

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局



(43) 国際公開日
2012年2月16日(16.02.2012)

PCT

(10) 国際公開番号
WO 2012/020471 A1

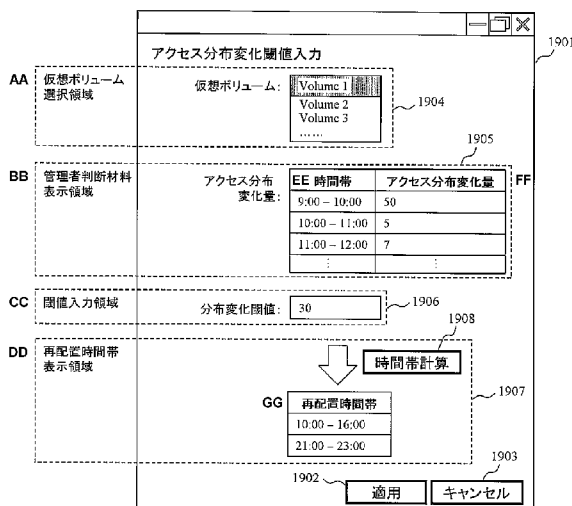
- (51) 国際特許分類:
G06F 12/00 (2006.01)
- (21) 国際出願番号: PCT/JP2010/063512
- (22) 国際出願日: 2010年8月10日(10.08.2010)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (71) 出願人(米国を除く全ての指定国について): 株式会社日立製作所(HITACHI, LTD.) [JP/JP]; 〒1008280 東京都千代田区丸の内一丁目6番6号 Tokyo (JP).
- (72) 発明者; および
- (75) 発明者/出願人(米国についてのみ): 宮本 佑樹(MIYAMOTO Yuuki) [JP/JP]; 〒2448555 神奈川県横浜市戸塚区戸塚町5030番地 株式会社日立製作所 ソフトウェア事業部内 Kanagawa (JP). 朝木 克利(ASAKI Katsutoshi) [JP/JP]; 〒2448555 神奈川県横浜市戸塚区戸塚町5030
- 番地 株式会社日立製作所 ソフトウェア事業部内 Kanagawa (JP).
- (74) 代理人: 平木 祐輔(HIRAKI Yusuke); 〒1050001 東京都港区虎ノ門4丁目3番20号 神谷町MTビル19階 Tokyo (JP).
- (81) 指定国(表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.
- (84) 指定国(表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア

[続葉有]

(54) Title: COMPUTER SYSTEM MANAGEMENT METHOD AND MANAGEMENT SYSTEM

(54) 発明の名称: 計算機システムの管理方法、及び管理システム

[図19]



- 1901 ACCESS DISTRIBUTION CHANGE THRESHOLD INPUT
1902 APPLY
1903 CANCEL
1904 VIRTUAL VOLUME
1905, FF ACCESS DISTRIBUTION CHANGE AMOUNT
1906 DISTRIBUTION CHANGE THRESHOLD
1908 TIME PERIOD CALCULATION
AA VIRTUAL VOLUME SELECTION REGION
BB MANAGER DECISION MATERIAL DISPLAY DOMAIN
CC THRESHOLD INPUT DOMAIN
DD RELOCATION TIME PERIOD DISPLAY DOMAIN
EE TIME PERIOD
GG RELOCATION TIME PERIOD

(57) Abstract: The present invention provides technology for efficiently relocating data in a real domain. A management system obtains information about changes to load concentration domains (the amount of change to access distribution) which indicates changes to positions in a virtual domain which become hotspots in the virtual volume of a storage subsystem. The management system then specifies an unchanged load position time period which indicates a time period that the position of the virtual domain that becomes a hotspot remains almost or completely unchanged (the time period that the hotspot position is stable), based on the information about changes in load concentration domains.

(57) 要約: 実領域のデータ再配置をより効率の良く実行するための技術を提供するものである。管理システムが、ストレージサブシステムの仮想ボリュームにおいてホットスポットとなる仮想領域の位置変化を示す負荷集中領域変化情報(アクセス分布の変化量)を取得する。そして、当該管理システムが、負荷集中領域変化情報に基づいて、ホットスポットとなる仮想領域の位置がほぼ又は完全に不変となる時間帯(ホットスポットの位置が安定した時間帯)を示す負荷位置不変時間帯を特定し、それを表示デバイスに表示する。



(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), ヨーロッパ
(AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR,
GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT,
NL, NO, PL, PT, RO, SE, SI, SK, SM, TR), OAPI (BF,
BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE,
SN, TD, TG).

添付公開書類:

— 国際調査報告 (条約第 21 条(3))

明 細 書

発明の名称： 計算機システムの管理方法、及び管理システム 技術分野

[0001] 本発明は、計算機システムの管理方法、及び管理システムに関し、例えば、ストレージサブシステム内の記憶領域を、複数の階層に分割して管理する方法、およびそのような管理を実行する管理システムに関する。

背景技術

[0002] 最近、仮想化技術とThin Provisioning技術とを組み合わせることにより、論理ボリュームよりも細かい単位でのストレージ階層管理することが考えられている（ストレージ階層管理技術）。

[0003] このストレージ階層管理技術では、Thin Provisioningに従う仮想的な論理ボリューム（以下、仮想ボリューム）の記憶領域を、複数の部分的な領域（以下、「仮想領域」という）に分割し、仮想領域の単位で、どの階層に属する記憶デバイスから実体的な記憶領域（以下、実領域）を割り当てるかが選択される。例えば、特許文献1によれば、仮想ボリュームの記憶領域と物理記憶デバイスの記憶領域を、それぞれ2つ以上の仮想領域および実領域に分割し、その各仮想領域に、実領域を1つ割り当てる。また、実領域ごとに、ホスト計算機からのアクセス数とアクセス数の許容範囲を記録し、許容範囲を超えたアクセス数の実領域のデータを、アクセス状況に応じた記憶デバイスの実領域に移行する。

先行技術文献

特許文献

[0004] 特許文献1：米国特許公開第2007/0283079号公報

発明の概要

発明が解決しようとする課題

[0005] 上記特許文献1記載の技術では、データの過去のアクセス履歴から当該データを格納するのに適した特性を持つデバイスを特定し、そのデバイスにデ

ータを移行する。

[0006] しかしながら、特許文献 1 に記載されるような従来のストレージ階層管理技術では、単にアクセス数に基づいてデータ移行を実行しているだけなので、データ移行に必要な時間又は性能負荷に代表されるコストを考慮して効率良くデータの移行を実現することが困難であった。

[0007] 本発明はこのような状況に鑑みてなされたものであり、より効率の良いデータ移行を実現するための技術を提供するものである。

課題を解決するための手段

[0008] 上記課題を解決するために、本発明では、管理システムが、ストレージサブシステムの仮想ボリュームにおいてホットスポットとなる仮想領域の位置変化を示す負荷集中領域変化情報（アクセス分布の変化量）を取得する。そして、当該管理システムが、負荷集中領域変化情報に基づいて、ホットスポットとなる仮想領域の位置が不変となる時間帯（ホットスポットの位置が安定した時間帯）を示す負荷位置不変時間帯を特定し、それを表示デバイスに表示する。

[0009] さらに本発明の特徴は、以下本発明を実施するための形態および添付図面によって明らかになるものである。

発明の効果

[0010] 本発明によれば、管理者は、実領域の再配置処理を効率よく実行できる時間帯を知る（予測）ことができるようになる。

図面の簡単な説明

[0011] [図1] 本発明の実施形態による計算機システム 10 の概略構成を示す図である。

[図2] ホスト計算機 101 の内部構成を示すブロック図である。

[図3] ストレージサブシステム 102 の内部構成を示すブロック図である。

[図4] RAID グループ管理表 308 の構成例を示す図である。

[図5] 実領域管理表 309 の構成例を示す図である。

[図6] 第 1 の実施形態における VVOL 管理表 311 の構成例を示す図である。

。

[図7]階層定義表 3 1 0 の構成例を示す図である。

[図8]管理計算機 1 0 3 の内部構成を示すブロック図である。

[図9]アクセス分布変化履歴表 8 0 5 の構成例を示す図である。

[図10]再配置時間帯管理表 8 0 6 の構成例を示す図である。

[図11]管理側 R A I D グループ管理表 8 0 7 の構成例を示す図である。

[図12]管理側実領域管理表 8 0 8 の構成例を示す図である。

[図13]第 1 の実施形態における管理側 V V O L 管理表 8 1 0 の構成例を示す図である。

[図14]管理側階層定義表 8 0 9 の構成例を示す図である。

[図15]リード要求に対応する処理を説明するためのフローチャートである。

[図16]ライト要求に対応する処理を説明するためのフローチャートである。

[図17]アクセス分布変化量計算処理を説明するためのフローチャートである。

。

[図18]第 1 の実施形態における再配置時間帯計算処理を説明するためのフローチャートである。

[図19]分布変化閾値入力画面 1 9 0 1 の構成例を示す図である。

[図20]画面 1 9 0 1 の管理者判断材料表示領域 1 9 0 5 に表示するアクセス分布変化量の、グラフによる表示態様例を示す図である。

[図21]再配置処理を説明するためのフローチャートである。

[図22]第 2 の実施形態における V V O L 管理表 3 1 1 の構成例を示す図である。

[図23]第 2 の実施形態における管理側 V V O L 管理表 8 1 0 の構成例を示す図である。

[図24]実領域の全データがゼロになった場合に当該実領域に対応する仮想領域を未割当ての領域とする処理（ゼロインクラメーション）の概念を示す図である。

[図25]第 2 の実施形態における再配置処理を説明するためのフローチャート

である。

発明を実施するための形態

[0012] 本発明は、Thin Provisioningに従う仮想的な論理ボリュームに対する実体的な記憶領域割り当ての見直し時間帯を制御するための技術に関するものである。下記実施形態では、仮想ボリューム毎に、ホスト計算機からのアクセスの、ボリューム内の仮想領域への分布に関するデータを収集する。その値を分析することで、仮想ボリュームごとに、アクセス分布が変化しない時間帯を特定する。仮想ボリュームごとに、アクセス分布が変化しない時間帯に合わせて、当該仮想ボリューム内の仮想領域への、割当実領域の見直し処理（以下、再配置処理）を行う。これにより、計算機システムの管理者は、仮想ボリュームに対してアクセス分布が変化しない時間帯を特定するための閾値を設定するだけで、再配置処理を行う時間帯を、その効果が長期間にわたって得られると予想される場合に限定することができ、効果が低い再配置処理を行うことによる性能の低下やストレージサブシステムの運用コストの増大を防ぐことができる。

[0013] 以下、添付図面を参照して本発明の実施形態について説明する。ただし、本実施形態は本発明を実現するための一例に過ぎず、本発明の技術的範囲を限定するものではないことに注意すべきである。また、各図において共通の構成については同一の参照番号が付されている。

[0014] なお、以後の説明では「aaa表」という表現にて本発明の各種情報を説明するが、これら情報は必ずしも表によるデータ構造で表現されていなくても良く、リスト、DB、キュー等のデータ構造やそれ以外で表現されていても良い。そのため、データ構造に依存しないことを示すために「aaa表」、「aaaリスト」、「aaaDB」、「aaaキュー」等について「aaa情報」と呼ぶことがある。

[0015] また、各情報の内容を説明する際に、「識別情報」、「識別子」、「名」、「名前」、「ID」という表現を用いることが可能であり、これらについてはお互いに置換が可能である。

[0016] さらに、以後の説明では「プログラム」を主語として説明を行うこともあるが、プログラムはプロセッサによって実行されることで定められた処理をメモリ及び通信ポート（通信制御装置）を用いながら行うため、プロセッサを主語とした説明としてもよい。また、プログラムを主語として開示された処理は管理サーバ等の計算機、情報処理装置が行う処理としてもよい。プログラムの一部または全ては専用ハードウェアで実現してもよく、また、モジュール化されていても良い。各種プログラムはプログラム配布サーバや記憶メディアによって各計算機にインストールされてもよい。

[0017] （１）第１の実施形態

＜特徴的処理の概要説明＞

本実施形態による計算機システムの構成及び動作を詳細に説明する前に、まず当該計算機システムの特徴的処理の概要について説明する。計算機システムは、ストレージサブシステムにホスト計算機および管理システムを有する。管理システムは、ストレージサブシステムを管理する。管理システムは、ストレージサブシステムまたはホスト計算機に含まれても良い。ストレージサブシステムは、プールと、複数の仮想ボリュームとを提供する。

[0018] プールは、複数の階層で構成されている。階層は、同じ性能を有する実領域の集合である。階層は例えば性能を表す特定の値によって定義され、その値が表す性能と同じ性能を有する実領域を内包する。階層は、他の階層と比較した高低の概念を持つ。階層の高さは、階層の性能の高さに依存する。なお、ここで言う「性能」とは、例えばアクセス性能である。アクセス性能としては、例えばレスポンスタイム、あるいはデータ転送速度、IOPS（単位時間あたりに処理したアクセス要求の数）がある。アクセス要求の一例としてはライト要求又はリード要求がある。

[0019] 各階層は、２以上の実領域で構成されている。したがって、プールは複数の実領域で構成されているとも言える。実領域は、物理記憶デバイスまたは／およびRAIDグループの記憶領域の一部である。階層の性能は、階層を構成する実領域の性能に依存する。実領域の性能は、例えば、当該実領域を

有する物理記憶デバイスの種別、RAIDグループのRAIDレベル、または／およびRAIDグループに参加する記憶デバイスの数に依存する。物理記憶デバイスの種別として、例えば、SSDと、SAS-HDD、およびSATA-HDDがある。

[0020] 仮想ボリュームは、複数の仮想領域（仮想的な記憶領域）で構成されている。各仮想領域には、それぞれ一つ以下の実領域が割り当てられている。なお、仮想領域には、実領域が割り当てられていなくてもよい。ストレージサブシステムは、そのような仮想領域（未割当の仮想領域）に対するライト要求をホスト計算機から受けた時、いずれの仮想領域にも割り当てられていない実領域（未割当の実領域）を、当該仮想領域に割り当てる。

[0021] 以下で、本実施形態において実行される処理の概要を説明する。本実施形態において実行される処理には、例えば、（i）アクセス分布変化量計算処理と、（ii）再配置時間帯計算処理と、（iii）再配置処理とがある。これらの概要について以下簡単に説明する。

[0022] （i）アクセス分布変化量計算処理

管理システムは、ストレージサブシステムが有する各仮想ボリュームのアクセス分布の変化量を計算し、記録する。ここで、仮想ボリュームのアクセス分布とは、例えば、ホスト計算機から当該仮想ボリューム内の各仮想領域へアクセスした数の集合である。これにより、仮想ボリューム全体のアクセス量のうちどの位の割合が各仮想領域にあるかが分かり、よって、ある時点でのホットスポット（負荷が掛かっている仮想領域）の位置が分かる。また、仮想ボリュームのアクセス分布の変化量とは、仮想ボリューム内において時間経過と共にどのようにホットスポットとなる仮想領域の位置が変化したかを示す指標（負荷集中領域変化指標）となるものである。例えば、「アクセス分布の変化量が多い」とは、1つの仮想ボリュームにおいて、アクセスされる仮想領域が激しく変化していることを意味し、「アクセス分布の変化量が少ない」とは、1つの仮想ボリュームにおいて、アクセス数に関係なく、アクセスされる仮想領域の変化が少ない（ホットスポットの位置に動きが

少ない)ことを意味している。

[0023] なお、「仮想領域へのアクセス数」とは対象仮想領域の全てまたは一部をアドレス範囲として指定し、処理が完了した（または処理中も含めても良い）アクセス要求の数を指す。例えば、アクセス要求のアドレス範囲がある仮想領域の一部または全てである（言い方を変えれば、仮想領域がアドレス範囲を含んでしまう）場合、当該要求の数だけ当該仮想領域のアクセス数を増加させる。他の例として、アクセス要求のアドレス範囲が複数の仮想領域それぞれの一部または全てを含む場合は、当該要求の数だけ、仮想領域それぞれのアクセス数を増加させる。なお、仮想領域のサイズが大きい場合は後者のようなケースはあまり発生しないため、アクセス要求で指定したアドレス範囲の先頭を含む仮想領域のカウント数だけ増やしても良い。

[0024] 仮想ボリュームのアクセス分布の変化量は、例えば、過去のある時点での当該仮想ボリュームのアクセス分布に関する情報と、現在の当該仮想ボリュームのアクセス分布に関する情報から計算できる。管理システムは、例えば定期的または繰り返し（周期的でなくても良く、取得間隔に多少の時間のずれがあっても良い）、ストレージサブシステムから各仮想ボリュームのアクセス分布に関する情報を取得し、前回取得した情報と合わせて当該仮想ボリュームのアクセス分布の変化量を計算し、時系列に記録する。この場合、仮想ボリュームのアクセス分布に関する情報は、前回取得した情報のみ記録すればよく、仮想ボリュームのアクセス分布に関する情報を時系列に記録する場合と比べて、管理システムがアクセス分布の変化量の計算に使用する情報を記録するために必要な記憶領域を少なくすることができる。

[0025] (ii) 再配置時間帯計算処理

管理システムは、各仮想ボリュームのアクセス分布の変化量の時系列データを分析し、各仮想ボリューム内の仮想領域に対する実領域の割当状況を見直す時間帯（以下、再配置時間帯）を計算して記録する。例えば、各仮想ボリュームに対して、1日を1時間ごとに区切った各時間帯（例えば、9:00-10:00、10:00-11:00など）におけるアクセス分布の変化量の平均値を、直近1

