(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2007-305012
(P2007-305012A)
(43)公開日　平成19年11月22日(2007.11.22)

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

審査請求 未請求　請求項の数 5　ＯＬ　（全 28 頁）

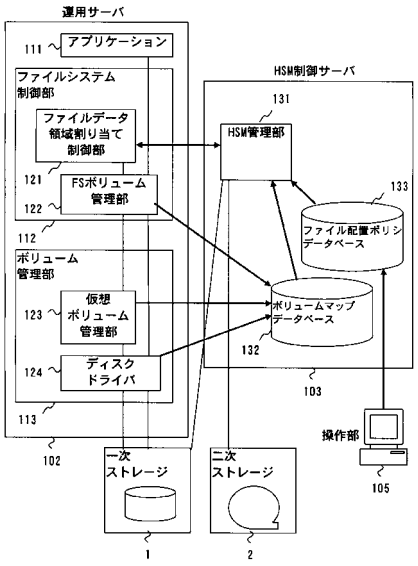
(21)出願番号　特願2006-134693 (P2006-134693)
(22)出願日　平成18年5月15日 (2006.5.15)

(71)出願人　000005223
富士通株式会社
神奈川県川崎市中原区上小田中４丁目１番１号
(74)代理人　100101856
弁理士　赤澤　日出夫
(72)発明者　伊藤　史昭
神奈川県川崎市中原区上小田中４丁目１番１号　富士通株式会社内
(72)発明者　村山　浩
神奈川県川崎市中原区上小田中４丁目１番１号　富士通株式会社内

Ｆターム(参考)　5B065 BA01　CA13　ZA01
5B082 CA13

(54)【発明の名称】ＨＳＭ制御プログラム、ＨＳＭ制御装置、ＨＳＭ制御方法

(57)【要約】
【課題】ＨＳＭにおける一次ストレージに適切なファイルデータ領域を割り当てるためのＨＳＭ制御プログラム、ＨＳＭ制御装置、ＨＳＭ制御方法を提供する。
【解決手段】一次ストレージ１と二次ストレージ２を有するＨＳＭの制御をＨＳＭ制御サーバ１０３に実行させるＨＳＭ制御プログラムであって、一次ストレージ１を構成する少なくとも一つのボリュームに関する情報をボリュームマップデータベース１３２として管理するボリューム情報管理ステップと、一次ストレージ１にファイルデータ領域の割り当てが行われる場合、ボリューム情報管理ステップにより管理されたボリュームマップデータベース１３２と予め設定されたファイル配置ポリシデータベース１３３とに基づいて、割り当ての対象となるボリューム、一次ストレージ１を構成するボリュームの中から選択するボリューム選択ステップとをコンピュータに実行させる。
【選択図】図７



【特許請求の範囲】
【請求項１】
一次ストレージと二次ストレージを有するＨＳＭの制御をコンピュータに実行させるＨＳＭ制御プログラムであって、
前記一次ストレージを構成する少なくとも一つのボリュームに関する情報をボリューム情報として管理するボリューム情報管理ステップと、
前記一次ストレージにファイルデータ領域の割り当てが行われる場合、前記ボリューム情報管理ステップにより管理されたボリューム情報と予め設定されたポリシとに基づいて、前記割り当ての対象となるボリューム、前記一次ストレージを構成するボリュームの中から選択するボリューム選択ステップと
をコンピュータに実行させるＨＳＭ制御プログラム。
【請求項２】
請求項１に記載のＨＳＭ制御プログラムにおいて、
前記ポリシは、前記ボリューム選択ステップにより選択されるボリュームの条件である対象ボリューム条件を含むことを特徴とするＨＳＭ制御プログラム。
【請求項３】
請求項２に記載のＨＳＭ制御プログラムにおいて、
前記ポリシは更に、前記対象ボリューム条件を適用するファイルの条件である対象ファイル条件を含み、
前記ボリューム選択ステップは、前記割り当ての対象のファイルが満たす前記対象ファイル条件を検索し、該対象ファイル条件に対応する前記対象ボリューム条件を選択することを特徴とするＨＳＭ制御プログラム。
【請求項４】
一次ストレージと二次ストレージを有するＨＳＭの制御を行うＨＳＭ制御装置であって、
前記一次ストレージを構成する少なくとも一つのボリュームに関する情報をボリューム情報として管理するボリューム情報管理部と、
前記一次ストレージにファイルデータ領域の割り当てが行われる場合、前記ボリューム情報管理部により管理されたボリューム情報と予め設定されたポリシとに基づいて、前記割り当ての対象となるボリューム、前記一次ストレージを構成するボリュームの中から選択するボリューム選択部と
を備えるＨＳＭ制御装置。
【請求項５】
一次ストレージと二次ストレージを有するＨＳＭの制御を行うＨＳＭ制御方法であって、
前記一次ストレージを構成する少なくとも一つのボリュームに関する情報をボリューム情報として管理するボリューム情報管理ステップと、
前記一次ストレージにファイルデータ領域の割り当てが行われる場合、前記ボリューム情報管理ステップにより管理されたボリューム情報と予め設定されたポリシとに基づいて、前記割り当ての対象となるボリューム、前記一次ストレージを構成するボリュームの中から選択するボリューム選択ステップと
を実行するＨＳＭ制御方法。
【発明の詳細な説明】
【技術分野】
【０００１】
本発明は、ＨＳＭにおける一次ストレージへのファイルデータ領域の割り当ての制御を行うＨＳＭ制御プログラム、ＨＳＭ制御装置、ＨＳＭ制御方法に関するものである。
【背景技術】
【０００２】
ＨＳＭ（Hierarchical Storage Management：階層記憶管理）は、テープライブラリな

どの低速なストレージ装置（二次ストレージ）とハードディスクなどの高速なストレージ装置（一次ストレージ）を組み合わせることにより、安価な大容量ファイルシステムを構築するものである。

【０００３】

ＨＳＭ制御装置においては、一次ストレージにおいて長時間アクセスされていないファイルを特定し、そのファイルを二次ストレージに書き出し、アクセスが要求された時点で一次ストレージに移動することが必要となる。従来、これを実現するために、従来のＨＳＭ制御装置は、階層構造を持つファイルシステムの名前空間を総なめし、ファイルシステムがファイル単位に保持するアクセス時刻を参照することにより、二次ストレージに書き出すファイルを特定する方式を用いている。

【０００４】

なお、本発明の関連ある従来技術として、例えば、下記に示す特許文献１が知られている。このデータ処理装置は、メタデータデータの内容が更新されると、ログが採取され、このログを用いてファイルシステムの不整合の修正を行うものである。

【特許文献１】特開２０００－４８４９９５号公報

【発明の開示】

【発明が解決しようとする課題】

【０００５】

しかしながら、上述した名前空間を総なめする方式のＨＳＭ制御装置には、以下の問題がある。

【０００６】

第１にファイルシステム総なめオーバヘッドの問題がある。従来のＨＳＭ制御装置では階層構造を持つファイル名前空間を定期的に総なめするために、オーバヘッドが大きくなってしまう。

【０００７】

第２に名前空間の排他問題がある。ＨＳＭ制御装置が名前空間を総なめしている間に、ｒｅｎａｍｅ操作などのファイル名変更操作が行われると、総なめの過程で求めたパス名が、実際には存在しない不当なものとなってしまう。このため、ＨＳＭ制御装置は、顧客が設定したポリシと矛盾するデータ移動操作を行ってしまう可能性がある。例えば、検索の途中で、上位ディレクトリがゴミ箱に移されたとすると、ゴミ箱全体を移動対象としてしまうようなことが起こる。こうした問題を防ごうとすると、ＨＳＭ制御装置は総なめの過程で、頻繁に矛盾をチェックし、矛盾があれば総なめをやり直すことが必要となり、論理が非常に複雑となるとともにオーバヘッドが大幅に増加する。

【０００８】

第３にＨＳＭポリシ制御の柔軟性がある。一般に階層構造の名前空間は格納されているファイル群の性格を表しているため、ＨＳＭポリシも名前空間に基づいて設定する（あるディレクトリ以下の全ファイルなど）のが自然である。しかし、上述した名前空間の排他問題により、名前空間に基づく複雑なポリシ制御を実現することが難しいという問題があった。

【０００９】

第４に二次ストレージに退避されたデータの属性情報不足の問題がある。また上述した名前空間の排他問題により、二次ストレージに格納されるデータに正しいパス名を付加することが難しい。このため、二次ストレージに格納されたデータはファイルシステムのメタデータのみからしかアクセスできないことになり、ファイルシステムのメタデータが壊れると、二次ストレージ上にデータは残っているのにもかかわらず、パス名と対応づけることができないため、ファイルデータを復旧することができないという問題があった。

【００１０】

本発明は上述した問題点を解決するためになされたものであり、ＨＳＭにおける一次ストレージに適切なファイルデータ領域を割り当てるためのＨＳＭ制御プログラム、ＨＳＭ制御装置、ＨＳＭ制御方法を提供することを目的とする。

【課題を解決するための手段】
【００１１】
上述した課題を解決するため、本発明は、一次ストレージと二次ストレージを有するＨＳＭの制御をコンピュータに実行させるＨＳＭ制御プログラムであって、前記一次ストレージを構成する少なくとも一つのボリュームに関する情報をボリューム情報として管理するボリューム情報管理ステップと、前記一次ストレージにファイルデータ領域の割り当てが行われる場合、前記ボリューム情報管理ステップにより管理されたボリューム情報と予め設定されたポリシとに基づいて、前記割り当ての対象となるボリューム、前記一次ストレージを構成するボリュームの中から選択するボリューム選択ステップとをコンピュータに実行させるものである。
【００１２】
また、本発明に係るＨＳＭ制御プログラムにおいて、前記ポリシは、前記ボリューム選択ステップにより選択されるボリュームの条件である対象ボリューム条件を含むことを特徴とするものである。
【００１３】
また、本発明に係るＨＳＭ制御プログラムにおいて、前記ポリシは更に、前記対象ボリューム条件を適用するファイルの条件である対象ファイル条件を含み、前記ボリューム選択ステップは、前記割り当ての対象のファイルが満たす前記対象ファイル条件を検索し、該対象ファイル条件に対応する前記対象ボリューム条件を選択することを特徴とするものである。
【００１４】
また、本発明に係るＨＳＭ制御プログラムにおいて、前記対象ファイル条件は、ファイル識別子とファイル名パターン指定の少なくともいずれかを含むことを特徴とするものである。
【００１５】
また、本発明に係るＨＳＭ制御プログラムにおいて、前記対象ボリューム条件は、新規のファイルが前記一次ストレージに作成されるケースであるファイル作成時に選択されるボリュームの条件であるファイル作成先ボリューム条件と、前記二次ストレージに退避済みのファイルが前記一次ストレージに復元されるケースであるファイル復元時に選択されるボリュームの条件であるファイル復元先ボリューム条件とを含むことを特徴とするものである。
【００１６】
また、本発明に係るＨＳＭ制御プログラムにおいて、前記ファイル作成先ボリューム条件は、前記一次ストレージを構成するボリュームのうち、信頼性の高いボリュームなど、ファイル作成時要件を満たすボリュームであることを特徴とするものである。
【００１７】
また、本発明に係るＨＳＭ制御プログラムにおいて、前記ファイル復元先ボリューム条件は、前記一次ストレージを構成するボリュームのうち、アクセス性能の高いボリュームなど、ファイル復元時要件を満たすボリュームであることを特徴とするものである。
【００１８】
また、本発明に係るＨＳＭ制御プログラムにおいて、前記ボリューム選択ステップは、前記割り当ての対象となるボリュームに対応するブロック範囲を出力することを特徴とするものである。
【００１９】
また、本発明に係るＨＳＭ制御プログラムにおいて、前記ボリュームは、仮想ボリュームまたはディスク装置で構成され、前記ボリューム情報は、前記仮想ボリュームの情報と前記ボリュームに含まれるディスク装置の情報とを含むことを特徴とするものである。
【００２０】
また、本発明は、一次ストレージと二次ストレージを有するＨＳＭの制御を行うＨＳＭ制御装置であって、前記一次ストレージを構成する少なくとも一つのボリュームに関する

情報をボリューム情報として管理するボリューム情報管理部と、前記一次ストレージにファイルデータ領域の割り当てが行われる場合、前記ボリューム情報管理部により管理されたボリューム情報と予め設定されたポリシとに基づいて、前記割り当ての対象となるボリューム、前記一次ストレージを構成するボリュームの中から選択するボリューム選択部とを備えたものである。

