

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2020-4186

(P2020-4186A)

(43) 公開日 令和2年1月9日(2020.1.9)

(51) Int.Cl.	F I	テーマコード (参考)
G06F 3/06 (2006.01)	G06F 3/06 301Z	
G06F 12/00 (2006.01)	G06F 3/06 304N	
	G06F 3/06 304Z	
	G06F 12/00 501B	

審査請求 未請求 請求項の数 10 O L (全 32 頁)

(21) 出願番号	特願2018-124348 (P2018-124348)	(71) 出願人	000005223 富士通株式会社 神奈川県川崎市中原区上小田中4丁目1番1号
(22) 出願日	平成30年6月29日 (2018.6.29)	(74) 代理人	100092152 弁理士 服部 毅巖
		(72) 発明者	小林 祐一 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	宮内 啓次 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	松田 純一 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

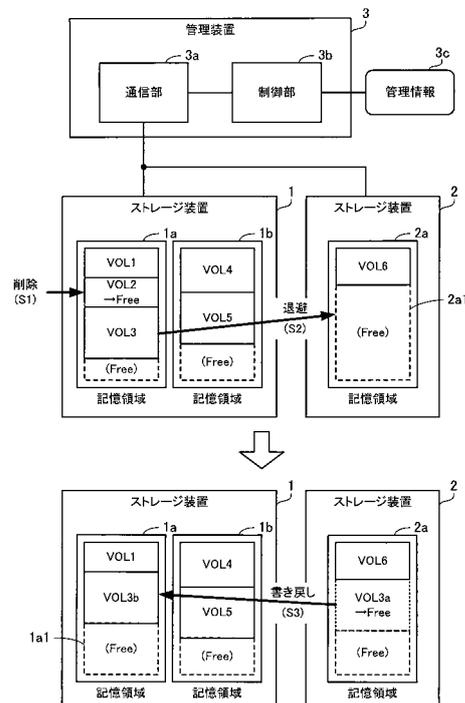
(54) 【発明の名称】 情報処理システム、管理装置および制御方法

(57) 【要約】

【課題】 ボリュームを削除した後の記憶領域の利用効率を向上させる。

【解決手段】 管理装置 3 は、ストレージ装置 1 の記憶領域 1 a に設定されたボリューム VOL 2 の削除が要求されたとき、管理情報 3 c に基づいて、記憶領域 1 a におけるボリューム VOL 2 の例えば下位アドレス側にボリューム VOL 3 が設定されていることを判定し、管理情報 3 c に基づいて、記憶領域 1 a、1 b、2 a の中からボリューム VOL 3 と同じ大きさ以上の連続した空き領域 2 a 1 を検出し、ボリューム VOL 2 を削除する削除処理と、ボリューム VOL 3 のデータを、空き領域 2 a 1 に退避させた後、記憶領域 1 a におけるボリューム VOL 2 が削除された領域を含む移動先領域に書き戻すことで、ボリューム VOL 3 を移動させる移動処理とを、ストレージ装置 1 に実行させる。

【選択図】 図 1



【特許請求の範囲】**【請求項 1】**

それぞれ 1 以上の記憶領域を有する複数のストレージ装置であって、前記複数のストレージ装置の前記記憶領域のそれぞれに対して 1 以上のボリュームを設定可能であり、設定されたボリュームは、前記複数のストレージ装置の前記記憶領域のうち当該ボリュームの設定先の記憶領域における当該ボリュームの設定領域と同じアドレス空間を有する、前記複数のストレージ装置と、管理装置とを有し、

前記管理装置は、

前記複数のストレージ装置の前記記憶領域のうち、第 1 のストレージ装置の第 1 の記憶領域に設定された第 1 のボリュームの削除が要求されたとき、前記複数のストレージ装置の前記記憶領域のそれぞれにおけるボリュームの設定状態を管理する管理情報に基づいて、前記第 1 の記憶領域に他のボリュームが設定されているかを判定し、

前記他のボリュームとして第 2 のボリュームが設定されていた場合、前記管理情報に基づいて、前記複数のストレージ装置の前記記憶領域の中から、前記第 2 のボリュームと同じ大きさ以上の連続した空き領域を検出し、

前記第 1 のボリュームを前記第 1 の記憶領域から削除する削除処理と、前記第 2 のボリュームのデータを、検出された前記空き領域に退避させた後、前記第 1 の記憶領域における前記第 1 のボリュームが削除された領域を含む移動先領域に書き戻すことで、前記第 2 のボリュームを移動させる移動処理を、前記第 1 のストレージ装置に実行させ、

前記第 1 のボリュームが削除され、前記第 2 のボリュームが移動したことを示すように前記管理情報を更新する、

情報処理システム。

【請求項 2】

前記記憶領域のそれぞれに対しては、上位アドレス側から順に前記 1 以上のボリュームを設定可能であり、

前記管理装置は、

前記第 1 のボリュームの削除が要求されたとき、前記第 1 の記憶領域における前記第 1 のボリュームの下位アドレス側に前記他のボリュームが設定されているかを判定し、

前記移動処理では、前記第 2 のボリュームのデータを、前記第 1 の記憶領域における前記第 1 のボリュームが削除された領域の先頭を起点とした前記移動先領域に書き戻す、

請求項 1 記載の情報処理システム。

【請求項 3】

前記記憶領域のそれぞれに対しては、下位アドレス側から順に前記 1 以上のボリュームを設定可能であり、

前記管理装置は、

前記第 1 のボリュームの削除が要求されたとき、前記第 1 の記憶領域における前記第 1 のボリュームの上位アドレス側に前記他のボリュームが設定されているかを判定し、

前記移動処理では、前記第 2 のボリュームのデータを、前記第 1 の記憶領域における前記第 1 のボリュームが削除された領域の末尾を後端とした前記移動先領域に書き戻す、

請求項 1 記載の情報処理システム。

【請求項 4】

前記移動処理は、検出された前記空き領域に前記第 2 のボリュームをコピーして第 3 のボリュームを生成し、前記第 2 のボリュームを削除し、前記移動先領域に前記第 3 のボリュームをコピーして移動後の前記第 2 のボリュームを生成し、前記第 3 のボリュームを前記空き領域から削除する処理を含む、

請求項 1 乃至 3 のいずれか 1 項に記載の情報処理システム。

【請求項 5】

前記情報処理システムは、前記複数のストレージ装置にアクセス可能な複数のサーバ装置をさらに有し、

前記管理情報には、前記第 2 のボリュームに対して、前記複数のサーバ装置のうち前記

10

20

30

40

50

第 2 のボリュームにアクセス可能な第 1 のサーバ装置が対応付けて登録され、

前記移動処理は、前記第 3 のボリュームの生成と前記第 2 のボリュームの削除との間に、前記第 1 のサーバ装置から前記第 2 のボリュームへのアクセスパスを前記第 3 のボリュームへのアクセスパスに切り替える第 1 の切り替え処理と、移動後の前記第 2 のボリュームの生成と前記第 3 のボリュームの削除との間に、前記第 1 のサーバ装置から前記第 3 のボリュームへのアクセスパスを移動後の前記第 2 のボリュームへのアクセスパスに切り替える第 2 の切り替え処理とを含み、

前記管理装置は、前記管理情報に基づいて前記第 1 の切り替え処理と前記第 2 の切り替え処理を前記第 1 のストレージ装置に実行させる、

請求項 4 記載の情報処理システム。

10

【請求項 6】

前記第 3 のボリュームの生成処理は、前記第 1 のサーバ装置から前記第 2 のボリュームへのアクセスを許容した状態のまま実行され、

移動後の前記第 2 のボリュームの生成処理は、前記第 1 のサーバ装置から前記第 3 のボリュームへのアクセスを許容した状態のまま実行される、

請求項 5 記載の情報処理システム。

【請求項 7】

前記第 1 のボリュームは、前記複数のサーバ装置のいずれかにおいて動作する仮想マシンに接続された仮想ディスクの記憶領域として使用され、

前記管理情報には、前記第 1 のボリュームと前記仮想ディスクとが対応付けて登録され

20

、前記管理装置は、前記仮想ディスクの削除が要求されたとき、前記管理情報に基づいて、前記仮想ディスクに対応付けられた前記第 1 のボリュームの削除が要求されたことを認識する、

請求項 5 または 6 記載の情報処理システム。

【請求項 8】

前記管理装置は、

前記移動処理を前記第 1 のストレージ装置に実行させる前に、検出された前記空き領域が予約されたことを示すように前記管理情報を更新し、

前記移動処理が完了すると、前記空き領域の予約が解除されたことを示すように前記管理情報を更新する、

30

請求項 1 乃至 7 のいずれか 1 項に記載の情報処理システム。

【請求項 9】

それぞれ 1 以上の記憶領域を有する複数のストレージ装置と通信する通信部であって、前記複数のストレージ装置の前記記憶領域のそれぞれに対して 1 以上のボリュームを設定可能であり、設定されたボリュームは、前記複数のストレージ装置の前記記憶領域のうち当該ボリュームの設定先の記憶領域における当該ボリュームの設定領域と同じアドレス空間を有する、前記通信部と、制御部とを有し、

前記制御部は、

前記複数のストレージ装置の前記記憶領域のうち、第 1 のストレージ装置の第 1 の記憶領域に設定された第 1 のボリュームの削除が要求されたとき、前記複数のストレージ装置の前記記憶領域のそれぞれにおけるボリュームの設定状態を管理する管理情報に基づいて、前記第 1 の記憶領域に他のボリュームが設定されているかを判定し、

40

前記他のボリュームとして第 2 のボリュームが設定されていた場合、前記管理情報に基づいて、前記複数のストレージ装置の前記記憶領域の中から、前記第 2 のボリュームと同じ大きさ以上の連続した空き領域を検出し、

前記第 1 のボリュームを前記第 1 の記憶領域から削除する削除処理と、前記第 2 のボリュームのデータを、検出された前記空き領域に退避させた後、前記第 1 の記憶領域における前記第 1 のボリュームが削除された領域を含む移動先領域に書き戻すことで、前記第 2 のボリュームを移動させる移動処理とを、前記第 1 のストレージ装置に実行させ、

50

前記第 1 のボリュームが削除され、前記第 2 のボリュームが移動したことを示すように前記管理情報を更新する、
管理装置。

【請求項 10】

コンピュータが、

それぞれ 1 以上の記憶領域を有する複数のストレージ装置と接続し、前記複数のストレージ装置の前記記憶領域のそれぞれに対して 1 以上のボリュームを設定可能であり、設定されたボリュームは、前記複数のストレージ装置の前記記憶領域のうち当該ボリュームの設定先の記憶領域における当該ボリュームの設定領域と同じアドレス空間を有し、

前記複数のストレージ装置の前記記憶領域のうち、第 1 のストレージ装置の第 1 の記憶領域に設定された第 1 のボリュームの削除が要求されたとき、前記複数のストレージ装置の前記記憶領域のそれぞれにおけるボリュームの設定状態を管理する管理情報に基づいて、前記第 1 の記憶領域に他のボリュームが設定されているかを判定し、

前記他のボリュームとして第 2 のボリュームが設定されていた場合、前記管理情報に基づいて、前記複数のストレージ装置の前記記憶領域の中から、前記第 2 のボリュームと同じ大きさ以上の連続した空き領域を検出し、

前記第 1 のボリュームを前記第 1 の記憶領域から削除する削除処理と、前記第 2 のボリュームのデータを、検出された前記空き領域に退避させた後、前記第 1 の記憶領域における前記第 1 のボリュームが削除された領域を含む移動先領域に書き戻すことで、前記第 2 のボリュームを移動させる移動処理を、前記第 1 のストレージ装置に実行させ、

前記第 1 のボリュームが削除され、前記第 2 のボリュームが移動したことを示すように前記管理情報を更新する、
制御方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、情報処理システム、管理装置および制御方法に関する。

【背景技術】

【0002】

ストレージシステムの一例として、ストレージ装置の記憶領域に対して複数の論理ボリュームが設定されて運用されるものがある。このようなストレージシステムでは、例えば、各論理ボリュームが個別のユーザに使用される、あるいは、各論理ボリュームに対して個別のサーバ装置からアクセスがなされる、といった運用を行うことができる。これにより、ストレージ装置の記憶領域を複数のユーザや複数のサーバ装置によって効率的に使用できる。

【0003】

また、複数の物理ディスクによって実現される論理記憶領域である R A I D (Redundant Arrays of Independent Disks) グループ上に、論理ボリュームが設定される場合もある。例えば、R A I D グループ上に複数の論理ボリュームが設定された状態から、格納されたデータを破壊することなく、物理ディスクの追加によって容量を拡張できるようにしたストレージ装置が提案されている。

【0004】

また、データ記憶に関連する技術として、記憶領域において断片化されたデータを連続領域に移動させる「デフラグメンテーション」と呼ばれる処理が知られている。例えば、H D D (Hard Disk Drive) がその内部の磁気ディスクについてのデフラグメンテーションを自律的に実行するようにした記録再生システムが提案されている。

【先行技術文献】

【特許文献】

【0005】

【特許文献 1】特開 2 0 0 8 - 3 8 5 7 号公報

10

20

30

40

50

【特許文献2】特開2005-165781号公報

【発明の概要】

【発明が解決しようとする課題】

【0006】

ところで、ある記憶領域にボリュームを設定する際に、そのボリューム上のアドレスを設定先の記憶領域上のアドレスと同一にする方法がある。この方法では、ボリュームのデータにアクセスする際に、そのデータのボリューム上のアドレスを、ボリュームが設定された記憶領域のアドレスに変換しなくてよい。このため、例えば、ボリュームのアドレスとして仮想アドレスを用いる場合と比較して、アクセスのための処理負荷が小さいという利点がある。

10

【0007】

その一方で、この方法を用いた場合には次のような問題がある。例えば、記憶領域の先頭から順に第1のボリューム、第2のボリュームが設定された状態から、第1のボリュームが削除されて、その分の空き領域が形成されたとする。この状態で、形成された空き領域に新たなボリュームを設定しようとする場合、削除された第1のボリュームの容量以下のボリュームしか新たに設定することはできず、削除された第1のボリュームの容量を超える容量のボリュームを新たに設定することはできない。このような制約により、記憶領域の利用効率が高いとはいえないという問題がある。

