小口研究室　研究紹介　(2008年度)
(お茶の水女子大学理学部情報科学科)

iSCSI遠隔ストレージアクセスにおける複数の層にまたがる最適化と性能劣化要因の解析(研究担当：比嘉　玲華)

研究背景・目的

☆iSCSI
- IP-SANのプロトコル
- SCSI over TCP/IP over Erthnet⇒複雑な階層構造

☆iSCSIの利点
- 相互接続性が高い
- 距離が延ばせる
- 管理が容易
- 安価にSANを構築可能

iSCSIを使って遠隔地へのデータのバックアップを従来の専用線よりも**より簡単に！より安価に！**

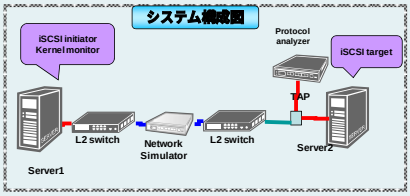
⇒ **広域ネットワークへの利用の期待が高まる！**

研究課題・アプローチ

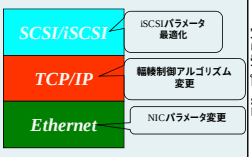
☆iSCSI問題点
- 高遅延環境下において性能が劇的に低下

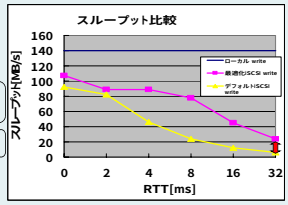
☆研究のアプローチ
- iSCSI最適化
  - 複雑な階層構造のため、単独ではなく複数の層にまたがる最適化が必要
  - 性能劣化要因となるボトルネックを特定

システム構成図



複数の層にまたがる最適化





☆RTT32msにおいて約4倍の性能向上が確認

☆しかし、高遅延環境における性能劣化は解消されず

☆送信できるデータ量が小さい

⇒ データ送信量を大きくし、その上で解析をしていく必要あり！

性能劣化要因となるボトルネックの特定

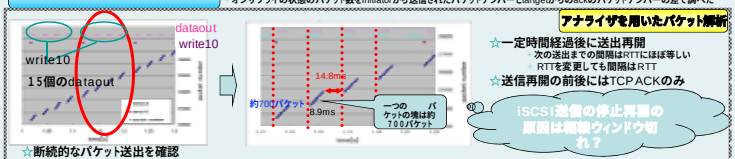
☆4MBの大きさのブロックサイズでアクセス　RTT 20ms　を設定
- SCSIレベルでの大きなブロックサイズのアクセスが可能になるようにカーネルを再構築

☆パケットの送出時間を測定し、iSCSIストレージアクセスの各処理を測定

高遅延環境におけるiSCSIアクセスの性能低下原因はRTTによらず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

データ転送部分を詳細に解析

ネットワーク上を飛び交うパケットをキャプチャし、アナライザをもちいて解析
自作のカーネルモニタを用いて輻輳ウィンドウの値を解析
オンザフライの状態のパケット数をinitiatorから送信されたパケットナンバーとtargetからのackのパケットナンバーの差で調べた



