

HDDシミュレーション用モデル及び性能評価支援システムの開発

久本 康司 五十嵐 真悟

株式会社 東芝 研究開発センター

我々はHDD(Hard Disk Drive)を対象とし、設計段階でシミュレーションによりシステムの性能を評価することを目的としたシミュレーション用HDDモデルを開発した。このモデルを用いたシミュレーションを行うことにより、HDDのスループットの傾向を評価することが可能となった。またシミュレーションから性能の評価までをサポートするための幾つかのツールを開発し、HDDモデルを中心としたHDD性能評価支援システムを構築した。本システムを運用することにより、設計段階においてHDDの性能評価および性能チューニングを効果的に行うことが可能となり、要求性能を満たすHDDを短期間で開発することが可能となる。

Development of HDD Simulation Model and Performance Evaluation Support System

Koji Hisamoto Masato Igarashi

Corporate Reserch & Development Center, TOSHIBA Corporation

In this paper we present the HDD(Hard Disk Drive) model for simulation-based system performance evaluation at a design stage. This simulation model enables evaluation of HDD throughput and its tendency. We also developed several tools which facilitate simulation and evaluation, and integrated them with the HDD model as a Performance Evaluation Support System. This system enables developers to evaluate and optimize HDD performance efficiently in short time at the design stages.

1 はじめに

今日、組み込み用マイクロプロセッサを使ったシステムは、プロセッサの高性能化と要求される機能の増大に伴い、システム全体が複雑化している。このため、システム完成後の性能テストによって性能未達であった場合、性能チューニングのための工程の後戻りが発生し、開発期間が延びてしまうという問題が生じている。このような問題を解決するアプローチの一つとして、シミュレータを用いて設計段階でシステム全体の要求性能を満たすように性能評価を行うという方法が考えられる。実際のシステムと同等に動作する性能評価用のシミュレータを用いれば、システムが未完成の状態でも性能チューニングを行うことが可能となる。

コンピュータの外部記憶装置として広く利用されているハードディスクドライブ(Hard Disk Drive, HDD)は高い性能が要求されるリアルタイムシステムであり、且つTime To Marketが非常に重要な製品でもある。そのため、要求性能を満たすためのチューニング作業を、試作機が完成する以前の設計段階で完了しておき、開発期間を短縮するアプローチが有効である。

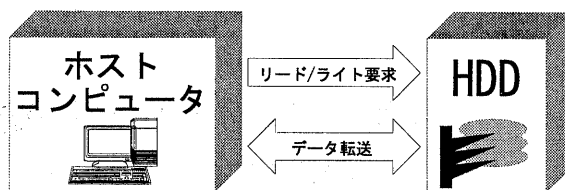


図 1: HDD とホストコンピュータとの関係

そこで我々は、HDD の性能を予測するためのシミュレーション用モデルの開発を行った。HDD は一般的に、ホストコンピュータ (ホスト) からの大量の命令¹に対するスループットが高いほど高性能であるとされるシステムである。また多くのハードウェア (HaradWare, HW) および HW 制御用ソフトウェア (FirmWare, FW) が複雑に影響し合って動作するため、実際の HDD (実機 HDD) をそのままモデル化することは非常に困難である。そこで我々は、HDD が命令を処理し終えるまでの時間 (命令処理時間) を性能評価の最も重要な指標として定め、命令処理時間の計測に最適な詳細度をもつシミュレーション用モデルを設計した。この HDD モデルを用いた様々な条件におけるシミュレーションを行うことによって、性能チューニングによる効果があるか否かを、実機 HDD を用いることなく設計段階で確認することができる。

本発表ではシミュレーション用 HDD モデルと、このモデルを中核とした HDD 性能評価支援システムについて述べる。

2 HDD の概要

2.1 HDD の特徴

HDD は、ホストからの命令を受け取ることによりデータのリード/ライト処理を開始し、ホストとの間でデータの受け渡しを完了すると、再びホストからの命令待ち状態となる。HDD は様々な HW および FW が複雑に影響し合って動作するシステムであり、命令処理時間はホストからの命令系列やシステムの内部状態により大きく変化する。例えば、HDD にはキャッシュと呼ばれる、以前にアクセスしたデータを一時的に保持する機構が組み込まれているが、キャッシュ内部に保持されたデータの状態により処理の分岐が行われる。リード命令で要求されるデータが既にキャッシュ内に存在する (ヒット) 場合は μs オーダーで高速に処理を完了する。そうでない場合 (ミスヒット) は磁気ディスクからデータをリードする処理を行う必要が生じるため、処理速度が ms オーダーまで低下する。キャッシュ内におけるデータの置換アルゴリズム、メモリサイズ、および判定時のキャッシュ内部の状態により、ヒットする確率は様々に変化し [1][2]、命令処理時間にも大きく影響を及ぼす。

