(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2015 年 11 月 19 日(19.11.2015)

WIPO | PCT

(10) 国際公開番号
WO 2015/173889 A1

(51) 国際特許分類:
G06F 12/08 (2006.01) *G06F 12/12* (2006.01)

(21) 国際出願番号: PCT/JP2014/062714

(22) 国際出願日: 2014 年 5 月 13 日(13.05.2014)

(25) 国際出願の言語: 日本語

(26) 国際公開の言語: 日本語

(71) 出願人: 株式会社日立製作所(HITACHI, LTD.) [JP/JP]; 〒1008280 東京都千代田区丸の内一丁目６番６号 Tokyo (JP).

(72) 発明者: 伊藤 悠二(ITO, Yuji); 〒1008280 東京都千代田区丸の内一丁目６番６号 株式会社日立製作所内 Tokyo (JP). 鈴木 彬史(SUZUKI, Akifumi); 〒1008280 東京都千代田区丸の内一丁目６番６号 株式会社日立製作所内 Tokyo (JP). 山本 彰(YAMAMOTO, Akira); 〒1008280 東京都千代田区丸の内一丁目６番６号 株式会社日立製作所内 Tokyo (JP).

(74) 代理人: 特許業務法人第一国際特許事務所(PATENT CORPORATE BODY DAI-ICHI KOKUSAI TOKKYO JIMUSHO); 〒1080014 東京都港区芝４丁目１０番５号 Tokyo (JP).

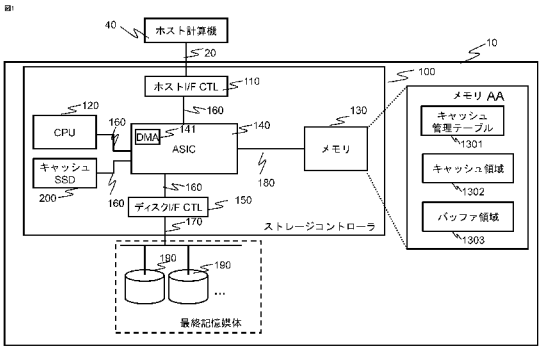
(81) 指定国 (表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) 指定国 (表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

添付公開書類:
— 国際調査報告（条約第 21 条(3)）

(54) **Title**: STORAGE DEVICE

(54) 発明の名称 : ストレージ装置



40 Host computer
100 Storage controller
110 Host interface CTL
130, AA Memory
150 Disk interface CTL
190 Final storage medium
200 Cache SSD
1301 Cache management table
1302 Cache region
1303 Buffer region

(57) **Abstract**: A storage device according to the present invention is provided with a cache device that caches data requested from an upper-level device such as a host computer. The cache device is provided with an internal controller that includes an internal cache configured from a memory capable of high speed access, such as a DRAM, and a storage unit configured from a storage medium having an erasure number limit, such as a flash memory (FM), and that has the function of controlling the internal cache and the storage unit and of performing a conversion and inverse-conversion process with respect to stored data. After the storage device stores write data in the cache device, the cache device, on the basis of an attribute or characteristics of the write data, determines the priority of an internal destage of the write data from the internal cache unit to the storage unit. The cache device then performs an internal destage from the internal cache unit to the storage unit preferentially from data with higher internal destage priority.

(57) 要約: 本発明に係るストレージ装置は、ホスト計算機などの上位装置から要求されたデータをキャッシュするキャッシュ装置を備える。キャッシュ装置は、ＤＲＡＭ等の高速アクセス可能なメモリで構成される内部キャッシュと、ＦＭ等の消去回数制限のある記憶媒体で構成される記憶部を有し、それらの制御および格納データに対する変換及び逆変換処理を行う機能を持つ内部コントローラを備える。ストレージ装置がキャッシュ装置にライトデータを格納した後、キャッシュ装置はライトデータの属性・特性に基づいて、ライトデータの内部キャッシュ部から記憶部への内部ディステージの優先度を決定する。そしてキャッシュ装置は、内部ディステージの優先度が高いデータを優先的に、内部キャッシュ部から記憶部に内部ディステージする。

