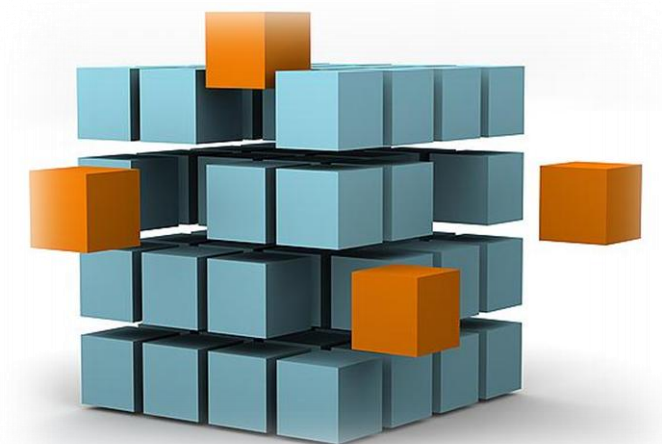


ハイエンド仮想化による新しいパラダイム versatile SMP (vSMP) アーキテクチャ概要



スケーラブルシステムズ株式会社

ハイエンド仮想化による新しいパラダイム

仮想化(バーチャリゼーション)とは、プロセッサやメモリ、ディスク、ネットワークなど、コンピュータシステムを構成するリソース(および、それらの組み合わせ)を、物理的構成に拠らず柔軟に分割・統合することを指します。一般的な仮想化としては、次のようなものが挙げられます。

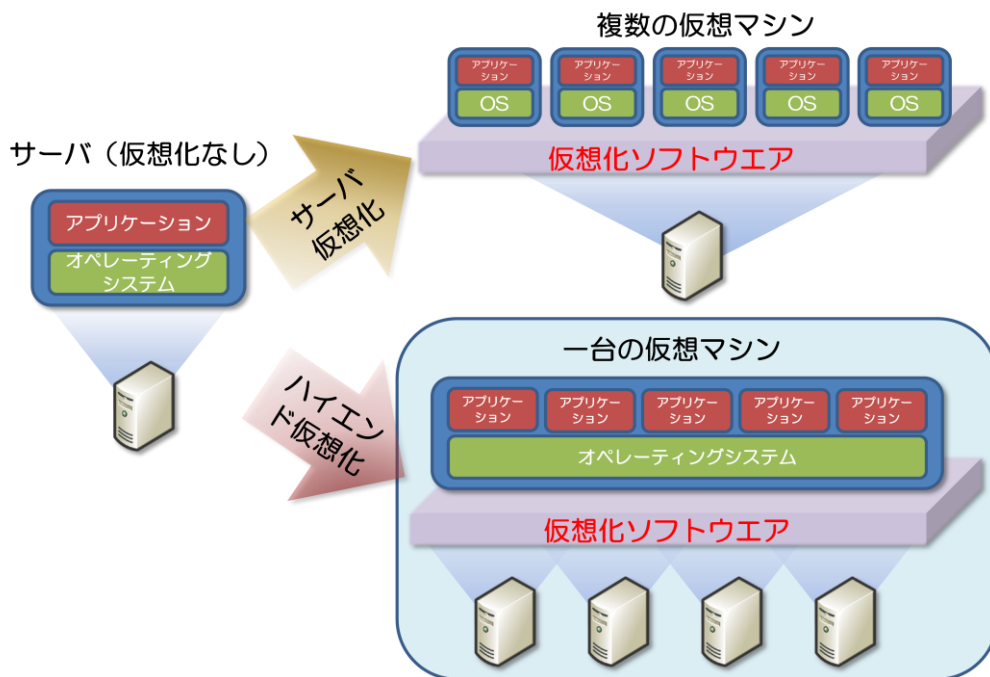
- **サーバ仮想化**: 1 台のサーバコンピュータを、あたかも複数台のコンピュータであるかのように論理的に分割し、それぞれに別の OS やアプリケーションソフトを動作させることができます。
- **ストレージ仮想化**: 複数のディスクをあたかも 1 台のディスクであるかのように扱い、大容量のデータを一括して保存したり耐障害性を高めたりする技術です。
- **デスクトップ仮想化**: 物理的な PC の設置場所に関係なく、リモートからのアクセスと操作を可能とする技術です。

