

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5182201号
(P5182201)

(45) 発行日 平成25年4月17日(2013.4.17)

(24) 登録日 平成25年1月25日(2013.1.25)

(51) Int. Cl. F I
G06F 3/06 (2006.01) G O 6 F 3/06 3 O 1 J
G06F 12/00 (2006.01) G O 6 F 12/00 5 O 1 M

請求項の数 5 (全 14 頁)

(21) 出願番号	特願2009-97916 (P2009-97916)	(73) 特許権者	000005223
(22) 出願日	平成21年4月14日(2009.4.14)		富士通株式会社
(65) 公開番号	特開2010-250476 (P2010-250476A)		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22年11月4日(2010.11.4)	(74) 代理人	100104190
審査請求日	平成23年12月5日(2011.12.5)		弁理士 酒井 昭徳
		(72) 発明者	内田 考介
			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	野口 泰生
			神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内
		審査官	木村 貴俊

最終頁に続く

(54) 【発明の名称】 ストレージ制御プログラム、ストレージシステム、およびストレージ制御方法

(57) 【特許請求の範囲】

【請求項1】

複数のサーバで共有可能なストレージ群を含む前記複数のサーバが使用する論理ディスクを制御する制御サーバにより前記ストレージ群の記憶領域の中から前記各サーバが使用できる記憶領域を設定するストレージシステムを構成するコンピュータを、

前記複数のサーバの中から選ばれた特定のサーバのみ使用する前記ストレージ群以外の特定のストレージの空き容量を検出する検出手段、

前記特定のストレージの中から前記検出手段によって検出された空き容量に応じた空き領域を特定する特定手段、

前記特定手段によって特定された空き領域を前記論理ディスクに組み込むことにより、前記空き領域を前記サーバ群で共有可能な共有記憶領域に設定する設定手段、

として機能させることを特徴とするストレージ制御プログラム。

【請求項2】

前記コンピュータを、

前記検出手段によって検出された空き容量が所定の空き容量以上であるか否かを判断する空き容量判断手段として機能させ、

前記特定手段は、

前記空き容量判断手段によって前記所定の空き容量以上であると判断された場合、前記特定のストレージの中から前記検出手段によって検出された空き容量に応じた空き領域を特定することを特徴とする請求項1に記載のストレージ制御プログラム。

【請求項 3】

前記コンピュータを、

前記特定のストレージのうち前記設定手段によって設定された共有記憶領域を除く残余の記憶領域の使用率が所定の使用率以上であるか否かを判断する使用率判断手段として機能させ、

前記設定手段は、

前記使用率判断手段によって前記所定の使用率以上であると判断された場合、前記共有記憶領域を、前記特定のサーバのみ使用する記憶領域に設定することを特徴とする請求項 1 または 2 に記載のストレージ制御プログラム。

【請求項 4】

複数のサーバで共有可能なストレージ群を含む前記複数のサーバが使用する論理ディスクを制御する制御サーバにより前記ストレージ群の記憶領域の中から前記各サーバが使用できる記憶領域を設定するストレージシステムにおいて、

前記複数のサーバの中から選ばれた特定のサーバのみ使用する前記ストレージ群以外の特定のストレージの空き容量を検出する検出手段と、

前記特定のストレージの中から前記検出手段によって検出された空き容量に応じた空き領域を特定する特定手段と、

前記特定手段によって特定された空き領域を前記論理ディスクに組み込むことにより、前記空き領域を前記サーバ群で共有可能な共有記憶領域に設定する設定手段と、

を備えることを特徴とするストレージシステム。

【請求項 5】

複数のサーバで共有可能なストレージ群を含む前記複数のサーバが使用する論理ディスクを制御する制御サーバにより前記ストレージ群の記憶領域の中から前記各サーバが使用できる記憶領域を設定するストレージシステムを構成するコンピュータが、

前記複数のサーバの中から選ばれた特定のサーバのみ使用する前記ストレージ群以外の特定のストレージの空き容量を検出する検出工程と、

前記特定のストレージの中から前記検出工程によって検出された空き容量に応じた空き領域を特定する特定工程と、

前記特定工程によって特定された空き領域を前記論理ディスクに組み込むことにより、前記空き領域を前記サーバ群で共有可能な共有記憶領域に設定する設定工程と、

を実行することを特徴とするストレージ制御方法。

【発明の詳細な説明】

【技術分野】

【0001】

本開示技術は、ストレージを制御するストレージ制御プログラム、ストレージシステム、およびストレージ制御方法に関する。

【背景技術】

【0002】

従来、ストレージ群を仮想領域として業務サーバで共有するストレージシステムが開示されている（たとえば、下記特許文献 1 を参照。）。

【0003】

図 1 2 は、従来のストレージシステムのシステム構成図である。ストレージシステム 100 は、データセンタ 101 と業務システム 102 群とを含む構成である。ストレージシステム 100 内では、LAN (Local Area Network)、WAN (Wide Area Network)、インターネットなどのネットワーク 103 により通信可能に接続されている。

【0004】

データセンタ 101 は、業務システム 102 群にストレージサービスを提供するシステムである。ストレージサービスとは、各業務システム 102 に対してデータの読み出しや書き込みを提供するコンピュータ処理である。データセンタ 101 は、管理ノード MP と

