(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号

特許第6005566号
(P6005566)

(45)発行日　平成28年10月12日(2016.10.12)　(24)登録日　平成28年9月16日(2016.9.16)

(51)Int.Cl.		F I		
G06F 12/16	(2006.01)	G06F	12/16	310Q
G06F 13/10	(2006.01)	G06F	13/10	340A
G06F 3/06	(2006.01)	G06F	3/06	301K

請求項の数 56　(全 105 頁)

(21)出願番号　特願2013-55548(P2013-55548)
(22)出願日　平成25年3月18日(2013.3.18)
(65)公開番号　特開2014-182503(P2014-182503A)
(43)公開日　平成26年9月29日(2014.9.29)
審査請求日　平成27年2月10日(2015.2.10)

(73)特許権者　000003078
株式会社東芝
東京都港区芝浦一丁目１番１号
(74)代理人　110002147
特許業務法人酒井国際特許事務所
(72)発明者　橋本　大輔
東京都港区芝浦一丁目１番１号　株式会社東芝内
(72)発明者　永井　宏一
東京都港区芝浦一丁目１番１号　株式会社東芝内
(72)発明者　渡部　孝紀
東京都港区芝浦一丁目１番１号　株式会社東芝内

最終頁に続く

(54)【発明の名称】情報処理システム、制御プログラムおよび情報処理装置

(57)【特許請求の範囲】
【請求項１】
リードおよびライト可能な第１のメモリを有する第１の記憶部と、引き継ぎ履歴格納部が記憶され、リードおよびライト可能な第２のメモリを有する第２の記憶部とを接続可能な情報処理装置であって、
前記第１の記憶部から取得した信頼性情報に基づき、前記第１の記憶部のディスクステータスを判定する第１の処理と、
前記第１の処理により前記第１の記憶部のディスクステータスが保護状態として認識された場合、書き込みは前記第２の記憶部に対し実行し、読み出しは前記引き継ぎ履歴格納部に読み出し対象アドレスが記録されている場合は、前記第２の記憶部からデータを読み出し、前記引き継ぎ履歴格納部に読み出し対象アドレスが記録されていない場合は、前記第１の記憶部からデータを読み出す第２の処理と、
前記第２の処理による書き込みに伴って、前記第２の記憶部に書き込まれたデータのアドレスを前記引き継ぎ履歴格納部に記録する第３の処理と、
前記第１の処理により前記第１の記憶部のディスクステータスが保護状態として認識されなかった場合、読み出しは前記引き継ぎ履歴格納部を読み出すこと無く前記第１の記憶部からデータを読み出す第４の処理と、
を実行する制御部
を備えることを特徴とする情報処理装置。
【請求項２】

前記制御部は、
前記第１の処理では、前記信頼性情報と閾値との比較を行い、この比較に基づいて前記第１の記憶部の寿命到達条件が成立したと判定したときに、前記第１の記憶部のディスクステータスを前記保護状態であると判定することを特徴とする請求項１に記載の情報処理装置。
【請求項３】
前記制御部は、
前記第１の処理では、前記信頼性情報と閾値との比較を行い、この比較に基づいて前記第１の記憶部の寿命到達条件が成立したと判定したときに、前記第１の記憶部のディスクステータスを信頼性劣化状態であると判定し、
前記第１の処理により前記第１の記憶部のディスクステータスが前記信頼性劣化状態として判定された場合、前記第２の記憶部を前記第１の記憶部に格納されているデータのデータ引き継ぎ先状態として認識し、前記第１の記憶部のディスクステータスを前記保護状態に変更することを特徴とする請求項１に記載の情報処理装置。
【請求項４】
リードおよびライト可能な第３のメモリを有する１乃至複数の第３の記憶部に接続可能であり、
前記制御部は、
同一アドレスのデータが前記第１および第２の記憶部のうちの一方に存在するように、前記第１および前記第２の記憶部での記憶データのアドレスを管理し、
前記第１の記憶部のディスクステータスが保護状態として認識された場合、前記第１の記憶部と前記第２の記憶部を同一の論理単位として認識し、
前記第１の記憶部のディスクステータスが保護状態として認識された場合、前記１乃至複数の第３の記憶部と前記論理単位とがＲＡＩＤを構成するように制御することを特徴とする請求項１に記載の情報処理装置。
【請求項５】
前記制御部は、
前記第１の記憶部に書き込まれているデータのうち前記第２の記憶部へ引き継ぎ対象とするデータを選び、前記引き継ぎ対象のデータに対応するアドレスが全て前記引き継ぎ履歴格納部に格納されている場合に、前記第１の記憶部の保護状態を終了すると判定することを特徴とする請求項１乃至４の何れか一つに記載の情報処理装置。
【請求項６】
前記制御部は、
前記第３の処理では、
前記引き継ぎ履歴格納部に第１のアドレスを記録する場合に、前記第１のアドレスと重複または連続する第２のアドレスがすでに前記引き継ぎ履歴格納部に格納されている場合には、前記第１のアドレスと前記第２のアドレスを合併した第３のアドレスを前記引き継ぎ履歴格納部に記録し、前記第２のアドレスを無効化することを特徴とする請求項１乃至５の何れか一つに記載の情報処理装置。
【請求項７】
前記制御部は、前記第１の処理により前記第１の記憶部のディスクステータスが保護状態として認識されなかった場合、前記引き継ぎ履歴格納部を消去することを特徴とする請求項１乃至６の何れか一つに記載の情報処理装置。
【請求項８】
リードおよびライト可能な不揮発性の半導体メモリである第１のメモリを有する第１の記憶部と、引き継ぎ履歴格納部が記憶され、リードおよびライト可能な不揮発性の半導体メモリである第２のメモリを有する第２の記憶部と、引き継ぎ履歴格納部とを接続可能な情報処理装置であって、
前記第１の記憶部から取得した信頼性情報に基づき、前記第１の記憶部のディスクステータスを判定する第１の処理と、

前記第１の処理により前記第１の記憶部のディスクステータスが保護状態として認識された場合、書き込みは前記第２の記憶部に対し実行し、読み出しは前記引き継ぎ履歴格納部に読み出し対象アドレスが記録されている場合は、前記第２の記憶部からデータを読み出し、前記引き継ぎ履歴格納部に読み出し対象アドレスが記録されていない場合は、前記第１の記憶部からデータを読み出す第２の処理と、

前記第２の処理による書き込みに伴って、前記第２の記憶部に書き込まれたデータのアドレスを前記引き継ぎ履歴格納部に記録する第３の処理と、

前記第１の処理により前記第１の記憶部のディスクステータスが保護状態として認識されなかった場合、読み出しは前記引き継ぎ履歴格納部を読み出すこと無く前記第１の記憶部からデータを読み出す第４の処理と、

を実行する制御部

を備えることを特徴とする情報処理装置。

【請求項９】

前記制御部は、

前記第１の処理では、前記信頼性情報と閾値との比較を行い、この比較に基づいて前記第１の記憶部の寿命到達条件が成立したと判定したときに、前記第１の記憶部のディスクステータスを前記保護状態であると判定することを特徴とする請求項８に記載の情報処理装置。

【請求項１０】

前記制御部は、

前記第１の処理では、前記信頼性情報と閾値との比較を行い、この比較に基づいて前記第１の記憶部の寿命到達条件が成立したと判定したときに、前記第１の記憶部のディスクステータスを信頼性劣化状態であると判定し、

前記第１の処理により前記第１の記憶部のディスクステータスが前記信頼性劣化状態として判定された場合、前記第２の記憶部を前記第１の記憶部に格納されているデータのデータ引き継ぎ先状態として認識し、前記第１の記憶部のディスクステータスを前記保護状態に変更することを特徴とする請求項８に記載の情報処理装置。

【請求項１１】

リードおよびライト可能な第３のメモリを有する１乃至複数の第３の記憶部に接続可能であり、

前記制御部は、

同一アドレスのデータが前記第１および第２の記憶部のうちの一方に存在するように、前記第１および前記第２の記憶部での記憶データのアドレスを管理し、

前記第１の記憶部のディスクステータスが保護状態として認識された場合、前記第１の記憶部と前記第２の記憶部を同一の論理単位として認識し、

前記第１の記憶部のディスクステータスが保護状態として認識された場合、前記１乃至複数の第３の記憶部と前記論理単位とがＲＡＩＤを構成するように制御することを特徴とする請求項８に記載の情報処理装置。

【請求項１２】

前記制御部は、

前記第１の記憶部に書き込まれているデータのうち前記第２の記憶部へ引き継ぎ対象とするデータを選び、前記引き継ぎ対象のデータに対応するアドレスが全て前記引き継ぎ履歴格納部に格納されている場合に、前記第１の記憶部の保護状態を終了すると判定することを特徴とする請求項８乃至１１の何れか一つに記載の情報処理装置。

【請求項１３】

前記制御部は、

前記第３の処理では、

前記引き継ぎ履歴格納部に第１のアドレスを記録する場合に、前記第１のアドレスと重複または連続する第２のアドレスがすでに前記引き継ぎ履歴格納部に格納されている場合には、前記第１のアドレスと前記第２のアドレスを合併した第３のアドレスを前記引き継

ぎ履歴格納部に記録し、前記第２のアドレスを無効化することを特徴とする請求項８乃至１２の何れか一つに記載の情報処理装置。

【請求項１４】

前記制御部は、前記第１の処理により前記第１の記憶部のディスクステータスが保護状態として認識されなかった場合、前記引き継ぎ履歴格納部を消去することを特徴とする請求項８乃至１３の何れか一つに記載の情報処理装置。

【請求項１５】

リードおよびライト可能な第１のメモリを有する第１の記憶部と、リードおよびライト可能な第２のメモリを有する第２の記憶部とを接続可能な情報処理装置であって、

前記第１の記憶部から取得した信頼性情報に基づき、前記第１の記憶部のディスクステータスを判定する第１の処理と、

前記第１の処理により前記第１の記憶部のディスクステータスが前記保護状態でないと判定された場合、前記第１の記憶部に書き込みを実行し、前記第１の処理により前記第１の記憶部のディスクステータスが保護状態として認識された場合、前記第２の記憶部に書き込みを実行する第２の処理と、

前記第１の処理により前記第１の記憶部のディスクステータスが保護状態として認識された後に前記第１の記憶部に対する読み出しが発生した場合は、前記第１の記憶部からデータを読み出し、前記読み出されたデータを前記第２の記憶部に書き込む第３の処理と、

を実行する制御部

を備えることを特徴とする情報処理装置。

【請求項１６】

前記制御部は、

前記第３の処理では、前記第１の記憶部からデータを読み出す処理と、前記読み出されたデータを前記第２の記憶部に書き込む処理を並行して実行することを特徴とする請求項１５に記載の情報処理装置。

【請求項１７】

メタデータを格納するメタデータ格納部に接続され、

前記制御部は、

前記第１の処理により前記第１の記憶部のディスクステータスが前記保護状態でないと判定された場合、前記第１の記憶部に書き込まれたデータのアドレスと前記第１の記憶部の識別情報との紐付けを前記メタデータに記録し、前記第１の処理により前記第１の記憶部のディスクステータスが前記保護状態と判定された場合、前記第２の記憶部に書き込まれたデータのアドレスと前記第２の記憶部の識別情報との紐付けを前記メタデータに記録し、

読み出し処理において、読み出し対象アドレスと前記第１の記憶部の識別情報との紐付けが前記メタデータに記録されている場合は、前記第１の記憶部からデータを読み出し、読み出し対象アドレスと前記第２の記憶部との紐付けが前記メタデータに記録されている場合は、前記第２の記憶部からデータを読み出すことを特徴とする請求項１５乃至１６の何れか一つに記載の情報処理装置。

【請求項１８】

引き継ぎ履歴格納部に接続され、

前記制御部は、

前記第３の処理では、

前記第１の処理により前記第１の記憶部のディスクステータスが保護状態と判定された場合、前記第２の記憶部に書き込まれたデータのアドレスを前記引き継ぎ履歴格納部に記録し、

読み出し処理において、前記引き継ぎ履歴格納部に読み出し対象アドレスが記録されている場合は、前記第２の記憶部からデータを読み出し、前記引き継ぎ履歴格納部に読み出し対象アドレスが記録されていない場合は、前記第１の記憶部からデータを読み出すことを特徴とする請求項１５乃至１６の何れか一つに記載の情報処理装置。

【請求項１９】
前記制御部は、
前記第１の処理では、前記信頼性情報と閾値との比較を行い、この比較に基づいて前記第１の記憶部の寿命到達条件が成立したと判定したときに、前記第１の記憶部のディスクステータスを前記保護状態であると判定することを特徴とする請求項１５に記載の情報処理装置。
【請求項２０】
前記制御部は、
前記第１の処理では、前記信頼性情報と閾値との比較を行い、この比較に基づいて前記第１の記憶部の寿命到達条件が成立したと判定したときに、前記第１の記憶部のディスクステータスを信頼性劣化状態であると判定し、
前記第１の処理により前記第１の記憶部のディスクステータスが前記信頼性劣化状態として判定された場合、前記第２の記憶部を前記第１の記憶部に格納されているデータのデータ引き継ぎ先状態として認識し、前記第１の記憶部のディスクステータスを前記保護状態に変更することを特徴とする請求項１５に記載の情報処理装置。
【請求項２１】
リードおよびライト可能な第３のメモリを有する１乃至複数の第３の記憶部に接続可能であり、
前記制御部は、
同一アドレスのデータが前記第１および第２の記憶部のうちの一方に存在するように、前記第１および前記第２の記憶部での記憶データのアドレスを管理し、
前記第１の記憶部のディスクステータスが保護状態として認識された場合、前記第１の記憶部と前記第２の記憶部を同一の論理単位として認識し、
前記第１の記憶部のディスクステータスが保護状態として認識された場合、前記１乃至複数の第３の記憶部と前記論理単位とがＲＡＩＤを構成するように制御することを特徴とする請求項１５に記載の情報処理装置。
【請求項２２】
リードおよびライト可能な第１のメモリを有する第１の記憶部と、リードおよびライト可能な第２のメモリを有する第２の記憶部と、引き継ぎ履歴格納部とを接続可能な情報処理装置であって、
前記第１の記憶部から取得した信頼性情報に基づき、前記第１の記憶部のディスクステータスを判定する第１の処理と、
前記第１の処理により前記第１の記憶部のディスクステータスが保護状態として認識された場合、書き込みは前記第２の記憶部に対し実行し、読み出しは前記引き継ぎ履歴格納部に読み出し対象アドレスが記録されている場合は、前記第２の記憶部からデータを読み出し、前記引き継ぎ履歴格納部に読み出し対象アドレスが記録されていない場合は、前記第１の記憶部からデータを読み出す第２の処理と、
前記第２の処理による書き込みに伴って、前記第２の記憶部に書き込まれたデータのアドレスを前記引き継ぎ履歴格納部に記録する第３の処理と、
前記第１の処理により前記第１の記憶部のディスクステータスが保護状態と判定された後に、前記第１の記憶部に格納されたデータの削除処理命令を受信した場合に、前記削除処理命令による削除対象データのアドレスを前記引き継ぎ履歴格納部に記録する第４の処理
を実行する制御部
を備えることを特徴とする情報処理装置。
【請求項２３】
前記制御部は、
前記第４の処理では、
前記第１の処理により前記第１の記憶部のディスクステータスが保護状態と判定された後に、前記第１または第２の記憶部に格納されたデータの削除処理命令を受信した場合に

、前記引き継ぎ履歴格納部に前記削除処理命令による削除対象データのアドレスが記録されている場合は、前記受信された削除処理命令を前記第２の記憶部に送信し、前記引き継ぎ履歴格納部に前記削除処理命令による削除対象データのアドレスを記録せず、
前記引き継ぎ履歴格納部に前記削除処理命令による削除対象データのアドレスが記録されていない場合は、前記削除処理命令による削除対象データのアドレスを前記引き継ぎ履歴格納部に記録する
ことを特徴とする請求項２２に記載の情報処理装置。
【請求項２４】
前記制御部は、
前記第１の処理では、前記信頼性情報と閾値との比較を行い、この比較に基づいて前記第１の記憶部の寿命到達条件が成立したと判定したときに、前記第１の記憶部のディスクステータスを前記保護状態であると判定することを特徴とする請求項２２に記載の情報処理装置。
【請求項２５】
前記制御部は、
前記第１の処理では、前記信頼性情報と閾値との比較を行い、この比較に基づいて前記第１の記憶部の寿命到達条件が成立したと判定したときに、前記第１の記憶部のディスクステータスを信頼性劣化状態であると判定し、
前記第１の処理により前記第１の記憶部のディスクステータスが前記信頼性劣化状態として判定された場合、前記第２の記憶部を前記第１の記憶部に格納されているデータのデータ引き継ぎ先状態として認識し、前記第１の記憶部のディスクステータスを前記保護状態に変更することを特徴とする請求項２２に記載の情報処理装置。
【請求項２６】
リードおよびライト可能な第３のメモリを有する１乃至複数の第３の記憶部に接続可能であり、
前記制御部は、
同一アドレスのデータが前記第１および第２の記憶部のうちの一方に存在するように、前記第１および前記第２の記憶部での記憶データのアドレスを管理し、
前記第１の記憶部のディスクステータスが保護状態として認識された場合、前記第１の記憶部と前記第２の記憶部を同一の論理単位として認識し、
前記第１の記憶部のディスクステータスが保護状態として認識された場合、前記１乃至複数の第３の記憶部と前記論理単位とがＲＡＩＤを構成するように制御することを特徴とする請求項２２に記載の情報処理装置。
【請求項２７】
前記制御部は、
前記第１の記憶部に書き込まれているデータのうち前記第２の記憶部へ引き継ぎ対象とするデータを選び、前記引き継ぎ対象のデータに対応するアドレスが全て前記引き継ぎ履歴格納部に格納されている場合に、前記第１の記憶部の保護状態を終了すると判定することを特徴とする請求項２２乃至２６の何れか一つに記載の情報処理装置。
【請求項２８】
前記制御部は、
前記第３の処理では、
前記引き継ぎ履歴格納部に第１のアドレスを記録する場合に、前記第１のアドレスと重複または連続する第２のアドレスがすでに前記引き継ぎ履歴格納部に格納されている場合には、前記第１のアドレスと前記第２のアドレスを合併した第３のアドレスを前記引き継ぎ履歴格納部に記録し、前記第２のアドレスを無効化することを特徴とする請求項２２乃至２７の何れか一つに記載の情報処理装置。
【請求項２９】
リードおよびライト可能な第１のメモリを有する第１の記憶部と、引き継ぎ履歴格納部が記憶され、リードおよびライト可能な第２のメモリを有する第２の記憶部とを接続可能

な情報処理装置にロードされる制御プログラムであって、
前記第１の記憶部から取得した信頼性情報に基づき、前記第１の記憶部のディスクステータスを判定する第１の処理と、
前記第１の処理により前記第１の記憶部のディスクステータスが保護状態として認識された場合、書き込みは前記第２の記憶部に対し実行し、読み出しは前記引き継ぎ履歴格納部に読み出し対象アドレスが記録されている場合は、前記第２の記憶部からデータを読み出し、前記引き継ぎ履歴格納部に読み出し対象アドレスが記録されていない場合は、前記第１の記憶部からデータを読み出す第２の処理と、
前記第２の処理による書き込みに伴って、前記第２の記憶部に書き込まれたデータのアドレスを前記引き継ぎ履歴格納部に記録する第３の処理と、
前記第１の処理により前記第１の記憶部のディスクステータスが保護状態として認識されなかった場合、読み出しは前記引き継ぎ履歴格納部を読み出すこと無く前記第１の記憶部からデータを読み出す第４の処理と、
を前記情報処理装置に実行させるように構成された制御プログラム。
【請求項３０】
前記制御部は、
前記第１の処理では、前記信頼性情報と閾値との比較を行い、この比較に基づいて前記第１の記憶部の寿命到達条件が成立したと判定したときに、前記第１の記憶部のディスクステータスを前記保護状態であると判定することを特徴とする請求項２９に記載の制御プログラム。
【請求項３１】
前記制御部は、
前記第１の処理では、前記信頼性情報と閾値との比較を行い、この比較に基づいて前記第１の記憶部の寿命到達条件が成立したと判定したときに、前記第１の記憶部のディスクステータスを信頼性劣化状態であると判定し、
前記第１の処理により前記第１の記憶部のディスクステータスが前記信頼性劣化状態として判定された場合、前記第２の記憶部を前記第１の記憶部に格納されているデータのデータ引き継ぎ先状態として認識し、前記第１の記憶部のディスクステータスを前記保護状態に変更することを特徴とする請求項２９に記載の制御プログラム。
【請求項３２】
リードおよびライト可能な第３のメモリを有する１乃至複数の第３の記憶部に接続可能であり、
前記制御部は、
同一アドレスのデータが前記第１および第２の記憶部のうちの一方に存在するように、前記第１および前記第２の記憶部での記憶データのアドレスを管理し、
前記第１の記憶部のディスクステータスが保護状態として認識された場合、前記第１の記憶部と前記第２の記憶部を同一の論理単位として認識し、
前記第１の記憶部のディスクステータスが保護状態として認識された場合、前記１乃至複数の第３の記憶部と前記論理単位とがＲＡＩＤを構成するように制御することを特徴とする請求項２９に記載の制御プログラム。
【請求項３３】
前記制御部は、
前記第１の記憶部に書き込まれているデータのうち前記第２の記憶部へ引き継ぎ対象とするデータを選び、前記引き継ぎ対象のデータに対応するアドレスが全て前記引き継ぎ履歴格納部に格納されている場合に、前記第１の記憶部の保護状態を終了すると判定することを特徴とする請求項２９乃至３２の何れか一つに記載の制御プログラム。
【請求項３４】
前記制御部は、
前記第３の処理では、
前記引き継ぎ履歴格納部に第１のアドレスを記録する場合に、前記第１のアドレスと重

複または連続する第２のアドレスがすでに前記引き継ぎ履歴格納部に格納されている場合には、前記第１のアドレスと前記第２のアドレスを合併した第３のアドレスを前記引き継ぎ履歴格納部に記録し、前記第２のアドレスを無効化することを特徴とする請求項２９乃至３３の何れか一つに記載の制御プログラム。

【請求項３５】

前記制御部は、前記第１の処理により前記第１の記憶部のディスクステータスが保護状態として認識されなかった場合、前記引き継ぎ履歴格納部を消去することを特徴とする請求項２９乃至３４の何れか一つに記載の制御プログラム。

【請求項３６】

リードおよびライト可能な不揮発性の半導体メモリである第１のメモリを有する第１の記憶部と、引き継ぎ履歴格納部が記憶され、リードおよびライト可能な不揮発性の半導体メモリである第２のメモリを有する第２の記憶部と、引き継ぎ履歴格納部とを接続可能な情報処理装置にロードされる制御プログラムであって、

前記第１の記憶部から取得した信頼性情報に基づき、前記第１の記憶部のディスクステータスを判定する第１の処理と、

前記第１の処理により前記第１の記憶部のディスクステータスが保護状態として認識された場合、書き込みは前記第２の記憶部に対し実行し、読み出しは前記引き継ぎ履歴格納部に読み出し対象アドレスが記録されている場合は、前記第２の記憶部からデータを読み出し、前記引き継ぎ履歴格納部に読み出し対象アドレスが記録されていない場合は、前記第１の記憶部からデータを読み出す第２の処理と、

前記第２の処理による書き込みに伴って、前記第２の記憶部に書き込まれたデータのアドレスを前記引き継ぎ履歴格納部に記録する第３の処理と、

前記第１の処理により前記第１の記憶部のディスクステータスが保護状態として認識されなかった場合、読み出しは前記引き継ぎ履歴格納部を読み出すこと無く前記第１の記憶部からデータを読み出す第４の処理と、

を前記情報処理装置に実行させるように構成された制御プログラム。

【請求項３７】

前記制御部は、

前記第１の処理では、前記信頼性情報と閾値との比較を行い、この比較に基づいて前記第１の記憶部の寿命到達条件が成立したと判定したときに、前記第１の記憶部のディスクステータスを前記保護状態であると判定することを特徴とする請求項３６に記載の制御プログラム。

【請求項３８】

前記制御部は、

前記第１の処理では、前記信頼性情報と閾値との比較を行い、この比較に基づいて前記第１の記憶部の寿命到達条件が成立したと判定したときに、前記第１の記憶部のディスクステータスを信頼性劣化状態であると判定し、

前記第１の処理により前記第１の記憶部のディスクステータスが前記信頼性劣化状態として判定された場合、前記第２の記憶部を前記第１の記憶部に格納されているデータのデータ引き継ぎ先状態として認識し、前記第１の記憶部のディスクステータスを前記保護状態に変更することを特徴とする請求項３６に記載の制御プログラム。

【請求項３９】

リードおよびライト可能な第３のメモリを有する１乃至複数の第３の記憶部に接続可能であり、

前記制御部は、

同一アドレスのデータが前記第１および第２の記憶部のうちの一方に存在するように、前記第１および前記第２の記憶部での記憶データのアドレスを管理し、

前記第１の記憶部のディスクステータスが保護状態として認識された場合、前記第１の記憶部と前記第２の記憶部を同一の論理単位として認識し、

前記第１の記憶部のディスクステータスが保護状態として認識された場合、前記１乃至

複数の第３の記憶部と前記論理単位とがＲＡＩＤを構成するように制御することを特徴とする請求項３６に記載の制御プログラム。

【請求項４０】

前記制御部は、

前記第１の記憶部に書き込まれているデータのうち前記第２の記憶部へ引き継ぎ対象とするデータを選び、前記引き継ぎ対象のデータに対応するアドレスが全て前記引き継ぎ履歴格納部に格納されている場合に、前記第１の記憶部の保護状態を終了すると判定することを特徴とする請求項３６乃至３９の何れか一つに記載の制御プログラム。

【請求項４１】

前記制御部は、

前記第３の処理では、

前記引き継ぎ履歴格納部に第１のアドレスを記録する場合に、前記第１のアドレスと重複または連続する第２のアドレスがすでに前記引き継ぎ履歴格納部に格納されている場合には、前記第１のアドレスと前記第２のアドレスを合併した第３のアドレスを前記引き継ぎ履歴格納部に記録し、前記第２のアドレスを無効化することを特徴とする請求項３６乃至４０の何れか一つに記載の制御プログラム。

【請求項４２】

前記制御部は、前記第１の処理により前記第１の記憶部のディスクステータスが保護状態として認識されなかった場合、前記引き継ぎ履歴格納部を消去することを特徴とする請求項３６乃至４１の何れか一つに記載の制御プログラム。

【請求項４３】

リードおよびライト可能な第１のメモリを有する第１の記憶部と、リードおよびライト可能な第２のメモリを有する第２の記憶部とを接続可能な情報処理装置にロードされる制御プログラムであって、

前記第１の記憶部から取得した信頼性情報に基づき、前記第１の記憶部のディスクステータスを判定する第１の処理と、

前記第１の処理により前記第１の記憶部のディスクステータスが前記保護状態でないと判定された場合、前記第１の記憶部に書き込みを実行し、前記第１の処理により前記第１の記憶部のディスクステータスが保護状態として認識された場合、前記第２の記憶部に書き込みを実行する第２の処理と、

前記第１の処理により前記第１の記憶部のディスクステータスが保護状態として認識された後に前記第１の記憶部に対する読み出しが発生した場合は、前記第１の記憶部からデータを読み出し、前記読み出されたデータを前記第２の記憶部に書き込む第３の処理と、

を前記情報処理装置に実行させるように構成された制御プログラム。

【請求項４４】

前記制御部は、

前記第３の処理では、前記第１の記憶部からデータを読み出す処理と、前記読み出されたデータを前記第２の記憶部に書き込む処理を並行して実行することを特徴とする請求項４３に記載の制御プログラム。

【請求項４５】

メタデータを格納するメタデータ格納部に接続され、

前記制御部は、

前記第１の処理により前記第１の記憶部のディスクステータスが前記保護状態でないと判定された場合、前記第１の記憶部に書き込まれたデータのアドレスと前記第１の記憶部の識別情報との紐付けを前記メタデータに記録し、前記第１の処理により前記第１の記憶部のディスクステータスが前記保護状態と判定された場合、前記第２の記憶部に書き込まれたデータのアドレスと前記第２の記憶部の識別情報との紐付けを前記メタデータに記録し、

読み出し処理において、読み出し対象アドレスと前記第１の記憶部の識別情報との紐付けが前記メタデータに記録されている場合は、前記第１の記憶部からデータを読み出し、

読み出し対象アドレスと前記第２の記憶部との紐付けが前記メタデータに記録されている場合は、前記第２の記憶部からデータを読み出すことを特徴とする請求項４３乃至４４の何れか一つに記載の制御プログラム。

【請求項４６】

引き継ぎ履歴格納部に接続され、

前記制御部は、

前記第３の処理では、

前記第１の処理により前記第１の記憶部のディスクステータスが保護状態と判定された場合、前記第２の記憶部に書き込まれたデータのアドレスを前記引き継ぎ履歴格納部に記録し、

読み出し処理において、前記引き継ぎ履歴格納部に読み出し対象アドレスが記録されている場合は、前記第２の記憶部からデータを読み出し、前記引き継ぎ履歴格納部に読み出し対象アドレスが記録されていない場合は、前記第１の記憶部からデータを読み出すことを特徴とする請求項４３乃至４４の何れか一つに記載の制御プログラム。

【請求項４７】

前記制御部は、

前記第１の処理では、前記信頼性情報と閾値との比較を行い、この比較に基づいて前記第１の記憶部の寿命到達条件が成立したと判定したときに、前記第１の記憶部のディスクステータスを前記保護状態であると判定することを特徴とする請求項４３に記載の制御プログラム。

【請求項４８】

前記制御部は、

前記第１の処理では、前記信頼性情報と閾値との比較を行い、この比較に基づいて前記第１の記憶部の寿命到達条件が成立したと判定したときに、前記第１の記憶部のディスクステータスを信頼性劣化状態であると判定し、

前記第１の処理により前記第１の記憶部のディスクステータスが前記信頼性劣化状態として判定された場合、前記第２の記憶部を前記第１の記憶部に格納されているデータのデータ引き継ぎ先状態として認識し、前記第１の記憶部のディスクステータスを前記保護状態に変更することを特徴とする請求項４３に記載の制御プログラム。

【請求項４９】

リードおよびライト可能な第３のメモリを有する１乃至複数の第３の記憶部に接続可能であり、

前記制御部は、

同一アドレスのデータが前記第１および第２の記憶部のうちの一方に存在するように、前記第１および前記第２の記憶部での記憶データのアドレスを管理し、

前記第１の記憶部のディスクステータスが保護状態として認識された場合、前記第１の記憶部と前記第２の記憶部を同一の論理単位として認識し、

前記第１の記憶部のディスクステータスが保護状態として認識された場合、前記１乃至複数の第３の記憶部と前記論理単位とがＲＡＩＤを構成するように制御することを特徴とする請求項４３に記載の制御プログラム。

【請求項５０】

リードおよびライト可能な第１のメモリを有する第１の記憶部と、リードおよびライト可能な第２のメモリを有する第２の記憶部と、引き継ぎ履歴格納部とを接続可能な情報処理装置にロードされる制御プログラムであって、

前記第１の記憶部から取得した信頼性情報に基づき、前記第１の記憶部のディスクステータスを判定する第１の処理と、

前記第１の処理により前記第１の記憶部のディスクステータスが保護状態として認識された場合、書き込みは前記第２の記憶部に対し実行し、読み出しは前記引き継ぎ履歴格納部に読み出し対象アドレスが記録されている場合は、前記第２の記憶部からデータを読み出し、前記引き継ぎ履歴格納部に読み出し対象アドレスが記録されていない場合は、前記

第１の記憶部からデータを読み出す第２の処理と、
前記第２の処理による書き込みに伴って、前記第２の記憶部に書き込まれたデータのアドレスを前記引き継ぎ履歴格納部に記録する第３の処理と、
前記第１の処理により前記第１の記憶部のディスクステータスが保護状態と判定された後に、前記第１の記憶部に格納されたデータの削除処理命令を受信した場合に、前記削除処理命令による削除対象データのアドレスを前記引き継ぎ履歴格納部に記録する第４の処理と、
を前記情報処理装置に実行させるように構成された制御プログラム。
【請求項５１】
前記制御部は、
前記第４の処理では、
前記第１の処理により前記第１の記憶部のディスクステータスが保護状態と判定された後に、前記第１または第２の記憶部に格納されたデータの削除処理命令を受信した場合に、前記引き継ぎ履歴格納部に前記削除処理命令による削除対象データのアドレスが記録されている場合は、前記受信された削除処理命令を前記第２の記憶部に送信し、前記引き継ぎ履歴格納部に前記削除処理命令による削除対象データのアドレスを記録せず、
前記引き継ぎ履歴格納部に前記削除処理命令による削除対象データのアドレスが記録されていない場合は、前記削除処理命令による削除対象データのアドレスを前記引き継ぎ履歴格納部に記録する
ことを特徴とする請求項５０に記載の制御プログラム。
【請求項５２】
前記制御部は、
前記第１の処理では、前記信頼性情報と閾値との比較を行い、この比較に基づいて前記第１の記憶部の寿命到達条件が成立したと判定したときに、前記第１の記憶部のディスクステータスを前記保護状態であると判定することを特徴とする請求項５０に記載の制御プログラム。
【請求項５３】
前記制御部は、
前記第１の処理では、前記信頼性情報と閾値との比較を行い、この比較に基づいて前記第１の記憶部の寿命到達条件が成立したと判定したときに、前記第１の記憶部のディスクステータスを信頼性劣化状態であると判定し、
前記第１の処理により前記第１の記憶部のディスクステータスが前記信頼性劣化状態として判定された場合、前記第２の記憶部を前記第１の記憶部に格納されているデータのデータ引き継ぎ先状態として認識し、前記第１の記憶部のディスクステータスを前記保護状態に変更することを特徴とする請求項５０に記載の制御プログラム。
【請求項５４】
リードおよびライト可能な第３のメモリを有する１乃至複数の第３の記憶部に接続可能であり、
前記制御部は、
同一アドレスのデータが前記第１および第２の記憶部のうちの一方に存在するように、前記第１および前記第２の記憶部での記憶データのアドレスを管理し、
前記第１の記憶部のディスクステータスが保護状態として認識された場合、前記第１の記憶部と前記第２の記憶部を同一の論理単位として認識し、
前記第１の記憶部のディスクステータスが保護状態として認識された場合、前記１乃至複数の第３の記憶部と前記論理単位とがＲＡＩＤを構成するように制御することを特徴とする請求項５０に記載の制御プログラム。
【請求項５５】
前記制御部は、
前記第１の記憶部に書き込まれているデータのうち前記第２の記憶部へ引き継ぎ対象とするデータを選び、前記引き継ぎ対象のデータに対応するアドレスが全て前記引き継ぎ履

歴格納部に格納されている場合に、前記第１の記憶部の保護状態を終了すると判定することを特徴とする請求項５０乃至５４の何れか一つに記載の制御プログラム。
【請求項５６】
前記制御部は、
前記第３の処理では、
前記引き継ぎ履歴格納部に第１のアドレスを記録する場合に、前記第１のアドレスと重複または連続する第２のアドレスがすでに前記引き継ぎ履歴格納部に格納されている場合には、前記第１のアドレスと前記第２のアドレスを合併した第３のアドレスを前記引き継ぎ履歴格納部に記録し、前記第２のアドレスを無効化することを特徴とする請求項５０乃至５５の何れか一つに記載の制御プログラム。
【発明の詳細な説明】
【技術分野】
【０００１】
本発明の実施形態は、第１の記憶部のデータを第２の記憶部に引き継ぐための情報処理システムおよび制御プログラムに関する。
【背景技術】
【０００２】
フラッシュメモリを用いてパーソナルコンピュータやサーバの二次記憶装置を構成する場合、誤りが多いなどの理由によって記憶領域として使用できなくなる不良ブロック(defective block)や、読み出せなくなる領域（不良領域）などが生じる場合がある。このような不良ブロック数や不良領域数が上限値を超えた場合には、新たな不良ブロックや不良領域を登録することができないので、ライト要求のあったデータをフラッシュメモリへ書き込むことを保障できない。このため、不良ブロック数や不良領域数が上限値を超えた場合には、フラッシュメモリに空き容量があるにも関わらず、突然データの書き込みが不能になるといった問題があった。
【０００３】
そこで、記憶装置の劣化をモニタし、記憶装置の寿命到達より前に、情報処理装置から記憶装置へのデータの書き込みを制限する方法がある。この方法により、記憶装置の寿命到達により前にユーザデータを他の記憶装置にバックアップして引き継ぎ、記憶装置の寿命到達によるデータ損失を未然に防ぐことが可能である。
【０００４】
しかしながら、この方法では、ユーザデータを他の記憶装置にバックアップする作業があり、煩雑である。また、情報処理装置によってバックアップ元の記憶装置へのデータ書き込みが制限されるので、情報処理装置にロードされている各種アプリケーションプログラムの動作が制限され、アプリケーションプログラムの処理速度が低下する可能性がある。
【先行技術文献】
【特許文献】
【０００５】
【特許文献１】米国特許出願公開第２０１２/０２８４４５３号明細書
【特許文献２】米国特許出願公開第２０１２/０２４６３８８号明細書
【特許文献３】米国特許出願公開第２００９/０２２２６１７号明細書
【特許文献４】米国特許第８,２３０,１６４号明細書
【特許文献５】米国特許第７,８６１,１２２号明細書
【特許文献６】米国特許第７,８４０,８３６号明細書
【特許文献７】米国特許出願公開第２００７／０２１４１８７号明細書
【特許文献８】米国特許第７,６０６,９４６号明細書
【特許文献９】米国特許第６,５２９,９９２号明細書
【特許文献１０】米国特許第７,２２２,２５９号明細書
【特許文献１１】米国特許第７,４２４,６３９号明細書

【特許文献１２】米国特許第７,７３９,５４４号明細書
【特許文献１３】米国特許第７,５２５,７４９号明細書
【特許文献１４】米国特許第８,０３７,３８０号明細書
【特許文献１５】米国特許出願公開第２０１０/０３２５１９９号明細書
【特許文献１６】米国特許出願公開第２０１１/０２３９０８３号明細書
【特許文献１７】特許第５０５２３７６号公報
【特許文献１８】特開２０１０-２２５０２１号公報
【特許文献１９】特許第３,５６５,６８７号公報
【特許文献２０】米国特許第８,１７６,３６７号明細書
【特許文献２１】米国特許出願公開第２０１１/０１９７０４５号明細書
【特許文献２２】特許第４,６４３,４５６号公報
【特許文献２３】特許第４,７６４,４９０号公報
【特許文献２４】特開２０１１-２０９８７９号公報
【特許文献２５】国際公開第２０１３／０２７６４２号パンフレット
【非特許文献】
【０００６】
【非特許文献１】Information technology ATA/ATAPI Command Set-3 (ACS-3), d2161r4-ATAATAPI_Command_Set_-_3.pdf, Sep.04.2012, http://www.t13.org/
【非特許文献２】NVM Express Revision 1.1, Oct.11.2012, NVM_Express_1_1.pdf, http://www.nvmexpress.org/
【非特許文献３】Serial ATA International Organization: Serial ATA Revision 3.1 Gold Revision, Jul.18.2011, http://www.serialata.org/
【非特許文献４】PCI Express Base Specification Revision 3.0, Nov.10.2010, http://www.pcisig.com/
【非特許文献５】Serial ATA Technical Proposal: SATA31_TPR_C108 Title: Device Sleep, http://www.serialata.org/
【非特許文献６】SCSI Primary Commands-4 (SPC-4), INCITS T10/1731-D, Revision 36e, http://www.t10.org/
【非特許文献７】喜連川優 編著、「よくわかるストレージネットワーキング」、第１版、株式会社オーム社、2011年5月20日、 p.6-9、p67-93
【発明の概要】
【発明が解決しようとする課題】
【０００７】
本発明の一つの実施形態は、ユーザによるバックアップ作業を行うことなくデータ引き継ぎをなし得る情報処理システムおよび制御プログラム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】
【０００８】
本発明の一つの実施形態によれば、情報処理システムは、リードおよびライト可能な第１のメモリを有する第１の記憶部と、リードおよびライト可能な第２のメモリを有する第２の記憶部と、前記第１の記憶部と前記第２の記憶部に接続可能な情報処理装置とを備える。前記情報処理装置は、前記第１の記憶部から取得した信頼性情報に基づき、前記第１の記憶部のディスクステータスを判定する第１の処理と、前記第１の処理により前記第１の記憶部のディスクステータスが保護状態として認識された場合、書き込みは前記第２の記憶部に対し実行し、読み出しは前記第１および前記第２の記憶部に対し実行する第２の処理と、前記第２の処理による書き込みに伴って、前記第１および前記第２の記憶部での記憶データのアドレスを更新する第３の処理とを実行する制御部を備えることを特徴とする。
【図面の簡単な説明】
【０００９】
【図１】図１は、第１の実施形態の情報処理システムの機能構成例を示すブロック図であ

る。
【図２】図２は、制御プログラムをＳＳＤに保存する場合の情報処理システムの機能構成例を示すブロック図である。
【図３】図３は、制御プログラムをその他の外部記憶装置に保存する場合の情報処理システムの機能構成例を示すブロック図である。
【図４】図４は、制御プログラムをＷＥＢからインストールする場合の情報処理システムの機能構成例を示すブロック図である。
【図５】図５は、制御プログラムを光学ドライブからインストールする場合の情報処理システムの機能構成例を示すブロック図である。
【図６】図６は、制御プログラムをＵＳＢメモリからインストールする場合の情報処理システムの機能構成例を示すブロック図である。
【図７】図７は、情報処理装置の階層的機能構成例を示すブロック図である。
【図８】図８は、情報処理システムの外観的構成を示す図である。
【図９】図９は、情報処理システムの他の外観的構成を示す図である。
【図１０】図１０は、ＮＡＮＤメモリチップの機能構成例を示すブロック図である。
【図１１】図１１は、ＮＡＮＤメモリチップに含まれる１個のプレーンの構成例を示す回路図である。
【図１２】図１２は、４値データ記憶方式での閾値分布を示す図である。
【図１３】図１３は、ＳＳＤの機能構成例を示すブロック図である。
【図１４】図１４は、ＳＳＤの管理情報を示す図である。
【図１５】図１５は、ＬＢＡとＳＳＤの管理単位との関係を示す図である。
【図１６】図１６は、ＬＢＡから物理アドレスを特定する手順を示すフローチャートである。
【図１７】図１７は、ＳＳＤの読み出し動作例を示すフローチャートである。
【図１８】図１８は、ＳＳＤの読み出し動作例を示すフローチャートである。
【図１９】図１９は、ＳＳＤの書き込み動作例を示すフローチャートである。
【図２０】図２０は、ＳＳＤの書き込み動作例を示すフローチャートである。
【図２１】図２１は、ＮＡＮＤ整理を行う場合の動作手順を示すフローチャートである。
【図２２】図２２は、削除通知を受信したときのＳＳＤの動作例を示すフローチャートである。
【図２３】図２３は、エラー発生時のＳＳＤの動作例を示すフローチャートである。
【図２４】図２４は、統計情報取得処理を示すフローチャートである。
【図２５】図２５は、バッド論理セクタテーブルを示す図である。
【図２６】図２６は、バッドクラスタテーブルを示す図である。
【図２７】図２７は、統計情報の例としてのＳＭＡＲＴ情報を示す図である。
【図２８】図２８は、統計情報のRaw ValueとＳＳＤの不良率の関係を示すグラフである。
【図２９】図２９は、コマンド応答エラー処理を示すフローチャートである。
【図３０】図３０は、他のコマンド応答エラー処理を示すフローチャートである。
【図３１】図３１は、寿命到達時処理を行うより以前における情報処理装置が管理するデータの構成例を示す図である。
【図３２】図３２は、Boot Loaderの内部データを示す図である。
【図３３】図３３は、メタデータの構成例を示す図である。
【図３４】図３４は、アプリケーションプログラムがＯＳに対して論理ドライブへのアクセス要求を行った場合のフローチャートである。
【図３５】図３５は、情報処理装置が起動した時などの制御プログラムの処理手順を示すフローチャートである。
【図３６】図３６は、記憶部のライフサイクルを示す状態遷移図である。
【図３７】図３７は、記憶部のライフサイクルを示す状態遷移図である。
【図３８】図３８は、データ引き継ぎ元記憶部及びデータ引き継ぎ先記憶部が接続された

状態での情報処理装置が管理するデータの構成例を示す図である。
【図３９】図３９は、記憶部のライフサイクルを示す状態遷移図である。
【図４０】図４０は、寿命到達時処理において、制御プログラムが行う処理を示すフローチャートである。
【図４１】図４１は、論理ドライブステータステーブルを示す図である。
【図４２】図４２は、アプリケーションプログラムからＯＳに書き込み要求が送信された時の、ＯＳの処理手順を示すフローチャートである。
【図４３】図４３は、アプリケーションプログラムからＯＳに削除要求が送信された時の、ＯＳの処理手順を示すフローチャートである。
【図４４】図４４は、アプリケーションプログラムからＯＳに読み出し要求が送信された時の、ＯＳの処理手順を示すフローチャートである。
【図４５】図４５は、バックグラウンドバックアップの処理手順を示すフローチャートである。
【図４６】図４６は、データ引き継ぎ完了時の制御プログラムの動作手順を示すフローチャートである。
【図４７】図４７は、データ引き継ぎ状態における論理ドライブからの読み出しを概念的に示す図である。
【図４８】図４８は、データ引き継ぎ状態における論理ドライブへの書き込みを概念的に示す図である。
【図４９】図４９は、第２の実施形態の情報処理システムの機能構成例を示すブロック図である。
【図５０】図５０は、引き継ぎ履歴を示す図である。
【図５１】図５１は、引き継ぎ履歴に対する履歴の書き込み例を示す図である。
【図５２】図５２は、寿命到達時処理を示すフローチャートである。
【図５３】図５３は、論理ドライブからの読み出しの動作手順を示すフローチャートである。
【図５４】図５４は、アプリケーションプログラムからＯＳにＬＢＡデータ削除要求が送信された時の、ＯＳの処理手順を示す図である。
【図５５】図５５は、論理ドライブへの書き込みの動作手順を示すフローチャートである。
【図５６】図５６は、引き継ぎ履歴を用いたデータ引き継ぎ状況の監視手順を示すフローチャートである。
【図５７】図５７は、記憶部が寿命到達する際の論理ドライブのステータスの遷移を示す図である。
【図５８】図５８は、第３の実施形態の情報処理システムの機能構成例を示すブロック図である。
【図５９】図５９は、通常状態のアレイＬＢＡ（ＡＬＢＡ）と記憶部ＬＢＡ（ＳＬＢＡ）の変換方法を示す図である。
【図６０】図６０は、引き継ぎ状態遷移直後のアレイＬＢＡ（ＡＬＢＡ）と記憶部ＬＢＡ（ＳＬＢＡ）の変換方法を示す図である。
【図６１】図６１は、引き継ぎ履歴を示す図である。
【図６２】図６２は、寿命到達時処理を示すフローチャートである。
【図６３】図６３は、ＲＡＩＤコントローラが起動した時などの制御部の処理手順を示すフローチャートである。
【図６４】図６４は、読み出し要求を実行する際の、ＲＡＩＤコントローラの処理手順を示すフローチャートである。
【図６５】図６５は、読み出し要求を実行する際の、ＲＡＩＤコントローラの他の処理手順を示すフローチャートである。
【図６６】図６６は、書き込みし要求を実行する際の、ＲＡＩＤコントローラの他の処理手順を示すフローチャートである。

【図６７】図６７は、書き込み処理の具体例をを示す図である。
【図６８】図６８は、引き継ぎ履歴を用いたデータ引き継ぎ状況の監視手順を示すフローチャートである。
【図６９】図６９は、各記憶部のステータスの遷移を示す図である。
【図７０】図７０は、１台のデータ引き継ぎ状態と他の１台の故障が同時に発生した場合の例を示す図である。
【図７１】図７１は、２台のデータ引き継ぎ状態と他の１台の故障が同時に発生した場合の例を示す図である。
【図７２】図７２は、１台のデータ引き継ぎ状態において、他の１台の記憶部でＥＣＣ訂正不能エラーが発生した場合の例を示すを示す図である。
【図７３】図７３は、第４の実施形態の情報処理システムの機能構成例を示すブロック図である。
【図７４】図７４は、第４の実施形態の他の情報処理システムの機能構成例を示すブロック図である。
【図７５】図７５は、ステータステーブルを示す図である。
【図７６】図７６は、第５の実施形態の情報処理システムの機能構成例を示すブロック図である。
【図７７】図７７は、第６の実施形態の情報処理システムの機能構成例を示すブロック図である。
【図７８】図７８は、第６の実施形態の情報処理システムの他の機能構成例を示すブロック図である。
【図７９】図７９は、制御部の通常状態から信頼性劣化モードへの遷移操作を示すフローチャートである。
【図８０】図８０は、制御部がＣＰＵから記憶部の記憶部情報要求を受信した場合の処理手順の例を示すフローチャートである。
【図８１】図８１は、制御部がＣＰＵから記憶部の記憶部情報要求を受信した場合の他の処理手順の例を示すフローチャートである。
【図８２】図８２は、情報処理装置でのドライブ表示画面例を示す図である。
【図８３】図８３は、情報処理装置での他のドライブ表示画面例を示す図である。
【図８４】図８４は、情報処理装置での他の表示画面例を示す図である。
【図８５】図８５は、第６の実施形態の情報処理システムの他の機能構成例を示すブロック図である。
【発明を実施するための形態】
【００１０】
以下に添付図面を参照して、実施形態にかかる情報処理システムおよび制御プログラムを詳細に説明する。なお、これら実施形態により本発明が限定されるものではない。
【００１１】
（第１の実施形態）
（システムの構成）
図１に情報処理システムの一例としてのコンピュータシステムの第１の実施形態の構成を示す。情報処理システム１は、情報処理装置１１１と、１乃至複数の記憶部と、情報処理装置１１１と記憶部とを接続するインタフェース１９とを含む。本実施形態では、記憶部として、不揮発性記憶装置であるＳＳＤ（Solid State Drive）を用いるが、たとえば、ハードディスクドライブ（ＨＤＤ）、ハイブリッドディスクドライブ、ＳＤカード、ＵＳＢメモリ、ＮＡＮＤ型フラッシュメモリチップ、磁気テープなど、他の記憶装置であってもよい。１つの記憶装置に含まれる複数の記憶領域をそれぞれ別の記憶部として用いてもよい。また、本実施形態では、インタフェース１９として、ＳＡＴＡ（Serial Advanced Technology Attachment）インタフェースを用いた場合について説明するが、ＰＣＩ Ｅｘｐｒｅｓｓ(Peripheral Component Interconnect Express, PCIe)、ＵＳＢ（Universal Serial Bus）、ＳＡＳ（Serial Attached SCSI）、Thunderbolt（登録商標）、イーサネ

ット（登録商標）、ファイバーチャネルなどが使用可能である。ＣＰＵ（制御回路）５が情報処理装置１１１における中央演算処理装置であり、情報処理装置１１１における種々の演算及び制御はＣＰＵ５によって行われる。ＣＰＵ５とチップセット７はたとえばＤＭＩ（Ｄｉｒｅｃｔ　Ｍｅｄｉａ　Ｉｎｔｅｒｆａｃｅ）などのインタフェースで接続され、ＣＰＵ５はチップセット７を介して記憶部２や、ＤＶＤドライブなどの光学ドライブ１０の制御を行う。ＣＰＵ５は主メモリ６の制御を行う。主メモリ６としては、たとえばＤＲＡＭ(Dynamic Random Access Memory)やＭＲＡＭ(Magnetoresistive Random Access Memory)やＲｅＲＡＭ(Resistance Random Access Memory)やＦｅＲＡＭ(Ferroelectric Random Access Memory)を採用してよい。

