

# VPN 遠隔ストレージアクセスの ハードウェアターゲット使用時の各層最適化に関する性能評価

浅田 菜那<sup>†</sup> 小口 正人<sup>††</sup>

<sup>†</sup> お茶の水女子大学 〒 112-8610 東京都文京区大塚 2-1-1

E-mail: <sup>†</sup>nana-asa@ogl.is.ocha.ac.jp, <sup>††</sup>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が、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、広域環境において IP-SAN を使用する場合の、実用的な環境を考慮した評価を行う。具体的には、ネットワークには遠隔サイトの接続に一般的に使用されている VPN を用いる。また iSCSI ターゲットは、ソフトウェアターゲットである iSCSI Enterprise Target だけでなく、FC ストレージと IP ブリッジを用いたハードウェアターゲットを使用する。このような実験システムを用い、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の検討を行う。

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

## An Evaluation of Each Layer Optimization of iSCSI Remote Storage Access Performance using a Hardware Target on VPN connection

Nana ASADA<sup>†</sup> and Masato OGUCHI<sup>††</sup>

<sup>†</sup> Ochanomizu Univesity 2-1-1 Otsuka, Bunkyo-ku Tokyo 112-8610 JAPAN

E-mail: <sup>†</sup>nana-asa@ogl.is.ocha.ac.jp, <sup>††</sup>oguchi@computer.org

### 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 は光ファイバ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 が期待されている [1]。iSCSI は、これまで DAS(Direct Attached Storage) で使われてきた SCSI コマンド [2] 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内線実現が期待され、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。これに対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 IP-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。

iSCSI ターゲットには、ソフトウェアターゲットとして iSCSI Enterprise Target [9] を用いた。それに加え市販されている FC ストレージと IP ブリッジを用いてハードウェアターゲット構成も導入することで、より現実的な環境における iSCSI 遠隔ストレージアクセスの性能向上を目指す。

## 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 [3]。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。

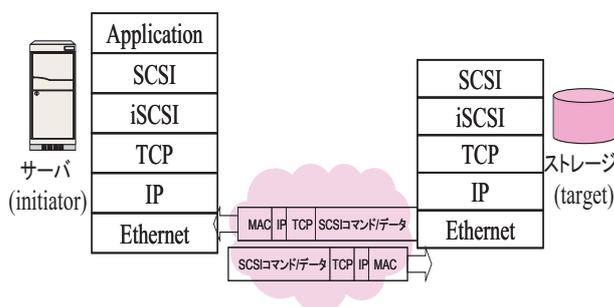


図 1 iSCSI のプロトコルスタック

### 2.2 iSCSI リモートストレージアクセス時における性能低下問題

iSCSI は接続距離に制約を受けないことや、安価に構築できるといった様々なメリットを活かして、高遅延環境での適用が望まれる。しかし iSCSI は、図 1 のように複雑な階層構成のプロトコルスタックで処理しており、パ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しいことが知られている。

本論文で用いた実験環境においても、後述する VPN 遠隔接続環境と VPN を使用しない遠隔接続環境の 2 種類の高遅延環境を構築し iSCSI の性能測定を行った。iSCSI を VPN 遠隔接続環境で測定した結果を示した図 2 によると、ソケット通信時においては高遅延環境下においてもほぼ性能を保っているのに対して、iSCSI 通信においては、高遅延環境になるにつれて性能が劇的に劣化していったという結果が得られた。

本実験環境においても、iSCSI 通信は高遅延環境下における急激な性能劣化という結果となった。リモートバックアップ

プロトコルとして iSCSI を使用するためには、この問題を解決しなければならない。

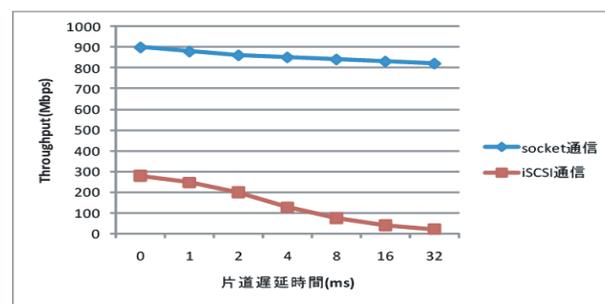


図 2 スループット比較

### 2.3 VPN(Virtual Private Network)

VPN の構成例を図 3 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網を用い、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。ただし通信品質については、利用しているインターネット等と同程度の低さとなる。本研究では広域ネットワーク環境における iSCSI の振舞いを解析するため、VPN ルータを使用する。