【0008】

1つの側面では、本発明は、ボリュームを削除した後の記憶領域の利用効率を向上させることが可能な情報処理システム、管理装置および制御方法を提供することを目的とする。

20

【課題を解決するための手段】

【0009】

1つの案では、複数のストレージ装置と、管理装置とを有する次のような情報処理システムが提供される。複数のストレージ装置は、それぞれ1以上の記憶領域を有する。また、複数のストレージ装置の記憶領域のそれぞれに対して1以上のボリュームを設定可能であり、設定されたボリュームは、複数のストレージ装置の記憶領域のうち当該ボリュームの設定先の記憶領域における当該ボリュームの設定領域と同じアドレス空間を有する。管理装置は、複数のストレージ装置の記憶領域のうち、第1のストレージ装置の第1の記憶領域に設定された第1のボリュームの削除が要求されたとき、複数のストレージ装置の記憶領域のそれぞれにおけるボリュームの設定状態を管理する管理情報に基づいて、第1の記憶領域に他のボリュームが設定されているかを判定し、他のボリュームとして第2のボリュームが設定されていた場合、管理情報に基づいて、複数のストレージ装置の記憶領域の中から、第2のボリュームと同じ大きさ以上の連続した空き領域を検出し、第1のボリュームを第1の記憶領域から削除する削除処理と、第2のボリュームのデータを、検出された空き領域に退避させた後、第1の記憶領域における第1のボリュームが削除された領域を含む移動先領域に書き戻すことで、第2のボリュームを移動させる移動処理とを、第1のストレージ装置に実行させ、第1のボリュームが削除され、第2のボリュームが移動したことを示すように管理情報を更新する。

30

40

【0010】

また、1つの案では、上記情報処理システムの管理装置と同様の処理を実行する管理装置が提供される。

さらに、1つの案では、上記情報処理システムの管理装置と同様の処理をコンピュータが実行する制御方法が提供される。

【発明の効果】

【0011】

1つの側面では、ボリュームを削除した後の記憶領域の利用効率を向上させることができる。

【図面の簡単な説明】

50

【 0 0 1 2 】

【 図 1 】 第 1 の実施の形態に係る情報処理システムの構成例および処理例を示す図である。

【 図 2 】 第 2 の実施の形態に係るクラウド基盤システムの構成例を示す図である。

【 図 3 】 管理サーバのハードウェア構成例を示す図である。

【 図 4 】 サーバ装置およびストレージ装置の処理機能の構成例を示す図である。

【 図 5 】 仮想ディスクとボリュームとの関係を説明するための図である。

【 図 6 】 ボリュームの削除時の処理例を示す図である。

【 図 7 】 データベースサーバおよび管理サーバの処理機能の構成例を示す図である。

【 図 8 】 ユーザテーブルの構成例を示す図である。

10

【 図 9 】 仮想サーバテーブルの構成例を示す図である。

【 図 1 0 】 仮想ディスクテーブルの構成例を示す図である。

【 図 1 1 】 ボリュームテーブルの構成例を示す図である。

【 図 1 2 】 ストレージ領域管理テーブルの構成例を示す図である。

【 図 1 3 】 予約データ管理テーブルの構成例を示す図である。

【 図 1 4 】 収集データ管理テーブルの構成例を示す図である。

【 図 1 5 】 ボリューム削除および再配置の処理例を示す図（その 1 ）である。

【 図 1 6 】 ボリューム削除および再配置の処理例を示す図（その 2 ）である。

【 図 1 7 】 ボリューム削除および再配置の処理例を示す図（その 3 ）である。

【 図 1 8 】 ボリューム操作指示の受け付け処理の手順を示すフローチャートの例である。

20

【 図 1 9 】 ボリューム削除処理の手順を示すフローチャートの例（その 1 ）である。

【 図 2 0 】 ボリューム削除処理の手順を示すフローチャートの例（その 2 ）である。

【 図 2 1 】 ボリューム削除処理の手順を示すフローチャートの例（その 3 ）である。

【 図 2 2 】 ボリューム削除処理の手順を示すフローチャートの例（その 4 ）である。

【 図 2 3 】 ボリューム削除処理の手順を示すフローチャートの例（その 5 ）である。

【 発明を実施するための形態 】

【 0 0 1 3 】

以下、本発明の実施の形態について図面を参照して説明する。

〔 第 1 の実施の形態 〕

図 1 は、第 1 の実施の形態に係る情報処理システムの構成例および処理例を示す図である。図 1 に示す情報処理システムは、ストレージ装置 1 , 2 と管理装置 3 を含む。なお、情報処理システムに含まれるストレージ装置は、図 1 のようにストレージ装置 1 , 2 の 2 台に限定されるものではなく、3 台以上であってもよい。

30

【 0 0 1 4 】

ストレージ装置 1 , 2 は、それぞれ 1 以上の記憶領域を有する。図 1 の例では、ストレージ装置 1 は記憶領域 1 a , 1 b を有し、ストレージ装置 2 は記憶領域 2 a を有しているものとする。記憶領域 1 a , 1 b , 2 a は、それぞれ個別の記憶装置によって実現される物理記憶領域であってもよいし、1 以上の記憶装置によって実現される論理記憶領域であってもよい。

【 0 0 1 5 】

40

また、記憶領域 1 a , 1 b , 2 a のそれぞれに対しては、1 以上のボリュームを設定可能になっている。図 1 の例では、記憶領域 1 a には、その先頭アドレス側から順にボリューム VOL 1 , VOL 2 , VOL 3 が設定されている。記憶領域 1 b には、その先頭アドレス側から順にボリューム VOL 4 , VOL 5 が設定されている。記憶領域 2 a には、その先頭アドレス側にボリューム VOL 6 が設定されている。なお、図 1 において「 F r e e 」と記載された領域は、ボリュームが設定されていない空き領域を示す。

【 0 0 1 6 】

記憶領域 1 a , 1 b , 2 a のそれぞれに設定されたボリュームは、その設定先の記憶領域におけるそのボリュームの設定領域と同じアドレス空間を有する。例えば、記憶領域 1 a のアドレス「 0 」からアドレス「 A 1 」までの領域にボリューム VOL 1 が設定されて

50

いる場合、ボリュームVOL1の先頭アドレスは「0」となり、末尾のアドレスは「A1」となる。また、記憶領域1aのアドレス「A1+1」からアドレス「A2」までの領域にボリュームVOL2が設定されている場合、ボリュームVOL2の先頭アドレスは「A1+1」となり、末尾のアドレスは「A2」となる。

【0017】

このようなボリュームに対するアドレスの割り当て方法が用いられることで、ボリュームにアクセスする際に、ボリュームに固有の論理アドレスをその設定先の記憶領域のアドレスに変換する必要がなくなる。このため、データアクセスのための処理負荷が軽減され、そのアクセス速度を向上させることができる。

【0018】

ただし、上記のようなアドレスの割り当て方法が用いられることで、1つのボリュームは、記憶領域1a, 1b, 2aにおける連続する空き領域に設定されなければならない。例えば、1つのボリュームを分割し、それらを離散した複数の空き領域に設定することはできない。また、離散した複数の空き領域に設定されたボリュームを1つのボリュームとして使用することはできない。

【0019】

管理装置3は、通信部3aと制御部3bを有する。

通信部3aは、例えば、通信インタフェース回路として実現される。通信部3aは、ストレージ装置1, 2との間で通信を行う。

【0020】

制御部3bは、例えば、プロセッサとして実現される。制御部3bは、通信部3aを介してストレージ装置1, 2と通信可能になっている。また、制御部3bは、管理情報3cを参照可能になっている。管理情報3cは、記憶領域1a, 1b, 2aにおけるボリュームの設定状態を管理するための情報である。管理情報3cは、例えば、管理装置3の内部に搭載された記憶装置に記憶されていてもよいし、管理情報3cとは別の装置に記憶されていてもよい。

【0021】

制御部3bは、設定されたボリュームの削除が要求されたとき、次のような処理を実行する。ここでは例として、記憶領域1aに設定されたボリュームVOL2の削除が要求されたとする。

【0022】

制御部3bは、管理情報3cに基づいて、記憶領域1aに他のボリュームが設定されているかを判定する。他のボリュームが設定されている場合、制御部3bは、管理情報3cに基づいて、記憶領域1a, 1b, 2aの中から、設定されている他のボリュームと同じ大きさ以上の連続した空き領域を検出する。

【0023】

上記で判定される他のボリュームとは、削除が要求されたボリュームVOL2の前側または後側に存在するボリュームである。例えば、記憶領域1aに対して順にボリュームが設定される際に、記憶領域1aの先頭アドレス側から順にボリュームが設定されるとする。この場合、制御部3bは上記判定において、記憶領域1aにおけるボリュームVOL2の下位アドレス側に他のボリュームが設定されているかを判定する。図1の例では、ボリュームVOL2の下位アドレス側にボリュームVOL3が設定されていると判定される。この場合、制御部3bは、管理情報3cに基づいて、記憶領域1a, 1b, 2aの中から、ボリュームVOL3と同じ大きさ以上の連続した空き領域を検出する。図1の例では、記憶領域2aから、条件を満たす空き領域2a1が検出されたとする。

【0024】

制御部3bは次に、ボリュームVOL2を記憶領域1aから削除する削除処理を、ストレージ装置1に実行させる(ステップS1)。これとともに、制御部3bは、ボリュームVOL3を移動させる移動処理を、ストレージ装置1に実行させる。移動処理では、次のような処理が実行される。

10

20

30

40

50

【 0 0 2 5 】

ストレージ装置 1 は、ボリューム V O L 3 のデータを検出された空き領域 2 a 1 に退避させる（ステップ S 2）。図 1 では、この退避によって、ボリューム V O L 3 と同一内容のボリューム V O L 3 a が記憶領域 2 a に設定されている。そして、ストレージ装置 1 は、退避されたボリューム V O L 3 a のデータを、記憶領域 1 a におけるボリューム V O L 2 が削除された領域を含む移動先領域に書き戻す（ステップ S 3）。

【 0 0 2 6 】

例えば、上記のように記憶領域 1 a の先頭アドレス側から順にボリュームが設定される場合、ストレージ装置 1 は、記憶領域 1 a におけるボリューム V O L 2 が削除された領域の先頭を起点とした移動先領域に、ボリューム V O L 3 a のデータを書き戻す。図 1 では、この書き戻しによって、ボリューム V O L 3 a と同一内容のボリューム V O L 3 b が記憶領域 1 a に設定されている。

10

【 0 0 2 7 】

このような移動処理により、ボリューム V O L 3 は、記憶領域 1 a においてボリューム V O L 2 が削除されたことで形成される空き領域の上位アドレス側に詰めて再配置される。その結果、記憶領域 1 a には、ボリューム V O L 2 より大きな連続する空き領域 1 a 1 が形成される。

【 0 0 2 8 】

この後、制御部 3 b は、記憶領域 1 a からボリューム V O L 2 が削除され、記憶領域 1 a においてボリューム V O L 3 が移動したことを示すように、管理情報 3 c を更新する。

20

ここで、例えば、単にボリューム V O L 3 が削除されて、ボリューム V O L 3 と同じ大きさの空き領域が形成された場合、この空き領域は、ボリューム V O L 3 の大きさ以下のボリュームを設定するためにしか利用できなくなる。このため、記憶領域 1 a の利用効率が悪化してしまう。

【 0 0 2 9 】

これに対して、上記の移動処理によって記憶領域 1 a にボリューム V O L 2 より大きな連続する空き領域 1 a 1 が形成されることで、この空き領域 1 a 1 に対してボリューム V O L 2 より大きなボリュームを新たに設定できるようになる。その結果、ボリューム V O L 2 を削除した後の記憶領域 1 a の利用効率を向上させることができる。

【 0 0 3 0 】

30

また、ボリュームの削除に伴う他のボリュームの移動の要否判定や、他のボリュームの退避先領域の検出は、ストレージ装置ではなく管理装置 3 の制御部 3 b によって、管理情報 3 c を参照しながら一元的に実行される。さらに、ストレージ装置における上記の削除処理および移動処理は、管理装置 3 の制御部 3 b の制御の下で実行される。これにより、ボリュームの移動処理を正確に実行できる。また、ストレージ装置の処理負荷を軽減でき、ストレージ装置に設定されたボリュームに対するアクセス速度の低下を抑制できる。

【 0 0 3 1 】

〔 第 2 の実施の形態 〕

次に、図 1 に示した情報処理システムの例として、クラウド基盤システムについて説明する。このクラウド基盤システムは、ユーザに対してクラウド基盤サービス（Infrastructure as a Service：I a a S）を提供する。

40

【 0 0 3 2 】

図 2 は、第 2 の実施の形態に係るクラウド基盤システムの構成例を示す図である。図 2 に示すクラウド基盤システムは、サーバ装置 1 1 0、1 2 0、ストレージ装置 2 1 0、2 2 0、データベースサーバ 3 1 0、管理サーバ 3 2 0 および操作端末 3 3 0 を有する。サーバ装置 1 1 0、1 2 0、ストレージ装置 2 1 0、2 2 0、データベースサーバ 3 1 0 および管理サーバ 3 2 0 は、ネットワーク 4 0 0 を介して相互に接続されている。また、管理サーバ 3 2 0 には、管理者によって操作される操作端末 3 3 0 が接続されている。

【 0 0 3 3 】

なお、ストレージ装置 2 1 0、2 2 0 は、図 1 に示したストレージ装置 1、2 の一例で

50

あり、管理サーバ320は、図1に示した管理装置3の一例である。

サーバ装置110, 120のそれぞれでは、1以上の仮想マシンが動作する。これらの仮想マシンは、ユーザによって例えば仮想サーバとして利用される。ユーザは、例えば、仮想マシン上でアプリケーションプログラムを実行させることができる。

【0034】

ストレージ装置210, 220は、サーバ装置110, 120上で動作する仮想マシンに接続される仮想ディスクの実記憶領域を実現する。ストレージ装置210は、コントローラ211とストレージ部212を備える。ストレージ装置220は、コントローラ221とストレージ部222を備える。