10

20

30

40

50

制御ノードC Pと複数のストレージ装置S Tとを含む構成である。管理ノードM Pは、サービス提供先の業務サーバSを管理するコンピュータである。制御ノードC Pは、全物理ディスクD aのディスク領域を一元管理するコンピュータである。

【0005】

ストレージ装置S Tは、ディスクノードD Pと物理ディスクD aとで構成されている。ディスクノードD Pは、制御ノードC Pからの制御により物理ディスクD aからデータを読み出したり、物理ディスクD aにデータを書き込んだりするコンピュータである。物理ディスクD aは、ハードディスクや不揮発性メモリ、磁気テープにより構成されている。

【0006】

物理ディスクD aは、スライスと呼ばれる単位で各業務サーバSにディスク領域を割り当てている。ある業務サーバSに割り当てられるスライスは、複数の物理ディスクD aにまたがっているが、それぞれの業務サーバSから見た場合、単一の物理ディスクであるかのように見える。このことから、全物理ディスクD aのディスク領域は仮想領域V（または仮想ディスク）と称される。

10

【0007】

業務システム102は、各種業務（たとえば、Webサービス、検索サービス、ストレージサービスなど）を提供するシステムであり、業務サーバSとその業務サーバS専用の物理ディスクD bとを有する。業務サーバSは、データセンタ101からストレージの割当サービスを受けるコンピュータである。各業務サーバSは、仮想領域Vに保存されたデータを読み出したり、仮想領域Vにデータを書き込むことができる。実際には、業務サーバSに割り当てられたディスクノードD Pの物理ディスクD a内のスライスにアクセスして、データの読み出し/書込みをおこなう。

20

【0008】

業務サーバSの物理ディスクD bは、その業務サーバSが利用する。各業務サーバSの物理ディスクD bに空き領域があると、空き領域が分散していることとなり、ストレージシステム100として非効率な状態である。

【先行技術文献】

【特許文献】

【0009】

【特許文献1】特開2005-4681号公報

30

【発明の概要】

【発明が解決しようとする課題】

【0010】

本開示技術では、業務サーバ自身が保有するストレージの空き領域が増加した場合に仮想領域Vに組み込むことで、ストレージシステム100におけるストレージの有効活用を図ることを課題とする。また、業務サーバ自身が保有するストレージの空き領域が減少した場合に仮想領域Vから切り離すことで、業務サーバSにおけるストレージの有効活用を図ることを課題とする。

【課題を解決するための手段】

【0011】

40

上述した課題を解決し、目的を達成するため、本開示技術は、複数のサーバで共有可能なストレージ群を制御する制御サーバにより前記ストレージ群の記憶領域の中から前記各サーバが使用できる記憶領域を設定するストレージシステムにおいて、前記複数のサーバの中から選ばれた特定のサーバのみ使用する前記ストレージ群以外の特定のストレージの空き容量を検出し、前記特定のストレージの中から、検出された空き容量に応じた空き領域を特定し、特定された空き領域を、前記サーバ群で共有可能な共有記憶領域に設定することを要件とする。

【発明の効果】

【0012】

本開示技術によれば、ストレージの有効活用を図ることができるという効果を奏する。

50

【図面の簡単な説明】

【0013】

【図1】本開示技術のストレージシステムのシステム構成図である。

【図2】コンピュータのハードウェア構成を示すブロック図である。

【図3】業務サーバ一覧テーブルの記憶内容を示す説明図である。

【図4】論理ディスク構成テーブルの記憶内容を示す説明図である。

【図5】ディスクノード一覧テーブルの記憶内容を示す説明図である。

【図6】仮想領域自動組込処理のシーケンスを示すシーケンス図である。

【図7】空きパーティション数特定処理（ステップS604）およびデータ退避処理（ステップS605）の具体例を示す説明図である。

10

【図8】変更後の論理ディスク構成テーブルの記憶内容を示す説明図である。

【図9】更新後のディスクノード一覧テーブルを示す説明図である。

【図10】仮想領域自動退避処理のシーケンスを示すシーケンス図である。

【図11】業務サーバ/ストレージシステムの機能的構成を示すブロック図である。

【図12】従来のストレージシステムのシステム構成図である。

【発明を実施するための形態】

【0014】

以下に添付図面を参照して、本開示技術のストレージ制御プログラム、ストレージシステム、およびストレージ制御方法の好適な実施の形態を詳細に説明する。

【0015】

20

(システム構成)

図1は、本開示技術のストレージシステムのシステム構成図である。図12に示した構成と同一構成には同一符号を付し、その説明を省略する。本開示技術では、業務サーバSが保有する物理ディスクDbについても空き領域がある場合には、データセンタ101内の物理ディスクDaと同様、空き領域について仮想領域Vに含める。これにより、業務サーバSの物理ディスクDbの空き領域も、データセンタ101内の物理ディスクDaと同様、仮想領域Vとなり、リソースの有効活用を図ることができる。たとえば、ある業務システム102（たとえば、A社の業務システム）の業務サーバSの物理ディスクDbに空き領域を仮想領域Vに含める。これにより、他の業務システム102（たとえば、B社の業務システム。図示しないC社、D社...の業務システムでもよい。）の業務サーバSは、増加した仮想領域Vを利用することができる。

