

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6194875号  
(P6194875)

(45) 発行日 平成29年9月13日(2017.9.13)

(24) 登録日 平成29年8月25日(2017.8.25)

(51) Int. Cl.		F I			
<b>G06F</b>	<b>3/06</b>	<b>(2006.01)</b>	G06F	3/06	302A
<b>G06F</b>	<b>12/08</b>	<b>(2016.01)</b>	G06F	12/08	513
<b>G06F</b>	<b>12/12</b>	<b>(2016.01)</b>	G06F	12/08	515Z
			G06F	12/12	501

請求項の数 9 (全 37 頁)

(21) 出願番号	特願2014-250580 (P2014-250580)	(73) 特許権者	000004237
(22) 出願日	平成26年12月11日(2014.12.11)		日本電気株式会社
(65) 公開番号	特開2016-114967 (P2016-114967A)		東京都港区芝五丁目7番1号
(43) 公開日	平成28年6月23日(2016.6.23)	(74) 代理人	100109313
審査請求日	平成28年4月15日(2016.4.15)		弁理士 机 昌彦
		(74) 代理人	100124154
			弁理士 下坂 直樹
		(72) 発明者	梶本 洋平
			東京都港区芝五丁目7番1号
			日本電気株式会社内
		審査官	梶井 茂行

最終頁に続く

(54) 【発明の名称】 キャッシュ装置、キャッシュシステム、キャッシュ方法、及びキャッシュプログラム

(57) 【特許請求の範囲】

【請求項1】

記憶装置に接続され、複数の要求元に接続されるキャッシュ装置であって、

第1の前記要求元から第1の読み出し要求を受信し、前記第1の読み出し要求において指定された第1のデータを保持していない場合に、前記記憶装置に前記第1の読み出し要求を送信し、前記記憶装置から第1の読み出し応答を受信し、前記第1の読み出し応答を前記第1の要求元へ送信する一方、前記第1のデータを保持している場合に、前記第1のデータを含む第2の読み出し応答を前記第1の要求元へ送信するキャッシュ手段と、

前記第1のデータの指定に対応付けて前記第1の要求元を第1の履歴として記憶し、前記キャッシュ手段により前記第1のデータが保持されておらず且つ前記キャッシュ手段に前記第1のデータを保持するために十分な記憶領域が存在する場合に、前記キャッシュ手段に前記第1のデータを記憶させる一方、前記キャッシュ手段により前記第1のデータが保持されておらず且つ前記キャッシュ手段に前記第1のデータを保持するために十分な記憶領域が存在しない場合に、前記第1のデータ又は前記キャッシュ手段に保持されたデータのうち、前記第1の履歴においてより少数の前記要求元に対応付けられたデータを前記第1の履歴においてより多数の前記要求元に対応付けられたデータに優先して第2のデータとして選択し、前記第2のデータと前記第1のデータとが異なるならば、前記キャッシュ手段に前記第2のデータを破棄させた後に前記第1のデータを記憶させる記憶領域制御手段と

を備えることを特徴とするキャッシュ装置。

## 【請求項 2】

請求項 1 に記載のキャッシュ装置と、

前記キャッシュ装置を介して前記記憶装置に接続され、それぞれ 1 台の要求元に接続される、複数の 2 次キャッシュ装置とを備えたキャッシュシステムあって、

前記 2 次キャッシュ装置は、前記第 1 の要求元から第 2 の読み出し要求を受信し、前記第 2 の読み出し要求において指定された第 3 のデータを保持していない場合に、前記キャッシュ装置を介して前記記憶装置に前記第 2 の読み出し要求を送信し、前記キャッシュ装置を介して前記記憶装置から第 3 の読み出し応答を受信し、前記第 3 の読み出し応答を前記第 1 の要求元へ送信する一方、前記第 3 のデータを保持している場合に、前記第 3 のデータを含む第 4 の読み出し応答を前記第 1 の要求元へ送信し、

10

前記 2 次キャッシュ装置は、前記第 3 のデータの指定に対応付けて前記第 1 の要求元を第 2 の履歴として記憶し、前記第 3 のデータが保持されておらず且つ前記第 3 のデータを保持するために十分な記憶領域が存在する場合に、前記第 3 のデータを記憶する一方、前記第 3 のデータが保持されておらず且つ前記第 3 のデータを保持するために十分な記憶領域が存在しない場合に、前記第 3 のデータ又は保持されたデータのうち、前記第 2 の履歴においてより少数回、前記要求元に対応付けられたデータを前記第 2 の履歴においてより多数回、前記要求元に対応付けられたデータに優先して第 4 のデータとして選択し、前記第 4 のデータと前記第 3 のデータとが異なるならば、前記第 4 のデータを破棄した後に前記第 3 のデータを記憶する

ことを特徴とするキャッシュシステム。

20

## 【請求項 3】

複数の前記要求元にそれぞれ 1 台の 2 次キャッシュ装置を介して接続され、

前記キャッシュ装置は、前記 2 次キャッシュ装置に前記第 1 の履歴の情報を送信する通信手段を更に備え、

前記 2 次キャッシュ装置は、前記第 1 の要求元から第 2 の読み出し要求を受信し、前記第 2 の読み出し要求において指定された第 3 のデータを保持していない場合に、前記キャッシュ装置を介して前記記憶装置に前記第 2 の読み出し要求を送信し、前記キャッシュ装置を介して前記記憶装置から第 3 の読み出し応答を受信し、前記第 3 の読み出し応答を前記第 1 の要求元へ送信する一方、前記第 3 のデータを保持している場合に、前記第 3 のデータを含む第 4 の読み出し応答を前記第 1 の要求元へ送信し、

30

前記 2 次キャッシュ装置は、前記第 3 のデータが保持されておらず且つ前記第 3 のデータを保持するために十分な記憶領域が存在する場合に、前記第 3 のデータを記憶する一方、前記第 3 のデータが保持されておらず且つ前記第 3 のデータを保持するために十分な記憶領域が存在しない場合に、前記キャッシュ装置から前記第 1 の履歴の情報を取得し、前記第 3 のデータ又は保持されたデータのうち、前記第 1 の履歴においてより多数の前記要求元に対応付けられたデータを前記第 1 の履歴においてより少数の前記要求元に対応付けられたデータに優先して第 5 のデータとして選択し、前記第 5 のデータと前記第 3 のデータとが異なるならば、前記第 5 のデータを破棄した後に前記第 3 のデータを記憶する

ことを特徴とする請求項 1 に記載のキャッシュ装置。

## 【請求項 4】

請求項 1 に記載のキャッシュ装置を介して記憶装置に接続され、それぞれ 1 台の要求元に接続された 2 次キャッシュ装置であって、

前記 2 次キャッシュ装置は、前記キャッシュ装置から前記第 1 の履歴の情報を取得し、前記第 1 の履歴に基づいて、破棄するデータを決定する

ことを特徴とする 2 次キャッシュ装置。

40

## 【請求項 5】

請求項 1 に記載のキャッシュ装置を介して記憶装置に接続され、それぞれ 1 台の要求元に接続された 2 次キャッシュ装置であって、

前記 2 次キャッシュ装置は、前記第 1 の要求元から第 2 の読み出し要求を受信し、前記第 2 の読み出し要求において指定された第 3 のデータを保持していない場合に、前記キャ

50

ッシュ装置を介して前記記憶装置に前記第2の読み出し要求を送信し、前記キャッシュ装置を介して前記記憶装置から第3の読み出し応答を受信し、前記第3の読み出し応答を前記第1の要求元へ送信する一方、前記第3のデータを保持している場合に、前記第3のデータを含む第4の読み出し応答を前記第1の要求元へ送信する2次キャッシュ手段と、

前記2次キャッシュ装置は、前記第3のデータが保持されておらず且つ前記第3のデータを保持するために十分な記憶領域が存在する場合に、前記第3のデータを記憶する一方、前記第3のデータが保持されておらず且つ前記第3のデータを保持するために十分な記憶領域が存在しない場合に、前記キャッシュ装置から前記第1の履歴の情報を取得し、前記第3のデータ又は保持されたデータのうち、前記第1の履歴においてより多数の前記要求元に対応付けられたデータを前記第1の履歴においてより少数の前記要求元に対応付けられたデータに優先して第5のデータとして選択し、前記第5のデータと前記第3のデータとが異なるならば、前記第5のデータを破棄した後に前記第3のデータを記憶する2次記憶領域制御手段と

を備えることを特徴とする2次キャッシュ装置。

【請求項6】

記憶装置に接続されたキャッシュ装置と、

前記キャッシュ装置を介して前記記憶装置に接続され、それぞれ1台の要求元に接続される、複数の2次キャッシュ装置とを備えたキャッシュシステムの制御方法であって、

前記キャッシュ装置により、

前記2次キャッシュ装置を介して第1の前記要求元から第1の読み出し要求を受信し、前記第1の読み出し要求において指定された第1のデータを保持していない場合に、前記記憶装置に前記第1の読み出し要求を送信し、前記記憶装置から第1の読み出し応答を受信し、前記第1の読み出し応答を前記2次キャッシュ装置を介して前記第1の要求元へ送信する一方、前記第1のデータを保持している場合に、前記第1のデータを含む第2の読み出し応答を前記2次キャッシュ装置を介して前記第1の要求元へ送信することと、

前記第1のデータの指定に対応付けて前記第1の要求元を第1の履歴として記憶し、前記キャッシュ手段により前記第1のデータが保持されておらず且つ前記キャッシュ手段に前記第1のデータを保持するために十分な記憶領域が存在する場合に、前記キャッシュ手段に前記第1のデータを記憶させる一方、前記キャッシュ手段により前記第1のデータが保持されておらず且つ前記キャッシュ手段に前記第1のデータを保持するために十分な記憶領域が存在しない場合に、前記第1のデータ又は前記キャッシュ手段に保持されたデータのうち、前記第1の履歴においてより少数の前記要求元に対応付けられたデータを前記第1の履歴においてより多数の前記要求元に対応付けられたデータに優先して第2のデータとして選択し、前記第2のデータと前記第1のデータとが異なるならば、前記キャッシュ手段に前記第2のデータを破棄させた後に前記第1のデータを記憶させることとを実行し、

前記2次キャッシュ装置により、

前記第1の要求元から第2の読み出し要求を受信し、前記第2の読み出し要求において指定された第3のデータを保持していない場合に、前記キャッシュ装置を介して前記記憶装置に前記第2の読み出し要求を送信し、前記キャッシュ装置を介して前記記憶装置から第3の読み出し応答を受信し、前記第3の読み出し応答を前記第1の要求元へ送信する一方、前記第3のデータを保持している場合に、前記第3のデータを含む第4の読み出し応答を前記第1の要求元へ送信することと、

前記第3のデータが保持されておらず且つ前記第3のデータを保持するために十分な記憶領域が存在する場合に、前記第3のデータを記憶する一方、前記第3のデータが保持されておらず且つ前記第3のデータを保持するために十分な記憶領域が存在しない場合に、前記キャッシュ装置から前記第1の履歴の情報を取得し、前記第3のデータ又は保持されたデータのうち、前記第1の履歴においてより多数の前記要求元に対応付けられたデータを前記第1の履歴においてより少数の前記要求元に対応付けられたデータに優先して第4のデータとして選択し、前記第4のデータと前記第3のデータとが異なるならば、前記

10

20

30

40

50

第4のデータを破棄した後に前記第3のデータを記憶することとを実行することを特徴とするキャッシュ方法。

【請求項7】

記憶装置に接続されたキャッシュ装置と、

前記キャッシュ装置を介して前記記憶装置に接続され、それぞれ1台の要求元に接続される、複数の2次キャッシュ装置とを備えたキャッシュシステムあって、

前記キャッシュ装置は、

前記2次キャッシュ装置を介して第1の前記要求元から第1の読み出し要求を受信し、前記第1の読み出し要求において指定された第1のデータを保持していない場合に、前記記憶装置に前記第1の読み出し要求を送信し、前記記憶装置から第1の読み出し応答を受信し、前記第1の読み出し応答を前記2次キャッシュ装置を介して前記第1の要求元へ送信する一方、前記第1のデータを保持している場合に、前記第1のデータを含む第2の読み出し応答を前記2次キャッシュ装置を介して前記第1の要求元へ送信するキャッシュ手段と、

前記第1のデータの指定に対応付けて前記第1の要求元を第1の履歴として記憶し、前記キャッシュ手段により前記第1のデータが保持されておらず且つ前記キャッシュ手段に前記第1のデータを保持するために十分な記憶領域が存在する場合に、前記キャッシュ手段に前記第1のデータを記憶させる一方、前記キャッシュ手段により前記第1のデータが保持されておらず且つ前記キャッシュ手段に前記第1のデータを保持するために十分な記憶領域が存在しない場合に、前記第1のデータ又は前記キャッシュ手段に保持されたデータのうち、前記第1の履歴においてより少数の前記要求元に対応付けられたデータを前記第1の履歴においてより多数の前記要求元に対応付けられたデータに優先して第2のデータとして選択し、前記第2のデータと前記第1のデータとが異なるならば、前記キャッシュ手段に前記第2のデータを破棄させた後に前記第1のデータを記憶させる記憶領域制御手段とを備え、

前記2次キャッシュ装置は、

前記第1の要求元から第2の読み出し要求を受信し、前記第2の読み出し要求において指定された第3のデータを保持していない場合に、前記キャッシュ装置を介して前記記憶装置に前記第2の読み出し要求を送信し、前記キャッシュ装置を介して前記記憶装置から第3の読み出し応答を受信し、前記第3の読み出し応答を前記第1の要求元へ送信する一方、前記第3のデータを保持している場合に、前記第3のデータを含む第4の読み出し応答を前記第1の要求元へ送信する2次キャッシュ手段と、

前記第3のデータが保持されておらず且つ前記第3のデータを保持するために十分な記憶領域が存在する場合に、前記第3のデータを記憶する一方、前記第3のデータが保持されておらず且つ前記第3のデータを保持するために十分な記憶領域が存在しない場合に、前記キャッシュ装置から前記第1の履歴の情報を取得し、前記第3のデータ又は保持されたデータのうち、前記第1の履歴においてより多数の前記要求元に対応付けられたデータを前記第1の履歴においてより少数の前記要求元に対応付けられたデータに優先して第4のデータとして選択し、前記第4のデータと前記第3のデータとが異なるならば、前記第4のデータを破棄した後に前記第3のデータを記憶する2次記憶領域制御手段とを備える

ことを特徴とするキャッシュシステム。

【請求項8】

記憶装置に接続され、複数の要求元に接続されるキャッシュ装置において、

第1の前記要求元から第1の読み出し要求を受信し、前記第1の読み出し要求において指定された第1のデータを保持していない場合に、前記記憶装置に前記第1の読み出し要求を送信し、前記記憶装置から第1の読み出し応答を受信し、前記第1の読み出し応答を前記第1の要求元へ送信する一方、前記第1のデータを保持している場合に、前記第1のデータを含む第2の読み出し応答を前記第1の要求元へ送信し、

前記第1のデータの指定に対応付けて前記第1の要求元を第1の履歴として記憶し、前

10

20

30

40

50

記第 1 のデータが保持されておらず且つ前記第 1 のデータを保持するために十分な記憶領域が存在する場合に、前記第 1 のデータを記憶する一方、前記第 1 のデータが保持されておらず且つ前記第 1 のデータを保持するために十分な記憶領域が存在しない場合に、前記第 1 のデータ又は保持されたデータのうち、前記第 1 の履歴においてより少数の前記要求元に対応付けられたデータを前記第 1 の履歴においてより多数の前記要求元に対応付けられたデータに優先して第 2 のデータとして選択し、前記第 2 のデータと前記第 1 のデータとが異なるならば、前記第 2 のデータを破棄した後に前記第 1 のデータを記憶することを特徴とするキャッシュ方法。

【請求項 9】

記憶装置に接続され、複数の要求元に接続されるキャッシュ装置であって、

10

第 1 の前記要求元から第 1 の読み出し要求を受信し、前記第 1 の読み出し要求において指定された第 1 のデータを保持していない場合に、前記記憶装置に前記第 1 の読み出し要求を送信し、前記記憶装置から第 1 の読み出し応答を受信し、前記第 1 の読み出し応答を前記第 1 の要求元へ送信する一方、前記第 1 のデータを保持している場合に、前記第 1 のデータを含む第 2 の読み出し応答を前記第 1 の要求元へ送信するキャッシュ処理と、