【００２１】

また、本発明に係るＨＳＭ制御装置において、前記ポリシは、前記ボリューム選択部により選択されるボリュームの条件である対象ボリューム条件を含むことを特徴とするものである。

【００２２】

また、本発明に係るＨＳＭ制御装置において、前記ポリシは更に、前記対象ボリューム条件を適用するファイルの条件である対象ファイル条件を含み、前記ボリューム選択部は、前記割り当ての対象のファイルが満たす前記対象ファイル条件を検索し、該対象ファイル条件に対応する前記対象ボリューム条件を選択することを特徴とするものである。

【００２３】

また、本発明に係るＨＳＭ制御プログラムにおいて、前記対象ファイル条件は、ファイル識別子とファイル名パターン指定の少なくともいずれかを含むことを特徴とするものである。

【００２４】

また、本発明に係るＨＳＭ制御装置において、前記対象ボリューム条件は、新規のファイルが前記一次ストレージに作成されるケースであるファイル作成時に選択されるボリュームの条件であるファイル作成先ボリューム条件と、前記二次ストレージに退避済みのファイルが前記一次ストレージに復元されるケースであるファイル復元時に選択されるボリュームの条件であるファイル復元先ボリューム条件とを含むことを特徴とするものである。

【００２５】

また、本発明に係るＨＳＭ制御装置において、前記ファイル作成先ボリューム条件は、前記一次ストレージを構成するボリュームのうち、信頼性の高いボリュームなど、ファイル作成時要件を満たすボリュームであることを特徴とするものである。

【００２６】

また、本発明に係るＨＳＭ制御装置において、前記ファイル復元先ボリューム条件は、前記一次ストレージを構成するボリュームのうち、アクセス性能の高いボリュームなど、ファイル復元時要件を満たすボリュームであることを特徴とするものである。

【００２７】

また、本発明に係るＨＳＭ制御装置において、前記ボリューム選択部は、前記割り当ての対象となるボリュームに対応するブロック範囲を出力することを特徴とするものである。

【００２８】

また、本発明に係るＨＳＭ制御装置において、前記ボリュームは、仮想ボリュームまたはディスク装置で構成され、前記ボリューム情報は、前記仮想ボリュームの情報と前記ボリュームに含まれるディスク装置の情報とを含むことを特徴とするものである。

【００２９】

また、本発明は、一次ストレージと二次ストレージを有するＨＳＭの制御を行うＨＳＭ制御方法であって、前記一次ストレージを構成する少なくとも一つのボリュームに関する情報をボリューム情報として管理するボリューム情報管理ステップと、前記一次ストレージにファイルデータ領域の割り当てが行われる場合、前記ボリューム情報管理ステップにより管理されたボリューム情報と予め設定されたポリシとに基づいて、前記割り当ての対象となるボリューム、前記一次ストレージを構成するボリュームの中から選択するボリューム選択ステップとを実行するものである。

【発明の効果】

【００３０】
本発明によれば、ＨＳＭにおける一次ストレージに適切なファイルデータ領域を割り当てることができる。
【発明を実施するための最良の形態】
【００３１】
以下、本発明の実施の形態を説明するため、その前提となる技術について図面を参照しつつ説明する。
【００３２】
（前提技術１）
前提技術１においては、ＨＳＭ制御装置であるサーバについて説明する。
【００３３】
まず、前提技術１に係るサーバを有するＨＳＭ装置の構成について説明する。
【００３４】
図１は、前提技術１に係るＨＳＭ装置の構成の一例を示すブロック図である。最近アクセスされたファイルを格納しているディスク装置などの高速ストレージ装置である一次ストレージ１、および長時間アクセスされていないファイルデータが格納されるテープライブラリ装置などの低速ストレージ装置である二次ストレージ２と、前提技術１に係るＨＳＭ制御装置であり、ファイルデータをアクセスするアプリケーションが動作するサーバ３から構成される。
【００３５】
また、サーバ３は、アプリケーション部１１、ファイルシステム制御部１２、名前空間複製部１３、名前空間追随部１４、名前空間複製ＤＢ（Database）１５、マイグレート決定部１６を備える。また、ファイルシステム制御部１２は、イベントデータ記録部２１を備える。
【００３６】
次に、サーバ３の各部について説明する。
【００３７】
イベントデータ記録部２１は、アプリケーションプログラムが発行したファイル操作要求の履歴をイベントデータとして蓄積するファイルシステム制御部１２内に配置されるプログラムである。イベントデータ記録部２１は、アプリケーション部１１が発行したファイル操作要求の内容をイベントデータに変換してメモリ上に蓄積しておき、一定量たまったところで名前空間複製部１３や名前空間追随部１４に渡す。イベントデータの受け渡しは、通信を使用してもよいし、専用のファイルを介して受け渡してもよい。
【００３８】
名前空間複製部１３は、アプリケーション部１１の動作と平行して、ファイルシステムの名前空間の複製を行うプログラムである。名前空間複製部１３は、ファイルシステムの名前空間をたどり、存在するファイルのファイル情報を取得する。このファイル情報と、ファイル情報取得中にイベントデータ記録部２１から受け取ったイベントデータを組み合わせて、名前空間の初期複製を名前空間複製ＤＢ１５として完成させる。
【００３９】
名前空間追随部１４は、名前空間の初期複製が完成した後、イベントデータ記録部２１から受け取ったイベントデータに従って複製を更新し、名前空間複製ＤＢ１５を最新の状態に維持する機能を受け持つ。また、名前空間追随部１４は、通知されたファイルアクセスやアーカイブ状態を名前空間複製ＤＢ１５に反映する役割も担う。
【００４０】
マイグレート決定部１６は、ポリシ制御の一例として、名前空間複製部１３が設定したファイルアクセス記録とユーザが設定したポリシに従い、一次ストレージ１において長時間アクセスされていないファイルを二次ストレージ２に追い出すため、ファイルシステム制御部１２に指示を出すプログラムである。通常、二次ストレージ２に追い出された（マイグレートされた）ファイルは、アプリケーション部１１がそのファイルをアクセスした

ときに、ファイルシステム制御部１２が二次ストレージ２から一次ストレージ１に戻す（リコール）。また、ファイルを更新したタイミングで、ファイルシステム制御部１２により二次ストレージ２上のデータ（アーカイブデータ）が無効化される。二次ストレージ２上のデータはこのタイミングでは消えず、二次ストレージ２が不足するまで、バックアップデータとして残され、ファイルシステム障害時などのリカバリで使われる。

【００４１】

次に、イベントデータ、ファイル情報、名前空間複製ＤＢ１５の詳細について説明する。

【００４２】

まず、イベントデータについて説明する。

【００４３】

イベントデータ記録部２１により作成されるイベントデータ（ｅｖｅｎｔ）はファイルやディレクトリの生成や削除、ファイル名の変更、ファイルアクセス、アーカイブ状態変化などのファイル操作の内容を表しており、操作名と操作が行われた時刻に加え、それぞれ以下のデータを含む。ここで、アーカイブ状態変化とは、アーカイブデータの有効化・無効化、マイグレート、リコールなどの事象を含む。

【００４４】

（１） ファイルあるいはディレクトリの作成

ｅｖｅｎｔ．ｒｅｃｔｙｐｅ ＝ ｃｒｅａｔｅ

ｅｖｅｎｔ．ｍ＿ｉｎｏｄｅ＃ ＝ 親ディレクトリのｉｎｏｄｅ番号

ｅｖｅｎｔ．ｆｔｙｐｅ ＝

ｄｉｒ （ｍｋｄｉｒ時）あるいはｆｉｌｅ （ｃｒｅａｔｅ時）

ｅｖｅｎｔ．ｆｎａｍｅ ＝ 作成されたファイルの名前

ｅｖｅｎｔ．ｉｎｏｄｅ＃ ＝

作成されたファイルあるいはディレクトリのｉｎｏｄｅ番号

ｅｖｅｎｔ．ｔｉｍｅ ＝ このイベントが発生した時刻

【００４５】

（２） ファイルあるいはディレクトリの削除

ｅｖｅｎｔ．ｒｅｃｔｙｐｅ ＝ ｄｅｌｅｔｅ

ｅｖｅｎｔ．ｍ＿ｉｎｏｄｅ＃ ＝ 親ディレクトリのｉｎｏｄｅ番号

ｅｖｅｎｔ．ｆｔｙｐｅ ＝ ｄｉｒ （ｒｍｄｉｒ時）あるいはｆｉｌｅ （ｒｅｍｏｖｅ時）

ｅｖｅｎｔ．ｉｎｏｄｅ＃ ＝ 削除されたファイルあるいはディレクトリのｉｎｏｄｅ番号

ｅｖｅｎｔ．ｔｉｍｅ ＝ このイベントが発生した時刻

【００４６】

（３） ファイル名の変更

ｅｖｅｎｔ．ｒｅｃｔｙｐｅ ＝ ｒｅｎａｍｅ

ｅｖｅｎｔ．ｍ＿ｉｎｏｄｅ＃ ＝ 親ディレクトリのｉｎｏｄｅ番号

ｅｖｅｎｔ．ｆｔｙｐｅ ＝

ｄｉｒ（対象がディレクトリの場合）

あるいはｆｉｌｅ（対象がファイルの場合）

ｅｖｅｎｔ．ｉｎｏｄｅ＃ ＝

対象のファイルあるいはディレクトリのｉｎｏｄｅ番号

ｅｖｅｎｔ．ｔａｒｇｅｔ．ｍ＿ｉｎｏｄｅ＃ ＝

移動先ディレクトリのｉｎｏｄｅ番号

ｅｖｅｎｔ．ｔａｒｇｅｔ．ｆｎａｍｅ ＝

変更後のファイルあるいはディレクトリ名

ｅｖｅｎｔ．ｔｉｍｅ ＝ このイベントが発生した時刻

【００４７】

（４）　ファイルアクセス（アプリケーションプログラムがファイルをｒｅａｄ／ｗｒｉｔｅ）
ｅｖｅｎｔ．ｒｅｃｔｙｐｅ　＝　ａｃｃｅｓｓ
ｅｖｅｎｔ．ｉｎｏｄｅ＃　＝　ファイルのｉｎｏｄｅ番号
ｅｖｅｎｔ．ｔｉｍｅ　＝　このイベントが発生した時刻
【００４８】
（５）　アーカイブ状態変化
ｅｖｅｎｔ．ｒｅｃｔｙｐｅ　＝　ａｒｃｈｉｖｅ
ｅｖｅｎｔ．ｉｎｏｄｅ＃　＝　ファイルのｉｎｏｄｅ番号
ｅｖｅｎｔ．ｍｉｇｒａｔｅ　＝
　オン（マイグレート状態となった）
　あるいはオフ（リコールが起動され、マイグレート状態でなくなった）
ｅｖｅｎｔ．ａｒｃｈｉｖｅ　＝
　オン（二次記憶へのファイルデータの書き出しが完了し、アーカイブデータが有効となった）
　あるいはオフ（ファイルが更新された結果アーカイブデータが無効となった）
ｅｖｅｎｔ．ｔｉｍｅ　＝　このイベントが発生した時刻
【００４９】
　次に、ファイル情報について説明する。
【００５０】
　名前空間複製復元中にファイルシステムから取得するファイル情報（ｆｓｔａｔ）には、以下のものがある。
ｆｓｔａｔ．ｍ＿ｉｎｏｄｅ＃　＝　親ディレクトリのｉｎｏｄｅ番号
ｆｓｔａｔ．ｆｔｙｐｅ　＝　ｄｉｒ（対象がディレクトリの場合）
　あるいはｆｉｌｅ　（対象がファイルの場合）
ｆｓｔａｔ．ｆｎａｍｅ　＝　ディレクトリあるいはファイルの名前
ｆｓｔａｔ．ｉｎｏｄｅ＃　＝
　ファイルあるいはディレクトリのｉｎｏｄｅ番号
ｆｓｔａｔ．ａｒｃｈｉｖｅ　＝　オン（アーカイブデータが有効なとき）
　あるいはオフ（アーカイブデータが無効なとき）
ｆｓｔａｔ．ｍｉｇｒａｔｅ　＝　オン（マイグレート状態のとき）
　あるいはオフ（マイグレートされていないとき）
ｆｓｔａｔ．ａｔｉｍｅ　＝　ファイルを最後にアクセスした時刻
ｆｓｔａｔ．ｔｉｍｅ　＝　ファイル情報取得時刻
【００５１】
　次に、名前空間複製ＤＢ１５の構成について説明する。
【００５２】
　名前空間複製ＤＢ１５は、以下のカラム（ｄｂｅ）を持つ、ディレクトリに設定されているファイルあるいはディレクトリ要素ごとにタプルを持つリレーショナルＤＢである。
【００５３】
ｄｂｅ．ｍ＿ｉｎｏｄｅ＃　＝　親ディレクトリのｉｎｏｄｅ番号
ｄｂｅ．ｆｔｙｐｅ　＝　ｄｉｒ（このタプルがディレクトリをあらわすとき）　あるいは　ｆｉｌｅ　（このタプルがファイルを表すとき）
ｄｂｅ．ｆｎａｍｅ　＝　ファイルあるいはディレクトリの名前
ｄｂｅ．ｉｎｏｄｅ＃　＝　ファイルあるいはディレクトリのｉｎｏｄｅ番号
ｄｂｅ．ａｒｃｈｉｖｅ　＝　オン（アーカイブデータが有効なとき）
　あるいはオフ（アーカイブデータが無効なとき）
ｄｂｅ．ｍｉｇｒａｔｅ　＝　オン（マイグレート状態のとき）
　あるいはオフ（マイグレートされていないとき）
ｄｂｅ．ａｔｉｍｅ　＝　ファイルを最後にアクセスした時刻