【００１２】

ユーザは、キーボード１４やマウス１５などの入力装置を通して情報処理装置１１１の制御を行い、キーボード１４やマウス１５からの信号は、たとえばＵＳＢ（Universal Serial Bus）コントローラ１３及びチップセット７を介してＣＰＵ５で処理される。ＣＰＵ５は、表示コントローラ８を介してディスプレイ（表示装置）９に画像データやテキストデータなどを送る。ユーザは、ディスプレイ９を介して情報処理装置１１１からの画像データやテキストデータなどを視認することができる。

【００１３】

ＣＰＵ５は、情報処理装置１１１の動作を制御するために設けられたプロセッサであり、たとえば記憶部２から主メモリ６にロードされるオペレーティングシステム（ＯＳ）１００を実行する。更に、光学ドライブ１０が、装填された光ディスクに対して読出し処理及び書込み処理の少なくとも１つの処理の実行を可能にした場合に、ＣＰＵ５は、それらの処理を実行する。また、ＣＰＵ５は、ＲＯＭ１１に格納されたＵＥＦＩ(Unified Extensible Firmware Interface)ファームウェアやシステムＢＩＯＳ (Basic Input/Output System)などを実行する。尚、ＵＥＦＩファームウェアやシステムＢＩＯＳは、情報処理装置１１１内のハードウェア制御のためのプログラムである。その他、ＣＰＵ５は、チップセット７を介してネットワークコントローラ１２を制御する。ネットワークコントローラ１２としては、たとえばＬＡＮ（Local Area Network）コントローラ、無線ＬＡＮコントローラなどがある。

【００１４】

主メモリ６は、プログラムやデータを一時的に記憶し、ＣＰＵ５の作業用メモリとして機能する。主メモリ６は、ＯＳ１００を格納する記憶領域６Ａと、制御プログラム２００を格納する記憶領域６Ｂを含んでいる。ＯＳ１００は、たとえばＷｉｎｄｏｗｓ（登録商標）として一般的に知られているように、情報処理装置１１１の入出力装置を管理し、ディスクやメモリを管理し、ソフトウェアが情報処理装置１１１のハードウェアや記憶装置を利用可能にするための制御を行うなど、情報処理装置１１１全体を管理するプログラムである。本実施形態では、制御プログラム２００はＯＳ１００の一部の場合である例を示すが、制御プログラム２００はＯＳ１００とは独立したプログラムであってもよい。また、本実施形態では制御プログラム２００はインタフェース１９を介して主メモリ１００に読み出されて実行されるプログラムとして説明するが、制御プログラム２００はＲＯＭ１１に保存され、ＵＥＦＩファームウェアやシステムＢＩＯＳの一部として動作してもよい。また制御プログラム２００はハードウェアとして情報処理装置１１１内に実装されてもよい。主メモリ６には、メタデータ３００と、論理ドライブステータステーブル４５０が格納されている。

【００１５】

表示コントローラ８は、情報処理装置１１１のディスプレイ９を制御するビデオ再生コントローラである。チップセット７は、ＣＰＵ５のローカルバスと接続するブリッジデバイスである。チップセット７は、各種ソフトウェア及びデータを格納する記憶装置である記憶部２や記憶部３を、インタフェース１９を介して制御する。なお、記憶部２および記憶部３は、チップセット７を経由してＣＰＵ５に接続されてもよいし、ＣＰＵ５に直接接続されても良い。

【００１６】
情報処理装置１１１では、論理セクタ単位で記憶部２や記憶部３へのアクセスを行う。インタフェース１９を介して、書き込みコマンド（ライトコマンド、ライト要求）、読み出しコマンド（リードコマンド、リード要求）、フラッシュコマンド等が記憶部２や記憶部３に入力される。
【００１７】
また、チップセット７は、ＲＯＭ１１、光学ドライブ１０、ネットワークコントローラ１２、ＵＳＢコントローラ１３をアクセス制御するための機能も有している。ＵＳＢコントローラ１３にキーボード１４、マウス１５が接続されている。
【００１８】
（制御プログラムの形態）
本実施形態では、制御プログラム２００は、例えば、図２に示すように、情報処理装置１１１が電源オフになっているときは記憶部２のＮＡＮＤ型フラッシュメモリ（ＮＡＮＤメモリ）１６の領域１６Ｂに格納されているが、情報処理装置１１１の起動時またはプログラム起動時に、ＮＡＮＤメモリ１６の領域１６Ｂから主メモリ６上の領域６Ｂにロードされる。一方、情報処理装置１１１に記憶部２や記憶部３とは別の記憶部２０が接続されている場合など、複数の記憶部が接続されている場合は、図３に示すように、制御プログラム２００は、記憶部２０の領域２０Ｂに格納されており、情報処理装置１１１の起動時またはプログラム起動時に、領域２０Ｂから主メモリ６上の領域６Ｂにロードされるようにしてもよい。特に、記憶部２０がＯＳを格納するシステムドライブとして使用されており、記憶部２が文書や静止画データや動画データなどユーザデータを格納するデータドライブとして使用されている場合は、システムドライブとしての記憶部２０にはＯＳやアプリケーションプログラムを主として格納する記憶ドライブとして使用し、データドライブとしての記憶部２にはユーザデータを格納する記憶ドライブとして使用するというように、記憶部２と記憶部２０の役割を明確に分ける観点で、システムドライブとしての記憶部２０に制御プログラム２００を格納することが望ましい。
【００１９】
制御プログラム２００のセットアップをユーザが行う労力を省く観点では、例えば図２や図３に示したように、制御プログラム２００が記憶部２や記憶部２０に格納された状態で情報処理システム１が製造元から出荷され、店頭に並び、ユーザの手に渡ることが望ましい。一方、ユーザが制御プログラム２００のインストールの可否を選択できるようにする観点、およびユーザに最新の制御プログラムを提供できるようにするという観点では、制御プログラム２００は、ＷＥＢからのダウンロード、またはＤＶＤ-ＲＯＭやＵＳＢメモリなど外部記憶媒体からのインストールにより記憶部２や記憶部２０に格納できるようにすることが望ましい。
【００２０】
図４はＷＥＢからのダウンロードの一例である。制御プログラム２００はＷＥＢサーバ２１内の記憶媒体２２の領域２２Ｂに格納されており、制御プログラム２００はインターネットやローカルネットワークや無線ＬＡＮなどのネットワークを介して、たとえばネットワークコントローラ１２を経由して記憶部２のＮＡＮＤメモリ１６上の領域１６Ｂにダウンロード（またはインストール）される。記憶媒体２２は、たとえばＳＳＤ、ＨＤＤ、ハイブリッドディスクドライブ、磁気テープなどが用いられる。なお、図３の場合は、制御プログラム２００は、記憶部２０上の領域２０Ｂにダウンロードまたはインストールされる。
【００２１】
図５は光学メディアからのインストールの一例である。制御プログラム２００はＤＶＤ-ＲＯＭやＣＤ-ＲＯＭやＤＶＤ-ＲＷやＢｌｕ－ｒａｙ（登録商標）Ｄｉｓｃなどの光学メディア２３に格納されており、光学メディア２３が光学ドライブ１０にセットされることで、制御プログラム２００は光学ドライブ１０を介して記憶部２のＮＡＮＤメモリ１６上の領域１６Ｂ（または領域２０Ｂ）にインストールされる。

【００２２】
図６はＵＳＢメモリからのインストールの一例である。制御プログラム２００はＵＳＢメモリ２４の領域２４Ｂに格納されており、ＵＳＢメモリ２４がＵＳＢコントローラ１３に接続されることで、制御プログラム２００はＵＳＢコントローラ１３を介して記憶部２のＮＡＮＤメモリ１６上の領域１６Ｂ（または領域２０Ｂ）にインストールされる。ＵＳＢメモリ２４のかわりに、ＳＤカードなどその他外部メモリであってもよいのはもちろんである。ユーザによる入手容易性の観点から、光学メディア２３やＵＳＢメモリ２４は、情報処理システム１または記憶部２の出荷時に、付属品として情報処理システム１または記憶部２と一緒に梱包されて販売されることが望ましい。一方、光学メディア２３やＵＳＢメモリ２４はソフトウェア商品として単独で販売されてもよいし、雑誌や書籍の付録として付属していてもよい。
【００２３】
（ソフトウェアの階層構造）
図７に、情報処理装置１１１のソフトウェアレベルでの階層構造を示す。主メモリ６上にロードされたアプリケーションプログラム４００は、通常は記憶部２や記憶部３などの記憶部と直接通信せずに、主メモリ６にロードされたＯＳ１００を経由して記憶部と通信する。また、ＯＳ１００はＵＥＦＩファームウェアやシステムＢＩＯＳを経由して記憶部と通信する。後述する寿命到達時処理前において、ＯＳ１００は記憶部２を論理単位としての論理ドライブ４（図３８参照）として認識し、アプリケーションプログラム４００に対して論理ドライブ４をアクセス可能な記憶ドライブとして通知する。アプリケーションプログラム４００が論理ドライブ４に対しリード要求、ライト要求などの命令を送信する必要が生じた場合、アプリケーションプログラム４００は、ＯＳ１００に論理ドライブ４に対するファイル単位でのアクセス要求を送信する。ＯＳ１００は主メモリ６に格納されたメタデータ３００を参照し、アクセス要求のあったファイルに対応する記憶部２の論理アドレス(ＬＢＡ：Logical Block Address)を特定し、記憶部２に対してインタフェース１９を介して命令およびＬＢＡおよびデータを送信する。記憶部２から応答が返ってきた場合、ＯＳ１００はアプリケーションプログラム４００に対して応答を返す。論理ドライブ４が後述する引き継ぎ状態になった時には、ＯＳ１００は記憶部２および記憶部３を論理ドライブ４として認識する。
【００２４】
（情報処理装置の構成）
次に、情報処理システム１の構成例について説明する。図８は、情報処理システム１としてのデスクトップコンピュータの概略図である。デスクトップコンピュータは、コンピュータ本体３１、ディスプレイ９、キーボード１４、及びマウス１５などを備えている。コンピュータ本体３１は、主要なハードウェアが搭載されたマザーボード３０、記憶部２、及び電源装置３２などを備えている。記憶部２は、ＳＡＴＡケーブルを介してマザーボード３０に物理的に接続され、マザーボード３０上に実装されたチップセット７を介して、同じくマザーボード上に実装されたＣＰＵ５に電気的に接続されている。電源装置３２は、デスクトップコンピュータで使用される各種電源を発生し、電源ケーブルを介してマザーボード３０や記憶部２などに電源を供給する。記憶部３は、ＳＡＴＡケーブルを介してマザーボード３０に物理的に接続可能であり、それによりマザーボード３０上に実装されたチップセット７を介して、同じくマザーボード上に実装されたＣＰＵ５に電気的に接続される。
【００２５】
図９は、情報処理システム１としてのポータブルコンピュータの概略図である。ポータブルコンピュータは、コンピュータ本体３４、ディスプレイユニット３５などから構成されている。ディスプレイユニット３５は、例えばＬＣＤ（Liquid Crystal Display）で構成される表示装置９が組み込まれている。ディスプレイユニット３５は、コンピュータ本体３４に対し、この本体３４の上面が露出される開放位置と本体３４の上面を覆う閉塞位置との間を回動自由に取り付けられている。本体３４は薄い箱形の筐体を有しており、そ

の上面には、電源スイッチ３６、キーボード１４、タッチパッド３３等が配置されている。また、本体３４も、デスクトップコンピュータと同様に、記憶部２、マザーボード、及び電源装置などを備えている。

【００２６】

本発明を適用する情報処理システム１としては、他に、スチルカメラ或いはビデオカメラなどの撮像装置などであってもよいし、タブレットコンピュータやスマートフォンやゲーム機器やカーナビゲーションシステムやプリンタ機器やスキャナ機器やサーバーシステムなどであってもよい。

【００２７】

（記憶装置の構成　その１）

次に、記憶部２の構成について説明する。本実施形態では、記憶部２の主な構成要素がＮＡＮＤメモリ１６である場合について説明する。本実施形態では、記憶部２と記憶部３が同一の構成である場合について説明するが、一方、記憶部２は主な構成要素をＮＡＮＤメモリ１６とするＳＳＤであり、記憶部３は主な構成要素を磁気ディスクとするハードディスクドライブである場合など、記憶部３が記憶部２とは異なる構成であっても本発明は適用可能である。記憶部２と記憶部３が統計情報６５を格納していることが望ましいが、記憶部２が統計情報６５を格納しており、記憶部３は統計情報６５を格納していない場合であっても本発明は適用可能である。データ引き継ぎ先である記憶部３の記憶可能容量は、データ引き継ぎ元である記憶部２の記憶可能容量と同じか、記憶部２の記憶可能容量よりも大きいことが望ましいが、記憶部２の記憶可能容量よりも小さい場合であっても本発明は適用可能である。

【００２８】

（ＮＡＮＤメモリの構成）

図１０に、本実施形態において記憶部２および記憶部３の構成要素として使用されるＮＡＮＤメモリ１６を構成するＮＡＮＤメモリチップ８０の内部構成例を示す。ＮＡＮＤメモリ１６は、１以上の数のＮＡＮＤメモリチップ８０よりなる。ＮＡＮＤメモリチップ８０は、複数のメモリセルがマトリクス状に配列されたメモリセルアレイを有する。メモリセルアレイを構成するメモリセルトランジスタは、半導体基板上に形成された積層ゲート構造を備えたＭＯＳＦＥＴ（Metal Oxide Semiconductor Field Effect Transistor）から構成される。積層ゲート構造は、半導体基板上にゲート絶縁膜を介在して形成された電荷蓄積層（浮遊ゲート電極）、及び浮遊ゲート電極上にゲート間絶縁膜を介在して形成された制御ゲート電極を含んでいる。メモリセルトランジスタは、浮遊ゲート電極に蓄えられる電子の数に応じて閾値電圧が変化し、この閾値電圧の違いに応じてデータを記憶する。

【００２９】

本実施形態では、個々のメモリセルが上位ページ及び下位ページを使用して2bit/cellの４値記憶方式の書き込み方式である場合について説明するが、個々のメモリセルが単一ページを使用しての1bit/cellの２値記憶方式の書き込み方式、または上位ページ及び中位ページ及び下位ページを使用しての3 bit/cellの８値記憶方式の書き込み方式である場合、あるいは4bit/cell以上の多値記憶方式の書き込み方式を採用する場合であっても本発明の本質は変わらない。また、メモリセルトランジスタは浮遊ゲート電極を有する構造に限らず、ＭＯＮＯＳ（Metal-Oxide-Nitride-Oxide-Silicon）型など、電荷蓄積層としての窒化界面に電子をトラップさせることで閾値電圧を調整可能な構造であってもよい。ＭＯＮＯＳ型のメモリセルトランジスタについても同様に、１ビットを記憶するように構成されていてもよいし、多値を記憶するように構成されていてもよい。また、不揮発性記憶媒体として、米国特許８,１８９,３９１号明細書や米国特許出願公開第２０１０/０２０７１９５号明細書や米国特許出願公開第２０１０/０２５４１９１号明細書に記述されるような３次元的にメモリセルが配置された半導体記憶媒体であってもよい。

【００３０】

図１０に示すように、ＮＡＮＤメモリチップ８０は、データを記憶するメモリセルをマ

トリックス状に配置してなるメモリセルアレイ８２を備えている。メモリセルアレイ８２は複数のビット線と複数のワード線と共通ソース線を含み、ビット線とワード線の交点に電気的にデータを書き換え可能なメモリセルがマトリクス状に配置されている。このメモリセルアレイ８２には、ビット線を制御するためのビット線制御回路８３、及びワード線電圧を制御するためのワード線制御回路８５が接続されている。すなわち、ビット線制御回路８３は、ビット線を介してメモリセルアレイ８２中のメモリセルのデータを読み出す一方、ビット線を介してメモリセルアレイ８２中のメモリセルに書き込み制御電圧を印加してメモリセルに書き込みを行う。

【００３１】

ビット線制御回路８３には、カラムデコーダ８４、データ入出力バッファ８９及びデータ入出力端子８８が接続されている。メモリセルアレイ８２から読み出されたメモリセルのデータは、ビット線制御回路８３、データ入出力バッファ８９を介してデータ入出力端子８８から外部へ出力される。また、外部からデータ入出力端子８８に入力された書き込みデータは、データ入出力バッファ８９を介して、カラムデコーダ８４によってビット線制御回路８３に入力され、指定されたメモリセルへの書き込みが行われる。

【００３２】

また、メモリセルアレイ８２、ビット線制御回路８３、カラムデコーダ８４、データ入出力バッファ８９、及びワード線制御回路８５は、制御回路８６に接続されている。制御回路８６は、制御信号入力端子８７に入力される制御信号に従い、メモリセルアレイ８２、ビット線制御回路８３、カラムデコーダ８４、データ入出力バッファ８９、及びワード線制御回路８５を制御するための制御信号及び制御電圧を発生させる。ＮＡＮＤメモリチップ８０のうち、メモリセルアレイ８２以外の回路部分をＮＡＮＤコントローラ（ＮＡＮＤＣ）８１と呼ぶ。

【００３３】

図１１は、図１０に示すメモリセルアレイ８２の構成を示している。メモリセルアレイ８２はＮＡＮＤセル型メモリセルアレイであり、複数のＮＡＮＤセルを含んで構成されている。１つのＮＡＮＤセルは、直列接続されたメモリセルからなるメモリストリングＭＳと、その両端に接続される選択ゲートＳ１、Ｓ２とにより構成されている。選択ゲートＳ１はビット線ＢＬに接続され、選択ゲートＳ２はソース線ＳＲＣに接続されている。同一のロウに配置されたメモリセルＭＣの制御ゲートはワード線ＷＬ０～ＷＬm-１に共通接続されている。また、第１の選択ゲートＳ１はセレクト線ＳＧＤに共通接続され、第２の選択ゲートＳ２はセレクト線ＳＧＳに共通接続されている。

【００３４】

メモリセルアレイ８２は、１または複数のプレーンを含んでおり、プレーンは複数のブロックを含んでいる。各ブロックは、複数のＮＡＮＤセルにより構成され、このブロック単位でデータが消去される。

【００３５】

また、１つのワード線に接続された複数のメモリセルは、１物理セクタを構成する。この物理セクタ毎にデータが書き込まれ、読み出される。この物理セクタは、後述するＬＢＡの論理セクタとは無関係である。１物理セクタには、２bit/cell書き込み方式（４値）の場合例えば２物理ページ（２ページ）分のデータが記憶される。一方、１bit/cell書き込み方式（２値）の場合は、１物理セクタに例えば１物理ページ（１ページ）分のデータが記憶され、３ bit/cell書き込み方式（８値）の場合、１物理セクタに例えば３物理ページ（３ページ）分のデータが記憶される。

【００３６】

リード動作、プログラムベリファイ動作及びプログラム動作時において、後述のＳＳＤＣ４１から受信した例えばRow Address等の物理アドレスに応じて、１本のワード線が選択され、１物理セクタが選択される。この物理セクタ内のページの切り替えは物理アドレスによって行われる。本実施形態ではＮＡＮＤメモリ１６は２bit/cell書き込み方式であり、ＳＳＤＣ４１は物理セクタには上位ページ（Upper Page）と下位ページ（Lower Page

）の２ページが物理ページとして割り当てられているとして取扱い、それら全ページに対して物理アドレスが割り当てられている。

【００３７】

２bit/cellの４値ＮＡＮＤメモリは、１つのメモリセルにおける閾値電圧が、４通りの分布を持ち得るように構成されている。図１２は、４値ＮＡＮＤセル型フラッシュメモリのメモリセルに記憶される２ビットの４値データ（データ“１１”、“０１”、“１０”、“００”）とメモリセルの閾値電圧分布との関係を示している。なお、図１２において、VA1は、下位ページのみ書き込み済みで上位ページが未書き込みの物理セクタについて、２つのデータを読み出す場合に選択ワード線に印加される電圧であり、VA1Vは、Ａ１への書き込みを行う場合において、書き込みが完了したかどうかを確認するために印加されるベリファイ電圧を示している。

【００３８】

また、VA2、VB2、VC2は、下位ページと上位ページが書き込み済みの物理セクタについて、４つのデータを読み出す場合に選択ワード線に印加される電圧であり、VA2V、VB2V、VC2Vは、各閾値電圧分布への書き込みを行う場合において、書き込みが完了したかどうかを確認するために印加されるベリファイ電圧を示している。また、Vread1、Vread2は、データの読み出しを行う場合に、ＮＡＮＤセル中の非選択メモリセルに対し印加され、その保持データに拘わらず当該非選択メモリセルを導通させる読み出し電圧を示している。さらに、Vev1、Vev2は、メモリセルのデータを消去する場合において、その消去が完了したか否かを確認するためメモリセルに印加される消去ベリファイ電圧であり、負の値を有する。その大きさは、隣接メモリセルの干渉の影響を考慮して決定される。上述の各電圧の大小関係は、

Vev1＜VA1＜VA1V＜Vread1

Vev2＜VA2＜VA2V＜VB2＜VB2V＜VC2＜VC2V＜Vread2

である。

【００３９】

なお、消去ベリファイ電圧Vev1、Vev2、Vev3は前述の通り負の値であるが、実際に消去ベリファイ動作においてメモリセルＭＣの制御ゲートに印加される電圧は、負の値ではなく、ゼロ又は正の値である。すなわち、実際の消去ベリファイ動作においては、メモリセルＭＣのバックゲートに正の電圧を与え、メモリセルＭＣの制御ゲートには、ゼロ又はバックゲート電圧より小さい正の値の電圧を印加している。換言すれば、消去ベリファイ電圧Vev1、Vev2、Vev3は、等価的に負の値を有する電圧である。

【００４０】

ブロック消去後のメモリセルの閾値電圧分布ＥＲは、その上限値も負の値であり、データ“１１”が割り当てられる。下位ページおよび上位ページ書き込み状態のデータ“１１”、“０１”、“１０”、“００”のメモリセルは、それぞれ正の閾値電圧分布ＥＲ２、Ａ２、Ｂ２、Ｃ２を有し（Ａ２、Ｂ２、Ｃ２の下限値も正の値である）、データ“０１”の閾値電圧分布Ａ２が最も電圧値が低く、データ“００”の閾値電圧分布Ｃ２が最も電圧値が高く、各種閾値電圧分布の電圧値はＡ２＜Ｂ２＜Ｃ２の関係を有する。下位ページ書き込みかつ上位ページ未書き込み状態のデータ“１０”のメモリセルは、正の閾値電圧分布Ａ１を有する（Ａ１の下限値も正の値である）。なお、図１２に示す閾値電圧分布はあくまでも一例であって、本発明はこれに限定されるものではない。例えば、図１２では閾値電圧分布Ａ２，Ｂ２，Ｃ２は全て正の閾値電圧分布であるとして説明したが、閾値電圧分布Ａ２は負の電圧の分布であり、閾値電圧分布Ｂ２、Ｃ２が正の電圧の分布であるような場合も、本発明の範囲に含まれる。また、閾値電圧分布ＥＲ１・ＥＲ２は正の値であったとしても、本発明はこれに限定されるものではない。また、本実施形態ではＥＲ２、Ａ２、Ｂ２、Ｃ２のデータの対応関係がそれぞれ“１１”、“０１”、“１０”、“００”であるとしているが、たとえばそれぞれ“１１”、“０１”、“００”、“１０”であるような他の対応関係であってもよい。

【００４１】

１つのメモリセルの２ビットデータは、下位ページデータと上位ページデータからなり、下位ページデータと上位ページデータは別々の書き込み動作、つまり、２回の書き込み動作により、メモリセルに書き込まれる。データを“＊@”と標記するとき、＊は上位ページデータを、@は下位ページデータを表している。

【００４２】

まず、下位ページデータの書き込みを、図１２の１段目～２段目を参照して説明する。全てのメモリセルは、消去状態の閾値電圧分布ＥＲを有し、データ“１１”を記憶しているものとする。図１２に示すように、下位ページデータの書き込みを行うと、メモリセルの閾値電圧分布ＥＲは、下位ページデータの値（“１”、或いは“０”）に応じて、２つの閾値電圧分布（ＥＲ１、Ａ１）に分けられる。下位ページデータの値が“１”の場合には、消去状態の閾値電圧分布ＥＲを維持するのでＥＲ１＝ＥＲであるが、ＥＲ１＞ＥＲであってもよい。

【００４３】

一方、下位ページデータの値が“０”の場合には、メモリセルのトンネル酸化膜に高電界を印加し、フローティングゲート電極に電子を注入して、メモリセルの閾値電圧Ｖthを所定量だけ上昇させる。具体的には、ベリファイ電位ＶA1Vを設定し、このベリファイ電圧ＶA1V以上の閾値電圧となるまで書き込み動作が繰り返される。その結果、メモリセルは、書き込み状態（データ“１０”）に変化する。書き込み動作を所定回繰り返しても閾値電圧に到達しなかった場合（または閾値電圧に到達しないメモリセル数が所定値以上の場合）、当該物理ページに対する書き込みは「書き込みエラー」(Program Error, Program Fail)となる。

【００４４】

次に、上位ページデータの書き込みを、図１２の２段目～３段目を参照して説明する。上位ページデータの書き込みは、チップの外部から入力される書き込みデータ（上位ページデータ）と、メモリセルに既に書き込まれている下位ページデータとに基づいて行われる。

【００４５】

即ち、図１２の２段目～３段目に示すように、上位ページデータの値が“１”の場合には、メモリセルのトンネル酸化膜に高電界がかからないようにし、メモリセルの閾値電圧Ｖthの上昇を防止する。その結果、データ“１１”（消去状態の閾値電圧分布ＥＲ１）のメモリセルは、データ“１１”をそのまま維持し（ＥＲ２）、データ“１０”（閾値電圧分布Ａ１）のメモリセルは、データ“１０”をそのまま維持する（Ｂ２）。ただし、各分布間の電圧マージンを確保するという点で、上述のベリファイ電圧ＶA1Vよりも大きい正のベリファイ電圧ＶB2Vを用いて閾値電圧分布の下限値を調整し、これにより閾値電圧分布の幅を狭めた閾値電圧分布Ｂ２を形成するのが望ましい。下限値調整を所定回繰り返しても閾値電圧に到達しなかった場合（または閾値電圧に到達しないメモリセル数が所定値以上の場合）、当該物理ページに対する書き込みは「書き込みエラー」となる。

【００４６】

一方、上位ページデータの値が“０”の場合には、メモリセルのトンネル酸化膜に高電界を印加し、フローティングゲート電極に電子を注入して、メモリセルの閾値電圧Ｖthを所定量だけ上昇させる。具体的には、ベリファイ電位ＶA2V、ＶC2Vを設定し、このベリファイ電圧ＶA1V以上の閾値電圧となるまで書き込み動作が繰り返される。その結果、データ“１１”（消去状態の閾値電圧分布ＥＲ１）のメモリセルは、閾値電圧分布Ａ２のデータ“０１”に変化し、データ“１０”（Ａ１）のメモリセルは、閾値電圧分布Ｃ２のデータ“００”に変化する。このとき、ベリファイ電圧ＶA2V、ＶC2Vが用いられて、閾値電圧分布Ａ２、Ｃ２の下限値が調整される。書き込み動作を所定回繰り返しても閾値電圧に到達しなかった場合（または閾値電圧に到達しないメモリセル数が所定値以上の場合）、当該物理ページに対する書き込みは「書き込みエラー」となる。

【００４７】

一方、消去動作においては、消去ベリファイ電位Ｖevを設定し、このベリファイ電圧Ｖ

ev以下の閾値電圧となるまで消去動作が繰り返される。その結果、メモリセルは、書き込み状態（データ“００”）に変化する。消去動作を所定回繰り返しても閾値電圧に到達しなかった場合（または閾値電圧に到達しないメモリセル数が所定値以上の場合）、当該物理ページに対する消去は「消去エラー」となる。

【００４８】

以上が、一般的な４値記憶方式におけるデータ書き込み方式の一例である。３ビット以上の多ビット記憶方式においても、上記の動作に更に上位のページデータに応じ、閾値電圧分布を８通り以上に分割する動作が加わるのみであるので、基本的な動作は同様である。

【００４９】

（記憶装置の構成　その２）

つぎに、記憶部２および記憶部３の構成例について説明する。本実施形態では、図１３に示すように、ＳＳＤとしての記憶部２および記憶部３は、不揮発性半導体メモリとしてのＮＡＮＤ型フラッシュメモリ（以下ＮＡＮＤメモリと略す）１６と、インタフェース１９を介して情報処理装置１１１との信号の送受信を行うインタフェースコントローラ（ＩＦＣ）４２と、ＩＦＣ４２とＮＡＮＤメモリ１６との中間バッファとして機能するキャッシュメモリ（ＣＭ）４６を有する半導体メモリとしてのＲＡＭ（Random Access Memory）４０と、ＮＡＮＤメモリ１６及びＲＡＭ４０の管理、制御、及びインタフェースコントローラ４２の制御を司るＳＳＤコントローラ（ＳＳＤＣ）４１と、これら構成要素を接続するバス４３を備える。

【００５０】

ＲＡＭ４０としては、たとえば、ＤＲＡＭ（Dynamic Random Access Memory）、ＳＲＡＭ（Static Random Access Memory）などの揮発性ＲＡＭや、ＦｅＲＡＭ（Ferroelectric Random Access Memory）、ＭＲＡＭ（Magnetoresistive Random Access Memory）、ＰＲＡＭ（Phase Change Random Access Memory）、ＲｅＲＡＭ（Resistance Random Access Memory）などの不揮発性ＲＡＭを採用することができる。ＲＡＭ４０はＳＳＤＣ４１に含まれてもよい。

【００５１】

ＮＡＮＤメモリ１６は、複数のＮＡＮＤメモリチップ８０からなり、情報処理装置１１１によって指定されたユーザデータを記憶したり、ユーザデータを管理する管理テーブルを記憶したり、ＲＡＭ４０で管理される管理情報をバックアップ用に記憶したりする。ＮＡＮＤメモリ１６は、複数のメモリセルがマトリクス状に配列されたメモリセルアレイ８２を有し、個々のメモリセルは上位ページ及び下位ページを使用して多値記憶が可能である。ＮＡＮＤメモリ１６は、複数のメモリチップによって構成され、各メモリチップは、データ消去の単位であるブロックを複数配列して構成される。また、ＮＡＮＤメモリ１６では、ページごとにデータの書き込み及びデータの読み出しが行われる。ブロックは、複数のページによって構成されている。

【００５２】

ＲＡＭ４０は、情報処理装置１１１とＮＡＮＤメモリ１６間でのデータ転送用キャッシュとして機能するキャッシュメモリ（ＣＭ）４６を有する。また、ＲＡＭ４０は、管理情報記憶用メモリ及び作業領域用メモリとして機能する。ＲＡＭ４０の領域４０Ａで管理される管理テーブルは、ＮＡＮＤメモリ１６の領域４０Ｍに記憶されている各種管理テーブルが記憶部２および記憶部３の起動時などに展開されたものであり、定期的、スタンバイコマンド受信時、フラッシュコマンド受信時、あるいは電源断時にＮＡＮＤメモリ１６の領域４０Ｍに退避保存される。

【００５３】

ＳＳＤＣ４１は、ＮＡＮＤメモリ１６に記憶されたシステムプログラム（ファームウエア）を実行するプロセッサと、各種ハードウェア回路などによってその機能が実現され、情報処理装置１１１からのライト要求、キャッシュフラッシュ要求、リード要求等の各種コマンドに対する情報処理装置１１１－ＮＡＮＤメモリ１６間のデータ転送制御、ＲＡＭ

４０及びＮＡＮＤメモリ１６に記憶された各種管理テーブルの更新・管理、ＮＡＮＤメモリ１６に書き込むデータのＥＣＣ符号化、ＮＡＮＤメモリ１６から読み出したデータのＥＣＣ復号化などを実行する。

【００５４】

情報処理装置１１１は記憶部２に対し、リード要求またはライト要求を発行する際には、インタフェース１９を介して論理アドレスとしてのＬＢＡを入力する。ＬＢＡは、論理セクタ（サイズ：例えば５１２Byte）に対して０からの通し番号をつけた論理アドレスである。また、情報処理装置１１１は記憶部２に対し、リード要求またはライト要求を発行する際には、ＬＢＡと併せて、リード要求またはライト要求の対象となる論理セクタサイズを入力する。

【００５５】

ＩＦＣ４２は、情報処理装置１１１からのリード要求、ライト要求、その他要求及びデータを受信し、受信した要求やデータをＳＳＤＣ４１に送信したり、ＳＳＤＣ４１の制御によりＲＡＭ４０にデータを送信したりする機能を持つ。

【００５６】

図１４に記憶部２および記憶部３で使用する管理情報４４の構成例を示す。これらの管理情報４４は、前述したように、ＮＡＮＤメモリ１６の領域４０Ｍで不揮発記憶されている。領域４０Ｍで記憶された管理情報が記憶部２の起動時にＲＡＭ４０の領域４０Ａに展開されて使用される。領域４０Ａの管理情報４４は、定期的あるいは電源断時に領域４０Ｍに退避保存される。ＲＡＭ４０がＭＲＡＭやＦｅＲＡＭなどのような不揮発ＲＡＭである場合には、この管理情報４４はＲＡＭ４０にのみ記憶されるようにしてもよく、この場合はこの管理情報４４はＮＡＮＤメモリ１６には記憶されない。ＮＡＮＤメモリ１６への書き込み量を少なくするためには、管理情報４４に記憶されるデータは、ＲＡＭ４０の領域４０Ａに記憶されているデータを圧縮したものであることが望ましい。また、ＮＡＮＤメモリ１６への書き込み頻度を少なくするためには、管理情報４４には、ＲＡＭ４０の領域４０Ａに記憶されている管理情報４４の更新情報（差分情報）を追記するようにすることが望ましい。

【００５７】

図１４に示すように、管理情報は、フリーブロックテーブル（ＦＢＴ）６０と、バッドブロックテーブル（ＢＢＴ）６１と、アクティブブロックテーブル（ＡＢＴ）６２と、トラックテーブル（トラック単位の論物変換テーブル）６３と、クラスタテーブル（クラスタ単位の論物変換テーブル）６４、統計情報６５を含む。

【００５８】

ＬＢＡは、図１５に示すように、論理セクタ（サイズ：例えば５１２Byte）に対して０からの通し番号をつけた論理アドレスである。本実施形態においては、記憶部２の論理アドレス（ＬＢＡ）の管理単位として、ＬＢＡの下位（ｓ＋１）ビット目から上位のビット列で構成されるクラスタアドレスと、ＬＢＡの下位（ｓ＋ｔ＋１）ビットから上位のビット列で構成されるトラックアドレスとを定義する。すなわち、論理セクタは、情報処理装置１１１からの最小アクセス単位である。クラスタは、ＳＳＤ内部で「小さなデータ」を管理する管理単位であり、論理セクタサイズの自然数倍がクラスタサイズとなるように定められる。また、トラックは、ＳＳＤ内部で「大きなデータ」を管理する管理単位であり、クラスタサイズの２以上の自然数倍がトラックサイズとなるように定められる。したがって、トラックアドレスはＬＢＡをトラックサイズで割ったものであり、トラック内アドレスはＬＢＡをトラックサイズで割った余りであり、クラスタアドレスはＬＢＡをクラスタサイズで割ったものであり、クラスタ内アドレスはＬＢＡをクラスタサイズで割った余りである。以下の説明では、便宜上、トラックのサイズは１物理ブロックに記録可能なデータのサイズに等しい（物理ブロックにＳＳＤＣ４１で行うＥＣＣ処理の冗長ビットが含まれる場合はこれを除いたサイズ）場合について述べ、クラスタのサイズは１物理ページに記録可能なデータのサイズに等しい（物理ページにＳＳＤＣ４１で行うＥＣＣ処理の冗長ビットが含まれる場合はこれを除いたサイズ）場合について述べる。

【００５９】
フリーブロックテーブル（ＦＢＴ）６０は、ＮＡＮＤメモリ１６に書き込みを行うときに書き込み用に新規に割り当てることのできるＮＡＮＤメモリの用途未割り当ての物理ブロック（フリーブロック：ＦＢ）のブロックアドレス（物理ブロックＩＤ）を管理する。また、物理ブロックＩＤ毎に消去回数を管理しており、物理ブロックが消去されたとき、当該ブロックの消去回数をインクリメントする。
【００６０】
バッドブロックテーブル（ＢＢＴ）６１は、誤りが多いなど記憶領域として使用できない物理ブロックとしてのバッドブロック（ＢＢ）のブロックＩＤを管理する。ＦＢＴ６０と同様に物理ブロックＩＤ毎に消去回数を管理するようにしてもよい。
【００６１】
アクティブブロックテーブル（ＡＢＴ）６２は、用途が割り当てられた物理ブロックであるアクティブブロック（ＡＢ）を管理する。また、物理ブロックＩＤ毎に消去回数を管理しており、物理ブロックが消去されたとき、当該ブロックの消去回数をインクリメントする。
【００６２】
トラックテーブル６３は、トラックアドレスと、このトラックアドレスに対応するトラックデータが記憶される物理ブロックＩＤとの対応関係を管理する。
【００６３】
クラスタテーブル６４は、クラスタアドレスと、このクラスタアドレスに対応するクラスタデータが記憶される物理ブロックＩＤと、このクラスタアドレスに対応するクラスタデータが記憶される物理ブロック内ページアドレスとの対応関係を管理する。
【００６４】
ＳＳＤＣ４１は、統計情報６５に、記憶部２の信頼性に関わる種々のパラメータ（Ｘ０１～Ｘ３２）を信頼性情報として格納する（図２７参照）。
【００６５】
信頼性情報の例として用いられる統計情報６５の値（Raw Value）としては、バッドブロック数総計（統計情報Ｘ０１）、バッド論理セクタ数総計（統計情報Ｘ０２）、消去回数総計（統計情報Ｘ０３）、消去回数平均値（統計情報Ｘ０４）、ＮＡＮＤメモリの書き込みエラー発生回数累積値（統計情報Ｘ０５）、ＮＡＮＤメモリの消去エラー発生回数累積値（統計情報Ｘ０６）、読み出し論理セクタ数総計（統計情報Ｘ０７）、書き込み論理セクタ数総計（統計情報Ｘ０８）、ＥＣＣ訂正不能回数総計（統計情報Ｘ０９）、リトライ読み出し発生回数（統計情報Ｘ１０）、nビット～mビットＥＣＣ訂正単位総計数（統計情報Ｘ１１）、インタフェース１９のデータ化けエラー発生回数（統計情報Ｘ１２）、インタフェース１９の通信速度ダウンシフト回数（統計情報Ｘ１３）、インタフェース１９のレーン数ダウンシフト回数（統計情報Ｘ１４）、インタフェース１９のエラー発生回数（統計情報Ｘ１５）、ＲＡＭ４０のエラー発生回数（統計情報Ｘ１６）、記憶部２の使用時間総計（統計情報Ｘ１７）、起動回数(統計情報Ｘ１８）、不正電源断発生回数(統計情報Ｘ１９)、温度が推奨動作温度の最高値を上回った時間累計（統計情報Ｘ２０）、温度が推奨動作温度の最低値を下回った時間累計（統計情報Ｘ２１）、コマンドの応答時間最大値（統計情報Ｘ２２）、コマンドの応答時間平均値（統計情報Ｘ２３）、ＮＡＮＤメモリの応答時間最大値（統計情報Ｘ２４）、ＮＡＮＤの応答時間平均値（統計情報Ｘ２５）、現在温度（統計情報Ｘ２６）、最高温度（統計情報Ｘ２７）、最低温度（統計情報Ｘ２８）、管理情報冗長度（統計情報Ｘ２９）、ＲＡＭ４０への書き込みデータ量合計（統計情報Ｘ３０）、統計情報増加率（統計情報Ｘ３１）、ＮＡＮＤ整理失敗フラグ（統計情報Ｘ３２）などが含まれる。
【００６６】
バッドブロック数総計（統計情報Ｘ０１）について説明する。記憶部２内のＮＡＮＤメモリ１６の物理ブロックが一つバッドブロックに追加されるごとに統計情報Ｘ０１が１インクリメントされる。統計情報Ｘ０１は記憶部２の製造時（検査工程前）にゼロにリセッ

トされていることが望ましく、検査工程でエラーが発生したり、閾値分布の分布間マージンが少ないことが判明したブロックは、あらかじめバッドブロックに加えておくことがさらに望ましい。ＳＳＤＣ４１は統計情報Ｘ０１を統計情報６５に格納せず、統計情報Ｘ０１をＢＢＴ６１から直接計算してもよい。統計情報Ｘ０１が大きいほど信頼性が悪化していることを示す。

【００６７】

バッド論理セクタ数総計（統計情報Ｘ０２）について説明する。ＳＳＤＣ４１は情報処理装置１１１から読み出し命令とＬＢＡを受信してＮＡＮＤ型フラッシュメモリ１６を読みだした時に読み出しデータをＥＣＣ訂正できなかった場合、当該ＬＢＡをバッド論理セクタとして管理情報４４内のバッド論理セクタテーブルに登録しても良い（図２５参照）。ＳＳＤＣ４１は、バッド論理セクタテーブルに登録されているＬＢＡの数をバッド論理セクタ数総計（統計情報Ｘ０２）として統計情報６５に格納する。情報処理装置１１１から読み出し命令を受信した場合には、ＳＳＤＣ４１はＲＡＭ４０上のバッド論理セクタテーブルを読みだして受信したＬＢＡをバッド論理セクタテーブルから検索し、バッド論理セクタテーブルから見つかった場合には、ＮＡＮＤ型フラッシュメモリ１６を読み出すことなく情報処理装置１１１に読み出しエラーを通知する。ＳＳＤＣ４１はバッド論理セクタのＬＢＡに対して情報処理装置１１１から書き込み命令を受信して書き込み処理を行った場合、ＳＳＤＣ４１は書き込まれたＬＢＡをバッド論理セクタテーブルから削除する。ＳＳＤＣ４１はバッド論理セクタのＬＢＡに対して情報処理装置１１１から削除通知を受信して削除通知処理を行った場合、ＳＳＤＣ４１は削除通知処理されたＬＢＡをバッド論理セクタテーブルから削除する。ＳＳＤＣ４１は情報処理装置１１１から記憶部２の消去コマンド（Secure Eraseコマンド）を受信した場合にはバッド論理セクタテーブルを消去する。記憶部２の消去コマンドとして、たとえばACS-3のF4h Security Erase Unitコマンドや、NVM Express Revision 1.1の80h Format NVMコマンドを用いてもよい。なお、ＬＢＡ単位（論理セクタ単位）でバッド論理セクタテーブルを管理する代わりに、図２６に示すように、クラスタ単位でバッド論理セクタテーブルをバッドクラスタテーブルとして管理しても良い。ＳＳＤＣ４１はバッド論理セクタテーブルに登録されているＬＢＡの個数またはバッドクラスタテーブルに登録されているクラスタアドレスの個数を統計情報Ｘ０２として管理する。ＳＳＤＣ４１は統計情報Ｘ０２を統計情報６５に格納せず、統計情報Ｘ０２をバッド論理セクタテーブルやバッドクラスタテーブルから直接計算してもよい。統計情報Ｘ０２が大きいほど信頼性が悪化していることを示す。

【００６８】

消去回数総計（統計情報Ｘ０３）について説明する。統計情報Ｘ０３は記憶部２内のＮＡＮＤメモリ１６の全ブロックの消去回数累計値を示す。ＳＳＤＣ４１は、記憶部２内のＮＡＮＤメモリ１６の物理ブロックが一つ消去されるごとに統計情報Ｘ０３を１インクリメントする。統計情報Ｘ０３は記憶部２の製造時（検査工程時）にゼロにリセットされていることが望ましい。ＳＳＤＣ４１は統計情報Ｘ０３を統計情報６５に格納せずＦＢＴ６０、ＢＢＴ６１、ＡＢＴ６２から直接計算してもよい。統計情報Ｘ０３が大きいほど信頼性が悪化していることを示す。

【００６９】

消去回数平均値（統計情報Ｘ０４）について説明する。ＳＳＤＣ４１は、記憶部２内のＮＡＮＤメモリ１６の全ブロックについて１ブロックあたりの消去回数平均値を計算し、統計情報Ｘ０４として統計情報６５に格納する。ＳＳＤＣ４１は、管理情報４４を格納するブロックなど一部のブロックを統計情報Ｘ０４の集計対象から除外してもよい。統計情報Ｘ０４は記憶部２の製造時（検査工程前）にゼロにリセットされていることが望ましい。ＳＳＤＣ４１は統計情報Ｘ０４を統計情報６５に格納せずＦＢＴ６０、ＢＢＴ６１、ＡＢＴ６２から直接計算してもよい。ＳＳＤＣ４１は、統計情報Ｘ０４として、消去回数平均値のかわりに、消去回数最大値や消去回数最小値を使用してもよい。統計情報Ｘ０４が大きいほど信頼性が悪化していることを示す。

【００７０】

ＮＡＮＤメモリの書き込みエラー発生回数累積値（統計情報Ｘ０５）について説明する。ＳＳＤＣ４１は、記憶部２内のＮＡＮＤメモリ１６で書き込みエラーが１書き込み単位で発生するごとに、統計情報Ｘ０５を１加算する（ブロック単位で加算してもよい）。統計情報Ｘ０５は記憶部２の製造時（検査工程前）にゼロにリセットされていることが望ましい。統計情報Ｘ０５が大きいほど信頼性が悪化していることを示す。

【００７１】

ＮＡＮＤメモリの消去エラー発生回数累積値（統計情報Ｘ０６）について説明する。統計情報Ｘ０６は記憶部２製造時（検査工程前）にゼロにリセットされていることが望ましい。ＳＳＤＣ４１はは、記憶部２内のＮＡＮＤメモリ１６で消去エラーが１ブロックで発生するごとに統計情報Ｘ０６を１加算する。ＳＳＤＣ４１は、複数のブロックをまとめて消去単位とし、この消去単位１つで消去エラーが発生するごとに統計情報Ｘ０６に１加算するようにしてもよい。統計情報Ｘ０６が大きいほど信頼性が悪化していることを示す。

【００７２】

読み出し論理セクタ数総計（統計情報Ｘ０７）について説明する。ＳＳＤＣ４１は、ＩＦＣ４２が読み出しデータとして情報処理装置１１１に送信したデータの論理セクタ数合計を、統計情報Ｘ０７として統計情報６５に格納する。統計情報Ｘ０７は記憶部２の製造時（検査工程前）にゼロにリセットされていることが望ましい。統計情報Ｘ０７が大きいほど信頼性が悪化していることを示す。

【００７３】

書き込み論理セクタ数総計（統計情報Ｘ０８）について説明する。ＳＳＤＣ４１は、ＩＦＣ４２が書き込みデータとして情報処理装置１１１から受信したデータの論理セクタ数合計を、統計情報Ｘ０８として統計情報６５に格納する。統計情報Ｘ０８は記憶部２の製造時（検査工程前）にゼロにリセットされていることが望ましい。統計情報Ｘ０８が大きいほど信頼性が悪化していることを示す。

【００７４】

ＥＣＣ訂正不能回数総計（統計情報Ｘ０９）について説明する。ＳＳＤＣ４１は、ＥＣＣ訂正によりエラービットが修復できなかった場合に、１読み出し単位ごとに統計情報Ｘ０９を１インクリメントする。ＳＳＤＣ４１は、エラー訂正できなかったエラービット数の推定値を加算するようにしてもよいし、エラー訂正できなかったブロックの数を加算するようにしてもよい。統計情報Ｘ０９は記憶部２製造時（検査工程前）にゼロにリセットされていることが望ましい。統計情報Ｘ０９が大きいほど信頼性が悪化していることを示す。

【００７５】

リトライ読み出し発生回数総計（統計情報Ｘ１０）について説明する。ＳＳＤＣ４１は、データ読み出し時にエラービット数が多く誤り訂正が不可能（ＥＣＣエラー）であった場合、リトライ読み出しを行い、再度ＥＣＣを用いた誤り訂正を実行することが望ましい。特に、ＳＳＤＣ４１は、リトライ読み出し時に図１２の読み出しレベルＶＡ１、ＶＡ２、ＶＢ２やＶＣ２をデフォルトの値からシフトして読み出しを行うことで、誤り訂正不可能であったデータが誤り訂正可能になる場合がある。ＳＳＤＣ４１はリトライ読み出し発生回数を統計情報Ｘ１０として統計情報Ｘ０９に格納し、寿命到達予測や寿命到達判定に用いてもよい。統計情報Ｘ１０は記憶部２の製造時（検査工程前）にゼロにリセットされていることが望ましい。統計情報Ｘ１０が大きいほど信頼性が悪化していることを示す。

【００７６】

ｎビット～ｍビットＥＣＣ訂正単位総計数（統計情報Ｘ１１）について説明する。ｎ，ｍは自然数で、０≦ｎ≦ｍ≦訂正可能最大ビット数である。ＳＳＤＣ４１は、ＥＣＣ訂正単位（たとえば物理ページ）に対してＥＣＣ訂正を行った時、全エラービットが正常に修復され、かつ修復されたエラービット数がｎ以上ｍ以下である場合、ＥＣＣ訂正単位１つにつき「ｎビット～ｍビットＥＣＣ訂正単位総計数」を１加算する。ＥＣＣ訂正により１訂正単位につき最大６４ビット訂正可能である場合、たとえば、ＳＳＤＣ４１は、「１ビット～８ビットＥＣＣ訂正単位総計数」「９ビット～１６ビットＥＣＣ訂正単位総計数」

「１７ビット～２４ビットＥＣＣ訂正単位総計数」「２５ビット～３２ビットＥＣＣ訂正単位総計数」「３３ビット～４０ビットＥＣＣ訂正単位総計数」「４１ビット～４８ビットＥＣＣ訂正単位総計数」「４９ビット～５６ビットＥＣＣ訂正単位総計数」「５７ビット～６４ビットＥＣＣ訂正単位総計数」の８つのパラメータを用意し、ＥＣＣ訂正が正常に行われた場合、１ＥＣＣ訂正単位のＥＣＣ訂正につきこれら８つのパラメータのうちいずれか１つに１をインクリメントする。統計情報Ｘ１１は記憶部２の製造時（検査工程前）にゼロにリセットされていることが望ましい。統計情報Ｘ１１が大きいほど信頼性が悪化していることを示す。

【００７７】

インタフェース１９のデータ化けエラー発生回数（統計情報Ｘ１２）について説明する。ＳＳＤＣ４１は、インタフェース１９上で信号のデータ化けを検出するごとに統計情報Ｘ１２を１インクリメントする。インタフェース１９上で送受信されるデータは、たとえばCyclic Redundancy Check(ＣＲＣ)符号やBose-Chaudhuri-Hocquenghem（ＢＣＨ)符号やReed-Solomon（ＲＳ)符号やLow-Density Parity-Check（ＬＤＰＣ)符号などを用いてＳＳＤＣ４１やＩＦＣ４２やチップセット７によって誤り検出や誤り訂正が行われており、誤りが検出された場合や誤り訂正できなかった場合にＳＳＤＣ４１は統計情報Ｘ１２を１インクリメントする。たとえば、インタフェース１９がＳＡＴＡ規格である場合、ＳＡＴＡ規格におけるＲエラー（Reception Error, R_ERR）が１回発生するごとにＳＳＤＣ４１は統計情報Ｘ１２を１インクリメントする。統計情報Ｘ１２として、ＳＡＴＡ規格のPhy Event Countersのカウンタのいずれかを採用してもよい。統計情報Ｘ１２は記憶部２の製造時（検査工程前）にゼロにリセットされていることが望ましい。統計情報Ｘ１２が大きいほど信頼性が悪化していることを示す。

