

(19) 日本国特許庁 (JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2008-140413  
(P2008-140413A)

(43) 公開日 平成20年6月19日(2008.6.19)

(51) Int.Cl.

F

### テーマコード（参考）

G 06 F 3/06 (2006.01)

G06F 3/06 301Z

5 B O 1 4

G06F 13/10 (2006.01)

G06F 13/10 340A

5B065

**G06F** 12/00 (2006, 01)

G06F 12/00 501A

5B082

審査請求 有 請求項の数 10 O.L. (全 18 頁)

(21) 出願番号	特願2008-2274 (P2008-2274)	(71) 出願人	000005108 株式会社日立製作所 東京都千代田区丸の内一丁目6番6号
(22) 出願日	平成20年1月9日 (2008.1.9)		
(62) 分割の表示	特願2002-125172 (P2002-125172) の分割	(74) 代理人	100079108 弁理士 稲葉 良幸
原出願日	平成14年4月26日 (2002.4.26)	(74) 代理人	100093861 弁理士 大賀 真司
		(72) 発明者	松並 直人 神奈川県川崎市麻生区王禅寺1099番地 株式会社日立製作所システム開発研究所 内
		(72) 発明者	北村 学 神奈川県川崎市麻生区王禅寺1099番地 株式会社日立製作所システム開発研究所 内

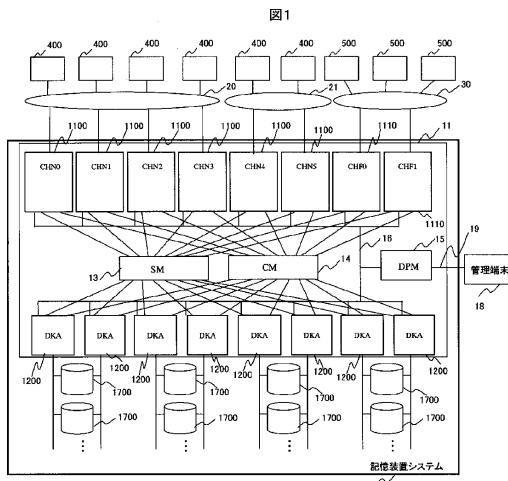
(54) 【発明の名称】 記憶装置システム

(57) 【要約】

【課題】N A Sの機能とS A Nの機能の両方を内蔵し、システム構成に自由度を有する記憶装置により管理運用コストを低減する。

【解決手段】記憶装置に、複数のインターフェース制御部を装填可能な複数のインターフェーススロットと、該スロットに装填可能なS A Nの機能を備えるブロックI/Oインターフェース制御部と、同N A Sの機能を備えるファイルI/Oインターフェース制御部と、任意のインターフェース制御部からアクセス可能な複数のディスク装置で構成された記憶容量プールと、記憶容量プールを制御する記憶容量プール制御部とを備える。

【選択図】 図 1



**【特許請求の範囲】****【請求項 1】**

種類の異なるインターフェース装置が装填される複数のスロットと、複数のディスク装置と、前記複数のディスク装置が有する記憶領域を管理する管理部とを有し、前記管理部は、前記複数のディスク装置が有する記憶領域と前記スロットに装填される前記種類の異なるインターフェース制御装置との対応関係を示す情報を保持し、前記種類の異なるインターフェース装置は、前記対応関係を示す情報に基づいて、前記記憶領域を使用することを特徴とする記憶装置システム。

**【請求項 2】**

前記種類の異なるインターフェース装置には、ファイルI/Oインターフェースに対応するチャネルアダプタボードと、ロックI/Oインターフェースに対応するチャネルアダプタボードとが含まれることを特徴とする請求項1記載の記憶装置システム。 10

**【請求項 3】**

前記種類の異なるインターフェース制御装置は、すべて同一の形状をしていることを特徴とする請求項2記載の記憶装置システム。

**【請求項 4】**

前記複数のディスク装置が有する記憶領域は、該ディスク装置毎の特性に基づいて、複数のグループに分割されていることを特徴とする請求項3記載の記憶装置システム。

**【請求項 5】**

前記管理部は、前記インターフェース制御装置に対応付けられている記憶領域を他の前記インターフェース制御装置にも対応付けたことを示す情報を保持し、前記他のインターフェース制御装置は、前記情報に基づいて前記記憶領域にアクセスすることを特徴とする請求項1記載の記憶装置システム。 20

**【請求項 6】**

前記複数のディスク装置が有する記憶領域には特定の属性が割り振られ、前記インターフェース制御装置を介して受信されるアクセス要求を前記属性に応じて処理することを特徴とする請求項1記載の記憶装置システム。

**【請求項 7】**

前記特定の属性とは、前記記憶領域に対応する前記インターフェース制御装置が、該記憶領域にデータを書き込めるか否かを示す属性であることを特徴とする請求項6記載の記憶装置システム。 30

**【請求項 8】**

前記特定の属性とは、前記記憶領域に対応する前記インターフェース制御装置が複数存在するか否かを示す属性であることを特徴とする請求項7記載の記憶装置システム。

**【請求項 9】**

前記インターフェース制御装置が受信したファイルアクセス要求に対して、前記インターフェース制御装置とは異なる他のインターフェース制御装置を介して前記ファイルアクセス要求に対応するデータを送信することを特徴とする請求項1記載の記憶装置システム。 40

**【請求項 10】**

共有メモリを有し、前記共有メモリには、前記複数のインターフェース制御装置と前記記憶領域の対応関係を示す情報が格納されることを特徴とする請求項1から6記載のうち、いずれか一つの記憶装置システム。

**【発明の詳細な説明】****【技術分野】****【0001】**

本発明は、計算機システムで用いられる記憶装置システムに関する。

**【背景技術】****【0002】**

記憶装置システムと計算機とを接続するためのインターフェースは、大きく分けて二つ

のタイプが存在する。第一は、記憶装置システムにおけるデータの管理単位であるブロックを単位として、記憶装置システムと計算機との間の入出力（以下「I/O」）が行われる「ブロックI/Oインターフェース」である。ブロックI/Oインターフェースには、ファイバチャネル（Fibre Channel）、SCSI（Small Computer Systems Interface）などの種類がある。

#### 【0003】

ブロックI/Oインターフェースを有する複数の記憶装置システム及び複数の計算機が相互に接続されたネットワークは、ストレージエリアネットワーク（SAN）と称される。SANは、ファイバチャネルを用いて構築されることが多い。尚、ファイバチャネルを用いたSANに接続される記憶装置システムの一例が、特開平10-333839号公報に開示されている。10

#### 【0004】

第二は、計算機で実行されるアプリケーションプログラムが認識するファイル単位で、記憶装置システムと計算機との間のI/Oが行われる「ファイルI/Oインターフェース」である。ファイルI/Oインターフェースでは、従来のファイルサーバで使用されているネットワークファイルシステムのプロトコルが用いられることが多い。特に、ファイルサーバの機能のうち、記憶装置システムに特化した機能と記憶装置システムとを組み合わせた装置は、ネットワークアタッチドストレージ（NAS）と称される。NASは、ローカルエリアネットワーク（LAN）等に直接接続される。

