

【特許請求の範囲】**【請求項 1】**

不揮発性の第 1 記憶装置および不揮発性の第 2 記憶装置を有する情報処理装置であって

、
データの書き込み指示を受け付ける第 1 通信手段と、
前記書き込み指示を前記第 1 記憶装置に送信する第 2 通信手段と、
前記書き込み指示を前記第 2 記憶装置に送信する第 3 通信手段と、
前記第 2 通信手段を介する前記第 1 記憶装置へのデータの送信および前記第 3 通信手段を介する前記第 2 記憶装置へのデータの送信を制御する制御手段と、
を有し、

10

前記第 1 記憶装置から前記書き込み指示に対する応答を受け付け、且つ、前記第 2 記憶装置から前記書き込み指示に対する応答を受け付けないことに基づいて、前記制御手段は、前記第 1 記憶装置への前記データの送信を継続し、前記第 2 記憶装置への前記データの送信を停止し、

前記第 1 記憶装置による前記データの記憶の完了通知を前記第 2 通信手段が受け付けることに基づいて、前記第 1 通信手段は、前記書き込み指示に対する処理の完了通知を送信することを特徴とする情報処理装置。

【請求項 2】

前記第 1 記憶装置への前記データの記憶の失敗通知を前記第 2 通信手段が受け付けることに基づいて、前記第 1 通信手段は、前記書き込み指示に対する処理の失敗通知を送信することを特徴とする請求項 1 に記載の情報処理装置。

20

【請求項 3】

前記第 1 記憶装置から前記書き込み指示に対する応答を受け付けることによって第 1 タイマーのカウントを行うカウント手段を有し、

前記第 2 記憶装置から前記書き込み指示に対する応答を受け付けないまま前記第 1 タイマーが第 1 閾値を超えることによって、前記制御手段は、前記第 1 記憶装置への前記データの送信を継続し、前記第 2 記憶装置への前記データの送信を停止することを特徴とする請求項 1 または 2 に記載の情報処理装置。

【請求項 4】

前記第 1 タイマーが第 1 閾値を超える前に前記第 2 記憶装置から前記書き込み指示に対する応答を受け付けることによって、前記カウント手段は、前記第 1 タイマーを解除することを特徴とする請求項 3 に記載の情報処理装置。

30

【請求項 5】

前記第 1 タイマーが第 1 閾値を超える前に前記第 2 記憶装置から前記書き込み指示に対する応答を受け付けることによって、前記制御手段は、前記第 1 記憶装置および前記第 2 記憶装置への前記データの送信を継続することを特徴とする請求項 3 または 4 に記載の情報処理装置。

【請求項 6】

前記カウント手段は、前記第 2 通信手段および前記第 3 通信手段が前記データの書き込み指示を送信することによって、第 2 タイマーのカウントを開始し、

40

前記第 2 タイマーが第 2 の閾値を超える前に、前記第 2 の通信手段は、前記第 1 記憶装置による前記完了通知を受け付けることによって、前記第 1 通信手段は、前記完了通知を送信することを特徴とする請求項 3 乃至 5 のいずれか 1 項に記載の情報処理装置。

【請求項 7】

前記第 2 タイマーが前記第 2 の閾値を超えた際に、前記第 2 の通信手段が前記第 1 記憶装置による前記完了通知を受け付けていないことによって、前記第 1 の通信手段は、前記書き込み指示に対する処理の失敗通知を送信することを特徴とする請求項 6 に記載の情報処理装置。

【請求項 8】

前記カウント手段は、前記第 2 通信手段および前記第 3 通信手段が前記データの書き込

50

み指示を送信することによって、第2タイマーのカウントを開始し、

前記第2タイマーが第2の閾値を超える前に、前記第2の通信手段が前記第1記憶装置による前記完了通知を受け付ける第1条件、または前記第3の通信手段が前記第2記憶装置による前記データの記憶の完了通知を受け付ける第2条件の少なくとも1つを満たすことによって、前記第1通信手段は、前記書き込み指示に対する処理の完了通知を送信することを特徴とする請求項3乃至5のいずれか1項に記載の情報処理装置。

【請求項9】

前記第2タイマーが第2の閾値を超えた際に、前記第2の通信手段が前記第1記憶装置による前記完了通知を受け付けず、且つ、前記第3の通信手段が前記第2記憶装置による前記データの記憶の前記完了通知を受け付けないことによって、前記第1通信手段は、前記書き込み指示に対する処理の失敗通知を送信することを特徴とする請求項8に記載の情報処理装置。

10

【請求項10】

前記第2記憶装置に前記第1記憶装置に記憶させた前記データを記憶させるリカバリ処理が必要であるか否かを判定する判定手段を有し、

前記リカバリ処理が必要である場合、制御手段は、前記第1記憶装置に記憶した前記データを読み出して、前記第2記憶装置に記憶させることを特徴とする請求項1乃至9のいずれか1項に記載の情報処理装置。

【請求項11】

第1のSATA回路と第2のSATA回路を有し、

20

前記第1通信手段、前記第2通信手段、前記第3通信手段は、前記第1のSATA回路に含まれ、

前記第2のSATA回路は、前記第1通信手段に前記データの書き込み指示を送信し、前記第1通信手段は、前記第2のSATA回路に前記書き込み指示に対する通知を送信することを特徴とする請求項1乃至10のいずれか1項に記載の情報処理装置。

【請求項12】

画像データを用紙に印刷する印刷部と、

前記画像データの画像処理を行う画像処理部と、

前記画像処理部を制御するCPUと、

前記第1通信手段、前記第2通信手段、前記第3通信手段を含むSATA回路と、を有し、

30

前記CPUは、前記第1通信手段に前記データの書き込み指示を送信し、

前記第1通信手段は、前記CPUに前記書き込み指示に対する通知を送信することを特徴とする請求項1乃至10のいずれか1項に記載の情報処理装置。

【請求項13】

不揮発性の第1記憶装置と、不揮発性の第2記憶装置と、データの書き込み指示を受け付ける第1通信部と、前記書き込み指示を前記第1記憶装置に送信する第2通信部と、前記書き込み指示を前記第2記憶装置に送信する第3通信部と、前記第2通信部を介する前記第1記憶装置へのデータの送信および前記第3通信部を介する前記第2記憶装置へのデータの送信を制御する制御部と、を有する情報処理装置の制御方法であって、

40

前記制御部が、前記第1記憶装置から前記書き込み指示に対する応答を受け付け、且つ、前記第2記憶装置から前記書き込み指示に対する応答を受け付けないことに基づいて、前記第1記憶装置への前記データの送信を継続し、前記第2記憶装置への前記データの送信を停止する第1工程と、

前記第1通信部が、前記第1記憶装置による前記データの記憶の完了通知を前記第2通信部が受け付けることに基づいて、前記書き込み指示に対する処理の完了通知を送信する第2工程と、を有することを特徴とする情報処理装置の制御方法。

【請求項14】

請求項13に記載の制御方法を、コンピュータに実行させるためのプログラム。

【請求項15】

50

請求項 1 4 に記載のプログラムを格納したコンピュータで読み取り可能な記憶媒体。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、情報処理装置および情報処理装置の制御方法に関する。

【背景技術】

【0002】

印刷装置などの情報処理装置において、複数のストレージデバイス（以後、ストレージ）を搭載してミラーリング機能を実装しているものがある。ここでミラーリング機能とは、複数台の HDD（ハードディスクドライブ）や SSD（ソリッドステートドライブ）などのストレージの全てに同一アドレスから同サイズのデータを書き込む機能である。ミラーリング機能の制御は、ミラーリング制御を行うコントローラ（以下、コントローラ）によって行われる。

10

【0003】

コントローラは、上流のホストコントローラから書き込み／読み出しの要求を受けると、ミラーリング制御している複数のストレージに受け付けた要求に応じた指示を出す。そして、コントローラは、ミラーリング機能で用いる複数のストレージの各々から応答を受け付けると、上流のホストコントローラに応答を返す。

【0004】

コントローラは、例えば 2 つのストレージのうちすべてのストレージから正常に書き込み／読み出しが出来た旨の応答を受けた場合には、ホストコントローラに正常に書き込み／読み出しが出来た旨の応答を返す。また、例えば 2 つのストレージのうち一方のストレージから正常に書き込み／読み出しが出来た旨の応答を受け、他方のストレージからエラーになった旨の応答を受けた場合にも、ホストコントローラに正常に書き込み／読み出しが出来た旨の応答を返す。そして、例えば 2 つのストレージのうち両方のストレージからエラーになった旨の応答を受けた場合には、ホストコントローラにエラーになった旨の応答を返す。

