

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2019-164659
(P2019-164659A)

(43) 公開日 令和1年9月26日(2019.9.26)

(51) Int. Cl.	F I	テーマコード (参考)
G06F 12/00 (2006.01)	G06F 12/00 560A	5B042
G06F 11/34 (2006.01)	G06F 11/34 147	5B160
G06F 12/06 (2006.01)	G06F 12/06 515K	
G06F 11/30 (2006.01)	G06F 11/30 140N	

審査請求 未請求 請求項の数 17 O L (全 39 頁)

(21) 出願番号 特願2018-52719 (P2018-52719)
(22) 出願日 平成30年3月20日 (2018.3.20)

(出願人による申告) 平成28年度 国立研究開発法人新エネルギー・産業技術総合開発機構「IoT推進のための横断技術開発プロジェクト/高速大容量ストレージデバイス・システムの研究開発」委託研究、産業技術力強化法第19条の適用を受ける特許出願

(71) 出願人 000003078
株式会社東芝
東京都港区芝浦一丁目1番1号
(74) 代理人 110002147
特許業務法人酒井国際特許事務所
(72) 発明者 城田 祐介
東京都港区芝浦一丁目1番1号 株式会社東芝内
(72) 発明者 肥塚 真由子
東京都港区芝浦一丁目1番1号 株式会社東芝内
(72) 発明者 金井 達徳
東京都港区芝浦一丁目1番1号 株式会社東芝内

最終頁に続く

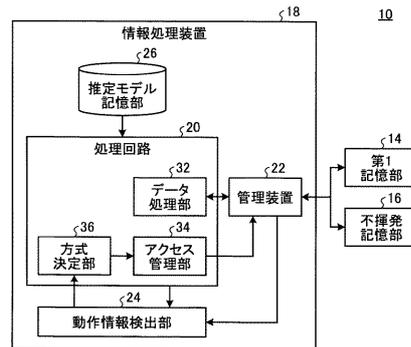
(54) 【発明の名称】 情報処理装置、情報処理方法およびプログラム

(57) 【要約】

【課題】 状況に応じた適切なメモリ制御方式で第1記憶部および不揮発記憶部にアクセスする。

【解決手段】 実施形態に係る情報処理装置は、処理回路と、第1記憶部と、不揮発記憶部と、管理装置とを備える。処理回路は、アクセス管理部と、取得部と、推定部と、選択部と、設定部とを有する。取得部は、動作情報を取得する。推定部は、取得した動作情報に基づき、管理装置が複数のメモリ制御方式のそれぞれにより動作する場合における、第1記憶部および不揮発記憶部に対するメモリアクセスの実行性能を推定する。選択部は、推定された複数のメモリ制御方式のそれぞれについての実行性能に基づき、複数のメモリ制御方式のうちの何れか1つのメモリ制御方式を選択する。設定部は、選択されたメモリ制御方式により管理装置が第1記憶部および不揮発記憶部にアクセスするようにアクセス管理部に対して設定処理をする。

【選択図】 図1



【特許請求の範囲】**【請求項 1】**

処理回路と、
第 1 記憶部と、
不揮発記憶部と、

前記処理回路から書き込みまたは読み出しの要求を受けたことに応じて、前記第 1 記憶部および前記不揮発記憶部にアクセスする管理装置と、

を備え、

前記処理回路は、

前記管理装置が前記第 1 記憶部および前記不揮発記憶部に対してアクセスするためのメモリ制御方式を管理するアクセス管理部と、

前記処理回路および前記管理装置の少なくとも一方における動作情報を取得する取得部と、

取得した前記動作情報に基づき、前記管理装置が複数のメモリ制御方式のそれぞれにより動作する場合における、前記第 1 記憶部および前記不揮発記憶部に対するメモリアクセスの実行性能を推定する推定部と、

推定された前記複数のメモリ制御方式のそれぞれについての前記実行性能に基づき、前記複数のメモリ制御方式のうちの何れか 1 つのメモリ制御方式を選択する選択部と、

選択されたメモリ制御方式により前記管理装置が前記第 1 記憶部および前記不揮発記憶部にアクセスするように前記アクセス管理部に対して設定処理をする設定部と、

を有する情報処理装置。

【請求項 2】

前記取得部は、検出期間における前記動作情報を取得し、

前記推定部は、前記検出期間において取得した前記動作情報に基づき、前記実行性能を推定する

請求項 1 に記載の情報処理装置。

【請求項 3】

前記処理回路は、一定期間毎に前記検出期間を開始する指示を前記取得部に与えるタイミング発生部をさら有する

請求項 2 に記載の情報処理装置。

【請求項 4】

前記処理回路は、動作中において、前記動作情報を取得し、取得した前記動作情報が所定の変化をした場合に、前記検出期間を開始する指示を前記取得部に与える変化検出部をさらに有する

請求項 2 に記載の情報処理装置。

【請求項 5】

前記設定部は、前記検出期間において、前記管理装置が基準のメモリ制御方式により前記第 1 記憶部および前記不揮発記憶部にアクセスするように設定処理をする

請求項 2 から 4 の何れか 1 項に記載の情報処理装置。

【請求項 6】

前記不揮発記憶部は、複数のページを含み、

前記アクセス管理部は、前記複数のページのそれぞれについて、前記不揮発記憶部から前記第 1 記憶部に転送したデータに対して書き込みおよび読み出しをする第 1 アクセス処理、または、前記不揮発記憶部に記憶されたデータに対して直接書き込みおよび読み出しをする第 2 アクセス処理の何れを実行するかを示すアクセス方法を示す変換テーブルを記憶し、

前記管理装置は、前記第 1 アクセス処理に設定されているページに対して書き込みまたは読み出しの要求を受けた場合、前記第 1 アクセス処理を実行し、前記第 2 アクセス処理に設定されているページに対して書き込みまたは読み出しの要求を受けた場合、前記第 2 アクセス処理を実行し、

10

20

30

40

50

前記設定部は、前記変換テーブルにおける前記複数のページのそれぞれのアクセス方法を、選択されたメモリ制御方式に応じて前記第1アクセス処理または前記第2アクセス処理に設定する

請求項2から5の何れか1項に記載の情報処理装置。

【請求項7】

前記複数のメモリ制御方式は、前記複数のページのうちの前記第1アクセス処理に設定されるページの数と前記第2アクセス処理に設定されるページの数との比率により区別され、

前記設定部は、前記変換テーブルにおける前記複数のページのそれぞれのアクセス方法を、選択されたメモリ制御方式の比率で前記第1アクセス処理または前記第2アクセス処理を設定する

10

請求項6に記載の情報処理装置。

【請求項8】

前記複数のメモリ制御方式は、さらに、前記第1記憶部のメモリ使用量の上限値で区別され、

前記設定部は、選択されたメモリ制御方式に示された前記上限値を、前記第1記憶部のメモリ使用量として前記管理装置に設定し、

前記管理装置は、前記第1アクセス処理に設定された第1ページにアクセスするために前記第1ページのデータを前記不揮発記憶部から前記第1記憶部に転送すると、前記第1記憶部のメモリ使用量が前記上限値を超えてしまう場合には、前記第1記憶部に記憶された何れかのページのデータを前記不揮発記憶部に書き戻した後に、前記第1ページのデータを前記不揮発記憶部から前記第1記憶部に転送する

20

請求項7に記載の情報処理装置。

【請求項9】

前記設定部は、前記検出期間において、アクセスするページの全てのアクセス方法を前記第2アクセス処理に設定する

請求項6から8の何れか1項に記載の情報処理装置。

【請求項10】

前記推定部は、前記実行性能として、前記管理装置が前記検出期間において実行した処理を、前記管理装置が対応するメモリ制御方式で実行した場合における実行時間を推定する

30

請求項2から9の何れか1項に記載の情報処理装置。

【請求項11】

前記取得部は、前記動作情報として、メモリアクセス回数およびアクセスしたページ数とを取得し、

前記推定部は、前記メモリアクセス回数および前記ページ数に基づき、前記実行時間を推定する

請求項10に記載の情報処理装置。

【請求項12】

前記選択部は、前記実行性能が予め定められた条件を満たすメモリ制御方式のうち、前記実行性能が最も小さいまたは最も大きいメモリ制御方式を選択する

40

請求項1から11の何れか1項に記載の情報処理装置。

【請求項13】

前記推定部は、

前記複数のメモリ制御方式のそれぞれについての前記実行性能として、前記検出期間において前記管理装置が実行した処理を、前記管理装置が対応するメモリ制御方式で実行した場合における実行時間を推定し、

前記選択部は、

前記管理装置が前記複数のページの全てを前記第1アクセス処理に設定するメモリ制御方式で実行した場合における実行時間より所定割合長い基準時間を算出し、

50

推定した実行時間が前記基準時間より短いメモリ制御方式のうちの、推定した実行時間が最も長いメモリ制御方式を選択する

請求項 6 から 9 の何れか 1 項に記載の情報処理装置。

【請求項 1 4】

前記処理回路は、

前記動作情報の変化内容と選択されたメモリ制御方式との関係に基づき、前記動作情報の変化内容からメモリ制御方式を決定するための決定モデルを生成する学習部と、

前記取得部から取得した前記動作情報の変化内容および前記決定モデルに基づき、メモリ制御方式を決定する決定部と、

をさらに有し、

前記実行性能を推定せずにメモリ制御方式を決定するモードが設定された場合、

前記変化検出部は、前記検出期間を開始する指示の出力を停止し、

前記設定部は、前記決定部により決定されたメモリ制御方式により前記管理装置が前記第 1 記憶部および前記不揮発記憶部にアクセスするように、前記アクセス管理部に対して設定処理をする

請求項 4 に記載の情報処理装置。

【請求項 1 5】

前記処理回路は、時系列に選択された複数のメモリ制御方式を、メモリ制御方式の切り替えにより生じるオーバヘッド実行性能を加えた合計実行性能が最適となるように修正する修正部と、

前記動作情報の変化パターンと修正後におけるメモリ制御方式との関係に基づき、前記動作情報の変化パターンからメモリ制御方式を決定するための決定モデルを生成する学習部と、

前記取得部から取得した前記動作情報の変化パターンおよび前記決定モデルに基づき、メモリ制御方式を決定する決定部と、

をさらに有し、

前記実行性能を推定せずにメモリ制御方式を決定するモードが設定された場合、

前記変化検出部は、前記検出期間を開始する指示の出力を停止し、

前記設定部は、前記決定部により決定されたメモリ制御方式により前記管理装置が前記第 1 記憶部および前記不揮発記憶部にアクセスするように、前記アクセス管理部に対して設定処理をする

請求項 4 に記載の情報処理装置。

【請求項 1 6】

処理回路と、第 1 記憶部と、不揮発記憶部と、前記処理回路から書き込みまたは読み出しの要求を受けたことに応じて、前記第 1 記憶部および前記不揮発記憶部にアクセスする管理装置と、を備える情報処理装置において実行される情報処理方法であって、

前記処理回路は、前記管理装置が前記第 1 記憶部および前記不揮発記憶部に対してアクセスするためのメモリ制御方式を管理するアクセス管理部を有し、

前記処理回路が、

前記処理回路および前記管理装置の少なくとも一方における動作情報を取得する取得ステップと、

取得した前記動作情報に基づき、前記管理装置が複数のメモリ制御方式のそれぞれにより動作する場合における、前記第 1 記憶部および前記不揮発記憶部に対するメモリアクセスの実行性能を推定する推定ステップと、

推定された前記複数のメモリ制御方式のそれぞれについての前記実行性能に基づき、前記複数のメモリ制御方式のうちの何れか 1 つのメモリ制御方式を選択する選択ステップと、

選択されたメモリ制御方式により前記管理装置が前記第 1 記憶部および前記不揮発記憶部にアクセスするように前記アクセス管理部に対して設定処理をする設定ステップと、

を実行する情報処理方法。

【請求項 17】

処理回路と、第1記憶部と、不揮発記憶部と、前記処理回路から書き込みまたは読み出しの要求を受けたことに応じて、前記第1記憶部および前記不揮発記憶部にアクセスする管理装置と、を備える情報処理装置により実行されるためのプログラムであって、

前記処理回路は、前記管理装置が前記第1記憶部および前記不揮発記憶部に対してアクセスするためのメモリ制御方式を管理するアクセス管理部を有し、

前記情報処理装置に、

前記処理回路および前記管理装置の少なくとも一方における動作情報を取得する取得ステップと、

取得した前記動作情報に基づき、前記管理装置が複数のメモリ制御方式のそれぞれにより動作する場合における、前記第1記憶部および前記不揮発記憶部に対するメモリアクセスの実行性能を推定する推定ステップと、

推定された前記複数のメモリ制御方式のそれぞれについての前記実行性能に基づき、前記複数のメモリ制御方式のうちの何れか1つのメモリ制御方式を選択する選択ステップと

、
選択されたメモリ制御方式により前記管理装置が前記第1記憶部および前記不揮発記憶部にアクセスするように前記アクセス管理部に対して設定処理をする設定ステップと、

を実行させるプログラム。

【発明の詳細な説明】**【技術分野】****【0001】**

本発明の実施形態は、情報処理装置、情報処理方法およびプログラムに関する。

【背景技術】**【0002】**

近年、ストレージクラスメモリと呼ばれる高速な不揮発メモリが開発されている。ストレージクラスメモリは、従来のブロック単位やページ単位でのデータの読み書きに加えて、ページよりも小さい例えばバイト単位でのデータの読み書きが可能 (byte-addressable: バイトアドレスサブル) な不揮発メモリである。このようにバイト単位でデータの読み書き可能となった不揮発メモリは、CPU (Central Processing Unit) の作業領域として機能する主記憶装置として用いることができる。

【0003】

また、近年、コンピュータシステムは、オンラインリアルタイム処理、ビッグデータ処理、ディープラーニング処理、AI (人工知能) 処理等の大規模データ処理を実行する。このような処理を実行する場合、コンピュータシステムは、巨大な容量の主記憶装置を用いなければならない。

【0004】

また、データセンタ、ロボット、ドローン、または、エッジサーバ等のエッジシステム等では、高速なデータ処理を実現するために、大規模データを主記憶装置上で処理するインメモリデータ処理を実行する。このような、インメモリデータ処理を実行するコンピュータシステムは、巨大容量の主記憶装置を用いなければならない。

【0005】

このような巨大容量の主記憶装置を全てDRAM (Dynamic Random Access Memory) で実現した場合、待機電力の大きさが課題となる。このため、高速かつ省電力なコンピュータシステムを実現するためには、低待機電力で高速なストレージクラスメモリを活用するのが望ましい。

【0006】

もっとも、ストレージクラスメモリは、DRAMよりも高集積化が可能であるものの、一般にはDRAMよりもアクセス速度が遅い。従って、全てのDRAMをストレージクラスメモリに置き換えたコンピュータシステムは、処理速度が低下する。また、一般に、ス

10

20

30

40

50

トレージクラスメモリは、待機電力は低い、アクセス時の動的電力は高い。従って、巨大容量の主記憶装置を用いるコンピュータシステムでは、DRAM等の高速な記憶装置と、ストレージクラスメモリ等の待機電力が低く高集積化が可能な不揮発記憶装置とを組み合わせ、状況に応じて適切なメモリ制御方式で主記憶装置にアクセスすることが好ましい。

【0007】

しかしながら、コンピュータシステムでは、状況に応じて適切なメモリ制御方式を選択することは非常に困難であった。

【先行技術文献】

【非特許文献】

【0008】

【非特許文献1】R. F. Freitas and W. W. Wilcke, "Storage-class Memory: The Next Storage System Technology", IBM Journal of Research and Development Vol.52 No.4, pp.439-447, 2008.

