

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第6146087号
(P6146087)

(45) 発行日 平成29年6月14日 (2017. 6. 14)

(24) 登録日 平成29年5月26日 (2017. 5. 26)

(51) Int. Cl.

F I

G 0 6 F 3/06 (2006.01)

G 0 6 F 12/08 (2016.01)

G 0 6 F 12/0802 (2016.01)

G 0 6 F 3/06 3 0 2 A

G 0 6 F 3/06 3 0 1 Z

G 0 6 F 3/06 3 0 4 N

G 0 6 F 12/08 5 4 3 B

G 0 6 F 12/08 5 0 1 Z

請求項の数 11 (全 24 頁) 最終頁に続く

(21) 出願番号 特願2013-68865 (P2013-68865)
 (22) 出願日 平成25年3月28日 (2013. 3. 28)
 (65) 公開番号 特開2014-191749 (P2014-191749A)
 (43) 公開日 平成26年10月6日 (2014. 10. 6)
 審査請求日 平成27年12月4日 (2015. 12. 4)

(73) 特許権者 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
 (74) 代理人 100094525
 弁理士 土井 健二
 (74) 代理人 100094514
 弁理士 林 恒徳
 (72) 発明者 小林 英一
 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 寺井 幸子
 神奈川県川崎市中原区上小田中4丁目1番
 1号 富士通株式会社内

最終頁に続く

(54) 【発明の名称】 ストレージ制御プログラム、ストレージ制御方法、ストレージシステム及びその階層制御装置

(57) 【特許請求の範囲】

【請求項 1】

第1のアクセス速度を有する第1の記憶装置と、前記第1のアクセス速度より遅い第2のアクセス速度を有する第2の記憶装置と、キャッシュメモリとを有し、前記第1または第2の記憶装置内のデータへのアクセス要求時に、前記アクセス要求のアドレスのデータが前記キャッシュメモリに記憶されている場合前記キャッシュメモリのデータにアクセスし、記憶されていない場合前記第1または第2の記憶装置内の前記アクセス要求のデータにアクセスしアクセスデータを前記キャッシュメモリに記憶すると共に前記キャッシュメモリから過去に記憶したデータを削除するキャッシュ制御処理が行われるストレージの制御処理をコンピュータに実行させるストレージ制御プログラムであって、

前記制御処理は、

前記第1及び第2の記憶装置内の再配置単位のデータへのアクセス頻度を監視する監視処理と、

前記第2の記憶装置内の前記アクセス頻度が第1の基準値を超える前記再配置単位のデータを前記第1の記憶装置に再配置する再配置処理と、

前記第2の記憶装置が過負荷状態の場合に、前記第2の記憶装置内の少なくとも一部のデータが前記キャッシュ制御処理により前記キャッシュメモリに記憶されたら前記監視処理中は前記一部のデータを前記キャッシュメモリに保持し、前記一部のデータへのアクセス要求時に前記キャッシュ制御処理により前記キャッシュメモリ内の前記一部のデータがアクセスされる過負荷処理とを有し、

10

20

前記一部のデータを前記キャッシュメモリに保持した状態で前記監視処理を実行し、前記再配置単位のデータへのアクセス頻度に基づいて前記再配置処理を実行するストレージ制御プログラム。

【請求項 2】

請求項 1 において、

前記監視処理と再配置処理は、評価期間毎に繰り返し実行され、

前記過負荷処理は、前記評価期間の間、前記キャッシュ制御処理にかかわらず前記一部のデータを前記キャッシュメモリに保持するストレージ制御プログラム。

【請求項 3】

請求項 1 において、

前記過負荷処理は、アクセス要求が発生した場合に、前記アクセス要求が発生した前記再配置単位内の前記一部のデータを前記キャッシュメモリに保持するストレージ制御プログラム。

【請求項 4】

請求項 3 において、

前記監視処理は、前記再配置単位内のデータへのアクセス頻度を、前記キャッシュメモリに保持されている前記一部のデータへのアクセスと前記再配置単位内のデータへのアクセスの合計数に基づいて計測するストレージ制御プログラム。

【請求項 5】

請求項 3 において、

前記過負荷処理は、前記過負荷状態にある第 2 の記憶装置内の少なくとも最もアクセス頻度が高い再配置単位に対して、前記アクセス要求が発生した前記一部のデータを前記キャッシュメモリに書き込み保持するよう前記ストレージに指示するストレージ制御プログラム。

【請求項 6】

請求項 5 において、

前記過負荷処理は、前記過負荷状態にある第 2 の記憶装置を通常負荷状態にできるようなアクセス頻度が計測された 1 つのまたは複数の再配置単位に対して、前記キャッシュメモリへの保持を指示するストレージ制御プログラム。

【請求項 7】

請求項 3 において、

前記キャッシュメモリ内の前記一部のデータを前記キャッシュ制御処理にかかわらず保持する容量は、少なくとも 1 つの再配置単位のデータを保持できる容量であるストレージ制御プログラム。

【請求項 8】

請求項 1 において、

前記第 2 の記憶装置は、複数の RAID グループを有し、

前記監視処理は、前記複数の RAID グループそれぞれの物理ディスクの使用率を監視し、

前記第 2 の記憶装置が過負荷状態になる場合は、前記使用率が第 2 の基準値を超える場合を含むストレージ制御プログラム。

【請求項 9】

第 1 のアクセス速度を有する第 1 の記憶装置と、前記第 1 のアクセス速度より遅い第 2 のアクセス速度を有する第 2 の記憶装置と、キャッシュメモリとを有し、前記第 1 または第 2 の記憶装置内のデータへのアクセス要求時に、前記アクセス要求のアドレスのデータが前記キャッシュメモリに記憶されている場合前記キャッシュメモリのデータにアクセスし、記憶されていない場合前記第 1 または第 2 の記憶装置内の前記アクセス要求のデータにアクセスしアクセスデータを前記キャッシュメモリに記憶すると共に前記キャッシュメモリから過去に記憶したデータを削除するキャッシュ制御処理が行われるストレージの制御方法であって、

前記第 1 及び第 2 の記憶装置内の再配置単位のデータへのアクセス頻度を監視する監視

10

20

30

40

50

処理工程と、

前記第2の記憶装置内の前記アクセス頻度が第1の基準値を超える前記再配置単位の前記データを前記第1の記憶装置に再配置する再配置処理工程と、

前記第2の記憶装置が過負荷状態の場合に、前記第2の記憶装置内の少なくとも一部のデータが前記キャッシュ制御処理により前記キャッシュメモリに記憶されたら前記監視処理工程中は前記一部のデータを前記キャッシュメモリに保持し、前記一部のデータへのアクセス要求時に前記キャッシュ制御処理により前記キャッシュメモリ内の前記一部のデータがアクセスされる過負荷処理工程とを有し、

前記一部のデータを前記キャッシュメモリに保持した状態で前記監視処理工程を実行し、前記再配置単位の前記データへのアクセス頻度に基づいて前記再配置処理工程を実行するストレージ制御方法。

10

【請求項10】

第1のアクセス速度を有する第1の記憶装置と、前記第1のアクセス速度より遅い第2のアクセス速度を有する第2の記憶装置と、キャッシュメモリとを有し、前記第1または第2の記憶装置内のデータへのアクセス要求時に、前記アクセス要求のアドレスのデータが前記キャッシュメモリに記憶されている場合前記キャッシュメモリのデータにアクセスし、記憶されていない場合前記第1または第2の記憶装置内の前記アクセス要求のデータにアクセスしアクセスデータを前記キャッシュメモリに記憶すると共に前記キャッシュメモリから過去に記憶したデータを削除するキャッシュ制御処理が行われるストレージと、

前記ストレージの階層制御を行うストレージ階層制御装置とを有し、

20

前記ストレージ階層制御装置は、

前記第1及び第2の記憶装置内の再配置単位の前記データへのアクセス頻度を監視する監視処理部と、

前記第2の記憶装置内の前記アクセス頻度が第1の基準値を超える前記再配置単位の前記データを前記第1の記憶装置に再配置する再配置処理部と、

前記第2の記憶装置が過負荷状態の場合に、前記第2の記憶装置内の少なくとも一部のデータが前記キャッシュ制御処理により前記キャッシュメモリに記憶されたら前記監視工程中は前記一部のデータを前記キャッシュメモリに保持し、前記一部のデータへのアクセス要求時に前記キャッシュ制御処理により前記キャッシュメモリ内の前記一部のデータがアクセスされる過負荷処理部とを有し、

30

前記一部のデータを前記キャッシュメモリに保持した状態で前記監視処理部が前記監視を実行し、前記再配置単位の前記データへのアクセス頻度に基づいて前記再配置処理部が前記再配置を実行するストレージシステム。

【請求項11】

第1のアクセス速度を有する第1の記憶装置と、前記第1のアクセス速度より遅い第2のアクセス速度を有する第2の記憶装置と、キャッシュメモリとを有し、前記第1または第2の記憶装置内のデータへのアクセス要求時に、前記アクセス要求のアドレスのデータが前記キャッシュメモリに記憶されている場合前記キャッシュメモリのデータにアクセスし、記憶されていない場合前記第1または第2の記憶装置内の前記アクセス要求のデータにアクセスしアクセスデータを前記キャッシュメモリに記憶すると共に前記キャッシュメモリから過去に記憶したデータを削除するキャッシュ制御処理が行われるストレージの階層制御を行うストレージ階層制御装置であって、

