

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第6320439号
(P6320439)

(45) 発行日 平成30年5月9日 (2018.5.9)

(24) 登録日 平成30年4月13日 (2018.4.13)

(51) Int. Cl.

F I

G 0 6 F 12/16 (2006.01)
 G 0 6 F 12/02 (2006.01)
 G 0 6 F 12/00 (2006.01)
 G 0 6 F 3/06 (2006.01)

G 0 6 F 12/16 3 1 0 A
 G 0 6 F 12/02 5 1 0 A
 G 0 6 F 12/00 5 9 7 U
 G 0 6 F 3/06 3 0 6 Z

請求項の数 8 (全 46 頁)

(21) 出願番号 特願2016-46934 (P2016-46934)
 (22) 出願日 平成28年3月10日 (2016.3.10)
 (65) 公開番号 特開2017-162260 (P2017-162260A)
 (43) 公開日 平成29年9月14日 (2017.9.14)
 審査請求日 平成29年3月10日 (2017.3.10)

(73) 特許権者 000003078
 株式会社東芝
 東京都港区芝浦一丁目1番1号
 (73) 特許権者 301063496
 東芝デジタルソリューションズ株式会社
 神奈川県川崎市幸区堀川町72番地34
 (74) 代理人 110001737
 特許業務法人スズエ国際特許事務所
 (72) 発明者 小石 真人
 東京都港区芝浦一丁目1番1号 株式会社
 東芝内

審査官 後藤 彰

最終頁に続く

(54) 【発明の名称】 メモリ制御装置、ストレージ装置、ストレージ装置の制御プログラム、及び、ストレージ装置の制御方法

(57) 【特許請求の範囲】

【請求項 1】

複数の不揮発性メモリを所定サイズで分けした複数の領域のそれぞれに対する書き込み量を管理する第1のデータと、前記複数の不揮発性メモリに対応する書き込み状態を管理する第2のデータとを格納するメモリと、

前記複数の不揮発性メモリに対する書き込み対象データの書き込みにおいて、前記メモリに格納されている前記第1のデータと前記第2のデータとに基づいて、前記複数の不揮発性メモリのそれぞれに対する第1の評価値を計算し、前記第1の評価値に基づいて、前記書き込み対象データの書き込み先の不揮発性メモリを選択する補正手段と、

前記補正手段によって選択された前記不揮発性メモリに対して、前記書き込み対象データを書き込み、前記メモリに格納されている前記第1のデータと前記第2のデータとを更新する書き込み手段と、

前記複数の不揮発性メモリの使用開始後に、前記メモリに格納されている前記第2のデータに基づいて、前記複数の不揮発性メモリのそれぞれに対する第2の評価値を計算し、前記第2の評価値に基づいて、前記複数の不揮発性メモリのうちの少なくとも1つに対して使用可能容量の設定を変更する設定変更手段と、

を具備するメモリ制御装置。

【請求項 2】

前記補正手段は、さらに、前記複数の不揮発性メモリに書き込まれている再配置対象データを前記複数の不揮発性メモリに再配置する場合に、前記第1のデータと前記第2のデ

10

20

ータとに基づいて、前記複数の不揮発性メモリのそれぞれに対する第3の評価値を計算し、前記第3の評価値に基づいて、前記再配置対象データの再配置先の不揮発性メモリを選択し、前記再配置先の不揮発性メモリに対して前記再配置対象データを書き込み、前記メモリに格納されている前記第1のデータと前記第2のデータとを更新する、ことを特徴とする請求項1に記載のメモリ制御装置。

【請求項3】

前記第1のデータは、前記複数の領域に対する第1の統計情報を含み、

前記第2のデータは、前記複数の不揮発性メモリに対する第2の統計情報を含み、

前記補正手段は、

前記複数の不揮発性メモリのそれぞれに対する性能評価値と寿命評価値とのうちの少なくとも1つが、前記複数の不揮発性メモリ間で平準化されるように、前記書き込み先の不揮発性メモリを選択し、及び、前記複数の不揮発性メモリのうちの少なくとも1つに対して前記使用可能容量の設定を変更する、

ことを特徴とする請求項1又は請求項2に記載のメモリ制御装置。

【請求項4】

前記補正手段によって実行される計算に用いられるパラメータの値を変更する変更手段をさらに具備する、

ことを特徴とする請求項1乃至請求項3のいずれか1項に記載のメモリ制御装置。

【請求項5】

使用開始後に使用可能容量の設定を変更可能な複数の不揮発性メモリ装置と、

前記複数の不揮発性メモリ装置の記憶領域を所定サイズで分けした複数の領域のそれぞれに対する書き込み量を管理する第1のデータと、前記複数の不揮発性メモリ装置に対応する書き込み状態を管理する第2のデータとを格納するメモリと、

前記複数の不揮発性メモリ装置に対する書き込み対象データの書き込みにおいて、前記メモリに格納されている前記第1のデータと前記第2のデータとに基づいて、前記複数の不揮発性メモリ装置のそれぞれに対する第1の評価値を計算し、前記第1の評価値に基づいて、前記書き込み対象データの書き込み先の不揮発性メモリ装置を選択する補正手段と、

前記補正手段によって選択された前記不揮発性メモリ装置に対して、前記書き込み対象データを書き込み、前記メモリに格納されている前記第1のデータと前記第2のデータとを更新する書き込み手段と、

前記複数の不揮発性メモリ装置の使用開始後に、前記メモリに格納されている前記第2のデータに基づいて、前記複数の不揮発性メモリ装置のそれぞれに対する第2の評価値を計算し、前記第2の評価値に基づいて、前記複数の不揮発性メモリ装置のうちの少なくとも1つに対して前記使用可能容量の設定を変更する設定変更手段と、を具備するストレージ装置。

【請求項6】

前記複数の不揮発性メモリ装置は、複数のディスクアレイであり、

前記複数のディスクアレイのそれぞれは、複数のソリッド・ステート・ドライブを含む、

ことを特徴とする請求項5に記載のストレージ装置。

【請求項7】

使用開始後に使用可能容量の設定を変更可能な複数の不揮発性メモリ装置を含むストレージ装置を、

前記複数の不揮発性メモリ装置の記憶領域を所定サイズで分けした複数の領域のそれぞれに対する書き込み量を管理する第1のデータと、前記複数の不揮発性メモリ装置に対応する書き込み状態を管理する第2のデータとを、メモリに格納する手段、

前記複数の不揮発性メモリ装置に対する書き込み対象データの書き込みにおいて、前記メモリに格納されている前記第1のデータと前記第2のデータとに基づいて、前記複数の不揮発性メモリ装置のそれぞれに対する第1の評価値を計算し、前記第1の評価値に基づ

10

20

30

40

50

いて、前記書き込み対象データの書き込み先の不揮発性メモリ装置を選択する補正手段、

前記補正手段によって選択された前記不揮発性メモリ装置に対して、前記書き込み対象データを書き込み、前記メモリに格納されている前記第1のデータと前記第2のデータとを更新する書き込み手段、

前記複数の不揮発性メモリ装置の使用開始後に、前記メモリに格納されている前記第2のデータに基づいて、前記複数の不揮発性メモリ装置のそれぞれに対する第2の評価値を計算し、前記第2の評価値に基づいて、前記複数の不揮発性メモリ装置のうちの少なくとも1つに対して前記使用可能容量の設定を変更する設定変更手段、
として機能させるための制御プログラム。

【請求項8】

10

使用開始後に使用可能容量の設定を変更可能な複数の不揮発性メモリ装置を含むストレージ装置を制御する制御方法であって、

前記複数の不揮発性メモリ装置の記憶領域を所定サイズで分けした複数の領域のそれぞれに対する書き込み量を管理する第1のデータと、前記複数の不揮発性メモリ装置に対応する書き込み状態を管理する第2のデータとを、メモリに格納することと、

前記複数の不揮発性メモリ装置に対する書き込み対象データの書き込みにおいて、前記メモリに格納されている前記第1のデータと前記第2のデータとに基づいて、前記複数の不揮発性メモリ装置のそれぞれに対する第1の評価値を計算し、前記第1の評価値に基づいて、前記書き込み対象データの書き込み先の不揮発性メモリ装置を選択することと、

選択された前記不揮発性メモリ装置に対して、前記書き込み対象データを書き込み、前記メモリに格納されている前記第1のデータと前記第2のデータとを更新することと、

20

前記複数の不揮発性メモリ装置の使用開始後に、前記メモリに格納されている前記第2のデータに基づいて、前記複数の不揮発性メモリ装置のそれぞれに対する第2の評価値を計算し、前記第2の評価値に基づいて、前記複数の不揮発性メモリ装置のうちの少なくとも1つに対して前記使用可能容量の設定を変更することと、
を具備する制御方法。

【発明の詳細な説明】

【技術分野】

【0001】

本実施形態は、メモリ制御装置、ストレージ装置、ストレージ装置の制御プログラム、及び、ストレージ装置の制御方法に関する。

30

【背景技術】

【0002】

フラッシュメモリの寿命を長期化する技術の一例として、オーバプロビジョニングがある。オーバプロビジョニングとは、例えばSSD（ソリッド・ステート・ドライブ）の論理容量を制限し、SSDの物理容量の特定領域を書き込み不能な領域として確保することで、データの書き込み可能な量を増加させる技術である。

【先行技術文献】

【特許文献】

【0003】

40

【特許文献1】特開2013-061847号公報

【特許文献2】特開2011-186561号公報

【発明の概要】

【発明が解決しようとする課題】

【0004】

本実施形態は、ストレージ装置の寿命を長期化することを目的とする。

【課題を解決するための手段】

【0005】

本実施形態によれば、メモリ制御装置は、メモリと、補正手段と、書き込み手段と、設定変更手段とを含む。メモリは、複数の不揮発性メモリを所定サイズで分けした複数の

50

領域のそれぞれに対する書き込み量を管理する第1のデータと、複数の不揮発性メモリに対応する書き込み状態を管理する第2のデータとを格納する。補正手段は、複数の不揮発性メモリに対する書き込み対象データの書き込みにおいて、メモリに格納されている第1のデータと第2のデータとに基づいて、複数の不揮発性メモリのそれぞれに対する第1の評価値を計算し、第1の評価値に基づいて、書き込み対象データの書き込み先の不揮発性メモリを選択する。書き込み手段は、補正手段によって選択された不揮発性メモリに対して、書き込み対象データを書き込み、メモリに格納されている第1のデータと第2のデータとを更新する。設定変更手段は、複数の不揮発性メモリの使用開始後に、メモリに格納されている第2のデータに基づいて、複数の不揮発性メモリのそれぞれに対する第2の評価値を計算し、第2の評価値に基づいて、複数の不揮発性メモリのうちの少なくとも1つに対して使用可能容量の設定を変更する。

10

【図面の簡単な説明】

【0006】

【図1】第1の実施形態に係るメモリ制御装置を含むストレージ装置の一例を示すブロック図。

【図2】第1の実施形態に係るストレージ装置のディスク構造の一例を示す概念図。

【図3】第1の実施形態に係るデバイス固有情報テーブルの一例を示すデータ構造図。

【図4】第1の実施形態に係る論理ディスク情報テーブルの一例を示すデータ構造図。

【図5】第1の実施形態に係る論理エクステンション情報テーブルの一例を示すデータ構造図。

20

【図6】第1の実施形態に係る物理エクステンション情報テーブルの一例を示すデータ構造図。

【図7】第1の実施形態に係るデバイス構成部の処理の一例を示すフローチャート。

【図8】第1の実施形態に係るデバイス構成部のレコード追加処理の一例を示すフローチャート。

【図9】第1の実施形態に係る論理ディスク構成部の処理の一例を示すフローチャート。

【図10】第1の実施形態に係る論理ディスク構成部によって実行される論理ディスクに関するレコード追加処理の一例を示すフローチャート。

【図11】第1の実施形態に係る読み出し部の処理の一例を示すフローチャート。

【図12】第1の実施形態に係るアドレス範囲抽出部の処理の一例を示すフローチャート。

30

【図13】第1の実施形態に係る書き込み部の処理の一例を示すフローチャート。

【図14】第1の実施形態に係る寿命補正部のアロケート書き込み先選択処理の一例を示すフローチャート。

【図15】第1の実施形態に係る寿命補正部のアロケート書き込み先評価処理の一例を示すフローチャート。

【図16】デバイス固有情報テーブルの一例を示すデータ構造図。

【図17】論理エクステンション情報テーブルの一例を示すデータ構造図。

【図18】物理エクステンション情報テーブルの一例を示すデータ構造図。

【図19】デバイスと評価値との関係の一例を示す図。

40

【図20】更新後の論理エクステンション情報テーブルの一例を示すデータ構造図。

【図21】第1の実施形態に係る寿命補正部の寿命補正処理の一例を示すフローチャート。

【図22】第1の実施形態に係る寿命補正部の寿命補正処理の変形例を示すフローチャート。

【図23】デバイスIDごとの累積書き込み量を例示する図。

【図24】デバイスIDごとの評価値を例示する図。

【図25】再配置後の論理エクステンション情報テーブルの一例を示すデータ構造図。

【図26】再配置後のデバイス固有情報テーブルの一例を示すデータ構造図。

【図27】オーバープロビジョニングの設定評価処理によって得られる評価結果の例を示

50

す図。

【図 2 8】第 1 の実施形態に係る寿命補正部の再配置処理の一例を示すフローチャート。

【図 2 9】第 1 の実施形態に係る寿命補正部の再配置先評価処理の一例を示すフローチャート。

【図 3 0】第 1 の実施形態に係る設定変更部によって実行されるオーバープロビジョニングの設定変更処理の一例を示すフローチャート。

【図 3 1】第 1 の実施形態に係る設定変更部によって実行されるオーバープロビジョニングの設定評価処理の一例を示すフローチャート。

【図 3 2】第 2 の実施形態に係るメモリ制御装置を含むストレージ装置の一例を示すブロック図。

10

【図 3 3】第 2 の実施形態に係るポリシー変更部の処理の一例を示すフローチャート。

【図 3 4】第 2 の実施形態に係るポリシー情報テーブルの一例を示すデータ構造図。

【図 3 5】第 3 の実施形態に係るメモリ制御装置を含むストレージ装置の一例を示すブロック図。

【図 3 6】第 3 の実施形態に係る R A I D グループ情報テーブルの一例を示すデータ構造図。

【図 3 7】第 3 の実施形態に係る S S D 固有情報テーブルの一例を示すデータ構造図。

【図 3 8】第 3 の実施形態に係る R A I D 処理部の処理の一例を示すフローチャート。

【発明を実施するための形態】

【 0 0 0 7 】

20

以下、図面を参照しながら本実施形態について説明する。なお、以下の各図において同一又はほぼ同一の要素には同一の符号を付して説明を省略するか又は簡単に説明し、異なる部分についてののみ詳しく説明する。

【 0 0 0 8 】

[第 1 の実施形態]

本実施形態においては、オーバープロビジョニングの設定を変更することにより、メモリの寿命を長期化するメモリ制御装置について説明する。

【 0 0 0 9 】

以下、識別情報は I D と表記する。

【 0 0 1 0 】

30

本実施形態において、メモリ装置は、デバイスという。

【 0 0 1 1 】

本実施形態において、ストレージ装置に含まれる複数のデバイス間の寿命、性能、負荷、容量の不均衡を低減することは、寿命、性能、負荷、容量の平準化という。

【 0 0 1 2 】

本実施形態において、デバイスは、デバイス起動後に、オーバープロビジョニングの設定を変更可能である。オーバープロビジョニングの設定変更によってデバイスの寿命を長期化する処理を、寿命補正処理という。

【 0 0 1 3 】

本実施形態において、アクセスとは、データ読み出し又はデータ書き込みの双方を含むとする。

40

【 0 0 1 4 】

本実施形態において、デバイスと当該デバイスの I D とは、同じ符号で表記する。エクステンと当該エクステンとの I D とは、同じ符号で表記する。

【 0 0 1 5 】

図 1 は、本実施形態に係るメモリ制御装置を含むストレージ装置の一例を示すブロック図である。

【 0 0 1 6 】

ストレージシステム 1 は、ホスト装置 2 1 ~ 2 3 と、ストレージ装置 3 とを含む。図 1 の例では、ストレージ装置 3 が 3 台のホスト装置 2 1 ~ 2 3 と通信可能な場合を例示して

50

いる。しかしながら、ストレージ装置 3 は、1 台、2 台、又は、4 台以上のホスト装置と通信可能でもよい。

【0017】

ホスト装置 2 1 ~ 2 3 は、書き込みコマンドと書き込み先アドレスと書き込みデータをストレージ装置 3 へ送信する。また、ホスト装置 2 1 ~ 2 3 は、読み出しコマンドと読み出し先アドレスとをストレージ装置 3 へ送信し、ストレージ装置 3 から読み出しデータを受信する。

【0018】

ストレージ装置 3 は、ホストインタフェース部 4、メモリ制御装置 5、デバイス $D_0 \sim D_{N-1}$ を含む。図 1 の例ではデバイスが N 台 (N は自然数) の場合を例示しているが、デ
バイスは 1 台以上でよい。本実施形態では、デバイス $D_0 \sim D_{N-1}$ のうちのそれぞれをデバ
イス D_i と表記する。 i は、 $0, 1, \dots, N-1$ である。

10

【0019】

例えば、ホストインタフェース部 4、メモリ制御装置 5、デバイス $D_0 \sim D_{N-1}$ は、互いにデータ、アドレス、コマンド、情報、信号、指示などを送受信可能である。

【0020】

ホストインタフェース部 4 は、ホスト装置 2 1 ~ 2 3 からのアクセスを受け付ける。

【0021】

メモリ制御装置 5 は、メモリ 5 1、プロセッサ 5 2 を含む。

【0022】

20

メモリ 5 1 は、例えば、主記憶装置として機能してもよい。メモリ 5 1 は、バッファメモリを含むとしてもよく、キャッシュメモリを含むとしてもよく、不揮発性メモリを含むとしてもよい。メモリ 5 1 としては、例えば、DRAM (Dynamic Random Access Memory) が用いられる。