【発明の概要】

【発明が解決しようとする課題】

【0009】

本発明が解決しようとする課題は、状況に応じた適切なメモリ制御方式で第1記憶部および不揮発記憶部にアクセスすることにある。

【課題を解決するための手段】

【0010】

実施形態に係る情報処理装置は、処理回路と、第1記憶部と、不揮発記憶部と、管理装置とを備える。前記管理装置は、前記処理回路から書き込みまたは読み出しの要求を受けたことに応じて、前記第1記憶部および前記不揮発記憶部にアクセスする。前記処理回路は、アクセス管理部と、取得部と、推定部と、選択部と、設定部とを有する。前記アクセス管理部は、前記管理装置が前記第1記憶部および前記不揮発記憶部に対してアクセスするためのメモリ制御方式を管理する。前記取得部は、前記処理回路および前記管理装置の少なくとも一方における動作情報を取得する。前記推定部は、取得した前記動作情報に基づき、前記管理装置が複数のメモリ制御方式のそれぞれにより動作する場合における、前記第1記憶部および前記不揮発記憶部に対するメモリアクセスの実行性能を推定する。前記選択部は、推定された前記複数のメモリ制御方式のそれぞれについての前記実行性能に基づき、前記複数のメモリ制御方式のうち何れか1つのメモリ制御方式を選択する。前記設定部は、選択されたメモリ制御方式により前記管理装置が前記第1記憶部および前記不揮発記憶部にアクセスするように前記アクセス管理部に対して設定処理をする。

【図面の簡単な説明】

【0011】

【図1】第1実施形態に係る情報処理システムの構成を示す図。

【図2】変換テーブルの一例を示す図。

【図3】第1アクセス処理および第2アクセス処理の内容を示す図。

【図4】動作情報の時系列データの一例を示す図。

【図5】メモリ制御方式の内容の第1例を示す図。

【図6】メモリ制御方式の内容の第2例を示す図。

【図7】第1実施形態に係る方式決定部の構成を推定モデル記憶部とともに示す図。

【図8】動作情報の波形図および選択されたメモリ制御方式の一例を示す図。

【図9】第1実施形態に係る方式決定部による処理の流れを示すフローチャート。

【図10】実行時間の算出処理の第1例を説明するための図。

【図11】実行時間の算出処理の第2例を説明するための図。

【図12】第1アクセス処理に設定されたページの比率に対する実行性能を示す図。

【図13】メモリ制御方式の選択方法の第1例を説明するための図。

【図14】メモリ制御方式の選択方法の第2例を説明するための図。

- 【図15】メモリ制御方式の選択方法の第3例を説明するための図。
- 【図16】第2実施形態に係る方式決定部の構成を示す図。
- 【図17】動作情報の波形図および選択されたメモリ制御方式の一例を示す図。
- 【図18】第3実施形態に係る情報処理システムの構成を示す図。
- 【図19】第3実施形態に係る方式決定部の構成を示す図。
- 【図20】動作情報の変化、実行性能および決定されたメモリ制御方式を示す図。
- 【図21】第1情報処理システムおよび第2情報処理システムを示す図。
- 【図22】第4実施形態に係る推定システムの構成を示す図。
- 【図23】第4実施形態に係る推定装置の構成を示す図。
- 【図24】第4実施形態に係る推定装置による処理の流れを示すフローチャート。 10
- 【図25】動作情報および部分情報毎に決定されたメモリ制御方式を示す図。
- 【図26】アプリケーションプログラム毎の処理時間の推定値の一例を示す図。
- 【図27】不揮発記憶部の種類毎の処理時間の推定値の一例を示す図。
- 【図28】第5実施形態に係る推定装置の構成を示す図。
- 【図29】オーバヘッド時間を考慮した場合の実行時間の一例を示す図。
- 【図30】選択後のメモリ制御方式および修正後のメモリ制御方式を示す図。
- 【図31】第6実施形態に係る方式決定部の構成を示す図。
- 【図32】選択されたメモリ制御方式および修正後のメモリ制御方式を示す図。
- 【図33】修正後のメモリ制御方式と動作情報の変化のパターンとの関係を示す図。 20
- 【図34】情報処理装置のハードウェアブロック図。
- 【発明を実施するための形態】

【0012】

以下、図面を参照しながら実施形態に係る情報処理システム10について詳細に説明する。なお、以下、複数の実施形態について説明をするが、略同一の機能および構成を有するブロックに同一の符号を付け、第2実施形態以降においては重複する内容の説明を省略する。

【0013】

(第1実施形態)

第1実施形態について説明する。

【0014】

図1は、第1実施形態に係る情報処理システム10の構成の一例を示す図である。情報処理システム10は、第1記憶部14と、不揮発記憶部16と、情報処理装置18とを備える。

【0015】

第1記憶部14は、情報処理装置18による作業領域として用いられる主記憶装置(メインメモリ)である。第1記憶部14は、例えば、電源の供給を停止すると記憶しているデータが消える揮発記憶部である。第1記憶部14は、例えば、DRAMである。なお、第1記憶部14は、DRAMと同様に高速アクセスが可能な、MRAM(Magnetoresistive Random Access Memory)等の不揮発メモリであってもよい。

【0016】

また、第1記憶部14は、不揮発記憶部16よりも書き込み可能回数が多。例えば、第1記憶部14は、書き込み可能回数を考慮した設計をしなくてよい程度(例えば、書き込み回数に制限が無いとして設計をすることが可能な程度)、書き込み可能回数が多。

【0017】

不揮発記憶部16は、電源の供給を停止してもデータを記憶し続けるメモリである。不揮発記憶部16は、第1記憶部14とともに情報処理装置18の主記憶装置として機能する。

【0018】

不揮発記憶部16は、例えば、DRAMより大容量な大容量高速不揮発メモリ(Non 50

- volatile Memory)を含む。不揮発記憶部16は、例えば、MRAM、PCM(Phase Change Memory)、PRAM(Phase Random Access Memory)、PCRAM(Phase Change Random Access Memory)、ReRAM(Resistance Change Random Access Memory)、FeRAM(Ferroelectric Random Access Memory)、3DXPointまたはMemristor等のいわゆるストレージクラスメモリ(SCM)と呼ばれるメモリである。また、不揮発記憶部16は、複数の半導体装置を1つの基板または筐体等に設けたモジュールであってもよい。

【0019】

10

不揮発記憶部16は、第1記憶部14と比較して容量が大きい。不揮発記憶部16の容量は、第1記憶部14と同じであってもよい。また、不揮発記憶部16は、第1記憶部14のアクセス速度と同等または少し遅い程度である。また、不揮発記憶部16は、待機電力がゼロか、または、第1記憶部14と比較して待機電力が非常に少ない。不揮発記憶部16は、一例として、アクセスレイテンシが10n秒から数μ秒程度の間のメモリである。

【0020】

不揮発記憶部16は、バイト単位等の小領域単位でデータを書き込みおよび読み出し可能である。従って、情報処理装置18は、ロード命令またはストア命令により不揮発記憶部16に対して直接アクセスが可能である。情報処理装置18は、例えばキャッシュライン単位等で不揮発記憶部16に直接アクセスすることが可能である。

20

【0021】

不揮発記憶部16は、複数のページを含む。ページは、情報処理装置18が管理するデータ単位に対応する。ページは、例えば、オペレーティングシステムが管理する仮想記憶装置のページに対応する。ページは、例えば、4Kバイトであってもよい。情報処理装置18は、第1記憶部14と不揮発記憶部16の間ではページ単位でもデータ転送が可能である。

【0022】

情報処理装置18は、処理回路20と、管理装置22と、動作情報検出部24と、推定モデル記憶部26とを有する。

30

【0023】

処理回路20は、1または複数のプロセッサを有する。プロセッサは、例えば、CPU(Central Processing Unit)である。プロセッサは、1または複数のCPUコアを含んでいてもよい。処理回路20は、プログラムを実行して、データを処理する。処理回路20は、プログラムの実行に応じて、第1記憶部14または不揮発記憶部16からデータを読み出したり、第1記憶部14または不揮発記憶部16にデータを書き込んだりする。

【0024】

また、処理回路20は、L1データキャッシュ、L1命令キャッシュ、L2キャッシュおよびL3キャッシュ等の階層的なキャッシュメモリを有する。処理回路20は、このようなキャッシュメモリを用いて、第1記憶部14または不揮発記憶部16に記憶されているデータを一時的に記憶する。処理回路20は、例えば、階層的なキャッシュにおける最下層のキャッシュ(ラストレベルキャッシュ)でキャッシュミスをした場合、キャッシュライン単位で第1記憶部14または不揮発記憶部16に対してアクセスして必要なデータを読み出したり書き込んだりする。

40

【0025】

なお、処理回路20は、データ処理を実行できれば、どのような回路であってもよい。例えば、処理回路20は、GPGPU(General-purpose computing on Graphics Processing Unit)で利用されるGPU(Graphics Processing Unit)であってもよい。また、処理

50

回路 20 は、FPGA (Field Programmable Gate Array) 等のアクセラレータであってもよい。

【0026】

管理装置 22 は、第 1 記憶部 14 および不揮発記憶部 16 に対する、処理回路 20 によるデータの読み出しおよび書き込みを制御する。管理装置 22 は、処理回路 20 からの第 1 記憶部 14 および不揮発記憶部 16 に対するアクセス要求を処理する。すなわち、管理装置 22 は、処理回路 20 からの書込命令に応じて、データを第 1 記憶部 14 または不揮発記憶部 16 に書き込む。また、管理装置 22 は、処理回路 20 からの読出命令に応じて、データを第 1 記憶部 14 または不揮発記憶部 16 から読み出し、読み出したデータを処理回路 20 に与える。

10

【0027】

また、管理装置 22 は、不揮発記憶部 16 に対するウェアレベリング制御を実行する。より具体的には、管理装置 22 は、不揮発記憶部 16 における領域毎のデータの書き換え回数を管理する。そして、管理装置 22 は、領域毎の書き換え回数に基づき、不揮発記憶部 16 における全体の領域に平均的にデータが書き込まれるように、データの書き込み位置を制御する。

【0028】

なお、管理装置 22 は、処理回路 20 と同一の半導体基板に形成された回路である。また、管理装置 22 は、処理回路 20 とは別体のハードウェアで構成されたメモリコントローラであってもよい。また、管理装置 22 は、処理回路 20 の一部のハードウェアとメモリコントローラとの組み合わせにより実現されてもよい。また、管理装置 22 は、処理回路 20 が実行するオペレーティングシステムの一部の機能により実現されてもよい。また、管理装置 22 は、処理回路 20 が実行するオペレーティングシステムの一部の機能により実現されてもよいし、オペレーティングシステムの一部の機能とメモリコントローラとの組み合わせにより実現されてもよい。

20

【0029】

また、管理装置 22 は、例えば、処理回路 20 とは別体のハードウェアで構成されたメモリ管理ユニット (MMU) であってもよい。また、管理装置 22 は、処理回路 20 の一部のハードウェアとメモリ管理ユニットとの組み合わせにより実現されてもよい。また、管理装置 22 は、処理回路 20 が実行するオペレーティングシステムの一部の機能とメモリ管理ユニットとの組み合わせにより実現されてもよい。

30

【0030】

また、管理装置 22 は、メモリコントローラと、MMU との組み合わせにより実現されてもよい。また、管理装置 22 は、処理回路 20 の一部のハードウェアと、メモリコントローラと、メモリ管理ユニットとの組み合わせにより実現されてもよい。また、管理装置 22 は、処理回路 20 が実行するオペレーティングシステムの一部の機能と、メモリコントローラと、メモリ管理ユニットとの組み合わせにより実現されてもよい。

【0031】

動作情報検出部 24 は、処理回路 20 および管理装置 22 の少なくとも一方における動作情報を検出する。動作情報検出部 24 は、処理回路 20 内のハードウェア回路として実現されてもよい。また、動作情報検出部 24 は、管理装置 22 内の一部のハードウェア回路または一部の機能として実現されてもよい。また、動作情報検出部 24 は、処理回路 20 が実行するオペレーティングシステムの一部の機能として実現されてもよい。動作情報検出部 24 は、いわゆるパフォーマンスカウンタと呼ばれる回路であってもよい。なお、動作情報については、図 4 を参照してさらに説明する。

40

【0032】

推定モデル記憶部 26 は、動作情報から、管理装置 22 による第 1 記憶部 14 および不揮発記憶部 16 に対するメモリアクセスの実行性能を推定するための推定モデルを記憶する。なお、推定モデルについては、図 10 および図 11 を参照してさらに説明する。

【0033】

50

処理回路 20 は、データ処理部 32 と、アクセス管理部 34 と、方式決定部 36 とを含む。データ処理部 32 は、例えば、処理回路 20 がアプリケーションプログラムを実行することにより実現される。また、方式決定部 36 およびアクセス管理部 34 は、例えば、処理回路 20 がオペレーティングシステムの一部の機能を実行することにより実現される。

【0034】

データ処理部 32 は、アプリケーションプログラムに従った情報処理を実行する。さらに、データ処理部 32 は、管理装置 22 に対して、第 1 記憶部 14 および不揮発記憶部 16 に対するアクセス要求を与える。すなわち、データ処理部 32 は、データを第 1 記憶部 14 または不揮発記憶部 16 に書き込む書込命令を管理装置 22 に与える。また、データ処理部 32 は、データを第 1 記憶部 14 または不揮発記憶部 16 から読み出す読出命令を管理装置 22 に与える。

10

【0035】

アクセス管理部 34 は、管理装置 22 が第 1 記憶部 14 および不揮発記憶部 16 に対してアクセスするためのメモリ制御方式を管理する。管理装置 22 は、第 1 記憶部 14 および不揮発記憶部 16 に対して、複数のメモリ制御方式によりアクセスすることが可能である。アクセス管理部 34 は、管理装置 22 が複数のメモリ制御方式のうち何れのメモリ制御方式によりアクセスするかを管理する。

【0036】

また、アクセス管理部 34 は、変換テーブルを記憶する。変換テーブルは、アクセス要求をするそれぞれのページについて、要求アドレスと、第 1 記憶部 14 または不揮発記憶部 16 における対応するページ番号（物理アドレス）との対応関係を記憶する。さらに、変換テーブルは、処理回路 20 がアクセス要求をするそれぞれのページについて、第 1 アクセス処理または第 2 アクセス処理の何れを実行するかを示すアクセス方法を記憶する。

20

【0037】

第 1 アクセス処理は、不揮発記憶部 16 から第 1 記憶部 14 に転送したデータに対して、書き込みおよび読み出しをする方法である。第 2 アクセス処理は、不揮発記憶部 16 に記憶されるデータに対して直接書き込みおよび読み出しをする方法である。

【0038】

そして、管理装置 22 は、データ処理部 32 がアクセス要求をしたページについて、変換テーブルに記憶されたアクセス方法により第 1 記憶部 14 および不揮発記憶部 16 にアクセスする。つまり、管理装置 22 は、第 1 アクセス処理に設定されているページに対して書き込みまたは読み出しの要求を受けた場合、第 1 アクセス処理を実行する。また、管理装置 22 は、第 2 アクセス処理に設定されているページに対して書き込みまたは読み出しの要求を受けた場合、第 2 アクセス処理を実行する。なお、変換テーブルの詳細については、図 2 を参照してさらに説明する。また、アクセス方法については、図 3 を参照してさらに説明する。

30

【0039】

方式決定部 36 は、動作情報検出部 24 により検出された動作情報、および、推定モデル記憶部 26 に記憶された推定モデルに基づき、複数のメモリ制御方式のうちの何れか 1 つのメモリ制御方式を決定する。そして、方式決定部 36 は、管理装置 22 が、決定したメモリ制御方式により第 1 記憶部 14 および不揮発記憶部 16 に対してアクセスするように、アクセス管理部 34 に対して設定処理をする。メモリ制御方式については、図 5 および図 6 を参照してさらに説明する。

40

【0040】

図 2 は、変換テーブルの一例を示す図である。変換テーブルは、処理回路 20 がアクセス要求をするそれぞれのページについて、要求アドレス（例えば論理アドレス）と、第 1 記憶部 14 または不揮発記憶部 16 における対応するページ番号（物理アドレス）との対応関係を記憶する。すなわち、変換テーブルは、処理回路 20 による要求アドレスに対応するデータが、第 1 記憶部 14 または不揮発記憶部 16 における何れのページに記憶され

50

ているのかを示すマッピング情報を記憶する。

【 0 0 4 1 】

例えば、図 2 の例では、変換テーブルにおけるページ番号の列に “ S x x x x ” と記述されたページは、不揮発記憶部 1 6 におけるページ番号が “ x x x x ” のページを示す。例えば、図 2 の例では、ページ番号の列に “ D x x x x ” と記述されたページは、第 1 記憶部 1 4 におけるページ番号が “ x x x x ” のページを示す。なお、ここでの x は、任意の値である。

【 0 0 4 2 】

さらに、変換テーブルは、処理回路 2 0 がアクセス要求をするそれぞれのページについて、第 1 アクセス処理または第 2 アクセス処理の何れを実行するかを示すアクセス方法を記憶する。なお、変換テーブルは、図 2 に示すような構成に限らず、他の構成であってもよい。

10

【 0 0 4 3 】

図 3 は、第 1 アクセス処理および第 2 アクセス処理の内容を示す図である。管理装置 2 2 は、第 1 アクセス処理に設定されている第 1 ページに対して書き込みまたは読み出しの要求を受けた場合、不揮発記憶部 1 6 に対して、第 1 アクセス処理を実行する。