前記第 1 のデータの指定に対応付けて前記第 1 の要求元を第 1 の履歴として記憶し、前記第 1 のデータが保持されておらず且つ前記第 1 のデータを保持するために十分な記憶領域が存在する場合に、前記第 1 のデータを記憶する一方、前記第 1 のデータが保持されておらず且つ前記第 1 のデータを保持するために十分な記憶領域が存在しない場合に、前記第 1 のデータ又は保持されたデータのうち、前記第 1 の履歴においてより少数の前記要求元に対応付けられたデータを前記第 1 の履歴においてより多数の前記要求元に対応付けられたデータに優先して第 2 のデータとして選択し、前記第 2 のデータと前記第 1 のデータとが異なるならば、前記第 2 のデータを破棄した後に前記第 1 のデータを記憶する記憶領域制御処理と

20

をコンピュータに実行させることを特徴とするキャッシュプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、キャッシュされたデータのヒット率を向上させるキャッシュ装置、キャッシュシステム、キャッシュ方法、及びキャッシュプログラムに関する。

30

【背景技術】

【0002】

データの読み出しを行う装置の多くでは、使用頻度の高いデータを高速な記憶装置に一時的に保持することにより、毎回、低速な記憶装置からデータを読み出す無駄を省いてデータの読み出しを高速化する「キャッシュ」技術が利用される。データの書き込みについても、「キャッシュ」技術が利用されることが多い。また、「キャッシュ」のために使われる高速な記憶装置や、複製されたデータも「キャッシュ」と呼ばれることがある。近年、キャッシュを高機能化又は高性能化するための技術が開発されている。

【0003】

キャッシュを高性能化するための技術の一例が、特許文献 1 に開示されている。

40

【0004】

特許文献 1 のストレージサブシステムは、書き込み速度に優劣がある記憶デバイスに対して、書き込み速度毎に書き込みキャッシュを分けることにより、速い記憶デバイスへの書き込み時間に対する、遅い記憶デバイスへの書き込み時間の影響を低減する。

【0005】

上記の動作の結果、特許文献 1 のストレージサブシステムでは、キャッシュから各記憶デバイスへのデータの書き込み速度に違いがある場合であっても、データの書き込み処理が停滞しない。

【0006】

キャッシュを高性能化するための技術の別の一例が、特許文献 2 に開示されている。

50

## 【 0 0 0 7 】

特許文献 2 の印刷文書処理システムは、複数のラスターイメージ生成部から出力された画像データをキャッシュに保持しつつ、プリンタに出力する。印刷文書処理システムは、キャッシュの残容量が少ない場合に、画像データをキャッシュに保持するか否かを、画像データのスコアにより決定する。画像データのスコアは、オブジェクトの種類若しくは複雑さ、又はラスターイメージ生成時の負荷、所要時間、データサイズ、色空間変換の要否、拡大率、回転の有無、グラデーション処理の有無、若しくはキャッシュ問い合わせ累積回数等により決定される。

## 【 0 0 0 8 】

上記の動作の結果、特許文献 2 の印刷文書処理システムは、画像データのスコアが適切に定義されることにより、データをキャッシュに保持するか否かを適切に判断する。

10

## 【 0 0 0 9 】

サーバがストレージからデータを読み出す場合、サーバとストレージのそれぞれがキャッシュを備えることがある。サーバ側のキャッシュを「サーバキャッシュ」、ストレージ側のキャッシュを「ストレージキャッシュ」と言う。

## 【 0 0 1 0 】

サーバキャッシュとストレージキャッシュとの両方を備える構成では、サーバにより読み出されたデータは、サーバキャッシュとストレージキャッシュの両方に一時的に保持される。そのため、1つのデータが、サーバキャッシュとストレージキャッシュとの両方に存在することがある。ところが、ストレージキャッシュ上のデータはサーバキャッシュ上のデータが消去されるまで参照されないため、ストレージのキャッシュが無駄に占有されるような状況が発生する可能性がある。

20

## 【 0 0 1 1 】

データが特定の 1 台のサーバのみにより読み出されるデータ（以下、「独立データ」という。）である場合には、特定のサーバのサーバキャッシュ上にあるデータが消去されるまで、ストレージキャッシュ上の対応するデータは読み出されない。一方、データが複数のサーバにより読み出されるデータ（以下、「共有データ」という。）である場合には、複数のサーバのうちの少なくとも一台のサーバのサーバキャッシュ上にあるデータが消去されると、ストレージキャッシュ上の対応するデータが読み出される。そこで、ストレージキャッシュでは、読み出される可能性が低い独立データを保持するよりも、読み出される可能性が高い共有データを保持する方が、限られたキャッシュ領域においてヒット率を向上させることができる。

30

## 【 0 0 1 2 】

独立データと共有データとを区別して扱うことにより、キャッシュを高性能化するための技術の一例が、特許文献 3 に開示されている。

## 【 0 0 1 3 】

特許文献 3 のディスクアレイ装置では、共有ディスクのキャッシュはディスクインターフェースユニット内の共有キャッシュに、占有ディスクのキャッシュはホストインターフェースユニット内の占有キャッシュに置かれる。

## 【 0 0 1 4 】

上記の動作の結果、特許文献 3 のディスクアレイ装置では、ホストインターフェースユニット及びディスクインターフェースユニットが増設された場合に、ディスクアレイ装置の性能が低下しない。

40

## 【 0 0 1 5 】

独立データと共有データとを区別して扱うことにより、キャッシュを高性能化するための技術の別の一例が、特許文献 4 に開示されている。

## 【 0 0 1 6 】

特許文献 4 のマルチコアプロセッサシステムでは、共有キャッシュコントローラは、プログラムコードに基づいて、変数がスレッド間で共有される共有データであるか、あるいは共有されない非共有データであるかを判別する。共有キャッシュコントローラは、共有

50

メモリ中の共有データを共有キャッシュメモリに、共有メモリ中の非共有データを分散キャッシュに保持する。

【0017】

上記の動作の結果、特許文献4のマルチコアプロセッサシステムは、共有キャッシュメモリの領域を効率的に利用する。

【先行技術文献】

【特許文献】

【0018】

【特許文献1】特開2010-049502

【特許文献2】特開2012-118745

【特許文献3】特開2007-206949

【特許文献4】WO2011/161829

【発明の概要】

【発明が解決しようとする課題】

【0019】

特許文献3及び特許文献4の技術では、事前に設定、分析等が行われることにより、キャッシュ動作前に独立データと共有データとの判別が行われる。そのため、特許文献3及び特許文献4の技術には、キャッシュ動作時に、動的に独立データと共有データとの判別を行うことができないという問題がある。

(発明の目的)

本発明の主たる目的は、キャッシュの動作時に、動的に独立データと共有データとの判別を行い、共有データを優先してキャッシュに保持することができるキャッシュ装置、キャッシュシステム、キャッシュ方法、及びキャッシュプログラムを提供することにある。

【課題を解決するための手段】

【0020】

本発明のキャッシュ装置は、記憶装置に接続され、複数の要求元に接続されるキャッシュ装置であって、第1の要求元から第1の読み出し要求を受信し、第1の読み出し要求において指定された第1のデータを保持していない場合に、記憶装置に第1の読み出し要求を送信し、記憶装置から第1の読み出し応答を受信し、第1の読み出し応答を第1の要求元へ送信する一方、第1のデータを保持している場合に、第1のデータを含む第2の読み出し応答を第1の要求元へ送信するキャッシュ手段と、第1のデータの指定に対応付けて第1の要求元を第1の履歴として記憶し、キャッシュ手段により第1のデータが保持されておらず且つキャッシュ手段に第1のデータを保持するために十分な記憶領域が存在する場合に、キャッシュ手段に第1のデータを記憶させる一方、キャッシュ手段により第1のデータが保持されておらず且つキャッシュ手段に第1のデータを保持するために十分な記憶領域が存在しない場合に、第1のデータ又はキャッシュ手段に保持されたデータのうち、第1の履歴においてより少数の要求元に対応付けられたデータを第1の履歴においてより多数の要求元に対応付けられたデータに優先して第2のデータとして選択し、第2のデータと第1のデータとが異なるならば、キャッシュ手段に第2のデータを破棄させた後に第1のデータを記憶させる記憶領域制御手段とを備えることを特徴とする。

【0021】

本発明のキャッシュシステムは、記憶装置に接続されたキャッシュ装置と、キャッシュ装置を介して記憶装置に接続され、それぞれ1台の要求元に接続される、複数の2次キャッシュ装置とを備えたキャッシュシステムであって、キャッシュ装置は、2次キャッシュ装置を介して第1の要求元から第1の読み出し要求を受信し、第1の読み出し要求において指定された第1のデータを保持していない場合に、記憶装置に第1の読み出し要求を送信し、記憶装置から第1の読み出し応答を受信し、第1の読み出し応答を2次キャッシュ装置を介して第1の要求元へ送信する一方、第1のデータを保持している場合に、第1のデータを含む第2の読み出し応答を2次キャッシュ装置を介して第1の要求元へ送信するキャッシュ手段と、第1のデータの指定に対応付けて第1の要求元を第1の履歴として記憶

10

20

30

40

50

し、キャッシュ手段により第1のデータが保持されておらず且つキャッシュ手段に第1のデータを保持するために十分な記憶領域が存在する場合に、キャッシュ手段に第1のデータを記憶させる一方、キャッシュ手段により第1のデータが保持されておらず且つキャッシュ手段に第1のデータを保持するために十分な記憶領域が存在しない場合に、第1のデータ又はキャッシュ手段に保持されたデータのうち、第1の履歴においてより少数の要求元に対応付けられたデータを第1の履歴においてより多数の要求元に対応付けられたデータに優先して第2のデータとして選択し、第2のデータと第1のデータとが異なるならば、キャッシュ手段に第2のデータを破棄させた後に第1のデータを記憶させる記憶領域制御手段とを備え、2次キャッシュ装置は、第1の要求元から第2の読み出し要求を受信し、第2の読み出し要求において指定された第3のデータを保持していない場合に、キャッシュ装置を介して記憶装置に第2の読み出し要求を送信し、キャッシュ装置を介して記憶装置から第3の読み出し応答を受信し、第3の読み出し応答を第1の要求元へ送信する一方、第3のデータを保持している場合に、第3のデータを含む第4の読み出し応答を第1の要求元へ送信する2次キャッシュ手段と、第3のデータが保持されておらず且つ第3のデータを保持するために十分な記憶領域が存在する場合に、第3のデータを記憶する一方、第3のデータが保持されておらず且つ第3のデータを保持するために十分な記憶領域が存在しない場合に、キャッシュ装置から第1の履歴の情報を取得し、第3のデータ又は保持されたデータのうち、第1の履歴においてより多数の要求元に対応付けられたデータを第1の履歴においてより少数の要求元に対応付けられたデータに優先して第4のデータとして選択し、第4のデータと第3のデータとが異なるならば、第4のデータを破棄した後に第3のデータを記憶する2次記憶領域制御手段とを備えることを特徴とする。

10

20

## 【0022】

本発明のキャッシュ方法は、記憶装置に接続され、複数の要求元に接続されるキャッシュ装置において、第1の要求元から第1の読み出し要求を受信し、第1の読み出し要求において指定された第1のデータを保持していない場合に、記憶装置に第1の読み出し要求を送信し、記憶装置から第1の読み出し応答を受信し、第1の読み出し応答を第1の要求元へ送信する一方、第1のデータを保持している場合に、第1のデータを含む第2の読み出し応答を第1の要求元へ送信し、第1のデータの指定に対応付けて第1の要求元を第1の履歴として記憶し、第1のデータが保持されておらず且つ第1のデータを保持するために十分な記憶領域が存在する場合に、第1のデータを記憶する一方、第1のデータが保持されておらず且つ第1のデータを保持するために十分な記憶領域が存在しない場合に、第1のデータ又は保持されたデータのうち、第1の履歴においてより少数の要求元に対応付けられたデータを第1の履歴においてより多数の要求元に対応付けられたデータに優先して第2のデータとして選択し、第2のデータと第1のデータとが異なるならば、第2のデータを破棄した後に第1のデータを記憶することを特徴とする。

30

## 【0023】

本発明のキャッシュプログラムは、記憶装置に接続され、複数の要求元に接続されるキャッシュ装置であって、第1の要求元から第1の読み出し要求を受信し、第1の読み出し要求において指定された第1のデータを保持していない場合に、記憶装置に第1の読み出し要求を送信し、記憶装置から第1の読み出し応答を受信し、第1の読み出し応答を第1の要求元へ送信する一方、第1のデータを保持している場合に、第1のデータを含む第2の読み出し応答を第1の要求元へ送信するキャッシュ処理と、第1のデータの指定に対応付けて第1の要求元を第1の履歴として記憶し、第1のデータが保持されておらず且つ第1のデータを保持するために十分な記憶領域が存在する場合に、第1のデータを記憶する一方、第1のデータが保持されておらず且つ第1のデータを保持するために十分な記憶領域が存在しない場合に、第1のデータ又は保持されたデータのうち、第1の履歴においてより少数の要求元に対応付けられたデータを第1の履歴においてより多数の要求元に対応付けられたデータに優先して第2のデータとして選択し、第2のデータと第1のデータとが異なるならば、第2のデータを破棄させた後に第1のデータを記憶させる記憶領域制御処理とをコンピュータに実行させることを特徴とする。

40

50

**【発明の効果】****【0024】**

本発明によれば、キャッシュの動作時に、動的に独立データと共有データとの判別を行い、共有データを優先してキャッシュに保持することができるという効果がある。

**【図面の簡単な説明】****【0025】**

【図1】本発明の第1の実施形態におけるキャッシュ装置100の構成の一例を示すブロック図である。

【図2】本発明の第1の実施形態におけるキャッシュ装置100の動作を示すフローチャートである。

【図3】本発明の第1の実施形態におけるキャッシュ装置100の動作の具体例を説明するための図である。

【図4】本発明の第2の実施形態におけるキャッシュ装置110の構成の一例を示すブロック図である。

【図5】本発明の第2の実施形態における2次キャッシュ装置400の動作を示すフローチャートである。

【図6】本発明の第2の実施形態におけるキャッシュ装置110及び2次キャッシュ装置400、401の動作の具体例を説明するための図である。

【図7】本発明の第3の実施形態におけるキャッシュ装置120の構成の一例を示すブロック図である。

【図8】本発明の第3の実施形態におけるキャッシュ装置120及び2次キャッシュ装置410、411の動作を示すフローチャートである。

【図9】本発明の第3の実施形態におけるキャッシュ装置120及び2次キャッシュ装置410、411の動作の具体例を説明するための図である。

【図10】本発明の第4の実施形態におけるストレージ130の構成の一例を示すブロック図である。

【図11】本発明の第4の実施形態におけるストレージ130の動作を示すフローチャートである。

【図12】本発明の第4の実施形態におけるストレージ130における「分割領域」の作成方法の具体例を説明するための図である。

【図13】本発明の第4の実施形態におけるストレージ130におけるストレージキャッシュ138から破棄されるデータの選択方法の具体例を説明するための図である。

【図14】本発明の第4の実施形態及び第5の実施形態におけるキャッシュデータの配置の一例を示す図である。

【図15】本発明の第5の実施形態におけるストレージ140及びサーバ340、341、342の構成の一例を示すブロック図である。

【図16】本発明の第5の実施形態におけるストレージ140及びサーバ340の動作を示すフローチャートである。

**【発明を実施するための形態】****【0026】**

以下、本発明の実施形態について、図面を参照して詳細に説明する。なお、すべての図面において、同等の構成要素には同じ符号を付し、適宜説明を省略する。

**(第1の実施形態)**

まず、本実施形態における構成について説明する。

**【0027】**

図1は、本発明の第1の実施形態におけるキャッシュ装置100の構成の一例を示すブロック図である。

**【0028】**

本実施形態のキャッシュ装置100は、記憶装置200と、要求元300と、要求元301とに接続される。キャッシュ装置100は、キャッシュ手段101と、記憶領域制御

10

20

30

40

50

手段 102 とを有する。なお、キャッシュ装置 100 は、複数の記憶装置に接続されてもよいし、3 台以上の要求元に接続されてもよい。

【0029】

キャッシュ手段 101 は、まず、要求元 300 からデータの読み出し要求を受信する。

【0030】

次に、キャッシュ手段 101 は、読み出し要求において指定された指定データを保持していない場合に、記憶装置 200 に読み出し要求を送信し、記憶装置 200 から読み出し要求に対する応答（読み出し応答）を受信し、読み出し応答を要求元 300 へ送信する。一方、キャッシュ手段 101 は、読み出し要求において指定された指定データを保持している場合に、その保持している指定データを含む読み出し応答を要求元 300 へ送信する。

