

 meta**SAN**

ホワイトペーパー

ここに記載された内容は更新される可能性があります。この文書に記載されている内容はこの文書の発行時点におけるエムアイシー・アソシエーツ株式会社の見解を述べたものです。エムアイシー・アソシエーツ株式会社が、この文書に記載された内容の実現に関して確約するものではありません。また発行日以降については、この文書に記載された内容の正確さは保証しません。

この文書は情報の提供のみを目的としており、明示的または黙示的に関わらず、この文書の内容についてエムアイシー・アソシエーツ株式会社はいかなる保証をするものでもありません。

エムアイシー・アソシエーツ株式会社は、本書に記載してあるすべて、または、一部の記載内容に関し、許可なく転載、または、引用することを禁じます。

metaSAN、metaLAN は Tiger Technology GmbH の登録商標です。

Windows、Windows 95、Windows 98、Windows NT、Windows Me、Windows 2000、Windows XP、Windows 2003 server NTFS は Microsoft Corp. の登録商標です。  
Power Mac、MacOS 9、MacOS X は Apple 社の商標登録です。

その他、記載されている会社名、製品名は各社が所有する商標の場合があります。

バージョン	作成日付	バージョンの内容	総ページ数
1	2008/6/19	新規	10

本書作成、編集、管理  
エムアイシー・アソシエーツ株式会社  
〒103-0004 東京都中央区東日本橋 3-12-12  
櫻正宗東日本橋ビル 9F  
Tel. 03-5614-3757 Fax. 03-5614-3752

# 目次

ストレージエリアネットワーク	4
ファイルレベル共有ファイルシステム	4
ファイルレベル共有とは	4
metaSAN について	4
metaSAN ファイル共有の仕組み	5
metaSAN の特長	5
metaSAN の働き	6
metaSAN の構成	6
メタデータマスターのフェイルオーバー	6
バンド帯域コントロール	7
SAN - LAN フェイルオーバー	7
metaLAN Server	8
metaSAN とアプリケーション	8
1. 映像のノンリニア編集	8
2. ビデオオンデマンド	8
3. HPC システム	8
4. NAS vs. metaLAN Server	9
むすび	9

## ストレージエリアネットワーク

SAN (ストレージエリアネットワーク) は、増加するデータに対してストレージ資源の集中を実現し、より高性能で、柔軟性、拡張性に富み、経済性に優れたストレージアーキテクチャーとして、高い価値をユーザーにもたらしました。

SANとは、複数のコンピュータがファイバーチャネルとファイバーチャネルスイッチを介して接続されたストレージを共有するネットワークです。従来はコンピュータ個々にストレージを接続し、用途に応じた容量を専用に使っていました。こうした接続方式をDAS (ダイレクトアタッチドストレージ) と言います。一方、1990年代初期にファイバーチャネルトポロジーが業界で標準になると、高性能なRAID装置で構築された高信頼性のストレージシステムを複数コンピュータに接続し、高速なデータ転送と、無駄のないキャパシティーマネージメントが可能になりました。ファイバーチャネルでネットワーク(SAN)化されたストレージとコンピュータが接続されることにより、数十メガからギガクラスの大容量のファイルをLANで搬送することなく、直接アクセスすることができます。しかしながら、この集中化したSANのストレージの利用はあくまでRAID装置内部に設定された論理ドライブをSANに接続されたコンピュータがローカルのドライブとしてマウントすることに限られ、複数のコンピュータから同時にSAN内の論理ドライブをマウントし、アクセスすることはできませんでした。万一、複数のコンピュータから同時にアクセスをしようとする、ディレクトリーが破壊され、データ消失という致命的結果をまねくこと

になりました。

こうしたことから、1990年代の後半、ハードウェアリソースとして複数のコンピュータから共有されたSANストレージ内部に保存されたデータを、SANを共有するコンピュータから同時にアクセスすることを可能にし、より効率のよいデータ共有環境を実現できないかということからファイルレベル共有のミドルウェアや、ファイルシステムが開発されてきました。

