI 並列ストレージアクセスの特性を解析，評価した．

本論文の構成は以下の通りである．2 章で VPN 複数経路接続 iSCSI ストレージアクセスについて述べ，3 章で並列ネットワークストレージシステムの説明をし，4 章で VPN のマルチルーティング機能を利用した本実験システムの概要を述べる．5 章で iSCSI 並列ストレージアクセスについての性能評価結果を示し，6 章で輻輳ウィンドウを用いて性能評価結果の解析を行い，最後に 7 章でまとめる．

2.　VPN 複数経路接続 iSCSI ストレージアクセス

2.1　iSCSI 複数コネクション

IP-SAN の代表的なプロトコルに iSCSI がある．iSCSI は SCSI コマンドを TCP/IP パケットでカプセル化する規格で，iSCSI により SAN を IP 機器だけで構築することが可能となる．一方で図 1 のように複雑な階層構成をとることになり，下位のプロトコルの限界性能を超えることはできない．また，iSCSI には長距離アクセスの実現が期待されているが，ギガビットクラスの太い回線を用いた場合の遅延帯域積の問題も指摘されている．そこで iSCSI における性能や信頼性を向上する手法の実現が求められている．

iSCSI は様々なチューニングを行うことが可能である．本実験で用いたニューハンプシャー大学が提供する UNH-iSCSI の実装では，1 つの iSCSI セッション内に複数の TCP コネクション

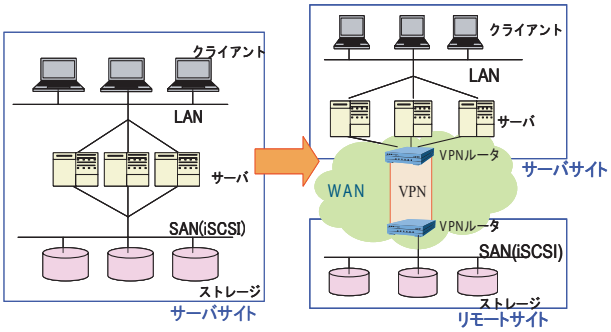

図 3　VPN 利用モデル

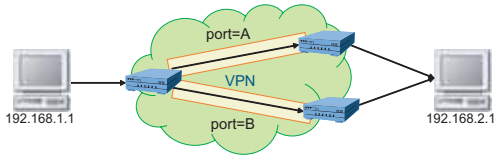

図 4　VPN マルチルーティング機能

を確立するように設定することができる [4]．さらにこのコネクションをポート番号と対応付けることができ，1 つの iSCSI セッションをポート番号が異なる複数のコネクションに分けることができる．つまり，図 2 に示すようにターゲットの 1 つの IP アドレス，1 つの iSCSI ドライブにポート番号の異なる複数のコネクションを接続することが可能である．

2.2　VPN 複数経路接続

VPN は，インターネットや通信事業者が持つ公衆ネットワークを使って，拠点間を仮想的に閉じたネットワークで接続する技術である．安価であるという公衆網のメリットを活かしつつ，機密性の低さを暗号化等の別の方法で補うことにより,「実質的な専用網」を実現できるということが VPN の利点である．一方，専用網と異なりネットワークの品質は保証されない場合が多い．

本稿では非常災害対策などを目的とした iSCSI による遠隔バックアップなどを行うため，図 3 に示すように VPN ルータで接続したリモート環境にネットワークストレージを設置し，iSCSI を広域ネットワーク環境に適用した．この場合，広域ネットワーク内の VPN 越しにアクセスを行うため，VPN ルータを通ることによってネットワークの帯域幅が制限され，スループットが著しく低下することが起こり得る [5]．さらに広域ネットワーク内は不安定な通信路であることが想定される．そこで本稿では，VPN 広域ネットワーク内を複数経路で接続することを考えた．これにより，データ転送の性能や信頼性，ネットワークの耐障害性なども向上すると考えられる．

ただし iSCSI 複数経路の構築は，アプリケーションなど上位層に対しては透過的に実現したい．本実験で用いた VPN ルータ Fujitsu Si-R570 はマルチルーティング機能を有している [6]．マルチルーティング機能を使用すると，図 4 に示すようにポート番号などの情報を利用して同じあて先 IP アドレスを持つネットワークへ複数の経路を用いて送信することが可能となる．それぞれの通信内容に通信経路を分離することが出来るため，片方

