

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第4433372号  
(P4433372)

(45) 発行日 平成22年3月17日(2010.3.17)

(24) 登録日 平成22年1月8日(2010.1.8)

(51) Int.Cl.		F I			
<b>G06F 3/06</b>	<b>(2006.01)</b>	G06F 3/06	301E		
<b>G06F 12/08</b>	<b>(2006.01)</b>	G06F 3/06	302A		
		G06F 3/06	540		
		G06F 12/08	523E		
		G06F 12/08	541C		

請求項の数 4 (全 31 頁) 最終頁に続く

(21) 出願番号	特願2003-173347 (P2003-173347)	(73) 特許権者	000005108
(22) 出願日	平成15年6月18日(2003.6.18)		株式会社日立製作所
(65) 公開番号	特開2005-10997 (P2005-10997A)		東京都千代田区丸の内一丁目6番6号
(43) 公開日	平成17年1月13日(2005.1.13)	(74) 代理人	110000279
審査請求日	平成18年6月15日(2006.6.15)		特許業務法人ウィルフォート国際特許事務所
		(74) 代理人	100095371
			弁理士 上村 輝之
		(74) 代理人	100089277
			弁理士 宮川 長夫
		(74) 代理人	100104891
			弁理士 中村 猛

最終頁に続く

(54) 【発明の名称】 データアクセスシステム及び方法

(57) 【特許請求の範囲】

【請求項1】

ホストコンピュータに接続された制御部と、  
複数の論理ボリュームと  
を備え、

前記複数の論理ボリュームには、前記ホストコンピュータに対してオンライン状態の論理ボリュームであるオンラインボリュームと、前記ホストコンピュータに対してオフライン状態の論理ボリュームであるオフラインボリュームとが含まれ、

前記オンラインボリュームに格納されたデータと、前記オフラインボリュームに格納されたデータは、それぞれ、別々に存在し、

前記制御部は、前記複数の論理ボリュームにそれぞれ用意された複数のアクセス制御テーブルを有し、

前記アクセス制御テーブルは、そのテーブルに対応する論理ボリュームに対するアクセスを制御するためのテーブルであり、

前記制御部は、前記ホストコンピュータから、論理ボリュームにアクセスするためのコマンドであるアクセスコマンドを受信し、

前記アクセスコマンドは、第1のサブコマンドと第2のサブコマンドとを含み、

前記第1のサブコマンドは、前記第2のサブコマンドよりも先に前記ホストコンピュータから送信されるパケットであり、

前記第1のサブコマンドが、コマンド対象ボリュームIDと、第1のアクセス種別コー

ドと、ターゲット情報とを含み、

前記コマンド対象ボリュームIDは、前記アクセスコマンドの発行先とする論理ボリュームであるコマンド対象ボリュームのIDであり、

前記第1のアクセス種別コードは、前記ホストコンピュータがどのようなアクセスを行いたいを示すコードであり、

前記ターゲット情報は、前記ホストコンピュータがアクセスしたい論理ボリュームであるターゲットボリュームのIDと、アクセス先のアドレスとを含んだ情報であり、

前記第2のサブコマンドが、前記コマンド対象ボリュームIDと、第2のアクセス種別コードとを含み、

前記第2のアクセス種別コードは、前記ホストコンピュータからのアクセスがリードとライトのどちらであるかを示すコードであり、

前記第1のサブコマンド内の前記コマンド対象ボリュームIDも前記第2のサブコマンド内の前記コマンド対象ボリュームIDも、前記オンラインボリュームのIDであり、

前記第1のサブコマンド内の前記第1のアクセス種別コードがリードを示す場合、前記ターゲットボリュームIDは前記コマンド対象ボリュームIDと同一であり、且つ、その第1のサブコマンドの後に送信される前記第2のサブコマンドは、前記第2のアクセス種別コードとしてリードを示し、

前記第1のサブコマンド内の前記第1のアクセス種別コードがライトを示す場合、前記ターゲットボリュームIDは前記コマンド対象ボリュームIDと同一であり、且つ、その第1のサブコマンドの後に送信される前記第2のサブコマンドは、前記第2のアクセス種別コードとしてライトを示し、

前記第1のサブコマンド内の前記第1のアクセス種別コードがリードコントロールを示す場合、前記ターゲットボリュームIDは、前記コマンド対象ボリュームIDと違って前記オフラインボリュームのIDであり、且つ、その第1のサブコマンドの後に送信される前記第2のサブコマンドは、前記第2のアクセス種別コードとしてリードを示し、

前記第1のサブコマンド内の前記第1のアクセス種別コードがライトコントロールを示す場合、前記ターゲットボリュームIDは、前記コマンド対象ボリュームIDと違って前記オフラインボリュームのIDであり、且つ、その第1のサブコマンドの後に送信される前記第2のサブコマンドは、前記第2のアクセス種別コードとしてライトを示し、

前記制御部が、

(A) 前記ホストコンピュータから前記第1のサブコマンドを受信し、前記受信した第1のサブコマンド内の前記第1のアクセス種別コードを解析し、

(B) 前記(A)で解析された第1のアクセス種別コードがリード又はライトを示す場合

:

(b1) その第1のサブコマンド内の前記コマンド対象ボリュームIDに対応した前記アクセス制御テーブルに、前記第1のアクセス種別コードが示すリード又はライトをアクセス種別として記録し、且つ、前記オンラインボリュームのアクセス先のアドレスを記録し、

(b2) 前記(b1)の後に、前記第2のサブコマンドを受信し、

(b3) 前記(b2)で受信した第2のサブコマンド内の前記コマンド対象ボリュームIDに対応したアクセス制御テーブル内のアクセス種別がリード又はライトであることを特定し、

(b4) そのコマンド対象ボリュームIDから識別される前記オンラインボリュームの、そのアクセス制御テーブルに記録されているアクセス先アドレスに対して、リード又はライトを行い、

(C) 前記(A)で解析された第1のアクセス種別コードがリードコントロール又はライトコントロールを示す場合:

(c1) その第1のサブコマンド内の前記コマンド対象ボリュームIDに対応した前記アクセス制御テーブルに、前記第1のアクセス種別コードが示すリードコントロール又はライトコントロールをアクセス種別として記録し、且つ、前記(A)で受信した第1の

10

20

30

40

50

サブコマンド内の前記ターゲット情報が有するターゲットボリュームIDである前記オフラインボリュームIDを記録し、

(c2) 前記(A)で受信した第1のサブコマンド内の前記ターゲットボリュームIDに対応したアクセス制御テーブルに、リード又はライトをアクセス種別として記録し、且つ、前記(A)で受信した第1のサブコマンド内の前記ターゲット情報が有するアクセス先アドレスを記録し、

(c3) 前記(c1)及び(c2)の後に、前記第2のサブコマンドを受信し、(c4) 前記(c3)で受信した第2のサブコマンド内の前記コマンド対象ボリュームIDに対応したアクセス制御テーブル内のアクセス種別がリードコントロール又はライトコントロールであることを特定し、

(c5) そのコマンド対象ボリュームIDに対応したアクセス制御テーブルに記録されているターゲットボリュームIDを特定し、そのターゲットボリュームIDから識別される前記オフラインボリュームの、そのターゲットボリュームIDに対応したアクセス制御テーブルに記録されているアクセス先アドレスに対して、リード又はライトを行う、記憶システム。

#### 【請求項2】

請求項1に記載の記憶システムにおいて、

前記制御部が、プライマリボリュームと、前記プライマリボリュームとペアを構成するセカンダリボリュームとの両方にデータをライトする二重化処理を行うようになっており、

前記オンラインボリュームが、前記プライマリボリュームであり、  
前記オフラインボリュームが、前記セカンダリボリュームである、  
記憶システム。

#### 【請求項3】

請求項1に記載の記憶システムにおいて、

第1のホストコンピュータにとっての前記ターゲットボリュームである前記オフラインボリュームは、第2のホストコンピュータに対してオンライン状態になっているオンラインボリュームである、  
記憶システム。

#### 【請求項4】

複数の論理ボリュームを有する記憶システムに対してホストコンピュータがデータアクセスするための方法において、

前記複数の論理ボリュームには、前記ホストコンピュータに対してオンライン状態の論理ボリュームであるオンラインボリュームと、前記ホストコンピュータに対してオフライン状態の論理ボリュームであるオフラインボリュームとが含まれ、

前記オンラインボリュームに格納されたデータと、前記オフラインボリュームに格納されたデータは、それぞれ、別々に存在し、

前記ホストコンピュータが、論理ボリュームにアクセスするためのコマンドであるアクセスコマンドを送信し、

前記アクセスコマンドは、第1のサブコマンドと第2のサブコマンドとを含み、  
前記第1のサブコマンドは、前記第2のサブコマンドよりも先に前記ホストコンピュータから送信されるパケットであり、

前記第1のサブコマンドが、コマンド対象ボリュームIDと、第1のアクセス種別コードと、ターゲット情報とを含み、

前記コマンド対象ボリュームIDは、前記アクセスコマンドの発行先とする論理ボリュームであるコマンド対象ボリュームのIDであり、

前記第1のアクセス種別コードは、前記ホストコンピュータがどのようなアクセスを行いたいかを示すコードであり、

前記ターゲット情報は、前記ホストコンピュータがアクセスしたい論理ボリュームであるターゲットボリュームのIDと、アクセス先のアドレスとを含んだ情報であり、

10

20

30

40

50

前記第2のサブコマンドが、前記コマンド対象ボリュームIDと、第2のアクセス種別コードとを含み、

前記第2のアクセス種別コードは、前記ホストコンピュータからのアクセスがリードとライトのどちらであるかを示すコードであり、

前記第1のサブコマンド内の前記コマンド対象ボリュームIDも前記第2のサブコマンド内の前記コマンド対象ボリュームIDも、前記オンラインボリュームのIDであり、

前記第1のサブコマンド内の前記第1のアクセス種別コードがリードを示す場合、前記ターゲットボリュームIDは前記コマンド対象ボリュームIDと同一であり、且つ、その第1のサブコマンドの後に送信される前記第2のサブコマンドは、前記第2のアクセス種別コードとしてリードを示し、

10

前記第1のサブコマンド内の前記第1のアクセス種別コードがライトを示す場合、前記ターゲットボリュームIDは前記コマンド対象ボリュームIDと同一であり、且つ、その第1のサブコマンドの後に送信される前記第2のサブコマンドは、前記第2のアクセス種別コードとしてライトを示し、

前記第1のサブコマンド内の前記第1のアクセス種別コードがリードコントロールを示す場合、前記ターゲットボリュームIDは、前記コマンド対象ボリュームIDと違って前記オフラインボリュームのIDであり、且つ、その第1のサブコマンドの後に送信される前記第2のサブコマンドは、前記第2のアクセス種別コードとしてリードを示し、

前記第1のサブコマンド内の前記第1のアクセス種別コードがライトコントロールを示す場合、前記ターゲットボリュームIDは、前記コマンド対象ボリュームIDと違って前記オフラインボリュームのIDであり、且つ、その第1のサブコマンドの後に送信される前記第2のサブコマンドは、前記第2のアクセス種別コードとしてライトを示し、

20

前記記憶システムは、前記複数の論理ボリュームにそれぞれ用意された複数のアクセス制御テーブルを有し、

前記アクセス制御テーブルは、そのテーブルに対応する論理ボリュームに対するアクセスを制御するためのテーブルであり、

前記記憶システムが、

(A) 前記ホストコンピュータから前記第1のサブコマンドを受信し、前記受信した第1のサブコマンド内の前記第1のアクセス種別コードを解析し、

(B) 前記(A)で解析された第1のアクセス種別コードがリード又はライトを示す場合

30

：  
(b1) その第1のサブコマンド内の前記コマンド対象ボリュームIDに対応した前記アクセス制御テーブルに、前記第1のアクセス種別コードが示すリード又はライトをアクセス種別として記録し、且つ、前記オンラインボリュームのアクセス先のアドレスを記録し、

(b2) 前記(b1)の後に、前記第2のサブコマンドを受信し、

(b3) 前記(b2)で受信した第2のサブコマンド内の前記コマンド対象ボリュームIDに対応したアクセス制御テーブル内のアクセス種別がリード又はライトであることを特定し、

(b4) そのコマンド対象ボリュームIDから識別される前記オンラインボリュームの、そのアクセス制御テーブルに記録されているアクセス先アドレスに対して、リード又はライトを行い、

40

(C) 前記(A)で解析された第1のアクセス種別コードがリードコントロール又はライトコントロールを示す場合：

(c1) その第1のサブコマンド内の前記コマンド対象ボリュームIDに対応した前記アクセス制御テーブルに、前記第1のアクセス種別コードが示すリードコントロール又はライトコントロールをアクセス種別として記録し、且つ、前記(A)で受信した第1のサブコマンド内の前記ターゲット情報が有するターゲットボリュームIDである前記オフラインボリュームIDを記録し、

(c2) 前記(A)で受信した第1のサブコマンド内の前記ターゲットボリュームI

50

Dに対応したアクセス制御テーブルに、リード又はライトをアクセス種別として記録し、且つ、前記(A)で受信した第1のサブコマンド内の前記ターゲット情報が有するアクセス先アドレスを記録し、

(c3)前記(c1)及び(c2)の後に、前記第2のサブコマンドを受信し、(c4)前記(c3)で受信した第2のサブコマンド内の前記コマンド対象ボリュームIDに対応したアクセス制御テーブル内のアクセス種別がリードコントロール又はライトコントロールであることを特定し、

(c5)そのコマンド対象ボリュームIDに対応したアクセス制御テーブルに記録されているターゲットボリュームIDを特定し、そのターゲットボリュームIDから識別される前記オフラインボリュームの、そのターゲットボリュームIDに対応したアクセス制御テーブルに記録されているアクセス先アドレスに対して、リード又はライトを行う、データアクセス方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、データ記憶領域に対するデータアクセスのための技術に関し、具体的には、例えば、アレイ状に配設された多数のディスク型記憶装置上に用意された論理ボリュームからデータをリードする又はその論理ボリュームにデータをライトするための技術に関する。

