

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号
特許第6142685号
(P6142685)

(45) 発行日 平成29年6月7日(2017.6.7)

(24) 登録日 平成29年5月19日(2017.5.19)

(51) Int.Cl.

G 0 6 F 3 / 0 6 (2 0 0 6 . 0 1)

F 1

G 0 6 F 3 / 0 6 3 0 1 Z

請求項の数 7 (全 26 頁)

(21) 出願番号	特願2013-124158 (P2013-124158)	(73) 特許権者	000005223
(22) 出願日	平成25年6月12日 (2013.6.12)		富士通株式会社
(65) 公開番号	特開2014-241117 (P2014-241117A)		神奈川県川崎市中原区上小田中4丁目1番1号
(43) 公開日	平成26年12月25日 (2014.12.25)	(74) 代理人	100089118
審査請求日	平成28年3月10日 (2016.3.10)		弁理士 酒井 宏明
		(72) 発明者	三浦 寛滋
			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
		(72) 発明者	酒井 基裕
			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
		(72) 発明者	小林 明人
			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
			最終頁に続く

(54) 【発明の名称】 ストレージシステム、運用管理方法及び運用管理プログラム

(57) 【特許請求の範囲】

【請求項 1】

性能の異なる複数のストレージのうち性能の高いストレージを優先させてアクセス頻度の高いデータから順にストレージを割り当てる割当部と、

前記割当部によりストレージが割り当てられたデータのうち、現在記憶されているストレージと異なるストレージが割り当てられたデータを再配置の対象として特定する特定部と、

前記特定部により再配置の対象として特定されたデータの再配置を行う複数の制御部とを有し、

性能が同じストレージは複数のグループにグループ化され、

所定の大きさの記憶単位のデータ及びグループには前記複数の制御部のうち1つの制御部が割り当てられ、

各制御部は、データの再配置を行う際に、該データを含む前記所定の大きさの記憶単位のデータに割り当てられた制御部と同じ制御部が割り当てられたグループに属するストレージをデータの移動先のストレージとすることを特徴とするストレージシステム。

【請求項 2】

前記割当部は、割合に関する設定がされている場合に、前記所定の大きさの記憶単位のデータが前記複数のストレージに記憶される割合に基づいて該所定の大きさの記憶単位のデータについてアクセス頻度の高いデータから順に性能の高いストレージに割り当てることを特徴とする請求項 1 に記載のストレージシステム。

【請求項 3】

前記割当部は、性能が同じであるストレージだけを使用する設定がされている場合に、性能が同じであるストレージだけにデータを割り当てることを特徴とする請求項 1 に記載のストレージシステム。

【請求項 4】

記憶領域の自動確保の設定がされている場合に、ストレージを選択し、選択したストレージを用いて前記記憶領域の作成を前記制御部に指示する作成指示部とをさらに有することを特徴とする請求項 1 に記載のストレージシステム。

【請求項 5】

前記特定部は、再配置の対象として特定したデータのうち、隣接する性能のストレージに記憶されるデータのうちアクセス頻度が最も近いデータからアクセス頻度が近い順に所定の割合のデータを再配置の対象から除外することを特徴とする請求項 1 に記載のストレージシステム。

【請求項 6】

性能の異なる複数のストレージのうち性能の高いストレージを優先させてアクセス頻度の高いデータから順にストレージを割り当て、

ストレージが割り当てられたデータのうち、現在記憶されているストレージと異なるストレージが割り当てられたデータを再配置の対象として特定し、

再配置の対象として特定したデータの再配置をストレージを制御する制御装置に指示する処理をコンピュータに実行させ、

前記制御装置は複数あり、

性能が同じストレージは複数のグループにグループ化され、

所定の大きさの記憶単位のデータ及びグループには前記複数の制御装置のうち 1 つの制御装置が割り当てられ、

各制御装置は、データの再配置を行う際に、該データを含む前記所定の大きさの記憶単位のデータに割り当てられた制御装置と同じ制御装置が割り当てられたグループに属するストレージをデータの移動先のストレージとすることを特徴とする運用管理プログラム。

【請求項 7】

性能の異なる複数のストレージのうち性能の高いストレージを優先させてアクセス頻度の高いデータから順にストレージを割り当て、

ストレージが割り当てられたデータのうち、現在記憶されているストレージと異なるストレージが割り当てられたデータを再配置の対象として特定し、

再配置の対象として特定したデータの再配置をストレージを制御する制御装置に指示する処理をコンピュータが実行し、

前記制御装置は複数あり、

性能が同じストレージは複数のグループにグループ化され、

所定の大きさの記憶単位のデータ及びグループには前記複数の制御装置のうち 1 つの制御装置が割り当てられ、

各制御装置は、データの再配置を行う際に、該データを含む前記所定の大きさの記憶単位のデータに割り当てられた制御装置と同じ制御装置が割り当てられたグループに属するストレージをデータの移動先のストレージとすることを特徴とする運用管理方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、ストレージシステム、運用管理方法及び運用管理プログラムに関する。

【背景技術】

【0002】

従来、ストレージシステムでは、複数の性能の異なるストレージを用いて、ストレージ自動階層制御 (AST: Automated Storage Tiering) が行われてきた。ここで、性能の異なるストレージとしては、HDD (Hard Disk Drive)、SSD (Solid State D

10

20

30

40

50

rive) などがある。また、HDDの中にも、アクセス性能の異なる装置がある。ASTは、情報処理システムの運用中に、性能の異なるストレージ間でデータの自動再配置を行う。このため、ストレージシステムは、ASTにより情報処理システムの運用中に動的にデータ配置を変更することで、システム運用中の性能状況変化などに追従することができる。

【0003】

ASTは、異なる性能特性のストレージを階層化したプールに存在するボリュームのアクセスを監視し、ストレージシステムの管理者が設定した階層化ポリシーに従ってサブプール間でデータの再配置を行う。ここで、プールとは、複数のHDDやSSDをまとめた1つの仮想的なディスクである。また、サブプールとは、性能に基づいてプールを分割したものである。複数のサブプールを複数の層(ティア:Tier)として有するプールはティアプール(又は、階層化プール)と呼ばれる。また、階層化ポリシーは、データの再配置のタイミング、データのアクセス頻度とデータを配置するサブプールとの関係など再配置に関する方針を定義する情報である。

10

【0004】

図25は、階層制御を説明するための図である。図25は、ストレージ装置92と運用管理装置93がLAN94で接続されたストレージシステムを示す。ストレージ装置92は、SSD、オンラインディスク、ニアラインディスクの3つのサブプールを有する。3つのサブプールの中でSSDが最も性能が高く、ニアラインディスクが最も性能が低い。

【0005】

20

運用管理装置93は、ストレージ装置92の運用を管理し、ASTの機能を有する。ASTは、アクセス頻度の高いデータを高性能なSSDに配置し、アクセス頻度の低いデータを低性能で安価なニアラインディスクに配置することで、コストを抑えつつ性能を向上させることができる。また、ASTは、システムの運用中に動的にデータ配置を変更するので、配置設計を不要とし、ストレージシステムの管理者の作業負荷の軽減、管理コストの低減を実現する。

【0006】

また、データアクセス頻度が予め指定した上限を上回るデータブロックを高性能のグループの記憶装置に移動し、予め指定した下限を下回るデータブロックを低性能のグループの記憶装置に移動するストレージシステムがある(例えば、特許文献1参照。)。また、2種以上の記憶媒体を有する記憶装置において、各データの使用頻度を、基準となる使用頻度と比較し、データの使用頻度の方が低い場合には、アクセス時間のより大きい記憶媒体にデータを移動する従来技術がある(例えば、特許文献2参照。)。

30

【先行技術文献】

【特許文献】

【0007】

【特許文献1】特開2003-108317号公報

【特許文献2】特開平3-48321号公報

【発明の概要】

【発明が解決しようとする課題】

40

【0008】

ストレージシステムの管理者は、階層化ポリシーでIOPS(Input/Output Per Second)値を指定することでデータの再配置の基準を設定する。例えば、ストレージシステムの管理者は、IOPSが501以上である場合にはデータをSSDに再配置する、というようにデータの再配置の基準を階層化ポリシーで指定する。

【0009】

ASTは、階層化ポリシーで指定されたIOPS値に基づいてデータの再配置を行う。このため、高性能なサブプールに空き領域が存在する場合でも高性能なサブプールが利用されないという問題がある。例えば、IOPSが501以上である場合にはデータをSSDに再配置するという階層化ポリシーでは、IOPSが501以上のデータがない場合に

50

は、SSDは利用されない。また、IOPSが501以上のデータが少ない場合には、SSDに空き領域が存在しても空き領域が利用されない。

【0010】

本発明は、1つの側面では、性能の異なる複数のストレージのうち高性能なストレージを有効に利用するストレージシステムを提供することを目的とする。

【課題を解決するための手段】

【0011】

本願の開示するストレージシステムは、1つの態様において、割当部と、特定部と、複数の制御部とを有する。割当部は、性能の異なる複数のストレージのうち性能の高いストレージを優先させてアクセス頻度の高いデータから順にストレージを割り当てる。特定部は、割当部によりストレージが割り当てられたデータのうち、現在記憶されているストレージと異なるストレージが割り当てられたデータを再配置の対象として特定する。制御部は、特定部により再配置の対象として特定されたデータの再配置を行う。また、性能が同じストレージは複数のグループにグループ化され、所定の大きさの記憶単位のデータ及びグループには複数の制御部のうち1つの制御部が割り当てられる。そして、各制御部は、データの再配置を行う際に、該データを含む所定の大きさの記憶単位のデータに割り当てられた制御部と同じ制御部が割り当てられたグループに属するストレージをデータの移動先のストレージとする。

