

ETERNUSディスクアレイの仮想化技術

Virtualization of ETERNUS Disk-array Subsystem

あらまし

ストレージシステムにおいても必要な資源を必要なだけ配分できる「仮想化」は重要なトレンドとなっている。ディスク装置においてはいわゆるストレージ仮想化だけではなく、様々な階層において仮想化の概念が取り入れられており、最近でも多くの「仮想化」の技術を利用した機能が生まれている。ETERNUSディスクアレイでもボリューム仮想化機能をはじめいくつかの技術が利用できる。

またストレージにとってはサーバ仮想化への対応も重要な要件となる。通常のI/Oに対してはサーバ仮想化には特殊な対応が不要である。しかし、最近のハイエンドのディスクアレイ装置では当然の機能となっているボリューム複製機能では仮想化されたサーバからの資源の見え方の変化に対応する必要がある。ETERNUSディスクアレイのアドバンスド・コピー機能ではETERNUS SF AdvancedCopy Managerでサーバ仮想化環境への対応を行っている。

Abstract

Recently, virtualization has become an important technology for storage systems. Many kinds of virtualization features are implemented in various layers of disk-array subsystems. This paper explains some virtualization features which have been introduced to ETERNUS disk-array subsystems recently. A thin-provisioning feature is one of them. It is also important to support server virtualization in storage systems. Normal read/write access does not need any additional feature to support server virtualization. But the volume replication feature, which has become commonplace in high-end disk-array subsystems, needs to adapt to server virtualization. Recognition of storage resources from operation systems may change in a server virtualization environment. ETERNUS SF AdvancedCopy Manager, which is a utility program for controlling ETERNUS disk-array's advanced copy feature, adapts that feature to server virtualization.



赤坂 勉 (あかさか つとむ)
ストレージシステム事業部 所属
現在、ディスクアレイ装置の開発に
従事。



安福洋一 (やすふく よういち)
ストレージシステム事業部 所属
現在、ディスクアレイ装置の開発に
従事。



川本睦史 (かわもと あつし)
第二プラットフォームソフトウェア
事業部 所属
現在、ストレージ管理ソフトウェア
の開発に従事。

ま え が き

情報システムの様々な分野において仮想化と呼ばれる技術が物理的なリソースの境界を隠蔽し、論理的なリソース分割を行うことで必要なだけのリソースを最適な形で提供するものとして注目されてきている。

そもそもサーバ、ストレージの分野においては広義での仮想化は決して目新しいものではない。サーバにおけるSMP (Symmetric Multi Processor) やマルチタスキング、仮想記憶、ストレージにおけるRAID、論理ディスク・ボリュームなど既に技術的にも当たり前なものとして利用されているが、物理リソースと論理的なリソースを切り離してリソースを分割し提供しているという意味では仮想化技術と考えることができる。

一方、現在サーバ仮想化という場合には一つのサーバを論理的に複数のサーバに分割するVM (Virtual Machine) の機能を指すことが多く、ストレージ仮想化という場合には、複数のストレージ装置をまたいであたかも一つのストレージ装置に見せる機能を指すことが多い。富士通ではストレージにおいて、この意味でのストレージ仮想化に対してVS900シリーズで対応している。その紹介は本特集掲載の「RAID仮想化技術」に譲ることとし、本稿では富士通のSAN対応ディスクアレイ装置であるETERNUSディスクアレイにおける仮想化への対応について二つ紹介する。

まず、はじめに上述の狭義のストレージ仮想化の定義には当てはまらないが、ETERNUSディスクアレイとして仮想化を実現している機能について述べる。つぎに、サーバ仮想化に対しては、論理分割されたゲストに対してETERNUSディスクアレイのボリューム複製機能であるアドバンスト・コピー機能が行っているゲストのディスクリソースのバックアップについて述べる。

ETERNUSにおける仮想化技術

ディスクアレイ装置では、従来より複数の物理ドライブを仮想的に一つのドライブに見せるRAID、一つあるいは複数のドライブを物理的な境界に制約されることなく複数のボリュームに見せる論理ボリュームなどの仮想化技術が実用に供されてきた。

本章においてはETERNUSディスクアレイに導入されている機能のうちでも論理的な構成を物理的な構成から分離するという意味での仮想化の技術を利用することで、ETERNUSディスクアレイ単体での利用において利便性を向上させている各種の機能について紹介する。