40

前記第1及び第2の記憶装置内の再配置単位の前記データへのアクセス頻度を監視する監視処理部と、

前記第2の記憶装置内の前記アクセス頻度が第1の基準値を超える前記再配置単位の前記データを前記第1の記憶装置に再配置する再配置処理部と、

前記第2の記憶装置が過負荷状態の場合に、前記第2の記憶装置内の少なくとも一部のデータが前記キャッシュ制御処理により前記キャッシュメモリに記憶されたら前記監視工程中は前記一部のデータを前記キャッシュメモリに保持し、前記一部のデータへのアクセス要求時に前記キャッシュ制御処理により前記キャッシュメモリ内の前記一部のデータがク

50

セスされる過負荷処理部とを有し、

前記一部のデータを前記キャッシュメモリに保持した状態で前記監視処理部が前記監視を実行し、前記再配置単位のデータへのアクセス頻度に基づいて前記再配置処理部が前記再配置を実行するストレージ階層制御装置。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、ストレージ制御プログラム、ストレージ制御方法、ストレージシステム及びその階層制御装置に関する。

【背景技術】

10

【0002】

ストレージシステムは、高速アクセスと大容量と低コストの要求に応えることが求められている。そのために、ストレージシステムは、ビット単価は高いが高速アクセス可能な階層の記憶装置と、低速アクセスだがビット単価が低い階層の記憶装置とを有し、アクセス頻度の高いデータは高速アクセス可能な階層の記憶装置に格納し、アクセス頻度の低いデータは低速だがビット単価が低い大容量の階層の記憶装置に格納する自動階層制御を行う。

【0003】

自動階層制御では、高速階層と低速階層の再配置単位のデータに対するアクセス頻度を監視し、低速階層内のアクセス頻度の高い再配置単位のデータを高速階層に再配置し、高速階層内のアクセス頻度の低い再配置単位のデータを低速階層に再配置する。

20

【0004】

自動階層制御については、特許文献に記載されている。

【先行技術文献】

【特許文献】

【0005】

【特許文献1】特開2011-227898号公報

【特許文献2】特開平8-77073号公報

【特許文献2】特開平6-309224号公報

【発明の概要】

30

【発明が解決しようとする課題】

【0006】

しかしながら、上記の自動階層制御では、低速階層の記憶装置へのアクセスが急増し性能限界に達すると、アクセスから応答までの応答時間が長くなるレスポンス遅延が発生する。その結果、低速階層内の再配置単位に対するアクセス頻度が、実際に要求されているアクセス頻度よりも低く抑えられてしまう。そのため、より高速な階層に再配置すべきデータが低速階層に留まり続けて、全体の応答時間が最適化されない場合がある。つまり、ストレージシステム全体では性能的に余裕があっても、局所的に過負荷状態になって性能限界に達すると、本来要求されているアクセス頻度を正しく評価することができず、望ましい再配置制御を行うことができない。

40

【0007】

そこで、本発明の目的は、ある一つの側面において、過負荷状態になっても適切に再配置制御を行うストレージ制御プログラム、ストレージ制御方法、ストレージシステム及びその階層制御装置を提供することにある。

【課題を解決するための手段】

【0008】

ストレージ制御プログラムの第1の側面によれば、第1のアクセス速度を有する第1の記憶装置と、前記第1のアクセス速度より遅い第2のアクセス速度を有する第2の記憶装置と、キャッシュメモリとを有するストレージの制御処理をコンピュータに実行させるストレージ制御プログラムであって、

50

前記制御処理は、
前記第 1 及び第 2 の記憶装置内のデータへのアクセス頻度を監視する監視処理と、
前記第 2 の記憶領域内の前記アクセス頻度が第 1 の基準値を超えるデータを前記第 1 の記憶装置に再配置する再配置処理と、
前記第 2 の記憶装置が過負荷状態の場合に、前記第 2 の記憶領域内の少なくとも一部のデータを前記キャッシュメモリに保持する過負荷処理とを有し、
前記一部のデータを前記キャッシュメモリに保持した状態で前記監視処理を実行し、前記一部のデータへのアクセス頻度に基づいて前記再配置処理を実行する。

【発明の効果】

【0009】

10

本実施の形態の第 1 の側面によれば、ストレージの階層制御を適切に行うことができる。
(具体的な効果は、実施の形態の最後に追記しました。)

【図面の簡単な説明】

【0010】

【図 1】本実施の形態におけるストレージシステムの概略構成図である。

【図 2】本実施の形態におけるストレージ制御の階層制御処理の概略フローチャート図である。

【図 3】本実施の形態におけるストレージシステムの構成図である。

【図 4】ストレージ装置 10 内の記憶装置群 14 の構成例を示す図である。

【図 5】ストレージ装置内のコントローラ 12 の構成を示す図である。

20

【図 6】ストレージ階層制御装置 20 の構成を示す図である。

【図 7】正常なストレージ階層制御を説明する図である。

【図 8】正常なストレージ階層制御を説明する図である。

【図 9】正常でないストレージ階層制御を説明する図である。

【図 10】正常でないストレージ階層制御を説明する図である。

【図 11】本実施の形態によるストレージ階層制御が適用された場合のアクセス頻度と負荷状態の一例を示す図である。

【図 12】評価期間 t_2 における記憶装置群内の RAID グループの負荷状態 LOAD と、再配置単位のアクセス頻度 AR の一例を示す図である。

【図 13】評価期間 t_3 における記憶装置群内の RAID グループの負荷状態 LOAD と、再配置単位のアクセス頻度 AR の一例を示す図である。

30

【図 14】過負荷制御処理 S4 のフローチャート図である。

【図 15】SSD キャッシュ制御部の処理を示すフローチャート図である。

【図 16】過負荷状態の RAID グループのテーブルの一例を示す図である。

【図 17】キャッシュ保持候補テーブルの一例を示す図である。

【図 18】SSD キャッシュ保持対象テーブルの一例を示す図である。

【図 19】ストレージ階層制御装置 20 の構成図である。

【発明を実施するための形態】

【0011】

[本実施の形態の概略説明]

40

図 1 は、本実施の形態におけるストレージシステムの概略構成図である。図 2 は、本実施の形態におけるストレージ制御の階層制御処理の概略フローチャート図である。図 1、2 を参照して、本実施の形態の概略を説明する。その後詳細を説明する。

【0012】

ストレージシステム 10 は、ストレージエリアネットワーク (SAN) であるファイバチャネル FC を介して、図示しない複数のサーバからアクセス要求を受信するコントローラ 12 と、キャッシュメモリユニット 13 と、ハードウェア資源である記憶装置群 14 とを有する。

【0013】

コントローラ 12 は、ハードウェア資源である記憶装置群 14 を仮想的な論理ボリューム

50

ムLB1, LB2に割り当て、論理ボリュームそれぞれに対するアクセス要求を受け、上記記憶装置群の割り当てにしたがって記憶装置群内のRAIDグループに対してアクセス要求を行う。また、コントローラ12は、アクセス要求に対してキャッシュメモリユニット13を利用してキャッシュ制御を行う。

【0014】

一般的なキャッシュ制御では、例えば、コントローラ12は、アクセス要求に対してそのアドレスのデータがキャッシュメモリユニット13内に記憶されているか否かをチェックし、キャッシュヒットすればキャッシュメモリユニット13内のデータを読み出して応答し、キャッシュヒットしなければ、記憶装置群14内にアクセスしてデータを読み出して応答する。そして、キャッシュヒットしないで読み出されたデータをキャッシュメモリユニット13内に格納する。その際、キャッシュメモリユニット13の空き容量が不足する場合は、過去に格納したデータを削除する。

10

【0015】

記憶装置群14は、物理的なハードディスクなどの階層構造を有する。記憶装置群14は、例えば、アクセス要求から応答までのレスポンス時間が短い高速アクセスの記憶装置を有する上位の階層L1のサブプールSP1と、それより遅い低速アクセスの記憶装置を有する下位の階層L2のサブプールSP2とを有する。上位の階層L1の記憶装置は、速度重視の記憶装置であり、ビット単価が高いのに対して、下位の階層L2の記憶装置は、コスト重視の記憶装置であり、低速であるがビット単価が低く、大容量である。したがって、階層L2は、階層L1よりも大容量になるように構成される。階層構造は、図1の2層に限らず、3層以上であってもよい。その場合、記憶装置群14は、より高速でビット単価がより高い上位の階層からより低速でビット単価がより低い下位の階層までが、3層以上を有する。

20

【0016】

記憶装置群14内の記憶装置は、物理ディスクの一例として、RAIDグループRAID11, RAID12を有する。ただし、高速アクセスの階層L1の記憶装置が、フラッシュメモリなどのSSD (Solid State Disk)であってもよい。また、物理ディスクとしてRAIDグループを構成しないで単一の物理ディスクであってもよい。

【0017】

このように、記憶装置群14は、速度重視の記憶装置の階層L1と、低速だが低コストで大容量化できる記憶装置の階層L2とによる階層構造を有する。そして、ストレージ階層制御装置20が、アクセス頻度が高いデータは速度重視の階層L1の記憶装置に再配置し、アクセス頻度が低いデータはコスト重視の階層L2の記憶装置に再配置するという自動階層制御を行う。したがって、記憶装置群14は、この自動階層制御により再配置される単位RU111-RU11m, RU211-RU21nでそのアクセス頻度が計測され、それぞれのアクセス頻度に応じて適切な階層L1, L2に再配置される。