10

【発明の効果】

【0012】

1実施態様によれば、性能の異なる複数のストレージのうち高性能なストレージを有効に利用することができる。

20

【図面の簡単な説明】

【0013】

【図1】図1は、実施例に係るストレージシステムの構成を示す図である。

【図2】図2は、ストレージ装置の構成を示す図である。

【図3】図3は、運用管理装置の機能構成を示す図である。

【図4】図4は、ストレージ管理部の機能構成を示す図である。

【図5】図5は、ポリシー管理テーブルの一例を示す図である。

【図6】図6は、プール管理テーブルの一例を示す図である。

30

【図7】図7は、性能情報記憶部が各ブロックについて記憶する情報の一例を示す図である。

【図8】図8は、評価結果記憶部が記憶する評価情報の一例を示す図である。

【図9A】図9Aは、アクセスリストの一例を示す図である。

【図9B】図9Bは、再配置先割当後のアクセスリストを示す図である。

【図10】図10は、比率指定に基づく割当を説明するための図である。

【図11】図11は、再配置リストの一例を示す図である。

【図12】図12は、再配置の対象とならないブロックを説明するための図である。

【図13】図13は、自動階層制御部の機能構成を示す図である。

【図14】図14は、CMの機能構成を示す図である。

40

【図15】図15は、担当CMを説明するための図である。

【図16】図16は、ストレージシステムの処理フローを示すフローチャートである。

【図17】図17は、ポリシー登録処理のフローを示すフローチャートである。

【図18】図18は、ティアプール登録処理のフローを示すフローチャートである。

【図19】図19は、ボリューム作成処理のフローを示すフローチャートである。

【図20】図20は、性能情報収集処理のフローを示すフローチャートである。

【図21】図21は、再配置処理のフローを示すフローチャートである。

【図22】図22は、再配置結果表示処理のフローを示すフローチャートである。

【図23】図23は、運用管理プログラムを実行するコンピュータのハードウェア構成を示す図である。

50

【図 2 4】図 2 4 は、C M のハードウェア構成を示す図である。

【図 2 5】図 2 5 は、階層制御を説明するための図である。

【発明を実施するための形態】

【0014】

以下に、本願の開示するストレージシステム、運用管理方法及び運用管理プログラムの実施例を図面に基づいて詳細に説明する。なお、この実施例は開示の技術を限定するものではない。また、以下では、HDD や SSD などのストレージの総称として「ディスク」を用いる。

【実施例】

【0015】

まず、実施例に係るストレージシステムの構成について説明する。図 1 は、実施例に係るストレージシステムの構成を示す図である。図 1 に示すように、ストレージシステム 1 は、ストレージ装置 2 と、運用管理装置 3 と、操作端末 4 とを有する。ストレージ装置 2 と、運用管理装置 3 と、操作端末 4 とは LAN (Local Area Network) 5 で接続される。

【0016】

ストレージ装置 2 は、SSD 及び HDD を有し、ティアプールを提供する。運用管理装置 3 は、ストレージ装置 2 の運用を管理し、AST 機能を有する。操作端末 4 は、ストレージシステム 1 の管理者 (以下、単に「管理者」という。) が運用管理装置 3 との対話に用いる装置である。

【0017】

図 2 は、ストレージ装置 2 の構成を示す図である。図 2 に示すように、ストレージ装置 2 は、2 台の CM (Controller Module) 2 1 と、DE (Drive Enclosure) 2 2 を有する。なお、ここでは、説明の便宜上 1 台の DE 2 2 を示すが、ストレージ装置 2 は複数の DE 2 2 を有する。

【0018】

CM 2 1 は、サーバ 7 からの要求に基づいてストレージ装置 2 を制御する装置であり、二重化されている。DE 2 2 は、SSD や HDD を搭載する機構であり、複数のディスク 6 を有し、低性能サブプール、中性能サブプール及び高性能サブプールの 3 つのサブプールを有するティアプールを提供する。

【0019】

各サブプールは複数の RAID (Redundant Arrays of Inexpensive Disks) グループを有する。例えば、低性能サブプールは、RAID グループ # 0 及び RAID グループ # 1 を有し、中性能サブプールは、RAID グループ # 2 及び RAID グループ # 3 を有する。また、高性能サブプールは、RAID グループ # 4 及び RAID グループ # 5 を有する。また、各 RAID グループは複数のディスク 6 を有する。

【0020】

次に、運用管理装置 3 の機能構成について説明する。図 3 は、運用管理装置 3 の機能構成を示す図である。図 3 に示すように、運用管理装置 3 は、ストレージ管理部 3 1 と、自動階層制御部 3 2 と、性能情報記憶部 3 3 と、評価結果記憶部 3 4 とを有する。

【0021】

ストレージ管理部 3 1 は、ストレージ装置 2 を管理する。具体的には、ストレージ管理部 3 1 は、階層化ポリシーの管理、ティアプールの管理、ボリュームの管理、階層化ポリシーに基づく再配置ブロックの特定などを行う。ここで、ブロックとは、再配置されるデータの単位であり、例えば、大きさが 1 . 3 ギガバイト (GB) のデータである。

【0022】

自動階層制御部 3 2 は、ストレージ管理部 3 1 の指示に基づいて AST を行う。具体的には、自動階層制御部 3 2 は、ストレージ装置 2 から IOPS 値などの性能情報の取得、ストレージ装置 2 への階層制御開始の通知、再配置の依頼などを行う。また、自動階層制御部 3 2 は、ストレージ装置 2 にティアプールの作成、ボリュームの作成などの依頼を行

10

20

30

40

50

う。

【 0 0 2 3 】

性能情報記憶部 3 3 は、自動階層制御部 3 2 がストレージ装置 2 から取得した I O P S 値などの性能情報を記憶する。また、性能情報記憶部 3 3 が記憶する情報は、ストレージ管理部 3 1 が再配置ブロックを特定する場合に用いられる。なお、性能情報記憶部 3 3 の詳細については後述する。

【 0 0 2 4 】

評価結果記憶部 3 4 は、性能情報記憶部 3 3 が記憶する性能情報をストレージ管理部 3 1 が解析して評価した結果を記憶する。なお、評価結果記憶部 3 4 の詳細については後述する。

10

【 0 0 2 5 】

次に、ストレージ管理部 3 1 の機能構成について説明する。図 4 は、ストレージ管理部 3 1 の機能構成を示す図である。図 4 に示すように、ストレージ管理部 3 1 は、ポリシー管理部 4 1 と、ポリシー管理テーブル 4 2 と、プール管理部 4 3 と、プール管理テーブル 4 4 と、ボリューム管理部 4 5 とを有する。また、ストレージ管理部 3 1 は、ボリューム管理テーブル 4 6 と、解析部 4 7 と、割当部 4 8 と、再配置特定部 4 9 と、再配置指示部 5 0 と、再配置確認部 5 1 と、通信部 5 2 と、装置間通信部 5 3 とを有する。

【 0 0 2 6 】

ポリシー管理部 4 1 は、操作端末 4 から階層化ポリシー作成指示を受け付けて、階層化ポリシーをポリシー管理テーブル 4 2 に登録し、登録した階層化ポリシーの情報を表示するように操作端末 4 に指示する。

20

【 0 0 2 7 】

ポリシー管理テーブル 4 2 は、ポリシー管理部 4 1 により登録された階層化ポリシーの情報を記憶する。図 5 は、ポリシー管理テーブル 4 2 の一例を示す図である。図 5 に示すように、ポリシー管理テーブル 4 2 は、階層化ポリシー名、実行モード、評価対象データ種別、評価基準、評価間隔、評価期間、評価対象時間帯、評価・再配置実行時間、再配置打切時間、低と中と高のレベルの範囲を、階層化ポリシー毎に記憶する。

【 0 0 2 8 】

階層化ポリシー名は、階層化ポリシーを識別する名前である。実行モードは、再配置を実行するモードを示し、再配置を自動で実行する「auto」と、管理者との対話に基づいて再配置を半自動で実行する「semi-auto」と、再配置を管理者の指示により実行する「manual」とがある。評価対象データ種別は、再配置の対象を評価する評価データの種別を示し、I O P S 値を示す「iops」がある。

30

【 0 0 2 9 】

評価基準は、評価対象データ種別で指定される評価データを評価する基準を示し、ピーク値を基準として評価する「peak」と平均値を基準として評価する「average」とがある。評価間隔は、評価データを評価する間隔の単位を示し、「day」は単位が「日」であることを示し、「time」は単位が「時間」であることを示す。

【 0 0 3 0 】

評価期間は、評価を行う間隔を示し、単位は評価間隔で指定される。例えば、評価間隔が「day」であり、評価期間が「7」である場合には、「7日」ごとに評価が行われる。評価対象時間帯は、データが評価対象となる時間帯を開始時刻と終了時刻で示す。評価・再配置実行時間は、評価及び再配置を行う時刻を示す。再配置打切時間は、再配置を打ち切る時間を示す。

40

【 0 0 3 1 】

低レベルの範囲は、低性能サブプールに配置すべきデータの I O P S 値の範囲を示し、中レベルの範囲は、中性能サブプールに配置すべきデータの I O P S 値の範囲を示し、高レベルの範囲は、高性能サブプールに配置すべきデータの I O P S 値の範囲を示す。

【 0 0 3 2 】

例えば、階層化ポリシー名が「Policy__Auto__01」である階層化ポリシーでは、評価は

50

自動で行われ、IOPS値を用いて評価が行われ、ピーク値を規準として評価が行われ、「7日」ごとに評価が行われる。データの評価対象時間帯の開始時刻は6時であり、終了時刻は20時である。評価・再配置は1時に実行され、24時間経過すると評価・再配置は打ち切られる。IOPS値が「100」以内のデータは低性能サブプールに記憶され、IOPS値が「101」～「500」のデータは中性能サブプールに記憶され、IOPS値が「501」以上のデータは高性能サブプールに記憶される。

【0033】

図4に戻って、プール管理部43は、操作端末4からティアプール作成指示を受け付けて、ティアプールの作成を自動階層制御部32に指示する。そして、プール管理部43は、作成されたティアプールをプール管理テーブル44に登録し、登録したティアプールの情報を表示するように操作端末4に指示する。