30

【0016】

(コンピュータのハードウェア構成)

図2は、コンピュータのハードウェア構成を示すブロック図である。図2において、コンピュータは、CPU(Central Processing Unit)201と、ROM(Read Only Memory)202と、RAM(Random Access Memory)203と、磁気ディスクドライブ204と、磁気ディスク205と、光ディスクドライブ206と、光ディスク207と、ディスプレイ208と、I/F(Interface)209と、キーボード210と、マウス211と、スキャナ212と、プリンタ213と、を備えている。また、各構成部はバス200によってそれぞれ接続されている。

40

【0017】

ここで、CPU201は、コンピュータの全体の制御を司る。ROM202は、ブートプログラムなどのプログラムを記憶している。RAM203は、CPU201のワークエリアとして使用される。磁気ディスクドライブ204は、CPU201の制御にしたがって磁気ディスク205に対するデータのリード/ライトを制御する。磁気ディスク205は、磁気ディスクドライブ204の制御で書き込まれたデータを記憶する。

【0018】

光ディスクドライブ206は、CPU201の制御にしたがって光ディスク207に対するデータのリード/ライトを制御する。光ディスク207は、光ディスクドライブ20

50

6の制御で書き込まれたデータを記憶したり、光ディスク207に記憶されたデータをコンピュータに読み取らせたりする。

【0019】

ディスプレイ208は、カーソル、アイコンあるいはツールボックスをはじめ、文書、画像、機能情報などのデータを表示する。このディスプレイ208は、たとえば、CRT、TFT液晶ディスプレイ、プラズマディスプレイなどを採用することができる。

【0020】

インターフェース(以下、「I/F」と略する。)209は、通信回線を通じてLAN、WAN、インターネットなどのネットワーク103に接続され、ネットワーク103を介して他の装置に接続される。そして、I/F209は、ネットワーク103と内部のインターフェースを司り、外部装置からのデータの入出力を制御する。I/F209には、たとえばモデムやLANアダプタなどを採用することができる。

10

【0021】

キーボード210は、文字、数字、各種指示などの入力のためのキーを備え、データの入力をおこなう。また、タッチパネル式の入力パッドやテンキーなどであってもよい。マウス211は、カーソルの移動や範囲選択、あるいはウィンドウの移動やサイズの変更などをおこなう。ポインティングデバイスとして同様に機能を備えるものであれば、トラックボールやジョイスティックなどであってもよい。

【0022】

スキャナ212は、画像を光学的に読み取り、コンピュータ内に画像データを取り込む。なお、スキャナ212は、OCR(Optical Character Reader)機能を持たせてもよい。また、プリンタ213は、画像データや文書データを印刷する。プリンタ213には、たとえば、レーザプリンタやインクジェットプリンタを採用することができる。

20

【0023】

(各種テーブルの記憶内容)

つぎに、本開示技術で使用する各種テーブルについて説明する。以下に示すテーブルは、具体的には、たとえば、図2に示したROM202、RAM203、磁気ディスク205などの記憶媒体によりその機能を実現する。

【0024】

図3は、業務サーバー一覧テーブルの記憶内容を示す説明図である。業務サーバー一覧テーブル300は、業務サーバーの所在(本例ではIP(Internet Protocol)アドレス)を登録するテーブルである。業務サーバー一覧テーブル300は、管理ノードMPが保有する。

30

【0025】

図4は、論理ディスク構成テーブルの記憶内容を示す説明図である。論理ディスク構成テーブル400とは、ストレージシステム100で定義された論理ディスクの内容を記憶するテーブルである。論理ディスク構成テーブル400は、各業務サーバーが保有するテーブルである。論理ディスク構成テーブル400は、ディスク番号と、パーティション番号と、割り当て状態と、容量とを、論理ディスクごとに記憶する。すなわち、各レコードはパーティションの情報を示している。ディスク番号とは、その業務サーバーが保有する物理ディスクDbを特定する識別情報である。

40

【0026】

パーティション番号とは、物理ディスクDb内のパーティションを特定する識別情報である。パーティションとは、物理ディスクDbを分割した部分的な記憶領域である。図4の例では、物理ディスクDb番号:1番の物理ディスクDbは、パーティション番号:1番~4番の4つのパーティションに分割されている。

【0027】

割り当て状態とは、パーティションの割当先を示す情報である。図4において、「業務」である場合は、その業務サーバーに割り当てられており、「仮想」である場合は、仮想

50

領域Vとして割り当てられている。容量とは、論理ディスクであるパーティションの記憶容量（サイズ）である。本例では、各パーティションの記憶容量は同一（図4では、100[GB]）とする。

【0028】

図5は、ディスクノード一覧テーブルの記憶内容を示す説明図である。ディスクノード一覧テーブル500は、制御ノードCPが保有するテーブルである。ディスクノード一覧テーブル500は、あらたにディスクノードDPとなる業務サーバSの物理ディスクDbに関する情報を記憶するテーブルである。

