

自律ディスクを用いたVPN広域分散接続ネットワークストレージの評価

浅田 菜那 (指導教員：小口 正人)

1 はじめに

近年、RAID は高性能ストレージのひとつとしてよく知られている。これは、複数台のハードディスクを組み合わせることで冗長性の高い仮想的な一台のハードディスクとして運用する技術である。これに対し、分散ディレクトリやチェインド・デクラスタリングなどを用い、負荷分散、故障対策、障害回復などにおいてRAID より更に高機能化された自律ディスクというシステムが研究されている。本研究においては、VPN を利用することにより自律ディスクを広域分散ネットワークに適用することを検討し、その状況における自律ディスクの振舞を解析することによって、さらに性能を高めるための手法を検討する。

2 実験環境

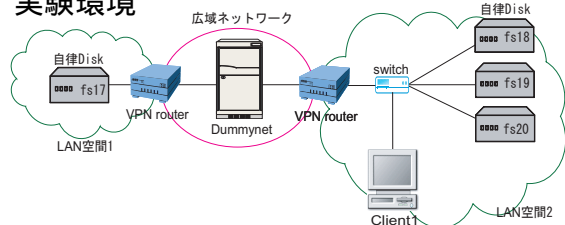


図 1: 実験システムの概要

本研究では、はじめに図 1 に示す実験環境を構築した。VPN ルーターを 2 台用いて一方の LAN 空間には 1 台の自律ディスク fs17 を、もう一方の LAN 空間には 3 台の自律ディスク fs18-20 を接続させた。2 台の VPN ルーターの間には、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。また fs18-20 をつなぐスイッチングハブに性能評価を行うクライアント端末 1 を接続した。自律ディスクとクライアント間は SMB (Server Message Block) プロトコルで接続されており、fs20 に自律ディスク向けに改変した Samba サーバを立ち上げてクライアントからのアクセスを受け付ける。

クライアント端末上では、Windows からの評価として Iometer を、Linux からの性能評価として自作ツールの HDDbench を使用した。Iometer においては Sequential Read (連続読み込み) と Sequential Write (連続書き込み) のそれぞれについてスループット、レスポンスタイム、CPU 使用率を、HDDbench においては Sequential Write のスループットとレスポンスタイムを測定した。

実験の手法は、まず遅延装置を使用し 0-64ms の遅延環境を作る。次に下記のように実験環境を 3 パターン構築し、それぞれにおいて測定した。

- (1) fs17-20 が 4 台そろった構成
- (2) fs18 を除いた構成
- (3) fs18, fs19 を除いた構成

はじめに 4 台を立ち上げた後、(2) の構成では fs18 を、(3) の構成では fs18 と fs19 を物理的に切り離す。このようにすることで、1 つまたは 2 つのノードが突然失われた場合でも残りのノードが速やかにチェイン

ド・デクラスタリングを構成しなおし正常に作動するかを調べることができる。

以上の実験を、クライアント端末を 1 台に設定した構成と 2 台に設定した構成それぞれについて測定する。

3 性能測定結果

3.1 クライアント端末 1 台の構成

3.1.1 Windows からの評価

上記実験環境において Windows から測定し比較した連続読み込みにおけるスループットのグラフを図 2、レスポンスタイムのグラフを図 3、連続書き込みにおけるスループットのグラフを図 4、レスポンスタイムのグラフを図 5 に示す。

Windows をクライアント端末に設定した状態での、連続読み込みと連続書き込みのスループットでは、3 パターンの構成において自律ディスクがほぼ同じような振舞をすることがわかった。またそれぞれのグラフは、遅延時間が大きくなるにつれて減少していく。連続書き込みのレスポンスタイムについてもスループット同様、3 台の振舞は似ているため 3 本のグラフは重なり合っている。一方、連続読み込みのレスポンスタイムでは、書き込みとは異なる結果が得られた。どのグラフも重なり合うことはなく、特に 4 台がそろった構成でのグラフは他の二つのグラフと比較したところ最も異なる動きをしている。CPU 使用率はどのグラフも片道遅延時間が増すにつれて減少した。これは実行中の自律ディスクが CPU を占有している時間の割合が少なくなるためだと考えられる。

Windows における、連続読み込みのスループット (図 2) と連続書き込みのスループット (図 4) の比較、また連続読み込みのレスポンスタイム (図 3) と連続書き込みのスループット (図 5) の比較より、連続読み込みのグラフは連続書き込みのグラフよりもばらつきがあることがわかる。よって自律ディスクは連続書き込みよりも連続読み込みにおいて多少性能が不安定となることがわかった。

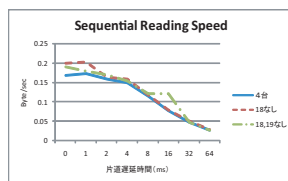


図 2: スループット

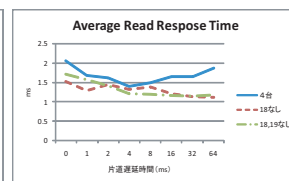


図 3: レスポンスタイム

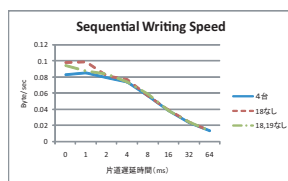


図 4: スループット

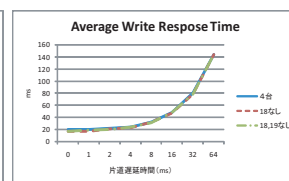


図 5: レスポンスタイム

3.1.2 Linux からの評価

スループットのグラフを図 6、レスポンスタイムのグラフを図 7 に示す。スループット、レスポンスタイムともにそれぞれの 3 つのグラフはほぼ重なり合い、どの遅延時間においてもあまり変化は見られなかった。スループットのグラフは片道遅延時間が大きくなるにつれて一定の割合で減少する。レスポンスタイムのグラフは図 5 の Windows からの評価におけるものと比較すると、形状が極めて似ていることが見て取れ、自律ディスクはどちらの OS からでも SMB プロトコルを介して同程度のアクセス性能を達成できることがわかった。