#### 【0005】

SANでは、計算機間のメッセージ交換用のネットワークとは別に設けられた記憶装置システム専用の高速ネットワークで、計算機と記憶装置システムとが接続される。このため、SANでは、計算機とLANで接続されるNASに比べて高速なデータの授受が可能である。また、SANではブロックI/Oインターフェースが採用されているので、NASに比べてプロトコル処理のオーバヘッドが小さく、高速なレスポンスを得ることができる。20

#### 【0006】

この反面、SANは、記憶装置システム専用のネットワークが必要となるので、ユーザにおいては、SANを導入する際のコストが大きい。したがって、SANは、主に企業の基幹システムで使用されており、データベースの構築を目的として利用されることが多い。30

#### 【0007】

一方、NASは既存のLANをそのまま使用できるので、導入の際のコストが小さく、導入が容易である。又、NASでは標準化されたネットワークファイルシステムが使用されるので、NASユーザは、一般的なファイルを単位としてデータを管理できる。このため、データの管理が容易で、かつ複数の計算機で容易にファイルを共有することができる。

#### 【0008】

その反面、NASは、計算機間の通信にも使用されるLANを介して計算機とデータのやり取りを行う。したがって、LANの負荷が増大するおそれがある。また、ネットワークファイルシステムで用いられるプロトコルの処理のオーバヘッドが大きく、レスポンスの性能はSANより低速となる。このことから、NASは、企業のファイル管理システムに主に使用されており、Webコンテンツの管理や、CADのデータファイル管理等のアプリケーションに利用されることが多い。40

#### 【0009】

上述のように、NASとSANはお互いに補完しあう関係にあり、また、適用される分野も異なる。したがって、一般に、両方は、適所に使い分けられている。

#### 【0010】

また、各計算機が、記憶装置システムが有する記憶容量を、必要な時に必要な分だけ利用できるように、記憶装置システム内の資源（以下「デバイス」）の設定、デバイスの各50

計算機への割り当てを容易に行えるようにする技術として、ストレージプールと呼ばれる技術がある。ストレージプールの一例は、特開2001-142648公報に開示されている。

【発明の開示】

【発明が解決しようとする課題】

【0011】

従来技術でS A NとN A Sとを含んだシステムを構築しようとした場合、S A Nの機能を有する記憶装置システムと、N A Sの両方が必要になる。この場合、構築されたシステムのユーザは、種類の異なる記憶装置システムを別個に管理する必要があり、システムを管理し運用するためのコストが増大してしまうという課題がある。

10

【0012】

また、S A Nの機能を備える記憶装置システムと、N A Sの各々を有するシステムでは、システムのユーザは、各々の記憶装置システムごとに記憶容量を管理する必要がある。具体的には、ユーザは、1台1台異なる環境設定や記憶容量の管理を行わなくてはならない。したがって、システムのユーザにとって、記憶装置システムの管理が煩雑かつ困難なものになるという課題がある。

【0013】

さらに、S A NとN A Sとを含んだシステムにおいて、S A Nに使用される記憶装置システムの記憶容量が不足したとき、たとえN A Sの記憶容量に余裕があったとしても、各々の記憶装置システムの管理体系が異なるので、ユーザは、N A Sに存在する記憶容量を自由に利用することはできない。つまり、ユーザは、システム全体が有する記憶容量を有效地に利用できないという課題がある。

20

【0014】

本発明の目的は、複数種類の記憶装置システムにかかる管理運用コストを低減できる記憶装置システムを提供することである。

【0015】

また、本発明のさらなる目的は、システム全体が有する記憶容量の管理を容易にするとともに、記憶容量の効率的な利用を実現する記憶装置システムを提供することである。

30

【課題を解決するための手段】

【0016】

上記目的を達成するために、本発明の記憶装置システムは、複数種類のインターフェース制御装置が装填される複数のスロット、複数のディスク装置、及び前記複数のディスク装置が有する記憶領域を前記スロットに割り当てる手段とを有する構成とする。

【0017】

また、複数種類のインターフェース制御装置には、S A Nの機能を備えるブロックI/Oインターフェース制御装置、N A Sの機能を備えるファイルI/Oインターフェース制御装置が含まれる構成が考えられる。

【0018】

さらに、本発明の好ましい実施態様として、各インターフェース制御装置は同一の形状を有する構成が考えられる。

40

【0019】

また、本発明の好ましい実施態様として、記憶装置システムは、あるインターフェース制御装置に割り当てた記憶領域を別の種類のインターフェース制御装置に割り当てる手段とを有する構成が考えられる。

【発明を実施するための最良の形態】

【0020】

図1は、本発明を適用した記憶装置システムの実施形態を示した図である。記憶装置システム1は、LAN20及び21を介して、NASクライアント400と接続される。本実施形態では、LAN20及び21はIPネットワーク(以下単に「ネットワーク」と称する)である。また、記憶装置システム1は、SAN30を介して、SANクライアント500と接続される。本実施形態

50

ではSAN30は、ファイバチャネルであるとする。

【0021】

又、記憶装置システム1は、管理端末18と接続されている。管理端末18は、ユーザに対し各種情報を提示するためのディスプレイ、及びユーザからの情報の入力を受け付けるためのキーボード、マウスなどの入出力装置を備える。管理端末18は、通信ネットワーク19を介して記憶装置システム1と通信することができ、ユーザが、記憶装置システム1の各種設定を行うために用いられる。

【0022】

このように、記憶装置システム1は、ファイバチャネルで構成され、ブロックI/Oインターフェースが用いられるSANと、IPネットワークで構成され、ファイルI/Oインターフェースが用いられるNASに適応するインターフェースの、両方のインターフェースを有する記憶装置システムである。尚、記憶装置システム1のインターフェースは、SAN対応とNAS対応の任意の組み合わせが可能である。特に、SANのみ又はNASのみに対応するインターフェースを有するシステムを構成することも可能である。

10

【0023】

記憶装置システム1は、ディスクコントローラ(以下、「DKC」と称する)11及び記憶装置(以下「ディスク」と称する)1700を有する。DKC11は、ネットワークチャネルアダプタ(以下CHN)1100、ファイバチャネルアダプタ(以下CHF)1110、ディスクアダプタ1200(以下、DKA)、共有メモリ13(以下SM)、キャッシュメモリ14(以下CM)、及びディスクプール管理部15(以下DPM)を有する。

20

【0024】

CHN1100は、ファイルI/OインターフェースでNASクライアント400と接続されるインターフェース制御部である。CHF1110は、ブロックI/OインターフェースでSANクライアント500と接続されるインターフェース制御部である。以下、CHN及びCHFをまとめてチャネルアダプタ(以下CH)と称する。

【0025】

DKA1200は、ディスク1700と接続されている。DKA1200は、接続されているディスク1700とクライアント間のデータ転送を制御する。SM13には、記憶装置システム1の構成を管理する構成管理情報や、CM14の制御情報等が格納される。SM13に格納されているデータは、全てのCHおよび全てのDKAに共有される。CM14には、ディスク1700のデータが一時的に保存(以下「キャッシングする」と称する)される。CM14のデータも、全てのCHおよび全てのDKA1200に共有される。記憶装置システム1では、すべてのCHが、CM14及びすべてのディスク1700へアクセスすることが出来る構成となっている。このような構成は、クロスバスイッチ等で実現できるディスクプール管理部15(以下DPM)は、全てのディスク1700を一括して管理する。具体的には、DPM15には、ディスク1700の記憶容量全てをまとめて一つのディスクプールとして扱うための情報が格納されている。尚、CH、DKA及びDPM15は、管理ネットワーク16によって各自接続される。又、DPM15は、通信ネットワーク19を介して、管理端末18と接続される。

