

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6437656号
(P6437656)

(45) 発行日 平成30年12月12日(2018.12.12)

(24) 登録日 平成30年11月22日(2018.11.22)

(51) Int.Cl.

F I

G 0 6 F 12/00 (2006.01)

G 0 6 F 12/00 5 3 1 D

請求項の数 15 (全 25 頁)

(21) 出願番号	特願2017-532241 (P2017-532241)	(73) 特許権者	000005108
(86) (22) 出願日	平成27年7月31日(2015.7.31)		株式会社日立製作所
(86) 国際出願番号	PCT/JP2015/071736		東京都千代田区丸の内一丁目6番6号
(87) 国際公開番号	W02017/022002	(74) 代理人	110002365
(87) 国際公開日	平成29年2月9日(2017.2.9)		特許業務法人サンネクト国際特許事務所
審査請求日	平成30年1月15日(2018.1.15)	(72) 発明者	長尾 尚
			東京都千代田区丸の内一丁目6番6号 株
			式会社日立製作所内
		(72) 発明者	杉本 定広
			東京都千代田区丸の内一丁目6番6号 株
			式会社日立製作所内
		(72) 発明者	川口 智大
			東京都千代田区丸の内一丁目6番6号 株
			式会社日立製作所内

最終頁に続く

(54) 【発明の名称】 ストレージ装置、ストレージシステム、ストレージシステムの制御方法

(57) 【特許請求の範囲】

【請求項 1】

第1ストレージ装置を含む複数のストレージ装置を備えるストレージシステムであって、

前記複数のストレージ装置の其々は、管理情報を格納する管理情報格納領域とキャッシュ情報を格納するキャッシュ領域とを有するメモリと、前記キャッシュ領域の状態を管理するプロセッサと、を備え、

前記第1ストレージ装置のメモリの使用を停止するとき、前記複数のストレージ装置の少なくとも一部のプロセッサは、前記第1ストレージ装置以外の前記ストレージ装置が其々管理する前記キャッシュ領域の状態に基づいて、前記使用を停止するメモリが格納する管理情報であるコピー対象管理情報のコピー先となるコピー先ストレージ装置を決定し、

前記コピー先ストレージ装置のプロセッサは、前記コピー先ストレージ装置のメモリのキャッシュ領域の少なくとも一部を解放して、前記解放したキャッシュ領域に、前記コピー対象管理情報を格納する

ことを特徴とするストレージシステム。

【請求項 2】

前記管理情報は、前記複数のストレージ装置が稼動しているとき、前記複数のストレージ装置の前記メモリの少なくともいずれかに格納され、前記複数のストレージ装置の少なくともいずれかのプロセッサが当該メモリから参照可能な情報であり、

前記キャッシュ情報は、前記複数のストレージ装置の少なくともいずれかが備える記憶

10

20

デバイスに同じデータが格納されている場合、前記複数のストレージ装置のいずれのメモリからも削除可能な情報であり、

前記キャッシュ領域の状態は、前記キャッシュ領域のアクセス頻度の情報を少なくとも含み、

前記複数のストレージ装置に、前記使用を停止するメモリが格納する管理情報のコピー先の候補となるコピー先候補ストレージ装置が複数ある場合、前記少なくとも一部のプロセッサは、

前記コピー先候補ストレージ装置が其々管理する前記キャッシュ領域のアクセス頻度に基づいて、前記コピー先候補ストレージ装置の其々に前記コピー対象管理情報をコピーした場合のアクセス性能の低下量を予測し、

前記複数のコピー先候補ストレージ装置のうち、前記予測したアクセス性能の低下量の低いストレージ装置を、前記コピー先ストレージ装置に決定する

ことを特徴とする請求項 1 に記載のストレージシステム。

【請求項 3】

前記複数のストレージ装置は、前記第 1 ストレージ装置と第 2 ストレージ装置と前記コピー先ストレージ装置を少なくとも備え、

前記第 1 ストレージ装置のメモリと前記第 2 ストレージ装置のメモリに、前記コピー対象管理情報を其々格納することで、前記コピー対象管理情報を冗長化しており、

前記第 1 ストレージ装置のメモリの使用を停止するとき、前記コピー先ストレージ装置のメモリに前記コピー対象管理情報を格納することで、前記第 2 ストレージ装置のメモリと前記コピー先ストレージ装置のメモリで前記コピー対象管理情報を冗長化し、

前記コピー先ストレージ装置は、前記コピー対象管理情報の格納先を、前記複数のストレージ装置の少なくとも一部に通知する

ことを特徴とする請求項 2 に記載のストレージシステム。

【請求項 4】

前記コピー先候補ストレージ装置は、前記複数のストレージ装置のうち、前記第 1 ストレージ装置と前記第 2 ストレージ装置以外の少なくとも 2 以上のストレージ装置であり、

前記少なくとも 2 以上のコピー先候補ストレージ装置は其々、前記アクセス性能の低下量を予測して、前記第 2 ストレージ装置に前記予測した低下量を送信し、

前記第 2 ストレージ装置は、前記少なくとも 2 以上のコピー先候補ストレージ装置其々で予測した低下量に基づき、前記予測したアクセス性能の低下量の低いコピー先候補ストレージ装置を、前記コピー先ストレージ装置に決定し、

前記コピー先ストレージ装置は、前記第 2 ストレージ装置から前記コピー対象管理情報をコピーする

ことを特徴とする請求項 3 に記載のストレージシステム。

【請求項 5】

前記複数のストレージ装置のそれぞれのメモリに格納する管理情報は、

前記複数のストレージ装置がそれぞれ有する統合管理情報と、

他ストレージ装置から参照可能な共有管理情報と、

各ストレージ装置で個別に管理する管理情報であって、当該ストレージ装置の前記キャッシュ領域のアクセス頻度の情報を含む個別管理情報と、

を有し、

前記コピー対象管理情報は、前記第 1 ストレージ装置及び前記第 2 ストレージ装置其々のメモリに格納された前記共有管理情報である

ことを特徴とする請求項 4 に記載のストレージシステム。

【請求項 6】

前記コピー先ストレージ装置のプロセッサは、前記コピー先ストレージ装置のメモリのキャッシュ領域を複数の小領域に分けて、各小領域のアクセス頻度を管理し、

前記各小領域のアクセス頻度に基づいて、アクセス頻度の低い小領域に格納されたキャッシュ情報の使用を停止することで、前記キャッシュ領域の少なくとも一部を解放する

10

20

30

40

50

ことを特徴とする請求項 5 に記載のストレージシステム。

【請求項 7】

前記コピー先ストレージ装置のプロセッサは、前記コピー対象管理情報を格納するキャッシュ領域が連続した領域となるように、前記アクセス頻度に基づいて、前記コピー先ストレージ装置のキャッシュ情報の再配置を実行する

ことを特徴とする請求項 6 に記載のストレージシステム。

【請求項 8】

前記コピー先ストレージ装置は、ホスト計算機に提供する論理ボリュームを構成する 1 以上の記憶デバイスをさらに有し、

前記コピー先ストレージ装置のプロセッサは、

前記論理ボリュームに設定されるキャッシュ保証量を管理し、

前記コピー先ストレージ装置のメモリのキャッシュ領域が有する小領域であって、当該小領域を解放しても前記キャッシュ保証量を満たす場合に、当該小領域に格納されたキャッシュ情報の使用を停止する

ことを特徴とする請求項 6 に記載のストレージシステム。

【請求項 9】

前記複数のストレージ装置の其々は、メモリを複数有し、

各ストレージ装置は、当該ストレージ装置で個別に管理する前記個別管理情報を、当該ストレージ装置内の異なるメモリに格納することで冗長化する

ことを特徴とする請求項 5 に記載のストレージシステム。

【請求項 10】

前記複数のストレージ装置の其々は、メモリを有するコントローラを複数備え、

各ストレージ装置は、当該ストレージ装置で個別に管理する前記個別管理情報を、当該ストレージ装置内の異なるコントローラのメモリに格納することで冗長化する

ことを特徴とする請求項 9 に記載のストレージシステム。

【請求項 11】

前記第 1 ストレージ装置のメモリの使用を停止するときは、前記第 1 ストレージ装置で障害が発生したとき、または第 1 ストレージ装置を減設するときである

ことを特徴とする請求項 5 に記載のストレージシステム。

【請求項 12】

各ストレージ装置の前記管理情報は、いずれかのプロセッサで実行されるプログラムと、前記いずれかのプログラムの実行に用いられる情報と、の少なくともいずれか一方を含み、

各ストレージ装置の前記キャッシュ情報は、当該ストレージ装置に接続するホスト計算機からのライト要求に基づくライトデータと、当該ホスト計算機からのリード要求への応答に用いるリードデータと、の少なくともいずれか一方を含む

ことを特徴とする請求項 5 に記載のストレージシステム。

【請求項 13】

前記ストレージシステムに 1 以上のストレージ装置を増設するとき、

前記増設するストレージ装置が有するメモリに、前記統合管理情報と、前記コピー対象管理情報を格納し、

前記コピー先ストレージ装置が有するメモリのうち、前記コピー対象管理情報が格納された領域を解放し、前記キャッシュ情報を格納する

ことを特徴とする請求項 5 に記載のストレージシステム。

【請求項 14】

管理情報を格納する管理情報格納領域とキャッシュ情報を格納するキャッシュ領域とを有するメモリと、前記キャッシュ領域の状態を管理するプロセッサと、を其々備える複数のストレージ装置に接続する接続部と、

前記複数のストレージ装置のうちの 1 つのストレージ装置のメモリの使用を停止するとき、前記 1 つのストレージ装置以外のストレージ装置であって、当該ストレージ装置のメ

10

20

30

40

50

モリのキャッシュ領域の少なくとも一部を解放して、前記使用を停止するメモリが格納する管理情報をコピーするストレージ装置を、前記１つのストレージ装置以外のストレージ装置が其々管理する前記キャッシュ領域の状態に基づいて決定する制御部と、
を備えることを特徴とするストレージ装置。

【請求項１５】

複数のストレージ装置と、前記複数のストレージ装置間を接続するネットワークと、を備えるストレージシステムの制御方法であって、

前記複数のストレージ装置は、第１ストレージ装置と第２ストレージ装置を含む４以上のストレージ装置を少なくとも備え、

前記複数のストレージ装置の其々は、管理情報を格納する管理情報格納領域とキャッシュ情報を格納するキャッシュ領域とを有するメモリと、前記キャッシュ領域のアクセス頻度を管理するプロセッサと、を備え、

前記第１ストレージ装置のメモリと前記第２ストレージ装置のメモリに、同一の管理情報を其々格納することで、前記同一の管理情報を冗長化しており、

前記第１ストレージ装置のメモリの使用を停止するとき、前記第１ストレージ装置と前記第２ストレージ装置以外の少なくとも２以上のストレージ装置が、前記同一の管理情報のコピー先の候補となるコピー先候補ストレージ装置であり、