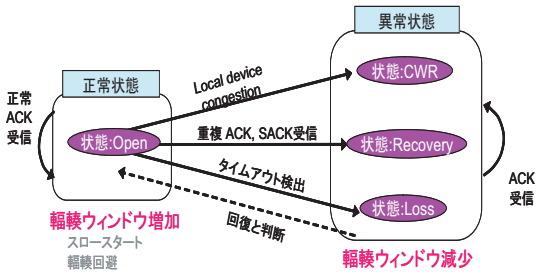

図 5　LinuxTCP の状態遷移

の回線をバックアップ用に用いたり，音声データは専用線を用いそのほかの通信は公衆網を用いるなどと設定することができる．本研究ではこの機能を利用し，iSCSI セッション内の TCP コネクションごとに異なる経路を構築することを可能にした．

iSCSI は通常ギガビットクラス以上の太いネットワーク上で用いられるが，途中に VPN ルータの暗号化処理速度などによりスループットが決まる細い回線が挟まることにより，トラフィックとして大いに性質の異なるものになる [7]．この場合，iSCSI の性能は下位層である TCP の振舞により決まってくるため，TCP パラメータなどを観測して iSCSI の性能との関連性を明らかにする必要がある．

2.3　TCP 輻輳ウィンドウ

TCP では，通信能力の制御にウィンドウサイズという概念を用いている．ウィンドウサイズとは，ホストが確認応答パケット (Acknowledgement:ACK) なしに一度に送信できるデータの量である．また，データの送信側では輻輳ウィンドウ，受信側では広告ウィンドウという値が決定され，このどちらか小さい方がウィンドウサイズとして用いられる．広告ウィンドウは現在の受信バッファの空き容量を示しており，ACK のヘッダにその情報が含められて送信側に送られる．一方，輻輳ウィンドウは送信側の制御パラメータで，ネットワークの混雑を回避するため送信側が自主的に制限する値である．

輻輳制御ではこの輻輳ウィンドウが利用されている．輻輳制御はネットワークの混雑解消の方法として TCP が行う機能である．通信開始時にはスロースタートと呼ばれるアルゴリズムに従って指数関数的に輻輳ウィンドウが大きくなる．これによりトラフィックが急激に増加するので，ネットワークが輻輳状態になる可能性がある．これを防ぐため，スロースタート閾値という値を用意し，輻輳ウィンドウがその大きさを超えると輻輳回避と呼ばれるフェーズに入り，一次関数的な増え方となる．そしてエラーが検出されると輻輳ウィンドウは急激に低下し，通常これらを繰り返すことで鋸型のグラフとなる．

また本実験で用いた LinuxOS における TCP の状態遷移を図 5 に示す．LinuxTCP においては，通信時の状態が正常であれば ACK の受信ごとに輻輳ウィンドウは増加するが，エラーが検出されると異常と判断され，輻輳ウィンドウは低下する．輻輳ウィンドウが低下する原因としては，送信側デバイスドライバのバッファが溢れることによる Local Congestion エラーを検出した場合 (CWR)，重複 ACK 又は SACK を受信した場合 (Recovery)，タイムアウトを検出した場合 (Loss) の 3 つが挙げられる．さらに Linux の TCP 実装では，通信中に一度設定された輻輳ウィンドウは，そのウィンドウ値を超えるデータ量が送られない限りは変化しないという特徴を持ち，この時スループットはほぼ一定の値で安定することが確認されている．