30

【0026】

図2は、記憶装置システム1の外観を示す図である。DKC筐体115は、CHN1100、CHF1110、DKA1200、SM13、及びCM14を格納する。ディスクユニット筐体(以下、「DKU」と称する)180は、ディスク1700を格納する。SM13は、実際には複数のコントローラボード1300で構成される。また、CM14も、複数のキャッシングボード1400で構成される。記憶装置システム1の使用者は、ボードやディスクの枚数を増減して、所望の記憶容量を得る。

40

【0027】

DKC筐体115は、複数のインターフェースアダプタスロット190を有する。各々のインターフェースアダプタスロット190には、CHN1100等が搭載されたアダプタボードが格納される。本実施形態においては、インターフェースアダプタスロット190の形状、アダプタボードのサイズ及びコネクタの形状を、インターフェースの種類を問わず同一にし、各種インターフェースのアダプタの互換性を保つようとする。したがって、DKC筐体115が有するインタ

50

フェースアダプタスロット190には、インターフェースの種類を問わず、任意のアダプタボードを装填することができる。

【0028】

また、記憶装置システム1の使用者は、CHN1100及びCHF1110のアダプタボードの数を自由に組み合わせて、記憶装置システム1のインターフェースアダプタスロット190に装填することが出来る。これにより、記憶装置システム1の使用者は、記憶装置システム1のインターフェースを自由に構成することができる。

【0029】

図3は、CHN1100が作りこまれたアダプタボードの構成を示す図である。コネクタ11007は、DKC筐体115が有するコネクタと接続される。本実施形態においては、上述したように、CHN1100及びCHF1110は同一形状のコネクタを有する。インターフェースコネクタ2001は、LANに対応している。尚、アダプタボードがCHF1110の場合、インターフェースコネクタ2001は、ファイバチャネルに対応する。

【0030】

図4は、CHN1100の内部構成を示す図である。CHN1100は、プロセッサ11001、LANコントローラ11002、管理ネットワークコントローラ11003、メモリ11004、SMI/F制御部11005及びCMI/F制御部11006を有する。

【0031】

プロセッサ11001は、CHN1100全体の制御を行う。LANコントローラ11002は、CHN1100とLANとの間の通信の制御を行う。管理ネットワークコントローラ11003は、CHN1100と管理ネットワーク16との間の通信の制御を行う。プロセッサ11001に接続されるメモリ11004には、プロセッサ11001が実行するプログラムや制御データが格納される。

【0032】

SMI/F制御部11005は、CHN1100とSM13との間のデータ転送を制御する。CMI/F制御部11006は、CHN1100とCM14との間のデータ転送を制御する。

【0033】

メモリ11004には、オペレーティングシステムプログラム110040、LANコントローラドライバプログラム110041、TCP/IPプログラム110042、ファイルシステムプログラム110043、ネットワークファイルシステムプログラム110044、ディスクボリューム制御プログラム110045、キャッシュ制御プログラム110046、ディスクプール情報取得プログラム110048、及びディスクプール情報管理テーブル110049が格納される。又、メモリ11004は、データバッファ110047を有する。

【0034】

オペレーティングシステムプログラム110040は、メモリ11004に格納されたプログラム全体の管理や入出力制御に用いられる。LANコントローラドライバプログラム110041は、LANコントローラ11002の制御に用いられる。TCP/IPプログラム110042は、LAN上の通信プロトコルであるTCP/IPの制御に用いられる。ファイルシステムプログラム110043は、ディスク1700に格納されるファイルの管理に用いられる。

【0035】

ネットワークファイルシステムプログラム110044は、ディスク1700に格納されるファイルをNASクライアント400に提供するためのネットワークファイルシステムのプロトコル制御に用いられる。ディスクボリューム制御プログラム110045は、ディスク1700に設定されたディスクボリュームへのアクセスを制御するために用いられる。キャッシュ制御プログラム110046は、CM14に格納されるデータの管理や、CM14におけるデータのヒット/ミス判定等の制御に用いられる。

【0036】

ディスクプール情報取得プログラム110048は、CHN1100が、SM13に格納されているディスクプールの情報を入手する際に実行される。ディスクプール情報管理テーブル110049には、入手されたディスクプール情報が格納される。データバッファ110047は、NASクライアント400とCM14との間のデータ転送速度の調整、及びファイルデータをキャッシュする

10

20

30

40

50

ために用いられる。

【0037】

なお、プロセッサ11001は、単一のプロセッサや、複数のプロセッサの集合体の構成も考えられる。例えば、制御に伴う処理を水平負荷分散させる対称型マルチプロセッサの構成でもよい。また、複数のプロセッサのうち、一方がファイルI/Oインターフェースのネットワークファイルシステムプロトコル処理を行い、他方のプロセッサがディスクボリュームの制御を行うという非対称型マルチプロセッサの構成としてもよい。またそれらの組み合わせの構成としてもよい。

【0038】

図5は、CHF1110の内部構造を示す図である。CHF1110は、プロセッサ11101、ファイバチャネルコントローラ11102、管理ネットワークコントローラ11103、メモリ11104、SM制御部11005、CM制御部11006及びデータバッファ11107を有する。

10

【0039】

プロセッサ11101は、CHF1110全体を制御する。ファイバチャネルコントローラ11102は、CHF1110とSANとの間の通信を制御する。管理ネットワークコントローラ11103は、CHF110と管理ネットワーク16との間の通信制御を行う。プロセッサ11101と接続されるメモリ11104には、プロセッサ11101で実行されるプログラムや制御データが格納される。

【0040】

SM制御部11005は、CHN1100とSM13との間のデータ転送を制御する。CM制御部11006は、CHN1100とCM14との間のデータ転送を制御する。データバッファ11107は、SANクライアント500とCM14との間のデータ転送速度の差を解消するための緩衝メモリの役割を果たす。

20

【0041】

メモリ11104には、ファイバチャネルコントローラドライバプログラム111041、オペレーティングシステムプログラム111040、ディスクボリューム制御プログラム111045、キャッシュ制御プログラム110046、ディスクプール情報取得プログラム111048、及びディスクプール情報管理テーブル111049が格納される。

【0042】

ファイバチャネルコントローラドライバプログラム111041は、ファイバチャネルコントローラ11002の制御に用いられる。その他のプログラムは、各々図4のCHN1100のメモリに格納されるプログラムと同じ機能を有するので、説明を省略する。

30

【0043】

なお、プロセッサ11101も、CHN1100のプロセッサ11001と同様、単一プロセッサではなく、マルチプロセッサの構成としてもよい。

【0044】

図6は、SM13の内容を示す図である。SM13は、構成管理情報格納エリア131を有する。構成管理情報格納エリア131には、チャネルアダプタ集中管理テーブル1310及びディスクプール情報集中管理テーブル1311が格納される。SM13には、その他、記憶装置システム1が有する全ての構成情報が格納されるが、本実施形態においては、説明を省略する。