前記少なくとも２以上のコピー先候補ストレージ装置は其々、当該コピー先候補ストレージ装置が管理する前記キャッシュ領域のアクセス頻度に基づいて、当該コピー先候補ストレージ装置に前記同一の管理情報をコピーした場合のアクセス性能の低下量を予測して、前記第２ストレージ装置に前記予測した低下量を送信し、

前記第２ストレージ装置は、前記少なくとも２以上のコピー先候補ストレージ装置其々で予測した低下量に基づき、前記予測したアクセス性能の低下量の低いコピー先候補ストレージ装置を、コピー先ストレージ装置に決定し、

前記コピー先ストレージ装置は、前記コピー先ストレージ装置のメモリのキャッシュ領域の少なくとも一部を解放して、前記第２ストレージ装置から前記同一の管理情報をコピーして、前記解放したキャッシュ領域に格納し、

前記同一の管理情報は、前記第２ストレージ装置のメモリと前記コピー先ストレージ装置のメモリの間で冗長化される

ことを特徴とするストレージシステムの制御方法。

【発明の詳細な説明】

【技術分野】

【０００１】

本発明は、ストレージ装置、ストレージシステム、ストレージシステムの制御方法に関する。

【背景技術】

【０００２】

本技術分野の背景技術として、特開２００６－２２１５２６号公報（特許文献１）に記載の技術がある。この公報には、複数のメモリを有するストレージ装置において、管理情報を異なるメモリ間で冗長化し、管理情報を格納するメモリに障害が発生したとき、他のメモリに管理情報をコピーして冗長性を確保する技術について記載されている。

【先行技術文献】

【特許文献】

【０００３】

【特許文献１】特開２００６－２２１５２６号公報

【発明の概要】

【発明が解決しようとする課題】

【０００４】

企業等がビジネスを遂行するための業務データを保存するストレージシステムについて、ビジネスの変化にあわせてシステム規模を柔軟に変更するため、複数のストレージ装置

10

20

30

40

50

を連携させてシステムを構築したいという要求がある。例えば、システム導入時は当面必要となるストレージリソースを備えるストレージ装置でシステムを構築し、ビジネスが成長してストレージリソースの利用量が増加してきたら、既存のストレージ装置を活用しつつストレージ装置を増設することで、システム全体のストレージリソースを拡張することができる。

【 0 0 0 5 】

このように、ストレージシステムを複数のストレージ装置で構築する場合、ストレージシステムの稼動に必要な管理情報を、異なるストレージ装置のメモリ間で冗長化することで、可用性を高めることができる。また、ストレージシステムを構築する一部のストレージ装置で障害が発生したとき、または減設するとき等、当該ストレージ装置の使用を停止する場合、当該ストレージ装置が格納する管理情報を他のストレージ装置にコピーすることで、冗長性を維持することができる。

10

【 0 0 0 6 】

特許文献 1 では、障害発生時に管理情報をコピーするために、コピー先のメモリに管理情報を格納可能な空き容量を確保する必要がある。しかし、コピー先のストレージ装置において、当該管理情報を格納可能な空き容量を定常的に確保する場合、通常時のメモリの使用率が低下して、コストが高くなる。これに対し、コピー先のストレージ装置に当該管理情報を格納可能な空き容量を動的に確保する場合、通常時のメモリの使用率は向上するが、動的に空き容量を確保した時のシステムへの影響を低減する必要がある。そのため、各ストレージ装置で個別に管理されるメモリの使用状態をふまえて、コピー先のストレージ装置を適切に決めることが重要である。

20

【 0 0 0 7 】

そこで、ストレージ装置の使用を停止する場合に備えて、当該ストレージ装置の管理情報を格納するための空き容量を、他のストレージ装置のメモリに定常的に確保しなくても、当該ストレージ装置の使用を停止するときに動的に空き容量を確保可能にし、また動的に空き容量を確保したときのシステムへの影響を低減可能なストレージシステムを提供する。

【課題を解決するための手段】

【 0 0 0 8 】

上記課題を解決するために、本発明のストレージシステムは、第 1 ストレージ装置を含む複数のストレージ装置を備える。複数のストレージ装置の其々は、管理情報を格納する管理情報格納領域とキャッシュ情報を格納するキャッシュ領域とを有するメモリと、キャッシュ領域の状態を管理するプロセッサと、を備える。第 1 ストレージ装置のメモリの使用を停止するとき、複数のストレージ装置の少なくとも一部のプロセッサは、第 1 ストレージ装置以外のストレージ装置が其々管理するキャッシュ領域の状態に基づいて、使用を停止するメモリが格納する管理情報であるコピー対象管理情報のコピー先となるコピー先ストレージ装置を決定する。コピー先ストレージ装置のプロセッサは、コピー先ストレージ装置のメモリのキャッシュ領域の少なくとも一部を解放して、解放したキャッシュ領域に、コピー対象管理情報を格納する。

30

【発明の効果】

40

【 0 0 0 9 】

ストレージ装置の使用を停止する場合に備えて、当該ストレージ装置の管理情報を格納するための空き容量を、他のストレージ装置のメモリに定常的に確保しなくても、当該ストレージ装置の使用を停止するときに動的に空き容量を確保でき、メモリの使用効率を向上できる。また、動的に空き容量を確保したときのストレージシステムへの影響を低減できる。上記した以外の課題、構成及び効果は、以下の発明を実施するための形態の説明により明らかにされる。

【図面の簡単な説明】

【 0 0 1 0 】

【図 1】管理情報の退避処理の概念を示す図である。

50

【図 2】ストレージ装置の構成を示す図である。

【図 3】メモリの構成を示す図である。

【図 4】管理情報配置テーブルの構成を示す図である。

【図 5】キャッシュ管理テーブルの構成を示す図である。

【図 6】論理ボリューム管理テーブルの構成である。

【図 7】論物変換テーブルの構成を示す図である。

【図 8】ストレージ装置の減設処理を示すフローチャートである。

【図 9】ストレージ装置の障害処理を示すフローチャートである。

【図 10】アクセス性能の低下量予測処理のフローチャートである

【図 11】キャッシュ情報の再配置処理のフローチャートである。

【図 12】共有管理情報へのアクセス処理を示すフローチャートである。

【発明を実施するための形態】

【0011】

以下、図面を用いて実施例について説明する。

【0012】

図 1 は、本実施例における管理情報の退避処理の概念を示す図である。

【0013】

本実施例のストレージシステムは、複数のストレージ装置 2 が、ネットワーク 4 1 を介して接続されている。各ストレージ装置 2 のメモリ 9 は、管理情報を格納する管理情報格納領域 11 と、キャッシュ情報を格納するキャッシュ領域 12 とを有する。また、各ストレージ装置 2 のプロセッサ 8 は、当該ストレージ装置 2 のキャッシュ領域 12 のアクセス頻度を管理している。

【0014】

管理情報は、複数のストレージ装置 2 が稼動しているとき、複数のストレージ装置 2 のメモリ 9 の少なくともいずれかに格納され、複数のストレージ装置 2 の少なくともいずれかのプロセッサ 8 が当該メモリ 9 から参照可能な情報である。特に、各ストレージ装置 2 が稼動しているとき、当該ストレージ装置 2 のプロセッサ 8 は、当該ストレージ装置 2 のメモリ 9 に格納された管理情報を参照できる。

【0015】

キャッシュ情報は、複数のストレージ装置 2 が稼動しているとき、複数のストレージ装置 2 の少なくともいずれかが備える記憶デバイス（ドライブ 4）に同じデータが格納されている場合、複数のストレージ装置 2 のいずれのメモリ 9 から削除可能な情報である。特に、各ストレージ装置 2 が稼動しているとき、当該ストレージ装置 2 が備える記憶デバイス（ドライブ 4）に格納されたデータと同じデータであるキャッシュ情報は、メモリ 9 のキャッシュ領域 12 から削除しても良い。ここで、削除とは、他のキャッシュ情報で上書きすること、当該キャッシュ情報を消失すること、当該キャッシュ情報にアクセス不能にすることのいずれであっても良い。

【0016】

図 1 では、複数のストレージ装置 2 の例として、ストレージ装置 A、B、C、D が示されている。ただし、ストレージシステムを構成するストレージ装置 2 の数はこの例に限定されず、少なくとも 2 以上のストレージ装置 2 であれば良い。

【0017】

ストレージ装置 A 及びストレージ装置 B それぞれのメモリ 9 の管理情報格納領域 11 は、同一の共有管理情報 X を格納しており、この共有管理情報 X を冗長化している。このように、異なるストレージ装置 2 の間で管理情報の冗長性をもたせることで、一方のストレージ装置 2 で障害が発生してメモリ 9 が使用不能になったとしても、他方のストレージ装置 2 のメモリ 9 に格納された管理情報を用いて、運用が継続可能となり、可用性を高めることができる。

【0018】

以下、ストレージ装置 B のメモリ 9 の使用を停止する場合に、ストレージシステムで実

10

20

30

40

50

行される管理情報の退避処理の例を説明する。ストレージ装置 B のメモリ 9 の使用を停止すると、共有管理情報 X の冗長性が失われ、信頼性が低下してしまう。そこで、本処理により、ストレージ装置 A、B 以外のストレージ装置 C、D の少なくともいずれかに共有管理情報 X をコピーすることで冗長化し、信頼性を回復する。ここで、ストレージ装置 2 のメモリ 9 の使用を停止するときとは、典型的にはストレージ装置 2 の使用を停止するときであり、例えば当該ストレージ装置 2 で障害が発生したときや、ストレージ装置 2 を減設するときである。例えば、ストレージシステムを構成する複数のストレージ装置 2 から一部のストレージ装置 2 を減設することで、運用管理コストが削減可能となる。この使用を停止するメモリ 9 に格納された共有管理情報 X を、コピー対象管理情報とも呼ぶ。

【0019】

10

ストレージ装置 B のメモリ 9 の使用を停止するとき、例えば、当該ストレージ装置 B と共有管理情報 X を冗長化しているストレージ装置 A が、ストレージ装置 B のメモリ 9 の使用停止を検知する。これは、ストレージ装置 A が定期的または不定期にストレージ装置 B にアクセスすることで検知しても良いし、ストレージシステムに接続する管理端末からの指示に基づいて検知しても良い。

【0020】

ストレージ装置 A、B 以外のストレージ装置 C、D が、共有管理情報 X のコピー先候補となるコピー先候補ストレージ装置である。以下の例では、コピー先候補ストレージ装置のいずれのメモリ 9 も、管理情報（共有管理情報 Y 等）とキャッシュ情報（キャッシュ情報 C 1、C 2、D 1、D 2 等）が格納されており、共有管理情報 X を格納するための空き容量がないとする。コピー先候補ストレージ装置に、他のストレージ装置 2 の管理情報（共有管理情報 X）を格納するための空き容量を定期的に確保せず、例えばその代わりにキャッシュ領域 1 2 を多く確保することで、メモリ 9 の使用効率を向上できる。

