

ホワイトペーパー

FUJITSU Server PRIMERGY & PRIMEQUEST ベンチマークの概要 vServCon

従来のベンチマークは、仮想化環境で実行されるオペレーティングシステムやアプリケーションの評価に適していません。そのため、仮想化環境に特化したベンチマークが開発されました。社内利用を目的として富士通が開発した「vServCon」では、仮想化ソリューションおよび PRIMERGY と PRIMEQUEST サーバでの測定と評価が可能となります。

このドキュメントでは、仮想化環境のベンチマークに関する問題、「vServCon」ベンチマークの基礎、および富士通での使用例について説明します。

バージョン

1.6a

2015-02-25



目次

はじめに	3
vServCon	4
vServCon ベンチマーク	4
vServCon スコア	5
負荷プロファイルとハードウェア環境	7
関連資料	9
お問い合わせ先	9

はじめに

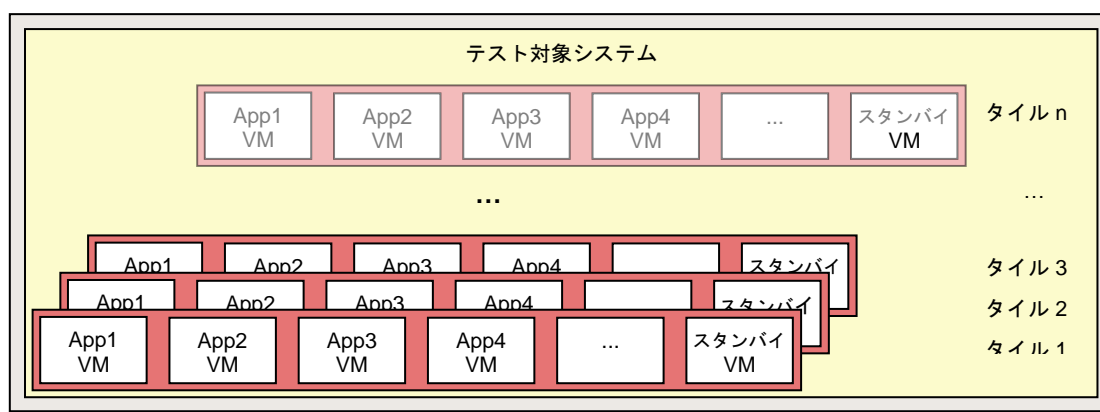
IT インフラストラクチャにおいてサーバ仮想化の流れが進行しています。仮想化により、ハードウェアの使用効率を高めると同時にハードウェアへの依存度を減らすことができます。1 つのオペレーティングシステム上でアプリケーションを使用する従来のサーバとは対照的に、仮想化環境では、複数のオペレーティングシステムとアプリケーションが並列に実行され、1 つのサーバ上に異種混合環境が構築されます。また、仮想マシンの管理はハイパーバイザーとも呼ばれる仮想化モニタによって行われます。

従来のベンチマークは、仮想化環境の性能測定や評価には適していません。仮想化環境の性能を適切に評価するには、複数の仮想マシンにさまざまな作業負荷をかけて同時に動作させ、1 台のサーバのハードウェアリソースを共有させる必要があります。仮想化ベンチマークの対象は 2 つあり、1 つはサーバ統合です。サーバ統合では、適切なレプリケーションを行うことにより、単一の仮想化ホスト上の複数の仮想マシン全体のスループットを最大化します。vConsolidate (Intel)、VMmark V1 (VMware)、および SPECvirt_sc2010 (SPEC) は、このタイプのベンチマークです。また、富士通は、「vConsolidate」の考え方に基づいて開発した内部ベンチマーク「vServCon」を使用して仮想環境の性能測定を行っています。仮想化ベンチマークのもう 1 つの対象は、データセンターの運用です。この場合は、複数の仮想化ホストのサーバ統合シナリオを想定しています。このタイプのベンチマークでは、仮想マシンのスループットに加えて、仮想マシンの再配置などの、一般的なデータセンター運用の効率性を測定します。VMmark V2 (VMware) はこのタイプのベンチマークです。