10

【0031】

なお、キャッシュ手段 101 は、保持可能なキャッシュデータの量を定める所定の上限値（キャッシュ容量）を有する。キャッシュ容量は、保持可能なキャッシュデータのバイト数の合計により規定される。あるいは、キャッシュ容量は、指定データを格納するために使われるメモリ領域のブロック数の上限、保持可能な指定データの個数の上限等により規定されてもよい。

【0032】

キャッシュ手段 101 は、要求元 301 に対しても、要求元 300 に対するのと同様に動作する。

20

【0033】

記憶領域制御手段 102 は、まず、読み出し要求における指定データの「指定」に対応付けて、要求元 300 を履歴に記憶する。以下、キャッシュ装置 100 の記憶領域制御手段 102 が有する前述の履歴を「要求履歴」と呼ぶ。なお、指定データの「指定」とは、指定データを特定する任意の情報である。指定データの指定は、例えば、データの識別子、ファイル名、開始アドレスと終了アドレスとの組、開始アドレスとデータ長との組、記憶装置の識別子、又はこれらの組み合わせである。また、要求履歴において、要求元 300 は、要求元の識別子、ネットワークアドレス等の要求元を特定する任意の情報により特定される。

【0034】

なお、記憶領域制御手段 102 は、要求履歴の要素として、所定の個数の最新の要素のみを保持してもよい。また、記憶領域制御手段 102 は、特定の指定データの指定に対する、要求元及び要求回数を、要求履歴として記憶してもよい。

30

【0035】

次に、記憶領域制御手段 102 は、キャッシュ手段 101 により指定データが保持されておらず、且つキャッシュ手段 101 に上述の指定データを保持するために十分な記憶領域が存在する場合に、キャッシュ手段 101 に指定データを記憶させる。一方、記憶領域制御手段 102 は、キャッシュ手段 101 により指定データが保持されておらず、且つキャッシュ手段 101 に指定データを保持するために十分な記憶領域が存在しない場合に、破棄候補のデータを決定する。具体的には、記憶領域制御手段 102 は、指定データ又はキャッシュ手段 101 に保持されたデータのうち、要求履歴においてより少数の要求元に対応付けられたデータを、より多数の要求元に対応付けられたデータに優先して破棄候補のデータとして選択する。続いて、記憶領域制御手段 102 は、選択した破棄候補のデータと指定データとが異なるならば、キャッシュ手段 101 に、破棄候補のデータを破棄させた後に指定データを記憶させる。

40

【0036】

なお、記憶領域制御手段 102 は、要求履歴の要素のうち、所定の個数の最新の要素のみを参照してもよい。また、最も少数の要求元に対応付けられた、キャッシュデータ又は指定データが複数存在する場合には、記憶領域制御手段 102 は、その中から任意のキャッシュデータ又は指定データを破棄候補として選択してもよい。キャッシュ装置 100 は

50

、例えば、その中から最も古い読み出し要求に対応するキャッシュデータを破棄候補として選択してもよい。

【0037】

記憶領域制御手段102は、要求元301に対しても、要求元300に対するのと同様に動作する。

【0038】

記憶装置200は、キャッシュデータの元データを保持する。記憶装置200は、例えば、メモリ、SSD (Solid State Drive)、ハードディスク、ディスクアレイ、データベース、サーバである。

【0039】

要求元300、要求元301は、記憶装置200が保持するデータを読み出す。要求元300、要求元301は、例えば、サーバ、端末、CPU (Central Processing Unit)、マルチコアCPUにおける各コア、マルチスレッドに対応したCPUにおける各スレッドである。

【0040】

次に、本実施形態における動作について説明する。

【0041】

図2は、本発明の第1の実施形態におけるキャッシュ装置100の動作を示すフローチャートである。図2のフローチャートは、1つの要求元からの1つのデータ読み出しに対する処理を示す。以下では、要求元300に対する処理を例に説明する。なお、図2に示すフローチャート及び以下の説明は一例であり、適宜求める処理に応じて、処理順等を入れ替えたり、処理を戻したり、又は処理を繰り返したりしてもよい。また、キャッシュ装置100が3つ以上の要求元に接続される場合にも、キャッシュ装置100は、以下に説明する動作と同様に動作する。

【0042】

キャッシュ手段101は、要求元300から特定のデータの読み出し要求を受信する (ステップS101)。

【0043】

記憶領域制御手段102は、読み出し要求における指定データの指定に対応付けて、要求元300を要求履歴として記憶する (ステップS102)。

【0044】

記憶領域制御手段102は、キャッシュ手段101により指定データが保持されているか否かを判定する (ステップS103)。

【0045】

キャッシュ手段101により指定データが保持されている場合に (ステップS103: Yes)、キャッシュ手段101は、その保持されている指定データを含む読み出し応答を要求元300へ送信し (ステップS104)、処理を終了する。

【0046】

キャッシュ手段101により指定データが保持されていない場合に (ステップS103: No)、キャッシュ手段101は、記憶装置200に読み出し要求を送信し (ステップS105)、ステップS106の処理へ進む。キャッシュ手段101は、記憶装置200から読み出し応答を受信し (ステップS106)、読み出し応答を要求元300へ送信し (ステップS107)、ステップS108の処理へ進む。

【0047】

記憶領域制御手段102は、キャッシュ手段101に上述の指定データを保持するために十分な記憶領域が存在するか否かを判定する (ステップS108)。

【0048】

キャッシュ手段101に上述の指定データを保持するために十分な記憶領域が存在する場合に (ステップS108: Yes)、記憶領域制御手段102は、キャッシュ手段101に指定データを記憶させ (ステップS109)、処理を終了する。

10

20

30

40

50

## 【 0 0 4 9 】

キャッシュ手段 1 0 1 に上述の指定データを保持するために十分な記憶領域が存在しない場合に (ステップ S 1 0 8 : N o )、記憶領域制御手段 1 0 2 は、指定データ又はキャッシュ手段 1 0 1 に保持されたデータのうち、要求履歴においてより少数の要求元に対応付けられたデータを、より多数の要求元に対応付けられたデータに優先して破棄候補のデータとして選択し (ステップ S 1 1 0)、ステップ S 1 1 1 の処理へ進む。記憶領域制御手段 1 0 2 は、選択した破棄候補のデータと指定データとが異なるならば、キャッシュ手段 1 0 1 に、破棄候補のデータを破棄させた後に指定データを記憶させ (ステップ S 1 1 1)、処理を終了する。なお、上記の要求元の数、要求履歴において特定のデータを要求した要求元の数である。例えば、特定のデータを、要求元 3 0 0 が 3 回、要求元 3 0 1 が 4 回要求した場合には、上記の要求元数は 2 である。

10

## 【 0 0 5 0 】

次に、本実施形態における処理の具体例について説明する。

## 【 0 0 5 1 】

図 3 は、本発明の第 1 の実施形態におけるキャッシュ装置 1 0 0 の動作の具体例を説明するための図である。

## 【 0 0 5 2 】

図 3 ( a ) は、要求元 3 0 0 及び要求元 3 0 1 からのデータ読み取り要求の時系列を示す。1 番目の要素は、時刻 “ 1 ” に、要求元 “ 要求元 A ” ( 要求元 3 0 0 ) から指定データ “ データ A ” の読み出し要求があることを示す。2 番目の要素は、時刻 “ 2 ” に、要求元 “ 要求元 B ” ( 要求元 3 0 1 ) から指定データ “ データ B ” の読み出し要求があることを示す。3 番目の要素は、時刻 “ 3 ” に、要求元 “ 要求元 A ” から指定データ “ データ C ” の読み出し要求があることを示す。4 番目の要素は、時刻 “ 4 ” に、要求元 “ 要求元 B ” から指定データ “ データ C ” の読み出し要求があることを示す。5 番目から 8 番目の要素は、時刻 “ 5 ” から時刻 “ 8 ” に、それぞれ、時刻 “ 1 ” から時刻 “ 4 ” における読み出し要求が繰り返されることを示す。

20

## 【 0 0 5 3 】

つまり、“ データ A ” は要求元 “ 要求元 A ” からのみ要求される独立データである。また、“ データ B ” は要求元 “ 要求元 B ” からのみ要求される独立データである。一方、“ データ C ” は要求元 “ 要求元 A ” 及び要求元 “ 要求元 B ” の両方から要求される共有データである。

30

## 【 0 0 5 4 】

図 3 ( b ) は、キャッシュ装置 1 0 0 のキャッシュデータに対する操作を示す。第 1 列は、時刻を示す。第 2 列、第 3 列、第 4 列は、それぞれ、指定データ “ データ A ”、指定データ “ データ B ”、指定データ “ データ C ” に対するキャッシュ装置 1 0 0 の操作を示す。操作 “ 更新 ” は、記憶装置 2 0 0 に対するデータ読み出し要求及びデータ読み出し応答を中継し、データ読み出し応答中のデータでキャッシュデータを更新することを示す。操作 “ ヒット ” は、保持するキャッシュデータを用いてデータ読み出し応答を送信することを示す。操作 “ 維持 ” は、キャッシュデータを変更しないことを示す。操作 “ 中継 ” は、記憶装置 2 0 0 に対するデータ読み出し要求及びデータ読み出し応答を中継するが、キャッシュデータを更新しないことを示す。なお、図 3 ( b ) では、キャッシュ手段 1 0 1 のキャッシュ容量は、指定データ 1 つ分である。また、時刻 “ 1 ” の前には、キャッシュは空であるものとする。

40

時刻 “ 1 ” )

キャッシュ装置 1 0 0 は、指定データ “ データ A ” に対する要求元 “ 要求元 A ” からの読み出し要求を受信すると (ステップ S 1 0 1)、要求履歴に記録する (ステップ S 1 0 2)。指定データ “ データ A ” のキャッシュデータは存在しないので (ステップ S 1 0 3 : N o)、キャッシュ装置 1 0 0 は、記憶装置 2 0 0 から指定データ “ データ A ” を取得し (ステップ S 1 0 5、S 1 0 6)、要求元 “ 要求元 A ” へ送信する (ステップ S 1 0 7)。キャッシュに必要な空き容量があるので (ステップ S 1 0 8 : Y e s)、キャッシュ

50

装置 100 は、キャッシュに指定データ “データ A” を記憶する (ステップ S 109)。  
時刻 “2”)

キャッシュ装置 100 は、指定データ “データ B” に対する要求元 “要求元 B” からの読み出し要求を受信すると (ステップ S 101)、要求履歴に記録する (ステップ S 102)。指定データ “データ B” のキャッシュデータは存在しないので (ステップ S 103: No)、キャッシュ装置 100 は、記憶装置 200 から指定データ “データ B” を取得し (ステップ S 105、S 106)、要求元 “要求元 B” へ送信する (ステップ S 107)。キャッシュに必要な空き容量がないので (ステップ S 108: No)、キャッシュ装置 100 は、1つの要求元 “要求元 A” に対応付けられたデータ “データ A” を破棄候補として選択し (ステップ S 110)、選択した破棄候補 “データ A” を指定データ “データ B” に更新する (ステップ S 111)。なお、本動作例では、最も少数台の要求元に対応付けられた、キャッシュデータ又は指定データが複数存在する場合には、キャッシュ装置 100 は、最も古い読み出し要求に対応するキャッシュデータを破棄候補として選択する。

10

時刻 “3”)

キャッシュ装置 100 は、指定データ “データ C” に対する要求元 “要求元 A” からの読み出し要求を受信すると (ステップ S 101)、要求履歴に記録する (ステップ S 102)。指定データ “データ C” のキャッシュデータは存在しないので (ステップ S 103: No)、キャッシュ装置 100 は、記憶装置 200 から指定データ “データ C” を取得し (ステップ S 105、S 106)、要求元 “要求元 A” へ送信する (ステップ S 107)。

20

時刻 “4”)

キャッシュ装置 100 は、指定データ “データ C” に対する要求元 “要求元 B” からの読み出し要求を受信すると (ステップ S 101)、要求履歴に記録する (ステップ S 102)。指定データ “データ C” のキャッシュデータは存在するので (ステップ S 103: Yes)、キャッシュ装置 100 は、キャッシュデータ “データ C” を含む読み出し応答を要求元 “要求元 B” へ送信する (ステップ S 104)。

30

時刻 “5”)

キャッシュ装置 100 は、指定データ “データ A” に対する要求元 “要求元 A” からの読み出し要求を受信すると (ステップ S 101)、要求履歴に記録する (ステップ S 102)。指定データ “データ A” のキャッシュデータは存在しないので (ステップ S 103: No)、キャッシュ装置 100 は、記憶装置 200 から指定データ “データ A” を取得し (ステップ S 105、S 106)、要求元 “要求元 A” へ送信する (ステップ S 107)。キャッシュに必要な空き容量がないので (ステップ S 108: No)、キャッシュ装置 100 は、1つの要求元 “要求元 A” に対応付けられたデータ “データ A” を破棄候補として選択し (ステップ S 110)、選択した破棄候補 “データ A” と指定データ “データ A” が同じなので処理を終了する (ステップ S 111)。

40

時刻 “6”)

キャッシュ装置 100 は、指定データ “データ B” に対する要求元 “要求元 B” からの読み出し要求を受信すると (ステップ S 101)、要求履歴に記録する (ステップ S 102)。指定データ “データ B” のキャッシュデータは存在しないので (ステップ S 103: No)、キャッシュ装置 100 は、記憶装置 200 から指定データ “データ B” を取得し (ステップ S 105、S 106)、要求元 “要求元 B” へ送信する (ステップ S 107)。キャッシュに必要な空き容量がないので (ステップ S 108: No)、キャッシュ装置 100 は、1台の要求元 “要求元 B” に対応付けられたデータ “データ B” を破棄候補として選択し (ステップ S 110)、選択した破棄候補 “データ B” と指定データ “データ B” が同じなので処理を終了する (ステップ S 111)。

50

時刻“ 7 ”)

キャッシュ装置 100 は、指定データ“ データ C ”に対する要求元“ 要求元 A ”からの読み出し要求を受信すると(ステップ S 101)、要求履歴に記録する(ステップ S 102)。指定データ“ データ C ”のキャッシュデータは存在するので(ステップ S 103: Yes)、キャッシュ装置 100 は、キャッシュデータ“ データ C ”を含む読み出し応答を要求元“ 要求元 A ”へ送信する(ステップ S 104)。

時刻“ 8 ”)

キャッシュ装置 100 は、指定データ“ データ C ”に対する要求元“ 要求元 B ”からの読み出し要求を受信すると(ステップ S 101)、要求履歴に記録する(ステップ S 102)。指定データ“ データ C ”のキャッシュデータは存在するので(ステップ S 103: Yes)、キャッシュ装置 100 は、キャッシュデータ“ データ C ”を含む読み出し応答を要求元“ 要求元 B ”へ送信する(ステップ S 104)。

【 0055 】

以上説明したように、本実施形態のキャッシュ装置 100 は、共有データを独立データに優先してキャッシュに保持する。特に、キャッシュ装置 100 は、共有データの全体を保持するために十分なキャッシュ容量を有する場合には、全ての共有データの読み出しをキャッシュに保持する。また、キャッシュ装置 100 は、要求履歴に基づき、指定データが共有データであるか否かをキャッシュ動作時に動的に判断する。従って、本実施形態のキャッシュ装置 100 は、キャッシュの動作時に、動的に独立データと共有データとの判別を行い、共有データを優先してキャッシュに保持することができる。

【 0056 】

なお、本実施形態のキャッシュ装置 100 では、共有データが独立データに優先してキャッシュに保持されるので、共有データの読み出し要求の独立データの読み出し要求に対する比率が大きい場合に、特にキャッシュのヒット率が高い。

