(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号
特許第5309043号
(P5309043)
(45)発行日　平成25年10月9日(2013.10.9)　(24)登録日　平成25年7月5日(2013.7.5)

(51)Int.Cl.　F I
G06F 3/06 (2006.01)　G06F 3/06 301A

請求項の数 14　外国語出願　(全 30 頁)

(21)出願番号　特願2010-16440(P2010-16440)
(22)出願日　平成22年1月28日(2010.1.28)
(65)公開番号　特開2010-182302(P2010-182302A)
(43)公開日　平成22年8月19日(2010.8.19)
審査請求日　平成24年7月24日(2012.7.24)
(31)優先権主張番号　12/365,566
(32)優先日　平成21年2月4日(2009.2.4)
(33)優先権主張国　米国(US)

(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　100093861
弁理士　大賀　眞司
(74)代理人　100129218
弁理士　百本　宏之
(72)発明者　兼田　泰典
アメリカ合衆国　カリフォルニア州　９５１２９　サンノゼ　コーデリア・アベニュー　１２０８

審査官　稲葉　崇

最終頁に続く

(54)【発明の名称】ストレージシステム及びストレージシステムでの重複データ削除のための方法

(57)【特許請求の範囲】
【請求項１】
ストレージ制御装置と、
チャンクプールを形成する複数のチャンクに分割されている複数のストレージ装置と、
前記ストレージ制御装置をホストコンピュータに接続するように構成したネットワークインターフェースと、
を備え、
前記ストレージ制御装置はメモリを備え、当該メモリのプログラムを実行してデータストレージボリュームを生成し、前記ネットワークインターフェースを経由して前記データストレージボリュームを前記ホストコンピュータに利用可能にするストレージシステムであって、
前記メモリは、
前記ホストコンピュータから前記データストレージボリュームに向けられた書込コマンドに関連したデータに対応する識別子の複数を夫々有する複数の識別子テーブルと、
前記ホストコンピュータから前記ストレージ制御装置への書込みコマンドに伴うメタデータの複数と、前記複数の識別子テーブルの夫々の情報とを有し、前記一つ又は複数のメタデータを前記複数の識別子テーブルの何れかに割り付けるメタデータテーブルと、
を有し、
前記複数の識別子テーブルの夫々は、前記識別子に対応するデータが格納された前記チャンクの情報が当該識別子に割り付けられて構成され、

前記ストレージ制御装置は、前記メモリのプログラムを実行して、
前記ホストコンピュータからの前記データストレージボリュームに向けられた書込コマンドを受信すると、前記識別子を計算し、
前記メタデータテーブルを参照して、前記複数の識別子テーブルのうち、前記書込みコマンドに伴う前記メタデータに割り付けられた識別子テーブルを選択し、
前記選択された識別子テーブを参照して、前記計算された識別子の固有性を前記チャンク毎に確認し、
前記計算された識別子が前記識別子テーブルにないものであれば、
前記書込コマンドに関連した前記データに対応する識別子を生成し、
前記チャンクプールから前記データストレージボリュームの前記書込みコマンドで指定されたデータ位置へ前記チャンクを割り付け、前記書込コマンドに関連した前記データを、前記割り付けられたチャンクに格納し、
前記生成された識別子と前記割り付けられたチャンクの情報とに基づいて前記識別子テーブルを更新し、
一方、前記計算された識別子が前記識別子テーブルに存在するものであれば、前記書込みコマンドに対応するデータを格納しない、
ことを特徴とするストレージシステム。

【請求項２】
前記識別子は、前記ホストコンピュータから受け取った前記データのSecure Hash Algorithm（ＳＨＡ）値を含む、
ことを特徴とする請求項１によるストレージシステム。

【請求項３】
前記識別子は、前記データのSecure Hash Algorithm（ＳＨＡ）値とハッシュ競合の数との組み合わせを含む、
ことを特徴とする請求項１によるストレージシステム。

【請求項４】
前記ストレージシステムは、前記ホストコンピュータから受け取った前記データが格納されているか否かを決定する前に、前記データを格納する、
ことを特徴とする請求項１によるストレージシステム。

【請求項５】
前記ストレージシステムは、前記データストレージボリュームに前記データを格納した後に、前記識別子の計算および前記決定を、非同期に実行する、
ことを特徴とする請求項４によるストレージシステム。

【請求項６】
前記ストレージシステムを管理コンピュータに接続するように構成したマネージメントネットワークインターフェースをさらに備え、
前記メタデータが前記管理コンピュータから登録される、
ことを特徴とする請求項１によるストレージシステム。

【請求項７】
前記識別子が、ハッシュ値に競合が検出された時に割り当てられる連続した番号をさらに含む、
ことを特徴とする請求項３によるストレージシステム。

【請求項８】
前記メタデータが前記ホストコンピュータのＯＳタイプである、
ことを特徴とする請求項１によるストレージシステム。

【請求項９】
前記ストレージ制御装置は、前記ホストコンピュータから受け取ったオブジェクト中のデータから識別子を計算する、
ことを特徴とする請求項１によるストレージシステム。

【請求項１０】

前記識別子は、前記オブジェクト中の前記データのSecure Hash Algorithm（ＳＨＡ）値と、前記オブジェクトのデータサイズとを含む、
ことを特徴とする請求項９によるストレージシステム。
【請求項１１】
前記オブジェクトは、データ、ファイル名を含み、前記メタデータは当該ファイル名のファイルに対するものである、
ことを特徴とする請求項９によるストレージシステム。
【請求項１２】
前記メタデータは前記ファイル名のファイルに対する特徴を表すものであり、前記ストレージ制御装置は、一つのファイル名に対して複数の当該メタデータを設定することができる、
ことを特徴とする請求項１１によるストレージシステム。
【請求項１３】
前記ストレージ制御装置は、複数の識別子テーブルの中から所定の２つ以上の識別子テーブルの相関性をチェックし、そのチェック結果に基づいて当該二つ以上の識別子テーブルを統合し、データの重複の削除に基づいて未使用チャンクを生成する、
ことを特徴とする請求項１によるストレージシステム。
【請求項１４】
ストレージ制御装置と、
チャンクプールを形成する複数のチャンクに分割されている複数のストレージ装置と、
前記ストレージ制御装置をホストコンピュータに接続するように構成したネットワークインターフェースと、
を備え、
前記ストレージ制御装置はメモリを備え、当該メモリのプログラムを実行してデータストレージボリュームを生成し、前記ネットワークインターフェースを経由して前記データストレージボリュームを前記ホストコンピュータに利用可能にし、
前記メモリは、
前記ホストコンピュータから前記データストレージボリュームに向けられた書込コマンドに関連したデータに対応する識別子の複数を夫々有する複数の識別子テーブルと、
前記ホストコンピュータから前記ストレージ制御装置への書込みコマンドに伴うメタデータの複数と、前記複数の識別子テーブルの夫々の情報とを有し、前記一つ又は複数のメタデータを前記複数の識別子テーブルの何れかに割り付けるメタデータテーブルと、
を有し、
前記複数の識別子テーブルの夫々は、前記識別子に対応するデータが格納された前記チャンクの情報が当該識別子に割り付けられて構成されているストレージシステムでの重複データ削除のための方法であって、
前記ストレージ制御装置は、前記メモリのプログラムを実行して、
前記ホストコンピュータからの前記データストレージボリュームに向けられた書込コマンドを受信すると、前記識別子を計算し、
前記メタデータテーブルを参照して、前記複数の識別子テーブルのうち、前記書込みコマンドに伴う前記メタデータに割り付けられた識別子テーブルを選択し、
前記選択された識別子テーブを参照して、前記計算された識別子の固有性を前記チャンク毎に確認し、
前記計算された識別子が前記識別子テーブルにないものであれば、
前記書込コマンドに関連した前記データに対応する識別子を生成し、
前記チャンクプールから前記データストレージボリュームの前記書込みコマンドで指定されたデータ位置へ前記チャンクを割り付け、前記書込コマンドに関連した前記データを、前記割り付けられたチャンクに格納し、
前記生成された識別子と前記割り付けられたチャンクの情報とに基づいて前記識別子テーブルを更新し、

一方、前記計算された識別子が前記識別子テーブルに存在するものであれば、前記書込みコマンドに対応するデータを格納しない、

ことを特徴とするストレージシステムでの重複データ削除のための方法。

【発明の詳細な説明】

【技術分野】

【０００１】

０００１　本発明は、全般的には、データストレージシステムにおいて効率的に記憶容量を使用するために、重複したデータブロックあるいはファイルを削除する方法と装置に関するものである。具体的には、本発明は、メタデータグルーピングによって重複したデータブロックおよびファイルを速やかに発見することに関連するものである。

【背景技術】

【０００２】

０００２　重複削除は、重複したデータストリーム、データブロックあるいはファイルを削除するよう設計された機能であり、データストレージシステムにおいて、非常に効率的なデータ保存を提供するためにデータストレージシステムとバックアップ装置に実施される。１つの実施例では、重複したデータを削除するために、識別子が、データ自体から生成される。識別子はハッシュまたはＭＤ５の、ＳＨＡ（セキュアハッシュアルゴリズム）として生成することができる。データストレージシステムがデータを受け取ると、識別子が計算される。それから、データストレージシステムは、同じ識別子が識別子テーブルに既に格納されているか否かをチェックする。同じ識別子が識別子テーブルにある場合、受信データはデータストレージシステムに保存されない。識別子テーブルに同じ識別子がない場合には、受信データが保存されることとなる。例えば２５６ビット以上の十分なビット長を備えたハッシュの利用は、競合をめったに引き起こさないことには注目すべきである。大きなデータ量を管理するには、多数の識別子の値を計算し別に管理せねばならない。この理由で、非常に大量の記憶容量を備えたデータストレージシステムにおいては、識別子を確認するのに長い時間がかかることがありうる。

【０００３】

０００３　例えば、米国特許６，９２８，５２６号、題名“効率的なデータストレージシステム”は、重複したデータを削除する方法を開示している。データは、データ自体を使用して生成される識別子によって削除される。

【先行技術文献】

【特許文献】

【０００４】

【特許文献１】米国特許６，９２８，５２６号公報

【発明の概要】

【発明が解決しようとする課題】

【０００５】

しかしながら、従来の技術は、メタデータに関連する複数のグループを定義することにより、データストレージシステムにおいて識別子を確認するためのＣＰＵ時間を縮小する方法及びシステムを提供していない。

【０００６】

０００４　本発明の実施形態は、先行技術における前述の１つ以上の不備に取り組み、データストレージシステムまたはバックアップの装置において、重複したデータを削除する機能によって、識別子を確認するためのＣＰＵ時間を縮小する方法と装置を提供する。

【課題を解決するための手段】

【０００７】

０００５　本発明の技法の一態様に従って、データストレージボリューム、データストレージボリュームに関連したメタデータを格納するメモリ、ストレージシステムをホストコンピュータに接続するように構成されたネットワークインターフェース、中央処理装

置、を備えたストレージシステムが準備される。ストレージシステムは、ホストコンピュータから受け取ったデータから識別子を計算し、データがデータストレージボリュームに格納されているか否かを識別子とメタデータによって決定する。

【０００８】

０００６　本発明の技法の別の態様に従って、データストレージボリューム、データストレージボリュームに関連したメタデータを格納するメモリ、ストレージシステムをホストコンピュータに接続するように構成されたネットワークインターフェース、中央処理装置、を備えたストレージシステムが準備される。ストレージシステムは、ホストコンピュータから受け取ったオブジェクト中のデータから識別子を計算し、識別子、メモリに格納されているメタデータ、およびオブジェクト中のメタデータによってデータストレージボリュームにデータが格納されているか否かを決定する。