週間の当該仮想ボリュームのアクセス分布の変化量データから計算し、その平均値がある閾値よりも小さい時間帯を、当該仮想ボリュームの再配置時間帯として記録する。

[0026] 本処理は「仮想ボリュームのアクセス分布の変化量が小さい時間帯では、当該仮想ボリューム内の各仮想領域に対するアクセス状況の変化も小さく、各仮想領域に割り当てべき物理記憶デバイスまたは／およびＲＡＩＤグループも変化しない可能性が高い。したがって、当該時間帯の初期に当該仮想ボリューム内の仮想領域に対する実領域の割り当て状況を見直せば、当該時間帯の末尾の時刻までは割り当て状況を変更しなくてもよい可能性が高い」という考えに基づいている。

[0027] なお、アクセス分布の変化量の平均値を計算する単位である時間帯の長さ（上記処理では１時間）は、例えば１０分や３時間でもよい。また、平均値を計算する期間（上記処理では１週間）は、例えば３日間や１ヶ月間でもよい。さらに、再配置時間帯は、例えば日付や曜日の情報を含んでもよい。そのように時間帯以外の情報を含む場合、時間帯の情報を含まなくてもよい。時間帯の情報を含まない場合、管理システムは、再配置時間帯が示す日付または／および曜日の特定の時刻に、ストレージサブシステムに再配置処理の実行を指示する。

[0028] (iii) 再配置処理

再配置時間帯計算処理によって、ホットスポットの動きが少ない（ほとんどない）時間帯が得られるが、この時間帯におけるホットスポットの位置は得られない。ホットスポットの位置に合わせて再配置処理を実行するには、当該時間帯における各仮想領域へのアクセス数を一定時間観測する必要がある。

[0029] そこで、管理システムは、再配置時間帯計算処理において計算し、記録した仮想ボリュームの再配置時間帯の先頭時刻から一定期間後、例えば３０分後に、当該仮想ボリューム内の仮想領域に対する実領域の割当状況を見直す。例えば、管理システムは、当該仮想ボリューム内の各仮想領域について、

仮想領域へのアクセス数が、当該仮想領域に割り当てられている実領域を有する階層に対応したアクセス数範囲を超えているか否かを判断する。管理システムは上記の判断の結果、範囲外であると判断した仮想領域に対して、当該仮想領域に割り当てられている実領域（以下、移行元の実領域）に代えて、当該仮想領域へのアクセス数が収まるアクセス数範囲に対応する階層内の実領域（以下、移行先の実領域）を割り当てる。具体的には、管理システムはストレージサブシステムに、移行元の実領域のデータを移行先の実領域に移行するよう指示する。

[0030] <計算機システムの構成>

図1は、本発明の実施形態に係わる計算機システム10の概略構成を示す図である。計算機システム10においては、ホスト計算機101とストレージサブシステム102とが、ストレージエリアネットワーク105によって接続されている。また、管理システム107とストレージサブシステム102とが、管理用ネットワーク106によって接続されている。管理システム107とホスト計算機101が管理用ネットワーク106を介して接続されていても良い。管理システム107は、管理計算機103とWebブラウザ用計算機104を含んでいる。

[0031] なお、管理計算機103とWebブラウザ用計算機104とは、同じ計算機でもよく、また、それぞれ複数の計算機で同等の機能を実現していてもよい。つまり、計算機システム10を管理し、本発明の表示用情報を表示する1つ以上の計算機の集合を管理システムと呼ぶことがある。管理計算機103が表示用情報を表示する場合は管理計算機103が管理システムであり、また、図1に示されるように、管理計算機103と表示用計算機（Webブラウザ用計算機）104の組み合わせも管理システム107である。また、管理処理の高速化や高信頼化のために複数の計算機で管理計算機と同等の処理を実現してもよく、この場合は当該複数の計算機（表示を表示用計算機が行う場合は表示用計算機も含め）が管理システムである。

[0032] また、ストレージエリアネットワーク105と管理用ネットワーク106

とは、同じネットワークであってもよい。

[0033] ホスト計算機 101 は、ストレージサブシステム 102 にアクセス要求を送信する。ストレージサブシステム 102 は、ホスト計算機 101 からストレージエリアネットワーク 105 を介してアクセス要求を受信し、そのアクセス要求を処理する。

[0034] 管理システム 107 は、ストレージサブシステム 102 に制御要求（例えば、再配置処理の実行要求）を送信する。ストレージサブシステム 102 は、管理システム 107 から管理ネットワーク 106 を介して制御要求を受信し、制御要求に従う処理を行う。

[0035] なお、計算機システム 10 は、複数のホスト計算機 101 を含んでもよい。同様に、計算機システム 10 は、複数のストレージサブシステム 102 を含んでもよい。

[0036] <ホスト計算機の構成>

図 2 は、ホスト計算機 101 の内部構成を示すブロック図である。ホスト計算機 101 は、ストレージエリアネットワーク 105 との接続を実現するための SAN ポート 203 と、メモリ 202 と、メモリ 202 および SAN ポート 203 に接続された CPU 201 と、を有する。

[0037] SAN ポート 203 は、ストレージエリアネットワーク 105 に接続されている。

[0038] メモリ 202 には、1 つ以上のアプリケーションプログラム 204 が格納されている。各アプリケーションプログラム 204 は、CPU 201 によって実行され、SAN ポート 203 およびストレージエリアネットワーク 105 を介して、ストレージサブシステム 102 にアクセス要求を送信する。

[0039] <ストレージサブシステムの構成>

図 3 は、ストレージサブシステム 102 の内部構成を示すブロック図である。ストレージサブシステム 102 は、性能の異なる複数の物理記憶デバイス 305 と、コントローラ 312 を有し、これらは内部バス等の回路を介して相互に接続されている。同種の物理記憶デバイス 305 が複数ある場合、

それらはRAID構成を組んでいてもよい。RAID構成を組んでいる複数の物理記憶デバイス305は、RAIDグループ306を構成する。

[0040] コントローラ312は、ストレージエリアネットワーク105に接続するためのSANポート302と、管理ネットワーク106に接続するためのLANポート303と、メモリ304と、それらに接続されたCPU301とを有する。

[0041] SANポート302は、ストレージエリアネットワーク105に接続されるポートである。SANポート302は、ホスト計算機101からのアクセス要求を受信する。

[0042] LANポート303は、管理ネットワーク106に接続されるポートである。LANポート303は、管理システム107からの制御要求を受信する。

[0043] メモリ304には、ストレージ制御プログラム307と、RAIDグループ管理表308と、実領域管理表309と、階層管理表310と、VVOL管理表311が格納されている。ストレージ制御プログラム307は、CPU301で実行されることにより、アクセス制御処理および再配置処理を行う。

[0044] <RAIDグループ管理表308>

図4は、ストレージサブシステム102が有するRAIDグループ管理表308の構成を示す図である。RAIDグループ管理表308は、RAIDグループ306の性能に関する情報を有する。

[0045] 例えば、RAIDグループ管理表は、RAIDグループの識別子を表すRAIDグループID401と、RAIDグループを構成する物理記憶デバイスの種類を表すデバイス種別402と、RAIDグループのRAIDレベルおよびコンビネーションを表すRAIDレベル403と、RAIDグループを構成する物理記憶デバイスの識別子を表すPDEV_ID404と、を構成項目として含んでいる。なお、情報401~404のうちの少なくとも1つに代えてまたは加えて、他種の情報が含まれていてもよい。

[0046] <実領域管理表 309>

図5は、ストレージサブシステム102が有する実領域管理表309の構成を示す図である。実領域管理表309は、各RAIDグループ306が持つ実領域が、仮想ボリュームに割当済か否かの情報を有する。

[0047] 例えば、実領域管理表309は、実領域を持つRAIDグループの識別子を表すRAIDグループID501と、実領域の識別子を表す実領域ID502と、実領域に対応するRAIDグループのLBA範囲を表すRG_LBA範囲503と、実領域が割当済か否かを表す割当状況504と、を構成項目として含んでいる。

[0048] <VVOL（仮想ボリューム）管理表 311>

図6は、ストレージサブシステム102が有するVVOL管理表311の構成を示す図である。VVOL管理表311は、仮想ボリューム内の各仮想領域と、その仮想領域に割り当てられている実領域に関する情報を有する。

[0049] 例えば、VVOL管理表311は、仮想ボリュームの識別子を表すVVOLL_ID601と、仮想ボリューム内の仮想領域に対応するLBA範囲を表すVVOLL_LBA範囲602と、仮想ボリューム内の仮想領域に割り当てられている実領域を表す実領域ID603と、仮想ボリューム内の仮想領域に対する、ホスト計算機からのアクセス数（累計のI/O数）を表すアクセス数604と、再配置処理において仮想ボリューム内の仮想領域にその階層内の実領域を割り当てべきであると判定された、当該階層の識別子を表す再配置判定結果605と、を構成項目として含んでいる。

[0050] 上記VVOL管理表311において、VVOLL_ID601は、ホスト計算機101から指定される識別子ではなく、ストレージサブシステム102の内部で認識される識別子である。アクセス数604は、当該仮想領域に対するアクセス回数の値を持っており、ストレージサブシステム102は一定期間、例えば24時間ごとに、アクセス数604の値を0にリセットする。また、再配置判定結果605には、例えば最近行われた再配置処理における判定結果が格納される。

[0051] <階層定義表 310>

図7は、ストレージサブシステム102が有する階層定義表310の構成を示す図である。階層定義表310は、階層の性能及び許容するアクセス数の範囲に関する情報を有する。

[0052] 例えば、階層定義表310は、階層毎に、階層の識別子を表す階層ID701と、階層の性能要件を表す性能要件702と、階層に対応した、単位時間当たりの許容アクセス数の範囲を表す許容IOPS範囲703と、を構成項目として含んでいる。階層定義表310において、階層の性能要件702は、例えば、物理記憶デバイスの種別と、RAIDグループのRAIDレベルとで定義される。

[0053] 階層IDで特定される各階層701は、性能要件702を満たす物理記憶デバイスまたは／およびRAIDグループに属する実領域を要素に持っている。なお、階層定義表310は、例えばストレージ管理者からの要求に応じて更新される。

[0054] <管理計算機の構成>

図8は、管理計算機103の内部構成を示すブロック図である。管理計算機103は、管理ネットワーク106に接続するためのLANポート802と、構成管理プログラム804を実行するCPU801と、CPU801が用いるメモリ803とを有し、これらは内部バス等の回路を介して相互に接続されている。

[0055] メモリ803は、構成管理プログラム804と、アクセス分布変化履歴表805と、再配置時間帯管理表806と、管理側RAIDグループ管理表807と、管理側実領域管理表808と、管理側階層管理表809と、管理側VVOL管理表810と、を格納している。

[0056] 管理表807～810は、ストレージサブシステム102の管理表308～311に対応しており、管理計算機103がストレージサブシステム102から構成情報を取得した際に、その情報を格納するために用いられる。

[0057] 構成管理プログラム804は、CPU801によって実行されることで、

以下の処理を行う。即ち、構成管理プログラム 804 は、定期的または繰り返しストレージサブシステム 102 から各表 308～311 が有する情報を取得し、その情報を基に、アクセス分布変化履歴表 805 を更新する。また、構成管理プログラム 804 は、定期的または繰り返し、アクセス分布変化履歴表 805 を基に、再配置時間帯計算処理を行う。そして、構成管理プログラム 804 は、その計算した結果を、ユーザ（例えば、管理者）に提示し、ユーザからの承認を受けた時に、あるいはユーザに提示することなく、再配置時間帯管理表 806 に反映する。なお、上記情報取得及び再配置時間帯計算処理は、必ずしも等時間間隔で行われる必要はなく、繰り返し行われれば処理時間間隔に多少のずれがあっても良い。さらに、構成管理プログラム 804 は、各仮想ボリューム毎に、再配置時間帯管理表 806 に記載された、当該仮想ボリュームに対応する再配置時間帯に入るたび、当該仮想ボリュームの再配置処理を実行する。

[0058] <アクセス分布変化履歴表 805>

図 9 は、管理計算機 103 が有するアクセス分布変化履歴表 805 の構成を示す図である。アクセス分布変化履歴表 805 は、仮想ボリューム毎に、アクセス分布の変化量の時系列情報を有する。

[0059] 例えば、アクセス分布変化履歴表 805 は、仮想ボリュームの識別子を表す VVOL_ID 901 と、仮想ボリュームのアクセス分布の変化量を観測した時間帯を表す観測時間 902 と、仮想ボリュームのアクセス分布の変化量を表すアクセス分布変化量 903 と、仮想ボリュームに対するホスト計算機からのアクセス数の平均値を表すアクセス平均 904 と、を構成項目として含んでいる。なお、観測時間は、例えば日付と時間帯の組で表される。また、上記アクセス平均 904 の単位は、例えば IOPS（1 秒あたりのアクセス（I/O）の回数）である。アクセス分布変化量 903 の演算については図 17 を用いて後述する。

[0060] <配置時間帯管理表 806>

図 10 は、再配置時間帯管理表 806 の構成を示す図である。再配置時間

帯管理表 806 は、仮想ボリューム毎に、当該仮想ボリュームに対して再配置処理を実施する時間帯に関する情報を有する。

[0061] 例えば、再配置時間帯管理表 806 は、仮想ボリュームの識別子を表す VVOL_ID1001 と、仮想ボリュームに対して再配置処理を実施する時間帯を決定するための、仮想ボリュームのアクセス分布の変化量に対する閾値を表す分布変化閾値 1002 と、ホットスポットの動きが少ない時間帯であって、仮想ボリュームに対して再配置処理を実施する時間帯を表す再配置時間帯 1003 と、を構成項目として含んでいる。

[0062] <管理側 RAID グループ管理表 807>

図 11 は、管理計算機 103 が有する管理側 RAID グループ管理表 807 の構成を示す図である。管理側 RAID グループ管理表 807 は、ストレージサブシステム 102 が有する RAID グループ管理表 308 に対応し、RAID グループ管理表 308 が有する情報を格納するために用いられる。ただし、管理側 RAID グループ管理表 807 が有する情報は、RAID グループ管理表 308 が有する情報と完全に一致している必要はなく、RAID グループ管理表 308 の一部の情報を、管理側 RAID グループ管理表 807 に格納しなくてもよい。

[0063] 例えば、管理側 RAID グループ管理表 807 は、RAID グループの識別子を表す RAID グループ ID 1101 と、RAID グループを構成する物理記憶デバイスの種類を表すデバイス種別 1102 と、RAID グループの RAID レベルおよびコンビネーションを表す RAID レベル 1103 と、を構成項目として含んでいる。

[0064] なお、以降で説明する各表 808～810 についても、それらが有する情報は、ストレージサブシステム 102 の対応する表 309～311 が有する情報と完全に一致している必要はない。

[0065] <管理側実領域管理表 808>

図 12 は、管理側計算機 103 が有する管理側実領域管理表 808 の構成を示す図である。管理側実領域管理表 808 は、ストレージサブシステム 1

０２が有する実領域管理表３０９に対応する。

[0066] 例えば、管理側実領域管理表８０８は、実領域を持つＲＡＩＤグループの識別子を表すＲＡＩＤグループＩＤ１２０１と、実領域の識別子を表す実領域ＩＤ１２０２と、実領域に対応するＲＡＩＤグループのＬＢＡ範囲を表すＲＧ_ＬＢＡ範囲１２０３と、実領域が割当済か否かを表す割当状況１２０４と、を構成項目として含んでいる。

[0067] <管理側ＶＶＯＬ管理表８１０>

図１３は、管理側計算機１０３が有する管理側ＶＶＯＬ管理表８１０の構成を示す図である。管理側ＶＶＯＬ管理表８１０は、ストレージサブシステム１０２が有するＶＶＯＬ管理表３１１に対応する。

[0068] 例えば、管理側ＶＶＯＬ管理表８１０は、仮想ボリュームの識別子を表すＶＶＯＬ_ＩＤ１３０１と、仮想ボリューム内の仮想領域に対応するＬＢＡ範囲を表すＶＶＯＬ_ＬＢＡ範囲１３０２と、仮想ボリューム内の仮想領域に割り当てられている実領域を表す実領域ＩＤ１３０３と、仮想ボリューム内の仮想領域に対する、ホスト計算機からのアクセス数を表すアクセス数１３０４と、再配置処理において、仮想ボリューム内の仮想領域にその階層内の実領域を割り当てるべきであると判定された、当該階層の識別子を表す再配置判定結果１３０５と、を構成項目として含んでいる。

[0069] なお、ＶＶＯＬ管理表３１１のアクセス数６０４には、当該仮想領域に対するアクセス回数を記録するのに対して、管理側ＶＶＯＬ管理表８１０のアクセス数１３０４には、管理計算機１０３が行う各処理において使用する、アクセス数に関連する値が記録される。例えば、アクセス数１３０４には、アクセス分布変化量計算処理にて使用する、前回アクセス分布変化量計算処理を行った時のアクセス数６０４の値（前回処理時のアクセス数）１３０６と、前回処理時に計算した当該仮想領域のアクセス数の平均値（前回処理時の平均値）１３０７が記録される。さらに、アクセス数１３０４には、例えば再配置処理にて使用する、再配置時間帯の先頭時刻におけるアクセス数６０４の値（先頭時刻でのアクセス数）１３０８と、再配置処理にて計算した

アクセス数の平均値（平均値）１３０９が記録される。この場合、アクセス数の平均値を表す１３０７および１３０９の単位は、例えばＩＯＰＳである。

[0070] また、例えば、カラム１３０７に記録する平均値を計算するための計測期間と、カラム１３０９に記録する平均値を計算するための計測期間が等しい場合には、カラム１３０６と１３０８、およびカラム１３０７と１３０９はそれぞれ片方のみであってもよい。さらに、アクセス分布変化量計算処理にて使う値として、カラム１３０６または／および１３０７に代えて、或いは加えて、他の値を記録するようにしてもよいし、再配置処理にて使う値として、カラム１３０８または／および１３０９に代えて、或いは加えて、他の値を記録するようにしてもよい。

[0071] <管理側階層定義表８０９>

図１４は、管理側計算機１０３が具備する管理側階層定義表８０９の構成を示す図である。管理側階層定義表８０９は、ストレージサブシステム１０２が有する階層定義表３１０（図７）に対応する。図１４の情報は管理者が定義しても良いが、図７の情報と同期が取られている必要がある。