20

【0005】

ミラーリング機能で用いる複数のストレージへの書き込み／読み出し速度は、ストレージごとに異なることがある。例えば、同一の種類ストレージ（HDD）を複数用いた場合には、一部の HDD の書き込み／読み出しの応答は早く、他部の HDD の書き込み／読み出しの応答は一部の HDD に比べて遅い場合（もしくは応答が帰って来ない場合）である。この事例は、例えば HDD の経年劣化や性能の差などによって生じる。また、特許文献 1 に示すように、ミラーリング構成として HDD と SSD を混在で搭載した場合には、SSD からの書き込み／読み出しの応答は早く、HDD の書き込み／読み出しの応答が遅くなる。

30

【先行技術文献】

【特許文献】

【0006】

【特許文献 1】特開 2018 73005 号公報

40

【発明の概要】

【発明が解決しようとする課題】

【0007】

ミラーリングを制御するコントローラは、複数のストレージに書き込みの指示／読み出しの指示を送ってから、所定時間がたってもコマンドを送ったすべてのストレージから応答が帰ってこない場合にホストコントローラにエラーが生じた旨の応答を行う。そのため、各々ストレージの応答速度の差が大きい場合や、応答を返さないストレージが存在した場合には、コントローラからの指示を正常に処理したストレージがあったとしても、コントローラは、ホストコントローラにエラーが生じた旨の応答を行ってしまう。

【0008】

50

そのためコマンドを送信したホストコントローラは、正常なストレージがあったとしても、すべてのストレージがエラー状態であると認識してしまうおそれがある。

【0009】

本発明は、書き込みコマンドに対して所定時間内に応答がないストレージがある場合に、ミラーリングを制御するコントローラがホストコントローラにエラー状態であると誤認識させないことを目的とする。

【課題を解決するための手段】

【0010】

本発明は、不揮発性の第1記憶装置および不揮発性の第2記憶装置を有する情報処理装置であって、データの書き込み指示を受け付ける第1通信手段と、前記書き込み指示を前記第1記憶装置に送信する第2通信手段と、前記書き込み指示を前記第2記憶装置に送信する第3通信手段と、前記第2通信手段を介する前記第1記憶装置へのデータの送信および前記第3通信手段を介する前記第2記憶装置へのデータの送信を制御する制御手段と、を有し、前記第1記憶装置から前記書き込み指示に対する応答を受け付け、且つ、前記第2記憶装置から前記書き込み指示に対する応答を受け付けないことに基づいて、前記制御手段は、前記第1記憶装置への前記データの送信を継続し、前記第2記憶装置への前記データの送信を停止し、前記第1記憶装置による前記データの記憶の完了通知を前記第2通信手段が受け付けることに基づいて、前記第1通信手段は、前記書き込み指示に対する処理の完了通知を送信することを特徴とする。

【発明の効果】

【0011】

本発明は、書き込みコマンドに対して所定時間内に応答がないストレージがあったとしてもホストコントローラにすべてのストレージがエラー状態であると誤認識させないことが可能である。

【図面の簡単な説明】

【0012】

【図1】メインコントローラのシステム構成例を示す図

【図2】SATAブリッジ構成の接続例を示す図

【図3】各SATA制御部の内部構成例を示す図

【図4】従来のリカバリーシーケンスを示す図

【図5】無応答ストレージの発生による課題が生じる構成図

【図6】実施例1における無応答ストレージ発生時の停止処理を説明する構成図

【図7】実施例1の書き込み処理フローを示す図

【図8】実施例2における無応答ストレージ発生時の停止処理を説明する構成図

【図9】実施例3における無応答ストレージ発生時の制御を説明する構成図

【図10】実施例3の書き込み処理フローを示す構成図

【図11】拡張コマンドの一例を示す図

【発明を実施するための形態】

【0013】

(第1実施形態)

添付図面を参照して本実施形態の各実施例を詳しく説明する。以下の実施例は特許請求の範囲に係る発明を限定するものではなく、また各実施例で説明されている特徴の組み合わせすべてが発明の解決手段に必須の構成とは限らない。例えば、各実施例では、ミラーリング構成として2台のストレージデバイスとSATA IF (Serial Advanced Technology Attachment Interface) を例として説明するが、この構成に限らない。ストレージデバイスが複数配された構成であれば何台でもよい。また、ストレージデバイスは、複数のHDDでもよいし、複数のSSDでもよいし、HDDとSSDが混在してもよい。なお、以下の実施例では、情報処理装置の例として画像処理装置を用いる。

【0014】

(実施例1)

図1は、画像処理装置におけるメインコントローラ120のシステム構成例である。メインCPU(中央処理演算器)101は、システム制御や各種演算処理を行う。メモリ制御部102は、各種メモリデバイスへの入出力制御やDMA(ダイレクト・メモリ・アクセス)制御を行う。

【0015】

FLASHメモリ103は、書き換え可能な不揮発性メモリであり、システム全体の制御プログラムや制御パラメータ等が格納される。DRAM(ダイナミック・ランダム・アクセス・メモリ)104は、DDR(Double-Data-Rate)メモリに代表される揮発性の書き換え専用メモリである。プログラムの作業領域や印刷データの格納領域、各種テーブル情報格納領域等の用途に用いられる。ここで、メモリ制御部102と各種メモリデバイスとの関係は、簡略化して表現したものであって、一般的には独立に制御される。

10

【0016】

LAN-IF制御部105は、印刷装置に接続されるローカル・エリア・ネットワーク106との入出力制御を行う。一般的にはTCP/IP(Transmission Control Protocol/Internet Protocol)プロトコルに対応する。ネットワークケーブルを介して外部HOSTコンピュータ107などのネットワーク対応機器と接続され、ネットワーク経由でのプリントを行うことができる。Reader IF制御部108は、スキャナ装置109との通信制御を行う。

20

【0017】

スキャナ装置109によってスキャンした入力画像データを印字させることでコピー機能を実現する。画像処理部110は、LAN-IF制御部105、Reader IF制御部108を介して取り込んだ画像データに対して各種画像処理を行う。

【0018】

SATAホスト制御部111(SATA回路)は、SATA規格に準拠したIFを有するデバイスとのデータ入出力制御を行う。SATAブリッジ制御部112(SATA回路)は、上流側としてSATAホスト制御部111にデバイスとして接続され、下流側としては複数のHost IFを有し、HDD又はSSD(ソリッドステートドライブ)113、114と接続される。SATAブリッジ制御部112では、RAID(Redundant Arrays of Inexpensive Disk)制御やデータ暗号化など付加価値としての機能が搭載されている。

30

【0019】

本実施例では、SATAホスト制御部111とSATAブリッジ制御部112は、それぞれ独立したASIC(特定用途向け集積回路)としてメインコントローラ120に搭載されていることを前提として説明を行う。

【0020】

パネルIF部115は、パネル装置116との通信制御を行う。ここでは図示しないがUI(ユーザー・インターフェイス)として、パネル上の液晶画面表示やボタン等を操作することにより印刷装置の各種設定及び状態の確認ができる。

40

【0021】

ビデオ出力IF部117は、印刷部118とのコマンド/ステータスの通信制御や印刷データの転送を行う。印刷部118は、ここでは図示しないが印刷装置本体と給紙系及び排紙系から構成され、主にビデオ出力IF部117からのコマンド情報に従い、印刷データを用紙に印刷する。メインバス119はバスコントローラ含み、制御バス、データバス及び任意ブロック間のローカルバスを便宜的にまとめて表現したものである。代表例としてPCIe(PCI Express)やASICの内部バスなども含まれる。

【0022】

図2は、SATAブリッジ構成としての接続例を示す図である。メインASIC201は、SATAホスト制御部111を含むメインコントローラ120のシステム全体を制御

50

する中心的なASICである。SATAホスト制御部111は、1個のSATA IP (Intellectual Property) 202をホストIFとして持つ。サブASICは、SATAブリッジ制御部112そのものであり、メインコントローラ120上に独立したIC (Integrated Circuit)として実装されている。

【0023】

SATAブリッジ制御部112は3個のSATA-IP 203~205を持つ。ブリッジ構成での上流側はSATA IP (Host) 202がH Host IF 206を介してSATA IP (Device) 203と接続される。下流側はSATA IP (Host 1) 204がB Host 1 IF 207を介してHDD 113 (SSDでもよい)と、SATA IP (Host 2) 205がB Host 2 IF 208を介してHDD 114 (SSDでもよい)と接続される。

10

【0024】

ここで、SATA IP 202~205は、SATAリンク層及び物理層から構成される。そして、SATA IP 202~205は、各種SATAレジスタの設定に応じてSATA-IF 206~208で接続されるSATAデバイスに対して物理的な(電気信号としての)SATA規格のコマンド発行やステータス受信を行う。つまり、SATA IP 202~205は、通信部として機能する。