【0045】

チャネルアダプタ集中管理テーブル1310には、記憶装置システム1が有する全てのCHの管理情報が格納される。チャネルアダプタ集中管理テーブル1310の詳細は後述する。ディスクプール情報集中管理テーブル1311には、記憶装置システム1が有する全てのディスク1700で構成される1つの仮想的な記憶領域であるディスクプール6の管理情報が格納される。

40

【0046】

図7は、DPM15の内部構造を示す図である。DPM15は、プロセッサ151、メモリ152、管理ネットワークコントローラ153及び通信コントローラ155を有する。

【0047】

プロセッサ151は、DPM15全体の制御を行う。メモリ152には、プロセッサ151が実行する制御プログラム、及び制御に使用されるデータが格納される。管理ネットワークコントロ

50

ーラ153は、DKC11が有する他の装置とDPM15との間のデータ転送の制御を行う。通信コントローラ19は、管理端末18とDPM15との間のデータ転送の制御を行う。

#### 【0048】

メモリ152には、ディスクプール管理プログラム1521、ディスクプール情報集中管理テーブル1522、チャネルアダプタ認識・認証管理プログラム1523及びチャネルアダプタ集中管理テーブル1524が格納される。ディスクプール管理プログラム1521は、記憶装置システム1が、ディスク1700を用いてディスクプール6を構築し、管理する際に実行される。チャネルアダプタ認識・認証管理プログラム1523は、記憶装置システム1が、CHN1100及びCHF1110の装填状況を検出し、それらが動作することを確認する際に実行される。

#### 【0049】

ディスクプール情報集中管理テーブル1522には、ディスクプール管理プログラム1521が構築したディスクプール6を管理するための情報と、ディスクプール6が有する記憶領域が各CHへどのように割り当てられているかを管理するための情報が格納される。ディスクプール情報集中管理テーブル1522の内容は、SM13に格納されるディスクプール情報集中管理テーブル1311と基本的に同一である。

#### 【0050】

チャネルアダプタ集中管理テーブル1524には、チャネルアダプタ認識・認証管理プログラム1523が検出し認証した全てのCHに関する情報が格納される。チャネルアダプタ集中管理テーブル1524の内容は、SM13に格納されたチャネルアダプタ集中管理テーブル1310と基本的に同一である。

#### 【0051】

図8は、ディスクプール6の構成を示す図である。ディスクプール6とは、複数のディスク1700の記憶容量を集合し、1つの仮想的な記憶領域として定義したものである。尚、ディスクプール6は記憶装置システム1に複数設定することもできる。具体的には、例えば、ディスク装置の性能特性（回転速度7200rpmと15000rpm等）やディスクグループの冗長度（RAID1とRAID5等）等、ディスク装置の性質の違いに基づいて、特性が一致するディスク装置群ごとにディスクプールを設定したり、記憶装置システムを使用するユーザのグループ（以下「ユーザグループ」と称する）毎に別々のディスクプールを設けることが考えられる。

#### 【0052】

ディスクプール6を構成するディスク1700の数は任意であり、必要に応じて、増設、減設される。このことにより、ディスクプール6の容量を変更することが可能である。

#### 【0053】

図8は、一つのディスクプール6と複数台のディスク1700、具体的にはRAIDグループ17000（以下RG）との対応関係の具体例を示す図である。RG17000は、複数台のディスク1700から構成されるRAID（Redundant Arrays of Inexpensive Disks）である。

#### 【0054】

一つのディスクプール6に割り当てられた記憶領域は、複数の論理デバイスLDEV170（以下LDEV）に分割されている。さらに、複数のLDEV170をまとめて論理的なボリュームとして定義したものを、論理ボリューム1750（以下LV）と称する。

#### 【0055】

図8では、4台のディスク1700で構成されたRG17000（RG0）が1つのディスクプール6に割り当てられている。そして、RG0が有する記憶容量は、記憶容量がLであるLDEV0～LDEV4に分割される。さらに、LDEV0～LDEV2で、1つのLV0が構成される。同様に、LDEV3及びLDEV4で、1つのLV1(1751)が構成される。以下、同様にいくつかのLVが構築される。尚、本図においては、LV0はCHN0に、LV1はCHN2に割り当てられる。また、LDEVは、RAIDにおける一つのストライプサイズに対応していても良い。

#### 【0056】

図9は、SM13に格納されているディスクプール情報集中管理テーブル1311の構成を示す図である。尚、DPM15に格納されているディスクプール情報集中管理テーブル1522もディ

10

20

30

40

50

スクプール情報集中管理テーブル1311と同一の構成を有する。以下ディスクプール情報集中管理テーブル1311について説明する。

【0057】

ディスクプール情報集中管理テーブル1311は、ディスクプール6の番号が登録されるエントリ13110、RG17000の番号が登録されるエントリ1311、LDEV170の番号が登録されるエントリ1312、LDEV170の容量についての情報が登録されるエントリ1313、LVの番号が登録されるエントリ1314、LVを構成するLDEV170の連結番号が登録されるCNエントリ1315、対応するLVが割り当てられているCHの番号が登録されるエントリ1316、ディスクプール6、RG、LVのCH間共有情報として、SAN(CHF)-NAS(CHN)間の共有の可否を示すS-N共有情報が登録されるエントリ1317、SAN(CHF)間の共有の可否を示すS-S共有情報が登録されるエントリ1318、NAS(CHN)間共有の可否を示すN-N共有情報が登録されるエントリ1319、及びユーザクラス共有情報が登録されるエントリ1320を有する。CH間共有情報等については後述する。尚、CN番号とは、複数のLDEV170を結合してLVを生成するときの結合の順番を示す番号のことである記憶装置システム1は、ディスクプール情報集中管理テーブル1311に登録された情報に基づいて、ディスクプール6を構成するRG17000及びLDEV170を確認する。また、記憶装置システム1は、ディスクプール情報集中管理テーブル1311に登録された情報に基づいて、ディスクプール6に属するLDEV170がどのCHに割り当てられているか、及びディスクプール6にどれだけ未使用の記憶容量が残っているか等を確認する。

10

【0058】

図10は、CM13に格納されているチャネルアダプタ集中管理テーブル1310の構成を示す図である。尚、DPM15が有するチャネルアダプタ集中管理テーブル1524もチャネルアダプタ集中管理テーブル1310と同一の構成を有する。チャネルアダプタ集中管理テーブル1310は、記憶装置システム1に装填されているCHの管理ネットワーク16におけるアドレスが登録されるエントリ13100、CHの識別番号が登録されるエントリ13101、CHの種類についての情報が登録されるエントリ13102、及びCHの稼働状態に関する情報が登録されるエントリ13103を有する。

20

【0059】

以下に、本発明の記憶装置システム1の本実施形態における動作を説明する。記憶装置システム1の主たる動作には次の2つがある。

30

(A)ボリューム割り当て処理 (B)ボリューム認識処理 (A)ボリューム割り当て処理とは、記憶装置システム1が、ディスクプール6に含まれる記憶容量をLVとしてCHに割り当てる処理である。

(B)ボリューム認識処理とは、記憶装置システム1のCHが、CHに割り当てられたLVを認識するための処理である。

【0060】

図11は、ボリューム割り当て処理の手順を示すフローチャートである。新しいLVをあるCH(以下CHNiとする)に割り当てる場合、記憶装置システム1の管理者は、管理端末18の管理ソフトウェアを起動する(ステップ6100)。