## ファイルレベル共有ファイルシステム

SGI社のCXFS、IBM社のGPFS、ADIC社のSNFS、Sanbolic社のMeliFS、更には、Red Hat社のGFS等はWindows、Linux、Mac OS XのOSネイティブファイルシステムに独自のファイルシステムを被せて、ファイルレベルの共有を実現しました。これらのユニークなファイルシステムはそれぞれジャーナリング機能やユーティリティーを付加して、ファイルシステムの信頼性や、使い勝手をより良くする努力をしてきました。

しかし、何れのファイルシステムも、あくまで開発元の独自のフォーマットのため、システム上に不整合なことが発生し、ファイルシステムにダメージが起きた場合に、通常のOSに標準装備されているチェックユーティリティーや、復元ユーティリティーを使用して、データの復旧を試みるのが難しいのが現状です。

## ファイルレベル共有とは

ここでファイル共有、又は、ファイルレベル共有ということを確認し、理解する必要があります。ファイルレベル共有はNAS(ネットワーク

アタッチドストレージ)では当たり前のことですが、SANのシステムが普及し始めて初期の頃は、ファイルレベルの共有の前にボリュームレベルの共有という方法が開発されました。SANのストレージ内に構築された論理的なストレージ領域をボリュームと言いますが、このボリュームを複数のコンピュータでマウントし、ボリューム内のデータにアクセスすることができる様にするということです。ボリューム内のフォルダー、データに対してはそのボリュームをマウントしたどのコンピュータからもいつでもアクセスすることはできますが、そのボリュームにデータを書き込んだり、ファイルを更新する書き込み動作は、1台のコンピュータに限定されて、他のコンピュータはファイルを更新したり、追加したりすることはできません。

このボリュームレベルの共有に対して、より、細かいファイルレベルでの読み書きの制御をするのをファイルレベル共有と言います。ファイルレベル共有の場合はボリューム内の複数のファイルに対する読み出しは勿論のこと、複数のファイルへの同時書き込みや更新を可能にします。ただ、1つのファイルに対する書き込みは同時に1台のコンピュータしか行うことができません。

## metaSANについて

metaSANは、このSANアーキテクチャーの価値を更に増進するためにデザインされ、従来に無い、独自のアプローチでSANストレージを管理し、システムの信頼性と性能の向上を図りました。

metaSANは独自のファイルシステムを使用しません。Windowsま

たは Mac OS X で充分その信頼性  
が実証された OS ネイティブのファ  
イルシステム [NTFS(Windows)、  
または、HFS+(Mac OS X)] で初期  
化されたストレージをそのままファ  
イル共有のボリュームとしてサポー  
トします。SAN メンバーのコン  
ピュータは SAN ストレージを OS  
と親和性の高いローカルディスクと  
してマウントし、共有ストレージで  
あることを意識せずにストレージに  
アクセスすることができます。

metaSAN では、SAN ストレ  
ージのボリューム情報を定義する  
Metadata を管理し、SAN に参加  
しているコンピュータにサービス  
を提供するコンピュータをメタデー  
タマスターと呼びます。metaSAN  
は、このメタデータマスターに障害  
が発生した場合、利用可能な SAN  
内のコンピュータを利用して、直ち  
にサービスを引き継ぐフェイルオー  
バーのメカニズム (ダイナミック  
マスター アービトレーション機能)  
を実装しており、より信頼性の高い  
SAN ファイル共有の環境を可能に  
します。

metaSAN は広く柔軟にアプリケー  
ションをサポートし、高速なファイ  
ル共有を実現し、SAN ストレージ  
をコアにした、クリエイティブな  
ワークグループのコラボレーション  
環境としての新しいスタンダードを  
提供しました。

### metaSAN ファイル共有の仕組み

metaSAN のファイル共有環境は、  
まず共有される SAN ストレージを  
中心にして設定されます。共有さ  
れるストレージは論理的なボリュー  
ムで各コンピュータがマウントす  
ることが可能な Windows NTFS、ま  
たは、Mac OSX HFS+ ファイル

システムで初期化されたボリュー  
ムです。共有されるストレージを  
設定すると、次に、そのストレ  
ージにアクセスするコンピュータを  
設定します。SAN に参加する総て  
のコンピュータの ID はネットワ  
クの IP アドレスで登録されます。  
metaSAN はインストールされたコ  
ンピュータ内の各ボードのユニーク  
な番号も管理するべきものとして登  
録します。

