

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4480479号
(P4480479)

(45) 発行日 平成22年6月16日(2010.6.16)

(24) 登録日 平成22年3月26日(2010.3.26)

(51) Int.Cl.

F I

G 0 6 F 12/00 (2006.01)

G 0 6 F 12/00 5 2 0 P

G 0 6 F 3/06 (2006.01)

G 0 6 F 3/06 3 0 1 A

請求項の数 2 (全 13 頁)

(21) 出願番号 特願2004-177482 (P2004-177482)
 (22) 出願日 平成16年6月15日(2004.6.15)
 (65) 公開番号 特開2006-4011 (P2006-4011A)
 (43) 公開日 平成18年1月5日(2006.1.5)
 審査請求日 平成19年3月26日(2007.3.26)

(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 100080001
 弁理士 筒井 大和
 (72) 発明者 徳田 晴介
 東京都国分寺市東恋ヶ窪一丁目280番地
 株式会社日立製作所中央研究所内
 (72) 発明者 中村 隆喜
 東京都国分寺市東恋ヶ窪一丁目280番地
 株式会社日立製作所中央研究所内
 審査官 桜井 茂行

最終頁に続く

(54) 【発明の名称】 ストレージシステム

(57) 【特許請求の範囲】

【請求項 1】

異なる特性を持つ複数のデバイスと、前記複数のデバイスを操作するデバイスドライバと、前記デバイスのうち、少なくとも2つ以上のデバイスから構成する1つ、または複数のファイルシステムとからなるストレージシステムであって、

前記ファイルシステムは、

ファイルシステム作成、拡張、マウントなどのファイルシステム操作時にデバイスごとの特性を取得するデバイス特性取得部と、

前記デバイス特性取得部が得るデバイスごとの特性を記憶するデバイス特性保持部と、

ファイル格納処理のファイル管理情報の準備を行うファイル操作時に、ファイルの用途、属性を解釈し、その特性を特性フラグとして前記ファイル管理情報に登録するファイル用途解釈部と、

ファイル格納処理の実際にファイル書き込みを行うファイル操作時に、前記ファイル管理情報の特性フラグと前記デバイス特性保持部の情報とからファイルの格納デバイスを判定、決定する格納デバイス判定部と、

前記ファイル管理情報内の特性フラグを参照、変更するインタフェースと、

前記ファイル管理情報内の特性フラグが変更された際、ファイルに対してロックを取り、特性フラグの変更内容に対応したデバイスへファイルをマイグレーションする手段とを有することを特徴とするストレージシステム。

【請求項 2】

10

20

異なる特性を持つ複数のデバイスと、前記複数のデバイスを操作するデバイスドライバと、前記複数のデバイスを統合し論理的なデバイスとして管理するオペレーティングシステム内の論理デバイス管理手段と、その論理デバイス管理手段により作成した論理的なデバイスから構成する１つ、または複数のファイルシステムとからなるストレージシステムであって、

前記論理デバイス管理手段は、

論理デバイスの作成などの論理デバイス操作時に、デバイスごとの特性を取得し論理デバイスの管理領域に記憶するデバイス特性取得部を有し、

前記ファイルシステムは、

ファイルシステム作成、拡張、マウントなどのファイルシステム操作時に、論理デバイスの管理領域からデバイスごとの特性情報を読み出す手段と、

読み出したデバイスごとの特性を記憶するデバイス特性保持部と、

ファイル格納処理のファイル管理情報の準備を行うファイル操作時に、ファイルの用途、属性を解釈し、その特性を特性フラグとして前記ファイル管理情報に登録するファイル用途解釈部と、

ファイル格納処理の実際にファイル書き込みを行うファイル操作時に、前記ファイル管理情報の特性フラグと前記デバイス特性保持部の情報とからファイルの格納デバイスを判定、決定する格納デバイス判定部と、

前記ファイル管理情報内の特性フラグを参照、変更するインタフェースと、

前記ファイル管理情報内の特性フラグが変更された際、ファイルに対してロックを取り、特性フラグの変更内容に対応したデバイスへファイルをマイグレーションする手段とを有することを特徴とするストレージシステム。

【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におけるファイルの格納技術に関し、特に、ファイル用途に応じた最適なデバイスへのファイルの格納に適用する有効な技術に関するものである。

【背景技術】

【０００２】

多くの計算機システムにおいて、データの格納、取り出しはファイル単位で行う。これは、オペレーティングシステム（ＯＳ）の一部であるファイルシステムが記憶装置（デバイス）を抽象化しデータの格納、読み込み処理を行うことによって実現する。

【０００３】

図８は、ファイルシステムが扱う一般的なファイルの構造を示した図である。

【０００４】

ファイル７０１は、ファイル本体７０３とその管理情報であるｉノード７０２からなる。ｉノードは、実際の記憶媒体であるデバイスの識別情報やファイルの格納デバイスに関する位置情報、ファイルの拡張子を含むファイル名、ファイルのサイズなどの情報を含む。