【0002】

【従来の技術】

例えば、大容量のデータを取り扱う基幹業務用のデータベースシステムでは、ホストコンピュータ(以下、単に「ホスト」と言う)とは別体に構成された記憶システムを用いてデータが管理されている。この記憶システムは、例えば、ディスクアレイ装置とも呼ばれ、多数のディスク型記憶装置をアレイ状に配設して構成されているRAID(Redundant Array of Independent Inexpensive Disks)である。

【0003】

この記憶システムには、例えば次のようなものがある。

【0004】

すなわち、その記憶システムは、ホストからリード/ライトコマンドを受けることができ、また、仮想的な記憶領域として仮想ディスクを持つことができる。記憶システムは、ホストからライトコマンドを受けたときは、その仮想ディスクが制御ボリュームとして定義されていれば、ミラーリングを行ない、その仮想ディスクが制御ボリュームとして定義されていなければ、通常のリード又はライトのための動作を行なう(例えば特許文献1)。

【0005】

【特許文献1】

特開2000-112666号公報。

【0006】

【発明が解決しようとする課題】

ところで、例えば、記憶システムには、仮想的な記憶領域である論理ボリュームを複数個用意することができるものがあり、ホストのオペレーティングシステム(以下、「ホストOS」と言う)には、記憶システム上の各論理ボリューム毎にその論理ボリュームがオンライン状態かオフライン状態かを管理するものがある。そのホストOSは、同ホストのアプリケーションプログラム(以下、「ホストAP」と言う)から、所定の又はユーザ所望の論理ボリューム(以下、「ターゲットボリューム」と言う)に対するデータの書込み又は読出し処理(以下、「I/O処理」と言う)の要求を受けたときは、そのターゲットボリュームがオンライン状態になっているか否かをチェックし、オンライン状態になっていれば、記憶システムに対してそのターゲットボリュームに対するI/O処理要求を送信し、オフライン状態であれば、ホストAPにエラーを返答する。すなわち、ホストOSは、ターゲットボリュームに対してオンライン状態になっていなければ、そのターゲットボリ

10

20

30

40

50

ュームに対してデータの読出し又は書込みを行なうことができない。

【0007】

しかし、ユーザは、オフライン状態になっているターゲットボリュームに対してデータの読出し又は書込みを行ないたい場合がある。

【0008】

具体的には、例えば、記憶システムでは、記憶装置内のディスクの故障に起因してデータが損失等してもシステムが停止することがないようにするために二重化処理（いわゆるミラーリング処理）が行われることがある。この処理では、記憶システムは、正ボリュームと副ボリュームという2つの論理ボリュームを有し、ホストからデータの書込み要求を受けたときは、書込み要求されたデータ（以下、「ライトデータ」と言う）を複製して、正

10

ボリュームと副ボリュームの両方に同一のライトデータを書込む。

【0009】

このような二重化処理において、正ボリュームと副ボリュームは、一般に、ペア状態で管理されていて、同じボリュームラベル（例えばボリューム名）を持っている。そのため、ホストOSは、ペア状態で管理されている正ボリュームと副ボリュームを別々のボリュームとして区別することはできない。故に、二重化処理では、一方の論理ボリューム、例えば、正ボリュームが強制的にオンライン状態となり、副ボリュームがオフライン状態となり、ホストOSは、正ボリュームに対してのみオンライン状態になる。

【0010】

しかし、ユーザは、ホストAPを操作し副ボリュームに対してI/O処理を行いたい場合がある。そのためには、ユーザは、副ボリュームのボリュームラベルを正ボリュームのそれと違えることにより副ボリュームを正ボリュームから切り離し、更に、その後、ホストを再起動することにより、ホストOSに対して副ボリュームを認識させるようにする。すなわち、ユーザは、手動で、副ボリュームを正ボリュームから切り離し、副ボリュームをホストOSとの間でオンライン状態にする必要がある。

20

【0011】

このように、オフライン状態になっている論理ボリュームに対してデータの読出し又は書込みを行なうためには、ユーザにとって面倒な処理が必要になる。

【0012】

また、例えばホストOSによっては、ユーザが上述のような作業を行なっても、オフライン状態になっている論理ボリュームをオンライン状態にすることができない場合もある。

30

【0013】

更に、例えば二重化処理において、ホストOSに対してオフライン状態である副ボリュームをオンライン状態にするためには、その副ボリュームを正ボリュームから切り離す必要があるが、副ボリュームが正ボリュームから切り離されている間は、二重化処理が行われないので、万一ライトデータの損失があった場合には対応が困難である。

【0014】

以上のような問題点は、仮想的な記憶領域である論理ボリュームに限らず、物理的な記憶領域についても存在し得るものである。

【0015】

従って、本発明の目的は、オンライン状態になっている物理的又は論理的なデータ記憶領域に対してデータアクセスを行うのと同様の方法でそのデータ記憶領域とは別のデータ記憶領域、例えばホストに対してオフライン状態になっているデータ記憶領域に対してデータアクセスを行うことができるようにすることにある。本発明の更なる目的は、後述する実施の形態の記載から明らかになるであろう。

40

【0016】

【課題を解決するための手段】

本発明に従うデータアクセスシステムは、記憶システムと、前記記憶システムと通信可能に接続されているホスト装置とを有するシステムであって、前記記憶システムは、前記ホスト装置に対してオンライン状態であるオンライン記憶領域を含む、物理的又は論理的な

50

複数のデータ記憶領域と、前記ホスト装置をして、所望の又は所定の前記オンライン記憶領域を介して、前記オンライン記憶領域とは別のデータ記憶領域であって、前記ホスト装置がアクセスしたいターゲットであるターゲット記憶領域にアクセスすることを可能にせしめるアクセス制御手段とを備える。

【0017】

なお、ここで言う「ターゲット記憶領域」には、ホスト装置に対してオフライン状態であるデータ記憶領域が含まれても良いし、或いは、ホスト装置に対する複数のオンライン記憶領域のうちの前記所望のオンライン記憶領域以外のオンライン記憶領域（すなわち、アクセス要求の発行先となっていない他のオンライン記憶領域）が含まれても良い。

【0018】

このデータアクセスシステムによれば、ターゲット記憶領域がオンライン状態か否かに関わらず、そのターゲット記憶領域を指すターゲット情報を含んだアクセス要求を、そのターゲット記憶領域とは別の所望のオンライン記憶領域に対して送信することができる。すなわち、ユーザは、オンライン状態である所望のデータ記憶領域に対してデータアクセスを行うのと同様の方法でそのデータ記憶領域とは別のデータ記憶領域に対してデータアクセスを行なうことができる。

【0019】

第1の好適な実施形態では、前記ターゲット記憶領域は、前記ホスト装置に対しオフライン状態であるオフライン記憶領域である。これにより、ユーザは、オンライン状態である所望のデータ記憶領域に対してデータアクセスを行うのと同様の方法で所望のオンライン状態のデータ記憶領域に対してデータアクセスを行なうことができる。

【0020】

第2の好適な実施形態では、前記ホスト装置は、前記ターゲット記憶領域を指すターゲット情報を含んだアクセス要求を、前記所望のオンライン記憶領域に対するアクセス要求として前記記憶システムに送信するアクセス要求手段を有し、前記アクセス制御手段は、前記ホスト装置から前記所望のオンライン記憶領域に対するアクセス要求を受信し、そのアクセス要求が持つターゲット情報に基づいて、前記ターゲット記憶領域からデータをリードして前記ホスト装置に提供する、又は、前記ホスト装置からデータを受けて前記ターゲット記憶領域にライトする。

【0021】

第3の好適な実施形態では、前記アクセス制御手段は、前記ホスト装置に対する通信インターフェースであるホスト通信手段と、前記複数のデータ記憶領域に対してアクセスする記憶領域通信手段と、前記データ記憶領域からリードされたリードデータ又は前記ホスト装置から受信したライトデータを前記ホスト通信手段と前記記憶領域通信手段との間で受け渡すためのキャッシュメモリとを備える。前記キャッシュメモリには、前記リードデータ又は前記ライトデータを一時的に格納するための記憶領域であるキャッシュスロットが確保されるようになっている。前記キャッシュスロットは、一つのデータ記憶領域に対して専用に設けられるようになっている。前記ホスト通信手段は、前記ホスト装置から前記所望のオンライン記憶領域に対するアクセス要求を受信した場合、前記所望のオンライン記憶領域に対してではなく、前記受信したアクセス要求が持つターゲット情報が指すターゲット記憶領域に専用のターゲット用キャッシュスロットを前記キャッシュメモリ上に確保し、それにより、前記ターゲット記憶領域から読み出されたリードデータ又は前記ターゲット記憶領域にライトするライトデータが前記確保したターゲット用キャッシュスロットに一時的に格納されるようにする。

【0022】

第4の好適な実施形態では、前記アクセス制御手段は、前記ホスト装置に対する通信インターフェースであるホスト通信手段と、前記複数のデータ記憶領域に対してアクセスする記憶領域通信手段と、前記複数のデータ記憶領域へのアクセスを制御するためのアクセス制御テーブルとを備えている。前記アクセス制御テーブルには、どのデータ記憶領域にアクセスするかアクセス対象情報を記録することができるようになっており、前記ホスト

10

20

30

40

50

通信手段は、前記ホスト装置から前記アクセス要求を受信した場合、前記アクセス制御テーブルに、前記受信したアクセス要求が持つターゲット情報が指す前記ターゲット記憶領域をアクセス対象としたアクセス対象情報を登録し、前記記憶領域通信手段は、前記アクセス制御テーブルに登録された前記アクセス対象に基づいて、前記ターゲット記憶領域からデータをリードしそのリードデータが前記ホスト通信手段から前記ホスト装置に送信されるようにする、又は、前記ホスト通信手段が前記ホスト装置から受信したライトデータを前記ターゲット記憶領域にライトする。

【0023】

より詳細には、例えば、前記アクセス制御手段は、前記複数のデータ記憶領域に対してそれぞれ用意される複数のアクセス制御テーブルを備えている。各データ記憶領域の前記アクセス制御テーブルには、そのデータ記憶領域のどこにアクセスするかのアクセスアドレスを記録することができるようになっている。前記ホスト通信手段は、前記ホスト装置から前記所望のオンライン記憶領域に対するアクセス要求を受信した場合、前記所望のオンライン記憶領域のアクセス制御テーブルに、前記受信したアクセス要求が持つターゲット情報が指す前記ターゲット記憶領域を識別するためのターゲット領域識別情報を前記アクセスアドレスとして登録し、且つ、前記ターゲット記憶領域の前記アクセス制御テーブルに、前記ターゲット情報に基づいて前記ターゲット記憶領域の前記アクセスアドレスであるターゲットアクセスアドレスを登録する。前記記憶領域通信手段は、前記所望のオンライン記憶領域の前記アクセス制御テーブルに登録された前記ターゲット領域識別情報から前記ターゲット記憶領域を識別して、そのターゲット記憶領域のアクセス制御テーブルを参照し、そのアクセス制御テーブルに登録されている前記ターゲットアクセスアドレスからデータをリードしそのリードデータが前記ホスト通信手段から前記ホスト装置に送信されるようにする、又は、前記ホスト通信手段が前記ホスト装置から受信したライトデータを前記ターゲットアクセスアドレスにライトする。

【0024】

更に詳細には、例えば、前記アクセス制御手段は、前記データ記憶領域からリードされたリードデータ又は前記ホスト装置から受信したライトデータを前記ホスト通信手段と前記記憶領域通信手段との間で受け渡すためのキャッシュメモリを更に備える。前記キャッシュメモリには、前記リードデータ又は前記ライトデータを一時的に格納するための記憶領域であるキャッシュスロットが確保されるようになっている。前記キャッシュスロットは、一つのデータ記憶領域に対して専用に設けられるようになっている。各データ記憶領域の前記アクセス制御テーブルには、更に、そのデータ記憶領域に専用に設けられた前記キャッシュスロットを特定するためのキャッシュポインタが記録されるようになっている。前記ホスト通信手段は、前記ホスト装置から前記所望のオンライン記憶領域に対するアクセス要求を受信した場合、前記所望のオンライン記憶領域に対してではなく、前記受信したアクセス要求が持つターゲット情報が指すターゲット記憶領域に専用のターゲット用キャッシュスロットを前記キャッシュメモリ上に確保して、前記確保したターゲット用キャッシュスロットのキャッシュポインタであるターゲット用キャッシュポインタも前記アクセス制御テーブルに登録し、そのターゲット用キャッシュポインタからリードデータを読み出して前記ホスト装置に送信する、又は、前記ホスト装置からのライトデータを前記ターゲット用キャッシュポインタに格納する。前記記憶領域通信手段は、前記ターゲット記憶領域のアクセス制御テーブルを参照してそのターゲット記憶領域のターゲット用アクセスポインタ及びターゲット用キャッシュポインタを特定し、前記特定されたターゲット用アクセスポインタからデータをリードしそのリードデータを前記特定されたターゲット用キャッシュポインタに格納する、又は、前記ホスト通信手段によって前記ターゲット用キャッシュポインタに格納された前記ライトデータを前記ターゲット用アクセスポインタから読み出して前記ターゲット用アクセスポインタにライトする。

【0025】

第5の好適な実施形態では、前記ホスト装置が送信するアクセス要求には、第1のサブアクセス要求と、第2のサブアクセス要求とが含まれている。前記第1のサブアクセス要求

10

20

30

40

50

は、前記所望のオンライン記憶領域に対する要求であって、前記ターゲット情報を含んでいる（更に、前記所望のオンライン記憶領域とそれとは別のデータ記憶領域のどちらに対してデータのリード又はライトを行なうかを表すアクセス種別を含んでいても良い）。前記第2のサブアクセス要求も、前記所望のオンライン記憶領域に対する要求であって、データのリードとライトのどちらを実行するのかを表すリード/ライト情報を含んでいる。前記アクセス制御手段は、前記第2のアクセス要求の前に前記第1のサブアクセス要求を受信することにより、前記ターゲット情報に基づいてリード/ライト準備処理を行い、続いて、前記第2のサブアクセス要求を受信することにより、前記ターゲット情報が指す前記ターゲット記憶領域に対するデータのリード又はライトを実行する。

【0026】

第6の好適な実施形態では、第3及び第5の好適な実施形態において、前記リード/ライト準備処理には、前記ホスト通信手段が前記キャッシュメモリ上に前記キャッシュスロットを確保することが含まれる。

