

(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2016-110305
(P2016-110305A)

(43) 公開日 平成28年6月20日 (2016. 6. 20)

(51) Int. Cl.	F I	テーマコード (参考)
G06F 3/06 (2006.01)	G06F 3/06 301Z	5B011
G06F 1/32 (2006.01)	G06F 3/06	
G06F 12/00 (2006.01)	G06F 3/06 304N	
	G06F 1/00 332Z	
	G06F 12/00 501B	

審査請求 未請求 請求項の数 6 O L (全 19 頁)

(21) 出願番号 特願2014-245504 (P2014-245504)
(22) 出願日 平成26年12月4日 (2014. 12. 4)

(71) 出願人 00005223
富士通株式会社
神奈川県川崎市中原区上小田中4丁目1番1号
(74) 代理人 100092152
弁理士 服部 毅巖
(72) 発明者 忽滑谷 淳史
神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内
(72) 発明者 橋本 剛
神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内
Fターム(参考) 5B011 EA02 EB07 FF01 FF04 LL11 MA04

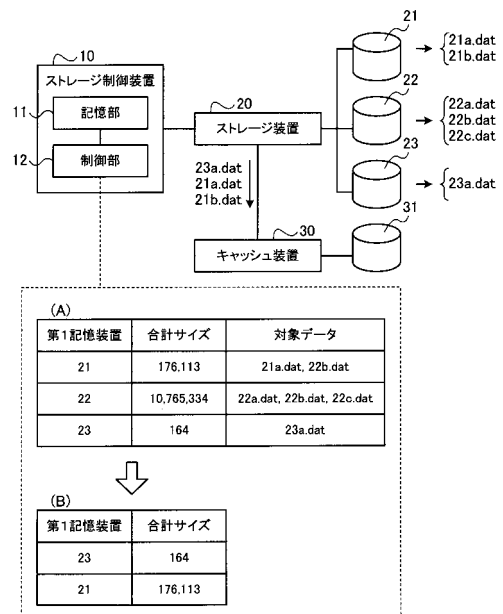
(54) 【発明の名称】 ストレージ制御装置、キャッシュ制御方法、キャッシュ制御プログラム、及びコンピュータシステム

(57) 【要約】

【課題】消費電力の低減効果を高めること。

【解決手段】複数の第1記憶装置21、22、23の各々に格納されたデータのうち設定された条件を満たすデータを第1記憶装置21、22、23とは異なる第2記憶装置31に格納させるストレージ制御装置10が提供される。ストレージ制御装置10は、第2記憶装置31に格納させるデータのサイズを取得し、第1記憶装置21、22、23毎に、第2記憶装置31に格納させるデータの合計サイズを計算し、合計サイズが小さい第1記憶装置23、21から順に、設定された条件を満たすデータを第2記憶装置31に格納させる制御部12を有する。

【選択図】 図1



【特許請求の範囲】**【請求項 1】**

複数の第 1 記憶装置の各々に格納されたデータのうち設定された条件を満たすデータを前記第 1 記憶装置とは異なる第 2 記憶装置に格納させるストレージ制御装置であって、
前記第 2 記憶装置に格納させるデータのサイズを取得し、
前記第 1 記憶装置毎に、前記第 2 記憶装置に格納させるデータの合計サイズを計算し、
前記合計サイズが小さい前記第 1 記憶装置から順に、前記設定された条件を満たすデータを前記第 2 記憶装置に格納させる制御部を有する
ストレージ制御装置。

【請求項 2】

前記制御部は、データに対するアクセス頻度が設定された閾値より大きい場合に該データが前記設定された条件を満たすと判定する
請求項 1 に記載のストレージ制御装置。

【請求項 3】

前記設定された条件を満たすデータを前記第 2 記憶装置に格納した前記第 1 記憶装置は、消費電力が小さい動作モードへと移行する
請求項 2 に記載のストレージ制御装置。

【請求項 4】

複数の第 1 記憶装置の各々に格納されたデータのうち設定された条件を満たすデータを前記第 1 記憶装置とは異なる第 2 記憶装置に格納させるキャッシュ制御方法であって、
コンピュータが、
前記第 2 記憶装置に格納させるデータのサイズを取得し、
前記第 1 記憶装置毎に、前記第 2 記憶装置に格納させるデータの合計サイズを計算し、
前記合計サイズが小さい前記第 1 記憶装置から順に、前記設定された条件を満たすデータを前記第 2 記憶装置に格納させる
キャッシュ制御方法。

【請求項 5】

複数の第 1 記憶装置の各々に格納されたデータのうち設定された条件を満たすデータを前記第 1 記憶装置とは異なる第 2 記憶装置に格納させるキャッシュ制御プログラムであって、
コンピュータに、
前記第 2 記憶装置に格納させるデータのサイズを取得し、
前記第 1 記憶装置毎に、前記第 2 記憶装置に格納させるデータの合計サイズを計算し、
前記合計サイズが小さい前記第 1 記憶装置から順に、前記設定された条件を満たすデータを前記第 2 記憶装置に格納させる
処理を実行させる、キャッシュ制御プログラム。

【請求項 6】

複数の第 1 記憶装置を制御するストレージ装置と、
前記第 1 記憶装置とは異なる第 2 記憶装置を制御するキャッシュ装置と、
前記ストレージ装置を制御して、前記複数の第 1 記憶装置の各々に格納されたデータのうち設定された条件を満たすデータを前記第 2 記憶装置に格納させるストレージ制御装置と、
を有し、
前記ストレージ制御装置は、前記第 2 記憶装置に格納させるデータのサイズを取得し、
前記第 1 記憶装置毎に、前記第 2 記憶装置に格納させるデータの合計サイズを計算し、
前記合計サイズが小さい前記第 1 記憶装置から順に、前記設定された条件を満たすデータを前記第 2 記憶装置に格納させる制御部を有する
コンピュータシステム。

【発明の詳細な説明】**【技術分野】**

10

20

30

40

50

【 0 0 0 1 】

本発明は、ストレージ制御装置、キャッシュ制御方法、キャッシュ制御プログラム、及びコンピュータシステムに関する。

【 背景技術 】

【 0 0 0 2 】

近年、R A I D (Redundant Arrays of Inexpensive Disks) 技術などにより複数の記憶装置 (例えば、H D D (Hard Disk Drive) や S S D (Solid State Drive)) を冗長化し、高速なデータアクセス及び大容量を実現したディスクアレイが広く利用されている。他方、ディスクアレイの容量を増やすために記憶装置の数が増えるとディスクアレイが消費する電力量が大きくなるため、電力コストや環境負荷を低減する見地から、ディスクアレイの省電力化に関する研究開発が進められている。

10

【 0 0 0 3 】

ディスクアレイの消費電力を実現する技術として M A I D (Massive Array of Idle Disk) と呼ばれるものがある。M A I D は、ディスクアレイの中で、データアクセスが発生していない記憶装置を省電力モードに遷移 (例えば、ディスク回転数の低下・停止、ディスクヘッドの退避、電源オフなど) させて消費電力を削減する技術である。M A I D を適用したディスクアレイではデータアクセスのある記憶装置が通常通り稼働するため、高速性や信頼性に対する影響が少ない。

【 0 0 0 4 】

但し、省電力モードにある記憶装置に対してデータアクセスが生じると、その記憶装置を通常の動作モードへと復帰させるための時間がかかり、データアクセスが遅延する。この遅延を抑制するために、書き込み対象のデータについてデータアクセスの確率を推定し、推定した確率に基づいて書き込み先の記憶装置を決定する技術が提案されている。また、ディスクアレイを含むストレージ装置と、データアクセスが集中したデータをキャッシュするためのキャッシュサーバとを有する分散ストレージシステムが提案されている。

20

【 先行技術文献 】

【 特許文献 】

【 0 0 0 5 】

【 特許文献 1 】 国際公開 2 0 1 1 / 1 2 1 8 6 9 号公報

【 特許文献 2 】 特開 2 0 1 0 - 0 0 2 9 4 2 号公報

30

【 発明の概要 】

【 発明が解決しようとする課題 】

【 0 0 0 6 】

データアクセスの確率が低いデータを多く格納した記憶装置を省電力モードへ遷移させれば、省電力モードにある記憶装置にデータアクセスが発生して高速性が損なわれるリスクを低減することはできる。しかし、推定精度が低ければリスク低減の効果は期待できない。また、データアクセスの頻度が高いデータをキャッシュサーバに退避しておけば、退避したデータに対するアクセスが生じた場合にキャッシュサーバが応答できるため、データアクセスの高速性が損なわれるリスクを低減することができる。