ストレージシステムには常にコストの低減が求められる。これには装置導入時のコストの低減と構成変更時のコストの低減がある。いかに安く導入し、いかに簡単に使えるかである。これを解決する技術として、ETERNUSディスクアレイではいくつかのボリュームの仮想化技術を実装している。

従来の方法では、ボリューム定義決定時に容量とディスクの割当てを決定する必要があり、後からそれを変更することは不可能であった。このためにしばしば導入コストの増大と管理コストの増大を招くという問題があった。つまり初期導入時に将来必要な容量を予測してボリュームの定義を行った場合、導入当初はほとんど使用されないディスクの導入が必要となるため、導入コストが増大する。また、初期導入時に必要最小容量でボリュームの定義を行った場合、ボリュームの容量追加時に特別な処理が必要となるため、管理コストが増大する。そもそも将来必要な容量を数年にわたって正確に予測することは困難であり、余裕を持って対処しようとするほど結果的には無駄なコストをかけなければならなかった。

また論理ボリュームを構成する物理ディスクの配置はしばしばアクセスの偏在による性能問題を引き起こすことがあったが、従来はこれを解決するためには構成する物理ディスクの負荷の高い論理ボリュームの内容を、構成する物理ディスクへの負荷が低い論理ボリュームへ移動させる必要があり、これらの操作には長時間の運用停止が必要であった。

ETERNUS8000およびETERNUS4000では、出荷当初よりこれらの問題を解決するために、論理ボリュームの容量拡張やRAID構成の変更を運用中のボリュームに対して動的に行う以下のような機能を提供している。

(1) RAIDマイグレーション

ある物理ディスク上にある論理ボリュームを別の物理ディスクへ動的に移行する機能である。

(2) LUNエクспанション

論理ボリュームの容量を空き領域と連結し動的に拡張する機能である。このとき、空き領域は物理的に連続している必要はなく、別なRAID上であっても構わない。

(3) LDE (Logical Device Expansion)

物理ボリュームの追加によりRAIDタイプを変更する(例: RAID5→RAID1+0など)。

これらの機能は、I/Oアクセスが継続している状態で運用中に使用することができる。これにより業務停止することなく必要なときに必要な容量を追加することが可能となる。またRAIDマイグレーションにより動的に装置内での負荷の最適化を図ることが可能となる。ただし、容量を追加する場合にはストレージシステムだけでなくサーバ側でも容量の増加に対応する必要があり、必ずしも自動的に容量が増加することにはならない場合もある。動的な容量変化にソフトウェアが対応していない場合には、OSなどの再起動が必要となることがある。

一方、RAIDマイグレーションやLDEで物理ディスク構成を変更する場合、ソフトウェアからは構成の変更は認識できないように実装されているためソフトウェアの機能とは無関係に実施できる。

ETERNUSディスクアレイではこれらの機能に加えてボリュームの仮想化機能を提供する。ボリュームの仮想化を行った場合、論理ボリュームの容量よりも少ない物理ディスク容量での運用が可能となる(図-1)。ボリュームの実際のデータ量が増加してもサーバから見える論理ボリューム容量は変更されないため、サーバ側の設定変更が不要となり、サーバ側の管理コストが削減できる。サーバから書き込みがあった(ボリュームに実際に意味のあるデータがある)領域だけにディスクを割り当てるため、使用されていないディスク容量分を搭載することが

不要となり、導入コストが低減できる。

以下にボリュームの仮想化の動作を説明する。

- (1) ボリュームのデータとディスクのデータのマッピング情報を管理する。
- (2) サーバから書き込みがあったときにディスクを割り当て、ボリューム定義時にはディスクの割り当てを行わない。
- (3) サーバへ見せるボリュームの容量はボリューム定義時に決定する。

ボリュームの仮想化により、初めて書き込みが生じた時点で実容量が割り当てられるため、実際に使用される容量だけを実装しておけばよい。また、事前にボリュームサイズを大き目に割り当てておけばソフトウェアから見た容量は実際に割り当てられた領域とは無関係に変更されることがないために、後からソフトの定義を変更する必要もなく管理コストを低減することが可能となる。実際に利用するだけの物理ディスクを必要に応じて動的に追加することで、実使用領域の増大を必要最小限の追加コストで実現できる。

このようにボリュームの仮想化を利用することで導入コスト、管理コストの低減を図れる一方で、自動で物理ディスクの割り当てが行われてしまうため、きめ細かな性能チューニングは困難になるというデメリットもある。ボリュームの仮想化を業務に適用する際は、メリット・デメリットを事前によく検討しておく必要がある。

