

**BladeSymphony**

バ タ ー ジ ュ

# サーバ仮想化機構「**Virtage**」

## 開発コンセプト

2008年9月発行

株式会社 日立製作所

## 目次

1. はじめに.....	3
2. 開発方針.....	3
3. Virtage の特長 .....	3
3.1 ハードウェア透過性 .....	3
3.2 I/O仮想化支援機構 .....	5
3.3 オーバヘッド削減のための制御方式.....	5
3.4 運用管理 .....	5
4. まとめ.....	6

## 1. はじめに

サーバハードウェア技術の進展に伴ない1台のサーバに搭載されるCPU数やメモリ容量が大きくなり、一つのOSでは高効率にサーバを利用することが難しくなっています。サーバ仮想化技術は1台のサーバを複数のOSで利用可能とすることによりこの課題を解決し、柔軟で効率性の良いシステム構成を実現する技術として必然的に要求されています。本解説では日立のブレードサーバ BladeSymphony BS1000 用に開発したサーバ仮想化機構 Virtage(バタージュ)の特長について紹介します。

## 2. 開発方針

企業などの基幹系システムに用いられることが多いブレードサーバにおいては、サーバ仮想化環境でも信頼性や運用効率の観点から物理サーバと同様の高可用・高運用性システムを構築できることが望まれています。

サーバ仮想化の実装方式には、ゲストOSに対して物理インタフェースを仮想化するハードウェア隠蔽方式と、物理インタフェースを保つ論理分割方式があります。ハードウェア隠蔽方式の実装では物理資源の制約に捉われずに論理サーバを定義・移動できる利点がありますが、ソフトウェアによる介入が多く必要になります。論理分割方式にはホットスタンバイシステムのように複雑な制御が必要なシステムを容易に組める利点がありますが、ハードウェア支援機構が必要になります。このようにいずれの方式を採用するかには技術開発上のトレードオフがあります。

1970年代、メインフレームコンピュータの分野では既に仮想計算機の高性能化に関する研究が行われてきましたが[1][2]、基幹系システムに用いられることが多いこの分野では現在でも主として論理分割方式の製品が使われています。

BladeSymphony は基幹系システム向けサーバ製品であり、これに適した仮想化機構が必要である点から、Virtage では論理分割方式を採用しています。

## 3. Virtage の特長

BladeSymphony は IPF プロセッサと x86 プロセッサをサポートしており Virtage も両プロセッサをサポートしています。以下、基幹系システム向けサーバ仮想化技術としての特長を解説します。

### 3.1 ハードウェア透過性

Virtage の特長は、論理サーバ上でのソフトウェアからI/O機器へのアクセスを物理サーバで動作する場合と同一のインタフェースで可能とするハードウェア透過性にあります。これにより以下のようなシステム構成が可能となります。

#### (1) クラスタシステム

現用系／待機系サーバによりホットスタンバイシステムを構築する場合、これを制御するクラスタ制御ソフトウェアがサーバ間の排他制御用共有ディスクを使用する場合があります。これはストレージハードウェアによる排他制御機能を利用するために必要となります。クラスタ制御ソフトウェアは

共有ディスクに対してリザーブやリセットなど制御系コマンドを一定の手順で発行しますが、I/Oインタフェースを仮想化するとコマンドインタフェースの変換が必要となり実装上クラスタソフトウェアが期待する動作に一致した変換が困難となります。Virtage では、ハードウェア透過性によりクラスタ制御ソフトウェアのコマンドが正常に実行可能となります(図1)。

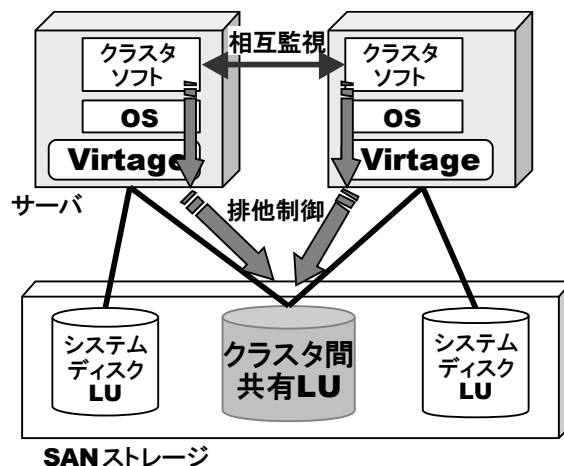


図1 I/O が透過なのでクラスタ構成が容易

## (2) 資源管理

仮想化の利点の一つはサーバを資源と見なしたときの柔軟な資源割当てによる運用性にあり、例えば負荷の軽いOSをある時間帯だけ特定の物理サーバ上に論理サーバとして集約し省電力運用を実施するといった使い方があります。ハードウェア透過性のない仮想化ソフトウェア上に集約する場合には、集約対象のサーバは同一の仮想化ソフトウェア上で動作する論理サーバに限定されます。