【 0 0 0 7 】

ただ、キャッシュサーバの記憶容量にも限りがあるため、全ての記憶装置におけるアクセス頻度の高いデータをキャッシュすることは現実的に難しい。アクセス頻度が高いデータを全てキャッシュした記憶装置を省電力モードへと遷移させ、他の記憶装置は通常の動作モードで稼働させると、データアクセスの高速性は維持されるが、省電力モードへ遷移できる記憶装置の数が限られる。その結果、消費電力の低減効果が限定的となる。

40

【 0 0 0 8 】

そこで、1つの側面によれば、本発明の目的は、消費電力の低減効果を高めることが可能なストレージ制御装置、キャッシュ制御方法、キャッシュ制御プログラム、及びコンピュータシステムを提供することにある。

【 課題を解決するための手段 】

50

【 0 0 0 9 】

本開示の1つの側面によれば、複数の第1記憶装置の各々に格納されたデータのうち設定された条件を満たすデータを第1記憶装置とは異なる第2記憶装置に格納させるストレージ制御装置が提供される。該ストレージ制御装置は、第2記憶装置に格納させるデータのサイズを取得し、第1記憶装置毎に、第2記憶装置に格納させるデータの合計サイズを計算し、合計サイズが小さい第1記憶装置から順に、設定された条件を満たすデータを第2記憶装置に格納させる制御部を有する。

【 発明の効果 】

【 0 0 1 0 】

本発明によれば、消費電力の低減効果を高めることが可能になる。

10

【 図面の簡単な説明 】

【 0 0 1 1 】

【 図 1 】 第1実施形態に係るコンピュータシステムの一例を示した図である。

【 図 2 】 第2実施形態に係るストレージシステムの一例を示した図である。

【 図 3 】 第2実施形態に係るストレージシステムで実行される処理の流れを示したシーケンス図である。

【 図 4 】 第2実施形態に係るメタサーバの機能を実現可能なハードウェアの一例を示した図である。

【 図 5 】 第2実施形態に係るメタサーバ、データサーバ、及びキャッシュサーバの主な機能を示したブロック図である。

20

【 図 6 】 第2実施形態に係る管理用データベースの一例を示した図である。

【 図 7 】 第2実施形態に係る対象ファイルリストの一例を示した図である。

【 図 8 】 第2実施形態に係る暫定対象ディスクリスト及び対象ディスクリストの一例を示した図である。

【 図 9 】 第2実施形態に係るメタサーバが実行する処理の流れを示した第1のフロー図である。

【 図 1 0 】 第2実施形態に係るメタサーバが実行する処理の流れを示した第2のフロー図である。

【 図 1 1 】 第2実施形態に係るデータサーバが実行する処理の流れを示したフロー図である。

30

【 図 1 2 】 第2実施形態に係るキャッシュ方法の効果の一例を示した図である。

【 発明を実施するための形態 】

【 0 0 1 2 】

以下に添付図面を参照しながら、本発明の実施形態について説明する。なお、本明細書及び図面において実質的に同一の機能を有する要素については、同一の符号を付することにより重複説明を省略する場合がある。

【 0 0 1 3 】

< 1 . 第1実施形態 >

図1を参照しながら、第1実施形態について説明する。

図1は、第1実施形態に係るコンピュータシステムの一例を示した図である。

40

【 0 0 1 4 】

図1に示すように、第1実施形態に係るコンピュータシステムは、ストレージ制御装置10、ストレージ装置20、及びキャッシュ装置30を有する。

ストレージ装置20は、直接又はネットワークを介して接続された第1記憶装置21、22、23に対するデータの読み書きを管理する。キャッシュ装置30は、直接又はネットワークを介して接続された第2記憶装置31に対するデータの読み書きを管理する。

【 0 0 1 5 】

なお、ストレージ装置20の数、第1記憶装置の数、キャッシュ装置30の数、第2記憶装置の数は任意に設定することができる。また、ストレージ制御装置10、ストレージ装置20、キャッシュ装置30は、有線又は無線のネットワークを介して接続されていて

50

もよいし、専用回線を介して接続されていてもよい。第1実施形態に係る技術は、小規模なRAID装置にも適用できるし、大規模な分散ストレージシステムにも適用できる。

【0016】

ストレージ制御装置10は、記憶部11及び制御部12を有する。

記憶部11は、RAM(Random Access Memory)などの揮発性記憶装置、或いは、HDDやフラッシュメモリなどの不揮発性記憶装置である。制御部12は、CPU(Central Processing Unit)やDSP(Digital Signal Processor)などのプロセッサである。但し、制御部12は、ASIC(Application Specific Integrated Circuit)やFPGA(Field Programmable Gate Array)などの電子回路であってもよい。制御部12は、例えば、記憶部11又は他のメモリに記憶されたプログラムを実行する。

10

【0017】

記憶部11には、第1記憶装置21、22、23に格納されたデータに関する情報が格納される。データに関する情報は、例えば、データを識別するための識別情報(ファイル名など)、データが格納された第1記憶装置の識別情報及び該第1記憶装置を管理するストレージ装置20の識別情報(IPアドレスなど)、データサイズなどを含む。

【0018】

制御部12は、ストレージ装置20を制御し、第1記憶装置21、22、23の各々に格納されたデータのうち設定された条件を満たすデータを第2記憶装置31に格納させる。設定された条件は、例えば、対象のデータに対するアクセス頻度が、予め設定された閾値より大きいことなどである。設定された条件を満たすデータを第2記憶装置31に格納させる制御は、例えば、新たなキャッシュ装置又は第2記憶装置が追加された場合や、制御の実施後に予め設定された時間が経過した場合に実施される。

20

【0019】

上記の制御に際し、制御部12は、第2記憶装置31に格納させるデータのサイズを取得する。このとき、制御部12は、データのサイズをストレージ装置20に問い合わせ、ストレージ装置20からデータのサイズを取得してもよいし、記憶部11に格納されたデータに関する情報からデータのサイズを取得してもよい。

【0020】

制御部12は、第1記憶装置21、22、23毎に、第2記憶装置31に格納させるデータの合計サイズを計算する。

30

例えば、図1(A)に示すように、制御部12は、第1記憶装置21に格納されているデータのうち上記設定された条件を満たすデータ(図1の例では21a.dat、21b.dat)を対象データとして特定し、対象データの合計サイズを計算する。また、制御部12は、第1記憶装置21の識別情報と、計算した対象データの合計サイズとを対応付けて記憶部11に記憶させる。同様に、制御部12は、第1記憶装置22、23についても対象データの合計サイズを計算し、計算した対象データの合計サイズを識別情報に対応付けて記憶部11に記憶させる。

【0021】

制御部12は、合計サイズが小さい第1記憶装置23、21から順に、設定された条件を満たすデータを第2記憶装置31に格納させる。

40

例えば、図1(B)に示すように、制御部12は、記憶部11に記憶させた合計サイズを参照し、合計サイズが小さい順に第1記憶装置(この例では第1記憶装置23、21)を特定する。このとき、制御部12は、特定した第1記憶装置に対応する合計サイズの計算値が第2記憶装置31の空き容量より小さくなるように、第1記憶装置を特定する。そして、制御部12は、第1記憶装置23、21の対象データを第2記憶装置31に格納するように制御する指示をストレージ装置20に送信する。

【0022】

上記指示を受信したストレージ装置20は、第1記憶装置23、21から対象データ(この例では23a.dat、21a.dat、21b.dat)を読み出し、キャッシュ装置30に送信する。キャッシュ装置30は、ストレージ装置20から受信した対象デー

50

データを第2記憶装置31に格納する。対象データの第2記憶装置31への格納が完了した場合、キャッシュ装置30は、格納の完了を示す応答をストレージ装置20に送信する。