40

【0061】

管理者は、CHNiを示す情報とCHNiに割り当たいLVに関する情報を管理端末18に入力する。具体的には、管理者1は、管理ソフトウェアの実行により、管理端末18が有するディスプレイに表示される、未使用を示すエリアに表示されたLDEV170を示すアイコンを、マウス等を用いて、画面上のCHNiを示すエリアにドラッグし移動する。このとき、管理者は、画面上で、必要な記憶容量分の複数のLDEVを同時に移動することができる。尚、本実施形態では、LDEVは、製品出荷時にあらかじめ記憶装置システム1に設定されているものとする(ステップ6101)。

【0062】

管理者1に選択されたCHNiとそれに割り当られるLDEV170の情報は、管理端末18からDPM15に送信される。DPM15は、送信された情報を受信する(ステップ6150)。

【0063】

50

DPM15は、管理端末18から情報を受け取ると、ディスクプール管理プログラム1521を実行する。DPM15は、指定されたLDEV170が属するプール番号を特定するとともに、ディスクプール情報集中管理テーブル1522を操作し、指定されたLDEV170の割り当て先CH情報エントリ1312の内容を、「未割り当て」から「CHNi」に変更する（ステップ6151）。

#### 【0064】

次に、DPM15は、指定されたLDEV170に新規の論理ボリューム番号LVjを割り当て、ディスクプール情報集中管理テーブル1522のLV番号エントリ1314やCN番号エントリ1315に、割り当てられたLVに関する情報を登録する。（ステップ6152）。

#### 【0065】

その後、DPM15は、管理ネットワーク16を介して、CHNiに対し、LDEV170が追加され、追加されたLDEV170が論理ボリュームLVjとして定義されたことを通知する（ステップ6153）。

10

#### 【0066】

CHNiは、この通知を受信すると（ステップ6160）、ディスクプール情報取得プログラム110048を実行して、以下の処理を行う。CHNiは、CHNi自身が有するディスクプール情報管理テーブル110049に登録された情報を、DPM15から通知された内容に従って変更する（ステップ6161）。また、CHNiは、SM13に格納されたディスクプール情報集中管理テーブル1311の該当する個所を変更するように、SM13を制御する（ステップ6162）。

#### 【0067】

次にCHNiは、ディスクボリューム制御プログラム110045を実行して、割り当てられたLDEV群を論理ボリュームLVjとして構成するための情報及び割り当てられたLDEV群をLVjとして使用するための情報（LV番号等）を、ディスクプール情報管理テーブル110049に登録する（ステップ6163）。

20

#### 【0068】

その後、CHNiはDPM15へ終了報告を送信する（ステップ6164）。終了報告を受信したDPM15は、管理端末18へ終了報告を送信する（ステップ6155）。以上でボリューム割り当て処理が終了する。

#### 【0069】

なお、上記では新規のボリューム割り当ての処理動作を説明したが、既存の論理ボリュームLVpに記憶容量を追加する処理も上記同様の処理で実現することができる。既存の論理ボリュームLVpに記憶容量を追加する場合、管理端末18の画面上でLVpに新規のLDEV170を追加するようにユーザが指示すること、ユーザのLDEV170追加の指示を受け、DPM15が、LVpにLDEV170を追加するようにディスクプール情報集中管理テーブル1522を変更するとともにCHNiに変更に関する情報を送信すること、及び変更に関する情報を受信したCHNiが、LVpにLDEV170を追加するようにディスクプール情報管理テーブル110049とSM13のディスクプール情報集中管理テーブルを変更することが、上記の処理と異なる点である。なお、上記の処理は、CHF1110へのLV追加にも同様に適用できる。

30

#### 【0070】

図12は、記憶装置システム1の起動時もしくはCH増設作業時等に行われる、CHによるボリューム認識処理の手順を示すフローチャートである。記憶装置システム1へ電源が投入されたこと、あるいは記憶装置システムにCHが増設されたこと（ステップ6200）を契機に、各CHは、CH自身の初期設定を行う（ステップ6201）。

40

#### 【0071】

初期設定を終了した後、各CHは、ディスクプール情報取得プログラム110048の実行を開始する（ステップ6202）。その後、各CHは、SM13のディスクプール情報集中管理テーブル1311を参照し、自身が使用するLVに関する情報を取得する（ステップ6203）。

#### 【0072】

使用するLVに関する情報を取得したCHは、自身のディスクプール情報管理テーブル110049にその情報を登録する（ステップ6204）。以上の処理により、各CHは、CHに割り当てられたLVを認識することができる（ステップ6205）。

50

## 【0073】

このような制御により、各CHは、記憶装置システム1全体のLVに関する情報が一括して登録されているディスクプール情報集中管理テーブル1311に登録された情報に基づいて、記憶装置システム1が有するLVを使用することが出来る。特に、本実施形態では、LVは各CHに重複しないようにディスクプール情報集中管理テーブル1311で排他的に割り当てられているので、各CHはLVを排他的に使用することができる。なお、上記の認識処理は、CHN1100やCHF1110といったCH種によらず同一である。

## 【0074】

また、上記の実施形態では、ディスク1700を論理デバイスLDEV170という形態に細分化し、その結合体として論理ボリュームLVを定義する例を説明した。しかし、ディスク1700をLDEV170に分割しない実施形態も考えられる。この場合、一つもしくは複数のRG17000の記憶容量全体でディスクプール6が生成され、そのディスクプール6から所定の記憶容量がLVとしてCHに割り当てられるという処理が行われる。

10

## 【0075】

上述した実施形態によれば、SANすなわちブロックI/Oインターフェースを備える複数のファイバチャネルアダプタ、及びNASすなわちファイルI/Oインターフェースを備える複数のネットワークチャネルアダプタとの任意の組み合わせを有する記憶装置システムを実現でき、スケーラビリティや構成自由度に優れた記憶装置システムを構築することが可能になる。

20

## 【0076】

また、記憶装置システムが有する記憶領域の設定や管理運用を单一の管理端末から統括的に行うことで、管理運用の容易性が向上し管理コストを低減することができる。

## 【0077】

さらに、SANとNASとが混在する記憶装置システムが有する記憶容量を、排他制御を行いながら、安全に使用することができるので、記憶容量の集中管理による容易利用可能な環境を提供するとともに、記憶容量の最適分配を実現でき、管理運用コストを低減することができる。

## 【0078】

次に、本発明の第二の実施形態を説明する。前述の実施形態においては、1台のCHF1110又は1台のCHN1100各々に、異なるLVを割り当てる方法を示したが、本実施形態では、複数のCHに同一のLVを割り当てる場合のディスクプールの管理方法、およびLVの共有の方法を説明する。

30

## 【0079】

本実施形態においては、図9に開示されているCH間共有情報1317～1319が用いられる。また、本実施形態においては、ディスクプール6、RG17000、LVの各々に対応してCH間共有情報1317～1319が設定される。