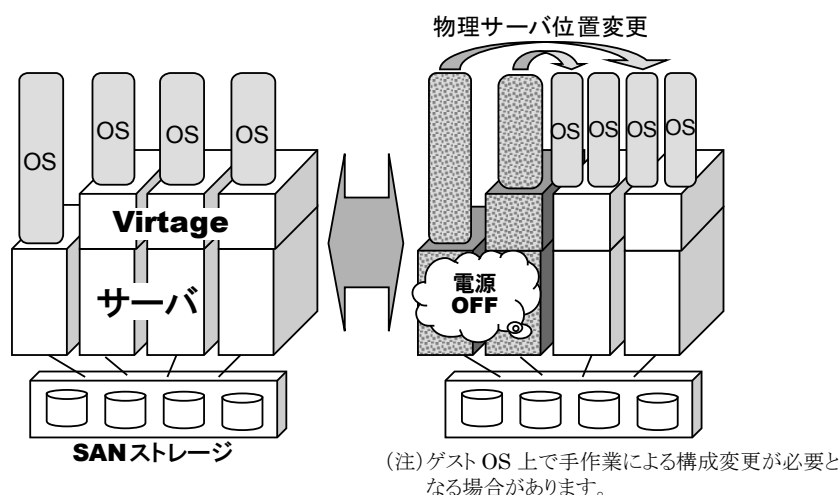


図2 物理/論理間移行による柔軟性向上

ハードウェア透過性がある場合には、物理サーバで動作しているOSを仮想化ソフトウェア上に集約することが可能となり資源管理としての柔軟性はさらに向上します(図2)。(移行の際には一旦ゲスト OS を停止する必要があります。物理サーバから論理サーバへの移行を実施する場合にはI/O 構成設定の変更などの操作が必要になる場合があります。)

### 3.2 I/O仮想化支援機構

ハードウェア透過性を実現するためにはゲストOSに対してI/Oインターフェースが物理サーバと等価に見えるI/Oパススルー方式の採用が必須となります。その実現のためにはハードウェア支援が必要であり、Virtage ではDMAリマッピング動作などを実行する独自のI/O仮想化支援機構を開発しました。

### 3.3 オーバヘッド削減のための制御方式

サーバ仮想化における大きな課題は、論理サーバと物理サーバを性能比較した場合の性能オーバヘッドをどれだけ小さくできるかという点にあります。オーバヘッド要因には主として、①命令エミュレーションオーバヘッド、②メモリアドレス変換オーバヘッド、③I/Oエミュレーションオーバヘッドなどがあります。

このうち①に関しては近年プロセッサに実装されてきたハードウェア支援機構[3]と連携、③に関しては前述の独自のI/O支援機構開発によりオーバヘッド低減を図っています。②に関しては効率的な制御方式の実装が必要となり、Virtage(IPF)では TLB (Translation Lookaside Buffer)制御、Virtage(x86)ではページテーブル制御において、論理サーバとして実用的な性能を実現する方式を開発しました。

### 3.4 運用管理

論理サーバでは電源管理、CPU やメモリなど資源の割当て、障害管理などが物理サーバと異なるため従来のサーバ管理ソフトウェアからは十分な管理ができず、別途、仮想化ソフトウェアベンダが提供するサーバ管理ソフトウェアが必要になるという課題がありました(図3(a))。

Virtage ではサーバ管理ソフトウェア(JP1/ServerConductor)上に論理サーバ管理機能を実装し物理サーバ管理機能と同等に監視・操作可能として、運用管理性の向上を図っています(図3(b))。単に統合管理可能であるだけでなく、N+1コールドスタンバイという高可用構成やシステムイメージのデプロイ機能なども物理サーバと同様にサポートしています。

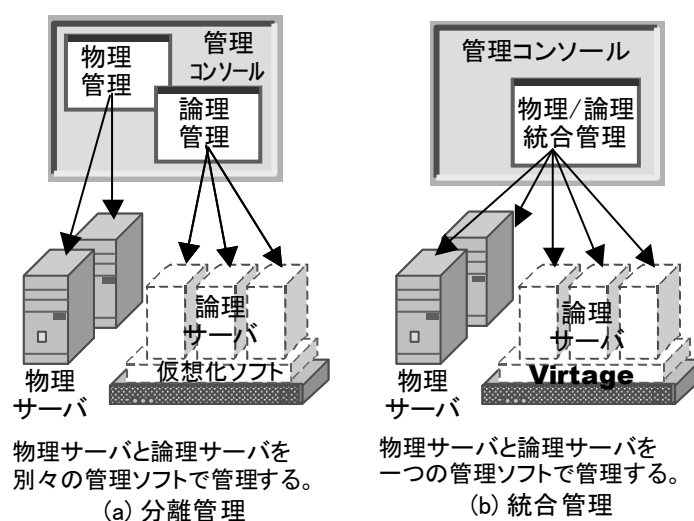


図3 物理/論理の統合管理による運用性向上

#### 4. まとめ

ハードウェア透過性を実現することにより基幹系システム構築に適したサーバ仮想化機構 Virtage を開発しました。

他社仮想化ソフトウェアではハードウェア支援を前提にできないので、ハードウェア隠蔽方式によるサーバ仮想化を採用しています。これに対して Virtage は I/O 仮想化支援機構を実装することによりハードウェア透過性を実現したという他社仮想化ソフトウェアにはない特長を持っており、基幹システムにも用いることができるシステム信頼性を実現しました。

## 参考文献

- [1]田口ほか:仮想計算機システムの制御効率を向上するための方式と実験結果,情報処理学会論文誌,Vol.20,No.4,pp.281-289(1979)
- [2]梅野英典:VM(Virtual Machine:仮想マシン)の発展,情報処理学会誌, Vol.48,No.9,pp.1026-1032(2007)
- [3]岩本成文:仮想マシンを支えるハードウェア技術(インテル),情報処理学会誌, Vol.48, No.12,pp.1421-1431(2007)

Microsoft, Windows, Windows NT , Windows Server, Windowsロゴは, 米国Microsoft Corporationの米国およびその他の国における商標または登録商標です。

Intel, インテル, Intel ロゴ, Intel Inside, Intel Inside ロゴ, Xeon, Xeon Inside, Itanium, Itanium Inside は, アメリカ合衆国およびその他の国における Intel Corporation の商標です。

その他記載の会社名・製品名は, それぞれの会社の商標もしくは登録商標です。

本書の記載事項および内容は予告なしに変更される場合があります。

Copyright (c) 2008 Hitachi, Ltd. All rights reserved

VWP-001	2008.09
---------	---------