30

【0018】

このような自動階層制御を行うために、コントローラ12は、アクセス要求を監視して、RAIDグループ(または物理ディスク。以下同様。)内の再配置単位RUのアクセス頻度を計測する。そして、ストレージ階層制御装置20が、アクセス頻度が高い再配置単位RUのデータをより低速の階層L2からより高速の階層L1に再配置し、アクセス頻度が低い再配置単位RUのデータをより高速の階層L1からより低速の階層L2に再配置する。

40

【0019】

このような自動階層制御において、低速階層L2のデータに対するアクセスが急増し、RAIDグループRAID21の物理ディスクの使用率が急増して過負荷状態になり性能限界に達すると、レスポンス遅延が発生する。それにより、低速階層L2のRAIDグループRAID21内の再配置単位RUへのアクセス頻度が抑えられてしまう。その結果、それらに対するアクセス頻度を適切に評価することができず、実際に要求されているサーバからのアクセス頻度が高いデータにもかかわらず、高速階層L1への再配置処理が行われず、低速階層L2にとどまり続ける。その結果、ストレージ装置全体では性能的に余裕があっても、上記のような局所的な過負荷状態が発生すると、望ましい階層制御による再配置処理ができなくなる。

50

【 0 0 2 0 】

そこで、本実施の形態では、低速階層L2内の過負荷状態になったRAIDグループ内のデータを、アクセス頻度の評価期間の間だけキャッシュメモリユニットに保持する。その場合、保持されたデータは、通常のキャッシュ制御によって追い出されることはない。このように過負荷状態のRAIDグループ内のデータをキャッシュメモリユニットに保持することで、そのデータに対するアクセス頻度を正しく計測することができる。また、そのRAIDグループの過負荷状態を一時的に解消することもでき、そのRAIDグループ内のデータへの実際に要求されているアクセス頻度も計測可能になる。

【 0 0 2 1 】

そして、上記の評価期間でアクセス頻度が高いと判定された場合は、その再配置単位RUのデータを、高速階層L1に再配置する。その結果、要求されているアクセス頻度は高いが所属しているRAIDグループが過負荷状態になって実際に実行されるアクセス頻度が上がらず、低速階層L2内に滞留し続けるという望ましくない状態を抑制することができる。

10

【 0 0 2 2 】

図2は、ストレージ階層制御装置20による階層制御処理の概略フローチャート図である。ストレージ階層制御装置20は、所定の評価期間のサイクルで図2の工程S1-S4を繰り返す。まず、ストレージ階層制御装置20は、監視処理S1でストレージ装置10内のコントローラ12から、前の評価期間での各再配置単位RU毎のアクセス頻度と、各RAIDグループの負荷（物理ディスクの使用率）とを有する監視データを取得して監視する。つまり、コントローラ12は、そのような再配置単位毎のアクセス頻度とRAIDグループ毎の負荷状態を計測しているので、その計測データを取得して監視する。

20

【 0 0 2 3 】

ここで、物理ディスクの使用率とは、単位時間当たりの物理ディスクに対するアクセス要求から応答までの使用時間が占める割合である。したがって、アクセス速度が遅い物理ディスクの場合は、アクセス頻度が低くても使用率が高くなり、使用率が例えば100%に達すると性能限界に達し過負荷状態になる。複数の物理ディスクで構成されるRAIDグループの場合は、全ての物理ディスクの使用率を計測して、RAIDグループの使用率を監視する。ただし、RAIDグループの構成上、アクセス要求に対応して複数の物理ディスクが並列にアクセスされる場合は、単一の物理ディスクの使用率を計測することでも、RAIDグループの使用率を監視することができる。

30

【 0 0 2 4 】

次に、ストレージ階層制御装置20は、監視データに基づく評価処理S2で、再配置単位RUのアクセス頻度に基づいて、再配置すべき再配置単位の候補を決定する。具体的には、アクセス頻度が第1の基準値を超えている下位階層の再配置単位が、上位階層への再配置の候補となり、アクセス頻度が第2の基準値以下の上位階層の再配置単位が、下位階層への再配置の候補となる。

【 0 0 2 5 】

そして、ストレージ階層制御装置20は、再配置処理S3にて、再配置の候補の再配置単位RUのデータを、それぞれ別の階層に再配置する。

【 0 0 2 6 】

また、ストレージ階層制御装置20は、過負荷制御処理S4にて、過負荷状態のRAIDグループ内の再配置単位のうち、相対的にアクセス頻度が高い再配置単位をキャッシュメモリユニット内に保持するキャッシュ候補として、ストレージ装置10のコントローラ12に通知する。

40

【 0 0 2 7 】

これに伴い、次の評価期間では、コントローラ12は、キャッシュ候補の再配置単位のデータがアクセスされた時に、そのアクセスされたデータをキャッシュメモリユニットに格納し、その評価期間の間、通常のキャッシュ制御にかかわらず保持を継続する。つまり、別のデータがキャッシュメモリに格納されたことに伴ってキャッシュ候補で保持されているデータがキャッシュメモリユニットからキャッシュあふれされることを禁止し、強制

50

的な保持状態を維持する。そして、コントローラ12は、キャッシュメモリユニットでヒットした場合も含めて、再配置単位毎のアクセス頻度を計測する。

【0028】

このように、過負荷状態のRAIDグループ内のアクセス頻度が高い再配置単位のデータを、評価期間の間キャッシュメモリユニット内に強制的に保持することで、RAIDグループが過負荷状態であっても、キャッシュメモリユニットを介してアクセス動作を行うことができるので、本来要求されているアクセス頻度を適切に計測することができる。また、過負荷状態であったRAIDグループが過負荷状態から抜け出すことができる場合は、キャッシュメモリユニットに保持されていないデータに対するアクセス要求はもはや抑制されないもので、そのアクセス頻度も適切に計測することができる。これにより、階層制御を適切に行うことができる。

10

【0029】

また、ストレージ階層制御装置20は、過負荷制御処理S4にて、評価期間の間キャッシュメモリユニットに保持していたデータに対する再配置単位をキャッシュ候補から削除する。つまり、評価期間を経過して適切にアクセス頻度が計測された再配置単位のデータは、通常のキャッシュ制御により適宜削除される。

【0030】

[実施の形態の詳細な説明]

次に、本実施の形態について詳細に説明する。上記概略説明と重複する説明は適宜省略する。

20

【0031】

図3は、本実施の形態におけるストレージシステムの構成図である。図1と同様に、ストレージシステムは、複数のサーバからファイバーチャネルFC(SAN)を介してアクセスされるストレージ装置10と、ストレージ階層制御装置20とを有する。図3において、図1と異なる構成は、ストレージ装置10において、キャッシュメモリユニット13が、DRAMなどのビット単価が高いが高速のキャッシュメモリ131と、SSDなどのビット単価が安い大容量フラッシュメモリのSSDキャッシュ132とを有することと、コントローラ12に加えてSSDキャッシュ132を制御するSSDキャッシュ制御部121を有することと、物理的な記憶装置群14の下位階層L2内に複数のRAIDグループRAID21,RAID22を有することである。

30

【0032】

SSDキャッシュ132は、キャッシュメモリ131を拡張するメモリであり、ストレージ自動階層制御と併用することができる。SSDキャッシュ制御部121は、キャッシュメモリ131に格納しているデータがあふれた場合に、SSDキャッシュ132にデータを書き込んで保持し、キャッシュメモリ131の容量があたかも拡大したように制御する。

【0033】

そして、通常のキャッシュ制御では、アクセスされたデータは最初にキャッシュメモリ131に格納され保持され、キャッシュ制御の結果保持していたデータがあふれる場合には、SSDキャッシュ132に格納され保持される。そして、SSDキャッシュ132からあふれる場合は、キャッシュメモリユニット13から削除される。

40

【0034】

また、アクセス要求に応答して、コントローラ12は、最初にキャッシュメモリ131内のアドレスをチェックしてキャッシュ判定を行い、ヒットすればそのデータを読み出して応答し、ヒットしなければ、SSDキャッシュ132内のアドレスをチェックしてキャッシュ判定を行い、ヒットすればそのデータを読み出して応答し、ヒットしなければ、記憶装置群14内の対応するアドレスのデータにアクセスを行う。

【0035】

SSDキャッシュ技術を併用する場合も、低速の下位階層L2のデータアクセスが急増しても、高速の上位階層L1の負荷に対する許容値が十分高いので、下位階層L2のアクセス頻度が高速の上位階層L1よりも低く抑えられる可能性が高い。つまり、低速の下位階層のデー

50

タがキャッシュメモリユニット13にキャッシュされても、高速の上位階層へのアクセスによりキャッシュメモリユニット13からあふれることがある。

【0036】

また、SSDキャッシュ132は大容量ではあるが、その容量には限りがあり、上記の理由により低速の下位階層のデータがキャッシュあふれによりキャッシュメモリユニットから削除されやすい。

【0037】

その結果、SSDキャッシュ技術を使用しない場合と同様に、ストレージ自動階層制御において、低速な下位階層へのアクセスの急増により物理ディスクが性能限界に近づくまたは達する過負荷状態になり、レスポンス遅延が発生して、適切なアクセス頻度を計測できないことが生じうる。