WO 2015/173889 A1

明　細　書

発明の名称 ： ストレージ装置

技術分野

[0001] 本発明は、ストレージ装置に関する。

背景技術

[0002] 近年、サーバやストレージシステム（ストレージ装置）には、記憶媒体として半導体デバイスを用いたＳＳＤ（Ｓｏｌｉｄ　Ｓｔａｔｅ　Ｄｒｉｖｅ）が盛んに利用されている。ＳＳＤは、従来のＨＤＤ（Ｈａｒｄ　Ｄｉｓｋ　Ｄｒｉｖｅ）に比べてランダムアクセス性能に優れ、特にＮＡＮＤフラッシュメモリを用いたＳＳＤについては大容量化、低価格化が進み、ＨＤＤとの置き換えが進んでいる。また、不揮発記憶媒体としてＳＣＭ（Ｓｔｏｒａｇｅ　Ｃｌａｓｓ　Ｍｅｍｏｒｙ）は、ＮＡＮＤフラッシュメモリに比べ小容量で高価格であるが、アクセス速度や書き換え回数に優れ、高性能・長寿命の記憶媒体として注目されている。ＳＣＭは、不揮発性半導体メモリの一種であり、例えば、ＭＲＡＭ（Ｍａｇｎｅｔｉｃ　Ｒａｎｄｏｍ　Ａｃｃｅｓｓ　Ｍｅｍｏｒｙ）、ＰＲＡＭ（Ｐｈａｓｅ　Ｃｈａｎｇｅ　Ｒａｎｄｏｍ　Ａｃｃｅｓｓ　Ｍｅｍｏｒｙ）、ＲｅＲＡＭ（Ｒｅｓｉｓｔａｎｃｅ　Ｒａｎｄｏｍ　Ａｃｃｅｓｓ　Ｍｅｍｏｒｙ）である。

[0003] サーバやストレージシステムでは、データへのアクセス頻度の偏りを利用し、サーバやストレージコントローラに高速な記憶媒体を用いたキャッシュを設けることで、システム性能を向上させている。データの最終格納先であるＨＤＤへのアクセスを逐一行うと、大きなオーバヘッドとなってしまうためである。キャッシュに高頻度でアクセスされるデータを格納しておくことで、再度同データにアクセスする際にはＨＤＤへアクセスを行わずに高速にアクセスすることができる。そのため、キャッシュ容量が大きいほど、高速にアクセスできるデータ量が増加する。つまりキャッシュでのデータのヒット率が向上して、システム性能の向上が期待できる。

[0004] サーバやストレージシステムにおけるキャッシュとしてＤＲＡＭが広く用いられている。ＤＲＡＭはＨＤＤより遥かに高速であるが、大容量化や価格に問題がある。近年では、ＤＲＡＭに比べ安価かつ大容量な特性を持つＳＳＤをストレージシステムのキャッシュとして適用する技術が公開されている（特許文献１）。ＤＲＡＭよりも低価格で大容量なＮＡＮＤフラッシュメモリをキャッシュとして用いることで、キャッシュヒット率を向上させ、システム性能、コストパフォーマンスの改善を可能としている。

先行技術文献

特許文献

[0005] 特許文献1：米国特許第８２１４５９５号明細書

発明の概要

発明が解決しようとする課題

[0006] フラッシュメモリでは、データ消去回数が限られている。さらにフラッシュメモリでは、同じ物理アドレスに対するデータの上書きはできず、一度リード／ライト単位（ページ）よりも大きい単位（ブロック）で消去してから書き込む必要がある。上書きのために逐一消去することはオーバヘッドが大きく、限られた消去回数を浪費してしまう。そのため、フラッシュメモリを用いたシステムでは、ユーザが認識できる容量よりも大きな容量を持ち、上書きされるデータ用に更新領域を確保して、ユーザが認識できる論理的な領域における論理アドレスと、実際の物理アドレスとの対応付けを行って更新処理を行う。例えば、論理アドレスＬに対して物理アドレスＰを割り当てているデータが更新された場合、論理アドレスＬに対して新たに物理アドレスＰ’を割り当て、その物理ページにデータを書き込み、物理アドレスＰのデータを無効化する。この後、フラッシュメモリの容量が足りなくなった場合には、新たに書き込める領域を確保するために無効化したデータの消去を行う必要がある。つまり、書き込みが行われるということは、ＳＳＤの寿命を消費することにつながる。

