(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号
特許第6094267号
(P6094267)
(45)発行日　平成29年3月15日(2017.3.15)　(24)登録日　平成29年2月24日(2017.2.24)

(51)Int.Cl.　　F I
G06F 12/00 (2006.01)　　G06F 12/00 501B
G06F 3/06 (2006.01)　　G06F 3/06 302J
G06F 3/06 301W

請求項の数 10　(全 26 頁)

(21)出願番号　特願2013-40842(P2013-40842)
(22)出願日　平成25年3月1日(2013.3.1)
(65)公開番号　特開2014-170304(P2014-170304A)
(43)公開日　平成26年9月18日(2014.9.18)
審査請求日　平成28年2月5日(2016.2.5)

(73)特許権者　000004237
日本電気株式会社
東京都港区芝五丁目７番１号
(74)代理人　100124811
弁理士　馬場　資博
(74)代理人　100088959
弁理士　境　廣巳
(72)発明者　山本　拓明
東京都港区芝五丁目７番１号　日本電気株式会社内

審査官　漆原　孝治

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】
記憶対象データを記憶装置に格納すると共に、当該記憶装置に既に記憶されている前記記憶対象データと同一のデータ内容の他の記憶対象データを前記記憶装置に格納する場合に、当該記憶装置に既に記憶されている前記記憶対象データを前記他の記憶対象データとして参照させるデータ格納制御部と、
前記記憶装置の所定の領域内においてデフラグ範囲とされた箇所に格納された前記記憶対象データを、前記記憶装置の他の領域内に格納し直すデフラグ処理部と、を備え、
前記データ格納制御部は、前記記憶装置に記憶されている前記記憶対象データ毎に、当該記憶対象データが他の記憶対象データとして参照されている数である参照数を記憶し、
前記デフラグ処理部は、前記記憶対象データの参照数に応じて当該記憶対象データを前記記憶装置の他の領域内において後にデフラグ範囲となる箇所に格納する、
ストレージシステム。
【請求項２】
請求項１に記載のストレージシステムであって、
前記デフラグ処理部は、前記参照数が予め設定された閾値未満である前記記憶対象データを、前記記憶装置の他の領域内において後にデフラグ範囲となる箇所に格納し、前記参照数が予め設定された閾値以上である前記記憶対象データを、前記記憶装置の他の領域内において後に非デフラグ範囲となる箇所に格納する、
ストレージシステム。

【請求項３】
請求項２に記載のストレージシステムであって、
前記記憶装置の他の領域は、予め設定されたデフラグ範囲と、データが未格納な範囲である未使用範囲と、が隣接して形成されており、
前記デフラグ処理部は、前記参照数が閾値未満である前記記憶対象データを、前記記憶装置の他の領域における前記未使用範囲内であり前記デフラグ範囲との隣接箇所からデータが連続して位置する箇所に格納し、前記未使用範囲内における前記デフラグ範囲との隣接箇所から連続して前記記憶対象データが格納された範囲と前記デフラグ範囲とを連結した範囲を新たなデフラグ範囲とする、
ストレージシステム。
【請求項４】
請求項３に記載のストレージシステムであって、
前記記憶装置の他の領域は、予め設定された非デフラグ範囲と、データが未格納な範囲である未使用範囲と、が隣接して形成されており、
前記デフラグ処理部は、前記参照数が閾値以上である前記記憶対象データを、前記記憶装置の他の領域における前記未使用範囲であり前記非デフラグ範囲との隣接箇所からデータが連続して位置する箇所に格納し、前記未使用範囲内における前記非デフラグ範囲との隣接箇所から連続して前記記憶対象データが格納された範囲と前記非デフラグ範囲とを連結した範囲を新たな非デフラグ範囲とする、
ストレージシステム。
【請求項５】
請求項２に記載のストレージシステムであって、
前記記憶装置の他の領域は、予め設定されたデフラグ範囲にデータが未格納な範囲である未使用範囲の一端側が隣接すると共に、当該未使用範囲の他端側が予め設定された非デフラグ範囲に隣接し、当該未使用範囲がデフラグ範囲と非デフラグ範囲とに挟まれて形成されており、
前記デフラグ処理部は、前記参照数が閾値未満である前記記憶対象データを、前記記憶装置の他の領域における前記未使用範囲内であり前記デフラグ範囲との隣接箇所からデータが連続して位置する箇所に格納し、前記未使用範囲内における前記デフラグ範囲との隣接箇所から連続して前記記憶対象データが格納された範囲と前記デフラグ範囲とを連結した範囲を新たなデフラグ範囲とし、前記参照数が閾値以上である前記記憶対象データを、前記記憶装置の他の領域における前記未使用範囲であり前記非デフラグ範囲との隣接箇所からデータが連続して位置する箇所に格納し、前記未使用範囲内における前記非デフラグ範囲との隣接箇所から連続して前記記憶対象データが格納された範囲と前記非デフラグ範囲とを連結した範囲を新たな非デフラグ範囲とする、
ストレージシステム。
【請求項６】
請求項１乃至５のいずれかに記載のストレージシステムであって、
前記記憶装置の所定の領域は、予め設定されたデフラグ範囲と、データが未格納な範囲である未使用範囲と、が隣接して形成されており、
前記データ格納制御部は、前記記憶対象データを新たに前記記憶装置に格納する際に、前記記憶装置の所定の領域における前記未使用範囲内であり前記デフラグ範囲との隣接箇所からデータが連続して位置する箇所に格納し、前記未使用範囲内における前記デフラグ範囲との隣接箇所から連続して前記記憶対象データが格納された範囲と前記デフラグ範囲とを連結した範囲を新たなデフラグ範囲とする、
ストレージシステム。
【請求項７】
請求項４に記載のストレージシステムであって、
前記記憶装置の他の領域は、前記デフラグ範囲と前記未使用範囲とが隣接して形成された第一領域と、当該第一領域とは異なり前記非デフラグ範囲と前記未使用範囲とが隣接し

て形成された第二領域と、を有する、
ストレージシステム。
【請求項８】
請求項７に記載のストレージシステムであって、
前記第一領域は、所定の記憶装置に形成されており、
前記第二領域は、前記所定の記憶装置よりもデータ読み出し処理が高速な他の記憶装置に形成されている、
ストレージシステム。
【請求項９】
情報処理装置に、
記憶対象データを記憶装置に格納すると共に、当該記憶装置に既に記憶されている前記記憶対象データと同一のデータ内容の他の記憶対象データを前記記憶装置に格納する場合に、当該記憶装置に既に記憶されている前記記憶対象データを前記他の記憶対象データとして参照させるデータ格納制御部と、
前記記憶装置の所定の領域内においてデフラグ範囲とされた箇所に格納された前記記憶対象データを、前記記憶装置の他の領域内に格納し直すデフラグ処理部と、を実現させると共に、
前記データ格納制御部は、前記記憶装置に記憶されている前記記憶対象データ毎に、当該記憶対象データが他の記憶対象データとして参照されている数である参照数を記憶し、
前記デフラグ処理部は、前記記憶対象データの参照数に応じて当該記憶対象データを前記記憶装置の他の領域内において後にデフラグ範囲となる箇所に格納する、
ことを実現させるためのプログラム。
【請求項１０】
記憶対象データを記憶装置に格納すると共に、当該記憶装置に既に記憶されている前記記憶対象データと同一のデータ内容の他の記憶対象データを前記記憶装置に格納する場合に、当該記憶装置に既に記憶されている前記記憶対象データを前記他の記憶対象データとして参照させてデータ格納制御を行うと共に、前記記憶装置に記憶されている前記記憶対象データ毎に、当該記憶対象データが他の記憶対象データとして参照されている数である参照数を記憶し、
前記記憶装置の所定の領域内においてデフラグ範囲とされた箇所に格納された前記記憶対象データを、前記記憶装置の他の領域内に格納し直すデフラグ処理を実行し、
前記デフラグ処理時に、前記記憶対象データの参照数に応じて当該記憶対象データを前記記憶装置の他の領域内において後にデフラグ範囲となる箇所に格納する、
デフラグ方法。

【発明の詳細な説明】
【技術分野】
【０００１】
本発明は、ストレージシステムにかかり、特に、同一内容のデータの重複記憶を排除するストレージシステムに関する。
【背景技術】
【０００２】
近年、コンピュータの発達及び普及に伴い、種々の情報がデジタルデータ化されている。このようなデジタルデータを保存しておく装置として、磁気テープや磁気ディスクなどの記憶装置がある。そして、保存すべきデータは日々増大し、膨大な量となるため、大容量なストレージシステムが必要となっている。また、記憶装置に費やすコストを削減しつつ、信頼性も必要とされる。これに加えて、後にデータを容易に取り出すことが可能であることも必要である。その結果、自動的に記憶容量や性能の増大を実現できると共に、重複記憶を排除して記憶コストを削減し、さらには、冗長性の高いストレージシステムが望まれている。

【０００３】
このような状況に応じて、近年では、特許文献１に示すように、コンテンツアドレスストレージシステムが開発されている。このコンテンツアドレスストレージシステムは、データを分散して複数の記憶装置に記憶すると共に、このデータの内容に応じて特定される固有のコンテンツアドレスによって、当該データを格納した格納位置が特定される。また、コンテンツアドレスストレージシステムの中には、所定のデータを複数のフラグメントに分割すると共に、冗長データとなるフラグメントをさらに付加して、これら複数のフラグメントをそれぞれ複数の記憶装置にそれぞれ格納する、というものもある。
【０００４】
そして、上述したようなコンテンツアドレスストレージシステムでは、後に、コンテンツアドレスを指定することにより、当該コンテンツアドレスにて特定される格納位置に格納されているデータつまりフラグメントを読み出し、複数のフラグメントから分割前の所定のデータを復元することができる。
【０００５】
また、上記コンテンツアドレスは、データの内容に応じて固有となるよう生成される値、例えばデータのハッシュ値、に基づいて生成される。このため、重複データであれば同じ格納位置のデータを参照することで、同一内容のデータを取得することができる。従って、重複データを別々に格納する必要がなく、重複記録を排除して、データ容量の削減を図ることができる。
【０００６】
特に、上述したような重複記憶を排除する機能を有するストレージシステムでは、ファイルなど書き込み対象となるデータを所定容量の複数のブロックデータに分割して圧縮し、記憶装置に書き込む。このように、ファイルを分割したブロックデータ単位で重複記憶を排除することで、重複率が増大し、データ容量の削減を図っている。そして、バックアップを行うストレージシステムに提供することで、バックアップの容量を節約したり、レプリケーション時の帯域の節約を図っている。
【０００７】
一方で、上述した重複排除を行うストレージシステムでは、格納したデータが上書きされることはない。つまり、新しい非重複データを追記するか、古く参照されていないデータを削除するか、といったどちらかしか行われない。このため、重複排除ストレージシステムを長期間運用すると、ストレージ内部で空き領域が断片化し、入出力性能（Ｉ／Ｏ性能）が低下する。そのため、重複排除ストレージシステムでは、定期的なデフラグ処理が必要になる。
【先行技術文献】
【特許文献】
【０００８】
【特許文献１】特開２００５－２３５１７１号公報
【特許文献２】特開２０１０－２８７０４９号公報
【発明の概要】
【発明が解決しようとする課題】
【０００９】
しかしながら、デフラグ処理は、ユーザによる運用とは関係のないストレージシステムにおけるＩ／Ｏ負荷を発生させるため、デフラグ期間中のＩ／Ｏ性能を低下させてしまう、という問題が生じる。ここで、特許文献２には、メモリ内におけるページ分割数が閾値を超えるブロックに対してデフラグ処理を行う、という技術が開示されている。かかる技術によると、デフラグ処理を行う領域が制限され、Ｉ／Ｏ性能の低下を抑制できる可能性があるが、データの内容を考慮してデフラグを行っていない。従って、重複排除ストレージシステムに格納される重複排除を伴うデータに対して、デフラグを効率よく行うことができるか不明であり、断片化を解消するためには結局のところ全ての領域に対してデフラグを行う必要がある。その結果、重複排除ストレージシステムにおけるＩ／Ｏ性能の低下

