
記事

[Tomohiro Iwamoto](#) · 2020年6月5日 18m read

InterSystemsデータプラットフォームとパフォーマンス -パート3: CPUに注目

今週は、ハードウェアの主な ” 食品群 ” (^) の1つであるCPUに注目します。お客様から、次のようなシナリオについてのアドバイスを求められました。お客様の本番サーバーはサポート終了に近づいており、ハードウェアを交換する時期が来ています。また、仮想化によってサーバーを一元的に管理できるようにし、ベアメタルか仮想化によりキャパシティを適正化したいとも考えています。今日はCPUに焦点を当てますが、後日の記事では、メモリやIOといったほかの主要食品群を適正化するアプローチについて説明したいと思います。

では、質問を整理しましょう。

- 5年前のプロセッサをもとに作られたアプリケーション要件を今日のプロセッサに変換するには？
- 現在ではどのプロセッサが適しているのか？
- 仮想化はCPUキャパシティプランニングにどのような影響を及ぼすのか？

2017年6月追記:

VMware CPUに

関する考慮事項と計画の詳細、お

よび一般的な質問と問題の詳細については、「[大規模データベースの仮想化 - VMware CPUのキャパシティ計画](#)

」の記事も参照してください。

[このシリーズの他の記事のリストはこちら](#)

spec.orgのベンチマークを使ったCPUパフォーマンスの比較

InterSystemsデータプラットフォーム（Caché、Ensemble、HealthShare）を使って構築されたアプリケーションにおいて、異なる種類のプロセッサのCPU使用率を解釈する際に、信頼できる大雑把な計算としてSPECintベンチマークを使用してプロセッサを比較できます。<http://www.spec.org> のWebサイトには、ハードウェアベンダーが標準的に実施した信頼性の高いベンチマークの結果が掲載されています。

具体的に言うと、SPECintは、プロセッサを同じベンダーのほかのプロセッサモデルや異なるベンダー(Dell、HP、Lenovo、およびIntel、AMD、IBM POWER、SPARC)のプロセッサと比較する方法です。ハードウェアをアップグレードする場合に、またはアプリケーションをさまざまな顧客ハードウェアに展開して、Intel Xeon E5-2680など選択したプロセッサのCPUコア当たりのピークトランザクションといったサイジングメトリックのベースラインを設定する必要がある場合に、SPECintを使って、アプリケーションに期待されるCPU要件を理解することができます。

SPECintのWebサイトではいくつかのベンチマークが使用されていますが、Cachéでは SPECint_{ratebase2006}の結果が最適であり、長年にわたって顧客データと独自のベンチマークでそのことが確認されています。

この記事の例として、Intel Xeon 5570プロセッサを実行する顧客のDell PowerEdgeサーバーとIntel Xeon E5-2680 V3プロセッサを実行する最新のDellサーバーを比較してみましょう。Intel Xeon V4サーバープロセッサが一般に利用できるようになったときにも（本記事を執筆した2016年初期時点で近日発売予定）、同じ方法を適用できます

例: プロセッサの比較

spec.orgデータベースでプロセッサの SPECint2006Rates を検索します。E5-2680 V3 などのプロセッサ名を検索し、対象サーバーの製造会社とモデル(Dell R730など)がわかっていればその名前で、そうでない場合は一般的なベンダーを使って検索結果を絞り込みます。私は、一般的なサーバーとしてDellやHPのモデルが優れたベースラインだと思っていますが、異なるベンダーのハードウェアに使われるプロセッサには大した差はありません。

この記事の最後に、spec.orgのWebサイトを使って結果を検索する例を段階的に説明します…

spec.orgを検索して、次のような既存のサーバーと可能性のある新しいサーバーを見つけたとします。

Existing: Dell PowerEdge R710 with Xeon 5570 2.93 GHz: 8 cores, 2 chips, 4 cores/chip, 2 threads/core:

SPECintratebase2006 = 251

New: PowerEdge R730 with Intel Xeon E5-2680 v3, 2.50 GHz: 24 cores, 2 chips, 12 cores/chip, 2 threads/core:

SPECintratebase2006 = 1030

当然のことながら、新しい方の24コアサーバーでは、クロック速度が遅いにも関わらず、古い8コアサーバーのSPECintratebase2006ベンチマークスループットが4倍以上増加しています。この例は、共に複数のプロセッサソケットが搭載された2プロセッササーバーであることに注意してください。

CachéにSPECintratebase2006を使用する理由

spec.orgのWebサイトに

はさまざまなベンチマークが説明されていますが

、要約すると、**SPECintrate2006**

ベンチマークは完全なシステムレベルのベンチマークで、すべてのCPUでハイパースレッディングがサポートされています。

特定のSPECintrate2006ベンチマークでは、baseとpeakの2つのメトリックが報告されています。ベース(base)は保守的なベ

ンチマークで、ピーク(peak)

は積極的なベンチマークです。キャパシティ計画では、**SPECintratebase2006** の結果を使用してください。

SPECintratebase2006の4倍とは、ユーザーまたはトランザクションの容量が4倍ということですか？

24個のコアをすべてを使用した場合、アプリケーションのスループットは古いサーバーの能力の4倍に拡張される可能性はありますが、この幅はいくつかの要因によって異なります。SPECintを使用すれば、可能性のあるサイジングとスループットについておおよその見当をつけることはできますが、注意点がいくつかあります。

SPECintでは、上記の例にある2つのサーバーの適切な比較を提供しますが、E5-2680 V3サーバーのピーク時の同時ユーザーまたはピーク時のトランザクションスループットにおけるキャパシティが古いXeon 5570ベースサーバーと比べて75%増になるという保証はありません。“食品群”のほかのハードウェアコンポーネントがアップグレードされるかどうかなど、ほかの要因もかかわってきます。たとえば、スループットの増加に対応できる新しいストレージや既存のストレージは存在するのかといったことです(ストレージについては後程詳細に説明します)。

Cachéをベンチマークし、顧客のパフォーマンスデータを調査した私の経験では、コンピューターのリソース(CPUコア)を追加するにつれ、

Cachéは単一サーバーで非常に高いスループットにリニアスケール

することができ、さらに年ごとのCaché自体の改善による恩恵もあります。言い換えると、最大アプリケーションスループットのリニアスケールは、CPUコアが追加されるにつれ、アプリケーショントランザクションやCachéに反映されるのがわかります。ただし、アプリケーションのボトルネックが存在する場合、トランザクションレートが高くなるとそれが表面化し始め、リニアスケールに影響を与えてしまいます。アプリケーションのボトルネックの症状を見分ける方法については、別の記事で説明します。アプリケーションのパフォーマンス機能を改善する最善の方法に、Cachéを最新バージョンにアップグレードすることが挙げられます。

注意: Cachéでは、64個を超える論理コアを搭載したWindows 2008サーバーはサポートされていません。たとえば、40コアサーバーのハイパースレッディングは無効にしておく必要があります。Windows 2012では、最大640個の論理プロセッサがサポートされています。Linuxには制限はありません。