【0027】

第7の好適な実施形態では、前記記憶システムが、物理的な1又は複数の記憶装置上に、プライマリの論理的なデータ記憶領域であるプライマリ記憶領域と、セコンダリの論理的なデータ記憶領域であるセコンダリ記憶領域とのペアを用意し、前記プライマリ記憶領域と前記セコンダリ記憶領域の両方にデータをライトする二重化処理を行う場合、前記所望のオンライン記憶領域は前記プライマリ記憶領域であり、前記ターゲット記憶領域（例えばオフライン記憶領域）は前記セコンダリ記憶領域である。

【0028】

更に好適には、例えば、前記プライマリ記憶領域内の第1のデータと前記セコンダリ記憶領域内の第2のデータが互いに一致しない場合、前記記憶システムは、前記セコンダリ記憶領域から前記第2のデータをリードし、その第2のデータを前記プライマリ記憶領域内の前記第1のデータに上書きする手段を有する。この場合、例えば、前記ホスト装置は、前記セコンダリ記憶領域の前記第2のデータをアクセス対象とした第2データアクセス要求を前記プライマリ記憶領域に対するものとして前記記憶システムに送信し、前記記憶システムは、前記ホスト装置からの前記第2データアクセス要求に基づいて前記セコンダリ記憶領域から前記第2のデータをリードし、そのリードした第2のデータを前記プライマリ記憶領域の第1のデータに上書きする。

【0029】

第8の好適な実施形態では、第1のホスト装置と第2のホスト装置が前記記憶システムに通信可能に接続されている場合、前記第1のホスト装置にとっての前記ターゲット記憶領域（例えばオフライン記憶領域）には、前記第2のホスト装置に対してオンライン状態であるデータ記憶領域が含まれている。

【0030】

第9の好適な実施形態では、第3の好適な実施形態において、第1のホスト装置と第2のホスト装置が前記記憶システムに通信可能に接続されている場合、前記キャッシュメモリには、第1のホスト装置用のキャッシュ領域である第1ホスト用領域と第2のホスト装置用のキャッシュ領域である第2ホスト用領域とが用意されており、前記第1のホスト装置にとっての前記ターゲット記憶領域（例えばオフライン記憶領域）には、前記第2のホスト装置に対してオンライン状態である第2ホストオンライン記憶領域が含まれており、前記ホスト通信手段は、前記第1のホスト装置から、前記第2ホストオンライン記憶領域をターゲット記憶領域とするターゲット情報を持ったアクセス要求を受けたときは、前記キャッシュメモリ上の前記第2ホスト用領域内に、前記第2ホストオンライン記憶領域用の前記キャッシュスロットを確保する。

【0031】

本発明に従う記憶システムは、通信接続可能なホスト装置に対してオンライン状態であるオンライン記憶領域を含む、物理的又は論理的な複数のデータ記憶領域と、前記ホスト装置をして、所望の又は所定の前記オンライン記憶領域を介して、前記オンライン記憶領域

10

20

30

40

50

とは別のデータ記憶領域であって、前記ホスト装置がアクセスしたいターゲットであるターゲット記憶領域にアクセスすることを可能にせしめるアクセス制御手段とを備える。

【0032】

具体的には、例えば、記憶システムは、物理的又は論理的な複数のデータ記憶領域と、前記ホスト装置の各データ記憶領域に対するデータアクセスを制御するアクセス制御手段とを備えている。前記複数のデータ記憶領域には、前記ホスト装置に対してオンライン状態であるオンライン記憶領域と、前記ホスト装置に対しオフライン状態であるオフライン記憶領域とが含まれている場合、前記アクセス制御手段は、アクセスしたいターゲットのデータ記憶領域であるターゲット記憶領域を指すターゲット情報を含んだアクセス要求を、前記ターゲット記憶領域とは別の所望のオンライン記憶領域に対するアクセス要求として前記ホスト装置から受信し、そのアクセス要求が持つターゲット情報に基づいて、前記ターゲット記憶領域からデータをリードして前記ホスト装置に提供する、又は、前記ホスト装置からデータを受けて前記ターゲット記憶領域にライトする。

10

【0033】

本発明に従うホスト装置は、物理的又は論理的な複数のデータ記憶領域を持つ記憶システムに対して通信可能に接続することができるホスト装置であって、前記複数のデータ記憶領域には、前記ホスト装置に対してオンライン状態であるオンライン記憶領域が含まれている場合、所望の又は所定の前記オンライン記憶領域を介して、前記オンライン記憶領域とは別のデータ記憶領域であって、前記ホスト装置がアクセスしたいターゲットであるターゲット記憶領域にアクセスする手段を備える。具体的には、例えば、ホスト装置は、アクセスしたいターゲットのデータ記憶領域であるターゲット記憶領域を指すターゲット情報を含んだアクセス要求を、前記ターゲット記憶領域とは別の所望のオンライン記憶領域に対するアクセス要求として生成する要求生成手段と、前記生成したアクセス要求を前記記憶システムに送信するアクセス要求手段とを備える。

20

【0034】

好適な実施形態では、ホスト装置は、更に、各データ記憶領域毎にそのデータ記憶領域がオンライン状態であるかオフライン状態であるかを示すオンライン/オフライン識別情報と、前記オンライン/オフライン識別情報に基づいて各データ記憶領域がオンライン状態であるかオフライン状態であるかを識別し、オンライン状態のときには前記アクセス要求手段を実行するようにし、オフライン状態のときにはエラーを出力する制御手段と、前記ターゲット記憶領域がオフライン記憶領域の場合には、そのターゲット記憶領域がオンライン状態であることを示すように前記オンライン/オフライン識別情報を更新する情報更新手段とを備える。この場合、前記制御手段が、前記情報更新手段によって更新された後の前記オンライン/オフライン識別情報に基づいて動作するようになっている。

30

【0035】

この実施形態によれば、ターゲット記憶領域がオフライン記憶領域であっても、制御手段が参照するオンライン/オフライン識別情報では、そのターゲット記憶領域はオンライン記憶領域であると更新され、制御手段は、その更新後のオンライン/オフライン識別情報を参照するので、制御手段によってエラーが出力されることはない。

【0036】

また、この実施形態によれば、例えば、ホスト装置のOSが、前記情報更新手段を有していないが故に上記のようにエラーを出力するようになっているとしても、そのOSに前記情報更新手段を組み込むことによりエラーが出力されないようにすることが可能になるので、既存のホスト装置のOSを簡単に設計変更するだけで、ホスト装置が、所望のオンライン記憶領域にデータアクセスするのと同様にオフライン記憶領域にアクセスすることができるようにすることが可能である。

40

【0037】

本発明に従うコンピュータプログラム（又はそれを記録した記録媒体）は、物理的又は論理的な複数のデータ記憶領域を持つ記憶システムに対して通信可能に接続することができるホスト装置として動作するためのコンピュータプログラムにおいて、前記複数のデータ

50

記憶領域には、前記ホスト装置に対してオンライン状態であるオンライン記憶領域が含まれている場合、所望の又は所定の前記オンライン記憶領域を介して、前記オンライン記憶領域とは別のデータ記憶領域であって、前記ホスト装置がアクセスしたいターゲットであるターゲット記憶領域にアクセスするステップをコンピュータに実行させるためのコンピュータプログラムである。

【0038】

本発明のデータアクセスシステム、記憶システム及びホスト装置は、専用ハードウェア、プログラムされたコンピュータ、又はそれらの組合せのいずれによっても実現することができる。

【0039】

【発明の実施の形態】

図1は、本発明の一実施形態に係るアクセスシステムの全体構成を示す。

【0040】

このアクセスシステム100には、ディスクアレイ装置(例えばRAID)3と、そのディスクアレイ装置3に通信可能に接続することができるホストシステム(以下、単に「ホスト」と言う)1とが備えられる(一つのホスト1に複数台のディスクアレイ装置3が接続されても良いし、一つのディスクアレイ装置3に複数台のホスト1が接続されても良い)。ディスクアレイ装置3は、アレイ状に配設された複数のディスク型記憶装置を有する1以上の物理ディスク群9を有し、それら物理ディスク群9により提供される物理的な記憶領域上には、論理的な記憶領域である複数の論理ボリューム(Logical Unit)11が設定される。

【0041】

ホスト1は、例えば、パーソナルコンピュータやワークステーション等であり、CPU(Central Processing Unit)やメモリ等を備えたコンピュータシステムである。ホスト1のCPUが各種プログラムを実行することにより、種々の機能が実現される。ホスト1は、例えば、LAN等の通信ネットワークを介してディスクアレイ装置3に接続される。

【0042】

このホスト1は、図示のように、1又は複数のアプリケーションプログラム(以下、「ホストAP」と言う)5と、オペレーティングシステム(以下、「ホストOS」と言う)7とを有している。

【0043】

ホストAP5は、所定の又はユーザから指定された論理ボリューム(以下、「ターゲットボリューム」と言う)11に対するデータの書込み処理又は読出し処理(以下、「I/O処理」と言う)の要求をホストOS7に発行する。

【0044】

ホストOS7は、ホストAP5からのI/O処理要求に応答して、ターゲットボリューム11がオンライン状態であるか否かのチェックを行ない、その結果に応じた内容を持つデータ書込み/読出しコマンド(以下、「アクセスコマンド」と言う)を生成し、そのコマンドをディスクアレイ装置3に送信する。

【0045】

ディスクアレイ装置3は、1又は複数のチャネル制御部13と、キャッシュメモリ15と、制御メモリ17と、1以上の物理ディスク群9と、1以上のディスク制御部19と、スイッチング制御部21とを備えている。

【0046】

1又は複数のチャネル制御部13は、1又は複数のホスト1と通信可能に接続される。各チャネル制御部13は、ハードウェア回路、ソフトウェア、又はそれらの組み合わせで構成することができ、このディスクアレイ装置3とそれに接続されているホスト1との間のデータ通信を行う。各チャネル制御部13は、接続されているホスト1と通信を行うための通信インターフェースを有し、また、ホスト1から受信した各種コマンドを解釈して処理するためのコマンドプロセッサ機能も有する。各チャネル制御部13は、制御メモリ1

10

20

30

40

50

7に登録されている情報を参照して、キャッシュメモリ15に格納されているデータを読み出してホスト1に送信したり、ホスト1から受信した書込み対象のデータ(以下、「ライトデータ」と言う)をキャッシュメモリ15に格納したりする。

【0047】

キャッシュメモリ15は、各チャンネル制御部13や各ディスク制御部19により共有されるメモリである。キャッシュメモリ15には、各チャンネル制御部13によって、1又は複数の仮想記憶領域(以下、「キャッシュスロット」と言う)を設定することができる。キャッシュスロットは、或る一つの論理ボリュームに対して設定される。具体的には、キャッシュスロットには、そのキャッシュスロットに対応している論理ボリュームにライトされるデータ、又は、その論理ボリュームからリードされたデータが一時的に格納される。

10

【0048】

制御メモリ17も、各チャンネル制御部13や各ディスク制御部19により共有されるメモリである。制御メモリ17には、1以上の物理ディスク群9上に用意される各論理ボリューム11毎に制御テーブル(図示せず)が用意され、各制御テーブルには後述する所定の種々の情報が格納される。キャッシュメモリ15上のキャッシュスロットや、各論理ボリューム11に対するデータの読出しや書込みは、その論理ボリュームに対応した制御テーブルに記録されている所定の情報に基づいて行なわれる。

【0049】

各物理ディスク群9は、複数のディスク型記憶装置をアレイ状に配設して構成されている。物理ディスク群9を構成するディスク型記憶装置としては、例えば、ハードディスク装置、フレキシブルディスク装置、半導体記憶装置等の種々のデバイスを用いることができる。1以上の物理ディスク群9により提供される物理的な記憶領域上には、前述したように、論理的な記憶領域である複数の論理ボリューム11が設定される。

20

【0050】

各ディスク制御部19は、各物理ディスク群9毎に備えられており、その所定の物理ディスク群9を制御するものである。例えば、ディスク制御部19は、制御メモリ17上のターゲットボリューム11の制御テーブルに基づいて、1以上の物理ディスク群9上のターゲットボリューム11に対してデータの読出し又は書込みを行なう。また、ディスク制御部19は、論理アドレスを物理アドレスに変換することにより、論理ボリューム11に対するデータアクセス要求を物理ディスクへのデータアクセス要求に変換する。

30

【0051】

スイッチング制御部21は、例えば、高速スイッチング動作によりデータ伝送を行う超高速クロスバススイッチ等のような高速バスとして構成することができる。スイッチング制御部21は、各チャンネル制御部13と、各ディスク制御部19と、制御メモリ17と、キャッシュメモリ15とを相互に通信可能に接続する。これらの各チャンネル制御部13と、各ディスク制御部19と、制御メモリ17と、キャッシュメモリ15との間のデータやコマンドの授受は、スイッチング制御部21を介して行われる。

【0052】

以上が、この実施形態に係るアクセスシステム100の概要である。なお、このアクセスシステム100では、例えば、各ディスク制御部19と各物理ディスク群9とは、直接的に接続してもよいし、ネットワークを介して間接的に接続してもよい。また、物理ディスク群9とディスク制御部19とを一体的に構成しても良い。また、各チャンネル制御部13毎にそれぞれ1つの論理ボリューム11が割り当てられてもよいし、複数のチャンネル制御部13で1つの論理ボリューム11が共有されても良い。

40

【0053】

さて、以下、本実施形態に係るアクセスシステム100をより詳細に説明する。

【0054】

まず、ホスト1がディスクアレイ装置3に対して発行するアクセスコマンドについて説明する。

【0055】

50

図2は、ホスト1が発行するアクセスコマンドを示す。

【0056】

ホスト1がディスクアレイ装置3に対して発行するアクセスコマンド19は、第1のサブコマンド19Aと、第2のサブコマンド19Bとを含んでいる。

【0057】

第1のサブコマンド19Aは、第2のサブコマンド19Bよりも先に（例えば最初に）ホスト1からディスクアレイ装置3に送信されるパケットである。このサブコマンド19Aには、コマンド対象ボリュームID、第1アクセス種別コード、及びターゲット情報が含まれている。