【 0 0 4 4 】

例えば、図 3 に示すように、第 1 アクセス処理において、管理装置 2 2 は、不揮発記憶部 1 6 における第 1 ページに記憶された全てのデータを転送して、第 1 記憶部 1 4 に記憶させる。これにより、第 1 記憶部 1 4 は、不揮発記憶部 1 6 における第 1 ページに記憶されたデータのコピーを記憶することができる。続いて、第 1 アクセス処理において、管理装置 2 2 は、第 1 記憶部 1 4 に記憶されている、不揮発記憶部 1 6 から転送されたデータに対して、読み出しおよび書き込みをする。例えば、管理装置 2 2 は、不揮発記憶部 1 6 から第 1 記憶部 1 4 に転送されたデータに対して、ページより小さいサイズ（例えばプロセッサのキャッシュラインサイズ）でデータの読み出しおよび書き込みをする。そして、第 1 アクセス処理において、管理装置 2 2 は、第 1 記憶部 1 4 の空き容量が無くなり不揮発記憶部 1 6 から第 1 記憶部 1 4 にデータを転送できなくなった場合、および、第 1 記憶部 1 4 に記憶させておくことが不要と判断された場合等において、第 1 記憶部 1 4 に転送されたデータを、不揮発記憶部 1 6 における第 1 ページに書き戻す。

20

【 0 0 4 5 】

なお、管理装置 2 2 は、第 1 記憶部 1 4 に転送されたデータを、第 1 ページ（同じ場所）以外に書き戻してもよい。例えば、管理装置 2 2 は、第 1 アクセス処理において、不揮発記憶部 1 6 から第 1 記憶部 1 4 に転送されたデータを、何れの要求アドレスも対応付けられていない未使用ページに書き戻してよい。これにより、管理装置 2 2 は、ページ毎の書き換え回数の隔たりを小さくし、特定のページの品質劣化を抑制することができる。

30

【 0 0 4 6 】

また、管理装置 2 2 は、第 2 アクセス処理に設定されている第 2 ページに対して書き込みまたは読み出しの要求を受けた場合、不揮発記憶部 1 6 に対して、第 2 アクセス処理を実行する。

【 0 0 4 7 】

例えば、図 3 に示すように、第 2 アクセス処理において、管理装置 2 2 は、不揮発記憶部 1 6 における第 2 ページに対して、直接読み出しおよび書き込みをする。例えば、管理装置 2 2 は、ページより小さいサイズ（例えばプロセッサのキャッシュラインサイズ）でデータの読み出しおよび書き込みをする。

40

【 0 0 4 8 】

このように、管理装置 2 2 は、2 種類のアクセス方法により不揮発記憶部 1 6 に対してアクセスする。例えば、メモリアクセスに局所性が高いアプリケーションを実行した場合、管理装置 2 2 は、第 1 アクセス処理により不揮発記憶部 1 6 に記憶されたページをアクセスする。これにより、管理装置 2 2 は、メモリアクセスに局所性が高いアプリケーションを実行した場合、同一のページに対してより高速に処理をすることができる。

50

【 0 0 4 9 】

また、例えば、ランダムアクセスのようにメモリアクセスに局所性が低い処理を実行する場合、管理装置 2 2 は、第 2 アクセス処理により不揮発記憶部 1 6 に記憶されたページをアクセスする。これにより、管理装置 2 2 は、局所性が低い処理を実行する場合、不揮発記憶部 1 6 から第 1 記憶部 1 4 への転送処理のオーバーヘッドを無くして、効率良く処理をすることができる。このように、管理装置 2 2 は、第 1 アクセス処理および第 2 アクセス処理の 2 種類のアクセス方法を用いることにより、処理の効率化を図ることができる。

【 0 0 5 0 】

図 4 は、動作情報の時系列データの一例を示す図である。動作情報検出部 2 4 は、一定時間毎に、処理回路 2 0 および管理装置 2 2 の少なくとも一方における動作情報を検出する。動作情報は、例えば、OS (オペレーティングシステム) 内のイベントの統計情報 (OS の挙動を表す情報) などのシステムレベルの情報であってもよい。

10

【 0 0 5 1 】

動作情報は、例えば、処理回路 2 0 による主記憶装置に対する書き込み回数、処理回路 2 0 による主記憶装置に対する読み出し回数、処理回路 2 0 のプリフェッチ成功率、処理回路 2 0 のキャッシュヒット数、処理回路 2 0 のキャッシュミス数、または TLB ミス回数等である。また、これらに限らず、動作情報は、例えば、処理回路 2 0 または管理装置 2 2 内での所定のイベントの回数、リタイアした命令数、コアクロック数およびキャッシュメモリにおける階層毎 (L 1 キャッシュ、L 2 キャッシュ、L 3 キャッシュおよびラストレベルキャッシュ) のキャッシュヒット数 / キャッシュミス数、ラストレベルキャッシュから追い出されたキャッシュライン数、各階層のキャッシュがキャッシュミスでストールするサイクル数等であってもよい。

20

【 0 0 5 2 】

動作情報検出部 2 4 は、このような動作情報を検出した時刻に対応付けて出力する。方式決定部 3 6 は、動作情報検出部 2 4 により検出された動作情報を時刻に対応付けた時系列データとして取得する。

【 0 0 5 3 】

図 5 は、メモリ制御方式の内容の第 1 例を示す図である。管理装置 2 2 は、複数のメモリ制御方式のうちの何れか 1 つのメモリ制御方式により第 1 記憶部 1 4 および不揮発記憶部 1 6 にアクセスする。

30

【 0 0 5 4 】

例えば、複数のメモリ制御方式のそれぞれは、複数のページのうちの第 1 アクセス処理に設定されるページの数と第 2 アクセス処理に設定されるページの数との比率により区別されてもよい。例えば、図 5 の例では、管理装置 2 2 は、比率が異なる第 1 から第 4 の方式の何れかにより、第 1 記憶部 1 4 および不揮発記憶部 1 6 にアクセスする。

【 0 0 5 5 】

例えば、図 5 の第 1 方式は、100%のページが第 2 アクセス処理に設定された方式である。また、図 5 の第 2 方式は、10%のページが第 1 アクセス処理に設定され、90%のページが第 2 アクセス処理に設定された方式である。また、図 5 の第 3 方式は、20%のページが第 1 アクセス処理に設定され、80%のページが第 2 アクセス処理に設定された方式である。また、図 6 の第 4 方式は、30%のページが第 1 アクセス処理に設定され、70%のページが第 2 アクセス処理に設定された方式である。なお、複数のメモリ制御方式は、図 5 よりもさらに細かい比率により区別されてもよい。

40

【 0 0 5 6 】

図 6 は、メモリ制御方式の内容の第 2 例を示す図である。また、複数のメモリ制御方式のそれぞれは、さらに、第 1 記憶部 1 4 のメモリ使用量の上限値で区別されてもよい。例えば、図 5 の例では、管理装置 2 2 は、比率およびメモリ使用量が異なる 8 つの方式の何れかにより、第 1 記憶部 1 4 および不揮発記憶部 1 6 にアクセスする。

【 0 0 5 7 】

例えば、図 6 の第 1 方式は、100%のページが第 2 アクセス処理に設定され、第 1 記

50

憶部 1 4 のメモリ使用量に制限がない方式である。図 6 の第 2 方式は、10%のページが第 1 アクセス処理に設定され、第 1 記憶部 1 4 のメモリ使用量の上限値が不揮発記憶部 1 6 の 5%に設定された方式である。図 6 の第 3 方式は、10%のページが第 1 アクセス処理に設定され、メモリ使用量の上限値が 10%に設定された方式である。

【0058】

図 6 の第 4 方式は、20%のページが第 1 アクセス処理に設定され、メモリ使用量の上限値が 10%に設定された方式である。図 6 の第 5 方式は、20%のページが第 1 アクセス処理に設定され、メモリ使用量の上限値が 20%に設定された方式である。

【0059】

図 6 の第 6 方式は、30%のページが第 1 アクセス処理に設定され、メモリ使用量の上限値が 15%に設定された方式である。図 6 の第 7 方式は、30%のページが第 1 アクセス処理に設定され、メモリ使用量の上限値が 30%に設定された方式である。なお、複数のメモリ制御方式は、比率および上限値がさらに細かく区別されていてもよい。

10

【0060】

処理回路 2 0 は、図 5 および図 6 に示したようなメモリ制御方式に限らず、他の方式を設定してもよい。また、例えば、情報処理システム 1 0 は、主記憶装置として、第 1 記憶部 1 4 および不揮発記憶部 1 6 とは特性が異なる、第 2 記憶部を備えてもよい。このような場合、管理装置 2 2 は、例えば、アクセス方法として、不揮発記憶部 1 6 から第 2 記憶部にデータをコピーしてからアクセスさせるような第 3 アクセス処理をさらに実行してもよい。そして、管理装置 2 2 は、複数のメモリ制御方式の中に、第 1 アクセス処理、第 2 アクセス処理および第 3 アクセス処理に設定されるページ数の比率等を設定する方式を含めてもよい。

20

【0061】

図 7 は、第 1 実施形態に係る方式決定部 3 6 の構成を推定モデル記憶部 2 6 とともに示す図である。方式決定部 3 6 は、タイミング発生部 4 2 と、取得部 4 4 と、推定部 4 6 と、選択部 4 8 と、設定部 5 0 とを有する。

【0062】

タイミング発生部 4 2 は、メモリ制御方式の切り替えタイミングとなった時に、検出期間を開始する指示を取得部 4 4 および設定部 5 0 に与える。例えば、タイミング発生部 4 2 は、一定時間毎に検出期間を開始する指示を取得部 4 4 および設定部 5 0 に与える。

30

【0063】

取得部 4 4 は、タイミング発生部 4 2 から開始指示を受け取る。取得部 4 4 は、開始指示を受け取ってから一定の期間（検出期間）における、処理回路 2 0 および管理装置 2 2 の少なくとも一方における動作情報を、動作情報検出部 2 4 から取得する。

【0064】

推定部 4 6 は、取得部 4 4 が取得した動作情報および推定モデル記憶部 2 6 に記憶された推定モデルに基づき、管理装置 2 2 が複数のメモリ制御方式のそれぞれにより動作する場合における、第 1 記憶部 1 4 および不揮発記憶部 1 6 に対するメモリアクセスの実行性能を推定する。推定部 4 6 は、実行性能として、例えば、管理装置 2 2 が検出期間において実行した処理を、管理装置 2 2 が対応するメモリ制御方式で実行した場合における実行時間を推定する。また、推定部 4 6 は、実行性能として、例えば、管理装置 2 2 が検出期間において実行した処理を、管理装置 2 2 が対応するメモリ制御方式で実行した場合における消費電力、または、不揮発記憶部 1 6 の寿命減少量（摩耗度）を推定してもよい。

40

【0065】

推定モデルは、動作情報から、複数のメモリ制御方式のそれぞれについての実行性能（例えば、実行時間、消費電力または寿命減少量）を算出するための、例えば演算式である。推定モデルは、例えば、決定木、線形回帰、ニューラルネットワークおよびサポートベクタマシン等であってもよい。なお、推定モデルについては、図 1 0 および図 1 1 を参照してさらに説明する。

【0066】

50

選択部 48 は、推定された複数のメモリ制御方式のそれぞれについての実行性能に基づき、複数のメモリ制御方式のうちの何れか 1 つのメモリ制御方式を選択する。例えば、選択部 48 は、実行性能が最も小さいまたは最も大きいメモリ制御方式を選択する。また、例えば、選択部 48 は、実行性能が予め定められた条件を満たすメモリ制御方式のうちの、実行性能が最も小さいまたは最も大きいメモリ制御方式を選択してもよい。メモリ制御方式の選択方法については、図 12 から図 15 を参照してさらに説明する。

【0067】

設定部 50 は、タイミング発生部 42 から開始指示を受け取る。設定部 50 は、開始指示を受け取ってから一定の期間（検出期間）において、管理装置 22 が基準のメモリ制御方式により第 1 記憶部 14 および不揮発記憶部 16 にアクセスするようにアクセス管理部 34 に対して設定処理をする。これにより、管理装置 22 は、検出期間において、管理装置 22 が基準のメモリ制御方式により第 1 記憶部 14 および不揮発記憶部 16 にアクセスすることができる。

10

【0068】

例えば、設定部 50 は、検出期間において、アクセスするページの全てのアクセス方法を第 2 アクセス処理に設定する。これにより、管理装置 22 は、検出期間において、第 2 アクセス処理により第 1 記憶部 14 および不揮発記憶部 16 にアクセスすることができる。なお、設定部 50 は、検出期間において、推定モデルにより実行性能を推定可能な動作情報が得られるメモリ制御方式に設定すればよい。もし、取得部 44 が取得する動作情報が、メモリ制御方式の違いによる影響を受けない場合には、設定部 50 は、検出期間において、何れのメモリ制御方式を設定してもよい。

20

【0069】

検出期間を経過した後の期間において、設定部 50 は、選択されたメモリ制御方式により管理装置 22 が第 1 記憶部 14 および不揮発記憶部 16 にアクセスするように、アクセス管理部 34 に対して設定処理をする。これにより、管理装置 22 は、検出期間が経過した後において、選択されたメモリ制御方式により第 1 記憶部 14 および不揮発記憶部 16 にアクセスすることができる。

【0070】

例えば、設定部 50 は、変換テーブルにおける複数のページのそれぞれのアクセス方法を、選択されたメモリ制御方式に応じて第 1 アクセス処理または第 2 アクセス処理に設定する。例えば、複数のメモリ制御方式が、第 1 アクセス処理に設定されるページの数と第 2 アクセス処理に設定されるページの数との比率で区別される場合、設定部 50 は、変換テーブルにおける複数のページのそれぞれのアクセス方法を、選択されたメモリ制御方式に示された比率に応じて第 1 アクセス処理または第 2 アクセス処理に設定する。

30

【0071】

なお、設定部 50 は、変換テーブルに管理されている全てのページのアクセス方法を変更するのではなく、直近（例えば、所定期間前まで）にアクセスしたページおよび検出期間後に新たにアクセスするページについてのアクセス方法を変更してもよい。また、設定部 50 は、検出期間後に新たにアクセスするページに対してアクセス方法を変更してもよい。

40

【0072】

また、設定部 50 は、処理回路 20 により管理される TLB (Translation Lookaside Buffer) と呼ばれる仮想管理機構により管理されるページのアクセス方法を変更してもよい。TLB は、仮想アドレスから物理アドレスへアドレス変換を行うために、ページにおける要求アドレス（論理アドレス）と物理アドレスとの対応を示す対応関係情報を記憶する。しかし、TLB は、保有しているエントリが限られているので、対応関係情報を必要に応じて入れ替える。TLB は、例えば、直近において高い頻度でアクセスされるページについての対応関係情報を優先して記憶する。

【0073】

さらに、複数のメモリ制御方式が、第 1 記憶部 14 のメモリ使用量の上限値で区別され

50

る場合、設定部 50 は、選択されたメモリ制御方式に示された上限値を、第 1 記憶部 14 のメモリ使用量として設定する。第 1 記憶部 14 のメモリ使用量の上限値が設定された場合、管理装置 22 は、第 1 記憶部 14 にコピーしたデータ量が上限値を超えないように、第 1 記憶部 14 および不揮発記憶部 16 に対してアクセス制御をする。具体的には、管理装置 22 は、第 1 アクセス処理に設定された第 1 ページにアクセスするために第 1 ページのデータを不揮発記憶部 16 から第 1 記憶部 14 に転送すると、第 1 記憶部 14 のメモリ使用量が上限値を超えてしまう場合には、第 1 記憶部 14 に記憶された何れかのページのデータを不揮発記憶部 16 に書き戻した後に、第 1 ページのデータを不揮発記憶部 16 から第 1 記憶部 14 に転送する。上限値を超えないようにするためのアクセス制御は、処理回路 20 (アクセス管理部 34 など) がおこなってもよい。

10

【0074】

図 8 は、動作情報の波形図および選択されたメモリ制御方式の一例を示す図である。図 8 に示すように、方式決定部 36 は、一定期間毎に、メモリ制御方式の選択および切り替えを行う。方式決定部 36 は、メモリ制御方式の切り替えタイミングの直前に検出期間を設定し、検出期間において動作情報を取得する。そして、方式決定部 36 は、取得した動作情報に基づきメモリ制御方式を決定し、新たなメモリ制御方式に切り替える。

【0075】

なお、図 8 の例では、方式決定部 36 は、検出期間において、基準のメモリ制御方式に設定している。しかし、方式決定部 36 は、メモリ制御方式の違いが、取得する動作情報に影響を与えない場合には、検出期間において、直前のメモリ制御方式を継続させてもよい。

20

【0076】

図 9 は、第 1 実施形態に係る方式決定部 36 による処理の流れを示すフローチャートである。第 1 実施形態に係る方式決定部 36 は、図 9 に示すフローチャートに従って処理を実行する。

【0077】

まず、S11において、方式決定部 36 は、動作情報の取得タイミングとなったか否かを判断する。取得タイミングとなっていない場合 (S11のNo)、方式決定部 36 は、処理をS11で待機する。取得タイミングとなった場合 (S11のYes)、方式決定部 36 は、処理をS12に進める。