10

【0038】

図4は、ストレージ装置20内の記憶装置群14の構成例を示す図である。図3と同様に、記憶装置群14は、高速の上位階層L1のサブプールSP1と、低速の下位階層L2のサブプールSP2とを有する。上位階層L1のサブプールSP1には、例えば5つのハードディスクHDDを有するRAIDグループRAID11を有する。また、下位階層L2のサブプールSP2には、同じく5つのハードディスクHDDを有する2つのRAIDグループRAID21,RAID22を有する。つまり、下位階層L2のほうが大容量である。

【0039】

RAIDグループは、アクセス要求に対して、5つのHDDに対して同時にアクセス動作を実行する。したがって、5つのHDDのいずれかのHDDに対するアクセス要求から応答までの使用時間を監視することで、各RAIDグループの物理ディスクに対する使用率（またはビジー率）を計測することができる。所定期間中の物理ディスクに対する使用時間の合計に対する所定期間の割合が高いほど、使用率が高くなる。使用率が100に達すると、アクセス要求は待ち行列内にとどまり、物理ディスクであるRAIDグループが過負荷状態、つまり性能限界に達することになる。

20

【0040】

一方、論理ボリュームへのアクセス要求は、サーバが要求しているアクセスであり、そのアクセス要求は、コントローラ12により、記憶装置群14内のRAIDグループへのアクセス要求へと変換される。

30

【0041】

なお、記憶装置群14は、3階層以上であってもよく、また、RAIDグループの数も図4と異なっても良い。また、上位の階層のサブプールがSSDで構成されても良い。

【0042】

図5は、ストレージ装置内のコントローラ12の構成を示す図である。コントローラ12は、論理ボリューム制御部122と、キャッシュメモリ制御部123と、再配置単位のアクセス頻度計測部124と、RAIDグループの負荷計測部125と、計測値通知部126とを有する。

【0043】

論理ボリューム制御部122は、物理的な記憶装置群14を仮想的な論理ボリュームに切り出して割り当て、論理ボリュームに対するアクセス要求を記憶装置群14の対応するRAIDグループのアドレスへのアクセス要求に変換するなどの制御を行う。キャッシュメモリ制御部123は、キャッシュメモリ131に対する前述のキャッシュ制御を行う。

40

【0044】

再配置単位のアクセス頻度計測部124は、評価期間中の各再配置単位へのアクセス頻度、つまり単位時間当たりのアクセス回数を計測する。前述のとおり、再配置単位内のデータがキャッシュメモリユニット13に格納されている状態でキャッシュヒットしたアクセス要求も、このアクセス頻度にカウントされる。

【0045】

RAIDグループの負荷計測部125は、評価期間中に各RAIDグループの負荷を示す、物理

50

ディスクの使用率を計測し、過負荷状態になっているか否かを計測する。そして、計測値通知部 126 は、ストレージ階層制御装置 20 からの問い合わせに回答して、計測した再配置単位のアクセス頻度と RAID グループの負荷状態を通知する。

【0046】

図 6 は、ストレージ階層制御装置 20 の構成を示す図である。ストレージ階層制御部 20 は、図 2 で説明した 4 つの処理 S1-S4 を行う。すなわち、ストレージ階層制御部 20 は、監視処理 S1 を実行する監視部 21 と、監視データに対する評価処理 S2 を実行する評価部 22 と、再配置処理を実行する再配置部 23 と、過負荷制御処理を実行する過負荷制御部 24 とを有する。それぞれが行う処理は前述のとおりであるが、後で具体的に説明する。

【0047】

図 19 は、ストレージ階層制御装置 20 の構成図である。ストレージ階層制御装置 20 は、プロセッサである CPU 201 と、メモリ 202 と、ストレージ階層制御プログラムなどのソフトウェアが格納される記憶媒体 204 と、外部とのインターフェース 203 とを有し、それらが内部バスを介して接続されている。プロセッサ 201 が記憶媒体 204 内のストレージ階層制御プログラムを実行することで、ストレージ装置 10 の階層制御を行い、図 6 の監視部 21、評価部 22、再配置部 23、過負荷制御部 24 を構築する。

【0048】

[正常なストレージ階層制御の例]

図 7、図 8 は、正常なストレージ階層制御を説明する図である。図 7 は、ストレージ装置内の記憶装置群内の RAID グループの負荷状態 LOAD、再配置単位のアクセス頻度 AR を示している。また、図 8 は、時間軸に対するアクセス頻度 AR と負荷 LOAD の状態 (性能限界値 LM) を示している。ここでは、下位階層 L2 内の RAID グループ RAID21 内の再配置単位 RU211 が注目されているものとする。

【0049】

図 8 には、下位階層 L2 の性能限界値 LM1 と、再配置単位 RU211 の性能限界値 LM2 とが示され、LM1-LM2 が再配置端子 RU211 以外の総負荷 LOAD1 になる。そして、評価期間 t1 において、再配置単位 RU211 のアクセス頻度 AR のピーク値が 10 であることが示されている。図 7 には、評価期間 t1 で計測した各再配置単位のアクセス頻度 AR が記入されている。これによれば、RAID グループ RAID21 内の再配置単位 RU211 ~ RU215 のアクセス頻度 AR は、10、5、7、0、3 であり、それ以外の再配置単位 RU216-21n の合計アクセス頻度 AR は 25 である。したがって、RAID グループ RAID21 の負荷率は、性能限界のアクセス回数が 100 とすると、 $50 / 100 = 50\%$ となる。この負荷率は過負荷状態ではなく通常負荷状態である。

【0050】

そして、再配置単位 RU211 のアクセス頻度 AR が 10 であり、図 8 に示されるとおり、再配置閾値 $V_{th} (= 8)$ を超えている。したがって、ストレージ階層制御により、次の評価期間 t2 において、再配置単位 RU211 のデータは上位階層 L1 に再配置される。その結果、図 8 に示されるとおり、評価期間 t2 において再配置単位 RU211 内のデータに対するアクセス頻度 AR は大きく増大している。

【0051】

このように、正常なストレージ階層制御によれば、各 RAID グループは過負荷状態になく、評価期間 t1 で計測されたアクセス頻度 AR が閾値 V_{th} を超えていれば、次の評価期間 t2 ではより上位の階層に再配置され、より高速アクセスが可能になり、要求されるアクセス要求の増加に応えることができる。

【0052】

[正常でないストレージ階層制御の例]

図 9、図 10 は、正常でないストレージ階層制御を説明する図である。図 9、図 10 は、前述の図 7、図 8 に対応する図である。

【0053】

図 9 では、下位の階層 L2 内の RAID グループ RAID21 が過負荷状態になり、各再配置単位 RU

10

20

30

40

50

のアクセス頻度ARが低下していることが示されている。各アクセス頻度ARには、過負荷状態の値と、括弧内の正常負荷状態の値（図7と同じ値）が示されている。すなわち、RAIDグループRAID21内の再配置単位RU211～RU215のアクセス頻度ARは、5、3、5、0、1と低下し、それ以外の再配置単位RU216-21nの合計アクセス頻度ARは86と急増している。逆に言えば、再配置単位RU216-21nの合計アクセス頻度ARは86と急増したために、再配置単位RU211～RU215のアクセス頻度ARが低下したともいえる。したがって、RAIDグループRAID21の負荷率は、性能限界のアクセス回数が100とすると、 $100 / 100 = 100\%$ となり、この負荷率は過負荷状態である。

【0054】

特に、再配置単位RU211は、本来要求されているアクセス頻度ARが10であったのに、属するRAIDグループRAID21が過負荷状態になりレスポンス遅延が生じて、実際に実行されたアクセス頻度ARは5に止まっている。このアクセス頻度は、本来サーバなどにより要求されているアクセス頻度とは異なっている。

10

【0055】

図10に示されるとおり、評価期間t1において、再配置単位RU211以外の他の再配置単位RUの総負荷LOAD1が増大したため、再配置単位RU211の性能限界値LM2が低下し、再配置単位RU211のアクセス頻度ARはその性能限界値LM2を超えることができない。その結果、計測されたアクセス頻度ARは、ピーク値で5となり、要求されるアクセス頻度よりも少なくなっている。

【0056】

20

このアクセス頻度AR=5は、再配置閾値Vth=8より少ない。そのため、次の評価期間t2において、再配置単位RU211のデータは上位階層に再配置されず、下位の階層L2内に留まり続ける。これでは、適切なストレージ階層制御を行うことができない。

【0057】

〔本実施の形態によるストレージ階層制御の例〕

図11は、本実施の形態によるストレージ階層制御が適用された場合のアクセス頻度と負荷状態の一例を示す図である。前述の図8、図10に対応する図である。また、図12は、評価期間t2における記憶装置群内のRAIDグループの負荷状態LOADと、再配置単位のアクセス頻度ARの一例を示す図である。図13は、評価期間t3における記憶装置群内のRAIDグループの負荷状態LOADと、再配置単位のアクセス頻度ARの一例を示す図である。

30

【0058】

図12に示した評価期間t1、t2、t3の順に、本実施の形態によるストレージ階層制御が適用された場合の動作について具体的に説明する。

【0059】

〔評価期間t1〕

評価期間t1は、正常でないストレージ階層制御で説明した評価期間t2と同じである。すなわち、RAID21にアクセス要求が急増して性能限界を超えて過負荷状態になり、各再配置単位RUのアクセス頻度が低下し、要求されているアクセス頻度よりも低くなっている。

