

(19) 日本国特許庁(JP)

再公表特許(A1)

(11) 国際公開番号

W02009/107213

発行日 平成23年6月30日 (2011.6.30)

(43) 国際公開日 平成21年9月3日 (2009.9.3)

(51) Int.Cl.	F I	テーマコード (参考)
G06F 12/16 (2006.01)	G06F 12/16 340Q	5B005
G06F 12/08 (2006.01)	G06F 12/08 541C	5B018
G06F 3/06 (2006.01)	G06F 12/08 541D	5B065
	G06F 12/08 557	
	G06F 3/06 304Z	

審査請求 有 予備審査請求 未請求 (全 65 頁) 最終頁に続く

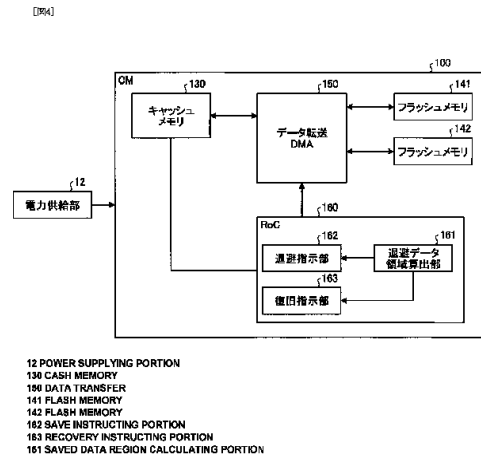
出願番号 特願2010-500488 (P2010-500488)	(71) 出願人 00005223 富士通株式会社 神奈川県川崎市中原区上小田中4丁目1番1号
(21) 国際出願番号 PCT/JP2008/053494	(74) 代理人 100089118 弁理士 酒井 宏明
(22) 国際出願日 平成20年2月28日 (2008.2.28)	(72) 発明者 塚本 新菜 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内
(81) 指定国 AP (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), EA (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), EP (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OA (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG), AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW	(72) 発明者 大山 貞之 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

最終頁に続く

(54) 【発明の名称】 ストレージ装置、ストレージ制御装置、データ転送集積回路、および、ストレージ制御方法

(57) 【要約】

不揮発性メモリから読み出し応答がなされない場合であっても、不揮発性メモリに退避させたデータを、キャッシュメモリに復旧させることを目的とする。データ退避処理時に、退避対象データのパリティデータを生成して、CRCやAIDを付与した退避対象データおよびパリティデータをフラッシュメモリ141または142に書き込み、データ復旧処理において、フラッシュメモリ141または142から所定の時間内にデータ読出処理が完了しない場合に、データ読出処理を中止して付加データを設定し、パリティデータを用いて誤りデータを訂正した退避対象データをキャッシュメモリ130に書き込む。



【特許請求の範囲】

【請求項 1】

異常終了時に、キャッシュメモリに記憶されているデータである退避対象データ、および、前記退避対象データの誤りデータを訂正するためのパリティデータに、誤りデータを検出するための誤り検出符号を付与した転送対象データを、不揮発性メモリへ記憶させるデータ転送部を有するストレージ装置であって、

当該のストレージ装置が異常終了後に電源投入された場合に、前記不揮発性メモリから、前記転送対象データが所定のサイズであるストライプサイズごとに区切られたデータであるストライプを読み出すデータ読出手段と、

前記データ読出手段によって前記ストライプが所定の時間内に読み出されない場合に、該ストライプに、該ストライプのサイズがストライプサイズになるまで、付加データを付与する付加データ付与手段と、

前記データ読出手段によって読み出されたストライプに付与されている誤り検出符号に基づいて、ストライプに誤りデータが存在しているか否かを検査する誤りデータ検査手段と、

前記誤りデータ検査手段によって誤りデータが検出された場合に、前記データ読出手段によって読み出されたストライプを用いて、誤りデータを訂正するデータ訂正手段と、

前記誤りデータ検査手段によって誤りデータが検出されなかったストライプに含まれる退避対象データを前記キャッシュメモリに書き込むとともに、前記データ訂正手段によって誤りデータが訂正されたストライプに含まれる退避対象データを前記キャッシュメモリに書き込むデータ書込手段と

を備えたことを特徴とするストレージ装置。

【請求項 2】

当該のストレージ装置が異常終了した場合に、前記退避対象データのキャッシュメモリにおける先頭の記憶位置を示す先頭アドレスから所定のサイズ分のデータを不揮発性メモリへ記憶させる第一のデータ退避手段と、

当該のストレージ装置が異常終了した場合に、前記退避対象データのキャッシュメモリにおける最終の記憶位置を示す最終アドレスから所定のサイズ分のデータを不揮発性メモリへ記憶させる第二のデータ退避手段とをさらに備えることを特徴とする請求項 1 に記載のストレージ装置。

【請求項 3】

当該のストレージ装置が異常終了した場合に、前記退避対象データの先頭アドレスを、前記第一のデータ退避手段によりキャッシュメモリから最初に読み出されるデータの記憶位置である第一の開始アドレスに決定し、所定のサイズであるバウンダリサイズに整数 N を乗算した値が、前記退避対象データを 2 で除算した値よりも大きくなる N の最小値である第一の基準値を算出して、該第一の基準値に前記バウンダリサイズを乗算した値を、前記第一のデータ退避手段によりキャッシュメモリから読み出されるデータのサイズである第一のデータ退避サイズに決定し、前記バウンダリサイズに整数 N を乗算した値が、前記退避対象データのサイズから前記第一のデータ退避サイズを減算した値よりも大きくなる N の最小値である第二の基準値を算出して、該第二の基準値に前記バウンダリサイズを乗算した値を、前記第二のデータ退避手段によりキャッシュメモリから読み出されるデータのサイズである第二のデータ退避サイズに決定し、前記退避対象データの最終アドレスと前記第一のデータ退避サイズとに基づいて、前記第二のデータ退避手段によりキャッシュメモリから最初に読み出されるデータの記憶位置である第二の開始アドレスを算出する退避データ領域算出手段をさらに備え、

前記第一のデータ退避手段は、前記退避対象データのうち、前記退避データ領域算出手段によって算出された第一の開始アドレスから、前記退避データ領域算出手段によって算出された第一のデータ退避サイズ分のデータを不揮発性メモリへ記憶させ、

前記第二のデータ退避手段は、前記退避対象データのうち、前記退避データ領域算出手段によって算出された第二の開始アドレスから、前記退避データ領域算出手段によって算

10

20

30

40

50

出された第二のデータ退避サイズ分のデータを不揮発性メモリへ記憶させることを特徴とする請求項2に記載のストレージ装置。

【請求項4】

前記誤りデータ検査手段は、前記ストライプに付与されているCRC (Cyclic Redundancy Check) に基づいて、前記ストライプに誤りデータが存在しているか否かを検査することを特徴とする請求項1～3のいずれか一つに記載のストレージ装置。

【請求項5】

前記誤りデータ検査手段は、前記ストライプに付与されている前記ストライプを識別するためのストライプ識別番号に基づいて、前記ストライプがすべて存在しているか否かを検査することを特徴とする請求項1～3のいずれか一つに記載のストレージ装置。

10

【請求項6】

異常終了時に、キャッシュメモリに記憶されているデータである退避対象データ、および、前記退避対象データの誤りデータを訂正するためのパリティデータに、誤りデータを検出するための誤り検出符号を付与した転送対象データを、不揮発性メモリへ記憶させるデータ転送部を有するストレージ装置を制御するストレージ制御装置であって、

前記ストレージ装置が異常終了後に電源投入された場合に、前記不揮発性メモリから、前記転送対象データが所定のサイズであるストライプサイズごとに区切られたデータであるストライプを読み出すデータ読出手段と、

前記データ読出手段によって前記ストライプが所定の時間内に読み出されない場合に、該ストライプに、該ストライプのサイズがストライプサイズになるまで、付加データを付与する付加データ付与手段と、

20

前記データ読出手段によって読み出されたストライプに付与されている誤り検出符号に基づいて、ストライプに誤りデータが存在しているか否かを検査する誤りデータ検査手段と、

前記誤りデータ検査手段によって誤りデータが検出された場合に、前記データ読出手段によって読み出されたストライプを用いて、誤りデータを訂正するデータ訂正手段と、

前記誤りデータ検査手段によって誤りデータが検出されなかったストライプに含まれる退避対象データを前記キャッシュメモリに書き込むとともに、前記データ訂正手段によって誤りデータが訂正されたストライプに含まれる退避対象データを前記キャッシュメモリに書き込むデータ書込手段と

30

を備えたことを特徴とするストレージ制御装置。

【請求項7】

前記ストレージ装置が異常終了した場合に、前記退避対象データのキャッシュメモリにおける先頭の記憶位置を示す先頭アドレスから所定のサイズ分のデータを不揮発性メモリへ記憶させる第一のデータ退避手段と、

前記ストレージ装置が異常終了した場合に、前記退避対象データのキャッシュメモリにおける最終の記憶位置を示す最終アドレスから所定のサイズ分のデータを不揮発性メモリへ記憶させる第二のデータ退避手段とをさらに備えることを特徴とする請求項6に記載のストレージ制御装置。

【請求項8】

40

前記ストレージ装置が異常終了した場合に、前記退避対象データの先頭アドレスを、前記第一のデータ退避手段によりキャッシュメモリから最初に読み出されるデータの記憶位置である第一の開始アドレスに決定し、所定のサイズであるバウンダリサイズに整数Nを乗算した値が、前記退避対象データを2で除算した値よりも大きくなるNの最小値である第一の基準値を算出して、該第一の基準値に前記バウンダリサイズを乗算した値を、前記第一のデータ退避手段によりキャッシュメモリから読み出されるデータのサイズである第一のデータ退避サイズに決定し、前記バウンダリサイズに整数Nを乗算した値が、前記退避対象データのサイズから前記第一のデータ退避サイズを減算した値よりも大きくなるNの最小値である第二の基準値を算出して、該第二の基準値に前記バウンダリサイズを乗算した値を、前記第二のデータ退避手段によりキャッシュメモリから読み出されるデータの

50

サイズである第二のデータ退避サイズに決定し、前記退避対象データの最終アドレスと前記第一のデータ退避サイズとに基づいて、前記第二のデータ退避手段によりキャッシュメモリから最初に読み出されるデータの記憶位置である第二の開始アドレスを算出する退避データ領域算出手段をさらに備え、

前記第一のデータ退避手段は、前記退避対象データのうち、前記退避データ領域算出手段によって算出された第一の開始アドレスから、前記退避データ領域算出手段によって算出された第一のデータ退避サイズ分のデータを不揮発性メモリへ記憶させ、

前記第二のデータ退避手段は、前記退避対象データのうち、前記退避データ領域算出手段によって算出された第二の開始アドレスから、前記退避データ領域算出手段によって算出された第二のデータ退避サイズ分のデータを不揮発性メモリへ記憶させることを特徴とする請求項7に記載のストレージ制御装置。

10

【請求項9】

前記誤りデータ検査手段は、前記ストライプに付与されているCRCに基づいて、前記ストライプに誤りデータが存在しているか否かを検査することを特徴とする請求項6～8のいずれか一つに記載のストレージ制御装置。

【請求項10】

前記誤りデータ検査手段は、前記ストライプに付与されている前記ストライプを識別するためのストライプ識別番号に基づいて、前記ストライプがすべて存在しているか否かを検査することを特徴とする請求項6～8のいずれか一つに記載のストレージ制御装置。

【請求項11】

キャッシュメモリに記憶されているデータである退避対象データ、および、前記退避対象データの誤りデータを訂正するためのパリティデータに、誤りデータを検出するための誤り検出符号を付与した転送対象データを、不揮発性メモリへ記憶させるデータ転送集積回路であって、

20

前記不揮発性メモリから、前記転送対象データが所定のサイズであるストライプサイズごとに区切られたデータであるストライプを読み出すデータ読出手段と、

前記データ読出手段によって前記ストライプが所定の時間内に読み出されない場合に、該ストライプに、該ストライプのサイズがストライプサイズになるまで、付加データを付与する付加データ付与手段と、

前記データ読出手段によって読み出されたストライプに付与されている誤り検出符号に基づいて、ストライプに誤りデータが存在しているか否かを検査する誤りデータ検査手段と、

30

前記誤りデータ検査手段によって誤りデータが検出された場合に、前記データ読出手段によって読み出されたストライプを用いて、誤りデータを訂正するデータ訂正手段と、

前記誤りデータ検査手段によって誤りデータが検出されなかったストライプに含まれる退避対象データを前記キャッシュメモリに書き込むとともに、前記データ訂正手段によって誤りデータが訂正されたストライプに含まれる退避対象データを前記キャッシュメモリに書き込むデータ書込手段と

を備えたことを特徴とするデータ転送集積回路。

【請求項12】

40

前記退避対象データのキャッシュメモリにおける先頭の記憶位置を示す先頭アドレスから所定のサイズ分のデータを不揮発性メモリへ記憶させる第一のデータ退避手段と、

前記退避対象データのキャッシュメモリにおける最終の記憶位置を示す最終アドレスから所定のサイズ分のデータを不揮発性メモリへ記憶させる第二のデータ退避手段とをさらに備えることを特徴とする請求項11に記載のデータ転送集積回路。

【請求項13】

異常終了時に、キャッシュメモリに記憶されているデータである退避対象データ、および、前記退避対象データの誤りデータを訂正するためのパリティデータに、誤りデータを検出するための誤り検出符号を付与した転送対象データを、不揮発性メモリへ記憶させるデータ転送部を有するストレージ装置を制御するストレージ制御方法であって、

50

前記ストレージ装置が、

異常終了後に電源投入された場合に、前記不揮発性メモリから、前記転送対象データが所定のサイズであるストライプサイズごとに区切られたデータであるストライプを読み出すデータ読出工程と、

前記データ読出工程によって前記ストライプが所定の時間内に読み出されない場合に、該ストライプに、該ストライプのサイズがストライプサイズになるまで、付加データを付与する付加データ付与工程と、

前記データ読出工程によって読み出されたストライプに付与されている誤り検出符号に基づいて、ストライプに誤りデータが存在しているか否かを検査する誤りデータ検査工程と、

前記誤りデータ検査工程によって誤りデータが検出された場合に、前記データ読出工程によって読み出されたストライプを用いて、誤りデータを訂正するデータ訂正工程と、

前記誤りデータ検査工程によって誤りデータが検出されなかったストライプに含まれる退避対象データを前記キャッシュメモリに書き込むとともに、前記データ訂正工程によって誤りデータが訂正されたストライプに含まれる退避対象データを前記キャッシュメモリに書き込むデータ書込工程と

を含んだことを特徴とするストレージ制御方法。

【請求項 14】

前記ストレージ装置は、

異常終了した場合に、前記退避対象データのキャッシュメモリにおける先頭の記憶位置を示す先頭アドレスから所定のサイズ分のデータを不揮発性メモリへ記憶させる第一のデータ退避工程と、

異常終了した場合に、前記退避対象データのキャッシュメモリにおける最終の記憶位置を示す最終アドレスから所定のサイズ分のデータを不揮発性メモリへ記憶させる第二のデータ退避工程とをさらに含んだことを特徴とする請求項 13 に記載のストレージ制御方法。

【請求項 15】

前記ストレージ装置は、

異常終了した場合に、前記退避対象データの先頭アドレスを、前記第一のデータ退避工程によりキャッシュメモリから最初に読み出されるデータの記憶位置である第一の開始アドレスに決定し、所定のサイズであるバウンダリサイズに整数 N を乗算した値が、前記退避対象データを 2 で除算した値よりも大きくなる N の最小値である第一の基準値を算出して、該第一の基準値に前記バウンダリサイズを乗算した値を、前記第一のデータ退避工程によりキャッシュメモリから読み出されるデータのサイズである第一のデータ退避サイズに決定し、前記バウンダリサイズに整数 N を乗算した値が、前記退避対象データのサイズから前記第一のデータ退避サイズを減算した値よりも大きくなる N の最小値である第二の基準値を算出して、該第二の基準値に前記バウンダリサイズを乗算した値を、前記第二のデータ退避工程によりキャッシュメモリから読み出されるデータのサイズである第二のデータ退避サイズに決定し、前記退避対象データの最終アドレスと前記第一のデータ退避サイズとに基づいて、前記第二のデータ退避工程によりキャッシュメモリから最初に読み出されるデータの記憶位置である第二の開始アドレスを算出する退避データ領域算出工程をさらに含み、

前記第一のデータ退避工程は、前記退避対象データのうち、前記退避データ領域算出工程によって算出された第一の開始アドレスから、前記退避データ領域算出工程によって算出された第一のデータ退避サイズ分のデータを不揮発性メモリへ記憶させ、

前記第二のデータ退避工程は、前記退避対象データのうち、前記退避データ領域算出工程によって算出された第二の開始アドレスから、前記退避データ領域算出工程によって算出された第二のデータ退避サイズ分のデータを不揮発性メモリへ記憶させることを特徴とする請求項 14 に記載のストレージ制御方法。

【請求項 16】

10

20

30

40

50

前記誤りデータ検査工程は、前記ストライプに付与されているCRCに基づいて、前記ストライプに誤りデータが存在しているか否かを検査することを特徴とする請求項13～15のいずれか一つに記載のストレージ制御方法。

【請求項17】

前記誤りデータ検査工程は、前記ストライプに付与されている前記ストライプを識別するためのストライプ識別番号に基づいて、前記ストライプがすべて存在しているか否かを検査することを特徴とする請求項13～15のいずれか一つに記載のストレージ制御方法。

【発明の詳細な説明】

【技術分野】

10

【0001】

本発明は、ストレージ装置、ストレージ制御装置、データ転送集積回路、および、ストレージ制御方法に関する。

【背景技術】

【0002】

従来、大規模なコンピュータシステムでは、ホストコンピュータと別体に構成された専用のデータ入出力装置を用いてデータを管理している。このデータ入出力装置の中でも、ストレージ装置は、複数のハードディスク装置（記憶媒体）によってRAID（Redundant Arrays of Independent (Inexpensive) Disks）グループを構成して、取り扱うデータの信頼性を向上させている。

20

【0003】

ストレージ装置は、一般に、ライトバック方式と呼ばれるライト処理（書込処理）方式を行うことで、アクセス性能を向上させている。具体的には、ストレージ装置は、上位装置であるホストコンピュータからデータのライト命令を受け付けた場合に、キャッシュメモリにデータを記憶した時点で、ライト処理が完了した旨をホストコンピュータへ通知する。その後、所定の条件を満たした時点で、キャッシュメモリに記憶したデータをハードディスク装置へ記憶させる。

【0004】

このようなライトバック方式を採用するストレージ装置には、停電などによって正常な終了操作が行われることなく電源が落とされた場合であっても（以下、正常な終了操作が行われることなく電源が落とされることを「異常終了」という）、揮発性のキャッシュメモリに記憶されているデータが消失しないように構成されているものがある。

30

【0005】

具体的には、異常終了した場合に、大容量のコンデンサやバッテリー等から電力供給されている間に、キャッシュメモリ上のデータを不揮発性メモリに退避させ、異常終了後に電源投入された場合に不揮発性メモリに退避したデータをキャッシュメモリに復旧させる。これにより、キャッシュメモリに記憶されているデータがハードディスク装置に記憶されていない状態で異常終了した場合であっても、データを消失させないことを実現している。

【0006】

40

ところで、データを退避させる不揮発性メモリは、一般的に、NAND型のフラッシュメモリが用いられることが多い。NAND型のフラッシュメモリは、セル劣化等によって記憶しているデータを破損や消失してしまうことがある。不揮発性メモリに記憶されているデータが破損や消失する状況が発生すると、ストレージ装置は、異常終了後の起動時に、不揮発性メモリから誤ったデータを読み出してしまい、正確にデータを復旧させることができなくなる。

【0007】

そこで、記憶媒体に記憶されているデータのエラーを検出し、エラー検出時にデータ訂正を行う技術（例えば、特許文献1および2を参照）を用いることが考えられる。このような技術を用いれば、異常終了後の起動時に、不揮発性メモリに退避させたデータを、キ

50

キャッシュメモリに復旧させることが可能になる。

【0008】

【特許文献1】特表2004-506256号公報

【特許文献2】特開2007-193449号公報

【発明の開示】

【発明が解決しようとする課題】

【0009】

しかしながら、上記特許文献1および2に開示されているような技術を用いても、不揮発性メモリから読み出し応答がなされない場合には、不揮発性メモリに退避させたデータを、正確にキャッシュメモリに復旧させることができないという問題があった。不揮発性メモリの劣化状態によっては、不揮発性メモリから読み出し応答がなされないことがあり、かかる状態では、ハングアップ状態となりデータ訂正処理が実行されないためである。

10

【0010】

ストレージ装置のような大規模なデータ入出力装置は、重要なコンピュータシステムに用いられることが多く、キャッシュメモリに記憶されているユーザデータやシステムデータを正確に復旧できないことは、大きな問題となる。このようなことから、不揮発性メモリから読み出し応答がなされない場合であっても、不揮発性メモリに退避させたデータを、いかにして正確にキャッシュメモリに復旧させるかが重要な課題となっていた。

【0011】

本発明は、上述した従来技術による課題を解消するためになされたものであり、不揮発性メモリから読み出し応答がなされない場合であっても、不揮発性メモリに退避させたデータを、キャッシュメモリに復旧させることができるストレージ装置、ストレージ制御装置、データ転送集積回路、および、ストレージ制御方法を提供することを目的とする。

20

【課題を解決するための手段】

【0012】

上述した課題を解決し、目的を達成するために、本願に開示するストレージ装置は、異常終了時に、キャッシュメモリに記憶されているデータである退避対象データ、および、前記退避対象データの誤りデータを訂正するためのパリティデータに、誤りデータを検出するための誤り検出符号を付与した転送対象データを、不揮発性メモリへ記憶させるデータ転送部を有するストレージ装置であって、当該のストレージ装置が異常終了後に電源投入された場合に、前記不揮発性メモリから、前記転送対象データが所定のサイズであるストライプサイズごとに区切られたデータであるストライプを読み出すデータ読出手段と、前記データ読出手段によって前記ストライプが所定の時間内に読み出されない場合に、該ストライプに、該ストライプのサイズがストライプサイズになるまで、付加データを付与する付加データ付与手段と、前記データ読出手段によって読み出されたストライプに付与されている誤り検出符号に基づいて、ストライプに誤りデータが存在しているか否かを検査する誤りデータ検査手段と、前記誤りデータ検査手段によって誤りデータが検出された場合に、前記データ読出手段によって読み出されたストライプを用いて、誤りデータを訂正するデータ訂正手段と、前記誤りデータ検査手段によって誤りデータが検出されなかったストライプに含まれる退避対象データを前記キャッシュメモリに書き込むとともに、前記データ訂正手段によって誤りデータが訂正されたストライプに含まれる退避対象データを前記キャッシュメモリに書き込むデータ書込手段とを備えたことを要件とする。

30

40

【0013】

なお、本願に開示するストレージ装置の構成要素、表現または構成要素の任意の組合せを、方法、装置、システム、コンピュータプログラム、記録媒体、データ構造などに適用したのも、他の態様として有効である。

【発明の効果】

【0014】

本願に開示したストレージ装置によれば、不揮発性メモリからデータの読出応答がなされない場合であっても、不揮発性メモリに退避させたデータを、キャッシュメモリに復旧

50

させることができるという効果を奏する。

【図面の簡単な説明】

【0015】

【図1】図1は、実施例1に係るストレージ装置の概略構成を示す図である。

【図2-1】図2-1は、フラッシュメモリ141の構成およびデータフォーマットを示す図である。

【図2-2】図2-2は、図2-1に示したpageの構成およびデータフォーマットを示す図である。

【図3-1】図3-1は、実施例1に係るストレージ装置によるデータ退避処理の概要を説明するための図である。

【図3-2】図3-2は、実施例1に係るストレージ装置によるデータ復旧処理の概要を説明するための図である。

【図3-3】図3-3は、実施例1に係るストレージ装置によるデータ復旧処理の概要を説明するための図である。

【図4】図4は、実施例1に係るストレージ装置が有するCMの構成を示す図である。

【図5】図5は、図4に示したデータ転送DMAの構成を示す図である。

【図6】図6は、図5に示したライトDMAの構成を示す図である。

【図7】図7は、図5に示したリードDMAの構成を示す図である。

【図8】図8は、実施例1に係るストレージ装置によるデータ退避処理の流れを示すシーケンス図である。

【図9】図9は、ライトDMAによるデータ退避処理手順を示すフローチャートである。

【図10】図10は、実施例1に係るストレージ装置によるデータ復旧処理の流れを示すシーケンス図である。

【図11】図11は、リードDMAによるデータ復旧処理手順を示すフローチャートである。

【図12-1】図12-1は、実施例2に係るストレージ装置によるデータ退避処理の概要を説明するための図である。

【図12-2】図12-2は、実施例2に係るストレージ装置によるデータ復旧処理の概要を説明するための図である。

【図13】図13は、実施例2に係るストレージ装置が有するCMの構成を示す図である。

【図14】図14は、図13に示した退避データ領域算出部による退避データ領域算出処理手順を示すフローチャートである。

【符号の説明】

【0016】

1、2	ストレージ装置
10	CE
11	HDD
12	電力供給部
20	DE
100、200	CM
110	I/F部
120、220	エクспанダ
130	キャッシュメモリ
141、142	フラッシュメモリ
150	データ転送DMA
160、260	ROC
161、261	退避データ領域算出部
162、262	退避指示部
163	復旧指示部

10

20

30

40

50

5 1 0	P C I バス I / F	
5 2 1、5 2 2	メモリー / F	
5 3 0	ディスクリプタ保持レジスタ	
5 4 1、5 4 2	ライト D M A	
5 4 1 a	データ読出部	
5 4 1 b	バウンダリ調整部	
5 4 1 c	パリティ生成部	
5 4 1 d	退避データバッファ	
5 4 1 e	パリティバッファ	
5 4 1 f	データ書込部	10
5 4 1 g	C R C 付与部	
5 4 1 h	A I D 付与部	
5 5 0	リード D M A	
5 5 0 a	セレクタ	
5 5 0 b	データ読出部	
5 5 0 c	タイムアウト検出部	
5 5 0 d	エラー検出部	
5 5 0 e	復旧データバッファ	
5 5 0 f	X O R 演算部	
5 5 0 g	C R C 検査部	20
5 5 0 h	A I D 検査部	
5 5 0 i	訂正データバッファ	
5 5 0 j	ログ保持レジスタ	
5 5 0 k	セレクタ	
5 5 0 l	バウンダリ調整部	
5 5 0 m	データ書込部	
5 6 1、5 6 2	イレース D M A	

【発明を実施するための最良の形態】

【0017】

以下に、本発明にかかるストレージ装置、ストレージ制御装置、データ転送集積回路、および、ストレージ制御方法の実施例を図面に基づいて詳細に説明する。なお、この実施例によりこの発明が限定されるものではない。 30

【実施例1】

【0018】

まず、実施例1に係るストレージ装置の概略構成について説明する。図1は、実施例1に係るストレージ装置1の概略構成を示す図である。同図に示したストレージ装置1は、高信頼性を確保するために二重化構成が採られており、各構成要素が2系統ずつ（同図中、参照符号に付加したA、Bによって区別する）備えられている。それぞれの系統の構成要素は互いに同等の機能を有するため、以下では一方の系統についてのみ構成要素の説明をする。 40

【0019】

同図に示すように、ストレージ装置1は、ストレージ装置1を制御する各部や、ストレージ装置1と上位装置であるホストコンピュータとの間の通信を担う通信モジュールなどを主に搭載するC E（Controller Enclosure：コントローラ エンクロージャ）10と、ハードディスク装置を主に搭載するD E（Device Enclosure：デバイス エンクロージャ）20とに、筐体が分けられている。