(第 2 の実施形態)

次に、上述した第 1 の実施形態において、キャッシュ装置が 2 次キャッシュ装置を介して要求元に接続される、本発明の第 2 の実施形態について説明する。なお、以下の説明において、第 1 の実施形態と同等の構成要素には同じ符号を付し、適宜説明を省略する。

【 0057 】

まず、本実施形態における構成について説明する。

【 0058 】

図 4 は、本発明の第 2 の実施形態におけるキャッシュ装置 110 の構成の一例を示すブロック図である。

【 0059 】

本実施形態のキャッシュ装置 110 は、2 次キャッシュ装置 400 を介して、要求元 300 に接続される。また、キャッシュ装置 110 は、2 次キャッシュ装置 401 を介して、要求元 301 に接続される。本実施形態のキャッシュ装置 110 の構成は、第 1 の実施形態のキャッシュ装置 100 の構成と同じである。

【 0060 】

2 次キャッシュ装置 400 及び 2 次キャッシュ装置 401 は、一般的なキャッシュ装置である。2 次キャッシュ装置 400、2 次キャッシュ装置 401 は、それぞれ、要求元 300、要求元 301 からのデータ読み出し要求を、キャッシュ装置 110 を介して、記憶装置 200 へ中継する。2 次キャッシュ装置 400、2 次キャッシュ装置 401 は、データ読み出し要求を中継する際に、中継する読み出しデータを内部のキャッシュに保持する。2 次キャッシュ装置 400、2 次キャッシュ装置 401 は、キャッシュデータがキャッシュ容量を超える場合には、読み出し頻度の低いキャッシュデータを優先して破棄する。

【 0061 】

次に、本実施形態における動作について説明する。

【 0062 】

図 5 は、本発明の第 2 の実施形態における 2 次キャッシュ装置 400、401 の動作を

示すフローチャートである。図5のフローチャートは、1つの要求元からの1つのデータ読み出しに対する処理を示す。以下では、要求元300に対する2次キャッシュ装置400処理を例に説明する。なお、図5に示すフローチャート及び以下の説明は一例であり、適宜求める処理に応じて、処理順等を入れ替えたり、処理を戻したり、又は処理を繰り返したりしてもよい。

【0063】

本実施形態のキャッシュ装置110の動作は、第1の実施形態のキャッシュ装置100の動作と同じである。

【0064】

本実施形態の2次キャッシュ装置400の動作は、ステップS110の処理がステップS210の処理に置き換えられることを除いて、第1の実施形態のキャッシュ装置100の動作と同じである。

【0065】

2次キャッシュ装置400は、指定データ又はキャッシュに保持されたデータのうち、要求履歴においてより少数回、要求元に対応付けられたデータを、より多数回、要求元に対応付けられたデータに優先して破棄候補のデータとして選択し(ステップS210)、ステップS111の処理へ進む。なお、上記の回数は、要求履歴において特定のデータが要求元から要求された累積回数である。例えば、特定のデータを、要求元300が2回要求した場合には、上記の回数は2である。

【0066】

なお、最も少数回、要求元に対応付けられた、キャッシュデータ又は指定データが複数存在する場合には、2次キャッシュ装置400は、その中から任意のキャッシュデータ又は指定データを破棄候補として選択してもよい。2次キャッシュ装置400は、例えば、その中から最も古い読み出し要求に対応するキャッシュデータを破棄候補として選択してもよい。

【0067】

次に、本実施形態における処理の具体例について説明する。

【0068】

図6は、本発明の第2の実施形態におけるキャッシュ装置110及び2次キャッシュ装置400、401の動作の具体例を説明するための図である。

【0069】

図6(a)は、要求元300及び要求元301からのデータ読み取り要求の時系列を示す。1番目の要素は、時刻“1”に、要求元“要求元A”から指定データ“データA”の読み出し要求があることを示す。2番目の要素は、時刻“2”に、要求元“要求元B”から指定データ“データB”の読み出し要求があることを示す。3番目から4番目の要素は、時刻“3”から時刻“4”に、それぞれ、時刻“1”から時刻“2”における読み出し要求が繰り返されることを示す。5番目の要素は、時刻“5”に、要求元“要求元A”から指定データ“データC”の読み出し要求があることを示す。6番目の要素は、時刻“6”に、要求元“要求元B”から指定データ“データC”の読み出し要求があることを示す。7番目から12番目の要素は、時刻“7”から時刻“12”に、それぞれ、時刻“1”から時刻“6”における読み出し要求が繰り返されることを示す。

【0070】

つまり、“データA”は要求元“要求元A”からのみ要求される独立データである。また、“データB”は要求元“要求元B”からのみ要求される独立データである。一方、“データC”は要求元“要求元A”及び要求元“要求元B”の両方から要求される共有データである。

【0071】

図6(b)は、2次キャッシュ装置400のキャッシュデータに対する操作を示す。前提条件及び操作の記載方法は、図3(b)の前提条件及び記載方法と同様である。

【0072】

10

20

30

40

50

図6(c)は、2次キャッシュ装置401のキャッシュデータに対する操作を示す。前提条件及び操作の記載方法は、図3(b)の前提条件及び記載方法と同様である。

【0073】

図6(d)は、キャッシュ装置110のキャッシュデータに対する操作を示す。前提条件及び操作の記載方法は、図3(b)の前提条件及び記載方法と同様である。

時刻“1”)

2次キャッシュ装置400は、指定データ“データA”に対する要求元“要求元A”からの読み出し要求を受信すると(ステップS101)、要求履歴に記録する(ステップS102)。指定データ“データA”のキャッシュデータは存在しないので(ステップS103: No)、2次キャッシュ装置400は、キャッシュ装置110を介して記憶装置200から指定データ“データA”を取得し(ステップS105、S106)、要求元“要求元A”へ送信する(ステップS107)。キャッシュに必要な空き容量があるので(ステップS108: Yes)、2次キャッシュ装置400は、キャッシュに指定データ“データA”を記憶する(ステップS109)。

10

【0074】

キャッシュ装置110は、指定データ“データA”に対する要求元“要求元A”からの読み出し要求を2次キャッシュ装置400を介して受信すると(ステップS101)、要求履歴に記録する(ステップS102)。指定データ“データA”のキャッシュデータは存在しないので(ステップS103: No)、キャッシュ装置110は、記憶装置200から指定データ“データA”を取得し(ステップS105、S106)、2次キャッシュ装置400を介して要求元“要求元A”へ送信する(ステップS107)。キャッシュに必要な空き容量があるので(ステップS108: Yes)、キャッシュ装置110は、キャッシュに指定データ“データA”を記憶する(ステップS109)。

20

時刻“2”)

2次キャッシュ装置401は、指定データ“データB”に対する要求元“要求元B”からの読み出し要求を受信すると(ステップS101)、要求履歴に記録する(ステップS102)。指定データ“データB”のキャッシュデータは存在しないので(ステップS103: No)、2次キャッシュ装置401は、キャッシュ装置110を介して記憶装置200から指定データ“データB”を取得し(ステップS105、S106)、要求元“要求元B”へ送信する(ステップS107)。キャッシュに必要な空き容量があるので(ステップS108: Yes)、2次キャッシュ装置401は、キャッシュに指定データ“データB”を記憶する(ステップS109)。

30

【0075】

キャッシュ装置110は、指定データ“データB”に対する要求元“要求元B”からの読み出し要求を2次キャッシュ装置401を介して受信すると(ステップS101)、要求履歴に記録する(ステップS102)。指定データ“データB”のキャッシュデータは存在しないので(ステップS103: No)、キャッシュ装置110は、記憶装置200から指定データ“データB”を取得し(ステップS105、S106)、2次キャッシュ装置401を介して要求元“要求元B”へ送信する(ステップS107)。キャッシュに必要な空き容量がないので(ステップS108: No)、キャッシュ装置110は、1台の要求元“要求元A”に対応付けられたデータ“データA”を破棄候補として選択し(ステップS110)、選択した破棄候補“データA”を指定データ“データB”に更新する(ステップS111)。

40

時刻“3”)

2次キャッシュ装置400は、指定データ“データA”に対する要求元“要求元A”からの読み出し要求を受信すると(ステップS101)、要求履歴に記録する(ステップS102)。指定データ“データA”のキャッシュデータは存在するので(ステップS103: Yes)、2次キャッシュ装置400は、キャッシュデータ“データA”を含む読み出し応答を要求元“要求元A”へ送信する(ステップS104)。

時刻“4”)

50

2次キャッシュ装置401は、指定データ“データB”に対する要求元“要求元B”からの読み出し要求を受信すると(ステップS101)、要求履歴に記録する(ステップS102)。指定データ“データB”のキャッシュデータは存在するので(ステップS103: Yes)、2次キャッシュ装置401は、キャッシュデータ“データB”を含む読み出し応答を要求元“要求元B”へ送信する(ステップS104)。

時刻“5”)

2次キャッシュ装置400は、指定データ“データC”に対する要求元“要求元A”からの読み出し要求を受信すると(ステップS101)、要求履歴に記録する(ステップS102)。指定データ“データC”のキャッシュデータは存在しないので(ステップS103: No)、2次キャッシュ装置400は、キャッシュ装置110を介して記憶装置200から指定データ“データC”を取得し(ステップS105、S106)、要求元“要求元A”へ送信する(ステップS107)。キャッシュに必要な空き容量がないので(ステップS108: No)、2次キャッシュ装置400は、1回、要求元“要求元A”に対応付けられたデータ“データC”を破棄候補として選択し(ステップS210)、選択した破棄候補“データC”を破棄する(ステップS111)。

【0076】

キャッシュ装置110は、指定データ“データC”に対する要求元“要求元A”からの読み出し要求を2次キャッシュ装置400を介して受信すると(ステップS101)、要求履歴に記録する(ステップS102)。指定データ“データC”のキャッシュデータは存在しないので(ステップS103: No)、キャッシュ装置110は、記憶装置200から指定データ“データC”を取得し(ステップS105、S106)、2次キャッシュ装置400を介して要求元“要求元A”へ送信する(ステップS107)。キャッシュに必要な空き容量がないので(ステップS108: No)、キャッシュ装置110は、1台の要求元“要求元B”に対応付けられたデータ“データB”を破棄候補として選択し(ステップS110)、選択した破棄候補“データB”を指定データ“データC”に更新する(ステップS111)。

時刻“6”)

2次キャッシュ装置401は、指定データ“データC”に対する要求元“要求元B”からの読み出し要求を受信すると(ステップS101)、要求履歴に記録する(ステップS102)。指定データ“データC”のキャッシュデータは存在しないので(ステップS103: No)、2次キャッシュ装置401は、キャッシュ装置110を介して記憶装置200から指定データ“データC”を取得し(ステップS105、S106)、要求元“要求元B”へ送信する(ステップS107)。キャッシュに必要な空き容量がないので(ステップS108: No)、2次キャッシュ装置401は、1回、要求元“要求元B”に対応付けられたデータ“データC”を破棄候補として選択し(ステップS210)、選択した破棄候補“データC”を破棄する(ステップS111)。

【0077】

キャッシュ装置110は、指定データ“データC”に対する要求元“要求元B”からの読み出し要求を2次キャッシュ装置401を介して受信すると(ステップS101)、要求履歴に記録する(ステップS102)。指定データ“データC”のキャッシュデータは存在するので(ステップS103: Yes)、キャッシュ装置110は、キャッシュデータ“データC”を含む読み出し応答を2次キャッシュ装置401を介して要求元“要求元B”へ送信する(ステップS104)。

時刻“7”)

2次キャッシュ装置400は、時刻“3”と同様に動作する。

時刻“8”)

2次キャッシュ装置401は、時刻“4”と同様に動作する。

時刻“9”)

2次キャッシュ装置400は、時刻“3”と同様に動作する。

時刻“10”)

10

20

30

40

50

2次キャッシュ装置401は、時刻“4”と同様に動作する。  
時刻“11”)

2次キャッシュ装置400は、時刻“5”と同様に動作する。

【0078】

キャッシュ装置110は、指定データ“データC”に対する要求元“要求元A”からの読み出し要求を2次キャッシュ装置400を介して受信すると(ステップS101)、要求履歴に記録する(ステップS102)。指定データ“データC”のキャッシュデータは存在するので(ステップS103: Yes)、キャッシュ装置110は、キャッシュデータ“データC”を含む読み出し応答を2次キャッシュ装置400を介して要求元“要求元A”へ送信する(ステップS104)。

10

時刻“12”)

2次キャッシュ装置401は、時刻“6”と同様に動作する。

【0079】

キャッシュ装置110は、時刻“6”と同様に動作する。

【0080】

以上説明したように、本実施形態のキャッシュ装置110は、共有データを独立データに優先してキャッシュに保持する。特に、キャッシュ装置110は、共有データの全体を保持するために十分なキャッシュ容量を有する場合には、全ての共有データの読み出しをキャッシュに保持する。また、キャッシュ装置110は、要求履歴に基づき、指定データが共有データであるか否かをキャッシュ動作時に動的に判断する。従って、本実施形態のキャッシュ装置110は、キャッシュの動作時に、動的に独立データと共有データとの判別を行い、共有データを優先してキャッシュに保持することができる。

20

【0081】

なお、2次キャッシュ装置400は、要求元300の要求頻度の高いデータを優先してキャッシュに保持する。特に、2次キャッシュ装置400は、要求頻度の高いデータの全体を保持するために十分なキャッシュ容量を有する場合には、要求頻度の高い全てのデータの読み出しをキャッシュに保持する。2次キャッシュ装置401も、2次キャッシュ装置400と同様である。つまり、2次キャッシュ装置400、401は、要求頻度の高いデータを優先してキャッシュに保持し、キャッシュ装置110は、共有データを優先してキャッシュに保持する。そのため、キャッシュ装置110では、個々の独立データの読み出し要求の個々の共有データの読み出し要求に対する比率が大きい場合に、特にキャッシュのヒット率が高い。

30

【0082】

また、2次キャッシュ装置400、2次キャッシュ装置401は、個々の独立データの読み出し要求の個々の共有データの読み出し要求に対する比率が大きい場合には、主に要求頻度の高いデータのキャッシュを保持する。一方、キャッシュ装置110は、個々の独立データの読み出し要求の個々の共有データの読み出し要求に対する比率が大きい場合には、主に共有データのキャッシュを保持する。その結果、本実施形態のキャッシュ装置110では、通常のキャッシュ装置を用いる場合に比べて、キャッシュ全体の利用効率が向上する。

40

(第3の実施形態)

次に、上述した第2の実施形態において、キャッシュ装置が2次キャッシュ装置に要求履歴の情報を渡す、本発明の第3の実施形態について説明する。なお、以下の説明において、第2の実施形態と同等の構成要素には同じ符号を付し、適宜説明を省略する。

【0083】

まず、本実施形態における構成について説明する。

【0084】

図7は、本発明の第3の実施形態におけるキャッシュ装置120の構成の一例を示すブロック図である。

【0085】

50

キャッシュ装置 120 は、第 2 の実施形態における 2 次キャッシュ装置 400 の代わりに、2 次キャッシュ装置 410 に接続される。また、キャッシュ装置 120 は、第 2 の実施形態における 2 次キャッシュ装置 401 の代わりに、2 次キャッシュ装置 411 に接続される。

【0086】

本実施形態のキャッシュ装置 120 は、通信手段 123 を有する。

【0087】

通信手段 123 は、2 次キャッシュ装置 410 又は 2 次キャッシュ装置 411 から要求履歴の情報を要求されると、2 次キャッシュ装置へ要求履歴の情報を送信する。

【0088】

2 次キャッシュ装置 410 は、第 2 の実施形態における 2 次キャッシュ装置 400 と動作が異なる。また、2 次キャッシュ装置 411 は、第 2 の実施形態における 2 次キャッシュ装置 410 と動作が異なる。本実施形態では、キャッシュ装置 120 は、第 2 の実施形態と同様に共有データを優先して保持する一方、2 次キャッシュ装置 410 及び 2 次キャッシュ装置 411 は、独立データを優先して保持する。それによってキャッシュ全体の利用効率を向上させる。