2.2 HDD のシステム構成と内部動作

HDD は図1 のようにホストと接続される。内部にはデータを保持するための磁気ディスク、およびデータをリード/ライトするためのヘッドが格納されている。ホストからリード命令が発行されると、ヘッドを通じて磁気ディスクからデータを読み出し、ホストへデータを転送する。逆にライト命令が発行されると、ホストから転送されたデータを、ヘッドを通じて磁気ディスクへ書き込む処理を行う。内部の HW は、FW

¹ 命令は、LBA(Logical Block Address, 論理アドレス)、セクタ数、リード/ライト要求で構成される。

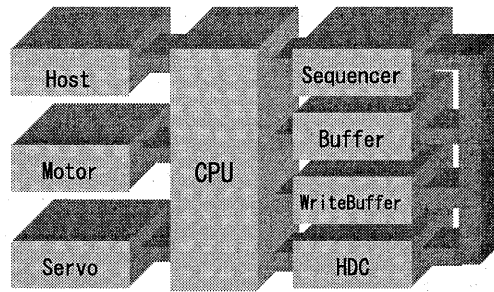


図 2: HDD モデルの HW 構成

により制御される。FW は、幾つかのプロセスおよび HW に付随する割り込みサービスルーチン (Interrupt Service Routine, ISR) で構成される。

3 シミュレーション用 HDD モデル

前節の実機 HDD の動作仕様をふまえ、シミュレーションによって命令処理時間を計測するための HDD モデルを設計した。

3.1 HDD モデルの設計方針

HDD のモデリングにおいて、以下の 5 項目の方針を立てた。

1. 大量の命令列を利用して HDD のスループットを計測するベンチマークテストを対象とする。
2. HDD の命令処理時間の算出に必要な実機の挙動をモデル上で実現する。
3. 実機 HDD の各処理に要する時間を利用し、全体として可能な限り実機 HDD と同等の実行時間をモデルへ組み込む。
4. 実機 HDD の仕様変更への容易な対応、およびモデルの保守/管理が容易に行えるようにする。
5. シミュレーション速度を落さないよう、扱う情報および内部処理は必要最低限とする。

3.2 HDD モデルの構成

HDD モデルの設計は、CARDtools² を用いた。CARDtools は DARTS 法 [4] を応用した表記法に基づいて HW/FW 設計を行うことができ、設計したモデルの動作テストが行えるシミュレータも備えたツールである。実機 HDD の分析結果から、HW 部は図 2 のようにモデル化した。各 HW 部品のモデル上の役割 (内部処理) を表 1 にまとめる。また FW は実機 HDD とほぼ同様の処理を行うプロセス、および ISR から構成される。HW および FW の内部処理の組込みは、次のように行った。

1. HW 処理は、ヘッドの移動、または磁気ディスクの回転など、実際の動作機構をモデルに組み込まず、各機構の動作に必要な処理時間のみを用いることによって、シミュレーション速度を向上させた。各処理時間の算出は次のように行った。

² CARDtools Systems Corp. 製 組込みソフトウェア設計用デザイン/シミュレーションツール [3]

表 1: HDD モデルを構成する HW 部品

Host	ホストの命令発行処理を担当する
Motor	モータの回転処理を担当する
Sequencer	シーケンサ処理 ³ を担当する
Buffer	磁気ディスクとバッファとの間のデータ転送処理を担当する
WriteBuffer	ホストとバッファとの間のデータ転送処理を担当する
HDC	Sequencer, Buffer, WriteBuffer の制御, および状態の管理を担当する
Servo	シーク処理 ⁴ を担当する

- シーク処理時間は、移動シリンダ数毎のシーク時間データをテーブルに持ち、移動シリンダ数に応じてテーブルを参照した
 - シーケンサ処理時間は、実機のデータ転送レートをもとに、ホストから受け取った命令毎に算出した
 - 回転待ち処理時間は、大量の命令列を用いたテストにおいては平均的な時間に収束するものとして考え、平均回転待ち時間を使用した
2. FW 処理は、実機 HDD と同等の順序で同じ内容の処理を行うように組み込み、HW の割込みとの相互作用により動作するように実装した
 3. FW 処理時間は設計段階では不確定な部分が多いため、処理時間がはっきりとしており、且つ命令処理時間への影響が比較的大きなものを以外は、モデルへ組み込まないようにした