ｄｂｅ．ａｃｔｉｖｅ　＝　オン（ファイル情報取得済みのとき）
あるいはオフ（まだファイル情報を取得していないとき）

【００５４】

次に、サーバ３の動作について説明する。

【００５５】

図２は、前提技術１に係るファイル情報取得処理の動作の一例を示すフローチャートである。サーバ３は、名前空間複製処理（Ｓ１１）、名前空間追随処理（Ｓ１２）、マイグレート処理（Ｓ１３）を実行する。

【００５６】

次に、サーバ３における動作の詳細について説明する。

【００５７】

まず、名前空間複製処理について説明する。

【００５８】

名前空間複製処理は、名前空間複製部１３が名前空間の初期複製を作成する処理であり、ファイル情報取得処理とイベントデータ反映処理からなる。また、名前空間複製処理は、障害発生後のサーバ再立ち上げ時など、メモリ上に蓄積されていたイベントデータが失われ、名前空間複製ＤＢ１５の内容がファイルシステムの最新状態を反映できなくなったときに、名前空間複製ＤＢ１５を再作成する目的で動作する。このように名前空間複製ＤＢ１５を動的に再作成する構成では、イベントデータをイベント発生時に不揮発化する必要がなく、小さい容量のメモリに蓄積するのみで良く、後の名前空間複製ＤＢの追随のオーバヘッドを削減することができる。

【００５９】

名前空間複製部１３は、まず、ファイル情報取得処理として、親ディレクトリをオープンし、子ファイル名あるいは子ディレクトリ名を引数として指定し、ファイルシステムの情報取得機能（ｇｅｔｉｎｆｏ）を発行することにより求める。また、名前空間複製部１３は、パス名昇順（あるいは降順）とした名前空間をたどることにより、ファイルシステム内に存在するディレクトリ、ファイルの情報を漏れなく求める。この過程で見逃したものは、イベントデータとして記録されているので、後で補正する。

【００６０】

図３は、名前空間におけるディレクトリの階層構造の一例を示す図である。この名前空間は、ディレクトリの階層構造を持ち、ディレクトリ名やファイル名を昇順に左から右へソートしたものである。図４は、前提技術１に係るファイル情報取得処理の動作の一例を示すフローチャートである。

【００６１】

まず、名前空間複製部１３は、対象ファイルシステムのルートディレクトリを基点とし、ディレクトリを左下方向（ディレクトリ名の昇順）に順にたどり、最も左下のディレクトリを見つける。見つけた最も左下のディレクトリを対象ディレクトリとし、検索の過程で求めた対象ディレクトリのパス名を対象ディレクトリパス名とする（Ｓ２０１）。次に、名前空間複製部１３は、対象ディレクトリのファイル情報および対象ディレクトリ内に存在する全ファイルのファイル情報をファイル名昇順にひとつずつ順に求め、ファイル情報記録ファイルの末尾に順に書き込む（Ｓ２０２）。次に、名前空間複製部１３は、対象ディレクトリがルートディレクトリであるか否かの判断を行う（Ｓ２０３）。対象ディレクトリがルートディレクトリである場合（Ｓ２０３，Ｙ）、全ファイルを処理し終わったことを意味するのでファイル情報取得処理を終了する。

【００６２】

一方、対象ディレクトリがルートディレクトリでない場合（Ｓ２０３，Ｎ）、名前空間複製部１３は、対象ディレクトリパス名から、対象ディレクトリのひとつ上のディレクトリパス名を求める、すなわち、パス名を構成する最終構成ディレクトリ名を取り除いたパス名を新しいパス名とする。次に、名前空間複製部１３は、求めたディレクトリパス名をルートディレクトリから下方に順に再度検索し、この検索で存在を確認できた最終ディレ

クトリを基点ディレクトリとする（Ｓ２０５）。パスの途中のディレクトリがｒｅｎａｍｅなどで名前空間の別の位置に動かされている場合、途中でみつからなくなるが、この部分は後続のファイル情報取得処理で見つかるか、イベントデータで必ず通知され、後で補正されるため、無視しても問題ない。

【００６３】

次に、名前空間複製部１３は、基点ディレクトリの内容を読み込み、基点ディレクトリ内に未処理のディレクトリがあるか否かの判断を行う（Ｓ２０６）。未処理のディレクトリがある場合（Ｓ２０６，Ｙ）、名前空間複製部１３は、未処理の最も左下のディレクトリを求め、これを対象ディレクトリとし（Ｓ２０７）、処理Ｓ２０２に移行する。未処理のディレクトリが存在しない、すなわち基点ディレクトリ内に対象ディレクトリパス名で示されるより大きなファイル名をもつディレクトリが存在しない場合（Ｓ２０６，Ｎ）、対象ディレクトリパス名を基点ディレクトリのパス名に設定し（Ｓ２０８）、処理Ｓ２０３に移行する。

【００６４】

次に、名前空間複製部１３は、対象ファイルシステムのファイル情報取得処理が全て終了すると、その間に発生したイベントデータをファイル情報に反映するイベントデータ反映処理を行う。ファイル情報記録ファイルを先頭から順によみ、ファイル情報記録ファイルに記録されている全てのファイル情報を処理したら、イベントデータ反映処理は終了する。

【００６５】

図５は、前提技術１に係るイベントデータ反映処理の動作の一例を示すフローチャートである。まず、名前空間複製部１３は、未処理のファイル情報を取り出し（Ｓ３０２）、ファイル情報に設定されていた情報取得時刻以前の時刻を持つ、イベントデータを順に取り出し、名前空間複製ＤＢ１５に反映する（Ｓ３０３）。

【００６６】

ここで、名前空間複製ＤＢ１５への反映について、イベントデータが、削除系、生成系、ファイル名の変更、ファイルアクセス、アーカイブ状態変化のそれぞれの場合について説明する。

【００６７】

イベントデータが削除系（ファイル削除，ディレクトリ削除）の場合、名前空間複製部１３は、削除対象ファイルあるいはディレクトリが既に名前空間複製ＤＢ１５に登録済みなら削除する。そうでなければ何もしない。ここで、以下の条件を全て満たすエントリが存在する場合、登録済みとみなす。

【００６８】

ｄｂｅ．ｉｎｏｄｅ＃　＝＝　ｅｖｅｎｔ．ｉｎｏｄｅ＃
ｄｂｅ．ｍ＿ｉｎｏｄｅ＃　＝＝　ｅｖｅｎｔ．ｍ＿ｉｎｏｄｅ＃
ｄｂｅ．ｆｎａｍｅ　＝＝　ｅｖｅｎｔ．ｆｎａｍｅ

【００６９】

イベントデータが生成系（ファイル生成，ディレクトリ生成）の場合、名前空間複製部１３は、作成されたファイルあるいはディレクトリが名前空間複製ＤＢ１５に登録済みでなければ情報取得済みで登録する。登録済みならこのイベントデータを無視し、何もしない。ここで、以下の条件を全て満たすエントリが存在する場合、登録済みとみなす。

【００７０】

ｄｂｅ．ｉｎｏｄｅ＃　＝＝　ｅｖｅｎｔ．ｉｎｏｄｅ＃
ｄｂｅ．ｍ＿ｉｎｏｄｅ＃　＝＝　ｅｖｅｎｔ．ｍ＿ｉｎｏｄｅ＃
ｄｂｅ．ｆｎａｍｅ　＝＝　ｅｖｅｎｔ．ｆｎａｍｅ

【００７１】

未登録時の設定内容を以下に示す。

【００７２】

ｄｂｅ．ｍ＿ｉｎｏｄｅ＃　＝　ｅｖｅｎｔ．　ｍ＿ｉｎｏｄｅ＃

dbe.ftype = event.ftype
dbe.fname = event.fname
dbe.inode# = event.inode#
dbe.archive = オフ
dbe.migrate = オフ
dbe.atime = event.time
dbe.active = オン

【００７３】

イベントデータがファイル名の変更（event.rectype == rename）の場合、名前空間複製部１３は、改名後と同じ名前をもつファイルあるいはディレクトリがすでに登録されていた場合（ファイル名と親ｉｎｏｄｅ番号で評価）、そのエントリを名前空間複製ＤＢ１５から削除する。ここで、以下の条件を全て満たすエントリが存在する場合、登録済みとみなす。

【００７４】

dbe.name == event.target.fname
dbe.m_inode# ==
event.target.m_inode#
dbe.fname == event.target.fname

【００７５】

ここで、対象ファイルが名前空間複製ＤＢ１５に既に登録されているならそのエントリの親情報とファイル名を変更する。ここで、以下の条件を全て満たすエントリが存在する場合、登録済みとみなす。

【００７６】

dbe.inode# == event.inode#
dbe.m_inode# == event.m_inode#
dbe.fname == event.fname

【００７７】

このときの変更内容を以下に示す。

【００７８】

dbe.m_inode# = event.target.m_inode#
dbe.name = event.target.fname

【００７９】

ここで、対象ファイルが未登録なら、改名後のファイルを名前空間複製ＤＢ１５に新しいエントリとして登録する。

【００８０】

dbe.inode# = event.inode#
dbe.m_inode# = event.target.m_inode#
dbe.name = event.target.fname
dbe.active = オフ

【００８１】

イベントデータがファイルアクセス（event.rectype == access）の場合、名前空間複製部１３は、対象ｉｎｏｄｅが未登録ならこのイベントデータを無視する。登録されていたら、登録済みのすべてのエントリのファイル最終アクセス時刻、アーカイブ情報、リコール情報を更新（ハードリンクがあるため）する。ここで、以下の条件を全て満たすエントリが存在する場合、登録済みとみなす。