【0035】

ストレージ部212, 222のそれぞれには、HDDやSSD (Solid State Drive) など、仮想ディスクの実記憶領域を実現する複数台の記憶装置が搭載されている。コントローラ211は、サーバ装置110, 120からの要求に応じて、ストレージ部212内の記憶装置にアクセスする。コントローラ221は、サーバ装置110, 120からの要求に応じて、ストレージ部222内の記憶装置にアクセスする。

【0036】

コントローラ211, 221は、それぞれストレージ部212, 222に搭載された記憶装置に対するデータの書き込みを、RAIDを用いて制御する。ストレージ部212, 222のそれぞれには、1以上のRAIDグループが設定される。RAIDグループとは、RAIDによって書き込みが制御される複数の記憶装置を用いて実現される論理記憶領域である。そして、RAIDグループに対して、各仮想ディスクに対応するボリューム (論理記憶領域) が割り当てられる。

【0037】

データベースサーバ310は、クラウド基盤システムに含まれる各種リソースの使用状況を管理するための情報を保持するサーバ装置である。例えば、データベースサーバ310は、クラウド基盤サービスを利用するユーザの情報、ユーザが利用する仮想マシンとサーバ装置110, 120との対応関係を示す情報、仮想ディスクとボリュームおよびRAIDグループとの対応関係を示す情報などを保持する。

【0038】

管理サーバ320は、データベースサーバ310に保持された情報を参照しながら、クラウド基盤システムを管理するサーバ装置である。例えば、管理サーバ320は、操作端末330を用いた管理者による入力操作に応じて、仮想マシン、仮想ディスク、RAIDグループ、ボリュームなどの設定処理を行う。また、管理サーバ320は、これらの設定に変更が生じた場合、その変更内容をデータベースサーバ310が保持する情報に反映させる。

【0039】

図3は、管理サーバのハードウェア構成例を示す図である。管理サーバ320は、例えば、図3に示すようなコンピュータとして実現される。

図3に示す管理サーバ320は、プロセッサ320a、RAM (Random Access Memory) 320b、HDD 320c、グラフィック処理装置320d、入力インタフェース320e、読み取り装置320fおよび通信インタフェース320gを有する。なお、プロセッサ320aは、図1に示した制御部3bの一例であり、通信インタフェース320gは、図1に示した通信部3aの一例である。

【0040】

プロセッサ320aは、管理サーバ320全体を統括的に制御する。プロセッサ320aは、例えば、CPU (Central Processing Unit)、MPU (Micro Processing Unit)、DSP (Digital Signal Processor)、ASIC (Application Specific Integrated Circuit) またはPLD (Programmable Logic Device) である。また、プロセッサ320aは、CPU、MPU、DSP、ASIC、PLDのうちの2以上の要素の組み合わせであってもよい。

10

20

30

40

50

【 0 0 4 1 】

R A M 3 2 0 b は、管理サーバ 3 2 0 の主記憶装置として使用される。R A M 3 2 0 b には、プロセッサ 3 2 0 a に実行させる O S (Operating System) プログラムやアプリケーションプログラムの少なくとも一部が一時的に格納される。また、R A M 3 2 0 b には、プロセッサ 3 2 0 a による処理に必要な各種データが格納される。

【 0 0 4 2 】

H D D 3 2 0 c は、管理サーバ 3 2 0 の補助記憶装置として使用される。H D D 3 2 0 c には、O S プログラム、アプリケーションプログラム、および各種データが格納される。なお、補助記憶装置としては、S S D などの他の種類の不揮発性記憶装置を使用することもできる。

10

【 0 0 4 3 】

グラフィック処理装置 3 2 0 d には、表示装置 3 2 0 d 1 が接続されている。グラフィック処理装置 3 2 0 d は、プロセッサ 3 2 0 a からの命令にしたがって、画像を表示装置 3 2 0 d 1 に表示させる。表示装置としては、液晶ディスプレイや有機 E L (Electroluminescence) ディスプレイなどがある。

【 0 0 4 4 】

入力インタフェース 3 2 0 e には、入力装置 3 2 0 e 1 が接続されている。入力インタフェース 3 2 0 e は、入力装置 3 2 0 e 1 から出力される信号をプロセッサ 3 2 0 a に送信する。入力装置 3 2 0 e 1 としては、キーボードやポインティングデバイスなどがある。ポインティングデバイスとしては、マウス、タッチパネル、タブレット、タッチパッド、トラックボールなどがある。

20

【 0 0 4 5 】

読み取り装置 3 2 0 f には、可搬型記録媒体 3 2 0 f 1 が脱着される。読み取り装置 3 2 0 f は、可搬型記録媒体 3 2 0 f 1 に記録されたデータを読み取ってプロセッサ 3 2 0 a に送信する。可搬型記録媒体 3 2 0 f 1 としては、光ディスク、光磁気ディスク、半導体メモリなどがある。

【 0 0 4 6 】

通信インタフェース 3 2 0 g は、ネットワーク 4 0 0 を介して、データベースサーバ 3 1 0 などの他の装置との間でデータの送受信を行う。

以上のようなハードウェア構成によって、管理サーバ 3 2 0 の処理機能を実現することができる。なお、データベースサーバ 3 1 0、サーバ装置 1 1 0、1 2 0、コントローラ 2 1 1、2 2 1 も、図 3 に示すようなコンピュータとして実現可能である。

30

【 0 0 4 7 】

図 4 は、サーバ装置およびストレージ装置の処理機能の構成例を示す図である。

サーバ装置 1 1 0 は、ハイパーバイザ 1 1 1 を有する。ハイパーバイザ 1 1 1 の処理は、例えば、サーバ装置 1 1 0 が備える図示しないプロセッサが所定のプログラムを実行することで実現される。ハイパーバイザ 1 1 1 は、サーバ装置 1 1 0 上に仮想マシンを構築し、構築された仮想マシンの動作を管理する。図 4 の例では、サーバ装置 1 1 0 では、ハイパーバイザ 1 1 1 の管理の下で仮想マシン V M 1、V M 2 が動作している。

40

【 0 0 4 8 】

サーバ装置 1 2 0 も同様に、ハイパーバイザ 1 2 1 を有する。ハイパーバイザ 1 2 1 の処理は、例えば、サーバ装置 1 2 0 が備える図示しないプロセッサが所定のプログラムを実行することで実現される。ハイパーバイザ 1 2 1 は、サーバ装置 1 2 0 上に仮想マシンを構築し、構築された仮想マシンの動作を管理する。図 4 の例では、サーバ装置 1 2 0 では、ハイパーバイザ 1 2 1 の管理の下で仮想マシン V M 3、V M 4 が動作している。

【 0 0 4 9 】

ストレージ装置 2 1 0 のコントローラ 2 1 1 は、R A I D 制御部 2 1 1 a を有する。R A I D 制御部 2 1 1 a の処理は、例えば、コントローラ 2 1 1 が備える図示しないプロセッサが所定のプログラムを実行することで実現される。R A I D 制御部 2 1 1 a は、サーバ装置 1 1 0、1 2 0 からの要求に応じて、ストレージ部 2 1 2 内の記憶装置に対するア

50

クセスを、RAIDを用いて制御する。また、RAID制御部211aは、管理サーバ320からの命令にしたがい、RAIDグループやボリュームの設定、後述するボリュームの再配置などの処理を実行する。

【0050】

ストレージ装置220のコントローラ221も同様に、RAID制御部221aを有する。RAID制御部221aの処理は、例えば、コントローラ221が備える図示しないプロセッサが所定のプログラムを実行することで実現される。RAID制御部221aは、サーバ装置110, 120からの要求に応じて、ストレージ部222内の記憶装置に対するアクセスを、RAIDを用いて制御する。また、RAID制御部221aは、管理サーバ320からの命令にしたがい、RAIDグループやボリュームの設定、後述するボリュームの再配置などの処理を実行する。

10

【0051】

なお、図4の例では、ストレージ装置210には、ストレージ部212内の記憶装置を用いてRAIDグループRG1, RG2, RG3が設定されている。また、ストレージ装置220には、ストレージ部222内の記憶装置を用いてRAIDグループRG4, RG5, RG6が設定されている。

【0052】

図5は、仮想ディスクとボリュームとの関係を説明するための図である。

前述のように、ストレージ装置210, 220に設定されたRAIDグループには、ボリュームが割り当てられる。そして、ボリュームは、サーバ装置110, 120で動作する仮想マシンに接続(マウント)された仮想ディスクに対応する記憶領域として利用される。

20

【0053】

図5の例では、サーバ装置110において仮想マシンVM1, VM2が動作し、仮想マシンVM1には仮想ディスクVD1, VD2が接続され、仮想マシンVM2には仮想ディスクVD3が接続されている。また、サーバ装置120において仮想マシンVM3, VM4が動作し、仮想マシンVM3には仮想ディスクVD4, VD5が接続され、仮想マシンVM4には仮想ディスクVD6が接続されている。

【0054】

一方、ストレージ装置210には、ストレージ部212内の記憶装置を用いてRAIDグループRG1, RG2, RG3が設定されている。また、ストレージ装置220には、ストレージ部222内の記憶装置を用いてRAIDグループRG4, RG5, RG6が設定されている。

30

【0055】

RAIDグループRG1には、ボリュームVOL11~VOL14が割り当てられている。RAIDグループRG2には、ボリュームVOL21, VOL22が割り当てられている。RAIDグループRG3には、ボリュームVOL31, VOL32が割り当てられている。RAIDグループRG4には、ボリュームVOL41, VOL42が割り当てられている。RAIDグループRG5には、ボリュームVOL51, VOL52が割り当てられている。RAIDグループRG6には、ボリュームVOL61が割り当てられている。

40

【0056】

そして、ボリュームVOL11が仮想ディスクVD1に対応付けられ、ボリュームVOL12が仮想ディスクVD3に対応付けられ、ボリュームVOL13が仮想ディスクVD2に対応付けられ、ボリュームVOL14が仮想ディスクVD4に対応付けられている。さらに、ボリュームVOL41が仮想ディスクVD5に対応付けられ、ボリュームVOL42が仮想ディスクVD6に対応付けられている。

【0057】

なお、図5に記載された「Free」とは、RAIDグループの論理記憶領域のうち、ボリュームが割り当てられていない空き領域を示す。

50

また、RAIDグループに対してボリュームが割り当てられる際には、RAIDグループの論理記憶領域の先頭側から順に、ボリュームが詰めて割り当てられる。したがって、通常、ボリュームが1つ以上割り当てられたRAIDグループでは、論理記憶領域の末端側には空き領域が存在し得るが、先頭側には空き領域が存在しない。

【0058】

仮想ディスクと、その仮想ディスクに対応付けられたボリュームとの間では、次のような制御が行われる。例えば、仮想マシンVM1により仮想ディスクVD1に対するデータの書き込みが行われると、サーバ装置110からストレージ装置210のRAID制御部211aに対して書き込み要求が送信される。この書き込み要求では、仮想マシンVM1上の書き込みアドレスが書き込み先として指定される。RAID制御部211aは、指定された書き込みアドレスをボリュームVOL11上のアドレスに変換して、書き込みが要求されたデータをボリュームVOL1に書き込む。このとき、ストレージ部212に搭載された、RAIDグループRG1に属する複数の記憶装置に対し、所定のRAIDレベルにしたがってデータの書き込み処理が実行される。

10

【0059】

このように、本実施の形態に係るクラウド基盤システムでは、ある仮想マシンに仮想ディスクが新たに接続されると、その仮想ディスクに対応するボリュームがいずれかのRAIDグループに割り当てられる。そして、仮想ディスクに対応するボリュームに対してアクセスが行われることで、仮想ディスクに対するアクセス処理が仮想的に実現される。

【0060】

20

また、クラウド基盤システムでは、管理サーバ320の制御の下で、ユーザの要求に応じた容量やアクセス性能を有する記憶領域を、仮想ディスクに対して割り当てることが可能となっている。

【0061】

例えば、RAIDグループRG1に属する記憶装置より、RAIDグループRG2、RG3に属する記憶装置の方が、アクセス性能が高いものとする。また、仮想マシンVM1を利用するユーザにより、当初は、仮想ディスクVD1の仕様としてアクセス性能が低く、利用料金が低い仕様が要求されたとする。この場合、管理サーバ320は、ユーザの要求にしたがって、仮想ディスクVD1に対応するボリュームVOL11を、例えば図5のようにRAIDグループRG1に割り当てる。

30

【0062】

その後、ユーザから仮想ディスクVD1のアクセス性能を高めたい旨が要求されると、管理サーバ320は例えば、ボリュームVOL11の割り当て先をRAIDグループRG1からRAIDグループRG2に移動させることができる。さらに、ユーザから仮想ディスクVD1の容量を拡張したい旨が要求されると、管理サーバ320は例えば、ボリュームVOL11の割り当て先を、RAIDグループRG2からより大きな空き領域を有するRAIDグループRG3に移動させることもできる。

【0063】

ところで、前述のように、RAIDグループは、RAIDによって書き込みが制御される複数の記憶装置を用いて実現される論理記憶領域である。RAIDグループの論理記憶領域には論理アドレス(Logical Block Address: LBA)が付与されている。そして、RAIDグループ上のある論理アドレスに対してデータの書き込みが要求されると、そのRAIDグループに属する複数の記憶装置に対して所定のRAIDレベルにしたがってデータの書き込みが行われる。

40

【0064】

これに加えて、本実施の形態のストレージ装置210、220では、RAIDグループに設定されたボリュームには、そのボリューム固有の論理アドレス(仮想アドレス)ではなく、RAIDグループと同じ論理アドレスが付与される。したがって、1つのボリュームは、RAIDグループの論理記憶領域のうち、論理アドレスが連続するひとまとめの領域に設定される。そして、設定されたボリュームは、RAIDグループの論理記憶領域の

50

うち、このボリュームが設定された領域と同じアドレス空間を有する。

【0065】

例えば、RAIDグループRG1の論理記憶領域における先頭領域に、ボリュームVOL11が設定されるとする。このとき、RAIDグループRG1の論理記憶領域における論理アドレス「0」から論理アドレス「X」までの領域にボリュームVOL11が設定されたとすると、ボリュームVOL11の先頭の論理アドレスは「0」となり、末尾の論理アドレスは「X」となる。さらに、RAIDグループRG1に対して続けてボリュームVOL12が設定されるとする。このとき、ボリュームVOL12の先頭の論理アドレスは「X+1」となる。なお、このようなボリューム設定方法が用いられるストレージ装置は、「物理ブロックストレージ」と呼ばれる場合もある。