【0023】

メモリ 5 1 は、例えば、制御プログラム CP と各種データを格納する。制御プログラム CP は、例えば、オペレーティングシステム、アプリケーションプログラム、ファームウェアなどでもよい。例えば、メモリ 5 1 は、各種のデータとして、デバイス固有情報テーブル 6、論理ディスク情報テーブル 7、論理エクステンション情報テーブル 8、物理エクステンション情報テーブル 9 を格納する。

30

【0024】

プロセッサ 5 2 は、メモリ 5 1 をアクセスし、各種処理を実行し、デバイス $D_0 \sim D_{N-1}$ に対するアクセス制御を実行し、オーバープロビジョニングの設定変更を行う。

【0025】

プロセッサ 5 2 は、例えば、メモリ 5 1 に格納されている制御プログラム CP を実行し、これによりデバイス構成部 1 0、論理ディスク構成部 1 1、読み出し部 1 2、書き込み部 1 3、アドレス範囲抽出部 1 4、寿命補正部 1 5、オーバープロビジョニングの設定変更部 1 6 として機能する。

【0026】

デバイス $D_0 \sim D_{N-1}$ は、それぞれ、コントローラ 1 7 と 1 以上の不揮発性メモリ 1 8 とを含む。

40

【0027】

コントローラ 1 7 は、プロセッサ 5 2 からのコマンドにしたがって、不揮発性メモリ 1 8 にデータを書き込み、又は、不揮発性メモリ 1 8 からデータを読み出す。

【0028】

本実施形態では、不揮発性メモリ 1 8 が NAND 型フラッシュメモリを含む場合について説明する。しかしながら、不揮発性メモリ 1 8 は、例えば、NOR 型フラッシュメモリ、MRAM (Magnetoresistive Random Access Memory)、PRAM (Phase change Random Access Memory)、ReRAM (Resistive Random Access Memory)、又は、FeRAM (Ferroelectric Random Access Memory) など NAND 型フラッシュメモリではない他

50

の種別のメモリを含むとしてもよい。また、不揮発性メモリ 18 は、3次元構造のフラッシュメモリを含むとしてもよい。

【0029】

デバイス $D_0 \sim D_{N-1}$ は、オーバプロビジョニングの設定の変更を当該デバイス $D_0 \sim D_{N-1}$ の運用後であっても可能である。本実施形態において、デバイス $D_0 \sim D_{N-1}$ は、例えば、半導体記憶装置であり、より具体的には、例えば SSD であるとする。しかしながら、デバイスは、例えば、SSD と HDD とを組み合わせたハイブリッド型のデバイスでもよい。

【0030】

デバイス $D_0 \sim D_{N-2}$ は、データ格納のための装置である。デバイス D_{N-1} は、ホットス

10

【0031】

デバイス構成部 10、論理ディスク構成部 11、読み出し部 12、書き込み部 13、アドレス範囲抽出部 14、寿命補正部 15、設定変更部 16 の具体的な説明は後述する。

【0032】

図2は、本実施形態に係るストレージ装置3のディスク構造の一例を示す概念図である。

【0033】

ストレージ装置3は、ホスト装置21~23に対して、ストレージ装置3内のデバイス $D_0 \sim D_{N-1}$ の記憶領域を、複数の論理ディスク $L D_0 \sim L D_{N-1}$ として提供する。この図2では、N個の論理ディスクが用いられる場合を例として説明するが、論理ディスクの数は、1以上で変更可能である。各論理ディスク $L D_0 \sim L D_{N-1}$ には、識別用の論理ディスク ID が割り当てられている。各論理ディスク $L D_0 \sim L D_{N-1}$ はセクタに分割され、各セクタには論理ブロックアドレス(論理アドレス)が割り当てられる。ホスト装置21~23は論理アドレスを用いて各論理ディスク $L D_0 \sim L D_{N-1}$ 内のセクタにアクセスする。

20

【0034】

論理アドレスは、ストレージ装置3に備えられているデバイス $D_0 \sim D_{N-1}$ のデータが格納されている各セクタの番号(物理アドレス)に一対一に対応している。物理アドレス及び論理アドレス1つに対応するセクタのサイズは一定であり、例えば512Byteでもよい。

【0035】

論理ディスク $L D_0 \sim L D_{N-1}$ 内の記憶領域は、固定長の論理エクステントと呼ぶ単位で管理する。論理エクステントのサイズ(エクステントサイズ)はストレージ装置3内で一定であり、例えば8MByteとしてもよい。論理エクステントは、寿命補正処理における処理単位として扱われる。

30

【0036】

上記の図2では、論理ディスク $L D_0 \sim L D_{N-1}$ のそれぞれが、論理エクステント $l e 1 \sim l e 9$ を含む。デバイス $D_0 \sim D_{N-1}$ のそれぞれが、物理エクステント $p e 1 \sim p e 9$ を含む。

【0037】

論理ディスク $L D_0$ の論理エクステント $l e 1$ は、デバイス D_0 の物理エクステント $p e 2$ に対応している。論理ディスク $L D_0$ の論理エクステント $l e 2$ は、デバイス D_0 の物理エクステント $p e 9$ に対応している。論理ディスク $L D_0$ の論理エクステント $l e 3$ は、デバイス D_{N-1} の物理エクステント $p e 3$ に対応している。論理ディスク $L D_0$ の論理エクステント $l e 9$ はどの物理エクステントとも対応していない。

40

【0038】

ストレージ装置3において、論理エクステントと物理エクステントは一対一に対応しており、論理エクステント内の論理アドレスそれぞれも、論理エクステントと対応する物理エクステント内の物理アドレスと一対一に対応している。

【0039】

ここで、ストレージ装置3のデータ読み出し、データ書き込みの概要を述べる。まず、

50

ホスト装置 2 1 ~ 2 3 は、論理ディスク I D 及び論理アドレスをストレージ装置 3 に送信する。次に、ストレージ装置 3 は、論理ディスク I D 及び論理アドレスを、デバイス D_0 ~ D_{N-1} の物理アドレスに変換し、物理アドレスに該当するデータを読み出し、又は、物理アドレスに対応する記憶領域へデータを書き込む。そして、ストレージ装置 3 は、応答をホスト装置 2 1 ~ 2 3 に送信し、処理を終了する。アドレスの変換は、後述する論理エクステンツ情報テーブル 8 を参照して行われる。

【 0 0 4 0 】

本実施形態においては、ストレージ装置 3 内の各デバイス D_0 ~ D_{N-1} に対して、ワークロードに適切なオーバープロビジョニング設定を行うための処理を説明する。

【 0 0 4 1 】

以下で、上記図 1 で例示されているストレージ装置 3 の機能について詳しく説明する。

【 0 0 4 2 】

デバイス構成部 1 0 は、ストレージ装置 3 にデバイスが追加又は交換される場合の設定を行う。デバイス構成部 1 0 は、例えば、デバイスに関するレコード追加処理を行う。

【 0 0 4 3 】

論理ディスク構成部 1 1 は、論理ディスクが追加される場合の設定を行う。論理ディスク構成部 1 1 は、例えば、論理ディスクに関するレコード追加処理を行う。

【 0 0 4 4 】

読み出し部 1 2 は、対象アドレス抽出処理、読み出し処理、テーブル更新処理を行う。

【 0 0 4 5 】

書き込み部 1 3 は、対象アドレス抽出処理、書き込み処理、テーブル更新処理を行う。

【 0 0 4 6 】

アドレス範囲抽出部 1 4 は、読み出し、書き込みの対象を抽出する。

【 0 0 4 7 】

寿命補正部 1 5 は、デバイス D_0 ~ D_{N-1} の寿命補正処理、再配置先選択処理、再配置処理、再配置先評価処理、アロケート書き込み先選択処理、アロケート書き込み先評価処理を行う。

【 0 0 4 8 】

設定変更部 1 6 は、オーバープロビジョニングの設定変更処理、設定評価処理を行う。

【 0 0 4 9 】

デバイス固有情報テーブル 6 は、ストレージ装置 3 のデバイス D_0 ~ D_{N-1} の論理容量等を管理する。

【 0 0 5 0 】

論理ディスク情報テーブル 7 は、論理ディスク $L D_0$ ~ $L D_{N-1}$ の論理容量等を管理する。

【 0 0 5 1 】

論理エクステンツ情報テーブル 8 は、論理エクステンツのアドレス情報と物理エクステンツとの対応関係を管理する。

【 0 0 5 2 】

物理エクステンツ情報テーブル 9 は、物理エクステンツの情報を管理する。

【 0 0 5 3 】

本実施形態の各図において、MbyteはMBと表記する。

【 0 0 5 4 】

図 3 は、本実施形態に係るデバイス固有情報テーブル 6 の一例を示すデータ構造図である。

【 0 0 5 5 】

デバイス固有情報テーブル 6 は、デバイス D_i の使用容量、書き込み量を管理するための情報である。図 3 の 1 エントリは、1 つのデバイスに対応している。デバイス固有情報テーブル 6 には、ストレージ装置 3 内のホットスペア用のデバイスを含むすべてのデバイス D_i の情報がエントリされている。

10

20

30

40

50

【 0 0 5 6 】

デバイス固有情報テーブル 6 の各エントリは、各デバイス D_i のデバイス ID、最大容量、論理容量、現在使用容量、残り書き込み量、保証書き込み量、読み出し量、更新量、追記量、性能仕様、最大負荷を含む。

【 0 0 5 7 】

デバイス ID は、デバイス D_i を一意に特定するための ID であり、例えば、対応するデバイスの SCSI (Small Computer System Interface) ID を用いてもよい。

【 0 0 5 8 】

最大容量は、ホスト装置 2 1 ~ 2 3 が使用できる容量の最大値を表す。

【 0 0 5 9 】

論理容量は、ホスト装置 2 1 ~ 2 3 の使用可能な容量であり、最大容量とは異なる。

【 0 0 6 0 】

現在使用容量は、論理容量のうち、実際にホスト装置 2 1 ~ 2 3 が使用している分の容量である。

【 0 0 6 1 】

残り書き込み量は、デバイス D_i のベンダが保証する総書き込み量から、デバイス D_i の使用開始からの書き込み量の合計を引いた値である。メモリ制御装置 5 は、残り書き込み量の値に基づいて、デバイス D_i に対して適切なオーバプロビジョニングの設定を行う。

【 0 0 6 2 】

保証書き込み量は、ベンダが保証する、単位時間あたりの書き込み量である。この保証書き込み量にそったペース以内で書き込みが実行されれば、ベンダの保証期間を過ぎるまでデバイス D_i は寿命に達しないことが保証される。本実施形態において、例えば、単位時間は一日としてもよい。

【 0 0 6 3 】

読み出し量は、デバイス D_i に対する単位時間あたりの読み出されたデータ量の累計値である。

【 0 0 6 4 】

更新量は、デバイス D_i に対する単位時間あたりに更新されたデータ量の累計値である。

【 0 0 6 5 】

追記量は、デバイス D_i に対する単位時間あたりのアロケート書き込みされたデータ量の累計値である。

【 0 0 6 6 】

アロケート書き込みとは、物理エクステントの割り当ての行われていない論理エクステントへの書き込みを指す。

【 0 0 6 7 】

性能仕様は、デバイス D_i の最大性能の仕様である。性能指数としては、IOPS (Input/Output Per Second)、応答時間(単位:millisecond)、スループット(単位:MB/sec)などが用いられてもよい。本実施形態では性能指数の一例として IOPS を使用する。

【 0 0 6 8 】

最大負荷は、デバイス D_i に対する単位時間中に記録された性能指数の最大値である。換言すれば、最大負荷は、単位時間中に記録されるデバイス D_i の最大性能を表す。

【 0 0 6 9 】

本実施形態におけるオーバプロビジョニングの設定では、論理容量を、最大容量以下の値で、デバイス D_i に対して設定することができる。

【 0 0 7 0 】

デバイス構成部 1 1 は、新しいデバイスが追加される場合に、デバイス固有情報テーブル 6 における各属性の初期化を行う。

【 0 0 7 1 】

10

20

30

40

50

ここで、デバイス構成部 1 1 によるデバイス固有情報テーブル 6 における各属性の初期化及び更新について説明する。デバイス固有情報テーブル 6 の更新は、デバイス Di の種類により適宜変更される。

【 0 0 7 2 】

デバイス構成部 1 1 は、デバイス固有情報テーブル 6 に、デバイス ID、最大容量、保証書き込み量、性能仕様として、固定値を設定する。

【 0 0 7 3 】

デバイス構成部 1 1 は、論理容量と書き込み量上限を初期化し、寿命補正部 1 5 は、論理容量を更新する。

【 0 0 7 4 】

デバイス構成部 1 1 は、現在使用容量、残り書き込み量、更新量、追記量を初期化し、書き込み部 1 3 は、現在使用容量、残り書き込み量、更新量、追記量を更新する。デバイス構成部 1 1 は、単位時間が経過すると、更新量、追記量をリセットする。

【 0 0 7 5 】

デバイス構成部 1 1 は、読み出し量を初期化し、読み出し部 1 2 は、読み出し量を更新する。デバイス構成部 1 1 は、単位時間が経過すると、読み出し量をリセットする。

【 0 0 7 6 】

デバイス構成部 1 1 は、最大負荷を、単位時間中に記録し続ける。デバイス構成部 1 1 は、単位時間が経過すると、最大負荷をリセットする。

【 0 0 7 7 】

図 4 は、本実施形態に係る論理ディスク情報テーブル 7 の一例を示すデータ構造図である。

【 0 0 7 8 】

論理ディスク情報テーブル 7 は、論理ディスク $LD_0 \sim LD_{N-1}$ の論理容量を管理するための情報である。図 4 の 1 エントリは、1 つの論理ディスクに対応している。論理ディスク情報テーブル 7 には、ストレージ装置 3 内のすべての論理ディスク $LD_0 \sim LD_{N-1}$ の情報がエントリされている。

【 0 0 7 9 】

論理ディスク情報テーブル 7 の各エントリは、論理ディスク ID と論理容量とを含む。

【 0 0 8 0 】

論理ディスク ID は、論理ディスクを一意に特定するための ID である。

【 0 0 8 1 】

論理容量は、論理ディスクに設定されている論理容量を表す。

【 0 0 8 2 】

論理ディスク構成部 1 1 は、論理ディスク情報テーブル 7 の初期化と更新を行う。

【 0 0 8 3 】

図 5 は、本実施形態に係る論理エクステント情報テーブル 8 の一例を示すデータ構造図である。

【 0 0 8 4 】

論理エクステント情報テーブル 8 は、論理エクステントを格納する位置を示すアドレス情報と、論理エクステント ID と、物理エクステント ID との対応関係を管理するための情報である。図 5 の 1 エントリは、1 つの論理エクステントに対応している。論理エクステント情報テーブル 8 の各エントリは、各論理エクステントの論理エクステント ID、論理ディスク ID、先頭論理アドレス、所属デバイス ID、物理エクステント ID、読み出し量、書き込み量を含む。論理エクステント情報テーブル 8 には、ストレージ装置 3 内の全ての論理エクステントがエントリされている。

【 0 0 8 5 】

論理エクステント ID は、論理ディスク内で論理エクステントを一意に特定するための識別情報である。

【 0 0 8 6 】

論理ディスクIDは、論理エクステントを含む論理ディスクを特定するためのIDである。

【0087】

先頭論理アドレスは、論理ディスク内で論理エクステントを管理する論理アドレス領域の先頭の論理アドレスである。

【0088】

所属デバイスIDは、論理エクステントと対応する物理エクステントを含むデバイスを特定するためのIDである。

【0089】

物理エクステントIDは、論理エクステントと対応する物理エクステントを特定するためのIDである。

10

【0090】

読み出し量は、論理エクステントに対する単位時間あたりの読み出されたデータ量の累計値である。

【0091】

書き込み量は、論理エクステントに対する単位時間あたりの更新されたデータ量の累計値である。

【0092】

デバイス構成部10は、論理エクステントID、論理ディスクID、先頭論理アドレスを初期化し、更新は行わない。

20

【0093】

書き込み部13は、各論理エクステントの所属デバイスIDと物理エクステントIDとの組を初期化する。寿命補正部15は、再配置処理に基づいて、各論理エクステントに対応する所属デバイスIDと物理エクステントIDとの組を更新する。

【0094】

論理ディスク構成部11は、読み出し量、書き込み量を初期化し、それぞれゼロとする。読み出し部12は、読み出し量を更新する。論理ディスク構成部11は、単位時間が経過すると、読み出し量をリセットする。書き込み部13は、書き込み量を更新する。論理ディスク構成部11は、単位時間が経過すると、書き込み量をリセットする。

【0095】

30

図6は、本実施形態に係る物理エクステント情報テーブル9の一例を示すデータ構造図である。

【0096】

物理エクステント情報テーブル9は、物理エクステントの情報を管理するための情報である。図6の1エントリは、1つの物理エクステントに対応している。物理エクステント情報テーブル9の各エントリは、各物理エクステントの物理エクステントID、所属デバイスID、先頭物理アドレス、空き情報を含む。物理エクステント情報テーブル9には、ストレージ装置3内のすべての物理エクステントがエントリされている。物理エクステントは、ストレージ装置3に備えられるデバイス $D_0 \sim D_{N-1}$ が提供する全記憶領域に対して定義される。

40

【0097】

物理エクステントIDは、物理エクステントをデバイス $MD_0 \sim MD_{N-1}$ 内で一意に特定するためのIDである。

【0098】

所属デバイスIDは、物理エクステントを含むデバイスを特定するためのIDである。

【0099】

先頭物理アドレスは、物理エクステントを管理する物理アドレス領域の先頭の物理アドレスである。

【0100】

空き情報は、物理エクステントの状態を示す情報である。空き情報は、物理エクステン

50

トを管理するアドレス領域に基づいて以下の値を持つ。

【 0 1 0 1 】

物理エクステンツのアドレス領域がすべて消去済みの場合、空き情報は 1 (「空き」) となる。

【 0 1 0 2 】

物理エクステンツのアドレス領域にデータが格納されている場合、空き情報は 0 (「使用」) となる。

【 0 1 0 3 】

物理エクステンツのアドレス領域がオーバプロビジョニングによりホスト装置 2 1 ~ 2 3 によって使用されない範囲となっている場合、空き情報は N U L L (「非使用領域」) となる。