【0058】

「コマンド対象ボリュームID」とは、アクセスコマンド19の発行先とする論理ボリュームの識別コードであり、この実施形態では、必ず、起動ボリュームのIDになる。「起動ボリューム」とは、ホスト1に対してオンライン状態である1又は複数の論理ボリューム（以下、「オンラインボリューム」と言う）の中からホスト1によって選択されたオンラインボリュームのことであり、アクセスコマンド19の発行先となるものである（起動ボリュームは、所定の規則に基づいてホスト1又はディスクアレイ装置3によって自動的に選択されても良いし、ユーザが任意に選択しても良い）。

【0059】

「第1アクセス種別コード」とは、ホスト1がどのような種別のアクセスを行ないたいかを示すコードである。第1のアクセス種別には、例えば、リード、ライト、リードコントロール及びライトコントロールの4種類がある（このような4種類に限定しなくても良い）。第1アクセス種別コードとして「リード」を示すコードが含まれている場合、起動ボリュームからデータをリードすることを意味する。同様に、第1のアクセス種別コードとして「ライト」を示すコードが含まれている場合、起動ボリュームにデータをライトすることを意味する。一方、「リードコントロール」を示すコードが含まれている場合、起動ボリュームとは別の論理ボリューム（例えば、起動ボリューム以外のオンラインボリューム、又は、ホスト1に対してオフライン状態である論理ボリュームであるオフラインボリューム）からデータをリードすることを意味する。また、「ライトコントロール」を示すコードが含まれている場合、起動ボリュームとは別の論理ボリュームにデータをライトすることを意味する。

【0060】

「ターゲット情報」とは、ホスト1がリードしたいデータを持っている論理ボリューム又はホスト1がデータをライトしたい論理ボリューム（つまりターゲットボリューム）に関する情報である。具体的には、例えば、第1のサブコマンド19Aにおいて、アクセス種別コードがリード又はライトを示すコードである場合、ターゲット情報には、起動ボリュームのどこにアクセスするかアクセス先を表すアドレス（以下、「アクセスアドレス」と言う）が含まれている。一方、第1のサブコマンド19Aにおいて、アクセス種別コードがリードコントロール又はライトコントロールを示すコードである場合、ターゲット情報には、起動ボリュームとは別の論理ボリュームのボリュームIDと、その別の論理ボリュームのどこにアクセスするかアクセス先を表すアクセスアドレスとが含まれている。

【0061】

第2のサブコマンド19Bは、第1のサブコマンド19Aがホスト1からディスクアレイ装置3に送信された後、所定のタイミングで、ホスト1からディスクアレイ装置3に送信されるパケットである。このサブコマンド19Bには、コマンド対象ボリュームID、第2アクセス種別コード、及び付随データが含まれている。

【0062】

「コマンド対象ボリュームID」は、第1のサブコマンド19Aに含まれるコマンド対象ボリュームIDと同じである。すなわち、アクセスコマンド19は、ホスト1が起動ボリュームにアクセスしたいのか或いは別の論理ボリュームにアクセスしたいのかに関わらず（換言すれば、ターゲットボリュームが起動ボリュームであるか或いは別の論理ボリュー

10

20

30

40

50

ムであるかに関わらず)、起動ボリュームをコマンドの対象としているため、第1のサブコマンド19Aと同様に、このサブコマンド19Bのコマンド対象ボリュームIDは、起動ボリュームのIDになる。

【0063】

「第2アクセス種別コード」は、第1アクセス種別コードと異なり、リードとライトのいずれかを示すコードになる。すなわち、この第2アクセス種別コードとして「リード」を示すコードが含まれている場合、第1のサブコマンド19Aのターゲット情報から特定可能なターゲットボリューム(すなわち、起動ボリューム又は起動ボリュームとは別の論理ボリューム)からデータをリードすることを意味する。同様に、第2のサブコマンド19Bにアクセス種別コードとして「ライト」を示すコードが含まれている場合、そのターゲットボリュームにデータをライトすることを意味する。

10

【0064】

「付属データ」とは、第2のサブコマンド19Bに付属するデータである。具体的には、例えば、付属データは、第2アクセス種別コードがライトを示すコードのときは、ホスト1がターゲットボリュームにライトしたいライトデータを含み、一方、第2アクセス種別コードがリードを示すコードのときは、ライトデータの代わりにヌルのデータを含む。

【0065】

以上が、ホスト1がディスクアレイ装置3に対して発行するアクセスコマンドについての説明である(どのような流れでホスト1が第1及び第2のサブコマンド19A、19Bを送信し、その第1及び第2のサブコマンド19A、19Bに基づいてディスクアレイ装置3がどのような流れで動作するかについては後で詳述する)。

20

【0066】

次に、ディスクアレイ装置3が有する制御メモリ17上に用意される制御テーブルについて説明する。

【0067】

図3は、制御メモリ17に用意される制御テーブルを示す。

【0068】

制御テーブル31は、前述したように、各論理ボリューム11毎に用意される。各制御テーブル31には、ボリュームID、アクセス種別、アクセス対象情報、及びキャッシュポインタ等が記録される。

30

【0069】

「ボリュームID」は、対応する論理ボリュームのIDである。各チャンネル制御部13及びディスク制御部19は、このボリュームIDを参照することにより、どの制御テーブル31がどの論理ボリューム11に対応しているのかを認識することができる。

【0070】

「アクセス種別」は、対応する論理ボリュームに対してどのような種別のアクセスが行なわれるのかを示すものである。具体的には、例えば、ターゲットボリュームが起動ボリュームの場合には、起動ボリュームの制御テーブル31に、アクセス種別として、リード又はライトが登録される。一方、ターゲットボリュームが起動ボリュームとは別の論理ボリューム(以下、「別ボリューム」と略記する)の場合には、起動ボリュームの制御テーブル31に、アクセス種別として、リードコントロール又はライトコントロールが登録され、別ボリュームの制御テーブル31には、アクセス種別として、リード又はライトが登録される。

40

【0071】

「アクセスアドレス」は、対応する論理ボリュームのどこにアクセスするかを表すアクセス先のアドレス(例えば論理アドレス)を示す。

【0072】

「キャッシュポインタ」とは、対応する起動ボリュームに対してキャッシュメモリ15上に専用に用意されたキャッシュスロットを指し示す情報である。各チャンネル制御部13及び各ディスク制御部19は、各制御テーブル31に登録されているキャッシュポインタを

50

参照することにより、その制御テーブル31に対応した論理ボリューム11からリードしたデータ又はその論理ボリューム11にライトするデータをどのキャッシュスロットに格納すれば良いのかを識別することができる。

【0073】

以上が、制御テーブル31についての説明である。

【0074】

次に、本実施形態に係るデータアクセスシステムにおいて、ホスト1が起動ボリュームとは別の論理ボリュームをターゲットとしそこからデータをリードする場合の流れを簡単に説明する。なお、以下の説明では、ターゲットボリュームを、ホスト1に対してオフライン状態であるオフラインボリュームであるとする。

10

【0075】

図4は、本実施形態に係るデータアクセスシステムにおいて、ホスト1が起動ボリュームとは別の論理ボリュームをターゲットとしそこからデータをリードする場合の流れを示す。

【0076】

ホストAP5は、所定の又はユーザから指定されたオフラインボリューム（つまりターゲットボリューム）11Tから所定の又はユーザ指定のデータをリードする旨を表したI/O処理要求をホストOS7に発行する。

【0077】

ホストOS7は、ホストAP5からのI/O処理要求にตอบสนองして、オフラインボリューム11を指すターゲット情報を持ったアクセスコマンド19を生成し、それを起動ボリューム11Kに対するコマンドとしてディスクアレイ装置3に送信する。

20

【0078】

ホスト1と通信可能に接続されているチャンネル制御部13は、ホスト1から、起動ボリューム11Kに対するアクセスコマンド19を受信して、そのアクセスコマンド19に含まれている第1アクセス種別コードを解析する。その結果、その第1アクセス種別コードが、起動ボリュームとは別の論理ボリュームからデータをリードするリードコントロールを意味している場合には、チャンネル制御部13は、次のように動作する。

【0079】

すなわち、チャンネル制御部13は、キャッシュメモリ15上にキャッシュスロット21を確保するが、その際、アクセスコマンド19が起動ボリューム11Kに対するものであっても、ターゲットがオフラインボリューム11Tなので、起動ボリューム11Kに対してではなくオフラインボリューム11Tに対してキャッシュスロット21を確保する。そして、チャンネル制御部13は、アクセスコマンド19に含まれているターゲット情報から特定できるオフラインボリューム11Tの制御テーブル31Tに、オフラインボリューム11Tに対して確保したキャッシュスロット21のキャッシュポイントを登録する（起動ボリューム11Kの制御テーブル31Kには、キャッシュポイントを登録しない）。また、チャンネル制御部13は、アクセスコマンド19に含まれているターゲット情報から特定できるオフラインボリューム11Tのアクセスアドレスも、制御テーブル31Tに登録する。

30

40

【0080】

各ディスク制御部19は、例えば、自分が接続されている物理ディスク群9上に用意された論理ボリューム11のIDを把握しており、その論理ボリューム11の制御テーブル（以下、「対象制御テーブル」と言う）を常に監視している。その結果、各ディスク制御部19は、対象制御テーブルに新たにキャッシュポイント、アクセスアドレス及びアクセス種別等が登録されたら、そのアクセス種別に応じた処理を行う。

【0081】

すなわち、この図4の例で言えば、オフラインボリューム11Tが用意されている物理ディスク群9に対応したディスク制御部（以下、「担当ディスク制御部」と言う）19は、制御テーブル31Tに新たにキャッシュポイント及びアクセスアドレス等が登録されたこ

50

とを検出したならば、そのアクセスアドレス（すなわちオフラインボリューム 1 1 T におけるアドレス）からデータをリードし、リードしたデータを（以下、「リードデータ」と言う）、制御テーブル 3 1 T に登録されているキャッシュポインタから識別できるキャッシュスロット 2 1 に格納する。

【 0 0 8 2 】

チャンネル制御部 1 3 は、所定の方法（例えば、上記確保したキャッシュスロット 2 1 を監視する方法）でキャッシュスロット 2 1 にリードデータが格納されたことを検出したら、そのリードデータをキャッシュスロット 2 1 から読出してホスト 1 に送信する。

【 0 0 8 3 】

このような流れにより、ホスト 1 は、起動ボリューム 1 1 K に対するアクセスコマンド 1 9 をディスクアレイ装置 3 に送信しても、起動ボリューム 1 1 K とは別の論理ボリューム（例えばオフラインボリューム 1 1 T ）からデータをリードすることができる。

【 0 0 8 4 】

なお、ホスト 1 及びディスクアレイ装置 3 についての動作については後で更に詳細に説明するので、ここでは詳しくは説明しないが、図 4 を参照した上記説明から、別ボリュームにデータをライトする場合の動作も理解することができる。

【 0 0 8 5 】

すなわち、チャンネル制御部 1 3 は、ターゲットボリューム 1 1 T に対してキャッシュスロット 2 1 を確保してそのキャッシュスロットを制御テーブル 3 1 T に登録し、且つ、そのキャッシュスロット 2 1 に、ホスト 1 からのライトデータを格納する。担当ディスク制御部 1 9 は、制御テーブル 3 1 T に登録されているキャッシュポインタからライトデータがあるキャッシュスロット 2 1 を識別し、そのキャッシュスロット 2 1 から、ライトデータを読み出して、そのライトデータを、制御テーブル 3 1 T に登録されているアクセスアドレスにライトする。

【 0 0 8 6 】

上記のことをより抽象的に表現すれば、例えば、ディスクアレイ装置 3 は、ホスト 1 をして、所望の又は所定のオンラインボリューム（すなわち起動ボリューム）1 1 K を介して、そのオンラインボリューム 1 1 K とは別の論理ボリュームであって、ホスト 1 がアクセスしたいターゲットであるターゲットボリューム 1 1 T にアクセスすることを可能にせしめる。別の観点から言えば、例えば、ホスト 1 は、ディスクアレイ装置 3 が有する複数の論理ボリュームのうちの所望の又は所定のオンラインボリューム 1 1 K を介して、そのオンラインボリューム 1 1 K とは別の論理ボリュームであるターゲットボリューム 1 1 T にアクセスする。

【 0 0 8 7 】

さて、以下、ホスト 1 及びディスクアレイ装置 3 の動作について更に詳細に説明する。

【 0 0 8 8 】

図 5 は、ホスト OS 7 の動作流れを示す。

【 0 0 8 9 】

まず、動作を説明する前に、ホスト OS 7 が管理している 2 種類の情報について説明する。

【 0 0 9 0 】

ホスト OS 7 は、各論理ボリューム 1 1 毎に、オンラインビットとオープンビットを管理している。

【 0 0 9 1 】

「オンラインビット」とは、対応する論理ボリュームがホスト 1 に対してオンライン状態か否かを表すものである。このオンラインビットがセットされていれば（例えば「1」であれば）、対応する論理ボリュームがオンライン状態であることを示し、セットされていなければ（例えば「0」であれば）、その論理ボリュームがオフライン状態であることを示す。ホスト OS 7 は、任意のタイミングで、各論理ボリュームがオンライン状態かオフライン状態かを監視し、その監視結果に応じて、各論理ボリュームのオンラインビットを

10

20

30

40

50

セットしたり、或いは、セットされているオンラインビットをはずしたりする（例えば「1」から「0」にしたりする）。

【0092】

「オープンビット」とは、対応する論理ボリュームに対し、I/O処理を要求するホストAP5以外の他のホストAP（図示せず）がアクセスしてはならないか否かをホストOS7に対して示すものである。このオープンビットがセットされていれば（例えば「1」であれば）、対応する論理ボリュームに、I/O処理を要求するホストAP5以外の他のホストAP（図示せず）がアクセスさせてはならないことをホストOS7に対して示し（つまり排他処理実行中であることを示し）、セットされていなければ（例えば「0」であれば）、その論理ボリュームにどのホストAP5がアクセスしても良いこと（つまり排他処理が行われていないこと）を示す。

10

【0093】

次に、このようなオンラインビット及びオープンビットを管理するホストOS7がどのような流れで動作するのかを説明する。

【0094】

ホストOS7は、ホストAP5から、ターゲットボリューム11Tの通知と共に、そのターゲットボリューム11Tに対してI/O処理を行うことのI/O処理要求を受ける（ステップS11）。