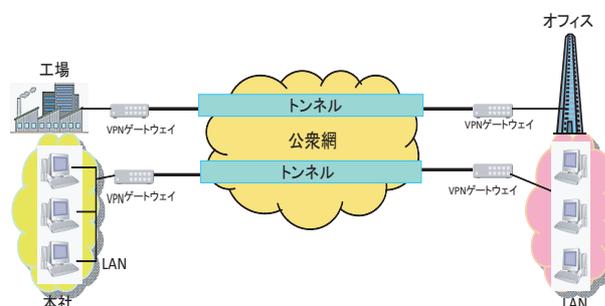


図 3 VPN 使用例

## 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 [4]。

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は研究目的が異なり、また実験環境も大きく異なる。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 のサーバ機を使用しており、iSCSI は UNH IOL reference implementation ver.3 on iSCSI Draft 18、VPN ルータは Fujitsu Si-570 を用いた [5]。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 [8], ターゲットに iSCSI Enterprise Target [9] 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 [6].

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、RTT=20ms においてデフォルト時よりも十分に高いスループットが観測され、大幅に性能を向上させることが出来た [6]. 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についての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環境において評価を行う。

#### 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。

##### 4.1 測定方法

本実験では、sg\_dd コマンドを用い [10], シーケンシャルライトアクセスのスループットを算出した。また、リモートバックアップを考えた場合、データの書き込み量と読み込み量を比較すると、圧倒的に書き込み量のほうが多い。以上のことを踏まえ、本研究においては、シーケンシャルライトアクセスとシーケンシャルリードアクセスのスループットに焦点を当てて実験を進めていく。

##### 4.2 システム構成

###### 4.2.1 iSCSI ソフトウェアターゲット構成 Type1, Type2

本実験では iSCSI ソフトウェアターゲットの実装環境として 2 種類のサーバ機を用いた。これをそれぞれ Type1, Type2 と呼ぶ。実験で使用した計算機の環境を表 1 に示す。なお Type2 で使用した計算機は既存研究 [6] で使用したものと同等のものである。

はじめに図 4 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 空間にはソフトウェアターゲットを接続させた。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。

図 5 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 4 の場合と同様に Initiator とソフトウェアターゲットの間に Dummynet を挟み、遠隔接続環境を模擬した。

表 1 iSCSI ソフトウェアターゲットシステム構成 Type1, Type2 における実験環境

OS	Red Hat Enterprise Linux 2.6.18-8.e15	
CPU	Type1	Intel Xeon 2.4GHZ
	Type2	Qued Core Intel Xeon 1.6GHZ
Main Memory	Type1	1GB
	Type2	2GB
NIC	Intel PRO/1000PT Server Adaptor on PCI Express	
iSCSI	Initiator	open-iSCSI-2.0-865
	Target	iSCSI Enterprise Target(IET)-0.4.15
VPN Router	Fujitsu Si-R570	
Dummynet	FreeBSB 4.9-RELEASE	