を解消することができない。

【００１０】

このため、本発明の目的は、重複排除ストレージシステムにおけるデフラグ処理によるＩ／Ｏ性能の低下、という問題を解決することができるストレージシステムを提供することにある。

【課題を解決するための手段】

【００１１】

本発明の一形態であるストレージシステムは、

記憶対象データを記憶装置に格納すると共に、当該記憶装置に既に記憶されている前記記憶対象データと同一のデータ内容の他の記憶対象データを前記記憶装置に格納する場合に、当該記憶装置に既に記憶されている前記記憶対象データを前記他の記憶対象データとして参照させるデータ格納制御部と、

前記記憶装置の所定の領域内においてデフラグ範囲とされた箇所に格納された前記記憶対象データを、前記記憶装置の他の領域内に格納し直すデフラグ処理部と、を備え、

前記データ格納制御部は、前記記憶装置に記憶されている前記記憶対象データ毎に、当該記憶対象データが他の記憶対象データとして参照されている数である参照数を記憶し、

前記デフラグ処理部は、前記記憶対象データの参照数に応じて当該記憶対象データを前記記憶装置の他の領域内において後にデフラグ範囲となる箇所に格納する、

という構成をとる。

【００１２】

本発明の他の形態であるプログラムは、

情報処理装置に、

記憶対象データを記憶装置に格納すると共に、当該記憶装置に既に記憶されている前記記憶対象データと同一のデータ内容の他の記憶対象データを前記記憶装置に格納する場合に、当該記憶装置に既に記憶されている前記記憶対象データを前記他の記憶対象データとして参照させるデータ格納制御部と、

前記記憶装置の所定の領域内においてデフラグ範囲とされた箇所に格納された前記記憶対象データを、前記記憶装置の他の領域内に格納し直すデフラグ処理部と、を実現させると共に、

前記データ格納制御部は、前記記憶装置に記憶されている前記記憶対象データ毎に、当該記憶対象データが他の記憶対象データとして参照されている数である参照数を記憶し、

前記デフラグ処理部は、前記記憶対象データの参照数に応じて当該記憶対象データを前記記憶装置の他の領域内において後にデフラグ範囲となる箇所に格納する、

ことを実現させるためのプログラムである。

【００１３】

本発明の他の形態であるデフラグ方法は、

記憶対象データを記憶装置に格納すると共に、当該記憶装置に既に記憶されている前記記憶対象データと同一のデータ内容の他の記憶対象データを前記記憶装置に格納する場合に、当該記憶装置に既に記憶されている前記記憶対象データを前記他の記憶対象データとして参照させてデータ格納制御を行うと共に、前記記憶装置に記憶されている前記記憶対象データ毎に、当該記憶対象データが他の記憶対象データとして参照されている数である参照数を記憶し、

前記記憶装置の所定の領域内においてデフラグ範囲とされた箇所に格納された前記記憶対象データを、前記記憶装置の他の領域内に格納し直すデフラグ処理を実行し、

前記デフラグ処理時に、前記記憶対象データの参照数に応じて当該記憶対象データを前記記憶装置の他の領域内において後にデフラグ範囲となる箇所に格納する、

という構成をとる。

【発明の効果】

【００１４】

本発明は、以上のように構成されることにより、ストレージシステムにおけるデフラグ

によるＩ／Ｏ性能の低下を抑制することができる。
【図面の簡単な説明】
【００１５】
【図１】本発明の実施形態１におけるストレージシステムを含むシステム全体の構成を示すブロック図である。
【図２】本発明の実施形態１におけるストレージシステムの構成の概略を示すブロック図である。
【図３】本発明の実施形態１におけるストレージシステムの構成を示す機能ブロック図である。
【図４】図３に開示したストレージシステムにおけるデータ書き込み処理の様子を説明するための説明図である。
【図５】図３に開示したストレージシステムにおけるデータ書き込み処理の様子を説明する説明図である。
【図６】図３に開示したストレージシステムにおけるデータ構造を説明する説明図である。
【図７】図３に開示したストレージシステムにおけるデータ書き込み処理の様子を説明する説明図である。
【図８】図３に開示したストレージシステムにおけるデータ書き込み処理の様子を説明する説明図である。
【図９】図３に開示したストレージシステムにおけるデータ書き込み処理の様子を説明する説明図である。
【図１０】図３に開示したストレージシステムにおけるデフラグ処理の様子を説明する説明図である。
【図１１】図３に開示したストレージシステムにおけるデフラグ処理の様子を説明する説明図である。
【図１２】図３に開示したストレージシステムにおけるデフラグ処理の様子を説明する説明図である。
【図１３】図３に開示したストレージシステムにおけるデフラグ処理の様子を説明する説明図である。
【図１４】図３に開示したストレージシステムにおけるデータ書き込み処理の動作を示すフローチャートである。
【図１５】図３に開示したストレージシステムにおけるデフラグ処理の動作を示すフローチャートである。
【図１６】本発明の実施形態２におけるストレージシステムに記憶されたデータの様子を示す図である。
【図１７】実施形態２のストレージシステムにおけるデフラグ処理の様子を説明する説明図である。
【図１８】実施形態２のストレージシステムにおけるデフラグ処理の様子を説明する説明図である。
【図１９】実施形態２のストレージシステムにおけるデフラグ処理の様子を説明する説明図である。
【図２０】実施形態２のストレージシステムにおけるデフラグ処理の動作を示すフローチャートである。
【図２１】実施形態２のストレージシステムの他の構成を示すブロック図である。
【図２２】本発明の付記１におけるストレージシステムの構成を示すブロック図である。
【発明を実施するための形態】
【００１６】
＜実施形態１＞
本発明の第１の実施形態を、図１乃至図１５を参照して説明する。図１は、システム全体の構成を示すブロック図である。図２は、ストレージシステムの概略を示すブロック図

であり、図３は、ストレージシステムの構成を示す機能ブロック図である。図４乃至図１３は、ストレージシステムにおけるデータ書き込み処理及びデフラグ処理の動作を説明するための説明図である。図１４乃至図１５は、ストレージシステムの動作を示すフローチャートである。

【００１７】

ここで、本実施形態は、後述する付記に記載のストレージシステム等の具体的な一例を示すものである。そして、以下では、ストレージシステムが、複数台のサーバコンピュータが接続されて構成されている場合を説明する。但し、本発明におけるストレージシステムは、複数台のコンピュータにて構成されることに限定されず、１台のコンピュータで構成されていてもよい。

【００１８】

［構成］

図１に示すように、本発明におけるストレージシステム１は、ネットワークＮを介してバックアップ処理を制御するバックアップシステム４に接続している。そして、バックアップシステム４は、ネットワークＮを介して接続されたバックアップ対象装置５に格納されているバックアップ対象データ（記憶対象データ）を取得し、ストレージシステム１に対して記憶するよう要求する。これにより、ストレージシステム１は、記憶要求されたバックアップ対象データをバックアップ用に記憶する。

【００１９】

そして、図２に示すように、本実施形態におけるストレージシステム１は、複数のサーバコンピュータが接続された構成を採っている。具体的に、ストレージシステム１は、ストレージシステム１自体における記憶再生動作を制御するサーバコンピュータであるアクセラレータノード２と、データを格納する記憶装置を備えたサーバコンピュータであるストレージノード３と、を備えている。なお、アクセラレータノード２の数とストレージノード３の数は、図２に示したものに限定されず、さらに多くの各ノード２，３が接続されて構成されていてもよい。

【００２０】

さらに、本実施形態におけるストレージシステム１は、データを分割及び冗長化し、分散して複数の記憶装置に記憶すると共に、記憶するデータの内容に応じて設定される固有のコンテンツアドレスによって、当該データを格納した格納位置を特定するコンテンツアドレスストレージシステムである。このコンテンツアドレスストレージシステムについては、後に詳述する。

【００２１】

なお、以下では、ストレージシステム１が１つのシステムであるとして、当該ストレージシステム１が備えている構成及び機能を説明する。つまり、以下に説明するストレージシステム１が有する構成及び機能は、アクセラレータノード２あるいはストレージノード３のいずれに備えられていてもよい。なお、ストレージシステム１は、図２に示すように、必ずしもアクセラレータノード２とストレージノード３とを備えていることに限定されず、いかなる構成であってもよく、例えば、１台のコンピュータにて構成されていてもよい。さらには、ストレージシステム１は、コンテンツアドレスストレージシステムであることにも限定されず、重複排除機能を有しているストレージシステムであればよい。

【００２２】

図３に、本実施形態におけるストレージシステム１の構成を示す。この図に示すように、ストレージシステム１は、サーバコンピュータにて構成され、所定の演算処理を行う演算装置（図示せず）と、記憶装置２０と、を備える。そして、ストレージシステム１は、上記演算装置にプログラムが組み込まれることで構築された、書き込み部１１（データ格納制御部）、読み出し部１２、削除部１３、デフラグ部１４（デフラグ処理部）、を備えている。

【００２３】

なお、実際には、上述したストレージシステム１が備える構成は、図２に示したアクセ

ラレータノード２及びストレージノード３がそれぞれ備えているＣＰＵ（Central Processing Unit）などの演算装置やハードディスクドライブなどの記憶装置にて構成されている。

【００２４】

ここで、上述したように、本実施形態におけるストレージシステム１は、コンテンツアドレスストレージシステムである。このため、ストレージシステム１は、コンテンツアドレスを利用してデータを記憶装置２０に格納する機能を有しており、以下に説明するように、データを分割及び分散し、かつ、コンテンツアドレスにて格納位置を特定して、データを格納する。以下、ストレージシステム１にてコンテンツアドレスを利用した書き込み部１１によるデータ書き込み処理の一例、及び、読み出し部１２による読み出し処理の一例について、図４及び図５を参照して説明する。但し、以下に説明する処理は、重複排除型のストレージシステムにおけるデータの書き込み処理及び読み出し処理の一例である。従って、本発明におけるストレージシステム１は、以下の方法でデータを書き込んだり読み出すことに限定されず、他の方法で重複記憶を排除してデータの書き込み処理及び読み出し処理を行ってもよい。

