

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4202026号  
(P4202026)

(45) 発行日 平成20年12月24日(2008.12.24)

(24) 登録日 平成20年10月17日(2008.10.17)

(51) Int.Cl.		F I			
<b>G06F 3/06</b>	<b>(2006.01)</b>	G06F	3/06	304F	
<b>G06F 12/00</b>	<b>(2006.01)</b>	G06F	3/06	304B	
		G06F	12/00	514E	
		G06F	12/00	531D	
		G06F	12/00	545A	

請求項の数 14 (全 20 頁)

(21) 出願番号	特願2002-23604 (P2002-23604)	(73) 特許権者	502038668
(22) 出願日	平成14年1月31日(2002.1.31)		株式会社アンソナ
(65) 公開番号	特開2003-223286 (P2003-223286A)		神奈川県川崎市高津区梶ヶ谷5-9-23
(43) 公開日	平成15年8月8日(2003.8.8)	(74) 代理人	100109081
審査請求日	平成16年4月15日(2004.4.15)		弁理士 三木 友由
前置審査		(72) 発明者	松葉 知則
			神奈川県川崎市高津区梶ヶ谷5-9-23
			株式会社アンソナ内
		(72) 発明者	佐藤 創
			神奈川県川崎市高津区梶ヶ谷5-9-23
			株式会社アンソナ内
		審査官	梅景 篤

最終頁に続く

(54) 【発明の名称】 ストレージシステムおよびストレージ装置

(57) 【特許請求の範囲】

【請求項1】

複数のストレージ装置をネットワークに接続して構成するストレージシステムであって、各ストレージ装置は、

ストレージ領域を複数のコピーユニットに分割した格納部と、

データを記録したコピーユニットが正常に存在していることを示す信号を送信する状況送信部と、

他のストレージ装置においてデータを記録したコピーユニットが正常に存在していることを示す信号を当該他のストレージ装置から受けて、他のストレージ装置におけるストレージ状況を監視し、システム全体におけるコピーユニットの多重度を検出して、検出した多重度と、予め設定されている必要な多重度とを比較する監視部と、

前記監視部において検出した多重度が、必要な多重度よりも小さいことが検知されると、他のストレージ装置に対するコピーユニットの複製要求を生成する複製要求生成部と、を備えることを特徴とするストレージシステム。

【請求項2】

前記複製要求生成部は、ストレージ装置がネットワークから切断される前に生成される複製コマンドを受けて、前記複製要求を生成することを特徴とする請求項1に記載のストレージシステム。

【請求項3】

前記ストレージ装置は、

データの書込要求を受け付ける要求受付部と、  
前記コピーユニットの空き容量に応じて前記書込要求に应答する应答部と、  
前記コピーユニットにデータを書き込む書込制御部とを備えることを特徴とする請求項  
1または2に記載のストレージシステム。

【請求項4】

前記ストレージ装置は、

前記コピーユニットの空き容量を検出する検出部と、  
その空き容量がデータの書込みに十分であるか否かを判断する容量判断部とを備えるこ  
とを特徴とする請求項3に記載のストレージシステム。

【請求項5】

前記ストレージ装置は、

自己が管理するコピーユニット内のデータの送信要求があった場合に、該コピーユニッ  
トと同一内容のコピーユニットを有する他のストレージ装置と協同してデータを送信する  
ようにデータを読み出す読出制御部を備えることを特徴とする請求項1から4のいずれか  
に記載のストレージシステム。

【請求項6】

ネットワークに接続してデータを格納するストレージ装置であって、

ストレージ領域を複数のコピーユニットに分割した格納部と、  
データを記録したコピーユニットが正常に存在していることを示す信号を送信する状況  
送信部と、

他のストレージ装置においてデータを記録したコピーユニットが正常に存在しているこ  
とを示す信号を当該他のストレージ装置から受けて、他のストレージ装置におけるストレ  
ージ状況を監視し、コピーユニットの多重度を検出して、検出した多重度と、予め設定さ  
れている必要な多重度とを比較する監視部と、

前記監視部において検出した多重度が、必要な多重度よりも小さいことが検知されると  
、他のストレージ装置に対するコピーユニットの複製要求を生成する複製要求生成部と、  
を備えることを特徴とするストレージ装置。

【請求項7】

前記複製要求生成部は、該ストレージ装置がネットワークから切断される前に生成される  
複製コマンドを受けて、前記複製要求を生成することを特徴とする請求項6に記載のスト  
レージ装置。

【請求項8】

データの書込要求を受け付ける要求受付部と、

前記コピーユニットの空き容量に応じて前記書込要求に应答する应答部と、  
前記コピーユニットにデータを書き込む書込制御部とを備えることを特徴とする請求項  
6または7に記載のストレージ装置。

【請求項9】

前記コピーユニットの空き容量を検出する検出部と、

その空き容量がデータの書込みに十分であるか否かを判断する容量判断部とを備えるこ  
とを特徴とする請求項8に記載のストレージ装置。

【請求項10】

自己が管理するコピーユニット内のデータの送信要求があった場合に、該コピーユニッ  
トと同一内容のコピーユニットを有する他のストレージ装置と協同してデータを送信するよ  
うにデータを読み出す読出制御部を備えることを特徴とする請求項6から9のいずれかに  
記載のストレージ装置。

【請求項11】

個々のコピーユニットを管理する複数のコピーユニットエージェントと、  
ユーザの識別情報に基づいて、複数のコピーユニットにまたがる1つ又は複数のファイ  
ルデータをユーザ毎に管理するファイルシステムエージェントと、  
を備えることを特徴とする請求項6から10のいずれかに記載のストレージ装置。

10

20

30

40

50

## 【請求項 1 2】

前記ファイルシステムエージェントは、コピーユニットを他のストレージ装置から探し出し、ユーザのファイルデータを記録するために利用することを可能とする請求項 1 1 に記載のストレージ装置。

## 【請求項 1 3】

コンピュータに、

ストレージ領域を複数のコピーユニットに分割する機能と、

データを記録したコピーユニットが正常に存在していることを示す信号をネットワークに送信する機能と、

他のストレージ装置においてデータを記録したコピーユニットが正常に存在していることを示す信号を当該他のストレージ装置から受けて、他のストレージ装置におけるストレージ状況を監視し、コピーユニットの多重度を検出して、検出した多重度と、予め設定されている必要な多重度とを比較する機能と、

検出した多重度が、必要な多重度よりも小さいことが検知されると、他のストレージ装置に対するコピーユニットの複製要求を生成する機能とを実現させて、ストレージ装置として機能させることを特徴とするプログラム。

## 【請求項 1 4】

複数のストレージ群を接続して構成するストレージシステムであって、各ストレージ群は、ゲートウェイ装置と、該ゲートウェイ装置にネットワーク接続される複数のストレージ装置とを備え、

各ストレージ装置は、

ストレージ領域を複数のコピーユニットに分割した格納部と、

データを記録したコピーユニットが正常に存在していることを示す信号を送信する状況送信部と、

他のストレージ装置においてデータを記録したコピーユニットが正常に存在していることを示す信号を当該他のストレージ装置から受けて、他のストレージ装置におけるストレージ状況を監視し、コピーユニットの多重度を検出して、検出した多重度と、予め設定されている必要な多重度とを比較する監視部と、

前記監視部において検出した多重度が、必要な多重度よりも小さいことが検知されると、他のストレージ装置に対するコピーユニットの複製要求を生成する複製要求生成部と、を有し、

少なくとも 2 つのストレージ群で、同一内容のデータを保持することを特徴とするストレージシステム。

## 【発明の詳細な説明】

## 【0001】

## 【発明の属する技術分野】

本発明は、データを記録するストレージ装置およびストレージシステムに関する。とくに本発明は、データに関する管理を自律的に行うことのできるストレージ装置およびその装置を利用するストレージシステムに関する。

## 【0002】

## 【従来の技術】

業務上発生する様々な情報は、ハードディスク等のストレージ装置に電子的なデータとして記録され管理される。大容量のデータを記録するために、複数のハードディスクを連携させて仮想的な大容量ドライブを構築する R A I D (Redundant Arrays of Independent Disks) と呼ばれるストレージ管理技術が開発されている。

## 【0003】

特に近年、コンピュータの処理速度の向上により、非常に複雑な演算処理が可能となってきたが、同時に企業において管理すべきデータ量は急速に増加する傾向にある。また映像のデジタル化に伴い、今後は家庭においても映画や TV 番組をハードディスクに保存することが一般的になると予想される。そのため、大量のデータを効率的に管理するストレ

10

20

30

40

50

ジシステムへの関心が高まり、現状ではS A N (Storage Area Network) やN A S (Network Attached Storage) に代表されるようにデータを統合化して一元管理する手法が盛んに研究されている。S A NやN A Sは既存のR A I D技術と組み合わせて利用することも可能であり、ストレージ管理技術の一つの方向性を示すものである。