【００７８】

インタフェース１９の通信速度ダウンシフト回数（統計情報Ｘ１３）について説明する。ＳＳＤＣ４１やＩＦＣ４２やチップセット７がインタフェース１９の通信速度が設計値よりも小さくなったことを検出した場合、ＳＳＤＣ４１はＸ１３を１インクリメントする。例えば、通信インタフェース１９やＩＦＣ４２やＳＳＤＣ４１がＳＡＴＡの通信速度が最大６Gbpsで設計されているにも関わらず、実際に記憶部２や情報処理装置１１１との間で確立された通信速度が３Gbpsのようなより低速の通信速度であることが検出された場合、ＳＳＤＣ４１はこれをＳＡＴＡ通信上のエラーとみなし、統計情報Ｘ１３を１インクリメントする。たとえば、インタフェース１９やＩＦＣ４２やＳＳＤＣ４１がＰＣＩ Expressの通信速度が最大８Gbpsとして設計されているにも関わらず、実際に記憶部２や情報処理装置１１１との間で確立された通信速度が５Gbpsのようなより低速の通信規格であることが検出された場合、ＳＳＤＣ４１はこれをＰＣＩExpress通信上のエラーとみなし、統計情報Ｘ１３を１インクリメントする。統計情報Ｘ１３は記憶部２の製造時（検査工程前）にゼロにリセットされていることが望ましい。この値が大きいほど信頼性が悪化していることを示す。

【００７９】

インタフェース１９のレーン数ダウンシフト回数（統計情報Ｘ１４）について説明する。ＳＳＤＣ４１やＩＦＣ４２やチップセット７がインタフェース１９のアクティブな伝送路数が設計値よりも小さくなったことを検出した場合、ＳＳＤＣ４１はＸ１４を１インクリメントする。たとえば、インタフェース１９やＩＦＣ４２やＳＳＤＣ４１がＰＣＩ Expressの伝送路数（Lane数）が最大８Laneとして設計されているにも関わらず、実際に記憶部２や情報処理装置１１１との間で確立された伝送路数が４Laneのようなより少数の伝送路数であることが検出された場合、ＳＳＤＣ４１はこれをPCI Express通信上のエラーとみなし、統計情報Ｘ１４を１インクリメントする。統計情報Ｘ１４は記憶部２の製造時（検査工程前）にゼロにリセットされていることが望ましい。この値が大きいほど信頼性が悪化していることを示す。

【００８０】

インタフェース１９のエラー発生回数（統計情報Ｘ１５）について説明する。ＳＳＤＣ

４１やＩＦＣ４２やチップセット７がその他の（Ｘ１２以外の）インタフェース１９での異常を１回検出するごとに、ＳＳＤＣ４１は統計情報Ｘ１５を１インクリメントする。統計情報Ｘ１５は記憶部２の製造時（検査工程前）にゼロにリセットされていることが望ましい。この値が大きいほど信頼性が悪化していることを示す。

【００８１】

ＲＡＭ４０のエラー発生回数（統計情報Ｘ１６）について説明する。ＳＳＤＣ４１がデータをＲＡＭ４０に書き込む場合、ＳＳＤＣ４１またはＲＡＭ４０のＥＣＣ符号化部または誤り検出符号作成部がデータを符号化してＲＡＭ４０に書き込み、ＳＳＤＣ４１がデータをＲＡＭ４０から読み出す場合、ＳＳＤＣ４１またはＲＡＭ４０のＥＣＣ復号部または誤り検出部がデータを誤り訂正または誤り検出してＲＡＭ４０からデータを読み出す。ＳＳＤＣ４１がＲＡＭ４０からデータを読み出す場合に誤り訂正できなかった時または誤り検出した時、ＳＳＤＣ４１は統計情報Ｘ１６を１インクリメントする。統計情報Ｘ１６は記憶部２の製造時（検査工程前）にゼロにリセットされていることが望ましい。この値が大きいほど信頼性が悪化していることを示す。

【００８２】

記憶部２の使用時間総計（統計情報Ｘ１７）について説明する。記憶部２の電源がＯＮになっている間、ＳＳＤＣ４１がクロックをカウントしたり、内部の時計回路から時刻情報を受信することで、ＳＳＤＣ４１は経過時間として統計情報Ｘ１７をインクリメントする。あるいは、ＳＳＤＣ４１が情報処理装置１１１から定期的に情報処理装置１１１の時刻情報を受信するようにし、その時刻情報の差分をインクリメントするようにしてもよい。統計情報Ｘ１７は記憶部２の製造時（検査工程前）にゼロにリセットされていることが望ましい。この値が大きいほど信頼性が悪化していることを示す。

【００８３】

起動回数(統計情報Ｘ１８）について説明する。記憶部２に電源が供給され起動されるたびにＳＳＤＣ４１はＸ１８を１インクリメントする。電源起動時にはＮＡＮＤ型フラッシュメモリ１６に対して読み出し動作が発生し、書き込み動作が発生することがあるため、この値が大きいほど信頼性が悪化していることを示す。統計情報Ｘ１８は記憶部２の製造時（検査工程前）にゼロにリセットされていることが望ましい。

【００８４】

不正電源断発生回数(統計情報Ｘ１９）について説明する。通常、記憶部２の電源をオフにするときには、情報処理装置１１１は記憶部２に対して、たとえばInformation technology ATA/ATAPI Command Set-3 (ACS-3)に記載されているE0h Standby Immediateコマンド等を発行したり、NVM Express Revision1.1に記載されているShutdown Notification (CC.SHN)を01bにセットしたりするなどして、あらかじめ記憶部２を電源遮断可能な状態に遷移させておいてから記憶部２の電源を遮断する。一方、記憶部２が電源遮断可能ではない状態において意図せず電源遮断が発生する場合があり、これは不正電源断（不正電源遮断、Ungraceful Power Down, Unsafe Shutdown, Unintentional Power Down）とよばれる。不正電源遮断後にはじめて記憶部２が起動する時、ＳＳＤＣ４１はＸ１９を１インクリメントする。不正電源遮断においてはユーザデータが破壊されたり、ＮＡＮＤメモリ１６に対して大量の読み出し・書き込み動作が発生して信頼性劣化要因にもなるため、Ｘ１９が大きいほど信頼性が悪化していることを示す。統計情報Ｘ１９は記憶部２の製造時（検査工程前）にゼロにリセットされていることが望ましい。

【００８５】

推奨動作温度の最高値を上回った時間累計（統計情報Ｘ２０）について説明する。たとえば、記憶部２の基板上、ＳＳＤＣ４１内、ＮＡＮＤメモリ１６内など、記憶部２内に温度計が実装されている場合、ＳＳＤＣ４１は温度計から定期的に温度情報を受信している。受信した温度が推奨動作温度（たとえば１００℃）を上回った場合、ＳＳＤＣ４１はクロックや内部の時計や情報処理装置１１１から取得する時刻情報をもとに、推定動作温度以上で動作している時間数をインクリメントしていく。統計情報Ｘ２０は記憶部２の製造時（検査工程前）にゼロにリセットされていることが望ましい。この値が大きいほど信頼

性が悪化していることを示す。

【００８６】

推奨動作温度の最低値を下回った時間累計（統計情報Ｘ２１）について説明する。記憶部２内に温度計が実装されている場合、ＳＳＤＣ４１は温度計から定期的に温度情報を受信している。受信した温度が推奨動作温度（たとえば－４０℃）を下回った場合、ＳＳＤＣ４１はクロックや内部の時計や情報処理装置１１１から取得する時刻情報をもとに、推定動作温度以上で動作している時間数をインクリメントしていく。記憶部２の製造時（検査工程前）にゼロにリセットされていることが望ましい。この値が大きいほど信頼性が悪化していることを示す。

【００８７】

コマンドの応答時間最大値（統計情報Ｘ２２）について説明する。ＳＳＤＣ４１は、情報処理装置１１１からコマンドを受信してから、情報処理装置１１１へ応答するまで（またはコマンド実行完了するまで）に要した時間（またはクロック数）を計測し、その最大値を統計情報Ｘ２２として統計情報６５に格納する。Ｘ２２を上回る応答時間が発生した場合は、ＳＳＤＣ４１はこの応答時間をＸ２２に上書きする。ＳＳＤＣ４１は、コマンドそれぞれに対して統計情報Ｘ２２を格納してもよい。記憶部２の製造時（検査工程前）または記憶部２の出荷時にはＸ２２がゼロにリセットされていることが望ましい。

【００８８】

コマンドの応答時間平均値（統計情報Ｘ２３）について説明する。ＳＳＤＣ４１は、情報処理装置１１１からコマンドを受信してから、情報処理装置１１１へ応答するまで（またはコマンド実行完了するまで）に要した時間（またはクロック数）を計測し、その平均値を統計情報Ｘ２３として統計情報６５に格納する。たとえば、ＳＳＤＣ４１は応答時間リストを一定数ＲＡＭ４０に保持しておき、その応答時間リストの平均値を算出することにより統計情報Ｘ２３を計算する。ＳＳＤＣ４１はコマンドそれぞれに対して統計情報Ｘ２３を格納してもよい。記憶部２の製造時（検査工程前）または記憶部２の出荷時にはＸ２３がゼロにリセットされていることが望ましい。

【００８９】

ＮＡＮＤメモリの応答時間最大値（統計情報Ｘ２４）について説明する。ＳＳＤＣ４１は、ＳＳＤＣ４１がＮＡＮＤメモリ１６に命令してから応答を得る（またはコマンド実行完了通知を受信する）までに要した時間（またはクロック数）を計算し、その最大値を統計情報Ｘ２４として統計情報６５に格納する。Ｘ２４を上回る応答時間が発生した場合は、ＳＳＤＣ４１はこの応答時間をＸ２４に上書きする。ＳＳＤＣ４１はコマンドそれぞれに対して統計情報Ｘ２４を保持してもよい。記憶部２の製造時（検査工程前）または記憶部２の出荷時にはＸ２４がゼロにリセットされていることが望ましい。

【００９０】

ＮＡＮＤの応答時間平均値（統計情報Ｘ２５）について説明する。ＳＳＤＣ４１は、ＮＡＮＤメモリ１６に命令してから応答を得る（またはコマンド実行完了通知を受信する）までに要した時間（またはクロック数）を計測し、その平均値を統計情報Ｘ２５として統計情報６５に格納する。たとえばＳＳＤＣ４１は応答時間リストを一定数ＲＡＭ４０に格納しておき、その応答時間リストの平均値を算出することにより統計情報Ｘ２５を得る。ＳＳＤＣ４１はコマンドそれぞれに対して統計情報Ｘ２５を保持してもよい。記憶部２の製造時（検査工程前）または記憶部２の出荷時にはＸ２５がゼロにリセットされていることが望ましい。

【００９１】

現在温度（統計情報Ｘ２６）について説明する。記憶部２内に温度計が実装されている場合、ＳＳＤＣ４１は温度計から定期的に温度情報を受信する。ＳＳＤＣ４１は温度計から最後に受信した温度を現在温度として統計情報Ｘ２６に格納する。ＳＳＤＣ４１は、この値が極端に大きいと（たとえば８５℃以上）、記憶部２の信頼性に悪影響があり、また、この温度が極端に小さいと（たとえば－１０℃以下）、記憶部２の信頼性に悪影響があると判定する。

【００９２】
最高温度（統計情報Ｘ２７）について説明する。ＳＳＤＣ４１は、現在温度Ｘ２６の最大値を最高温度として統計情報Ｘ２７に格納する。この値が極端に大きいと（たとえば８５℃以上）、記憶部２の信頼性に悪影響がある。ＳＳＤＣ４１は、Ｘ２７よりも大きい現在温度を温度計から受信した時、Ｘ２７を現在温度に書き換える。記憶部２の製造時（検査工程前）または記憶部２の出荷時にはＸ２７が記憶部２の動作温度にくらべて十分に小さい温度（たとえば－４０℃）にリセットされていることが望ましい。
【００９３】
最低温度（統計情報Ｘ２８）について説明する。ＳＳＤＣ４１は、現在温度Ｘ２６の最小値を最低温度として統計情報Ｘ２８に格納する。この値が極端に小さいと（たとえば－４０℃以下）、記憶部２の信頼性に悪影響がある。ＳＳＤＣ４１は、Ｘ２８よりも小さい現在温度を温度計から受信した時、Ｘ２８を現在温度に書き換える。記憶部２の製造時（検査工程前）または記憶部２の出荷時にはＸ２８が記憶部２の動作温度にくらべて十分に大きい温度（たとえば１２０℃）にリセットされていることが望ましい。
【００９４】
管理情報冗長度（統計情報Ｘ２９）について説明する。ＮＡＮＤメモリ１６の管理情報領域４０Ｍのデータ やＮＡＮＤメモリ１６に記憶されたシステムプログラム（ファームウエア）などのシステムデータが壊れて読みだし不能になると、記憶部２が正常な動作ができなくなる可能性がある。記憶部２の信頼性向上のため、ＳＳＤＣ４１はシステムデータをＲＡＩＤ１やＲＡＩＤ５やＲＡＩＤ６により複数物理ブロックや複数チャネルをまたいで冗長化して領域４０Ｍに格納することが望ましい。ＳＳＤＣ４１は、システムデータの冗長度を数値化して管理情報冗長度（統計情報Ｘ２９）として統計情報６５に格納する。冗長度Ｘ２９＝Ｒの時、最大（Ｒ－１）ブロックのデータ消失まではデータを復元することができる。たとえば、ＳＳＤＣ４１が管理情報４５を４ブロックにわたってＲＡＩＤ１で管理する場合、管理情報４５はたとえばブロックＡとブロックＢとブロックＣとブロックＤそれぞれにクローンとして格納される。この場合、管理情報４５は合計４つのクローンを保持するため、管理情報４５の冗長度Ｘ２９は４である。たとえばブロックＡのデータが壊れて読みだし不能になった場合、ＳＳＤＣ４１はブロックＢまたはブロックＣまたはブロックＤからデータ読み出しを行うことで管理情報４５を読み出すことができる。この場合、管理情報４５は合計３つのクローンを保持するため、管理情報４５の冗長度Ｘ２９は３である。たとえば、ＳＳＤＣ４１が管理情報４５を４ブロックにわたってＲＡＩＤ５で管理する場合、管理情報４５はたとえばブロックＡとブロックＢとブロックＣとブロックＤそれぞれにおいて４台のＲＡＩＤ５で格納される。この場合、最大１ブロックのデータが消失したとしてもデータを復元可能であるため、管理情報４５の冗長度Ｘ２９は２である。１ブロック分のデータが消失している状態においては、冗長度Ｘ２９は１となる。冗長度Ｘ２９が下がると、システムデータが復元不可能になる確率が上がり記憶部２の故障率が増大するため、冗長度Ｘ２９が小さいほど信頼性が悪化していることを示す。冗長度Ｘ２９が減少している場合、ＳＳＤＣ４１は復元したデータをデータ消失したブロックに再書き込みすることで冗長度を回復させることが望ましい。
【００９５】
ＲＡＭ４０への書き込みデータ量合計（統計情報Ｘ３０）について説明する。ＳＳＤＣ４１は、記憶部２内のＲＡＭ４０へのデータ書き込み量の累計値を統計情報Ｘ３０として統計情報６５に格納する。ＳＳＤＣ４１は、たとえばＲＡＭ４０に１ページのデータを書き込むごとに統計情報Ｘ３０を１インクリメントする。統計情報Ｘ３０は記憶部２の製造時（検査工程前）にゼロにリセットされていることが望ましい。統計情報Ｘ３０が大きいほど信頼性が悪化していることを示す。
【００９６】
統計情報増加率（統計情報Ｘ３１）について説明する。ＳＳＤＣ４１は、統計情報Ｘ０１～Ｘ２５の最新でない情報（たとえば一定時刻前や、記憶部２をパワーオンした時の値や前回記憶部２をパワーダウンしたときの値など）を別途管理情報４４に格納しておく。

ＳＳＤＣ４１は、例えば、下記のいずれかの式で統計情報Ｘ３１を計算する。
統計情報増加率＝（最新統計情報）－（旧情報）
統計情報増加率＝（（最新統計情報）－（旧情報））／（旧情報を取得してからの経過時刻）
統計情報増加率＝（（最新統計情報）－（旧情報））／（旧情報を取得してからのＮＡＮＤアクセス回数）
記憶部２の製造時（検査工程前）にゼロにリセットされていることが望ましい。この値が大きいほど信頼性が悪化していることを示す。

【００９７】
ＮＡＮＤ整理失敗フラグ（統計情報Ｘ３２）について説明する。統計情報Ｘ３２が１になっていると、ＮＡＮＤ整理によっても動作に十分な数のフリーブロック数を確保できないことになる。記憶部２の製造時（検査工程前）にゼロにリセットされていることが望ましい。この値が大きいほど信頼性が悪化していることを示す。

【００９８】
ＳＳＤＣ４１は、統計情報６５に上述したすべてのパラメータを格納してもよいし、これらの一部あるいはどれか一つのみを格納してもよい。ＳＳＤＣ４１は、統計情報６５を最新情報をＲＡＭ４０上の領域４０Ａに保持し、定期的にＮＡＮＤメモリ１６上の領域４０Ｍにバックアップデータとして退避保存することが望ましい。一方、ＳＳＤＣ４１は、ＲＡＭ４０やＮＡＮＤメモリ１６のどちらか一方にのみ保存するようにしてもよいし、当該統計情報を情報処理装置１１１に送信して、情報処理装置１１１や情報処理装置１１１に接続された記憶装置に保存するようにしてもよい。

【００９９】
（ＬＢＡ正引き変換）
つぎに、図１６を用いてＳＳＤＣ４１がＬＢＡから物理アドレスを特定する手順（ＬＢＡ正引き変換）について説明する。ＬＢＡが指定されたとき、ＳＳＤＣ４１はＬＢＡからトラックアドレスとクラスタアドレスとクラスタ内アドレスを計算する。

【０１００】
ＳＳＤＣ４１は、まずトラックテーブル６３を検索し、計算したトラックアドレスに対応する物理ブロックＩＤを特定する（ステップＳ１００、Ｓ１０１）。ＳＳＤＣ４１は、特定した物理ブロックＩＤが有効であるか否かを判定し（ステップＳ１０２）、物理ブロックＩＤがヌルではなく有効な値である場合は（ステップＳ１０２：Ｙｅｓ）、この物理ブロックＩＤがＡＢＴ６２にエントリされているか否かを検索する（ステップＳ１０３）。ＡＢＴ６２に物理ブロックＩＤがエントリされている場合は（ステップＳ１０４：Ｙｅｓ）、この物理ブロックＩＤの指定する物理ブロックの先頭位置からトラック内アドレス分だけシフトした位置が指定されたＬＢＡに対応するＮＡＮＤメモリ１６上の物理的な位置となる（ステップＳ１０５）。このような場合には、ＬＢＡに対応するＮＡＮＤメモリ１６上の物理的な位置の特定にはクラスタテーブル６４を必要とせず、このようなＬＢＡを「トラック管理されているＬＢＡ」と呼ぶ。ステップＳ１０４において、ＡＢＴ６２に物理ブロックＩＤがエントリされていない場合は（ステップＳ１０４：Ｎｏ）、指定されたＬＢＡは対応する物理アドレスを持たないことになり、このような状態を「未書き込み状態」と呼ぶ（ステップＳ１０６）。

【０１０１】
ステップＳ１０２において、指定されたトラックアドレスに対応する物理アドレスがヌルであり無効な値の場合は（ステップＳ１０２：Ｎｏ）、ＳＳＤＣ４１はＬＢＡからクラスタアドレスを計算し、クラスタテーブル６４を検索し、計算したクラスタアドレスに対応する物理ブロックＩＤ及び対応する物理ブロック内ページアドレスをクラスタテーブル６４から取得する（ステップＳ１０７）。物理ブロックＩＤと物理ブロック内ページアドレスが指定する物理ページの先頭位置からクラスタ内アドレス分だけシフトした位置が指定されたＬＢＡに対応するＮＡＮＤメモリ１６上の物理的な位置となる。このような場合は、ＬＢＡに対応するＮＡＮＤメモリ１６上の物理的な位置は、トラックテーブル６３の

みからは特定できず、クラスタテーブル６４の参照を必要とすることになり、このようなＬＢＡを「クラスタ管理されているＬＢＡ」という（ステップＳ１０８）。

【０１０２】

（読み出し動作）

つぎに、図１７、図１８を用いて情報処理装置１１１による記憶部２および記憶部３からの読み出し動作を説明する。本実施形態で説明する読み出し動作は、読み出しコマンドとしてInformation technology ATA/ATAPI Command Set-3 (ACS-3)に記載されている60h READ FPDMA QUEUEDを用いた場合であるが、25h READ DMA EXTなどその他書き込みコマンドを採用してもよく、読み出しコマンドの種類の違いは発明の本質には影響しない。たとえば、読み出しコマンドとしてNVM Express Revision 1.1に記載されている02h Readを用いても良い。記憶部２が情報処理装置１１１から読み出し命令を受信した場合は（ステップＳ１１０）、ＳＳＤＣ４１がこの読み出し命令をＲＡＭ４０上の読み出し命令待ち行列に追加し（ステップＳ１１１）、読み出し命令を受理した旨を情報処理装置１１１に返信する。

【０１０３】

一方、ＳＳＤＣ４１は、ＲＡＭ４０上の読み出し命令待ち行列に命令が存在している場合、読み出し処理を実行可能な状態になっているか否かを判定し（ステップＳ１２０）、読み出し処理を実行可能な状態になったと判定すると、先の図１６に示したＬＢＡ正引き変換手順にしたがって情報処理装置１１１から受信したＬＢＡからデータの物理的な位置を特定する（ステップＳ１２１）。ＳＳＤＣ４１は、特定した位置の物理ページからデータを読み出し（ステップＳ１２３）、読み出したデータのうちＥＣＣ冗長ビットを用いてＥＣＣ復号化し（ステップＳ１２４）、復号化したデータをＩＦＣ４２を介して情報処理装置１１１に送信し（ステップＳ１２５）、統計情報６５を更新する。なお、ＮＡＮＤメモリ１６から読み出したデータは、いったんＲＡＭ４０に書き込み、ＲＡＭ４０に書き込んだデータを復号化して情報処理装置１１１に送信するようにしてもよいし、復号化したデータをいったんＲＡＭ４０に書き込み、ＲＡＭ４０に書き込んだデータを情報処理装置１１１に送信するようにしてもよい。

【０１０４】

ステップＳ１２４において、ＳＳＤＣ４１はＥＣＣによる復号化を試みるが、復号化できなかった場合、復号化できなかったページを含む物理ブロックをＡＢＴ６２から削除してＢＢＴ６１に登録し、統計情報６５のＥＣＣ訂正できなかったＥＣＣ訂正単位数総計（統計情報Ｘ０９）を加算する。その際、当該ブロックのデータをＦＢＴ６０から割り当てたフリーブロックにコピーし、当該フリーブロックの物理ブロックＩＤをＡＢＴ６２に登録してトラックテーブル６３およびクラスタテーブル６４の物理ブロックをコピー元物理ブロックＩＤからコピー先物理ブロックＩＤに書き換えることが望ましい。

【０１０５】

（書き込み動作）

つぎに、図１９、図２０を用いて情報処理装置１１１による記憶部２および記憶部３への書き込み動作を説明する。本実施形態で説明する書き込み動作は、書き込みコマンドとしてInformation technology ATA/ATAPI Command Set-3 (ACS-3)に記載されている61h WRITE FPDMA QUEUEDを用いた場合であるが、35h WRITE DMA EXTなどその他書き込みコマンドを採用してもよく、書き込みコマンドの種類の違いは発明の本質には影響しない。たとえば、書き込みコマンドとしてNVM Express Revision 1.1に記載されている01h Writeを用いても良い。たとえば、記憶部２が情報処理装置１１１から書き込み命令を受信した場合は（ステップＳ１３０）、ＳＳＤＣ４１がこの書き込み命令をＲＡＭ４０上の読み出し命令待ち行列に追加し（ステップＳ１３１）、書き込み命令を受理した旨を情報処理装置１１１に返信する。

【０１０６】

一方、ＳＳＤＣ４１は、ＲＡＭ４０上の書き込み命令待ち行列に命令が存在している場合、書き込み処理を実行可能な状態になっているか否かを判定し（ステップＳ１４０）、

書き込み処理を実行可能な状態になったと判定すると、情報処理装置１１１に書き込み可能であることを通知し、情報処理装置１１１から書き込みデータを受信し、受信したデータをＥＣＣ符号化し、符号化したデータをＲＡＭ４０のキャッシュメモリ４６に記憶する。なお、符号化しないデータをキャッシュメモリ４６に記憶し、ＮＡＮＤメモリ１６に書き込む時に符号化するようにしてもよい。

【０１０７】

つぎに、ＳＳＤＣ４１はＦＢＴ６０を読み出し（ステップＳ１４１）、ＦＢＴ６０からフリーブロックの物理ブロックＩＤを取得する。フリーブロックが存在しない場合は（ステップＳ１４２：Ｎｏ）、ＳＳＤＣ４１は後述するＮＡＮＤメモリ１６の整理（ＮＡＮＤ整理）を行い（ステップＳ１４３）、この整理の後、ＦＢＴ６０を読み出し（ステップＳ１４４）、ＦＢＴ６０からフリーブロックの物理ブロックＩＤを取得する。ＳＳＤＣ４１は、物理ブロックＩＤを取得したフリーブロックに対し、消去動作を行う。消去エラーが発生した場合は、当該物理ブロックＩＤをＢＢＴ６１に追加し、ＦＢＴ６０から削除し、Ｓ１４１からやり直してフリーブロックを再取得する。なお、一度消去エラーが発生した物理ブロックであっても、再度消去動作を行うと消去エラーが発生せずに正常に消去できることがあるため、バッドブロック数の不必要な増大を防止するという観点では、ＦＢＴ６０やＡＢＴ６２に統計情報Ｘ０６としてのブロックごと消去エラー発生回数の項目を各ブロックごとに設け、ブロックの消去エラーが発生した場合にこれをインクリメントするようにし、ブロックごと消去エラー発生回数が所定値以上になった場合に当該ブロックをＢＢＴ６１に登録するようにするのが望ましい。さらに望ましくは、連続して消去エラーが発生する物理ブロックのみをバッドブロック化するために、ＳＳＤＣ４１は、前記「ブロックごと消去エラー発生回数」のかわりに「ブロックごと消去連続エラー回数」の項目をもうけ、ブロックの消去エラーが発生した場合にこれをインクリメントするようにし、消去をエラー無く行えた場合にこれをゼロにリセットするようにし、「ブロックごと消去連続エラー回数」が所定値以上になった場合に当該ブロックをＢＢＴ６１に登録するようにするのが望ましい。

【０１０８】

つぎに、ＳＳＤＣ４１は、書き込み命令で指定されているＬＢＡが未書き込み状態であるか否かを検索するために、先の図１６に示した正引き変換手順にしたがってＬＢＡに対応する有効なデータがＮＡＮＤメモリ１６に記憶済みであるか否か判定する（ステップＳ１４５、Ｓ１４６）。

【０１０９】

ＬＢＡが未書き込み状態である場合は（ステップＳ１４６：Ｙｅｓ）、ＳＳＤＣ４１は、キャッシュメモリ４６に記憶している受信データをフリーブロックに書き込み（ステップＳ１４７）、書き込みを行ったフリーブロック（新物理ブロック）のＩＤ及びその消去回数をＡＢＴ６２に登録し、書き込みを行った物理ブロックのＩＤをＦＢＴ６０から削除する（ステップＳ１５１）。この際、受信データのＬＢＡをトラック単位の区画（トラック区画）で区切り、トラック区画内がデータで埋め尽くされているか否かを判定することで、トラック管理するかクラスタ管理するかを判定する（ステップＳ１５２）。すなわち、トラック区画内がデータで埋め尽くされている場合は、トラック管理となり、トラック区画内がデータで埋め尽くされていない場合は、クラスタ管理となる。クラスタ管理の場合は、クラスタテーブル６４を書き換えて、ＬＢＡに新物理ブロックＩＤを関連付け（ステップＳ１５３）、さらにトラックテーブル６３を書き換えて、ＬＢＡに無効な物理ブロックＩＤ（例えば、ヌル）を関連付ける。トラック管理の場合は、トラックテーブルを書き換えて、ＬＢＡに新物理ブロックＩＤを関連付ける（ステップＳ１５４）。

【０１１０】

一方、ステップＳ１４６において、ＬＢＡが未書き込み状態でない場合は、ＳＳＤＣ４１は正引き変換により得られた物理ブロックＩＤをもとに、対応する物理ブロック内全データをＮＡＮＤメモリ１６から読み出して、ＲＡＭ４０に書き込む（ステップＳ１４８）。そして、キャッシュメモリ４６に記憶されているデータとＮＡＮＤメモリ１６から読み

出してＲＡＭ４０に書き込んだデータとをＲＡＭ４０上で合成し（ステップＳ１４９）、合成したデータをフリーブロックに書き込む（ステップＳ１５０）。

【０１１１】

なお、ステップＳ１５０で書き込みエラーが発生した場合は、当該物理ブロックＩＤをＢＢＴ６１に追加し、ＦＢＴ６０から削除し、ステップＳ１４１からやり直してフリーブロックを再取得する。なお、一度書き込みエラーが発生した物理ブロックであっても、再度書き込み動作を行うと書き込みエラーが発生せずに正常に書き込みできることがあるため、バッドブロック数の不必要な増大を防止するという観点では、ＦＢＴ６０やＡＢＴ６２に統計情報Ｘ０５としてのブロックごと書き込みエラー発生回数の項目を各ブロックごとに設け、ブロックの書き込みエラーが発生した場合にこれをインクリメントするようにし、「ブロックごと書き込みエラー発生回数」が所定値以上になった場合に当該ブロックをＢＢＴ６１に登録するようにするのが望ましい。さらに望ましくは、連続して書き込みエラーが発生する物理ブロックのみをバッドブロック化するために、ＳＳＤＣ４１は、前記「ブロックごと書き込みエラー発生回数」のかわりに「ブロックごと書き込み連続エラー回数」の項目をもうけ、ブロックの書き込みエラーが発生した場合にこれをインクリメントするようにし、書き込みをエラー無く行えた場合にこれをゼロにリセットするようにし、「ブロックごと書き込み連続エラー回数」が所定値以上になった場合に当該ブロックをＢＢＴ６１に登録するようにするのが望ましい。

【０１１２】

ＳＳＤＣ４１は、書き込みを行ったフリーブロック（新物理ブロック）のＩＤ及びその消去回数をＡＢＴ６２に登録し、書き込みを行った物理ブロックのＩＤをＦＢＴ６０から削除する（ステップＳ１５１）。ＬＢＡがクラスタ管理である場合は、クラスタテーブル６４の旧物理ブロックＩＤを新物理ブロックＩＤに書き換える（ステップＳ１５２、Ｓ１５３）。トラック管理の場合は、トラックテーブルの旧物理ブロックＩＤを新物理ブロックＩＤに書き換える（ステップＳ１５２、Ｓ１５４）。さらに、ＳＳＤＣ４１は、旧物理ブロックＩＤ及びその消去回数をＦＢＴ６０に追加し、旧物理ブロックＩＤ及びその消去回数をＡＢＴ６２から削除する（ステップＳ１５５）。ＳＳＤＣ４１は以上の書き込み処理の内容を統計情報６５に反映する。

【０１１３】

（ＮＡＮＤ整理）

通常、記憶部２の全ＬＢＡの容量（全論理セクタ数）は、記憶部２のＮＡＮＤメモリ１６の全容量よりも小さく設計されているため（オーバープロビジョニング）、書き込み動作がトラック単位で書き込まれ続ける限りはフリーブロックが枯渇することはない。一方、未書き込みＬＢＡに対してクラスタ単位の書き込みが多数発生した場合、クラスタ単位の書き込み一つに対してクラスタよりも容量の大きい物理ブロックが割り当てられることになるため、書き込まれるデータ容量よりも多くのＮＡＮＤメモリ１６物理ブロックを必要とすることになり、それによりフリーブロックが枯渇する可能性がある。フリーブロックが枯渇した場合は、以下に示すＮＡＮＤメモリ１６の整理によってフリーブロックを新たに確保することができる。

【０１１４】

図２１を用いてＳＳＤＣ４１によるＮＡＮＤ整理を説明する。物理ブロックに記憶されている全てのクラスタが有効クラスタであるとは限らず、有効クラスタに該当しない無効クラスタはＬＢＡに対応付けられていない。有効クラスタとは最新のデータを記憶しているクラスタであり、無効クラスタとは同一ＬＢＡのデータが他の場所に書きこまれ、参照されることがなくなったクラスタである。物理ブロックは無効クラスタの分だけデータに空きがあることになり、有効クラスタのデータを集めて違うブロックに書き直すＮＡＮＤ整理を実行することでフリーブロックを確保することができる。

【０１１５】

まず、ＳＳＤＣ４１は、選択物理ブロックＩＤ＝ｉを０にセットし、空き領域累積量Ｓを０にセットする（ステップＳ１６０）。ＳＳＤＣ４１は、このＩＤがｉ＝０の物理ブロ

ックがトラックテーブル６３にエントリされているか否かを判定する（ステップＳ１６１）。トラックテーブルにエントリされている場合はｉを＋１し（ステップＳ１６２）、つぎの番号のＩＤを持つ物理ブロックについて同様の判定を行う（ステップＳ１６１）。すなわち、物理ブロックＩＤがトラックテーブル６３に含まれている場合は、この物理ブロックのデータはトラック管理であるため、ＮＡＮＤ整理対象に含めない。

【０１１６】

ＳＳＤＣ４１は、ＩＤ＝ｉの物理ブロックがトラック管理でない場合は（ステップＳ１６１：Ｎｏ）、つぎにクラスタテーブル６４を参照し、ＩＤ＝ｉの物理ブロックに含まれる有効クラスタのアドレスを全て取得する（ステップＳ１６３）。そして、ＳＳＤＣ４１は、取得した有効クラスタの総容量分のサイズｖを求め（ステップＳ１６４）、ｖ＜物理ブロックサイズであるときは（ステップＳ１６５）、現在物理ブロックのＩＤをＮＡＮＤ整理対象ブロックリストに加える（ステップＳ１６６）。さらに、ＳＳＤＣ４１は、取得クラスタ累計量Ｓに現在物理ブロックの取得クラスタ容量ｖを加算し、取得クラスタ累計量Ｓを更新する（ステップＳ１６７）。

【０１１７】

ステップＳ１６５で、ｖ＜物理ブロックサイズでないとき、あるいはステップＳ１６８で取得クラスタ累計量Ｓが物理ブロックサイズに到達していない場合は、ＳＳＤＣ４１は、ｉを＋１し（ステップＳ１６２）、つぎの番号のＩＤを持つ物理ブロックについて、ステップＳ１６１～Ｓ１６７の手順を前記同様に実行する。そして、ステップＳ１６８で、取得クラスタ累計量Ｓが物理ブロックサイズに到達するまで、ステップＳ１６１～Ｓ１６７の手順を繰り返す。

【０１１８】

そして、ステップＳ１６８において、取得クラスタ累計量Ｓが物理ブロックサイズに到達した場合は、ＳＳＤＣ４１は、ＮＡＮＤ整理対象ブロックリスト上の全物理ブロックについての全有効クラスタのデータをＮＡＮＤメモリ１６から読み出してＲＡＭ４０に書き込み（ステップＳ１６９）、さらに、ＮＡＮＤ整理対象ブロックリスト上の全物理ブロックに対して消去処理を行い（ステップＳ１７０）、消去処理を行った全物理ブロックをＡＢＴ６２から削除してＦＢＴ６０に追加する（ステップＳ１７１）。その際、消去回数をインクリメントする。なお、ステップＳ１７０で行う消去動作の対象は、ステップＳ１７２でデータを書き込む対象のブロックに限定してもよく、ブロックの消去回数を抑制するという観点ではそのように行うことが望ましい。

【０１１９】

消去エラーが発生した場合は、当該物理ブロックＩＤをＢＢＴ６１に追加し、ＦＢＴ６０から削除する。なお、一度消去エラーが発生した物理ブロックであっても、再度消去動作を行うと消去エラーが発生せずに正常に消去できることがあるため、バッドブロック数の不必要な増大を防止するという観点では、ＦＢＴ６０やＡＢＴ６２に「ブロックごと消去エラー発生回数」の項目を各ブロックごとに設け、ブロックの消去エラーが発生した場合にこれをインクリメントするようにし、ブロックごと消去エラー発生回数が所定値以上になった場合に当該ブロックをＢＢＴ６１に登録するようにするのが望ましい。さらに望ましくは、連続して消去エラーが発生する物理ブロックのみをバッドブロック化するために、ＳＳＤＣ４１は、前記「ブロックごと消去エラー発生回数」のかわりに「ブロックごと消去連続エラー回数」の項目をもうけ、ブロックの消去エラーが発生した場合にこれをインクリメントするようにし、消去をエラー無く行えた場合にこれをゼロにリセットするようにし、「ブロックごと消去連続エラー回数」が所定値以上になった場合に当該ブロックをＢＢＴ６１に登録するようにするのが望ましい。

【０１２０】

そして、ＳＳＤＣ４１は、ＦＢＴ６０から新たなフリーブロックを取得し、取得したフリーブロックに対しＲＡＭ４０に書き込んだデータを書き込み（ステップＳ１７２）、データを書き込んだフリーブロックの物理ブロックＩＤ及び当該ブロックの消去回数をＡＢＴ６２に追加し、さらにデータが書き込まれたブロックのブロックＩＤをＦＢＴ６０から

削除する（ステップＳ１７３）。さらに、ＳＳＤＣ４１は、今回のＮＡＮＤ整理に対応するように、クラスタテーブル６４におけるクラスタアドレス、物理ブロックＩＤ及び物理ブロック内ページアドレスを更新する（ステップＳ１７４）。ＳＳＤＣ４１は上記ＮＡＮＤ整理の処理内容を統計情報６５に反映する。

【０１２１】

なお、ステップＳ１７２で書き込みエラーが発生した場合は、当該物理ブロックＩＤをＢＢＴ６１に追加し、ＦＢＴ６０から削除し、フリーブロックを再取得する。なお、一度書き込みエラーが発生した物理ブロックであっても、再度書き込み動作を行うと書き込みエラーが発生せずに正常に書き込みできることがあるため、バッドブロック数の不必要な増大を防止するという観点では、ＦＢＴ６０やＡＢＴ６２に「ブロックごと書き込みエラー発生回数」の項目を各ブロックに設け、ブロックの書き込みエラーが発生した場合にこれをインクリメントするようにし、「ブロックごと書き込みエラー発生回数」が所定値以上になった場合に当該ブロックをＢＢＴ６１に登録するようにするのが望ましい。さらに望ましくは、連続して書き込みエラーが発生する物理ブロックのみをバッドブロック化するために、ＳＳＤＣ４１は、前記「ブロックごと書き込みエラー発生回数」のかわりに「ブロックごと書き込み連続エラー回数」の項目をもうけ、ブロックの書き込みエラーが発生した場合にこれをインクリメントするようにし、書き込みをエラー無く行えた場合にこれをゼロにリセットするようにし、「ブロックごと書き込み連続エラー回数」が所定値以上になった場合に当該ブロックをＢＢＴ６１に登録するようにするのが望ましい。

【０１２２】

なお、図２１の手順では、フリーブロックにデータを詰め込むことを優先するＮＡＮＤ整理を行ったが、ステップＳ１６４で、ｖを物理ブロックサイズから取得したクラスタの容量を減算することによって求め、ステップＳ１６５でｖ＞０か否かを判定し、ｖ＞０の場合はステップＳ１６８に移行し、ｖ＞０でない場合はステップＳ１６２に移行させるようにすることで、フリーブロックを確保することを優先するＮＡＮＤ整理を行うようにしてもよい。

【０１２３】

（削除通知）

つぎに、図２２を用いてＳＳＤＣ４１による削除通知処理について説明する。削除通知は、情報処理装置１１１上のＯＳ１００によってデータの削除が行われた場合に、情報処理装置１１１から記憶部２や記憶部３に対して送信される命令である。削除通知処理に使用されるコマンドは通称トリムコマンドと呼ばれる。トリムコマンドの例として、例えば、Information technology ATA/ATAPI Command Set-3 (ACS-3)に記載されている06h Data Set Managementコマンドや、NVM Express Revision 1.1に記載されている09h Dataset ManagementコマンドのDeallocateがあげられる。これは、ＯＳ１００上でデータが削除された場合、削除されたデータの存在する論理アドレス領域（ＬＢＡ領域）を、ＬＢＡ及びセクタ数の組みよりなるＬＢＡ Range Entryとして外部記憶装置に通知することにより、記憶部２や記憶部３上でもその領域を空き領域として扱うことができる方式である。削除通知により、ＳＳＤＣ４１はフリーブロックを新たに確保することができる。なお、トリムコマンドの機能は、前記コマンドだけでなく、たとえば、Information technology ATA/ATAPI Command Set-3 (ACS-3)で記述されているSCT Command Transportや、NVM Express Revision 1.1に記載されている08h Write Zeroesコマンドや、その他ベンダー独自のコマンドなどその他コマンドによって実現してもよい。

【０１２４】

記憶部２および記憶部３が情報処理装置１１１から削除通知を受信した場合は（ステップＳ１８０）、ＳＳＤＣ４１は、削除通知に含まれるＬＢＡを先の図１６に示した手順に従ってＬＢＡ正引き変換する。ＳＳＤＣ４１は、削除通知に含まれるＬＢＡがトラック管理である場合は（ステップＳ１８１：Ｙｅｓ）、物理ブロックＩＤをＦＢＴ６０に追加してＡＢＴ６２から削除する（ステップＳ１８４）。一方、ＳＳＤＣ４１は、削除通知に含まれるＬＢＡがクラスタ管理である場合は（ステップＳ１８１：Ｎｏ）、物理ブロックに

対応する全クラスタをクラスタテーブル６４から削除し（ステップＳ１８２）、トラックテーブル６３において、ＬＢＡに対応するトラックに対応する物理ブロックＩＤに適当な有効値（例えばＦＦＦＦ）を記入し（ステップＳ１８３）、物理ブロックＩＤをＦＢＴ６０に追加してＡＢＴ６２から削除する（ステップＳ１８４）。ＳＳＤＣ４１は、ＮＡＮＤ整理以外に、削除通知処理によってもフリーブロックを確保することができる。

【０１２５】

このようなＮＡＮＤ整理により、書き込みに対し十分な数のフリーブロックを確保できるのが通常である。ＮＡＮＤ整理によっても、書き込みに対し十分な数のフリーブロックを確保できなかった場合は、統計情報６５のＮＡＮＤ整理失敗フラグを１にして、情報処理装置１１１による統計情報６５の取得を通じてＳＳＤＣ４１がフリーブロックを確保できなかったことを情報処理装置１１１に通知できるようにすることが望ましい。たとえば、ＮＡＮＤ整理失敗フラグが１になってから実際に記憶部２が動作しなくなるまでの時間に猶予をもたせるという観点では、

（ＮＡＮＤ整理をして確保できたフリーブロック数）＜（書き込みに必要なフリーブロック数）＋（マージン）

の条件を満たす場合にＮＡＮＤ整理失敗フラグを１にセットして信頼性情報として情報処理装置１１１に通知することが望ましい。

【０１２６】

上記ＮＡＮＤ整理は、情報処理装置１１１からライト要求を受信した時だけでなく、情報処理装置から最後に命令を受信してから所定時間経過した時、または情報処理装置１１１からスタンバイやアイドルやスリープ状態に移行するコマンドを受信した時などに実行してもよいし、ACS-3に記載のSCT Command Transportやその他ベンダーコマンドなどを通じて、ＮＡＮＤ整理を開始する命令をＳＳＤＣ４１が情報処理装置１１１から受信した時などに実行してもよい。

【０１２７】

（エラー処理）

次に、図２３を用いてＳＳＤＣ４１によるＮＡＮＤメモリ１６に関するエラー処理について説明する。情報処理装置１１１からのライト要求に対する処理やＮＡＮＤ整理処理など各種処理は通常は上記のように行われるが、ＮＡＮＤメモリ１６に対する書き込み動作（プログラム動作）で書き込みエラーが発生する場合、ＮＡＮＤメモリ１６に対する消去動作（イレーズ動作）で消去エラーが発生する場合、ＮＡＮＤメモリ１６に対する読み出し動作の際にＥＣＣエラー（誤り訂正処理の失敗）が生じる場合などがあり、これらに対する例外処理が必要となる。

【０１２８】

ＳＳＤＣ４１は、上記の何れかのエラーが発生した場合（ステップＳ１９０）、エラーが発生した物理ブロックをＢＢＴ６１に追加し（ステップＳ１９１）、エラーが発生した物理ブロックをＡＢＴ６２及びＦＢＴ６０から削除することで（ステップＳ１９２）、以後は、エラーが発生した物理ブロックにアクセスできないようにする。この際、エラーが発生した物理ブロックのデータを別の物理ブロックにコピーしてもよい。ＳＳＤＣ４１は上記エラー処理を統計情報６５に反映させる。

【０１２９】

上記では、エラー処理の一例を、読み出し処理、書き込み処理、ＮＡＮＤ整理処理に関して示したが、エラー処理はこれらの例に限らず、ＮＡＮＤメモリ１６に対する全ての読み出し処理、書き込み処理、消去処理に対して適用可能である。

【０１３０】

（寿命到達処理）

情報処理装置１１１が記憶部２を利用しているうちに、統計情報６５内に格納された値は悪化していき、記憶部２は寿命に到達する。たとえば、情報処理装置１１１が記憶部２を利用しているうちに、ＮＡＮＤメモリ１６の各ブロックの信頼性が劣化していき、バッドブロックの数が増えていき、フリーブロック数とアクティブブロック数の和が減ってい

くことになる。さらに情報処理装置１１１が記憶部２を使用していると、ＳＳＤＣ４１がＮＡＮＤ整理を実行しても、書き込み処理を行うのに十分なフリーブロック数を確保できなくなる。これが記憶部２の寿命到達の一例である。以下では、記憶部２の寿命が到達した場合の制御プログラム２００の処理を示す。

【０１３１】

制御プログラム２００は起動されると、主メモリ６に常駐して記憶部２の統計情報６５などの信頼性情報を監視する。記憶部２の統計情報６５を常に監視するためには、ＯＳ１００が領域１６Ｂ（または２０Ｂ）から領域６Ａに読み出されるときまたはその直後に、制御プログラム２００が領域１６Ｂ（または領域２０Ｂ）から領域６Ｂに読み出されるのが望ましい（図２、図３参照）。

【０１３２】

制御プログラム２００は、例えば図２４に示すように、一定時間おき（たとえば１分おき）または一定処理ごと（たとえば記憶部２へのアクセスが１０００アクセスごとや、記憶部２と授受したデータが１０ＧＢごと、など）に記憶部２から統計情報６５などの信頼性情報を取得する。なお、記憶部２が寿命に近づくほど信頼性情報を取得する頻度を上げることで、より堅牢なユーザデータの保護が可能になる。統計情報を取得する方式として、例えば、INCITS ACS-3に記述されている、メモリの自己診断機能であるS.M.A.R.T（Self-Monitoring Analysis and Reporting Technology）のコマンドである、B0h/D0h SMART READ DATAコマンドやB0h/D5h SMART READ LOGコマンドを用いてもよい。又は、NVM Express Revision 1.1に記述されている02h Get Log Pageコマンドを用いても良い。又はACS-3に記述されているSCT Command Transportやその他ベンダー独自のコマンドを用いてもよい。又は、SCSI Primary Commands-4 (SPC-4), INCITS T10/1731-D, Revision 36e (http://www.t10.org/)に記述されている4Dh LOG SENSEコマンドを用いても良い。

【０１３３】

図２７は、統計情報６５をもとに生成されるか、あるいは統計情報６５に直接格納されて、ＳＳＤＣ４１から情報処理装置１１１に信頼性情報として送信されるテーブルデータ例を示すものである。ＳＳＤＣ４１は、統計情報６５としてS.M.A.R.Tを用いる場合、図２７に示すように、統計情報６５の構成要素それぞれに対し、属性ＩＤ（attribute ＩＤ）を割り当てる。ＳＳＤＣ４１は、統計情報６５の構成要素の一部のみに属性ＩＤを割り当ててもよい。ＳＳＤＣ４１は、これら統計情報６５の構成要素において、信頼性が悪化すればするほど値が増加する要素に関しては、たとえば以下のようにして規格化を行った値であるattribute valueを計算する。

attribute value = SMAL + SMAB×（1-AMALR）×(RMAX－Raw Value)／ RMAX

【０１３４】

Raw Valueは統計情報６５に格納される値そのものである。RMAXは信頼性保証できるRaw Valueの上限値である。SMABはattribute valueの初期値としてあらかじめ設定されているパラメータであり、たとえば100が設定される。SMAL（=attribute Threshold）は、Raw Valueが上限値であるRMAXに等しくなった時にattribute valueが到達するべき値であり、あらかじめ設定されているパラメータである。SMALとし、たとえば30が設定される。AMALRは、AMALR=SMAL/SMABの関係から導かれるパラメータであり、０≦AMALR＜１となる。このようにして、ＳＳＤＣ４１は、smart情報のattribute value（図２７の「Value」）を計算して制御プログラム２００に送信する。attribute Thresholdは、図２７の「Threshold」であり、Raw Valueは、図２７の「Raw Data」である。

【０１３５】

たとえば、ＳＳＤＣ４１がRaw Valueとして消去回数平均値（統計情報Ｘ０４）を用いる場合、現在の消去回数平均値が1000回ならばRaw Data=1000となり、NANDメモリ１６の信頼性保証できる消去回数が10000回の時はRMAX=10000となり、消去回数=0の初期状態でattribute valueの初期値が100となっているようＳＳＤＣ４１が設計されている場合にはSMAB=100となり、消去回数がRMAX=10000に到達するとattribute valueはSMALに到達する。

【０１３６】

ＳＳＤＣ４１は、統計情報６５の構成要素において、信頼性が悪化すればするほど値が減少する要素に関しては、たとえば以下のようにして規格化を行った値であるattribute valueを計算する。

attribute value = SMAL + SMAB×（1-AMALR）×(Raw Value - RMIN)／（RINIT-RMIN)

RMINは信頼性保証できるRaw Valueの下限値である。RINITはRaw Valueの初期値である。

【０１３７】

RMAX、AMALRおよびSMABは各Ｘ０１～Ｘ３２に対して、それぞれ別の値を採用することができる。SMAB＝１００で、AMALR＝０．３を採用した時、採用対象の統計情報に関してattribute valueの最良値が１００（たとえば出荷直後に１００）になっており、信頼性が劣化するにつれて徐々に減少していき、記憶部２が信頼性保証できなくなったとき（統計情報のRaw ValueがRMAXに等しいかそれ以上になったとき）または記憶部２が寿命到着直前になったとき、attribute valueが３０かそれ以下の値に到達することになる。なお、制御プログラム２００は、Attribute ValueがThresholdを超過しているか否かを検知する手段として、ACS-3に記載のコマンドであるB0h/DAh SMART RETURN STATUSを用い、当該コマンドのOutputからAttribute ValueがThresholdを超過しているか否かで寿命到達を判定してもよい。

【０１３８】

記憶部２の製造業者は、たとえば図２８のように統計情報のRaw Valueと記憶部２の不良率の関係を開発段階に導き出し、不良率が許容値を超える時のRaw ValueをRMAXとして採用することが望ましい。たとえば、記憶部２の開発段階で、多数個（たとえば１００個）の試験用記憶部２群に対して高温で書き込み動作を繰り返しながら、書き込んだデータが一定時間以上正しく記憶され続けるかを検証する摩耗試験を行い、同時に統計情報をモニタしつづけ、不良率が一定割合に到達する時点での統計情報のRaw ValueをRMAXとして採用すればよい。たとえば、摩耗した記憶部２を高温状態である時間以上放置して、その後記憶部２の温度を下げ、記憶部２に対して読み出し動作を行い、読み出したデータがＥＣＣ訂正できない場合（またはＥＣＣ訂正できないデータが一定数以上の場合）、これを記憶部２の不良として定義し、不良数を、同等の試験を行った記憶部２の個数で割った値を不良率として採用すればよい。この不良率が、前記許容できる不良率を統計的に有意に下回るRaw ValueをRMAXとして採用すればよい。RMAXにある程度マージンをもたせて、