【００２５】

まず、図４及び図５の矢印Ｙ１に示すように、ストレージシステム１が書き込み要求されたファイルＡの入力を受ける。すると、図４及び図５の矢印Ｙ２に示すように、ファイルＡを所定容量（例えば、６４ＫＢ）のブロックデータＤに分割する。

【００２６】

続いて、分割されたブロックデータＤのデータ内容に基づいて、当該データ内容を代表する固有のハッシュ値Ｈを算出する（図５の矢印Ｙ３）。例えば、ハッシュ値Ｈは、予め設定されたハッシュ関数を用いて、ブロックデータＤのデータ内容から算出する。

【００２７】

続いて、ファイルＡのブロックデータＤのハッシュ値Ｈを用いて、当該ブロックデータＤが既に格納されているか否かを調べる。具体的には、まず、既に格納されているブロックデータＤは、そのハッシュ値Ｈと格納位置を表すコンテンツアドレスＣＡとが、関連付けられてＭＦＩ（Ｍａｉｎ　Ｆｒａｇｍｅｎｔ　Ｉｎｄｅｘ）ファイルに登録されている。従って、格納前に算出したブロックデータＤのハッシュ値ＨがＭＦＩファイル内に存在している場合には、既に同一内容のブロックデータＤが格納されていると判断できる（図５の矢印Ｙ４）。この場合には、格納前のブロックデータＤのハッシュ値Ｈと一致したＭＦＩ内のハッシュ値Ｈに関連付けられているコンテンツアドレスＣＡを、当該ＭＦＩファイルから取得する。そして、このコンテンツアドレスＣＡを、書き込み要求されたブロックデータＤのコンテンツアドレスＣＡとして返却する。

【００２８】

そして、返却されたコンテンツアドレスＣＡが参照する既に格納されているデータを、書き込み要求されたブロックデータＤとして使用する。つまり、書き込み要求されたブロックデータＤの格納先として、返却されたコンテンツアドレスＣＡが参照する領域を指定することで、当該書き込み要求されたブロックデータＤを記憶したこととする。これにより、書き込み要求にかかるブロックデータＤを、実際に記憶装置２０内に記憶する必要がなくなる。

【００２９】

また、書き込み要求にかかるブロックデータＤがまだ記憶されていないと判断された場合には、以下のようにして書き込み要求にかかるブロックデータＤの書き込みを行う。まず、書き込み要求にかかるブロックデータＤを圧縮して、図５の矢印Ｙ５に示すように、複数の所定の容量のフラグメントデータに分割する。例えば、図４の符号Ｄ１～Ｄ９に示すように、９つのフラグメントデータ（分割データ９１）に分割する。そしてさらに、分割したフラグメントデータのうちいくつかが欠けた場合であっても、元となるブロックデータを復元可能なよう冗長データを生成し、上記分割したフラグメントデータ９１に追加する。例えば、図４の符号Ｄ１０～Ｄ１２に示すように、３つのフラグメントデータ（冗

長データ９２）を追加する。これにより、９つの分割データ９１と、３つの冗長データ９２とにより構成される１２個のフラグメントデータからなるデータセット９０を生成する。

【００３０】

続いて、上述したように生成されたデータセットを構成する各フラグメントデータを、記憶装置に形成された各記憶領域に、それぞれ分散して格納する。例えば、図４に示すように、１２個のフラグメントデータＤ１～Ｄ１２を生成した場合には、複数の記憶装置内にそれぞれ形成したデータ格納ファイルに、各フラグメントデータＤ１～Ｄ１２を１つずつそれぞれ格納する（図５の矢印Ｙ６参照）。

【００３１】

続いて、ストレージシステム１は、上述したように格納したフラグメントデータＤ１～Ｄ１２の格納位置、つまり、当該フラグメントデータＤ１～Ｄ１２にて復元されるブロックデータＤの格納位置を表すコンテンツアドレスＣＡを生成して管理する。具体的には、格納したブロックデータＤの内容に基づいて算出したハッシュ値Ｈの一部（ショートハッシュ）（例えば、ハッシュ値Ｈの先頭８Ｂ（バイト））と、論理格納位置を表す情報と、を組み合わせて、コンテンツアドレスＣＡを生成する。そして、このコンテンツアドレスＣＡを、ストレージシステム１内のファイルシステムに返却する（図５の矢印Ｙ７）。すると、ストレージシステム１は、バックアップ対象データのファイル名などの識別情報と、コンテンツアドレスＣＡとを関連付けてファイルシステムで管理する。

【００３２】

また、ブロックデータＤのコンテンツアドレスＣＡと、当該ブロックデータＤのハッシュ値Ｈと、を関連付けて、各ストレージノード３がＭＦＩファイルにて管理する。このように、上記コンテンツアドレスＣＡは、ファイルを特定する情報やハッシュ値Ｈなどと関連付けられて、アクセラレータノード２やストレージノード３の記憶装置に格納される。

【００３３】

さらに、ストレージシステム１は、上述したように格納したファイルを読み出す制御を行う。例えば、ストレージシステム１に対して、特定のファイルを指定して読み出し要求があると、まず、ファイルシステムに基づいて、読み出し要求にかかるファイルに対応するハッシュ値の一部であるショートハッシュと論理位置の情報からなるコンテンツアドレスＣＡを指定する。そして、コンテンツアドレスＣＡがＭＦＩファイルに登録されているか否かを調べる。登録されていなければ、要求されたデータは格納されていないため、エラーを返却する。

【００３４】

一方、読み出し要求にかかるコンテンツアドレスＣＡが登録されている場合には、上記コンテンツアドレスＣＡにて指定される格納位置を特定し、この特定された格納位置に格納されている各フラグメントデータを、読み出し要求されたデータとして読み出す。このとき、各フラグメントが格納されているデータ格納ファイルと、当該データ格納ファイルのうち１つのフラグメントデータの格納位置が分かれば、同一の格納位置から他のフラグメントデータの格納位置を特定することができる。

【００３５】

そして、読み出し要求に応じて読み出した各フラグメントデータからブロックデータＤを復元する。さらに、復元したブロックデータＤを複数連結し、ファイルＡなどの一群のデータに復元して返却する。

【００３６】

次に、上述したように記憶装置２０内に書き込まれたデータに対するデフラグ処理を行うための構成を、さらに説明する。図３に示すように、ストレージシステム１は、領域管理表１５、ブロック管理表１６、ファイル管理表１７を備える。これらの表１５，１６，１７は、図示しない主記憶装置に形成されている。また、補助記憶装置にて構成されている記憶装置２０の記憶領域は、複数の領域２１に分割されて管理されている。

【００３７】

図６は、領域管理表１５、ブロック管理表１６、ファイル管理表１７の詳細、及び、これら各表と記憶装置２０及び各領域２１との関係を示している。

【００３８】

領域管理表１５の「領域番号」の列は、各領域２１を識別するための一意な番号を表す。そして、記憶装置２０内の領域２１は、書き込み時において同時に１つだけ書き込み可能となり、その書き込み可能な領域２１のみを「使用中フラグ」を「yes」とすることで特定する。「領域サイズ」は、各領域２１に格納可能なデータの合計サイズを表す。「デフラグ範囲」は、後述するようにデフラグ部１４がデフラグを行う際の対象とすべき範囲を表す。「未使用範囲」は、一切のデータが格納されていない範囲であり、後述するように書き込み部１１やデフラグ部１４によってデータが格納される範囲を表す。「長寿命範囲」は、後述するようにデフラグ部１４によって重複率が高いブロックが格納される範囲を表す。「最終書込時刻」は、領域２１に最後にデータが書き込まれたときの時刻である。

【００３９】

ブロック管理表１６の「ハッシュ値」の列は、ファイルを分割した各ブロックデータのハッシュ値を表している。このハッシュ値は、ブロックデータのデータ内容を代表する値であり、書き込み部１１は、このハッシュ値を用いて、上述したように書き込み対象であるブロックデータが既に記憶装置２０に格納されているか否かを調べる。「重複率」の列は、各ブロックデータの重複率を表している。重複率は、ブロックデータが他のブロックデータとして参照されている数（参照数）の値である。「領域番号」、「物理アドレス」及び「ブロックサイズ」は、各ブロックデータが記憶装置２０において格納されている物理的な位置を指し示している。

【００４０】

ファイル管理表１７は、ファイルを構成するブロックデータのリストを、ハッシュ値のリストとして保持している。

【００４１】

そして、本実施形態における書き込み部１１は、ファイルを記憶装置２０の領域２１に格納する際には、当該ファイルを分割したブロックデータのハッシュ値を、上記ファイル管理表１７及びブロック管理表１６に登録する。このとき、ブロックデータのハッシュ値から、当該ブロックデータが既に記憶装置２０内の領域２１に記憶されているか否かを判断する。そして、記憶対象となるブロックデータが既に記憶装置２０内に存在している場合には、当該存在しているブロックデータを参照し、ブロック管理表１６の該当ハッシュ値の行における重複率（参照する）の値をインクリメントする。

【００４２】

一方、記憶対象となるブロックデータのハッシュ値が未知のものである場合は、領域管理表１５を参照し、「使用中フラグ」が「yes」となっている領域２１を見つけ出し、該当領域２１の未使用範囲の先頭に該当ブロックデータを書き込む。そして、ブロックデータのサイズ分だけ未使用範囲の先頭を縮小し、デフラグ範囲の末尾を拡大し、領域管理表１５の該当領域２１の最終書き込み時刻を更新する。加えて、ブロック管理表１６に新規エントリを追加し、書き込みの際に使用した領域番号、物理アドレス、およびブロックサイズを格納する。最後に、ファイル管理表１７の該当ファイル名のリストに書き込んだブロックのハッシュ値を追記する。

【００４３】

また、読み出し部１２は、ファイル管理表１７を参照し、読み出し対象となるファイルを構成するブロックデータのハッシュ値を列挙する。そして、ブロック管理表１６を参照し、該当ブロックが格納されている領域、物理アドレスおよびブロックサイズを割り出し、該当する格納場所からブロックデータの読み出しを行う。

【００４４】

また、削除部１３は、削除対象のファイルを構成するブロックデータに該当するブロック管理表１６のエントリにおける「重複率」をデクリメントする。重複率が０になった場

合、該当するエントリをブロック管理表１６から削除する。ただし、重複率は、削除時に即座にデクリメントされる必要はなく、ファイルつまりブロックデータ削除後の適切なタイミングでデクリメントすることも可能である。

【００４５】

次に、本実施形態におけるデフラグ部１４について説明する。デフラグ部１４は、所定のタイミングで、記憶装置２０の所定の領域２１であるデフラグ元領域内に格納されたブロックデータを、他の領域２１であるデフラグ先領域内に格納し直すデフラグ処理を実行する。このとき、デフラグ部１４は、デフラグ元領域のデフラグ範囲にあるブロックデータをデフラグ対象とし、当該ブロックデータの重複率に応じて、デフラグ先領域に格納する箇所を決定する。