【0025】

また、SATAブリッジ制御部112は、電源制御部209と制御信号214で接続されている。電源制御部209はメインボード上に搭載され、メインコントローラ120に含まれる各機能モジュールやそれに接続される各種装置への電力供給の有無を決定し、印刷装置システム全体としての電源制御を担っている。

20

【0026】

電源制御部209から出ている一点斜線210~213は、システム全体の部分であるSATAブリッジ部の各構成要素にたいする電源ラインを示したものである。従って、制御信号214を介した電源制御部209とのやり取りによってSATAブリッジ制御部112は意図したタイミングでHDD 113及び114の電源をOFF・ONすることが可能となる。

【0027】

図3は、SATAホスト制御部111及びSATAブリッジ制御部112の内部構成例を示す図である。H CPU 301は、SATAコマンド発行処理、送受信データの転送処理及びステータス受信処理等SATAコントローラとしての全般的な制御を行う。また、H CPU 301は、メインバス119を介してメインコントローラ120に含まれるCPU 101と通信を行う。

30

【0028】

メモリ制御部302は、Flashメモリ303やSRAM (スタティック・ランダム・アクセス・メモリ) 304との入出力制御を行う。Flashメモリ303には、ブートプログラムやSATAコントローラとしての制御プログラムが格納されている。SRAM 304にはH CPU 301の作業領域、各種制御テーブルやパラメータ格納領域及びデータバッファなどに使用される。ここで、SRAM 304は1ポートRAM、2ポートRAM、FIFO (First IN First OUT) メモリ等の制御を簡略化して記載しているのであって、それぞれ独立に制御され複数個所にSRAMが存在しても構わない。

40

【0029】

割り込み制御部305は、H CPU 301に対する割り込み信号の入力や出力処理、割り込み信号に対するマスク処理などを行う。SATAブリッジ制御部112からの割り込み信号318も接続されている。レジスタH 306は、制御パラメータやステータス情報などを一時的に記憶するためのレジスタである。DMAC (ダイレクト・メモリ・アクセス・コントローラ) 307は、ここでは図示しないがH CPU 301によって所定のレジスタに転送元及び転送先の先頭アドレス及びサイズが設定され、起動が掛けられると所定

50

のメモリ間でデータ転送を行う。

【0030】

Hバス308はバスコントローラ含み、制御バス、データバス及び任意ブロック間のローカルバスを便宜的にまとめて表現したものである。バスブリッジ回路309は、メインバス119とHバス308間のバスプロトコルを相互に変換するバスブリッジである。これによってDMAC307は、メインバス119に接続されているメインコントローラ120に含まれるDRAM104にアクセスすることが可能となる。

【0031】

BCPU310は、SATAコマンド発行処理、送受信データの転送処理及びステータス受信処理等SATAコントローラとしての全般的な制御を行う。メモリ制御部311は、Flashメモリ312やSRAM313との入出力制御を行う。Flashメモリ312には、ブートプログラムやミラーリング処理プログラムなどSATAコントローラとしての制御プログラムが格納されている。SRAM313にはBCPU310の作業領域、各種制御テーブルやパラメータ格納領域及びデータバッファなどに使用される。

【0032】

ここで、SRAM313は1ポートRAM、2ポートRAM、FIFOメモリ等の制御を簡略化して記載しているのであって、それぞれ独立に制御され複数個所にSRAMが存在しても構わない。レジスタB314は、制御パラメータやステータス情報などを一時的に記憶するためや割り込み信号318を発行するために利用される各種レジスタをまとめて記載したものである。

【0033】

電源IF部315は、電源制御部209と制御信号214で接続され、HDD113、114への電源OFF/ON要求信号の制御を行う。割り込み制御部316は、BCPU310に対する各SATA IP204、205などからの割り込み信号の入力や出力処理、割り込み信号に対するマスク処理などを行う。Bバス317はバスコントローラ含み、制御バス、データバス及び任意ブロック間のローカルバスを便宜的にまとめて表現したものである。また、図2で説明したようにSATAホスト制御部111のSATA IP(Host)202とSATAブリッジ制御部112のSATA IP(Device)203とは、H Host IF206とで接続される。さらにSATA IP(Host1/2)204及び205は、B Host1/2 IF207及び208を介して、HDD113及び114とに接続されている。

【0034】

図11は、SATAブリッジ制御部112の設定や情報取得に利用される拡張コマンドの一例を示す図である。拡張コマンドは、SATAブリッジ制御部112で用いることが出来る独自コマンドであり、SATA規格で準備されているユーザー定義のベンダユニーク・コマンド(例えば、F0h)を利用して作成する。

【0035】

例えば、図11の第1行目左から示されるように拡張コマンド名称1001、CMD(サブコマンド)番号1002、転送タイプ1003、指定内容1004として各種拡張コマンドを独自に定義する。ここで、CMD番号はベンダユニーク・コマンド(例えば、F0h)に対するFeatureレジスタに設定されるサブコマンド番号を示す。また、SATA規格では基本的な転送タイプとして、データを伴わないNon Data(ND)転送、単発データ転送を実行するPIO In(PI)又はPIO-Out(PO)転送、連続データ転送を実行するDMA転送などの転送タイプが定義されている。

【0036】

図11での転送タイプ1003は、CMD番号1002に対する転送タイプを定義している。たとえば、SetupBridgeコマンド1005はCMD番号:01h(1006)、転送タイプ:PO(1007)で構成される。同様にGetBridgeInfoコマンド1008はCMD番号:02h(1009)、転送タイプ:PI(1010)として定義されている。SetupBridgeコマンド1005は、SATAブリッジ

10

20

30

40

50

制御部 1 1 2 の各種設定を行うコマンドである。以後に現れるパラメータ、例えばタイムアウト値のようなものはミラーリング運用開始前に本コマンドを用いて予め設定されていることを前提とする。Get Bridge Info コマンド 1 0 0 8 は、SATAブリッジ制御部 1 1 2 全般のステータス（内部情報）取得のための拡張コマンドである。HDD 1 1 3、1 1 4 との接続状態やエラー内容は本コマンドを用いて取得することができる。

【 0 0 3 7 】

図 4 は従来のライト（W）コマンド & ステータス及び異常時のリカバリーシーケンスを示す図である。

【 0 0 3 8 】

図 4（a）は、画像処理装置がミラー状態であることを示す図である。SATAホスト制御部 1 1 1 は、上位からの書き込み要求に対してライト（W）コマンド 4 0 1 を SATAブリッジ制御部 1 1 2 に送信する。そして、Wコマンド 4 0 1 を受信した SATAブリッジ制御部 1 1 2 はコマンド解釈を実行する。そして SATAブリッジ制御部 1 1 2 は、HDD 1 1 3 及び HDD 1 1 4 に対して Wコマンド 4 0 1 と同じ内容の Wコマンド 4 0 2、4 0 3 を送信し、タイムのカウントを開始する。

10

【 0 0 3 9 】

Wコマンドを受け付けた HDD 1 1 3 及び HDD 1 1 4 は、SATAブリッジ制御部 1 1 2 に応答（後述の DMA act パケット）を返す。そして、ここでは詳細な説明は省略するが、SATA規格のフロー制御に基づいて SATAホスト制御部 1 1 1 は、1つのライト（W）データ 4 0 4 を一定のデータパケット単位 4 1 0 で分割して SATAブリッジ制御部 1 1 2 に送信する。

20

【 0 0 4 0 】

ここで、SATA規格でのデータ単位について説明する。SATA規格で送受信される1個のコマンドに対するデータ転送の最大サイズはWデータ 4 0 4 に示すように 3 2 M B で、伝送路で送受信される1回のデータパケット単位 4 1 0 は 8 K B である。

【 0 0 4 1 】

SATAブリッジ制御部 1 1 2 は、SATA-IP 2 0 3、2 0 4、2 0 5 の内部にバッファを持つ。本実施例では、各バッファサイズの例としてデータパケット単位 8 K B × 2 個分の 1 6 K B を保持可能な容量としている。

【 0 0 4 2 】

一般的に1個のコマンド処理でのデータサイズを大きくすることで処理効率は向上する。本例では各バッファ 4 1 1、4 1 2、4 1 3 を 1 6 K B としているが、本サイズは引き出したパフォーマンスによって決められるべき値であり、この値に限定するものではない。

30

【 0 0 4 3 】

図 4（a）の説明に戻る。SATAホスト制御部 1 1 1 から SATAブリッジ制御部 1 1 2 に送信された Wデータ（Wデータ 4 0 4 の一部のデータ）が、バッファ 4 1 1 に保持される。バッファ 4 1 1 に保持された Wデータを、SATAブリッジ制御部 1 1 2 が、バッファ 4 1 2、4 1 3 に保持させる。そして、SATAブリッジ制御部 1 1 2 は、バッファ 4 1 2、4 1 3 に保持されたデータを Wデータ 4 0 5 および Wデータ 4 0 6 として HDD 1 1 3 及び 1 1 4 に逐次送信する。一方、バッファ 4 1 1 は、前の Wデータを消去して次の Wデータを受信して保持する。なお、以下では、前の Wデータの消去については説明を省く。これは、バッファ 4 1 2 と 4 1 3 においても同様である。