☆断続的なパケット送出を確認

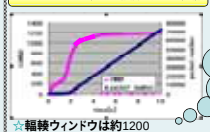
アナライザを用いたパケット解析

☆一定時間経過後に送出再開
- 次の送出までの間隔はRTTにほぼ等しい
- RTTを変更しても間隔はRTT

☆送信再開の前後にはTCP ACKのみ

iSCSI送信の停止再開の原因は輻輳ウィンドウ切れ？

カーネルモニタを用いた輻輳ウィンドウ解析

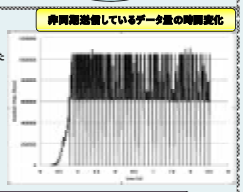


☆輻輳ウィンドウは約1200
- 4MBを送出するのに十分な輻輳ウィンドウは4096KB/1.448KB＝2828　4MBには足りないが、1200＜700

☆輻輳ウィンドウを使い果たしているわけではない

パケット送出停止は輻輳ウィンドウが原因ではなさそう

非同期送信しているデータ量の時間変化



☆本当に輻輳ウィンドウが余っているのかを調べた

☆オンザフライの状態のパケット数を調べる

☆ネットワーク上を飛び交うパケット数は、最大でも約1.1MB

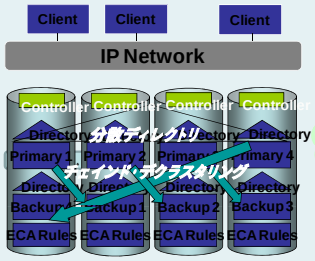
輻輳ウィンドウは使い果たされていない状態である！

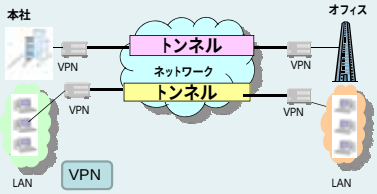
まとめと今後の課題

☆iSCSI送信において停止・再開の制御はTCPレベルで行なわれているがその制御は輻輳ウィンドウの値によるものではない

☆今後はカーネルの内部をより詳細に解析していくことで、なぜ送信が止まってしまうのかを特定する

自律ディスクを用いたVPN接続広域分散ネットワークストレージの評価（研究担当：浅田菜那）





VPN
- 公衆網を使い拠点間を仮想的に閉じたワークで接続
- 安価に実質的な専用網を実現

自律ディスク

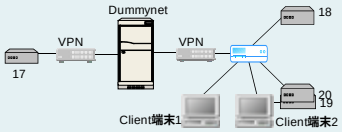
- ストレージをストレージ上で自律的に行う
- データの冗長配置による対策や効率のよい障害回復、故障対策、負荷分散を行う
- 一台を失っても残りのノードが自動的にチェインド・デクラスタリングを組み直す

⇒ **自律ディスクを広域分散ネットワーク上で安定性・信頼性のあるネットワークストレージの構築**

実験環境

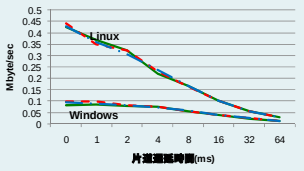
	端末OS	端末台数
■fs17-20がそろった構成	■Windows	■1台
■fs18を除いた構成	■Linux	■2台
■fs18,19を除いた構成		

すべての組み合わせにおいて遅延装置により0-64msの遅延環境を作り、測定



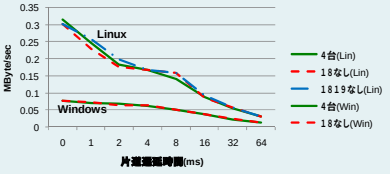
実験結果

クライアント端末1台



- 自律ディスクは広域分散ネットワーク環境でも正常に動作する
- 1つのノードを失っても残りのノードがチェーンを組み直し、正常に機能している

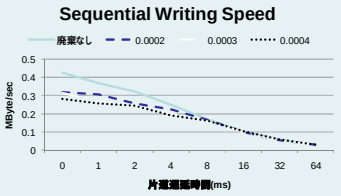
クライアント端末2台



- 自律ディスクは広域分散ネットワーク環境でも正常に動作する
- 1つのノードを失っても残りのノードがチェーンを組み直し、正常に機能している

現実的なネットワーク上での評価

- 一定の率でパケット廃棄を起こす
- 片道遅延時間が短い時にはパケット廃棄率の違いがスループットにある程度大きな影響を及ぼす
- パケット廃棄によるオーバーヘッドの差異が、遅延時間の増大による性能低下により隠蔽される



センサネットワークを用いたストリームデータ処理実験環境の構築(研究担当：川口菜々)

研究目的

・近年の無線通信技術の発展に伴い、無線通信端末で構築される無線センサネットワークに注目が集まっている

⇒ 離れた場を確認したり、過去の時系列をデータベースから参照することでより便利なアプリケーションの実現が期待されている

実験住宅『OCHA HOUSE』内においてネットワーク技術を駆使した生活に便利なアプリケーションの作成に向けた実験環境の基盤を構築する

•実験住宅内で無線センサネットワークを構築し、取得されたセンサデータをリモートの情報解析端末から観測するといった情報解析のプラットホームを構築

実験環境の構築

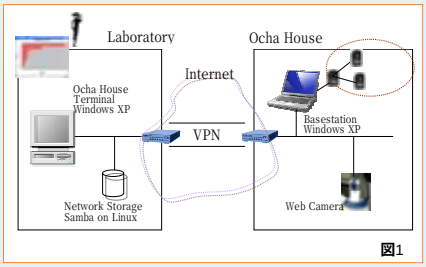


図1

•実験住宅と解析拠点を結ぶインターネット上にVPN環境を構築

•実験住宅内にはWebカメラも導入し、実験住宅内で観測されたセンサデータと画像データの双方をネットワークストレージに記録(図1)

•VPN環境において、VPNを挟んだ反対側のLAN空間へもパケットが届くように、各ノードに経路制御のエントリを書き加える(図2)

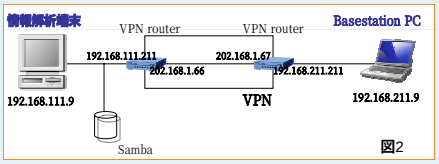


図2

•センサデータと画像データの双方を取得して比較することでデータの相関を調べるといった解析処理が可能

•また、VPN環境を構築することで、記録したデータを遠隔ネットワーク越しに、リモートの端末で観測することが可能

http://ogl.is.ocha.ac.jp/

oguchi@computer.org