【００４６】

例えば、予め設定された一定値（閾値）未満の重複率であるブロックデータは、デフラグ先領域の未使用範囲のうち、デフラグ範囲に隣接する箇所から格納データが連続するよう格納する。そして、ブロックデータ格納後やデフラグ終了時には、デフラグ先領域のデフラグ範囲から未使用範囲に格納したブロックデータまでの連続する範囲を新たなデフラグ範囲とし、その分、未使用範囲を縮小する。

【００４７】

一方、予め設定された一定値（閾値）以上の高い重複率であるブロックデータは、デフラグ先領域の未使用範囲のうち、長寿命範囲（非デフラグ範囲）に隣接する箇所から格納データが連続するよう格納する。そして、ブロックデータ格納後やデフラグ終了時には、デフラグ先領域の長寿命範囲から未使用範囲に格納したブロックデータまでの連続する範囲を新たな長寿命範囲とし、その分、未使用範囲を縮小する。

【００４８】

［動作］

次に、上述した構成のストレージシステムの動作を、図７乃至図１５を参照して説明する。まず、ストレージシステムにファイルを書き込むときの動作を、図７乃至図９と、図１４のフローチャートを参照して説明する。

【００４９】

まず、書き込み部１１は、書き込み対象のファイル（データストリーム）をフィンガープリントに基づいて所定容量のブロックデータに分割する（ステップＳ１）。次に、分割された各ブロックデータのハッシュ値を計算する（ステップＳ２）。

【００５０】

その後、書き込み部１１は、ブロック管理表１６を確認し、該当するハッシュ値を持つブロックデータが記憶装置２０内に存在しているか否かを確認する（ステップＳ３）。ブロックデータが既に記憶装置２０内に存在する場合は（ステップＳ３：Ｙｅｓ）、書き込み対象であるブロックデータとして既に存在するブロックデータを参照し、ブロック管理表１６内の参照されたブロックデータに該当するハッシュ値の行における重複率の値をインクリメントする（ステップＳ４）。

【００５１】

一方、書き込み部１１は、書き込み対象である該当するブロックデータのハッシュ値が未知のもの、つまり、記憶装置２０に格納されていない場合は（ステップＳ３：Ｎｏ）、領域管理表１５を参照し、使用中フラグが「yes」となっている領域２１（所定の領域）を見つけ出す。ここでは、該当する領域として、図７の書き込み領域３０が見つかったとする。なお、書き込み領域３０は、それぞれ所定の容量を有するデフラグ範囲３１、未使用範囲３２、長寿命範囲３３が、この順に連なって設定された領域である。つまり、未使用範囲３２がデフラグ範囲３１と長寿命範囲３３とに挟まれており、未使用範囲３２の先頭（一端側）にデフラグ範囲３１が隣接しており、未使用範囲３２の末尾（他端側）に長寿命範囲３３が隣接している。なお、書き込み領域３０には、必ずしも長寿命範囲３３が設定されていなくてもよい。

【００５２】

そして、書き込み部１１は、図８に示すように、書き込み領域３０の未使用範囲３２の先頭、つまり、デフラグ範囲３１との隣接箇所に、ブロックデータを書き込む（ステップＳ５、図８の網掛け部分参照）。なお、未使用範囲３２のデフラグ範囲３１との隣接箇所側に既に他のブロックデータが格納されている場合には、当該デフラグ範囲３１からデータが連続して位置する箇所に、ブロックデータを書き込む。

【００５３】

その後、書き込み部１１は、図９に示すように、書き込んだブロックデータのサイズ分だけデフラグ範囲３１の末尾を拡大すると共に、当該書き込んだブロックデータのサイズ分だけ未使用範囲３２の先頭を縮小する。換言すると、未使用範囲３１の先頭側に位置するデフラグ範囲３１との隣接箇所から連続してブロックデータが格納された範囲（図９の網掛け部分）と、デフラグ範囲（図８の符号３１）と、を連結した範囲を、新たなデフラグ範囲（図９の符号３１）とする。そして、領域管理表１５における書き込み領域３０の最終書込時刻を更新する。

【００５４】

加えて、書き込み部１１は、ブロック管理表１６に新規エントリを追加し、ブロックデータを書き込んだ領域の領域番号、物理アドレス、および、ブロックデータのサイズを格納する。最後に、ファイル管理表１７の該当ファイル名のリストに、書き込んだブロックデータのハッシュ値を追記する。

【００５５】

次に、読み出し部１２による動作を説明する。ファイルの読み出し要求があると、読み出し部１２は、ファイル管理表１７を参照し、ファイルを構成する各ブロックデータのハッシュ値を読み出して列挙する。そして、ブロック管理表１６を参照し、該当ブロックデータが格納されている領域および物理アドレスを割り出し、かかる格納場所からブロックデータの読み出し、及び、ファイルの復元を行うことで、要求されたファイルを読み出す。

【００５６】

次に、削除部１３による動作を説明する。ファイルの削除要求があると、削除部１３は、削除対象のブロックデータに該当するブロック管理表１６のエントリにおいて、重複率をデクリメントする。重複率が０になった場合、該当するエントリをブロック管理表１６から削除する。ただし、削除時に即座に重複率をデクリメントする必要はなく、データ削除後の適切なタイミングでデクリメントすることも可能である。

【００５７】

次に、デフラグ部１４によるデフラグ処理の動作を、図１０乃至図１３と、図１５のフローチャートを参照して説明する。書き込み実施後、使用中フラグが「yes」となっている領域２１について未使用領域が空になった場合、デフラグを実行する。デフラグには、２つの領域を用いる。このとき、図１３に示すように領域管理表１５を用いて、デフラグ元となる領域２１は、デフラグ範囲が空でない領域のうち最終書き込み時刻が最も古い領域を選択する。また、デフラグ先となる領域２１は、デフラグ範囲が空（０～０）である領域を選択する。これによって、ここでは図１０に示すように、デフラグ元となる領域２１としてデフラグ元領域４０（図１３の領域２）が選択され、デフラグ先となる領域２１としてデフラグ先領域５０（図１３の領域１）が選択されたこととする。そして、デフラグ元領域４０からデフラグ先領域５０へブロックデータを移動することでデフラグを行なう。

【００５８】

なお、デフラグ元領域４０及びデフラグ先領域５０は、それぞれ所定の容量を有するデフラグ範囲４１，５１、未使用範囲４２，５２、長寿命範囲４３，５３が、この順に連なって設定された領域である。つまり、未使用範囲４２，５２がデフラグ範囲４１，５１と長寿命範囲４３，５３とに挟まれており、未使用範囲４２，５２の先頭（一端側）にデフラグ範囲４１，５１が隣接しており、未使用範囲４２，５２の末尾（他端側）に長寿命範囲４３，５３が隣接している。

【００５９】
まず、デフラグ部１４は、デフラグ元領域４０内のデフラグ範囲４１に含まれるブロックデータを列挙する（ステップＳ１１）。次に、デフラグ部１４は、各ブロックデータについて、ブロック管理表１６を参照し、該当ブロックデータの重複率を確認する（ステップＳ１２）。
【００６０】
そして、デフラグ部１４は、該当ブロックデータの重複率が一定値未満である場合は（ステップＳ１３：Ｙｅｓ）、図１１の矢印Ｍ２，Ｍ４に示すように、デフラグ先領域５０の未使用領域５２の先頭にブロックデータを移動する（ステップＳ１４）。つまり、デフラグ先領域５０の未使用範囲５２のうち、デフラグ範囲５１との隣接箇所にブロックデータを書き込む（図１１の網掛け部分参照）。なお、未使用範囲５２のデフラグ範囲５１との隣接箇所側に既に他のブロックデータが格納されている場合には、当該デフラグ範囲５１からデータが連続して位置する箇所に、ブロックデータを書き込む。
【００６１】
その後、デフラグ部１４は、図１２に示すように、デフラグ先領域５０において、書き込んだブロックデータのサイズ分だけデフラグ範囲５１の末尾を拡大すると共に、当該書き込んだブロックデータのサイズ分だけ未使用範囲５２の先頭を縮小する。換言すると、未使用範囲５２の先頭側に位置するデフラグ範囲５１との隣接箇所から連続してブロックデータが格納された範囲（図１１の網掛け部分）と、デフラグ範囲（図１１の符号５１）と、を連結した範囲を、新たなデフラグ範囲（図１２の符号５１）とする。このデフラグ先領域５０の新たなデフラグ範囲５１は、後にデフラグ対象となる。そして、領域管理表１５におけるデフラグ先領域５０の最終書込時刻を更新する。
【００６２】
また、デフラグ部１４は、該当ブロックデータの重複率が一定値以上である場合は（ステップＳ１３：Ｎｏ）、図１１の矢印Ｍ１，Ｍ３に示すように、デフラグ先領域５０の未使用領域５２の末尾にブロックデータを移動する（ステップＳ１５）。つまり、デフラグ先領域５０の未使用範囲５２のうち、長寿命範囲５３との隣接箇所にブロックデータを書き込む（図１１の網掛け部分参照）。なお、未使用範囲５２の長寿命範囲５３との隣接箇所側に既に他のブロックデータが格納されている場合には、当該長寿命範囲５１からデータが連続して位置する箇所に、ブロックデータを書き込む。
【００６３】
その後、デフラグ部１４は、図１２に示すように、デフラグ先領域５０において、書き込んだブロックデータのサイズ分だけ長寿命範囲５３の先頭を拡大すると共に、当該書き込んだブロックデータのサイズ分だけ未使用範囲５２の末尾を縮小する。換言すると、未使用範囲５２の末尾側に位置する長寿命範囲５３との隣接箇所から連続してブロックデータが格納された範囲（図１１の網掛け部分）と、長寿命範囲（図１１の符号５３）と、を連結した範囲を、新たな長寿命範囲（図１２の符号５３）とする。
【００６４】
上述したデフラグ処理により、デフラグ元領域４０のデフラグ範囲４１内のブロックデータの重複率が「０」となった場合、ブロック管理表１６の該当する行を削除することで、該当ブロックデータを破棄する。この時点で、デフラグ先領域５０の未使用範囲５２が空になった場合、デフラグを中断する。また、ブロック管理表１６の該当するハッシュ値をもつ行について、領域番号と物理アドレスを、移動先を指し示すように更新する。これらの処理がすべて完了した時点で、デフラグ元領域４０のデフラグ範囲４１を未使用範囲４２に統合し、当該デフラグ範囲４１を「0-0」とする。
【００６５】
デフラグ後、デフラグ部１４あるいは書き込み部１１は、デフラグ先領域５０の使用中フラグを「yes」に変更し、それ以外の領域２１すべてについて使用中フラグを「no」にする。
【００６６】