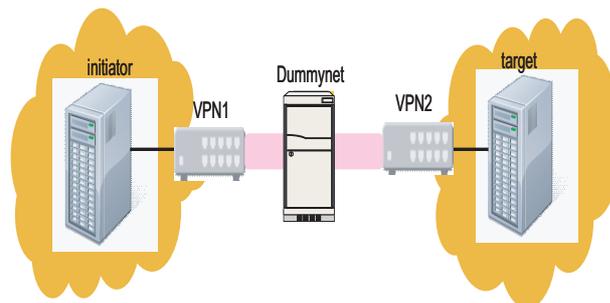


図 4 iSCSI ソフトウェアターゲットシステム構成 Type1, Type2 における VPN 遠隔接続環境

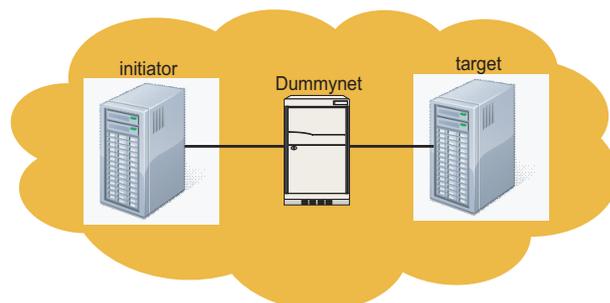


図 5 iSCSI ソフトウェアターゲットシステム構成 Type1, Type2 における VPN を使用しない遠隔接続環境

###### 4.2.2 iSCSI ハードウェアターゲット構成

実験で使用した計算機の環境を表 2 に示す。

表 2 iSCSI ハードウェアターゲットシステム構成における実験環境

OS	Linux2.6.18-8.e15
CPU	Intel Xeon 2.4GHZ
Main Memory	1GB
NIC	Intel PRO/1000PT Server Adaptor on PCI Express
iSCSI	Initiator : open-iSCSI-2.0-865
FC Storage	HP Storage Works2012fc
IP Bridge	ATTO iPBridge2700
VPN Router	Fujitsu Si-R570
Dummynet	FreeBSD 4.9-RELEASE

図 6 のように 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 IP ブリッジ, FC ストレージを接続させた。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。

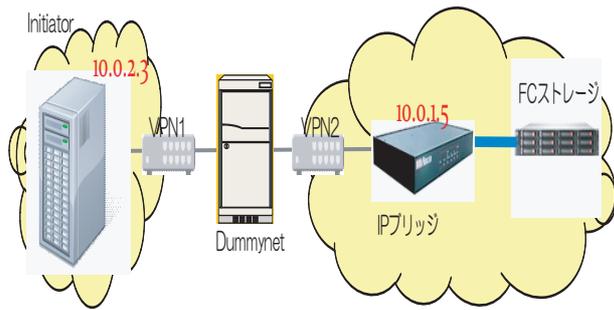


図 6 iSCSI ハードウェアターゲットシステム構成における VPN 遠隔接続環境

次に図 7 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 IP ブリッジ、FC ストレージを接続させた。図 6 の場合と同様に Initiator と IP ブリッジの間に Dummynet を挟み、遠隔接続環境を模擬した。

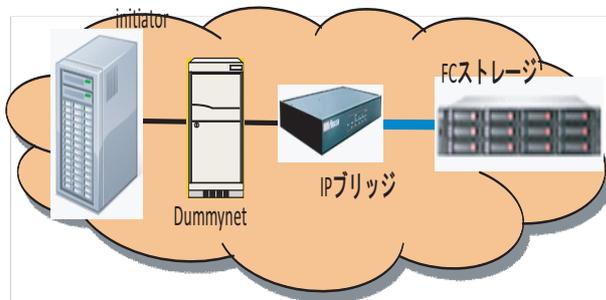


図 7 iSCSI ハードウェアターゲットシステム構成における VPN を使用しない遠隔接続環境

また使用している FC ストレージは RAID 構成を RAID0, RAID1, RAID5 に変えることができる。本研究では RAID0 をデフォルトとし測定し RAID1, RAID5 に変えた実験も行った。

## 5. iSCSI リモートアクセスの基本通信性能

図 8 は iSCSI ソフトウェアターゲットシステム構成 Type1 における、図 9 は iSCSI ハードウェアターゲットシステム構成におけるシーケンシャルライトアクセスのスループットの結果である。どちらの構成でも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、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。

しかし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、高遅延になった場合は iSCSI の性能向上が期待される結果となった。