【０００９】

０００７　さらに、本発明の技法の別の態様に従って、チャンクプールを形成する多数のチャンクに分割されている多数のデータストレージ装置、ストレージシステムをホストコンピュータに接続するように構成されたネットワークインターフェース、中央処理装置とメモリを含むストレージコントローラ、を備えたストレージシステムにおいて実行される方法が提供される。発明の方法は、データストレージボリュームを供給し、ネットワークインターフェースによってホストコンピュータにデータストレージボリュームを利用可能にするステップと、ホストコンピュータからのデータストレージボリュームに向けられた書込コマンドを受信すると、書込コマンドに関連したデータに対応する識別子を計算するステップと、少なくとも１つの識別子グループへメタデータに基づいて識別子をグループ化するステップと、メタデータに関連した識別子グループ内で識別子の固有性を確認するステップと、識別子が識別子グループ内で固有な場合、チャンクプールからデータストレージボリュームに少なくとも１つのチャンクを割り付け、割り付けられた少なくとも１つのチャンクに書込コマンドに関連したデータを格納するステップと、を含んでいる。

【００１０】

０００８　本発明に関連するさらなる態様は、一部分は以後の記述で述べられ、一部分は記述から明白になるか、もしくは発明の実施によって知るところとなるであろう。本発明の態様、特に、次の詳細な記述および最後尾に設けた請求項において指摘した態様は、要素と諸要素の組合せによって実現し達成することが出来る。

【００１１】

０００９　先の記述および以下の記述の両方ともに、典型的なものであって説明のためだけのものであり、主張している発明や応用を如何なる方法においても制限するようには意図されていないことは、当然理解すべきことである。

【００１２】

００１０　添付の図面は、この明細書に組み入れられ、この明細書の一部を構成して本発明の実施形態を例証し、記述と共に、本発明の技術の原理について説明し例示する役割を担っている。

具体的には：

【図面の簡単な説明】

【００１３】

【図１ａ】００１１　図１（ａ）は、本発明の概念が適用され得る情報システムの典型的な実施形態を示す。

【図１ｂ】図１（ｂ）は、本発明の概念が適用され得る情報システムの典型的な実施形態を示す。

【図２】００１２　図２は、データボリューム管理テーブルの典型的な実施形態を示す。

【図３】００１３　図３は、チャンクテーブルの典型的な実施形態を示す。

【図４】００１４　図４は固有のチャンク識別番号を使用したチャンクの識別を示す。

【図５】００１５　図５は、チャンク状態テーブルの典型的な実施形態を示す。

【図６】００１６　図６（ａ）および６（ｂ）は、メタデータマッピングテーブルの典型的な実施形態を示す。
【図７】００１７　図７は、識別子テーブルの典型的な実施形態を示す。
【図８】００１８　図８は、ボリューム生成プロセスの典型的な実施形態を示す。
【図９】００１９　図９は、メタデータマッピングテーブルの初期化プロセスの典型的な実施形態を示す。
【図１０】００２０　図１０は、書込み動作の典型的な実施形態を示す。
【図１１】００２１　図１１（ａ）は、書き込みコマンドの典型的な実施形態を示す。００２２　図１１（ｂ）は、読み出しコマンドの典型的な実施形態を示す。
【図１２】００２３　図１２は、読み出し動作の典型的な実施形態を示す。
【図１３ａ】００２４　図１３（ａ）は、本発明の方法が適用することができる情報システムの別の典型的な実施形態を示す。
【図１３ｂ】図１３（ｂ）は、本発明の方法が適用することができる情報システムの別の典型的な実施形態を示す。
【図１４】００２５　図１４は、ファイル管理テーブルの別の典型的な実施形態を示す。
【図１５ａ】００２６　図１５（ａ）は、メタデータマッピングテーブルの別の典型的な実施形態を示す。
【図１５ｂ】図１５（ｂ）は、メタデータマッピングテーブルの別の典型的な実施形態を示す。
【図１５ｃ】図１５（ｃ）は、メタデータマッピングテーブルの別の典型的な実施形態を示す。
【図１５ｄ】図１５（ｄ）は、メタデータマッピングテーブルの別の典型的な実施形態を示す。
【図１６】００２７　図１６は、書き込み動作の別の典型的な実施形態を示す。
【図１７】００２８　図１７（ａ）は、書き込みコマンドの別の典型的な実施形態を示す。００２９　図１７（ｂ）は、読み出しコマンドの別の典型的な実施形態を示す。
【図１８】００３０　図１８は、読み出し動作の別の典型的な実施形態を示す。
【図１９】００３１　図１９は、書き込み動作の別の典型的な実施形態を示す。
【図２０】００３２　図２０は、ファイル管理テーブルの別の典型的な実施形態を示す。
【図２１】００３３　図２１は、識別子テーブル統合プログラムによって表示されたスクリーンの典型的な実施形態を示す。
【図２２】００３４　図２２は、識別子テーブル統合プロセスの典型的な実施形態を示す。
【図２３】００３５　図２３は、書き込み動作のための修正済のチャンクテーブルを示す。
【図２４】００３６　図２４は書き込みプロセスを示す。ここでの欄は、書き込みコマンドを受け取った後に非同期的に実行される。
【図２５】００３７　図２５は、バックグラウンドでの重複削除のプロセスの典型的な実施形態を示す。
【図２６】００３８　図２６は、本発明のシステムが実施され得るコンピュータプラットフォームの典型的な実施形態を示す。
【発明を実施するための形態】
【００１４】
００３９　以下の詳細な説明において、添付の図面が参照されるであろうが、その中において、同一の要素は類似の数字で示されている。前述の添付の図面は、本発明の原理と一致する特定の実施形態および実施例を図示の形で示しているのであって、限定するために示しているのではない。これらの実施例は、当分野の業者が発明を実施することを可能にするほどに十分に詳細に記述されている。また、これ以外の実施例も可能であり、構造

の変更および／または諸要素の置換が本発明の範囲および精神から外れることなく行なう事が可能なことも理解すべきである。したがって、次の詳細な説明は、限定する意味において解釈されるべきではない。

【００１５】

００４０　発明の第１実施形態のシステム構成を以下記述する。図１は、本発明の方法が適用できる情報システムの典型的な実施形態を示す。第１の実施形態の情報システムは、少なくともホストコンピュータ１０、ストレージ装置１００、管理コンピュータ５００、データネットワーク５０およびマネジメントネットワーク９０からなる。

【００１６】

００４１　ホストコンピュータ１０について以下記述する。少なくとも１台のホストコンピュータ１０がデータネットワーク５０によってストレージ装置１００に接続される。特にこの実施形態では、６台のホストコンピュータ１０ａ、１０ｂ、１０ｃ、１０ｄ、１０ｅおよび１０ｆが接続されている。少なくとも１つのＯＳ１３がホストコンピュータ上で実行される。アプリケーションプログラム１４はＯＳ１３の上で実行することができる。ＯＳ１３およびアプリケーションプログラム１４のためのファイルおよびデータは、ストレージ装置１００によって提供されるデータボリュームに格納される。ＯＳ１３およびアプリケーションプログラム１４は，ストレージ装置１００へ書き込みおよび／または読み出しコマンドを出す。ホストコンピュータ１０ａ、１０ｂおよび１０ｃはタイプＡのＯＳを実行し、ホストコンピュータ１０ｄ、１０ｅおよび１０ｆはタイプＢのＯＳを実行する。ＯＳタイプはベンダーのＯＳ名およびバージョン番号を使用して定義することが出来る。

【００１７】

００４２　ストレージ装置１ＯＯについて以下記述する。情報システムは、ストレージコントローラ１５０および１台以上のＨＤＤ１０１から構成される少なくとも１つのストレージ装置１００を備えている。ストレージ装置１００は、１つ以上のデータボリューム１１１をホストコンピュータ１０に供給する。

【００１８】

００４３　管理コンピュータ５００について以下記述する。情報システムは、マネジメントネットワーク９０によってストレージ装置１００を接続している、少なくとも１台の管理コンピュータ５００を備えている。

【００１９】

００４４　データネットワーク５０について以下記述する。ホストコンピュータ１０およびストレージ装置１００はデータネットワーク５０によって接続している。
この実施形態においては、データネットワーク５０はファイバーチャンネルプロトコルを使用し実装されている。しかしながら、イーサネット（登録商標）やインフィニバンドのような他のネットワーク接続も、この目的に同様に使用することができる。ネットワークスイッチとハブはデータネットワーク５０の構成要素を互いに接続するために使用することができる。
図１では、ファイバーチャンネルスイッチ５５（ＦＣＳＷ５５）が、データネットワーク５０の構成要素を互いに接続するために使用されている。この目的のために、ホストコンピュータ１０およびストレージ装置１００はファイバーチャンネルデータネットワーク５０に接続するために、１つ以上のファイバーチャンネルインターフェースボード（ＦＣＩＦ）を組込んでいる。

【００２０】

００４５　マネジメントネットワーク９０について以下記述する。ストレージ装置１００はマネジメントネットワーク９０によって管理コンピュータ５００に接続されている。この実施形態においてマネジメントネットワーク９０はイーサネット（登録商標）プロトコルを使用し実装されている。しかしながら、他のネットワーク相互接続や、他の接続方法が、この目的に同様に使用することができる。ネットワークスイッチとハブはマネジメントネットワーク９０の構成要素を互いに接続するために使用することができ、本発明

の本実施形態においては、ストレージ装置１００および管理コンピュータ５００は、イーサネット（登録商標）マネジメントネットワーク９０に接続するために１つ以上のイーサネット（登録商標）インタフェースボード（イーサＩＦ）を持っている。

【００２１】

００４６　ホストコンピュータ１０について以下詳細に記述する。ホストコンピュータ１０は、メモリ１２に格納したプログラムを実行するためのＣＰＵ１１、プログラムとデータを格納するためのメモリ１２、データネットワーク５０に接続するためのＦＣＩＦ１５、から構成されている。この実施形態では、ＣＰＵ１１は、メモリ１２に格納された少なくとも３つのプログラムを実行する。

【００２２】

００４７　本発明のこの実施形態においては、メモリ１２は、オペレーティングシステムプログラム１３（ＯＳ１３）、アプリケーションプログラム１４、ＯＳ１３及び／またはアプリケーションプログラム１４をインストールするためのインストーラプログラム１５、を格納している。

【００２３】

００４８　管理コンピュータ５００について以下詳細に記述する。管理コンピュータ５００は、メモリ５２０に格納したプログラムを実行するためのＣＰＵ５１０、プログラムとデータを格納するためのメモリ５２０、マネジメントネットワーク９０に接続するためのイーサＩＦ５９０、から構成されている。

【００２４】

００４９　ＣＰＵ５１０は、メモリ５２０に格納された少なくとも３つのプログラムを実行する。この実施形態において、メモリ５２０は、ストレージ装置１００に、データボリュームを供給することを要求するための、データボリューム供給要求プログラム５２１、ストレージ装置１００への重複削除の範囲の定義の要求を出すための、重複削除範囲定義要求プログラム５２２、管理スクリーンを表示し、２台の識別子テーブルの違いの比率を計算し、識別子テーブル統合の要求を、ストレージ装置１００に出すための、識別子テーブル統合プログラム５２３、を格納している。

