

## ストレージ仮想化による仮想サーバ移行方式の検討 Study of Virtual Server Migration by Storage Virtualization

寺山 充実†

永見 明久†

兼田 泰典†

Atsumi Terayama

Akihisa Nagami

Yasunori Kaneda

### 1. はじめに

仮想サーバ台数の増加に伴い、運用管理コストの増大が課題となっている。そのため、統合管理ソフトウェアによる高効率な大規模サーバ仮想化環境へ、既設の仮想サーバを移行し、集約する動きがある。集約には、仮想サーバが使用する仮想ディスクを新たなストレージ装置へ転送する作業が伴う。

しかし、従来の移行方式では、仮想サーバ移行と同時に仮想ディスクの転送が生じ、全ての転送が完了するまでの長期間、移行元の装置を管理し続ける必要があった(図1)。そこで、本研究では、外部接続ストレージ仮想化機能により移行元ストレージ装置を移行先装置へ統合し、仮想ディスクを転送する前に移行先管理ソフトウェアの下で仮想サーバを再稼働させる方式を提案する。さらに、ストレージ仮想化により失われる仮想ディスクの接続構成を維持する構成管理技術によりこの方式を実現し、移行元・移行先両方の装置を管理しなければならない時間を短縮する。

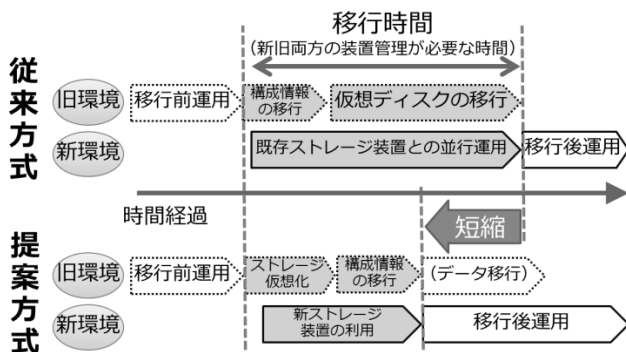


図1. 移行時間の概念

### 2. 従来方式による仮想サーバの移行

#### 2.1. 仮想サーバの移行

サーバ仮想化は、ハイパバイザの機能により、物理サーバのリソースを一つ以上の仮想サーバ(ゲストOS)に割り当てる技術である。仮想サーバに割り当てられるリソース量や設定などの構成情報は、物理サーバ上のハイパバイザにより管理される。一方、ゲストOSが使用するデータは仮想ディスクに保存され、ファイル形式でストレージ装置に格納される。物理サーバやストレージなどの各装置は、装置管理ソフトウェアの管理対象に登録され、仮想サーバの構成情報や仮想ディスクが管理される。

近年、仮想サーバ台数の増加傾向を受け、大規模仮想サーバ環境を管理する技術が注目されている。その実装の一

つとして、装置管理ソフトウェアの上位にあって、データセンタ全体を統合的に管理し、運用管理コストを低減する統合管理ソフトウェアが提供されている。

そのような統合管理ソフトウェアの管理下へ仮想サーバを移行するには、移行対象となる仮想サーバの構成情報および仮想ディスクをそれぞれ移行先環境へ移行し、さらに、移行元で構築されていたそれら仮想サーバと仮想ディスクの接続構成を維持しなければならない(図2)。

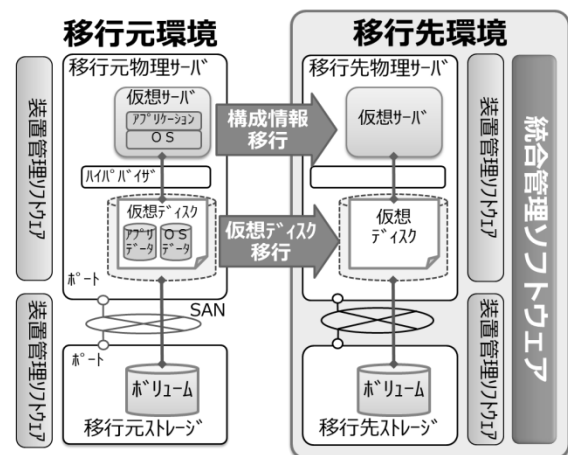


図2. 仮想サーバの移行

移行時間の観点では、実運用を想定すると、システムの移行作業は長くとも1日で完了することが望ましい。これを実現することで、例えば、土曜日に移行作業を実施、日曜日は失敗時の切り戻しのための予備日とし、営業日の業務に差し支えないように作業を行える。

#### 2.2. ホストベース移行方式

従来の仮想サーバ移行方式として、ホストベース移行方式がある。ホストベース移行方式では、移行元と移行先ハイパバイザが連携し、仮想サーバの構成情報と仮想ディスクをそれぞれ移行する。ハイパバイザを経由することで、仮想サーバの状態や設定、仮想ディスクとの接続関係などの構成情報を維持し、無停止で移行することができる。

ただし、この方式では、対象の仮想サーバを移行先管理ソフトウェアの管理対象に組み込むために、構成情報と仮想ディスクの移行完了を待つ必要がある。仮想ディスクの転送には、そのデータ容量に比例して長い時間が必要となる。その間、管理者は移行元および移行先環境の両方を管理し続けなくてはならなかった。