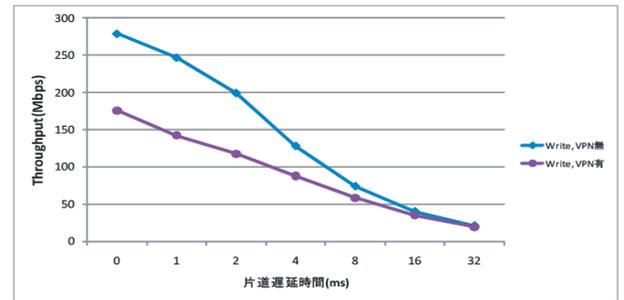


図 8 iSCSI ソフトウェアターゲットシステム構成 Type1【default】

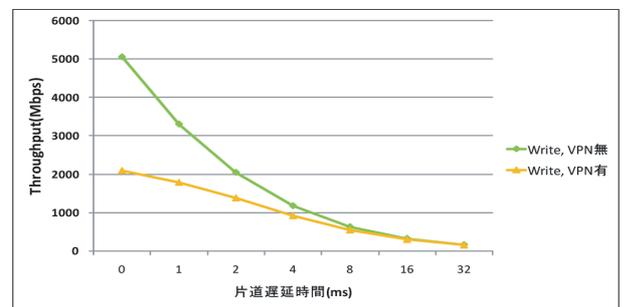


図 9 iSCSI ハードウェアターゲットシステム構成【default】

## 6. iSCSI 層における最適化

### 6.1 iSCSI パラメータ最適化

iSCSI はパラメータを最適化することによりスループットが向上することが確認されている。本実験においては、iSCSI のパラメータ設定をライトアクセス時における最適な状態になるように調整した。変更内容は表 3 の通りである。

表 3 iSCSI パラメータ設定

Writeに関するパラメータ		default	1024KB
Target	Initiator	YES	NO
	Immediate Data	NO	YES
	FirstBurstLength	65536	1048576
	MaxBurstLength	262144	1048576
	MaxRecvDataSegmentLength	8192	1048576
Initiator	node.conn[0].iscsiMaxRecvDataSegmentLength	131072	1048576
	node.session.iscsiFirstBurstLength	262144	1048576

### 6.2 iSCSI パラメータ最適化考察

#### 6.2.1 iSCSI ソフトウェアターゲットシステム構成 Type1

iSCSI パラメータ最適化を行った実験結果を図 10 に示す。iSCSI パラメータの最適化によって、VPN を使用しない遠隔接続環境ではデフォルト時と比較して片道遅延時間 32ms において約 2 倍の性能向上という結果を得ることが出来た。VPN 遠隔接続環境でも VPN を使用しない遠隔接続環境においても全体を通して性能が向上する結果となった。

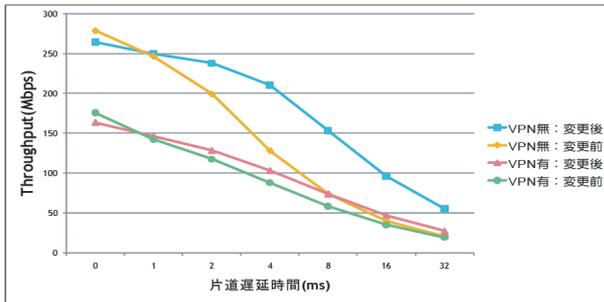


図 10 iSCSI ソフトウェアターゲットシステム構成 Type1【iSCSI 層最適化】

### 6.2.2 iSCSI ハードウェアターゲットシステム構成

iSCSI パラメータ最適化を行った実験結果を図 11 に示す。iSCSI パラメータの最適化によって、VPN 遠隔接続環境でも VPN を使用しない遠隔接続環境においても全体を通して性能が向上する結果となった。

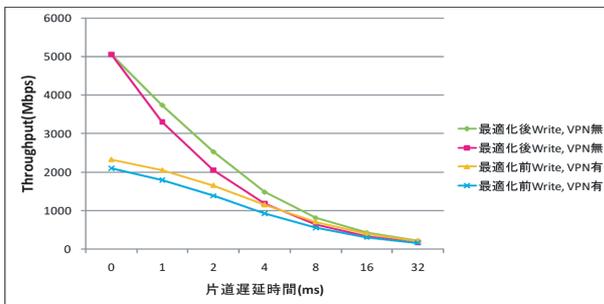


図 11 iSCSI ハードウェアターゲットシステム構成【iSCSI 層最適化】

全ての環境において iSCSI 層の最適化による性能向上は見られたが、まだなお高遅延環境下での劇的な性能劣化という問題は解決されていない。ソケット通信の場合は、高遅延環境においても高いスループットを保っていることが確認されている。したがって iSCSI 通信の場合も、高遅延環境において性能の劣化を極力防ぎ、スループットを保つことが目標となる。

## 7. RAID 構成変化

### 7.1 RAID0, RAID1, RAID5

iSCSI ハードウェアターゲットシステム構成において、RAID 構成を RAID1, RAID5 に変えスループットを測定し、比較した。結果を図 12 に示す。RAID 構成の違いが性能に影響する場合、RAID1 構成の性能が RAID0 構成の場合より低くなると考えられる。しかし本実験では RAID0, RAID1, RAID5 の実験結果に大きな差は見られなかった。

