

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第4704161号
(P4704161)

(45) 発行日 平成23年6月15日 (2011.6.15)

(24) 登録日 平成23年3月18日 (2011.3.18)

(51) Int.Cl.

F I

G O 6 F 12/00 (2006.01)

G O 6 F 12/00 5 O 1 H

G O 6 F 3/06 (2006.01)

G O 6 F 3/06 3 O 1 Z

G O 6 F 12/00 5 2 O P

請求項の数 14 (全 34 頁)

(21) 出願番号 特願2005-265093 (P2005-265093)
 (22) 出願日 平成17年9月13日 (2005.9.13)
 (65) 公開番号 特開2007-79774 (P2007-79774A)
 (43) 公開日 平成19年3月29日 (2007.3.29)
 審査請求日 平成20年4月25日 (2008.4.25)

(73) 特許権者 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 110000028
 特許業務法人明成国際特許事務所
 (72) 発明者 野中 裕介
 神奈川県川崎市麻生区王禅寺1099番地
 株式会社日立製作所 システム開発研究
 所内
 (72) 発明者 山本 彰
 神奈川県川崎市麻生区王禅寺1099番地
 株式会社日立製作所 システム開発研究
 所内

最終頁に続く

(54) 【発明の名称】 ファイルシステムの構築方法

(57) 【特許請求の範囲】

【請求項1】

少なくとも1つの第1種のストレージ媒体と少なくとも1つの第2種のストレージ媒体とを備える記憶装置に接続されたファイルサーバであって、

クライアントコンピュータと通信するためのネットワークインターフェースと、

前記記憶装置に格納されたファイルを管理するためのファイル管理部と、
 を備え、

前記ファイルサーバは、前記少なくとも1つの第1種のストレージ媒体と前記少なくとも1つの第2種のストレージ媒体とによって構成された論理装置内にファイルシステムを構成し、

前記ネットワークインターフェースは、前記ファイルシステムのファイルの書き込み要求を前記クライアントコンピュータから受信し、前記ファイルは、データとメタデータとを含み、

前記記憶装置は、ストレージ媒体の種類に応じて前記論理装置内のアドレスレンジを管理するテーブルを備え、

前記ファイル管理部は、ファイルの書き込み要求を前記クライアントコンピュータから受信した時に、前記テーブルによって管理されている前記アドレスレンジを考慮して、データ格納のためにストレージ媒体の種類を選択し、前記ファイルの前記メタデータおよび前記データを前記少なくとも1つの第1種のストレージ媒体または前記少なくとも1つの第2種のストレージ媒体の内の1つに書き込むことで、前記メタデータおよびデータが、

前記少なくとも１つの第１種のストレージ媒体と前記少なくとも１つの第２種のストレージ媒体とに跨って格納されないように、前記論理装置に前記メタデータおよびデータを格納する、ファイルサーバ。

【請求項２】

請求項１に記載のファイルサーバであって、前記ファイルサーバが、ファイルの書き込み要求を前記クライアントコンピュータから受信した時に、前記ファイル管理部は、データ格納のためにストレージ媒体の種類を選択し、第１種が選択された場合、前記メタデータおよび前記データの両方が、前記少なくとも１つの第１種のストレージ媒体に格納される、ファイルサーバ。

【請求項３】

請求項１に記載のファイルサーバであって、データ格納のために前記少なくとも１つの第２種のストレージ媒体を選択した場合、前記メタデータおよび前記データは、それぞれ、前記少なくとも１つの第１種のストレージ媒体および前記少なくとも１つの第２種のストレージ媒体に格納される、ファイルサーバ。

【請求項４】

請求項３に記載のファイルサーバであって、前記ファイル管理部は、前記少なくとも１つの第１種のストレージ媒体から前記少なくとも１つの第２種のストレージ媒体に前記データを移動し、前記少なくとも１つの第１種のストレージ媒体にすでに格納されているメタデータを、前記少なくとも１つの第２種のストレージ媒体の実データの位置を示すように変更する、ファイルサーバ。

【請求項５】

少なくとも１つの第１種のストレージ媒体によって構成された第１の記憶領域と、少なくとも１つの第２種のストレージ媒体によって構成された第２の記憶領域とを備える記憶装置に接続されたファイルサーバであって、

クライアントコンピュータと通信するためのネットワークインターフェースと、

前記記憶装置に格納されたファイルを管理するためのファイル管理部と、
を備え、

前記ファイルサーバは、前記第１の記憶領域と前記第２の記憶領域とによって構成された論理装置内にファイルシステムを構成し、

前記ネットワークインターフェースは、前記ファイルシステムのファイルの書き込み要求を前記クライアントコンピュータから受信し、前記ファイルは、データとメタデータとを含み、

前記記憶装置は、ストレージ媒体の種類に応じて、前記論理装置内で、前記第１の記憶領域によって占められた第１のアドレスレンジと、前記第２の記憶領域によって占められた第２のアドレスレンジとを管理するテーブルを備え、

前記ファイル管理部は、前記ファイルサーバが、ファイルの書き込み要求を前記クライアントコンピュータから受信した時に、前記テーブルによって管理されている前記アドレスレンジを考慮して、データ格納のためにストレージ媒体の種類を選択し、前記ファイルの前記データを格納するために、前記少なくとも１つの第１の記憶領域または前記第２の記憶領域の内の１つにおける格納位置を選択し、前記選択された格納位置に前記ファイルの前記データおよび前記メタデータを格納することで、前記メタデータおよび前記データが、前記少なくとも１つの第１種のストレージ媒体と前記少なくとも１つの第２種のストレージ媒体とに跨らないように、前記論理装置に前記メタデータおよび前記データを格納する、ファイルサーバ。

【請求項６】

請求項５に記載のファイルサーバであって、前記ファイルサーバが、ファイルの前記書き込み要求を前記クライアントコンピュータから受信した時に、前記ファイル管理部は、データ格納のためにストレージ媒体の種類を選択し、第１種が選択された場合、メタデータおよびデータの両方が、前記少なくとも１つの第１種のストレージ媒体に格納される、ファイルサーバ。

10

20

30

40

50

【請求項 7】

請求項 5 に記載のファイルサーバであって、前記ファイル管理部は、前記少なくとも 1 つの第 1 種のストレージ媒体から前記少なくとも 1 つの第 2 種のストレージ媒体に前記データを移動する時に、前記少なくとも 1 つの第 2 種の記憶領域にメタデータを作成する、ファイルサーバ。

【請求項 8】

少なくとも 1 つの第 1 種のストレージ媒体と少なくとも 1 つの第 2 種のストレージ媒体とを備える記憶装置に接続されたファイルサーバにおける方法であって、

ネットワークインターフェースを介してクライアントコンピュータと通信する工程と、

前記記憶装置に格納されたファイルをファイル管理部によって管理する工程と、

前記少なくとも 1 つの第 1 種のストレージ媒体と前記少なくとも 1 つの第 2 種のストレージ媒体とによって構成された論理装置内にファイルシステムを構成する工程と、

前記ファイルシステムのファイルの書き込み要求を前記クライアントコンピュータから受信する工程であって、前記ファイルは、データとメタデータとを含む、工程と、

ストレージ媒体の種類に応じて前記論理装置内のアドレスレンジを管理するテーブルを保持する工程と、

ファイルの書き込み要求を前記クライアントコンピュータから受信した時に、前記テーブルによって管理されている前記アドレスレンジを考慮してデータ格納のためにストレージ媒体の種類を選択し、前記ファイルの前記メタデータおよび前記データを前記ファイルの前記データを前記少なくとも 1 つの第 1 種のストレージ媒体または前記少なくとも 1 つの第 2 種のストレージ媒体の内の 1 つに書き込むことで、前記メタデータおよびデータが、前記少なくとも 1 つの第 1 種のストレージ媒体と前記少なくとも 1 つの第 2 種のストレージ媒体とに跨って格納されないように、前記論理装置に前記メタデータおよびデータを格納する工程と、

を備える、方法。

【請求項 9】

請求項 8 に記載の方法であって、前記ファイルサーバが、ファイルの書き込み要求を前記クライアントコンピュータから受信した時に、前記ファイル管理部は、データ格納のためにストレージ媒体の種類を選択し、第 1 種が選択された場合、前記メタデータおよび前記データの両方が、前記少なくとも 1 つの第 1 種のストレージ媒体に格納される、方法。

【請求項 10】

請求項 8 に記載の方法であって、データ格納のために前記少なくとも 1 つの第 2 種のストレージ媒体を選択した場合、前記メタデータおよび前記データは、それぞれ、前記少なくとも 1 つの第 1 種のストレージ媒体および前記少なくとも 1 つの第 2 種のストレージ媒体に格納される、方法。

【請求項 11】

請求項 8 に記載の方法であって、前記少なくとも 1 つの第 1 種のストレージ媒体から前記少なくとも 1 つの第 2 種のストレージ媒体に前記データを移動する時に、前記少なくとも 1 つの第 1 種のストレージ媒体にすでに格納されているメタデータを、前記少なくとも 1 つの第 2 種のストレージ媒体の実データの位置を示すように変更する工程を備える、方法。

【請求項 12】

少なくとも 1 つの第 1 種のストレージ媒体によって構成された第 1 の記憶領域と、少なくとも 1 つの第 2 種のストレージ媒体によって構成された第 2 の記憶領域とを備える記憶装置に接続されたファイルサーバにおける方法であって、

ネットワークインターフェースを介してクライアントコンピュータと通信する工程と、

前記記憶装置に格納されたファイルをファイル管理部によって管理する工程と、

前記第 1 の記憶領域と前記第 2 の記憶領域とによって構成された論理装置内にファイルシステムを構成する工程と、

前記ファイルシステムのファイルの書き込み要求を前記クライアントコンピュータから

受信する工程であって、前記ファイルは、データとメタデータとを含む、工程と、

ストレージ媒体の種類に応じて、前記論理装置内で、前記第1の記憶領域によって占められた第1のアドレスレンジと、前記第2の記憶領域によって占められた第2のアドレスレンジとを管理するテーブルを保持する工程と、

ファイルの書き込み要求を前記クライアントコンピュータから受信した時に、前記テーブルによって管理されている前記アドレスレンジを考慮してデータ格納のためにストレージ媒体の種類を選択し、前記ファイルの前記データを格納するために、前記少なくとも1つの第1の記憶領域または前記第2の記憶領域の内の1つにおける格納位置を選択し、前記選択された格納位置に前記ファイルの前記データおよび前記メタデータを格納することで、前記メタデータおよびデータが、前記少なくとも1つの第1種のストレージ媒体と前記少なくとも1つの第2種のストレージ媒体とに跨って格納されないように、前記論理装置に前記メタデータおよびデータを格納する工程と、
を備える、方法。

【請求項13】

請求項12に記載の方法であって、ファイルの前記書き込み要求を前記クライアントコンピュータから受信した時に、データ格納のためにストレージ媒体の種類を選択する工程を備え、第1種が選択された場合、メタデータおよびデータの両方が、前記少なくとも1つの第1種のストレージ媒体に格納される、方法。

【請求項14】

請求項12に記載の方法であって、前記少なくとも1つの第1種のストレージ媒体から前記少なくとも1つの第2種のストレージ媒体に前記データを移動する時に、前記少なくとも1つの第2種の記憶領域にメタデータを作成する工程を備える、方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、ストレージシステムの技術に関する。

【背景技術】

【0002】

NAS (Network Attached Storage) と呼ばれるネットワークに直接接続する形式のストレージシステムが多く利用されている。このストレージシステムを利用すれば、ネットワークを介して、複数のクライアント間でファイルシステムを共有可能であり、複数のクライアントは同じファイルにアクセス可能である。

【0003】

近年、データを、重要度などに基づくデータの価値に応じて、最適なコストのストレージ媒体に格納することによって、効果的なデータの活用と、効率的なストレージの投資と、の実現を図ることが試みられている。このような考え方は、情報ライフサイクル管理 (ILM (information lifecycle management)) と呼ばれている。例えば、価値の比較的高いデータは、高価なFC (Fibre Channel) 対応のハードディスクに格納され、価値の比較的低いデータは、安価なATA対応のハードディスクに格納される。なお、FC対応のハードディスクは、ATA対応のハードディスクよりも、信頼性が高いと共に、アクセスが速いという特徴を有している。

【0004】

FC対応のハードディスクやATA対応のハードディスクなどの複数種のストレージ媒体を用いてILMを実現する場合には、通常、ストレージ媒体の種類毎に論理ユニットが設定され、論理ユニット毎にファイルシステムが構築される。すなわち、従来では、ILMを実現する場合には、ストレージ媒体の種類毎に複数のファイルシステムが構築されていた。

【0005】

【特許文献1】米国特許出願公開第2004/0133540号公報

【特許文献2】特表2005-512171号公報

10

20

30

40

50

【発明の開示】

【発明が解決しようとする課題】

【0006】

既存の技術を適用すれば、複数種のストレージ媒体に1つの論理ユニットを設定し、該論理ユニット上に1つのファイルシステムを構築することも可能である。しかしながら、この場合には、ILMの適切な運用を図ること、すなわち、複数種のストレージ媒体を効率よく利用することが困難となってしまう。

【0007】

この発明は、従来技術における上述の課題を解決するためになされたものであり、複数種のストレージ媒体上に1つのファイルシステムを構築する場合にも、複数種のストレージ媒体を効率よく利用することのできる技術を提供することを目的とする。

10

【課題を解決するための手段】

【0008】

上述の課題の少なくとも一部を解決するため、本発明の第1の装置は、第1種のストレージ媒体と第2種のストレージ媒体とを管理するための管理装置であって、

前記第1種のストレージ媒体の少なくとも一部で構成された第1の記憶領域と、前記第2種のストレージ媒体の少なくとも一部で構成された第2の記憶領域と、を含む統合記憶領域内のファイルを管理するためのファイル管理部を備え、

