

iSCSI 遠隔ストレージアクセスにおける VPN 経路接続を用いた場合の性能に与える影響の評価

浅田 菜那[†] 比嘉 玲華[†] 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

E-mail: †{nana-asa,reika}@ogl.is.ocha.ac.jp, ††oguchi@computer.org

あらまし コンピュータシステムにおけるデータ量の増大に伴い、効率的にストレージを管理したいという要望が高まっている。またストレージの管理コスト低減などの目的で SAN の導入が進んでいる。現状では、SAN は主にローカル環境のみで用いられている。そこで IP ネットワークを利用した IP-SAN として iSCSI が注目されている。iSCSI を用いることにより広域環境における IP-SAN を低コストで構築でき、遠隔地のデータセンタなどにデータをバックアップすることが容易となるため、ストレージのアウトソーシングといったサービスへの利用が可能になるためである。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させ、高遅延環境における iSCSI ストレージアクセスの特性、解析を評価しスループット向上の方法について検討する。

キーワード iSCSI, 広域ネットワーク, VPN, ネットワークストレージ

An Examination of Performance of iSCSI Remote Storage Access using VPN connection

Nana ASADA[†], Reika HIGA[†], and Masato OGUCHI[†]

[†] Ochanomizu University Otsuka 2-1-1, Bunkyo-ku, Tokyo, 112-8610 Japan

E-mail: †{nana-asa,reika}@ogl.is.ocha.ac.jp, ††oguchi@computer.org

Abstract As the volume of data computer systems process increases, it is important that storage is managed efficiently. The introduction of SAN progresses for the purpose of the storage management cost reduction. SAN is mostly used in the server site currently. iSCSI is expected as IP-SAN that uses IP network. Because iSCSI can configure the wide area IP-SAN with low cost and perform data backup to remote place easily, for example data center, iSCSI is expected to be used as Storage Outsourcing service. In this research, we have first applied iSCSI used in a local environment in the WAN using VPN router. Second, we evaluated performance in the case of iSCSI storage access in WAN and examined the method to improve throughput.

Key words iSCSI, Wide Area Network, VPN, NetworkStorage

1. はじめに

近年、インターネット技術の進展などにより、個人のユーザや企業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 とは、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、また距離に制約がある。これに対し、SAN に IP ネットワークを利用した IP-SAN として iSCSI[3] が期待されている。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の発展により、ギガビットクラスの回線実現が期待され、iSCSI の有効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使用されている。そこで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 で接続することが期待されて

いる。そのような背景をふまえて本研究では、拠点間接続等に用いられる 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。

VPN 環境ではルータにおけるパケット暗号化処理等により、通常のネットワークとはトラフィックの性質が異なったものになると考えられる。そこで本研究では、広域ネットワークを遅延装置で模擬したネットワーク上に VPN を張った環境において iSCSI 遠隔ストレージアクセスを実行した際の性能について検討を行う。

本稿の構成は以下の通りである。2 章で研究背景として iSCSI, VPN について述べ、3 章で既存研究についてまとめる。4 章で本研究の実験概要をについて記述する。5 章で測定結果と考察を示し、6 章で iSCSI パラメータ最適化について、7 章でソケットバッファに関するカーネルコード書き換えについての結果と考察を述べる。最後に 8 章でまとめと今後の課題を述べる。

2. 研究背景

2.1 iSCSI

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 機器だけで構成することが可能となる。また iSCSI は図 1 のように、SCSI over TCP/IP over Ethernet という複雑な階層構造のプロトコルスタックとなる。そのオーバーヘッドなどが影響し、iSCSI による通信は特に高遅延環境においては大幅に性能が劣化することがわかっている [1]。そこでそのような環境におけるスループットの低下を改善することが求められている。

