(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号

特許第6136460号
(P6136460)

(45)発行日　平成29年5月31日(2017.5.31)　(24)登録日　平成29年5月12日(2017.5.12)

(51)Int.Cl.　F I
G06F 3/06 (2006.01)　G06F 3/06 304P
G06F 3/08 (2006.01)　G06F 3/08 H
G06F 3/06 540
G06F 3/06 304Z

請求項の数 9　(全 24 頁)

(21)出願番号　特願2013-70675(P2013-70675)
(22)出願日　平成25年3月28日(2013.3.28)
(65)公開番号　特開2014-194667(P2014-194667A)
(43)公開日　平成26年10月9日(2014.10.9)
審査請求日　平成27年12月4日(2015.12.4)

(73)特許権者　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　100089118
弁理士　酒井　宏明
(72)発明者　小野　貴継
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

審査官　田上　隆一

最終頁に続く

(54)【発明の名称】情報処理装置、情報処理装置の制御プログラムおよび情報処理装置の制御方法

(57)【特許請求の範囲】
【請求項１】
データが読み書きされる半導体記憶装置と、前記半導体記憶装置の予備である予備半導体記憶装置とを有する情報処理装置において、
前記半導体記憶装置について、書き込まれたデータを消去できる残余回数に基づいた余命指標値、および、前記半導体記憶装置に対する読出しと書き込みとの読み書きの比率を収集する収集部と、
前記収集部が収集した前記比率においてデータの読出しの比率が多い場合、第１の閾値を前記半導体記憶装置の交換基準となる閾値として決定し、データの書き込みの比率が多い場合、前記第１の閾値よりも値が大きい第２の閾値を前記半導体記憶装置の交換基準となる閾値として決定する決定部と、
前記収集部が収集した前記余命指標値が、前記決定部が決定した閾値よりも短い場合は、前記半導体記憶装置を前記予備半導体記憶装置と交換する交換部と
を有することを特徴とする情報処理装置。
【請求項２】
前記情報処理装置はさらに、
前記予備半導体記憶装置を複数有し、
前記決定部は、
収集した前記比率においてデータの読出しの比率が多い場合、前記複数の予備半導体記憶装置のうち、余命指標値が前記第１の閾値よりも大きく、かつ、前記第２の閾値よりも

小さい予備半導体記憶装置を選択し、
前記交換部は、
前記半導体記憶装置を前記決定部が選択した予備半導体記憶装置と交換することを特徴とする請求項１に記載の情報処理装置。
【請求項３】
前記決定部は、
収集した前記比率においてデータの書き込みの比率が多い場合、前記複数の予備半導体記憶装置のうち、余命が前記第２の閾値よりも大きい予備半導体記憶装置を選択することを特徴とする請求項２に記載の情報処理装置。
【請求項４】
前記収集部は、
前記比率として、単位時間当たりに前記半導体記憶装置から読出されたデータ量と前記半導体記憶装置に書き込まれたデータ量との比率を収集し、
前記決定部は、
単位時間当たりの前記半導体記憶装置から読み出したデータ量が所定の閾値よりも多い場合、データの読出しの比率が多いと判定することを特徴とする請求項１～３のいずれか１つに記載の情報処理装置。
【請求項５】
前記収集部は、
前記比率として、単位時間当たりに前記半導体記憶装置からデータを読出した回数とデータを書き込んだ回数との比率を収集し、
前記決定部は、
単位時間当たりの前記データの読出し回数が所定の閾値よりも多い場合、データの読出しの比率が多いと判定することを特徴とする請求項１～３のいずれか１つに記載の情報処理装置。
【請求項６】
前記情報処理装置はさらに、
データが読み書きされる複数の半導体記憶装置と、
前記半導体記憶装置が、二重化されている場合、前記交換部が、前記半導体記憶装置を前記予備の半導体記憶装置と交換する間、前記半導体記憶装置に対する書込みデータを保持する保持部と、
前記交換部が、交換対象の半導体記憶装置が記憶するデータと、前記保持部が保持したデータとを、前記予備半導体記憶装置と、前記複数の半導体記憶装置のうち前記予備半導体記憶装置と二重化される半導体記憶装置とに複製する複製部と
を有することを特徴とする請求項１～５のいずれか１つに記載の情報処理装置。
【請求項７】
前記情報処理装置はさらに、
データが読み書きされる複数の半導体記憶装置を有し、
前記交換部は、
前記複数の半導体記憶装置に、データとデータを復元するためのパリティとが分散して格納されている場合、前記複数の半導体記憶装置のうち、いずれかの半導体記憶装置を前記予備半導体記憶装置と交換したとき、前記複数の半導体記憶装置のうち、前記予備半導体記憶装置以外の半導体記憶装置が記憶するデータとパリティとを用いて、前記予備半導体記憶装置が記憶するデータを生成することを特徴とする請求項１～５のいずれか１つに記載の情報処理装置。
【請求項８】
データが読み書きされる半導体記憶装置と、前記半導体記憶装置の予備である予備半導体記憶装置とを有する情報処理装置の制御プログラムにおいて、
前記情報処理装置に、
前記半導体記憶装置について、書き込まれたデータを消去できる残余回数に基づいた余

命指標値、および、前記半導体記憶装置に対する読出しと書き込みとの読み書きの比率を収集させ、
収集された前記比率においてデータの読出しの比率が多い場合、第１の閾値を前記半導体記憶装置の交換基準となる閾値として決定し、データの書き込みの比率が多い場合、前記第１の閾値よりも値が大きい第２の閾値を前記半導体記憶装置の交換基準となる閾値として決定させ、
収集された前記余命指標値が、決定された前記閾値よりも短い場合は、前記半導体記憶装置を前記予備半導体記憶装置と交換させる
ことを特徴とする情報処理装置の制御プログラム。
【請求項９】
データが読み書きされる半導体記憶装置と、前記半導体記憶装置の予備である予備半導体記憶装置とを有する情報処理装置の制御方法において、
前記情報処理装置が、
前記半導体記憶装置について、書き込まれたデータを消去できる残余回数に基づいた余命指標値、および、前記半導体記憶装置に対する読出しと書き込みとの読み書きの比率を収集し、
収集された前記比率においてデータの読出しの比率が多い場合、第１の閾値を前記半導体記憶装置の交換基準となる閾値として決定し、データの書き込みの比率が多い場合、前記第１の閾値よりも値が大きい第２の閾値を前記半導体記憶装置の交換基準となる閾値として決定し、
収集された前記余命指標値が、決定された前記閾値よりも短い場合は、前記半導体記憶装置を前記予備半導体記憶装置と交換する
ことを特徴とする情報処理装置の制御方法。
【発明の詳細な説明】
【技術分野】
【０００１】
本発明は、情報処理装置、情報処理装置の制御プログラムおよび情報処理装置の制御方法に関する。
【背景技術】
【０００２】
従来、不揮発性の半導体を含む記憶装置をストレージとして用いる情報処理装置が知られている。このような情報処理装置の一例として、ＮＡＮＤ型のフラッシュメモリを有するＳＳＤ（Solid State Drive）をストレージとして用いる情報処理装置が知られている。
【０００３】
ここで、ＮＡＮＤ型のフラッシュメモリでは、データを保持する複数のメモリセルを含むページ単位で書き込み及び読出しが行われ、複数のページを含むブロック単位でデータの消去が行われる。しかし、メモリセルは、データの書き換えを行う度に劣化するので、データの書き換えを何度も実行すると、正常に情報を記録できなくなる。このため、同じメモリセルに対するデータの書き換えが頻繁に発生すると、メモリセルが正常に情報を記録できなくなり、正常に情報を記録できなくなったメモリセルを含むブロックが不良ブロックとなる。
【０００４】
そこで、メモリセルへの書き込み回数や消去回数を平坦化することで、不良ブロックの発生を防ぎ、ＳＳＤの寿命を改善するウェアレベリングの技術が知られている。例えば、ウェアレベリングの技術が適用されたＳＳＤは、頻繁に更新されるブロックに格納されたデータを更新回数が少ないブロックに移動することで、ＳＳＤが有するメモリセル全体の更新回数を平坦化する。
【先行技術文献】
【特許文献】

【０００５】
【特許文献１】特開２００７－３２３２２４号公報
【発明の概要】
【発明が解決しようとする課題】
【０００６】
しかしながら、上述したウェアレベリングの技術は、１つのＳＳＤ内における書き込み回数や消去回数の制限を根本的に解決するものではなく、ＳＳＤの寿命を延命するに過ぎない。このため、情報処理装置は、フラッシュメモリの書き込み回数や消去回数の制限を考慮せずに１つのＳＳＤを使用し続けることができない。
【０００７】
本願は、上述した問題に鑑みてなされたものであり、１つの側面では、寿命を考慮せずにＳＳＤをストレージとして利用することを目的とする。
【課題を解決するための手段】
【０００８】
１つの側面では、データが読み書きされる半導体記憶装置と、半導体記憶装置の予備である予備半導体記憶装置とを有する情報処理装置である。また、情報処理装置は、半導体記憶装置について、書き込まれたデータを消去できる残余回数に基づいた余命指標値、および、半導体記憶装置に対する読出しと書き込みとの読み書きの比率を収集する。また、情報処理装置は、収集した比率においてデータの読出しの比率が多い場合、第１の閾値を前記半導体記憶装置の交換基準となる閾値として決定し、データの書き込みの比率が多い場合、前記第１の閾値よりも値が大きい第２の閾値を前記半導体記憶装置の交換基準となる閾値として決定する。そして、情報処理装置は、収集した余命指標値が、決定した閾値よりも短い場合は、半導体記憶装置を予備半導体記憶装置と交換する。
【発明の効果】
【０００９】
１つの側面では、寿命を考慮せずにＳＳＤをストレージとして利用することができる。
【図面の簡単な説明】
【００１０】
【図１】図１は、実施例１に係る情報処理装置を説明する図である。
【図２】図２は、実施例１に係る情報処理装置が実行する処理を説明する図である。
【図３】図３は、実施例１に係る管理サーバの機能構成を説明する図である。
【図４】図４は、余命管理テーブルの一例を説明する図である。
【図５】図５は、ワークロード管理テーブルの一例を説明する図である。
【図６】図６は、実施例１に係る情報処理装置が実行する処理の流れを説明するフローチャートである。
【図７】図７は、選択部が実行する処理の流れを説明するための第１のフローチャートである。
【図８】図８は、選択部が実行する処理の流れを説明するための第２のフローチャートである。
【図９】図９は、実施例２に係る情報処理装置が実行する処理の流れを説明するフローチャートである。
【図１０】図１０は、制御プログラムを実行するコンピュータを説明する図である。
【発明を実施するための形態】
【００１１】
以下に添付図面を参照して本願に係る情報処理装置、情報処理装置の制御プログラムおよび情報処理装置の制御方法の実施例を図面に基づいて詳細に説明する。なお、この実施例により開示技術が限定されるものではない。また、各実施例は、矛盾しない範囲で適宜組みあわせてもよい。
【実施例１】
【００１２】