【0060】

40

本実施の形態のストレージ階層制御は、図2のフローチャートに沿って実行される。評価期間t1の後半で、ストレージ階層制御装置20内の監視部21は、監視処理S1を実行する。すなわち、ストレージ装置10内のコントローラ12から、評価期間t1での各再配置単位RUのアクセス頻度ARと、各RAIDグループの負荷率（物理ディスクの使用率）LOADの計測結果を取得して監視する。

【0061】

取得したアクセス頻度と負荷率は、図9に示されるのと同じである。すなわち、下位の階層L2内の再配置単位RU211は過負荷状態によりそのアクセス頻度ARが5と低く抑えられて計測され、この再配置単位RU211を有するRAIDグループRAID21が過負荷状態であることが計測される。

50

【 0 0 6 2 】

次に、ストレージ階層制御装置 2 0 内の評価部 2 2 は、評価処理S2を実行し、計測された再配置単位のアクセス頻度ARを評価して、再配置の候補リストを作成する。すなわち、下位の階層L2内の第 1 の閾値を超えるアクセス頻度が計測された再配置単位RUを、より上位の階層に再配置する候補に指定し、上位の階層L1内の第 2 の閾値を下回るアクセス頻度が計測された再配置単位RUを、より下位の階層に再配置する候補に指定して、候補リストを作成する。しかし、RAIDグループRAID21内の全ての再配置単位のアクセス頻度が抑制されて計測されているため、再配置単位RU211は再配置候補リストには入らず再配置対象にはならない。

【 0 0 6 3 】

10

そして、ストレージ階層制御装置 2 0 内の再配置部 2 3 は、再配置処理S3を実行する。再配置部 2 3 は、評価部 2 2 で作成した再配置候補リストにしたがって、候補の再配置単位内のデータについて再配置を実行する。ただし、評価期間t1での計測結果によっては再配置候補リストには再配置単位RU211などは入っていない。よって、再配置部 2 3 は評価期間 t 1 では再配置処理を行わない。

【 0 0 6 4 】

次に、ストレージ階層制御装置 2 0 内の過負荷制御部 2 4 は、過負荷制御処理S4を実行する。すなわち、過負荷制御部 2 4 は、過負荷状態のRAIDグループRAID21内に存在する再配置単位RU211-21nのうち、相対的にアクセス頻度ARが高いものをキャッシュ候補としてSSDキャッシュ制御部 1 2 1 に通知する。また、既に、評価期間 t 1 においてキャッシュメモリユニット 1 3 内に強制的に保持されていたデータがあれば、強制保持の終了をSSDキャッシュ制御部 1 2 1 に通知する。

20

【 0 0 6 5 】

図 1 4 は、過負荷制御処理S4のフローチャート図である。過負荷制御部 2 4 は、SSDキャッシュ制御部 1 2 1 にSSDキャッシュ 1 3 2 内での強制的な保持の終了を指示する (S11)。この処理の意味については後述する。そして、過負荷制御部 2 4 は、評価期間 t 1 において過負荷状態になっていたRAIDグループのテーブル (図 1 6) を作成する (S12)。

【 0 0 6 6 】

図 1 6 は、過負荷状態のRAIDグループのテーブルの一例を示す図である。図 9 に示した評価期間 t 1 では、RAIDグループRAID21が過負荷状態と評価されているので、図 1 6 の過負荷状態RAIDグループテーブルには、RAIDグループRAID21の「 2 1 」が記載されている。

30

【 0 0 6 7 】

次に、過負荷制御部 2 4 は、過負荷状態のRAIDグループ毎に、アクセス頻度順に再配置単位を並べたキャッシュ保持候補テーブル (図 1 7) を作成する (S13)。このキャッシュ保持候補テーブルには、少なくとも各RAIDグループの過負荷状態を解消できるだけの再配置単位がリストアップされる。

【 0 0 6 8 】

図 1 7 は、キャッシュ保持候補テーブルの一例を示す図である。図 1 7 には、RAIDグループRAID21の再配置単位RU211, RU213の番号 2 1 1 , 2 1 3 , 2 1 2 , 2 1 5 がリストアップされている。つまり、RAIDグループRAID21は過負荷状態で各再配置単位のアクセス頻度は抑制されてはいるが、その中でもアクセス頻度が高い順に再配置単位がキャッシュ保持候補としてリストアップされる。

40

【 0 0 6 9 】

図 9 に示されるとおり、RAIDグループRAID21内のアクセス頻度ARが高い再配置単位RU211, RU213のアクセス頻度AR, 5 , 5 , を性能限界 1 0 0 から減算すると 9 0 となり、過負荷状態の閾値、例えば 9 2 を下回ることになる。したがって、少なくとも再配置単位RU211, 213をキャッシュ保持候補としてキャッシュ保持候補テーブルにリストアップしておけば、後述するSSDキャッシュ制御でキャッシュ保持される。

【 0 0 7 0 】

次に、過負荷制御部 2 4 は、キャッシュ保持候補テーブルにしたがい、キャッシュ保持

50

候補の再配置単位をSSDキャッシュ制御部 1 2 1 に指示する (S14)。この場合、過負荷制御部 2 4 は、SSDキャッシュ 1 2 1 内のSSDキャッシュ保持領域を満たすまで、キャッシュ保持候補テーブル内の再配置単位をキャッシュ保持候補として指示する。

【 0 0 7 1 】

ここで、SSDキャッシュ 1 2 1 内のSSDキャッシュ保持領域として確保すべき容量について説明する。このSSDキャッシュ保持領域の容量は、少なくとも 1 つの再配置単位の容量を有することが望ましい。SSDキャッシュ保持領域の容量が、1 つの再配置単位の容量と同じ場合は、少なくとも最もアクセス頻度が高かった再配置単位がキャッシュ保持候補としてSSDキャッシュ制御部 1 2 1 に指示される。その結果、後述するSSDキャッシュ制御部による制御で、指定された 1 つの再配置単位内のデータがアクセスの度にSSDキャッシュ 1 3 2 内に保持される。このように保持されれば、少なくとも 1 つの再配置単位のデータに対するアクセスを適正に発生させることができ、適正なアクセス頻度の計測を行うことができる。そして、評価期間毎にキャッシュ保持される再配置単位を変更していけば、時間がかかるかもしれないが、やがて要求されるアクセス頻度が高いデータをより上位の階層に再配置することができる。

10

【 0 0 7 2 】

図 1 5 は、SSDキャッシュ制御部の処理を示すフローチャート図である。最初に、SSDキャッシュ制御部 1 2 1 は、ストレージ階層制御装置 2 0 内の過負荷制御部 2 4 からキャッシュ保持候補の再配置単位の指示を受信する (S21)。ここでの例では、少なくとも再配置単位RU211, RU213がキャッシュ保持候補として指示されるものとする。

20

【 0 0 7 3 】

これに応答して、SSDキャッシュ制御部 1 2 1 は、SSDキャッシュ保持対象テーブル (図 1 8) に、そのキャッシュ保持候補を追加する (S22)。また、SSDキャッシュ制御部 1 2 1 は、過負荷制御部 2 4 からキャッシュ保持終了の指示を受信する (S23)。受信した場合は、SSDキャッシュ制御部 1 2 1 は、SSDキャッシュ保持対象テーブルからそのキャッシュ保持終了対象の再配置単位を削除する (S24)。

【 0 0 7 4 】

図 1 8 は、SSDキャッシュ保持対象テーブルの一例を示す図である。ここの例では、SSDキャッシュ保持対象テーブル内には、RAIDグループRAID21内の再配置単位RU211, RU213がリストアップされている。

30

【 0 0 7 5 】

[評価期間 t_2]

このようにして作成されたSSDキャッシュ保持対象テーブルに基づいて、SSDキャッシュ制御部 1 2 1 は、キャッシュメモリユニット 1 3 内のキャッシュメモリ 1 3 1 からあふれて書き込まれたデータのうち、SSDキャッシュ保持対象テーブルで指定されている再配置単位RU211, RU213内のデータについては、SSDキャッシュ保持領域内に強制的に保持する。そして、次の評価期間 t_2 の間は、他のデータの書き込みによりSSDキャッシュ 1 3 2 からあふれることがないように制御する (S25)。

【 0 0 7 6 】

具体的には、SSDキャッシュ制御部 1 2 1 は、SSDキャッシュ保持対象テーブルにリストアップされている再配置単位RU211, RU213内へのアクセスが発生した場合に、キャッシュメモリ 1 3 1 に書き込み、評価期間 t_2 の間は、少なくともSSDキャッシュ 1 3 2 からあふれて削除されないように制御する。つまり、通常のキャッシュ制御により、アクセスされたデータはキャッシュメモリ 1 3 1 に格納され、そこからあふれたときにSSDキャッシュ 1 3 2 に格納される。しかし、上記の制御により、SSDキャッシュ 1 3 2 からはあふれて削除されない。再配置単位RU211, RU213内へのアクセスが発生しない場合は、そもそもキャッシュ制御の対象にはならない。

40

【 0 0 7 7 】

図 1 2 に評価期間 t_2 の状態が示されている。すなわち、評価期間 t_1 で作成されたSSDキャッシュ保持対象テーブル (図 1 8) に基づいて、SSDキャッシュ制御部 1 2 1 が、再