20

【0021】

本実施例では、コピー先候補ストレージ装置のいずれかのメモリ 9 の情報を削除して、共有管理情報 X を格納する。管理情報はストレージシステムの動作に必要な情報であり削除できないため、キャッシュ情報を削除する。ストレージ装置 2 のメモリ 9 のキャッシュ領域 1 2 は、ホスト計算機 1 からのデータアクセス要求時の応答性能を高速化するために、キャッシュ情報（キャッシュデータ）が格納されており、コピーが必要な管理情報を格納するための容量が空いていないことが多い。キャッシュ領域 1 2 には、アクセス頻度の高いキャッシュ領域とアクセス頻度の低いキャッシュ領域とが混在している。このため、無作為にキャッシュ情報を削除してキャッシュ領域を解放すると、キャッシュヒット率が低下してアクセス性能が大きく低下する。本実施例では、アクセス性能をキャッシュヒット率とし、アクセス性能の低下量をキャッシュヒット率の低下量として説明する。ただし、アクセス性能及びアクセス性能の低下量は、この例に限定されない。

30

【0022】

そこで、キャッシュヒット率の低下量が小さくなるように、キャッシュ領域 1 2 のアクセス頻度に基づいて、少なくとも一部のキャッシュ領域 1 2 を解放する。これにより、あらかじめメモリ容量の空きを確保していなくても、ストレージ装置 2 の障害時または減設時等に、当該ストレージ装置 2 の管理情報を他のストレージ装置 2 のメモリ 9 に格納可能になる。

40

【0023】

以下の説明では、コピー先候補ストレージ装置のいずれか一つに、一つのコピー対象管理情報である共有管理情報 X を格納する例を示す。しかし、コピー対象管理情報が複数あって、異なるストレージ装置 2 に分けて格納可能な場合においては、コピー対象管理情報毎に以下の処理を実行しても良い。その場合、容量の大きいコピー対象管理情報から優先的に、コピー先ストレージ装置を決定しても良い。

【0024】

まず、ストレージ装置 A は、コピー先候補ストレージ装置 C、D の其々に、共有管理情報 X の容量を通知する。コピー先候補ストレージ装置 C、D は、各々が個別に管理するキ

50

キャッシュ領域 1 2 のアクセス頻度に基づいて、共有管理情報 X をコピーした場合のアクセス性能の低下量を予測して、ストレージ装置 A に予測した低下量を送信する。ストレージ装置 A は、コピー先候補ストレージ装置 C、D 其々で予測した低下量に基づき、前記予測したアクセス性能の低下量の低いコピー先候補ストレージ装置を、前記コピー先ストレージ装置に決定する (S 1 0 0)。これにより、キャッシュ領域を解放したときにおけるシステム全体のアクセス頻度への影響を小さくできる。図 1 の例では、コピー先ストレージ装置は、ストレージ装置 C である。各コピー先候補ストレージ装置 C、D におけるアクセス性能の低下量予測処理の詳細は、図 1 0 を用いて後述する。

【 0 0 2 5 】

ここで、コピー先ストレージ装置は、アクセス性能の低下量の低いコピー先候補ストレージ装置とする例について説明したが、必ずしもアクセス性能の低下量だけで決定しなくても良い。例えば、複数のコピー先候補ストレージ装置のうち、コピー対象管理情報が優先的に配置されるストレージ装置 2 がある場合は、その優先順位に基づいて、コピー先ストレージ装置を決定しても良い。この場合、各コピー先候補ストレージ装置から、必要な情報を取得しても良い。例えば、プロセッサ 8 が高負荷のストレージ装置 2 がコピー処理を実施するとアクセス性能が低下する恐れがあるため、各ストレージ装置 2 のプロセッサ 8 の負荷を取得し、プロセッサ 8 の負荷が低いストレージ装置 2 をコピー先ストレージ装置に決定しても良い。

【 0 0 2 6 】

コピー先ストレージ装置は、コピー先候補ストレージ装置が其々管理するキャッシュ領域 1 2 の状態に基づいて、決定されて良い。キャッシュ領域 1 2 の状態には、キャッシュ領域 1 2 のアクセス頻度の情報を含んでも良い。キャッシュ領域 1 2 の状態として、キャッシュ領域 1 2 の最終アクセス時刻 1 8 5 や、使用可能か否かの情報、キャッシュ領域 1 2 が使用されているか否かの情報 (使用中フラグ 1 8 3) を含んでも良い。

【 0 0 2 7 】

ストレージ装置 A は、コピー先ストレージ装置 C に共有管理情報 X をコピーするための容量を確保するように指示する。コピー先ストレージ装置 C のプロセッサは、共有管理情報 X を格納するキャッシュ領域が連続した領域となるように、アクセス頻度に基づいて、当該コピー先ストレージ装置 C のキャッシュ情報 (C 1、C 2 等) の再配置を実行する (S 1 0 1)。このキャッシュ情報の再配置処理 (S 1 0 1) の詳細は、図 1 1 を用いて後述する。

【 0 0 2 8 】

コピー先ストレージ装置 C は、当該コピー先ストレージ装置 C のメモリ 9 のキャッシュ領域 1 2 の少なくとも一部を解放することで、共有管理情報 X を格納するための空き容量を確保する (S 1 0 2)。コピー先ストレージ装置 C は、ストレージ装置 A から共有管理情報 X をコピーして、解放したキャッシュ領域 1 2 に格納する (S 1 0 3)。したがって、共有管理情報 X は、ストレージ装置 A とコピー先ストレージ装置 C 其々のメモリ 9 の間で冗長化される。

【 0 0 2 9 】

コピー先ストレージ装置 C は、共有管理情報 X の格納先を、前記複数のストレージ装置 2 の少なくとも一部または全部に通知する。例えば、コピー先ストレージ装置 C は、共有管理情報 X の格納先を、ストレージ装置 A とストレージ装置 D に通知する。このように、共有管理情報 X の格納先を、他のストレージ装置 2 に通知するので、通知されたストレージ装置 2 はコピー先ストレージ装置 C に格納された共有管理情報 X にアクセスすることができる。

【 0 0 3 0 】

ストレージ装置 2 の使用を停止する場合に備えて、当該ストレージ装置 2 のコピー対象管理情報を格納するための空き容量を、他のストレージ装置 2 のメモリ 9 に定常的に確保しなくても、当該ストレージ装置 2 の使用を停止するときに動的に空き容量を確保でき、メモリ 9 の使用効率を向上できる。また、動的に空き容量を確保したときのストレージシ

10

20

30

40

50

ステムへの影響を低減できる。

【0031】

図2は、ストレージ装置2の構成を示す図である。

【0032】

ストレージ装置2は、データの記憶装置であるドライブ4と、コマンドに応じた処理を実行するストレージコントローラ5とを有する。ドライブ4は、不揮発性の記憶デバイスである。ドライブ4は、例えば、磁気ディスクやSSD(Solid State Drive)といった半導体メモリ装置であっても良い。ドライブ4とストレージコントローラ5は、それぞれ1または2以上であっても良い。

【0033】

ストレージコントローラ5は、サーバI/F(インターフェイス)6、ドライブI/F7、装置間結合I/F40、プロセッサ8、メモリ9を有し、これらは内部ネットワークによって相互に接続される。サーバI/F6は、ネットワーク3を介してホスト計算機1と接続し、ホスト計算機1に対してコマンドやデータの送受信処理を実行する。ドライブI/F7は、ドライブ4に対してコマンドやデータの送受信処理を実行する。装置間結合I/F40は、ネットワーク41を介して他のストレージ装置2と接続し、他のストレージ装置2に対してコマンドやデータの送受信処理を実行する。一つのストレージコントローラ5内にメモリ9は、1または2以上であっても良い。ネットワーク3とネットワーク41は、それぞれに接続された装置(ホスト計算機1やストレージ装置2)間でコマンドやデータをやり取りするための通信路であり、例えば、SAN(Storage Area Network)である。

【0034】

ネットワーク3とネットワーク41は、同じネットワークであっても良いし、互いに独立したネットワークであっても良い。なお、ネットワーク3とネットワーク41とが互いに独立したネットワークである場合、複数のストレージ装置2間でネットワーク41を介して行われる通信が、ホスト計算機1とストレージ装置2との間でネットワーク3を介して行われる通信に、影響を及ぼさない利点がある。

【0035】

プロセッサ8は、メモリ9内のプログラムを実行して、コマンドに応じた各種の処理を実行する。プロセッサ8は、プログラムを実行する演算部または制御部であれば良い。以降では、プロセッサ8がメモリ9上のプログラムを実行して行われる処理を、ストレージ装置2やストレージコントローラ5を処理の主体として記載することがある。

【0036】

メモリ9は、揮発性の記憶装置である。メモリ9は、データを記憶する記憶部であれば良い。複数のストレージ装置2は結合し、互いにメモリ9を共有する。複数のストレージ装置2の間で共有されるメモリ9は、複数のストレージ装置2の間で一意的識別子である論理メモリIDと、当該メモリ9が存在するストレージ装置2内で一意的識別子である物理メモリIDが付与される。

【0037】

論理ボリューム10は、ストレージコントローラ5によって、複数のドライブ4の物理的な記憶領域から構成される。この構成方法には、例えば、RAID(Redundant Arrays of Inexpensive Disks)があり、複数のドライブ4を用いることで、データ冗長化による信頼性向上やドライブ4の並列動作による性能向上が期待できる。ストレージコントローラ5は、論理ボリューム10をホスト計算機1に提供する。ストレージ装置2は、1または2以上の論理ボリューム10を有する。

【0038】

ホスト計算機1は、プロセッサとメモリとを有する物理計算機、または物理計算機上で稼働する仮想計算機であっても良い。ホスト計算機1は、サーバとも呼ばれる。ホスト計算機1は、例えばオンライン通販サービスを提供するワークステーションである。ホスト計算機1は、ホスト計算機1が提供するサービスに必要なデータの少なくとも一部を、スト

10

20

30

40

50

レージ装置 2 に格納する。ホスト計算機 1 は、ストレージ装置 2 に対して、データアクセス要求を送信することで、ストレージ装置 2 に格納されたデータを読み書きする。データアクセス要求として、例えば、SCSI (Small Computer System Interface) 規格のリードコマンド (Read command) やライトコマンド (Write command) が用いられる。データアクセス要求は、I/O (Input/output) 要求とも呼ばれる。

【0039】

図 3 は、メモリ 9 の構成を示す図である。プロセッサ 8 は、後述する管理情報配置テーブル 13 及びキャッシュ管理テーブル 18 を用いて、メモリ 9 の記憶領域を、管理情報格納領域 11 とキャッシュ領域 12 に分けて管理する。

10

【0040】

管理情報格納領域 11 には、管理情報が格納される。管理情報には、統合管理情報と、共有管理情報と、個別管理情報がある。管理情報は、その管理情報の内容によっては、構成情報や制御情報とも呼ばれる。