以下の実施例１では、図１を用いて、本願に係る情報処理装置を説明する。図１は、実施例１に係る情報処理装置を説明する図である。図１に示すように、情報処理装置１は、ＣＰＵ（Central Processing Unit）プール２、ディスクエリアネットワーク３、ディスクプール４、管理ネットワーク７、管理サーバ１０を有する。

【００１３】

また、ＣＰＵプール２は、複数のノード５ａ～５ｅを有する。ここで、ノード５ａは、それぞれＣＰＵとメモリとを有し、各種アプリケーションプログラムを独立、または、連携して実行可能な装置であり、例えば、ＣＰＵとメモリとを搭載したサーバボードである。なお、ノード５ｂ～５ｅも、ノード５ａと同様の機能を発揮するものとして、以下の説明を省略する。

【００１４】

一方、ディスクプール４は、複数のＳＳＤ６ａ～６ｆを有する。なお、図１に示すように、ディスクプール４は、ＳＳＤ６ａ～６ｆ以外にも複数のＳＳＤを有するものとする。各ＳＳＤ６ａ～６ｆは、ノード５ａ～５ｅが実行するアプリケーションプログラムが使用する半導体記憶装置であり、データの消去回数に基づく余命を有する。例えば、各ＳＳＤ６ａ～６ｆのうち、データの消去が行われていない半導体記憶装置については、余命の指標値である余命指標値が「１００」となり、データの消去が何度も行われ、寿命となったＳＳＤについては、余命指標値が「０」となる。

【００１５】

ディスクエリアネットワーク３は、各ノード５ａ～５ｅと、各ノード５ａ～５ｅがアプリケーションプログラムを実行する際に使用するＳＳＤとを接続するネットワークである。例えば、ディスクエリアネットワーク３は、管理サーバ１０からの指定に従って、ノード５ａと、ディスクプール４が有する１つ以上の任意のＳＳＤとを接続、もしくは接続の解除を行う。

【００１６】

このような構成の元、ノード５ａは、ディスクプール４が有する各ＳＳＤのうち、任意のＳＳＤを用いてアプリケーションプログラムを実行する。例えば、ノード５ａは、ＳＳＤ６ａとＳＳＤ６ｂとを用いて、アプリケーションプログラムを実行する。

【００１７】

管理サーバ１０は、管理ネットワーク７を介して、各ノード５ａ～５ｅと接続されている。そして、管理サーバ１０は、例えば、ノード５ｂが、新たなサービスを提供するため、アプリケーションプログラムを実行する場合は、ＳＳＤ６ｄとノード５ｂとを接続するようディスクエリアネットワーク３に指示する。この結果、ノード５ｂは、ＳＳＤ６ｄを用いて、アプリケーションプログラムを実行し、各種サービスを提供することができる。

【００１８】

このように、情報処理装置１は、複数のノード５ａ～５ｅとディスクプール４が有する各ＳＳＤ６ａ～６ｆとを用いて、任意の数のアプリケーションプログラムを実行し、各種サービスを提供することができる。例えば、情報処理装置１は、ノード５ａ、５ｂをＳＳＤ６ａ、６ｂと組み合わせてＷｅｂサービスを提供させ、ノード５ｃ～５ｅとＳＳＤ６ｃ、６ｄとを組み合わせて大規模データ処理のサービスを提供させることができる。

【００１９】

ここで、ＳＳＤ６ａは、ＮＡＮＤ型のフラッシュメモリや相変化メモリ（ＰＣＭ:Phase Change Memory）等、不揮発性の半導体メモリを有する。ここで、不揮発性の半導体メモリは、データの消去回数やデータの書き込み回数に制限があるデータセルを有する。例えば、１つのデータセルが１ビットのデータを記憶するＳＬＣ（Single Level Cell）方式では、データの消去回数が約１万回に制限される。また、１つのデータセルが複数ビットのデータを記憶するＭＬＣ（Multi Level Cell）方式では、データの消去回数が約１０万回に制限される。

【００２０】

このように、ＳＳＤ６ａが有するメモリセルには、データの消去回数に制限が存在する

ため、ＳＳＤには寿命が存在する。このため、情報処理装置１は、各ＳＳＤ６ａ～６ｆがデータを消去することができる残余回数に基づく余命指標値を収集する。例えば、情報処理装置１は、smartmontoolsを利用してMedia_Wearout_Indicatorの値を余命として収集する。そして、情報処理装置１は、各ノード５ａ～５ｅが利用するＳＳＤの余命が所定の閾値を下回った場合は、利用中のＳＳＤを他のＳＳＤに変更する。すなわち、情報処理装置１は、ディスクプール４が有する各ＳＳＤ６ａ～６ｆについてのウェアレベリングを実行する。このため、情報処理装置１は、寿命を考慮せずにＳＳＤをストレージとして利用することができる。

【００２１】

例えば、図２は、実施例１に係る情報処理装置が実行する処理を説明する図である。例えば、図２中（Ａ）に示すＳＳＤは、ノード５ａ、５ｂがアプリケーションプログラムの実行に利用するＳＳＤである。すなわち、情報処理装置１は、ノード５ａ、５ｂと図２中（Ａ）に示すＳＳＤとを組み合わせて、１つのシステムとして動作させている。また、図２中（Ｂ）に示すＳＳＤは、利用されていないＳＳＤである。

【００２２】

ここで、図２中（Ｃ）に示すＳＳＤの余命が所定の閾値を下回った場合は、情報処理装置１は、図２中（Ｃ）に示すＳＳＤを図２中（Ｄ）に示すＳＳＤと論理的な交換を行う。すなわち、情報処理装置１は、ディスクエリアネットワーク３の接続を変更し、ノード５ａ、５ｂと図２中（Ｃ）に示すＳＳＤとの接続を切断し、ノード５ａ、５ｂと図２中（Ｄ）に示すＳＳＤとを接続する。この結果、ノード５ａ、５ｂは、図２中（Ｅ）に示すＳＳＤを用いて、アプリケーションプログラムを継続して実行する。

【００２３】

また、例えば、ＳＳＤ６ａの余命が短い際に、データの書き込み等、メモリセルにおけるデータの消去を伴うアクセスを行うのは、不具合が生じる可能性があるものの、データの読出し等、データの消去を伴わないアクセスについては、許容してもよい。そこで、情報処理装置１は、各ノード５ａ～５ｅが実行するアプリケーションプログラムが、単位時間当たりに読出したデータ量と書き込んだデータ量との比率を収集する。

【００２４】

そして、情報処理装置１は、収集した比率に基づいて、アプリケーションプログラムの性質を判定する。その後、情報処理装置１は、判定したアプリケーションプログラムの性質に応じて、ＳＳＤを交換するかを判定するための閾値を決定し、ＳＳＤ６ａ～６ｆの余命が決定した閾値よりも短い場合は、ＳＳＤの交換を行う。

【００２５】

例えば、情報処理装置１は、ノード５ａが実行するアプリケーションプログラムがデータの書き込みよりもデータの読出しを頻繁に実行するリードインテンシブな性質である場合は、閾値の値「１０」を決定する。そして、情報処理装置１は、ノード５ａが使用するＳＳＤ６ａの余命が「１０」よりも短い場合は、ノード５ａが利用するＳＳＤ６ａを他のＳＳＤと交換する。

【００２６】

若しくは、情報処理装置１は、収集した比率においてデータの読み出しの比率が多い場合、直接、ＳＳＤを交換するかを判定するための閾値「１０」を決定し、ＳＳＤ６ａ～６ｆの余命が決定した閾値よりも短い場合は、ＳＳＤの交換を行うようにしても良い。

【００２７】

一方、情報処理装置１は、ノード５ａが実行するアプリケーションプログラムがデータの読出しよりもデータの書き込みを頻繁に実行するライトインテンシブな性質である場合は、閾値の値「３０」を決定する。すなわち、情報処理装置１は、アプリケーションプログラムの性質がライトインテンシブである場合は、アプリケーションプログラムの性質がリードインテンシブである場合の閾値よりも、大きな値を閾値とする。そして、情報処理装置１は、ノード５ａが使用するＳＳＤ６ａの余命が「３０」よりも短い場合は、ノード５ａが利用するＳＳＤ６ａを他のＳＳＤと交換する。