[0007] サーバやストレージシステムのキャッシュにおいて、キャッシュへのデータの読み上げをステージング、キャッシュからデータの追い出しをディステージングと呼ぶ。サーバやストレージシステムのキャッシュとして用いられる記憶媒体では、キャッシュデータのステージング、ディステージングにより通常の最終記憶装置（データが最終的に管理される記憶装置）よりもデータ消去回数が非常に多くなる。なぜなら、システム全体の容量に対して小容量であるキャッシュでは、多数の最終記憶装置が処理するＩ／Ｏを少数のキャッシュ用の記憶媒体で処理することになるため、記憶媒体１台あたりの負荷が高くなって更新頻度が増加してしまうからである。また、リード処理時にデータのステージングを行う際には、キャッシュ用の記憶媒体に対してはリードデータを格納するためのライト処理が行われる。つまりリードのみのアクセスしかない場合でも、キャッシュ用記憶媒体の内容の書き換えが行われ、消去回数消費につながる。特にＭＬＣ（Ｍｕｌｔｉ Ｌｅｖｅｌ Ｃｅｌｌ）タイプのＮＡＮＤフラッシュメモリは、データ消去回数が限られているため、キャッシュとして用いると、最終記憶装置として用いるよりもさらに短寿命となってしまう。一方で、ＭＬＣタイプのＮＡＮＤフラッシュメモリより長寿命なＳＬＣ（Ｓｉｎｇｌｅ Ｌｅｖｅｌ Ｃｅｌｌ）タイプのフラッシュメモリやＳＣＭなどは、前述した通り、ＭＬＣタイプのフラッシュメモリよりも容量単価が高いため、キャッシュ容量を増やすことが難しく、システム性能の向上に効果的でない。

[0008] 寿命に課題のあるＳＳＤをサーバやストレージシステムといった上位装置のキャッシュとして用いる場合、キャッシュする効果の小さいデータについては、高コスト、小容量だが長寿命なＤＲＡＭのみにキャッシュ、またはＳＳＤにキャッシュしないことが考えられる。しかし、データをＳＳＤにキャッシュしない場合、キャッシュ装置として用いるＳＳＤの機能が利用不可となってしまい、上位装置のＣＰＵ、または別途搭載したハードウェアによって同等の機能を実現する必要がある。上位装置のＣＰＵを用いる場合、処理オーバヘッドが大きくなり、システム性能が低下してしまう。別途ハードウ

ェアを搭載する場合はシステムのコストが上昇してしまう。

[0009] 例えば、ＳＳＤが格納したデータを内部で暗号化し、暗号化後のデータをリードできる機能を持つとする。そのＳＳＤをキャッシュ装置として用い、データをキャッシュする際にＳＳＤ内部で暗号化し、ＳＳＤからディステージされるデータを暗号化したまま最終記憶媒体に格納することで、上位装置のＣＰＵなどのリソースを消費することなく全データの暗号化するシステムを実現できる。データのリード時は、ＳＳＤに一度格納することで、ＳＳＤ内部で復号させ、さらにはキャッシュとしてヒットした場合にＨＤＤよりも高速にデータを読み出せる。しかし、ＳＳＤの寿命を考慮して、前述したキャッシュする効果の小さいデータについてＳＳＤへキャッシュしない場合、暗号化および復号化ができず、同等の処理を行う上位装置のＣＰＵまたは別ハードウェアなどのリソースが必要となる。