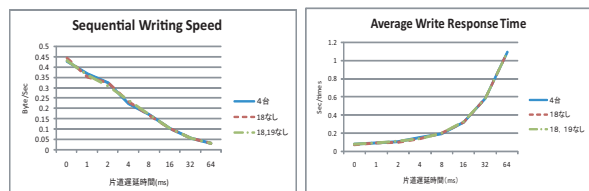


図 6: スループット 図 7: レスポンスタイム

3.2 クライアント端末 2 台の構成

次にクライアント端末を LAN 空間 2 に 2 台設置し、同時に測定した。

3.2.1 Windows からの評価

Windows をクライアント端末に設定した構成では、4 台揃った場合と fs18 を除いた場合の 2 パターンで実験を行った。2 台の端末からはほぼ同じような結果が得られた為、グラフは酷似したものになった。例として端末 1 における結果を図 8 に示す。

連続読み込みでは端末 1 台の実験と同様に連続書き込みのグラフと比較した際、特にレスポンスタイムにおいて不安定な結果になった。

連続書き込みの 2 本のグラフはほぼ重なり合い、クライアント端末が 1 台の場合 (図 4, 図 5) と比較したところ、ほぼ同じような数値を示していた。

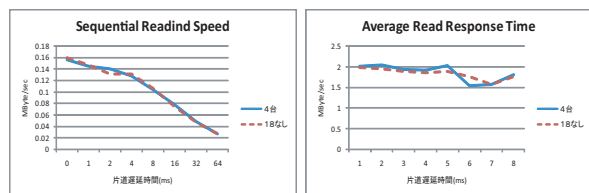


図 8: スループット 図 9: レスポンスタイム

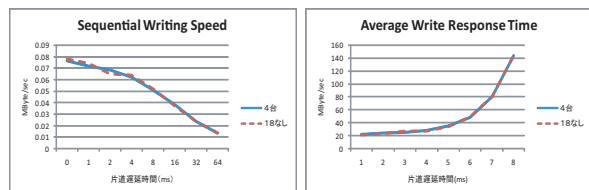


図 10: スループット 図 11: レスポンスタイム

3.2.2 Linux からの評価

次に、端末 2 台を Linux に設定し、4 台が揃っている構成、fs18 を外した構成、fs18, fs19 を外した構成において測定した。Windows における測定の結果と同様に、2 台の端末からは同じような結果が得られた為、例として端末 1 における結果のみ図 12, 13 に示す。

グラフの形状はクライアント端末が 1 台の場合と似ているが、2 台に設定した場合はスループット、レスポンスタイム共に数値に変化が見られた。

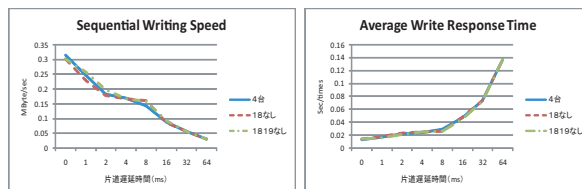


図 12: スループット 図 13: レスポンスタイム

4 考察

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ことがわかった。4 台のうち 1 台または 2 台がダウンしても残りのノードでチェインド・デクラスタリングを構成しなおし、機能することを確かめることが出来た。また、クライアント端末が 1 台の場合でも 2 台の場合でも、3 つの異なる構成におけるスループットやレスポンスタイムの結果が大きく変わるということにはなかった。このことより、自律ディスクのチェインが直接つながっている fs17, 20 がそろっている場合において fs18, 19 はそれほど大きく関わらないのではないかと考えることが出来る。理由としてはデフォルトでは fs20 宛に SMB でアクセスし、それが自律ディスクの仕組みによって fs17 にバックアップとして書きこまれているためであろう。また片道遅延時間が大きくなるにつれスループットが低下し、レスポンスタイムが長くなっているため、性能は劣化していることがわかった。

5 まとめと今後の課題

本研究では VPN 広域分散ネットワーク環境で自律ディスクの振舞を解析、評価した。結果、このような環境でも動作しノードが欠けてしまった場合でも残りのノードで構成を組みなおし正常に機能することが確認できた。今後はより詳細な解析、評価を行っていくと共に、自律ディスク 4 台を同じ LAN 空間に置いたり、クライアント端末を LAN 空間 1 と LAN 空間 2 に分けて置く等今までと異なる構成において測定する。

参考文献

- [1] 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
- [2] 横田 治夫: 機能追加によるストレージの高付加価値化のアプローチ, MR2006-79, 2007 年 3 月
- [3] 浅田 菜那, 小口 正人: "自律ディスクを用いた VPN 広域分散接続ネットワークストレージの評価", DEIM2009, 2009 年 3 月発表予定