

【特許請求の範囲】**【請求項 1】**

計算機とネットワークを介して接続するファイルサーバと、
前記ファイルサーバとネットワークを介して接続するストレージ装置とを含み、
前記ファイルサーバは、第 1 のコントローラを含み、
前記ストレージ装置は、複数の記憶領域を構成する複数の記憶デバイスと、前記複数の記憶領域へのアクセスを制御する第 2 のコントローラと、を含み、
前記複数の記憶領域はそれぞれ、省電力モードからレディモードへの移行時間が異なる複数の省電力モードのうち、少なくとも 1 つの省電力モードを備え、
前記第 1 のコントローラは、
前記計算機からデータを受信すると、前記データに、前記計算機からのアクセスへの応答性能に関する指標を設定し、
前記データの前記指標を参照して、前記指標を満たす省電力モードを備える第 1 の記憶領域を選択し、
前記第 2 のコントローラは、
前記第 1 の記憶領域に、前記データを格納する、
ストレージシステム。

10

【請求項 2】

前記指標とは、計算機がアクセス要求を送信してから完了報告を受けるまでの時間に関連する指標である、
請求項 1 に記載のストレージシステム。

20

【請求項 3】

前記複数の省電力モードとは、少なくとも前記記憶デバイスの電源を遮断するモードを含む、
請求項 2 に記載のストレージシステム。

【請求項 4】

前記ファイルサーバは、さらに計算機から送信されたデータにメタデータを格納するメモリを含み、
前記第 1 のコントローラは、前記メタデータに前記指標を設定する、
請求項 3 に記載のストレージシステム。

30

【請求項 5】

前記ファイルサーバは、さらにデータが格納されるフォルダを設定するファイルシステムを含み、
前記フォルダには、前記指標が設定され、
前記計算機から送信されたデータを前記フォルダに格納すると、前記フォルダに設定された指標が、前記データの前記メタデータに設定される、
請求項 4 に記載のストレージシステム。

【請求項 6】

前記第 2 のコントローラは、最終アクセス時間から、一定の時間が経過した後、前記レディモードになった記憶領域を省電力モードへ移行する、
請求項 3 に記載のストレージシステム。

40

【請求項 7】

前記第 1 のコントローラは、前記第 1 の記憶領域が、省電力モードである場合、レディモードである第 2 の記憶領域を選択し、
前記第 2 のコントローラは、前記第 2 の記憶領域に前記データを格納し、
前記第 2 のコントローラは、前記第 1 の記憶領域を省電力モードから、レディモードへ移行し、
前記第 2 のコントローラは、前記第 2 の記憶領域から、前記第 1 の記憶領域にデータを移動する、
請求項 1 に記載のストレージシステム。

50

【請求項 8】

前記ファイルサーバは、さらに記憶領域と記憶デバイスとの関係を示すテーブルを格納するメモリを有し、

前記第 1 のコントローラは、前記第 1 の領域を構成する第 1 の記憶デバイスを特定し、前記ストレージ装置に第 1 の記憶デバイスへの前記データの格納指示を送信する、

請求項 1 に記載のストレージシステム。

【請求項 9】

前記テーブルは、記憶デバイスの現在の動作モードを管理し、

前記第 1 のコントローラは、前記第 1 の記憶デバイスが省電力モードである場合には、前記ストレージ装置へ前記第 1 の記憶デバイスをレディモードに移行する指示を送信する

10

、
請求項 8 に記載のストレージシステム。

【請求項 10】

計算機とネットワークを介して接続するファイルサーバと、

前記ファイルサーバとネットワークを介して接続し、

複数の記憶領域を構成する複数の記憶デバイスを備え、

前記複数の記憶領域へのアクセスを制御するストレージ装置と、

を含む、ストレージシステムの制御方法であって、

前記ストレージ装置が、前記複数の記憶領域に、それぞれ、省電力モードからレディモードへの移行時間が異なる複数の省電力モードのうち、少なくとも 1 つの省電力モードを設定し、

20

前記ファイルサーバが、前記計算機から受信したデータに、前記計算機からのアクセスへの応答性能に関する指標を設定し、

前記ファイルサーバが、前記データの前記指標を参照して、前記指標を満たす省電力モードを備える第 1 の記憶領域を選択し、

前記ストレージ装置は、前記第 1 の記憶領域に、前記データを格納する、ストレージシステムの制御方法。

【請求項 11】

前記指標とは、計算機がアクセス要求を送信してから完了報告を受けるまでの時間に関連する指標である、

30

請求項 10 に記載のストレージシステムの制御方法。

【請求項 12】

前記ストレージ装置は、前記複数の記憶領域の少なくとも 1 つに、前記記憶デバイスの電源を遮断するモードを設定する、請求項 11 に記載のストレージシステムの制御方法。

【請求項 13】

前記ファイルサーバは、前記計算機から送信されたデータにメタデータを設定し、前記メタデータに前記指標を設定する、請求項 12 に記載のストレージシステムの制御方法。

【請求項 14】

前記ファイルサーバのファイルシステムの前記データが格納されるフォルダに前記指標を設定し、前記計算機から送信されたデータを前記フォルダに格納すると、前記フォルダに設定された指標を前記データの前記メタデータに設定する、

40

請求項 13 に記載のストレージシステムの制御方法。

【請求項 15】

前記ストレージ装置は、最終にアクセスされた時間から、一定の時間が経過した後、前記レディモードになった記憶領域を省電力モードへ移行させる、請求項 12 に記載のストレージシステムの制御方法。

【請求項 16】

前記ファイルサーバは、前記第 1 の記憶領域が、前記省電力モードである場合、前記レディモードである第 2 の記憶領域を選択し、

前記ストレージ装置は、前記第 2 の記憶領域に前記データを格納し、前記第 1 の記憶領

50

域を前記省電力モードから、前記レディモードへ移行し、次いで、前記第 2 の記憶領域から、前記第 1 の記憶領域にデータを移動する、

請求項 10 に記載のストレージシステムの制御方法。

【請求項 17】

前記ファイルサーバは、前記複数の記憶領域と前記複数の記憶デバイスとの関係を示すテーブルに基づいて、前記第 1 の記憶領域を構成する第 1 の記憶デバイスを特定し、前記ストレージ装置に当該第 1 の記憶デバイスへの前記データの格納指示を送信する、

請求項 10 に記載のストレージシステムの制御方法。

【請求項 18】

前記ファイルサーバは、前記テーブルに基づいて、前記第 1 の記憶デバイスの現在の動作モードを認識し、前記第 1 の記憶デバイスが前記省電力モードである場合には、前記ストレージ装置へ前記第 1 の記憶デバイスを前記レディモードに移行する指示を送信する、
請求項 17 に記載のストレージシステムの制御方法。

10

【発明の詳細な説明】

【技術分野】

【0001】

本発明はストレージシステムに係わり、特に、記憶デバイスに対する省電力化機能を備えたストレージシステム及びその制御方法に関するものである。

【背景技術】

【0002】

従来から、ストレージシステムとして、サーバと、複数のハードディスクドライブ（以下：HDD）をアレイ状に配置し、サーバに対して記憶領域を提供するディスクアレイサブシステムと、を備えたものが知られている。

20

【0003】

サーバがディスクアレイサブシステムを利用する用途には複数あり、その 1 つとして、サーバがアーカイブの用途でディスクアレイサブシステムを利用する形態がある。ところで、ディスクアレイサブシステムにおいて、アクセス頻度が少ない HDD を常時オンしておく、省電力の観点から好ましくない。そこで、MAID（Massive Array of Inactive Disks）と称される省電力技術がストレージシステムに適用されている。アクセスには、リード要求、ライト要求などが含まれる。

30

【0004】

MAID とは、ディスクアレイの中で、データアクセスが発生していない HDD のディスクの回転、または HDD の電源を遮断することによって、消費電力を削減するという技術である。MAID を採用したディスクアレイは、データアクセスに必要な HDD だけが稼働するように制御されるため、消費電力を削減することができる。

【0005】

一方、MAID 機能を用いた省電力モードには、HDD が複数個設置されている筐体への電源供給を遮断するモード、HDD への電源供給を遮断するモード、HDD のディスクの回転を停止させるモードなど、省電力機能の程度によっていくつかのモードがあることが知られている。そして、省電力機能の程度が階層化された複数の論理ユニットを備え、論理ユニットに対するアクセス要求の頻度に応じて、格納されたデータを、他の省電力モードを適用する論理ユニットに移動するディスクアレイシステムが知られている（特開 2008 - 90352 号）。

40

【特許文献 1】特開 2008 - 90352 号公報

【発明の開示】

【発明が解決しようとする課題】

【0006】

従来のディスクアレイサブシステムは、サーバから書き込まれるアーカイブ用データが、サーバから高いアクセス性能を求められるオンラインデータであるか、サーバから比較的低いアクセス性能を求められるオフラインデータであるか、サーバからオンラインデー

50

タとオフラインデータとの間のアクセス性能を求められるニアラインデータであるか、をデータ格納前に判断できないため、適切な省電力機能を持ったHDDにデータを格納することができなかった。

【0007】

したがって、従来のディスクレイサブシステムは、例えば、オンラインデータを、省電力効果は高いが応答までの時間がかかる省電力モードである、HDDへの電源供給を遮断するモードを適用したHDDに格納してしまう可能性がある。そして、サーバからオンラインデータへのアクセス要求が発生した時、オンラインデータへのアクセスに対する応答性が低下し、サーバからのアクセスがタイムアウトになってしまうおそれがあった。

【0008】

そこで、本発明は、記憶デバイスを省電力化する機能を備えたストレージシステムであっても、記憶デバイスに対する、サーバからのアクセスに対する応答性が低下しない、ストレージシステム及びその制御方法を提供することを目的とするものである。