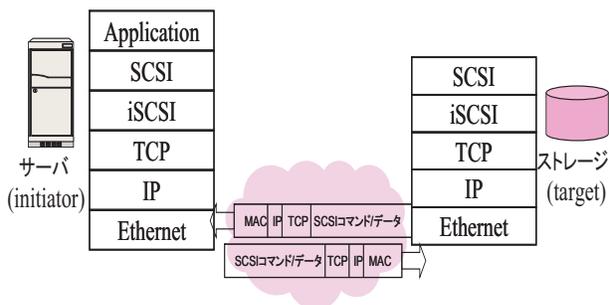


図 1 iSCSI のプロトコルスタック

2.2 VPN(Virtual Private Network)

VPN の構成例を図 2 に示す。VPN は公衆回線をあたかも専用回線のように利用できる仕組みである。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用い、拠点間を仮想的に閉じたネットワークで接続する。今日では、企業内ネットワークの拠点間接続などに使用されている。VPN の大きなメリットは専用回線を導入するよりもコストを抑えられることである。一方、公衆網のデメリットとして機密性が低いことがあげられるが、VPN では暗号化や認証などを用いて、実質的な専用網を実現する。本研究では広域ネットワーク環境における iSCSI の振舞を解析するため、VPN ルータを使用する。

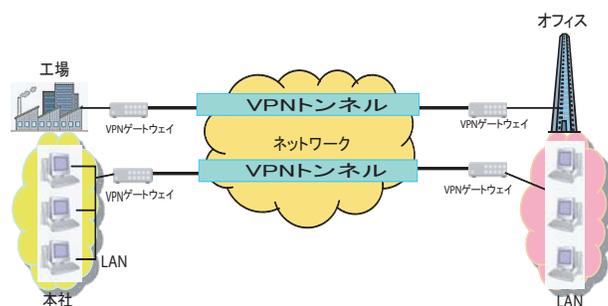


図 2 VPN 使用例

3. 既存研究

我々は、これまでに 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ときのスループットの違いを観察するという実験を行った [2]。

この実験によって、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ことがわかっている。

この既存研究と本研究との最も大きく異なる点は実験環境のソフトウェア環境である。この既存研究において Initiator と Target には、OS は Linux2.4.18-3, CPU は Intel Xeon 2.4GHz, Main Memory は 512MB を使用しており、iSCSI は UNH IOL reference implementation ver.3 on iSCSI Draft 18, VPN ルータは Fujitsu Si-750 を用いた [4]。既存研究は iSCSI の性質を調べることにに関して成果を挙げたが、使用した iSCSI ドライバは試験的な実装であり、その性能は高いものではなかった。それに対し、本研究では Initiator と Target の OS は Linux2.6.18-8.e15, iSCSI は Initiator に Open-iSCSI, ターゲットに iSCSI Enterprise Target を使用しており、ローカル環境では十分に高い性能を発揮できる iSCSI 実装である [5]。

また、我々は iSCSI 遠隔ストレージアクセス時のパケット送信に関する考察についても研究を行ってきた。この実験において、iSCSI パラメータ、輻輳制御アルゴリズム変更、NIC パラメータの複数の層にまたがる最適化とデータ転送部分における解析、ソケットバッファ解析という 3 つの実験を行った結果、RTT20ms においてデフォルト時よりも約 7 倍スループットがあがり、大幅に性能を向上させることが出来た。ただしこの実験において、Initiator と Target の間には遅延装置しか挟んでおらず、遠隔アクセスの往復遅延時間が長くなると性能にどのような影響が見られるかという評価およびその際の性能向上のみに焦点が当てられてきた。しかし現実の広域ネットワーク環境においては遠隔アクセスの両端が長いケーブルのみで接続されている訳ではなく、間の接続に VPN などを用いることが一般的である。そこで、本研究では遅延装置だけでなく VPN を使用することによる、より実用性を考慮した手法の提案を行う。