### 7.2 RAID 構成についての考察

RAID 構成の違いが性能に影響しているのであれば、RAID1, と RAID5 構成の場合は RAID0 構成の場合より性能が低くなると考えられる。しかし、本実験において RAID 構成の違いがスループットに大きな影響を与えることはなかった。このことから本実験環境においては RAID 構成の違いが性能に影響を与えていないと言える。この理由としては、ハードウェアター

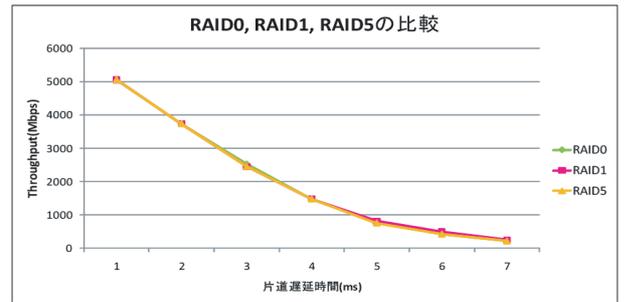


図 12 FC ストレージに RAID0, RAID1, RAID5 を用いた場合の比較

ゲットでは FC ストレージ内部で仮想化やマッピングなど複雑な構成を取っており (図 13)、それらの構成でどこかで性能が決まっています。RAID の性能差が見えなくなっているからであると考えられる。

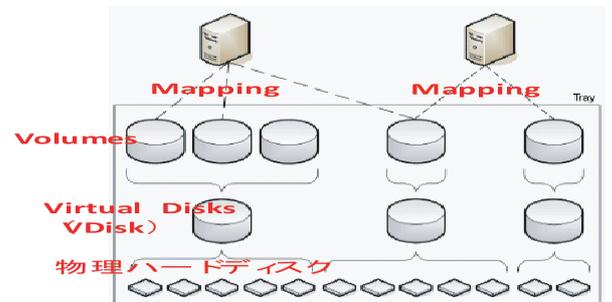


図 13 ハードウェアターゲット構成

## 8. TCP 層における最適化

まず、iSCSI ソフトウェアターゲットシステム構成 Type1 において RTT80ms の iSCSI アクセスを tcpdump コマンドを使用し解析した。VPN を使用しない遠隔接続環境における結果を図 14 に、VPN 遠隔接続環境における結果を図 15 に示す。

VPN を使用しない遠隔接続環境においては、短い時間に連続してパケットが送信された後、突然パケットの送出が止まっており、一定時間の後再びパケットの連続送信が行われている。その間隔は RTT である 80ms、または RTT の約 2 倍の値の 160ms であることが分かった。

VPN 遠隔接続環境における結果も同様に、短い時間に連続してパケットが送信されたあと突然パケットの送出が止まり、一定時間の後再びパケットの連続送信が行われているが、その間隔は RTT より 10ms 大きい 90ms と、RTT の 2 倍より 10ms 大きい 170ms であることが確認された。これは VPN ルータにおいて暗号化の為に一時的にパケットがバッファに蓄えられることによる効果であると考えられる。また一度に送られるパケットの量も少し増えたことも、バッファによる効果だと考えられる。

いずれの場合も、このように送信すべきデータがあるにもかかわらず、送信されることなくパケット送信の断続が生じていることが、高遅延環境における性能劣化の原因として考えることが出来る。

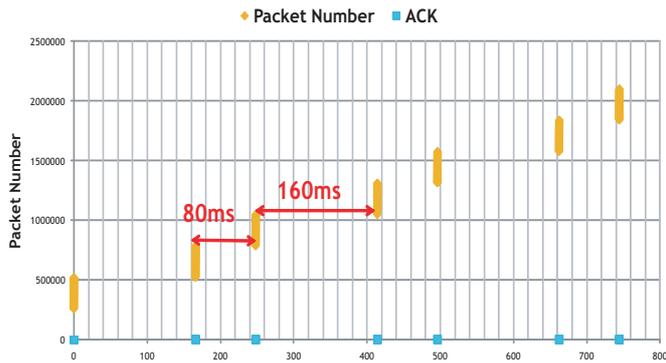


図 14 RTT80ms におけるパケット解析 (VPN 無)