【0020】

C E 10は、図示しないチャネルアダプタ経由でファイバチャネルを介してホストコンピュータと接続され、エキスパンダ120Aおよび120B経由でD E 20と接続されている。D E 20は、ハードディスク装置（以下、「H D D：Hard Disk Drive」と略記 50

する) 21を有し、エキスパンダ220Aおよび220Bを直接、もしくは、ルータ(Router)を経由するなどして、他のDEと複数個接続することが可能である。

【0021】

また、CE10は、ストレージ装置1を制御するためのCM(Controller Module:コントローラモジュール)100Aおよび100Bと、CM100Aおよび100Bに接続される複数のハードディスク装置11とを有する。

【0022】

HDD11は、CM100Aによってデータが冗長に読み書きされる記憶装置である。なお、CM100Aに接続される記憶装置は、ハードディスク装置に限られず、熱磁気ディスク装置や光磁気ディスク装置、または、半導体不揮発メモリを用いた半導体ディスクのような他の記憶装置であってもよい。

10

【0023】

CM100Aは、ストレージ装置1を制御するストレージ制御装置であり、インタフェース(以下、「I/F」と略記する)部110Aと、エキスパンダ120Aと、キャッシュメモリ130Aと、フラッシュメモリ141Aおよび142Aと、データ転送DMA(Direct Memory Access)150Aと、RAID-on-Chip(以下、「RoC」と略記する)160Aとを有する。

【0024】

I/F部110Aは、LAN(Local Area Network)などのネットワークへ接続するためのインタフェースである。エキスパンダ120Aは、RoC160AとHDD11との間で送受信されるデータを中継する中継装置である。

20

【0025】

キャッシュメモリ130Aは、ホストコンピュータとHDD11との間で転送が行われるデータを一時的に記憶するメモリである。フラッシュメモリ141Aおよび142Aは、不揮発性メモリであり、例えば、NAND型のフラッシュメモリである。実施例1に係るストレージ装置1では、フラッシュメモリ141Aおよび142Aに、OneNAND(ワンナンド)型のフラッシュメモリを用いることとする。

【0026】

ここで、OneNAND型のフラッシュメモリ141Aおよび142Aの構成およびデータフォーマットについて説明する。なお、フラッシュメモリ141Aおよび142Aの構成およびデータフォーマットは同一であるため、ここでは、フラッシュメモリ141Aの構成およびデータフォーマットについてのみ説明する。

30

【0027】

図2-1は、フラッシュメモリ141Aの構成およびデータフォーマットを示す図である。なお、フラッシュメモリ141Aの記憶容量は、2[GB(ギガバイト)]であるとす。同図に示すように、フラッシュメモリ141Aは、4つのメモリバンク#00~メモリバンク#03を有する。メモリバンク#00~メモリバンク#03は、メモリの集合体であり、それぞれがダイ#0およびダイ#1を有する。

【0028】

ダイ#0およびダイ#1は、数ミリ角のフラッシュメモリであり、それぞれがBlock#0~Block#2047を有する。また、Block#0~#2047は、それぞれがpage(ページ)#0~page#63を有する。pageが、フラッシュメモリ141への最小アクセス単位となる。

40

【0029】

pageの構成およびデータフォーマットについて説明する。図2-2は、図2-1に示したpageの構成およびデータフォーマットを示す図である。同図に示すように、pageは、メイン領域と、スペア領域とを有する。

【0030】

メイン領域は、ユーザデータが記憶される2[KB(キロバイト)]の記憶領域であり、4個のセクタ#0~セクタ#3を有する。スペア領域は、主にデータを管理するための

50

システムデータ等が記憶される64 [byte (バイト)] の記憶領域であり、4個のスペア#0~#3を有する。

【0031】

スペア領域を構成する1つのスペア(例えば、スペア#0)は、同図に示すように、システム領域と、ECC(Error Correcting Code: 誤り訂正符号)領域と、リザーブ領域と、ユーザ領域とを有する。

【0032】

システム領域は、不良セクタ情報等の各種情報が記憶される記憶領域である。ECC領域は、フラッシュメモリ141Aによって生成されるECCが記憶される領域である。フラッシュメモリ141Aは、このECCを用いてエラーチェックを行い、エラーを検出した場合に、ECCチェックエラーが発生した旨を通知する。リザーブ領域は、予備領域である。ユーザ領域は、フラッシュメモリ141Aにデータを記憶させるシステムによってデータ管理上必要となるデータが記憶される記憶領域である。

10

【0033】

図1の説明に戻って、データ転送DMA150Aは、キャッシュメモリ130Aと、フラッシュメモリ141Aおよび142Aとの間におけるデータ送受信をDMAにより行うDMA回路である。このデータ転送DMA150Aは、例えば、FPGA(Field Programmable Gate Array)によって実現される。

【0034】

RoC160Aは、CM100Aを全体制御する制御装置であり、ホストコンピュータとのインタフェースを制御したり、キャッシュメモリ130Aの管理を行ったりする。また、RoC160Aは、ホストコンピュータからデータのライト命令を受け付けた場合に、かかるデータをキャッシュメモリ130Aに記憶する。そして、RoC160Aは、キャッシュメモリ130Aにデータを記憶した時点で、ライト処理が完了した旨をホストコンピュータへ通知する。その後、所定の条件を満たした時点で、キャッシュメモリ130Aに記憶したデータをHDD11に書き込みする。

20

【0035】

また、RoC160Aは、ストレージ装置1が異常終了した場合に、データ転送DMA150Aに対して、キャッシュメモリ130Aに記憶されているデータをフラッシュメモリ141Aおよび142Aに退避させる処理(以下、「データ退避処理」という)を行うように指示する。また、RoC160Aは、異常終了後のストレージ装置1が起動した場合に、データ転送DMA150Aに対して、フラッシュメモリ141Aおよび142Aに退避したデータをキャッシュメモリ130Aに復旧させる処理(以下、「データ復旧処理」という)を行うように指示する。

30

【0036】

次に、図3-1~図3-3を用いて、実施例1に係るストレージ装置1によるデータ退避処理およびデータ復旧処理の概要について説明する。図3-1は、実施例1に係るストレージ装置1によるデータ退避処理の概要を説明するための図である。なお、以下では、図1中に付加した参照符号のAおよびBを付与せずに説明する。すなわち、例えば、図3-1に示したキャッシュメモリ130は、図1に示したキャッシュメモリ130Aまたは130Bに対応する。

40

【0037】

図3-1に示すように、キャッシュメモリ130には、53 [KB] のデータD1が記憶されているものとする。この状態で、ストレージ装置1が異常終了したため、RoC160が、データ転送DMA150に対して、データD1を退避するように指示したとする。かかる指示を受け付けたデータ転送DMA150は、キャッシュメモリ130からデータD1を読み出す。なお、以下では、ストレージ装置1が異常終了した時に、データ転送DMA150によって退避されるキャッシュメモリ130上のデータを「退避対象データ」と呼ぶこととする。

【0038】

50

そして、データ転送DMA150は、データD1を、28[KB]ごとに区切る。具体的には、同図に示すように、データ転送DMA150は、データD1を、28[KB]のデータD10と、25[KB]のデータD20とに区切る。なお、データ転送DMA150が退避対象データ(上記例ではデータD1)を分割するサイズ(上記例では、28[KB])は、予めシステムで決められており、以下では、このサイズを「バウンダリサイズ」と呼び、バウンダリサイズによって区切られた退避対象データの1つのデータを「バウンダリデータ」と呼ぶこととする。実施例1に係るストレージ装置1では、バウンダリサイズが28[KB]に決定されていることとする。

【0039】

そして、データ転送DMA150は、データD10を、2[KB]ごとに区切る。具体的には、同図に示すように、データ転送DMA150は、データD10を、14個のデータD101~D114に区切る。なお、データ転送DMA150がバウンダリデータ(上記例では、データD10およびD20)を分割するサイズ(上記例では、2[KB])は、予めシステムで決められており、実施例1に係るストレージ装置1では、このサイズが、2[KB]に決定されていることとする。以下では、2[KB]ごとに区切られたバウンダリデータの1つのデータ(データD101等)を、「退避対象単位データ」と呼ぶこととし、退避対象単位データを識別するための番号を「退避対象単位データ番号」と呼ぶこととする。本明細書では、退避対象単位データに付した参照符号の数字部分を退避対象単位データ番号とする。例えば、データD101の退避対象単位データ番号は、「101」である。

10

20

【0040】

そして、データ転送DMA150は、退避対象単位データ番号が奇数である7個のデータD101、D103、・・・、D113の排他的論理和(以下、「XOR」という)演算を行ってパリティデータP115を生成する。また、データ転送DMA150は、退避対象単位データ番号が偶数である7個のデータD102、D104、・・・、D114のXOR演算を行ってパリティデータP116を生成する。

【0041】

このパリティデータP115およびP116のサイズは、データD101~D114のサイズと同様の2[KB]となる。すなわち、バウンダリサイズ28[KB]に区切られたデータD10のサイズと、かかるデータD10から生成されたパリティデータP115およびP116のサイズの合計は、32[KB]となる。なお、以下では、パリティデータも「退避対象単位データ」と呼ぶこととする。

30

【0042】

その後、データ転送DMA150は、データD101~D114ごとに、CRC(Cyclic Redundancy Check)を生成する。同様に、データ転送DMA150は、パリティデータP115およびP116ごとに、CRCを生成する。そして、データ転送DMA150は、生成したCRCを、データD101~D114と、パリティデータP115およびP116とに付与する。

【0043】

また、データ転送DMA150は、データD101~D114と、パリティデータP115およびP116とを識別するための識別番号(以下、「AID:Area Identification」という)をデータD101~D114と、パリティデータP115およびP116とに付与する。このAIDについては、後に詳述する。

40

【0044】

そして、データ転送DMA150は、CRCおよびAIDが付与されたデータD101~D114と、CRCおよびAIDが付与されたパリティデータP115およびP116とをフラッシュメモリ141に書き込む。すなわち、データ転送DMA150は、16個の退避対象単位データ(14個のデータD101~D114と、2個のパリティデータP115およびP116)を、フラッシュメモリ141に書き込む。

【0045】

50

このとき、データ転送DMA150は、フラッシュメモリ141が有する4個のメモリバンク#00～メモリバンク#03に対して、退避対象単位データを4個ずつ書き込む。また、このとき、データ転送DMA150は、メモリバンク#00～メモリバンク#03が有する2個のダイ#0およびダイ#1に対して、退避対象単位データを2個ずつ書き込む。また、このとき、データ転送DMA150は、ダイ#0およびダイ#1が有する所定の2個のBlockに対して、退避対象単位データ番号が奇数である退避対象単位データと、退避対象単位データ番号が偶数である退避対象単位データとを1個ずつ書き込む。

【0046】

以下では、データ転送DMA150が、16個の退避対象単位データを書き込むフラッシュメモリ141上の領域を「小エリア」と呼ぶこととする。この「小エリア」は、例えば、図2-1に示した小エリアSA1に該当する。

10

【0047】

ここで、データ転送DMA150のデータ書込処理について、データD101～D114と、パリティデータP115およびP116とを、図2-1に示した小エリアAS1に書き込む場合を例に挙げて説明する。なお、以下では、「メモリバンク#Pのダイ#QのBlock#Rが有するpage#S」を示す場合に、「メモリバンク#P/ダイ#Q/Block#R/page#S」と記載することとする。

【0048】

かかる場合、データ転送DMA150は、例えば、データD101を、小エリアSA1内のメモリバンク#00/ダイ#0/Block#2/page#1のメイン領域に書き込む。そして、データ転送DMA150は、データD102を、メモリバンク#00/ダイ#0/Block#3/page#1のメイン領域に書き込み、データD103を、メモリバンク#00/ダイ#1/Block#2/page#1のメイン領域に書き込む。このようにして、データ転送DMA150は、データD104～D113を書き込んでいき、データD114を、メモリバンク#03/ダイ#0/Block#3/page#1に書き込み、パリティデータP115を、メモリバンク#03/ダイ#1/Block#2/page#1に書き込み、パリティデータP116を、メモリバンク#03/ダイ#1/Block#3/page#1に書き込む。

20

【0049】

そして、データ転送DMA150は、各退避対象単位データに付与したCRCをpageのスペア領域内のユーザ領域に書き込む。上記例のように、データD101を、メモリバンク#00/ダイ#0/Block#2/page#1のメイン領域に書き込んだ場合、データ転送DMA150は、データD101のCRCを、メモリバンク#00/ダイ#0/Block#2/page#1のスペア#0のユーザ領域に書き込む。同様に、データ転送DMA150は、データD102～D114と、パリティデータP115およびP116のCRCについても、pageのユーザ領域に書き込む。

30

【0050】

また、データ転送DMA150は、各退避対象単位データに付与したAIDをpageのスペア領域内のユーザ領域に書き込む。上記例のように、データD101のCRCを、メモリバンク#00/ダイ#0/Block#2/page#1のスペア#0のユーザ領域に書き込んだ場合、データ転送DMA150は、データD101のAIDを、メモリバンク#00/ダイ#0/Block#2/page#1のスペア#1のユーザ領域に書き込む。同様に、データ転送DMA150は、データD102～D114と、パリティデータP115およびP116のAIDについても、pageのユーザ領域に書き込む。

40

【0051】

このようにして、データ転送DMA150は、16個の退避対象単位データを、小エリアに書き込む。以下では、小エリア内の1つのpageに記憶されている全てのデータ(退避対象単位データやCRCやAID等)を「ストライプ」と呼ぶこととし、小エリア内でストライプを識別するための番号を「ストライプ番号」と呼ぶこととする。本明細書では、退避対象単位データに付した参照符号の数字下2桁部分をストライプ番号とする。例

50

例えば、データD101を記憶するpageに記憶されているストライプのストライプ番号は「1」であり、パリティデータP115を記憶するpageに記憶されているストライプのストライプ番号は「15」である。以下では、ストライプ番号「N」のストライプを、単に、ストライプS「N」と呼ぶこととする。

【0052】

なお、図3-1に示したフラッシュメモリ141内の「D101」や「P115」等を記載した矩形は、pageのメイン領域を示し、かかるメイン領域の後に斜線を付して示した矩形は、pageのスペア領域を示す。すなわち、同図では、例えば、「D101」を記載したメイン領域と、かかるメイン領域の後に示したスペア領域との組合せがストライプS1であることを示す。

10

【0053】

同様にして、データ転送DMA150は、データD20をフラッシュメモリ142に書き込む。具体的には、データ転送DMA150は、データD20を2[KB]ごとに区切る。このとき、データ転送DMA150は、データD20のサイズが25[KB]であり、バウンダリサイズの28[KB]に満たないので、データD20のサイズが28[KB]になるように、3[KB]の所定の値を加える。以下では、この所定の値を「バウンダリ調整値」と呼ぶこととする。バウンダリデータのサイズがバウンダリサイズに満たない場合に、バウンダリ調整値を加える理由は、パリティデータを生成できるようにするためである。

20

【0054】

そして、データ転送DMA150は、バウンダリ調整値を加えたデータD20を2[KB]ごとに区切る。具体的には、同図に示すように、データ転送DMA150は、3[KB]のバウンダリ調整値を加えたデータD20を、14個のデータD201~D214に区切る。

【0055】

その後の処理は、上述したデータD10をフラッシュメモリ141に書き込む処理と同様であり、データ転送DMA150は、パリティデータP215およびP216を生成して、データD201等のCRCを生成して、データD201~D214と、パリティデータP215およびP216と、CRCと、AIDとをフラッシュメモリ142に書き込む。

30

【0056】

このようにして、データ転送DMA150は、キャッシュメモリ130に記憶されているデータD1のうち、データD10をフラッシュメモリ141に書き込み、データD20をフラッシュメモリ142に書き込む。

【0057】

なお、データ転送DMA150は、複数の小エリアに対して、データを並行して書き込むことが可能である。実施例1に係るストレージ装置1では、データ転送DMA150が、最大で64個の小エリアに対して同時にデータを書き込むこととする。以下では、この64個の小エリアを「エリア」と呼ぶこととする。この「エリア」は、例えば、図2-1に示したエリアA1に該当する。上述したAIDは、1個のエリア内に含まれるストライプを識別するための識別番号である。このAIDは、Roc160によって算出され、データ転送DMA150は、Roc160からAIDを受け付ける。

40

【0058】

次に、実施例1に係るストレージ装置1によるデータ復旧処理の概要について説明する。図3-2および図3-3は、実施例1に係るストレージ装置1によるデータ復旧処理の概要を説明するための図である。図3-2および図3-3に示したフラッシュメモリ141には、図3-1に示したデータ転送DMA150によって退避されたデータが記憶されているものとする。

【0059】

まず、図3-2を用いて、図3-1に示したデータD10をキャッシュメモリ130に

50

復旧させる処理について説明する。ここでは、図3-2に示したフラッシュメモリ141に記憶されているデータD103の一部が、フラッシュメモリ141のセル劣化等の原因により破損しているものとする。そして、フラッシュメモリ141は、データD103の読出要求を受け付けると、かかる読出要求に対する応答ができなくなるものとする。

【0060】

かかる状態で、ストレージ装置1が異常終了後に起動した場合、データ転送DMA150は、フラッシュメモリ141から、小エリア単位でストライプを読み出す。具体的には、データ転送DMA150は、フラッシュメモリ141から、ストライプS1～S16を読み出す。

【0061】

そして、データ転送DMA150は、ストライプS3を読み出している途中で、フラッシュメモリ141から読出応答を受け付けなくなる。所定の時間が経過してもストライプS3を読み出せない場合、データ転送DMA150は、ストライプS3に対してタイムアウトが発生したことを検出する。

【0062】

そして、データ転送DMA150は、タイムアウトが発生したストライプS3の読出処理を中止して、ストライプS3に、読み出しできなかったサイズ分の所定の値(以下、この所定の値を「付加データ」という)を設定する。例えば、1[KB]分のストライプS3を読み出した時点で、フラッシュメモリ141から読み出し応答を受け付けなくなった場合、データ転送DMA150は、読み出し済みの1[KB]のデータに1[KB]の付加データを加えた2[KB]のデータをストライプS3とする。その後、データ転送DMA150は、データの読出処理を再開して、フラッシュメモリ141から、ストライプS4～S16を読み出す。

【0063】

また、データ転送DMA150は、ストライプS1～16を読み出しながら、CRC検査、AID検査およびXOR演算を行う。具体的には、データ転送DMA150は、読み出したストライプS1～16内のCRCを用いて、ストライプS1～16に誤りデータが存在するか否かを検査する。また、データ転送DMA150は、ストライプS1～16内のAIDを用いて、ストライプS1～16がすべて揃っているか否かを検査する。また、データ転送DMA150は、ストライプ番号が奇数であるストライプS1、S3、・・・、S15のXOR演算を行い、訂正データT11を生成するとともに、ストライプ番号が偶数であるストライプS2、S4、・・・、S16のXOR演算を行い、訂正データT12を生成する。

【0064】

図3-2に示した例では、データ転送DMA150は、付加データが設定されているストライプS3のCRC検査において、誤りデータが存在することを検出する。なお、ストライプS3に含まれるAIDを読み出す前に、タイムアウトが発生した場合には、AIDには付加データが設定されているため、データ転送DMA150は、ストライプS3のAID検査において、データD103が欠如していると検出する可能性もある。ここでは、データ転送DMA150は、ストライプS3に対してのみデータ誤りを検出し、ストライプS1、S2、S4～S16に対しては、CRC検査およびAID検査においてエラーを検出しなかったものとする。

【0065】

すべてのストライプS1～16を読み出した後、データ転送DMA150は、CRC検査およびAID検査においてエラーが検出されなかったストライプS1およびS2に含まれるデータD101およびD102をキャッシュメモリ130に書き込む。そして、データ転送DMA150は、エラーが検出されたストライプS3については、ストライプ番号が奇数であるストライプS1、S5、S7、・・・、S15および訂正データT11を用いて、データD103の誤りデータを訂正して、訂正済みのデータD103をキャッシュメモリ130に書き込む。そして、データ転送DMA150は、エラーが検出されなかった

10

20

30

40

50

ストライプ S 4 ~ S 1 4 に含まれるデータ D 1 0 4 ~ D 1 1 4 をキャッシュメモリ 1 3 0 に書き込む。なお、ストライプ S 1 5 および S 1 6 はパリティデータであるため、データ転送 D M A 1 5 0 は、ストライプ S 1 5 および S 1 6 に含まれるパリティデータ P 1 1 5 および P 1 1 6 をキャッシュメモリ 1 3 0 に書き込まない。このようにして、データ転送 D M A 1 5 0 は、データ D 1 0 をキャッシュメモリ 1 3 0 に復旧する。

【 0 0 6 6 】

続いて、図 3 - 3 を用いて、図 3 - 1 に示したデータ D 2 0 をキャッシュメモリ 1 3 0 に復旧させる処理について説明する。同図に示すように、データ転送 D M A 1 5 0 は、データ D 1 0 の復旧処理が終了した後、フラッシュメモリ 1 4 2 から、ストライプ S 1 ~ S 1 6 を読み出しながら、C R C 検査、A I D 検査および X O R 演算を行い、訂正データ T 2 1 および T 2 2 を生成する。ここでは、データ転送 D M A 1 5 0 は、フラッシュメモリ 1 4 2 からストライプ S 1 ~ S 1 6 の読出処理中に、タイムアウトが発生しなかったものとする。また、データ転送 D M A 1 5 0 は、ストライプ S 1 ~ S 1 6 の C R C 検査および A I D 検査においてエラーを検出しなかったものとする。

10

【 0 0 6 7 】

すべてのストライプ S 1 ~ 1 6 を読み出した後、データ転送 D M A 1 5 0 は、ストライプ S 1 ~ S 1 2 に含まれるデータ D 2 0 1 ~ D 2 1 2 をキャッシュメモリ 1 3 0 に書き込む。そして、データ転送 D M A 1 5 0 は、ストライプ S 1 3 に含まれるデータ D 2 1 3 から 1 [K B] のバウンダリ調整値を除去したデータ D 2 1 3 をキャッシュメモリ 1 3 0 に書き込む。なお、ストライプ S 1 4 に含まれるデータ D 2 1 4 は、すべてバウンダリ調整値であるため、データ転送 D M A 1 5 0 は、データ D 2 1 4 をキャッシュメモリ 1 3 0 に書き込まない。

20

【 0 0 6 8 】

以上のように、実施例 1 に係るストレージ装置 1 は、データ退避処理時に、退避対象データのパリティデータを生成して、C R C や A I D を付与した退避対象データおよびパリティデータをフラッシュメモリ 1 4 1 または 1 4 2 に書き込み、データ復旧処理において、フラッシュメモリ 1 4 1 または 1 4 2 から所定の時間内にデータ読出処理が完了しない場合に、データ読出処理を中止して付加データを設定し、パリティデータを用いて誤りデータを訂正したデータをキャッシュメモリ 1 3 0 に書き込むので、フラッシュメモリ 1 4 1 または 1 4 2 からデータの読出応答がなされない場合であっても、フラッシュメモリ 1 4 1 または 1 4 2 に退避させたデータを、キャッシュメモリ 1 3 0 に復旧させることができる。

30

【 0 0 6 9 】

また、実施例 1 に係るストレージ装置 1 は、退避対象単位データ番号が奇数である退避対象単位データを用いてパリティデータを生成するとともに、退避対象単位データ番号が偶数である退避対象単位データを用いてパリティデータを生成し、小エリア内の 1 つのダイには、退避対象単位データ番号が奇数である退避対象単位データと、退避対象単位データ番号が偶数である退避対象単位データとを書き込むように構成したので、1 個のダイが故障した場合であっても、かかる故障したダイに記憶されていたデータを訂正することができる。すなわち、1 個のダイが故障した場合であっても、フラッシュメモリ 1 4 1 または 1 4 2 に退避させたデータを、キャッシュメモリ 1 3 0 に復旧させることができる。

40

【 0 0 7 0 】

次に、実施例 1 に係るストレージ装置 1 が有する C M 1 0 0 の構成について説明する。図 4 は、実施例 1 に係るストレージ装置 1 が有する C M 1 0 0 の構成を示す図である。なお、同図では、データ退避処理およびデータ復旧処理に関連する構成のみを図示している。

【 0 0 7 1 】

同図に示すように、C M 1 0 0 は、電力供給部 1 2 と接続されており、キャッシュメモリ 1 3 0 と、フラッシュメモリ 1 4 1 および 1 4 2 と、データ転送 D M A 1 5 0 と、R o C 1 6 0 とを有する。

50

【 0 0 7 2 】

電力供給部 1 2 は、ストレージ装置 1 が異常終了した場合に、C M 1 0 0 に対して電力供給を行う装置であり、例えば、大容量のコンデンサやバッテリー等である。なお、同図では、C M 1 0 0 が外部に備えられた電力供給部 1 2 によって電力供給される例を示しているが、C M 1 0 0 は、電力供給部 1 2 を有してもよい。

【 0 0 7 3 】

キャッシュメモリ 1 3 0 は、上述したように、ホストコンピュータと H D D 1 1 との間で転送が行われるデータを一時的に記憶するメモリである。フラッシュメモリ 1 4 1 および 1 4 2 は、不揮発性メモリであり、ストレージ装置 1 の異常終了時にデータ転送 D M A 1 5 0 によって、キャッシュメモリ 1 3 0 に記憶されているデータが退避される。

10

【 0 0 7 4 】

データ転送 D M A 1 5 0 は、R o C 1 6 0 に指示に従って、キャッシュメモリ 1 3 0 と、フラッシュメモリ 1 4 1 および 1 4 2 との間におけるデータ送受信を行う D M A 回路である。具体的には、データ転送 D M A 1 5 0 は、ストレージ装置 1 が異常終了した場合に、キャッシュメモリ 1 3 0 に記憶されているデータを、フラッシュメモリ 1 4 1 および 1 4 2 に退避させるデータ退避処理を行う。また、データ転送 D M A 1 5 0 は、ストレージ装置 1 が異常終了後に電源投入された場合に、フラッシュメモリ 1 4 1 および 1 4 2 に退避したデータをキャッシュメモリ 1 3 0 に復旧するデータ復旧処理を行う。なお、データ転送 D M A 1 5 0 の構成、データ転送 D M A 1 5 0 によるデータ退避処理およびデータ復旧処理については、後に詳述する。

20

【 0 0 7 5 】

R o C 1 6 0 は、C M 1 0 0 を全体制御する制御装置であり、退避データ領域算出部 1 6 1 と、退避指示部 1 6 2 と、復旧指示部 1 6 3 とを有する。退避データ領域算出部 1 6 1 は、ストレージ装置 1 が異常終了した場合に、データ転送 D M A 1 5 0 にデータ退避処理を行うように指示する際に必要となる各種情報を算出する処理部である。

【 0 0 7 6 】

具体的には、退避データ領域算出部 1 6 1 は、退避対象データのサイズや、退避対象データのキャッシュメモリ 1 3 0 における記憶位置（アドレス）に基づいて、データ転送 D M A 1 5 0 が退避するデータのサイズや、データ転送 D M A 1 5 0 が退避するデータの記憶位置や、バウンダリ調整値のサイズ等を算出する。

30

【 0 0 7 7 】

退避指示部 1 6 2 は、データ転送 D M A 1 5 0 に対して、データ退避処理を行うように指示する処理部である。具体的には、退避指示部 1 6 2 は、ストレージ装置 1 が異常終了した場合に、データ退避処理を行うように指示する命令（以下、「データ退避ディスクリプタ」という）を生成して、データ転送 D M A 1 5 0 に送信する。