本モデルでは、FW 処理は現在の状態から内部処理の実行順序を決定する役目を果し、HW 処理は各内部処理が必要とする処理時間を算出し、割込みを発行することによって FW に内部状態の変化を伝える役目を果している。またアドレス計算などの複雑な計算処理、バッファ内部の状態の管理などは、内部処理が複雑になることを避けるため、直接モデル内に組み込まず、外部関数として持たせている。図3にそれぞれの役割および関係を示す。

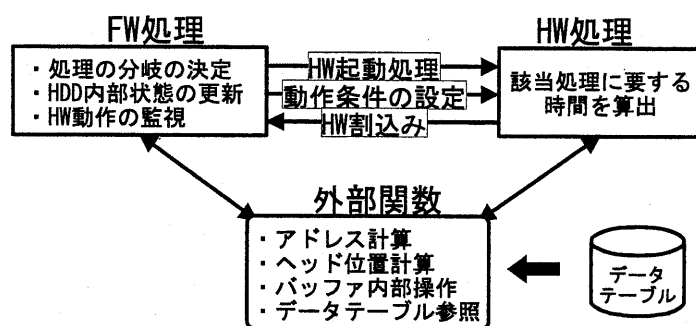


図 3: HW, FW および外部関数の役割とそれぞれの関係

³ヘッドが磁気ディスクに対して、データのリード/ライト動作を行い、バッファとの間でデータ転送を行う処理

⁴ヘッドを、リード/ライトを行うセクタの位置を含むトラックへ移動する処理

3.3 シミュレーション結果

HDD モデルを用いて行ったシミュレーション結果を以下に示す。シミュレーション用入力データは、実機 HDD がホストから受け取る情報と同様に、{ LBA, セクタ数, リード/ライト } のデータ列を抽出したものをを用いた。本シミュレーションは PentiumII 350MHz, 128MB のメモリ, OS に WindowsNT4.0 を搭載したマシンを使用し, 5000 命令を含むデータ列を処理するのに約 50 分要した。

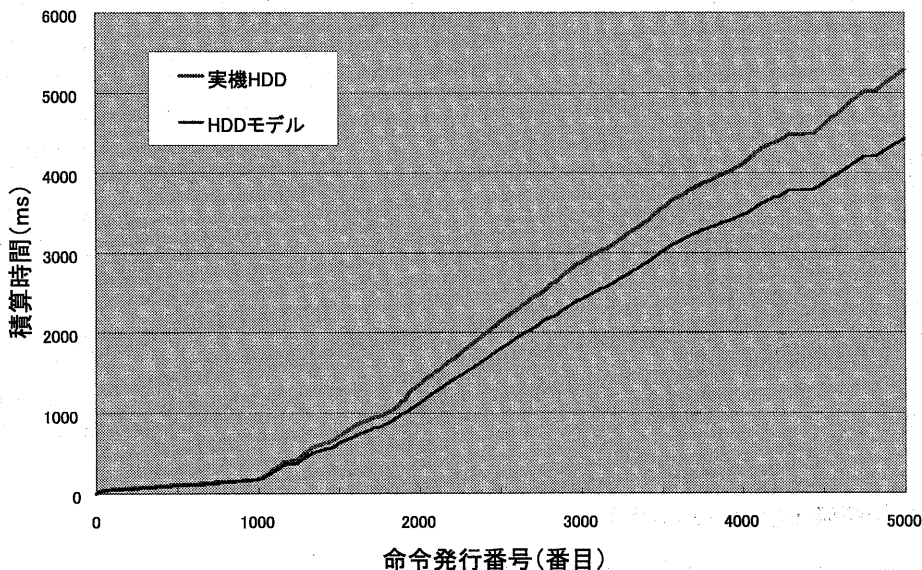


図 4: 実機 HDD と HDD モデルを用いたシミュレーション結果との積算時間グラフによる比較

図4 は, 5000 命令を含む命令列を処理した場合の, 実機 HDD と HDD モデルを用いたシミュレーション結果との比較を, 命令処理時間の積算時間グラフで示したものである。これを見ると, 実機 HDD の処理時間と HDD モデルの処理時間の比較結果は, 定常的な誤差が見られるものの, グラフの傾向がほぼ同じように現れていることが確認できる。