仮想化ベンチマークの目的を達成するには、サーバ統合の観点から実際のデータセンターを見直す必要があります。言い方を換えると、現在サーバ上で稼働しているアプリケーションシナリオを考慮する必要があるということです。これらのサーバは稼働率が低いので、できる限り多くのサーバを仮想マシン (VM) として集約することが目的になります。そのため、ベンチマークでは、さまざまなアプリケーション VM の全体的なスループットと、効率よく運用できる仮想マシンの個数を評価することが必要です。

これらの 2 つの目的のために次のようなソリューションコンセプトが確立されています。ベンチマークではアプリケーションシナリオの代表的な VM のグループが選択されます。これらは、測定を行うときに仮想ホスト上で同時に実行されます。各 VM には、適切な負荷ツールを使用して低い負荷がかけられるようになっています。一般的に仮想化ベンチマークでは、オペレーティングシステム、アプリケーション、および「アイドル」VM または「スタンバイ」VM (仮想化環境の稼働していないフェーズを表し、ハイパーバイザーによって管理される VM の数を増加させます) をグループ化したものを「タイル」と呼びこの単位で実行します。

検討対象のシステムがパフォーマンスの限界に到達するまで、この仮想マシンのグループ (タイル) によって生成される負荷を段階的に増やす必要があります。次の図は、複数のタイルを稼働させることによるテスト対象システム上の VM 負荷の増加を示しています。



アプリケーションは各仮想マシン上で実行され、確立されたベンチマークによって負荷がかけられます。これらのアプリケーションに加えて、インフラストラクチャーコンポーネントが実行される場合もあります。個々の測定結果は、全体結果を示す 1 つのスコアに集約されます。このスコアが仮想化環境の処理能力を示すこととなります。

vServCon

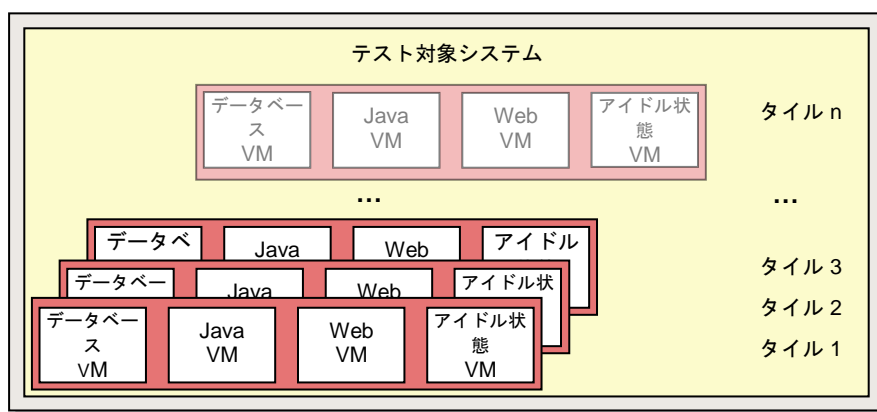
富士通では、現在のところ社内ベンチマークとして後述の「vServCon」（Intel の「vConsolidate」の考え方を基にしています）を使用して仮想化環境のスケーラビリティ測定を行っています。「vServCon」は、「Virtualization enables SERVER CONSolidation：仮想化はサーバ統合を可能にする」を意味します。

vServCon ベンチマーク

vServCon は、実際には新しいベンチマークではありません。VMmark は、すでに確立されたベンチマークをワークロードとして統合するフレームワークであり、必要に応じて、ワークロードに変更を加えることができます。データベース、アプリケーションサーバ、Web サーバというアプリケーションシナリオを対象とする 3 つの実証済みのベンチマークが使用されます。

アプリケーションシナリオ	負荷ツール	VM の数
データベース	Sysbench	1
Java アプリケーションサーバ	SPECjbb2005	1
Web サーバ	WebBench	1
（アイドル状態）	-	1

3 つのアプリケーションシナリオのそれぞれが、1 つの専用の仮想マシン（VM）に割り当てられます。さらに、「アイドル VM」という第 4 の VM がこれらに追加されます。これらの 4 つの VM が「タイル」を形成します。タイルは、vConsolidate の用語では「Consolidation Stack Unit」（CSU）に相当します。測定対象となるサーバの処理能力によっては、全体として最大のパフォーマンスを達成するために複数のタイルを並列して開始する必要があります。



vServCon には、Web サーバ VM 用に外部の負荷ジェネレーター（タイルごとに 1 つのクライアントシステム）が必要です。負荷ジェネレーターと「テスト対象システム」は、適切な数のネットワークを介して接続されます。