【0004】

【発明が解決しようとする課題】

しかしながら、S A NやN A Sを利用してデータを統合管理するためには、複数のストレージ装置を統括する管理サーバが必須の構成となる。この管理サーバは、データ統合管理システムの中核をなしており、ストレージ装置の故障には対応することができるが、管理サーバ自身が故障するとシステム全体が機能不能な状態となる危険性がある。特に昨今のコンピュータウイルスが蔓延している状況下では、管理サーバのセキュリティ確保などメンテナンス作業に多くの労力を費やす必要があり、そのため管理コストも大きくなる。また、ストレージ装置を増設する場合には管理サーバに対してユーザが登録作業を行う必要があるが、それにはある程度の専門的な知識が要求される場所、そのような知識を有する者を多くの企業が雇用できるわけではない。

10

【0005】

従来のストレージシステムは、管理サーバの能力および機能に大きく依存しているため、管理コストが増加することや、ストレージ装置の増設に柔軟に対応できないという問題があった。そこで、本発明は、上記の課題を解決することのできるストレージ装置およびストレージシステムを提供することを目的とする。これらの目的は特許請求の範囲における独立項に記載の特徴の組合せにより達成される。また従属項は本発明の更なる有利な具体例を規定する。

20

【0006】

【課題を解決するための手段】

上記目的を達成するために、本発明の第1の形態に係るストレージシステムは、複数のストレージ装置をネットワークに接続して構成され、各ストレージ装置は、ストレージ領域を複数のコピーユニットに分割した格納部と、自身のストレージ状況を示す信号を送信する状況送信部と、他のストレージ装置におけるストレージ状況を監視する監視部と、監視結果に基づいて、他のストレージ装置に対するコピーユニットの複製要求を生成する複製要求生成部とを備える。他のストレージ装置のストレージ状況を監視することによって、他のストレージ装置におけるコピーユニットに障害が発生すると、ストレージ装置が自律的にその状態を検知してコピーユニットの複製を行うことを可能とする。ここで、コピーユニットの複製とは、コピーユニットに記録されているデータを、別のコピーユニットにそのまま複製することを意味する。概念的には、コピーユニットをデータを受け入れる容器として捉えることも可能であるが、本明細書では前述のとおりコピーユニットの複製という表現を、その容器内のデータをそのまま別の容器に複製するという意味で用いる。コピーユニットを他のストレージ装置に複製することによって、同一のデータを記録したコピーユニットが生成されることになる。以下では、同一のデータを記録したコピーユニットを、同一内容のコピーユニットと表現することもある。

30

【0007】

監視部は、システム全体におけるコピーユニットの多重度を監視する。複製要求生成部は、監視部においてコピーユニットの多重度が不足していることが検知されると、複製要求を生成する。このストレージシステムにおいては、同一内容のデータを格納する2つ以上のコピーユニットが存在しており、この同一内容のコピーユニットの数を「多重度」と呼ぶ。コピーユニットは必要な多重度を定められており、監視した多重度がこの必要な多重度よりも小さい場合、コピーユニットの多重度が不足していることが検知される。複製要求生成部は、ストレージ装置がネットワークから切断される前に生成される複製コマンドを受けて複製要求を生成してもよい。

40

【0008】

また、第1の形態に係るストレージシステムにおいて、ストレージ装置は、データの書込

50

要求を受け付ける要求受付部と、コピーユニットの空き容量に応じて書込要求に応答する応答部と、コピーユニットにデータを書き込む書込制御部とを備えてもよい。またストレージ装置は、コピーユニットの空き容量を検出する検出部と、その空き容量がデータの書込みに十分であるか否かを判断する容量判断部とを備えてもよい。ストレージ装置は、自己が管理するコピーユニット内のデータに対する送信要求を外部から受けた場合に、該コピーユニットと同一内容のコピーユニットを有する他のストレージ装置と協同してデータを送信するようにデータを読み出す読出制御部を有してもよい。

【0009】

本発明の第2の形態に係るストレージ装置は、ネットワークに接続してデータを格納するものであって、ストレージ領域を複数のコピーユニットに分割した格納部と、自身のストレージ状況を示す信号をネットワークに送信する状況送信部と、他のストレージ装置におけるストレージ状況を監視する監視部と、監視結果に基づいて、他のストレージ装置に対するコピーユニットの複製要求を生成する複製要求生成部とを備える。他のストレージ装置のストレージ状況を監視することによって、他のストレージ装置におけるコピーユニットに障害が発生すると、ストレージ装置が自律的にその状態を検知してコピーユニットの複製を行うことを可能とする。

10

【0010】

監視部は、コピーユニットの多重度を監視する。複製要求生成部は、監視部においてコピーユニットの多重度が不足していることが検知されると、複製要求を生成してもよい。また複製要求生成部は、該ストレージ装置がネットワークから切断される前に生成される複製コマンドを受けて、複製要求を生成してもよい。

20

【0011】

第2の形態に係るストレージ装置は、データの書込要求を受け付ける要求受付部と、コピーユニットの空き容量に応じて書込要求に応答する応答部と、コピーユニットにデータを書き込む書込制御部とをさらに備えてもよい。また、コピーユニットの空き容量を検出する検出部と、その空き容量がデータの書込みに十分であるか否かを判断する容量判断部とを備えてもよい。さらにストレージ装置は、自己が管理するコピーユニット内のデータに対する送信要求を外部から受けた場合に、該コピーユニットと同一内容のコピーユニットを有する他のストレージ装置と協同してデータを送信するようにデータを読み出す読出制御部を有してもよい。

30

【0012】

本発明の第3の形態に係るプログラムは、コンピュータに、ストレージ領域を複数のコピーユニットに分割する機能と、自身のストレージ状況を示す信号をネットワークに送信する機能と、他のストレージ装置におけるストレージ状況を監視する機能と、監視結果に基づいて、他のストレージ装置に対するコピーユニットの複製要求を生成する機能とを実現させて、ストレージ装置として機能させることを特徴とする。このプログラムをハードディスク等の格納部を備えたコンピュータにインストールすることにより、コンピュータをストレージ装置として利用することが可能となる。

【0013】

本発明の第4の形態に係るストレージシステムは、複数のストレージ群を接続して構成され、各ストレージ群は、ゲートウェイ装置と、該ゲートウェイ装置にネットワーク接続される複数のストレージ装置とを備え、ストレージ装置は、他のストレージ装置のストレージ状況を監視して、自律的にデータに関する管理を行う。複数のストレージ装置が接続されたLAN同士をさらに接続することによって、異なる場所にストレージ群を分散したストレージシステムを実現することが可能となる。また、LAN同士を接続することにより、1つのLANにおいて接続可能なストレージ装置の制約を超えて、システムにおける容量上の制約を無くすることができる。このストレージシステムにおいては、少なくとも2つのストレージ群で、同一内容のデータを保持することが好ましい。同一内容の複数のデータの格納場所を空間的に離し、災害などの不慮の事態によって同一内容のデータの全てが一度に消失する確率を低減することができる。

40

50

## 【 0 0 1 4 】

本発明の第5の形態に係るストレージシステムは、ネットワークに接続された複数のストレージ装置を備え、2以上のストレージ装置が同一内容のコピーユニットを有しており、各ストレージ装置は、自身の管理するコピーユニットの多重度をシステム全体として所定の値以上に保つように自律的に他のストレージ装置を監視し、あるコピーユニットの多重度が不足する場合には他のストレージ装置に対して該コピーユニットの複製要求を行うことを特徴とする。

## 【 0 0 1 5 】

本発明の第6の形態に係るストレージシステムは、複数のストレージ装置をネットワークに接続して構成され、各ストレージ装置は、自身のストレージ状況を示す信号をネットワーク上にブロードキャストする状況送信部と、他のストレージ装置からブロードキャストされるストレージ状況を示す信号を監視する監視部と、この監視結果に応じて、他のストレージ装置に対する要求を生成する要求生成部とを備え、各ストレージ装置は記録したデータに関する管理を自律的に行い、システム全体を管理する管理サーバを不要としたことを特徴とする。ブロードキャストとは、ネットワークに接続されている複数の端末に対して送信する方式であり、信号を一斉に複数の端末に送信する場合や、複数の端末に順次スイッチングして送信する場合の双方を含む。各ストレージ装置は、データの多重度を監視して、他のストレージ装置において障害が発生している場合には、そのデータの複製要求を生成するなどデータに関する管理を行うことができる。各ストレージ装置が自律的に機能することによって管理サーバを不要とし、システムの管理コストを小さくすることが可能となる。