10

【0034】

また、プール管理部43は、ティアプールを作成する際、ディスク6が指定されているか否かを判定し、ディスク6が指定されていない場合には、管理者により指定された容量でティアプールを作成するように自動階層制御部32経由でストレージ装置2に指示する。プール管理部43が、容量指定でティアプールを作成するようにストレージ装置2に指示することによって、管理者は、ディスク6の搭載状況を把握することなく、簡単にティアプールを作成することができる。

【0035】

また、プール管理部43は、管理者からの指定に基づいて、1つのサブプールから構成されるティアプールを作成し、プール管理テーブル44に登録する。プール管理部43が、1つのサブプールから構成される階層化ポリシーをプール管理テーブル44に登録することによって、管理者は、AST運用を行った場合と比較するための情報を得ることができる。

20

【0036】

プール管理テーブル44は、プール管理部43により登録されたティアプールの情報を記憶する。図6は、プール管理テーブル44の一例を示す図である。図6に示すように、プール管理テーブル44は、ティアプール名、階層化ポリシー名、装置IPアドレス、警告閾値、注意閾値、暗号化、担当CM番号、低性能サブプール情報、中性能サブプール情報及び高性能サブプール情報をティアプール毎に記憶する。

30

【0037】

ティアプール名は、ティアプールを識別する名前である。階層化ポリシー名は、ティアプールに適用されるポリシーを示す。装置IPアドレスは、ティアプールを提供するストレージ装置2のIPアドレスを示す。警告閾値は、越えると警告を出力する閾値を示し、単位は%である。注意閾値は、越えると注意を出力する閾値を示し、単位は%である。暗号化は、ティアプールに記憶するデータを暗号化するか否かを示し、担当CM番号は、ティアプールの管理を担当するCM21の番号を示す。なお、担当CMの詳細については後述する。

【0038】

低性能サブプール情報は、ティアプールに含まれる低性能サブプールの情報を示し、中性能サブプール情報は、ティアプールに含まれる中性能サブプールの情報を示し、高性能サブプール情報は、ティアプールに含まれる高性能サブプールの情報を示す。

40

【0039】

サブプールの情報には、サブプール名、RAIDレベル、ディスク情報、担当CMがある。サブプール名は、サブプールを識別する名前である。RAIDレベルは、サブプールのRAIDレベルを示し、「RAID5」、「RAID1+0」などがある。ディスク情報は、サブプールを構成するディスクのRAIDグループの番号とグループ内の番号である。担当CMは、サブプールを担当するCM21の番号を示す。

【0040】

図4に戻って、ボリューム管理部45は、操作端末4からボリューム作成指示を受け付

50

けて、自動階層制御部 3 2 にボリュームの作成を指示する。また、ボリューム管理部 4 5 は、ストレージ装置 2 に作成されたボリュームをボリューム管理テーブル 4 6 に登録し、登録したボリュームの情報を表示するように操作端末 4 に指示する。また、ボリューム管理部 4 5 は、ティアプール内のサブプールの比率が管理者により指定された場合には、ボリューム管理テーブル 4 6 に比率を登録する。ボリューム管理テーブル 4 6 は、ボリューム管理部 4 5 により登録されたボリュームの情報を記憶する。

【 0 0 4 1 】

解析部 4 7 は、性能情報記憶部 3 3 が記憶する性能情報の解析及び評価を行い、評価結果を評価結果記憶部 3 4 に保存する。図 7 は、性能情報記憶部 3 3 が各ブロックについて記憶する情報の一例を示す図である。図 7 に示すように、性能情報記憶部 3 3 が各ブロックについて記憶する情報には、BNo、OLUN、RLUN、ReadIOPSPeak、WriteIOPSPeak、TotalIOPSPeakが含まれる。BNoは、ブロック番号である。OLUNは、ブロックが属しているボリューム番号であり、RLUNはブロックが属している R A I D グループ番号である。ReadIOPSPeakは、ReadIO回数のピーク値であり、WriteIOPSPeakは、WriteIO回数のピーク値であり、TotalIOPSPeakは、ReadIO回数とWriteIO回数の和のピーク値である。解析部 4 7 は、例えば、TotalIOPSPeakを用いて各ブロックのアクセス頻度を判定する。

【 0 0 4 2 】

性能情報記憶部 3 3 は、1 回に収集された性能情報をIOPS-PK__ii__YYYYMMDD@HHMM-yyyymmdd@hhmm.CSVの名前のファイルに記憶する。ここで、iiは、性能監視間隔を分で示す。また、YYYYMMDDは、性能情報採取開始年月日であり、HHMMは、性能情報採取開始時刻である。また、yyyymmddは、性能情報採取終了年月日であり、hhmmは、性能情報採取終了時刻である。

【 0 0 4 3 】

図 8 は、評価結果記憶部 3 4 が記憶する評価情報の一例を示す図である。図 8 に示すように、評価情報には、評価実施日時と、ティアプール名と階層化ポリシー名と評価データが含まれる。評価データには、VolNo、UpgradeToHigh、UpgradeToMiddle、DowngradeToMiddle、DowngradeToLow、KeepLow、KeepMiddle、KeepHigh、Statusが含まれる。

【 0 0 4 4 】

VolNoは、ボリュームを識別するボリューム番号である。UpgradeToHighは、上位階層であるティアレベル = 高のサブプールすなわち高性能サブプールにマイグレーションすべきであると評価されたブロックの割合である。UpgradeToMiddleは、上位階層であるティアレベル = 中のサブプールすなわち中性能サブプールにマイグレーションすべきであると評価されたブロックの割合である。DowngradeToMiddleは、下位階層であるティアレベル = 中のサブプールにマイグレーションすべきであると評価されたブロックの割合である。DowngradeToLowは、下位階層であるティアレベル = 低のサブプールすなわち低性能サブプールにマイグレーションすべきであると評価されたブロックの割合である。

【 0 0 4 5 】

KeepLowは、ティアレベル = 低でマイグレーションを実施しない評価となったブロックの割合である。KeepMiddleは、ティアレベル = 中でマイグレーションを実施しない評価となったブロックの割合である。KeepHighは、ティアレベル = 高でマイグレーションを実施しない評価となったブロックの割合である。Statusは、ボリュームの再配置実行状況である。

【 0 0 4 6 】

また、解析部 4 7 は、評価結果に基づき、アクセス頻度の高い順にブロックがならんだアクセスリストを作成する。図 9 A は、解析部 4 7 が作成するアクセスリストの一例を示す図である。図 9 A に示すように、アクセスリスト 5 5 は、アクセス頻度の高い順にブロック番号とブロックに現在の配置場所を示す再配置元と再配置先とを対応させたリストである。図 9 A では、ブロック番号が「 5 」のブロックのアクセス頻度が最も高く、再配置元は「高」であり、ブロック番号が「 9 」のブロックのアクセス頻度が次に高く、再配置元は「中」である。なお、再配置先は、割当部 4 8 により割り当てられる。

【 0 0 4 7 】

割当部 4 8 は、階層化ポリシー及びボリュームの情報に基づいてブロックの再配置先を割り当て、アクセスリスト 5 5 に書き込む。図 9 B は、図 9 A に示したアクセスリスト 5 5 に再配置先が割当部 4 8 により割り当てられた後のアクセスリスト 5 5 を示す図である。図 9 B に示すように、ブロック番号が「 5 」のブロックは、再配置先として「高」が割り当てられ、ブロック番号が「 9 」のブロックは、再配置先として「高」が割り当てられている。アクセスリスト 5 5 の配置先は、上から「高」が並び、次に「中」が並び、最後に「低」が並び。

【 0 0 4 8 】

割当部 4 8 は、ポリシー判定部 4 8 a と、第 1 割当部 4 8 b と、第 2 割当部 4 8 c と、第 3 割当部 4 8 d とを有する。ポリシー判定部 4 8 a は、ポリシー管理テーブル 4 2 が記憶する階層化ポリシーの情報及びボリュームの情報に基づいてブロックの割当方法を判定する。割当方法としては、容量指定に基づく割当、比率指定に基づく割当、1つのサブプール指定に基づく割当、IOPS 値指定に基づく割当がある。

10

【 0 0 4 9 】

第 1 割当部 4 8 b は、容量指定に基づく割当を行う。具体的には、第 1 割当部 4 8 b は、各サブプールの容量の範囲内で高性能のサブプールを優先させてアクセス頻度の高いブロックから順にサブプールに割り当てる。第 1 割当部 4 8 b が、容量指定に基づく割当を行うことによって、ストレージシステム 1 は、高性能サブプールを常に無駄なく使用することができる。なお、ポリシー判定部 4 8 a は、ポリシー管理テーブル 4 2 の低レベルの範囲、中レベルの範囲及び高レベルの範囲に値が指定されていないときに容量指定に基づく割当であると判定する。容量指定に基づく割当であるか否かは、ポリシー管理部 4 1 が階層化ポリシー作成時にポリシー管理テーブル 4 2 に登録する。

20

【 0 0 5 0 】

第 2 割当部 4 8 c は、ボリューム単位に比率指定に基づく割当を行う。具体的には、第 2 割当部 4 8 c は、指定されたボリュームに対して指定された比率及びアクセス頻度に基づいて、ブロックを低性能サブプール、中性能サブプール、高性能サブプールに割当を行う。図 1 0 は、比率指定に基づく割当を説明するための図である。図 1 0 に示すように、10 個のブロック # 0 1 ~ ブロック # 1 0 を有するボリューム # 1 に対して低割当率 2 0 %、中割当率 5 0 %、高割当率 3 0 % が指定されると、低性能サブプールには、アクセス頻度が最も低い方からブロック # 0 1 及び # 0 9 が割り当てられる。また、中性能サブプールには、次にアクセス頻度の低い 5 つのブロック # 0 3、# 0 4、# 0 6、# 0 8 及び # 1 0 が割り当てられ、高性能サブプールには、最もアクセス頻度の高い 3 つのブロック # 0 2、# 0 5 及び # 0 7 が割り当てられる。第 2 割当部 4 8 c が、ボリューム単位に比率指定に基づく割当を行うことによって、管理者は、ボリューム毎にサブプールを使い分けることができ、業務に適したティアプールを作成することができる。なお、ポリシー判定部 4 8 a は、ボリューム管理テーブル 4 6 が記憶するボリューム情報に比率指定があるときに比率指定に基づく割当であると判定する。