【０００５】

ファイル７０１を扱うファイルシステムの作成処理、およびファイルシステムによるファイル７０１の格納処理について、図９を用いて説明する。

【０００６】

まず、ファイルシステムの作成処理について説明する。

【０００７】

ユーザ空間でファイルシステム作成コマンドＣ８１１が実行されると、カーネル空間において、初期化／運用準備部８１２が、ファイルシステム作成対象のデバイスのファイルシステム制御ブロック初期化と運用準備を行い、ファイルシステム８０１を作成する。

【０００８】

次に、ファイル格納処理について説明する。

【0009】

その格納処理は、ファイル管理情報であるiノードの準備と、実際のファイル書き込みとの2段階の処理からなる。第1段階の処理として、openシステムコールSC821が発行されると、カーネル空間のファイルシステム801で、準備部822がファイルiノードの準備を行う。そして、準備部822は、そのファイルiノードと関連付けられたファイルディスクリプタFS1をユーザ空間に返す。

【0010】

第2段階の処理として、ユーザ空間から、writeシステムコール831が、先のopenシステムコールの戻り値であったファイルディスクリプタFS2を引数として発行する。

10

【0011】

すると、空き領域確保部833は、カーネル空間のファイルシステム801でファイルディスクリプタFS2を通じてiノードを参照して、デバイスの空き領域を確保し、書き込み要求部834に実際の書き込み要求を出す。以上2段階の処理によりファイル格納を行う。

【発明の開示】

【発明が解決しようとする課題】

【0012】

ところが、上記のような計算機システムでは、次のような問題点があることが本発明者により見いだされた。

20

【0013】

近年、データ格納を行うストレージシステムが処理するファイルの種別は、電子メールなどのテキストファイル、電子文書ファイル、科学技術計算による大サイズの出力ファイル、データベースファイルなどと多岐に及ぶ。そして、ファイルを格納するデバイスは、ファイルの用途に応じて多様な特性をもつようになった。

【0014】

例えば、その特性として、法により規制されているWORM用、高速なファイルアクセス(ランダム、シーケンシャル)用、ビジネスの継続が要求されるデータを扱う同期コピー用、外部接続する比較的安価なデバイス用などがある。

30

【0015】

そこで、複数のファイルを格納する場合、図10に示すように、ファイルの用途に応じた特性を持つデバイス902ごとに1つつファイルシステム801を構成する形態で計算機システム901を運用していた。このとき、デバイスドライバ903は、各ファイルシステムの要求に対応するデバイスに発行する。

【0016】

この形態による運用では、ユーザは手間をかけて、ファイル用途に応じたファイルシステム801を選択し、ファイルを格納しなければならなかった。つまり、データ管理コストの増大という問題が生じている。

【0017】

40

本発明の目的は、特性の異なる複数のデバイスを1つのファイルシステム(マウントポイント)として提供し、ユーザが意識することなく、ファイルシステムがファイル用途に応じた最適なデバイスへファイルを格納することにより、データ管理コストを低減する技術を提供することにある。

【0018】

本発明の前記ならびにその他の目的と新規な特徴は、本明細書の記述および添付図面から明らかになるであろう。

【課題を解決するための手段】

【0019】

本願において開示される発明のうち、代表的なものの概要を簡単に説明すれば、次のと

50

おりである。

【 0 0 2 0 】

本発明は、異なる特性を持つ複数のデバイスと、前記複数のデバイスを操作するデバイスドライバと、前記デバイスのうち、少なくとも2つ以上のデバイスから構成する1つ、または複数のファイルシステムとからなるストレージシステムであって、前記ファイルシステムは、ファイルシステム作成、拡張、マウントなどのファイルシステム操作時にデバイスごとの特性を取得するデバイス特性取得部と、前記デバイス特性取得部が得るデバイスごとの特性を記憶するデバイス特性保持部とを有するものである。

【 0 0 2 1 】

そして、前記ファイルシステムは、ファイル格納処理のファイル管理情報の準備を行うファイル操作時に、ファイルの用途、属性を解釈し、その特性を特性フラグとしてファイル管理情報に登録するファイル用途解釈部と、ファイル格納処理の実際にファイル書き込みを行うファイル操作時に、ファイル管理情報の特性フラグと前記デバイス特性保持部の情報とからファイルの格納デバイスを判定、決定する格納デバイス判定部とを有するものである。

10

【 0 0 2 2 】

また、本願の発明の内容をより具体的に説明すれば以下の通りである。

【 0 0 2 3 】

本発明は、ファイルシステム作成、拡張、マウントなどのファイルシステム操作時と、ファイル格納処理時に新たな処理を追加する。具体的にその新たな処理は、ファイルシステム操作時に、デバイスごとの特性を取得、決定し記憶する。