【課題を解決するための手段】

【0009】

本発明にかかるストレージシステムはこの目的を達成するために、計算機とネットワークを介して接続するファイルサーバと、前記ファイルサーバとネットワークを介して接続するストレージ装置とを含み、前記ファイルサーバは、第1のコントローラを含み、前記ストレージ装置は、複数の記憶領域を構成する複数の記憶デバイスと、前記複数の記憶領域へのアクセスを制御する第2のコントローラと、を含み、前記複数の記憶領域はそれぞれ、省電力モードからレディモードへの移行時間が異なる複数の省電力モードのうち、少なくとも1つの省電力モードを備え、前記第1のコントローラは、前記計算機からデータを受信すると、前記データに、前記計算機からのアクセスへの応答性能に関する指標を設定し、前記データの前記指標を参照して、前記指標を満たす省電力モードを備える第1の記憶領域を選択し、前記第2のコントローラは、前記第1の記憶領域に、前記データを格納する、ストレージシステムであることを特徴とする。

【0010】

さらに本発明に係わるストレージシステムの制御方法は、計算機とネットワークを介して接続するファイルサーバと、前記ファイルサーバとネットワークを介して接続し、複数の記憶領域を構成する複数の記憶デバイスを備え、前記複数の記憶領域へのアクセスを制御するストレージ装置と、を含む、ストレージシステムの制御方法であって、前記ストレージ装置が、前記複数の記憶領域に、それぞれ、省電力モードからレディモードへの移行時間が異なる複数の省電力モードのうち、少なくとも1つの省電力モードを設定し、前記ファイルサーバが、前記計算機から受信したデータに、前記計算機からのアクセスへの応答性能に関する指標を設定し、前記ファイルサーバが、前記データの前記指標を参照して、前記指標を満たす省電力モードを備える第1の記憶領域を選択し、前記ストレージ装置は、前記第1の記憶領域に、前記データを格納する、ことを特徴とするものである。

【発明の効果】

【0011】

本発明によれば、記憶デバイスを省電力化する機能を備えたストレージシステムであっても、記憶デバイスに対する、サーバからのアクセスに対する応答性が低下しない、ストレージシステム及びその制御方法を提供することができる。

【発明を実施するための最良の形態】

【0012】

次に本発明の実施形態について説明する。図1は、本発明に係るストレージシステムの一例を備えた記憶制御システムのブロック図である。記憶制御システムは、ストレージ装置に対して上位として位置づけられる業務サーバ10と、ファイルサーバ16とストレージ装置30とからなるストレージシステムと、を備えている。なお、ファイルサーバ16とストレージ装置30とは同一の筐体内に設置されていてもよい。

【0013】

業務サーバ１０は、ＣＰＵやメモリ等の情報処理資源を備えたコンピュータ装置である。業務サーバは、キーボードスイッチやポインティングデバイス、マイクロフォン等の情報入力装置と、モニタやスピーカ等の情報出力装置と、を備えている。

【００１４】

さらに、業務サーバ１０は、ファイルサーバ１６のファイルシステムによる記憶構造を認識して、ファイルシステムに対してデータのライト及びリードを実行する。業務サーバ１０は、例えば、アーカイブ用ソフトウェアを備えることにより、クライアントにアーカイブソリューションを提供する。業務サーバ１０は、アーカイブソフトウェア以外の業務アプリケーションを管理クライアントに提供するものでもよい。

【００１５】

ネットワーク１２は、複数の業務サーバそれぞれのＬＡＮインターフェース制御部のポート１４に接続している。また、ファイルサーバ１６のＬＡＮインターフェースのポート１８がネットワーク１２に接続している。業務サーバ１０は、ネットワーク１２を介してファイルサーバ１６に接続する。

【００１６】

ファイルサーバ１６は、既述のポート１８に加えて、マイクロプロセッサ２４、キャッシュメモリ２１、ローカルメモリ２２を備えている。これらの各機器は、接続部２０を介して互いに接続されている。キャッシュメモリ２１は、業務サーバ１０から送られたデータを一時的に保存するメモリである。ローカルメモリ２２は、ファイルサーバ１６の各種制御機能を実現するプログラムや管理情報を格納する。

【００１７】

図２はファイルサーバ１６の各種機能を達成するプログラムの一例である。これらのプログラムは、ローカルメモリ２２に格納される。そして、マイクロプロセッサ２４は、これプログラムを実行して、ファイルサーバの各機能を実現する。ゲートウェイ（Gateway）は、業務サーバ１０との通信に必要な通信プロトコルをマイクロプロセッサ２４に提供する。ファイルシステム（File System）は、業務サーバ１０からのファイルアクセスを認識して、業務サーバ１０からのデータのアクセス先となる記憶領域を決定する。

【００１８】

リクエストマネージャ（Request Manage）は、業務サーバ１０からのアクセスを処理し、さらに、メタデータの作成や更新などの処理を実施する。ストレージマネージャ（Storage Manager）はストレージ装置３０に対するリード・ライトのための処理を実施する。メタデータマネージャ（Metadata Manager）は、メタデータに対するリード・ライトのための処理を実施する。

【００１９】

ファイルサーバ１６はさらに、ファイバチャネルインターフェース制御部のポート２６を備え、ポート２６は、ネットワーク２８に接続している。ネットワーク２８はＳＡＮによって構成されている。ファイルサーバ１６は、ファイバチャネルプロトコルに従って、ＨＤＤが持つ記憶領域の管理単位であるブロックを基準としてデータの入出力をストレージ装置に対して要求する。

【００２０】

ストレージ装置３０は、記憶デバイスとして、ＨＤＤを備える。そして、複数のＨＤＤを備えた、ディスクアレイサブシステムとして構成される。但し、これに限らず、記憶デバイスの一部を、フラッシュメモリなどの半導体メモリから構成することもできる。

【００２１】

ストレージ装置３０は、記憶デバイスとして複数のＨＤＤ５４Ａ－５４Ｄと、ＨＤＤに対するデータの入出力を制御するディスクアダプタ（ＤＫＡ）４０を複数有している。ＤＫＡは、ＨＤＤが接続されるポート４２を有している。ポート４０とＨＤＤとの間は、ファイバチャネルのＦＣ－ＡＬやファブリック、またはＳＡＳなどによって接続されている。

【００２２】

10

20

30

40

50

ストレージ装置 30 は、ファイルサーバ 16 に対するインターフェース制御部となる、複数のチャネルアダプタ (CHA) 32 を備えている。CHA はポート 34 を有しており、このポート 34 がネットワーク 28 に接続されることにより、ストレージ装置 30 がファイルサーバ 16 に接続する。

【0023】

CHA 32 と DKA 40 とは接続部 (接続回路) 36 を介して接続されている。さらに、接続部 36 には、共有メモリ (SM) 38 とキャッシュメモリ (CM) 34 が接続されている。CHA 32 は、マイクロプロセッサ (MP) と、MP がファイルサーバ 16 から送られてくるコマンドを処理するためのマイクロプログラムが格納されたローカルメモリ (LM) と、を備えている。DKA 40 は、MP と、MP が複数の HDD を制御するためのマイクロプログラムを実行するための LM を備えている。

10

【0024】

複数ある CHA 32 および複数ある DKA 40 との間で、連携した処理を行うために、それぞれが共有すべき制御情報は共有メモリ 38 に存在する。CHA 32 は、ファイルサーバ 16 からデータの読出しあるいは書き込みコマンドを受信すると、コマンドを共有メモリ 38 に記憶する。DKA 40 は、共有メモリ 38 を随時参照しており、未処理の読出しコマンドを発見すると、HDD からデータを読み出して、キャッシュメモリ 34 に記憶させる。

【0025】

CHA 32 は、キャッシュメモリ 34 に移されたデータを読み出し、コマンド発行元のファイルサーバ 16 に送信する。CHA 32 が、ファイルサーバからデータの書き込み要求を受信すると、書き込みコマンドを共有メモリ 38 に記憶させると共に、受信データをキャッシュメモリ 34 に記憶させる。DKA 40 は、共有メモリ 38 に記憶されたコマンドに従って、キャッシュメモリ 34 に記憶されたデータを HDD に記憶する。

20

【0026】

DKA 40 は、HDD との間でデータ入出力を行う際、ファイルサーバからのコマンドに付帯した論理的なアドレスを、HDD の物理的なアドレスへ変換する。各 DKA は、HDD の RAID 構成に応じたデータアクセスを行う。

【0027】

各 DKA は、HDD の省電力モードを含む各種のモードを随時監視しており、この監視結果は接続部 36 につながれた LAN インターフェース 48 を介して SV P (サービスプロセッサ) 50 に送信される。SV P 50 は、ストレージ装置の管理及び監視を行うコンピュータ装置 (管理装置) である。SV P は、接続部 36 を介して各 CHA 32 及び DKA 40 等から各種の環境情報や性能情報等を収集する。

30

【0028】

共有メモリ 38 には、ワーク領域が設定されるほか、以下に説明する第 1 の管理テーブル (図 3)、第 2 の管理テーブル (図 4) も格納される。

【0029】

接続部 36 は、例えば、高速スイッチング動作によって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。

40

【0030】

ストレージ装置 30 は、ディスク駆動装置 54 A - 54 D への電源の供給を制御する電源回路 (PDB: Power Distribution Block) 52 に接続された制御信号 44 に接続する制御用インターフェースポート 46 を備えている。電源制御回路 52 は、複数あるディスク駆動装置 54 A - 54 D それぞれへの電源をオン・オフするスイッチ 51 A - 51 D を備えている。

【0031】

各スイッチ 51 A - 51 D は 1 つのディスク駆動装置 54 A - 54 D と接続されており、各ディスク駆動装置は、複数の HDD 60 A - 60 N (62 A - 62 N, 64 A - 64 N, 66 A - 66 N) を備えている。電源制御回路 52 のスイッチが開放されると、この

50

スイッチに接続されたディスク駆動装置に属する全てのHDDへの電源供給が遮断されるため、HDDの電源がオフされる（ディスク駆動装置オフモード）。電源制御回路のスイッチが閉じられている状態では、このスイッチの接続するディスク駆動装置に含まれる複数のHDDに電源を供給することができるため、DKA40によって、各省電力モードに制御される。省電力モードとは、HDDの電源がオフされているモード（HDDオフモード）、HDDのディスクの回転が停止しているモード（スピンドアウンモード）、HDDのディスクの回転が低回転速度のモード（低回転数モード）、HDDのヘッドがディスク上から待避しているモード（ヘッドアンロードモード）などがある。電源制御回路のスイッチの制御は、CHAのMP、又はDKAのMPによって行われる。