スケーラブルシステムズでは、このような仮想化の新しいパラダイムとして、複数の物理的なサーバをひとつの論理サーバとして利用可能とするサーバ仮想化技術を「ハイエンド仮想化」と定義しています。

- **ハイエンド仮想化**: ScaleMP 社が開発したハイエンド仮想化技術である Versatile SMP (vSMP) アーキテクチャは、1 台のサーバを複数のサーバに分割するこれまでのサーバ仮想化とは反対に、複数システムを 1 台のシステムに統合することを実現します。

この革新的な vSMP アーキテクチャでは、複数の x86 システムを 1 台の x86 システムに統合し、ハイエンド SMP システムとして利用可能となります。このハイエンド SMP システムは、ソフトウェアによって特別なハードウェアコンポーネントを不要とするスケーラブルな SMP システムを実現し、コンピューティングに関するパラダイムの変革を可能とします。

ハイパーバイザーによる仮想化は、今日のクラウドコンピューティングにおいて重要な技術となっています。マイクロプロセッサの性能とその価格性能比の向上、ネットワーク帯域の拡大、そして仮想化技術の進化に伴って、「コンソリデーション」と「アグリゲーション」されたリモートシステム上で、ワークロードとアプリケーションの処理が可能となります。



サーバ仮想化とハイエンド仮想化

Versatile SMP (vSMP) アーキテクチャ

Versatile SMP (vSMP) アーキテクチャは、ScaleMP 社が開発したハイエンド仮想化を実現するアーキテクチャです。vSMP アーキテクチャでは、複数の x86 ベースのシステムを一つの x86 システムに統合し、x86 ベースの SMP (Symmetric Multiprocessor) として利用可能となります。

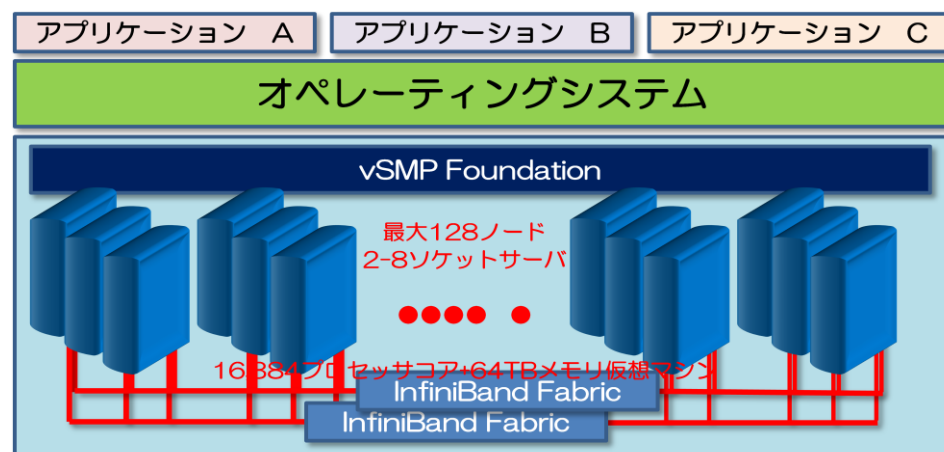
従来の SMP アーキテクチャは、独自プロセッサとチップセット、インターコネクトなどを必要としたのに対して、Versatile SMP アーキテクチャはクラスタシステムで一般的なシステム構築のコンポーネント上に実装することが可能です。ScaleMP 社の Versatile SMP (vSMP) アーキテクチャの実装が、vSMP Foundation です。vSMP Foundation によって、複数の x86 システム上に一つの大規模な SMP システムの構築が可能となります。

vSMP Foundation : vSMP アーキテクチャ実装

技術革新のスピードが加速している現在、コンピュータシステムの各コンポーネントは 18 カ月程度で世代交代がなされています。従来の SMP システム開発手法では、このペースに合わせて専用プロセッサとチップセットインターコネクトを同時に開発するのは非常に困難となっています。vSMP Foundation は、この限界を打ち破り、ハイエンドシステムとして常に最新技術が利用可能な SMP システムの提供を可能にします。

vSMP Foundation は、標準的な x86 システムをソフトウェアによって統合し、シングルシステムによるハイエンド仮想化を実現します。このシステムは、最大 128 ノード (16,384 プロセッサ・コア、65,536GB 共有メモリ) までの比類ない拡張性を備えています。

vSMP Foundation は、一般的なブレードサーバでも、ラックマウント型のサーバシステムでも利用可能です。インターコネクトとしては、現在クラスタシステムのインターコネクトとして広く利用されている InfiniBand を利用します。従って、システムはクラスタシステムを構築する場合と同様のコストでの構築が可能です。従来の共有メモリ・システムは、専用プロセッサと独自のハードウェア技術 (チップセットとインターコネクト) を必要とするため、その開発期間とコストが大きな問題となっていました。vSMP Foundation は、多くのプロセッサ・コアと大きなメモリ空間を持つシステムを、クラスタと同様のハードウェアで構成し