20

【 0 0 2 4 】

その取得、決定方法は、(1) デバイスドライバによる特性取得、(2) 汎用デバイスの特性について予め登録したデータベースへの問い合わせによる特性取得、(3) 指定したデバイス特性へのRAID (Redundant Array of Inexpensive Disks) キャッシュなどの制御変更指示による特性決定、(4) 実計測によるデバイス特性評価、などによりデバイス特性の取得、決定を行う。

【 0 0 2 5 】

(1)の方法は、デバイス特性情報が登録されており、その情報をデバイスドライバが取得可能であるデバイスについて行う。(3)の方法は、デバイスが、OSが認識するRAID装置内の論理的なデバイスであり、かつそのデバイスごとにRAIDキャッシュなどの制御変更可能である場合に行う。(1)~(3)のデバイスに該当しない場合、(4)の方法で、数種のアクセスパターン要求を行い、その応答時間を計測することによりデバイス特性を決定する。

30

【 0 0 2 6 】

また、ファイル格納処理時の新たな処理は、ファイル管理情報の準備を行うファイル操作時に、ファイルの用途、属性を解釈し、ファイルの特性を表す特性フラグをファイル管理情報に拡張属性として登録する。続いて、実際にファイル書き込みを行うファイル操作時に、ファイル管理情報内の特性フラグと保持しているデバイス特性の情報からファイルの格納先デバイスを判定する。

40

【発明の効果】

【 0 0 2 7 】

本願において開示される発明のうち、代表的なものによって得られる効果を簡単に説明すれば以下のとおりである。

【 0 0 2 8 】

ユーザが意識することなく、用途に応じたファイル格納を行うことができ、それによって管理コストの低減が可能となる。

【発明を実施するための最良の形態】

【 0 0 2 9 】

以下、本発明の実施の形態を図面に基づいて詳細に説明する。なお、実施の形態を説明

50

するための全図において、同一の部材には原則として同一の符号を付し、その繰り返しの説明は省略する。

【 0 0 3 0 】

図 1 は、本発明の一実施の形態によるストレージシステムの構成を示すブロック図、図 2 は、図 1 のストレージシステムのファイルシステムが扱うファイルの構造を示す説明図、図 3 は、図 1 のストレージシステムに設けられたデバイス特性保持部における構成の一例を示す説明図、図 4 は、図 1 のストレージシステムに設けられたデバイス特性データベースにおける構成の一例を示す説明図、図 5 は、図 1 のストレージシステムに設けられたデバイス特性取得部における処理を示すフローチャート、図 6 は、図 1 のストレージシステムに設けられたファイル用途解釈部の処理を示すフローチャート、図 7 は、図 1 のストレージシステムに設けられた格納デバイス判定部の処理を示すフローチャートである。

10

【 0 0 3 1 】

本実施の形態において、ストレージシステムは、図 1 に示すように、計算機システム 1 0 1、および複数のデバイス 1 0 2 a ~ 1 0 2 z から構成されている。計算機システム 1 0 1 のオペレーティングシステムは、ユーザ空間、およびカーネル空間からなる。

【 0 0 3 2 】

カーネル空間では、オペレーティングシステムの機能の一部としてファイルシステム 1 0 3、デバイスドライバ 1 0 4 が動作し、ファイルシステム 1 0 3 からデバイスドライバ 1 0 4 へデバイス特性の取得要求を出すデバイス特性取得インタフェースが存在する。

【 0 0 3 3 】

20

ファイルシステム 1 0 3 は、初期化 / 運用準備部 8 1 2、準備部 8 2 2、書き込み要求部 8 3 4、デバイス特性取得部 1 1 2、デバイス特性保持部 1 1 3、デバイス特性データベース 1 1 4、ファイル用途解釈部 1 2 2、および格納デバイス判定部 1 3 2 から構成されている。

【 0 0 3 4 】

デバイス 1 0 2 a ~ 1 0 2 z は、例えばハードディスクなどからなり、異なる特性情報 1 1 6 a ~ 1 1 6 z を持ち、これらデバイス 1 0 2 a ~ 1 0 2 z を 1 つのファイルシステム 1 0 3 として提供する。

【 0 0 3 5 】

ファイルシステム 1 0 3 において、ユーザ空間からファイルシステム作成コマンド C 1 1 1 が発行されると、初期化 / 運用準備部 8 1 2 によって図 8 に示したファイルシステム 1 0 3 を作成する。その後、デバイス特性取得部 1 1 2 が処理を行う。

30

【 0 0 3 6 】

その処理は、デバイス特性データベース 1 1 4、デバイス特性取得インタフェース 1 1 5 などを用いてデバイス 1 0 2 ごとの特性情報 1 1 6 を取得し、その情報をデバイス特性保持部 1 1 3 に記憶する。

