

## ① 性能評価の意義と定義

### ① システム評価の意義

経費の有効利用や節約の目的から、コストパフォーマンスの良好な計算機利用が追求される。計算機システムの複雑化に伴い、システムの振る舞いや性能を的確に把握するための手段が求められる。計算機の使用法の多様化により、種々の立場からの評価が必要になっている。新しいシステムの設計に従来のシステムの評価データが不可欠である。

### ② システムの性能

性能を評価する場合、コンピュータシステムの性能について明確に定義する必要がある。性能はあるプログラムをコンピュータシステム上で実行した場合の定量的な指標である。

### ③ システム評価のための定量的な性能指標

- ㊦ コンパイルなどの言語処理時間を含めたコンピュータシステムでの実行時間
- ㊧ 主記憶上に割り付けられたオブジェクトプログラムが結果を出すまでの実行時間
- ㊨ プロセッサが使用される実質時間、CPU時間
- ㊩ そのプログラムを実行するのに必要なメインメモリ量

### ④ システム評価の方法

システムの評価はシステムの開発計画の段階から、システムの運用、新しいシステムの計画に至るライフサイクルの各段階で種々の方法が利用される。評価手法の代表的なものに、解析的方法、シミュレーション法、モニタリングなどがある。

開発・導入段階の評価では、システムのモデルを作り解析的方法やシミュレーション法を使用して事前評価を行う。ベンチマークプログラムを利用して、処理スピードやスループットを実測あるいは実測データにもとづく推定が行われる。稼働中のシステムの評価には、モニタリング手法が用いられる。モニタリング手法は、運転中の計算機システムの動作状況を観測し、定量的にその動作パターンを把握するための手法である。

## ② 評価の測定法

### ① カタログデータ

カタログは容易に得られる資料で、システムの概要を知る上で役立つ。CPUのサイクル時間、基本演算時間、主記憶のアクセス時間、周辺装置の速さなどをカタログ性能として得ることができる。機種選択、機器構成決定の参考資料となる。

### ② インストラクションミックス

インストラクションミックスは、CPUの処理速度を単一の数値で表すために、使用命令を分類して使用頻度を重み付けし、加算平均により平均実行速度を求める。

### ③ ベンチマーク

ベンチマークは、処理能力を調べるために、計算機システムに標準的なテストプログラムを使用する方法である。入出力や制御プログラムの性能を含んだテストが実施できる。計算機システムのスループットなどの総合的性能のデータが得られる。

### ④ モニタリング

モニタリング手法は、運転中の計算機システムの動作状況を観測し、定量的にその動作パターンを把握するための手法で、計算機システム内を走る各プログラムの振る舞いや各装置の利用状況の変化をとらえるための測定である。

機器構成、ソフトウェア構成の改良、TSSの応答性能の測定や改良に用いる。ボトルネックの発見、プログラムの振る舞いを調べ改善する。

### ⑤ トレーサ

各種プログラムの振る舞いを調べるツールで、モジュール内あるいはモジュール間の走行経路をリアルタイムまたは実行後の一括出力として確認するために用いる。

トレーサ上でプログラムを実行すると、その過程を追跡・記録し、実行した命令を順番に、レジスタや変数の値などの状態と共に表示する。プログラムが意図したとおりに動作しているかを調べたり、どこに誤りがあるかを発見するのが容易になる。