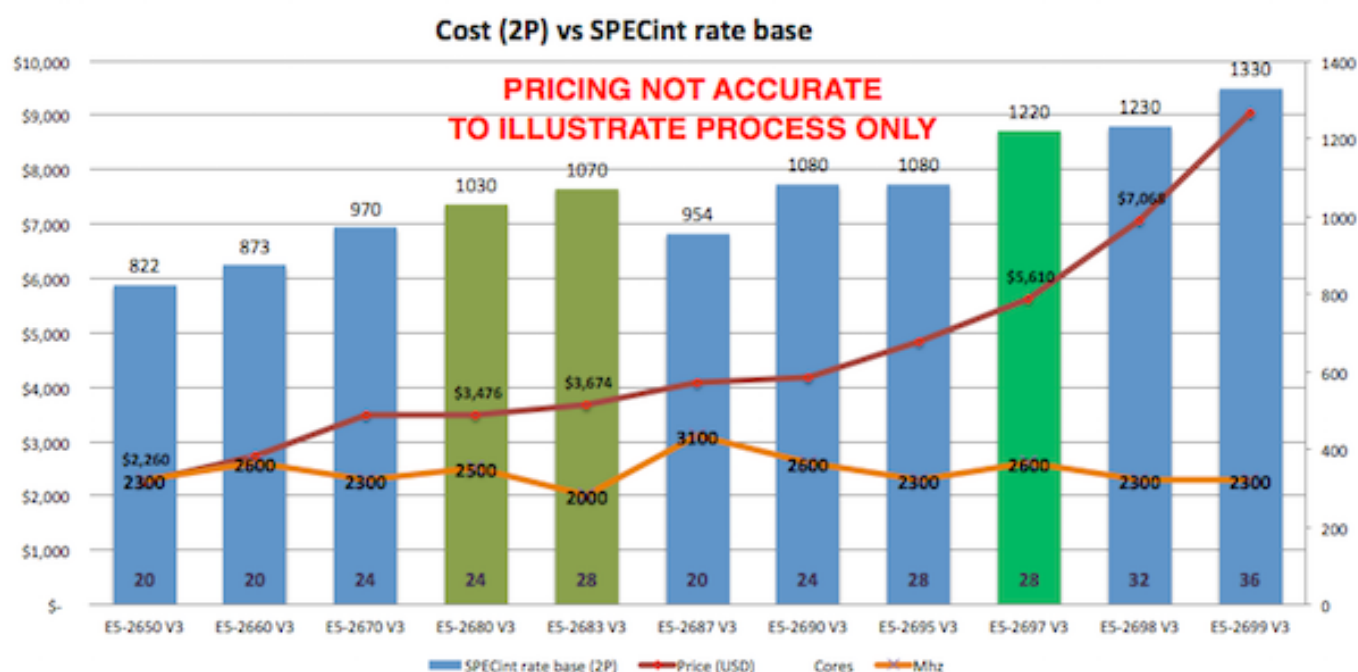
アプリケーションに必要なコア数は？

アプリケーションは多様であり、あなたのアプリケーションプロファイルはあなた自身にしかわからないことではありますが、サーバー(または仮想マシン)のCPUキャパシティプランニングを行う場合に私が共通して使用するアプローチは、入念なシステム監視活動です。ある「標準」プロセッサの特定の数のCPUコアが、1分当たり n 個のトランザクションというピーク時トランザクションレートを維持できることを確認する方法です。これは、エピソードであっても、受診時対応やラボテストなど、あなたの環境で意味を成すものであっても構いません。ポイントは、標準プロセッサのスループットは、現在のシステムまたは顧客のシステムで収集したメトリックに基づくものであるということです。

n 個のコアを備えた既知のプロセッサにおけるピーク時のCPUリソース使用率を知っているのであれば、SPECintを使用して、同じトランザクションレートでより新しいプロセッサまたは別のプロセッサに必要なコア数に変換することができます。期待されるリニアスケーリングにおいて($2 \times$ 毎秒 n 個のトランザクション)はざっと2倍のコア数が必要ということになります。

プロセッサの選択

spec.org Webサイトで、もしくは選択したベンダー製品を見てわかるように、プロセッサの選択肢はたくさんあります。この例のお客様はIntelで満足しているため、私が現在のIntelサーバーを推奨するとした場合、「値段に見合う価値」または「コアあたりの1ドルあたりのSPECint rate base 2006」を探るのが1つの選択方法として挙げられます。たとえば、次のグラフではDellコモディティサーバーについて表示しています。価格の有用性は人によって異なりますが、これは仮想化を使ったサーバーの統合には、価格と高いコア数にスイートスポットがあることを示しています。このグラフは、Dell R730などの本番グレードのサーバーの価格を設定してから、別のプロセッサオプションを検討して作成したものです。



グラフのデータと顧客サイトにおける経験に基づく、E5-2680 V3プロセッサは、コアあたりSPECintあたりの良好のパフォーマンスと価格ポイントを示しています。

また、ほかの要因も関係してきます。たとえば、仮想化デプロイメント向けのサーバープロセッサを検討している場合、プロセッサあたりのコア数を増やせばコストは高くなりますが、すべてのVMをサポートするために必要なサーバーホストの合計数が減るため、プロセッサソケット単位でライセンス提供されるソフトウェア（VMwareやオペレーティングシステムなど）を節約できるため、より安価に収めることができます。また、高可用性(HA)要件に対するホスト数のバランスも考慮する必要があります。VMwareとHAについては、後の記事でもう一度取り上げることにします。