サーバ仮想化技術への対応

本章では、ETERNUSディスクアレイが備えるストレージベースでのボリューム複製機能である、アドバンスト・コピー機能のサーバ仮想化への対応について述べる。

● 仮想サーバ環境におけるアドバンスト・コピーの実施方法

アドバンスト・コピーとは、サーバのCPUを使用せずにストレージ装置だけで高速にデータコピーを行うETERNUSディスクアレイのボリューム複製機能である。

オープンシステムでアドバンスト・コピーを利用する場合、ETERNUS SF AdvancedCopy Manager (以下、ETERNUS SF ACM) というソフトウェアが必要になる。コピー対象サーバが物理

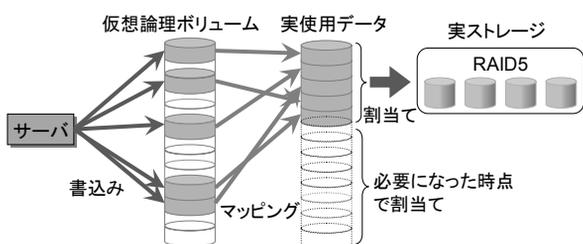


図-1 ボリューム仮想化技術
Fig.1-Virtualization of logical volume.

サーバの場合は、対象サーバにETERNUS SF ACMのエージェント機能をインストールして使用する。ETERNUS SF ACMは、ファイバチャネル経由でアドバンスト・コピーの各種操作をETERNUSディスクアレイに対して指示する(図-2)。コピー対象サーバが仮想サーバの場合、二通りのコピー方式が存在する。

【仮想サーバからのコピー】

第一のコピー方式は物理サーバの場合と同様に対象サーバにETERNUS SF ACMをインストールして使用する方式である(図-3)。本方式では、仮想化ソフトウェアが以下の条件を満足することを要求する。

- (1) コピー対象ディスクは仮想ディスクではなく物理ディスクであること。
- (2) 仮想サーバからアドバンスト・コピーの指示ができること。

現時点では、上記条件を満足するのは以下のいずれかの場合に限られる。

- ・PRIMEQUEST仮想マシン機能において、ゲストOSがLinuxの場合、かつ、仮想SCSIデバイスを使用している場合
- ・VMwareにおいてRDM (Raw Device Mapping)の物理互換モードを使用している場合

本方式のメリットは、物理サーバの場合と同様の運用が仮想サーバの場合にも実現されることである。一方、本方式のデメリットは、すでに述べたように適用対象サーバが限定されてしまうことである。

【コピー制御サーバからのコピー】

つぎに第二のコピー方式について説明する。本方式は、物理サーバ(コピー制御サーバ)^(注)を別途用意して、そのサーバからアドバンスト・コピーを行う方式である(図-4)。本方式はETERNUS SF ACM 13.4以降で利用可能である。本方式では業務サーバとは別のサーバからアドバンスト・コピーを実行するため、業務サーバのOSやアプリケーションとの連携が複雑になるものの、あらゆる仮想サーバに適用可能な方式であること、および多数の仮想サーバのコピーをコピー制御サーバに集約できるメ

(注) コピー制御サーバにはコピー元、コピー先ディスクは接続されておらず、アドバンスト・コピー制御用ディスク(アクセスバス用ディスク)が1本接続されているのみである。

リットがあり、運用環境や目的に応じてこの2方式を使い分けることが望ましい。

● 仮想サーバの災害対策ソリューション

本節では仮想サーバの災害対策ソリューションとして、VMware vCenter Site Recovery Manager(以下、Site Recovery Manager)について説明する。Site Recovery Managerは災害復旧プランの設定、フェイルオーバー、およびテストを自動化する災害復旧用ワークフロー製品である。Site Recovery Managerにより以下のメリットがもたらされる。

- ・自動化による仮想環境のリカバリの迅速化
- ・ダウンタイムなしのテストが可能のため、確実に信頼性の高いリカバリを実現
- ・複雑だったマニュアル作業でのリカバリ手順を排除し、リカバリプランの管理を統合することで、リカバリ作業を簡素化

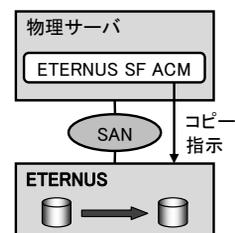


図-2 物理サーバでのアドバンスト・コピー
Fig.2-Advanced copy from physical server.