【００２５】

００５０　ストレージ装置１００について以下詳細に記述する。ストレージ装置１００は、データを格納するための１台以上のＨＤＤ１０１、およびホストコンピュータへのデータボリュームを提供するための１台以上のストレージコントローラ１５０から構成されている。ストレージコントローラ１５０は、メモリ１５２に格納したプログラムを実行するためのＣＰＵ１５１、プログラムおよびデータを格納するためのメモリ１５２、データネットワーク５０に接続するためのＦＣＩＦ１５５、ＨＤＤ１０１に接続するためのＳＡＴＡ　ＩＦ１５６（ＨＤＤにＦＣ、ＳＣＳＩ　ＳＡＳのような別のインターフェースがある場合、適切なインターフェースが実装されねばならない。）、ホストコンピュータから受け取り、ＨＤＤから読み出されたデータを格納するためのキャッシュ１５３、マネジメントネットワーク９０に接続するためのイーサＩＦ１５９、から構成されている。

【００２６】

００５１　ＣＰＵ　１５１は、メモリ１５２に格納された少なくとも４つのプログラムを実行する。この実施形態においては、メモリ１５２は、ホストコンピュータ１０からの、少なくとも読み出し／書込みコマンドに応答するための応答プログラム１６１、ボリュームを生成し、ホストコンピュータ１０にそれを割り付けるためのデータボリューム割付けプログラム１６２、メタデータマッピングテーブルを更新するための重複削除範囲定義プログラム１６３、ホストコンピュータ１０から転送されるデータからメッセージダイジェスト５（ＭＤ５）を計算し、同じ識別子が識別子テーブルに既に格納されているか否かを確認するための重複削除プログラム１６５を格納している。同じ識別子が見つからない場合には、重複削除プログラム１６５はデータを格納する。同じ識別子が見つかった場合には、重複削除プログラム１６５はデータを格納しない。この技術分野における同業者には、本発明のシステムは、ＭＤ５アルゴリズムにのみ限定されるものではないことは理

解されよう。ＳＨＡのような、他の適切なハッシュあるいはダイジェスト機能も、本発明に同様に適用することができる。

【００２７】

００５２　重複削除プログラム１６５におけるＭＤ５を計算し識別子機能を確認する機能は、計算と確認の加速のためにゲートアレイあるいはＦＰＧＡ（フィールドプログラマブルゲートアレイ）上にハードウェアロジックとして実装することもできる。

【００２８】

００５３　データボリューム管理テーブルについて以下詳細に記述する。データボリューム管理テーブル１６６は生成されたボリュームを管理する。図２に示されるように、データボリューム管理テーブル１６６は、ボリューム番号を格納するための“ボリューム番号”列１６６０１、データボリュームサイズ（ブロックの数）の格納のための“サイズ”列１６６０２、ボリューム生成時に、管理コンピュータ５００から付与するメタデータの格納のための“メタデータ”列１６６０３、チャンクテーブル番号の格納のための“チャンクテーブル番号”列１６６０４、から構成されている。

【００２９】

００５４　チャンクテーブルについて以下詳細に記述する。チャンクテーブル１６７は、データボリュームのＬＢＡ（論理的ブロックアドレス）と識別子の関係についての情報を格納する。識別子はＭＤ５および別のシーケンシャル番号から構成される。シーケンシャル番号は、ＭＤ５に競合が生じた場合に使用される。各データボリュームはそれぞれ自分のチャンクテーブル１６７を持つ。ＬＢＡは，アクセスブロック位置を指定するために、ホストコンピュータによって読み／書き動作に使用される。図３に示されるように、チャンクテーブル１６７は、ＬＢＡを格納するための“ＬＢＡ”列１６７０１および識別子を格納するための“識別子”列１６７０２から構成されている。

【００３０】

００５５　チャンク状態表について以下詳細に説明する。ストレージ装置１００のＨＤＤは多数のデータブロックで構成されている。データブロックのサイズは、通常、最近の製品では５１２バイトである。この実施形態では、チャンクは１つのデータブロックからなる。したがって、この実施形態では、チャンクのサイズは５１２バイトである。しかし、他のどのようなチャンクサイズも同様に使用することができる。チャンクはそれぞれ図４に示すように、一つずつ識別するために固有の番号を持っている。チャンク状態テーブル１７０は、チャンクの使用状況を管理する。図５は、この実施形態におけるチャンク状態表１７０の典型的な実施例を示す。チャンク状態表の典型的な実施形態はチャンク番号を格納するための“チャンク番号”列１７００１、チャンクが使用されているか否かを示す格納状態情報のための“状態”列１７００２、から構成されている。データが保存されている場合、重複削除プログラム１６５はチャンク状態表１７０を使用して、未使用のチャンクを探す。

【００３１】

００５６　メタデータマッピングテーブルについて以下詳細に記述する。メタデータマッピングテーブル１８０は、メタデータと識別子テーブル１８５の関係についての情報を格納する。データボリュームに書かれたデータは、メタデータマッピングテーブル１８０で指定された識別子テーブルによって確認される。図６に示されるように、メタデータマッピングテーブル１８０は、ボリュームに割り当てられたメタデータを保持するための“メタデータ”列１８００１と、識別子テーブル番号を保持するための“識別子テーブル番号”列１８００２と、から構成されている。

【００３２】

００５７　識別子テーブルについて以下詳細に記述する。識別子テーブル１８５は、チャンクの識別子、参照カウントおよびチャンク番号を管理する。識別子はＭＤ５および別のシーケンシャル番号から構成される。ＭＤ５に競合が生じた場合に、シーケンシャル番号が使用される。図７に示すように、識別子テーブル１８５は、識別子を格納するための「識別子」列１８５０１、チャンクテーブル１６７からの参照の数である参照カウント

の格納のための“参照カウント”列１８５０２、データが保存されているチャンク番号を格納するための“チャンク番号”列１８５０３、から構成されている。

【００３３】

００５８　ボリューム生成プロセスについて以下詳細に記述する。ボリューム生成プロセス８００は、図８を参照して説明する。

【００３４】

００５９　ステップ８１０：データボリュームの供給要求プログラム５２１は、ストレージコントローラ１５０上のデータボリューム割付けプログラム１６２に、サイズ（ブロックの数）およびメタデータとともに、データボリューム供給要求を発行する。（メタデータは、“ＯＳタイプＡ”、“ＯＳタイプＢ”などのようなものである。）任意のタイプのメタデータを管理者が指定することができる。

【００３５】

００６０　ステップ８２０：データボリューム割付けプログラム１６２はデータボリューム管理テーブル１６６を更新する。図２は、６つのボリューム生成が完了した場合を示す。データボリューム１１１ａ、１１１ｂおよび１１１ｃは“ＯＳタイプＡ”メタデータを伴って生成される。データボリューム１１１ｄ、１１１ｅおよび１１１ｆは“ＯＳタイプＢ”メタデータを伴って生成される。

【００３６】

００６１　メタデータマッピングテーブルを初期化するプロセスについて以下詳細に記述する。メタデータマッピングテーブル初期化プロセス９００は、図９を参照して説明する。

【００３７】

００６２　ステップ９１０：重複削除範囲定義要求プログラム５２２は、ストレージコントローラ１５０上の重複削除範囲定義プログラム１６３へメタデータと共に重複削除範囲定義要求を出す。

【００３８】

００６３　ステップ９２０：重複削除範囲定義プログラム１６３は新しいメタデータを受け取り、重複削除範囲定義プログラム１６３は受信したメタデータのために新しい識別子テーブル１８５を割り付ける。

【００３９】

００６４　ステップ９３０：重複削除範囲定義プログラム１６３はメタデータマッピングテーブルを更新する。図６（ａ）は、２つのメタデータ（“ＯＳタイプＡ”と“ＯＳタイプＢ”）を受け取った場合を示す。識別子テーブル１８５ａはメタデータ“ＯＳタイプＡ”に割り付けられる。識別子テーブル１８５ｂはメタデータ“ＯＳタイプＢ”に割り付けられる。

【００４０】

００６５　書き込み動作について以下詳細に記述する。書き込み動作１０００を、図１０を参照して説明する。書き込み動作は、ＯＳとアプリケーションソフトウエアがインストールされ実行されている間に実行される。この実施形態では、ＯＳタイプＡがデータボリューム１１１ａ、１１１ｂおよび１１１ｃにインストールされている。ＯＳタイプＢが、データボリューム１１１ｄ、１１１ｅおよび１１１ｆにインストールされている。非常に多数の書き込み動作が、実際のＯＳのインストール動作の間に実行されている事は注意すべきである。図１０は、応答プログラム１６１および重複削除プログラム１６５における典型的な処理フローを示す。ホストコンピュータ１０にあるインストーラプログラム１５は、書き込コマンドおよびデータをボリューム１１１へ発行する。この実施形態では、ホストコンピュータ１０ａがデータボリューム１１１ａを使用し、ホストコンピュータ１０ｂがデータボリューム１１１ｂを使用している、等。

【００４１】

００６６　ステップ１０１０：書き込みコマンドとデータを受け取る。書き込みコマンドはＬＢＡ、およびブロックの数についての情報を含んでいる。図１１（ａ）は書き込

みコマンドを示す。書き込みコマンドはコマンド種別（＝書き込み）、ＬＢＡ（＝データの位置）、およびブロックの数（＝データのサイズ）についての情報を含んでいる。

【００４２】

００６７　ステップ１０１２：ボリュームのメタデータはデータボリューム管理テーブル１６６から得られる。（書き込みコマンドを受け取る現在のデータボリュームが１１１ａである場合、メタデータは“ＯＳタイプＡ”である（図２を参照）。）

【００４３】

００６８　ステップ１０１４：識別子テーブルはメタデータによりメタデータマッピングテーブル１８０から選ばれる。（書き込みコマンドを受け取る現在のデータボリュームが１１１ａである場合、メタデータが“ＯＳタイプＡ”であるので、識別子テーブル１８５ａが選択されている。）下記ステップが各ブロックに対し実行される。

【００４４】

００６９　ステップ１０１６：ＭＤ５値が受け取ったデータから計算される。

【００４５】

００７０　ステップ１０１８：書き込みコマンドによって指定された現在のデータ位置が割り付けられたチャンクを有するか否かのチェックが、チャンクテーブルを参照することにより行なわれる。チャンクが既に割り付けられている場合、プロセスはステップ１０６０に進む。チャンクが割り付けられない場合、プロセスはステップ１０２０に進む。

【００４６】

００７１　ステップ１０２０：選択された識別子テーブルからの同じＭＤ５の値を持っている欄が列挙される。同じＭＤ５の値が見つかった場合、プロセスはステップ１０４０に進む。同じＭＤ５の値が見つからなかった場合、プロセスはステップ１０２２に進む。

【００４７】

００７２　ステップ１０２２：識別子が生成される。識別子はＭＤ５の値と０の組合せである。

【００４８】

００７３　ステップ１０２４：チャンク状態表１７０によって未使用のチャンクを得る。

【００４９】

００７４　ステップ１０２６：使用状態を示すためのチャンク状態テーブル１７０を更新する。

【００５０】

００７５　ステップ１０２８：チャンクへデータを格納する。

【００５１】

００７６　ステップ１０３０：選択された識別子テーブル１８５を更新する。（ステップ１０２２において）生成された識別子が格納される。参照カウントは１にセットされる。チャンク番号（ステップ１０２２で得られた）が格納される。

【００５２】

００７７　ステップ１０３２：チャンクテーブル１６７を更新する。（書込コマンドを受け取る現在データボリュームが１１１ａである場合、チャンクテーブル１６７ａが更新される。）生成された識別子は、現在のＬＢＡ欄に格納される。

【００５３】

００７８　ステップ１０４０：ＭＤ５の競合を回避するために、バイト単位データチェックを実行する。データが異ならない場合、動作はステップ１０５０に進む。データが異なる場合（ＭＤ５の競合）、ステップ１０４２に進む。（複数の同じＭＤ５が見つかった場合、バイト単位チェックを各々に実行する。）