提供するため、従来の共有メモリ・システムの抱えていた問題をすべて解決することができます。

また、クラスタシステムの構築と運用管理を複雑にしていた複数のオペレーティングシステムの問題も解決し、プロセッサ数とメモリサイズに関係なくシングル OS でのシステム利用が可能となります。

vSMP Foundation アーキテクチャ実装の特徴

- 一般の x86 チップセットと標準インターコネクトを活用
- 専用ハードウェアや特別なコンポーネントを必要としない
- OS は通常の Linux ディストリビューションが利用可能
- 一般の x86 システムと 100% のバイナリ互換
- 特別なミドルウェアやソフトウェア無しでシステムの利用が可能
- システムの全リソース (プロセッサ、メモリ、IO) へのアクセスが可能

vSMP Foundation に必要なコンポーネント

vSMP Foundation では、特別なハードウェアで必要とせずに広く一般的に利用されている汎用製品 “off-the-shelf” によって SMP システムを構築することができます。vSMP Foundation では、ソフトウェアによって最新技術を利用した SMP システムの構築が可能で、キャッシュコヒレンシ、共有 I/O、システムインターフェイス(BIOS 及び ACPI)を提供することで、オペレーティングシステムの変更も必要ありません。

- **ノード**：複数のメモリ、CPU、ハードディスクを搭載した標準的な x86 システム(ブレードサーバ、ラックマウントサーバなどの選択が可能で、また各システムのリソース構成は必ずしも同等である必要はありません)
- **インターコネクト**：InfiniBand インフラストラクチャ(HCA、スイッチ、ケーブル)
- **vSMP Foundation デバイス**：各システムを vSMP Foundation でブートするために利用するストレージデバイス

vSMP Foundation による SMP システムの構築

vSMP Foundation は、主要なコンピュータベンダーのブレードサーバやラックマウントサーバでの検証がすでに行なわれています。スケーラブルシステムズでは、これらの認証プラットフォーム上での vSMP Foundation による SMP システムの構築とサポートを行います。スケーラブルシステムズは、お客様のニーズに最適な HPC(ハイパフォーマンス・コンピューティングシステム)の導入を強力に支援していきます。クラスタシステムの機能拡張や、大規模なメモリ空間を必要とするアプリケーションのための最適なプラットフォームとしてご活用いただけます。

クラスタ構成で導入したシステムのアップグレードのオプションとして、SMP システムとしてお使いいただけるようご提案も可能です。vSMP Foundation は、標準コンポーネントを最大限に活用し、従来の RISC や IA64 プロセッサを利用した共有メモリ・システムと同等のオペレーション環境を、コストメリットの高い標準のクラスタ環境で実現します。



標準コンポーネントによるハイエンド x86 SMP システムの実現

*) vSMP Foundation for Cloud では、vSMP Foundation モジュールをネットワークから指定されたシステムにロードして利用することが可能であり、物理的なデバイスモジュールを必要としません。詳しくは、「Versatile SMP(vSMP)アーキテクチャと vSMP Foundation for Cloud」をご覧ください。