30

【 0 0 5 1 】

第 3 割当部 4 8 d は、IOPS 値指定に基づく割当を行う。すなわち、第 3 割当部 4 8 d は、従来と同様に、階層化ポリシーで指定された IOPS 値に基づいて、ブロックをサブプールに割り当てる。

40

【 0 0 5 2 】

なお、サブプールが 1 つだけ指定された場合には、再配置は行われない。管理者は、1つのサブプールに全てのブロックを割り当てることによって、AST 運用を行った場合と比較するための情報を得ることができる。

【 0 0 5 3 】

再配置特定部 4 9 は、アクセスリスト 5 5 から再配置元と再配置先が異なるブロックを取り出して再配置リストを作成する。図 1 1 は、再配置リストの一例を示す図である。図 1 1 に示すように、再配置リスト 5 6 は、再配置の対象となるブロック毎に、ブロック番

50

号と再配置元と再配置先を対応付けたリストである。例えば、ブロック番号「9」のブロックは再配置対象であり、高性能サブプールから中性能サブプールに再配置される。

【0054】

再配置リスト56に登録されるブロックの個数は64個である。再配置特定部49は、再配置リスト56に登録されたブロックの数が64になると、再配置指示部50に再配置を要求し、再配置リスト56をクリアする。そして、再配置特定部49は、アクセスリスト55から再配置元と再配置先が異なるブロックがなくなるまで、再配置リスト56の作成、再配置の要求、再配置リスト56のクリアを繰り返す。

【0055】

また、再配置特定部49は、アクセスリスト55において、隣接するサブプールの容量の5%に含まれるブロックについては隣接するサブプールへの再配置の対象としない。図12は、再配置の対象とならないブロックを説明するための図である。図12は、中性能サブプールの容量が130GBであり、高性能サブプールに再配置されるブロックのうち中性能サブプールに隣接するブロック番号が「101」、「105」、「103」、「107」及び「110」である5個のブロックを示している。ここで、5個のブロックは、隣接する中性能サブプールの容量130GBの5%すなわち6.5GBに含まれるブロックである。

【0056】

図12において、ブロック番号が「105」であるブロックは、中性能サブプールの容量の5%に含まれるブロックであるので、中性能サブプールから高性能サブプールへの再配置の対象とならない。一方、ブロック番号が「110」であるブロックは、中性能サブプールの容量の5%に含まれるブロックであるが、低性能サブプールから高性能サブプールへの再配置であるので、再配置の対象となる。

【0057】

このように、再配置特定部49は、アクセスリスト55において、隣接するサブプールの容量の5%に含まれるブロックについては隣接するサブプールへの再配置の対象としない。したがって、ストレージ装置2は、隣接するサブプールに記憶されるブロックとIOPS数にあまり差がないブロックを隣接するサブプールに移動するような、効果の少ない再配置の実行を防ぐことができる。なお、ここでは、5%を用いたが、再配置特定部49は、任意の割合を5%の代わりに用いることができる。

【0058】

再配置指示部50は、再配置特定部49からの要求に基づいて自動階層制御部32に再配置リスト56に登録されているブロックの再配置を指示する。また、再配置指示部50は、自動階層制御部32から再配置に関する情報を受け取り、操作端末4に再配置に関する情報を表示するように指示する。

【0059】

再配置確認部51は、操作端末4からの指示に基づいて、自動階層制御部32に再配置の進捗情報の取得を指示し、進捗情報を受け取ると、端末装置に進捗情報を送信し、再配置結果として表示するように指示する。

【0060】

通信部52は、自動階層制御部32と通信を行い、装置間通信部53は、操作端末4及びストレージ装置2とLAN5を介して通信を行う。

【0061】

次に、自動階層制御部32の機能構成について説明する。図13は、自動階層制御部32の機能構成を示す図である。図13に示すように、自動階層制御部32は、割当状態確認部61と、プール作成依頼部62と、ボリューム作成依頼部63と、開始通知部64と、性能情報取得部65と、再配置依頼部66と、進捗情報取得部67とを有する。また、自動階層制御部32は、通信部68と、装置間通信部69とを有する。

【0062】

割当状態確認部61は、プール管理部43からの指示に基づいて、ディスク6の割当状

10

20

30

40

50

態をストレージ装置 2 に確認し、ディスク 6 の割当状態に関する情報をプール管理部 4 3 に応答する。

【 0 0 6 3 】

プール作成依頼部 6 2 は、プール管理部 4 3 からの指示に基づいて、ストレージ装置 2 にティアプールの作成を依頼し、作成されたティアプールに関する情報をプール管理部 4 3 に応答する。

【 0 0 6 4 】

ボリューム作成依頼部 6 3 は、ボリューム管理部 4 5 からの指示に基づいて、ストレージ装置 2 にボリュームの作成を依頼し、作成されたボリュームに関する情報をボリューム管理部 4 5 に応答する。

10

【 0 0 6 5 】

開始通知部 6 4 は、操作端末 4 からの階層制御の開始指示を受け付けて、ストレージ装置 2 に階層制御の開始を通知する。性能情報取得部 6 5 は、例えば、5 分間隔でストレージ装置 2 に性能情報の取得を要求し、ストレージ装置 2 から性能情報を取得する。また、性能情報取得部 6 5 は、取得した性能情報を性能情報記憶部 3 3 に保存する。性能情報記憶部 3 3 は、階層化ポリシーで指定された評価期間内の性能情報を蓄積する。評価期間は、例えば、1 時間、1 日などである。なお、1 階層のティアプールについては、性能情報取得部 6 5 は、自動的に性能情報の取得を行い、性能情報記憶部 3 3 は、最大 7 日分の性能情報を保持する。

【 0 0 6 6 】

20

再配置依頼部 6 6 は、再配置指示部 5 0 からの再配置指示に基づいて、ストレージ装置 2 に再配置を依頼し、再配置先の R A I D グループなどの情報をストレージ装置 2 から受け取って再配置指示部 5 0 に応答する。

【 0 0 6 7 】

進捗情報取得部 6 7 は、再配置確認部 5 1 からの進捗確認指示に基づいて、ストレージ装置 2 に再配置の進捗情報の送信を依頼し、ストレージ装置 2 から送信された進捗情報を再配置確認部 5 1 に応答する。

【 0 0 6 8 】

通信部 6 8 は、ストレージ管理部 3 1 と通信を行い、装置間通信部 6 9 は、ストレージ装置 2 及び操作端末 4 と L A N 5 を介して通信を行う。

30

【 0 0 6 9 】

次に、C M 2 1 の機能構成について説明する。図 1 4 は、C M 2 1 の機能構成を示す図である。図 1 4 に示すように、C M 2 1 は、ディスク情報応答部 8 1 と、プール作成部 8 2 と、ボリューム作成部 8 3 と、性能情報収集部 8 4 と、再配置部 8 5 と、装置間通信部 8 6 とを有する。

【 0 0 7 0 】

ディスク情報応答部 8 1 は、割当状態確認部 6 1 からのディスク割当状態の確認指示に基づいて、ディスクの割当情報を割当状態確認部 6 1 に応答する。

【 0 0 7 1 】

プール作成部 8 2 は、プール作成依頼部 6 2 からのティアプール作成依頼に基づいて、ティアプールを作成し、作成したティアプールに関する情報をプール作成依頼部 6 2 に応答する。

40

【 0 0 7 2 】

ボリューム作成部 8 3 は、ボリューム作成依頼部 6 3 からのボリューム作成依頼に基づいて、ボリュームを作成し、作成したボリュームに関する情報をボリューム作成依頼部 6 3 に応答する。

【 0 0 7 3 】

性能情報収集部 8 4 は、性能情報取得部 6 5 からの性能情報の取得依頼に基づいて、性能情報を収集し、収集した情報を性能情報取得部 6 5 に応答する。性能情報には、ブロック毎の I O P S 値が含まれる。

50

【 0 0 7 4 】

再配置部 8 5 は、再配置依頼部 6 6 からの再配置依頼に基づいて、再配置リスト 5 6 で指定されたブロックを再配置し、再配置対象の情報を再配置依頼部 6 6 に応答する。再配置部 8 5 は、ブロックをサブプール間で移動する際に、移動先のディスク領域をボリュームを担当する C M 2 1 と同じ C M 2 1 が担当する R A I D グループから確保する。

【 0 0 7 5 】

図 1 5 は、担当 C M 2 1 を説明するための図である。図 1 5 に示すように、C M # 0 は、ボリュームについてはボリューム # 0 ~ # 3 を担当し、R A I D グループについては低性能サブプールに含まれる R A I D グループ # 0 と中性能サブプールに含まれる R A I D # 2 を担当する。C M # 1 は、ボリュームについてはボリューム # 4 ~ # 7 を担当し、R A I D グループについては低性能サブプールに含まれる R A I D グループ # 1 と中性能サブプールに含まれる R A I D グループ # 3 を担当する。

10

【 0 0 7 6 】

したがって、図 1 5 に示す C M 担当で、例えば、ボリューム # 0 のブロックを低性能サブプールに再配置する場合、再配置部 8 5 は、ボリューム # 0 を担当する C M # 0 が担当する R A I D グループ # 0 からディスク領域を取得してブロックを再配置する。