RMAX´＝RMAX－マージン

としたRMAX´をRMAXとして採用してもよい。

【０１３９】

図２７の「Worst」を制御プログラム２００が記憶部２の寿命を診断する指標として採用してもよい。「Worst」はattribute valueの最悪値としてＳＳＤＣ４１により計算される。たとえば、Worstは、たとえば記憶部２の出荷後の（または製造後の）、attribute valueの最小値である。あるいは、Worstとして、過去ある一定時間範囲内のattribute valueの最小値をWorst Valueとして採用してもよいし、ある一定回数（一定データ量）通信または処理が行われた過去にさかのぼって、その過去から現在に至るまでの最小値をworst valueとして採用してもよい。

【０１４０】

図２７の「Raw Data」（Raw Value）を制御プログラム２００が記憶部２の寿命を診断する仕様として採用してもよい。統計情報のRaw ValueがRaw Dataとして記憶部２から制御プログラム２００に送信される。この場合、制御プログラム２００は、RMAXを制御プログラム２００内にすでに保持しているか記憶部２から別途読み出すかその他記憶装置から読み出すかしてRMAXを取得し、RMAXとRaw Dataを比較し、Raw Data＞RMAXまたはRaw Data≧RMAXとなった時、記憶部２が寿命に到達したと判定する。たとえば、ＮＡＮＤ整理失敗フラグの場合、これが１である場合に記憶部２が寿命に到達したと判定する。たとえば、バッドブロック数総計の場合、これが所定値を上回った場合に記憶部２が寿命に到達したと判定する。記憶部２から情報処理装置１１１に送信されるRaw Dataとして、必ずしも統計情報のRaw Valueを出力する必要はなく、たとえば、統計情報のRaw Valueを四則演算し

た値をRaw DataとしてＳＳＤＣ４１が制御プログラム２００に送信し、同じくRMAXを四則演算した値と比較することで判定を行ってもよい。また、統計情報のRaw Valueを暗号化するなど難読化したデータをRaw DataとしてＳＳＤＣ４１が制御プログラム２００に送信し、ＳＳＤＣ４１がこれを復号化して復号化後のデータをRMAXと比較することで判定を行ってもよい。

【０１４１】

上記のようにして、制御プログラム２００は、記憶部２が寿命に到達したか否か（記憶部２が異常状態であるか否か）を判定し、記憶部２が寿命に到達したと判定された場合（記憶部２が異常状態であると判定された場合）、後述する寿命到達時処理（ステップＳ２０５）に移行する。統計情報６５は、統計情報Ｘ０１～Ｘ３２以外にも種々の形態をとりうるが、本発明はこれらに対しても適用可能である。また、統計情報と不良率の関係に正の相関関係が存在する場合にだけでなく、統計情報と不良率の関係に負の相関関係が存在する場合に対しても本発明は適用可能である。たとえば、記憶部２が出荷後に体験した最低温度などである。その場合、RMAXの代わりに、信頼性保証できる下限値RMINを採用し、統計情報がRMINを下回った場合に記憶部２が寿命に到達したと判定すればよい。

【０１４２】

本実施形態では、制御プログラム２００は図２４に示すようにS.M.A.R.Tを用いて一定時間おき（たとえば１分おき）に統計情報の取得を行う（ステップＳ２００：Yes）。制御プログラム２００は、統計情報取得コマンドであるACS-3に記載のB0h/D0h SMART READ DATAを発行し（ステップＳ２０１）、記憶部２から統計情報を含むデータを受信し（ステップＳ２０２）、この受信したデータを診断する（ステップＳ２０３）。診断方法は前述したとおりである。ステップＳ２０４において、記憶部２が寿命に到達したと制御プログラム２００が判定した時、または記憶部２が寿命到達目前であると制御プログラム２００が判定した時（ステップＳ２０４：Ｙｅｓ）、制御プログラムは寿命到達時処理に移行する（ステップＳ２０５）。記憶部２が寿命に到達していなくても、統計情報があらかじめ定められたRMAXを上回ったり、あるいは通常動作ではありえない異常値を示した場合などにおいても、ステップＳ２０５の処理に移行することが望ましい。

【０１４３】

特許文献２２や特許文献２３の寿命予測技術を用いて記憶部２の寿命予測を行い、記憶部２の寿命がある一定期間後に迫っていると判定された場合に、寿命到達時処理に移行してもよい。

【０１４４】

統計情報６５以外の信頼性情報を用いて、寿命到達時処理に移行してもよい。たとえば、図２９に示すように、制御プログラム２００は、ＯＳ１００が記憶部２から受信する応答情報（図７参照）をＯＳ１００から取得（監視）して信頼性情報として用い（ステップＳ２１０）、エラー応答であった場合（ステップＳ２１１）、記憶部２が異常状態に至ったと判定し、寿命到達時処理に移行する（ステップＳ２０５）。監視する応答はどのようなコマンドに対する応答でもよいが、たとえばACS-3記載の61h WRITE FPDMA QUEUEDや35h WRITE DMA EXTや、NVM Express Revision 1.1記載の01h Writeコマンドなど、記憶部２に対する書き込みコマンドに対する応答のみを監視することが、ＣＰＵ５への負荷低減の観点から望ましい。特に、記憶部２が特許文献３の発明を採用したＳＳＤである場合、記憶部２が寿命に至ると、記憶部２への書き込みコマンドに対する応答がエラーで返ってくるようになるため、統計情報の取得無しに寿命到達を判定することができる。なお、記憶部２が特許文献３の発明を採用したＳＳＤではない場合であっても本発明が適用できることはもちろんである。

【０１４５】

なお、特許文献３の発明を採用した記憶部２である場合、記憶部２が書き込みコマンドに対してエラーを返す状態である時には、後述する記憶部２のBoot Loader領域の書き換えを行う場合は、特許文献３のRead Onlyモード状態において特殊な書き込みコマンド（例えばACS-3に記述されているSCT Command Transportやその他ベンダー独自コマンドなど

）に対してはエラーを返さないようＳＳＤＣ４１を構成しておき、前述した特殊な書き込みコマンドを用いて記憶部２に書き込みを行うことが望ましい。なお、この特殊な書き込みコマンドは、記憶部２以外の記憶装置への書き込みには用いなくてもよい。あるいは、ＯＳ１００が、書き込みコマンドとしてある書き込みコマンド（たとえば６1h WRITE FPDMA QUEUED）のみを使うようなＯＳである場合には、ＳＳＤＣ４１が特許文献３のRead Onlyに至ると、当該書き込みコマンド（たとえば６1h WRITE FPDMA QUEUED）に対してエラーを返すようにＳＳＤＣ４１を構成し、別の書き込みコマンド（たとえば３０h WRITE SECTOR(S)）に対してはエラーを返さないようにＳＳＤＣ４１を構成し、別の書き込みコマンド（たとえば３０h WRITE SECTOR(S)）を用いて記憶部２に対してBoot Loader領域の書き込みを行うようにしてもよい。

【０１４６】

監視対象のコマンドは、書き込みコマンド以外のコマンドであってもよいのはもちろんである。たとえば、コマンド応答としてACS-3記載のB0H/D4H SMART EXECUTE OFF-LINE IMMEDIATEの応答(Outputs)やレポートを監視してもよいし、90h EXECUTE DEVICE DIAGNOSTICの応答を監視してもよい。たとえば、制御プログラム２００は、記憶部２から取得されるSMARTのself-testの結果を用いて、寿命到達時処理に移行してもよい。制御プログラム２００は、ACS-3に記載のB0h/D4h SMART EXECUTE OFF-LINE IMMEDIATEコマンドを記憶部２に送信することでＳＳＤＣ４１はself-testを実行し、ACS-3に記載のB0h/D0h SMART READ DATAやB0h/D5h SMART Read Logコマンドを記憶部２に送信することでself-testの結果を信頼性情報として取得する。たとえば、制御プログラム２００は取得したself-testの結果にエラーが含まれていることを検知したとき、記憶部２が寿命に到達したと判定する。

【０１４７】

あるコマンド応答がエラーであっても、同コマンドを再度送信するとエラーでない可能性があり、この場合は記憶部２が寿命に到達していない可能性があるため、再現性のあるコマンドエラーが発生した場合にのみ寿命到達時処理を行う観点では、コマンドエラーが複数回発生した場合に寿命到達時処理を行うことが望ましい。さらに、エラー再現性を厳密に判定する観点では、コマンドエラーが複数回連続で発生した場合に寿命到達時処理を行うことが望ましい。あるいは、図３０に示すように、記憶部２へのコマンド監視中にコマンド応答がエラーで返ってきた場合（ステップＳ２２０、Ｓ２２１:Yes）、制御プログラム２００またはＯＳ１００が同コマンドを記憶部２に再度送信（コマンドリトライ）し（ステップＳ２２２）、リトライしたコマンドがエラーした場合に（ステップＳ２２３:Yes）、寿命到達時処理を行うようにしてもよい（ステップＳ２０５）。

【０１４８】

また、制御プログラム２００は、情報処理装置１１１から取得される信頼性情報を用いて、寿命到達時処理に移行してもよい。たとえば、情報処理装置１１１内に温度計が設置されているとき、制御プログラム２００は温度計から出力される温度を監視し、温度が上限値を上回った場合や下限値を下回った場合に、信頼例劣化時処理としての寿命到達時処理を行うようにしてもよい。

【０１４９】

（通常状態のデータ構成）

図３１は寿命到達時処理Ｓ２０５を行うより以前の、情報処理装置１１１が管理するデータの構成例である。前述のとおり、情報処理装置１１１はＮＡＮＤメモリ１６の物理アドレスを直接指定してデータの読み出し、書き込みを記憶部２に要求するのではなく、論理アドレスであるＬＢＡを指定してデータの読み出し、書き込み命令を記憶部２に送信する。記憶部２内のＳＳＤＣ４１は、ＬＢＡと、ＮＡＮＤメモリ１６の物理アドレスの対応を管理情報４４によって動的に割り当てる。このように、情報処理装置１１１が直接管理できるデータはＬＢＡによって管理されており、情報処理装置１１１によって管理可能なアドレス空間として、ＬＢＡ領域２００１が記憶部２に割り当てられている。ＬＢＡ領域２００１は、Boot Loader領域２００２と、メタデータ領域２００３と、ユーザデータ領

域２００４からなる。なお、こららの領域のうち一部の領域をACS-3のSMART Read Logコマンドや Read Logコマンドでアクセス可能なログページアドレスに割り当ててもよい。記憶部２は、論理ドライブ４に割り当てられている。

【０１５０】

Boot Loader領域２００２は、情報処理装置１１１の起動時に読み出される領域である。本実施形態ではBoot Loader領域２００２はＬＢＡの固定領域に割り当てられているとするが、情報処理装置１１１はBoot Loader領域２００２を動的に割り当て可能としてもよい。なお、Boot Loader領域の例として、たとえばMaster Boot Record（ＭＢＲ)があり、ＭＢＲではたとえばＬＢＡ=0x0000の合計１論理セクタ（合計５１２Byte）の領域が固定Boot Loader領域として割り当てられている。また、Boot Loader領域の例として、たとえばＧＵＩＤパーティションテーブル(ＧＰＴ)が存在する。本実施形態では、Boot Loader領域２００２は、図３２に示すように、メタデータ領域２００３の先頭ＬＢＡを格納したメタデータへのポインタ領域２００５と、ディスクステータスを格納したステータス格納領域２００６と、データ引き継ぎ対象記憶装置アドレスを格納したアドレス領域２００７よりなる。例えば、記憶部２がデータ引き継ぎ元になり、記憶部３がデータ引き継ぎ先になる場合、記憶部２のアドレス領域２００７には、記憶部３のディスク識別ＩＤが格納され、記憶部３のアドレス領域２００７には、記憶装置のディスク識別ＩＤが格納される。本実施形態では、ステータス格納領域２００６に格納されるデータは、０～５の値を取りうる。それぞれの値は、格納先の記憶部がそれぞれ

０：初期ディスク状態
１：通常状態
２：信頼性劣化状態
３：データ引き継ぎ元状態（保護状態）
４：データ引き継ぎ先状態
５：廃棄対象状態

の状態であることを示す。

【０１５１】

情報処理装置１１１は起動時にポインタ２００５を読み出してメタデータ領域２００３のＬＢＡを特定し、メタデータ３００をＬＢＡ領域２００１のメタデータ領域２００３から主メモリ６のメタデータ領域６Ｃに読み出す。ＯＳ１００はファイルの書き換えが発生すると主メモリ６のメタデータ領域６Ｃのメタデータ３００を書き換え、さらにメタデータ領域６Ｃのメタデータ３００を定期的に記憶部２のメタデータ領域２００３に退避保存したり、メタデータ３００のジャーナルを記憶部２のメタデータ領域２００３に逐次記録したりする。

【０１５２】

図３３はメタデータ３００の構成例である。ファイル識別ＩＤはアプリケーションプログラム４００がデータを識別するために用いられるデータのアドレスまたはファイル名であり、論理ドライブアドレスは論理ドライブ４（図３１、図３８参照）を特定するために用いられる各論理ドライブに割り当てられたアドレスである。また、ディスク識別ＩＤは、記憶部２や記憶部３などの物理的な記憶装置を特定するために用いられる、物理的記憶装置に割り当てられたアドレスである。本実施形態においては、ディスク識別ＩＤとして、ＷＷＮ (World Wide Name)を用いる。ＷＷＮは各記憶装置それぞれに異なる値が割り当てられているため、物理的な記憶装置を区別するのに用いることができる。ＷＷＮは、たとえば、ACS-3に記述されているECh Identify Deviceコマンドで読み出されるデータのWord108-111に格納される。あるいは、ディスク識別ＩＤとして、たとえばACS-3に記述されているECh Identify Deviceコマンドで読み出されるデータのWord10-19に割り当てられたシリアル番号を用いてもよいし、NVM Express Revision 1.1に記述されている06h Identifyコマンドで読み出されるデータのByte23:04に割り当てられたSerial Number（ＳＮ)を用いてもよいし、ネットワーク・プロトコルにおけるＭＡＣ（Media Access Control）アドレスやＩＰ(Internet Protcol)アドレスを用いてもよい。ディスク識別ＩＤは記憶装置

の製造時にあらかじめ記憶装置に書き込まれていてもよいし、情報処理装置１１１との接続時に情報処理装置１１１によって新規に割り当てられたりしてもよい。メタデータ３００のＬＢＡにはメタデータ３００によって紐付けされているユーザデータ領域２００４のＬＢＡアドレスが格納されている。メタデータ３００のＬＢＡにメタデータ領域２００３やBoot Loader領域２００２のＬＢＡなど、ユーザデータ領域２００４以外のＬＢＡが格納されてもよい。セクタカウントは、データ長を示している。

【０１５３】

メタデータ３００は、ＯＳ１００がファイル識別ＩＤおよび論理ドライブアドレスからディスク識別ＩＤとＬＢＡとセクタカウントに正引きしたり、ＯＳ１００がディスク識別ＩＤとＬＢＡとセクタカウントから論理ドライブアドレスとファイル識別ＩＤを逆引きしたりするのに用いられる。通常、アプリケーションプログラム４００は記憶部２やＬＢＡを直接指定して記憶装置への読み出し、書き込みを行わない。ＯＳ１００は記憶部２を論理ドライブ４として認識（マウント）し、アプリケーションプログラム４００が論理ドライブアドレスおよびファイルＩＤをＯＳ１００に送信すると、ＯＳ１００がメタデータ３００を読みだして論理ドライブおよびファイルＩＤに対応する記憶装置とＬＢＡを特定し、当該記憶装置に命令とＬＢＡを送信する。

【０１５４】

論理ドライブは、図７におけるＯＳ１００および制御プログラム２００によって、下位の階層である１乃至複数台の物理ドライブやあるいはそれらの一部のＬＢＡ領域に割り当てられ、それにより上位の階層であるアプリケーションプログラム４００は論理ドライブを仮想的に１台のドライブとして認識する。本実施形態においては、寿命到達時処理Ｓ２０５以前の状態において、論理ドライブ４が１つの物理記憶装置である記憶部２に割り当てられる場合について述べる。なお、寿命到達時処理Ｓ２０５以前の状態であっても、複数の物理記憶装置を用いてRedundant Arrays of Inexpensive Disks（ＲＡＩＤ)のディスクアレイ、たとえばＲＡＩＤ０やＲＡＩＤ５のディスクアレイを構成し、これを一つの論理ドライブ４として認識してもよく、その場合であっても本発明は適用可能である。

【０１５５】

図３４はアプリケーションプログラム４００がＯＳ１００に対して論理ドライブ４へのアクセス要求を行った場合のフローチャートである。アプリケーションプログラム４００がファイル読み出し命令と論理ドライブアドレスとファイル識別ＩＤをＯＳ１００に送信すると（ステップＳ３００）、ＯＳ１００はメタデータ領域６Ｃからメタデータ３００を読みだす（ステップＳ３０１）。ＯＳ１００は、論理ドライブアドレスとファイル識別ＩＤをディスク識別ＩＤとＬＢＡに正引き変換し（ステップＳ３０２）、ディスク識別ＩＤに対応する記憶装置に対して命令およびＬＢＡを送信する（ステップＳ３０３）。記憶装置は、命令に従ってユーザデータ領域２００４への書き込みまたはユーザデータ領域２００４からの読み出し動作を行い、ＯＳ１００に応答またはデータを送信する（ステップＳ３０４）。ＯＳ１００は、記憶装置から応答およびデータを受信し、さらにアプリケーションプログラム４００に応答およびデータを送信して処理が終了する（ステップＳ３０５）。

【０１５６】

本実施形態においては、記憶部２に関する寿命到達時処理Ｓ２０５を行う以前の状態においては、論理ドライブ４は記憶部２のみから構成されているため、メタデータ３００における論理ドライブ４に対応するディスク識別ＩＤは全て記憶部２のＷＷＮとなっている。一方、論理ドライブにおいてＲＡＩＤアレイを構築する場合のように、寿命到達時処理Ｓ２０５を行う以前であっても、メタデータ３００において一つの論理ドライブアドレスに対して複数のディスク識別ＩＤが割り当てられていてもよい。

【０１５７】

ディスクステータス領域２００６には、ＯＳ１００に対する記憶装置の情報が格納される。図３５は情報処理装置１１１が起動した時、およびインタフェース１９に記憶装置が接続された時の制御プログラム２００の処理手順を示すものである。制御プログラム２０

０はインタフェース１９を経由して記憶装置のBoot Loader領域２００２のディスクステータス領域２００６を読み出し、読み出された値に応じてＯＳ１００に通知する記憶装置の状態を変更する。ディスクステータス＝０であるとき（ステップＳ３１１）、制御プログラム２００はＯＳ１００に対して記憶装置が初期ディスクであると通知し、ＯＳ１００は記憶装置を初期ディスク状態として認識する（ステップＳ３１２）。記憶装置の工場出荷時や、ACS-3のF4h Security Erase Unitコマンドや、NVM Express Revision 1.1の80h Format NVMコマンドなどをもちいることによって情報処理装置１１１が記憶装置を消去したとき、当該記憶装置のディスクステータス領域２００６の値はディスクステータス＝０に変更される。また、情報処理装置１１１が記憶装置をフォーマットした時、当該記憶装置のディスクステータス領域２００６の値はディスクステータス＝０に変更される。

【０１５８】

ディスクステータス＝１であるとき（ステップＳ３１３）、制御プログラム２００はＯＳ１００に対して記憶装置が通常状態であると通知し、ＯＳ１００は記憶装置を通常状態として認識する（ステップＳ３１４）。寿命到達時処理Ｓ２０５を行う前の記憶部２のディスクステータスはディスクステータス＝１となっている。

【０１５９】

ディスクステータス＝２であるとき（ステップＳ３１５）、制御プログラム２００はＯＳ１００に対して記憶装置が信頼性劣化状態であると通知し、制御プログラム２００は記憶装置が信頼化劣化状態として認識し（ステップＳ３１６）、寿命到達時処理Ｓ２０５を行う。

【０１６０】

ディスクステータス＝３であるとき（ステップＳ３１７）、制御プログラム２００はＯＳ１００に対して記憶装置がデータ引き継ぎ元としてデータ引き継ぎ作業中であると通知し、ＯＳ１００は記憶装置を保護状態として認識する（ステップＳ３１８）。

【０１６１】

ディスクステータス＝４であるとき（ステップＳ３１９）、制御プログラム２００はＯＳ１００に対して記憶装置がデータ引き継ぎ先としてデータ引き継ぎ作業中であると通知し、ＯＳ１００は記憶装置をデータ引き継ぎ先状態として認識する（ステップＳ３２０）。

【０１６２】

ディスクステータス＝５であるとき（ステップＳ３２１）、制御プログラム２００はＯＳ１００に対して記憶装置が廃棄対象状態であると通知し、ＯＳ１００は記憶装置を廃棄対象状態として認識する（ステップＳ３２２）。ディスクステータスが０～５以外であるときは、不正なディスクとみなし、マウントしないことが望ましい（ステップＳ３２３）。

【０１６３】

図３６および図３７は記憶装置のライフサイクルを示す状態遷移図である。図３６に示すように、記憶装置製造直後、記憶装置購入直後、記憶装置消去直後、記憶装置フォーマット直後の記憶装置のディスクステータス領域２００６の値は初期ディスク状態を示すディスクステータス＝０となっており（ステップＳ３３０）、その後通常状態を示すディスクステータス＝１となることによって、ＯＳ１００に通常の記憶装置として利用され（ステップＳ３３１）、記憶装置が信頼性劣化するにつれて信頼性劣化状態を示すディスクステータス＝２へと遷移し（ステップＳ３３２）、さらに保護状態を示すディスクステータス＝３を経由した（ステップＳ３３３）後、廃棄対象状態を示すディスクステータス＝５となった（ステップＳ３３４）後、最終的に廃棄される。

【０１６４】

図３７は、情報処理装置１１１にデータ引き継ぎ先状態としてマウントされた場合の記憶装置のライフサイクルを示すもので、この場合は、ステップＳ３３０のディスクステータス＝０の初期ディスク状態の後、データ引き継ぎ先状態（ディスクステータス＝４）を経由した（ステップＳ３３０ｂ）後、データ引き継ぎ元状態の記憶装置が脱着されること

で、ディスクステータス＝１の通常状態として利用されている（ステップＳ３３１）。
【０１６５】
図３８は、記憶部２が信頼性劣化状態として認識されたときに、記憶部２とは異なる別の記憶部３が接続された状態を示すものである。新たに接続された記憶部３にも、情報処理装置１１１によって管理可能なアドレス空間として、ＬＢＡ領域３００１が割り当てられている。ＬＢＡ領域３００１は、Boot Loader領域３００２と、メタデータ領域３００３と、ユーザデータ領域３００４からなる。
【０１６６】
なお、記憶部２が信頼性劣化状態として認識されたとき、初期ディスク状態の記憶装置が情報処理装置１１１に接続されると該初期ディスク状態の記憶装置をデータ引き継ぎ先状態として認識するのではなく、図３９の破線部のように、すでに情報処理装置１１１に接続されている通常状態の記憶部３をデータ引き継ぎ先状態として認識するようにしてもよい。
【０１６７】
（寿命到達時処理）
図４０に寿命到達時処理Ｓ２０５の際に、制御プログラム２００が行う処理のフローチャートを示す。制御プログラム２００は、接続されている記憶部２が寿命到達するかあるいは寿命到達直前になって寿命到達時処理が開始されると、記憶部２のBoot Loader領域２００２のディスクステータス領域２００６がディスクステータス＝２（信頼性劣化状態）であるか否かを判定する（ステップＳ３４０）。この判定結果が否である場合、制御プログラム２００は、ディスクステータス領域２００６をディスクステータス＝２に書き換え（ステップＳ３４１）、手順をステップＳ３４２に移行する。ステップＳ３４０の判定結果がＹｅｓの場合は、、手順をステップＳ３４２に移行する。
【０１６８】
ステップＳ３４２のように、制御プログラム２００は、「新しいディスクをインタフェース１９に接続してください」などの新しいディスクの接続を促すメッセージをディスプレイ９に表示することが望ましい。制御プログラム２００は、ディスクステータス＝０（初期ディスク状態）の記憶装置が接続されたか否かを判定する（ステップＳ３４３）。
【０１６９】
新たな記憶装置が接続されると、制御プログラム２００は、記憶部２のBoot Loader領域２００２を、新たに接続されたディスクステータス＝０（初期ディスク状態）の記憶部３のBoot Loader領域３００２にコピーする（ステップＳ３４４、図３２参照）。制御プログラム２００は、記憶部３のBoot Loader領域３００２のディスクステータス領域３００６をディスクステータス＝４（データ引き継ぎ先状態）に書き換える（ステップＳ３４５）。
【０１７０】
なお、ステップＳ３４０：ＹｅｓやＳ３４１の時点で、すでにディスクステータス＝０（初期ディスク状態）の記憶装置がすでに接続されている場合や、データ引き継ぎ先として割り当て可能なディスクステータス＝１（通常状態）の記憶装置がすでに接続されている場合には、制御プログラム２００は、この記憶装置を記憶部３とし、記憶部２のBoot Loader領域２００２を、記憶部３のBoot Loader領域３００２にコピーし、記憶部３のBoot Loader領域３００２のディスクステータス領域３００６をディスクステータス＝４に書き換えてもよい。
【０１７１】
制御プログラム２００は、記憶部３のBoot Loader領域３００２のアドレス領域３００７に記憶部２のディスク識別ＩＤを書き込む（ステップＳ３４６）。制御プログラム２００は、記憶部２のBoot Loader領域２００２のディスクステータス領域２００６をディスクステータス＝３（保護状態）に書き換える（ステップＳ３４７）。制御プログラム２００は、記憶部２のBoot Loader領域２００２のアドレス領域２００７に記憶部３のディスク識別ＩＤを書き込む（ステップＳ３４６）。記憶部２のアドレス領域２００７および記

憶部３のアドレス領域を見ることで、記憶部２と記憶部３がデータ引き継ぎ処理のペアであることを認識することができる。

【０１７２】

制御プログラム２００は、最新のメタデータ３００を主メモリ６からまたは記憶部２のメタデータ格納領域２００３から読み出して、読み出した最新のメタデータを記憶部３のメタデータ格納領域３００３に書き込んでコピーする(ステップＳ３４９)。制御プログラム２００は、記憶部２＋記憶部３を１つの論理ドライブ４としてＯＳ１００に認識させる（ステップＳ３５０）。制御プログラム２００は、主メモリ６上の領域６Ｄの論理ドライブステータステーブル４５０を、ステータスが「通常状態」から「データ引き継ぎ状態」になるように書き換える（ステップＳ３５１）。

【０１７３】

図４１は、主メモリ６の領域６Ｄに格納されている論理ドライブステータステーブル４５０を示すものである。テーブル４５０では、複数の論理ドライブとステータスの対応が管理されている。制御プログラム２００は、論理ドライブの状態（通常状態またはデータ引き継ぎ状態）に応じて、テーブル４５０を随時書き換える。

【０１７４】

本実施形態では、寿命到達時処理Ｓ２０５によって、記憶部２と記憶部３は、図３８に示したように、単一の論理ドライブである論理ドライブ４として認識される。新しい記憶部３が接続されてから、記憶部３を用いた論理ドライブ４が構築されるまでのデータの読み出し、書き込みは、高々、メタデータ３００のデータ量しか発生しないため、ＲＡＩＤディスクアレイでディスクを交換して論理ドライブにマウントするまでの時間と比べ、記憶部３のマウントが非常に高速に行われる。

【０１７５】

寿命到達時処理Ｓ２０５以前においては、主メモリ６のメタデータ領域６Ｃのメタデータ３００やメタデータ３００のジャーナルはＯＳ１００によって定期的にデータ引き継ぎ元の記憶部２のメタデータ領域２００３に退避保存されていたが、寿命到達時処理Ｓ２０５以後においては、主メモリのメタデータ領域６Ｃのメタデータ３００やメタデータ３００のジャーナルはＯＳ１００によって定期的にデータ引き継ぎ先の記憶部３のメタデータ領域３００３に退避保存される。これにより、メタデータ領域３００３には最新に近いメタデータが格納され、メタデータ領域２００３には寿命到達時処理Ｓ２０５以前の古いメタデータが格納されることになる。

【０１７６】

なお、上記の説明では、ディスクステータス＝２の信頼性劣化状態を定義することにしたが、制御プログラム２００は、信頼性情報と閾値との比較の結果、記憶部２が寿命に到達したと判定した場合、ディスクステータス＝２の信頼性劣化状態を経ることなく、ディスクステータスをディスクステータス＝３のデータ引き継ぎ元状態（保護状態）に即座に移行させるようにしてもよい。すなわち、この場合は、制御プログラム２００は、寿命到達時処理が開始されると、図４０のステップＳ３４０において、記憶部２のBoot Loader領域２００２のディスクステータス領域２００６がディスクステータス＝３（データ引き継ぎ元状態）であるか否かを判定する。この判定結果が否である場合、制御プログラム２００は、図４０のステップＳ３４１において、ディスクステータス領域２００６をディスクステータス＝３に書き換え、手順をステップＳ３４２に移行する。それ以降の手順は、ステップＳ３４７が削除される以外は、図４０と同様である。

【０１７７】

（論理ドライブへの書き込み）

図４２はアプリケーションプログラム４００からＯＳ１００にファイルデータ書き込み要求が送信された時の、ＯＳ１００の処理手順を示すものである。ＯＳ１００は、アプリケーションプログラム４００から書き込み命令と、論理ドライブアドレスと、ファイルＩＤと、データを受信する（ステップＳ３６０）。ＯＳ１００は、主メモリ６から論理ドライブステータステーブル４５０を読み出し（ステップＳ３６１）、主メモリ６からメタデ

ータ３００を読み出し（ステップＳ３６２）、メタデータ３００を参照してデータ書き込み用のＬＢＡを割り当てる（ステップＳ３６３）。

【０１７８】

ＯＳ１００は、論理ドライブステータステーブル４５０に基づいて書き込み命令で指定された論理ドライブが通常状態かデータ引き継ぎ状態であるかを判定する（ステップＳ３６４）。論理ドライブが通常状態である場合、ＯＳ１００は、記憶部２に対し、書き込み命令、ＬＢＡおよび書き込みデータを送信する（ステップＳ３６５）。ＯＳ１００は、記憶部２から応答を受信する（ステップＳ３６６）。ＯＳ１００は、主メモリ６上のメタデータを書き換えて、書き込みファイルＩＤと記憶部２とＬＢＡとセクタカウントを紐付けする（ステップＳ３６７）。ＯＳ１００は、アプリケーションプログラム４００に応答を送信する（ステップＳ３７１）。

【０１７９】

論理ドライブがデータ引き継ぎ状態である場合、ＯＳ１００は、データ引き継ぎ先の記憶部３に対し、書き込み命令、ＬＢＡおよび書き込みデータを送信する（ステップＳ３６８）。ＯＳ１００は、記憶部３から応答を受信する（ステップＳ３６９）。ＯＳ１００は、主メモリ６上のメタデータを書き換えて、書き込みファイルＩＤと記憶部３とＬＢＡとセクタカウントを紐付けする（ステップＳ３７０）。ＯＳ１００は、アプリケーションプログラム４００に応答を送信する（ステップＳ３７１）。すなわち、ＯＳ１００は、論理ドライブがデータ引き継ぎ状態である場合、データ引き継ぎ先の記憶部３への書き込みに伴って、主メモリ６上のメタデータを書き換えて、記憶部２および記憶部３の記憶データのアドレスを更新する。書き込みに伴う記憶データのアドレスの更新とは、記憶部３への書き込み処理と同時並行的に行うこと、記憶部３への書き込み処理の中で行うこと、記憶部３への書き込み処理を行う前に行うこと、記憶部３への書き込み処理を行った後に行うことを全て含む。

【０１８０】

（論理ドライブへのファイル削除命令）

図４３はアプリケーションプログラム４００からＯＳ１００にファイル削除要求が送信された時の、ＯＳ１００の処理手順を示すものである。ＯＳ１００は、アプリケーションプログラム４００から削除命令と、論理ドライブアドレスと、ファイルＩＤを受信する（ステップＳ９００）。ＯＳ１００は、主メモリ６から論理ドライブステータステーブル４５０を読み出し（ステップＳ９０１）、主メモリ６からメタデータ３００を読み出し（ステップＳ９０２）、メタデータ３００を参照して論理ドライブアドレスとファイルＩＤをディスク識別ＩＤとＬＢＡに正引き変換する（ステップＳ９０３）。ＯＳ１００は、主メモリ６内のメタデータから削除対象ファイルのファイルＩＤが含まれる行を削除するか、主メモリ６内のメタデータ上の削除対象ファイルのファイルＩＤを無効なＩＤに書き換えるかすることで、削除対象ファイルＩＤをメタデータ３００から削除する（ステップＳ９０４）。

【０１８１】

ＯＳ１００は、論理ドライブステータステーブル４５０に基づいて削除命令で指定された論理ドライブが通常状態かデータ引き継ぎ状態であるかを判定する（ステップＳ９０５）。論理ドライブが通常状態である場合、ＯＳ１００は、記憶部２に対し、削除通知およびＬＢＡを送信する（ステップＳ９０６）。ＯＳ１００は、記憶部２から応答を受信する。ＯＳ１００は、アプリケーションプログラム４００に応答を送信する（ステップＳ９１０）。

【０１８２】

論理ドライブがデータ引き継ぎ状態である場合、ＯＳ１００は、正引き変換後のディスク識別ＩＤがデータ引き継ぎ元である記憶部２であるかデータ引き継ぎ先の記憶部３であるかを判定する（ステップＳ９０７）。正引き変換後のディスク識別ＩＤが記憶部２である場合、ＯＳ１００は、記憶部２に対し、削除通知およびＬＢＡを送信し（ステップＳ９０８）、記憶部２から応答を受信し、アプリケーションプログラム４００に応答を送信す

る（ステップＳ９１０）。正引き変換後のディスク識別ＩＤが記憶部３である場合、ＯＳ１００は、記憶部３に対し、削除通知およびＬＢＡを送信し（ステップＳ９０９）、記憶部３から応答を受信し、アプリケーションプログラム４００に応答を送信する（ステップＳ９１０）。

【０１８３】

（論理ドライブからの読み出し）

図４４はアプリケーションプログラム４００からＯＳ１００にファイルデータ読み出し要求が送信された時の、ＯＳ１００の処理手順を示すものである。ＯＳ１００は、アプリケーションプログラム４００から読み出し命令と、論理ドライブアドレスと、ファイルＩＤとを受信する（ステップＳ３８０）。ＯＳ１００は、主メモリ６から論理ドライブステータステーブル４５０を読み出し（ステップＳ３８１）、主メモリ６からメタデータ３００を読み出し（ステップＳ３８２）、メタデータ３００を参照して、論理ドライブアドレスとファイルＩＤをデータ読み出し用のディスク識別ＩＤとＬＢＡとセクタカウントに正引き変換する（ステップＳ３８３）。

【０１８４】

正引き変換後のディスク識別ＩＤが記憶部２を指定している場合（ステップＳ３８４）、ＯＳ１００は、記憶部２に対し、読み出し命令、ＬＢＡおよびセクタカウントを送信する（ステップＳ３８５）。ＯＳ１００は、記憶部２から応答および読み出しデータを受信する（ステップＳ３８６）。ＯＳ１００は、アプリケーションプログラム４００に読み出しデータおよび応答を送信する（ステップＳ３８９）。

【０１８５】

正引き変換後のディスク識別ＩＤが記憶部３を指定している場合（ステップＳ３８４）、ＯＳ１００は、記憶部３に対し、読み出し命令、ＬＢＡおよびセクタカウントを送信する（ステップＳ３８７）。ＯＳ１００は、記憶部３から応答および読み出しデータを受信する（ステップＳ３８８）。ＯＳ１００は、アプリケーションプログラム４００に読み出しデータおよび応答を送信する（ステップＳ３８９）。例えば、ＬＢＡ＝０がデータ引き継ぎ済み、ＬＢＡ＝１が未データ引き継ぎの状態で、ＬＢＡ＝０、セクタカウント＝１で読み出しを行うと、記憶部３から読み出しが行われ、ＬＢＡ＝１、セクタカウント＝１で読み出しを行うと、記憶部２から読み出しが行われ、ＬＢＡ＝０、セクタカウント＝２で読み出しを行うと、記憶部２および記憶部３から読み出しが行われる。

【０１８６】

このように、データ引き継ぎ元記憶装置への書き込みを禁止し、データ引き継ぎ先の記憶部３への書き込みを利用してデータ引き継ぎ元記憶装置からデータ引き継ぎ先記憶装置へのデータ引き継ぎを実現しているので、ユーザ自身によるバックアップ作業は不要となる。また、データ引き継ぎ中において、ユーザデータ２００４のコピーそのものは行われず、データ引き継ぎはユーザデータ２００４の新規書き込み処理を利用して行われるため、データ引き継ぎ中であってもアプリケーションプログラム４００の書き込み処理性能は低下しない。また、寿命到達時処理Ｓ２０５後において、記憶部２に発生する書き込み処理はたかだかディスクステータス領域２００６に対する書き込みのみであるため、記憶部２への書き込み処理はほとんど発生しない。このようにして、記憶部２に対する寿命到達時処理Ｓ２０５以後であっても、アプリケーションプログラム４００にとって論理ドライブ４そのものは読み出しおよび書き込み可能なドライブとして認識されるが、情報処理装置１１１にとって実際には記憶部２はリードオンリーデバイスのように扱われることになる。

【０１８７】

（書き戻しバックアップ）

なお、データ引き継ぎ元である記憶部２のデータを主メモリ６内のキャッシュメモリ領域に読みだした時には、キャッシュメモリ領域に読みだしたデータをデータ引き継ぎ先である記憶部３に書き込み（書き戻し）、書き込み先ＬＢＡとファイルＩＤを関連付けるようにメタデータ３００を書き換えてもよい。以下、図４５を用いて説明する。

【０１８８】
ＯＳ１００は、アプリケーションプログラム４００から読み出し命令と、論理ドライブアドレスと、ファイルＩＤとを受信する（ステップＳ４００）。ＯＳ１００は、主メモリ６から論理ドライブステータステーブル４５０を読み出し（ステップＳ４０１）、主メモリ６からメタデータ３００を読み出し（ステップＳ４０２）、メタデータ３００を参照して、論理ドライブアドレスとファイルＩＤをデータ読み出し用のディスク識別ＩＤとＬＢＡとセクタカウントに正引き変換する（ステップＳ４０３）。
【０１８９】
正引き変換後のディスク識別ＩＤが記憶部３を指定している場合（ステップＳ４０４）、ＯＳ１００は、記憶部３に対し、読み出し命令、ＬＢＡおよびセクタカウントを送信する（ステップＳ４０９）。ＯＳ１００は、記憶部３から応答および読み出しデータを受信する（ステップＳ４１０）。ＯＳ１００は、アプリケーションプログラム４００に記憶部３から読み出したデータおよび応答を送信する（ステップＳ４１１）。
【０１９０】
正引き変換後のディスク識別ＩＤが記憶部２を指定している場合（ステップＳ４０４）、ＯＳ１００は、記憶部２に対し、読み出し命令、ＬＢＡおよびセクタカウントを送信する（ステップＳ４０５）。さらに、ＯＳ１００は、記憶部３に対し、書き込み命令、ＬＢＡおよびセクタカウントを送信する（ステップＳ４０６）。ＯＳ１００は、記憶部２から応答および読み出しデータを受信する（ステップＳ４０７）。ＯＳ１００は、記憶部３に対し、書き込み命令、ＬＢＡおよび記憶部２から読み出したデータを送信し、これにより記憶部２から読み出したデータを記憶部３に書き込むバックグラウンド書き込みを行う（ステップＳ４０８）。ＯＳ１００は、アプリケーションプログラム４００に記憶部２から受信したデータおよび応答を送信する（ステップＳ４１２）。ＯＳ１００は、主メモリ６上のメタデータを書き換えて、書き込みファイルＩＤと記憶部３とＬＢＡとセクタカウントを紐付けする（ステップＳ４１３）。ＯＳ１００は、アプリケーションプログラム４００に応答を送信する（ステップＳ３７２）。
【０１９１】
このようにすることで、論理ドライブ４から情報処理装置１１１へのデータ読み出しのバックグラウンドで、記憶部３へのデータ移行が可能になり、後述するバックグラウンドバックアップするべきＬＢＡ領域量を削減し、データ引き継ぎ状態開始から完了までの期間がより短縮化される。特に、論理ドライブ４の読み出し動作において、記憶部２のデータ読み出しと記憶部３へのデータ書き戻しを並列して行うことで、より高速にデータ引き継ぎを行うことができる。
【０１９２】
（バックグラウンドバックアップ）
論理ドライブ４が論理ドライブステータステーブル４５０においてデータ引き継ぎ状態である場合、アプリケーションプログラム４００やＯＳ１００による論理ドライブ４へのアクセスがほとんど発生しない時に（アイドル時に）、データ引き継ぎ元の記憶部２からデータ引き継ぎ先の記憶部３にバックグラウンドでバックアップを行う（バックグラウンドバックアップ）。制御プログラム２００は主メモリ６からメタデータ３００を読み出し、記憶部２に紐付けられているファイルＩＤを検索し、記憶部２に紐付けられているファイルが存在すれば、当該ファイルのＬＢＡに対して記憶部２に読み出し命令を送信し、記憶部３の当該ＬＢＡに対して書き込み命令と読み出されたデータを送信して、書き込みを行い、主メモリ６上のメタデータ３００を書き換えて当該ファイルＩＤを記憶部３に紐付ける。
【０１９３】
（データ引き継ぎ完了時）
図４６は、データ引き継ぎ完了時の制御プログラムの動作手順を示すものである。論理ドライブステータステーブル４５０において、論理ドライブ４のステータスが「データ引き継ぎ状態」の時（ステップＳ４２０）、制御プログラム２００は、主メモリ６上のメタ

データ３００を定期的に読み出し（ステップＳ４２１）、記憶部２に紐付けられている引き継ぎ対象ファイルＩＤが存在するか否かを定期的にチェックする（ステップＳ４２２）。たとえば、制御プログラム２００は、論理ドライブ４に格納されている全ファイルのファイルＩＤのうち、記憶部２に紐付けられている引き継ぎ対象ファイルＩＤが存在するか否かを定期的にチェックする。存在する場合には、まだデータ引き継ぎが完了していないため、データ引き継ぎ状態のステータスを続行する。

【０１９４】

一方、存在しない場合には、制御プログラム２００はデータ引き継ぎ先である記憶部３のディスクステータス領域３００６をディスクステータス＝１（通常状態）に書き換え（ステップＳ４２３）、データ引き継ぎ元である記憶部２の領域２００６をディスクステータス＝５（廃棄対象状態）に書き換える（ステップＳ４２４）。制御プログラム２００は、記憶部２を論理ドライブ４から切り離し、記憶部３を論理ドライブ４として認識（マウント）し（ステップＳ４２５）、論理ドライブステータステーブル４５０で論理ドライブ４のステータスを「データ引き継ぎ状態」から「通常状態」に書き換える（ステップＳ４２６）。

【０１９５】

これにより、記憶部２はいつでも物理的（機械的）に取り外し可能な状態で破棄可能な状態になり、かつ記憶部３が寿命到達時処理Ｓ２０５以前の記憶部２の役割を担うことになる。以後、記憶部３を記憶部２とみなしてよく、情報処理装置１１１のデータ構成は寿命到達時処理Ｓ２０５以前のデータ構成である図３１の状態に戻る。

【０１９６】

なお、記憶部２を情報処理装置１１１から安全に取り外すためには、記憶部２を論理ドライブ４から切り離し後に、Information technology ATA/ATAPI Command Set-3 (ACS-3)に記載されているE0h Standby Immediateコマンド等を発行したり、NVM Express Revision1.1に記載されているShutdown Notification (CC.SHN)を01bにセットしたりするなどして、あらかじめ記憶部２を電源遮断可能な状態に遷移させておくことが望ましい。

【０１９７】

また、廃棄可能な状態の記憶部２の消費電力を低減するためには、記憶部２を論理ドライブ４から切り離し後に、Information technology ATA/ATAPI Command Set-3 (ACS-3)に記載されているE0h Standby ImmediateコマンドやE6h SLEEPコマンド等を発行したり、記憶部２への供給電源を遮断したり、記憶部２をSerial ATA Revision 3.1 Gold Revisionに記載のPartial状態やSlumber状態に遷移させたり、”Serial ATA Technical Proposal: SATA31_TPR_C108 Title: Device Sleep”に記載のDEVSLP信号をアクティブにして記憶部２をDevSleep状態に遷移させたり、記憶部２をPCI Express Base Specification Revision 3.0に記載のD1状態やD2状態やD3状態に遷移させたり、PCI Express Base Specification Revision 3.0に記載のL1状態やL2状態やL3状態に遷移させたりしてもよい。

【０１９８】

図４７は、記憶部２がディスクステータス＝３（保護状態）で、記憶部３がディスクステータス＝４（データ引き継ぎ先状態）である状態における情報処理装置１１１による論理ドライブ４からの読み出し状態を示すものである。記憶部２，３で、ＬＢＡは重複していない。この状態の時には、指定されるＬＢＡに応じて、記憶部２，３の少なくとも一方からデータが読み出される。

【０１９９】

図４８は、記憶部２がディスクステータス＝３（保護状態）で、記憶部３がディスクステータス＝４（データ引き継ぎ先状態）である状態における情報処理装置１１１による論理ドライブ４に対する書き込み状態を示すものである。この状態の時には、書き込みは、記憶部３に対してのみ行われる。すなわち、記憶部２は、リードオンリーデバイスとし機能する。記憶部３に書き込みが行われたＬＢＡに関しては、記憶部２から割り当てが解除される。

【０２００】

以上のようにして、本実施形態により、記憶部２が寿命到達した場合または寿命到達直前になった場合、記憶部２の書き換えは、ディスクステータス領域２００６の書き換えしか発生しないため、書き込み処理がほとんど行われなくなり、あたかもリードオンリーデバイスであるかのように取り扱われる。一方、論理ドライブは読み出しと書き込み両方可能なドライブとして振る舞うため、アプリケーションプログラム４００にとっての論理ドライブ４は寿命到達以前と同様の振る舞いになる。記憶部２から記憶部３へのデータのデータ引き継ぎは、アプリケーションプログラム４００やＯＳ１００から論理ドライブ４への書き込みが要求された場合に発生し、記憶部２から記憶部３へのデータのコピーではなく、アプリケーションプログラム４００やＯＳ１００から記憶部３へのデータ書き込み処理とメタデータ書き換えによる論理的なデータ移行という形で行われる。それにより、記憶部２から記憶部３へのデータのデータ引き継ぎは、アプリケーションプログラム４００やＯＳ１００から記憶部２への通常のデータ書き込みのバックグラウンドで実行可能であり、記憶部２からデータを読み出し記憶部３に当該データを書き込む処理をアプリケーションプログラム４００やＯＳ１００からの論理ドライブ４へのアクセスとは独立して行う比較例のバックアップ処理に比べて飛躍的に高速に行われる。つまり、アプリケーションプログラム４００やＯＳ１００からの書き込みが発生するＬＢＡに対しては実質的にデータ引き継ぎ時間がゼロとなる。

【０２０１】

アプリケーションプログラム４００やＯＳ１００からの書き込みが発生しないようなＬＢＡに対しては、別途バックアップ処理が必要となるが、新しい記憶装置のマウント前にデータをコピーしなければならない比較例のバックアップ処理やＲＡＩＤアレイの再構築と異なり、記憶部２と記憶部３をマウントした後に、アイドル時にバックグラウンドで行うことができるため、アプリケーションプログラム４００の性能劣化を抑制することができる。新しい記憶装置のマウント前にユーザデータのコピーが必要となる比較例のバックアップ処理や、新しい記憶装置のマウント前にユーザデータおよびパリティデータの再構築が必要となるＲＡＩＤなどによる論理ドライブの再構築と異なり、本実施形態によるデータ引き継ぎ先記憶装置接続に伴う論理ドライブ再構築は、図４０に示したように、ディスクステータス領域およびディスク識別ＩＤ領域の書き換えとメタデータ領域のコピーのみ必要となるため、非常に高速に行うことが可能になる。

【０２０２】

（第２の実施形態）

第１の実施形態では、記憶部３に引き継がれたデータを検索するための情報として、主メモリ６に格納されるメタデータ３００が使用される例について説明した。これにより、たとえばＯＳ１００がアプリケーションプログラム４００からファイルＩＤの指定をうけて論理ドライブ４のデータ読み出しを命令されたとき、メタデータ３００を読み出すことで、記憶部２と記憶部３のどちらからデータを読み出すべきかの情報と、どのＬＢＡからデータを読み出すべきかの情報を取得することができる。第２の実施形態では、記憶部３に引き継がれたデータを検索するための情報として、記憶部３に格納される引き継ぎ履歴５５０が使用される例について説明する。たとえばＯＳ１００がアプリケーションプログラム４００から論理ドライブ４のデータ読み出しを命令されたとき、引き継ぎ履歴５５０を読み出すことで、記憶部２と記憶部３のどちらからデータを読み出すべきかの情報を取得することができる。なお、本実施形態においては、アプリケーションプログラム４００はＬＢＡを直接指定してＯＳ１００に読み出し命令、書き込み命令を送信する場合について説明する。アプリケーションプログラム４００が第１の実施形態のようにファイルＩＤを指定してＯＳ１００に読み出し命令・書き込み命令を送信する場合についても本実施形態の発明は適用可能であり、その場合、制御プログラム２００またはＯＳ１００がメタデータ３００を読みだすことでファイルＩＤをＬＢＡに変換することができ、変換したＬＢＡについての読み出し、書き込み処理は本実施形態と同様に行われる。

【０２０３】

図４９に第２の実施形態の情報処理システム１の構成を示す。情報処理システム１の基

本構成は第１の実施形態１と同様である。論理ドライブは、ＯＳ１００が認識しうる論理的なドライブであり、論理ドライブＩＤ（ドライブ名やボリューム番号や論理ユニット番号など）が割り当てられ、ＯＳ１００は１乃至複数の物理デバイスとしての記憶部を論理ドライブとして認識する。論理ドライブは、論理セクタ（論理ブロック）に分割され、各論理セクタには、ＬＢＡが割り当てられる。論理ドライブは、図７におけるＯＳ１００および制御プログラム２００によって、下位の階層である１乃至複数台の物理ドライブやあるいはそれらの一部のＬＢＡ領域に割り当てられ、ＯＳ１００は論理ドライブのＬＢＡと物理ドライブのＬＢＡを相互変換する。上位の階層であるアプリケーションプログラム４００は論理ドライブを仮想的に１台のドライブとして認識する。本実施形態においては、寿命到達時処理Ｓ２０５以前の状態において、論理ドライブ４が１つの物理記憶装置である記憶部２に割り当てられる場合について述べる。この場合、論理ドライブのＬＢＡと物理ドライブのＬＢＡは同じ値となる。なお、寿命到達時処理Ｓ２０５以前の状態であっても、複数の物理記憶装置を用いてRedundant Arrays of Inexpensive Disks (ＲＡＩＤ)のディスクアレイ、たとえばＲＡＩＤ０やＲＡＩＤ５のディスクアレイを構成し、これを一つの論理ドライブ４として認識してもよく、その場合であっても本実施形態は適用可能である。アプリケーションプログラム４００は、論理ドライブＩＤ及びＬＢＡからなる論理アドレスを含むコマンドをＯＳ１００に与えることにより、特定の論理ドライブにおける特定の論理セクタに対してアクセスすることができる。なお、論理ドライブＩＤは記憶部の全ＬＢＡ領域ではなく一部のＬＢＡ領域に割り当てられても良いし、そうすることで、記憶部２および記憶部３は複数の論理ドライブに分割して管理されそれぞれの論理ドライブに別々の論理ドライブＩＤが割り当てられてもよい。

【０２０４】

本実施形態は、例として、記憶部２として第１の実施形態に記載の記憶部２であるＳＳＤを用い、記憶部３として第１の実施形態に記載の記憶部３であるＳＳＤを用いる場合について説明する。信頼性劣化後に廃棄して設置スペースを削減しかつシステム１全体の消費電力を削減するという観点では、記憶部２は情報処理装置１１１に対して物理的に着脱可能であることが望ましい。