以上のように、本実施形態におけるストレージシステムでは、デフラグ時に、重複率の高いブロックデータと重複率の低いブロックデータとを、別々の場所（デフラグ範囲と長寿命範囲）に格納している。これにより、次回のデフラグ時には、重複率の高いブロックを格納している長寿命範囲をデフラグの対象から除外し、重複率の低いブロックデータを格納しているデフラグ範囲のみをデフラグ対象としている。このため、領域２１全体をデフラグする場合と比較して、デフラグ処理に要するＩ／Ｏ負荷を低減させることができる。

【００６７】

また、データ書き込み時には、領域管理表１５における使用中フラグを用いることで、書き込み可能な領域を限定している。これにより、他の書き込み可能でない領域についてはデータの書き込みが発生せず、ブロックの削除のみが行なわれる。そのため、時間経過とともに、書き込みが行なわれていない領域は空き領域に変わることが期待できる。

【００６８】

さらに、重複率の低いデータは、時間経過とともに大半が削除され、連続した空き領域に変わることが期待できる。そのため、重複率の低いデータをまとめて一箇所に格納し、しばらく経ってから空き領域として利用することができる。一方で、重複率の高いデータは、以後削除される可能性が低い。そのため、一箇所にまとめて格納しておき、空き領域の断片化の発生を抑制することができる。

【００６９】

＜実施形態２＞

次に、本発明の第２の実施形態を、図１６乃至図２０を参照して説明する。図１６は、領域管理表の一例を示す図である。図１７乃至図１９は、ストレージシステムにおけるデフラグ処理の動作を説明するための説明図である。図２０は、ストレージシステムの動作を示すフローチャートである。

【００７０】

本実施形態におけるストレージシステムは、上述した実施形態１とほぼ同様の構成をとっているが、デフラグ時におけるブロックデータを格納するデフラグ先領域の構成が異なる。以下、実施形態１と異なる構成について主に説明する。

【００７１】

まず、本実施形態における記憶装置２０内の領域２１は、図１６の領域管理表１５に示すように、主に重複率の低いブロックデータを格納する「短寿命領域」と、主に重複率の高いブロックデータを格納するための「長寿命範囲」といった専用領域が設定されている。

【００７２】

そして、図１７に示すように、「短寿命領域」は、デフラグ時に、一次デフラグ先領域６０（第一領域）として選択される。この「短寿命領域」である一次デフラグ先領域６０は、所定の容量を有するデフラグ範囲６１、未使用範囲６２、長寿命範囲６３が、この順に連なって設定された領域である。つまり、未使用範囲６２がデフラグ範囲６１と長寿命範囲６３とに挟まれており、未使用範囲６２の先頭（一端側）にデフラグ範囲６１が隣接しており、未使用範囲６２の末尾（他端側）に長寿命範囲６３が隣接している。

【００７３】

また、図１７に示すように、「長寿命領域」は、デフラグ時に、二次デフラグ先領域７０（第二領域）として選択される。この「長寿命領域」である二次デフラグ先領域７０は、所定の容量を有するデフラグ範囲７１、未使用範囲７２、長寿命範囲７３が、この順に連なって設定された領域である。つまり、未使用範囲７２がデフラグ範囲７１と長寿命範囲７３とに挟まれており、未使用範囲７２の先頭（一端側）にデフラグ範囲７１が隣接しており、未使用範囲７２の末尾（他端側）に長寿命範囲７３が隣接している。

【００７４】

次に、デフラグ時の動作を説明する。まず、図１６に示す領域管理表１５より、デフラグ元となるデフラグ元領域４０として、デフラグ範囲が空でない領域のうち、最終書き込

み時刻が最も古い領域を選択する。また、デフラグ先となるデフラグ先領域６０，７０は、短寿命用と長寿命用との２種類を選択する。短寿命用の一次デフラグ領域６０として、特性が短寿命の領域のうち未使用範囲が最大である領域を選択する。長寿命用の二次デフラグ領域７０として、特性が長寿命の領域のうちデフラグ範囲が空である領域を選択する。ここでは、図１７に示すように、短寿命用の一次デフラグ領域６０と、長寿命用の二次デフラグ領域７０と、が選択されたこととする。そして、デフラグ元領域４０から一次デフラグ先領域６０、二次デフラグ領域７０へブロックデータを移動することでデフラグを行う。

【００７５】

まず、デフラグ部１４は、デフラグ元領域４０内のデフラグ範囲４１に含まれるブロックデータを列挙する（ステップＳ２１）。次に、デフラグ部１４は、各ブロックデータについて、ブロック管理表１６を参照し、該当ブロックデータの重複率を確認する（ステップＳ２２）。

【００７６】

そして、デフラグ部１４は、該当ブロックデータの重複率が一定値未満である場合は（ステップＳ２３：Ｙｅｓ）、図１８に示すように、一次デフラグ先領域６０の未使用領域６２の先頭にブロックデータを移動する（ステップＳ２４）。つまり、デフラグ先領域６０の未使用範囲６２のうち、デフラグ範囲６１との隣接箇所にブロックデータを書き込む（図１８の網掛け部分参照）。

【００７７】

その後、デフラグ部１４は、図１９に示すように、一次デフラグ先領域６０において、書き込んだブロックデータのサイズ分だけデフラグ範囲６１の末尾を拡大すると共に、当該書き込んだブロックデータのサイズ分だけ未使用範囲６２の先頭を縮小する。換言すると、未使用範囲６２の先頭側に位置するデフラグ範囲６１との隣接箇所から連続してブロックデータが格納された範囲（図１８の網掛け部分）と、デフラグ範囲（図１８の符号６１）と、を連結した範囲を、新たなデフラグ範囲（図１９の符号６１）とする。これにより、一次デフラグ先領域６０の新たなデフラグ範囲６１は、後にデフラグ対象となる。そして、領域管理表１５における一次デフラグ先領域６０の最終書込時刻を更新する。

【００７８】

また、デフラグ部１４は、該当ブロックデータの重複率が一定値以上である場合は（ステップＳ２３：Ｎｏ）、図１８に示すように、二次デフラグ先領域７０の未使用領域７２の末尾にブロックデータを移動する（ステップＳ２５）。つまり、二次デフラグ先領域７０の未使用範囲７２のうち、長寿命範囲７３との隣接箇所にブロックデータを書き込む（図１８の網掛け部分参照）。

【００７９】

その後、デフラグ部１４は、図１９に示すように、二次デフラグ先領域７０において、書き込んだブロックデータのサイズ分だけ長寿命範囲７３の先頭を拡大すると共に、当該書き込んだブロックデータのサイズ分だけ未使用範囲７２の末尾を縮小する。換言すると、未使用範囲７２の末尾側に位置する長寿命範囲７３との隣接箇所から連続してブロックデータが格納された範囲（図１８の網掛け部分）と、長寿命範囲（図１８の符号７３）と、を連結した範囲を、新たな長寿命範囲（図１９の符号７３）とする。

【００８０】

上述したデフラグ処理により、デフラグ元領域４０のデフラグ範囲４１内のブロックデータの重複率が「０」となった場合、ブロック管理表１６の該当する行を削除することで、該当ブロックデータを破棄する。この時点で、一次デフラグ先領域６０の未使用範囲６２が空になった場合、デフラグを中断し、二次デフラグ先領域７０の未使用範囲７２が空になった場合、他の特性が超寿命となっている領域を選びなおし続行する。また、ブロック管理表１６の該当するハッシュ値をもつ行について、領域番号と物理アドレスを、移動先を指し示すように更新する。これらの処理がすべて完了した時点で、デフラグ元領域４０のデフラグ範囲４１を未使用範囲４２に統合し、当該デフラグ範囲４１を「0-0」とす

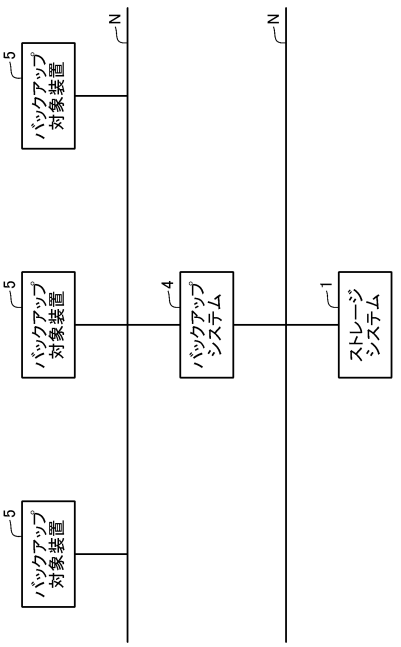
る。
【００８１】
デフラグ後、デフラグ部１４あるいは書き込み部１１は、一次デフラグ先領域６０の使用中フラグを「yes」に変更し、それ以外の領域２１すべてについて使用中フラグを「no」にする。
【００８２】
以上のように、重複率に応じてブロックデータを格納するデフラグ先として、「短寿命領域」と「長寿命範囲」といった専用領域を設定することで、デフラグ時におけるデフラグ先領域の確保が容易になり、デフラグ処理を実行する手段の実装が簡潔になる。
【００８３】
ここで、上記記憶装置２０として、データのアクセス速度、特に、読み出し速度が異なる２種類の装置を用いてもよい。例えば、記憶装置２０は、図２１に示すように、ハードディスクドライブといった一次記憶装置２０Ａと、当該一次記憶装置２０Ａよりもデータのアクセスが高速なＳＳＤ（Solid State Drive）といった二次記憶装置２０Ｂと、を備える。そして、一次記憶装置２０Ａ内の領域２１は、主に重複率の低いブロックデータを格納する上記一次デフラグ先領域６０として使用し、二次記憶装置２０Ｂ内の領域２１は、主に重複率の高いブロックデータを格納する上記二次デフラグ先領域７０として使用する。
【００８４】
これにより、重複率の高いブロックデータは高頻度で参照されることとなるが、かかるブロックデータがデフラグ処理により集約される上記二次デフラグ先領域７０を、読み出し速度が高速な二次記憶装置２０Ｂに設けることで、読み込み性能を改善することができる。
【００８５】
＜付記＞
上記実施形態の一部又は全部は、以下の付記のようにも記載されうる。以下、本発明におけるストレージシステム（図２２参照）、プログラム、デフラグ方法の構成の概略を説明する。但し、本発明は、以下の構成に限定されない。
【００８６】
（付記１）
記憶対象データを記憶装置１２０に格納すると共に、当該記憶装置に既に記憶されている前記記憶対象データと同一のデータ内容の他の記憶対象データを前記記憶装置に格納する場合に、当該記憶装置に既に記憶されている前記記憶対象データを前記他の記憶対象データとして参照させるデータ格納制御部１１１と、
前記記憶装置の所定の領域内においてデフラグ範囲とされた箇所に格納された前記記憶対象データを、前記記憶装置の他の領域内に格納し直すデフラグ処理部１１２と、を備え、
前記データ格納制御部１１１は、前記記憶装置に記憶されている前記記憶対象データ毎に、当該記憶対象データが他の記憶対象データとして参照されている数である参照数を記憶し、
前記デフラグ処理部１１２は、前記記憶対象データの参照数に応じて当該記憶対象データを前記記憶装置の他の領域内において後にデフラグ範囲となる箇所に格納する、
ストレージシステム１００。
【００８７】
（付記２）
付記１に記載のストレージシステムであって、
前記デフラグ処理部は、前記参照数が予め設定された閾値未満である前記記憶対象データを、前記記憶装置の他の領域内において後にデフラグ範囲となる箇所に格納し、前記参照数が予め設定された閾値以上である前記記憶対象データを、前記記憶装置の他の領域内において後に非デフラグ範囲となる箇所に格納する、