【 0 0 7 7 】

このように、再配置部 8 5 が、移動先のディスク領域をボリュームを担当する C M 2 1 と同じ C M 2 1 が担当する R A I D グループから確保することで、担当 C M 2 1 が異なることに起因する I O 性能の劣化を抑えることができる。

20

【 0 0 7 8 】

装置間通信部 8 6 は、運用管理装置 3 との間で L A N 5 を介して通信を行う。例えば、装置間通信部 8 6 は、運用管理装置 3 からの指示を受信し、指示に対する応答を運用管理装置 3 に送信する。

【 0 0 7 9 】

次に、ストレージシステム 1 の処理フローについて説明する。図 1 6 は、ストレージシステム 1 の処理フローを示すフローチャートである。図 1 6 に示すように、まず、ストレージシステム 1 は、管理者からの指示に基づいて階層化ポリシーを作成し、ポリシー管理テーブル 4 2 に登録する（ステップ S 9 1）。

【 0 0 8 0 】

30

そして、ストレージシステム 1 は、管理者からの指示に基づいてティアプールを作成し、プール管理テーブル 4 4 に登録する（ステップ S 9 2）。そして、ストレージシステム 1 は、管理者からの指示に基づいてボリュームを作成し、ボリューム管理テーブル 4 6 に登録する（ステップ S 9 3）。

【 0 0 8 1 】

その後、サーバ 7 で動作する業務ソフトウェアがボリュームを使用して業務処理を開始する。そして、ストレージシステム 1 は、管理者からの指示に基づいて階層制御を開始し、性能情報を収集する（ステップ S 9 4）。

【 0 0 8 2 】

そして、ストレージシステム 1 は、ボリュームの情報、階層化ポリシー及び性能情報に基づいてブロックの再配置を実行する（ステップ S 9 5）。なお、ステップ S 9 4 及びステップ S 9 5 の処理は、業務処理中に、階層化ポリシーに基づいて実行される。そして、ストレージシステム 1 は、管理者から要求されると、再配置の進捗状況を確認して、再配置結果を表示する（ステップ S 9 6）。

40

【 0 0 8 3 】

このように、ボリューム情報、階層化ポリシー及び性能情報に基づいてブロックの再配置を実行することによって、ストレージシステム 1 は、I O 性能を向上することができる。

【 0 0 8 4 】

次に、ポリシー登録処理のフローについて説明する。図 1 7 は、ポリシー登録処理のフ

50

ローを示すフローチャートである。なお、図 17 の処理は、図 16 に示したステップ S 9 1 の処理に対応する。

【0085】

図 17 に示すように、操作端末 4 が管理者の指示に基づいて階層化ポリシーの作成を運用管理装置 3 に指示する（ステップ S 1）。すると、ストレージ管理部 31 は、管理者が指定したポリシー指定値を確認し（ステップ S 2）、ポリシー指定値をポリシー管理テーブル 42 に登録して保存する（ステップ S 3）。なお、ポリシー指定値には、容量指定に基づく割当であるか否かを示す情報が含まれる。そして、ストレージ管理部 31 は、操作端末 4 に登録した階層化ポリシーの情報を表示するように指示し、操作端末 4 が、階層化ポリシーの情報を表示する（ステップ S 4）。

10

【0086】

このように、管理者は階層化ポリシーを指定することで、ストレージシステム 1 を効率よく利用することができる。

【0087】

次に、ティアプール登録処理のフローについて説明する。図 18 は、ティアプール登録処理のフローを示すフローチャートである。なお、図 18 の処理は、図 16 に示したステップ S 9 2 の処理に対応する。

【0088】

図 18 に示すように、操作端末 4 が管理者の指示に基づいてティアプールの作成を運用管理装置 3 に指示する（ステップ S 5）。すると、ストレージ管理部 31 は、管理者が指示した内容を確認する（ステップ S 6）。管理者は、複数のティアから構成されるティアプールに加えて 1 つのティアから構成されるティアプールを指定することができる。

20

【0089】

そして、ストレージ管理部 31 は、担当 CM 21 が指定されているか否かを判定する（ステップ S 7）。その結果、担当 CM 21 が指定されている場合には、ストレージ管理部 31 は、指定された担当 CM 21 を指示情報に設定する（ステップ S 8）。一方、担当 CM 21 が指定されていない場合には、ストレージ管理部 31 は、担当 CM 21 は自動選択であることを示す「Auto」を指示情報に設定する（ステップ S 9）。

【0090】

そして、ストレージ管理部 31 は、ディスク指定か否かを判定し（ステップ S 10）、ディスク指定でない場合には、ディスク自動選択を指示情報に設定する（ステップ S 11）。そして、ストレージ管理部 31 は、指示情報とともにティアプール作成指示を自動階層制御部 32 に送信する（ステップ S 12）。

30

【0091】

一方、ディスク指定の場合には、ストレージ管理部 31 は、空きディスクの確認指示を自動階層制御部 32 に送信し（ステップ S 13）、自動階層制御部 32 は、ディスク割当状態の確認をストレージ装置 2 に指示する（ステップ S 14）。そして、ストレージ装置 2 がディスク割当状態についてのディスク情報を返信し（ステップ S 15）、自動階層制御部 32 は、ディスク情報をストレージ管理部 31 に送信する。そして、ストレージ管理部 31 は、ディスク指定の確認を行って（ステップ S 16）、選択の対象ディスクの情報を表示するように操作端末 4 に指示する。そして、操作端末 4 が対象ディスクの情報を表示し、管理者にディスクを選択させる（ステップ S 17）。そして、ストレージ管理部 31 は、選択された対象ディスクを指定して（ステップ S 18）、指示情報とともにティアプール作成指示を自動階層制御部 32 に送信する（ステップ S 19）。

40

【0092】

指示情報とともにティアプール作成指示を受信すると、自動階層制御部 32 は、ティアプールの作成をストレージ装置 2 に依頼し（ステップ S 20）、ストレージ装置 2 がティアプールを作成する（ステップ S 21）。そして、ストレージ装置 2 は、作成したティアプールの情報を自動階層制御部 32 に送信し、自動階層制御部 32 は、ティアプールの情報をストレージ管理部 31 に送信する。そして、ストレージ管理部 31 は、ティアプール

50

の情報を用いてティアプールをプール管理テーブル 4 4 に登録して保存し（ステップ S 2 2）、操作端末 4 にティアプールの情報を表示するように指示する。そして、操作端末 4 が、ティアプールの情報を表示する（ステップ S 2 3）。

【0093】

このように、ディスク指定でない場合に、ストレージ管理部 3 1 が、ディスク自動選択を指示情報に設定し、指示情報とともにティアプール作成を指示することによって、管理者は、簡単にティアプールを作成することができる。

【0094】

次に、ボリューム作成処理のフローについて説明する。図 1 9 は、ボリューム作成処理のフローを示すフローチャートである。なお、図 1 9 の処理は、図 1 6 に示したステップ S 9 3 の処理に対応する。

【0095】

図 1 9 に示すように、操作端末 4 が管理者の指示に基づいてボリュームの作成を運用管理装置 3 に指示する（ステップ S 2 4）。すると、ストレージ管理部 3 1 は、管理者が指示した内容を確認し（ステップ S 2 5）、指示内容で指定された対象ティアプールの情報をプール管理テーブル 4 4 から取得する（ステップ S 2 6）。なお、管理者が指定する内容にはサブプールの割当比率が含まれる場合がある。そして、ストレージ管理部 3 1 は、対象ティアプールの情報とともにボリューム作成指示を自動階層制御部 3 2 に送信し、自動階層制御部 3 2 は、ボリューム作成をストレージ装置 2 に依頼する（ステップ S 2 7）。

【0096】

そして、ストレージ装置 2 は、ボリュームを作成し（ステップ S 2 8）、作成したボリュームの情報を自動階層制御部 3 2 に送信し、自動階層制御部 3 2 は、ボリュームの情報をストレージ管理部 3 1 に送信する。

【0097】

そして、ストレージ管理部 3 1 は、ボリューム管理テーブル 4 6 にボリュームの情報を登録して保存し（ステップ S 2 9）、操作端末 4 にボリュームの情報を表示するように指示する。そして、操作端末 4 が、ボリュームの情報を表示する（ステップ S 3 0）。

【0098】

このように、ストレージ管理部 3 1 は、対象ティアプールの情報をプール管理テーブル 4 4 から取得して、対象ティアプールの情報とともにボリュームの作成を指示することによって、階層化ポリシーに基づくボリュームを作成することができる。

【0099】

次に、性能情報収集処理のフローについて説明する。図 2 0 は、性能情報収集処理のフローを示すフローチャートである。なお、図 2 0 の処理は、図 1 6 に示したステップ S 9 4 の処理に対応する。

【0100】

図 2 0 に示すように、操作端末 4 が管理者の指示に基づいて階層制御の開始を運用管理装置 3 に指示する（ステップ S 3 1）。すると、自動階層制御部 3 2 は、階層制御の開始をストレージ装置 2 に通知する（ステップ S 3 2）。すると、ストレージ装置 2 は、性能情報の収集を開始する（ステップ S 3 3）。

【0101】

そして、自動階層制御部 3 2 は、性能情報の取得をストレージ装置 2 に要求し（ステップ S 3 4）、ストレージ装置 2 は、収集した性能情報を返信する（ステップ S 3 5）。そして、自動階層制御部 3 2 は、性能情報を性能情報記憶部 3 3 に保存する（ステップ S 3 6）。なお、ステップ S 3 4 ~ ステップ S 3 6 の処理は、例えば、5 分間隔で定期的に実行される。

【0102】

このように、自動階層制御部 3 2 が、ストレージ装置 2 から受信した性能情報を性能情報記憶部 3 3 に保存することによって、ストレージ管理部 3 1 は、性能情報記憶部 3 3 を

10

20

30

40

50

参照してブロックを各サブプールに割り当てることができる。

【 0 1 0 3 】

次に、再配置処理のフローについて説明する。図 2 1 は、再配置処理のフローを示すフローチャートである。なお、図 2 1 の処理は、図 1 6 に示したステップ S 9 5 の処理に対応する。