50

配置単位RU211, RU213へのアクセスがあった場合に、そのアクセス領域を含む所定領域のデータをSSDキャッシュ132内に強制的に保持する。この所定領域のデータは、キャッシュ制御単位の領域のデータである。

【0078】

上記のキャッシュ制御は、具体的には、図12に示されるとおり、再配置単位RU211, RU213にアクセスが発生すると、そのデータが一旦はキャッシュメモリ131に格納され、その後通常のキャッシュ制御によりSSDキャッシュ132に格納され保持される。なお、キャッシュの単位は、必ずしも再配置単位ではなく、それより少ないアクセスデータを含むキャッシュ制御単位の領域のデータである。

【0079】

そして、評価期間t2でも、評価期間t1と同様に、ストレージ階層制御装置20が、図2の監視処理S1と評価処理S2と再配置処理S3と過負荷制御処理S4を実行する。評価期間t2では、再配置単位RU211, 213へのアクセス要求は最初の要求を除いてキャッシュメモリユニットで全てキャッシュヒットし応答処理される。したがって、それらのアクセス頻度ARは要求されるアクセス頻度に対応した頻度になる。図12によれば、評価期間t2において、RAIDグループRAID21内の再配置単位RU211, RU213に対するアクセス頻度は最初の1回だけに止まり、その後のアクセス要求はSSDキャッシュ132で処理されて、そこでのアクセス頻度ARはそれぞれ9, 6となっている。

【0080】

したがって、評価期間t2での再配置単位RU211, 213へのアクセス頻度は、それぞれ1 + 9 = 10, 1 + 6 = 7と適切に計測される。すなわち、図11に示されるとおり、評価期間t2において、SSDキャッシュ保持されている再配置単位RU211へのアクセス頻度は10となり、再配置の閾値Vthを超えることになる。

【0081】

ストレージ階層制御装置20内の監視部21が、評価期間t2での再配置単位毎のアクセス頻度と、RAIDグループ毎の負荷状態の計測結果をコントローラ12から取得して監視する。この計測結果によれば、RAIDグループRAID21はもはや過負荷状態ではないことが判明し、再配置単位RU211のアクセス頻度が再配置の閾値を超えていることが判明する。

【0082】

そして、評価部22が、上記の計測データを評価して、再配置単位RU211を再配置候補にリストアップする。その結果、ストレージ階層制御装置20内の再配置部23が、再配置単位RU211のデータを上位の階層L1内のRAIDグループRAID11内のいずれかの再配置単位内に格納する。この再配置処理により、次の評価期間t3では、再配置単位RU211のデータは上位の階層L1でアクセス処理されるので、サーバが要求する高いアクセス頻度を満たすことができる。

【0083】

さらに、過負荷制御部24は、図14に示した処理を実行する。これにより、評価期間t2でSSDキャッシュ保持されていた再配置単位RU211, RU213の保持状態が解除され(S11)、過負荷状態のRAIDグループのテーブル(図16)からRAIDグループRAID21が削除され(S12)、SSDキャッシュ保持候補テーブルにもリストアップされない(S13)。したがって、SSDキャッシュ制御部へのSSDキャッシュ保持候補の指示は行われない(S14)。

【0084】

[評価期間t3]

図13に示されるとおり、評価期間t3では、SSDキャッシュ制御部121が、過負荷制御部24からの終了指示にตอบสนองして、再配置単位RU211のSSDキャッシュ保持制御を終了し、通常のキャッシュ制御によって、アクセス頻度が下がればSSDキャッシュ132からあふれて削除される。またSSDキャッシュ制御部121は、過負荷制御処理部24から新たなSSDキャッシュ保持候補が指示されないのので、評価期間t3ではSSDキャッシュ保持する対象はない。

【0085】

10

20

30

40

50

そして、下位の階層L2内の再配置単位RU211のデータは、上位の階層L1に再配置されているので、上位の階層L1の高いアクセス速度でアクセス処理されるので、サーバから要求されるアクセス頻度を満たすことができる。そのため、図11に示されるとおり、評価期間t3において、再配置単位RU211内に格納されていたデータのアクセス頻度は急増している。

【0086】

以上の通り、本実施の形態によれば、過負荷状態になっている物理ディスク内の再配置単位のデータを、評価期間の間、通常のSSDキャッシュ制御にかかわらずSSDキャッシュ132内に保持されるようにして、サーバから要求されるアクセス頻度を適正に計測できるようにする。それにより、過負荷状態の物理ディスク内に要求されるアクセス頻度が高い再配置単位のデータが下位の階層L2内に留まり続けることが回避され、アクセス頻度を適正に計測でき、適切にストレージ階層制御を行うことができる。

10

【0087】

したがって、閑散期と繁忙期を繰り返すような業務や、長期間停止していた業務が急に再開されるような季節性業務に対して、ストレージ自動階層制御を適用する際に、慣らし運転や手動による再配置などの煩雑な運用を回避することができる。

【0088】

上記の実施の形態において、キャッシュメモリユニット13は、DRAMのキャッシュメモリ131だけで構成されてもよい。その場合は、SSDキャッシュ制御部の代わって、コントローラ12のキャッシュ制御において、候補に指定された再配置領域RU211のアクセスデータをキャッシュメモリ131内の強制的に保持することで、同様にアクセス頻度を適切に計測し、適切なストレージ階層制御を行う。

20

【0089】

上記の実施の形態において、物理ディスクは、RAID構成でなく、1個のHDDで構成されていてもよい。いずれにしても、物理ディスクの使用率に基づいて、それが過負荷状態か否かが判定される。

【0090】

以上の実施の形態をまとめると、以下の付記のとおりである。

【0091】

(付記1)

第1のアクセス速度を有する第1の記憶装置と、前記第1のアクセス速度より遅い第2のアクセス速度を有する第2の記憶装置と、キャッシュメモリとを有するストレージの制御処理をコンピュータに実行させるストレージ制御プログラムであって、

30

前記制御処理は、

前記第1及び第2の記憶装置内のデータへのアクセス頻度を監視する監視処理と、

前記第2の記憶領域内の前記アクセス頻度が第1の基準値を超えるデータを前記第1の記憶装置に再配置する再配置処理と、

前記第2の記憶装置が過負荷状態の場合に、前記第2の記憶領域内の少なくとも一部のデータを前記キャッシュメモリに保持する過負荷処理とを有し、

前記一部のデータを前記キャッシュメモリに保持した状態で前記監視処理を実行し、前記一部のデータへのアクセス頻度に基づいて前記再配置処理を実行するストレージ制御プログラム。

40

【0092】

(付記2)

付記1において、

さらに、前記第1または第2の記憶装置内のアクセス要求が行われた第1のデータを前記キャッシュメモリに書き込み、第2のデータの前記キャッシュメモリへの書き込みにより前記第1のデータを前記キャッシュメモリから削除するキャッシュ制御処理を有し、

前記監視処理と再配置処理は、評価期間毎に繰り返し実行され、

前記過負荷処理は、前記評価期間の間、前記キャッシュ制御処理にかかわらず前記一部

50

のデータを前記キャッシュメモリに保持するストレージ制御プログラム。

【 0 0 9 3 】

(付記 3)

付記 1 において、

前記再配置処理は、前記第 1 及び第 2 の記憶装置内の再配置単位内のデータを、前記再配置単位のデータへのアクセス頻度に応じて再配置し、

前記過負荷処理は、アクセス要求が発生した場合に、前記アクセス要求が発生した前記再配置単位内の前記一部のデータを前記キャッシュメモリに保持するストレージ制御プログラム。

【 0 0 9 4 】

10

(付記 4)

付記 3 において、

前記監視処理は、前記再配置単位内のデータへのアクセス頻度を、前記キャッシュメモリに保持されている前記一部のデータへのアクセスと前記再配置単位内のデータへのアクセスの合計数に基づいて計測するストレージ制御プログラム。

【 0 0 9 5 】

(付記 5)

付記 3 において、

前記過負荷処理は、前記過負荷状態にある第 2 の記憶装置内の少なくとも最もアクセス頻度が高い再配置単位に対して、前記アクセス要求が発生した前記一部のデータを前記キャッシュメモリに書き込み保持するよう前記ストレージに指示するストレージ制御プログラム。

20

【 0 0 9 6 】

(付記 6)

付記 5 において、

前記過負荷処理は、前記過負荷状態にある第 2 の記憶装置を通常負荷状態にできるようにアクセス頻度が計測された 1 つのまたは複数の再配置単位に対して、前記キャッシュメモリへの保持を指示するストレージ制御プログラム。

【 0 0 9 7 】

(付記 7)

30

付記 3 において、

前記キャッシュメモリ内の前記一部のデータを前記キャッシュ制御処理にかかわらず保持する容量は、少なくとも 1 つの再配置単位のデータを保持できる容量であるストレージ制御プログラム。

【 0 0 9 8 】

(付記 8)

付記 1 において、

前記第 2 の記憶装置は、複数の RAID グループを有し、

前記監視処理は、前記複数の RAID グループそれぞれの物理ディスクの使用率を監視し、

前記第 2 の記憶装置が過負荷状態になる場合は、前記使用率が第 2 の基準値を超える場合を含むストレージ制御プログラム。

40

【 0 0 9 9 】

(付記 9)