また表2 に, 10 パターンの命令列に対するそれぞれの積算時間の数値比較を示す。表2 における相対誤差とは, 命令列 1 を基準とした場合の各命令列の処理時間の比較である。実機 HDD と HDD モデルの処理時間を比較した結果, 処理時間の絶対値に 15% 程度の誤差が見られた。これはモデルへの組み込み対象外とした FW 処理に要する時間が影響していることが原因の一つとして挙げられる。FW 処理時間を詳細にモデルに組み込むことによって, シミュレーションにより算出される命令処理時間を実機 HDD に限りなく近づけることが可能であるが, 先述した通り, 設計段階で実機 HDD と同等の FW 処理時間をモデルに組み込むことは困難である。また試作機の完成後に正確な FW 処理時間を計測し, モデルに組み込むことは, 本 HDD モデル設計の目的から逸脱したアプローチとなる。一方, 表2 における命令列に対する相対誤差は, 実機 HDD と HDD モデルとの間で, それぞれの命令列が約 1% 前後の差に留まっているという結果を得た。従って, HDD モデルを用いたシミュレーションの結果より, 異なる命令列に対するスループットを相対的に評価することが十分に可能であることが確認できた。

表 2: 実機 HDD と HDD モデルを用いたシミュレーション結果との命令処理時間の比較

	実機 HDD (ms)	HDD モデル (ms)	誤差 (%)	実機 HDD (%)	HDD モデル (%)	相対誤差 (%)
命令列 1	5292.9	4427.0	16.4	命令列 1 を 100%とした場合		
命令列 2	4759.3	4027.4	15.4	89.1	91.0	1.1
命令列 3	5278.0	4501.8	14.7	99.7	100.2	0.5
命令列 4	4909.1	4166.5	15.1	92.8	94.1	1.3
命令列 5	5302.4	4502.6	15.1	101.8	101.7	0.1
命令列 6	4595.5	3882.4	15.5	86.8	87.7	0.9
命令列 7	5217.2	4377.8	16.1	98.6	98.9	0.3
命令列 8	3003.7	2523.2	16.0	56.7	57.0	0.3
命令列 9	3640.2	3085.0	15.3	68.8	69.7	1.1
命令列 10	5153.8	4378.1	15.1	97.4	98.9	1.5

4 HDD 性能評価支援システム

我々は、HDD モデルを用いたシミュレーションを効率良く行い、さらにシミュレーション結果を評価し易い形式で表現するためのツールを開発し、HDD モデルを中心とした HDD の性能評価環境を構築した。

4.1 性能評価支援ツールの概要

性能評価支援ツールは、シミュレータ/ビューア起動用のランチャ、シミュレーション結果から必要なデータを抽出するスクリプト、および抽出したデータを表示するビューアから構成される。図5に性能評価支援ツールの画面構成を示し、以下に各機能について簡単に説明する。

- ランチャ

シミュレータやビューアを起動する。回転数、磁気ディスク上のトラック当りのセクタ数、およびキャッシュ条件など、HDD の仕様を示すパラメータの数値を設計者が設定しておき、自動シミュレーションによって計算結果を蓄積する機能を備える。また複数の異なるパラメータセットに対する連続シミュレーションも可能である。

- データ抽出用スクリプト

シミュレーション結果から、ビューアで使用する性能評価の対象となる数値を抽出し、シミュレーション結果の冗長な部分を削除して、結果の運用/管理を容易にする。

- ビューア

シミュレーション結果を評価し易いグラフなどで表示するビューア機能を備える。ビューアには以下の3つを用意した。

- － 積算時間グラフビューア

各命令毎の処理時間の累積を、折れ線グラフで表示する。異なる条件のシミュレーション結果と比較することにより、全体的な命令処理時間の傾向を把握することができる。

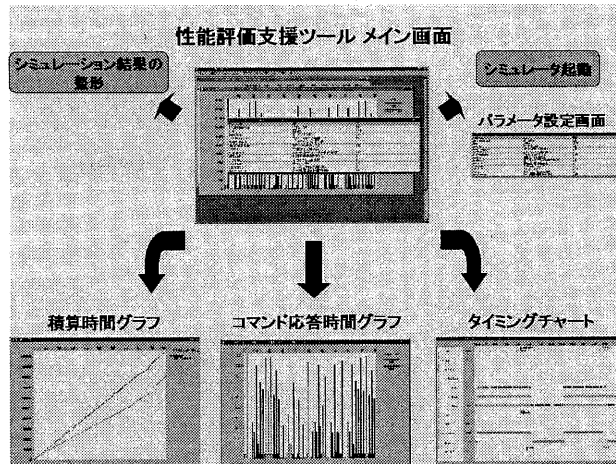


図 5: 性能評価支援ツールの画面構成

－ コマンド応答時間グラフビューア

各命令毎の処理時間を、棒グラフで表示する。棒グラフはリード/ライト命令の処理によって色分けし、キャッシュヒットした場合はマークすることにより、結果を確認し易くしている。