【0032】

10

図1において、HDDの“Active”とは、HDDがオンで高回転速度であるモードであり、ライトコマンドに対して直ちにデータを書き込むことができ、或いはリードコマンドに対して直ちにデータを出力できる、レディモードにあることを示している。“Inactive”とはHDDオフモード、スピンドアウンモード、低回転数モード、ヘッドアンロードモード、ディスク駆動装置オフモードなど、HDDが省電力モードにあることを示している。

【0033】

ストレージ装置のDKA40は、HDDをレディモードから省電力モードにするための電源制御プログラムを、DKA40のローカルメモリに備える。そして、DKAのMPがこのプログラムを実行する。さらに、CHA32及びDKA40は、HDDに対するリード又はライト処理のためのマネージャプログラムを実行する。

20

【0034】

HDDが省電力モードであると、ライトコマンドやリードコマンドに応答するために、ストレージ装置は一旦HDDをレディモードに移行しなければならず、ライトコマンド及びリードコマンドに対するHDDの応答性が低下し、タイムアウトになる可能性がある。

【0035】

そこで、図1に示す業務サーバ10からファイルサーバ30に書き込み要求が送信されたデータに、業務サーバ又はファイルサーバは、業務サーバ10からアクセス応答性能の指標としての、許容されるタイムアウト値（Time Out Value）に準拠した指標（以下、この指標をTOV指標、という。）を設定する。一方、少なくとも1つのHDDにより構成される論理ユニットを、この指標に基づいて分類する。例えば、省電力機能としてHDDオフモードを備えているLUは、アクセス性能に関する指標として、比較的長いTOV値として「オフライン」に分類される。そして、ファイルサーバ16は、このデータのTOV指標と、論理ユニットのTOV指標とを比較し、データの指標に合致する論理ユニットのHDDに該データを格納する。

30

【0036】

ストレージ装置又はファイルサーバは、各ライトアクセス又はリードアクセスへの応答に対する許容時間としてTOV値を設定しており、TOV値を過ぎてもHDDから応答が無い場合には、ライトエラー又はリードエラーと判定する。

【0037】

上位装置からアクセスに対して高いアクセス応答性能を求められず、アクセスに対する応答において比較的大きなTOV値（数分程度）が許容されるデータを「オフラインデータ」という。オフラインデータを格納するための記憶デバイスには、省電力モードとして、省電力モードからレディモードへ移行するまでに数分の時間を要する、ディスク駆動装置オフモード、HDDオフモードなどを設定する。これらの記憶デバイスは、TOV指標として、「オフライン」を有する。本実施例では、HDDオフモードを設定する。

40

【0038】

一方、上位装置からのアクセスに対して高いアクセス応答性能を求められ、アクセスに対する応答において許容されるTOV値が小さい（数秒程度）データを、「オンラインデータ」という。オンラインデータを格納するための記憶デバイスには、常時レディモードを設定したり、省電力モードの中でも数秒程度の短時間でレディモードに移行できる、ヘ

50

ッドアンロードモード、低回転モードを設定する。本実施例では、常時レディモードを設定する。これらの記憶デバイスは、T O V 指標として、「オンライン」を有する。

【 0 0 3 9 】

また、ストレージ装置は、求められるアクセス応答性能がオンラインデータとオフラインデータとの中間（数十秒程度）のデータであり、「ニアラインデータ」という。ニアラインデータを格納する記憶デバイスには、省電力モードとして、省電力モードからレディモードに変更するのに数十秒の時間を要する、スピンドアウンモードを設定する。これらの記憶デバイスは、T O V 指標として、「ニアライン」を有する。

【 0 0 4 0 】

なお、オンラインデータ、ニアラインデータ、オフラインデータを格納するための記憶デバイスが備える動作モードは、これに限定されるものではなく、適宜T O V 値と、省電力モードからレディモードに移行する時間との関係で適宜変更することができる。

【 0 0 4 1 】

図 3 は、ディスク駆動装置、H D D、L U の対応関係と、ディスク駆動装置の電源状態、H D D の現在の動作モードと、H D D のサポートする省電力モードとを示す第 1 の管理テーブルである。図 4 は、L U の現在の動作モードと、各 L U のアクセス応答の性能を示すT O V 指標と、各 L U がサポートする省電力モードとを示す、第 2 の管理テーブルである。以下、図面をもとに説明する。

【 0 0 4 2 】

図 3 において、“ディスク駆動装置 #”とは、図 1 のディスク駆動装置 5 4 A - 5 4 D の各々を識別する記号であり、“ディスク駆動装置動作モード”とは、ディスク駆動装置への電源供給がされている（オン）又は電源供給が遮断されている（オフ）モードであり、“H D D”とは各ディスク駆動装置に属するH D D の識別番号であり、“H D D 動作モード”とは各H D D の現在の動作モード（レディモード又は各省電力モード）を示すものであり、“L U”とは論理ユニット（L U）の識別記号であり、“H D D に対する電源制御指定可能モード”とは、各H D D が省電力機能のうち、どの省電力化機能をサポートしているかを示す情報である。“常時ON”とは、H D D の電源が常時オンするモードをサポートしているか否か、“低回転”とは低回転数モードをサポートしているか否か、“スピンドアウン”とはスピンドアウンモードをサポートしているか否か、“OFF”とはH D D オフモードをサポートしているか否かを示している。

【 0 0 4 3 】

“H D D に対する電源制御指定可能モード”において、“○”はH D D が、各省電力機能をサポートしていることを示し、“×”はH D D が各省電力機能をサポートしていないことを示している。

【 0 0 4 4 】

また、各L U をどのH D D により構成するかは、S V P から管理クライアントが指定するようにしてもよいし、ストレージ装置がH D D の情報を基に指定してもよい。この時、同一の省電力モードをサポートしている、複数のH D D によりL U 構成するように設定するのが好ましい。ここで、同一の省電力モードを備えているかどうかは、必ずしも省電力モードの全項目が完全に同一である必要はないため、少なくとも1つの同一の省電力モードを備えているかにより判断される。

【 0 0 4 5 】

ここで、S V P の管理クライアントは、複数のディスク駆動装置をそれぞれ区別して認識でき、かつ、複数のH D D をそれぞれ区別して認識することができる。さらに、D K A 4 0 は、ディスク駆動装置及びH D D からのステータス情報に基づいて、ディスク駆動装置の動作モードと、H D D の省電力モードを認識することができる。

【 0 0 4 6 】

一方、図 4 の“L U #”は、L U の識別記号であり、“L U 動作モード”は、L U を構成するH D D が、レディモードであるか、いずれかの省電力モードであるか、を示している。通常、L U を構成する複数のH D D は同一の動作モードである。例えばディスクスピ

10

20

30

40

50

ンダウンのモードであるLU1を構成するHDDは、図3においてすべてスピンドウンのモードである。“TOV指標”は、省電力モード時に業務サーバよりアクセスコマンドを受領した場合の、アクセス応答性能に関する指標である。“LUに対する電源制御指定可能モード”は、LU#がサポートする省電力化機能の項目である。したがって、LUを構成する複数のHDDがサポートしている省電力機能により決定される。各項目の意味は図3と同じである。なお、“ディスク駆動装置OFF”とは、LUに属するディスク駆動装置がディスク駆動装置オフモードをサポートしているか否かを意味する。

【0047】

ファイルサーバ16が、第2の管理テーブルの“TOV指標”は、LUを構成する複数のHDDの省電力モードの設定により決定される。すなわち、本実施例においては、TOV指標として「オンライン」を設定しようとする際、LUを構成するHDDは、少なくとも常時レディモードを備えていなければならない。同様に、「ニアライン」を設定しようとする際、LUを構成するHDDは、少なくともスピンドウンモードを備えていなければならない。同様に、「オフライン」を設定しようとする際、LUを構成するHDDは、少なくともHDDオフモードを備えていなければならない。

【0048】

図5はメタデータの構造を示す。メタデータは、ファイル毎で、かつファイルの作成・更新毎に作成或いは更新される。また、更新の契機はこれに限定されるものではなく、管理クライアントにより適宜更新してもよいし、一定期間たつと自動に構成されるようにしてもよい。メタデータは、作成日時、保存期間、TOV指標、最終アクセス日時から構成される。TOV指標としては、TOV(〇〇秒、〇〇分等)そのものであってもよいし、或いは、オンライン、オフライン、又はニアラインであってもよい。

【0049】

図6は、ファイルサーバ16が、ストレージ装置30にデータを格納する際の動作概要の一例を示すブロック図である。サーバ10がネットワーク12を介してファイルサーバ16にファイルに格納すると(S1)、ファイルシステムは、データ領域(キャッシュメモリ等)21Aにファイル60を記憶する。一方、ファイルシステムは、ファイル情報からメタデータを作成或いは決定して、これをメタデータ領域62に格納する。メタデータ領域は、ローカルメモリ22に格納される。ファイルサーバ16のストレージマネージャは、ファイルのデータをメタデータに含まれるTOV指標に基づいて、この指標に合致する指標をもつストレージ装置30の論理ボリューム(LU)を指定する(S2)。また、管理クライアントは、ファイルの属性を変更することができ、このファイルの属性の変更がTOV指標に関係する場合には、この変更に基づいて、ファイルのデータを、格納されている記憶領域から、変更後のTOV指標に合致する記憶領域に移動する(S4)。

【0050】

ファイルサーバが、業務サーバよりデータのライトコマンド又はリードコマンドを受信した際に、データを格納するLUを決定する。そして、格納先LUが省電力モードで動作している場合には、ファイルサーバの指示に従い、ストレージ装置が省電力モードにあるHDDをレディモードに移行する。そして、図5の最終アクセス日時を参照し、一定時間経過した後に、ファイルサーバの指示に従い、ストレージ装置がレディモードのHDDを省電力モードに変更する。また、管理クライアントの指示によりレディモードのHDDを省電力モードに変更してもよい。