【００２８】
若しくは、情報処理装置１は、収集した比率においてデータの書き込みの比率が多い場合、直接、ＳＳＤを交換するかを判定するための閾値「３０」を決定し、ＳＳＤ６ａ～６ｆの余命が決定した閾値よりも短い場合は、ＳＳＤの交換を行うようにしても良い。
【００２９】
また、例えば、ノード５ａがＳＳＤ６ａのみを使用する際に、ＳＳＤ６ａの寿命が尽きてしまうと、ＳＳＤ６ａを他のＳＳＤと交換する間、ノード５ａが処理を実行できなくなる。また、ノード５ａが使用するＳＳＤ６ａを単にＳＳＤ６ｄと交換しただけでは、ノード５ａが使用していたデータがなくなるので、ノード５ａが継続して処理を実行できない。
【００３０】
そこで、情報処理装置１においては、各ノード５ａ～５ｅは、２つのＳＳＤを用いて、ＲＡＩＤ（Redundant Arrays of Inexpensive Disks）１を構成し、データをミラーリングする。例えば、ノード５ａは、ＲＡＩＤ１によりデータがミラーリングされたＳＳＤ６ａとＳＳＤ６ｂとを使用する。そして、情報処理装置１は、ノード５ａが使用するＳＳＤ６ａをＳＳＤ６ｄと交換する場合は、ＳＳＤ６ａとノード５ａとの接続を解除し、ＳＳＤ６ａのデータをＳＳＤ６ｄに複製する。そして、情報処理装置１は、データを複製したＳＳＤ６ｄとノード５ａとを接続する。
【００３１】
また、ノード５ａは、ＳＳＤ６ａとの接続が解除されてからＳＳＤ６ｄが接続されるまでの間、以下の処理を実行する。すなわち、ノード５ａは、データの書き込みを行う場合は、書き込み対象のデータをメモリ上にバッファリングし、データの読出しを行う場合は、接続されているＳＳＤ６ｂからデータの読出しを行う。そして、ノード５ａは、ＳＳＤ６ｄが接続されると、メモリ上にバッファリングしたデータをＳＳＤ６ｂ、およびＳＳＤ６ｄに反映させる。このため、情報処理装置１は、各ノード５ａ～５ｅにアプリケーションプログラムの実行を一時停止させずとも、ＳＳＤの交換を行うことができる。
【００３２】
次に、図３を用いて、管理サーバ１０が有する機能構成について説明する。図３は、実施例１に係る管理サーバの機能構成を説明する図である。なお、図３に示す例では、情報処理装置１が有するノード５ａ～５ｅのうち、ノード５ａ、５ｂを記載し、情報処理装置１が有するＳＳＤのうち、ＳＳＤ６ａ～６ｄを記載した。
【００３３】
図３に示すように、管理サーバ１０は、記憶部１１、通信部１４、更新部１５、判定部１６、決定部１７、選択部１８、設定部１９を有する。また、記憶部１１は、余命管理テーブル１２、ワークロード管理テーブル１３を記憶する。
【００３４】
以下、図４、５を用いて、記憶部１１が記憶する余命管理テーブル１２、ワークロード管理テーブル１３に格納された情報について説明する。まず、図４を用いて、余命管理テーブル１２に格納された情報の一例を説明する。図４は、余命管理テーブルの一例を説明する図である。図４に示すように、余命管理テーブル１２には、各ＳＳＤ６ａ～６ｆの余命が格納されている。詳細には、余命管理テーブル１２には、ＳＳＤ　ＩＤ（Ientifier）、接続先、余命が対応付けて格納されている。ここで、ＳＳＤ　ＩＤとは、各ＳＳＤ６ａ～６ｆを識別するための識別子である。また、接続先とは、対応付けられたＳＳＤ　ＩＤが示すＳＳＤが接続されているノードを示す情報である。また、余命とは、対応付けられたＳＳＤ　ＩＤの余命である。
【００３５】
例えば、図４に示す例では、余命管理テーブル１２は、ＳＳＤ　ＩＤが「ＳＳＤ＃０」のＳＳＤについて、ノード５ａが接続されており、余命が「１００」であることを記憶する。また、余命管理テーブル１２は、ＳＳＤ　ＩＤが「ＳＳＤ＃１」のＳＳＤについて、ノード５ｂが接続されており、余命が「８０」であることを記憶する。また、余命管理テ

ーブル１２は、ＳＳＤ　ＩＤが「ＳＳＤ＃２」のＳＳＤについては、どのノード５ａ～５ｅも接続されておらず、余命が「５０」であることを記憶する。

【００３６】

次に、図５を用いて、ワークロード管理テーブル１３に格納された情報について説明する。図５は、ワークロード管理テーブルの一例を説明する図である。図５に示すように、ワークロード管理テーブル１３には、各ノード５ａ～５ｅが実行するアプリケーションプログラムが、使用中のＳＳＤから読みだしたデータと書き込んだデータとの比率が格納されている。

【００３７】

詳細には、ワークロード管理テーブル１３には、ノードＩＤとプログラム実行フラグと比率とが対応付けて格納されている。ここで、ノードＩＤとは、各ノード５ａ～５ｅを識別する識別子である。またプログラム実行フラグとは、対応付けられたノードＩＤが示すノードがプログラムを実行中か否かを示すフラグである。また、比率とは、対応付けられたノードＩＤが示すノードが、１秒間に読出したデータ量（Ｂｙｔｅ）と書き込んだデータ量（Ｂｙｔｅ）との比率を示す情報である。

【００３８】

例えば、ワークロード管理テーブル１３は、ノード５ａがアプリケーションプログラムを実行しており、アプリケーションプログラムが読出したデータ量と書き込んだデータ量との比率が「５０：５０」である旨を記憶する。また、ワークロード管理テーブル１３は、ノード５ｂがアプリケーションプログラムが読出したデータ量と書き込んだデータ量との比率が「９０：１０」である旨を記憶する。また、ワークロード管理テーブル１３は、ノード５ｃがアプリケーションプログラムを実行していないので、比率が「ｎｕｌｌ」である旨を記憶する。

【００３９】

図３に戻り、通信部１４は、管理ネットワーク７を介して、管理サーバ１０と各ノード５ａ～５ｅとの通信、およびディスクエリアネットワーク３に対する指示を制御する。詳細には、通信部１４は、更新部１５および設定部１９と各ノード５ａ～５ｅとの間の通信を制御する。また、通信部１４は、設定部１９からディスクエリアネットワーク３に対する指示を制御する。

【００４０】

更新部１５は、所定の時間間隔で、記憶部１１が記憶する余命管理テーブル１２とワークロード管理テーブル１３とを更新する。例えば、各ノード５ａ～５ｅは、所定の時間間隔で、smartmontoolsを利用してMedia_Wearout_Indicatorの値を取得し、取得したMedia_Wearout_Indicatorの値を余命指標値とし、ＳＳＤ　ＩＤと対応付けて管理サーバ１０へ送信する。また、各ノード５ａ～５ｅは、読出したデータの量と書き込んだデータの量との比率をＬｉｎｕｘ（登録商標）のｓａｒコマンド等、任意の手法で取得し、取得した比率を管理サーバ１０へ送信する。また、各ノード５ａ～５ｅは、アプリケーションプログラムを起動した場合は、アプリケーションプログラムの起動を管理サーバ１０に通知し、アプリケーションプログラムの実行を終了した場合は、アプリケーションプログラムの終了を管理サーバ１０へ送信する。

【００４１】

更新部１５は、各ノード５ａ～５ｅからＳＳＤ　ＩＤ、余命指標値、比率を受信する。そして、更新部１５は、受信したＳＳＤ　ＩＤと余命指標値とに基づいて、余命管理テーブル１２を更新し、受信した比率を用いて、ワークロード管理テーブル１３を更新する。

【００４２】

例えば、更新部１５は、ＳＳＤ　ＩＤ「ＳＳＤ＃０」と余命「５０」とをノード５ａから受信すると、余命管理テーブル１２からＳＳＤ　ＩＤ「ＳＳＤ＃０」が格納されたエントリを抽出する。そして、更新部１５は、抽出したエントリの接続先を「ノード５ａ」に更新し、余命を「５０」に更新する。また、更新部１５は、ノード５ａから比率「４０：６０」を受信した場合には、ワークロード管理テーブル１３からノードＩＤ「ノード５ａ

」が格納されたエントリを抽出する。そして、更新部１５は、抽出したエントリの比率を「４０：６０」に更新する。

【００４３】

また、更新部１５は、ノード５ａからアプリケーションの起動を通知された場合は、ワークロード管理テーブル１３からノードＩＤ「ノード５ａ」が格納されたエントリを抽出し、抽出したエントリのプログラム実行フラグを「１」に更新する。また、更新部１５は、ノード５ａからアプリケーションの終了を通知された場合は、ワークロード管理テーブル１３からノードＩＤ「ノード５ａ」が格納されたエントリを抽出し、抽出したエントリのプログラム実行フラグを「０」に更新する。

【００４４】

判定部１６は、ワークロード管理テーブル１３に格納された読出しと書き込みの比率に応じて、各ノード５ａ～５ｅが実行するアプリケーションプログラムの性質を判定する。例えば、判定部１６は、ワークロード管理テーブル１３を参照し、各ＳＳＤ６ａ～６ｆに対する読出しと書き込みの比率を取得する。

【００４５】

そして、判定部１６は、読出したデータ量を算出し、算出したデータ量が所定の閾値よりも大きい場合は、アプリケーションプログラムがリードインテンシブであると判定する。また、判定部１６は、算出したデータ量が所定の閾値よりも小さい場合は、アプリケーションプログラムがライトインテンシブであると判定する。なお、判定部１６は、アプリケーションプログラムによる読出しの比率が書き込みの比率よりも多い場合は、アプリケーションプログラムがリードインテンシブであると判定してもよい。

【００４６】

例えば、ワークロード管理テーブル１３に図５に示す情報が格納されている場合には、判定部１６は、以下の処理を実行する。例えば、判定部１６は、ノード５ｂが実行するアプリケーションプログラムの読出し比率が書き込み比率よりも大きいので、ノード５ｂが実行するアプリケーションプログラムがリードインテンシブであると判定する。また、判定部１６は、ノード５ｃが実行するアプリケーションプログラムの書き込み比率が読出し比率よりも大きいので、ノード５ｃが実行するアプリケーションプログラムがライトインテンシブであると判定する。

【００４７】

ここで、ノード５ａが実行するアプリケーションプログラムについては、読出しと書き込みの比率が同じである。このような場合には、判定部１６は、システムの安定性を考慮し、ノード５ａが実行するアプリケーションプログラムがライトインテンシブであると判定してもよい。なお、判定部１６は、アプリケーションプログラムの性質を判定した場合は、判定結果を決定部１７に通知する。

【００４８】

なお、判定部１６は、アプリケーションプログラムの性質を判定せず、各ＳＳＤ６ａ～６ｆに対する読出しと書込みのどちらかが多いかを判定し、判定結果を決定部１７に通知してもよい。