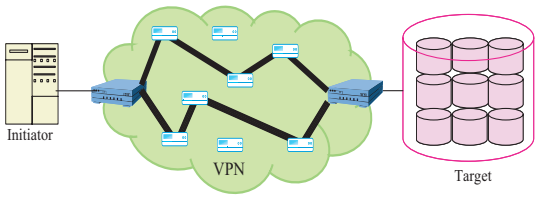

図 6　iSCSI 並列ネットワークストレージシステム

3.　並列ネットワークストレージシステム

3.1　並列ストレージシステム

近年，大規模なストレージシステムは多数のストレージノードをネットワーク結合した並列ストレージシステムとして構成される．そしてデータ郡をファイル，エクステント，ページ，ブロック等の粒度で分割を行い，分散格納する．データ配置管理の複雑さを利用者から隠蔽し，アクセス性能を最適化するため，システムはメタデータサーバや分散ディレクトリなどによるストレージ仮想化機構により仮想化され，単一の巨大ボリュームとして提供される．各クライアントからのアクセスリクエストは仮想化機構を通じて適切なストレージノードへと送られて処理される．並列ストレージシステムでは，ノード故障によるデータ喪失を防ぐため，他のノードのバックアップデータを保持する構成をとるものが多い．

並列ストレージシステムの研究において，これまでに可用性やスケーラビリティに優れた高機能並列ストレージシステムである自律ディスクが提案されている [8] [9]．自律ディスクのデータ配置戦略において，データマイグレーションに起因するノード負荷上昇を効果的に分散するために複製データを利用し，マイグレーション移動元移動先ノードが性能要件を満たさない状態に陥ることを防ぐ手法 Replica-assisted Migration が提案されている．自律ディスクでは，Primary-Backup の one-copy 複製により管理する障害復旧用バックアップデータを Chained Declustering 複製配置戦略に基づきストレージノード間で保持し合うことで信頼性を高めている．

3.2　iSCSI 並列ネットワークストレージシステム

前節で述べた並列ストレージシステムでは，複数のストレージデバイスが束ねられ単一の巨大なボリュームを持つストレージとして提供される．これを iSCSI のようなネットワークストレージで Target として利用しようとした場合，実際には複数の Target が存在するが，サーバ側からは単一のデバイスを持つストレージとして見えるように構成することが可能である．

Target に並列ストレージのように性能の高いストレージシステムを用いた場合，ネットワークストレージのボトルネックはネットワーク部分となる可能性が高い．従ってネットワーク部分の能力を高めることが，全体の性能を上げるために重要である．このためには，2.2 節で述べた VPN マルチルーティングの

ように複数経路接続の手法が有効である．

従って本論文では図 6 に示すように，Target に並列ストレージシステムを利用した際に，ネットワークを VPN 複数経路接続とすることにより性能向上を図る手法を提案し，その性能評価を行う．各 Target は単一のデバイスを持ち，サーバも単一のインタフェースからアクセスを行っているにも関わらず，ネットワーク部分は複数経路化され，Target も複数の Target の集合から構成される並列ストレージシステムとなっている．

3.3 本研究の評価内容と関連研究

本研究では以上に紹介した知識と技術を利用し、VPN 複数経路接続時における iSCSI 並列ストレージアクセスの性能評価を行う．関連する研究としては，以下のようなものが挙げられる．

iSCSI と TCP の関係を評価した研究としては，iSCSI のソフトウェア実装を使用した場合と，TCP の性能を向上させる TOE (TCP Offload Engine) や HBA (Host Bas Adaptor) を使用した場合との比較などが行われている [10]．またファイル操作やベンチマークプログラム実行時における iSCSI と NFS との比較も報告されている [11]．一方 iSCSI 層と TCP 層の間の処理を最適化することにより，特別なハードウェアを使用せず iSCSI プロトコル処理スループットの性能を向上させる手法が提案されている [12]．iSCSI の Initiator や Target の実装について，詳細に解析を行った結果なども報告されている [13] [14]．

iSCSI ストレージアクセスにおいて TCP 輻輳ウィンドウを制御する研究としては，輻輳ウィンドウ値を動的にコントロールする手法がある [15]．この手法は，まず Target の OS のカーネルに輻輳ウィンドウモニタ関数を挿入し，これによりモニタした輻輳ウィンドウの変化を観察して，Initiator にその値を通知する．通知を受けた Initiator は輻輳ウィンドウの値に基づきブロックサイズを再指定して，シーケンシャルリードアクセスを行うというものである．この手法を適用し輻輳ウィンドウを限界値で一定に保った場合には，高遅延環境において最大 28% のスループットの向上が確認されている．またトランスポート層を標準的な TCP Reno から他の TCP に置き換えて iSCSI に適用した際のシミュレーション結果なども報告されている [16]．