前記ファイル管理部は、

前記第1の記憶領域の範囲と前記第2の記憶領域の範囲とを考慮して、前記統合記憶領域上に1つのファイルシステムを構築することを特徴とする。

20

【0009】

第1の記憶領域の範囲と第2の記憶領域の範囲とを考慮せずに、統合記憶領域上にファイルシステムが構築される場合には、ファイルが第1の記憶領域と第2の記憶領域とに跨って格納されてしまう恐れがある。しかしながら、この装置では、第1の記憶領域の範囲と第2の記憶領域の範囲とを考慮して、統合記憶領域上にファイルシステムが構築される。このため、ファイルが第1の記憶領域と第2の記憶領域とに跨って格納されずに済むと共に、1つのファイルシステム内で第1の記憶領域と第2の記憶領域とにファイル単位でファイルが格納されるため、第1種のストレージ媒体と第2種のストレージ媒体とを効率よく利用することが可能となる。

30

【0010】

上記の装置において、

前記ファイル管理部は、

前記統合記憶領域内に、前記第1の記憶領域と前記第2の記憶領域とに跨らない複数のグループ領域を設定することが好ましい。

【0011】

このように、第1の記憶領域と前記第2の記憶領域とに跨らないグループ領域を予め設定しておけば、ファイルが第1の記憶領域と第2の記憶領域とに跨って格納されずに済む。

【0012】

上記の装置において、

前記ファイル管理部は、

前記統合記憶領域内の前記第1の記憶領域に、メタデータが格納される第1種のグループ領域を設定し、

前記統合記憶領域内の前記第2の記憶領域に、メタデータが格納されない第2種のグループ領域を設定するようにしてもよい。

40

【0013】

ファイルにアクセスする際には、メタデータが検索される。上記のようにすれば、メタデータを検索する際に、第1の記憶領域のみを検索すれば済み、この結果、目的のファイルに迅速にアクセスすることが可能となる。

50

【 0 0 1 4 】

なお、この効果は、第 1 種のストレージ媒体が第 2 種のストレージ媒体よりも高速にアクセス可能な場合に顕著となる。

【 0 0 1 5 】

上記の装置において、

前記ファイル管理部は、

第 1 のファイルを前記統合記憶領域内の前記第 1 の記憶領域に格納する場合には、前記第 1 のファイルの実データを前記第 1 の記憶領域に書き込むと共に、前記第 1 のファイルに関するメタデータを前記第 1 の記憶領域に作成し、

第 2 のファイルを前記統合記憶領域内の前記第 2 の記憶領域に格納する場合には、前記第 2 のファイルの実データを前記第 2 の記憶領域に書き込むと共に、前記第 2 のファイルに関するメタデータを前記第 1 の記憶領域に作成するようにしてもよい。

10

【 0 0 1 6 】

上記の装置において、

前記ファイル管理部は、

第 1 のファイルを前記統合記憶領域内の前記第 1 の記憶領域から前記第 2 の記憶領域に移動させる場合には、前記第 1 のファイルの実データを前記第 2 の記憶領域に書き込むと共に、前記第 1 の記憶領域に既に格納されている前記第 1 のファイルに関するメタデータを、前記第 2 の記憶領域に書き込まれた前記第 1 のファイルの実データの位置を示すように、変更し、

20

第 2 のファイルを前記統合記憶領域内の前記第 2 の記憶領域から前記第 1 の記憶領域に移動させる場合には、前記第 2 のファイルの実データを前記第 1 の記憶領域に書き込むと共に、前記第 1 の記憶領域に既に格納されている前記第 2 のファイルに関するメタデータを、前記第 1 の記憶領域に書き込まれた前記第 2 のファイルの実データの位置を示すように、変更するようにしてもよい。

【 0 0 1 7 】

このように、第 1 の記憶領域のみにメタデータが格納される場合には、ファイルを移動させる際に、該ファイルに関するメタデータを新たに作成せずに変更するだけで済む。

【 0 0 1 8 】

上記の装置において、

前記ファイル管理部は、

前記統合記憶領域内の前記第 1 の記憶領域と前記第 2 の記憶領域とのそれぞれに、メタデータが格納されるグループ領域を設定するようにしてもよい。

30

【 0 0 1 9 】

こうすれば、第 1 種のストレージ媒体が第 2 種のストレージ媒体よりも高価であり、かつ、第 1 の記憶領域のみにメタデータが格納される場合と比較して、メタデータの格納に要するビット当たりのコストを低減させることができる。

【 0 0 2 0 】

上記の装置において、

前記ファイル管理部は、

第 1 のファイルを前記統合記憶領域内の前記第 1 の記憶領域に格納する場合には、前記第 1 のファイルの実データを前記第 1 の記憶領域に書き込むと共に、前記第 1 のファイルに関するメタデータを前記第 1 の記憶領域に作成し、

第 2 のファイルを前記統合記憶領域内の前記第 2 の記憶領域に格納する場合には、前記第 2 のファイルの実データを前記第 2 の記憶領域に書き込むと共に、前記第 2 のファイルに関するメタデータを前記第 2 の記憶領域に作成することが好ましい。

40

【 0 0 2 1 】

上記の装置において、

前記ファイル管理部は、

第 1 のファイルを前記統合記憶領域内の前記第 1 の記憶領域から前記第 2 の記憶領域に

50

移動させる場合には、前記第 1 のファイルの実データを前記第 2 の記憶領域に書き込むと共に、前記第 1 のファイルに関するメタデータを前記第 2 の記憶領域に作成し、

第 2 のファイルを前記統合記憶領域内の前記第 2 の記憶領域から前記第 1 の記憶領域に移動させる場合には、前記第 2 のファイルの実データを前記第 1 の記憶領域に書き込むと共に、前記第 2 のファイルに関するメタデータを前記第 1 の記憶領域に作成するようにしてもよい。

【 0 0 2 2 】

上記の装置において、

前記ファイル管理部は、

前記統合記憶領域内に、前記第 1 の記憶領域と前記第 2 の記憶領域とに跨るグループ領域を設定し、

前記ファイル管理部は、

前記第 1 の記憶領域の範囲と前記第 2 の記憶領域の範囲とを示す領域情報に基づいて、前記第 1 の記憶領域または前記第 2 の記憶領域にファイルの実データが格納されるように、前記ファイルシステムを構築することが好ましい。

【 0 0 2 3 】

こうすれば、第 1 の記憶領域と前記第 2 の記憶領域とに跨るグループ領域が設定される場合にも、ファイルが第 1 の記憶領域と第 2 の記憶領域とに跨って格納されずに済む。

【 0 0 2 4 】

上記の装置において、

前記ファイル管理部は、さらに、

前記領域情報に基づいて、前記第 1 の記憶領域にメタデータが格納され、前記第 2 の記憶領域にメタデータが格納されないように、前記ファイルシステムを構築することが好ましい。

【 0 0 2 5 】

こうすれば、メタデータを検索する際に、第 1 の記憶領域のみを検索すれば済み、この結果、目的のファイルに迅速にアクセスすることが可能となる。

【 0 0 2 6 】

上記の装置において、

前記ファイル管理部は、

第 1 のファイルを前記統合記憶領域内の前記第 1 の記憶領域に格納する場合には、前記領域情報に基づいて、前記第 1 のファイルの実データを前記第 1 の記憶領域に書き込むと共に、前記第 1 のファイルに関するメタデータを前記第 1 の記憶領域に作成し、

第 2 のファイルを前記統合記憶領域内の前記第 2 の記憶領域に格納する場合には、前記領域情報に基づいて、前記第 2 のファイルの実データを前記第 2 の記憶領域に書き込むと共に、前記第 2 のファイルに関するメタデータを前記第 1 の記憶領域に作成することが好ましい。

【 0 0 2 7 】

上記の装置において、

前記ファイル管理部は、

第 1 のファイルを前記統合記憶領域内の前記第 1 の記憶領域から前記第 2 の記憶領域に移動させる場合には、前記領域情報に基づいて、前記第 1 のファイルの実データを前記第 2 の記憶領域に書き込むと共に、前記第 1 の記憶領域に既に格納されている前記第 1 のファイルに関するメタデータを、前記第 2 の記憶領域に書き込まれた前記第 1 のファイルの実データの位置を示すように、変更し、

第 2 のファイルを前記統合記憶領域内の前記第 2 の記憶領域から前記第 1 の記憶領域に移動させる場合には、前記領域情報に基づいて、前記第 2 のファイルの実データを前記第 1 の記憶領域に書き込むと共に、前記第 1 の記憶領域に既に格納されている前記第 2 のファイルに関するメタデータを、前記第 1 の記憶領域に書き込まれた前記第 2 のファイルの実データの位置を示すように、変更することが好ましい。

10

20

30

40

50

【 0 0 2 8 】

このように、第 1 の記憶領域のみにメタデータが格納される場合には、ファイルを移動させる際に、該ファイルに関するメタデータを新たに作成せずに変更するだけで済む。

【 0 0 2 9 】

本発明の第 2 の装置は、ストレージシステムであって、
前記第 1 種のストレージ媒体と、
前記第 2 種のストレージ媒体と、
上記のいずれかに記載の前記管理装置と、
を備えることを特徴とする。

【 0 0 3 0 】

この発明は、種々の形態で実現することが可能であり、例えば、異種ストレージ媒体を管理するための管理装置、該管理装置を備えるストレージシステム、該ストレージシステムを備えるネットワークシステム、これらの装置における異種ストレージ媒体の管理方法、これらの方法または装置の機能を実現するためのコンピュータプログラム、そのコンピュータプログラムを記録した記録媒体、そのコンピュータプログラムを含み搬送波内に具現化されたデータ信号、等の形態で実現することができる。

【発明を実施するための最良の形態】

【 0 0 3 1 】

次に、本発明の実施の形態を実施例に基づき以下の順序で説明する。

A . 第 1 実施例 :

A - 1 . ネットワークシステムの構成 :

A - 1 - 1 . ストレージシステムの内部構成 :

A - 1 - 2 . 管理サーバの内部構成 :

A - 2 . ファイルシステムの作成 :

A - 3 . 統合記憶領域上のファイルシステムの構造 :

A - 4 . ファイルの新規作成 :

A - 5 . ファイルの移動 :

B . 第 2 実施例 :

C . 第 3 実施例 :

【 0 0 3 2 】

A . 第 1 実施例 :

A - 1 . ネットワークシステムの構成 :

図 1 は、ネットワークシステムの概略構成を示す説明図である。ネットワークシステムは、ストレージシステム (NAS) 100 と、管理サーバ 200 と、管理端末 300 と、複数のクライアント 400 と、を備えている。各デバイスは、IP ベースのネットワーク NW に接続されており、各デバイス間の通信は、TCP/IP に従って実行される。

【 0 0 3 3 】

ストレージシステム 100 は、ファイルサーバ (NAS ヘッドとも呼ばれる) 110 と、ストレージ装置 130 と、を備えている。本実施例では、ファイルサーバ 110 とストレージ装置 130 とは、異なる筐体に収納されているが、共通の筐体に収納されていてもよい。

【 0 0 3 4 】

A - 1 - 1 . ストレージシステムの内部構成 :

図 2 は、ストレージシステム 100 の内部構成を示す説明図である。上記のように、ストレージシステム 100 は、ファイルサーバ 110 とストレージ装置 130 とを備えている。そして、ストレージ装置 130 は、ディスクアレイコントローラ 140 と、ディスクアレイ装置 160 と、を備えている。なお、ファイルサーバ 110 とディスクアレイコントローラ 140 とが同一の装置 (例えば同一のプロセッサ) でも良い。

【 0 0 3 5 】

ファイルサーバ 110 は、CPU 112 と、ROM や RAM などのメモリ 114 と、ネ

10

20

30

40

50

ットワークインタフェース（ＩＦ）部１１６と、ストレージインタフェース（ＩＦ）部１１８と、を備えている。なお、ネットワークＩＦ部１１６は、ネットワークＮＷを介して、外部のデバイスと通信を行うためのインタフェースであり、ストレージＩＦ部１１８は、ストレージ装置１３０と通信を行うためのインタフェースである。

【００３６】

メモリ１１４には、プロトコル制御部１２２として機能するコンピュータプログラムと、第１構成管理部１２４として機能するコンピュータプログラムと、ファイルシステム制御部１２６として機能するコンピュータプログラムと、が格納されている。また、メモリ１１４には、伝送されるデータを一時的に格納するためのキャッシュ領域１２８が設けられている。

10

【００３７】

プロトコル制御部１２２は、ネットワークＮＷを介して、外部のデバイスとの通信を実現するために、プロトコルに従った種々の処理を実行する際にＣＰＵ１１２で実行されるプログラムである。

【００３８】

第１構成管理部１２４は、ディスクアレイ装置１６０に設定された１以上の論理ユニット上にファイルシステムを作成し、論理ユニットとファイルシステムとの対応関係を管理する際にＣＰＵ１１２で実行されるプログラムである。

【００３９】

ファイルシステム制御部１２６は、ファイルの新規作成や移動などの種々の処理を制御する際にＣＰＵ１１２で実行されるプログラムである。例えば、ファイルシステム制御部１２６は、ファイルの新規作成や移動などに伴って、ファイルシステムの内容を変更する。

20

【００４０】

ディスクアレイコントローラ１４０は、ＣＰＵ１４２と、ＲＯＭやＲＡＭなどのメモリ１４４と、ホストインタフェース（ＩＦ）部１４６と、ドライブインタフェース（ＩＦ）部１４８と、を備えている。なお、ホストＩＦ部１４６は、ストレージ装置１３０のホストとして機能するファイルサーバ１１０と通信を行うためのインタフェースであり、ドライブＩＦ部１４８は、ディスクアレイ装置１６０を駆動してデータを入出力するためのインタフェースである。