## 【 0 0 1 6 】

本発明の第7の形態に係るストレージシステムは、複数のストレージ装置をネットワークに接続して構成され、各ストレージ装置は、ストレージ領域を複数のコピーユニットに分割し、個々のコピーユニットを管理する複数のコピーユニットエージェントと、ユーザの識別情報に基づいて、複数のコピーユニットにまたがる1つ又は複数のファイルデータをユーザ毎に管理するファイルシステムエージェントとを備える。このファイルシステムエージェントは、ユーザのコピーユニットエージェントを管理することによって、コピーユニットに含まれるファイルデータをユーザ毎に管理することを可能とする。コピーユニットエージェントが多重度検知などコピーユニットの自律的な管理を行うことにより、管理サーバを不要とするストレージシステムを実現することができる。ファイルシステムエージェントは、コピーユニットをユーザに割り当てる際に、コピーユニットを自身のストレージ装置からだけでなく、他のストレージ装置から探し出し、ユーザのファイルデータを記録するために利用することを可能とする。

## 【 0 0 1 7 】

なお、上記の発明の概要は、本発明の必要な特徴の全てを列挙したものではなく、これらの特徴の組合わせも又発明となりうる。また、本発明の表現を装置、方法、システム、コンピュータプログラムの間で変換したのもまた、本発明の態様として有効である。

## 【 0 0 1 8 】

## 【 発明の実施の形態 】

以下、発明の実施の形態を通じて本発明を説明するが、以下の実施形態は特許請求の範囲に係る発明を限定するものではなく、又実施形態の中で説明されている特徴の組合せの全てが発明の解決手段に必須であるとは限らない。

## 【 0 0 1 9 】

図1は、本発明の第1の実施形態に係るストレージシステム100の構成図である。ストレージシステム100は、複数のストレージ装置10a~10d(以下、符号「10」で代表させることもある)をネットワーク30に接続することにより構成される。ネットワーク30は、LAN(local Area Network)やWAN(Wide Area Network)であってもよく、それぞれのストレージ装置10a~10dは、有線または無線で接続される。外部装置20a、20bは、ストレージシステム100に対してデータの書込みや読出しを要

10

20

30

40

50

求する外部端末である。ストレージ装置10a~10dおよび外部装置20a~20bの通信は、TCP/IPなどの各種プロトコルを用いて実現される。このストレージ装置10は、ハードウェアコンポーネントでいえば、任意のコンピュータのCPU、メモリ、メモリにロードされたプログラムなどによって実現されるが、ここではそれらの連携によって実現される機能ブロックを描いている。したがって、これらの機能ブロックがハードウェアのみ、ソフトウェアのみ、またはそれらの組合せによっていろいろな形で実現できることは、当業者には理解されるところである。

#### 【0020】

ストレージ装置10は、必要なプログラムを起動するプロセッサと、ハードディスクや不揮発性メモリなどの大容量格納部を備えている。このストレージ装置10は、ハードディスク等を組み込んだ市販の汎用コンピュータの形態であってもよく、またこのストレージシステム100の実現に必要な機能のみを持った専用機器の形態であってもよい。汎用コンピュータに対しては、ユーザが必要なプログラムをインストールすることにより、ストレージ装置10としての機能を持たせることができる。

10

#### 【0021】

このストレージシステム100は、ストレージ装置10a~10dを統括管理する管理サーバを有しない。各ストレージ装置10は、ネットワーク30に接続されている他のストレージ装置の状態を監視し、記録したデータに関する管理を自律的に行う機能を有する。ストレージシステム100のストレージ容量を増やすためには、新たなストレージ装置10をネットワーク30に追加接続するだけでよく、SANやNASなどのシステムにおいてストレージの増設時に必要であった管理サーバへの登録作業を不要とする。またストレージ装置10は自己が管理するデータの複製機能を有しており、ストレージ装置10を交換する場合には、ネットワークから切断する前にこの複製機能を利用して他のストレージ装置10にデータを自律的に複製することも可能とする。

20

#### 【0022】

図2は、ストレージ装置10の概念的な構成を示すブロック図である。ストレージ装置10は、ファイルシステムエージェント200、コピーユニットエージェント300a~300n(以下、符号「300」で代表させることもある)および格納部400を備える。格納部400は、データを記録するストレージ領域を有しており、このストレージ領域は複数の記録単位に分割されている。以下では、この記録単位のことを「コピーユニット」と呼ぶ。図2では、複数のコピーユニット500a~500n(以下、符号「500」で代表させることもある)が示されている。後述するが、これらのコピーユニット500はユーザに割り当てられたものであり、本実施の形態では、ユーザに割り当てられたコピーユニットのことを単にコピーユニットと表現し、ユーザに割り当てる前のコピーユニットのことを、便宜上割当て済みのコピーユニットと区別するために、空のコピーユニットとして表現することもある。なお、割当て済みのコピーユニットと割当て前のコピーユニットとを特に区別することなく説明する場合には、単にコピーユニットとして表現する。このコピーユニット500は、可変の容量を有してもよいが、管理の容易性や、ユーザに対する課金などのビジネス戦略上の観点から、例えば100Mバイトなどの所定容量であってもよい。コピーユニット500には1つ又は複数のファイルデータを格納することができ、1つのファイルデータは、複数のコピーユニット500にまたがって格納されてもよい。ファイルシステムエージェント200およびコピーユニットエージェント300は、ストレージ装置10においてエージェントプログラムとして組み込まれる。コピーユニットエージェント300は、各コピーユニット500に対して起動され、コピーユニット500に格納されているデータに関する管理を行う。

30

40

#### 【0023】

ここで、図2は、ストレージ装置10の物理的な構成を示すものではなく、実際には1人の、または1人とみなすことにしたユーザに対するストレージのソフトウェア的な構成を示している。ファイルシステムエージェント200は、ユーザの識別情報(以下、ユーザアカウントという)に基づいて、複数のコピーユニット500をユーザ毎に管理する。し

50

たがって、1台のストレージ装置10が複数のユーザのデータを記録する場合には、ユーザアカウント毎にファイルシステムエージェント200が起動されることになる。コピーユニット500は、ユーザアカウントに対して予め割当てられてもよいし、データの記録を行う度に動的に割当てられてもよい。ユーザアカウントに割当てられると、コピーユニット500には識別情報(以下、コピーユニットIDという)が付与される。このコピーユニットIDは、ストレージシステム100において重複することがないように決定される。コピーユニットエージェント300は、自身が管理するコピーユニット500をこのコピーユニットIDに基づいて識別し、ファイルシステムエージェント200からのリクエストに応答することができる。

【0024】

格納部400は、一つのストレージ装置10におけるストレージ領域だけでなく、複数のストレージ装置10におけるストレージ領域も含む概念である。すなわち、ユーザの複数のコピーユニット500は、一つのストレージ装置10内だけでなく複数のストレージ装置10にまたがって存在してもよい。複数のストレージ装置10にまたがる場合には、各ストレージ装置10におけるコピーユニットエージェント300が、自身のストレージ装置10に存在するコピーユニット500に対して起動される。ファイルシステムエージェント200は、複数のストレージ装置10のうちいずれか1つに存在している。

【0025】

ファイルシステムエージェント200は、自己が管理するコピーユニット500の存在位置を記憶しておくことが好ましいが、管理対象であるコピーユニット500のIDを認識していれば、コピーユニットエージェント300に対して存在位置を問い合わせることが可能である。コピーユニットエージェント300またはコピーユニット500がユーザアカウントを保持する場合には、ファイルシステムエージェント200が、このユーザアカウントに基づいて自己が管理するコピーユニット500の存在位置をコピーユニットエージェント300に対して問い合わせるように構成してもよい。

【0026】

ストレージ装置10が上記したエージェントプログラムを備えることにより、本ストレージシステム100は、ストレージ装置10の枠組みにしばられることなく、複数のストレージ装置間の障壁を取り除いた柔軟なストレージ構成を実現することが可能となる。

【0027】

また、本ストレージシステム100の重要な特徴として、このシステムは複数の同一内容のコピーユニット500を保持している。必要な多重度、すなわちシステム中に存在しなければならない同一内容のコピーユニット500の数は、コピーユニット500またはユーザアカウントに応じて設定される。上記のとおり、コピーユニット500はユーザアカウント毎に管理されており、したがって図2に示したストレージの構成は、ストレージシステム100においてユーザアカウント毎に複数存在し、ファイルシステムエージェント200が必要な多重度の分だけ立ち上がっている。コピーユニット500に多重度を持たせる一つの理由はストレージ装置10の故障に備えることにあり、また更なる理由として、本ストレージシステム100によるデータ送信効率の向上がある。