【0041】

統合管理情報は、ストレージシステムを構成する複数のストレージ装置 2 が其々有する情報である。統合管理情報の例として、管理情報配置テーブル 13 と、論物変換テーブル 35 と、がある。

【0042】

共有管理情報は、自ストレージ装置 2 だけでなく、他ストレージ装置 2 から参照可能な管理情報である。以下では、共有管理情報は、ストレージシステムを構成する複数のストレージ装置 2 のうち、一部の異なるストレージ装置 2 の間において冗長化される管理情報として、主に説明する。ただし、共有管理情報によっては、同じストレージ装置 2 内の異なるメモリ 9 間で、冗長化されていても良い。

20

【0043】

共有管理情報の位置 (アドレス) は、統合管理情報である管理情報配置テーブル 13 によって管理される。このため、あるストレージ装置 2 内のメモリ 9 に格納された共有管理情報は、他のストレージ装置 2 と共有できる。共有管理情報は、複数のストレージ装置 2 のいずれかのプロセッサで実行されるプログラムと、いずれかのプログラムの実行に用いられる管理情報と、の少なくともいずれか一方を含む。プロセッサ 8 が実行すべきプログラムの例として、データのバックアップ機能を実現するバックアップ機能プログラムがある。プログラムの実行に用いられる管理情報の例として、バックアップ機能プログラムの実行に用いられるバックアップ機能制御情報、論理ボリューム管理テーブル 26、またセキュリティ設定情報、課金情報等がある。

30

【0044】

個別管理情報は、各ストレージ装置 2 で個別に管理する管理情報である。個別管理情報の例として、当該ストレージ装置 2 のキャッシュ領域 12 のアクセス頻度の情報を含むキャッシュ管理テーブル 18 がある。個別管理情報の内容は、異なるストレージ装置 2 の間では異なっていて良い。例えば、各ストレージ装置 2 のキャッシュ管理テーブル 18 は、自身のストレージ装置 2 のキャッシュ領域 12 のアクセス頻度を管理する管理情報である。異なるストレージ装置 2 間のキャッシュ管理テーブル 18 はその内容が異なる。また、キャッシュ管理テーブル 18 は、ホスト計算機 1 のデータアクセス要求によって、その内容が適宜変更される情報である。異なるストレージ装置 2 間でキャッシュ管理テーブル 18 を冗長化すると、その内容が変更される度に同期が必要となるので、ストレージシステムの負荷が向上し、またアクセス応答性能が劣化する可能性がある。キャッシュ管理テーブル 18 は、通常、当該ストレージ装置 2 内で使用される情報であるので、異なるストレージ装置 2 の間で冗長化しなくても良い。

40

【0045】

このように管理情報の用途や内容によって、アクセスするストレージ装置 2 が異なる。複数のストレージ装置 2 からアクセスされる管理情報はストレージ装置 2 間での冗長化が

50

必要であり、コピー対象管理情報をこれらの管理情報に限定することで、コピーする管理情報の量を削減しても良い。例えば、コピー対象管理情報を共有管理情報に限定しても良い。

【0046】

また、キャッシュ領域12には、ホスト計算機1からのデータアクセス要求に基づく入出力データが一時的に格納される。キャッシュ領域12に格納されるデータを、キャッシュ情報と呼ぶ。キャッシュ情報として、ライトデータやリードデータがある。キャッシュ情報は、メモリ9よりも低性能（低速）なドライブ4内のデータへのアクセスを、高性能化（高速化）するための情報である。すなわち、各ストレージ装置2のキャッシュ情報は、ストレージ装置2に接続するホスト計算機1からのライト要求に基づくライトデータと、ホスト計算機1からのリード要求への応答に用いるリードデータと、の少なくともいずれか一方を含む。

10

【0047】

例えば、ストレージ装置2がホスト計算機1からライト要求を受信すると、ストレージ装置2はライト要求にかかるライトデータをメモリ9に一旦格納し、ホスト計算機1にライト完了の応答を返す。ストレージ装置2は、メモリ9から低性能なドライブ4へのライトデータのライト処理（以後、デステージ処理とも言う。）を非同期に実行できるので、ホスト計算機1のライト要求に対して、短時間でライト完了の応答ができる。

【0048】

また、例えば、ストレージ装置2がホスト計算機1からリード要求を受信したときリード要求にかかるデータがメモリ9に格納されている場合、ストレージ装置2はメモリ9から当該データを取得して、ホスト計算機1に応答する。リード要求にかかるデータがメモリ9に格納されていない場合、ストレージ装置2は、ドライブ4から当該データを取得してメモリ9に格納し、ホスト計算機1に応答する。リード要求にかかるデータをリードデータと呼ぶ。ストレージ装置2は、メモリ9のキャッシュ領域12にリードデータが格納されている場合、メモリ9より低性能なドライブ4からリードデータを読み出さなくても良く、ホスト計算機1のリード要求に対して短時間で応答できる。

20

【0049】

各ストレージ装置2は、メモリ9に障害が発生した場合等に個別管理情報やキャッシュ情報の消失を防ぐため、同じストレージ装置2の異なるメモリ9間で個別管理情報及びキャッシュ情報をコピーして多重化し、冗長性を持たせる。ただし、リードデータについては、メモリ9のキャッシュに格納されたリードデータが消失しても、ドライブ4からリードデータを読み出してホスト計算機1へ応答できるため、複数のメモリ9間で多重化しなくても良い。ライトデータについては、後述の通りに冗長化を継続または回復しても良いし、デステージ処理を実行することで、冗長化不要としても良い。また、ストレージ装置2が、メモリ9を有するストレージコントローラ5を複数備える場合、個別管理情報やキャッシュ情報を、当該ストレージ装置2内の異なるストレージコントローラ5のメモリ9に格納することで冗長化しても良い。

30

【0050】

また、各ストレージ装置2は、メモリ9に障害が発生した場合等、メモリ9の使用を停止するとき、使用を停止するメモリ9が格納する管理情報を、同じストレージ装置2内の別のメモリ9にコピーすることで、冗長性を回復しても良い。この場合の管理情報（コピー対象管理情報）には、個別管理情報、共有管理情報、統合管理情報の少なくともいずれかが含まれていて良い。

40

【0051】

このとき、当該ストレージ装置2は、自ストレージ装置2が有する複数のメモリ9を対象として、アクセス性能の低下量予測処理（図10）によって、削除するキャッシュ情報を決定する。これにより、キャッシュ領域12を解放したときにおける当該ストレージ装置2のアクセス頻度への影響を小さくできる。そして、当該ストレージ装置2は、コピー対象管理情報を格納するために、解放するキャッシュ領域12が連続した領域となるよう

50

に、アクセス頻度に基づいて、キャッシュ情報の再配置処理（S101、図11）を実行する。このとき、解放するキャッシュ領域12を選択する際、所定のポリシーや優先度に基づいて選択されても良い。例えば、当該ストレージ装置2内の異なるストレージコントローラ5間のメモリ9で管理情報を冗長化するポリシーや、当該ストレージ装置2内の異なるメモリ9の間で管理情報を冗長化するポリシー等によって、解放するキャッシュ領域12を有するメモリ9が適切に選択されて良い。当該ストレージ装置2は、当該ストレージ装置2のメモリ9のキャッシュ領域12の少なくとも一部を解放することで、コピー対象管理情報を格納するための空き容量を確保する。当該ストレージ装置2は、コピー対象管理情報を有するメモリ9から、コピー対象管理情報をコピーして、解放したキャッシュ領域12に格納することで、冗長性を回復する。

10

【0052】

図4は、管理情報配置テーブル13の構成を示す図である。ストレージ装置2は、管理情報配置テーブル13を用いて、共有管理情報が格納されるメモリ9の管理情報格納領域11を管理する。管理情報配置テーブル13は、管理情報名130と、論理メモリID（A）131と、論理アドレス（A）132と、論理メモリID（B）133と、論理アドレス（B）134と、サイズ135と、を対応づける管理情報である。格納情報名130で識別される管理情報を格納するメモリ9の記憶領域として、論理メモリID（A）131で識別されるメモリ9の論理アドレス（A）132からサイズ135分の領域と、論理メモリID（B）133で識別されるメモリ9の論理アドレス（A）134からサイズ135分の領域とが確保されていることを示す。

20

【0053】

例えば、図4に示す一例によると、管理情報名130が「セキュリティ設定情報」のエントリでは、「セキュリティ設定情報」と呼ばれる管理情報の格納先の記憶領域として、論理メモリID 1の論理アドレス0×1000と、論理メモリID 2の論理アドレス0×1000のそれぞれから、0×1000のサイズの記憶領域が確保されている。

【0054】

各管理情報について、格納先のメモリ9のうち一方が使用不能となっても、管理情報が消失しないように、物理的なメモリ9が異なるように、論理メモリID（A）131と論理メモリID（B）133には異なるIDが設定される。

【0055】

なお、本実施例では、前述のように、論理ボリューム管理テーブル26が共有管理情報であるので、この管理情報配置テーブル13で管理されるが、図4には図示していない。各ストレージ装置2で、それぞれ自ストレージ装置2の論理ボリューム10を対象として、論理ボリューム管理テーブル26が作成される。したがって、各ストレージ装置2の論理ボリューム管理テーブル26は、それぞれ管理情報配置テーブル13で管理されて良い。

30

【0056】

図5は、キャッシュ管理テーブル18の構成を示す図である。ストレージ装置2のプロセッサ8は、キャッシュ管理テーブル14を用いて、キャッシュ情報を格納するメモリ9のキャッシュ領域12の状態を管理する。キャッシュ領域12は、ホスト計算機1のデータアクセスに応じて、格納されるキャッシュ情報が動的に変化する。そのため、物理メモリID181と物理メモリアドレス182が示すキャッシュ領域12の小領域を、所定のサイズごとに管理する。図5の一例では、所定サイズは「0×1000」として表されている。

40

【0057】

キャッシュ管理テーブル18は、エントリを識別するキャッシュID180と、物理メモリID181と、物理アドレス182と、使用中フラグ183と、使用開始時刻184と、最終アクセス時刻185と、ダーティデータフラグ186と、キャッシュヒット回数187を対応づける管理情報である。キャッシュID180は、各エントリの識別子である。物理メモリID181及び物理アドレス182は、物理メモリID181で識別され

50

るメモリ9において、物理アドレス182から所定サイズのキャッシュ領域12が、そのエントリの対象となる記憶領域であることを示している。

【0058】

使用中フラグ183は、当該エントリの対象となるキャッシュ領域12に格納されたキャッシュ情報の有無を示す。使用開始時刻184は、当該キャッシュ領域12に、現在のキャッシュ情報が格納された時刻を示す。最終アクセス時刻185は、ホスト計算機1が当該エントリのキャッシュ領域12を最後にアクセスした時刻を示す。使用開始時刻184と最終アクセス時刻185は、時分秒の形式で例示しているものの、年月日があっても良い。