30

【００４１】

メモリ１４４には、入出力処理部１５２として機能するコンピュータプログラムと、第２構成管理部１５４として機能するコンピュータプログラムと、が格納されている。また、メモリ１４４には、伝送されるデータを一時的に格納するためのキャッシュ領域１５６が設けられている。

【００４２】

入出力処理部１５２は、ホストＩＦ部１４６を介してファイルサーバ１１０から与えられるＩ／Ｏコマンドを処理する。例えば、入出力処理部１５２は、読み出しコマンドや書き込みコマンドなどに従って、ディスクアレイ装置１６０に対するデータの読み出しや書き込み処理を実行する。

40

【００４３】

第２構成管理部１５４は、ディスクアレイ装置１６０に対して論理ユニットを設定し、論理ユニットと実際のディスク（ストレージ媒体）との対応関係を管理する際にＣＰＵ１４２で実行されるプログラムである。なお、第２構成管理部１５４は、ディスクアレイ装置１６０内に上記の対応関係を示すボリューム管理テーブルＴＢ（後述する）を作成する。

【００４４】

ディスクアレイ装置１６０は、複数のストレージ媒体（例えば複数のハードディスクドライブ）で構成されたディスクアレイを備えている。なお、ディスクアレイは、ＲＡＩＤ構成であっても良い。本実施例では、高価なＦＣ対応のハードディスクと、廉価なＡＴＡ

50

対応のハードディスクと、の２種類のハードディスクが利用されている。なお、以下では、ＦＣ対応のハードディスクを「第１種の媒体」とも呼び、ＡＴＡ対応のハードディスクを「第２種の媒体」とも呼ぶ。

【００４５】

ディスクアレイ装置１６０は、通常、論理ユニット（ＬＵ）に区分されて利用される。図２では、システム用のＬＵと、３つのデータ用のＬＵと、が準備されている。データ用の第１ＬＵは、第１種の媒体（ＦＣ）の一部で構成された記憶領域と、第２種の媒体（ＡＴＡ）の一部で構成された記憶領域と、を含む統合記憶領域（統合ＬＵ）であり、論理ユニット番号（ＬＵＮ）「０」が割り当てられている。第２ＬＵは、第１種の媒体（ＦＣ）の一部で構成された記憶領域（ＬＵ）であり、ＬＵＮ「１」が割り当てられている。第３ＬＵは、第２種の媒体（ＡＴＡ）の一部で構成された記憶領域（ＬＵ）であり、ＬＵＮ「２」が割り当てられている。システム用ＬＵは、本実施例では、第１種の媒体（ＦＣ）の一部で構成された記憶領域（ＬＵ）である。図示するように、システム用ＬＵには、ボリューム管理テーブルＴＢが格納される。なお、各記憶領域は、それぞれ、対応する種類の１つのディスクの一部または全部で構成されていてもよいし、対応する種類の複数のディスクで構成されていてもよい。

10

【００４６】

ボリューム管理テーブルＴＢは、ネットワークシステムの管理者が管理端末３００を操作することによって予め作成される。具体的には、管理者からの指示に従って、ストレージ装置１３０内の第２構成管理部１５４がボリューム管理テーブルＴＢを作成する。

20

【００４７】

図３は、予め準備されたボリューム管理テーブルＴＢの内容を示す説明図である。なお、このボリューム管理テーブルＴＢは、データ用の各ＬＵ上にファイルシステムが作成される前に、作成される。図示するように、ボリューム管理テーブルＴＢには、データ用の各ＬＵの情報が含まれている。具体的には、ボリューム管理テーブルＴＢには、記憶領域を構成する媒体の種類毎に、該記憶領域を構成する媒体の種類と、該記憶領域の開示番地および終了番地と、が設定されている。

【００４８】

例えば、第１ＬＵ（ＬＵＮ＝０）は、図２で説明したように、第１種の媒体（ＦＣ）で構成された記憶領域と、第２種の媒体（ＡＴＡ）で構成された記憶領域と、を含んでいる。このため、ボリューム管理テーブルＴＢでは、第１ＬＵに関して、第１の記憶領域の媒体の種類として「ＦＣ」が登録されていると共に、第１の記憶領域の開始番地「０」および終了番地「９９９」が登録されている。また、第２の記憶領域の媒体の種類として「ＡＴＡ」が登録されていると共に、第２の記憶領域の開始番地「１０００」および終了番地「１９９９」が登録されている。

30

【００４９】

同様に、第２ＬＵ（ＬＵＮ＝１）は、第１種の媒体（ＦＣ）で構成された記憶領域のみを含んでいる。このため、ボリューム管理テーブルＴＢでは、第２ＬＵに関して、第１の記憶領域の媒体の種類として「ＦＣ」が登録されていると共に、第１の記憶領域の開始番地「０」および終了番地「３０００」が登録されている。また、第３ＬＵ（ＬＵＮ＝２）は、第２種の媒体（ＡＴＡ）で構成された記憶領域のみを含んでいる。このため、ボリューム管理テーブルＴＢでは、第３ＬＵに関して、第１の記憶領域の媒体の種類として「ＡＴＡ」が登録されていると共に、第１の記憶領域の開始番地「０」および終了番地「５０００」が登録されている。

40

【００５０】

なお、本実施例におけるファイルサーバ１１０が本発明における管理装置に相当し、第１構成管理部１２４とファイルシステム制御部１２６とが本発明におけるファイル管理部に相当する。

【００５１】

A - 1 - 2 . 管理サーバの内部構成：

50

図4は、管理サーバ200の内部構成を示す説明図である。管理サーバ200は、CPU202と、ROMやRAMなどのメモリ204と、ネットワークインタフェース(IF)部206と、を備えている。なお、ネットワークIF部206は、ネットワークNWを介して、外部のデバイスと通信を行うためのインタフェースである。

【0052】

メモリ204には、プロトコル制御部212として機能するコンピュータプログラムと、ファイル分配制御部214として機能するコンピュータプログラムと、が格納されている。

【0053】

プロトコル制御部212は、図2のプロトコル制御部122と同様に、ネットワークNWを介して、外部のデバイスとの通信を実現するために、プロトコルに従った種々の処理を実行する際にCPU202で実行されるプログラムである。

【0054】

ファイル分配制御部214は、ストレージシステム100内の統合記憶領域(統合LU)に格納されるファイルが、いずれの種類の媒体に格納されるべきかを決定する際にCPU202で実行されるプログラムである。具体的には、ファイル分配制御部214は、ファイルサーバ110から問い合わせがあった場合に、予め管理者によって設定された分配ルールに従って、統合LU内におけるファイルの格納先の媒体を決定する。分配ルールは、例えば、ファイル名や、ファイルの最終更新日時、ファイルのサイズ、ファイルの作成者名などのファイルに付随する種々の情報(属性)を利用して、設定される。分配ルールに従えば、ファイルの価値に応じて格納先の媒体の種類を決定することができる。なお、分配ルールの情報は、管理サーバ200のメモリに格納される。

【0055】

例えば、ファイル分配制御部214は、ファイル名に所定の文字列が含まれる場合や、ファイルの最終更新日時が所定の日時以降である場合、ファイルのサイズが所定のサイズ以上である場合、ファイルの作成者名がリストに登録された名前と一致する場合などに、ファイルの価値が高いとみなし、ファイルの格納先を第1種の媒体(FC)と決定する。

【0056】

なお、分配ルールとしては、ファイルに付随する種々の情報(属性)だけでなく、ファイルに付随しない他の情報を利用することも可能である。例えば、ファイルがストレージシステム100に格納される時間帯が利用されてもよい。この場合には、例えば、多数のファイルが格納される時間帯(平日の就業時間)には、ファイルの格納先は、書き込みに要する時間が比較的短い第1種の媒体(FC)と決定され、少数のファイルが格納される時間帯(平日の就業時間以外)には、ファイルの格納先は、書き込みに要する時間が比較的長い第2種の媒体(ATA)と決定されればよい。

【0057】

A-2. ファイルシステムの作成:

図5は、ファイルシステムの作成手順を示すフローチャートである。ステップS102では、管理者は、管理端末300を操作して、ファイルサーバ110に、対象LU上へのファイルシステムの作成を指示する。なお、この際、管理者は、管理端末300を介して、ファイルサーバ110に、記憶領域へのメタデータの格納の可否を指定する情報を送信する。ここで、メタデータの格納の可否については、管理者は、例えば、高速・高コストなFCと低速・低コストなATAとを統合する場合、「高いコストを払っても、メタデータはFCだけに格納して高速にアクセスしたい」ときは、FCのみにメタデータを格納し、「メタデータをFCとATAとの両方に格納し、性能は優れていなくても低コストにしたい」ときは、FCとATAとにメタデータを格納するといった情報に基づいて、判断する。そして、管理者は、この判断に基づいて、上述のメタデータの格納の可否を指定する情報を作成する。

【0058】

ステップS104では、指示を受けたファイルサーバ110内の第1構成管理部124

10

20

30

40

50

は、ストレージ装置 130 内の第 2 構成管理部 154 に、対象 LU を構成する媒体の種類を問い合わせる。このとき、第 2 構成管理部 154 は、ボリューム管理テーブル TB を参照して、対象 LU を構成する媒体の種類を、第 1 構成管理部 124 に通知する。

【0059】

ステップ S106 では、第 1 構成管理部 124 は、第 2 構成管理部 154 からの通知に基づいて、対象 LU が複数種の媒体で構成されるか否かを判断する。対象 LU が複数種の媒体で構成される場合（例えば対象 LU が第 1 LU（ $LUN = 0$ ）である場合）には、ステップ S108 に進む。一方、対象 LU が複数種の媒体で構成されず 1 種類の媒体で構成される場合（例えば対象 LU が第 2 LU（ $LUN = 1$ ）または第 3 LU（ $LUN = 2$ ）である場合）には、ステップ S112 に進む。

10

【0060】

ステップ S108 では、第 1 構成管理部 124 は、管理端末 300 から、媒体の種類毎の用法（設定）の指定が有るか否かを判断する。本実施例では、第 1 種の媒体にはメタデータを格納し、第 2 種の媒体にはメタデータを格納しない、という用法が指定されている。なお、メタデータは、ディレクトリ情報と、ファイルの実データの格納位置を示すデータポインタを含む *inode* と、を含んでいる。メタデータについては、さらに、後述する。

【0061】

用法の指定が有る場合には、ステップ S110 に進む。一方、用法の指定が無い場合には、ステップ S112 に進む。

20

【0062】

ステップ S110 では、第 1 構成管理部 124 は、対象 LU（統合 LU）内に、指定された用法に基づいて、ブロックグループを作成する。この際、第 1 構成管理部 124 は、ボリューム管理テーブル TB（図 3）を参照し、各記憶領域のアドレス範囲を考慮して、ブロックグループを作成する。これにより、対象 LU 内にファイルシステムが作成される。なお、ブロックグループが、本発明におけるグループ領域に相当する。

【0063】

図 6 は、ステップ S110（図 5）で第 1 LU（ $LUN = 0$ ）上に作成されるファイルシステムの構造を示す説明図である。

【0064】

30

LU 上にファイルシステムが作成される場合には、通常、LU が複数種の媒体で構成されているか否かに拘わらず、LU に対して 1 以上のブロックグループが作成される。このため、LU が複数種の媒体で構成されている場合には、ブロックグループが 2 種類以上の媒体に跨って作成される恐れがある。

【0065】

しかしながら、本実施例では、各ブロックグループは、各記憶領域のアドレス範囲を考慮して、記憶領域毎に、換言すれば媒体の種類毎に、作成される。このため、各ブロックグループは、2 種類以上の媒体を跨ることなく、1 つの媒体内に作成される。

【0066】

具体的には、本実施例では、図 6 に示すように、第 1 種の媒体（FC）で構成された第 1 の記憶領域には、3 つのブロックグループが作成され、第 2 種の媒体（ATA）で構成された第 2 の記憶領域にも、3 つのブロックグループが作成される。そして、各ブロックグループは、第 1 種の媒体と第 2 種の媒体とを跨ることなく、換言すれば、第 1 の記憶領域と第 2 の記憶領域とを跨ることなく、作成される。なお、ブロックグループは、例えば、所定のサイズ（容量）毎に作成されてもよいし、所定の個数だけ作成されてもよい。

40

【0067】

ところで、図 6 に示すように、第 1 の記憶領域に作成される各ブロックグループ（以下「第 1 種のブロックグループ」とも呼ぶ）の構成と、第 2 の記憶領域に作成される各ブロックグループ（以下「第 2 種のブロックグループ」とも呼ぶ）の構成とは、異なっている。

50

【 0 0 6 8 】

具体的には、第1種のブロックグループは、通常の構成を有しており、ブロックグループ配置情報と、`inode` (`i` ノード) カウンタと、`inode` リストと、フリーデータブロックカウンタと、フリーデータブロックリストと、データブロックと、を含んでいる。第2種のブロックグループは、第1種のブロックグループとほぼ同様の構成を有しているが、`inode` カウンタと、`inode` リストと、を含んでいない。また、第1種のブロックグループ内のデータブロックには、メタデータが格納されるが、第2種のブロックグループ内のデータブロックには、メタデータは格納されない。なお、ブロックグループに含まれる上記の各要素については、後述する。

【 0 0 6 9 】

10

このように、第1の記憶領域に作成されるブロックグループの構成と、第2の記憶領域に作成されるブロックグループの構成と、が異なるのは、用法の相違に起因している。すなわち、本実施例では、ステップ S 1 0 8 において、第1種の媒体にはメタデータを格納し、第2種の媒体にはメタデータを格納しない、という用法が指定されている。このため、本実施例では、第1種の媒体で構成される第1の記憶領域には、`inode` カウンタと `inode` リストとを含み、メタデータを格納可能な第1種のブロックグループが作成される。また、第2種の媒体で構成される第2の記憶領域には、`inode` カウンタと `inode` リストとを含まず、メタデータの格納が考慮されない第2種のブロックグループが作成される。