vSMP Foundation による SMP システムの特徴

最新のサーバとネットワークテクノロジーから革新的なソフトウェアによって構築された SMP システムは次のような特徴を持ちます。

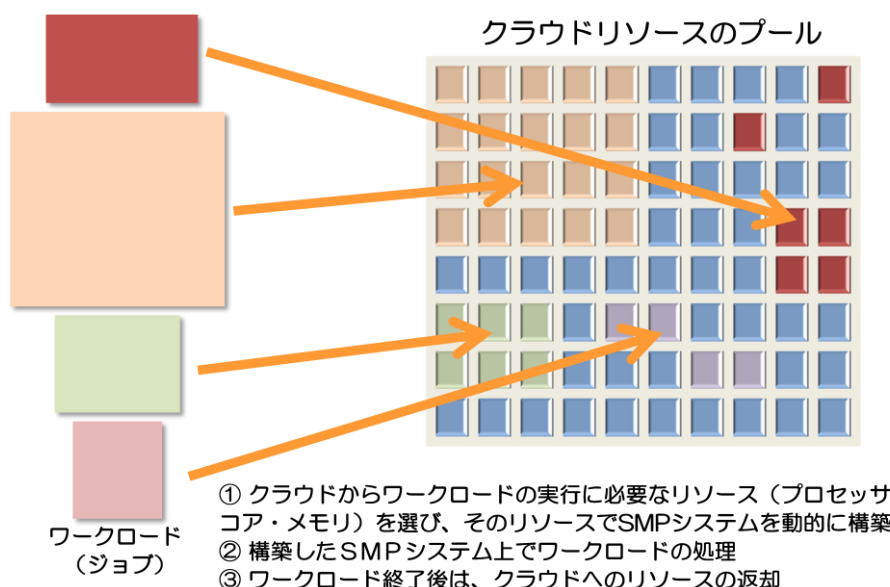
シングルシステム：各システムのメモリ上にロードされた後、vSMP Foundation は各システムの計算リソース、メモリリソース、I/O リソースの統合を行い、オペレーティングシステムとそのオペレーティングシステム上で実行されるアプリケーションに対して仮想的な一つのシステムとなります。

コヒレントメモリ：vSMP Foundation は、複数システム間でのキャッシュコヒレンシを提供します。キャッシュコヒレンシの維持に最適化されたコヒレンシアルゴリズムを採用しています。システムのメモリアクセスパターンをリアルタイムで解析し、システム内のメモリをキャッシュとして活用することでシステムインターコネクトによる通信時のレイテンシの低減を可能とします。

共有 I/O：vSMP Foundation は、システム内の全ての I/O リソースを一つの PCI 階層として構築し、オペレーティングシステムに対して一つの I/O プールを提供します。オペレーティングシステムは、すべてのシステムストレージとネットワークコントロールを利用することが可能であり、シングルシステムとしてこれらの I/O リソースへのパラレルアクセスが可能となります。

自由度の高いシステム構成 (Versatile System)：vSMP Foundation では、統合する各システムのプロセッサ速度、メモリサイズ、I/O デバイスが異なってもシステム構成が可能です。従って、システム構成の自由度が非常に高くなります。例えば、最大限の計算処理能力が求められる場合、32 プロセッサ (192 プロセッサ・コア)、各システムで均等メモリを搭載したシステム構成が可能です (例えば、1.5TFLOPS、1TB メモリのような構成)。一方、計算処理能力よりもメモリサイズを求められるアプリケーションでは、計算処理に必要なプロセッサは高速なものを搭載し、残りは低速で省電力プロセッサを搭載することでコストと消費電力を削減したシステムの構築も可能となります。また、アプリケーションが必要とする I/O 拡張オプションについても、同様に柔軟なシステム構成が可能となります。このようにシステム構成の自由度と柔軟性に優れていることから、この技術と実装を Versatile SMP と呼んでいます。

オンデマンドで利用可能な SMP システム：vSMP Foundation for Cloud は、HPC の高い処理能力をオンデマンドで提供するアーキテクチャによって、ワークロードに合わせて柔軟にシステムを構成することができます。この Versatile SMP (vSMP) アーキテクチャは、ジョブやプロジェクト、あるいは顧客毎に仮想マシンを構築、プロビジョニングすることが可能で、従来の拡張性に制限のあった専用の共有メモリ・システムに代わる、お客様のニーズに最適な選択肢となります。ワークロードの拡張に柔軟に対応可能であり、クラウドコンピューティング向けのインフラ構築が求められる場合には最適です。ジョブ、プロジェクト、ユーザ毎に柔軟に仮想マシンの構築を可能として、それぞれの用途に共有メモリ・システムを提供します。vSMP Foundation for Cloud では、クラウドインフラストラクチャに容易に適用出来る、ネットワークからブート可能なイメージを提供します。



vSMP Foundation の利点：OPEX と CAPEX の削減

vSMP Foundation では、最新システムをコンポーネントとして SMP システムの構築を行うため、技術革新のスピードに合わせてスケーラブルな SMP システムの提供が可能となります。また、コストパフォーマンスに優れた 2 ソケットのサーバを利用して、高いコストパフォーマンスを実現することも可能です。

