SAN (Storage Area Network)
ー分散システムでの効率的データ管理のソリューションー

西山 智
(株) KDD研究所

近年，銀行・証券・通信等のミッションクリティカルな基幹システムを，メインフレームではなくUNIXサーバ等からなる分散システムとして構築する事例が増えている．この背景には，サーバ単体の性能向上に加えてTUXEDO，ENCINA，CORBAといった分散ミドルウエアやWEBサーバを用いた3階層クライアントサーバ構成の分散処理による処理能力の向上が容易となったことが大きい．また運用管理のミドルウエアを用いることで，メインフレーム並みとはいえないが，複数のサーバを統合的に管理できることも理由の1つである．基幹システムの分散化によって，システムの構築コストや運用コストの低減化が実現されてきている．しかしながら，これらのサーバが使用する2次記憶装置は各々のサーバに接続されてばらばらに存在することとなり，記憶容量管理やデータバックアップ等のデータ管理を効率的に実施することが困難となっている．このような問題に対するソリューションとしてSAN (Storage Area Network) が登場してきた．

●コンピュータの2次記憶装置専用のネットワーク

SANはコンピュータと2次記憶装置の間，あるいは2次記憶装置間を直接接続し，コンピュータからの記憶装置の読み書きや，2次記憶装置間でのデータの移動を行うための専用のネットワークである (図-1)．これに対し従来から存在するコンピュータにSCSI (Small Computer System Interface) などのチャネルで直結した2次記憶装置をSAS (Server Attached Storage) ☆1，またNFS (Network File System) プロトコル等によりLANを介してアクセスする形態の2次記憶装置をNAS (Network Attached Storage) と呼ぶ (図-2)．従来サーバ間でばらばらだった2次記憶装置を専用のネットワークで結びつけることにより，自由に記憶装置をサーバに割り当てたり，またLANに負荷をかけずにバックアップが行える．

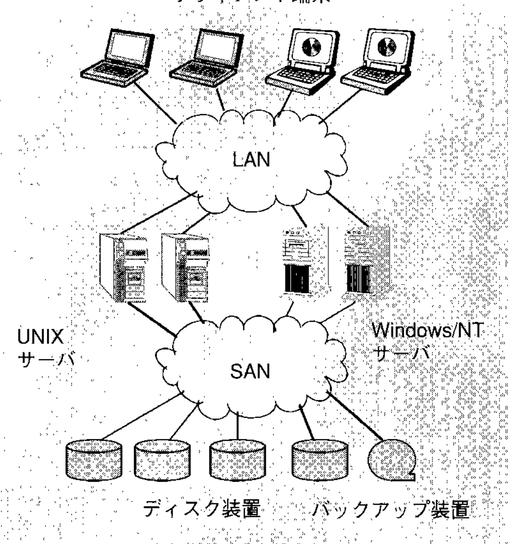

図-1 SAN (Storage Area Network)