【００４９】

決定部１７は、各ノード５ａ～５ｅが実行するアプリケーションプログラムの性質に基づいて、ＳＳＤの交換基準となる閾値を決定する。例えば、決定部１７は、ノード５ａが実行するアプリケーションプログラムがライトインテンシブである旨の通知を判定部１６から受信する。このような場合は、決定部１７は、ノード５ａが使用するＳＳＤの交換基準となる閾値を「３０」とし、ノード５ａが使用するＳＳＤの交換基準となる閾値が「３０」である旨を選択部１８に通知する。

【００５０】

一方、決定部１７は、ノード５ｃが実行するアプリケーションプログラムがリードインテンシブである旨の通知を判定部１６から受信する。このような場合は、決定部１７は、ノード５ｃが使用するＳＳＤの交換基準となる閾値を「１０」とし、ノード５ｃが使用す

るＳＳＤの交換基準となる閾値が「１０」である旨を選択部１８に通知する。すなわち、決定部１７は、各ノード５ａ～５ｅが実行するアプリケーションプログラムがリードインテンシブである場合は、アプリケーションプログラムがライトインテンシブである際よりも小さい値の閾値を決定する。

【００５１】

なお、上述した例では、決定部１７は、各ノード５ａ～５ｅが実行するアプリケーションプログラムがリードインテンシブである場合は、閾値を「１０」とし、ライトインテンシブである場合は、閾値を「３０」とした。しかし、実施例はこれに限定されるものではない。例えば、決定部１７は、ワークロード管理テーブル１３に格納された比率に応じて、閾値を決定してもよい。例えば、決定部１７は、ライトインテンシブなアプリケーションプログラムについて、ワークロード管理テーブル１３に格納された書き込みの比率が所定の値よりも大きい場合は、閾値を「３０」よりも大きい値にしてもよい。

【００５２】

なお、決定部１７は、各ＳＳＤ６ａ～６ｆに対する読出しと書込みのどちらか多いかを示す判定結果を受信した場合は、受信した判定結果に基づいて、ＳＳＤ６ａ～６ｆの交換基準となる閾値を決定しても良い。例えば、決定部１７は、ＳＳＤ６ａに対して読出しが多い旨の判定結果を受信した場合は、ＳＳＤ６ａの交換基準となる閾値を「１０」に決定する。また、決定部１７は、ＳＳＤ６ａに対して書込みが多い旨の判定結果を受信した場合は、ＳＳＤ６ａの交換基準となる閾値を「３０」に決定する。

【００５３】

選択部１８は、アプリケーションプログラムの性質に応じて決定した閾値に応じて、各ノード５ａ～５ｅが使用するＳＳＤを交換するか否かを判定する。例えば、選択部１８は、決定部１７からノード５ａが使用するＳＳＤの交換基準となる閾値が「３０」である旨を受信する。すると、選択部１８は、余命管理テーブル１２を参照し、ノード５ａが使用する１つ以上のＳＳＤを識別する。そして、選択部１８は、識別したＳＳＤの余命指標値が閾値「３０」よりも小さいか否かを判定し、余命指標値が閾値「３０」よりも小さい場合は、ＳＳＤを交換すると判定する。

【００５４】

また、選択部１８は、ＳＳＤを交換すると判定した場合は、以下の処理を実行する。例えば、選択部１８は、余命管理テーブル１２を参照し、接続先が未接続のＳＳＤを抽出する。そして、選択部１８は、ノード５ａが実行するアプリケーションがリードインテンシブである場合は、抽出したＳＳＤのうち、余命指標値が「１０」より大きく「３０」以下であるＳＳＤを検索する。その後、選択部１８は、余命指標値が「１０」より大きく「３０」以下であるＳＳＤを検出した場合は、ノード５ａが使用するＳＳＤのうち交換対象となるＳＳＤのＳＳＤ　ＩＤと余命管理テーブル１２から検出したＳＳＤのＳＳＤ　ＩＤとを設定部１９に通知する。

【００５５】

また、選択部１８は、余命指標値が「１０」より大きく「３０」以下であるＳＳＤを検出できなかった場合は、抽出したＳＳＤのうち、余命指標値が「３０」より大きなＳＳＤを検索する。そして、選択部１８は、余命指標値が「３０」より大きなＳＳＤを検出した場合は、ノード５ａが使用するＳＳＤのうち交換対象となるＳＳＤのＳＳＤ　ＩＤと余命管理テーブル１２から検出したＳＳＤのＳＳＤ　ＩＤとを設定部１９に通知する。なお、選択部１８は、余命指標値が「３０」より大きなＳＳＤを検出できなかった場合は、交換可能なＳＳＤがない旨を管理者等に通知する。

【００５６】

一方、選択部１８は、ノード５ａが実行するアプリケーションがライトインテンシブである場合は、余命指標値が「３０」より大きなＳＳＤを検索する。そして、選択部１８は、余命指標値が「３０」より大きなＳＳＤを検出した場合は、ノード５ａが使用するＳＳＤのうち交換対象となるＳＳＤのＳＳＤ　ＩＤと余命管理テーブル１２から検出したＳＳＤのＳＳＤ　ＩＤとを設定部１９に通知する。

【００５７】
なお、選択部１８は、判定部１６がアプリケーションの性質を判定しない場合であっても、同様の処理を行うことで、交換するＳＳＤを選択できる。例えば、選択部１８は、ＳＳＤ６ａについて閾値が「１０」であり、ＳＳＤ６ａの余命指標値が「１０」よりも小さい場合は、ＳＳＤ６ａと交換するＳＳＤとして、余命指標値が「１０」より大きく「３０」以下となるＳＳＤを選択する。また、選択部１８は、ＳＳＤ６ａについて閾値が「３０」であり、ＳＳＤ６ａの余命指標値が「３０」より小さい場合は、ＳＳＤ６ａと交換するＳＳＤとして、余命指標値が「３０」より大きいＳＳＤを選択する。
【００５８】
設定部１９は、各ノード５ａ～５ｅが使用するＳＳＤの設定変更を行う。例えば、設定部１９は、選択部１８から、ＳＳＤ６ａのＳＳＤ　ＩＤとＳＳＤ６ｄのＳＳＤ　ＩＤとを受信する。すると、設定部１９は、ワークロード管理テーブル１３を参照し、ＳＳＤ６ａを使用中のノード５ａを識別する。そして、設定部１９は、識別したノード５ａに対してＳＳＤ６ａの接続解除手続きを依頼する。このような場合は、ノード５ａは、ＳＳＤ６ａの接続解除手続きを実行し、ＳＳＤ６ａの接続を解除する。
【００５９】
また、設定部１９は、ノード５ａとＳＳＤ６ａとの接続を解除するようディスクエリアネットワーク３に指示する。この結果、ノード５ａとＳＳＤ６ａとの接続は解除される。また、設定部１９は、ワークロード管理テーブル１３を参照し、アプリケーションプログラムを実行していないノードを識別する。例えば、設定部１９は、ワークロード管理テーブル１３に図５に示す情報が格納されている場合は、ノード５ｃを識別する。そして、設定部１９は、ディスクエリアネットワーク３にノード５ｃとＳＳＤ６ａ、６ｄとを接続するよう指示する。そして、設定部１９は、ノード５ｃにＳＳＤ６ａのデータをＳＳＤ６ｄに複製するよう依頼する。
【００６０】
すると、ノード５ｃは、ＳＳＤ６ａのデータをＳＳＤ６ｄに複製し、複製完了を設定部１９に通知する。すると、設定部１９は、ノード５ｃとＳＳＤ６ｄとの接続を解除し、ノード５ａとＳＳＤ６ｄとを接続するようにディスクエリアネットワーク３に指示する。そして、設定部１９は、ノード５ａにＳＳＤ６ｄの接続手続きを依頼する。この結果、ノード５ａは、ＳＳＤ６ａと同じデータを記憶するＳＳＤ６ｄを用いて、アプリケーションプログラムの実行を継続する。
【００６１】
なお、ノード５ａは、ＳＳＤ６ａとの接続が解除されてからＳＳＤ６ｄが接続されるまでの間、ＳＳＤ６ｂからデータの読出しを行う。また、ノード５ａは、データの書き込みを行う場合は、メモリ上に書き込み対象のデータをバッファし、ＳＳＤ６ｄが接続された後に、バッファしたデータをＳＳＤ６ｂ、６ｄに反映させる。なお、ノード５ａは、バッファオーバーフローが発生した場合は、ＳＳＤ６ｂに対してデータの書き込みを行い、ＳＳＤ６ｄが接続された後に、ＳＳＤ６ｂからＳＳＤ６ｄへデータを複製してもよい。
【００６２】
次に、図６を用いて、実施例１に係る情報処理装置１が実行する処理の流れについて説明する。図６は、実施例１に係る情報処理装置が実行する処理の流れを説明するフローチャートである。なお、以下の説明では、ノード５ａがＳＳＤ６ａ、６ｂを使用しており、ＳＳＤ６ａをＳＳＤ６ｄに交換する処理の流れについて説明する。また、以下の説明では、ノード５ｃがアプリケーションプログラムを実行していないものとする。
【００６３】
例えば、ノード５ａは、ＳＳＤへのアクセスを監視し、読出しと書き込みの比率であるＲｅａｄ／Ｗｒｉｔｅ比率を管理サーバ１０に通知する（ステップＳ１０１）。また、ノード５ａは、使用中のＳＳＤ６ａ、６ｂの余命を確認し、余命指標値を管理サーバ１０へ送信する（ステップＳ１０２）。すると、管理サーバ１０は、受信したＲｅａｄ／Ｗｒｉｔｅ比率、および余命指標値に従って、ワークロード管理テーブル１３と余命管理テーブ