【００８２】

dbe.inode# == event.inode#

【００８３】

このときの変更内容を以下に示す。

【００８４】

ｄｂｅ．ａｔｉｍｅ　＝　ｅｖｅｎｔ．ｔｉｍｅ
【００８５】
　イベントデータがアーカイブ状態変化（ｅｖｅｎｔ．ｒｅｃｔｙｐｅ　＝＝　ａｒｃｈｉｖｅ）の場合、対象ｉｎｏｄｅが未登録ならこのイベントデータを無視。登録されていたら、すべてのエントリのアーカイブ情報を更新（ハードリンクがあるため）する。ここで、以下の条件を全て満たすエントリが存在する場合、登録済みとみなす。
【００８６】
ｄｂｅ．ｉｎｏｄｅ＃　＝＝　ｅｖｅｎｔ．ｉｎｏｄｅ＃
【００８７】
　このときの変更内容を以下に示す。
【００８８】
ｄｂｅ．ａｒｃｈｉｖｅ　＝　ｅｖｅｎｔ．ａｒｃｈｉｖｅ
ｄｂｅ．ｍｉｇｒａｔｅ　＝　ｅｖｅｎｔ．ｍｉｇｒａｔｅ
【００８９】
　次に、名前空間複製部１３は、ファイル情報の内容を名前空間複製ＤＢ１５に未登録なら情報取得済みとして登録する（Ｓ３０５）。同一ｉｎｏｄｅ番号を持つタプルが登録されていた場合には、登録されている全てのエントリの内容を変更する。ここで、以下の条件をすべて満たすエントリが存在するとき、登録済みとみなす。
【００９０】
ｄｂｅ．ｉｎｏｄｅ＃　＝＝　ｆｓｔａｔ．ｉｎｏｄｅ＃
ｄｂｅ．ｆｎａｍｅ　＝＝　ｆｓｔａｔ．ｆｎａｍｅ
ｄｂｅ．ｍ＿ｉｎｏｄｅ＃　＝＝　ｆｓｔａｔ．ｍ＿ｉｎｏｄｅ＃
【００９１】
　また、未登録時の設定内容を以下に示す。
【００９２】
ｄｂｅ．ｍ＿ｉｎｏｄｅ＃　＝　ｆｓｔａｔ．　ｍ＿ｉｎｏｄｅ＃
ｄｂｅ．ｆｔｙｐｅ　＝　ｆｓｔａｔ．　ｆｔｙｐｅ
ｄｂｅ．ｆｎａｍｅ　＝　ｆｓｔａｔ．ｆｎａｍｅ
ｄｂｅ．ｉｎｏｄｅ＃　＝　ｆｓｔａｔ．ｉｎｏｄｅ＃
ｄｂｅ．ａｒｃｈｉｖｅ　＝　ｆｓｔａｔ．ａｒｃｈｉｖｅ
ｄｂｅ．ｍｉｇｒａｔｅ　＝　ｆｓｔａｔ．ｍｉｇｒａｔｅ
ｄｂｅ．ａｔｉｍｅ　＝　ｆｓｔａｔ．ａｔｉｍｅ
ｄｂｅ．ａｃｔｉｖｅ　＝　オン
【００９３】
　また、同一ｉｎｏｄｅ番号が既に登録済み（すなわちｄｂｅ．ｉｎｏｄｅ＃　＝　ｆｓｔａｔ．ｉｎｏｄｅ＃の場合）の設定内容を以下に示す。
【００９４】
ｄｂｅ．ａｒｃｈｉｖｅ　＝　ｆｓｔａｔ．ａｒｃｈｉｖｅ
ｄｂｅ．ｍｉｇｒａｔｅ　＝　ｆｓｔａｔ．ｍｉｇｒａｔｅ
ｄｂｅ．ａｔｉｍｅ　＝　ｆｓｔａｔ．ａｔｉｍｅ
ｄｂｅ．ａｃｔｉｖｅ　＝　オン
【００９５】
　次に、名前空間複製部１３は、記録されていた全ファイル情報の処理を終了すると、名前空間の変更との競合のため情報取得で見逃した名前空間のセグメント（情報取得済みが表示されていないディレクトリ）が存在するか否かの判断を行う（Ｓ３１１）。存在しない場合（Ｓ３１１，Ｎ）、このフローを終了する。一方、存在する場合（Ｓ３１１，Ｙ）、そのディレクトリをルートとするファイル情報取得処理、およびその間に発生したイベントデータ反映処理を行い（Ｓ３１２）、処理Ｓ３１１へ戻り、次の情報取得済みが表示されていないディレクトリを見つけ、この処理を繰り返す。
【００９６】

次に、名前空間追随処理について説明する。

【００９７】

名前空間追随部１４は、名前空間複製処理が完了した後に発生したイベントデータをイベントデータ記録部２１から受け取り、名前空間複製ＤＢ１５に順次反映していく。イベントデータ反映処理は名前空間複製処理とほぼ同じだが、ファイル情報を用いない分、単純となる。

【００９８】

イベントデータが削除系ファイル操作イベント（ファイル削除、ディレクトリ削除）である場合、名前空間追随部１４は、イベントデータで示されるｉｎｏｄｅ番号、親ｉｎｏｄｅ番号、ファイル名を全て含むエントリを名前空間複製ＤＢ１５上から削除する。

【００９９】

イベントデータが生成系ファイル操作イベント（ファイル生成、ディレクトリ生成）である場合、名前空間追随部１４は、イベントデータで示されるｉｎｏｄｅ番号を含むエントリを名前空間複製ＤＢ１５上に登録し、イベントデータで伝えられた属性（タイプ）、および親ｉｎｏｄｅ番号を設定する。

【０１００】

イベントデータがファイル名の変更（ｒｅｎａｍｅ）でターゲットと同じファイルがあれば、名前空間追随部１４は削除する。また、名前空間追随部１４は、ソースの親属性を変更する。

【０１０１】

イベントデータがファイルアクセスイベントである場合、名前空間追随部１４は、イベントデータで伝えられたアクセス時刻をｉｎｏｄｅ番号で特定し、名前空間複製ＤＢ１５に設定する。

【０１０２】

イベントデータがアーカイブ状態変化である場合、名前空間追随部１４は、アーカイブ情報を更新する。

【０１０３】

次に、マイグレート処理について説明する。

【０１０４】

マイグレート決定部１６は、ファイルシステムが提供するコマンドなどを使い、一次ストレージ１の空きスペース状況を定期的に調べ、空きスペース量がユーザにより指定された量以下になった場合、名前空間複製ＤＢ１５に設定されている情報を使って、マイグレートの対象ファイルを決定し、ファイルシステム制御部１２にマイグレートを要求する。この際、マイグレート決定部１６は、名前空間複製ＤＢ１５から求めたファイルのパス名をファイルシステム制御部１２に渡し、ファイルデータとともに二次ストレージ２に書き出してもらう。マイグレート決定処理は、ユーザポリシに応じて様々な実装を行うことができるが、一例を以下に示す。

【０１０５】

図６は、前提技術１に係るマイグレート決定処理の動作の一例を示すフローチャートである。まず、マイグレート決定部１６は、一次ストレージ１の不足が深刻であるか否かの判断を行う（Ｓ４０１）。

【０１０６】

一次ストレージ１の不足が深刻である場合（Ｓ４０１，Ｙ）、マイグレート決定部１６は、名前空間複製ＤＢ１５を検索し、アーカイブ済みでかつマイグレート済みではないファイルを見つけ（Ｓ４１１）、見つけた全てのファイルに対し以下のリリース処理（一次ストレージ域の解放）を行う。次に、マイグレート決定部１６は、見つけたファイルのうち未処理のファイルがあるか否かの判断を行う（Ｓ４１２）。

【０１０７】

未処理のファイルがなければ（Ｓ４１２，Ｎ）、このフローを終了する。一方、未処理のファイルがあれば（Ｓ４１２，Ｙ）、マイグレート決定部１６は、名前空間複製ＤＢ１

５に設定されているｉｎｏｄｅ番号を引数としてファイルシステム制御部１２に対象ファイルのリリース（一次ストレージ解放）を要求する（Ｓ４１３）。次に、マイグレート決定部１６は、ファイルシステム制御部１２からの応答を得ると、処理Ｓ４１２へ戻り、次の対象ファイルの処理を行う。

【０１０８】

ここで、マイグレート決定部１６は、名前空間複製ＤＢ１５はファイルシステムに遅れて追随するため、実際にはファイルが存在しなくなっている場合や、アーカイブが無効になっている場合があり、この場合にはファイルシステム制御部１２がエラー応答を返す。ファイルがアーカイブ済みであった場合、ファイルシステム制御部１２はそのファイルに割り当てていた一次ストレージ領域を解放して正常応答を返す。

【０１０９】

一方、一次ストレージ１が不足しているがそれほど深刻ではない場合（Ｓ４０１、Ｎ）、マイグレート決定部１６は、深刻な不足が発生したときに、事態をただちに改善できるようにするため、一定時間以上アクセスされていないファイルをアーカイブする。このため、マイグレート決定部１６は、名前空間複製ＤＢ１５を検索し、最終アクセス時刻が所定の時刻（例えば現時刻－１日）以前でかつ、アーカイブ無効（アーカイブ済みでない）なものを見つける（Ｓ４２１）。次に、マイグレート決定部１６は、見つけたファイルのうち未処理のファイルがあるか否かの判断を行う（Ｓ４２２）。

【０１１０】

未処理のファイルがなければ（Ｓ４２２，Ｎ）、このフローを終了する。一方、未処理のファイルがあれば（Ｓ４２２，Ｙ）、マイグレート決定部１６は、名前空間複製ＤＢ１５に設定されている親ｉｎｏｄｅ番号をキーとして、繰り返し名前空間複製ＤＢ１５を検索することで、ファイルのパス名を求める（Ｓ４２３）。次に、マイグレート決定部１６は、ｉｎｏｄｅ番号、ファイルパス名を引数としたアーカイブ要求をファイルシステム制御部１２に出す（Ｓ４２４）。ここで、ファイルシステム制御部１２は、指定されたファイルのデータとファイルパス名、ｉｎｏｄｅ番号を一括して二次ストレージ上に書き出し、処理Ｓ４２２へ戻り、次の対象ファイルの処理を行う。ここで、要求されたファイルが存在しなくなっている場合、ファイルシステム制御部１２はエラーを応答し、要求を無視する。

【０１１１】

次に、その他の各部の動作について説明する。

【０１１２】

まず、ファイルシステム制御部１２について説明する。

【０１１３】

まず、マイグレート決定部１６からのリリース要求があった場合、ファイルシステム制御部１２は、リリース要求を処理し、二次ストレージにファイルデータのコピーが存在する（アーカイブ済み）なら、一次ストレージを返却し、マイグレート済みとする。このとき、イベントデータ記録部２１はアーカイブ状態変化イベントを作成する。

【０１１４】

ｅｖｅｎｔ．ｒｅｃｔｙｐｅ　＝　ａｒｃｈｉｖｅ
ｅｖｅｎｔ．ａｒｃｈｉｖｅ　＝　オン
ｅｖｅｎｔ．ｍｉｇｒａｔｅ　＝　オン

【０１１５】

また、マイグレート決定部１６からのアーカイブ要求があった場合、ファイルシステム制御部１２は、アーカイブ要求を処理し、ファイルデータの二次ストレージ２への書き出しを起動し、マイグレート決定部１６に復帰する。この際、二次ストレージ２に書き出すデータのヘッダ部にファイルのマイグレート決定部１６から通知されたファイルのパス名を付加して書き出す。二次ストレージ２への書き出しが完了すると、イベントデータ記録部２１はアーカイブ状態変化イベントを作成する。

【０１１６】

ｅｖｅｎｔ．ｒｅｃｔｙｐｅ　＝　ａｒｃｈｉｖｅ
ｅｖｅｎｔ．ａｒｃｈｉｖｅ　＝　オン
ｅｖｅｎｔ．ｍｉｇｒａｔｅ　＝　オフ

【０１１７】

また、アプリケーション部１１がマイグレート済みファイルをアクセスしようとした場合、ファイルシステム制御部１２は、アプリケーション部１１がアクセスしようとしたタイミングで、一次ストレージ１上に領域を新たに割り当て、二次ストレージ２上のデータをその領域に読み込む。その後、イベントデータ記録部２１は、リコール完了を表示したアーカイブ状態変化イベントを作成する。

【０１１８】

ｅｖｅｎｔ．ｒｅｃｔｙｐｅ　＝　ａｒｃｈｉｖｅ
ｅｖｅｎｔ．ａｒｃｈｉｖｅ　＝　オン
ｅｖｅｎｔ．ｍｉｇｒａｔｅ　＝　オフ

【０１１９】

また、アプリケーション部１１がファイル操作（ファイル生成・削除、ディレクトリ生成・削除、ファイルｒｅａｄ／ｗｒｉｔｅ）を要求した場合、ファイルシステム制御部１２は要求を処理し、正常に完了した時点で、イベントデータ記録部２１は対応するイベントデータを作成する。

【０１２０】

名前空間複製部１３からｇｅｔｉｎｆｏでファイル情報を要求された場合、ファイルシステム制御部１２は、指定されたファイルが親ディレクトリに存在することを確認した上で、指定されたファイルのファイル情報を返す。存在しなければ、エラーを応答。エラーが返された場合の名前空間複製部１３はそのファイルがなかったものとして処理を続ける。

【０１２１】

次に、イベントデータ記録部２１について説明する。

【０１２２】

イベントデータ記録部２１は、ファイルシステム制御部１２内に存在し、ファイルシステム制御部１２の説明で述べたタイミングでイベントデータを作成し、メモリ上に蓄積する部分である。また、イベントデータ記録部２１は、メモリ上に蓄積されたイベントデータが一定量以上となった、あるいは最後に通知してから、一定時間経過したときに、メモリ上に蓄積されていたイベントデータを一括して、名前空間追随部１４あるいは名前空間複製部１３に通知する。また、システム停止時にも、イベントデータ記録部２１が蓄積していたイベントデータを名前空間追随部１４に通知し、名前空間追随部１４がメモリ上に蓄積されているイベントデータを名前空間複製ＤＢ１５に全て反映する、システム停止処理を行う。