【0089】

本実施形態のキャッシュ装置 120 のその他の構成は、第 2 の実施形態のキャッシュ装置 110 の構成と同じである。

【0090】

次に、本実施形態における動作について説明する。

【0091】

図 8 は、本発明の第 3 の実施形態におけるキャッシュ装置 120 及び 2 次キャッシュ装置 410、411 の動作を示すフローチャートである。

【0092】

図 8 (a) は、キャッシュ装置 120 の動作を示すフローチャートである。図 8 (a) のフローチャートは、1 台の 2 次キャッシュ装置から 1 つの要求履歴の取得要求を受信した際の処理を示す。以下では、2 次キャッシュ装置 410 から取得要求を受信した際の処理を例に説明する。なお、図 8 (a) に示すフローチャート及び以下の説明は一例であり、適宜求める処理に応じて、処理を繰り返してもよい。

【0093】

図 8 (b) は、2 次キャッシュ装置 410 及び 2 次キャッシュ装置 411 の動作を示すフローチャートである。図 8 (b) のフローチャートは、1 つの要求元からの 1 つのデータ読み出しに対する処理を示す。以下では、要求元 300 に対する 2 次キャッシュ装置 410 の処理を例に説明する。なお、図 8 (b) に示すフローチャート及び以下の説明は一例であり、適宜求める処理に応じて、処理順等を入れ替えたり、処理を戻したり、又は処理を繰り返したりしてもよい。

【0094】

本実施形態のキャッシュ装置 120 の動作は、ステップ S401 の処理が追加されることを除いて、第 2 の実施形態のキャッシュ装置 110 の動作と同じである。

【0095】

キャッシュ装置 120 は、2 次キャッシュ装置 410 から、要求履歴の取得要求を受信すると、キャッシュ装置 120 が有する要求履歴の情報を送信する (ステップ S401)。

【0096】

本実施形態の 2 次キャッシュ装置 410 の動作は、ステップ S210 の処理がステップ S310 及びステップ S311 の処理に置き換えられることを除いて、第 2 の実施形態の 2 次キャッシュ装置 400 の動作と同じである。

【0097】

2 次キャッシュ装置 410 は、キャッシュ装置 120 からキャッシュ装置 120 の要求

10

20

30

40

50

履歴を取得し（ステップS 3 1 0）、ステップS 3 1 1の処理へ進む。

【0098】

2次キャッシュ装置410は、指定データ又はキャッシュに保持されたデータのうち、キャッシュ装置120の要求履歴においてより多数の要求元に対応付けられたデータを、より少数の要求元に対応付けられたデータに優先して破棄候補のデータとして選択し（ステップS 3 1 1）、ステップS 1 1 1の処理へ進む。なお、上記の要求元の数は、要求履歴において特定のデータを要求した要求元の数である。例えば、特定のデータを、要求元300が2回、要求元301が3回要求した場合には、上記の要求元の数は2である。

【0099】

なお、最も多数の要求元に対応付けられた、キャッシュデータ又は指定データが複数存在する場合には、2次キャッシュ装置410は、その中から任意のキャッシュデータ又は指定データを破棄候補として選択してもよい。2次キャッシュ装置410は、例えば、その中から最も新しい読み出し要求に対応するキャッシュデータを破棄候補として選択してもよい。

10

【0100】

次に、本実施形態における処理の具体例について説明する。

【0101】

図9は、本発明の第3の実施形態におけるキャッシュ装置120及び2次キャッシュ装置410、411の動作の具体例を説明するための図である。

【0102】

図9(a)は、要求元300及び要求元301からのデータ読み取り要求の時系列を示し、図3(a)と同様である。

20

【0103】

図9(b)は、2次キャッシュ装置410のキャッシュデータに対する操作を示す。前提条件及び操作の記載方法は、図6(b)の前提条件及び記載方法と同様である。

【0104】

図9(c)は、2次キャッシュ装置411のキャッシュデータに対する操作を示す。前提条件及び操作の記載方法は、図6(c)の前提条件及び記載方法と同様である。

【0105】

図9(d)は、キャッシュ装置120のキャッシュデータに対する操作を示す。前提条件及び操作の記載方法は、図6(d)の前提条件及び記載方法と同様である。  
時刻“1”)

30

2次キャッシュ装置410は、指定データ“データA”に対する要求元“要求元A”からの読み出し要求を受信すると（ステップS 1 0 1）、要求履歴に記録する（ステップS 1 0 2）。指定データ“データA”のキャッシュデータは存在しないので（ステップS 1 0 3：No）、2次キャッシュ装置410は、キャッシュ装置120を介して記憶装置200から指定データ“データA”を取得し（ステップS 1 0 5、S 1 0 6）、要求元“要求元A”へ送信する（ステップS 1 0 7）。キャッシュに必要な空き容量があるので（ステップS 1 0 8：Yes）、2次キャッシュ装置410は、キャッシュに指定データ“データA”を記憶する（ステップS 1 0 9）。

40

【0106】

キャッシュ装置120は、指定データ“データA”に対する要求元“要求元A”からの読み出し要求を2次キャッシュ装置410を介して受信すると（ステップS 1 0 1）、要求履歴に記録する（ステップS 1 0 2）。指定データ“データA”のキャッシュデータは存在しないので（ステップS 1 0 3：No）、キャッシュ装置120は、記憶装置200から指定データ“データA”を取得し（ステップS 1 0 5、S 1 0 6）、2次キャッシュ装置410を介して要求元“要求元A”へ送信する（ステップS 1 0 7）。キャッシュに必要な空き容量があるので（ステップS 1 0 8：Yes）、キャッシュ装置120は、キャッシュに指定データ“データA”を記憶する（ステップS 1 0 9）。

時刻“2”)

50

2次キャッシュ装置411は、指定データ“データB”に対する要求元“要求元B”からの読み出し要求を受信すると(ステップS101)、要求履歴に記録する(ステップS102)。指定データ“データB”のキャッシュデータは存在しないので(ステップS103:No)、2次キャッシュ装置411は、キャッシュ装置120を介して記憶装置200から指定データ“データB”を取得し(ステップS105、S106)、要求元“要求元B”へ送信する(ステップS107)。キャッシュに必要な空き容量があるので(ステップS108:Yes)、2次キャッシュ装置411は、キャッシュに指定データ“データB”を記憶する(ステップS109)。

【0107】

キャッシュ装置120は、指定データ“データB”に対する要求元“要求元B”からの読み出し要求を2次キャッシュ装置411を介して受信すると(ステップS101)、要求履歴に記録する(ステップS102)。指定データ“データB”のキャッシュデータは存在しないので(ステップS103:No)、キャッシュ装置120は、記憶装置200から指定データ“データB”を取得し(ステップS105、S106)、2次キャッシュ装置411を介して要求元“要求元B”へ送信する(ステップS107)。キャッシュに必要な空き容量がないので(ステップS108:No)、キャッシュ装置120は、1台の要求元“要求元A”に対応付けられたデータ“データA”を破棄候補として選択し(ステップS110)、選択した破棄候補“データA”を指定データ“データB”に更新する(ステップS111)。

時刻“3”)

2次キャッシュ装置410は、指定データ“データC”に対する要求元“要求元A”からの読み出し要求を受信すると(ステップS101)、要求履歴に記録する(ステップS102)。指定データ“データC”のキャッシュデータは存在しないので(ステップS103:No)、2次キャッシュ装置410は、キャッシュ装置120を介して記憶装置200から指定データ“データC”を取得し(ステップS105、S106)、要求元“要求元A”へ送信する(ステップS107)。

【0108】

キャッシュ装置120は、指定データ“データC”に対する要求元“要求元A”からの読み出し要求を2次キャッシュ装置410を介して受信すると(ステップS101)、要求履歴に記録する(ステップS102)。指定データ“データC”のキャッシュデータは存在しないので(ステップS103:No)、キャッシュ装置120は、記憶装置200から指定データ“データC”を取得し(ステップS105、S106)、2次キャッシュ装置410を介して要求元“要求元A”へ送信する(ステップS107)。キャッシュに必要な空き容量がないので(ステップS108:No)、キャッシュ装置120は、1台の要求元“要求元B”に対応付けられたデータ“データB”を破棄候補として選択し(ステップS110)、選択した破棄候補“データB”を指定データ“データC”に更新する(ステップS111)。

【0109】

2次キャッシュ装置410は、キャッシュに必要な空き容量がないので(ステップS108:No)、キャッシュ装置120の要求履歴を取得する(ステップS310)。

【0110】

キャッシュ装置120は、2次キャッシュ装置410から、要求履歴の取得要求を受信すると、キャッシュ装置120が有する要求履歴の情報を送信する(ステップS401)。

【0111】

2次キャッシュ装置410は、キャッシュ装置120の要求履歴に基づいて、1台の要求元“要求元A”に対応付けられたデータ“データC”を破棄候補として選択し(ステップS311)、選択した破棄候補“データC”を破棄する(ステップS111)。

時刻“4”)

2次キャッシュ装置411は、指定データ“データC”に対する要求元“要求元B”か

10

20

30

40

50

らの読み出し要求を受信すると(ステップS 1 0 1)、要求履歴に記録する(ステップS 1 0 2)。指定データ“データC”のキャッシュデータは存在しないので(ステップS 1 0 3: No)、2次キャッシュ装置4 1 1は、キャッシュ装置1 2 0を介して記憶装置2 0 0から指定データ“データC”を取得し(ステップS 1 0 5、S 1 0 6)、要求元“要求元B”へ送信する(ステップS 1 0 7)。

【0 1 1 2】

キャッシュ装置1 2 0は、指定データ“データC”に対する要求元“要求元B”からの読み出し要求を2次キャッシュ装置4 1 1を介して受信すると(ステップS 1 0 1)、要求履歴に記録する(ステップS 1 0 2)。指定データ“データC”のキャッシュデータは存在するので(ステップS 1 0 3: Yes)、キャッシュ装置1 2 0は、キャッシュデータ“データC”を含む読み出し応答を2次キャッシュ装置4 1 1を介して要求元“要求元B”へ送信する(ステップS 1 0 4)。

10

【0 1 1 3】

2次キャッシュ装置4 1 1は、キャッシュに必要な空き容量がないので(ステップS 1 0 8: No)、キャッシュ装置1 2 0の要求履歴を取得する(ステップS 3 1 0)。

【0 1 1 4】

キャッシュ装置1 2 0は、2次キャッシュ装置4 1 1から、要求履歴の取得要求を受信すると、キャッシュ装置1 2 0が有する要求履歴の情報を送信する(ステップS 4 0 1)。

【0 1 1 5】

2次キャッシュ装置4 1 1は、キャッシュ装置1 2 0の要求履歴に基づいて、2台の要求元“要求元A”及び“要求元B”に対応付けられたデータ“データC”を破棄候補として選択し(ステップS 3 1 1)、選択した破棄候補“データC”を破棄する(ステップS 1 1 1)。

20

時刻“5”)

2次キャッシュ装置4 1 0は、指定データ“データA”に対する要求元“要求元A”からの読み出し要求を受信すると(ステップS 1 0 1)、要求履歴に記録する(ステップS 1 0 2)。指定データ“データA”のキャッシュデータは存在するので(ステップS 1 0 3: Yes)、2次キャッシュ装置4 1 0は、キャッシュデータ“データA”を含む読み出し応答を要求元“要求元A”へ送信する(ステップS 1 0 4)。

30

時刻“6”)

2次キャッシュ装置4 1 1は、指定データ“データB”に対する要求元“要求元B”からの読み出し要求を受信すると(ステップS 1 0 1)、要求履歴に記録する(ステップS 1 0 2)。指定データ“データB”のキャッシュデータは存在するので(ステップS 1 0 3: Yes)、2次キャッシュ装置4 1 1は、キャッシュデータ“データB”を含む読み出し応答を要求元“要求元B”へ送信する(ステップS 1 0 4)。

時刻“7”)

2次キャッシュ装置4 1 0は、時刻“3”と同様に動作する。

【0 1 1 6】

キャッシュ装置1 2 0は、指定データ“データC”に対する要求元“要求元A”からの読み出し要求を2次キャッシュ装置4 1 0を介して受信すると(ステップS 1 0 1)、要求履歴に記録する(ステップS 1 0 2)。指定データ“データC”のキャッシュデータは存在するので(ステップS 1 0 3: Yes)、キャッシュ装置1 2 0は、キャッシュデータ“データC”を含む読み出し応答を2次キャッシュ装置4 1 0を介して要求元“要求元A”へ送信する(ステップS 1 0 4)。

40

【0 1 1 7】

キャッシュ装置1 2 0は、2次キャッシュ装置4 1 0から、要求履歴の取得要求を受信すると、キャッシュ装置1 2 0が有する要求履歴の情報を送信する(ステップS 4 0 1)。

時刻“8”)

50

2次キャッシュ装置411は、時刻“4”と同様に動作する。

【0118】

キャッシュ装置120は、時刻“4”と同様に動作する。

【0119】

以上説明したように、本実施形態のキャッシュ装置120は、共有データを独立データに優先してキャッシュに保持する。特に、キャッシュ装置120は、共有データの全体を保持するために十分なキャッシュ容量を有する場合には、全ての共有データの読み出しをキャッシュに保持する。また、キャッシュ装置120は、要求履歴に基づき、指定データが共有データであるか否かをキャッシュ動作時に動的に判断する。従って、本実施形態のキャッシュ装置120は、キャッシュの動作時に、動的に独立データと共有データとの判別を行い、共有データを優先してキャッシュに保持することができる。

10

【0120】

なお、2次キャッシュ装置410は、要求元300の独立データを優先してキャッシュに保持する。特に、2次キャッシュ装置410は、要求元300の独立データの全体を保持するために十分なキャッシュ容量を有する場合には、要求元300の全ての独立データの読み出しをキャッシュに保持する。2次キャッシュ装置411も、2次キャッシュ装置410と同様である。つまり、2次キャッシュ装置410、2次キャッシュ装置411は、それぞれ、要求元300、要求元301の独立データを優先してキャッシュに保持し、キャッシュ装置120は、共有データを優先してキャッシュに保持する。そのため、キャッシュ装置120では、キャッシュ装置120が共有データを保持するために十分なキャッシュ容量を有し、2次キャッシュ装置410及び2次キャッシュ装置411が独立データを保持するために十分なキャッシュ容量を有する場合には、独立データの読み出し要求の共有データの読み出し要求に対する比率によらずに、キャッシュのヒット率が高い。

20

【0121】

また、2次キャッシュ装置410及び2次キャッシュ装置411が独立データを保持するために十分なキャッシュ容量を有する場合には、2次キャッシュ装置410、2次キャッシュ装置411は、専ら独立データのキャッシュを保持する。一方、キャッシュ装置120が共有データを保持するために十分なキャッシュ容量を有する場合には、キャッシュ装置120は専ら共有データのキャッシュを保持する。その結果、本実施形態のキャッシュ装置120では、キャッシュ装置110を用いる場合に比べて、キャッシュ全体の利用効率が向上する。

30

(第4の実施形態)

次に、上述した第2の実施形態におけるキャッシュ装置を内蔵するストレージである、本発明の第4の実施形態について説明する。なお、以下の説明において、第2の実施形態と同等の構成要素には同じ符号を付し、適宜説明を省略する。

【0122】

まず、本実施形態における構成について説明する。

【0123】

図10は、本発明の第4の実施形態におけるストレージ130の構成の一例を示すブロック図である。

40

【0124】

ストレージ130は、サーバキャッシュ430と、サーバキャッシュ431と、サーバキャッシュ432とに接続される。ストレージ130は、サーバキャッシュ430、サーバキャッシュ431、サーバキャッシュ432を介して、それぞれ、サーバ330、サーバ331、サーバ332に接続される。なお、サーバキャッシュ430、サーバキャッシュ431、サーバキャッシュ432は、それぞれ、サーバ330、サーバ331、サーバ332に内蔵される。ストレージ130は、サーバ330、サーバ331、サーバ332により読み出されるデータを保持する、キャッシュを内蔵する記憶装置である。

【0125】

ストレージ130は、ストレージキャッシュ情報制御手段131と、分割領域判定手段

50

132と、アクセス情報制御手段133と、データ破棄判定手段134と、分割領域作成手段135と、ストレージキャッシュ138と、HDD(Hard Disk Drive)139とを有する。