【 0 1 0 4 】

図 2 1 に示すように、ストレージ管理部 3 1 は、性能情報記憶部 3 3 を参照し（ステップ S 3 7）、階層化ポリシーに従って性能情報を評価する（ステップ S 3 8）。そして、ストレージ管理部 3 1 は、比率指定に基づく割当か否かを判定する（ステップ S 3 9）。その結果、比率指定に基づく割当の場合には、ストレージ管理部 3 1 は、比率が指定されたボリュームの性能情報のみを性能情報記憶部 3 3 から抽出して I O P S 値でソートし、アクセスリスト 5 5 を作成する（ステップ S 4 0）。そして、ストレージ管理部 3 1 は、各サブプールの空き容量を確認し、配置可能なブロック数を算出する（ステップ S 4 1）。そして、ストレージ管理部 3 1 は、指定された比率に応じて再配置対象のブロックを特定し（ステップ S 4 2）、ステップ S 4 8 に進む。

【 0 1 0 5 】

一方、比率指定に基づく割当てない場合には、ストレージ管理部 3 1 は、性能情報記憶部 3 3 を参照し、ティアプール内のブロックを I O P S 値でソートし、アクセスリスト 5 5 を作成する（ステップ S 4 3）。そして、ストレージ管理部 3 1 は、容量指定に基づく割当であるか否かを判定し（ステップ S 4 4）、容量指定に基づく割当である場合には、各サブプールの空き容量を規準に、配置可能ブロック数を算出し（ステップ S 4 5）、ステップ S 4 8 へ進む。一方、容量指定に基づく割当てない場合には、ストレージ管理部 3 1 は、I O P S 値を規準に、再配置対象のブロックを特定し（ステップ S 4 6）、対象のブロックが再配置可能か空き容量を確認する（ステップ S 4 7）。

【 0 1 0 6 】

そして、ストレージ管理部 3 1 は、再配置対象のブロックと配置先のサブプールの情報を決定し（ステップ S 4 8）、決定した情報に基づいて、再配置の対象を決定する（ステップ S 4 9）。そして、ストレージ管理部 3 1 は、再配置リスト 5 6 を作成し（ステップ S 5 0）、自動階層制御部 3 2 に再配置を指示する。そして、自動階層制御部 3 2 がストレージ装置 2 に再配置を依頼する（ステップ S 5 1）。

【 0 1 0 7 】

そして、ストレージ装置 2 の C M 2 1 は、再配置を行うボリュームの担当 C M 2 1 を確認し（ステップ S 5 2）、再配置先のサブプールを確認する（ステップ S 5 3）。そして、ストレージ装置 2 は、ボリュームの担当 C M 2 1 と同一の担当 C M 2 1 の R A I D グループがサブプールにあるか否かを判定する（ステップ S 5 4）。

【 0 1 0 8 】

その結果、同一の担当 C M 2 1 の R A I D グループがサブプールにない場合には、ストレージ装置 2 は、サブプール内で最も割当量の少ない R A I D グループを検索し（ステップ S 5 5）、ステップ S 5 8 に進む。一方、同一の担当 C M 2 1 の R A I D グループがサブプールにある場合には、ストレージ装置 2 は、同一の担当 C M 2 1 の R A I D グループを抽出し（ステップ S 5 6）、抽出した R A I D グループで、最も割当量の少ない R A I D グループを検索する（ステップ S 5 7）。

【 0 1 0 9 】

そして、ストレージ装置 2 は、検索した R A I D グループを移動先 R A I D グループとして決定し（ステップ S 5 8）、運用管理装置 3 に移動先 R A I D グループの情報を送信し、決定した R A I D グループにデータを移動する（ステップ S 5 9）。

【 0 1 1 0 】

一方、ストレージ管理部 3 1 は、再配置の対象情報を操作端末 4 に通知し（ステップ S 6 0）、操作端末 4 は、再配置対象の情報を表示する（ステップ S 6 1）。なお、再配置されるブロックの数が 6 4 個を超える場合には、6 4 個のブロック毎にステップ S 5 0 ~

10

20

30

40

50

ステップ S 6 1 の処理が繰り返される。また、階層化ポリシーの指定に基づいて再配置処理が繰り返される。

【 0 1 1 1 】

このように、階層化ポリシーに基づいて再配置の対象を決定することにより、ストレージシステム 1 は、管理者のポリシーに基づいてブロックの再配置を制御することができる。

【 0 1 1 2 】

次に、再配置結果表示処理のフローについて説明する。図 2 2 は、再配置結果表示処理のフローを示すフローチャートである。なお、図 2 2 の処理は、図 1 6 に示したステップ S 9 6 の処理に対応する。

10

【 0 1 1 3 】

図 2 2 に示すように、操作端末 4 が管理者の指示に基づいて再配置結果の表示を運用管理装置 3 に指示する（ステップ S 6 2）。すると、ストレージ管理部 3 1 は、再配置の進捗状況の確認を自動階層制御部 3 2 に指示する（ステップ S 6 3）。

【 0 1 1 4 】

すると、自動階層制御部 3 2 が、再配置の情報取得をストレージ装置 2 に依頼し（ステップ S 6 4）、ストレージ装置 2 が、再配置の進捗情報を運用管理装置 3 に返信する（ステップ S 6 5）。すると、ストレージ管理部 3 1 は、操作端末 4 に再配置の状況を通知し（ステップ S 6 6）、操作端末 4 が、再配置結果を表示する（ステップ S 6 7）。

【 0 1 1 5 】

20

このように、操作端末 4 が、再配置結果を表示することによって、管理者は、階層化ポリシーやボリューム情報に基づく再配置を評価することができる。

【 0 1 1 6 】

上述してきたように、実施例では、ポリシー判定部 4 8 a が、ポリシー管理テーブル 4 2 が記憶する階層化ポリシーの情報及びボリューム管理テーブル 4 6 が記憶するボリュームの情報に基づいてブロックの割当方法を判定する。そして、ポリシー判定部 4 8 a の判定に基づいて、第 1 割当部 4 8 b は容量指定に基づく割当を行う。したがって、ストレージシステム 1 は、性能の異なる複数のストレージのうち高性能なストレージを有効に利用することができる。また、ポリシー判定部 4 8 a の判定に基づいて、第 2 割当部 4 8 c は、ボリューム単位に比率指定に基づく割当を行う。したがって、管理者は、ボリューム毎にサブプールを使い分けることができ、業務に適したティアプールを作成することができる。

30

【 0 1 1 7 】

また、実施例では、プール管理部 4 3 は、管理者からの指定に基づいて、1つのサブプールから構成されるティアプールを作成し、プール管理テーブル 4 4 に登録する。そして、性能情報取得部 6 5 は、1つのサブプールから構成されるティアプールについては、自動的に性能情報を取得する。したがって、管理者は、A S T 運用を行った場合と比較するための情報を得ることができる。

【 0 1 1 8 】

また、実施例では、C M 2 1 は、再配置を行うボリュームの担当 C M 2 1 と同一の担当 C M 2 1 の R A I D グループから再配置先の R A I D グループを決定する。したがって、ボリュームの担当 C M 2 1 と R A I D グループの担当の C M 2 1 が異なることに起因する I O 性能の劣化を抑えることができる。

40

【 0 1 1 9 】

また、実施例では、プール管理部 4 3 は、ティアプールを作成する際、ディスク 6 が指定されていない場合には、管理者により指定された容量でティアプールを作成するようにストレージ装置 2 に指示する。したがって、管理者は、ディスク 6 の搭載状況を把握することなく、簡単にティアプールを作成することができる。

【 0 1 2 0 】

また、実施例では、再配置特定部 4 9 は、アクセスリスト 5 5 において、隣接するサブ

50

プールの容量の5%に含まれるブロックについては隣接するサブプールへの再配置の対象としない。したがって、ストレージ装置2は、隣接するサブプールに記憶されるブロックとIOPS数にあまり差がないブロックを隣接するサブプールに移動するような、効果の少ない再配置の実行を防ぐことができる。

【0121】

なお、実施例では、運用管理装置について説明したが、運用管理装置が有する構成をソフトウェアによって実現することで、同様の機能を有する運用管理プログラムを得ることができる。そこで、運用管理プログラムを実行するコンピュータについて説明する。

【0122】

図23は、運用管理プログラムを実行するコンピュータのハードウェア構成を示す図である。図23に示すように、コンピュータ200は、メインメモリ210と、CPU(Central Processing Unit)220と、LAN(Local Area Network)インタフェース230と、HDD(Hard Disk Drive)240とを有する。また、コンピュータ200は、スーパーIO(Input Output)250と、DVI(Digital Visual Interface)260と、ODD(Optical Disk Drive)270とを有する。

10

【0123】

メインメモリ210は、プログラムやプログラムの実行途中結果などを記憶するメモリである。CPU220は、メインメモリ210からプログラムを読み出して実行する中央処理装置である。CPU220は、メモリコントローラを有するチップセットを含む。

【0124】

20

LANインタフェース230は、コンピュータ200をLAN経由で他のコンピュータに接続するためのインタフェースである。HDD240は、プログラムやデータを格納するディスク装置であり、スーパーIO250は、マウスやキーボードなどの入力装置を接続するためのインタフェースである。DVI260は、液晶表示装置を接続するインタフェースであり、ODD270は、DVDの読み書きを行う装置である。

【0125】

LANインタフェース230は、PCIエクスプレスによりCPU220に接続され、HDD240及びODD270は、SATA(Serial Advanced Technology Attachment)によりCPU220に接続される。スーパーIO250は、LPC(Low Pin Count)によりCPU220に接続される。