【0051】

ファイルサーバの指示は、ストレージ装置にコマンドを発行することにより実現される。ファイルサーバが、ストレージ装置に省電力モードへの移行もしくはレディモードへの移行のコマンドを発行する際、記憶領域をLUだけでなく、LUを構成するHDDまで特定し、コマンドを発行する。

【0052】

一方、後述する変形例では、ファイルサーバが、図4に示す第2の管理テーブルのみを保持し、参照することによって、記憶領域をLUとして特定し、該LU対してコマンドを

10

20

30

40

50

発行する。コマンドを受領したストレージ装置は、第 1 の管理テーブルを用いて L U を構成する H D D を決定する。詳細については後述する。

【 0 0 5 3 】

先ず、図 3 に示す第 1 の管理テーブル及び図 4 に示す第 2 の管理テーブルをファイルサーバに設定する動作について説明する。図 7 は、第 1 の管理テーブル及び第 2 の管理テーブルを設定する動作を説明するフローチャートである。

【 0 0 5 4 】

ファイルサーバ 1 6 のリクエストマネージャが、ファイルサーバの管理装置からストレージ装置の L U ディスカバリ要求を受領すると、第 1 の管理テーブル及び第 2 の管理テーブルをストレージ装置 3 0 に要求する (7 0 0)。ストレージ装置の C H A 3 2 がこの要求を受領すると (7 0 2)、共有メモリ 3 8 から第 1 の管理テーブル及び第 2 の管理テーブルを読み出し、第 1 の管理テーブル及び第 2 の管理テーブルをファイルサーバのリクエストマネージャに送信する (7 0 4)。リクエストマネージャは、第 1 の管理テーブル及び第 2 の管理テーブルを受領すると (7 0 6)、ローカルメモリ 2 2 に格納する (7 0 8)。

【 0 0 5 5 】

なお、ファイルサーバが、ストレージ装置より、第 1 の管理テーブル及び第 2 の管理テーブルの各項目に相当する情報を受領し、ファイルサーバにおいて、第 1 の管理テーブル及び第 2 の管理テーブルを作成し、格納してもよい。

【 0 0 5 6 】

H D D の省電力モードのサポート情報は、H D D をストレージ装置に導入する際、或いは H D D を交換、あるいは H D D の増設時に D K A が H D D の製品情報を入手しサポート情報を取得する。なお、管理クライアントが S V P から入力してもよい。

【 0 0 5 7 】

次に、ファイルサーバにおける、データの T O V 指標を基にしたファイルシステムの設定方法について説明する。業務サーバ 1 0 がファイルサーバのファイルシステムにデータを格納すると、図 5 に示したメタデータに T O V 指標が設定される。本実施例でのファイルシステムは、フォルダのディレクトリ構造であるがこれに限定されるものではない。図 8 は、ファイルサーバがファイルシステムにフォルダを作成するフローチャートを示す。先ず、業務サーバ 1 0 がファイルサーバ 1 6 にフォルダ作成のためのコマンドを発行する (8 0 0) と、これを受領したリクエストマネージャ (8 0 2) は、フォルダ作成コマンドに T O V 指標が含まれているか否かをチェックする (8 0 4)。

【 0 0 5 8 】

リクエストマネージャがフォルダ作成コマンドに T O V 指標が含まれていると判定すると (8 0 4 : Y の場合)、第 2 の管理テーブルを確認して、コマンドに含まれている T O V 指標に合致する L U が存在するか否かを判定する (8 0 6)。合致する L U が存在しない場合 (8 0 6 : N の場合) には、業務サーバにエラー表示を行うと共に (8 0 8)、業務サーバに再度フォルダの作成を促してもよい。また、ファイルサーバがコマンドに含まれる T O V 指標に対応する指標を持つ L U を作成するようにしてもよい。一方、合致する L U が存在する場合 (8 0 6 : Y の場合)、対応する L U にフォルダを作成 (8 1 0) し、フォルダのメタデータを作成する (8 1 2)。その後、ファイルサーバのリクエストマネージャが完了報告を発行し (8 1 4)、業務サーバがこれを受領すると処理が終了する (8 1 5)。

【 0 0 5 9 】

一方、フォルダ作成要求コマンドに T O V 指標が含まれていない場合は (8 0 4 : N の場合)、作成しようとするフォルダに対して上位のフォルダがあるか否かを判定する (8 1 6)。上位のフォルダが存在する場合 (8 1 6 : Y の場合) には、作成しようとしているフォルダのメタデータに、上位のフォルダのメタデータの T O V 指標を設定する (8 1 8)。一方、作成しようとするフォルダに上位フォルダが存在しない場合 (8 1 6 : N の場合) には、作成しようとしているフォルダにデフォルトの T O V 指標 (例えばオンライン) を設定し (8 2 0)、フォルダのメタデータを作成する。

【 0 0 6 0 】

業務サーバの管理クライアントが、ファイルを、管理クライアントが目的とする T O V 指標を持ったフォルダに格納すると、ファイルサーバは該フォルダのメタデータに設定された T O V 指標をファイルのメタデータの T O V 指標に反映する。フォルダとして、例えば、オンラインデータ用フォルダ、ニアライン用フォルダ、又はオフライン用フォルダなどがある。

【 0 0 6 1 】

次に、ファイルサーバがファイルのデータをストレージ装置に格納する動作、そして、ストレージ装置の H D D を省電力化する動作について説明する。図 9 は、ファイルサーバ 1 6 が、業務サーバ 1 0 からファイルを受領し、ストレージ装置にファイルを格納する動作のフローチャートである。ファイルサーバ 1 6 のリクエストマネージャが、業務サーバからファイルを受領する (9 0 0) と、データをキャッシュメモリ 2 1 に格納して (9 0 2) 、業務サーバからのライトコマンドに対して完了を報告する (9 0 4) 。

10

【 0 0 6 2 】

リクエストマネージャは、ファイルが格納された、ファイルシステムのフォルダのメタデータを参照して、T O V 指標をファイルのメタデータに設定し (9 0 6) 、このメタデータをメタデータフォルダに保存する。

【 0 0 6 3 】

次に、リクエストマネージャは、ファイルのメタデータの T O V 指標を参照し、この T O V 指標に合致する指標を持った L U を第 2 の管理テーブルから選択する (9 0 8) 。次いで、当該 L U の動作モードを参照して、当該 L U の動作モードが、データを直ちに格納できるレディモードであるか否かを判定する (9 1 0) 。レディモードの場合には、ストレージマネージャは、L U を構成する H D D を第 1 の管理テーブルから決定して、ストレージ装置に対して、データと、データの格納先である H D D の識別情報を送信する。その後、ストレージ装置の D K A は、H D D に対してデータを格納する (9 1 4) 。

20

【 0 0 6 4 】

一方、L U がレディモードでない場合は、ストレージマネージャは、これをレディモードにするための H D D の電源制御動作 (9 1 2) に移行する。図 1 0 は、電源制御動作のフローチャートである。リクエストマネージャは、第 1 の管理テーブルより、該 L U を構成する H D D と、H D D に含まれるディスク駆動装置とを特定する。

30

【 0 0 6 5 】

リクエストマネージャは、第 1 の管理テーブルのディスク駆動装置の動作モードを参照し、動作モードがオンの場合にはディスク駆動装置がレディと判定し、動作モードがオフの場合にはレディでないと判定する (1 0 0 0) 。レディでないと判定する場合、リクエストマネージャは、ディスク駆動装置をレディモードにするコマンドをストレージ装置の電源制御プログラムに要求する (1 0 0 2) 。この要求を受け取ったストレージ装置は、ディスク駆動装置をレディモードにする処理を実行する回数の上限值を超えたか否かを判定するための制御記号 (X) に 0 を設定し (1 0 0 4 , 1 0 0 6) 、ディスク駆動装置に接続されるスイッチをオンする指示を電源制御回路 5 2 に出力する (1 0 0 8) 。

【 0 0 6 6 】

電源制御プログラムは、ディスク駆動装置の電源がオンされたか否かを判定し (1 0 0 0) 、オンされていない場合には、制御記号 (X) をインクリメントし (1 0 0 6) 、次いでインクリメント後の制御記号が最大値 Y を以上か否かを判定する (1 0 0 8) 。この判定を否定する場合は、再度ディスク駆動装置の電源をオンする。インクリメント後の制御記号が Y を越える場合は、ディスク駆動装置のレディモードにする処理の上限值を超えたため、ファイルサーバに障害通知を行う (1 0 2 0) 。

40

【 0 0 6 7 】

L U に対応するディスク駆動装置が複数存在する場合 (例えば、第 1 の管理テーブル中の L U 1 のような場合) は、リクエストマネージャ及び電源管理プログラムは、全てのディスク駆動装置の電源がオンであるかを判定し、電源がオンでない場合には、ディスク駆

50

動装置の電源をオンする動作を行う。

【 0 0 6 8 】

ストレージ装置の電源制御プログラムが、ディスク駆動装置の電源のオン処理が成功したと判定する場合には、電源制御プログラムは、第 1 の管理テーブルのディスク駆動装置の動作モードをオフからオンに更新し (1 0 1 2)、完了報告 (1 0 1 4) をリクエストマネージャに送信する。リクエストマネージャは完了報告を受領した後、同様にローカルメモリに格納されている第 1 の管理テーブル及び第 2 の管理テーブルを更新する (1 0 2 2 , 1 0 2 4)。