40

【 0 0 4 4 】

少なくとも一方の HDD において Wデータの書き込みが終わると、SATAブリッジ制御部 1 1 2 は、バッファ 4 1 1 から書き込みが完了した HDD に対応するバッファに次の Wデータ（Wデータ 4 0 4 の一部）を保持させる。例えば、HDD 1 1 3 への書き込みが先に終わり、HDD 1 1 4 への書き込みが後に終わる場合には、バッファ 4 1 2 に次の Wデータが保持された後にバッファ 4 1 3 に次の Wデータが保持される。同時に書き込みが終わった場合には、バッファ 4 1 2 とバッファ 4 1 3 に同時に次の Wデータが保持される

50

。なお、バッファは、前述したように、前のWデータを消去して次のWデータを保持する。

【0045】

そして、SATAホスト制御部からSATAブリッジ制御部112に送信されたWデータ411の全てが、SATAブリッジ制御部112を介してHDD113およびHDD114に記憶されるまで上述の送信が繰り返される。

【0046】

ここで、予め設定されたタイムアウト時間内にWデータ404の全ての送信がHDD113、HDD114の両方に正常に転送完了すると、SATAブリッジ制御部112は両デバイスのそれぞれからOK（正常）ステータス407、408を受信する。

10

【0047】

このような場合にSATAブリッジ制御部112は、SATAホスト制御部111に対して書き込み処理正常終了としてOKステータス409を送信する。SATAブリッジ制御部112からOKステータスを受信したSATAホスト制御部111はWコマンド401の一連の処理を終了する。

【0048】

このように、HDD113及び114は双方のデータ保持状態が同一レベルであり整合性が取れている状態を、ミラーリングの正常状態としてミラーリング状態と定義する。

【0049】

なお、本実施例ではHDD113をマスターストレージ、HDD114をスレーブストレージとして説明する。ここでは図示しないが、リード（R）コマンド発行時はマスターからのみ読み出し処理が実行される。

20

【0050】

図4（b）においてSATAブリッジ制御部112が、書き込みに失敗した旨のステータス（NGステータス）を一方のHDDから受け付けた状況の一例を説明している。

【0051】

具体的には、HDD114のスレーブ側が書き込み処理途中に何らかの原因で失敗した場合を示す。

【0052】

このとき、SATAブリッジ制御部112がNG（失敗）ステータス408を受信し、BCPU310はスレーブ側のHDD114に対して停止処理を実行する。停止処理を実行すると、SATAブリッジ制御部112は、SATAホスト制御部111から受け付けたWデータ404の一部をバッファ413に送信しなくなり、バッファ412にのみ送信する。

30

【0053】

つまり、BCPU310は、正常なHDD113が接続されたマスター側に対して書き込み処理を継続し、書き込みに失敗したHDD114が接続されたスレーブ側のSATAIP205を初期化して現行の書き込み処理を実行しなくなる。このSATAIP205の初期化処理によって、SATAブリッジ制御部112は、SATAブリッジ制御部112に接続されたストレージがHDD113のみであると認識する。なお、初期化処理は実施せず、ただ無視することができるのであればそれでもよい。

40

【0054】

タイムアウト時間内にマスター側の書き込み処理が完了すると、SATAブリッジ制御部112はマスター側のHDD113からOKステータス407を受信する。これによりSATAブリッジ制御部112は、本書き込み処理は正常に完了したとして上位SATAホスト制御部111にOKステータス409を送信する。OKステータス409を受信したSATAホスト制御部111は本書き込み処理を終了し、次のコマンドが送信可能となる。

【0055】

一方、SATAブリッジ制御部112のBCPU310は、未使用処理時に前回書き込

50

み処理時の開始アドレス及びデータサイズをS R A M 3 1 3に記録した情報に基づき、失敗した側のリカバリ処理をバックグラウンドで開始する。

【 0 0 5 6 】

図 4 (c) を用いて図 4 (b) の後のリカバリ処理について説明する。S A T A ブリッジ制御部 1 1 2 は、書き込み処理に失敗したH D D 1 1 4の初期化処理を行い、前回書き込み処理に成功したH D D 1 1 3から記録した開始アドレス及びデータサイズのデータの読み出し処理 (R) 4 1 4を開始する。

【 0 0 5 7 】

読み出したデータを、書き込み処理に失敗した側H D D 1 1 4に対して同一アドレスから逐次書き込み処理 (W) 4 1 5を実行する。

10

【 0 0 5 8 】

リカバリ処理中に書き込みエラーが生じた場合、画像形成装置は、ここでは図示しないが予め設定された回数のリカバリ処理を繰り返す。予め設定された回数のリカバリ処理の中でリカバリ処理に成功すれば、画像形成装置は、ミラー状態に遷移する。

【 0 0 5 9 】

予め設定された回数のリカバリ処理の全てで失敗した場合には、H D D 1 1 4はシステム的使用不能になりH D D 交換通知がなされ、片側H D D 1 1 3のみ (デグレード状態) で装置の運用が継続される。

【 0 0 6 0 】

この時、H D D 交換通知は割り込み信号 3 1 8によりS A T A ホスト制御部 1 1 1に通知される。ここでは省略するが、S A T A ホスト制御部 1 1 1は、S A T A ブリッジ制御部 1 1 2の内部状態を取得するG e t B r i d g e I n f oコマンド 1 0 0 8を有し、それによってどちらのH D Dを交換すべきかを判断することができる。

20

【 0 0 6 1 】

なお、S R A M 3 1 3に記録した情報は、B C P U 3 1 0によって利用後消去される。また、ミラーリング状態 (例えば、ミラー状態、デグレード状態) は、S A T A ブリッジ制御部 1 1 2内のS R A M 3 1 3やレジスタB 3 1 4などに記録されている。

【 0 0 6 2 】

図 5 は、課題が生じる構成としてWコマンド 5 0 1に対して無応答のH D Dが生じる構成について説明する。

30

【 0 0 6 3 】

図 5 は、S A T A ホスト制御部 1 1 1がS A T A ブリッジ制御部 1 1 2にWコマンド 5 0 1とWデータ 5 0 3の一部のデータ (データパケット単位 4 1 0で送付されたデータ) を送信する。S A T A ブリッジ制御部 1 1 2は、一部のデータをバッファ 4 1 1に保持し、S A T A ブリッジ制御部 1 1 2は、保持した一部のデータをバッファ 4 1 2および 4 1 3に保持する。

【 0 0 6 4 】

そしてS A T A ブリッジ制御部 1 1 2は、H D D 1 1 3およびH D D 1 1 4にWコマンドを送信し、WコマンドをH D D 1 1 3およびH D D 1 1 4に送信することを起点にタイマー (経過時間の閾値) を設定し、時間のカウントを開始する。

40

【 0 0 6 5 】

図 5 においてH D D 1 1 3は、Wコマンドに対応してデータ転送開始割り込みを示すD M A a c t パケット 5 0 2をS A T A ブリッジ制御部 1 1 2に返す。D M A a c t パケット 5 0 2を受け付けたS A T A ブリッジ制御部 1 1 2は、バッファ 4 1 1に保持した一部のデータをH D D 1 1 3に送信する。H D D 1 1 3が受信した一部のデータを書き込むと、次の一部のデータをバッファ 4 1 1から受信するためにD M A a c t パケット 5 0 2をS A T A ブリッジ制御部 1 1 2に出力する。

【 0 0 6 6 】

なお、D M A a c t パケット 5 0 2は、S A T A ブリッジ制御部 1 1 2に対してバッファ 4 1 2に保持された分割されたデータを送信させるためにH D D 1 1 3が送付するコマ

50

ンドである。

【0067】

一方、図5においてHDD114は、設定されたタイマーの時間内にWコマンドに 응답しない。例えば、HDDの寿命が近い場合や寿命が尽きている場合などにこのような無 응답が生じる。このような状態では、HDD114は、DMAactバケット502をSATAブリッジ制御部112に返さない。

【0068】

そのため、SATAブリッジ制御部112のバッファ412が2回目にフル状態になり、HDD113に2回目の分割データの書き込みを終えた時点で、無 응답側のバッファ413には1回目の分割データが保持されているという状況が発生する。このような状況が発生すると、バッファ413に2回目の分割データを保持させていないため、バッファ412に3回目の分割データを保持させることが出来ない。言い換えると、HDD114に1回目のデータを記憶させていないことで、HDD113に2回目の分割データまで記憶させることはできるが、3回目以降の分割データを記憶させることはできない。