【００５４】

００７９　ステップ１０４２：識別子を生成する。識別子は、ＭＤ５と現在の最大番号の次の番号となる新しいシーケンシャル番号の組合せである。（データが異なっても、

１つの同じＭＤ５が見つかった場合、シーケンシャル番号は１である。データが異なっても、２つの同じＭＤ５が見つかった場合、シーケンシャル番号は２である。）ステップ１０２４に進む。ＭＤ５はめったに競合しない。したがって、ステップ１０４２はまれな場合である。

【００５５】

００８０　ステップ１０５０：選択された識別子テーブル１８５を更新する。識別子に従い識別子テーブル１８５中の参照カウントを増加させる。

【００５６】

００８１　ステップ１０５２：チャンクテーブル１６７を更新する。識別子はＬＢＡに対応して格納される。

【００５７】

００８２　ステップ１０５４：データを廃棄する。

【００５８】

００８３　ステップ１０６０：識別子に従い識別子テーブル１８５の参照カウントを減少させる。

【００５９】

００８４　ステップ１０６２：参照カウントが０であるか否かをチェックする。参照カウントが０である場合には、ステップ１０６４に進む。参照カウントが０でない場合には、ステップ１０２０に進む。

【００６０】

００８５　ステップ１０６４：チャンクを空ける。

【００６１】

００８６　ステップ１０６６：選択された識別子テーブル１８５を更新する。識別子の欄は選択された識別子テーブルから消去される。ステップ１０２０に進む。

【００６２】

００８７　上に言及したように、特定の識別子テーブルはデータボリュームに割り当てられるメタデータに従って使用される。１つの識別子テーブルは複数のデータボリュームに使用することができる。データボリュームは、種々のＯＳを保存することができる。しかしながら、種々のＯＳは種々のデータを含んでいるので、１つの識別子テーブルでは巨大になってしまう。１つの巨大な識別子テーブルで識別子確認を実行するのには長い時間がかかる。この実施形態では、複数の識別子テーブルが使用される。１つの識別子テーブルはＯＳタイプＡに使用され、もう一方はＯＳタイプＢに使用される。各識別子テーブルは、２つのＯＳに対し１つの識別子テーブルの場合より小さくなる。メタデータによる分離の結果、識別子確認を実行するためのＣＰＵ時間はより短くなる。本発明を備えたデータストレージ装置の読み／書きアクセスの性能は改善される。識別子を確認するためのどのようなアルゴリズムでも、例えば二分木のようなものでも、使用することができる。

【００６３】

００８８　読み出し動作について以下詳細に記述する。読み出し動作１２００を図１２で説明する。ＯＳとアプリケーションが実行されている間に読み出し動作は実行される。ホストコンピュータ１０上の、ＯＳ１３、およびアプリケーションプログラム１４がボリューム１１１へ読み出しコマンドを発行する。この実施形態では、ホストコンピュータ１０ａはデータボリューム１１１ａを、ホストコンピュータ１０ｂはデータボリューム１１１ｂを、・・・使用する。

【００６４】

００８９　ステップ１２１０：読み出しコマンドを受け取る。読み出しコマンドは、ＬＢＡおよびブロックの数を含んでいる。図１１（ｂ）は読み出しコマンドを示す。読み出しコマンドは、コマンド種別（＝読み出し）、ＬＢＡ（＝データの位置）、ブロックの数（＝データのサイズ）を含んでいる。

【００６５】

００９０　ステップ１２１２：データボリューム管理テーブル１６６からボリューム

のメタデータを得る。（読み出しコマンドを受け取る現在データボリュームが１１１ａである場合、メタデータは“ＯＳタイプＡ”である（図２）。）

【００６６】

００９１　ステップ１２１４：メタデータに従いメタデータマッピングテーブル１８０から識別子テーブルを選択する。（読み出しコマンドを受け取る現在データボリュームが１１１ａである場合、メタデータが“ＯＳタイプＡ”であるので、識別子テーブル１８５ａが選択される。）

【００６７】

００９２　ステップ１２１６：チャンクテーブル１６７を選択する。（読み出しコマンドを受け取る現在データボリュームが１１１ａである場合、チャンクテーブル１６７ａが選択される。）各ブロックに対し下記のステップが実行される。

【００６８】

００９３　ステップ１２２０：ＬＢＡに対応する選択されたチャンクテーブル１６７から識別子を得る。

【００６９】

００９４　ステップ１２２２：識別子に対応している識別子テーブル１８５からチャンク番号を得る。

【００７０】

００９５　ステップ１２２４：チャンク番号が指定されたチャンクのデータを、ホストコンピュータへ転送する。

【００７１】

００９６　本発明の第２の実施形態について以下詳細に説明する。図１３は、本発明の方法が適用される情報システムの全体図の例を示す。第１の実施形態と第２の実施形態の違いは以下の通りである。

【００７２】

００９７　ストレージ装置１００はファイル管理プログラム１６４を有する。

【００７３】

００９８　ホストコンピュータ１０は、ファイルアクセスコマンド（図１７に示される）によってファイルストレージ装置へのアクセスを行う。

【００７４】

００９９　データネットワークはイーサネット（登録商標）である。イーサネット（登録商標）スイッチ８５（Ether　ＳＷ　８５）が相互接続のために使用されている。

【００７５】

０１００　ホストコンピュータ１０がイーサネット（登録商標）データネットワーク８０の接続のためにイーサネット（登録商標）インタフェースボード１８（イーサＩＦ１８）を持っている。

【００７６】

０１０１　ストレージ装置１００がイーサネット（登録商標）データネットワーク８０の接続のためにイーサネット（登録商標）インタフェースボード１５８（イーサＩＦ１５８）を持っている。

【００７７】

０１０２　ファイル管理プログラム１６４はファイル管理テーブル１９０を使用する。

ファイル管理テーブルはファイル名、ファイルサイズ、識別子およびメタデータの管理のために使用する。

【００７８】

０１０３　識別子テーブル１８５中の各欄はそれぞれ複数のチャンク番号を保持することができる。

【００７９】

０１０４　第２の実施形態のファイル管理テーブルについて以下詳細に説明する。図

１４に示すように、ファイル管理テーブル１９０は、ファイル名を保持するための“ファイル名”列１９００１（この実施形態では、ファイル名はディレクトリー名とファイル名からなる）、ファイルサイズを保持するための“ファイルサイズ”列１９００２、ＭＤ５および別の連続番号からなる識別子を保持するための“識別子”列１９００３、メタデータの保持のための“メタデータ”列１９００４、とから構成されている。１つのファイル名に対応する複数のメタデータを保持することができる。

【００８０】

０１０５　メタデータマッピングテーブルの初期化について以下詳細に説明する。メタデータマッピングテーブルの初期化のプロセスは第１の実施形態と同じである。しかしながら、生成者名、グループ名、組織名称、作成の時間スタンプ、生成されたアプリケーション名、ファイルタイプ、ホストコンピュータ名（物理的な及び／または仮想の）などなど、多様なメタデータを使用することができる。図１５は、この実施形態におけるメタデータマッピングテーブル１８０の例を示す。図１５（ａ）はファイルタイプによるメタデータマッピングテーブルを示す。図１５（ｂ）は組織名称によるメタデータマッピングテーブルを示す。図１５（ｃ）は作成時の時間スタンプによるメタデータマッピングテーブルを示す。

【００８１】

０１０６　第２の実施形態の書き込み動作について以下詳細に記述する。

【００８２】

０１０７　書き込み動作１６００を図１６により説明する。書き込み動作はＯＳ１３およびアプリケーションソフト１４から実行される。

【００８３】

０１０８　ステップ１６１０：書き込みコマンドとデータを受け取る。書き込みコマンドはファイル名、ファイルサイズおよびメタデータを含んでいる。図１７（ａ）は書き込みコマンドを示す。

【００８４】

０１０９　ステップ１６１６：データからのＭＤ５を計算する。

【００８５】

０１１０　ステップ１６１８：同じファイル名が既にファイル管理テーブル１９０に存在するか否かチェックする。同じファイル名が見つかった場合には、１６６０に進む。同じファイル名が見つからなかった場合には、ステップ１６２０に進む。

【００８６】

０１１１　ステップ１６２０：書き込みコマンドに保持されているメタデータに従いメタデータマッピングテーブル１８０から識別子テーブルを選ぶ。

【００８７】

０１１２　ステップ１６２１：選択された識別子テーブルから同じＭＤ５の値を持っている欄を列挙する。同じＭＤ５が見つかった場合には、ステップ１６４０に進む。同じＭＤ５が見つからなかった場合には、ステップ１６２２に進む。

【００８８】

０１１３　ステップ１６２２：識別子を生成する。識別子はＭＤ５とゼロの組合せである。

【００８９】

０１１４　ステップ１６２４：チャンク状態テーブルおよびファイルのサイズ（サイズはチャンクの数に変換されている）によって未使用のチャンクを得る。

【００９０】

０１１５　ステップ１６２６：使用状態を示すためにチャンク状態表１７０を更新する。

【００９１】

０１１６　ステップ１６２８：チャンクへデータを格納する。

【００９２】

０１１７　ステップ１６３０：選択された識別子テーブル１８５を更新する。生成された識別子を格納する。参照カウントは１にセットされる。チャンク番号（ステップ１６２４で得られた）を格納する。

【００９３】

０１１８　ステップ１６３２：ファイル管理テーブル１９０を更新する。ファイル名、ファイルサイズ、生成された識別子（ステップ１６２２において）およびメタデータを格納する。

【００９４】

０１１９　ステップ１６４０：ＭＤ５の競合を回避するためにバイト単位データチェックを行う。データが異ならない場合、ステップ１６５０に進む。データが異なる場合には（ＭＤ５競合）、ステップ１６４２に進む。（複数の同じＭＤ５が見つかった場合、各々に対しバイト単位チェックを行う。）

【００９５】

０１２０　ステップ１６４２：識別子を生成する。識別子は、ＭＤ５と現在の最大番号の次のシーケンシャル番号との組合せである。（データが異なっても、１つの同じＭＤ５が見つかった場合、シーケンシャル番号は１である。データが異なっても、２つの同じＭＤ５が見つかった場合、シーケンシャル番号は２である。）

【００９６】

０１２１　ステップ１６５０：選択された識別子テーブル１８５を更新する。識別子に従い識別子テーブル１８５の参照カウントを増加させる。

【００９７】

０１２２　ステップ１６５２：ファイル管理テーブル１９０を更新する。ファイル名、ファイルサイズ、生成された識別子およびメタデータを格納する。

【００９８】

０１２３　ステップ１６５４：データを廃棄する。

【００９９】

０１２４　ステップ１６６０：ファイル管理テーブルに保持されたメタデータによって識別子テーブルを選択する。

【０１００】

０１２５　ステップ１６６１：ファイル名に従いファイル管理テーブルから識別子を得る。

【０１０１】

０１２６　ステップ１６６２：識別子に従い識別子テーブル１８５の参照カウントを減少させる。

【０１０２】

０１２７　ステップ１６６３：参照カウントが０であるか否かチェックする。参照カウントが０である場合には、ステップ１６６４に進む。参照カウントが０でない場合は、ステップ１６２０に進む。

【０１０３】

０１２８　ステップ１６６４：識別子とともに欄に保持されているチャンク番号のチャンクを空ける。

【０１０４】

０１２９　ステップ１６６６：選択された識別子テーブル１８５を更新する。識別子を有する欄は選択された識別子テーブルから消去する。ステップ１６２０に進む。

【０１０５】

０１３０　上に述べたように、特定の識別子テーブルはファイルに割り当てられたメタデータに従って使用される。この実施形態では、複数の識別子テーブルが同様にメタデータマッピングテーブル１８０に従い使用されている。各識別子テーブルは、すべてのファイルに対し１つの識別子テーブルの場合よりもより小さくなる。メタデータによる分離の結果、識別子確認を実行するためのＣＰＵ時間はより短くなる。本発明を備えたデータ