【０２０５】

データ引き継ぎ先である記憶部３の記憶可能容量は、データ引き継ぎ元である記憶部２の記憶可能容量と同じか、記憶部２の記憶可能容量よりも大きいことが望ましいが、記憶部２の記憶可能容量よりも小さい場合であっても本発明は適用可能である。

【０２０６】

本実施形態では、記憶部３は記憶部２が寿命到達または寿命到達目前であると判定された後に、新たに情報処理装置１１１に接続される記憶部である。記憶部２が寿命到達または寿命到達目前であると判定された後に、すでに情報処理装置１１１に接続されている通常状態の記憶部３を新たな接続無しに引き継ぎ先として利用する場合についても本発明は適用可能である。記憶部３の接続前の設置スペースを削減しかつシステム１全体の消費電力を削減するという観点、および記憶部３の信頼性劣化後に記憶部３を廃棄して設置スペースを削減しかつシステム１全体の消費電力を削減するという観点では、記憶部３は情報処理装置１１１に対して物理的に着脱可能であることが望ましい。

【０２０７】

主メモリ６に格納される制御プログラム２００は、記憶部２や記憶部３の統計情報、それぞれのステータス記憶領域５１０、論理ドライブＩＤ記憶領域５２０、および引き継ぎ履歴記憶領域５５０の制御および管理を行い、統計情報に基づく寿命到達時処理やデータ引き継ぎ処理等を行う。

【０２０８】

記憶部２および記憶部３はそれぞれステータス格納領域５１０と論理ドライブＩＤ格納領域５２０を具備し、記憶部３は引き継ぎ履歴格納領域５５０を具備する。

【０２０９】

本実施形態では、ステータス記憶領域５１０に格納されるデータは、０乃至５の値を取

りうる。それぞれの値は、対応する記憶部がそれぞれ
０：初期ディスク状態
１：通常状態
２：信頼性劣化状態
３：データ引き継ぎ元状態（保護状態）
４：データ引き継ぎ先状態
５：廃棄対象状態
の状態であることを示す。なお、ステータス記憶領域５１０や論理ドライブＩＤ記憶領域５２０や引き継ぎ履歴記憶領域５５０は各記憶部内に格納されるのではなく、主メモリ６に格納されるようにしてもよい。データ引き継ぎ元の記憶部２とデータ引き継ぎ先の記憶部３の論理ドライブＩＤ記憶領域５２０には、同じ論理ドライブＩＤが格納される。

【０２１０】

引き継ぎ履歴格納領域５５０は、データ引き継ぎ先である記憶部３に具備され、データ引き継ぎ状態の記憶部３のＬＢＡに対して書き込みが行われたり、情報処理装置１から論理ドライブ４に削除通知が発行されて記憶部２のＬＢＡのデータが無効化されたりすると、その更新履歴（データ引き継ぎ履歴）として、図５０に示すように、書き込みＬＢＡと書き込みセクタサイズを格納する。引き継ぎ履歴格納領域５５０は書き込み時や削除通知時や記憶部３のアイドル時に随時整理・最適化される。たとえば、ＬＢＡ＝ｃＬＢＡ～ｃＬＢＡ＋Ｘ－１のセクタサイズＸの領域（以後(ｃＬＢＡ,Ｘ)と表記する）に対し、オーバーラップするＬＢＡや連続するＬＢＡが引き継ぎ履歴５５０にすでに記録されている場合には、それらを合併（統合）したログを引き継ぎ履歴５５０に記録し、引き継ぎ履歴５５０から合併されたログを削除することが望ましい。たとえば、Ａ＝(ｃＬＢＡ,Ｘ)のＬＢＡ領域Ａを新たに引き継ぎ履歴格納領域５５０に記入する場合であって、Ｂ＝(ｃＬＢＡ－ａ,ａ)のＬＢＡ領域Ｂがすでに引き継ぎ履歴５５０に格納されている時、ＬＢＡ領域ＡとＬＢＡ領域Ｂは連続しているＬＢＡ領域であるため、Ａ＋ＢのＬＢＡ領域である(ｃＬＢＡ－ａ,ａ＋Ｘ)のデータを引き継ぎ履歴５５０におけるＢ＝(ｃＬＢＡ－ａ,ａ)の履歴格納部に上書きすることで、ログデータ量を増大させることなく引き継ぎ履歴５５０を更新することができる。引き継ぎ履歴５５０は後述するデータ引き継ぎが完了すると消去されることにより、ユーザデータ格納など他の用途に使用されてもよい。

【０２１１】

図５１は引き継ぎ履歴５５０に対する履歴の書き込み例である。寿命到達前の論理ドライブ４のＬＢＡ＝ｃＬＢＡのデータＤ(ｃＬＢＡ)は記憶部２のＬＢＡ＝ｃＬＢＡに格納されている。記憶部２が寿命到達すると、たとえば論理ドライブ４のＬＢＡ＝ｃＬＢＡへの１セクタサイズのデータＤ(ｃＬＢＡ)の書き込みは記憶部３のＬＢＡ＝ｃＬＢＡに対して行われ、引き継ぎ履歴５５０にＬＢＡ＝ｃＬＢＡとセクタカウント＝１が格納される。寿命到達後の論理ドライブ４のＬＢＡ＝ｃＬＢＡへのセクタサイズ＝ＸのデータＤ(ｃＬＢＡ),Ｄ(ｃＬＢＡ＋１),...,Ｄ(ｃＬＢＡ＋Ｘ－１)の書き込みは、記憶部３のＬＢＡ＝ｃＬＢＡ,cＬＢＡ＋１,...,ｃＬＢＡ＋Ｘ－１に対して行われ、引き継ぎ履歴５５０にＬＢＡ＝ｃＬＢＡとセクタカウント＝Ｘが格納される。

【０２１２】

図５１では、ＬＢＡ＝１、ＬＢＡ＝３、ＬＢＡ＝４に対する書き込みは記憶部３のＬＢＡ＝１、ＬＢＡ＝３、ＬＢＡ＝４に行われ、引き継ぎ履歴として、ＬＢＡ＝１およびセクタカウント＝１の履歴が記録され、ＬＢＡ＝３およびセクタカウント＝１と、ＬＢＡ＝４およびセクタカウント＝１の履歴はＬＢＡ＝３およびセクタカウント＝２の履歴に統合されて記録される。

【０２１３】

ステータス格納領域５１０と論理ドライブＩＤ格納領域５２０と引き継ぎ履歴格納領域５５０にはＬＢＡの領域が割り当てられてもよい。あるいは、ステータス格納領域５１０と論理ドライブＩＤ格納領域５２０と引き継ぎ履歴格納領域５５０にはＬＢＡ領域ではなくログページアドレスが割り当てられたログ領域が割り当てられてもよく、この場合には

、たとえば、非特許文献１のACS-3に記載されている2Fh Read Log Extにより該ログページ領域の読み出しが行われ、非特許文献１のACS-3に記載されている3Fh Write Log Extにより該領域への書き込みが行われる。

【０２１４】

制御プログラム２００は、第１の実施形態の図２４や図２９と同様にして、ＣＰＵ５に接続されたそれぞれの記憶部が寿命に到達しているか否か、寿命到達目前であるか否か、あるいは故障目前であるか否かを判定し、寿命に到達しているか場合、寿命到達目前である場合、あるいは故障目前である場合に、該記憶部の寿命到達時処理を行う。寿命到達判定は、第１の実施形態と同様に、図２４のような一定時間ごと、または一定処理数おき、または一定データ送受信おき、または図２９や図３０のように記憶部からのコマンド応答がエラーであった場合に実施される。

【０２１５】

（寿命到達処理）

図５２は、たとえば、記憶部２が寿命に到達していると判定された場合の、本実施例における記憶部２の寿命到達時処理を示す。記憶部２が寿命に到達していると判定された場合（ステップＳ４３０）、制御プログラム２００は記憶部２のステータスを１の通常状態から２の信頼性劣化状態に書き換える（ステップＳ４３１）。制御プログラム２００はインタフェース１９の空きポートに新しい記憶部である記憶部３を接続するようディスプレイ装置やポート近傍に設置されたＬＥＤ等を経由してユーザや管理者に通知することが望ましい（ステップＳ４３２）。あるいは、ディスクロード・アンロード装置（図示せず）として記憶部２や記憶部３のインタフェース１９に対する物理的な着脱を自動的に行う機械装置がシステム１に搭載されている場合、新しい記憶部である記憶部３をインタフェース１９に接続するよう該ディスクロード・アンロード装置に対して命令してもよい。

【０２１６】

新たな記憶部として記憶部３が接続された時（ステップＳ４３３）、制御プログラム２００は記憶部３のステータス５１０を４のデータ引き継ぎ先状態に書き換え（ステップＳ４３４）、記憶部３の論理ドライブＩＤ格納領域５２０に記憶部２の論理ドライブＩＤ格納領域５２０の値をコピーすることで、記憶部２、３の論理ドライブＩＤを一致させる（ステップＳ４３５）。今回の例では、記憶部２は図４９に示すように論理ドライブ４として割り当てられているため、記憶部３の論理ドライブＩＤ格納領域５２０には論理ドライブ４のＩＤが書き込まれる。制御プログラム２００は記憶部２のステータス５１０を３のデータ保護状態（データ引き継ぎ元状態）に書き換え（ステップＳ４３６）、ＯＳ１００に記憶部２と記憶部３を同一の論理ドライブである論理ドライブ４として認識させる（ステップＳ４３７）。寿命到達処理後、記憶部２のステータスは３に、記憶部３のステータスは４になっており、論理ドライブ４および記憶部２および記憶部３はデータ引き継ぎ状態に遷移していることになる。

【０２１７】

本実施形態では、各記憶部のステータスは、各記憶部のステータス格納領域５１０に不揮発に格納されており、図３５に示すように、ＯＳ１００の起動ごとに制御プログラム２００はステータス格納領域５１０を読み出すことで各記憶部のステータスを認識する。制御プログラム２００は、各記憶部のステータスを認識し、かつ各記憶部の論理ドライブＩＤを論理ドライブＩＤ格納領域５２０から読み出すことで、論理ドライブ４がデータ引き継ぎ状態か否かを認識する。

【０２１８】

（論理ドライブからの読み出し）

制御プログラム２００は、アプリケーションプログラム４００からの読み出し命令に対し、図５３に示すようにしてデータを読み出す。制御プログラム２００は、アプリケーションプログラム４００から読み出し命令と、読み出し対象論理ドライブＩＤと、読み出しＬＢＡと、セクタカウントを受信する（ステップＳ４４０）。制御プログラム２００は、論理ドライブＩＤ格納部５２０の値が、読み出し対象論理ドライブＩＤに等しい記憶部を

全て検索して、記憶部２や記憶部３を特定する（ステップＳ４４１）。制御プログラム２００は、検索した記憶部のステータス格納部５１０の値を読み出し、ステータスを判定することで、検索した記憶部それぞれが記憶部２と記憶部３のいずれであるかを特定する（ステップＳ４４２）。ステータス格納部５１０の読み出しにともなう性能劣化を抑制するためには、記憶部２や記憶部３のステータス格納部５１０の各データは情報処理装置１１１起動時に主メモリ６に読み出され、以後ステータス格納部５１０のデータは主メモリ６から読み出されるようにすることが望ましい。

【０２１９】

対象論理ドライブに属する記憶部のステータスが１の場合、論理ドライブのステータスは通常状態であるということになる（ステップＳ４４３：Ｙｅｓ）。制御プログラム２００は、記憶部２に、読み出し命令と、読み出しＬＢＡと、セクタカウントを送信する（ステップＳ４４４）。制御プログラム２００は、記憶部２から応答および読み出しデータを受信する（ステップＳ４４５）。制御プログラム２００は、アプリケーションプログラム４００に読み出しデータおよび応答を送信する（ステップＳ４４６）。

【０２２０】

対象論理ドライブに属する記憶部のステータスが１ではない場合、論理ドライブのステータスはデータ引き継ぎ状態ということになる（ステップＳ４４３：Ｎｏ）。制御プログラム２００は、記憶部３の引き継ぎ履歴５５０を読み出し（ステップＳ４４７）、読み出しＬＢＡが引き継ぎ履歴に含まれるか否かを判定する（ステップＳ４４８）。制御プログラム２００は、読み出しＬＢＡが引き継ぎ履歴に含まれる場合は（ステップＳ４４８：Ｙｅｓ）、記憶部３に読み出し命令と、読み出しＬＢＡと、セクタカウントを送信する（ステップＳ４５２）。制御プログラム２００は、記憶部３から応答および読み出しデータを受信する（ステップＳ４５３）。制御プログラム２００は、アプリケーションプログラム４００に読み出しデータおよび応答を送信する（ステップＳ４５５）。

【０２２１】

制御プログラム２００は、読み出しＬＢＡに引き継ぎ履歴が含まれない場合は（ステップＳ４４８：Ｎｏ）、記憶部２に読み出し命令と、読み出しＬＢＡと、セクタカウントを送信する（ステップＳ４４９）。制御プログラム２００は、記憶部２から応答および読み出しデータを受信する（ステップＳ４５０）。制御プログラム２００は、アプリケーションプログラム４００に読み出しデータおよび応答を送信する（ステップＳ４５５）。

【０２２２】

（書き戻しバックアップ）

なお、例えば図５３において、ステップＳ４５１の書き戻しバックアップを行うようにしてもよいし、行わないようにしてもよい。ステップＳ４５１では、データ引き継ぎ中のデータ引き継ぎ元である記憶部２のデータを主メモリ６内のキャッシュメモリ領域に読みだした時には、キャッシュメモリ領域に読みだしたデータをデータ引き継ぎ中のデータ引き継ぎ先である記憶部３に書き込み、書き込み先ＬＢＡとセクタサイズを引き継ぎ履歴５５０に書きこむ。そうすることで、論理ドライブ４から情報処理装置１１１へのデータ読み出しのバックグラウンドで記憶部３へのデータ移行が可能になり、バックグラウンドバックアップするべきＬＢＡ領域量を削減し、データ引き継ぎ状態開始から完了までの期間がより短縮化される。特に、論理ドライブ４の読み出し動作において、記憶部２のデータ読み出しと記憶部３へのデータ書き戻しを並列して行うことで、より高速にデータ引き継ぎを行うことができる。

【０２２３】

（論理ドライブへの削除命令）

図５４はアプリケーションプログラム４００からＯＳ１００にＬＢＡデータ削除要求が送信された時の、ＯＳ１００の処理手順を示すものである。ＯＳ１００は、アプリケーションプログラム４００から削除命令と、論理ドライブアドレスと、削除対象ＬＢＡを受信する（ステップＳ９２０）。制御プログラム２００は、論理ドライブＩＤ格納部５２０の値が、ＬＢＡデータ削除対象の論理ドライブＩＤに等しい記憶部を全て検索して、記憶部

２や記憶部３を特定する（ステップＳ９２１）。制御プログラム２００は、検索した記憶部のステータス格納部５１０の値を読み出し（ステップＳ９２２）、ステータスを判定することで、検索した記憶部それぞれが記憶部２と記憶部３のいずれであるかを特定する。

【０２２４】

対象論理ドライブに属する記憶部のステータスが１の場合（ステップＳ９２３：Ｙｅｓ）、論理ドライブのステータスは通常状態であるということになる。ＯＳ１００は、記憶部２に対し、削除通知およびＬＢＡを送信する（ステップＳ９２４）。ＯＳ１００は、記憶部２から応答を受信する。ＯＳ１００は、アプリケーションプログラム４００に応答を送信する（ステップＳ９３２）。

【０２２５】

対象論理ドライブに属する記憶部のステータスが１ではない場合（ステップＳ９２３：Ｎｏ）、論理ドライブのステータスはデータ引き継ぎ状態ということになる。制御プログラム２００は、記憶部３の引き継ぎ履歴５５０を読み出し（ステップＳ９２５）、データ削除対象ＬＢＡが引き継ぎ履歴に含まれるか否かを判定する（ステップＳ９２６）。削除ＬＢＡが引き継ぎ履歴に含まれる場合は、削除対象のデータは記憶部３に格納されていることになる。制御プログラム２００は、記憶部３に対し、削除通知およびＬＢＡを送信し（ステップＳ９２７）、記憶部３は削除通知対象ＬＢＡのデータを無効化し、記憶部３から応答を受信し、アプリケーションプログラム４００に応答を送信する（ステップＳ９３０）。

【０２２６】

読み出しＬＢＡに引き継ぎ履歴が含まれない場合は（ステップＳ９２６）、削除対象のデータは記憶部２に格納されていることになる。制御プログラム２００は記憶部２に対し削除通知およびＬＢＡを送信し（ステップＳ９２８）、記憶部２から応答を受信する。記憶部２に対する削除通知は送信しなくても良い。アプリケーションプログラム４００からの削除命令の対象となるＬＢＡはアプリケーションプログラム４００やＯＳ１００にとって今後読み出す必要の無いデータであり、記憶部３に引き継ぐ必要のないデータであるため、制御プログラム２００は、引き継ぎ履歴５５０に削除対象ＬＢＡとセクタカウントを記録することにより、削除対象ＬＢＡと記憶部２の紐付けを無効にする（ステップＳ９２９）。制御プログラム２００はアプリケーションプログラム４００に応答を送信する（ステップＳ９３０）。

【０２２７】

このようにして、データ削除命令によって、引き継ぎ履歴５５０の更新により、記憶部２から記憶部３にデータが論理的に引き継がれていくことになり、データ削除がデータ引き継ぎ動作を兼用することになる。

【０２２８】

（論理ドライブへの書き込み）

制御プログラム２００は、アプリケーションプログラム４００からの書き込み命令に対し、図５５に示すようにしてデータを書き込む。制御プログラム２００は、アプリケーションプログラム４００から書き込み命令と、書き込み対象論理ドライブＩＤと、書き込みＬＢＡと、セクタカウントを受信する（ステップＳ４６０）。制御プログラム２００は、論理ドライブＩＤ格納部５２０の値が、書き込み対象論理ドライブＩＤに等しい記憶部を全て検索して、記憶部２や記憶部３を特定する（ステップＳ４６１）。制御プログラム２００は、検索した記憶部のステータス格納部５１０の値を読み出し、ステータスを判定することで、検索した記憶部それぞれが記憶部２と記憶部３のいずれであるかを特定する（ステップＳ４６２）。

【０２２９】

対象論理ドライブに属する記憶部のステータスが通常状態である場合、論理ドライブのステータスは通常状態であるということになる（ステップＳ４６３：Ｙｅｓ）。制御プログラム２００は、記憶部２に、書き込み命令と、書き込みＬＢＡと、セクタカウントを送信する（ステップＳ４６４）。また、制御プログラム２００は、アプリケーションから受

信した書き込みデータを記憶部２に送信する（ステップＳ４６５）。
【０２３０】
対象論理ドライブに属する記憶部のステータスがデータ引き継ぎ状態である場合、論理ドライブのステータスはデータ引き継ぎ状態ということになる（ステップＳ４４３：Ｎｏ）。制御プログラム２００は、データ引き継ぎ先の記憶部３に書き込み命令と、書き込みＬＢＡと、セクタカウントを送信する（ステップＳ４６６）。また、制御プログラム２００は、アプリケーションから受信した書き込みデータを記憶部３に送信する（ステップＳ４６７）。制御プログラム２００は、記憶部３の引き継ぎ履歴５５０を読み出し（ステップＳ４６８）、書き込みＬＢＡが引き継ぎ履歴に含まれるか否かを判定する（ステップＳ４６９）。書き込みＬＢＡが引き継ぎ履歴に含まれる場合は、書き込みＬＢＡはすでに引き継ぎ済みということになるので、制御プログラム２００は引き継ぎ履歴５５０を特に更新しない。書き込みＬＢＡが引き継ぎ履歴に含まれない場合は、書き込みＬＢＡは新たに引き継ぎ完了したＬＢＡということになるので、制御プログラム２００は引き継ぎ履歴５５０に書き込みＬＢＡとセクタカウントを記録する（ステップＳ４７０）。記憶部３では、書き込みＬＢＡに対し、書き込みデータを書き込む。
【０２３１】
このようにして、データ引き継ぎ中には、データ引き継ぎ元の記憶部２に対して書き込み命令を送信せず、データ引き継ぎ先の記憶部３に対して書き込み命令を送信し、記憶部３の引き継ぎ履歴５５０に引き継ぎ履歴を記録する。論理ドライブ４において、アプリケーションプログラム４００から書き込み要求を受信するごとに、記憶部２に格納される有効なデータが記憶部３に徐々に移行されていくことになり、新規データ書き込みがデータ引き継ぎ動作を兼用することになる。
【０２３２】
アプリケーションプログラム４００からのデータ書き込み命令が記憶部２の全ＬＢＡに対して均等な確率分布で送信される場合、十分な量のデータが書き込まれると、記憶部２のほぼ全ての有効データが記憶部３に移行され、記憶部２内の有効データはほとんどなくなる。たとえば記憶部２の記憶容量である全論理セクタ数がＣ２で、記憶部３の全論理セクタ数がＣ３とし、Ｃ２＝Ｃ３＝Ｃである場合を考える。モデルケースとして、全ＬＢＡに対する書き込み分布が均一な確率分布である場合を考えると、ある書き込みコマンドであるＬＢＡ＝ｃＬＢＡが書き込まれる確率は１／Ｃである。ｎ回の書き込み命令が処理された場合、ＬＢＡ＝ｃＬＢＡが一度も書き込まれない確率は(１－（１/Ｃ）)^ｎとなる。^ｎはｎ乗である。よって、ｎ回の書き込み命令が処理されたあとに書き込みが完了した論理セクタ数の期待値はＣ－Ｃ×(１－（１/Ｃ）)^ｎとなる。
【０２３３】
１回の書き込み命令で１論理セクタの書き込みが行われる場合には、記憶部２と記憶部３の記憶容量のＮ倍の容量のデータが書き込まれたとき、処理された書き込みコマンド数はｎ＝ＮＣであるため、書き込みが行われていない論理セクタ数の期待値ＥはＥ＝Ｃ×（１－（１/Ｃ）^(ＮＣ)となる。たとえばGbyte単位での記憶部２の記憶容量ＧがIDEMA (International Disk Drive Equipment and Materials Association)規格に基づきＧ＝５１２GB (＝４７６.９GiByte)であるとき、Ｃ＝９７,６９６,３６８＋１,９５３,５０４ ×(Ｇ＝５１２-５０)＝ １，０００，２１５,２１６であり、一般にＣは十分に大きい整数であるため、Ｅ＝Ｃ×ｅ^(－Ｎ)と近似可能である（ｅは自然対数の底）。よって、期待値ＥはＮの増加に対して指数関数的に減少する。たとえば、Ｇ＝５１２GByte (＝４７６.９GiByte)の容量の論理ドライブ４に対して、１周分である４７６.９GiByteのデータ書き込みが発生した場合、論理ドライブ４の約６３.２％のＬＢＡが書き込み完了したことになり、論理ドライブ４の半数以上のデータが記憶部２から記憶部３に移行完了したといえる。たとえば、Ｇ＝５１２GByte (＝４７６.９GiByte)の容量の論理ドライブ４に対して、４.６周分である１３TiByteのデータ書き込みが発生した場合、論理ドライブ４の約９９％のＬＢＡが書き込み完了したことになり、論理ドライブ４のほぼ全てのデータが記憶部２から記憶部３に移行完了したといえる。Ki＝１０２４,Mi ＝１０２４×Ki,Gi＝１０２

４×Mi,Ti＝１０２４×Giである。

【０２３４】

（データ引き継ぎ状況の監視）

制御プログラム２００は引き継ぎ履歴５５０を読み出すことでデータ引き継ぎ状態の論理ドライブ４のデータ引き継ぎ状況を監視する。図５６は、引き継ぎ履歴を用いたデータ引き継ぎ状況の監視手順を示すものである。たとえば、制御プログラム２００は所定時間経過毎に引き継ぎ履歴５５０を読み出すことでデータ引き継ぎ状況を監視する（ステップＳ４８０、Ｓ４８１）。制御プログラム２００は、全ての引き継ぎ対象ＬＢＡが引き継ぎ履歴５５０に含まれる場合、データ引き継ぎが完了したと判定する。たとえば、データ引き継ぎ元である記憶部２の全ＬＢＡが引き継ぎ履歴５５０に含まれる場合、データ引き継ぎが完了したと判定される（ステップＳ４８２）。あるいは、データ引き継ぎ完了の判定として、例えば、データ引き継ぎ先である記憶部３の全ＬＢＡが引き継ぎ履歴５５０に含まれるか否かを判定するようにしてもよい。

【０２３５】

データ引き継ぎが完了したと判定された場合、制御プログラム２００はデータ引き継ぎ元である記憶部２のステータスを５の廃棄対象状態に変更し、データ引き継ぎ先である記憶部３のステータスを１の通常状態に変更することで（ステップＳ４８３）、論理ドライブ４のデータ引き継ぎ状態を終了し、論理ドライブ４のデータ引き継ぎ状況監視を終了する。記憶部２が消費する電力を低減するという観点では、制御プログラム２００は記憶部２に低消費電力モードへの遷移要求を送信することが望ましい。データ引き継ぎ状態終了後、制御プログラム２００はディスプレイ９やポート近傍に設置されたＬＥＤを通じて、記憶部２をインタフェース１９から取り外すようユーザや管理者に通知したり、記憶部２をインタフェース１９から取り外すよう前記ディスクロード・アンロード装置に命令を送信することが望ましい。

【０２３６】

（全体のステータスの遷移）

図５７は、記憶部２が寿命到達する際の論理ドライブ４のステータスの遷移を示したものである。記憶部２が通常状態である場合には記憶部２のステータスは１になっており、記憶部３はＣＰＵ５に未接続状態である（ステップ１）。制御プログラム２００は記憶部２が寿命到達したと判定すると、制御プログラム２００は記憶部２のステータスを２に変更する（ステップ２）。制御プログラム２００からの通知または命令にもとづき、新しい記憶部として記憶部３がインタフェース１９に接続される（ステップ３）。記憶部３が接続されると、制御プログラム２００はデータ引き継ぎ元として記憶部２のステータスを３に変更し、データ引き継ぎ先として記憶部３のステータスを４に変更することで、論理ドライブ４はデータ引き継ぎ状態に移行完了する（ステップ４）。引き継ぎ履歴５５０の情報をもとに、記憶部２のすべての有効データが記憶部３にデータ引き継ぎされたと判定されると、制御プログラム２００は、記憶部２のステータスを５に変更し、記憶部３のステータスを１に変更する。以後、記憶部３が、元の記憶部２であるかのように振る舞う（ステップ１に戻る）。その後さらに記憶部３＝新記憶部が寿命到達したと判定された場合、同様のステップ２～５が繰り返される。このようにして、情報処理装置１１１の記憶部のいずれか一つが寿命到達した場合、寿命到達目前である場合、故障寸前である場合となっても、本実施形態により記憶部２のデータを新しい記憶部に容易にデータ引き継ぎすることが可能になる。

【０２３７】

（バックグラウンドバックアップ）

制御プログラム２００は、たとえば、アプリケーションプログラム４００から命令を一定時間以上受信しない場合などのアイドル時、またはスタンバイモード遷移要求をアプリケーションプログラム４００から受信した時、または情報処理装置１１１やＯＳ１００のシャットダウン時などに、記憶部３へのデータ引き継ぎ未完了のＬＢＡに対して記憶部２からデータを自動的に読み出し記憶部３へデータを自動的に書き込むバックアップ動作を

行う、バックグラウンドバックアップを行うことが望ましい。たとえばバックグラウンドバックアップ動作は、制御プログラム２００が、記憶部３の引き継ぎ履歴５５０を読み出し、引き継ぎ履歴５５０に含まれないｃＬＢＡに対して、記憶部２からデータ読み出しを行い、記憶部３に該データを書き込み、引き継ぎ履歴５５０に該ｃＬＢＡと書き込みデータのセクタサイズを格納することで実行される。引き継ぎ履歴５５０へのｃＬＢＡの格納の際、連続したＬＢＡ領域が存在する場合、またはオーバーラップしたＬＢＡ領域が存在する場合には、これらを統合したＬＢＡ領域が引き継ぎ履歴５５０に格納され、統合前のＬＢＡ領域は引き継ぎ履歴５５０から削除されることが望ましい。

【０２３８】

引き継ぎ履歴５５０のデータサイズを低減するという観点、および後述するデータ引き継ぎ終了判定を高速に行うという観点では、引き継ぎ履歴５５０に登録されたｃＬＢＡ領域以外のｃＬＢＡ領域のうち小さく断片化されたｃＬＢＡ領域（断片化した未データ引き継ぎｃＬＢＡ領域）に対してバックグラウンドバックアップが優先的に行われることが望ましい。断片化した未データ引き継ぎｃＬＢＡ領域をバックアップすることにより、断片化未データ引き継ぎ領域の前記断片化未データ引き継ぎ領域の前後の既データ引き継ぎｃＬＢＡ領域と、新たにデータ引き継ぎされたｃＬＢＡ領域が引き継ぎ履歴５５０に統合されて格納される。例えば、領域ＬＢＡ＝０（（ＬＢＡ＝０，セクタサイズ＝１））および領域ＬＢＡ＝２（（ＬＢＡ＝２，セクタサイズ＝１））のデータがバックアップ済みで引き継ぎ履歴に登録されている場合、ＬＢＡ１のデータをバックグラウンドバックアップすることで、連続したＬＢＡ０からＬＢＡ２のＬＢＡ領域（ＬＢＡ＝０，セクタサイズ＝３）が引き継ぎ完了したことになり、引き継ぎ履歴５５０の情報量が削減される。

【０２３９】

たとえば、前述のように、５１２GByte（＝４７６．９GiByte)のＳＳＤに対してアプリケーションプログラム４００から合計４７６．９GiByteのデータが書き込まれた時に、制御プログラム２００が記憶容量の３６．８％の容量である１７５．５GiBの領域に対しバックグラウンドバックアップ動作を行うことで、記憶部２の全ＬＢＡ領域がデータ引き継ぎ完了したことになる。ＳＳＤである記憶部２と、ＳＳＤである記憶部３の典型的な読み出し速度と書き込み速度はたとえば、約４００MiB/秒であり、前述の４７６．９GiBのＬＢＡ領域の記憶部２からの読み出しは約４４９秒で完了し、記憶部３への書き込みは約４４９秒で完了する。よって、このような状況下では、前記バックグラウンドバックアップは高々約１５分で完了し、記憶部２からの読み出しと記憶部３への書き込みが並列に行われる場合には約８分で完了する。

【０２４０】

さらには、たとえば、前述のように、５１２GByte（＝４７６．９GiByte)のＳＳＤに対してアプリケーションプログラム４００から合計１３TiByteのデータが書き込まれた時に、制御プログラム２００が記憶容量の１％の容量である４．８GiBの領域に対しバックグラウンドバックアップ動作を行うことで、記憶部２の全ＬＢＡ領域がデータ引き継ぎ完了したことになる。ＳＳＤである記憶部２と、ＳＳＤである記憶部３の典型的な読み出し速度と書き込み速度はたとえば、約４００MiB/秒であり、前述の４．８GiBのＬＢＡ領域の記憶部２からの読み出しは約１２秒で完了し、記憶部３への書き込みは約１２秒で完了する。よって、このような状況下では、前記バックグラウンドバックアップは高々２４秒で完了し、記憶部２からの読み出しと記憶部３への書き込みが並列に行われる場合には１２秒で完了する。

【０２４１】

一方、本実施形態を適用せずに、５１２GBの容量の記憶部２から全データを読みだして記憶部３に書き込むことむことによるデータのバックアップには、２０分～４１分の時間を要する。すなわち、本実施形態の適用によって、実質的なバックアップに要する時間は、記憶容量×１に相当するデータ書き込みの後には６３％が減少し、記憶容量×４．６に相当するデータ書き込みの後には９９％が減少することになる。

【０２４２】

このように、本実施形態の適用により、ユーザ自身によるデータバックアップ作業は不要となる。また、バックグラウンドバックアップによる情報処理装置１１１への処理の負荷は大幅に低減され、アプリケーションプログラム４００は記憶部２から記憶部３へのデータバックアップの影響をほとんど受けることなく、論理ドライブ４を利用可能である。信頼性劣化した記憶部２のデータ書き換えは、ディスクステータス格納部の書き換えしか発生しないため、記憶部２へのさらなるデータ書き込みによる記憶部２の故障の可能性を低減することができる。論理ドライブ４への新規データ書き込みは記憶部２ではなく信頼性のよい記憶部３に書き込まれるため、書き込みデータの損失を防止することができる。記憶部２が寿命到達してさらなるデータ書き込みが阻止される場合であっても、記憶部の上位レイヤである論理ドライブ４は読み出しと書き込みの両方が可能なドライブとして振る舞うため、アプリケーションプログラムなど上位のソフトウェアレイヤは論理ドライブ４が寿命到達モードであっても通常状態であっても同等の論理ドライブとして取り扱うことができるため、本実施形態導入のためのアプリケーションプログラム修正を必要とせず、本実施形態採用のシステムへの移行が容易である。

【０２４３】

（第３の実施形態）

第３の実施形態では、本発明をディスクアレイからなる情報処理システム１に適用した場合について説明する。図５８は、第３の実施形態の情報処理システム１を示すものである。情報処理システム１はストレージアレイ装置１００３、記憶部２Ａ～２Ｄ、記憶部３、ストレージアレイ装置１００３と記憶部２Ａ～２Ｄおよび記憶部３とを接続するインタフェース１９，クライアント１００２、クライアント１００２とストレージアレイ装置１００３とを接続するストレージネットワーク１０００とを具備する。情報処理システム１において、記憶部２Ａ～２Ｄはストレージアレイ装置１００３に接続されてそれぞれが論理単位としての論理スロットとして認識され、それら論理スロットを用いてＲＡＩＤ (Redundant Arrays of Inexpensive Disks)アレイが構築される。ストレージアレイ装置１００３には、さらにデータ引き継ぎ先としての記憶部３が接続可能である。本実施形態においては寿命到達時処理前のＲＡＩＤアレイを構築する記憶部は４つである場合について説明するが、２乃至複数の任意の数の記憶部を用いてＲＡＩＤアレイを構築してもよい。また、本実施形態ではＲＡＩＤアレイとしてＲＡＩＤ５を用いるが、ＲＡＩＤ０,ＲＡＩＤ２,ＲＡＩＤ３,ＲＡＩＤ４,ＲＡＩＤ６など他のＲＡＩＤ技術や他のディスクアレイ実装形態を用いてディスクアレイを構築した場合についても本実施形態は適用可能である。

【０２４４】

ネットワーク１０００は、ストレージアクセスのためのストレージネットワークであり、たとえばファイバーチャネル(Fibre Channel)やイーサネット（登録商標）が用いられる。特に、ストレージネットワーク１０００として、例えばＳＡＮ(Storage Area Network)やＮＡＳ(Network Attached Storage)が用いられる。ＳＡＮには、たとえば、ＦＣ-ＳＡＮ(Fibre Channel Storage Area Network)やＩＰ-ＳＡＮ(Internet Protocol Storage Area Network)が用いられ、その上位プロトコルとして、たとえば、ＳＣＳＩ(Small Computer System Interface)が用いられる。本実施形態では、ストレージネットワーク１０００としてＩＰ-ＳＡＮを採用した例を示し、その上位プロトコルとしてはｉＳＣＳＩ(Internet Small Computer System Interface)が利用されるものとする。ストレージネットワーク１０００にはネットワークスイッチ１００００１やハブ（図示せず）が含まれる。

【０２４５】

クライアント１００２は、ストレージネットワーク１０００に接続され、所望の処理を遂行するコンピュータである。クライアント１００２は、典型的には、プロセッサと、メインメモリと、通信インタフェースと、ローカル入出力装置等のハードウェア資源を備え、また、デバイスドライバやオペレーティングシステム（ＯＳ）、アプリケーションプログラム等のソフトウェア資源を備える（図示せず）。これによって、クライアント１００２は、プロセッサの制御の下、各種のプログラムを実行して、ハードウェア資源との協働作用により処理を実現する。例えば、クライアント１００２は、プロセッサの制御の下、

業務アプリケーションプログラムを実行することにより、ストレージネットワーク１０００を経由してストレージアレイ装置１００３にＩ／Ｏアクセスし、所望の業務システムを実現する。または、クライアント１００２は、データベース管理システム(Database management system,ＤＢＭＳ)が稼働しているサーバであるデータベースサーバ(ＤＢサーバ)であってもよく、その場合、ストレージネットワーク１０００またはその他のネットワーク（図示せず）を経由してＤＢサーバに接続されたクライアント（図示せず）からのデータ読み出し要求を受信すると、ストレージアレイ装置１００３からデータを読みだしてそのクライアントに読み出しデータを送信し、そのクライアントからのデータ書き込み要求を受信するとそのクライアントから書き込みデータを受信してストレージアレイ装置１００３にデータを書き込む。

【０２４６】

ストレージアレイ装置１００３は論理スロット０～３をＲＡＩＤの構成単位として使用する。論理スロットは、第２の実施形態における論理ドライブに対応する。記憶部２Ａ～２Ｄのいずれか一つが寿命到達する前である通常状態において、インタフェース１９経由で記憶部２Ａ～２Ｄがストレージアレイ装置１００３に接続されており、論理スロット０には記憶部２Ａが、論理スロット１には記憶部２Ｂが、論理スロット２には記憶部２Ｃが、論理スロット３には記憶部２Ｄが割り当てられている。これにより、ストレージアレイ装置１００３は、４台の記憶部２Ａ～２Ｄに対応する４つの論理スロットをＲＡＩＤ５により仮想的な１台の論理デバイスとしてクライアント１００２に通知する。クライアント１００２はストレージアレイ装置にアクセスするためのＬＢＡ（以下、「アレイＬＢＡ」「ＡＬＢＡ」と呼ぶ）を送信する。ＲＡＩＤコントローラ１００５内の制御部２００はアレイＬＢＡを論理スロット番号と記憶部２Ａ～２ＤにアクセスするためのＬＢＡ（以下、「記憶部ＬＢＡ」「ＳＬＢＡ」と呼ぶ）に変換し、制御部２００は論理スロット番号で特定された記憶部２Ａ～２Ｄのうち少なくとも１台の記憶部のＳＬＢＡに対してアクセス命令を送信する。

【０２４７】

ストレージアレイ装置１００３は、それ単体で、クライアント１００２に対するデータストレージサービスを提供してもよいし、ストレージアレイ装置１００３と図示しない別のストレージアレイ装置から仮想的に構成される１つのストレージ装置として、クライアント１００２に対するデータストレージサービスを提供してもよい。ストレージアレイ装置１００３には、クライアント１００２に提供するための１つ以上の論理デバイス（ＬＤＥＶ）が形成される。

【０２４８】

論理デバイスは、クライアント１００２が認識しうる論理的な記憶装置であり、論理ユニット（ＬＵ）が割り当てられ、クライアント１００２は、物理デバイス上に形成された論理デバイスを論理ユニットとして認識する。各論理ユニットには、論理ユニット番号（ＬＵＮ）が付与される。また、論理ユニットは、論理セクタ（論理ブロック）に分割され、各論理セクタには、アレイＬＢＡが割り当てられる。クライアント１００２は、論理ユニット番号及びアレイＬＢＡからなる論理アドレスを含むコマンドをストレージアレイ装置１００３に与えることにより、特定の論理ユニットにおける特定の論理セクタに対してアクセスすることができる。ｉＳＣＳＩを利用した本実施形態では、クライアント１００２、ストレージアレイ装置１００３はそれぞれ、ｉＳＣＳＩネームが割り当てられたｉＳＣＳＩノードとしてのイニシエータ及びターゲットとして機能し、従って、クライアント１００２及びストレージアレイ装置１００３は、ＩＰアドレス及びＴＣＰポート番号の組み合わせで特定されるネットワークポータルを介して、ｉＳＣＳＩ ＰＤＵを送受する。従って、クライアント１００２は、ｉＳＣＳＩネーム、ＩＰアドレス及びＴＣＰポート番号を指定することにより、ネットワーク１０００上のストレージアレイ装置１００３を認識し、その論理ユニット内の論理セクタに対してアクセスする。

【０２４９】

記憶部２Ａ～２Ｄは、インタフェース１９を経由してストレージアレイ装置１００３に

対して接続される記憶部である。記憶部２Ａ～２Ｄは、たとえば、第１の実施形態に記載の記憶部２と同等の記憶部をそれぞれ用いることができる。本実施形態においては、例として、記憶部２Ａ～２Ｄは第１の実施形態に記載のＳＳＤを用いた場合について説明する。信頼性劣化後に廃棄して設置スペースを削減しかつ情報処理システム１全体の消費電力を削減するという観点では、記憶部２Ａ～２Ｄはストレージアレイ装置１００３に対して物理的に着脱可能であることが望ましい。

【０２５０】

記憶部３は記憶部２Ａ～２Ｄのいずれか一つが寿命到達または寿命到達目前であると判定された後に、新たにストレージアレイ装置１００３に接続される記憶部であり、たとえば第１の実施形態に記載の記憶部３と同等の記憶部を用いることができる。本実施形態においては、記憶部３は第１の実施形態に記載のＳＳＤを用いた場合について説明する。記憶部３の接続前の設置スペースを削減しかつシステム１全体の消費電力を削減するという観点、および記憶部３の信頼性劣化後に記憶部３を廃棄して設置スペースを削減しかつ情報処理システム１全体の消費電力を削減するという観点では、記憶部３はストレージアレイ装置１００３に対して物理的に着脱可能であることが望ましい。

【０２５１】

ＲＡＩＤコントローラ１００５はディスクインタフェース１００７に接続された記憶部のＲＡＩＤアレイの構築・管理をつかさどり、制御部２００を具備する。制御部２００は、ＲＡＩＤコントローラ内のメモリに格納されたファームウェアやソフトウェア、あるいはＲＡＩＤコントローラ１００５内のハードウェアなどの実装形態をとる。制御部２００は、ネットワークインタフェース１００４経由でクライアント１００２やネットワークスイッチ６００９等から命令を受信すると、ディスクインタフェース１００７経由で各記憶部に読み出し命令や書き込み命令やその他命令やデータを送信したり、各記憶部から応答やデータを受信したり、ネットワークインタフェース１００４経由でクライアント１００２に応答やデータを送信したりする。制御部２００は記憶部２Ａ～２Ｄや記憶部３の統計情報やステータス記憶領域５１０やスロット番号記憶領域５３０や引き継ぎ履歴記憶領域５５０の制御および管理を行い、統計情報に基づく寿命到達時処理やデータ引き継ぎ処理等を行う。なお、ステータス記憶領域５１０、スロット番号記憶領域５３０および引き継ぎ履歴記憶領域５５０は各記憶部内に格納されるのではなく、ＲＡＩＤコントローラ１００５内のメモリ領域（図示せず）など、情報処理システム１内の記憶領域に格納されるようにしてもよい。本実施形態では、ステータス記憶領域５１０に格納されるデータは、０乃至５の値を取りうる。それぞれの値は、対応する記憶部がそれぞれ

０：初期ディスク状態
１：通常状態
２：信頼性劣化状態
３：データ引き継ぎ元状態（保護状態）
４：データ引き継ぎ先状態
５：廃棄対象状態

の状態であることを示す。

【０２５２】

図５９に本実施形態における通常状態のアレイＬＢＡ（ＡＬＢＡ）と記憶部ＬＢＡ（ＳＬＢＡ）の変換方法を示す。ＲＡＩＤ５を採用する制御部２００は、３つの連続する論理セクタであるＡＬＢＡ＝３ｑ,３ｑ＋１,３ｑ＋２(ｑは０以上の任意の整数)を組として、それぞれのデータであるアレイデータＤ(ＡＬＢＡ＝３ｑ),Ｄ(ＡＬＢＡ＝３ｑ+１),Ｄ(ＡＬＢＡ＝３ｑ+２)に対して１論理セクタ分のパリティデータであるＰ(３ｑ,３ｑ+２)を生成する。

【０２５３】

パリティデータＰ(３ｑ,３ｑ＋２)は、

Ｐ(３ｑ,３ｑ+２)＝(Ｄ(ＡＬＢＡ＝３ｑ) XOR Ｄ(ＡＬＢＡ＝３ｑ+１) XOR Ｄ(ＡＬＢＡ＝３ｑ+２)

のように、D(ＡＬＢＡ＝３ｑ),D(ＡＬＢＡ＝３ｑ+１),D(ＡＬＢＡ＝３ｑ+２)内の論理セクタ内オフセットが同じである各ビットに対して排他的論理和をとって得られたビットをパリティデータP(3q,3q+２)の論理セクタ内の各オフセットのビットデータとすることにより得られる。例えば、ＡＬＢＡ＝０のデータであるD(０)およびＡＬＢＡ＝１のデータであるD(１)およびＡＬＢＡ＝２のデータであるD(２)からパリティデータP(０,２)が制御部２００によって計算される。通常状態においては、アレイデータD(ＡＬＢＡ＝３ｑ),D(ＡＬＢＡ＝３ｑ+１),D(ＡＬＢＡ＝３ｑ+２)およびパリティデータP(３ｑ,３ｑ+２)は、図５９に示すように、記憶部２Ａ～記憶部２Ｄに分散管理される。たとえばＡＬＢＡ＝１の場合、対応するデータD(１)は論理スロット１に割り当てられている記憶部２ＢのＬＢＡ＝ＳＬＢＡ＝１に格納されており、クライアント１００２からＡＬＢＡ＝１を受信すると制御部２００は論理スロット番号＝１およびＳＬＢＡ＝１を特定する。

【０２５４】

たとえば、制御部２００がクライアント１００２からＡＬＢＡ＝１の読み出し命令を受信すると、制御部２００はD(１)の格納先である論理スロット番号＝１およびＳＬＢＡ＝１を特定し、論理スロット番号＝１に接続されている記憶部２Ｂに対してＳＬＢＡ＝１の読み出し命令を送信し、受信した読み出しデータをクライアント１００２に送信する。もし記憶部２Ｂから応答が無かった場合または記憶部２Ｂからエラーが返信されてきた場合、制御部２００は論理スロット１以外のスロットである論理スロット０および論理スロット２～３に接続された記憶部２Ａと記憶部２Ｃと記憶部２ＤにＳＬＢＡ＝１の読み出し命令を送信し、受信したデータであるD(０),D(２),P(０,２)から排他的論理和によってD(１)を復元し、復元したD(１)をクライアント１００２に送信する。なお、D(１)の読み出し時にD(０),D(２),P(０,２)のデータ読み出しを並列して行なっても良い。

【０２５５】

たとえば、制御部２００がクライアント１００２からＡＬＢＡ＝１の書き込み命令および書き込みデータを受信すると、制御部２００はD(１)の格納先である論理スロット番号＝１およびＳＬＢＡ＝１を特定し、論理スロット番号＝１の記憶部２ＢにＳＬＢＡ＝１の書き込み命令を送信して書き込みデータを書き込むとともに、論理スロット番号１以外のスロットに接続されておりかつＳＬＢＡ＝１にパリティデータが格納されていない記憶部である記憶部２Ａおよび記憶部２ＣのＳＬＢＡ＝１からデータD(０)とD(２)を読み出し、D(０),D(１),D(２)からパリティデータP(０,２)を生成して記憶部２ＤのＳＬＢＡ＝１に書き込む。

【０２５６】

図６０に、論理スロット１の引き継ぎ状態遷移直後のアレイＬＢＡと記憶部ＬＢＡの変換方法を示す。引き継ぎ状態遷移直後においては、アレイＬＢＡと記憶部ＬＢＡの変換方法は、図５９に示す通常状態と同様である。

【０２５７】

図５８に示したように、記憶部２Ａ～２Ｄおよび記憶部３はそれぞれステータス格納領域５１０と論理スロット番号格納領域５３０を具備し、記憶部３は引き継ぎ履歴格納領域５５０を具備する。引き継ぎ履歴格納領域５５０は後述するデータ引き継ぎが完了すると消去されることにより、ユーザデータ格納など他の用途に使用されてもよい。ステータス格納領域５１０と論理スロット番号格納領域５３０と引き継ぎ履歴格納領域５５０には、ＬＢＡの領域が割り当てられても良い。あるいは、ステータス格納領域５１０と論理スロット番号格納領域５３０と引き継ぎ履歴格納領域５５０にはＬＢＡではなくログページアドレスが割り当てられたログ領域が割り当てられてもよく、この場合には、たとえば、非特許文献１のACS-3に記載されている2Fh Read Log Extにより該領域の読み出しが行われ、非特許文献１のACS-3に記載されている3Fh Write Log Extにより該領域への書き込みが行われる。引き継ぎ履歴格納領域５５０には後述するデータ引き継ぎ状態のログが格納される。

【０２５８】

引き継ぎ履歴格納領域５５０に記録される履歴データとして、本実施形態では図６１に

示すようなテーブルデータを使用する。データ引き継ぎ状態に記憶部３のＬＢＡ(＝ＳＬＢＡ)にセクタ長Ｘのデータが書き込まれると、制御部２００は引き継ぎ履歴格納領域５５０に(ＳＬＢＡ,Ｘ)を追記する。ＬＢＡ＝ＳＬＢＡ～ＳＬＢＡ＋Ｘ－１の領域に対し、オーバーラップするＬＢＡや連続するＬＢＡが引き継ぎ履歴格納領域５５０にすでに記録されている場合には、それらを合併したログを引き継ぎ履歴格納領域５５０に記録し、引き継ぎ履歴格納領域５５０から合併されたログを削除することが望ましい。たとえば、(ＳＬＢＡ,Ｘ)のＬＢＡ領域Ａを新たに引き継ぎ履歴格納領域５５０に記入する場合であって、(ＳＬＢＡ－ａ,ａ)のＬＢＡ領域Ｂがすでに引き継ぎ履歴格納領域５５０に記録されている時、ＬＢＡ領域ＡとＬＢＡ領域Ｂは連続しているため、Ａ＋Ｂの領域を示す(ＳＬＢＡ－ａ,ａ＋Ｘ)のデータを(ＳＬＢＡ－ａ,ａ)の履歴部に上書きすることで、ログデータ量を増大させることなく引き継ぎ履歴格納領域５５０を更新することができる。

【０２５９】

制御部２００は、第１の実施形態の図２４や図２９と同様にして、ディスクインタフェース１００７に接続されたそれぞれの記憶部が寿命に到達しているか否か、寿命到達目前であるか否か、あるいは故障目前であるか否かを判定し、寿命に到達しているか場合、寿命到達目前である場合、あるいは故障目前である場合に寿命到達時処理を行う。寿命到達判定は、第１の実施形態と同様に、図２４のような一定時間ごと、または一定処理数おき、または一定データ送受信おき、または図２９や図３０のように記憶部からのコマンド応答がエラーであった場合などに実施される。

【０２６０】

図６２は、たとえば、記憶部２Ｂが寿命に到達していると判定された場合の、記憶部２Ｂの寿命到達時処理を示す。記憶部２Ｂが寿命に到達していると判定された場合（ステップＳ５００）、制御部２００は記憶部２Ｂのステータスを１から２に書き換える（ステップＳ５０１）。制御部２００は空きスロットに新しい記憶部をディスクインタフェース１００７に接続するようディスプレイ装置やＬＥＤ等を経由してネットワーク管理者に通知することが望ましい（ステップＳ５０２）。あるいは、ディスクロード・アンロード装置（図示せず）として記憶部２Ａ～２Ｄや記憶部３のディスクインタフェース１００７に対する物理的な着脱を自動的に行う機械装置がシステム１に搭載されている場合、新たな記憶部をインタフェース１００７に接続するよう該ディスクロード・アンロード装置に対して命令してもよい（ステップＳ５０２）。