【０１２３】

また、イベントデータ記録部２１では、通知するデータ量を削減するため、以下の最適化を施す。まず、イベントデータ記録部２１がファイルアクセスイベントを作成する場合、メモリ上に蓄積されている未通知のイベントデータの中に同じファイルに対するファイルアクセスイベントが含まれているなら、後続のファイルアクセスイベントは捨てる。すなわち、メモリ上に蓄積しない。また、イベントデータ記録部２１がファイル削除イベントの作成を依頼されたときに、対応するファイル生成イベントが未通知のイベントデータとして含まれるなら、ファイル生成イベントをメモリ上で無効化し、イベントデータ通知の対象から取り除く。

【０１２４】

次に、サーバ３におけるシステム立ち上げの処理について説明する。

【０１２５】

システムを正常終了した場合、上述したように名前空間追随部１４がメモリ上に滞留していたイベントデータを一括して名前空間複製ＤＢ１５に反映する正常終了処理を行うた

め、次回立ち上げ時に名前空間複製部１３を動作させる必要はない。一方、障害発生の場合、その後の再立ち上げ時には、名前空間複製部１３を動作させ、名前空間複製ＤＢ１５を再初期化するシステム異常終了後起動処理を行う。なお、この場合でも、障害発生直前の名前空間情報は残っているので、名前空間複製の再初期化が完了するまでの間にマイグレーション対象を決定する必要が発生した場合には、マイグレート決定部は古い複製を使って処理を行う。

【０１２６】

なお、前提技術１においては、名前空間複製ＤＢ１５に基づくポリシ制御の例としてマイグレート決定部１６について説明したが、ＨＳＭ制御における他のポリシ制御を名前空間複製ＤＢ１５に基づいて行う構成としても良い。

【０１２７】

上述した前提技術１のようなＨＳＭ装置において、二次ストレージ２への退避先のボリュームを指定することはできるが、一次ストレージ１（ファイルシステム）への復元先のボリュームを指定することができない。また、一次ストレージ１におけるファイルデータ領域の割り当ては、すべてファイルシステム制御部１２により決定される。ここで、ファイルシステム制御部１２は、利用しているボリューム構成およびその配下のディスク構成の情報を保持していないため、常に一次ストレージ１全体の中から均一に割り当てる。また、一次ストレージ１へのファイル作成（新規ファイルの書き込み）時と二次ストレージ２から一次ストレージ１へのファイル復元（退避済みファイルの読み込み）時とでボリュームに対する性能や信頼性の要件が異なる場合であっても、一次ストレージ１のボリューム構成（ディスク構成）は、どちらか一方の要件を満たす必要がある。従って、どちらか一方の要件をあきらめる、あるいは、両方の要件を満たすために一次ストレージ１用に用意するディスク容量（ディスク数）を多くする、などの制約が発生する。

【０１２８】

以下、本発明の実施の形態について図面を参照しつつ説明する。

【０１２９】

実施の形態１．

本実施の形態では、前提技術１と同様にして名前空間（メタデータ）の複製を作成、更新するＨＳＭシステムにおいて、一次ストレージにおいて最適なファイルデータ領域の割り当てを行うＨＳＭシステムについて説明する。

【０１３０】

まず、本実施の形態に係るＨＳＭシステムの構成について説明する。

【０１３１】

図７は、本実施の形態に係るＨＳＭシステムの構成の一例を示すブロック図である。この図において、図１と同一符号は図１に示された対象と同一又は相当物を示しており、ここでの説明を省略する。このＨＳＭシステムは、一次ストレージ１、二次ストレージ２、運用サーバ１０２、ＨＳＭ制御サーバ１０３（ＨＳＭ制御装置）、操作部１０５を備える。

【０１３２】

運用サーバ１０２は、アプリケーション１１１、ファイルシステム制御部１１２、ボリューム管理部１１３を備える。アプリケーション１１１は、ユーザが利用するプログラムである。ファイルシステム制御部１１２は、ファイルデータ領域割り当て制御部１２１、ＦＳ（File System）ボリューム管理部１２２を備える。ボリューム管理部１１３は、仮想ボリューム管理部１２３、ディスクドライバ１２４を備える。ＨＳＭ制御サーバ１０３は、ＨＳＭ管理部１３１、ボリュームマップデータベース１３２、ファイル配置ポリシデータベース１３３を備える。操作部１０５は、ユーザによる操作を受け付ける端末であり、ＨＳＭ制御サーバ１０３の内部に備えられても良い。

【０１３３】

なお、前提技術１におけるファイルシステム制御部１２は、本実施の形態におけるファイルシステム制御部１１２に対応する。また、前提技術１における名前空間複製部１３、

名前空間追随部１４、マイグレート決定部１６は、本実施の形態におけるＨＳＭ管理部１３１に対応する。

【０１３４】

また、ボリューム情報管理ステップは、本実施の形態におけるボリューム情報取得処理に対応する。また、ボリューム選択ステップは、本実施の形態における対象ＦＳボリューム選択処理に対応する。

【０１３５】

一次ストレージ１は、単一または複数のディスク装置を有する。ファイルデータ領域割り当て制御部１２１は、ファイルシステム（一次ストレージ１）におけるファイルデータの格納先であるファイルデータ領域の割り当てと解放を行う。ＦＳボリューム管理部１２２は、ファイルシステム上のボリュームであるＦＳボリュームを構成するボリューム（仮想ボリュームまたはディスク装置）を管理する。仮想ボリューム管理部１２３は、仮想ボリュームを構成するディスク装置を管理する。仮想ボリュームは、複数のディスク装置の連結、ストライピング構成、ミラー構成などを実現する。ＦＳボリューム管理部１２２、仮想ボリューム管理部１２３は、ディスクドライバ１２４を介して一次ストレージ１におけるディスク装置にアクセスする。

【０１３６】

ＨＳＭ管理部１３１は、ＨＳＭに関する一次ストレージ１と二次ストレージ２のファイル操作を行う。ボリュームマップデータベース１３２は、ＦＳボリュームを構成する要素の情報（ボリューム情報）を保存し、管理する。ファイル配置ポリシデータベース１３３は、ファイルデータ領域を割り当てるためのポリシを保存し、管理する。

【０１３７】

次に、ボリュームマップデータベース１３２について説明する。

【０１３８】

図８は、本実施の形態に係るボリュームマップデータベース１３２の内容の一例を示す表である。ボリュームマップデータベース１３２は、ＦＳボリューム構成情報テーブル、仮想ボリューム構成情報テーブル、ディスク情報テーブルを備える。

【０１３９】

ＦＳボリューム構成情報テーブルは、ＦＳボリュームを構成するボリュームの情報であるＦＳボリューム構成情報を保存する。ここでは、ＦＳボリューム毎に、ＦＳボリューム通番、ＦＳ内ブロック範囲、割り当てボリューム、ボリューム内ブロック範囲が指定される。ＦＳ内ブロック範囲は、該当するＦＳボリュームに対応するファイルシステム内のブロック範囲を表す。割り当てボリュームは、該当するＦＳボリュームとして割り当てられた仮想ボリュームまたはディスク装置を表す。ボリューム内割り当て範囲は、該当するＦＳボリュームに対応するボリューム内のブロック範囲を表す。

【０１４０】

仮想ボリューム構成情報テーブルは、仮想ボリュームを構成するディスク装置の情報である仮想ボリューム構成情報を保存する。ここでは、仮想ボリューム毎に、仮想ボリュームＩＤ、構成タイプ、割り当てディスク群が指定される。構成タイプは例えば、ミラー（二重化）、ストライプ、コンカチ（結合）などを示す。割り当てディスク群は、仮想ボリュームに割り当てられるディスクＩＤを示す。

【０１４１】

ディスク情報テーブルは、ディスク装置の情報であるディスク情報を保存する。ここでは、ディスク装置毎に、ディスクＩＤ、ディスクタイプ、ディスク装置情報が指定される。ディスクタイプは例えば、単体のディスク（シングル構成）、ＲＡＩＤ－０（ストライプ）などを示す。ディスク装置情報は例えば、冗長性なし、高速ディスクなどの特性を示す。

【０１４２】

ボリュームマップデータベース１３２は、ＨＳＭシステム起動時に行われるボリューム情報取得処理により設定される。図９は、本実施の形態に係るボリューム情報取得処理の

動作の一例を示すフローチャートである。まず、ＨＳＭ制御サーバ１０３は、運用サーバ１０２に接続する（Ｓ５５１）。次に、ＨＳＭ管理部１３１は、ファイルシステム制御部１１２のＦＳボリューム管理部１２２から、ＦＳボリューム構成情報を取得する（Ｓ５５２）。次に、ＨＳＭ管理部１３１は、ボリューム管理部１１３の仮想ボリューム管理部１２３から、仮想ボリューム構成情報を取得する（Ｓ５５３）。次に、ＨＳＭ管理部１３１は、ボリューム管理部１１３のディスクドライバ１２４、一次ストレージ１におけるディスク装置から、ディスク情報を取得する（Ｓ５５４）。次に、ＨＳＭ管理部１３１は、取得したＦＳボリューム構成情報、ディスク構成情報、ディスク情報を統合し、ボリュームマップ情報としてボリュームマップデータベース１３２に保存し（Ｓ５５５）、このフローは終了する。

【０１４３】

次に、ファイル配置ポリシデータベース１３３について説明する。

【０１４４】

図１０は、本実施の形態に係るファイル配置ポリシデータベース１３３の内容の一例を示す表である。ファイル配置ポリシデータベース１３３は、ファイル条件指定ポリシ設定テーブルと個別ファイルポリシ設定テーブルを備える。

【０１４５】

ファイル条件指定ポリシ設定テーブルは、対象ファイルを選別するための条件である対象ファイル条件が指定される。ここでは、対象ファイル条件毎に、ファイル作成先ＦＳボリューム条件とファイル復元先ＦＳボリューム条件が指定される。１つの対象ファイル条件は、ファイル所有者、ファイルサイズ、ファイル名、ファイル名パターン指定などで表され、１つの条件でも良いし、複数の条件でも良いし、複数の種類の条件が組み合わせられたものでも良い。また、条件を指定しないものは、対象ファイル条件を「その他すべて」とする。

【０１４６】

ファイル作成先ＦＳボリューム条件は、新規のファイルが一次ストレージ１に作成されるケースにおいて、割り当てるファイルデータ領域として適するボリュームの条件を表す。同様に、ファイル復元先ＦＳボリューム条件は、二次ストレージ２に退避済みのファイルが一次ストレージ１に復元されるケースにおいて、割り当てるファイルデータ領域として適するボリュームの条件を表す。これらは、ボリュームの信頼性が高いか低いか、ボリュームのアクセス性能が高いか低いか、指定するＦＳボリュームの識別子などで表される。ここで、対象ファイル条件がファイルサイズで表された場合、ファイル作成先ＦＳボリューム条件はファイルサイズが未定であるため指定不可能である。

【０１４７】

個別ファイルポリシ設定テーブルは、ファイル作成後に、ファイル毎にポリシを設定する場合のテーブルである。ファイル毎に、ファイルｉｎｏｄｅ番号、ファイルｉｎｏｄｅ番号毎にファイル復元先ＦＳボリューム条件が指定される。ここで、ファイル作成後に設定されるものであるため、ファイル作成先ＦＳボリューム条件は、指定不可能である。

【０１４８】

ファイル配置ポリシデータベース１３３は、ポリシ設定処理により設定される。図１１は、本実施の形態に係るポリシ設定処理の動作の一例を示すフローチャートである。まず、操作部１０５は、ユーザの指示に従って、ＨＳＭ制御サーバ１０３に接続する（Ｓ５１１）。次に、操作部１０５は、ユーザの指示に従って、ファイルやディレクトリの種別などの条件を指定し、ファイル毎に利用すべきボリューム（ディスク）の選択ポリシを入力する（Ｓ５１２）。次にＨＳＭ管理部１３１は、入力されたポリシをファイル配置ポリシデータベース１３３に保存し（Ｓ５１３）、このフローは終了する。

【０１４９】

次に、ファイルデータ領域割り当て処理について説明する。

【０１５０】

ファイルデータ領域割り当て処理は、ファイル作成またはファイル復元のために、一次

ストレージ１におけるファイルデータ領域の割り当てを行う処理である。図１２は、本実施の形態に係るファイルデータ領域割り当て処理の動作の一例を示すフローチャートである。まず、ファイルシステム制御部１１２は、一次ストレージ１におけるファイルデータ領域の割り当てが必要な処理が発生すると（Ｓ６１１）、その処理がファイル復元であるか否かの判断を行う（Ｓ６１２）。