【0028】

本ストレージシステム100は、同一内容のコピーユニット500を管理する複数のコピーユニットエージェント300が、互いに協同してコピーユニット500中のファイルデータを送信することを可能とする。例えば、同一のコピーユニット500が2つ存在する場合に、2つのコピーユニットエージェント300が、ファイルデータを半分に分けてそれぞれ送信することができる。ファイルデータの更新頻度が少なく且つデータの送信要求が非常に多いファイルデータについては、送信効率の向上という観点から多くのコピーユニット500に格納しておくことが好ましい。

【0029】

図3は、ストレージシステム100における多重度を説明するための概念図である。ストレージシステム100においては、同一のコピーユニットIDを有するコピーユニット5

10

20

30

40

50

00が少なくとも多重度分だけ存在する。図3では、多重度が2の場合であって、同一IDをもつ2つのコピーユニット500がストレージシステム100に存在している。コピーユニット500のヘッダには、ストレージ領域の空き位置と空き容量とが記録されており、ファイルデータが更新または追加等されると、空き位置および空き容量は随時書き換えられる。

#### 【0030】

各コピーユニットエージェント300はファイルシステムエージェント200により管理され、各コピーユニット500は、コピーユニットエージェント300およびファイルシステムエージェント200により管理される。図3の例では便宜上、ファイルシステムエージェント200がコピーユニット500を管理するものとして説明する。

10

#### 【0031】

ユーザに対して計12の所定容量のコピーユニット500が割当てられ、ファイルシステムエージェント200aが、ID01~ID06で特定される6つのコピーユニット500a~500fの一系列を管理し、同様にファイルシステムエージェント200bが、ID01~ID06で特定される6つのコピーユニット500g~500lの一系列を管理する。

#### 【0032】

これらのコピーユニット500には、ファイルA、ファイルBおよびファイルCの3つのデータが格納されている。このうち、ファイルCのデータは、ID02~ID04のコピーユニット500にまたがって格納されている。また図中、斜線が引かれた領域は空き領域を示し、この例ではID04のコピーユニット500d(500j)の一部と、ID05のコピーユニット500e(500k)およびID06のコピーユニット500f(500l)の全体が空き領域となっている。同一ユーザの新たなファイルデータは、この空き領域に書き込まれる。新たなファイルデータの容量が各コピーユニット500の空き容量の総和よりも大きいときには、ファイルシステムエージェント200が新たにコピーユニット500を生成して対応してもよい。コピーユニットエージェント300は、自身が管理するコピーユニット500の空き容量を認識できることが好ましいが、他のコピーユニット500の空き容量を認識しなくてもよい。ファイルシステムエージェント200は、各コピーユニットエージェント300に対してコピーユニット500の空き容量を報告させることにより、システム全体の空き容量を演算することが可能である。

20

30

#### 【0033】

図4は、ファイルシステムエージェント200の構成を示す機能ブロック図である。ファイルシステムエージェント200は、入出力処理部210a、210b、要求受付部212、応答部214、応答受付部216、データ受付部218、データ送信部220、容量処理部230、コピーユニット指定部238および書込制御部240を備える。容量処理部230は、容量検出部232、容量判断部234および容量拡大部236を有する。入出力処理部210aは、他のストレージ装置10または外部装置20との間で信号を送受し、入出力処理部210bは、コピーユニットエージェント300との間で信号を送受する機能を有する。入出力処理部210aは、外部から入力された信号に含まれるユーザアカウントに基づいて、自身が管理するユーザからの要求であるか否かを判断する。別のユーザからの要求である場合にはその信号を無視し、ファイルシステムエージェント200は動作しない。コピーユニットID格納部250は、管理するコピーユニット500のIDを格納する。ID以外にも、コピーユニット500の存在位置が格納されることが好ましい。

40

#### 【0034】

要求受付部212は、他のストレージ装置10または外部装置20から送信される要求を受け付ける。新規ファイルデータの書込時には外部装置20からデータの書込要求が供給され、ファイルデータの送信時には外部装置20からデータの送信要求が供給され、またコピーユニット500の複製時には他のストレージ装置10からデータの複製要求が供給される。

50

## 【 0 0 3 5 】

新規ファイルデータの書込時、要求受付部 2 1 2 が受け付けた書込要求が容量処理部 2 3 0 に供給される。この書込要求を受けて、容量検出部 2 3 2 が、管理する 1 つ又は複数のコピーユニット 5 0 0 におけるストレージ領域の空き容量を検出する。この検出方法には様々な手法があるが、一つの例として容量検出部 2 3 2 が各コピーユニットエージェント 3 0 0 にコピーユニット 5 0 0 の空き容量を問い合わせ、コピーユニットエージェント 3 0 0 が、コピーユニット 5 0 0 のヘッダに記録された空き容量を読み出して容量検出部 2 3 2 に回答し、容量検出部 2 3 2 がその総和を演算する。別の手法としては、容量検出部 2 3 2 が最新の空き容量をデータとして常に保持するようによい。

## 【 0 0 3 6 】

容量判断部 2 3 4 は、空き容量が新規ファイルデータの書込みに十分であるか否かを判断する。この判断は、新規ファイルデータの大きさと、空き容量とを比較することにより行われる。新規ファイルデータの大きさは、書込要求に含まれている。新規ファイルデータの大きさが空き容量を超える場合には、容量拡大部 2 3 6 が、データのストレージ領域を拡大するべく、空のコピーユニットを探し出し、ユーザに割り当てる。このとき、容量拡大部 2 3 6 は、ファイルシステムエージェント 2 0 0 が存在するストレージ装置 1 0 内で空のコピーユニットを探してもよいが、別のストレージ装置 1 0 から探し出してもよい。図 3 に関連して説明したように、同一内容のコピーユニット 5 0 0 を管理するファイルシステムエージェント 2 0 0 は多重度の数だけ存在しており、新たなコピーユニット 5 0 0 をユーザに割り当てる場合には、ファイルシステムエージェント 2 0 0 同士が連絡をとりあって、同数のコピーユニット 5 0 0 をユーザに割り当て、互いに同一のコピーユニット ID を付すようにする。割り当てたコピーユニット 5 0 0 の ID は、コピーユニット ID 格納部 2 5 0 に追加される。容量拡大部 2 3 6 のこの機能は、空き容量が不足したときに自動的に実行されてもよいが、一旦ユーザに対して容量不足の旨を通知し、容量を拡大することの指示がユーザからなされることを条件として実行されてもよい。

## 【 0 0 3 7 】

コピーユニット指定部 2 3 8 は、新規ファイルデータを格納するコピーユニット 5 0 0 の ID を指定する。この指定が完了すると、応答部 2 1 4 が、書込要求に対する応答をネットワーク上にブロードキャストする。ブロードキャストすることによって、この応答は、外部装置 2 0 だけでなく、同一内容のコピーユニットを管理する他のファイルシステムエージェント 2 0 0 にも供給される。十分な容量を確保できない場合、応答部 2 1 4 は、新規ファイルデータの書込みを要求する外部装置 2 0 に対して容量不足の旨を通知する。空き容量が十分である場合には、書込みを受け入れる状態にある旨を通知する。

## 【 0 0 3 8 】

このとき、同一内容のコピーユニット 5 0 0 を管理する他のファイルシステムエージェント 2 0 0 も書込要求に回答し、この応答はネットワーク上にブロードキャストされる。応答受付部 2 1 6 は、他のストレージ装置 1 0 からの応答を受け付け、他のストレージ装置 1 0 が正常に動作していることを認識する。応答部 2 1 4 による応答と、応答受付部 2 1 6 による応答受付の時間的な関係により、格納したデータを外部に送信する場合の優先順位などが決定されてもよい。例えば、ファイルシステムエージェント 2 0 0 a が、自身の応答後に他のファイルシステムエージェント 2 0 0 b からの応答を受け付け、一方、他のファイルシステムエージェント 2 0 0 b においては、応答する前にファイルシステムエージェント 2 0 0 a からの応答を受け付けた場合、ファイルシステムエージェント 2 0 0 a の優先順位を他のファイルシステムエージェント 2 0 0 b よりも高く設定することができる。この優先順位は、コピーユニットエージェント 3 0 0 に供給され、データの読出制御を行う際に利用されてもよい。

## 【 0 0 3 9 】

外部装置 2 0 は、ストレージ装置 1 0 から書込要求に対する応答を受けると、新規ファイルデータを送信する。データ受付部 2 1 8 がこのファイルデータを受け付ける。書込制御部 2 4 0 は、コピーユニット指定部 2 3 8 により指定されたコピーユニット 5 0 0 の空き