課題を解決するための手段

[0010] 本発明の一実施形態に係るストレージ装置は、サーバやストレージコントローラといった上位装置から要求されたデータをキャッシュするキャッシュ装置を備え、キャッシュ装置は、高速・小容量・長寿命な内部キャッシュ部と、内部キャッシュ部より寿命の短い記憶部を有し、それらの制御を行う内部コントローラを備える。内部コントローラは、キャッシュ装置の内部キャッシュ部に格納された各データの属性・特性に応じて、内部キャッシュ部から記憶部への内部ディステージの優先度を制御する。内部コントローラは、内部ディステージ優先度の高いデータを優先的に、内部キャッシュ部から記憶部へとディステージし、内部ディステージ優先度が低いデータの記憶部へのライトを抑制する。

[0011] また本発明の一実施形態に係るストレージ装置が備えるキャッシュ装置は、キャッシュ装置に格納されたデータに対する、暗号化や圧縮などのデータ変換処理を行う機能を備え、上位装置から受信したライトデータにデータ変換処理を施してキャッシュ装置に格納し、上位装置に対してデータを送信する際には、データ逆変換を施して上位装置にデータを送信する。

発明の効果

[0012] 本発明によれば、フラッシュメモリなどの、データ消去回数の限られている記憶媒体を記憶部として用いたキャッシュ装置において、記憶部へのデータ書き込み回数を適切に抑制することができ、記憶部を長寿命化することができる。また、キャッシュ装置がデータ変換等の機能を備えている場合、短寿命な記憶部へのデータ書き込み回数を抑制しつつ、ライトデータに対するキャッシュ装置の機能の適用を可能にすることができる。

図面の簡単な説明

[0013] [図1]図１は、実施例１に係るストレージシステムの一例の構成図である。

[図2]図２は、実施例１に係るキャッシュ管理テーブルの一例を示す図である。

[図3]図３は、実施例１に係るキャッシュＳＳＤの一例の構成図である。

[図4]図４は、実施例１に係る内部キャッシュ管理テーブルの一例を示す図である。

[図5]図５は、実施例１に係る論理物理アドレス変換テーブルの一例を示す図である。

[図6]図６は、実施例１に係るストレージコントローラとキャッシュＳＳＤコントローラとの間のライト要求コマンドとその応答コマンドの一例を示す図である。

[図7]図７は、実施例１に係るストレージコントローラとキャッシュＳＳＤコントローラとの間のリードおよびデータ無効化要求コマンドとその応答コマンドの一例を示す図である。

[図8]図８は、実施例１に係るストレージコントローラによるライト処理のフローチャートである。

[図9]図９は、実施例１に係るストレージコントローラによるキャッシュＳＳＤへのキャッシュ格納処理のフローチャートである。

[図10]図１０は、実施例１に係るストレージコントローラによるキャッシュＳＳＤから最終記憶媒体へのディステージ処理を説明する図である。

[図11]図１１は、実施例１に係るキャッシュＳＳＤでのリード要求を受領した時の処理のフローチャートである。

[図12]図１２は、実施例１に係るキャッシュＳＳＤでのデータ無効化要求を受領した時の処理のフローチャートである。

[図13]図１３は、実施例１に係るキャッシュＳＳＤでのライト要求を受領した時の処理のフローチャートである。

[図14]図１４は、実施例１に係るキャッシュＳＳＤでの内部ディステージ処理のフローチャートである。

[図15]図１５は、実施例１に係るストレージコントローラによるリード処理のフローチャートである。

[図16]図１６は、実施例２に係るキャッシュＳＳＤでの内部ディステージ処理のフローチャートである。

[図17]図１７は、実施例２に係るストレージコントローラによるキャッシュＳＳＤへのライト処理のフローチャートである。

[図18]図１８は、実施例３に係るストレージコントローラによるライト処理のフローチャートである。

[図19]図１９は、実施例３に係るストレージコントローラによるリード処理のフローチャートである。

[図20]図２０は、実施例４に係るストレージコントローラによる、ライト処理のフローチャートである。

[図21]図２１は、実施例５に係るキャッシュＳＳＤで実行されるライト処理のフローチャートである。

発明を実施するための形態

[0014] 以下、図面を用いて、本発明の実施形態について述べる。

実施例 1

[0015] 図１は、本発明の実施例１に係る計算機システムの一例の構成図である。計算機システムは、キャッシュ制御装置の一例としてのストレージ装置１０と、ホスト計算機４０とを有する。ストレージ装置１０とホスト計算機４０