【０１５１】

発生した処理がファイル復元である場合（Ｓ６１２，Ｙ）、ファイルデータ領域割り当て制御部１２１は、ファイル復元を行うこととその対象ファイルの情報をＨＳＭ管理部１３１に通知する（Ｓ６１３）。

【０１５２】

ＨＳＭ管理部１３１は、ファイル配置ポリシデータベース１３３とボリュームマップデータベース１３２に基づいて、対象ファイルのファイルデータ領域割り当てに使用すべきＦＳボリュームである対象ＦＳボリュームを選択する対象ＦＳボリューム選択処理を行い、対象ＦＳボリュームをファイルシステム内ブロック範囲として、ファイルデータ領域割り当て制御部１２１に通知する（Ｓ６１４）。次に、ファイルデータ領域割り当て制御部１２１は、通知されたブロック範囲の中で対象ファイルのファイルデータ領域を割り当て、ＨＳＭ管理部１３１に通知する（Ｓ６１５）。次に、ＨＳＭ管理部１３１は、二次ストレージ２に退避済みの対象ファイルのファイルデータを、通知されたファイルデータ領域へ復元し（Ｓ６１６）、このフローは終了する。

【０１５３】

処理Ｓ６１２において、発生した処理がファイル作成である場合（Ｓ６１２，Ｎ）、ファイルデータ領域割り当て制御部１２１は、ファイル作成を行うこととその対象ファイルの情報をＨＳＭ管理部１３１に通知する（Ｓ６２３）。

【０１５４】

次に、ＨＳＭ管理部１３１は、ファイル配置ポリシデータベース１３３とボリュームマップデータベース１３２に基づいて、対象ファイルのファイルデータ領域割り当てに使用すべきＦＳボリュームである対象ＦＳボリュームを選択する対象ＦＳボリューム選択処理を行い、対象ＦＳボリュームをファイルシステム内ブロック範囲として、ファイルデータ領域割り当て制御部１２１に通知する（Ｓ６２４）。次に、ファイルデータ領域割り当て制御部１２１は、通知されたブロック範囲の中で対象ファイルのファイルデータ領域を割り当て（Ｓ６２５）、このフローは終了する。

【０１５５】

次に、ファイル復元時の対象ＦＳボリューム選択処理（Ｓ６１４）の詳細について説明する。

【０１５６】

図１３は、本実施の形態に係るファイル復元時の対象ＦＳボリューム選択処理の第１の動作の一例を示すフローチャートである。図１４は、本実施の形態に係るファイル復元時の対象ＦＳボリューム選択処理の第２の動作の一例を示すフローチャートである。まず、ＨＳＭ管理部１３１は、ファイルデータ領域割り当て制御部１２１からファイル復元の対象ファイルの情報として、ファイルｉｎｏｄｅ番号、ファイル所有者、ファイルサイズ、ファイル名を受け取る（Ｓ７１１）。次に、ＨＳＭ管理部１３１は、ファイル配置ポリシデータベース１３３の個別ファイルポリシ設定テーブルから対象ファイルのｉｎｏｄｅ番号を検索し（Ｓ７１２）、同一ｉｎｏｄｅ番号があるか否かの判断を行う（Ｓ７１３）。

【０１５７】

同一ｉｎｏｄｅ番号がある場合（Ｓ７１３，Ｙ）、処理Ｓ７２１へ移行する。一方、同一ｉｎｏｄｅ番号がない場合（Ｓ７１３，Ｎ）、ＨＳＭ管理部１３１は、ファイル配置ポリシデータベース１３３のファイル条件指定ポリシテーブルから対象ファイルの所有者、サイズ、名前に該当する対象ファイル条件を順に検索し（Ｓ７１４）、一致する対象ファイル条件があるか否かの判断を行う（Ｓ７１５）。一致する対象ファイル条件がある場合（Ｓ７１５，Ｙ）、処理Ｓ７２１へ移行する。一方、一致する対象ファイル条件がない場

合（Ｓ７１５，Ｎ）、ＨＳＭ管理部１３１は、ファイル条件指定ポリシ設定テーブルにおける「その他すべて」のファイル復元先ＦＳボリューム条件を取得する（Ｓ７１６）。また、処理Ｓ７２１において、ＨＳＭ管理部１３１は、個別ファイルポリシ設定テーブルまたはファイル条件指定ポリシ設定テーブルにおいて該当するポリシのファイル復元先ＦＳボリューム条件を取得する（Ｓ７２１）。

【０１５８】

次に、ＨＳＭ管理部１３１は、ボリュームマップデータベース１３２のＦＳボリューム構成情報テーブルからＦＳボリューム構成情報を取得し（Ｓ７５１）、仮想ボリューム構成情報テーブルから仮想ボリューム構成情報を取得し（Ｓ７５２）、ディスク情報テーブルからディスク情報を取得する（Ｓ７５３）。次に、ＨＳＭ管理部１３１は、取得した情報からＦＳボリューム毎の特性であるＦＳボリューム特性を決定する（Ｓ７５４）。次に、ＨＳＭ管理部１３１は、ＦＳボリューム特性がファイル復元先ＦＳボリューム条件に適合するＦＳボリュームを対象ＦＳボリュームとして選択し（Ｓ７５５）、対象ＦＳボリュームに対応するファイルシステム内ブロック範囲をファイルデータ領域割り当て制御部１２１に通知し（Ｓ７５６）、このフローは終了する。

【０１５９】

次に、ファイル作成時の対象ＦＳボリューム選択処理（Ｓ６２４）の詳細について説明する。

【０１６０】

図１５は、本実施の形態に係るファイル作成時の対象ＦＳボリューム選択処理の第１の動作の一例を示すフローチャートである。図１６は、本実施の形態に係るファイル作成時の対象ＦＳボリューム選択処理の第２の動作の一例を示すフローチャートである。図１５において、図１３と同一符号は図１３に示された対象と同一又は相当物を示しており、ここでの説明を省略する。また、図１６において、図１４と同一符号は図１４に示された対象と同一又は相当物を示しており、ここでの説明を省略する。まず、ＨＳＭ管理部１３１は、ファイルデータ領域割り当て制御部１２１からファイル作成の対象ファイルの情報として、ファイル所有者、ファイル名を受け取る（Ｓ８１１）。次に、ＨＳＭ管理部１３１は、ファイル配置ポリシデータベース１３３のファイル条件指定ポリシテーブルから対象ファイルの所有者、名前に該当する対象ファイル条件を順に検索し（Ｓ８１４）、一致する対象ファイル条件があるか否かの判断を行う（Ｓ８１５）。

【０１６１】

一致する対象ファイル条件がある場合（Ｓ８１５，Ｙ）、処理Ｓ８２１へ移行する。一方、一致する対象ファイル条件がない場合（Ｓ８１５，Ｎ）、ＨＳＭ管理部１３１は、ファイル条件指定ポリシ設定テーブルにおける「その他すべて」のファイル作成先ＦＳボリューム条件を取得する（Ｓ８１６）。また、処理Ｓ８２１において、ＨＳＭ管理部１３１は、ファイル条件指定ポリシ設定テーブルにおける該当ポリシのファイル作成先ＦＳボリューム条件を取得する（Ｓ８２１）。以後、上述した処理Ｓ７５１～Ｓ７５６が実行され、このフローは終了する。

【０１６２】

ここで、ファイル復元時またはファイル作成時の対象ＦＳボリューム選択処理について、図８の例を用いて説明する。処理Ｓ７５１において、ＦＳボリューム構成情報として、各ＦＳボリュームに割り当てられた仮想ボリュームまたはディスク装置（割り当てボリューム）がｖｏｌ＿００１，ｄｉｓｋ＿００３であることが取得される。処理Ｓ７５２において、仮想ボリューム構成情報として、ｖｏｌ＿００１はミラーであってｄｉｓｋ＿００１とｄｉｓｋ＿００２を使用していることが取得される。処理Ｓ７５３において、ｄｉｓｋ＿００１は「冗長性なし」であること、ｄｉｓｋ＿００２は「冗長性なし」であること、ｄｉｓｋ＿００３は「冗長性なし」、「高速」であることが取得される。

【０１６３】

処理Ｓ７５４において、ＦＳボリュームに対応するｖｏｌ＿００１は「冗長性あり」、「低速」であること、ＦＳボリュームに対応するｄｉｓｋ＿００３は「冗長性なし」、「

高速」であることが取得される。処理Ｓ７５５において、該当ポリシのファイル復元先ＦＳボリューム条件が「信頼性高」であればｖｏｌ＿００１が選択され、該当ポリシのファイル復元先ＦＳボリューム条件が「アクセス性能高」であればｄｉｓｋ＿００３が選択される。処理Ｓ７５６において、ｖｏｌ＿００１が選択された場合はファイルシステム内ブロック範囲０～９９９９がファイルデータ領域割り当て制御部１２１に通知され、ｄｉｓｋ＿００３が選択された場合はファイルシステム内ブロック範囲１００００～４９９９９がファイルデータ領域割り当て制御部１２１に通知される。ファイルデータ領域割り当て制御部１２１は、このブロック範囲の中で対象ファイルのファイルデータ領域の割り当てを行う。

【０１６４】

本実施の形態によれば、予めユーザにより設定されたポリシに従って、一次ストレージ１へのファイル作成時と二次ストレージ２から一次ストレージ１へのファイル復元時とで異なる特性のボリュームを用いるなど、一次ストレージ１における最適なファイルデータの配置を行うことができる。同様に、予めユーザにより設定されたポリシに従って、ファイル毎やファイル種別毎に配置を分けることができる。

【０１６５】

例えば、一次ストレージ１が、ミラー構成のボリュームとシングル構成のボリュームで構成されている場合、ファイル作成時は信頼性の高いミラー構成のボリュームを用い、ファイル復元時は二次ストレージ２にデータが退避済みであることから信頼性は低くてもよく、シングル構成のボリュームを用いるように設定する。これにより、一次ストレージ１用のディスク容量の節約や有効活用を実現することができる。

【０１６６】

また、本実施の形態に係るＨＳＭ制御装置は、サーバに容易に適用することができ、サーバの性能をより高めることができる。

【０１６７】

更に、ＨＳＭ制御装置を構成するコンピュータにおいて上述した各ステップを実行させるプログラムを、ＨＳＭ制御プログラムとして提供することができる。上述したプログラムは、コンピュータにより読取り可能な記録媒体に記憶させることによって、ＨＳＭ制御装置を構成するコンピュータに実行させることが可能となる。ここで、上記コンピュータにより読取り可能な記録媒体としては、ＲＯＭやＲＡＭ等のコンピュータに内部実装される内部記憶装置、ＣＤ－ＲＯＭやフレキシブルディスク、ＤＶＤディスク、光磁気ディスク、ＩＣカード等の可搬型記憶媒体や、コンピュータプログラムを保持するデータベース、或いは、他のコンピュータ並びにそのデータベースや、更に回線上の伝送媒体をも含むものである。

【０１６８】

（付記１）　一次ストレージと二次ストレージを有するＨＳＭの制御をコンピュータに実行させるＨＳＭ制御プログラムであって、

前記一次ストレージを構成する少なくとも一つのボリュームに関する情報をボリューム情報として管理するボリューム情報管理ステップと、

前記一次ストレージにファイルデータ領域の割り当てが行われる場合、前記ボリューム情報管理ステップにより管理されたボリューム情報と予め設定されたポリシとに基づいて、前記割り当ての対象となるボリューム、前記一次ストレージを構成するボリュームの中から選択するボリューム選択ステップと

をコンピュータに実行させるＨＳＭ制御プログラム。

（付記２）　付記１に記載のＨＳＭ制御プログラムにおいて、

前記ポリシは、前記ボリューム選択ステップにより選択されるボリュームの条件である対象ボリューム条件を含むことを特徴とするＨＳＭ制御プログラム。

（付記３）　付記２に記載のＨＳＭ制御プログラムにおいて、

前記ポリシは更に、前記対象ボリューム条件を適用するファイルの条件である対象ファイル条件を含み、

前記ボリューム選択ステップは、前記割り当ての対象のファイルが満たす前記対象ファイル条件を検索し、該対象ファイル条件に対応する前記対象ボリューム条件を選択することを特徴とするＨＳＭ制御プログラム。

（付記４）　付記３に記載のＨＳＭ制御プログラムにおいて、

前記対象ファイル条件は、ファイル識別子とファイル名パターン指定の少なくともいずれかを含むことを特徴とするＨＳＭ制御プログラム。

（付記５）　付記２乃至付記４のいずれかに記載のＨＳＭ制御プログラムにおいて、

前記対象ボリューム条件は、新規のファイルが前記一次ストレージに作成されるケースであるファイル作成時に選択されるボリュームの条件であるファイル作成先ボリューム条件と、前記二次ストレージに退避済みのファイルが前記一次ストレージに復元されるケースであるファイル復元時に選択されるボリュームの条件であるファイル復元先ボリューム条件とを含むことを特徴とするＨＳＭ制御プログラム。