【0095】

ホストOS7は、ホストAP5からのI/O処理要求に回答して、ターゲットボリューム11Tに対してオープン処理を行うことを開始する（S12）。この「オープン処理」とは、ターゲットボリューム11Tに対して他のホストAP（図示せず）がアクセスしないようにする排他処理のことである。

20

【0096】

まず、ホストOS7は、ターゲットボリューム11Tにオンラインビットがセットされているか否かに関わらず、ターゲットボリューム11Tのオンラインビットをセットする（S13）。

【0097】

次に、ホストOS7は、ターゲットボリューム11Tにオンラインビットがセットされているか否かを確認する（S14）。ここで、ホストOS7は、S13でオンラインビットがセットされたので、オンラインビットがセットされていることを検出し（S14でYES）、そのターゲットボリューム11Tに対してオープンビットをセットする（S16）。そして、ホストOS7は、そのターゲットボリューム11Tを起動ボリューム11Kとする、或いは、そのターゲットボリューム11Tとは別のオンラインボリュームを起動ボリューム11Kとするアクセスコマンド19（厳密に言うと、この段階では第1のサブコマンド19A）を生成してディスクアレイ装置3に送信する。

30

【0098】

以上が、ホストOS7の動作流れである。

【0099】

なお、上記流れにおいて、ホストOS7は、もし、S14の結果、オンラインビットがセットされていないことを検出したならば（S14でNO）、ターゲットボリューム11Tがオフラインボリュームということであり、エラーをホストAP5に通知するようになっている（S15）。しかし、前述した工夫、すなわち、ターゲットボリューム11Tにオンラインビットがセットされているか否かを確認する処理（S14の処理）の前に、ターゲットボリューム11Tにオンラインビットがセットされているか否かに関わらずターゲットボリューム11Tのオンラインビットをセットする処理（S13の処理）を行うという工夫がホストOS7に施されていることにより、ターゲットボリューム11Tがオフラインボリュームであっても、ホストOS7からホストAP5にエラーが通知されることなく、ターゲットボリューム11Tがオンラインボリュームであるときと同様に、ホストOS7からディスクアレイ装置3にアクセスコマンド19が送信される。

40

50

## 【 0 1 0 0 】

また、上述の S 1 2 ~ S 1 6 の処理は、マクロによって行なわれても良い。従って、S 1 4 の処理の前に、ターゲットボリューム 1 1 T にオンラインビットがセットされているか否かに関わらず強制的にターゲットボリューム 1 1 T のオンラインビットをセットする処理 ( S 1 3 の処理 ) を行なうようになったマクロをホスト O S 7 に搭載すれば、ホスト O S 7 は前述した図 5 の動作を実行することができる。また、ホスト O S 7 に既に搭載されているマクロが、S 1 2 の処理の後、S 1 3 の処理を行わずに S 1 4 の処理を行うようになっていても、S 1 2 の処理と S 1 4 の処理との間に S 1 3 の処理を入れ込むことで、前述した S 1 2 ~ S 1 6 の動作を行なうことができるマクロに変更することができる。

## 【 0 1 0 1 】

また、図 5 の処理は、ターゲットボリューム 1 1 T がオフラインボリュームの場合にのみ行なわれるようにしても良い。具体的には、例えば、S 1 2 において、ホスト O S 7 は、一次的にターゲットボリューム 1 1 T にオンラインビットがセットされているか否かを確認し、その結果、オンラインビットがセットされていないときのみ、S 1 3 の処理を行い、オンラインビットが既にセットされているときは ( つまりターゲットボリューム 1 1 T がオンラインボリュームの場合には )、S 1 3 の処理を省略して S 1 4 の処理 ( 二次的にターゲットボリューム 1 1 T にオンラインビットがセットされているか否かを確認する処理 ) を行っても良い。

## 【 0 1 0 2 】

また、図示しないが、ホスト O S 7 は、第 1 のサブコマンド 1 9 A を送信した後、任意のタイミングで ( 例えば、ディスクアレイ装置 3 から所定内容の返答を受けたときに )、第 2 のサブコマンド 1 9 B を生成し送信することができる。具体的には、例えば、ホスト O S 7 は、送信した第 1 のサブコマンド 1 9 A に含まれる第 1 アクセス種別コードが、リード又はリードコントロールを示すコードであれば、リードを示す第 2 アクセス種別コードを持った第 2 のサブコマンド 1 9 B を生成し送信し、一方、上記第 1 アクセス種別コードがライト又はライトコントロールを示すコードであれば、ライトを示す第 2 アクセス種別コードを持った第 2 のサブコマンド 1 9 B を生成し送信する。

## 【 0 1 0 3 】

図 6 ~ 図 1 0 は、ディスクアレイ装置 3 の動作流れを示す。具体的には、図 6 は、ホスト 1 から第 1 のサブコマンド 1 9 A を受信したときのディスクアレイ装置 3 の動作流れを示し、図 7 ~ 図 1 0 は、ホスト 1 から第 2 のサブコマンド 1 9 B を受信したときのディスクアレイ装置 3 の動作流れを示す。

## 【 0 1 0 4 】

まず、図 6 を参照して、ホスト 1 から第 1 のサブコマンド 1 9 A を受信したときのディスクアレイ装置 3 の動作流れを説明する。

## 【 0 1 0 5 】

図 6 に示すように、ディスクアレイ装置 3 のチャンネル制御部 1 3 は、ホスト 1 から第 1 のサブコマンド 1 9 A を受信した場合 ( S 2 1 )、そのサブコマンド 1 9 A に含まれているコマンド対象ボリューム I D から起動ボリューム 1 1 K を特定すると共に、そのサブコマンド 1 9 A に含まれている第 1 アクセス種別コードを解析する ( S 2 2 )。

## 【 0 1 0 6 】

S 2 2 の結果、第 1 アクセス種別コードがリード又はライトを示している場合、ターゲットが起動ボリューム 1 1 K であることを意味するので、チャンネル制御部 1 3 は、上記特定された起動ボリューム 1 1 K の I D を持つ制御テーブル ( 以下、「起動制御テーブル」と言う ) 3 1 K に、アクセス種別としてリード又はライトを登録し、且つ、起動ボリューム 1 1 K のアクセスアドレスを登録する ( S 2 3 )。そして、チャンネル制御部 1 3 は、第 1 のサブコマンド 1 9 A についての処理が完了した旨を表すコマンド処理終了をホスト 1 に返答する ( S 2 6 )。

## 【 0 1 0 7 】

一方、S 2 2 の結果、第 1 アクセス種別コードがリードコントロール又はライトコントロ

10

20

30

40

50

ールを示している場合、ターゲットは起動ボリューム 1 1 Kとは別のボリューム（例えばオフラインボリューム）1 1 Tであることを意味するので、チャンネル制御部 1 3は、起動制御テーブル 3 1 Kに、アクセス種別としてリードコントロール又はライトコントロールを登録し、且つ、受信したサブコマンド 1 1 Aに含まれるターゲット情報が示すターゲットボリューム 1 1 TのボリュームIDを、制御テーブル 3 1 Kに登録する（S 2 4）。また、チャンネル制御部 1 3は、ターゲットボリューム 1 1 Tの制御テーブル（以下、「ターゲット制御テーブル」と言う）3 1 Tに、アクセス種別としてリード又はライトを登録し、且つ、ターゲットボリューム 1 1 Tのアクセスアドレスを登録する（S 2 5）。それが完了したら、チャンネル制御部 1 3は、上述したコマンド処理終了をホスト 1 に返答する（S 2 6）。

10

## 【0108】

ホストOS 7は、チャンネル制御部 1 3からコマンド処理終了の返答を受けたら、第2のサブコマンド 1 9 Bを生成し送信する。このとき送信されるサブコマンド 1 9 Bに含まれる第2アクセス種別コードは、前述したように、先に送信した第1のサブコマンド 1 9 Aに含まれる第1アクセス種別コードがリード又はリードコントロールを示すコードのときは、リードを示すコードであり、上記第1アクセス種別コードがライト又はライトコントロールを示すコードのときは、ライトを示すコードである。

## 【0109】

さて、次に、図7～10を参照して、ホスト1から第2のサブコマンド 1 9 Bを受信したときのディスクアレイ装置3の動作流れを説明する。なお、その際、説明を分かり易くするために、以下の4つの場合、

20

（1）起動ボリューム 1 1 Kからデータをリードする場合（換言すれば、先に受信した第1のサブコマンド 1 9 Aが持つ第1アクセス種別コードがリードを示す場合、以下、「通常リードの場合」と言う）、

（2）起動ボリューム 1 1 Kとは別のボリューム 1 1 Tからデータをリードする場合（換言すれば、先に受信した第1のサブコマンド 1 9 Aが持つ第1アクセス種別コードがリードコントロールを示す場合、以下、単に「リードコントロールの場合」と言う）、

（3）起動ボリューム 1 1 Kとは別のボリューム 1 1 Tにデータをライトする場合（換言すれば、先に受信した第1のサブコマンド 1 9 Aが持つ第1アクセス種別コードがライトコントロールを示す場合、以下、単に「ライトコントロールの場合」と言う）、

30

（4）起動ボリューム 1 1 Kにデータをライトする場合（換言すれば、先に受信した第1のサブコマンド 1 9 Aが持つ第1アクセス種別コードがライトを示す場合、以下、「通常ライトの場合」と言う）、

に分けて説明する。

## 【0110】

（1）通常リードの場合。

## 【0111】

図7に示すように、チャンネル制御部 1 3は、ホスト1から第2のサブコマンド 1 9 Bを受信すると（S 3 1）、そのサブコマンド 1 9 Bに含まれているコマンド対象ボリュームIDから起動ボリューム 1 1 Kを特定すると共に、そのサブコマンド 1 9 Bに含まれている第2アクセス種別コードを解析する（S 3 2）。

40

## 【0112】

S 3 2の結果、チャンネル制御部 1 3は、第2アクセス種別コードがリードを示しているときは、第2のサブコマンド 1 9 Bの対象となっている起動ボリューム 1 1 Kの制御テーブル 3 1 Kを参照して、アクセス種別としてリードとリードコントロールのどちらが登録されているかを調べる（S 3 3）。

## 【0113】

S 3 3の結果、リードが登録されていれば、チャンネル制御部 1 3は、起動制御テーブル 3 1 Kにキャッシュポイントが登録されているか否か（以下、これを「キャッシュヒットしたか否か」と言う）を確認する（S 3 4）。

50

## 【 0 1 1 4 】

S 3 4の結果、キャッシュヒットしなければ(S 3 5でN)、チャンネル制御部 1 3は、キャッシュメモリ 1 5上に起動ボリューム 1 1 Kに専用のキャッシュスロット 2 1を確保し、そのキャッシュスロット 2 1のキャッシュポイントを起動制御テーブル 3 1 Kに登録する(S 3 6)。その後、チャンネル制御部 1 3は、ホスト 1に切り離し要求(具体的には、例えば、複数の通信チャンネルのうちの所定の通信チャンネルを切断する要求)をする(S 3 7)。それにより、ホスト 1によって、ホスト 1とディスクアレイ装置 3との間で切り離しが行なわれる。

## 【 0 1 1 5 】

S 3 7の後、図 8に示すように、起動ボリューム 1 1 Kが用意されている物理ディスク群 9に対応のディスク制御部 1 9が、起動制御テーブル 3 1 Kにキャッシュポイント等が新たに登録されたことを検出したら、その起動制御テーブル 3 1 Kに登録されているアクセスアドレス(つまり起動ボリューム 1 1 Kのアクセスアドレス)からデータをリードし、そのリードデータを、起動制御テーブル 3 1 Kに登録されているキャッシュポイントが示すキャッシュポイント 2 1に格納する(S 4 4)。そして、そのディスク制御部 1 9は、起動ボリューム 1 1 Kのキャッシュスロット 2 1にリードデータを格納したことを意味するデータ準備完了をチャンネル制御部 1 3に通知する(S 4 5)。なお、どのチャンネル制御部 1 3に通知するかは、例えば、起動制御テーブル 3 1 Kにキャッシュポイント等を登録するチャンネル制御部 1 3が自分の識別コードも一緒に登録し、ディスク制御部 1 9は、その識別コードから、上記データ準備完了の通知先とするチャンネル制御部 1 3を識別することができる。

## 【 0 1 1 6 】

チャンネル制御部 1 3は、ディスク制御部 1 9からデータ準備完了の通知を受けたときは、図 7の S 3 7で切り離しを要求したホスト 1に対し再接続を要求する(S 4 6)。

## 【 0 1 1 7 】

ホスト OS 7は、チャンネル制御部 1 3から再接続の要求を受けたときは、ディスクアレイ装置 3に対して再接続を行なって(例えば、切断した所定の通信チャンネルを形成し)、先に送信した第 2のサブコマンド 1 9 Bを再送信する。そのため、再び、上述した図 7の S 3 1 ~ S 3 4の処理が行われる。

## 【 0 1 1 8 】

このとき S 3 4では、S 3 6で起動制御テーブル 3 1 Kにキャッシュポイントが登録されたので、キャッシュヒットする(S 3 5でY)。その場合、チャンネル制御部 1 3は、キャッシュヒットしたキャッシュスロット 2 1(すなわち、起動制御テーブル 3 1 Kに登録されたキャッシュポイントが示すキャッシュスロット 2 1)から、上記格納されたリードデータを読み出してホスト 1に送信し、且つ、ホスト 1に、正常にリード処理が行われた旨を示す正常ステータスを通知して、起動制御テーブル 3 1 K上のレコード(例えば、キャッシュポイント、アクセスアドレス、アクセス種別等)を消去する(S 3 8)。

## 【 0 1 1 9 】

以上が、通常リードの場合の流れである。

## 【 0 1 2 0 】

( 2 ) リードコントロールの場合(すなわち、起動ボリューム 1 1 Kとは別のボリューム 1 1 Tからデータをリードする場合)。

## 【 0 1 2 1 】

この場合、まず、図 7に示すように、通常リードの場合と同様に S 3 1 ~ S 3 3が行われる。そして、この場合では、S 3 3の結果として、起動制御テーブル 3 1 Kにリードコントロールが登録されていることが検出される。そのため、チャンネル制御部 1 3は、起動制御テーブル 3 1 Kに登録されているターゲットボリュームの ID から、ターゲットボリューム 1 1 T及びその制御テーブル(つまりターゲット制御テーブル) 3 1 Tを認識し、その認識された制御テーブル 3 1 Tを参照して、キャッシュヒットしたか否かを確認する(S 3 9)。