【 0 0 7 8 】

このデータ退避ディスクリプタには、退避データ領域算出部 1 6 1 によって算出された各種情報が含まれる。具体的には、データ退避ディスクリプタには、退避対象データのキャッシュメモリ 1 3 0 における先頭の記憶位置（先頭アドレス）や、退避対象データのサイズや、バウンダリ調整値のサイズや、A I D 等が含まれる。

40

【 0 0 7 9 】

復旧指示部 1 6 3 は、データ転送 D M A 1 5 0 に対して、データ復旧処理を行うように指示する処理部である。具体的には、復旧指示部 1 6 3 は、ストレージ装置 1 が異常終了後に電源投入された場合に、データ復旧処理を行うように指示する命令（以下、「データ復旧ディスクリプタ」という）を生成して、データ転送 D M A 1 5 0 に送信する。このデータ復旧ディスクリプタには、除去するバウンダリ調整値のサイズ等が含まれる。

【 0 0 8 0 】

また、復旧指示部 1 6 3 は、データ転送 D M A 1 5 0 によるデータ復旧処理が完了した場合に、フラッシュメモリ 1 4 1 および 1 4 2 に記憶されているデータを消去する命令（以下、「イレースディスクリプタ」という）を生成して、データ転送 D M A 1 5 0 に送信

50

する。

【0081】

次に、データ転送DMA150の構成について説明する。図5は、図4に示したデータ転送DMA150の構成を示す図である。同図に示すように、データ転送DMA150は、PCI(Peripheral Component Interconnect)バスI/F510と、メモリI/F521および522と、ディスクリプタ保持レジスタ530と、ライトDMA541および542と、リードDMA550と、イレースDMA561および562とを有する。

【0082】

PCIバスI/F510は、データ転送DMA150とキャッシュメモリ130との間のデータ送受信を行うインタフェースである。メモリI/F521は、データ転送DMA150とフラッシュメモリ141との間のデータ送受信を行うインタフェースであり、メモリI/F522は、データ転送DMA150とフラッシュメモリ142との間のデータ送受信を行うインタフェースである。

10

【0083】

ディスクリプタ保持レジスタ530は、RoC160から受信した各種ディスクリプタを記憶するレジスタである。具体的には、ディスクリプタ保持レジスタ530は、データ退避ディスクリプタや、データ復旧ディスクリプタや、イレースディスクリプタ等を記憶する。

【0084】

ライトDMA541および542は、データ退避処理を行うDMA回路である。具体的には、ライトDMA541は、ディスクリプタ保持レジスタ530からデータ退避ディスクリプタを受け取った場合に、PCIバスI/F510を介してキャッシュメモリ130に記憶されているデータを取得する。そして、ライトDMA541は、取得したデータを、メモリI/F521を介してフラッシュメモリ141に書き込む。同様に、ライトDMA542は、ディスクリプタ保持レジスタ530からデータ退避ディスクリプタを受け取った場合に、キャッシュメモリ130に記憶されているデータを取得し、取得したデータを、フラッシュメモリ142に書き込む。

20

【0085】

ライトDMA541および542は、キャッシュメモリ130からフラッシュメモリ141または142にデータを転送する処理を並行して行う。図3-1に示した例において、例えば、ライトDMA541がデータD10をフラッシュメモリ141へ書き込む処理を行い、ライトDMA542がデータD20をフラッシュメモリ142へ書き込む処理を行う。そして、このライトDMA541および542によるデータ退避処理が並行して行われる。

30

【0086】

このように、ストレージ装置1は、2系統のフラッシュメモリ141および142と、2系統のライトDMA541および542とを有して、ライトDMA541および542に対してデータ退避処理を並行して行わせる。これにより、ストレージ装置1は、データ退避処理を効率よく行うことが可能となり、電力供給部12から電力供給される短い時間内にデータ退避処理を完了させることを実現している。

40

【0087】

リードDMA550は、データ復旧処理を行うDMA回路である。具体的には、リードDMA550は、ディスクリプタ保持レジスタ530から、データ復旧ディスクリプタを受け取った場合に、フラッシュメモリ141に記憶されているデータを取得し、取得したデータを、キャッシュメモリ130に記憶させる。同様に、リードDMA550は、データ復旧ディスクリプタを受け取った場合に、フラッシュメモリ142に記憶されているデータを取得し、取得したデータを、キャッシュメモリ130に記憶させる。

【0088】

実施例1に係るストレージ装置1が、データ退避処理を2系統のライトDMA541および542に行わせるのに対して、データ復旧処理を1系統のリードDMA550に行わ

50

せる理由について説明する。これは、データ退避処理には、電力供給部 1 2 から電力供給される短い時間内に完了させなければならないという制約があるのに対して、データ復旧処理には、このような制約がないためである。

【 0 0 8 9 】

ただし、データ転送 DMA 1 5 0 が有するライト DMA やリード DMA の数は、図 5 に示した例に限られない。例えば、データ転送 DMA 1 5 0 は、例えば、3 個のライト DMA を有してもよいし、1 0 個のリード DMA を有してもよい。

【 0 0 9 0 】

イレース DMA 5 6 1 は、ディスクリプタ保持レジスタ 5 3 0 からイレースディスクリプタを受け取った場合に、フラッシュメモリ 1 4 1 に記憶されているデータを消去する DMA 回路である。同様に、イレース DMA 5 6 2 は、ディスクリプタ保持レジスタ 5 3 0 からイレースディスクリプタを受け取った場合に、フラッシュメモリ 1 4 2 に記憶されているデータを消去する DMA 回路である。

10

【 0 0 9 1 】

次に、ライト DMA 5 4 1 および 5 4 2 の構成について説明する。図 6 は、図 5 に示したライト DMA 5 4 1 および 5 4 2 の構成を示す図である。なお、ライト DMA 5 4 1 および 5 4 2 の構成は同一であるため、ここでは、ライト DMA 5 4 1 の構成についてのみ説明する。

【 0 0 9 2 】

同図に示すように、ライト DMA 5 4 1 は、データ読出部 5 4 1 a と、バウンダリ調整部 5 4 1 b と、パリティ生成部 5 4 1 c と、退避データバッファ 5 4 1 d と、パリティバッファ 5 4 1 e と、データ書込部 5 4 1 f とを有する。

20

【 0 0 9 3 】

データ読出部 5 4 1 a は、キャッシュメモリ 1 3 0 から所定のサイズのデータを読み出す処理部である。具体的には、データ読出部 5 4 1 a は、データ退避ディスクリプタに含まれている退避対象データの先頭アドレスから、退避対象データのサイズ分のデータを読み出す。

【 0 0 9 4 】

バウンダリ調整部 5 4 1 b は、データ読出部 5 4 1 a によって読み出された退避対象データにバウンダリ調整値を加える処理部である。具体的には、バウンダリ調整部 5 4 1 b は、退避対象データに、データ退避ディスクリプタに含まれているバウンダリ調整値のサイズ分のバウンダリ調整値を加える。

30

【 0 0 9 5 】

パリティ生成部 5 4 1 c は、データ読出部 5 4 1 a によって読み出された退避対象データを用いて、パリティデータを生成する処理部である。具体的には、パリティ生成部 5 4 1 c は、退避対象データを、バウンダリサイズ 2 8 [K B] ごとのバウンダリデータに区切って、さらに、バウンダリデータを、予めシステムで決められているサイズである 2 [K B] ごとに区切る。

【 0 0 9 6 】

そして、パリティ生成部 5 4 1 c は、退避対象単位データ番号が奇数である 7 個の退避対象単位データを用いてパリティデータを生成するとともに、退避対象単位データ番号が偶数である 7 個の退避対象単位データを用いてパリティデータを生成する。そして、データ読出部 5 4 1 a によって読み出された退避対象データを、退避データバッファ 5 4 1 d に記憶させるとともに、生成したパリティデータをパリティバッファ 5 4 1 e に記憶させる。

40

【 0 0 9 7 】

退避データバッファ 5 4 1 d は、データを一次記憶する記憶装置であり、上述したように、パリティ生成部 5 4 1 c によって退避対象データが記憶される。パリティバッファ 5 4 1 e は、データを一次記憶する記憶装置であり、上述したように、パリティ生成部 5 4 1 c によってパリティデータが記憶される。

50

【0098】

データ書込部541fは、退避データバッファ541dおよびパリティバッファ541eから、退避対象単位データおよびパリティデータを読み出してフラッシュメモリ141に書き込む処理部であり、CRC付与部541gとAID付与部541hとを有する。

【0099】

CRC付与部541gは、退避データバッファ541dから読み出した退避対象単位データごとにCRCを生成するとともに、パリティバッファ541eから読み出したパリティデータごとにCRCを生成する。そして、CRC付与部541gは、生成したCRCを、退避対象単位データおよびパリティデータに付与する。

【0100】

AID付与部541hは、退避データバッファ541dから読み出した退避対象単位データに、AIDを付与するとともに、パリティバッファ541eから読み出したパリティデータに、AIDを付与する。なお、AID付与部541hは、データ退避ディスクリプタからAIDを取得する。

【0101】

データ書込部541fは、このようにしてCRCおよびAIDが付与された退避対象単位データおよびパリティデータを、フラッシュメモリ141に書き込む。

【0102】

次に、リードDMA550の構成について説明する。図7は、図5に示したリードDMA550の構成を示す図である。同図に示すように、リードDMA550は、セクタ550aと、データ読出部550bと、タイムアウト検出部550cと、エラー検出部550dと、復旧データバッファ550eと、XOR演算部550fと、CRC検査部550gと、AID検査部550hと、訂正データバッファ550iと、ログ保持レジスタ550jと、セクタ550kと、バウンダリ調整部550lと、データ書込部550mとを有する。

【0103】

セクタ550aは、複数の入力信号の中から1つの入力信号のみを選択して出力するセクタ回路であり、フラッシュメモリ141に記憶されているストライプをデータ読出部550bに出力したり、付加データをデータ読出部550bに出力したりする。セクタ550aが、ストライプまたは付加データのいずれを出力するかは、後に説明する。

【0104】

データ読出部550bは、セクタ550aを介して、フラッシュメモリ141から、小エリアのストライプを読み出して、読み出したストライプを復旧データバッファ550eに記憶させる。このとき、データ読出部550bは、データ読出処理を開始した時刻をタイムアウト検出部550cに通知する。

【0105】

タイムアウト検出部550cは、データ読出部550bから受け付けたデータ読出処理の開始時刻に基づいて、データ読出部550bによるデータ読出処理が所定の時間内に完了するか否かを監視する処理部である。

【0106】

具体的には、タイムアウト検出部550cは、データ読出部550bによるデータ読出処理が、データ読出処理の開始時刻から所定の時間を経過しても完了しない場合に、データ読出部550bが読取処理中のストライプに対してタイムアウトが発生したことを検出する。そして、タイムアウト検出部550cは、タイムアウトを検出したストライプのストライプ番号を含むタイムアウト検出通知をセクタ550aに対して行う。

【0107】

ここで、タイムアウト検出通知を受け付けたセクタ550aの処理について説明する。タイムアウト検出通知を受け付けたセクタ550aは、タイムアウト検出通知に含まれるストライプ番号が示すストライプの読出処理を中止して、かかるストライプの読み出しできなかったサイズ分の付加データをデータ読出部550bに出力する。その後、セ

10

20

30

40

50

クタ550aは、次のストライプの読出処理を再開する。

【0108】

エラー検出部550dは、フラッシュメモリ141からECCチェックエラーが発生した旨の通知を受け付けた場合に、ECCエラー通知をセクタ550aに対して行う処理部である。

【0109】

ここで、ECCエラー通知を受け付けたセクタ550aの処理について説明する。ECCエラー通知を受け付けたセクタ550aは、読出処理中のストライプの読出処理を中止して、かかるストライプの読み出しできなかったサイズ分の付加データをデータ読出部550bに出力する。その後、セクタ550aは、次のストライプの読出処理を再開する。

10

【0110】

復旧データバッファ550eは、データを一次記憶する記憶装置であり、上述したように、データ読出部550bによって、ストライプまたは付加データが設定されているストライプが記憶される。

【0111】

XOR演算部550fは、データ読出部550bによって読み出されたストライプを用いて訂正データを生成する処理部である。具体的には、XOR演算部550fは、ストライプ番号が奇数である7個のストライプを用いて訂正データを生成するとともに、ストライプ番号が偶数である7個のストライプを用いて訂正データを生成する。そして、XOR演算部550fは、生成した訂正データを訂正データバッファ550iに記憶させる。

20

【0112】

CRC検査部550gは、データ読出部550bによって読み出されたストライプに誤りデータが存在するか否かを検査する処理部である。具体的には、CRC検査部550gは、データ読出部550bによって読み出されたストライプ内のCRCを用いて、誤りデータが存在するか否かを、ストライプごとに検査する。そして、誤りデータが存在する場合、すなわち、エラーを検出した場合、CRC検査部550gは、エラーを検出したストライプのストライプ番号をログ保持レジスタ550jに記憶させる。

【0113】

AID検査部550hは、データ読出部550bによって読み出されたストライプがすべて揃っているか否かを検査する処理部である。具体的には、AID検査部550hは、データ読出部550bによって読み出されたストライプ内のAIDを用いて、小エリア内のストライプがすべて揃っているか否かを検査する。そして、ストライプがすべて揃っていないことを検出した場合、すなわち、エラーを検出した場合、AID検査部550hは、欠如しているストライプのストライプ番号をログ保持レジスタ550jに記憶させる。

30

【0114】

訂正データバッファ550iは、データを一次記憶する記憶装置であり、上述したように、XOR演算部550fによって、訂正データが記憶される。ログ保持レジスタ550jは、データを一次記憶する記憶装置であり、上述したように、CRC検査部550gまたはAID検査部550hによって、エラーが検出されたストライプのストライプ番号が記憶される。

40

【0115】

セクタ550kは、セクタ回路であり、復旧データバッファ550eと、訂正データバッファ550iと、ログ保持レジスタ550jとに記憶されている各種データに基づいて、復旧データバッファ550eに記憶されているストライプを、データ書込部550mに出力したり、誤りデータを訂正したストライプをデータ書込部550mに出力したりする。

【0116】

具体的には、セクタ550kは、ログ保持レジスタ550jにストライプ番号が記憶されていないストライプについては、復旧データバッファ550eに記憶されているスト

50

ライブを、データ書込部 550 m に出力する。また、セクタ 550 k は、ログ保持レジスタ 550 j にストライプ番号が記憶されているストライプ（以下、「エラー検出ストライプ」という）については、復旧データバッファ 550 e および訂正データバッファ 550 i に記憶されている各種データを用いて、かかるエラー発生ストライプの誤りデータを訂正してデータ書込部 550 m に出力する。

【0117】

バウンダリ調整部 550 l は、セクタ 550 k から受け付けた小エリア単位のストライプからバウンダリ調整値を除去する処理部である。具体的には、バウンダリ調整部 550 l は、小エリア単位のストライプから、データ復旧ディスクリプタに含まれるバウンダリ調整値のサイズ分のバウンダリ調整値を除去する。

10

【0118】

データ書込部 550 m は、バウンダリ調整部 550 l によってバウンダリ調整値が除去された小エリア単位のストライプに含まれる退避対象単位データをキャッシュメモリ 130 に書き込む処理部である。なお、データ書込部 550 m は、パリティデータについてはキャッシュメモリ 130 に書き込まない。

【0119】

次に、実施例 1 に係るストレージ装置 1 によるデータ退避処理について説明する。図 8 は、実施例 1 に係るストレージ装置 1 によるデータ退避処理の流れを示すシーケンス図である。

【0120】

同図に示すように、R o C 160 の退避データ領域算出部 161 は、ストレージ装置 1 が異常終了したことを検出した場合（ステップ S 101）、データ退避ディスクリプタを生成するために必要となる各種情報を算出する。

20

【0121】

そして、退避指示部 162 は、退避データ領域算出部 161 によって算出された各種情報に基づいて、ライトDMA 541 にデータ退避処理を行わせるためのデータ退避ディスクリプタを生成して、ディスクリプタ保持レジスタ 530 に記憶させる（ステップ S 102）。

【0122】

例えば、図 3 - 1 に示した例のように、キャッシュメモリ 130 にデータ D 1 が記憶されている場合、退避指示部 162 は、ライトDMA 541 にデータ退避処理をさせるためのデータ退避ディスクリプタとして、データ D 10 の先頭アドレスと、データ D 10 のサイズ 28 [KB] と、バウンダリ調整値のサイズ 0 [KB] と、A I D 等を含むデータ退避ディスクリプタを生成する。

30

【0123】

同様に、退避指示部 162 は、ライトDMA 542 にデータ退避処理をさせるためのデータ退避ディスクリプタを生成して、ディスクリプタ保持レジスタ 530 に記憶させる（ステップ S 103）。

【0124】

上述例の場合、退避指示部 162 は、ライトDMA 542 にデータ退避処理をさせるためのデータ退避ディスクリプタとして、データ D 20 の先頭アドレスと、データ D 20 のサイズ 25 [KB] と、バウンダリ調整値のサイズ 3 [KB] と、A I D 等を含むデータ退避ディスクリプタを生成する。

40

【0125】

その後、退避指示部 162 は、ライトDMA 541 を起動させるとともに（ステップ S 104）、ライトDMA 542 を起動させて（ステップ S 105）、ライトDMA 541 およびライトDMA 542 に対して、データ退避処理を並行して行わせる。

【0126】

退避指示部 162 によって起動されたライトDMA 541 は、データ退避処理を行い（ステップ S 106）、データ退避処理が完了した場合に、データ退避処理が完了した旨を

50

R o C 1 6 0 に対して通知する（ステップ S 1 0 7）。同様に、ライト D M A 5 4 2 は、データ退避処理を行い（ステップ S 1 0 8）、データ退避処理が完了した場合に、データ退避処理が完了した旨を R o C 1 6 0 に対して通知する（ステップ S 1 0 9）。ライト D M A 5 4 1 および 5 4 2 によるデータ退避処理は、並行して行われる。

【 0 1 2 7 】

なお、上述したデータ退避処理の流れでは、ストレージ装置 1 が異常終了した後に、退避指示部 1 6 2 がデータ退避ディスクリプタを生成する例を説明したが、退避指示部 1 6 2 は、ストレージ装置 1 が異常終了する前に、データ退避ディスクリプタを生成してもよい。例えば、ストレージ装置 1 の起動中に、退避対象データのサイズ等を予め取得しておけば、ストレージ装置 1 の異常終了前にデータ退避ディスクリプタを生成することができる。

10

【 0 1 2 8 】

次に、ライト D M A 5 4 1 および 5 4 2 によるデータ退避処理について説明する。図 9 は、ライト D M A 5 4 1 および 5 4 2 によるデータ退避処理手順を示すフローチャートである。なお、ライト D M A 5 4 1 および 5 4 2 によるデータ退避処理手順は同一であるため、ここでは、ライト D M A 5 4 1 によるデータ退避処理手順についてのみ説明する。

【 0 1 2 9 】

同図に示すように、R o C 1 6 0 によって起動されたライト D M A 5 4 1 は、ディスクリプタ保持レジスタ 5 3 0 からデータ退避ディスクリプタを読み出す（ステップ S 2 0 1）。そして、ライト D M A 5 4 1 のデータ読出部 5 4 1 a は、キャッシュメモリ 1 3 0 から、所定のサイズのデータを読み出す（ステップ S 2 0 2）。このとき、データ読出部 5 4 1 a は、データ退避ディスクリプタに含まれている退避対象データの先頭アドレスから、退避対象データのサイズ分のデータを読み出す。

20

【 0 1 3 0 】

続いて、バウンダリ調整部 5 4 1 b は、データ読出部 5 4 1 a によって読み出された退避対象データに、データ退避ディスクリプタに含まれているバウンダリ調整値のサイズ分のバウンダリ調整値を加える（ステップ S 2 0 3）。

【 0 1 3 1 】

続いて、パリティ生成部 5 4 1 c は、退避対象データを、バウンダリサイズごとのバウンダリデータに区切って、さらに、バウンダリデータを、予めシステムで決められているサイズである 2 [K B] ごとに区切る。そして、パリティ生成部 5 4 1 c は、退避対象単位データ番号が奇数である 7 個の退避対象単位データを用いてパリティデータを生成するとともに、退避対象単位データ番号が偶数である 7 個の退避対象単位データを用いてパリティデータを生成する（ステップ S 2 0 4）。

30

【 0 1 3 2 】

続いて、C R C 付与部 5 4 1 g は、退避対象単位データごとに C R C を生成するとともに、パリティデータごとに C R C を生成して、生成した C R C を、退避対象単位データおよびパリティデータに付与する（ステップ S 2 0 5）。続いて、A I D 付与部 5 4 1 h は、退避対象単位データおよびパリティデータに、A I D を付与する（ステップ S 2 0 6）。そして、データ書込部 5 4 1 f は、C R C および A I D が付与された退避対象単位データおよびパリティデータを、フラッシュメモリ 1 4 1 に書き込む（ステップ S 2 0 7）。

40

【 0 1 3 3 】

次に、実施例 1 に係るストレージ装置 1 によるデータ復旧処理について説明する。図 1 0 は、実施例 1 に係るストレージ装置 1 によるデータ復旧処理の流れを示すシーケンス図である。

【 0 1 3 4 】

同図に示すように、R o C 1 6 0 の復旧指示部 1 6 3 は、異常終了後にストレージ装置 1 が起動したことを検出した場合（ステップ S 3 0 1）、リード D M A 5 5 0 に対して、フラッシュメモリ 1 4 1 に記憶されているデータをキャッシュメモリ 1 3 0 へ復旧する旨のデータ復旧ディスクリプタを生成して、ディスクリプタ保持レジスタ 5 3 0 に記憶させ

50

る（ステップS302）。その後、復旧指示部163は、リードDMA550を起動させる（ステップS303）。

【0135】

復旧指示部163によって起動されたリードDMA550は、フラッシュメモリ141に記憶されているデータをキャッシュメモリ130へ復旧させるデータ復旧処理を行う（ステップS304）。かかるデータ復旧処理が完了した場合に、リードDMA550は、データ復旧処理が完了した旨をROC160に対して通知する（ステップS305）。

【0136】

続いて、復旧指示部163は、リードDMA550に対して、フラッシュメモリ142に記憶されているデータをキャッシュメモリ130へ復旧する旨のデータ復旧ディスクリプタを生成して、ディスクリプタ保持レジスタ530に記憶させる（ステップS306）。その後、復旧指示部163は、リードDMA550を起動させる（ステップS307）。

10

【0137】

復旧指示部163によって起動されたリードDMA550は、フラッシュメモリ142に記憶されているデータをキャッシュメモリ130へ復旧させるデータ復旧処理を行う（ステップS308）。かかるデータ復旧処理が完了した場合に、リードDMA550は、データ復旧処理が完了した旨をROC160に対して通知する（ステップS309）。

【0138】

続いて、ROC160は、キャッシュメモリ130に復旧されたデータを、HDD11に記憶させるリストア処理（フラッシュバック）を行う（ステップS310）。これにより、HDD11に書き込みされていないキャッシュメモリ130上のデータはなくなる。

20

【0139】

続いて、復旧指示部163は、イレースディスクリプタを生成して、データ転送DMA150に記憶させる（ステップS311）。そして、復旧指示部163は、イレースDMA561および562を起動させて（ステップS312）、イレースDMA561および562に対して、フラッシュメモリ141および142に記憶されているデータを消去させる。

【0140】

復旧指示部163によって起動されたイレースDMA561および562は、ディスクリプタ保持レジスタ530から、イレースディスクリプタを読み出して（ステップS313）、イレース処理を行う（ステップS314）。かかるイレース処理が完了した場合に、イレースDMA561および562は、イレース処理が完了した旨をROC160に対して通知する（ステップS315）。そして、イレース処理が完了した旨の通知を受け付けたROC160は、ストレージ装置1を起動させる（ステップS316）。

30

【0141】

次に、リードDMA550によるデータ復旧処理について説明する。図11は、リードDMA550によるデータ復旧処理手順を示すフローチャートである。同図に示すように、ROC160によって起動されたリードDMA550は、ディスクリプタ保持レジスタ530からデータ復旧ディスクリプタを読み出す（ステップS401）。

40

【0142】

そして、リードDMA550のデータ読出部550bは、セクタ550aを介して、フラッシュメモリ141から、小エリアのストライプを読み出す（ステップS402）。このデータ読出部550bによるデータ読出処理が所定の時間内に完了しない場合（ステップS403否定）、タイムアウト検出部550cは、データ読出部550bによって読取処理されているストライプに対して、タイムアウトが発生したことを検出する。そして、タイムアウト検出部550cは、タイムアウト検出通知をセクタ550aに対して行う。

【0143】

タイムアウト検出通知を受け付けたセクタ550aは、タイムアウト検出通知に含ま

50

れるストライプ番号が示すストライプの読出処理を中止して、かかるストライプの読み出しできなかったサイズ分の付加データをデータ読出部550bに出力する(ステップS404)。このようにして、データ読出部550bは、フラッシュメモリ141から小エリアのストライプを読み出す。

【0144】

続いて、XOR演算部550fは、データ読出部550bによって読み出されたストライプを用いて訂正データを生成する(ステップS405)。続いて、CRC検査部550gは、データ読出部550bによって読み出されたストライプに誤りデータが存在するかどうかを検査し、AID検査部550hは、データ読出部550bによって読み出されたストライプがすべて揃っているかどうかを検査する(ステップS406)。

10

【0145】

CRC検査部550g、または、AID検査部によってエラーが検出された場合(ステップS407肯定)、セクタ550kは、データ読出部550bによって読み出されたストライプ、および、XOR演算部550fによって生成された訂正データに基づいて、エラーが検出されたストライプの誤りデータを訂正する(ステップS408)。

【0146】

続いて、バウンダリ調整部550lは、データ読出部550bによって読み出された小エリアのストライプから、データ復旧ディスクリプタに含まれるバウンダリ調整値のサイズ分のバウンダリ調整値を除去する(ステップS409)。

【0147】

そして、データ書込部550mは、バウンダリ調整部550lによってバウンダリ調整値が除去された小エリアのストライプに含まれる退避対象単位データを、キャッシュメモリ130に書き込む(ステップS410)。なお、このとき、データ書込部550mは、パリティデータについてはキャッシュメモリ130に書き込まない。

20

【0148】

リードDMA550は、フラッシュメモリ141に記憶されている全てのデータを読み出していない場合(ステップS411否定)、フラッシュメモリ141から全てのデータを読み出すまで、上述した処理手順(ステップS401~S410)を繰り返し行う。

【0149】

上述してきたように、実施例1に係るストレージ装置1は、データ退避処理時に、退避対象データのパリティデータを生成して、CRCやAIDを付与した退避対象データおよびパリティデータをフラッシュメモリ141または142に書き込み、データ復旧処理において、フラッシュメモリ141または142から所定の時間内にデータ読出処理が完了しない場合に、データ読出処理を中止して付加データを設定し、パリティデータを用いて誤りデータを訂正したデータをキャッシュメモリ130に書き込むように構成したので、フラッシュメモリ141または142からデータの読出応答がなされない場合であっても、フラッシュメモリ141または142に退避させたデータを、キャッシュメモリ130に復旧させることができる。

30

【0150】

なお、上記実施例1では、データ退避処理時に、ライトDMA541および542が、7個の退避対象単位データに対して、1個のパリティデータを生成する例を示したが、退避対象単位データの数と、パリティデータの数の比率はこれに限らない。例えば、ライトDMA541および542は、2個退避対象単位データに対して、1個のパリティデータを生成してもよいし、10個退避対象単位データに対して、1個のパリティデータを生成してもよい。