[0072] 例えば、階層定義表８０９は、階層毎に、階層の識別子を表す階層ＩＤ１４０１と、階層の性能要件を表す性能要件１４０２と、階層に対応した、単位時間当たりの許容アクセス数範囲を表す許容ＩＰＯＳ範囲１４０３と、を構成項目として含んでいる。

[0073] <計算機システム内における処理の内容>

本実施形態によるストレージサブシステム１０２が行う処理として、（ｉ）リード処理、（ii）ライト処理、および（iii）アクセス数計算処理がある。また、本実施形態による管理計算機１０３が行う処理として、（iv）アクセス分布変化量計算処理、（v）再配置時間帯計算処理、及び（vi）再配置処理がある。以下、これらについて詳細に説明する。

[0074] （Ａ）ストレージサブシステム１０２における処理

リード処理、ライト処理、及びそれらの処理終了時のアクセス数計算処理

に関し、ホスト計算機 101 は、ストレージサブシステム 102 に記録したデータにアクセスするために、ストレージサブシステム 102 にアクセス要求を送信する。そして、ストレージサブシステム 102 は、ホスト計算機 101 からアクセス要求を受信し、受信した要求に応じた処理を行う。以下で、アクセス要求の一例としてリード要求を受信した場合の処理と、ライト要求を受信した場合の処理をそれぞれ説明する。なお、(iii) アクセス数計算処理は (i) リード処理及び (ii) ライト処理のそれぞれの最後の処理として含まれている。

[0075] (i) リード処理

図 15 は、ストレージサブシステム 102 におけるリード処理を説明するためのフローチャートである。ホスト計算機 101 からリード要求が送信された場合、例えば下記の処理が行われる。

[0076] ストレージ制御プログラム 307 は、リード要求をホスト計算機 101 から受信する（ステップ 1501）。そして、ストレージ制御プログラム 307 は、リード要求が有するアクセス先情報を基に、データを読み出す元の仮想領域（以下、リード元仮想領域）を特定する（ステップ 1502）。

[0077] 次に、ストレージ制御プログラム 307 は、リード対象のデータがキャッシュメモリ領域に残っているか否かを判断する（ステップ 1503）。

[0078] 対象データがキャッシュにある場合（ステップ 1503 で Yes の場合）、ストレージ制御プログラム 307 は、キャッシュメモリ領域のリード対象データを、ホスト計算機 101 に送信する（ステップ 1504）。

[0079] 対象データがキャッシュにない場合（ステップ 1503 で No の場合）、ストレージ制御プログラム 307 は、VVOL 管理表 311 を基に、1502 で特定されたリード元仮想領域に割り当てられている実領域（以下、リード元実領域）を特定する（ステップ 1505）。続いて、ストレージ制御プログラム 307 は、リード元実領域からデータを読み出し、そのデータをキャッシュメモリ領域に書き込み、かつホスト計算機 101 に送信する（ステップ 1506）。

[0080] 最後に、ストレージ制御プログラム307は、リード元仮想領域に対応するアクセス数604の値を更新する（ステップ1507）。この時、ステップ1507の直前に実行した処理がステップ1504であった場合には、ストレージ制御プログラム307は、リード元仮想領域に対応するアクセス数604の値を更新しなくてもよい。

[0081] (ii) ライト処理

図16は、ストレージサブシステム102におけるライト処理を説明するためのフローチャートである。ホスト計算機101からライト要求が送信された場合、例えば下記の処理が行われる。

[0082] ストレージ制御プログラム307は、ライト要求をホスト計算機101から受信する（ステップ1601）。そして、ストレージ制御プログラム307は、ライト要求が有するアクセス先情報を基に、データを書き込む先の仮想領域（以下、ライト先仮想領域）を特定する（ステップ1602）。

[0083] 次に、ストレージ制御プログラム307は、ライト先仮想領域に実領域が割り当てられているか否かを判断する（ステップ1603）。具体的には、ライト先仮想領域がVVOL管理表311に登録されているか否かを判断する。

[0084] ライト先仮想領域に実領域が割り当てられている場合（ステップ1603でYesの場合）、ストレージ制御プログラム307は、ライト対象データをライト先仮想領域に割り当てられている実領域に書き込む（ステップ1604）。

[0085] ライト先仮想領域に実領域が割り当てられていない場合（ステップ1603でNoの場合）、ストレージ制御プログラム307は、ライト先仮想領域に割り当てられる未割当の実領域が存在するか否かを判断する（ステップ1605）。具体的には、ストレージ制御プログラム307は、実領域管理表309の割当状況504が「未割当」である実領域が存在するか否かを判断する。

[0086] ライト先仮想領域に未割当の実領域が存在する場合（ステップ1605で

Yesの場合)、ストレージ制御プログラム307は、ライト先仮想領域に未割当の実領域を割り当てて、ライト対象データを当該実領域に書き込む(ステップ1606)。

[0087] ライト先仮想領域に未割当の実領域が存在しない場合(ステップ1605でNoの場合)、ストレージ制御プログラム307は、ホスト計算機101にエラーを送信する(ステップ1607)。

[0088] 最後に、ストレージ制御プログラム307は、ライト元仮想領域に対応するアクセス数604の値を更新する(ステップ1608)。

[0089] (B)管理計算機103における処理

(iv) アクセス分布変化量計算処理

図17は、アクセス分布変化量計算処理の詳細を説明するためのフローチャートである。アクセス分布変化量計算処理は、一定の時間間隔、例えば1時間毎に定期的・繰返しに行われる。この時間間隔は、正確な間隔でなくてもよく、多少の時間のずれがあっても良い。

[0090] 構成管理プログラム804は、管理側VVOL管理表810に登録されている全ての仮想ボリュームに対して、以降の1701~1704の処理を行う。以下、1つの仮想ボリューム(以下、図17の説明において「対象VVOL」という)を例にして、以降の処理を説明する。

[0091] (ステップ1701の処理内容)

構成管理プログラム804は、対象VVOL内の各仮想領域に対応するアクセス数の情報を、ストレージサブシステム102から取得する。具体的には、構成管理プログラム804は、ストレージサブシステム102に対して、対象VVOL内の各仮想領域に対応するアクセス数の情報を要求するリクエストを送信する。リクエストは、例えば対象VVOLの識別子であるVVOL_IDの情報を有する。

[0092] ストレージサブシステム102はリクエストを受信し、リクエストが有するVVOL_IDで特定される仮想ボリューム内の各仮想領域について、VVOL管理表311の当該仮想領域に対応するVVOL_LBA範囲602の値

とアクセス数604の値を、例えば2つ組にして、管理計算機103に送信する。なお、この時ストレージサブシステム102が管理計算機103に送信する情報は、仮想領域の識別情報とアクセス数の情報の対応が分かれば、どのような形式であってもよい。

[0093] (ステップ1702の処理内容)

構成管理プログラム804は、前回処理時から現在までの期間における、対象VVOLに対するアクセス数の平均値(IOPS)を計算する。具体的には、対象VVOL内のすべての仮想領域に対して、ステップ1701で取得したアクセス数604の値と前回処理時のアクセス数1306の値との差分を求め、その差分値の総和を計算する。差分を求めるのは本実施形態ではアクセス数カウントにインクリメントカウンタを用いているからであり、差分の総和によって、当該期間におけるアクセス総数(累計値)が算出される。ただし、この時、前回処理時のアクセス数1306には、前回アクセス分布変化量計算処理を行った時に取得した、各仮想領域のアクセス数604の値が記録されているものとする。

[0094] 構成管理プログラム804は算出した総和の値を基に、前回処理時から現在までの期間における対象VVOLのアクセス数の単位時間平均値(例えばIOPS)を求める(以下、図17の説明において「VVOLアクセス平均」という)。

[0095] (ステップ1703の処理内容)

構成管理プログラム804は、対象VVOLに対するアクセス分布の、前回処理時から現在までの期間における変化量を計算する。具体的には、まず、構成管理プログラム804は、対象VVOL内のすべての仮想領域に対して、ステップ1701で取得したアクセス数604の値と前回処理時のアクセス数1306の値との差分を基に、前回処理時から現在までの期間における、当該仮想領域に対するアクセス数の単位時間平均(例えばIOPS)を求める(以下、図17の説明において「仮想領域アクセス平均」という)。

[0096] 次に構成管理プログラム804は、対象VVOL内のすべての仮想領域に

対して、当該仮想領域の仮想領域アクセス平均をVVOLアクセス平均で割った値と、当該仮想領域に対応する前回処理時の平均値1307を対象VVOLに対応するアクセス平均904の値で割った値との差分を計算する。ただし、この時前回処理時の平均値1307には、前回処理時に計算した仮想領域アクセス平均の値が記録されているものとする。ここで計算した差分の値は、当該仮想領域へのアクセス数の、対象VVOLへのアクセス数に対する割合が、前回処理時からどれだけ変化したかを表している。

[0097] 構成管理プログラム804は、例えば対象VVOL内のすべての仮想領域に対する上記差分の二乗合計の平方根を計算し、対象VVOLに対応するアクセス分布の変化量903として、アクセス分布変化履歴表805に記録する。この時、アクセス分布変化履歴表805の観測日時902には、前回のアクセス分布変化量計算処理を実行した時刻から、現在の時刻までの時間帯を記録する。また、アクセス分布変化履歴表805のアクセス平均904には、ステップ1702で計算したVVOLアクセス平均を記録する。

[0098] 上記処理にて計算したアクセス分布変化量903の値を分析することで、対象VVOL内のホットスポット（その仮想領域に対するアクセス数が、同じ仮想ボリューム内の他の仮想領域に対するアクセス数よりも高い仮想領域）の位置が、その時間帯において変化したことを検知することができる。

[0099] ここで、「その指標の分析により検知できる情報」という観点で、上記処理の計算値であるアクセス分布変化量と他の指標を比較する。例えば、対象VVOLへのアクセス数を指標に用いた場合、個々の仮想領域へのアクセス数が指標に現れないため、「対象VVOLへのアクセス数は変化していないが、対象VVOLのホットスポットの位置は変化した」ような場合を検知できない。ホットスポットの位置の変化を検知するには、対象VVOLのアクセス分布の変化に着目すればよい。そこで、対象VVOL内の各仮想領域に対するアクセス数の差分を取り、その統計値を指標とすることが考えられる。しかし、この値を用いると、ホットスポットの位置の変化以外にも、例えば「対象VVOL内の各仮想領域へのアクセス数が、平均的に増加した」よ

うな場合（以下、図１７の説明において「全体増加ケース」という）も検知してしまう。

[0100] 本発明は、個々の仮想ボリュームごとにその内部の各仮想領域に対する最適な実領域割り当てを実現するものであり、全体増加ケースに対する解決策ではないため、ホットスポットの位置が変化した場合と全体増加ケースを切り分ける必要がある。そこで、上記処理の計算値のような、対象VVOL内の各仮想領域に対するアクセス数を正規化した値の統計値を指標とする。正規化を行うことで、全体増加ケースにおいては指標が変化しないため検知せずに済む。これにより、ホットスポットの位置が変化したケースに的を絞って検知を行うことができる。

[0101] なお、前述の全体増加ケースに対する解決策としては、従来の定期的かつ仮想ボリューム横断（プール単位）で行う再配置処理が適切である。本発明と従来技術を組み合わせて用いることで、論理ボリュームよりも細かい領域単位での再配置処理を、より効果的に利用することができる。ただし、本発明と従来技術を組み合わせる場合には、再配置を実施する時間が衝突しないように、例えば再配置処理を行ってから一定期間（例えば１時間）以内には再配置処理を行わないようにしたり、本特許による再配置処理と従来技術による再配置処理の優先順を決定したりするなどの制御を行う必要がある。

[0102] また、上述のアクセス分布の変化量の計算方法は一例であり、他の計算方法を用いてもよい。

[0103] （ステップ１７０４の処理内容）

構成管理プログラム８０４は、対象VVOL内のすべての仮想領域に対して、ステップ１７０１で取得した各仮想領域に対するアクセス数６０４の値をカラム１３０６に記録し、ステップ１７０３で計算した各仮想領域に対する仮想領域アクセス平均の値をカラム１３０７に記録する。

[0104] なお、対象VVOL内の仮想領域の中に、ストレージサブシステム１０２のVVOL管理表３１１には登録されているが、管理側VVOL管理表８１０には登録されていないものがあった場合、構成管理プログラム８０４は、

上記ステップ 1701～1703 の処理において、当該仮想領域に対応するアクセス数に関する値 1306～1309 を、例えば全て 0 として計算する。また、構成管理プログラム 804 は、上記ステップ 1704 の処理において、当該仮想領域を管理側 VVOL 管理表 810 に登録した後、対応するコラム 1306 および 1307 の値を更新する。

[0105] (v) 再配置時間帯計算処理

図 18 は、再配置時間帯計算処理の詳細を説明するためのフローチャートである。再配置時間帯計算処理は、例えば定期的・繰り返し行われる。ただし、周期的に実行する必要はなく、実行時間間隔に多少のバラツキがあっても構わない。

[0106] (ステップ 1801 の処理内容)

管理計算機 103 内の構成管理プログラム 804 は、再配置時間帯管理表 806 における再配置時間帯 1003 の値をクリアする。

[0107] 次に、構成管理プログラム 804 は、管理側 VVOL 管理表 810 に登録されているすべての仮想ボリュームに対して、以降ステップ 1802～1805 の処理を行う。以下、1 つの仮想ボリューム（以下、図 18 の説明において「対象 VVOL」という）を例にして説明する。

[0108] 構成管理プログラム 804 は、例えば 1 日を 1 時間毎に区切った各時間帯（以下、対象時間帯）に対して、以降ステップ 1802～1804 の処理を行う。

[0109] (ステップ 1802 の処理内容)

構成管理プログラム 804 は、アクセス分布変化履歴表 805 を基に、対象 VVOL の対象時間帯におけるアクセス分布変化量 903 の平均値、例えば直近 1 週間のアクセス分布変化量 903 の平均値を計算する。なお、平均値を計算する期間は、例えば 2 日間や 1 ヶ月間でもよい。

[0110] また、ストレージ管理者が、平均値を計算する期間を設定できるようにしてもよい。その場合、構成管理プログラム 804 は、例えばストレージ管理者からの指示に応じて、平均値を計算する期間を設定する画面を Web ブラ

ウス用計算機 104 に表示する。さらに、平均値を計算する期間は、仮想ボリュームごとに異なってもよいし、複数の仮想ボリュームで共有してもよい。

[0111] (ステップ 1803 の処理内容)

構成管理プログラム 804 は、計算したアクセス分布変化量 903 の平均値（以下、分布変化平均という）が再配置時間帯管理表 806 の対象 VVOL に対応する分布変化閾値 1002 より小さいか否かを判定する。なお、分布変化閾値 1002 の値は、1803 の処理が実行される段階で設定してもよいし、それより前に設定されていてもよい。分布変化閾値 1002 の設定の詳細に関しては、後述する。

[0112] (ステップ 1804 の処理内容)

分布変化平均が分布変化閾値 1002 よりも小さい場合（ステップ 1803 で Yes の場合）、構成管理プログラム 804 は、対象時間帯を再配置時間帯管理表 806 に登録する。分布変化平均が分布変化閾値 1002 以上の場合（ステップ 1803 で No の場合）、構成管理プログラム 804 は、次の時間帯について処理（ステップ 1802 ～ 1804）を継続する。以上の処理を、各時間帯に対して繰り返し行う。

[0113] (ステップ 1805 の処理内容)

構成管理プログラム 804 は、再配置時間帯管理表 806 の対象 VVOL に対応する再配置時間帯 1003 に対して、連続する時間帯を 1 つにまとめるという処理を行う。具体的には、構成管理プログラム 804 は、対象 VVOL の VVOL_ID と、VVOL_ID が一致するレコードを、再配置時間帯管理表 806 から特定する。2 以上のレコードが特定された場合、構成管理プログラム 804 は特定された 2 以上のレコードの、2 以上の再配置時間帯が連続していれば、それらを 1 つにまとめる。例えば、再配置時間帯が 9:00-10:00、10:00-11:00、11:00-12:00 のように連続していれば、9:00-10:00、10:00-11:00、11:00-12:00 を、1 つの時間帯 9:00-12:00 にまとめる。構成管理プログラム 804 は、再配置時間帯をまとめられた 2 以上のレコードをす

べて削除し、まとめた後の時間帯を登録する。

[0114] (上述の再配置時間帯計算処理の変形例について)

以上が再配置時間帯計算処理の説明であるが、構成管理プログラム 804 は、再配置時間帯計算処理において、対象時間帯を再配置時間帯管理表 806 に登録するか否かの判定時に、アクセス分布変化量に加えて対象 VVOL の対象時間帯におけるアクセス数の情報を考慮してもよい。具体的には、ステップ 1802 において、構成管理プログラム 804 はアクセス分布変化履歴表 805 を基に、対象 VVOL の対象時間帯におけるアクセス数 904 の平均値を計算する。

[0115] ステップ 1803 において、構成管理プログラム 804 は、例えば、分布変化平均が分布変化閾値 1002 よりも小さくかつアクセス数 904 の平均値が一定値（例えば、分布変化閾値 1002 とは別に設定または／および計算した閾値）よりも小さい場合にのみ、対象時間帯を再配置時間帯管理表 806 に登録すると判定してもよい。

[0116] (分布変化閾値について)

ステップ 1803 で使用する分布変化閾値 1002 は、ストレージ管理者が入力してもよく、管理計算機 103 が計算しても良い。ここでは一例として、ストレージ管理者が入力する形態について説明する。