とは、ＳＡＮ（Ｓｔｏｒａｇｅ　Ａｒｅａ　Ｎｅｔｗｏｒｋ）やＬＡＮ（Ｌｏｃａｌ　Ａｒｅａ　Ｎｅｔｗｏｒｋ）などのネットワーク２０を介して接続される。ホスト計算機４０は、例えば、汎用的な計算機であり、ストレージ装置１０に格納されたデータを読み出したり、ストレージ装置１０にデータを格納させたりして、所定の業務処理を実行する。

[0016] ストレージ装置１０は、ストレージコントローラ１００と、１以上の最終記憶媒体１９０とを含む。最終記憶媒体１９０は、ホスト計算機４０からストレージ装置１０に対して書き込まれたデータを最終的に格納しておく記憶媒体であり、一例としてＨＤＤが用いられる。ただしＨＤＤ以外の記憶媒体、たとえばＳＳＤを用いる構成であってもよい。ストレージコントローラ１００は、バス１７０を介して、最終記憶媒体１９０と接続される。

[0017] ストレージコントローラ１００は、ホストインタフェースコントローラ（Ｉ／Ｆ　ＣＴＬ）１１０、ＣＰＵ１２０、例えばＤＲＡＭに代表されるメモリ１３０、ディスクインタフェースコントローラ（Ｉ／Ｆ　ＣＴＬ）１５０、２次キャッシュとして用いられるキャッシュＳＳＤ２００、およびＡＳＩＣ１４０を備える。ホストＩ／Ｆ　ＣＴＬ１１０、ＣＰＵ１２０、メモリ１３０、およびディスクＩ／Ｆ　ＣＴＬ１５０、キャッシュＳＳＤ２００は、ＰＣＩなどの専用内部バス１６０、メモリバス１８０を介してＡＳＩＣ１４０に接続される。ホストＩ／Ｆ　ＣＴＬ１１０は、ネットワーク２０を介しての通信を行うためのデバイスである。ＣＰＵ１２０は、ストレージ装置１０全体の動作制御を行う。メモリ１３０は、ストレージコントローラ１００のキャッシュ全体を管理するキャッシュ管理テーブル１３０１とデータをキャッシュするキャッシュ領域１３０２、データ転送のための一時格納領域として用いられるバッファ領域１３０３を含む。以降、キャッシュ領域１３０２のことをキャッシュＤＲＡＭ１３０２と呼ぶこともある。ＡＳＩＣ１４０は、ＣＰＵ１２０を介さずにメモリ１３０上のデータを転送するＤＭＡ（Ｄｉｒｅｃｔ　Ｍｅｍｏｒｙ　Ａｃｃｃｅｓｓ）１４１を備える。ディスクＩ／Ｆ　ＣＴＬ１５０は、内部バス１７０を介して、最終記憶媒体１９０と通

信を行うためのデバイスである。

[0018] ストレージコントローラ１００は、ホスト計算機４０がリードまたはライトするデータを、メモリ１３０上のキャッシュ領域１３０２またはキャッシュＳＳＤ２００上にキャッシュする。キャッシュ領域１３０２またはキャッシュＳＳＤ２００へのデータの格納及び読み出し処理については後述する。

[0019] 図２は、実施例１に係るストレージコントローラ１００のキャッシュを管理するキャッシュ管理テーブルの一例である。

[0020] キャッシュ管理テーブル１３０１は、ホスト計算機４０から書き込まれたデータが格納されている、キャッシュ領域１３０１またはキャッシュＳＳＤ２００上の位置情報等を管理する情報で、ストレージコントローラ１００のメモリ１３０上に格納されている。そしてキャッシュ管理テーブル１３０１の内容は、キャッシュへのデータの格納時等に、ストレージコントローラのＣＰＵ１２０によって更新される。

