

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4354023号
(P4354023)

(45) 発行日 平成21年10月28日(2009.10.28)

(24) 登録日 平成21年8月7日(2009.8.7)

(51) Int.Cl. F I
G 0 6 F 3/06 (2006.01) G O 6 F 3/06 5 4 0
 G O 6 F 3/06 3 0 1 Z

請求項の数 4 (全 14 頁)

<p>(21) 出願番号 特願平10-89812 (22) 出願日 平成10年4月2日(1998.4.2) (65) 公開番号 特開平11-39105 (43) 公開日 平成11年2月12日(1999.2.12) 審査請求日 平成17年1月28日(2005.1.28) (31) 優先権主張番号 832367 (32) 優先日 平成9年4月2日(1997.4.2) (33) 優先権主張国 米国(US)</p> <p>前置審査</p>	<p>(73) 特許権者 591030868 コンパック・コンピューター・コーポレーション COMPAQ COMPUTER CORPORATION アメリカ合衆国テキサス州77070, ヒューストン, ステイト・ハイウェイ 249, 20555 20555 State Highway 249, Houston, Texas 77070, United States of America</p> <p>(74) 代理人 100089705 弁理士 社本 一夫</p> <p style="text-align: right;">最終頁に続く</p>
--	--

(54) 【発明の名称】 記憶サブシステム内でのドライブ移動及び記憶サブシステムへのボリューム追加のための装置及び方法

(57) 【特許請求の範囲】

【請求項1】

コンピュータ・システムにおいて、
 アレイ・コントローラと、
 前記アレイ・コントローラに接続され、第1の順番に配置されている複数のハード・ドライブを含んでいる記憶サブシステムであって、これらハード・ドライブにまたがって複数の論理ドライブ装置を指定可能であり、各論理ドライブ装置は、複数のハード・ドライブを含み、かつハード・ドライブの位置を昇順で指定した場合にのみ構成可能である、記憶サブシステムと

を備えており、

前記アレイ・コントローラは、

前記複数のハード・ドライブの識別番号、位置を示す位置番号、及びどの論理ドライブ装置に関連するかを示す情報を含む構成情報に基づき、前記複数のハード・ドライブが前記記憶サブシステムの中で前記第1の順番から第2の順番に配列変更されたかどうかを判定するよう構成され、

前記第2の順番に配列変更されたと判定した場合、前記複数のハード・ドライブの前記構成情報を、前記第2の順番となるように再編成するよう構成され、

再編成された前記構成情報に基づいて、第1の論理ドライブ装置に関連する複数のハード・ドライブが前記記憶サブシステム内に非昇順で配置されているかどうかを判定するよう構成され、

前記第 1 の論理ドライブ装置に関連する前記複数のハード・ドライブが非昇順であると判定した場合、これらハード・ドライブが昇順に配置されるようにするための情報を、ユーザに提供するように構成されていることを特徴とするコンピュータ・システム。

【請求項 2】

請求項 1 記載のコンピュータ・システムにおいて、前記複数のハード・ドライブの各々は、前記構成情報を格納するための記憶領域を備えていることを特徴とするコンピュータ・システム。

【請求項 3】

請求項 1 又は 2 記載のコンピュータ・システムにおいて、前記アレイ・コントローラは更に、

前記構成情報に基づいて、前記記憶サブシステムに第 2 の論理ドライブ装置が追加されたかどうかを判定するように構成されており、

前記第 2 の論理ドライブ装置が追加されたと判定した場合、前記構成情報を、前記第 1 の論理ドライブ装置及び前記第 2 の論理ドライブ装置それぞれにおける複数のハード・ドライブの配列を示すように再編成するように構成されていることを特徴とするコンピュータ・システム。

【請求項 4】

請求項 1 ~ 3 いずれかに記載のコンピュータ・システムにおいて、前記アレイ・コントローラは、前記記憶サブシステムの電源投入時に、複数のハード・ドライブの配列変更が生じたかどうかを判定するように構成されていることを特徴とするコンピュータ・システム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、アレイ・コントローラ及び記憶サブシステムを備えたコンピュータ・システムに関する。よりに詳しくは、本発明は、記憶サブシステム内での複数台のハード・ディスク・ドライブ (hard disk drive) の配列変更及び自動的再構成 (reconfiguration) を、記憶内容を喪失することなく行えるようにした、アレイ・コントローラに関するものである。ここでいうハード・ディスク・ドライブの配列変更には、記憶サブシステム内での複数台のハード・ドライブの配列変更 (rearrangement) も含まれ、また、記憶サブシステムへの 1 台ないし複数台の論理ドライブ装置の追加も含まれ、更には、それらだけに限られずその他の配列変更も含まれる。

【0002】

【従来の技術】

ドライブ・アレイとは、簡単にいえば、複数台のハード・ディスク・ドライブ等の物理ドライブの集合体であり、それらドライブを編成して 1 つの物理ドライブ・アレイとしたものである。1 つのドライブ・アレイは、1 つまたは複数の、論理ドライブ (論理ドライブ装置) と呼ばれる部分集合から成る。論理ドライブ装置は、ときに論理ボリュームと呼ばれることもある。1 台の論理ドライブ装置は、そのドライブ・アレイに含まれている複数台の物理ドライブのうち 1 台の物理ドライブ上に存在していることもあり、また、2 台以上の物理ドライブにまたがって存在していることもある。ただしオペレーティング・システムは、複数台の物理ドライブにまたがって存在している論理ドライブ装置であっても、1 つの連続した記憶空間として認識する。

【0003】

記憶サブシステムに用いられる一般的なドライブ・アレイの一例を図 1 に示した。この図 1 のドライブ・アレイは、3 台の物理ドライブにまたがって存在している 1 台の論理装置として構成されている。図 2 及び図 3 は、記憶サブシステムに用いられるドライブ・アレイの構成例として、更に別の例を示したものである。これらの図のうち、図 2 に示した記憶サブシステムは 6 台の物理ドライブから成り、それら 6 台の物理ドライブは、2 台の論