10

20

30

40

50

## 【 0 1 2 2 】

S 3 9の結果、キャッシュヒットしなければ(S 4 0でN)、チャンネル制御部 1 3は、キャッシュメモリ 1 5上にターゲットボリューム 1 1 Tに専用のキャッシュスロット 2 1を確保し、そのキャッシュスロット 2 1のキャッシュポインタをターゲット制御テーブル 3 1 Tに登録する(S 4 1)。その後、チャンネル制御部 1 3は、ホスト 1に切り離し要求をする(S 4 2)。それにより、ホスト 1によって、ホスト 1とディスクアレイ装置 3との間で切り離しが行なわれる。

## 【 0 1 2 3 】

S 4 2の後、図 8に示すように、ターゲットボリューム 1 1 Tが用意されている物理ディスク群 9に対応のディスク制御部 1 9が、ターゲット制御テーブル 3 1 Tにキャッシュポインタ等が新たに登録されたことを検出したら、そのターゲット制御テーブル 3 1 Tに登録されているアクセスアドレス(つまりターゲットボリューム 1 1 Tのアクセスアドレス)からデータをリードし、そのリードデータを、ターゲット制御テーブル 3 1 Tに登録されているキャッシュポインタが示すキャッシュポインタ 2 1に格納する(S 4 4)。そして、そのディスク制御部 1 9は、ターゲットボリューム 1 1 Tのキャッシュスロット 2 1にリードデータを格納したことを意味するデータ準備完了をチャンネル制御部 1 3に通知する(S 4 5)。

10

## 【 0 1 2 4 】

チャンネル制御部 1 3は、ディスク制御部 1 9からデータ準備完了の通知を受けたときは、図 7の S 3 7で切り離しを要求したホスト 1に対し再接続を要求する(S 4 6)。

20

## 【 0 1 2 5 】

ホスト O S 7は、チャンネル制御部 1 3から再接続の要求を受けたときは、ディスクアレイ装置 3に対して再接続を行なって、先に送信した第 2のサブコマンド 1 9 Bを再送信する。そのため、再び、上述した S 3 1 ~ S 3 3及び S 3 9の処理が行われる。

## 【 0 1 2 6 】

このとき S 3 9では、S 4 1でターゲット制御テーブル 3 1 Tにキャッシュポインタが登録されたので、キャッシュヒットする(S 4 0でY)。その場合、チャンネル制御部 1 3は、キャッシュヒットしたキャッシュスロット 2 1(すなわち、ターゲット制御テーブル 3 1 Tに登録されたキャッシュポインタが示すキャッシュスロット 2 1)から、上記格納されたリードデータを読み出してホスト 1に送信し、且つ、ホスト 1に、正常にリード処理が行われた旨を示す正常ステータスを通知して、起動制御テーブル 3 1 K及びターゲット制御テーブル 3 1 T上のレコード(例えば、キャッシュポインタ、アクセスアドレス、アクセス種別等)を消去する(S 4 3)。

30

## 【 0 1 2 7 】

以上が、リードコントロールの場合の流れである。

## 【 0 1 2 8 】

( 3 )ライトコントロールの場合(すなわち、起動ボリューム 1 1 Kとは別のボリューム 1 1 Tにデータをライトする場合)。

## 【 0 1 2 9 】

この場合、まず、図 7に示すように、通常リードの場合と同様に S 3 1 ~ S 3 2が行われる。そして、この場合では、S 3 2の結果として、第 2のサブコマンド 1 9 Bが持つ第 2アクセス種別コードはライトを示すことが検出される。

40

## 【 0 1 3 0 】

その後、図 9に示すように、チャンネル制御部 1 3は、第 2のサブコマンド 1 9 Bの対象となっている起動ボリューム 1 1 Kの制御テーブル 3 1 Kを参照して、アクセス種別としてライトとライトコントロールのどちらが登録されているかを調べる(S 4 7)。

## 【 0 1 3 1 】

S 4 7の結果、ライトコントロールが登録されていれば、チャンネル制御部 1 3は、起動制御テーブル 3 1 Kに登録されているターゲットボリュームの I Dから、ターゲットボリューム 1 1 T及びその制御テーブル 3 1 Tを認識し、その認識された制御テーブル 3 1 T

50

を参照して、キャッシュヒットしたか否かを確認する（S 4 9）。

【0 1 3 2】

S 4 9の結果、キャッシュヒットしなければ（S 5 0でN）、チャンネル制御部 1 3は、キャッシュメモリ 1 5上にターゲットボリューム 1 1 Tに専用のキャッシュスロット 2 1を確保し、そのキャッシュスロット 2 1のキャッシュポインタをターゲット制御テーブル 3 1 Tに登録する（S 5 1）。その後、チャンネル制御部 1 3は、ホスト 1に切り離し要求をする（S 5 2）。

【0 1 3 3】

S 5 2の後、図 1 0に示すように、ターゲットボリューム 1 1 Tが用意されている物理ディスク群 9に対応のディスク制御部 1 9が、ターゲット制御テーブル 3 1 Tにキャッシュポインタ等が新たに登録されたことを検出したら、そのターゲット制御テーブル 3 1 Tに登録されているアクセスアドレス（つまりターゲットボリューム 1 1 Tのアクセスアドレス）からデータをリードし、そのリードデータを、ターゲット制御テーブル 3 1 Tに登録されているキャッシュポインタが示すキャッシュポインタ 2 1に格納する（S 5 4）。そして、そのディスク制御部 1 9は、データ準備完了をチャンネル制御部 1 3に通知する（S 5 5）。

10

【0 1 3 4】

チャンネル制御部 1 3は、ディスク制御部 1 9からデータ準備完了の通知を受けたときは、図 9の S 5 2で切り離しを要求したホスト 1に対し再接続を要求する（S 5 6）。

【0 1 3 5】

ホスト OS 7は、チャンネル制御部 1 3から再接続の要求を受けたときは、ディスクアレイ装置 3に対して再接続を行なって、先に送信した第 2のサブコマンド 1 9 Bを再送信する。そのため、再び、上述した S 3 1 ~ S 3 2、S 4 7及び S 4 9の処理が行われる。

20

【0 1 3 6】

このときの S 4 9では、S 5 1でターゲット制御テーブル 3 1 Tにキャッシュポインタが登録されたので、キャッシュヒットする（S 5 0でY）。その場合、チャンネル制御部 1 3は、キャッシュヒットしたキャッシュスロット 2 1（すなわち、ターゲット制御テーブル 3 1 Tに登録されたキャッシュポインタが示すキャッシュスロット 2 1）に、上記受信した第 2のサブコマンド 1 9 Bに含まれているライトデータを格納する（S 5 3）。

【0 1 3 7】

S 5 3の後、図 1 0に示すように、ディスク制御部 1 9は、ターゲット制御テーブル 3 1 Tに登録されているキャッシュポインタが示すキャッシュポインタ 2 1にライトデータがあることを検出してそのライトデータを読み出し、それを、ターゲット制御テーブル 3 1 Tに登録されているアクセスアドレス（つまりターゲットボリューム 1 1 Tのアクセスアドレス）にライトする。その後、そのディスク制御部 1 9は、ライトデータの格納が完了したことをチャンネル制御部 1 3に通知する（S 5 8）。

30

【0 1 3 8】

ライトデータの格納が完了した通知を受けたチャンネル制御部 1 3は、正常にライト処理が行われた旨を示す正常ステータスをホスト 1に通知すると共に、起動制御テーブル 3 1 K及びターゲット制御テーブル 3 1 T上のレコード（例えば、キャッシュポインタ、アクセスアドレス、アクセス種別等）を消去する（S 5 9）。

40

【0 1 3 9】

以上が、ライトコントロールの場合の動作流れである。

【0 1 4 0】

（4）通常ライトの場合（起動ボリューム 1 1 Kにデータをライトする場合）。

【0 1 4 1】

この場合、まず、ライトコントロールの場合と同様に S 3 1 ~ S 3 2及び S 4 7が行われる。そして、この場合では、S 4 7の結果として、起動制御テーブル 3 1 Kにライトが登録されていることが検出されるので、通常のライト処理が行われる（S 4 8）。通常のライト処理の流れについては特別図示しないが、通常リード、リードコントロール、及びラ

50

イトコントロールの場合の動作流れについての上記説明から、当業者であれば、通常のライト処理は下記のような流れで行なわれることを理解することができる。

【0142】

すなわち、チャンネル制御部13は、起動制御テーブル31Kを参照して、キャッシュヒットしたか否かを確認する。その結果、キャッシュヒットしなければ、チャンネル制御部13は、キャッシュメモリ15上に起動ボリューム11Kに専用のキャッシュスロット21を確保し、そのキャッシュスロット21のキャッシュポインタを起動制御テーブル31Kに登録する。その後、チャンネル制御部13は、ホスト1に切り離し要求をする。

【0143】

起動ボリューム11Kが用意されている物理ディスク群9に対応のディスク制御部19が、起動制御テーブル31Kにキャッシュポインタ等が新たに登録されたことを検出したら、その制御テーブル31Kに登録されているアクセスアドレスからデータをリードし、そのリードデータを、起動制御テーブル31Kに登録されているキャッシュポインタが示すキャッシュポインタ21に格納する。そして、ディスク制御部19は、データ準備完了をチャンネル制御部13に通知する。

【0144】

チャンネル制御部13は、ディスク制御部19からデータ準備完了の通知を受けたときは、ホスト1に対し再接続を要求する。

【0145】

ホストOS7は、チャンネル制御部13から再接続の要求を受けたときは、ディスクアレイ装置3に対して再接続を行なって、先に送信した第2のサブコマンド19Bを再送信する。そのため、再び、キャッシュヒットしたか否かが確認される。

【0146】

この場合の確認では、前述の通り、起動制御テーブル31Kにキャッシュポインタが登録されたので、キャッシュヒットする。その場合、チャンネル制御部13は、キャッシュヒットしたキャッシュスロット21に、上記受信した第2のサブコマンド19Bに含まれているライトデータを格納する。

【0147】

その後、ディスク制御部19は、起動制御テーブル31Kに登録されているキャッシュポインタが示すキャッシュポインタ21にライトデータがあることを検出してそのライトデータを読み出し、それを、起動制御テーブル31Kに登録されているアクセスアドレス（つまり起動ボリューム11Kのアクセスアドレス）にライトする。その後、そのディスク制御部19は、ライトデータの格納が完了したことをチャンネル制御部13に通知する。

【0148】

ライトデータの格納が完了した通知を受けたチャンネル制御部13は、正常にライト処理が行われた旨を示す正常ステータスをホスト1に通知すると共に、起動制御テーブル31K上のレコードを消去する。

【0149】

以上が、通常ライトの場合の動作流れである。

【0150】

上述した実施形態によれば、ホスト1が、アクセスコマンド19を起動ボリューム（つまりそのコマンド19の対象となる論理ボリューム）11Kに送信するようになっており、その場合、アクセスしたいターゲットの論理ボリューム11に関するターゲット情報をアクセスコマンド19に含めて送信するようになっている。ディスクアレイ装置3は、アクセスコマンド19を受けると、そのアクセスコマンド19に含まれているターゲット情報が示すターゲットボリュームが起動ボリューム11Kとは別のボリューム11Tであっても、その別のボリューム11Tからデータをリードしてホスト1に提供したり、或いは、ホスト1が書き込みたいデータをその別のボリューム11Tにライトしたりする。このように、この実施形態によれば、オンライン状態になっている起動ボリュームに対してデータアクセスを行うのと同様の方法でそれとは別のターゲットボリュームに対してデータア

10

20

30

40

50

クセスを行なうことができる。それ故、ユーザにとって便利である。なお、このことは、ターゲットボリュームがオフラインボリュームである場合には特に利便性が高い。

【0151】

ところで、上述した説明では、ホストAP5がどのようにしてターゲットボリュームをホストOS7に通知するかについては具体的に説明していないが、これについては幾つかの方法が考えられる。一例として、ユーザが、予め、ディスクアレイ装置3に対して直接的に又は遠隔から所定の操作を行なって、ディスクアレイ装置3が持つ全ての論理ボリュームと各論理ボリュームにはどのようなデータが格納されているかということ（以下、これを「ボリューム情報」と言う）をディスクアレイ装置3から通知してもらい、その通知されたボリューム情報に基づいて、ホストAP5に対し所望のターゲットボリュームを入力することで、ホストAP5がホストOS7にターゲットボリュームを通知する方法が考えられる。別の一例として、ディスクアレイ装置3が、ユーザからの要求に回答して或いは任意のタイミングで自発的に、上記ボリューム情報をホスト1に送信し、ホスト1がそのボリューム情報を所定の記憶領域に格納しておき、ユーザがホストAP5を操作してターゲットボリュームを入力するときは、ホストAP5が、そのボリューム情報を読み出してディスプレイ画面に表示し、ユーザに選択された論理ボリュームをターゲットボリュームとしてホストOS7に通知する方法も考えられる。

10

【0152】

また、起動ボリューム11Kは、例えば、ホスト1に対してオンライン状態である複数のオンラインボリュームの中から、ホスト1又はディスクアレイ装置3によって自動的に又はユーザによって任意に選択されたものである。

20

【0153】

さて、上述した実施形態には、幾つか応用例、例えば下記のような第1～第3の応用例が考えられる。

【0154】

図11は、本発明の一実施形態に係るアクセスシステムの第1の応用例を示す。

【0155】

第1の応用例は、上述した実施形態を、正ボリューム（プライマリのボリューム）と副ボリューム（セカンダリのボリューム）とをペアで管理して二重化処理を行うことができるようになったディスクアレイ装置A1を有するデータアクセスシステム200に適用したものである。

30

【0156】

この第1の応用例では、起動ボリューム11Kは、正ボリュームであり、起動ボリューム11Kとは別のターゲットボリューム（例えばオフラインボリューム）11Tは、副ボリュームである。