10

【0066】

上記のようなボリューム設定方法が用いられることで、RAID制御部211a, 221aは、ボリューム内のデータにアクセスする際に、ボリュームに固有の論理アドレスをRAIDグループの論理記憶領域の論理アドレスに変換する必要がなくなる。このため、データアクセスのための処理負荷が軽減され、そのアクセス速度を向上させることができる。

【0067】

その一方で、上記のようなボリューム設定方法を用いてボリュームの削除が行われる場合には、次の図6に示すような課題がある。

図6は、ボリュームの削除時の処理例を示す図である。

20

【0068】

図6の状態1では、RAIDグループRG1の論理記憶領域の先頭から順に、ボリュームVOL11, VOL12, VOL13, VOL14が設定されているとする。さらに、ボリュームVOL14より末端側(下位アドレス側)には空き領域FR1が形成されている。

【0069】

そして、この状態1から、ボリュームVOL12が削除されるものとする。この場合、図6の状態2に示すように、ボリュームVOL11とボリュームVOL13との間に空き領域FR2が形成される。なお、空き領域FR2は、空き領域FR1より大きいものとする。

30

【0070】

この状態2では、RAIDグループRG1の論理記憶領域には、空き領域FR1と空き領域FR2とが存在している。状態2では、例えば、空き領域FR2の容量以下の新規のボリュームを、ボリュームVOL11とボリュームVOL13との間に追加設定できる。しかしながら、ボリュームの記憶領域が仮想化されない物理ブロックストレージでは、状態2では、RAIDグループRG1に対して、空き領域FR2より大きい容量のボリュームを追加設定できない。例えば、この後に空き領域FR2より大きい容量のボリュームの追加が要求され続けると、空き領域FR2(および空き領域FR1)は利用されないままになってしまう。このように、物理ブロックストレージでは、ボリュームが削除されると、記憶領域の利用効率が低下する可能性がある、という課題がある。

40

【0071】

そこで、本実施の形態では、管理サーバ320の制御の下で、あるボリュームより先頭側に空き領域が生じた場合に、その空き領域が埋まるようにボリュームを先頭側(上位アドレス側)に詰める「再配置」が実行される。図6の例では、ボリュームVOL12が削除されると、状態3に示すように、ボリュームVOL13, VOL14がボリュームVOL11の隣接領域に再配置される。これにより、利用効率の低い空き領域FR2がなくなるとともに、末端側の空き領域FR1が拡張されて、連続するひとまとめのより大きな空き領域FR1が形成されて、状態2より大きなボリュームを追加設定できるようになる。

【0072】

ところで、実際の再配置では、ボリュームVOL13, VOL14のデータは、他のR

50

RAIDグループに対して一旦退避された後、RAIDグループRG1に書き戻される。この一連の処理には、例えば、退避先の他のRAIDグループの探索、探索された他のRAIDグループへのボリュームVOL13, VOL14のデータコピー、元のボリュームVOL13, VOL14の削除、退避先からRAIDグループRG1へのボリュームVOL13, VOL14のデータコピー、退避先のボリュームVOL13, VOL14の削除といった処理が含まれる。

【0073】

また、サーバ装置110, 120からそれぞれボリュームVOL13, VOL14へのアクセスをできる限り継続したままで（活性状態で）、これら一連の処理が実行される。このため、一連の処理には、退避先のボリュームVOL13, VOL14へのパスの変更や、RAIDグループRG1上の再配置後のボリュームVOL13, VOL14へのパスの変更も含まれる。

10

【0074】

このように、ボリューム再配置のためには複雑な処理が必要となる。このため、このような一連の処理をストレージ装置210のコントローラ211だけに実行させた場合、コントローラ211の処理負荷が大きく、サーバ装置110, 120からストレージ装置210内のボリュームへのアクセス性能が低下する可能性がある。

【0075】

また、このクラウド基盤システムでは、複数のサーバ装置や複数のストレージ装置のリソースが複数のユーザによって共用されるため、リソースの使用状況が複雑である。さらに、予測できない突発的な使用状況の変更が生じ得るなど、リソースの使用状況が流動的である。このため、ストレージ装置210のコントローラ211だけに上記の一連の処理を実行させた場合、その実行中に、退避先のRAIDグループにおける容量不足、通信断、データ破壊などの様々な障害が発生する可能性がある。特に、活性状態のままボリュームを再配置する処理の難度は非常に高いと言える。

20

【0076】

そこで、本実施の形態では、クラウド基盤システムの使用状況を一元的に管理する管理サーバ320により、上記の一連の処理を実行するための命令を生成し、それらの命令をストレージ装置210に送信する。ストレージ装置210のコントローラ211は、単に受信した命令にしたがって処理を実行する。これにより、コントローラ211の処理負荷を軽減できるとともに、ボリューム再配置の実行中における障害発生率を低減できる。

30

【0077】

図7は、データベースサーバおよび管理サーバの処理機能の構成例を示す図である。データベースサーバ310は、記憶部311と入出力部312を有する。

記憶部311は、例えば、データベースサーバ310が備える図示しない記憶装置の記憶領域によって実現される。記憶部311には、ユーザテーブル311a、仮想サーバテーブル311b、仮想ディスクテーブル311c、ボリュームテーブル311dおよびストレージ領域管理テーブル311eが記憶される。以下、図8～図12を用いてこれらのテーブル情報について説明する。

40

【0078】

図8は、ユーザテーブルの構成例を示す図である。ユーザテーブル311aは、ユーザの情報を管理するための管理テーブルである。ユーザテーブル311aには、会社名など、ユーザの名称を示すユーザ名と、ユーザを識別するユーザID (Identification) とが対応付けて登録されている。

【0079】

図9は、仮想サーバテーブルの構成例を示す図である。仮想サーバテーブル311bは、ユーザが使用する仮想サーバ（仮想マシン）に関する情報を管理するための管理テーブルである。仮想サーバテーブル311bは、ユーザ名、ホスト名、仮想サーバ名、物理サーバ名、クラスタグループ、タイプ、マシンタイプ、OSタイプの各項目を有する。

50

【 0 0 8 0 】

ユーザ名は、仮想サーバを使用するユーザの名称を示す。ホスト名は、仮想サーバに対応するホストの名称を示す。仮想サーバ名は、仮想サーバ（仮想マシン）を識別するための名称を示す。物理サーバ名は、仮想サーバが動作するサーバ装置（サーバ装置 1 1 0 , 1 2 0）を識別するための名称を示す。クラスタグループは、仮想サーバが属するクラスタグループを識別するための情報を示す。

【 0 0 8 1 】

タイプは、仮想サーバのタイプを示す。例えば、仮想サーバが単独で運用される場合、タイプとして「1」が登録される。また、仮想サーバが他の仮想サーバと冗長化されて運用される場合（例えば、一方が他方の予備機として運用される場合）、タイプとして「冗長」が登録される。マシンタイプは、仮想マシンの性能や構成などの仕様を識別するための情報を示す。OSタイプは、仮想マシンで実行されるOSプログラムを識別するための情報を示す。

10

【 0 0 8 2 】

図 1 0 は、仮想ディスクテーブルの構成例を示す図である。仮想ディスクテーブル 3 1 1 c は、仮想サーバ（仮想マシン）に接続された仮想ディスクに関する情報を管理するための管理テーブルである。仮想ディスクテーブル 3 1 1 c は、仮想サーバ名、LUN（Logical Unit Number）番号、仮想ディスクID、ボリューム名、ディスクタイプの各項目を有する。

【 0 0 8 3 】

20

仮想サーバ名は、仮想ディスクが接続された仮想サーバを識別するための名称を示す。LUN番号は、仮想ディスクに対応するボリュームを識別するための情報を示す。仮想ディスクIDは、仮想ディスクを識別するための情報を示す。ボリューム名は、仮想ディスクに対応するボリュームを識別するための情報を示す。例えば、LUN番号がユーザ側でボリュームを識別するために使用されるのに対し、ボリューム名はクラウド基盤システム側でボリュームを識別するために使用される。ディスクタイプは、仮想ディスクの仕様に関する情報を示す。例えば、仮想ディスクが単独の仮想サーバに接続される場合、ディスクタイプとして「シングル」が登録され、仮想ディスクが複数の仮想サーバに接続されてこれらの中で共有される場合、ディスクタイプとして「共有」が登録される。

【 0 0 8 4 】

30

図 1 1 は、ボリュームテーブルの構成例を示す図である。ボリュームテーブル 3 1 1 d は、RAIDグループに設定されたボリュームに関する情報を管理するための管理テーブルである。ボリュームテーブル 3 1 1 d は、ストレージ装置名、RAIDグループ名、ボリューム番号、ボリューム名、サイズの各項目を有する。

【 0 0 8 5 】

ストレージ装置名は、ボリュームが設定されるストレージ装置（ストレージ装置 2 1 0 , 2 2 0）を識別するための名称を示す。RAIDグループ名は、ボリュームが設定されるRAIDグループを識別するための名称を示す。ボリューム番号は、ボリュームの識別番号を示す。ボリューム名は、ボリュームを識別するための名称を示す。サイズは、ボリュームの容量を示す。

40

【 0 0 8 6 】

図 1 2 は、ストレージ領域管理テーブルの構成例を示す図である。ストレージ領域管理テーブル 3 1 1 e は、ストレージ装置 2 1 0 , 2 2 0 のストレージ部 2 1 2 , 2 2 2 の記憶領域における使用状況を管理するための管理テーブルである。ストレージ領域管理テーブル 3 1 1 e は、ストレージ装置名、RAIDグループ名、RAIDサイズ、ボリューム名、サイズ、ステータスの各項目を有する。

【 0 0 8 7 】

ストレージ装置名は、RAIDグループが設定されたストレージ装置（ストレージ装置 2 1 0 , 2 2 0）を識別するための名称を示す。RAIDグループ名は、設定されたRAIDグループを識別するための名称を示す。RAIDサイズは、RAIDグループの論理

50

記憶領域の総容量を示す。ボリューム名は、RAIDグループに設定されたボリュームを識別するための名称を示す。サイズは、RAIDグループに設定されたボリュームの容量を示す。

【0088】

ステータスは、RAIDグループの論理記憶領域における該当領域の状態を示す。例えば、該当領域がボリュームとして使用中の場合、ステータスとして「Used」が登録される。また、該当領域がボリュームとして使用するよう予約されている場合、ステータスとして「Reserved」が登録される。ステータスが「Reserved」の記憶領域は、未使用状態であっても空き領域ではないと認識されるため、ボリュームの退避先などとして誤って使用されなくなる。なお、ステータスが「Free」の場合、ボリューム名とサイズの各項目には、情報が登録されない。

10

【0089】

以上のストレージ領域管理テーブル311eにより、RAIDグループ上にどのようなボリュームが設定され、どれだけ空き領域があるかが管理される。なお、図12の例では、RAIDグループ「RG000」に対してその上位アドレス側から順に、300ギガバイトのボリューム「vd-1101」と、120ギガバイトのボリューム「vd-1102」とが設定されていることが示されている。また、RAIDグループ「RG000」におけるそれより下位アドレス側の領域が空き領域であることが示されている。

【0090】

以下、図7に戻って説明を続ける。

20

入出力部312の処理は、例えば、データベースサーバ310が備える図示しないプロセッサが所定のプログラムを実行することで実現される。入出力部312は、管理サーバ320からの要求に応じて、記憶部311内のテーブル情報に対するデータの登録や、テーブル情報からのデータの読み出しを行う。

【0091】

次に、管理サーバ320は、記憶部321、管理部322および制御部323を有する。

記憶部321は、例えば、RAM320b、HDD320cなど、管理サーバ320が備える記憶装置の記憶領域によって実現される。管理部322および制御部323の処理は、例えば、プロセッサ320aが所定のプログラムを実行することで実現される。

30

【0092】

記憶部321には、予約データ管理テーブル321aと収集データ管理テーブル321bが記憶される。

ここで、図13は、予約データ管理テーブルの構成例を示す図である。予約データ管理テーブル321aは、ボリューム操作予約に関する情報を管理するための管理テーブルである。

【0093】

ボリューム操作予約は、仮想ディスクの追加や削除、あるいは仮想ディスクの仕様変更など、ボリュームの操作が必要となる処理がユーザから要求された場合に発生する。ボリューム操作予約が発生すると、それに対応する予約データが生成される。なお、仮想ディスクの仕様変更の例としては、仮想ディスクのアクセス速度の変更があり、その場合、対応するボリュームがよりアクセス速度の高い記憶装置によって実現されるRAIDグループに移動される。

40

【0094】

予約データ管理テーブル321aの1つのレコードが、1つのボリューム操作予約に関する予約データを示す。予約データは、ユーザ情報、仮想サーバ名、物理サーバ名、操作種別、ボリューム番号を含む。

【0095】

ユーザ情報は、ユーザ名とユーザIDを含む。ユーザ名は、ボリューム操作を予約したユーザの名称を示し、ユーザIDは、そのユーザを識別するための識別番号を示す。仮想

50

サーバ名は、操作対象のボリュームに対応付けられた仮想ディスクに接続する仮想サーバ（仮想マシン）を識別するための名称を示す。物理サーバ名は、その仮想サーバが動作するサーバ装置（サーバ装置 110, 120）を識別するための名称を示す。操作種別は、操作内容の種別を示す。操作種別としては、例えば、ボリュームの削除、ボリュームの追加、ボリュームの移動などがある。なお、図示しないが、例えば仮想ディスクの仕様変更の要求に応じてボリュームの移動を行う場合、変更後の仕様を示す情報が予約データに登録される。ボリューム番号は、操作対象のボリュームを識別するための識別番号を示す。

【0096】

図7に戻り、管理部322は、仮想ディスクの追加や削除、仮想ディスクの仕様変更の要求を操作端末330から受け付ける。管理部322は、このようなボリュームの操作が必要となる処理が要求されると、上記の予約データを生成して、ボリューム操作予約を行う。管理部322は、ボリューム操作予約が発生するたびに、そのボリューム操作予約に対応する予約データを予約データ管理テーブル321aに登録する。

