

6ZB-04

Linux クラスタシステムへの SAN 及び 共有ファイルシステム適用有効性の検証

佐藤 清司 佐藤 政一 野澤 雅之 監物 文乃

エヌ・ティ・ティ・コムウェア株式会社

表 2.1 SAN の特徴

	特徴
メリット	複数ホストからの Storage の共有が可能 システムの大規模化に対応可能 ストレージの統合、LAN の負荷軽減、大容量（柔軟な拡張性） 高性能（高速データ転送、効率的なバックアップ）、高信頼性（2 重化） TCO 削減（一元管理によるコストダウン）
問題点	ファイル（ディスク）の共有ができない

表 2.2 共有ファイルシステムの特徴と製品

特徴	SAN 環境での複数ホストからのファイル（ディスク）の共有が可能。 NFS のような Network のボトルネックは解消される。 ファイルが 1 つでいい。（総ディスク容量の削減、ファイルをサーバごとで持つ必要がない。） 1 つのファイルを複数サーバから読み書き可能
製品	GFS (Global File System) (Sistina Software 社) SANergy (日本デボリシステムズ社) Centraviton (ADIC 社) 等

1. はじめに

近年、Web やメールのサーバに Linux クラスタが数多く用いられている。それらのシステムでは一般的に NFS によるファイル共有が行われているが、システムを大規模化した場合ネットワークトラフィックの増加・NFS サーバの負荷の増加が発生し、Storage へのファイルのアクセス性能の低下につながり、システムを構築する上で大きな問題となる。

そこで、今回我々はこの問題を解決する手法の一つとして SAN (Storage Area Network) と共有ファイルシステムを用いてシステムを構築する手法を試みた。

本稿では、この検証結果を元に Linux クラスタ環境において SAN を用いた共有ファイルシステムの有効性について報告する。

2. 現状の問題点とその解決方法

我々が用いた解決方法を以下に説明する。

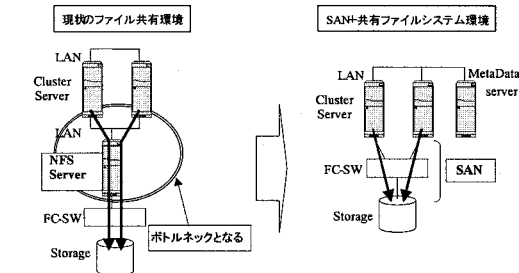


図 2.1 今回適用したモデル

2.1 問題点

現状の NFS によるファイル共有環境では、システムが大規模化した場合 NFS のための LAN のネットワークトラフィックと NFS サーバの負荷が増加するため、これらがボトルネックとなり Storage へのファイルアクセス性能の低下の原因となる。そのため大規模システムにおけるファイル共有には適していない。

2.2 解決方法

SAN と共有ファイルシステムを用いてファイル共有環境を構築する事により NFS のための LAN と NFS サーバのボトルネックを解消し大規模なファイル共有を実現する。

表 2.1, 2.2 に SAN と共有ファイルシステムの特徴を示す。

SAN と共有ファイルシステムを両方利用することにより、SAN 環境でのファイル共有が可能となり、Storage へのファイルアクセス性能が改善される。また NFS サーバ等も不要となり TCO の削減にもつながる。

3. 検証環境

本検証では NFS を用いたファイル共有環境と SAN+共有ファイルシステムの環境を構築し、Storage へのファイルアクセス性能を比較する。図 3.1 に NFS 環境、図 3.2 に SAN+共有ファイルシステム環境を示す。

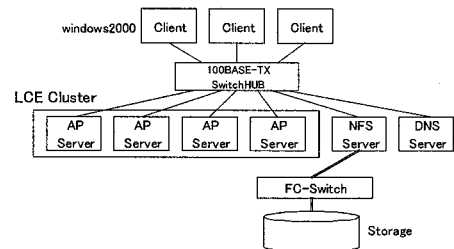


図 3.1 NFS 環境

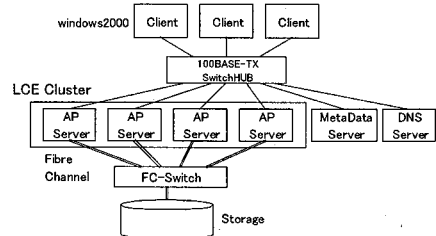


図 3.2 SAN+共有ファイルシステム環境

表 3.1 システム構成

Server	VA LINUX VA2240 : CPU : PentiumⅢ 850MHzX2 メモリ : 512MB OS : RedHat6.2		
Client	ThinkPadX21 : PentiumⅢ 700MHz OS : Windows2000		
LAN	100Base-T		
HBA	QLogic2100		
FC-Switch	Brocade SW2800		
Storage	e-Storage DF25030 キャッシュ : 512MB RAID : 5 (100GB)		
ClusterSoft	NTTコムウェア LCE (Local Cluster Enterprise)		
Kernel構成	NFS環境	2.2.19SMP	
	SAN+共有ファイルシステム	製品A	2.2.19SMP+GFS (V4.1.1)
		製品B	2.2.14

4. 検証方法

4.1 Iozone による性能測定

Storage へのファイルアクセスの基本の性能を測定するために NFS 環境及び SAN+共有ファイルシステム環境において性能測定ツール (Iozone:Version3.53) を使用し Sequential-Read/Write, Random-Read/Write 性能を測定した。図 3.1,3.2 の AP サーバ上で Iozone を起動し、サーバの台数を 1~4 台からの同時アクセス性能を測定した。なお、試験ファイルサイズは 1 GB を使用した。