10

【 0 1 0 4 】

デバイス構成部 1 0 は、新しいデバイスが追加された場合に、物理エクステンツ情報テーブル 9 を設定する。

【 0 1 0 5 】

書き込み部 1 3 は、物理エクステンツ情報テーブル 9 に対して、物理エクステンツ I D、所属デバイス I D、先頭物理アドレスに対するレコード追加及び初期値の設定を行う。書き込み部 1 3 は、物理エクステンツ情報テーブル 9 に対して、物理エクステンツ I D、所属デバイス I D、先頭物理アドレスの更新を行わない。書き込み部 1 3 は、空き情報に初期値として 1 を設定し、その後更新を行う。

20

【 0 1 0 6 】

メモリ制御装置 5 は、論理エクステンツ情報テーブル 8 と物理エクステンツ情報テーブル 8 とを参照することにより、論理アドレスに対応する物理アドレスを求める。この論理アドレスから物理アドレスへの変換は、後述する。

【 0 1 0 7 】

以下で、メモリ制御装置 5 で実行される各種の処理について説明する。

【 0 1 0 8 】

図 7 は、本実施形態に係るデバイス構成部 1 0 の処理の一例を示すフローチャートである。

【 0 1 0 9 】

30

デバイス構成部 1 0 は、ストレージ装置 3 に追加されるデバイスのリストを受け付けた場合に、ステップ S 7 0 1 において、リストの全てのデバイスに対する設定が完了したか否かを判断する。受け付けたリストは D L と表記する。

【 0 1 1 0 】

リスト D L の全てのデバイスに対する設定が完了している場合、処理は終了する。

【 0 1 1 1 】

リスト D L の全てのデバイスに対する設定が完了していない場合、ステップ S 7 0 2 において、デバイス構成部 1 0 は、未設定のデバイス D i を選択する。

【 0 1 1 2 】

ステップ S 7 0 3 において、デバイス構成部 1 0 は、デバイス D i に関するレコード追加処理を実行する。その後、処理はステップ S 7 0 1 に戻る。

40

【 0 1 1 3 】

図 8 は、本実施形態に係るデバイス構成部 1 0 のレコード追加処理の一例を示すフローチャートである。

【 0 1 1 4 】

ステップ S 8 0 1 において、デバイス構成部 1 0 は、デバイス固有情報テーブル 6 へデバイス D i に関するレコードを新規生成する。

【 0 1 1 5 】

具体的には、デバイス構成部 1 0 は、デバイス I D、最大容量、保証書込量、性能仕様を以下の方式で設定する。

50

【 0 1 1 6 】

デバイス構成部 1 0 は、デバイス D_iへコマンドを送り、デバイス D_iからのコマンドに対する応答にしたがって、デバイス I D、最大容量、保証書込量、性能仕様を設定する。コマンドは、例えば、S C S I の INQUIRY コマンドである。

【 0 1 1 7 】

デバイス構成部 1 0 は、デバイス D_iがこのようなコマンドを扱わない場合に、例えばユーザなどの人手で設定された追加されるデバイス D_iの初期値を設定する。

【 0 1 1 8 】

デバイス構成部 1 0 は、デバイス D_iの種類ごとに設定値の定義されたテーブルを参照することで、デバイス D_iの設定値を設定する。

10

【 0 1 1 9 】

デバイス構成部 1 0 は、論理容量の初期値を、最大容量の 5 0 % で設定する。しかしながら、論理容量は、最大容量に対して 5 0 % ではない他の割合で設定されてもよい。

【 0 1 2 0 】

デバイス構成部 1 0 は、現在使用容量、読み出し量、更新量、追記量、最大負荷をゼロで初期化する。

【 0 1 2 1 】

ステップ S 8 0 2 において、デバイス構成部 1 0 は、デバイス D_iの最大容量 / エクステントサイズにより、追加物理エクステント数を計算する。

【 0 1 2 2 】

ステップ S 8 0 3 において、デバイス構成部 1 0 は、追加物理エクステント数分の物理エクステントのレコード追加が完了したか否か判断する。

20

【 0 1 2 3 】

追加物理エクステント数分の物理エクステントの設定 (レコード追加) が完了した場合、処理は終了する。

【 0 1 2 4 】

追加物理エクステント数分の物理エクステントの設定が完了していない場合、ステップ S 8 0 4 において、デバイス構成部 1 0 は、物理エクステント情報テーブル 9 に物理エクステントのレコードを新規生成する。

【 0 1 2 5 】

具体的には、デバイス構成部 1 0 は、物理エクステント I D として、物理エクステント情報テーブル 9 内で一意な物理エクステント I D を設定する。デバイス構成部 1 0 は、所属デバイス I D として、デバイス D_iのデバイス I D を設定する。デバイス構成部 1 0 は、先頭物理アドレスとして、デバイス D_iに対してすでに設定したレコード数 × エクステントサイズ分のセクタ数、を設定する。デバイス構成部 1 0 は、デバイス D_iの空き情報を「空き」とする。

30

【 0 1 2 6 】

なお、ストレージ装置 3 の運用中にデバイス D_iが故障した場合、デバイス構成部 1 0 は、所属デバイス I D が故障したデバイス D_iを示す全ての論理エクステントと物理エクステントとを削除する。

40

【 0 1 2 7 】

図 9 は、本実施形態に係る論理ディスク構成部 1 1 の処理の一例を示すフローチャートである。

【 0 1 2 8 】

論理ディスク構成部 1 1 は、論理ディスクが設定される場合に、論理ディスク情報テーブル 7、論理エクステント情報テーブル 8 を更新する。論理ディスクの生成、論理ディスク I D の付与、論理容量の付与、及びパーティション分け等の設定は、例えば、人手によりすでに行われているとする。

【 0 1 2 9 】

論理ディスク構成部 1 1 は、追加される論理ディスクのリストを受け付けた場合に、ス

50

ステップS901において、リストの全ての論理ディスクに対する設定が完了したか否か判断する。追加される論理ディスクのリストはLLと表記する。

【0130】

リストLLの全ての論理ディスクに対する設定が完了している場合、処理は終了する。

【0131】

リストLLの全ての論理ディスクに対する設定が完了していない場合、ステップS902において、論理ディスク構成部11は、未設定の論理ディスクを選択する。選択された論理ディスクはLiと表記する。

【0132】

ステップS903において、論理ディスク構成部11は、論理ディスクLiに関するレコード追加処理を実行する。その後、処理はステップS901に戻る。

10

【0133】

図10は、本実施形態に係る論理ディスク構成部11によって実行される論理ディスクLiに関するレコード追加処理の一例を示すフローチャートである。

【0134】

論理ディスク構成部11は、論理ディスクLiの論理ディスクIDと論理容量とを入力する。入力される論理ディスクIDはLID、入力される論理容量はLVと表記する。論理ディスクID LIDと論理容量LVとの入力値は、論理ディスクLiが設定される場合に参照される構成情報の一部とする。

【0135】

20

ステップS1001において、論理ディスク構成部11は、論理ディスク情報テーブル7へ未設定の論理ディスクLiに関するレコードを新規生成し、レコードに、入力した論理ディスクID LIDと入力した論理容量LVとを設定する。

【0136】

ステップS1002において、論理ディスク構成部11は、論理ディスクLiの入力した論理容量LV / エクステントサイズにより、追加論理エクステント数を計算する。

【0137】

ステップS1003において、論理ディスク構成部11は、論理ディスクLiに所属する全論理エクステントの設定が完了したか否か判断する。

【0138】

30

論理ディスクLiに所属する全論理エクステントの設定が完了した場合、処理は終了する。

【0139】

論理ディスクLiに所属する全論理エクステントの設定が完了していない場合、ステップS1004において、論理ディスク構成部11は、論理エクステント情報テーブル8に、追加論理エクステントサイズ数分の論理エクステントのレコードを新規生成する。さらに、論理ディスク構成部11は、論理エクステント情報テーブル8のレコード1つに対応する論理エクステントの設定を、以下の方式で行う。

【0140】

論理ディスク構成部11は、論理エクステントIDとして、論理エクステント情報テーブル8内で一意なIDを設定する。

40

【0141】

論理ディスク構成部11は、入力した論理ディスクID LIDを、論理ディスクIDとして設定する。

【0142】

論理ディスク構成部11は、ゼロから始まり論理エクステント情報テーブル8中の同じ論理ディスクIDに所属する論理エクステントの数×エクステントサイズ分のアドレス数により、論理エクステントの先頭論理アドレスを設定する。

【0143】

論理ディスク構成部11は、所属デバイスIDを、NULLに設定する。所属デバイス

50

ＩＤの値は、データ書き込み時に書き込み部１３が設定する。

【０１４４】

論理ディスク構成部１１は、物理エクステンツＩＤを、ＮＵＬＬに設定する。物理エクステンツＩＤの値は、データ書き込み時に書き込み部１３が設定する。

【０１４５】

論理ディスク構成部１１は、読み出し量及び書き込み量を、ゼロに設定する。

【０１４６】

図１１は、本実施形態に係る読み出し部１２の処理の一例を示すフローチャートである。ただし、図１１の一部、具体的にはステップＳ１１０２は、アドレス範囲抽出部１４の処理に相当する。

10

【０１４７】

ホスト装置２１～２３のいずれかは、論理ディスクＩＤ、論理アドレス、データサイズをパラメータとする読み出しコマンドを、ストレージ装置３に送信する。ストレージ装置３は、パラメータとデバイス情報とに基づいてデータ読み出しを行い、読み出しデータをホスト装置２１～２３のいずれかに返す。

【０１４８】

読み出し部１２は、対象アドレス抽出処理、読み出し処理、テーブル更新処理を実行する。

【０１４９】

図１１において、ステップＳ１１０１～Ｓ１１０３は、対象アドレス抽出処理に相当する。

20

【０１５０】

ステップＳ１１０１において、読み出し部１２は、ホスト装置２１～２３のいずれかから、読み出し先である論理ディスクの論理ディスクＩＤ、論理アドレス、データサイズを受信する。受信した論理アドレスはＡＤＤＲと表記し、データサイズはＳＩＺＥと表記する。

【０１５１】

ステップＳ１１０２において、アドレス範囲抽出部１４は、受信した論理ディスクＩＤ、論理アドレスＡＤＤＲ、データサイズＳＩＺＥに基づいて、読み出し対象の論理エクステンツの論理エクステンツＩＤと、当該読み出し対象の論理エクステンツ内の読み出し対象の先頭アドレスと、当該読み出し対象の論理エクステンツ内のデータのサイズとの組を含むアドレス範囲のリストを計算する。読み出し対象の論理エクステンツはＥ_jと表記し、読み出し対象の先頭アドレスはＡ_jと表記し、サイズはＳ_jと表記し、アドレス範囲のリストはＬＲと表記する。

30

【０１５２】

ステップＳ１１０３において、読み出し部１２は、読み出しデータを格納するためのバッファメモリを、データサイズＳＩＺＥ分確保する。

【０１５３】

図１１において、ステップＳ１１０４～Ｓ１１１０は、読み出し処理とテーブル更新処理とに相当する。

40

【０１５４】

ステップＳ１１０４において、読み出し部１２は、リストＬＲ内から、読み出し対象の論理エクステンツＩＤ、Ｅ_jと先頭アドレスＡ_jとサイズＳ_jとの組（Ｅ_j，Ａ_j，Ｓ_j）を選択する。

【０１５５】

ステップＳ１１０５において、読み出し部１２は、論理エクステンツ情報テーブル８を参照し、読み出し対象の論理エクステンツＩＤ、Ｅ_jに対応する読み出し対象の物理エクステンツＩＤを選択する。読み出し対象の物理エクステンツＩＤはＰ_kと表記する。

【０１５６】

ステップＳ１１０６において、読み出し部１２は、物理エクステンツＩＤ、Ｐ_kの示す物

50

理エクステントにおける先頭アドレス A_j からサイズ S_j 分までのデータを読み出し、読み出しデータとしてバッファメモリに格納する。

【0157】

ステップ S_{1107} において、読み出し部 12 は、論理エクステント情報テーブル 8 を参照し、読み出し対象の論理エクステント ID_{Ej} に対応する読み出し量を、サイズ S_j 分増加する。

【0158】

ステップ S_{1108} において、読み出し部 12 は、デバイス固有情報テーブル 6 を参照し、データ読み出しの行われたデバイス Di を示すデバイス ID に対応する読み出し量を、サイズ S_j 分増加する。

10

【0159】

ステップ S_{1109} において、読み出し部 12 は、リスト LR 内から、全ての論理エクステント ID_{Ej} と先頭アドレス A_j とサイズ S_j との組 (E_j, A_j, S_j) が選択されたか否か判断する。

【0160】

リスト LR 内から、全ての組 (E_j, A_j, S_j) が選択されていない場合には、処理はステップ S_{1104} に移動する。

【0161】

リスト LR 内から全ての組 (E_j, A_j, S_j) が選択されている場合には、ステップ S_{1110} において、読み出し部 12 は、ホスト装置 21 ~ 23 のいずれかにバッファメモリの読み出しデータを送信し、処理を終了する。データは $DATA$ と表記する。

20

【0162】

図 12 は、本実施形態に係るアドレス範囲抽出部 14 の処理の一例を示すフローチャートである。この図 12 の処理は、上記図 11 のステップ S_{1102} に相当する。

【0163】

アドレス範囲抽出部 14 は、アドレス範囲を抽出する対象の論理ディスク ID_{LID} 、対象のアドレス $ADDR$ 、データサイズ $SIZE$ を入力し、論理エクステント ID_{Ej} と先頭アドレス A_j とサイズ S_j とを含むリスト LR を出力する。

【0164】

図 12 において、ステップ S_{1201} 及びステップ S_{1202} は、初期化処理に相当する。

30

【0165】

ステップ S_{1201} において、アドレス範囲抽出部 14 は、アドレス範囲のリスト LR を空リストとして初期化する。

【0166】

ステップ S_{1202} において、アドレス範囲抽出部 14 は、変数 $NS = SIZE$ とし、変数 $NA = ADDR$ とする。

【0167】

図 12 において、ステップ S_{1203} 以降は、リスト LR の生成処理に相当する。

【0168】

ステップ S_{1203} において、アドレス範囲抽出部 14 は、論理エクステント情報テーブル 8 を参照し、論理エクステント ID のうち、論理エクステント ID_{Ej} の先頭論理アドレス $NA < (E_j \text{ の先頭論理アドレス} + \text{エクステントサイズ分のセクタ数})$ 、を満たす論理エクステント ID_{Ej} を計算する。

40

【0169】

ステップ S_{1204} において、アドレス範囲抽出部 14 は、先頭アドレス $A_j = NA$ とする。

【0170】

ステップ S_{1205} において、アドレス範囲抽出部 14 は、論理エクステント ID_{Ej} の残りサイズ $RS = E_j \text{ の (終端論理アドレス} - ADDR) \times \text{セクタサイズ}$ 、とする。ア

50

ドレス範囲抽出部 14 は、 $RS > SIZE$ の場合、 $S_j = NS$ 、 $NA = 0$ とする。アドレス範囲抽出部 14 は、 $RS > SIZE$ ではない場合、 $S_j = RS$ 、 $NA = (E_j \text{ の先頭論理アドレス} + \text{エクステントサイズ分のセクタ数})$ 、とする。

【0171】

ステップ S 1206 において、アドレス範囲抽出部 14 は、リスト LR に、 (E_j, A_j, S_j) を追加する。

【0172】

ステップ S 1207 において、アドレス範囲抽出部 14 は、 $NS = NA - S_j$ とする。

【0173】

ステップ S 1208 において、アドレス範囲抽出部 14 は、 $NS = 0$ か否か判断する。

10

【0174】

$NS = 0$ ではない場合、処理はステップ S 1205 に移る。

【0175】

$NS = 0$ の場合、ステップ S 1209 において、アドレス範囲抽出部 14 は、リスト LR を、読み出し部 12 又は書き込み部 13 に返す。

【0176】

図 13 は、本実施形態に係る書き込み部 13 の処理の一例を示すフローチャートである。図 13 におけるステップ S 1302 は、上記図 12 のアドレス範囲抽出部 14 の処理に相当し、ステップ S 1304 は、寿命補正部 15 のアロケート書き込み先選択処理に相当する。

20

【0177】

ホスト装置 21 ~ 23 のいずれかは、論理ディスク ID、論理アドレス、データサイズ、書き込みデータをパラメータとする書き込みコマンドを、ストレージ装置 3 に送信する。ストレージ装置 3 は、パラメータとデバイス情報とに基づいてデータ書き込みを行う。

【0178】

図 13 において、ステップ S 1301 ~ S 1304 は、対象アドレス抽出処理に相当する。

【0179】

ステップ S 1301 において、書き込み部 13 は、ホスト装置 21 ~ 23 のいずれかから、書き込み先である論理ディスクの論理ディスク ID、論理アドレス ADDR、データサイズ SIZE、書き込みデータ DATA を受信する。書き込み部 13 は、書き込みデータ DATA を格納するためのバッファメモリを、データサイズ SIZE 分確保し、バッファメモリに書き込みデータ DATA を格納する。

30

【0180】

ステップ S 1302 において、アドレス範囲抽出部 14 は、受信した論理ディスク ID、論理アドレス ADDR、データサイズ SIZE に基づいて、書き込み対象の論理エクステントの論理エクステント ID E_j と、当該書き込み対象の論理エクステント内の書き込み対象の先頭アドレス A_j と、当該書き込み対象の論理エクステント内のデータのサイズ S_j との組 (E_j, A_j, S_j) を含むアドレス範囲のリスト LR を計算する。

【0181】

40

ステップ S 1303 において、書き込み部 13 は、論理エクステント情報テーブル 8 を参照し、リスト LR 内の論理エクステント ID E_j のうち、所属デバイス ID が設定されていない論理エクステント ID があるか否か、すなわち、論理エクステント ID と物理エクステント ID との新規の対応付けが必要か否か判断する。