4. 実験概要

本章では、本実験で使用した測定ツール、実験環境および実験手順を示す。

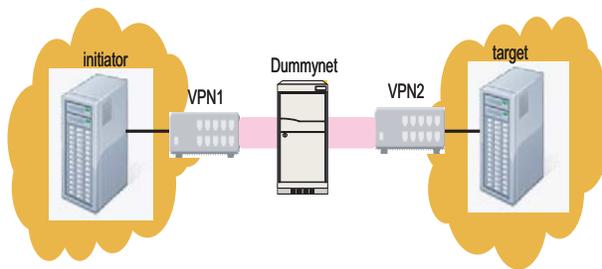


図 3 VPN 遠隔接続環境

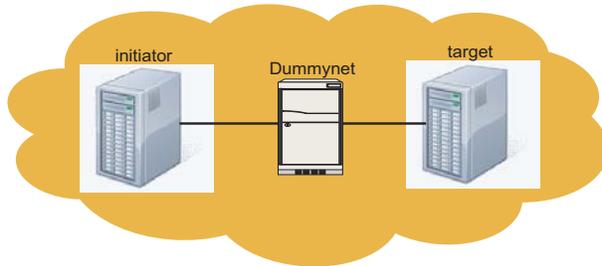


図 4 VPN を使用しない遠隔接続環境

Initiator と Target の OS は Linux2.6.18-8.e15, CPU は Intel Xeon 2.4GHz, Main Memory は 512MB を使用しており, iSCSI は Initiator に Open-iSCSI, ターゲットに iSCSI Enterprise Target を用いた。また VPN ルータには Fujitsu Si-750 を用いた。これは 3DES 暗号化速度が最大 500Mbps である。

4.1 Bonnie++

ハードディスクベンチマークツールとして, Bonnie++1.03 を用いた [6]。これはデータベースのような大規模なファイル操作のスループットを測定することができる。また比較的小さなファイルの作成, 読み込み, 削除のスループットも測定可能である。本研究では, これを用いて Sequential Write(連続書き込み), Sequential Read(連続読み込み) のスループットを測定した。また, ローカルハードディスクのアクセス性能を測定し, iSCSI アクセスの場合と比較した。

4.2 システム構成

本実験では VPN 遠隔接続環境と VPN を使用しない遠隔接続環境の 2 つのシステム構成を作り, それぞれにおいて連続書き込み, 連続読み込みを測定した。

4.2.1 VPN 遠隔接続環境

はじめに図 3 に示す実験環境を構築した。VPN ルータを 2 台用いて一方の LAN 空間には Initiator を, もう一方の LAN 空間には Target を接続させた。2 台の VPN ルータの間には, 広域ネットワークを想定して人工的な遅延装置である Dummynet を挿入した。Dummynet には FreeBSD 4.9-RELEASE を用いた。

4.2.2 VPN を使用しない遠隔接続環境

次に図 4 のように VPN ルータを外し, 同じアドレス空間のネットワークで Initiator と Target を接続させた。図 3 の場合と同様に Initiator と Target の間に Dummynet を挟み, 測定した。

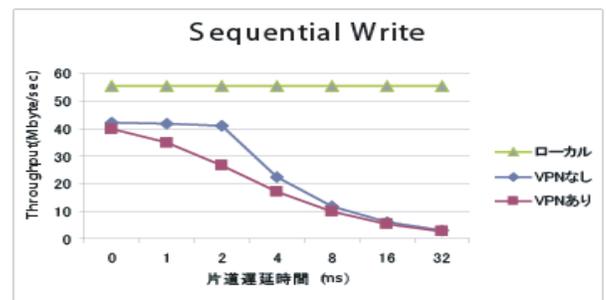


図 5 連続書き込みのスループット

5. 実験結果と考察

実験概要で述べたような環境において Sequential Write におけるスループットの比較を図 5 に, Sequential Read におけるスループットの比較を図 6 に示す。遅延時間が大きくなるにつれスループットは低下しているため, 性能が落ちていることがわかる。VPN ルータを挟んだ場合と挟まない場合を比較したところ, 連続書き込み, 読み込みどちらの場合においても全体的にルータを設定しない環境におけるスループットが高くなった。ただし連続書き込みの方は VPN ルータを挟まない場合にもローカルアクセスの場合と比較してスループットは低めであり, VPN ルータを挟んだ場合もあまり大きな違いは出ていない。また片道遅延時間が大きくなるにつれ, ルータを設定する場合としない場合の差は縮まっている。片道時間 32ms における測定結果は連続書き込みと連続読み込みのどちらにおいてもほぼ同程度の数値となり, それぞれ 2 本のグラフにも差が見られない。これは遅延時間が短いときは Target 側の VPN ルータでの処理がボトルネックとなっていたため, VPN 経路接続をしない環境における結果の方がスループットが高いが, 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, VPN 接続環境でない場合における結果との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低遅延の場合には VPN ルータの性能向上が期待され, 高遅延になったら iSCSI の性能向上が期待される結果となった。