10

20

30

40

50

理ドライブ装置を備えた1つのドライブ・アレイを構成するように構成 (c o n f i g u r e) されている。図3に示した記憶サブシステムは、1つのドライブ・アレイを構成する3台の物理ドライブから成り、それら3台の物理ドライブにまたがって2台の論理ドライブ装置が存在している。当業者には容易に理解されるように、1つまたは複数の記憶サブシステムに含まれている1つまたは複数のドライブ・アレイの構成 (c o n f i g u r a t i o n) に際しては、様々な形に構成することができる。

【0004】

【発明が解決しようとする課題及び課題を解決するための手段】

本発明の1つの目的は、次のような記憶サブシステムを提供することにある。その記憶サブシステムとは、複数台のハード・ドライブを備えており、電源投入時に、この記憶サブシステム10の複数のハード・ドライブ位置の間でそれら複数台のハード・ドライブが配列変更されたか否かを判定できるように構成されていると共に、ハード・ドライブ位置が変更されていたならばその新たなハード・ドライブ位置に応じて、この記憶サブシステムに含まれている複数台のハード・ドライブを自動的に再構成 (r e c o n f i g u r a t i o n) することができるように構成されているものである。

【0005】

本発明の更なる目的は、次のような記憶サブシステム及びアレイ・コントローラを提供することにある。それら記憶サブシステム及びアレイ・コントローラにおいては、記憶サブシステムが複数台のハード・ドライブを備えており、それらハード・ドライブが第1状態に構成されており、この第1状態とは、論理ドライブ装置を含んでいる状態である。また、アレイ・コントローラは、記憶サブシステムに新たな論理ドライブ装置が追加されたか否かを判定できるように構成されていると共に、記憶サブシステムに含まれている全てのハード・ドライブを第2状態に自動的に再構成することができるように構成されているものである。

【0006】

【発明の実施の形態】

本発明の以上の目的及び利点、並びにその他の目的及び利点は、以下に示す、現時点で好適であると考えられる実施の形態についての説明を、添付図面と共に参照することによって、より明瞭に且つ容易に理解することができる。

【0007】

アレイ・コントローラは、一般的に、1つまたは複数のドライブ・アレイに接続して、そのドライブ・アレイの制御や構成を実行するために用いられている。アレイ・コントローラは、ネットワーク・サーバやパーソナル・コンピュータに接続され、また、その他の様々な種類のコンピュータ・システムにも接続される。図4は、コンピュータ・システムの一部を成しているサーバを示した図である。サーバは一般的に、中央処理装置 (C P U)、マザーボード、モニタ等で構成されている。

【0008】

図4のサーバ10は、様々な構成要素で構成されているが、それら構成要素のうちでも特に、本発明にかかるアレイ・コントローラ12を備えている。このアレイ・コントローラ12は、サーバ10のマザーボードに一体化した回路として構成してもよく、或いは、マザーボードとは別体の、回路カードないしサブシステムとして構成してもよい。アレイ・コントローラ12は、その設計次第で、様々な方式のデータ及びデータ転送構成に対応させることができ、例えば、E I S A、P C I、S C S I、I D E、I P I等々に対応可能なものとすることができる。これらのデータ転送方式ないしデータ・バス方式は、アレイ・コントローラ回路の設計の分野では周知のものである。

【0009】

アレイ・コントローラ12は、記憶サブシステム14に接続されている。記憶サブシステム14は、サーバ10の中に構築してもよく、或いは、サーバ10とは別体の装置の中に構築してもよい。記憶サブシステム14は、通常は複数台のハード・ドライブ16で構成される。それらハード・ドライブ16は「ベイ」ないしスロットの中に装備される。図4

10

20

30

40

50

に示した記憶サブシステム 14 は、6つのベイを備えている。従ってこの記憶サブシステム 14 には、最大で6台までの物理ドライブ 16 を装備することができる。物理ドライブであるハード・ドライブ 16 の容量は、一般的に入手可能なドライブ装置の容量でよく、例えば1ギガバイト、4ギガバイト、或いは9ギガバイト等でもよい。

【0010】

アレイ・コントローラ 12 を用いて、記憶サブシステム 14 の全体で1台の論理装置を構成するようにすることも可能であり、或いは、物理ドライブのうちの2台の物理ドライブで1台の論理装置を構成し、残りの4台の物理ドライブでもう1台の論理装置を構成するようにすることも可能である。1台の論理装置（論理ドライブ装置）を構成するのに、1台のハード・ドライブで構成することもあれば、複数台のハード・ドライブで構成することもあり、更には、1台または複数台のハード・ドライブの一部分で構成することもあるが、いずれの場合にも、論理装置は、あたかも1台のハード・ドライブであるかのように取扱われる。

10

【0011】

図5は、1つの記憶サブシステムの説明図である。この図5に示した記憶サブシステム 18 は、本発明の幾つかの特征的局面のうちの、1つを備えたものである。記憶サブシステム A (18) は、7台の物理ドライブで構成された記憶サブシステムであり、それらのうち第1物理ドライブ～第3物理ドライブで第0論理装置 20 が構成されており、また第4物理ドライブ～第7物理ドライブで第1論理装置 22 が構成されている。

【0012】

それら物理ドライブは、ハード・ドライブであってもよく、また、書換可能レーザ・ドライブや、テープ・ドライブ、或いは更にその他の適当な、任意の書換可能な記憶媒体であってもよい。この好適な実施の形態では、それら物理ドライブはハード・ドライブである。