【0059】

ダーティデータフラグ186は、当該エントリのキャッシュ情報がダーティデータであるか否かを示す。ダーティデータフラグ186がONの場合、すなわち当該エントリのキャッシュ情報がダーティデータである場合、当該キャッシュ情報が最新のライトデータであり、ドライブ4には更新前の古いデータがあることを示す。ダーティデータフラグがOFFの場合、すなわち当該エントリのキャッシュ情報がダーティデータでない場合、当該キャッシュ情報が最新データであり、またドライブ4にも最新データがあることを示す。

【0060】

キャッシュヒット回数187は、所定の単位時間内において、ホスト計算機1からのリード要求に対して、当該リード要求にかかるデータがキャッシュ領域12に存在した回数である。このキャッシュヒット回数187は、各ストレージ装置2が有するカウンタにより計測しても良い。所定の単位時間におけるキャッシュヒット回数187は、カウンタの回数と、カウンタの計測時間から算出しても良い。

【0061】

本実施例では、アクセス頻度を、所定の単位時間におけるキャッシュヒット回数187として説明する。また、本実施例で、アクセス頻度が低いとは、所定の単位時間におけるキャッシュヒット回数187が小さいこととして説明する。比較対象は、他ストレージ装置2または自ストレージ装置2の他メモリ9である。ただし、アクセス頻度は、この例に限定されない。例えば、アクセス頻度を、最終アクセス時刻185としても良い。この場合、アクセス頻度が低いとは、最終アクセス時刻185が古いことである。

【0062】

図6は、論理ボリューム管理テーブル26の構成を示す図である。論理ボリューム管理テーブル26は、各論理ボリューム10を識別する論理ボリュームID260と、キャッシュ保証量261と、論理ボリュームアドレス262と、キャッシュID(A)263と、キャッシュID(B)264と、ドライブID265と、ドライブアドレス266と、を対応づける管理情報である。

【0063】

ストレージ装置2は、論理ボリュームID260と論理ボリュームアドレス262が示す論理ボリューム10の記憶領域を、所定のサイズごとに管理する。図6の一例では、所定サイズは「0×1000」として表されている。論理ボリュームID260は、論理ボリューム10の識別子である。キャッシュ保証量261は、当該論理ボリューム10に設定されたキャッシュ保証量を示す。論理ボリュームアドレス262は、当該論理ボリューム10におけるアドレスを示す。

【0064】

キャッシュID(A)263と、キャッシュID(B)264とは、当該論理ボリュームアドレス262に対応するキャッシュ領域12である。キャッシュID(A)263と、キャッシュID(B)264とがいずれも設定されている場合、これらのキャッシュ領域12にライトデータがあることを示している。キャッシュID(B)263は、ライトデータの二重化先のキャッシュ領域12を示しており、リードデータの場合には、設定されない。また、キャッシュID(A)263と、キャッシュID(B)264とがいずれも設定されている場合、キャッシュ領域12にデータがないことを示す。

10

20

30

40

50

【 0 0 6 5 】

ドライブ I D 2 6 5 は、ドライブ 4 を識別する識別子である。ドライブアドレス 2 6 6 は、当該ドライブ 4 における物理アドレスであり、論理ボリューム I D 2 6 0 及びアドレス 2 6 2 で指定されるデータが格納されるドライブ 4 上の物理記憶領域を示す。

【 0 0 6 6 】

以下、ホスト計算機 1 がストレージ装置 2 へのデータアクセスしたとき、ストレージ装置 2 が実行するデータアクセス処理について述べる。ストレージ装置 2 は、ホスト計算機 1 から、論理ボリューム 1 0 の I D とアドレスを指定するデータアクセス要求を受信する。ストレージ装置 2 は、指定された論理ボリューム 1 0 のアドレスに対応づけられたキャッシュのアドレス、ドライブ 4 のアドレスを特定する。具体的には、論理ボリューム管理
10
テーブル 2 6 を参照して、データアクセス要求で指定された論理ボリューム I D 2 6 0 と論理ボリュームアドレス 2 6 2 に対応づけられたキャッシュ I D (A) 2 6 3 とキャッシュ I D (B) 2 6 4、ドライブ I D 2 6 5 とドライブアドレス 2 6 6 を取得する。取得したキャッシュ I D (A) 2 6 3 とキャッシュ I D (B) 2 6 4 はそれぞれ、キャッシュ管理テーブル 1 8 におけるキャッシュ I D 1 8 0 に対応する。

【 0 0 6 7 】

ここで、ホスト計算機 1 からのリード要求に対して、当該リード要求にかかるデータがキャッシュ領域 1 2 に存在する確率（以降、キャッシュヒット率と呼ぶ）が高いほど、リード要求に対する応答は高性能となる。ただし、キャッシュ領域 1 2 は、複数の論理ボリューム 1 0 間の共有資源であり、全てのキャッシュ I D 1 8 0 のエントリについての使用中フラグ 1 8 3 が O N（使用中）の場合には、任意のエントリのキャッシュ情報を削除して、新しいキャッシュ情報を格納する。この結果、ある論理ボリューム 1 0 へのデータ
20
アクセスによって、他の論理ボリューム 1 0 のキャッシュヒット率が低下する場合がある。ユーザはサービス毎に異なる論理ボリューム 1 0 を使用する場合があり、サービス間での性能の干渉を抑えたい要求がある。

【 0 0 6 8 】

そこで、ストレージ装置 2 は、論理ボリューム 1 0 毎に使用可能なキャッシュ領域 1 2 の最低容量を保証する機能を提供する。例えば、図 6 のように、ユーザによって設定されるキャッシュ保証量 2 6 1 を、論理ボリューム 1 0 毎に記憶し、削除するキャッシュ情報を選択する際、削除後もキャッシュ保証量 2 6 1 を満たせるか否かを判定する。満たせる
30
場合には当該キャッシュ情報を削除し、満たせない場合には他のキャッシュ情報を削除対象とする。

【 0 0 6 9 】

図 7 は、論物変換テーブル 3 5 の構成を示す図である。論物変換テーブル 3 5 は、各種プログラムが認識するデータの格納先である論理メモリ I D 3 5 0 及び論理アドレス 3 5 1 と、これらが示すデータが実際に格納されている格納先として、ストレージ装置 2 を識別するストレージ I D 3 5 2 と、当該ストレージ装置 2 における物理メモリ I D 3 5 3 と物理アドレス 3 5 4 とを対応づける管理情報である。各ストレージ装置 2 のプロセッサ 8
40
は、論理メモリ I D 3 5 0 と論理アドレス 3 5 1 と、ストレージ I D 3 5 2 と物理メモリ I D 3 5 3 と物理アドレス 3 5 4 への変換、または逆変換を、論物変換テーブル 3 5 を用いて、適宜実行する。

【 0 0 7 0 】

前述の S 1 0 2 及び S 1 0 3 で、コピー先ストレージ装置 2 が、当該コピー先ストレージ装置 2 のメモリ 9 のキャッシュ領域 1 2 の少なくとも一部を解放し、コピー対象管理情報を格納する処理について述べた。本処理について、以下具体的に説明する。コピー先ストレージ装置 2 は、論理ボリューム管理テーブル 2 6 において、削除対象のキャッシュ情報が格納されたキャッシュ領域 1 2 のキャッシュ I D (A) 2 6 3 とキャッシュ I D (B) 2 6 4 を「なし」に設定し、当該キャッシュ領域 1 2 の使用を停止する。さらに、コピー先ストレージ装置 2 は、キャッシュ管理テーブル 1 8 を参照し、削除対象のキャッシュ
50
I D 1 8 0 について、物理メモリ I D 1 8 1 と物理アドレス 1 8 2 が示す領域を管理情報

のコピー先として記憶して、当該エントリを削除する。

【 0 0 7 1 】

そして、コピー先ストレージ装置 2 は、S 1 1 のステップで記憶した管理情報のコピー先に、管理情報をコピーする。具体的には、ストレージ装置 2 は、管理情報配置テーブル 1 3 を参照し、論理メモリ ID (A) 1 3 1 又は論理メモリ ID (B) 1 3 3 の値が使用不能となるメモリ ID となっているエントリを検索する。ストレージ装置 2 は、当該エントリのアクセス可能なメモリアドレスを使用して管理情報をリードし、S 1 1 のステップで記録した管理情報のコピー先にライトする。その後、ストレージ装置 2 は、コピー先のメモリ ID とメモリアドレスを当該エントリの使用不能となる論理メモリ ID と論理メモリアドレスに上書きする。

10

【 0 0 7 2 】

以降、ストレージ装置 2 の減設処理 (図 8) 及び障害処理 (図 9) について述べる。これらの処理の主体となるストレージ装置 2 は、ストレージシステムのいずれのストレージ装置 2 でも良い。ただし、障害が発生したストレージ装置 2 以外のストレージ装置 2 で行われるのが望ましい。例えば、減設するストレージ装置 2 または障害が発生したストレージ装置 2 が有する管理情報のコピーを有しているストレージ装置 2 でこれらの処理を実行しても良い。また、各ストレージ装置 2 と接続し、ストレージシステムを管理する管理計算機がある場合、管理計算機が各ストレージ装置 2 と連携して、これらの処理を実行しても良い。

【 0 0 7 3 】

20

また、減設処理または障害処理を開始する契機は、これらの処理を実行するストレージ装置 2 が、減設するストレージ装置 2 または障害が発生したストレージ装置 2 から、減設通知または障害通知を受領した時や、ユーザの管理端末から減設処理または障害処理の実行を指示された時であっても良い。また、各ストレージ装置 2 間で通信時にあるストレージ装置 2 が、他のストレージ装置 2 の障害を検出したときに、障害処理を実行しても良い。

【 0 0 7 4 】

図 8 は、ストレージ装置 2 の減設処理を示すフローチャートである。複数のストレージ装置 2 のうち、1 以上のストレージ装置 2 を減設する場合の処理について述べる。

【 0 0 7 5 】

30

ストレージ装置 2 は、各ストレージ装置 2 における管理情報の容量 (以後、管理情報量という。) とキャッシュ保証量の合計、また減設後のメモリ量の合計 (以後、減少後のメモリ量という。) を算出する。ストレージ装置 2 は、減少後のメモリ量が、管理情報量とキャッシュ保証量の合計 (以後、必要なメモリ量という。) を上回るか否かを判定する (S 1)。ここで、必要なメモリ量と減少後のメモリ量について、ストレージ装置 2 は、各ストレージ装置 2 の管理情報配置テーブル 1 3 とキャッシュ管理テーブル 1 8 と論理ボリューム管理テーブル 2 6 とに基づいて算出しても良い。または、ストレージ装置 2 は、各ストレージ装置 2 に対して、メモリ量と管理情報量とキャッシュ保証量を返却するように要求し、各ストレージ装置 2 から返却されたメモリ量、管理情報量とキャッシュ保証量に基づいて、必要なメモリ量と減少後のメモリ量を算出しても良い。