一方 iSCSI 複数コネクションに関する研究として，広域 IP 網を介した長距離アクセス向けに iSCSI および関連プロトコルレイヤのプロトコルチューニングの検討が行われ，その有効性が確認されている [17]．また TCP 複数コネクションにおいて，各々の TCP 制御情報をコネクション間で共有することにより複数コネクション間の公平性を保つ Fair-TCP を，iSCSI 複数コネクションに適用した結果も報告されており，情報を共有することによって複数コネクションにおける制御がより適切なものとなり，性能向上が見られた [18]．しかし iSCSI 複数コネクションを複数の経路に割り当てて利用する手法に関しては，これまで検討が行われてこなかった．

そこで我々はこれまでに，iSCSI 複数コネクションと VPN 複数経路を利用することで，1 つの iSCSI アクセスを 2 つの経路を用いて通信させたときの性能評価と TCP 輻輳ウィンドウの振舞を観察，解析した [19]．その結果，iSCSI 複数コネクション VPN 複数経路で接続したときのスループットは向上し，また各コネクションごとの輻輳ウィンドウは独立の振舞をすることが確認された．

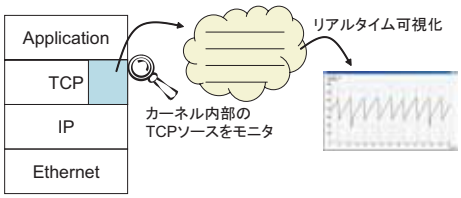

図 7　TCP 輻輳ウィンドウモニタツール

本研究では，より高機能，高信頼なネットワークストレージを実現することを目的に，iSCSI ストレージノードである Target 2 台を RAID0 で構築し，この iSCSI 並列ストレージシステムにおいて iSCSI セッションを VPN 複数経路に乗せてアクセスする手法の評価を行った．このような組み合わせを実現することにより，より高い性能が期待できると共に，VPN 複数経路 iSCSI アクセスにおけるボトルネックなどの詳細を明らかにすることが可能となる．

4. 実験システム

4.1 TCP 輻輳ウィンドウモニタツール

本実験では，TCP 輻輳ウィンドウをモニタするツールを構築した．図 7 に示すように，カーネル内部の TCP ソースにモニタ関数を挿入しカーネルを再コンパイルした．ここでモニタできるようになったものには，輻輳ウィンドウの値の他，各種エラーイベントの発生 (Local device congestion，重複 ACK，SACK 受信，タイムアウト検出) などがある．また X11 ウィンドウシステムライブラリ関数を用いて，モニタした値をリアルタイムに可視化することもできる．このツールを用い，TCP 輻輳ウィンドウの振舞などを解析することにより，iSCSI 複数経路アクセスの評価を行った．

4.2 複数経路並列ストレージアクセス実験システム

本実験では，VPN ルータを用いて複数経路を構成し，iSCSI 並列ストレージアクセスを実行した時の，性能と輻輳ウィンドウを評価するために図 8 に示す実験環境を構築した．iSCSI ストレージアクセスを行う Initiator とストレージを提供する 2 台の Target の間に VPN ルータを 4 台を挟み，複数経路アクセスが実行できるように構築した．さらにそれぞれの経路に，遠距離アクセスを想定して人工的な遅延装置である FreeBSD Dummynet を挿入した [20]．

Initiator と Target には，OS は Linux2.4.18-3，CPU は Intel Xeon 2.4GHz，Main Memory は 512MB DDR SDRAM，NIC は Intel Pro/1000MT Server Adapter on PCI-X (64bit,133MHz)，iSCSI は UNH IOL reference implementation ver.3 on iSCSI Draft 18 を用いた．Dummynet1 には FreeBSD4.9-RELEASE，Dummynet2 には FreeBSD6.2-RELEASE を用いた．また VPN ルータには Fujitsu Si-R570 を用いた．これは 3DES 暗号化速度最大 500Mbps を実現する．