【 0 0 7 0 】

20

図5の説明に戻る。ステップ S 1 1 2 (図5) では、第1構成管理部 1 2 4 は、対象 L U 内に、通常の方法で、換言すれば、媒体の種類を意識せずに、ブロックグループを作成する。本実施例では、第2 L U (`LUN` = 1) と第3 L U (`LUN` = 2) とのそれぞれに、1以上のブロックグループが作成される。なお、各ブロックグループは、`inode` カウンタと `inode` リストとを含み、メタデータを格納可能な第1種のブロックグループである。

【 0 0 7 1 】

ステップ S 1 1 4 では、第1構成管理部 1 2 4 は、第2構成管理部 1 5 4 にボリューム管理テーブル T B を更新させる。

【 0 0 7 2 】

30

図7は、図5の処理が完了したときに得られるボリューム管理テーブル T B の内容を示す説明図である。図7は、図3とほぼ同じであるが、記憶領域を構成する媒体の種類毎に、メタデータの格納の可否が追加されている。

【 0 0 7 3 】

図7では、メタデータを格納可に設定された記憶領域には「☐」印が付与されており、メタデータを格納不可に設定された記憶領域には「☒」印が付与されている。

【 0 0 7 4 】

例えば、第1 L U (`LUN` = 0) は、第1種の媒体 (`FC`) と第2種の媒体 (`ATA`) とで構成されている。このため、ステップ S 1 0 8 の用法の指定に基づいて、第1 L U 内の第1種の媒体 (`FC`) で構成された第1の記憶領域は、メタデータを格納可に設定されており、第2種の媒体 (`ATA`) で構成された第2の記憶領域は、メタデータを格納不可に設定されている。また、第2 L U (`LUN` = 1) と第3 L U (`LUN` = 2) とは、1種類の媒体で構成されている。このため、第2 L U 内の記憶領域と第3 L U 内の記憶領域とは、それぞれメタデータを格納可に設定されている。

40

【 0 0 7 5 】

A - 3 . 統合記憶領域上のファイルシステムの構造 :

図8は、第1 L U (`LUN` = 0) 上のファイルシステムの構造を示す説明図である。図8は、図6とほぼ同じであるが、図8では、データブロック内部のデータの図示が追加されている。なお、図8では、既にファイルが格納されている。

【 0 0 7 6 】

50

前述したように、第1種および第2種のブロックグループは、ブロックグループ配置情報と、フリーデータブロックカウンタと、フリーデータブロックリストと、データブロックと、を含んでいる。そして、第1種のブロックグループは、さらに、`inode`カウンタと、`inode`リストと、を含んでいる。

【0077】

また、第1種および第2種のブロックグループは、共に、データブロックを含んでいるが、第1種のブロックグループ内のデータブロックと第2種のブロックグループ内のデータブロックとは、異なる情報を含んでいる。具体的には、第1種のブロックグループ内のデータブロックには、ディレクトリ情報および`inode`で構成されるメタデータと、ファイルの実データが格納されるデータ領域と、が含まれているが、第2種のブロックグループ内のデータブロックには、データ領域のみが含まれている。

10

【0078】

ディレクトリ情報は、ファイル名と`inode`番号との対応関係が登録された情報である。なお、図8では、1つのディレクトリが作成されているが、複数のディレクトリが作成され得る。

【0079】

各`inode`は、ファイルのサイズや、最終更新日時、データポインタなどの種々の情報を含んでいる。データポインタは、ファイルの実データが格納されるデータ領域の位置（アドレス）を示す。

【0080】

20

例えば、図8では、第1の記憶領域内の「ブロックグループ1」において、「ファイルA」には「`inode1`」が対応付けられており、「`inode1`」のデータポインタは、「ファイルA」を構成するデータブロック（以下、「実データ」と称する）が格納されたデータ領域の位置を特定する。同様に、「ブロックグループ1」において、「ファイルB」には「`inode2`」が対応付けられており、「`inode2`」のデータポインタは、「ファイルB」の実データが格納されたデータ領域の位置を示す。ただし、図8では、「ファイルA」の実データは、第1の記憶領域内の「ブロックグループ1」内に格納されており、「ファイルB」の実データは、第2の記憶領域内の「ブロックグループ4」内に格納されている。

【0081】

30

このように、本実施例では、ディレクトリ情報と`inode`とで構成されるメタデータは、第1の記憶領域のみに格納され、ファイルの実データは、データの価値に応じて、2つの記憶領域のうちのいずれか一方に格納される。

【0082】

ブロックグループ配置情報は、自己のブロックグループの範囲（アドレス）を示す情報である。

【0083】

フリーデータブロックカウンタは、`inode`やファイルの実データをデータブロック内の空き領域に格納する際に、各領域に番号を順次割り当てるためのカウンタである。フリーデータブロックリストは、データブロック内の空き領域の範囲（アドレス）を示す情報を含んでいる。

40

【0084】

`inode`カウンタは、各`inode`に番号を順次割り当てるためのカウンタである。`inode`リストは、データブロック内の各`inode`の位置（アドレス）を示す情報を含んでいる。

【0085】

なお、図8では、既にファイルが格納されているため、第1種のブロックグループ内のデータブロックには、メタデータが格納されているが、図5の処理の完了直後には、メタデータは格納されていない。図5の処理では、第1種のブロックグループ内のデータブロックには、ルートディレクトリのみが作成される。

50

【 0 0 8 6 】

また、図 5 でブロックグループが作成される際には、各ブロックグループの初期設定も同時に行われる。例えば、ブロックグループ配置情報に自己のブロックグループのアドレス範囲が登録され、i n o d e カウンタおよびフリーデータブロックカウンタの値が初期値に設定される。

【 0 0 8 7 】

A - 4 . ファイルの新規作成 :

図 9 は、L U 内にファイルを新規に作成する際の手順を示すフローチャートである。

【 0 0 8 8 】

ステップ S 2 0 2 では、ファイルサーバ 1 1 0 は、クライアント 4 0 0 から格納対象ファイルを受信する。具体的には、クライアント 4 0 0 のユーザは、コンピュータ (クライアント 4 0 0) の表示画面を確認しつつ、ネットワークに接続された対象ディスクへの対象ファイルの格納を指示する。このとき、クライアント 4 0 0 は、該対象ディスクを示すディスク情報と、対象ファイルと、をファイルサーバ 1 1 0 に送信する。そして、ファイルサーバ 1 1 0 内のファイルシステム制御部 1 2 6 は、受信したディスク情報に基づいて、受信した対象ファイルを格納すべき L U (格納先 L U) を特定する。

10

【 0 0 8 9 】

ステップ S 2 0 3 では、ファイルシステム制御部 1 2 6 は、ボリューム管理テーブル T B を参照して、格納先 L U が複数種の媒体で構成されているか否かを判断する。格納先 L U が複数種の媒体で構成されている場合 (例えば格納先 L U が第 1 L U (L U N = 0) である場合) には、ステップ S 2 0 4 に進み、格納先 L U が複数種の媒体で構成されていない場合 (例えば格納先 L U が第 2 L U (L U N = 1) または第 2 L U (L U N = 2) である場合) には、ステップ S 2 0 8 に進む。

20

【 0 0 9 0 】

ステップ S 2 0 4 では、ファイルシステム制御部 1 2 6 は、管理サーバ 2 0 0 に対象ファイルを格納すべき媒体 (格納先媒体) の種類を問い合わせる。この際、ファイルシステム制御部 1 2 6 は、対象ファイルについての情報も併せて管理サーバ 2 0 0 に送信する。このとき、管理サーバ 2 0 0 内のファイル分配制御部 2 1 0 は、受信した対象ファイルのファイル名や最終更新日時などの情報に基づいて、分配ルールに従って格納先媒体の種類を決定し、ファイルシステム制御部 1 2 6 に通知する。

30

【 0 0 9 1 】

ステップ S 2 0 6 では、ファイルシステム制御部 1 2 6 は、ステップ S 2 0 4 で指定された格納先媒体の種類に対応する格納先記憶領域にメタデータを格納可能か否かを判断する。この判断は、ボリューム管理テーブル T B (図 7) を参照することによって行われる。格納先記憶領域にメタデータを格納可である場合 (例えば格納先記憶領域が第 1 種の媒体 (F C) で構成された第 1 の記憶領域である場合) には、ステップ S 2 0 8 に進む。一方、格納先記憶領域にメタデータを格納不可である場合 (例えば格納先記憶領域が第 2 種の媒体 (A T A) で構成された第 2 の記憶領域である場合) には、ステップ S 2 1 0 に進む。

40

【 0 0 9 2 】

ステップ S 2 0 8 では、ファイルシステム制御部 1 2 6 は、格納先記憶領域 (例えば図 8 の第 1 の記憶領域) に、i n o d e を作成すると共にファイルの実データを書き込む。なお、格納先記憶領域のアドレス範囲は、ボリューム管理テーブル T B を参照することによって特定される。

【 0 0 9 3 】

具体的には、まず、格納先記憶領域に作成された 1 以上のブロックグループのうちの 1 つのブロックグループ (例えば図 8 の「ブロックグループ 1」) が対象ブロックグループとして選択される。なお、ブロックグループが 2 以上存在する場合には、対象ブロックグループは、例えば、ファイルシステム制御部 1 2 6 によって 2 以上のブロックグループの中から循環的に順次選択される。

50

【 0 0 9 4 】

そして、格納先記憶領域の対象ブロックグループにおいて、フリーデータブロックリストから新規 *i n o d e* 領域が取得され、*i n o d e* リストに該新規 *i n o d e* 領域のアドレスが登録される。なお、新規 *i n o d e* 領域の取得の際には、*i n o d e* カウンタによって新規 *i n o d e* 番号が決定され、*i n o d e* カウンタの値が 1 だけインクリメントされる。また、新規 *i n o d e* 領域の取得の際には、フリーデータブロックカウンタの値が 1 だけインクリメントされる。

【 0 0 9 5 】

次に、格納先記憶領域の対象ブロックグループにおいて、フリーデータブロックリストから新規データ領域が取得され、*i n o d e* のデータポインタに該新規データ領域のアドレスが登録される。なお、新規データ領域の取得の際には、フリーデータブロックカウンタの値が 1 だけインクリメントされる。そして、新規データ領域にファイルの実データが書き込まれる。

10

【 0 0 9 6 】

例えば、ステップ S 2 0 8 において、「ファイル A」が第 1 L U (L U N = 0) の第 1 種の媒体 (F C) で構成された第 1 の記憶領域に格納される場合には、図 8 に示すように、第 1 の記憶領域に「*i n o d e* 1」が作成されると共に「ファイル A」の実データが格納される。なお、図 8 では、「ファイル A」が格納される際に、「ブロックグループ 1」が対象ブロックグループとして選択されている。

【 0 0 9 7 】

20

一方、ステップ S 2 1 0 では、ファイルシステム制御部 1 2 6 は、格納先記憶領域 (例えば図 8 の第 2 の記憶領域) と異なるメタデータを格納可能な他の記憶領域 (例えば図 8 の第 1 の記憶領域) に *i n o d e* を作成すると共に、格納先記憶領域にファイルの実データを書き込む。なお、格納先記憶領域のアドレス範囲と他の記憶領域のアドレス範囲は、ボリューム管理テーブル T B を参照することによって特定される。

【 0 0 9 8 】

具体的には、まず、他の記憶領域 (例えば図 8 の第 1 の記憶領域) に作成された 1 以上のブロックグループのうちの 1 つのブロックグループ (例えば「ブロックグループ 1」) が第 1 対象ブロックグループとして選択される。また、格納先記憶領域 (例えば図 8 の第 2 の記憶領域) に作成された 1 つのブロックグループ (例えば「ブロックグループ 4」) が第 2 対象ブロックグループとして選択される。

30

【 0 0 9 9 】

そして、他の記憶領域の第 1 対象ブロックグループにおいて、フリーデータブロックリストから新規 *i n o d e* 領域が取得され、*i n o d e* リストに該新規 *i n o d e* 領域のアドレスが登録される。

【 0 1 0 0 】

次に、格納先記憶領域の第 2 対象ブロックグループにおいて、フリーデータブロックリストから新規データ領域が取得され、*i n o d e* のデータポインタに該新規データ領域のアドレスが登録される。そして、新規データ領域にファイルの実データが書き込まれる。

【 0 1 0 1 】

40

例えば、ステップ S 2 1 0 において、「ファイル B」が第 1 L U (L U N = 0) の第 2 種の媒体 (A T A) で構成された第 2 の記憶領域に格納される場合には、図 8 に示すように、第 1 の記憶領域に「*i n o d e* 2」が作成され、第 2 の記憶領域に「ファイル B」の実データが格納される。なお、図 8 では、「ファイル B」が格納される際に、「ブロックグループ 1」が第 1 対象ブロックグループとして選択されており、「ブロックグループ 4」が第 2 対象ブロックグループとして選択されている。

【 0 1 0 2 】

ステップ S 2 1 2 では、ファイルシステム制御部 1 2 6 は、対象ファイルが作成されたディレクトリのディレクトリ情報に、ファイル名と *i n o d e* 番号との対応関係を登録する。

50

【0103】

例えば、ステップS208で「ファイルA」が第1LU(LUN=0)の第1種の媒体(FC)で構成された第1の記憶領域に格納される場合には、図8に示すように、第1の記憶領域内のディレクトリ情報に「ファイルA」と「inode1」との対応関係が登録される。また、ステップS210で「ファイルB」が第1LU(LUN=0)の第2種の媒体(ATA)で構成された第2の記憶領域に格納される場合には、第1の記憶領域内のディレクトリ情報に「ファイルB」と「inode2」との対応関係が登録される。

【0104】

ステップS214では、ファイルシステム制御部126は、クライアント400に対して、ファイルの新規作成が正常に終了したことを通知する。これにより、ファイルの新規作成処理が完了する。