## ③ プロセッサ・メモリの評価

### ① 絶対評価

評価の対象とするコンピュータの処理速度を示すことによって評価する。命令ミックスによるMIPS値やFLOPS値がある。

## ② 相対評価

ある基準のコンピュータを定め、それとの相対的な性能を示すことで評価する。速度向上比やベンチマークによる評価がある。

## ③ 静的評価

コンピュータを稼働させなくても評価データを提示できるものである。仕様として示される項目による評価である。

## ④ 動的評価

あるプログラムを実行させたときに示される項目による評価である。ハードウェアの実行速度を性能と定義する場合、意味があるのは動的評価である。

## ⑤ プロセッサおよびメモリの評価項目

- ㊦ 命令ごとの実行速度
- ㊧ ユーザが利用できるレジスタ個数
- ㊨ パイプラインの深さ
- ㊩ キャッシュなどのプロセッサに最も近いメモリ階層の容量や速度
- ㊪ ALU、ALU以外の演算器、命令フェッチ／デコード、メモリアクセス機能ユニットなどの種類、演算／処理の単位、並列度、個数
- ㊫ 命令の種類や命令セットの大きさ
- ㊬ メモリの容量
- ㊭ メモリのアクセス時間とサイクル時間
- ㊮ スループット、メモリの転送性能
- ㊯ メモリ階層の種類

## ④ ベンチマークテスト

### ㊰ ベンチマーク

ある応用分野の代表的なプログラムをサンプルプログラムとして評価対象のコンピュータでコンパイル・実行して得た性能を基準コンピュータの性能との相対値で示す指標である。ハードウェアによる実行だけでなくコンパイラやOSによる処理過程まで含めて評価できるので、コンピュータシステム全体を対象として評価することが可能である。動的評価手法としてベンチマークによる相対評価が使われる。

## ⑥ 代表的なベンチマークテスト

### ㊦ SPEC

複数本のベンチマークプログラムの実行速度をある基準コンピュータとの相対値でそれぞれ表し、それらの幾何平均をとる。ベンチマークプログラムには整数演算用と浮動小数点数演算用があり、それぞれのベンチマークをSPECint、SPECfp という。

SPECintは、C言語で作成したメモリアクセス、整数演算の速度を反映させるベンチマークテストである。SPECmarkは、Fortranで作成した倍精度浮動小数点ベンチマークテストである。SPECint92は、整数演算の占める割合の多い6種類のC言語で書かれた整数演算の性能を測るベンチマークテストである。SPECfp92は、浮動小数点演算の性能を測るベンチマークテストで、浮動小数点演算の占める割合の多い14種のテストから成る。SPECint92/SPECfp92の基準コンピュータはVAX-11/780であったが、SPECint95およびSPECfp95では、SPARCstation10 になっている。

### ㊧ ドライストン

整数演算性能を評価するベンチマークで、コンパイラの性能評価に用いる。

### ㊨ ウェットストーン

もともとは浮動小数点数演算性能を評価するものであったが、事務計算も考慮したベンチマークプログラムを加え汎用コンピュータ向きとしている。ベンチマークプログラムが小規模計算であるために、大規模キャッシュを装備するコンピュータでは正確な評価はできない。

### ㊩ リバモアループ

複数本のループプログラムによるベンチマークで、ベルトコンピュータの評価に用いる。

### ㊪ リムパック

浮動小数点数演算用ベンチマークで、FLOPS値の計測に利用される。

## ⑤ OLTPのベンチマーク

### ㊰ TPCベンチマーク

システムの性能を評価する方法にMIPS値があるが、OLTPシステムの場合、コンピュータシステム、通信装置および関係するソフトウェアについての性能の評価も必要となるため、MIPS値の代わりにTPCベンチマークを用いる。TPCベンチマークテストはアメリカのトランザクション処理性能評議会(TPC)が定めたものである。