ル１２とを更新する（ステップＳ１０３）。
【００６４】
　また、管理サーバ１０は、ワークロード管理テーブル１３に格納されたＲｅａｄ／Ｗｒｉｔｅ比率に基づいて、ノード５ａが実行するアプリケーションプログラムがリードインテンシブであるかライトインテンシブであるかを判定する（ステップＳ１０４）。そして、管理サーバ１０は、ノード５ａに接続されているＳＳＤ６ａ、６ｂの余命が、ノード５ａが実行するアプリケーションプログラムの性質に応じた閾値を超えていないか確認する（ステップＳ１０５）。
【００６５】
　また、管理サーバ１０は、アプリケーションプログラムを実行中の全てのノードについて、アプリケーションプログラムの性質に応じた閾値とＳＳＤの余命とのチェックを行ったか判定する（ステップＳ１０６）。そして、管理サーバ１０は、アプリケーションプログラムを実行中の全てのノードについて、アプリケーションプログラムの性質に応じた閾値とＳＳＤの余命とのチェックを行っていない場合は（ステップＳ１０６Ｎｏ）、以下の処理を実行する。すなわち、管理サーバ１０は、アプリケーションプログラムを実行中の各ノードについて、ステップＳ１０５を実行する。
【００６６】
　また、管理サーバ１０は、アプリケーションプログラムを実行中の全てのノードについて、アプリケーションプログラムの性質に応じた閾値とＳＳＤの余命とのチェックを行った場合は（ステップＳ１０６Ｙｅｓ）、以下の処理を実行する。すなわち、管理サーバ１０は、余命指標値が閾値を超えたＳＳＤ、すなわち余命指標値が閾値よりも小さいＳＳＤがあるか否かを判定する（ステップＳ１０７）。また、管理サーバ１０は、余命指標値が閾値よりも小さいＳＳＤがある場合は（ステップＳ１０７Ｙｅｓ）、使用されていないＳＳＤにアプリケーションプログラムの性質に応じた閾値を満たすＳＳＤが存在するか否かを判定する（ステップＳ１０８）。
【００６７】
　また、管理サーバ１０は、使用されていないＳＳＤにアプリケーションプログラムの性質に応じた閾値を満たすＳＳＤ、例えばＳＳＤ６ｄが存在する場合は（ステップＳ１０８Ｙｅｓ）、以下の処理を実行する。すなわち、管理サーバ１０は、余命が閾値を超えたＳＳＤを使用するノード、例えばノード５ａに対して、余命が閾値を超えたＳＳＤ、例えばＳＳＤ６ａの接続解除手続きを依頼する（ステップＳ１０９）。すると、ノード５ａは、ＳＳＤ６ａの接続解除手続きを実行し、手続き完了を管理サーバ１０に通知する（ステップＳ１１０）。
【００６８】
　また、管理サーバ１０は、ノード５ａからＳＳＤ６ａの接続解除手続きの完了通知を受信すると、以下のＳＳＤ交換処理を実行する（ステップＳ１１１）。すなわち、管理サーバ１０は、ディスクエリアネットワーク３に指示し、ノード５ａからＳＳＤ６ａの接続を解除し、ノード５ｃとＳＳＤ６ａ、６ｄとを接続し、接続完了をノード５ｂに通知する。すると、ノード５ｃは、ＳＳＤ６ａからＳＳＤ６ｄへデータのコピーを実行し、実行完了を管理サーバ１０に通知する（ステップＳ１１２）。
【００６９】
　すると、管理サーバ１０は、ノード５ｃからＳＳＤ６ｄの接続を解除し、ＳＳＤ６ｄをノード５ａに接続してノード５ａにＳＳＤ６ｄを接続した旨を通知する（ステップＳ１１３）。すると、ノード５ａは、ＳＳＤ６ｄの接続手続きを実行する（ステップＳ１１４）。また、管理サーバ１０は、ノード５ｃからＳＳＤ６ａの接続を解除する（ステップＳ１１５）。
【００７０】
　また、管理サーバ１０は、閾値を超えたＳＳＤで未交換のＳＳＤが存在するか判定し（ステップＳ１１６）、閾値を超えたＳＳＤで未交換のＳＳＤが存在しない場合は（ステップＳ１１６Ｎｏ）、処理を終了する。一方、管理サーバ１０は、閾値を超えたＳＳＤで未

交換のＳＳＤが存在する場合は（ステップＳ１１６Ｙｅｓ）、ステップＳ１０８を実行する。なお、管理サーバ１０は、使用していないＳＳＤでアプリケーションプログラムの性質に応じた閾値を満たすＳＳＤが存在しない場合は（ステップＳ１０８Ｎｏ）、そのまま処理を終了する。

【００７１】

次に、図７、図８を用いて、図６中ステップＳ１０８にて使用されていないＳＳＤからアプリケーションプログラムの性質に応じた閾値を満たすＳＳＤがあるか否かを判定する処理の流れについて説明する。まず、図７を用いて、アプリケーションプログラムの性質がリードインテンシブである際に、選択部１８が実行する処理の流れについて説明する。

【００７２】

図７は、選択部が実行する処理の流れを説明するための第１のフローチャートである。例えば、選択部１８は、使用されていないＳＳＤに、余命指標値が「１０」よりも大きく、「３０」以下のＳＳＤが存在するか否かを判定する（ステップＳ２０１）。ここで、選択部１８は、余命指標値が「１０」よりも大きく、「３０」以下のＳＳＤが存在しない場合は（ステップＳ２０１Ｎｏ）、余命指標値が「３０」より大きく「１００」以下のＳＳＤが存在するか否かを判定する（ステップＳ２０２）。

【００７３】

そして、選択部１８は、余命指標値が「３０」より大きく「１００」以下のＳＳＤが存在する場合は（ステップＳ２０２Ｙｅｓ）、閾値を満たすＳＳＤを選択し（ステップＳ２０３）、処理を終了する。また、選択部１８は、使用されていないＳＳＤに、余命指標値が「１０」よりも大きく、「３０」以下のＳＳＤが存在する場合は（ステップＳ２０１Ｙｅｓ）、閾値を満たすＳＳＤを選択し（ステップＳ２０３）、処理を終了する。また、選択部１８は、余命指標値が「３０」より大きく「１００」以下のＳＳＤが存在しない場合は（ステップＳ２０２Ｎｏ）、ＳＳＤの選択を行わずに、処理を終了する。

【００７４】

次に、図８を用いて、アプリケーションプログラムの性質がライトインテンシブである際に、選択部１８が実行する処理の流れについて説明する。図８は、選択部が実行する処理の流れを説明するための第２のフローチャートである。例えば、選択部１８は、アプリケーションプログラムの性質がライトインテンシブである場合は、使用されていないＳＳＤに余命指標値が「３０」より大きく「１００」以下のＳＳＤが存在するか判定する（ステップＳ３０１）。

【００７５】

そして、選択部１８は、使用されていないＳＳＤに余命指標値が「３０」より大きく「１００」以下のＳＳＤが存在する場合は（ステップＳ３０１Ｙｅｓ）、閾値を満たすＳＳＤを選択し（ステップＳ３０２）、処理を終了する。一方、選択部１８は、使用されていないＳＳＤに余命指標値が「３０」より大きく「１００」以下のＳＳＤが存在しない場合は（ステップＳ３０１Ｎｏ）、ＳＳＤを選択せずに処理を終了する。

【００７６】

[情報処理装置１の効果]

上述したように、情報処理装置１は、ＳＳＤ６ａと、例えばＳＳＤ６ａの予備であるＳＳＤ６ｆとを有する。また、情報処理装置１は、ＳＳＤ６ａについて、余命指標値と、ＳＳＤ６ａに対する書込みと読み出しとの読み書きの比率を収集する。そして、情報処理装置１は、収集した比率に基づいて、ＳＳＤ６ａの交換基準となる閾値を決定する。その後、情報処理装置１は、収集した余命指標値が決定した閾値よりも短い場合は、ＳＳＤ６ａをＳＳＤ６ｆと交換する。このため、情報処理装置１は、ノード５ａにＳＳＤ６ａ～６ｆをストレージ装置として継続して利用させることができる。

【００７７】

また、情報処理装置１は、ＳＳＤ６ａに対する読み書きの比率に応じた閾値を用いて、ＳＳＤ６ａの交換を判定するので、各ＳＳＤ６ａ～６ｆの余命を使いきることができる。この結果、情報処理装置１は、各ＳＳＤ６ａ～６ｆをストレージとして利用した際の寿命

を最大限利用することができる。

【００７８】

なお、情報処理装置１は、収集した比率に応じて、ノード５ａが実行するアプリケーションプログラムの性質を判定し、判定したアプリケーションプログラムの性質に応じて、ノード５ａが使用するＳＳＤ６ａの交換基準となる閾値を決定してもよい。その後、情報処理装置１は、収集した余命指標値が、決定した閾値よりも短い場合は、ＳＳＤ６ａを他のＳＳＤと交換する。このような場合にも、情報処理装置１は、ノード５ａにＳＳＤ６ａ～６ｆをストレージ装置として継続して利用させることができる。

【００７９】

また、情報処理装置１は、ノード５ａが実行するアプリケーションプログラムの性質に応じた閾値を用いて、ＳＳＤの交換を行ってもよい。このような場合にも、情報処理装置１は、各ＳＳＤ６ａ～６ｆの余命を使いきり、各ＳＳＤ６ａ～６ｆをストレージとして利用した際の寿命を最大限利用することができる。

【００８０】

また、情報処理装置１は、収集した比率においてデータの書き込みの比率が多い場合は、データの読み出しの比率が多い場合に設定する第１の閾値よりも値が大きい第２の閾値をＳＳＤ６ａの交換基準となる閾値とする。このため、情報処理装置１は、書き込みの頻度が多いアプリケーションプログラムを実行するノードに対して余命が長いＳＳＤを割り当て、読出しの頻度が多いアプリケーションプログラムを実行するノードに対して余命が短いＳＳＤを割り当てる。この結果、情報処理装置１は、各ＳＳＤ６ａ～６ｆの寿命を最大限利用することができる。

【００８１】

なお、情報処理装置１は、ノード５ａが実行するアプリケーションプログラムの性質がリードインテンシブであるかライトインテンシブであるかを判定してもよい。そして、情報処理装置１は、アプリケーションプログラムがライトインテンシブである場合は、リードインテンシブである際に用いる閾値よりも値が大きい閾値を用いることとしてもよい。