20

【0013】

図5は更に第1物理ドライブ 24 の説明図の部分を含んでおり、これについて以下に説明する。ハード・ドライブの容量は、利用可能なデータ記憶空間の大きさで表される。ハード・ドライブは、セクタと呼ばれる記憶領域に分割することができる。その場合、1台のハード・ドライブに含まれるセクタの総数は、例えば3万セクタほどになることもある。ハード・ドライブにおいては、その記憶領域の一部分を、重要なシステム情報を記憶させるための予約領域とするということも行われている。この好適な実施の形態では、個々のハード・ドライブ 24 の予約領域 26 は、その大きさを1088セクタとしてあり、また、個々のハード・ドライブ 24 の先頭の1088セクタをこの予約領域 26 に割当ててある。予約領域 26 の大きさ及び割当位置はこれらに限られず、様々な大きさとすることができ、ハード・ドライブ上の任意の位置に割当てることができ、更には、そのハード・ドライブとは別体の、データ情報格納用の略々同様の装置に割当てることが可能である。この実施の形態に関しては、予約領域 26 のことを予約情報セクタ (RIS) と呼ぶことにする。

30

【0014】

RIS 26 に、物理ドライブ及び論理ドライブ装置についての構成情報を記憶させるようにしている。構成情報を生成して個々の物理ドライブの RIS 26 に記憶させるのは、即ち格納するのは、アレイ・コントローラ 12 である。RIS 26 には様々な種類の情報が格納され、それら情報には、例えば、パフォーマンス情報、フォールト・トレランス情報、ハードウェア・トポロジー情報、物理ドライブ番号、ドライブ位置番号、論理装置関連情報、それに、個々の物理ドライブがどの論理ドライブ装置に所属するか（1台の物理ドライブが同時に2台以上の論理ドライブ装置に所属することもある）を表す情報等があり、更にその他の情報も含まれる。

40

【0015】

この実施の形態においては、記憶サブシステム 14 に含まれている複数台の物理ドライブは、それら物理ドライブの RIS 26 の内容が、物理ドライブ番号及びドライブ位置番号

50

を表す情報以外は、互いに同一となるように構成されている。

【0016】

本発明の第1の実施の形態では、アレイ・コントローラ12が、記憶サブシステム14に含まれている複数台の物理ドライブの位置が変更された(別のドライブ・ベイへ移動された)場合に、ソフトウェアとユーザとの間の情報交換を必要とすることなく、それら物理ドライブのRIS26を自動的に再構成できるように構成されている。

【0017】

これについて、図5の記憶サブシステムA(18)及び記憶サブシステムA'(28)と、図4とを参照して以下に説明する。ここでは、第4物理ドライブ及び第5物理ドライブを記憶サブシステムA(18)から一旦取外して再装着する際に、物理ドライブ位置を間違えて装着したために、それまでは記憶サブシステムA(18)として構成されていたものが、記憶サブシステムA'(28)になったものとする。この場合、本発明のこの実施の形態にかかるアレイ・コントローラ12は、それら物理ドライブのドライブ位置が、記憶サブシステム(18、28、14)の内部で変化したことを認識する。アレイ・コントローラ12は、この物理ドライブの移動を認識するために、システムの電源投入時に、個々の物理ドライブのRIS26からの読取りを実行する。それらRIS26から、ドライブ位置に対するドライブ番号が変化していることが分かる。続いて、この実施の形態のアレイ・コントローラ12のファームウェアが、記憶サブシステム14の再構成を自動的に実行し、この再構成によって、個々の物理ドライブのRIS26が、ドライブ位置と物理ドライブ番号との関係を正しく表示するようにする。

【0018】

この実施の形態は、物理ドライブ番号が記憶サブシステムの複数のドライブ位置(ベイ)の間でどのように置換された場合でも、記憶サブシステム14に再構成を施すことができる。更に、この実施の形態は、論理装置が、記憶サブシステムの複数台の物理ドライブに対してどのように割当てられた場合でも、記憶サブシステム14に再構成を施すことができるように構成されている。

【0019】

例えば、ある1つの記憶サブシステムが6台の物理ドライブから成り、最初の状態では、それら6台の物理ドライブが、図1と図3を組合せたように構成されていて、第1ドライブ位置~第3ドライブ位置に装着されている第1物理ドライブ~第3物理ドライブで第0論理ドライブ装置が構成され、第4ドライブ位置~第6ドライブ位置に装着されている第4物理ドライブ~第6物理ドライブで第1論理ドライブ装置及び第2論理ドライブ装置が構成されていたものとする。そして、第1論理ドライブ装置及び第2論理ドライブ装置が、第4物理ドライブ~第6物理ドライブを同等に分け合い、更に、以上の6台の物理ドライブのRISの内容が、物理ドライブ番号及びドライブ位置に関する情報を除いて互いに同一であったものとする。

【0020】

更に、この記憶サブシステムの電源が落とされている間に、以上の6台の物理ドライブの位置が置換されたものとする。ここでは、第1、第2、第3物理ドライブが、夫々、第5、第2、第3ドライブ位置へ移動され、一方、第4、第5、第6物理ドライブが、夫々、第4、第6、第1ドライブ位置へ移動されたものとする。記憶サブシステムへ電源が再投入されたならば、この実施の形態におけるアレイ・コントローラは、個々の物理ドライブのRISからの読取りを実行してそれら物理ドライブのドライブ位置が変化しているか否かを判定する。アレイ・コントローラは、それら物理ドライブのRISから、ドライブ位置が置換される前のそれら物理ドライブの位置関係を表す情報を得る。この情報に基づいて、この実施の形態におけるアレイ・コントローラのファームウェアが、記憶サブシステムに含まれるそれらハード・ドライブ(物理ドライブ)の再構成を実行して、ドライブ位置が置換された後の正しい関係の情報をそれら物理ドライブに付与する。即ち、個々の物理ドライブのRIS部分が、アレイ・コントローラによって再構成され(この場合、少なくとも、個々の論理ドライブ装置に対応するドライブ位置を表したテーブルが更新される