4.2 応用 AP を利用した性能検証

Client-Server のシステム構成にした場合のシステム全体を通した場合の性能を NFS 環境と SAN+共有ファイルシステム環境について比較検証した。図 4.1 に検証 AP の動作イメージを示す。Client において検証 AP を起動し、Client の台数を 1~3 台にした場合に、処理が完了するまでの時間を測定した。本検証では画像や映像の配信を行うシステムを想定し、Storage に原本を持ち Client に配信する際にファイルを圧縮し Client に転送する。Read 系、Write 系それぞれについて測定した。

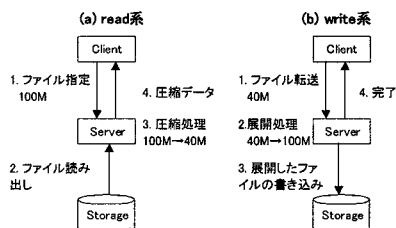


図 4.1 応用 AP の構成

5. 検証結果と考察

5.1 Iozone による性能測定結果

図 5.1 に Iozone による性能測定結果を示す。

基本的なアクセス性能では、SAN+共有ファイルシステムの方がよいことがわかる。これは Ethernet の NFS に対して FibreChannel の SAN 利用している効果が現れている。特に 1 台のホストからのアクセスの場合では約 2.2 倍(Sequential-Read/Write)の性能向上となっている。一方複数からの同時アクセスでは、性能差が少なくなっている。これは複数ホストからのアクセスでは 1 つの Disk へのアクセスが集中するため Disk アクセスのボトルネックにより性能が低下していると思われる。また、製品 B では Random アクセスでは全く性能がでなかった。これは本検証の条件のような大きなファイルへのアクセスでは安定動作しなかったためである。

5.2 応用検証結果

図 5.2 に応用検証における性能測定結果を示す。

全てにおいて SAN+共有ファイルシステムの性能が良いことがわかる。特に Read において約 10%、Write では約 5~20%性能向上しており本手法が有効であるといえる。図 5.2(c),(d)より、この性能差はファイルの読出し(Storage から Server)とファイルの移動(mv: Server から Storage)の処理により発生していることがわかる。特に、Client→Server(Write)では、NFS は右上がりのグラフになっているが、製品 B においてはほとんど変化がなく効果が十分に発揮されていることがわかる。

今回は Client が 3 台までの検証であったが、システムが大規模化するにつれて今回の応用 AP のようなシステムにおいては Disk アクセスの割合が大きくなり(NFS:Read:33%,Write:43%)、システムパフォーマンスに大きな影響を与えており Disk アクセスの性能向上

が重要と言える。したがって SAN+共有ファイルシステムを有効に活用する必要がある。

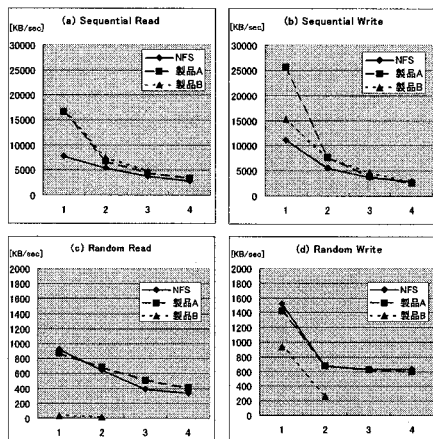


図 5.1 Iozone による性能測定結果

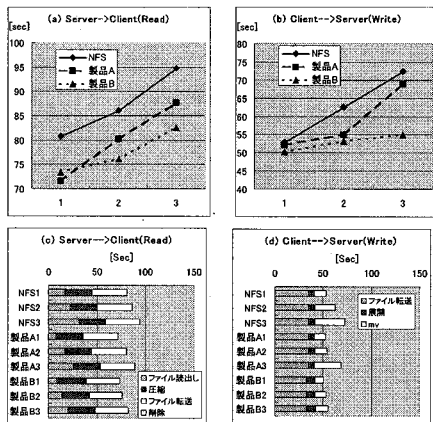


図 5.2 応用検証結果

6. まとめ

Linux クラスタ環境に対して SAN+共有ファイルシステムを用いることがファイル処理性能の向上に有効であることがわかった。これにより大規模システムにおいて問題となる性能低下を防ぐことが可能となる。近年ブロードバンド化が進み、大容量ファイルを扱うことが多くなっており、これに対して各ノード間で複製を持つことは資源の無駄遣いであり、共有ファイルソフトウェア等を使用して有効に活用するべきである。また、Disk 単体の大容量化により今後ストレージの総容量も増大することになり、NFS 等によるファイル共有では限界になると思われる。従って、本検証が示すように共有ファイルシステムを有効に活用することが重要になる。

参考文献

- [1] GFS http://www.sistina.com/products_gfs.htm
- [2] SANergy <http://www.tivoli.com/products/index/sanergy/>
- [3] Centravision <http://www.adic.com/ibeCCtPltmDspRte.jsp?minisite=10000&respid=22372&item=14906§ion=10111>
- [4] LocalClusterEnterprise <http://www.linux.nttcom.ne.jp/solution/product/progartlce/index.html>

※各会社名・製品名等は、それぞれ各社の商標または登録商標です。