[0117] 構成管理プログラム 804 は、上記の再配置時間帯計算処理において、Web ブラウズ用計算機 104 に、図 19 に示す分布変化閾値入力画面 1901 を表示する。画面 1901 は、分布変化閾値の設定対象である仮想ボリュームを選択するための領域（仮想ボリューム選択領域）1904 と、選択された仮想ボリュームに設定する分布変化閾値を入力するための領域（閾値入力領域）1906 を有する。

[0118] 仮想ボリューム選択領域 1904 には、例えば再配置時間帯管理表 806 に登録されている各仮想ボリュームに対応する VVOL_ID 1001 の一覧が表示される。ストレージ管理者は仮想ボリュームの ID の一覧から、例えば 1 つを選択可能である。

- [0119] 閾値入力領域 1906 には、例えば分布変化閾値を入力するためのテキストボックスが表示される。テキストボックスには、選択された ID で識別される仮想ボリュームに対応する分布変化閾値の入力が可能である。
- [0120] ストレージ管理者が、画面 1901 を用いて分布変化閾値を入力し、適用ボタン 1902 を押すと、構成管理プログラム 804 は入力された値を、画面 1901 で選択されている仮想ボリュームの ID に対応する分布変化閾値 1002 に格納し、上記再配置時間帯計算処理の 1803 以降を実行することになる。
- [0121] なお、画面 1901 には、ストレージ管理者が分布変化閾値を決定するための判断材料となる情報を表示する領域（管理者判断材料表示領域）1905 があっても良い。例えば、画面 1901 には、上述の再配置時間帯計算処理のステップ 1802 で計算した、1 日を 1 時間に区切った時間帯毎のアクセス分布変化量の平均値を表示しても良い。表示の方法は、例えば図 19 の管理者判断材料表示領域 1905 のような表形式でもよく、図 20 のような時系列のグラフ形式で表示してもよい。図 20 のようにグラフで表示した場合には、管理者は視覚でアクセス分布の変化を捉えることができ、感覚的な理解が得られやすいという効果がある。
- [0122] また、構成管理プログラム 804 は、分布変化閾値が入力された時に、入力された閾値を用いて計算される再配置時間帯を、再配置時間帯管理表 806 に登録する前に画面 1901 に表示し、それらの時間帯を再配置時間帯管理表に登録するか否かを、ストレージ管理者に選択させてもよい。この場合、画面 1901 は計算した再配置時間帯を表示するための領域（再配置時間帯表示領域）1907 を持つ。具体的には、構成管理プログラム 804 は、上記再配置時間帯計算処理の開始時に、メモリ 803 上に中間表を準備する。ストレージ管理者は画面 1901 において、分布変化閾値を閾値入力領域 1906 に入力した後、例えば適用ボタン 1902 の代わりに時間帯計算ボタン 1908 を押す。
- [0123] 時間帯計算ボタン 1908 が押されると、構成管理プログラム 804 は、

選択されている仮想ボリュームに対応する分布変化閾値 1002 の代わりに、閾値入力領域 1906 に入力されている値を用いて、ステップ 1803 ~ 1805 の処理を行う。この場合、構成管理プログラム 804 は、ステップ 1804 および 1805 で計算した再配置時間帯を、再配置時間帯管理表 806 ではなく中間表に記録する。構成管理プログラム 804 は、中間表に記録した再配置時間帯の情報を画面 1901 の再配置時間帯表示領域 1907 に表示する。ストレージ管理者は分布変化閾値を入力し直し、再度時間帯計算ボタン 1908 を押してもよい。

[0124] 構成管理プログラム 804 は、ストレージ管理者が分布変化閾値を入力し直すたびに、新しく入力された閾値を用いて再配置時間帯を計算し、中間表に記録して再配置時間帯表示領域 1907 に表示する。ストレージ管理者が適用ボタン 1902 を押下すると、構成管理プログラム 804 は、閾値入力領域 1906 に入力された値を分布変化閾値 1002 に格納し、また中間表に記録されている再配置時間帯の情報を再配置時間帯管理表 806 に登録し、中間表をメモリ 803 から削除して再配置時間帯計算処理を終了する。

[0125] さらに、ストレージ管理者が構成管理プログラム 804 に指示することで、再配置時間帯計算処理を実行する前に、分布変化閾値を入力しても良い。その場合、再配置時間帯計算処理の間では、ストレージ管理者に分布変化閾値の入力を求めなくても良い。ストレージ管理者が構成管理プログラム 804 に、分布変化閾値の入力を指示した場合、構成管理プログラム 804 は Web ブラウズ用計算機 104 に画面 1901 を表示する。

[0126] また、ストレージ管理者が構成管理プログラム 804 に指示することで、分布変化閾値を入力するのではなく、アクセス分布に関する情報を閲覧するための画面を表示できてもよい。この場合構成管理プログラム 804 は、例えば画面 1901 に仮想ボリューム選択領域 1904 と管理者判断材料表示領域 1905 のみを表示する。

[0127] さらに、構成管理プログラム 804 は、分布変化閾値を自動的に計算しても良い。その場合、ストレージ管理者に分布変化閾値の入力を求めなくても

良い。例えば、ストレージ管理者は分布変化閾値の代わりに、１日の間で行う再配置処理の上限数を入力する。構成管理プログラム８０４は、入力された上限数を超えない数の再配置時間帯が登録されるような分布変化閾値を計算する。分布変化閾値を、構成管理プログラムに計算させるか、自分で入力するかは、ストレージ管理者が選択できて良い。

[0128] (vi) 再配置処理

図２１は、再配置処理の詳細を説明するためのフローチャートである。再配置処理は、再配置時間帯管理表８０６に再配置時間帯１００３が登録されている仮想ボリューム（以下、図２１の説明において対象ＶＶＯＬ）に対して、対象ＶＶＯＬに対応する再配置時間帯１００３毎に行う処理である。再配置処理は、例えば、上述の（v）再配置時間帯計算処理によって、再配置時間帯管理表８０６に登録された各再配置時間帯の先頭時刻から一定期間後、例えば３０分後に行われる。

[0129] (ステップ２１００の処理内容)

管理計算機１０３の構成管理プログラム８０４は、ストレージサブシステム１０２から対象ＶＶＯＬ内のすべての仮想領域に対するアクセス数６０４の値を取得し、管理側ＶＶＯＬ管理表８１０の先頭時刻でのアクセス数１３０８に記録する。

[0130] 構成管理プログラム８０４はステップ２１００を実行した後、所定時間（例えば３０分）が経過するまで待機する。なお、構成管理プログラム８０４は、待機中に、他の処理を行ってもよい。即ち、ステップ２１００を行った後の所定時間経過後にステップ２１０１の処理を行うことができれば、当該待機時間内にどのような処理を何回行っても構わない。

[0131] (ステップ２１０１の処理内容)

構成管理プログラム８０４は、ストレージサブシステム１０２から対象ＶＶＯＬ内のすべての仮想領域に対するアクセス数６０４の値を取得し、各仮想領域の例えば過去３０分間におけるアクセス数の平均値を計算して平均値１３０９に記録する。具体的には、構成管理プログラム８０４は、次の（式

1) にしたがってアクセス数の平均値を計算し、平均値 1309 に記録する。

[0132] (アクセス数 604 - 先頭時刻でのアクセス数 1308) / 計測期間の秒数 (式 1)

例えば、VVOL_ID が Volume 1 であり、VVOL_LBA 範囲が 0 - 999 である仮想領域に対する図 6 および図 13 の値を例に取り、また計測期間を 30 分間として上記 (式 1) に当てはめて計算すると、 $(792000 - 621000) / (30 \times 60) = 95$ となる。

[0133] このように、再配置時間帯の先頭時刻から一定期間 (例えば 30 分間) のアクセス数の情報を基に再配置処理を行うことで、当該仮想ボリュームの、アクセス分布が変化しない時間帯のアクセス数の情報に基づいた再配置処理を行うことができる。

[0134] また、構成管理プログラム 804 は、ストレージサブシステム 102 から各表 308 ~ 310 の情報を取得し、取得した情報を基に管理計算機 103 の各表 807 ~ 809 の情報を更新する。

[0135] (ステップ 2102 の処理内容)

次に、構成管理プログラム 804 は、管理側 VVOL 管理表 810 に登録されている、対象 VVOL 内のすべての仮想領域に対して、対象仮想領域に割り当てべき階層を特定する。具体的には、構成管理プログラム 804 は、対象仮想領域に対応するアクセス数の平均値 1309 の値を含むアクセス数範囲を表す許容 IOPS 範囲 1403 を管理側階層定義表 809 から特定し、その許容アクセス数範囲 1403 に対応する階層 ID 1401 を、対象仮想領域の再配置判定結果 1305 の値として記録する。この結果、記録された階層 ID は、当該仮想領域に割り当てられている実領域を含む階層の階層 ID と同じであることもあれば異なることもある。なお、アクセス数 1304 に、再配置処理にて再配置判定結果の階層を特定するために用いる値として複数の種類の値を記録している場合には、許容 IOPS 範囲 1403 を特定するために用いる値として、どれか 1 つを選択してもよいし、計算によ

り導出してもよい。また、許容 I O P S 範囲 1 4 0 3 を特定する際に複数の値の組を用いてもよい。その場合、階層定義表 3 1 0 および管理側階層定義表 8 0 9 の許容 I O P S 範囲 7 0 3 および 1 4 0 3 には、複数の値の組を設定する。

[0136] (繰り返し処理)

次に、構成管理プログラム 8 0 4 は、管理側 V V O L 管理表 8 1 0 に登録されている、対象 V V O L 内のすべての仮想領域に対してステップ 2 1 0 3 ~ 2 1 0 8 の処理を実行する。ただし、ここで構成管理プログラム 8 0 4 は、例えば、各仮想領域を、対応するアクセス数の平均値 1 3 0 9 の値が高い順に処理する。以下、1つの仮想領域(「対象仮想領域」という)を例にして、以降のステップ 2 1 0 3 ~ 2 1 0 8 の処理について説明する。

[0137] (ステップ 2 1 0 3 の処理内容)

構成管理プログラム 8 0 4 は、現在対象仮想領域に割り当てられている実領域に対応する階層(「対象領域配置元階層」という)の I D と対象仮想領域に対応する再配置判定結果 1 3 0 5 の値とが一致するか否かを判定する。具体的には、構成管理プログラム 8 0 4 は、対象仮想領域に割り当てられている実領域の I D 1 3 0 3 を基に、その実領域を含む R A I D グループの I D 1 2 0 1 を特定し、その R A I D グループに対応するデバイス種別 1 1 0 2 または / および R A I D レベル 1 1 0 3 が適合する性能要件 1 4 0 2 に対応する階層 I D 1 4 0 1 を特定する。特定した階層 I D 1 4 0 1 と、再配置判定結果 1 3 0 5 の値が一致するか否かを判定する。

[0138] 対象領域配置元階層が配置されるべき階層の場合(ステップ 2 1 0 3 で Y e s の場合)、対象の仮想領域のデータ移行は不要のため、対象の仮想領域に対する処理は終了する。一方、対象領域配置元階層が配置されるべき階層でない場合(ステップ 2 1 0 3 で N o の場合)、処理はステップ 2 1 0 4 に移行する。

[0139] (ステップ 2 1 0 4 の処理内容)

構成管理プログラム 8 0 4 は、管理側実領域管理表 8 0 8 を基に対象仮想

領域に対応する再配置判定結果 1305 の階層 ID によって示される階層（「対象領域配置先階層」という）内に、未割当の実領域が存在するか否かを判定する。

[0140] 配置されるべき階層に空きがある場合（ステップ 2104 で Yes の場合）、処理はステップ 2106 に移行し、配置されるべき階層に空きがない場合（ステップ 2104 で No の場合）、処理はステップ 2105 に移行する。

[0141] （ステップ 2106 の処理内容）

構成管理プログラム 804 は、対象仮想領域に現在割り当てられている実領域（「データ移行元実領域」という）に代えて、対象領域配置先階層内の未割当の実領域（「データ移行先実領域」という）を割り当てる。具体的には、構成管理プログラム 804 は、管理側実領域管理表 808 の、データ移行元実領域に対応する割り当て状況 1204 の値を未割当に更新し、データ移行先実領域に対応する割り当て状況 1204 の値を割当済みに更新する。なお、本実施形態では、例えば、実領域と仮想領域は同じ容量で区切られるように管理され、移行元と移行先の容量を考慮する必要がないようになっている。

[0142] また、構成管理プログラム 804 は、管理側 VVOL 管理表 810 の、対象仮想領域に対応する実領域 ID 1303 の値を、データ移行先実領域の ID に更新する。また、構成管理プログラム 804 は、データ移行元実領域からデータ移行先実領域にデータを移行するようにストレージサブシステム 102 に指示する。ストレージサブシステム 102 のストレージ制御プログラム 307 は、構成管理プログラム 804 からの指示を受信し、データ移行元実領域からデータ移行先実領域にデータを移行する。

[0143] なお、ステップ 2106 の処理を行うと、他の仮想ボリュームの性能に影響を及ぼす場合がある。例えば、ある仮想ボリュームの再配置処理において、ステップ 2106 の処理を行った結果、性能が高い階層内の未割当実領域を全て使い切ってしまう場合がある。この場合、この後に再配置処理を行う

仮想ボリューム内の仮想領域には、当該階層内の未割当実領域を割り当てることができない。したがって、当該階層内の実領域を割り当てべきと判定された仮想領域が存在する場合、後述するステップ 2107 および 2108 の処理のように、当該階層内の実領域が割り当てられている他の仮想領域と割り当て状況およびデータを入れ替えたり、性能に近い他の階層の実領域を割り当てたりすることとなり、性能に影響を及ぼす場合がある。本事象に対処するためには、例えば、再配置処理において、対象 VVOL の仮想領域に未割当の実領域を割り当てることは行わず、常に対象 VVOL 内の他の仮想領域との入れ替えのみによって、対象 VVOL 内仮想領域への実領域割り当て状況の最適化を実現する方法が考えられる。再配置処理において未割当の実領域を使用するか否かは、例えば、各階層内の未割当の実領域の個数または／および割合によって、構成管理プログラム 804 が判断する。具体的には、構成管理プログラム 804 は、ステップ 2104 の処理において、対象領域配置先階層内に未割当実領域が存在したとしても、その個数が対象領域配置先階層内の全実領域数の 1 割よりも少ない場合は、ステップ 2106 の処理ではなくステップ 2105 の処理に移行する。

[0144] (ステップ 2105 の処理内容)

構成管理プログラム 804 は、対象仮想領域に割り当てられている実領域と、データを入れ替え可能な実領域が存在するか否かを判定する。具体的には、構成管理プログラム 804 は、対象領域配置先階層内の割当済の実領域の中で、その実領域が割り当てられている仮想領域に対応する再配置判定結果 1305 の階層 ID が、対象領域配置元階層の階層 ID と一致するものが存在するか否かを判定する。

[0145] 入れ替え可能な実領域が存在する場合（ステップ 2105 で Yes の場合）、処理はステップ 2107 に移行し、入れ替え可能な実領域が存在しない場合（ステップ 2105 で No の場合）、処理はステップ 2108 に移行する。

[0146] (ステップ 2107 の処理内容)

構成管理プログラム 804 は、対象仮想領域に割り当てられている実領域（以下、ステップ 2107 の説明において「対象実領域 1」という）と、対象領域配置先階層が有する割当済実領域（以下、ステップ 2107 の説明において「対象実領域 2」という）との間で、仮想領域への割り当て状況を入れ替える。具体的には、構成管理プログラム 804 は、管理側 VVOL 管理表 810 の実領域 ID 1303 において、対象実領域 1 の ID が格納されている箇所に対象実領域 2 の ID を格納し、対象実領域 2 の ID が格納されている箇所に対象実領域 1 の ID を格納する。また、構成管理プログラム 804 は、対象実領域 1 と対象実領域 2 の間のデータの入れ替えを行うよう、ストレージサブシステム 102 に指示する。ストレージサブシステム 102 のストレージ制御プログラム 307 は、構成管理プログラム 804 からの指示を受信し、実領域間のデータの入れ替えを行う。

[0147] 例えば、ストレージ制御プログラム 307 は、下記の処理を行うことによりデータの入れ替えを実現する。なお、下記処理におけるキャッシュメモリ領域は、ストレージサブシステム 102 が有する未割当実領域でもよい。

[0148] 手順 1：ストレージ制御プログラム 307 は、対象実領域 1 内のデータをキャッシュメモリ領域に書く。

[0149] 手順 2：ストレージ制御プログラム 307 は、対象実領域 2 内のデータをキャッシュメモリ領域に書く。

[0150] 手順 3：ストレージ制御プログラム 307 は、対象実領域 1 のデータをキャッシュメモリ領域から対象実領域 2 に書き込む。

[0151] 手順 4：ストレージ制御プログラム 307 は、対象実領域 2 のデータをキャッシュメモリ領域から対象実領域 1 に書き込む。

[0152] （ステップ 2108 の処理内容）

構成管理プログラム 804 は、対象階層に性能が最も近い階層内の未割当の実領域に、対象仮想領域に割り当てられている実領域内のデータを移行する。また、構成管理プログラム 804 は、管理側 VVOL 管理表 810 と管理側実領域管理表 808 を更新する。

[0153] (ステップ2109の処理内容)

構成管理プログラム804は、ストレージサブシステム102に、上記の再配置処理によって更新された各表の情報を取得し、ストレージサブシステム102の各表の情報を更新するよう指示する。

[0154] ストレージサブシステム102のストレージ制御プログラム307は、構成管理プログラム804からの指示を受信し、管理計算機103の各表807～809の情報を取得し、取得した情報を基にストレージサブシステム102の各表308～310の情報を更新する。