（付記６）　付記５に記載のＨＳＭ制御プログラムにおいて、

前記ファイル作成先ボリューム条件は、前記一次ストレージを構成するボリュームのうち、信頼性の高いボリュームなど、ファイル作成時要件を満たすボリュームであることを特徴とするＨＳＭ制御プログラム。

（付記７）　付記５または付記６に記載のＨＳＭ制御プログラムにおいて、

前記ファイル復元先ボリューム条件は、前記一次ストレージを構成するボリュームのうち、アクセス性能の高いボリュームなど、ファイル復元時要件を満たすボリュームであることを特徴とするＨＳＭ制御プログラム。

（付記８）　付記１乃至付記７のいずれかに記載のＨＳＭ制御プログラムにおいて、

前記ボリューム選択ステップは、前記割り当ての対象となるボリュームに対応するブロック範囲を出力することを特徴とするＨＳＭ制御プログラム。

（付記９）　付記１乃至付記８のいずれかに記載のＨＳＭ制御プログラムにおいて、

前記ボリュームは、仮想ボリュームまたはディスク装置で構成され、

前記ボリューム情報は、前記仮想ボリュームの情報と前記ボリュームに含まれるディスク装置の情報とを含むことを特徴とするＨＳＭ制御プログラム。

（付記１０）　一次ストレージと二次ストレージを有するＨＳＭの制御を行うＨＳＭ制御装置であって、

前記一次ストレージを構成する少なくとも一つのボリュームに関する情報をボリューム情報として管理するボリューム情報管理部と、

前記一次ストレージにファイルデータ領域の割り当てが行われる場合、前記ボリューム情報管理部により管理されたボリューム情報と予め設定されたポリシとに基づいて、前記割り当ての対象となるボリューム、前記一次ストレージを構成するボリュームの中から選択するボリューム選択部と

を備えるＨＳＭ制御装置。

（付記１１）　付記１０に記載のＨＳＭ制御装置において、

前記ポリシは、前記ボリューム選択部により選択されるボリュームの条件である対象ボリューム条件を含むことを特徴とするＨＳＭ制御装置。

（付記１２）　付記１１に記載のＨＳＭ制御装置において、

前記ポリシは更に、前記対象ボリューム条件を適用するファイルの条件である対象ファイル条件を含み、

前記ボリューム選択部は、前記割り当ての対象のファイルが満たす前記対象ファイル条件を検索し、該対象ファイル条件に対応する前記対象ボリューム条件を選択することを特徴とするＨＳＭ制御装置。

（付記１３）　付記１２に記載のＨＳＭ制御プログラムにおいて、

前記対象ファイル条件は、ファイル識別子とファイル名パターン指定の少なくともいずれかを含むことを特徴とするＨＳＭ制御プログラム。

（付記１４）　付記１１乃至付記１３のいずれかに記載のＨＳＭ制御装置において、

前記対象ボリューム条件は、新規のファイルが前記一次ストレージに作成されるケース

であるファイル作成時に選択されるボリュームの条件であるファイル作成先ボリューム条件と、前記二次ストレージに退避済みのファイルが前記一次ストレージに復元されるケースであるファイル復元時に選択されるボリュームの条件であるファイル復元先ボリューム条件とを含むことを特徴とするＨＳＭ制御装置。

（付記１５）　付記１４に記載のＨＳＭ制御装置において、

前記ファイル作成先ボリューム条件は、前記一次ストレージを構成するボリュームのうち、信頼性の高いボリュームなど、ファイル作成時要件を満たすボリュームであることを特徴とするＨＳＭ制御装置。

（付記１６）　付記１４または付記１５に記載のＨＳＭ制御装置において、

前記ファイル復元先ボリューム条件は、前記一次ストレージを構成するボリュームのうち、アクセス性能の高いボリュームなど、ファイル復元時要件を満たすボリュームであることを特徴とするＨＳＭ制御装置。

（付記１７）　付記１０乃至付記１６のいずれかに記載のＨＳＭ制御装置において、

前記ボリューム選択部は、前記割り当ての対象となるボリュームに対応するブロック範囲を出力することを特徴とするＨＳＭ制御装置。

（付記１８）　付記１０乃至付記１７のいずれかに記載のＨＳＭ制御装置において、

前記ボリュームは、仮想ボリュームまたはディスク装置で構成され、

前記ボリューム情報は、前記仮想ボリュームの情報と前記ボリュームに含まれるディスク装置の情報とを含むことを特徴とするＨＳＭ制御装置。

（付記１９）　一次ストレージと二次ストレージを有するＨＳＭの制御を行うＨＳＭ制御方法であって、

前記一次ストレージを構成する少なくとも一つのボリュームに関する情報をボリューム情報として管理するボリューム情報管理ステップと、

前記一次ストレージにファイルデータ領域の割り当てが行われる場合、前記ボリューム情報管理ステップにより管理されたボリューム情報と予め設定されたポリシとに基づいて、前記割り当ての対象となるボリューム、前記一次ストレージを構成するボリュームの中から選択するボリューム選択ステップと

を実行するＨＳＭ制御方法。

（付記２０）　付記１９に記載のＨＳＭ制御方法において、

前記ポリシは、前記ボリューム選択部により選択されるボリュームの条件である対象ボリューム条件を含むことを特徴とするＨＳＭ制御方法。

【図面の簡単な説明】

【０１６９】

【図１】前提技術１に係るＨＳＭ装置の構成の一例を示すブロック図である。

【図２】前提技術１に係るファイル情報取得処理の動作の一例を示すフローチャートである。

【図３】前提技術１に係る名前空間におけるディレクトリの階層構造の一例を示す図である。

【図４】前提技術１に係るファイル情報取得処理の動作の一例を示すフローチャートである。

【図５】前提技術１に係るイベントデータ反映処理の動作の一例を示すフローチャートである。

【図６】前提技術１に係るマイグレート決定処理の動作の一例を示すフローチャートである。

【図７】本実施の形態に係るＨＳＭシステムの構成の一例を示すブロック図である。

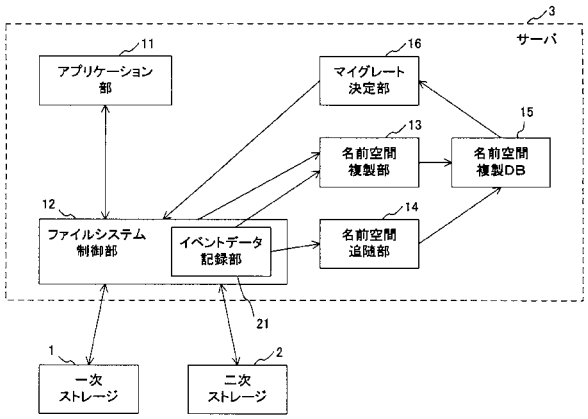
【図８】本実施の形態に係るボリュームマップデータベース１３２の内容の一例を示す表である。

【図９】本実施の形態に係るボリューム情報取得処理の動作の一例を示すフローチャートである。

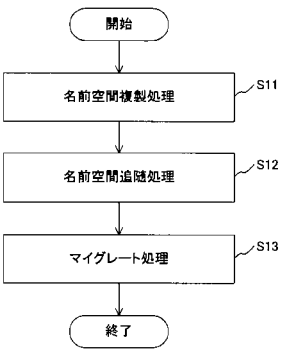
【図１０】本実施の形態に係るファイル配置ポリシデータベース１１３の内容の一例を示

す表である。
【図１１】本実施の形態に係るポリシ設定処理の動作の一例を示すフローチャートである。
【図１２】本実施の形態に係るファイルデータ領域割り当て処理の動作の一例を示すフローチャートである。
【図１３】本実施の形態に係るファイル復元時の対象ＦＳボリューム選択処理の第１の動作の一例を示すフローチャートである。
【図１４】本実施の形態に係るファイル復元時の対象ＦＳボリューム選択処理の第２の動作の一例を示すフローチャートである。
【図１５】本実施の形態に係るファイル作成時の対象ＦＳボリューム選択処理の第１の動作の一例を示すフローチャートである。
【図１６】本実施の形態に係るファイル作成時の対象ＦＳボリューム選択処理の第２の動作の一例を示すフローチャートである。
【符号の説明】
【０１７０】
１　一次ストレージ、２　二次ストレージ、３　サーバ、１１　アプリケーション部、１２　ファイルシステム制御部、１３　名前空間複製部、１４　名前空間追随部、１５　名前空間複製ＤＢ、１６　マイグレート決定部、２１　イベントデータ記録部、１０２　運用サーバ、１０３　ＨＳＭ制御サーバ、１０５　操作部、１１１　アプリケーション、１１２　ファイルシステム制御部、１１３　ボリューム管理部、１１２　ファイルシステム制御部は、１２１　ファイルデータ領域割り当て制御部、１２２　ＦＳボリューム管理部、１２３　仮想ボリューム管理部、１２４　ディスクドライバ、１３１　ＨＳＭ管理部、１３２　ボリュームマップデータベース、１３３　ファイル配置ポリシデータベース。

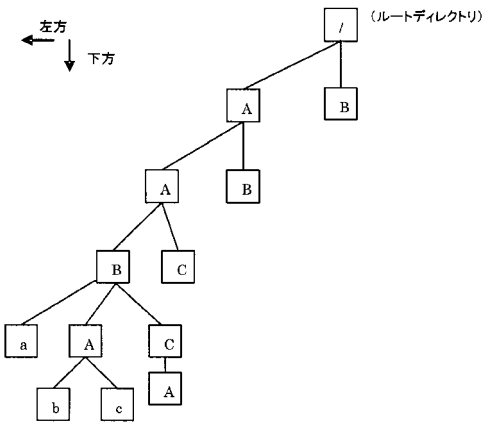
【図１】



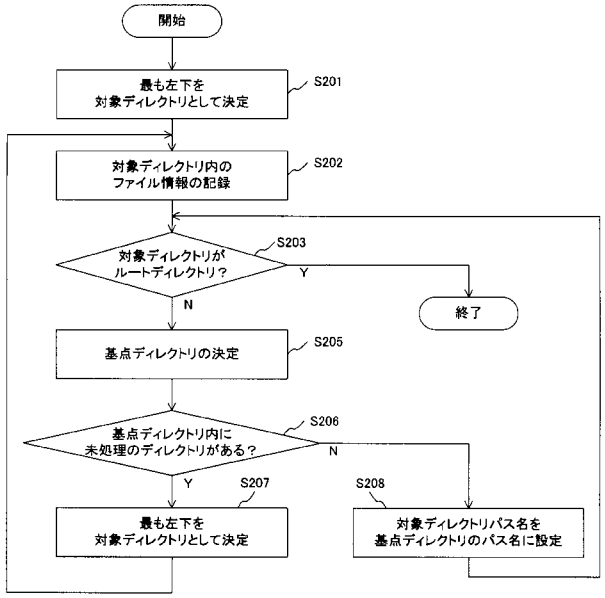
【図２】



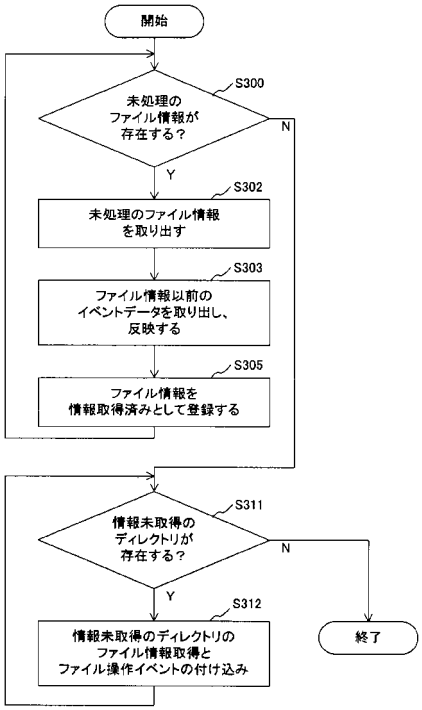
【図３】



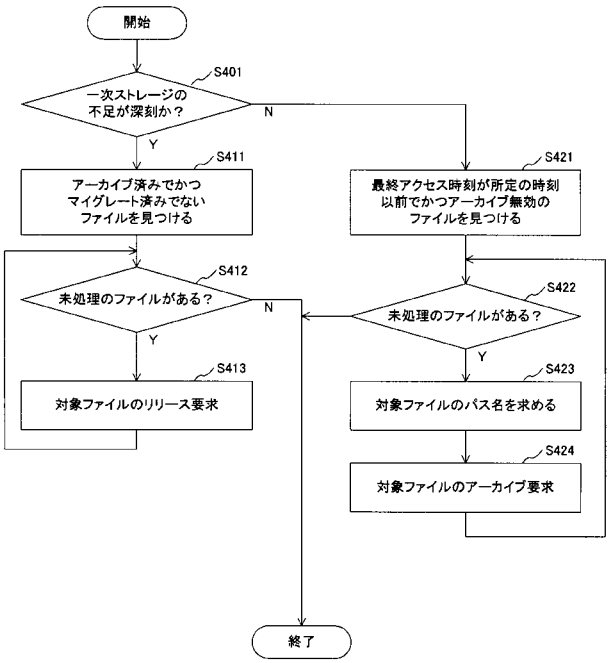
【図４】


開始
最も左下を
対象ディレクトリとして決定
S201
対象ディレクトリ内の
ファイル情報の記録
S202
対象ディレクトリが
ルートディレクトリ？
S203
Y
N
終了
基点ディレクトリの決定
S205
基点ディレクトリ内に
未処理のディレクトリがある？
S206
N
Y
S207
最も左下を
対象ディレクトリとして決定
S208
対象ディレクトリパス名を
基点ディレクトリのパス名に設定