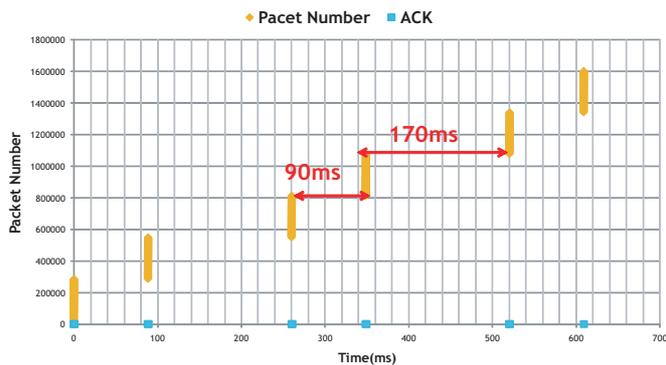


図 15 RTT80ms におけるパケット解析 (VPN 有)

### 8.1 TCP 層における解析

TCP ACK の受信をきっかけとしてパケット送信の再開が行われていることからトランスポート層のどこかにパケット送信断続の原因があるということになる。性能劣化の原因が TCP 層のどこかであると想定した場合、その要因は広告ウィンドウ、輻輳ウィンドウ、ソケットバッファの 3 点が考えられる。本研究では広告ウィンドウを初めから通信の妨げにならない程度の大きさとして、4MB を設定し測定、解析を行った。

輻輳ウィンドウ、ソケットバッファについては以下に示す、カーネルモニタというオリジナルツールを使用し解析を行った。

### 8.2 カーネルモニタ

本実験では TCP カーネルの振舞をモニタするため、Initiator 側の TCP ソースコードの中にモニタ関数を挿入しカーネルを再コンパイルした。これによりユーザ空間からは見ることが出来ないカーネル内の情報を可視化することが可能になる。カーネルモニタによって、カーネル内部にログのトレース、タイムスタンプ、CWND、ソケットバッファキュー長といった TCP パラメータ情報を取得することができる。

### 8.3 既存研究における TCP 層最適化

既存研究より、iSCSI 通信中とソケット通信中とでは、明らかにソケットバッファの振舞が異なることが確認されている [6]。両者のカーネル中の処理がどのように行われているのかを調べたところ、通過する場所は異なるものの、どちらも timeout が生じる。ソケット通信においては、図 17 における 1415 行目の schedule() を通過し、iSCSI 通信においては 1439 行目の

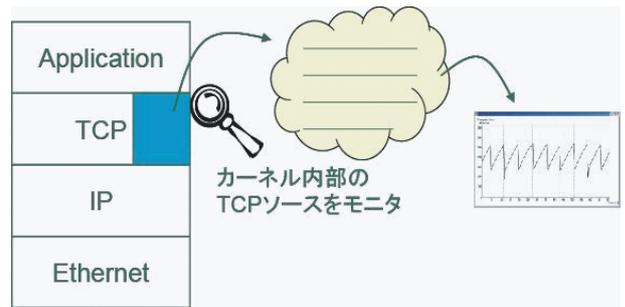


図 16 カーネルモニタ

schedule() を通過する際にタイムアウト待ちに入り RTT 後の ACK で再開されており、iSCSI 通信においては不必要なタイムアウト待ちが頻繁に生じていることが分かっている。

両者ともにタイムアウト待ちに入っているのにも関わらず、ソケット通信においては性能は維持されたままで、iSCSI 通信においては性能低下を引き起こしている。その原因としては、ソケット通信の場合はソケットバッファのキューに十分なデータが補遺されているからパケット送信は途切れないということが考えられる。

```

1400 fastcall signed long schedule_timeout(signed long timeout)
1401 {
1402     struct timer_list timer;
1403     unsigned long expire;
1404
1405     switch (timeout)
1406     {
1407     case MAX_SCHEDULE_TIMEOUT:
1408         schedule(O);
1409         goto out;
1410     default:
1411         if (timeout < 0)
1412             ;
1413         printk(KERN_ERR "schedule_timeout: wrong timeout
1414             " value %lx from %fn", timeout,
1415             _builtin_return_address(0));
1416         current->state = TASK_RUNNING;
1417         goto out;
1418     }
1419     expire = timeout + jiffies;
1420     setup_timer(&timer, process_timeout_unsigned, current);
1421     mod_timer(&timer, expire);
1422     del_timer_sync(&timer);
1423     out:
1424     ;