30

【0078】

S12において、方式決定部 36 は、管理装置 22 が基準のメモリ制御方式により第 1 記憶部 14 および不揮発記憶部 16 にアクセスするように、アクセス管理部 34 に対して設定処理をする。続いて、S13において、方式決定部 36 は、動作情報検出部 24 から動作情報を取得する。続いて、S14において、方式決定部 36 は、検出期間が経過したか否かを判断する。検出期間が経過していない場合 (S14のNo)、方式決定部 36 は、処理をS13に戻して、検出期間が経過するまで、動作情報の取得を継続する。検出期間が経過した場合 (S14のYes)、方式決定部 36 は、処理をS15に進める。

【0079】

S15、S16およびS17のループ処理において、方式決定部 36 は、推定モデルおよび検出期間における動作情報に基づき、複数のメモリ制御方式のそれぞれについての実行性能を推定する。全てのメモリ制御方式について実行性能を推定した場合、方式決定部 36 は、処理をS18に進める。

40

【0080】

S18において、方式決定部 36 は、実行性能を推定した複数のメモリ制御方式の中から、最適メモリ制御方式を選択する。続いて、S19において、方式決定部 36 は、管理装置 22 が、選択した最適メモリ制御方式により第 1 記憶部 14 および不揮発記憶部 16 にアクセスするように、アクセス管理部 34 に対して設定処理をする。方式決定部 36 は、S19の処理を終了すると本フローを終了する。

【0081】

50

図10は、実行時間の算出処理の第1例を説明するための図である。実行性能として、実行時間を推定する場合、取得部44は、検出期間におけるメモリアクセス回数、検出期間にアクセスしたページ数、および、実行時間の実測値を取得する。そして、推定部46は、予め定められた推定モデル（例えば演算式）に、メモリアクセス回数、ページ数および実行時間の実測値を代入して、複数のメモリ制御方式のそれぞれについての実行時間の推定値を算出する。

【0082】

例えば、100%のページを第2アクセス処理に設定するメモリ制御方式により管理装置22を動作させ、100%のページを第1アクセス処理に設定するメモリ制御方式により管理装置22を動作させた場合における実行時間を推定する場合、推定部46は、下記の式(1)および式(2)に表される式を演算して、実行時間を算出する。

$$T1 = T2 - t \dots (1)$$

$$t = (N_A \times (L_S - L_D)) - (N_P \times T) \dots (2)$$

【0083】

式(1)において、T1は、100%のページを第1アクセス処理に設定するメモリ制御方式により管理装置22を動作させた場合における実行時間の推定値である。T2は、100%のページを第2アクセス処理に設定するメモリ制御方式により管理装置22を動作させた場合における実行時間の実測値である。

【0084】

また、式(2)において、N_Aは、メモリアクセス回数である。N_Pは、アクセスしたページ数である。

【0085】

また、式(2)において、L_Dは、第1記憶部14に対するメモリアクセスのレイテンシである。L_Sは、不揮発記憶部16に対するメモリアクセスのレイテンシである。また、Tは、1ページ分のデータを不揮発記憶部16から第1記憶部14に転送する転送時間である。L_D、L_SおよびTは、予め測定済みの値であり、推定モデルに定数として組み込まれる。

【0086】

従って、例えば、推定モデル記憶部26は、上述の式(1)および式(2)により表される演算式を推定モデルとして記憶する。推定部46は、100%のページを第2アクセス処理に設定するメモリ制御方式により管理装置22を動作させた場合における、検出期間におけるメモリアクセス回数、ページ数および実行時間の実測値を取得する。そして、推定部46は、上述の式(1)および式(2)により表される演算式と、検出期間におけるメモリアクセス回数、ページ数および実行時間の実測値とに基づき、100%のページを第1アクセス処理に設定するメモリ制御方式により管理装置22を動作させた場合の実行時間を推定する。

【0087】

また、X%のページ(Xは、0より大きく100より小さい値)を第1アクセス処理に設定するメモリ制御方式により管理装置22を動作させた場合の実行時間を推定する場合、推定部46は、次のような処理を行う。

【0088】

まず、推定部46は、100%のページを第2アクセス処理に設定するメモリ制御方式により管理装置22を動作させた場合における、検出期間におけるメモリアクセス回数、ページ数および実行時間の実測値を取得する。続いて、推定部46は、検出期間にアクセスしたページの中から、アクセス数の多い上位のX%のページを選択する。推定部46は、選択したX%のページのページ数を、式(2)におけるN_Pに代入する。

【0089】

さらに、推定部46は、選択したX%のページに対するメモリアクセス回数を算出する。推定部46は、算出したX%のページに対するメモリアクセス回数を、式(2)におけるN_Aに代入する。そして、推定部46は、上記の式(1)および式(2)に表される式

10

20

30

40

50

を演算して、実行時間を算出する。これにより、推定部 4 6 は、X % のページを第 1 アクセス処理に設定するメモリ制御方式により管理装置 2 2 を動作させた場合の実行時間を推定することができる。

【 0 0 9 0 】

図 1 1 は、実行時間の算出処理の第 2 例を説明するための図である。図 1 0 に示した例は、アクセス数の多い上位 X % のページ数 (N_p) に対する、上位 X % のページへのメモリアクセス回数 (N_A) の割合が大きい場合、 t が正の値となる場合を示している。しかし、図 1 1 に示すように、アクセス数の多い上位 X % のページ数 (N_p) に対する、上位 X % のページ数へのメモリアクセス回数 (N_A) の割合が小さい場合、 t が負の値となる。つまり、メモリアクセスが多数のページに分散している場合、 t が負の値となる。この場合、ページ転送によるオーバーヘッドが相対的に大きくなり、管理装置 2 2 は、第 1 アクセス処理を実行しても、実行時間を短くすることができない。

10

【 0 0 9 1 】

従って、推定部 4 6 は、第 1 アクセス処理に設定するページのパーセンテージ (つまり、X) を変えた複数のメモリ制御方式についての実行時間を推定する。そして、選択部 4 8 は、複数のメモリ制御方式の中から、より適切な実行時間となるような 1 つのメモリ制御方式を選択する。

【 0 0 9 2 】

図 1 2 は、第 1 アクセス処理に設定されたページの比率に対する実行性能を示す図である。情報処理システム 1 0 は、複数のアプリケーションプログラムを同時に並行して実行することができる。

20

【 0 0 9 3 】

複数のアプリケーションプログラムを同時に実行する場合、アプリケーションプログラム毎に、メモリ制御方式に対する実行性能を表す特性が異なる。例えば、図 1 2 の例では、第 1 のアプリケーションプログラムは、第 1 アクセス処理に設定されたページの比率が低いメモリ制御方式で動作させた場合でも、実行性能は高い。これは、第 1 のアプリケーションプログラムが局所性の高いメモリアクセス (同一の領域に繰り返して書き込みをするようなアクセス) を実行するためであると考えられる。

【 0 0 9 4 】

また、例えば、図 1 2 の例では、第 2 のアプリケーションプログラムは、第 1 アクセス処理に設定されたページの比率が小さいと、実行性能が著しく低い。これは、第 1 のアプリケーションプログラムが局所性の低いメモリアクセス (多数の領域に分散して書き込みがされるようなアクセス) を実行するためであると考えられる。

30

【 0 0 9 5 】

従って、推定部 4 6 は、複数のアプリケーションプログラムを実行する場合、それぞれのアプリケーションプログラムがアクセスする記憶領域毎に、実行性能を推定する。選択部 4 8 は、それぞれのアプリケーションプログラムがアクセスする記憶領域毎に、メモリ制御方式を選択する。そして、設定部 5 0 は、それぞれのアプリケーションプログラムがアクセスする記憶領域毎に、メモリ制御方式を設定する。

【 0 0 9 6 】

図 1 3 は、メモリ制御方式の選択方法の第 1 例を説明するための図である。例えば、選択部 4 8 は、実行性能が予め定められた条件を満たすメモリ制御方式のうち、実行性能が最も小さいまたは最も大きいメモリ制御方式を選択する。

40

【 0 0 9 7 】

例えば、選択部 4 8 は、管理装置 2 2 が基準のメモリ制御方式で動作した場合における実行性能の実測値を所定数倍して基準値を算出する。例えば、選択部 4 8 は、1 0 0 % のページが第 2 アクセス処理に設定されたメモリ制御方式により管理装置 2 2 が動作した場合における実行性能の実測値を所定数倍して基準値を算出する。

【 0 0 9 8 】

そして、推定された実行性能が基準値より大きいことが条件である場合、選択部 4 8 は

50

、推定された実行性能が基準値より大きいメモリ制御方式のうちの、最小の実行性能のメモリ制御方式を選択する。あるいは、推定された実行性能が基準値より小さいことが条件である場合、選択部 48 は、推定された実行性能が基準値より小さいメモリ制御方式のうちの、最大の実行性能のメモリ制御方式を選択する。これにより、選択部 48 は、少なくとも基準の条件を満たしながら、最低のコストで管理装置 22 を動作させることができる。

【0099】

図 14 は、メモリ制御方式の選択方法の第 2 例を説明するための図である。また、選択部 48 は、実行性能が予め定められた条件を満たす制御方式が存在しない場合には、実行性能が最も小さいまたは最も大きいメモリ制御方式を選択する。

10

【0100】

例えば、選択部 48 は、管理装置 22 が基準のメモリ制御方式で動作した場合における実行性能の実測値の所定数倍の値を基準値とする。そして、推定された実行性能が基準値より大きいことが条件である場合において、推定された実行性能が基準値より大きいメモリ制御方式が存在しない場合、選択部 48 は、最大の実行性能のメモリ制御方式を選択する。あるいは、推定された実行性能が基準値より小さいことが条件である場合において、推定された実行性能が基準値より小さいメモリ制御方式が存在しない場合には、選択部 48 は、最小の実行性能のメモリ制御方式を選択する。これにより、選択部 48 は、最も実行性能の良いメモリ制御方式により管理装置 22 を動作させることができる。

【0101】

図 15 は、メモリ制御方式の選択方法の第 3 例を説明するための図である。例えば、選択部 48 は、管理装置 22 が、最もコストの高いメモリ制御方式で動作した場合における実行性能の推定値を所定割合低下させて基準値を算出してもよい。例えば、選択部 48 は、100%のページが第 1 アクセス処理に設定されたメモリ制御方式により管理装置 22 が動作した場合における実行性能の推定値における所定割合低下した値を基準値として算出してもよい。

20

【0102】

例えば、選択部 48 は、複数のページの全てを第 1 アクセス処理に設定するメモリ制御方式により管理装置 22 が動作した場合における実行時間の推定値より所定割合長い基準値（基準時間）を算出する。そして、選択部 48 は、推定した実行時間が基準値（基準時間）より短いメモリ制御方式のうちの、推定した実行時間が最も長いメモリ制御方式を選択してもよい。これにより、選択部 48 は、基準のコストを超えない最も実行性能の良いメモリ制御方式を選択することができる。

30

【0103】

以上のように、第 1 実施形態に係る情報処理システム 10 は、方式決定部 36 が、動作情報に基づき適切なメモリ制御方式を選択し、選択したメモリ制御方式により管理装置 22 に第 1 記憶部 14 および不揮発記憶部 16 にアクセスさせる。これにより、第 1 実施形態に係る情報処理システム 10 によれば、例えば実行するアプリケーションプログラム毎に適切な実行性能が得られるように、管理装置 22 を動作させることができる。

【0104】

（第 2 実施形態）

つぎに、第 2 実施形態について説明する。第 2 実施形態は、第 1 実施形態の変形例である。

【0105】

図 16 は、第 2 実施形態に係る方式決定部 36 の構成を推定モデル記憶部 26 とともに示す図である。第 2 実施形態に係る方式決定部 36 は、第 1 実施形態に係る方式決定部 36 と比較して、タイミング発生部 42 に代えて、変化検出部 52 を有する。

【0106】

変化検出部 52 は、情報処理システム 10 の動作中の全期間において、動作情報を動作情報検出部 24 から取得する。変化検出部 52 は、取得した動作情報が所定の変化をした

40

50

場合に、取得部 4 4 に対して検出期間を開始する指示を与える。所定の変化とは、例えば、動作情報における特定のパラメータが一定時間内に所定値以上増加した場合または減少した場合である。

【 0 1 0 7 】

取得部 4 4 は、変化検出部 5 2 から開始指示を受け取る。そして、取得部 4 4 は、開始指示を受け取ってから一定の期間（検出期間）における、動作情報を動作情報検出部 2 4 から取得する。

【 0 1 0 8 】

図 1 7 は、動作情報の波形図、動作情報の変化点および選択されたメモリ制御方式の一例を示す図である。第 2 実施形態において、方式決定部 3 6 は、動作情報に所定の変化が検出される毎に、メモリ制御方式の選択および設定を行う。より具体的には、方式決定部 3 6 は、所定の変化が検出された後に検出期間を設定し、検出期間において動作情報を取得する。続いて、方式決定部 3 6 は、取得した動作情報に基づきメモリ制御方式を選択する。そして、方式決定部 3 6 は、選択したメモリ制御方式により管理装置 2 2 が第 1 記憶部 1 4 および不揮発記憶部 1 6 にアクセスするように、アクセス管理部 3 4 に対して設定処理をする。

10

【 0 1 0 9 】

変化検出部 5 2 が取得する動作情報と取得部 4 4 が取得する動作情報とは、同一であっても異なってもよい。変化検出部 5 2 が取得する動作情報と取得部 4 4 が取得する動作情報とが同一であり、且つ、検出期間において基準のメモリ制御方式に切り換えない場合、方式決定部 3 6 は、所定の変化を検出した時点までに取得している動作情報を用いて、メモリ制御方式を選択してもよい。これにより、方式決定部 3 6 は、検出期間を設定せずに、メモリ制御方式の切り換えをすることができる。

20

【 0 1 1 0 】

また、変化検出部 5 2 が取得する動作情報と取得部 4 4 が取得する動作情報とが異なる場合において、動作情報検出部 2 4 は、制約により、変化検出部 5 2 が取得する動作情報と取得部 4 4 が取得する動作情報とを同時に出力することができない場合がある。このような場合、変化検出部 5 2 は、情報処理システム 1 0 の動作中において、変化点を検出するために必要な動作情報を動作情報検出部 2 4 に出力させる。続いて、変化検出部 5 2 は、所定の変化を検出した場合、動作情報検出部 2 4 に切り換え指示を与え、メモリ制御方式を選択するために必要な動作情報を動作情報検出部 2 4 に出力させる。そして、変化検出部 5 2 は、検出期間が経過した後、動作情報検出部 2 4 に切り換え指示を与え、変化点を検出するために必要な動作情報を動作情報検出部 2 4 に出力させる。

30

【 0 1 1 1 】

変化検出部 5 2 は、例えば、プリフェッチ成功率が一定期間内に所定値以上増加または減少する点を、所定の変化が生じた点として検出してもよい。プリフェッチ成功率とは、プリフェッチのための先読みアルゴリズムが、アプリケーションプログラムにどの程度適合しているかを表す値である。具体的には、第 1 記憶部 1 4 または不揮発記憶部 1 6 からラストレベルキャッシュ（例えば、L 3）へロードされたデータ総量（総メモリアクセス量）を、M とする。プリフェッチにより、アプリケーションプログラムによる実際の使用前にラストレベルキャッシュへロードされたデータ量を P とする。プリフェッチ成功率は、M に対する P の割合である。

40

【 0 1 1 2 】

なお、変化検出部 5 2 は、プリフェッチ成功率に限らず、各種キャッシュミス率（L 1 キャッシュミス率、L 2 キャッシュミス率、L 3 キャッシュミス率またはラストレベルキャッシュミス率）等に所定の変化が生じたか否かを判断してもよい。

【 0 1 1 3 】

以上のように、第 2 実施形態に係る情報処理システム 1 0 は、方式決定部 3 6 が、動作情報の変化点においてメモリ制御方式を切り換える。これにより、第 2 実施形態に係る情報処理システム 1 0 によれば、適切な位置でメモリ制御方式を切り換えて、適切な実行性

50

能を得られるように管理装置 2 2 を動作させることができる。

【 0 1 1 4 】

(第 3 実施形態)

つぎに、第 3 実施形態について説明する。第 3 実施形態は、第 2 実施形態の変形例である。

【 0 1 1 5 】

図 1 8 は、第 3 実施形態に係る情報処理システム 1 0 の構成を示す図である。第 3 実施形態に係る情報処理装置 1 8 は、第 2 実施形態に係る情報処理装置 1 8 と比較して、結果記憶部 6 2 と、決定モデル記憶部 6 4 とをさらに有する。

【 0 1 1 6 】

結果記憶部 6 2 は、過去の動作情報の変化内容を表す変化情報を記憶する。さらに、結果記憶部 6 2 は、過去に選択されたメモリ制御方式を記憶する。結果記憶部 6 2 は、動作情報の変化内容と選択されたメモリ制御方式とを時系列に対応付けて記憶する。