【0029】

具体的には、業務サーバSごとに、業務サーバSのIPアドレスと、業務サーバSの物理ディスクDbのディスク番号と、そのディスク番号で特定される物理ディスクDb内のパーティション番号を記憶する。すなわち、ディスクノード一覧テーブル500のレコードは、仮想領域Vに組み込まれた業務サーバSの物理ディスクDbのパーティションを示している。

【0030】

たとえば、図5のレコードは、IPアドレス：10.77.205.57で特定される業務サーバSが保有するディスク番号：1番の物理ディスクDb内のパーティション番号：4番のパーティションが仮想領域Vに組み込まれたことを示している。他のパーティションがあらたに仮想領域Vに組み込まれるとそのレコードが追加され、仮想領域Vから解放されるとそのレコードは削除されることとなる。

【0031】

（仮想領域自動組込処理）

図6は、仮想領域自動組込処理のシーケンスを示すシーケンス図である。まず、管理ノードMPが、所定のタイミングで自動的にまたは管理者の操作指示により、バッチジョブである組込プログラムを対象となる業務サーバS（以下、「対象業務サーバS」）に送信する（ステップS601）。具体的には、たとえば、リナックスのsshコマンドを実行する。

【0032】

対象業務サーバSは組込プログラムを受信すると、組込プログラムにより、使用パーティション特定処理（ステップS602）、使用量測定処理（ステップS603）、空きパーティション数特定処理（ステップS604）、データ退避処理（ステップS605）を実行する。

【0033】

使用パーティション特定処理（ステップS602）とは、業務サーバSが使用できるパーティション、換言すれば、仮想領域Vに割り当てられていないパーティションを特定する処理である。具体的には、論理ディスク構成テーブル400にアクセスして、割り当て状態が「業務」であるパーティションを特定する。図4の例では、パーティション番号：1番～3番のパーティションが特定される。

【0034】

使用量測定処理（ステップS603）とは、使用パーティション特定処理（ステップS602）で特定されたパーティションの現在の使用量を測定する処理である。具体的には、たとえば、リナックスのdfコマンドを実行する。

【0035】

空きパーティション数特定処理（ステップS604）では、空きパーティションの数を特定する処理である。空きパーティション数の特定に際し、使用パーティション特定処理（ステップS602）で特定されたパーティションの総記憶容量から、使用量測定処理（ステップS603）で測定された使用量を引く。この計算結果が空き容量となる。つぎに、計算された空き容量がパーティションサイズ以上か否かを判断する。

【0036】

パーティションサイズとは、1パーティションの記憶容量（図4の例では、100[GB]

10

20

30

40

50

B]) である。空き容量が1パーティションサイズ以上である場合、1パーティションサイズの空き容量があることとなる。同様に、空き容量が2パーティションサイズ以上である場合、1パーティションサイズ以上の空き容量があることとなる。このように、最大でいくつのパーティションサイズ分の空き容量があるか特定されるまで判断を続ける。

【 0 0 3 7 】

データ退避処理 (ステップ S 6 0 5) は、パーティション内でデータを集約し、空きパーティション数特定処理 (ステップ S 6 0 4) により特定された空きパーティション数分の空きパーティションを作る。ここで具体例を説明する。

【 0 0 3 8 】

図7は、空きパーティション数特定処理 (ステップ S 6 0 4) およびデータ退避処理 (ステップ S 6 0 5) の具体例を示す説明図である。図7中、P 1 ~ P 3 は、図4に示したパーティション番号1~3のパーティションである。本例では、パーティションP 1の使用量が70 [GB]、パーティションP 2の使用量が60 [GB]、パーティションP 3の使用量が50 [GB]とする。(A)において、パーティションP 1~P 3の総記憶容量は300 [GB]、総使用量は180 [GB]であるため、パーティションP 1~P 3の空き容量は、120 [GB]となる。

10

【 0 0 3 9 】

そして、空き容量120 [GB]が1パーティションサイズである100 [GB]以上か否かを判断する。この場合は100 [GB]以上であるため、1パーティションサイズ分の空き容量があることがわかる。つぎに、空き容量120 [GB]が2パーティションサイズである200 [GB]以上か否かを判断する。この場合は200 [GB]以上でないため、2パーティションサイズ分の空き容量ではなく、1パーティションサイズ分の空き容量で確定する。

20

【 0 0 4 0 】

(B)では、パーティションP 3に格納されている50 [GB]分のデータを、パーティションP 1, P 2に移行する。これにより、パーティションP 3は空となり、仮想領域Vとして組み込むことができる。

【 0 0 4 1 】

図6に戻って、データ退避処理 (ステップ S 6 0 5) が完了すると、組込プログラムにより組込指示をおこなう (ステップ S 6 0 6)。組込指示には、空きパーティションのパーティション番号が含まれる。このあと、論理ディスク制御プログラムを起動する。論理ディスク制御プログラムは、組込処理を実行する (ステップ S 6 0 7)。組込処理 (ステップ S 6 0 7) とは、図4に示した論理ディスク構成テーブル400にアクセスして、組込指示に含まれているパーティション番号の割り当て状態の項目を「仮想」に変更する。