【 0 0 3 7 】

そして、ファイル格納処理の第 1 段階で、ユーザ空間から open システムコール S C 1 2 1 が発行されると、従来の処理である i ノード準備部による処理後、ファイル用途解釈部 1 2 2 が動作する。その動作により、図 2 に示したファイル 2 0 1 のファイル管理情報 (i ノード) 7 0 2 に拡張属性としてファイルの特性を現す特性フラグ 2 0 2 を登録する。

40

【 0 0 3 8 】

次に、ファイル格納処理の第 2 段階である write システムコール S C 1 3 1 の発行時、空き領域確保部 8 3 3 (図 8) の処理を含む格納デバイス判定部 1 3 2 が処理を行う。

【 0 0 3 9 】

その処理後、ファイル書き込み要求部がファイル書き込み要求を出す。以上の処理を新たに追加することにより、ファイル用途に応じて格納先デバイスを変更するストレージシステムを実現する。

50

【 0 0 4 0 】

ここで、本実施の形態では、計算機システム 1 0 1 が提供するファイルシステム 1 0 3 は 1 つだが、ファイルシステム 1 0 3 が 2 つ以上であっても構わない。

【 0 0 4 1 】

また、デバイス特性処理部 1 1 2 の動作は、ファイルシステム作成時のみではなくファイルシステム拡張、マウントなどのファイルシステム操作時に行うことも考えられる。同じくファイル格納処理について、この実施形態では `open` システムコール、`write` システムコールを用いて説明したが、その他のファイル操作であっても構わない。

【 0 0 4 2 】

以下にデバイス特性取得部 1 1 2、ファイル用途解釈部 1 2 2、格納デバイス判定部 1 3 2 の詳細な処理についてそれぞれ述べる。 10

【 0 0 4 3 】

デバイス特性取得部 1 1 2 の処理フローを図 5 に示す。

【 0 0 4 4 】

まず、ステップ 5 0 1 の処理において、デバイス 1 0 2 は、OS が認識する RAID 装置内の論理的なデバイス(以後、論理ユニットと呼ぶ)であり、かつその論理ユニットごとに RAID キャッシュなどの制御が変更可能であるか判定を行う。

【 0 0 4 5 】

条件に該当しない場合、ステップ 5 0 2 に処理を移す。条件に該当する場合、ステップ 5 1 1 で、ファイルシステム作成コマンド C 1 1 1 実行時にデバイス 1 0 2 ごとに特性の指定があるか判定を行う。指定がない場合、ステップ 5 0 2 に処理を移す。 20

【 0 0 4 6 】

指定がある場合、ステップ 5 1 2 で指定された特性に RAID キャッシュなどの制御を変更するよう RAID 装置に指示を出す。そしてステップ 5 0 7 へ進む。

【 0 0 4 7 】

ステップ 5 0 2 の処理においては、OS が `proc` ファイルシステムを搭載しているか判定する。搭載していない場合、ステップ 5 0 3 へ処理を移す。

【 0 0 4 8 】

搭載している場合、ステップ 5 2 1 の処理で、`proc` ファイルシステム内にデバイス 1 0 2 それぞれの特性情報があるか判定を行う。特性情報がない場合、ステップ 5 0 3 へ処理を移す。特性情報がある場合、ステップ 5 0 7 へ進む。 30

【 0 0 4 9 】

ここで、`proc` ファイルシステムは OS のデバイス認識時に得る情報を保持している。例えば SCSI デバイスの場合、`/proc/scsi` ディレクトリ以下に OS が認識している SCSI デバイスに関する情報がある。

【 0 0 5 0 】

ステップ 5 0 3 では、デバイス特性取得インタフェース 1 1 5 を用いデバイス 1 0 2 ごとの特性情報取得をデバイスドライバ 1 0 4 に指示し、ステップ 5 0 4 に進む。デバイス特性取得インタフェース 1 1 5 として `ioctl` システムコールなどを用いる。

【 0 0 5 1 】

ステップ 5 0 4 の処理において、デバイスドライバ 1 0 4 は、デバイス 1 0 2 ごとの特性情報 1 1 6 の取得要求を出し、取得に成功したか判定する。 40

【 0 0 5 2 】

取得に失敗した場合、ステップ 5 0 5 へ進む。取得に成功した場合、ステップ 5 0 7 へ進む。この成功した場合とは、デバイス 1 0 2 に特性情報 1 1 6 が登録されており、デバイスドライバ 1 0 4 よりその特性情報 1 1 6 の取得が可能である場合である。

【 0 0 5 3 】

例えば、SCSI デバイスの場合、特性情報 1 1 6 は `inquiry` コマンドの応答領域に登録されているとすると、デバイスドライバ 1 0 4 は `inquiry` コマンドを発行し、その応答としてデバイス特性情報 1 1 6 を得る。 50

