

(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2007-226596

(P2007-226596A)

(43) 公開日 平成19年9月6日(2007.9.6)

(51) Int. Cl.		F I		テーマコード (参考)
G06F 3/06	(2006.01)	G06F 3/06	301T	5B065
G06F 12/00	(2006.01)	G06F 12/00	501B	5B082

審査請求 未請求 請求項の数 11 O L (全 30 頁)

(21) 出願番号 特願2006-47799 (P2006-47799)
 (22) 出願日 平成18年2月24日 (2006.2.24)

(71) 出願人 000005108
 株式会社日立製作所
 東京都千代田区丸の内一丁目6番6号
 (74) 代理人 110000279
 特許業務法人ウィルフォート国際特許事務所
 (72) 発明者 室谷 暁
 神奈川県小田原市中里322番2号 株式会社日立製作所RAIDシステム事業部内
 (72) 発明者 檜垣 誠一
 神奈川県小田原市中里322番2号 株式会社日立製作所RAIDシステム事業部内
 Fターム(参考) 5B065 BA01 CE14
 5B082 CA13

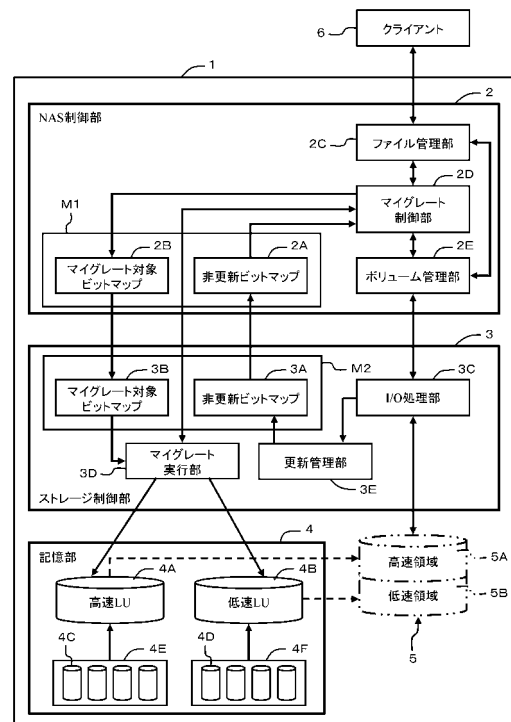
(54) 【発明の名称】 記憶制御装置及び記憶制御装置を用いたデータマイグレーション方法

(57) 【要約】

【課題】 NAS制御部の負荷を増大させずに、データマイグレーションを実行する。

【解決手段】 FCディスク4Cに基づくボリューム4AとATAディスク4Dに基づくボリューム4Bを結合させて、一つの仮想ボリューム5が生成されている。更新管理部3Eは、高速領域5Aの更新を監視し、更新ビットマップを生成する。更新管理部3Eは、各更新ビットマップのORを演算して、所定期間更新されていないセグメントを検出するための非更新ビットマップを生成する。マイグレート制御部2Dは、非更新ビットマップに基づいて、マイグレート対象となるセグメントを抽出したマイグレート対象ビットマップ2Bを生成する。マイグレート実行部3Dは、マイグレート対象ビットマップ3Bに基づいて、マイグレート対象のセグメントを1つずつ、または複数個まとめて、データマイグレーションを実行する。

【選択図】 図1



【特許請求の範囲】

【請求項 1】

ファイルコントローラ及びブロックコントローラを備えた記憶制御装置であって、各ファイルに関する情報をそれぞれ管理し、上位装置からのファイルアクセス要求をキャッシュメモリを利用して処理するファイル管理部と、

第 1 物理的記憶デバイス上に設けられる第 1 論理的記憶デバイスと前記第 1 物理的記憶デバイスとは異なる性能を有する第 2 物理的記憶デバイス上に設けられる第 2 論理的記憶デバイスとを一つの仮想ボリュームとして結合させ、この仮想ボリュームを前記ファイル管理部に提供するボリューム管理部と、

前記第 1 論理的記憶デバイス及び前記第 2 論理的記憶デバイスの更新状態を、所定の管理単位でそれぞれ管理する更新管理部と、 10

前記管理単位での更新状態及び前記ファイル管理部により管理される前記各ファイルの情報に基づいて、マイグレート対象となるデータを特定し、この特定されたデータを前記第 1 論理的記憶デバイスと前記第 2 論理的記憶デバイスとの間で移動させるための指示を発行するマイグレート制御部と、

前記マイグレート制御部からの指示に基づき、前記第 1 論理的記憶デバイスと前記第 2 論理的記憶デバイスとの間で、前記管理単位で、データを移動させるマイグレート実行部と、を備えた記憶制御装置。

【請求項 2】

前記管理単位には、前記キャッシュメモリにおけるデータ管理単位であるセグメント単位が用いられ、 20

前記更新管理部は、前記セグメント単位で、前記各論理的記憶デバイスの更新状態を、更新ビットマップによってそれぞれ管理する請求項 1 に記載の記憶制御装置。

【請求項 3】

前記更新管理部は、所定期間毎に、前記各論理的記憶デバイス毎に前記更新ビットマップをそれぞれ作成し、かつ、所定期間内に作成された前記複数の更新ビットマップの論理和に基づいて、前記所定期間内に更新されなかった非更新セグメントを検出するための非更新ビットマップを生成し、

前記マイグレート制御部は、前記非更新ビットマップ及び前記ファイル管理部により管理される前記各ファイルの情報に基づいて、前記非更新セグメントに記憶されているデータをマイグレート対象のデータとして特定するためのマイグレート対象ビットマップを作成し、このマイグレート対象ビットマップに基づいて、前記マイグレート対象のデータを前記第 1 論理的記憶デバイスと前記第 2 論理的記憶デバイスとの間で移動させるための指示を前記マイグレート実行部に発行する請求項 2 に記載の記憶制御装置。 30

【請求項 4】

前記マイグレート制御部は、前記非更新ビットマップに含まれる非更新セグメントのうち、予め設定されたマイグレーションポリシーに合致する非更新セグメントのみを、前記マイグレート対象ビットマップに登録する請求項 3 に記載の記憶制御装置。

【請求項 5】

前記マイグレート制御部は、前記マイグレート対象ビットマップに登録された非更新セグメント毎に、前記指示を発行する請求項 3 に記載の記憶制御装置。 40

【請求項 6】

前記マイグレート制御部は、前記マイグレート対象ビットマップを複数のセグメント範囲に分割し、これらセグメント範囲毎に、当該セグメント範囲に含まれている前記非更新セグメントについて一括して、前記指示を発行する請求項 3 に記載の記憶制御装置。

【請求項 7】

前記ファイルコントローラは、前記ファイル管理部と、前記ボリューム管理部及び前記マイグレート制御部を備えており、

前記ブロックコントローラは、前記更新管理部及び前記マイグレート実行部を備えており、

前記非更新ビットマップ及び前記マイグレート対象ビットマップは、前記ファイルコントローラと前記ブロックコントローラとにより共有されている請求項 3 に記載の記憶制御装置。

【請求項 8】

前記ファイルコントローラの前記キャッシュメモリが接続される内部バスと前記ブロックコントローラの他のメモリが接続される内部バスとは、ブリッジ回路を介して結合されており、

前記非更新ビットマップ及び前記マイグレート対象ビットマップは、前記内部バスを介したメモリ間コピーを用いることにより、前記ファイルコントローラと前記ブロックコントローラとに共有されている請求項 7 に記載の記憶制御装置。

10

【請求項 9】

前記更新管理部は、前記ファイルコントローラ内に設けられている請求項 3 に記載の記憶制御装置。

【請求項 10】

ファイルコントローラ及びブロックコントローラを備えた記憶制御装置を用いて、データをマイグレートさせるためのデータマイグレーション方法であって、

前記ファイルコントローラによって、第 1 物理的記憶デバイス上に設けられる第 1 論理的記憶デバイスと第 2 物理的記憶デバイス上に設けられる第 2 論理的記憶デバイスとを一つの仮想ボリュームとして結合させ、この仮想ボリュームを、上位装置からのファイルアクセス要求をキャッシュメモリを利用して処理するためのファイル管理部に提供するステップと、

20

前記ブロックコントローラによって、前記各論理的記憶デバイスの更新状態を、所定期間毎に更新ビットマップをそれぞれ作成することにより、前記キャッシュメモリにおけるデータ管理単位であるセグメント単位で管理するステップと、

前記ブロックコントローラによって、所定期間内に作成された前記複数の更新ビットマップの論理和に基づいて、前記所定期間内に更新されなかった非更新セグメントを検出するための非更新ビットマップを生成して記憶させるステップと、

前記非更新ビットマップを前記ファイルコントローラと前記ブロックコントローラとが共有するステップと、

前記ファイルコントローラによって、前記非更新ビットマップ及び前記ファイル管理部により管理されている各ファイルの情報に基づいて、前記非更新セグメントに記憶されているファイルをマイグレート対象のデータとして特定するためのマイグレート対象ビットマップを作成して記憶させるステップと、

30

前記マイグレート対象ビットマップを前記ファイルコントローラと前記ブロックコントローラとが共有するステップと、

前記ファイルコントローラによって、前記マイグレート対象ビットマップに登録された非更新セグメントに記憶されているファイルへの更新を禁止させるステップと、

前記ファイルコントローラによって、前記マイグレート対象ビットマップに登録された非更新セグメントに記憶されているデータを前記第 1 論理的記憶デバイスから前記第 2 論理的記憶デバイスに移動させるための指示を、前記ブロックコントローラに発行させるステップと、

40

前記ブロックコントローラによって、前記発行された指示及び前記マイグレート対象ビットマップに基づいて、前記第 1 論理的記憶デバイスに記憶されたデータを、一つまたは複数のセグメント単位で、前記第 2 論理的記憶デバイスに移動させるステップと、

前記ファイルコントローラによって、前記データの移動が完了した場合には、更新が禁止されたファイルの更新禁止を解除するステップと、
を実行するデータマイグレーション方法。

【請求項 11】

ファイルアクセスを担当するファイルコントローラと、ブロックアクセスを担当するブロックコントローラと、このブロックコントローラによりそれぞれ使用される高速物理的

50

記憶デバイス及び低速物理的記憶デバイスとを備えた記憶制御装置であって、

前記ファイルコントローラの有するキャッシュメモリと前記ブロックコントローラの有する他のメモリとは、前記ファイルコントローラの内部バスと前記ブロックコントローラの内部バスとをブリッジ回路を介して結合することにより接続されており、

前記ファイルコントローラは、ファイルシステムと、ボリューム管理部及びマイグレート制御部を備え、

前記ブロックコントローラは、更新管理部及びマイグレート実行部を備え、

前記ファイルシステムは、前記上位装置からのファイルアクセス要求を前記キャッシュメモリを利用して処理し、

前記ボリューム管理部は、前記高速物理的記憶デバイス上に設けられる高速論理的記憶デバイスが前半部分に、前記低速物理的記憶デバイス上に設けられる低速論理的記憶デバイスが後半部分にそれぞれ位置するようにして、前記高速論理的記憶デバイスと前記低速論理的記憶デバイスとから一つの仮想ボリュームを生成し、この仮想ボリュームを前記ファイル管理部に提供し、

前記更新管理部は、前記高速論理的記憶デバイス及び前記低速論理的記憶デバイスの更新状態を前記キャッシュメモリのデータ管理単位であるセグメント毎にそれぞれ管理し、所定期間毎に、前記各論理的記憶デバイスのそれぞれについて更新ビットマップを生成し、かつ、所定期間内に生成された前記各更新ビットマップの論理和に基づいて、前記所定期間内に更新されなかった非更新セグメントを検出するための非更新ビットマップを生成して前記他のメモリに記憶させ、

前記他のメモリに記憶された前記非更新ビットマップは、メモリ間コピーによって、前記他のメモリから前記キャッシュメモリにコピーされ、

前記マイグレート制御部は、前記キャッシュメモリに記憶された前記非更新ビットマップに含まれる前記非更新セグメントに記憶されているファイルの属性を前記ファイルシステムに問い合わせることにより、マイグレートさせるべきマイグレート対象セグメントを特定するためのマイグレート対象ビットマップを生成して、このマイグレート対象ビットマップを前記キャッシュメモリに記憶させ、

前記キャッシュメモリに記憶された前記マイグレート対象ビットマップは、前記メモリ間コピーによって、前記キャッシュメモリから前記他のメモリにコピーされ、

前記マイグレート制御部は、前記マイグレート対象セグメントに登録された前記非更新セグメントに記憶されているデータを、前記高速論理的記憶デバイスから前記低速記憶デバイスに移動させるための指示を前記マイグレート実行部に発行し、

前記マイグレート実行部は、前記発行された指示及び前記マイグレート対象ビットマップに基づいて、前記高速論理的記憶デバイスに記憶されたデータを、一つまたは複数のセグメント単位で、前記低速論理的記憶デバイスに移動させ、

前記ファイルシステムは、前記マイグレート対象データの移動が完了するまでの間、前記上位装置から当該マイグレート対象データへの更新を禁止する、
記憶制御装置。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、記憶制御装置及び記憶制御装置を用いてデータをマイグレートさせる方法に関する。

【背景技術】

【0002】

企業等では、管理すべきデータが日々増大するが、一般的に、生成後一定時間の経過したデータは、その使用頻度が低下する。データの価値は、時間の経過につれて減少するため、利用価値の少なくなったデータは、高速な記憶装置から低速な記憶装置に移されるべきである。高速な記憶装置は、一般的に高価であり、使用可能な容量には限りがあるためである。

10

20

30

40

50

【0003】

そこで、ディスク装置上のファイルを、所定のタイミングで、磁気テープ装置に移動させるようにしたデータマイグレーション処理方法も提案されている（特許文献1）。この文献に記載の技術では、マイグレート候補を検索してリストを作成する処理と、作成されたリストに基づいてディスク装置内のファイルを一定量だけ磁気テープ装置に移動させる処理とを分離し、それぞれの処理を非同期で実行させている。