40

【0151】

また、上記実施例1では、ライトDMA541および542が、退避対象単位データ番号が奇数である退避対象単位データを用いてパリティデータを生成するとともに、退避対象単位データ番号が偶数である退避対象単位データを用いてパリティデータを生成する例を示したが、ライトDMA541および542は、退避対象単位データ番号に関わらず、

50

所定の数の退避対象単位データを用いてパリティデータを構成してもよい。

【実施例 2】

【0152】

ところで、上記実施例 1 では、データサイズがバウンダリサイズに満たないバウンダリデータに、バウンダリ調整値を加える処理であるバウンダリ処理を行う例を示したが、バウンダリ処理を行わないようにしてもよい。そこで、実施例 2 では、キャッシュメモリ 130 から読み出したデータをバウンダリサイズごとに区切れない場合であっても、バウンダリ処理を行わない例について説明する。

【0153】

まず、実施例 2 に係るストレージ装置 2 によるデータ退避処理の概要を説明する。なお、実施例 2 に係るストレージ装置 2 の概略構成は、図 1 に示したストレージ装置 1 の概略構成と同様であるため、その説明を省略する。

10

【0154】

図 12 - 1 は、実施例 2 に係るストレージ装置 2 によるデータ退避処理の概要を説明するための図である。同図に示すように、キャッシュメモリ 130 には、210 [KB] のデータ D3 が記憶されているものとする。この状態で、ストレージ装置 2 が異常終了した場合、データ D3 は、バウンダリサイズ 28 [KB] ごとにデータ D31 ~ D38 に区切られる。

【0155】

ここで、実施例 2 に係るストレージ装置 2 の RoC (RoC 260 とする) は、ライト DMA 541 に対して、キャッシュメモリ 130 に記憶されているデータ D3 のうち、データ D4 (データ D31 ~ D34) について、データ退避処理を行うように指示する。

20

【0156】

さらに、RoC 260 は、ライト DMA 542 に対して、キャッシュメモリ 130 に記憶されているデータ D3 のうち、データ D5 (データ D34 - 2 と、データ D35 ~ D38) について、データ退避処理を行うように指示する。

【0157】

つまり、RoC 260 は、ライト DMA 541 に対して、データ D3 のキャッシュメモリ 130 における先頭の記憶位置である先頭アドレスから 4 個分のバウンダリデータ (同図に示したデータ D31、D32、D33、D34) について、データ退避処理を行うように指示する。また、RoC 260 は、ライト DMA 542 に対して、データ D3 のキャッシュメモリ 130 における最後の記憶位置である最終アドレスから、先頭アドレス方向に 4 個分のバウンダリデータ (同図に示したデータ D54、D53、D52、D51) について、データ退避処理を行うように指示する。

30

【0158】

このような指示を受け付けたライト DMA 541 は、データ D4 を、バウンダリサイズ 28 [KB] ごとにデータ D31 ~ D34 に区切ることができる。また、ライト DMA 542 は、データ D5 を、バウンダリサイズ 28 [KB] ごとにデータ D51 ~ D54 に区切ることができる。すなわち、実施例 2 に係るストレージ装置 2 は、バウンダリ処理を行うことなく、データ退避処理を行うことができる。

40

【0159】

次に、実施例 2 に係るストレージ装置 2 によるデータ復旧処理の概要について説明する。図 12 - 2 は、実施例 2 に係るストレージ装置 2 によるデータ復旧処理の概要を説明するための図である。図 12 - 2 に示したフラッシュメモリ 141 および 142 には、図 12 - 1 に示したライト DMA 541 およびライト DMA 542 によって退避されたデータが記憶されているものとする。

【0160】

ストレージ装置 2 が異常終了後に起動した場合、リード DMA 550 は、フラッシュメモリ 141 からデータ D31 を読み出しながら各種検査 (CRC 検査等) を行った後、キャッシュメモリ 130 に書き込む。また、リード DMA 550 は、同様の処理を、データ

50

D 3 2 ~ D 3 4 についても行い、キャッシュメモリ 1 3 0 にデータ D 4 を復旧させる。

【 0 1 6 1 】

その後、リード DMA 5 5 0 は、フラッシュメモリ 1 4 2 からデータ D 5 1 を読み出しながら各種検査 (CRC 検査等) を行った後、キャッシュメモリ 1 3 0 に書き込む。このとき、キャッシュメモリ 1 3 0 には、フラッシュメモリ 1 4 1 から読み出したデータ D 3 4 - 2 が書き込み済みであるので、リード DMA 5 5 0 は、データ D 5 1 を書き込む場合、データ D 3 4 - 2 の部分については上書きすることとなる。

【 0 1 6 2 】

そして、リード DMA 5 5 0 は、フラッシュメモリ 1 4 2 からデータ D 5 2 ~ D 5 4 を順に読み出して、キャッシュメモリ 1 3 0 にデータ D 5 を復旧させる。リード DMA 5 5 0 が、フラッシュメモリ 1 4 1 および 1 4 2 に記憶されているデータをすべてキャッシュメモリ 1 3 0 に書き込むと、キャッシュメモリ 1 3 0 にデータ D 3 を復旧することができる。

10

【 0 1 6 3 】

このように、実施例 2 に係るストレージ装置 2 は、2 系統のライト DMA 5 4 1 および 5 4 2 のうち、一方のライト DMA 5 4 1 に対して、退避対象データのキャッシュメモリ 1 3 0 における先頭の記憶位置である先頭アドレスから所定の数のバウンダリデータについて、データ退避処理を行わせるとともに、他方のライト DMA 5 4 2 に対して、退避対象データのキャッシュメモリ 1 3 0 における最後の記憶位置である最終アドレスから、先頭アドレス方向に所定の数のバウンダリデータについて、データ退避処理を行わせるので、DMA 5 4 1 および 5 4 2 が、フラッシュメモリ 1 4 1 または 1 4 2 に退避するデータを、バウンダリデータで割り切れることができ、その結果、バウンダリ処理を行うことなく、データ退避処理を行うことができる。

20

【 0 1 6 4 】

次に、実施例 2 に係るストレージ装置 2 が有する CM 2 0 0 の構成について説明する。図 1 3 は、実施例 2 に係るストレージ装置 2 が有する CM 2 0 0 の構成を示す図である。ここでは、図 4 に示した構成部位と同様の機能を有する部位には同一符号を付すこととして、その詳細な説明を省略する。

【 0 1 6 5 】

同図に示した退避データ領域算出部 2 6 1 は、ストレージ装置 2 が異常終了した場合に、キャッシュメモリ 1 3 0 に記憶されている退避対象データのサイズ等に基づいて、まず、ライト DMA 5 4 1 が退避するデータの開始アドレス BA 1 およびサイズ DS 1 を算出し、次に、ライト DMA 5 4 2 が退避するデータの開始アドレス BA 2 およびサイズ DS 2 を算出する。

30

【 0 1 6 6 】

以下に、図 1 2 - 1 に示した例を用いつつ、退避データ領域算出部 2 6 1 による退避データ領域算出処理について、より具体的に説明する。まず、退避データ領域算出部 2 6 1 は、ストレージ装置 2 が異常終了した場合に、退避対象データのキャッシュメモリ 1 3 0 における記憶位置を示す先頭アドレスおよび最終アドレスを取得する。そして、退避データ領域算出部 2 6 1 は、取得した先頭アドレスを、ライト DMA 5 4 1 が退避するデータの開始アドレス BA 1 に決定する。

40

【 0 1 6 7 】

続いて、退避データ領域算出部 2 6 1 は、退避対象データのサイズを 2 で除算したサイズ「X」を求める。図 1 2 - 1 に示した例では、退避対象データのサイズが 2 1 0 [KB] であるので、退避データ領域算出部 2 6 1 は、サイズ X として、2 1 0 [KB] を 2 で除算した 1 0 5 [KB] を求める。

【 0 1 6 8 】

続いて、退避データ領域算出部 2 6 1 は、式 (1) 「 2 8 [KB] × N > サイズ X [KB] 」を満たす整数 N の最小値に 2 8 [KB] を乗算した値を、ライト DMA 5 4 1 が退避するデータのサイズ DS 1 に決定する。上記例では、サイズ X が 1 0 5 [KB]

50

]であり、式(1)を満たすNの最小値は「4」となるので、退避データ領域算出部261は、4に28[KB]を乗算した値である112[KB]を、ライトDMA541が退避するデータのサイズDS1に決定する。

【0169】

続いて、退避データ領域算出部261は、退避対象データの先頭アドレスに、ライトDMA541が退避するデータサイズDS1を加算した値「Y」を求める。このYは、ライトDMA541が退避するデータの最終アドレスに該当する。そして、退避データ領域算出部261は、退避対象データの最終アドレスからYを減算した値「Z」を求める。このZは、退避対象データのサイズから、ライトDMA541が退避するデータのサイズDS1を減算したサイズに該当する。

10

【0170】

すなわち、上記例では、退避対象データのサイズが210[KB]であり、サイズDS1が112[KB]であるので、退避データ領域算出部261は、サイズZとして、210[KB]から112[KB]を減算した98[KB]を求める。

【0171】

続いて、退避データ領域算出部261は、式(2)「 $28[\text{KB}] \times N > \text{サイズ}Z[\text{KB}]$ 」を満たす整数Nの最小値に28[KB]を乗算した値を、ライトDMA542が退避するデータのサイズDS2に決定する。上記例では、サイズZが98[KB]であり、式(2)を満たすNの最小値は「4」となるので、退避データ領域算出部261は、4に28[KB]を乗算した値である112[KB]を、ライトDMA542が退避するデータのサイズDS2に決定する。

20

【0172】

続いて、退避データ領域算出部261は、退避対象データの最終アドレスから、ライトDMA542が退避するデータのサイズDS2を減算した値を、ライトDMA542が退避するデータの先頭アドレスBA2に決定する。

【0173】

このようにして、退避データ領域算出部261は、ライトDMA541が退避するデータの開始アドレスBA1およびサイズDS1と、ライトDMA542が退避するデータの開始アドレスBA2およびサイズDS2を算出する。

【0174】

図13に示した退避指示部262は、データ退避ディスクリプタを生成して、データ転送DMA150に送信することで、ライトDMA541および542に対して、データ退避処理を行うように指示する。

30

【0175】

具体的には、退避指示部262は、上述した開始アドレスBA1およびサイズDS1を含むデータ退避ディスクリプタを生成して、かかるデータ退避ディスクリプタに基づいてデータ退避処理を行うようにライトDMA541に対して指示をする。また、退避指示部262は、上述した開始アドレスBA2およびサイズDS2を含むデータ退避ディスクリプタを生成して、かかるデータ退避ディスクリプタに基づいてデータ退避処理を行うようにライトDMA542に対して指示をする。なお、退避指示部262が生成するデータ退避ディスクリプタには、バウンダリ調整値のサイズを含まない。

40

【0176】

次に、図13に示した退避データ領域算出部261による退避データ領域算出処理について説明する。図14は、図13に示した退避データ領域算出部261による退避データ領域算出処理手順を示すフローチャートである。

【0177】

図14に示すように、退避データ領域算出部261は、退避対象データのキャッシュメモリ130における記憶位置を示す先頭アドレスおよび最終アドレスを取得する(ステップS501)。

【0178】

50

そして、退避データ領域算出部 261 は、取得した先頭アドレスを、ライトDMA 541 が退避するデータの開始アドレスに決定する（ステップ S502）。続いて、退避データ領域算出部 261 は、退避対象データのサイズを 2 で除算したサイズ X を求める（ステップ S503）。

【0179】

続いて、退避データ領域算出部 261 は、式(1)「 $28[\text{KB}] \times N > \text{サイズ} X [\text{KB}]$ 」を満たす整数 N の最小値を求める（ステップ S504）。以下では、ここで求めた最小値 N を「第一の基準値」と呼ぶこととする。そして、退避データ領域算出部 261 は、第一の基準値に $28[\text{KB}]$ を乗算した値を、ライトDMA 541 が退避するデータのサイズに決定する（ステップ S505）。

10

【0180】

続いて、退避データ領域算出部 261 は、退避対象データのサイズから、ライトDMA 541 が退避するデータのサイズを減算したサイズ Z を求める（ステップ S506）。続いて、退避データ領域算出部 261 は、式(2)「 $28[\text{KB}] \times N > \text{サイズ} Z [\text{KB}]$ 」を満たす整数 N の最小値を求める（ステップ S507）。以下では、ここで求めた最小値 N を「第二の基準値」と呼ぶこととする。そして、退避データ領域算出部 261 は、第二の基準値に $28[\text{KB}]$ を乗算した値を、ライトDMA 542 が退避するデータのサイズに決定する（ステップ S508）。

【0181】

続いて、退避データ領域算出部 261 は、退避対象データの最終アドレスから、ライトDMA 542 が退避するデータのサイズを減算した値を求め（ステップ S509）、求めた値を、ライトDMA 542 が退避するデータの先頭アドレスに決定する（ステップ S510）。

20

【0182】

上述してきたように、実施例 2 に係るストレージ装置 2 は、退避データ領域算出部 261 が、ライトDMA 541 および 542 が退避するデータの開始アドレスおよびサイズを求めるように構成したので、DMA 541 および 542 は、フラッシュメモリ 141 または 142 に退避するデータを、バウンダリデータで割り切れることができ、その結果、フラッシュメモリ 141 または 142 からデータの読出応答がなされない場合であっても、バウンダリ処理を行うことなく、フラッシュメモリ 141 または 142 に退避させたデータを、キャッシュメモリ 130 に復旧させることができる。

30

【0183】

また、実施例 2 に係るストレージ装置 2 では、バウンダリ処理を行う必要がないので、退避データ領域算出部 261 がバウンダリ調整値のサイズを算出する必要がなくなる。その結果、退避データ領域算出部 261 にかかる処理負荷を軽減することができ、データ退避処理の高速化を図ることができる。すなわち、異常終了時に、電力供給部 12 から電力供給される短い時間内に、大容量のデータについて、データ退避処理を行うことができる。

【0184】

また、実施例 2 に係るストレージ装置 2 では、バウンダリ処理を行う必要がないので、ライトDMA 541 および 542 と、リードDMA 550 から、バウンダリ処理を行うための回路を削除することができ、ライトDMA 541 および 542 と、リードDMA 550 の規模を小さくすることができる。例えば、実施例 2 に係るストレージ装置 2 では、図 6 に示したライトDMA 541 からバウンダリ調整部 541b を削除することができ、図 7 に示したリードDMA 550 からバウンダリ調整部 5501 を削除することができる。その結果、ライトDMA 541 および 542 と、リードDMA 550 にかかるコストを軽減することができる。

40

【0185】

なお、上記実施例 2 では、ストレージ装置 2 が 2 系統のライトDMA 541 および 542 を有することを前提として、退避データ領域算出部 261 が、ライトDMA 541 およ

50

び542が退避するデータの開始アドレスおよびサイズを算出する例を示したが、3系統以上のライトDMAを有するストレージ装置においても、バウンダリ処理を行わないように、退避データ領域算出部261が、各ライトDMAが退避するデータの開始アドレスおよびサイズを算出してもよい。例えば、ストレージ装置が4個のライトDMAを有する場合、退避データ領域算出部261は、4個目のライトDMAに対して、退避対象データの最終アドレスから、先頭アドレス方向に所定の数のバウンダリデータについて、データ退避処理を行わせるように、4個目のライトDMAが退避するデータの開始アドレスおよびサイズを算出する。

【0186】

また、上記実施例1および2では、バウンダリサイズを28[KB]とし、バウンダリデータを区切るサイズを2[KB]とし、エリア内のpageの数を1024個とし、小エリア内のpageの数を16個とする例を示したが、これらの規定値は、上記例に限られない。例えば、バウンダリサイズを56[KB]とし、バウンダリデータを区切るサイズを4[KB]とし、エリア内のpageの数を2048個とし、小エリア内のpageの数を32個としてもよい。

10

【0187】

また、上記実施例1および2では、異常終了時にデータを退避する記憶領域として、OneNAND型のフラッシュメモリを用いる例を示したが、異常終了時にデータを退避する記憶領域は、OneNAND型のフラッシュメモリ以外の不揮発性メモリであってもよい。

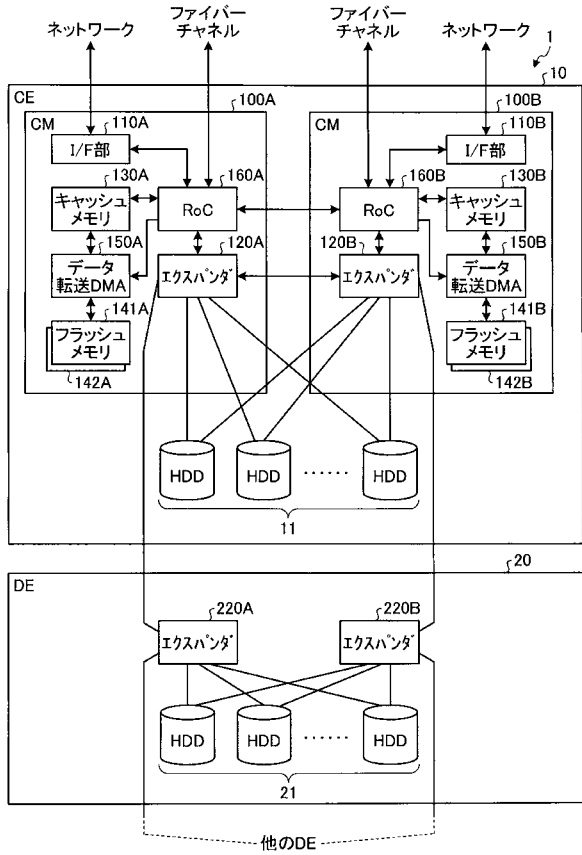
20

【0188】

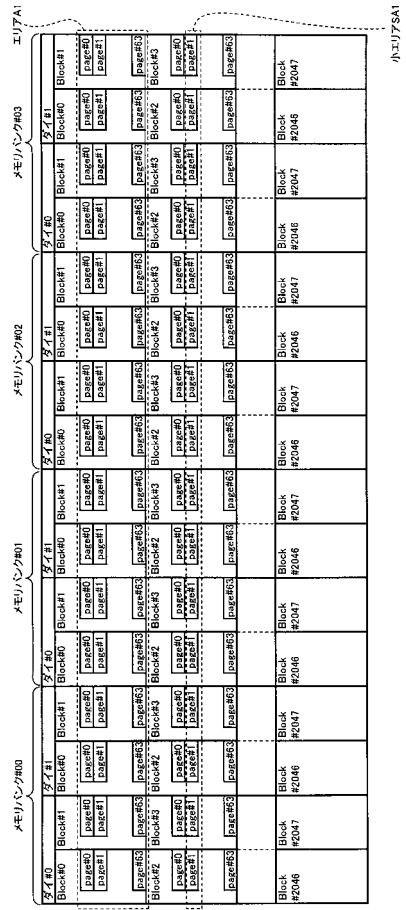
また、上記文書中や図面中で示した処理手順、制御手順、具体的名称、各種のデータやパラメータを含む情報については、特記する場合を除いて任意に変更することができる。また、図示した各装置の各構成要素は機能概念的なものであり、必ずしも物理的に図示の如く構成されていることを要しない。すなわち、各装置の分散、統合の具体的形態は図示のものに限られず、その全部または一部を、各種の負荷や使用状況などに応じて、任意の単位で機能的または物理的に分散、統合して構成することができる。さらに、各装置にて行なわれる各処理機能は、その全部または任意の一部が、CPU(Central Processing Unit)および当該CPUにて解析実行されるプログラムにて実現され、あるいは、ワイヤードロジックによるハードウェアとして実現され得る。