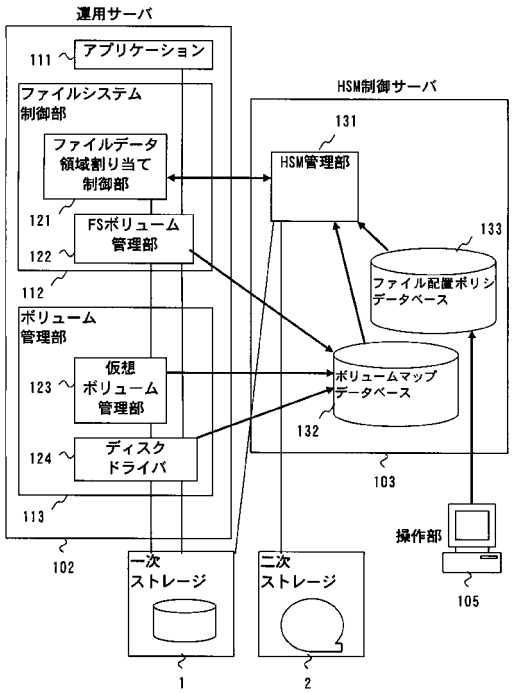
【図５】


開始
未処理の
ファイル情報が
存在する？
S300
N
Y
未処理のファイル情報
を取り出す
S302
ファイル情報以前の
イベントデータを取り出し、
反映する
S303
ファイル情報を
情報取得済みとして登録する
S305
情報未取得の
ディレクトリが
存在する？
S311
N
Y
情報未取得のディレクトリの
ファイル情報取得と
ファイル操作イベントの付け込み
S312
終了

【図６】


開始
一次ストレージの
不足が深刻か？
S401
N
Y
アーカイブ済みでかつ
マイグレート済みでない
ファイルを見つける
S411
未処理のファイルがある？
S412
N
Y
対象ファイルのリリース要求
S413
最終アクセス時刻が所定の時刻
以前でかつアーカイブ無効の
ファイルを見つける
S421
未処理のファイルがある？
S422
Y
対象ファイルのパス名を求める
S423
対象ファイルのアーカイブ要求
S424
終了

【図７】


運用サーバ
アプリケーション
111
ファイルシステム
制御部
ファイルデータ
領域割り当て
制御部
121
FSボリューム
管理部
122
112
ボリューム
管理部
仮想
ボリューム
管理部
123
ディスク
ドライバ
124
113
102
HSM制御サーバ
131
HSM管理部
133
ファイル配置ポリシ
データベース
ボリュームマップ
データベース
132
103
操作部
105
一次
ストレージ
1
二次
ストレージ
2

【図８】

FSボリューム構成情報テーブル

FS ボリューム 通番	FS内 ブロック範囲	割り当て ボリューム	ボリューム内 ブロック範囲
1	0 ～ 9999	vol_001	0 ～ 9999
2	10000 ～ 49999	disk_003	0 ～ 39999

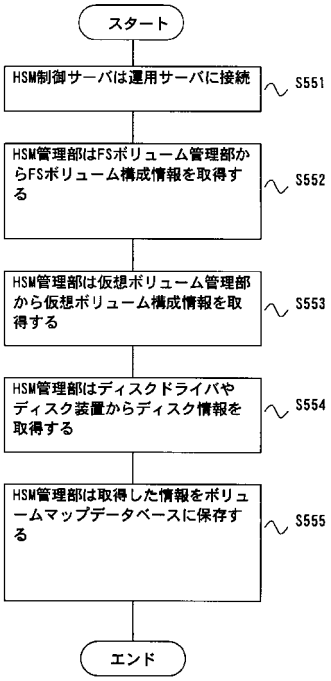
仮想ボリューム構成情報テーブル

仮想 ボリューム ID	構成タイプ	割り当てディスク群
vol_001	ミラー	disk_001, disk_002
vol_002	ストライプ	disk_004, disk_005, disk_006, disk_007
vol_003	コンカチ	disk_008, disk_009, disk_010

ディスク情報テーブル

ディスク ID	ディスクタイプ	ディスク装置情報
disk_001	単体	冗長性なし
disk_002	単体	冗長性なし
disk_003	RAID-0	高速ディスク、冗長性なし
・ ・ ・		

【図９】



【図１０】

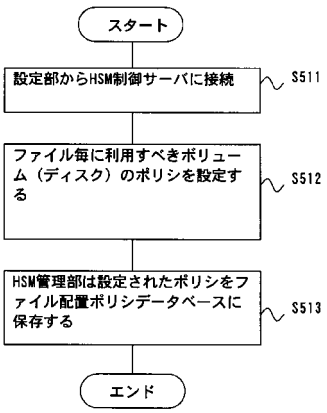
ファイル条件指定ポリシ設定テーブル

対象ファイル条件	ファイル作成先 FSボリューム条件	ファイル復元先 FSボリューム条件
ファイル所有者 = root	信頼性高	信頼性高
ファイル名 = "～.doc"	vol_001	vol_001
ファイル名 = "～.tmp"	アクセス性能高	アクセス性能高
ファイルサイズ ≦ 1MB	指定不可	vol_001
その他すべて	信頼性高	アクセス性能高

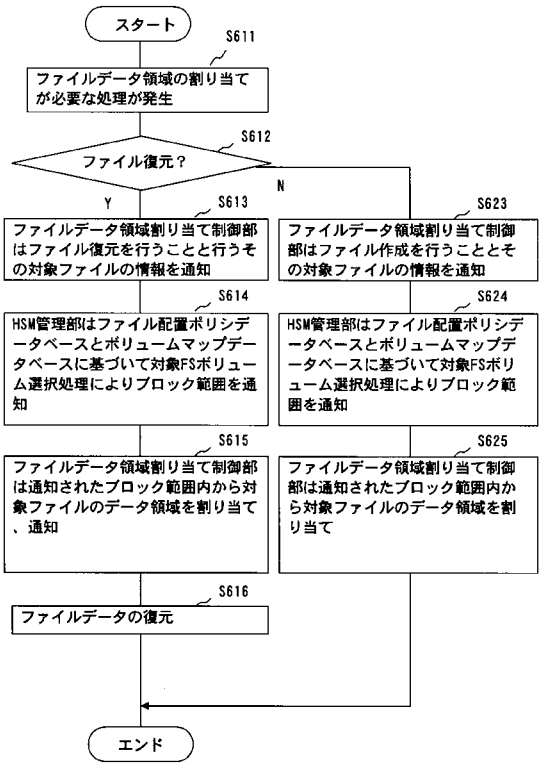
個別ファイルポリシ設定テーブル

ファイル inode 番号	ファイル作成先 FSボリューム条件	ファイル復元先 FSボリューム条件
inode=201 (FileA)	指定不可	信頼性高
inode=310 (TempData)	指定不可	アクセス性能高

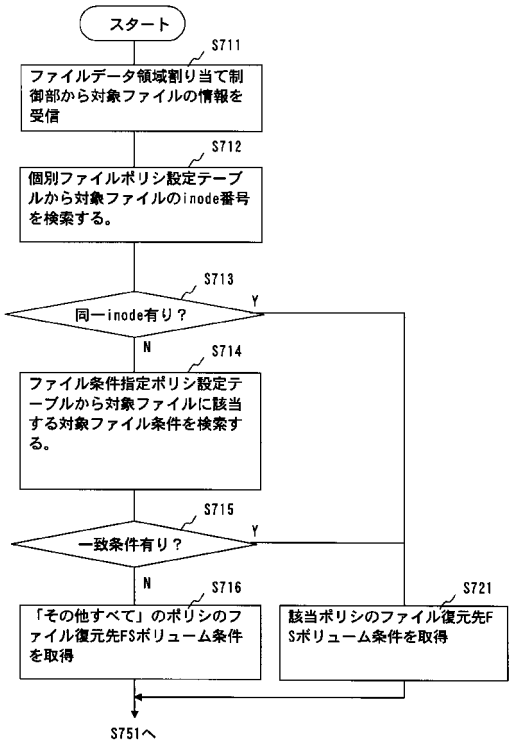
【図１１】



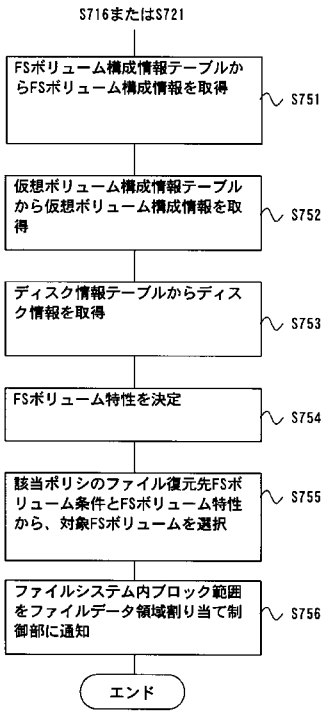
【図１２】



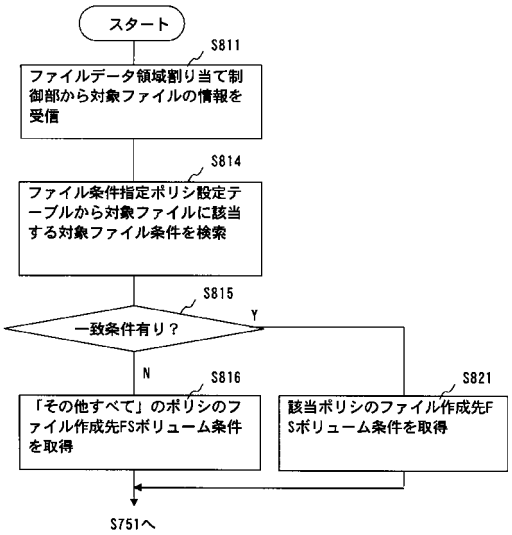
【図１３】



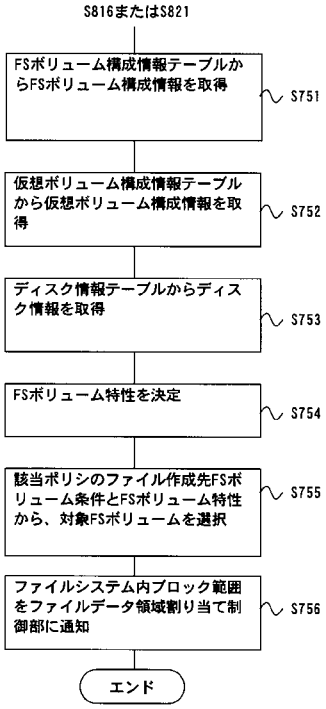
【図１４】



【図１５】



【図１６】


S816またはS821
FSボリューム構成情報テーブルからFSボリューム構成情報を取得
S751
仮想ボリューム構成情報テーブルから仮想ボリューム構成情報を取得
S752
ディスク情報テーブルからディスク情報を取得
S753
FSボリューム特性を決定
S754
該当ポリシのファイル作成先FSボリューム条件とFSボリューム特性から、対象FSボリュームを選択
S755
ファイルシステム内ブロック範囲をファイルデータ領域割り当て制御部に通知
S756
エンド