たとえば、3つの24コアホストサーバーで構成されるVMware HAクラスターは、優れた可用性と大幅な処理能力(コア数)を提供するため、本番と非本番のVMの構成を柔軟に行うことができます。VMware HAのサイズはN+1サーバーで決めることができるため、3つの24コアサーバーの場合、VMで利用できる合計コアは48コアに相当します。

コアとGHz-Cachéに最適なものは？

CPUコアの速度と数のどちらかを選択する場合、次の点を考慮する必要があります。

- アプリケーションに多くのcache.exeスレッド/プロセスが必要な場合は、コア数を増やすことで、より多くのスレッド/プロセスをまったく同時に実行することができる。
- アプリケーションのプロセスが少ない場合は、それぞれができる限り高速に動作する方が望ましい。

別の見方をすると、たとえば同時ユーザーあたり1つ(以上)のプロセスといったように、多数のプロセスがあるクライアント/サーバーアプリケーションがある場合、利用できるコア数がさらに必要ということになります。CSPを使ったブラウザベースのアプリケーションの場合、ユーザーが少数の非常にビジーなCSPサーバーにバンドルされるため、アプリケーションに必要なコア数は少ない代わりにより高速である方がメリットを得られる可能性があります。

理想的には、複数のcache.exeプロセスがそれらすべてのコアで同時に実行している場合にリソースの競合がないと仮定すれば、いずれのアプリケーションタイプにおいても、高速コアを多く使用の方がメリットが高いと言えます。上で述べたように、CachéはリリースされるたびにCPUリソースの使用が改善されているため、最新バージョンのCachéにアプリケーションをアップグレードすることで、より多くのコアを利用できるようになります。

もう1つの重要な考慮事項は、仮想化を使用する際にホストあたりのコア数を最大化することです。個別のVMのコア数が多くない場合でも、コア数を増やすことで可用性に必要なホスト数と、管理とコストを考慮してホスト数を最小限にすることのバランスを取る必要があります。