【 0 0 5 4 】

ステップ 5 0 5 の処理で、デバイス 1 0 2 ごとのペンダモデルなどのデバイス識別情報が得られるか判定する。得られない場合、ステップ 5 0 6 へ処理を移す。得られた場合、ステップ 5 3 1 の処理で、それらデバイス識別情報を用いてデバイス特性データベース 1 1 4 に問い合わせ、デバイス 1 0 2 ごとの特性情報が登録されているか判定する。

【 0 0 5 5 】

登録されていない場合、ステップ 5 0 6 へ処理を移す。登録されている場合、ステップ 5 0 7 の処理へ進む。ステップ 5 0 5 の処理でのデバイス識別情報取得は、p r o c ファイルシステムなどを利用する。

【 0 0 5 6 】

ステップ 5 0 6 では、ランダムアクセスの R e a d、シーケンシャルアクセスの w r i t e などの複数のアクセスパターンをデバイスドライバ 1 0 4 に発行し、その応答時間を計測することによりデバイス 1 0 2 ごとの特性を評価する(デバイス特性評価)。そしてステップ 5 0 7 の処理へ進む。

【 0 0 5 7 】

ステップ 5 0 7 の処理では、上述のステップで得たデバイス 1 0 2 ごとの特性情報をデバイス特性保持部 1 1 3 へ記憶する。以上がデバイス特性取得部 1 1 2 の処理フローである。

【 0 0 5 8 】

この処理フローについて、ファイルシステム作成コマンド C 1 1 1 実行時にデバイス 1 0 2 ごとの特性を指定されていれば、ステップ 5 0 2 以降へ移ることなくステップ 5 0 7 へ進み、指定された特性をデバイス特性保持部 1 1 3 に記憶することも可能である。

【 0 0 5 9 】

また、デバイス 1 0 2 は、R A I D 装置内の論理ユニットである場合、デバイス 1 0 2 に特性情報 1 1 6 を登録することも考えられる。論理ユニットは、1 つまたは複数の R A I D 装置内の実デバイスから作成する。

【 0 0 6 0 】

よって、その論理ユニット作成時に、デバイスドライバによるデバイス情報取得要求の応答領域に特性情報 1 1 6 を登録する。例えば、S C S I デバイスドライバが発行する i n q u i r y コマンド応答領域に登録する。

【 0 0 6 1 】

この論理ユニット作成時、デバイスの特性が指定されれば、R A I D 装置はその指定された特性を登録する。また、R A I D 装置内の論理ユニットごとの制御を変更可能ならば、指定された特性となるよう制御を変更し、その特性を登録する。特性の指定がなければ、論理ユニット作成時のデバイス制御情報を特性として登録する。

【 0 0 6 2 】

図 3 は、デバイス特性保持部 1 1 3 を示している。

【 0 0 6 3 】

保持部の行は、デバイス特性情報であり、W O R M 用、高速なファイルアクセス、スループット性能などが考えられる。また列は、ファイルシステムを構成するデバイスであり、個々のデバイスを表すデバイス識別子、またデバイスごとのブロック番号の範囲などで表す。そして、デバイスごとに有するそれぞれ特性情報を Y、有していないものは N で表している。

【 0 0 6 4 】

スループット性能の場合は性能値により表す。このデバイス特性保持部 1 1 3 は、ファイルシステムを構成するデバイス内のファイルシステム制御ブロックに記憶し、ファイルシステムをマウントするときに読み出す。もしくはマウントするときに動的に生成する。

【 0 0 6 5 】

図 4 は、デバイス特性データベース 1 1 4 を示している。

【 0 0 6 6 】

10

20

30

40

50

このデータベースは汎用的なデバイスについて、その特性情報を登録している。データベースの行は、図3のデバイス特性保持部113と同様にデバイス特性情報であり、列は、ベンダモデルなどのデバイスを識別する情報である。

【0067】

デバイスごとの有する特性の表示は、デバイス特性保持部113と同様である。ここでデバイス特性データベース114は、ファイルシステムから参照できればファイルシステム内になくとも構わない。

【0068】

本実施の形態では、ファイルシステムが複数のデバイスを統合し、1つのファイルシステムとして提供している。そこで、計算機システム101のオペレーティングシステムが複数のデバイスを統合して単一の論理的なデバイスとして管理する手段(論理デバイス管理手段と呼ぶ)を有している場合、論理デバイス管理手段内部でデバイス特性取得部112の処理を実行することも可能である。

【0069】

その実施形態は、論理デバイス管理手段により特性の異なる複数のデバイスを統合し、1つの論理的なデバイスを作成する。そしてその論理的なデバイスからファイルシステムを作成する。