【 0 0 6 9 】

リクエストマネージャが、ディスク駆動装置がレディモードであると判定する場合 (1 0 0 0)、もしくはディスク駆動装置の電源オン処理を完了した場合、L U を構成する全ての H D D について、アクセス要求への応答が可能なモードであるレディモードであるかを判定する (1 0 2 6)。この判定を否定した場合には、少なくともレディモードに無い H D D をレディモードに移行する要求 (1 0 2 8) をストレージ装置の電源制御プログラムに送信する。この要求を受領したストレージ装置の電源制御プログラムは (1 0 3 0)、ディスク駆動装置の電源をオンする動作と同様に、複数の全ての H D D をレディモードにした後、管理テーブルを更新し、これをリクエストマネージャに送信する (1 0 3 0 - 1 0 4 6)。リクエストマネージャは第 1 の管理テーブルを更新した後 (1 0 4 8 , 1 0 5 0)、図 9 のデータ格納処理 (9 1 4) に移行する。

【 0 0 7 0 】

図 9 と図 1 0 のフローチャートによって、ファイルサーバが、業務サーバから受領したデータの T O V 指標に応じて、当該指標のアクセス応答性能を満たしつつ、十分な省電力モードを備える L U にデータを格納することができる。即ち、オンラインデータは、常時レディモードが設定された L U に格納され、ニアラインデータは、スピンダウンモードが設定された L U に格納され、オフラインデータは、H D D オフモードが設定された L U に格納される。

【 0 0 7 1 】

したがって、ファイルサーバは業務サーバから高いアクセス応答性能 (小さい T O V 指標) を求められるオンラインデータのリード要求があった場合でも、オンラインデータが格納されている L U を構成する H D D は数秒以内にアクセス応答することができ、オンラインデータを素早く業務サーバに送ることができる。一方、ファイルサーバは業務サーバからニアラインデータ或いはオフラインデータのリード要求があっても、比較的長い T O V の範囲内で、これらデータを業務サーバに確実に送ることができる。したがって、T O V 値などのデータのアクセス応答性能に応じて、適当な省電力機能を備えた L U にデータを格納することができるため、業務サーバから求められるアクセス性能を満足しつつ、十分な省電力効果を得ることのできる省電力モードを使用し消費電力の削減を図ることができる。

【 0 0 7 2 】

次に、図 1 1 は、ストレージ装置の H D D をレディモードから省電力モードにするためのフローチャートである。ファイルサーバのリクエストマネージャは、このフローチャートを所定の時間間隔で繰り返し実行する。リクエストマネージャはメタデータフォルダを走査して、ファイルへの最終アクセス時間が所定時間を越えるメタデータがあるか否かをチェックする (1 1 0 0)。

【 0 0 7 3 】

ファイルへの最終アクセス時間から所定時間を越えるメタデータが存在すると判定した場合、メタデータの T O V 指標がオンラインか否かを判定する (1 1 0 2)。T O V 指標がオンラインの場合は、ディスク駆動装置の電源や、H D D の電源や回転数を制御する必要がないので処理を終了する。

【 0 0 7 4 】

T O V 指標がオンラインでない場合には、T O V 指標がニアラインか否かを判定する。

ニアラインでない場合には、リクエストマネージャは、HDDへの電源供給を遮断させるために、ファイルが格納されたフォルダに対応するLUを特定し、第1の管理テーブルを参照して、LUに対応する複数のHDDへの電源供給を遮断する要求(1106)を、ストレージ装置の電源制御プログラムに対して送信する。

【0075】

ストレージ装置の電源制御プログラムは、HDDへの電源供給を遮断する要求を受けると(1110)、DKAは、HDDの電源をオフするために、HDDのモードを変更し(1114)、第1の管理テーブル及び第2の管理テーブルの動作モードを更新し(1114)、完了報告をリクエストマネージャに対して行う(1116)。完了報告を受領した、リクエストマネージャは第1の管理テーブル及び第2の管理テーブルの動作モードを更新して処理を終了する(1118、1120)。

10

【0076】

一方、メタデータのTOV指標がニアラインの場合には、ストレージマネージャは、複数のHDDをスピンドアウンモードにするための要求(1108)を電源管理プログラムに送信する。この要求を受けた電源管理プログラムは(1110)、HDDの回転速度を下げるべく、DKAを介してHDDに対して省電力モードを変更するコマンドを発行する。このコマンドを受けることにより、スピンドアウンモードでないHDDはスピンドアウンモードに変更し(1112)、電源制御プログラムは第1の管理テーブル及び第2の管理テーブルの動作モードをスピンドアウンモードに更新する(1114)。ファイルサーバも同様に管理1の管理テーブル及び第2の管理テーブルの動作モードを更新する(1116、1120)。

20

【0077】

以上のように、ファイルサーバ及びストレージ装置は、一定期間アクセスがない場合には、レディモードから、設定された省電力モードに変更することができる。

【0078】

次に、メタデータのTOV指標が変更された場合での、ファイルサーバの処理動作について、図12に基づいて説明する。LUに格納されているデータに対応するメタデータのTOV指標が変更された場合、業務サーバがファイルサーバにファイル移動要求を発行する。なお、メタデータのTOV指標が変更された場合にファイルサーバが自動的に1200以降の処理をするようにしてもよい。

30

【0079】

ファイルサーバがファイルを格納されているフォルダから他のフォルダに移動する要求を受領すると(1200)、移動元フォルダと移動先フォルダの情報から、移動元LUの特定情報(1202)と、移動先LUの特定情報(1204)と、を認識する。ここで、ファイルの移動は、例えば、ニアライン・オフライン用のフォルダからオンライン用のフォルダに移動する場合であるか、又はその逆であり、それに伴って格納先LUを変更する処理である。

【0080】

次いで、リクエストマネージャは、第2の管理テーブルを用いて、移動元LUと移動先LUのそれぞれがレディモードか否かを判定し(1206、1208)、レディモードでない場合には、LUを構成する全てのHDDをレディモードになるように既述の電源制御フロー(図10)を実行する(1210、1212)。

40

【0081】

次いで、リクエストマネージャの要求に基づいて、ストレージ装置のDKAは、移動元LUのデータをHDDから読み出して、キャッシュメモリ34に一時記憶し(1214)、一時記憶したデータを移動先LUの記憶領域に格納し(1216)、移動元LUの記憶領域に格納されたデータを削除する(1218)。そして、リクエストマネージャは、移動元データのメタデータをフォルダから削除する(1220)。

【0082】

既述の実施形態は、ファイルサーバが、第1の管理テーブル及び第2の管理テーブルを

50

保持しているため、ストレージ装置のＬＵと、このＬＵに対応するＨＤＤを認識することができる。そして、業務サーバがデータにアクセスする傾向をデータに対して設定し、この傾向に最適な省電力特性を持ったＨＤＤにデータを格納する。

【００８３】

上記の実施例に対して以下に述べる変形例では、ファイルサーバは、第１の管理テーブルを保持せず、第２の管理テーブルを保持するため、記憶領域としてＬＵを特定することはできるが、ＬＵを構成するＨＤＤを特定することはできない。ファイルサーバは、ＬＵに対してレディモードへの移行や省電力モードへの移行を要求すれば、ストレージ装置がＬＵを構成するＨＤＤを特定して制御する。この構成により、ファイルサーバの負荷を低減することができる。また、複数のストレージ装置がファイルサーバに接続する場合には、処理を分散できるため、よりストレージシステム全体の性能を向上させることができる。以下に、具体的に処理を説明する。

10

【００８４】

まず、ファイルサーバが、第２の管理テーブルを設定する動作を説明する。図７のフローと異なる点は、ファイルサーバとストレージ装置間で送受信するデータに、第１の管理テーブルは含まれず、第２の管理テーブルである点である。

【００８５】

具体的には、ファイルサーバ１６のリクエストマネージャが、ファイルサーバの管理装置からストレージ装置のＬＵディスカバリ要求を受領すると、第２の管理テーブルをストレージ装置３０に要求する（７００）。ストレージ装置のＣＨＡ３２がこの要求を受領すると（７０２）、共有メモリ３８から第２の管理テーブルを読み出し、第２の管理テーブルをファイルサーバのリクエストマネージャに送信する（７０４）。リクエストマネージャは、第２の管理テーブルを受領すると（７０６）、ローカルメモリ２２に格納する（７０８）。

20

【００８６】

なお、ファイルサーバが、ストレージ装置より、第２の管理テーブルの各項目に相当する情報を受領し、ファイルサーバにおいて、第２の管理テーブルを作成し、格納してもよい。

【００８７】

フォルダの作成処理は、図８のフローと同様である。ファイルサーバがファイルシステムにフォルダを作成するフローチャートは、図９と同様である。ただし、電源制御処理（９１２）は、下記の図１３の電源制御処理を用いる。また、データ格納（９１４）時には、ファイルサーバのリクエストマネージャは、格納先ＬＵを指定し、ストレージ装置においてＬＵを構成するＨＤＤを特定する。

30

【００８８】

図１３は、電源制御処理のフローチャートであり、既述の図１０の変形例である。既述の実施形態では、ＬＵが含まれるディスク駆動装置と、ＬＵを構成するＨＤＤとをファイルサーバ側で認識して、ストレージ装置がファイルサーバから指示に基づいてディスク駆動装置やＨＤＤの動作モードを制御する。これに対して、図１３では、ファイルサーバは電源制御を行うＬＵを指定してストレージ装置にコマンドを発行する（１３００）。そして、コマンドを受領したストレージ装置は、ストレージ装置が保持する第１の管理テーブルに従い、ＬＵが属するディスク駆動装置を特定する。そして、ストレージ装置のＤＫＡは、ディスク駆動装置がレディモードであるかを判定する（１０００）。判定後の処理は、図１０と同様である（１００６ - １０１８）。次に、ストレージ装置が保持する第１の管理テーブルに従い、ＬＵを構成するＨＤＤを特定する。そして、ストレージ装置は、ＬＵを構成するＨＤＤがレディモードであるかを判定する（１０２６）。判定後の処理は、図１０と同様である（１０３２ - １０４６）。そして、ファイルサーバが、ストレージ装置から完了報告を受信すると、ファイルサーバは第２の管理テーブルの動作モードを更新する（１０５０）。