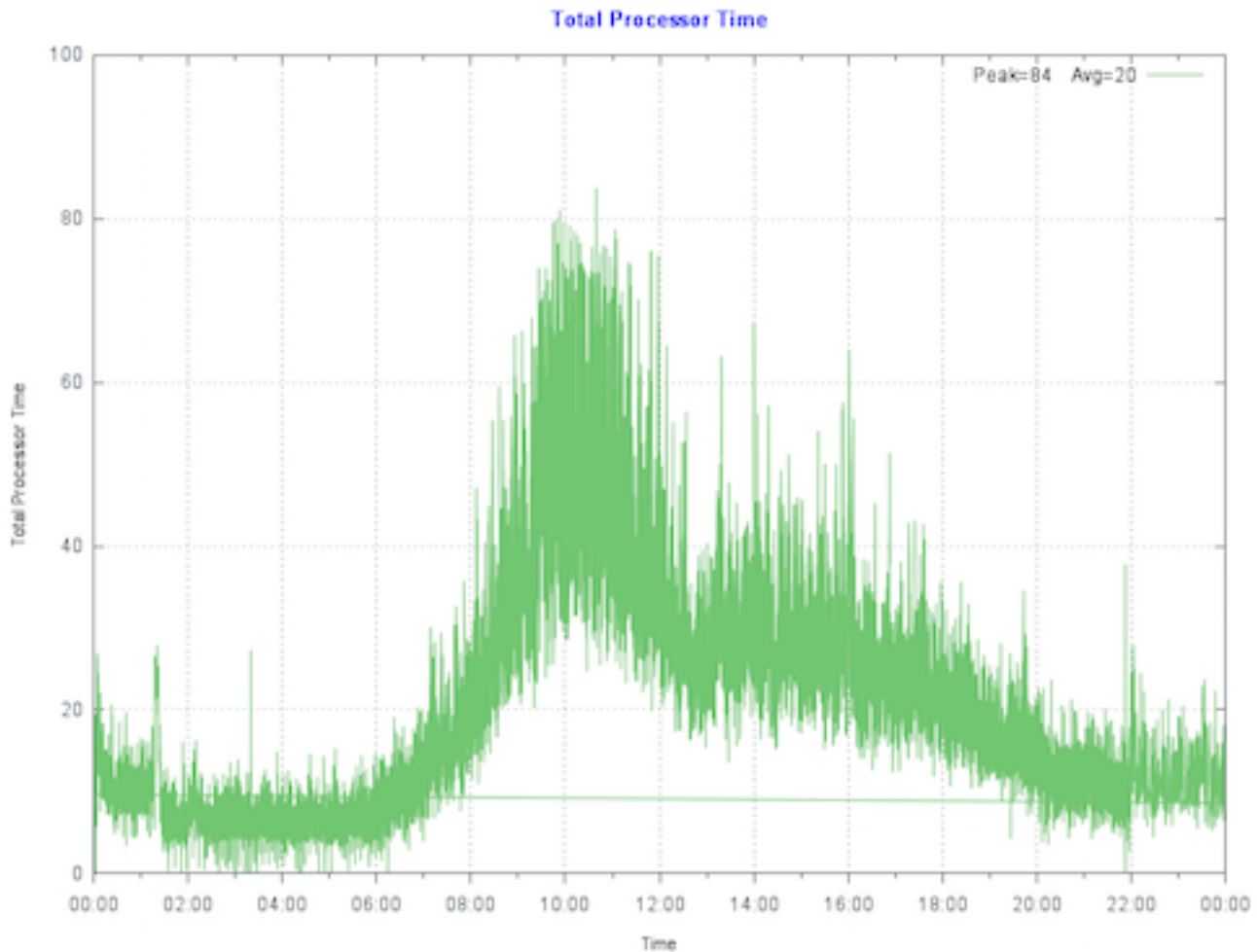
VMware仮想化とCPU

VMware仮想化は、最新のサーバーとストレージコンポーネントと合わせて使用した場合に、Cachéでうまく機能します。物理キャパシティプランニングと同じルールに従えば、適切に構成されたストレージ、ネットワーク、およびサーバーでVMware仮想化を実施しても、パフォーマンスに大きな影響はありません。仮想化サポートは、より新しいモデルのIntel Xeonプロセッサではるかに優れています。具体的には、Intel Xeon 5500（Nehalem）以降、つまりIntel Xeon 5500、5600、7500、E7シリーズ、およびE5シリーズのみでの仮想化を検討する必要があります。

例: ハードウェアの更新 - 最低CPU要件の計算

上で述べたヒントと手順をまとめ、例として、8コア(2つの4コアXeon 5570プロセッサ)のDell PowerEdge R710で実行するワークロードのサーバーアップグレードを検討します。

顧客のプライマリ本番サーバーにおける現在のCPU使用率を図に表すと、1日の最も忙しい時間帯におけるサーバーのピークは、80%未満であることがわかります。実行キューは圧迫されていません。IOとアプリケーションにも問題がないので、人為的にCPUを抑制しているボトルネックはありません。



経験則: 予測される成長(ユーザー、トランザクションの増加など)を考慮し、ハードウェアの使用期間の終わりに最大80%のCPU使用率となるようにシステムをサイジングします。こうすることで、予想外の成長、異常なイベント、または予定外のアクティビティの急増に対応することができます。

計算をわかりやすくするために、新しいハードウェアの使用期間中にスループットの増加がないと仮定します。

コアあたりのスケーリングは、 $(251/8):(1030/24)$ またはコアあたりのスループット26%増として計算できます。

8コアを使用する古いサーバーのCPU使用率80%は、ざっと、**6コア**を使用する新しいE5-2680 V3プロセッサのCPU 80%に相当します。
したがって、6つのコアで同じ数のトランザクションをサポートできます。

お客様にはいくつかの選択肢があり、最低CPU要件である6個のE5-2680 V3または同等のCPUコアを満たす新しいベアメタルサーバーを購入するか、本番ワークロードをVMwareで仮想化する計画を進めることができます。