【0157】

この第1の応用例によれば、ホスト1に対して正ボリューム11Kのみがオンライン状態になっていて副ボリューム11Tはオフライン状態でも、ホスト1は、正ボリューム11Kに対してデータアクセスする方法と同様の方法で副ボリューム11Tにデータアクセスすることができる。具体的には、図示のように、ホスト1は、正ボリューム11Kを対象とするアクセスコマンド19を発行しても、そのアクセスコマンド19に、リードコントロールを示す第1アクセス種別コードと、副ボリューム11TのID及びアクセスアドレスを含んだターゲット情報とを含めれば、副ボリューム11Tから所望のデータを取得することができる。また、ホスト1は、正ボリューム11Kを対象とするアクセスコマンド19を発行しても、そのアクセスコマンド19に、ライトコントロールを示す第1アクセス種別コードと、副ボリューム11TのID及びアクセスアドレスを含んだターゲット情報とを含めれば、副ボリューム11Tに所望のデータをライトすることができる。

40

【0158】

なお、この第1の応用例において、例えば、正ボリューム11K内の第1のデータと副ボリューム11T内の第2のデータ（第1のデータとペアになったもの）が互いに一致しな

50

い場合、ディスクアレイ装置 3 は、副ボリューム 1 1 T から第 2 のデータをリードし、その第 2 のデータを正ボリューム 1 1 K 内の第 1 のデータに上書きすることにより、第 1 のデータの回復を行なうことができる。この処理は、例えば、ディスクアレイ装置 3 が任意のタイミングで上記第 1 のデータと第 2 のデータの比較を行ない一致しないときに自動的に行なっても良いし、或いは、上記一致しないことをホスト 1 に通知し、ホスト 1 が、ユーザからの要求に応じて又は自発的に、副ボリューム 1 1 T の上記第 2 のデータをターゲットとしたアクセスコマンド 1 9 を、正ボリューム 1 1 K を起動ボリュームとして送信し、ディスクアレイ装置 3 が、ホスト 1 からのそのコマンド 1 9 に基づいて、上述の上書き処理をおこなっても良い。

【 0 1 5 9 】

10

図 1 2 は、本発明の一実施形態に係るアクセスシステムの第 2 の応用例を示す。

【 0 1 6 0 】

第 2 の応用例は、上述した実施形態を、複数のホスト 1 A、1 B と、それら複数のホスト 1 A、1 B に通信可能に接続されるディスクアレイ装置 A 2 とを有するデータアクセスシステム 3 0 0 に適用したものである。

【 0 1 6 1 】

この第 2 の応用例では、ホスト 1 A にとっての起動ボリューム 1 1 K は、ホスト 1 A に対してオンライン状態になっている 1 又は複数のオンラインボリューム 1 1 K A の中から自動で又は手動で選択された一つの論理ボリュームである。一方、ホスト 1 A にとってのターゲットボリューム（例えばオフラインボリューム）1 1 T は、ホスト 1 A に対してはオンライン状態になっていないが別のホスト 1 B に対してオンライン状態になっている 1 又は複数のオンラインボリューム 1 1 K B の中からホスト 1 A で指定された論理ボリュームである。

20

【 0 1 6 2 】

この第 2 の応用例によれば、ホスト 1 A は、自機に対してオンライン状態になっている起動ボリューム 1 1 K に対してデータアクセスする方法と同様の方法で、他のホスト 1 B に対してオンライン状態になっている論理ボリューム 1 1 K B にデータアクセスすることができる。別の言い方をすれば、各ホスト 1 A、1 B 毎に論理ボリュームが割当てられている場合、各ホスト 1 A、1 B は、自機に割当てられているオンラインボリュームに対してデータアクセスする方法と同様の方法で、他のホストに対して割当てられている論理ボリュームにデータアクセスすることができる。このため、ホスト 1 A、1 B はホスト種類（例えば、ホストのメーカー、ホストの OS の種類）が異なっていて、ホスト 1 A、1 B がそれぞれ認識している論理ボリューム 1 1 K A、1 1 K B のデータ形式が異なっていたとしても、各ホスト 1 A、1 B は、自機に割当てられているオンラインボリュームに対してデータアクセスする方法と同様の方法で、他のホストに割当てられているオンラインボリューム（オフラインボリュームでも良い）にデータアクセスすることができる。

30

【 0 1 6 3 】

図 1 3 は、本発明の一実施形態に係るアクセスシステムの第 3 の応用例を示す。

【 0 1 6 4 】

第 3 の応用例は、上記の第 2 の応用例を変形例であり、複数のホスト 1 A、1 B、1 C と、それら複数のホスト 1 A、1 B、1 C に通信可能に接続されるディスクアレイ装置 A 3 とを有するデータアクセスシステム 4 0 0 に適用したものである。

40

【 0 1 6 5 】

この第 3 の応用例において、ディスクアレイ装置 A 3 のキャッシュメモリ 1 5 の記憶領域は、各ホスト毎に仕切られている。具体的に言えば、キャッシュメモリ 1 5 には、第 1 のホスト 1 A に対して割当てられた第 1 のキャッシュ領域 1 5 A と、第 2 のホスト 1 B に対して割当てられた第 2 のキャッシュ領域 1 5 B と、第 3 のホスト 1 C に対して割当てられた第 3 のキャッシュ領域 1 5 C とがある。

【 0 1 6 6 】

この第 3 の応用例では、ホスト 1 A にとっての起動ボリューム 1 1 K は、第 2 の応用例と

50

同様に、ホスト 1 A に対してオンライン状態になっている 1 又は複数のオンラインボリューム 1 1 K A の中から自動で又は手動で選択された一つの論理ボリュームである。一方、ホスト 1 A にとってのターゲットボリューム（例えばオフラインボリューム）1 1 T は、ホスト 1 A に対してはオンライン状態になっていないが別のホスト 1 C に対してオンライン状態になっている 1 又は複数のオンラインボリューム 1 1 K C の中からホスト 1 A で指定された論理ボリュームである。

【0167】

この第 3 の応用例では、ホスト 1 A からアクセスコマンド 1 9 を受けたチャネル制御部 1 3 は、そのコマンド 1 9 に含まれているターゲット情報が、第 3 のホスト 1 C の或る一つのオンラインボリューム 1 1 K C を指しているため、その第 3 のホスト 1 C のキャッシュ領域 1 5 C 内にキャッシュスロット 2 1 を確保し、且つ、そのオンラインボリューム 1 1 K C に対応した制御テーブルに、そのキャッシュスロット 2 1 のキャッシュポインタを記入する。

10

【0168】

この第 3 の応用例によれば、各ホスト毎にキャッシュ領域が割当てられている場合、ホスト 1 A が、他のホスト 1 C のオンラインボリューム 1 1 K C に対してデータアクセスするときは、ホスト 1 A と通信するチャネル制御部 1 3 は、ホスト 1 A に対応したキャッシュ領域 1 5 A ではなく、ホスト 1 A のターゲットボリューム 1 1 T に対してオンライン状態になっている他のホスト 1 C のキャッシュ領域 1 5 C にキャッシュスロット 2 1 を確保し、そのキャッシュスロット 2 1 を介して、ホスト 1 A と、他のホスト 1 C に対するオンラインボリュームとの間でデータがやり取りされる。すなわち、各ホスト毎にキャッシュ領域が割当てられている場合であっても、各ホスト 1 A、1 B、1 C は、自機に割当てられているオンラインボリュームに対してデータアクセスする方法と同様の方法で、他のホストに対して割当てられているオンラインボリューム（オフラインボリュームでも良い）にデータアクセスすることができる。

20

【0169】

以上、本発明の好適な実施形態及びその応用例を説明したが、これらは本発明の説明のための例示であって、本発明の範囲をこの実施形態及び応用例にのみ限定する趣旨ではない。本発明は、他の種々の形態でも実施することが可能である。例えば、ディスクアレイ装置 3 は、複数台存在してもよく、その場合、ホスト 1 の起動ボリューム 1 1 K は、ホスト 1 に接続されている第 1 のディスクアレイ装置内にあり、ホスト 1 のターゲットボリューム 1 1 T は、そのディスクアレイ装置とは別の第 2 のディスクアレイ装置内に存在しても良い。この場合、ホスト 1 は、自機が接続されているディスクアレイ装置内の所望の起動ボリュームに対してデータアクセスすると同様の方法で、他のディスクアレイ装置内の論理ボリュームにデータアクセスすることができる。

30

【図面の簡単な説明】

【図 1】本発明の一実施形態に係るアクセスシステムの全体構成を示す。

【図 2】ホスト 1 が発行するアクセスコマンドを示す。

【図 3】制御メモリ 1 7 に用意される制御テーブルを示す。

【図 4】本実施形態に係るデータアクセスシステムにおいて、ホスト 1 が起動ボリュームとは別の論理ボリュームをターゲットとしそこからデータをリードする場合の流れを示す。

40

【図 5】ホスト OS 7 の動作流れを示す。

【図 6】ホスト 1 から第 1 のサブコマンド 1 9 A を受信したときのディスクアレイ装置 3 の動作流れを示す。

【図 7】ホスト 1 から第 2 のサブコマンド 1 9 B を受信したときのディスクアレイ装置 3 の動作流れを示す。

【図 8】ホスト 1 から第 2 のサブコマンド 1 9 B を受信したときのディスクアレイ装置 3 の動作流れを示す。

【図 9】ホスト 1 から第 2 のサブコマンド 1 9 B を受信したときのディスクアレイ装置 3

50

の動作流れを示す。

【図10】ホスト1から第2のサブコマンド19Bを受信したときのディスクアレイ装置3の動作流れを示す。

【図11】本発明の一実施形態に係るアクセスシステムの第1の応用例を示す。

【図12】本発明の一実施形態に係るアクセスシステムの第2の応用例を示す。

【図13】本発明の一実施形態に係るアクセスシステムの第3の応用例を示す。

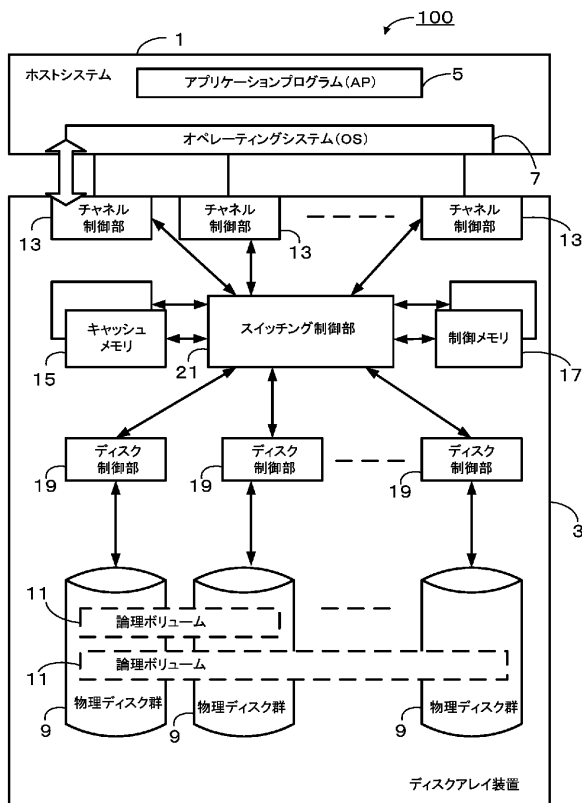
【符号の説明】

- 1 ホストシステム
- 3 ディスクアレイ装置
- 5 アプリケーションプログラム(ホストAP)
- 7 オペレーティングシステム(ホストOS)
- 9 物理ディスク群
- 11 論理ボリューム
- 11K 起動ボリューム
- 11T ターゲットボリューム
- 13 チャンネル制御部
- 15 キャッシュメモリ
- 17 制御メモリ
- 19 ディスク制御部
- 21 スイッチング制御部
- 31 制御テーブル
- 31K 起動ボリュームの制御テーブル
- 31T ターゲットボリュームの制御テーブル