【0023】

上記応答を受信したストレージ装置20は、第2記憶装置31に格納した対象データを格納している第1記憶装置23、21を省電力モードへと遷移させる。省電力モードへの遷移は、例えば、ディスク回転数の低下・停止、ディスクヘッドの退避、電源オフなどである。そして、ストレージ装置20は、省電力モードへの遷移が完了したことを示す通知をストレージ制御装置10へと送信する。この通知を受信した制御部12は、省電力モードにある第1記憶装置23、21の情報を記憶部11に記憶させる。

【0024】

第1実施形態に係る技術を適用すれば、合計サイズが小さい順に第2記憶装置31へと対象データがキャッシュされるため、より多くの第1記憶装置を省電力モードに遷移させることができ、システム全体の電力消費量をより多く削減することが可能になる。また、対象データ(アクセス頻度が高いデータなど)が第2記憶装置31にキャッシュされているため、省電力モードから通常の動作モードへ復帰する処理に伴う遅延のリスクが低く、データアクセスの高速性も維持される。

【0025】

以上、第1実施形態について説明した。

<2. 第2実施形態>

次に、第2実施形態について説明する。

【0026】

[2-1. システム]

図2は、第2実施形態に係るストレージシステムの一例を示した図である。図2に示したストレージシステム50は、第2実施形態に係るストレージシステムの一例である。

【0027】

図2に示すように、ストレージシステム50は、メタサーバ100、データサーバ201、202、キャッシュサーバ300、及びクライアント装置400を含む。

データサーバ201には、記憶装置211、212、213が接続されている。データサーバ202には、記憶装置221、222、223が接続されている。キャッシュサーバ300には記憶装置301が接続されている。記憶装置211、212、213、221、222、223、301は1台又は複数台のHDDやSSDなどである。

【0028】

クライアント装置400は、メタサーバ100に対し、データの書き込みやデータの読み出しを要求する。メタサーバ100は、クライアント装置400から要求を受け、要求の対象となるデータの格納場所をクライアント装置400に通知する。

【0029】

データサーバ201は、データの格納場所となる記憶装置211、212、213を管理する。データサーバ202は、データの格納場所となる記憶装置221、222、223を管理する。クライアント装置400は、メタサーバ100から通知された格納場所を管理するデータサーバへアクセスし、データの書き込みやデータの読み出しを実行する。

【0030】

キャッシュサーバ300は、キャッシュディスクとして機能する記憶装置301を管理する。記憶装置301には、記憶装置211、212、213、221、222、223に格納されたデータの一部がコピーされる。コピー元の記憶装置が省電力モード(例えば、ディスク回転数の低下・停止、ディスクヘッドの退避、電源オフなど)にあるとき、その記憶装置に代えて記憶装置301が利用される。

【0031】

ここで、キャッシュディスクへのデータコピー及びキャッシュサーバ300へのアクセスに関する処理の流れについて、図3を参照しながら説明する。図3は、第2実施形態に係るストレージシステムで実行される処理の流れを示したシーケンス図である。

10

20

30

40

50

【 0 0 3 2 】

図 3 に示した処理は、キャッシュディスクとして利用可能な記憶装置が追加された場合、或いは、当該処理を前回実行してから一定の期間が経過した場合に実行される。

(S 1 1) メタサーバ 1 0 0 は、データサーバ 2 0 1 に対し、記憶装置 2 1 1、2 1 2、2 1 3 に格納されたデータをキャッシュサーバ 3 0 0 の記憶装置 3 0 1 にコピーするように指示する(キャッシュ指示)。このとき、メタサーバ 1 0 0 は、データサーバ 2 0 1 が管理する記憶装置 2 1 1、2 1 2、2 1 3 に格納されたデータのうちアクセス頻度が高いデータを記憶装置 3 0 1 にコピーするように指示する。

【 0 0 3 3 】

(S 1 2) メタサーバ 1 0 0 は、データサーバ 2 0 2 に対し、記憶装置 2 2 1、2 2 2、2 2 3 に格納されたデータをキャッシュサーバ 3 0 0 の記憶装置 3 0 1 にコピーするように指示する(キャッシュ指示)。このとき、メタサーバ 1 0 0 は、データサーバ 2 0 2 が管理する記憶装置 2 2 1、2 2 2、2 2 3 に格納されたデータのうちアクセス頻度が高いデータを記憶装置 3 0 1 にコピーするように指示する。

【 0 0 3 4 】

なお、S 1 1、S 1 2 の処理は順序を入れ替えることができる。

(S 1 3) キャッシュ指示を受けたデータサーバ 2 0 1 は、メタサーバ 1 0 0 から指示されたアクセス頻度の高いデータを記憶装置 2 1 1、2 1 2、2 1 3 から読み出す。そして、データサーバ 2 0 1 は、読み出したデータをキャッシュサーバ 3 0 0 へ送信し、記憶装置 3 0 1 に格納させる(データコピー)。

【 0 0 3 5 】

(S 1 4) データコピーを完了したデータサーバ 2 0 1 は、記憶装置 2 1 1、2 1 2、2 1 3 のうち、アクセス頻度の高いデータを全てコピーし終えた記憶装置を省電力モードへと遷移させる。このように、記憶装置を省電力モードへと遷移させることで、ストレージシステム 5 0 の消費電力を抑制することができる。

【 0 0 3 6 】

(S 1 5) キャッシュ指示を受けたデータサーバ 2 0 2 は、メタサーバ 1 0 0 から指示されたアクセス頻度の高いデータを記憶装置 2 2 1、2 2 2、2 2 3 から読み出す。そして、データサーバ 2 0 2 は、読み出したデータをキャッシュサーバ 3 0 0 へ送信し、記憶装置 3 0 1 に格納させる(データコピー)。

【 0 0 3 7 】

(S 1 6) データコピーを完了したデータサーバ 2 0 2 は、記憶装置 2 2 1、2 2 2、2 2 3 のうち、アクセス頻度の高いデータを全てコピーし終えた記憶装置を省電力モードへと遷移させる。このように、記憶装置を省電力モードへと遷移させることで、ストレージシステム 5 0 の消費電力を抑制することができる。

【 0 0 3 8 】

(S 1 7、S 1 8) クライアント装置 4 0 0 は、データを読み出す際に、読み出し対象のデータがキャッシュサーバ 3 0 0 の記憶装置 3 0 1 に格納済みであるか(つまり、キャッシュがあるか)を確認する。この例では、読み出し対象のデータが記憶装置 3 0 1 に格納済みであるとする。この場合、クライアント装置 4 0 0 は、キャッシュサーバ 3 0 0 へアクセスし、記憶装置 3 0 1 から読み出し対象のデータを取得する。

【 0 0 3 9 】

上記のように、キャッシュサーバ 3 0 0 の記憶装置 3 0 1 にアクセス頻度の高いデータをコピーすることで、データサーバ 2 0 1、2 0 2 が管理する記憶装置を省電力モードにすることができ、ストレージシステム 5 0 の消費電力を抑制することができる。

【 0 0 4 0 】

また、省電力モードにした記憶装置へのアクセスが発生しても、データがキャッシュされていれば、クライアント装置 4 0 0 は、キャッシュサーバ 3 0 0 にアクセスして読み出し対象のデータを取得することができる。つまり、多くの場合に、省電力モードの記憶装置を通常の動作モードへと復帰させずに所望のデータが得られるため、アクセス速度が低

10

20

30

40

50

下するリスクを抑制することができる。

【 0 0 4 1 】

上記のように、キャッシュサーバ 3 0 0 を利用することでアクセス性能の低下を抑制しつつ省電力化を図れるが、より多くの記憶装置を省電力モードへ遷移させることができれば、さらなる電力消費の抑制が可能になる。

【 0 0 4 2 】

但し、省電力化とアクセス性能の向上とはトレードオフの関係にあるため、通常は省電力化を進めるとアクセス性能が低下する傾向にある。そこで、第 2 実施形態は、キャッシュ方法を工夫することによってアクセス性能の低下を抑えつつ、ストレージシステム 5 0 における消費電力を低減する技術を提供する。

10

【 0 0 4 3 】

[2 - 2 . ハードウェア]

ここで、図 4 を参照しながら、メタサーバ 1 0 0 のハードウェアの一例について説明する。図 4 は、第 2 実施形態に係るメタサーバの機能を実現可能なハードウェアの一例を示した図である。