【００８２】

また、情報処理装置１は、使用されていないＳＳＤ、すなわち予備のＳＳＤを複数有し、ＳＳＤ６ａを交換する場合は、使用されていないＳＳＤのうち、以下の条件を満たすＳＳＤを選択する。すなわち、情報処理装置１は、ＳＳＤ６ａから収集した比率についてデータ読み出しの比率が多い場合、予備のＳＳＤのうち、余命指標値が第１の閾値よりも大きく、かつ、余命指標値が第２の閾値よりも小さいＳＳＤを選択する。そして、情報処理装置１は、ＳＳＤ６ａを選択したＳＳＤと交換する。このため、情報処理装置１は、使用されていないＳＳＤのうち、余命が短いＳＳＤを優先して使用することができる。

【００８３】

なお、情報処理装置１は、ノード５ａが実行するアプリケーションプログラムの性質がリードインテンシブである場合は、使用されていないＳＳＤのうち、以下の条件を満たすＳＳＤを選択してもよい。すなわち、情報処理装置１は、余命指標値が、アプリケーションプログラムの性質がリードインテンシブである際の閾値よりも大きく、かつ、アプリケーションプログラムの性質がライトインテンシブである際の閾値よりも短いＳＳＤを選択してもよい。

【００８４】

また、情報処理装置１は、ＳＳＤ６ａから収集した比率についてデータ書込みの比率が多い場合、予備のＳＳＤのうち、余命指標値が第２の閾値よりも大きいＳＳＤを選択する。そして、情報処理装置１は、ＳＳＤ６ａを選択したＳＳＤと交換する。このため、情報処理装置１は、書き込みの頻度が多いアプリケーションプログラムを実行するノードに対し、余命に余裕があるＳＳＤを割り当てることができる。この結果、情報処理装置１は、ＳＳＤ６ａ～６ｆの余命が尽き、各ノード５ａ～５ｅが使用するＳＳＤが急に使用できなくなるといった事態を防ぐことができるので、アプリケーションプログラムを安定して実行することができる。

【００８５】
なお、情報処理装置１は、ノード５ａが実行するアプリケーションプログラムの性質がライトインテンシブである場合は、使用されていないＳＳＤのうち、以下の条件を満たすＳＳＤを選択してもよい。すなわち、情報処理装置１は、余命指標値が、アプリケーションプログラムの性質がライトインテンシブである際の閾値よりも大きいＳＳＤを選択する。そして、情報処理装置１は、ノード５ａが使用するＳＳＤを選択したＳＳＤと交換してもよい。
【００８６】
また、情報処理装置１は、単位時間あたりにＳＳＤ６ａから読み出されたデータ量が所定の閾値よりも多い場合は、データの読み出しの比率が多いと判定する。このため、情報処理装置１は、多くのデータを読出すアプリケーションプログラムに対し、余命指標値が小さいＳＳＤを割り当てることができる。
【００８７】
なお、情報処理装置１は、ノード５ａが実行するアプリケーションプログラムが読出したデータ量が所定の閾値よりも多い場合は、ノード５ａが実行するアプリケーションプログラムがリードインテンシブであると判定する。そして、情報処理装置１は、リードインテンシブなアプリケーションプログラムに余命指標値が小さいＳＳＤを割当ててもよい。
【００８８】
また、情報処理装置１は、例えば、ノード５ａがＳＳＤ６ａとＳＳＤ６ｂとをＲＡＩＤ１によりミラーリングした際に、ＳＳＤ６ａをＳＳＤ６ｄと交換する場合は、以下の処理を実行する。すなわち、情報処理装置１は、ＳＳＤ６ａをＳＳＤ６ｄとを交換するまでの間、ノード５ａからＳＳＤ６ａ、６ｂに書き込もうとしたデータをメモリ上に保持する。そして、情報処理装置１は、ＳＳＤ６ａのデータをＳＳＤ６ｄに複製し、メモリ上に保持したデータをＳＳＤ６ｂおよび６ｄに複製する。このため、情報処理装置１は、ノード５ａが実行するアプリケーションプログラムを動作させたままでＳＳＤの交換を行うことができる。
【００８９】
このように、情報処理装置１は、各ＳＳＤ６ａ～６ｆの余命が尽きるまで、各ＳＳＤ６ａ～６ｆを交換するか否かを判定する。ここで、情報処理装置１には、複数のノード５ａ～５ｅが設置され、各ノード５ａ～５ｅが実行するアプリケーションには、それぞれ異なる性質を有する。このため、各ノード５ａ～５ｅがＲＡＩＤ１によるミラーリングを構築した場合にも、各ＳＳＤ６ａ～６ｆの余命に偏りが発生することとなる。この結果、情報処理装置１は、ノード５ａがＲＡＩＤ１によりミラーリングされたＳＳＤを用いる場合であっても、各ＳＳＤ６ａ～６ｆを長期に渡り効率よく利用することができる。
【実施例２】
【００９０】
これまで本発明の実施例について説明したが実施例は、上述した実施例以外にも様々な異なる形態にて実施されてよいものである。そこで、以下では実施例２として本発明に含まれる他の実施例を説明する。
【００９１】
（１）他のＲＡＩＤについて
上述した情報処理装置１では、ノード５ａがＲＡＩＤ１により、ＳＳＤ６ａ、６ｂのデータをミラーリングしていた。しかし、実施例はこれに限定されるものではない。例えば、ノード５ａは、５台のＳＳＤに対し、データとデータを復元するためのパリティとを分散して格納するＲＡＩＤ５を構築してもよい。また、ノード５ａは、６台のＳＳＤに対し、データとデータを復元するためのパリティとを分散して格納するＲＡＩＤ６を構築してもよい。
【００９２】
このように、ノード５ａがＲＡＩＤ５、またはＲＡＩＤ６を構築中には、情報処理装置１は、交換対象となるＳＳＤが記憶するデータの複製を行わなくともよい。例えば、情報

処理装置１は、ノード５ａがＳＳＤ６ａ～６ｅを用いてＲＡＩＤ５を構築中に、ＳＳＤ６ａをＳＳＤ６ｆと交換する場合は、ＳＳＤ６ｂ～６ｅからＳＳＤ６ｆが記憶すべきデータを復元することができる。

【００９３】

このため、情報処理装置１は、ノード５ａが使用中のＳＳＤ６ａをＳＳＤ６ｆと交換する場合は、ＳＳＤ６ａとノード５ａとの接続を解除する。そして、情報処理装置１は、ＳＳＤ６ａからＳＳＤ６ｆにデータを複製せずにＳＳＤ６ｆとノード５ｂとの接続を行う。この結果、ノード５ａは、ＲＡＩＤ５の機能により、ＳＳＤ６ｂ～６ｅからＳＳＤ６ｆが記憶すべきデータを復元する。また、ＲＡＩＤ５の機能により、ノード５ａは、ＳＳＤ６ａとＳＳＤ６ｆとを交換する間、データの読出しを継続することができ、データの書き込みもバッファする必要はない。このため、ノード５ａは、ＳＳＤ６ａとＳＳＤ６ｆとを交換する間、アプリケーションプログラムを継続して実行することができる。

【００９４】

次に、図９を用いて、ＲＡＩＤ５を構築中のノード５ａに対して情報処理装置１が実行する処理の流れについて説明する。図９は、実施例２に係る情報処理装置が実行する処理の流れを説明するフローチャートである。なお、図９に示す処理のうちステップＳ４０１～Ｓ４０８については、図６に示すステップＳ１０１～Ｓ１０８と同様の処理であるものとして、説明を省略する。また、以下の説明では、ノード５ａがＳＳＤ６ａ～６ｅを用いてＲＡＩＤ５を構築しており、ノード５ａが実行するアプリケーションプログラムの性質に応じた閾値を満たすＳＳＤ６ｆとＳＳＤ６ａとを交換する処理の流れについて説明する。

【００９５】

例えば、管理サーバ１０は、ノード５ａからＳＳＤ６ａの接続を解除し、ＳＳＤ６ｆをノード５ａに接続し、余命管理テーブル１２を更新してノード５ａにＳＳＤ６ｆを接続した旨を通知する（ステップＳ４０９）。すると、ノード５ａは、ＳＳＤ６ｂ～６ｅを用いてＳＳＤ６ｆが記憶すべきデータを復元し、ＳＳＤ６ｂ～６ｆを用いてＲＡＩＤ５を再構築する。そして、ノード５ａは、ＲＡＩＤ５の再構築が終了した旨を管理サーバ１０に通知する（ステップＳ４１０）。また、管理サーバ１０は、閾値を超えたＳＳＤで未交換のＳＳＤが存在するか判定し（ステップＳ４１１）、閾値を超えたＳＳＤで未交換のＳＳＤが存在しない場合は（ステップＳ４１１Ｎｏ）、処理を終了する。一方、管理サーバ１０は、閾値を超えたＳＳＤで未交換のＳＳＤが存在する場合は（ステップＳ４１１Ｙｅｓ）、ステップＳ４０８を実行する。

【００９６】

このように、情報処理装置１は、ノード５ａが使用するＳＳＤ６ａ～６ｅにデータとデータを復元するためのパリティとが格納されている場合に、ＳＳＤ６ａ～６ｅのいずれかを交換する時は、以下の処理を実行する。例えば、情報処理装置１は、ＳＳＤ６ａとＳＳＤ６ｆとを交換し、ＳＳＤ６ｂ～６ｅが記憶するデータとパリティとを用いて、ＳＳＤ６ｆが記憶するデータを復元する。このため、情報処理装置１は、交換対象となるＳＳＤのデータを複製せずとも、ノード５ａがアプリケーションプログラムを実行したままで、ＳＳＤの交換を行うことができる。

【００９７】

（２）情報処理装置１の構成について

上述した実施例１では、ディスクエリアネットワーク３を介して、ＣＰＵプール２が有する各ノード５ａ～５ｅに対し、ディスクプール４が有する各ＳＳＤ６ａ～６ｅのうち、任意のＳＳＤを接続する例について説明した。しかし、実施例はこれに限定されるものではない。例えば、情報処理装置１は、ディスクエリアネットワーク３の代わりに、ＳＡＮ（Storage Area Netowrk）を用いて、各ノード５ａ～５ｅと各ＳＳＤ６ａ～６ｅとを接続してもよい。また、情報処理装置１は、ＳＳＤだけではなく、ＨＤＤ（Hard Disk Drive）等任意の装置をストレージとしてディスクプール４に含めてもよい。