30

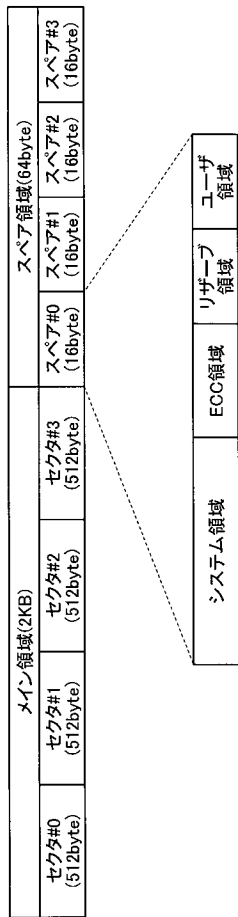
【 図 1 】



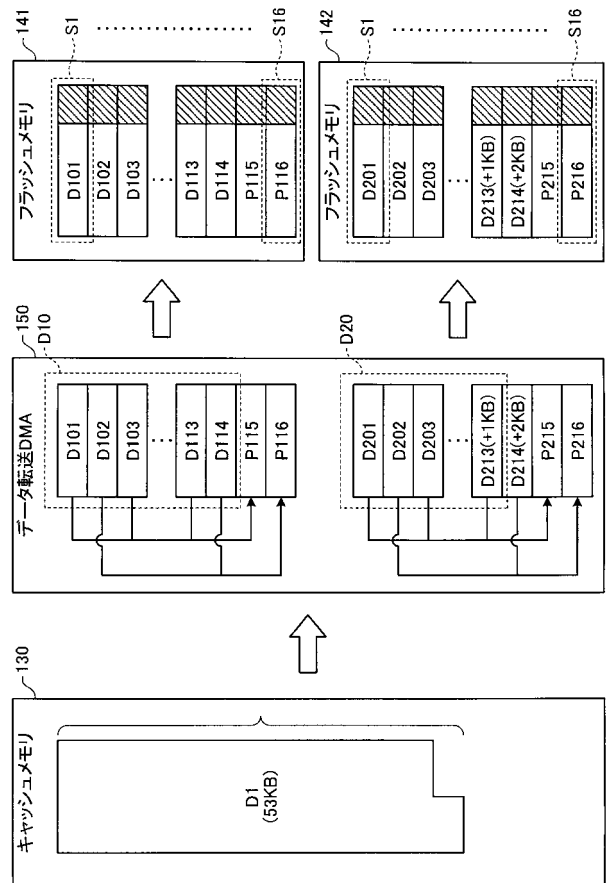
【 図 2 - 1 】



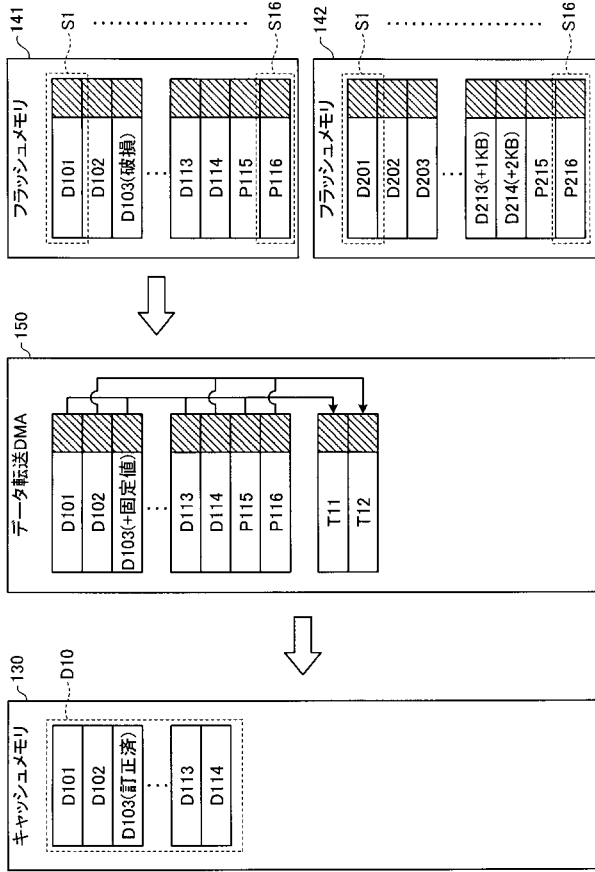
【 図 2 - 2 】



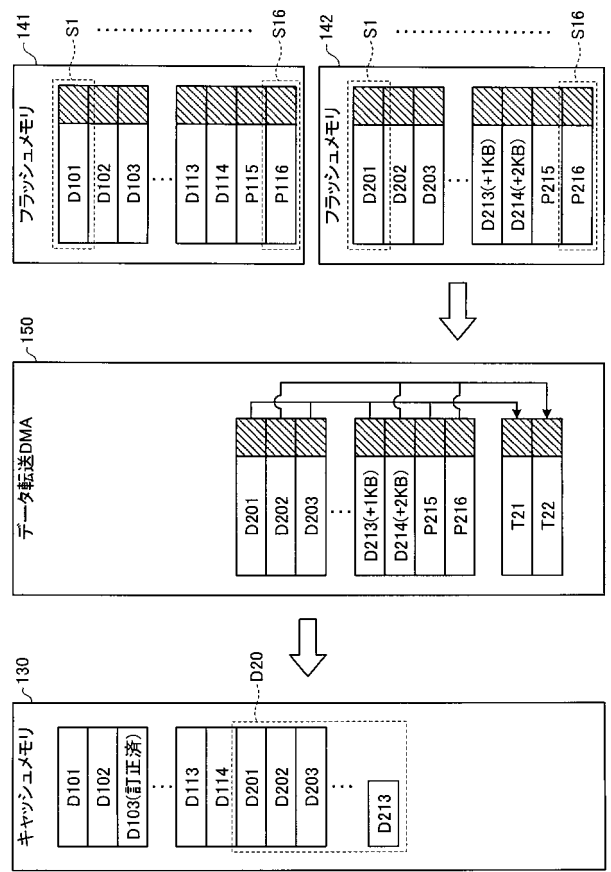
【 図 3 - 1 】



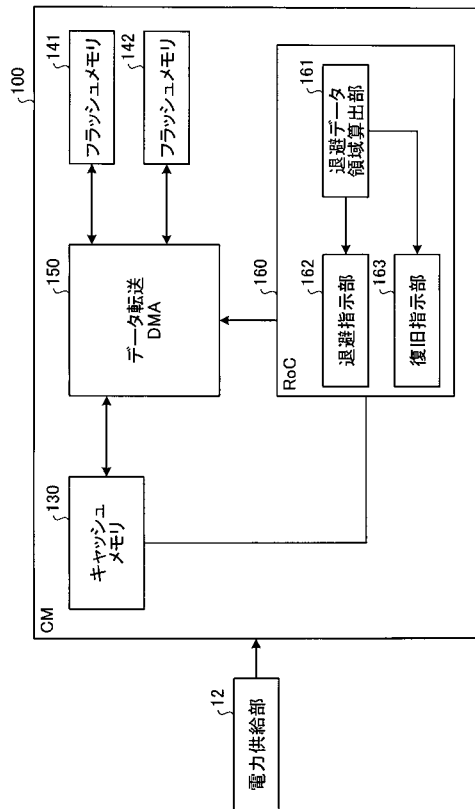
【 図 3 - 2 】



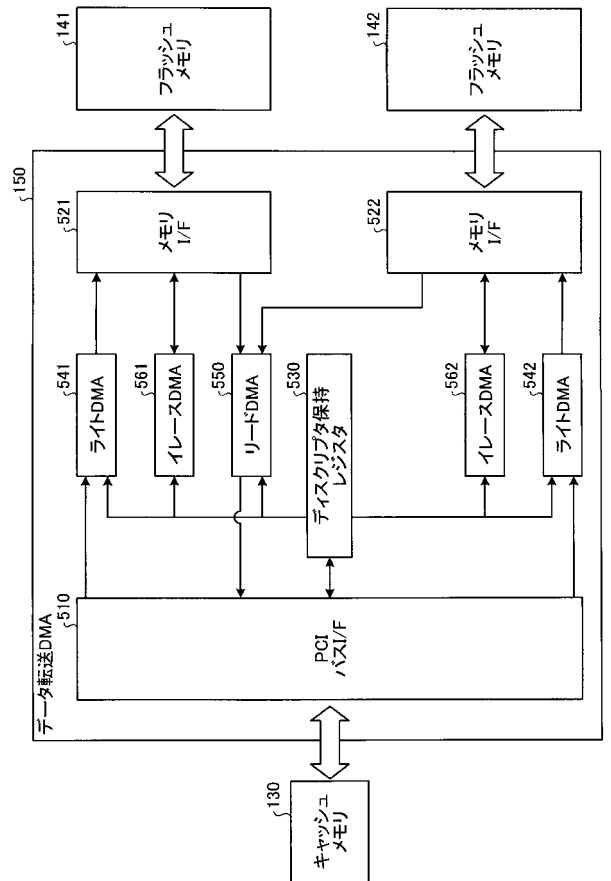
【 図 3 - 3 】



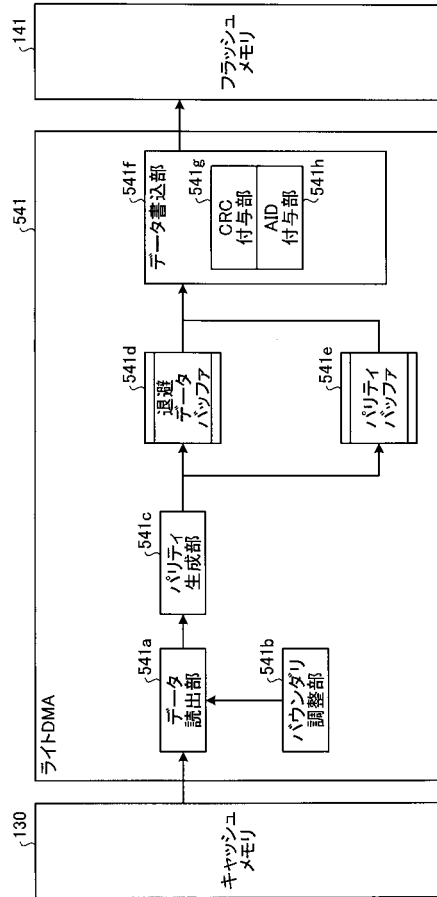
【 図 4 】



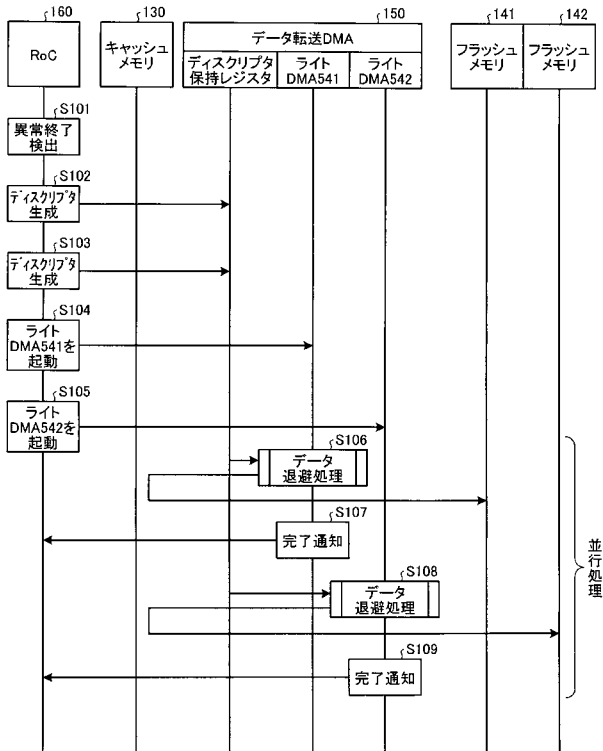
【 図 5 】



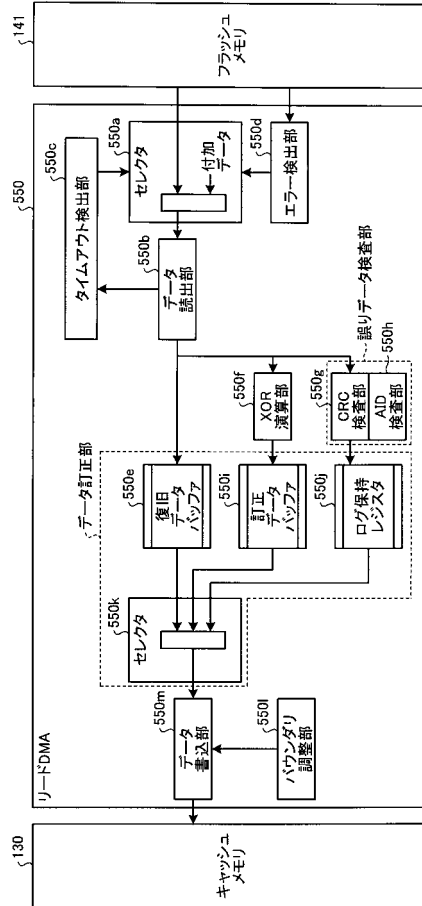
【 図 6 】



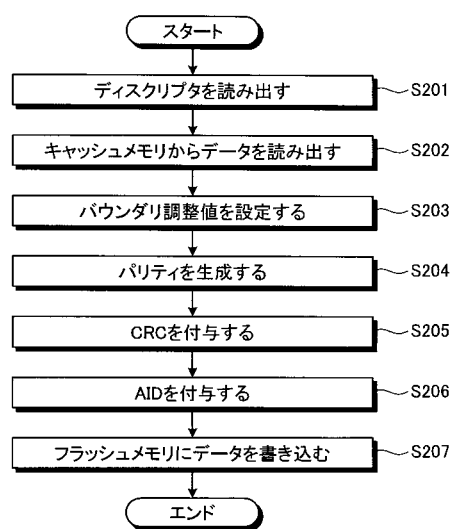
【 図 8 】



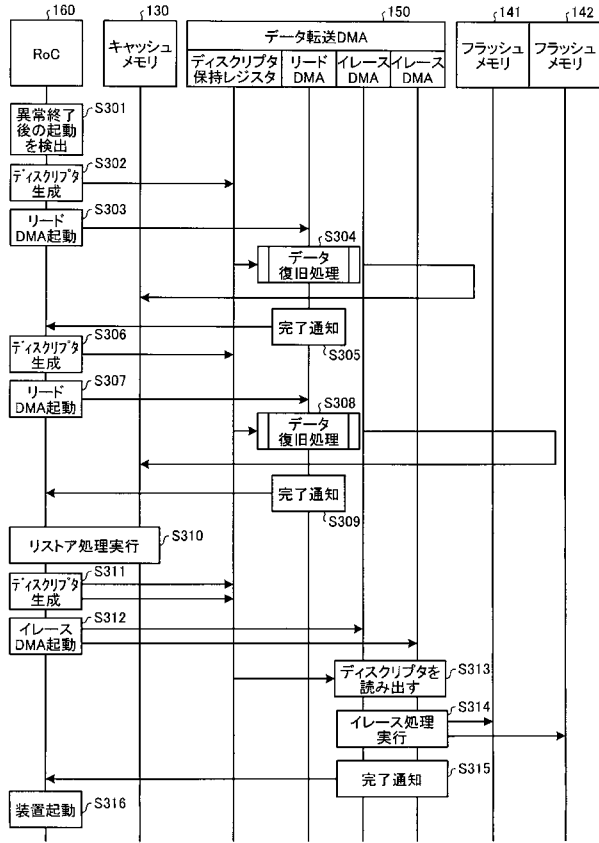
【 図 7 】



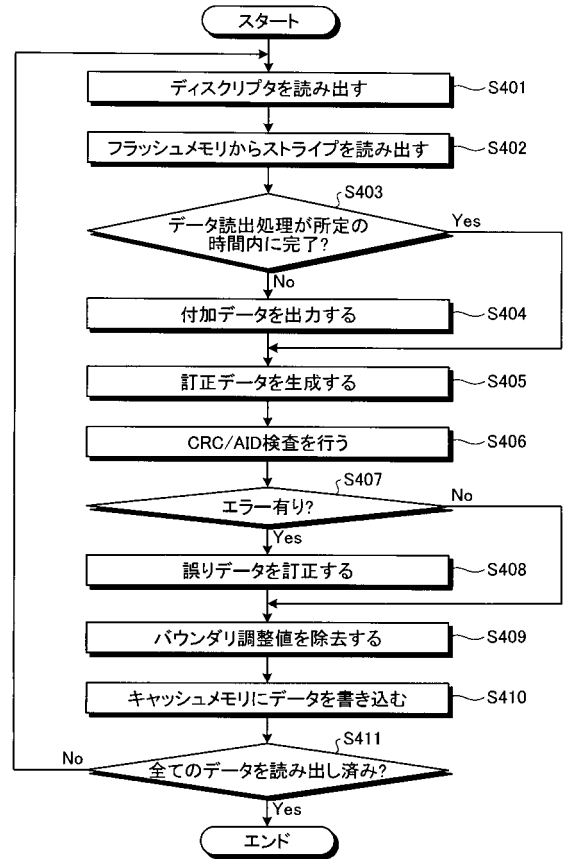
【 図 9 】



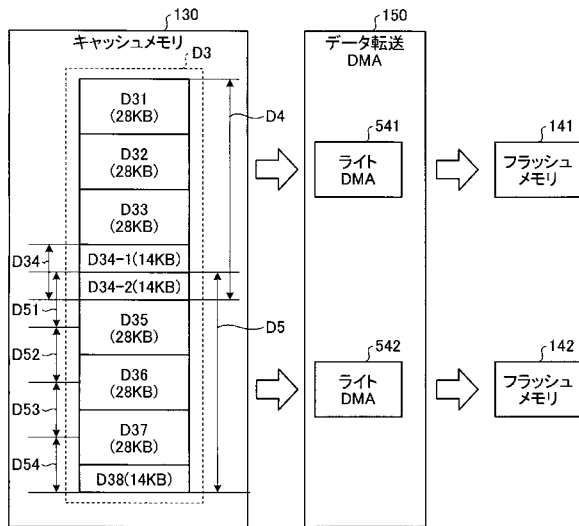
【 図 1 0 】



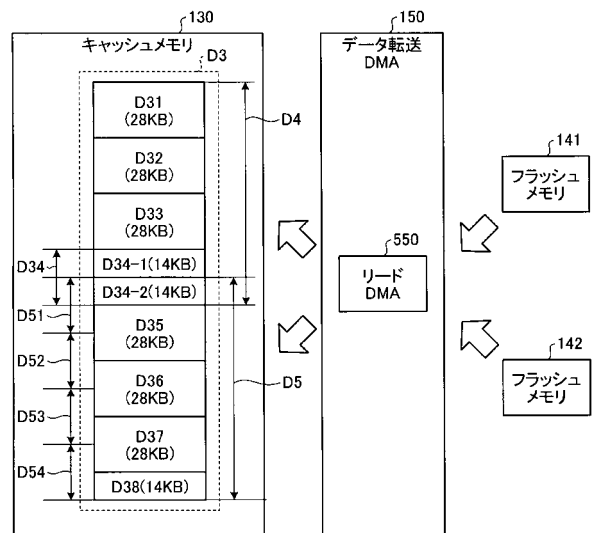
【 図 1 1 】



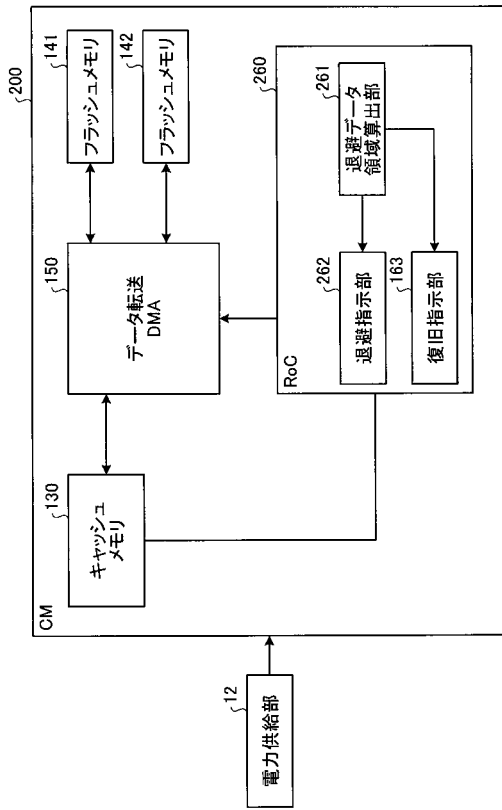
【 図 1 2 - 1 】



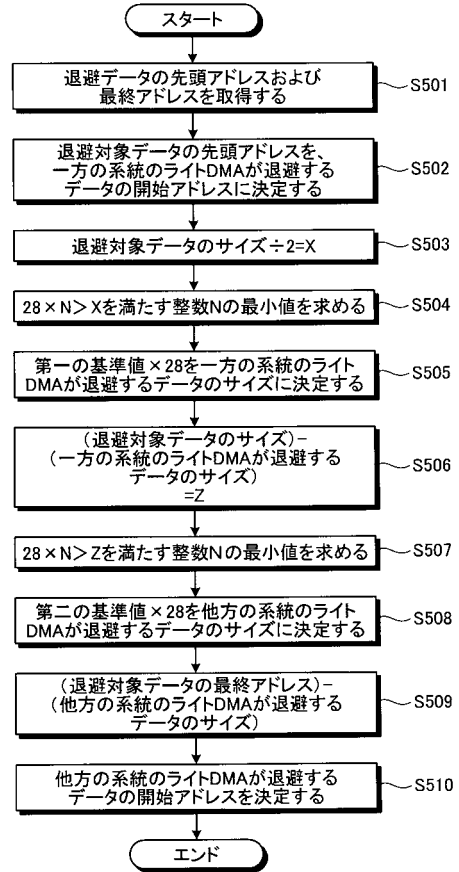
【 図 1 2 - 2 】



【 図 1 3 】



【 図 1 4 】



【 手続 補正書 】

【 提出日 】平成22年6月17日 (2010.6.17)

【 手続 補正 1 】

【 補正対象書類名 】特許請求の範囲

【 補正対象項目名 】全文

【 補正方法 】変更

【 補正の内容 】

【 特許請求の範囲 】

【 請求項 1 】

異常終了時に、キャッシュメモリに記憶されているデータである退避対象データ、および、前記退避対象データの誤りデータを訂正するためのパリティデータに、誤りデータを検出するための誤り検出符号を付与した転送対象データを、不揮発性メモリへ記憶させるデータ転送部を有するストレージ装置であって、

当該のストレージ装置が異常終了後に電源投入された場合に、前記不揮発性メモリから、前記転送対象データが所定のサイズであるストライプサイズごとに区切られたデータであるストライプを読み出すデータ読出手段と、

前記データ読出手段によって前記ストライプが所定の時間内に読み出されない場合に、該ストライプに、該ストライプのサイズがストライプサイズになるまで、付加データを付与する付加データ付与手段と、

前記データ読出手段によって読み出されたストライプに付与されている誤り検出符号に基づいて、ストライプに誤りデータが存在しているか否かを検査する誤りデータ検査手段と、

前記誤りデータ検査手段によって誤りデータが検出された場合に、前記データ読出手段によって読み出されたストライプを用いて、誤りデータを訂正するデータ訂正手段と、

前記誤りデータ検査手段によって誤りデータが検出されなかったストライプに含まれる

退避対象データを前記キャッシュメモリに書き込むとともに、前記データ訂正手段によって誤りデータが訂正されたストライプに含まれる退避対象データを前記キャッシュメモリに書き込むデータ書込手段と

を備えたことを特徴とするストレージ装置。

【請求項 2】

当該のストレージ装置が異常終了した場合に、前記退避対象データのキャッシュメモリにおける先頭の記憶位置を示す先頭アドレスから所定のサイズ分のデータを不揮発性メモリへ記憶させる第一のデータ退避手段と、

当該のストレージ装置が異常終了した場合に、前記退避対象データのキャッシュメモリにおける最終の記憶位置を示す最終アドレスから所定のサイズ分のデータを不揮発性メモリへ記憶させる第二のデータ退避手段とをさらに備えることを特徴とする請求項 1 に記載のストレージ装置。

【請求項 3】

異常終了時に、キャッシュメモリに記憶されているデータである退避対象データ、および、前記退避対象データの誤りデータを訂正するためのパリティデータに、誤りデータを検出するための誤り検出符号を付与した転送対象データを、不揮発性メモリへ記憶させるデータ転送部を有するストレージ装置を制御するストレージ制御装置であって、

前記ストレージ装置が異常終了後に電源投入された場合に、前記不揮発性メモリから、前記転送対象データが所定のサイズであるストライプサイズごとに区切られたデータであるストライプを読み出すデータ読出手段と、

前記データ読出手段によって前記ストライプが所定の時間内に読み出されない場合に、該ストライプに、該ストライプのサイズがストライプサイズになるまで、付加データを付与する付加データ付与手段と、

前記データ読出手段によって読み出されたストライプに付与されている誤り検出符号に基づいて、ストライプに誤りデータが存在しているか否かを検査する誤りデータ検査手段と、

前記誤りデータ検査手段によって誤りデータが検出された場合に、前記データ読出手段によって読み出されたストライプを用いて、誤りデータを訂正するデータ訂正手段と、

前記誤りデータ検査手段によって誤りデータが検出されなかったストライプに含まれる退避対象データを前記キャッシュメモリに書き込むとともに、前記データ訂正手段によって誤りデータが訂正されたストライプに含まれる退避対象データを前記キャッシュメモリに書き込むデータ書込手段と

を備えたことを特徴とするストレージ制御装置。

【請求項 4】

前記ストレージ装置が異常終了した場合に、前記退避対象データのキャッシュメモリにおける先頭の記憶位置を示す先頭アドレスから所定のサイズ分のデータを不揮発性メモリへ記憶させる第一のデータ退避手段と、

前記ストレージ装置が異常終了した場合に、前記退避対象データのキャッシュメモリにおける最終の記憶位置を示す最終アドレスから所定のサイズ分のデータを不揮発性メモリへ記憶させる第二のデータ退避手段とをさらに備えることを特徴とする請求項 3 に記載のストレージ制御装置。

【請求項 5】

キャッシュメモリに記憶されているデータである退避対象データ、および、前記退避対象データの誤りデータを訂正するためのパリティデータに、誤りデータを検出するための誤り検出符号を付与した転送対象データを、不揮発性メモリへ記憶させるデータ転送集積回路であって、

前記不揮発性メモリから、前記転送対象データが所定のサイズであるストライプサイズごとに区切られたデータであるストライプを読み出すデータ読出手段と、

前記データ読出手段によって前記ストライプが所定の時間内に読み出されない場合に、該ストライプに、該ストライプのサイズがストライプサイズになるまで、付加データを付

与する付加データ付与手段と、

前記データ読出手段によって読み出されたストライプに付与されている誤り検出符号に基づいて、ストライプに誤りデータが存在しているか否かを検査する誤りデータ検査手段と、

前記誤りデータ検査手段によって誤りデータが検出された場合に、前記データ読出手段によって読み出されたストライプを用いて、誤りデータを訂正するデータ訂正手段と、

前記誤りデータ検査手段によって誤りデータが検出されなかったストライプに含まれる退避対象データを前記キャッシュメモリに書き込むとともに、前記データ訂正手段によって誤りデータが訂正されたストライプに含まれる退避対象データを前記キャッシュメモリに書き込むデータ書込手段と

を備えたことを特徴とするデータ転送集積回路。

【請求項 6】

前記退避対象データのキャッシュメモリにおける先頭の記憶位置を示す先頭アドレスから所定のサイズ分のデータを不揮発性メモリへ記憶させる第一のデータ退避手段と、

前記退避対象データのキャッシュメモリにおける最終の記憶位置を示す最終アドレスから所定のサイズ分のデータを不揮発性メモリへ記憶させる第二のデータ退避手段とをさらに備えることを特徴とする請求項 5 に記載のデータ転送集積回路。

【請求項 7】

異常終了時に、キャッシュメモリに記憶されているデータである退避対象データ、および、前記退避対象データの誤りデータを訂正するためのパリティデータに、誤りデータを検出するための誤り検出符号を付与した転送対象データを、不揮発性メモリへ記憶させるデータ転送部を有するストレージ装置を制御するストレージ制御方法であって、

前記ストレージ装置が、

異常終了後に電源投入された場合に、前記不揮発性メモリから、前記転送対象データが所定のサイズであるストライプサイズごとに区切られたデータであるストライプを読み出すデータ読出工程と、

前記データ読出工程によって前記ストライプが所定の時間内に読み出されない場合に、該ストライプに、該ストライプのサイズがストライプサイズになるまで、付加データを付与する付加データ付与工程と、

前記データ読出工程によって読み出されたストライプに付与されている誤り検出符号に基づいて、ストライプに誤りデータが存在しているか否かを検査する誤りデータ検査工程と、

前記誤りデータ検査工程によって誤りデータが検出された場合に、前記データ読出工程によって読み出されたストライプを用いて、誤りデータを訂正するデータ訂正工程と、

前記誤りデータ検査工程によって誤りデータが検出されなかったストライプに含まれる退避対象データを前記キャッシュメモリに書き込むとともに、前記データ訂正工程によって誤りデータが訂正されたストライプに含まれる退避対象データを前記キャッシュメモリに書き込むデータ書込工程と

を含んだことを特徴とするストレージ制御方法。

【請求項 8】

前記ストレージ装置は、

異常終了した場合に、前記退避対象データのキャッシュメモリにおける先頭の記憶位置を示す先頭アドレスから所定のサイズ分のデータを不揮発性メモリへ記憶させる第一のデータ退避工程と、

異常終了した場合に、前記退避対象データのキャッシュメモリにおける最終の記憶位置を示す最終アドレスから所定のサイズ分のデータを不揮発性メモリへ記憶させる第二のデータ退避工程とをさらに含んだことを特徴とする請求項 7 に記載のストレージ制御方法。

【手続補正 2】

【補正対象書類名】明細書

【補正対象項目名】全文

【補正方法】変更

【補正の内容】

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、ストレージ装置、ストレージ制御装置、データ転送集積回路、および、ストレージ制御方法に関する。

【背景技術】

【0002】

従来、大規模なコンピュータシステムでは、ホストコンピュータと別体に構成された専用のデータ入出力装置を用いてデータを管理している。このデータ入出力装置の中でも、ストレージ装置は、複数のハードディスク装置（記憶媒体）によって R A I D（Redundant Arrays of Independent (Inexpensive) Disks）グループを構成して、取り扱うデータの信頼性を向上させている。

【0003】

ストレージ装置は、一般に、ライトバック方式と呼ばれるライト処理（書込処理）方式を行うことで、アクセス性能を向上させている。具体的には、ストレージ装置は、上位装置であるホストコンピュータからデータのライト命令を受け付けた場合に、キャッシュメモリにデータを記憶した時点で、ライト処理が完了した旨をホストコンピュータへ通知する。その後、所定の条件を満たした時点で、キャッシュメモリに記憶したデータをハードディスク装置へ記憶させる。

【0004】

このようなライトバック方式を採用するストレージ装置には、停電などによって正常な終了操作が行われることなく電源が落とされた場合であっても（以下、正常な終了操作が行われることなく電源が落とされることを「異常終了」という）、揮発性のキャッシュメモリに記憶されているデータが消失しないように構成されているものがある。

【0005】

具体的には、異常終了した場合に、大容量のコンデンサやバッテリー等から電力供給されている間に、キャッシュメモリ上のデータを不揮発性メモリに退避させ、異常終了後に電源投入された場合に不揮発性メモリに退避したデータをキャッシュメモリに復旧させる。これにより、キャッシュメモリに記憶されているデータがハードディスク装置に記憶されていない状態で異常終了した場合であっても、データを消失させないことを実現している。

【0006】

ところで、データを退避させる不揮発性メモリは、一般的に、N A N D型のフラッシュメモリが用いられることが多い。N A N D型のフラッシュメモリは、セル劣化等によって記憶しているデータを破損や消失してしまうことがある。不揮発性メモリに記憶されているデータが破損や消失する状況が発生すると、ストレージ装置は、異常終了後の起動時に、不揮発性メモリから誤ったデータを読み出してしまい、正確にデータを復旧させることができなくなる。

【0007】

そこで、記憶媒体に記憶されているデータのエラーを検出し、エラー検出時にデータ訂正を行う技術（例えば、特許文献1および2を参照）を用いることが考えられる。このような技術を用いれば、異常終了後の起動時に、不揮発性メモリに退避させたデータを、キャッシュメモリに復旧させることが可能になる。

【先行技術文献】

【特許文献】

【0008】

【特許文献1】特表2004-506256号公報

【特許文献2】特開2007-193449号公報

【発明の概要】

【発明が解決しようとする課題】

【0009】

しかしながら、上記特許文献1および2に開示されているような技術を用いても、不揮発性メモリから読み出し応答がなされない場合には、不揮発性メモリに退避させたデータを、正確にキャッシュメモリに復旧させることができないという問題があった。不揮発性メモリの劣化状態によっては、不揮発性メモリから読み出し応答がなされないことがあり、かかる状態では、ハングアップ状態となりデータ訂正処理が実行されないためである。

【0010】

ストレージ装置のような大規模なデータ入出力装置は、重要なコンピュータシステムに用いられることが多く、キャッシュメモリに記憶されているユーザデータやシステムデータを正確に復旧できないことは、大きな問題となる。このようなことから、不揮発性メモリから読み出し応答がなされない場合であっても、不揮発性メモリに退避させたデータを、いかにして正確にキャッシュメモリに復旧させるかが重要な課題となっていた。

【0011】

本発明は、上述した従来技術による課題を解消するためになされたものであり、不揮発性メモリから読み出し応答がなされない場合であっても、不揮発性メモリに退避させたデータを、キャッシュメモリに復旧させることができるストレージ装置、ストレージ制御装置、データ転送集積回路、および、ストレージ制御方法を提供することを目的とする。

【課題を解決するための手段】

【0012】

上述した課題を解決し、目的を達成するために、本願に開示するストレージ装置は、異常終了時に、キャッシュメモリに記憶されているデータである退避対象データ、および、前記退避対象データの誤りデータを訂正するためのパリティデータに、誤りデータを検出するための誤り検出符号を付与した転送対象データを、不揮発性メモリへ記憶させるデータ転送部を有するストレージ装置であって、当該のストレージ装置が異常終了後に電源投入された場合に、前記不揮発性メモリから、前記転送対象データが所定のサイズであるストライプサイズごとに区切られたデータであるストライプを読み出すデータ読出手段と、前記データ読出手段によって前記ストライプが所定の時間内に読み出されない場合に、該ストライプに、該ストライプのサイズがストライプサイズになるまで、付加データを付与する付加データ付与手段と、前記データ読出手段によって読み出されたストライプに付与されている誤り検出符号に基づいて、ストライプに誤りデータが存在しているか否かを検査する誤りデータ検査手段と、前記誤りデータ検査手段によって誤りデータが検出された場合に、前記データ読出手段によって読み出されたストライプを用いて、誤りデータを訂正するデータ訂正手段と、前記誤りデータ検査手段によって誤りデータが検出されなかったストライプに含まれる退避対象データを前記キャッシュメモリに書き込むとともに、前記データ訂正手段によって誤りデータが訂正されたストライプに含まれる退避対象データを前記キャッシュメモリに書き込むデータ書込手段とを備えたことを要件とする。

【0013】

なお、本願に開示するストレージ装置の構成要素、表現または構成要素の任意の組合せを、方法、装置、システム、コンピュータプログラム、記録媒体、データ構造などに適用したものも、他の態様として有効である。

【発明の効果】

【0014】

本願に開示したストレージ装置によれば、不揮発性メモリからデータの読出応答がなされない場合であっても、不揮発性メモリに退避させたデータを、キャッシュメモリに復旧させることができるという効果を奏する。