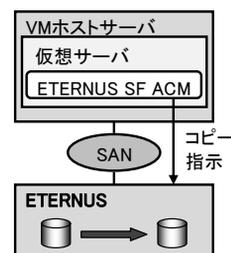


図-3 仮想サーバからのコピー指示
Fig.3-Advanced copy from virtual server.

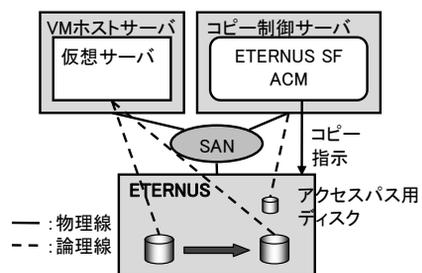


図-4 コピー制御サーバからのコピー指示
Fig.4-Advanced copy from copy controlling server.

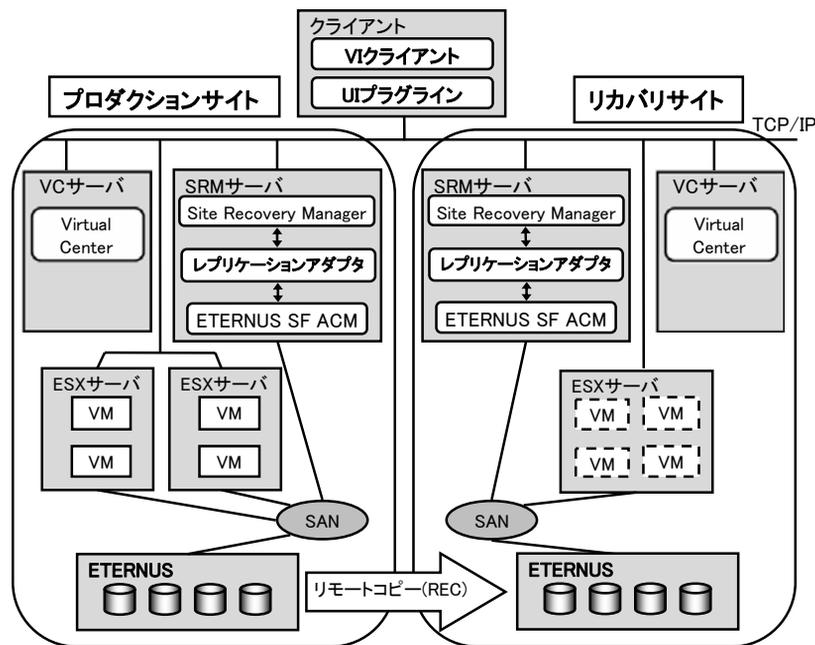


図-5 VMware vCenter Site Recovery Managerの構成
Fig.5-Configuration of VMware vCenter Site Recovery Manager.

Site Recovery Managerはストレージ装置の複製機能を利用して、仮想マシンの複製を行う。Site Recovery Managerとストレージ装置の連携は、ストレージレプリケーションアダプタと呼ばれるプログラムによって実現される(図-5)。

ストレージレプリケーションアダプタはストレージ装置固有のプログラムであり、ストレージベンダから提供されるものである。ETERNUSディスクアレイの場合、ETERNUS4000、ETERNUS8000に対応したストレージレプリケーションアダプタはETERNUS SF ACMから提供されている。本ストレージレプリケーションアダプタでは、ETERNUSディスクアレイにおけるボリュームのリモートレプリケーション機能であるREC(Remote Equivalent Copy)により、ディスクアレイ装置間のデータコピーを実施する。また、RECの転送方式としては、災害対策ソリューションに最適な非同期Consistencyモードを採用した。非同期Consistencyモードの特徴を以下に示す。

- ・データコピーをコピー元へのWriteと非同期に行

うことにより、コピー元へのWrite性能に影響を与えることなくコピーを行うことができる。

- ・転送順序性が保証される。そのため、被災が発生した場合、コピー先データがコピー元データの過去のある一時点のイメージになっていることが保証される。

む す び

本稿ではETERNUSディスクアレイにおける仮想化への取組みとして装置内部での仮想化技術の利用として、RAIDマイグレーション、LUNエクスパンション、LDE、ボリュームの仮想化機能を紹介した。またアドバンスド・コピー機能のサーバ仮想化環境対応として、サーバ仮想化環境でのアドバンスド・コピー実施方法および仮想サーバの災害対策ソリューションについて紹介した。これらによりETERNUSディスクアレイが、今後ますます利用が広がると推測される仮想化技術に対しても最適なプラットフォームとなることの一端を示せたものと思う。