30

【 0 0 4 2 】

図8は、変更後の論理ディスク構成テーブル400の記憶内容を示す説明図である。図8では、図4に示した状態から、パーティション番号: 3番のパーティションP 3の割り当て状態を「仮想」に変更する。

【 0 0 4 3 】

図6に戻って、組込処理 (ステップ S 6 0 7) のあと、組込プログラムは、ディスクノードプログラムにパーティション情報 (「仮想」に変更されたパーティション番号) を与え (ステップ S 6 0 8)、組込対象 (本例ではパーティションP 3) のフォーマット処理を実行する (ステップ S 6 0 9)。具体的には、たとえば、パーティションP 3の全域にオールゼロの値を書き込む。これにより、物理ディスクD bのパーティションP 3に仮想領域Vが生成されることとなる (ステップ S 6 1 0)。

40

【 0 0 4 4 】

仮想領域Vが生成されると、ディスクノードプログラムは、制御ノードCPに対し仮想領域情報を送信する (ステップ S 6 1 1)。仮想領域情報とは、仮想領域Vが設定された業務サーバSのIPアドレス、仮想領域Vが設定されたディスク番号、仮想領域Vが設定された物理ディスクD b内のパーティション番号である。制御ノードCPでは、ディスク

50

ノード一覧テーブル更新処理（ステップS 6 1 2）を実行する。具体的には、ステップS 6 1 1で送信されてきた仮想領域情報をディスクノード一覧テーブル5 0 0の新規レコードとして追加する。

【0 0 4 5】

図9は、更新後のディスクノード一覧テーブル5 0 0を示す説明図である。図5と比較すると、レコードが新たに追加されていることがわかる。

【0 0 4 6】

（仮想領域自動退避処理）

図10は、仮想領域自動退避処理のシーケンスを示すシーケンス図である。まず、対象業務サーバSが、所定のタイミングで自動的にまたは操作指示により、使用パーティション特定処理（ステップS 1 0 0 1）、使用量測定処理（ステップS 1 0 0 2）、パーティション返却判断処理（ステップS 1 0 0 3）を実行する。

【0 0 4 7】

使用パーティション特定処理（ステップS 1 0 0 1）は、図6に示した使用パーティション特定処理（ステップS 6 0 2）と同一処理内容であり、使用量測定処理（ステップS 1 0 0 2）は、使用量測定処理（ステップS 6 0 3）と同一処理内容である。

【0 0 4 8】

パーティション返却判断処理（ステップS 1 0 0 3）では、仮想領域Vに設定されたパーティションを返却すべきか否か、すなわち、仮想領域Vから解放して対象業務サーバSにのみ割り当てるか否かを判断する。具体的には、使用パーティション特定処理（ステップS 1 0 0 1）で特定されたパーティションの総記憶容量から、使用量測定処理（ステップS 1 0 0 2）で測定された使用量を引く。この計算結果が空き容量となる。

【0 0 4 9】

つぎに、計算された空き容量がしきい値となる空き容量より大きいか否かを判断する。大きければ、対象業務サーバSの容量不足となるため、制御ノードCPに対し仮想領域Vに設定されたパーティションの返却要求を送信する（ステップS 1 0 0 4）。返却要求には、対象業務サーバSのIPアドレス、仮想領域Vに設定されたパーティションのパーティション番号とそのパーティションを含む物理ディスクDbのディスク番号が含まれている。計算された空き容量がしきい値となる空き容量以下である場合、まだ余裕があるとして、返却要求をしない。

【0 0 5 0】

制御ノードCPは、返却要求を受信すると、ディスクノードプログラムに対し、返却指示をおこなう（ステップS 1 0 0 5）。返却指示には、返却要求と同一の情報（対象業務サーバSのIPアドレス、仮想領域Vに設定されたパーティションのパーティション番号とそのパーティションを含む物理ディスクDbのディスク番号）が含まれている。ディスクノードプログラムは、返却指示を制御ノードCPから受けると、解放処理を実行する（ステップS 1 0 0 6）。

【0 0 5 1】

具体的には、返却指示に埋め込まれている仮想領域Vに設定されたパーティション番号により、ディスクノード一覧テーブル5 0 0内の該当するレコードを特定し、特定されたレコードをディスクノード一覧テーブル5 0 0から削除する。これにより、制御ノードCPからは、該当するパーティションが仮想領域Vであると認識できず、解放されることとなる（ステップS 1 0 0 7）。このあと、ディスクノードプログラムは、解放情報を制御ノードCPに送信する（ステップS 1 0 0 8）。解放情報とは、仮想領域Vから開放されたパーティション情報（「業務」に変更されたパーティション番号）である。

【0 0 5 2】

制御ノードCPでは、解放情報を受けると、仮想領域Vの減少により、データの2重保存処理を実行する（ステップS 1 0 0 9）。2重保存処理については、上述した特許文献1の技術であるため、詳細な説明を省略する。制御ノードCPは、このあと、解放情報に含まれているパーティション情報を組込プログラムに与える（ステップS 1 0 1 0）。組

10

20

30

40

50

込プログラムは、パーティション情報を受けると、論理ディスク制御プログラムに対し、パーティション情報により特定されるパーティションの仮想領域Vからの切離し指示を与える(ステップS1011)。