【図面の簡単な説明】

【0015】

【図1】図1は、実施例1に係るストレージ装置の概略構成を示す図である。

【図 2 - 1】図 2 - 1 は、フラッシュメモリ 1 4 1 の構成およびデータフォーマットを示す図である。

【図 2 - 2】図 2 - 2 は、図 2 - 1 に示した page の構成およびデータフォーマットを示す図である。

【図 3 - 1】図 3 - 1 は、実施例 1 に係るストレージ装置によるデータ退避処理の概要を説明するための図である。

【図 3 - 2】図 3 - 2 は、実施例 1 に係るストレージ装置によるデータ復旧処理の概要を説明するための図である。

【図 3 - 3】図 3 - 3 は、実施例 1 に係るストレージ装置によるデータ復旧処理の概要を説明するための図である。

【図 4】図 4 は、実施例 1 に係るストレージ装置が有する CM の構成を示す図である。

【図 5】図 5 は、図 4 に示したデータ転送 DMA の構成を示す図である。

【図 6】図 6 は、図 5 に示したライト DMA の構成を示す図である。

【図 7】図 7 は、図 5 に示したリード DMA の構成を示す図である。

【図 8】図 8 は、実施例 1 に係るストレージ装置によるデータ退避処理の流れを示すシーケンス図である。

【図 9】図 9 は、ライト DMA によるデータ退避処理手順を示すフローチャートである。

【図 10】図 10 は、実施例 1 に係るストレージ装置によるデータ復旧処理の流れを示すシーケンス図である。

【図 11】図 11 は、リード DMA によるデータ復旧処理手順を示すフローチャートである。

【図 12 - 1】図 12 - 1 は、実施例 2 に係るストレージ装置によるデータ退避処理の概要を説明するための図である。

【図 12 - 2】図 12 - 2 は、実施例 2 に係るストレージ装置によるデータ復旧処理の概要を説明するための図である。

【図 13】図 13 は、実施例 2 に係るストレージ装置が有する CM の構成を示す図である。

【図 14】図 14 は、図 13 に示した退避データ領域算出部による退避データ領域算出処理手順を示すフローチャートである。

【発明を実施するための形態】

【0016】

以下に、本発明にかかるストレージ装置、ストレージ制御装置、データ転送集積回路、および、ストレージ制御方法の実施例を図面に基づいて詳細に説明する。なお、この実施例によりこの発明が限定されるものではない。

【実施例 1】

【0017】

まず、実施例 1 に係るストレージ装置の概略構成について説明する。図 1 は、実施例 1 に係るストレージ装置 1 の概略構成を示す図である。同図に示したストレージ装置 1 は、高信頼性を確保するために二重化構成が採られており、各構成要素が 2 系統ずつ（同図中、参照符号に付加した A、B によって区別する）備えられている。それぞれの系統の構成要素は互いに同等の機能を有するため、以下では一方の系統についてのみ構成要素の説明をする。

【0018】

同図に示すように、ストレージ装置 1 は、ストレージ装置 1 を制御する各部や、ストレージ装置 1 と上位装置であるホストコンピュータとの間の通信を担う通信モジュールなどを主に搭載する CE (Controller Enclosure: コントローラ エンクロージャ) 10 と、ハードディスク装置を主に搭載する DE (Device Enclosure: デバイス エンクロージャ) 20 とに、筐体が分けられている。

【0019】

CE 10 は、図示しないチャネルアダプタ経由でファイバチャネルを介してホストコン

コンピュータと接続され、エキスパンダ 1 2 0 A および 1 2 0 B 経由で D E 2 0 と接続されている。D E 2 0 は、ハードディスク装置（以下、「H D D : Hard Disk Drive」と略記する）2 1 を有し、エキスパンダ 2 2 0 A および 2 2 0 B を直接、もしくは、ルータ（R o u t e r）を経由するなどして、他の D E と複数個接続することが可能である。

【 0 0 2 0 】

また、C E 1 0 は、ストレージ装置 1 を制御するための C M（Controller Module：コントローラ モジュール）1 0 0 A および 1 0 0 B と、C M 1 0 0 A および 1 0 0 B に接続される複数のハードディスク装置 1 1 とを有する。

【 0 0 2 1 】

H D D 1 1 は、C M 1 0 0 A によってデータが冗長に読み書きされる記憶装置である。なお、C M 1 0 0 A に接続される記憶装置は、ハードディスク装置に限られず、熱磁気ディスク装置や光磁気ディスク装置、または、半導体不揮発メモリを用いた半導体ディスクのような他の記憶装置であってもよい。

【 0 0 2 2 】

C M 1 0 0 A は、ストレージ装置 1 を制御するストレージ制御装置であり、インタフェース（以下、「I / F」と略記する）部 1 1 0 A と、エキスパンダ 1 2 0 A と、キャッシュメモリ 1 3 0 A と、フラッシュメモリ 1 4 1 A および 1 4 2 A と、データ転送 D M A（Direct Memory Access）1 5 0 A と、R A I D - o n - C h i p（以下、「R o C」と略記する）1 6 0 A とを有する。

【 0 0 2 3 】

I / F 部 1 1 0 A は、L A N（Local Area Network）などのネットワークへ接続するためのインタフェースである。エキスパンダ 1 2 0 A は、R o C 1 6 0 A と H D D 1 1 との間で送受信されるデータを中継する中継装置である。

【 0 0 2 4 】

キャッシュメモリ 1 3 0 A は、ホストコンピュータと H D D 1 1 との間で転送が行われるデータを一時的に記憶するメモリである。フラッシュメモリ 1 4 1 A および 1 4 2 A は、不揮発性メモリであり、例えば、N A N D 型のフラッシュメモリである。実施例 1 に係るストレージ装置 1 では、フラッシュメモリ 1 4 1 A および 1 4 2 A に、O n e N A N D（ワンナンド）型のフラッシュメモリを用いることとする。

【 0 0 2 5 】

ここで、O n e N A N D 型のフラッシュメモリ 1 4 1 A および 1 4 2 A の構成およびデータフォーマットについて説明する。なお、フラッシュメモリ 1 4 1 A および 1 4 2 A の構成およびデータフォーマットは同一であるため、ここでは、フラッシュメモリ 1 4 1 A の構成およびデータフォーマットについてのみ説明する。

【 0 0 2 6 】

図 2 - 1 は、フラッシュメモリ 1 4 1 A の構成およびデータフォーマットを示す図である。なお、フラッシュメモリ 1 4 1 A の記憶容量は、2 [G B（ギガバイト）] であるとす。同図に示すように、フラッシュメモリ 1 4 1 A は、4 つのメモリバンク # 0 0 ~ メモリバンク # 0 3 を有する。メモリバンク # 0 0 ~ メモリバンク # 0 3 は、メモリの集合体であり、それぞれがダイ # 0 およびダイ # 1 を有する。

【 0 0 2 7 】

ダイ # 0 およびダイ # 1 は、数ミリ角のフラッシュメモリであり、それぞれが B l o c k # 0 ~ B l o c k # 2 0 4 7 を有する。また、B l o c k # 0 ~ # 2 0 4 7 は、それぞれが p a g e（ページ）# 0 ~ p a g e # 6 3 を有する。p a g e が、フラッシュメモリ 1 4 1 への最小アクセス単位となる。

【 0 0 2 8 】

p a g e の構成およびデータフォーマットについて説明する。図 2 - 2 は、図 2 - 1 に示した p a g e の構成およびデータフォーマットを示す図である。同図に示すように、p a g e は、メイン領域と、スペア領域とを有する。

【 0 0 2 9 】

メイン領域は、ユーザデータが記憶される2 [K B (キロバイト)] の記憶領域であり、4 個のセクタ # 0 ~ セクタ # 3 を有する。スペア領域は、主にデータを管理するためのシステムデータ等が記憶される64 [b y t e (バイト)] の記憶領域であり、4 個のスペア # 0 ~ # 3 を有する。

【 0 0 3 0 】

スペア領域を構成する1つのスペア (例えば、スペア # 0) は、同図に示すように、システム領域と、E C C (Error Correcting Code : 誤り訂正符号) 領域と、リザーブ領域と、ユーザ領域とを有する。

【 0 0 3 1 】

システム領域は、不良セクタ情報等の各種情報が記憶される記憶領域である。E C C 領域は、フラッシュメモリ 1 4 1 A によって生成される E C C が記憶される領域である。フラッシュメモリ 1 4 1 A は、この E C C を用いてエラーチェックを行い、エラーを検出した場合に、E C C チェックエラーが発生した旨を通知する。リザーブ領域は、予備領域である。ユーザ領域は、フラッシュメモリ 1 4 1 A にデータを記憶させるシステムによってデータ管理上必要となるデータが記憶される記憶領域である。

【 0 0 3 2 】

図 1 の説明に戻って、データ転送 D M A 1 5 0 A は、キャッシュメモリ 1 3 0 A と、フラッシュメモリ 1 4 1 A および 1 4 2 A との間におけるデータ送受信を D M A により行う D M A 回路である。このデータ転送 D M A 1 5 0 A は、例えば、F P G A (Field Programmable Gate Array) によって実現される。

【 0 0 3 3 】

R o C 1 6 0 A は、C M 1 0 0 A を全体制御する制御装置であり、ホストコンピュータとのインタフェースを制御したり、キャッシュメモリ 1 3 0 A の管理を行ったりする。また、R o C 1 6 0 A は、ホストコンピュータからデータのライト命令を受け付けた場合に、かかるデータをキャッシュメモリ 1 3 0 A に記憶する。そして、R o C 1 6 0 A は、キャッシュメモリ 1 3 0 A にデータを記憶した時点で、ライト処理が完了した旨をホストコンピュータへ通知する。その後、所定の条件を満たした時点で、キャッシュメモリ 1 3 0 A に記憶したデータを H D D 1 1 に書き込みする。

【 0 0 3 4 】

また、R o C 1 6 0 A は、ストレージ装置 1 が異常終了した場合に、データ転送 D M A 1 5 0 A に対して、キャッシュメモリ 1 3 0 A に記憶されているデータをフラッシュメモリ 1 4 1 A および 1 4 2 A に退避させる処理 (以下、「データ退避処理」という) を行うように指示する。また、R o C 1 6 0 A は、異常終了後のストレージ装置 1 が起動した場合に、データ転送 D M A 1 5 0 A に対して、フラッシュメモリ 1 4 1 A および 1 4 2 A に退避したデータをキャッシュメモリ 1 3 0 A に復旧させる処理 (以下、「データ復旧処理」という) を行うように指示する。

【 0 0 3 5 】

次に、図 3 - 1 ~ 図 3 - 3 を用いて、実施例 1 に係るストレージ装置 1 によるデータ退避処理およびデータ復旧処理の概要について説明する。図 3 - 1 は、実施例 1 に係るストレージ装置 1 によるデータ退避処理の概要を説明するための図である。なお、以下では、図 1 中に付加した参照符号の A および B を付与せずに説明する。すなわち、例えば、図 3 - 1 に示したキャッシュメモリ 1 3 0 は、図 1 に示したキャッシュメモリ 1 3 0 A または 1 3 0 B に対応する。

【 0 0 3 6 】

図 3 - 1 に示すように、キャッシュメモリ 1 3 0 には、53 [K B] のデータ D 1 が記憶されているものとする。この状態で、ストレージ装置 1 が異常終了したため、R o C 1 6 0 が、データ転送 D M A 1 5 0 に対して、データ D 1 を退避するように指示したとする。かかる指示を受け付けたデータ転送 D M A 1 5 0 は、キャッシュメモリ 1 3 0 からデータ D 1 を読み出す。なお、以下では、ストレージ装置 1 が異常終了した時に、データ転送 D M A 1 5 0 によって退避されるキャッシュメモリ 1 3 0 上のデータを「退避対象データ

」と呼ぶこととする。

【0037】

そして、データ転送DMA150は、データD1を、28[KB]ごとに区切る。具体的には、同図に示すように、データ転送DMA150は、データD1を、28[KB]のデータD10と、25[KB]のデータD20とに区切る。なお、データ転送DMA150が退避対象データ(上記例ではデータD1)を分割するサイズ(上記例では、28[KB])は、予めシステムで決められており、以下では、このサイズを「バウンダリサイズ」と呼び、バウンダリサイズによって区切られた退避対象データの1つのデータを「バウンダリデータ」と呼ぶこととする。実施例1に係るストレージ装置1では、バウンダリサイズが28[KB]に決定されていることとする。

【0038】

そして、データ転送DMA150は、データD10を、2[KB]ごとに区切る。具体的には、同図に示すように、データ転送DMA150は、データD10を、14個のデータD101~D114に区切る。なお、データ転送DMA150がバウンダリデータ(上記例では、データD10およびD20)を分割するサイズ(上記例では、2[KB])は、予めシステムで決められており、実施例1に係るストレージ装置1では、このサイズが、2[KB]に決定されていることとする。以下では、2[KB]ごとに区切られたバウンダリデータの1つのデータ(データD101等)を、「退避対象単位データ」と呼ぶこととし、退避対象単位データを識別するための番号を「退避対象単位データ番号」と呼ぶこととする。本明細書では、退避対象単位データに付した参照符号の数字部分を退避対象単位データ番号とする。例えば、データD101の退避対象単位データ番号は、「101」である。

【0039】

そして、データ転送DMA150は、退避対象単位データ番号が奇数である7個のデータD101、D103、・・・、D113の排他的論理和(以下、「XOR」という)演算を行ってパリティデータP115を生成する。また、データ転送DMA150は、退避対象単位データ番号が偶数である7個のデータD102、D104、・・・、D114のXOR演算を行ってパリティデータP116を生成する。

【0040】

このパリティデータP115およびP116のサイズは、データD101~D114のサイズと同様の2[KB]となる。すなわち、バウンダリサイズ28[KB]に区切られたデータD10のサイズと、かかるデータD10から生成されたパリティデータP115およびP116のサイズの合計は、32[KB]となる。なお、以下では、パリティデータも「退避対象単位データ」と呼ぶこととする。

【0041】

その後、データ転送DMA150は、データD101~D114ごとに、CRC(Cyclic Redundancy Check)を生成する。同様に、データ転送DMA150は、パリティデータP115およびP116ごとに、CRCを生成する。そして、データ転送DMA150は、生成したCRCを、データD101~D114と、パリティデータP115およびP116とに付与する。

【0042】

また、データ転送DMA150は、データD101~D114と、パリティデータP115およびP116とを識別するための識別番号(以下、「AID:Area Identification」という)をデータD101~D114と、パリティデータP115およびP116とに付与する。このAIDについては、後に詳述する。

【0043】

そして、データ転送DMA150は、CRCおよびAIDが付与されたデータD101~D114と、CRCおよびAIDが付与されたパリティデータP115およびP116とをフラッシュメモリ141に書き込む。すなわち、データ転送DMA150は、16個の退避対象単位データ(14個のデータD101~D114と、2個のパリティデータP

115およびP116)を、フラッシュメモリ141に書き込む。

【0044】

このとき、データ転送DMA150は、フラッシュメモリ141が有する4個のメモリバンク#00～メモリバンク#03に対して、退避対象単位データを4個ずつ書き込む。また、このとき、データ転送DMA150は、メモリバンク#00～メモリバンク#03が有する2個のダイ#0およびダイ#1に対して、退避対象単位データを2個ずつ書き込む。また、このとき、データ転送DMA150は、ダイ#0およびダイ#1が有する所定の2個のBlockに対して、退避対象単位データ番号が奇数である退避対象単位データと、退避対象単位データ番号が偶数である退避対象単位データとを1個ずつ書き込む。

【0045】

以下では、データ転送DMA150が、16個の退避対象単位データを書き込むフラッシュメモリ141上の領域を「小エリア」と呼ぶこととする。この「小エリア」は、例えば、図2-1に示した小エリアSA1に該当する。

【0046】

ここで、データ転送DMA150のデータ書込処理について、データD101～D114と、パリティデータP115およびP116とを、図2-1に示した小エリアAS1に書き込む場合を例に挙げて説明する。なお、以下では、「メモリバンク#Pのダイ#QのBlock#Rが有するpage#S」を示す場合に、「メモリバンク#P/ダイ#Q/Block#R/page#S」と記載することとする。

【0047】

かかる場合、データ転送DMA150は、例えば、データD101を、小エリアSA1内のメモリバンク#00/ダイ#0/Block#2/page#1のメイン領域に書き込む。そして、データ転送DMA150は、データD102を、メモリバンク#00/ダイ#0/Block#3/page#1のメイン領域に書き込み、データD103を、メモリバンク#00/ダイ#1/Block#2/page#1のメイン領域に書き込む。このようにして、データ転送DMA150は、データD104～D113を書き込んでいき、データD114を、メモリバンク#03/ダイ#0/Block#3/page#1に書き込み、パリティデータP115を、メモリバンク#03/ダイ#1/Block#2/page#1に書き込み、パリティデータP116を、メモリバンク#03/ダイ#1/Block#3/page#1に書き込む。

【0048】

そして、データ転送DMA150は、各退避対象単位データに付与したCRCをpageのスペア領域内のユーザ領域に書き込む。上記例のように、データD101を、メモリバンク#00/ダイ#0/Block#2/page#1のメイン領域に書き込んだ場合、データ転送DMA150は、データD101のCRCを、メモリバンク#00/ダイ#0/Block#2/page#1のスペア#0のユーザ領域に書き込む。同様に、データ転送DMA150は、データD102～D114と、パリティデータP115およびP116のCRCについても、pageのユーザ領域に書き込む。

【0049】

また、データ転送DMA150は、各退避対象単位データに付与したAIDをpageのスペア領域内のユーザ領域に書き込む。上記例のように、データD101のCRCを、メモリバンク#00/ダイ#0/Block#2/page#1のスペア#0のユーザ領域に書き込んだ場合、データ転送DMA150は、データD101のAIDを、メモリバンク#00/ダイ#0/Block#2/page#1のスペア#1のユーザ領域に書き込む。同様に、データ転送DMA150は、データD102～D114と、パリティデータP115およびP116のAIDについても、pageのユーザ領域に書き込む。

【0050】

このようにして、データ転送DMA150は、16個の退避対象単位データを、小エリアに書き込む。以下では、小エリア内の1つのpageに記憶されている全てのデータ(退避対象単位データやCRCやAID等)を「ストライプ」と呼ぶこととし、小エリア内

でストライプを識別するための番号を「ストライプ番号」と呼ぶこととする。本明細書では、退避対象単位データに付した参照符号の数字下2桁部分をストライプ番号とする。例えば、データD101を記憶するpageに記憶されているストライプのストライプ番号は「1」であり、パリティデータP115を記憶するpageに記憶されているストライプのストライプ番号は「15」である。以下では、ストライプ番号「N」のストライプを、単に、ストライプS「N」と呼ぶこととする。

【0051】

なお、図3-1に示したフラッシュメモリ141内の「D101」や「P115」等を記載した矩形は、pageのメイン領域を示し、かかるメイン領域の後に斜線を付して示した矩形は、pageのスペア領域を示す。すなわち、同図では、例えば、「D101」を記載したメイン領域と、かかるメイン領域の後に示したスペア領域との組合せがストライプS1であることを示す。

【0052】

同様にして、データ転送DMA150は、データD20をフラッシュメモリ142に書き込む。具体的には、データ転送DMA150は、データD20を2[KB]ごとに区切る。このとき、データ転送DMA150は、データD20のサイズが25[KB]であり、バウンダリサイズの28[KB]に満たないので、データD20のサイズが28[KB]になるように、3[KB]の所定の値を加える。以下では、この所定の値を「バウンダリ調整値」と呼ぶこととする。バウンダリデータのサイズがバウンダリサイズに満たない場合に、バウンダリ調整値を加える理由は、パリティデータを生成できるようにするためである。

【0053】

そして、データ転送DMA150は、バウンダリ調整値を加えたデータD20を2[KB]ごとに区切る。具体的には、同図に示すように、データ転送DMA150は、3[KB]のバウンダリ調整値を加えたデータD20を、14個のデータD201~D214に区切る。

【0054】

その後の処理は、上述したデータD10をフラッシュメモリ141に書き込む処理と同様であり、データ転送DMA150は、パリティデータP215およびP216を生成して、データD201等のCRCを生成して、データD201~D214と、パリティデータP215およびP216と、CRCと、AIDとをフラッシュメモリ142に書き込む。

【0055】

このようにして、データ転送DMA150は、キャッシュメモリ130に記憶されているデータD1のうち、データD10をフラッシュメモリ141に書き込み、データD20をフラッシュメモリ142に書き込む。

【0056】

なお、データ転送DMA150は、複数の小エリアに対して、データを並行して書き込むことが可能である。実施例1に係るストレージ装置1では、データ転送DMA150が、最大で64個の小エリアに対して同時にデータを書き込むこととする。以下では、この64個の小エリアを「エリア」と呼ぶこととする。この「エリア」は、例えば、図2-1に示したエリアA1に該当する。上述したAIDは、1個のエリア内に含まれるストライプを識別するための識別番号である。このAIDは、Roc160によって算出され、データ転送DMA150は、Roc160からAIDを受け付ける。

【0057】

次に、実施例1に係るストレージ装置1によるデータ復旧処理の概要について説明する。図3-2および図3-3は、実施例1に係るストレージ装置1によるデータ復旧処理の概要を説明するための図である。図3-2および図3-3に示したフラッシュメモリ141には、図3-1に示したデータ転送DMA150によって退避されたデータが記憶されているものとする。

【 0 0 5 8 】

まず、図 3 - 2 を用いて、図 3 - 1 に示したデータ D 1 0 をキャッシュメモリ 1 3 0 に復旧させる処理について説明する。ここでは、図 3 - 2 に示したフラッシュメモリ 1 4 1 に記憶されているデータ D 1 0 3 の一部が、フラッシュメモリ 1 4 1 のセル劣化等の原因により破損しているものとする。そして、フラッシュメモリ 1 4 1 は、データ D 1 0 3 の読出要求を受け付けると、かかる読出要求に対する応答ができなくなるものとする。

【 0 0 5 9 】

かかる状態で、ストレージ装置 1 が異常終了後に起動した場合、データ転送 D M A 1 5 0 は、フラッシュメモリ 1 4 1 から、小エリア単位でストライプを読み出す。具体的には、データ転送 D M A 1 5 0 は、フラッシュメモリ 1 4 1 から、ストライプ S 1 ~ S 1 6 を読み出す。

【 0 0 6 0 】

そして、データ転送 D M A 1 5 0 は、ストライプ S 3 を読み出している途中で、フラッシュメモリ 1 4 1 から読出応答を受け付けなくなる。所定の時間が経過してもストライプ S 3 を読み出せない場合、データ転送 D M A 1 5 0 は、ストライプ S 3 に対してタイムアウトが発生したことを検出する。

【 0 0 6 1 】

そして、データ転送 D M A 1 5 0 は、タイムアウトが発生したストライプ S 3 の読出処理を中止して、ストライプ S 3 に、読み出しできなかったサイズ分の所定の値（以下、この所定の値を「付加データ」という）を設定する。例えば、1 [K B] 分のストライプ S 3 を読み出した時点で、フラッシュメモリ 1 4 1 から読み出し応答を受け付けなくなった場合、データ転送 D M A 1 5 0 は、読み出し済みの 1 [K B] のデータに 1 [K B] の付加データを加えた 2 [K B] のデータをストライプ S 3 とする。その後、データ転送 D M A 1 5 0 は、データの読出処理を再開して、フラッシュメモリ 1 4 1 から、ストライプ S 4 ~ S 1 6 を読み出す。

【 0 0 6 2 】

また、データ転送 D M A 1 5 0 は、ストライプ S 1 ~ 1 6 を読み出しながら、C R C 検査、A I D 検査および X O R 演算を行う。具体的には、データ転送 D M A 1 5 0 は、読み出したストライプ S 1 ~ 1 6 内の C R C を用いて、ストライプ S 1 ~ 1 6 に誤りデータが存在するか否かを検査する。また、データ転送 D M A 1 5 0 は、ストライプ S 1 ~ 1 6 内の A I D を用いて、ストライプ S 1 ~ 1 6 がすべて揃っているか否かを検査する。また、データ転送 D M A 1 5 0 は、ストライプ番号が奇数であるストライプ S 1、S 3、・・・、S 1 5 の X O R 演算を行い、訂正データ T 1 1 を生成するとともに、ストライプ番号が偶数であるストライプ S 2、S 4、・・・、S 1 6 の X O R 演算を行い、訂正データ T 1 2 を生成する。

【 0 0 6 3 】

図 3 - 2 に示した例では、データ転送 D M A 1 5 0 は、付加データが設定されているストライプ S 3 の C R C 検査において、誤りデータが存在することを検出する。なお、ストライプ S 3 に含まれる A I D を読み出す前に、タイムアウトが発生した場合には、A I D には付加データが設定されているため、データ転送 D M A 1 5 0 は、ストライプ S 3 の A I D 検査において、データ D 1 0 3 が欠如していると検出する可能性もある。ここでは、データ転送 D M A 1 5 0 は、ストライプ S 3 に対してのみデータ誤りを検出し、ストライプ S 1、S 2、S 4 ~ S 1 6 に対しては、C R C 検査および A I D 検査においてエラーを検出しなかったものとする。

【 0 0 6 4 】

すべてのストライプ S 1 ~ 1 6 を読み出した後、データ転送 D M A 1 5 0 は、C R C 検査および A I D 検査においてエラーが検出されなかったストライプ S 1 および S 2 に含まれるデータ D 1 0 1 および D 1 0 2 をキャッシュメモリ 1 3 0 に書き込む。そして、データ転送 D M A 1 5 0 は、エラーが検出されたストライプ S 3 については、ストライプ番号が奇数であるストライプ S 1、S 5、S 7、・・・、S 1 5 および訂正データ T 1 1 を用い

て、データD103の誤りデータを訂正して、訂正済みのデータD103をキャッシュメモリ130に書き込む。そして、データ転送DMA150は、エラーが検出されなかったストライプS4～S14に含まれるデータD104～D114をキャッシュメモリ130に書き込む。なお、ストライプS15およびS16はパリティデータであるため、データ転送DMA150は、ストライプS15およびS16に含まれるパリティデータP115およびP116をキャッシュメモリ130に書き込まない。このようにして、データ転送DMA150は、データD10をキャッシュメモリ130に復旧する。

【0065】

続いて、図3-3を用いて、図3-1に示したデータD20をキャッシュメモリ130に復旧させる処理について説明する。同図に示すように、データ転送DMA150は、データD10の復旧処理が終了した後、フラッシュメモリ142から、ストライプS1～S16を読み出しながら、CRC検査、AID検査およびXOR演算を行い、訂正データT21およびT22を生成する。ここでは、データ転送DMA150は、フラッシュメモリ142からストライプS1～S16の読出処理中に、タイムアウトが発生しなかったものとする。また、データ転送DMA150は、ストライプS1～S16のCRC検査およびAID検査においてエラーを検出しなかったものとする。

【0066】

すべてのストライプS1～16を読み出した後、データ転送DMA150は、ストライプS1～S12に含まれるデータD201～D212をキャッシュメモリ130に書き込む。そして、データ転送DMA150は、ストライプS13に含まれるデータD213から1[KB]のバウンダリ調整値を除去したデータD213をキャッシュメモリ130に書き込む。なお、ストライプS14に含まれるデータD214は、すべてバウンダリ調整値であるため、データ転送DMA150は、データD214をキャッシュメモリ130に書き込まない。