【0182】

新規の対応付けが必要でない場合、処理は、ステップ S 1305 に移動する。

【0183】

新規の対応付けが必要な場合、ステップ S 1304 において、寿命補正部 15 は、アロケート書き込み先選択処理により、論理エクステント ID と一対一で対応する物理エクステント ID を設定する。

50

【0184】

図13において、ステップS1305～S1314は、書き込み処理とテーブル更新処理とに相当する。

【0185】

ステップS1305において、書き込み部13は、OFFSET = 0とする。

【0186】

ステップS1306において、書き込み部13は、リストLR内から、読み出し対象の論理エクステンツID E_jと先頭アドレスA_jとサイズS_jとの組(E_j, A_j, S_j)を選択する。

【0187】

ステップS1307において、書き込み部13は、論理エクステンツ情報テーブル8を参照し、書き込み対象の論理エクステンツID E_jに対応する書き込み対象の物理エクステンツID P_kを選択する。

【0188】

ステップS1308において、書き込み部13は、書き込み対象の物理エクステンツID P_kで示される物理エクステンツに対して、先頭アドレスA_jからサイズS_j分までに、バッファメモリのOFFSETの位置からサイズS_j分のデータを書き込む。

【0189】

ステップS1309において、書き込み部13は、OFFSET = OFFSET + S_jにより、バッファメモリの参照位置を更新する。

【0190】

ステップS1310において、書き込み部13は、論理エクステンツ情報テーブル8を参照し、書き込み対象の論理エクステンツID E_jに対応するレコードの書き込み量を、サイズS_j分増加する。

【0191】

ステップS1311において、書き込み部13は、物理エクステンツ情報テーブル9を参照し、書き込み対象の物理エクステンツID P_kに対応するレコードの空き情報を「使用中」に設定する。

【0192】

ステップS1312において、書き込み部13は、デバイス固有情報テーブル6を参照し、データ書き込みの行われたデバイスD_iの残り書き込み量を、サイズS_j分減少する。書き込み部13は、エクステンツサイズを現在使用容量に加え、デバイスD_iの現在使用容量を更新する。書き込み部13は、アップデート書き込みの場合は更新量に、アロケート書き込みの場合は追記量に書き込みデータのデータサイズを加算する。アップデート書き込みとは、物理エクステンツに対する割り当てが行われている論理エクステンツへの書き込みを指す。

【0193】

ステップS1313において、書き込み部13は、リストLR内の全ての論理エクステンツID E_jと先頭アドレスA_jとサイズS_jとの組(E_j, A_j, S_j)が選択されたか否か判断する。

【0194】

リストLR内の全ての組(E_j, A_j, S_j)が選択されていない場合、処理はステップS1306に移動する。

【0195】

リストLR内の全ての組(E_j, A_j, S_j)が選択された場合、ステップS1314において、書き込み部13は、ホスト装置13のいずれかに書き込み完了通知を送信し、処理は終了する。

【0196】

図14は、本実施形態に係る寿命補正部15のアロケート書き込み先選択処理の一例を示すフローチャートである。この図14のアロケート書き込み先選択処理は、上記図13

10

20

30

40

50

のステップ S 1 3 0 4 に相当する。

【 0 1 9 7 】

寿命補正部 1 5 によって実行されるアロケート書き込み先選択処理は、デバイス $D_0 \sim D_{N-1}$ の寿命及び性能を後述するアロケート書き込み先評価処理によって評価し、評価値が最も大きいデバイスを適切なアロケート書き込み先として選択する。

【 0 1 9 8 】

図 1 4 において、ステップ S 1 4 0 2 は、初期化処理に相当する。

【 0 1 9 9 】

ステップ S 1 4 0 1 において、寿命補正部 1 5 は、リスト L R から、論理エクステン
I D E j を 1 つ選択する。

10

【 0 2 0 0 】

ステップ S 1 4 0 2 において、寿命補正部 1 5 は、評価リスト E L を空リストとして初
期化する。

【 0 2 0 1 】

図 1 4 において、ステップ S 1 4 0 3 ~ S 1 4 0 6 は、アロケート書き込みデバイス決
定処理に相当する。

【 0 2 0 2 】

ステップ S 1 4 0 3 において、寿命補正部 1 5 は、デバイス固有情報テーブル 6 を参照
し、デバイス I D の昇順にデバイス D i の情報を選択する。

【 0 2 0 3 】

20

ステップ S 1 4 0 4 において、寿命補正部 1 5 は、後述するアロケート書き込み先評価
処理により、論理エクステン
I D E j とデバイス D i の情報とに基づいて、評価値を計
算する。計算した評価値は F と表記する。そして、寿命補正部 1 5 は、デバイス D i の情
報と評価値 F との組をリスト E L に追加する。

【 0 2 0 4 】

ステップ S 1 4 0 5 において、寿命補正部 1 5 は、デバイス固有情報テーブル 6 から全
てのデバイス D i の情報が選択されたか否か判断する。

【 0 2 0 5 】

デバイス固有情報テーブル 6 から全てのデバイス D i の情報が選択されていない場合、
処理はステップ S 1 4 0 3 に移動する。

30

【 0 2 0 6 】

デバイス固有情報テーブル 6 から全てのデバイス D i の情報が選択された場合、ステッ
プ S 1 4 0 6 において、寿命補正部 1 5 は、リスト E L から、評価値 F が最大であるデバ
イス D i のデバイス I D を選択する。

【 0 2 0 7 】

図 1 4 において、ステップ S 1 4 0 7 ~ S 1 4 0 9 は、論理エクステン
I D と物理エ
クステン
I D との対応付け処理に相当する。

【 0 2 0 8 】

ステップ S 1 4 0 7 において、寿命補正部 1 5 は、選択されたデバイス I D で示される
デバイス D i に所属し、かつ、空き情報「空き」に対応する物理エクステン
I D から、
I D の昇順に、1 つの物理エクステン
I D P k を選択する。このステップ S 1 4 0 7 の
処理は、必要な情報として、物理エクステン
I D P k とを格納し、論理エクステン
I D E j と物理エクステン
I D P k とを対応付ける。

40

【 0 2 0 9 】

ステップ S 1 4 0 8 において、寿命補正部 1 5 は、論理エクステン
I D E j に対応するレコードに、デバイス D i を示すデバイス
I D と物理エクステン
I D P k とを格納し、論理エクステン
I D E j と物理エクステン
I D P k とを対応付ける。

【 0 2 1 0 】

50

ステップS 1 4 0 9において、寿命補正部 1 5は、リストL Rから全ての論理エクステントIDが選択されたか否か判断する。

【0 2 1 1】

リストL Rから全ての論理エクステントIDが選択されていない場合、処理はステップS 1 4 0 1に移動する。

【0 2 1 2】

リストL Rから全ての論理エクステントIDが選択された場合、処理は終了する。

【0 2 1 3】

図1 5は、本実施形態に係る寿命補正部 1 5のアロケート書き込み先評価処理の一例を示すフローチャートである。この図1 5のアロケート書き込み先評価処理は、上記図1 4のステップS 1 4 0 4に相当する。この図1 5において、ステップS 1 5 0 1～S 1 5 0 4の実行順序は、自由に変更可能であり、並列に実行されてもよい。

【0 2 1 4】

寿命補正部 1 5は、デバイスD_iのデバイス情報、論理エクステントID E_j、重みW₁～W₄($W_i = 1, i = 0 \sim W_i$)を入力し、評価値Fを出力する。例えば、重みW₁～W₄は、ストレージ装置3のパラメータに含まれるポリシー情報としてユーザが設定可能な値とする。例えば、重みW₁～W₄の初期値は、すべて0.25としてもよい。例えば、重みW₁～W₄などのようなデバイス又は論理エクステントに関わる情報は、デバイス固有情報テーブル6又は論理エクステント情報テーブル8に格納されていてもよい。

【0 2 1 5】

ステップS 1 5 0 1において、寿命補正部 1 5は、性能値V₁=W₁×(デバイスD_iの性能仕様-デバイスD_iの現在IOPS)/デバイスD_iの性能仕様、を計算する。例えば、デバイスD_iの現在IOPSは、このアロケート書き込み先評価処理を実行中に計測されたIOPS値とする。

【0 2 1 6】

ステップS 1 5 0 2において、寿命補正部 1 5は、平準化後保証書き込み量EW=(残り寿命/ストレージ装置3全体のデバイスの残り寿命平均)×デバイスD_iの保証書き込み量)、を計算し、寿命値V₂=W₂×(EW-(デバイスD_iに含まれる論理エクステントの書き込み量))/EW、を計算する。

【0 2 1 7】

ステップS 1 5 0 3において、寿命補正部 1 5は、容量値V₃=W₃×(デバイスD_iの論理容量-デバイスD_iの現在使用容量-エクステントサイズ)/デバイスD_iの論理容量、を計算する

ステップS 1 5 0 4において、寿命補正部 1 5は、追記率AR=デバイスD_iの追記量/(デバイスD_iの追記量+デバイスD_iの読み出し量+デバイスD_iの更新量)、を計算し、追記値V₄=W₄×(1-AR)、を計算する。

【0 2 1 8】

ステップS 1 5 0 5において、寿命補正部 1 5は、評価関数F=V₁+V₂+V₃+V₄に基づいて、評価値Fを計算し、評価値Fを返す。もし、V₁<0又はV₃<0の場合はF=0とする。

【0 2 1 9】

上記のような評価関数を用いて、アロケート書き込み先を管理することにより、以下の効を得ることができる。

【0 2 2 0】

例えば、デバイスの性能仕様に対して現在の負荷が小さく、負荷に余裕が有るほど、デバイスに対する評価値Fは高くなる。したがって、負荷に余裕のあるデバイスに対してエクステントがアロケート書き込みされる割合を増やすことができる。

【0 2 2 1】

デバイスに対する書き込み量が保証書き込み量に対して少ない場合に、このデバイスの寿命に余裕が生じる。書き込み量が少ない場合、評価関数によって、寿命に余裕のあるデ

10

20

30

40

50

バスの評価値Fは高くなる。したがって、寿命に余裕のあるデバイスに対してエクステントがアロケート書き込みされる割合を増やすことができ、寿命の少ないデバイスへエクステントがアロケート書き込みされる割合を減らすことができる。

【0222】

アロケート書き込み先評価処理における寿命値の計算では、保証書き込み量に対して、デバイスの寿命とデバイス全体の残り寿命平均との比を、掛けている。これにより、ストレージ装置3に既に設置済みのデバイス全体と比較し、残り寿命が大きい新規設置のデバイスは、評価値Fが高くなる。このため、新規設置のデバイスに対してエクステントがアロケート書き込みされる割合を増やすことができる。一方、残り寿命が少ないデバイスへエクステントがアロケート書き込みされる割合は減少し、デバイスの寿命平準化を行うことができる。

10

【0223】

アロケート書き込み先評価処理では、デバイスの最大容量に対して現在使用容量が少なほど、デバイスに対する評価値Fが高くなり、アロケート書き込みされ易くなる。したがって、現在使用容量が論理容量に近いデバイスへのアロケート書き込みの割合を減らすことができる。

【0224】

アロケート書き込み先評価処理では、追記率の高いデバイスの評価値Fは低くなる。したがって、現在のアロケート書き込みの多いデバイスに対してアロケート書き込みが行われる割合は減少し、過剰なアロケート書き込みによる容量圧迫を抑え、デバイス間で容量平準化を行うことができる。

20

【0225】

ここで、図16乃至図20を用いてアロケート書き込み先選択処理の具体例を説明する。

【0226】

図16は、デバイス固有情報テーブル6の一例を示すデータ構造図である。

【0227】

デバイス固有情報テーブル6は、デバイスID D1～D3を含む。

【0228】

図17は、論理エクステント情報テーブル7の一例を示すデータ構造図である。

30

【0229】

論理エクステント情報テーブル7は、5つの論理エクステントID E1～E5を含む。本実施形態において、論理エクステントID D1～D3のそれぞれは、一般化して論理エクステントID E_jと表記する。

【0230】

図18は、物理エクステント情報テーブル9の一例を示すデータ構造図である。

【0231】

物理エクステント情報テーブル8は、6つの物理エクステントID P1～P6を含む。本実施形態において、物理エクステントID P1～P6のそれぞれは、一般化して物理エクステントID P_kと表記する。

40

【0232】

例えば、論理エクステント情報テーブル8における論理エレメントID E5を新たにアロケート書き込みする場合について説明する。

【0233】

まず、寿命補正部15は、上記図14のステップS1403からステップS1405までのループにより、各デバイスD1～D3の評価値Fを計算する。

【0234】

図19は、デバイスと評価値との関係の一例を示す図である。

【0235】

寿命補正部15は、上記図14のステップS1406において、評価値Fが最大のデバ

50

イス D3 に対してアロケート書き込みを行うことを決定する。

【 0 2 3 6 】

寿命補正部 15 は、上記図 14 のステップ S 1 4 0 7 及びステップ S 1 4 0 8 において、論理エクステント I D E5 を、評価値 F が最大のデバイス I D D3 の空き状態の物理エクステント I D P6 と対応付け、論理エクステント情報テーブル 8 を更新する。

【 0 2 3 7 】

図 20 は、更新後の論理エクステント情報テーブル 8 の一例を示すデータ構造図である。

【 0 2 3 8 】

論理エクステント I D E5 に対応するレコードに、所属デバイス I D として、デバイス I D D3 が格納され、物理エクステント I D として、物理エクステント I D P6 が格納される。

10

【 0 2 3 9 】

図 21 は、本実施形態に係る寿命補正部 15 の寿命補正処理の一例を示すフローチャートである。ただし、図 21 におけるステップ S 2 1 1 8 は、後述する設定変更部 16 の処理に相当する。

【 0 2 4 0 】

寿命補正部 15 は、各デバイス $D_0 \sim D_{N-1}$ 内の論理エクステントごとに、デバイス固有情報テーブル 6 に基づいて評価を行い、各論理エクステントを評価値の高いデバイスに再配置する。

20

【 0 2 4 1 】

また、寿命補正部 15 は、各デバイス $D_0 \sim D_{N-1}$ の論理容量を、デバイス固有情報テーブル 6 に基づいて計算されたオーバプロビジョニングの設定に、変更する。

【 0 2 4 2 】

寿命補正処理は、ストレージ装置 3 内のデバイス $D_0 \sim D_{N-1}$ の寿命を平準化する。

【 0 2 4 3 】

寿命補正処理では、まず、初期化処理によって、再配置を行う候補となる論理エクステントのリストを初期化する。

【 0 2 4 4 】

次に、寿命補正処理では、移動候補となる論理エクステントの追加処理によって、保証書き込み量を上回るデバイス上の論理エクステントを 1 つ又は複数選択し、移動候補とする。

30

【 0 2 4 5 】

次に、寿命補正処理では、1 つ又は複数の論理エクステントの再配置を行い、各デバイス $D_0 \sim D_{N-1}$ の寿命及び性能の無駄をなくす。

【 0 2 4 6 】

そして、寿命補正処理では、最後に、各デバイス $D_0 \sim D_{N-1}$ のオーバプロビジョニングの設定を、後述する設定変更部 16 によって実行されるオーバプロビジョニングの設定評価処理によって適切な値に設定し、処理を終了する。

【 0 2 4 7 】

40

図 21 において、ステップ S 2 1 0 1 は、初期化処理に相当する。

【 0 2 4 8 】

ステップ S 2 1 0 1 において、寿命補正部 15 は、再配置を行う候補となる論理エクステントのリストを空リストとして生成する。再配置を行う候補となる論理エクステントのリストは C L と表記する。

【 0 2 4 9 】

図 21 において、ステップ S 2 1 0 2 ~ S 2 1 0 8 は、移動候補となる論理エクステントの追加処理に相当する。

【 0 2 5 0 】

ステップ S 2 1 0 2 において、寿命補正部 15 は、デバイス固有情報テーブル 6 からデ

50

バイスIDの昇順で、デバイスID D_i を選択する。

【0251】

ステップS2103において、寿命補正部15は、論理エクステント情報テーブル8を参照し、所属デバイスID D_i に対応付けられている論理エクステントID群を、論理エクステント情報テーブル8から全て選択する。

【0252】

ステップS2104において、寿命補正部15は、選択された論理エクステントID群を書き込み量の降順にソートする。

【0253】

ステップS2105において、寿命補正部15は、論理エクステントID群の書き込み量の総和を累積書き込み量AWとする。

10

【0254】

ステップS2106において、寿命補正部15は、累積書き込み量AW < 保証書き込み量が否か判断する。

【0255】

累積書き込み量AW < 保証書き込み量ではない場合、ステップS2107において、寿命補正部15は、書き込み量が最大の論理エクステントIDを、リストCLに追加し、論理エクステントID群から省き、処理はステップS2105へ移動する。

【0256】

累積書き込み量AW < 保証書き込み量の場合、ステップS2108において、寿命補正部15は、上記のステップS2102において、デバイス固有情報テーブル6から全てのデバイスID D_i が選択されたか否か判断する。

20

【0257】

デバイス固有情報テーブル6から全てのデバイスID D_i が選択されていない場合、処理はステップS2102に移動する。

【0258】

デバイス固有情報テーブル6から全てのデバイスID D_i が選択された場合、処理はステップS2109に移動する。

【0259】

図21におけるステップS2109～S2117は、寿命補正部15の論理エクステント単位の再配置先選択処理及び再配置処理に相当する。

30

【0260】

ステップS2109において、寿命補正部15は、リストCL内の論理エクステントIDを、書き込み量の降順にソートする。

【0261】

ステップS2110において、寿命補正部15は、リストCL内から書き込み量が大きい順で論理エクステントID E_j を選択する。

【0262】

ステップS2111において、寿命補正部15は、初期化された空の評価値リストELを生成する。

40

【0263】

ステップS2112において、寿命補正部15は、デバイス固定情報テーブル6から昇順でデバイスID D_i を選択する。

【0264】

ステップS2113において、寿命補正部15は、選択された論理エクステントID E_j とデバイスID D_i と評価関数とに基づいて、評価値Fを計算し、デバイスID D_i と評価値Fの組(デバイスID D_i , 評価値F)を評価値リストELに追加する。