【0126】

キャッシュ情報制御手段131は、各サーバからストレージ130に対してデータ読み出し要求が発行された際に、データ読み出しの指定データがストレージキャッシュ138に存在するか否かを判定し、ストレージキャッシュ138の使用量を計測する。

【0127】

分割領域判定手段132は、データ読み出しの指定データがストレージキャッシュ138上のどの「分割領域」に属するかの判定を行う。なお、「分割領域」については、後述する。

10

【0128】

アクセス情報制御手段133は、各サーバからのデータ読み出しの要求の要求履歴を記憶する。

【0129】

データ破棄判定手段134は、ストレージキャッシュ138に存在するデータのうち、破棄対象のデータの判定を行う。なお、データ破棄判定手段134は、ストレージキャッシュ138に存在するデータのうち、独立データを共有データに優先して、破棄対象のデータとして選択する。

【0130】

20

分割領域作成手段135は、ストレージキャッシュ138に新たにデータを保持させる際に、新たな「分割領域」を作成する。

【0131】

ストレージキャッシュ138は、HDD139から読み出したデータを一時的に保持する。

【0132】

HDD139は、サーバ330、サーバ331、サーバ332により読み出されるデータを常に保持する。

【0133】

次に、本実施形態における動作について説明する。

30

【0134】

図11は、本発明の第4の実施形態におけるストレージ130の動作を示すフローチャートである。なお、図11のフローチャートでは、説明を簡単にするために、ストレージキャッシュにおけるキャッシュデータの更新処理のみが示され、その他の処理は省略されている。また、図11に示すフローチャート及び以下の説明は一例であり、適宜求める処理に応じて、処理順等を入れ替えたり、処理を戻したり、又は処理を繰り返したりしてもよい。

【0135】

ストレージキャッシュ情報制御手段131は、各サーバからのデータ読み出し要求の指定データがストレージキャッシュ138に存在するか否かを判定する(ステップS501)。ストレージキャッシュ情報制御手段131は、存在する場合(ステップS501:Yes)にはステップS502の処理へ推移し、存在しない場合(ステップS501:No)にはステップS503の処理へ推移する。

40

【0136】

分割領域判定手段132は、データ読み出し要求の指定データがどの「分割領域」に対応するのかを判定し(ステップS502)、ステップS506の処理へ推移する。

【0137】

データ破棄判定手段134は、ストレージキャッシュ情報制御手段131からの情報に基づいて、ストレージキャッシュの使用量が所定の閾値を超えているか否かを判定する(ステップS503)。閾値を超えている場合(ステップS503:Yes)には、データ

50

破棄判定手段134は、ステップS504の処理へ推移する。閾値を超えていない場合（ステップS503：No）には、分割領域作成手段135は、ステップS505の処理へ推移する。

【0138】

データ破棄判定手段134は、独立データである「分割領域」を優先して特定の「分割領域」のデータを破棄し（ステップS504）、ステップS505の処理へ推移する。なお、データ破棄判定手段134は、「分割領域」に関連付けられた「サーバ優先度」にも基づいて破棄対象を決定してもよい。破棄対象の選択方法の具体例については、後述する。

【0139】

分割領域作成手段135は、データ読み出し要求の指定データが存在する範囲を覆うように、論理ディスクにおける開始アドレス及び終了アドレスにより指定された「分割領域」を決定し、その「分割領域」内に存在するデータをストレージキャッシュ138に保持させる（ステップS505）。新たな「分割領域」の作成方法の具体例については、後述する。

【0140】

アクセス情報制御手段133は、データ読み出し要求元のサーバから特定の「分割領域」の読み出しがあったことを要求履歴に登録し（ステップS506）、ステップS501の処理へ戻る。

【0141】

次に、本実施形態における処理の具体例について説明する。

【0142】

まず、「分割領域」について説明する。

【0143】

図12は、本発明の第4の実施形態におけるストレージ130における「分割領域」の作成方法の具体例を説明するための図である。

【0144】

「分割領域」とは、ストレージ130がHDD139上のデータをキャッシュデータとして管理するためのデータ指定方法である。分割領域は、HDD139が抽象化された「論理ディスク」における、特定のデータを含むブロックの、開始アドレス、終了アドレスの組である。以下では、分割領域におけるブロックの最小単位が“1MB”である場合の例を説明するが、最小単位は“1MB”に限定されない。

【0145】

最初に、ストレージ130は、あるサーバから論理ディスクのアドレス“2MB”からアドレス“2.7MB”の範囲のデータを読み出すコマンドを受信する。

【0146】

ストレージ130の分割領域作成手段135は、分割領域におけるブロックの最小単位は“1MB”なので、開始アドレスが“2MB”であり終了アドレスが“3MB”であるブロックを1つの分割領域として作成し、ストレージキャッシュ138にデータを保持する。

【0147】

次に、ストレージ130は、あるサーバから論理ディスクのアドレス“5.2MB”からアドレス“8.6MB”の範囲のデータを読み出すコマンドを受信する。

【0148】

分割領域作成手段135は、開始アドレスが“5MB”であり終了アドレスが“9MB”であるブロックを1つの分割領域として作成し、ストレージキャッシュ138にデータを保持する。

【0149】

続いて、ストレージ130は、他のサーバから同一論理ディスクのアドレス“8.7MB”からアドレス“8.9MB”の範囲のデータを読み出すコマンドを受信する。

10

20

30

40

50

## 【0150】

最新のデータ読み出しコマンドにおいて指定されたデータは、直前のデータ読み出しコマンドにおいて指定されたデータとは異なるデータである。しかしながら、最新のデータ読み出しコマンドにおいて指定されたデータは、直前のデータ読み出しコマンドに対応してストレージキャッシュ138上に保持されている分割領域に含まれる。そこで、ストレージ130は、最新のデータ読み出しコマンドにおいて指定されたデータを直前のデータ読み出しコマンドにおいて指定されたデータと同一のデータであるものとみなす。

## 【0151】

次に、ストレージキャッシュ138から破棄されるデータの選択方法の具体例について説明する。

## 【0152】

図13は、本発明の第4の実施形態におけるストレージ130におけるストレージキャッシュ138から破棄されるデータの選択方法の具体例を説明するための図である。

## 【0153】

最初に、ストレージ130は、サーバAからストレージキャッシュ138上に保持されていないデータAに対するデータ読み出し要求を受信する。

## 【0154】

ストレージ130は、ストレージキャッシュ138にデータAを含む分割領域Aを保持する。また、ストレージ130は、サーバAからの分割領域Aに対する読み出し要求を要求履歴に登録する。この時点では、ストレージ130は、分割領域AがサーバAの独立データであると判別する。

## 【0155】

次に、ストレージ130は、サーバBからストレージキャッシュ138上に存在する分割領域Aに含まれるデータに対するデータ読み出し要求を受信する。

## 【0156】

ストレージ130は、既に存在する分割領域Aについての要求履歴に対しサーバBからのデータ読み出し要求があった旨を追加する。その結果、ストレージ130は、分割領域AがサーバA及びサーバBの共有データであると判別する。

## 【0157】

ストレージ130は、同様に、サーバA、サーバB、サーバCから、データB、データC、データD、又はデータEに対するデータ読み出し要求を受信する。その結果、ストレージ130は、図13に示される、分割領域B、分割領域C、分割領域D、及び分割領域Eを生成し、各分割領域に関する要求履歴を保持する。

## 【0158】

図13に示されるように、複数のサーバからのデータ読み出し要求が要求履歴に登録された分割領域A、分割領域B、分割領域Eは共有データである。一方、単一のサーバからのデータ読み出し要求が要求履歴に登録された分割領域Cと分割領域Dは独立データである。

## 【0159】

ストレージ130は、サーバ毎の優先度にも基づいて、ストレージキャッシュ138からキャッシュデータを破棄してもよい。例えば、どのサーバのデータ読み出し要求を優先してストレージキャッシュ138に残すかに関するサーバ毎の優先度が、優先度の高いものから、サーバA、サーバB、サーバCの順で定義されるものとする。この場合、ストレージ130は、ストレージキャッシュ138の使用量が閾値を超えた場合に、独立データでありサーバの優先度の低い分割領域Dを最優先してストレージキャッシュ138から破棄する。

## 【0160】

また、ストレージ130は、ストレージキャッシュ138上に共有データしか存在しない場合に、共有データの中で、データの「共有度」が低く、更にサーバ毎の優先度が低いキャッシュデータを優先して、ストレージキャッシュ138から破棄してもよい。なお、

10

20

30

40

50

「共有度」とは、共有データを共有しているサーバの台数である。

【0161】

図13に示される状況では、キャッシュデータが破棄される優先度は、優先度の高いものから順に、分割領域D、分割領域C、分割領域E、分割領域A、分割領域Bである。

【0162】

なお、サーバ毎の優先度は、ユーザにより予め設定されるか、データ読み出しの優先度の統計を取ることにより取得可能である。例えば、「SCSI Architecture Model」に準拠するデータ読み出しコマンドが利用される場合には、ストレージ130は、データ読み出しコマンド中の「Command priority」の値を参照することにより、サーバ毎の優先度を取得することが可能である。

10

【0163】

以上説明したように、本実施形態のストレージ130は、共有データを独立データに優先してストレージキャッシュに保持する。特に、ストレージ130は、共有データの全体を保持するために十分なキャッシュ容量を有する場合には、全ての共有データの読み出しをストレージキャッシュに保持する。また、ストレージ130は、要求履歴に基づき、指定データが共有データであるか否かをキャッシュ動作時に動的に判断する。従って、本実施形態のストレージ130は、キャッシュの動作時に、動的に独立データと共有データとの判別を行い、共有データを優先してキャッシュに保持することができる。

【0164】

また、本実施形態のストレージ130は、各サーバの共有データをストレージキャッシュ138に保持する可能性が高い。そのため、本実施形態のストレージ130では、共有データの独立データに対する比率が高い場合には、キャッシュヒット率が高いという効果がある。

20

【0165】

また、本実施形態のストレージ130では、共有データが優先的にストレージキャッシュ138に保持されるので、共有データの独立データに対する比率が高い場合に、キャッシュ容量が効率的に利用される。そのため、本実施形態のストレージ130では、共有データの独立データに対する比率が高い場合には、読み出しキャッシュのメモリ容量を減らして、その分のメモリ容量を書き込みキャッシュに割り当て可能である。

【0166】

また、本実施形態のストレージ130では、データの共有度及びサーバ毎の優先度に応じて破棄するキャッシュデータが決定される場合には、優先サーバから要求されたデータがストレージキャッシュ138に保持される可能性が高い。そのため、本実施形態のストレージ130では、高負荷時にも、優先サーバからのデータ読み出し要求に対する応答性能の低下を抑制することができる。

30

【0167】

なお、上述の説明において、分割領域の最低単位は1MBであるが、最低単位は1MBに限定されない。また、同一のデータであるか否かの判定のために分割領域を使用したか、別の方法を用いて同一のデータであるか否かの判定が行われてもよい。例えば、ストレージキャッシュ138に保持されるデータは、データ読み出し要求における指定データの開始アドレスとデータ長の組を用いて同定されてもよい。

40

【0168】

また、上述の説明において、破棄対象データの判定のために、データの共有度及びサーバ毎の優先度が使用された。しかしながら、破棄対象データの判定のために、別の要素が使用されてもよい。例えば、破棄対象データの判定の際に、データの読み出しが要求された最近の時刻が古いキャッシュデータから優先的に破棄する、又は要求頻度が低いキャッシュデータから優先的に破棄する等の、新しい要素が追加されてもよい。さらに、各要素に基づく優先度の重み付き総和を各分割領域について算出し、重み付き総和が低いキャッシュデータが優先されて破棄されてもよい。

(第5の実施形態)

50

次に、上述した第4の実施形態において、サーバキャッシュとストレージキャッシュが連携してキャッシュデータを保持する、本発明の第5の実施形態について説明する。なお、以下の説明において、第4の実施形態と同等の構成要素には同じ符号を付し、適宜説明を省略する。

【0169】

まず、本実施形態におけるキャッシュデータの配置について説明する。

【0170】

図14は、本発明の第4の実施形態及び第5の実施形態におけるキャッシュデータの配置の一例を示す図である。本発明の第4の実施形態では、同一の共有データが各サーバキャッシュに保持される可能性が高い(図14(a))。一方、本実施形態では、ストレージキャッシュには共有データのみが保持され、サーバキャッシュには独立データのみが保持される可能性が高い(図14(b))。

10

【0171】

図14(b)の詳細について説明する。

【0172】

サーバAからのみ参照される独立データAは、サーバキャッシュAのみにより保持される可能性が高い。すなわち、ストレージキャッシュにより一旦は保持された独立データAは、サーバキャッシュにより保持された後に、優先的に破棄される。独立データB、独立データCについても同様である。

【0173】

複数のサーバA、サーバB、サーバCから参照される共有データDは、ストレージキャッシュのみにより保持される可能性が高い。すなわち、サーバキャッシュA、サーバキャッシュB、又はサーバキャッシュCにより一旦は保持された共有データDは、ストレージキャッシュにより保持された後に、優先的に破棄される。

20

【0174】

次に、本実施形態における構成について説明する。

【0175】

図15は、本発明の第5の実施形態におけるストレージ140及びサーバ340、341、342の構成の一例を示すブロック図である。

【0176】

本実施形態のストレージ140の構成は、第4の実施形態のストレージ130の構成と同じである。ただし、本実施形態のアクセス情報制御手段143は、第4の実施形態のアクセス情報制御手段133と動作が異なる。

30

【0177】

アクセス情報制御手段143は、ストレージ140の要求履歴の情報を取得するための「取得コマンド」を受信すると、ストレージ140の要求履歴の情報を送信する。

【0178】

また、アクセス情報制御手段143は、ストレージ140の特定の分割領域(キャッシュデータ)を破棄するための「破棄コマンド」を受信すると、指定された分割領域が独立データである場合には分割領域を破棄し、指定された分割領域が共有データである場合には要求履歴から対象サーバを削除する。

40

【0179】

アクセス情報制御手段143のその他の動作は、第4の実施形態のアクセス情報制御手段133の動作と同じである。

【0180】

サーバ340の構成について説明する。なお、サーバ341、サーバ342の構成は、サーバ340の構成と同じである。また、サーバ340、サーバ341、サーバ342は、それぞれ、サーバキャッシュ440、サーバキャッシュ441、サーバキャッシュ442を有する。また、サーバキャッシュ440、サーバキャッシュ441、サーバキャッシュ442は、それぞれ、サーバキャッシュ情報制御手段450、サーバキャッシュ情報制

50

御手段 4 5 1、サーバキャッシュ情報制御手段 4 5 2 を有する。

【 0 1 8 1 】

サーバ 3 4 0 のサーバキャッシュ 4 4 0 は、サーバキャッシュ情報制御手段 4 5 0 を有する。

【 0 1 8 2 】

また、サーバキャッシュ情報制御手段 4 5 0 は、ストレージ 1 4 0 に取得コマンドを送信して、ストレージ 1 4 0 の要求履歴の情報を受信する。

【 0 1 8 3 】

また、サーバキャッシュ情報制御手段 4 5 0 は、取得した要求履歴の情報に基づいて、ストレージキャッシュ 1 3 8 上の各分割領域（キャッシュデータ）が、サーバキャッシュ 4 4 0 により保持されているか否かを判定する。

10

【 0 1 8 4 】

また、サーバキャッシュ情報制御手段 4 5 0 は、各キャッシュデータをサーバキャッシュ 4 4 0 又はストレージキャッシュ 1 3 8 のどちらにより保持されるべきか否かを判定する。

【 0 1 8 5 】

また、サーバキャッシュ情報制御手段 4 5 0 は、ストレージ 1 4 0 にストレージキャッシュ 1 3 8 上の特定の分割領域（キャッシュデータ）を破棄させる破棄コマンドを送信する。

【 0 1 8 6 】

本実施形態のサーバ 3 4 0 のその他の構成は、第 4 の実施形態のサーバ 3 3 0 の構成と同じである。

20

【 0 1 8 7 】

なお、取得コマンドに対する応答は、ストレージ 1 4 0 の要求履歴の情報のうち少なくとも、サーバ 3 4 0 からの要求履歴が存在する分割領域に関する要求履歴の情報を含む。1 つの分割領域に関する要求履歴の情報は、例えば、分割領域の開始アドレス、分割領域のデータ長、分割領域データに対する各サーバからのデータ読み出し要求の有無の情報である。