ストレージシステム。
【００８８】
（付記３）
付記２に記載のストレージシステムであって、
前記記憶装置の他の領域は、予め設定されたデフラグ範囲と、データが未格納な範囲である未使用範囲と、が隣接して形成されており、
前記デフラグ処理部は、前記参照数が閾値未満である前記記憶対象データを、前記記憶装置の他の領域における前記未使用範囲内であり前記デフラグ範囲との隣接箇所からデータが連続して位置する箇所に格納し、前記未使用範囲内における前記デフラグ範囲との隣接箇所から連続して前記記憶対象データが格納された範囲と前記デフラグ範囲とを連結した範囲を新たなデフラグ範囲とする、
ストレージシステム。
【００８９】
（付記４）
付記３に記載のストレージシステムであって、
前記記憶装置の他の領域は、予め設定された非デフラグ範囲と、データが未格納な範囲である未使用範囲と、が隣接して形成されており、
前記デフラグ処理部は、前記参照数が閾値以上である前記記憶対象データを、前記記憶装置の他の領域における前記未使用範囲であり前記非デフラグ範囲との隣接箇所からデータが連続して位置する箇所に格納し、前記未使用範囲内における前記非デフラグ範囲との隣接箇所から連続して前記記憶対象データが格納された範囲と前記非デフラグ範囲とを連結した範囲を新たな非デフラグ範囲とする、
ストレージシステム。
【００９０】
（付記５）
付記２に記載のストレージシステムであって、
前記記憶装置の他の領域は、予め設定されたデフラグ範囲にデータが未格納な範囲である未使用範囲の一端側が隣接すると共に、当該未使用範囲の他端側が予め設定された非デフラグ範囲に隣接し、当該未使用範囲がデフラグ範囲と非デフラグ範囲とに挟まれて形成されており、
前記デフラグ処理部は、前記参照数が閾値未満である前記記憶対象データを、前記記憶装置の他の領域における前記未使用範囲内であり前記デフラグ範囲との隣接箇所からデータが連続して位置する箇所に格納し、前記未使用範囲内における前記デフラグ範囲との隣接箇所から連続して前記記憶対象データが格納された範囲と前記デフラグ範囲とを連結した範囲を新たなデフラグ範囲とし、前記参照数が閾値以上である前記記憶対象データを、前記記憶装置の他の領域における前記未使用範囲であり前記非デフラグ範囲との隣接箇所からデータが連続して位置する箇所に格納し、前記未使用範囲内における前記非デフラグ範囲との隣接箇所から連続して前記記憶対象データが格納された範囲と前記非デフラグ範囲とを連結した範囲を新たな非デフラグ範囲とする、
ストレージシステム。
【００９１】
（付記６）
付記１乃至５のいずれかに記載のストレージシステムであって、
前記記憶装置の所定の領域は、予め設定されたデフラグ範囲と、データが未格納な範囲である未使用範囲と、が隣接して形成されており、
前記データ格納制御部は、前記記憶対象データを新たに前記記憶装置に格納する際に、前記記憶装置の所定の領域における前記未使用範囲内であり前記デフラグ範囲との隣接箇所からデータが連続して位置する箇所に格納し、前記未使用範囲内における前記デフラグ範囲との隣接箇所から連続して前記記憶対象データが格納された範囲と前記デフラグ範囲とを連結した範囲を新たなデフラグ範囲とする、

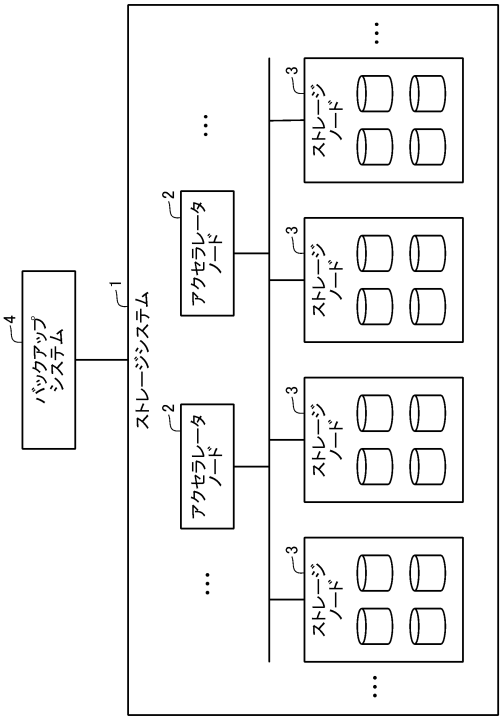
ストレージシステム。
【００９２】
（付記７）
付記４に記載のストレージシステムであって、
前記記憶装置の他の領域は、前記デフラグ範囲と前記未使用範囲とが隣接して形成された第一領域と、当該第一領域とは異なり前記非デフラグ範囲と前記未使用範囲とが隣接して形成された第二領域と、を有する、
ストレージシステム。
【００９３】
（付記８）
付記７に記載のストレージシステムであって、
前記第一領域は、所定の記憶装置に形成されており、
前記第二領域は、前記所定の記憶装置よりもデータ読み出し処理が高速な他の記憶装置に形成されている、
ストレージシステム。
【００９４】
（付記９）
情報処理装置に、
記憶対象データを記憶装置に格納すると共に、当該記憶装置に既に記憶されている前記記憶対象データと同一のデータ内容の他の記憶対象データを前記記憶装置に格納する場合に、当該記憶装置に既に記憶されている前記記憶対象データを前記他の記憶対象データとして参照させるデータ格納制御部と、
前記記憶装置の所定の領域内においてデフラグ範囲とされた箇所に格納された前記記憶対象データを、前記記憶装置の他の領域内に格納し直すデフラグ処理部と、を実現させると共に、
前記データ格納制御部は、前記記憶装置に記憶されている前記記憶対象データ毎に、当該記憶対象データが他の記憶対象データとして参照されている数である参照数を記憶し、
前記デフラグ処理部は、前記記憶対象データの参照数に応じて当該記憶対象データを前記記憶装置の他の領域内において後にデフラグ範囲となる箇所に格納する、
ことを実現させるためのプログラム。
【００９５】
（付記９－２）
付記９に記載のプログラムであって、
前記デフラグ処理部は、前記参照数が予め設定された閾値未満である前記記憶対象データを、前記記憶装置の他の領域内において後にデフラグ範囲となる箇所に格納し、前記参照数が予め設定された閾値以上である前記記憶対象データを、前記記憶装置の他の領域内において後に非デフラグ範囲となる箇所に格納する、
プログラム。
【００９６】
（付記１０）
記憶対象データを記憶装置に格納すると共に、当該記憶装置に既に記憶されている前記記憶対象データと同一のデータ内容の他の記憶対象データを前記記憶装置に格納する場合に、当該記憶装置に既に記憶されている前記記憶対象データを前記他の記憶対象データとして参照させてデータ格納制御を行うと共に、前記記憶装置に記憶されている前記記憶対象データ毎に、当該記憶対象データが他の記憶対象データとして参照されている数である参照数を記憶し、
前記記憶装置の所定の領域内においてデフラグ範囲とされた箇所に格納された前記記憶対象データを、前記記憶装置の他の領域内に格納し直すデフラグ処理を実行し、
前記デフラグ処理時に、前記記憶対象データの参照数に応じて当該記憶対象データを前記記憶装置の他の領域内において後にデフラグ範囲となる箇所に格納する、

デフラグ方法。
【００９７】
（付記１０－２）
付記１０に記載のデフラグ方法であって、
前記デフラグ処理時に、前記参照数が予め設定された閾値未満である前記記憶対象データを、前記記憶装置の他の領域内において後にデフラグ範囲となる箇所に格納し、前記参照数が予め設定された閾値以上である前記記憶対象データを、前記記憶装置の他の領域内において後に非デフラグ範囲となる箇所に格納する、
デフラグ方法。
【００９８】
なお、上述したプログラムは、記憶装置に記憶されていたり、コンピュータが読み取り可能な記録媒体に記録されている。例えば、記録媒体は、フレキシブルディスク、光ディスク、光磁気ディスク、及び、半導体メモリ等の可搬性を有する媒体である。
【００９９】
以上、上記実施形態等を参照して本願発明を説明したが、本願発明は、上述した実施形態に限定されるものではない。本願発明の構成や詳細には、本願発明の範囲内で当業者が理解しうる様々な変更をすることができる。
【符号の説明】
【０１００】
１　ストレージシステム
２　アクセラレータノード
３　ストレージノード
４　バックアップシステム
５　バックアップ対象装置
１１　書き込み部
１２　読み出し部
１３　削除部
１４　デフラグ部
１５　領域管理表
１６　ブロック管理表
１７　ファイル管理表
２０　記憶装置
２０Ａ　一次記憶装置
２０Ｂ　二次記憶装置
２１　領域
３０　書き込み領域
４０　デフラグ元領域
５０　デフラグ先領域
６０　一次デフラグ領域
７０　二次デフラグ領域
３１，４１，５１，６１，７１　デフラグ範囲
３２，４２，５２，６２，７２　未使用範囲
３３，４３，５３，６３，７３　長寿命範囲
１００　ストレージシステム
１１１　データ格納制御部
１１２　デフラグ処理部
１２０　記憶装置