10

【0097】

また、管理部322は、予約データ管理テーブル321aから予約データを登録順に読み出し、読み出した予約データに基づいて、操作対象のボリュームに関する情報をデータベースサーバ310から収集する。管理部322は、データベースサーバ310から収集したデータを収集データとして結合し、収集データ管理テーブル321bに登録する。

【0098】

ここで、図14は、収集データ管理テーブルの構成例を示す図である。収集データ管理テーブル321bは、収集データを一時的に保持するための管理テーブルである。収集データ管理テーブル321bの1つのレコードが、1つのボリューム操作予約に関する収集データを示す。収集データは、予約データに含まれていたユーザ情報、仮想サーバ名、物理サーバ名、操作種別、ボリューム番号に加えて、ストレージ装置名、LUN番号、ボリューム情報を含んでいる。なお、予約データ中のボリューム番号は、ボリューム情報内に格納される。

20

【0099】

ストレージ装置名は、操作対象のボリュームが設定されているストレージ装置（ストレージ装置210, 220）を識別するための名称を示す。LUN番号は、操作対象のボリュームを識別するための番号を示す。RAIDグループ名は、操作対象のボリュームが設定されているRAIDグループを識別するための名称を示す。ボリューム名は、操作対象のボリュームを識別するための名称を示す。サイズは、操作対象のボリュームの容量を示す。ディスクタイプは、操作対象のボリュームに対応する仮想ディスクの仕様に関する情報を示す。ステータスは、操作対象のボリュームの状態を示す。

30

【0100】

図7に戻り、管理部322は、収集データに対応するボリューム操作を実行するための制御情報を生成する。例えば、ボリュームの再配置を伴うボリューム削除が実行される場合、削除対象および再配置対象の各ボリューム、再配置対象のボリュームの退避先のRAIDグループを示すRAIDグループ名などを含む制御情報が生成される。管理部322は、生成した制御情報を制御部323に受け渡す。

40

【0101】

制御部323は、管理部322から受け渡された制御情報に基づいて、ボリューム操作を実行するための命令を生成し、生成した命令を記述した命令情報を生成する。制御部323は、生成した命令情報を、操作対象のボリュームが設定されたストレージ装置に送信して、命令情報に基づくボリューム操作を実行させる。制御部323は、命令情報に基づくボリューム操作の完了通知をストレージ装置から受信すると、ボリューム操作に応じた管理テーブルの更新を管理部322に依頼する。

【0102】

管理部322は、この依頼に応じて、ボリューム操作に応じたリソース使用状況の変化を、データベースサーバ310に登録された管理テーブルに反映させる。そして、管理部

50

3 2 2 は、完了したボリューム操作に対応する予約データおよび収集データを、それぞれ予約データ管理テーブル 3 2 1 a、収集データ管理テーブル 3 2 1 b から削除する。

【 0 1 0 3 】

次に、図 1 5 ~ 図 1 7 を用いて、ボリュームの再配置を伴うボリューム削除の処理例について説明する。

図 1 5 ~ 図 1 7 は、ボリューム削除および再配置の処理例を示す図である。図 1 5 ~ 図 1 7 では、図 5 に示した状態から、ボリューム VOL 1 2 を削除するボリューム操作が要求された場合の処理について例示する。

【 0 1 0 4 】

管理サーバ 3 2 0 の管理部 3 2 2 は、ボリューム VOL 1 2 を削除するボリューム操作が要求されると、RAID グループ RG 1 の論理記憶領域におけるボリューム VOL 1 2 の下位アドレス側に他のボリュームが設定されているかを判定する。管理部 3 2 2 は、下位アドレス側にボリューム VOL 1 3 , VOL 1 4 が設定されていることを判定すると、これらのボリューム VOL 1 3 , VOL 1 4 の再配置が必要と判定する (ステップ S 1 1)。

【 0 1 0 5 】

また、管理部 3 2 2 は、ストレージ装置 2 1 0 , 2 2 0 に設定された RAID グループ RG 1 ~ RG 6 の中から、ボリューム VOL 1 3 , VOL 1 4 の退避先を決定する (ステップ S 1 2)。この処理では、RAID グループ RG 1 ~ RG 6 の中から、ボリューム VOL 1 3 , VOL 1 4 の合計容量以上の連続する空き領域を有する RAID グループが退避先として決定される。図 1 5 では例として、RAID グループ RG 1 と同一のストレージ装置 2 1 0 に設定された RAID グループ RG 3 が、退避先として決定されたとする。なお、ボリューム VOL 1 3 の退避先とボリューム VOL 1 4 の退避先とが、異なる RAID グループに決定されてもよい。

【 0 1 0 6 】

次に、管理部 3 2 2 は、ボリューム VOL 1 2 の削除およびボリューム VOL 1 3 , VOL 1 4 の再配置を実行するための制御情報を生成し、制御部 3 2 3 に受け渡す。制御部 3 2 3 は、これらの処理の実行のための命令情報を生成し、生成した命令情報をストレージ装置 2 1 0 のコントローラ 2 1 1 に送信して、命令情報に基づく処理の実行を要求する。

【 0 1 0 7 】

コントローラ 2 1 1 の RAID 制御部 2 1 1 a は、受信した命令情報にしたがって以下のような処理を実行する。まず、RAID 制御部 2 1 1 a は、ボリューム VOL 1 2 を削除する (ステップ S 1 3)。この時点では、RAID グループ RG 1 には、ボリューム VOL 1 1 とボリューム VOL 1 3 との間に空き領域 FR 3 が形成される。次に、RAID 制御部 2 1 1 a は、ボリューム VOL 1 3 , VOL 1 4 を退避先の RAID グループ RG 3 の空き領域にコピー (ミラーリング) する (ステップ S 1 4)。

【 0 1 0 8 】

このコピー処理では、仮想マシン VM 3 , VM 1 からそれぞれ RAID グループ RG 1 上のボリューム VOL 1 3 , VOL 1 4 に対するアクセスが許容された状態 (活性状態) のまま、ボリューム VOL 1 3 , VOL 1 4 のコピーが実行される。例えば、RAID 制御部 2 1 1 a は、ボリューム VOL 1 3 , VOL 1 4 にそれぞれ対応するビットマップを生成する。各ビットマップは、ボリューム VOL 1 3 , VOL 1 4 上の論理ブロックごとのビットを有し、初期状態では全ビットの値が「1」に設定される。

【 0 1 0 9 】

ボリューム VOL 1 3 のコピー処理は、例えば、対応するビットマップを用いて次のように実行される。RAID 制御部 2 1 1 a は、ボリューム VOL 1 3 におけるビットの値が「1」の論理ブロックのデータを退避先にコピーし、そのデータのコピーが完了するとビットの値を「0」に更新する。これとともに、RAID 制御部 2 1 1 a は、仮想マシン VM 3 からボリューム VOL 1 3 に対するデータの書き込みが要求されると、データをボ

10

20

30

40

50

リユームVOL13に書き込むとともに、書き込み先の論理ブロックに対応するビットの値を「1」に更新する。RAID制御部211aは、ビットマップの全ビットの値が「0」になるまで、退避先へのデータのコピーと、仮想マシンVM3からの書き込み要求の受け付けとを継続する。ビットマップの全ビットの値が「0」になったとき、コピー処理は終了する。ボリュームVOL14についても、対応するビットマップを用いて同様の処理が実行される。

【0110】

退避先へのコピー処理が終了すると、RAID制御部211aは、仮想マシンVM3からRAIDグループRG1上のボリュームVOL13への接続パスを、RAIDグループRG3上のボリュームVOL13への接続パスに切り替える。また、RAID制御部211aは、仮想マシンVM1からRAIDグループRG1上のボリュームVOL14への接続パスを、RAIDグループRG3上のボリュームVOL14への接続パスに切り替える（ステップS15）。

10

【0111】

パス切り替えが完了すると、RAID制御部211aは、RAIDグループRG3上のボリュームVOL13、VOL14に対するアクセスの受け付けを開始する。これにより、これ以後の仮想マシンVM3、VM1からそれぞれボリュームVOL13、VOL14へのアクセスは、退避先のボリュームVOL13、VOL14に対して行われるようになる。

20

【0112】

ステップS15のパス切り替えにより、仮想ディスクVD4、VD2に対応するアクセス先のボリュームは、RAIDグループRG3上のボリュームVOL13、VOL14にそれぞれ切り替えられている。このため、RAID制御部211aは、仮想ディスクVD4上のアドレスを指定したアクセス要求を受信すると、指定されたアドレスを、RAIDグループRG1上のボリュームVOL13の論理アドレスではなく、RAIDグループRG3上のボリュームVOL13の論理アドレスに変換する。同様に、RAID制御部211aは、仮想ディスクVD2上のアドレスを指定したアクセス要求を受信すると、指定されたアドレスを、RAIDグループRG1上のボリュームVOL14の論理アドレスではなく、RAIDグループRG3上のボリュームVOL14の論理アドレスに変換する。

30

【0113】

また、上記のパス切り替えの完了後、RAID制御部211aは、仮想マシンVM3、VM1からそれぞれRAIDグループRG1上のボリュームVOL13、VOL14への接続パスを閉塞する。そして、RAID制御部211aは、RAIDグループRG1上のボリュームVOL13、VOL14を削除する（ステップS16）。

【0114】

次に、RAID制御部211aは、RAIDグループRG3上のボリュームVOL13、VOL14を、RAIDグループRG1におけるボリュームVOL11に続く位置にコピー（ミラーリング）する（ステップS17）。このコピー処理では、上記のステップS14の処理と同様に、仮想マシンVM3、VM1からそれぞれRAIDグループRG3上のボリュームVOL13、VOL14に対するアクセスが許容された状態（活性状態）のまま、コピーが実行される。

40

【0115】

コピー処理が完了すると、RAID制御部211aは、仮想マシンVM3からRAIDグループRG3上のボリュームVOL13への接続パスを、RAIDグループRG1上のボリュームVOL13への接続パスに切り替える。また、RAID制御部211aは、仮想マシンVM1からRAIDグループRG3上のボリュームVOL14への接続パスを、RAIDグループRG1上のボリュームVOL14への接続パスに切り替える（ステップS18）。

【0116】

パス切り替えが完了すると、RAID制御部211aは、RAIDグループRG1上の

50

ボリュームVOL13, VOL14に対するアクセスの受け付けを開始する。このとき、RAID制御部211aは、仮想ディスクVD4上のアドレスを指定したアクセス要求を受信すると、指定されたアドレスを、RAIDグループRG3上のボリュームVOL13の論理アドレスではなく、RAIDグループRG1上の移動後のボリュームVOL13の論理アドレスに変換する。同様に、RAID制御部211aは、仮想ディスクVD2上のアドレスを指定したアクセス要求を受信すると、指定されたアドレスを、RAIDグループRG3上のボリュームVOL14の論理アドレスではなく、RAIDグループRG1上の移動後のボリュームVOL14の論理アドレスに変換する。これにより、これ以後の仮想マシンVM3, VM1からそれぞれボリュームVOL13, VOL14へのアクセスは、RAIDグループRG1上のボリュームVOL13, VOL14に対して行われるようになる。

10

【0117】

また、上記のパス切り替えの完了後、RAID制御部211aは、仮想マシンVM3, VM1からそれぞれRAIDグループRG3上のボリュームVOL13, VOL14への接続パスを閉塞する。そして、RAID制御部211aは、RAIDグループRG3上のボリュームVOL13, VOL14を削除する(ステップS19)。

【0118】

以上の処理により、RAIDグループRG1には、図15に示した空き領域FR3より大きな連続する空き領域FR4が形成される。これにより、RAIDグループRG1に対して空き領域FR3より大きなボリュームを新たに設定できるようになり、その結果、RAIDグループRG1に対応する記憶領域の利用効率が向上する。

20

【0119】

また、上記の処理では、パス切り替えの処理中を除き、仮想マシンVM3, VM1からそれぞれボリュームVOL13, VOL14に対するアクセスの受け付けが継続される。このため、再配置処理によりボリュームVOL13, VOL14にアクセスできなくなる期間が最小限に抑えられ、再配置処理が仮想マシンVM3, VM1の動作に与える影響を小さくすることができる。

【0120】

上記の処理が完了すると、RAID制御部211aは、完了通知を管理サーバ320に送信する。管理サーバ320では、制御部323によって完了通知を受信されると、管理部322は、ボリューム操作に応じたリソース使用状況の変化を、データベースサーバ310に登録された管理テーブルに反映させる。

30

【0121】

具体的には、管理部322は、仮想ディスクテーブル311c、ボリュームテーブル311dおよびストレージ領域管理テーブル311eから、ボリュームVOL12が登録されたレコードを削除する。なお、ストレージ領域管理テーブル311eでは、ボリュームVOL12が登録されたレコードが削除されることで、RAIDグループRG1に対してボリュームVOL11, VOL13, VOL14にそれぞれ対応するレコードが順に登録された状態となる。これにより、ボリュームVOL11とボリュームVOL13との間に空き領域が存在しないことが示されるようになるので、ストレージ領域管理テーブル311eについては、ボリュームVOL12が登録されたレコードの削除のみが行われればよい。

40

【0122】

なお、前述のように、本実施の形態では、RAIDグループに対するボリュームの割り当ては、RAIDグループの論理記憶領域の先頭側から順に、ボリュームが詰めて割り当てられる、というルールにしたがって行われる。上記の図15~図17でも、このルールが用いられている場合の処理例を示している。しかし、他の方法として、RAIDグループの論理記憶領域の末端側から順に、ボリュームが詰めて割り当てられる、というルールが用いられる場合もある。

【0123】

50

後者の方法が用いられる場合には、図15のステップS11においてボリュームVOL12の削除が要求されたとき、管理部322は、ボリュームVOL12より上位アドレス側に他のボリュームが設定されているかを判定する。そして、管理部322は、上位アドレス側に他のボリュームが設定されている場合に、他のボリュームの再配置が必要と判定する。図15の例では、ボリュームVOL11が再配置の対象となる。