【 0 1 8 8 】

また、破棄コマンドは、例えば、分割領域の開始アドレスとデータ長を含む。

30

【 0 1 8 9 】

もしくは、取得コマンドの応答において、各分割領域の識別子がサーバキャッシュ 4 4 0 に通知され、破棄コマンドにおいて、分割領域の識別子により各分割領域が識別されてもよい。

【 0 1 9 0 】

次に、本実施形態における動作について説明する。

【 0 1 9 1 】

図 1 6 は、本発明の第 5 の実施形態におけるストレージ 1 4 0 及びサーバ 3 4 0 の動作を示すフローチャートである。サーバ 3 4 1、サーバ 3 4 2 の動作は、サーバ 3 4 0 の動作と同じである。なお、図 1 6 のフローチャートでは、説明を簡単にするために、ストレージキャッシュ及びサーバキャッシュにおける 1 つのキャッシュデータの更新処理のみが示され、その他の処理は省略されている。また、図 1 6 に示すフローチャート及び以下の説明は一例であり、適宜求める処理に応じて、処理順等を入れ替えたり、処理を戻したり、又は処理を繰り返したりしてもよい。

40

【 0 1 9 2 】

サーバ 3 4 0 のサーバキャッシュ情報制御手段 4 5 0 は、ストレージ 1 4 0 へストレージ 1 4 0 の要求履歴の情報を取得するコマンドを送信し（ステップ S 6 0 1）、ステップ S 6 0 2 の処理へ推移する。

【 0 1 9 3 】

ストレージ 1 4 0 のアクセス情報制御手段 1 4 3 は、サーバ 3 4 0 から取得コマンドを

50

受信し、アクセス情報制御手段 1 4 3 にて保持される分割領域の要求履歴の情報のうち、サーバ 3 4 0 から要求された分割領域に関する要求履歴の情報を送信し（ステップ S 6 0 2）、ステップ S 6 0 3 の処理へ推移する。なお、分割領域に対する扱いは本発明の第 4 の実施形態における扱いと同じである。

【 0 1 9 4 】

サーバ 3 4 0 のサーバキャッシュ情報制御手段 4 5 0 は、ストレージキャッシュ 1 3 8 の要求履歴の情報に基づいて、ストレージキャッシュ 1 3 8 の各分割領域がサーバキャッシュ 4 4 0 に保持されているか否かを判定する（ステップ S 6 0 3）。サーバキャッシュ情報制御手段 4 5 0 は、存在しない場合（ステップ S 6 0 3 : Y e s）にはステップ S 6 0 4 の処理へ推移し、存在する場合（ステップ S 6 0 3 : N o）にはステップ S 6 0 6 の

10

【 0 1 9 5 】

サーバ 3 4 0 のサーバキャッシュ情報制御手段 4 5 0 は、各分割領域（キャッシュデータ）の共有度や要求頻度等に基づいて、特定の分割領域がストレージキャッシュ 1 3 8 により保持されるべきか否かを判定する（ステップ S 6 0 4）。サーバキャッシュ情報制御手段 4 5 0 は、分割領域がストレージキャッシュ 1 3 8 により保持されるべき場合（ステップ S 6 0 4 : Y e s）にはステップ S 6 0 1 の処理へ戻り、サーバキャッシュ 4 4 0 により保持されるべき場合（ステップ S 6 0 4 : N o）にはステップ S 6 0 5 の処理へ推移する。

【 0 1 9 6 】

サーバ 3 4 0 のサーバキャッシュ情報制御手段 4 5 0 は、特定の分割領域をサーバキャッシュ 4 4 0 に保持し（ステップ S 6 0 5）、ステップ S 6 0 8 の処理へ推移する。

20

【 0 1 9 7 】

サーバ 3 4 0 のサーバキャッシュ情報制御手段 4 5 0 は、各分割領域（キャッシュデータ）の共有度や要求頻度等に基づいて、特定の分割領域がサーバキャッシュ 4 4 0 により保持されるべきか否かを判定する（ステップ S 6 0 6）。サーバキャッシュ情報制御手段 4 5 0 は、分割領域がサーバキャッシュ 4 4 0 により保持されるべき場合（ステップ S 6 0 6 : Y e s）にはステップ S 6 0 8 の処理へ推移し、ストレージキャッシュ 1 3 8 により保持されるべき場合（ステップ S 6 0 6 : N o）にはステップ S 6 0 7 の処理へ推移する。

30

【 0 1 9 8 】

サーバ 3 4 0 のサーバキャッシュ情報制御手段 4 5 0 は、サーバキャッシュ 4 4 0 の特定の分割領域を破棄し（ステップ S 6 0 7）、ステップ S 6 0 1 の処理へ戻る。

【 0 1 9 9 】

サーバ 3 4 0 サーバキャッシュ情報制御手段 4 5 0 は、は、ストレージキャッシュ 1 3 8 の特定の分割領域（キャッシュデータ）を破棄させるためのコマンドをストレージ 1 4 0 へ送信し（ステップ S 6 0 8）、ステップ S 6 0 9 の処理へ推移する。

【 0 2 0 0 】

ストレージ 1 4 0 のアクセス情報制御手段 1 4 3 は、破棄コマンドを受信し、指定された分割領域が独立データであるか否かを判定する（ステップ S 6 0 9）。アクセス情報制御手段 1 4 3 は、独立データである場合（ステップ S 6 0 9 : Y e s）にはステップ S 6 1 0 の処理へ推移し、共有データである場合（ステップ S 6 0 9 : N o）にはステップ S 6 1 1 の処理へ推移する。

40

【 0 2 0 1 】

ストレージ 1 4 0 のアクセス情報制御手段 1 4 3 は、特定の分割領域をストレージキャッシュ 1 3 8 から破棄し（ステップ S 6 1 0）、ステップ S 6 0 1 の処理へ戻る。

【 0 2 0 2 】

ストレージ 1 4 0 のアクセス情報制御手段 1 4 3 は、サーバ 3 4 0 から特定の分割領域に関する要求履歴からサーバ 3 4 0 を消去し（ステップ S 6 1 1）、ステップ S 6 0 1 の処理へ戻る。

50

## 【0203】

以上説明したように、本実施形態のサーバ340は、サーバキャッシュ440に保持されている共有データを破棄できる。従って、本実施形態のサーバ340では、本発明の第4の実施形態における効果に加えて、限られたサーバキャッシュの容量の利用効率を向上することができるか、又はサーバキャッシュの容量を低減することができるという効果がある。

## 【0204】

また、本実施形態のサーバ340、341、342及びストレージ140では、サーバキャッシュとストレージキャッシュに同一のデータが保持される可能性が低い。従って、本実施形態のサーバ340、341、342又はストレージ140では、本発明の第4の実施形態における効果に加えて、限られたサーバキャッシュ若しくはストレージキャッシュの容量の利用効率を向上することができるか、又はサーバキャッシュ若しくはストレージキャッシュの容量を低減することができるという効果がある。

## 【0205】

なお、上述した各実施形態におけるキャッシュ装置は、専用の装置によって実現してもよいが、コンピュータ（情報処理装置）によっても実現可能である。この場合、係るコンピュータは、メモリ（不図示）に格納されたソフトウェア・プログラムをCPU（Central Processing Unit、不図示）に読み出し、読み出したソフトウェア・プログラムをCPUにおいて実行することにより、実行結果を、例えば、ユーザ・インタフェースに出力する。上述した各実施形態及び変形例の場合、係るソフトウェア・プログラムには、上述したところの、図1に示したキャッシュ装置100の各手段、図4に示したキャッシュ装置110の各手段、又は図7に示したキャッシュ装置120の各手段の機能を実現可能な記述がなされていればよい。但し、キャッシュ装置100のキャッシュ手段101は、キャッシュ装置110のキャッシュ手段101、又はキャッシュ装置120のキャッシュ手段101は、適宜ハードウェアを含むことも想定される。そして、このような場合、係るソフトウェア・プログラム（コンピュータ・プログラム）は、本発明を構成すると捉えることができる。更に、係るソフトウェア・プログラムを格納した、コンピュータ読み取り可能な記憶媒体も、本発明を構成すると捉えることができる。

## 【0206】

以上、本発明を、上述した各実施形態およびその変形例によって例示的に説明した。しかしながら、本発明の技術的範囲は、上述した各実施形態およびその変形例に記載した範囲には限定されない。当業者には、係る実施形態に対して多様な変更又は改良を加えることが可能であることは明らかである。そのような場合、係る変更又は改良を加えた新たな実施形態も、本発明の技術的範囲に含まれ得る。そしてこのことは、特許請求の範囲に記載した事項から明らかである。

## 【産業上の利用可能性】

## 【0207】

本発明は、複数のサーバを含むシステムにおいて、各サーバにより、共有される、メモリ、SSD、ハードディスク、ディスクアレイ、データベース、サーバ等がアクセスされる際のキャッシュデータを制御する用途に利用できる。

## 【0208】

本発明は、複数の端末を含むシステムにおいて、各端末により、共有される、メモリ、SSD、ハードディスク、ディスクアレイ、データベース、サーバ等がアクセスされる際のキャッシュデータを制御する用途に利用できる。

## 【0209】

本発明は、複数のCPUを含む装置において、各CPUにより、共有される、メモリ、SSD、ハードディスク、ディスクアレイ、データベース、サーバ等がアクセスされる際のキャッシュデータを制御する用途に利用できる。

## 【0210】

本発明は、マルチコアCPUにおいて、各コアにより、共有される、メモリ、SSD、

10

20

30

40

50

ハードディスク、ディスクアレイ、データベース、サーバ等がアクセスされる際のキャッシュデータを制御する用途に利用できる。

【 0 2 1 1 】

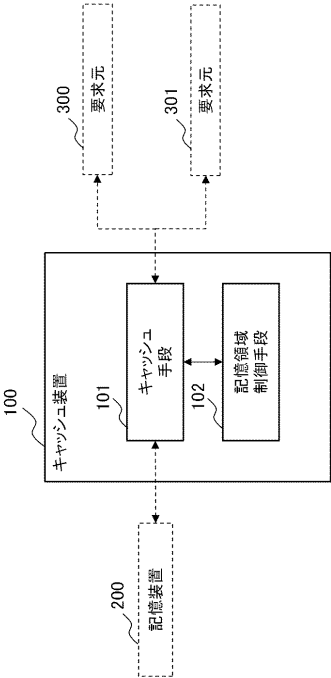
本発明は、マルチスレッドに対応したCPUにおいて、各スレッドにより、共有される、メモリ、SSD、ハードディスク、ディスクアレイ、データベース、サーバ等がアクセスされる際のキャッシュデータを制御する用途に利用できる。

【符号の説明】

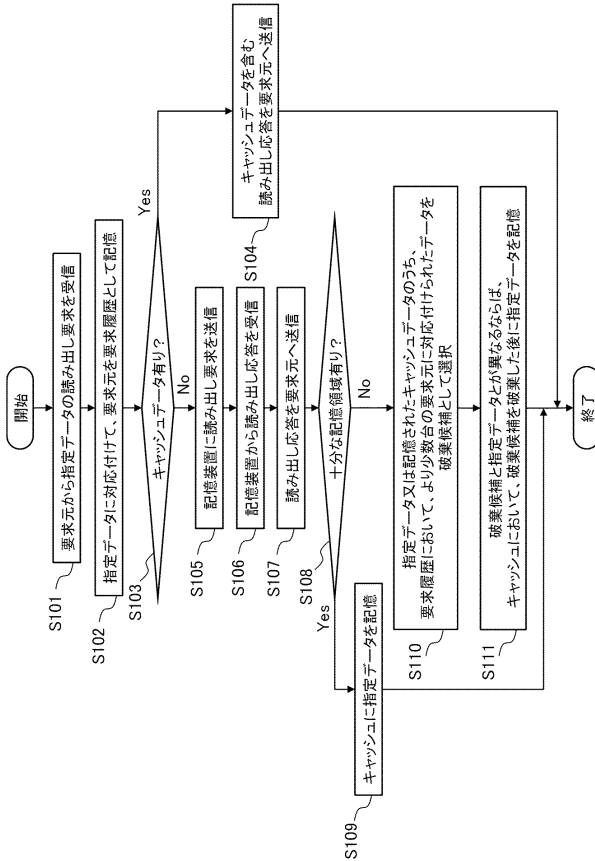
【 0 2 1 2 】

1 0 0	キャッシュ装置	
1 0 1	キャッシュ手段	10
1 0 2	記憶領域制御手段	
2 0 0	記憶装置	
3 0 0、3 0 1	要求元	
1 1 0	キャッシュ装置	
4 0 0、4 0 1	2次キャッシュ装置	
1 2 0	キャッシュ装置	
1 2 3	通信手段	
4 1 0、4 1 1	2次キャッシュ装置	
1 3 0	ストレージ	
1 3 1	ストレージキャッシュ情報制御手段	20
1 3 2	分割領域判定手段	
1 3 3	アクセス情報制御手段	
1 3 4	データ破棄判定手段	
1 3 5	分割領域作成手段	
1 3 8	ストレージキャッシュ	
1 3 9	HDD	
3 3 0、3 3 1、3 3 2	サーバ	
4 3 0、4 3 1、4 3 2	サーバキャッシュ	
1 4 0	ストレージ	
1 4 3	アクセス情報制御手段	30
3 4 0、3 4 1、3 4 2	サーバ	
4 4 0、4 4 1、4 4 2	サーバキャッシュ	
4 5 0、4 5 1、4 5 2	サーバキャッシュ情報制御手段	

【図 1】



【図 2】



【図 3】

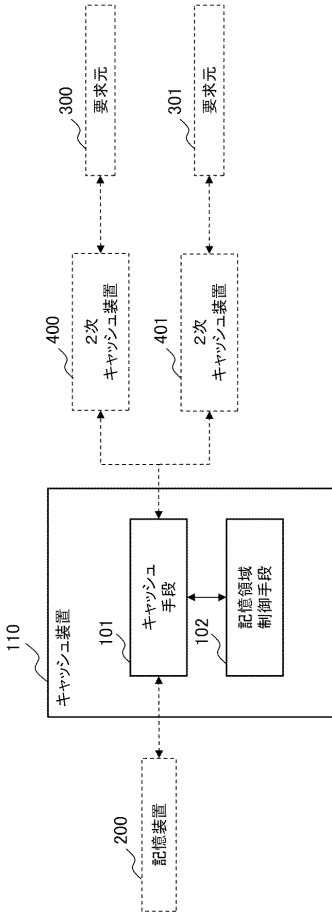
(b) キャッシュに対する操作

時刻	データA	更新	データB	データC
1				
2	更新			
3		更新		
4			更新	
5				更新
6	中継			
7			中継	
8				中継

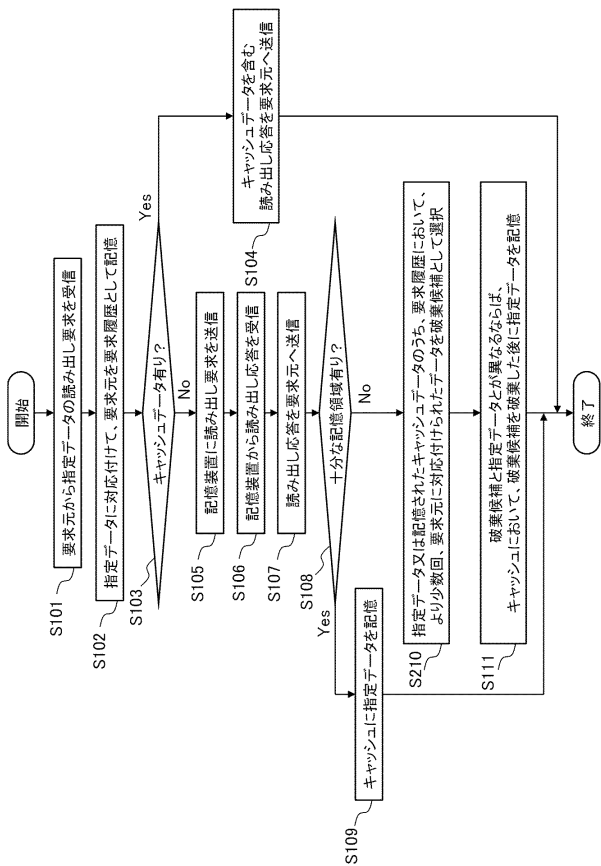
(a) 読み出し要求の時系列

時刻	指定データ	要求元
1	データA	要求元A
2	データB	要求元B
3	データC	要求元A
4	データC	要求元B
5	データA	要求元A
6	データB	要求元B
7	データC	要求元A
8	データC	要求元B