こうして、共有ボリュームと共有に  
参加するコンピュータが設定する  
ことにより、SAN のファイル共有の  
条件が整います。metaSAN 環境で  
のファイルレベル共有はファイルに  
対する書き込みを制御することが基本  
的な機能です。ファイルには SAN  
のメンバーのコンピュータの内1台  
だけがその時点で書き込むことがで  
きます。他のメンバーからの書き込み  
はその時点ではブロックされていま  
す。この様にファイル共有ソフトの  
仕事は反対に排他制御をすることな  
のです。

metaSAN の場合、ネイティブファ  
イルシステムのデータベースに代わ  
り、metaSAN のファイル構造を定  
義するデータベースがあります。こ  
のデータはディスクの内部に書き込  
まれていますが、同時に各メンバ  
ーのコンピュータのメモリー中に展開  
されています。

### metaSAN の特長

1. クロスプラットフォーム環境をサ  
ポート：Windows XP、Windows  
2003、Mac OS X、 及 び、RH  
Linux のコンピュータから SAN 内  
でコンテンツを共有することによ  
り、クォータの高度なコラボレー  
ションを実現します。
2. ファイバーチャネル、iSCSI、

Ethernet、及び、InfiniBand イン  
ターフェースで接続されたストレ  
ージに対して、複数のコンピュータ  
からの共有が可能です。

3. 高速なファイルシェアリング：  
複数のユーザーが同じボリューム、  
フォルダーにアクセスし、読み書き  
が可能です。

4. 高いデータ転送性能：最低限の  
オーバーヘッドが最大のデータス  
ループットを保証します。

5. OS ネイティブファイルシステ  
ム：NTFS と HFS+ のもつ完全性  
とセキュリティーをそのまま保持し  
ます。

6. ネイティブ OS でのアクセス権  
アカウントをそのまま参照：パス  
ワードの管理を追加する必要があり  
ません。

7. 透過的で、簡単な使用方法：共  
有ストレージはローカルドライブと  
して常にマウントされます。

8. ノード毎へのバンド幅割当：利  
用可能な SAN のファイバーチャ  
ネルのバンド幅を各 SAN メンバ  
ーのコンピュータに対して、個別に優  
先的に分配したり、平等化させたり  
することができます。

9. LAN クライアントから SAN 共  
有ボリュームへのアクセス：SAN  
メンバーに接続された LAN クライ  
アントに、SAN ボリュームへアク  
セスすることを可能にします。

10. LAN クライアントに対して、  
SAN に直接接続されたメンバーが  
ゲートウェイとして、メンバー間で  
ロードバランシングを行い、LAN  
クライアントのトラフィックを集中  
を分散します。

11. ユーザークォータの割当：ス  
トレージに対するアクセスの競合を  
避けるために個々にクォータを割り  
当てることができます。

12. 24 / 7 のミッションクリティ  
カル業務対応：リアルタイムフェイ

ルオーバー機能によるフルに冗長化されたクラスター型の SAN ストレージ環境を提供します。

## metaSAN の働き

metaSAN は、複数のコンピュータからアクセスの衝突を防ぎ、データを破壊から守り、ファイルの真の共有を可能にします。

図 1 は、metaSAN がサポートする SAN の一例です。大容量の SAN ストレージに SAN スイッチ、ファイバーチャネルを介して、Windows 機や Mac OSX 機、または、Linux 機が接続されます。

これらのコンピュータは NTFS、または、HFS+ で初期化された SAN ボリュームをローカルディスクとしてマウントし、読み書きすることができます。

更に、図-1. 右下にある FC インターフェースを持つテープバックアップは、SAN メンバーのコンピュータから metaSAN のストレージ共有

とバックアップユーティリティを使用して、データのバックアップをテープ装置にとることができます。この方法は、通常の Ethernet を介してデータを転送するバックアップサーバーソリューションに比較して、高速、且つ、低コストな LAN フリーバックアップソリューションです。

