

iSCSI 遠隔ストレージアクセスにおける パケット送信の断続性の解析

B-6 Analysis of intermittent packets on iSCSI Remote Storage Access

比嘉 玲華¹
Reika Higa

松原 幸助²
Kosuke Matsubara

岡廻 隆生²
Takao Okamawari

山口 実靖³
Saneyasu Yamaguchi

小口 正人¹
Masato Oguchi

お茶の水女子大学¹
Ochanomizu University

ソフトバンクテレコム株式会社²
SOFTBANK TELECOM Corp

工学院大学³
Kogakuin University

1 はじめに

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期待されている。しかし現状において iSCSI は、高遅延環境においては性能の劣化が著しく、iSCSI 層だけではなく複数の層にまたがる制御を施すことによる性能向上が期待される。既存研究において、複数の層にまたがる最適化を行ったが、高遅延環境においては、なお性能低下が著しい [1]。iSCSI ブロックアクセスのモデル化を行い、アナライザを用いて解析を行った結果、高遅延環境における iSCSI アクセスが低下する理由は、RTT によらずに一定であるはずのデータ転送時間が RTT に比例する値になっていることが原因であることがわかった。

そこで本稿では、パケット解析、輻輳ウィンドウ解析を行うことで、より詳しく検討していく。

2 実験システム

本研究において、Initiator と Target 間は Gigabit Ethernet で接続し、TCP/IP コネクションを確立した。Target のストレージには SAS ディスクを用い RAID コントローラによる RAID0 構成で接続した。本実験において使用した OS は、Red Hat Enterprise Linux 2.6.18-8.e.15、CPU は Intel Quad Core Xeon 1.6GHz である。iSCSI は Initiator に open-iscsi-2.0-865、Target に iSCSI Enterprise Target(IET)-0.4.15 を使用した。

3 輻輳ウィンドウとパケット解析

3.1 パケット送出と CWND 解析

スループットと輻輳ウィンドウには密接な関係があることが知られている。そこで、Initiator 側でカーネルモニタと tcpdump を使って輻輳ウィンドウの値とパケット送出量の関係を調べた。RTT50ms、ブロックサイズ 4MB の iSCSI アクセスを実行した。その結果、4096 MB を非同期に送信するには、輻輳ウィンドウは約 3000 が必要であるが、輻輳ウィンドウは約 1200 であり十分な大きさではないことがわかった。詳細なパケット解析を行うべくアナライザを用いて測定、解析を行った。

3.2 データ送出のより詳細な解析

アナライザによる解析結果を図 1 に示す。この図から以下のことがわかる。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いる。一定時間の後に再びパケットの連続送信が行われており、それらの間隔は RTT に等しい約 50ms である。パケットの送出量は約 600 であり、輻輳ウィンドウの 1200 を使い果たす値ではない。またアナライザによりパケットを詳細に調べた結果、送信再開の前後には TCP ACK のみが受信されていることもわかった。このことから、パケットの送出制御が TCP レベルで行なわれていると考えられる。

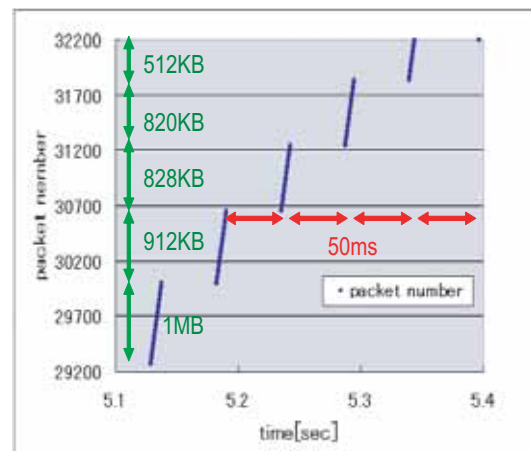


図 1 RTT50ms におけるパケット解析

4 まとめと今後の課題

パケット解析、輻輳ウィンドウ解析を行い更なる検討を行なった結果、iSCSI の送信において停止・開始の制御は TCP レベルで行われているが、しかし、その制御は輻輳ウィンドウの値だけによってはいけないということが考えられる。今後の課題としては、カーネル解析等の更なる解析を行っていく。

参考文献

- [1] 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の複数レイヤにまたがる最適化”, インターネットコンファレンス 2008, pp.120, 沖縄, 2008 年 10 月