個々の負荷ツールの実行は、フレームワークコントローラーと呼ばれる独立したシステムによって制御されます。このシステムは、測定を監視し、VM の個々のパフォーマンスデータを収集します。

標準ベンチマークのうち Sysbench と SPECjbb の 2 つは、vServCon フレームワーク内で測定できるように、仮想化環境での使用に合わせた調整が必要でした。

vServCon スコア

vServCon の結果は「スコア」と呼ばれる数値であり、測定した仮想化ホストのパフォーマンスを表します。このスコアは、目的上、多数のアプリケーション VM を組み合わせた定義済みの構成を実行することで達成可能な最大合計スループットを反映する必要があります。

このスコアは、各 VM の個々の結果から決定されます。3 つの vServCon アプリケーションシナリオのそれぞれが、各 VM のアプリケーション固有のトランザクションレートという形で具体的なベンチマーク結果を提供します。正規化されたスコアを取得するために、1 つのタイルのそれぞれのベンチマーク結果をいつも同一のリファレンスシステム (Xeon 5130 プロセッサ × 2、16 GB RAM を搭載した PRIMERGY RX300 S3) の結果と比較した値 (比率) を求めます。その結果の相対パフォーマンス値に適切な重み付けを行います。すべての VM とすべてのタイルについて同じ手順で求めた値を足し合わせたものがこのタイル数に対する vServCon スコアになります (スコア計算の詳細については、[下記](#)を参照)。

原則として、1 つのタイルから始めて、vServCon スコアの大幅な増加が見られなくなるまで、タイル数を増やしながらかこの手順が実行されます。最終的な vServCon スコアは、すべてのタイル数から得られた vServCon スコアのうちの最大値になります。

タイル数の増加に対する vServCon スコアの伸びは、テスト対象システムのスケーリング特性についての有益な情報となります。

オプションとして、vServCon の測定中に、ホストの合計 CPU 負荷 (VM および他のすべての CPU 処理) の記録、可能な場合は電力消費量も記録することができます。

詳細情報 : スコアの計算方法

本書の以降の内容を理解する上で、必ずしも、ここに示す詳細なスコア計算について理解する必要はありません。このトピックに特に興味がなければ読み飛ばして構いません。

現在のタイル数 n のスコアは下記の式では $Score_n$ と表記されます。これは、個々の VM のスコアを求め、すべてのタイルの個々のスコアを合計して算出します。

$$Score_n = \sum_i \sum_{VM} Score_{VM_i}$$

VM タイルの 3 つのベンチマーク VM の 1 つを表します (アイドル VM を除く)。
i タイルに 1~n の番号を割り当てます。

ベンチマークの測定値をそのまま VM のスコアとして使用することはできない点に注意してください。個々のベンチマークにおける測定値は、まったく異なる側面について、異なる測定単位で求められるものだからです。このため、個々のベンチマーク結果とリファレンスシステムで単一のタイルを実行した結果との比率を求めます。つまり、後者にスコア 1 が与えられることになります。これですべてのベンチマークのスコアの次元が同一になり、測定単位の違いがなくなるので、値を合計することができます。

相対スコア

$$\frac{R_{VM_i}}{R_{VM_{ref}}}$$

R_{VM_i} 測定対象 VM のベンチマーク結果
 $R_{VM_{ref}}$ リファレンスシステムで 1 タイルだけ使用して測定した場合の VM のベンチマーク結果