10

20

30

40

50

)、この再構成は、ユーザ(ネットワーク管理者)からの指示を必要とすることなく、即ち、ユーザの介在なしに実行される。こうして再構成が施された記憶サブシステムにおいては、第0論理装置が、第5、第2、及び第3ドライブ位置にまたがって存在し、第1論理装置及び第2論理装置が、第4、第6、及び第1ドライブ位置にまたがって存在することになる。

【0021】

本発明の発明者が知る限りでは、現在、多くの記憶サブシステムにおいて、記憶サブシステムの中の複数の物理ドライブ位置にまたがって存在する論理ドライブ装置を構成する際に、その論理ドライブ装置に対応する複数の物理ドライブ位置を昇順で指定することが要求されている。例えば、1つの論理ドライブ装置に対応する複数の物理ドライブ位置を指定する際に、第1ドライブ位置、第3ドライブ位置、第7ドライブ位置という順序で指定することはできても、第3ドライブ位置、第1ドライブ位置、第7ドライブ位置という順序で指定することはできず、なぜならば、後者ではドライブ位置が昇順になっていないからである。そこで、本発明のこの実施の形態の別の特徴的局面として、この実施の形態のアレイ・コントローラのファームウェアが、ドライブ位置が置換されたならば、1台の論理ドライブ装置を構成している複数台の物理ドライブのドライブ位置が昇順になっているか否かを判定するようにしたということがある。そして、それらドライブ位置が昇順になっていなかったならば、この実施の形態では、その順序が許容されないものであることをユーザ(システム管理者)に通知できるようにしている。更に、ソフトウェアがシステム管理者に対して、どのようにすれば状況を回復できるかを提示するようにしてもよい。即ち、コンピュータ・システムがシステム管理者に対して、記憶サブシステムの複数のドライブ位置の間でハード・ドライブをどのように移動させれば、記憶サブシステムの複数のベイ位置に1台の論理装置に対応した複数台の物理ドライブを装着する際にはそれら物理ドライブを昇順に並べなければならないという標準的規約を満足することができるかを提示するようにしてもよい。

【0022】

図6は、この実施の形態にかかるアレイ・コントローラのファームウェアが実行する個々のステップのロジックを示したフローチャートである。先ずステップS10では、記憶サブシステム及び/またはアレイ・コントローラに電源が投入され、アレイ・コントローラが電源投入時自己診断テストを実行する。ステップS12では、アレイ・コントローラが、それが接続されている記憶サブシステムに装着されている個々の物理ドライブの現在ドライブ位置を読取る。続いてステップS14では、アレイ・コントローラが、RIS(個々の物理ドライブの予約領域部分)からの読取りを実行して、記憶サブシステム及び/またはアレイ・コントローラの電源が落とされる前の個々の物理ドライブの位置である旧ドライブ位置を判別する。尚、RISからの読取りと、ドライブ番号のチェックとの間には、その他のプロセスのステップが幾つか存在しており、それらプロセスには、例えば、整合性チェックを実行するプロセスや、カウンタ更新チェックを実行するプロセスが含まれ、更にその他のプロセスも含まれる。

【0023】

ステップS16では、アレイ・コントローラが、第1ドライブ位置についての調査を実行して、この第1ドライブ位置に装着されているハード・ドライブのドライブ番号を調べる。ステップS18では、アレイ・コントローラのファームウェアが、第1ドライブ位置に現在装備されている物理ドライブが、記憶サブシステム及び/またはアレイ・コントローラの電源が落とされる前に第1ドライブ位置に装着されていた物理ドライブと同一のものであるか否かを判定する。即ち、現在ドライブ番号が旧ドライブ番号と同一であったならば、これに続いてステップS20を実行し、このステップでは、現在ドライブ位置(ここでは、第1ドライブ位置)が、ハード・ドライブが装備されているドライブ位置のうちの最後のものか否かをファームウェアが判定する。もしそれが最後のものでなかったならば(即ち、他にも調査すべきドライブが存在していたならば)ファームウェアが、次のドライブ(ここでは、第2ドライブ位置に装着されているドライブ)の調査を開始する。こ

10

20

30

40

50

でも、第1ドライブ位置の場合と同様に、第2ドライブ位置に現在装備されている物理ドライブが、電源が落とされる前に第2ドライブ位置に装着されていた物理ドライブと同一のものであるか否かを判定する。もし仮に、第2ドライブ位置に現在装備されている物理ドライブが、第2ドライブ位置に以前装備されていた物理ドライブと同一のものでなかったならば、全ての物理ドライブのR I Sに対して再構成を施して、この矛盾を解消する必要があり、これをステップS 2 6で実行する。

【0024】

全てのドライブの調査が完了したならば、ファームウェアが、それら全てのディスクのR I S領域部分を更新し、即ちR I Sの内容に再構成を施すことで、記憶サブシステムに含まれている全てのハード・ドライブについて、論理装置と、ドライブ番号と、ドライブ位置との間の正しい関係が表示されるようにする。

10

【0025】

以上の作業は、記憶サブシステム及び/またはアレイ・コントローラの電源が落とされる前に個々のドライブがどのドライブ位置に装着されていたかを判定し、その情報を、記憶サブシステム及び/またはアレイ・コントローラに電源が再投入されたときに個々のドライブが装備されているドライブ位置と比較し、もしドライブ位置に変化が生じていたならばその変化に応じて自動的にドライブを更新し、即ちドライブに再構成を施すという作業であるが、当業者であれば、この作業を達成するための様々なソフトウェア・アルゴリズムを作成することは容易である。そのための技法としては、ドライブ番号、ドライブ通し番号、ドライブ位置、それに論理装置の割当てを表す情報を、アレイ・コントローラに付