【 0 0 4 4 】

メタサーバ 1 0 0 が有する機能は、例えば、図 4 に示す情報処理装置のハードウェア資源を用いて実現することが可能である。つまり、メタサーバ 1 0 0 が有する機能は、コンピュータプログラムを用いて図 4 に示すハードウェアを制御することにより実現される。

【 0 0 4 5 】

図 4 に示すように、このハードウェアは、主に、CPU 9 0 2 と、ROM (Read Only Memory) 9 0 4 と、RAM 9 0 6 と、ホストバス 9 0 8 と、ブリッジ 9 1 0 とを有する。さらに、このハードウェアは、外部バス 9 1 2 と、インターフェース 9 1 4 と、入力部 9 1 6 と、出力部 9 1 8 と、記憶部 9 2 0 と、ドライブ 9 2 2 と、接続ポート 9 2 4 と、通信部 9 2 6 とを有する。

20

【 0 0 4 6 】

CPU 9 0 2 は、例えば、演算処理装置又は制御装置として機能し、ROM 9 0 4、RAM 9 0 6、記憶部 9 2 0、又はリムーバブル記録媒体 9 2 8 に記録された各種プログラムに基づいて各構成要素の動作全般又はその一部を制御する。ROM 9 0 4 は、CPU 9 0 2 に読み込まれるプログラムや演算に用いるデータなどを格納する記憶装置の一例である。RAM 9 0 6 には、例えば、CPU 9 0 2 に読み込まれるプログラムや、そのプログラムを実行する際に変化する各種パラメータなどが一時的又は永続的に格納される。

30

【 0 0 4 7 】

これらの要素は、例えば、高速なデータ伝送が可能なホストバス 9 0 8 を介して相互に接続される。一方、ホストバス 9 0 8 は、例えば、ブリッジ 9 1 0 を介して比較的データ伝送速度が低速な外部バス 9 1 2 に接続される。また、入力部 9 1 6 としては、例えば、マウス、キーボード、タッチパネル、タッチパッド、ボタン、スイッチ、及びレバーなどが用いられる。さらに、入力部 9 1 6 としては、赤外線やその他の電波を利用して制御信号を送信することが可能なリモートコントローラが用いられることもある。

【 0 0 4 8 】

出力部 9 1 8 としては、例えば、CRT (Cathode Ray Tube)、LCD (Liquid Crystal Display)、PDP (Plasma Display Panel)、又はELD (Electro-Luminescence Display) などのディスプレイ装置が用いられる。また、出力部 9 1 8 として、スピーカやヘッドホンなどのオーディオ出力装置、又はプリンタなどが用いられることもある。つまり、出力部 9 1 8 は、情報を視覚的又は聴覚的に出力することが可能な装置である。

40

【 0 0 4 9 】

記憶部 9 2 0 は、各種のデータを格納するための装置である。記憶部 9 2 0 としては、例えば、HDD などの磁気記憶デバイスが用いられる。また、記憶部 9 2 0 として、SSD や RAM ディスクなどの半導体記憶デバイス、光記憶デバイス、又は光磁気記憶デバイスなどが用いられてもよい。

50

【 0 0 5 0 】

ドライブ 9 2 2 は、着脱可能な記録媒体であるリムーバブル記録媒体 9 2 8 に記録された情報を読み出し、又はリムーバブル記録媒体 9 2 8 に情報を書き込む装置である。リムーバブル記録媒体 9 2 8 としては、例えば、磁気ディスク、光ディスク、光磁気ディスク、又は半導体メモリなどが用いられる。

【 0 0 5 1 】

接続ポート 9 2 4 は、例えば、U S B (Universal Serial Bus) ポート、I E E E 1 3 9 4 ポート、S C S I (Small Computer System Interface)、R S - 2 3 2 C ポート、又は光オーディオ端子など、外部接続機器 9 3 0 を接続するためのポートである。外部接続機器 9 3 0 としては、例えば、プリンタなどが用いられる。

10

【 0 0 5 2 】

通信部 9 2 6 は、ネットワーク 9 3 2 に接続するための通信デバイスである。通信部 9 2 6 としては、例えば、有線又は無線 L A N (Local Area Network) 用の通信回路、W U S B (Wireless USB) 用の通信回路、光通信用の通信回路やルータ、A D S L (Asymmetric Digital Subscriber Line) 用の通信回路やルータ、携帯電話ネットワーク用の通信回路などが用いられる。通信部 9 2 6 に接続されるネットワーク 9 3 2 は、有線又は無線により接続されたネットワークであり、例えば、インターネット、L A N、放送網、衛星通信回線などを含む。

【 0 0 5 3 】

以上、メタサーバ 1 0 0 のハードウェアの一例について説明した。なお、データサーバ 2 0 1、2 0 2、キャッシュサーバ 3 0 0、及びクライアント装置 4 0 0 の機能もそれぞれ図 4 に示したハードウェア資源を利用して実現可能である。そのため、データサーバ 2 0 1、2 0 2、キャッシュサーバ 3 0 0、及びクライアント装置 4 0 0 のハードウェアについては説明を省略する。

20

【 0 0 5 4 】

[2 - 3 . 機能]

次に、メタサーバ 1 0 0、データサーバ 2 0 1、2 0 2、及びキャッシュサーバ 3 0 0 の機能について説明する。図 5 は、第 2 実施形態に係るメタサーバ、データサーバ、及びキャッシュサーバの主な機能を示したブロック図である。

【 0 0 5 5 】

(メタサーバ 1 0 0)

図 5 に示すように、メタサーバ 1 0 0 は、記憶部 1 0 0 a、制御部 1 0 0 b、及び通信部 1 0 0 c を有する。記憶部 1 0 0 a の機能は、上述した R A M 9 0 6 や記憶部 9 2 0 などを用いて実現できる。制御部 1 0 0 b の機能は、上述した C P U 9 0 2 などを用いて実現できる。通信部 1 0 0 c の機能は、上述した接続ポート 9 2 4 や通信部 9 2 6 などを用いて実現できる。

30

【 0 0 5 6 】

記憶部 1 0 0 a には、図 6 に示すような管理用データベースが格納される。図 6 は、第 2 実施形態に係る管理用データベースの一例を示した図である。この管理用データベースには、記憶装置 2 1 1、2 1 2、2 1 3、2 2 1、2 2 2、2 2 3 に格納されたデータファイルに関する情報が記録される。

40

【 0 0 5 7 】

例えば、管理用データベースには、データファイルのファイル名、データサーバ 2 0 1、2 0 2 の I P アドレス、ディスク識別子、ファイルサイズ、キャッシュの有無を示す情報が記録される。I P アドレスは、データサーバ 2 0 1、2 0 2 を識別する識別情報の一例である。ディスク識別子は、記憶装置 2 1 1、2 1 2、2 1 3、2 2 1、2 2 2、2 2 3 をそれぞれ識別する識別情報の一例である。キャッシュの有無を示す情報は、キャッシュサーバ 3 0 0 の記憶装置 3 0 1 に格納済みか否かを示す情報である。

【 0 0 5 8 】

図 6 の例において、ファイル名「 A A A . d a t 」のデータファイルは、I P アドレス

50

「192.168.1.10」、ディスク識別子「A」、ファイルサイズ「690kB」、キャッシュ「あり」という情報に対応付けられている。

【0059】

この情報から、制御部100bは、データファイルAAA.datが、IPアドレスが192.168.1.10のデータサーバ(データサーバ201)が管理するディスク識別子Aの記憶装置(記憶装置211)に格納されていることを認識できる。また、制御部100bは、データファイルAAA.datのデータサイズが690kBであること、及びデータファイルAAA.datがキャッシュサーバ300の記憶装置301に格納済みであることを認識できる。

【0060】

記憶装置211、212、213、221、222、223に格納されているデータファイルが更新されると、制御部100bは、管理用データベースを更新する。

以下、説明の都合上、記憶装置211のディスク識別子が「A」、記憶装置212のディスク識別子が「B」、記憶装置213のディスク識別子が「C」とする。また、記憶装置221のディスク識別子が「D」、記憶装置222のディスク識別子が「E」、記憶装置223のディスク識別子が「F」とする。また、データサーバ201のIPアドレスが「192.168.1.10」、データサーバ202のIPアドレスが「192.168.1.3」とする。