【0124】

この後の再配置の処理過程では、図17のステップS17のように退避されたボリュームがRAIDグループRG1に書き戻される際に、そのボリュームのデータは、RAIDグループの論理記憶領域の先頭側でなく、末端側に詰めて書き込まれる。これにより、RAIDグループの論理記憶領域における先頭側の空き領域が拡張される。

10

【0125】

次に、管理サーバの処理について、図18～図23のフローチャートを用いて説明する。

図18は、ボリューム操作指示の受け付け処理の手順を示すフローチャートの例である。

【0126】

[ステップS101] 管理部322は、仮想ディスクの追加や削除、仮想ディスクの仕様変更など、ボリュームの操作が必要となる処理の指示を、操作端末330から受け付ける。このとき、管理部322は、追加、削除、使用変更などの対象の仮想ディスクを示す仮想ディスクIDや、その要求元のユーザ名を受け付ける。管理部322は、受け付けた仮想ディスクIDに対応付けられたボリュームのボリューム番号、仮想ディスクIDが示す仮想ディスクが接続された仮想サーバ(仮想マシン)を示す仮想サーバ名、その仮想サーバが動作するサーバ装置を示す物理サーバ名を、データベースサーバ310に問い合わせる。データベースサーバ310の入出力部312は、問い合わせに応じて、ユーザテーブル311a、仮想サーバテーブル311bおよび仮想ディスクテーブル311cを参照して上記各情報を抽出し、管理サーバ320に返信する。

20

【0127】

[ステップS102] 管理部322は、受け付けた仮想ディスクIDおよびユーザ名と、データベースサーバ310から返信された情報とに基づいて、指示された処理内容に対応するボリューム操作予約を示す予約データを生成し、予約データ管理テーブル321aに登録する。

30

【0128】

管理部322は、以上の図18の手順により予約データを予約データ管理テーブル321aに登録する。この登録処理と並行して、管理部322は、予約データ管理テーブル321aから予約データを登録順に読み込み、予約データに対応する処理を実行する。次の図19～図23では、管理部322が、ボリューム削除を指示する予約データを予約データ管理テーブル321aから読み込んだ場合の処理の例を示す。

【0129】

図19～図23は、ボリューム削除処理の手順を示すフローチャートの例である。

[ステップS111] 管理部322は、ボリューム削除を指示する予約データを予約データ管理テーブル321aから読み込む。

40

【0130】

[ステップS112] 管理部322は、読み込んだ予約データに基づいて、操作対象のボリュームに関する情報をデータベースサーバ310から収集する。管理部322は、データベースサーバ310から収集したデータを収集データとして結合し、収集データ管理テーブル321bに登録する。

【0131】

続いて、管理部322は、ボリューム設定状況の確認処理(ステップS113, S114, S116)を実行する。

[ステップS113] 管理部322は、削除対象のボリュームが設定されたRAIDグ

50

ループにおけるボリュームの設定状況を、データベースサーバ310に問い合わせる。データベースサーバ310の入出力部312は、ストレージ領域管理テーブル311eを参照して、該当するRAIDグループにおけるボリュームの設定状況を管理サーバ320に返信する。

【0132】

管理部322は、データベースサーバ310から返信された情報に基づき、削除対象のボリュームが設定されたRAIDグループにおいて、このボリュームより下位アドレス側に別のボリュームが設定されているかを判定する。管理部322は、下位アドレス側に別のボリュームが設定されている場合、ステップS114の処理を実行する。この場合、下位アドレス側のボリュームについての再配置が必要と判定される。一方、管理部322は、下位アドレス側にボリュームが設定されていない場合、図22のステップS141の処理を実行する。この場合、ボリューム再配置の必要はなく、削除対象のボリュームの削除処理のみが実行されることになる。

10

【0133】

[ステップS114] 管理部322は、削除対象のボリュームが設定された同一のストレージ装置内の各RAIDグループにおける空き領域を、データベースサーバ310に問い合わせる。データベースサーバ310の入出力部312は、ストレージ領域管理テーブル311eを参照して、該当ストレージ装置内の各RAIDグループにおける空き領域を管理サーバ320に返信する。

【0134】

管理部322は、データベースサーバ310から返信された情報に基づき、該当ストレージ装置内のRAIDグループの中に、退避先となる空き領域があるかを判定する。退避先の空き領域とは、再配置対象のボリュームの容量以上の連続した空き領域である。管理部322は、空き領域がある場合、ステップS115の処理を実行し、空き領域がない場合、ステップS116の処理を実行する。

20

【0135】

[ステップS115] 管理部322は、削除対象のボリュームが設定された同一のストレージ装置内のRAIDグループを、再配置対象のボリュームの退避先に決定する。このとき、管理部322は、退避先のRAIDグループと、そのRAIDグループ上の退避先領域とを決定する。退避先領域は、再配置対象のボリュームと同じ大きさの領域である。再配置対象のボリュームが複数存在する場合、それらの合計の大きさの領域が退避先領域となる。

30

【0136】

管理部322は、退避先のRAIDグループ上の退避先領域を予約するようにデータベースサーバ310に依頼する。データベースサーバ310の入出力部312は、この依頼に応じてストレージ領域管理テーブル311eにアクセスし、退避先領域のステータスを「Reserved」に設定する。これにより、ボリュームの再配置が完了するまで、退避先領域が他の用途で誤って使用されなくなり、再配置を確実に実行できるようになる。この後、処理は図20のステップS121に進められる。

【0137】

[ステップS116] 管理部322は、他のストレージ装置内の各RAIDグループにおける空き領域を、データベースサーバ310に問い合わせる。他のストレージ装置とは、削除対象のボリュームが設定されたストレージ装置以外のストレージ装置である。データベースサーバ310の入出力部312は、ストレージ領域管理テーブル311eを参照して、該当ストレージ装置内の各RAIDグループにおける空き領域を管理サーバ320に返信する。

40

【0138】

管理部322は、データベースサーバ310から返信された情報に基づき、他のストレージ装置内のRAIDグループの中に、退避先となる空き領域があるかを判定する。上記と同様に、退避先の空き領域とは、再配置対象のボリュームの容量以上の連続した空き領

50

域である。管理部 3 2 2 は、空き領域がある場合、ステップ S 1 1 7 の処理を実行し、空き領域がない場合、図 2 2 のステップ S 1 4 1 の処理を実行する。後者の場合、ボリュームの退避が不可能なので、ボリューム再配置は実行されず、削除対象のボリュームの削除処理のみが実行されることになる。

【 0 1 3 9 】

[ステップ S 1 1 7] 管理部 3 2 2 は、他のストレージ装置内の R A I D グループを、再配置対象のボリュームの退避先に決定する。このとき、管理部 3 2 2 は、退避先のストレージ装置および R A I D グループと、その R A I D グループ上の退避先領域とを決定する。

【 0 1 4 0 】

ステップ S 1 1 5 と同様に、管理部 3 2 2 は、退避先の R A I D グループ上の退避先領域を予約するようにデータベースサーバ 3 1 0 に依頼する。データベースサーバ 3 1 0 の入出力部 3 1 2 は、この依頼に応じてストレージ領域管理テーブル 3 1 1 e にアクセスし、退避先領域のステータスを「 R e s e r v e d 」に設定する。これにより、ボリュームの再配置が完了するまで、退避先領域が他の用途で誤って使用されなくなり、再配置を確実に実行できるようになる。この後、処理は図 2 0 のステップ S 1 2 1 に進められる。

【 0 1 4 1 】

以上の図 1 9 の処理では、ステップ S 1 1 4 またはステップ S 1 1 6 でボリュームの退避先となる空き領域があると判定されると、ボリュームの再配置が可能と判定される。この場合、管理部 3 2 2 は次に、図 2 0 に示す手順により、ステップ S 1 1 2 で登録した収集データに対応するボリューム操作（ボリューム削除および再配置）を実行するための制御情報を生成する。

【 0 1 4 2 】

以下、図 2 0 を用いて説明を続ける。

[ステップ S 1 2 1] 管理部 3 2 2 は、収集データに基づき、操作対象（削除対象および再配置対象）のボリュームが設定されたストレージ装置を示すストレージ装置名を、制御情報に指定する。

【 0 1 4 3 】

[ステップ S 1 2 2] 管理部 3 2 2 は、収集データに基づき、操作対象のボリュームが設定された R A I D グループを示す R A I D グループ名を、制御情報に指定する。

[ステップ S 1 2 3] 管理部 3 2 2 は、収集データに基づき、削除対象のボリュームを示すボリューム名を制御情報に指定する。

【 0 1 4 4 】

[ステップ S 1 2 4] 管理部 3 2 2 は、再配置対象のボリュームを示すボリューム名を制御情報に指定する。再配置対象のボリュームは、図 1 9 のステップ S 1 1 3 で削除対象のボリュームより下位アドレス側に設定されていると判定されたボリュームである。

【 0 1 4 5 】

[ステップ S 1 2 5] 管理部 3 2 2 は、再配置対象のボリュームの退避先となる R A I D グループを示す R A I D グループ名を、制御情報に指定する。退避先の R A I D グループは、図 1 9 のステップ S 1 1 5 またはステップ S 1 1 7 で決定された R A I D グループである。

【 0 1 4 6 】

[ステップ S 1 2 6] 管理部 3 2 2 は、生成された制御情報を制御部 3 2 3 に受け渡す。この後、図 2 1 のステップ S 1 3 1 の処理が実行される。

以下、図 2 1 を用いて説明を続ける。まず、図 2 1 のステップ S 1 3 1 ~ S 1 3 8 では、制御情報に基づき、ボリュームの削除および再配置を実行するための命令情報が生成される。

【 0 1 4 7 】

[ステップ S 1 3 1] 制御部 3 2 3 は、制御情報に基づき、削除対象のボリュームを削除するための削除命令を生成する。

10

20

30

40

50

【ステップ S 1 3 2】制御部 3 2 3 は、制御情報に基づき、再配置対象のボリュームを退避先領域へコピーするためのボリュームコピー命令を生成する。以下、この命令によって退避先領域に生成されるボリュームを「退避先ボリューム」と記載する。

【0 1 4 8】

【ステップ S 1 3 3】制御部 3 2 3 は、制御情報に基づき、元の R A I D グループ上の再配置対象のボリュームにアクセスするためのパスを退避先ボリュームへのパスに切り替えるためのパス切り替え命令と、切り替え前のボリュームへのパスを閉塞するためのパス閉塞命令とを生成する。なお、退避先ボリュームが設定されたストレージ装置が、コピー元のボリュームが設定されたストレージ装置と異なる場合がある。この場合、パス切り替え命令には、ボリュームへのアクセス元の仮想マシンが動作するサーバ装置から、退避先ボリュームが設定されたストレージ装置に対してアクセスパスを切り替えるための命令も含まれる。

10

【0 1 4 9】

【ステップ S 1 3 4】制御部 3 2 3 は、制御情報に基づき、コピー元のボリュームを削除するための削除命令を生成する。

【ステップ S 1 3 5】制御部 3 2 3 は、制御情報に基づき、退避先ボリュームを元の R A I D グループにおける空き領域の上位アドレス側に詰めてコピーする（書き戻す）ためのボリュームコピー命令を生成する。以下、この命令によって元の R A I D グループ上のコピー先領域に生成されるボリュームを「再配置先ボリューム」と記載する。

【0 1 5 0】

20

【ステップ S 1 3 6】制御部 3 2 3 は、制御情報に基づき、退避先ボリュームにアクセスするためのパスを再配置先ボリュームへのパスに切り替えるためのパス切り替え命令と、退避先ボリュームへのパスを閉塞するためのパス閉塞命令とを生成する。なお、再配置先ボリュームが設定されたストレージ装置が、退避先ボリュームが設定されたストレージ装置と異なる場合がある。この場合、パス切り替え命令には、ボリュームへのアクセス元の仮想マシンが動作するサーバ装置から、再配置先ボリュームが設定されたストレージ装置に対してアクセスパスを切り替えるための命令も含まれる。

【0 1 5 1】

【ステップ S 1 3 7】制御部 3 2 3 は、制御情報に基づき、退避先ボリュームを削除するための削除命令を生成する。

30

【ステップ S 1 3 8】制御部 3 2 3 は、ステップ S 1 3 1 ~ S 1 3 7 で生成された命令の実行シーケンスを生成し、各命令が実行シーケンスにしたがって実行されるように記述された命令情報を生成する。

【0 1 5 2】

【ステップ S 1 3 9】制御部 3 2 3 は、削除対象および再配置対象のボリュームが設定されたストレージ装置のコントローラに対して、生成された命令情報を発行（送信）する。この後、処理は図 2 3 のステップ S 1 5 1 に進められる。

【0 1 5 3】

命令情報を受信したコントローラの R A I D 制御部（R A I D 制御部 2 1 1 a , 2 2 1 a のいずれか）は、受信した命令情報にしたがってボリュームの削除および再配置の処理を実行する。すなわち、R A I D 制御部は、命令情報にしたがい、削除対象のボリュームの削除、再配置対象のボリュームの退避先領域へのコピー、パス切り替えおよび閉塞、コピー元のボリュームの削除、退避先ボリュームの元の R A I D グループへのコピー（書き戻し）、パス切り替えおよび閉塞、退避先ボリュームの削除を実行する。これらの実行が完了すると、R A I D 制御部は、完了通知を管理サーバ 3 2 0 に送信する。

40

【0 1 5 4】

次に、図 2 2 を用いて説明を続ける。

図 2 2 に示す処理は、上記の図 1 9 のステップ S 1 1 3 においてボリュームの再配置の必要がないと判定された場合、または、ステップ S 1 1 6 でボリュームの退避先となる空き領域がなく、再配置が不可能と判定された場合に実行される。

50

【 0 1 5 5 】

[ステップ S 1 4 1] 管理部 3 2 2 は、収集データに基づき、操作対象（削除対象）のボリュームが設定されたストレージ装置を示すストレージ装置名を、制御情報に指定する。

【 0 1 5 6 】

[ステップ S 1 4 2] 管理部 3 2 2 は、収集データに基づき、操作対象のボリュームが設定された R A I D グループを示す R A I D グループ名を、制御情報に指定する。

