

ホワイトペーパー

FUJITSU x86 Servers & Workstations

ベンチマークの概要 SPECcpu2006

このドキュメントでは、Standard Performance Evaluation Corporation（SPEC）によって開発および公開されているベンチマーク SPECcpu2006 について説明します。

このベンチマークはプロセッサやメモリのパフォーマンスを測定するもので、業界で広く受け入れられているベンチマークです。SPECcpu2006 で実現されているベンチマークのコンセプトにより、多機種を一律に比較できます。

バージョン

1.3
2015-09-16



目次

ドキュメントの履歴	2
SPECcpu2006 – 概要	3
ベンチマークの目的	3
ベンチマークの構成要素	4
測定結果とその解釈	6
関連資料	8
お問い合わせ先	8

ドキュメントの履歴

バージョン 1.0

初版

バージョン 1.1

レイアウト変更

バージョン 1.2

レイアウト変更

バージョン 1.3

測定結果を有効数字 3 桁に丸めることに関するパラグラフの挿入

SPECcpu2006 – 概要

SPEC (Standard Performance Evaluation Corporation) は、コンピュータシステムのパフォーマンスの評価を目的とする、標準化されたベンチマークの開発と提供に特化した組織です。SPEC コンソーシアムのメンバーは、富士通など、コンピュータ業界の有力企業です。SPEC が提供したベンチマークの 1 つに CPU2006 があり、CPU2000 の後継ベンチマークです。このドキュメントの目的は、このベンチマークの重要性とその結果の有意性について、概要を示すことです。

ベンチマークの目的

システムを比較するときは、原則として、個々のお客様のアプリケーションが評価の基準になります。しかし、時間の不足や複雑さなどの理由のため、お客様のアプリケーションでさまざまなシステムをテストして最適解を見つけ出すことができない場合もあります。ベンチマークとは、お客様がシステムを選択するときに客観的な基準を提供するために開発されたものです。これらは、個々のシステムコンポーネントの速度やシステム全体の速度を測定する、標準化されたツールです。

CPU2006 は、計算処理性能のシステム間比較を行えるように開発されたベンチマークです。厳密には、CPU2006 は単一のベンチマークではなく、次の 2 つのベンチマークグループで構成されています。

- 整数演算時のシステムパフォーマンスを測定するための CINT2006
- 浮動小数点演算時のシステムパフォーマンスを測定するための CFP2006

ベンチマークグループの名前の「C」は「コンポーネント」を表しています。したがって、これはシステムベンチマークではなく、コンポーネントベンチマークとすることができます。計算処理ベンチマークとして、このベンチマークはプロセッサに加え、メモリアーキテクチャとコンパイラを評価対象としています。重要：最後に挙げたコンポーネント（メモリとコンパイラ）は無視しないでください。パフォーマンスは、プロセッサの周波数だけの問題ではありません。重要：I/O、グラフィックス、またはネットワークなど、他のシステムコンポーネントは負荷テストの対象ではありません。アプリケーションのパフォーマンスボトルネックがこれらのいずれかのコンポーネントに起因している場合、CPU2006 ベンチマークではこれを認識できません。したがって、CPU2006 が個々のケースで適切なベンチマークであるかどうかは、アプリケーションの負荷プロファイルによって決まります。

ベンチマークの構成要素

ベンチマークグループ CINT2006 は、C および C++ で記述された次の 12 のアプリケーションで構成されています。

ベンチマーク	アプリケーション分野
400.perlbench	プログラミング言語 Perl V5.8.7 から派生。ワークロードには、SpamAssassin、MHonArc（電子メールインデクサー）、および specdiff（ベンチマーク出力をチェックする SPEC のツール）があります。
401.bzip2	圧縮 Julian Seward の bzip2 バージョン 1.0.3。I/O を実行するのではなく、ほとんどの作業をメモリで実行するように変更されています。
403.gcc	C コンパイラ gcc バージョン 3.2 に基づき、Opteron 用のコードを生成します。
429.mcf	組み合わせの最適化 車両スケジュール。ネットワークシンプレックス法アルゴリズム（市販製品でも使用されています）を使って公共交通機関のスケジュールを作成します。
445.gobmk	人工知能：囲碁 簡素なルールでありながら、複雑なゲームである囲碁をプレイします。
447.dealll	遺伝子配列の検索 プロファイル隠れ Markov モデル（プロファイル HMMs）を使用するたんぱく質配列分析
458.sjeng	人工知能：チェス 高ランクのチェスプログラムで、チェスの複数の変種もプレイできます。
462.libquantum	物理学／量子計算 ショアの多項式時間因数分解アルゴリズムを実行して、量子コンピュータをシミュレーションします。
464.h264ref	ビデオ圧縮 H.264/AVC の基準実装。2つのパラメータセットを使用してビデオストリームをエンコードします。MPEG2 は、H.264/AVC 標準に置き換えられる予定です。
471.omnetpp	離散イベントシミュレーション OMNet++ 離散イベントシミュレータを使用して、大規模な Ethernet キャンパスネットワークをモデル化します。
473.astar	経路探索アルゴリズム よく知られている A* アルゴリズムなど、2D マップ用の経路探索ライブラリ。
483.xalancbmk	XML 処理 Xalan-C++ の変更バージョン。XML ドキュメントを他のドキュメントタイプに変換します。