【0061】

制御部100bは、キャッシュディスクとして利用可能な記憶装置が追加された場合、或いは、前回キャッシュ処理を実行してから一定の期間が経過した場合にキャッシュ処理を実行する。キャッシュ処理とは、データサーバ201、202が管理する記憶装置211、212、213、221、222、223に格納されているアクセス頻度の高いデータファイルをキャッシュサーバ300の記憶装置301にコピーする処理である。

【0062】

キャッシュ処理を実行するため、制御部100bは、データファイル毎のアクセス状況(データの読み出し及び書き込みの回数)をデータサーバ201、202から取得し、記憶装置301にコピーするデータファイルを特定するためのアクセス頻度を計算する。例えば、制御部100bは、キャッシュ処理を実行する度にアクセス頻度を計算する。キャッシュ処理を実行する時刻tにおけるアクセス頻度AW(t)は、例えば、下記の式(1)に基づいて計算される。

【0063】

$$AW(t) = RA(t) + AW(t-1) \times WT$$

... (1)

AW(t-1)は、前回キャッシュ処理を実行した時刻t-1におけるアクセス頻度を表す。また、RA(t)は、前回キャッシュ処理を実行した後、今回キャッシュ処理を実行するまでの期間に、対象のデータファイルに対して生じたアクセスの回数を表す。また、WTは重みであり、条件「 $0 < WT < 1$ 」を満たす値(例えば、 $1/2$)が予め設定される。上記の式(1)は、右辺第1項で最新の一定期間内に生じたアクセス頻度を表現し、右辺第2項で時刻t-1以前のアクセス頻度による影響を表現している。

【0064】

例えば、制御部100bは、キャッシュ処理の実行時にデータサーバ201、202から取得したデータファイル毎のアクセス数を記憶部100aに格納しておき、今回取得したアクセス数から前回取得したアクセス数を減算することでRA(t)を計算できる。また、制御部100bは、上記の式(1)に基づいて計算したアクセス頻度AW(t)を記憶部100aに格納し、次回計算時に利用する。

【0065】

制御部100bは、アクセス頻度AW(t)が予め設定された基準値より大きなデータファイルを特定し、特定したデータファイルを記載した対象ファイルリストを作成する。このとき、制御部100bは、管理用データベースを参照し、キャッシュサーバ300の

10

20

30

40

50

記憶装置 301 に格納されていないデータファイルを対象ファイルリストに記載する。

【0066】

制御部 100b は、図 7 に示すように、記憶装置 211、212、213、221、222、223 のそれぞれについて対象ファイルリストを作成する。そして、制御部 100b は、対象ファイルリストを記憶部 100a に格納する。図 7 は、第 2 実施形態に係る対象ファイルリストの一例を示した図である。

【0067】

制御部 100b は、管理用データベースを参照し、対象ファイルリストに記載したデータファイルの合計サイズ（ファイルサイズの合計）を計算する。また、制御部 100b は、管理用データベースを参照し、計算した合計サイズとディスク識別子とを対応付けた暫定対象ディスクリスト（図 8 を参照）を作成する。図 8 は、第 2 実施形態に係る暫定対象ディスクリスト及び対象ディスクリストの一例を示した図である。

10

【0068】

制御部 100b は、暫定対象ディスクリストの記載情報を合計サイズが小さい順にソートし、選択した記載情報の合計サイズの合算値がキャッシュサーバ 300 の記憶装置 301 の空き容量を超えない範囲で、上位から順に記載情報を選択する。そして、制御部 100b は、暫定対象ディスクリストから選択した記載情報を有する対象ディスクリスト（図 8 を参照）を作成し、記憶部 100a に格納する。

【0069】

制御部 100b は、対象ディスクリストのディスク識別子に対応する対象ファイルリスト（図 7 を参照）を記憶部 100a から読み出す。そして、制御部 100b は、通信部 100c を用いて、読み出した対象ファイルリストに記載のデータファイルをキャッシュサーバ 300 の記憶装置 301 にコピーするようにデータサーバ 201、202 に指示する（キャッシュ指示）。このキャッシュ指示を受けたデータサーバ 201、202 は、図 3 に示したようにデータコピー、省電力モードへの遷移などの処理を実行する。

20

【0070】

上記のように、記憶装置毎にアクセス頻度の高いデータファイルの合計サイズを計算し、合計サイズが小さい順にキャッシュ対象とすることで、キャッシュ対象のデータファイルを全てキャッシュできる記憶装置の数を増やすことができる。その結果、省電力モードへ遷移できる記憶装置の数が増加し、ストレージシステム 50 の消費電力をより大きく低減することが可能になる。

30

【0071】

（データサーバ 201、202）

データサーバ 201 は、記憶部 201a、制御部 201b、及び通信部 201c を有する。記憶部 201a の機能は、上述した RAM 906 や記憶部 920 などを用いて実現できる。制御部 201b の機能は、上述した CPU 902 などを用いて実現できる。通信部 201c の機能は、上述した接続ポート 924 や通信部 926 などを用いて実現できる。

【0072】

記憶部 201a には、記憶装置 211、212、213 に格納されたデータファイルに関する情報が格納される。例えば、制御部 201b は、データファイル毎に読み出し及び書き込みの回数を監視し、その回数をアクセス数として記憶部 201a に記録する。このとき、制御部 201b は、データファイルが格納されている記憶装置とアクセス数とを対応付けて記憶部 201a に記録する。そして、制御部 201b は、メタサーバ 100 から受けた要求に応じて記憶部 201a に記録したアクセス数を応答する。

40

【0073】

記憶部 201a には、省電力モードに関する設定情報が格納される。この設定情報は、例えば、省電力モードが、ディスク回転数の低下・停止、ディスクヘッドの退避、又は電源オフのいずれであることを示すモード情報を含む。また、この設定情報は、省電力モード時のディスクの回転数などのパラメータ情報を含む。制御部 201b は、省電力モードへと遷移する際、記憶部 201a に格納された設定情報に基づいて記憶装置 211、212

50

、 2 1 3 を制御する。

【 0 0 7 4 】

また、上記の設定情報は、通常動作モードに関する情報も含む。制御部 2 0 1 b は、省電力モードから通常動作モードへと復帰する際、上記の設定情報に基づいて記憶装置 2 1 1、2 1 2、2 1 3 を制御する。例えば、制御部 2 0 1 b は、キャッシュサーバ 3 0 0 の記憶装置 3 0 1 に格納されていないデータファイルへのアクセスを受け、そのデータファイルを格納した記憶装置が省電力モードである場合、上記の設定情報に基づいて通常動作モードへと復帰させる。

【 0 0 7 5 】

制御部 2 0 1 b は、通信部 2 0 1 c を介して、キャッシュ対象のデータファイルを指定したキャッシュ指示をメタサーバ 1 0 0 から受信した場合、指定されたデータファイルをキャッシュサーバ 3 0 0 に送信する。制御部 2 0 1 b は、送信したデータファイルをキャッシュサーバ 3 0 0 が記憶装置 3 0 1 に格納し終えたことを確認し、指定されたデータファイルを格納していた記憶装置を省電力モードへと遷移させる。

10

【 0 0 7 6 】

データサーバ 2 0 2 は、記憶部 2 0 2 a、制御部 2 0 2 b、及び通信部 2 0 2 c を有する。記憶部 2 0 2 a の機能は、上述した R A M 9 0 6 や記憶部 9 2 0 などを用いて実現できる。制御部 2 0 2 b の機能は、上述した C P U 9 0 2 などを用いて実現できる。通信部 2 0 2 c の機能は、上述した接続ポート 9 2 4 や通信部 9 2 6 などを用いて実現できる。データサーバ 2 0 2 の動作はデータサーバ 2 0 1 と同じであるため説明を省略する。

20

【 0 0 7 7 】

(キャッシュサーバ 3 0 0)

キャッシュサーバ 3 0 0 は、記憶部 3 0 0 a、制御部 3 0 0 b、及び通信部 3 0 0 c を有する。記憶部 3 0 0 a の機能は、上述した R A M 9 0 6 や記憶部 9 2 0 などを用いて実現できる。制御部 3 0 0 b の機能は、上述した C P U 9 0 2 などを用いて実現できる。通信部 3 0 0 c の機能は、上述した接続ポート 9 2 4 や通信部 9 2 6 などを用いて実現できる。