10

【0105】

なお、格納先LUが1種類の媒体で構成される場合(例えば格納先LUが第2LU(LUN=1)または第3LU(LUN=2)である場合)には、ステップS208の処理が実行され、格納先記憶領域にinodeが作成されると共にファイルの実データが書き込まれる。

【0106】

A-5. ファイルの移動:

図10は、統合LU内でファイルを移動させる際の手順を示すフローチャートである。統合LU内でファイルを移動させるのは、ファイルの価値は変化するためである。例えば、ファイルが格納された時にはファイルの価値が高くても、現在ではファイルの価値が低くなっている可能性がある。逆に、ファイルが格納された時にはファイルの価値が低くても、現在ではファイルの価値が高くなっている可能性がある。また、管理サーバ200内のファイル分配制御部214が有する分配ルールが、ファイルが格納された時と現在とで、異なる場合もある。

20

【0107】

ステップS302では、ファイルサーバ110は、管理サーバ200から、移動対象ファイルと、移動元媒体の種類と、移動先媒体の種類と、を受信する。なお、移動元媒体の種類は、移動対象ファイルが格納されていた媒体の種類であり、移動先媒体の種類は、移動対象ファイルが格納されるべき媒体の種類である。なお、管理サーバ200が移動対象ファイル等の情報を送信する契機(管理サーバ200がファイルシステムを調べる契機)は、例えば、負荷の低い夜間や休日等といった契機が考えられる。

30

【0108】

具体的には、管理サーバ200内のファイル分配制御部210は、ボリューム管理テーブルTBを参照しつつ、ストレージシステム100内のファイルシステムを調べ、複数種の媒体で構成される統合LU内の各ファイルが分配ルールに従って格納されているか否かを調査する。そして、ファイル分配制御部210は、分配ルールと整合しないファイルを移動対象ファイルとして決定すると共に、該ファイルの移動元媒体の種類を特定する。また、ファイル分配制御部210は、分配ルールに従って、移動先媒体の種類を決定する。

【0109】

例えば、図8の第1LU(LUN=0)内の第1種の媒体(FC)で構成される第1の記憶領域に格納されている価値の低くなったファイルは、分配ルールと整合せず、移動対象ファイルとして決定され、第1種の媒体(FC)が移動元媒体であると特定される。そして、分配ルールに従って、第2種の媒体(ATA)が移動先媒体に決定される。

40

【0110】

ステップS304では、ファイルシステム制御部126は、移動先記憶領域に移動対象ファイルの実データを書き込むと共に、移動対象ファイルに対応するinodeの内容を変更する。なお、移動先記憶領域のアドレス範囲は、ボリューム管理テーブルTBを参照することによって特定される。

【0111】

50

具体的には、まず、移動先記憶領域に作成された１以上のブロックグループのうちの１つのブロックグループが対象ブロックグループとして選択される。

【０１１２】

次に、移動先記憶領域の対象ブロックグループにおいて、フリーデータブロックリストから新規データ領域が取得される。また、対象ファイルに対応する *inode* のデータポインタが示すアドレスが、旧データ領域のアドレスから該新規データ領域のアドレスに変更される。なお、対象ファイルに対応する *inode* は、移動先記憶領域が図８の第１の記憶領域である場合には移動先記憶領域に含まれているが、移動先記憶領域が図８の第２の記憶領域である場合には移動元記憶領域（すなわち図８の第１の記憶領域）に含まれている。

10

【０１１３】

そして、移動対象ファイルの実データが移動元記憶領域内の旧データ領域から読み出され、移動先記憶領域の新規データ領域に書き込まれる。なお、移動対象ファイルの実データが格納された旧データ領域のアドレスは、移動処理前のファイル名に対応する *inode* のデータポインタを参照することによって特定される。

【０１１４】

ステップＳ３１０では、ファイルシステム制御部１２６は、移動元記憶領域内の移動対象ファイルの実データが依然として格納されている旧データ領域を開放する。具体的には、旧データ領域が含まれるブロックグループにおいて、フリーデータブロックリストに旧データ領域を空き領域として追加する。これにより、旧データ領域に他のデータを書き込むことが可能となる。

20

【０１１５】

ステップＳ３１２では、ファイルシステム制御部１２６は、ファイルの移動が正常に終了したことを管理サーバ２００内のファイル分配制御部２１０に通知する。これにより、ファイルの移動処理が完了する。

【０１１６】

なお、ステップＳ３０２における移動元媒体の種類の受信は、省略可能である。また、統合ＬＵ内が２種類の媒体で構成される場合には、さらに移動先媒体の種類の受信も、省略可能である。この場合には、ファイルシステム制御部１２６は、管理サーバ２００内のファイル分配制御部２１０から移動対象ファイルの情報のみを受け取って、移動対象ファ

30

【０１１７】

図１１は、ファイルの移動処理の前後のファイルシステムの構造を示す説明図である。図１１は、図８とほぼ同じであるが、移動処理後のデータブロック内部の図示が追加されている。

【０１１８】

図１１では、「ファイルＡ」が第１種の媒体（ＦＣ）で構成された第１の記憶領域から第２種の媒体（ＡＴＡ）で構成された第２の記憶領域に移動している。移動処理後には、第２の記憶領域内に「ファイルＡ」の実データを格納する新規データ領域が設けられていると共に、第１の記憶領域から「ファイルＡ」の実データが格納されていた旧データ領域が開放されている。また、「*inode* １」のデータポインタが示すアドレスが、旧データ領域のアドレスから新規データ領域のアドレスに変更されている。なお、ディレクトリ情報は、変更されていない。

40

【０１１９】

同様に、「ファイルＢ」を第２種の媒体（ＡＴＡ）で構成された第２の記憶領域から第１種の媒体（ＦＣ）で構成された第１の記憶領域に移動させる場合には、移動処理後には、第１の記憶領域内に「ファイルＢ」の実データを格納する新規データ領域が設けられると共に、第２の記憶領域から「ファイルＢ」の実データが格納されていた旧データ領域が開放される。また、「*inode* ２」のデータポインタが示すアドレスが、旧データ領域のアドレスから新規データ領域のアドレスに変更される。なお、ディレクトリ情報は、変

50

更されない。

【0120】

「ファイルB」の実データを格納する新規データ領域は、「inode2」が格納された「ブロックグループ1」に設けられることが好ましい。こうすれば、「ファイルB」に対応する「inode2」と「ファイルB」の実データを格納する新規データ領域とが、物理的に比較的近い位置に配置される可能性が高いため、「ファイルB」を比較的迅速に読み出すことができる。ただし、「ファイルB」の実データを格納する新規データ領域は、「inode2」が格納されていない他のブロックグループ（例えば「ブロックグループ2」または「ブロックグループ3」）に設けられていてもよい。この場合には、「inode2」のデータポインタが示すアドレスが、他のブロックグループ内に設けられた新規データ領域のアドレスに変更されればよい。

10

【0121】

上記のように、本実施例では、第1種の媒体（FC）で構成された第1の記憶領域のみにメタデータが格納されるため、ファイルを移動させる際に、該ファイルに関するメタデータを新たに作成せずに変更するだけで済む。

【0122】

以上説明したように、本実施例では、第1種の媒体（FC）の一部で構成された第1の記憶領域と、第2種の媒体（FC）の一部で構成された第2の記憶領域と、を含む統合LU上に、1つのファイルシステムが構築される。具体的には、統合LU内には、複数のブロックグループが設定されるが、各ブロックグループは、第1の記憶領域と第2の記憶領域とに跨らないように設定される。すなわち、本実施例では、第1の記憶領域の範囲と第2の記憶領域の範囲とを考慮して、ブロックグループが設定される。このため、ファイルの新規作成や移動（マイグレーション）の際に、ファイルの実データが、第1の記憶領域と第2の記憶領域とに跨って（換言すれば第1種の媒体と第2種の媒体とに跨って）格納されずに済むと共に、1つのファイルシステム内で第1の記憶領域と第2の記憶領域とにファイル単位でファイルが分配される。したがって、ファイルを価値に応じた適切な種類の媒体に確実に格納することができ、この結果、複数種の媒体を効率よく利用することが可能となる。

20

【0123】

また、本実施例では、2つの記憶領域に統合LUが設定されて該統合LU上に1つのファイルシステムが構築されるため、2つの記憶領域に2つのLUが設定されて該2つのLU上に2つのファイルシステムが構築される場合と比較して、容量管理を容易に行うことができる。これは、1つのファイルシステムが構築される場合には、該1つのファイルシステムのみを対象として容量を調べれば済むが、2つのファイルシステムが構築される場合には、該2つのファイルシステムの双方を対象として容量を調べる必要があるためである。

30

【0124】

任意のファイルにアクセスする際には、メタデータが検索される。本実施例では、メタデータは、第1の記憶領域のみに格納されている。このため、本実施例では、メタデータを検索する際に、第1の記憶領域のみを検索すれば済み、この結果、目的のファイルに迅速にアクセスすることが可能となる。

40

【0125】

特に、本実施例では、メタデータは、第2種の媒体（ATA）よりも高速にアクセスが可能な第1種の媒体（FC）で構成される第1の記憶領域のみに格納されている。このため、本実施例では、メタデータが第1種の媒体で構成される第1の記憶領域と第2種の媒体で構成される第2の記憶領域との双方に格納される場合と比較して、ファイルの実データの位置を迅速に特定することができるという利点がある。

【0126】

ところで、一方の記憶領域のみにメタデータが格納され、かつ、双方の記憶領域に個別にファイルシステムが構築される場合には、ファイルシステム単位でデータを他のLUへ

50

複製（リモートコピー）することが困難となる。しかしながら、本実施例では、統合ＬＵ上に１つのファイルシステムが構築されるため、以下に説明するように、ファイルシステム単位のデータの複製を容易に行うことができる。

【０１２７】

図１２は、第１実施例におけるファイルシステムと比較例におけるファイルシステムとを模式的に示す説明図である。

【０１２８】

本実施例では、図１２（Ａ）に示すように、第１種の媒体（ＦＣ）で構成された第１の記憶領域と第２種の媒体（ＡＴＡ）で構成された第２の記憶領域とに統合ＬＵが設定されており、該統合ＬＵ上に１つのファイルシステムが構築されている。そして、メタデータは、第１の記憶領域のみに格納されている。

10

【０１２９】

一方、比較例では、図１２（Ｂ）に示すように、第１種の媒体（ＦＣ）で構成された第１の記憶領域に第１ＬＵが設定されており、第２種の媒体（ＡＴＡ）で構成された第２の記憶領域に第２ＬＵが設定されている。また、２つのＬＵ上に２つのファイルシステムが構築されている。そして、メタデータは、第１の記憶領域（第１ＬＵ）のみに格納されている。したがって、比較例では、ファイルの実データが第２の記憶領域（第２ＬＵ）内に格納される場合には、メタデータ内の *inode* のデータポインタは、第２ＬＵの *LUN* を含む必要がある。

【０１３０】

20

このため、仮に、ファイルシステム単位でデータの複製を行う場合には、例えば第２ＬＵ内のデータのみを他のＬＵにコピーする場合には、*inode* のデータポインタに含まれる *LUN* を変更する必要がある。このため、データの複製が困難となる。この問題は、第１ＬＵ内のデータのみを他のＬＵに複製する場合、第１および第２ＬＵ内のデータを他の２つのＬＵに複製する場合にも同様に生じる。

【０１３１】

一方、本実施例では、２つの記憶領域に統合ＬＵが設定されているため、ファイルの実データが第２の記憶領域内に格納される場合にも、メタデータ内の *inode* のデータポインタには *LUN* が含まれていない。このため、本実施例では、統合ＬＵ内のデータを他のＬＵに容易に複製することができる。すなわち、本実施例では、比較例と比べ、ファイルシステム単位のデータの複製を容易に行うことができる。

30

【０１３２】

また、本実施例では、２つの記憶領域に統合ＬＵが設定されて該統合ＬＵ上に１つのファイルシステムが構築されるため、２つの記憶領域に２つのＬＵが設定されて該２つのＬＵ上に２つのファイルシステムが構築される場合と比較して、ファイルシステムのイメージを作成するスナップショット機能を容易に適用することができる。

【０１３３】

B．第２実施例：

図１３は、第２実施例における第１ＬＵ（*LUN* = 0）上のファイルシステムの構造を示す説明図であり、図８に対応する。図１３では、図８と比較して、第２の記憶領域に作成されるブロックグループが変更されている。具体的には、第２の記憶領域には、第１の記憶領域に設定されるブロックグループと同様に、*inode* カウンタと *inode* リストとを含み、メタデータを格納可能な第１種のブロックグループが作成されている。

40

【０１３４】

本実施例では、ファイルシステムの作成は、第１実施例と同様に、図５に示す手順に従って実行可能である。ただし、本実施例では、第１種の媒体（ＦＣ）および第２種の媒体（ＡＴＡ）の双方にメタデータを格納する、という用法が指定されている。この結果、本実施例では、第１ＬＵ（*LUN* = 0）内に図１３に示すファイルシステムが作成される。また、本実施例では、ステップ *S* 114 で得られるボリューム管理テーブルの内容が、第１実施例（図７）と異なっている。具体的には、第１ＬＵ（*LUN* = 0）に関し、第１種

50

の媒体（ＦＣ）で構成される第１の記憶領域と第２種の媒体（ＡＴＡ）で構成される第２の記憶領域との双方が、メタデータを格納可に設定される。

【０１３５】

また、本実施例では、ファイルの新規作成は、第１実施例と同様に、図９に示す手順に従って実行可能である。ただし、本実施例では、統合ＬＵ内に作成されるブロックグループは、すべてメタデータを格納可能な第１種のブロックグループであるため、ステップＳ２０３，Ｓ２０４，Ｓ２０６，Ｓ２１０の処理は省略可能である。