ここで iSCSI ストレージの 2 台の Target を用いて RAID0 を構成し iSCSI 並列ストレージシステムを構築した．この実験シ

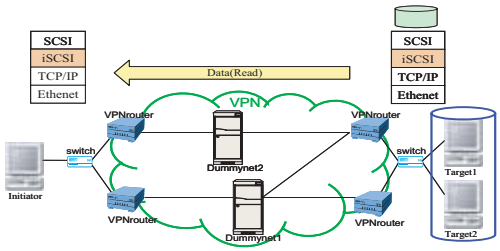

図 8　実験システムの概要

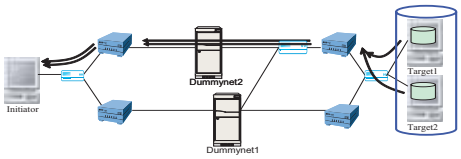

図 9　VPN 単数経路

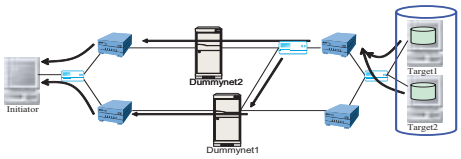

図 10　VPN 複数経路

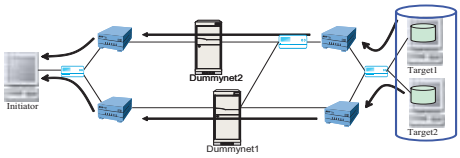

図 11　VPN 独立複数経路

ステムにおいて VPN ルータのマルチルーティング機能を用いて遠隔ストレージアクセスの通信制御を行った．以下では iSCSI リードアクセス，すなわち Target から Initiator へデータが転送される場合について説明しているが，逆も基本的に同じである．

まず VPN 単数経路通信の場合は，図 9 に示す通り同一経路上を 2 つのセッションが張れるように設定した．一方 VPN 複数経路通信の場合は図 10 に示す通りセッションごとに経路が異なるように VPN ルータの設定を行った．このとき Target から送られるパケットは，右上の VPN ルータに送られるように経路を設定する．そして VPN ルータのマルチルーティング機能により，左の 2 つの VPN ルータ宛にパケットが転送される．このとき，ポート番号の違いにより iSCSI セッションごとに上下の VPN ルータに分かれるように設定した．また通常並列ネットワークストレージにおいて通信の口（アクセスインタフェース）は 1 つであり，単一のルータが入口となっていることが想定されるが，参考として図 11 に示す通り 2 つの送受信口がある独立複数経路の場合の性能も測定し比較対象とした．

5.　性能測定結果

iSCSI 並列ストレージアクセスにおいてブロックサイズを変化させた時の各アクセス手法によるスループット比較のグラフを図 12 に示す．この実験において遅延は設定していない．またグラフ中の Target1 は図 8 における Target1 単一のストレージへ

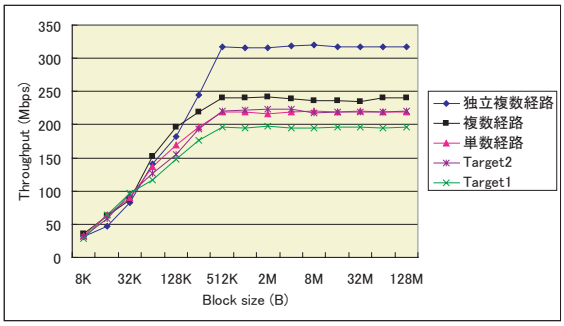

図 12　ブロックサイズとスループット比較

のアクセスを行った場合の性能を示し，Target2 も同様である．

どの場合でもブロックサイズを増加していくとスループットは増加していくが，ブロックサイズ 512KB を過ぎたところからスループットはほぼ一定となる．

次にアクセス方法の違いによる性能変化を見てみる．Target1 や Target2 と単数経路の性能はほぼ同一となっている．このように 2 台の Target を RAID0 としストレージ側では性能が向上しているにも関わらず，ネットワークなど他の部分がボトルネックとなってしまっていることがわかる．