ストレージ装置の読み／書きアクセスの性能が改善される。

【０１０６】

０１３１　第２の実施形態の読み出し動作について以下詳細に記述する。読み出し動作１８００は図１８で説明する。読み出し動作はＯＳ１３およびアプリケーションソフトウエア１４から実行される。

【０１０７】

０１３２　ステップ１８１０：読み出しコマンドを受け取る。読み出しコマンドはファイル名を含んでいる。図１７（ｂ）は読み出しコマンドを示す。

【０１０８】

０１３３　ステップ１８１２：ファイル名に従いファイル管理テーブル１９０からファイルのメタデータを得る。

【０１０９】

０１３４　ステップ１８１４：メタデータに従いメタデータマッピングテーブル１８０から識別子テーブルを選ぶ。

【０１１０】

０１３５　ステップ１８２０：ファイル名に従いファイル管理テーブル１９０から識別子を得る。

【０１１１】

０１３６　ステップ１８２２：識別子に従い識別子テーブル１８５からチャンク番号を得る。

【０１１２】

０１３７　ステップ１８２４：ファイル名、ファイルサイズおよびメタデータをホストコンピュータへ転送する。

【０１１３】

０１３８　ステップ１８２６：チャンク番号を指定したチャンクのデータをホストコンピュータへ転送する。

【０１１４】

０１３９　書き込み動作の変形について以下詳細に記述する。書き込み動作１９００の別の実施形態を図１９を参照して説明する。書き込み動作はＯＳ１３およびアプリケーションソフト１４から実行される。この書き込み動作では、ファイルはＭＤ５およびファイルのサイズによって重複削除される。識別子の一部であるシーケンシャル番号は使用しない。識別子はＭＤ５およびファイルのサイズから構成される。図２０に示すファイル管理テーブル１９０を書き込み動作に使用する。

【０１１５】

０１４０　ステップ１９１０：書き込みコマンドとデータを受け取る。書き込みコマンドはファイル名、ファイルサイズおよびメタデータを含んでいる。図１７（ａ）は書き込みコマンドを示す。

【０１１６】

０１４１　ステップ１９１６：データからのＭＤ５を計算する。

【０１１７】

０１４２　ステップ１９１８：同じファイル名がファイル管理テーブル１９０に既に存在するか否かチェックする。同じファイル名が見つかった場合には、ステップ１９６０に進む。同じファイル名が見つからなかった場合には、ステップ１９２０に進む。

【０１１８】

０１４３　ステップ１９２０：書き込コマンドに保持されているメタデータに従いメタデータマッピングテーブル１８０から識別子テーブルを選ぶ。

【０１１９】

０１４４　ステップ１９２１：同じ識別子（ＭＤ５の値、ファイルのサイズ）が選択された識別子テーブルに存在しているか否かを確認する。同じ識別子が見つかった場合には、ステップ１９５０に進む。同じ識別子が見つからなかった場合には、ステップ１９２

２に進む。
【０１２０】
０１４５　ステップ１９２２：識別子を生成する。識別子はＭＤ５とファイルサイズの組合せである。
【０１２１】
０１４６　ステップ１９２４：チャンク状態テーブルおよびファイルのサイズ（サイズはチャンクの数に変換されている）に従い未使用のチャンクを得る。
【０１２２】
０１４７　ステップ１９２６：使用状況を示すためにチャンク状態表１７０を更新する。
【０１２３】
０１４８　ステップ１９２８：チャンクへデータを格納する。
【０１２４】
０１４９　ステップ１９３０：選択した識別子テーブル１８５を更新する。識別子（ＭＤ５の値、ファイルのサイズ）を格納し、また、チャンク番号も格納する。参照カウントを１にセットする。
【０１２５】
０１５０　ステップ１９３２：ファイル管理テーブル１９０を更新する。ファイル名、ファイルサイズ、ＭＤ５およびメタデータを格納する。
【０１２６】
０１５１　ステップ１９５０：選択された識別子テーブル１８５を更新する。識別子に従い識別子テーブル１８５の参照カウントを増加させる。
【０１２７】
０１５２　ステップ１９５２：ファイル管理テーブル１９０を更新する。ファイル名、ファイルサイズ、ＭＤ５およびメタデータを格納する。
【０１２８】
０１５３　ステップ１９５４：データを廃棄する。
【０１２９】
０１５４　ステップ１９６０：ファイル管理テーブルに保持されたメタデータに従い識別子テーブルを選択する。
【０１３０】
０１５５　ステップ１９６１：ファイル名に従って識別子（ＭＤ５の値およびファイルのサイズ）を得る。
【０１３１】
０１５６　ステップ１９６２：識別子に従って識別子テーブル１８５参照カウントを減少させる。
【０１３２】
０１５７　ステップ１９６３：参照カウントが０であるか否かチェックする。参照カウントが０の場合には、ステップ１９６４に進む。参照カウントが０でない場合には、ステップ１９２０に進む。
【０１３３】
０１５８　ステップ１９６４：識別子と共に、欄に保持されているチャンク番号のチャンクを空ける。
【０１３４】
０１５９　ステップ１９６６：選択された識別子テーブル１８５を更新する。識別子を有する欄を選択した識別子テーブルから消去する。ステップ１９２０に進む。
【０１３５】
０１６０　識別子テーブル統合について以下詳細に記述する。管理コンピュータ５００は識別子テーブル統合プログラム５２３を持つことが出来る。識別子テーブル統合プログラム５２３は図２１に示すスクリーン２１００を表示する。メタデータおよび識別子テ

ーブル番号はスクリーン２１００上のテーブル２１０１に表示される。管理者はチェックボックスのチェックにより違いを比較するために２つのテーブルを選ぶことができる。図２１では、メタデータ“アカウンティング”および“ＨＲ”が選択されている。識別子テーブル統合プログラムは２つの識別子テーブル間の相関性をチェックし、違いの比率を計算する。違いの比率は表示枠２１０２に表示される。違いが大きい場合には、２つのテーブルを統合すべきでない。統合はより長い待ち時間を引き起こすかも知れない。小さな違いの場合には、２つのテーブルを統合してもよい。統合は重複削除により未使用チャンクを生み出す。管理者は２つのテーブルの統合のためにボタン２１０３を押すことができる。識別子テーブル統合プログラム５２３は、ストレージ装置１００へ識別子テーブル統合要求を出す。識別子テーブルの統合プロセスは、管理者からの要求なしの自動化にすることもでき得る。予め決定のしきい値を管理者がセットし、識別子テーブル統合プロセスの起動を決定するための比率と比較することができる。また、ストレージに十分なデータがあれば、比率が安定した量になるので、ストレージの使用状況のレベルを考慮すべきであろう。ストレージ装置１００は、図２２に示すように識別子テーブル統合プロセスをスタートする。

【０１３６】

０１６１　ステップ２２１０：識別子テーブル統合要求で指定された第１の識別子テーブルを選択する。

【０１３７】

０１６２　ステップ２２１２：識別子テーブル統合リクエストで指定された第２の識別子テーブルを選択する。

【０１３８】

０１６３　ステップ２２１４：ｉ＝０と初期化する。

【０１３９】

０１６４　ステップ２２２０：もし欄［ｉ］が最後の欄でなければ、２２５０へスキップする。

【０１４０】

０１６５　ステップ２２２２：欄［ｉ］の識別子（ＭＤ５の値およびファイルのサイズ）が第１の識別子テーブルに存在することを確認する。同じ識別子が存在する場合、ステップ２２３０に進む。（第１の識別子テーブルの識別子を保持する欄の番号にｊをセットする）。

【０１４１】

０１６６　ステップ２２２４：欄［ｉ］を第１の識別子テーブルに加える。

【０１４２】

０１６７　ステップ２２３０：欄［ｉ］に、チャンク番号が保持されているチャンクを空ける。

【０１４３】

０１６８　ステップ２２３２：次の方法で、第１の識別子テーブルの欄［ｊ］の参照カウントに第２の識別子テーブルの欄［ｉ］の参照カウント値を加える。（欄［ｊ］参照カウント＝欄［ｊ］参照カウント＋欄［ｉ］参照カウント）。

【０１４４】

０１６９　ステップ２２４０：ｉを増加させる。ステップ２２２０に進む。

【０１４５】

０１７０　ステップ２２５０：メタデータ管理テーブルを更新する（例えば、図１５（ａ）は図１５（ｄ）として更新される）

【０１４６】

０１７１　ステップ２２５２：第２の識別子テーブルを削除する。

【０１４７】

０１７２　第１の実施形態の書き込み動作の代替実施形態について以下記述する。書き込み動作２３００の別の実施形態は図２３、２４および２５を参照して説明する。書き

込み動作について、修正済のチャンクテーブルを、図２３に示している。チャンクテーブルはチャンク番号情報を格納する。ＬＢＡ欄にチャンク番号がある場合、それは、チャンク番号で指定したチャンクにＬＢＡに対するデータが一時的に格納されたことを意味する。この欄は、書込コマンドを受け取った後に非同期に実行されるであろう。対応する書き込みプロセスは図２４に示されている。

【０１４８】

０１７３　ステップ２４１０：書込コマンドを受け取る。書込コマンドは、ＬＢＡおよびブロックの番号を含んでいる。

【０１４９】

０１７４　ステップ２４１２：チャンク状態テーブル１７０に従って未使用のチャンクを得る。

【０１５０】

０１７５　ステップ２４１４：チャンクが使用されていることを示すためにチャンク状態テーブル１７０を更新する。

【０１５１】

０１７６　ステップ２４１６：チャンクへデータを格納する。

【０１５２】

０１７７　ステップ２４１８：チャンクテーブルを更新する。この時に、ステップ２４１２で受け取ったチャンク番号だけを、チャンクテーブルに格納する。

【０１５３】

０１７８　重複削除の評価は、書込コマンドを受け取った後に非同期に実行される。バックグラウンド重複削除プロセスを図２５に示す。バックグラウンド重複削除は、各データボリュームで周期的に実行される。図１０と図２５との差はステップ２５１５の中のみにある。

【０１５４】

０１７９　ステップ２５１５：テーブルから、一時的に格納したデータを列挙する。一時的に格納したデータは、チャンクテーブルのチャンク番号のチェックにより確認される。一時的に格納したデータは、チャンクテーブルにチャンク番号を有する。例えば、図２３において、ＬＢＡ２のデータはチャンク１０００３に一時的に格納される。ステップ２５１５に続くステップは、各ブロックに対し実行される。

【０１５５】

典型的なコンピュータプラットフォーム

０１８０　図２６は、発明の技法の実施形態を実装することが出来るコンピュータ／サーバシステム２６００の実施形態を示すブロック図である。システム２６００は、コンピュータ／サーバプラットフォーム２６０１、周辺機器２６０２およびネットワークリソース２６０３を備えている。

【０１５６】

０１８１　コンピュータプラットフォーム２６０１は、コンピュータプラットフォーム２６０１の様々な部分の間で情報通信するための、データバス２６０４あるいは他のコミュニケーション機構、ならびに、バス２６０４と接続した、情報処理や、他の計算処理および制御タスクを行なうプロセッサー２６０５を備えている。コンピュータプラットフォーム２６０１は、また、プロセッサー２６０５によって実行される命令とともに種々の情報を格納するための、バス２６０４に接続した、ランダムアクセスメモリ（ＲＡＭ）あるいは他の動的ストレージ装置のような、揮発性ストレージ装置２６０６を備えている。揮発性ストレージ装置２６０６は、また、プロセッサー２６０５による命令の実行の間に一時変数あるいは他の中間情報を格納するために使用することも出来る。コンピュータプラットフォーム２６０１は、さらに、基本入出力システム（ＢＩＯＳ）や様々なシステム構成パラメーターの様な、プロセッサー２６０５のための静的な情報や命令を格納するために、バス２６０４と接続した読み取り専用メモリ（ＲＯＭまたはＥＰＲＯＭ）２６０７あるいは他のスタティックストレージ装置を備えることがある。磁気ディスク、光ディス