10

20

30

40

50

領域にデータを書き込み、ヘッダには変更された空き位置と空き容量とを書き込む。

【 0 0 4 0 】

容量処理部 2 3 0 の上記説明ではファイルシステムエージェント 2 0 0 が容量を判断することとしたが、コピーユニットエージェント 3 0 0 が書込要求に対して自律的にコピーユニット 5 0 0 の空き容量を検出して、新規ファイルデータを格納できるか否かを判断してもよい。この場合、コピーユニットエージェント 3 0 0 が容量検出部 2 3 2 および容量判断部 2 3 4 を有し、要求受付部 2 1 2 から供給される書込要求に応答する。この応答は各コピーユニットエージェント 3 0 0 からコピーユニット指定部 2 3 8 に供給され、コピーユニット指定部 2 3 8 がデータの書込みに最適なコピーユニット 5 0 0 を指定する。1つのコピーユニット 5 0 0 にデータを格納できる場合には、コピーユニット指定部 2 3 8 は、そのコピーユニット 5 0 0 を指定する。一方、1つのコピーユニット 5 0 0 にデータを格納しきれない場合には、コピーユニット指定部 2 3 8 が容量処理部 2 3 0 を有して、適切なコピーユニット 5 0 0 の組み合わせを定めることが好ましい。

10

【 0 0 4 1 】

図 5 は、ファイルデータをコピーユニット 5 0 0 に書き込むフローを示すフローチャートである。複数のストレージ装置 1 0 のうちストレージ装置 1 0 a および 1 0 b が、外部装置 2 0 から送信されるデータを受け付けるファイルシステムエージェント 2 0 0 を有している。この例では、必要な多重度が 2 の場合を想定している。

【 0 0 4 2 】

まず、外部装置 2 0 において新規のファイルデータが生成される ( S 1 0 0 )。それからユーザがデータ書込要求を生成し ( S 1 0 2 )、書込要求をブロードキャストする ( S 1 0 4 )。それから、ユーザアカウントに基づいて、ストレージ装置 1 0 a および 1 0 b が、書込要求を受け付け ( S 1 0 6、S 1 0 8 )、ファイルデータの書込みを受入可能な状態とする。ストレージ装置 1 0 c は、このユーザのコピーユニットを管理していないため、この書込要求を受け付けない。このように本ストレージシステム 1 0 0 においては、各ストレージ装置 1 0 が自律的に書込要求の取捨をするため、ユーザは自分のファイルデータをどのストレージ装置 1 0 に格納するべきかを知る必要がない。ストレージ装置 1 0 a および 1 0 b は、書込要求に対する応答をブロードキャストで送信し ( S 1 1 0、S 1 1 2 )、外部装置 2 0 は、それぞれからの応答を受け付ける ( S 1 1 4 )。ストレージ装置 1 0 a および 1 0 b も、他のストレージ装置からの応答を受け付ける ( S 1 1 6、S 1 1 8 )。

20

30

【 0 0 4 3 】

外部装置 2 0 は、多重度が 2 であることと、ストレージ装置 1 0 からの応答を 2 つ受けたことを確認して、ファイルデータをブロードキャストする ( S 1 2 0 )。このとき、ストレージ装置 1 0 a と 1 0 b のみにデータをマルチキャストしてもよい。ストレージ装置 1 0 a および 1 0 b はデータを受け付け ( S 1 2 2、1 2 4 )、同一の ID を有するコピーユニットにデータを同時に書き込んでいく ( S 1 2 6、S 1 2 8 )。データをブロードキャスト又はマルチキャストで送信することにより複数のコピーユニットへの同時書込を実現し、コピーユニットをリアルタイムに多重化することが可能となる。

【 0 0 4 4 】

なお、後で詳述するが、本ストレージシステム 1 0 0 は、コピーユニットの多重度が不足する場合に、そのコピーユニットを複製する機能を有している。そのため外部装置 2 0 は、多重度が 2 の場合であっても、ストレージ装置 1 0 a または 1 0 b の一方からの応答を受け付けた段階でファイルデータを送信し、データ受信後この一方のストレージ装置は、他方のストレージ装置に複製要求をだして、コピーユニットの多重度を 2 に保つように複製作業を行うことができる。

40

【 0 0 4 5 】

次に、外部装置 2 0 が、コピーユニット 5 0 0 に格納したファイルデータの送信を要求する場合について説明する。図 4 において、要求受付部 2 1 2 が外部装置 2 0 からデータの送信要求を受け付け、コピーユニット指定部 2 3 8 に供給する。コピーユニット指定部 2

50

38は、このファイルデータを格納しているコピーユニットを特定して、対応するコピーユニットエージェント300に対して、入出力処理部210bを介してデータの送信指示を供給する。図3の例でファイルデータAの送信要求を受けた場合、コピーユニット指定部238は、コピーユニット500aを管理するコピーユニットエージェントに対して送信要求を供給する。ファイルデータCの送信要求を受けた場合には、コピーユニット500b、500cおよび500dを管理する3つのコピーユニットエージェントに対して送信要求を供給する。

#### 【0046】

図6は、コピーユニットエージェント300の構成を示す機能ブロック図である。コピーユニットエージェント300は、入出力処理部310、送信要求受付部312、読出制御部314、同一コピーユニット探索部316、送信ルール決定部318、状況送信部320、監視部330、複製コマンド受付部340および複製要求生成部350を備える。監視部330は状況受付部332および多重度判断部334を有する。

10

#### 【0047】

送信要求受付部312が、ファイルシステムエージェント200より供給される送信要求を受け付ける。この要求の受け付けを契機として、同一コピーユニット探索部316が、他のストレージ装置10における同一内容のコピーユニットを探索する。コピーユニットエージェント300は、同一内容のコピーユニットの存在位置を予め記録しておいてもよい。同一内容のコピーユニットが見つかり、送信ルール決定部318が、同一内容のコピーユニットを管理するコピーユニットエージェントとデータの送信ルールを決定する。送信ルール決定部318は、図4の応答受付部216において生成された優先順位を用いて送信ルールを決定してもよい。読出制御部314は、同一内容のコピーユニットを有する他のストレージ装置と協同してデータを送信するようにデータを読み出す。読み出されたデータは、ファイルシステムエージェント200のデータ送信部220からネットワーク30上に送信される。

20

#### 【0048】

図3において、ファイルAの送信要求があった場合、コピーユニット500aを管理するコピーユニットエージェント300aと、コピーユニット500gを管理するコピーユニットエージェント300gとの間で、ファイルAの送信ルールを決定する。単純には、ファイルAの半分をコピーユニットエージェント300aが送信し、ファイルAの残り半部分をコピーユニットエージェント300gが送信するというルールが考えられる。例えば、数十キロバイト程度の所定長のデータブロック毎にファイルAを分割して、ブロックを交互に送信するルールを採用することも可能である。データの送信においては、ハードディスクからの読出しがボトルネックとなるため、小さいデータを交互に送信することによりデータを連続的に送信することが可能となる。コピーユニット500の多重度をさらに大きくすることにより、データの送信効率は向上する。また、複数のストレージ装置10からデータを送信するため、あるストレージ装置に障害が生じた場合であっても、残りのストレージ装置でその分を補うことができ、耐障害性に優れたデータ送信を確立することが可能となる。

30

#### 【0049】

図7は、ファイルデータをコピーユニット500から読み出して送信するフローを示すフローチャートである。複数のストレージ装置10のうちストレージ装置10aおよび10bが、そのファイルデータを格納したコピーユニット500を管理するコピーユニットエージェント300を有している。データ送信要求の受付の可否は、ユーザアカウントによって決定される。なお前述したとおり、このコピーユニット500は、必ずしもファイルシステムエージェント200を有するストレージ装置10に位置する必要はなく、別のストレージ装置10に存在してもよい。

40

#### 【0050】

外部装置20が、データ送信要求を生成し(S200)、送信要求をブロードキャストする(S202)。ストレージ装置10aおよび10bが、この送信要求を受け付け(S2

50

04、S206)、それに対する応答をブロードキャストする(S208、S210)。外部装置20は、各ストレージ装置10a、10bからの応答を受け付け(S212)、またストレージ装置10aおよび10bも、他方からの応答を受け付ける(S214、S216)。外部装置20は、ファイルデータの書込みを可能な状態とする。それから、ストレージ装置10aと10bとの間で送信ルールを決定する(S218)。ストレージ装置10aおよび10bは、送信ルールに従ってファイルデータを読み出し(S220、S222)、互いに協同してデータを送信して(S224、S226)、外部装置20はこのデータを受け付ける(S226)。

#### 【0051】

上記の説明では、便宜上ストレージ装置10aおよび10bがファイルデータを読み出すこととしたが、実際には、ファイルデータの読出しは、対象となるコピーユニット500を管理するコピーユニットエージェント300が行う。したがって、コピーユニット500がファイルシステムエージェント200とは異なるストレージ装置10に格納されている場合には、そのストレージ装置10内のコピーユニットエージェント300がファイルデータの読出しを行うことになる。