【０２６１】

新たな記憶部として記憶部３が接続された時（ステップＳ５０３）、制御部２００は記憶部３のステータス５１０を４に書き換え（ステップＳ５０４）、記憶部３の論理スロット番号格納領域５３０に記憶部２Ｂの論理スロット番号格納領域５３０のデータをコピーする（ステップＳ５０５）。今回の例では、記憶部２Ｂは図５７に示すように論理スロット１として割り当てられているため、記憶部３の論理スロット番号格納領域５３０には１が書き込まれる。制御部２００は記憶部２Ｂのステータス５１０を３に書き換え（ステップＳ５０６）、ＲＡＩＤコントローラ１００５に記憶部２Ｂと記憶部３を同一の論理スロットである論理スロット１として認識させる（ステップＳ５０７）。該寿命到達処理後、記憶部２Ｂのステータスは３に、記憶部３のステータスは４になっており、論理スロット１はデータ引き継ぎ状態に遷移していることになる。

【０２６２】

本実施形態では、各記憶部のステータスは、各記憶部のステータス格納領域５１０に不揮発に格納されている。図６３はＲＡＩＤコントローラ１００５が起動した時、およびディスクインタフェース１００７に記憶部が接続された時の制御部２００の処理手順を示すものである。ＲＡＩＤコントローラ１００５の起動ごとに制御部２００はステータス格納領域５１０を読み出すことで各記憶部のステータスを認識する。制御部２００は、各記憶部のステータスを認識し、かつ各記憶部の論理スロット番号を論理スロット番号格納領域５３０から読み出すことで、各論理スロット０～３それぞれがデータ引き継ぎ状態か否かを判定する。

【０２６３】
すなわち、制御部２００は、ディスクステータス＝０であるとき（ステップＳ５１１）、記憶部が初期ディスク状態であると認識し（ステップＳ５１２）、ディスクステータス＝１であるとき（ステップＳ５１３）、記憶部が通常状態であると認識し（ステップＳ５１４）。ディスクステータス＝２であるとき（ステップＳ５１５）、記憶部が信頼性劣化状態であると認識し（ステップＳ５１６）、ディスクステータス＝３であるとき（ステップＳ５１７）、記憶部がデータ引き継ぎ作業中のデータ引き継ぎ元状態（保護状態）として認識し（ステップＳ５１８）、ディスクステータス＝４であるとき（ステップＳ５１９）、記憶部がデータ引き継ぎ作業中のデータ引き継ぎ先状態として認識し（ステップＳ５２０）、ディスクステータス＝５であるとき（ステップＳ５２１）、記憶部が廃棄対象状態として認識し（ステップＳ５２２）、ディスクステータスが０～５以外であるときは、不正なディスクとみなす（ステップＳ５２３）。

【０２６４】
（論理ドライブからの読み出し１）
図６４はクライアント１００２からストレージアレイ装置１００３に読み出し要求が送信された時の、制御部２００の処理手順を示すものである。制御部２００は、クライアント１００２から読み出し対象アレイＬＢＡであるＡＬＢＡ＝ｃＡＬＢＡの読み出し命令を受信する（ステップＳ５３０）。制御部２００は、ｃＡＬＢＡから論理スロット番号ｃＳＬＯＴと、読み出し対象記憶部ＬＢＡであるＳＬＢＡ＝ｃＳＬＢＡを計算する（ステップＳ５３１）。制御部２００は、ｃＳＬＯＴの論理スロットの記憶部が故障中であるか否かを判定し（ステップＳ５３２）、ｃＳＬＯＴの論理スロットの記憶部が故障中でない場合、つぎにｃＳＬＯＴの論理スロットの記憶部がデータ引き継ぎ状態であるか否かを判定する（ステップＳ５３３）。

【０２６５】
ｃＳＬＯＴの論理スロットの記憶部がデータ引き継ぎ状態である場合（ステップＳ５３３）、制御部２００は、ｃＳＬＯＴのスロット以外からパリティを含むデータを読み出して、それらのデータを使用してｃＳＬＯＴのｃＳＬＢＡのデータを復元し、復元したデータをクライアント１００２に送信する（ステップＳ５３４）。そして、制御部２００は、ｃＳＬＯＴのｃＳＬＢＡの復元データをデータ引き継ぎ先の記憶部３に書き戻し、引き継ぎ履歴格納領域５５０に引き継ぎ履歴を記録する（ステップＳ５３５）。ｃＳＬＯＴの論理スロットの記憶部がデータ引き継ぎ中でない場合（ステップＳ５３３）、制御部２００は、ｃＳＬＯＴの記憶部からデータＤ（ｃＳＬＢＡ）を読み出し、読み出したデータをクライアント１００２に送信する。このように、読み出し対象の論理スロットがデータ引き継ぎ状態である場合は、読み出し対象以外のスロットから読み出し対象データを復元することで、データ引き継ぎ状態の記憶部のデータを損失した場合であっても救済可能であるとともに、データ引き継ぎ状態の記憶部からの読み出しを減らし、リードディスターブ（データを読み出したメモリセルと同一のブロックに含まれる非選択メモリセルのフローティングゲートに微妙な電荷が蓄えられることによって、記憶したデータに誤りが生じる現象）を抑制する。

【０２６６】
ｃＳＬＯＴの論理スロットの記憶部が故障中である場合、制御部２００は、データ引き継ぎ状態のスロットが存在するか否かを判定し（ステップＳ５３７）、データ引き継ぎ状態のスロットが存在しない場合、ｃＳＬＯＴのスロット以外からデータを読み出して、それらのデータを使用してｃＳＬＯＴのｃＳＬＢＡのデータを復元し、復元したデータをクライアント１００２に送信する（ステップＳ５３８）。このように読み出し対象のスロットが故障中であり、かつ他のスロットにデータ引き継ぎ状態のスロットが存在しない場合は、読み出し対象以外のスロットから読み出し対象データを復元する。

【０２６７】
ｃＳＬＯＴの論理スロットの記憶部が故障中であって、データ引き継ぎ状態のスロットが存在する場合（ステップＳ５３７）、制御部２００は、引き継ぎ履歴格納領域５５０を

読み出して読み出し対象ＳＬＢＡのデータが引き継ぎ元と引き継ぎ先のどちらに存在するかを判定し（ステップＳ５３９）、データ引き継ぎ先記憶部、データ引き継ぎ元記憶部および通常状態記憶部から読みだしたデータからｃＳＬＯＴのｃＳＬＢＡのデータを復元し、復元したデータをクライアント１００２に送信する（ステップＳ５４０）。データ引き継ぎ元のデータを使用した場合は、使用したデータ引き継ぎ元データをデータ引き継ぎ先の記憶部３に書き戻し、データ引き継ぎ先の引き継ぎ履歴格納領域５５０に引き継ぎ履歴を記録する（ステップＳ５４１）。

【０２６８】

（論理ドライブからの読み出し２）

図６５はクライアント１００２からストレージアレイ装置１００３に読み出し要求が送信された時の、制御部２００の他の処理手順を示すものである。図６５に示す処理手順では、図６４のステップ５３３の判定がＹｅｓのときの処理を、ステップＳ５３４、Ｓ５３５からステップＳ５５０～Ｓ５５４に置換している。また、図６５では、図６４のステップＳ５４１の処理を削除している。ｃＳＬＯＴの論理スロットの記憶部がデータ引き継ぎ状態である場合（ステップＳ５３３）、制御部２００は、データ引き継ぎ先の記憶部３から引き継ぎ履歴を読み出し、ｃＳＬＢＡのデータが引き継ぎ履歴に含まれるか否かを判定する（ステップＳ５５１）。制御部２００は、ｃＳＬＢＡのデータが引き継ぎ履歴に含まれる場合は、データ引き継ぎ先の記憶部３からｃＳＬＢＡのデータを読み出し、読み出したデータをクライアント１００２に送信する（ステップＳ５５４）。制御部２００は、ｃＳＬＢＡのデータが引き継ぎ履歴に含まれない場合は、データ引き継ぎ元の記憶部からｃＳＬＢＡのデータを読み出し、読み出したデータをクライアント１００２に送信する（ステップＳ５５２）。そして、データ引き継ぎ元から読み出したデータをデータ引き継ぎ先の記憶部３に書き戻し、データ引き継ぎ先の記憶部３の更新履歴５５０を更新する（ステップＳ５５３）。

【０２６９】

（論理ドライブへの書き込み）

図６６は、本実施形態におけるクライアント１００２からの書き込み命令に対する処理を示すフローチャートである。制御部２００は、クライアント１００２からアレイＬＢＡであるｃＬＢＡとセクタ長を含む書き込み命令を受信すると（ステップＳ５６０）、クライアント１００２から書き込みデータを受信する（ステップＳ５６１）。制御部２００は、データを書き込むべき論理スロット番号ｃＳＬＯＴをｃＡＬＢＡから計算し、データを書き込むべき記憶部ＬＢＡであるｃＳＬＢＡをｃＡＬＢＡから検索し、ｃＳＬＢＡのパリティデータ格納先の論理スロット番号ｃＰＳＬＯＴをｃＡＬＢＡから計算する（ステップＳ５６２）。制御部２００は、全スロットのｃＳＬＢＡからデータを並列に読み出す。故障スロットがある場合は、故障スロットのデータは、故障スロット以外からデータを読み出して、故障スロットのデータを復元する（ステップＳ５６３）。以後、制御部２００は、本体データ書き込みタスクとパリティデータ書き込みタスクを並列に処理する。

【０２７０】

本体データ書き込みタスクは次のように実行される。制御部２００は、ｃＳＬＯＴが故障中かどうかを判定し（ステップＳ５６４）、故障中の場合にはデータを書き込まずタスクを終了する。故障中でない場合には、ｃＳＬＯＴがデータ引き継ぎ状態かどうかを判定し（ステップＳ５６５）、データ引き継ぎ状態でない場合には、スロットｃＳＬＯＴのｃＳＬＢＡに対してクライアント１００２からの受信データを書き込む（ステップＳ５６６）。データ引き継ぎ状態の場合には、スロットｃＳＬＯＴに割り当てられたデータ引き継ぎ元記憶部とデータ引き継ぎ先記憶部のうちデータ引き継ぎ先記憶部のｃＳＬＢＡに受信データを書き込み（ステップＳ５６７）、データ引き継ぎ先記憶部の引き継ぎ履歴格納部５５０にｃＳＬＢＡとセクタサイズをデータ引き継ぎ履歴として記録する（ステップＳ５６８）。引き継ぎ履歴格納部５５０に、ｃＳＬＢＡの前後の連続したＬＢＡや重複したＬＢＡの引き継ぎ履歴が存在する場合には、それらのログを合併した引き継ぎ履歴を引き継ぎ履歴５５０に記入し、合併元の引き継ぎ履歴を削除する（無効化する）（ステップＳ５

６８）。
【０２７１】
なお、引き継ぎ履歴５５０からの引き継ぎ履歴削除は論理的な削除でよく、ＮＡＮＤメモリ１６のブロック消去によって物理的に消去しなくてもよい。たとえば、ログを無効化するフラグを引き継ぎ履歴５５０に記入したり、更新後の引き継ぎ履歴５５０を更新前の引き継ぎ履歴５５０とは別の領域に格納し引き継ぎ履歴５５０の格納位置を示すポインタを更新したりすることで、引き継ぎ履歴５５０から引き継ぎ履歴が論理的に削除される。
【０２７２】
パリティデータ書き込みタスクは次のように実行される。制御部２００は、全論理スロットからＲＡＩＤコントローラ１００５内のメモリに読み出されたデータに対してクライアント１００２から受信した書き込みデータを上書きし、パリティデータを再計算する（ステップＳ５７０）。制御部２００は、ｃＰＳＬＯＴが故障中かどうかを判定し（ステップＳ５７１）、故障中の場合にはパリティデータを書き込まずタスクを終了する。故障中でない場合には、ｃＰＳＬＯＴがデータ引き継ぎ状態かどうかを判定し（ステップＳ５７２）、データ引き継ぎ状態でない場合には、スロットｃＰＳＬＯＴのｃＳＬＢＡに対してパリティデータを書き込む（ステップＳ５７３）。データ引き継ぎ状態の場合には、スロットｃＰＳＬＯＴに割り当てられたデータ引き継ぎ元記憶部とデータ引き継ぎ先記憶部のうちデータ引き継ぎ先記憶部のｃＳＬＢＡにパリティデータを書き込み（ステップＳ５７４）、データ引き継ぎ先記憶部の引き継ぎ履歴格納部５５０にｃＳＬＢＡとセクタサイズを引き継ぎ履歴として記録する。引き継ぎ履歴５５０に、ｃＳＬＢＡの前後の連続したＬＢＡや重複したＬＢＡの引き継ぎ履歴が存在する場合には、それらの引き継ぎ履歴を合併した引き継ぎ履歴を引き継ぎ履歴格納領域５５０に記入し、合併元の引き継ぎ履歴を引き継ぎ履歴５５０から削除する（ステップＳ５７５）。
【０２７３】
このようにして、制御部２００は、クライアントからの書き込み命令に対し、データ引き継ぎ状態ではない各論理スロット０,２,３に割り当てられた各記憶部２Ａ,２Ｃ,２Ｄに対してはＳＬＢＡの書き込み命令を送信する一方で、データ引き継ぎ中である論理スロット１に対しては、データ引き継ぎ元である記憶部２Ｂに対して書き込み命令を送信せず、データ引き継ぎ先である記憶部３に対して書き込み命令を送信し、記憶部３の引き継ぎ履歴格納領域５５０に引き継ぎ履歴を記録する。
【０２７４】
図６７に更に詳細な例を示す。図６７は、論理スロット１がデータ引き継ぎ状態で、データ引き継ぎ元として記憶部２Ｂが割り当てられ、データ引き継ぎ先として記憶部３が割り当てられており、論理スロット０,２,３は通常状態である場合に、ストレージアレイ装置１００３のＡＬＢＡ＝１にＤ(１)を、ＡＬＢＡ＝７にＤ(７)を、ＡＬＢＡ＝１６にＤ(１６)を書き込む命令がクライアント１００２から送信された場合の例を示すものである。
【０２７５】
（例１）
制御部２００はＡＬＢＡ＝１に対する新たなデータＤ(１)ｎｅｗの書き込み命令をクライアント１００２から受信すると、ＡＬＢＡ＝１からＳＬＢＡ＝０および論理スロット１を特定する。論理スロット１がデータ引き継ぎ状態ではない場合には、Ｄ(１) ｎｅｗは記憶部２ＢのＳＬＢＡ＝０に書き込まれるが、今回の例では論理スロット１はデータ引き継ぎ状態であるため、制御部２００はＤ(１) ｎｅｗを記憶部３に書き込む。制御部２００は前述した読み出し動作に基づき論理スロット０の記憶部２ＡからＤ（０）を読み出し、論理スロット２の記憶部２ＣからＤ(２)を読み出し、Ｄ(０)とＤ(１) ｎｅｗとＤ(２)から排他的論理和で新しいパリティデータＰ(０,２)newを計算し、Ｐ(０,２)newを論理スロット３の記憶部２ＤのＳＬＢＡ＝０に格納する。なお、処理速度向上のためには、記憶部３に対するＤ(１) ｎｅｗの書き込み命令と、記憶部２Ａに対するＤ(０)の読み出し命令と、記憶部２Ｃに対するＤ(２)の読み出し命令は、制御部２００から同時に送信される

ことが望ましく、さらにはD(０)とD(２)の読み出しが完了した時点で、D(１)の書き込み完了を待たずしてP(０,２)newを計算し記憶部２Ｄに書き込み命令を送信することが望ましい。制御部２００は記憶部３の引き継ぎ履歴格納領域５５０に、記憶部３に書き込みを行ったＳＬＢＡとセクタカウントである(ＳＬＢＡ,セクタカウント)を記録する。たとえば、この例では記憶部３のＳＬＢＡ＝０から１セクタのＳＬＢＡ領域に書き込みを行ったので、引き継ぎ履歴格納領域５５０に(０,１)を追記する。前述のとおり、連続するＬＢＡ領域やオーバーラップするＬＢＡ領域がすでに引き継ぎ履歴格納領域５５０に記録されている場合には、これらを統合したＳＬＢＡ領域を引き継ぎ履歴格納領域５５０に記録する。データ引き継ぎ状態ではない論理スロットに書き込みを行った場合には、引き継ぎ履歴５５０の更新は行わない。

【０２７６】

（例２）

制御部２００はＡＬＢＡ＝７に対する新たなデータD(７)ｎｅｗの書き込み命令をクライアント１００２から受信すると、ＡＬＢＡ＝７からＳＬＢＡ＝２および論理スロット２を特定する。論理スロット２はデータ引き継ぎ状態ではないためD(７) ｎｅｗは記憶部２ＢのＳＬＢＡ＝０に書き込まれる。制御部２００は、前述の読み出し動作に基づき論理スロット０、論理スロット３からD(６)とD(８)の読み出しを行い、パリティデータP(６,８)newを計算し、論理スロット１に格納する。論理スロット１はデータ引き継ぎ状態であるため、制御部２００はP(６,８)newを記憶部２Ｂではなく記憶部３に書き込み、記憶部３の引き継ぎ履歴格納領域５５０に、記憶部３に書き込みを行ったＳＬＢＡとセクタカウントである(ＳＬＢＡ,セクタカウント)＝（２, １）を記録する。

【０２７７】

（例３）

制御部２００はＡＬＢＡ＝１６に対する新たなデータD(１６) ｎｅｗの書き込み命令をクライアント１００２から受信すると、ＡＬＢＡ＝１６からＳＬＢＡ＝５および論理スロット１を特定する。今回の例では論理スロット１はデータ引き継ぎ状態であるため、制御部２００はD(１６) ｎｅｗを記憶部３に書き込む。制御部２００は前述した読み出し動作に基づき論理スロット０の記憶部２AからD（１５）を読み出し、論理スロット２の記憶部２DからD(１７)を読み出し、D(１５)とD(１６) ｎｅｗとD(１７)から排他的論理和で新しいパリティデータP(１５,１７)newを計算し、論理スロット２の記憶部２ＣのＳＬＢＡ＝５に格納する。制御部２００は記憶部３の引き継ぎ履歴格納領域５５０に、記憶部３に書き込みを行ったＳＬＢＡとセクタカウントである(ＳＬＢＡ,セクタカウント)を記録する。たとえば、この例では記憶部３のＳＬＢＡ＝５から１セクタのＬＢＡ領域に書き込みを行ったので、引き継ぎ履歴領域５５０に(５,１)を追記する。前述のとおり、連続するＳＬＢＡ領域やオーバーラップするＳＬＢＡ領域がすでに引き継ぎ履歴５５０に記録されている場合には、これらを統合したＳＬＢＡ領域を引き継ぎ履歴格納領域５５０に記録する。データ引き継ぎ状態ではないスロットに書き込みを行った場合には、引き継ぎ履歴格納領域５５０の更新は行わない。このようにして、第２の実施形態と同じように、クライアント１００２から書き込み要求を受信するごとに、対応するＳＬＢＡのデータをデータ引き継ぎ元の記憶部２Ｂではなくデータ引き継ぎ先の記憶部３に書き込みデータ引き継ぎ履歴を記録することで、新規データ書き込みとデータ引き継ぎ動作を同時に行うことが可能になる。

【０２７８】

（バックグラウンドバックアップ）

制御部２００はクライアント１００２から命令を一定時間以上受信しない場合などのアイドル時、またはスタンバイモード遷移要求をクライアント１００２から受信した場合などに、記憶部２Ｂから記憶部３へのバックグラウンドバックアップを行う。制御部２００は、記憶部３の引き継ぎ履歴格納領域５５０を読み出し、引き継ぎ履歴格納領域５５０に記録されていないＳＬＢＡに対して、記憶部２Ｂからデータ読み出しを行い、記憶部３に該データを書き込み、記憶部３の引き継ぎ履歴格納領域５５０に該ＳＬＢＡと書き込みデ

ータのセクタサイズを格納することでバックグラウンドバックアップが行われる。引き継ぎ履歴５５０へのＳＬＢＡの格納の際、連続したＳＬＢＡ領域が存在する場合、またはオーバーラップしたＳＬＢＡ領域が存在する場合には、このＳＬＢＡ領域と統合したＳＬＢＡ領域が引き継ぎ履歴５５０に格納され、統合前のＳＬＢＡ領域は引き継ぎ履歴格納領域５５０から削除されることが望ましい。

【０２７９】

引き継ぎ履歴格納領域５５０に格納される引き継ぎ履歴のデータサイズを低減するという観点、および後述するデータ引き継ぎ終了判定を高速に行うという観点では、引き継ぎ履歴５５０で示されるＳＬＢＡ領域以外のＳＬＢＡ領域のうち小さく断片化されたＳＬＢＡ領域（断片化した未データ引き継ぎＳＬＢＡ領域）に対してバックグラウンドバックアップが優先的に行われることが望ましい。断片化した未データ引き継ぎＳＬＢＡ領域をバックアップすることにより、前記断片化未データ引き継ぎ領域の前記断片化未データ引き継ぎ領域の前後の既データ引き継ぎＳＬＢＡ領域と、新たにデータ引き継ぎされたＳＬＢＡ領域が引き継ぎ履歴５５０に統合されて格納される。

【０２８０】

（データ引き継ぎ状況の監視）

制御部２００は引き継ぎ履歴格納領域５５０を読み出すことでデータ引き継ぎ中の論理スロットのデータ引き継ぎ状況を監視する。図６８は、引き継ぎ履歴を用いたデータ引き継ぎ状況の監視手順を示すものである。たとえば、制御部２００は所定時間経過毎に引き継ぎ履歴格納領域５５０からデータ引き継ぎ履歴を読み出すことでデータ引き継ぎ状況を監視する（ステップＳ６００、Ｓ６０１）。制御部２００は、全ての引き継ぎ対象ＳＬＢＡが引き継ぎ履歴５５０に含まれる場合、データ引き継ぎが完了したと判定する。たとえば、データ引き継ぎ元である記憶部２Ｂの全ＳＬＢＡが引き継ぎ履歴５５０に含まれる場合、データ引き継ぎが完了したと判定される（ステップＳ６０２）。なお、データ引き継ぎ完了の判定として、例えば、データ引き継ぎ先である記憶部３の全ＳＬＢＡが引き継ぎ履歴５５０に含まれるか否かを判定するようにしてもよい。

【０２８１】

データ引き継ぎが完了したと判定された場合、制御部２００はデータ引き継ぎ元である記憶部２のステータスを５の廃棄対象状態に変更し、データ引き継ぎ先である記憶部３のステータスを１の通常状態に変更することで（ステップＳ６０３）、論理スロット１のデータ引き継ぎ状態を終了し、論理スロット１のデータ引き継ぎ状況監視を終了する。記憶部２Ｂが消費する電力を低減するという観点では、制御部２００は記憶部２に低消費電力モードへの遷移要求を送信することが望ましい。データ引き継ぎ状態終了後、制御部２００はディスプレイ９やポート近傍に設置されたＬＥＤを通じて、記憶部２Ｂをインタフェース１９から取り外すようユーザや管理者に通知したり、記憶部２Ｂをインタフェース１９から取り外すよう前記ディスクロード・アンロード装置に命令を送信することが望ましい。

【０２８２】

（全体のステータスの遷移）

図６９は記憶部２Ｂが寿命到達する際の各記憶部のステータスの遷移を示したものである。記憶部２Ａ～２Ｄすべてが通常状態である場合には、記憶部２Ａ～２Ｄすべてのステータスは１になっており、記憶部３はディスクインタフェース１００７に未接続状態である（ステップ１）。記憶部２Ｂが寿命到達したと制御部２００が判定すると、制御部２００は記憶部２Ｂのステータスを２に変更する（ステップ２）。制御部２００からの通知または命令にもとづき、新しい記憶部として記憶部３がインタフェース１００７に接続される（ステップ３）。記憶部３が接続されると、制御部２００はデータ引き継ぎ元としての記憶部２Ｂのステータスを３に変更し、データ引き継ぎ先としての記憶部３のステータスを４に変更することで、論理スロット１の状態をデータ引き継ぎ状態に移行させる（ステップ４）。引き継ぎ履歴格納領域５５０から読み出される情報をもとに、記憶部２Ｂのすべての有効データが記憶部３にデータ引き継ぎされたと判定されると、制御部２００は、

記憶部２Ｂのステータスを５に変更し、記憶部３のステータスを１に変更する。以後、記憶部２Ａ，記憶部３,記憶部２Ｃ,記憶部２Ｄが、元の記憶部２Ａ,記憶部２Ｂ,記憶部２Ｃ,記憶部２Ｄであるかのように振る舞う（ステップ１に戻る）。その後さらに記憶部２Ａ,記憶部３,記憶部２Ｃ,記憶部２Ｄのいずれかが寿命到達したと判定された場合、同様のステップ２～５が繰り返される。このようにして、ストレージアレイ装置１００３の記憶部のいずれか一つが寿命到達した場合、寿命到達目前である場合、故障寸前である場合となっても、本実施形態により記憶部のデータを新しい記憶部に容易にデータ引き継ぎすることが可能になる。本実施形態において、データ引き継ぎ動作はバックグラウンドで行われるため、バックアップ動作による性能劣化を抑制することが可能になる。

【０２８３】

（１台のデータ引き継ぎ状態と１台の故障が同時に発生した場合）

図７０は、１台のデータ引き継ぎ状態と他の１台の故障が同時に発生した場合の例を示すものである。図７０では、論理スロット１がデータ引き継ぎ状態で、データ引き継ぎ元として記憶部２Ｂが割り当てられ、データ引き継ぎ先として記憶部３が割り当てられており、論理スロット０,２,３は通常状態である場合に、論理スロット３の記憶部２Ｄに故障が発生した場合の例を示すものである。

【０２８４】

ＳＬＢＡ＝０からの読み出しの場合、論理スロット０のＤ（０）、論理スロット２のＤ（２）に関しては復元不要であり、論理スロット１のＤ（１）に関しては、データ引き継ぎ先の記憶部３から最新のデータであるＤ（１）ｎｅｗを読み出し可能である。ＳＬＢＡ＝２からの読み出しの場合、論理スロット０のＤ（６）、論理スロット２の２のＤ（７）に関しては復元不要であり、論理スロット３のＤ（８）に関しては、Ｄ（６）とＤ（７）と、データ引き継ぎ先の記憶部３からＰ（６，８）ｎｅｗとの排他的論理和をとることで復元可能である。ＳＬＢＡ＝５からの読み出しの場合、論理スロット０のＤ（１５）に関しては復元不要であり、論理スロット１のＤ（１６）に関しては、データ引き継ぎ先の記憶部３から最新のデータであるＤ（１６）ｎｅｗを読み出し可能であり、論理スロット３のＤ（１７）に関しては、Ｄ（１５）とデータ引き継ぎ先の記憶部３のＤ（１６）newと、論理スロット２のＰ（１５，１７）ｎｅｗとの排他的論理和をとることで復元可能である。

【０２８５】

（２台のデータ引き継ぎ状態と１台の故障が同時に発生した場合）

図７１は、２台のデータ引き継ぎ状態と他の１台の故障が同時に発生した場合の例を示すものである。図６８では、論理スロット１および論理スロット２がデータ引き継ぎ状態で、データ引き継ぎ元として記憶部２Ｂ、２Ｃが割り当てられ、データ引き継ぎ先として記憶部３Ｂ、３Ｃが割り当てられており、論理スロット０,３は通常状態である場合に、論理スロット３の記憶部２Ｄに故障が発生した場合の例を示すものである。

【０２８６】

ＳＬＢＡ＝０からの読み出しの場合、論理スロット０のＤ（０）、論理スロット２のＤ（２）に関しては復元不要であり、論理スロット１のＤ（１）に関しては、データ引き継ぎ先の記憶部３ＢからＤ（１）ｎｅｗを読み出し可能である。ＳＬＢＡ＝２からの読み出しの場合、論理スロット０のＤ（６）に関しては復元不要であり、論理スロット２のＤ（７）に関しては、データ引き継ぎ先の記憶部３Ｃから最新のデータであるＤ（７）ｎｅｗを読み出し可能であり、論理スロット３のＤ（８）に関しては、Ｄ（６）と、Ｄ（７）newと、Ｐ（６，８）newの排他的論理和をとることで復元可能である。ＳＬＢＡ＝５からの読み出しの場合、論理スロット０のＤ（１５）に関しては復元不要であり、論理スロット１のＤ（１６）に関しては、データ引き継ぎ先の記憶部３Ｂから最新のデータであるＤ（１６）newを読み出し可能であり、論理スロット３のＤ（１７）に関しては、Ｄ（１５）とデータ引き継ぎ先の記憶部３ＢのＤ（１６）newと、論理スロット２のデータ引き継ぎ先の記憶部３ＣのＰ（１５，１７）newとの排他的論理和をとることで復元可能である。

【０２８７】

（１台のデータ引き継ぎ状態とデータ読み出し不能エラーが同時に発生した場合）
図７２は、データ引き継ぎ中に１台の記憶部でＥＣＣ訂正不能エラーなどのアンコレクタブル リードエラーが発生した場合の例を示すものである。
（例１）
制御部２００はＡＬＢＡ＝２に対する新たなデータD(２)の読み出し命令をクライアント１００２から受信すると、ＡＬＢＡ＝２からＳＬＢＡ＝０および論理スロット２を特定する。この論理スロット２の記憶部２Ｃからのデータ（２）の読み出しの際にＥＣＣ訂正不能エラーが発生した場合は、Ｄ（０）と、データ引き継ぎ先の記憶部３のＤ（１）newと、論理スロット３のＰ（０，２）newとの排他的論理和をとることでＤ（２）を復元可能である。
（例２）
制御部２００はＡＬＢＡ＝７に対する新たなデータD(７)の読み出し命令をクライアント１００２から受信すると、ＡＬＢＡ＝７からＳＬＢＡ＝２および論理スロット３を特定する。この論理スロット３の記憶部２Ｄからのデータ（８）の読み出しの際にＥＣＣ訂正不能エラーが発生した場合は、Ｄ（６）と、Ｄ（７）newと、記憶部３のＰ（６，８）newとの排他的論理和をとることでＤ（８）を復元可能である。
（例３）
制御部２００はＡＬＢＡ＝９に対する新たなデータD(９)の読み出し命令をクライアント１００２から受信すると、ＡＬＢＡ＝９からＳＬＢＡ＝３および論理スロット１を特定する。この論理スロット１の記憶部２Ｂからのデータ（９）の読み出しの際にＥＣＣ訂正不能エラーが発生した場合は、Ｄ（１０）と、Ｄ（１１）と、Ｐ（９，１０）との排他的論理和をとることでＤ（９）を復元可能である。
【０２８８】
（第４の実施形態）
第４の実施形態では、本発明をストレージネットワークからなる情報処理システム１に適用した場合について説明する。本実施形態においては、情報処理システム１は図７３に示すように構成される。なお、情報処理システム１は図７４に示すように構成されてもよい。図７３に示すように、本実施形態における情報処理システム１は、記憶部２、記憶部３、１乃至複数のその他の記憶部である記憶部６００４、１乃至複数のメタデータサーバ６００３、サーバ６００１、記憶部２と記憶部３と記憶部６００４とメタデータサーバ６００３とサーバ６００１を相互接続するストレージネットワーク１０００、１乃至複数のクライアント６００２、サーバ６００１とクライアント６００２を相互接続するネットワーク６０００からなる。図７４の場合は、ストレージネットワーク１０００は、チャネルエクステンダ、ＷＡＮなどの長距離ネットワークを介してストレージネットワーク６００１ｂに接続され、このストレージネットワーク６００１ｂに記憶部３が接続されている。
【０２８９】
記憶部２、３は、ステータス格納領域５１０を具備する。記憶部２、３には、第１の実施形態と同様のデータ引き継ぎ対象記憶部アドレス格納領域２００７、３００７（図３２参照）を具備させてもよいし、第２の実施形態と同様の論理ドライブＩＤ格納領域５２０を具備させてもよい。ステータス格納領域５１０は、前述と同様、対応する記憶部がそれぞれ
０：初期ディスク状態
１：通常状態
２：信頼性劣化状態
３：データ引き継ぎ元状態（保護状態）
４：データ引き継ぎ先状態
５：廃棄対象状態
の状態であることを示す。
【０２９０】
メタデータサーバ６００３は、第１の実施形態と同様のＯＳ１００と制御部２００とメ

タデータ３００と論理ドライブステータステーブル４５０を格納する主メモリ６と、ＯＳ１００や制御部２００を実行する演算装置であるＣＰＵ５を具備し、第１の実施形態の情報処理装置１１１のＯＳ１００や制御部２００に相当する機能を担う。メタデータサーバ６００３の主メモリ６に格納されるメタデータ３００はたとえば図３３に示したような構造であり、主メモリ６上のメタデータやメタデータのジャーナルは、メタデータサーバ６００３内の不揮発性記憶部や、ストレージネットワーク１０００内の不揮発性記憶部や、記憶部２、記憶部３などに退避保存される。また、論理ドライブステータステーブル４５０は、例えば図４０に示したような構造であり、論理ドライブが通常状態かデータ引き継ぎ状態であることを示す。メタデータサーバ６００３は、たとえばサーバ６００１からファイルＩＤを受信すると、メタデータ領域３００からファイルＩＤを検索し、該ファイルＩＤに対応する論理ドライブアドレスや、記憶部アドレスであるディスク識別ＩＤやＬＢＡやセクタカウントなどを特定する。記憶部アドレスであるディスク識別ＩＤにはたとえばＩＰ(Internet Protcol)アドレスやＭＡＣ(Media Access Control)アドレスやＷＷＮ(World Wide Name)を用いてもよい。ステータステーブル５１０は、ストレージネットワーク１０００に接続される各記憶部２、３のステータスを記憶する。

【０２９１】

なお、論理ドライブステータステーブル４５０を図４１に示した構造とする他に、メタデータサーバ６００３の主メモリ６に図７５に示すようなステータステーブル６５０を格納し、論理ドライブステータステーブル４５０を格納しないようにしてもよい。このステータステーブル６５０においては、論理ドライブアドレスと、ディスク識別ＩＤと、論理ドライブステータスと、ディスクステータを管理するようにしており、ステータステーブル６５０は論理ドライブステータステーブル４５０およびステータス格納領域５１０の代わりに使用される。この場合は、記憶部２，３でのステータス格納領域５１０は不要となる。図７５では、論理ドライブアドレスＢは、ディスク識別ＩＤがｂ１とｂ２の２つの記憶部を有し、これら記憶部がデータ引き継ぎ状態で、ｂ１のディスク識別ＩＤを有する記憶部がデータ引き継ぎ元で、ｂ２のディスク識別ＩＤを有する記憶部がデータ引き継ぎ先であることが示されている。

【０２９２】

ストレージネットワーク１０００は、ストレージアクセスのためのネットワークであり、たとえばファイバーチャネル(Fibre Channel)やイーサネット（登録商標）が用いられる。特に、ストレージネットワーク１０００として、例えばＳＡＮ(Storage Area Network)やＮＡＳ(Network Attached Storage)が用いられる。ＳＡＮには、たとえば、ＦＣ-ＳＡＮ(Fibre Channel Storage Area Network)やＩＰ-ＳＡＮ(Internet Protocol Storage Area Network)が用いられ、その上位プロトコルとして、たとえば、ＳＣＳＩ(Small Computer System Interface)が用いられる。本実施形態では、ストレージネットワーク１０００としてＩＰ-ＳＡＮを採用した例を示し、その上位プロトコルとしてはｉＳＣＳＩが利用されるものとする。ストレージネットワーク１０００にはネットワークスイッチ６００９やハブ（図示せず）が含まれる。

【０２９３】

ネットワーク６０００は、クライアントがサーバ６００１にアクセスして種々のサービスを利用するためのネットワークであり、たとえばファイバーチャネル(Fibre Channel)やイーサネット（登録商標）が用いられる。ネットワーク６０００として、たとえばＷＡＮやＬＡＮなどがある。ネットワーク６０００にはネットワークスイッチ（図示せず）やハブ（図示せず）が含まれる。

【０２９４】

クライアント１００２は、ネットワーク６０００に接続され、所望の処理を遂行するコンピュータである。クライアント１００２は、典型的には、プロセッサと、メインメモリと、通信インタフェースと、ローカル入出力装置等のハードウェア資源を備え、また、デバイスドライバやオペレーティングシステム（ＯＳ）、アプリケーションプログラム４００等のソフトウェア資源を備える（図示せず）。これによって、クライアント１００２は

、プロセッサの制御の下、各種のプログラムを実行して、ハードウェア資源との協働作用により処理を実現する。例えば、クライアント６００２は、プロセッサの制御の下、業務アプリケーションプログラムを実行することにより、ネットワーク６０００を経由してサーバ６００１にＩ／Ｏアクセスし、所望の業務システムを実現する。

【０２９５】

サーバ６００１は、ストレージネットワーク１０００とネットワーク６０００に接続され、クライアント１００２からの要求に応じて所望の処理を遂行するコンピュータである。サーバ６００１は、典型的には、プロセッサと、メインメモリと、通信インタフェースと、ローカル入出力装置等のハードウェア資源を備え、また、デバイスドライバやオペレーティングシステム（ＯＳ）、アプリケーションプログラム等のソフトウェア資源を備える（図示せず）。これによって、サーバ６００１は、プロセッサの制御の下、各種のプログラムを実行して、ハードウェア資源との協働作用により処理を実現する。例えば、サーバ６００１は、クライアント１００２からの要求に応じて、プロセッサの制御の下、アプリケーションサービスプログラムを実行することにより、ストレージネットワーク６０００を経由して記憶部２または記憶部３または記憶部６００４にＩ/Ｏアクセスし、所望のアプリケーションサービスプログラムを実現する。

【０２９６】

サーバ６００１は、たとえば、クライアント６００２からのファイルデータ読み出し要求とファイルＩＤを受信すると、メタデータサーバ６００３にファイルＩＤを送信し、メタデータサーバ６００３からファイルが格納されている記憶部のＩＰアドレスやＭＡＣアドレスやＷＷＮなどの記憶部アドレスとＬＢＡを受け取り、記憶部アドレスを指定したパケットをストレージネットワーク１０００に送信することで記憶部に読み出し命令を送信し、記憶部から読み出しデータを受信し、クライアント１００２に読み出しデータを送信する。サーバ６００１は、たとえば、クライアント１００２からのファイルのデータ書き込み要求とファイルＩＤを受信すると、メタデータサーバ６００３にファイルＩＤを送信し、メタデータサーバ６００３からファイルを格納すべき記憶部の記憶部アドレスとＬＢＡを受け取り、クライアント６００２から書き込みデータを受信し、記憶部にデータを送信することでデータを書き込む。サーバ６００１は、データベース管理システム(Database management system，ＤＢＭＳ)が稼働しているサーバであるデータベースサーバ(ＤＢサーバ)であってもよい。

【０２９７】

（寿命到達処理）

制御部２００は起動されると、記憶部２についての前述した統計情報６５などの信頼性情報を監視する。制御部２００は、例えば図２４に示すように、一定時間おき（たとえば１分おき）または一定処理ごとに記憶部２から統計情報６５を取得する。制御部２００は、第１の実施形態と同様、取得した信頼性情報に基づいて、記憶部２が寿命に到達したか否かを判定し、記憶部２が寿命に到達したと判定された場合、後述する寿命到達時処理を実行する。

【０２９８】

（寿命到達時処理）

制御部２００は、接続されている記憶部２が寿命到達するかあるいは寿命到達直前になって寿命到達時処理が開始されると、記憶部２のステータス５１０をディスクステータス＝２（信頼性劣化状態）に変更した後、新しいディスクの接続を促すメッセージをメタデータサーバ６００３やサーバ６００１やクライアント６００２などのディスプレイ９などに表示する。新たな記憶部３が接続されると、制御部２００は、ステータステーブル４５０の記憶部３のステータス５１０をディスクステータス＝４（データ引き継ぎ先状態）に書き換え、さらに記憶部２のステータス５１０をディスクステータス＝３（保護状態）に書き換える。そして、制御部２００は、記憶部２＋記憶部３を１つの論理ドライブ４としてＯＳ１００に認識させる。制御部２００は、主メモリ６上の論理ドライブステータステーブル４５０またはステータステーブル６５０に格納された論理ドライブ４のステータス

を、ステータスが「通常状態」から「データ引き継ぎ状態」になるように書き換える。
【０２９９】
（データ引き継ぎ状態での論理ドライブからの読み出し）
サーバ６００１は、クライアント６００２からの読み出し要求とファイルＩＤを受信すると、メタデータサーバ６００３にファイルＩＤを送信する。メタデータサーバ６００３は、メタデータ３００からファイルＩＤに対応する論理ドライブ４を特定し、さらに主メモリ６から論理ドライブステータステーブル４５０またはステータステーブル６５０を読み出し、特定した論理ドライブ４のステータスがデータ引き継ぎ状態であることを認識する。メタデータサーバ６００３は、メタデータ３００からファイルＩＤで指定されたファイルが格納されている記憶部２または３の記憶部アドレスとＬＢＡを取得し、取得した記憶部アドレスとＬＢＡをサーバ６００１に送信する。サーバ６００１は、受信した記憶部アドレスおよびＬＢＡを指定したパケットをストレージネットワーク１０００に送信することで記憶部２または３に読み出し命令を送信し、記憶部から読み出しデータを受信し、クライアント１００２に読み出しデータを送信する。
【０３００】
（論理ドライブへの書き込み）
サーバ６００１は、たとえば、クライアント１００２からのファイルのデータ書き込み要求とファイルＩＤを受信すると、メタデータサーバ６００３にファイルＩＤを送信する。メタデータサーバ６００３は、論理ドライブステータステーブル４５０またはステータステーブル６５０から論理ドライブ４のステータスを判定し、論理ドライブ４が通常状態であることを認識すると、主メモリ６からメタデータ３００を読み出し、メタデータ３００を参照してデータ書き込み用のＬＢＡを割り当てる。メタデータサーバ６００３は、ＬＢＡおよび記憶部２の記憶部アドレスとＬＢＡをサーバ６００１に送信する。サーバ６００１は、受信した記憶部アドレスおよびＬＢＡを指定したパケットをストレージネットワーク１０００に送信することで記憶部２に書き込み命令を送信し、記憶部２に書き込みデータを記憶する。制御部２００はメタデータ３００を書き換え、書き込みデータのＬＢＡおよびセクタカウントを、記憶部２および書き込みファイルＩＤを紐付ける。
【０３０１】
制御部２００は、論理ドライブ４がデータ引き継ぎ状態であることを認識すると、主メモリ６からメタデータ３００を読み出し、メタデータ３００を参照してデータ書き込み用のＬＢＡを割り当てる。メタデータサーバ６００３は、ＬＢＡおよび記憶部３の記憶部アドレスとＬＢＡをサーバ６００１に送信する。サーバ６００１は、受信した記憶部３の記憶部アドレスおよびＬＢＡを指定したパケットをストレージネットワーク１０００に送信することで記憶部２に書き込み命令を送信し、記憶部２に書き込みデータを記憶する。制御部２００はメタデータ３００を書き換え、書き込みデータのＬＢＡおよびセクタカウントの記憶部２への紐付けを無効化すると共に、書き込みデータのＬＢＡおよびセクタカウントを、記憶部３および書き込みファイルＩＤを紐付けることで、記憶部３への書き込みを利用して記憶部２から記憶部３へのデータ引き継ぎを実現する。
【０３０２】
（バックグラウンドバックアップ）
制御部２００は、論理ドライブステータステーブル４５０において論理ドライブ４がデータ引き継ぎ状態のステータスである場合、クライアント６００２による論理ドライブ４へのアクセスがほとんど発生しない時に（アイドル時に）、データ引き継ぎ元の記憶部２からデータ引き継ぎ先の記憶部３にバックグラウンドバックアップを行うようにしてもよい。制御部２００は主メモリ６からメタデータ３００を読み出し、記憶部２に紐付けられているファイルＩＤを検索し、記憶部２に紐付けられているファイルが存在すれば、サーバ６００１を介して当該ファイルのＬＢＡに対して記憶部２に読み出し命令を送信し、記憶部３の当該ＬＢＡに対して書き込み命令と読み出されたデータを送信して、書き込みを行い、主メモリ６上のメタデータ３００を書き換えて、当該ファイルＩＤと記憶部２の間の紐付けを無効化すると共に、当該ファイルＩＤを記憶部３に紐付ける。

【０３０３】
制御部２００のバックグラウンドバックアップ動作としてサーバフリーバックアップを採用してもよく、その場合たとえば拡張コピーコマンドを用いてもよい。拡張コピーコマンドとして、たとえばSCSI Primary Commands-４（SPC-４），INCITS Ｔ１０/１７３１-D，Revision ３６e（http://www.t１０.org/）に記述されている８３h EXTENDED COPYコマンドを用いてもよい。制御部２００はバックアップ対象ＬＢＡと記憶部３のアドレスとを含む拡張コピーコマンドを記憶部２に送信すると、記憶部２が該ＬＢＡからデータを読み出し、記憶部２が該読み出しデータを記憶部３に送信し、記憶部３は受信データを該ＬＢＡに書き込む。
【０３０４】
（データ引き継ぎ完了時）
論理ドライブステータステーブル４５０において、論理ドライブ４のステータスが「データ引き継ぎ状態」の時、制御部２００は、主メモリ６上のメタデータ３００を定期的に読み出し、記憶部２に紐付けられている引き継ぎ対象ファイルＩＤが存在するか否かを定期的にチェックする。たとえば、制御部２００は、論理ドライブ４に格納されている全ファイルのファイルＩＤのうち、記憶部２に紐付けられている引き継ぎ対象ファイルＩＤが存在するか否かを定期的にチェックする。存在しない場合には、制御部２００はデータ引き継ぎ先である記憶部３のステータス５１０をディスクステータス＝１（通常状態）に書き換え、データ引き継ぎ元である記憶部２のステータス５１０をディスクステータス＝５（廃棄対象状態）に書き換える。制御部２００は、記憶部２を論理ドライブ４から切り離し、記憶部３を論理ドライブ４として認識し、論理ドライブステータステーブル４５０またはステータステーブル６５０内の論理ドライブ４のステータスを「データ引き継ぎ状態」から「通常状態」に書き換える。
【０３０５】
このように、データ引き継ぎ状態においては、論理ドライブ４へのデータ書き込みに際して、データ引き継ぎ元の記憶部２に対して書き込み命令を送信せず、データ引き継ぎ先の記憶部３に対して書き込み命令を送信する。論理ドライブ４からのデータ読み出しは、記憶部２，３の何れかから実行される。論理ドライブ４においては、クライアント６００２から書き込み要求を受信するごとに、記憶部２に格納される有効なデータが記憶部３に徐々に引き継がれていくことになり、新規データ書き込み動作がデータ引き継ぎ動作を兼用することになる。
【０３０６】
（第５の実施形態）
第５の実施形態では、複数のデータセンタおよびデータセンタ間を接続する長距離ネットワークからなる情報処理システム１に本発明を適用した場合について説明する。本実施形態においては、情報処理システム１は図７６に示すように構成される。本実施形態では、情報処理システム１は、データ引き継ぎ元のデータセンタ３００２、データ引き継ぎ先のデータセンタ３００３、その他のデータセンタであるデータセンタ３００５、サーバ３００６、データセンタ管理サーバ３００７、これらを接続する長距離ネットワーク３０００からなる。本実施形態では、データセンタ管理サーバ３００７はデータセンタ３００２を論理単位としての論理データセンタ３００４として認識し、データセンタ３００２が信頼性劣化するとデータセンタ３００２とデータセンタ３００３とを論理データセンタ３００４として認識する。
【０３０７】
データセンタ３００２、３００３は、ステータス格納領域５１０を具備する。記憶部３００２、３００３には、第１の実施形態のデータ引き継ぎ対象記憶部アドレス格納領域２００７、３００７（図３２参照）に対応するデータ引き継ぎ対象データセンタＩＤ格納領域２００７、３００７を具備させてもよいし、第２の実施形態の論理ドライブＩＤ格納領域５２０に対応する論理データセンタＩＤ格納領域５２０を具備させてもよい。ステータス格納領域５１０は、前述と同様、対応する記憶部がそれぞれ

０：初期データセンタ状態
１：通常状態
２：信頼性劣化状態
３：データ引き継ぎ元状態（保護状態）
４：データ引き継ぎ先状態
５：廃棄対象状態
の状態であることを示す。

【０３０８】

データセンタ管理サーバ３００７は、第１の実施形態と同様のＯＳ１００と制御部２００とメタデータ３００と論理データセンタステータステーブル４５０を格納する主メモリ６と、ＯＳ１００や制御部２００を実行する演算装置であるＣＰＵ５を具備し、第１の実施形態の情報処理装置１１１のＯＳ１００や制御部２００に相当する機能を担う。メタデータサーバ６００３の主メモリ６に格納されるメタデータ３００はたとえば図３３に示したような構造であり、主メモリ６上のメタデータやメタデータのジャーナルは、データセンタ管理サーバ３００７内の不揮発性記憶部や、長距離ネットワーク３０００内のデータセンタ３００２，３００３、３００５などに退避保存される。また、論理データセンタステータステーブル４５０は、例えば図４１に示したような構造であり、論理データセンタ３００４が通常状態かデータ引き継ぎ状態であることを示す。データセンタ管理サーバ３００７は、たとえばサーバ３００６からファイルＩＤを受信すると、メタデータ領域３００からファイルＩＤを検索し、該ファイルＩＤに対応する論理データセンタＩＤや、データセンタＩＤやＬＢＡやセクタカウントなどを特定する。ステータステーブル５１０は、データセンタ３００２、３００３のステータスを記憶する。

【０３０９】

なお、論理データセンタステータステーブル４５０を図４１に示した構造とする他に、データセンタ管理サーバ３００７に図７５に示したようなステータステーブル６５０を採用するようにしてもよい。このステータステーブル６５０においては、論理データセンタＩＤと、データセンタＩＤと、論理データセンタステータスと、データセンタステータスを管理するようにしており、この場合は、記憶部２，３でのステータス格納領域５１０は不要となる。制御部２００は、メタデータ領域３００の代わりに、第２の実施例のようにデータ引き継ぎ履歴格納領域を用いてデータセンタ３００２からデータ３００３へのデータ引き継ぎを管理するようにしてもよい。

【０３１０】

（寿命到達処理）

制御部２００は起動されると、データセンタ３００２についての前述した信頼性情報を監視する。制御部２００は、例えば図２４に示すように、一定時間おき（たとえば１分おき）または一定処理ごとにデータセンタ３００２から統計情報６５を取得する。制御部２００は、第１の実施形態と同様、取得した統計情報６５に基づいて、データセンタ３００２が寿命に到達したか否かを判定し、データセンタ３００２が寿命に到達したと判定された場合、後述する寿命到達時処理を実行する。

【０３１１】

（寿命到達時処理）

制御部２００は、接続されているデータセンタ３００２が寿命到達するかあるいは寿命到達直前になって寿命到達時処理が開始されると、データセンタ３００２のステータス５１０をデータセンタステータス＝２（信頼性劣化状態）に変更した後、新しいデータセンタの接続を促すメッセージをデータセンタ管理サーバ３００７のディスプレイなどに表示する。新たなデータセンタ３００３が接続されると、制御部２００は、ステータステーブル４５０のデータセンタ３００３のステータス５１０をデータセンタステータス＝４（データ引き継ぎ先状態）に書き換え、さらにデータセンタ３００２のステータス５１０をデータセンタステータス＝３（保護状態）に書き換える。そして、制御部２００は、データセンタ３００２＋データセンタ３００３を１つの論理データセンタ３００４としてＯＳ１

００に認識させる。制御部２００は、主メモリ６上の論理データセンタステータステーブル４５０を、ステータスが「通常状態」から「データ引き継ぎ状態」になるように書き換える。

【０３１２】

（データ引き継ぎ中での論理データセンタからの読み出し）

データセンタ管理サーバ３００７は、サーバ３００６からの読み出し要求とファイルＩＤを受信すると、メタデータ３００からファイルＩＤに対応する論理データセンタ３００４を特定し、さらに主メモリ６から論理データセンタステータステーブル４５０を読み出し、特定した論理データセンタ３００４のステータスがデータ引き継ぎ状態であることを認識する。データセンタ管理サーバ３００７は、メタデータ３００からファイルＩＤで指定されたファイルが格納されているデータセンタ３００２または３００３のデータセンタアドレスとＬＢＡを取得し、取得したデータセンタアドレスとＬＢＡをサーバ３００６に送信する。サーバ３００１は、受信したデータセンタアドレスおよびＬＢＡを指定したパケットをネットワーク３０００に送信することでデータセンタ３００２または３００３に読み出し命令を送信し、データセンタ３００２または３００３から読み出しデータを受信する。