また単数経路の場合と複数経路の場合を比較すると複数経路の性能が向上しており，本研究で用いた VPN 複数経路接続によるアクセス手法は有効な手段であることがわかる．性能向上の理由としては，VPN 内の経路が 2 倍の容量を持つようになり，さらに Initiator 側の VPN ルータでの処理が分散されたためであろう．一方で，独立複数経路を用いた場合に比べ性能が抑えられてしまっている理由としては，Target 側での VPN ルータにおける暗号化処理等がボトルネックとなっていると考えられる．しかし現在 VPN ルータは急速に進歩しており，VPN ルータの性能を十分に利用できる本アクセス手法は今後より有効な手段になると考えられる．

次に，片道遅延時間を変化させた時の iSCSI 並列ストレージアクセススループット比較のグラフを図 13 に示す．この実験においてブロックサイズは 2MB に設定した．また複数経路において経路ごとの遅延時間は同じ値を設定した．

どの場合も片道遅延時間を長くするとスループットは急激に減少した．また，各アクセス方法における変化の様子を見てみると，遅延時間が短い時には複数経路は単数経路の性能に近い値をとっているが，遅延時間を長くすると独立複数経路の性能に近い値をとるようになっている．これは遅延時間が短い時は複数経路にしても Target 側の VPN ルータでの処理がボトルネックとなってしまっていたため，独立複数経路の場合ほど性能は良くなかったが，高遅延環境の場合になると VPN ルータでの処理より経路長による性能劣化の方がボトルネックとなり，独立複数経路の性能と差がなくなってきたと考えられる．このように本実験環境においては VPN ルータの性能がボトルネックとなっていることが確認でき，今後 VPN ルータの進歩により，独立複数経路に近い性能を出すことが可能になるだろう．

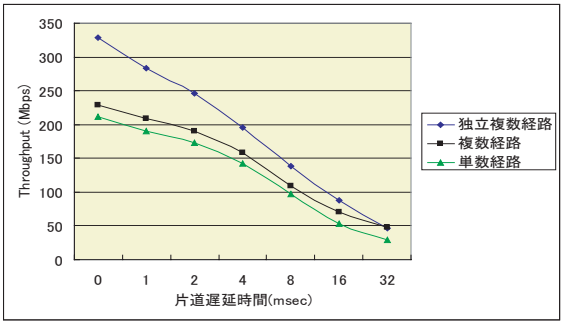

図 13　片道遅延時間とスループット比較

6.　輻輳ウィンドウの比較による解析

TCP 輻輳ウィンドウのグラフを元に，VPN 複数経路における iSCSI 並列ストレージアクセスの解析を行う．本実験における複数経路で輻輳ウィンドウを確認する場合，iSCSI セッションごとに Target1 と Target2 の 2 つの輻輳ウィンドウがそれぞれ独立に変化する．そこで図 8 の実験システムにおける Target1 と Target2 において TCP 輻輳ウィンドウモニタツールを実行し，2 台の Target から得たグラフを合わせて考える必要がある．またこの時のブロックサイズは 2MB に設定し，遅延は入れていない．

図 14，図 15 は複数経路を構築した時の輻輳ウィンドウの振舞の様子を示している．Target1，Target2 はほぼ同一の変化をしており，初め急速に輻輳ウィンドウは増加するがその後なだらかな増加になり，輻輳ウィンドウが約 190 パケットのところで一定の値をとるようになった．ここで LinuxTCP 輻輳ウィンドウの特徴により，輻輳ウィンドウが一定となるのは輻輳ウィンドウ値を超えるデータ量が送られていないからである．つまり輻輳ウィンドウが約 190 パケットになると輻輳ウィンドウ値は十分に大きく，本実験システムにおいて TCP プロトコル処理を行っている Target ではない他の部分での性能が通信のボトルネックとなっていることがわかる．

この時の通信のボトルネックを調べるため，独立複数経路の場合である図 16，図 17 との比較を行う．独立複数経路の場合は複数経路の場合と比べると実行時間は短くなっており輻輳ウィンドウの増え方も早くなっているが，ほぼ同様な変化をしている．つまり Target 側の VPN ルータによる処理の分散により輻輳ウィンドウはより早く増加するようになるが，一方で輻輳ウィンドウを使い切るほどのデータ送信が行われていないことがわかる．すなわち Target におけるプロトコル処理速度に比べて Target 側の VPN ルータを複数にしてもなおネットワークの転送速度は十分ではなく，これは複数経路のトラフィックによる負荷が Initiator におけるプロトコル処理や RAID 処理などにボトルネックを生じさせていることによると考えられる．