【0265】

ステップS2114において、寿命補正部15は、ステップS2112において、全てのデバイスIDが選択されたか否か判断する。

50

【0266】

全てのデバイスIDが選択されていない場合、処理はステップS2112に移動する。これにより、全デバイスのFが計算される。

【0267】

全てのデバイスIDが選択された場合、ステップS2115において、リストEL内で、評価値Fが最大のデバイスID D_i を選択する。

【0268】

ステップS2116において、寿命補正部15は、後述する論理エクステンツの再配置処理を行い、論理エクステンツEjを選択された D_i に再配置する。

【0269】

ステップS2117において、寿命補正部15は、上記ステップS2110において、リストCL内から全ての論理エクステンツID E_j が選択されたか否か判断する。

【0270】

リストCL内から全ての論理エクステンツIDが選択されていない場合、処理はステップS2110に移動する。すなわち、リストCLから全論理エクステンツIDが選択されるまでステップS2110～ステップS2117が繰り返される。

【0271】

リストCL内から全ての論理エクステンツIDが選択された場合、ステップS2118において、設定変更部16は、オーバプロビジョニングの設定評価処理を実行し、全デバイス $D_0 \sim D_{N-1}$ のオーバプロビジョニングの設定を適切な値に変更する。

【0272】

なお、上記の図21の処理のうちステップS2109～S2118の処理は、適宜変更可能である。

【0273】

図22は、本実施形態に係る寿命補正部15の寿命補正処理の変形例を示すフローチャートである。上記図21では評価値Fを計算するごとに論理エクステンツを再配置する。これに代えて、この図22では、全ての論理エクステンツとデバイスとの組み合わせに対する評価値Fを計算し、その全ての論理エクステンツとデバイスとの組み合わせに対する評価値Fに基づいて、再配置の優先順位を判断する。

【0274】

ステップS2209において、寿命補正部15は、評価値を格納するための評価値リストを空リストとして初期化する。評価値リストは、ELと表記する。

【0275】

ステップS2210において、寿命補正部15は、リストCL内の論理エクステンツID E_j を選択する。

【0276】

ステップS2211において、寿命補正部15は、デバイス固定情報テーブル6から昇順でデバイスID D_i を選択する。

【0277】

ステップS2212において、寿命補正部15は、選択された論理エクステンツID E_j とデバイスID D_i と評価関数とに基づいて、評価値Fを計算し、デバイスID D_i と評価値Fの組(デバイスID D_i , 評価値F)を評価値リストELに追加する。

【0278】

ステップS2213において、寿命補正部15は、ステップS2211において、全てのデバイスIDが選択されたか否か判断する。

【0279】

全てのデバイスIDが選択されていない場合、処理はステップS2211に移動する。これにより、全デバイスの評価値Fが計算される。

【0280】

全てのデバイスIDが選択された場合、寿命補正部15は、ステップS2214におい

10

20

30

40

50

て、リストC L内の全ての論理エクステントIDが選択されたか否か判断する。

【0281】

リストC L内の全ての論理エクステントIDが選択されていない場合、処理はステップS 2 2 1 0に移動する。これにより、リストC L内の全ての論理エクステントIDが選択されるまでステップS 2 2 1 0～S 2 2 1 4までが繰り返される。

【0282】

図22におけるステップS 2 2 1 5～S 2 2 2 1は、寿命補正部15によって実行される論理エクステント単位の再配置先選択処理及び再配置処理に相当する。

【0283】

ステップS 2 2 1 5において、寿命補正部15は、デバイス固有情報テーブル6から、昇順でデバイスDiを選択する。

10

【0284】

ステップS 2 2 1 6において、寿命補正部15は、デバイスID Diに対するオーバープロビジョニングの設定変更処理を行う。設定変更処理の入力値となる論理容量は、デバイスID Diで示されるデバイスに再配置される論理エクステント分の容量である。すなわち、設定変更の入力値となる論理容量は、例えば、今の論理容量と、評価値リストE L内のデバイスID Diを含む組の個数×エクステントサイズ分の容量との和である。

【0285】

ステップS 2 2 1 7において、寿命補正部15は、ステップS 2 2 1 5において、全てのデバイスIDが選択されたか否か判断する。

20

【0286】

全てのデバイスIDが選択されていない場合、処理はステップS 2 2 1 5に移動する。これにより、全デバイスに対してステップS 2 2 1 6の処理が実行される。

【0287】

全てのデバイスIDが選択された場合、ステップS 2 2 1 8において、寿命補正部15は、評価値リストE Lを評価値Fの降順にソートする。

【0288】

ステップS 2 2 1 9において、寿命補正部15は、評価値リストE L内で評価値Fが最大の論理エクステントID EjとデバイスID Diを選択する。

【0289】

30

ステップS 2 2 2 0において、寿命補正部15は、選択された論理エクステントID Ejを、選択されたデバイスID Diで示されるデバイスに再配置する。

【0290】

ステップS 2 2 2 1において、寿命補正部15は、上記ステップS 2 2 1 9において、リストE L内から全ての論理エクステントIDとデバイスIDとが選択されたか否か判断する。

【0291】

リストE L内から全ての論理エクステントIDとデバイスIDとが選択されていない場合、処理はステップS 2 2 1 9に移動する。すなわち、リストE Lから全ての論理エクステントIDとデバイスIDとが選択されるまでステップS 2 2 1 9～ステップS 2 2 2 1が繰り返される。

40

【0292】

リストE L内から全ての論理エクステントIDとデバイスIDとが選択された場合、ステップS 2 2 2 2において、設定変更部16は、オーバープロビジョニングの設定評価処理を実行し、全デバイスD₀～D_{N-1}のオーバープロビジョニングの設定を適切な値に変更する。

【0293】

例えば、上記図21の処理では、論理エクステントの再配置時に、評価値Fが大きいデバイスの残り容量が不足し、評価値Fが大きいデバイスが再配置先とならない場合がある。すると、例えば残り寿命が多いので再配置先としたいデバイスであっても、容量不足の

50

ため再配置先として選択されない場合がある。しかしながら、上記図 22 の処理では、各デバイスに、再配置される論理エクステント分の論理容量を予め確保することにより、容量不足によって評価値 F が大きいデバイスに再配置されないことを防止することができる。

【0294】

ここで、図 23 乃至図 27 を用いて寿命補正処理の具体例を説明する。寿命補正処理前の各テーブルのレコードは、例えば、上述の図 16 乃至図 18、図 20 と同様である。

【0295】

ステップ S 2101 において、寿命補正部 15 は、再配置候補のリスト CL を初期化する。

10

【0296】

ステップ S 2102 ~ S 2104 において、寿命補正部 15 は、デバイス ID Di ごとに、デバイス ID Di に対応する論理エクステント ID Ej をソートする。

【0297】

ステップ S 2105 において、寿命補正部 15 は、デバイス ID Di ごとに、論理エクステントの書き込み量の総和である累積書き込み量 AW を計算する。

【0298】

図 23 は、デバイス ID ごとの累積書き込み量を例示する図である。

【0299】

この図 23 では、デバイス ID D1 で示されるデバイスが、保証書き込み量より累積書き込み量が多い。したがって、このデバイス ID D1 に対する処理は、ステップ S 2106 からステップ S 2107 へ移動し、寿命補正部 15 は、候補リスト CL に、デバイス ID D1 に対応付けられている論理エクステントのうち、最大の書き込み量に対応する論理エクステント ID を追加する。

20

【0300】

寿命補正部 15 は、再びステップ S 2105 で累積書き込み量を計算する。この場合、ステップ S 2106 において、累積書き込み量 AW が保証書き込み量を下回るため、処理はステップ S 2108 に移動する。同様の処理が全てのデバイス ID に対して実行される。この図 23 の例では、リスト CL は論理エクステント ID E2 のみを含む。この例では、リスト CL 内に 1 つの論理エクステント ID しか含まれないため、ステップ S 2109 におけるソートは必要ない。ステップ S 2110 において、寿命補正部 15 は、論理エクステント ID E2 を選択する。

30

【0301】

ステップ S 2112 ~ S 2114 のループで、寿命補正部 15 は、各デバイス ID に対応する評価値 F を、再配置先評価処理によって計算する。

【0302】

図 24 は、デバイス ID ごとの評価値 F を例示する図である。この図 24 において、評価値 F が最大なのはデバイス ID D3 である。

【0303】

図 25 は、再配置後の論理エクステント情報テーブル 8 の一例を示すデータ構造図である。

40

【0304】

寿命補正部 15 は、ステップ S 2115 で、デバイス ID D3 を選択し、ステップ S 2116 において、論理エクステント ID E2 で示される論理エクステントを、デバイス ID D3 で示されるデバイス内の、物理エクステント ID P5 で示される物理エクステントに再配置する。

【0305】

図 26 は、再配置後のデバイス固有情報テーブル 6 の一例を示すデータ構造図である。寿命補正部 15 は、デバイス固有情報テーブル 6 のうち、デバイス ID D1 に対応する現在使用容量を減少させる。また、寿命補正部 15 は、デバイス固有情報テーブル 6 のうち

50

、デバイスID D3に対応する現在使用容量を増加させ、残り書き込み量を増加させ、更新量を増加させる。

【0306】

寿命補正部15は、ステップS2118において、オーバプロビジョニングの設定評価処理を実行し、全デバイスD₀～D_{N-1}のオーバプロビジョニングの設定を変更する。

【0307】

図27は、オーバプロビジョニングの設定評価処理によって得られる評価結果の例を示す図である。

【0308】

評価結果は、デバイスID D1～D3ごとに、性能値V1、平準化後保証書き込み量EW、累積書き込み量AW、寿命値V2、容量値V3、追記率AR、追記値V4、評価値F、容量増加量又は容量減少量V、計算前論理容量、計算後論理容量、計算後論理容量と計算前論理容量との差分を含む。

【0309】

計算前論理容量は、各デバイスの元の論理容量である。計算後論理容量は、オーバプロビジョニング設定の変更後の論理容量である。この図27の評価結果においては、書き込みの多いデバイスID D1は、あまり論理容量が増えず、アロケート書き込みされにくくなる。

【0310】

図28は、本実施形態に係る寿命補正部15の再配置処理の一例を示すフローチャートである。再配置処理において、寿命補正部15は、論理エクステントID E_jと再配置先のデバイスID D_iとを入力し、エクステントを再配置する。

【0311】

図28において、ステップS2801～S2806は、再配置前の元のデバイスからのデータの読み出し処理と配置先のデバイスへの書き込み処理に相当する。

【0312】

ステップS2801において、寿命補正部15は、論理エクステント情報テーブル8を参照し、論理エクステントID E_jに対応する物理エクステントID P_kを選択する。

【0313】

ステップS2802において、寿命補正部15は、物理エクステント情報テーブル9を参照し、物理エクステントID P_kに対応する元のデバイスのデバイスIDを選択する。

【0314】

ステップS2803において、寿命補正部15は、物理エクステントID P_kで示される物理エクステントからデータを読み出し、バッファメモリに格納する。

【0315】

ステップS2804において、寿命補正部15は、物理エクステント情報テーブル9を参照し、配置先のデバイスID D_iに対応し、かつ、空き情報が「空き」を示す任意の再配置先の物理エクステントIDを選択する。

【0316】

ステップS2805において、寿命補正部15は、選択された再配置先の物理エクステントIDで示される物理エクステントへ、バッファメモリに格納されているデータを書き込む。

【0317】

ステップS2806において、寿命補正部15は、物理エクステントID P_kで示される物理エクステントのデータを消去する。元のデバイスがSATA (Serial ATA) に準拠している場合には、trimコマンドを用いてデータが消去される。元のデバイスがSAS (Serial Attached SCSI) に準拠している場合には、unmapコマンドを用いてデータが消去される。

【0318】

図28において、ステップS2807～S2809は、テーブル更新処理に相当する。

このステップS 2 8 0 7 ~ S 2 8 0 9の実行順序は、自由に変更可能であり、並列に実行されてもよい。

【 0 3 1 9 】

ステップS 2 8 0 7において、寿命補正部 1 5は、論理エクステント情報テーブル 8を参照し、論理エクステントID E jに対応する所属デバイスIDを再配置先のデバイスを示すデバイスIDに変更し、論理エクステントID E jに対応する物理エクステントIDを再配置先の物理エクステントを示す物理エクステントIDに変更する。

【 0 3 2 0 】

ステップS 2 8 0 8において、寿命補正部 1 5は、物理エクステント情報テーブル 9を参照し、再配置前の元の物理エクステントID P kに対応する空き情報を「空き」に変更し、再配置先の物理エクステントIDに対応する空き情報を「使用」に変更する。

10

【 0 3 2 1 】

ステップS 2 8 0 9において、寿命補正部 1 5は、デバイス固有情報テーブル 6を参照し、元のデバイスを示すデバイスIDに対応する現在使用容量をエクステントサイズ分減少し、配置先のデバイスID D iに対応する現在使用容量をエクステントサイズ分増加し、配置先のデバイスID D iの残り書き込み量をエクステントサイズ分減少する。

【 0 3 2 2 】

図 2 9は、本実施形態に係る寿命補正部 1 5の再配置先評価処理の一例を示すフローチャートである。

【 0 3 2 3 】

20

寿命補正部 1 5の再配置先評価処理は、寿命補正部 1 5のアロケート書き込み先評価処理と類似する。

【 0 3 2 4 】

ステップS 2 9 0 1において、寿命補正部 1 5は、性能値 $V1 = W1 \times (\text{デバイス } D_i \text{ の性能仕様} - \text{デバイス } D_i \text{ の最大負荷}) / \text{デバイス } D_i \text{ の性能仕様}$ 、を計算する。このステップS 2 9 0 1は、デバイスD iの最大負荷を用いる点で、デバイスD iのI O P Sを用いるアロケート書き込み先評価処理のステップS 1 5 0 1と相違している。

【 0 3 2 5 】

ステップS 2 9 0 2において、寿命補正部 1 5は、上記アロケート書き込み先評価処理のステップS 1 5 0 1と同様に、寿命値V 2を計算する。

30

【 0 3 2 6 】

ステップS 2 9 0 3において、寿命補正部 1 5は、容量値 $V3 = W3 \times (\text{デバイス } D_i \text{ の最大容量} - \text{デバイス } D_i \text{ の現在使用容量} - \text{エクステントサイズ}) / \text{デバイス } D_i \text{ の論理容量}$ 、を計算する。上記アロケート書き込み先評価処理のステップS 1 5 0 3において、デバイスD iの論理容量を用いていた部分が、このステップS 2 9 0 3では、デバイスD iの最大容量を用いている。アロケート書き込み先評価処理では、現在の論理容量を超えて論理エクステントをデバイスに新たに配置することができないため、デバイスD iの論理容量とデバイスD iの現在使用容量が等しい場合は評価値をゼロとしている。一方、寿命補正処理では、デバイスD iの論理容量をオーバプロビジョニングの設定変更により後から増加可能であるため、デバイスD iの論理容量を超えて論理エクステントを再配置することが可能である。ただし、最大容量を超えた再配置はできない。

40

【 0 3 2 7 】

ステップS 2 9 0 4において、寿命補正部 1 5は、追記値V 4を計算する。

【 0 3 2 8 】

ステップS 2 9 0 5において、寿命補正部 1 5は、評価関数 $F = V1 + V2 + V3 + V4$ に基づいて、評価値Fを計算し、評価値Fを返す。

【 0 3 2 9 】

このような再配置先評価処理を実行することにより、アロケート書き込み先評価処理とほぼ同様の効果を得ることができる。

【 0 3 3 0 】

50

図30は、本実施形態に係る設定変更部16によって実行されるオーバプロビジョニングの設定変更処理の一例を示すフローチャートである。

【0331】

設定変更部16は、デバイスID Diと論理容量LVとを入力する。

【0332】

ステップS3001において、設定変更部16は、デバイスDiへオーバプロビジョニングの設定変更をSCSIコマンドなどで指示し、デバイスDiで示されるデバイスの論理容量をLVの値に変更する。

【0333】

ステップS3002において、設定変更部16は、デバイスDiで示されるデバイスの論理容量が増加するか否か判断する。

10

【0334】

デバイスDiの論理容量が増加する場合、処理はステップS3003に移動し、デバイスDiの論理容量が増加しない場合、処理はステップS3004に移動する。

【0335】

ステップS3003において、設定変更部16は、物理エクステント情報テーブル9を参照し、容量増加量 $V = LV - \text{現在のデバイスDiの論理容量}$ を計算する。設定変更部16は、デバイス固有情報テーブル6を参照し、デバイスDiの論理容量を得る。設定変更部16は、 $(\text{容量増加量} V / \text{エクステントサイズ})$ 個の、空き情報が「非使用領域」となっている物理エクステント群を物理エクステントIDの昇順に選択する。設定変更部16は、空き情報が「非使用領域」となっている物理エクステント群の空き情報を「空き」に設定する。

20

【0336】

ステップS3004において、設定変更部16は、物理エクステント情報テーブル9を参照し、容量減少量 $V = \text{現在のデバイスDiの論理容量} - \text{論理容量} LV$ を計算する。設定変更部16は、デバイス固有情報テーブル6を参照し、デバイスDiの論理容量を得る。設定変更部16は、 $(V / \text{エクステントサイズ})$ 個の、空き情報が「空き」となっている物理エクステント群を物理エクステントIDの昇順に選択する。設定変更部16は、空き情報が「空き」となっている物理エクステント群の空き情報を「未使用領域」に設定する。