### 3. 移行方式の検討

#### 3.1. ストレージ仮想化を用いた仮想サーバ移行方式

外部接続ストレージ仮想化は機種異なる複数のディスクアレイを、一つのディスクアレイであるかのように統合し、管理を一元化する技術である。これを移行に適用すると、移行元ストレージを移行先ストレージの一部であるかのように仮想化できる。より具体的には、ストレージ仮想化は、ホストから受けたアクセスをストレージ間の経路にリダイレクトすることで実現されている。これにより、データ転送を行わずとも、移行先ホストから移行元ストレージの読み書き操作が可能となる。

ストレージ仮想化の設定は、データ容量によらずボリューム数に比例し、かつデータ転送と比べてごく短時間で完了する。本研究では、このストレージ仮想化機能に着目し、データ転送を行うことなく対象の仮想ディスクを移行先ストレージ装置の一部へ仮想化し、移行先管理ソフトウェアの管理対象へ組み込む移行方式(図3)を検討した。

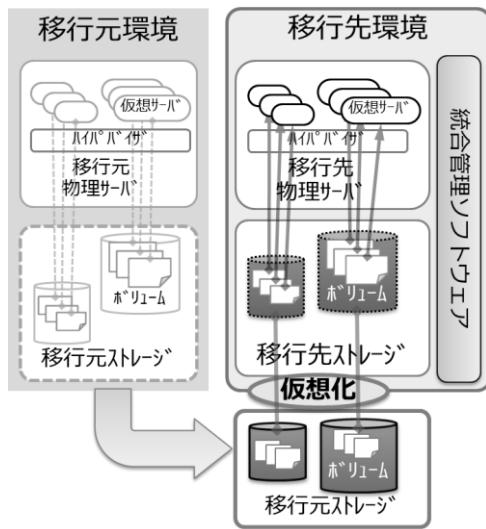


図3. ストレージ仮想化による移行方式

#### 3.2. 実現にあたっての課題とその解決

ストレージ仮想化機能は、ストレージ管理ソフトウェアにより設定される。ストレージ管理ソフトウェアは、ハイパバイザが保持する情報を持たないため、仮想サーバと仮想ディスクとの接続構成や、仮想サーバの稼働状況を知りえない。さらに、ストレージ仮想化はボリュームという仮想ディスクとは異なる単位ごとに設定されるため、単に仮想サーバの構成情報を補間するだけでは移行できない。仮に、それらの構成を手作業で確認する場合、非常に長い時間を要し、ストレージ仮想化を用いたことによる時間短縮効果を活かせない。

そこで、ハイパバイザ側のSANポート情報と、ストレージ側のSANポート情報を照合することで、移行対象の仮想ディスクが格納されているボリュームを特定する接続構成管理方式を開発した。これにより、仮想サーバの移行指定と連動したストレージ仮想化機能の自動適用を可能とした。

### 4. 評価

ホストベース移行方式による移行時間 $T_H$ は、次式で見積もることができる。

$$T_H = \frac{N_S}{C_S} \cdot \frac{p_R}{v_R} + \frac{N_D}{C_D} \cdot \frac{p_D}{v_D} + N_H t_H + \frac{2N_V}{C_P} t_V + N_S m_H \quad (1)$$

上式は、第一項から順に仮想サーバの構成情報(メモリを含む)の転送時間、仮想ディスクの転送時間、ボリューム共有構成にかかる時間、ボリュームのパス設定時間、および余剰項を示す。各記号が表す意味は次表1に示した。

一方、ストレージ仮想化を用いた移行方式による移行時間 $T_S$ は、次式で見積もることができる。

$$T_S = \frac{2N_V}{C_P} t_V + N_S m_S \quad (2)$$

上式では、第一項がボリュームのパス設定時間、第二項が余剰項を示す。

表1. 記号の定義

記号	説明
$N_S, N_D$	仮想サーバ数, 仮想ディスク数
$N_H, N_V$	物理サーバ数, ボリューム数
$C_S, C_D$	仮想サーバ, 仮想ディスクの移行多重度
$C_P$	アクセスパス設定の同時実行可能数
$p_R, v_R$	平均メモリ転送量および転送速度
$p_D, v_D$	平均仮想ディスク容量および転送速度
$t_V$	ボリュームあたり共有構成の所要時間
$t_H$	アクセスパスあたり設定の所要時間
$m_H, m_S$	余剰係数

各移行方式について小規模評価環境にて実装したプロトタイプを開発し、式(1),(2)中の各転送速度、多重度、および所要時間の計測を行った。さらに、大規模環境(仮想サーバ3200台、物理サーバ400台、移行対象ボリューム容量計800TB)を想定し、移行時間を算出した結果を次表2に示す。これにより、大規模環境において1日以内での移行を達成する見通しを得た。

表2. 評価結果

評価項目	提案方式 ([]内は手動構成の場合)	従来方式
移行時間	3.05時間 [108.3時間]	36.2時間

### 5. まとめ

本研究では、ストレージ仮想化機能を活用した仮想サーバ移行方式を検討した。さらに、同方式の実現に不可欠な仮想サーバと仮想ディスクの接続を維持する構成管理技術を開発した。これにより、移行時間を短縮した。

#### 参考文献

- [1] Gartner, "Vertically Integrated Computing Systems Will Change Storage Purchasing Choices and Decisions"
- [2] ACM JCDL '10, Luan, Nygard, Mestl, "A Mathematical Framework for Modeling and Analyzing Migration Time", 2010

† (株) 日立製作所 横浜研究所