付記 1 において、

前記第 1 及び第 2 の記憶装置は、複数の論理ボリュームに割り当てられ、

前記監視処理は、前記論理ボリュームへのアクセス処理に基づいて、前記再配置単位のアクセス頻度を計測するストレージ制御プログラム。

【 0 1 0 0 】

(付記 1 0)

第 1 のアクセス速度を有する第 1 の記憶装置と、前記第 1 のアクセス速度より遅い第 2

50

のアクセス速度を有する第２の記憶装置と、キャッシュメモリとを有するストレージの制御方法であって、

前記第１及び第２の記憶装置内のデータへのアクセス頻度を監視する監視処理工程と、
前記第２の記憶領域内の前記アクセス頻度が第１の基準値を超えるデータを前記第１の記憶装置に再配置する再配置処理工程と、

前記第２の記憶装置が過負荷状態の場合に、前記第２の記憶領域内の少なくとも一部のデータを前記キャッシュメモリに保持する過負荷処理工程とを有し、

前記一部のデータを前記キャッシュメモリに保持した状態で前記監視処理工程を実行し、前記一部のデータへのアクセス頻度に基づいて前記再配置処理工程を実行するストレージ制御方法。

10

【０１０１】

（付記１１）

第１のアクセス速度を有する第１の記憶装置と、前記第１のアクセス速度より遅い第２のアクセス速度を有する第２の記憶装置と、キャッシュメモリとを有するストレージと、
前記ストレージの階層制御を行うストレージ階層制御装置とを有し、

前記ストレージ階層制御装置は、

前記第１及び第２の記憶装置内のデータへのアクセス頻度を監視する監視処理部と、

前記第２の記憶領域内の前記アクセス頻度が第１の基準値を超えるデータを前記第１の記憶装置に再配置する再配置処理部と、

前記第２の記憶装置が過負荷状態の場合に、前記第２の記憶領域内の少なくとも一部のデータを前記キャッシュメモリに保持する過負荷処理部とを有し、

20

前記一部のデータを前記キャッシュメモリに保持した状態で前記監視処理部が前記監視を実行し、前記一部のデータへのアクセス頻度に基づいて前記再配置処理部が前記再配置を実行するストレージシステム。

【０１０２】

（付記１２）

付記１１において、

前記キャッシュメモリは、第３のアクセス速度を有する第１のキャッシュメモリと、前記第３のアクセス速度より遅い第４のアクセス速度を有するSSDキャッシュとを有するストレージシステム。

30

【０１０３】

（付記１３）

第１のアクセス速度を有する第１の記憶装置と、前記第１のアクセス速度より遅い第２のアクセス速度を有する第２の記憶装置と、キャッシュメモリとを有するストレージの階層制御を行うストレージ階層制御装置であって、

前記第１及び第２の記憶装置内のデータへのアクセス頻度を監視する監視処理部と、

前記第２の記憶領域内の前記アクセス頻度が第１の基準値を超えるデータを前記第１の記憶装置に再配置する再配置処理部と、

前記第２の記憶装置が過負荷状態の場合に、前記第２の記憶領域内の少なくとも一部のデータを前記キャッシュメモリに保持する過負荷処理部とを有し、

40

前記一部のデータを前記キャッシュメモリに保持した状態で前記監視処理部が前記監視を実行し、前記一部のデータへのアクセス頻度に基づいて前記再配置処理部が前記再配置を実行するストレージ階層制御装置。