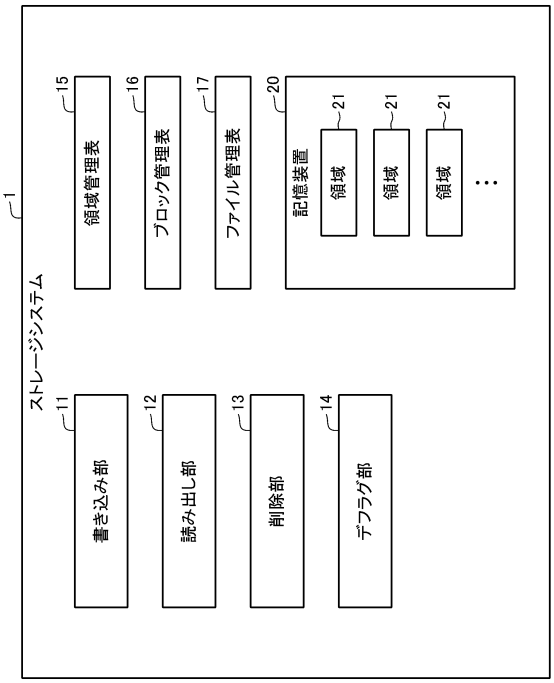
【図１】


ストレージシステム
1
バックアップシステム
4
バックアップ対象装置
5
N

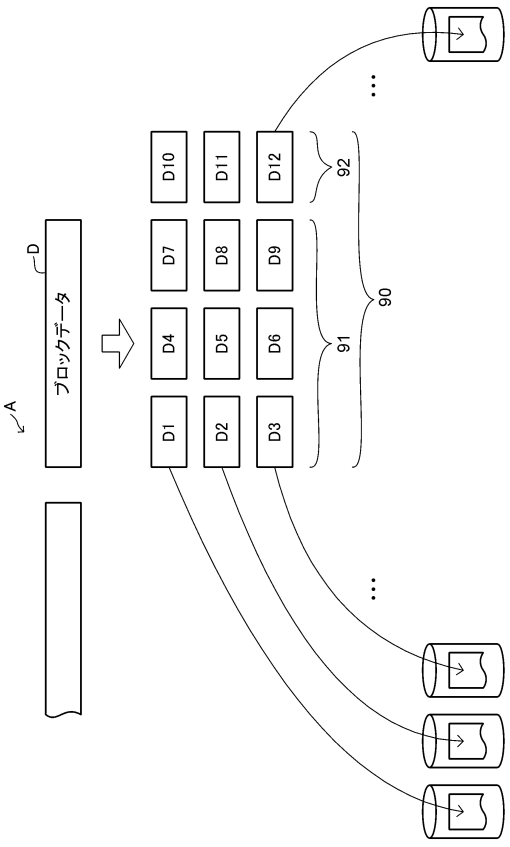
【図２】


ストレージシステム
1
バックアップシステム
4
アクセラレータノード
2
ストレージノード
3

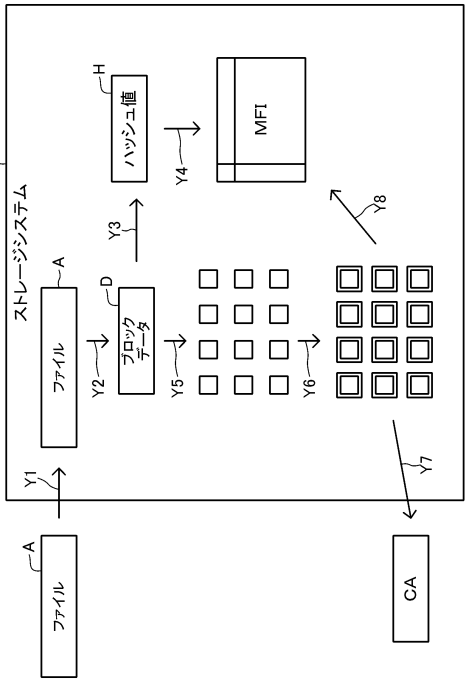
【図３】


ストレージシステム
1
書き込み部
11
読み出し部
12
削除部
13
デフラグ部
14
領域管理表
15
ブロック管理表
16
ファイル管理表
17
記憶装置
20
領域
21

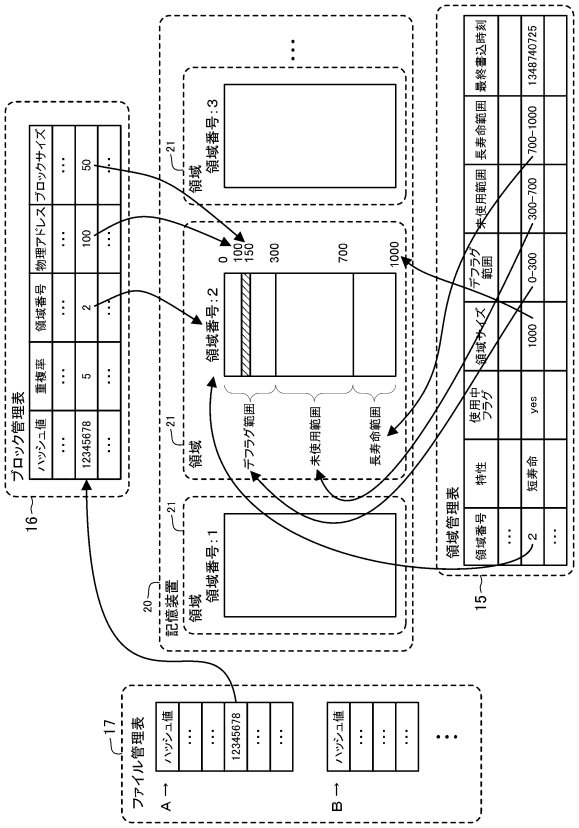
【図４】


A
ブロックデータ
D
D1
D2
D3
D4
D5
D6
D7
D8
D9
D10
D11
D12
91
92
90

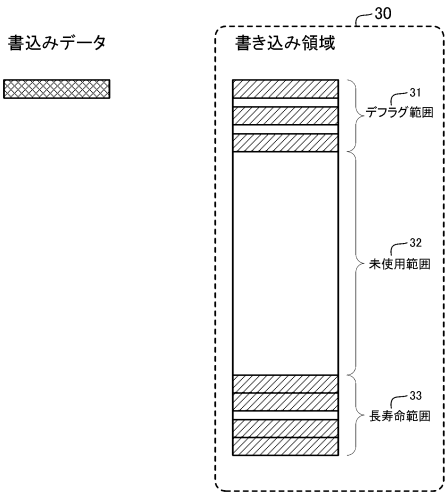
【図５】


1
ストレージシステム
A
ファイル
D
ブロックデータ
H
ハッシュ値
MFI
Y1
Y2
Y3
Y4
Y5
Y6
Y7
Y8
A
ファイル
CA

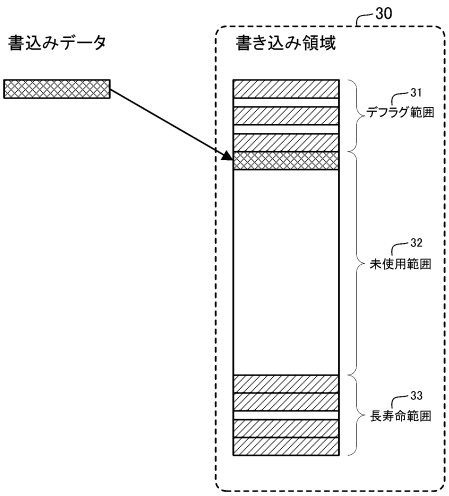
【図６】


ブロック管理表
ハッシュ値
重複率
領域番号
物理アドレス
ブロックサイズ
12345678
5
2
100
50
16
領域
領域番号:3
領域番号:2
領域番号:1
21
0
100
150
300
700
1000
デフラグ範囲
未使用範囲
長寿命範囲
20
記憶装置
領域管理表
領域番号
特性
使用中フラグ
領域サイズ
デフラグ範囲
未使用範囲
長寿命範囲
最終書込時刻
2
短寿命
yes
1000
0-300
300-700
700-1000
1348740725
15
ファイル管理表
17
A→
B→
12345678