## ⑥ TPCベンチマークテストの種類

### ア TPC-A

銀行業務のATMによる入出力をモデル化したもので、通信処理やデータベースの更新処理を含むオンライントランザクションシステムの基本性能評価用である。

### イ TPC-B

バッチ処理のデータベースシステムをモデル化したもので、データベース評価用である。

### ウ TPC-C

オーダエントリをモデル化したもので、業務処理性能評価用である。

### エ TPC-D

意志決定支援アプリケーションをモデル化したものである。

### オ TPC-H

意思決定支援システムをシミュレートするベンチマークテストである。

### カ TPC-R

データベースシステム用のベンチマークテストである。

### キ TPC-W

WebサーバやWebアプリケーションのフレームワークの性能評価に利用する。

## 例題演習

システムの性能評価に関する指標のうち、主記憶の競合状態を最もよく表すものはどれか。

ア 実行待ち時間

イ トランザクション応答時間

ウ ページング発生頻度

エ メモリ使用率

## 解答解説

主記憶の競合状態に関する問題である。

競合状態は複数のプロセスがある資源に同時にアクセスする状態を呼ぶ。

アの実行待ち時間は、プロセスが入出力を完了するまでCPUを使用するのを待っている状態であり、競合状態ではない。

イのトランザクション応答時間は、トランザクションがネットワークで送信され、コンピュータで処理されて、再びネットワークで送信されてくるまでの時間である。主記憶の競合ではない。

ウのページングは、主記憶と補助記憶の間でページイン・ページアウトされる現象で、頻度の大小は主記憶の競合状態を表す。求める答えはウとなる。

エのメモリ使用率は、主記憶の領域の使用割合を表すものであり、競合ではない。

### 例題演習

コンピュータシステムのベンチマークテストの説明として、最も適切なものはどれか。

- ア 1命令の実行に要する平均時間から、コンピュータの性能を測る。
- イ システムが連続して稼働する時間の割合を測定し、他の製品と比較する。
- ウ 想定されるトランザクション量にシステムが耐えられるかどうかを判定する。
- エ 測定用のソフトウェアを実行し、システムの処理性能を数値化して、他の製品と比較する。

### 解答解説

ベンチマークテストに関する問題である。

ベンチマークテストは、ある応用分野の代表的なプログラムをサンプルプログラムとして評価対象のコンピュータでコンパイル・実行して得た性能を基準コンピュータの性能との相対値で示す指標である。ハードウェアによる実行だけでなくコンパイラやOSによる処理過程まで含めて評価できるので、コンピュータシステム全体を対象として評価することが可能である。動的評価手法としてベンチマークによる相対評価が使われる。システムの性能評価に用いる。

アは、命令実行時間を使用して、CPUの性能を評価する。

イは、システムの稼働率を使用して、システムや装置の信頼性を評価する。

ウは、トランザクション処理能力からシステムの最大処理能力を測定し、評価する。

エは、ベンチマークテストである。求める答えはエとなる。

### 例題演習

性能評価法の一つであるTPCベンチマークによって評価できるものはどれか。

- ア OLTPの性能
- イ ディスク装置の性能
- ウ ネットワークの性能
- エ プロセッサの性能

### 解答解説

TPCベンチマークに関する問題である。

TPCベンチマークは、CPU性能だけでなく、ハードディスク装置、OS、データベース管理システム、ネットワークなどのシステム全体の性能を評価できる特徴がある。

OLTPはオンライン・トランザクション処理で、ホストコンピュータにオンライン接続された複数の端末がホストコンピュータにメッセージを送信し、そのメッセージに従って、ホストコンピュータで一連のデータベースアクセスを含む処理を行い、処理結果を即座に端末に返信する処理システムである。従って、OLTPの性能はTPCベンチマークテストによって評価することが可能になる。求める答えはアとなる。

### 例題演習

コンピュータの性能評価の基準に関する記述のうち、SPECint に関するものはどれか。

- ア 1秒間に実行可能な浮動小数点演算の回数、主に科学技術計算の性能尺度として用いられるが、超並列コンピュータの評価指数としても用いられる。
- イ 1秒間の平均命令実行回数。アーキテクチャが異なるコンピュータ間の性能比較には適さない。
- ウ UNIXが動作するコンピュータを主対象とした整数演算ベンチマーク。システム性能評価協会が開発し、標準的なベンチマークとして普及している。
- エ オンライントランザクション処理システム用ベンチマーク。対象とするモデル別にA、B、C、Dの4種のベンチマーク仕様が開発されている。