20

【0026】

本発明の実施の形態のもう1つの特徴的局面として、記憶サブシステムに1台ないし複数台の新たな論理ドライブ装置が追加されたならば、記憶サブシステム及びそれに付随したアレイ・コントローラに電源が投入されたときに、人手の介入を必要とすることなく、その記憶サブシステムに含まれるドライブに自動的に再構成が施されるということがある。換言すれば、アレイ・コントローラ回路のファームウェアが、記憶サブシステムに新たな論理ドライブ装置が追加されたことを自動的に認識し、そして、記憶サブシステムの再構成を実行する(例えば、構成情報を更新する)ことで、その新たな論理ドライブ装置を組み込むようにしている。本発明のこの特徴的局面を、ボリューム追加といい、また特に、オフライン・ボリューム追加という。

30

【0027】

図7は、以上の実施の形態における現況の組込みを行うという特徴的局面を説明するための図である。アレイ・コントローラA(30)は、記憶サブシステムA(32)に接続されている。記憶サブシステムA(32)は、3台のハード・ドライブ(第1、第2、及び第3ハード・ドライブ)によって、単一の論理ドライブ装置(第0論理ドライブ装置)34を構成するように構成されている。第1、第2、及び第3ハード・ドライブは、夫々、第1、第2、及び第3ベイに装備されている。更に、別の論理ドライブ装置(第0'論理ドライブ装置)36が、記憶サブシステムB(38)の中に構成されている。この第0'論理ドライブ装置36は、第1'ハード・ドライブ及び第2'ハード・ドライブによって構成されている。更に、第1'ハード・ドライブ及び第2'ハード・ドライブは、夫々、記憶サブシステムB(38)の第2ベイと第3ベイとに装備されている。記憶サブシステムB(38)は、アレイ・コントローラBに接続されている。

40

【0028】

ここで例えば、システム管理者が何らかの理由で、第0論理ドライブ装置34と第0'論

50

理ドライブ装置36とをまとめて1つの記憶サブシステムの中に包含させたいと考えたものとする。この場合、本発明のこの実施の形態によれば、システム管理者は単に、記憶サブシステムA(32)の電源を落とし、第0'論理ドライブ装置(36)の全体を、記憶サブシステムA(32)の空きベイへ移動させて装備するだけでよい。そのようにした状態を、図7の下側部分に示してあり、図示の如く、第1'ハード・ドライブと第2'ハード・ドライブとは、記憶サブシステムA(32)の第5ベイと第6ベイとに装備されている。

【0029】

記憶サブシステムA(32)に電源が投入されたならば、アレイ・コントローラA(30)が、記憶サブシステムA(32)に含まれている個々のハード・ドライブの予約領域部分からの読取りを実行し、それによって、この記憶サブシステムA(32)に新たな論理ドライブ装置(ここでは、第0'論理ドライブ装置36)が追加されたと判定する。続いて、アレイ・コントローラA(30)は、この記憶サブシステムA(32)に含まれている個々のハード・ドライブの再構成を実行し、それによって、第0'論理ドライブ装置36の組み込みが完了する。アレイ・コントローラA(30)が実行するこの再構成は、電源投入時に自動的に実行されるものである。

10

【0030】

既に述べたことであるが、記憶サブシステムB(38)から記憶サブシステムA(32)へ第0'論理ドライブ装置36が移設される前の時点では、記憶サブシステムA(32)の中に構成されている第0論理ドライブ装置34に所属している複数台の物理ドライブのRIS部分(「オリジナルRIS部分」)は、物理ドライブ番号及びドライブ位置を表しているデータ以外は、互いに完全に同一になっている。更に、この第0論理ドライブ装置34を構成しているそれら物理ドライブのRIS部分は、記憶サブシステムA(32)の第4、第5、第6ベイには物理ドライブが装備されていないということを表示している。

20

【0031】

そして、第0'論理ドライブ装置36の全体が記憶サブシステムA(32)に追加されて、記憶サブシステムA(32)に電源が投入されたならば、アレイ・コントローラA(30)が、記憶サブシステムA(32)に含まれている個々の物理ドライブのRIS部分からの読取りを実行し、そして、それら物理ドライブのRIS部分に含まれているデータのうち、そのRISの書込みを行ったアレイ・コントローラのIDを表すデータ、ドライブ位置のデータ、論理ドライブ装置番号のデータ、及び物理ドライブ番号のデータのうちの何れか1つを、当該物理ドライブの、現在ドライブ位置及び物理ドライブ番号を表しているデータと比較する。この比較及び検証の作業が完了したならば、その時点で、アレイ・コントローラA(30)は、記憶サブシステムA(32)に含まれている全ての物理ドライブに対して適切に再構成を施すことができる状態になっており、その再構成によって、現在ドライブ位置、物理ドライブ番号、及び当該物理ドライブが所属する論理ドライブ装置が表示されるようになる。

30

【0032】

容易に理解されるように、本発明のこの特徴的局面が機能するのは、記憶サブシステムに、論理ドライブ装置の全体が追加された場合である。また、本発明のこの特徴的局面が良好に機能するのは、その論理ドライブ装置が追加されることによって、その論理ドライブ装置に所属する物理ドライブと、その記憶サブシステムに元から存在していた物理ドライブとが、混成して配列される場合である。従って本発明は、オフライン・ドライブ移動と、オフライン論理ドライブ装置追加との両方を、システムへの電源投入時に自動的に実行するものであり、その際にデータが失われることがなく、またその際に人手の介入を必要としない。

40

【0033】

本明細書において先に述べたように、本発明の発明者が知る限り、現在、多くの記憶サブシステムにおいて、1台の論理ドライブ装置に所属する複数台のハード・ドライブ(物理ドライブ)を記憶サブシステムの複数のベイに装備する際には、昇順で装備することが要