30

【0337】

ステップS3005において、設定変更部16は、デバイス固有情報テーブル6を参照し、デバイスID Diに対応する論理容量を、論理容量LVに変更する。

【0338】

図31は、本実施形態に係る設定変更部16によって実行されるオーバプロビジョニングの設定評価処理（余剰容量評価処理）の一例を示すフローチャートである。

【0339】

オーバプロビジョニングの設定評価処理は、上記図21のステップS2119及び上記図22のステップS2222などで実行される。

【0340】

設定評価処理は、エクステントを再配置した後に、寿命が平準化されるように各デバイスの論理容量を設定する。

40

【0341】

ステップS3101において、設定変更部16は、オーバプロビジョニングの設定を変更する対象のデバイスを示すデバイスID Diを選択する。

【0342】

ステップS3102において、設定変更部16は、デバイスID Diの最大容量の特定の割合（例えば10%）を基準余剰容量として計算する。設定変更部16は、論理容量 $LV = \text{デバイスDiの現在使用容量} + \text{基準余剰容量} + \text{デバイスDiの追記量} \times 1.2$ 、を計算する。ここで、追記量を1.2倍しているのは、追記量が現状から20%上昇しても余裕

50

のある論理容量 L_V を計算するためである。論理容量 L_V は、最大容量を超えない範囲とする。この論理容量 L_V の計算式では、基準余剰容量を加えることで、容量不足となることを防止する。また、論理容量 L_V の計算式では、追記量を加えることで、寿命補正処理後に、同程度のアロケート書き込みが発生しても容量不足になることを防止する。

【0343】

S3103において、設定変更部16は、デバイスID D_i と論理容量 L_V とに基づいて、上記図30で説明されたオーバープロビジョニングの設定変更処理を行う。

【0344】

S3104において、設定変更部16は、全てのデバイスのオーバープロビジョニングの設定が実行されたか否か判断する。

10

【0345】

全てのデバイスのオーバープロビジョニングの設定が実行されていない場合、すなわち、全てのデバイスの再配置が完了していない場合、処理はステップS3101に移動し、上記の処理を繰り返す。

【0346】

全てのデバイスのオーバープロビジョニングの設定が実行された場合、処理は終了する。

【0347】

このように、設定評価処理は、アロケート書き込み量（追記量）の多いデバイスの論理容量を上げ、これによりアロケート書き込み量の多いデバイスが容量不足のためにアロケート書き込み先に選択されないことを防止することができる。

20

【0348】

以上説明した本実施形態においては、個々のデバイス $D_0 \sim D_{N-1}$ に対して、評価値 F に基づいてアロケート書き込み及び再配置、オーバープロビジョニングの設定を行い、ストレージ装置3におけるデバイス $D_0 \sim D_{N-1}$ の寿命、容量、性能を平準化することができる。

【0349】

本実施形態においては、アロケート書き込み先の選択により、残り寿命の少ないデバイス、単位時間の書き込み量が保証書き込み量に近いデバイスに対して、アロケート書き込みが発生することを抑制することができ、デバイス $D_0 \sim D_{N-1}$ の寿命を長くすることができる。

30

【0350】

本実施形態においては、書き込み量の多いエクステントを、評価関数を用いて残り寿命が大きいデバイス、書き込み量が少ないデバイスに再配置することができ、寿命の少ないデバイスに対する将来的な書き込み量を抑えることができる。

【0351】

本実施形態においては、未使用のデバイスの追加又は故障したデバイスの交換によりストレージ装置3内のデバイスの寿命が不均等になった場合であっても、アロケート書き込み先評価処理、再配置先評価処理、オーバープロビジョニングの設定評価処理により、寿命の短いデバイスへの書き込みを抑制し、寿命の長いデバイスへの書き込みを増加させ、ストレージ装置3全体で寿命を平準化することができる。

40

【0352】

本実施形態においては、アロケート書き込み先評価処理、再配置先評価処理、オーバープロビジョニングの設定評価処理により、負荷が高いデバイスへのアロケート書き込みを減らし、負荷の少ないデバイスのアロケート書き込みを増やし、ストレージ装置3全体で性能平準化を行うことができる。新規データの書き込みがある場合、新規データに対しては短期間に読み出し又は書き込みが発生しやすい傾向がある。本実施形態では、負荷が高いデバイスへの新規データの書き込みを減らすことができ、これにより負荷が高いデバイスに対する書き込み及び読み出しを抑制することができる。

【0353】

50

本実施形態においては、評価値 F に基づいてデータの再配置が実行される。これにより、本実施形態においては、高負荷のデバイス、過剰に書き込まれているデバイス、残り寿命の短いデバイスへの将来的な負荷、書き込みを減らすことができ、これとは逆に、低負荷のデバイス、書き込みの少ないデバイス、残り寿命に余裕のあるデバイスに対する書き込みを増やすことができ、性能平準化及び寿命平準化を行うことができる。

【 0 3 5 4 】

本実施形態においては、デバイスの論理容量が変更される。例えば、一度も書き込みのされていない領域に対する書き込みを新規書き込みとする。本実施形態においては、高負荷のデバイス、過剰に書き込まれているデバイス、残り寿命の短いデバイスの論理容量を減らすことができ、評価値 F を下げることで新規書き込みを抑制することができる。これとは逆に、本実施形態においては、低負荷のデバイス、書き込みの少ないデバイス、残り寿命に余裕のあるデバイスの論理容量を上げることができ、評価値 F を上げ、新規書き込みを増やすことができる。このように、デバイスの論理容量の増減が評価値 F に影響を与え、性能平準化及び寿命平準化を実現することができる。

10

【 0 3 5 5 】

本実施形態においては、新規書き込み時に評価値 F に基づいて書き込み先を選択することで、寿命平準化及び性能平準化を行うことができる。本実施形態では、例えば、書き込み量が多い場合に評価値を下げることで、基準値より過剰に書き込まれているデバイスに対する新規書き込みの割合を減らすことができる。これとは逆に、本実施形態では、書き込み量の少ないデバイスの評価値 F を上げ、新規書き込み先となる割合を増やし、ストレージ装置 3 の寿命平準化を行うことができる。同様に、本実施形態では、負荷が高いデバイスの評価値 F を下げ、新規書き込み先としないことで、負荷が高いデバイスに対する新規書き込みを減らし、反対に、負荷の少ないデバイスの評価値 F を上げ、新規書き込み先となる割合を増やすことで、デバイスの性能平準化を行うことができる。

20

【 0 3 5 6 】

[第 2 の実施形態]

本実施形態においては、上記第 1 の実施形態の変形例について説明する。本実施形態においては、例えば、アロケート書き込み先評価処理、再配置先評価処理、オーバプロビジョニングの設定評価処理などの各種の評価処理で用いられるパラメータを変更可能とする。本実施形態では、ホスト装置 2 1 ~ 2 3 からの指示にしたがって、評価関数に含まれているストレージ装置 3 のパラメータを調整可能とする。これにより、書き込み量の抑制より性能負荷の削減を重視する設定とすること、又は、性能負荷の削減より書き込み量の抑制を重視する設定とすること、など、重視する指標を調整することで、ストレージ装置 3 をユーザのポリシーにそって最適化可能である。

30

【 0 3 5 7 】

本実施形態では、パラメータは、各種の評価処理で用いられる重みの場合を例として説明する。しかしながら、パラメータは、例えば、重みではない他の数値でもよく、例えば評価関数に含まれている定数などでもよい。

【 0 3 5 8 】

図 3 2 は、本実施形態に係るメモリ制御装置 5 を含むストレージ装置 3 の一例を示すブロック図である。

40

【 0 3 5 9 】

本実施形態において、プロセッサ 5 2 は、メモリ 5 1 に格納されている制御プログラム C P を実行し、ポリシー変更部 1 9 としての機能を実行する。メモリ 5 1 は、ポリシー情報テーブル 2 0 を格納する。

【 0 3 6 0 】

ポリシー変更部 1 9 は、ホスト装置 2 1 ~ 2 3 のいずれかからホストインタフェース 4 経由で受信したポリシー設定コマンドにしたがって、ポリシー情報テーブル 2 0 を更新する。

【 0 3 6 1 】

50

ポリシー情報テーブル 20 は、各種の評価処理で用いられる重みとその値との組み合わせを管理する。

【0362】

図 33 は、本実施形態に係るポリシー変更部 19 の処理の一例を示すフローチャートである。

【0363】

ステップ S3301 において、ポリシー変更部 19 は、ホスト装置 21 ~ 23 のいずれかからホストインタフェース 4 経由で、新規のポリシー設定として、重み W1 ~ W4 のうちの少なくとも 1 つを受信する。

【0364】

ステップ S3302 において、ポリシー変更部 19 は、新規のポリシー設定に基づいて、ポリシー情報テーブル 20 を更新する。

【0365】

ステップ S3303 において、ポリシー変更部 19 は、ポリシー変更終了を示す応答を、ホストインタフェース 4 経由でホスト装置 21 ~ 23 のいずれかへ送信する。

【0366】

図 34 は、本実施形態に係るポリシー情報テーブル 20 の一例を示すデータ構造図である。

【0367】

ポリシー情報テーブル 20 では、評価処理に用いられる係数として重み W1 ~ W4 が管理されており、重み W1 ~ W4 のそれぞれに対してユーザによって設定された値が関連付けられている。

【0368】

以上説明した本実施形態においては、例えば、アロケート書き込み先評価処理、再配置先評価処理などの重みをユーザが自由にポリシーとして設定することができる。これにより、ユーザのニーズに応じて、性能、寿命、容量、追記量に基づいて行われる評価処理を最適化することができる。

【0369】

例えば、性能を重視する場合には、性能値 V1 に関わる重み W1 の値を増やすことで、アロケート書き込みの集中を防ぎ、性能平準化を行うことができる。例えば、アロケート書き込み先評価処理において、重み W1 の値を増加すると、性能負荷が少ないデバイスは評価値 F が上がり、アロケート書き込み先となりやすくなる。したがって、性能負荷の少ないデバイスへのアロケート書き込みの割合を増加させ、性能平準化を達成することができる。

【0370】

本実施形態の効果をより具体的に説明する。

【0371】

上記第 1 の実施形態では、各種の評価処理で用いられる重み W1 ~ W4 は、評価項目の間で同じ値であり、定数としている。この場合、例えば、新規に追加されたデバイスは、残り寿命が大きいため、評価処理で計算される評価値が高まり、アロケート書き込み先として選択されやすくなる。この結果、新規に追加されたデバイスは、アロケート書き込みが集中し、性能劣化が生じる可能性がある。

【0372】

本実施形態では、例えば、デバイスが新規に追加された場合に、アロケート書き込み先評価処理の計算の性能値 V1 の重み W1 を増やす。すると、重み W1 の増加により、ストレージ装置 3 における性能が重視され、性能負荷の低いデバイスの評価値 F が上がり、アロケート書き込み先として選択されやすくなる。このように、重み W1 を増加することにより、新規に追加されたデバイスに対するアロケート書き込みの集中を防ぐことができ、性能負荷の少ないデバイスに対するアロケート書き込みの割合を増加させることができ、デバイス $D_0 \sim D_{N-1}$ の性能平準化を行うことができる。

10

20

30

40

50

【0373】

例えば、ユーザは残り寿命を重視したい場合に、寿命値V2に関する重みW2を増やすことで、寿命の少ないデバイスへのアロケート書き込みを減らすことができる。

【0374】

以上説明したように、重みW1～W4の値を変えることで、ユーザのニーズに応じて、性能、寿命、容量、追記量の影響を最適化することができる。

【0375】

[第3の実施形態]

本実施形態においては、上記第1及び第2の実施形態の変形例である。上記第1及び第2の実施形態では、各デバイス $D_0 \sim D_{N-1}$ が例えばSSD単体の場合について説明している。これに対して、本実施形態においては、複数のデバイス $D_0 \sim D_{N-1}$ のそれぞれが複数のRAID (Redundant Arrays of Inexpensive Disks) グループ (RAIDデバイス) である場合について説明する。本実施形態に係るメモリ制御装置5は、RAIDグループ上の記憶領域 (物理領域) を、エクステント (物理エクステント) 単位で管理する。ストレージ装置3は、ホスト装置21～23に対して、複数のRAIDグループの記憶領域を、複数の論理ディスク $LD_0 \sim LD_{N-1}$ として提供する。

10

【0376】

上記第1及び第2の実施形態のストレージ装置3は、RAID0に相当する。冗長化の考慮されていないRAID0に代えて、本実施形態では、冗長化を考慮したRAID1～RAID5に相当するストレージ装置3を説明する。

20

【0377】

図35は、本実施形態に係るメモリ制御装置5を含むストレージ装置3の一例を示すブロック図である。

【0378】

本実施形態において、プロセッサ52は、メモリ51に格納されている制御プログラムCPを実行し、RAID処理部24としての機能を実行する。メモリ51は、RAIDグループ情報テーブル25と、SSD固有情報テーブル26とを格納する。

【0379】

本実施形態において、デバイス $D_0 \sim D_{N-1}$ は、例えば、それぞれ、RAIDグループであり、複数のSSDを含む。

30

【0380】

デバイス $D_0 \sim D_{N-1}$ に備えられている複数のSSDに対しては、オーバープロビジョニングの設定の変更を運用後からでも可能である。

【0381】

デバイス $D_0 \sim D_{N-1}$ のRAIDレベルは、ストレージ装置3内ですべて同一とする。本実施形態は、各種のRAIDレベルに適用可能であるが、RAID1の場合を例に説明する。

【0382】

本実施形態において、デバイス $MD_0 \sim MD_{N-1}$ のそれぞれに備えられるSSDの個数は、異なってもよい。RAIDのストライプサイズは、例えば、256Kbyte又は512Kbyteでもよい。

40

【0383】

本実施形態において、デバイス $D_0 \sim D_{N-1}$ 内のSSDは、同一のオーバープロビジョニングの設定を持つとする。RAID0又はRAID5などでは、SSDへの書き込みは、デバイス $D_0 \sim D_{N-1}$ 内のSSD間でほぼ均一である。このため、デバイス $D_0 \sim D_{N-1}$ 内の各SSDの残り書き込み量も、同程度と仮定する。また、ストレージ装置3は、デバイス $D_0 \sim D_{N-1}$ に加えて、例えばホットスペア用のSSDなどのような、1以上の予備のメモリ装置27を備える。

【0384】

上記第1及び第2の実施形態においては、個々のSSDがデバイス $D_0 \sim D_{N-1}$ に相当し

50

、個々のデバイス $D_0 \sim D_{N-1}$ に対して読み出し処理、書き込み処理、寿命補正処理を実行している。これに対して、本実施形態においては、RAIDグループ単位の個々のデバイス $D_0 \sim D_{N-1}$ を仮想デバイスとして同様の処理を行う。

【0385】

RAID処理部24は、ストレージ装置3でRAIDを適用する場合に、RAIDグループ情報テーブル25及びSSD固有情報テーブル26に設定値を格納する。例えば、RAID処理部24は、RAIDグループ情報テーブル25及びSSD固有情報テーブル26の初期化を行い、また、RAIDグループ情報テーブル25及びSSD固有情報テーブル26への情報の登録及び更新を行う。

【0386】

RAIDグループ情報テーブル25は、各デバイス $D_0 \sim D_{N-1}$ に所属するSSDを管理する。

【0387】

SSD固有情報テーブル26は、各デバイス $D_0 \sim D_{N-1}$ に所属するSSDの個々の情報を管理する。

【0388】

図36は、本実施形態に係るRAIDグループ情報テーブル25の一例を示すデータ構造図である。

【0389】

RAIDグループ情報テーブル25は、デバイスID、RAIDレベル、所属SSD IDを互いに関連付けて格納する。RAIDグループ情報テーブル25には、ストレージ装置3内の全てのデバイス $D_0 \sim D_{N-1}$ に対する情報がエントリされる。

【0390】

デバイスIDは、上記のデバイス固有情報テーブル6におけるデバイスIDと同一である。

【0391】

RAIDレベルは、ストレージ装置3におけるRAIDレベルを表す。このRAIDレベルは、ストレージ装置3で既に実行しているRAID設定を参照して設定されてもよい。

【0392】

所属SSD IDは、デバイス $D_0 \sim D_{N-1}$ 内に含まれるSSDを示すIDである。

【0393】

図37は、本実施形態に係るSSD固有情報テーブル26の一例を示すデータ構造図である。

【0394】

SSD固有情報テーブル26には、ストレージ装置3内の全てのSSDの情報がエントリされている。

【0395】

この図37では、SSDデバイスIDに対して、最大容量、論理容量、現在使用容量、残り書き込み量、保証書き込み量が関連付けられている。

【0396】

図38は、本実施形態に係るRAID処理部24の処理の一例を示すフローチャートである。

【0397】

RAID処理部24は、ストレージ装置3へ追加されたRAIDグループに関する各テーブルの初期化を行う。この図38において、各SSDとRAIDグループとの関連付けは終わっているとする。

【0398】

ステップS3801において、RAID処理部24は、ストレージ装置3内の全てのSSDの設定が完了したか否か判断する。

10

20

30

40

50

【 0 3 9 9 】

ストレージ装置 3 内の全ての S S D の設定が完了した場合、処理はステップ S 3 8 0 4 に移動する。

【 0 4 0 0 】

ストレージ装置 3 内の全ての S S D の設定が完了していない場合、ステップ S 3 8 0 2 において、R A I D 処理部 2 4 は、未設定の S S D を選択する

ステップ S 3 8 0 3 において、R A I D 処理部 2 4 は、S S D 固有情報テーブル 2 6 に対して、選択された S S D に対応するレコードを生成する。その後、処理はステップ S 3 8 0 1 に移動する。

【 0 4 0 1 】

ステップ S 3 8 0 4 において、R A I D 処理部 2 4 は、ストレージ装置 3 内の全ての R A I D グループ、すなわちデバイス $D_0 \sim D_{N-1}$ の設定が完了したか否か判断する。

【 0 4 0 2 】

ストレージ装置 3 内の全ての R A I D グループの設定が完了した場合、処理はステップ S 3 8 0 7 に移動する。

【 0 4 0 3 】

ストレージ装置 3 内の全ての R A I D グループの設定が完了していない場合、ステップ S 3 8 0 5 において、R A I D 処理部 2 4 は、未設定の R A I D グループを選択する。

【 0 4 0 4 】

ステップ S 3 8 0 6 において、R A I D 処理部 2 4 は、R A I D グループ情報テーブル 2 5 に対して、選択された R A I D グループに対応するレコードを生成する。