【０１３６】

図１３では、「ファイルＡ」は、図８と同様に格納されている。一方、「ファイルＢ」の実データは、図８と同様に、第２の記憶領域内の「ブロックグループ４」内に格納されているが、「ファイルＢ」に関するメタデータ（ディレクトリ情報とinode）は、図８と異なり、第２の記憶領域内の「ブロックグループ４」内に格納されている。

10

【０１３７】

図１４は、第２実施例における統合ＬＵ内でファイルを移動させる際の手順を示すフローチャートであり、図１０に対応する。なお、図１４のステップＳ３０２，Ｓ３１０，Ｓ３１２は、図１０と同じである。

【０１３８】

ステップＳ３０６では、ファイルシステム制御部１２６は、移動先記憶領域に、inodeを作成すると共に移動対象ファイルの実データを書き込む。なお、ステップＳ３０６の処理は、図９のステップＳ２０８の処理とほぼ同様である。

20

【０１３９】

具体的には、まず、移動先記憶領域に作成された１以上のブロックグループのうちの１つのブロックグループが対象ブロックグループとして選択される。

【０１４０】

そして、移動先記憶領域の対象ブロックグループにおいて、フリーデータブロックリストから新規inode領域が取得され、inodeリストに該新規inode領域のアドレスが登録される。

【０１４１】

次に、移動先記憶領域の対象ブロックグループにおいて、フリーデータブロックリストから新規データ領域が取得され、inodeのデータポインタに該新規データ領域のアドレスが登録される。そして、移動対象ファイルの実データが移動元記憶領域内の旧データ領域から読み出され、移動先記憶領域の新規データ領域に書き込まれる。

30

【０１４２】

ステップＳ３０８では、ファイルシステム制御部１２６は、移動元記憶領域内のディレクトリ情報からファイル名と旧inodeとの対応関係を削除する。具体的には、移動元記憶領域内の移動対象ファイルが依然として格納されているブロックグループにおいて、ディレクトリ情報から上記の対応関係を削除する。

【０１４３】

また、ステップＳ３０８では、ファイルシステム制御部１２６は、移動元記憶領域内の移動対象ファイルのinodeが依然として格納されている旧inode領域を開放する。具体的には、旧inode領域が含まれるブロックグループにおいて、フリーデータブロックリストに旧inode領域を空き領域として追加する。これにより、旧inode領域に他のデータを書き込むことが可能となる。

40

【０１４４】

なお、ステップＳ３１０では、図１０で説明したように、移動元記憶領域内の移動対象ファイルの実データが依然として格納されている旧データ領域が開放される。

【０１４５】

図１５は、第２実施例におけるファイルの移動処理の前後のファイルシステムの構造を示す説明図である。図１５は、図１３とほぼ同じであるが、移動処理後のデータブロック内部の図示が追加されている。

50

【0146】

図15では、「ファイルA」が第1種の媒体(FC)で構成された第1の記憶領域から第2種の媒体(ATA)で構成された第2の記憶領域に移動している。移動処理後には、第2の記憶領域内のディレクトリ情報に「ファイルA」と「inode2」との対応関係が追加されていると共に、第1の記憶領域内のディレクトリ情報から「ファイルA」と「inode1」との対応関係が削除されている。また、第2の記憶領域内に「ファイルA」の新規inode領域が設けられていると共に、第1の記憶領域から「ファイルA」の旧inode領域が開放されている。さらに、第2の記憶領域内に「ファイルA」の新規データ領域が設けられていると共に、第1の記憶領域から「ファイルA」の旧データ領域が開放されている。

10

【0147】

同様に、「ファイルB」を第2種の媒体(ATA)で構成された第2の記憶領域から第1種の媒体(FC)で構成された第1の記憶領域に移動させる場合には、移動処理後には、第1の記憶領域内のディレクトリ情報に「ファイルB」と新規inodeとの対応関係が追加されると共に、第2の記憶領域内のディレクトリ情報から「ファイルB」と旧inodeとの対応関係が削除される。また、第1の記憶領域内に「ファイルB」の新規inode領域が設けられると共に、第2の記憶領域から「ファイルB」の旧inode領域が開放される。さらに、第1の記憶領域内に「ファイルB」の新規データ領域が設けられると共に、第2の記憶領域から「ファイルB」の旧データ領域が開放される。

【0148】

以上説明したように、本実施例でも、第1実施例と同様に、統合LU上に1つのファイルシステムが構築される。そして、統合LU内には、複数のブロックグループが設定されるが、各ブロックグループは、第1の記憶領域と第2の記憶領域とに跨らないように設定される。このため、ファイルの新規作成や移動の際に、ファイルの実データが、第1の記憶領域と第2の記憶領域とに跨って(換言すれば第1種の媒体と第2種の媒体とに跨って)格納されずに済むと共に、1つのファイルシステム内で第1の記憶領域と第2の記憶領域とにファイル単位でファイルが分配される。したがって、ファイルを価値に応じた適切な種類の媒体に確実に格納することができ、この結果、複数種の媒体を効率よく利用することが可能となる。

20

【0149】

また、本実施例では、第1実施例と同様に、容量管理や、ファイルシステム単位でのデータの複製、スナップショット機能の適用を容易に行うことができる。

30

【0150】

特に、本実施例では、比較的高価な第1の媒体(FC)で構成された第1の記憶領域と比較的安価な第2種の媒体(ATA)で構成された第2の記憶領域との双方に、メタデータを格納可能な第1種のブロックグループが設定される。一方、第1実施例では、比較的高価な第1種の媒体(FC)で構成された第1の記憶領域のみにメタデータを格納可能な第1種のブロックグループが設定される。このため、本実施例では、第1実施例と比較して、メタデータの格納に要するビット当たりのコストを低減させることができる。

【0151】

C. 第3実施例:

図16は、第3実施例における第1LU(LUN=0)上のファイルシステムの構造を示す説明図であり、図8に対応する。図16は、図8と比較して、第1LU内に作成されるブロックグループが変更されている。具体的には、第1LUには、1つのブロックグループ(第1種のブロックグループ)のみが作成される。すなわち、本実施例では、ブロックグループは、2つの記憶領域に跨って設定されている。

40

【0152】

ただし、本実施例では、後述するように、ブロックグループ配置情報等のデータブロック以外の情報(以下「非データブロック情報」)は、第1の記憶領域のみに格納される。また、データブロックは、2つの記憶領域に跨っているが、メタデータは、第1の記憶領

50

域のみに格納される。

【0153】

本実施例では、ファイルシステムの作成は、第1実施例と同様に、図5に示す手順に従って実行可能である。本実施例では、第1実施例と同様に、第1種の媒体（FC）にはメタデータを格納し、第2種の媒体（ATA）にはメタデータを格納しない、という用法が指定されている。さらに、本実施例では、1つのブロックグループのみが作成され、第1種の媒体（FC）には非データブロック情報を格納し、第2の媒体（ATA）には非データブロック情報を格納しない、という用法が指定されている。この結果、本実施例では、第1LU（LUN=0）内に図16に示すファイルシステムが作成される。なお、本実施例では、ステップS114で得られるボリューム管理テーブルの内容は、第1実施例（図7）と同じである。

10

【0154】

ところで、第1および第2実施例では、各ブロックグループは、2種類の媒体に跨らないように設けられている。そして、各ブロックグループのフリーデータブロックリストは、自己のブロックグループのアドレス範囲内の空き領域のみをinode領域やデータ領域として割り当て可能である。すなわち、第1および第2実施例では、格納すべき媒体の種類に応じてブロックグループが選択され、対応するフリーデータブロックリストからinode領域やデータ領域が取得されれば、該領域は、目的の種類の媒体で構成される領域である。しかしながら、本実施例では、唯一のブロックグループは、2種類の媒体に跨るように設けられている。このため、本実施例では、唯一のブロックグループのフリーデータブロックリストからinode領域やデータ領域が取得されても、該領域は、いずれの種類の媒体で構成される領域であるか不明である。そこで、本実施例では、以下に説明するように、ファイルの新規作成処理やファイルの移動処理が変更されている。具体的には、本実施例では、ボリューム管理テーブルの「第1の記憶領域の媒体の種類」（メタデータ格納可能な記憶領域）の開始番地と終了番地と、「第2の記憶領域の媒体の種類」（メタデータが格納されない記憶領域）の開始番地と終了番地をファイルサーバが意識して見分ける。これにより、2種類の媒体に跨ってブロックグループが設定されても、ファイルサーバはファイルを特定の領域に格納することができる。

20

【0155】

図17は、第3実施例におけるLU内にファイルを新規に作成する際の手順を示すフローチャートである。図17は、図9とほぼ同じであるが、ステップS208C、S210Cが変更されている。また、ステップS203で否定判断がなされた場合には、図9と同じステップS208に進む。

30

【0156】

ステップS208Cは、格納先LUが複数種の媒体で構成された統合LUであり、かつ、格納先記憶領域がメタデータを格納可である場合（例えば格納先LUが図16の第1LU（LUN=0）であり、格納先記憶領域が第1種の媒体（FC）で構成された第1の記憶領域である場合）に実行される。

【0157】

ステップS208Cでは、ファイルシステム制御部126は、格納先記憶領域（例えば図16の第1の記憶領域）に、inodeを作成すると共にファイルの実データを書き込む。

40

【0158】

具体的には、まず、統合LUの唯一のブロックグループにおいて、フリーデータブロックリストから新規inode領域が取得され、inodeリストに該新規inode領域のアドレスが登録される。ただし、新規inode領域は、メタデータを格納可能な記憶領域内に確保される必要がある。このため、ファイルシステム制御部126は、ボリューム管理テーブルTBを参照し、メタデータを格納可能な格納先記憶領域のアドレス範囲内に含まれる領域を新規inode領域として取得する。

【0159】

50

次に、統合LU内の唯一のブロックグループにおいて、フリーデータブロックリストから新規データ領域が取得され、inodeのデータポインタに該新規データ領域のアドレスが登録される。この際、ファイルシステム制御部126は、ボリューム管理テーブルTBを参照し、格納先記憶領域のアドレス範囲内に含まれる領域を新規データ領域として取得する。そして、該新規データ領域にファイルの実データが書き込まれる。

【0160】

一方、ステップS210Cは、格納先LUが複数種の媒体で構成された統合LUであり、かつ、格納先記憶領域がメタデータを格納不可である場合（例えば格納先LUが図16の第1LU（LUN=0）であり、格納先記憶領域が第2種の媒体（ATA）で構成された第1の記憶領域である場合）に実行される。

10

【0161】

ステップS210Cでは、ファイルシステム制御部126は、格納先記憶領域（例えば図16の第2の記憶領域）と異なるメタデータを格納可能な他の記憶領域（例えば図16の第1の記憶領域）にinodeを作成すると共に、格納先記憶領域にファイルの実データを書き込む。

【0162】

具体的には、まず、統合LU内の唯一のブロックグループにおいて、フリーデータブロックリストから新規inode領域が取得され、inodeリストに該新規inode領域のアドレスが登録される。この際、ファイルシステム制御部126は、ボリューム管理テーブルTBを参照し、メタデータを格納可能な他の記憶領域のアドレス範囲内に含まれる領域を新規inode領域として取得する。

20

【0163】

次に、統合LU内の唯一のブロックグループにおいて、フリーデータブロックリストから新規データ領域が取得され、inodeのデータポインタに該新規データ領域のアドレスが登録される。この際、ファイルシステム制御部126は、ボリューム管理テーブルTBを参照し、格納先記憶領域のアドレス範囲内に含まれる領域を新規データ領域として取得する。そして、該新規データ領域にファイルの実データが書き込まれる。

【0164】

図16では、「ファイルA」の実データとメタデータとは、図8と同様に、第1の記憶領域に格納されている。また、「ファイルB」の実データとメタデータとは、図8と同様に、第2の記憶領域に格納されている。しかしながら、「ファイルB」の実データとメタデータとは、図8では異なるブロックグループ内に格納されているが、図16では唯一のブロックグループ内に格納されている。

30

【0165】

図18は、第3実施例における統合LU内でファイルを移動させる際の手順を示すフローチャートであり、図10に対応する。図18は、図10とほぼ同じであるが、ステップS304Cが変更されている。

【0166】

ステップS304Cでは、ファイルシステム制御部126は、移動先記憶領域に移動対象ファイルの実データを書き込むと共に、移動対象ファイルに対応するinodeの内容を変更する。

40

【0167】

具体的には、統合LU内の唯一のブロックグループにおいて、フリーデータブロックリストから新規データ領域が取得される。この際、ファイルシステム制御部126は、ボリューム管理テーブルTBを参照し、移動先記憶領域のアドレス範囲内に含まれる領域を新規データ領域として取得する。また、対象ファイルに対応するinodeのデータポインタが示すアドレスが、旧データ領域のアドレスから新規データ領域のアドレスに変更される。

【0168】

そして、移動対象ファイルの実データが移動元記憶領域内の旧データ領域から読み出さ

50

れ、移動先記憶領域の新規データ領域に書き込まれる。

【 0 1 6 9 】

なお、ステップ S 3 1 0 では、図 1 0 で説明したように、移動元記憶領域内の移動対象ファイルの実データが依然として格納されている旧データ領域が開放される。

【 0 1 7 0 】

図 1 9 は、第 3 実施例におけるファイルの移動処理の前後のファイルシステムの構造を示す説明図である。図 1 9 は、図 1 6 とほぼ同じであるが、移動処理後のデータブロック内部の図示が追加されている。

【 0 1 7 1 】

図 1 9 では、「ファイル A」が第 1 種の媒体 (F C) で構成された第 1 の記憶領域から第 2 種の媒体 (A T A) で構成された第 2 の記憶領域に移動している。移動処理後には、第 2 の記憶領域内に「ファイル A」の実データを格納する新規データ領域が設けられていると共に、第 1 の記憶領域から「ファイル A」の実データが格納されていた旧データ領域が開放されている。また、「 i n o d e 1 」のデータポインタが示すアドレスが、旧データ領域のアドレスから新規データ領域のアドレスに変更されている。なお、ディレクトリ情報は、変更されていない。