[0021] ストレージコントローラ１００（のＣＰＵ１２０）は、キャッシュＤＲＡＭ１３０２及びキャッシュＳＳＤ２００上の記憶領域を、６４ＫＢの固定領域に分割して管理しており、本発明の実施例ではこの６４ＫＢの固定領域のことをキャッシュセグメントと呼ぶ（または「セグメント」とも呼ばれる）。ＣＰＵ１２０は各セグメントにストレージ装置１０内で一意な識別番号（ＩＤ）を付して管理しており、この識別番号はキャッシュセグメントＩＤと呼ばれる（または「セグメントＩＤ」とも呼ばれる）。

[0022] 図２の例では、キャッシュセグメントＩＤ１３０１１の０ｘ０～０ｘ１ＦＦＦＦをキャッシュＤＲＡＭ１３０２の各セグメント（計１２８ＧＢの記憶領域）に付し、キャッシュセグメントＩＤ１３０１１の０ｘ２００００～０ｘ９ＦＦＦＦＦをキャッシュＳＳＤ２００の各セグメント（計５１２ＧＢの記憶領域）に付して管理している。ただしキャッシュ管理単位、キャッシュセグメントＩＤの割り当て方について、これに限定されない。

[0023] また、ストレージコントローラ１００は、キャッシュＤＲＡＭ１３０２及びキャッシュＳＳＤ２００上の記憶領域を、６４ＫＢのセグメント単位に管

理しているが、セグメントとは、ストレージコントローラ１００がキャッシュＤＲＡＭ１３０２及び／またはキャッシュＳＳＤ２００内の記憶領域を確保、解放する時の最小単位であり、ストレージコントローラ１００がキャッシュＤＲＡＭ１３０２及びキャッシュＳＳＤ２００に対してデータのアクセス（読み書き）を行う場合の、最小アクセス単位ではない。ストレージコントローラ１００はホスト計算機４０のストレージ装置１０に対する最小アクセス単位であるディスクブロック（５１２バイト。セクタとも呼ばれる）単位で、キャッシュセグメントへのデータの読み書きを行うことが可能である。ただし、本発明の実施例に係るストレージ装置１０において、後述するキャッシュＳＳＤ２００が有するデータ処理機能により、データ圧縮や暗号化を行う場合、データ圧縮や暗号化を行うデータ単位である４ＫＢが、最小アクセス単位になる。

[0024] キャッシュ管理テーブル１３０１には、キャッシュセグメントＩＤ１３０１１で特定される各キャッシュセグメントについて、当該キャッシュセグメントに格納されているデータが本来格納されるべき最終記憶媒体１９０（あるいは複数の最終記憶媒体１９０から構成される論理ボリューム）上の格納位置（アドレス）であるＬＢＡ１３０１２、当該キャッシュセグメント内に、実際にデータが格納されている位置を表すビットマップ１３０１３、キャッシュ状態１３０１４、最終記憶媒体１９０へのディステージ時に参照する直近アクセス時刻１３０１５（または最終アクセス時刻１３０１５とも呼ばれる）が格納されている。

[0025] ビットマップ１３０１３は、キャッシュセグメント上の、実際にデータが格納されている位置を表す１２８ビットの情報である。本発明の実施例では、キャッシュセグメントサイズは６４ＫＢだが、６４ＫＢの領域すべてにデータが格納されている（キャッシュされている）場合もあれば、必ずしもそうでない場合もある。ビットマップ１３０１３の各ビットは、キャッシュセグメント内の各ディスクブロック（１ディスクブロックは５１２バイトとする。そのため、１キャッシュセグメントは１２８個のディスクブロックから

なる）にデータが格納されているか否かを表す。ビットマップ１３０１３の先頭ビット（ＭＳＢ）は、キャッシュセグメント内の先頭ディスクブロックにデータが格納されているか表しており、以下順に２、３、…番目のビットはそれぞれ、キャッシュセグメント内の２番目、３番目、…のディスクブロックにデータが格納されているか表している。各ビットの内容が１の場合、当該ビットに対応するキャッシュセグメント上ディスクブロックにデータが格納されていることを意味し、０の場合には、当該ビットに対応するキャッシュセグメント上ディスクブロックにデータが格納されていないことを意味する。