【特許文献1】特開2003-15917号公報

【発明の開示】

【発明が解決しようとする課題】

【0004】

前記文献に記載の技術では、ファイル情報を収集する処理とマイグレートを実行する処理とを分離し、マイグレートに要する時間の短縮を図っている。しかし、ボリュームサイズが大きくなればなるほど、ファイルツリーの全検索等を行うために時間がかかり、マイクロプロセッサの負荷も増大する。

【0005】

特に、例えば、高速な記憶領域と低速な記憶領域とを結合させて一つの仮想ボリュームを構築する場合、この仮想ボリュームのサイズは、通常のボリュームよりも大きくなる。そして、例えば、NAS（Network Attached Storage）と呼ばれるファイルサーバでは、この大サイズの仮想ボリューム上にファイルシステムを構築可能なため、マイグレート処理に過大な負荷がかかるという問題がある。

【0006】

本発明は、上記問題点に鑑みてなされたもので、その目的は、仮想ボリュームを構成する各論理的記憶デバイスの更新状態を管理することにより、マイグレート対象のデータを比較的簡単に特定し、効率的にデータをマイグレートできるようにした記憶制御装置及び記憶制御装置を用いたデータマイグレーション方法を提供することにある。本発明の他の目的は、仮想ボリュームを構成する各論理的記憶デバイスの更新状態を、キャッシュメモリ上の管理単位で管理することにより、ファイルコントローラに過大な負荷を与えることなく、マイグレート対象のファイルを間接的に特定し、この特定されたファイルのデータを移動させることができるようにした記憶制御装置及び記憶制御装置を用いたデータマイグレーション方法を提供することにある。本発明の更なる目的は、ブロックコントローラ内で非更新ビットマップを生成し、ファイルコントローラ内でマイグレート対象ビットマップを生成し、これら各ビットマップをファイルコントローラ及びブロックコントローラがメモリ間コピーを用いて共有することにより、特別なコマンドを使用せずに、比較的簡単にデータのマイグレートを行うことができる記憶制御装置及び記憶制御装置を用いたデータマイグレーション方法を提供することにある。本発明のさらなる目的は、後述する実施形態の記載から明らかになるであろう。

【課題を解決するための手段】

【0007】

上記課題を解決すべく、本発明に従う記憶制御装置は、ファイルコントローラ及びブロックコントローラを備えた記憶制御装置であって、各ファイルに関する情報をそれぞれ管理し、上位装置からのファイルアクセス要求をキャッシュメモリを利用して処理するファイル管理部と、第1物理的記憶デバイス上に設けられる第1論理的記憶デバイスと第1物理的記憶デバイスとは異なる性能を有する第2物理的記憶デバイス上に設けられる第2論理的記憶デバイスとを一つの仮想ボリュームとして結合させ、この仮想ボリュームをファイル管理部に提供するボリューム管理部と、第1論理的記憶デバイス及び第2論理的記憶デバイスの更新状態を、所定の管理単位でそれぞれ管理する更新管理部と、管理単位での更新状態及びファイル管理部により管理される各ファイルの情報に基づいて、マイグレート対象となるデータを特定し、この特定されたデータを第1論理的記憶デバイスと第2論理的記憶デバイスとの間で移動させるための指示を発行するマイグレート制御部と、マイグレート制御部からの指示に基づき、第1論理的記憶デバイスと第2論理的記憶デバイス

10

20

30

40

50

との間で、管理単位で、データを移動させるマイグレート実行部と、を備える。

【0008】

本発明の一態様では、管理単位には、キャッシュメモリにおけるデータ管理単位であるセグメント単位が用いられ、更新管理部は、セグメント単位で、各論理的記憶デバイスの更新状態を、更新ビットマップによってそれぞれ管理する。

【0009】

本発明の一態様では、更新管理部は、所定期間毎に、各論理的記憶デバイス毎に更新ビットマップをそれぞれ作成し、かつ、所定期間内に作成された複数の更新ビットマップの論理和に基づいて、所定期間内に更新されなかった非更新セグメントを検出するための非更新ビットマップを生成し、マイグレート制御部は、非更新ビットマップ及びファイル管理部により管理される各ファイルの情報に基づいて、非更新セグメントに記憶されているデータをマイグレート対象のデータとして特定するためのマイグレート対象ビットマップを作成し、このマイグレート対象ビットマップに基づいて、マイグレート対象のデータを第1論理的記憶デバイスと第2論理的記憶デバイスとの間で移動させるための指示をマイグレート実行部に発行する。

10

【0010】

本発明の一態様では、マイグレート制御部は、非更新ビットマップに含まれる非更新セグメントのうち、予め設定されたマイグレーションポリシーに合致する非更新セグメントのみを、マイグレート対象ビットマップに登録する。

【0011】

本発明の一態様では、マイグレート制御部は、マイグレート対象ビットマップに登録された非更新セグメント毎に、前記指示を発行する。

20

【0012】

本発明の一態様では、マイグレート制御部は、マイグレート対象ビットマップを複数のセグメント範囲に分割し、これらセグメント範囲毎に、当該セグメント範囲に含まれている非更新セグメントについて一括して、前記指示を発行する。

【0013】

本発明の一態様では、ファイルコントローラは、ファイル管理部と、ボリューム管理部及びマイグレート制御部を備えており、ブロックコントローラは、更新管理部及びマイグレート実行部を備えており、非更新ビットマップ及びマイグレート対象ビットマップは、ファイルコントローラとブロックコントローラとにより共有されている。

30

【0014】

本発明の一態様では、ファイルコントローラのキャッシュメモリが接続される内部バスとブロックコントローラの他のメモリが接続される内部バスとは、ブリッジ回路を介して結合されており、非更新ビットマップ及びマイグレート対象ビットマップは、内部バスを介したメモリ間コピーを用いることにより、ファイルコントローラとブロックコントローラとに共有されている。

【0015】

本発明の一態様では、更新管理部は、ファイルコントローラ内に設けられている。

【0016】

本発明の他の観点に従う記憶制御装置を用いたデータマイグレーション方法は、ファイルコントローラ及びブロックコントローラを備えた記憶制御装置を用いて、データをマイグレートさせるためのデータマイグレーション方法であって、ファイルコントローラによって、第1物理的記憶デバイス上に設けられる第1論理的記憶デバイスと第2物理的記憶デバイス上に設けられる第2論理的記憶デバイスとを一つの仮想ボリュームとして結合させ、この仮想ボリュームを、上位装置からのファイルアクセス要求をキャッシュメモリを利用して処理するためのファイル管理部に提供するステップと、ブロックコントローラによって、各論理的記憶デバイスの更新状態を、所定期間毎に更新ビットマップをそれぞれ作成することにより、キャッシュメモリにおけるデータ管理単位であるセグメント単位で管理するステップと、ブロックコントローラによって、所定期間内に作成された複数の更

40

50

新ビットマップの論理和に基づいて、所定期間内に更新されなかった非更新セグメントを検出するための非更新ビットマップを生成して記憶させるステップと、非更新ビットマップをファイルコントローラとブロックコントローラとが共有するステップと、ファイルコントローラによって、非更新ビットマップ及びファイル管理部により管理されている各ファイルの情報に基づいて、非更新セグメントに記憶されているファイルをマイグレート対象のデータとして特定するためのマイグレート対象ビットマップを作成して記憶させるステップと、マイグレート対象ビットマップをファイルコントローラとブロックコントローラとが共有するステップと、ファイルコントローラによって、マイグレート対象ビットマップに登録された非更新セグメントに記憶されているファイルへの更新を禁止させるステップと、ファイルコントローラによって、マイグレート対象ビットマップに登録された非更新セグメントに記憶されているデータを第1論理的記憶デバイスから第2論理的記憶デバイスに移動させるための指示を、ブロックコントローラに発行させるステップと、ブロックコントローラによって、発行された指示及びマイグレート対象ビットマップに基づいて、第1論理的記憶デバイスに記憶されたデータを、一つまたは複数のセグメント単位で、第2論理的記憶デバイスに移動させるステップと、ファイルコントローラによって、データの移動が完了した場合には、更新が禁止されたファイルの更新禁止を解除するステップと、を実行する。

10

20

30

40

50

【0017】

本発明のさらに別の観点に従う記憶制御装置は、ファイルアクセスを担当するファイルコントローラと、ブロックアクセスを担当するブロックコントローラと、このブロックコントローラによりそれぞれ使用される高速物理的記憶デバイス及び低速物理的記憶デバイスとを備える。この記憶制御装置では、ファイルコントローラの有するキャッシュメモリとブロックコントローラの有する他のメモリとは、ファイルコントローラの内部バスとブロックコントローラの内部バスとをブリッジ回路を介して結合することにより接続されている。ファイルコントローラは、ファイルシステムと、ボリューム管理部及びマイグレート制御部を備え、ブロックコントローラは、更新管理部及びマイグレート実行部を備えている。

【0018】

そして、ファイルシステムは、上位装置からのファイルアクセス要求をキャッシュメモリを利用して処理し、ボリューム管理部は、高速物理的記憶デバイス上に設けられる高速論理的記憶デバイスが前半部分に、低速物理的記憶デバイス上に設けられる低速論理的記憶デバイスが後半部分にそれぞれ位置するようにして、高速論理的記憶デバイスと低速論理的記憶デバイスとから一つの仮想ボリュームを生成し、この仮想ボリュームをファイル管理部に提供する。

【0019】

更新管理部は、高速論理的記憶デバイス及び低速論理的記憶デバイスの更新状態をキャッシュメモリのデータ管理単位であるセグメント毎にそれぞれ管理し、所定周期毎に、各論理的記憶デバイスのそれぞれについて更新ビットマップを生成し、かつ、所定期間内に生成された各更新ビットマップの論理和に基づいて、所定期間内に更新されなかった非更新セグメントを検出するための非更新ビットマップを生成して他のメモリに記憶させ、他のメモリに記憶された非更新ビットマップは、メモリ間コピーによって、他のメモリからキャッシュメモリにコピーされる。

【0020】

マイグレート制御部は、キャッシュメモリに記憶された非更新ビットマップに含まれる非更新セグメントに記憶されているファイルの属性をファイルシステムに問い合わせることにより、マイグレートさせるべきマイグレート対象セグメントを特定するためのマイグレート対象ビットマップを生成して、このマイグレート対象ビットマップをキャッシュメモリに記憶させ、キャッシュメモリに記憶されたマイグレート対象ビットマップは、メモリ間コピーによって、キャッシュメモリから他のメモリにコピーされる。

【0021】

マイグレート制御部は、マイグレート対象セグメントに登録された非更新セグメントに記憶されているデータを、高速論理的記憶デバイスから低速記憶デバイスに移動させるための指示をマイグレート実行部に発行し、マイグレート実行部は、発行された指示及びマイグレート対象ビットマップに基づいて、高速論理的記憶デバイスに記憶されたデータを、一つまたは複数のセグメント単位で、低速論理的記憶デバイスに移動させ、ファイルシステムは、マイグレート対象データの移動が完了するまでの間、上位装置から当該マイグレート対象データへの更新を禁止する。

【0022】

本発明の機能、手段、ステップの全部または一部は、例えば、マイクロコンピュータにより実行されるコンピュータプログラムとして構成可能な場合がある。そして、このコンピュータプログラムは、例えば、ハードディスク、光ディスク、半導体メモリ等の記憶媒体に固定して配布することができる。または、コンピュータプログラムをインターネット等の通信ネットワークを介して、配信することもできる。

10

【発明を実施するための最良の形態】

【0023】

以下、図面に基づき、本発明の実施の形態を説明する。図1は、本実施形態の全体概念を示す説明図である。この記憶制御装置1は、例えば、「ファイルコントローラ」としてのNAS制御部2と、「ブロックコントローラ」としてのストレージ制御部3と、記憶部4とを備えて構成される。

【0024】

20

NAS制御部2は、例えば、第1メモリM1と、非更新ビットマップ2Aと、マイグレート対象ビットマップ2Bと、ファイル管理部2Cと、マイグレート制御部2D及びボリューム管理部2Eを備えて構成することができる。

【0025】

第1メモリM1は、ストレージ制御部3内の第2メモリM2と接続されており、コマンド等を用いずに、メモリ間コピーによって記憶内容を転送可能となっている。第1メモリM1には、非更新ビットマップ2A及びマイグレート対象ビットマップ2Bがそれぞれ記憶される。ここで、非更新ビットマップ2Aは、ストレージ制御部3によって生成され、第1メモリM1にも記憶される。マイグレート対象ビットマップ2Bは、NAS制御部2によって生成され、第2メモリM2にも記憶される。各ビットマップ2A、2Bについては、後述する。

30

【0026】

ファイル管理部2Cは、「ファイルシステム」に該当する。ファイル管理部2Cは、「上位装置」としてのクライアントマシン6から発行されたファイルアクセス要求を処理するものである。なお、クライアントマシン6は、ホストコンピュータ（「ホスト」と略記する場合がある）と呼ぶこともできる。

【0027】

マイグレート制御部2Dは、それぞれ異なるタイミングで作成された複数の非更新ビットマップ2Aに基づいて、一度も更新されていない非更新セグメントを特定し、マイグレート対象ビットマップ2Bを作成する。マイグレート制御部2Dは、非更新セグメントに記憶されているファイルの属性情報をファイル管理部2Cに問合せ、マイグレートさせるべきセグメントを決定する。そして、マイグレート制御部2Dは、決定されたセグメントについてのマイグレート実行を、ストレージ制御部3に要求する。

40

【0028】

ボリューム管理部2Eは、高速領域5A及び低速領域5Bを有する仮想ボリューム5を生成し、ファイル管理部2Cに提供するものである。仮想ボリューム5の構成は、さらに後述する。