[0155] (2) 第2の実施形態

第2の実施形態では、仮想領域に割り当てられた実領域にゼロデータ（すべてのビットがゼロであるデータ）が格納された場合に、ストレージサブシステムが当該仮想領域を未割当の仮想領域に変更する（以下で、このような仮想領域を「ゼロデータ仮想領域」という）。より具体的には、実領域の全てのデータがゼロとなった場合、対応する仮想領域に対してNULL_P o i n t e r（何も割り当てられていないことを示すポインタ）が付与され、当該仮想領域が未割当の仮想領域となる。ストレージサブシステムは、ホスト計算機から未割当の仮想領域に対するリード要求を受信した場合、エラーではなくゼロデータを送信する。このような技術を「ゼロリクレーション」という。図24は、このゼロインクレーションの概念を示す図である。

[0156] この場合、ストレージサブシステムは未割当の仮想領域の内、ゼロデータ仮想領域に対しては例えば割当済仮想領域と同様にアクセス数を計測する。

[0157] <システム及び装置の構成、並びに、管理情報について>

第2の実施形態において、計算機システム、ホスト装置、ストレージサブシステム、及び管理システムの構成は、第1の実施形態のそれらと同様であるが、第1の実施形態の構成と比較して、VVOL管理表311および管理側VVOL管理表810が有する情報が異なる。以下では、第1の実施形態と異なる部分のみ説明する。

[0158] 図22は、第2の実施形態におけるVVOL管理表311の構成を示す図

である。図 23 は、第 2 の実施形態における管理側 VVOL 管理表 810 の構成を示す図である。

[0159] 第 2 の実施形態の VVOL 管理表 311 および管理側 VVOL 管理表 810 は、実領域 ID 2203 および 2303、および再配置判定結果 2205 および 2305 に対して、値が存在しないことを示す NULL 値（ゼロ値）を格納することにより、未割当の仮想領域に対応する情報を格納することができる点で、第 1 の実施形態の VVOL 管理表 311 および管理側 VVOL 管理表 810 と異なる。このように、ゼロデータ仮想領域に対応する実領域には NULL 値を格納しておき、当該仮想領域に対するアクセスについては、常にゼロデータが返されることになる。

[0160] 第 2 の実施形態では、ゼロデータ仮想領域についてはそれが未割当の仮想領域であっても、VVOL 管理表 311 および管理側 VVOL 管理表 810 に、当該仮想領域に対応する情報を登録したままにする。こうすることにより、アクセス分布変化量計算処理において、ゼロデータ仮想領域に対するアクセス数の情報を含めて、仮想ボリュームのアクセス分布変化量を正確に計算することができる。また、再配置時間帯計算処理において、ゼロデータ仮想領域に対するアクセス数の情報を含めて計算した仮想ボリュームのアクセス分布変化量の情報に基づいて、仮想ボリュームの再配置時間帯を計算することができる。

[0161] なお、上述のように、第 2 の実施形態では、ゼロデータが格納された仮想領域を未割当にしているが、その代わりに、ストレージシステムが管理する所定の実領域（以下で、基底実領域という）を当該仮想領域に割り当てても良い。ストレージサブシステムはホスト計算機から当該仮想領域へのリード要求を受信した時、当該仮想領域を未割当にしていた場合は、上述のように「未割当の仮想領域に対するリード要求にはゼロデータを返す」という規則に従い、ゼロデータを生成して返す必要がある。しかし、基底実領域がある場合は、事前にゼロデータを基底実領域に格納しておけば、ホスト計算機に返すデータは他の割当済みの仮想領域と同様の処理で取得可能である。

[0162] <処理内容>

次に、第2の実施形態において行われる処理について説明する。第2の実施形態において行われる処理は第1の実施形態の処理と比較して、以下の処理が異なる。

[0163] (i) ストレージ制御プログラム307は、ホスト計算機101が送信したリード要求を受信した時、リード対象の仮想領域がVVOL管理表311に登録されていなかった場合は、ホスト計算機101に対してエラーを送信する。また、リード対象の仮想領域はVVOL管理表311に登録されているが、当該仮想領域に対応する実領域ID2203がNULL値だった（当該仮想領域がゼロデータ仮想領域だった）場合、ストレージ制御プログラム307はホスト計算機101に、実領域内のデータに代えてゼロデータを送信する。

[0164] (ii) ストレージ制御プログラム307は、ホスト計算機101が送信したライト要求を受信した時、ライト対象データがゼロデータだった場合は、ライト対象の仮想領域に対応する実領域ID2203で識別される実領域に対して、実領域管理表309の当該実領域に対応する割当状況504に未割当の値を格納する。また、ライト対象の仮想領域に対応する実領域ID2203および再配置判定結果2205にNULL値を格納する。

[0165] (iii) 再配置処理において、VVOL管理表311および管理側VVOL管理表810に登録されている仮想領域の内、未割当の仮想領域に対しては、再配置判定処理を行わない。以下では、再配置処理のみ詳細に説明する。

[0166] <再配置処理の詳細>

図25は、第2の実施形態における再配置処理の詳細について説明するためのフローチャートである。当該再配置処理は、管理側VVOL管理表810に登録されている仮想領域の内、割当済の仮想領域のみに対してステップ2502～2509の処理を行う。具体的には、構成管理プログラム804は、管理側VVOL管理表810の実領域ID2303または／および再配置判定結果2305がNULL値のみが割り当てられている仮想領域に対し

ては、ステップ2502～2509の処理を行わず、次の仮想領域の処理に移る。

[0167] (3) まとめ

(i) 上述したように、特許文献1では、データのアクセス数に応じて記憶デバイスの移行処理を実行している。これは、データへのアクセス要求に求められるストレージデバイスの性能を保ち、かつストレージサブシステム内の物理記憶デバイスを効率的に利用するためである。各データが、そのアクセス状況に応じたデバイスに配置されている時間が長いほど、ストレージサブシステム内の物理記憶デバイスを効率的に利用していると言える。したがって、移行した後のデータへのアクセス状況は、なるべく長い期間移行先のデバイスの特性に適した状況であることが期待される。

[0168] しかしながら、仮想領域へのアクセス状況は、必ずしも一定ではない。例えば、オンライン処理アプリケーションは、多くのリクエストを高速に処理するため、オンライン処理アプリケーションに処理されるリクエストが短い時間間隔で切り替わる場合がある。そのような場合、リクエストによって、処理中にアクセスするデータは異なるため、オンライン処理アプリケーションがアクセスするデータは、短い時間間隔で変化する。したがって、オンライン処理アプリケーションがアクセスする、ストレージサブシステム内の仮想領域も、短い時間間隔で変化するため、ストレージサブシステム内の各仮想領域へのアクセス頻度も、短い時間間隔で変化すると考えられる。

[0169] このような時に、特許文献1にあるように、データの過去のアクセス履歴から当該データを格納するのに適した特性を持つデバイスを特定し、そのデバイスにデータを移行しても、データのアクセス状況がすぐに変化し、データが不適切なデバイスに格納された状態になってしまう。また、アクセス状況が変化するたびに、それに合わせてデータを移行すると、データ移行が頻繁に発生し、ストレージサブシステム全体の性能に影響を及ぼす恐れがある。

[0170] そこで、本発明は、仮想ボリューム内のホットスポットの仮想領域の位置

の変化を捉え、変化量が設定閾値よりも小さい時間帯をデータ移行時間帯候補として管理者に提示するようにしている。そして、データを移行する際には、長期間にわたってその効果が得られると予想される場合に限り、データのアクセス履歴（アクセス分布変化量）に基づいて、当該データの記録により適したデバイス内の実領域へのデータ移行を実施するようにしている。

[0171] (ii) これを実現するために、本発明の第1の実施形態では、管理システムが、各仮想ボリューム内でホットスポットとなる仮想領域の位置変化を示すアクセス分布変化量（負荷集中領域変化情報）を演算する。そして、アクセス分布変化量に基づいて、ホットスポットとなる仮想領域の位置が不変（位置変化が少ない、或いはホットスポットが安定している）となる時間帯を示す負荷位置不変時間帯を特定し、それを表示デバイスに表示する。このようにすることにより、管理者は、各仮想ボリューム内の各仮想領域に対応する実領域の割当状況を見直す時間帯（再配置時間帯）の候補を知ることができるようになる。

[0172] より具体的には、管理システムは、仮想ボリュームの複数の仮想領域に対するホスト計算機からのアクセス負荷量を、複数のタイミング（定期的に繰り返して：ただし、必ずしも周期的なタイミングである必要はない）において取得する。そして、複数の仮想領域のそれぞれについて、複数のタイミングのうち隣り合う異なるタイミングにおけるアクセス負荷量の差分値をそれぞれ計算し、その差分値を統計処理することにより、アクセス変化量の情報を演算する（図17参照）。このようにすることにより、ホットスポットとなる仮想領域の変化を定量的に測定することができるようになる。

[0173] また、上記管理システムは、一日を複数の時間帯に区切り、各時間帯について、設定された期間におけるアクセス分布変化量の平均値を算出し、当該平均値が設定された閾値（管理者が設定可能）より小さいか判定する。そして、閾値よりも小さい平均値に対応する時間帯が連続する場合には、当該連続する時間帯を結合して負荷不変時間帯とする。このようにすることにより、アクセス分布変化量に幅を持たせてホットスポットとなる仮想領域の位置

が不変となる時間帯を特定することができ、また、アクセスに対応する処理の種類（例えば、昼間のアプリケーションによる処理や夜間のバッチ処理等）によって負荷位置不変時間帯にばらつきがある場合（比較的短時間であったり、ある程度長時間に亘る場合であったり）にも柔軟に対応して再配置時間帯を特定することができるようになる。

[0174] さらに、管理システムは、負荷位置不変時間帯に、実領域に格納されたデータの配置変更を実行する。具体的には、管理システムは、負荷位置不変時間帯が開始した後の所定時間内（例えば、30分間）におけるアクセス数の平均値を算出する。そして、アクセス数の平均値とストレージサブシステムの構成情報から、対象の仮想領域に対応する実領域のデータが配置されるべき階層を特定する。さらに、管理システムは、対象の仮想領域に対応する実領域のデータが配置された現在の階層が特定された階層と異なる場合に、対象に対応する実領域の仮想領域のデータを現在の階層とは異なる階層（特定された階層）に移行する。このようにすることにより、ホットスポットの位置が安定した状態で、アクセス数が多い仮想領域に対応する実領域を本来配置されるべき階層の記憶デバイスに移行することが可能となる。なお、特定された階層に空き領域がない場合には、当該特定された階層に格納されたデータと対象の仮想領域に対応する実領域のデータとを入れ替えることにより、データの配置変更を実行するようにしても良い。このように階層レベルを下げて比較的影響の少ない実領域にホットスポットでアクセス数が多い仮想領域に対応する実領域のデータを移行することにより、容量に余裕がないストレージサブシステムであっても効率的に運用することが可能となる。

[0175] 以上のように、第1の実施形態によれば、ストレージ管理者はホスト計算機から仮想ボリュームへのアクセス状況を継続的に監視することなく、定期的にアクセス分布変化量閾値を入力するだけで、移行による効率性が高く得られる時間帯に再配置処理を行うことができ、ストレージサブシステム内の物理記憶デバイスを効率的に利用することができる。

[0176] (iii) 本発明の第2の実施形態では、管理システムは、仮想ボリュームにお

いて、データが未割り当ての仮想領域にゼロデータを割り当てて、それをゼロデータ仮想領域として管理している。このようにすることにより、未割り当ての仮想領域に対するアクセス数もカウントすることができ、より正確にアクセス分布及びその変化量をモニタリングすることができるようになる。

[0177] そして、管理システムは、データの配置変更を実行する際に、仮想ボリュームのゼロデータ仮想領域を除く仮想領域について、再配置時間帯（負荷位置不変時間帯）開始後の所定時間内（例えば、30分間）におけるアクセス数の平均値を取得する。そして、管理システムは、アクセス数の平均値とストレージサブシステムの構成情報から、対象の仮想領域に対応する実領域のデータが配置されるべき階層を特定し、対象の仮想領域のデータが配置された現在の階層が特定された階層と異なる場合に、対象の仮想領域に対応する実領域のデータを特定された階層に移行する。このようにすることにより、ゼロデータ仮想領域を再配置処理の対象から除くことができ、効率的に再配置処理が実行できる。また、ホットスポットの位置が安定した状態で、アクセス数が多い仮想領域に対応する実領域を本来配置されるべき階層の記憶デバイスに移行することが可能となる。第2の実施形態でも、特定された階層に空き領域がない場合には、当該特定された階層に格納されたデータと対象の仮想領域に対応する実領域のデータとを入れ替えることにより、データの配置変更を実行するようにしても良い。

[0178] 以上のように、第2の実施形態では、管理側VVOL管理表810に未割り当ての仮想領域が登録されている場合があるが、未割り当ての仮想領域は、割り当てられている実領域が存在しないため、実領域間のデータ移行を行うことができない。そのため、上記のように未割り当ての実領域には再配置判定を行わないようにしている。

[0179] (iv) 第1及び第2の実施形態では、アクセス数をカウントする場合に、ストレージサブシステム102内の各仮想ボリュームのLBA範囲に対するI/O数をカウントするようにしているが、アクセス数をカウントする方法はこれに限られるものではない。例えば、ホスト計算機101にエージェント

を設け、エージェントがホスト計算機の例えばアプリケーションから発行される I/O 要求をキャッチし、それを管理計算機 103 に通知する。そして、管理計算機 103 が、自身が保持している管理側実領域管理表 808 及び管理側 VVOL 管理表 810 と通知された I/O 要求とを照合して、どの仮想領域へのアクセスかを検知し、アクセス数をカウントするようにしても良い。この場合、図において、管理計算機 103 とホスト計算機 101 はネットワークを介して接続されるようにするのが好ましい。

[0180] (v) なお、本発明は、実施形態の機能を実現するソフトウェアのプログラムコードによっても実現できる。この場合、プログラムコードを記録した記憶媒体をシステム或は装置に提供し、そのシステム或は装置のコンピュータ（又は CPU や MPU）が記憶媒体に格納されたプログラムコードを読み出す。この場合、記憶媒体から読み出されたプログラムコード自体が前述した実施形態の機能を実現することになり、そのプログラムコード自体、及びそれを記憶した記憶媒体は本発明を構成することになる。このようなプログラムコードを供給するための記憶媒体としては、例えば、フレキシブルディスク、CD-ROM、DVD-ROM、ハードディスク、光ディスク、光磁気ディスク、CD-R、磁気テープ、不揮発性のメモリカード、ROMなどが用いられる。

[0181] また、プログラムコードの指示に基づき、コンピュータ上で稼動している OS（オペレーティングシステム）などが実際の処理の一部又は全部を行い、その処理によって前述した実施の形態の機能が実現されるようにしてもよい。さらに、記憶媒体から読み出されたプログラムコードが、コンピュータ上のメモリに書きこまれた後、そのプログラムコードの指示に基づき、コンピュータの CPU などが実際の処理の一部又は全部を行い、その処理によって前述した実施の形態の機能が実現されるようにしてもよい。

[0182] また、実施の形態の機能を実現するソフトウェアのプログラムコードを、ネットワークを介して配信することにより、それをシステム又は装置のハードディスクやメモリ等の記憶手段又は CD-RW、CD-R 等の記憶媒体に

格納し、使用時にそのシステム又は装置のコンピュータ（又はCPUやMPU）が当該記憶手段や当該記憶媒体に格納されたプログラムコードを読み出して実行するようにしても良い。

符号の説明

- [0183] 10…計算機システム
- 101…ホスト計算機
- 102…ストレージサブシステム
- 103…管理計算機
- 104…Webブラウザ用計算機
- 105…ストレージエリアネットワーク
- 106…管理用ネットワーク
- 107…管理システム

請求の範囲

[請求項1]

ホスト計算機と、実体的な記憶領域である実領域とそれに関連付けられる仮想ボリュームを有するストレージサブシステムと、当該ストレージサブシステムを管理する管理システムと、を有し、前記ストレージサブシステムが前記仮想ボリュームを複数の部分的な領域に分割して構成される複数の仮想領域を含む、計算機システムの管理方法であって、

前記管理システムが、前記仮想ボリュームにおいてホットスポットとなる仮想領域の位置変化を示す負荷集中領域変化情報を取得し、

前記管理システムが、前記負荷集中領域変化情報に基づいて、前記ホットスポットとなる仮想領域の位置が不変となる時間帯を示す負荷位置不変時間帯を特定し、

前記管理システムが、前記特定された負荷位置不変時間帯を表示デバイスに表示する、

ことを特徴とする管理方法。

[請求項2]

請求項1において、

前記管理システムは、前記負荷集中領域変化情報を取得する際に、

前記仮想ボリュームの前記複数の仮想領域に対する前記ホスト計算機からのアクセス負荷量を複数の時間において取得し、

前記複数の仮想領域のそれぞれについて、前記複数の時間のうち異なる時間における前記アクセス負荷量の差分値をそれぞれ計算し、

前記各差分値を統計処理して前記負荷集中領域変化情報を算出する、

ことを特徴とする管理方法。

[請求項3]

請求項2において、

前記管理システムは、前記負荷位置不変時間帯を特定する際に、

一日を複数の時間帯に区切り、各時間帯について、設定された期間における前記負荷集中領域変化情報の平均値を算出し、

前記平均値が設定された閾値より小さいか判定し、

前記閾値よりも小さい前記平均値に対応する時間帯が連続する場合には、当該連続する時間帯を結合して前記負荷不変時間帯とする、ことを特徴とする管理方法。

[請求項4]

請求項 1 において、

さらに、前記管理システムが、前記負荷位置不変時間帯に、前記実領域に格納されたデータの配置変更を実行することを特徴とする管理方法。

[請求項5]

請求項 4 において、

前記管理システムは、前記データの配置変更を実行する際に、

前記負荷位置不変時間帯におけるアクセス数の情報を取得し、

前記アクセス数の情報と前記ストレージサブシステムの構成情報から、対象の仮想領域に対応する実領域のデータが配置されるべき階層を特定し、

前記対象の仮想領域に対応する実領域のデータが配置された現在の階層が前記特定された階層と異なる場合に、前記対象の仮想領域に対応する実領域の仮想領域のデータを前記現在の階層とは異なる階層に移行する、