## metaSAN の構成

典型的な metaSAN の構成は SAN に参加するコンピュータ (SAN メンバーと呼ばれます。) と、これらのコンピュータとファイバーチャネル、iSCSI、または、InfiniBand インターフェースで接続された共有ストレージから構成されます。

metaSAN では、共有されるストレージの内部情報、及び、参加メンバーのアクセス状況は Metadata として管理されます。この Metadata を SAN メンバーを代表して管理する役割をメタデータマスター が果たします。メタデータ

マスターは共有ボリュームを管理し、他のメンバーが共有ボリュームにアクセスする場合に排他制御を行うことにより、データ破壊を防ぎます。共有ストレージにメンバーがアクセスしようとする、メタデータマスター は必要なサービスデータを提供し、オペレーションをうまく完了させます。このサービスデータは LAN で渡されますが、実際のユーザーデータは高性能なファイバーチャネルを介してアクセス、転送されます。

## メタデータマスターのフェイルオーバー

メタデータマスターのフェイルオーバーとは、メタデータマスターに障害が発生した場合に、その管理権限が自動的に他のメンバーのコンピュータに譲渡される機能のことです。

Mac OS X や、Windows が稼動するメンバーのコンピュータで、事前にメタデータマスター にならないように設定したコンピュータ以外、どのコンピュータもメタデータマスター になることができます。

metaSAN においては、メタデータマスター及び、他の SAN のメンバーは常時相互をチェックしており、一定の時間以内にメタデータマスター がレスポンスを返してこない場合、メタデータマスター に障害が発生したと判断し、Metadata のマスター権限を他の正常な SAN メンバーにフェイルオーバーさせます。この機能により、ボリュームを管理するメタデータマスター はダイナミックに交替することができます。フェイルオーバーが起こる状況になった場合は、他の SAN メンバーがメタデータマスター を引き継ぎ、他のメンバーからの Metadata に対するリ

図 1



クエストを処理し続けます。この意味で、metaSAN はノンストップのオペレーションを可能にするクラスター型のストレージネットワークと言えます。

### バンド帯域コントロール

metaSAN にはファイバーチャネル等の高性能ストレージインターフェースの利用可能バンド帯域を各 metaSAN メンバーに指定した量割り当てることができます。例えば、A のメンバーには 50MB/ 秒、B のメンバーには 100MB/ 秒、更に、C のメンバーには 200MB/ 秒という様に、それぞれ許容されるバンド帯域を指定し、早いもの勝ちに無秩序にバンド帯域が取られる状況を回避することができます。バンド帯域コントロール機能により、各ワークステーション、サーバーにおいて、どうしても確保すべきデータ転送帯域が必要な場合、プリセット方式で帯域を確保することが可能となり、ミッションクリティカルなサーバー、ワークステーションの高いサービス性能を実現することができます。

### SAN - LAN フェイルオーバー

SAN メンバーで万一、ストレージインターフェースに障害が発生した場合、metaSAN はそれまでの共有ボリュームに対するアクセス経路を Ethernet にフェイルオーバーさせ、それまでのオペレーションの止めること無く継続することを可能にします。このことを SAN-LAN フェイルオーバーと呼びます。ストレージインターフェースによって転送されていたデータストリームはその時点でネットワーク経由で流されることとなります。この時、ユーザーには

SAN - LAN フェイルオーバーを意識させること無くオペレーションは継続されますので、高いストレージの可用性を可能にします。

また、この機能は metaSAN のもつ他のユニークな機能 metaLAN で更にその有用性を発揮します。

### LAN クライアントによる SAN ボリューム共有

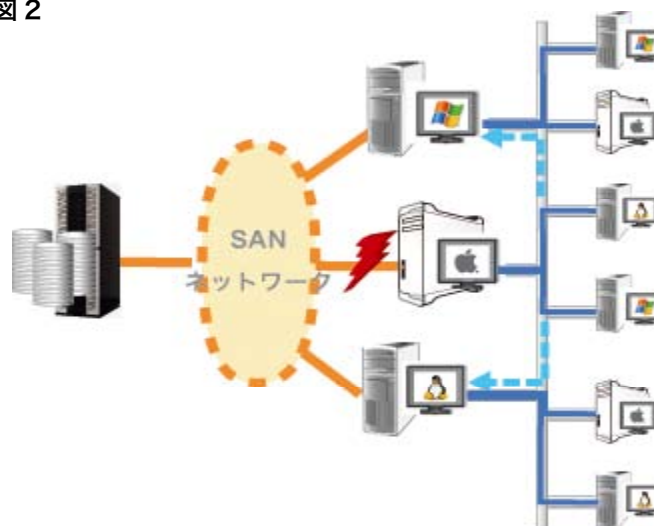
metaSAN には更にユニークな機能としてネットワーククライアントから、SAN ボリュームをあたかもローカルディスクの様にマウントすることを可能にする metaLAN ゲートウェイという機能があります。