40

【００８９】

50

図 1 4 は図 1 1 で示されているストレージ装置の HDD を省電力モードに移行するフローチャートの変形例である。図 1 1 では、ファイルサーバのリクエストマネージャが、データの格納先である LU と、LU を構成する HDD を特定して、ストレージ装置に省電力モードへの移行要求をする。一方、図 1 5 では、ファイルサーバのリクエストマネージャが、データの格納先である LU を特定し、ストレージ装置に対してして省電力モードへの移行要求をする。

【 0 0 9 0 】

1 1 0 0、1 1 0 2、1 1 0 4 の処理は図 1 1 と同様である。

【 0 0 9 1 】

次に、ファイルサーバのリクエストマネージャが、ストレージ装置の D K A に対し、LU の動作モードをオフする要求を送信する (1 4 0 0)。D K A が LU の動作モードをオフする要求を受領すると (1 4 0 2)、第 1 の管理テーブルに従い LU を構成する HDD を特定し、該 HDD の電源をオフする要求を行う (1 4 0 4)。そして、ファイルサーバに完了を報告する (1 4 1 8)。この完了を受けて、ファイルサーバのリクエストマネージャは、第 2 の管理テーブルの動作モードを更新する (1 4 2 0、1 4 2 2)。一方、ストレージ装置は、HDD のモードを HDD オフモードに変更し (1 4 0 8)、第 1 の管理テーブル及び第 2 の管理テーブルを更新した後 (1 4 1 0)、HDD が搭載されているディスク駆動装置の全ての HDD がオフされたか否かを判定し (1 4 1 2)、これを肯定した場合、これら HDD を有するディスク駆動装置の電源をオフし、これを第 1 の管理テーブルに更新登録する (1 4 1 4、1 4 1 6)。

10

20

【 0 0 9 2 】

また、ファイルサーバのリクエストマネージャがデータを格納する LU に、HDD のスピンドアウンモードに変更する要求 (1 4 2 4) をストレージ装置に送信すると、ストレージ装置の D K A がこの要求を受領する (1 4 2 6)。受領後に、D K A はファイルサーバに対し完了報告を送信し、ファイルサーバは、第 2 の管理テーブルを更新登録する (1 4 3 0)。ストレージ装置の D K A が LU に対するスピンモードへの移行要求を受けると、第 1 の管理テーブルをもちいて LU を構成する HDD を特定し、LU を構成する HDD をスピンドアウンモードにする要求を行う (1 4 3 4)。HDD をスピンドアウンモードへ変更させた (1 4 3 8) 後、共有メモリの第 1 の管理テーブルと第 2 の管理テーブルを更新登録する (1 4 4 0)。

30

【 0 0 9 3 】

次に、データ格納の方法の変形例について以下に説明する。これらの変形例は、上記 2 つの実施形態のいずれにも適用可能である。

【 0 0 9 4 】

図 1 5 は、図 9 の変形例であって、業務サーバからファイルサーバに送られたファイルデータを目的の HDD に格納するためのフローチャートである。

【 0 0 9 5 】

ステップ 9 0 0 - 9 0 6 までは図 9 と同じであるので、同一の符号を付した。ステップ 9 0 8 において、ファイルサーバのリクエストマネージャは、第 2 の管理テーブルを参照し、LU に対する T O V 指標がオンラインか否かをチェックする (1 5 0 0)。

40

【 0 0 9 6 】

リクエストマネージャが、T O V 指標がオンラインであると判定すると、データ格納処理 (1 5 0 2) に移行する。データ格納処理においては、ファイルサーバのローカルメモリが第 1 の管理テーブル及び第 2 の管理テーブルを格納している場合、ファイルサーバが格納先 HDD を特定して、ストレージ装置に指示を出す。ファイルサーバのローカルメモリに第 2 の管理テーブルが格納されている場合、ファイルサーバが格納先 LU を特定して、ストレージ装置に指示を出し、ストレージ装置の D K A が LU の HDD を特定しデータを格納する。

【 0 0 9 7 】

ステップ 1 5 0 0 において、マネージャプログラムは、T O V 指標がオンラインでない

50

場合には、L Uがレディモードか否かを判定し(1504)、レディモードである場合には、上記のデータ格納処理(1502)に移行する。一方、レディモードでないと判定された場合には、HDD電源制御処理(1506)に移行する。HDD電源制御処理においては、ファイルサーバのローカルメモリが第1の管理テーブル及び第2の管理テーブルを格納している場合、図10の電源制御処理を実行する。ファイルサーバのローカルメモリに第2の管理テーブルが格納されている場合、図13の電源制御処理を実行する。その後、上記のデータ格納処理(1502)に移行する。このように、L Uに対するTOV指標がオンラインデータであるかを判断することにより、既にレディモードであるかをチェックすることができ、効率的なデータ格納を実行できる。

【0098】

10

図16は、図9の変形例として、さらに他の実施形態を示すものである。ステップ900 - 908までは図9と同じであるので、同一の符号を付した。その後の処理を、ファイルサーバのローカルメモリが第1の管理テーブル及び第2の管理テーブルを格納している場合、とファイルサーバのローカルメモリが第2の管理テーブルを格納している場合、とに分けて以下に説明する。

【0099】

ファイルサーバのローカルメモリが第1の管理テーブル及び第2の管理テーブルを格納している場合には、ファイルサーバのリクエストマネージャが、データを格納するL Uが属するディスク起動装置がレディモードであるかどうかを判断する(1600)。レディモードでない場合には、他のディスク駆動装置に他のレディモードのL Uがあるかを判定する(1606)。レディモードのL Uがない場合には、図15の1506以降の処理をして、格納先L Uを構成するHDDをレディモードに移行する。レディモードの他のL Uがある場合には、他のL Uに格納するため、ファイルサーバが第1の管理テーブルを用いて格納先HDDを特定して、ストレージ装置にデータ格納指示を出す(1608)。その後、データ格納とは非同期に、ファイルの格納先であるL Uをレディモードに移行し、ファイルを他のL Uから格納先のL Uへと移行する(1610)。

20

【0100】

データを格納するL Uが属するディスク起動装置がレディモードの場合(1600)、ファイルサーバのリクエストマネージャが、第1の管理テーブルを用いて格納先L Uを構成するHDDがレディ状態であるかを確認する。レディ状態でない場合には、上記の1606の処理に移行する。レディ状態である場合には、格納先L Uにデータを格納するため、ファイルサーバが第1の管理テーブルを用いて格納先HDDを特定して、ストレージ装置にデータ格納指示を出す。

30

【0101】

次に、ファイルサーバのローカルメモリが第2の管理テーブルを格納している場合について説明する。ストレージ装置のDKAが、データを格納するL Uが属するディスク起動装置がレディモードであるかどうかを判断する(1600)。レディモードでない場合には、他のディスク駆動装置に他のレディモードのL Uがあるかを判定する(1606)。レディモードのL Uがない場合には、図15の1506以降の処理をして、格納先L Uを構成するHDDをレディモードに移行する。レディモードの他のL Uがある場合には、他のL Uに格納するため、ストレージ装置が第1の管理テーブルを用いて格納先HDDを特定してデータ格納する(1608)。その後、データ格納とは非同期に、ファイルの格納先であるL Uをレディモードに移行し、ファイルを他のL Uから格納先のL Uへと移行する(1610)。

40

【0102】

データを格納するL Uが属するディスク起動装置がレディモードの場合(1600)、ストレージ装置のDKAが、第1の管理テーブルを用いて格納先L Uを構成するHDDがレディ状態であるかを確認する。レディ状態でない場合には、上記の1606の処理に移行する。レディ状態である場合には、格納先L Uにデータを格納するため、ストレージ装置のDKAが第1の管理テーブルを用いて格納先HDDを特定して、ストレージ装置にデ

50

ータ格納指示を出す。

【 0 1 0 3 】

上記の 1 6 1 0 におけるファイル移行処理には、図 1 7 のファイル管理テーブルを用いる。ファイル管理テーブルは、図 1 7 に示すように、ファイル名と、ニアラインデータ・オフラインデータが本来格納されるべき、ニアライン又はオフライン用のフォルダ 1 の情報(或いは、このフォルダに対応する L U とその論理アドレス)、及び、これらデータが仮に格納されているオンライン用のフォルダ 2 の情報(或いは、このフォルダに対応する L U とその論理アドレス)から構成される。

【 0 1 0 4 】

ファイル管理テーブルはファイルサーバの L M (2 2) に格納されている。ファイルサーバのストレージマネージャは、ファイル管理テーブルを定期的にレビューし、例えば、ファイル名が所定値以上になった際に、オンライン用の L U からニアライン又はオフライン用の L U にデータを移動させる。或いは所定時間経過後にこの処理を行ってもよい。

【図面の簡単な説明】

【 0 1 0 5 】

【図 1】本発明に係るストレージシステムを備える記憶制御システムのブロック図である。

【図 2】ファイルサーバの各種プログラムを説明するブロック図である。

【図 3】第 1 の管理テーブルである。

【図 4】第 2 の管理テーブルである。

【図 5】メタデータの構造を示すブロック図である。

【図 6】ファイルサーバがストレージ装置にデータを格納する際の動作の概要を説明するブロック図である。

【図 7】管理テーブルのファイルサーバへの設定を説明するブロック図である。

【図 8】ファイルサーバがフォルダをファイルシステムに作成する際のフローチャートである。

【図 9】業務サーバから受信したデータをファイルサーバがストレージ装置に格納する動作を説明するフローチャートである。

【図 1 0】電源制御動作を説明するフローチャートである。

【図 1 1】ストレージ装置の記憶デバイスを省電力モードにするためのフローチャートである。

【図 1 2】メタデータの T O V 指標が変更された場合での、ファイルサーバの処理動作を説明するフローチャートである。

【図 1 3】図 1 0 の変形例であって電源制御動作を説明するフローチャートである。

【図 1 4】図 1 1 の変形例であって、ストレージ装置の記憶デバイスを省電力モードにするためのフローチャートである。

【図 1 5】図 9 の変形例であって、業務サーバから受信したデータをファイルサーバがストレージ装置に格納する動作を説明するフローチャートである

【図 1 6】図 9 の変形例であって、業務サーバから受信したデータをファイルサーバがストレージ装置に格納する動作を説明するフローチャートである

【図 1 7】ファイルの移動を管理するテーブルである。

【符号の説明】

【 0 1 0 6 】

1 0 計算機、1 2 ネットワーク、1 6 ファイルサーバ、2 0 接続部、2 1 キャッシュメモリ、2 2 ローカルメモリ、2 4 マイクロプロセッサ、3 0 ストレージ装置、3 2 チャンネルアダプタ、3 4 キャッシュメモリ、3 6 接続部、4 0 ディスクアダプタ、5 2 Power Distribution Block、5 4 A - 5 4 D ディスク駆動装置、6 0 ハードディスクドライブ。