ことを特徴とする管理方法。

[請求項6]

請求項 5 において、

前記管理システムは、前記特定された階層に空き領域がない場合には、当該特定された階層に格納されたデータと前記対象の仮想領域に対応する実領域のデータとを入れ替えることにより、データの配置変更を実行することを特徴とする管理方法。

[請求項7]

請求項 4 において、

前記管理システムは、前記仮想ボリュームにおいて、データが未割り当ての仮想領域にゼロデータを割り当てて、それをゼロデータ仮想領域として管理しており、

前記管理システムは、前記データの配置変更を実行する際に、前記仮想ボリュームの前記ゼロデータ仮想領域を除く仮想領域について、

前記負荷位置不変時間帯におけるアクセス数の情報を取得し、

前記アクセス数の情報と前記ストレージサブシステムの構成情報から、対象の仮想領域に対応する実領域のデータが配置されるべき階層を特定し、

前記対象の仮想領域に対応する実領域のデータが配置された現在の階層が前記特定された階層と異なる場合に、前記対象の仮想領域に対応する実領域のデータを前記現在の階層とは異なる階層に移行する

、

ことを特徴とする管理方法。

[請求項8]

請求項 1 において、

i) 前記管理システムは、前記負荷集中領域変化情報を取得する際に、

、

前記仮想ボリュームの前記複数の仮想領域に対する前記ホスト計算機からのアクセス負荷量を複数の時間において取得し、

前記複数の仮想領域のそれぞれについて、前記複数の時間のうち異なる時間における前記アクセス負荷量の差分値をそれぞれ計算し、

前記各差分値を統計処理して前記負荷集中領域変化情報を算出し

、

ii) 前記管理システムは、前記負荷位置不変時間帯を特定する際に、

一日を複数の時間帯に区切り、各時間帯について、設定された期間における前記負荷集中領域変化情報の平均値を算出し、

前記平均値が設定された閾値より小さいか判定し、

前記閾値よりも小さい前記平均値に対応する時間帯が連続する場合には、当該連続する時間帯を結合して前記負荷不変時間帯とし、

iii) 前記管理方法は、さらに、

前記負荷位置不変時間帯開始から所定時間内におけるアクセス数

の平均値を算出し、

前記アクセス数の平均値と前記ストレージサブシステムの構成情報から、対象の仮想領域に対応する実領域のデータが配置されるべき階層を特定し、

前記対象の仮想領域に対応する実領域のデータが配置された現在の階層が前記特定された階層と異なる場合に、前記対象の仮想領域のデータを前記現在の階層とは異なる階層に移行し、

配置変更後の前記ストレージサブシステムの構成情報を更新する、
ことを特徴とする管理方法。

[請求項9] 実体的な記憶領域である実領域とそれに関連付けられる仮想ボリュームを有するストレージサブシステムを管理する管理システムであって、

演算を実行するプロセッサと、

情報を格納するメモリと、

前記ストレージサブシステムと通信を行うための通信ポートと、を有し、

前記プロセッサは、前記仮想ボリュームにおいてホットスポットとなる仮想領域の位置変化を示す負荷集中領域変化情報を取得し、

前記負荷集中領域変化情報に基づいて、前記ホットスポットとなる仮想領域の位置が不変となる時間帯を示す負荷位置不変時間帯を特定し、

前記特定された負荷位置不変時間帯を前記メモリに格納し、表示デバイスに表示する、

ことを特徴とする管理システム。

[請求項10] 請求項9において、

前記プロセッサは、前記負荷集中領域変化情報を取得する際に、

前記仮想ボリュームの前記複数の仮想領域に対する前記ホスト計

算機からのアクセス負荷量を複数の時間において取得し、

前記複数の仮想領域のそれぞれについて、前記複数の時間のうち異なる時間における前記アクセス負荷量の差分値をそれぞれ計算し、

前記各差分値を統計処理して前記負荷集中領域変化情報を算出する、

ことを特徴とする管理システム。

[請求項11]

請求項10において、

前記プロセッサは、前記負荷位置不変時間帯を特定する際に、

一日を複数の時間帯に区切り、各時間帯について、設定された期間における前記負荷集中領域変化情報の平均値を算出し、

前記平均値が設定された閾値より小さいか判定し、

前記閾値よりも小さい前記平均値に対応する時間帯が連続する場合には、当該連続する時間帯を結合して前記負荷不変時間帯とする、ことを特徴とする管理システム。

[請求項12]

請求項9において、

前記プロセッサは、

前記負荷位置不変時間帯におけるアクセス数の情報を取得し、

前記アクセス数の情報と前記ストレージサブシステムの構成情報から、対象の仮想領域に対応する実領域のデータが配置されるべき階層を特定し、

前記対象の仮想領域に対応する実領域のデータが配置された現在の階層が前記特定された階層と異なる場合に、前記対象の仮想領域に対応する実領域のデータを前記現在の階層とは異なる階層に移行する、

ことを特徴とする管理システム。

[請求項13]

請求項12において、

前記プロセッサは、前記特定された階層に空き領域がない場合には、当該特定された階層に格納されたデータと前記対象の仮想領域に対

応する実領域のデータとを入れ替えることにより、データの配置変更を実行することを特徴とする管理方法。

[請求項14]

請求項9において、

前記プロセッサは、前記仮想ボリュームにおいて、データが未割り当ての仮想領域にゼロデータを割り当てて、それをゼロデータ仮想領域として管理しており、

前記プロセッサは、前記仮想ボリュームの前記ゼロデータ仮想領域を除く仮想領域について、

前記負荷位置不変時間帯におけるアクセス数の情報を取得し、

前記アクセス数の情報と前記ストレージサブシステムの構成情報から、対象の仮想領域に対応する実領域のデータが配置されるべき階層を特定し、

前記対象の仮想領域に対応する実領域のデータが配置された現在の階層が前記特定された階層と異なる場合に、前記対象の仮想領域に対応する実領域のデータを前記現在の階層とは異なる階層に移行すること、
ことを特徴とする管理システム。

[請求項15]

請求項9において、

前記プロセッサは、

i) 前記負荷集中領域変化情報を取得する際に、

前記仮想ボリュームの前記複数の仮想領域に対する前記ホスト計算機からのアクセス負荷量を複数の時間において取得し、

前記複数の仮想領域のそれぞれについて、前記複数の時間うち異なる時間における前記アクセス負荷量の差分値をそれぞれ計算し、

前記各差分値を統計処理して前記負荷集中領域変化情報を算出し、

ii) 前記管理システムは、前記負荷位置不変時間帯を特定する際に、

一日を複数の時間帯に区切り、各時間帯について、設定された期

間における前記負荷集中領域変化情報の平均値を算出し、

前記平均値が設定された閾値より小さいか判定し、

前記閾値よりも小さい前記平均値に対応する時間帯が連続する場合には、当該連続する時間帯を結合して前記負荷不変時間帯とし、
iii) さらに、

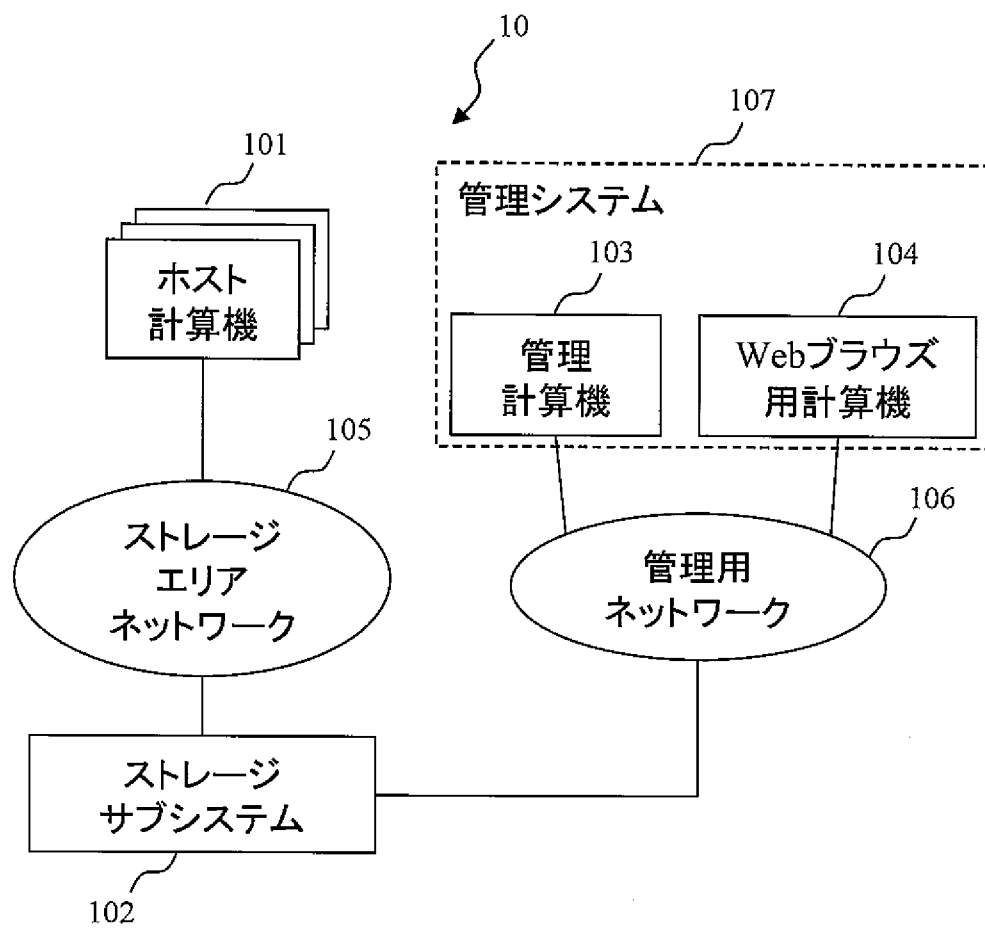
前記負荷位置不変時間帯開始から所定時間内におけるアクセス数の平均値を算出し、

前記アクセス数の平均値と前記ストレージサブシステムの構成情報から、対象の仮想領域に対応する実領域のデータが配置されるべき階層を特定し、

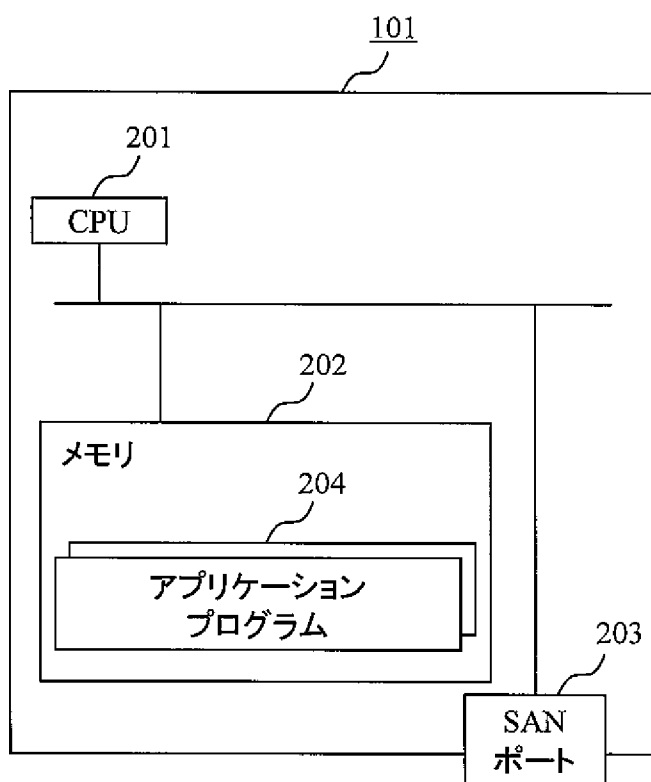
前記対象の仮想領域に対応する実領域のデータが配置された現在の階層が前記特定された階層と異なる場合に、前記対象の仮想領域のデータを前記現在の階層とは異なる階層に移行し、

配置変更後の前記ストレージサブシステムの構成情報を更新する、
ことを特徴とする管理システム。

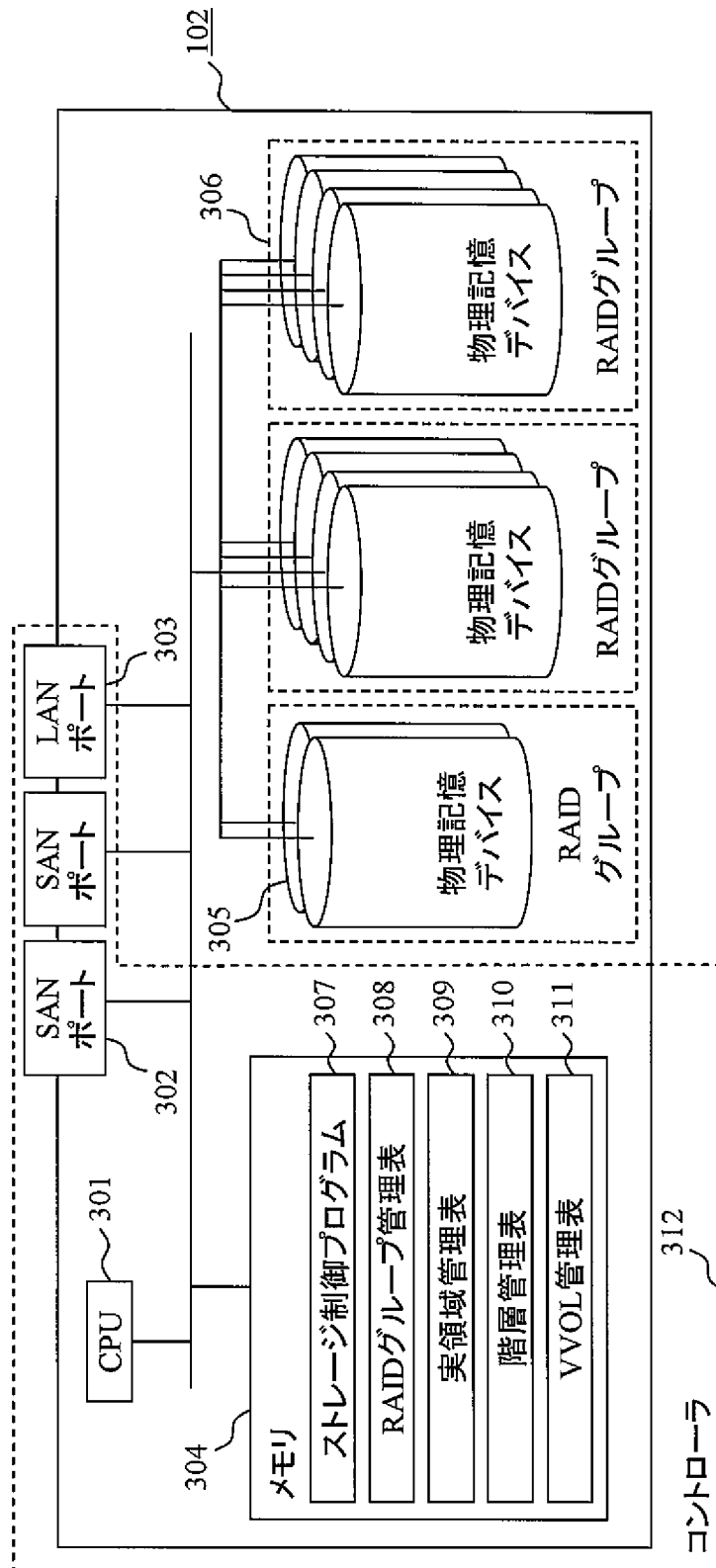
[図1]



[図2]



[図3]



[図4]

401 RAIDグループID	402 デバイス種別	403 RAIDレベル	404 対応PDEV ID	308
RAIDグループ1	SSD	RAID 5 (3D+1P)	PDEV1	
			PDEV2	
			PDEV3	
			PDEV4	
RAIDグループ2	FC	RAID 5 (3D+1P)	PDEV5	
			PDEV6	
			PDEV7	
			PDEV8	
RAIDグループ3	SATA	RAID 0	PDEV9	
			PDEV10	

[図5]

501 RAIDグループID	502 実領域ID	503 RG LBA範囲	504 割当状況	309
RAIDグループ1	SEG1	0 – 999	割当済み	
	SEG2	1000 – 1999	割当済み	
	SEG3	2000 – 2999	割当済み	
RAIDグループ2	SEG4	0 – 999	割当済み	
	SEG5	1000 – 1999	未割当	
RAIDグループ3	SEG6	0 – 999	割当済み	
	SEG7	1000 – 1999	未割当	
⋮	⋮	⋮	⋮	

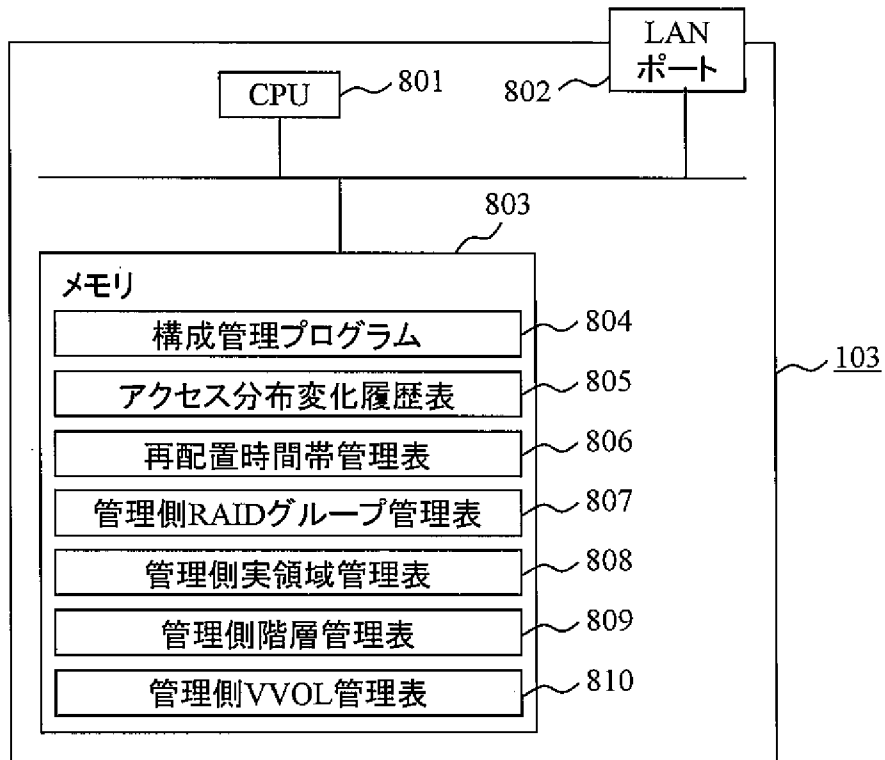
[図6]