【0067】

以上のように、実施例1に係るストレージ装置1は、データ退避処理時に、退避対象データのパリティデータを生成して、CRCやAIDを付与した退避対象データおよびパリティデータをフラッシュメモリ141または142に書き込み、データ復旧処理において、フラッシュメモリ141または142から所定の時間内にデータ読出処理が完了しない場合に、データ読出処理を中止して付加データを設定し、パリティデータを用いて誤りデータを訂正したデータをキャッシュメモリ130に書き込むので、フラッシュメモリ141または142からデータの読出応答がなされない場合であっても、フラッシュメモリ141または142に退避させたデータを、キャッシュメモリ130に復旧させることができる。

【0068】

また、実施例1に係るストレージ装置1は、退避対象単位データ番号が奇数である退避対象単位データを用いてパリティデータを生成するとともに、退避対象単位データ番号が偶数である退避対象単位データを用いてパリティデータを生成し、小エリア内の1つのダイには、退避対象単位データ番号が奇数である退避対象単位データと、退避対象単位データ番号が偶数である退避対象単位データとを書き込むように構成したので、1個のダイが故障した場合であっても、かかる故障したダイに記憶されていたデータを訂正することができる。すなわち、1個のダイが故障した場合であっても、フラッシュメモリ141または142に退避させたデータを、キャッシュメモリ130に復旧させることができる。

【0069】

次に、実施例1に係るストレージ装置1が有するCM100の構成について説明する。図4は、実施例1に係るストレージ装置1が有するCM100の構成を示す図である。なお、同図では、データ退避処理およびデータ復旧処理に関連する構成のみを図示している。

【0070】

同図に示すように、CM100は、電力供給部12と接続されており、キャッシュメモ

リ 1 3 0 と、フラッシュメモリ 1 4 1 および 1 4 2 と、データ転送 D M A 1 5 0 と、R o C 1 6 0 とを有する。

【 0 0 7 1 】

電力供給部 1 2 は、ストレージ装置 1 が異常終了した場合に、C M 1 0 0 に対して電力供給を行う装置であり、例えば、大容量のコンデンサやバッテリー等である。なお、同図では、C M 1 0 0 が外部に備えられた電力供給部 1 2 によって電力供給される例を示しているが、C M 1 0 0 は、電力供給部 1 2 を有してもよい。

【 0 0 7 2 】

キャッシュメモリ 1 3 0 は、上述したように、ホストコンピュータと H D D 1 1 との間で転送が行われるデータを一時的に記憶するメモリである。フラッシュメモリ 1 4 1 および 1 4 2 は、不揮発性メモリであり、ストレージ装置 1 の異常終了時にデータ転送 D M A 1 5 0 によって、キャッシュメモリ 1 3 0 に記憶されているデータが退避される。

【 0 0 7 3 】

データ転送 D M A 1 5 0 は、R o C 1 6 0 に指示に従って、キャッシュメモリ 1 3 0 と、フラッシュメモリ 1 4 1 および 1 4 2 との間におけるデータ送受信を行う D M A 回路である。具体的には、データ転送 D M A 1 5 0 は、ストレージ装置 1 が異常終了した場合に、キャッシュメモリ 1 3 0 に記憶されているデータを、フラッシュメモリ 1 4 1 および 1 4 2 に退避させるデータ退避処理を行う。また、データ転送 D M A 1 5 0 は、ストレージ装置 1 が異常終了後に電源投入された場合に、フラッシュメモリ 1 4 1 および 1 4 2 に退避したデータをキャッシュメモリ 1 3 0 に復旧するデータ復旧処理を行う。なお、データ転送 D M A 1 5 0 の構成、データ転送 D M A 1 5 0 によるデータ退避処理およびデータ復旧処理については、後に詳述する。

【 0 0 7 4 】

R o C 1 6 0 は、C M 1 0 0 を全体制御する制御装置であり、退避データ領域算出部 1 6 1 と、退避指示部 1 6 2 と、復旧指示部 1 6 3 とを有する。退避データ領域算出部 1 6 1 は、ストレージ装置 1 が異常終了した場合に、データ転送 D M A 1 5 0 にデータ退避処理を行うように指示する際に必要となる各種情報を算出する処理部である。

【 0 0 7 5 】

具体的には、退避データ領域算出部 1 6 1 は、退避対象データのサイズや、退避対象データのキャッシュメモリ 1 3 0 における記憶位置（アドレス）に基づいて、データ転送 D M A 1 5 0 が退避するデータのサイズや、データ転送 D M A 1 5 0 が退避するデータの記憶位置や、バウンダリ調整値のサイズ等を算出する。

【 0 0 7 6 】

退避指示部 1 6 2 は、データ転送 D M A 1 5 0 に対して、データ退避処理を行うように指示する処理部である。具体的には、退避指示部 1 6 2 は、ストレージ装置 1 が異常終了した場合に、データ退避処理を行うように指示する命令（以下、「データ退避ディスクリプタ」という）を生成して、データ転送 D M A 1 5 0 に送信する。

【 0 0 7 7 】

このデータ退避ディスクリプタには、退避データ領域算出部 1 6 1 によって算出された各種情報が含まれる。具体的には、データ退避ディスクリプタには、退避対象データのキャッシュメモリ 1 3 0 における先頭の記憶位置（先頭アドレス）や、退避対象データのサイズや、バウンダリ調整値のサイズや、A I D 等が含まれる。

【 0 0 7 8 】

復旧指示部 1 6 3 は、データ転送 D M A 1 5 0 に対して、データ復旧処理を行うように指示する処理部である。具体的には、復旧指示部 1 6 3 は、ストレージ装置 1 が異常終了後に電源投入された場合に、データ復旧処理を行うように指示する命令（以下、「データ復旧ディスクリプタ」という）を生成して、データ転送 D M A 1 5 0 に送信する。このデータ復旧ディスクリプタには、除去するバウンダリ調整値のサイズ等が含まれる。

【 0 0 7 9 】

また、復旧指示部 1 6 3 は、データ転送 D M A 1 5 0 によるデータ復旧処理が完了した

場合に、フラッシュメモリ 141 および 142 に記憶されているデータを消去する命令（以下、「イレスディスクリプタ」という）を生成して、データ転送 DMA 150 に送信する。

【0080】

次に、データ転送 DMA 150 の構成について説明する。図 5 は、図 4 に示したデータ転送 DMA 150 の構成を示す図である。同図に示すように、データ転送 DMA 150 は、P C I (Peripheral Component Interconnect) バス I / F 510 と、メモリ I / F 521 および 522 と、ディスクリプタ保持レジスタ 530 と、ライト DMA 541 および 542 と、リード DMA 550 と、イレス DMA 561 および 562 とを有する。

【0081】

P C I バス I / F 510 は、データ転送 DMA 150 とキャッシュメモリ 130 との間のデータ送受信を行うインタフェースである。メモリ I / F 521 は、データ転送 DMA 150 とフラッシュメモリ 141 との間のデータ送受信を行うインタフェースであり、メモリ I / F 522 は、データ転送 DMA 150 とフラッシュメモリ 142 との間のデータ送受信を行うインタフェースである。

【0082】

ディスクリプタ保持レジスタ 530 は、R o C 160 から受信した各種ディスクリプタを記憶するレジスタである。具体的には、ディスクリプタ保持レジスタ 530 は、データ退避ディスクリプタや、データ復旧ディスクリプタや、イレスディスクリプタ等を記憶する。

【0083】

ライト DMA 541 および 542 は、データ退避処理を行う DMA 回路である。具体的には、ライト DMA 541 は、ディスクリプタ保持レジスタ 530 からデータ退避ディスクリプタを受け取った場合に、P C I バス I / F 510 を介してキャッシュメモリ 130 に記憶されているデータを取得する。そして、ライト DMA 541 は、取得したデータを、メモリ I / F 521 を介してフラッシュメモリ 141 に書き込む。同様に、ライト DMA 542 は、ディスクリプタ保持レジスタ 530 からデータ退避ディスクリプタを受け取った場合に、キャッシュメモリ 130 に記憶されているデータを取得し、取得したデータを、フラッシュメモリ 142 に書き込む。

【0084】

ライト DMA 541 および 542 は、キャッシュメモリ 130 からフラッシュメモリ 141 または 142 にデータを転送する処理を並行して行う。図 3 - 1 に示した例において、例えば、ライト DMA 541 がデータ D 10 をフラッシュメモリ 141 へ書き込む処理を行い、ライト DMA 542 がデータ D 20 をフラッシュメモリ 142 へ書き込む処理を行う。そして、このライト DMA 541 および 542 によるデータ退避処理が並行して行われる。

【0085】

このように、ストレージ装置 1 は、2 系統のフラッシュメモリ 141 および 142 と、2 系統のライト DMA 541 および 542 とを有して、ライト DMA 541 および 542 に対してデータ退避処理を並行して行わせる。これにより、ストレージ装置 1 は、データ退避処理を効率よく行うことが可能となり、電力供給部 12 から電力供給される短い時間内にデータ退避処理を完了させることを実現している。

【0086】

リード DMA 550 は、データ復旧処理を行う DMA 回路である。具体的には、リード DMA 550 は、ディスクリプタ保持レジスタ 530 から、データ復旧ディスクリプタを受け取った場合に、フラッシュメモリ 141 に記憶されているデータを取得し、取得したデータを、キャッシュメモリ 130 に記憶させる。同様に、リード DMA 550 は、データ復旧ディスクリプタを受け取った場合に、フラッシュメモリ 142 に記憶されているデータを取得し、取得したデータを、キャッシュメモリ 130 に記憶させる。

【0087】

実施例 1 に係るストレージ装置 1 が、データ退避処理を 2 系統のライト DMA 5 4 1 および 5 4 2 に行わせるのに対して、データ復旧処理を 1 系統のリード DMA 5 5 0 に行わせる理由について説明する。これは、データ退避処理には、電力供給部 1 2 から電力供給される短い時間内に完了させなければならないという制約があるのに対して、データ復旧処理には、このような制約がないためである。

【 0 0 8 8 】

ただし、データ転送 DMA 1 5 0 が有するライト DMA やリード DMA の数は、図 5 に示した例に限られない。例えば、データ転送 DMA 1 5 0 は、例えば、3 個のライト DMA を有してもよいし、1 0 個のリード DMA を有してもよい。

【 0 0 8 9 】

イレース DMA 5 6 1 は、ディスクリプタ保持レジスタ 5 3 0 からイレースディスクリプタを受け取った場合に、フラッシュメモリ 1 4 1 に記憶されているデータを消去する DMA 回路である。同様に、イレース DMA 5 6 2 は、ディスクリプタ保持レジスタ 5 3 0 からイレースディスクリプタを受け取った場合に、フラッシュメモリ 1 4 2 に記憶されているデータを消去する DMA 回路である。

【 0 0 9 0 】

次に、ライト DMA 5 4 1 および 5 4 2 の構成について説明する。図 6 は、図 5 に示したライト DMA 5 4 1 および 5 4 2 の構成を示す図である。なお、ライト DMA 5 4 1 および 5 4 2 の構成は同一であるため、ここでは、ライト DMA 5 4 1 の構成についてのみ説明する。

【 0 0 9 1 】

同図に示すように、ライト DMA 5 4 1 は、データ読出部 5 4 1 a と、バウンダリ調整部 5 4 1 b と、パリティ生成部 5 4 1 c と、退避データバッファ 5 4 1 d と、パリティバッファ 5 4 1 e と、データ書込部 5 4 1 f とを有する。

【 0 0 9 2 】

データ読出部 5 4 1 a は、キャッシュメモリ 1 3 0 から所定のサイズのデータを読み出す処理部である。具体的には、データ読出部 5 4 1 a は、データ退避ディスクリプタに含まれている退避対象データの先頭アドレスから、退避対象データのサイズ分のデータを読み出す。

【 0 0 9 3 】

バウンダリ調整部 5 4 1 b は、データ読出部 5 4 1 a によって読み出された退避対象データにバウンダリ調整値を加える処理部である。具体的には、バウンダリ調整部 5 4 1 b は、退避対象データに、データ退避ディスクリプタに含まれているバウンダリ調整値のサイズ分のバウンダリ調整値を加える。

【 0 0 9 4 】

パリティ生成部 5 4 1 c は、データ読出部 5 4 1 a によって読み出された退避対象データを用いて、パリティデータを生成する処理部である。具体的には、パリティ生成部 5 4 1 c は、退避対象データを、バウンダリサイズ 2 8 [K B] ごとのバウンダリデータに区切って、さらに、バウンダリデータを、予めシステムで決められているサイズである 2 [K B] ごとに区切る。

【 0 0 9 5 】

そして、パリティ生成部 5 4 1 c は、退避対象単位データ番号が奇数である 7 個の退避対象単位データを用いてパリティデータを生成するとともに、退避対象単位データ番号が偶数である 7 個の退避対象単位データを用いてパリティデータを生成する。そして、データ読出部 5 4 1 a によって読み出された退避対象データを、退避データバッファ 5 4 1 d に記憶させるとともに、生成したパリティデータをパリティバッファ 5 4 1 e に記憶させる。

【 0 0 9 6 】

退避データバッファ 5 4 1 d は、データを一次記憶する記憶装置であり、上述したように、パリティ生成部 5 4 1 c によって退避対象データが記憶される。パリティバッファ 5

4 1 e は、データを一次記憶する記憶装置であり、上述したように、パリティ生成部 5 4 1 c によってパリティデータが記憶される。

【0097】

データ書込部 5 4 1 f は、退避データバッファ 5 4 1 d およびパリティバッファ 5 4 1 e から、退避対象単位データおよびパリティデータを読み出してフラッシュメモリ 1 4 1 に書き込む処理部であり、CRC 付与部 5 4 1 g と A I D 付与部 5 4 1 h とを有する。

【0098】

CRC 付与部 5 4 1 g は、退避データバッファ 5 4 1 d から読み出した退避対象単位データごとに CRC を生成するとともに、パリティバッファ 5 4 1 e から読み出したパリティデータごとに CRC を生成する。そして、CRC 付与部 5 4 1 g は、生成した CRC を、退避対象単位データおよびパリティデータに付与する。

【0099】

A I D 付与部 5 4 1 h は、退避データバッファ 5 4 1 d から読み出した退避対象単位データに、A I D を付与するとともに、パリティバッファ 5 4 1 e から読み出したパリティデータに、A I D を付与する。なお、A I D 付与部 5 4 1 h は、データ退避ディスクリプタから A I D を取得する。

【0100】

データ書込部 5 4 1 f は、このようにして CRC および A I D が付与された退避対象単位データおよびパリティデータを、フラッシュメモリ 1 4 1 に書き込む。

【0101】

次に、リードDMA 5 5 0 の構成について説明する。図 7 は、図 5 に示したリードDMA 5 5 0 の構成を示す図である。同図に示すように、リードDMA 5 5 0 は、セクタ 5 5 0 a と、データ読出部 5 5 0 b と、タイムアウト検出部 5 5 0 c と、エラー検出部 5 5 0 d と、復旧データバッファ 5 5 0 e と、X O R 演算部 5 5 0 f と、CRC 検査部 5 5 0 g と、A I D 検査部 5 5 0 h と、訂正データバッファ 5 5 0 i と、ログ保持レジスタ 5 5 0 j と、セクタ 5 5 0 k と、バウンダリ調整部 5 5 0 l と、データ書込部 5 5 0 m とを有する。

【0102】

セクタ 5 5 0 a は、複数の入力信号の中から 1 つの入力信号のみを選択して出力するセクタ回路であり、フラッシュメモリ 1 4 1 に記憶されているストライプをデータ読出部 5 5 0 b に出力したり、付加データをデータ読出部 5 5 0 b に出力したりする。セクタ 5 5 0 a が、ストライプまたは付加データのいずれを出力するかは、後に説明する。

【0103】

データ読出部 5 5 0 b は、セクタ 5 5 0 a を介して、フラッシュメモリ 1 4 1 から、小エリアのストライプを読み出して、読み出したストライプを復旧データバッファ 5 5 0 e に記憶させる。このとき、データ読出部 5 5 0 b は、データ読出処理を開始した時刻をタイムアウト検出部 5 5 0 c に通知する。

【0104】

タイムアウト検出部 5 5 0 c は、データ読出部 5 5 0 b から受け付けたデータ読出処理の開始時刻に基づいて、データ読出部 5 5 0 b によるデータ読出処理が所定の時間内に完了するか否かを監視する処理部である。

【0105】

具体的には、タイムアウト検出部 5 5 0 c は、データ読出部 5 5 0 b によるデータ読出処理が、データ読出処理の開始時刻から所定の時間を経過しても完了しない場合に、データ読出部 5 5 0 b が読取処理中のストライプに対してタイムアウトが発生したことを検出する。そして、タイムアウト検出部 5 5 0 c は、タイムアウトを検出したストライプのストライプ番号を含むタイムアウト検出通知をセクタ 5 5 0 a に対して行う。

【0106】

ここで、タイムアウト検出通知を受け付けたセクタ 5 5 0 a の処理について説明する。タイムアウト検出通知を受け付けたセクタ 5 5 0 a は、タイムアウト検出通知に含ま

れるストライプ番号が示すストライプの読出処理を中止して、かかるストライプの読み出しできなかつたサイズ分の付加データをデータ読出部 550 b に出力する。その後、セクタ 550 a は、次のストライプの読出処理を再開する。

【0107】

エラー検出部 550 d は、フラッシュメモリ 141 から ECC チェックエラーが発生した旨の通知を受け付けた場合に、ECC エラー通知をセクタ 550 a に対して行う処理部である。

【0108】

ここで、ECC エラー通知を受け付けたセクタ 550 a の処理について説明する。ECC エラー通知を受け付けたセクタ 550 a は、読出処理中のストライプの読出処理を中止して、かかるストライプの読み出しできなかつたサイズ分の付加データをデータ読出部 550 b に出力する。その後、セクタ 550 a は、次のストライプの読出処理を再開する。

【0109】

復旧データバッファ 550 e は、データを一次記憶する記憶装置であり、上述したように、データ読出部 550 b によって、ストライプまたは付加データが設定されているストライプが記憶される。

【0110】

XOR 演算部 550 f は、データ読出部 550 b によって読み出されたストライプを用いて訂正データを生成する処理部である。具体的には、XOR 演算部 550 f は、ストライプ番号が奇数である 7 個のストライプを用いて訂正データを生成するとともに、ストライプ番号が偶数である 7 個のストライプを用いて訂正データを生成する。そして、XOR 演算部 550 f は、生成した訂正データを訂正データバッファ 550 i に記憶させる。

【0111】

CRC 検査部 550 g は、データ読出部 550 b によって読み出されたストライプに誤りデータが存在するか否かを検査する処理部である。具体的には、CRC 検査部 550 g は、データ読出部 550 b によって読み出されたストライプ内の CRC を用いて、誤りデータが存在するか否かを、ストライプごとに検査する。そして、誤りデータが存在する場合、すなわち、エラーを検出した場合、CRC 検査部 550 g は、エラーを検出したストライプのストライプ番号をログ保持レジスタ 550 j に記憶させる。

【0112】

AID 検査部 550 h は、データ読出部 550 b によって読み出されたストライプがすべて揃っているか否かを検査する処理部である。具体的には、AID 検査部 550 h は、データ読出部 550 b によって読み出されたストライプ内の AID を用いて、小エリア内のストライプがすべて揃っているか否かを検査する。そして、ストライプがすべて揃っていないことを検出した場合、すなわち、エラーを検出した場合、AID 検査部 550 h は、欠如しているストライプのストライプ番号をログ保持レジスタ 550 j に記憶させる。

【0113】

訂正データバッファ 550 i は、データを一次記憶する記憶装置であり、上述したように、XOR 演算部 550 f によって、訂正データが記憶される。ログ保持レジスタ 550 j は、データを一次記憶する記憶装置であり、上述したように、CRC 検査部 550 g または AID 検査部 550 h によって、エラーが検出されたストライプのストライプ番号が記憶される。

【0114】

セクタ 550 k は、セクタ回路であり、復旧データバッファ 550 e と、訂正データバッファ 550 i と、ログ保持レジスタ 550 j とに記憶されている各種データに基づいて、復旧データバッファ 550 e に記憶されているストライプを、データ書込部 550 m に出力したり、誤りデータを訂正したストライプをデータ書込部 550 m に出力したりする。

【0115】

具体的には、セクタ550kは、ログ保持レジスタ550jにストライプ番号が記憶されていないストライプについては、復旧データバッファ550eに記憶されているストライプを、データ書込部550mに出力する。また、セクタ550kは、ログ保持レジスタ550jにストライプ番号が記憶されているストライプ（以下、「エラー検出ストライプ」という）については、復旧データバッファ550eおよび訂正データバッファ550iに記憶されている各種データを用いて、かかるエラー発生ストライプの誤りデータを訂正してデータ書込部550mに出力する。

【0116】

バウンダリ調整部550lは、セクタ550kから受け付けた小エリア単位のストライプからバウンダリ調整値を除去する処理部である。具体的には、バウンダリ調整部550lは、小エリア単位のストライプから、データ復旧ディスクリプタに含まれるバウンダリ調整値のサイズ分のバウンダリ調整値を除去する。

【0117】

データ書込部550mは、バウンダリ調整部550lによってバウンダリ調整値が除去された小エリア単位のストライプに含まれる退避対象単位データをキャッシュメモリ130に書き込む処理部である。なお、データ書込部550mは、パリティデータについてはキャッシュメモリ130に書き込まない。

【0118】

次に、実施例1に係るストレージ装置1によるデータ退避処理について説明する。図8は、実施例1に係るストレージ装置1によるデータ退避処理の流れを示すシーケンス図である。

【0119】

同図に示すように、R o C 1 6 0の退避データ領域算出部161は、ストレージ装置1が異常終了したことを検出した場合（ステップS101）、データ退避ディスクリプタを生成するために必要となる各種情報を算出する。

【0120】

そして、退避指示部162は、退避データ領域算出部161によって算出された各種情報に基づいて、ライトDMA541にデータ退避処理を行わせるためのデータ退避ディスクリプタを生成して、ディスクリプタ保持レジスタ530に記憶させる（ステップS102）。

【0121】

例えば、図3-1に示した例のように、キャッシュメモリ130にデータD1が記憶されている場合、退避指示部162は、ライトDMA541にデータ退避処理をさせるためのデータ退避ディスクリプタとして、データD10の先頭アドレスと、データD10のサイズ28[KB]と、バウンダリ調整値のサイズ0[KB]と、AID等を含むデータ退避ディスクリプタを生成する。

【0122】

同様に、退避指示部162は、ライトDMA542にデータ退避処理をさせるためのデータ退避ディスクリプタを生成して、ディスクリプタ保持レジスタ530に記憶させる（ステップS103）。

【0123】

上述例の場合、退避指示部162は、ライトDMA542にデータ退避処理をさせるためのデータ退避ディスクリプタとして、データD20の先頭アドレスと、データD20のサイズ25[KB]と、バウンダリ調整値のサイズ3[KB]と、AID等を含むデータ退避ディスクリプタを生成する。

【0124】

その後、退避指示部162は、ライトDMA541を起動させるとともに（ステップS104）、ライトDMA542を起動させて（ステップS105）、ライトDMA541およびライトDMA542に対して、データ退避処理を並行して行わせる。

【0125】

退避指示部 162 によって起動されたライトDMA541は、データ退避処理を行い(ステップS106)、データ退避処理が完了した場合に、データ退避処理が完了した旨をROC160に対して通知する(ステップS107)。同様に、ライトDMA542は、データ退避処理を行い(ステップS108)、データ退避処理が完了した場合に、データ退避処理が完了した旨をROC160に対して通知する(ステップS109)。ライトDMA541および542によるデータ退避処理は、並行して行われる。

【0126】

なお、上述したデータ退避処理の流れでは、ストレージ装置1が異常終了した後に、退避指示部162がデータ退避ディスクリプタを生成する例を説明したが、退避指示部162は、ストレージ装置1が異常終了する前に、データ退避ディスクリプタを生成してもよい。例えば、ストレージ装置1の起動中に、退避対象データのサイズ等を予め取得しておけば、ストレージ装置1の異常終了前にデータ退避ディスクリプタを生成することができる。

【0127】

次に、ライトDMA541および542によるデータ退避処理について説明する。図9は、ライトDMA541および542によるデータ退避処理手順を示すフローチャートである。なお、ライトDMA541および542によるデータ退避処理手順は同一であるため、ここでは、ライトDMA541によるデータ退避処理手順についてのみ説明する。

【0128】

同図に示すように、ROC160によって起動されたライトDMA541は、ディスクリプタ保持レジスタ530からデータ退避ディスクリプタを読み出す(ステップS201)。そして、ライトDMA541のデータ読出部541aは、キャッシュメモリ130から、所定のサイズのデータを読み出す(ステップS202)。このとき、データ読出部541aは、データ退避ディスクリプタに含まれている退避対象データの先頭アドレスから、退避対象データのサイズ分のデータを読み出す。

【0129】

続いて、バウンダリ調整部541bは、データ読出部541aによって読み出された退避対象データに、データ退避ディスクリプタに含まれているバウンダリ調整値のサイズ分のバウンダリ調整値を加える(ステップS203)。

【0130】

続いて、パリティ生成部541cは、退避対象データを、バウンダリサイズごとのバウンダリデータに区切って、さらに、バウンダリデータを、予めシステムで決められているサイズである2[KB]ごとに区切る。そして、パリティ生成部541cは、退避対象単位データ番号が奇数である7個の退避対象単位データを用いてパリティデータを生成するとともに、退避対象単位データ番号が偶数である7個の退避対象単位データを用いてパリティデータを生成する(ステップS204)。

【0131】

続いて、CRC付与部541gは、退避対象単位データごとにCRCを生成するとともに、パリティデータごとにCRCを生成して、生成したCRCを、退避対象単位データおよびパリティデータに付与する(ステップS205)。続いて、AID付与部541hは、退避対象単位データおよびパリティデータに、AIDを付与する(ステップS206)。そして、データ書込部541fは、CRCおよびAIDが付与された退避対象単位データおよびパリティデータを、フラッシュメモリ141に書き込む(ステップS207)。

【0132】

次に、実施例1に係るストレージ装置1によるデータ復旧処理について説明する。図10は、実施例1に係るストレージ装置1によるデータ復旧処理の流れを示すシーケンス図である。

【0133】

同図に示すように、ROC160の復旧指示部163は、異常終了後にストレージ装置1が起動したことを検出した場合(ステップS301)、リードDMA550に対して、

フラッシュメモリ 141 に記憶されているデータをキャッシュメモリ 130 へ復旧する旨のデータ復旧ディスクリプタを生成して、ディスクリプタ保持レジスタ 530 に記憶させる (ステップ S302)。その後、復旧指示部 163 は、リード DMA 550 を起動させる (ステップ S303)。

【0134】

復旧指示部 163 によって起動されたリード DMA 550 は、フラッシュメモリ 141 に記憶されているデータをキャッシュメモリ 130 へ復旧させるデータ復旧処理を行う (ステップ S304)。かかるデータ復旧処理が完了した場合に、リード DMA 550 は、データ復旧処理が完了した旨を R o C 160 に対して通知する (ステップ S305)。

【0135】

続いて、復旧指示部 163 は、リード DMA 550 に対して、フラッシュメモリ 142 に記憶されているデータをキャッシュメモリ 130 へ復旧する旨のデータ復旧ディスクリプタを生成して、ディスクリプタ保持レジスタ 530 に記憶させる (ステップ S306)。その後、復旧指示部 163 は、リード DMA 550 を起動させる (ステップ S307)。

。

【0136】

復旧指示部 163 によって起動されたリード DMA 550 は、フラッシュメモリ 142 に記憶されているデータをキャッシュメモリ 130 へ復旧させるデータ復旧処理を行う (ステップ S308)。かかるデータ復旧処理が完了した場合に、リード DMA 550 は、データ復旧処理が完了した旨を R o C 160 に対して通知する (ステップ S309)。