50

求されている。これに関して発明者が特に述べておきたいことは、そのような複数台の物理ドライブを昇順で並べねばならないという要求それ自体は本発明にとって必要なものではないが、1台の論理ドライブ装置を所属する複数台の物理ドライブが記憶サブシステムの複数のベイ位置に昇順に並べられていることをチェックして確かめる機能をアレイ・コントローラ回路のファームウェアに組み込んでおいてもよいということである。

【0034】

図8は、オフライン・ボリューム追加の方法の一例を示したフローチャートである。ステップS100では、電源投入時自己診断テストが実行される。この電源投入時自己診断テストの実行中に、全てのハード・ドライブの予約領域部分からの読取りが行われ、そして様々な判定が実行される。電源投入時自己診断テストの実行中に完了されるそれら作業のうちには、コントローラIDデータが、現在、物理ドライブからの読取りを実行しているアレイ・コントローラによって書込まれたものか否かを判定する作業がある。この情報を利用して、記憶サブシステムに新たな論理ドライブ装置が追加されたか否かを容易に判定することができる。電源投入時自己診断テストの実行中には更に、複数台のハード・ドライブの予約領域部分のうちどの部分が互いに同一でどの部分が互いに異なっているかも調べられる。

【0035】

ステップS112では、アレイ・コントローラがオフライン・ドライブ移動に関する機能を実行し、この作業は先に説明したようにして行われる。即ち、この時点で、ファームウェアが、複数台のハード・ドライブのうち、どのドライブが「オリジナルの」構成の一部を構成するドライブであるか、また、どのドライブが新たに追加されたドライブであるかを判定する。この判定を行うために、オリジナルの構成を表す情報のコピーを、アレイ・コントローラに付設されているメモリに格納するようにしている。このメモリを「グローバルRIS」と呼ぶことにする。

【0036】

ステップS114では、ファームウェア及びハードウェアが、第1ドライブ位置のドライブの調査を実行する。ここでは、第1ドライブ位置のドライブの予約領域部分(RIS)の有効性検証を実行する。RISの有効性検証の方法には様々なものがあるが、その一例は、当該RISの内容が、アレイの電源が落とされる前に現在アレイ・コントローラが書込んだRISの内容と同一であったならば(ただしドライブ位置及び物理ドライブ番号(ドライブID)を除く)、そのRISが有効であると判定するという方法である。またそのためには、例えば、当該RISをグローバルRISと比較して、両者が同一であるか否かをチェックするようにすればよい。第1ドライブ位置のハード・ドライブ(物理ドライブ)のRISが有効であったならば、システムは続いてステップS118において、そのハード・ドライブ(物理ドライブ)が、記憶サブシステムに装備されている物理ドライブのうち最後のものか否かを判定する。もしそのハード・ドライブが、記憶サブシステムの物理ドライブのうち最後のものではなかったならば、ステップS120においてファームウェアが、次のドライブ位置にあるドライブの調査を実行する。続いてステップS116において、そのドライブのRISの有効性検証を実行し、以下、最後のドライブに至るまでこれを反復する。

【0037】

ステップS116における有効性検証は、調査対象のハード・ドライブのRISが有効RISであるか否かを判定するものである。これは、アレイ・コントローラの電源が落とされた時点ではそのハード・ドライブが存在していなかったということが、そのハード・ドライブの予約領域部分の情報から判明した場合である。そのような情報は、次のようなことを表す情報であり、それは、当該RISとグローバルRISとが一致していないこと、或いは、当該RISに、現在アレイ・コントローラとは異なったアレイ・コントローラのIDないしシグニチャが書込まれていること、或いは、記憶サブシステムに含まれているその他のハード・ドライブのRISの中に、当該RISの論理ドライブ装置のボリューム番号が存在していないこと、或いは、当該RISが、記憶サブシステムのその他のベイに

10

20

30

40

50

装着されているハード・ドライブのR I Sと、ある重要な点において異なっていることである。

【 0 0 3 8 】

当該ハード・ドライブのR I Sが無効であると判定された場合には、続いてステップS 1 2 2において、そのハード・ドライブのR I Sの内容を再び読取る。ここで読取った内容は、例えばアレイ・コントローラに付属するローカル・メモリに格納しておけばよい。続いてステップS 1 2 4では、そのR I Sを、それまでに格納しておいた複数のR I Sの夫々と比較する。この比較ステップは、現在ハード・ドライブが、それまでに格納しておいた複数のR I Sのうちのいずれか1つに対応した論理ドライブ装置に所属するものか、それとも、新たな論理ドライブ装置に所属するものかを判定するためのステップである。

10

【 0 0 3 9 】

ステップS 1 2 6では、論理ドライブ装置が完全なものか（即ち、論理ドライブ装置の全体が揃っているか）否かをシステムが判定する。論理ドライブ装置が完全なものか否かを判定するための方法の一例は、その論理ドライブ装置に所属する個々の物理ドライブのR I S部分を調査するというものである。R I Sは、その他のドライブのうちどのドライブがその論理ドライブと同じ論理ドライブ装置に所属しているのかを表示しており、また、その論理ドライブ装置に所属する複数の物理ドライブどうしの間正しい順序も表示している。その論理ドライブ装置が完全なものであったならば（即ち、その論理ドライブ装置に所属する全ての物理ドライブが存在していたならば）、ステップS 1 2 8において、その完全な論理ドライブ装置に関する情報を、新たな有効R I Sの中に組み込む。この新たな有効R I Sは、ボリューム追加の判定に関する全作業が完了したときに（これについては以下に説明する）、記憶サブシステムに含まれている全てのドライブにコピーされる。その後、ステップS 1 1 8において、調査対象のドライブが、最後のドライブか否かを判定する。

20

【 0 0 4 0 】

一方、ステップS 1 2 6において、論理ドライブ装置（ボリューム）が完全なものではないと判定されたならば、プログラムはそこから直接ステップS 1 1 8へ進み、そこで、調査対象の物理ドライブが、記憶サブシステムに含まれている最後の物理ドライブか否かを判定する。