【００９８】

（３）アプリケーションプログラムの性質について
上述した実施例１では、管理サーバ１０は、アプリケーションプログラムによる読出しと書き込みの比率に基づいて、アプリケーションプログラムの性質を判定した。しかし、実施例はこれに限定されるものではない。例えば、管理サーバ１０は、アプリケーションプログラムによる読出しと書き込みの比率だけではなく、例えば、平均的な書き込みデータ量を考慮に加えてアプリケーションプログラムの性質を判定してもよい。
【００９９】
また、管理サーバ１０は、アプリケーションプログラムを実行するノードの数を考慮に加えてもよい。また、管理サーバ１０は、必ずしも、アプリケーションプログラムの性質を判定する必要はない。すなわち、管理サーバ１０は、各ＳＳＤ６ａ～６ｅに対する読み出しと書込みの比率に応じて、交換の指標となる閾値を直接設定し、設定した閾値と余命指標値とに応じて、各ＳＳＤ６ａ～６ｅの交換を行っても良い。
【０１００】
（４）ＲＡＩＤについて
上述した実施例１では、ノード５ａがＲＡＩＤ１を構築する例について説明し、実施例２では、ノード５ａがＲＡＩＤ６を構築する例について説明した。しかしながら、実施例はこれに限定されるものではない。例えば、情報処理装置１は、ノード５ａがＲＡＩＤ１を構築し、ノード５ｂがＲＡＩＤ５、またはＲＡＩＤ６を構築中には、ノード５ａに対して図６に示した処理を実行し、ノード５ｂに対して図９に示した処理を実行してもよい。
【０１０１】
（５）収集する情報について
上述した実施例１では、管理サーバ１０は、単位時間当たりに読出したデータ量と書き込んだデータ量との比率を収集した。しかし、実施例はこれに限定されるものではなく、例えば、管理サーバ１０は、単位時間当たりにデータを読出した回数とデータを書き込んだ回数とを各ノード５ａ～５ｅから収集してもよい。また、管理サーバ１０は、データを読出した回数が所定の閾値よりも多い場合は、リードインテンシブであると判断しても良い。すなわち、管理サーバ１０は、データを読出した回数とデータを書き込んだ回数との比率に応じて、アプリケーションプログラムがリードインテンシブであるかライトインテンシブであるかを判断して良い。
【０１０２】
このように、情報処理装置１は、ノード５ａが実行するアプリケーションプログラムがＳＳＤ６ａからデータを読出した回数が所定の閾値よりも多い場合は、データの書き込みの比率が多いと判定する。例えば、情報処理装置１は、ノード５ａが実行するアプリケーションプログラムがリードインテンシブであると判定する。このため、情報処理装置１は、データを何度も読出すアプリケーションプログラムに対し、余命指標値が小さいＳＳＤを割り当てることができる。
【０１０３】
（６）プログラム
ところで、実施例１に係る管理サーバ１０は、ハードウェアを利用して各種の処理を実現する場合を説明した。しかし、実施例はこれに限定されるものではなく、あらかじめ用意されたプログラムをコンピュータで実行することによって実現するようにしてもよい。そこで、以下では、図１０を用いて、実施例１に示した管理サーバ１０と同様の機能を有するプログラムを実行するコンピュータの一例を説明する。図１０は、制御プログラムを実行するコンピュータを説明する図である。
【０１０４】
図１０に例示されたコンピュータ１００は、ＲＯＭ（Read Only Memory）１１０、ＨＤＤ（Hard Disk Drive）１２０、ＲＡＭ（Random Access Memory）１３０、ＣＰＵ１４０がバス１６０で接続される。また、図１０に例示されたコンピュータ１００は、パケットを送受信するためのＩ／Ｏ（Input Output）１５０を有する。
【０１０５】

ＨＤＤ１２０は、図３に示した記憶部１１が記憶する余命管理テーブル１２と同様の情報である余命管理テーブル１２１、および、ワークロード管理テーブル１３と同様の情報であるワークロード管理テーブル１２２を記憶する。また、ＲＡＭ１３０には、制御プログラム１３１があらかじめ保持される。ＣＰＵ１４０が制御プログラム１３１をＲＡＭ１３０から読出して実行することによって、図１０に示す例では、制御プログラム１３１は、制御プロセス１４１として機能するようになる。なお、制御プロセス１４１は、図３に示した更新部１５、判定部１６、決定部１７、選択部１８、設定部１９と同様の機能を発揮する。

【０１０６】

なお、本実施例で説明した制御プログラムは、あらかじめ用意されたプログラムをパーソナルコンピュータやワークステーションなどのコンピュータで実行することによって実現することができる。このプログラムは、インターネットなどのネットワークを介して配布することができる。また、このプログラムは、ハードディスク、フレキシブルディスク（ＦＤ）、ＣＤ－ＲＯＭ（Compact Disc Read Only Memory）、ＭＯ（Magneto Optical Disc）、ＤＶＤ（Digital Versatile Disc）などのコンピュータで読取可能な記録媒体に記録される。また、このプログラムは、コンピュータによって記録媒体から読出されることによって実行することもできる。

【０１０７】

以上の各実施例を含む実施形態に関し、さらに以下の付記を開示する。

【０１０８】

（付記１）データが読み書きされる半導体記憶装置と、前記半導体記憶装置の予備である予備半導体記憶装置とを有する情報処理装置において、

前記半導体記憶装置について、書き込まれたデータを消去できる残余回数に基づいた余命指標値、および、前記半導体記憶装置に対する読出しと書き込みとの読み書きの比率を収集する収集部と、

前記収集部が収集した前記比率に基づいて、前記半導体記憶装置の交換基準となる閾値を決定する決定部と、

前記収集部が収集した前記余命指標値が、前記決定部が決定した閾値よりも短い場合は、前記半導体記憶装置を前記予備半導体記憶装置と交換する交換部と

を有することを特徴とする情報処理装置。

【０１０９】

（付記２）前記決定部は、

収集した前記比率においてデータの読出しの比率が多い場合、第１の閾値を前記半導体記憶装置の交換基準となる閾値とし、データの書き込みの比率が多い場合、前記第１の閾値よりも値が大きい第２の閾値を前記半導体記憶装置の交換基準となる閾値とすることを特徴とする付記１に記載の情報処理装置。

【０１１０】

（付記３）前記情報処理装置はさらに、

前記予備半導体記憶装置を複数有し、

前記決定部は、

収集した前記比率においてデータの読出しの比率が多い場合、前記複数の予備半導体記憶装置のうち、余命指標値が前記第１の閾値よりも大きく、かつ、前記第２の閾値よりも小さい予備半導体記憶装置を選択し、

前記交換部は、

前記半導体記憶装置を前記選択部が選択した予備半導体記憶装置と交換することを特徴とする付記２に記載の情報処理装置。

【０１１１】

（付記４）前記選択部は、

収集した前記比率においてデータの書き込みの比率が多い場合、前記複数の予備半導体記憶装置のうち、余命が前記第２の閾値よりも大きい予備半導体記憶装置を選択すること

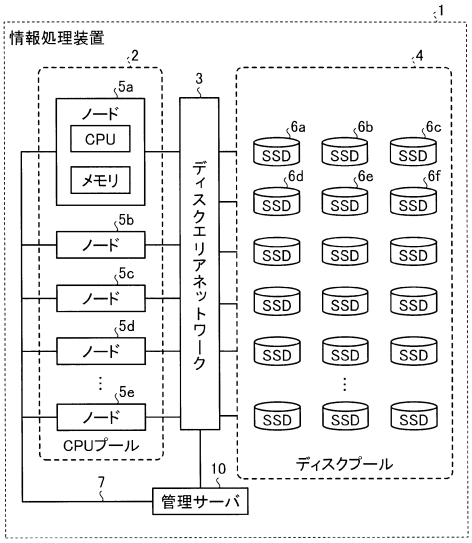
を特徴とする付記３に記載の情報処理装置。
【０１１２】
（付記５）前記収集部は、
前記比率として、単位時間当たりに前記半導体記憶装置から読出されたデータ量と前記半導体記憶装置に書き込まれたデータ量との比率を収集し、
前記判定部は、
単位時間当たりの前記半導体記憶装置から読み出したデータ量が所定の閾値よりも多い場合、データの読出しの比率が多いと判定することを特徴とする付記１～４のいずれか１つに記載の情報処理装置。
【０１１３】
（付記６）前記収集部は、
前記比率として、単位時間当たりに前記半導体記憶装置からデータを読出した回数とデータを書き込んだ回数との比率を収集し、
前記判定部は、
単位時間当たりの前記データの読出し回数が所定の閾値よりも多い場合、データの読出しの比率が多いと判定することを特徴とする付記１～４のいずれか１つに記載の情報処理装置。
【０１１４】
（付記７）前記情報処理装置はさらに、
データが読み書きされる複数の半導体記憶装置と、
前記半導体記憶装置が、二重化されている場合、前記交換部が、前記半導体記憶装置を前記予備の半導体記憶装置と交換する間、前記半導体記憶装置に対する書込みデータを保持する保持部と、
前記交換部が、交換対象の半導体記憶装置が記憶するデータと、前記保持部が保持したデータとを、前記予備半導体記憶装置と、前記複数の半導体記憶装置のうち前記予備半導体記憶装置と二重化される半導体記憶装置とに複製する複製部と
を有することを特徴とする付記１～６のいずれか１つに記載の情報処理装置。
【０１１５】
（付記８）前記情報処理装置はさらに、
データが読み書きされる複数の半導体記憶装置を有し、
前記交換部は、
前記複数の半導体記憶装置に、データとデータを復元するためのパリティとが分散して格納されている場合、前記複数の半導体記憶装置のうち、いずれかの半導体記憶装置を前記予備半導体記憶装置と交換したとき、前記複数の半導体記憶装置のうち、前記予備半導体記憶装置以外の半導体記憶装置が記憶するデータとパリティとを用いて、前記予備半導体記憶装置が記憶するデータを生成することを特徴とする付記１～６のいずれか１つに記載の情報処理装置。
【０１１６】
（付記９）データが読み書きされる半導体記憶装置と、前記半導体記憶装置の予備である予備半導体記憶装置とを有する情報処理装置の制御プログラムにおいて、
前記情報処理装置に、
前記半導体記憶装置について、書き込まれたデータを消去できる残余回数に基づいた余命指標値、および、前記半導体記憶装置に対する読出しと書き込みとの読み書きの比率を収集させ、
収集された前記比率に基づいて、前記半導体記憶装置の交換基準となる閾値を決定させ、
収集された前記余命指標値が、決定された前記閾値よりも短い場合は、前記半導体記憶装置を前記予備半導体記憶装置と交換させる
ことを特徴とする情報処理装置の制御プログラム。
【０１１７】