【0137】

続いて、R o C 160 は、キャッシュメモリ 130 に復旧されたデータを、HDD 11 に記憶させるリストア処理 (フラッシュバック) を行う (ステップ S310)。これにより、HDD 11 に書き込みされていないキャッシュメモリ 130 上のデータはなくなる。

【0138】

続いて、復旧指示部 163 は、イレーズディスクリプタを生成して、データ転送 DMA 150 に記憶させる (ステップ S311)。そして、復旧指示部 163 は、イレーズ DMA 561 および 562 を起動させて (ステップ S312)、イレーズ DMA 561 および 562 に対して、フラッシュメモリ 141 および 142 に記憶されているデータを消去させる。

【0139】

復旧指示部 163 によって起動されたイレーズ DMA 561 および 562 は、ディスクリプタ保持レジスタ 530 から、イレーズディスクリプタを読み出して (ステップ S313)、イレーズ処理を行う (ステップ S314)。かかるイレーズ処理が完了した場合に、イレーズ DMA 561 および 562 は、イレーズ処理が完了した旨を R o C 160 に対して通知する (ステップ S315)。そして、イレーズ処理が完了した旨の通知を受け付けた R o C 160 は、ストレージ装置 1 を起動させる (ステップ S316)。

【0140】

次に、リード DMA 550 によるデータ復旧処理について説明する。図 11 は、リード DMA 550 によるデータ復旧処理手順を示すフローチャートである。同図に示すように、R o C 160 によって起動されたリード DMA 550 は、ディスクリプタ保持レジスタ 530 からデータ復旧ディスクリプタを読み出す (ステップ S401)。

【0141】

そして、リード DMA 550 のデータ読出部 550 b は、セクタ 550 a を介して、フラッシュメモリ 141 から、小エリアのストライプを読み出す (ステップ S402)。このデータ読出部 550 b によるデータ読出処理が所定の時間内に完了しない場合 (ステップ S403 否定)、タイムアウト検出部 550 c は、データ読出部 550 b によって読取処理されているストライプに対して、タイムアウトが発生したことを検出する。そして、タイムアウト検出部 550 c は、タイムアウト検出通知をセクタ 550 a に対して行う。

【0142】

タイムアウト検出通知を受け付けたセクタ550aは、タイムアウト検出通知に含まれるストライプ番号が示すストライプの読出処理を中止して、かかるストライプの読み出しできなかったサイズ分の付加データをデータ読出部550bに出力する(ステップS404)。このようにして、データ読出部550bは、フラッシュメモリ141から小エリアのストライプを読み出す。

【0143】

続いて、XOR演算部550fは、データ読出部550bによって読み出されたストライプを用いて訂正データを生成する(ステップS405)。続いて、CRC検査部550gは、データ読出部550bによって読み出されたストライプに誤りデータが存在するかどうかを検査し、AID検査部550hは、データ読出部550bによって読み出されたストライプがすべて揃っているかどうかを検査する(ステップS406)。

【0144】

CRC検査部550g、または、AID検査部によってエラーが検出された場合(ステップS407肯定)、セクタ550kは、データ読出部550bによって読み出されたストライプ、および、XOR演算部550fによって生成された訂正データに基づいて、エラーが検出されたストライプの誤りデータを訂正する(ステップS408)。

【0145】

続いて、バウンダリ調整部550lは、データ読出部550bによって読み出された小エリアのストライプから、データ復旧ディスクリプタに含まれるバウンダリ調整値のサイズ分のバウンダリ調整値を除去する(ステップS409)。

【0146】

そして、データ書込部550mは、バウンダリ調整部550lによってバウンダリ調整値が除去された小エリアのストライプに含まれる退避対象単位データを、キャッシュメモリ130に書き込む(ステップS410)。なお、このとき、データ書込部550mは、パリティデータについてはキャッシュメモリ130に書き込まない。

【0147】

リードDMA550は、フラッシュメモリ141に記憶されている全てのデータを読み出していない場合(ステップS411否定)、フラッシュメモリ141から全てのデータを読み出すまで、上述した処理手順(ステップS401~S410)を繰り返し行う。

【0148】

上述してきたように、実施例1に係るストレージ装置1は、データ退避処理時に、退避対象データのパリティデータを生成して、CRCやAIDを付与した退避対象データおよびパリティデータをフラッシュメモリ141または142に書き込み、データ復旧処理において、フラッシュメモリ141または142から所定の時間内にデータ読出処理が完了しない場合に、データ読出処理を中止して付加データを設定し、パリティデータを用いて誤りデータを訂正したデータをキャッシュメモリ130に書き込むように構成したので、フラッシュメモリ141または142からデータの読出応答がなされない場合であっても、フラッシュメモリ141または142に退避させたデータを、キャッシュメモリ130に復旧させることができる。

【0149】

なお、上記実施例1では、データ退避処理時に、ライトDMA541および542が、7個の退避対象単位データに対して、1個のパリティデータを生成する例を示したが、退避対象単位データの数と、パリティデータの数の比率はこれに限らない。例えば、ライトDMA541および542は、2個退避対象単位データに対して、1個のパリティデータを生成してもよいし、10個退避対象単位データに対して、1個のパリティデータを生成してもよい。

【0150】

また、上記実施例1では、ライトDMA541および542が、退避対象単位データ番号が奇数である退避対象単位データを用いてパリティデータを生成するとともに、退避対

象単位データ番号が偶数である退避対象単位データを用いてパリティデータを生成する例を示したが、ライトDMA541および542は、退避対象単位データ番号に関わらず、所定の数の退避対象単位データを用いてパリティデータを構成してもよい。

【実施例2】

【0151】

ところで、上記実施例1では、データサイズがバウンダリサイズに満たないバウンダリデータに、バウンダリ調整値を加える処理であるバウンダリ処理を行う例を示したが、バウンダリ処理を行わないようにしてもよい。そこで、実施例2では、キャッシュメモリ130から読み出したデータをバウンダリサイズごとに区切れない場合であっても、バウンダリ処理を行わない例について説明する。

【0152】

まず、実施例2に係るストレージ装置2によるデータ退避処理の概要を説明する。なお、実施例2に係るストレージ装置2の概略構成は、図1に示したストレージ装置1の概略構成と同様であるため、その説明を省略する。

【0153】

図12-1は、実施例2に係るストレージ装置2によるデータ退避処理の概要を説明するための図である。同図に示すように、キャッシュメモリ130には、210[KB]のデータD3が記憶されているものとする。この状態で、ストレージ装置2が異常終了した場合、データD3は、バウンダリサイズ28[KB]ごとにデータD31~D38に区切られる。

【0154】

ここで、実施例2に係るストレージ装置2のROC(ROC260とする)は、ライトDMA541に対して、キャッシュメモリ130に記憶されているデータD3のうち、データD4(データD31~D34)について、データ退避処理を行うように指示する。

【0155】

さらに、ROC260は、ライトDMA542に対して、キャッシュメモリ130に記憶されているデータD3のうち、データD5(データD34-2と、データD35~D38)について、データ退避処理を行うように指示する。

【0156】

つまり、ROC260は、ライトDMA541に対して、データD3のキャッシュメモリ130における先頭の記憶位置である先頭アドレスから4個分のバウンダリデータ(同図に示したデータD31、D32、D33、D34)について、データ退避処理を行うように指示する。また、ROC260は、ライトDMA542に対して、データD3のキャッシュメモリ130における最後の記憶位置である最終アドレスから、先頭アドレス方向に4個分のバウンダリデータ(同図に示したデータD54、D53、D52、D51)について、データ退避処理を行うように指示する。

【0157】

このような指示を受け付けたライトDMA541は、データD4を、バウンダリサイズ28[KB]ごとにデータD31~D34に区切ることができる。また、ライトDMA542は、データD5を、バウンダリサイズ28[KB]ごとにデータD51~D54に区切ることができる。すなわち、実施例2に係るストレージ装置2は、バウンダリ処理を行うことなく、データ退避処理を行うことができる。

【0158】

次に、実施例2に係るストレージ装置2によるデータ復旧処理の概要について説明する。図12-2は、実施例2に係るストレージ装置2によるデータ復旧処理の概要を説明するための図である。図12-2に示したフラッシュメモリ141および142には、図12-1に示したライトDMA541およびライトDMA542によって退避されたデータが記憶されているものとする。

【0159】

ストレージ装置2が異常終了後に起動した場合、リードDMA550は、フラッシュメモ

メモリ141からデータD31を読み出しながら各種検査(CRC検査等)を行った後、キャッシュメモリ130に書き込む。また、リードDMA550は、同様の処理を、データD32~D34についても行い、キャッシュメモリ130にデータD4を復旧させる。

【0160】

その後、リードDMA550は、フラッシュメモリ142からデータD51を読み出しながら各種検査(CRC検査等)を行った後、キャッシュメモリ130に書き込む。このとき、キャッシュメモリ130には、フラッシュメモリ141から読み出したデータD34-2が書き込み済みであるので、リードDMA550は、データD51を書き込む場合、データD34-2の部分については上書きすることとなる。

【0161】

そして、リードDMA550は、フラッシュメモリ142からデータD52~D54を順に読み出して、キャッシュメモリ130にデータD5を復旧させる。リードDMA550が、フラッシュメモリ141および142に記憶されているデータをすべてキャッシュメモリ130に書き込むと、キャッシュメモリ130にデータD3を復旧することができる。

【0162】

このように、実施例2に係るストレージ装置2は、2系統のライトDMA541および542のうち、一方のライトDMA541に対して、退避対象データのキャッシュメモリ130における先頭の記憶位置である先頭アドレスから所定の数のバウンダリデータについて、データ退避処理を行わせるとともに、他方のライトDMA542に対して、退避対象データのキャッシュメモリ130における最後の記憶位置である最終アドレスから、先頭アドレス方向に所定の数のバウンダリデータについて、データ退避処理を行わせるので、DMA541および542が、フラッシュメモリ141または142に退避するデータを、バウンダリデータで割り切れることができ、その結果、バウンダリ処理を行うことなく、データ退避処理を行うことができる。

【0163】

次に、実施例2に係るストレージ装置2が有するCM200の構成について説明する。図13は、実施例2に係るストレージ装置2が有するCM200の構成を示す図である。ここでは、図4に示した構成部位と同様の機能を有する部位には同一符号を付すこととして、その詳細な説明を省略する。

【0164】

同図に示した退避データ領域算出部261は、ストレージ装置2が異常終了した場合に、キャッシュメモリ130に記憶されている退避対象データのサイズ等に基づいて、まず、ライトDMA541が退避するデータの開始アドレスBA1およびサイズDS1を算出し、次に、ライトDMA542が退避するデータの開始アドレスBA2およびサイズDS2を算出する。

【0165】

以下に、図12-1に示した例を用いつつ、退避データ領域算出部261による退避データ領域算出処理について、より具体的に説明する。まず、退避データ領域算出部261は、ストレージ装置2が異常終了した場合に、退避対象データのキャッシュメモリ130における記憶位置を示す先頭アドレスおよび最終アドレスを取得する。そして、退避データ領域算出部261は、取得した先頭アドレスを、ライトDMA541が退避するデータの開始アドレスBA1に決定する。

【0166】

続いて、退避データ領域算出部261は、退避対象データのサイズを2で除算したサイズ「X」を求める。図12-1に示した例では、退避対象データのサイズが210[KB]であるので、退避データ領域算出部261は、サイズXとして、210[KB]を2で除算した105[KB]を求める。

【0167】

続いて、退避データ領域算出部261は、式(1)「 $28[KB] \times N > \text{サイ}$

ズ $X [KB]$ 」を満たす整数 N の最小値に $28 [KB]$ を乗算した値を、ライトDMA541が退避するデータのサイズ $DS1$ に決定する。上記例では、サイズ X が $105 [KB]$ であり、式(1)を満たす N の最小値は「4」となるので、退避データ領域算出部261は、4に $28 [KB]$ を乗算した値である $112 [KB]$ を、ライトDMA541が退避するデータのサイズ $DS1$ に決定する。

【0168】

続いて、退避データ領域算出部261は、退避対象データの先頭アドレスに、ライトDMA541が退避するデータサイズ $DS1$ を加算した値「 Y 」を求める。この Y は、ライトDMA541が退避するデータの最終アドレスに該当する。そして、退避データ領域算出部261は、退避対象データの最終アドレスから Y を減算した値「 Z 」を求める。この Z は、退避対象データのサイズから、ライトDMA541が退避するデータのサイズ $DS1$ を減算したサイズに該当する。

【0169】

すなわち、上記例では、退避対象データのサイズが $210 [KB]$ であり、サイズ $DS1$ が $112 [KB]$ であるので、退避データ領域算出部261は、サイズ Z として、 $210 [KB]$ から $112 [KB]$ を減算した $98 [KB]$ を求める。

【0170】

続いて、退避データ領域算出部261は、式(2)「 $28 [KB] \times N > \text{サイズ} Z [KB]$ 」を満たす整数 N の最小値に $28 [KB]$ を乗算した値を、ライトDMA542が退避するデータのサイズ $DS2$ に決定する。上記例では、サイズ Z が $98 [KB]$ であり、式(2)を満たす N の最小値は「4」となるので、退避データ領域算出部261は、4に $28 [KB]$ を乗算した値である $112 [KB]$ を、ライトDMA542が退避するデータのサイズ $DS2$ に決定する。

【0171】

続いて、退避データ領域算出部261は、退避対象データの最終アドレスから、ライトDMA542が退避するデータのサイズ $DS2$ を減算した値を、ライトDMA542が退避するデータの先頭アドレス $BA2$ に決定する。

【0172】

このようにして、退避データ領域算出部261は、ライトDMA541が退避するデータの開始アドレス $BA1$ およびサイズ $DS1$ と、ライトDMA542が退避するデータの開始アドレス $BA2$ およびサイズ $DS2$ を算出する。

【0173】

図13に示した退避指示部262は、データ退避ディスクリプタを生成して、データ転送DMA150に送信することで、ライトDMA541および542に対して、データ退避処理を行うように指示する。

【0174】

具体的には、退避指示部262は、上述した開始アドレス $BA1$ およびサイズ $DS1$ を含むデータ退避ディスクリプタを生成して、かかるデータ退避ディスクリプタに基づいてデータ退避処理を行うようにライトDMA541に対して指示をする。また、退避指示部262は、上述した開始アドレス $BA2$ およびサイズ $DS2$ を含むデータ退避ディスクリプタを生成して、かかるデータ退避ディスクリプタに基づいてデータ退避処理を行うようにライトDMA542に対して指示をする。なお、退避指示部262が生成するデータ退避ディスクリプタには、バウンダリ調整値のサイズを含まない。

【0175】

次に、図13に示した退避データ領域算出部261による退避データ領域算出処理について説明する。図14は、図13に示した退避データ領域算出部261による退避データ領域算出処理手順を示すフローチャートである。

【0176】

図14に示すように、退避データ領域算出部261は、退避対象データのキャッシュメモリ130における記憶位置を示す先頭アドレスおよび最終アドレスを取得する(ステッ

プ S 5 0 1)。

【 0 1 7 7 】

そして、退避データ領域算出部 2 6 1 は、取得した先頭アドレスを、ライト D M A 5 4 1 が退避するデータの開始アドレスに決定する (ステップ S 5 0 2)。続いて、退避データ領域算出部 2 6 1 は、退避対象データのサイズを 2 で除算したサイズ X を求める (ステップ S 5 0 3)。

【 0 1 7 8 】

続いて、退避データ領域算出部 2 6 1 は、式 (1) 「 $28 [KB] \times N > \text{サイズ} X [KB]$ 」を満たす整数 N の最小値を求める (ステップ S 5 0 4)。以下では、ここで求めた最小値 N を「第一の基準値」と呼ぶこととする。そして、退避データ領域算出部 2 6 1 は、第一の基準値に $28 [KB]$ を乗算した値を、ライト D M A 5 4 1 が退避するデータのサイズに決定する (ステップ S 5 0 5)。

【 0 1 7 9 】

続いて、退避データ領域算出部 2 6 1 は、退避対象データのサイズから、ライト D M A 5 4 1 が退避するデータのサイズを減算したサイズ Z を求める (ステップ S 5 0 6)。続いて、退避データ領域算出部 2 6 1 は、式 (2) 「 $28 [KB] \times N > \text{サイズ} Z [KB]$ 」を満たす整数 N の最小値を求める (ステップ S 5 0 7)。以下では、ここで求めた最小値 N を「第二の基準値」と呼ぶこととする。そして、退避データ領域算出部 2 6 1 は、第二の基準値に $28 [KB]$ を乗算した値を、ライト D M A 5 4 2 が退避するデータのサイズに決定する (ステップ S 5 0 8)。

【 0 1 8 0 】

続いて、退避データ領域算出部 2 6 1 は、退避対象データの最終アドレスから、ライト D M A 5 4 2 が退避するデータのサイズを減算した値を求め (ステップ S 5 0 9)、求めた値を、ライト D M A 5 4 2 が退避するデータの先頭アドレスに決定する (ステップ S 5 1 0)。

【 0 1 8 1 】

上述してきたように、実施例 2 に係るストレージ装置 2 は、退避データ領域算出部 2 6 1 が、ライト D M A 5 4 1 および 5 4 2 が退避するデータの開始アドレスおよびサイズを求めるように構成したので、D M A 5 4 1 および 5 4 2 は、フラッシュメモリ 1 4 1 または 1 4 2 に退避するデータを、バウンダリデータで割り切れることができ、その結果、フラッシュメモリ 1 4 1 または 1 4 2 からデータの読出応答がなされない場合であっても、バウンダリ処理を行うことなく、フラッシュメモリ 1 4 1 または 1 4 2 に退避させたデータを、キャッシュメモリ 1 3 0 に復旧させることができる。

【 0 1 8 2 】

また、実施例 2 に係るストレージ装置 2 では、バウンダリ処理を行う必要がないので、退避データ領域算出部 2 6 1 がバウンダリ調整値のサイズを算出する必要がなくなる。その結果、退避データ領域算出部 2 6 1 にかかる処理負荷を軽減することができ、データ退避処理の高速化を図ることができる。すなわち、異常終了時に、電力供給部 1 2 から電力供給される短い時間内に、大容量のデータについて、データ退避処理を行うことができる。

【 0 1 8 3 】

また、実施例 2 に係るストレージ装置 2 では、バウンダリ処理を行う必要がないので、ライト D M A 5 4 1 および 5 4 2 と、リード D M A 5 5 0 から、バウンダリ処理を行うための回路を削除することができ、ライト D M A 5 4 1 および 5 4 2 と、リード D M A 5 5 0 の規模を小さくすることができる。例えば、実施例 2 に係るストレージ装置 2 では、図 6 に示したライト D M A 5 4 1 からバウンダリ調整部 5 4 1 b を削除することができ、図 7 に示したリード D M A 5 5 0 からバウンダリ調整部 5 5 0 1 を削除することができる。その結果、ライト D M A 5 4 1 および 5 4 2 と、リード D M A 5 5 0 にかかるコストを軽減することができる。

【 0 1 8 4 】

なお、上記実施例 2 では、ストレージ装置 2 が 2 系統のライト D M A 5 4 1 および 5 4 2 を有することを前提として、退避データ領域算出部 2 6 1 が、ライト D M A 5 4 1 および 5 4 2 が退避するデータの開始アドレスおよびサイズを算出する例を示したが、3 系統以上のライト D M A を有するストレージ装置においても、バウンダリ処理を行わないように、退避データ領域算出部 2 6 1 が、各ライト D M A が退避するデータの開始アドレスおよびサイズを算出してもよい。例えば、ストレージ装置が 4 個のライト D M A を有する場合、退避データ領域算出部 2 6 1 は、4 個目のライト D M A に対して、退避対象データの最終アドレスから、先頭アドレス方向に所定の数のバウンダリデータについて、データ退避処理を行わせるように、4 個目のライト D M A が退避するデータの開始アドレスおよびサイズを算出する。

【 0 1 8 5 】

また、上記実施例 1 および 2 では、バウンダリサイズを 2 8 [K B] とし、バウンダリデータを区切るサイズを 2 [K B] とし、エリア内の p a g e の数を 1 0 2 4 個とし、小エリア内の p a g e の数を 1 6 個とする例を示したが、これらの規定値は、上記例に限られない。例えば、バウンダリサイズを 5 6 [K B] とし、バウンダリデータを区切るサイズを 4 [K B] とし、エリア内の p a g e の数を 2 0 4 8 個とし、小エリア内の p a g e の数を 3 2 個としてもよい。

【 0 1 8 6 】

また、上記実施例 1 および 2 では、異常終了時にデータを退避する記憶領域として、O n e N A N D 型のフラッシュメモリを用いる例を示したが、異常終了時にデータを退避する記憶領域は、O n e N A N D 型のフラッシュメモリ以外の不揮発性メモリであってもよい。

【 0 1 8 7 】

また、上記文書中や図面中で示した処理手順、制御手順、具体的名称、各種のデータやパラメータを含む情報については、特記する場合を除いて任意に変更することができる。また、図示した各装置の各構成要素は機能概念的なものであり、必ずしも物理的に図示の如く構成されていることを要しない。すなわち、各装置の分散、統合の具体的形態は図示のものに限られず、その全部または一部を、各種の負荷や使用状況などに応じて、任意の単位で機能的または物理的に分散、統合して構成することができる。さらに、各装置にて行なわれる各処理機能は、その全部または任意の一部が、C P U (Central Processing Unit) および当該 C P U にて解析実行されるプログラムにて実現され、あるいは、ワイヤードロジックによるハードウェアとして実現され得る。

【 符号の説明 】

【 0 1 8 8 】

1、2	ストレージ装置
1 0	C E
1 1	H D D
1 2	電力供給部
2 0	D E
1 0 0、2 0 0	C M
1 1 0	I / F 部
1 2 0、2 2 0	エクスパンダ
1 3 0	キャッシュメモリ
1 4 1、1 4 2	フラッシュメモリ
1 5 0	データ転送 D M A
1 6 0、2 6 0	R o C
1 6 1、2 6 1	退避データ領域算出部
1 6 2、2 6 2	退避指示部
1 6 3	復旧指示部
5 1 0	P C I バス I / F

5 2 1、5 2 2	メモリ I / F
5 3 0	ディスクリプタ保持レジスタ
5 4 1、5 4 2	ライト D M A
5 4 1 a	データ読出部
5 4 1 b	バウンダリ調整部
5 4 1 c	パリティ生成部
5 4 1 d	退避データバッファ
5 4 1 e	パリティバッファ
5 4 1 f	データ書込部
5 4 1 g	C R C 付与部
5 4 1 h	A I D 付与部
5 5 0	リード D M A
5 5 0 a	セレクタ
5 5 0 b	データ読出部
5 5 0 c	タイムアウト検出部
5 5 0 d	エラー検出部
5 5 0 e	復旧データバッファ
5 5 0 f	X O R 演算部
5 5 0 g	C R C 検査部
5 5 0 h	A I D 検査部
5 5 0 i	訂正データバッファ
5 5 0 j	ログ保持レジスタ
5 5 0 k	セレクタ
5 5 0 l	バウンダリ調整部
5 5 0 m	データ書込部
5 6 1、5 6 2	イレース D M A

【 国際調査報告 】

INTERNATIONAL SEARCH REPORT		International application No. PCT/JP2008/053494
A. CLASSIFICATION OF SUBJECT MATTER G06F12/16(2006.01)i, G06F3/06(2006.01)i, G06F12/08(2006.01)i According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) G06F12/16, G06F3/06, G06F12/08 Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Jitsuyo Shinan Koho 1922-1996 Jitsuyo Shinan Toroku Koho 1996-2008 Kokai Jitsuyo Shinan Koho 1971-2008 Toroku Jitsuyo Shinan Koho 1994-2008 Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2004-531814 A (International Business Machines Corp.), 14 October, 2004 (14.10.04), Par. Nos. [0017] to [0020] & WO 2002/086721 A1 & US 2002/0156983 A1 & GB 2391095 A	1-17
A	JP 07-160594 A (Hitachi, Ltd., Hitachi Video and Information System, Inc.), 23 June, 1995 (23.06.95), Par. Nos. [0030] to [0033] (Family: none)	1-17
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/> See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search 11 July, 2008 (11.07.08)		Date of mailing of the international search report 22 July, 2008 (22.07.08)
Name and mailing address of the ISA/ Japanese Patent Office		Authorized officer
Facsimile No.		Telephone No.

国際調査報告		国際出願番号 PCT/JP2008/053494									
A. 発明の属する分野の分類 (国際特許分類 (IPC)) Int.Cl. G06F12/16(2006.01)i, G06F3/06(2006.01)i, G06F12/08(2006.01)i											
B. 調査を行った分野 調査を行った最小限資料 (国際特許分類 (IPC)) Int.Cl. G06F12/16, G06F3/06, G06F12/08											
最小限資料以外の資料で調査を行った分野に含まれるもの <table border="0"> <tr> <td>日本国実用新案公報</td> <td>1922-1996年</td> </tr> <tr> <td>日本国公開実用新案公報</td> <td>1971-2008年</td> </tr> <tr> <td>日本国実用新案登録公報</td> <td>1996-2008年</td> </tr> <tr> <td>日本国登録実用新案公報</td> <td>1994-2008年</td> </tr> </table>				日本国実用新案公報	1922-1996年	日本国公開実用新案公報	1971-2008年	日本国実用新案登録公報	1996-2008年	日本国登録実用新案公報	1994-2008年
日本国実用新案公報	1922-1996年										
日本国公開実用新案公報	1971-2008年										
日本国実用新案登録公報	1996-2008年										
日本国登録実用新案公報	1994-2008年										
国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)											
C. 関連すると認められる文献											
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号									
A	JP 2004-531814 A (インターナショナル・ビジネス・マシーンズ・コーポレーション) 2004.10.14, 段落【0017】～【0020】 & WO 2002/086721 A1 & US 2002/0156983 A1 & GB 2391095 A	1-17									
A	JP 07-160594 A (株式会社日立製作所, 株式会社日立画像情報システム) 1995.06.23, 段落【0030】～【0033】 (ファミリーなし)	1-17									
<input type="checkbox"/> C欄の続きにも文献が列挙されている。 <input type="checkbox"/> パテントファミリーに関する別紙を参照。											
* 引用文献のカテゴリー		の日の後に公表された文献									
「A」特に関連のある文献ではなく、一般的技術水準を示すもの		「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの									
「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの		「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの									
「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)		「Y」特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの									
「O」口頭による開示、使用、展示等に言及する文献		「&」同一パテントファミリー文献									
「P」国際出願日前で、かつ優先権の主張の基礎となる出願											
国際調査を完了した日 11.07.2008		国際調査報告の発送日 22.07.2008									
国際調査機関の名称及びあて先 日本国特許庁 (ISA/JP) 郵便番号100-8915 東京都千代田区霞が関三丁目4番3号		特許庁審査官 (権限のある職員) 堀江 義隆	5N 9172								
		電話番号 03-3581-1101	内線 3586								

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 3/06 3 0 2 A

(72)発明者 花岡 祐司

神奈川県川崎市中原区上小田中4丁目1番1号 株式会社富士通コンピュータテクノロジーズ内

Fターム(参考) 5B005 JJ01 MM11 UU23 WW16 WW17

5B018 GA02 GA04 HA11 KA03 KA21 KA22 MA03 MA14 NA06 QA06

5B065 BA01 CA30 CH01 EA03 EA23 EA26 EA27

(注)この公表は、国際事務局(WIPO)により国際公開された公報を基に作成したものである。なおこの公表に係る日本語特許出願(日本語実用新案登録出願)の国際公開の効果は、特許法第184条の10第1項(実用新案法第48条の13第2項)により生ずるものであり、本掲載とは関係ありません。