10

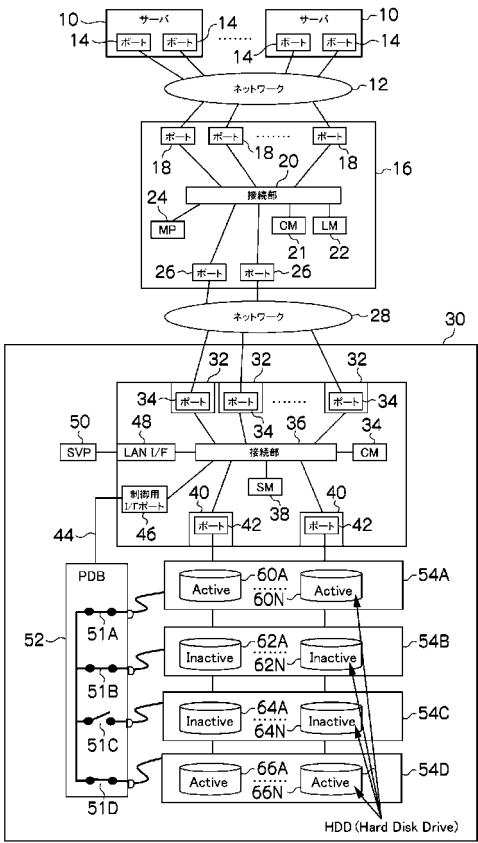
20

30

40

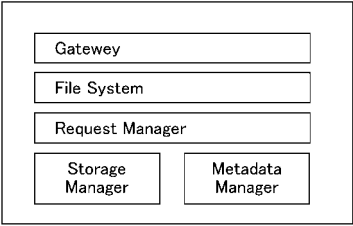
【 図 1 】

図1



【 図 2 】

図2



【 図 3 】

図3

ディスク 駆動装置 #	ディスク 駆動装置 電源状態	HDD#	HDD 動作 モード	LU#	HDDに対する 指定可能モード			
					常時 ON	低回転	スピン ダウン	OFF
01	ON	1	レディ	0	○	×	×	×
		2	レディ	0	○	×	×	×
		...	...	...	...	...	...	...
		14	レディ	0	○	×	×	×
02	ON	15	スピンダウン	1	○	○	○	×
		1	スピンダウン	1	○	○	○	×
		2	スピンダウン	1	○	○	○	×
		...	...	...	...	...	...	...
		10	スピンダウン	1	○	○	○	×
		11	OFF	2	○	○	○	○
		12	OFF	2	○	○	○	○
...	...	...	...	...	...	...	...	...
		15	OFF	2	○	...	...	○

【 図 4 】

図4

LU#	LU動作 モード	TOV指標	LUに対する 電源制御指定可能モード			
			低回転	スピン ダウン	OFF	ディスク 駆動装置 OFF
0	レディ	オンライン	×	×	×	×
1	スピンダウン	ニアライン	○	○	○	×
2	OFF	オフライン	○	○	○	○
3	レディ	オンライン	×	×	×	×
4	レディ	オンライン	×	×	×	×
...	...	...	...	...	...	...

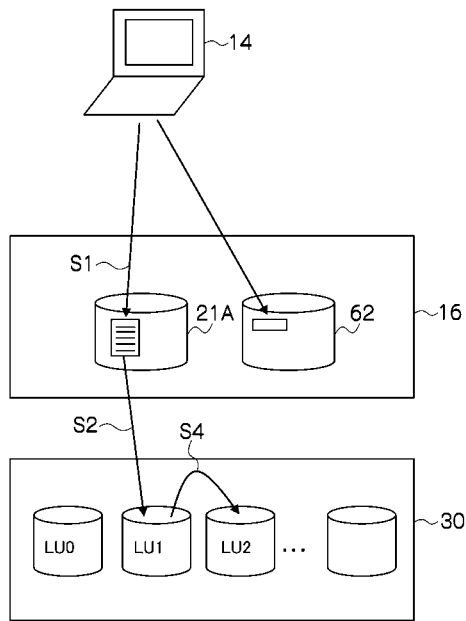
【 図 5 】

図5

作成日時	保存期間	TOV指標	最終アクセス日時
xxxx/xx/xx yy:yy:yy	xxxx/xx/xx yy:yy:yy	オンライン	xxxx/xx/xx yy:yy:yy