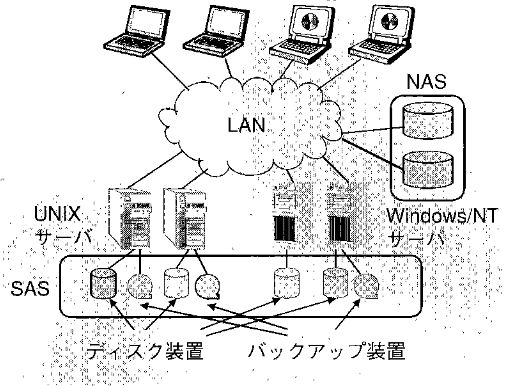

図-2 SASとNAS：従来の2次記憶装置の接続形態

☆1 SASという単語は場合によりSAN Attached Storageという正反対の単語の略として使われる場合もあるので注意が必要である．

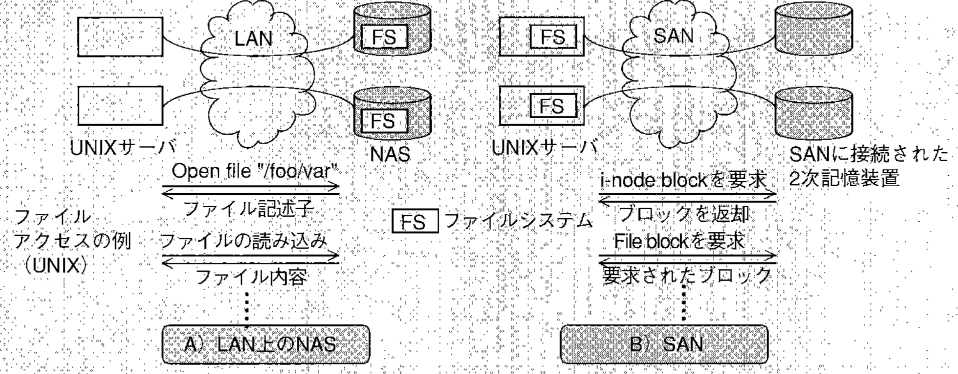

図-3 SANとNASの違い

ではNAS専用のLANを設けるとSANと呼べるだろうか？ 通常NASはファイルシステムを提供するサーバとして動作し，コンピュータはそのクライアントとしてファイルシステムにアクセスする．これに対し，SANではファイルシステムはコンピュータ側に存在し，コンピュータと2次記憶装置の間はより低水準な「ブロック」を単位とするコマンドがやりとりされる．したがってコンピュータは，NASへのアクセスと比較してよりオーバヘッドの少ない方法でSAN上の2次記憶装置にアクセスできる．**図-3**はUNIXでの例であるが，NASへはファイル"/foo/var"をオープンしてファイル記述子をもらい，そのファイル記述子を指定して"/foo/var"ファイルの内容を読み書きするという手順でアクセスする．一方SAN上の2次記憶装置に対しては，SASの場合と同様に"/foo/var"ファイルのi-nodeブロックを指定してその内容を読むことで，次にファイル内容のブロックに直接アクセスできる．

●ファイバチャネル－SANを支えるネットワーク技術

SANが提唱されるようになった技術的基盤として，ファイバチャネル（Fibre Channel）☆2と呼ばれる高速な光シリアルインタフェースの登場がある．今から10年前，UNIXワークステーションが一般化した頃を思い起こす．当時，ネットワークの標準的な転送速度は約1Mバイト／秒（イーサネット），一方記憶装置の転送速度は10Mバイト／秒（SCSI）であり，ネットワークのほうが10倍遅い状態であった．NFS等によりネットワークを介した2次記憶装置（主にディスク）の共有は実現されていたが，データベースなど高負荷の応用に利用することは性能的に困難であった☆3．現在，UltraSCSI2により2次記憶装置の転送速度も80Mバイト／秒まで増加した．しかしネットワーク側の転送速度はATM（Asynchronous Transfer Mode），ギガビットイーサネット，HIPPI（High-Performance Parallel Interface），そしてファイバチャネルなどにより100Mバイト／秒近くまで到達している．ネットワーク速度がディスク転送速度とほぼ同等となることで，ネットワークを介してのディスクアクセスが高負荷の応用に対しても実用的になり，SANという概念が実現できるようになった．SANは特定のメディアによるネットワークではなく，2次記憶装置の専用ネットワークという概念にすぎない．したがってSCSIスイッチを中心としたSCSIネットワークやESCON（Enterprise System Connection）やSSA（Serial Storage Architecture）など従来のチャネル接続によるネットワーク，さらにはATMやIEEE1394等の新しいネットワークでもSANは構築可能である．しかし現状ではファイバチャネルのネットワークが拡張性やケーブル延伸距離などの点で有利なことから，SANはファイバチャネルを中心としたネットワークとして考えられている．

ファイバチャネルとは

ファイバチャネルは，SCSIコマンド，IPやオーディオ・ビデオといった複数の上位プロトコルをサポートするANSI標準のプロトコルであり，物理媒体までの下位プロトコルを包含している1)．ファイバチャネルはチャネルという名前の通り，ハードウェアにより高速（現在100Mバ

☆2 当初Fiber Channelという単語が使用されていたが，後にシールド付き撚り線等ファイバ以外の物理インタフェースが追加されたため，現在はFibreという単語を使用している．
☆3 NFSの性能問題の原因は，ネットワーク帯域の不足のみならず，IP上のアプリケーションとして転送していることも挙げられる．