10

【0069】

このような状況が発生するとバッファ412側を制御しているSATA-IP204は、HDD113に対してHoldバケット506を発行する。それを受けたHDD113側はHoldA507を返し、待機状態になる。

【0070】

一方、HDD114側は無 응답のままであるため、タイマーの時間内にSATAブリッジ制御部112は両方のHDDからのOKステータスを受信することなくタイムアウトとなる(設定した経過時間の閾値を超える)。

20

【0071】

SATAブリッジ制御部112は、双方のHDD113、114にWコマンドを発行したにも係わらずタイムアウトし、書き込み処理が完了しなかったため、SATAホスト制御部111に対してNG(失敗)ステータス409を送信する。

【0072】

NGステータスを受信したSATAホスト制御部111は、再書き込み処理(同一Wコマンド501を再度発行する)をSATAブリッジ制御部112に指示する。

【0073】

前述したようにHDDが無 응답になる現象は、ストレージデバイスの経年変化が進むことによって生じる。そして、無 응답になったHDDは、電源OFF ONすることにより一時的に回復する可能性があるが、すぐに無 응답になる現象が生じてしまう。

30

【0074】

つまり、SATAホスト制御部111が指示した再書き込み処理も失敗する可能性が高い。その結果、使用可能なHDDがあっても、SATAホスト制御部111は、すべてのHDDが使用不能であると把握してしまい、システム的な破綻に至る可能性がある。さらに、SATAホスト制御部111は、SATAブリッジ制御部112からNGのステータスを取得してもどのHDDが無 응답状態のであるかを判定することが出来ない。

【0075】

図6は、本実施例における無 응답ストレージ発生時の停止処理を説明する構成である。本実施例では、一方のHDDが無 응답状態に陥った際に一方のHDDへのデータの送信を停止し、他方へのHDDにデータの送信を継続する。そして、他方のHDDが書き込み処理に成功すれば、上位回路(SATAホスト制御部111)に書き込み処理に成功した旨の通知を行う。以下で詳細を説明する。なお、図5の説明と同様の部分については説明を省略する。

40

【0076】

図6の場合もSATAブリッジ制御部112は、SATAホスト制御部111からWコマンドを受け付け、1回目の分割データを受け付けてバッファ411に保持してからバッファ412とバッファ413に保持する。SATAブリッジ制御部112は、HDD11

50

3 および HDD 1 1 4 に W コマンド 4 0 2 及び W コマンド 4 0 3 を送信し、W コマンドの送信を起点として第 1 タイマー（時間経過の第 1 閾値）を設定し、第 1 タイマーに対する時間のカウントを開始する。

【 0 0 7 7 】

S A T A ブリッジ制御部 1 1 2 は、H D D 1 1 3 および H D D 1 1 4 の少なくとも一方から W コマンドへの応答（例えば、D M A a c t パケット 5 0 2（データ転送開始割り込み））を受け付ける。応答を受信した B C P U 3 1 0 は H D D 1 1 3 側がデータ転送開始したことを S R A M 3 1 3 又はレジスタ B 3 1 4 などに記憶する。そして、S A T A ブリッジ制御部 1 1 2 は、応答（データ転送開始割り込み）を起点として第 2 タイマー（経過時間の第 2 閾値）を設定し、第 2 タイマーに対する時間のカウントを開始する。この第 2 閾値は、第 1 閾値よりも短い値となる。

10

【 0 0 7 8 】

S A T A ブリッジ制御部 1 1 2 は、応答（データ転送開始割り込み）を受け付けない状態で第 2 タイマーがタイムアップした場合（経過時間が第 2 閾値を超えた場合）には、H D D 1 1 4 側の停止処理を行う。停止処理およびその後のリカバリ処理は、図 4（b）、図 4（c）と同様であるため説明を省略する。

【 0 0 7 9 】

S A T A ブリッジ制御部 1 1 2 は、第 2 タイマーの時間内に H D D 1 1 4 から応答（データ転送開始割り込み）を受け付けた場合には、H D D 1 1 4 側への書込み処理を継続する。

20

【 0 0 8 0 】

なお、S A T A ブリッジ制御部 1 1 2 から各デバイスに対してコマンドが発行できるのは、各 H D D がレディ状態（ビジーではない）であることを意味する。一旦コマンドを発行すると後は H D D 側の応答を待つしかすべがない。

【 0 0 8 1 】

第 2 タイマーを設定することによって、W コマンド送付後の無応答などの故障発生時に正常なデバイスが異常なデバイスに引きずられてコマンド処理失敗となることを回避することができる。

【 0 0 8 2 】

いずれかの H D D へのデータ転送が開始されるとそれを通知する割り込み信号を設ける。その割り込み信号を割り込み制御部 3 1 6 から受信した B C P U 3 1 0 は、受信した S A T A - I P 側を記憶しておき、第 2 タイマーを設定する。

30

【 0 0 8 3 】

なお、ここではデータ転送開始を S A T A - I P からの割り込み信号で検知する例を示しているが、この方法に限定するものではない。他の例としてデータ転送回数のカウンタ値の変化から検知するなどであっても構わない。

【 0 0 8 4 】

次に図 7 を用いて図 6 で説明した本実施例の構成を実現するための書込み処理フローの一例を説明する。本実施例のフローチャートは、B C P U 3 1 0 によって実行され、H D D へのアクセスが生じると開始する。

40

【 0 0 8 5 】

ステップ S 6 0 1 において、S A T A ブリッジ制御部 1 1 2 を制御している B C P U 3 1 0 は、ステップ S 6 0 1 において S A T A ホスト制御部 1 1 1 から W コマンドを受信する。

【 0 0 8 6 】

ステップ S 6 0 2 において B C P U 3 1 0 は、受信したコマンドの内容を解釈し、同一内容の W コマンドを接続されている H D D 1 1 3 及び 1 1 4 に送信する。そして、ステップ S 6 0 2 を終わるとステップ S 6 0 3 において、第 1 タイマー（経過時間の第 1 閾値）をタイマー回路に設定し、本コマンド処理全体に対する制限時間のカウントを開始する。なお、第 1 タイマーの設定およびカウント開始は、W コマンドを H D D 1 1 3 および H D

50

D 1 1 4 に送信することが起点であればよいためステップ S 6 0 2 と S 6 0 3 は同時でも逆でもよい。なお、タイマー回路は、S A T Aブリッジ制御部 1 1 2 に含まれる。

【 0 0 8 7 】

ステップ S 6 0 4 において B C P U 3 1 0 は、H D D の一方から応答（データ転送開始割り込み）を待つ。一方の H D D から応答を受け付けると、H C P U 3 0 1 は、応答を返してきた H D D へのデータ転送を開始する。一方の H D D へのデータ転送を開始したらステップ S 6 0 5 に遷移する。ステップ S 6 0 5 において第 2 タイマー（経過時間の第 2 閾値）をタイマー回路に設定し、まだデータ転送開始していないデバイスに対する制限時間のカウンタを開始する。なお、第 2 タイマーの設定およびカウンタ開始は、応答を H D D 1 1 3 または H D D 1 1 4 から受信することが起点である。

10

【 0 0 8 8 】

ここで、第 1 タイマー及び第 2 タイマーを設定するタイマー回路はひとつ以上あればよい。ひとつの場合には第 1 タイマーの残り時間を記録しておき第 2 タイマーを設定し、第 2 タイマーの経過後、再び第 1 タイマーの残りの時間を設定すればよい。タイマー回路の使い方は特に限定しない。

【 0 0 8 9 】

次に、ステップ S 6 0 6 において B C P U 3 1 0 は、他方の H D D 側がデータ転送を開始したか否かの判定を実施する。

【 0 0 9 0 】

ステップ S 6 0 6 において B C P U 3 1 0 は、H D D の他方から応答（データ転送開始割り込み）を受け付けると、B C P U 3 1 0 は、応答を返してきた H D D へのデータ転送を開始する。他方の H D D へのデータ転送を開始したらステップ S 6 0 8 に遷移する。

20

【 0 0 9 1 】

ステップ S 6 0 8 で第 2 タイマーを解除し、ステップ S 6 0 9 において両 H D D に対する図 4 で説明した通常処理を実行し、ステップ S 6 1 2 に遷移する。なお、第 2 タイマーの解除とは、例えば、第 2 タイマーの設定を解除し、且つ、第 2 タイマーに対する経過時間のカウンタを停止することである。

【 0 0 9 2 】

ステップ S 6 0 6 の説明に戻る。ステップ S 6 0 6 において、他方の H D D から応答を受け付けていない場合には、データ転送を開始していないためステップ S 6 0 7 に遷移する。ステップ S 6 0 7 において第 2 タイマーがタイムアウトしているか否かの判定を行う。言い換えると、経過時間が第 2 閾値を超えたか否かの判定を行う。