クあるいはソリッドステートフラッシュメモリ装置のような持続形ストレージ装置２６０８は、情報と命令の格納のためにバス２６０４に接続し提供される。

【０１５７】

０１８２　コンピュータプラットフォーム２６０１は、コンピュータプラットフォーム２６０１のシステム管理者またはユーザへの情報の表示のために、陰極線管（ＣＲＴ）、プラズマディスプレイあるいは液晶ディスプレイ（ＬＣＤ）のようなディスプレイ２６０９に、バス２６０４によってつなぐことができる。アルファニューメリックや他のキーを含む入力装置２６１０が、プロセッサー２６０５に情報やコマンドの選択を伝えるためにバス２６０４に接続されている。別のタイプのユーザ入力デバイスは、プロセッサー２６０５に命令情報とコマンドの選択を伝えるため、あるいはディスプレイ２６０９上でカーソル移動をコントロールするための、マウス、トラックボールあるいはカーソル指示キーのようなカーソル制御デバイス２６１１である。この入力装置は、２つの軸、第一軸（例えばｘ）および第二軸（例えばｙ）によって、典型的には２つの自由度を持っているが、これによって入力装置は平面上での位置を指定することが可能となる）。

【０１５８】

０１８３　外部ストレージ装置２６１２は、コンピュータプラットフォーム２６０１に追加のまたはリムーバブルストレージの容量を供給するために、バス２６０４によってコンピュータプラットフォーム２６０１につなぐ事が出来る。計算機システム２６００の実施形態では、外部リムーバブルストレージ装置２６１２は、他の計算機システムとのデータの交換を容易にするために使用することが出来る。

【０１５９】

０１８４　本発明は、ここに記述された技術を実装するのに対し計算機システム２６００を使用することに関係する。ある実施形態では、本発明のシステムはコンピュータプラットフォーム２６０１のようなマシン上で存在することができる。本発明の１つの実施形態によれば、ここに記述された技術は、揮発性メモリ２６０６に含まれている１つ以上の命令の１つ以上のシーケンスを実行するプロセッサー２６０５に応答して計算機システム２６００によって行なわれる。そのような命令は、持続型ストレージ装置２６０８のような別のコンピュータ可読媒体から揮発性メモリ２６０６に読み込む事が出来る。揮発性メモリ２６０６に含まれている一連の命令の実行は、プロセッサー２６０５に、ここに記述されたプロセスのステップを実行させる。他の実施例においては、ハードワイヤードの回路類が、本発明を実装するためにソフトウェア命令の代わりに、あるいはソフトウェア命令と組み合わせて使用する事が出来る。このように、本発明の実施形態はハードウェア回路とソフトウェアのどんな特定の組合せにも制限されてはいない。

【０１６０】

０１８５　ここに使用されている用語“コンピュータ可読媒体”は、実行するプロセッサー２６０５に命令を提供することに関与するあらゆる媒体を表す。コンピュータ可読媒体は機械可読媒体の単に１つの例であり、ここに記述された方法及び／または技術のうちのどれであっても実装するために命令を運ぶことが出来る。そのような媒体は、不揮発性のメディアおよび揮発性のメディアを含み、これらに限定されることなく、多くの形式をとることが出来る。不揮発性のメディアは、例えば、ストレージ装置２６０８のように、光ディスクまたは磁気ディスクを含んでいる。揮発性のメディアは、揮発性ストレージ装置２６０６のように、ダイナミックメモリを含んでいる。

【０１６１】

０１８６　コンピュータ可読媒体の一般の形態は、例えば、フロッピー（登録商標）ディスク、フレキシブルディスク、ハードディスク、磁気テープあるいは何らかの他の磁気メディア、ＣＤ－ＲＯＭ、何らかの他の光学媒体、パンチカード、紙テープ、何らかの他の物理的な孔部のパターンを備えたメディア、ＲＡＭ、ＰＲＯＭ、ＥＰＲＯＭ、フラッシュＥＰＲＯＭ、フラッシュドライブ、メモリカード、何らかの他のメモリチップあるいはカートリッジ、搬送波、など以下に記述されるように、あるいはコンピュータが読むことができる如何なる他のメディアをも含んでいる。

【０１６２】
０１８７　コンピュータ読取り可能なメディアの様々な形式は実行用プロセッサー２６０５への１つ以上の命令の１つ以上のシーケンスを運ぶことに関係する。例えば、命令は、当初はリモートコンピュータから磁気ディスク上で運ばれるかも知れない。あるいは、リモートコンピュータはそのダイナミックメモリに命令をロードし、モデムを使用して、電話回線上で命令を送ることができる。計算機システム２６００に対して直接接続のモデムは、電話回線上のデータを受け取ることができ、データを赤外線信号に変換するために赤外線送信機を使用することができる。赤外線検知器は、赤外線信号で運ばれたデータを受け取ることができ、適切な回路類がデータバス２６０４にデータを乗せることができる。バス２６０４は、データを揮発性ストレージ装置２６０６に運び、プロセッサー２６０５が命令を検索し実行する。揮発性メモリ２６０６によって受け取られた命令は、プロセッサー２６０５によって実行する前に、あるいはその実行の後に持続型記憶装置２６０８上に格納することも選択肢である。命令は、業界では良く知られた様々なネットワークデータ通信プロトコルを使用して、インターネットを通じてコンピュータプラットフォーム２６０１へダウンロードすることも出来る。

【０１６３】
０１８８　コンピュータプラットフォーム２６０１は、また、データバス２６０４につながれたネットワークインターフェースカード２６１３のような通信インターフェースを有する。通信インターフェース２６１３は、ローカルネットワーク２６１５につながっているネットワークリンク２６１４に双方向のデータ通信接続を提供する。例えば、通信インターフェース２６１３は、総合サービスディジタルネットワーク（ＩＳＤＮ）カードか、あるいは対応するタイプの電話回線にデータ通信接続を提供するモデムかもしれない。別の例として、通信インターフェース２６１３は、互換性をもつＬＡＮにデータ通信接続を提供するローカルエリアネットワークインターフェースカード（ＬＡＮ　ＮＩＣ）かもしれない。良く知られている、８０２．１１ａ、８０２．１１ｂ、８０２．１１ｇおよびブルートゥースのような無線リンク、もまたネットワーク実装に使用することが出来る。すべてそのような実装においては、通信インターフェース２６１３は、多様な形式の情報を表すデジタルデータストリームを運ぶ、電気的か、電磁気的か、光学的な信号を送受信する。

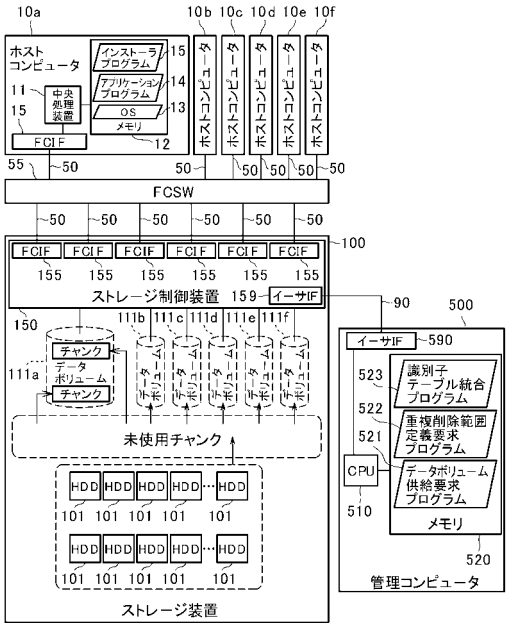
【０１６４】
０１８９　ネットワークリンク２６１３は、典型的には、１つ以上のネットワークを通して他のネットワークリソースにデータ通信を提供する。例えば、ネットワークリンク２６１４は、ローカルネットワーク２６１５を通して、ホストコンピュータ２６１６、あるいはネットワークストレージ／サーバ２６２２への接続を提供することが出来る。さらに、あるいは代わりに、ネットワークリンク２６１３は、ゲートウェイ／ファイアウォール２６１７を通してインターネットのような広域またはグローバルネットワーク２６１８に接続することも出来る。このように、コンピュータプラットフォーム２６０１は、遠隔ネットワークストレージ／サーバ２６１９のように、インターネット２６１８のいかなる場所に位置したネットワークリソースにもアクセス可能である。一方では、コンピュータプラットフォーム２６０１はまた、ローカルエリアネットワーク２６１５及び／またはインターネット２６１８のいかなる場所に位置したクライアントによってもアクセスすることが出来る。ネットワーククライアント２６２０および２６２１は、プラットフォーム２６０１に類似したコンピュータプラットフォームに基づいて、それら自身実装することが出来る。

【０１６５】
０１９０　ローカルネットワーク２６１５およびインターネット２６１８は両方とも、デジタルデータストリームを運ぶ電気的、電磁気的、あるいは光学的信号を使用する。様々なネットワークを通る信号、およびコンピュータプラットフォーム２６０１の間のデジタルデータを運ぶ、通信インターフェース２６１３を通りネットワークリンク２６１４上の信号は、情報を運ぶ搬送波の典型的な形態である。

【０１６６】
０１９１　コンピュータプラットフォーム２６０１は、インターネット２６１８、ＬＡＮ２６１５、ネットワークリンク２６１４および通信インターフェース２６１３を含む多様なネットワーク、を通して、プログラムコードを含み、メッセージを送信しデータを受信することができる。インターネットの例において、システム２６０１がネットワークサーバとして働く場合、クライアント２６２０及び／または２６２１の上で走っているアプリケーションプログラムのために要求されるコードあるいはデータを、インターネット２６１８、ゲートウェイ／ファイアウォール２６１７、ローカルエリアネットワーク２６１５および通信インターフェース２６１３を通して送信することもありうる。同様に、他のネットワークリソースからコードを受け取ることもできる。
【０１６７】
０１９２　受け取られると、受信したコードはプロセッサー２６０５によって実行されるかも知れないし、及び／または、それぞれ持続型または揮発性ストレージ装置である２６０８および２６０６に各々格納され、あるいは後での実行のために他の不揮発性のストレージに格納されるかも知れない。
【０１６８】
０１９３　本発明が、どのような特定のファイアウォールシステムにも制限されていないことは注目すべきである。本発明の方針に基づく内容の処理システムは、３つのファイアウォールオペレーティングモード、のうちのどれででも使用できる、具体的には、ＮＡＴ、ルーテッド、透過型である。
【０１６９】
０１９４　最後に、ここに記述されたプロセスと技術は、どの様な特別の装置とも本質的には関係がなく、構成要素のいかなる適切な組合せによっても実行され得ることは理解すべきである。さらに、多様な型式の汎用目的の装置が、ここに記述された教えに従って使用することも出来る。あるいはまた、ここに記述された方法ステップを実行する専用の装置を構築することが有利であると判明するかも知れない。本発明は、特定の例に関して記述されているが、それは限定をするというよりむしろ全ての関連での例証となることを意図している。当分野での業者は、多くの違ったハードウェア、ソフトウェアおよびファームウェアの組合せが本発明を実施するのに適していることを認識するであろう。例えば、記述されるソフトウェアは、アセンブラー、Ｃ／Ｃ＋＋、ｐｅｒｌ、シェル、ＰＨＰ、Ｊａｖａ（登録商標）などのような、種々様々なプログラミング言語あるいはスクリプト言語で実施されてもよい。
【０１７０】
０１９５　さらに、本発明のその他の実施も、ここに開示された本発明の明細書および実施の考察から、当分野の業者には明白になるであろう。記述された実施形態の種々の態様及び／または構成要素は、データ重複削除機能を備えたコンピュータ化されたストレージシステムにおいて、単独であるいは任意の組合せで使用することが出来る。明細書と実例は、典型的なものとしてのみ考慮されるよう意図されている。