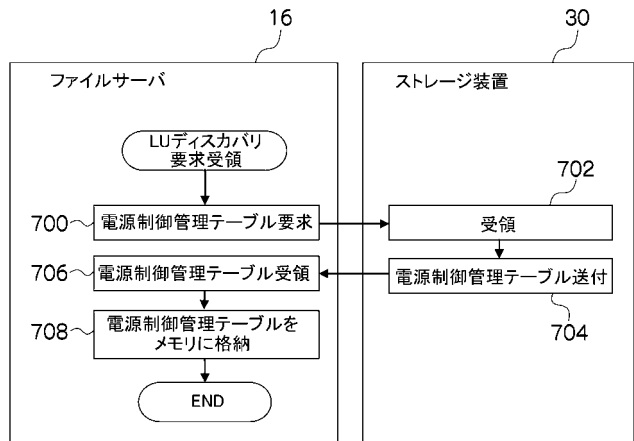
【図 6】

図6



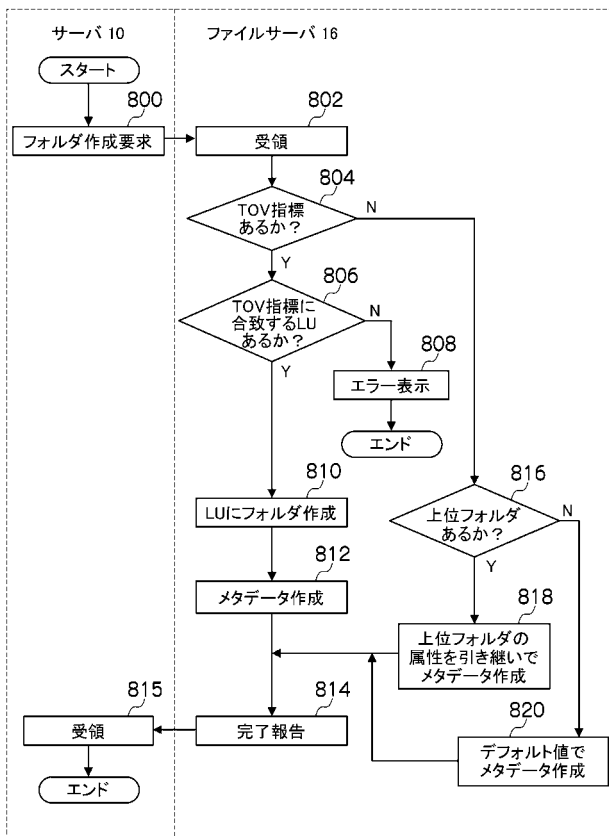
【図 7】

図7



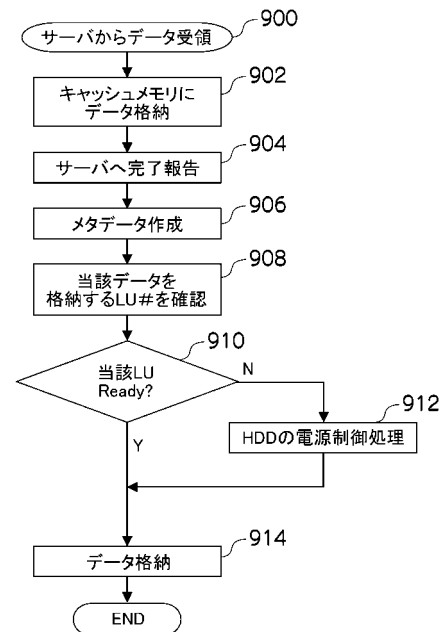
【図 8】

図8



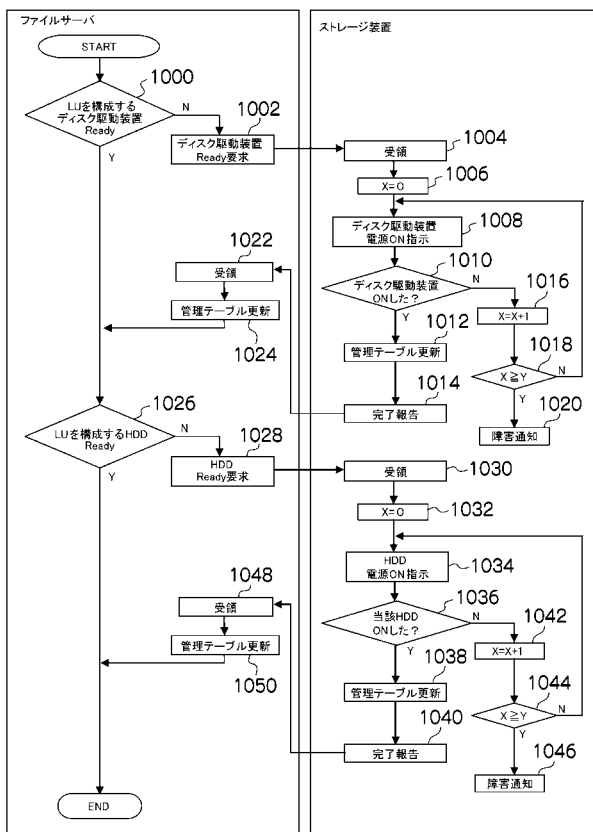
【図 9】

図9



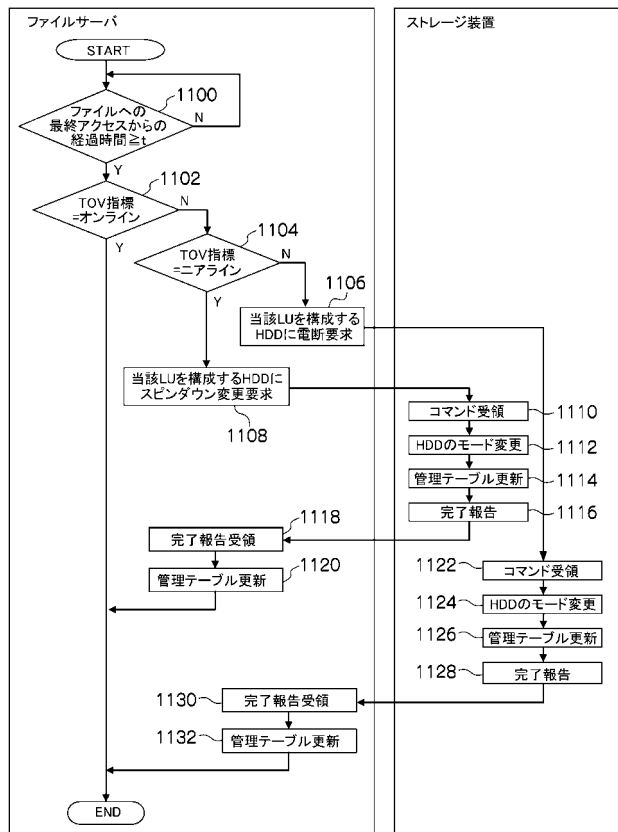
【図 10】

図10



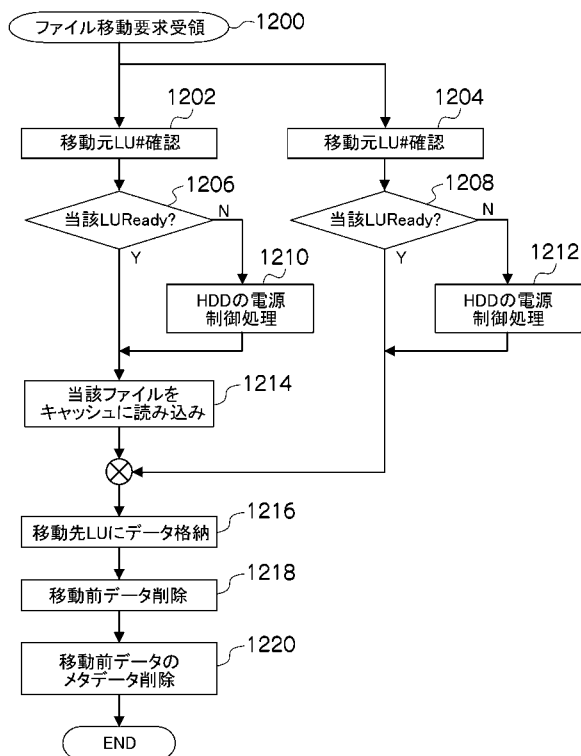
【図 11】

図11



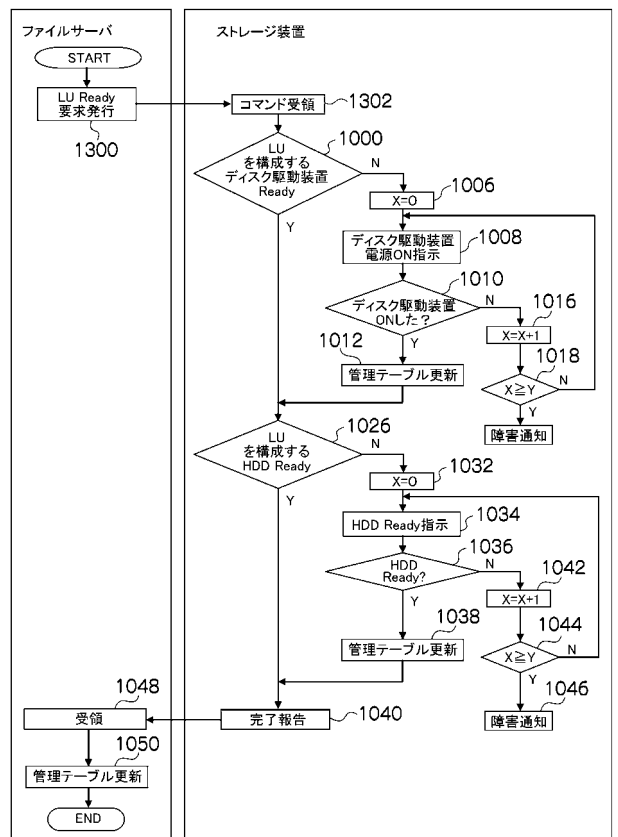
【図 12】

図12



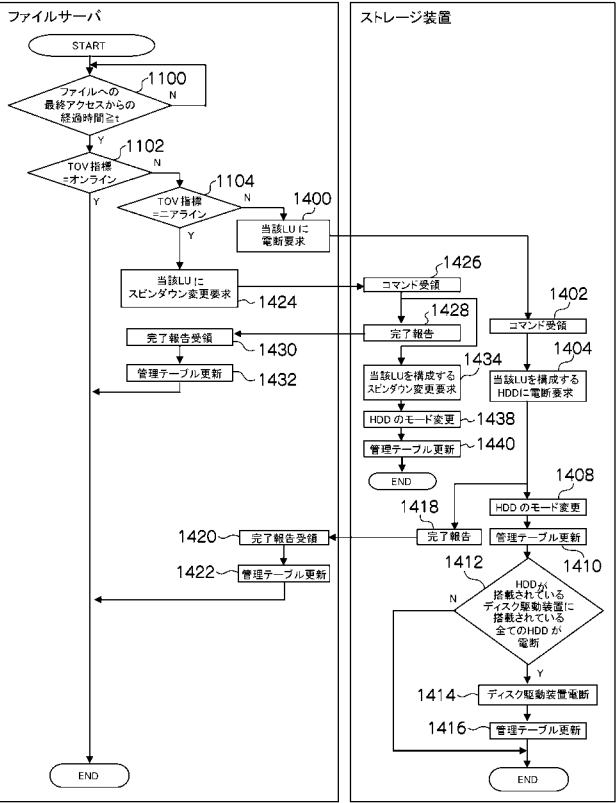
【図 13】

図13



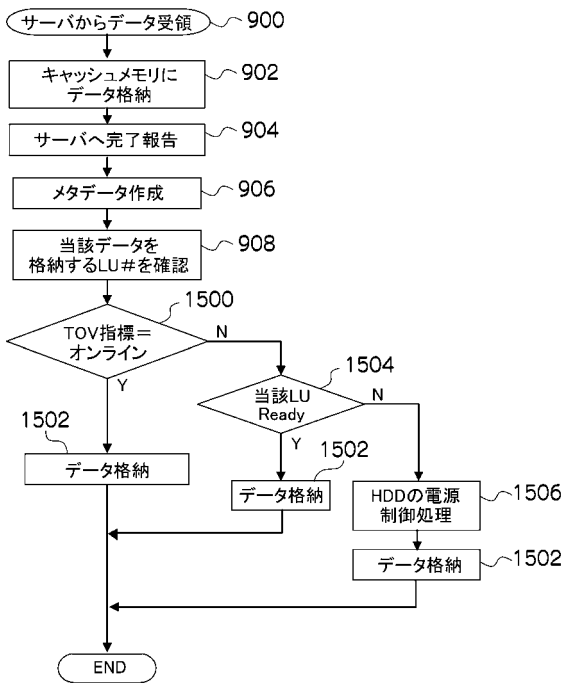
【図 14】

図14



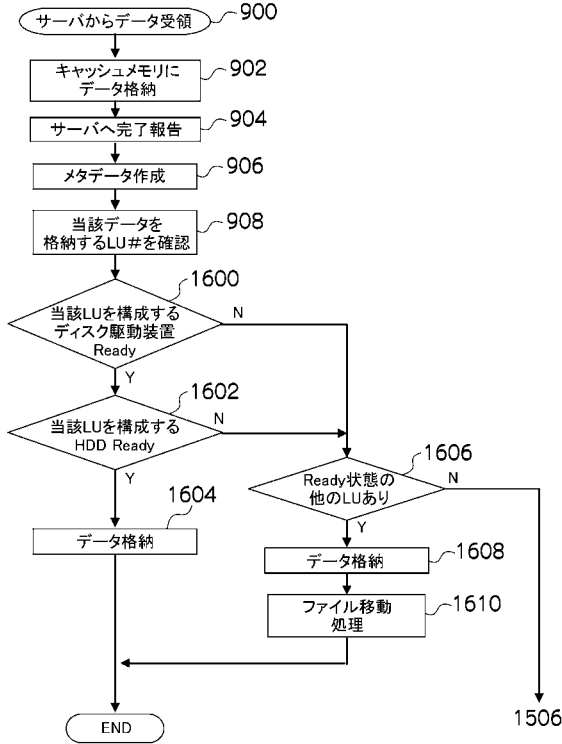
【図 15】

図15



【図 16】

図16



【図 17】

図17

ファイル名	フォルダ1	フォルダ2
xxxxxxx.xxx	¥¥¥¥¥¥¥¥¥¥¥¥¥¥	¥¥¥¥¥¥¥zzz¥zzz
xxxxxxx.xxx	¥¥¥¥¥¥¥¥¥¥¥¥¥¥	¥¥¥¥¥¥¥zzz¥zzz
...	...	...
...	...	...

フロントページの続き

(72)発明者 田淵 英夫

神奈川県小田原市中里 3 2 2 番 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) 5B011 EB07 LL06 LL14

5B065 BA01 CA30 ZA14

5B082 CA11