【 0 1 1 7 】

決定モデル記憶部 6 4 は、動作情報の変化内容から、直接、メモリ制御方式を決定するための決定モデルを記憶する。決定モデルは、方式決定部 3 6 により生成されて、決定モデル記憶部 6 4 に書き込まれる。

【 0 1 1 8 】

図 1 9 は、第 3 実施形態に係る方式決定部 3 6 の構成を推定モデル記憶部 2 6 、結果記憶部 6 2 および決定モデル記憶部 6 4 とともに示す図である。第 3 実施形態に係る情報処理装置 1 8 は、第 2 実施形態に係る方式決定部 3 6 と比較して、学習部 6 8 と、決定部 7 0 と、モード設定部 7 2 とをさらに有する。

【 0 1 1 9 】

変化検出部 5 2 は、検出した動作情報の変化内容を表す変化情報を結果記憶部 6 2 に書き込む。例えば、変化検出部 5 2 は、所定の変化を検出したそれぞれの時点における、動作情報の変化の方向（例えば立ち上がりまたは立ち下り）および変化の大きさを表す変化情報を結果記憶部 6 2 に書き込む。

【 0 1 2 0 】

また、選択部 4 8 は、所定の変化を検出したことにより選択したメモリ制御方式を結果記憶部 6 2 に書き込む。結果記憶部 6 2 は、動作情報の変化内容を表す変化情報と選択されたメモリ制御方式とを対応付けて記憶する。

【 0 1 2 1 】

学習部 6 8 は、結果記憶部 6 2 に記憶された、動作情報の変化内容を表す変化情報と選択されたメモリ制御方式との関係に基づき、動作情報の変化内容からメモリ制御方式を決定するための決定モデルを生成する。例えば、学習部 6 8 は、学習処理またはクラスタリング処理等を実行することにより動作情報の複数の変化内容のそれぞれと、選択されたメモリ制御方式との対応関係を検出する。そして、学習部 6 8 は、複数の変化内容のそれぞれと選択されたメモリ制御方式との対応関係に基づき、決定モデルを生成する。学習部 6 8 は、生成した決定モデルを決定モデル記憶部 6 4 に書き込む。

【 0 1 2 2 】

決定部 7 0 は、変化検出部 5 2 が動作情報から所定の変化を検出した時点における、動作情報の変化の内容を表す変化情報を取得する。例えば、決定部 7 0 は、変化情報として、変化検出部 5 2 が動作情報から所定の変化を検出した時点における、動作情報の変化の方向および動作情報の変化の大きさを取得する。そして、決定部 7 0 は、取得した変化情報と決定モデル記憶部 6 4 に記憶された決定モデルとに基づき、メモリ制御方式を決定する。決定部 7 0 は、決定したメモリ制御方式を設定部 5 0 に通知する。

【 0 1 2 3 】

モード設定部 7 2 は、実行性能を推定してメモリ制御方式を選択する第 1 のモードと、実行性能を推定せずにメモリ制御方式を決定する第 2 のモードとを切り換える。モード設定部 7 2 は、情報処理システム 1 0 が動作を開始した段階においては、第 1 のモードに設

10

20

30

40

50

定する。モード設定部 7 2 は、結果記憶部 6 2 に十分な変化情報および選択情報が蓄積されたことにより精度の良い決定モデルが生成された後に、第 2 のモードに設定する。また、モード設定部 7 2 は、ユーザの指示により第 1 のモードと第 2 のモードとを切り換えてもよい。

【 0 1 2 4 】

第 1 のモードに設定されている場合、取得部 4 4、推定部 4 6、選択部 4 8、設定部 5 0 および変化検出部 5 2 は、第 2 実施形態と同様に動作する。また、第 1 のモードに設定されている場合、決定部 7 0 は、機能を停止している。

【 0 1 2 5 】

第 2 のモードに設定されている場合、取得部 4 4、推定部 4 6 および選択部 4 8 は、機能を停止している。また、第 2 のモードに設定されている場合、変化検出部 5 2 は、取得部 4 4 への検出期間を開始する指示の出力を停止し、決定部 7 0 へ変化情報を与える。第 2 のモードに設定されている場合、決定部 7 0 は、取得した変化情報と決定モデル記憶部 6 4 に記憶された決定モデルとに基づき、メモリ制御方式を決定する。そして、第 2 のモードに設定されている場合、設定部 5 0 は、決定部 7 0 により決定されたメモリ制御方式により管理装置 2 2 が第 1 記憶部 1 4 および不揮発記憶部 1 6 にアクセスするように、アクセス管理部 3 4 に対して設定処理をする。

10

【 0 1 2 6 】

図 2 0 は、動作情報の変化、実行性能および決定されたメモリ制御方式を示す図である。学習部 6 8 が学習した結果、決定モデルは、動作情報が第 1 閾値より大きい変化で立ち上がった場合に第 1 方式を選択し、動作情報が第 2 閾値より大きい変化で立ち下がった場合および第 1 閾値以下の変化で立ち上がった場合に第 2 方式を選択するモデルとなったとする。

20

【 0 1 2 7 】

このような決定モデルを用いると、図 2 0 の時刻 t_1 のように、方式決定部 3 6 は、動作情報が第 1 閾値より大きい変化で立ち上がった場合、第 1 方式に決定する。また、図 2 0 の時刻 t_2 および t_4 のように、方式決定部 3 6 は、動作情報が第 2 閾値より大きい変化で立ち下がった場合、第 2 方式に決定する。また、図 2 0 の時刻 t_3 のように、方式決定部 3 6 は、動作情報が第 1 閾値以下の変化で立ち上がった場合、第 2 方式に決定する。

【 0 1 2 8 】

以上のように、第 3 実施形態に係る情報処理システム 1 0 は、方式決定部 3 6 が、過去の動作情報の変化を表す変化情報と過去に選択されたメモリ制御方式を表す選択情報との関係に基づき、変化情報から直接メモリ制御方式を決定するための決定モデルを生成する。そして、方式決定部 3 6 は、第 2 モードに設定された場合、動作情報の変化を表す変化情報から、メモリ制御方式を決定する。これにより、第 3 実施形態に係る情報処理システム 1 0 によれば、実行性能を推定せずにメモリ制御方式を決定することができる。従って、情報処理システム 1 0 によれば、少ない演算量でメモリ制御方式を高速に決定することができる。

30

【 0 1 2 9 】

(第 4 実施形態)

つぎに、第 4 実施形態について説明する。

40

【 0 1 3 0 】

図 2 1 は、第 1 情報処理システム 1 1 0 および第 2 情報処理システム 1 2 0 を示す図である。例えば、社会インフラシステム等に用いられるエッジシステムは、長期に運用がされる。このようなエッジシステムは、高機能化および高精度化等を図るために、運用途中にアプリケーションプログラムが変更される場合がある。このような場合、設計者は、エッジシステムのハードウェア構成を再設計しなければならない。

【 0 1 3 1 】

例えば、現在、現行アプリケーションプログラムを実行する第 1 情報処理システム 1 1 0 を運用しているとする。現行アプリケーションプログラムをバージョンアップして、新

50

アプリケーションプログラムを実行しようとした場合、第1情報処理システム110のハードウェア能力が足りないと予測される。この場合、設計者は、第1情報処理システム110のハードウェア構成を変更することにより、新アプリケーションプログラムを実行するために十分なハードウェア構成を有する第2情報処理システム120を設計しなければならない。

【0132】

例えば、第1情報処理システム110は、情報処理装置18と、第1記憶部14とを備える。設計者は、このような構成の第1情報処理システム110に、さらに、主記憶装置として不揮発記憶部16を加えた第2情報処理システム120を設計する。第2情報処理システム120は、主記憶装置の容量が増えることにより、二次記憶装置へのデータの回避が少なくなったり、あるいは、データが不揮発記憶部16上に保存されるようになるため二次記憶装置が不要になりデータの移動が削減されるため、より短い実行時間で処理を完了させることができる。

10

【0133】

ここで、設計者は、第2情報処理システム120がどの程度の仕様の不揮発記憶部16を備えればよいかを判断しなければならない。例えば、第2情報処理システム120は、高速な不揮発記憶部16を備えれば、高速に動作することはできるが、コストは高くなる。反対に、第2情報処理システム120は、低速な不揮発記憶部16を備えれば、コストは低い、高速に動作することはできない。従って、設計者は、新アプリケーションプログラムを実行した場合の実行性能を考慮しながら、第2情報処理システム120を設計しなければならない。

20

【0134】

しかし、第2情報処理システム120を実際に製造して実行性能を測定した場合、第2情報処理システム120の設計時間が長くなり、設計コストが高くなってしまう。

【0135】

図22は、第4実施形態に係る推定システム130の構成を示す図である。推定システム130は、第1情報処理システム110と、推定装置140とを備える。第1情報処理システム110は、第2情報処理システム120により実行される予定のアプリケーションプログラムを実行する。

【0136】

推定装置140は、第2情報処理システム120がアプリケーションプログラムを実行した場合におけるメモリアクセスの実行性能を推定する。第2情報処理システム120は、第1情報処理システム110のハードウェア構成を変更したシステムである。例えば、第2情報処理システム120は、情報処理装置18と第1記憶部14とを備える第1情報処理システム110に、不揮発記憶部16を加えた構成である。第2情報処理システム120は、第1実施形態において説明した情報処理システム10と同一の構成である。推定装置140は、通常のコンピュータ等であってもよい。

30

【0137】

推定装置140は、アプリケーションプログラムを実行した第1情報処理システム110の動作情報の時系列データを取得する。そして、推定装置140は、取得した動作情報の時系列データに基づき、第2情報処理システム120がアプリケーションプログラムを実行した場合における、第1記憶部14および不揮発記憶部16に対するメモリアクセスの実行性能を推定して出力する。

40

【0138】

図23は、第4実施形態に係る推定装置140の構成を示す図である。推定装置140は、取得部44と、動作情報記憶部142と、分割部144と、推定モデル記憶部26と、推定部46と、選択部48と、出力部146とを備える。

【0139】

取得部44は、アプリケーションプログラムを実行した第1情報処理システム110の動作情報の時系列データを取得する。例えば、取得部44は、アプリケーションプログラ

50

ムの全実行期間の動作情報の時系列データを取得する。取得部 4 4 は、取得した動作情報の時系列データを動作情報記憶部 1 4 2 に書き込む。動作情報記憶部 1 4 2 は、取得部 4 4 により取得された動作情報の時系列データを記憶する。

【 0 1 4 0 】

分割部 1 4 4 は、動作情報記憶部 1 4 2 に記憶された動作情報の時系列データを時間方向に分割することにより複数の部分情報を生成する。分割部 1 4 4 は、例えば、動作情報の時系列データを一定時間毎に区切って複数の部分情報を生成する。また、分割部 1 4 4 は、動作情報の時系列データを不均等な間隔毎に区切って複数の部分情報を生成してもよい。また、分割部 1 4 4 は、第 2 実施形態と同様に、動作情報の時系列データから動作情報の変化点を検出し、変化点において動作情報の時系列データを区切って複数の部分情報を生成してもよい。

10

【 0 1 4 1 】

推定モデル記憶部 2 6 は、動作情報から、第 1 記憶部 1 4 および不揮発記憶部 1 6 に対するメモリアクセスの実行性能を推定するための推定モデルを記憶する。推定モデルは、第 1 実施形態において説明した推定モデルと同一である。

【 0 1 4 2 】

推定部 4 6 は、複数の部分情報のそれぞれについて、第 1 情報処理システム 1 1 0 が実行した処理を、第 2 情報処理システム 1 2 0 が複数のメモリ制御方式のそれぞれにより実行した場合における、第 1 記憶部 1 4 および不揮発記憶部 1 6 に対するメモリアクセスの実行性能を推定する。この場合において、推定部 4 6 は、対応する部分情報および推定モデルに基づき、複数のメモリ制御方式のそれぞれについての実行性能を推定する。

20

【 0 1 4 3 】

例えば、推定部 4 6 は、実行性能として、第 1 情報処理システム 1 1 0 が実行した処理を、第 2 情報処理システム 1 2 0 が対応するメモリ制御方式で実行した場合における実行時間を推定する。また、推定部 4 6 は、実行性能として、第 1 情報処理システム 1 1 0 が実行した処理を、第 2 情報処理システム 1 2 0 が対応するメモリ制御方式で実行した場合における消費電力または寿命減少量（摩耗度）を推定してもよい。1つの部分情報に対する推定部 4 6 による推定処理は、第 1 実施形態において説明した推定処理と同一である。

【 0 1 4 4 】

選択部 4 8 は、複数の部分情報のそれぞれについて、推定された複数のメモリ制御方式のそれぞれについての実行性能に基づき複数のメモリ制御方式のうちの何れか1つの最適メモリ制御方式を選択する。1つの部分情報についての選択部 4 8 による選択処理は、第 1 実施形態において説明した選択処理と同一である。

30

【 0 1 4 5 】

出力部 1 4 6 は、選択部 4 8 から、複数の部分情報のそれぞれについて選択された複数の最適メモリ制御方式を取得する。さらに、出力部 1 4 6 は、複数の部分情報のそれぞれについて選択された複数の最適メモリ制御方式のそれぞれについて、推定部 4 6 から、対応する実行性能を取得する。

【 0 1 4 6 】

出力部 1 4 6 は、複数の部分情報のそれぞれについて選択された複数の最適メモリ制御方式についての複数の実行性能を合計した合計実行性能を算出する。そして、出力部 1 4 6 は、算出した合計実行性能を出力する。例えば、出力部 1 4 6 は、算出した合計実行性能を表示装置等に表示させる。

40

【 0 1 4 7 】

さらに、出力部 1 4 6 は、複数の部分情報のそれぞれについて選択された複数の最適メモリ制御方式の時系列の並びを示す方式選択情報を出力してもよい。さらに、出力部 1 4 6 は、メモリ制御方式の切り換え時刻も併せて出力してもよい。

【 0 1 4 8 】

図 2 4 は、第 4 実施形態に係る推定装置 1 4 0 による処理の流れを示すフローチャートである。第 1 実施形態に係る推定装置 1 4 0 は、図 2 4 に示すフローチャートに従って処

50

理を実行する。

【 0 1 4 9 】

まず、S 1 1 1において、推定装置 1 4 0は、アプリケーションプログラムを実行した第 1 情報処理システム 1 1 0の動作情報の時系列データを取得する。例えば、推定装置 1 4 0は、アプリケーションプログラムの全実行期間の動作情報の時系列データを取得する。そして、推定装置 1 4 0は、取得した動作情報の時系列データを記憶する。

【 0 1 5 0 】

続いて、S 1 1 2において、推定装置 1 4 0は、記憶した動作情報の時系列データを時間方向に分割して複数の部分情報を生成する。続いて、推定装置 1 4 0は、複数の部分情報のそれぞれ毎に、S 1 1 3とS 1 1 8との間のループ処理を実行する。

10

【 0 1 5 1 】

S 1 1 3とS 1 1 8との間のループ処理内において、まず、S 1 1 4、S 1 1 5およびS 1 1 6のサブループ処理を実行する。サブループ処理において、推定装置 1 4 0は、対応する部分情報および推定モデルに基づき、複数のメモリ制御方式のそれぞれについての実行性能を推定する。全てのメモリ制御方式について実行性能を推定した場合、推定装置 1 4 0は、処理をS 1 1 7に進める。

【 0 1 5 2 】

S 1 1 7において、推定装置 1 4 0は、実行性能を推定した複数のメモリ制御方式の中から、1つの最適メモリ制御方式を選択する。そして、複数の部分情報の全てについて、最適メモリ制御方式を選択した場合、推定装置 1 4 0は、処理をS 1 1 9に進める。

20

【 0 1 5 3 】

S 1 1 9において、推定装置 1 4 0は、複数の部分情報について選択した複数の最適メモリ制御方式についての複数の実行性能を取得し、取得した複数の実行性能を合計した合計実行性能を算出する。そして、推定装置 1 4 0は、算出した合計実行性能を出力する。さらに、推定装置 1 4 0は、選択された複数の最適メモリ制御方式を時系列に並べた方式選択情報を出力してもよい。さらに、推定装置 1 4 0は、メモリ制御方式の切り換え時刻も出力してもよい。推定装置 1 4 0は、S 1 1 9の処理を終えると、本フローを終了する。

【 0 1 5 4 】

図 2 5は、動作情報、実行性能、および、部分情報毎に決定されたメモリ制御方式を示す図である。推定装置 1 4 0は、動作情報の時系列データを時間方向に分割して、複数の部分情報を生成する。さらに、推定装置 1 4 0は、複数の部分情報毎に、複数のメモリ制御方式のそれぞれについての実行性能を推定する。図 2 5の例においては、推定装置 1 4 0は、第 1 方式、第 2 方式および第 3 方式のそれぞれについて、実行性能を推定している。