－ タイミングチャートビューア

HDD モデルの内部プロセス、および HW 部品の動作履歴を、時間系列で表示する。各プロセスもしくは HW 部品の内部処理が実行されたタイミングや、処理に要した時間を確認する必要がある場合に使用する。

4.2 HDD 性能評価支援システムの全体構成

図6に、HDDモデルを中心としたHDD性能評価支援システムの全体構成を示す。設計者は性能評価支援ツールからシミュレーション設定、および結果のビューを確認し、通常はHDDモデルを意識することなく本システムを運用することができる。シミュレーションは、HDDの様々な仕様を示す複数の異なるパラメータセットを予め準備し、ランチャを用いて自動的に設計者が必要とする結果を揃えることができる。また3種類のビューアは、(1)積算時間グラフビューアにより全体の傾向を検討し、(2)グラフの傾きから詳しい処理情報が知りたい命令番号付近のコマンド応答時間グラフビューアを確認し、(3)さらに詳細に評価したい命令番号付近の各HW/FWの動作を、タイミングチャートビューアから検証する、という順序でそれぞれを使用することにより、性能評価を進めていく運用方法を想定している。

なおHDDの仕様が大幅に変更する場合は、直接HDDモデルを変更して対応する必要がある。

5 まとめ

我々は、対象とするシステムの性能に関する評価項目を設定し、最適な詳細度を持つシミュレーション用HDDモデルを設計した。本モデルは高速にシミュレーションを行うことができ、且つ実機HDDの仕様変更にも柔軟に対応できるように、HDDの仕様を表す主要な数値はパラメータとして与えることができ、モ

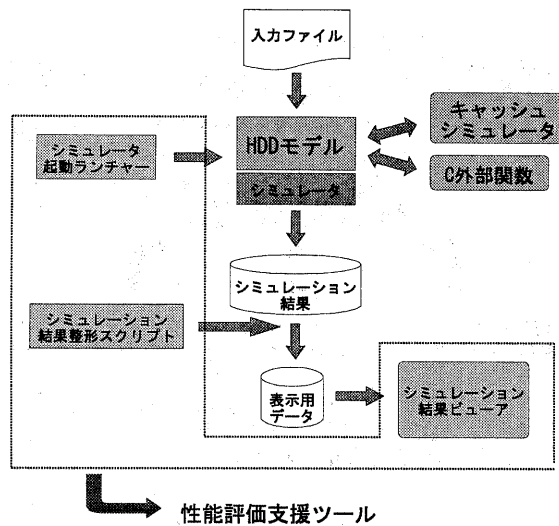


図 6: HDD 性能評価支援システムの構成

デルを直接変更する必要が生じないように設計している。またモデルの詳細度が適度であるため、モデルの保守/管理も容易に行うことができる。

HDD モデルを用いたシミュレーションを通じて、これまでは試作機が完成してから行っていたチューニング作業を、実機 HDD の設計とコンカレントに進めていくことが可能となる。さらに我々は、これらの性能評価作業を円滑に進めていくことができるように性能評価支援ツールを開発し、HDD モデルおよびシミュレータと統合した HDD 性能評価支援システムを構築した。本システムの運用により、HDD の性能評価および性能チューニング情報を、効果的にモデルへフィードバックすることができるため、高性能な HDD を短時間で開発することが可能となる。

今後は、本 HDD 性能評価支援システムを性能評価の手段として、実際の HDD システム設計段階にて試験的に適用し、HDD 性能評価支援システムを評価していくことが課題として挙げられる。また様々な性能に関する評価の要求に対する HDD モデルおよび HDD 性能評価支援システムの対応、および異なる HW/FW 処理に対するパフォーマンスの変化の測定など、実際の運用を通じて設計者からフィードバックされた要求に対して HDD 性能評価支援システムを改善していくことも、課題の一つとして挙げられる。

参考文献

- [1] R.Karedla, J.S.Love, and B.Wherry, "Caching Strategies to Improve Disk Systems Performance", *Computer*, Mar. 1994, pp.38-46.
- [2] A.J.Smith, "Disk-Cache Miss-Ratio Analysis and Design Considerations", *ACM Trans. Computer Systems*, Vol.3, No.3, Aug. 1985, pp.161-203.
- [3] "Task Map Builder User's Guide", *CARDtools Systems Corp.*, 1999.
- [4] Hassan Gomaa, "Software Design methods for Concurrent and Real-Time Systems", *ADDISON-WESLEY PUBLISHING COMPANY*, 1993.