10

#### 【0052】

続いて、本ストレージシステム100におけるコピーユニット500の複製機能について説明する。本ストレージシステム100においては、ストレージ装置10の耐障害性およびデータ送信効率の向上などの観点から、コピーユニット500の多重度を一定以上に保つことを特徴の一つとしている。そのため前述したように、新規ファイルデータは原則として複数のストレージ装置10に同時に書き込まれ、多重度に変化のない状態が理想である。

20

#### 【0053】

しかしながら、ストレージ装置10が物理的な構成をとる以上、劣化は避けられず、使用期間が長くなるにつれて故障する確率も高くなる。また、ストレージシステム100の柔軟性を高めるためにも、新規書込みを要求する外部装置20に対して、必ずしも複数のストレージ装置10に同時書込を行わなければならないという制約を課したくない。さらに、ハードディスク等の記録密度は年々飛躍的に向上しており、省スペース化を実現するためにも、一定周期でストレージ装置10を簡易に交換できることが望ましい。以上の理由から、本ストレージシステム100は、多重度を一定以上に保つべく、自律的にコピーユニット500を複製する機能を有している。

30

#### 【0054】

図6において、コピーユニットエージェント300において、状況送信部320が、自身のストレージ状況を示す信号を周期的にブロードキャストする。このストレージ状況を示す信号は、コピーユニットが正常に存在していることを示す信号である。このストレージ状況を示す信号は、キープアライブパケットとして送信されてもよい。また、同一内容のコピーユニット500を管理する他のコピーユニットエージェント300も、ストレージ状況を示す信号を周期的にブロードキャストする。監視部330は、他のコピーユニットエージェント300から送信されるストレージ状況を示す信号を受けて、ストレージ状況を監視する。具体的には、状況受付部332が、他のコピーユニットエージェント300から送信されるストレージ状況を示す信号を受け付ける。周期的に受け付けている場合には、他のストレージ装置10が正常に動作していることを認識する。自身も正常に動作している場合には、多重度判断部334が、管理するコピーユニット500が所定の多重度を保ってストレージシステム100に存在することを判断する。

40

#### 【0055】

一方、他のコピーユニットエージェント300から周期的にストレージ状況を示す信号を受け付けられない場合には、多重度判断部334は、コピーユニットの多重度が所定に保たれていないことを判断する。例えばコピーユニット500の必要な多重度が2と定められている場合に、自身は正常であるのに、他からストレージ状況を示す信号を受けられなくなったときには、多重度判断部334は、コピーユニット500の多重度が1になった

50

ことを検出し、多重度の不足分を複製要求生成部 350 に供給する。複製要求生成部 350 は、多重度の不足分（このケースでは 1）だけコピーユニットを複製するように他のストレージ装置 10 に対する複製要求を生成する。

#### 【0056】

図 8 は、複製要求を生成するフローを示すフローチャートである。まず、周期的に他のストレージ装置 10 から送信されるストレージ状況を示す信号を監視する（S300）。それから多重度を検出し（S302）、検出した多重度と必要な多重度とを比較する（S304）。検出した多重度が必要な多重度以上である場合（S304 の N）には、コピーユニット 500 が正常に保存されていることを認識し、S300 から S304 までのステップを繰り返す。一方、検出した多重度が必要な多重度よりも小さい場合（S304 の Y）には、コピーユニット 500 の多重度が不足していることを検知し、コピーユニット 500 を格納する他のストレージ装置 10 に故障が生じたと判断して、コピーユニット 500 の複製要求を生成する（S306）。なお、S304 において、他のストレージ装置 10 に故障が生じたことを判断するのは、ストレージ状況を示す信号の送信周期に比して十分に長い期間ストレージ状況を示す信号を受け付けなかったことを条件とするのが好ましい。ストレージ装置 10 に故障がない場合であっても、ネットワーク 30 上のトラフィックの問題で、一時的に受信するタイミングが遅れることもあるからである。

#### 【0057】

図 6 において、故障の場合でなく、ストレージ装置 10 を交換する場合には、元のストレージ装置 10 がネットワーク 30 から接続を切断される前に複製コマンドが生成され、複製コマンド受付部 340 が受け取る。複製コマンド受付部 340 はその旨を複製要求生成部 350 に通知し、複製要求生成部 350 が複製要求を他のストレージ装置 10 に送信する。この複製コマンドは、ユーザにより入力されてもよいが、ユーザからの切断要求を受けた際に、複製コマンドを自動的に生成して複製コマンド受付部 340 に供給するようにしてもよい。なお、複製コマンドを供給しない場合、例えばネットワークとの接続が突然切断されたような場合であっても、監視部 330 がストレージ状況を監視することにより多重度が保たれることは前述したとおりである。

#### 【0058】

図 4 において、他のストレージ装置 10 から供給される複製要求は、要求受付部 212 にて受け付けられる。複製要求を受けて、コピーユニット指定部 238 は、空のコピーユニットに ID を付与して新たなコピーユニットをユーザに割り当てる。この ID は、複製されるコピーユニット 500 と同一のものである。応答部 214 がその旨を、複製要求を生成したコピーユニットエージェント 300 に送信する。要求元のコピーユニットエージェント 300 は、その応答を受けて、当該コピーユニット 500 のデータを送信し、要求先のコピーユニットエージェント 300 は、そのデータをそのまま新たなコピーユニットに複製する。

#### 【0059】

複数のコピーユニット 500 のうちの 1 つは、ファイルシステムエージェント 200 を起動可能なデータを有している。ストレージ装置 10 が突然ネットワーク 30 から切断されて、ユーザのコピーユニットの一系列が全て消失した場合、生き残っている他の系列のコピーユニット 500 がそれぞれ複製要求を生成して、他のストレージ装置 10 にコピーユニット 500 を複製する。それから、複製されたコピーユニット 500 のうちの 1 つがファイルシステムエージェント 200 を起動し、起動されたファイルシステムエージェント 200 が複製された全てのコピーユニット 500 を管理する。このようにして、コピーユニット 500 の一系列を復元することが可能となる。

#### 【0060】

図 9 は、本発明の第 2 の実施形態に係るストレージシステム 700 の構成図である。このストレージシステム 700 は、複数のストレージ群 600 a、600 b、600 c を接続して構成され、各ストレージ群 600 は、ゲートウェイ装置 12 と、このゲートウェイ装置にネットワーク接続される複数のストレージ装置 10 を備える。ストレージ装置 10 お

10

20

30

40

50

よびネットワーク30は、第1の実施形態において説明したストレージシステム100の構成をとる。第2の実施形態のように複数のストレージ群を接続することによって、ストレージ群内の容量の制約にとらわれず、システム全体の容量を制限なく増加させることが可能となる。ストレージ群600間の通信は、ゲートウェイ装置12を介して行われる。

【0061】

また、同一内容のコピーユニット500は、少なくとも2つのストレージ群で保持されることが好ましい。災害などによりストレージ群600が消失する危険性を回避するために、コピーユニット500を物理的に離れたストレージ群600に格納する。

【0062】

図10は、本発明の実施形態にかかるストレージ装置10のハードウェアコンポーネントを示す構成図である。ストレージ装置10は、表示装置50、入力部52、CPU54、ハードディスク56、RAM(Random Access Memory)58およびドライブ装置60を備える。これらの構成は、バス62などの信号伝送路により電氣的に接続されている。入力部52は、ユーザからストレージ装置10をネットワークから切り離す切断コマンドなどを受け付ける。このストレージ装置10がストレージシステム100の専用機器として構成される場合、ユーザに可視的な情報を提示する必要やユーザからの直接入力を受け付ける必要がないことも考えられ、この場合には表示装置50、入力装置52、ドライブ装置60などは存在しなくてもよい。あるストレージ装置10にウェブサーバ機能を持たせて、同じネットワークに接続された別のストレージ装置10からの入力を受け付けるように構成してもよい。

【0063】

ハードディスク56は、大容量の磁気記録装置であり、ストレージ領域は複数のコピーユニットに分割される。ハードディスク56には、このコピーユニットを記録単位としてデータが記録される。ここで、ハードディスク56は、他の大容量記録装置への置換えが可能である。記録媒体70は、実施形態に関連して説明したストレージ装置10の様々な機能を、CPU54に実現させるためのプログラムを記録する。記録媒体70がドライブ装置60に挿入されると、そのプログラムは、ハードディスク56またはRAM58に読み出され、CPU54は、読み出されたプログラムによりハードディスク56の分割機能やストレージ管理機能などを実現する。この記録媒体70は、CD-ROM、DVD、FDなどのコンピュータ読み取り可能な媒体である。

【0064】

ここでは、ストレージ管理用のプログラムが記録媒体70に記録されている例について説明したが、別の例においては、プログラムは、無線、有線を問わず、外部のコンピュータまたは端末から送信されてもよい。図10に示したハードウェア構成において、プログラムは、コンピュータにストレージ管理機能を実現させればよいのであって、外部から供給される場合だけでなく、予めハードディスク56に格納されていて、ストレージシステム100用の専用機器を構成するものとして存在してよいことも当業者には理解されることである。