30

【 0 1 5 5 】

続いて、推定装置 1 4 0は、複数の部分情報毎に、推定した実行性能に基づき最適メモリ制御方式を選択する。そして、推定装置 1 4 0は、複数の部分情報について選択した複数の最適メモリ制御方式についての複数の実行性能を合計した合計実行性能を算出する。

【 0 1 5 6 】

図 2 6は、アプリケーションプログラムの種類毎および不揮発記憶部 1 6の種類毎の処理時間の推定値の一例を示す図である。

40

【 0 1 5 7 】

例えば、現在、第 1 記憶部 1 4のみを主記憶装置として備える第 1 情報処理システム 1 1 0は、現行アプリケーションプログラムを実行している。例えば、画像認識処理のアプリケーションを例に説明する。設計者は、将来、現行アプリケーションプログラムに代えて、現行アプリケーションより認識精度が高い（高性能な）結果が得られるアルゴリズムが実装されている第 1 の新アプリケーションプログラム、または、第 1 の新アプリケーションプログラムよりも高性能の第 2 の新アプリケーションプログラムの採用を検討している。さらに、アプリケーションプログラムの変更に伴い、ハードウェアを、不揮発記憶部

50

16を備えた第2情報処理システム120へ変更することも検討している。このような場合、設計者は、推定装置140を用いて、第1の新アプリケーションプログラムまたは第2の新アプリケーションプログラムの何れを採用するか、および、第2情報処理システム120の構成をどうすればよいかを検討することができる。

【0158】

以上の検討をする場合、まず、推定装置140は、第1情報処理システム110に第1の新アプリケーションプログラムを実行させ、第1情報処理システム110から動作情報の時系列データを取得する。そして、推定装置140は、低速の不揮発記憶部16を備えた第2情報処理システム120が第1の新アプリケーションプログラムを実行した場合の実行時間と、高速の不揮発記憶部16を備えた第2情報処理システム120が第1の新アプリケーションプログラムを実行した場合の実行時間とを推定する。

10

【0159】

続いて、推定装置140は、第1情報処理システム110に第2の新アプリケーションプログラムを実行させ、第1情報処理システム110から動作情報の時系列データを取得する。そして、推定装置140は、低速の不揮発記憶部16を備えた第2情報処理システム120が第2の新アプリケーションプログラムを実行した場合の実行時間と、高速の不揮発記憶部16を備えた第2情報処理システム120が第2の新アプリケーションプログラムを実行した場合の実行時間とを推定する。

【0160】

そして、設計者は、このように推定された各実行時間を参照して、第1の新アプリケーションプログラムまたは第2の新アプリケーションプログラムの何れを採用するかを決定する。さらに、設計者は、低速の不揮発記憶部16を備えた第2情報処理システム120または高速の不揮発記憶部16を備えた第2情報処理システム120の何れを採用するかを決定する。

20

【0161】

ここで、設計者は、第1情報処理システム110により現行アプリケーションプログラムを実行させた場合の実行時間を基準の処理時間とし、この基準の処理時間よりも推定された処理時間が短いことを第1の条件とした。さらに、設計者は、第1の条件を満たすなかで、コストが低く認識性能が高い(高性能)ことを第2の条件とした。

【0162】

第2の新アプリケーションプログラムは、第1の新アプリケーションプログラムよりも認識性能が高い(高性能)ので条件をクリアすれば、前者がのぞましい。低速の不揮発記憶部16は、高速の不揮発記憶部16よりもコストが低い。従って、図26の例のような結果が得られた場合、第1の条件および第2の条件を満たすのは、第2の新アプリケーションプログラムを、低速の不揮発記憶部16を備える第2情報処理システム120により実行した場合となる。これは、例えば、第2の新アプリケーションプログラムの方が使用するメモリが大きいため、第1記憶部14のみの処理では、第1記憶部14とOSの仮想記憶のスワップ領域との間のデータの入れ替え処理であるスワップ処理が多く、第2の新アプリケーションプログラムの方が遅かった。しかし、第2の新アプリケーションプログラムの処理の方が、大容量メモリを利用すると性能が出やすい処理だった。このため、第2の新アプリケーションプログラムの方が、第1の新アプリケーションプログラムよりも認識性能が高い(高性能)のにもかかわらず、不揮発記憶部16と組み合わせて処理する場合は、逆に処理時間が短く、低速の不揮発記憶部16の利用でも処理時間の基準を満たしたという例である。このように設計者は、推定装置140により推定された結果を参照することにより、将来実行するアプリケーションプログラムおよびシステムのハードウェア構成を決定することができる。

30

40

【0163】

図27は、不揮発記憶部16の種類毎およびメモリ制御方式毎の処理時間の推定値の一例を示す図である。推定装置140は、メモリ制御方式毎に、合成実行性能を算出してもよい。この場合、推定装置140は、部分情報毎に最適メモリ制御方式を選択する処理は

50

実行せずに、メモリ制御方式毎に合計の実行性能を算出する。

【0164】

図27において、推定された実行時間を表す棒グラフに付けられている比率は、下記のように、第1アクセス処理に設定されたページ数の割合(a)と第2アクセス処理に設定されたページ数の割合(b)との比率(a:b)を表す。

【0165】

例えば、設計者は、第1情報処理システム110により現行アプリケーションプログラムを実行させた場合の実行時間を基準の処理時間とし、この基準の処理時間よりも推定された処理時間が短いことを第1の条件とした。さらに、設計者は、第1の条件を満たすなかで、コストが最も低いことを第2の条件とした。

10

【0166】

第2情報処理システム120は、第1アクセス処理に設定されたページの比率が多い方が、使用する第1記憶部14の容量(DRAM量)が大きくなるため消費電力が大きくなり、実行コストが高い。また、低速の不揮発記憶部16は、高速の不揮発記憶部16よりもコストが低い。従って、図27の例のような結果が得られた場合、第1の条件および第2の条件を満たすのは、新アプリケーションプログラムを、低速の不揮発記憶部16を備える第2情報処理システム120により、比率が2:8のメモリ制御方式で実行した場合となる。このように設計者は、推定装置140により推定された結果を参照することにより、将来実行するシステムのハードウェア構成およびメモリ制御方式を決定することができる。

20

【0167】

以上のように、第4実施形態に係る推定システム130は、第1情報処理システム110がアプリケーションプログラムを実行した場合に得られる動作情報から、第1情報処理システム110のハードウェア構成を変更した第2情報処理システム120の実行性能を推定する。この場合において、推定システム130は、動作情報の時系列データを時間方向に分割した複数の部分情報を生成し、複数の部分情報のそれぞれ毎に実行性能を推定する。これにより、推定システム130は、精度良く第2情報処理システム120の実行性能を推定することができる。

【0168】

(第5実施形態)

つぎに、第5実施形態について説明する。第5実施形態は、第4実施形態の変形例である。

30

【0169】

図28は、第5実施形態に係る推定装置140の構成を示す図である。第5実施形態に係る推定装置140は、第5実施形態に係る推定装置140と比較して、修正部152をさらに有する。

【0170】

修正部152は、選択部48により、複数の部分情報について選択された複数の最適メモリ制御方式を取得する。修正部152は、メモリ制御方式の切り替えにより生じるオーバヘッド実行性能を加えた合計実行性能が最適となるように、複数の部分情報について選択された複数の最適メモリ制御方式を修正する。オーバヘッド実行性能は、メモリ制御方式の切り替えにより生じるオーバヘッドの実行性能(例えば、実行時間、消費電力または寿命減少量)である。そして、修正部152は、修正した後における、複数の最適メモリ制御方式を出力部146に与える。

40

【0171】

出力部146は、修正後における複数の最適メモリ制御方式についての複数の実行性能を取得する。そして、出力部146は、修正後における複数の最適メモリ制御方式についての複数の実行性能を合計した合計実行性能を算出する。

【0172】

図29は、メモリ制御方式の切り替え時におけるオーバヘッド時間を考慮した場合の実

50

行時間の一例を示す図である。例えば、推定部 4 6 は、実行性能として、第 2 情報処理システム 1 2 0 においてアプリケーションプログラムを対応するメモリ制御方式で実行した場合における実行時間を推定する。この場合、修正部 1 5 2 は、メモリ制御方式の切り替えにより生じるオーバヘッド時間を加えた合計実行時間が短くなるように、複数の部分情報について選択された複数の最適メモリ制御方式を修正する。

【 0 1 7 3 】

例えば、第 1 方式 第 2 方式 第 1 方式と切り替わる場合について考える。オーバヘッドを考慮せずにメモリ制御方式の切り換えをした場合には、合計実行時間は、図 2 9 の A のようになる。

【 0 1 7 4 】

一方、切り替えをしない場合、つまり、第 1 方式を継続する場合、合計実行時間は、図 2 9 の B のようになる。第 1 方式よりも第 2 方式の方が実行時間が短いので、A と B とを比較すると、A の方が合計実行時間は短くなる。

【 0 1 7 5 】

ここで、第 1 方式から第 2 方式への切り替わりにより生じるオーバヘッド時間および第 2 方式から第 1 方式への切り替わりにより生じるオーバヘッド時間の合計値が、第 1 方式の実行時間と第 2 方式の実行時間の差より小さい場合（第 1 オーバヘッド時間の場合）、図 2 9 の C のように、切り替えを行った方が合計実行時間が短くなる。しかし、合計値が第 1 方式の実行時間と第 2 方式の実行時間の差より大きい場合（第 2 オーバヘッド時間の場合）、図 2 9 の D のように、切り替えを行った方が合計実行時間が長くなる。

【 0 1 7 6 】

従って、修正部 1 5 2 は、第 1 方式 第 2 方式 第 1 方式と切り替わる場合、第 1 方式から第 2 方式への切り替わりにより生じるオーバヘッド時間および第 2 方式から第 1 方式への切り替わりにより生じるオーバヘッド時間の合計値が、第 1 方式の実行時間から第 2 方式の実行時間を減じた値より大きい場合、図 2 9 の B のように、切り替えを行わずに、第 1 方式を継続する。これにより、修正部 1 5 2 は、メモリ制御方式の切り替えにより生じるオーバヘッド時間を加えた合計実行時間が短くなるように、複数の部分情報について選択された複数の最適メモリ制御方式を修正することができる。

【 0 1 7 7 】

図 3 0 は、選択部 4 8 による選択後のメモリ制御方式および修正部 1 5 2 による修正後のメモリ制御方式を示す図である。例えば、推定部 4 6 は、実行性能として、第 2 情報処理システム 1 2 0 においてアプリケーションプログラムを対応するメモリ制御方式で実行した場合における実行時間を推定する。この場合、修正部 1 5 2 は、選択された複数の最適メモリ制御方式の中で、2 つの第 1 メモリ制御方式に挟まれた第 2 メモリ制御方式の実行時間が、後段の第 1 メモリ制御方式の実行時間の所定数倍より短い場合、第 2 メモリ制御方式を第 1 メモリ制御方式に置き換える。

【 0 1 7 8 】

例えば、図 3 0 における、R 1、R 3、R 5、R 7、R 9 および R 1 1 のメモリ制御方式は、第 1 方式である。R 2、R 4、R 6、R 8 および R 1 0 のメモリ制御方式は、第 2 方式である。R 2、R 3、R 4、R 5、R 6、R 7、R 8、R 9 および R 1 1 の実行時間は、1 である。R 1 0 の実行時間は、8 である。

【 0 1 7 9 】

ここで、修正部 1 5 2 は、2 つの第 1 メモリ制御方式に挟まれた第 2 メモリ制御方式の実行時間が、後段の第 1 メモリ制御方式の実行時間の 2 倍より短い場合、第 2 メモリ制御方式を第 1 メモリ制御方式に置き換えるとする。この場合、R 1 および R 3 に挟まれた R 2 の実行時間は 1 であり、R 1 および R 3 のうちの後段の R 3 の実行時間は 1 である。従って、R 2 の実行時間は、R 3 の実行時間の 2 倍よりも短いので、修正部 1 5 2 は、R 2 のメモリ制御方式を第 2 方式から第 1 方式に置き換える。同様に、修正部 1 5 2 は、R 4、R 6 および R 8 のメモリ制御方式も、第 2 方式から第 1 方式に置き換える。これにより、修正部 1 5 2 は、切り替えにより生じるオーバヘッド時間を加えた場合の合計実行時間

10

20

30

40

50

が、切り替えをしない場合の実行時間より長くなる場合には、切り換えしないようにメモリ制御方式を修正することができる。

【0180】

一方、R9およびR11に挟まれたR10の実行時間は8であり、R9およびR11のうちの後段のR11の実行時間は1である。従って、R10の実行時間は、R1の実行時間の2倍よりも長いので、修正部152は、R10のメモリ制御方式を、そのまま第2方式とする。これにより、修正部152は、切り替えにより生じるオーバヘッド時間を加えた場合の合計実行時間が、切り替えをしない場合の実行時間より短い場合には、切り換えするようにメモリ制御方式を修正することができる。

【0181】

従って、修正部152は、メモリ制御方式の切り替えにより生じるオーバヘッド時間を加えた合計実行時間が短くなるように、複数の部分情報のそれぞれについて選択されたメモリ制御方式を修正することができる。

【0182】

以上のように、第5実施形態に係る推定システム130は、メモリ制御方式の切り替えにより生じるオーバヘッド実行性能を考慮して、第2情報処理システム120の実行性能を推定する。これにより、第5実施形態に係る推定システム130によれば、より精度良く第2情報処理システム120の実行性能を推定することができる。

【0183】

(第6実施形態)

つぎに、第6実施形態について説明する。第6実施形態は、第3実施形態の変形例である。

【0184】

図31は、第6実施形態に係る方式決定部36の構成を推定モデル記憶部26、結果記憶部62および決定モデル記憶部64とともに示す図である。第6実施形態に係る情報処理装置18は、第3実施形態に係る方式決定部36と比較して、修正部152をさらに有する。

【0185】

結果記憶部62は、過去の動作情報の変化の時系列データを記憶する。さらに、結果記憶部62は、過去に選択された複数のメモリ制御方式の履歴を記憶する。結果記憶部62は、動作情報の変化の時系列データと、選択された複数のメモリ制御方式とを時系列に対応付けて記憶する。

【0186】

修正部152は、結果記憶部62に記憶された、選択された複数のメモリ制御方式の履歴を読み出す。修正部152は、時系列に選択された複数のメモリ制御方式を、メモリ制御方式の切り替えにより生じるオーバヘッド実行性能を加えた合計実行性能が最適となるように修正する。例えば、修正部152は、第5実施形態と同様の処理を実行して、メモリ制御方式を修正する。そして、修正部152は、修正後における時系列に選択された複数のメモリ制御方式を学習部68に与える。

【0187】

学習部68は、動作情報の変化パターン(変化の周期性)と、修正後におけるメモリ制御方式との関係に基づき、動作情報の変化パターンからメモリ制御方式を決定するための決定モデルを生成する。例えば、学習部68は、学習処理またはクラスタリング処理等を実行することにより動作情報の複数の変化パターン(変化の周期性)のそれぞれと、メモリ制御方式との対応関係を検出する。そして、学習部68は、複数の変化パターン(周期性)のそれぞれとメモリ制御方式との対応関係に基づき、決定モデルを生成する。学習部68は、生成した決定モデルを決定モデル記憶部64に書き込む。

【0188】

決定部70は、変化検出部52が動作情報から所定の変化を検出した時点における、動作情報の変化パターンを取得する。例えば、決定部70は、変化検出部52が動作情報か

10

20

30

40

50

ら所定の変化を検出した時点における、動作情報の変化方向のパターンを取得する。そして、決定部 70 は、取得した動作情報の変化パターン（変化の周期性）が、決定モデルに示された変化パターンと一致した場合には、メモリ制御方式を決定する。決定部 70 は、決定したメモリ制御方式を設定部 50 に通知する。

【0189】

図 3 2 は、実際に選択されたメモリ制御方式および修正後のメモリ制御方式を示す図である。例えば、アプリケーションプログラムを実行した場合、情報処理システム 10 は、複数の種類のデータに対して処理を実行する。例えば、顔認識処理をする場合、情報処理システム 10 は、複数人の顔画像データに対して処理を実行する。

【0190】

このような場合、情報処理システム 10 は、複数の顔画像データ毎に近似した動作情報を出力する。この結果、情報処理システム 10 は、同一の時系列パターンで複数のメモリ制御方式を選択する。従って、情報処理システム 10 が複数のデータ（例えば、第 1 データおよび第 2 データ）に対する処理を実行した場合、修正部 152 は、複数のデータのそれぞれについて選択された複数のメモリ制御方式を、同様に修正することができる。

【0191】

図 3 3 は、修正後のメモリ制御方式と動作情報の変化のパターンとの関係を示す図である。学習部 68 は、動作情報の変化パターン（変化の周期性）と、修正後におけるメモリ制御方式との関係を検出する。図 3 3 の例であれば、学習部 68 は、第 1 方式に設定された状態において、動作情報に 5 個の変化点が発生した場合、次に選択されるメモリ制御方式が第 2 方式であることを学習できる。また、図 3 3 の例であれば、学習部 68 は、第 2 方式に設定された状態において、動作情報に 1 個の変化点が発生した場合、次に選択されるメモリ制御方式が第 1 方式であることが学習できる。