レイヤ	機　能	備　考
FC-4	上位プロトコルマッピング	SCSI，HIPPI，IP，ビデオ等
FC-3	共通制御	
FC-2	信号制御	フレーム検出，順序検査，フロー制御
FC-1	転送符号化	8B/10B 符号化
FC-0	物理媒体	STP (27m), MMF (500m), SMF (10Km) 他 伝送速度 133Mbps ～ 1Gbps (4Gbps 標準化中)

表-1　ファイバチャネルのプロトコル構造

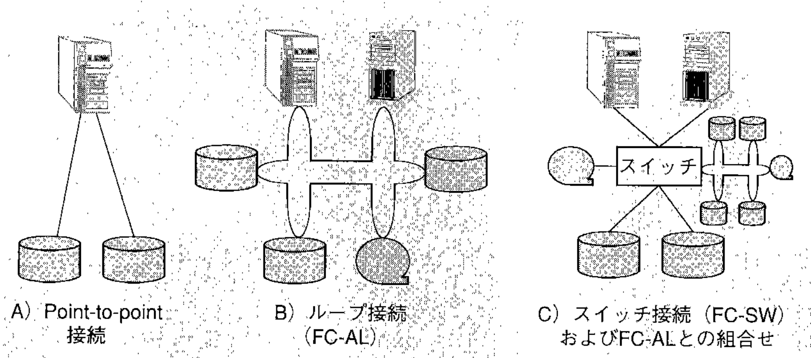

図-4　ファイバチャネルの接続形態

イト／秒)，高信頼かつ伝送遅延の少ないネットワークを実現する．ファイバチャネルは**表-1**に示すプロトコル階層からなる．

ファイバチャネルは**図-4**に示すように必要に応じて，ポイント・ツー・ポイント接続，ループ接続 (FC-AL)，スイッチ接続 (FC-SW) の接続形態が選択できる．接続できるノード数はFC-ALで126，FC-SWも使用すると最大1600万ノードまで可能である．伝送速度は現在1Gbps (100Mバイト／秒相当) であるが，SCSIが半2重なのに対し，ファイバチャネルは全2重であるため実効上その倍あるということになる．

使用できる物理媒体は多数あり，シングルモードファイバ (SMF，最大10Kmまで)，マルチモードファイバ (MMF，最大500mまで) およびシールド付き撚り線 (STP，1Gbpsでは最大27mまで) など，必要な延伸距離に応じて適切な物理媒体を選択できる．また，符号化にはIBM社が考案した8B/10B方式[2) を使用し，誤りの少ないデータ伝送を実現している．

ファイバチャネルの提供するサービスは接続形態や信頼性に応じていくつかのクラスに分かれている．代表的なクラスは，クラス1 (特定のノードとの間を回線接続するクラス)，クラス2 (コネクションレス型で，リンクごとに送達確認を行い，不特定多数のノードと通信できるクラス)，クラス3 (クラス2と同様にコネクションレス型で不特定多数と通信できるが，送達確認が行われないクラス) である．SANでは一般にクラス2または3が使用される．またファイバチャネルはSCSIコマンド，IPなど複数の上位プロトコルを同一のチャネル上に伝送できるので，IP上のプロトコルによるリモートバックアップなどもLANを経由せずSANの上で実施できる．

ファイバチャネルによるアクセスの高速化

ファイバチャネルはその高速性が注目されることが多いが，ファイバチャネルで高速なSANを作ればすべての問題は解決するのではない．まずアクセス時間について考える．一般に2次記憶装置に対するアクセス時間において，接続インタフェースでの転送時間のオーバヘッドは意外と小さい．たとえばハードディスク装置から4Kバイト程度のデータを読み出す場合，10Mバイト／秒のSCSI接続の場合でも転送時間は0.4ミリ秒程度であり，通常のディスクアクセス時間の1/10以下である．したがって，アクセス速度が遅い2次記憶装置に対してはファ