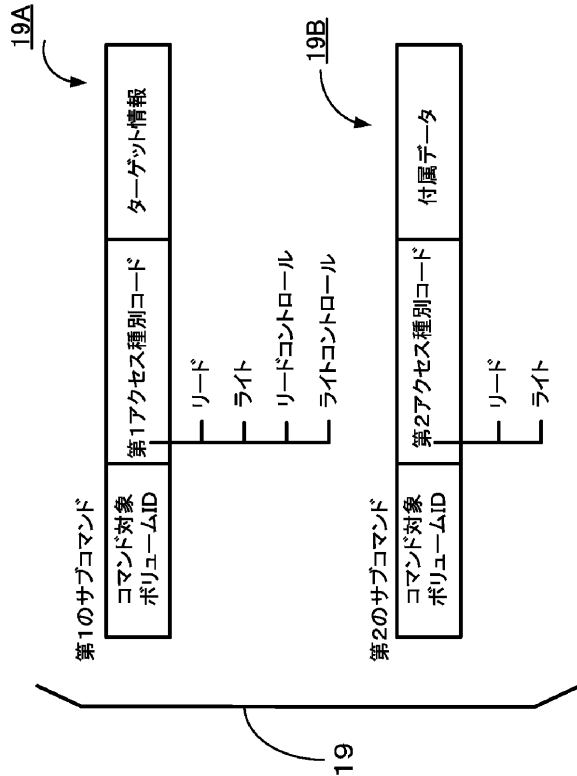
10

20

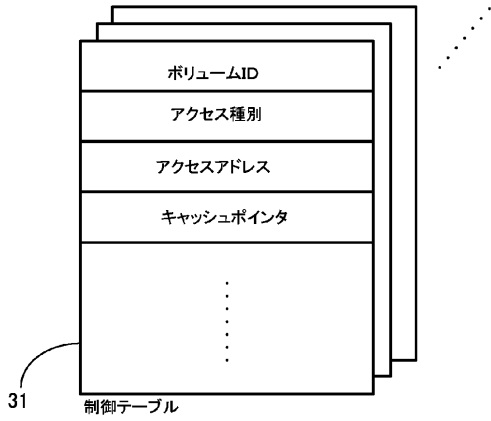
【図1】



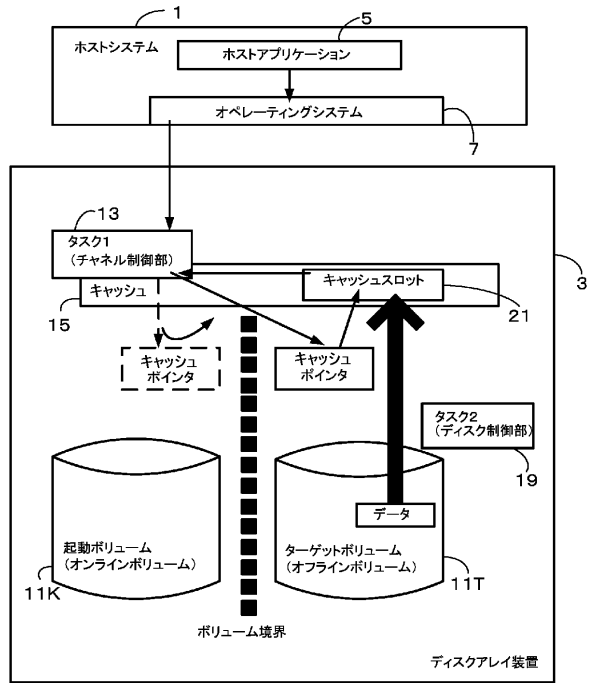
【図2】



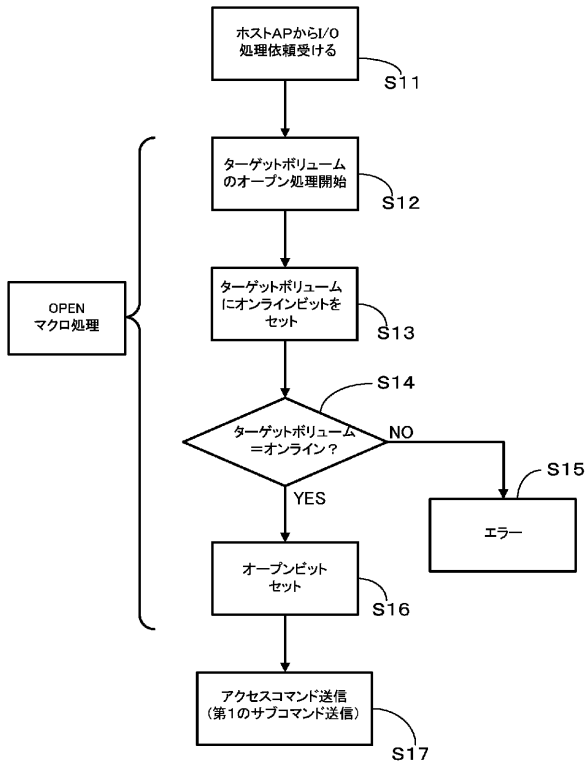
【図3】



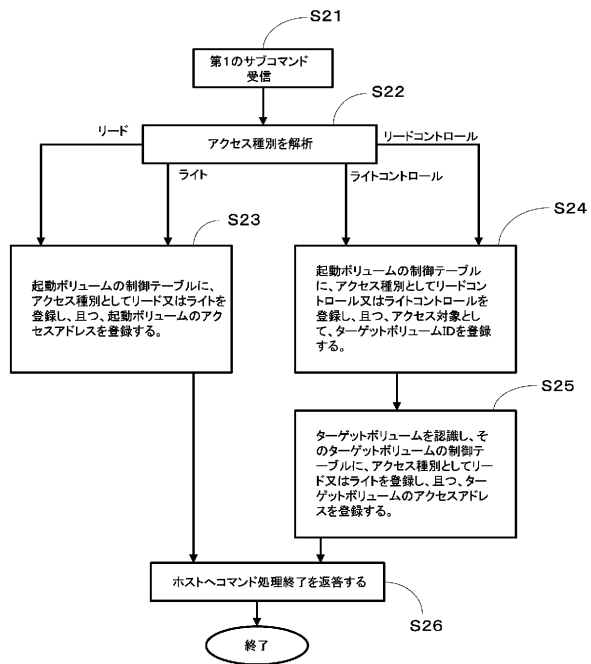
【図4】



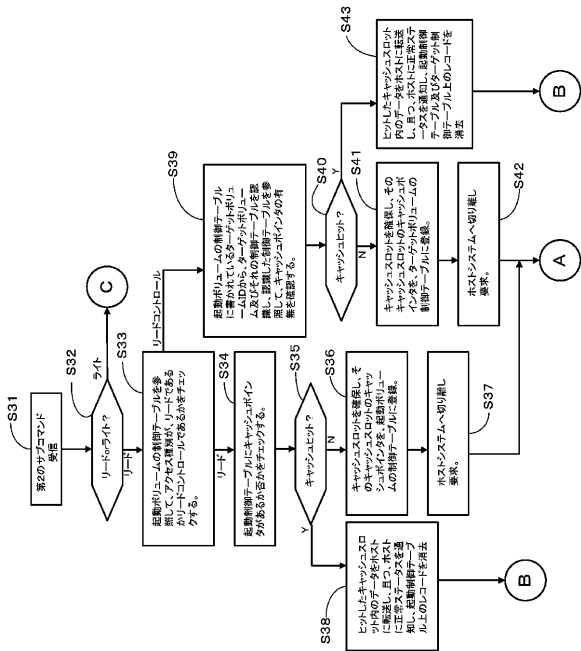
【図5】



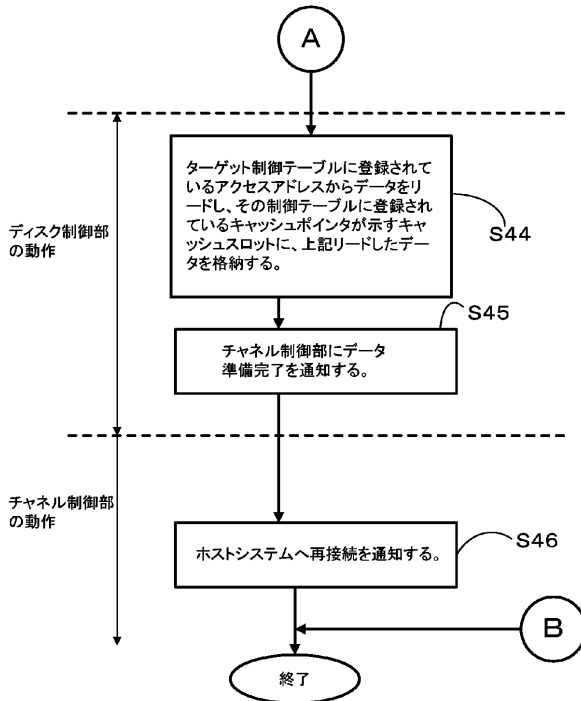
【図6】



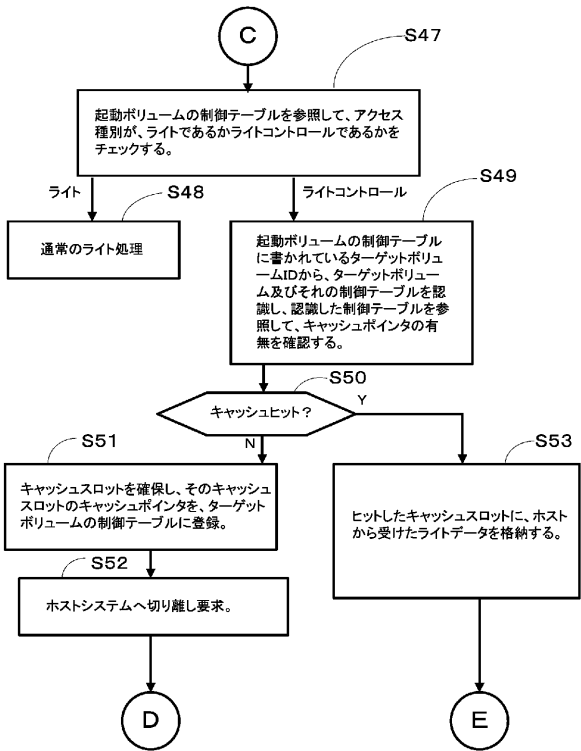
【図7】



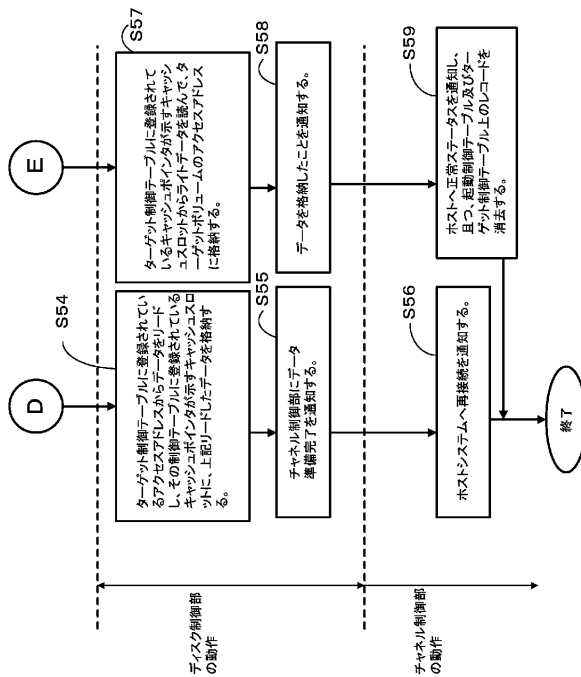
【図8】



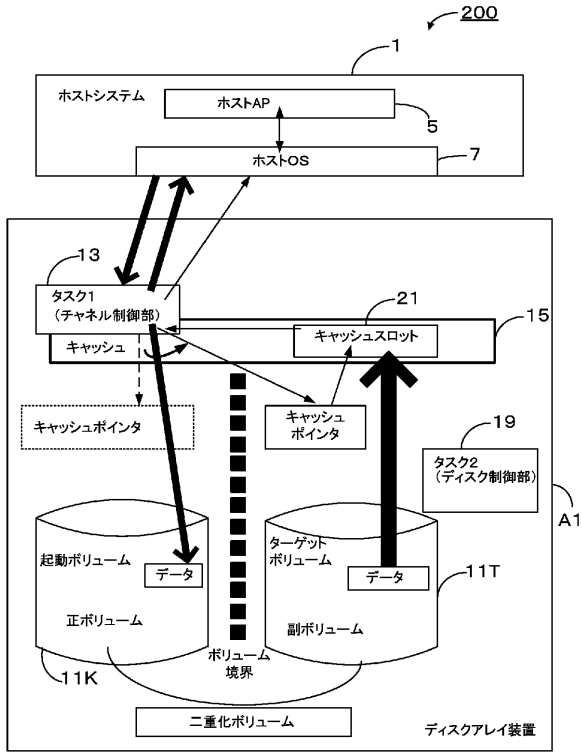
【図9】



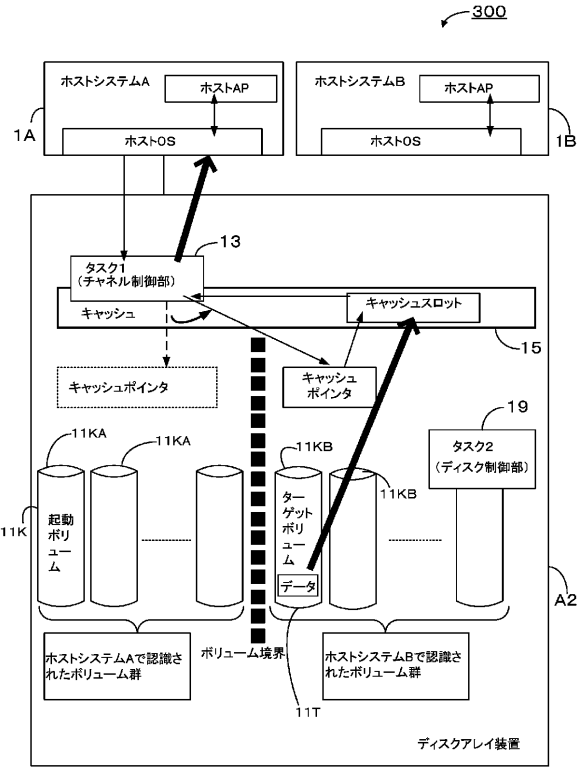
【図10】



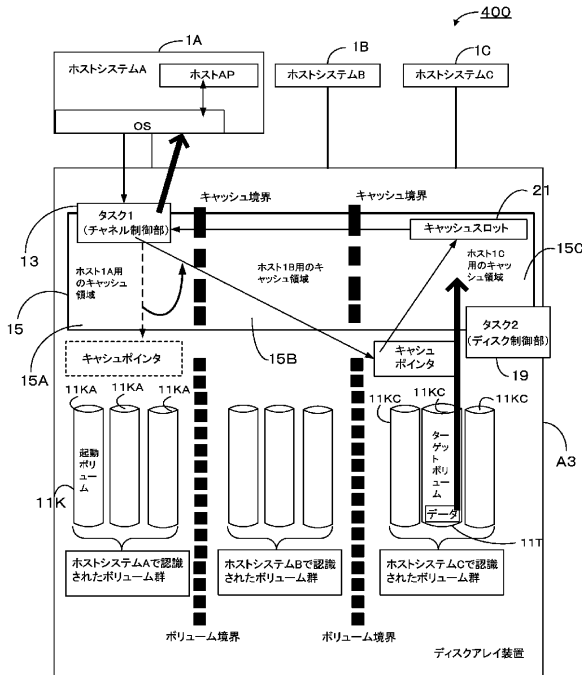
【図11】



【図12】



【図13】



---

フロントページの続き

(51)Int.Cl. F I  
G 0 6 F 12/08 5 5 1 Z  
G 0 6 F 12/08 5 5 7

(72)発明者 森 昭洋  
神奈川県小田原市中里3 2 2番地2号 株式会社日立製作所 R A I Dシステム事業部内  
(72)発明者 宮本 清久  
神奈川県小田原市中里3 2 2番地2号 株式会社日立製作所 R A I Dシステム事業部内  
(72)発明者 木村 雅史  
神奈川県小田原市中里3 2 2番地2号 株式会社日立製作所 R A I Dシステム事業部内

審査官 木村 雅也

(56)参考文献 特開2 0 0 3 - 1 5 7 1 5 2 ( J P , A )  
特開2 0 0 0 - 3 4 7 8 1 1 ( J P , A )  
特表平 1 1 - 5 0 9 6 5 9 ( J P , A )

(58)調査した分野(Int.Cl. , D B名)  
G06F 3/06  
G06F 12/08