30

【 0 0 9 3 】

ステップ S 6 0 7 において第 2 タイマーがタイムアウトしていない場合には、ステップ S 6 0 6 に戻る。

【 0 0 9 4 】

ステップ S 6 0 7 において第 2 タイマーがタイムアウト時間を過ぎているか場合には、ステップ S 6 1 0 に遷移する。ステップ S 6 1 0 において B C P U 3 1 0 は、W コマンドに対して無応答である（データの転送を開始できない）H D D 1 1 4 に対して停止処理を実行する。具体的には、H D D 1 1 4 側のバッファ 4 1 3 へのデータ転送を停止する停止処理を実行する。ステップ S 6 1 0 を終わるとステップ S 6 1 1 に遷移する。ステップ S 6 1 1 において、ステップ S 6 0 4 でデータ転送を開始した H D D 1 1 3 のデータ転送を継続して行い、ステップ S 6 1 2 に遷移する。

40

【 0 0 9 5 】

ステップ S 6 1 2 において、B C P U 3 1 0 は、H D D 1 1 3 および H D D 1 1 4 の両方の H D D の書き込み処理のステータスが所定のステータスであるか否かの判定を行う。所定のステータスは、2 通りある。ステップ S 6 1 1 からステップ S 6 1 2 に遷移した場合、所定のステータスは、H D D 1 1 3 の書き込み処理のステータスが正常に完了、または、エラーを示すステータスであることである。

【 0 0 9 6 】

50

ステップ S 6 0 9 からステップ S 6 1 2 に遷移した場合、少なくとも 1 つの HDD の書き込み処理が正常に完了したことを示すステータスであること、または、すべての HDD の書き込み処理がエラーになったことを示すステータスであることである。なお、少なくとも 1 つの HDD の書き込み処理が正常に完了したことを示すステータスであるとは、例えば、HDD 1 1 3 は正常なステータスを示し、HDD 1 1 4 は、エラーもしくは書き込み処理実行中である場合である。

【 0 0 9 7 】

ステップ S 6 1 2 において、所定のステータスではない場合には、ステップ S 6 1 3 に遷移する。ステップ S 6 1 3 において、第 1 タイマーがタイムアウトしていない場合には、ステップ S 6 0 9 またはステップ S 6 1 1 に戻る。一方、ステップ S 6 1 3 において、第 1 タイマーがタイムアウトした場合（経過時間が第 1 閾値を超えた場合）には、ステップ S 6 1 6 に遷移する。具体的には、ステップ S 6 1 1 を介した場合、HDD 1 1 3 の書き込み処理が第 1 タイマーの時間内で終了しない場合である。ステップ S 6 0 9 を介した場合、HDD 1 1 3 及び HDD 1 1 4 の一方がエラーで他方の書き込み処理が第 1 タイムアウト時間内に終了していない場合、または、両方の HDD の書き込み処理が第 1 タイムアウト時間内に終了していない場合である。

10

【 0 0 9 8 】

これらの場合、ステップ S 6 1 6 において B C P U 3 1 0 は、書き込み処理がエラー状態であると判断する。そして、SATAブリッジ制御部 1 1 2 は、SATAホスト制御部 1 1 1 にエラーである旨のステータス（失敗通知）を送信する。ステップ S 6 1 6 を終わるとフローが終了する。

20

【 0 0 9 9 】

ステップ S 6 1 2 の説明に戻る。ステップ S 6 1 2 において、所定のステータスである場合には、ステップ S 6 1 4 に移行する。ステップ S 6 1 4 において B C P U 3 1 0 は、第 1 タイマーを解除する。なお、第 1 タイマーを解除するとは、例えば、第 1 タイマーの設定を解除し、第 1 タイマーに対する経過時間のカウンタを停止することである。ステップ S 6 1 4 を終わるとステップ S 6 1 5 に移行する。

【 0 1 0 0 】

次のステップ S 6 1 5 において、HDD 1 1 3 および HDD 1 1 4 の少なくとも一方の書き込み処理の結果が成功しているか否かを判定する。ステップ S 6 1 5 が No の場合、すなわち両 HDD とともに失敗の場合にはステップ S 6 1 6 に進む。ステップ S 6 1 6 は既に説明済みなので省略する。

30

【 0 1 0 1 】

ステップ S 6 1 5 が Yes の場合には、すなわち HDD 1 1 3 および HDD 1 1 4 の少なくとも一方の書き込み処理の結果が成功している場合には、ステップ S 6 1 7 に遷移する。ステップ S 6 1 7 において要求元である SATAホスト制御部 1 1 1 に対して W コマンドに基づく書き込み処理が正常に完了した旨のステータス（完了通知）を送信する。これにより、SATAブリッジ制御部 1 1 2 と SATAホスト制御部 1 1 1 間の処理が終了する。なお、S 6 1 7 の後に S 6 1 4 を行ってもよい。

【 0 1 0 2 】

ステップ S 6 1 8 において B C P U 3 1 0 は、リカバリ処理の有無を判定する。ステップ S 6 1 8 においてリカバリ処理がない場合、すなわち両 HDD とともに成功の場合には、フローを終了し、次のコマンドを受付可能な状態に戻る。

40

【 0 1 0 3 】

ステップ S 6 1 8 において B C P U 3 1 0 が、リカバリ処理がある状態と判断するとステップ S 6 1 9 に遷移する。なおリカバリ処理が有る状態とは、例えば、停止処理をした HDD が有る場合、または、エラーステータスの HDD が有る場合である。

【 0 1 0 4 】

ステップ S 6 1 9 において B C P U 3 1 0 は、NG側のリカバリ（修復）処理を開始する。リカバリ処理は、既に図 4（c）で説明した内容であるので省略する。リカバリ処理

50

は必要に応じて予め設定されたリトライ回数分繰り返される。

【0105】

ステップS620においてBCPU310は、リカバリ処理のn-1回目までが成功したか否かの判定を行う。ステップS620においてn-1回目までのどこかで成功(Yes)すれば、その後次のコマンド受付状態に戻る。

【0106】

ステップS620において、リカバリ処理をn-1回目まで行っても成功しなかった場合には、ステップS622に進む。ステップS622においてBCPU310は、リカバリ処理に電源OFF/ONを実行するかどうかの判定を行う。ここで、電源OFF/ON条件は予め設定されているものとする。

10

【0107】

ステップS622がYesの場合には、BCPU310は電源IF部315を介して電源制御部209に所望のデバイス側の電源をOFF ON制御を行う。例えば、図5(b)で説明した無応答でのリカバリ処理時に電源OFF/ON実行が有効であれば、失敗しているHDD114の電源をOFF ONして修復作業を試みる。そして、ステップS624の判定において修復成功(Yes)であれば次のコマンド受付状態に戻る。ステップS622がNo又はステップS624が失敗(No)した場合にはステップS625に進む。なお、ステップS622およびステップS623は、なくてもよい。

【0108】

ステップS625においてBCPU310は、リカバリ側のデバイスをシステム的に未使用にするためにデグレード状態として運用を継続する。さらにNG側デバイスを特定して交換作業の警告を、例えばパネルIF部115を介してパネル装置116に表示する。

20

【0109】

このような構成によれば、書き込みコマンドに対して所定時間内に応答がないストレージがあったとしてもホストコントローラにすべてのストレージがエラー状態であると誤認識させないことが可能である。

【0110】

(実施例2)

図8では、実施例2における無応答ストレージ発生時の停止処理を説明する構成図を説明する。

30

【0111】

図8は、SATAブリッジ制御部112を搭載せず、SATAホスト制御部111に2個のHDD113、114を接続してミラーリング処理を実施する場合の例である。図8においてSATAホスト制御部111は、SATA-IP704とSATA-IP705を有し、これらは、図3のSATA-IP204とSATA-IP205と同等の機能を有する。そのため、バッファ702及び703は、Host1及び2に対するデータ送受信バッファであり、図4で説明したバッファ412、413に該当するものと同等の機能となる。

【0112】

なお、図8の書き込み処理については図5の処理と同じであるため同等の機能を有する構成には図4、5と同じ番号を用いて説明する。

40

【0113】

バッファ702及び703は、Host1及び2に対するデータ送受信バッファであり、図4で説明したバッファ412、413に該当するものと同等の機能となる。バッファ702はSATA-IP(Host1)204に含まれ、バッファ703は、SATA-IP(Host2)205に含まれる。

【0114】

SATAホスト制御部111は、CPU101からWコマンドを受け付ける。SATAホスト制御部111は、Wコマンドを受け付けると、DMAC701によってメインバス上に接続されるDRAM104から書き込みを行うデータを取得する。DMAC701は