metaSAN をインストールされたコンピュータをゲートウェイとして LAN クライアントが SAN のメンバーになることを可能にします。

同 ドメイン上に存在する

metaSAN をインストールされたサーバーは、metaLAN をインストールされたノード (metaLAN client) からの SAN の共有ボリュームへのアクセスを仲介します。metaLAN client は SAN の共有ボリュームを他のサーバーのサービスプロトコルに依存するネットワークボリュームとしてではなく、ローカルディスクとして、マウントすることができます。また、ゲートウェイとして機能する metaSAN メンバーは同一ドメイン上に複数存在することができ、各メンバー負荷をバランスを図ることができます。更には、ゲートウェイとして機能する metaSAN メンバーに障害が発生し、metaLAN client からのアクセス仲介を行うことができなくなった場合、自動的にデータのストリームを他の metaSAN ゲートウェイが引継ぎ、metaLAN からのアクセスを継続させることができます。

図 2



## metaLAN Server

以上の様に、metaSANにはmetaLAN clientに対するゲートウェイの機能があります。このゲートウェイの機能だけをサーバー機能としたソフトウェア製品がmetaLAN Serverとして販売されています。metaLAN ServerはmetaLAN clientに対するサーバー間での負荷バランス、および、フェールオーバー機能を有しています。その意味ではクラスター型のゲートウェイサーバーということができます。

## metaSAN とアプリケーション

metaSANはファイル共有を実現するミドルウェアです。ファイルを複数のコンピュータで同時に共有することが可能になることで、従来の業務システムでは実現が難しかった高い生産性や、効率性を可能にしました。以下に挙げるのはその一例です。

### 1. 映像のノンリニア編集

metaSANを使用することにより最も効果的な業務アプリケーションのひとつです。いうまでもなく、映像データはデータそのものが大変大きく、近年の高精細化に伴い、ますますファイルサイズが大型化しています。HD非圧縮データ(HDTV, 1080/60i, 4:2:2非圧縮)の場合、約180MB/秒のデータをストレージに取り込み、2時間の映像の場合はファイルサイズが約1.3TBになります。このような大容量データを取込み、編集して完パケとしてデータを次行程に渡すのは通常のネットワーク転送を行った場合、それぞれの工程で長時間の転送時間を必要と

していました。

しかし、ドライブの大容量化、RAID装置の高速化、ストレージインターフェースの高速化が進み、ストレージをコアにしたSANを構成することにより、映像編集ワークフローの個々の工程をリアルタイムに行うことができるようになりました。しかし、各工程間のデータの転送はその容量の大きさからままなりません。そこで、metaSANを使用した、各工程間でのファイル共有環境を実現することで従来にない効率的なワークフローを実現し、大幅な処理工程の短縮を図ることが可能になります。

### 2. ビデオオンデマンド

近年、映画や、TV局が作成したドラマ/アニメ/ドキュメント等をテレビ放送や、レンタルビデオだけではなくインターネットによるパソコンや、セットボックスを使用したビデオオンデマンドストリーミングサービスが始まっています。ビデオオンデマンドサービスの場合、不特定多数の視聴者がネットワークを通して、ストリーミングサーバーにアクセスし、好みの映画や、ドラマを視聴するため、ピークデマンドに対応可能なサーバーシステムを構築する必要があります。しかし、2時間の映画1本はファイルサイズは2GBを超える容量です。複数のサーバーにそれぞれ大容量ストレージを接続し、数千、数万のコンテンツをコピーするのは大変な時間を要することになり、且つ、ピークデマンドに合わせてサーバー/ストレージを用意するのも、無視できないコストがかかります。そこで、ピークユーザーのログインに対応可能な複数のサーバーと、それらのサーバーが共有するスト