[ステップ S 1 4 3] 管理部 3 2 2 は、収集データに基づき、削除対象のボリュームを示すボリューム名を制御情報に指定する。

【 0 1 5 7 】

[ステップ S 1 4 4] 管理部 3 2 2 は、生成された制御情報を制御部 3 2 3 に受け渡す。

[ステップ S 1 4 5] 制御部 3 2 3 は、制御情報に基づき、削除対象のボリュームを削除するための削除命令を生成する。

【 0 1 5 8 】

[ステップ S 1 4 6] 制御部 3 2 3 は、生成された命令の実行シーケンスを生成し、この実行シーケンスにしたがって実行されるように記述された命令情報を生成する。

[ステップ S 1 4 7] 制御部 3 2 3 は、削除対象のボリュームが設定されたストレージ装置のコントローラに対して、生成された命令情報を発行（送信）する。この後、処理は図 2 3 のステップ S 1 5 1 に進められる。

【 0 1 5 9 】

命令情報を受信したコントローラの R A I D 制御部（R A I D 制御部 2 1 1 a , 2 2 1 a のいずれか）は、受信した命令情報にしたがってボリュームの削除処理を実行する。この処理の実行が完了すると、R A I D 制御部は、完了通知を管理サーバ 3 2 0 に送信する。

【 0 1 6 0 】

以下、図 2 3 を用いて説明を続ける。

[ステップ S 1 5 1] 制御部 3 2 3 は、図 2 1 のステップ S 1 3 9 または図 2 2 のステップ S 1 4 7 で発行した命令情報に対応する完了通知を受信するまで、待機状態となる。そして、制御部 3 2 3 は、完了通知を受信すると、ステップ S 1 5 2 の処理を実行する。

【 0 1 6 1 】

[ステップ S 1 5 2] 制御部 3 2 3 は、図 2 0 のステップ S 1 2 6 または図 2 2 のステップ S 1 4 4 で受け渡された制御情報に基づくボリューム操作が完了したことを管理部 3 2 2 に通知する。管理部 3 2 2 は、データベースサーバ 3 1 0 に管理テーブルの更新を依頼して、ボリューム操作に応じたリソース使用状況の変化を管理テーブルに反映させる。

【 0 1 6 2 】

データベースサーバ 3 1 0 の入出力部 3 1 2 は、管理部 3 2 2 からの依頼に応じて、仮想ディスクテーブル 3 1 1 c、ボリュームテーブル 3 1 1 d およびストレージ領域管理テーブル 3 1 1 e から、削除対象のボリュームが登録されたレコードを削除する。また、ボリュームの再配置が実行された場合、入出力部 3 1 2 は、管理部 3 2 2 からの依頼に応じて、図 1 9 のステップ S 1 1 5 またはステップ S 1 1 7 で予約された退避先領域について、予約が解除されるようにストレージ領域管理テーブル 3 1 1 e を更新する。これにより、ストレージ領域管理テーブル 3 1 1 e における退避先領域のステータスは、「R e s e r v e d」から「F r e e」に変更される。

【 0 1 6 3 】

[ステップ S 1 5 3] 管理部 3 2 2 は、図 1 9 のステップ S 1 1 1 で読み込んだ予約データを予約データ管理テーブル 3 2 1 e から削除する。また、管理部 3 2 2 は、ステップ S 1 1 2 で収集データ管理テーブル 3 2 1 b に登録した収集データを削除する。

【 0 1 6 4 】

以上の図 1 9 ~ 図 2 3 の処理によれば、削除対象のボリュームが設定された R A I D グ

10

20

30

40

50

ループにおいて、そのボリュームより下位アドレス側に他のボリュームが存在する場合、下位アドレス側のボリュームの再配置が実行される。これにより、RAIDグループの論理記憶領域において、削除されたボリュームより大きな連続する空き領域が形成される。そのため、そのRAIDグループに対して、削除されたボリュームより大きなボリュームを新たに設定できるようになり、その結果、そのRAIDグループに対応する記憶領域の利用効率を向上させることができる。

【0165】

また、ボリューム削除および再配置のための前処理（例えば、再配置の要否判定や退避先の決定）や、ボリューム削除および再配置のための一連の命令の生成処理は、ボリューム操作が行われる個々のストレージ装置でなく、管理サーバ320によって、データベースサーバ310に登録された管理テーブルに基づいて一元的に実行される。そして、生成された命令がストレージ装置に発行され、ストレージ装置は発行された命令にしたがってボリューム操作を実行する。このような処理により、ボリュームの削除や退避、書き戻し、パス変更を含む一連の処理を、ストレージ装置に正確に実行させることができる。特に、上記のボリューム操作を、再配置対象のボリュームに対する仮想マシンからのアクセスを可能にした状態のまま正確に実行できる。例えば、ボリューム操作の間、ボリュームの退避先領域の容量不足、意図しない通信断、データ破壊などの事態の発生を防止できる。

10

【0166】

また、上記の前処理や命令生成処理が不要になる分、ストレージ装置210、220のコントローラ211、221の処理負荷を軽減できる。そのため、ボリューム操作のための処理が、各仮想マシンから各仮想ディスクに対するアクセス速度に与える影響を低減できる。

20

【0167】

なお、上記の各実施の形態に示した装置（例えば、管理装置3、サーバ装置110、120、コントローラ211、221、データベースサーバ310、管理サーバ320）の処理機能は、コンピュータによって実現することができる。その場合、各装置が有すべき機能の処理内容を記述したプログラムが提供され、そのプログラムをコンピュータで実行することにより、上記処理機能がコンピュータ上で実現される。処理内容を記述したプログラムは、コンピュータで読み取り可能な記録媒体に記録しておくことができる。コンピュータで読み取り可能な記録媒体としては、磁気記憶装置、光ディスク、光磁気記録媒体、半導体メモリなどがある。磁気記憶装置には、ハードディスク装置（HDD）、磁気テープなどがある。光ディスクには、CD（Compact Disc）、DVD（Digital Versatile Disc）、ブルーレイディスク（Blu-ray Disc：BD、登録商標）などがある。光磁気記録媒体には、MO（Magneto-Optical disk）などがある。

30

【0168】

プログラムを流通させる場合には、例えば、そのプログラムが記録されたDVD、CDなどの可搬型記録媒体が販売される。また、プログラムをサーバコンピュータの記憶装置に格納しておき、ネットワークを介して、サーバコンピュータから他のコンピュータにそのプログラムを転送することもできる。

【0169】

プログラムを実行するコンピュータは、例えば、可搬型記録媒体に記録されたプログラムまたはサーバコンピュータから転送されたプログラムを、自己の記憶装置に格納する。そして、コンピュータは、自己の記憶装置からプログラムを読み取り、プログラムにしたがった処理を実行する。なお、コンピュータは、可搬型記録媒体から直接プログラムを読み取り、そのプログラムにしたがった処理を実行することもできる。また、コンピュータは、ネットワークを介して接続されたサーバコンピュータからプログラムが転送されるごとに、逐次、受け取ったプログラムにしたがった処理を実行することもできる。