## 【0080】

CH間共有情報には、主に以下の3つの属性に関する情報が含まれる。

- (1)リード/ライト可能共有属性（以下「R/W」）
- (2)リードオンリ共有属性（以下「R」）
- (3)共有禁止属性（以下「P」）

40

これらの属性は、P>R>R/Wの順に順位付けされている。そして、これらの属性は、一度LVに設定されると、設定された属性の順位よりも上位の順位の属性にしか変更することができない。つまり、一度Rの属性が設定されたLVは、Pの属性には変更されうるが、R/Wの属性には変更できない。これは、一旦共有が禁止されたLVを共有可能とすると、あるCHがあるLVを専有していることを前提としてそのLVを使用していた場合、そのLVの属性がR等に変更されると、公開したくないデータが公開されたり、データが改変され、論理的な矛盾が生じる可能性があるからである。

## 【0081】

また、上位の属性が設定されているLVに下位の属性を設定したい場合には、ユーザは、

50

一旦LVを削除して、LVに割り当てられていたLDEV等の記憶領域をディスクプール6に戻し、新たにLVを再構成する必要がある。なお、この3つ以外の属性を設けることも可能である。

#### 【0082】

また、CH間共有情報は、ディスクプール6 > RG17000 > LV又はLDEV170の順で継承される。すなわち、上位（ディスクプールが最上位）の共有情報が、下位のRG17000とLDEV170が有する共有情報に優先する。またRG17000の共有情報がLDEV170が有する共有情報に優先する。

#### 【0083】

例えば、ディスクプール6の共有属性がPであるとすると、その下位のRG17000等の属性もすべてPとなる。よって、記憶装置システム1内のディスクプール6が有する記憶領域は、一切共有することができなくなる。また、P属性の順位が高く設定されているので、下位のRG17000等にPよりも順位の低い属性を個別に付与又は変更することができない。また、ディスクプール6の共有属性がR/Wだとすると、その下位のRG17000等の属性もすべてR/Wとなる。よって、記憶装置システム1が有する記憶領域すべて共有することができる。またこの場合、下位のRG17000等に任意の属性R、Pを個別に付与することが可能となる。

10

#### 【0084】

このように、ディスクプール6や、RG17000といった、記憶装置システム1が有する記憶領域の論理的レベルの各々で共有の可否を設定することで、LV割り当ての際に、記憶装置システム1の記憶領域の一切の共有を認めないようにしたり、記憶装置システム1の記憶領域の一切を自由に共有することができるようにしたりすることができ、ボリューム割り当ての容易さと、セキュリティの強度をコントロールすることができる。

20

#### 【0085】

以下、いくつかの場合に分け、CH間の共有ボリュームの設定について説明する。ここでは、CHF1110とCHN1100、すなわちSANとNASの間で単一の論理ボリュームを共有する場合について、以下の二つの形態に分けて説明する。

#### 【0086】

第一の例は、NAS over SANと呼ばれる形態である。これは、CHN1100に割り当てられたLV上に作成されたファイルシステムに対し、記憶装置システム1の外部にある計算機がLAN2経由でファイルアクセス要求をCHN1100に送信し、アクセス要求をした計算機が、SAN30を経由してデータを送受信するというものである。

30

#### 【0087】

第二の例は、SAN経由バックアップと呼ばれる形態である。これは、CHN1100に割り当てられたファイルシステムが構築されるLVに格納されているデータを、CHF1110に接続されたSAN経由でバックアップするというものである。

#### 【0088】

上述の例は、SAN - NAS間の計算機同士における論理ボリュームの共有に関するので、上記CH間共有情報のうち、S-N共有情報1317が使用される。

#### 【0089】

図9において、ディスクプール6（番号1）、RG17000（番号2）、LDEV170（番号6～8）で構築されたLV8が、SAN対応の計算機とNAS対応の計算機とで共有されるLVとして定義されている。これは、CH番号エントリ1316に複数種類のCHが登録されていることから明らかである。

40

#### 【0090】

また、LV8に対応するS-N共有情報エントリ1317には、ディスクプール1がR/W、RG2がR/W、LV8がR/Wであるという情報が登録されている。尚、同一ディスクプールで同一RG17000内のLV8以外のLVはすべてP属性に変更されており、共有が禁止されている。さらに、未使用のLDEV170には、全てR/W属性が設定されている。

#### 【0091】

このような共有属性の設定は、ディスクプール6に関してはディスクプール定義時に、R

50

G17000に関しては、RG17000をディスクプール6に登録する時、LVやLDEVに関しては、ボリューム割り当て処理動作時に、それぞれ行われる。具体的な設定情報の指示は、ユーザが管理端末18を介して行う。SAN-SAN間の共有情報（S-S共有情報）や、NAS - NAS間の共有情報（N-N共有情報）も上記と同様に各々のエントリ1318、1319に設定される。

#### 【0092】

上述した第二の実施形態においては、LVの共有について複数レベルの属性を設定する例を示した。次の第三の実施形態においては、これらの属性に加え、LVを使用するユーザについて複数レベルの属性を設定して、各レベルのユーザに対して異なるアクセス制限をする例について説明する。本実施形態においては、NAS - NAS間の共有について説明する。

#### 【0093】

図9において、ディスクプール情報集中管理テーブル1311（1522）のユーザクラス共有情報エントリ1320は、ユーザに関する属性（ユーザクラス）ごとのサブエントリを有する。ユーザクラスとしては次の5つの属性がある。

- (1)Everyone : 全ユーザ（以下「E」）
- (2)Company : 同一企業内ユーザ（以下「C」）
- (3)Domain : 同一ドメイン内ユーザ（以下「E」）
- (4)WorkGroup : 同一ワークグループ内ユーザ（以下「W」）
- (5)UserGroup : 同一ユーザグループ内ユーザ（以下「U」）

尚、一般的に、(1)から(5)の順で、対象となるユーザの数は少なくなる。

#### 【0094】

各ユーザクラスには、さらに、第二の実施形態において説明したCH間ボリューム共有属性と同様の属性、すなわち、R/W、R、Pの3種類の属性が付与される。これらの属性の意味、順位、継承、変更可否等の考え方は第二の実施形態と同一であるので、説明は省略する。

#### 【0095】

図9に示されるLV0を例に説明する。ユーザクラス共有情報エントリ1320の各サブエントリには、E=P、C=R、D=R、W=R/W、U=R/Wの情報が登録されている。すなわち、LV0は、同一ユーザグループおよび同一ワークグループのユーザがR/WできるLVであるが、これらに属さない同一ドメイン内および同一企業のユーザはこのLVの内容を変更することが出来ない。

#### 【0096】

このようなユーザクラス共有属性の設定は、ディスクプール6に関してはディスクプール定義時に、RG17000に関しては、RG17000がディスクプール6へ登録される時に、LVやLDEVに関してはボリューム割り当て処理動作時に、それぞれ行われる。この設定の指示は、ユーザが管理端末16を介して行う。

#### 【0097】

尚、記憶装置システム1は、ユーザの管理を行うために、ユーザが、上記1～5のユーザクラスの、何という名前のグループに属しているか等の情報が格納されたユーザクラス定義情報テーブルを有する必要がある。このユーザクラス定義情報テーブルには、LV毎にユーザの情報が登録されている。

#### 【0098】

具体的には、NASにはファイルシステム毎にユーザを管理する機構が備わっているため、ユーザは、CHN1100にユーザ管理情報を設定する必要がある。そのユーザ情報を設定する際に、ユーザクラス定義情報テーブルをあわせて作成すればよい。