【0029】

ストレージ制御部3は、例えば、第2メモリM2と、非更新ビットマップ3Aと、マイグレート対象ビットマップ3Bと、入出力処理部（図中「I/O処理部」）3Cと、マイグ

50

レート実行部 3 D 及び更新管理部 3 E を備えて構成することができる。

【0030】

第 2 メモリ M 2 は、例えば、他のキャッシュメモリとして構成される。NAS 制御部 2 の内部バスとストレージ制御部 3 の内部バスとは接続されているため、第 2 メモリ M 2 と第 1 メモリ M 1 との間で情報を共有することができる。第 2 メモリ M 2 には、非更新ビットマップ 3 A 及びマイグレート対象ビットマップ 3 B が記憶されている。上述のように、第 2 メモリ M 2 内の非更新ビットマップ 3 A と第 1 メモリ M 1 内の非更新ビットマップ 2 A とは同一内容であり、同様に、第 2 メモリ M 2 内のマイグレート対象ビットマップ 3 B と第 1 メモリ M 1 内のマイグレート対象ビットマップ 2 B とは同一内容である。

【0031】

入出力処理部 3 C は、NAS 制御部 2 からの要求に従って、記憶部 4 内のボリューム 4 A または 4 B にアクセスすることにより、所定のデータを各ボリューム 4 A , 4 B から読み出したり、所定のデータを各ボリューム 4 A , 4 B に書き込んだりする。

【0032】

マイグレート制御部 2 D は、マイグレート対象ビットマップ 3 B に基づいて、セグメント単位で、データのマイグレートを実行するものである。データマイグレーション方法については後述する。

【0033】

更新管理部 3 E は、入出力処理部 3 C による仮想ボリューム 5 へのアクセスに基づいて、即ち、クライアント 6 による仮想ボリューム 5 の使用状況に基づいて、各ボリューム 4 A , 4 B 毎に非更新ビットマップ 3 A を作成する。更新管理部 3 E は、例えば、毎日所定の時刻に、非更新ビットマップ 3 A を作成することができる。また、更新管理部 3 E は、所定期間内に作成された非更新ビットマップ 3 A のみを第 2 メモリ M 2 に保存させることができ、所定期間の過ぎた非更新ビットマップ 3 A は第 2 メモリ M 2 内から削除させることもできる。

【0034】

更新管理部 3 E は、例えば、ボリュームレプリケーション機能やスナップショット作成機能に使用されることもできる。また、例えば、NAS 制御部 2 がスナップショット作成を行うような場合、更新管理部 3 E を NAS 制御部 2 内に設けることもできる。

【0035】

記憶部 4 は、例えば、複数のディスクドライブ 4 C , 4 D を備えている。第 1 のディスクドライブ 4 C は、例えば、FC (Fibre Channel) ディスクのような比較的高速のデバイスとして構成される。第 2 のディスクドライブ 4 D は、例えば、ATA (AT Attachment) ディスクのような比較的低速のデバイスとして構成される。この場合、ディスクドライブ 4 C の方が、ディスクドライブ 4 D よりも高速かつ高性能である。なお、ディスクドライブ 4 C , 4 D は、ハードディスクドライブに限定されない。例えば、半導体メモリドライブ、光ディスクドライブ、光磁気ディスクドライブ等の種々のデバイスを利用することができる。

【0036】

例えば、一つまたは複数の第 1 ディスクドライブ 4 C の有する記憶領域によって、第 1 物理的記憶デバイス 4 E が生成される。RAID (Redundant Array of Independent Disks) 構成のレベルによっても相違するが、複数の記憶領域を集めて冗長化することにより、物理的な記憶デバイス 4 E が生成される。この記憶デバイス 4 E を、第 1 物理的記憶デバイス 4 E と呼ぶ。そして、例えば、この第 1 物理的記憶デバイス 4 E から一つまたは複数の第 1 論理的記憶デバイス 4 A を生成することができる。

【0037】

上記同様に、一つまたは複数の第 2 ディスクドライブ 4 D の有する記憶領域によって、第 2 物理的記憶デバイス 4 F が生成される。そして、第 2 物理的記憶デバイス 4 F に基づいて、一つまたは複数の第 2 論理的記憶デバイス 4 B を生成することができる。これらの各論理的記憶デバイス 4 A , 4 B は、論理ボリューム (Logical Unit) とも呼ばれる。

10

20

30

40

50

【 0 0 3 8 】

仮想ボリューム 5 は、第 1 論理的記憶デバイス 4 A と第 2 論理的記憶デバイス 4 B とを仮想的に結合させることにより、構成されている。この仮想化は、ボリューム管理部 2 E によって行われる。ファイル管理部 2 C は、ボリューム管理部 2 E を介して、この仮想ボリューム 5 の構成を認識することができる。

【 0 0 3 9 】

仮想ボリューム 5 は、その前半部分に位置する高速領域 5 A と、その後半部分に位置する低速領域 5 B とに大別される。高速領域 5 A は、高速な論理的記憶デバイスである第 1 論理的記憶デバイス 4 A に対応付けられている。低速領域 5 B は、低速な論理的記憶デバイスである第 2 論理的記憶デバイス 4 B に対応付けられている。高速領域 5 A は、クライアント 6 によって現在使用されている情報価値の高いデータ群が記憶されており、低速領域 5 B には、情報価値の低下したデータ群が記憶されている。

10

【 0 0 4 0 】

このように、仮想ボリューム 5 は、それぞれ性能の異なる複数種類の論理的記憶デバイス 4 A , 4 B から構成されているが、クライアント 6 は、仮想ボリューム 5 の詳細な構造まで認識することなく、一つのボリュームとして利用可能である。

【 0 0 4 1 】

次に、この記憶制御装置 1 の動作を説明する。クライアント 6 から仮想ボリューム 5 に記憶されている所定のファイルの読出しを要求された場合、NAS制御部 2 は、論理的記憶デバイス 4 A または 4 B における格納先アドレス等を明示して、ストレージ制御部 3 に所定のファイルの読出しを要求する。ストレージ制御部 3 の入出力処理部 3 C は、所定のファイルに対応するデータを論理的記憶デバイス 4 A または 4 B のいずれかから読出して、NAS制御部 2 内の第 1 メモリ M 1 に記憶させる。

20

【 0 0 4 2 】

クライアント 6 から仮想ボリューム 5 へ所定のファイル (ファイルデータ) の書込みが要求された場合、NAS制御部 2 は、論理的記憶デバイス 4 A の書込み先アドレス等を明示して、ストレージ制御部 3 に所定のファイルの書込みを要求する。入出力処理部 3 C は、ファイルデータを論理的記憶デバイス 4 A に記憶させる。

【 0 0 4 3 】

更新管理部 3 E は、高速領域 5 A (即ち、論理的記憶デバイス 4 A) へのデータ書込みを監視しており、更新されたセグメントの更新フラグを " 1 " にセットする。セグメントとは、キャッシュメモリ (この例では、第 2 メモリ M 2) 上でデータを管理するための単位である。更新管理部 3 E は、第 1 論理的記憶デバイス 4 A の各セグメント毎に、更新されたか否かを識別するための更新フラグをセットすることにより、更新ビットマップを生成する。同様に、更新管理部 3 E は、第 2 論理的記憶デバイス 4 B における更新状況を管理することもできる。このように、更新管理部 3 E は、例えば、毎日の更新状況を、各論理的記憶デバイス 4 A , 4 B の各セグメント毎にそれぞれ管理する。なお、キャッシュメモリ上のデータ管理単位としては、セグメントに限らず、他のサイズ (ブロック単位等) で管理してもよい。各セグメントは、それぞれ複数のブロックから構成可能である。

30

【 0 0 4 4 】

例えば、一週間等の所定期間分の更新ビットマップが蓄積されると、更新管理部 3 E は、各更新ビットマップの OR を演算することにより、所定期間更新されていないセグメントを検出する。ここでは、この所定期間更新されていないセグメントを非更新セグメントと呼ぶ。更新管理部 3 E は、各更新ビットマップの OR 演算することにより、非更新ビットマップ 3 A を生成し、第 2 メモリ M 2 に記憶させる。この非更新ビットマップ 3 A は、メモリ間コピーにより、第 1 メモリ M 1 にも記憶される。なお、NAS制御部 2 が、各更新ビットマップに基づいて、非更新ビットマップ 2 A を生成し、この非更新ビットマップ 2 A を第 2 メモリ M 2 にコピーさせる構成でもよい。

40

【 0 0 4 5 】

マイグレート制御部 2 D は、第 1 メモリ M 1 に記憶された非更新ビットマップに基づい

50

て、非更新セグメントに記憶されているデータの属性を、ファイル管理部 2 C に問い合わせることができる。ファイル管理部 2 C は、問い合わせされた非更新セグメント内に、データの全部または一部が記憶されているファイルを特定し、この特定されたファイルの属性を回答する。属性としては、例えば、ファイルサイズ、ファイルの所有者、参照日時等を挙げるができる。

【 0 0 4 6 】

マイグレート制御部 2 D は、非更新セグメントに記憶されているデータ（ファイルデータ）の属性に基づいて、非更新ビットマップ 2 A に登録されている各非更新セグメントのうち、マイグレート対象となるセグメントを決定する。例えば、マイグレート制御部 2 D は、所定期間更新されていないセグメントであっても、最近参照されているセグメントの場合、マイグレート対象から除外することができる。

10

【 0 0 4 7 】

なお、マイグレート制御部 2 D は、データの属性を問わずに、全ての非更新セグメントをマイグレート対象として選択することもできる。非更新ビットマップ 2 A に登録されている非更新セグメントのうち、いずれの非更新セグメントをマイグレート対象セグメントとして選択するかは、マイグレーションポリシーによって定めることができる。そして、マイグレーションポリシーは、ユーザが自由に定義することもできる。

【 0 0 4 8 】

このようにして、マイグレート制御部 2 D により、マイグレート対象のセグメントを特定したビットマップ 2 B が生成される。このマイグレート対象ビットマップ 2 B は、第 2

20

メモリ M 2 にもコピーされる。

【 0 0 4 9 】

マイグレート実行部 3 D は、第 2 メモリ M 2 内のマイグレート対象ビットマップ 3 B に基づいて、マイグレート対象のセグメントを 1 つずつ、または複数個まとめて、データマイグレーションを実行する。即ち、マイグレート実行部 3 D は、マイグレート対象のセグメントに記憶されているデータを、第 1 論理的記憶デバイス 4 A から第 2 論理的記憶デバイス 4 B にコピーする。これにより、仮想ボリューム 5 において、高速領域 5 A 内のデータが低速領域 5 B に移され、高速領域 5 A の空き容量が増大する。そして、空き容量の増加した高速領域 5 A（第 1 論理的記憶デバイス 4 A）には、情報価値の高い新たなファイルデータが書き込まれる。

30

【 0 0 5 0 】

本実施形態は、このように構成されるため、以下の効果を奏する。本実施形態では、ファイルそのものをファイル単位で直接移動させるのではなく、論理的記憶デバイス 4 A、4 B に記憶されているファイルデータをセグメント単位で移動させ、結果的に、擬似的にファイル単位の移動を実現する。従って、従来技術のように、ファイルツリーの全てを検索してマイグレート対象を特定し、ファイル毎に移動させる必要がない。これにより、NAS制御部 2 がマイグレート対象を特定するための負荷を軽減することができ、NAS制御部 2 のコンピュータ資源を本来のNASサービスに割り当てることができ、使い勝手も向上する。

【 0 0 5 1 】

本実施形態では、セグメント単位で移動の可否を判断し、セグメント単位でデータを移動させるため、移動前後の整合性を維持するための更新排他制御を行う期間を短くすることができる。従って、移動対象のデータを速やかに利用することができ、使い勝手が向上する。

40

【 0 0 5 2 】

本実施形態では、マイグレート対象のデータを特定するための処理を、ストレージ制御部 3 による非更新セグメントを検出する処理と、NAS制御部 2 による非更新セグメントの中からマイグレート対象セグメントを選択する処理との 2 つの処理に分割し、これら非更新セグメントの検出処理とマイグレート対象セグメントの選択処理とを、非同期で実行させる構成とした。従って、NAS制御部 2 とストレージ制御部 3 とが協働で、マイグレート

50

対象の特定処理を行うことができ、NAS制御部2に負荷が集中するのを防止することができる。

【0053】

本実施形態では、NAS制御部2は、マイグレート対象のデータを特定し、実際のデータマイグレーションはストレージ制御部3で実行する。従って、NAS制御部2がファイルのデータを読み出して、この読出したデータを移動先の論理的記憶デバイス4Bに書き込む必要がない。即ち、データマイグレーションは、NAS制御部2を経由することなく、ストレージ制御部3内で実行される。これにより、データマイグレーションに関するNAS制御部2の負荷をより一層低減することができる。

【0054】

本実施形態では、スナップショット作成等に利用される既存の更新管理部3Eを用いて、マイグレート対象を特定するための前処理（非更新セグメントの検出処理）を行う構成とした。従って、ストレージ制御部3の構成を大幅に変更することなく、セグメント単位のデータマイグレーションを実現することができる。

【0055】

本実施形態では、NAS制御部2内の第1メモリM1とストレージ制御部3内の第2メモリM2との間で、非更新ビットマップ2A、3A及びマイグレート対象ビットマップ2B、3Bを共有する構成とした。従って、例えば、コマンドを用いて情報交換を行う必要がなく、情報共有のための構成を簡素化することができる。そして、各制御部2、3に負担をかけずに、非更新ビットマップ2A、3A及びマイグレート対象ビットマップ2B、3Bを共有することができる。

【0056】

本実施形態では、各更新ビットマップはストレージ制御部3内で保持し、これら各更新ビットマップから生成される非更新ビットマップ3A(2A)を、NAS制御部2と共有する構成とした。従って、第1メモリM1の記憶領域を効率的に使用できる。なお、NAS制御部2内で非更新ビットマップ2Aを作成する構成でもよい。以下、本実施形態を詳細に説明する。

【実施例1】

【0057】

図2は、記憶制御装置10の全体構成を示す構成説明図である。先に、図1との対応関係を説明すると、図2中の記憶制御装置10は図1中の記憶制御装置1に、図2中のNASノード100は図1中のNAS制御部2に、図2中のストレージコントローラ200は図1中のストレージ制御部3に、図2中の記憶部300は図1中の記憶部4に、図2中のクライアント20は図1中のクライアント6に、それぞれ対応する。