40

【符号の説明】

【0170】

1, 2 ストレージ装置

50

1 a , 1 b , 2 a 記憶領域

1 a 1 , 2 a 1 空き領域

3 管理装置

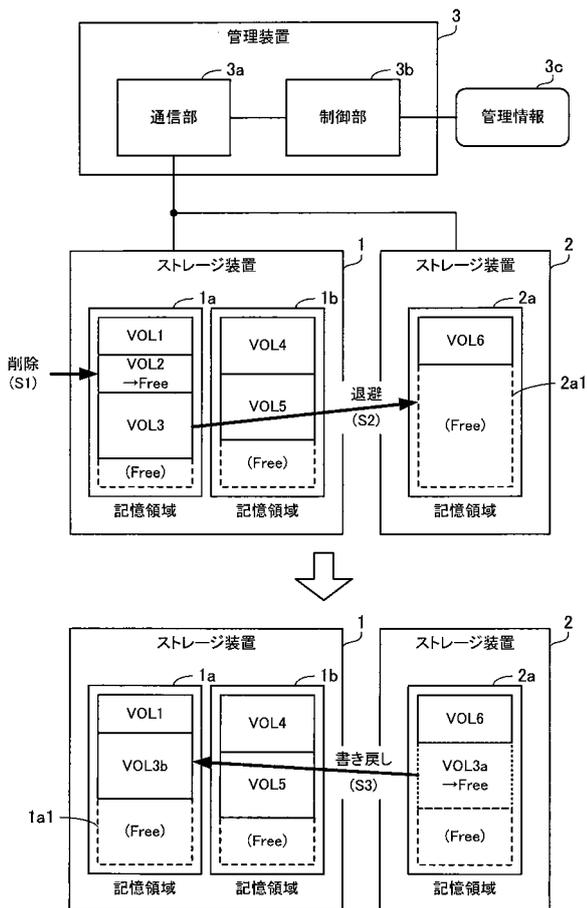
3 a 通信部

3 b 制御部

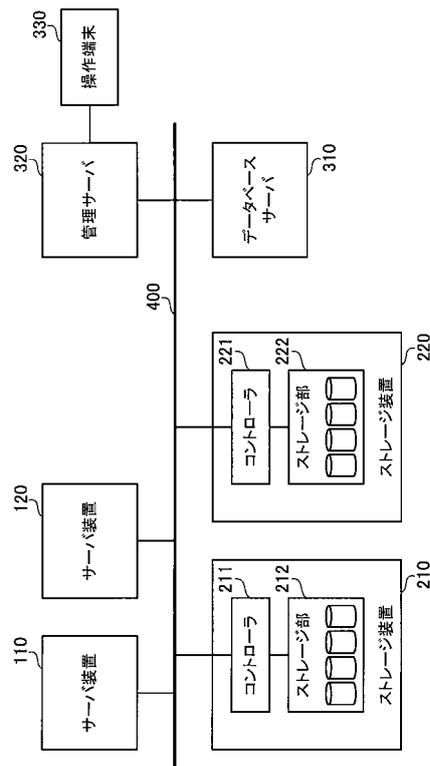
3 c 管理情報

VOL 1 ~ VOL 3 , VOL 3 a , VOL 3 b , VOL 4 ~ VOL 6 ボリューム

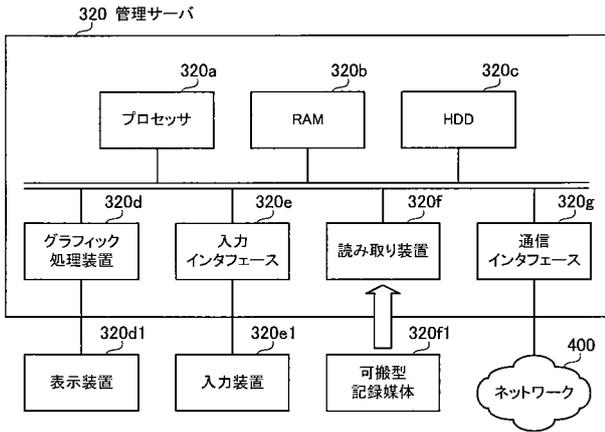
【 図 1 】



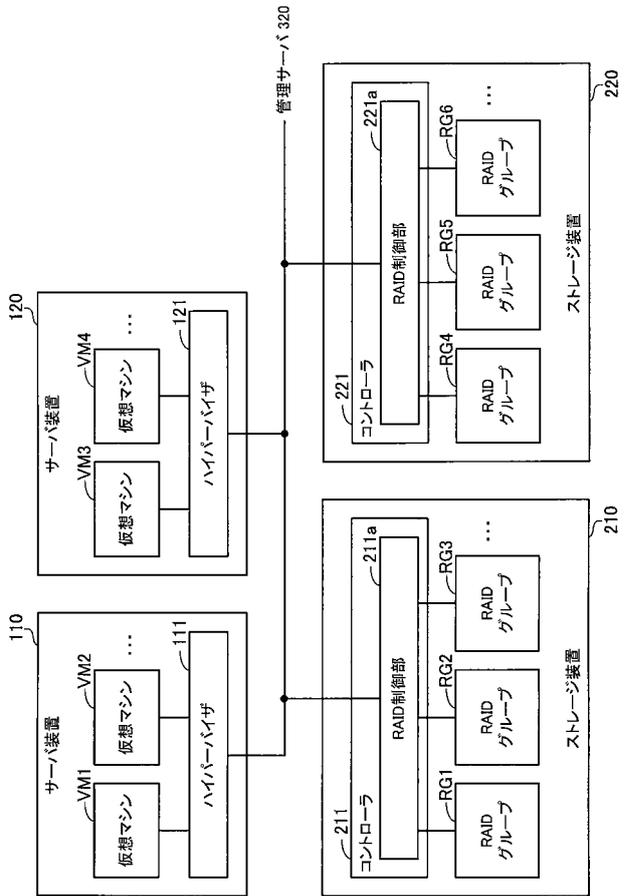
【 図 2 】



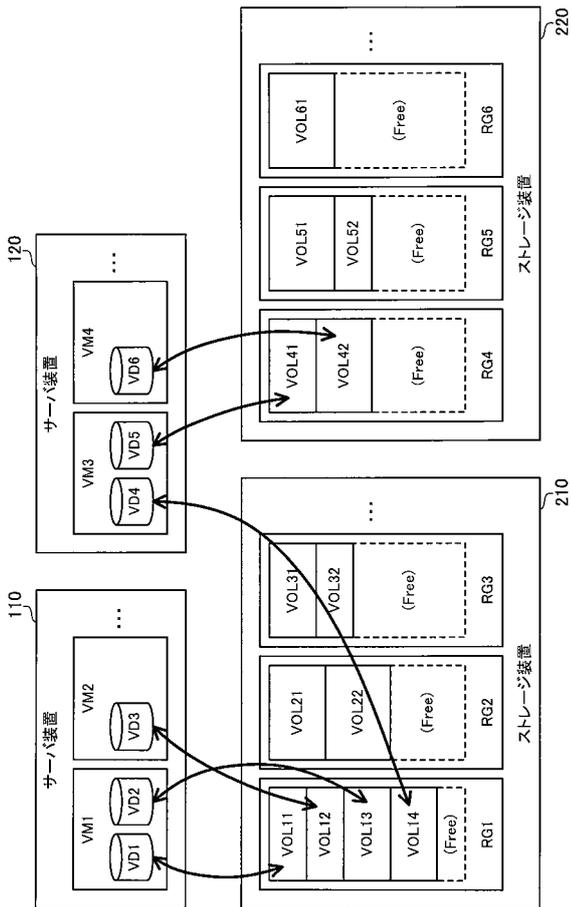
【図3】



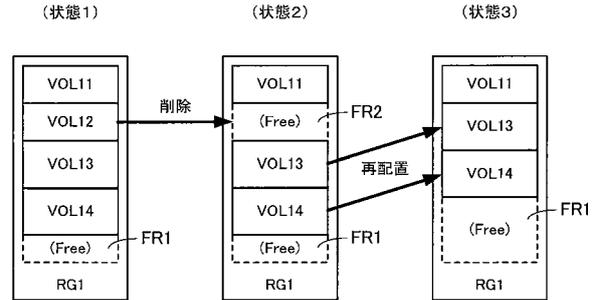
【図4】



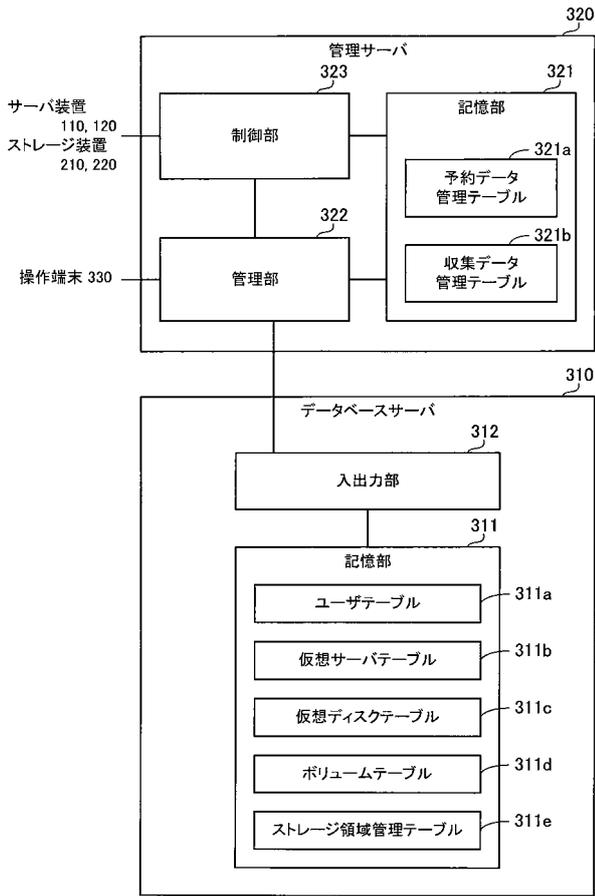
【図5】



【図6】



【 図 7 】



【 図 8 】

311a ユーザテーブル

ユーザ名	ユーザID
AA株式会社	0001
BB株式会社	0002
...	...

【 図 9 】

311b 仮想サーバテーブル

ユーザ名	ホスト名	仮想サーバ名	物理サーバ名	クラスタグループ	タイプ	マシンのタイプ	OSタイプ
0001	AAA	Dom0001	HW001	A	冗長	2	OS11u3
0001	AAB	Dom0002	HW002	B	冗長	2	OS11u3
0002	CCC	Dom0003	HW003	-	1	1	OS11u3
...

【 図 10 】

311c 仮想ディスクテーブル

仮想サーバ名	LUN番号	仮想ディスクID	ボリューム名	ディスクタイプ
Dom0001	1101	1101	vd-1101	共有
Dom0002	1101	1101	vd-1101	共有
Dom0003	1102	1102	vd-1102	シングル
...

【 図 11 】

311d ボリュームテーブル

ストレージ装置名	RAIDグループ名	ボリューム番号	ボリューム名	サイズ
ET200-1	RG020	1101	vd-1101	1TB
ET200-1	RG020	1101	vd-1101	1TB
ET200-2	RG000	1102	vd-1102	120GB
...

【図 1 2】

311e ストレージ領域管理テーブル

ストレージ装置名	RAIDグループ名	RAIDサイズ	ボリューム名	サイズ	ステータス
ET200-2	RG000	1TB	vd-1101	300G	Used
			vd-1102	120G	Used
	RG001	1TB	vd-1103	300G	Used
			vd-1104	300G	Used
			vd-1105	300G	Reserved
...

【図 1 3】

321a 予約データ管理テーブル

ユーザ情報		仮想サーバ名	物理サーバ名	操作種別	ボリューム番号
ユーザ名	ユーザID				
AA株式会社	00001	Dom0001	HW001	削除	1101
AA株式会社	00001	Dom0002	HW002	削除	1101
BB株式会社	00002	Dom0003	HW003	削除	1102

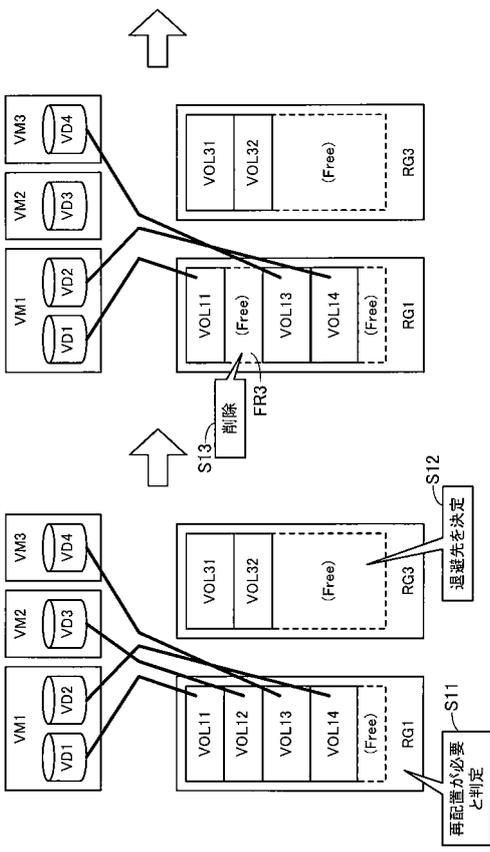
【図 1 4】

321b 収集データ管理テーブル

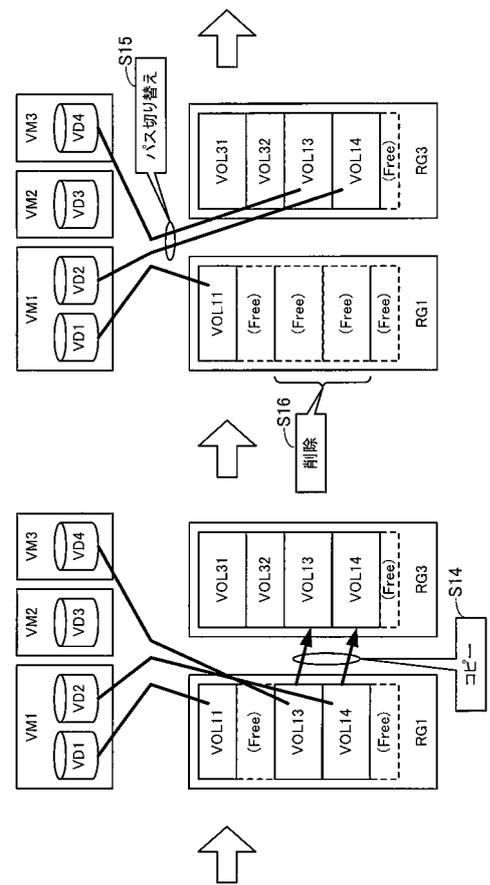
ユーザ情報	仮想サーバ名			物理サーバ名	操作種別	ストレージ装置名	LUN番号
ユーザ名	ユーザID
AA株式会社	00001	Dom0001	Dom0001	HW001	削除	ET200-1	1101
AA株式会社	00001	Dom0002	Dom0002	HW002	削除	ET200-1	1101
BB株式会社	00002	Dom0003	Dom0003	HW003	削除	ET200-2	1102

ボリューム情報					
RAIDグループ名	ボリューム番号	ボリューム名	サイズ	ディスクタイプ	ステータス
RG020	1101	vd-1101	1TB	共有	Used
RG020	1101	vd-1101	1TB	共有	Used
RG000	1102	vd-1102	120GB	シングル	Reserved

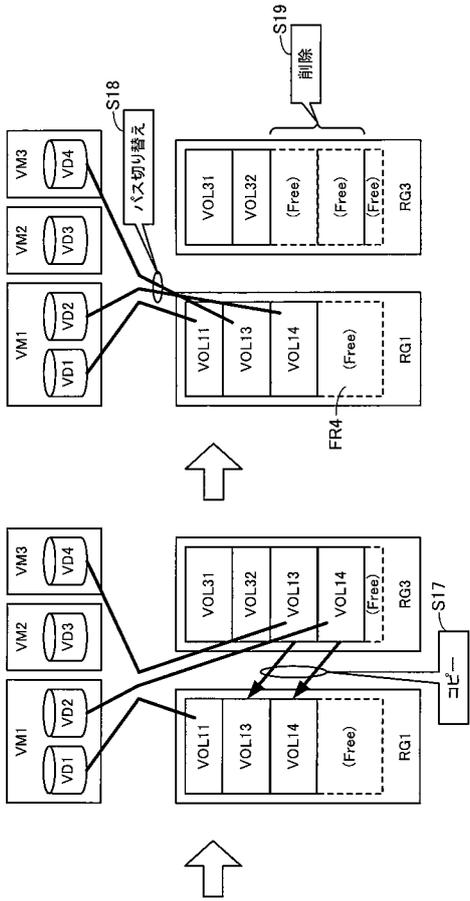
【図 1 5】



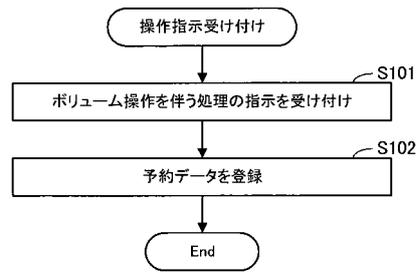
【図 1 6】



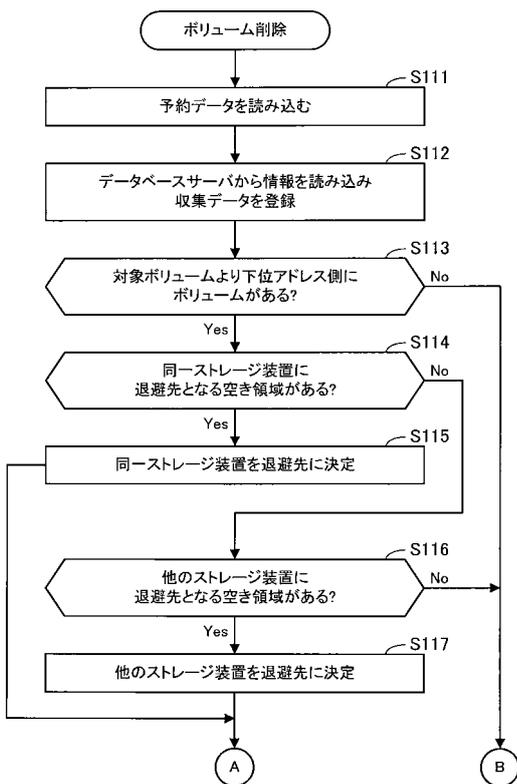
【図17】



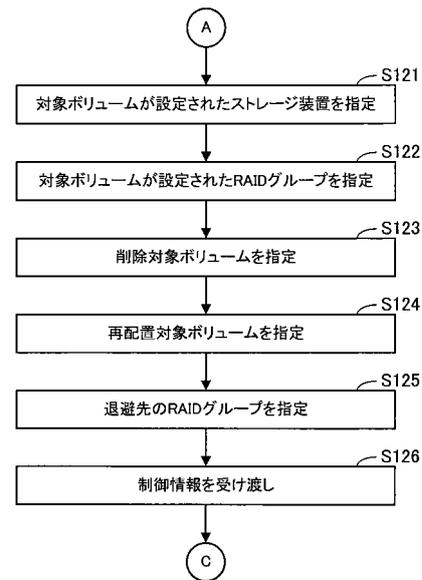
【図18】



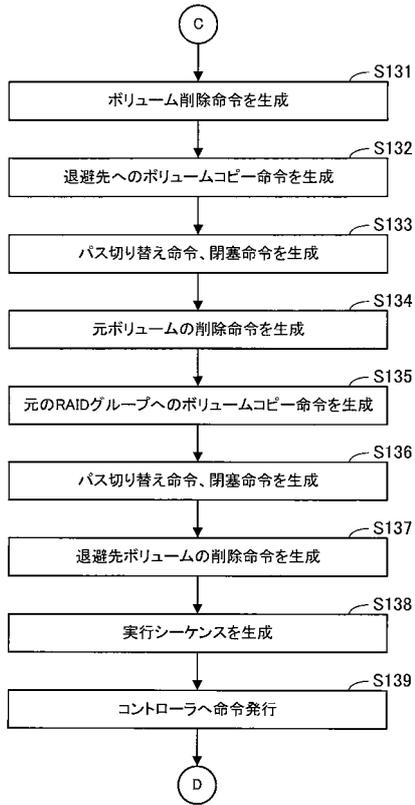
【図19】



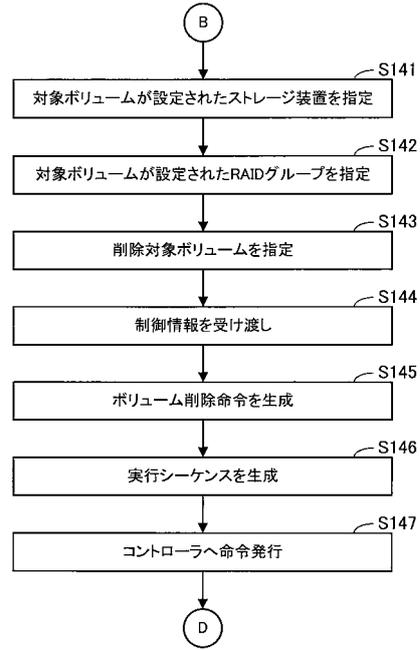
【図20】



【 図 2 1 】



【 図 2 2 】



【 図 2 3 】