```

図 17 kernel ソースコード

### 8.4 ソケットバッファ最適化

ソケットバッファ最適化をしたところ、iSCSI ソフトウェアターゲットシステム構成 Type1 と iSCSI ハードウェアターゲットシステム構成においては既存研究 [6] のような劇的な性能向上は見られなかった。そこで以下に示すカーネルモニタというオリジナルツールを用い、輻輳ウィンドウの解析を行った。

### 8.5 輻輳ウィンドウ解析

#### 8.5.1 iSCSI ソフトウェアターゲットシステム構成 Type1 における解析

iSCSI ソフトウェアターゲットシステム構成 Type1 においてカーネルモニタを用い、輻輳ウィンドウをモニタした結果のグラフを図 18 に示す。片道遅延時間 32ms の iSCSI アクセスにおいて、輻輳ウィンドウは急激に伸びるものの最大値 290 で頭打ちとなった。輻輳ウィンドウが 290、送信可能データが Ethernet のフレームサイズが 1.5KB の場合、約  $1.5 \times 290 = 435$ (KB) であるのに対し、ACK が返るまでに送信すべきデータ量は通信速度が 300Mbps、RTT=64ms の場合、

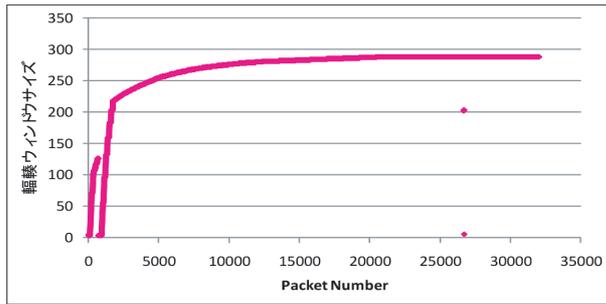


図 18 iSCSI ソフトウェアターゲットシステム構成 Type1 における解析

$300/8 \times 10^3 \times 64/10^3 = 2400(\text{KB})$  である。

このように、iSCSI ソフトウェアターゲットシステム構成 Type1 では高遅延環境において送信すべきデータ量より輻輳ウィンドウが小さいということが確認された。

### 8.5.2 iSCSI ハードウェアターゲットシステム構成における解析

iSCSI ハードウェアターゲットシステム構成においてカーネルモニタを用いて輻輳ウィンドウをモニタした結果のグラフを図 19 に示す。片道遅延時間 32ms の iSCSI アクセスにおいて、輻輳ウィンドウは急激に伸びるものの最大値 251 で頭打ちとなった。輻輳ウィンドウが 251、送信可能データが Ethernet のフレームサイズが 1.5KB の場合、約  $1.5 \times 251 = 377(\text{KB})$  であるのに対し、ACK が返るまでに送信すべきデータ量は通信速度が 300Mbps、RTT=64ms の場合、 $300/8 \times 10^3 \times 64/10^3 = 2400(\text{KB})$  である。このように、iSCSI ハードウェアターゲットシステム構成においても高遅延環境において送信すべきデータ量より輻輳ウィンドウが小さいということが確認された。この輻輳ウィンドウ切れがこの実験におけるスループット低下の原因であるということが考えられる。

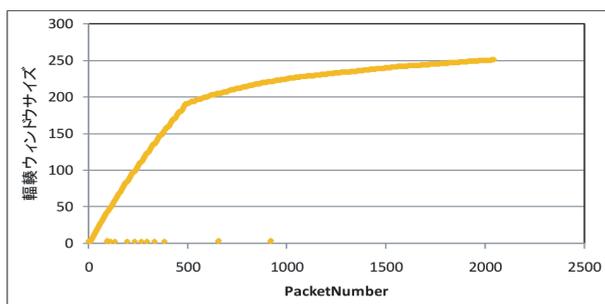


図 19 iSCSI ハードウェアターゲットシステム構成における解析

## 9. iSCSI ソフトウェアターゲット構成 Type2 での TCP 層最適化

iSCSI ソフトウェアターゲット構成 Type2 において iSCSI 層最適化をした後、TCP 層のソケットバッファ最適化を行った。各層最適化後のスループットの比較を、VPN 遠隔接続環境における結果は図 20 に、VPN を使用しない遠隔接続環境における結果は図 21 に示す。VPN 遠隔接続環境において、RTT=32ms