50

、取得したデータを各バッファ702、703に実施例1で説明した構成と同様に分割して保持させる。

【0115】

また、DMAC701はひとつのチャンネルしか持たないのでWコマンド501に対するWデータ503を各バッファに保持させる際には、各バッファに分割データを送信する必要がある。従って、HDD114がWコマンドに対して無応答状態に陥るとバッファ703が2回目の分割データから送信できなくなる。

【0116】

そこで、本実施例においても、一方のHDDが無応答状態に陥った際に一方のHDDへのデータの送信を停止し、他方へのHDDにデータの送信を継続する。そして他方のHDDが書き込み処理に成功すれば、上位回路(CPU101)に書き込み処理に成功した旨の通知を行う。図8を用いて詳細を説明する。

10

【0117】

まず上位のCPU101からSATAホスト制御部111がWコマンドを受け付ける。SATAホスト制御部111がHDD113およびHDD114にWコマンド402及びWコマンド403を送信し、Wコマンドの送信を起点として第1タイマーを設定し、第1タイマーに対するタイムカウントを開始する。

【0118】

SATAブリッジ制御部112は、HDD113およびHDD114の一方からWコマンドへの応答(例えば、DMAactパケット502)を受け付ける。応答を受信したHCPU301はHDD113側がデータ転送開始したことをSRAM304又はレジスタH306などに記憶する。そして、SATAホスト制御部111は、応答(データ転送開始割り込み)を起点として第2タイマーを設定し、第2タイマーのタイムカウントを開始する。この第2タイマー(経過時間の第2閾値)は、第1タイマー(経過時間の第1閾値)よりも短い値となる。

20

【0119】

SATAホスト制御部111は、第2タイマーがタイムアウトしてもHDD114から応答(データ転送開始割り込み)を受け付けなかった場合には、HDD114側に対して停止処理を行う。そして、第1タイマーがタイムアウトする前に、HDD113から書き込み処理に成功した旨の通知を受け付けると、CPU101に書き込み処理に成功した旨の通知を送信する。なお、停止処理およびその後のリカバリ処理は説明を省略する。

30

【0120】

SATAホスト制御部111は、第2タイマーがタイムアウトする前にHDD114から応答(データ転送開始割り込み)を受け付けた場合には、HDD114側への書き込み処理を継続する。そして、第1タイマーがタイムアウトする前に、HDD113とHDD114の少なくとも1つから書き込み処理に成功した旨の通知を受け付けると、CPU101に書き込み処理に成功した旨の通知を送信する。

【0121】

なお、SATAホスト制御部111から各デバイスに対してコマンドが発行できるのは、各HDDがレディ状態(ビジーではない)であることを意味する。一旦コマンドを発行すると後はHDD側の応答を待つしかすべがない。

40

【0122】

なお、実施例1と同様に、ここではデータ転送開始をSATA-IPからの割り込み信号で検知する例を示しているが、この方法に限定するものではない。他の例としてデータ転送回数のカウンタ値の変化から検知するなどであっても構わない。

【0123】

このような構成においても、書き込みコマンドに対して所定時間内に応答がないストレージがあったとしてもホストコントローラ(ここではCPU101)にすべてのストレージがエラー状態であると誤認識させないことが可能である。

【0124】

50

(第2実施形態)

図面9および図10を参照して本実施形態の実施例を詳しく説明する。第1実施形態と同様の構成については同じ番号を付して説明を省略する。

【0125】

(実施例3)

図9では、無応答ストレージの発生によって生じる課題に対応するための本実施例の制御構成を説明する。

【0126】

図9のDMAC801は、独立に処理可能なチャンネルを2個搭載し、DMAC801のCH0はバッファ702と、CH1はバッファ703と接続されている。

10

【0127】

本構成においても、図に示すようにHDD114がWコマンド805を送信した後に無応答809に陥っても、すべてのHDDがエラー状態であると誤認識させないことが可能である。

【0128】

SATAホスト制御部111は、CPU101からWコマンドを受け付けると、HDD113およびHDD114にWコマンドを出力する。そして、SATAIP704がHDD113から応答(DMAact806)を受け付けると、HCPU301は、応答を起点として第3タイマーを設定し、制限時間のカウントを開始する。

【0129】

20

図9の構成においては、HDD113の書き込み処理とHDD114の書き込み処理は、独立して実行可能である。そのため、SATAホスト制御部111は、HDD114の書き込み処理の状況に係わらず、HDD113側からのDMAact806を毎回の起点としたデータパケット単位での転送を継続して実行可能である。

【0130】

そのため、HDD114側が無応答状態となってもHDD113側は書き込み処理に成功することが出来る。具体的には、SATAホスト制御部111は、HDD113側において、第3タイマーがタイムアウトする前に、書き込み処理に成功した旨のステータス808を受信することが出来る。一方、無応答状態のHDD114側は、第2タイマーがタイムアウトしても書き込み処理を完了させることができない。その停止処理が実行され、HCPU301が、HDD114側がエラーステータスであると判断する。

30

【0131】

このような状況において、HCPU301は、CPU101に対して書き込み処理成功として書き込み完了割り込み信号を発行する。一方、停止処理されたHDD114の以後のバックグラウンドでのリカバリ処理は図4(c)で説明済みなので説明を省略する。

【0132】

図10を用いて図9で説明した本実施例の構成を実現するための書き込み処理フローの一例を説明する。図10は、図9で説明した独立した2チャンネルを有するDMACを搭載したミラーリング構成での書き込み処理に対する処理フローを示す図である。図7と同様の構成については同様の番号を付して説明を省略する。図10のフローチャートは、HCPU301によって実行される。

40

【0133】

HCPU301は、ステップS901において上位アプリケーション(CPU101)からWコマンドを受信する。ステップS902においてHCPU301は、受信したコマンドの内容を解釈し、同一内容のWコマンドを接続されているHDD113及び114に送信する。

【0134】

ステップS903においてHCPU301は、HDD113またはHDD114の少なくとも一方の応答(データ転送開始割り込み)を待つ。少なくとも一方のHDDから応答を受け付けると、HCPU301は、応答を返してきたHDDへのデータ転送を開始する

50

。少なくとも一方のHDDへのデータ転送を開始したらステップS904に遷移する。

【0135】

ステップS904において第3タイマー（時間経過の第3閾値）をタイマー回路に設定し、本コマンド処理全体に対する制限時間のカウントを開始する。

【0136】

ステップS905において第3タイマーを設定後もデータ転送およびHDDの書き込み処理は継続される。

【0137】

ステップS906において、HCPU301は、HDD113およびHDD114の両方のHDDの書き込み処理のステータスが所定のステータスであるか否かの判定を行う。

10

【0138】

所定のステータスとは、少なくとも1つのHDDの書き込み処理が正常に完了したことを示すステータスであること、すべてのHDDの書き込み処理がエラーになったことを示すステータスであること、のいずれかのステータスとなる。なお、少なくとも1つのHDDの書き込み処理が正常に完了したことを示すステータスであるとは、例えば、HDD113は正常なステータスを示し、HDD114は、エラーもしくは書き込み処理実行中である場合である。

【0139】

ステップS906において、所定のステータスではない場合には、ステップS907に遷移する。

20

【0140】

ステップS907においてHCPU301は、第3タイマーがタイムアウトしたかどうかの判定を行う。第3タイマーがタイムアウトしていない場合には、ステップS905に戻り、処理を継続する。ステップS907において、第3タイマーがタイムアウトした場合には、ステップS910に遷移する。ステップS910において、HCPU301は、CPU101にエラーステータスを送信する。

【0141】

ステップS906の説明に戻る。ステップS906において、所定のステータスではある場合には、ステップS909に遷移する。

【0142】

ステップS909においてHCPU301は、HDD113およびHDD114の書き込み処理の結果（ステータス）を確認する。HDD113およびHDD114の両方のステータスが、書き込み処理の失敗を示している場合には、ステップS910に遷移する。

30

【0143】

ステップS909において、すべてのHDDのステータスが、書き込み処理の成功を示す場合、または少なくとも1つのHDDのステータスが書き込み処理の成功を示す場合には、ステップS911に遷移する。

【0144】

ステップS911において、HCPU301は、要求元であるCPU101に対してWコマンド処理正常としてステータスを送信し、CPU101とSATAホスト制御部111の間の処理は終了する。

40

【0145】

ステップS912においてHCPU301は、第3タイマーがタイムアウトしたかどうかの判定を行う。ステップS907において、第3タイマーがタイムアウトした場合には、ステップS910に遷移する。ここでは詳細な説明は省略するが、この場合ステップS910においてHCPU301は、CPU101とのやり取りは完了しているため、その他所定のエラー処理を実行する。第3タイマーがタイムアウトしていない場合には、ステップS913に進む。ステップS913においてHCPU301は、もう一方のHDDが処理継続中かどうかの判定を行う。ステップS913の判定結果がYesの場合には、ステップS902に戻り、S912及びS913の間を繰り返す。ステップS913の判定