最後に，RAID 構築を行っていない単一の Target1 へのアクセスの場合である図 18 と比較を行う．この時輻輳ウィンドウは約 380 パケットになるまで急速に増加し，その後一定になる．このグラフと複数経路の場合のグラフを比べると，一定となる値がおよそ 2 倍になっていることがわかる．つまり，RAID 構築

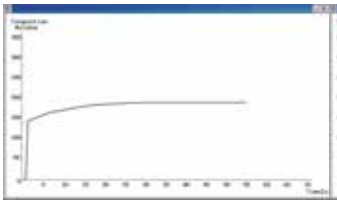

図 14　複数経路 (Target1)

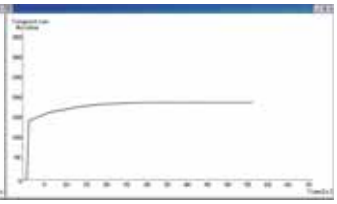

図 15　複数経路 (Target2)

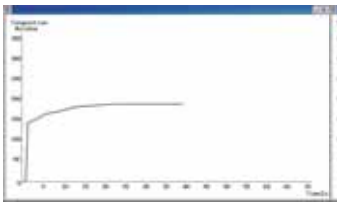

図 16　独立複数経路 (Target1)

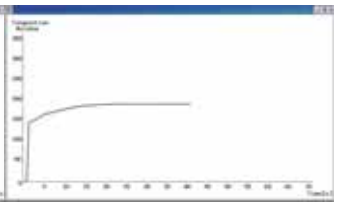

図 17　独立複数経路 (Target2)

図 18　Target1

複数経路における Target1 と Target2 の輻輳ウィンドウ値を足し合わせた値になっている．複数経路の場合においては，Target1 と Target2 の 2 つの iSCSI セッションを用いて RAID 構築しているため，輻輳ウィンドウの成長は止まりネットワークの帯域幅を十分に使い切れていない．したがって，より高機能な並列ストレージシステムを使うことにより更なる性能向上が期待できる．

7.　ま　と　め

本研究では，iSCSI 並列ストレージアクセスにおいて VPN 接続を単数経路，複数経路に変化させた時のスループットの違いと輻輳ウィンドウの振舞を観察し，比較した．

その結果，単一の Target を用いた性能と 2 台の Target を RAID で構築し単数経路でアクセスした性能はほぼ同一となっていた．一方本研究で用いた VPN 複数経路接続時においては性能が向上し，複数経路にすることで信頼性だけでなく性能向上も図れることがわかった．

また遅延時間の短い時には複数経路は単数経路の性能に近い値をとっているが，遅延時間を長くすると独立複数経路の性能に近い値をとるようになった．これは高遅延にすることで VPN ルータでの処理より経路のほうがボトルネックとなってきたためであると考えられる．このように，VPN ルータの進歩により本アクセス手法は今後より有効な手段になると考えられる．

輻輳ウィンドウの振舞を比較すると，複数経路の場合輻輳ウィンドウはなだらかに増加して一定になっている．ここで独立複数経路の場合と比較することにより，Initiator での処理や RAID での処理によって輻輳ウィンドウを使い切るほどのデータ送信

が行われなくなっていることがわかる．また単一の Target へのアクセスの場合における輻輳ウィンドウ値は，複数経路の場合の Targe1 と Target2 の輻輳ウィンドウ値を足した値とほぼ同じにあることがわかった．複数経路の場合においては，Target1 と Target2 の 2 つの iSCSI セッションを用いて RAID 構築しているため，ネットワークの帯域幅を十分に使い切れていないということがわかる．したがって，より高機能な並列ストレージシステムを使うことにより更なる性能向上が期待できる．