【 0 4 0 5 】

R A I D グループ情報テーブル 2 5 のレコード生成において、各種の I D は、R A I D グループ情報テーブル 2 5 内で一意な情報とする。

【 0 4 0 6 】

R A I D レベルとしては、ストレージ装置 3 の R A I D 構成情報を参照して得られた設定が設定される。

【 0 4 0 7 】

所属 S S D デバイス I D には、R A I D 構成情報に含まれる S S D デバイス I D が設定される。

【 0 4 0 8 】

選択された R A I D グループに対応するレコードの生成後、処理はステップ S 3 8 0 4 に移動する。

【 0 4 0 9 】

ステップ S 3 8 0 7 において、R A I D 処理部 2 4 は、ストレージ装置 3 内の全ての R A I D グループに関するレコード追加処理が完了したか否か判断する。

【 0 4 1 0 】

ストレージ装置 3 内の全ての R A I D グループに関するレコード追加処理が完了した場合、処理は終了する。

【 0 4 1 1 】

ストレージ装置 3 内の全ての R A I D グループに関するレコード追加処理が完了していない場合、ステップ S 3 8 0 8 において、R A I D 処理部 2 4 は、レコード追加処理を未実行の R A I D グループを選択する。

【 0 4 1 2 】

ステップ S 3 8 0 9 において、R A I D 処理部 2 4 は、デバイス固有情報テーブル 6 に対して、選択された R A I D グループに対応するレコード追加処理を実行する。

【 0 4 1 3 】

例えば、R A I D 処理部 2 4 は、デバイス固有情報テーブル 6 におけるデバイス I D として、R A I D グループ情報テーブル 2 5 のデバイス I D を設定する。

【 0 4 1 4 】

10

20

30

40

50

例えば、RAID処理部24は、デバイス固有情報テーブル6における最大容量及び論理容量として、RAIDレベルによって異なる値を設定する。例えば、RAIDレベルがRAID1の場合、RAID処理部24は、RAIDグループに所属するSSDのうちのいずれかの値を設定する。例えば、RAIDレベルがRAID0の場合、RAID処理部24は、最大容量に対して、各SSDの最大容量の合計を設定する。論理容量も同様に、RAIDレベルに応じて、異なる値が設定される。

【0415】

例えば、RAID処理部24は、デバイス固有情報テーブル6における保証書き込み量、性能仕様として、RAIDグループに所属するSSDのいずれかの値を設定する。

【0416】

例えば、RAID処理部24は、デバイス固有情報テーブル6における現在使用容量、読み出し量、更新量、追記量、最大負荷をゼロで初期化する。

【0417】

レコード追加処理の実行後、処理はステップS3807に移動する。

【0418】

以上説明した本実施形態では、デバイス $D_0 \sim D_{N-1}$ がRAIDグループとして機能する場合に、デバイス $D_0 \sim D_{N-1}$ のそれぞれを、1つの仮想的なデバイスとして扱うことができ、上記第1及び第2の実施形態と同様の効果を得ることができる。すなわち、本実施形態においては、ストレージ装置3が備える複数台のディスクアレイの間で、性能平準化及び寿命平準化を実現することができる。

【0419】

[第4の実施形態]

本実施形態においては、上記第1乃至第3の実施形態に係るストレージ装置3と、比較例のストレージ装置とを対比し、上記第1乃至第3の実施形態に係るストレージ装置3の有効性を説明する。

【0420】

比較例のストレージ装置は、SSD内のフラッシュメモリの寿命劣化を防ぐため、書き込み寿命の異なる複数台のSSDを備え、データの書き込み先となるSSDを選択することにより、フラッシュメモリの寿命劣化を防ぐ。書き込み寿命とは、ベンダの保証するSSDの書き込み総量から現在までに書き込まれた量を引いた残りである。比較例のストレージ装置では、1ビットあたりの単価が高いが、書き込み寿命の長いSLC (Single Level Cell) を備えるSSDと、1ビットあたりの単価は低いが、書き込み寿命の短いMLC (Multi Level Cell) を備えるSSDとの間で、書き込み量の多いデータをSLCのSSDへ書き込み、書き込み量の少ないデータをMLCのSSDへ書き込む。この比較例のストレージ装置では、比較例のストレージ装置のワークロード、すなわち、比較例のストレージ装置に対する書き込み量とユーザの使用容量を予め想定し、ワークロードに対して適切な寿命と容量になるようにSLCのSSD及びMLCのSSDの台数を決定する。

【0421】

ワークロードに対して比較例のストレージ装置の寿命と容量とを適切にするために、以下のような検討がなされる。

【0422】

SLCのSSDの容量を S_s とする。MLCのSSDの容量を S_m とする。SLCのSSDの書き込み寿命を W_s とする。MLCのSSDの書き込み寿命を W_m とする。SLCのSSDの個数を P_s とする。MLCのSSDの個数を P_m とする。

【0423】

比較例のストレージ装置で想定される使用可能な最大容量を S_t 、想定される書き込み総量を W_t とする。

【0424】

比較例のストレージ装置は、以下の(1)式、(2)式を、使用期間中、常に満たすことが望まれる。

【 0 4 2 5 】

$$S_m \times P_m + S_s \times P_s > S_t \quad \dots \quad (1)$$

$$W_m \times P_m + W_s \times P_s > W_t \quad \dots \quad (2)$$

(1) 式が成り立たない場合は容量が不足していることを示す。(2) が成り立たない場合は書き込み寿命が不足していることを示す。

【 0 4 2 6 】

比較例のストレージ装置は、上記(1)式及び(2)式を満たしつつ、当該比較例のストレージ装置内のSSDのコスト合計が最小に成るように、 S_t と W_t とを事前に想定し、SLCのSSDの個数 P_s とMLCのSSDの個数 P_m とを決定する。

【 0 4 2 7 】

しかしながら、比較例のストレージ装置では、当該比較例のストレージ装置に対するワークロードを運用する前から正確に把握することは困難である。また、運用後に、ワークロードは刻々と変化する。よって、想定しているワークロードに最適と思われる構成をしても、比較例のストレージ装置の運用後に、現在の書き込み量が想定した書き込み量から変化し、寿命劣化又は寿命過剰となる場合がある。より具体的には、実際のワークロードに対して、比較例のストレージ装置内にSSDを必要以上に備えた場合、(2)式の W_t が想定より非常に小さく、書き込み寿命に達する前の運用終了自体は問題ないが、過剰にSSDを用意したことになり、コストが高くなる場合がある。また、実際のワークロードに対して、比較例のストレージ装置内のSSDの数が不足した場合、(2)式の W_t が想定より大きくなり、ベンダの保証する期間より前に書き込み寿命が尽きてしまい、運用途中でのSSDの交換が必要となる場合がある。

【 0 4 2 8 】

また、比較例のストレージ装置においては、容量について、(1)式で想定した場合よりも S_t が大きくなり、容量不足になる場合がある。また、比較例のストレージ装置においては、容量について、(1)式で想定した場合よりも S_t が小さくなり、容量過剰となる場合がある。

【 0 4 2 9 】

これに対して、上記第1乃至第3の実施形態に係るストレージ装置3は、オーバプロビジョニングの設定を運用後に変更可能なデバイス $D_0 \sim D_{N-1}$ を用いて、ワークロードに応じて、寿命、容量を最適化する。オーバプロビジョニングの効果として、書き込み不能な領域を変化させることで、デバイス $D_0 \sim D_{N-1}$ の書き込み寿命を変化(増減)させることが可能である。上記第1乃至第3の実施形態では、オーバプロビジョニングの設定を運用後でも変更することが可能なデバイス $D_0 \sim D_{N-1}$ を用い、オーバプロビジョニングで論理容量を適切に設定する。このため、実際のワークロードの変化に応じて、デバイス $D_0 \sim D_{N-1}$ の書き込み量を適切化することができる。

【 0 4 3 0 】

上記第1乃至第3の実施形態に係るストレージ装置3では、ストレージ装置3の運用中のワークロードの変化にそってオーバプロビジョニングの設定及びデータの配置が実行されるため、ワークロードに対してデバイス $D_0 \sim D_{N-1}$ の寿命、容量を適切に設定することができる。

【 0 4 3 1 】

上記第1乃至第3の実施形態に係るストレージ装置3は、運用中に、デバイス $D_0 \sim D_{N-1}$ のうちのいずれかの故障、寿命、使用する容量の増加等の理由により、新しくデバイスを増設、交換してもよい。この場合であっても、上記第1乃至第3の実施形態に係るストレージ装置3では、ストレージ装置3内の各デバイスの書き込み寿命は、デバイス間で不均衡にならない。上記第1乃至第3の実施形態に係るストレージ装置3では、書き込み寿命に達しそうなデバイスに対する書き込みを削減し、書き込み寿命が十分あるデバイスに優先的に書き込みを行うことで、書き込み寿命に達しそうなデバイスの寿命低下を抑えることができる。その結果、デバイスの書き込み寿命が尽きることを防止することができる。

【 0 4 3 2 】

上記第 1 乃至第 3 の実施形態によれば、論理ディスクを固定サイズの領域に分割し、所定サイズの領域単位で書き込み量を計測する。ストレージ装置 3 は、書き込み量及びデバイス $D_0 \sim D_{N-1}$ の性能値 $V1$ 、寿命値 $V2$ 、容量値 $V3$ 、追記値 $V4$ に基づいて評価値 F を計算し、まだデータの書き込まれていない所定サイズの領域に対するデータの書き込み時に、評価値 F に基づいて書き込み先デバイスを選択する。また、ストレージ装置 3 は、評価値 F に基づいてデバイス $D_0 \sim D_{N-1}$ に対する各所定サイズの領域の再配置と、デバイス $D_0 \sim D_{N-1}$ のオーバープロビジョニングの設定を行う。これにより、上記第 1 乃至第 3 の実施形態においては、ワークロードが変化した場合でも、ワークロードに応じて、デバイス $D_0 \sim D_{N-1}$ の寿命、負荷、性能、容量を平準化することができる。

10

【 0 4 3 3 】

本発明のいくつかの実施形態を説明したが、これらの実施形態は、例として提示したものであり、発明の範囲を限定することは意図していない。これら新規な実施形態は、その他の様々な形態で実施されることが可能であり、発明の要旨を逸脱しない範囲で、種々の省略、置き換え、変更を行うことができる。これら実施形態やその変形は、発明の範囲や要旨に含まれるとともに、特許請求の範囲に記載された発明とその均等の範囲に含まれる

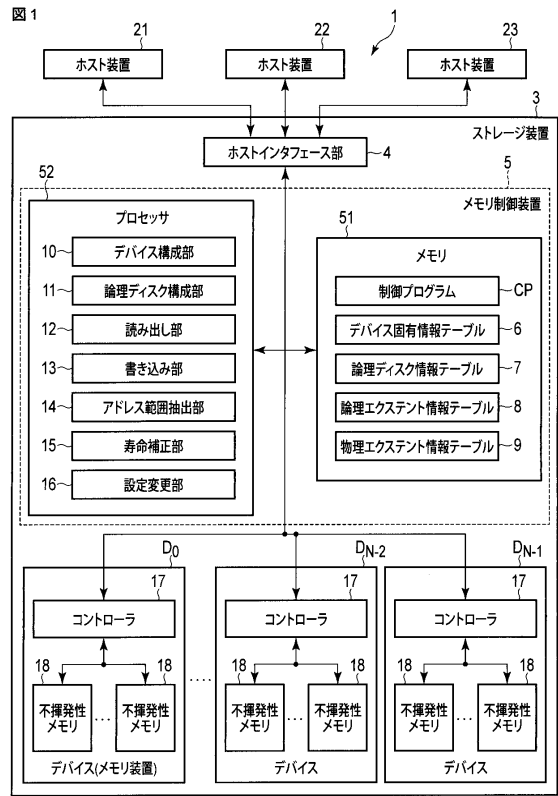
【 符号の説明 】

【 0 4 3 4 】

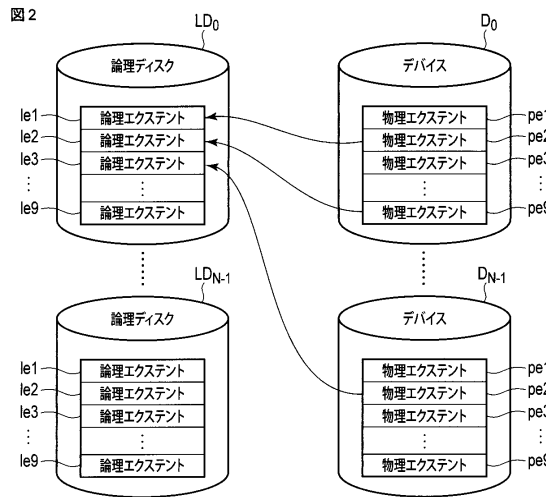
1 ... ストレージシステム、2 1 ~ 2 3 ... ホスト装置、4 ... ホストインタフェース部、5 ... メモリ制御装置、5 1 ... メモリ、5 2 ... プロセッサ、C P ... 制御プログラム、6 ... デバイス固定情報テーブル、7 ... 論理ディスク情報テーブル、8 ... 論理エクステント情報テーブル、9 ... 物理エクステント情報テーブル、1 0 ... デバイス構成部、1 1 ... 論理ディスク構成部、1 2 ... 読み出し部、1 3 ... 書き込み部、1 4 ... アドレス範囲抽出部、1 5 ... 寿命補正部、1 6 ... 設定変更部、1 7 ... コントローラ、1 8 ... 不揮発性メモリ、 $D_0 \sim D_{N-1}$... デバイス、 $L D_0 \sim L D_{N-1}$... 論理ディスク、1 e 9 ... 論理エクステント、p e 1 ... 物理エクステント、1 9 ... ポリシー変更部、2 0 ... ポリシー情報テーブル、2 4 ... R A I D 処理部、2 5 ... R A I D グループ情報テーブル、2 6 ... S S D 固有情報テーブル、2 7 ... 予備のメモリ装置。

20

【図 1】



【図 2】



【図 3】

図 3

6 デバイス固有情報テーブル

デバイス ID	最大容量 (MB)	論理容量 (MB)	現在使用量 (MB)	残り容量 (MB)	保証書き込み量 (MB/日)	読み出し量 (MB)	書き出し量 (MB)	更新量 (MB)	追記量 (MB)	性能仕様 (IOPS)	最大負荷 (IOPS)
D ₀	102400.0	20000.0	20000.0	400000.0	20480	1233	1233	123	12	40000	3200
D ₁	102400.0	40000.0	30000.0	500000.0	20480	1032	1032	100	11	40000	1000
...
D _{N-1}	102400.0	80000.0	40000.0	450000.0	20480	330	330	0	0	11	11

【図 4】

図 4

7 論理ディスク情報テーブル

論理ディスクID	論理容量
LD ₀	10240
LD ₁	5120
...	...
LD _{N-1}	10240

【図 5】

図 5

8 論理エクステント情報テーブル

論理エクステント ID	論理ディスク ID	先頭論理アドレス	所属デバイス ID	物理エクステント ID	読み出し量 (MB)	書き込み量 (MB)
0	LD ₀	0000000000	D1	61	19	0.5
1	LD ₀	0000016384	D1	62	4	1.3
...	D1
10240	LD ₀	00167772159	D1	661	0.0	0.0
10241	LD ₁	00000000000	D1	771	0.0	0.0
...
102399	LD _{N-1}	00167772159	D10	991	0.0	0.0

【図 6】

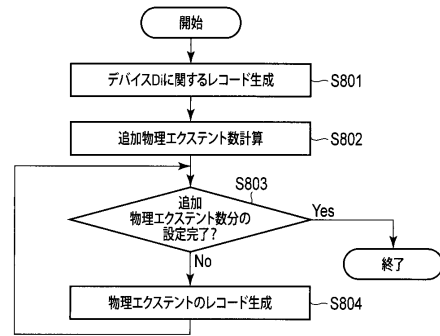
図 6

9 物理エクステント情報テーブル

物理 エクステント ID	所属 デバイス ID	先頭物理アドレス	空き情報
0	D ₁	00000000000	0
1	D ₁	00000016384	1
...
10240	D ₁	00167772159	NULL
10241	D ₂	00000000000	1
...
102399	D ₁₀	00167772159	NULL