レージを構成し、共有ストレージに収納されたコンテンツを総てのサーバーが共有することにより、効率的なストレージの構成と、大容量コンテンツを運営、管理することが容易になります。metaSANは国内で大きな会員を誇る商用のストリーミングサービスのシステムの中で使用されています。

### 3. HPC システム

HPCシステムとは高性能コンピュータを複数台使用して、一つのジョブを並列に処理するシステムのことで、High Performance Computingの意味です。または、クラスターシステムとも呼ばれます。

通常、複数のコンピュータで一つのジョブを並列に処理する場合、大容量のデータを同時に複数のノードコンピュータが読み込み、それぞれのノードが担当する部分の計算結果を自分のワークファイルとして自分のローカルディスクに書き込みます。最後に、クラスターコントローラによって、それらのデータが集められ、一つの計算結果として書き出されず。

通常の場合では、クラスターのコントローラであるコンピュータの指示により、それぞれのノードがネットワーク経由で共有ボリュームからデータを取込み、処理をしますが、画像や、映像といった大型のデータの場合、ネットワークでデータを転送してしまうと、分散して処理をするより、データの転送に時間を要してしまい、分散処理をする意味が薄れてきます。metaSANの様な共有ファイルシステムを使用することにより、HPCに参加するクラスターノードのコンピュータにとって、大容量のデータを収納したストレージ

をローカルディスクとしてマウントしておけば、クラスターコントローラからの演算の命令を受けた段階で、一斉に共有ディスク内のデータを読み込み、更に、演算結果を同じくローカルディスクに書き出すことにより、並列でのコンピュータ演算はデータ転送を待つことなく高速に行うことができます。

#### 4. NAS vs. metaLAN Server

NASはネットワークを介して共有されるストレージのことです。近年NASベースのストレージが中小規模のグループストレージとして多く採用されるようになりました。限られたユーザーのためのバックアップや、データ共有を実現することが可能で比較的アクセスの少ないストレージとしては大変有効なストレージです。しかし、NASはあくまでEthernetを経由して複数のクライアントからアクセスされますので、多数のクライアントから同時に大量のデータを転送した場合、ネットワークが大きなボトルネックになります。

metaLAN Server/metaLAN clientはこの問題を解決します。従来、NASベースで大容量データを読み書きしているワークグループにおいて、データ転送が多重化すると、個々の転送が相乗的に劣化してゆきます。

metaLAN Server/metaLAN clientの環境では、metaLAN clientからmetaLAN Serverをゲートウェイとして、SANのボリュームを直接アクセスすることができます。複数のmetaLAN Serverが存在する場合、metaLAN clientはダイナミックにSANへのゲートウェイサーバーを選択します。そして、あるクラ

イアントとSANストレージとのアクセス経路に障害が発生した場合、それまでそのクライアントとストレージのアクセスのゲートウェイとなったmetaLAN Serverから、他のmetaLAN ServerやmetaSANワークフローを代替のゲートウェイとして、オペレーションをフェイルオーバーし、クライアントからのSANストレージへのアクセスは停止することなく継続することが可能です。

以上のことから、NAS vs. metaLAN Server/metaLAN clientというテーマにかんしては、NASは低いアクティビティのストレージとして不特定多数のクライアントがデータを共有する上で有効なストレージといえます。また、metaLAN Server/metaLAN clientのシステムでは比較的大容量データを同時に複数転送することができ、大容量画像、映像などの処理、チェックをネットワーク経由で行う場合に有効なストレージシステムといえます。

#### むすび

metaSANは以上の様に、WindowsやMac OS Xを搭載したコンピュータがストレージを共有するSAN環境でこれたのコンピュータによるファイルレベルの共有を可能にし、ワークグループの生産性を向上させ、ネットワーククライアントへの透過的、効率的なサービスを実現し、コンピュータの有効な運用を可能にする、クラスター型ファイル共有システムといえます。metaSANは世界中に多くのミッションクリティカルなコンピュータシステムサイトにインストールされ、システムの生産性向上に役立っています。