### 解答解説

ベンチマークテストに関する問題である。

SPECのベンチマークテストには次のものがある。

- ① SPECintはC言語で作成したメモリアクセス、整数演算の速度を反映させるベンチマークである。
- ② SPECmarkはFortranで作成した倍精度浮動小数点ベンチマークテストである。
- ③ SPECint92は整数演算の占める割合の多い6種類のC言語で書かれた整数演算の性能を測るベンチマークテストである。
- ④ SPECfp92は浮動小数点演算の性能を測るベンチマークテストで、浮動小数点演算の占める割合の多い14種のテストから成る。

アはFLOPSである。

イは100万回単位で表すとMIPSとなる。

ウはSPECintに関するものである。求める答えはウとなる。

エはTPCベンチマークテストである。

### 例題演習

システムの性能評価に関する記述のうち、最も適切なものはどれか。

- ア システムに組み込まれたスーパバイザ機能呼び出すことによって、性能評価に必要なデータをすべて測定することができる。
- イ システムをモデル化して性能評価を行う方法として、解析的手法によるものとシミュレーションによるものがある。
- ウ 性能評価のためにモニタリングを行うときは、ハードウェアモニタによるモニタリングよりも被測定系に影響が少ないソフトウェアによるモニタリングの方がよい。
- エ ベンチマークテストに使用するプログラムは、JISによって仕様が規定されているので、その仕様を満たすプログラムを使うことで公平なテストを行うことができる。

### 解答解説

システムの性能評価に関する問題である。

システムの評価手法の代表的なものに解析的方法、シミュレーション法、モニタリングなどの方法がある。開発・導入段階ではシステムのモデルを作り、解析的方法やシミュレーション法を使用して事前評価を行う。ベンチマークプログラムを利用して、処理スピードやスループットを実測あるいは実測データに基づく推定が行われる。ベンチマークは処理能力を調べるために計算機システムに標準的なテストプログラムを使用して、入出力や制御プログラムの性能を含んだテストを実施する。モニタリング手法は運転中の計算機システムの動作状況を観測し、定量的にその動作パターンを把握するための手法である。モニタリングは計算機システム内を走る各プログラムの振る舞いや各装置の利用状況の変化をとらえるための測定で、機器構成やソフトウェア構成などに用いる。

アのシステムのスーパバイザ機能は周辺機器の制御・管理に用いられるものであり、性能評価に必要なデータを測定するものではない。

イのモデル化して性能評価する方法に解析的手法とシミュレーションがある。イは正しい答である。求める答えはイとなる。

ウのモニタリングはハードウェアによるモニタリングとソフトウェアによるモニタリングが一体になってシステムの性能を評価するもので個別の優劣が問題ではない。

エのベンチマークテストはJ I SではなくS P E C、T P Cによって仕様が規定されている。

### 例題演習

オンラインシステムの性能を評価するとき、特に業務処理性能を評価する指標として、最も適切なものはどれか。

- |                |            |
|----------------|------------|
| ア 実行待ち時間       | イ チャンネル使用率 |
| ウ トランザクション応答時間 | エ ページング回数  |

### 解答解説

オンラインシステムの性能評価に関する問題である。

オンライントランザクションシステムの業務処理性能にはスループットと応答時間がある。生産性を表す性能としては単位時間当たりの処理件数即ちスループットに相当するものがあり、利用者へのサービスとしてはトランザクションを投入してから応答が返されるまでのトランザクション応答時間がある。

アの実行待ち時間は、サービスを受けるまでの待ち時間であり、サービス時間は考慮されていないため業務処理性能にはならない。

イのチャンネル使用率は、チャンネルの使用割合を示すものである。

ウのトランザクション応答処理は、利用者へのサービスを表す業務処理性能になる。求める答えはウとなる。

エのページング回数は、実記憶空間に所定のページが存在しないページフォルトに対応する回数を示すもので、どちらも業務処理割合とは調節に関係ないものである。