【符号の説明】

【０１０４】

１０：ストレージ装置

１２：コントローラ

１３：キャッシュメモリユニット（キャッシュメモリ）

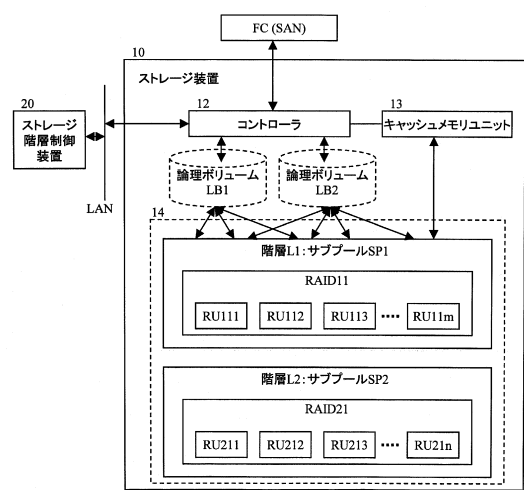
１４：記憶装置群

RAID：RAIDグループ

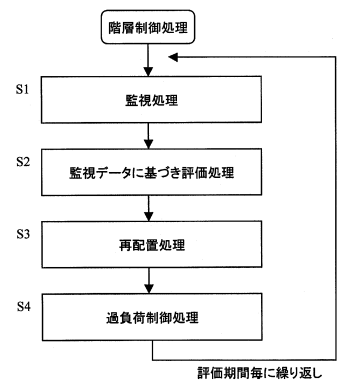
50

RU：再配置単位
20：ストレージ階層制御装置

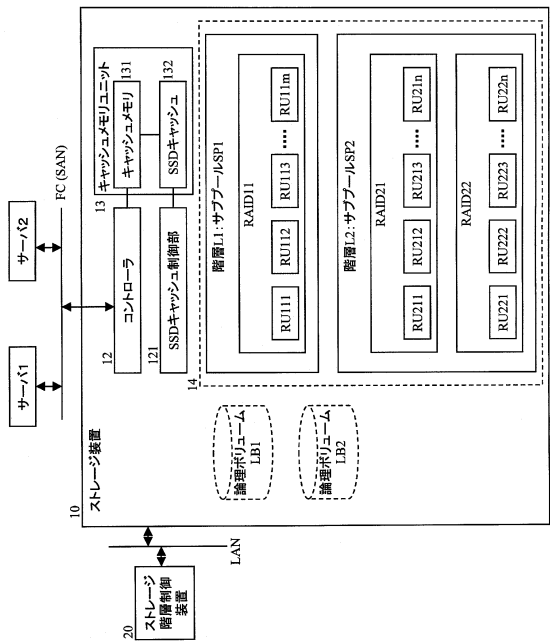
【図 1】



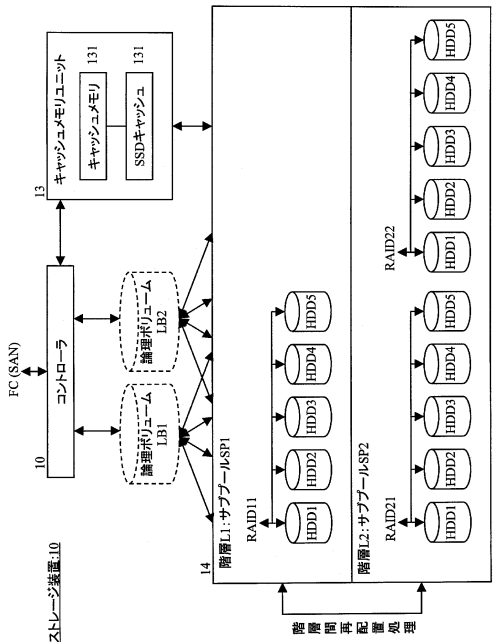
【図 2】



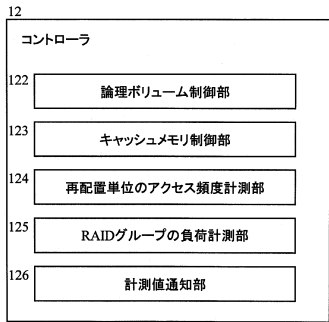
【図 3】



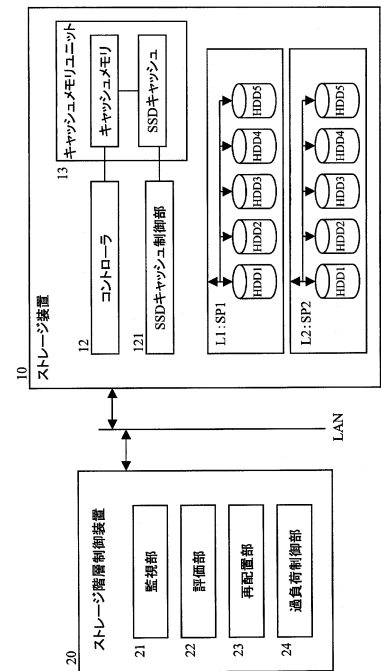
【図 4】



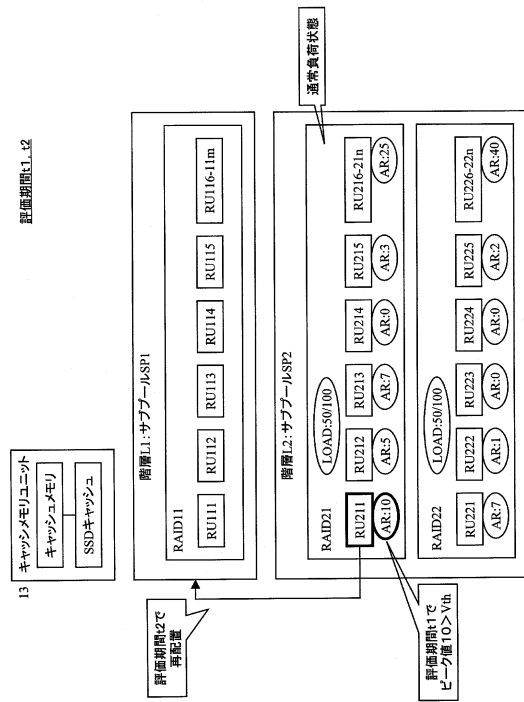
【図 5】



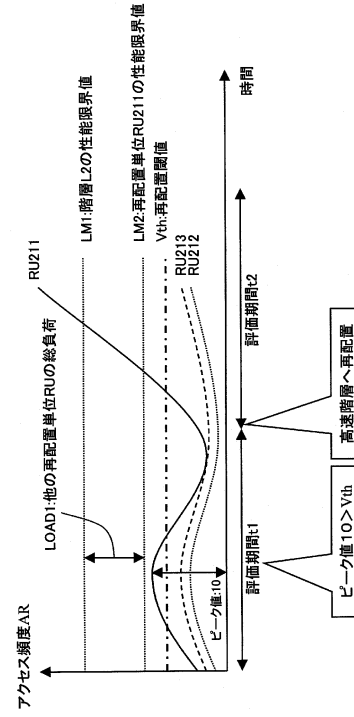
【図 6】



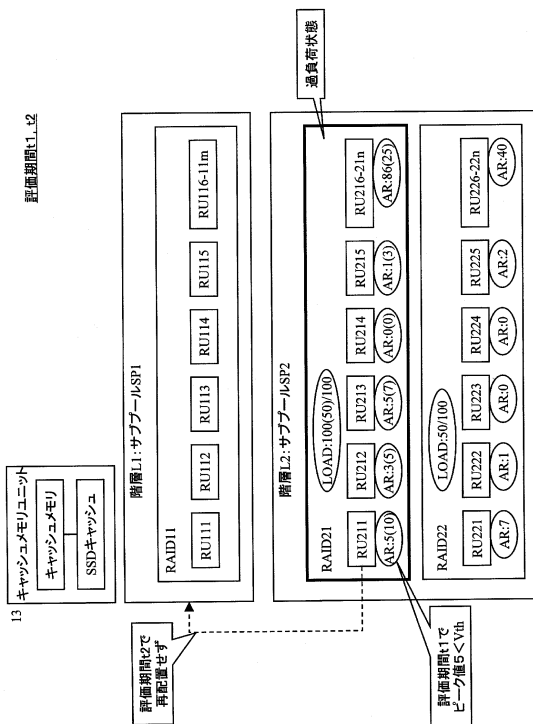
【図 7】



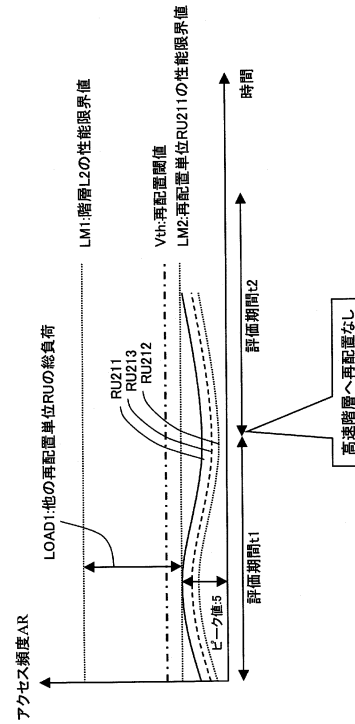
【図 8】



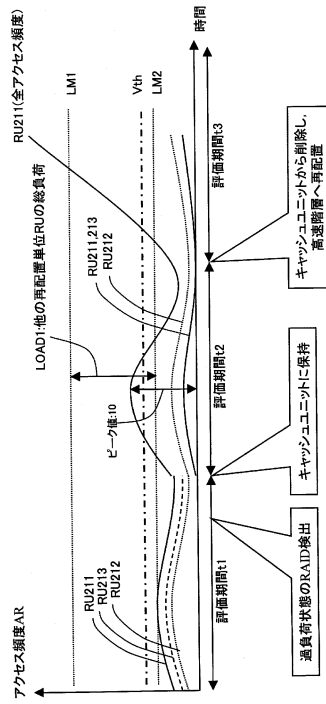
【図 9】



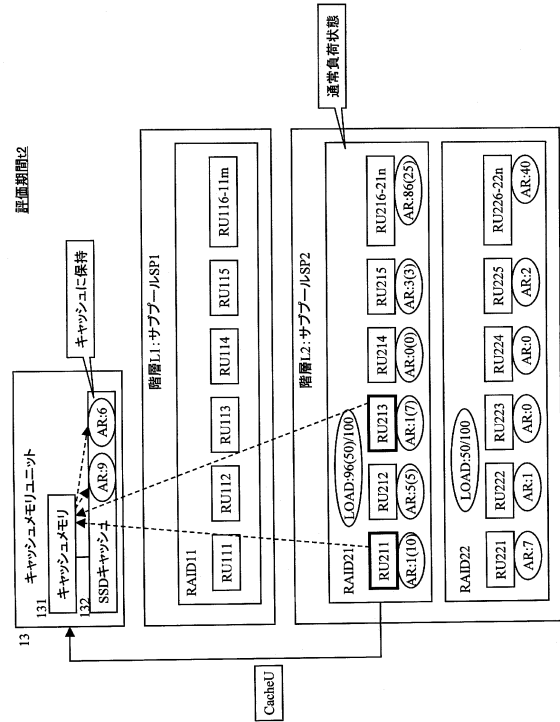
【図 10】



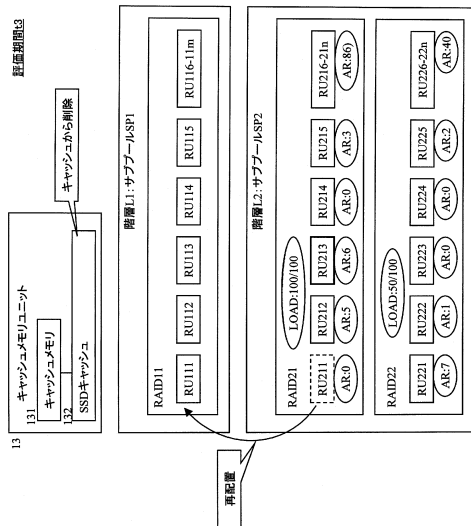
【図 1 1】



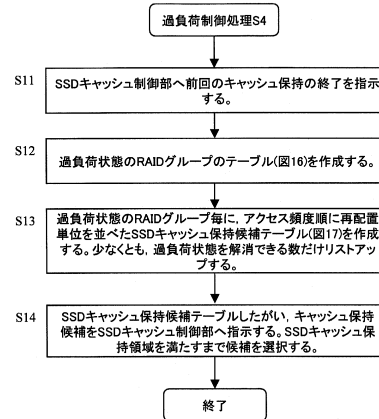
【図 1 2】



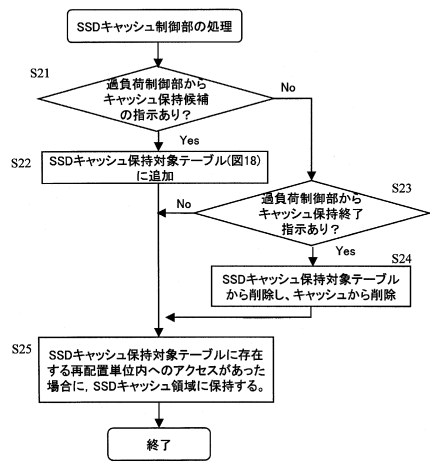
【図 1 3】



【図 1 4】



【図 15】



【図 16】

過負荷状態RAIDグループテーブル

RAIDグループのID
21
y

z

【図 17】

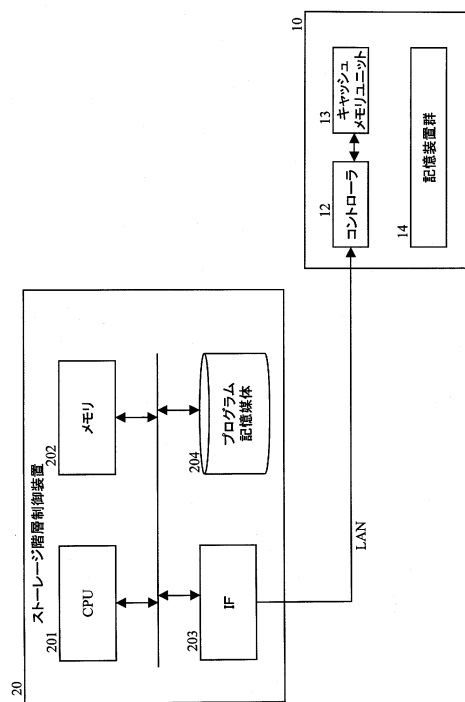
SSDキャッシュ保持候補テーブル

RAIDグループのID	保持候補の再配置単位のID
21	211
21	213
21	212
21	215
--	--
z	z2

【図 18】

SSDキャッシュ保持対象テーブル

RAIDグループのID	保持対象の再配置単位のID
21	211
21	213
y	--
y	--
--	--
z	--



フロントページの続き

(51)Int.Cl. F I
G 0 6 F 12/08 5 5 7

(72)発明者 邵 春友
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 リーボ ピエール ローラン
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 田中 啓介

(56)参考文献 米国特許出願公開第2012/0246386(US,A1)
特開平05-019981(JP,A)

(58)調査した分野(Int.Cl.,DB名)
G 0 6 F 3 / 0 6 - 3 / 0 8
1 2 / 0 8 - 1 2 / 1 2