【0192】

従って、学習部 68 は、このような変化点の発生パターンと次に選択されるメモリ制御方式との関係を定義した決定モデルを生成することができる。

【0193】

以上のように、第 6 実施形態に係る情報処理システム 10 は、過去の動作情報の変化パターンと、修正後におけるメモリ制御方式との関係に基づき、動作情報の変化パターンからメモリ制御方式を決定するための決定モデルを生成する。そして、情報処理システム 10 は、動作情報の変化パターンからメモリ制御方式を決定する。これにより、第 6 実施形態に係る情報処理システム 10 によれば、実行性能を推定せずにメモリ制御方式を決定することができる。従って、情報処理システム 10 によれば、少ない演算量でメモリ制御方式を決定することができる。

【0194】

（情報処理装置 200 のハードウェア構成）

図 3 4 は、情報処理装置 200 のハードウェアブロック図である。情報処理装置 200 は、一例として、一般のコンピュータと同様のハードウェア構成により実現される。情報処理装置 200 は、所定プログラムを実行することにより、推定装置 140 として機能することができる。

【0195】

情報処理装置 200 は、CPU 202 と、ROM (Read Only Memory) 204 と、RAM (Random Access Memory) 206 と、操作部 208 と、表示部 210 と、通信装置 212 と、記憶装置 214 とを備える。各部は、バスにより接続される。

【0196】

CPU 202 は、情報処理を実行するプロセッサであって、記憶装置 214 に記憶されたプログラムを RAM 206 に展開して実行し、各部を制御して入出力を行ったり、データの加工を行ったりする。CPU 202 は、1 または複数のプロセッサにより構成されていてもよい。また、情報処理装置 200 は、プログラムを実行することが可能であれば、

10

20

30

40

50

C P U 2 0 2に限らず他のプロセッサを備えてもよい。R O M 2 0 4には、起動用プログラムを記憶装置 2 1 4 から R A M 2 0 6 に読み出すスタートプログラムが記憶されている。R A M 2 0 6 は、C P U 2 0 2 の作業領域としてデータを記憶する。

【 0 1 9 7 】

操作部 2 0 8 は、マウスまたはキーボード等の入力デバイスであって、ユーザから操作入力された情報を指示信号として受け付け、その指示信号を C P U 2 0 2 に出力する。表示部 2 1 0 は、例えば L C D (L i q u i d C r y s t a l D i s p l a y) 等の表示装置である。表示部 2 1 0 は、C P U 2 0 2 からの表示信号に基づいて、各種情報を表示する。通信装置 2 1 2 は、ネットワーク等を介して、外部機器と情報をやり取りする。記憶装置 2 1 4 は、例えば、ハードディスクドライブまたはフラッシュメモリ等である。記憶装置 2 1 4 は、情報処理装置 2 0 0 で実行されるプログラム、および、オペレーティングシステムを記憶している。

10

【 0 1 9 8 】

本実施形態の情報処理装置 2 0 0 で実行されるプログラムは、インストール可能な形式または実行可能な形式のファイルで C D - R O M 、フレキシブルディスク (F D) 、 C D - R 、 D V D 等のコンピュータで読み取り可能な記録媒体に記録されて提供される。また、本実施形態の情報処理装置 2 0 0 で実行されるプログラムを、インターネット等のネットワークに接続されたコンピュータ上に格納し、ネットワーク経由でダウンロードさせることにより提供するように構成してもよい。また、本実施形態の情報処理装置 2 0 0 で実行されるプログラムをインターネット等のネットワーク経由で提供または配布するように構成してもよい。また、本実施形態のプログラムを、R O M 2 0 4 等に予め組み込んで提供するように構成してもよい。

20

【 0 1 9 9 】

情報処理装置 2 0 0 を推定装置 1 4 0 として機能させるためのプログラムは、取得モジュールと、分割モジュールと、推定モジュールと、選択モジュールと、出力モジュールとを有する。情報処理装置 2 0 0 は、プロセッサ (C P U 2 0 2) が記憶媒体 (記憶装置 2 1 4 等) からプログラムを読み出して実行することにより各モジュールが主記憶装置 (R A M 2 0 6) 上にロードされ、プロセッサ (C P U 2 0 2) が、取得部 4 4 、分割部 1 4 4 、推定部 4 6 、選択部 4 8 および出力部 1 4 6 として機能する。また、R A M 2 0 6 または記憶装置 2 1 4 等が、動作情報記憶部 1 4 2 および推定モデル記憶部 2 6 として機能する。なお、これらの一部または全部がプロセッサ以外のハードウェアにより実現されてもよい。

30

【 0 2 0 0 】

本発明のいくつかの実施形態を説明したが、これらの実施形態は、例として提示したものであり、発明の範囲を限定することは意図していない。これら新規な実施形態は、その他の様々な形態で実施されることが可能であり、発明の要旨を逸脱しない範囲で、種々の省略、置き換え、変更を行うことができる。これら実施形態やその変形は、発明の範囲や要旨に含まれるとともに、請求の範囲に記載された発明とその均等の範囲に含まれる。

【 符号の説明 】

【 0 2 0 1 】

40

- 1 0 情報処理システム
- 1 4 第 1 記憶部
- 1 6 不揮発記憶部
- 1 8 情報処理装置
- 2 0 処理回路
- 2 2 管理装置
- 2 4 動作情報検出部
- 2 6 推定モデル記憶部
- 3 2 データ処理部
- 3 4 アクセス管理部

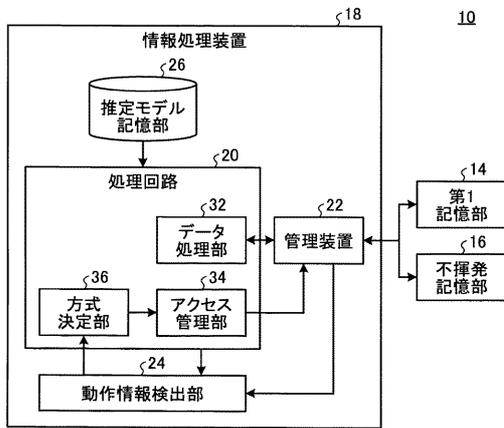
50

- 3 6 方式決定部
- 4 2 タイミング発生部
- 4 4 取得部
- 4 6 推定部
- 4 8 選択部
- 5 0 設定部
- 5 2 変化検出部
- 6 2 結果記憶部
- 6 4 決定モデル記憶部
- 6 8 学習部
- 7 0 決定部
- 7 2 モード設定部
- 1 1 0 第1情報処理システム
- 1 2 0 第2情報処理システム
- 1 3 0 推定システム
- 1 4 0 推定装置
- 1 4 2 動作情報記憶部
- 1 4 4 分割部
- 1 4 6 出力部
- 1 5 2 修正部

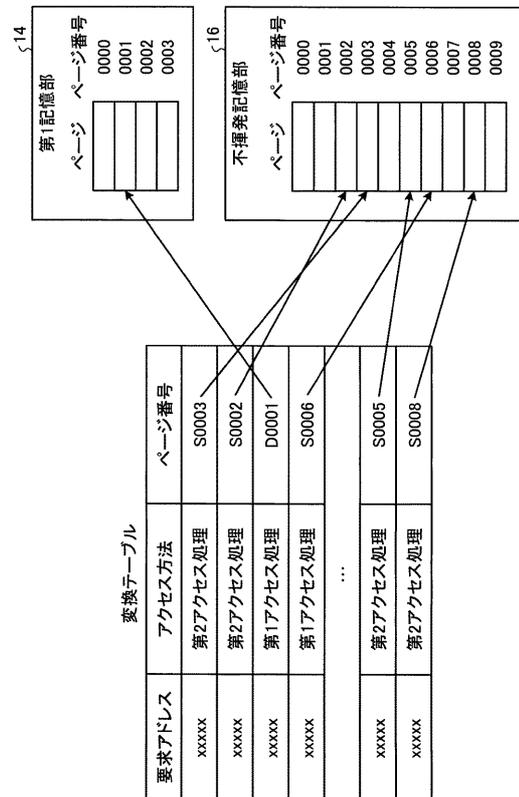
10

20

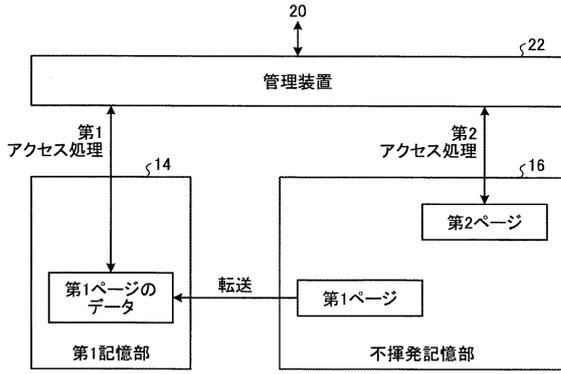
【図1】



【図2】



【 図 3 】



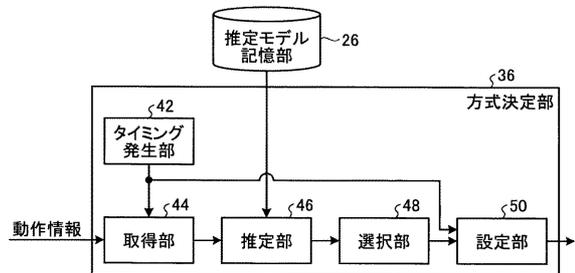
【 図 4 】

動作情報#6 (TLBミス回数)							
動作情報#5 (キャッシュミス数)							
動作情報#4 (キャッシュヒット数)							
動作情報#3 (プリフェッチ成功率)							
動作情報#2 (読み出し回数)							
動作情報#1 (書き込み回数)							
時刻	t_0	t_1	t_2	t_3	t_4	t_5	...

【 図 5 】

メモリ制御方式	制御内容
第1方式	100%のページが第2アクセス処理
第2方式	10%のページが第1アクセス処理 90%のページが第2アクセス処理
第3方式	20%のページが第1アクセス処理 80%のページが第2アクセス処理
第4方式	30%のページが第1アクセス処理 70%のページが第2アクセス処理

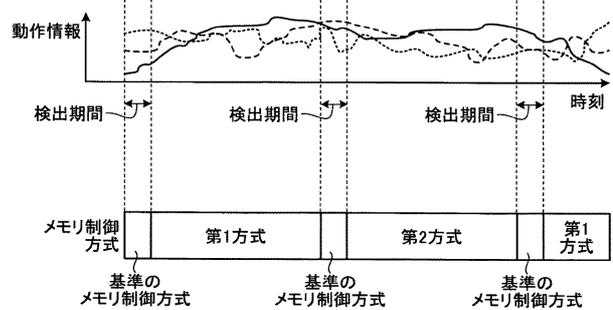
【 図 7 】



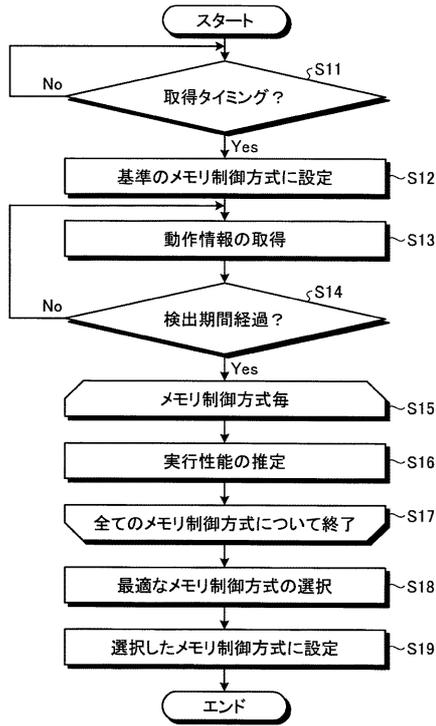
【 図 6 】

メモリ制御方式	第1制御内容	第2制御内容
第1方式	100%のページが第2アクセス処理	
第2方式	10%のページが第1アクセス処理 90%のページが第2アクセス処理	第1記憶部のメモリ使用量が不揮発記憶部の5%
第3方式	10%のページが第1アクセス処理 90%のページが第2アクセス処理	第1記憶部のメモリ使用量が不揮発記憶部の10%
第4方式	20%のページが第1アクセス処理 80%のページが第2アクセス処理	第1記憶部のメモリ使用量が不揮発記憶部の10%
第5方式	20%のページが第1アクセス処理 80%のページが第2アクセス処理	第1記憶部のメモリ使用量が不揮発記憶部の20%
第6方式	30%のページが第1アクセス処理 70%のページが第2アクセス処理	第1記憶部のメモリ使用量が不揮発記憶部の15%
第7方式	30%のページが第1アクセス処理 70%のページが第2アクセス処理	第1記憶部のメモリ使用量が不揮発記憶部の30%