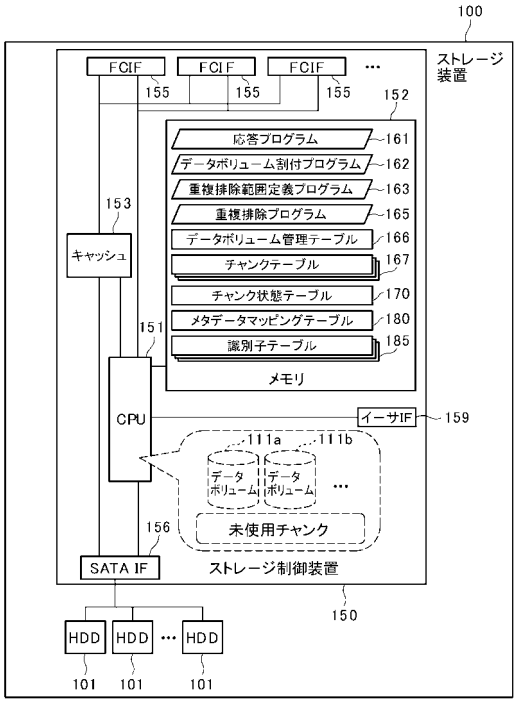
【図１ａ】

図１（ａ）



【図１ｂ】

図１（ｂ）



【図２】

図２

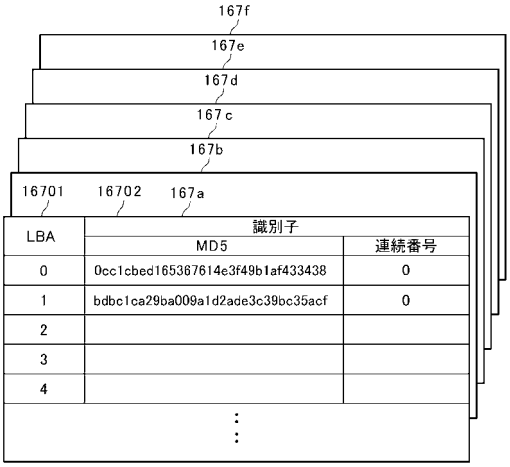
166

16601 ボリューム番号	16602 サイズ	16603 メタデータ	16604 チャンクテーブル番号
111a	2,00 0,000, 000	OSタイプA	167a
111b	2,00 0,000, 000	OSタイプA	167b
111c	2,00 0,000, 000	OSタイプA	167c
111d	2,00 0,000, 000	OSタイプB	167d
111e	2,00 0,000, 000	OSタイプB	167e
111f	2,00 0,000, 000	OSタイプB	167f
⋮			

データボリューム管理テーブル

【図３】

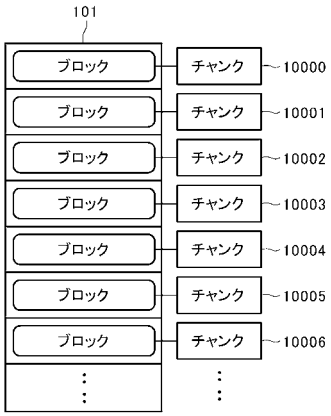
図３



チャンクテーブル

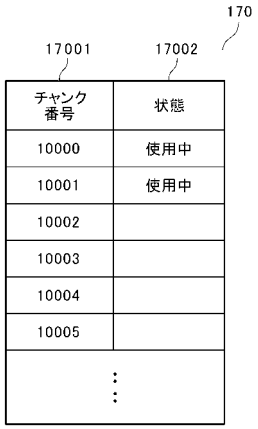
【図４】

図４



【図５】

図５

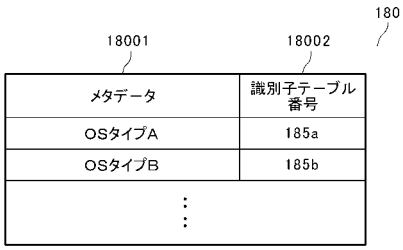


チャンク状態テーブル

【図６】

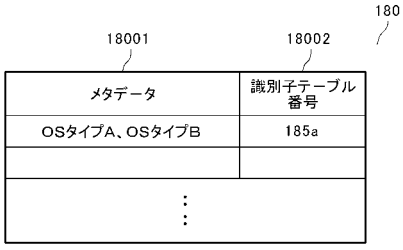
図６

(a)



メタデータマッピングテーブル

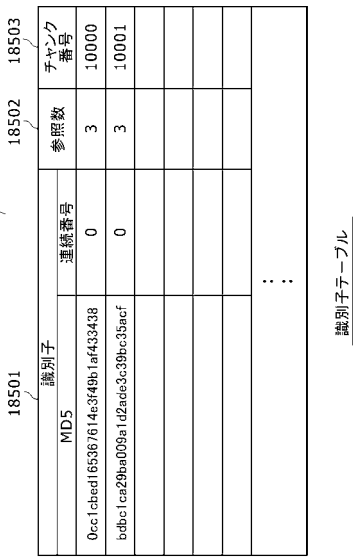
(b)



メタデータマッピングテーブル

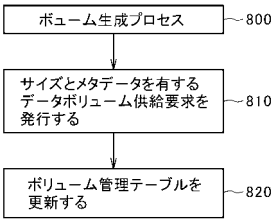
【図７】

図７



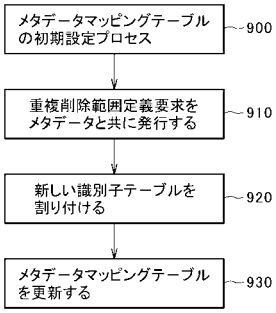
【図８】

図８



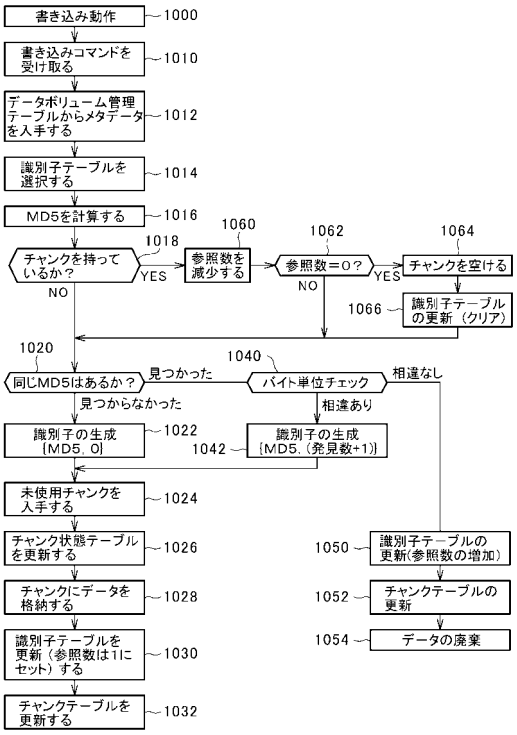
【図９】

図９



【図１０】

図１０



【図１１】

図１１

(a)

タイプ	書き込み
LBA	0
ブロック数	2

書き込みコマンド

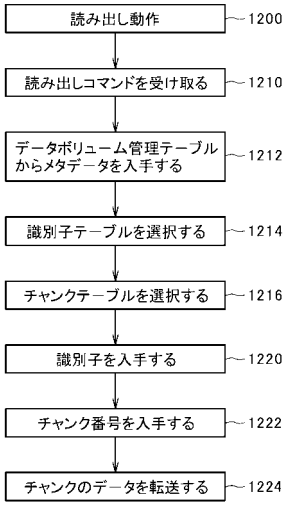
(b)

タイプ	読み出し
LBA	0
ブロック数	2

読み出しコマンド

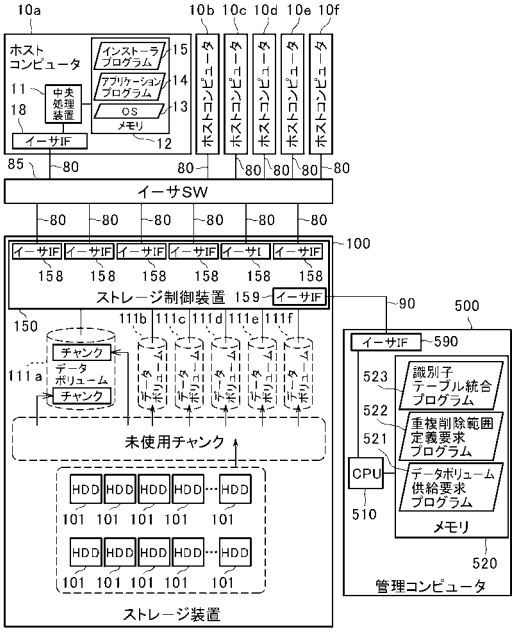
【図１２】

図１２



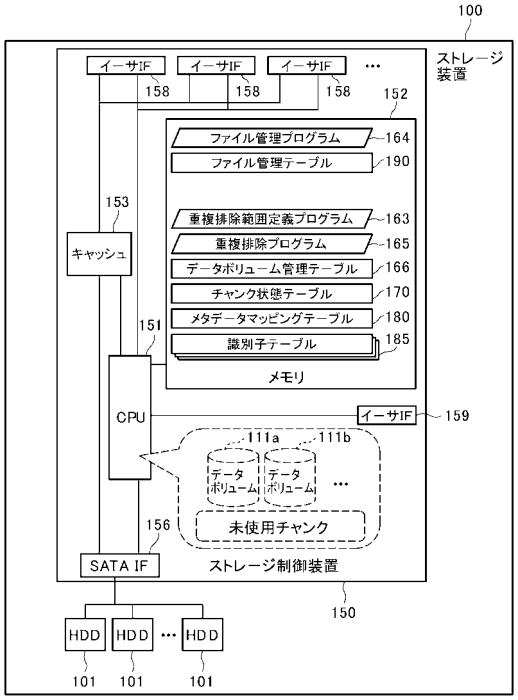
【図１３ａ】

図１３(ａ)



【図１３ｂ】

図１３(ｂ)



【図１４】

図１４

ファイル名	ファイルサイズ	識別子 MD5	識別子 連続番号	メタデータ
/foo/sample.txt	2048	0cc1cbed165367614e3f49b1af433438	0	Tom, アカウンティング, .txt, texteditorX, 10/31/2008
/foo/mail.txt	1024	bdbc1ca29ba009a1d2ade3c39bc35aef	0	Tom, アカウンティング, .txt, mailZ, 10/1/2008
/boo/mail.txt	1024	bdbc1ca29ba009a1d2ade3c39bc35aef	0	Tom, アカウンティング, .txt, mailZ, 10/1/2008
/boo/document.txt	4096	098f6bcd4621d373cade4e832627b4f6	0	John, HR, .txt, texteditorY, 10/2/2008
...				