【 0 0 7 8 】

記憶部 3 0 0 a は、記憶装置 3 0 1 の空き容量など、記憶装置 3 0 1 に関する情報を記憶する。制御部 3 0 0 b は、メタサーバ 1 0 0 から記憶装置 3 0 1 の空き容量に関する問い合わせを受けた場合に記憶部 3 0 0 a を参照して空き容量を応答する。

30

【 0 0 7 9 】

制御部 3 0 0 b は、通信部 3 0 0 c を介して、データサーバ 2 0 1、2 0 2 からキャッシュ対象のデータファイルを受信した場合、受信したデータファイルを記憶装置 3 0 1 に格納する。そして、制御部 3 0 0 b は、データファイルの格納が完了した場合に完了通知をデータファイルの送信元へと送信する。また、制御部 3 0 0 b は、クライアント装置 4 0 0 からデータファイルの読み出し要求を受けた場合、そのデータファイルを記憶装置 3 0 1 から読み出してクライアント装置 4 0 0 に送信する。

【 0 0 8 0 】

以上、メタサーバ 1 0 0、データサーバ 2 0 1、2 0 2、及びキャッシュサーバ 3 0 0 の機能について説明した。上記のように、記憶装置毎にアクセス頻度の高いデータファイルの合計サイズが計算され、合計サイズが小さい順にキャッシュされることで、キャッシュ対象のデータファイルを全てキャッシュできる記憶装置の数を増やすことができる。その結果、省電力モードへ遷移できる記憶装置の数が増加し、ストレージシステム 5 0 の消費電力をより大きく低減することが可能になる。

40

【 0 0 8 1 】

[2 - 4 . 処理フロー]

次に、キャッシュ処理に際してメタサーバ 1 0 0、データサーバ 2 0 1、2 0 2 が実行する処理の流れについて説明する。なお、データサーバ 2 0 1、2 0 2 が実行する処理は実質的に同じであるため、ここではデータサーバ 2 0 1 の処理について説明する。

50

【 0 0 8 2 】

(メタサーバの処理)

図 9 及び図 10 を参照しながら、メタサーバ 100 が実行する処理の流れについて説明する。図 9 は、第 2 実施形態に係るメタサーバが実行する処理の流れを示した第 1 のフロー図である。図 10 は、第 2 実施形態に係るメタサーバが実行する処理の流れを示した第 2 のフロー図である。

【 0 0 8 3 】

図 9 及び図 10 に示した処理は、例えば、キャッシュディスクとして利用可能な記憶装置が追加された場合、或いは、前回キャッシュ処理を実行後、一定の期間が経過したタイミングで実行される。なお、ストレージシステム 50 の管理者などが任意のタイミングで該処理を実行してもよいし、予め設定された不定期のタイミングで実行されるように設定することもできる。

10

【 0 0 8 4 】

(S 1 0 1) キャッシュ処理を開始すると、制御部 100 b は、データサーバ 201、202 が管理する記憶装置の中に未確認の記憶装置 (S 1 0 2 以降の処理が未実行の記憶装置) があるかを判定する。未確認の記憶装置がある場合、制御部 100 b により未確認の記憶装置が 1 つ選択され、処理は S 1 0 2 へと進む。一方、未確認の記憶装置がない場合、処理は図 10 の S 1 0 8 へと進む。

【 0 0 8 5 】

(S 1 0 2) 制御部 100 b は、S 1 0 1 で選択した未確認の記憶装置に未確認のファイル (S 1 0 3 以降の処理が未実行のデータファイル) があるかを判定する。未確認のファイルがある場合、制御部 100 b により未確認のファイルが 1 つ選択され、処理は S 1 0 3 へと進む。一方、未確認のファイルがない場合、処理は S 1 0 6 へと進む。

20

【 0 0 8 6 】

(S 1 0 3) 制御部 100 b は、S 1 0 2 で選択した未確認のファイルについてアクセス頻度を計算する。アクセス頻度は、上記の式 (1) に基づいて計算される。

例えば、制御部 100 b は、データサーバから未確認のファイルのアクセス数を取得し、今回取得したアクセス数から前回キャッシュ処理時に取得したアクセス数を減算することで、上記の式 (1) に含まれる $RA(t)$ を計算する。また、制御部 100 b は、前回キャッシュ処理時に計算したアクセス頻度 $AW(t-1)$ を利用し、未確認のファイルに関する今回のアクセス頻度 $AW(t)$ を計算する。

30

【 0 0 8 7 】

(S 1 0 4) 制御部 100 b は、S 1 0 3 で計算したアクセス頻度が予め設定された基準値より大きいかを判定する。アクセス頻度が基準値より大きい場合、処理は S 1 0 5 へと進む。一方、アクセス頻度が基準値より大きくない場合、処理は S 1 0 2 へと進む。なお、基準値は、ストレージシステム 50 の管理者などが標準的な運用実態に即した経験則に基づいて予め設定してもよいし、シミュレーションや実験により得られた好適な値を設定してもよい。

【 0 0 8 8 】

(S 1 0 5) 制御部 100 b は、S 1 0 1 で未確認の記憶装置として選択された記憶装置に対応する対象ファイルリスト (図 7 を参照) に、S 1 0 2 で未確認のファイルとして選択されたデータファイルを追加する。S 1 0 5 の処理が完了すると、処理は S 1 0 2 へと進む。

40

【 0 0 8 9 】

(S 1 0 6) 制御部 100 b は、記憶部 100 a に格納された管理用データベースを参照し、記憶装置毎の対象ファイルリストに記載されたデータファイルについて、ファイルサイズの合計値である合計サイズを計算する。

【 0 0 9 0 】

(S 1 0 7) 制御部 100 b は、S 1 0 6 で計算した合計サイズに対応する記憶装置のディスク識別子と共に暫定対象ディスクリスト (図 8 を参照) に追加する。S 1 0 7 の処

50

理が完了すると、処理は S 1 0 1 へと進む。

【 0 0 9 1 】

(S 1 0 8) 制御部 1 0 0 b は、暫定対象ディスクリストに基づいて対象ディスクリストを生成する。例えば、制御部 1 0 0 b は、暫定対象ディスクリストの記載情報を合計サイズが小さい順にソートし、選択した記載情報の合計サイズの合算値がキャッシュサーバ 3 0 0 の記憶装置 3 0 1 の空き容量を超えない範囲で、上位から順に記載情報を選択する。そして、制御部 1 0 0 b は、暫定対象ディスクリストから選択した記載情報を有する対象ディスクリスト (図 8 を参照) を生成する。

【 0 0 9 2 】

(S 1 0 9) 制御部 1 0 0 b は、対象ディスクリストに未処理の記憶装置 (S 1 1 0 以降の処理が未実行の記憶装置) があるかを判定する。未処理の記憶装置がある場合、制御部 1 0 0 b により未処理の記憶装置が 1 つ選択され、処理は S 1 1 0 へと進む。未処理の記憶装置がない場合、処理は S 1 1 2 へと進む。

10

【 0 0 9 3 】

(S 1 1 0) 制御部 1 0 0 b は、S 1 0 1 で選択した未処理の記憶装置に対応する対象ファイルリストに未処理のファイル (S 1 1 1 の処理が未実行のデータファイル) があるかを判定する。未処理のファイルがある場合、制御部 1 0 0 b により未処理のファイルが 1 つ選択され、処理は S 1 1 1 へと進む。一方、未処理のファイルがない場合、処理は S 1 0 8 へと進む。

【 0 0 9 4 】

(S 1 1 1) 制御部 1 0 0 b は、S 1 1 0 で選択した未処理のファイルのキャッシュ処理をデータサーバに指示する。このとき、制御部 1 0 0 b は、管理用データベースを参照し、未処理のファイルを格納した記憶装置を管理するデータサーバを特定し、特定したデータサーバに対して未処理のファイルを指定してキャッシュ処理の実行を指示する。

20

【 0 0 9 5 】

この指示 (キャッシュ指示) により、データサーバの記憶装置から未処理のファイルが読み出され、キャッシュサーバ 3 0 0 の記憶装置 3 0 1 に書き込まれる (図 3 を参照) 。S 1 1 1 の処理が完了すると、処理は S 1 1 0 へと進む。

