

次世代の分散型NASシステム

シーティーシー・エスピー株式会社
苫米地 秀之助 (toma@ctc-g.co.jp)

SpinServer 3300/4100シリーズ・ネットワークストレージ(NAS)サーバは、EthernetベースのNASサーバの使いやすさと、独自の2ステージ分散型アーキテクチャ・ファイルシステムであるSpinFS™のクラスタリング能力を兼ね備えた、次世代NASサーバである。SpinFSアーキテクチャにより、ハイパフォーマンス性を維持したまま、多数のサーバをクラスタリングし、シングル・ネーム・スペースで大容量な分散型ファイルシステムを実現している。

SpinServer3300は、単体性能として、最大ファイル容量22TB、最大ファイルシステム容量22TBをサポートし、処理能力は23,400 IOPS (TCP)を実現している。さらに、クラスタリング構成時は最大512ノード、11,000TB/SingleFileSystemをサポートし、6ノード・クラスタでは117,500 IOPS (TCP)と、サーバを増設する毎に処理能力が直線的に増加する。

2ステージ分散型ファイルシステム

SpinFSアーキテクチャはNetwork機能、ディスク機能そしてSpinFSプロトコル3つの主要素機能で構成されている。(次ページの図参照)

Network機能はClientプロトコル(NFS/CIFS)を終端して、クライアントから送られてくるNFSあるいはCIFSコール(ストレージ・コール)をSpinFSプロトコルへ変換し、自分自身を含めたSpinCluster内の適切なSpinServerのDisk機能へ転送する。SpinFSプロトコルは軽量かつ高効率に設計されており、レイヤ3のルーティングが可能なので、LAN接続だけでなく、MANあるいはWANを介して接続したSpinServerのクラスタ化も可能である。

プロトコル変換時には、SpinCluster内の全てのSpinServerで共有されているVFS Location Data Base(VLDB)を参照して、ファイルのパス名からファイルを実際に管理しているSpinServerを検索する。これにより、SpinFSプロトコルは1回のクラスタ・ホップで目的のSpinServerに到達できる。ファイルアクセスの速度を向上させるためにクラスタ内にミラーリングによるファイルの複製を持つ場合においても、VLDBではオリジナルなファイルの情報だけでなく、ミラーリングしたファイルの情報も管理しているので、書き込み / 読み込みを考慮した、最適な分散処理が可能である。

Disk機能はSpinFSプロトコルをblock I/Oコマンドに変換し、実際のディスクI/Oの全てを実行する。新たにファイルが作られたり、削除されたり、名前を変更されたりした場合には、Disk機能はReplication Data Base Service(RDB)機能を使って、VLDBの内容を更新する。

Network機能とDisk機能が異なるSpinServerであった場合の効率上の負担も15%程度に抑えられている。マルチキャストやブロードキャストもほとんど使わない設計になっているので、従来のNASサーバが持つ容量とパフォーマンスの限界を超えた、複数SpinServerのスケラブルなクラスタリングが可能となっている。

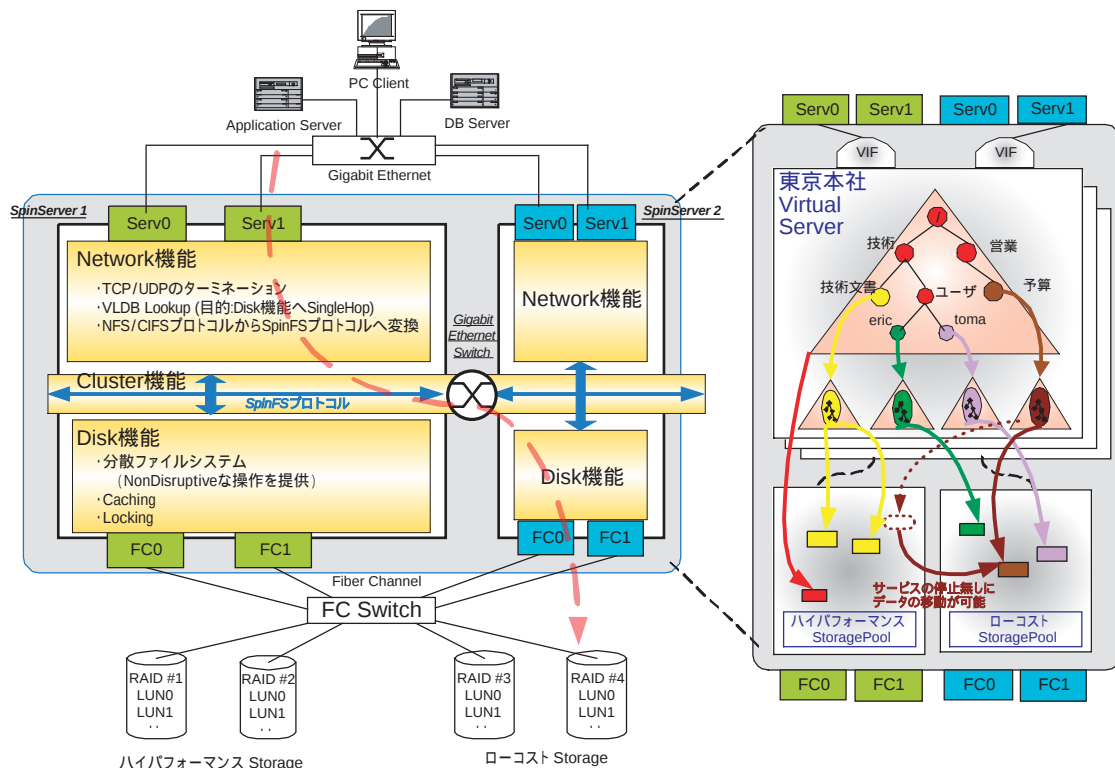
仮想サーバと仮想ファイルシステム

クライアント・アクセスから見えるのは仮想サーバで、これは一つのNISドメインまたはNTドメインに関連付けられた単一の名前空間を持ち、仮想インターフェイス(VI)、ストレージ・プール、仮想ファイルシステム(VFS)の3要素から構成される。N台のSpinServerからなるクラスタ上にはM台の仮想サーバを定義することができる。

ストレージ管理者は様々なストレージ・リソースを仮想サーバ上で論理的にグループ分けすることが可能である。クライアント・アクセスは仮想サーバ上の、指定されたリソースにのみにアクセスが限定されるため、セキュアにリソースを共有し、ストレージの稼働率を向上させることができる。

一方、仮想ファイルシステム(VFS)はSpinServerにおけるストレージの基本ユニットであり、ディレクトリ及びファイルと論理的に結び付けられた概念上の入れ物である。VFSはコピー、移動、ミラーの単位となるばかりでなく、Quotaとアクセス権設定の対象にもなる。論理的には、ファイルはSpinServer内のVFSに保存される。

仮想サーバは一つだけルートとなるVFSを持ち、ストレージ管理者がルートVFSに対して複数のVFSをマウントすることによって、シングル・ネーム・スペースのファイルシステムを構成する。



【参考文献】

- Spinnaker Networks, Inc. 「WHITE PAPER: Spinnaker SpinServer Architecture」 2003/06/25
- Spinnaker Networks, Inc. 「WHITE PAPER: Evaluating Spinnaker NAS Performance」 2003/08/25