【0053】

このあと、論理ディスク制御プログラムを起動する。論理ディスク制御プログラムは、切離し処理を実行する(ステップS1012)。切離し処理(ステップS1012)とは、図4に示した論理ディスク構成テーブル400にアクセスして、パーティション情報(ステップS1010)に含まれているパーティション番号の割り当て状態の項目を「業務」に変更する。

【0054】

パーティションP3が仮想領域Vから解放されたとすると、論理ディスク構成テーブル400のパーティション番号:3番のパーティションP3の割り当て状態を、図8に示した「仮想」から「業務」に変更する。この場合、図4に示した割り当て状態に戻る。これにより、仮想領域Vから解放されたパーティションが業務領域として対象業務サーバSのみが使用する記憶領域となる(ステップS1013)。

【0055】

(業務サーバS/ストレージシステム100の機能的構成)

図11は、業務サーバS/ストレージシステム100の機能的構成を示すブロック図である。図11において、業務サーバS/ストレージシステム100は、検出部1101と、空き容量判断部1102と、特定部1103と、使用率判断部1104と、設定部1105とを備える。この制御部となる機能(検出部1101~設定部1105)は、具体的には、たとえば、図2に示したROM202、RAM203、磁気ディスク205、光ディスク207などの記憶装置に記憶されたプログラムをCPU201に実行させることにより、または、I/F209により、その機能を実現する。

【0056】

検出部1101は、複数のサーバの中から選ばれた特定のサーバのみ使用するストレージ群以外の特定のストレージの空き容量を検出する機能を有する。具体的には、たとえば、業務サーバS群の中から選ばれた対象業務サーバS専用の物理ディスクDbの空き容量を検出する。より具体的には、図6に示した仮想領域自動組込処理では、使用パーティション特定処理(ステップS602)および使用量測定処理(ステップS603)に相当する。また、図10に示した仮想領域自動退避処理では、使用パーティション特定処理(ステップS1001)および使用量測定処理(ステップS1002)に相当する。

【0057】

空き容量判断部1102は、検出部1101によって検出された空き容量が所定の空き容量以上であるか否かを判断する機能を有する。具体的には、たとえば、空き容量が、1パーティションサイズ以上であるか否か、2パーティションサイズ以上であるか否か、...、nパーティションサイズ以上であるか否かを判断する。より具体的には、図6に示した仮想領域自動組込処理では、空きパーティション数特定処理(ステップS604)に相当する。

【0058】

特定部1103は、特定のストレージの中から検出部1101によって検出された空き容量に応じた空き領域を特定する機能を有する。具体的には、たとえば、対象業務サーバS専用の物理ディスクDbの中から、検出された空き容量に応じた空き領域を特定する。より具体的には、nパーティションサイズ以上でないかと判断された場合、パーティション数を(n-1)個であると特定する。パーティションサイズをm[GB]とすると、空き容量は(n-1)×mとなる。図6に示した仮想領域自動組込処理では、空きパーティション数特定処理(ステップS604)に相当する。

【0059】

使用率判断部1104は、特定のストレージのうち共有記憶領域を除く残余の記憶領域の使用率が所定の使用率以上であるか否かを判断する機能を有する。具体的には、たとえ

10

20

30

40

50

ば、対象業務サーバS専用の物理ディスクDbのうち仮想領域Vを除く残余のパーティションの使用率が所定の使用率以上であるか否かを判断する。たとえば、残余のパーティションの総記憶容量のうち使用量が所定のしきい値以上であるか否かを判断する。図10に示した仮想領域自動退避処理では、パーティション返却判断処理(ステップS1003)に相当する。

【0060】

設定部1105は、特定部1103によって特定された空き領域を、サーバ群で共有可能な共有記憶領域に設定する機能を有する。具体的には、たとえば、空き領域に相当するパーティションを業務サーバS群で共有可能な仮想領域Vに設定する。図6に示した仮想領域自動組込処理では、組込処理(ステップS607)、ディスクノード一覧テーブル更新処理(ステップS612)に相当する。

10

【0061】

また、設定部1105は、使用率判断部1104によって所定の使用率以上であると判断された場合、共有記憶領域を、特定のサーバのみ使用する記憶領域に設定する機能を有する。具体的には、たとえば、残余のパーティションの総記憶容量のうち使用量が所定のしきい値以上である場合、仮想領域Vである対象業務サーバS専用の物理ディスクDb内のパーティションを、対象業務サーバSにのみ使用されるパーティションに設定する。図10に示した仮想領域自動退避処理では、解放処理(ステップS1006)および切離し処理(ステップS1012)に相当する。

【0062】

20

以上説明したように、本開示技術によれば、業務サーバS自身が保有するストレージの空き領域が増加した場合に仮想領域Vに組み込むことで、ストレージシステム100におけるストレージの有効活用を図ることができる。

【0063】

また、空き容量を判断することで、空き容量に応じたパーティションを空き領域として特定することができるため、業務サーバSの物理ディスクDb内において、仮想領域Vとなるパーティションと業務サーバSが専有するパーティションが混在していても問題はなく、リソースの効率的な活用を実現することができる。