イバチャネル接続を行っても，アクセス時間の短縮はあまり期待できない．より高速なディスク装置に置換する方が容易にアクセス時間を短縮できる．次に伝送容量はどうであろうか．ファイバチャネルは100Mバイト／秒の伝送速度を持つが，これは40Mバイト／秒のUltraSCSIの2.5倍にしかすぎない．したがって，アクセス量の多い2次記憶装置を複数台FC-ALによりループ接続した場合，各々の記憶装置に対してはUltraSCSI接続の場合と同程度の伝送容量しか得られないこととなる．このような場合，より高伝送容量を得るためには，ループ内の台数を抑えて，サーバとはスイッチ接続するなど，処理負荷に応じたネットワーク設計が必要となる．

●SANがもたらす利点

拡張性

一般に情報処理機器は短期間に性能が向上する．したがって基幹システムなどの構築では，構築コストの観点から必要最低限の処理能力でスタートし徐々にその能力の向上が図れること，すなわち拡張性を実現することが望ましい．2次記憶装置についても同様である．しかし従来のSCSI接続されたSASの形態では，SCSIインタフェース当たり接続できる2次記憶装置数に制限（最大で16ノード）があり拡張性を実現することは難しい．またケーブル長が25メートルまでに制限されており，設置場所の制限が厳しい．特に大規模システムでは電力や空調など，設置面積以外にも追加の2次記憶装置をサーバの近くに設置する際に問題となることは多く，ケーブル長制限は大きな問題となりやすい．ファイバチャネルによるSANでは，接続できる記憶装置の数や最大延長距離の制限がゆるいため，記憶装置の拡張性を実現することが容易である．

データの共用

複数のコンピュータ間で2次記憶装置を共用できる．これにより企業内で，あるシステムから別のシステムへデータを単にファイル転送するという処理が省ける場合が多い．したがってたとえば，オンライン処理を行っている基幹システムに蓄積された情報を，バックオフィスにあるデータウェアハウスシステムで加工する，といった利用が可能となる．

データの一元管理

SAN内の2次記憶装置を管理するためのツールが開発されており，そのツールによりSAN上に分散した2次記憶装置を一元的に管理することができる．これにより記憶領域の割り当てにおける分割損が減少することが期待できる．アクセス制御，データのバックアップなど，データ管理は専門のソフトウェアや高速なバックアップシステムで効率的に実施でき，セキュリティやデータの信頼性の向上が期待できる．また個々のコンピュータシステムの管理者はデータに関する運用管理から解放され，全体としてシステム運用コストの低下が期待できる．

SANの管理では，従来のSCSI装置で使用されていたSES（SCSI Enclosure Services）やSAF-TE（SCSI Accessed Fault-Tolerant Enclosure）と呼ばれる管理プロトコルに加えて，SNMP（Simple Network Management Protocol）も使用され，2次記憶装置のみならずSANを構成する機器が一元的に管理できるようになっている．ほとんどのツールがWebによるユーザインタフェースを提供し，管理を容易としている．

非常災害対策

ファイバチャネルによるSANは10Kmまで延長できるため，別の建物など物理的に離れた場所に2次記憶装置を設置しデータのコピーを置くことができる．これにより，重要な情報の災害に対する安全性を向上することができる．さらに，現在ファイバチャネルをATMあるいはSDH（Synchronous Digital Hierarchy）等で延長する研究が進められており，**図-5**に示すように将来的には広域災害にも耐え得る距離（たとえば東京／大阪間）でのSANの構築も可能となる．また，専門業者による遠隔データバックアップサービスも可能となる．

LANの負荷軽減

一般的にLANでは負荷が高いとされているNFSやファイル転送のパケットを取り除けるので，LANのトラフィックの軽減が図れる．従来，大学や企業でNFSサーバなどのNASによりどの端末からでも同一の環境が得られるように構築している場合があるが，バックボーンLANに重い負荷をかけている場合が多い．SANを用いることでLANに負荷をかけずにこのような分散環境が実現できる．また，バックアップもSAN内に閉じて行われるのでLANに負荷をかけることがない．

●SANの今後の展開

試しに米国のサーチエンジンAltaVistaでStorage Area Networkをキーワードに検索すると約2000件のページがヒットした．これらのページに対して，さらに".com"を含むページを除外すると約170件と1/10以下になる．この中には日本や欧州等，米国以外の企業や団体のページ，それに以下に述べる標準化団体のページも含まれており，それらも除くと10件程度にすぎない．これはメーカ主導

