

基幹系システム向け仮想化技術「Virtage」の開発(その1)

開発方針と特長

上野仁^{†1} 新井利明^{†2} 對馬雄次^{†3} 井形博之^{†4}

(株)日立製作所 エンタープライズサーバ事業部^{†1} 同システム開発研究所^{†2} 同中央研究所^{†3} 同ソフトウェア事業部^{†4}

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る。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る。本論文では日立のブレードサーバ BladeSymphony 用に開発したサーバ仮想化機構 Virtage(バタージュ)の特長について報告する。

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しい。

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る。ハードウェア隠蔽方式の実装では物理資源の制約に捕われずに論理サーバを定義・移動できる利点があるがソフトウェアによる介入が多く必要であり、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が組める利点があるがハードウェア支援機構が必要であるなど、方式実現にはトレードオフがある。

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る。

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た。

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する。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する。

3.1 機器透過性

Virtage の特長は、論理サーバ上での I/O 機器へのアクセスを物理サーバと同一インタフェースで可能とする機器透過性である。これにより以下のようなシステム構成が可能となる。

(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必要とする場合がある。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である。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するが、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が難しい。機器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る(図1)。

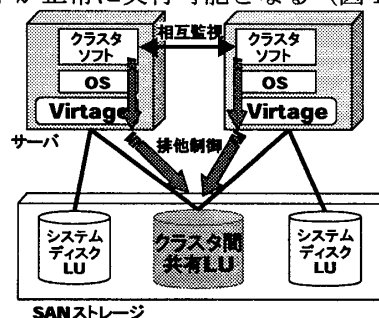


図1 I/Oが透過なのでクラスタ構成が可能

(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

Development of Server Virtualization Feature "Virtage" (1), Concepts

†1 Hitoshi Ueno, Enterprise Server Division, Hitachi, Ltd.

†2 Toshiaki Arai, Systems Development Laboratory, Hitachi, Ltd.

†3 Yuji Tsushima, Central Research Laboratory, Hitachi, Ltd.

†4 Hiroyuki Igata, Software Division, Hitachi, Ltd.

運用を実施するといった使い方があり。機器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る。

機器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する(図2)。

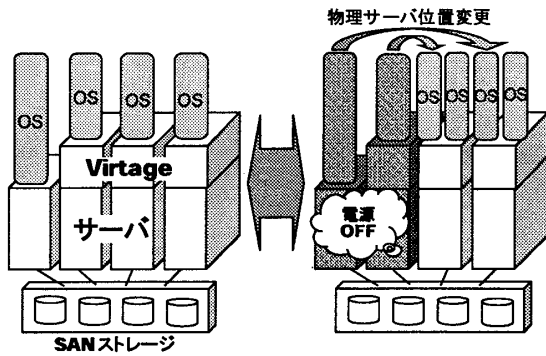


図2 物理/論理間移行による柔軟性向上

3.2 仮想化アシスト機構

機器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る。その実現のためにはハードウェア支援が必要であり、Virtageでは独自の仮想化アシスト機構でありDMAマッピング動作など実行するI/O支援機構を開発した。

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ーヘッドである。オーバー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ーヘッド、②メモリアドレス変換オーバーヘッド、③I/Oエミュレーションオーバーヘッドなどがある。

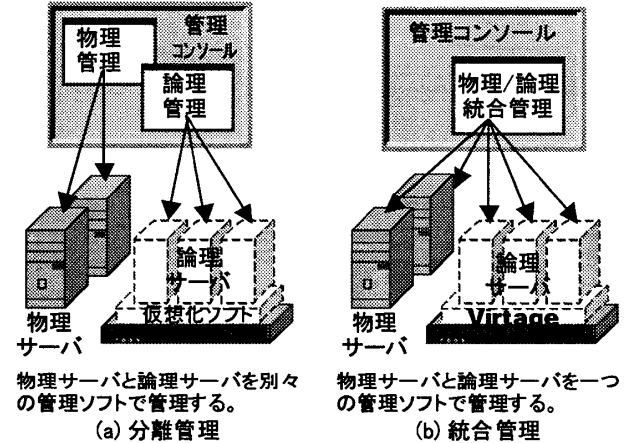
このうち①に関しては近年プロセッサに実装されてきたハードウェア支援機構[3]と連携、③に関しては前述の独自のI/O支援機構開発によりオーバーヘッド低減を図った。②に関しては効率的な制御方式の実装が必要となり、Virtage(IPF)ではTLB(Translation Lookaside Buffer)制御、Virtage(x86)ではページテーブル制御において、論理サーバとして実用的な性能を実現する方式を開発した。

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る(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る(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る。



物理サーバと論理サーバを別々の管理ソフトで管理する。

(a) 分離管理

物理サーバと論理サーバを一つの管理ソフトで管理する。

(b) 統合管理

図3 物理/論理の統合管理による運用性向上

4. おわりに

機器透過性を実現することにより基幹系システムに適したサーバ仮想化機構 Virtage を開発した。これによりハードウェア隠蔽方式の仮想化ソフトウェアとは異なる機能を実現できた。

今後はさらなる性能改善、リソースプール管理のためのプラットフォームとしての機能拡張を開発していく。

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)

IPF: Itanium Processor Family

Itanium Processor Familyは、アメリカ合衆国およびその他の国におけるIntel Corporationまたはその子会社の商標または登録商標です。