サーバーの統合、柔軟性、および高可用性を活用するには、仮想化を選択するのが当然です。CPU要件の算出が終われば、お客様は自信をもってVMware上の本番VMの適正化に進むことができます。補足として、コア数の少ない最新のサーバーを購入するのは調達が困難であるかコストがかかるため、仮想化のオプションはさらに魅力的と言えます。

大幅な成長が見込まれる場合にも、仮想化が有利と言えます。CPU要件は、最初の数年間の成長に基づいて計算できます。継続的な監視により、必要に応じて必要となる前に必要なリソースのみを追加する、というのも有効な戦略といえるでしょう。

CPUと仮想化に関する考慮事項

これまで見てきたように、本番Cachéシステムのサイジングは実稼働の顧客サイトのベンチマークと測定に基づきます。また、ベアメタルの監視からVMwareの仮想CPU(vCPU)要件をサイジングすることも有効です。共有ストレージを使用した仮想化は、ベアメタルに比べてCPUオーバーヘッドがほとんど付加されません^{**}。本番システムには、ベアメタルのCPUコアと同じようにシステムのサイズを最初に決定する戦略を使用してください。

****注意:** VMware VSANデプロイメントについては、VSAN処理用に10%のホストレベルCPUバッファを追加する必要があります。

仮想CPU割り当てについては、次の重要なルールを考慮する必要があります。

推奨: 余裕を持ったパフォーマンスに必要な数以上のvCPUを割り当てないようにしてください。

- 仮想マシンには大量のvCPUを割り当てることができますが、必要以上のvCPUを割り当てないことがベストプラクティスです。使用されていないvCPUを管理するために、(通常は小さい)パフォーマンスのオーバーヘッドが生じる可能性があります。ここで重要なのは、システムを定期的に監視し、VMの適正化を確認することです。

推奨: 本番システム、特にデータベースサーバーは、初めに1つの物理CPU = 1つの仮想CPUでサイジングしてください。

- 本番サーバー、特にデータベースサーバーの使用率は非常に高くなることが予想されます。6つの物理コアが必要な場合は、仮想コアも6つにサイジングしてください。下のハイパースレッディングに関する注意も参照してください。

オーバーサブスクリプション

オーバーサブスクリプションとは、物理ホストで利用可能なリソースよりも多くのリソースをそのホストがサポートする仮想サーバーに割り当てるさまざまな方法を指します。一般に、仮想マシンの処理、メモリ、およびストレージリソースをオーバーサブスクライブすることで、サーバーを統合することができます。

ホストのオーバーサブスクリプションは、本番Cachéデータベース環境においても可能ですが、本番システムの初期サイジングでは、各vCPUが完全に専用のコアを持つと想定します。たとえば、24コア(2 12コア)E5-2680 V3サーバーを使用している場合は、統合に利用できるヘッドルームがあることを考慮しつつ、最大でも合計24vCPUキャパシティになるようにサイジングします。この構成では、ホストレベルでハイパースレッディングが有効化されていることが前提です。ピーク処理期間のアプリケーション、オペレーティングシステム、およびVMwareのパフォーマンスをしばらく監視したら、より高い統合が可能であるかどうかを判断することができます。

非本番VMを混在させている場合、合計CPUコアを計算するためのシステムサイジングにおいて私がよく使用する経験則は、非本番の物理対仮想CPUのサイズを最初は 2:1で設定することです。ただし、これは確実に有効性が異なる領域であり、キャパシティプランニングには監視が必要となります。疑問がある場合や経験がない場合は、ホストレベルか、ワークロードを理解するまでvSphere構成を使用して、本番VMと非本番VMを分けることができます。

VMware vRealize Operationsとほかのサードパーティツールには、経時的にシステムを監視し、統合を提案したり、VMにリソースを追加する必要があることを警告したりするファシリティがあります。監視に利用できるツールについては、今後の記事で取り上げることにします。

結論として、お客様の例では、6 vCPUの本番VMでうまく機能することを確認できます。もちろん、IOやストレージなどのほかの主要 ” 食品群 ” コンポーネントに十分な容量があると想定した上です。