ベンチマークグループ CFP2006 は、C、C++、および FORTRAN で記述された次の 17 のアプリケーションで構成されています。

ベンチマーク	アプリケーション分野
410.bwaves	流体力学 3D 遷音速遷移粘性層流の計算を実行します。
416.gamess	量子化学 Gamess は広範囲な量子化学計算を実装します。SPEC ワークロードの場合、自己無撞着場の計算は Restricted Hartree Fock 法、Restricted open-shell Hartree-Fock、および多配置自己無撞着法を使って実行されます。
433.milc	物理学／量子色力学 動的クォークによる格子ゲージ理論プログラムのゲージ場生成プログラム。
434.zeusmp	物理学／CFD ZEUS-MP は、Laboratory for Computational Astrophysics（米国、イリノイ大学、Urbana-Champaign 校、NCSA）で宇宙現象のシミュレーション用に開発された計数流体力学コードです。
435.gromacs	生化学／分子動力学 分子動力学。つまり、数百から数百万の粒子の動きのニュートニアン運動方程式をシミュレーションします。このテストケースは、溶液中のたんぱく質リゾチームをシミュレーションします。
436.cactusADM	物理学／一般相対性理論 互い違い蛙跳び（staggered leapfrog）法を使って、アインシュタインの発展方程式を解きます。
437.leslie3d	流体力学 3 次元リニアエディモデルによるラージェディシミュレーションを使った数値流体力学（CFD）。MacCormack 予測子修正子時間積分スキームを使用します。
444.namd	生物学／分子動力学 大規模な生体分子系をシミュレーションします。このテストケースには、アポリポたんぱく質 A-I の 92,224 個の原子が含まれています。
447.dealII	有限要素解析 deal.II は、適応型有限要素と誤差推定を対象とする C++ プログラムライブラリです。このテストケースは、非定数の係数を使ってヘルムホルツ方程式を解きます。
450.soplex	線形計画法、最適化 シンプレックスアルゴリズムとスパース線形代数を使用して線形プログラムを解きます。テストケースには、鉄道計画および軍用空輸モデルがあります。
453.povray	画像レイトレーシング 画像レンダリング。パーリンノイズ関数を使用するテクスチャ付き抽象オブジェクトを含む風景の 400x400 アンチエイリアシング画像です。
454.calculix	構造力学 線形および非線形 3D 構造アプリケーション用の有限要素コード。SPOOLES ソルバライブラリを使用します。
459.GemsFDTD	計算電磁気学 時間領域差分（FDTD : finite-difference time-domain）法を使用して、3 次元マクスウェル方程式を解きます。
465.tonto	量子化学 Fortran 95 のオブジェクト指向デザインを使用する、オープンソースの量子化学パッケージです。このテストケースは、X 線解析実験データと整合させるために、分子のハートリーフォック波動関数の計算に制約を加えます。
470.lbm	流体力学 3 次元の非圧縮性流体をシミュレーションする格子ボルツマン法（Lattice Boltzmann Method）を実装します。

481.wrf	気象 数メートルから数千キロメートルまでの気象モデルです。このテストケースは、2 日間にわたる 30 km の地域からのものです。
482.sphinx3	音声認識 Carnegie Mellon University が提供している有名な音声認識システムです。

すべての個々のベンチマークは、入力ファイルが読み込まれたときに発生するファイル I/O のみを共有します。計算では、キャッシュと RAM のみを使用します。ただし、このためにシステムに必要な最小限度のメモリ容量は、32 ビットシステムでは 1 * プロセッサコア数 GB、64 ビットシステムでは 2 * プロセッサコア数 GB です。これには、オペレーティングシステム用の容量、および SPECcpu2006 以外のタスクに必要な容量は含まれません。

測定結果とその解釈

CPU2006 では、手段や方法論が異なる次の 8 つの測定を実行できます。

ベンチマーク	単一ベンチマークの数	演算	タイプ	コンパイラ最適化	測定結果	アプリケーション
SPECint2006	12	整数	ピーク	アグレッシブ	速度	単体実行
SPECint_base2006	12	整数	ベース	標準		
SPECint_rate2006	12	整数	ピーク	アグレッシブ	スループット	多重実行
SPECint_rate_base2006	12	整数	ベース	標準		
SPECfp2006	17	浮動小数点	ピーク	アグレッシブ	速度	単体実行
SPECfp_base2006	17	浮動小数点	ベース	標準		
SPECfp_rate2006	17	浮動小数点	ピーク	アグレッシブ	スループット	多重実行
SPECfp_rate_base2006	17	浮動小数点	ベース	標準		