#### 【0099】

本実施形態によれば、記憶装置システム1にLVを設定する際に、管理者が適切なユーザクラスを選択する。こうすることにより、LVレベルでユーザクラスに応じたセキュリティを設定することができる。この機能により、万一、許可されていないユーザクラスのユーザがこのLVをアクセスしようとしても、LVごとにユーザの認証が行われるため、アクセス権のないユーザのアクセスは抑止される。

10

20

30

40

50

**【 0 1 0 0 】**

このようなセキュリティ機能により、一台の大規模記憶装置システム1をデータセンタに設置し、複数の企業が単一のディスクプールを共有しながら、アクセスセキュリティを維持することが可能になる。

**【 0 1 0 1 】**

NASにはもともとユーザ管理の機能が備わっているので、本実施形態をNASへ適用するは容易である。一方、SANにも同様にユーザ管理の機能を実装することにより、まったく同様にユーザクラス間のアクセスセキュリティを実現したディスクプール管理およびボリューム割り当てを実現することができる。

**【 0 1 0 2 】**

また、本実施形態では、単一のCH内でのユーザクラス間アクセスセキュリティの実現を説明したが、第二実施形態のCH間ボリューム共用の構成と第三実施形態のユーザクラス間アクセスセキュリティの構成をまとめた実施形態も可能である。すなわち、記憶装置システム1は、CH間でボリュームを共用しながら、適切なユーザクラスのユーザのみにアクセスを許可するというような使い方をユーザに提供できる。

**【 0 1 0 3 】****[ 発明の効果 ]**

本発明によれば、記憶装置システムの管理運用コストを低減できる。また、本発明によれば、記憶容量の管理が容易な記憶装置システムを提供できる。

**【 0 1 0 4 】**

さらに、本発明によれば、記憶装置システムの運用の自由度をあげることができる。

**[ 図面の簡単な説明 ]****【 0 1 0 5 】**

【図1】本発明の実施形態のシステム構成図である。

【図2】本発明の実施形態のシステム概観図である。

【図3】チャネルアダプタの概観図である。

【図4】ネットワークチャネルアダプタ(CHN)の構成図である。

【図5】ファイバチャネルアダプタボード(CHF)の構成図である。

【図6】SMの構成図である。

【図7】ディスクプール管理部の構成図である。

【図8】ディスクプールの構成図である。

【図9】ディスクプール情報集中管理テーブルの構成図である。

【図10】チャネルアダプタ集中管理テーブルの構成図である。

【図11】ボリューム割り当て処理フローチャートである。

【図12】ボリューム認識処理フローチャートである。

**[ 符号の説明 ]****【 0 1 0 6 】**

1...記憶装置システム、11...ディスクコントローラ、1700...ディスク、1100...ネットワークチャネルアダプタ(CHN)、400...NASクライアント、1110...ファイバチャネルアダプタ(CHF)、500...SANクライアント、13...共有メモリ(SM)、14...キャッシュメモリ(CM)、15...ディスクプール管理部(DPM)、16...管理ネットワーク、18...管理端末。