601 VVOL ID	602 VVOL LBA範囲	603 実領域ID	604 アクセス数	605 再配置判定結果	311
Volume1	0 - 999	SEG1	792000	Tier1	
	1000 - 1999	SEG4	342000	Tier2	
	2000 - 2999	SEG5	64800	Tier3	
Volume2	0 - 999	SEG2	558000	Tier2	
Volume3	0 - 999	SEG3	144000	Tier3	

[図7]

701 階層ID	702 性能要件	703 許容IOPS範囲	310
Tier1	デバイス種別=SSD, RAIDレベル=RAID5	100以上	
Tier2	デバイス種別=FC, RAIDレベル=RAID5	30以上100未満	
Tier3	デバイス種別=SATA, RAIDレベル=ANY	30未満	

[図8]



[図9]

VVOL ID	観測日時		アクセス分布 変化量	アクセス 平均 (IOPS)
	日付	時間帯		
VVOL1	2010/5/6	9:00 – 10:00	50	175
		10:00 – 11:00	5	145
		11:00 – 12:00	7	160
	⋮	⋮	⋮	⋮
VVOL2	2010/5/6	9:00 – 10:00	42	80
		10:00 – 11:00	57	85
		11:00 – 12:00	39	75
	⋮	⋮	⋮	⋮
VVOL3	2010/5/6	9:00 – 10:00	2	20
		10:00 – 11:00	4	15
		11:00 – 12:00	5	10
	⋮	⋮	⋮	⋮

[図10]

VVOL ID	分布変化閾値	再配置時間帯
VVOL1	50	9:00 – 16:00
		21:00 – 23:00
VVOL2	30	0:00 – 3:00
VVOL3	10	

[図11]

1101 RAIDグループID	1102 デバイス種別	1103 RAIDレベル	807
RAIDグループ1	SSD	RAID 5 (3D+1P)	
RAIDグループ2	FC	RAID 5 (3D+1P)	
RAIDグループ3	SATA	RAID 0	

[図12]

1201 RAIDグループID	1202 実領域ID	1203 RG LBA範囲	1204 割当状況	808
RAIDグループ1	SEG1	0 – 999	割当済み	
	SEG2	1000 – 1999	割当済み	
	SEG3	2000 – 2999	割当済み	
RAIDグループ2	SEG4	0 – 999	割当済み	
	SEG5	1000 – 1999	未割当	
RAIDグループ3	SEG6	0 – 999	割当済み	
	SEG7	1000 – 1999	未割当	
⋮	⋮	⋮	⋮	

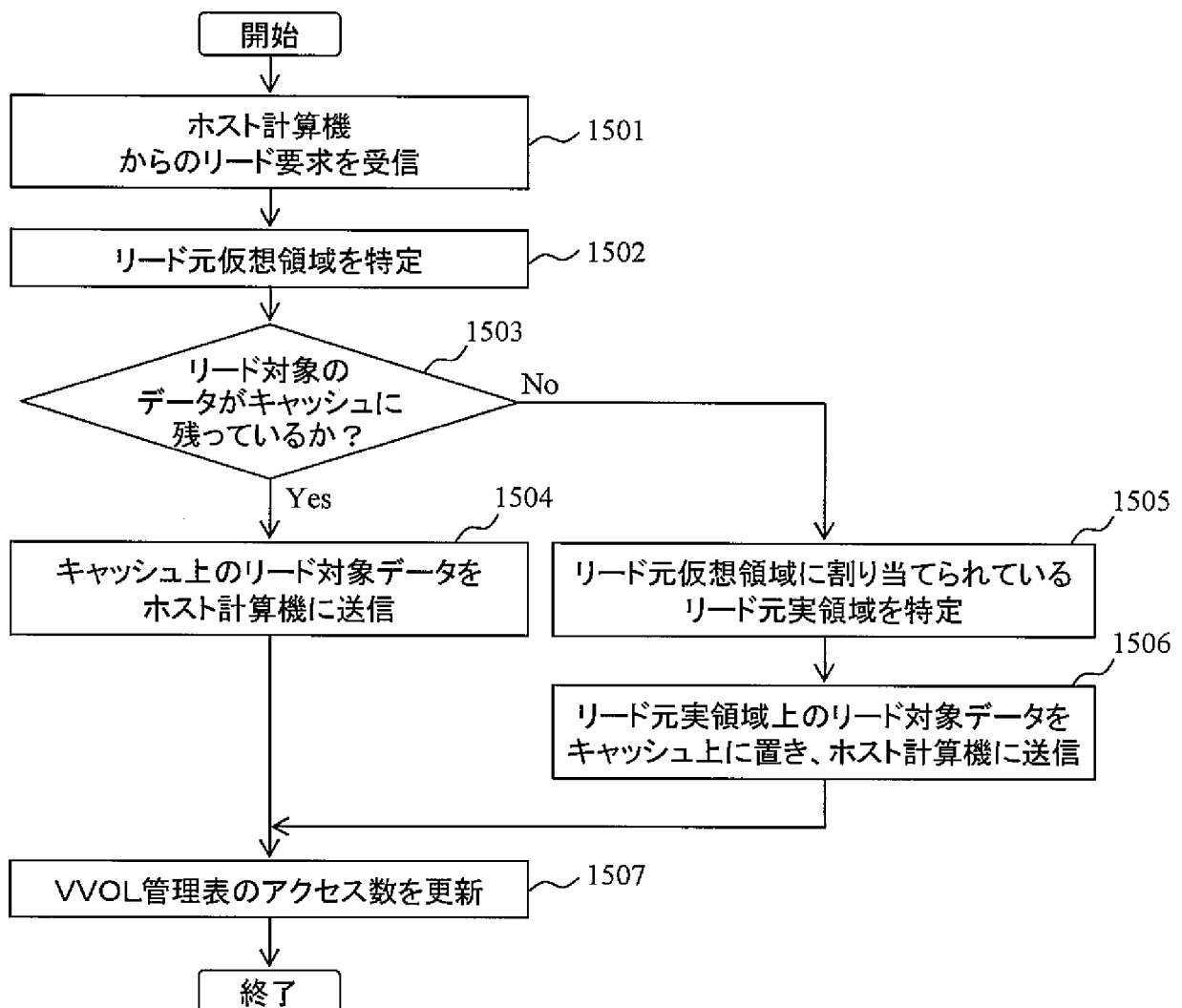
[図13]

1301	1302	1303	1304				1305	810
VVOL ID	VVOL LBA範囲	実領域 ID	アクセス数				再配置 判定結果	
			アクセス分布変化量 計算処理にて使う値	前回処理時の 平均値 (IOPS)	先頭時刻での アクセス数	再配置処理にて 使う値		
Volume1	0 - 999	SEG1	360000	120	621000	95	Tier1	
	1000 - 1999	SEG4	180000	45	255600	48	Tier2	
	2000 - 2999	SEG5	28800	10	37800	15	Tier3	
Volume2	0 - 999	SEG2	270000	80	486000	40	Tier2	
Volume3	0 - 999	SEG3	72000	20	108000	20	Tier3	

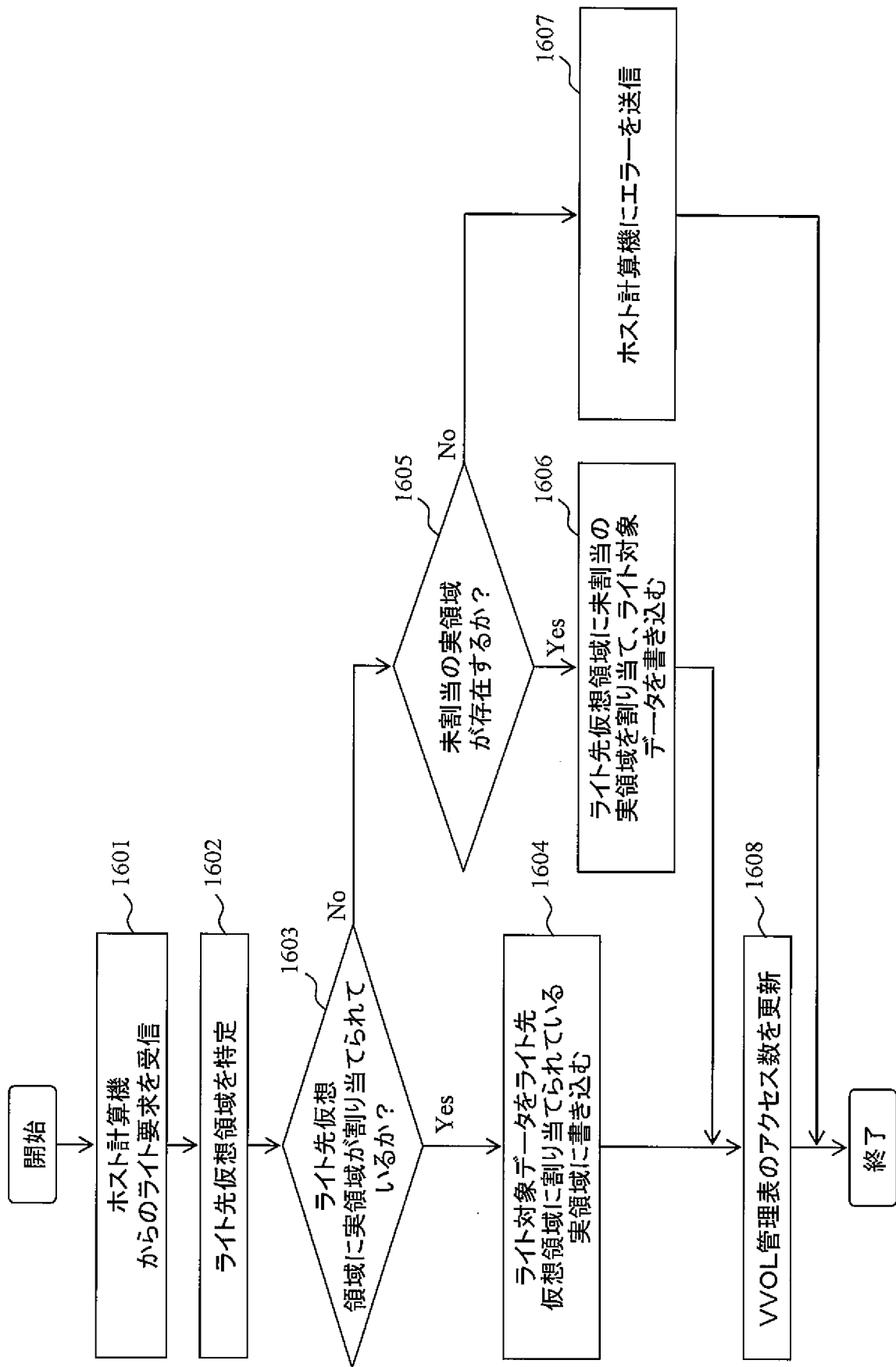
[図14]

階層ID	性能要件	許容IOPS範囲
Tier1	デバイス種別=SSD, RAIDレベル=RAID5	100以上
Tier2	デバイス種別=FC, RAIDレベル=RAID5	30以上100未満
Tier3	デバイス種別=SATA, RAIDレベル=ANY	30未満

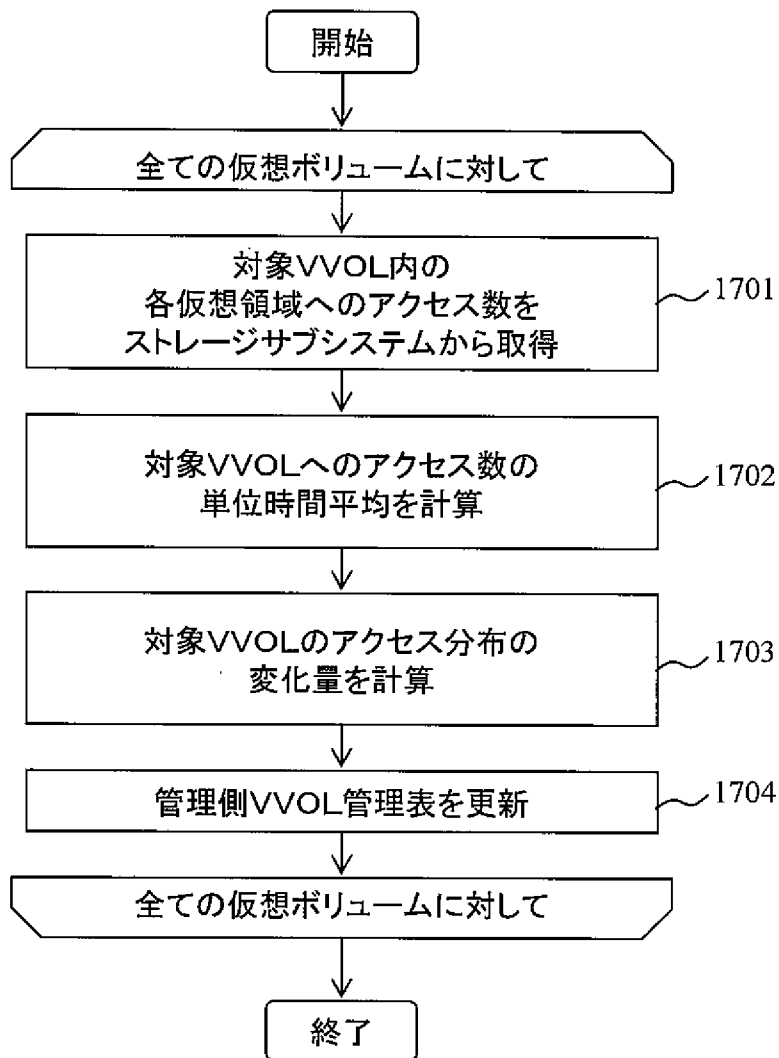
[図15]



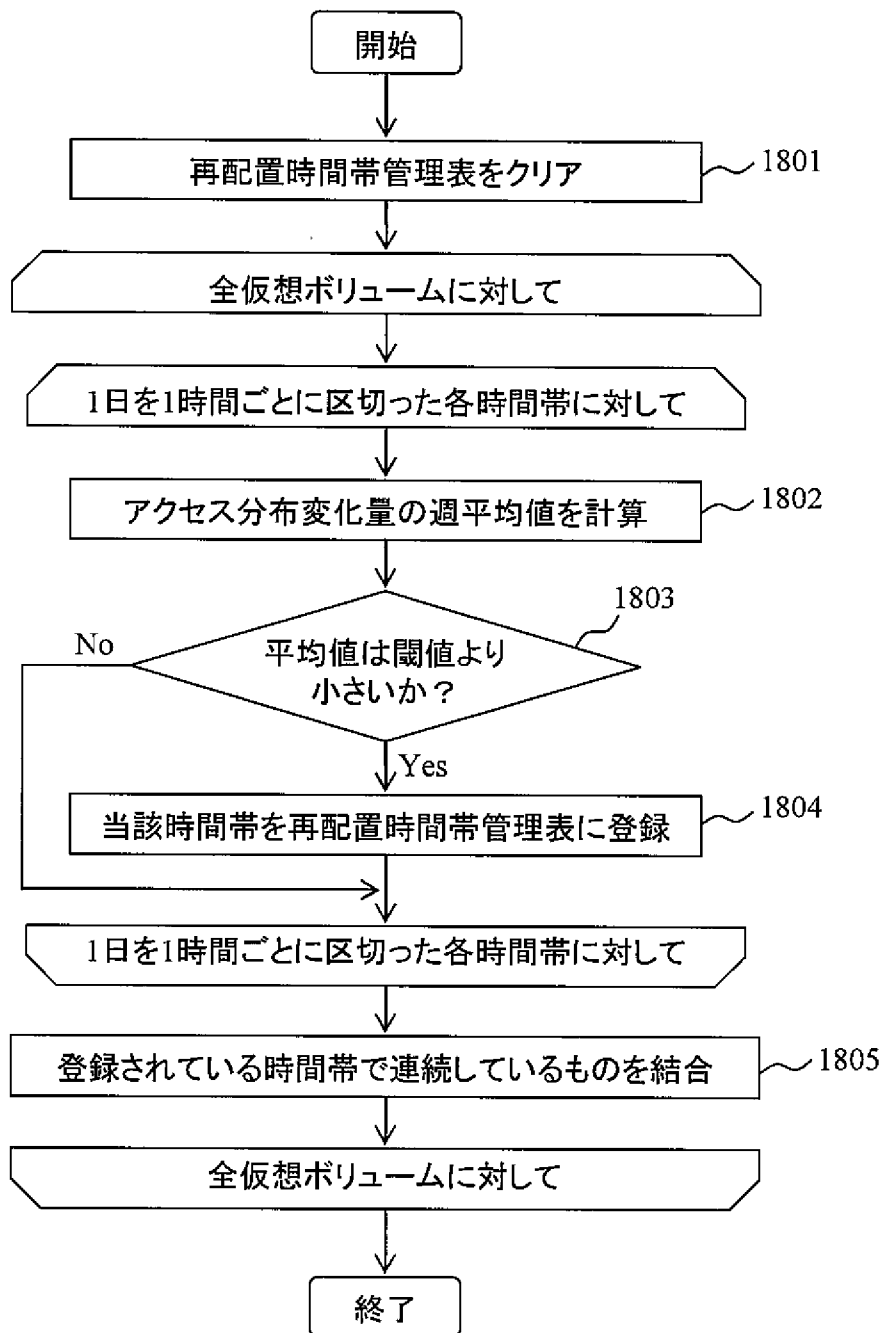
[図16]