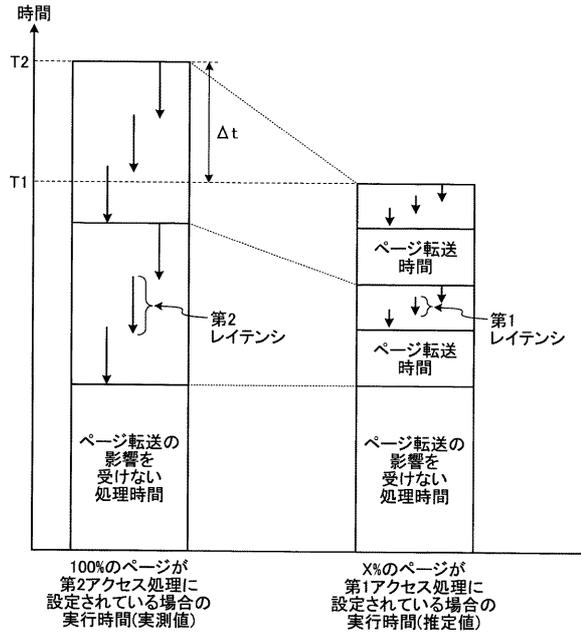
【 図 8 】



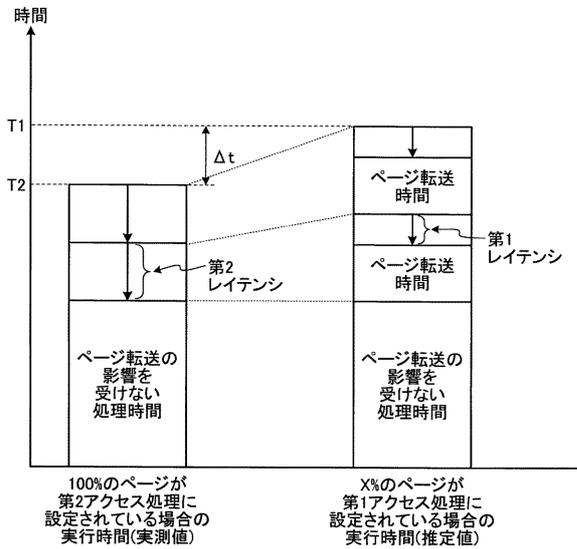
【 図 9 】



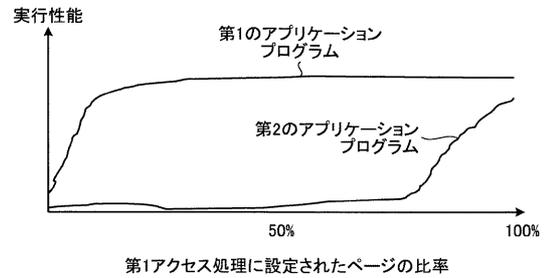
【 図 10 】



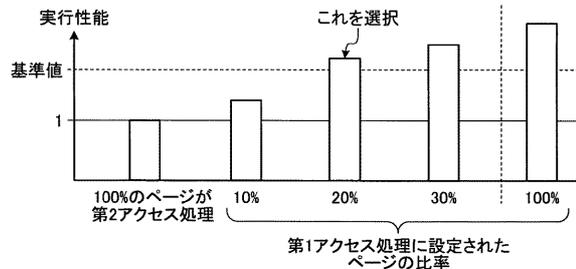
【 図 11 】



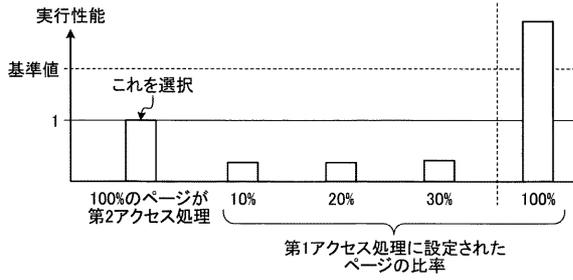
【 図 12 】



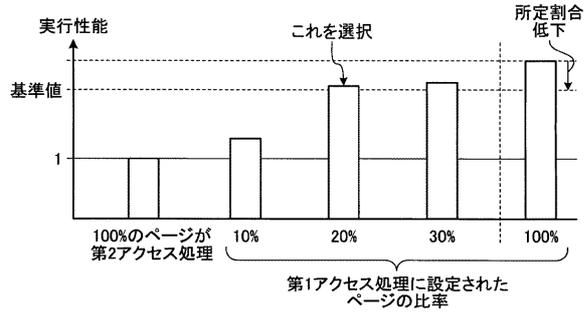
【 図 13 】



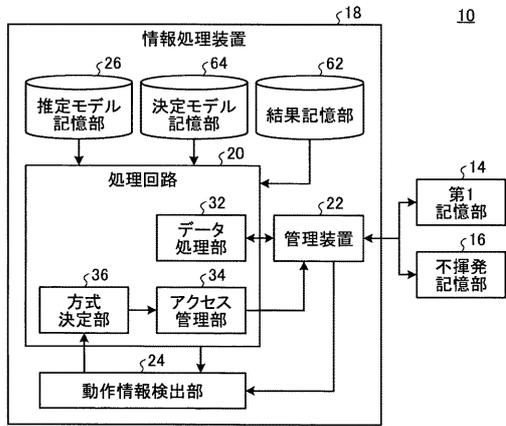
【図14】



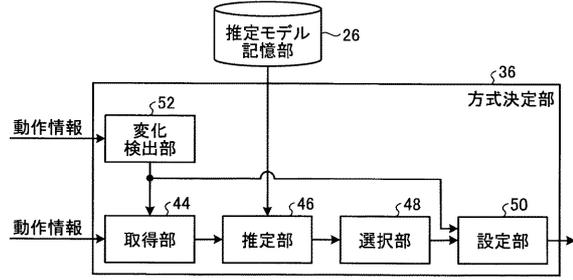
【図15】



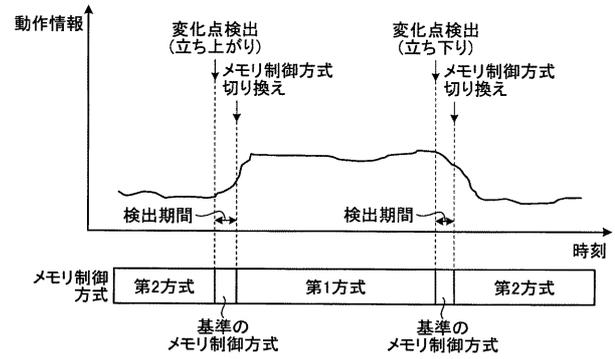
【図18】



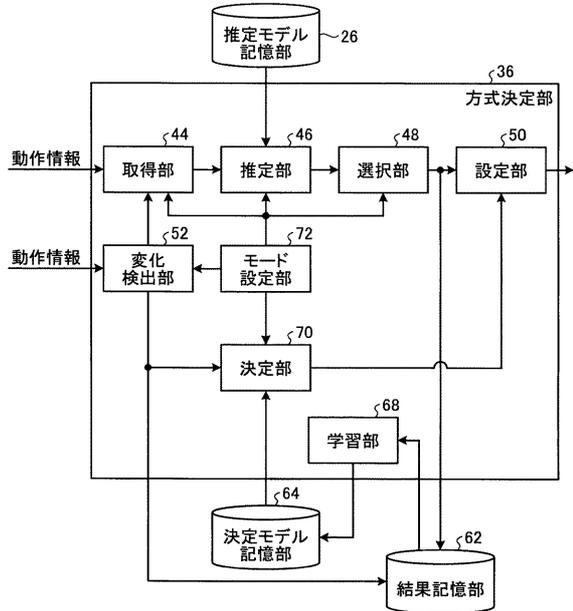
【図16】



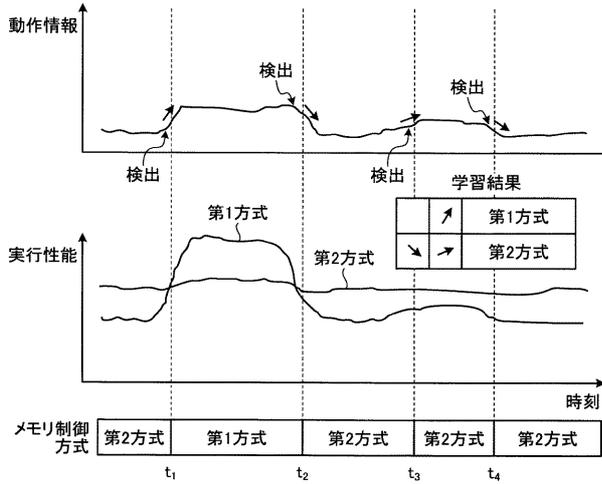
【図17】



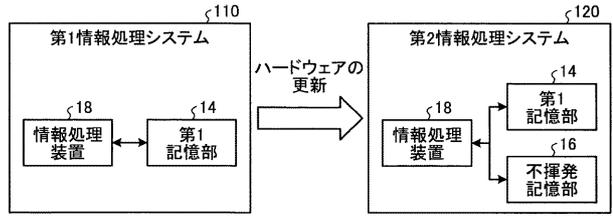
【図19】



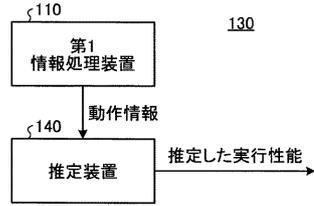
【図20】



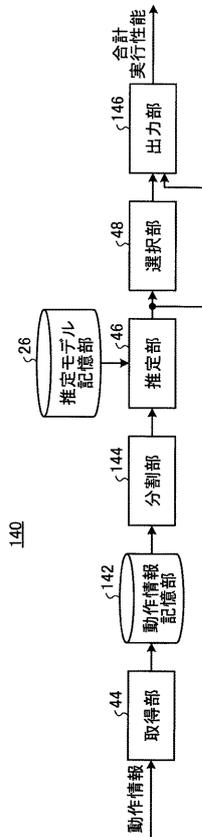
【図21】



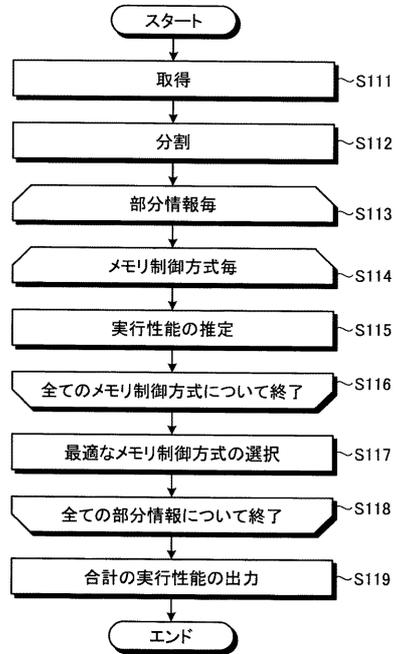
【図22】



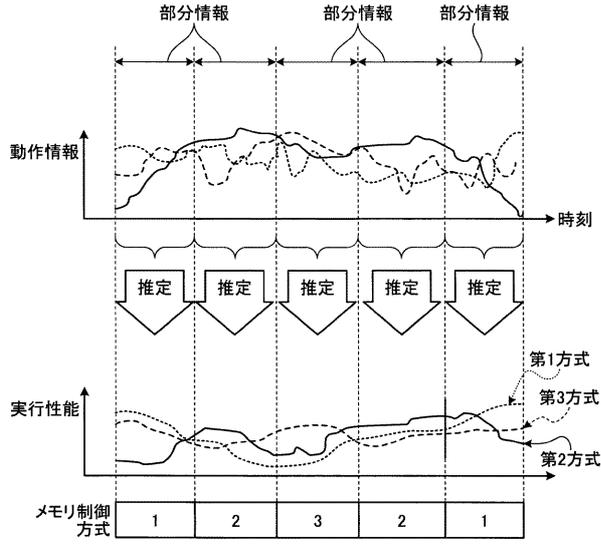
【図23】



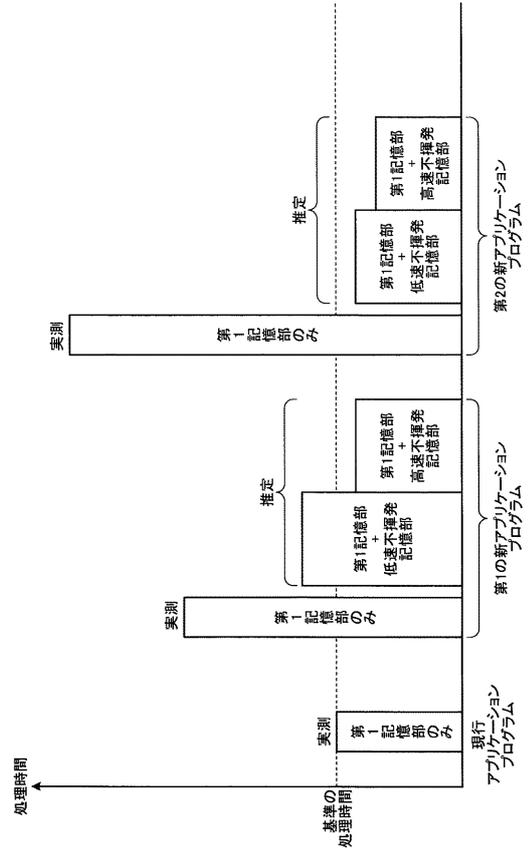
【図24】



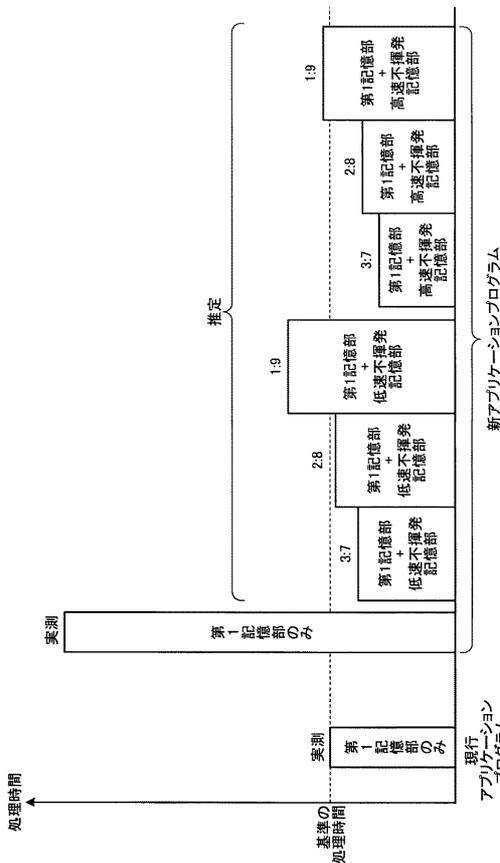
【図25】



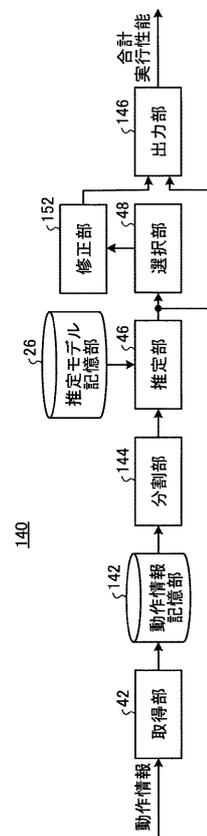
【図26】



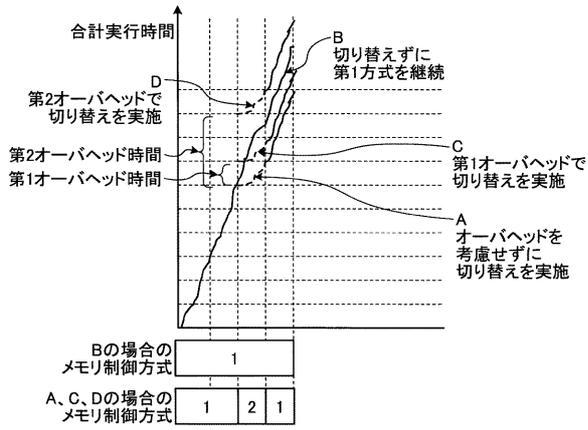
【図27】



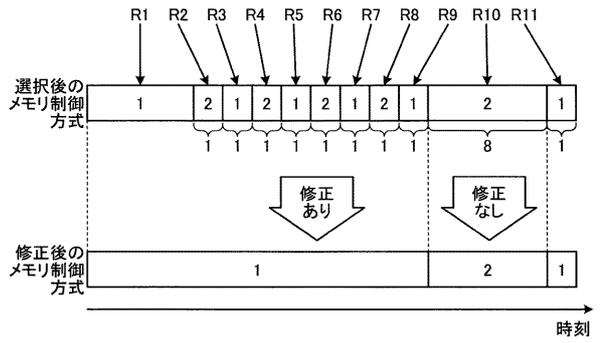
【図28】



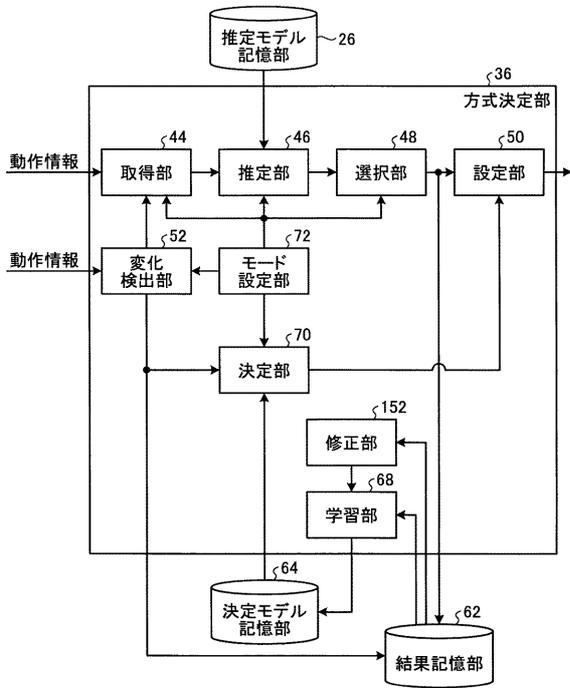
【 図 2 9 】



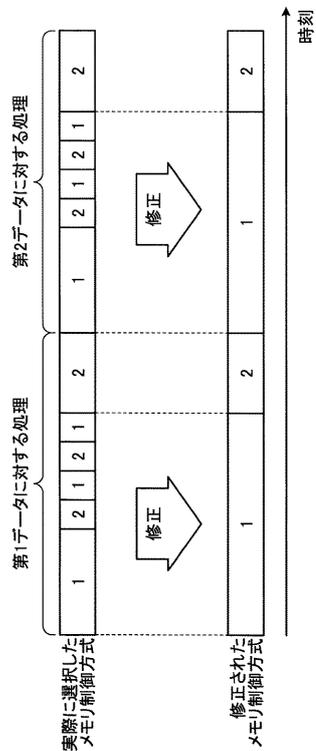
【 図 3 0 】



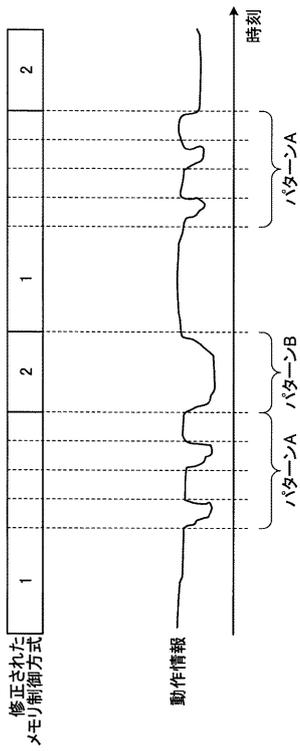
【 図 3 1 】



【 図 3 2 】

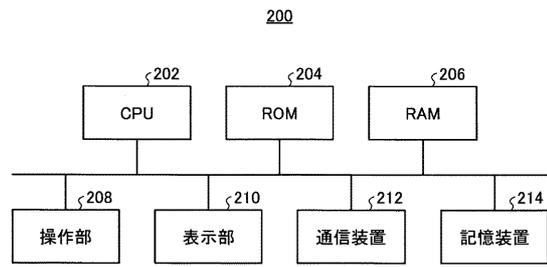


【 図 3 3 】



・パターンAを検出した場合、第2方式から第1方式に切り換え
 ・パターンBを検出した場合、第2方式から第1方式に切り換え

【 図 3 4 】



フロントページの続き

Fターム(参考) 5B042 GA15 GA25 GA33 JJ41 MA08 MA11 MA14 MC25 MC27
5B160 AA14 CA01 MM02