【0058】

記憶制御装置10は、例えば、NASノード100と、ストレージコントローラ200及び記憶部300を備えて構成される。記憶制御装置10は、例えば、LAN等の通信ネットワークCN1を介して、一つまたは複数のクライアントマシン20と、管理端末30とに接続されている。クライアントマシン20は、記憶制御装置10を利用してファイルの入出力を行うためのコンピュータ装置である。管理端末30は、記憶制御装置10を管理するためのコンピュータ装置である。管理端末30は、例えば、記憶制御装置10の構成変更等を指示したり、記憶制御装置10の各種状態を確認することができる。

【0059】

NASノード100は、例えば、マイクロプロセッサ（以下、「プロセッサ」）110と、メモリ120と、ネットワークインターフェース（以下、インターフェースを「I/F」と略記する場合がある）部130と、内部バス140と、ブリッジ回路150とを備えて構成される。

【0060】

プロセッサ110は、所定のコンピュータプログラムを読み込んで、所定の機能を実現するものである。プロセッサ110は、NASとしてのデータ処理サービスを実現するほか

10

20

30

40

50

に、後述のように、データマイグレーション処理を行う。

【0061】

メモリ120は、例えば、RAM(Random Access Memory)やフラッシュメモリ等から構成される。メモリ120内には、LVM定義テーブルT1と非更新ビットマップT2及びマイグレート対象ビットマップT3が記憶されている。メモリ120は、プロセッサ110等とともに内部バス140に接続されている。また、メモリ120は、内部バス140及びブリッジ回路150を介して、ストレージコントローラ200の内部バス240にも接続されている。従って、メモリ120は、ストレージコントローラ200内のキャッシュメモリ220にも接続されており、コマンドを用いることなく、各メモリ120, 220間で、各種情報T1~T3を共有可能となっている。

10

【0062】

ネットワークI/F部130は、通信ネットワークCN1を介して、各クライアント20や管理端末30と通信を行うためのものである。ブリッジ回路150は、NASノード100の内部バス140とストレージコントローラ200の内部バス240とを接続するための回路である。

【0063】

ストレージコントローラ200は、例えば、プロセッサ210と、キャッシュメモリ220と、ドライブI/F部230及び内部バス240とを備えて構成される。プロセッサ210は、所定のコンピュータプログラムを読み込んで、所定の機能を実現する。所定の機能としては、ブロック単位のデータの入出力処理や、更新管理処理、マイグレート実行処理等が挙げられる。

20

【0064】

キャッシュメモリ220は、クライアント20によって使用されるユーザデータや、各種の制御情報、管理情報を記憶するものである。キャッシュメモリ220には、上述のメモリ120と同様に、LVM定義テーブルT1と非更新ビットマップT2及びマイグレート対象ビットマップT3がそれぞれ記憶されている。

【0065】

記憶部300は、例えば、ディスクアレイエンクロージャとして構成されるもので、複数のディスクドライブ310, 311を備えている。第1の種類ディスクドライブ310は、例えば、FCディスク等のように、比較的高速かつ高性能のドライブである。第2の種類ディスクドライブ311は、例えば、ATAディスク等のように、比較的低速のドライブである。

30

【0066】

本実施例では、FCディスク310から構成される論理的記憶デバイス330(図3参照)内に、新しいユーザデータを格納し、ATAディスク311から構成される論理的記憶デバイス331(図3参照)内に、情報価値の低下した旧データを保存する。情報価値の低下した旧データを直ちに削除することはできないため、FCディスク310の数よりも多い数のATAディスク311が記憶部300に設けられている。

【0067】

図3は、NASノード100のソフトウェア構成及び階層記憶の構成を模式的に示す説明図である。NASノード100は、例えば、ファイルシステム(NASのOS)111と、論理ボリュームマネージャ(「LVM」と略記)112及びマイグレートツール113とを備えている。

40

【0068】

ファイルシステム111は、ファイル群を管理し、ファイルの共有サービスをクライアント20に提供するものである。LVM112は、仮想ボリューム340を管理するためのものである。マイグレートツール113は、マイグレート制御を行うものである。

【0069】

本記憶制御装置10の提供する階層記憶について説明する。第1論理的記憶デバイス330と第2論理的記憶デバイス331とを論理的に結合させることにより、仮想的な一つ

50

のボリューム 340 が構成されている。

【0070】

第1論理的記憶デバイス330は、第1ディスクドライブ310（以下、FCディスク310とも呼ぶ）に由来する。一つまたは複数のFCディスク310の提供する記憶領域をRAID構成に基づいて仮想化することにより、第1物理的記憶デバイス320（図中「FC領域320」）が形成される。この第1物理的記憶デバイス320の全部または一部の記憶領域から、第1論理的記憶デバイス330が生成される。

【0071】

同様に、第2論理的記憶デバイス331は、第2ディスクドライブ311（以下、ATAディスクとも呼ぶ）に由来する。一つまたは複数のATAディスク311の提供する記憶領域をRAID構成に基づいて仮想化することにより、第2物理的記憶デバイス321（図中「ATA領域321」）が生成される。この第2物理的記憶デバイス321の全部または一部の記憶領域から、第2論理的記憶デバイス331が生成される。なお、以下の説明では、各論理的記憶デバイス330、331を論理ボリューム330、331と呼ぶ場合がある。

【0072】

上述の通り、仮想ボリューム340は、その前半部分に論理ボリューム330を、その後半部分に論理ボリューム331を配置するようにして、構成されている。前半、後半とは、仮想ボリューム340の先頭アドレスに近い領域を前半部分と呼び、前半部分に続く他の領域を後半部分と呼ぶ。

【0073】

図4は、仮想ボリューム340内の構成を模式的に示す説明図である。上述の通り、仮想ボリューム340の前半部分には、FCディスク310に由来する論理ボリューム330が対応付けられている。この論理ボリューム330には、iノード管理領域330A及び第1ファイルデータ領域330Bが設けられている。iノード管理領域330Aの詳細は、さらに後述する。第1ファイルデータ領域330Bは、ユーザデータを記憶するための記憶領域である。第1ファイルデータ領域330Bには、クライアント20によって使用されている比較的情報価値の高いユーザデータが記憶されている。

【0074】

仮想ボリューム340の後半部分には、ATAディスク311に由来する論理ボリューム331が対応付けられている。この論理ボリューム331には、第2ファイルデータ領域331Aが設けられている。第2ファイルデータ領域331Aには、情報価値の低下したユーザデータが保存されている。

【0075】

第1ファイルデータ領域330Bの先頭アドレスをSA1、第2ファイルデータ領域331Aの先頭アドレスをSA2とする。LVM112は、仮想ボリューム340の構成を認識しており、仮想ボリューム340のアドレス空間が、実際にはどの論理ボリューム330、331のどのアドレス空間に割り当てられているのかを把握している。LVM112は、仮想ボリューム340内の前半部分のアドレス空間と後半部分のアドレス空間が連続するように、論理ブロックアドレス（LBA）やセグメント番号を管理する。なお、LVM112は、スナップショット機能等を備えることもできる。スナップショットとは、指示された時点における記憶内容の静止化イメージである。

【0076】

図5は、iノード管理領域330A等の構成を示す説明図である。iノード管理領域330Aには、仮想ボリューム340内に記憶されたファイル群を管理するための情報が記憶されている。例えば、iノード管理領域330Aには、iノード番号と、他のiノード情報と、ATAフラグと、セグメント番号（図中では、番号を「#」で表記する場合がある）と、オフセットアドレスが対応付けられて記憶されている。

【0077】

iノード番号とは、ファイルシステム上でファイルを識別するための番号である。仮想

10

20

30

40

50

ボリューム 3 4 0 内の各ファイル（ユーザデータ）は、i ノード番号によって、その格納位置が特定される。

【 0 0 7 8 】

他の i ノード情報とは、各 i ノード番号で特定される各ファイル（ディレクトリを含む場合もある）の属性を示す情報である。ファイルの属性情報としては、例えば、所有者名、ファイルサイズ、アクセスタイム等を挙げるができる。

【 0 0 7 9 】

ATAフラグとは、その i ノード番号で特定されるファイルが、各論理ボリューム 3 3 0 , 3 3 1 のいずれに記憶されているかを識別するための情報である。本実施例では、データマイグレーションが実行されると、高速な論理ボリューム 3 3 0 内のデータは、低速な論理ボリューム 3 3 1 に移される。ファイルデータが論理ボリューム 3 3 0 に記憶されている場合、そのATAフラグには " 0 " がセットされる。ファイルデータが論理ボリューム 3 3 1 に記憶されている場合、そのATAフラグには " 1 " がセットされる。従って、ATAフラグは、データ格納先の論理ボリュームを特定する役割と、マイグレートの実行の有無を識別する役割を担う。

【 0 0 8 0 】

セグメント番号及びオフセットアドレスは、i ノード番号で特定されたファイルの格納先を示すためのアドレス情報である。本実施例では、ATAフラグとセグメント番号及びオフセットアドレスによって、ファイルデータの格納先を特定する。上述のように、ATAフラグの値によって、ファイルデータが格納されている論理ボリュームを特定することができる。そして、セグメント番号とオフセットアドレスによって、その論理ボリュームのどこにファイルデータが格納されているかを特定することができる。

【 0 0 8 1 】

なお、i ノード管理領域 3 3 0 A で管理されるセグメント番号と、各ビットマップ T 2 , T 3 , T 4 のセグメント番号とは完全に一致している。即ち、NASノード 1 0 0 で管理するセグメント番号と、ストレージコントローラ 2 0 0 で管理するセグメント番号とは一致しており、NASノード 1 0 0 からストレージコントローラ 2 0 0 にセグメント番号等を通知するだけで、セグメント単位のデータ移動を行えるようになっている。

【 0 0 8 2 】

セグメント番号にセグメントサイズ（例えば、1 MB）を乗算した値が、ATAフラグで示された論理ボリューム内での格納先を示す。図 5 に示す例では、i ノード番号 " 2 " のルートディレクトリは、ATAフラグが " 0 " 、セグメント番号が " 000001h " 、セグメントサイズが 1 MB、オフセットアドレスが " aa " であるから、このルートディレクトリは、論理ボリューム 3 3 0 内のセグメント 000001h の先頭から aa 分だけオフセットした位置を先頭アドレスとして、記憶されていることになる。即ち、このルートディレクトリは、第 1 ファイルデータ領域 3 3 0 B に記憶されている。

【 0 0 8 3 】

同様に、i ノード番号 " 1 0 0 " を有する " /usr/dir1/file1.txt " は、ATAフラグが " 0 " 、セグメント番号が " 000025h " 、オフセットアドレスが " dd " であるから、このファイルのデータは、第 1 ファイルデータ領域 3 3 0 B の先頭アドレス S A 1 から、1MB × 2 5 + dd だけ離れた箇所に記憶されている。このファイル " /usr/dir1/file1.txt " を論理ボリューム 3 3 0 から論理ボリューム 3 3 1 に移動させる場合、このファイルのATAフラグには " 1 " がセットされる。また、このファイルのセグメント番号及びオフセットアドレスは、コピー先のセグメント番号及びオフセットアドレスに、それぞれ変更される。

【 0 0 8 4 】

論理ボリューム 3 3 1 の先頭アドレスは、S A 2 である。従って、第 2 ファイルデータ領域 3 3 1 A 内では、S A 2 + セグメント番号 × セグメントサイズ + オフセットアドレスによって、格納先が特定される。

【 0 0 8 5 】

図 6 は、LVM定義テーブル T 1 の構成を示す説明図である。LVM定義テーブル T 1 は、仮

10

20

30

40

50

想ボリューム340を管理するためのものである。上述のように、NASノード100及びストレージコントローラ200は、同一内容のLVM定義テーブルT1をそれぞれ保持している。LVM定義テーブルT1は、例えば、FC領域管理テーブルT1AとATA領域管理テーブルT1Bとを備えて構成される。図6中、ファイルシステム111を「FS」と略記する。

【0086】

FC領域管理テーブルT1Aは、仮想ボリューム340を構成するFC領域を管理するためのテーブルである。このテーブルT1Aは、例えば、ファイルシステム111が認識するLUN(Logical Unit Number)と、NASノード100が認識するFC領域を識別するための情報と、NASノード100が認識するFC領域上のLUNと、ストレージコントローラ200が認識するFC領域上のLUNとを対応付けることにより、構成されている。

10

【0087】

ファイルシステム111が認識するLUNとは、仮想ボリューム340に設定されているLUNである。図3では一つの仮想ボリューム340のみを示しているが、本実施例では、複数の仮想ボリューム340を設けることができる。

【0088】

NASノード100が認識するFC領域を識別するための情報とは、仮想ボリューム340におけるFC領域部分、即ち、論理ボリューム330に対応付けられている前半部分を識別するための情報である。

【0089】

NASノード100が認識するFC領域上のLUNとは、仮想ボリューム340の前半部分を構成するFC領域に対応付けられている論理ボリューム330にアクセスするための情報である。論理ボリューム330, 331は、それぞれ2種類のLUNを備える。一つのLUNは、NASノード100から認識される値であり、他の一つは、ストレージコントローラ200から認識される値である。NASノード100が認識するFC領域上のLUNは、例えば、"c#t#d#"の形式で表現される。"c#"は、ドライブI/Fの番号を示し、"t#"は、SCSI(Small Computer System Interface)ターゲット番号を示す。"d#"は、LUNの値を示す。

20

【0090】

ストレージコントローラ200が認識するFC領域上のLUNは、上述のように、ストレージコントローラ200内において使用されるLUNを示す。

30

【0091】

ATA領域管理テーブルT1Bは、仮想ボリューム340を構成するATA領域を管理するためのテーブルである。このATA領域管理テーブルT1Bも、上述のFC領域管理テーブルT1Aと同様に、構成されている。即ち、ATA領域管理テーブルT1Bは、例えば、ファイルシステム111が認識するLUNと、NASノード100が認識するATA領域を識別するための情報と、NASノード100が認識するATA領域上のLUNと、ストレージコントローラ200が認識するATA領域上のLUNとを対応付けることにより、構成されている。