10

【 0 1 7 2 】

同様に、「ファイル B」を第 2 種の媒体 (A T A) で構成された第 2 の記憶領域から第 1 種の媒体 (F C) で構成された第 1 の記憶領域に移動させる場合には、移動処理後には、第 1 の記憶領域内に「ファイル B」の実データを格納する新規データ領域が設けられると共に、第 2 の記憶領域から「ファイル B」の実データが格納されていた旧データ領域が開放される。また、「 i n o d e 2 」のデータポインタが示すアドレスが、旧データ領域のアドレスから新規データ領域のアドレスに変更される。なお、ディレクトリ情報は、変更されない。

20

【 0 1 7 3 】

上記のように、本実施例では、第 1 実施例と同様に、第 1 種の媒体 (F C) で構成された第 1 の記憶領域のみにメタデータが格納されるため、ファイルを移動させる際に、該ファイルに関するメタデータを新たに作成せずに変更するだけで済む。

【 0 1 7 4 】

以上説明したように、本実施例でも、第 1 実施例と同様に、統合 L U 上に 1 つのファイルシステムが構築される。ただし、本実施例では、統合 L U 内には、1 つのブロックグループのみが設定され、該ブロックグループは、第 1 の記憶領域と第 2 の記憶領域とに跨るように設定される。そこで、本実施例では、ボリューム管理テーブル T B を参照することによって、ファイルに関するメタデータが第 1 の記憶領域のみに格納され、ファイルの実データが適切な一方の記憶領域のみに格納される。このため、本実施例でも、ファイルの新規作成や移動の際に、ファイルの実データが、第 1 の記憶領域と第 2 の記憶領域とに跨って (換言すれば第 1 種の媒体と第 2 種の媒体とに跨って) 格納されずに済むと共に、1 つのファイルシステム内で第 1 の記憶領域と第 2 の記憶領域とにファイル単位でファイルが分配される。したがって、ファイルを価値に応じた適切な種類の媒体に確実に格納することができ、この結果、複数種の媒体を効率よく利用することが可能となる。

30

40

【 0 1 7 5 】

また、本実施例では、第 1 実施例と同様に、容量管理や、メタデータの迅速な検索、ファイルシステム単位でのデータの複製、スナップショット機能の適用を容易に行うことができる。

【 0 1 7 6 】

特に、本実施例では、汎用のファイルシステムと同様の構造を有するファイルシステムが統合 L U 上に構築されるため、クライアント 4 0 0 は、ファイルサーバ 1 1 0 を介さずに直接的にストレージ装置 1 3 0 内の構築されたファイルシステムを参照することができるという利点もある。

【 0 1 7 7 】

50

なお、この発明は上記の実施例や実施形態に限られるものではなく、その要旨を逸脱しない範囲において種々の態様で実施することが可能であり、例えば次のような変形も可能である。

【0178】

(1) 上記実施例では、複数種の媒体で構成される複数の記憶領域をストレージ装置130(より具体的には第2構成管理部154)の機能を利用して統合し、得られた統合記憶領域(統合LU)上にファイルシステムを構築している。しかしながら、これに代えて、他の手法を採用してもよい。例えば、ファイルサーバ110に論理ボリュームマネージャ(コンピュータプログラム)が備えられている場合には、複数種の媒体で構成される複数の記憶領域を論理ボリュームマネージャの機能を利用して統合し、得られた統合記憶領域上にファイルシステムを構築するようにしてもよい。また、バーチャリゼーションスイッチが利用される場合には、複数台の記憶装置内の複数種の媒体で構成される複数の記憶領域をバーチャリゼーションスイッチの機能を利用して統合し、得られた統合記憶領域上にファイルシステムを構築するようにしてもよい。

10

【0179】

また、上記実施例では、媒体としてFC対応のハードディスクとATA対応のハードディスクが利用されているが、他の媒体が利用されてもよい。例えば、光ディスクや、光磁気ディスク、磁気テープなどが利用されてもよい。

【0180】

一般には、第1種のストレージ媒体の少なくとも一部で構成された第1の記憶領域と、第2種のストレージ媒体の少なくとも一部で構成された第2の記憶領域と、を少なくとも含む統合記憶領域上に1つのファイルシステムが構築されればよい。

20

【0181】

(2) 上記実施例では、ファイル分配制御部214は、管理サーバ200に設けられているが、これに代えて、ファイルサーバ110に設けられていてもよい。こうすれば、ファイルサーバ110が、ファイルの新規作成処理(またはファイルの移動処理)の際に、ファイルの格納先媒体の種類(または移動対象ファイルおよび移動先媒体の種類)を決定することができる。

【0182】

また、上記実施例では、管理サーバ200がファイル分配制御部214を備えており、ファイルサーバ110は、ファイルの新規作成処理またはファイルの移動処理の際(より具体的にはステップS204、S302)に、管理サーバ200からファイルの格納先媒体の種類または移動対象ファイル等を受け取っているが、これに代えて、管理サーバから分配ルールを受け取るようにしてもよい。

30

【0183】

さらに、上記実施例では、ファイルの新規作成処理の際には、ファイル分配制御部214によってファイルの格納先媒体の種類が決定されているが、これに代えて、クライアント400からファイルが送信される際に、クライアント400から格納先媒体の種類が指示されるようにしてもよい。

【0184】

また、上記実施例では、ファイルの移動処理の際には、ファイル分配制御部214によって移動対象ファイル等が決定されているが、これに代えて、クライアント400によって移動対象ファイル等が指示されるようにしてもよい。

40

【0185】

なお、上記実施例では、管理サーバ200と管理端末300とは異なる装置であるが、これに代えて、管理サーバと管理端末とは同一の装置であっても良い。

【図面の簡単な説明】

【0186】

【図1】ネットワークシステムの概略構成を示す説明図である。

【図2】ストレージシステム100の内部構成を示す説明図である。

50

【図 3】予め準備されたボリューム管理テーブル T B の内容を示す説明図である。

【図 4】管理サーバ 2 0 0 の内部構成を示す説明図である。

【図 5】ファイルシステムの作成手順を示すフローチャートである。

【図 6】ステップ S 1 1 0 (図 5) で第 1 L U (L U N = 0) 上に作成されるファイルシステムの構造を示す説明図である。

【図 7】図 5 の処理が完了したときに得られるボリューム管理テーブル T B の内容を示す説明図である。

【図 8】第 1 L U (L U N = 0) 上のファイルシステムの構造を示す説明図である。

【図 9】L U 内にファイルを新規に作成する際の手順を示すフローチャートである。

【図 1 0】統合 L U 内でファイルを移動させる際の手順を示すフローチャートである。

10

【図 1 1】ファイルの移動処理の前後のファイルシステムの構造を示す説明図である。

【図 1 2】第 1 実施例におけるファイルシステムと比較例におけるファイルシステムとを模式的に示す説明図である。

【図 1 3】第 2 実施例における第 1 L U (L U N = 0) 上のファイルシステムの構造を示す説明図である。

【図 1 4】第 2 実施例における統合 L U 内でファイルを移動させる際の手順を示すフローチャートである。

【図 1 5】第 2 実施例におけるファイルの移動処理の前後のファイルシステムの構造を示す説明図である。

【図 1 6】第 3 実施例における第 1 L U (L U N = 0) 上のファイルシステムの構造を示す説明図である。

20

【図 1 7】第 3 実施例における L U 内にファイルを新規に作成する際の手順を示すフローチャートである。

【図 1 8】第 3 実施例における統合 L U 内でファイルを移動させる際の手順を示すフローチャートである。

【図 1 9】第 3 実施例におけるファイルの移動処理の前後のファイルシステムの構造を示す説明図である。

【符号の説明】

【 0 1 8 7 】

1 0 0 ... ストレージシステム

30

1 1 0 ... ファイルサーバ

1 1 2 ... C P U

1 1 4 ... メモリ

1 1 6 ... ネットワーク I F 部

1 1 8 ... ストレージ I F 部

1 2 2 ... プロトコル制御部

1 2 4 ... 第 1 構成管理部

1 2 6 ... ファイルシステム制御部

1 2 8 ... キャッシュ領域

1 3 0 ... ストレージ装置

40

1 4 0 ... ディスクアレイコントローラ

1 4 2 ... C P U

1 4 4 ... メモリ

1 4 6 ... ホスト I F 部

1 4 8 ... ドライブ I F 部

1 5 2 ... 入出力処理部

1 5 4 ... 第 2 構成管理部

1 5 6 ... キャッシュ領域

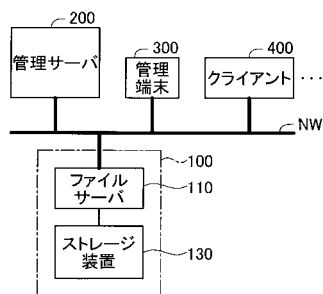
1 6 0 ... ディスクアレイ装置

2 0 0 ... 管理サーバ

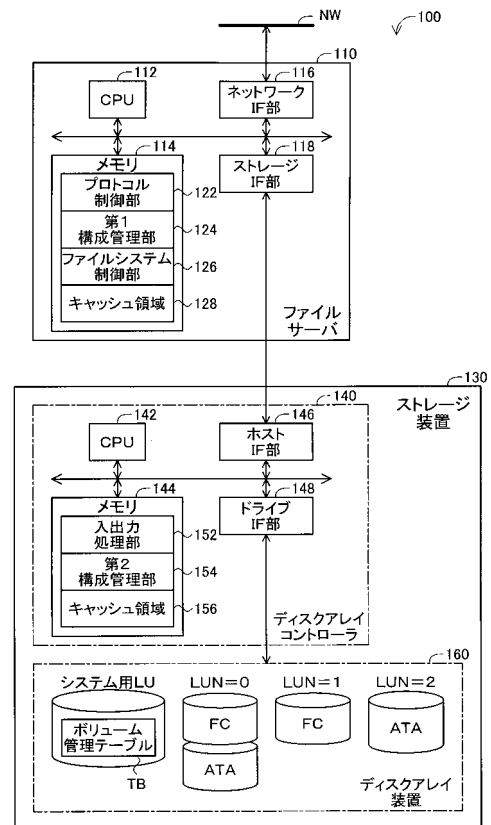
50

202 ... CPU
 204 ... メモリ
 206 ... ネットワーク I/F 部
 212 ... プロトコル制御部
 214 ... ファイル分配制御部
 300 ... 管理端末
 400 ... クライアント
 TB ... ボリューム管理テーブル

【図 1】



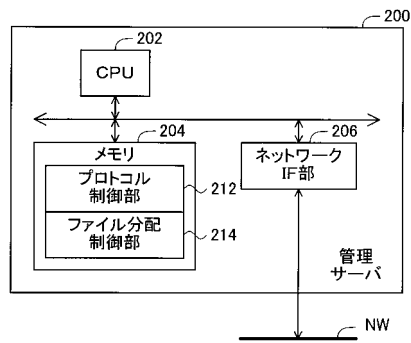
【図 2】



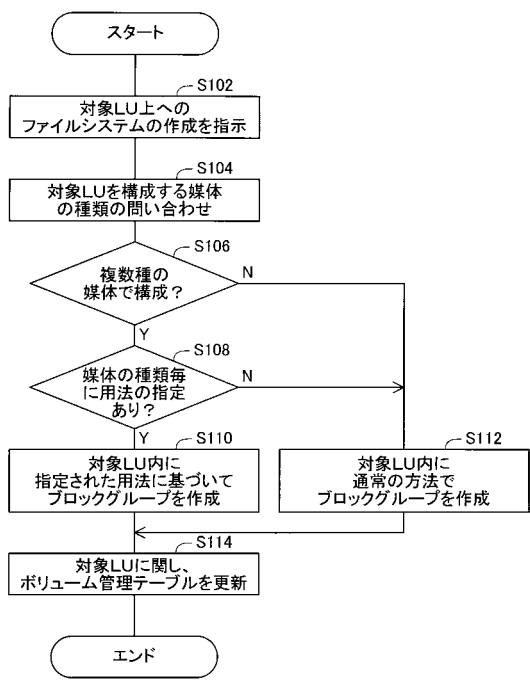
【図3】

ボリューム管理テーブル TB										
LUN	第1記憶領域 の媒体の種類	開始 番地	終了 番地	第2記憶領域 の媒体の種類	開始 番地	終了 番地	...	第3記憶領域 の媒体の種類	開始 番地	終了 番地
0	FC	0	999	ATA	1000	1999	...	—	—	—
1	FC	0	3000	—	—	—	...	—	—	—
2	ATA	0	5000	—	—	—	...	—	—	—