ハイパースレッディングとキャパシティプランニング

物理サーバーに関する既知のルールに基づいてVMのサイジングに着手するには、まず、ハイパースレッディング

が有効化された状態でプロセッサあたりのターゲットに対する物理サーバーのCPU要件を計算し、それを単に変換します。

1つの物理CPU(ハイパースレッディングを含む)= 1つのvCPU(ハイパースレッディングを含む)

ハイパースレッディングによってvCPUキャパシティが2倍になるという誤解が一般的にありますが、これは、物理サーバーまたは論理vCPUには当てはまりません。経験則として、ベアメタルサーバーでのハイパースレッディングは、ハイパースレッディングを使用しない同じサーバーよりもパフォーマンスキャパシティが増加する率は30%です。同じ30%のルールは、仮想サーバーに適用されます。

ライセンスとvCPU

vSphereでは、特定の数のソケットまたはコアを持つようにVMを構成できます。たとえば、デュアルプロセッサVMを使用している場合、2つのCPUソケット、または2つのCPUコアを持つ単一のソケットに構成することができます。VMが1つの物理ソケットで実行するのか2つの物理ソケットで実行するのかは、最終的にハイパーバイザーが決定するため、実行の観点からは、大きな違いはありません。しかし、デュアルCPU VMに実際には2つのソケットではなく2つのコアがあると指定すると、Cache以外のソフトウェアライセンスに違いが生じる可能性があります。

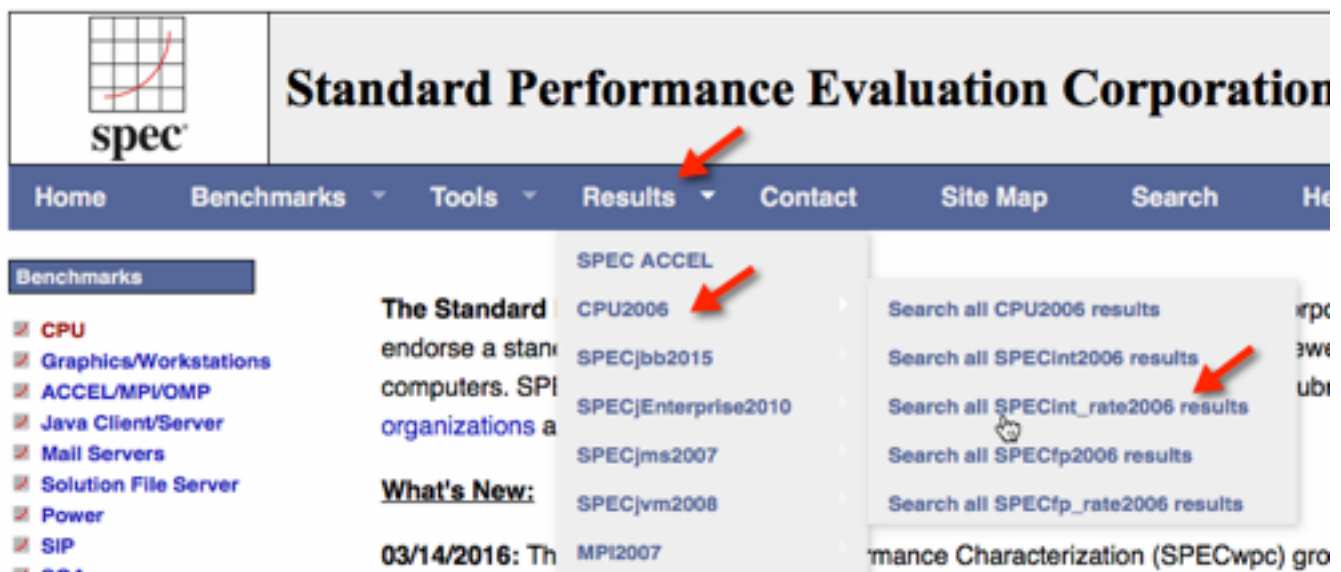
最後に

この記事では、SPECintベンチマークの結果を使用して、プロセッサをベンダー、サーバー、またはモデル間で比較する方法を説明しました。また、仮想化の使用状況に関係なく、パフォーマンスとアーキテクチャに基づいて、キャパシティを計画する方法とプロセッサを選択する方法についても説明しました。