【0092】

LVM定義テーブルT1の構成から明らかなように、本実施例の仮想ボリューム340は、そのFC領域及びATA領域の構成を変更可能となっている。即ち、仮想ボリューム340には、その前半部分のFC領域に、FCディスク310にそれぞれ由来する複数の論理ボリューム330を対応付けることができる。また、仮想ボリュームの後半部分のATA領域にも、ATAディスク311にそれぞれ由来する複数の論理ボリューム331を対応付けることができる。

40

【0093】

つまり、後述の実施例に示すように、複数の論理ボリューム330を結合して、仮想ボリューム340内のFC領域を拡張することができる。また、複数の論理ボリューム331を結合して、仮想ボリューム340内のATA領域を拡張することもできる。

【0094】

FC領域及びATA領域のサイズを拡張する場合、それぞれの領域内で、セグメント番号が

50

連続するように、セグメント番号が管理される。即ち、ある論理ボリューム330(1)と別の論理ボリューム330(2)とを結合させて仮想ボリューム340内のFC領域を構成する場合、論理ボリューム330(1)の終端セグメント番号の次の値を、論理ボリューム330(2)の先頭セグメント番号とする。

【0095】

FC領域内及びATA領域内で、セグメント番号がそれぞれ連続していればよく、FC領域とATA領域との間で、セグメント番号が連続している必要はない。ビットマップT2~T3も、各領域単位でそれぞれ管理される。

【0096】

次に、記憶制御装置10の動作について説明する。まず、図7は、非更新ビットマップT2を生成するための処理を示すフローチャートである。この処理は、ストレージコントローラ200によって実行される。なお、図中ではステップを「S」と略記する。また、各フローチャートは、本発明の理解及び実施に必要な程度で処理の概要を示しており、実際のプログラムとは相違する。

【0097】

図8は、図7に示す処理の流れを模式的に示す説明図である。以下、図7及び図8を適宜参照しながら、非更新ビットマップ生成処理を説明する。ストレージコントローラ200は、所定の周期で(例えば、毎日)、仮想ボリューム340の更新状態を管理する更新ビットマップT4を作成する(S1)。

【0098】

この更新ビットマップT4は、FC領域及びATA領域内のセグメント毎に、“1”または“0”の値をセットすることにより、生成される。更新フラグに“1”がセットされた場合は、そのセグメントが更新されたことを示す。更新フラグに“0”がセットされている場合は、そのセグメントが更新されていないことを示す。ストレージコントローラ200は、クライアント20によって更新されたセグメントについて、更新フラグを“1”にセットする。

【0099】

このようにして作成された更新ビットマップT4は、キャッシュメモリ220内に記憶される(S2)。図8において、符号T4に添えられた()内の数字は、作成の順番(作成日)を示す。ストレージコントローラ200は、所定期間であるn日間が経過すると、このn日間に作成された各更新ビットマップT4(1)~T4(n)の各セグメント毎にそれぞれOR演算を行い、非更新ビットマップT2を作成する(S3)。

【0100】

非更新ビットマップT2では、あるセグメントがn日間内に1回でも更新されている場合、そのセグメントの非更新フラグには“0”がセットされる。別のセグメントがn日間内に1回も更新されなかった場合、そのセグメントの非更新フラグには“1”がセットされる。

【0101】

このようにして生成された非更新ビットマップT2は、キャッシュメモリ220に記憶され、そして、キャッシュメモリ220からNASノード100内のメモリ120にもコピーされる(S4)。ストレージコントローラ200は、n+1日間以上経過した古い更新ビットマップT4をキャッシュメモリ220から削除する(S5)。

【0102】

図8を参照する。詳細は別のフローチャートと共に説明するが、マイグレートツール113は、メモリ120に記憶された非更新ビットマップT2に基づいて、非更新セグメントに記憶されているデータの属性をファイルシステム111に問い合わせる。マイグレートツール113は、データの属性が予め設定されたポリシー113Aに合致するか否かを判断する。

【0103】

マイグレートツール113は、ポリシー113Aに合致する非更新セグメントをマイグ

10

20

30

40

50

レート対象のセグメントとして選択し、このマイグレート対象のセグメントに " 1 " をセットする。このようにして、マイグレート対象ビットマップ T 3 が生成され、メモリ 1 2 0 に記憶される。そして、このマイグレート対象ビットマップ T 3 は、メモリ 1 2 0 からキャッシュメモリ 2 2 0 にコピーされる。

【 0 1 0 4 】

図 9 は、マイグレーションポリシー 1 1 3 A を設定するために使用されるマイグレーションポリシー設定テーブル T 5 の一例を示す説明図である。マイグレーションポリシー設定テーブル T 5 は、例えば、予め登録されている一つまたは複数のポリシー内容と、各ポリシーを選択するか否かを示す選択フラグとを対応付けることにより構成される。

【 0 1 0 5 】

ポリシーの内容としては、例えば、「 n 2 (n 2 n) 日間以上参照されていないファイルを移動させる。」、「アーカイブ対象として予め指定されているファイルを移動させる。」、「ファイルサイズが D S 1 以下のファイルを移動させる。」、「ファイルサイズが D S 2 以上のファイルを移動させる。」、「予め指定された特定の所有者のファイルを移動させる。」等を挙げることができる。ポリシーの内容は、予め設定しておくこともできるし、ユーザによる自由な設定を許可することもできる。また、クライアント 2 0 または管理端末 3 0 からポリシーの内容を設定可能な構成とすることもできる。

【 0 1 0 6 】

ユーザは、マイグレーションポリシー設定テーブル T 5 に登録されているポリシーの中から、所望のポリシーをいずれか一つまたは複数選択することができる。ユーザにより選択されたポリシーには、その選択フラグに " 1 " がセットされる。選択されたポリシーを結合することにより、マイグレーションポリシー 1 1 3 A が生成される。なお、ユーザは、必ずしも一つ以上のポリシーを選択する必要はない。一つもポリシーが選択されなかった場合、非更新ビットマップ T 2 内の全ての非更新セグメントが、データマイグレーションの対象となる。

【 0 1 0 7 】

図 1 0 は、仮想ボリューム 3 4 0 内に設けられるファイルツリーに、新たなファイルを追加する場合の処理を示すフローチャートである。クライアント 2 0 がファイルの追加を指示すると (S 1 1)、NAS ノード 1 0 0 のファイルシステム 1 1 1 は、新たに追加されるファイルのための i ノードを i ノード管理領域 3 3 0 A に確保する (S 1 2)。

【 0 1 0 8 】

ファイルシステム 1 1 1 は、新たに追加されるファイルのデータを格納する範囲 (記憶領域) を、仮想ボリューム 3 4 0 内の FC 領域内に確保する (S 1 3)。即ち、ファイルシステム 1 1 1 は、格納すべきデータのサイズに応じて、仮想ボリューム 3 4 0 の有する FC 領域内の空きセグメントまたは使用されているセグメント内の未使用領域を、必要なだけ確保する。

【 0 1 0 9 】

続いて、ファイルシステム 1 1 1 は、ファイルデータを格納する範囲に関する情報 (図中、格納先情報と示す) を特定する (S 1 4)。即ち、ファイルシステム 1 1 1 は、ファイルデータが記憶される先頭セグメントの番号とオフセットアドレス及びファイルサイズをそれぞれ特定する。

【 0 1 1 0 】

次に、ファイルシステム 1 1 1 は、 S 1 4 で特定された格納先情報を仮想ボリューム 3 4 0 上の格納先情報に変換し (S 1 5)、LVM 1 1 2 にファイルデータの格納を指示する (S 1 6)。この格納指示には、仮想ボリューム 3 4 0 におけるデータ格納先の先頭アドレスと、ファイルサイズ及びファイルデータが含まれている。

【 0 1 1 1 】

LVM 1 1 2 は、仮想ボリューム 3 4 0 における格納先情報を、LVM 定義テーブル T 1 を参照して、論理ボリューム 3 3 0 に記憶させるための情報に変換し、ストレージコントローラ 2 0 0 にファイルデータの格納を指示する (S 1 7)。この指示には、格納先の論理ボ

10

20

30

40

50

リユーム 3 3 0 を特定する情報と、格納先の論理ボリューム 3 3 0 における格納先先頭アドレスと、ファイルサイズ及びファイルデータが含まれる。

【 0 1 1 2 】

ストレージコントローラ 2 0 0 は、LVM 1 1 2 からの指示を受領すると、LVM 1 1 2 から受け取ったファイルデータを、LVM 1 1 2 から指示された所定の領域に記憶させる (S 1 8)。ストレージコントローラ 2 0 0 は、ファイルデータの記憶されたセグメントについて、更新ビットマップ T 4 の更新フラグを " 1 " にセットし (S 1 9)、LVM 1 1 2 に処理完了を通知する (S 2 0)。

【 0 1 1 3 】

LVM 1 1 2 は、ストレージコントローラ 2 0 0 からの処理完了を受領すると、ファイルシステム 1 1 1 に処理完了を報告する (S 2 1)。ファイルシステム 1 1 1 は、LVM 1 1 2 からファイルデータの記憶が正常に完了した旨の報告を受け取ると、この新たに書き込まれたファイルデータに関する i ノード情報を更新する (S 2 2)。これにより、クライアント 2 0 は、ファイルの追加が完了した旨を認識する (S 2 3)。

【 0 1 1 4 】

図 1 1 は、仮想ボリューム 3 4 0 内のファイルを更新する場合の処理を示すフローチャートである。ファイル更新の前に、クライアント 2 0 は、仮想ボリューム 3 4 0 から対象のファイルデータを読み出し、そのファイルを部分的に書き換える等して、ファイルの更新を指示する。ファイルデータの読出し処理は、ファイルの更新処理から容易に理解されるため、説明を省略する。

【 0 1 1 5 】

クライアント 2 0 がファイルの更新を要求すると (S 3 1)、ファイルシステム 1 1 1 は、更新を要求されたファイルの i ノード情報に基づいて、ファイルデータの更新範囲を特定する (S 3 2)。以下、ファイルデータ中の更新される範囲のデータを更新データと呼ぶ。ファイルシステム 1 1 1 は、更新データのセグメント番号及びオフセットアドレスを、仮想ボリューム 3 4 0 におけるアドレス情報に変換し (S 3 3)、LVM 1 1 2 に更新データの格納を指示する (S 3 4)。この指示には、仮想ボリュームにおけるアドレス情報とデータサイズ及び更新データが含まれている。

【 0 1 1 6 】

LVM 1 1 2 は、ファイルシステム 1 1 1 からの指示に基づいて、仮想ボリューム 3 4 0 上のアドレス情報を論理ボリューム 3 3 0 におけるアドレス情報に変換し、ストレージコントローラ 2 0 0 に更新データの格納を指示する (S 3 5)。この指示には、論理ボリューム 3 3 0 におけるアドレス情報とデータサイズ及び更新データが含まれる。

【 0 1 1 7 】

ストレージコントローラ 2 0 0 は、LVM 1 1 2 からの指示に従って、論理ボリューム 3 3 0 の指定されたアドレスに更新データを格納させる (S 3 6)。そして、ストレージコントローラ 2 0 0 は、更新データが格納されたセグメントについて、更新ビットマップ T 4 の更新フラグを " 1 " にセットし (S 3 7)、LVM 1 1 2 に処理完了を通知する (S 3 8)。

【 0 1 1 8 】

LVM 1 1 2 は、ストレージコントローラ 2 0 0 から処理完了を通知されると、ファイルシステム 1 1 1 に処理完了を報告する (S 3 9)。ファイルシステム 1 1 1 は、LVM 1 1 2 からの処理完了報告を受領すると、更新データに関するファイルの i ノード情報を更新する (S 4 0)。そして、クライアント 2 0 は、ファイルの更新が正常に完了したことを確認する (S 4 1)。

【 0 1 1 9 】

図 1 2 は、データマイグレーション処理を示すフローチャートである。この処理は、例えば、一日に一回実行させることができる。まず、データマイグレーションの実施に先立って、マイグレート対象ビットマップ T 3 が生成される (S 5 0)。

【 0 1 2 0 】

10

20

30

40

50

即ち、マイグレートツール 1 1 3 は、非更新ビットマップ T 2 及びマイグレーションポリシー 1 1 3 A に基づいて、非更新セグメントのうちマイグレーションポリシー 1 1 3 A に合致するセグメントを抽出し、マイグレート対象ビットマップ T 3 を生成する。マイグレート対象ビットマップ T 3 を生成する際に、マイグレートツール 1 1 3 は、非更新セグメントにデータが記憶されているファイルの属性について、ファイルシステム 1 1 1 に問い合わせることができる。このようにして生成されたマイグレート対象ビットマップ T 3 は、NAS ノード 1 0 0 内のメモリ 1 2 0 に記憶され、メモリ 1 2 0 からストレージコントローラ 2 0 0 内のキャッシュメモリ 2 2 0 にコピーされる。

【 0 1 2 1 】

マイグレートツール 1 1 3 は、メモリ 1 2 0 内に記憶されているマイグレート対象ビットマップ T 3 を参照し (S 5 1)、最初のセグメントを一つ指定する (S 5 2)。ファイルシステム 1 1 1 は、マイグレートツール 1 1 3 によりマイグレート対象として指定されたセグメントに関して、i ノード管理領域 3 3 0 A を参照し (S 5 3)、指定されたセグメントにデータを格納しているファイルを特定する (S 5 4)。ファイルシステム 1 1 1 は、指定されたセグメントにデータを格納しているファイルについて、「更新排他ロック」を設定する (S 5 5)。更新排他ロックとは、クライアント 2 0 によるファイルへの更新を禁止する処理である。データマイグレート中に、移動元のデータが更新されるのを防止し、移動の前後でデータの整合性を維持するために、更新を禁止する。

【 0 1 2 2 】

マイグレートツール 1 1 3 は、マイグレート対象のファイルへの更新が禁止されたことを確認すると、ストレージコントローラ 2 0 0 に向けて、データのマイグレート実行を指示する (S 5 6)。この指示には、ストレージコントローラ 2 0 0 によって認識されている FC 領域上の LUN 及びセグメント番号が含まれる。