【図 4】



【図5】



【図6】

(b) 要求元A側2次キャッシュに対する操作

時刻	指定データ	要求元
1	データA	要求元A
2	データB	要求元B
3	データA	要求元A
4	データB	要求元B
5	データC	要求元A
6	データC	要求元B
7	データA	要求元A
8	データB	要求元B
9	データA	要求元A
10	データB	要求元B
11	データC	要求元A
12	データC	要求元B

時刻	2次キャッシュデータ(容量1)		
	データA	データB	データC
1	更新		
2	維持		
3	ヒット		
4	維持		
5	維持		中継
6	維持		
7	ヒット		
8	維持		
9	ヒット		
10	維持		
11	維持		中継
12	維持		

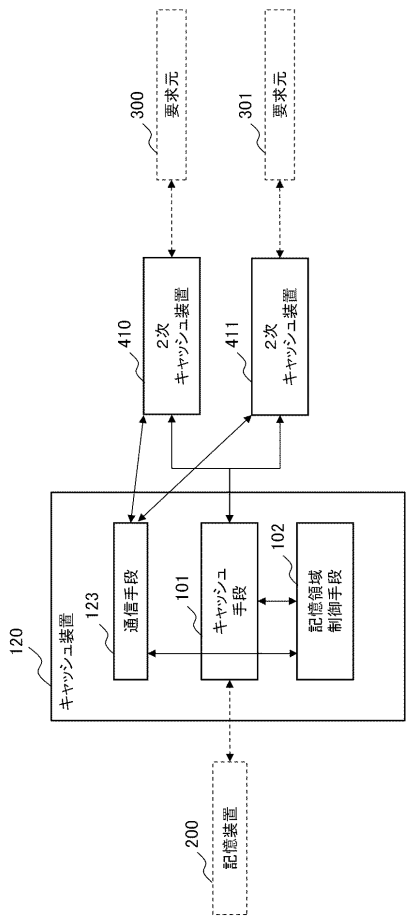
(c) 要求元B側2次キャッシュに対する操作

時刻	2次キャッシュデータ(容量1)		
	データA	データB	データC
1			
2		更新	
3		維持	
4		ヒット	
5		維持	
6		維持	中継
7		維持	
8		ヒット	
9		維持	
10		ヒット	
11		維持	
12		維持	中継

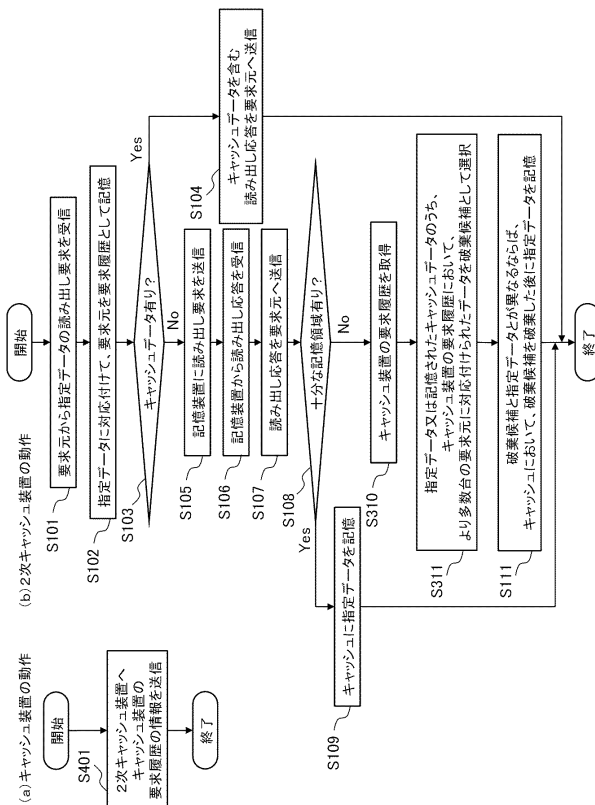
(d) キャッシュに対する操作

時刻	キャッシュデータ(容量1)		
	データA	データB	データC
1	更新		
2		更新	
3		維持	
4		維持	
5			更新
6			ヒット
7			維持
8			維持
9			維持
10			維持
11			ヒット
12			ヒット

【図7】



【図8】



【図9】

(a) 読み出し要求の時系列

時刻	指定データ	要求元
1	データA	要求元A
2	データB	要求元B
3	データC	要求元A
4	データC	要求元B
5	データA	要求元A
6	データB	要求元B
7	データC	要求元A
8	データC	要求元B

(b) 要求元A側 2次キャッシュに対する操作

時刻	2次キャッシュデータ(容量1)		
	データA	データB	データC
1	更新		
2	維持		
3	維持		中継
4	維持		
5	ヒット		
6	維持		
7	維持		中継
8	維持		

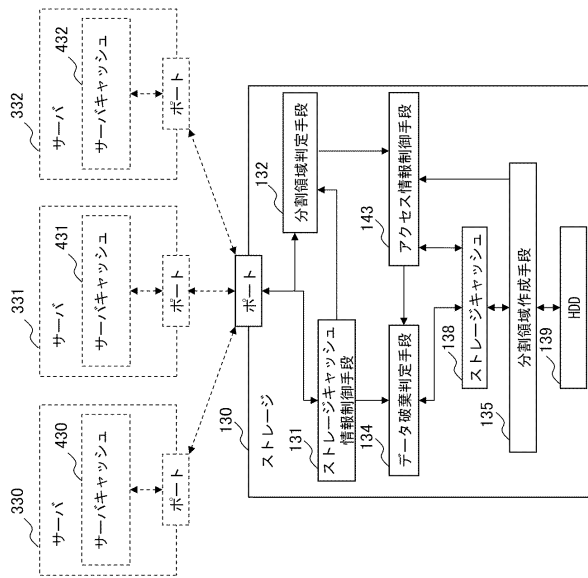
(c) 要求元B側 2次キャッシュに対する操作

時刻	2次キャッシュデータ(容量1)		
	データA	データB	データC
1			
2		更新	
3		維持	
4		維持	中継
5		維持	
6		ヒット	
7		維持	
8		維持	中継

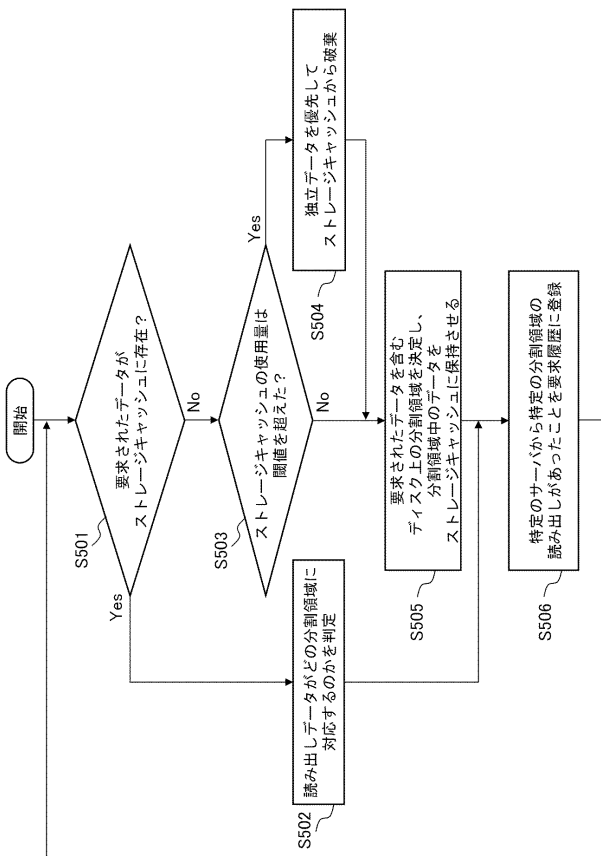
(d) キャッシュに対する操作

時刻	キャッシュデータ(容量1)		
	データA	データB	データC
1	更新		
2		更新	
3			更新
4			ヒット
5			維持
6			維持
7			ヒット
8			ヒット

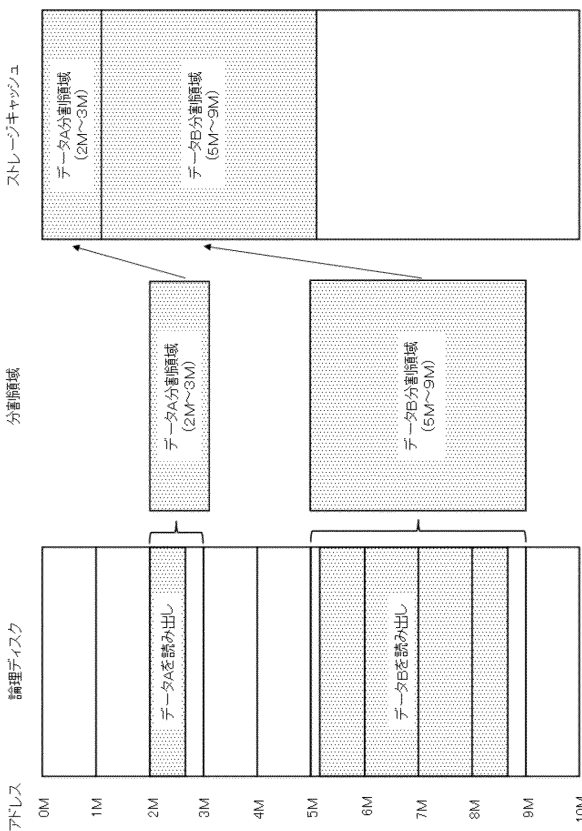
【図10】



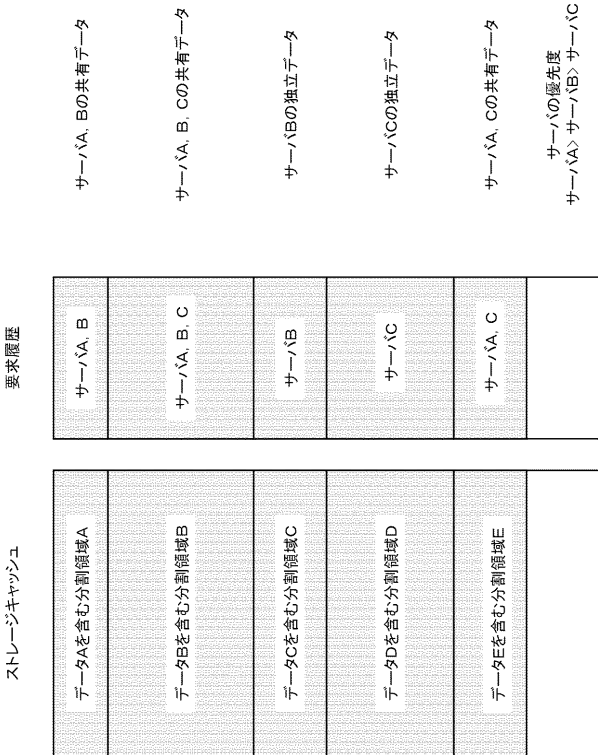
【図11】



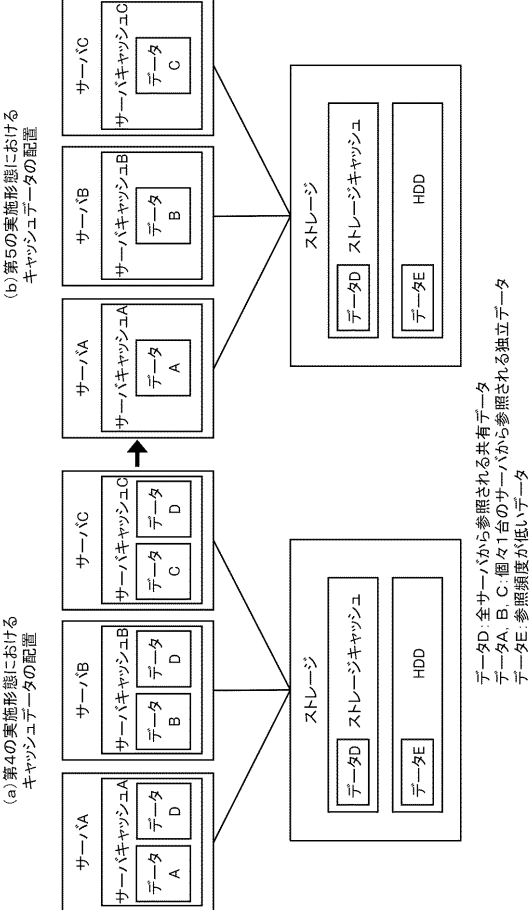
【図12】



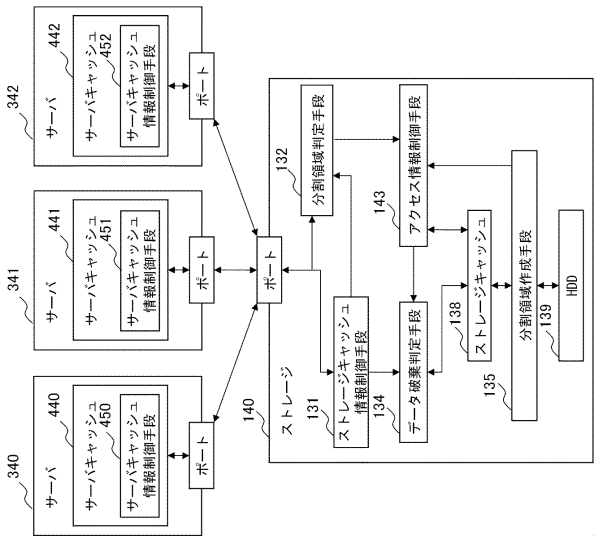
【図 13】



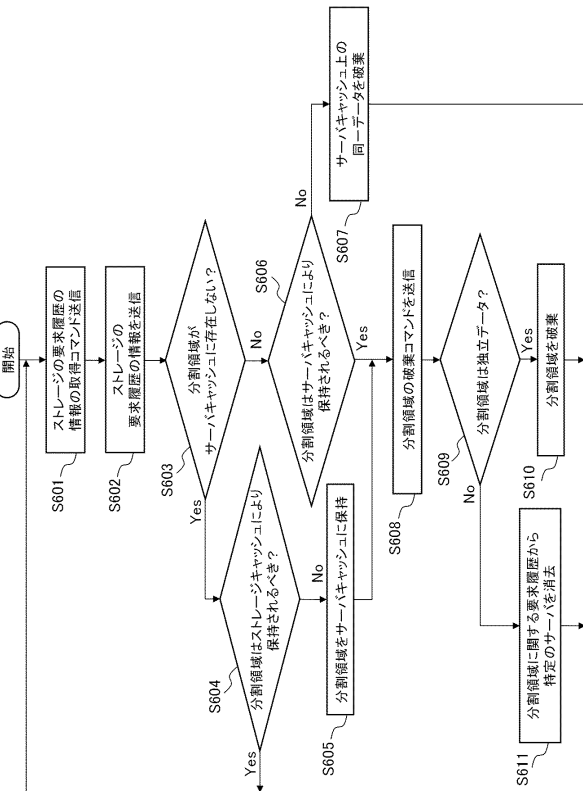
【図 14】



【図 15】



【図 16】



---

フロントページの続き

- (56)参考文献 国際公開第2011/161829(WO, A1)  
米国特許出願公開第2013/0111138(US, A1)  
特開2001-117817(JP, A)  
米国特許第6629200(US, B1)  
長廻 雄介 外, “上位キャッシュの動作を考慮した下位キャッシュの置換手法”, 情報処理学会第73回(平成23年)全国大会 講演論文集(1) アーキテクチャ ソフトウェア科学・工学 データベースとメディア, 日本, 一般社団法人情報処理学会, 2011年 3月 2日, pp. 1-179~1-180

(58)調査した分野(Int.Cl., DB名)

G06F 3/06 - 3/08  
G06F 12/08 - 12/12