ベンチマークは、種々のハードウェアモデルで実行できるようにソースコードとして SPEC から提供されています。このため、まずコンパイルする必要があります。ただし、プログラムのコンパイルは標準化された手順ではありません。複数の最適化オプションがあり、あるプログラムが使用しても、別のプログラムが無視する場合があります。この問題を考慮して、SPEC は 厳密なガイドラインが存在する標準（ベース）最適化による測定（原則として個々のベンチマークすべてに同じ順序で同じ最適化フラグを使用）と、ガイドラインがより緩やかなアグレッシブ（ピーク）最適化による測定（個々のベンチマークを個別に最適化することが可能）の両方を許可しています。前者は必須であり、後者は省略可能です。

SPEC は、2 つの異なるパフォーマンス測定方法を提供しています。1 つの方法（速度）では、1 つのタスクの実行に必要な時間を測定します。もう 1 つの方法（レート）ではスループットを測定します。つまり、事前に定義された数の複数のタスクを並列で実行するために必要な時間です。

ベンチマークの結果はすべて、個々のベンチマークで判定された標準化比の幾何平均を表しています。「標準化」とは、テストシステムがリファレンスシステムと比較してどの程度高速に実行されるかを測定することをいいます。SPEC は、296 MHz UltraSPARC II プロセッサを搭載した Sun の Ultra Enterprise 2 システムをリファレンスシステムとして選択しています。各ベンチマークはこのシステムで実行され、基準時間が測定されています。すべてのベースベンチマークで、リファレンスシステムに対して値「1」が決められています。テストシステムの結果は、テストシステムがリファレンスシステムと同数の作業を、リファレンスシステムが必要とする時間と同じ時間で実行できることを意味します。たとえば、テストシステムで SPECint_base2006 の値が 2 の場合は、テストシステムがこのベンチマークをリファレンスシステムの約 2 倍の速度で実行したことを意味します。または、テストシステムの SPECfp_rate_base2006 の値が 7.8 である場合は、テストシステムがこのベンチマークをリファレンスシステムよりも約 7.8/[ユーザ数] 倍の速度で実行し、[ユーザ数] のユーザを並列で実行したことを意味します。一般に、2 つのシステムでユーザ数が同じ場合、高い結果のシステムの方が高速です。ユーザ数が異なる場合、高い結果のシステムがより多くを処理します。この理由は、そのシステムの方が高速であるか、並列でより多くの作業を実行できるためです。

幾何平均は、1 つ 1 つのベンチマークの結果から全体のベンチマークの結果を計算するために使用されます。この方法の効果は、算術平均と比較して、低い単一結果の影響が大きくなることです。この背景にある意図は、特定のベンチマーク向けにコンパイラを最適化して不釣り合いに高い重みをそのベンチマークに与えることを防止するためです。さらに、これにより向上しないままになっているコンポーネントへのシフトが生じます。優れたプロセッサは、キャッシュアクセスと比較して多数の RAM アクセスを必要とする単一ベンチマークに強い影響を受けません。他の単一ベンチマークと比較して劣る結果は、幾何平均の使用により強調されます。

測定ラン間でパフォーマンスの小さな変動があるのは当然であるため、すべての SPEC 測定結果は意図的に有効数字 3 桁に丸められます。例えば、測定結果の SPECint_rate 値 1234.999 と SPECint_rate 値 1225.000 は両方とも、SPECint_rate 値 1230 に丸められます。html、pdf、ps、txt 形式の結果ファイルには、この丸められた値が示されます。SPEC の結果（1230 より大きい場合など）のリクエストも、常に丸められた値を参照します。

ベンチマーク結果の解釈は、競合する製品に対し、システム全体を比較するか個別のコンポーネント（プロセッサ、メモリ、コンパイラ）を比較するかによって異なります。1 つのコンポーネントに焦点を当てる場合は、その他のコンポーネントが同じであると考えする必要があります。そうでない場合、個別のコンポーネントに値を明確に割り当てることはできません。したがって、公開する際はシステム構成の記述の正確性に万全を期しています。

弊社は、SPECcpu2006 を使用した測定結果を SPEC での公開用に提供しています。しかし、弊社のすべての結果が SPEC の Web サイトで公開されているわけではありません。弊社は、すべての測定値のログデータをアーカイブしているため、提供していないものについても、測定値が正しく達成されていることをいつでも証明できます。

関連資料

PRIMERGY サーバ

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

PRIMEQUEST サーバ

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

CELSIUS ワークステーション

<http://www.fmworld.net/biz/celsius/>

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/>

SPECcpu2006

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>

<http://www.spec.org/osg/cpu2006>

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>