VM のベンチマーク結果に、各 VM タイプに固有の重み係数 α_{VM} をかけます。VM のスコアは次のようにして求められます。

$$Score_{VM_i} = \frac{R_{VM_i}}{R_{VM_{ref}}} \alpha_{VM}$$

α_{VM} VM タイプの重み係数

VM タイプの重み係数 α_{VM} は、タイル全体の CPU およびメモリ構成と VM の CPU およびメモリ構成との比率を考慮に入れ、評価するものです。

$$\alpha_{VM} = \frac{CPU_{VM}}{CPU_{CSU}} \times 0.8 + \frac{MEM_{VM}}{MEM_{CSU}} \times 0.2$$

$$CPU_{CSU} = \sum_{VM} CPU_{VM} \quad MEM_{CSU} = \sum_{VM} MEM_{VM}$$

CPU_{VM} : 仮想マシン VM に割り当てられた仮想 CPU 数

CPU_{Tile} : 1 タイルの 仮想 CPU 数

MEM_{VM} : 仮想マシン VM に割り当てられたメモリサイズ

MEM_{Tile} : 1 タイルのメモリサイズ合計

CPU 構成に関して、仮想 CPU (vCPU) が仮想化ホストの物理 CPU と必ずしも同じにはならない点に注意してください。仮想 CPU は常に物理 CPU の 1 論理実行単位のみから構成されます。例えば、マルチコア CPU のコア 1 つを使用している場合や、ハイパースレッディング機能搭載 CPU の 1 スレッドを使用している場合などです。仮想システムの構成は、物理 CPU の内部構造を直接反映するものではありません。実行時に最も効率的な方法で動的な CPU スケジューリングが行われるようにするのは、各仮想化ハイパーバイザーの仕事です。

負荷プロファイルとハードウェア環境

原理的には、各仮想マシンの CPU およびメモリのリソース、オペレーティングシステムやアプリケーションソフトウェアを最高のスコアが出るように最適化することは可能です。ただし、この場合、比較作業が大幅に複雑化します。比較こそがスケーリング特性を語る上で最も重要であるにもかかわらずです。したがって、富士通では、標準化されたベンチマーク環境を定義し、vServCon フレームワーク内で「プロファイル」と呼ばれるものを使用します。プロファイルには、リソースと、オペレーティングシステムおよびアプリケーションソフトウェアのソフトウェアバージョンが厳密に指定されます。VMmark が提供するゲストオペレーティングシステム、アプリケーション、およびそれらが指定するリソース要件の集まりは、複雑な仮想化環境の代表的な選択肢とみなすことができます。ベンチマークに含まれるソフトウェアに、新しくよりパフォーマンスの高いバージョンが存在する場合でも、比較可能性を確保するために、負荷プロファイルの継続性をできる限り長く維持する必要があります。PRIMERGY と PRIMEQUEST「vServCon」測定では次のプロファイルが使用されます。

富士通 プロファイル v1.0	リソース	Web WebBench	データベース Sysbench	Java SPECjbb	アイドル状態
	vCPU 数	1	2	2	1
	メモリ	1.5 GB	1.5 GB	2.0 GB	0.4 GB
	OS	Linux 32 ビット SUSE SLES 10	Windows Server 2003 R2 EE SP2 64 ビット	Linux 64 ビット RedHat RHEL 4.5	Windows Server 2003 R2 EE SP2 32 ビット
	アプリケーション	Apache	Microsoft SQL Server 2005 SP2	BEA JRockit(6)	
	ベンチマーク	WebBench 5.0	Sysbench v0.4.0 (Intel バイナリ)	SPECjbb 2005	-
	ディスクサブシステム	10 GB	10 GB	5 GB	5 GB

オペレーティングシステムの組み合わせについては、ヨーロッパで普及している Linux ディストリビューションが考慮されています。

また、ディスクサブシステムも定義しました。サイズと RAID レベルに関する論理設計だけでなく、物理的な実装も定義されています。ローカルディスクサブシステムは、通常、仮想化および統合向けの実用的なソリューションではありません。それが SAN ベースのディスクサブシステムを使用している理由です。最終的な目的は仮想化ソリューションとサーバハードウェアのパフォーマンス性能を知ることなので、ディスク I/O サブシステムは、長期的にボトルネックになることを回避するように設計されています。