【 0 1 2 3 】

ストレージコントローラ 2 0 0 は、マイグレートツール 1 1 3 からの指示を受領すると、LVM 定義テーブル T 1 を参照し (S 5 7)、指定された FC 領域に対応する ATA 領域を特定する (S 5 8)。そして、ストレージコントローラ 2 0 0 は、その ATA 領域において、次の空きセグメントの番号を確保する (S 5 9)。ATA 領域では、空いているセグメントから順番に使用され、隙間が空かないようにデータが記憶されていく。

【 0 1 2 4 】

ストレージコントローラ 2 0 0 は、移動元のセグメントからデータを読み出し、このデータを S 5 9 で確保した空きセグメントに記憶させる (S 6 0)。そして、ストレージコントローラ 2 0 0 は、データのコピーされたセグメントの番号、即ち、データの移動先のセグメント番号をマイグレートツール 1 1 3 に通知する (S 6 1)。

【 0 1 2 5 】

マイグレートツール 1 1 3 は、ストレージコントローラ 2 0 0 からデータ移動先のセグメント番号を受領すると、移動されたファイルに関する情報の切替を、ファイルシステム 1 1 1 に指示する (S 6 2)。この指示には、S 6 1 で通知されたデータ移動先のセグメント番号が含まれる。

【 0 1 2 6 】

ファイルシステム 1 1 1 は、移動されたファイル、即ち、S 5 5 で更新排他ロックの設定されたファイルについて、i ノード領域 3 3 0 A に登録されている情報を一部書き換える (S 6 3)。具体的には、ファイルシステム 1 1 1 は、移動されたファイルの ATA フラグを " 0 " から " 1 " に変更し、格納先のセグメント番号を移動先のセグメント番号に変更する。また、ファイルシステム 1 1 1 は、移動元のセグメント番号、即ち、FC 領域上でマイグレート対象として指定されたセグメント番号のステータスを、空きセグメントに変更する。そのセグメントは、別のデータの格納に直ちに使用可能となる。

【 0 1 2 7 】

ファイルシステム 1 1 1 が、S 5 5 で設定した更新排他ロックを解除すると (S 6 4)、マイグレートツール 1 1 3 は、S 5 2 で選択された一つのセグメントのデータ移動が完

10

20

30

40

50

了したことを確認する（S 6 5）。マイグレート対象ビットマップ T 3 に登録されたマイグレート対象のセグメントを全て移動させるまで、S 5 2 ~ S 6 5 のステップが繰り返される。

【0 1 2 8】

本実施例は上述のように構成されるので、以下の効果を奏する。本実施例では、ファイルデータをセグメント単位で移動させるため、移動元のボリュームに格納されたファイルツリーを全検索等する必要がない。これにより、NASノード 1 0 0 の負荷を低減することができ、データマイグレーション時のNASノード 1 0 0 の性能低下を抑制して、使い勝手を高めることができる。

【0 1 2 9】

本実施例の記憶制御装置 1 0 では、複数の論理ボリューム 3 3 0 , 3 3 1 を結合させて大きな仮想ボリューム 3 4 0 を構築可能であるから、仮想ボリューム 3 4 0 内に、多量のファイルを有するファイルツリーが構築されやすい。しかし、比較的大きなファイルツリーが形成された場合でも、本実施例によれば、NASノード 1 0 0 の負荷増大を抑制して、効率的にデータマイグレートを行うことができる。

【0 1 3 0】

本実施例では、セグメント単位で移動の可否を判断し、セグメント単位でデータを移動させるため、更新排他ロックの設定時間を短縮できる。従って、クライアント 2 0 は、移動後のデータを速やかに利用することができ、使い勝手が向上する。

【0 1 3 1】

本実施例では、セグメント単位でデータを移動させるため、データの移動が完了した場合に、移動元のセグメントを空きセグメントとして直ちに再利用することができ、使い勝手が向上する。

【0 1 3 2】

本実施例では、ストレージコントローラ 2 0 0 による非更新セグメントの検出処理（非更新ビットマップ T 2 の生成処理）と、NASノード 1 0 0 によるマイグレート対象セグメントの選択処理（マイグレート対象ビットマップ T 3 の生成処理）とを非同期に実行させることにより、データマイグレート対象を決定する。従って、NASノード 1 0 0 とストレージコントローラ 2 0 0 とが協働で、マイグレート対象を特定することができ、NASノード 1 0 0 に負荷が集中するのを防止することができる。

【0 1 3 3】

本実施例では、NASノード 1 0 0 は、マイグレート対象のデータを特定し、実際のデータマイグレーションはストレージコントローラ 2 0 0 で実行する。従って、NASノード 1 0 0 の負荷をより一層軽減することができる。

【0 1 3 4】

本実施例では、スナップショット作成等にも利用される既存の更新ビットマップ T 4 を利用して、マイグレート対象を特定するための前処理（非更新セグメントの検出処理）を行う。従って、ストレージコントローラ 2 0 0 の構成を大幅に変更することなく、セグメント単位のデータマイグレーションを実現することができる。

【0 1 3 5】

本実施例では、NASノード 1 0 0 内のメモリ 1 2 0 とストレージコントローラ 2 0 0 内のキャッシュメモリ 2 2 0 との間で、非更新ビットマップ T 2 及びマイグレート対象ビットマップ T 3 を共有する構成とした。従って、例えば、コマンドを用いて各ビットマップ T 2 , T 3 を転送する必要がない。これにより、比較的簡易な構成で、かつ、NASノード 1 0 0 及びストレージコントローラ 2 0 0 の負荷を増大させずに、各ビットマップ T 2 , T 3 を共有することができる。

【0 1 3 6】

本実施例では、各更新ビットマップ T 4 はストレージコントローラ 2 0 0 内で保持し、これら各更新ビットマップから生成される非更新ビットマップ T 2 を、NASノード 1 0 0 と共有する。従って、NASノード 1 0 0 のメモリ資源を効率的に使用できる。

10

20

30

40

50

【実施例 2】

【0137】

図 13, 図 14 に基づいて、第 2 実施例を説明する。本実施例では、マイグレート対象ビットマップ T3 を複数のエリアに分割し、各エリア毎にセグメント単位でのデータ移動を行う。本実施例及び後述の各実施例は、第 1 実施例の変形例に相当する。

【0138】

図 13 は、マイグレート対象ビットマップ T3 について、複数のセグメント範囲を設定する様子を示す説明図である。図 13 の上側に示すように、マイグレート対象ビットマップ T3 に、セグメント範囲 AS1 ~ AS3 を設定し、各セグメント範囲 AS1 ~ AS3 毎に、セグメント単位でのデータマイグレートを実行する。

10

【0139】

図 13 の下側には、セグメント範囲のサイズを決定するための処理を示すフローチャートが示されている。この処理は、図 14 中の S71 の詳細を示す。まず、マイグレートツール 113 は、クライアント 20 からファイルシステム 111 への現在のアクセス要求の負荷を検出する (S711)。

【0140】

マイグレートツール 113 は、検出された負荷に基づいて、閾値セグメント数を決定する (S712)。閾値セグメント数とは、セグメント範囲 AS1 ~ AS3 に含まれるマイグレート対象セグメントの上限値である。マイグレート対象のセグメントが多いほど、そのセグメント範囲のデータマイグレーションが完了するまでの時間が長くなるため、各セグメント範囲 AS1 ~ AS3 に含まれるマイグレート対象セグメントの数を制限する。

20

【0141】

そして、マイグレートツール 113 は、各セグメント範囲 AS1 ~ AS3 内のマイグレート対象セグメントの数が、閾値セグメント数以下となるように、各セグメント範囲 AS1 ~ AS3 を決定する (S713)。なお、上記の例では、NAS ノード 100 の現在の負荷状態に基づいて、閾値セグメント数を算出するものと述べたが、これに限らず、閾値セグメント数を固定値としてもよいし、ユーザによって指定可能としてもよい。また、セグメント範囲の数は、3 個に限られず、2 個以上であればよい。

【0142】

図 14 は、本実施例によるデータマイグレーション処理を示す。この処理は、例えば、一日に一回実行させることができる。前記同様に、データマイグレーションの実施に先立って、マイグレート対象ビットマップ T3 が生成される (S70)。マイグレートツール 113 は、図 13 と共に述べたように、マイグレート対象ビットマップ T3 に複数のセグメント範囲 AS1 ~ AS3 を設定する (S71)。

30

【0143】

マイグレートツール 113 が最初のセグメント範囲 AS1 を一つ指定すると (S72)、ファイルシステム 111 は、i ノード管理領域 330A を参照し (S73)、指定されたセグメント範囲 AS1 に含まれるマイグレート対象セグメントにファイルデータが記憶されているファイル群を特定する (S74)。ファイルシステム 111 は、その特定されたファイル群について、更新排他ロックを設定する (S75)。

40

【0144】

マイグレートツール 113 は、ストレージコントローラ 200 に向けて、データのマイグレート実行を指示する (S76)。この指示には、ストレージコントローラ 200 によって認識されている FC 領域上の LUN 及びセグメント範囲 AS1 が含まれる。マイグレートツール 113 は、セグメント範囲 AS1 の先頭セグメント番号及び最終セグメント番号を示すことにより、または、先頭セグメント番号及びセグメント数を示すことにより、セグメント範囲 AS1 をストレージコントローラ 200 に通知することができる。

【0145】

前記実施例と同様に、ストレージコントローラ 200 は、マイグレートツール 113 からの指示を受領すると、LVM 定義テーブル T1 を参照し (S77)、指定された FC 領域に

50

対応するATA領域を特定する（S78）。ストレージコントローラ200は、指示されたセグメント範囲AS1に含まれるマイグレート対象セグメントの中から、データを移動させるセグメントを降順で決定する（S79）。

【0146】

ストレージコントローラ200は、ATA領域において、次の空きセグメントの番号を確保し（S80）、移動元のセグメントのデータを確保した空きセグメントに記憶させる（S81）。ストレージコントローラ200は、このデータがコピーされたセグメントの番号を、新セグメント番号リストに登録する（S82）。新セグメント番号リストは、データ移動先のセグメントの番号を特定するための情報である。

【0147】

ストレージコントローラ200は、指定されたセグメント範囲AS1について、各マイグレート対象セグメントのデータ移動が完了したか否かを判定する（S83）。指定されたセグメント範囲AS1内の全てのマイグレート対象セグメントについて、ATA領域への移動が完了するまで、S79～S82のステップが繰り返され、新セグメント番号リストが更新されていく。指定されたセグメント範囲AS1についてデータの移動が完了すると、ストレージコントローラ200は、新セグメント番号リストをマイグレートツール113に通知する（S84）。

【0148】

マイグレートツール113は、ストレージコントローラ200から受領した新セグメント番号リストをファイルシステム111に通知し、移動されたファイルに関する情報の切替を指示する（S85）。

【0149】

ファイルシステム111は、指定されたセグメント範囲AS1に含まれるマイグレート対象セグメントを降順で選択し（S86）、選択されたセグメントに関するATAフラグ及びセグメント番号をそれぞれ更新する（S87）。移動前のセグメント番号は、新セグメント番号リストに登録された移動先セグメント番号に書き換えられる。また、ファイルシステム111は、ATAフラグを"1"に更新したセグメントを、空きセグメントとして登録する。

【0150】

ファイルシステム111は、ATAフラグ及びセグメント番号を更新したセグメントについて、更新排他ロックを解除する（S88）。ファイルシステム111は、指定されたセグメント範囲AS1内の全ての更新排他ロックが解除されるまで、S86～S88のステップを繰り返す。

【0151】

ファイルシステム111が、セグメント範囲AS1についての処理を完了すると、マイグレートツール113は、S72で選択されたセグメント範囲AS1内のデータ移動が完了したことを確認する（S90）。全てのセグメント範囲についてデータ移動が完了するまで、S72～S90のステップが繰り返される。

【0152】

このように構成される本実施例も前記第1実施例と同様の効果を奏する。これに加えて、本実施例では、マイグレート対象ビットマップT3に複数のセグメント範囲AS1～AS3を設定し、セグメント範囲毎にデータを移動させる構成とした。従って、複数のマイグレート対象セグメントを一括して処理することができる。これにより、FC領域内における空きセグメントの数を速やかに増加させることができる。

【0153】

また、複数のマイグレート対象セグメントを一括して扱うため、NASノード100とストレージコントローラ200との間の情報交換頻度を低減し、NASノード100の負荷をより一層低減することができる。

【0154】

さらに、複数のセグメントのデータを一括して移動させるため、データマイグレーション

10

20

30

40

50

ンに要する時間を低減することができる。

【0155】

また、本実施例では、各セグメント範囲AS1～AS3に含まれるマイグレート対象セグメントの数を調整するため、各セグメント範囲のデータ移動が完了するまでの時間を、比較的短くすることができる。

【実施例3】

【0156】

図15～図17に基づいて、第3実施例を説明する。本実施例では、仮想ボリューム340の構成を変更する場合について説明する。LVM定義テーブルT1の説明で述べたように、仮想ボリューム340の構成は変更可能である。

10

【0157】

図15は、NASノード100のソフトウェア構成及び階層記憶の構成を模式的に示す説明図である。仮想ボリューム340のFC領域には、それぞれFCディスク310に由来する複数の論理ボリューム330(LU0, LU2)が対応付けられている。仮想ボリューム340のATA領域には、それぞれATAディスク311に由来する複数の論理ボリューム331(LU1, LU3)が対応付けられている。

【0158】

図16は、FC領域及びATA領域を複数の論理ボリュームから構成する場合のセグメントの管理状態を模式的に示す説明図である。図16(a)は、FC領域を一つの論理ボリューム(LU0)から構成し、ATA領域を一つの論理ボリューム331(LU1)から構成する場合を示す。この場合、FC領域の先頭セグメント番号は、論理ボリューム330(LU0)の先頭セグメント番号に一致し、FC領域の終端セグメント番号は、論理ボリューム330(LU0)の終端セグメント番号に一致する。同様に、ATA領域の先頭セグメント番号は、論理ボリューム331(LU1)の先頭セグメント番号に一致し、ATA領域の終端セグメント番号は、論理ボリューム331の終端セグメント番号に一致する。