また Sequential Read についてはローカル環境におけるハードディスクアクセス性能のグラフが極めて高くなっているが, iSCSI を用いたグラフは VPN を使用したグラフ, 使用しないグラフのどちらも低くなっている。これは iSCSI のネットワーク通信処理のオーバーヘッドによるものであると考えられる。これに対し, Sequential Write では, 片道遅延時間が短い場合は iSCSI の性能がかなり高くなり, ローカルアクセスと比較しても性能の落ち込みは少なくなっている。しかし遅延時間が大きくなると性能の急激な低下が見られた。これらのことから, 特に VPN 遠隔接続環境での性能低下が著しいことがわかった。

6. iSCSI パラメータ最適化

次に本研究では, 高遅延環境での性能向上のため Write に関するパラメータをさまざまに変更し測定した。変更内容の表を

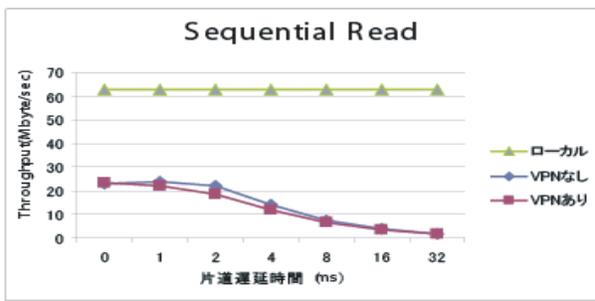


図 6 連続読み込みのスループット

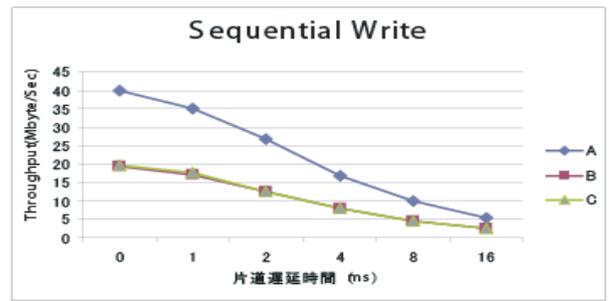


図 7 VPN 遠隔接続環境

表 1 iSCSI パラメータ最適化

		Writeに関するパラメータ		
		A	B	C
Target	Initiator	YES	NO	NO
	Immediate Data	NO	YES	YES
	FirstBurstLength	65536	524288	1048576
	MaxBurstLength	262144	524288	1048576
	MaxRecvDataSegmentLength	8192	524288	1048576
Initiator	node.conn[0].iscsiMaxRecvDataSegmentLength	131072	524288	1048576
	node.session.iscsiFirstBurstLength	262144	524288	1048576

表 1 に示す . A,B,C それぞれの場合において Bonnie++ で測定し、スループットが一番高くなる iSCSI パラメータ最適値を探した .

実験結果のグラフが図 8 と図 9 である . 図 8 は VPN 遠隔接続環境でのスループットの結果、図 9 は VPN を使用しない遠隔接続環境における結果である .

まず VPN 遠隔接続環境では全体を通して、バースト長などを長くした B と C は A の約 1/2 の値のスループットとなった . 高遅延環境になるにつれ A と、BC の差は縮むもののパラメータ A のグラフが常に高い値であった .

次に VPN を使用しない遠隔接続環境において片道遅延時間 0ms ではパラメータ A,B,C はほぼ同じ値であり、高遅延環境になるにつれどのグラフも急激に性能が下がった . 全体を通してパラメータ A のグラフが常に高いという結果が出た .

VPN 遠隔接続環境と VPN を使用しない遠隔接続環境のどちらの場合でも、BC のグラフが A のグラフを越えることはどの遅延時間においてもなかったという結果からパラメータ A が最適値であるといえる . またパラメータ A はデフォルトで設定されていた値であることから、本研究で用いた iSCSI 実装はデフォルトでの状態において最も性能が良くなるということがわかった .