40

【 0 0 7 6 】

減少後のメモリ量が必要なメモリ量を上回る場合、減設対象のストレージ装置 2 のメモリ 9 に格納された管理情報を、他のストレージ装置 2 のメモリ 9 に退避する退避処理を実行する (S 2)。この管理情報の退避処理 (S 2) は、図 1 を用いて前述したとおりである。その後、ユーザの管理端末に対して、ストレージ装置 2 の減設が可能であることを示す情報を表示しても良い (S 3)。ユーザは、その表示内容を確認した上で、ストレージ装置 2 を減設可能となる。

【 0 0 7 7 】

S 1 の判定で、減少後のメモリ量が必要なメモリ量を上回らない場合、減設すると必要な情報を残りのストレージ装置 2 のメモリ 9 に格納できないために減設できない旨を、ユ

50

ーザの管理端末に表示しても良い(S4)。このとき、ストレージ装置2は、例えば、ユーザに対して減設後にどの程度メモリ量が不足するかを示す情報を表示しても良い。ユーザは、その表示内容を参照することで、キャッシュ保証量28の見直しなどを実行して必要なメモリ量を減少させたうえで、ストレージ装置2に再び減設処理を実行させることができる。

【0078】

図9は、ストレージ装置2の障害処理を示すフローチャートである。複数のストレージ装置2のうち、1以上のストレージ装置2で障害が発生した場合の処理について述べる。

【0079】

ストレージ装置2は、他ストレージ装置2における障害を検出すると、図8のS1のステップと同じ処理を実行する。S1の判定で、必要なメモリ量を上回る場合、S2の管理情報の退避処理を実行する。

【0080】

S1の判定で、必要なメモリ量を上回らない場合、障害時のポリシーを参照する(S5)。障害時のポリシーとは、管理情報の二重化とキャッシュ保証量の両立ができない場合に信頼性と性能のどちらを優先するかを定めた情報であり、ユーザによって設定される。

【0081】

性能優先のポリシーの場合(S5 No)、管理情報を一重状態のままとし、信頼性を回復させない(S8)。このステップS8では、例えば、ストレージ装置2は、ユーザの管理端末に対して、メモリ量の不足によって信頼性を回復できない旨を表示しても良い。また、ストレージ装置2は、キャッシュ保証量の見直しや、早期に障害部品の交換作業を促す情報を、ユーザの管理端末に表示しても良い。

【0082】

S5において信頼性優先のポリシーであった場合(S5 Yes)、キャッシュ保証量のうち、実際に使用しているキャッシュ量のみを必要なメモリ量と見なして、必要なメモリ量を再計算する(S6)。具体的には、論理ボリューム管理テーブル26の論理ボリュームID260ごとに、キャッシュID(A)263とキャッシュID(B)264の有無を集計し、使用中のキャッシュ量を算出する。算出した使用中のキャッシュ量とキャッシュ保証量261を比較し、小さい値を当該論理ボリューム10で保証すべきメモリ量とする。管理情報配置テーブル13のサイズ135の2倍と、論理ボリューム10ごとの保証すべきメモリ量との合計を必要なメモリ量とし、減少後のメモリ量と比較する。

【0083】

減少後のメモリ量が再計算した必要なメモリ量を上回る場合(S6 Yes)、S2の管理情報の退避処理を実行する。減少後のメモリ量が再計算した必要なメモリ量を上回らない場合(S6 No)、S8のステップを実行する。

【0084】

これにより、サービスの性能よりもサービスの信頼性を優先するユーザの要求を満たせる。また、ホスト計算機1のデータアクセスパターンが変わらない限り、使用するキャッシュ量は増加しないため、性能を低下させることなく、信頼性を維持できる。

【0085】

S2のステップ実行後、管理情報の信頼性回復を完了する(S7)。S7のステップでは、例えば、ストレージ装置2は、障害の発生と信頼性が回復されたことを示す情報を、ユーザの管理端末に表示しても良い。また、ストレージ装置2は、障害の発生していないストレージ装置2をそのまま動作させても良いことを、ユーザの管理端末に表示しても良い。また、ストレージ装置2は、ユーザの管理端末に、障害部品の交換作業を促す情報を表示しても良い。

【0086】

図10は、アクセス性能の低下量予測処理のフローチャートである。以下、アクセス性能の低下量が、キャッシュヒット率の低下量であるときの例を述べる。本処理は、コピー先候補ストレージ装置の其々で実行する。そのため、コピー先候補ストレージ装置の其々

10

20

30

40

50

のメモリ9のキャッシュ領域12の使用状態に基づいて、キャッシュヒット率の低下を抑えるように、コピー先ストレージ装置及び解放するキャッシュ領域12を決定することができる。

【0087】

まず、コピー先候補ストレージ装置2は、コピー対象管理情報の容量を取得する(S14)。図1の例で説明したように、ストレージ装置Aが、コピー先候補ストレージ装置C、Dの其々に、コピー対象管理情報(共有管理情報X)の容量を通知することにより実行されても良い。または、コピー先候補ストレージ装置が、管理情報配置テーブル13を参照して、論理メモリID(A)131又は論理メモリID(B)133の値が使用不能となるメモリIDとなっているエントリを検索する。当該エントリのサイズ135の合計値を、コピー対象管理情報の容量としても良い。コピー対象管理情報の容量を、この時点での「削除するキャッシュ情報量」とする。

10

【0088】

次に、コピー先候補ストレージ装置2は、個別管理情報であるキャッシュ管理テーブル18を用いて、自身が有する全メモリ9のキャッシュ領域12を使用頻度の低い順序に並び替える(S15)。これは、キャッシュヒット率を低下させにくい順にキャッシュ情報を並び替えることを目的としており、例えば、キャッシュ管理テーブル18の最終アクセス時刻185の順に並び替える。他にも、最終アクセス時刻185が同じ場合には、使用開始時刻184の古い方を優先して並び替える方法、キャッシュヒット回数187の少ない方を優先して並び替える方法、などがあり、過去のキャッシュの使用状況を使用する限り、その方法はS15のステップに記載に限定されない。

20

【0089】

コピー先候補ストレージ装置2は、S15のステップで並び替えた使用頻度の低い順にキャッシュ領域12を特定し(S16)、当該キャッシュ領域12を確保する必要があるか否かを判定する(S17)。具体的には、コピー先候補ストレージ装置2は、キャッシュ管理テーブル18を参照して、取り出したキャッシュ領域12に合致するエントリを特定し、当該エントリID180を持つ論理ボリューム管理テーブル26のエントリを特定する。コピー先候補ストレージ装置2は、特定したエントリにおいて、キャッシュID(A)263とキャッシュID(B)264が示すキャッシュ領域12の割り当て量とキャッシュ保証量261を比較する。コピー先候補ストレージ装置2は、当該キャッシュ領域12を解放しても、論理ボリューム10のキャッシュ領域12の割り当て量がキャッシュ保証量261を下回らない場合(S17 No)、保証不要と判定してS18のステップを実行する。一方、論理ボリューム10のキャッシュ領域12の割り当て量がキャッシュ保証量261を下回る場合(S17 Yes)、コピー先候補ストレージ装置2は、保証要と判断してS16のステップを実行する。

30

【0090】

コピー先候補ストレージ装置2は、S17のステップで確保不要と判定した後、当該キャッシュ領域12を削除対象に設定(S18)し、「削除するキャッシュ情報量」から当該キャッシュ領域12のサイズを減算して、更新する(S19)。更新した「削除するキャッシュ情報量」が0より大きい場合、さらにキャッシュ情報を削除する必要があると判定してS16のステップを実行する(S20 Yes)。

40

【0091】

更新した「削除するキャッシュ情報量」が0以下の場合、(S20 No)、コピー先候補ストレージ装置2は、キャッシュ管理テーブル18を参照して、削除対象キャッシュ領域12における所定の単位時間あたりのキャッシュヒット回数187の合計を、当該ストレージ装置が所定の単位時間あたりに受信したリード回数、すなわちリード要求の数で割ることで、アクセス性能の低下量の予測値を算出する(S200)。ただし、アクセス性能の低下量の予測値は、他の方法で算出しても良く、S200の記載に限定されない。

【0092】

図11は、キャッシュ情報の再配置処理のフローチャートである。本処理は、コピー先

50

ストレージ装置で実行する。本処理により、コピー対象管理情報を格納するために必要な連続したメモリ領域を確保し、またキャッシュ領域１２の解放後のアクセス性能の低下を抑えることができる。

【００９３】

以下の説明では、簡単化のために削除対象のキャッシュ領域１２を一か所に再配置する処理を説明するものの、一か所に限定されるものではない。また、コピー対象管理情報を複数のメモリ領域に分けて管理できるならば、複数のメモリ領域に再配置しても良いし、コピー対象管理情報をキャッシュ領域１２の連続したメモリ領域に配置できるならば再配置そのものを実行しなくても良い。ただし、複数のメモリ領域に細かく分けてコピー対象管理情報を管理するほど、細分化された管理情報の格納先メモリを記憶するためのメモリ消費量が増大する。例えば、管理情報配置テーブル１３のエントリが増加する。このため、ある程度の大きな連続領域で管理情報を格納することで、メモリ９の使用効率を向上できる。

【００９４】

まず、コピー先ストレージ装置２は、使用可能なメモリ９を選択（Ｓ２１）し、削除対象のキャッシュ領域１２を列挙する（Ｓ２２）。

【００９５】

次に、コピー先ストレージ装置２は、削除対象のキャッシュ領域１２の密度が高い領域を再配置先のキャッシュ領域１２に設定する（Ｓ２３）。これは、再配置するキャッシュ情報の数を少なくすることを目的としている。

【００９６】

コピー先ストレージ装置２は、列挙したキャッシュ領域１２の先頭を取り出し（Ｓ２４）、再配置先のメモリ領域外にあるか否かを判定する（Ｓ２５）。コピー先ストレージ装置２は、再配置先のメモリ領域外にある場合（Ｓ２５　Ｙｅｓ）、削除対象でなくかつ領域内にあるキャッシュ情報と取り出したキャッシュ情報とのデータを交換する（Ｓ２６）。具体的には、コピー先ストレージ装置２は、キャッシュ管理テーブル１８のエントリ間で物理アドレス１８２が示す領域のキャッシュ情報を相互に移動した上で、エントリ間で物理アドレス１８２、使用開始時刻１８４、最終アクセス時刻１８５、ダーティデータフラグ１８６、アクセスヒット回数１８７も相互に移動する。なお、この処理はキャッシュ情報の配置変更であり、キャッシュ領域１２へのデータアクセスではないため、使用開始時刻１８４や最終アクセス時刻１８５は変更しない。

【００９７】

コピー先ストレージ装置２は、Ｓ２５のステップにて再配置先のメモリ領域内と判定された場合（Ｓ２５　Ｎｏ）、又はＳ２６のステップを実行した場合、キャッシュ管理テーブル１８を参照して、Ｓ２４のステップで取り出したキャッシュ領域１２に対応するエントリのダーティデータフラグ１８６を参照する（Ｓ２７）。