20

【0159】

図16(b)は、ATA領域に、別の論理ボリューム331(LU3)を追加した場合を示す。ATA領域の先頭セグメント番号に変化はない。ATA領域の終端セグメント番号は、追加された論理ボリューム331(LU3)の終端セグメント番号となる。ここで、留意すべきは、ATA領域の前側に位置する論理ボリューム331(LU1)の終端セグメント番号に続けて、追加された論理ボリューム331(LU3)の先頭セグメント番号が位置する点である。即ち、ATA領域内のセグメントは、セグメント番号が連続するように管理されている。

30

【0160】

同様に、FC領域に別の論理ボリューム330(LU2)を追加する場合も、FC領域内でセグメント番号が連続するようにして管理される。なお、各ビットマップT2, T3, T4は、FC領域とATA領域のそれぞれに対応するように管理される。

【0161】

本実施例では、FC領域のサイズには限界値が存在する。即ち、FC領域は、ATA領域の先頭アドレスとして設定されたSA2まで、サイズを拡張することができる。仮に、SA2の値が2TBであるとすると、現在のFC領域に1TBのボリュームを割り当てている場合は、さらに1TBのボリュームを追加可能である。SA2の値をより大きく設定すれば、その分だけFC領域のサイズを拡張することができる。なお、ATA領域は、記憶制御装置10が管理可能な最大サイズまで拡張可能である。

40

【0162】

図17は、仮想ボリューム340の容量を増加するための処理を示すフローチャートである。この処理は、例えば、管理端末30から指示して実行させることができる。まず、記憶制御装置10は、FC領域またはATA領域のいずれに対して容量の増加が指示されたのかを判定する(S111)。

【0163】

FC領域の容量増加が指示された場合、記憶制御装置10は、FC領域に容量を追加可能で

50

あるか否かを判定する (S 1 1 2)。FC領域のサイズが限界値 (S A 2) に達している場合 (S112:NO)、記憶制御装置 1 0 は、容量の増加ができない旨を管理端末 3 0 に通知する (S 1 1 3)。

【 0 1 6 4 】

FC領域に容量を追加可能な場合 (S112:YES)、記憶制御装置 1 0 は、新たな容量をFC領域に追加し、FC領域内のセグメント番号を連続させ (S 1 1 4)、LVM定義テーブル T 1 を更新させる (S 1 1 5)。

【 0 1 6 5 】

ATA領域の容量追加が指示された場合、記憶制御装置 1 0 は、新たな容量をATA領域に追加し、ATA領域内のセグメント番号を連続させ (S 1 1 6)、LVM定義テーブル T 1 を更新させる (S 1 1 7)。

【 0 1 6 6 】

このように構成される本実施例も前記第 1 実施例と同様の効果を奏する。これに加えて、本実施例では、クライアント 2 0 による仮想ボリューム 3 4 0 の利用状況等に応じて、仮想ボリューム 3 4 0 の構成を変更することができ、使い勝手が向上する。

【 0 1 6 7 】

なお、本発明は、上述した実施の形態に限定されない。当業者であれば、本発明の範囲内で、種々の追加や変更等を行うことができる。例えば、各実施例を適宜組み合わせることができる。

【 図面の簡単な説明 】

【 0 1 6 8 】

【 図 1 】 本発明の実施形態の概念を示す説明図である。

【 図 2 】 記憶制御装置のハードウェア構成を示す説明図である。

【 図 3 】 NASノードのソフトウェア構成及び階層記憶の構成を示す説明図である。

【 図 4 】 仮想ボリュームの構成を示す説明図である。

【 図 5 】 i ノード管理領域の構成を示す説明図である。

【 図 6 】 LVM定義テーブルの構成を示す説明図である。

【 図 7 】 非更新ビットマップを生成する処理を示すフローチャートである。

【 図 8 】 更新ビットマップから非更新ビットマップを経てマイグレート対象ビットマップを生成する様子を示す説明図である。

【 図 9 】 マイグレーションポリシー設定テーブルを示す説明図である。

【 図 1 0 】 ファイル追加処理を示すフローチャートである。

【 図 1 1 】 ファイル更新処理を示すフローチャートである。

【 図 1 2 】 データマイグレーション処理を示すフローチャートである。

【 図 1 3 】 第 2 実施例に係り、(a) は、マイグレート対象ビットマップを複数のセグメント範囲毎に処理する様子を、(b) はセグメント範囲を設定するためのフローチャートを、それぞれ示す。

【 図 1 4 】 データマイグレーション処理を示すフローチャートである。

【 図 1 5 】 本発明の第 3 実施例に係り、仮想ボリュームの各領域を複数の論理ボリュームから構成可能であることを示す階層記憶の説明図である。

【 図 1 6 】 領域の容量を追加する場合のセグメント番号の様子を示す説明図である。

【 図 1 7 】 領域に容量を追加する場合の処理を示すフローチャートである。

【 符号の説明 】

【 0 1 6 9 】

1 ... 記憶制御装置、 2 ... 制御部、 2 A , 3 A ... 非更新ビットマップ、 2 B , 3 B ... マイグレート対象ビットマップ、 2 C ... ファイル管理部、 2 D ... マイグレート制御部、 2 E ... ボリューム管理部、 3 ... ストレージ制御部、 3 C ... 入出力処理部、 3 D ... マイグレート実行部、 3 E ... 更新管理部、 4 ... 記憶部、 4 A , 4 B ... 論理ボリューム (論理的記憶デバイス)、 4 C , 4 D ... ディスクドライブ、 4 E , 4 F ... 物理的記憶デバイス、 5 ... 仮想ボリューム、 5 A ... 高速領域、 5 B ... 低速領域、 6 ... クライアントマシン、 1 0 ... 記憶制御装

10

20

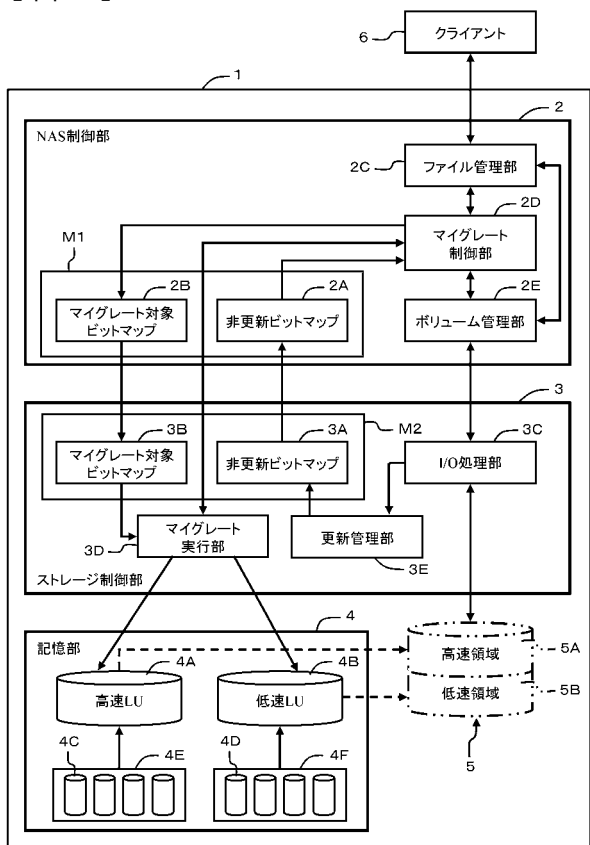
30

40

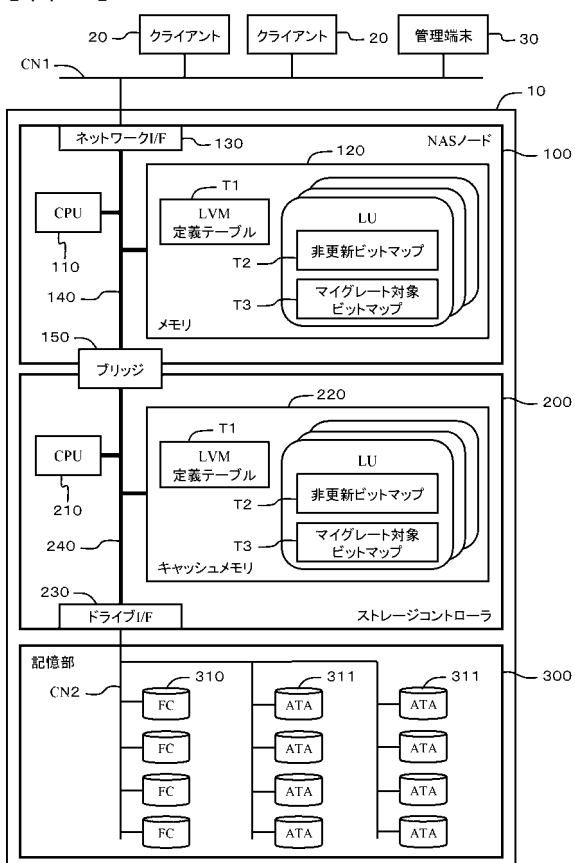
50

置、20...クライアントマシン、30...管理端末、100...NASノード、110...プロセッサ、111...ファイルシステム、112...論理ボリュームマネージャ(LVM)、113...マイグレートツール、113A...マイグレーションポリシー、120...メモリ、130...ネットワークインターフェース部、140...内部バス、150...ブリッジ回路、200...ストレージコントローラ、210...プロセッサ、220...キャッシュメモリ、230...ドライブインターフェース部、240...内部バス、300...記憶部、310, 311...ディスクドライブ、320, 321...物理的記憶デバイス、330, 331...論理ボリューム(論理的記憶デバイス)、330A...iノード管理領域、330B, 331A...ファイルデータ領域、340...仮想ボリューム、AS1~AS3...セグメント範囲、M1, M2...メモリ、T1...LVM定義テーブル、T2...非更新ビットマップ、T3...マイグレート対象ビットマップ、T4...更新ビットマップ、T5...マイグレーションポリシー設定テーブル