【 0 0 4 1 】

プログラムによる以上の作業が、記憶サブシステムに含まれている全ての物理ドライブについて完了したならば、続いてこのプログラムは、ステップS 1 3 0へ進み、ここでは、記憶サブシステムに装備されている全てのハード・ドライブのR I Sが更新されて新たな有効R I Sとされる。この新たな有効R I Sには、追加されたことが判明した論理ドライブ装置や、記憶サブシステムの電源が落とされる前に存在していた論理ドライブ装置が、全てが含まれている。

30

【 0 0 4 2 】

尚、図8に示した以上のフローチャートは、記憶サブシステムに論理ドライブ装置（ボリューム）が追加されたか否かを判定するための様々な技法の一例として、ファームウェアによる技法を示したものである。このフローチャートに示した技法に対しては、小さな変更であるか大きな変更であるかを問わず、様々な変更を施すことができ、それによって、ボリューム追加が行われたことを自動的に判定して関連データの更新を行うための、以上に説明した機能と略々同一の機能を達成することのできる、様々な方法が可能となる。

40

【 0 0 4 3 】

【 発明の効果 】

以上から明らかなように、本発明の1つの具体的な実施の形態によれば、記憶サブシステムの電源が断たれている間に、その記憶サブシステム内の複数台のハード・ドライブ（物理ドライブ）のドライブ位置が変更（即ち「置換」）されたり、その記憶サブシステムに新たな完全なボリュームが追加されたならば、その記憶サブシステムに電源が再投入されたときに、その記憶サブシステムにドライブ移動やボリューム追加という変更が発生した

50

ということがその記憶サブシステムに付随するアレイ・コントローラによって自動的に判定されるようにすることができる。

【0044】

また、本発明の別の具体的な実施の形態によれば、論理ドライブ装置を、その論理ドライブ装置に所属する複数台の物理ドライブが正しく順序付けられた状態で、追加することが可能となる。即ち、この実施の形態では、論理ドライブ装置が追加されたか否かが判定されると共に、複数台の物理ドライブが正しい順序に並べられているか否かも判定される。

【0045】

以上に本発明の好適な実施の形態を幾つか図示して説明したが、当業者には容易に理解されるように、それら実施の形態に対しては本発明の概念及び範囲から逸脱することなく様々な改変を加えることができ、また、本発明の範囲は請求項に記載したとおりのものである。

10

【図面の簡単な説明】

【図1】一般的なドライブ・アレイの一例を示した図である。

【図2】一般的なドライブ・アレイの別の一例を示した図である。

【図3】一般的なドライブ・アレイの更に別の一例を示した図である。

【図4】ドライブ・アレイに接続されたアレイ・コントローラの一例を示した図である。

【図5】記憶サブシステムの説明図である。

【図6】本発明の1つの実施の形態におけるオフライン・ドライブ移動を表したフローチャートである。

20

【図7】本発明の1つの実施の形態にかかるオフライン・ボリューム移動を説明するためのブロック図である。

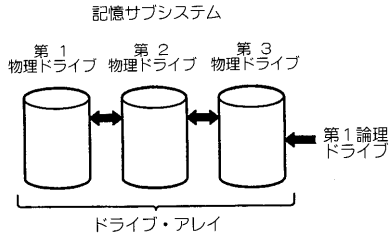
【図8】本発明の1つの実施の形態にかかるオフライン・ボリューム追加を表したフローチャートである。

【符号の説明】

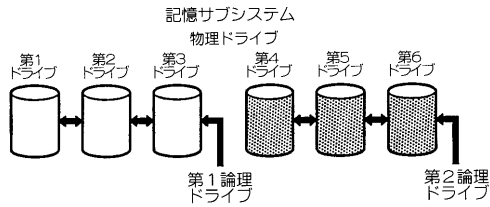
- 10 サーバ
- 12 アレイ・コントローラ
- 14 記憶サブシステム
- 16 ハード・ドライブ
- 18 記憶サブシステム
- 20、22 論理ドライブ装置
- 26 R I S
- 28 記憶サブシステム
- 30 アレイ・コントローラ
- 32 記憶サブシステム
- 34、36 論理ドライブ装置
- 38 記憶サブシステム