【 0 0 9 6 】

(S 1 1 2) 制御部 1 0 0 b は、全データサーバ (データサーバ 2 0 1 、 2 0 2) へ完了通知を送信する。S 1 1 2 の処理が完了すると、図 9 及び図 1 0 に示した一連の処理は終了する。

30

【 0 0 9 7 】

以上、メタサーバ 1 0 0 が実行する処理の流れについて説明した。

(データサーバの処理)

次に、図 1 1 を参照しながら、データサーバ 2 0 1 が実行する処理の流れについて説明する。図 1 1 は、第 2 実施形態に係るデータサーバが実行する処理の流れを示したフロー図である。

【 0 0 9 8 】

(S 1 2 1) 制御部 2 0 1 b は、メタサーバ 1 0 0 から完了通知を受信する。

40

(S 1 2 2) 制御部 2 0 1 b は、管理下に未調査の記憶装置 (S 1 2 3 以降の処理が未実行の記憶装置) があるかを判定する。未調査の記憶装置がある場合、制御部 2 0 1 b により未調査の記憶装置が 1 つ選択され、処理は S 1 2 3 へと進む。一方、未調査の記憶装置がない場合、図 1 1 に示した一連の処理は終了する。

【 0 0 9 9 】

(S 1 2 3) 制御部 2 0 1 b は、S 1 2 2 で選択された未調査の記憶装置に、アクセス頻度が基準値より大きい対象ファイル (メタサーバ 1 0 0 から指定されたデータファイル) が 1 つ以上存在するかを判定する。対象ファイルが 1 つ以上存在する場合、処理は S 1 2 4 へと進む。対象ファイルが存在しない場合、処理は S 1 2 5 へと進む。

【 0 1 0 0 】

50

(S124) 制御部201bは、全ての対象ファイルがキャッシュされているか(キャッシュサーバ300の記憶装置301にコピー済みか)を判定する。全ての対象ファイルがキャッシュされている場合、処理はS125へと進む。キャッシュされていない対象ファイルがある場合、処理はS122へと進む。

【0101】

(S125) 制御部201bは、S122で選択された未調査の記憶装置を省電力モードへと遷移させる。S125の処理が完了すると、処理はS122へと進む。

以上、データサーバ201が実行する処理の流れについて説明した。

【0102】

(省電力化について)

10

ここで、図12を参照しながら、図9から図11に示したキャッシュ方法を適用した場合の効果について説明する。図12は、第2実施形態に係るキャッシュ方法の効果の一例を示した図である。

【0103】

図12中のA、B、...、Fはデータサーバが管理する記憶装置のディスク識別子を表し、ディスク識別子の横に(...)と表記された数字は記憶装置に格納されているアクセス頻度の高いデータファイルの合計サイズを表す。また、キャッシュディスクに付された(...)の数字は空き容量を表す。キャッシュ方法#1は、ディスク識別子の順にデータファイルをキャッシュする方法を示す。一方、キャッシュ方法#2は、第2実施形態に係る技術を適用したキャッシュ方法を示す。

20

【0104】

キャッシュ方法#1の場合、ディスク識別子の順にキャッシュディスクへデータファイルがコピーされる。ディスク識別子がAの記憶装置におけるキャッシュ対象の合計サイズは10、ディスク識別子がBの記憶装置におけるキャッシュ対象の合計サイズは20である。キャッシュディスクの空き容量が30であるため、ディスク識別子がA、Bの2つの記憶装置についてキャッシュ対象のデータファイルをキャッシュディスクにキャッシュすることができる。

【0105】

キャッシュ方法#2の場合、キャッシュ対象の合計サイズが小さい順にキャッシュディスクへデータファイルがコピーされる。ディスク識別子がAの記憶装置におけるキャッシュ対象の合計サイズは10、ディスク識別子がDの記憶装置におけるキャッシュ対象の合計サイズは10、ディスク識別子がE及びFの記憶装置におけるキャッシュ対象の合計サイズは5である。つまり、ディスク識別子がB、Cの記憶装置に比べ、ディスク識別子がA、D、E、Fの記憶装置の方が、キャッシュ対象の合計サイズが小さい。そのため、ディスク識別子がA、D、E、Fの4つの記憶装置についてキャッシュ対象のデータファイルがキャッシュディスクにキャッシュされる。

30

【0106】

キャッシュ方法#1の場合には2つの記憶装置(A、B)を省電力モードに遷移できる。一方、キャッシュ方法#2の場合には4つの記憶装置(A、D、E、F)を省電力モードに遷移できる。つまり、キャッシュ方法#2の方が、より多くの記憶装置を省電力モードへ遷移できる分だけ、ストレージシステム50の電力消費を低減することが可能になる。なお、アクセス頻度の高いデータファイルはキャッシュディスクに格納されているため、そのデータファイルへのアクセスがあっても素早く応答することができる。

40

【0107】

このように、第2実施形態に係る技術を適用することにより、アクセス速度の低下を抑制しつつ、より消費電力を低減することが可能になる。

以上、第2実施形態について説明した。

【符号の説明】

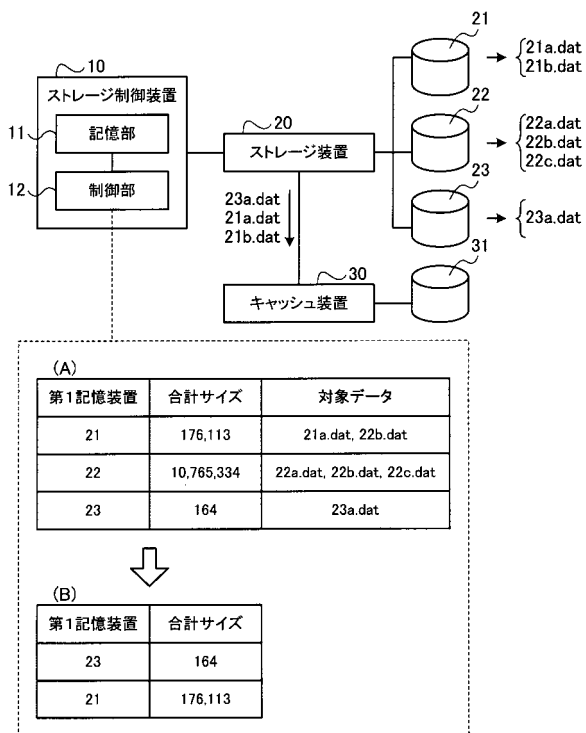
【0108】

10 ストレージ制御装置

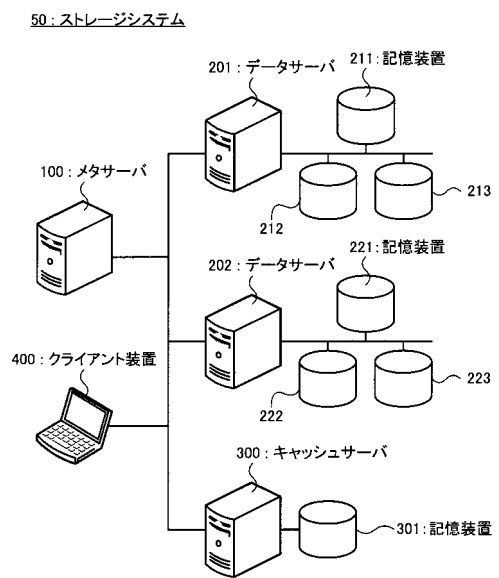
50

- 1 1 記憶部
- 1 2 制御部
- 2 0 ストレージ装置
- 2 1、2 2、2 3 第 1 記憶装置
- 3 0 キャッシュ装置
- 3 1 第 2 記憶装置