[図17]



[図18]



[図19]

図19は、アクセス分布変化閾値入力画面のスクリーンショットを示す。画面は、仮想ボリューム選択領域、管理者判断材料表示領域、閾値入力領域、再配置時間帯表示領域、および適用/キャンセルボタン領域に構成されている。

仮想ボリューム選択領域 (1901) は、仮想ボリューム (1904) のリストを表示する。リストには Volume 1、Volume 2、Volume 3、および省略記号 (.....) が含まれる。Volume 1 は選択されている。

管理者判断材料表示領域 (1905) は、アクセス分布変化量 (1906) の表を表示する。表の列は「時間帯」と「アクセス分布変化量」である。

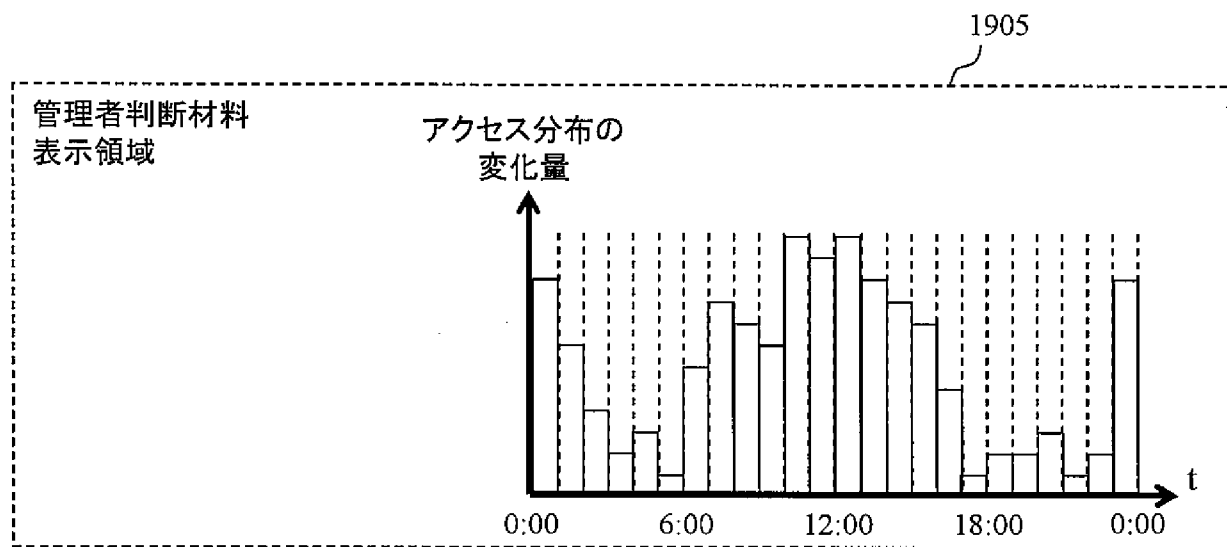
時間帯	アクセス分布変化量
9:00 - 10:00	50
10:00 - 11:00	5
11:00 - 12:00	7
⋮	⋮

閾値入力領域 (1907) は、分布変化閾値 (1908) の入力フィールドを表示する。入力値は 30 である。

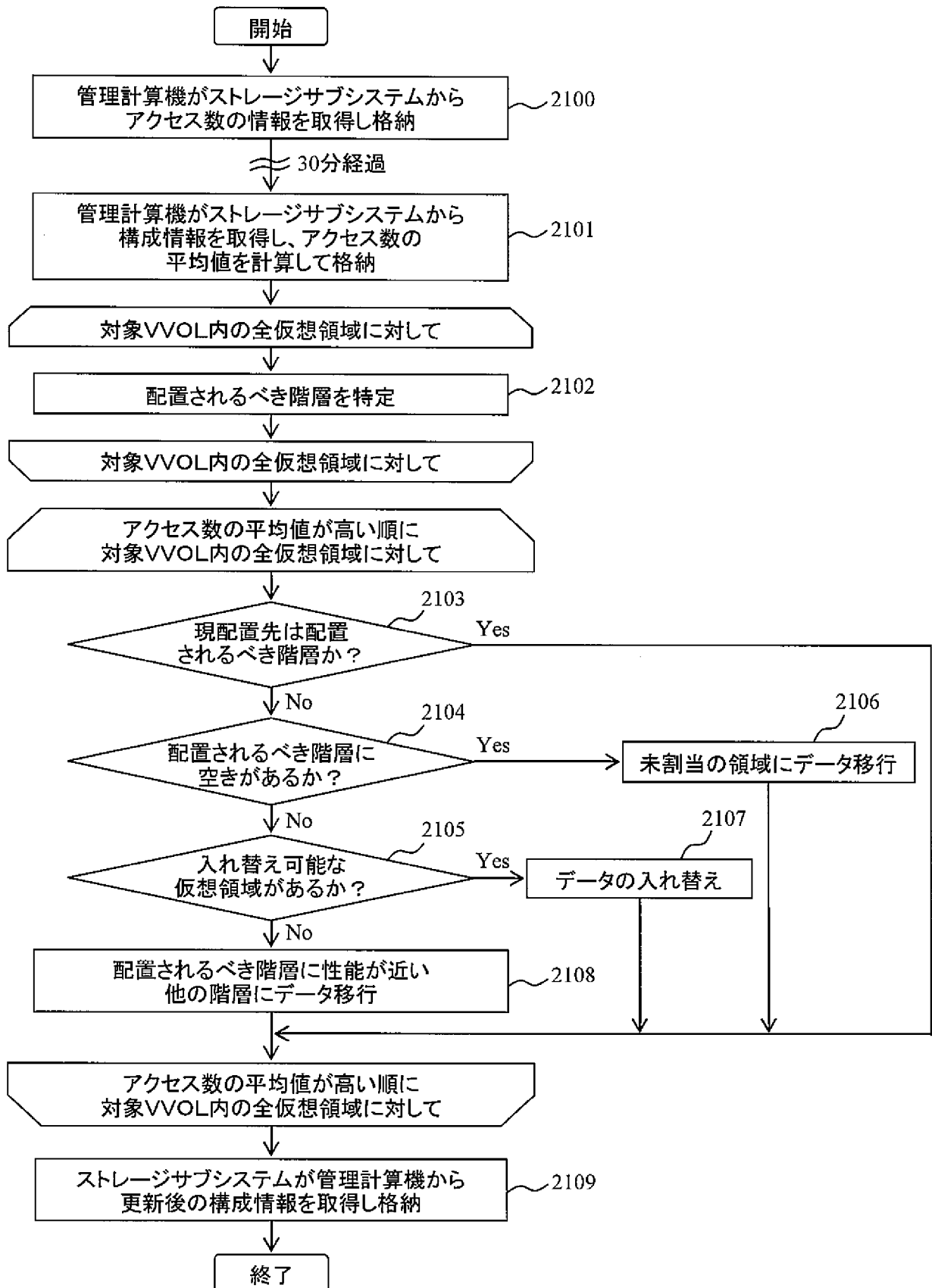
再配置時間帯表示領域 (1903) は、再配置時間帯 (1902) のリストを表示する。リストには 10:00 - 16:00、21:00 - 23:00、および省略記号 (.....) が含まれる。10:00 - 16:00 は選択されている。

画面下部には、適用 (1902) とキャンセル (1903) のボタンがある。

[図20]



[図21]



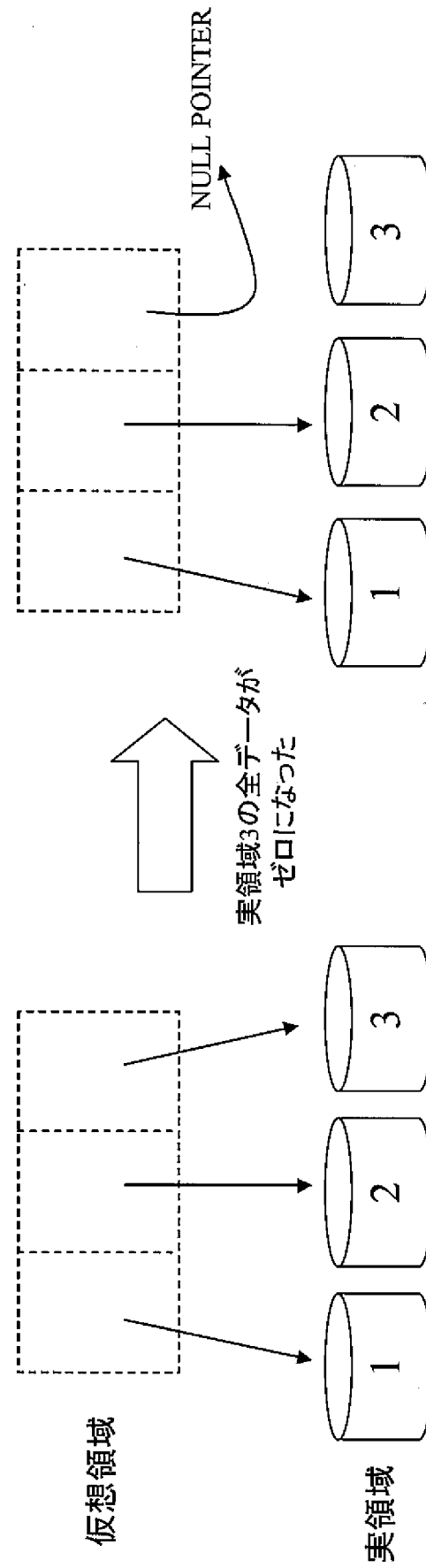
[図22]

VVOL ID	VVOL LBA範囲	実領域ID	アクセス数	再配置判定結果
Volume1	0 – 999	SEG1	792000	Tier1
	1000 – 1999	SEG4	342000	Tier2
	2000 – 2999	NULL	64800	NULL
Volume2	0 – 999	SEG2	558000	Tier2
Volume3	0 – 999	SEG3	144000	Tier3

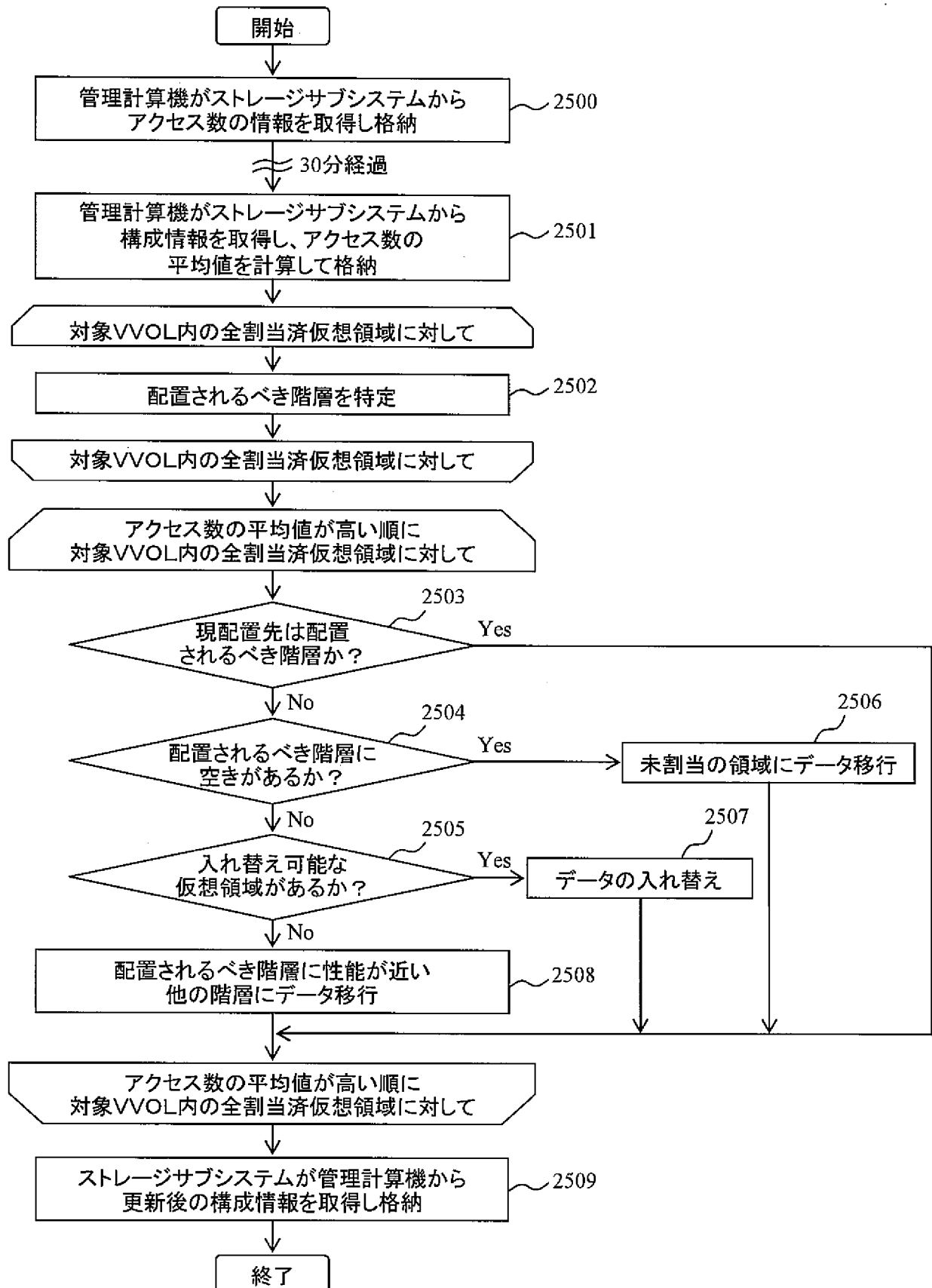
[図23]

2301	2302	2303	2304				2305	810
VVOL ID	VVOL LBA範囲	実領域 ID	アクセス数				再配置 判定結果	
			アクセス分布変化量計算 処理にて使う値		再配置処理にて 使う値			
			前回処理時の アクセス数	前回処理時の 平均値 (IOPS)	先頭時刻での アクセス数	平均値 (IOPS)		
Volume1	0 - 999	SEG1	360000	120	621000	95	Tier1	
	1000 - 1999	SEG4	180000	45	255600	48	Tier2	
	2000 - 2999	NULL	28800	10	37800	15	NULL	
Volume2	0 - 999	SEG2	270000	80	486000	40	Tier2	
Volume3	0 - 999	SEG3	72000	20	108000	20	Tier3	

[図24]



[図25]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2010/063512

A. CLASSIFICATION OF SUBJECT MATTER

G06F12/00 (2006.01) i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06F12/00

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho	1922-1996	Jitsuyo Shinan Toroku Koho	1996-2010
Kokai Jitsuyo Shinan Koho	1971-2010	Toroku Jitsuyo Shinan Koho	1994-2010

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2007-66259 A (Hitachi, Ltd.), 15 March 2007 (15.03.2007), paragraphs [0039] to [0055]; fig. 12 to 15 & US 2007/0055713 A1	1-15
A	JP 2010-26919 A (Hitachi, Ltd.), 04 February 2010 (04.02.2010), entire text; all drawings & US 2010/0023566 A1	1-15
A	JP 2006-277723 A (Hitachi, Ltd.), 12 October 2006 (12.10.2006), entire text; all drawings & US 2006/0224844 A1 & US 2007/0067588 A1	1-15

☐

Further documents are listed in the continuation of Box C.

☐

See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search
03 September, 2010 (03.09.10)

Date of mailing of the international search report
14 September, 2010 (14.09.10)

Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

A. 発明の属する分野の分類 (国際特許分類 (I P C))

Int.Cl. G06F12/00 (2006.01) i

B. 調査を行った分野

調査を行った最小限資料 (国際特許分類 (I P C))

Int.Cl. G06F12/00

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	1 9 2 2 - 1 9 9 6 年
日本国公開実用新案公報	1 9 7 1 - 2 0 1 0 年
日本国実用新案登録公報	1 9 9 6 - 2 0 1 0 年
日本国登録実用新案公報	1 9 9 4 - 2 0 1 0 年

国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
A	JP 2007-66259 A (株式会社日立製作所) 2007.03.15, 第 39-55 段落, 第 12-15 図 & US 2007/0055713 A1	1-15
A	JP 2010-26919 A (株式会社日立製作所) 2010.02.04, 全文, 全図 & US 2010/0023566 A1	1-15
A	JP 2006-277723 A (株式会社日立製作所) 2006.10.12, 全文, 全図 & US 2006/0224844 A1 & US 2007/0067588 A1	1-15

C 欄の続きにも文献が列挙されている。

パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

「A」特に関連のある文献ではなく、一般的技術水準を示すもの
「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)
「O」口頭による開示、使用、展示等に言及する文献
「P」国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献
「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの
「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
「Y」特に関連のある文献であって、当該文献と他の 1 以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
「&」同一パテントファミリー文献

国際調査を完了した日

0 3 . 0 9 . 2 0 1 0

国際調査報告の発送日

1 4 . 0 9 . 2 0 1 0

国際調査機関の名称及びあて先

日本国特許庁 (I S A / J P)

郵便番号 1 0 0 - 8 9 1 5

東京都千代田区霞が関三丁目 4 番 3 号

特許庁審査官 (権限のある職員)

北村 学

5 M

4 5 3 5

電話番号 0 3 - 3 5 8 1 - 1 1 0 1 内線 3 5 9 9