30

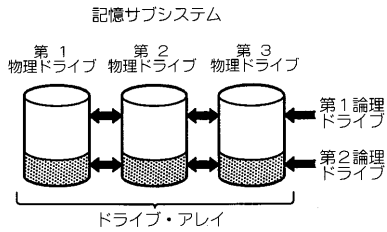
【図1】



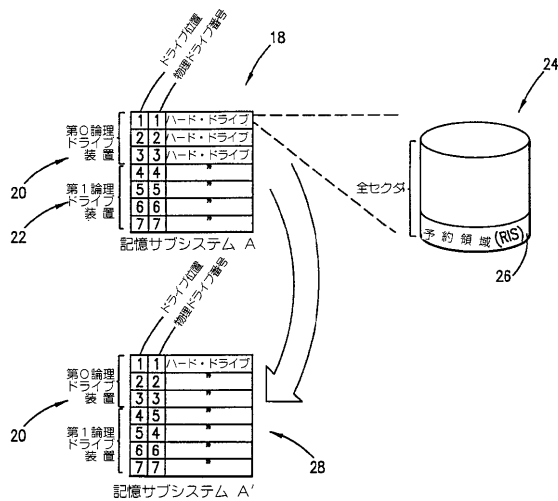
【図2】



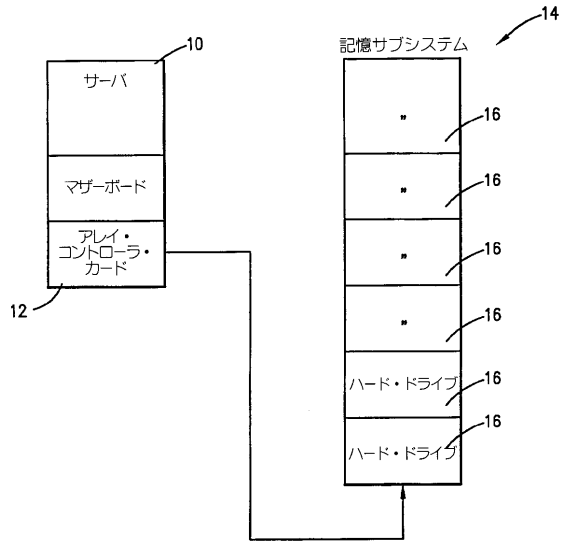
【図3】



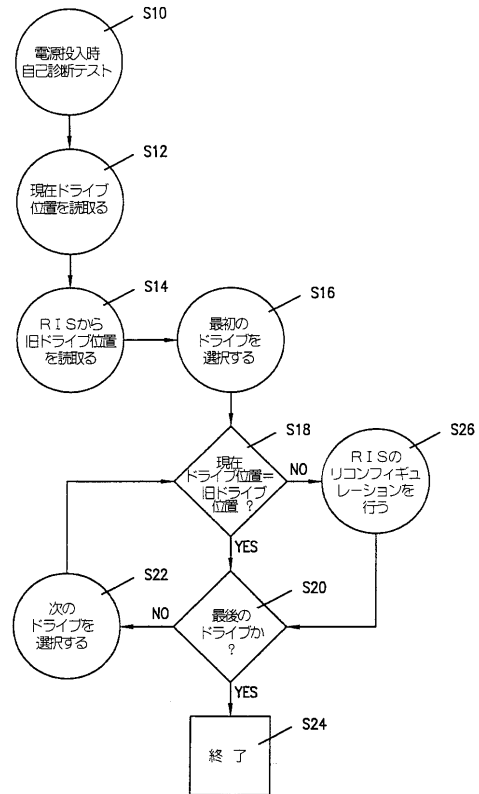
【図5】



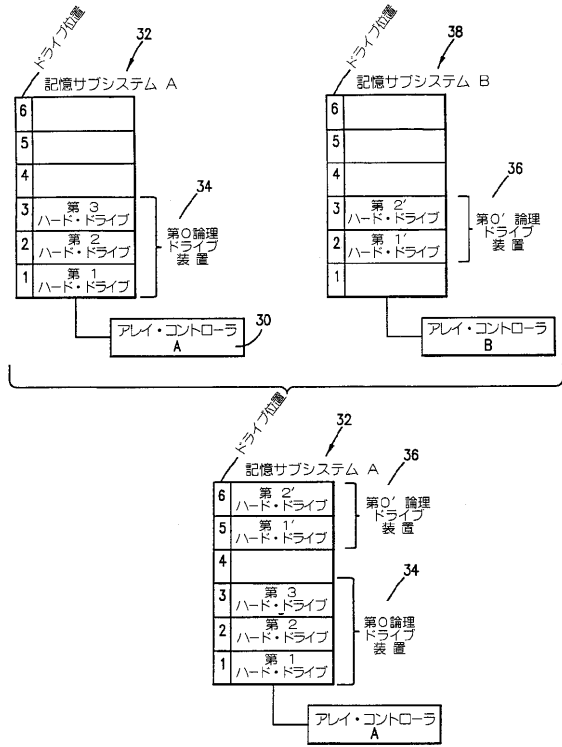
【図4】



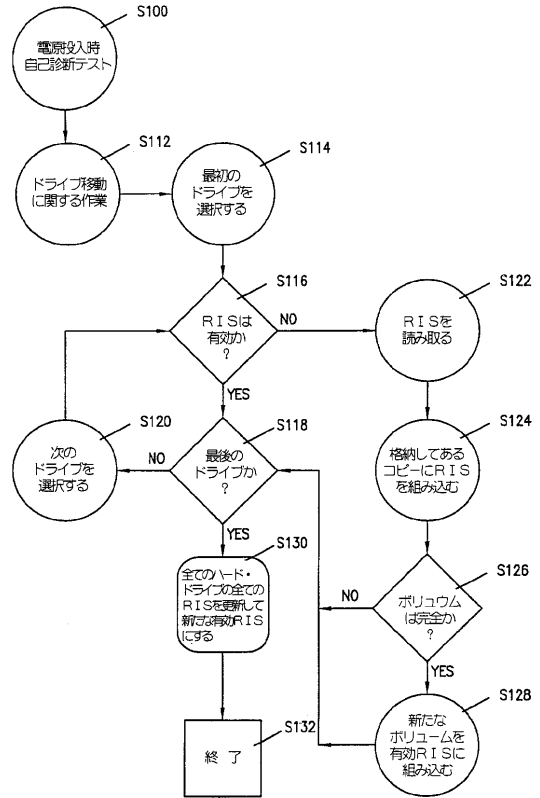
【図6】



【図7】



【図8】



フロントページの続き

- (74)代理人 100071124
弁理士 今井 庄亮
- (74)代理人 100076691
弁理士 増井 忠式
- (74)代理人 100075270
弁理士 小林 泰
- (74)代理人 100080137
弁理士 千葉 昭男
- (72)発明者 パルナ・シー・マーシー
アメリカ合衆国テキサス州77084, ヒューストン, グリーン・ストーン・コート 6710
- (72)発明者 マーク・ジェイ・トンプソン
アメリカ合衆国テキサス州77379, スプリング, パイン・ディストル・レーン 8206

審査官 梅景 篤

- (56)参考文献 特開平04-049413(JP, A)
特開平05-210464(JP, A)
特開平08-221219(JP, A)
特開平05-334010(JP, A)
特開平07-225652(JP, A)
特開昭63-098720(JP, A)

- (58)調査した分野(Int.Cl., DB名)
G06F 3/06