【0065】

以上、本発明を実施の形態をもとに説明したが、本発明の技術的範囲は上記実施の形態に記載の範囲には限定されない。上記実施の形態は例示であり、それらの各構成要素や各処理プロセスの組合せに、さらにいろいろな変形例が可能で、またそうした変形例も本発明の範囲にあることは当業者に理解されることである。

【0066】

例えば、上述の実施形態ではストレージ装置10がファイルシステムエージェント200とコピーユニットエージェント300を用いて実現されているが、一方のエージェントの機能を他方のエージェントが有していてもよく、または両者が有していてもよい。本ストレージシステム100において、各ストレージ装置10は自律的に動作する機能を有しており、データの複製等を行う際に管理サーバを必要としないが、他の何らかの機能を実現するための形式的なサーバを備えることも可能である。また、上記例では、ファイルデー

10

20

30

40

50

タをユーザごとに管理することとしているが、ユーザの制限を外して、複数のユーザのデータを1つのファイルシステムエージェントに管理させてもよい。

【0067】

また本実施形態の説明において、データの書込要求を行う外部装置20と、データを格納するストレージ装置10とを別の構成として記載しているが、データの書込要求を行う装置とストレージ装置10とは一体として構成されてもよい。例えば、一つのLANに接続する複数のパーソナルコンピュータがストレージ装置10としての機能も有している場合、あるコンピュータのユーザが、自身で作成したファイルデータを一旦該コンピュータに備わるハードディスクに格納する。その後、コンピュータ内で起動されるコピーユニットエージェントが、そのデータを他のコンピュータのハードディスクにおいて複製するように複製要求を生成してもよい。ユーザは自身のハードディスクにデータを保存することにより、データの更新などの作業を簡易に行うことが可能となり、また自身のハードディスクには常に最新のデータが格納されることとなるため、自身のハードディスク内のデータと複製先のデータが異なるものである場合には、複製先のデータが古いデータであることを認識して、最新のデータを複製先のデータにオーバーライトすることが容易に実現可能となる。

10

【0068】

また、ストレージシステム100の運用上、ユーザに割当てたコピーユニットを開放したい場合がある。この場合には、例えばストレージ装置10自身ないしはその他の装置でコピーユニットの開放コマンドが生成され、対応するコピーユニットエージェントがこれを受けてコピーユニットを不活性にし、コピーユニット内のデータを更新できない状態にする。本システム100において同一内容のコピーユニットは複数存在しており、これら複数のコピーユニット全てが同時期に不活性にされる。なお、開放コマンドの受付時期の関係で、一つのコピーユニットが不活性にされた後、まだ他のコピーユニットが不活性にされていない場合に、当該他のコピーユニットが自律的に複製要求を生成しないようにするために、複製要求の生成タイミングは、不活性にされたコピーユニットからストレージ状況を示す信号を受け取らなくなった時から十分に時間が経過したときであることが好ましい。このように複製要求の生成タイミングを設定することで、当該他のコピーユニットは、複製要求を生成する前に開放コマンドを受け付けることが可能となり、実質的に同時期に複数の同一内容のコピーユニットを不活性にすることが可能となる。コピーユニットエージェントは、コピーユニットを不活性にした後、該コピーユニット内のデータおよびIDを消去して、元の空のコピーユニットの状態に戻してもよい。

20

30

【0069】

【発明の効果】

本発明によれば、自律的に機能するストレージ装置およびこのストレージ装置を利用したストレージシステムを提供することができる。

【図面の簡単な説明】

【図1】本発明の第1の実施形態に係るストレージシステム100の構成図である。

【図2】ストレージ装置10の概念的な構成を示すブロック図である。

【図3】ストレージシステム100における多重度を説明するための概念図である。

40

【図4】ファイルシステムエージェント200の構成を示す機能ブロック図である。

【図5】ファイルデータをコピーユニット500に書き込むフローを示すフローチャートである。

【図6】コピーユニットエージェント300の構成を示す機能ブロック図である。

【図7】ファイルデータをコピーユニット500から読み出して送信するフローを示すフローチャートである。

【図8】複製要求を生成するフローを示すフローチャートである。

【図9】本発明の第2の実施形態に係るストレージシステム700の構成図である。

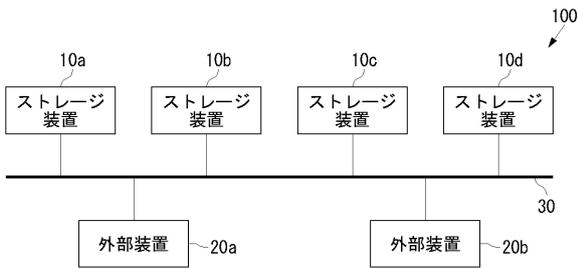
【図10】本発明の実施形態にかかるストレージ装置10のハードウェアコンポーネントを示す構成図である。

50

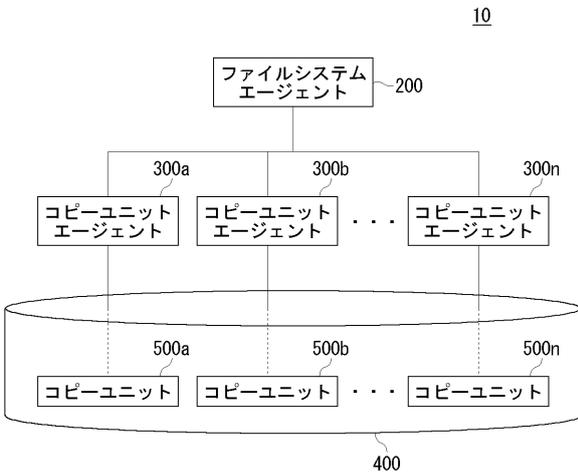
【符号の説明】

10・・・ストレージ装置、20・・・外部装置、30・・・ネットワーク、100・・・ストレージシステム、200・・・ファイルシステムエージェント、300・・・コピーユニットエージェント、400・・・格納部、500・・・コピーユニット