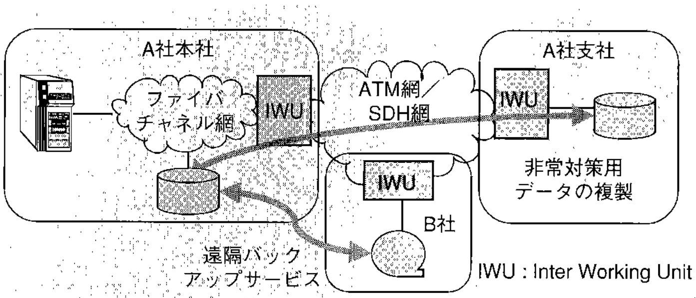

図-5 非常災害対策の例

で進んでいるSANの研究開発状況を端的に表している．

相互接続を容易とするための活動

現状ではSANを構成する機器間での相互接続性に問題がある場合もある．このため，相互接続性を保証するための活動が進められている．現在標準化活動を行っている有力団体としてSNIA (Storage Networking Industry Association) がある[3]．SNIAは1997年後半から活動を開始し，100程度の団体（主に企業）が参加して，標準化活動を実施している．現在はファイルシステム，ストレージネットワーク管理，バックアップの3つの作業部会があり，ベンダ間の相互接続性を目的とした業界標準を策定している．

一方，SANの中核となるファイバチャネル技術については，FCA (Fibre Channel Association) という団体が現在活動を行っている[4]．ファイバチャネルは1992年に3つのコンピュータメーカにより設立されたFCSI (Fibre Channel Systems Initiative) という団体がANSIに標準案として提案したのがはじまりである．その後FCSIが解散した後，FCAが標準化活動や相互接続試験などを進めている．

FCAには，約150以上の団体が加盟している．FCAにはFCA-Japanと呼ばれる日本支部もあり，現在28の団体（やはり大半が企業であるが，官公庁等も含む）が参加して相互接続試験や展示会でのデモを実施している．

そのほか，テストベッドの運営で有名なニューハンプシャー大やミネソタ州立大がファイバチャネルについてもテストベッドを運用し，相互接続性の確認を行っている．

よりSANを使いやすくするための研究

SANに関連する研究は，SANそのものではなくファイバチャネルによるネットワークの研究が中心であった．ファイバチャネルやスイッチの性能見積もり，設計等に関する発表が1995年頃に集中しており，その後ファイバチャネルをATM網やSDH網により中継する方式の検討が進められてきた．特にファイバチャネルとATM網のゲートウェイ (IWU) については，1990年代前半から研究が始まり，国内でも超高速ネットワーク技術研究所やNTTが1990年代後半にIWUを試作している[5), 6)]．SANが使用するファイバチャネルのクラス（クラス2や3）ではリンクごとのフロー制御が行われる．このためIWUでは伝送速度の違いを吸収するためのバッファ制御や，IWU間でのATM網のフロー制御が重要となる．NTTのIWUは155MbpsのATM回線を使用した場合，最大約100Mbpsのスループットが得られたと報告されている．将来IPバックボーンの高速化に伴い，IPバックボーンの上でファイバチャネルを中継することも考えられる．この場合，IP網での遅延に対するSCSIコマンドの応答制御，帯域保証，データの改ざんや傍受を防ぐための暗号化などが研究課題となると考えられる．

今後SANが一般化すると，企業や大学などではPCやUNIXサーバのデータはすべて，SAN上の2次記憶装置に安全に管理されるようになるかもしれない．その場合，「パソコンのディスクが壊れまして」という言い訳は過去のせりふとなってしまうだろう．

参考文献
1) ANSI X3T11 Fibre Channel Standards and Draft Standards.
2) Widmer, A. and Franaszek, P.: A DC-Balanced, Partitioned-Block, 8B/10B Transmission Code, IBM J. Res. Develop., Vol.27, No.5 (1993).
3) http://www.snia.org
4) http://www.fibrechannel.com
5) Anzaloni, A. et al.: Fibre Channel FCS/ATM Interworking : Design and Performance Study, GLOBECOM'94.
6) 藤岡他: Fibre Channel-ATM接続の性能評価, 1998年電子情報通信学会総合大会, p.320 (1998).

（平成11年8月5日受付）