【0064】

さらに、業務サーバS自身が保有するストレージの空き領域が減少した場合に仮想領域Vから切り離すことで、業務サーバSの空き容量が逼迫した場合にも対処することができる。したがって、業務サーバSにおけるストレージの有効活用を図ることができる。

30

【0065】

なお、本実施の形態で説明したストレージ制御方法は、予め用意されたプログラムをパーソナル・コンピュータやワークステーション等のコンピュータで実行することにより実現することができる。ストレージ制御プログラムは、ハードディスク、フレキシブルディスク、CD-ROM、MO、DVD等のコンピュータで読み取り可能な記録媒体に記録され、コンピュータによって記録媒体から読み出されることによって実行される。またストレージ制御プログラムは、インターネット等のネットワークを介して配布してもよい。

【符号の説明】

40

【0066】

CP 制御ノード

DP ディスクノード

Da, Db 物理ディスク

MP 管理ノード

S 業務サーバ

V 仮想領域

100 ストレージシステム

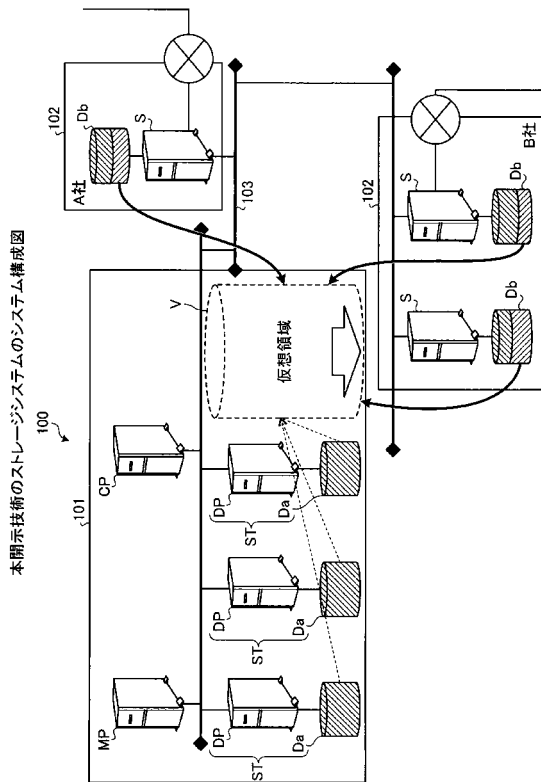
101 データセンタ

102 業務システム

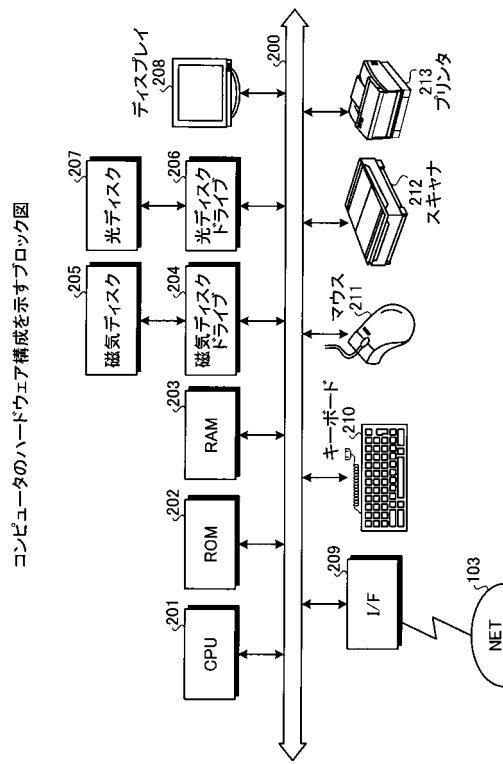
50

- 3 0 0 業務サーバー一覧テーブル
- 4 0 0 論理ディスク構成テーブル
- 5 0 0 ディスクノード一覧テーブル
- 1 1 0 1 検出部
- 1 1 0 2 空き容量判断部
- 1 1 0 3 特定部
- 1 1 0 4 使用率判断部
- 1 1 0 5 設定部

【 図 1 】



【 図 2 】



【 図 3 】

業務サーバ一覧テーブルの記憶内容を示す説明図

300

業務サーバ名	IPアドレス
A社ウェブサーバ	10.77.205.57
B社DBサーバ	10.77.205.58
B社アプリケーションサーバ	10.77.206.12