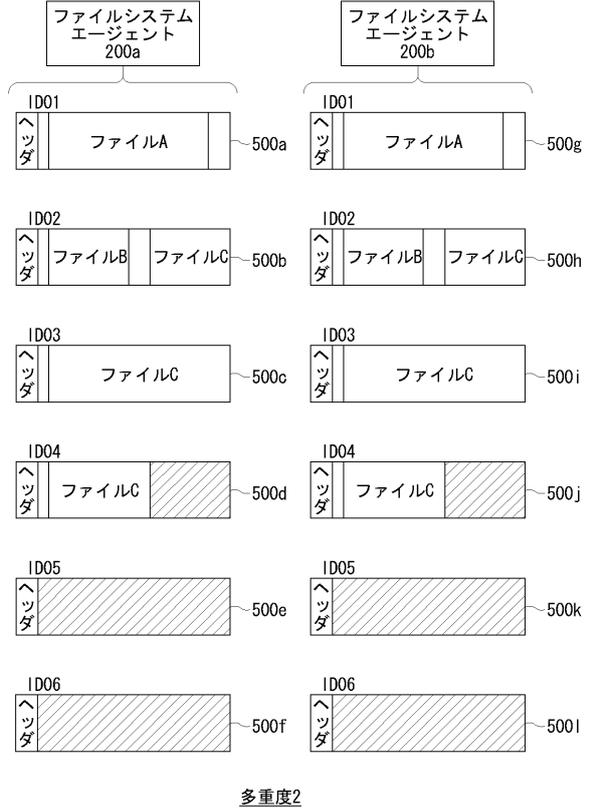
【図1】



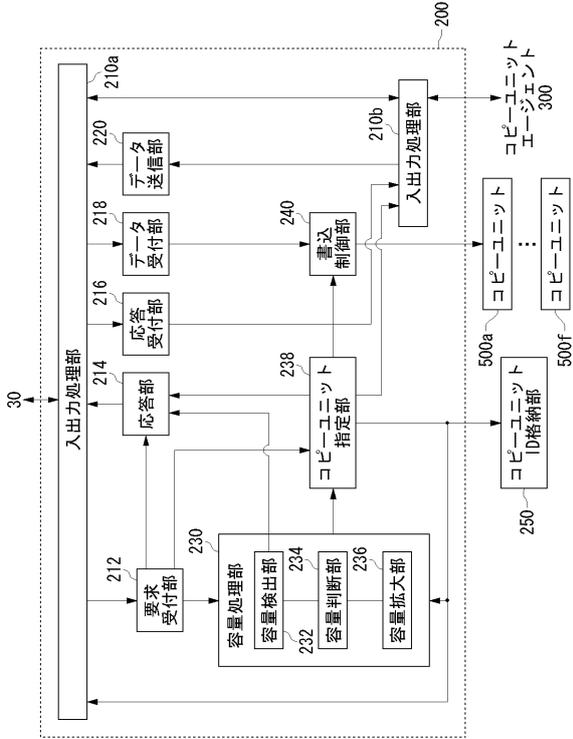
【図2】



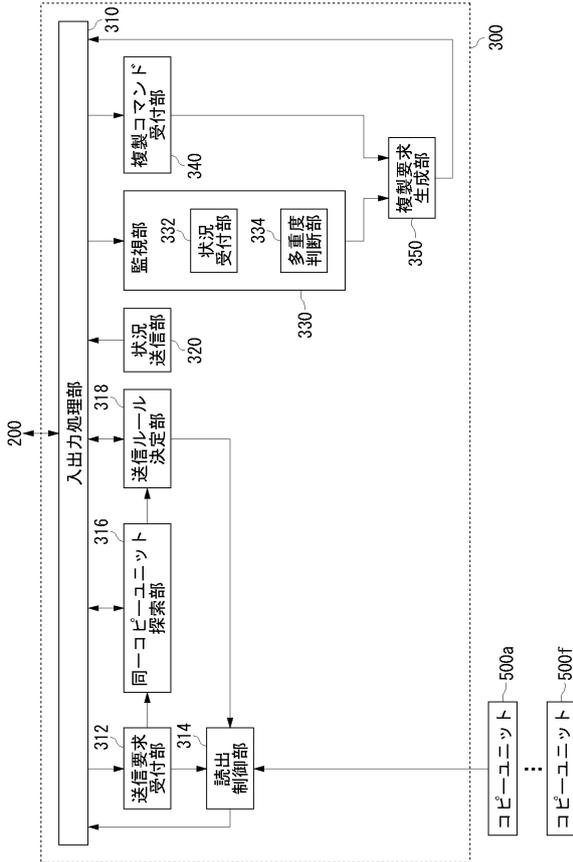
【図3】



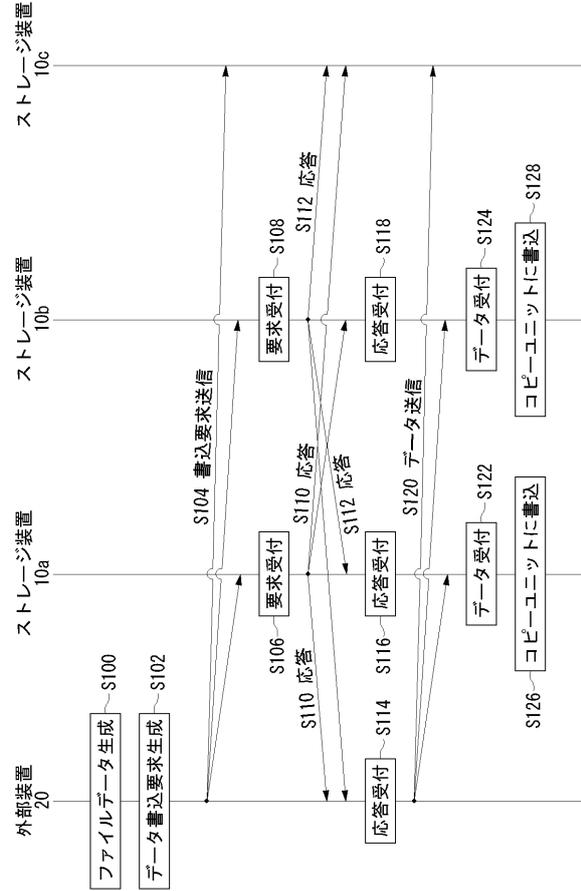
【 図 4 】



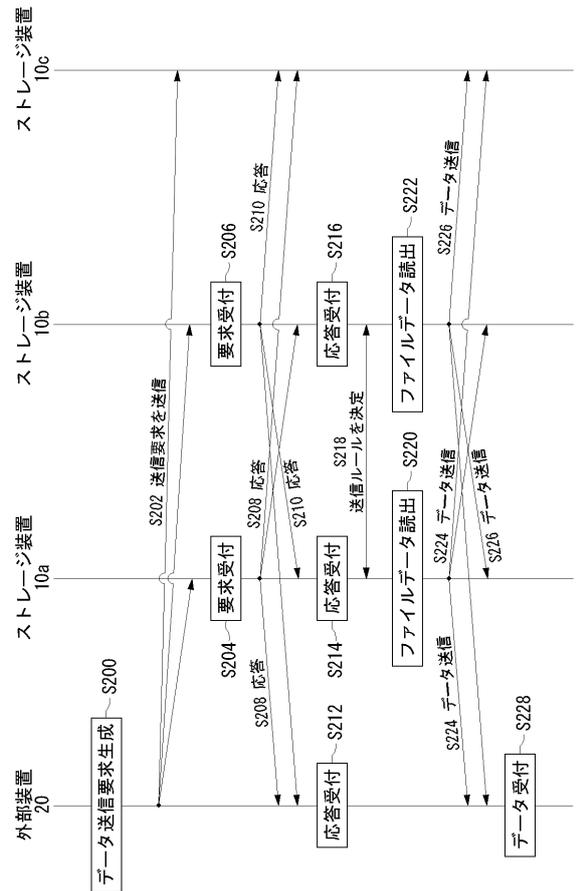
【 図 6 】



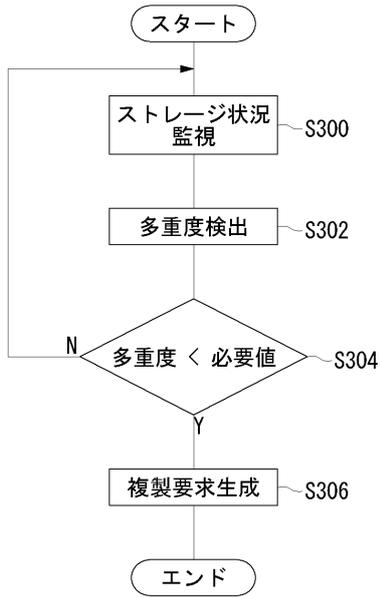
【 図 5 】



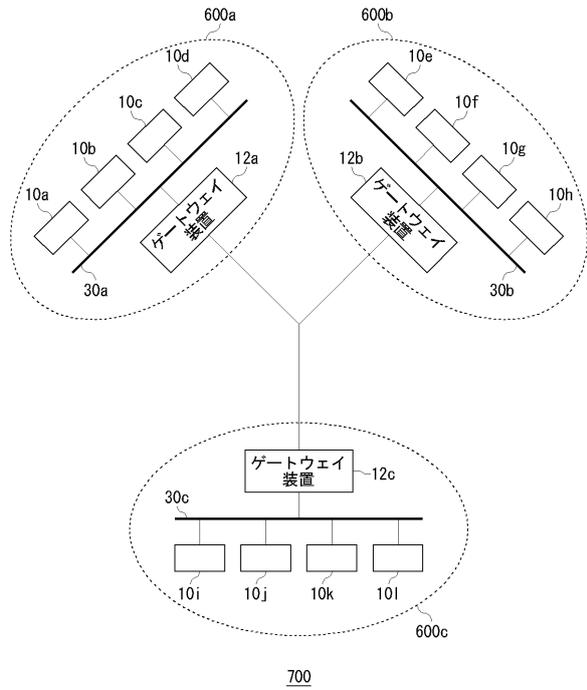
【 図 7 】



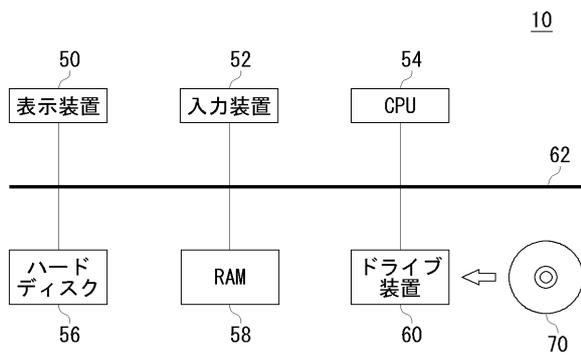
【図8】



【図9】



【図10】



---

フロントページの続き

(56)参考文献 特表2001-508900(JP,A)  
特開平11-053122(JP,A)  
特開2000-267979(JP,A)  
特開平07-253917(JP,A)  
特開平04-288638(JP,A)  
特開平11-328038(JP,A)  
特開2000-089996(JP,A)  
特開2001-051890(JP,A)  
特表平11-502644(JP,A)  
特開2000-112675(JP,A)  
特開平06-348527(JP,A)  
特開平06-077993(JP,A)

(58)調査した分野(Int.Cl., DB名)

G06F 3/06  
G06F 12/00  
G06F 13/00