【０３１３】

（論理データセンタへの書き込み）

データセンタ管理サーバ３００７は、サーバ３００６からの書き込み要求とファイルＩＤを受信すると、論理データセンタステータステーブル４５０から論理データセンタ３００４のステータスを判定し、論理データセンタ３００４が通常状態であることを認識すると、主メモリ６からメタデータ３００を読み出し、メタデータ３００を参照してデータ書き込み用のＬＢＡを割り当てる。データセンタ管理サーバ３００７は、ＬＢＡおよびデータセンタ３００２のデータセンタＩＤとＬＢＡをサーバ３００６に送信する。サーバ３００６は、受信したデータセンタＩＤおよびＬＢＡを指定したパケットをネットワーク３０００に送信することでデータセンタ３００２に書き込み命令を送信し、データセンタ３００２に書き込みデータを記憶する。

【０３１４】

データセンタ管理サーバ３００７は、論理データセンタ３００４がデータ引き継ぎ状態であることを認識すると、主メモリ６からメタデータ３００を読み出し、メタデータ３００を参照してデータ書き込み用のＬＢＡを割り当てる。データセンタ管理サーバ３００７は、ＬＢＡおよびデータセンタ３００３のデータセンタＩＤとＬＢＡをサーバ３００６に送信する。サーバ３００６は、受信したデータセンタ３００３のデータセンタＩＤおよびＬＢＡを指定したパケットをネットワーク３０００に送信することでデータセンタ３００３に書き込み命令を送信し、データセンタ３００３に書き込みデータを記憶する。

【０３１５】

（バックグラウンドバックアップ）

論理データセンタステータステーブル４５０において、論理データセンタ４がデータ引き継ぎ中のステータスである場合、サーバ３００６による論理データセンタ４へのアクセスがほとんど発生しない時に（アイドル時に）、データ引き継ぎ元のデータセンタ３００２からデータ引き継ぎ先のデータセンタ３００３にバックグラウンドバックアップを行うようにしてもよい。制御部２００は主メモリ６からメタデータ３００を読み出し、データセンタ３００２に紐付けられているファイルＩＤを検索し、データセンタ３００２に紐付けられているファイルが存在すれば、サーバ３００６を介して当該ファイルのＬＢＡに対してデータセンタ３００２に読み出し命令を送信してデータ３００２から読み出しデータを受信し、データセンタ３００３の当該ＬＢＡに対して書き込み命令と読み出されたデータを送信して、書き込みを行い、主メモリ６上のメタデータ３００を書き換えて当該ファイルＩＤをデータセンタ３００３に紐付ける。

【０３１６】

制御部２００のバックグラウンドバックアップ動作としてサーバフリーバックアップを

採用してもよく、その場合たとえば拡張コピーコマンドを用いてもよい。拡張コピーコマンドとして、たとえばSCSI Primary Commands-４（SPC-４）, INCITS Ｔ１０/１７３１-D, Revision ３６e (http://www.t１０.org/)に記述されている８３h EXTENDED COPYコマンドを用いてもよい。制御部２００はバックアップ対象ＬＢＡとのデータセンタ３００３のＩＤとを含む拡張コピーコマンドをデータセンタ３００２に送信すると、データセンタ３００２が該ＬＢＡからデータを読み出し、データセンタ３００２が該読み出しデータをデータセンタ３００３に送信し、データセンタ３００３は受信データを該ＬＢＡに書き込む。

【０３１７】

（データ引き継ぎ完了時）

論理データセンタステータステーブル４５０において、論理データセンタ４のステータスが「データ引き継ぎ状態」の時、制御部２００は、主メモリ６上のメタデータ３００を定期的に読み出し、データセンタ３００２に紐付けられている引き継ぎ対象ファイルＩＤが存在するか否かを定期的にチェックする。たとえば、制御部２００は、論理データセンタ４に格納されている全ファイルのファイルＩＤのうち、データセンタ３００２に紐付けられている引き継ぎ対象ファイルＩＤが存在するか否かを定期的にチェックする。存在しない場合には、制御部２００はデータ引き継ぎ先であるデータセンタ３００３のステータス５１０をデータセンタステータス＝１（通常状態）に書き換え、データ引き継ぎ元であるデータセンタ３００２のステータス５１０をデータセンタステータス＝５（廃棄対象状態）に書き換える。制御部２００は、データセンタ３００２を論理データセンタ４から切り離し、データセンタ３００３を論理データセンタ４として認識し、論理データセンタステータステーブル４５０で論理データセンタ３００４のステータスを「データ引き継ぎ状態」から「通常状態」に書き換える。

【０３１８】

このように、データ引き継ぎ中には、論理データセンタ３００４へのデータ書き込みに際し、データ引き継ぎ元のデータセンタ３００２に対して書き込み命令を送信せず、データ引き継ぎ先のデータセンタ３００３に対して書き込み命令を送信する。論理データセンタ３００４からのデータ読み出しは、データセンタ３００２、３００３の何れかから実行される。論理データセンタ４においては、サーバ３００６から書き込み要求を受信するごとに、データセンタ３００２に格納される有効なデータがデータセンタ３００３に徐々に引き継がれていくことになり、新規データ書き込みがデータ引き継ぎ動作を兼用することになる。

【０３１９】

（第６の実施形態）

（中継装置）

第１及び第２の実施形態では、制御部２００が、記憶部２を寿命到達したと判定した場合、寿命到達目前であると判定した場合、故障目前であると判定した場合、信頼性劣化したと判定した場合に、記憶部２に対する書き込みを行わないよう書き込み命令を制御することで、データ書き込みによる記憶部２のさらなる信頼性劣化を抑制するとともに、新規書き込みデータの損失を防止することを可能にしている。記憶部２に対するデータ書き込みを安全に制限するという観点では、情報処理装置やクライアントやサーバが記憶部２に対する書き込みを自発的に行わないようにすることが望ましい。本実施形態では、情報処理システムは中継装置５０００を具備し、中継装置５０００内の制御部２００が記憶部２の統計情報などの信頼性情報を監視し、中継装置５０００が記憶部２を寿命到達したと判定した場合、寿命到達目前であると判定した場合、故障目前であると判定した場合、信頼性劣化したと判定した場合に、記憶部２が読み出し専用デバイスであるという情報を情報処理装置１１１やクライアントやサーバに通知することにより、情報処理装置１１１やクライアントやサーバが記憶部２に対する書き込みを自発的に行わないようにさせている。本実施形態は単独で実施することが可能であるし、一方、第１の実施形態や第２の実施形態等と組み合わせることで、信頼性劣化した記憶部２に対する書き込みをより堅牢に抑制

することが可能になる。たとえば、本実施形態を第１の実施形態や第２の実施形態等を組み合わせることで、中継装置は記憶部２が読み出しと書き込みのうち読み出しのみをサポートする記憶部であるという記憶部識別情報を送信し、ＣＰＵ５や主メモリ６に格納された制御部２００は、記憶部２を読み出し専用記憶部として認識し、記憶部２を保護状態（データ引き継ぎ元状態）として認識する。

【０３２０】

図７７は、デスクトップパソコンやノートパソコンなどの情報処理システム１に中継装置が搭載された例である。中継装置５０００は、情報処理装置１内に搭載されてもよいし、情報処理装置１外に搭載されてもよい。中継装置５０００は、インタフェース１９を介して記憶部２と接続され、インタフェース５００１、チップセット７を経由してＣＰＵ５と接続される。なお、中継装置５０００は、チップセット７を経由せずＣＰＵ５に直接接続されてもよい。また、中継装置５０００は、チップセット７に含まれていてもよい。

【０３２１】

中継装置５０００は制御部２００を具備する。制御部２００は、図７７に示すように、全てが中継装置５０００に含まれても良いし、図７８に示すように、一部が中継装置５０００に含まれ一部が主メモリ６に含まれても良いし、一部が中継装置５０００に含まれ一部がＲＯＭ１１など情報処理装置１１１内のその他のメモリ部分に含まれても良い。制御部２００は、ファームウェアやソフトウェアの形態で実装されてもよいし、ハードウェアの形態で実装されてもよい。

【０３２２】

インタフェース１９およびインタフェース５００１として、ＳＡＴＡ（Serial Advanced Technology Attachment）、ＰＣＩExpress (Peripheral Component Interconnect Express, PCIe)、ＵＳＢ（Universal Serial Bus）、ＳＡＳ（Serial Attached SCSI）、Thunderbolt（登録商標）、イーサネット（登録商標）、ファイバーチャネルなどが使用可能である。インタフェース１９とインタフェース５００１は同じ規格のインタフェースであってもよいし、異なる規格のインタフェースであってもよい。本実施形態では、インタフェース１９とインタフェース５０００１がＳＡＴＡインタフェースである場合について説明する。

【０３２３】

制御部２００は、第１の実施形態の図２４や図２９と同様にして、記憶部２が寿命に到達しているか否か、寿命到達目前であるか否か、あるいは故障目前であるか否かを判定し、寿命に到達しているか場合、寿命到達目前である場合、あるいは故障目前である場合に寿命到達時処理として、図７９に示すように、通常状態から信頼性劣化モードに遷移する（ステップＳ８００）。通常状態と信頼性劣化モードは制御部２００のモードであり、記憶部２が正常状態であるとき制御部２００は通常状態で動作し、記憶部２を寿命到達したと判定した場合、寿命到達目前であると判定した場合、故障目前であると判定した場合、信頼性劣化したと判定した場合に、制御部２００は信頼性劣化モードで動作する。寿命到達判定は、第１の実施形態と同様に、図２４のような一定時間ごと、または一定処理数おき、または一定データ送受信おき、または図２９や図３０のように記憶部からのコマンド応答がエラーであった場合に実施される。

【０３２４】

制御部２００は、インタフェース５００１を経由してＣＰＵ５から受信した命令・データを、インタフェース１９を経由して記憶部２に送信する。また、制御部２００は、インタフェース１９を経由して記憶部２から受信した応答・データを、インタフェース５００１を経由してＣＰＵ５に送信する。インタフェース５００１とインタフェース１９のプロトコルが異なる場合には、制御部２００がプロトコル変換を行った後に、変換後の命令・応答・データをＣＰＵ５や記憶部２に送信する。制御部２００は記憶部２の記憶部情報をＣＰＵ５に送信する時、制御部２００が通常状態であるか信頼性劣化モードであるかによって、ＣＰＵ５に送信する記憶部２の記憶部情報を切り替える。すなわち、制御部２００は、通常状態では記憶部２が読み出しおよび書き込み可能な記憶部であるという記憶部情

報をＣＰＵ５に送信し、信頼性劣化モードでは記憶部２が読み出しと書き込みのうち読み出しのみ可能な記憶部であるという記憶部情報をＣＰＵ５に送信する。

【０３２５】

制御部２００は、ＣＰＵ５から記憶部２の記憶部情報要求を受信すると、記憶部情報要求に対する応答として、通常状態では記憶部２が読み出しおよび書き込み可能な記憶部であるという記憶部情報をＣＰＵ５に送信し、信頼性劣化モードでは記憶部２が読み出しと書き込みのうち読み出しのみ可能な記憶部であるという記憶部情報をＣＰＵ５に送信する。記憶部情報要求としては、たとえば、ACS-3記載のECh IDENTIFY DEVICEコマンドや、SCSI Primary Commands-4 (SPC-4)記載のA3h REPORT IDENTIFYING INFORMATIONコマンドや、INFORMATION TECHNOLOGY Multi-Media Commands-6 (MMC-6)(http://www.t１０.org/)に記載の46h GET CONFIGURATIONコマンドや、NVM Express Revision 1.1記載の06h Identifyコマンドなどが用いられる。

【０３２６】

図８０は、制御部２００がＣＰＵ５から記憶部２の記憶部情報要求を受信した場合の処理手順の例を示すものである。制御部２００がＣＰＵ５から記憶部２の記憶部情報要求を受信すると（ステップＳ８０１）、制御部２００は、記憶部２が信頼性劣化モードであるか否かを、制御部２００自身が通常状態にあるか信頼性劣化モードにあるかに基づいて判定する（ステップＳ８０２）。通常状態である場合、制御部２００は、記憶部２に記憶部情報要求を送信し（ステップＳ８０３）、記憶部２から記憶部情報を受信し、ＣＰＵ５に受信した記憶部情報を送信する（ステップＳ８０４）。信頼性劣化モードである場合、制御部２００は、記憶部２に記憶部情報要求を送信し（ステップＳ８０５）、記憶部２から記憶部情報を受信し、受信データを書き換えて記憶部２がＲＯＭデバイスであるかのように記憶部情報を変更し、ＣＰＵ５に変更した記憶部情報を送信する（ステップＳ８０６）。なお、制御部２００は、ステップＳ８０１のＣＰＵ５からの記憶部情報要求を受信すること無しに、制御部２００が自発的にＳ８０２～Ｓ８０６の処理を実施してもよい。また、制御部２００は、記憶部２に対する記憶部情報要求送信をステップＳ８０１とステップＳ８０２の間に行い、Ｓ８０３およびＳ８０５での記憶部２に対する記憶部情報要求送信を行わないようにしてもよい。

【０３２７】

図８１に示すように、制御部２００は、信頼性劣化モードのとき、記憶部２に記憶部情報要求を送信することなく記憶部２がＲＯＭデバイスであるかのような記憶部情報を生成し、ＣＰＵ５に送信するようにしてもよい。すなわち、制御部２００は、記憶部２の記憶部情報要求をＣＰＵ５から受信すると（ステップＳ８１０）、記憶部２が信頼性劣化モードであるか否かを、制御部２００自身が通常状態にあるか信頼性劣化モードにあるかに基づいて判定する（ステップＳ８１１）。通常状態である場合、制御部２００は、記憶部２に記憶部情報要求を送信し（ステップＳ８１２）、ＣＰＵ５に受信した記憶部情報を送信する（ステップＳ８１３）。信頼性劣化モードである場合、制御部２００は、記憶部２に記憶部情報要求を送信することなく、記憶部２がＲＯＭデバイスであるとＣＰＵ５に見せかける記憶部情報を生成し、ＣＰＵ５に生成した記憶部情報を送信する（ステップＳ８１４）。

【０３２８】

通常状態での動作時に、記憶部２が読み出しおよび書き込み可能な記憶部であるという記憶部情報として、制御部２００はたとえば記憶部２がＡＴＡデバイスであることをＣＰＵ５に明示的に通知することが望ましい。たとえば、ATA/ATAPI Command Set-3 (ACS-3)で記述されているDevice Signatureにおいて、ＬＢＡ(7:0)を01hに、ＬＢＡ(15:8)を00hに、ＬＢＡ(23:16)を00hとしてＣＰＵ５に出力することで、記憶部２がＡＴＡデバイスであることをＣＰＵ５に通知することができる。

【０３２９】

信頼性劣化モードでの動作時に、記憶部２が読み出しと書き込みのうち読み出しのみ可能な記憶部であるという記憶部情報として、制御部２００はたとえば記憶部２がＡＴＡＰ

Ｉデバイスであることを ＣＰＵ５に明示的に通知する。たとえば、ACS-3で記述されているDevice Signatureにおいて、ＬＢＡ(7:0)を01hに、ＬＢＡ(15:8)を14hに、ＬＢＡ(23:16)をEBhとしてＣＰＵ５に出力することで、記憶部２がＡＴＡＰＩデバイスであることをＣＰＵ５に通知することができる。さらに、信頼性劣化モードでの動作時に、記憶部２が読み出しと書き込みのうち読み出しのみ可能な記憶部であるという記憶部情報として、制御部２００は、たとえばＣＰＵ５からINCITS Multi-Media Commands-6 (MMC-6)記載の46h GET CONFIGURATIONコマンドを受信したとき、Random Writable(Feature Number＝0020h)、Incremental Streaming Writable (Feature Number＝0021h)、Write Once(Feature Number＝0025h)などのFeaturesにおいて、書き込み系機能は全て非サポートであることをＣＰＵに返信する。これにより、ＯＳ１００がWindows（登録商標）などである場合でも、記憶部２を読み出し可能なデバイスとしてＯＳ１００に認識させることが可能となり、ＯＳ１００や、ＯＳ１００より上位のレイヤのアプリケーションプログラム４００に対しては記憶部２があたかもリードオンリーデバイスであるかのように見え、ＯＳ１００やアプリケーションプログラム４００が記憶部２に対して誤って書き込み命令を送信してしまうことを防止できる。

【０３３０】

あるいは、信頼性劣化モードでの動作時に、記憶部２が読み出しと書き込みのうち読み出しのみ可能な記憶部であるという記憶部情報として、制御部２００はたとえば記憶部２がＡＴＡデバイスであることをＣＰＵ５に明示的に通知し、かつＣＰＵ５からACS-3記載のコマンドECh IDENTIFY DEVICEを受信したとき、書き込み系機能は全て非サポートであるという情報をＣＰＵ５に通知してもよい。

【０３３１】

記憶部２が読み出しと書き込みのうち読み出しのみ可能な記憶部であるか否かの通知方法はこれ以外の様々な形態をとってもよい。ＣＰＵ５は、記憶部２が読み出しと書き込みのうち読み出しのみ可能な記憶部であるという情報を受信することにより、ＯＳ１００は、記憶部２に適用されるドライバソフトウェアとして読み出し専用の例えばＡＴＡＰＩの読み出し専用記憶部用ドライバを適用し、ＯＳ１００上で記憶部２はたとえばＣＤ－ＲＯＭやＤＶＤ－ＲＯＭやＢＤ－ＲＯＭなどのような読み出し専用記憶部として認識される。たとえば、図８２や、図８３に示すように、情報処理装置１１１のユーザは、ＯＳ１００により、ディスプレイ装置９経由で記憶部２がたとえばＣＤ-ＲＯＭやＤＶＤ-ＲＯＭやＢＤ－ＲＯＭなどのような読み出し専用デバイスであるとアイコンのグラフィック等により明示的に通知されることが望ましい。さらに望ましくは、ユーザが記憶部２にファイル等を書き込もうと試みた場合、図８４のようにＯＳ１００がディスプレイ装置９を経由して記憶部２がライトプロテクトされていることを通知することが望ましい。

【０３３２】

ＯＳ１００が記憶部２に読み出し専用記憶部用ドライバを適用した場合など、ＣＰＵ５やＯＳ１００が記憶部２を読み出しと書き込みのうち読み出しのみ可能な記憶部として認識した場合であっても、ＣＰＵ５は記憶部２に対して読み出し命令を送信可能である。ＣＰＵ５が記憶部２に読み出し命令を送信した時、制御部２００は記憶部２に読み出し命令を送信し、制御部２００は記憶部２から読み出しデータを受信し、制御部２００は読み出しデータをＣＰＵ５に送信する。このようにして、ＣＰＵ５は、制御部２００が通常状態と信頼性劣化モードのどちらの場合であっても、記憶部２からデータを読み出すことができる。

【０３３３】

記憶部２のデータ破壊や故障によるユーザデータ損失を防止する観点では、信頼性劣化モードにおける制御部２００は、記憶部２に対する書き込みコマンドを一切送信しないよう構成されることが望ましい。しかし、制御部２００は、たとえばＯＳ１００のシステム情報など一部のデータを記憶部２に書き込む必要がある場合は、例外として記憶部２に当該データを書き込むことを許可してもよいが、当該書き込みデータはデータ量が記憶部２の容量に対して十分に小さいことが望ましい。さらに望ましくは、ユーザが誤って書き込

みコマンドを送信してしまい記憶部２に対してデータを書き込んでしまうことを防止するために、記憶部２に対する通常の書き込みコマンドをＣＰＵ５から受信しても記憶部２に書き込みコマンドを送信せず、記憶部２に対して例外的にデータを書き込む必要がある場合として、特殊なコマンドを用いた書き込みコマンドをＣＰＵ５から受信した場合に限り、記憶部２に対して書き込みコマンドを送信するようにすることが望ましい。たとえば、ＣＰＵ５の記憶部２に対する通常の書き込みコマンドとして、たとえばACS-3記載の35h WRITE DMA EXTや61h WRITE FPDMA QUEUEDのような書き込みコマンドが使われる場合、通常状態の制御部２００が35h WRITE DMA EXTコマンドや61h WRITE FPDMA QUEUEDコマンドを通常書き込みコマンドとしてＣＰＵ５から受信すると、通常状態の制御部２００は該コマンドを記憶部２に転送し、信頼性劣化モードの制御部２００が35h WRITE DMA EXTコマンドや61h WRITE FPDMA QUEUEDコマンドを通常書き込みコマンドとしてＣＰＵ５から受信すると、信頼性劣化モードの制御部２００は書き込みコマンドを記憶部２に送信しない。一方、信頼性劣化モードの制御部２００がたとえばINCITS ACS-3に記述されている30h Write Sectorsコマンドや3Fh Write Log ExtコマンドやSCT Command Transportやその他ベンダー独自のコマンドなどを特殊書き込みコマンドとしてＣＰＵ５から受信すると、信頼性劣化モードの制御部２００は該コマンドを記憶部２に転送する。

【０３３４】

以上、情報処理システム１がデスクトップパソコンやノートパソコンなどのコンピュータシステムである場合について説明したが、情報処理システム１はたとえば図８５のようなストレージネットワークを有する情報処理システムであってもよい。図８５では、インタフェース１９やインタフェース５００１としてのストレージネットワーク１０００が用いられ、中継装置５０００としてのネットワークスイッチ６００９が用いられる。

【０３３５】

ストレージネットワーク１０００として、例えばＳＡＡＮ (Storage Area Network)やＮＡＳ(Network Attached Storage)が用いられる。ＳＡＮには、たとえば、ＦＣ-ＳＡＮ(Fibre Channel Storage Area Network)やＩＰ-ＳＡＮ(Internet Protocol Storage Area Network)が用いられ、その上位プロトコルとして、たとえば、ＳＣＳＩ(Small Computer System Interface)が用いられる。たとえば、ストレージネットワーク１０００としてＩＰ-ＳＡＮが採用可能であり、その上位プロトコルとしてはｉＳＣＳＩ (Internet Small Computer System Interface)が採用可能である。

【０３３６】

ネットワークスイッチ６００９はストレージネットワーク１０００上においてクライアントやサーバや記憶部など複数のネットワーク機器間の接続を行うネットワークデバイスであり、ネットワーク機器からパケットを受信すると、受信パケットの宛先アドレスをもとに宛先となるネットワーク機器に受信パケットを送信する。

【０３３７】

ネットワークスイッチ６００９は制御部２００を具備する。制御部２００は、図８５に示すように全てがネットワークスイッチ６００９に含まれても良いし、一部がネットワークスイッチ６００９に含まれ一部がたとえばクライアント７０００Ａの主メモリ６に含まれても良いし、一部がネットワークスイッチ６００９に含まれ一部がたとえばクライアント７０００ＡのＲＯＭなどシステム１内のその他の部分に含まれても良い。制御部２００は、ファームウェアやソフトウェアの形態で実装されてもよいし、ハードウェアの形態で実装されてもよい。

【０３３８】

制御部２００は、第１の実施形態の図２４や図２９と同様にして、ストレージネットワーク１０００に接続された１乃至複数の記憶部が寿命に到達しているか否か、寿命到達目前であるか否か、あるいは故障目前であるか否かを判定し、寿命に到達しているか場合、寿命到達目前である場合、あるいは故障目前である場合に寿命到達時処理として、寿命到達時処理対象の記憶部のみを対象として通常状態から信頼性劣化モードに遷移する。通常状態と信頼性劣化モードは、ストレージネットワーク１０００に接続された１乃至複数の

記憶部それぞれに対応して存在する制御部２００のモードであり、たとえば記憶部２Ａが正常状態であるとき制御部２００の記憶部２Ａ用モードは通常状態で動作し、たとえば記憶部２Ａを寿命到達したと判定した場合、寿命到達目前であると判定した場合、故障目前であると判定した場合、信頼性劣化したと判定した場合に、制御部２００は記憶部２Ａ用モードとして信頼性劣化モードで動作する。制御部２００は、寿命到達時処理対象の記憶部を対象として通常状態から信頼性劣化モードに遷移した場合であっても、寿命到達時処理対象ではない記憶部に対しては通常状態で動作する。寿命到達判定は、第１の実施形態と同様に、図２４のような一定時間ごと、または一定処理数おき、または一定データ送受信おき、または図２９や図３０のように記憶部からのコマンド応答がエラーであった場合に実施される。

【０３３９】

本実施形態では例としてストレージネットワーク１０００に、２つのクライアント７０００Ａと７０００Ｂと、２つの記憶部２Ａと２Ｂが接続された場合について説明するが、システム１のネットワーク機器の構成は他の任意の構成を採用しても良い。また、クライアント７０００Ａの代わりにサーバ７０００Ａが用いられてもよいし、クライアント７０００Ｂの代わりにサーバ７０００Ｂが用いられてもよい。ネットワーク機器であるクライアント７０００Ａ、クライアント７０００Ｂ、記憶部２Ａ、記憶部２Ｂにはそれぞれストレージネットワーク上でアドレスが割り当てられる。アドレスとして、たとえばＩＰ(Internet Protcol)アドレスやＭＡＣ(Media Access Control)アドレスを使用することができる。たとえば、クライアント７０００ＡのＣＰＵ５が、記憶部２Ａのアドレスのみを指定する命令およびデータをストレージネットワーク１０００に送信すると、ネットワークスイッチ６００９がこの命令やデータを構成するパケット内のアドレスから記憶部２Ａを特定し、ネットワークスイッチ６００９はパケットを記憶部２Ａにのみ送信する。たとえば、記憶部２Ａが、クライアント７０００Ａのアドレスのみを指定する応答およびデータをストレージネットワーク１０００に送信すると、ネットワークスイッチ６００９が前記応答やデータを構成するパケット内のアドレスからクライアント７０００Ａを特定し、ネットワークスイッチ６００９はそのパケットをクライアント７０００Ａにのみ送信する。このアドレス指定は、単一のネットワーク機器のみでなく、複数のネットワーク機器を指定することもできる。

【０３４０】

制御部２００はたとえば記憶部２Ａの記憶部情報をたとえばクライアント７０００Ａに送信する時、制御部２００の記憶部２Ａ用モードが通常状態であるか信頼性劣化モードであるかによって、クライアント７０００Ａに送信する記憶部２Ａの記憶部情報を切り替える。すなわち、制御部２００は、制御部２００の記憶部２Ａ用モードが通常状態のとき記憶部２Ａが読み出しおよび書き込み可能な記憶部であるという記憶部情報をクライアント７０００Ａに送信し、制御部２００の記憶部２Ａ用モードが信頼性劣化モードのとき記憶部２Ａが読み出しと書き込みのうち読み出しのみ可能な記憶部であるという記憶部情報をクライアント７０００Ａに送信する。記憶部２Ａの記憶部情報は同様にしてクライアント７０００Ｂにも送信される。また、記憶部２Ｂの記憶部情報も同様にしてクライアント７０００Ａやクライアント７０００Ｂにも送信される。

【０３４１】

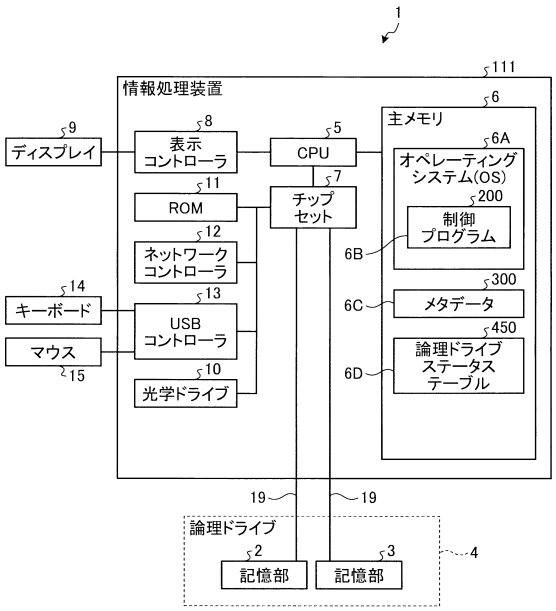
このように本実施形態では、制御部２００は、記憶部２を寿命到達したと判定した場合、寿命到達目前であると判定した場合、故障目前であると判定した場合、信頼性劣化したと判定した場合に、ＣＰＵ５に送信される記憶部２の記憶部情報を加工または生成することにより、記憶部２が読み出しと書き込みのうち読み出しのみ可能な記憶部であるとＣＰＵ５やＯＳ１００に認識させることで、記憶部２に対するデータ書き込みを防止することができる。また、記憶部２が読み出しと書き込みのうち読み出しのみ可能な記憶部であるという認識処理はＯＳ１００など下位のソフトウェアレイヤで行われるため、アプリケーションプログラム４００など上位のソフトウェアレイヤやユーザは記憶部２への特別な制御を必要としない。

【０３４２】
本発明のいくつかの実施形態を説明したが、これらの実施形態は、例として提示したものであり、発明の範囲を限定することは意図していない。これら新規な実施形態は、その他の様々な形態で実施されることが可能であり、発明の要旨を逸脱しない範囲で、種々の省略、置き換え、変更を行うことができる。これら実施形態やその変形は、発明の範囲や要旨に含まれるとともに、特許請求の範囲に記載された発明とその均等の範囲に含まれる。

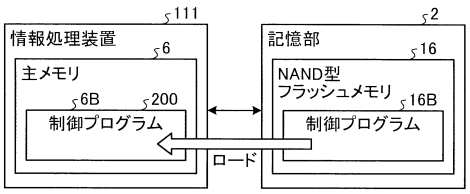
【符号の説明】

【０３４３】
１　情報処理システム、２，２Ａ-２Ｄ　記憶部（記憶装置）、３　記憶部（記憶装置）、４　論理ドライブ、６　主メモリ、１６ＮＡＮＤ型フラッシュメモリ、１９　インタフェース、１１１　情報処理装置、２００　制御プログラム（制御部）、３００　メタデータ、４００　アプリケーションプログラム、４５０　論理ドライブステータステーブル、５１０　ステータス記憶領域（ステータステーブル）、５２０　論理ドライブＩＤ記憶領域、５３０　スロット番号記憶領域、５５０　引き継ぎ履歴（引き継ぎ履歴記憶領域）、１０００　ストレージネットワーク、１００１　ネットワークスイッチ、１００３　ストレージアレイ装置。

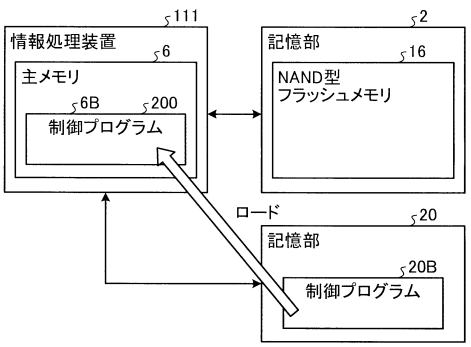
【図１】



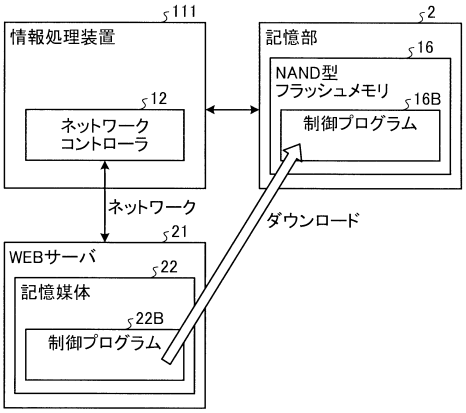
【図２】



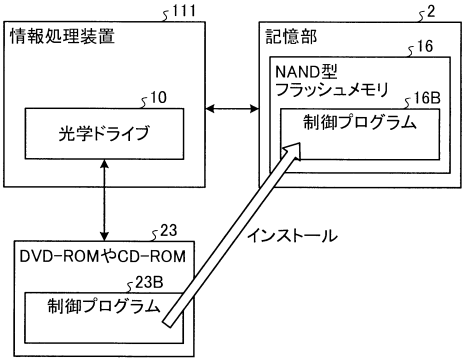
【図３】



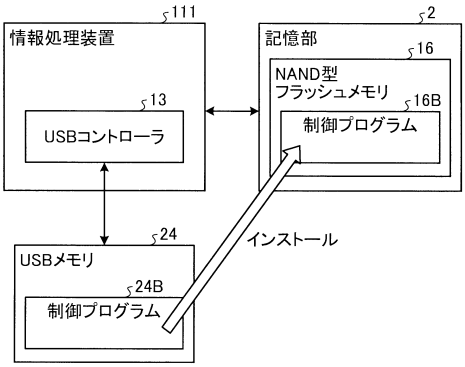
【図４】



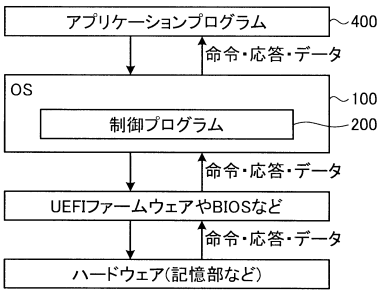
【図５】



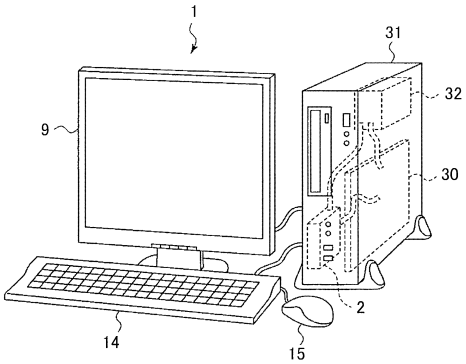
【図６】



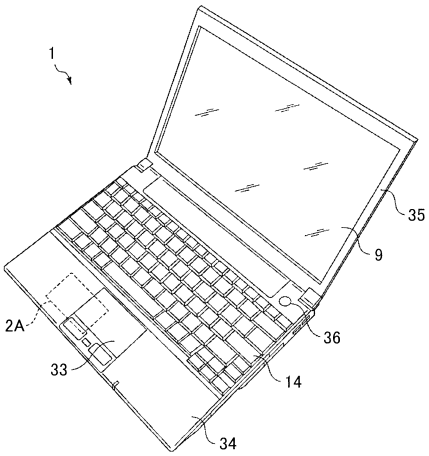
【図７】



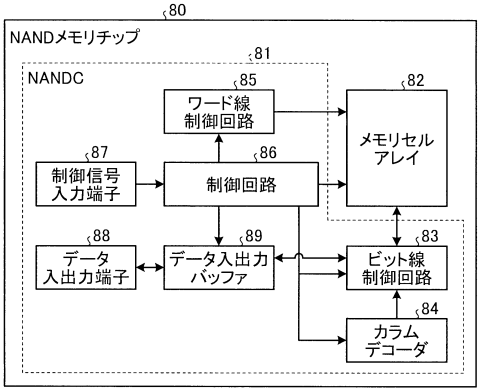
【図８】



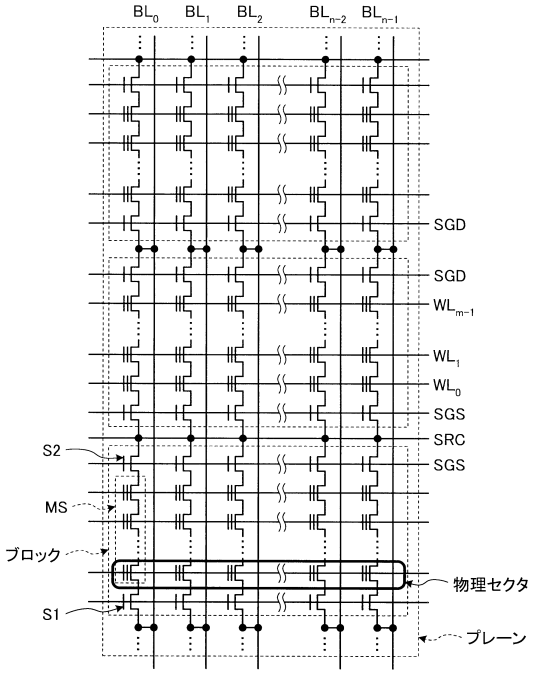
【図９】



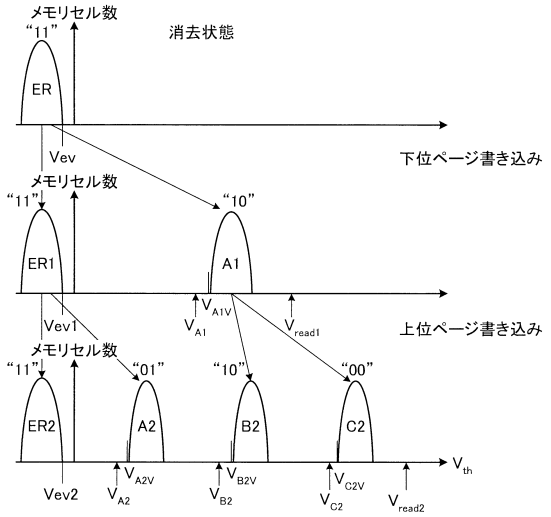
【図１０】



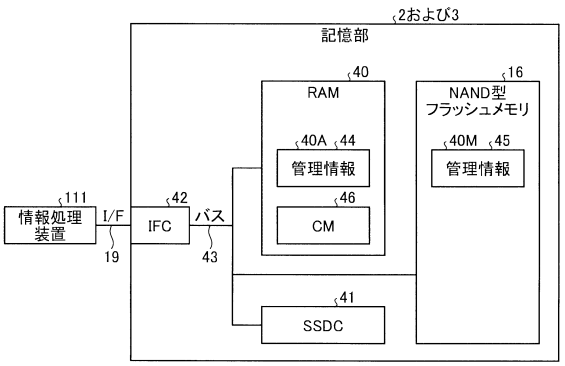
【図１１】



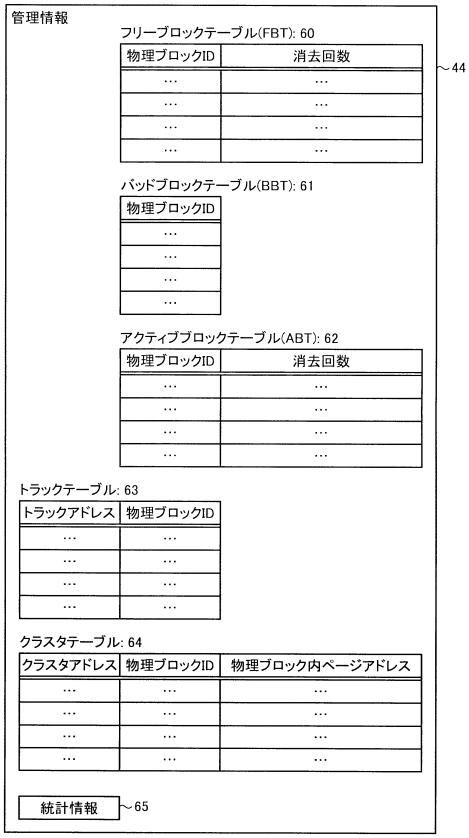
【図１２】



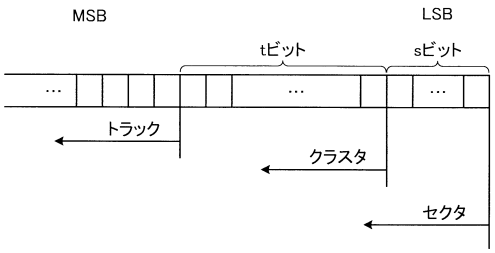
【図１３】



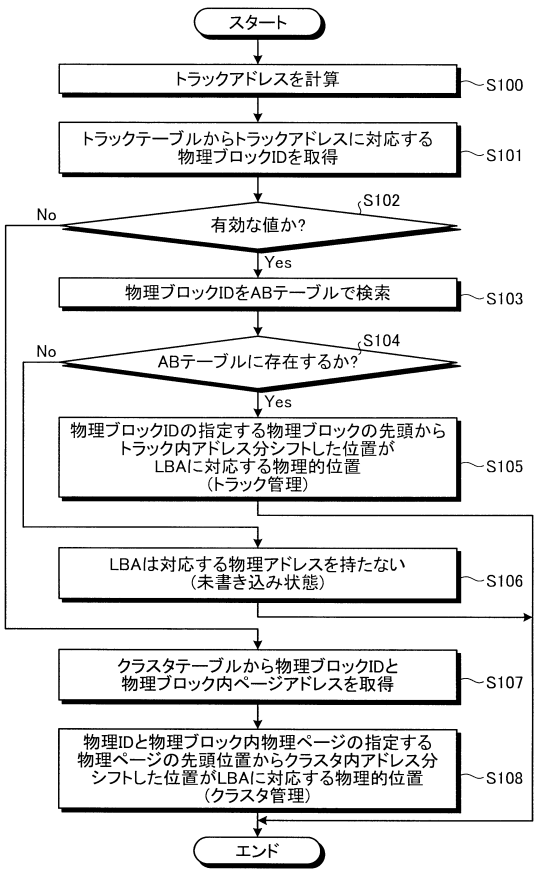
【図１４】



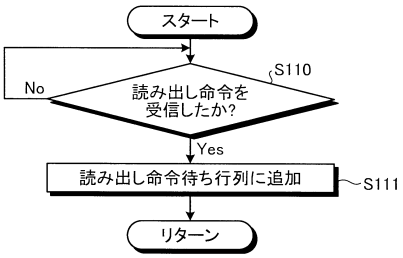
【図１５】



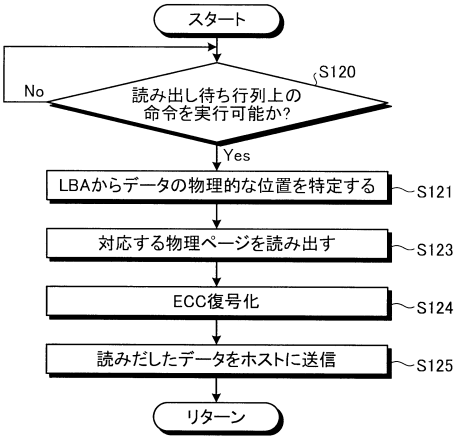
【図１６】



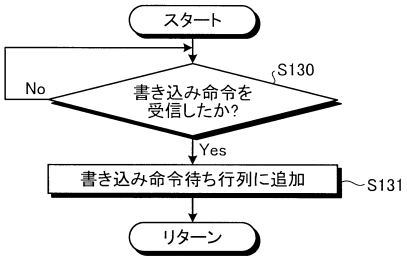
【図１７】



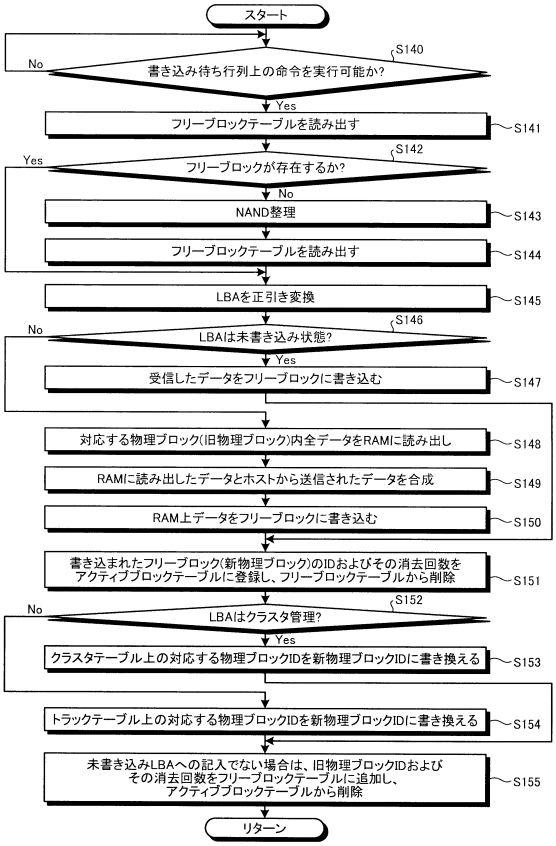
【図１８】



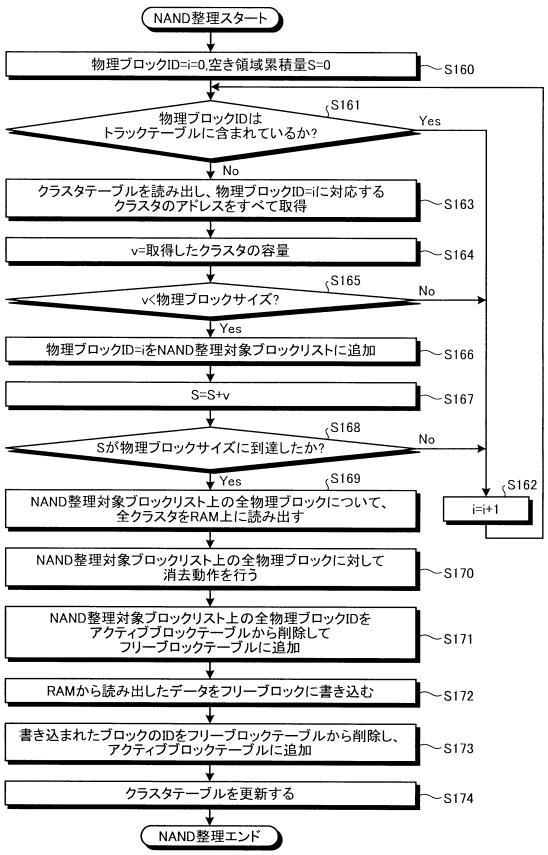
【図１９】



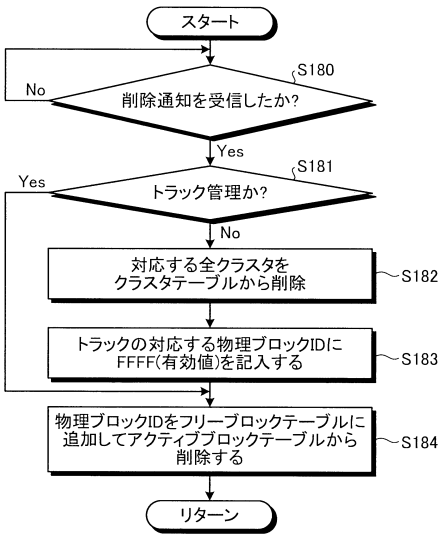
【図２０】



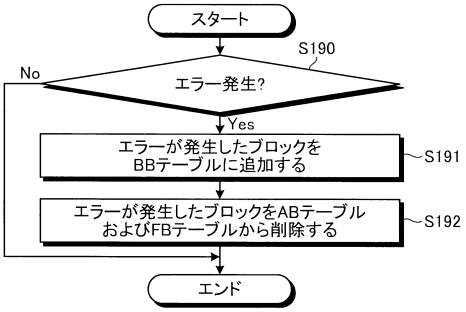
【図２１】



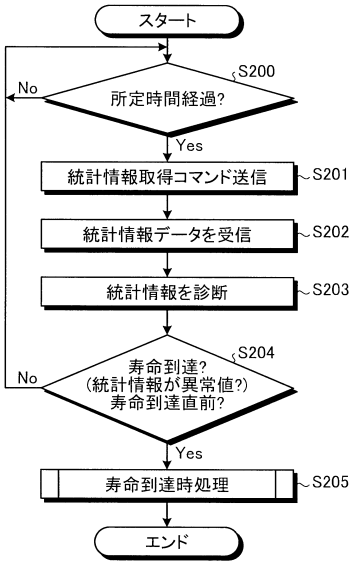
【図２２】



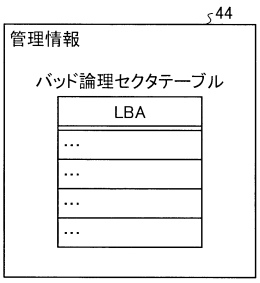
【図２３】



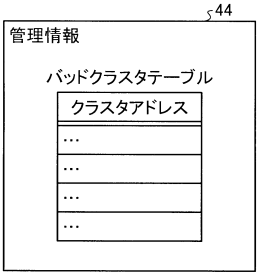
【図２４】



【図２５】



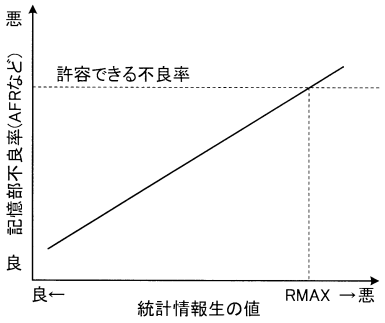
【図２６】



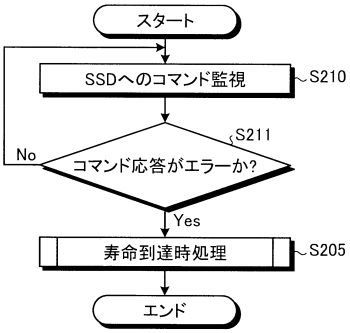
【図２７】

attribute ID	attribute名	Value	Threshold	Worst	Raw Data
…	…	…	…	…	…
…	バッドブロック数総計(統計情報X01)	…	…	…	…
…	バッド論理セクタ数総計(統計情報X02)	…	…	…	…
…	バッド論理セクタ数総計(統計情報X03)	…	…	…	…
…	消去回数総計(統計情報X03)	…	…	…	…
…	消去回数平均値(統計情報X04)	…	…	…	…
…	NANDメモリの書き込みエラー発生回数累積値(統計情報X05)	…	…	…	…
…	NANDメモリの消去エラー発生回数累積値(統計情報X06)	…	…	…	…
…	読み出し論理セクタ数総計(統計情報X07)	…	…	…	…
…	書き込み論理セクタ数総計(統計情報X08)	…	…	…	…
…	ECC訂正不能回数総計(統計情報X09)	…	…	…	…
…	リトライ読み出し発生回数(統計情報X10)	…	…	…	…
…	nビット～mビットECC訂正単位総計数(統計情報X11)	…	…	…	…
…	インターフェースI/OのデータCRCエラー発生回数(統計情報X12)	…	…	…	…
…	インターフェースI/Oの通信速度ダウンシフト回数(統計情報X13)	…	…	…	…
…	インターフェースI/Oのレーン数ダウンシフト回数(統計情報X14)	…	…	…	…
…	インターフェースI/Oのエラー発生回数(統計情報X15)	…	…	…	…
…	RAM40のエラー発生回数(統計情報X16)	…	…	…	…
…	記憶部20の使用時間総計(統計情報X17)	…	…	…	…
…	起動回数(統計情報X18)	…	…	…	…
…	不正電源断発生回数(統計情報X19)	…	…	…	…
…	温度が推奨動作温度の最高値を上回った時間累計(統計情報X20)	…	…	…	…
…	温度が推奨動作温度の最低値を下回った時間累計(統計情報X21)	…	…	…	…
…	コマンドの応答時間最大値(統計情報X22)	…	…	…	…
…	コマンドの応答時間平均値(統計情報X23)	…	…	…	…
…	NANDメモリの応答時間最大値(統計情報X24)	…	…	…	…
…	NANDの応答時間平均値(統計情報X25)	…	…	…	…
…	現在温度(統計情報X26)	…	…	…	…
…	最高温度(統計情報X27)	…	…	…	…
…	最低温度(統計情報X28)	…	…	…	…
…	管理情報冗長度(統計情報X29)	…	…	…	…
…	RAM40への書き込みデータ量合計(統計情報X30)	…	…	…	…
…	統計情報増加率(統計情報X31)	…	…	…	…
…	NAND整理失敗フラグ(統計情報X32)	…	…	…	…