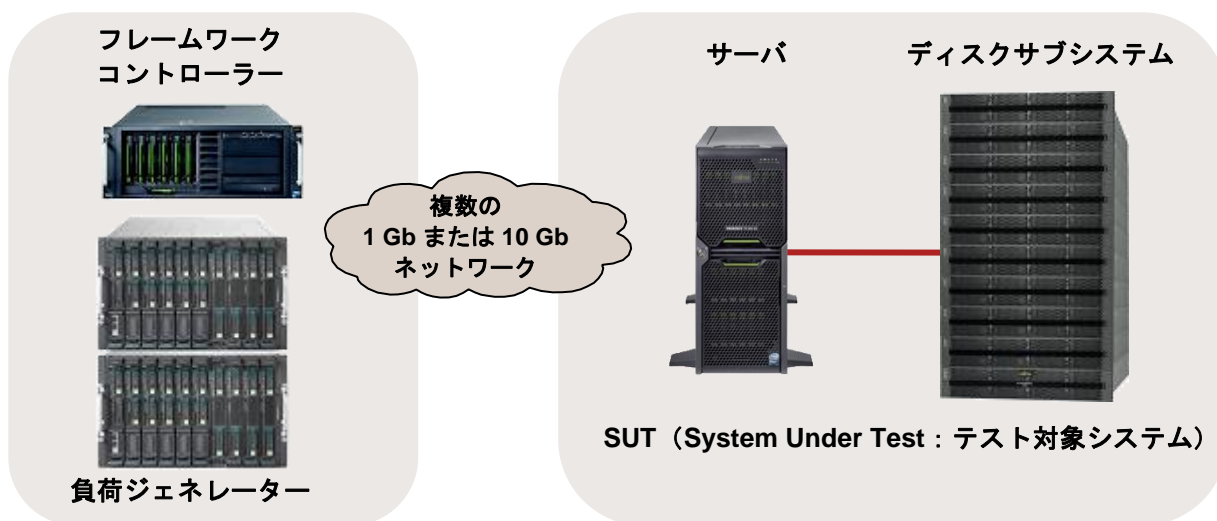
また、テスト対象システムのメインメモリも vServCon 測定のボトルネックにならないようにする必要があります。そのため、最大タイル数でもホストレベルでスワップが発生しないように適切な量のメモリが搭載されます。

ベンチマーク「WebBench」には、外部負荷ジェネレーターが必要です。負荷ジェネレーターとテスト対象システムは、十分な帯域幅を持った専用 LAN を介して接続されます。

2 本目の専用ネットワークは、vServCon ベンチマークの制御に使用します。

フレームワークコントローラーとして、追加の PRIMERGY システム が使用されます。

以下に一般的な測定環境を示します。



個々の VM の構成は、定義された仕様に従い、常に同一になります。具体的には、オペレーティングシステムも各ベンチマークソフトウェアも最適化されません。高スコアを出すための最適化は不可能ではありませんが、比較が目的であるため、あまり重要ではありません。使用されるハイパーバイザーに応じて必要な修正は常に可能で、それによって文書化されます。

VM の外側、BIOS レベルまたは仮想化ソフトウェア内で、仮想化プラットフォームのパフォーマンス最適化を緊急で行う必要がある場合は、仮想化プラットフォームのパフォーマンスでより良いスコアを出すために、設定を変更できます。ハイパーバイザーメーカーによって推奨または規定される設定も vServCon では考慮されています。適切であれば標準設定が保持されます。これは通常、パフォーマンスとエネルギー効率とがバランス良く調整された形になります。

関連資料

PRIMERGY & PRIMEQUEST サーバ

<http://jp.fujitsu.com/platform/server/>




PRIMEQUEST のパフォーマンス

<http://jp.fujitsu.com/platform/server/primequest/products/2000/benchmark/>

PRIMERGY のパフォーマンス

<http://jp.fujitsu.com/platform/server/primergy/performance/>

このホワイトペーパー

 <http://docs.ts.fujitsu.com/dl.aspx?id=1a427c16-12bf-41b0-9ca3-4cc360ef14ce>
 <http://docs.ts.fujitsu.com/dl.aspx?id=04351fd2-8a69-42a3-ba1c-4342dcc89b89>
 <http://docs.ts.fujitsu.com/dl.aspx?id=00b0bf10-8f75-435f-bb9b-3eceb5ce0157>

ベンチマーク

SPECjbb2005
<http://www.spec.org/jbb2005/>
SPECvirtualization
<http://www.spec.org/benchmarks.html#virtual>
VMmark
<http://www.vmmark.com/>

お問い合わせ先

富士通

Web サイト : <http://jp.fujitsu.com/>

PRIMERGY のパフォーマンスとベンチマーク

<mailto:primergy.benchmark@ts.fujitsu.com>