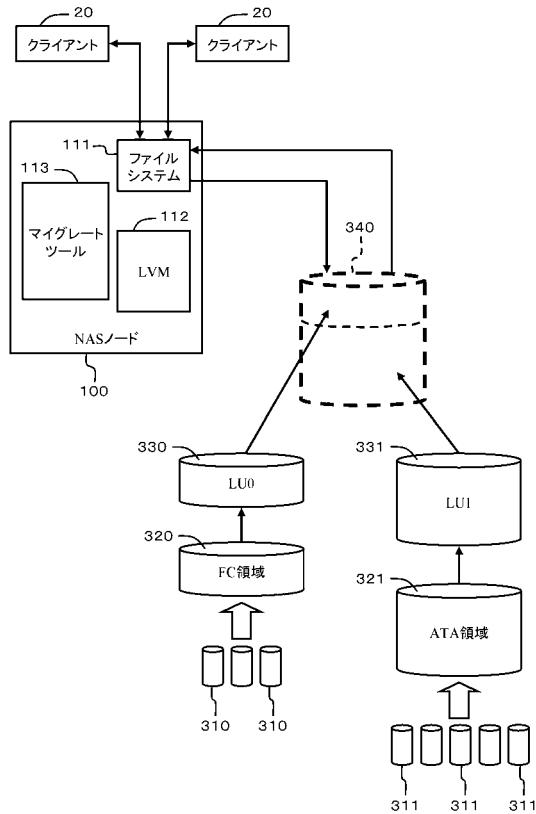
【図1】



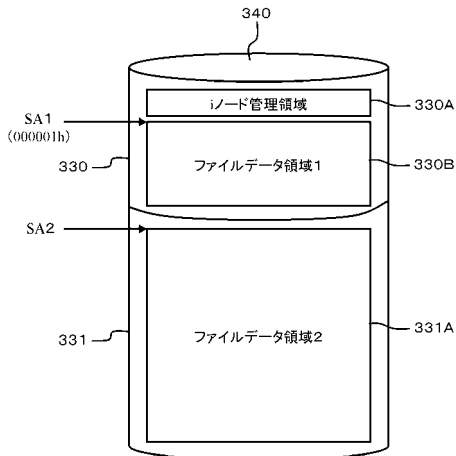
【図2】



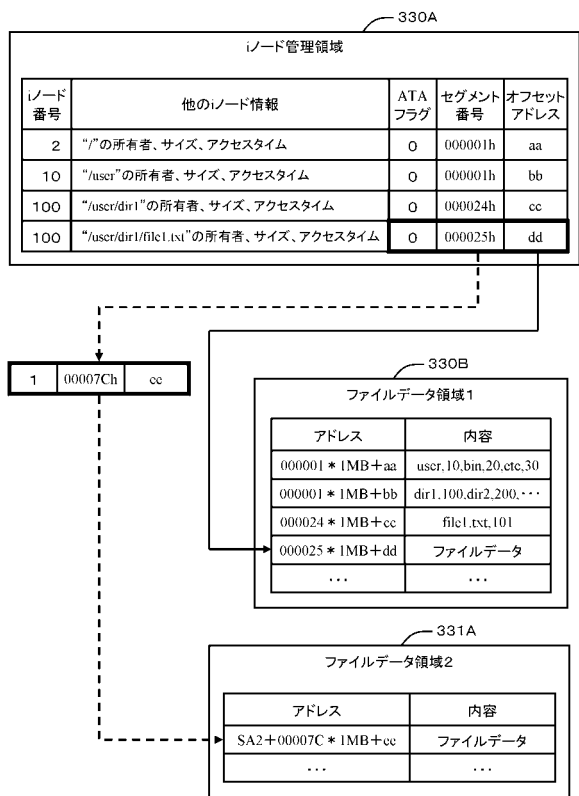
【 図 3 】



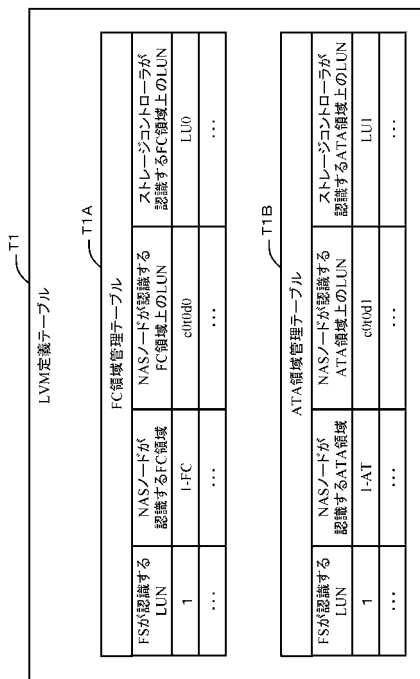
【 図 4 】



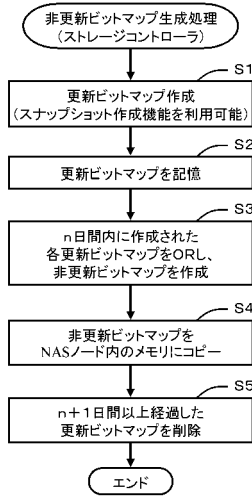
【 図 5 】



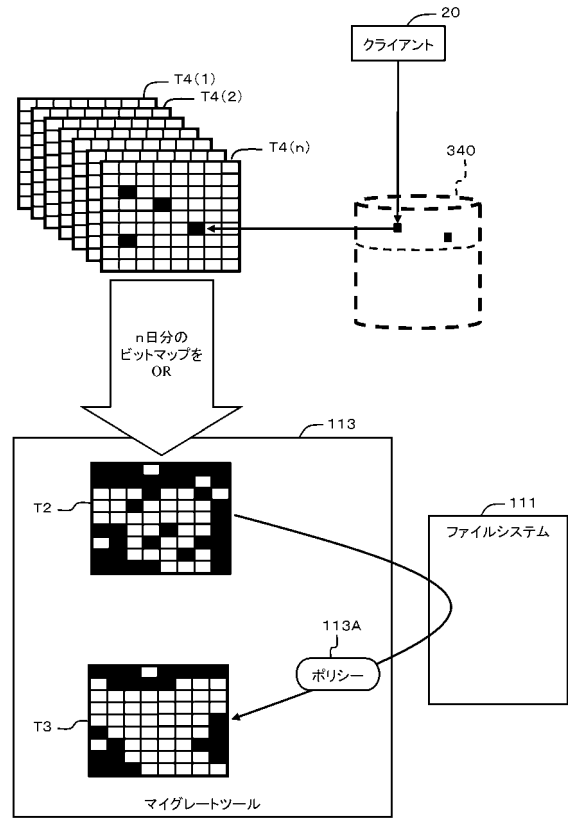
【 図 6 】



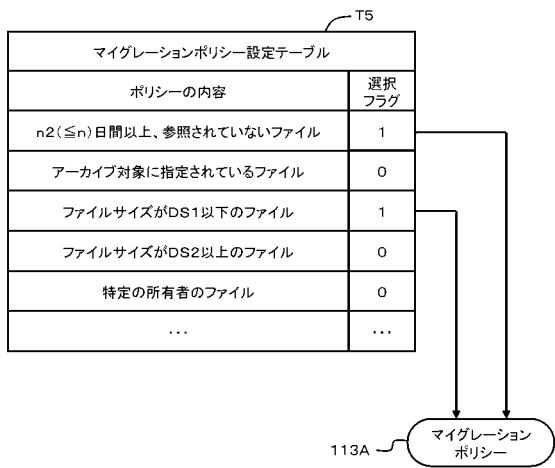
【 図 7 】



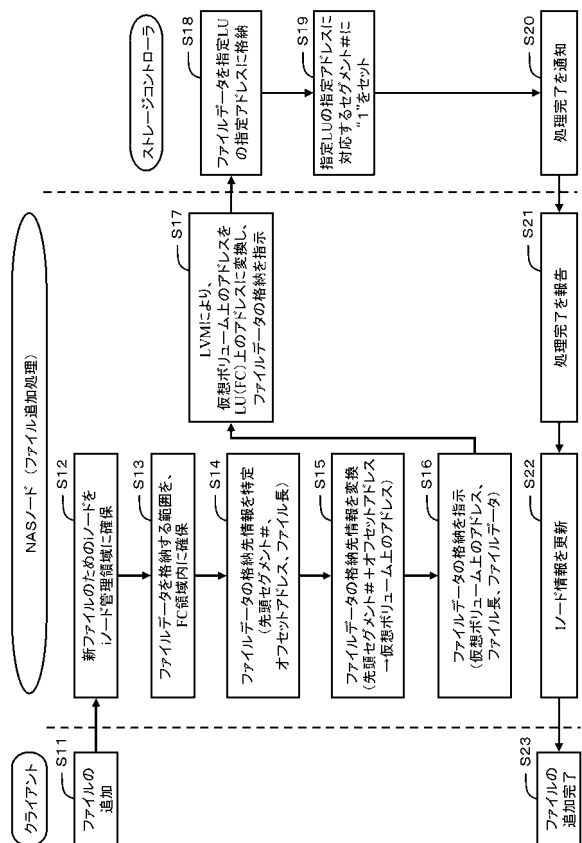
【 図 8 】



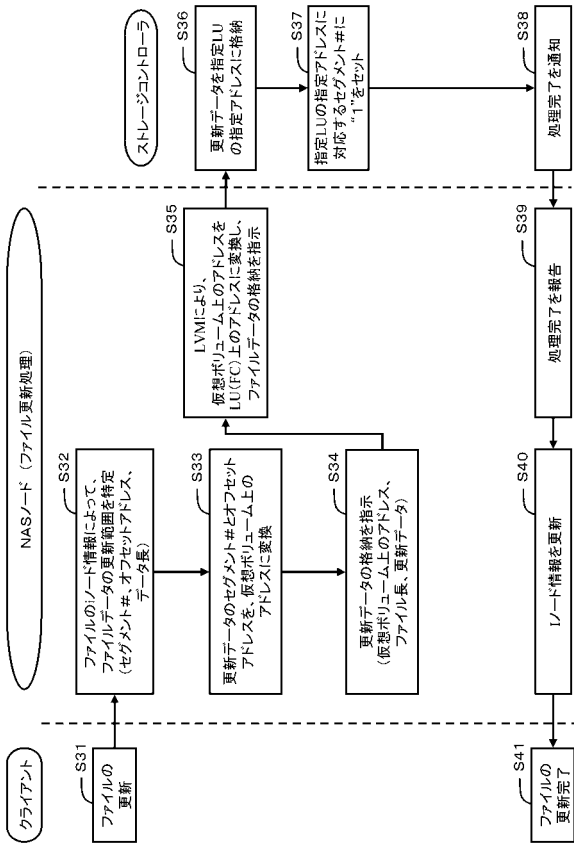
【 図 9 】



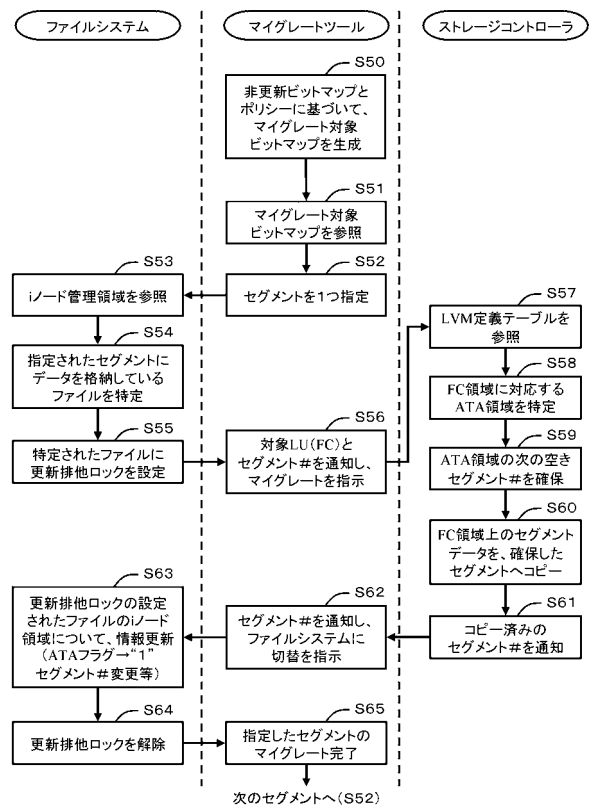
【 図 10 】



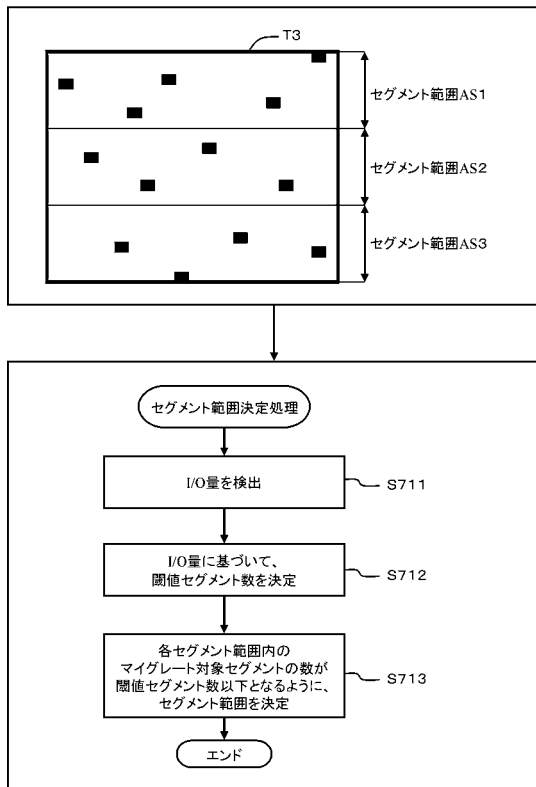
【図 1 1】



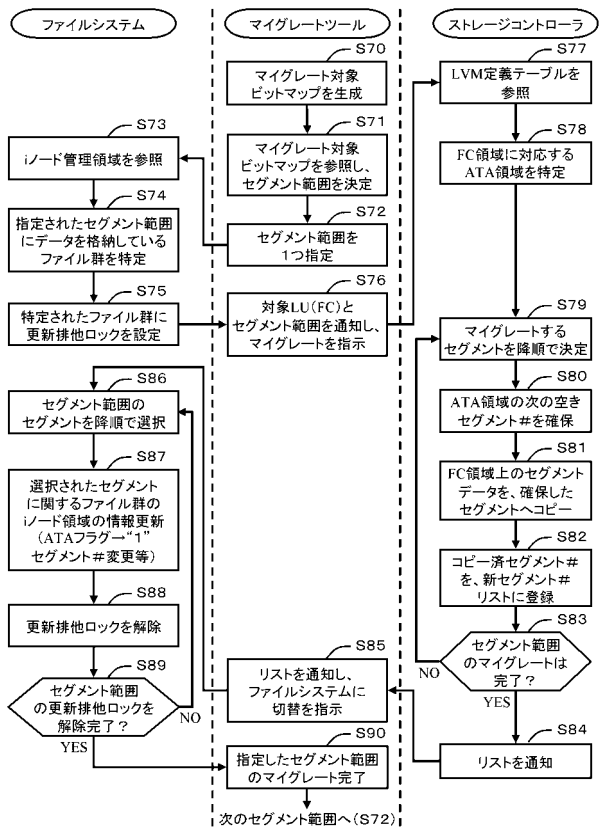
【図 1 2】



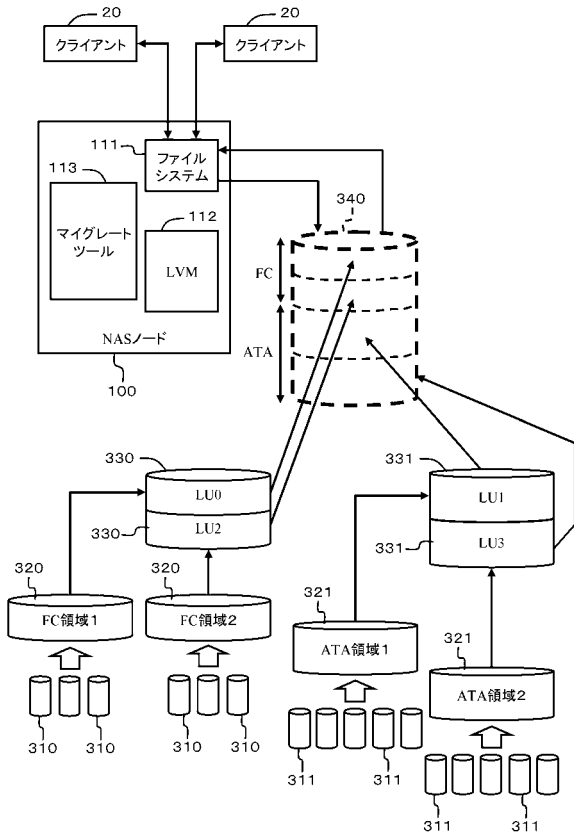
【図 1 3】



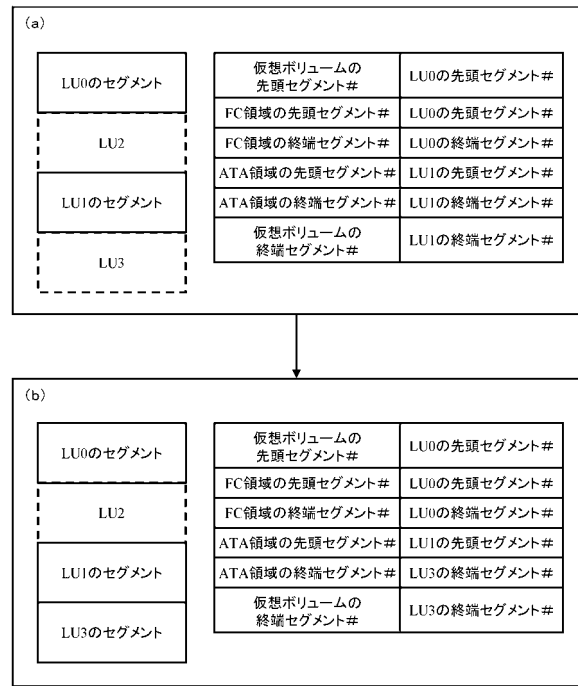
【図 1 4】



【 図 1 5 】



【 図 1 6 】



【 図 1 7 】