（付記１０）データが読み書きされる半導体記憶装置と、前記半導体記憶装置の予備である予備半導体記憶装置とを有する情報処理装置の制御方法において、
前記情報処理装置が、
前記半導体記憶装置について、書き込まれたデータを消去できる残余回数に基づいた余命指標値、および、前記半導体記憶装置に対する読出しと書き込みとの読み書きの比率を収集し、
収集された前記比率に基づいて、前記半導体記憶装置の交換基準となる閾値を決定し、
収集された前記余命指標値が、決定された前記閾値よりも短い場合は、前記半導体記憶装置を前記予備半導体記憶装置と交換する
ことを特徴とする情報処理装置の制御方法。

【符号の説明】
【０１１８】
１　情報処理装置
２　ＣＰＵプール
３　ディスクエリアネットワーク
４　ディスクプール
５ａ～５ｅ　ノード
６ａ～６ｆ　ＳＳＤ
７　管理ネットワーク
１０　管理サーバ
１１　記憶部
１２　余命管理テーブル
１３　ワークロード管理テーブル
１４　通信部
１５　更新部
１６　判定部
１７　決定部
１８　選択部
１９　設定部

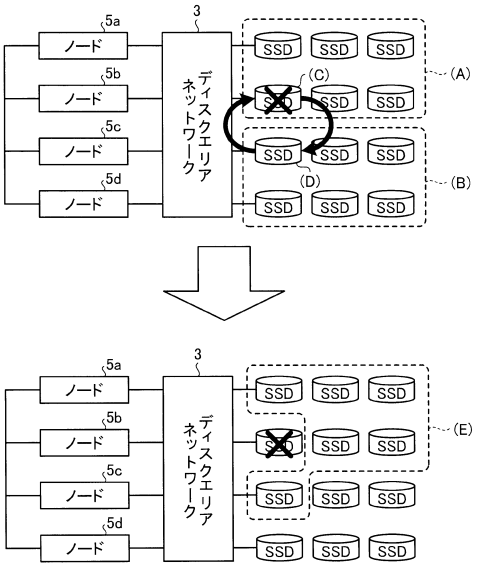
【図１】

実施例1に係る情報処理装置を説明する図



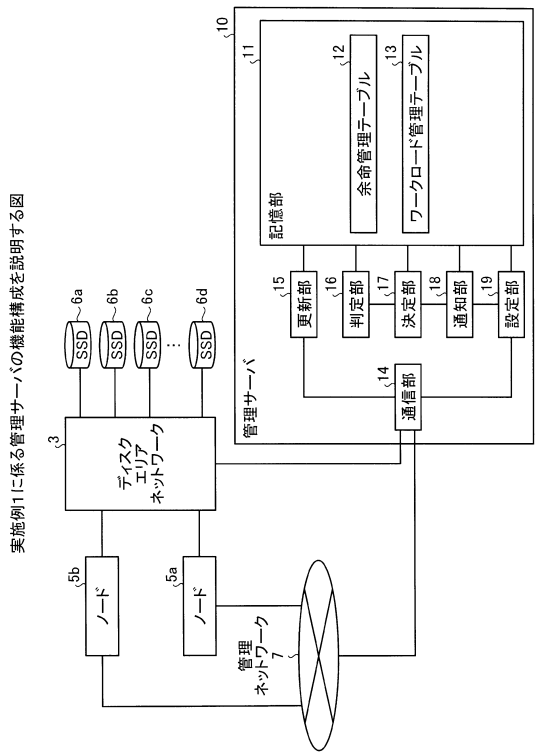
【図２】

実施例1に係る情報処理装置が実行する処理を説明する図



【図３】

実施例1に係る管理サーバの機能構成を説明する図



【図４】

余命管理テーブルの一例を説明する図

12

SSD ID	接続先	余命
SSD#0	ノード5a	100
SSD#1	ノード5b	80
SSD#2	-(未接続)	50
…		

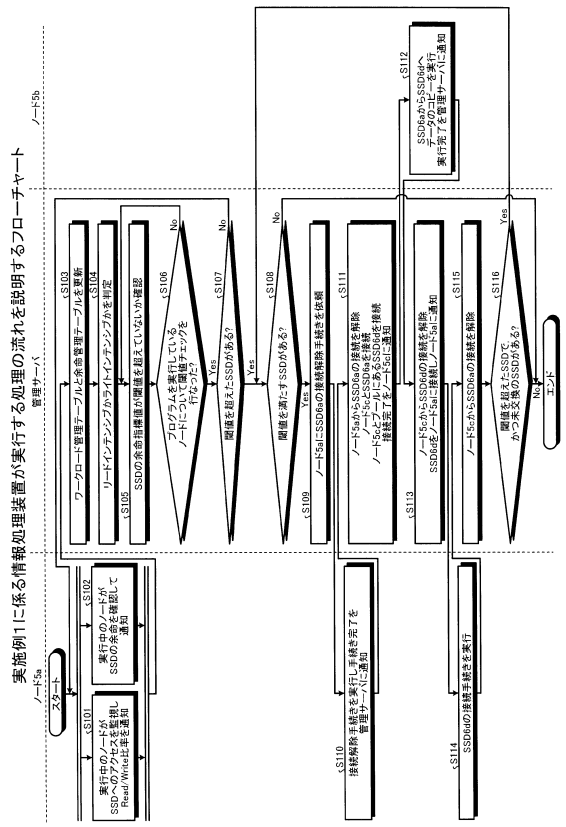
【図５】

ワークロード管理テーブルの一例を説明する図

13

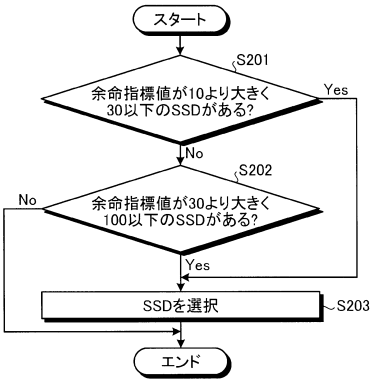
ノードID	プログラム実行フラグ	比率 (Byte/s)
ノード5a	1(実行している)	50:50
ノード5b	1(実行している)	90:10
ノード5c	0(実行していない)	null
…		

【図６】



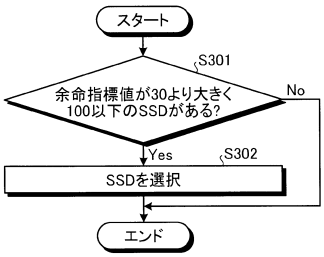
【図７】

選択部が実行する処理の流れを説明するための第１のフローチャート



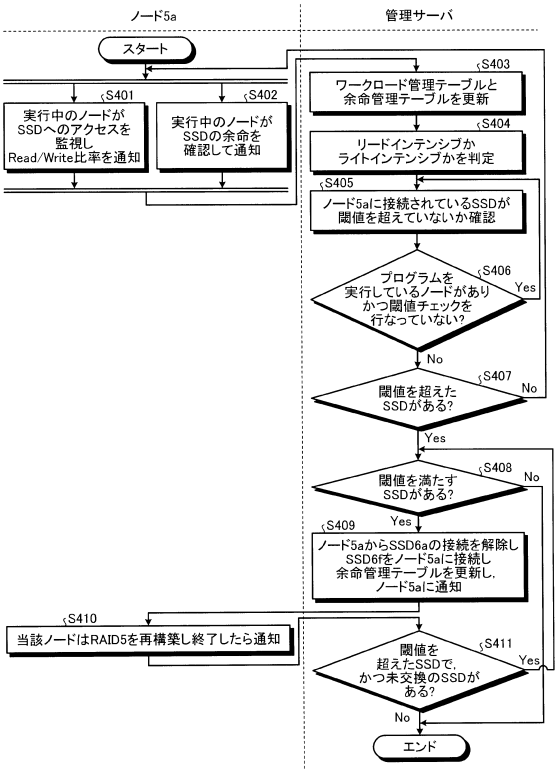
【図８】

選択部が実行する処理の流れを説明するための第２のフローチャート



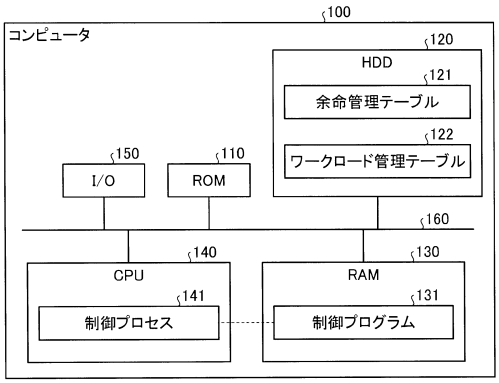
【図９】

実施例2に係る情報処理装置が実行する処理の流れを説明するフローチャート



【図１０】

制御プログラムを実行するコンピュータを説明する図



フロントページの続き

(56)参考文献 特開２０１３－０２０５４４（ＪＰ，Ａ）
特開２００７－１１５２３２（ＪＰ，Ａ）
特開２００７－３２３２２４（ＪＰ，Ａ）
特開２００８－０９７２３７（ＪＰ，Ａ）

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