【0070】

この形態において、論理デバイス管理手段が論理的なデバイスを作成するときに、デバイス特性取得部の処理を行う。その処理は、デバイスごとの特性情報を取得し、論理デバイスの管理領域に保存する。そして、ファイルシステム作成処理時に、論理デバイスの管理領域からデバイスごとの特性情報を読み出し、読み出した特性情報をデバイス特性保持部に記憶する。

【0071】

次に、ファイル用途解釈部122の処理におけるフローチャートを図6に示し、格納デバイス判定部132の処理におけるフローチャートを図7に示す。

【0072】

まず、図6に示すように、ステップ601の処理において、openシステムコールSC121発行時に引数としてファイルの特性を表すフラグが指定されているか判定する。指定されている場合、ステップ603へ進む。指定されていない場合、ステップ602の処理へ移る。ここでファイルの特性とは、そのファイルの格納に適するデバイスの特性を意味する。

【0073】

ステップ602の処理で、ファイルの用途、属性を解釈し、ファイルの特性を決定する。そしてステップ603の処理へ進む。このとき、ファイルiノード702内のファイル拡張子、ファイルサイズ、またキーワード情報などを解釈する。

【0074】

具体的にその解釈方法は、ファイルの拡張子が、.txt, .docであればランダムアクセス、.emlならばWORMの特性、ファイルサイズがあるしきい値より大きければシーケンシャルアクセス、小さければランダムアクセスの特性、またキーワードとしてデータベースファイルの記述があればランダムアクセスを有すると解釈する。

【0075】

ステップ603の処理では、上述の手段により決定したファイル特性をファイルのiノード702に特性フラグ202という拡張属性として登録する。以上がファイル用途解釈部の処理フローである。

【0076】

続いて、格納デバイス判定部132における処理のフローチャートを図7を用いて説明する。

【0077】

ステップ611で、ファイルiノード702内の特性フラグ202とデバイス特性保持

10

20

30

40

50

部 1 1 3 の情報とからファイルの格納先デバイスを判定する。そして、判定結果であるデバイスについて、空き領域確保部 8 3 3 (図 9) によるデバイスの空き領域確保処理を実行する。

【 0 0 7 8 】

次のステップ 6 1 2 の処理で、空き領域の確保に成功した場合、格納デバイス判定部 1 3 2 の処理を終了する。空き領域の確保に失敗した場合、再度ステップ 6 1 1 の処理に戻り、別デバイスに書き込むかエラー終了するか判定を行う。以上が格納デバイス判定部 1 3 2 の処理である。

【 0 0 7 9 】

ファイル i ノード 7 0 2 内の特性フラグ 2 0 2 を利用してファイルマイグレーションも可能である。その方法は、ファイル i ノードの拡張属性を変更する `setxattr` システムコールなどのファイル操作を行い、特性フラグ 2 0 2 を変更する。

【 0 0 8 0 】

そして、ファイルに対してロックを取り、特性フラグの変更に応じたデバイスへファイルをマイグレーションする。このとき、ファイルシステムビューは変化しない。

【 0 0 8 1 】

この実施の形態におけるデバイス特性保持部 1 0 2 、デバイス特性データベース 1 0 3 の情報、またファイル用途解釈部 1 1 1 の解釈基準データ、格納デバイス判定部の判定基準データは、参照、変更、追記可能である。その方法は、`proc` ファイルシステムや `ioctl` システムコールなどを利用し行う。

【 0 0 8 2 】

それにより、本実施の形態では、用途に応じたファイル格納を行うことができ、管理コストを低減することができる。

【 0 0 8 3 】

以上、本発明者によってなされた発明を実施の形態に基づき具体的に説明したが、本発明は前記実施の形態に限定されるものではなく、その要旨を逸脱しない範囲で種々変更可能であることはいうまでもない。