今後はより詳細な解析，評価を行うとともに，ネットワーク遅延やストレージ性能などに対応した有効なアクセスコントロール手法を提案していきたい．

謝　　辞

本研究は一部，独立行政法人科学技術振興機構戦略的創造研究推進事業 CREST によるものである．

文　　献

[1] Internet Small Computer Systems Interface (iSCSI), http://www.ietf.org/rfc/rfc3720.txt?number=3270

[2] SCSI Specification, http://www.danbbs.dk/ ˜dino/SCSI/

[3] 山口実靖，小口正人，喜連川優: 高遅延広帯域ネットワーク環境下における iSCSI プロトコルを用いたシーケンシャルストレージアクセスの性能評価ならびにその性能向上手法に関する考察，電子情報通信学会論文誌，Vol.J87-D-I，No.2，pp.216-231，2004

[4] InterOperability Lab, Univ,of New Hampshire, http://www.iol.unh.edu/consortiums/iscsi/

[5] 千島 望，豊田 真智子，山口 実靖，小口 正人: VPN 接続環境における TCP パラメータと通信性能の相関関係評価，FIT2006，L-042，2006

[6] 富士通 IP アクセスルータ GeoStream Si-R シリーズ GeoStream Si-R570，http://fenics.fujitsu.com/products/sir/sir570/index.html

[7] 千島 望，豊田 真智子，山口 実靖，小口 正人: iSCSI アクセス時の VPN 環境における TCP 輻輳ウィンドウ制御手法の検討，DBWS2006，pp.709-712，2006

[8] Haruo Yokota: Autonomous Disks for Advansed Database Applications, DANTE'99, pp.441-448, 1999

[9] 小林 大，田口 亮，横田 治夫: 並列ストレージにおけるサービス性能を保った複製利用負荷均衡化に対する更新リクエストの影響，DEWS2007，L2-1，2007

[10] P.Sarkar, S.Uttamchandani, and K.Voruganti: Storage over IP: When Does Hardware Support help?, *Proc. 2003 USENIX Conference on File and Storage Technologies (FAST2003)*, pp.231-244, 2003

[11] P.Radkov, L.Yin, P.Goyal, P.Sarkar, and P.Shenoy: Performance Comparison of NFS and iSCSI for IP-Networked Storage, *Proc. 2004 USENIX Conference on File and Storage Technologies (FAST2004)*, pp.101-114, 2004

[12] A. Joglekar, M. E. Kounavis, adn F. L. Berry: A Scalable and High Performance Software iSCSI Implementation, *Proc. 2005 USENIX Conference on File and Storage Technologies (FAST2005)*, pp.267-280, 2005

[13] 藤田智成，矢田浩二: iSCSI イニシエータの設計についての一考察，情報処理学会研究報告，2005-OS-99(21)，pp.135-140，2005

[14] 藤田智成，小河原成哲: iSCSI ターゲットソフトウェアシステムの解析，情報処理学会論文誌コンピュータシステム，Vol.46，No.SIG3(ACS8)，pp.38-50，2005

[15] 豊田 真智子，山口 実靖，小口 正人: iSCSI ストレージアクセスにおける TCP 輻輳ウィンドウコントロール手法の提案と性能評価，電子情報通信学会論文誌 Vol.J90-D，No.2，pp.359-372，2007

[16] G. Motwani and K. Gopinath: Evaluation of Advanced TCP Stacks in the iSCSI Environment using Simulation Model, *Proc. 22nd IEEE / 13th NASA Goddard Conference on Mass Storage Systems and Technologies (MSST2005)*, pp.210-217, 2005

[17] 藤原 啓成，若宮 直紀，志賀 賢太: 広域 IP 網を介した iSCSI 通信におけるプロトコルチューニングの一検討，第 68 回情報処理学会全国大会，pp.155-156，2006

[18] B. K. Kancherla, G. M. Narayan, and K. Gopinath: Performance Evaluation of Multiple TCP connections in iSCSI, *Proc. 24th IEEE Conference on Mass Storage Systems and Technologies (MSST2007)*, pp.239-244, 2007

[19] 千島 望，山口 実靖，小口 正人: VPN 複数経路接続時における iSCSI ストレージアクセスの特性解析，DICOMO2007，3H-4，2007

[20] L.Rizzo: dummynet, http://info.iet.unipi.it/~luigi/ip_dummynet/