このトピックは非常に深く、洞窟のさらに奥へと進みがちですが、別の方向性を探りたい方は、ほかの記事と同様にコメントやご質問をお寄せください。

SPECintrate2006の結果の検索例

下の図は、SPECintrate2006の結果の選択方法を示します。



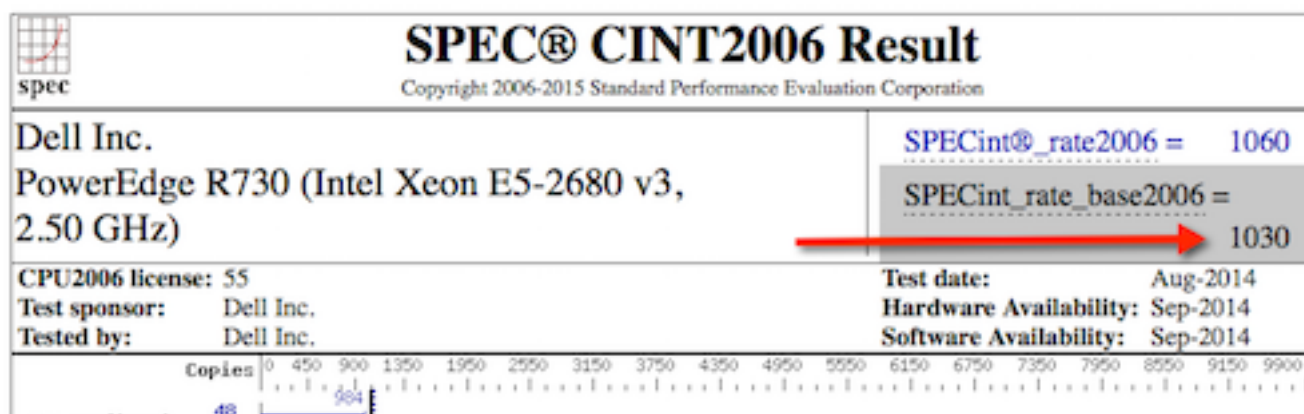
検索画面で結果を絞り込みます。

Excelなどを使ってローカルで処理できるように、最大20 MBの .csvファイルにすべてのレコードをダンプすることもできます。

検索結果には、Dell R730が表示されます。

Hardware Vendor	System	# Cores	# Chips	# Cores Per Chip	Base Copies	Result	Baseline	Published	Disclosure
Dell Inc.	PowerEdge R730 (Intel Xeon E5-2680 v3, 2.50 GHz)	24	2	12	48	1060	1030	Sep-2014	PS Text Config HTML CSV PDF
Dell Inc.	PowerEdge T810 (Intel Xeon E5-2680 v3, 2.50 GHz)	24	2	12	48	1070	1030	Sep-2014	PS Text Config HTML CSV PDF

HTMLを選択すると、ベンチマークの全結果が表示されます。



この記事の例で使ったプロセッサを搭載するサーバーについて、次の結果を確認できます。

Dell PowerEdge R710 with 2.93 GHz: 8 cores, 2 chips, 4 cores/chip, 2 threads/core
Xeon 5570: **SPECint_rate_base2006 = 251**

PowerEdge R730 (Intel Xeon E5-2680 v3, 2.50 GHz) 24 cores, 2 chips, 12 cores/chip, 2 threads/core
Xeon E5-2680 v3: **SPECint_rate_base2006 = 1030**

[#システム管理](#) [#パフォーマンス](#) [#Caché](#) [#InterSystems IRIS](#) [#InterSystems IRIS for Health](#)

ソースURL:

<https://jp.community.intersystems.com/post/intersystems%E3%83%87%E3%83%BC%E3%82%BF%E3%83%97%E3%83%A9%E3%83%83%E3%83%88%E3%83%95%E3%82%A9%E3%83%BC%E3%83%A0%E3%81%A8%E3%83%91%E3%83%95%E3%82%A9%E3%83%BC%E3%83%9E%E3%83%B3%E3%82%B9%C2%A0%E2%80%93%E3%83%91%E3%83%BC%E3%83%88%C2%A0cpu%E3%81%AB%E6%B3%A8%E7%9B%AE>