【産業上の利用可能性】

【 0 0 8 4 】

本発明は、ストレージシステムにおいて、特性の異なる複数のデバイスを 1 つのファイルシステムとして提供し、データ管理コストを低減する技術に適している。

【図面の簡単な説明】

【 0 0 8 5 】

【図 1】本発明の一実施の形態によるストレージシステムの構成を示すブロック図である。

【図 2】図 1 のストレージシステムのファイルシステムが扱うファイルの構造を示す説明図である。

【図 3】図 1 のストレージシステムに設けられたデバイス特性保持部における構成の一例を示す説明図である。

【図 4】図 1 のストレージシステムに設けられたデバイス特性データベースにおける構成の一例を示す説明図である。

【図 5】図 1 のストレージシステムに設けられたデバイス特性取得部における処理を示すフローチャートである。

【図 6】図 1 のストレージシステムに設けられたファイル用途解釈部の処理を示すフローチャートである。

【図 7】図 1 のストレージシステムに設けられた格納デバイス判定部の処理を示すフローチャートである。

【図 8】本発明者が検討した計算機システムのファイルシステムが扱う一般的なファイルの構造を示した説明図である。

【図 9】本発明者が検討したオペレーティングシステムにおけるファイルシステムの作成

10

20

30

40

50

処理とファイル格納処理とを示すブロック図である。

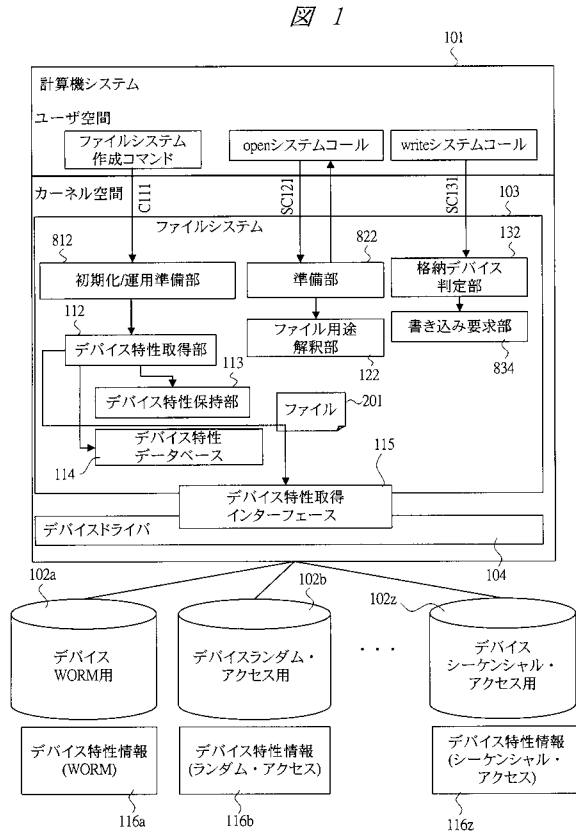
【図 10】図 9 のオペレーティングシステムにおいて特性の異なるデバイスを用いる場合の運用形態を示す説明図である。

【符号の説明】

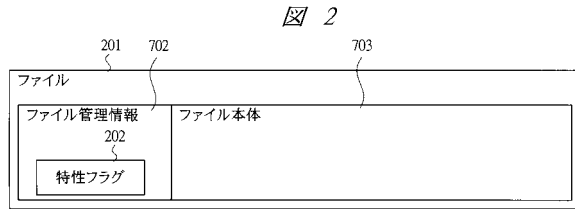
【 0 0 8 6 】

1 0 1	計算機システム	
1 0 2 a ~ 1 0 2 z	デバイス	
1 0 3	ファイルシステム	
1 0 4	デバイスドライバ	
1 1 2	デバイス特性取得部	10
1 1 3	デバイス特性保持部	
1 1 4	デバイス特性データベース	
1 1 5	デバイス特性取得インタフェース	
1 1 6 a ~ 1 1 6 z	特性情報	
1 2 2	ファイル用途解釈部	
1 3 2	格納デバイス判定部	
2 0 1	ファイル	
2 0 2	特性フラグ	
8 0 2	ファイル管理情報 (i ノード)	
8 1 2	初期化 / 運用準備部	20
8 2 2	準備部	
8 3 3	空き領域確保部	
8 3 4	書き込み要求部	
7 0 1	ファイル	
7 0 2	i ノード	
7 0 3	ファイル本体	
8 0 1	ファイルシステム	
9 0 1	計算機システム	