30

【0126】

そして、コンピュータ200において実行される運用管理プログラムは、DVDに記憶され、ODD270によってDVDから読み出されてコンピュータ200にインストールされる。あるいは、運用管理プログラムは、LANインタフェース230を介して接続された他のコンピュータシステムのデータベースなどに記憶され、これらのデータベースから読み出されてコンピュータ200にインストールされる。そして、インストールされた運用管理プログラムは、HDD240に記憶され、メインメモリ210に読み出されてCPU220によって実行される。

【0127】

次に、CM21のハードウェア構成について説明する。図24は、CM21のハードウェア構成を示す図である。図24に示すように、CM21は、CPU21aと、RAM21bと、フラッシュメモリ21cとを有する。

40

【0128】

CPU21aは、フラッシュメモリ21cからプログラムを読み出して実行する処理装置である。RAM(Random Access Memory)21bは、データを記憶する揮発性のメモリである。フラッシュメモリ21cは、図14に示したCM21の機能を実現するためのプログラムを記憶したメモリである。

【0129】

なお、実施例では、ストレージ装置2と運用管理装置3とを有するストレージシステム1について説明したが、本発明はこれに限定されるものではなく、運用管理装置3の機能

50

をストレージ装置が有するストレージシステムにも同様に適用することができる。

【0130】

また、実施例では、ティアプールが3つのサブプールを有する場合について説明したが、本発明はこれに限定されるものではなく、3より大きな任意の個数のサブプールを有するティアプールにも同様に適用することができる。

【0131】

また、実施例では、HDD及びSSDをディスクとして用いる場合について説明したが、本発明はこれに限定されるものではなく、他の記憶装置をディスクとして用いる場合にも同様に適用することができる。

【符号の説明】

10

【0132】

- 1 ストレージシステム
- 2 ストレージ装置
- 3 運用管理装置
- 4 操作端末
- 5 LAN
- 6 ディスク
- 7 サーバ
- 21 CM
- 21a CPU
- 21b RAM
- 21c フラッシュメモリ
- 22 DE
- 31 ストレージ管理部
- 32 自動階層制御部
- 33 性能情報記憶部
- 34 評価結果記憶部
- 41 ポリシー管理部
- 42 ポリシー管理テーブル
- 43 プール管理部
- 44 プール管理テーブル
- 45 ボリューム管理部
- 46 ボリューム管理テーブル
- 47 解析部
- 48 割当部
- 48a ポリシー判定部
- 48b 第1割当部
- 48c 第2割当部
- 48d 第3割当部
- 49 再配置特定部
- 50 再配置指示部
- 51 再配置確認部
- 52 通信部
- 53 装置間通信部
- 55 アクセスリスト
- 56 再配置リスト
- 61 割当状態確認部
- 62 プール作成依頼部
- 63 ボリューム作成依頼部
- 64 開始通知部

20

30

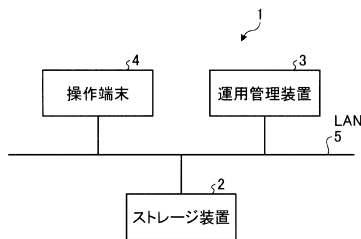
40

50

6 5	性能情報取得部
6 6	再配置依頼部
6 7	進捗情報取得部
6 8	通信部
6 9	装置間通信部
8 1	ディスク情報応答部
8 2	プール作成部
8 3	ボリューム作成部
8 4	性能情報収集部
8 5	再配置部
8 6	装置間通信部
2 0 0	コンピュータ
2 1 0	メインメモリ
2 2 0	C P U
2 3 0	L A N インタフェース
2 4 0	H D D
2 5 0	スー パー I O
2 6 0	D V I
2 7 0	O D D

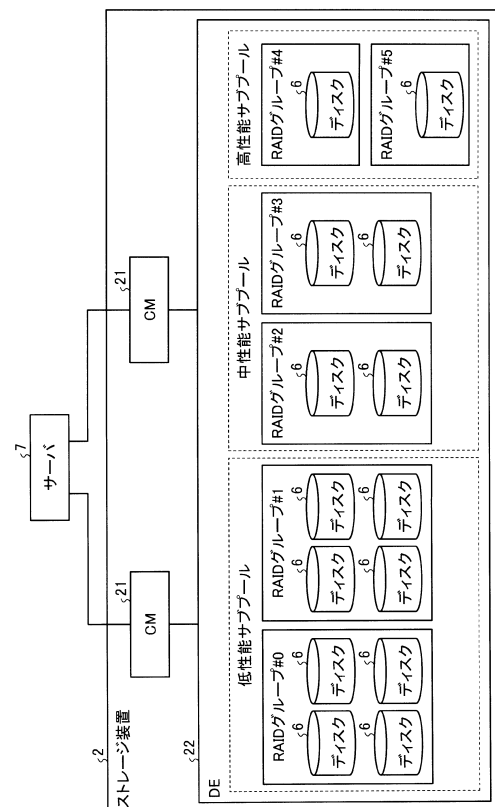
【図 1】

実施例に係るストレージシステムの構成を示す図



【図 2】

ストレージ装置の構成を示す図



【図 7】

性能情報記憶部が各ブロックについて記憶する情報の一例を示す図

情報	説明
BNo	ブロック番号
OLUN	ブロックが属しているボリューム番号
RLUN	ブロックが属しているRAIDグループ番号
ReadIOPSPeak	Read IO回数のピーク値(単位:回)
WriteIOPSPeak	Write IO回数のピーク値(単位:回)
TotalIOPSPeak	Read IO回数とWrite IO回数の和のピーク値(単位:回)

【図 8】

評価結果記憶部が記憶する評価情報の一例を示す図

情報	値
評価実施日時	文字列(yyyy/MM/ddHH:mm:ss)
ティアプール名	文字列
階層化ポリシー名	文字列
VolNo	当該ボリュームのボリューム番号
UpgradeToHigh	当該ボリュームに割り当てられた全ブロックのうち、上位階層であるティアレベル=高のサブプールへマイグレーションすべきであると評価されたブロックの割合(%)
UpgradeToMiddle	当該ボリュームに割り当てられた全ブロックのうち、上位階層であるティアレベル=中のサブプールへマイグレーションすべきであると評価されたブロックの割合(%)
DowngradeToMiddle	当該ボリュームに割り当てられた全ブロックのうち、下位階層であるティアレベル=中のサブプールへマイグレーションすべきであると評価されたブロックの割合(%)
DowngradeToLow	当該ボリュームに割り当てられた全ブロックのうち、下位階層であるティアレベル=低のサブプールへマイグレーションすべきであると評価されたブロックの割合(%)
KeepLow	当該ボリュームに割り当てられた全ブロックのうち、ティアレベル=低でマイグレーションを実施しない評価となったブロックの割合(%)
KeepMiddle	当該ボリュームに割り当てられた全ブロックのうち、ティアレベル=中でマイグレーションを実施しない評価となったブロックの割合(%)
KeepHigh	当該ボリュームに割り当てられた全ブロックのうち、ティアレベル=高でマイグレーションを実施しない評価となったブロックの割合(%)
Status	当該ボリュームの再配置実行状況