では約 2 倍の性能向上となった。VPN を使用しない遠隔接続環境では RTT=32ms においてデフォルト時と比較して約 12 倍の性能向上が達成されたことから、iSCSI 遠隔ストレージアクセスの高遅延環境における性能劣化問題は解決されることが出来る。このことからこの実験環境では、VPN 技術の性能向上が期待される結果となった。

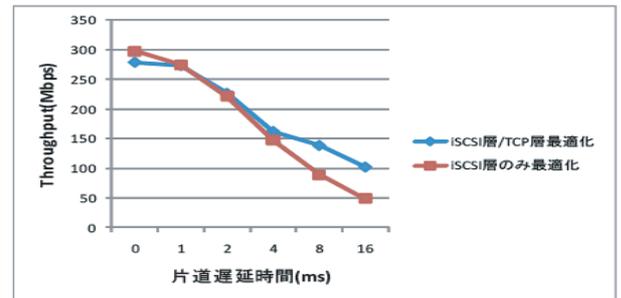


図 20 VPN 遠隔接続環境における性能向上

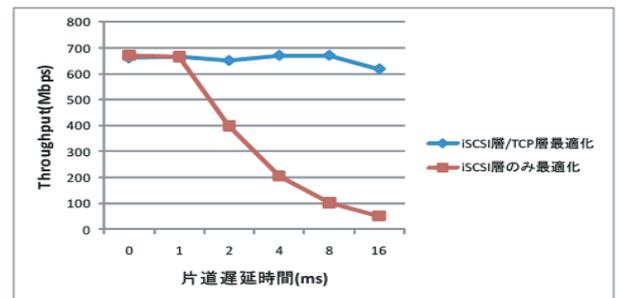


図 21 VPN を使用しない遠隔接続環境における性能向上

## 10. まとめと今後の課題

遠隔 iSCSI アクセスは需要の高い技術であるが、iSCSI の階層構成と遠隔アクセスによる遅延が合わさり振舞が複雑で、性能向上のためには各層を丁寧に最適化する必要がある。実環境になると多種多様な構成が加わりさらに複雑になるため、システム内部の各層で詳細な振舞を調べる必要があり、本研究ではこれらに様々なタイプの最適化を適用した。

ソフトウェアターゲット構成 Type1、ハードウェアターゲット構成の性能低下原因は輻輳ウィンドウ切れであることが分かった。一方、ソフトウェアターゲット構成 Type2 においては VPN 遠隔接続環境ではソケットバッファ最適化により全体的に性能が向上したものの、高遅延環境での性能劣化問題はまだ残っており、VPN 技術特有の問題解決が求められる結果となった。VPN を使用しない遠隔接続環境における結果はソケットバッファ最適化により片道遅延時間 16ms において約 12 倍の性能向上となった。

これらの実験結果より、環境が変わると、どの最適化が効果があるかが変わる事がわかり、最適化が効く程度も大きく変化する。実環境において遠隔 iSCSI アクセスを実行したい場合には、環境ごとに適用する最適化手法を変える必要があり、その効果も環境次第で程度が異なる事がわかった。

今後はハードウェアターゲット構成でも高遅延環境における性能向上を実現させるため、組み合わせの検証を行っていき  
たい。

#### 文 献

- [1] iSCSI Specification ,  
<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: iSCSI 解析システムの構築  
と高遅延環境におけるシーケンシャルライトアクセスの性能向  
上に関する考察, 電子情報通信学会論文誌 Vol.J87-D-I, No.2,  
pp.216-231, 2004 年 2 月
- [4] 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性能評価, 情報処理学会論文誌,  
Vol.49, No.10, pp.3645-3656, 2008 年 10 月
- [5] 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>
- [6] 比嘉 玲華, 松原 幸助, 岡廻 隆生, 山口 実靖, 小口 正人: iSCSI  
遠隔ストレージアクセス性能向上を実現する最適化手法の異な  
る iSCSI 実装による評価, DEIM2010, E6-5, 2010 年 3 月
- [7] 浅田菜那, 小口正人: iSCSI ストレージの各層最適化時の VPN  
遠隔接続環境における性能向上の検討, SWoPP2010, 2010 年 8  
月
- [8] Open-iSCSI, <http://www.open-iscsi.org/>
- [9] iSCSI Enterprise Target, <http://iscsitarget.sourceforge.net/>
- [10] The Linux sg\_dd utility, [http://sg.danny.cz/sg/sg\\_dd.html](http://sg.danny.cz/sg/sg_dd.html)