7. カーネルコード最適化

本研究ではソケットバッファに関するカーネルコードを書き換え最適化することにより、高遅延環境における性能向上を図った . その結果を図 9 に示す . グラフは VPN 接続環境において測定したシーケンシャルライトのカーネルコード書き換え前と書き換え後の比較である . 片道遅延時間 16ms においてスループットが 2 倍以上になることを確認できた . カーネルコードを最適化したことにより、データをバッファに入れる作業を

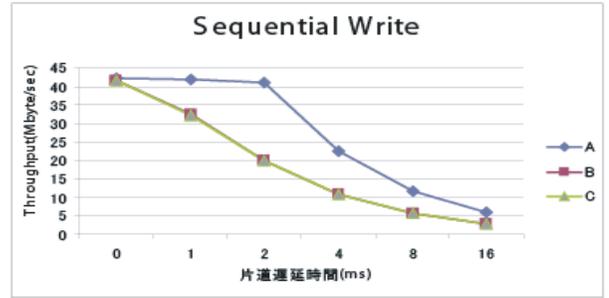


図 8 VPN を使用しない遠隔接続環境

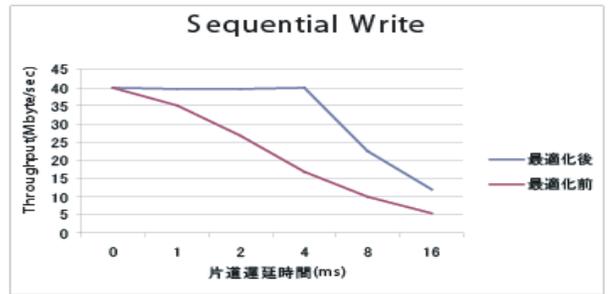


図 9 ソケットバッファに関するカーネルコードの書き換え

常に滞らないようになり TCP において ACK 待ちの状態を減らすことができたからであると考えられる .

8. まとめと今後の課題

本研究では、VPN 接続環境における iSCSI ストレージアクセス時のスループットの測定を行った . また VPN 接続をしない環境においても同様に測定し、それらの結果にどの程度差があるか調べた . 連続書き込み、連続読み込みどちらのスループットにおいても片道遅延時間に反比例するが、遅延時間が長くなると 2 つの異なる環境におけるスループットの差はほとんど無くなった . これより VPN 接続環境における iSCSI のリモートアクセスは VPN 接続しない場合に比べ、遠隔アクセス時の性能低下は少なく、遠隔アクセス時には iSCSI の通信プロトコル処理による性能低下が支配的であることが明らかになった .

また iSCSI のネットワーク通信処理のオーバーヘッドにより、Sequential Read において iSCSI を用いた場合にアクセス性能は下がった . 一方 Sequential Write では、片道遅延時間が短い場合は iSCSI の性能がかなり高くなるが、高遅延環境では特に VPN 接続環境において性能の大きな低下がみられた .

パラメータ最適化については様々な値を設定し測定した結果、最適値を見つけることが出来た。今後は本研究で見つけた最適値をパラメータに設定し実験を行っていく。

今後はより詳細な測定を行うと共に 1 対 1 通信のみではなく、1 対多、多対多の場合の性能評価も行いたい。Write の場合と同様、iSCSI の Read に関するパラメータを最適化することにより、今回測定した環境でどの程度スループットが変化するかを特定する。また既存研究でわかっている高遅延環境での iSCSI 性能向上の方法を本研究に取り入れることにより iSCSI 性能を向上させる手法を検討していく。

文 献

- [1] 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 年 2 月
- [2] 千島 望, 山口 実靖, 小口 正人: VPN 複数経路接続を用いた iSCSI ストレージアクセス性能評価, DICOMO2007, 3H-4, 2007 年
- [3] 比嘉 玲華, 岡廻 隆生, 山口 実靖, 小口 正人: Linux カーネルのログ解析によるパースト的 iSCSI 遠隔ストレージアクセス時のパケット送信に関する考察 DICOMO2009
- [4] 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>
- [5] 比嘉 玲華, 松原 幸助, 岡廻 隆生, 山口 実靖, 小口 正人: パケット解析と輻輳ウィンドウ解析による遠隔 iSCSI アクセスの断続的パケット送出に関する考察, DEIM2009, E1-1, 2009 年
- [6] bonnie++, <http://www.textuality.com/bonnie/intro.html>