【図 9 A】

アクセスリストの一例を示す図

55

ブロック番号	再配置元	再配置先
5	高	-
9	中	-
⋮	⋮	⋮

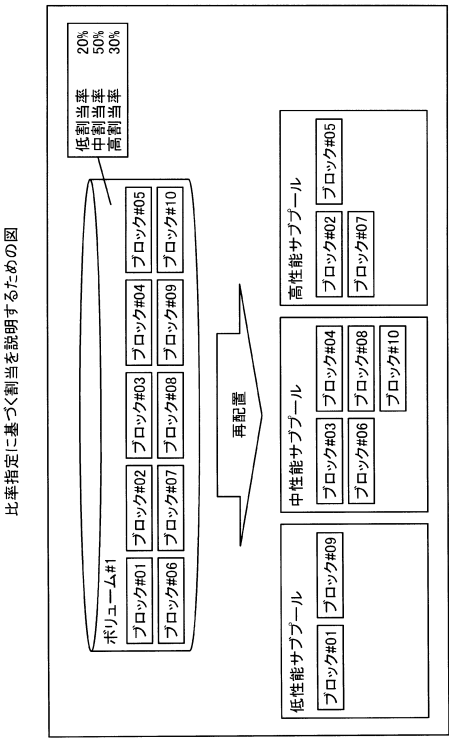
【図 9 B】

再配置先割当後のアクセスリストを示す図

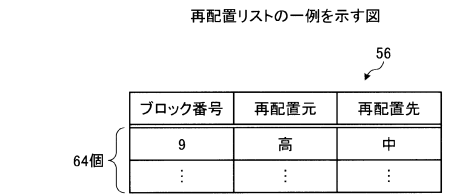
55

ブロック番号	再配置元	再配置先
5	高	高
9	中	高
⋮	⋮	⋮

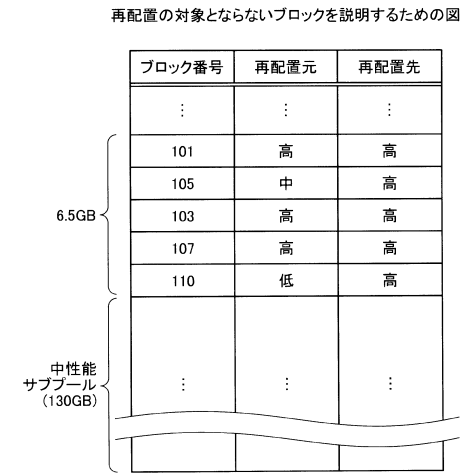
【図 10】



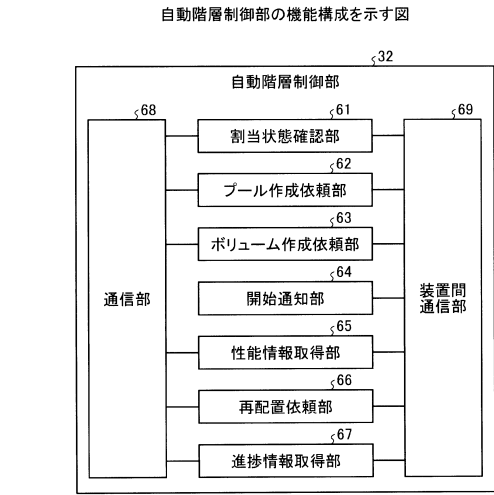
【図 1 1】



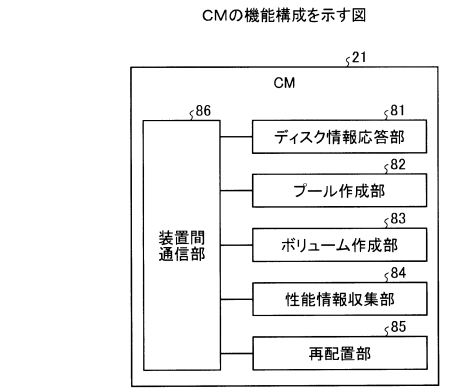
【図 1 2】



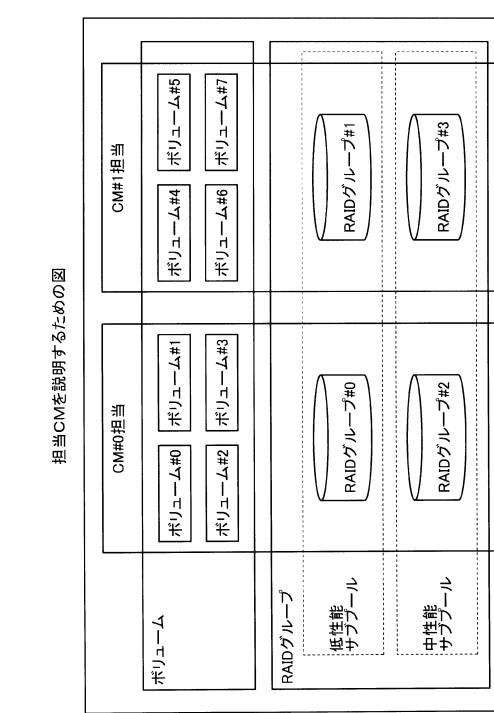
【図 1 3】



【図 1 4】

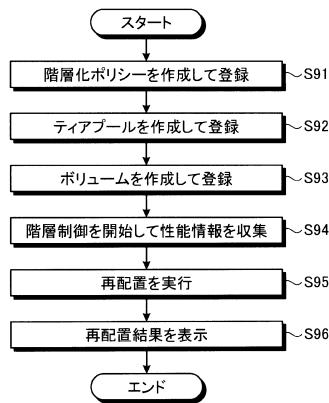


【図 1 5】



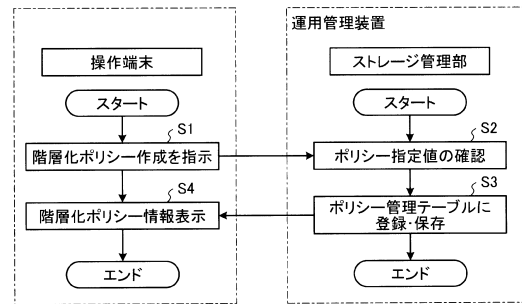
【 図 1 6 】

ストレージシステムの処理フローを示すフローチャート



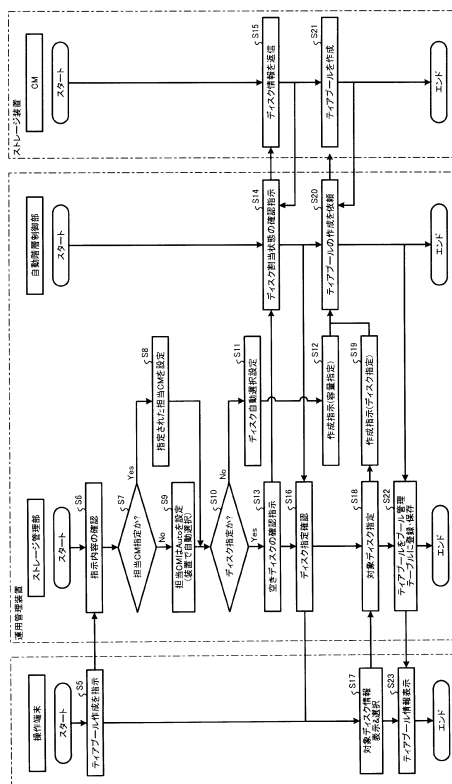
【 図 1 7 】

ポリシー登録処理のフローを示すフローチャート



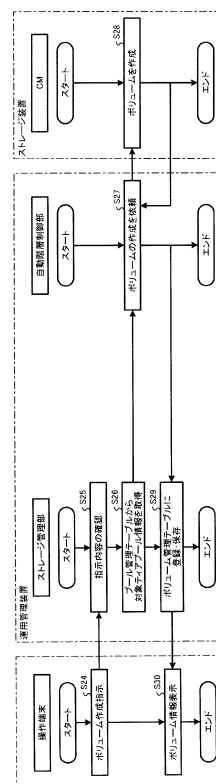
【 図 1 8 】

ティアプールの登録処理のフローを示すフローチャート



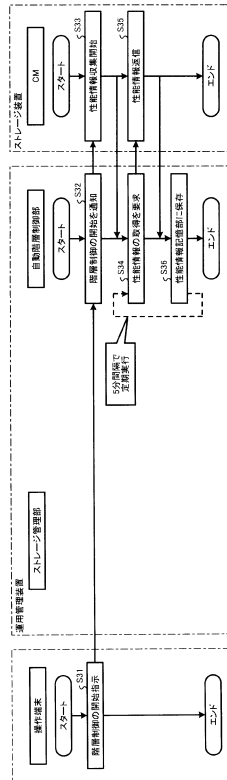
【 図 1 9 】

ポリューム作成処理のフローを示すフローチャート



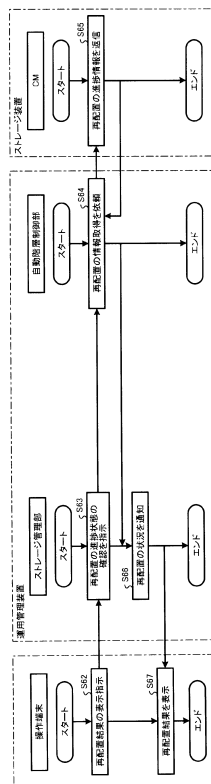
【図 20】

性能情報収集処理のフローを示すフローチャート



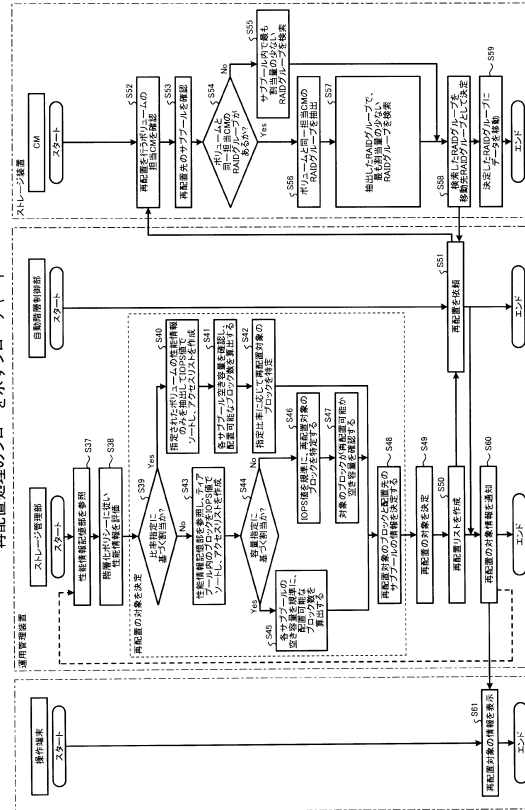
【図 22】

再配置結果表示処理のフローを示すフローチャート



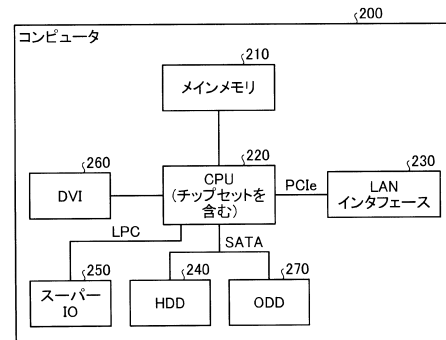
【図 21】

再配置処理のフローを示すフローチャート



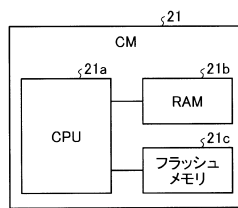
【図 23】

運用管理プログラムを実行するコンピュータのハードウェア構成を示す図



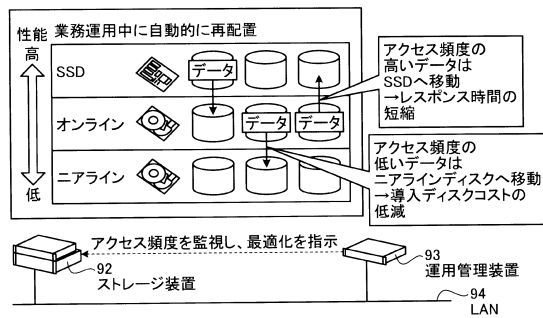
【図 24】

CMのハードウェア構成を示す図



【図 25】

階層制御を説明するための図



フロントページの続き

審査官 桜井 茂行

- (56)参考文献 国際公開第2013/018132(WO, A1)
特表2014-516442(JP, A)
国際公開第2012/104912(WO, A1)
特表2014-500530(JP, A)
国際公開第2013/018130(WO, A1)
特表2014-516426(JP, A)
特開2011-227898(JP, A)
特開平11-296313(JP, A)

(58)調査した分野(Int.Cl., DB名)

G06F 3/06 - 3/08
G06F 12/00