

学 位 論 文 題 名

ディスク・サブシステムの高スループット化に関する研究

学位論文内容の要旨

近年、磁気ディスク装置の記憶容量は大容量化の傾向にあり、記憶データの共有率が高くなっている。ディスク・サブシステムは複数の入出力チャンネルから同時にアクセスされる必要があり、高スループットが望まれる。

機械動作を含む磁気ディスク装置の処理速度は主記憶装置の電子的処理速度に比べて遅い。主記憶装置の処理速度とディスク装置の処理速度の差を緩和し、処理性能を向上させるための緩衝メモリとしてディスクキャッシュが用いられる。ディスク・サブシステムの高スループット化には、ディスクキャッシュのデータ転送幅を上げる必要がある。すでに、マルチポート・ページメモリ（MPPM）と呼ばれるメモリをディスクキャッシュとして用いることが提案されている。MPPMは、複数の入出力ポートをもち、どのポートからも並行にブロック単位でデータの入出力が行えるメモリである。だが、その処理性能については報告されていない。また、MPPMをバッファメモリとして用いた知識ベースマシンの研究もある。MPPMを用いることにより高スループットが得られることが示される。

本論文では、ディスク・サブシステムの高スループット化を目標とし、MPPMをディスク・キャッシュとして階層的に用いたアーキテクチャを提案し、いくつかのキャッシュ動作において性能評価を行い、それらの特性を明らかにする。MPPMをディスクキャッシュとして階層的に用いるといった提案は今までない。その評価性能についての報告もない。

本研究の遂行において性能評価は不可欠である。種々の構成の対象システムについてそれらのモデリングが容易に行え、シミュレーション中の対象モデルの状態が容易に把握でき、対話的にシミュレーションが行えるシミュレータが必要である。このようなシミュレータを対話型グラフィカルシミュレータと呼ぶことにする。このような機能をもつシミュレーションシステムの開発を本研究の初期目的として、対話的グラフィカルシミュレータを開発するツール群を提供するシステムとしてFESを開発した。

VLSI設計用のCADをはじめとして対話型グラフィカルシミュレータは多数開発されてい

る。だが、これらではその対象分野が固定されており、種々の分野をシミュレーション対象としては扱えない。FES は、MV モダリングと呼ぶモデリング手法を採用したことにより、種々の分野をシミュレーション対象として扱える。MV モデリングとは、画面上に表示される形状を定義した視覚表示部（View）と内部動作機構を定義した動作記述部（Model）に分けて基本モデルの定義を行うモデリング手法である。すでに、MVC モデリングと呼ばれるモデリング手法を提供するオブジェクト指向プログラミング言語もあるが、これではシミュレーションのための機能が無い。

本論文では、MV モデリングをはじめとして、FES が提供する種々の機能と特徴を述べ、対話型グラフィカルシミュレータの開発においてそれらがどのように用いられているかを述べる。まず、2 章では、FES シミュレーションシステムの仕様および機能について述べる。MV モデリングについて説明し、FES の特徴であるビジュアル合成機能とビジュアルインスペクタ機能の実現に、それがどのように用いられているかを示す。また、FES を用いて開発されたシミュレータの例を挙げ、FES の有用性を示す。特に、ビジュアル合成による対象モデルの構成の視覚的定義機能と、ビジュアルインスペクタによる対象モデルの状態・動作のグラフィカルな表示機能が、本システムの応用例として開発した 3 つのシミュレータにどう活かされているかを示す。

3 章以下では、本研究の主目的であるディスク・サブシステムの高スループット化について述べる。まず、3 章では、MPPM をディスクキャッシュとして階層的に用いたディスク・サブシステム（DIMP と呼ぶことにする）のアーキテクチャについて述べる。DIMP の構成と基本的なキャッシュ動作について解説する。また、従来のキャッシュ付きディスク・サブシステムの構成と動作についても解説し、そのボトルネックを指摘する。FES を用いて開発されたシミュレータにより性能評価を行った結果から、従来のキャッシュ付きディスク・サブシステムの性能と DIMP の性能を比較する。従来のキャッシュ付きディスク・サブシステムでは、ディスク装置とディスクキャッシュ間のデータ転送幅がボトルネックとなることを示す。DIMP は、チャネルの転送幅によって制限される、あるいは、ディスク装置の処理性能によって制限されるまでの高いスループットが得られることを示す。また、MPPM を階層的に用いたことにより、構成上の自由度が高くシステム拡張が容易であることを述べる。

4 章では、上下層に同数の MPPM を配した二層全接続型構成と呼ぶ DIMP の構成について、キャッシュデータのコヒレンシを考慮した 4 種のキャッシュ動作を述べ、それらの処理性能を示す。キャッシュデータの置き換え方式として、上下層の MPPM それぞれについてライトバック（WB）方式とライトスルー（WT）方式を採用でき、4 種（WB/WB, WT/WT, WT/WB,

WB/WT)のキャッシュ動作が考えられる。解析的に得た処理性能により、限界スループットと平均応答時間を比較する。限界スループットについては、WB/WB, WT/WB, WB/WT, WT/WTの順番で性能が良くなることを示す。低トランザクション域での平均応答時間については、WB/WT, WT/WT, WT/WB, WB/WBの順番で性能が良くなることを示す。