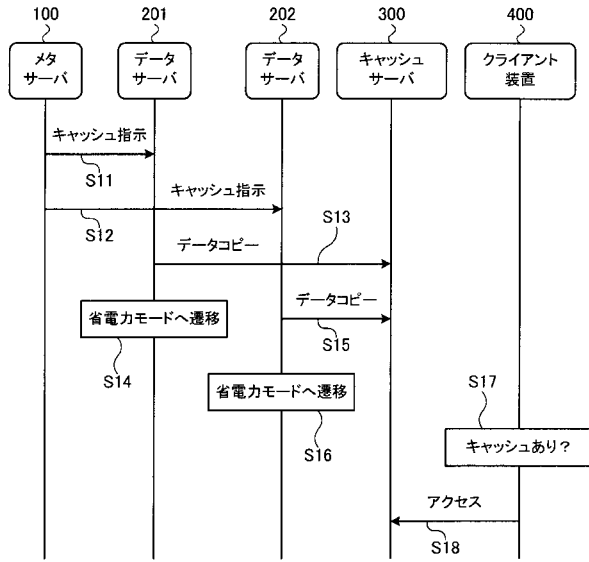
【 図 1 】



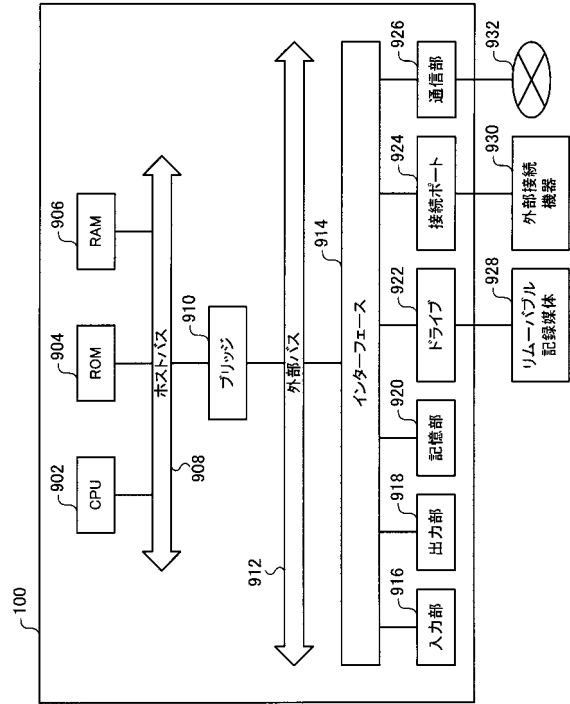
【 図 2 】



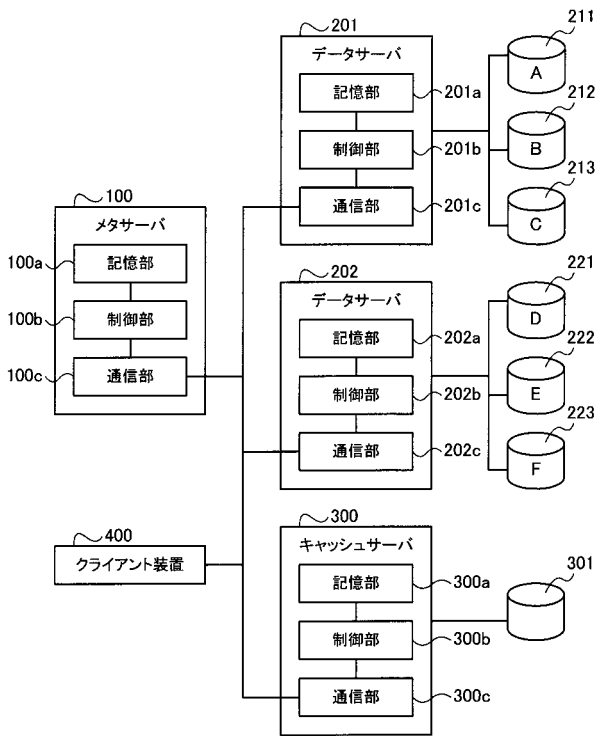
【図3】



【図4】



【図5】

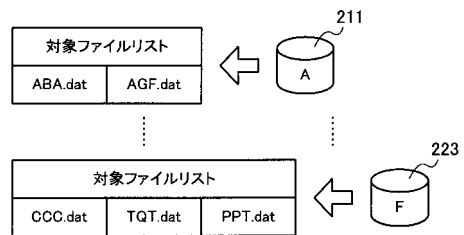


【図6】

(管理用データベース)

ファイル名	データサーバ IPアドレス	ディスク 識別子	ファイル サイズ [kB]	キャッシュ
AAA.dat	192.168.1.10	A	690	あり
BBB.dat	192.168.1.10	C	1,020,776	
CCC.dat	192.168.1.3	F	90,176	あり
DDD.dat	192.168.1.3	D	120	あり
FFF.dat	192.168.1.10	C	56,000	
GGG.dat	192.168.1.10	B	620,778	
⋮	⋮	⋮	⋮	⋮

【図7】



【 図 8 】

(暫定対象ディスクリスト)

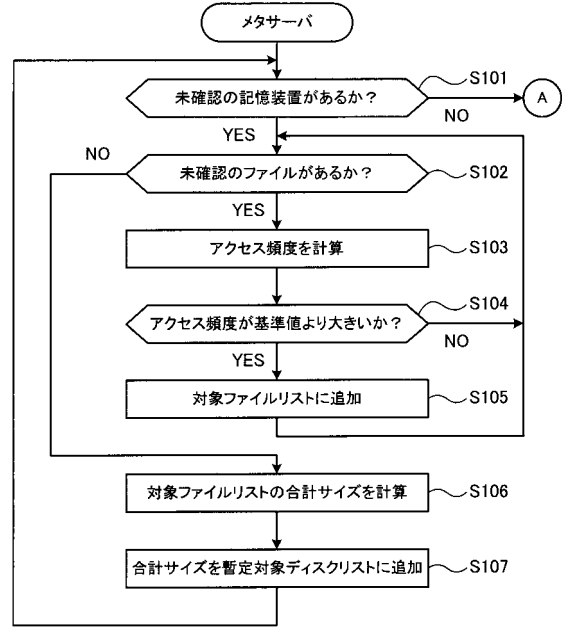
ディスク識別子	対象ファイルリスト合計サイズ
A	176,113
B	10,765,334
C	164
D	2,789,110
E	35,789
F	662,115

↓

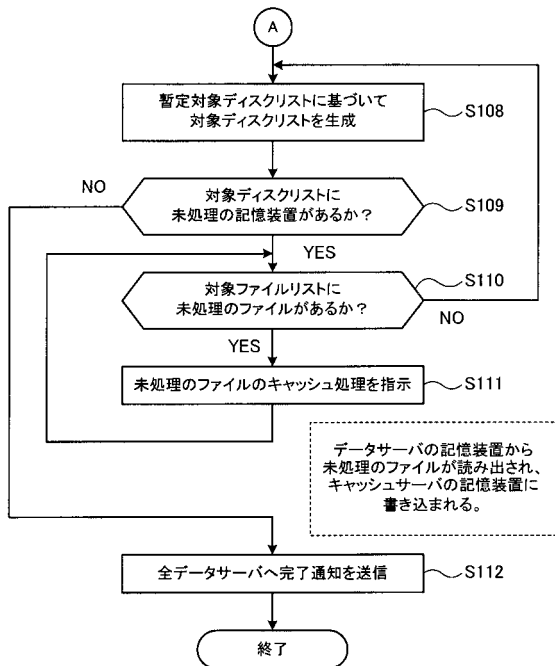
(対象ディスクリスト)

ディスク識別子	対象ファイルリスト合計サイズ
C	164
E	35,789
A	176,113

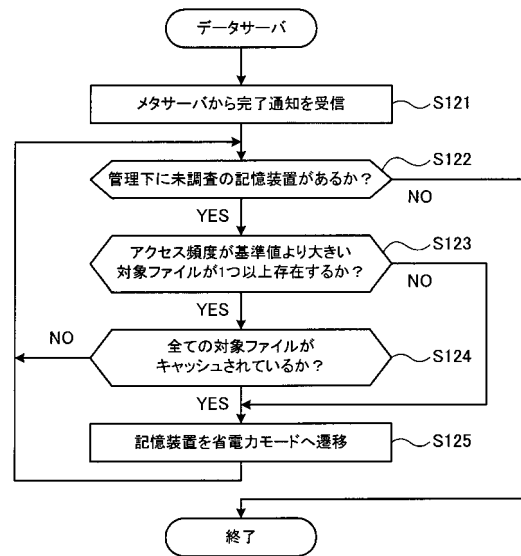
【 図 9 】



【 図 10 】



【 図 11 】



【 図 1 2 】