19001 19002 19003 190 19004

ファイル管理テーブル

【図１５ａ】

図１５(a)

メタデータ	識別子テーブル番号
.txt	185a
.doc	185b
.exe	185c
⋮	

18001 18002 180

メタデータマッピングテーブル

【図１５ｂ】

図１５(b)

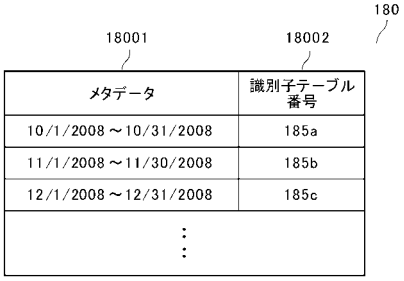
メタデータ	識別子テーブル番号
アカウンティング	185a
HR	185b
R&D	185c
⋮	

18001 18002 180

メタデータマッピングテーブル

【図１５ｃ】

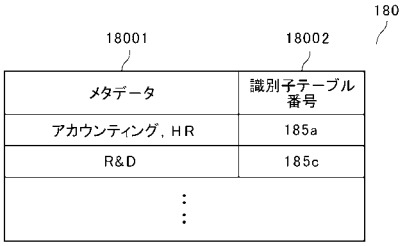
図１５(c)



メタデータマッピングテーブル

【図１５ｄ】

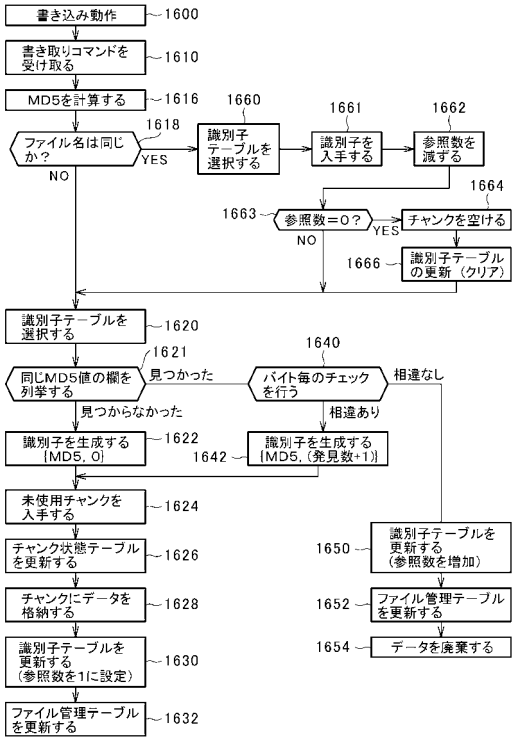
図１５(d)



メタデータマッピングテーブル

【図１６】

図１６



【図１７】

図１７

(a)

タイプ	書き込み
ファイル名	/foo/sample.txt
ファイルサイズ	2048 byte
メタデータ	Tom, アカウンティング, .txt, texteditorX, 10/31/2008

書き込みコマンド

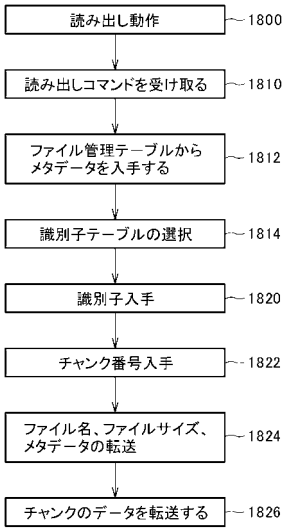
(b)

タイプ	読み出し
ファイル名	/foo/sample.txt

読み出しコマンド

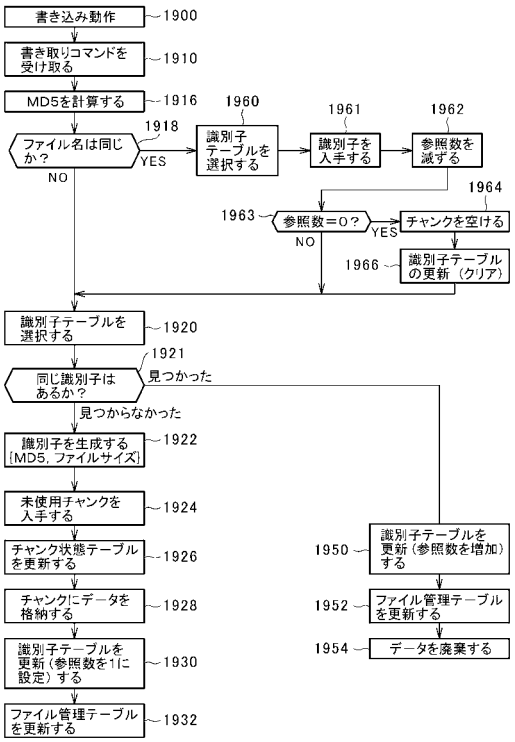
【図１８】

図１８



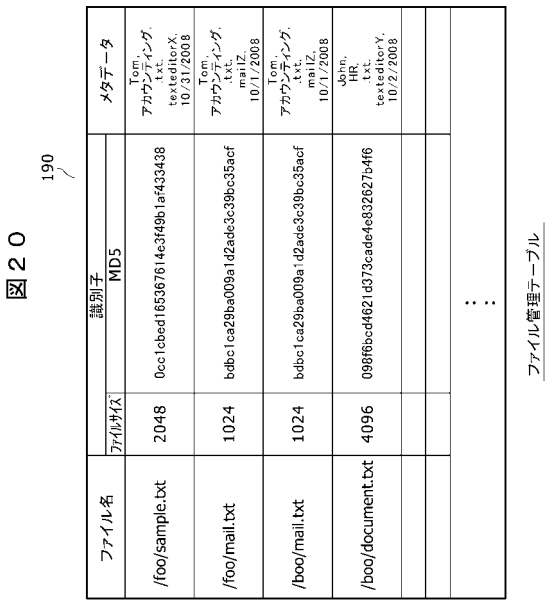
【図１９】

図１９



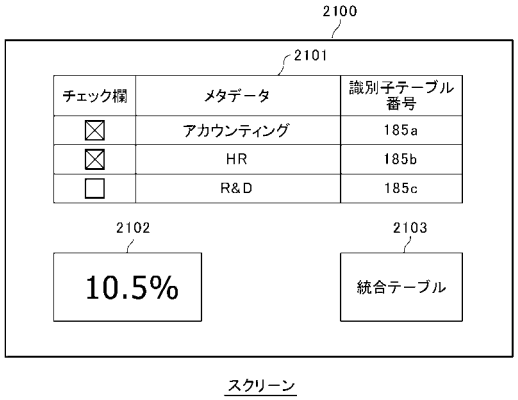
【図２０】

図２０



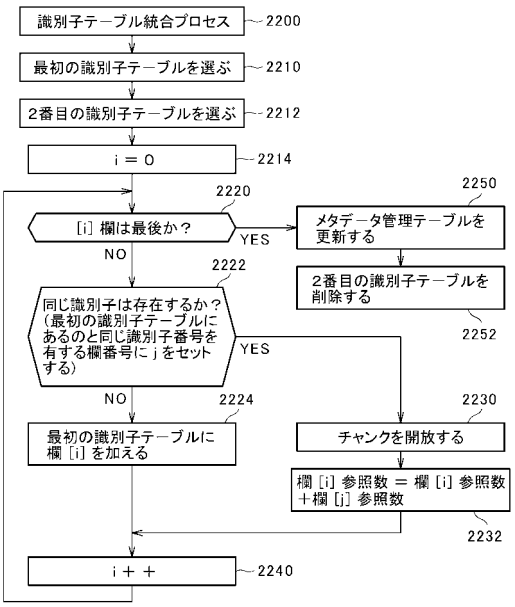
【図２１】

図２１



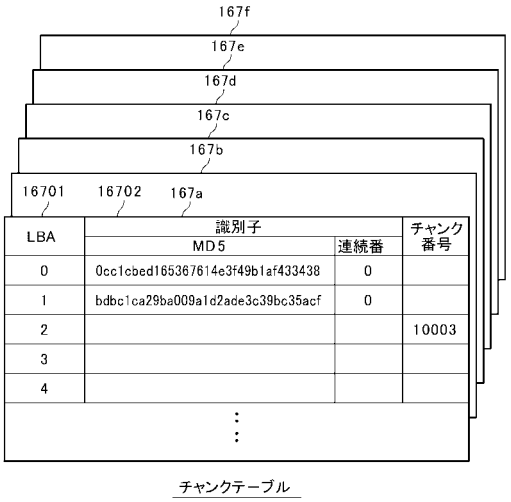
【図２２】

図２２



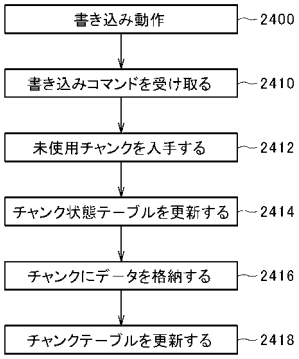
【図２３】

図２３



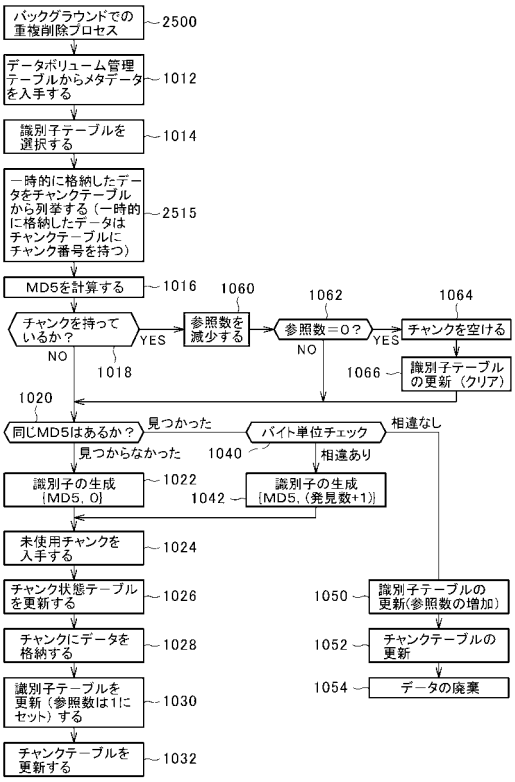
【図２４】

図２４



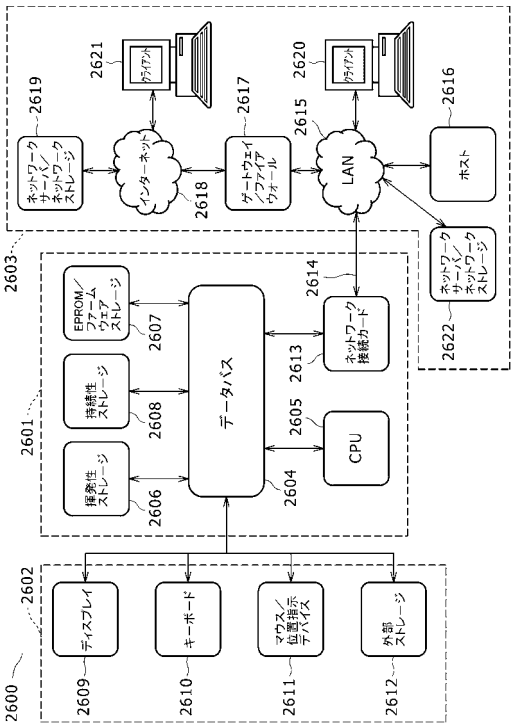
【図２５】

図２５



【図２６】

図２６



フロントページの続き

(56)参考文献　国際公開第２００８／００５２１１（WO，Ａ１）
国際公開第２００９／００６２７８（WO，Ａ１）

(58)調査した分野(Int.Cl.，ＤＢ名)
Ｇ０６Ｆ　３／０６－３／０８
Ｇ０６Ｆ　１２／００