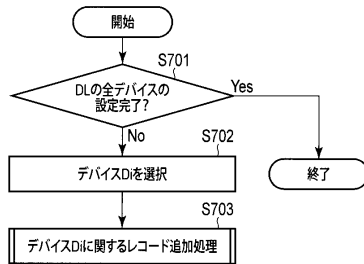
【図 8】

図 8



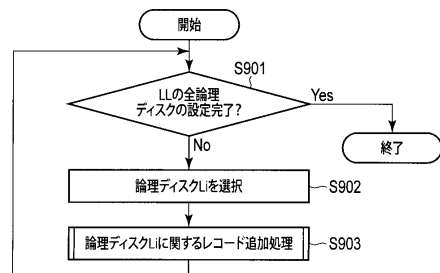
【図 7】

図 7



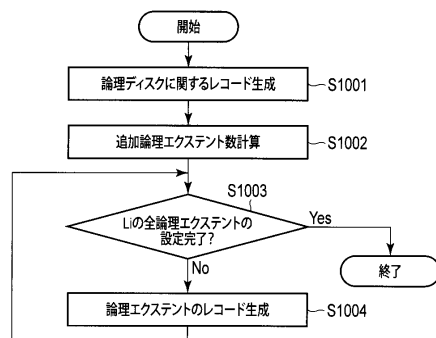
【図 9】

図 9



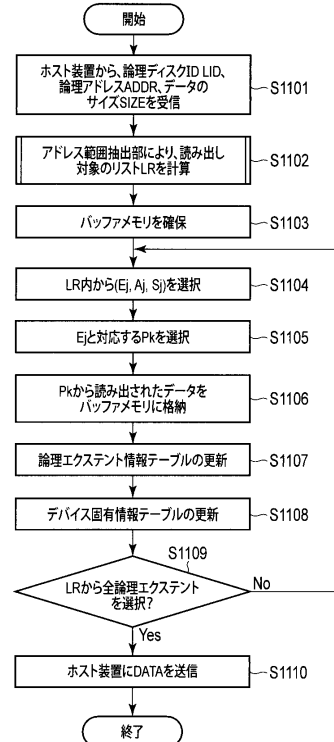
【図 10】

図 10



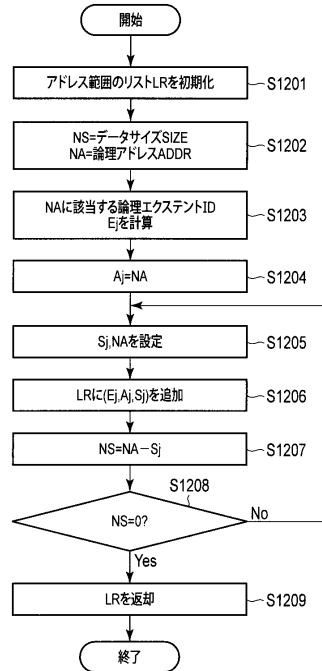
【図 11】

図 11



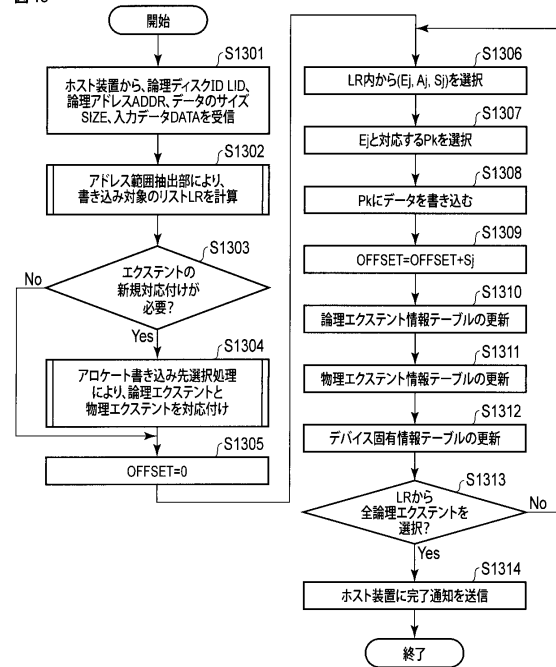
【図 12】

図 12



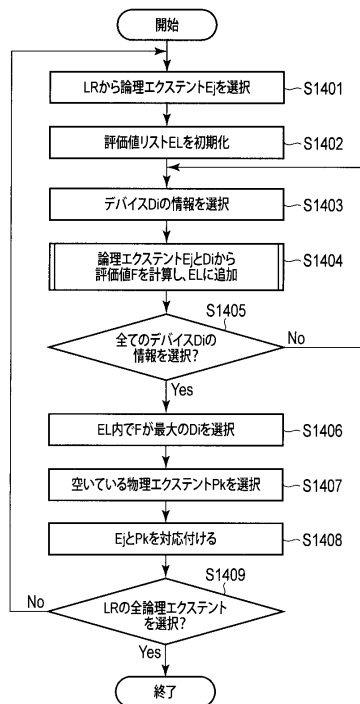
【図 13】

図 13



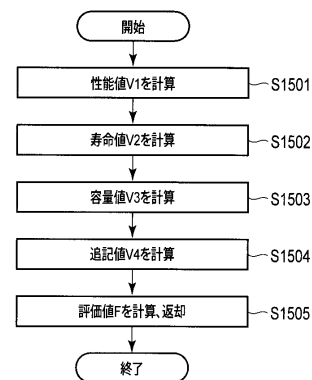
【図 14】

図 14



【図 15】

図 15



【図 16】

図 16

8 デバイス固有情報テーブル

デバイス ID	最大容量 (MB)	論理容量 (MB)	現在使用容量 (MB)	残り書き込み量 (MB)	保証書き込み量 (MB/日)	読み出し量 (MB)	更新量 (MB)	追加量 (MB)	性能仕様 (IOPS)	最大負荷 (IOPS)
D ₁	10	7	5	4000	10	12	12	1	40000	3200
D ₂	10	5	4	5000	10	10	10	1	40000	1000
D ₃	10	5	1	4500	10	1	0	0	30000	0

【図 17】

図 17

8 論理エクステント情報テーブル

論理エクステント ID	論理ディスク ID	先頭論理アドレス	所属デバイス ID	物理エクステント ID	読み出し量 (MB)	書き込み量 (MB)
E1	0	0	D ₁	0	19	2
E2	0	2048	D ₁	P1	2	8
E3	0	4096	D ₁	P2	0	2
E4	0	6144	D ₂	0	0	1
E5	0	8192	NULL	NULL	0	2

【図 18】

図 18

9 物理エクステント情報テーブル

物理エクステント ID	所属デバイス ID	先頭物理アドレス	空き情報
P1	D ₁	0	0
P2	D ₁	2048	0
P3	D ₂	0	0
P4	D ₂	2048	0
P5	D ₃	0	0
P6	D ₃	2048	0

【図 19】

図 19

デバイス ID	V1	EW	AW	V2	V3	AR	V4	F
D ₁	0.23	8.888889	12	-0.0875	0.035714	0.04	0.24	0.418214
D ₂	0.24375	11.111111	1	0.2275	0	0.047619	0.238095	0.709345
D ₃	0.25	10	0	0.25	0.15	0	0.25	0.9

【図 20】

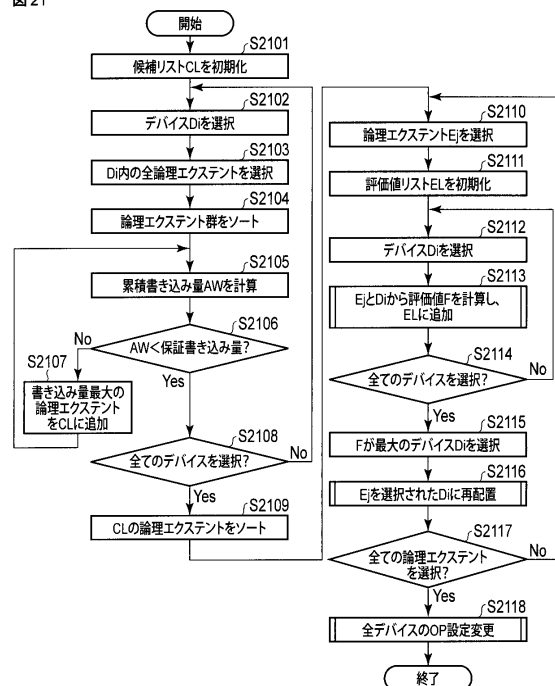
図 20

8 論理エクステント情報テーブル

論理エクステント ID	論理ディスク ID	先頭論理アドレス	所属デバイス ID	物理エクステント ID	読み出し量 (MB)	書き込み量 (MB)
E1	0	0	D ₁	0	19	2
E2	0	2048	D ₁	P1	2	8
E3	0	4096	D ₁	P2	0	2
E4	0	6144	D ₂	0	0	1
E5	0	8192	D ₃	P6	0	2

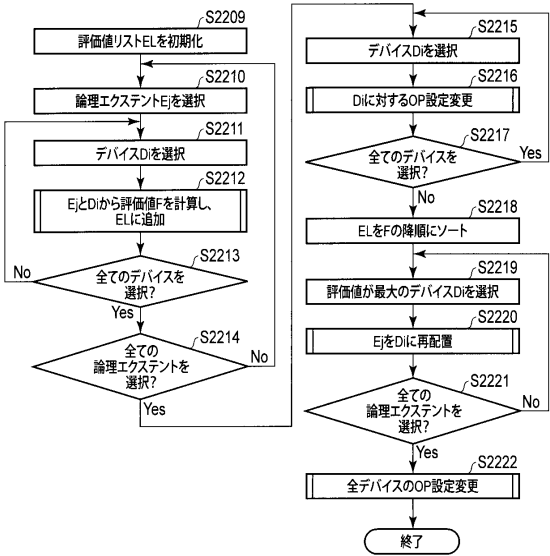
【図 21】

図 21



【図 2 2】

図 22



【図 2 3】

図 23

デバイスID	保証書き込み量 (MB/日)	AW
D ₁	10	12
D ₂	10	1
D ₃	10	0

【図 2 4】

図 24

デバイスID	V1	EW	AW	V2	V3	AR	V4	F
D ₁	0.23	8.888889	12	-0.0875	0.035714	0.04	0.24	0.418214
D ₂	0.24375	11.11111	1	0.2275	0	0.047619	0.238095	0.709345
D ₃	0.25	10	0	0.25	0.15	0	0.25	0.9

【図 2 5】

図 25

8 論理エクステント情報テーブル

論理 エクステント ID	論理 ディスク ID	先頭論理 アドレス	所属 デバイス ID	物理 エクステント ID	読み出し量 (MB)	書き込み量 (MB)
E1	0	0	D ₁	0	19	2
E2	0	2048	D ₃	P5	2	8
E3	0	4096	D ₁	P2	0	2
E4	0	6144	D ₂	0	0	1
E5	0	8192	NULL	NULL	0	2

【図 2 6】

図 26

6 デバイス固有情報テーブル

デバイスID	最大容量 (MB)	論理容量 (MB)	現在 使用容量 (MB)	残り 書き込み量 (MB)	保証 書き込み量 (MB/日)	読み出し量 (MB)	更新量 (MB)	追記量 (MB)	性能仕様 (IOPS)	最大負荷 (IOPS)
D ₁	10	10	4	4000	10	12	12	1	40000	3200
D ₂	10	10	4	5000	10	10	10	1	40000	1000
D ₃	10	10	2	4499	10	1	1	0	30000	0

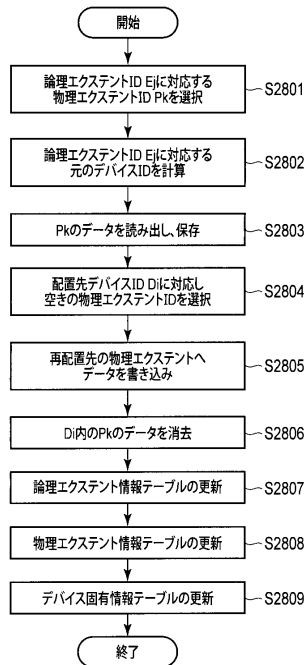
【図 2 7】

図 27

デバイスID	V1	EW	AW	V2	V3	AR	V4	F	C	V	計算前 論理容量 (MB)	計算後 論理容量 (MB)	差分
D ₁	0.23	8.888889	12	-0.0875	0.125	0.04	0.01	0.2775	5	2.3875	7	7.3875	0.3875
D ₂	0.24375	11.11111	1	0.2275	0.15	0.047619	0.011905	0.633155	6	4.798929	5	8.798928571	3.798929
D ₃	0.25	10	0	0.25	0.225	0	0	0.725	9	6.525	5	7.525	2.525

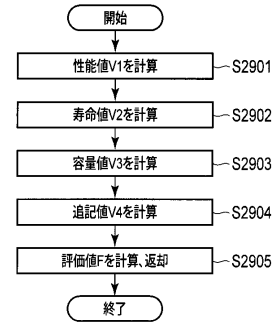
【図 28】

図 28



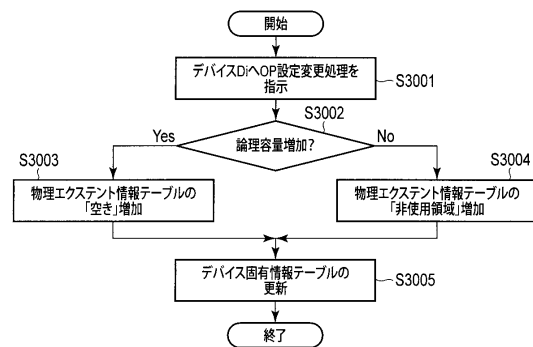
【図 29】

図 29



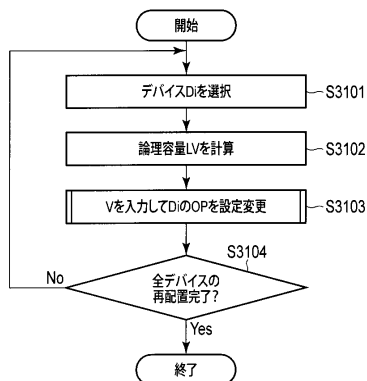
【図 30】

図 30



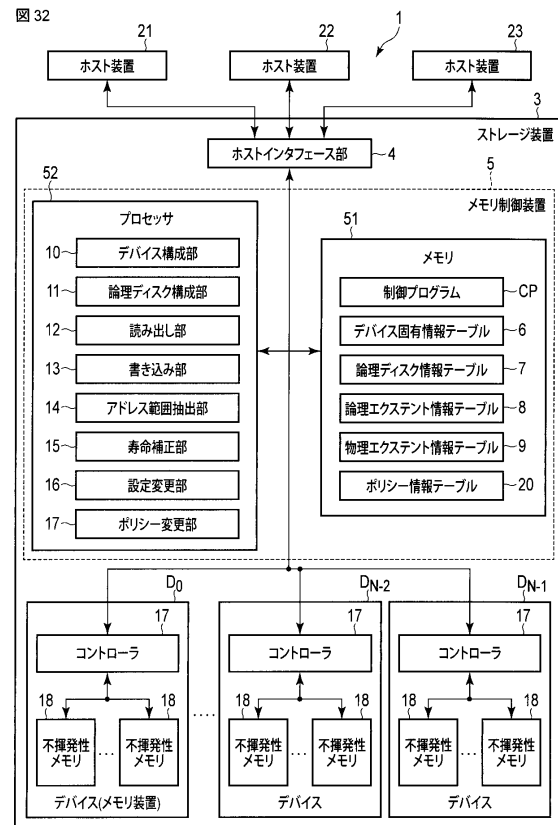
【図 31】

図 31



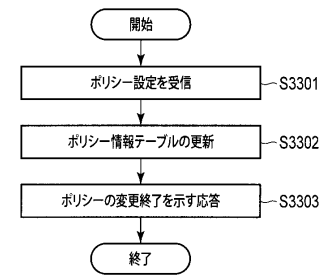
【図 32】

図 32



【図 33】

図 33



【図 34】

図 34

20 ポリシー情報テーブル

係数	値
W1	0.1
W2	0.4
W3	0.3
W4	0.2

【図 36】

図 36

25 RAIDグループ情報テーブル

デバイスID	RAIDレベル	所属SSD ID
D ₀	1	1,2,3,4
D ₁	1	5,6,7,8
...
D _{N-1}	1	37,38,39,40

【図 37】

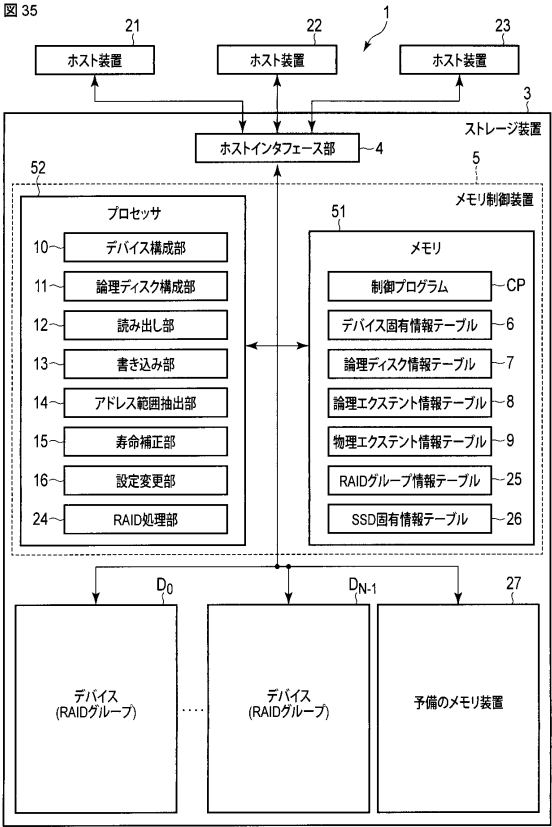
図 37

26 SSD固有情報テーブル

SSDデバイス ID	最大容量 (MB)	論理容量 (MB)	現在使用容量 (MB)	残り書き込み量 (MB)	保証書き込み量 (MB/日)
0	102400.0	20000.0	20000.0	102400	4000000.0
1	102400.0	40000.0	30000.0	51200	5000000.0
...
10	102400.0	80000.0	40000.0	4500000.0	20480

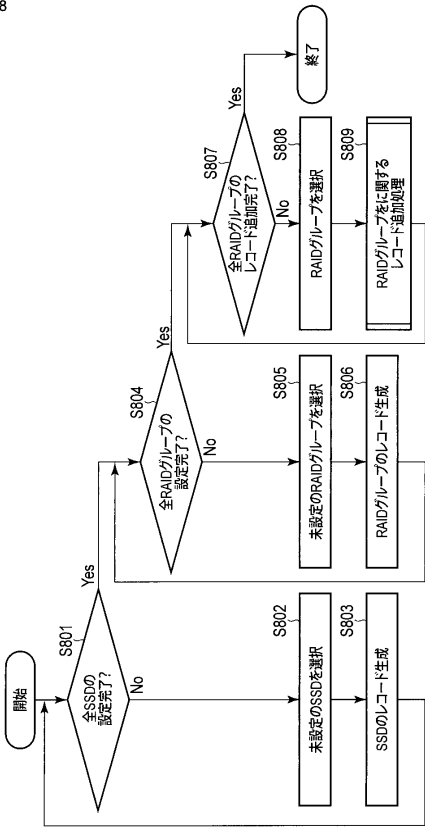
【図 35】

図 35



【図 38】

図 38



フロントページの続き

(56)参考文献 国際公開第2016/013202(WO, A1)

特表2015-505078(JP, A)

特表2014-513356(JP, A)

特表2011-505046(JP, A)

特開2008-102900(JP, A)

(58)調査した分野(Int.Cl., DB名)

G06F 12/00

G06F 12/02

G06F 12/16

G06F 3/06