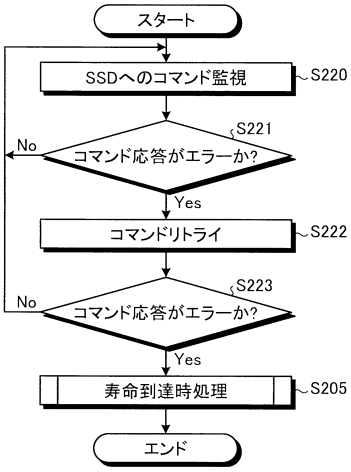
【図２８】



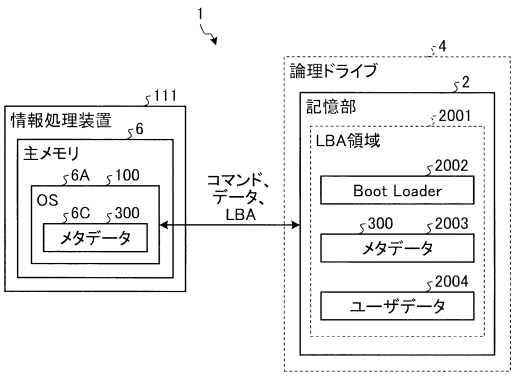
【図２９】



【図３０】

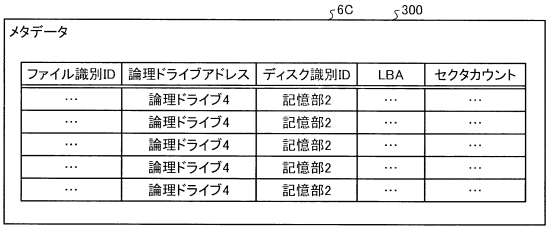


【図３１】



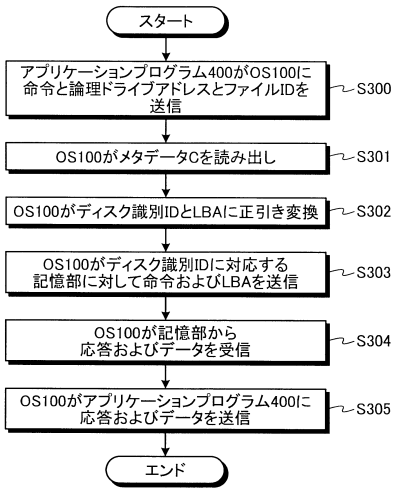
【図３２】

【図３３】

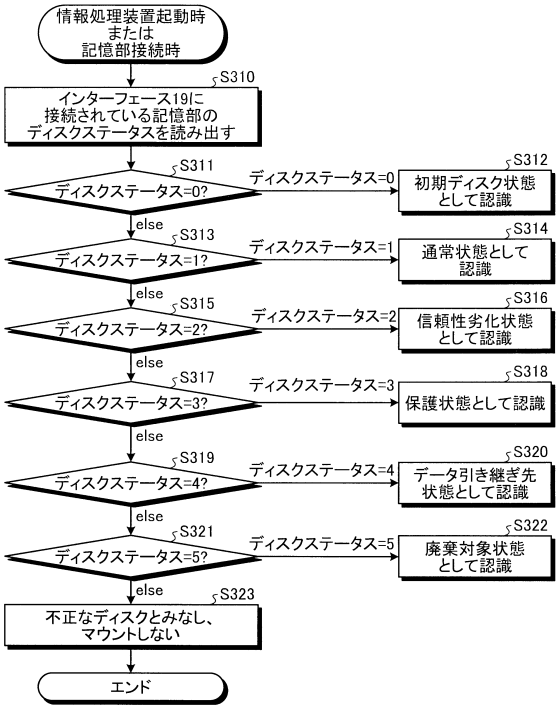


ファイル識別ID	論理ドライブアドレス	ディスク識別ID	LBA	セクタカウント
…	論理ドライブ4	記憶部2	…	…
…	論理ドライブ4	記憶部2	…	…
…	論理ドライブ4	記憶部2	…	…
…	論理ドライブ4	記憶部2	…	…
…	論理ドライブ4	記憶部2	…	…

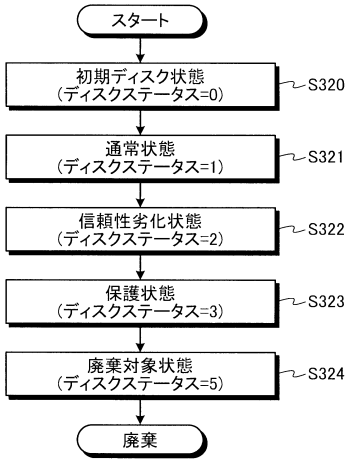
【図３４】



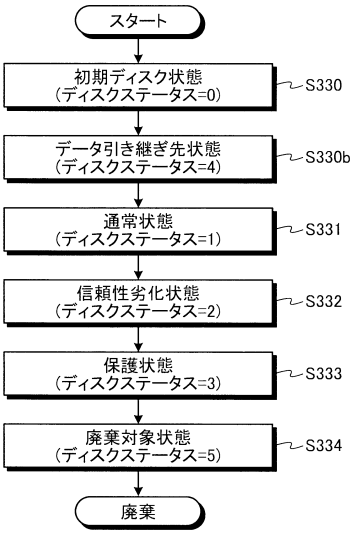
【図３５】



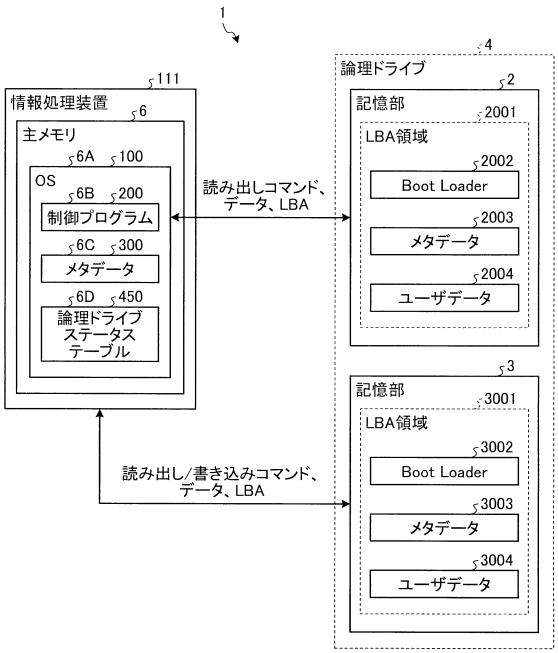
【図３６】



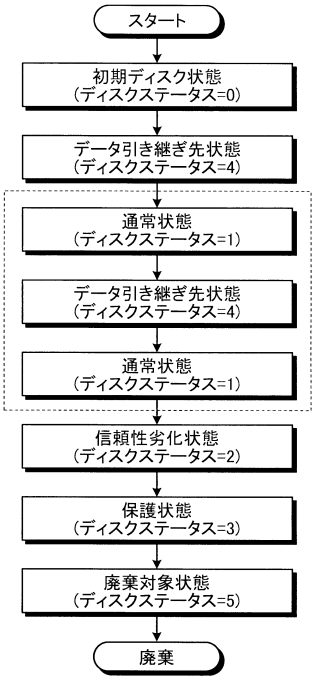
【図３７】



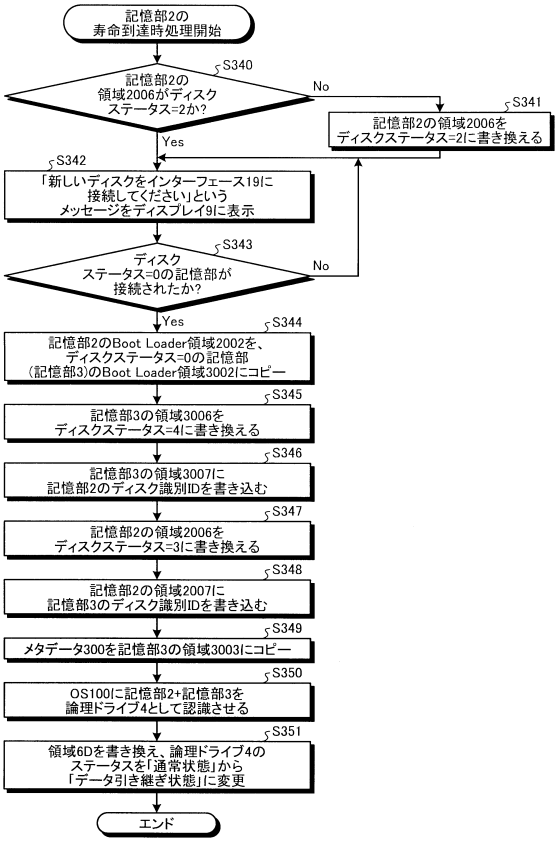
【図３８】



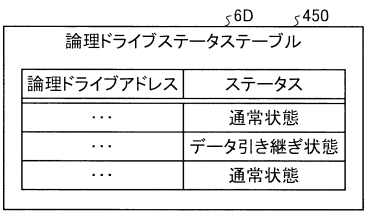
【図３９】



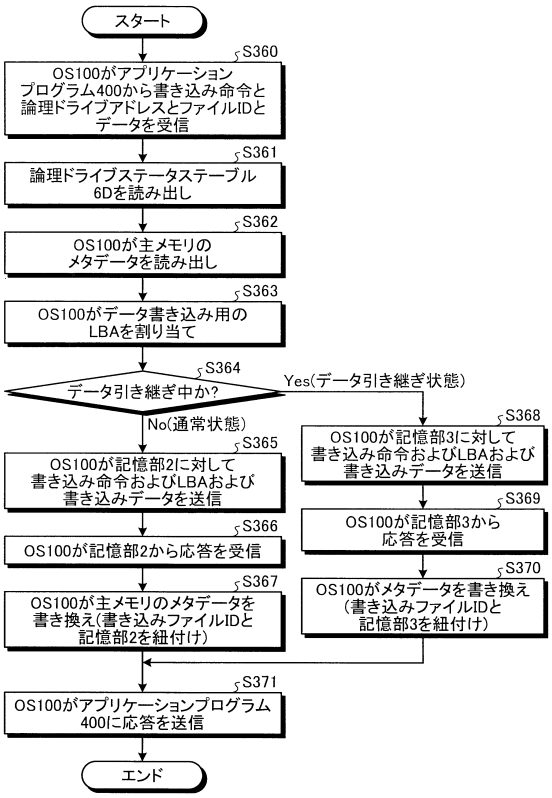
【図４０】



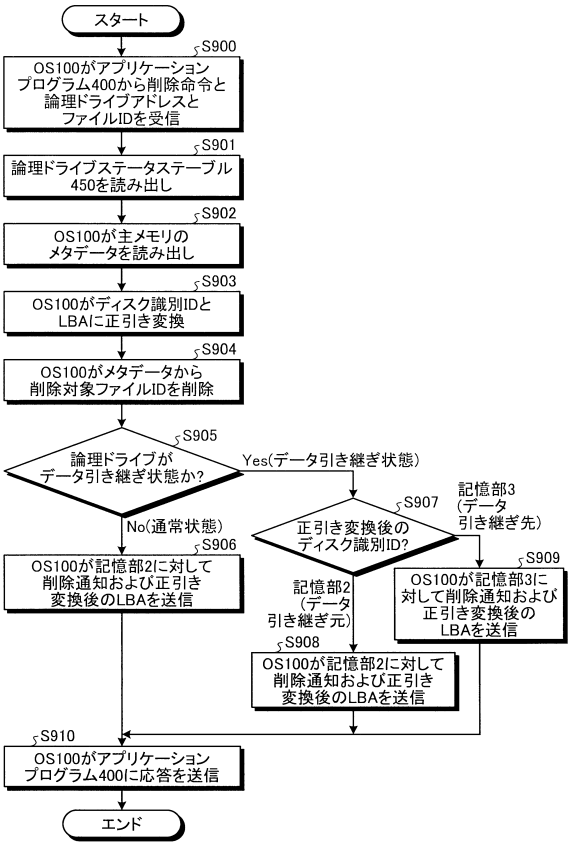
【図４１】



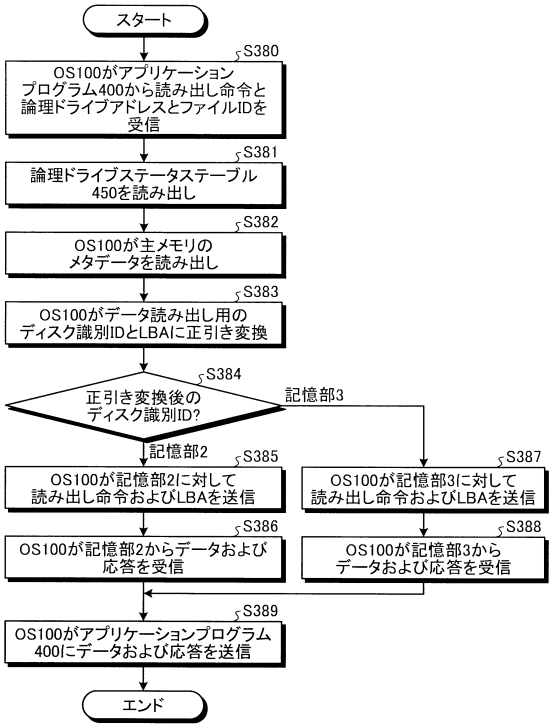
【図４２】



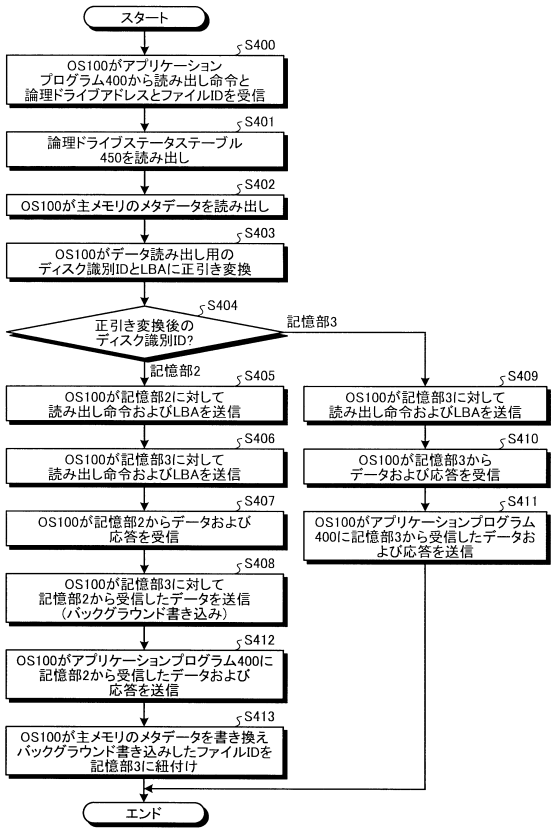
【図４３】



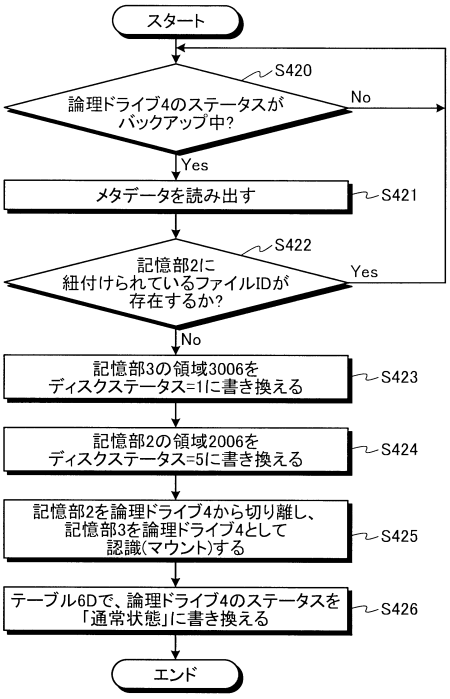
【図４４】



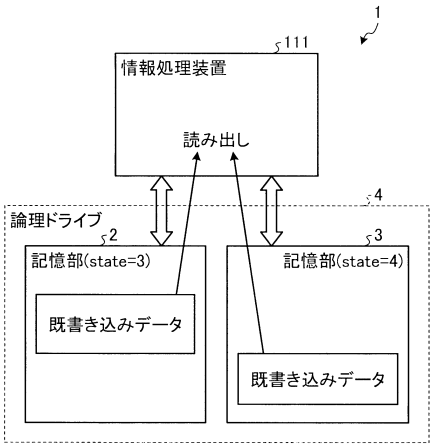
【図４５】



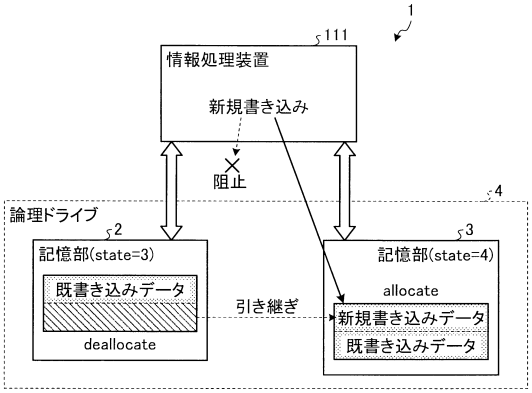
【図４６】



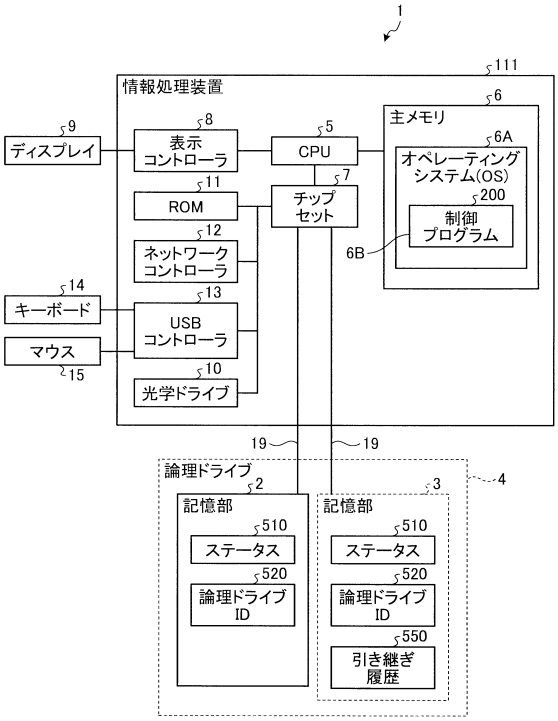
【図４７】



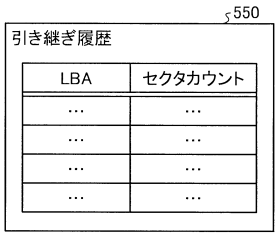
【図４８】



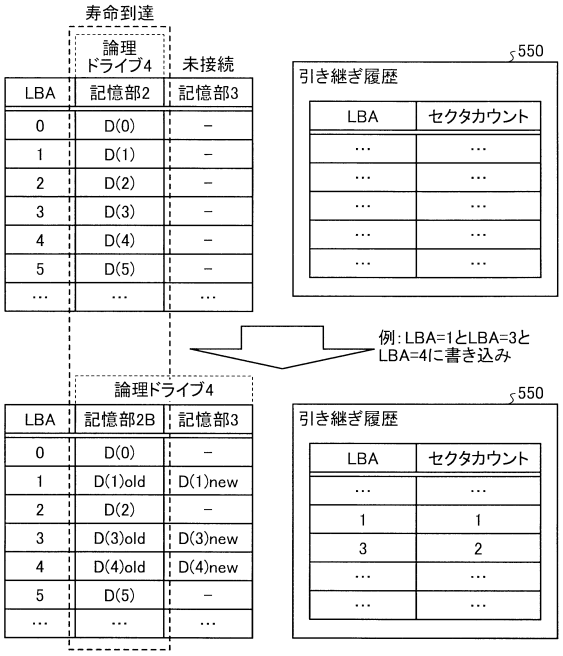
【図４９】



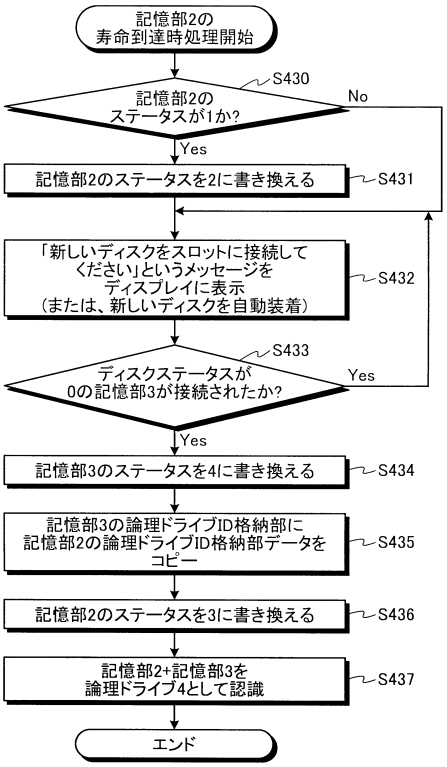
【図５０】



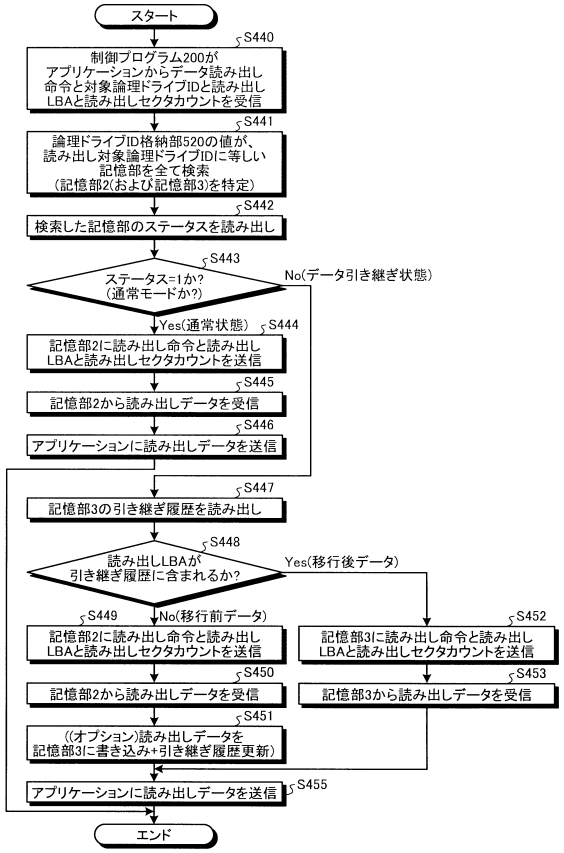
【図５１】



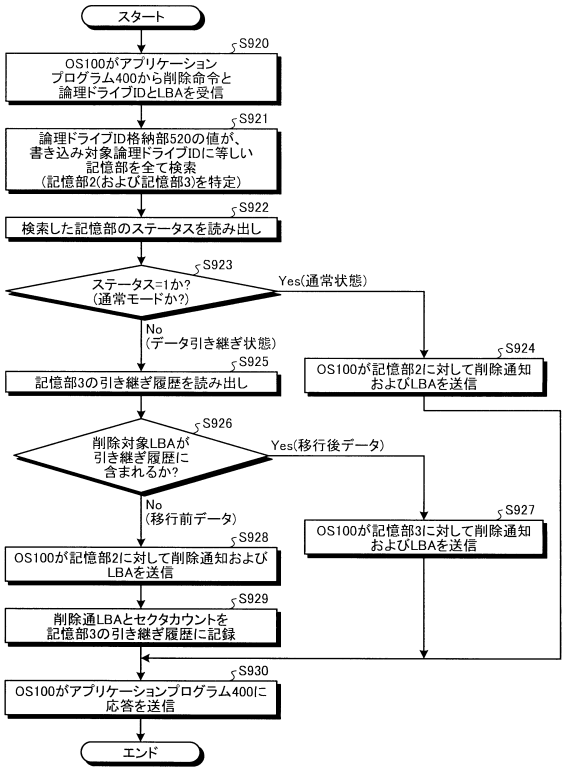
【図５２】



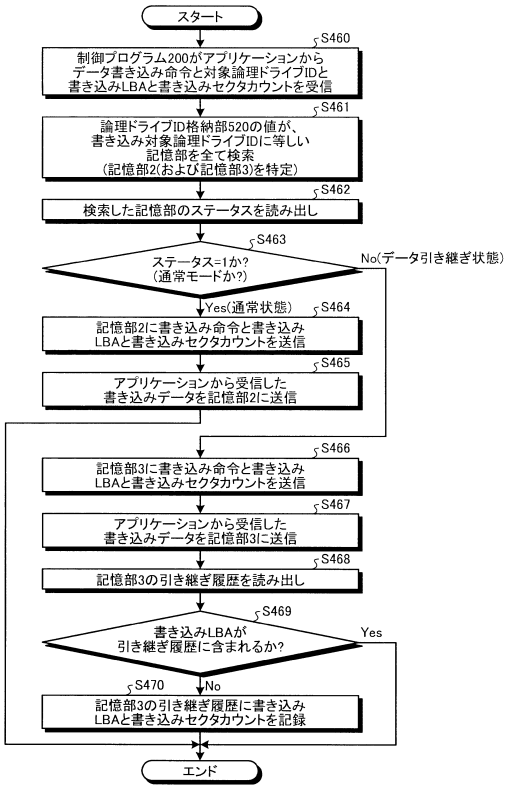
【図５３】



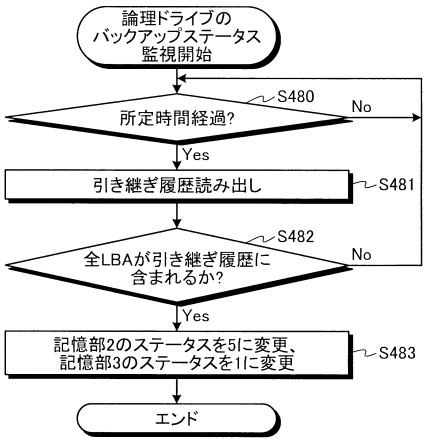
【図５４】



【図５５】



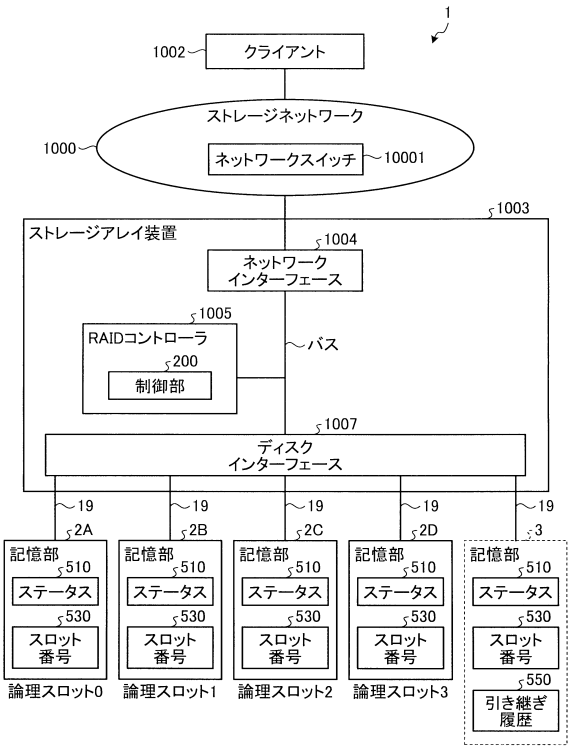
【図５６】



【図５７】

	論理ドライブ4		
ステップ	記憶部2	記憶部3	状態
1	1	未接続	信頼性劣化前
2	2	未接続	記憶部3を接続するよう要求
3	3	0	記憶部3を接続直後
4	4	4	データ引き継ぎ中
5	5	1	データ引き継ぎ完了

【図５８】



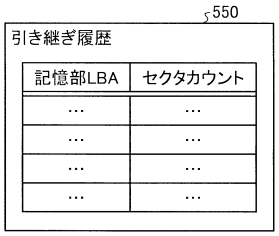
【図５９】

	論理スロット0	論理スロット1	論理スロット2	論理スロット3
記憶部LBA	記憶部2A	記憶部2B	記憶部2C	記憶部2D
0	D(0)	D(1)	D(2)	P(0,2)
1	D(3)	D(4)	P(3,5)	D(5)
2	D(6)	P(6,8)	D(7)	D(8)
3	P(9,11)	D(9)	D(10)	D(11)
4	D(12)	D(13)	D(14)	P(12,14)
5	D(15)	D(16)	P(15,17)	D(17)
…	…	…	…	…

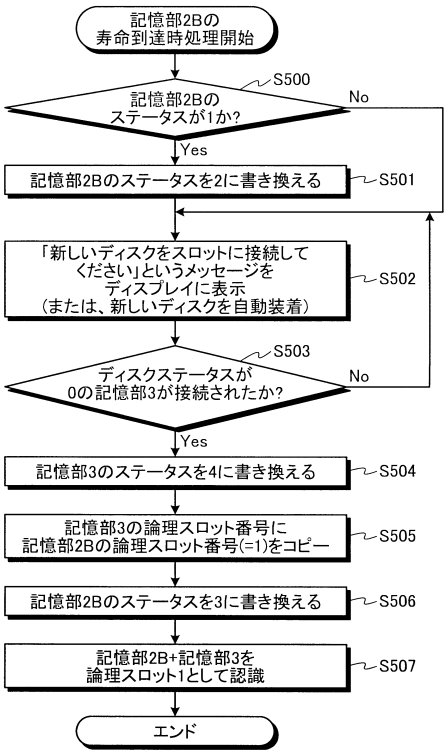
【図６０】

	論理スロット0	論理スロット1	論理スロット2	論理スロット3	論理スロット1
記憶部LBA	記憶部2A	記憶部2B	記憶部2C	記憶部2D	記憶部3
0	D(0)	D(1)	D(2)	P(0,2)	–
1	D(3)	D(4)	P(3,5)	D(5)	–
2	D(6)	P(6,8)	D(7)	D(8)	–
3	P(9,11)	D(9)	D(10)	D(11)	–
4	D(12)	D(13)	D(14)	P(12,14)	–
…	…	…	…	…	…

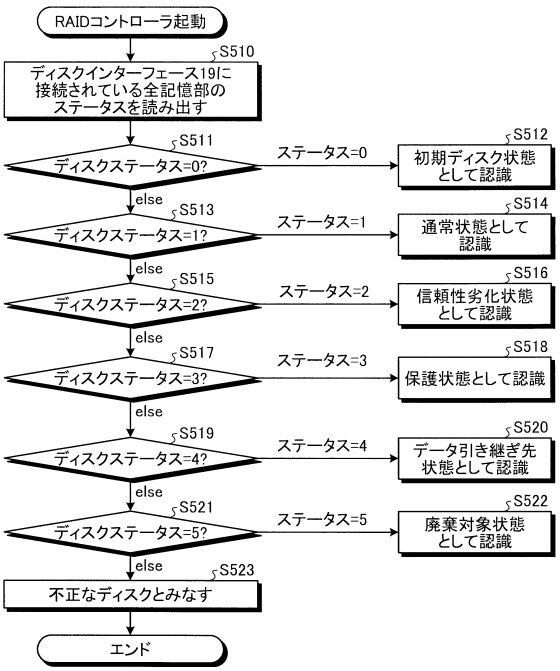
【図６１】



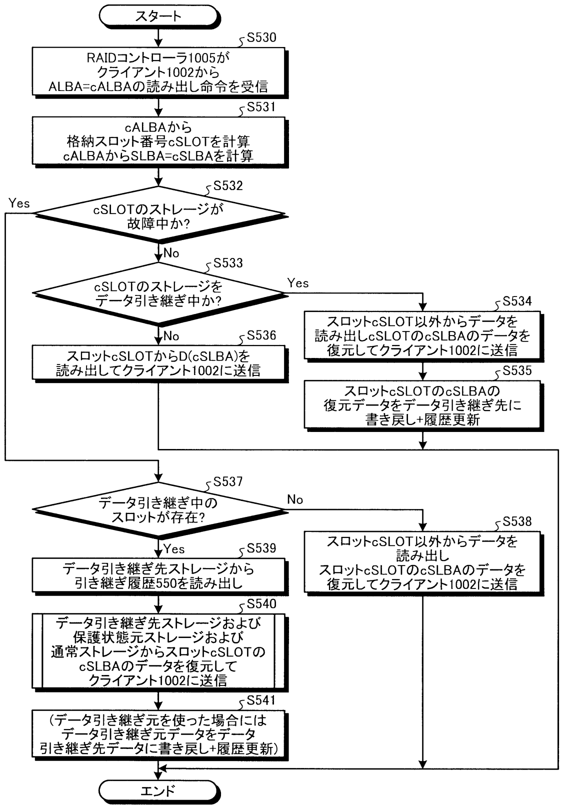
【図６２】



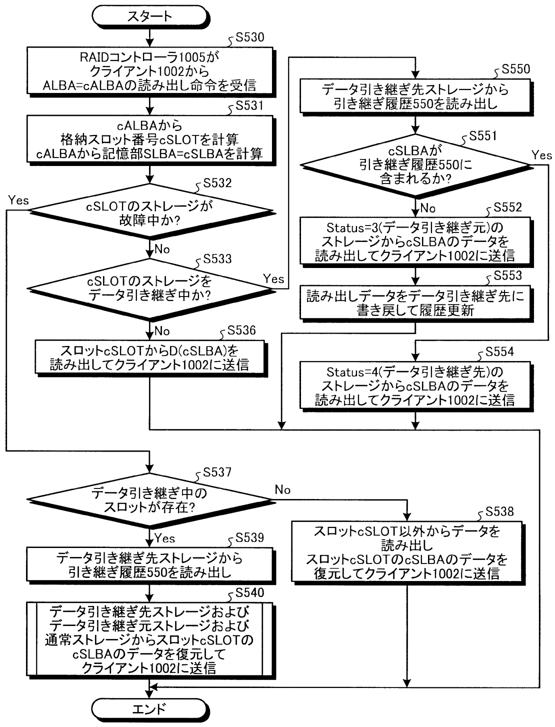
【図６３】



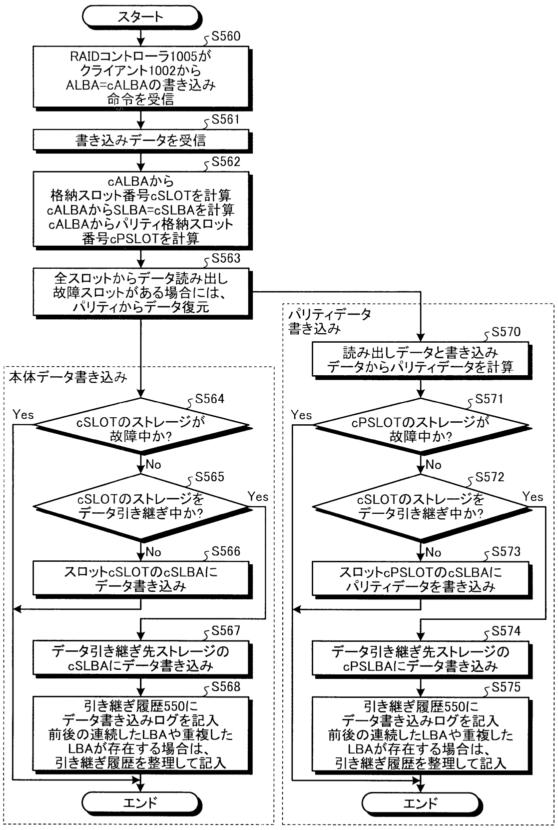
【図６４】



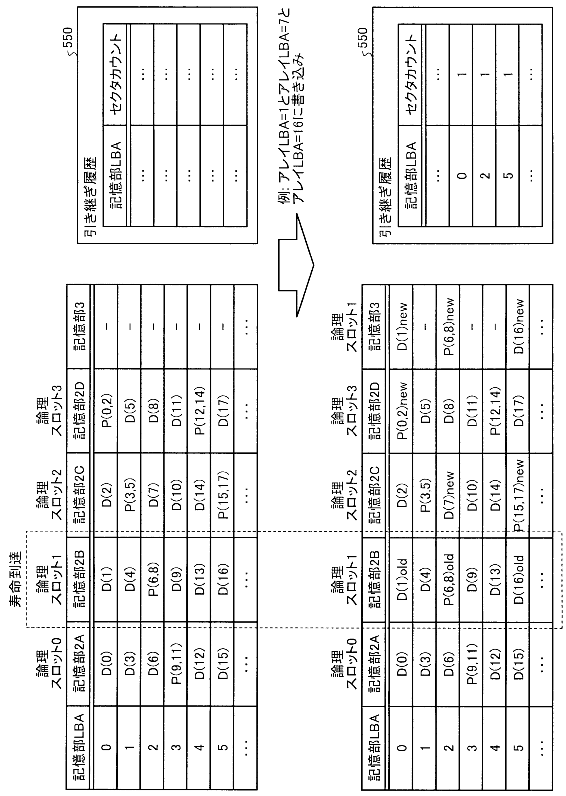
【図６５】



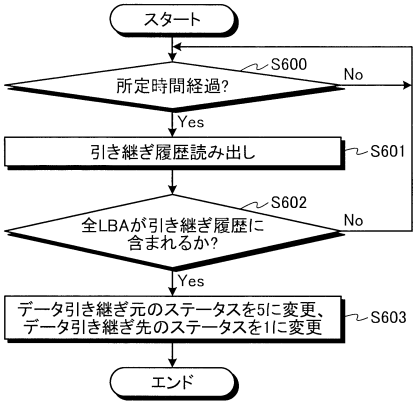
【図６６】



【図６７】



【図６８】



【図６９】

ステップ	論理スロット0 記憶部2A	論理スロット1 記憶部2B	論理スロット2 記憶部2C	論理スロット3 記憶部2D	論理スロット-1 記憶部3	状態
1	1	1	1	1	未接続	信頼性劣化前
2	1	2	1	1	未接続	記憶部3を接続するよう要求
3	1	2	1	1	0	記憶部3接続直後
4	1	3	1	1	4	データ引き継ぎ中 読み出しは記憶部2A/2C/2Dを優先して行い、記憶部2Bの LBAデータはパリティから適宜復元
5	1	5	1	1	1	データ引き継ぎ完了

【図７０】

記憶部LBA	論理スロット0 記憶部2A	信頼性劣化 データ引き継ぎ元 論理スロット1 記憶部2B	論理スロット2 記憶部2C	論理スロット-3 記憶部2D(故障)	新規ディスク データ引き継ぎ先 論理スロット-1 記憶部3	読み出し時の挙動
0	D(0)	D(1)old	D(2)	P(0,2)new (故障)	D(1)new	D(0), D(2)復元不要 D(1)はデータ引き継ぎ先である記憶部3から読み出し
1	D(3)	D(4)	P(3,5)	D(5) (故障)	–	D(3),D(4)は復元不要 D(5)はD(3)とD(4)とP(3,5)をXORして復元
2	D(6)	P(6,8)old	D(7)new	D(8) (故障)	P(6,8)new	D(6),D(7)は復元不要 D(8)はD(6)とD(7)newとP(6,8)newをXORして復元
3	P(9,11)	D(9)	D(10)	D(11) (故障)	–	D(9),D(10)は復元不要 D(11)はD(9)とD(10)とP(9,11)をXORして復元
4	D(12)	D(13)	D(14)	P(12,14) (故障)	–	復元不要
5	D(15)	D(16)old	P(15,17)new	D(17) (故障)	D(16)new	D(15)は復元不要 D(16)はデータ引き継ぎ先である記憶部3から読み出し D(17)はD(15)とD(16)newとP(15,17)newをXORして復元
…	…	…	…	…	…	…

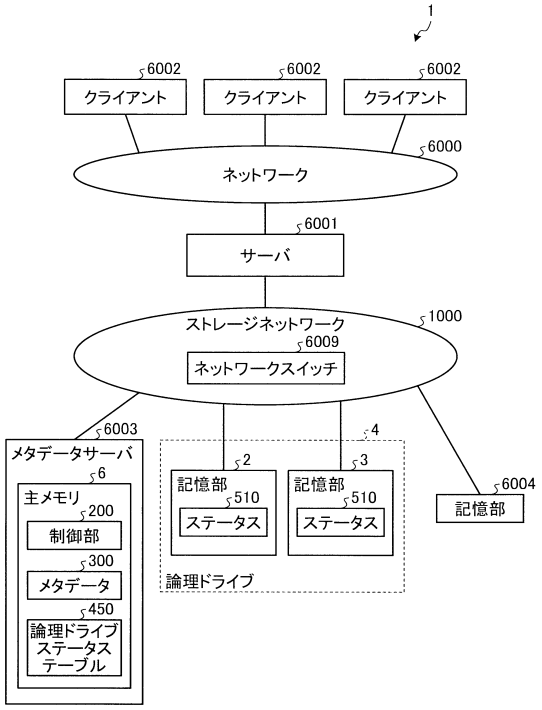
【図７１】

記憶部LBA	論理スロット0 記憶部2A	信頼性劣化 バックアップ元 論理スロット1 記憶部2B	論理スロット2 記憶部2C	論理スロット-3 記憶部2D(故障)	新規ディスク データ引き継ぎ先 論理スロット1 記憶部3B	新規ディスク データ引き継ぎ先 論理スロット2 記憶部3C	読み出し時の挙動
0	D(0)	D(1)old	D(2)	P(0,2)new (故障)	D(1)new	–	D(0), D(2)復元不要 D(1)は2Bのデータ引き継ぎ先である3Bから読み出し
1	D(3)	D(4)	P(3,5)	D(5) (故障)	–	–	D(3),D(4)は復元不要 D(5)はD(3)とD(4)とP(3,5)をXORして復元
2	D(6)	P(6,8)old	D(7)old	D(8) (故障)	P(6,8)new	D(7)new	D(6)は復元不要。 D(7)は2Bのデータ引き継ぎ先である3Cから読み出し D(8)はD(6)とD(7)newとP(6,8)newをXORして復元
3	P(9,11)	D(9)	D(10)	D(11) (故障)	–	–	D(9),D(10)は復元不要 D(11)はD(9)とD(10)とP(9,11)をXORして復元
4	D(12)	D(13)	D(14)	P(12,14) (故障)	–	–	復元不要
5	D(15)	D(16)old	P(15,17)old	D(17) (故障)	D(16)new	P(15,17) new	D(15)は復元不要 D(16)はデータ引き継ぎ先である記憶部3Bから読み出し D(17)はD(15)とD(16)newとP(15,17)newをXORして復元
…	…	…	…	…	…	…	…

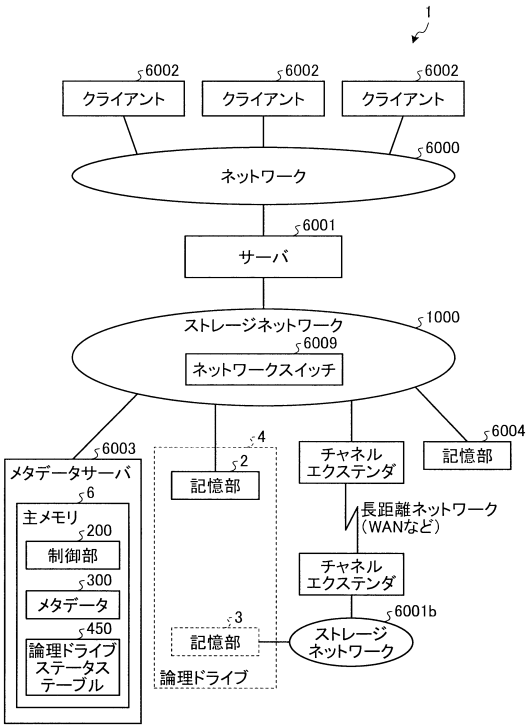
【図７２】

記憶部LBA	論理スロット0 データ引き継ぎ元 論理スロット1 記憶部2A	記憶部2B	論理スロット2 記憶部2C	論理スロット3 データ引き継ぎ先 論理スロット1 記憶部2D	記憶部3	読み出し時の挙動
0	D(0)	D(1)old	D(2)UECC	P(0,2)new	D(1)new	ALBA=2の読み出しでECC訂正不能エラー →D(0)とD(1)newとP(0,2)newをXORして復元
1	D(3)	D(4)	P(3,5)	D(5)	–	
2	D(6)	P(6,8)old	D(7)new	D(8)UECC	P(6,8)new	ALBA=7の読み出しでECC訂正不能エラー →D(6)とD(7)newとP(6,8)newをXORして復元
3	P(9,11)	D(9)UECC	D(10)	D(11)	–	ALBA=9の読み出しでECC訂正不能エラー →D(10)とD(11)とP(9,11)をXORして復元
4	D(12)	D(13)	D(14)	P(12,14)	–	
5	D(15)	D(16)old	P(15,17)new	D(17)	D(16)new	
…	…	…	…	…	…	

【図７３】



【図７４】



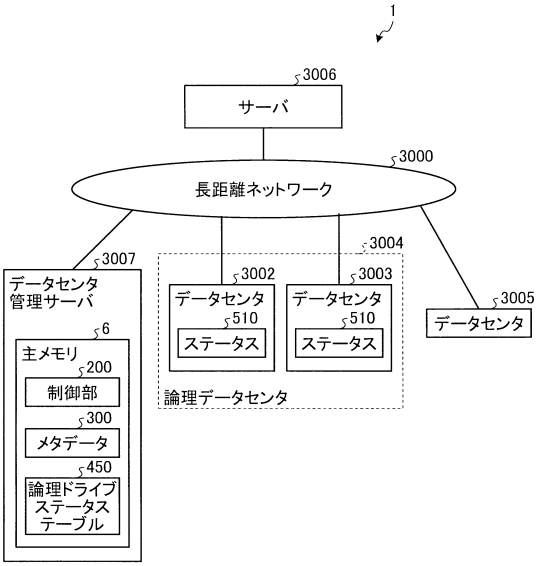
【図７５】

650

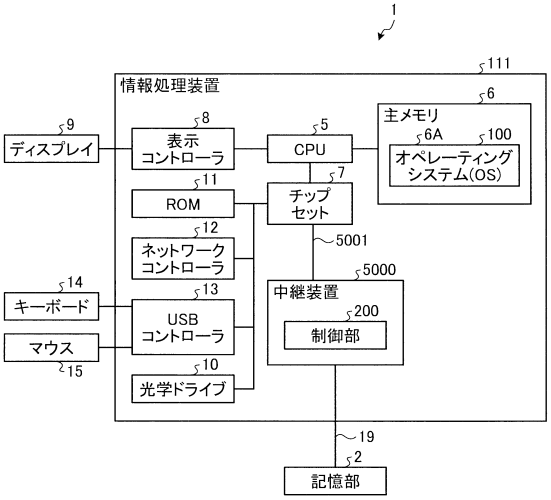
ステータステーブル

論理ドライブアドレス	ディスク認識ID	論理ドライブステータス	ディスクステータス
A	a	通常状態	1(通常状態)
B	b1	データ引き継ぎ中	3(データ引き継ぎ元)
B	b2	データ引き継ぎ中	4(データ引き継ぎ先)
C	c	通常状態	1(通常状態)

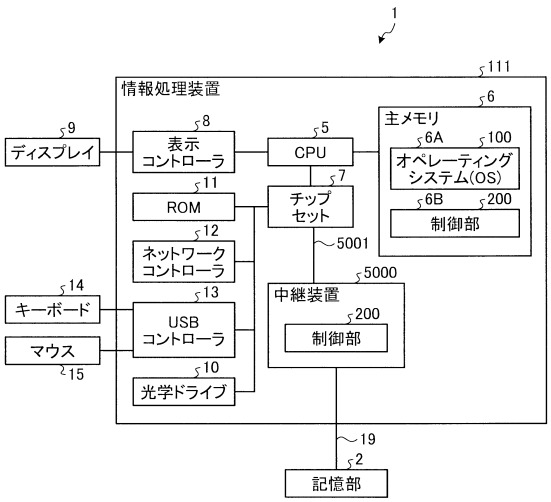
【図７６】



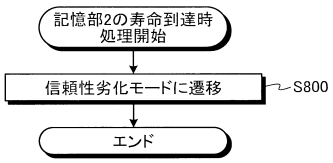
【図７７】



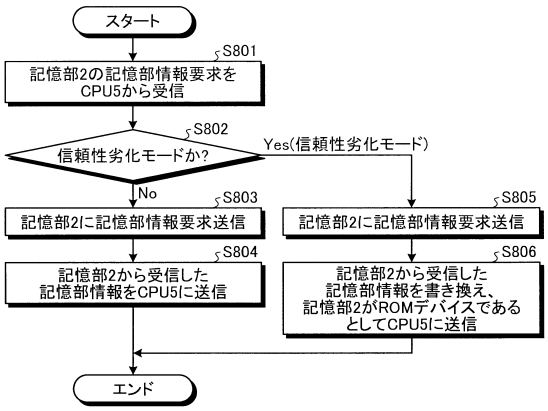
【図７８】



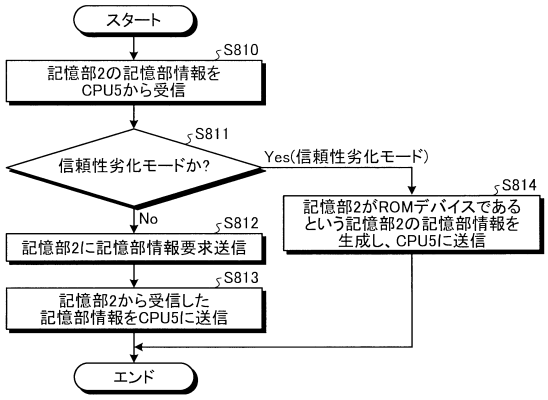
【図７９】



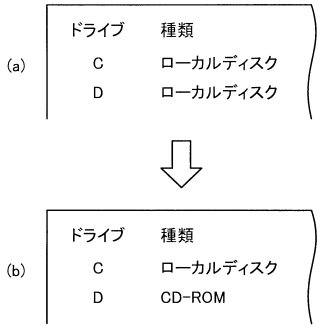
【図８０】



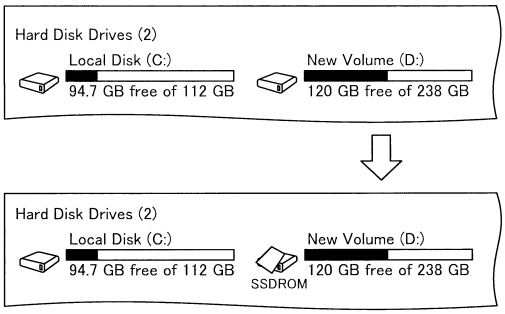
【図８１】



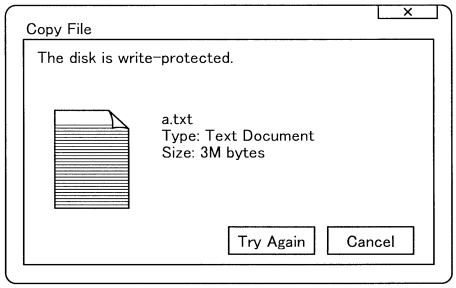
【図８２】



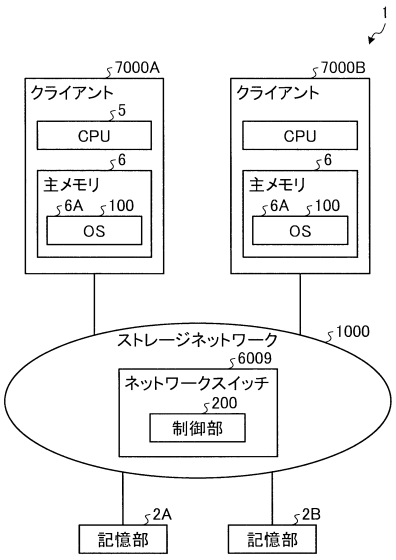
【図８３】



【図８４】



【図８５】



フロントページの続き

審査官　滝谷　亮一

(56)参考文献　米国特許出願公開第２０１０／０３０６５７７（ＵＳ，Ａ１）
特開２０１２－０２２６１９（ＪＰ，Ａ）
特表２０１２－５０９５２１（ＪＰ，Ａ）

(58)調査した分野(Int.Cl.，ＤＢ名)
Ｇ０６Ｆ　１２／１６
Ｇ０６Ｆ　３／０６
Ｇ０６Ｆ　１３／１０