- **シングルアドレス空間**：メモリリソースについては、アプリケーションがオペレーティングシステムを利用できる限界を大幅に拡張することが可能です。スレッドを利用したアプリケーションが、より多くのプロセッサ・コアを利用することを可能とします。
- **シングルオペレーティングシステム**：オペレーティングシステムがシングルシステムとなることで、運用・管理のための特別なツールやソフトウェアを必要としません。また、クラスタシステムで問題となるファイルシステムの問題も、vSMP Foundation が解決します。全プロセッサ・コアは、全ての I/O リソースへのアクセスが可能です。
- **シングルシステム**：vSMP Foundation では、オペレーティングシステムやアプリケーションに変更を加えることなく利用することができます。2 ソケットや 4 ソケットのサーバシステムを、より多くのソケットが利用可能なシステムに拡張したものと同等になります。vSMP Foundation のインストールと立ち上げは数十分で完了するため、非常に短時間での導入が可能です。シングルシステムイメージではシステムがよりシンプルなものになるため、システムの運用管理業務が大幅に簡素化されます。
- **シングルインフラストラクチャ**：vSMP Foundation は、一般的なブレードサーバやラックマウントサーバで構築することができます。これらの汎用サーバは、各ベンダーが競ってシステム開発を続けることで常に改良されており、最新システムはより優れたエネルギー効率、より高い性能、より高度なシステム実装が可能となるため、vSMP Foundation はその技術革新のメリットを 100%活用することが可能となります。



物理的な計算リソースの要求に合ったシステムを構築
OPEX と CAPEX の削減を目的

まとめとして：vSMP Foundation によるソリューション

ハイパフォーマンスシステムでのフレームワークにおいては、幾つかの課題があります。これらの課題に対して、vSMP Foundation は次のようなソリューションの提案が可能です。

システムの課題	vSMP Foundation でのソリューション
一般の SMP システムとクラスタシステムで利用されるアプリケーションを効率良く実行できる	一般の SMP システムに対して、様々なアプリケーションの実行において、同等以上の性能で実行出来ます。また、アプリケーションの実行モデル（マルチスレッド、スループット、メッセージパッシング、大規模メモリへのアクセス）に対して、シングルシステムの構成で、効率の良い処理が可能です。
アプリケーションの実行において、一般のクラスタや SMP に比べて遜色のない性能を示すことが可能である	vSMP Foundation は、メモリ配置の局所性を最大限に活用しています。このため、分散アプリケーションの実行を最適に行うことが可能です。同時に SMP システムとして高いメモリバンド幅を実現しているため、共有メモリ・システムとしても高い性能を発揮することが可能です。
技術革新の速度が速いマイクロプロセッサやチップセット、インターコネクトについて、常に最新技術を利用することが可能である	vSMP Foundation の最もユニークな特徴は、ソフトウェアベースの SMP システムの構築アーキテクチャであることです。利用するシステムのコンポーネントについては、常に最新技術を採用したサーバやインターコネクトを利用することが可能となります。
システムの運用管理は SMP システムと同様に簡便に低コストで行うことが可能である	vSMP Foundation を利用したシステムは、シングルシステムとして運用することが可能であり、オペレーションコストを大幅に削減することができます。
購入コストはクラスタシステムと同等であり、また、特別なハードウェアの購入も可能な限り避けることができる	vSMP Foundation では、必要とするハードウェアはクラスタシステムと同じで、一般商用サーバとインターコネクトでのシステム構築が可能です。

ハイパフォーマンスなクラウド環境の構築のためのソリューションとして導入実績のあるハイエンド仮想化技術 vSMP Foundation により、クラウドでの運用管理費用を大幅に削減すると同時に、導入を躊躇する最も大きな理由と言える面倒な利用環境の問題を解消し、クリティカルなエンタープライズ・アプリケーションの実行に最適なシステムを構築することが可能となります。

スケーラブルシステムズ株式会社

〒102-0083 東京都千代田区麹町 3-5-2 BUREX 麹町 11 階

電話：03-5875-4718 FAX：03-3237-7612 www.sstc.co.jp

©Scalable Systems, Ltd. All rights reserved.

社名、製品名などは、一般に各社の商標または登録商標です。無断での引用、転載を禁じます。