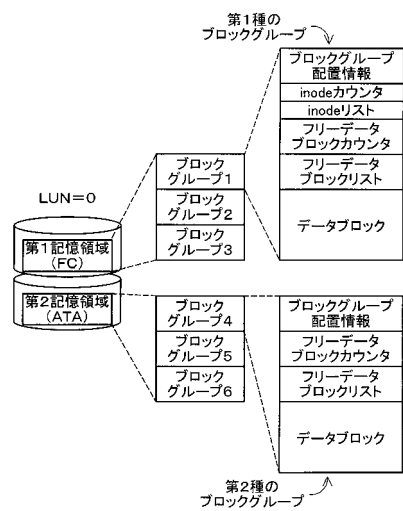
【図4】



【図5】



【図6】

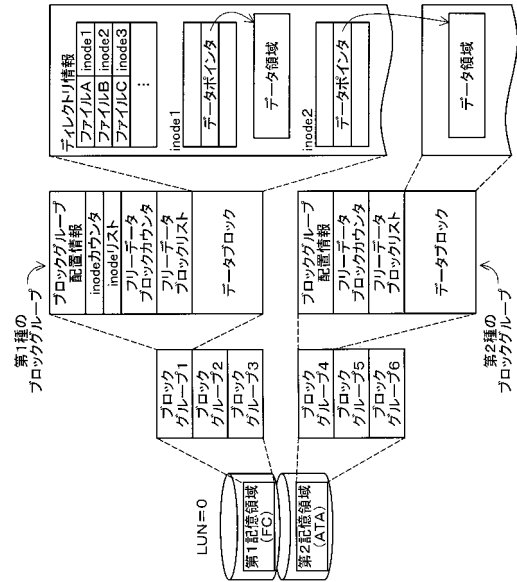


【図 7】

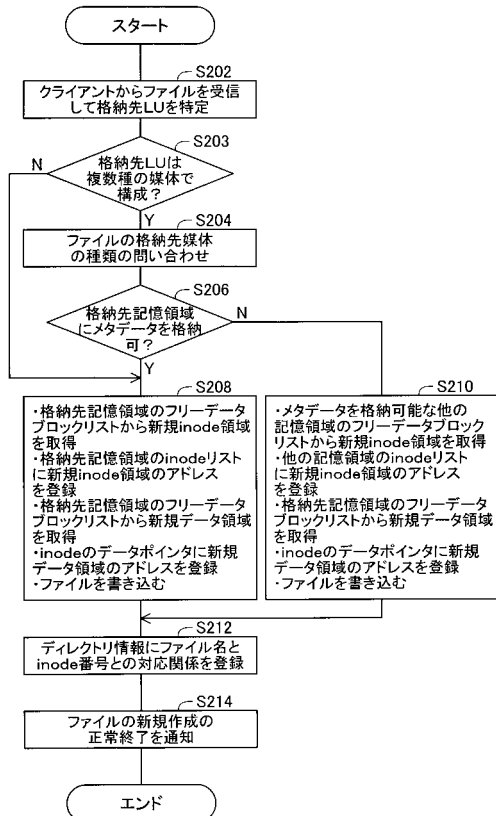
ボリューム管理テーブル TB

LUN	第1記憶領域の媒体の種類	開始番地	終了番地	格納	第2記憶領域の媒体の種類	開始番地	終了番地	格納
0	FC	0	999	×	ATA	1000	1999	×
1	FC	0	3000	○	—	—	—	—
2	ATA	0	5000	○	—	—	—	—

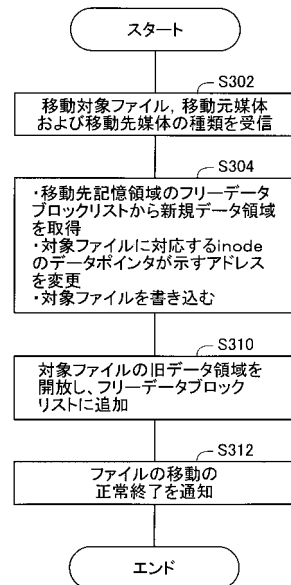
【図 8】



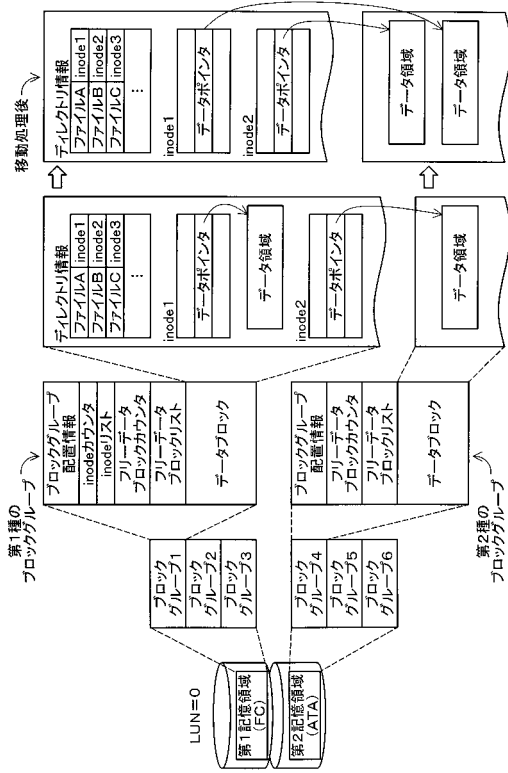
【図 9】



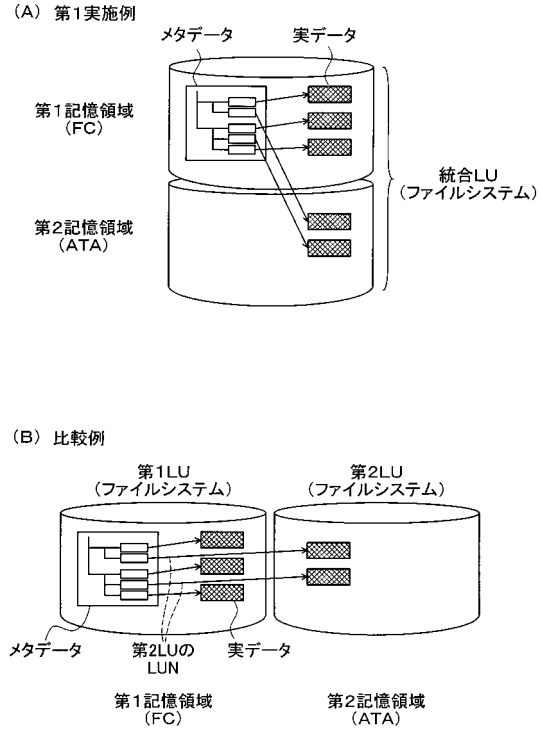
【図 10】



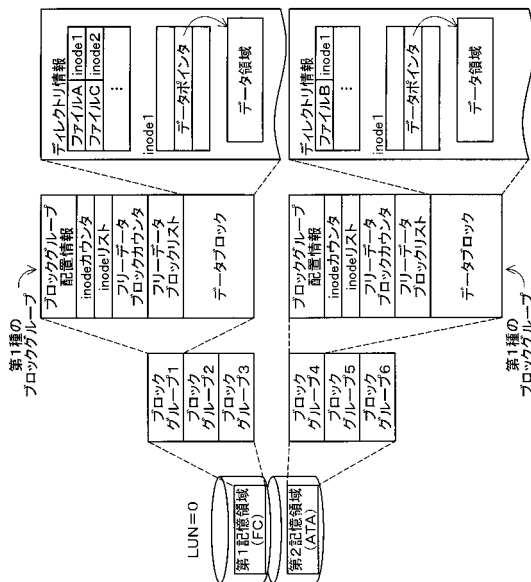
【図 1 1】



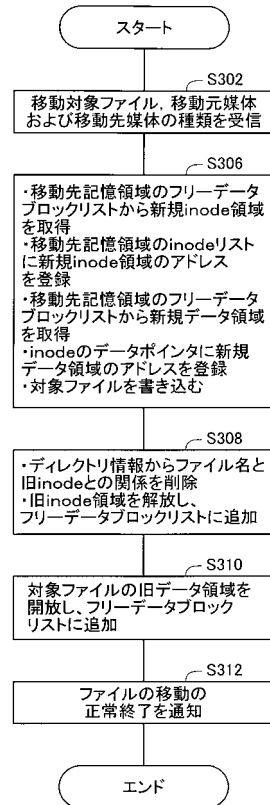
【図 1 2】



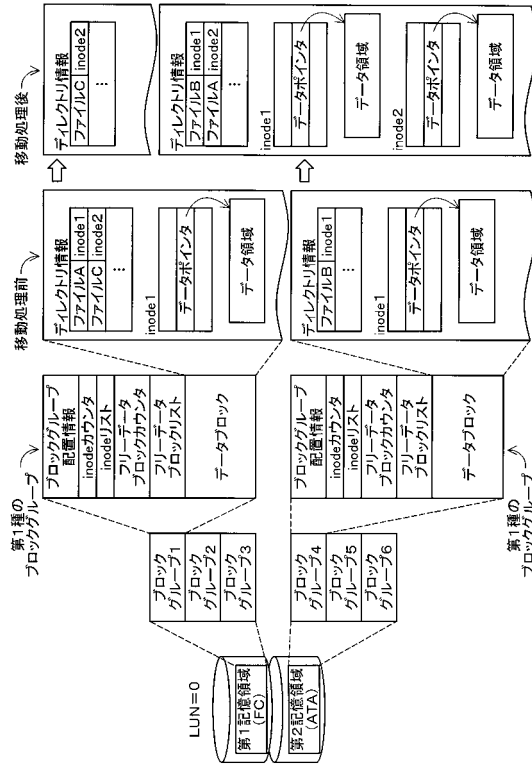
【図 1 3】



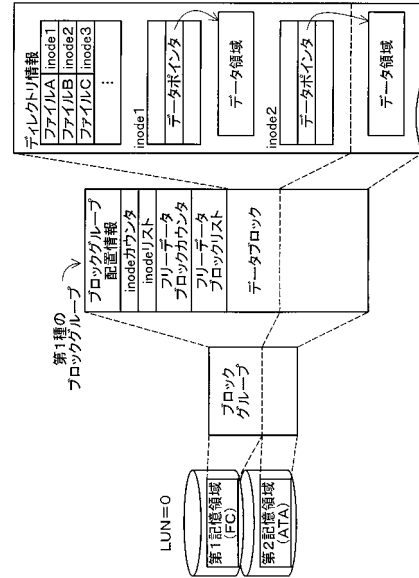
【図 1 4】



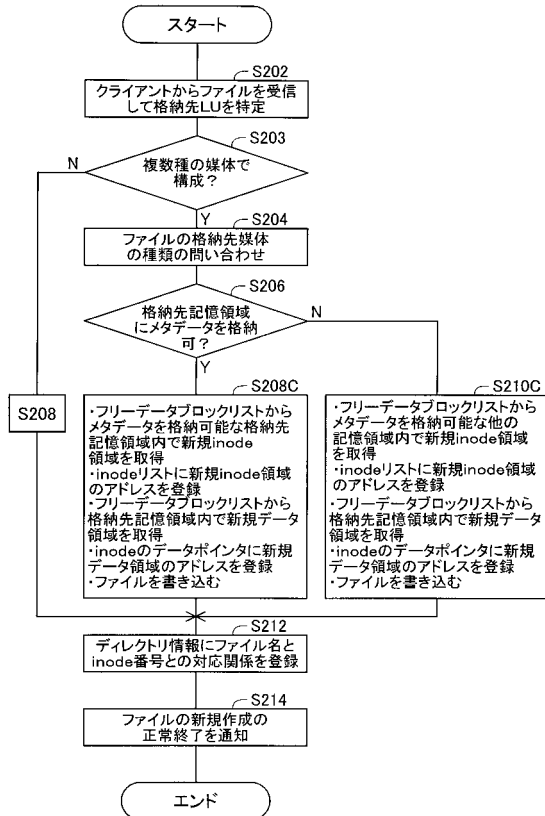
【図 15】



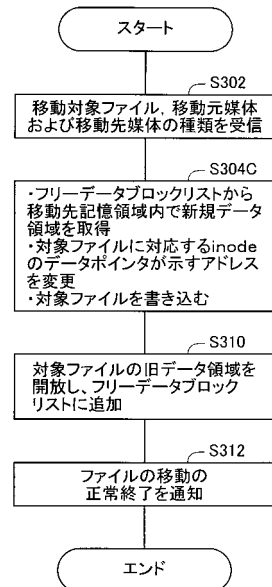
【図 16】



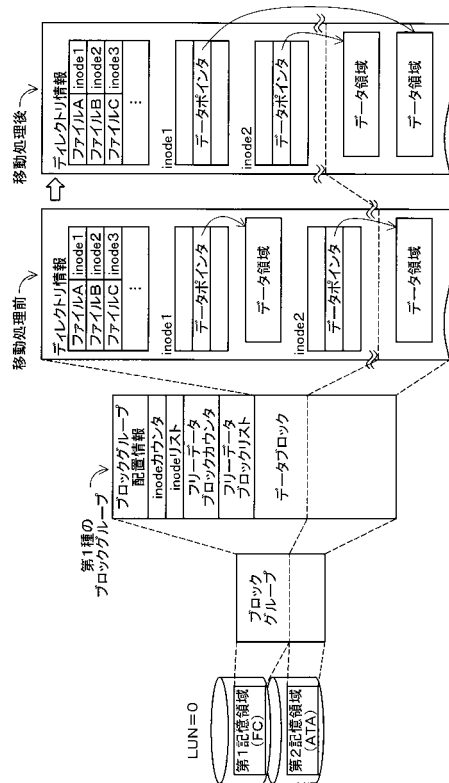
【図 17】



【図 18】



【図 19】



フロントページの続き

(72)発明者 松並 直人

神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所 システム開発研究所内

(72)発明者 園田 浩二

神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所 システム開発研究所内

審査官 工藤 嘉晃

(56)参考文献 特開 2 0 0 5 - 2 2 8 1 7 0 (J P , A)

特開 2 0 0 5 - 0 5 0 0 2 4 (J P , A)

特開 2 0 0 1 - 0 7 5 8 5 8 (J P , A)

特開 2 0 0 3 - 2 5 6 1 4 4 (J P , A)

特開 2 0 0 2 - 0 8 2 7 7 5 (J P , A)

特表 2 0 0 3 - 5 1 6 5 8 2 (J P , A)

特開 2 0 0 4 - 1 7 8 2 8 9 (J P , A)

特開 2 0 0 4 - 2 3 4 5 5 8 (J P , A)

特開 2 0 0 4 - 1 4 5 9 0 1 (J P , A)

(58)調査した分野(Int.Cl. , D B 名)

G 0 6 F 1 2 / 0 0

G 0 6 F 3 / 0 6

J S T P l u s (J D r e a m I I)