【００９８】

コピー先ストレージ装置２は、ダーティデータフラグ１８６がＯＮの場合（Ｓ２７　Ｙｅｓ）、ドライブ４のデータは旧データであり、当該キャッシュ領域１２のキャッシュ情報が最新データであるため、キャッシュ情報を削除する前に、キャッシュ情報をドライブ４に格納する（Ｓ２８）。具体的には、コピー先ストレージ装置２は、論理ボリューム管理テーブル２６を参照して、キャッシュＩＤ（Ａ）２６３と（Ｂ）２６４の値が当該キャッシュのＩＤ１８０と合致するエントリを特定する。特定したエントリのドライブＩＤ２６５とドライブアドレス２６６が示す領域に、当該キャッシュ情報を格納する。その後、ダーティデータフラグ１８６をＯＦＦに設定する。これにより、当該キャッシュ領域１２のキャッシュ情報を削除できる。

【００９９】

コピー先ストレージ装置２は、ダーティデータフラグ１８６がＯＦＦの場合（Ｓ２７　Ｎｏ）、ドライブ４のデータに最新データであるため、当該キャッシュ領域１２のキャッシュ情報を削除できる。

10

20

30

40

50

【 0 1 0 0 】

そして、列挙したキャッシュ領域 1 2 の残りがあるか否かを確認する (S 2 9)。残があれば (S 2 9 Y e s)、S 2 4 のステップを実行する。残りがなければ (S 2 9 N o)、処理を終了する。

【 0 1 0 1 】

なお、本実施例では、キャッシュ情報の再配置処理において、再配置先のメモリ領域内外のキャッシュ情報の交換 (S 2 6)、ドライブへの最新データとなるキャッシュ情報の格納 (S 2 8)、キャッシュ情報の削除 (S 1 0 2) の順で処理を説明した。しかし、S 2 8 のステップ、S 1 0 2 のステップを先に実行すると、S 2 6 のステップでは再配置先のメモリ領域内の削除対象でないキャッシュ情報を領域外の空き領域に移動すれば良く、
10
キャッシュ情報の交換に比べて、データコピー回数を減らすことができる。削除するキャッシュ領域 1 2 をキャッシュヒット率が低下しにくいように選択し、削除した領域にコピー対象管理情報の二重化するという限り、処理の順序は記載のみに限定されるものではない。

【 0 1 0 2 】

図 1 0 で説明したアクセス性能の低下量予測処理と、図 1 1 のキャッシュ情報の再配置処理を実行することにより、コピー先ストレージ装置 2 は、以下の処理を実行する。コピー先ストレージ装置 2 のプロセッサ 8 は、キャッシュ管理テーブル 1 8 を用いて、当該コピー先ストレージ装置のメモリ 9 のキャッシュ領域 1 2 を複数の小領域に分けて、各小領域のアクセス頻度を管理する。ここで、小領域とは、キャッシュ管理テーブルのキャッシュ
20
I D 1 8 0 の各エントリで管理されるキャッシュ領域 1 2 である。コピー先ストレージ装置 2 は、各小領域のアクセス頻度に基づいて、アクセス頻度の低い小領域に格納されたキャッシュ情報の使用を停止することで、キャッシュ領域 1 2 の少なくとも一部を解放する。ここで、コピー先ストレージ装置 2 のプロセッサ 8 は、論理ボリューム管理テーブル 2 6 を用いて、論理ボリューム 1 0 に設定されるキャッシュ保証量 2 6 1 を管理し、キャッシュ領域 1 2 のアクセス頻度の低い小領域を解放してもキャッシュ保証量を満たす場合に、当該小領域に格納されたキャッシュ情報の使用を停止する。これにより、コピー先ストレージ装置において、キャッシュ領域 1 2 を解放した後のアクセス性能の低下を抑えることができる。

【 0 1 0 3 】

また、コピー先ストレージ装置を決定する処理 (S 1 0 0) において、過去にメモリ障害が発生したストレージ装置 2 をコピー先ストレージ装置としないことで、今後の障害における管理情報の退避処理を省略できる可能性がある。例えば、サービスの成長に合わせて、ストレージ装置 2 を徐々に増設した場合、新旧のメモリ 9 が混在するストレージ装置となる。ここで、初期購入した比較的古いストレージ装置 2 のあるメモリ 9 に障害が発生した場合に、当該ストレージ装置 2 の別のメモリ 9 のキャッシュ領域 1 2 を解放して管理情報をコピーしても、当該メモリ 9 も故障したメモリ 9 と同様に比較的古いため、障害が発生する可能性が高い。このため、比較的新しい他のストレージ装置 2 のメモリ 9 のキャッシュ領域 1 2 を優先的に解放して管理情報をコピーしても良い。
30

【 0 1 0 4 】

S 1 0 3 の管理情報のコピー処理では、前述のように、コピー先ストレージ装置は、解放したキャッシュ領域 1 2 にコピー対象管理情報を格納する。このとき、管理情報配置テーブル 1 3 の論理メモリ I D 1 3 1 と論理アドレス 1 3 2 を変更しても良い。しかし、ストレージ装置 2 で実行される各種プログラムが共有管理情報の格納先である論理メモリ I D 1 3 1 と論理アドレス 1 3 2 を静的に記憶している場合、又はプログラム初期起動時に管理情報配置テーブル 1 3 から過去に読み込んだ値を再利用している場合、S 1 0 3 のステップでコピー対象管理情報の格納先が変わると、当該コピー対象管理情報へのアクセスができなくなる。
40

【 0 1 0 5 】

そこで、本実施例では、S 1 0 3 のステップでコピー対象管理情報の格納先が変わった
50

場合に、管理情報配置テーブル 13 の論理メモリ ID 131 と論理アドレス 132 を変更せず、論物変換テーブル 35 を更新する。具体的には、コピー先ストレージ装置 2 は、構成情報配置テーブル 13 及び論物変換テーブル 35 を用いて、障害が発生したストレージ装置 2 または減設するストレージ装置 2 において、コピー対象管理情報が格納されたメモリ 9 の論理メモリ ID 131 と論理アドレス 132 を特定する。コピー先ストレージ装置 2 は、論物変換テーブル 35 において、特定した論理メモリ ID 131 と論理アドレス 132 に対するエントリに対し、自身のストレージ ID 352、コピー対象管理情報の格納先である自身のメモリ 9 の物理メモリ ID 353 及び物理アドレス 354 に更新する。

【0106】

また、論物変換テーブル 35 は、全てのストレージ装置 2 がそれぞれ有する統合管理情報であるため、各ストレージ装置 2 の論物変換テーブル 35 を更新する。例えば、コピー先ストレージ装置 2 は、コピー対象管理情報の格納先（ストレージ ID 352、物理メモリ ID 353、物理アドレス 354）を、複数のストレージ装置 2 の少なくとも一部または全部に通知することで、各ストレージ装置 2 の論物変換テーブル 35 を更新する。例えば、通知先のストレージ装置 2 は、使用を停止するストレージ装置 2 と当該通知先のストレージ装置 2 自身以外のストレージ装置 2 に、コピー対象管理情報の格納先を通知する。通知されたストレージ装置 2 は、其々論物変換テーブル 35 を更新する。

【0107】

図 12 は、論物変換テーブル 35 を用いた共有管理情報へのアクセス処理を示すフローチャートである。本処理は、各ストレージ装置 2 で実行される。例えば、ストレージ装置 2 のプロセッサ 8 で実行されるプログラムが論理メモリ ID と論理アドレスを指定して共有管理情報にアクセスしようとする、以下に示す共有管理情報へのアクセス処理を実施する。

【0108】

まず、ストレージ装置 2 は、論物変換テーブル 35 を参照して、指定された論理メモリ ID と論理アドレスに合致するエントリを、管理情報配置テーブル 13 から検索する（S30）。ストレージ装置 2 は、合致するエントリがあれば（S30 Yes）、指定された論理メモリ ID 131 と論理アドレス 132 は管理情報格納領域と判定して S31 のステップを実施する。ストレージ装置 2 は、合致するエントリがなければ（S30 No）、管理情報格納領域でないと判定してデータアクセス失敗とする（S34）。

【0109】

S31 のステップでは、ストレージ装置 2 は、論物変換テーブル 35 を参照して、指定された論理メモリ ID と論理アドレスに合致する論理メモリ ID 350 と論理アドレス 351 を持つエントリを検索する。

【0110】

次に、ストレージ装置 2 は、S31 のステップで検索したエントリにおけるストレージ ID 352 と、物理メモリ ID 353 と物理アドレス 354 が示すメモリ領域のデータにアクセスし（S32）、データアクセス成功とする（S33）。

【0111】

以上のように、論物変換テーブル 35 を設けることで、ストレージ装置 2 のプロセッサ 8 で実行されるプログラムは、S103 により共有管理情報の格納場所が変更されても、同じ論理メモリ ID 131 と論理アドレス 132 を使って管理情報にアクセスできる。

【0112】

これまで、ストレージ装置 2 の減設や障害によって、使用可能なメモリ 9 が減少する場合に、管理情報の二重化を継続する方法を述べた。ここで、ストレージ装置 2 の増設や、障害回復によって使用可能なメモリ 9 が増加する場合には、ストレージ装置 2 は増加したメモリ 9 にキャッシュ情報を格納することで性能向上を図る。ただし、メモリ障害時処理の S8 のステップを実施して管理情報が二重化されていない状態であれば、増加したメモリ 9 を管理情報の退避先を選択し、S103 のステップにより管理情報を二重化して信頼性を回復しても良い。

10

20

30

40

50

【 0 1 1 3 】

また、ストレージシステムの管理情報配置テーブル 1 3 及び論物変換テーブル 3 5 の通常時（ストレージ装置障害発生前または減設前等）の状態を記憶しておき、ストレージ装置障害発生後または減設後に、ストレージシステムに 1 以上のストレージ装置 2 を増設するとき、通常時の状態に復旧しても良い。例えば、ストレージシステムに 1 以上のストレージ装置 2 を増設するとき、増設するストレージ装置 2 が有するメモリ 9 に、前記統合管理情報と、コピー対象管理情報を格納し、コピー先ストレージ装置 2 が有するメモリ 9 のうち、コピー対象管理情報が格納された領域を解放し、前記キャッシュ情報を格納しても良い。

【 0 1 1 4 】

10

以上の説明では、ストレージシステムが有する各情報について、テーブルなどの表現で説明する場合があるが、各情報のデータ構造は限定されず、他のデータ構造であってもよい。各情報はデータ構造に依存しないため、例えば、「k k k テーブル」を「k k k 情報」と呼ぶことができる。

【 0 1 1 5 】

また、以上の説明において、プロセッサ 8 は、プログラムを実行し、記憶資源（例えば、メモリ 9）及び／又は通信インターフェイスデバイス（例えば、通信ポート）等を用いながら処理を行う。プロセッサ 8 を主体とする処理は、1 以上のプログラムを実行する事によりおこなわれると解釈されてもよい。また、ストレージ装置 2 を主体とする処理は、当該ストレージ装置 2 のプロセッサ 8 を主体として実行されると解釈されてもよい。プロセッサ 8 は、典型的には、CPU（Central Processing Unit）のようなマイクロプロセッサであるが、処理の一部（例えば、暗号化／復号化、圧縮／伸張）を実行するハードウェア回路を含んでもよい。