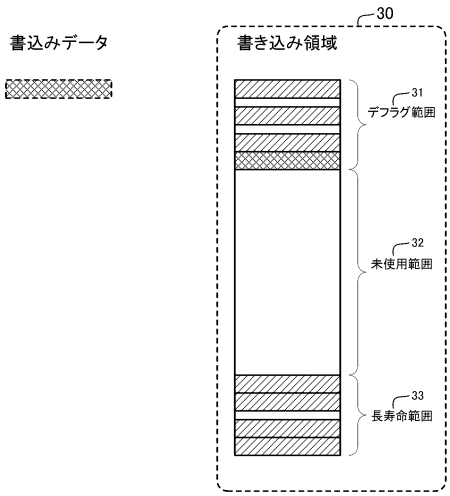
【図７】


書込みデータ
30
書き込み領域
31
デフラグ範囲
32
未使用範囲
33
長寿命範囲

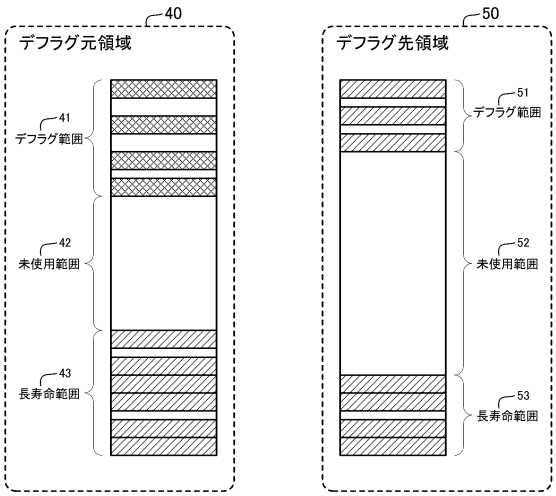
【図８】


書込みデータ
30
書き込み領域
31
デフラグ範囲
32
未使用範囲
33
長寿命範囲

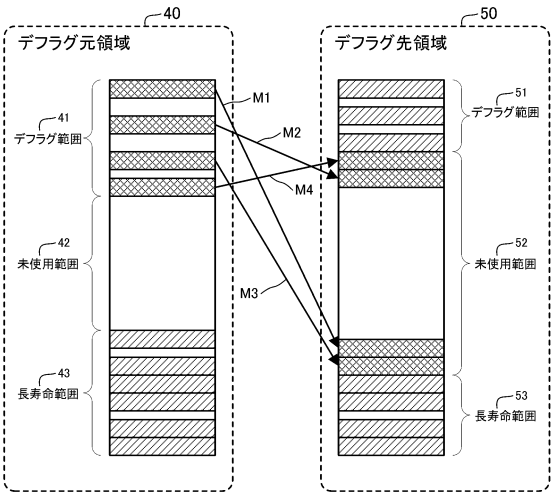
【図９】



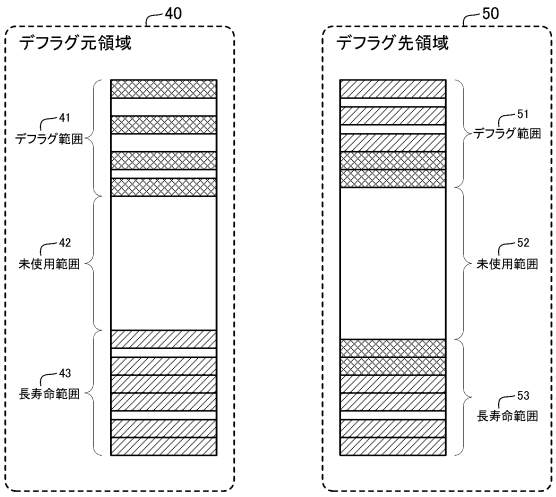
【図１０】



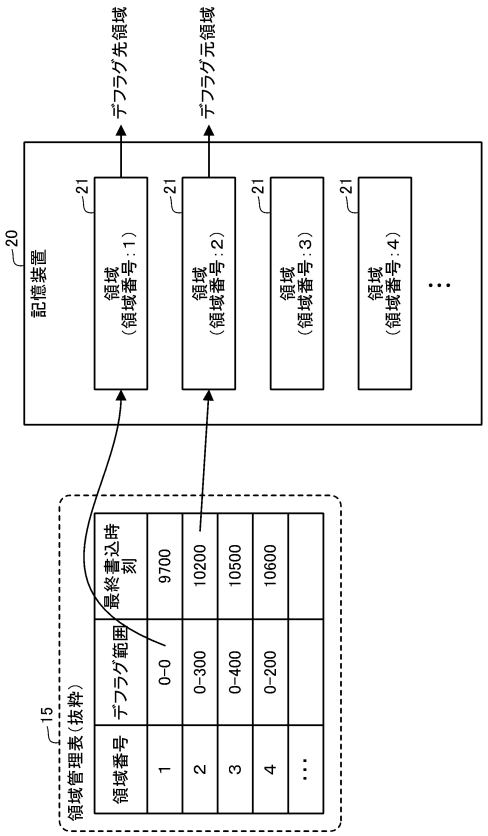
【図１１】



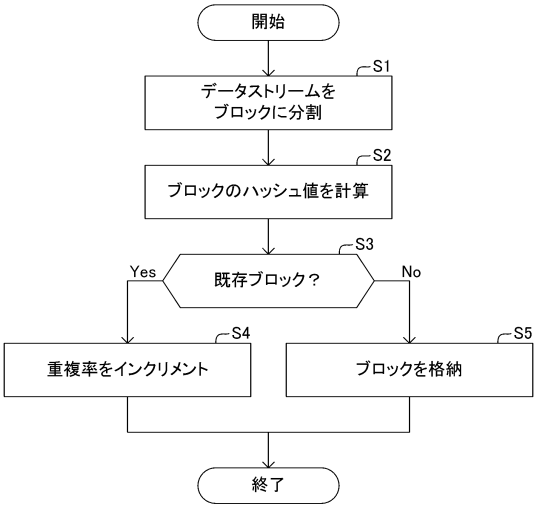
【図１２】



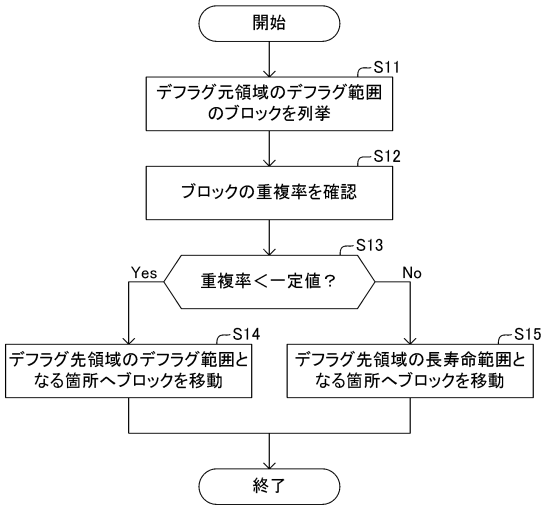
【図１３】



【図１４】



【図１５】

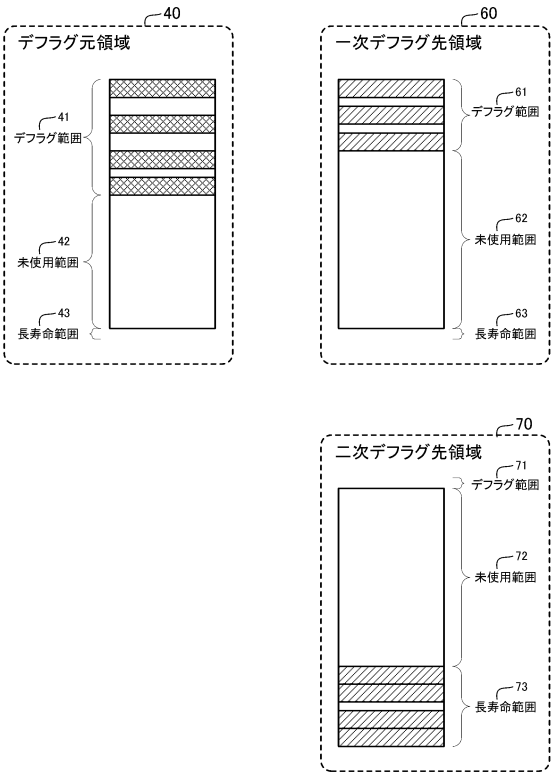


【図１６】

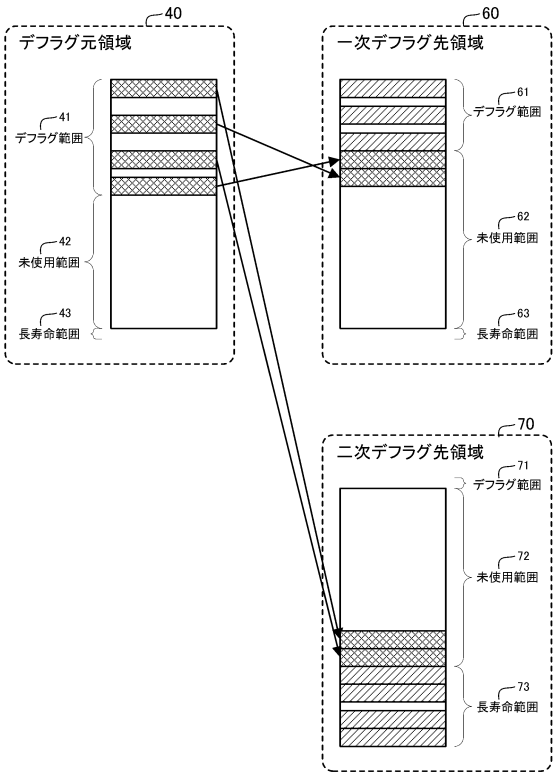
15 領域管理表

領域番号	特性	使用中フラグ	領域サイズ	デフラグ範囲	未使用範囲	長寿命範囲	最終書込時刻
…							
2	短寿命	yes	1000	0-300	300-700	700-1000	1348740725
…							
5	長寿命	-	1000	0-0	0-600	600-1000	-
…							

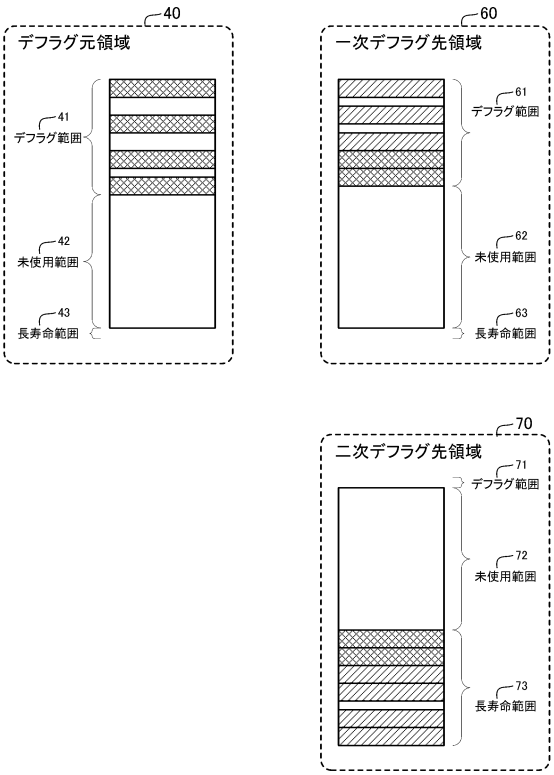
【図１７】



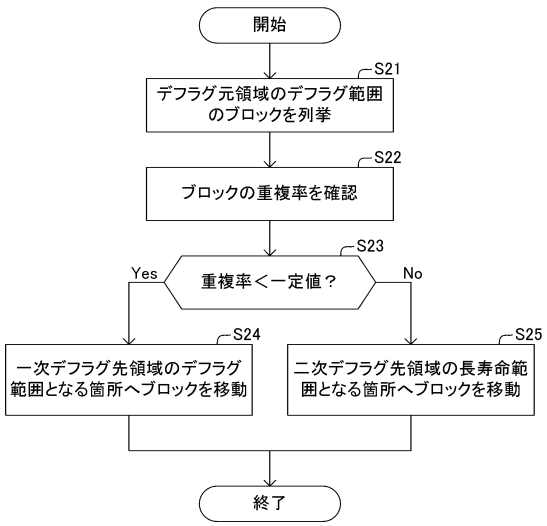
【図１８】



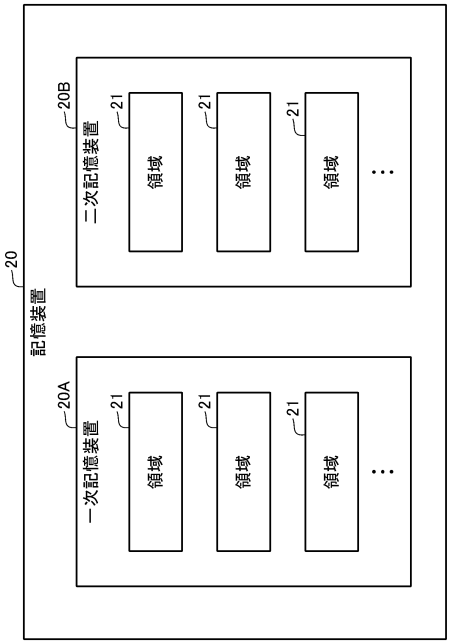
【図１９】



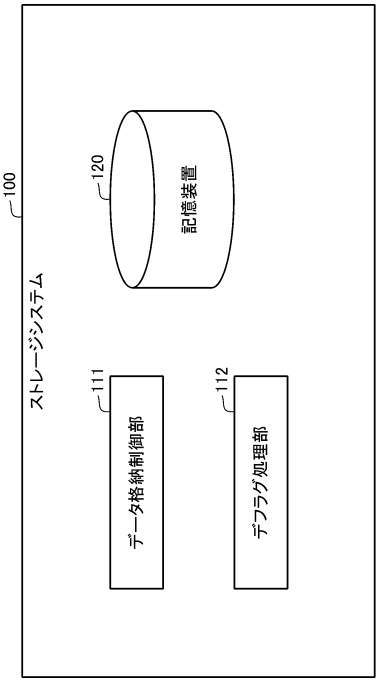
【図２０】



【図２１】


20
記憶装置
20A
一次記憶装置
21
領域
20B
二次記憶装置

【図２２】


100
ストレージシステム
120
記憶装置
111
データ格納制御部
112
デフラグ処理部

フロントページの続き

(56)参考文献 特開２０１１－１７０６６５（ＪＰ，Ａ）
特開２００９－２０５２０１（ＪＰ，Ａ）
特開２０１０－２１８１９４（ＪＰ，Ａ）
米国特許出願公開第２０１３／０２２６８８１（ＵＳ，Ａ１）

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