

# iSCSI 遠隔ストレージアクセスの 実環境向けシステム構成における性能評価

浅田 菜那 (指導教員: 小口 正人)

## 1 はじめに

近年, ストレージ管理コスト低減などの目的で SAN(Storage Area Network) の導入が進んでおり, その中でも, より経済的でコストパフォーマンスの高い IP-SAN のプロトコルである iSCSI は大いに期待されている. 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が広く使われることが期待され, 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. また iSCSI ターゲットには, オープンソースソフトウェアである iSCSI Enterprise Target の他, 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とする, 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 実験システム

本研究において, Initiator と Target 間は GigabitEthernet で接続し, TCP/IP コネクションを確立した. iSCSI Target としては, iSCSI Software Target 構成と, FC ストレージと IP ブリッジを用いた iSCSI Hardware Target 構成の 2 種類の構成を構築した. またそれぞれの環境において VPN 遠隔接続環境 (図 1, 図 2) と VPN を使用しない遠隔接続環境 (図 1, 図 2 から VPN ルータを除いたもの) 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を測定した.

## 3 iSCSI 層における最適化

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するため表 1 のように調整した.

iSCSI パラメータ最適化を行った iSCSI Software Target 構成における実験結果を図 3 に, iSCSI Hard-

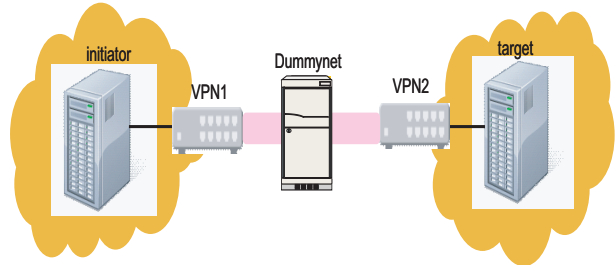


図 1: iSCSI Software Target 構成における VPN 遠隔接続環境

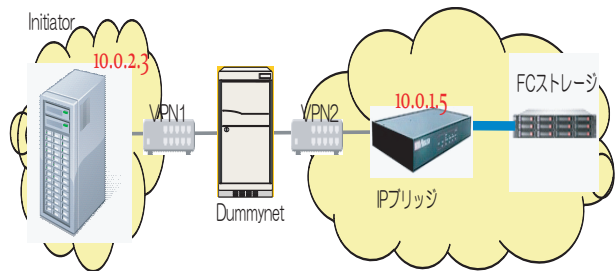


図 2: iSCSI Hardware Target 構成における VPN 遠隔接続環境

表 1: iSCSI パラメータ設定

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

ware Target 構成における実験結果を図 4 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わずかではあるが全体を通して性能が向上する結果となった. しかし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.

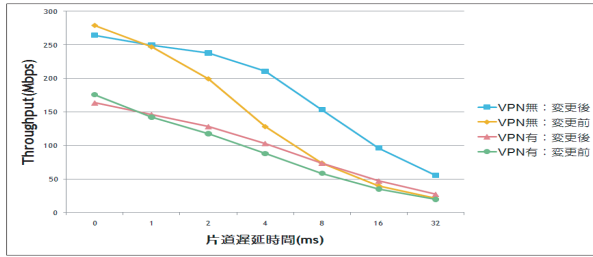


図 3: iSCSI Software Target 構成での iSCSI 層最適化

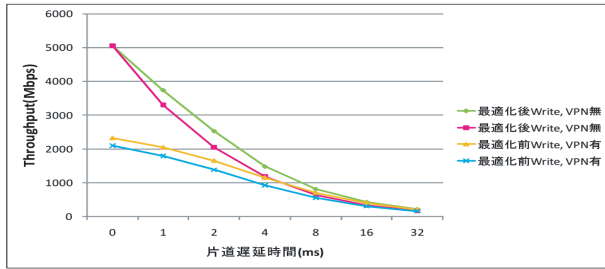


図 4: iSCSI Hardware Target 構成での iSCSI 層最適化

## 4 TCP 層における解析と最適化

iSCSI Software Target 構成において, RTT=80ms における iSCSI アクセスを tcpdump コマンドを使用し解析した。その結果, VPN 遠隔接続環境においても VPN を使用しない遠隔接続環境においても, 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ことがわかった。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できる。

また 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, 性能劣化の原因が TCP 層のどこかにあると言える。そのように想定した場合, 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。

### 4.1 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, 全ての環境において既存研究 [2] で示された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。

### 4.2 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 パラメータ情報を記録し, これを取得して解析することができる。

### 4.3 輻輳ウィンドウサイズ解析

カーネルにモニタ関数を挿入し, 輻輳ウィンドウの振舞をモニタしたところ, SoftwareTarget の場合は最大値 290, HardwareTarget の場合は最大値 251 であった。これは送信すべきデータ量より輻輳ウィンドウが小さい値となっている。このような輻輳ウィンドウ切れがスループット低下の原因であることが分かった。

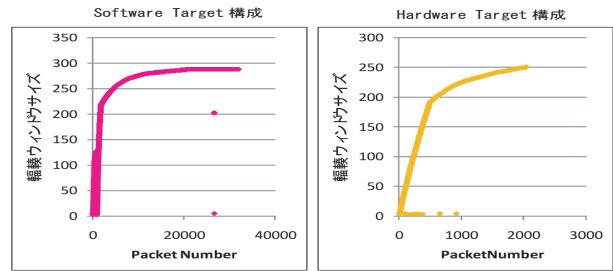


図 5: 輻輳ウィンドウサイズ解析

## 5 まとめと今後の課題

本研究では, 様々な環境における, 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。iSCSI 層における最適化については, パラメータを最適化し, 高遅延環境における性能向上を確認することが出来た。スループットは片道遅延時間に反比例するが, 遅延時間が長くなると VPN 遠隔接続環境と VPN を使用しない遠隔接続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。

TCP 層における最適化については, 既存研究 [2] では性能向上の鍵となったソケットバッファに関するカーネルコードの書き換えを行ったが, 本研究の実験環境では性能向上とはならなかった。そこで輻輳ウィンドウを解析したところ, 性能低下原因は輻輳ウィンドウ切れであることがわかった。

今後は VPN 有無, Target, Initiator, RAID 構成などの組み合わせを変え, より多くの実験環境における測定と解析を行いたい。例としては, 既存研究で TCP 層最適化が成功した Initiator とハードウェアターゲットとを組み合わせ構成し, 各層最適化を行い, 結果を検証していく。

## 参考文献

- [1] 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
- [2] 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