【 図 5 】

ディスクノード一覧テーブルの記憶内容を示す説明図

500

IPアドレス	ディスク番号	パーティション番号
10.77.205.57	1	4

【 図 4 】

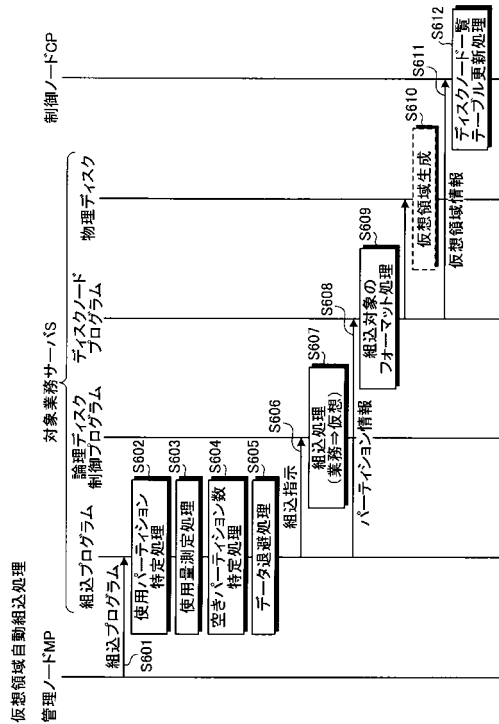
論理ディスク構成テーブルの記憶内容を示す説明図

400

ディスク番号	パーティション番号	割り当て状態	容量
1	1	業務	100GB
1	2	業務	100GB
1	3	業務	100GB
1	4	仮想	100GB

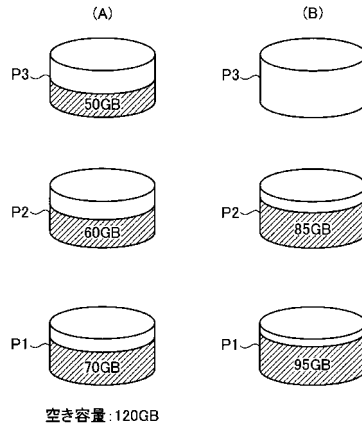
【 図 6 】

仮想領域自動組込処理のシーケンスを示すシーケンス図



【 図 7 】

空きパーティション数特定処理(ステップS604)およびデータ退避処理(ステップS605)の具体例を示す説明図



【 図 8 】

変更後の論理ディスク構成テーブルの記憶内容を示す説明図

400

ディスク番号	パーティション番号	割り当て状態	容量
1	1	業務	100GB
1	2	業務	100GB
1	3	仮想	100GB
1	4	仮想	100GB

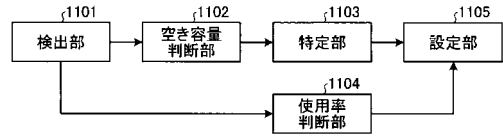
【図9】

更新後のディスクノード一覧テーブルを示す説明図

IPアドレス	ディスク番号	パーティション番号
10.77.205.57	1	4
10.77.205.57	1	3

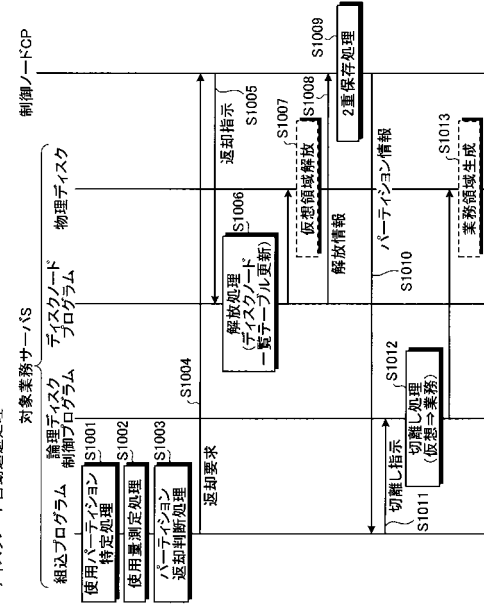
【図11】

業務サーバ/ストレージシステムの機能的構成を示すブロック図



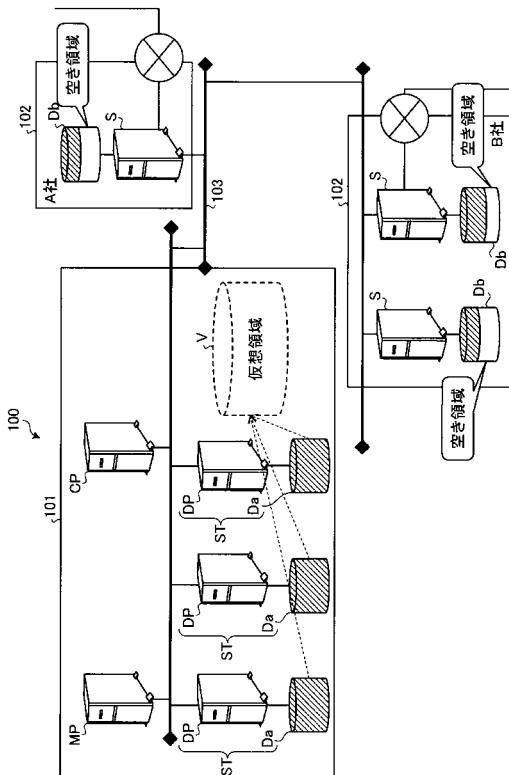
【図10】

仮想領域自動退避処理のシーケンスを示すシーケンス図
ディスクノード自動退避処理



【図12】

従来のストレージシステムのシステム構成図



フロントページの続き

- (56)参考文献 特開平05 - 334006 (JP, A)
特開2007 - 087083 (JP, A)
特開2003 - 323329 (JP, A)

(58)調査した分野(Int.Cl., DB名)

G06F 3/06 - 3/08
G06F 12/00 - 12/16
G06F 13/00 - 13/42