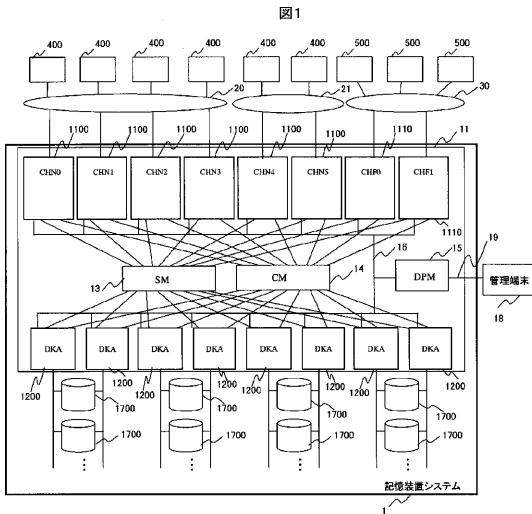
10

20

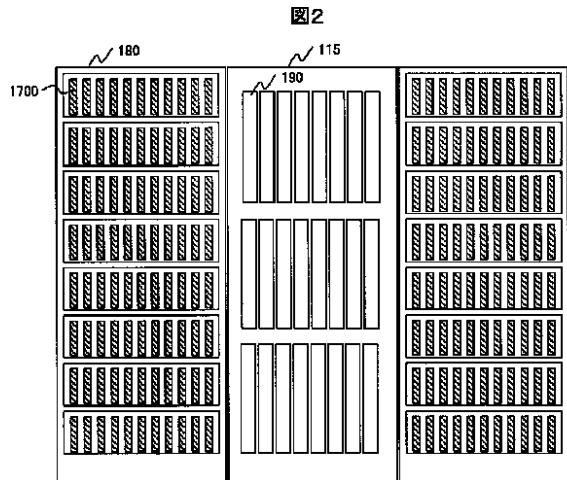
30

40

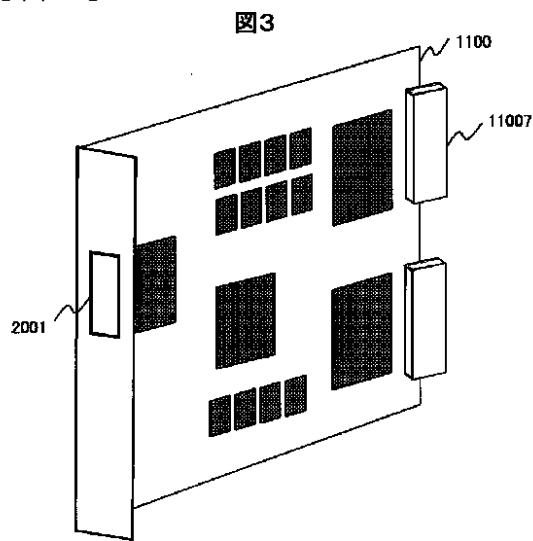
【図1】



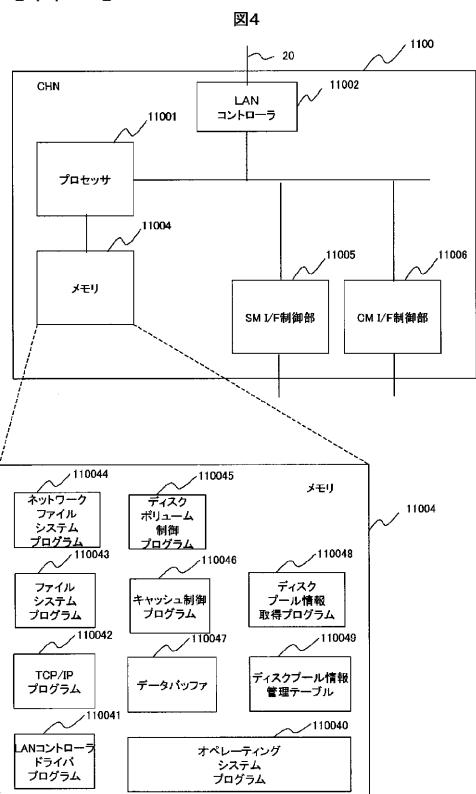
【図2】



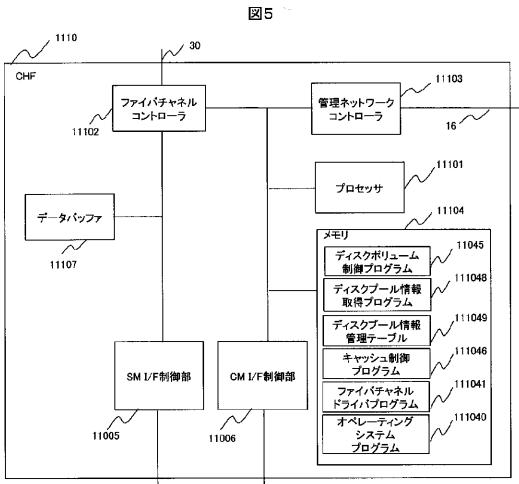
【図3】



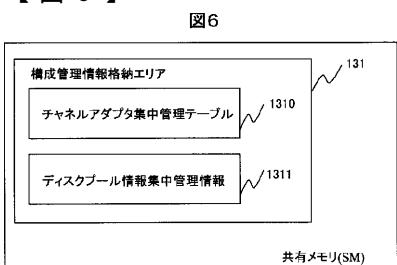
【図4】



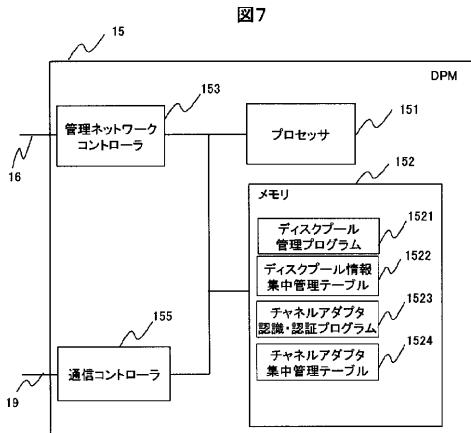
【図5】



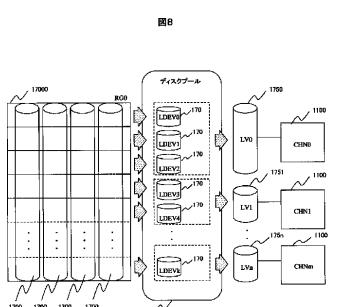
【図6】



【図7】



【図8】



【図9】

Pool# (1310)	RG# (1311)	LDEV# (1312)	容量 (1313)	LV# (1314)	CN# (1315)	CH# (1316)	S-N共有 (1317)		S-S共有 (1318)		N-N共有 (1319)		ユーパークラス (1320)	
							R/W	P	R/W	P	R/W	P	R/W	P
0	0	0	10	LV0	1	CHN0	P	P	P	P	R/W	P	R/W	P
		1	10		2									
		2	10		3									
		3	10	LV1	1	CHN1	P	P	P	P	R/W	P	R/W	P
		4	10		2									
		5	10	LV2	1	CHN2	P	P	P	P	R/W	P	R/W	P
		6	10		2									
		7	10		3									
		8	10		4									
		9	10	LV4	1	CHN4	P	P	R/W	P	P	R/W	P	R/W
1	1	0	10	LV4	1	CHN0	P	P	R/W	P	P	R/W	P	R/W
		1	10		2	CHN5								
		2	10		3									
		3	10		4									
		4	10		5									
		5	4	LV8	1	CHN5	R/W	P	P	P	R/W	R/W	R/W	P
		6	4	LV8	2	CHP1								
		7	4		2									
		8	4		3									
		9	4		4									
2	2	0	4	LV7	1	CHF1	P	P	P	P	R/W	R/W	R/W	P
		1	4		2									
		2	4		3									
		3	4		4									
		4	4	LV8	1	CHF1	R/W	P	P	P	R/W	R/W	R/W	P
		5	4	LV8	2	CHF1								
		6	4	LV8	3	CHF1								
		7	4		4									
		8	4		5									
		9	4		6									
3	3	0	10	未使用	-	-	R/W	R/W	R/W	P	R/W	R/W	R/W	P
		1	10	未使用	-	-	R/W	R/W	R/W	P	R/W	R/W	R/W	P
		2	10	未使用	-	-	R/W	R/W	R/W	P	R/W	R/W	R/W	P
		3	10	未使用	-	-	R/W	R/W	R/W	P	R/W	R/W	R/W	P
		4	10	未使用	-	-	R/W	R/W	R/W	P	R/W	R/W	R/W	P
		5	10		6									
		6	10		7									
		7	10		8									
		8	10		9									
		9	10		10									

【図10】

図10

スロットアドレス	チャネルアダプタ識別番号	チャネルアダプタ種	状態
0	CHN100B0013	CHN	正常
1	CN100B0059	CHN	正常
2	CHN100B0153	CHN	正常
3	CHN500A1936	CHN	正常
4	CHN500A1940	CHN	正常
5	CHN500A2002	CHN	正常
6	CHF500A1234	CHF	正常
7	CHF500A2199	CHF	正常

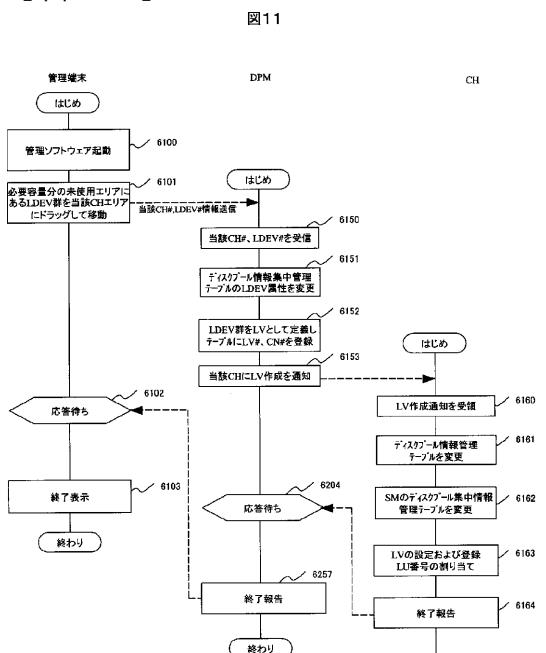
13100

13101

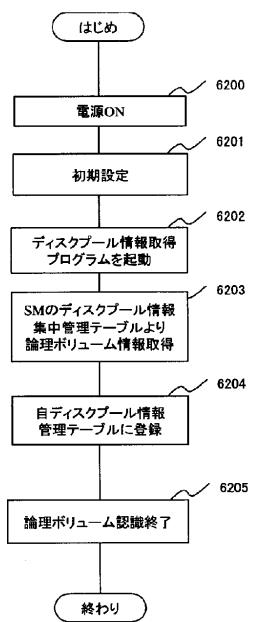
13102

13103  
チャネルアダプタ管理テーブル

【図11】



【図12】  
図12



---

フロントページの続き

(72)発明者 蘭田 浩二

神奈川県川崎市麻生区玉禅寺1099番地 株式会社日立製作所システム開発研究所内

(72)発明者 横畠 静生

神奈川県小田原市中里332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

F ターム(参考) 5B014 EB04

5B065 BA01 CA11 ZA02 ZA08

5B082 CA19