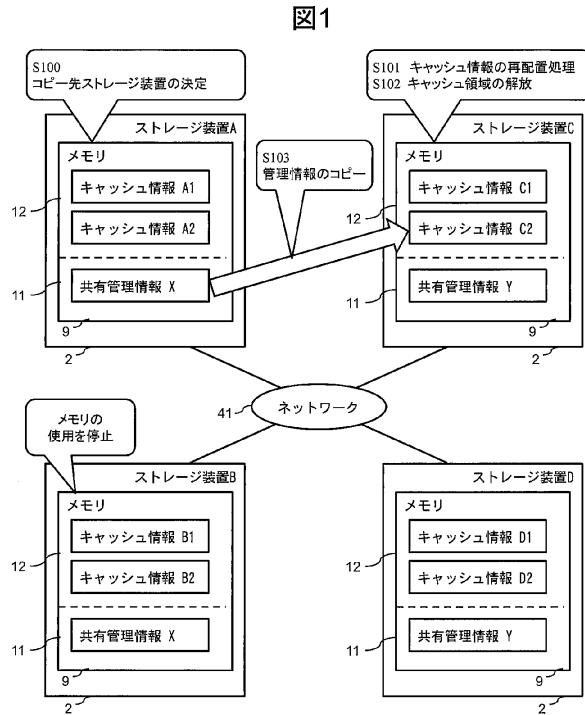
20

【 符号の説明 】

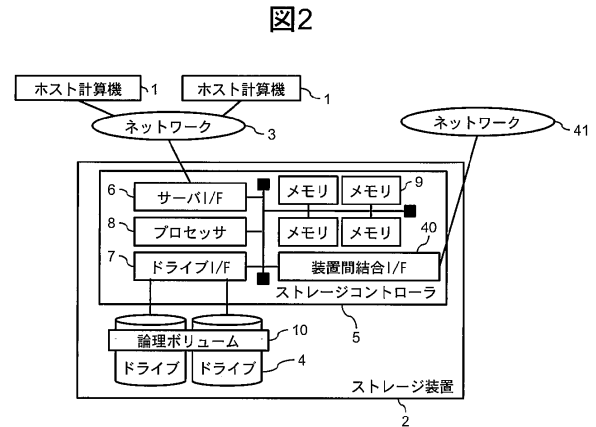
【 0 1 1 6 】

1 ホスト計算機、2 ストレージ装置、3 ネットワーク、4 ドライブ、5 ストレージコントローラ、6 サーバ I / F、7 ドライブ I / F、8 プロセッサ、9 メモリ、10 論理ボリューム、11 管理情報格納領域、12 キャッシュ領域、40 装置間結合 I / F、41 ネットワーク

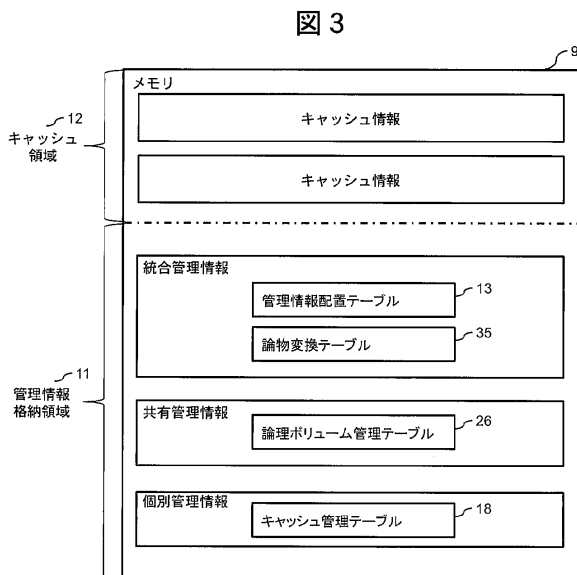
【図 1】



【図 2】



【図 3】



【図 4】

図4は、管理情報配置テーブル13を示す。

管理情報名	論理メモリID (A)	論理アドレス (A)	論理メモリID (B)	論理アドレス (B)	サイズ
セキュリティ設定情報	1	0x1000	2	0x1000	0x1000
バックアップ制御情報	1	0x2000	2	0x3000	0x2000
課金情報	2	0x8000	4	0x8000	0x100

【図 5】

図5は、キャッシュ管理テーブル18を示す。

キャッシュID	物理メモリID	物理アドレス	使用中フラグ	使用開始時刻	最終アクセス時刻	ダーティデータフラグ	キャッシュヒット回数
1	1	0x0000	ON	10:00:00	12:00:00	ON	40
2	1	0x4000	OFF	00:00:00	00:00:00	OFF	0
3	2	0x2000	ON	09:59:59	10:00:00	OFF	70

【図 6】

図 6
論理ボリューム管理テーブル 26

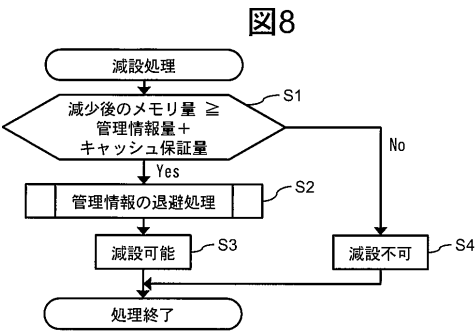
260 論理 ボリューム ID	261 キャッシュ 保証量	262 論理 ボリューム アドレス	263 キャッ シュ ID (A)	264 キャッ シュ ID (B)	265 ドライブ ID	266 ドライブ アドレス
1	500MB	0x0000	1	7	1	0x2000
		0x1000	3	なし	5	0x2000
2	100MB	0x0000	なし	なし	3	0x3000

【図 7】

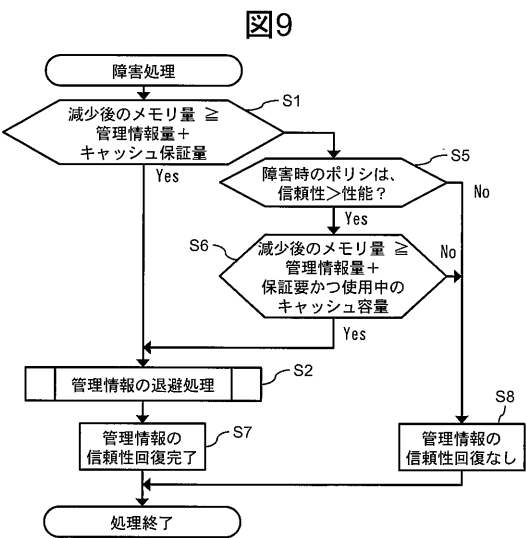
図7
論物変換テーブル 35

350 論理 メモリ ID	351 論理 アドレス	352 ストレージ ID	353 物理 メモリ ID	354 物理 アドレス
1	0x1000	1	3	0x5000
1	0x2000	2	4	0x5000
2	0x8000	3	2	0x8000

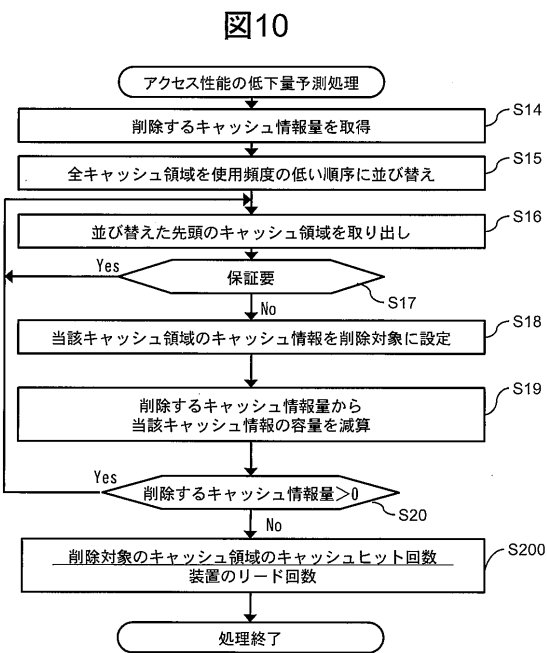
【図 8】



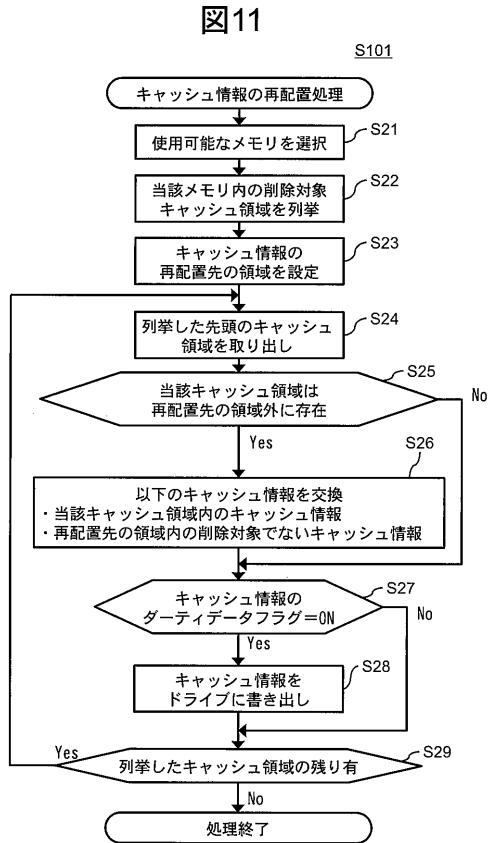
【図 9】



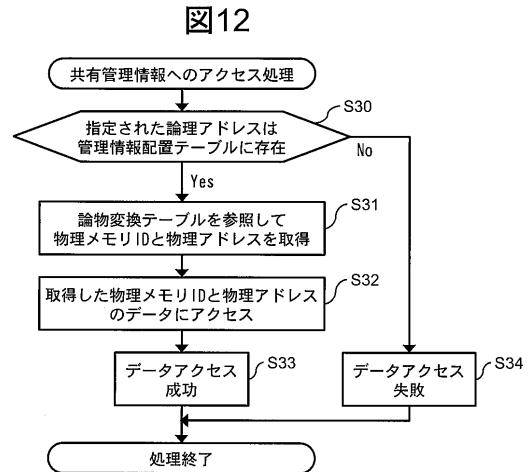
【図 10】



【図 11】



【図 12】



フロントページの続き

(72)発明者 下園 紀夫

東京都千代田区丸の内一丁目6番6号 株式会社日立製作所内

審査官 小林 哲雄

(56)参考文献 特開2006-221526(JP,A)

特表2015-517697(JP,A)

特開2005-285058(JP,A)

(58)調査した分野(Int.Cl., DB名)

G06F 12/00