50

結果がN oの場合には、ステップS 9 1 4に進む。ステップS 9 1 4においてH C P U 3 0 1は、第3タイマーの設定を解除し、ステップS 6 1 8に遷移する。

【0146】

ステップS 6 1 8 ~ S 6 2 5については、説明を省略する。なお、本フローチャートでは、図7のS 6 2 1 ~ S 6 2 3を省略している。

【0147】

このような構成においても、書き込みコマンドに対して所定時間内に応答がないストレージがあったとしてもホストコントローラにすべてのストレージがエラー状態であると誤認識させないことが可能である。

【0148】

(その他の実施形態)

以上、本発明の様々な例と実施形態を示して説明したが、本発明の趣旨と範囲は、本明細書内の特定の説明に限定されるものではない。

【0149】

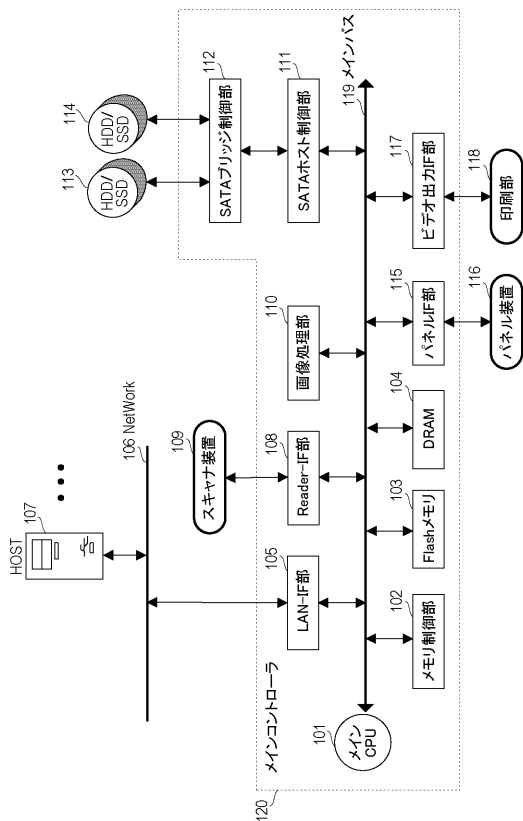
本発明は、上述の実施形態の1以上の機能を実現するプログラムを、ネットワーク又は記憶媒体を介してシステム又は装置に供給し、そのシステム又は装置のコンピュータにおける1つ以上のプロセッサがプログラムを読み出し実行する処理でも実現可能である。また、1以上の機能を実現する回路(例えば、ASIC)によっても実現可能である。

【符号の説明】

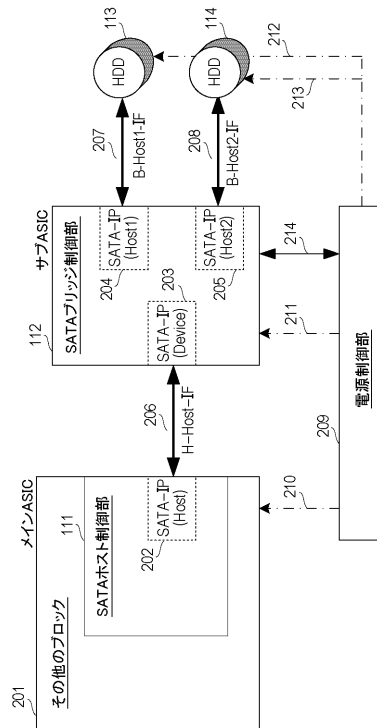
【0150】

- 111 SATAホスト制御部
- 112 SATAブリッジ制御部
- 113 HDD
- 114 HDD

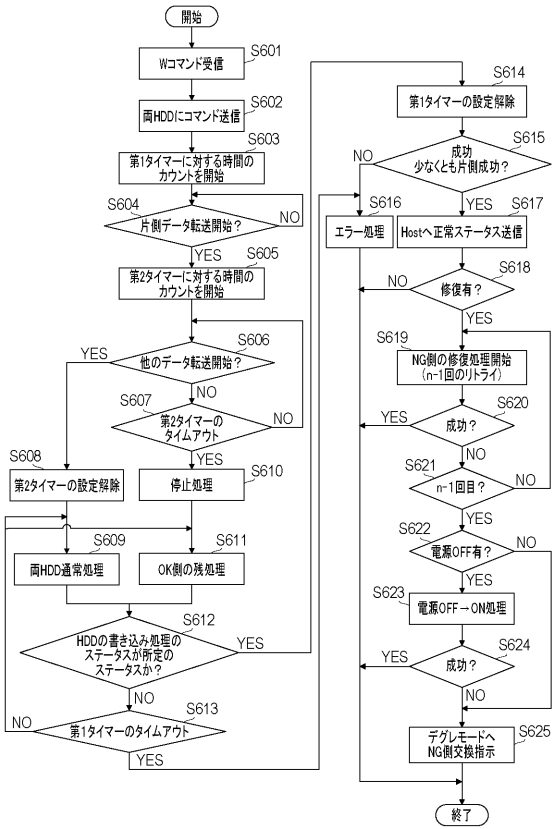
【図1】



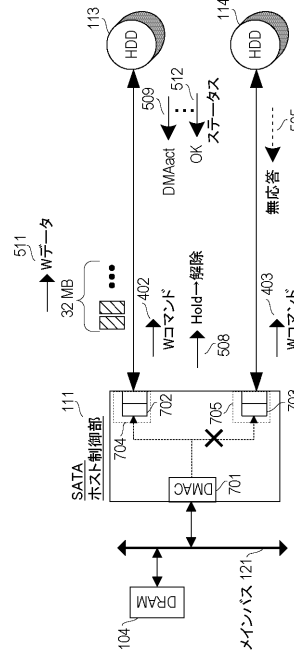
【図2】



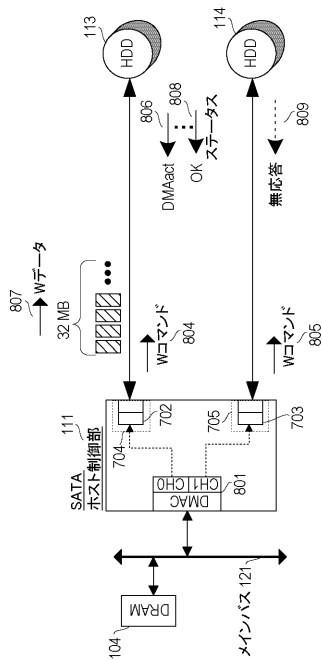
【 図 7 】



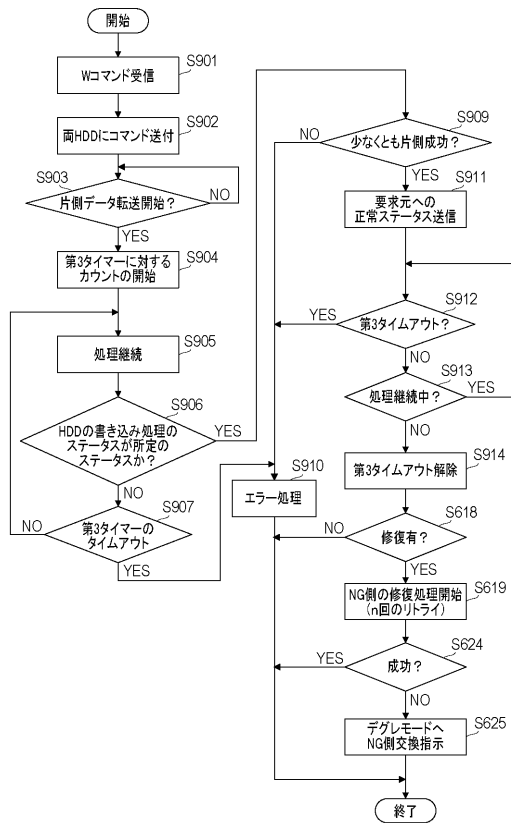
【 図 8 】



【 図 9 】



【 図 10 】



【 図 1 1 】

1001 拡張コマンド名称	1002 CMD番号	1003 転送タイプ	1004 指定内容
1005 SetupBridge	01h	PO	SATAブリッジ制御部の各種設定 (モード設定やその他タイムアウトなどの条件設定等)
1008 GetBridgeInfo	02h	PI	SATAブリッジ制御部のステータス取得
.	.	.	.

1006 1009 1007 1010

フロントページの続き

(51)Int.Cl.

F I

テーマコード(参考)

G 0 6 F 11/07 1 9 6