5章では、ディスク・サブシステムの信頼性を向上させる構成であるディスクアレイに関して、DIMPの処理性能を示す。ディスク装置1台あたりのスループットについて、DIMPの性能と従来のディスク・サブシステムにおけるディスクアレイの性能とを比較する。DIMPは従来のディスク・サブシステムの5倍程度のスループットが得られることを示す。

最後に、6章で、これまでの章を総括し、本研究の成果について要約し、今後に残された課題について述べる。

本研究は、ディスク・サブシステムの高スループット化を目的とし、ディスクキャッシュとしてMPPM階層的に用いることを提案している。種々の構成についてキャッシュ動作を考察し、それらの評価性能を示し特徴を明らかにした。本研究の遂行上不可欠である性能評価のために、FESシミュレーションシステムの開発を本研究の初期目的とした。MVモデリングと呼ぶモデリング手法を採用し、種々の機能をもつ視覚表示部を提供したことが、対話型グラフィカルシミュレータの開発において、どのように有益であったかを示した。

学位論文審査の要旨

主査	教授	田中	譲
副査	教授	宮本	衛市
副査	教授	青木	由直
副査	教授	栃内	香次

近年、磁気ディスク装置単体の記憶容量が著しく増大しつつある。これに対してディスク装置単体の入出力の動作速度と転送速度に関しては著しい改善が望めない。このため、多数のディスク装置単体を用いて構成されるディスク・サブシステムの全体システムとしての入出力アクセスのスループットをいかに高めるかが重要な研究開発課題となっている。

ディスク・サブシステムの入出力アクセスの高スループット化に関しては、従来、大きく分け

て、ディスク・キャッシュの採用、ディスク装置内蔵バッファメモリの採用、ディスク装置からディスクキャッシュへの転送を過去のアクセス履歴によって最適化する方法の3種類の方法が提案されている。ディスク・キャッシュの採用は、ディスク装置へのアクセス頻度を下げ、キャッシュのヒット率が高い場合には顕著な性能の向上を期待することができる。しかし、そのような場合においてもキャッシュメモリの入出力アクセスのスループットがボトルネックとなり、ディスク装置の限界アクセス・スループットは達成することができない。さらに、ヒット率の低い場合にはディスク・キャッシュの効果は望めない。ディスク装置内蔵のバッファメモリの採用は転送パスがビジー状態の場合にも、一旦バッファにデータを保持することによりディスクの回転待ちを必要としないため、RPS ミスを低減することができるが、著しい性能向上を目指す手法ではない。アクセス履歴によって転送を最適化する方法はディスク・キャッシュと併用する方法で、ディスク・キャッシュ自体の性能によって制限される。

本論文は、ディスク・キャッシュとして用いるメモリ装置のアクセス・スループットを著しく改善することにより、ディスク装置とディスク・キャッシュの間と、ディスク・キャッシュと処理装置群との間の各々の転送パスの容量を著しく高めるとともに、ディスク装置内蔵バッファメモリを使用した場合と同様に RPS ミスも低減させることを目指した研究である。具体的には、同時にブロック単位のアクセスが可能な複数の入出力ポートを持つマルチポート・ページメモリ（MPPM）を2段に階層化してディスクキャッシュとして用いたディスク・サブシステムのアーキテクチャを初めて提案し、いくつかのキャッシュ・プロトコルの性能比較を行い、最適なプロトコルを明らかにした。その処理性能について解析的手法と計算機シミュレーションによる詳細な評価を与え、ディスク装置単体のスループットにその台数を乗じて得られる限界性能との比較を行い、従来は達成することが難しいとされていた限界性能が提案方式によって初めてほぼ達成できることを示した。

その主要な成果は次の3点に要約される。

- (1) ディスク・サブシステムの詳細は性能評価のための視覚的シミュレータを開発した。本シミュレータは数十台のディスク装置と16ポートの MPPM 16台を含むディスク・サブシステムのシミュレーションを個々のディスクと MPPM の詳細な内部動作を正確に追いつながらシミュレーションすることができるシステムで、ビジュアルにモデリングができる。
- (2) マルチポート・ページメモリをディスク・キャッシュとして階層的に用いたディスク・サブシステムのアーキテクチャを初めて提案した。
- (3) 提案アーキテクチャにおける種々のキャッシュ・プロトコルの性能評価を行い、WT/WB

(ライト・スルー／ライト・バック) プロトコルが最適であることを示し、その際のディスク・サブシステムの性能が、ディスク装置単体のスループットにその台数を乗じて得られる限界性能をほぼ達成することを明らかにした。

これを要するに、著者は、マルチポート・ページメモリをディスクキャッシュとして階層的に用いることを提案し、その性能評価を行うことにより、ディスク・サブシステムの著しい高スループット化に有益を知見を与えている。この成果は計算機工学の進歩に寄与するところ大である。よって著者は、博士（工学）の学位を授与される資格あるものと認める。