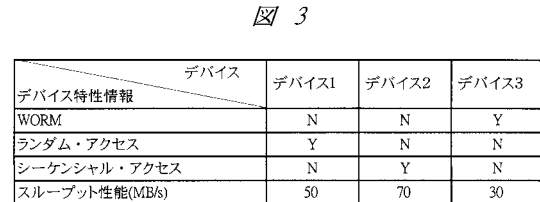
【 図 1 】



【圖 2】



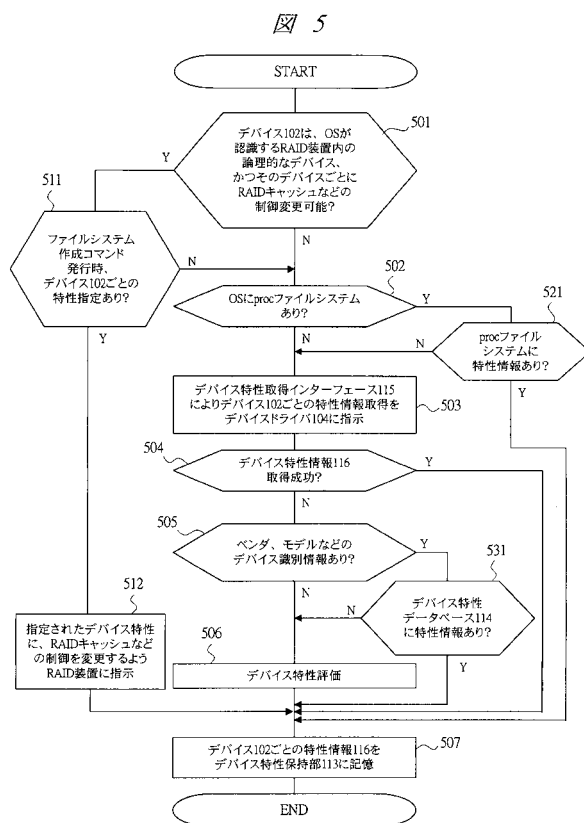
【圖 3】



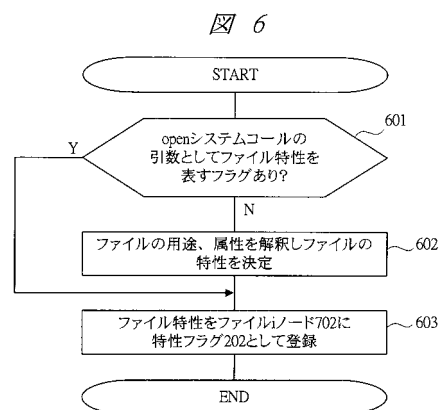
【 図 4 】



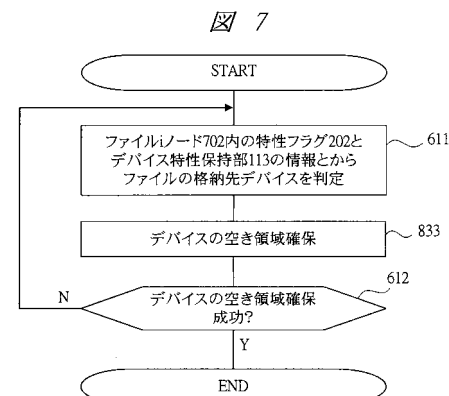
【 図 5 】



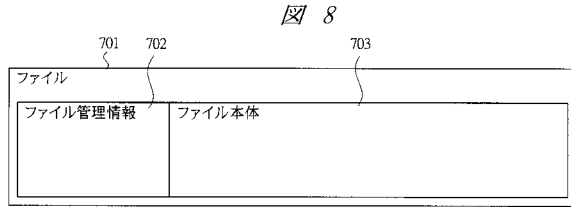
【 図 6 】



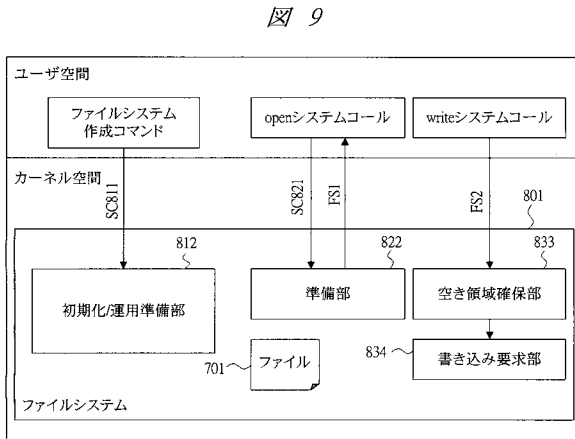
【圖 7】



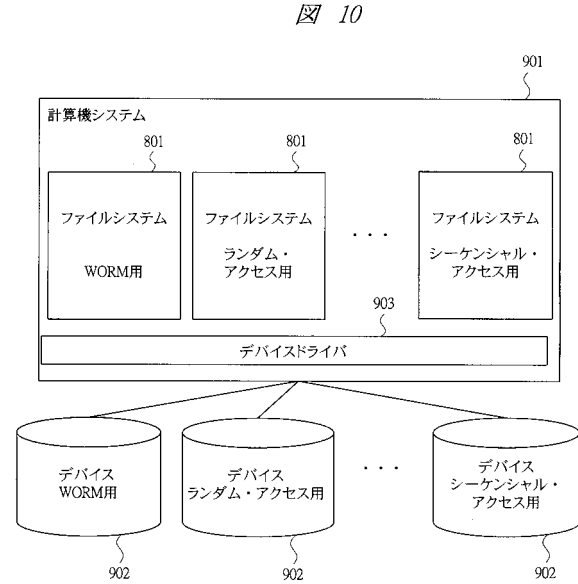
【図 8】



【図 9】



【図 10】



フロントページの続き

(56)参考文献 特開2003-131908(JP,A)
特開2004-070403(JP,A)
特開2003-345514(JP,A)
特開2004-070971(JP,A)
国際公開第2004/109517(WO,A1)

(58)調査した分野(Int.Cl., DB名)
G06F 12/00
G06F 3/06