[0026] キャッシュ状態１３０１４は、キャッシュされているデータが最終記憶媒体１９０に格納されているものに対して更新されているか否かを表すものである。キャッシュされているデータが最終記憶媒体１９０に格納されているものに対して更新されている場合にはキャッシュ状態１３０１４には「ｄｉｒｔｙ」が格納され、更新されていない場合（つまり、キャッシュセグメントに格納されているデータと最終記憶媒体１９０に格納されているデータとが同じ内容である場合）には、キャッシュ状態１３０１４には「ｃｌｅａｎ」が格納される。ｃｌｅａｎなデータである場合にはディステージ時に最終記憶媒体１９０へ追い出す必要はない。一方、ｄｉｒｔｙなデータは、最終記憶媒体１９０へ書き出す（ディステージする）必要がある。なお、１キャッシュセグメント内にｃｌｅａｎなデータとｄｉｒｔｙなデータが混在している場合があるが、１キャッシュセグメント内にｃｌｅａｎなデータとｄｉｒｔｙなデータが混在している場合には、キャッシュ状態１３０１４にはｄｉｒｔｙが格納される。

[0027] 直近アクセス時刻１３０１５には、直近アクセス時刻１３０１５に対応するキャッシュセグメントに最後にアクセスがあった時刻の情報が格納される。時刻情報としてたとえば、ある時点（２０１０年１月１日など）からの経過秒数が格納されるが、これ以外の形式で時刻情報を格納してもよい。直近アクセス時刻１３０１５は、ストレージコントローラ１００が最終記憶媒体

１９０へディステージするデータを選択する際に、ＬＲＵ（Ｌｅａｓｔ　Ｒｅｃｅｎｔｌｙ　Ｕｓｅｄ）アルゴリズムを用いて選択する場合に必要となる情報である。ただし、ディステージ時のデータ選択ポリシはＬＲＵに限定されない。最終記憶媒体１９０へのディステージするデータの選択を別ポリシで行う場合は、キャッシュ管理テーブル１３０１には直近アクセス時刻１３０１５に代えて、データ選択ポリシに必要な情報を格納するようにする。

[0028] なお、本発明の実施例１に係るストレージ装置１０では、キャッシュＤＲＡＭ１３０２を１次キャッシュとし、キャッシュＳＳＤ２００を２次キャッシュとして用いる。そのため、ホスト計算機４０からのライトデータが、キャッシュＤＲＡＭ１３０２、キャッシュＳＳＤ２００の両方に格納されている場合もある。その場合、キャッシュセグメントＩＤ１３０１１が０ｘ０～０ｘ１ＦＦＦＦＦ（キャッシュＤＲＡＭ１３０２内セグメント）のエントリと、キャッシュセグメントＩＤ１３０１１が０ｘ２００００～０ｘ９ＦＦＦＦＦ（キャッシュＳＳＤ２００内セグメント）のエントリの中に、ＬＢＡ１３０１２が同一であるエントリが存在する。

[0029] 図３は、本実施例におけるキャッシュＳＳＤ２００の構成図の一例である。本実施例におけるキャッシュＳＳＤ２００は、不揮発記憶媒体であるフラッシュメモリ（以下ではＦＭと略記することもある）を主たる記憶媒体とする。キャッシュＳＳＤ２００は、内部に１または複数のＳＳＤコントローラ（ＳＳＤ　ＣＴＬ）２０１と複数のＦＭ２０１１を備える。

[0030] ＳＳＤコントローラ２０１は、その内部にインタフェースコントローラ（Ｉ／Ｆ　ＣＴＬ）２００１、プロセッサ２００３、内部メモリ（ＤＲＡＭ）２００５、データ処理アシストユニット（Ａｓｓｉｓｔ）２００６、ＦＭコントローラ（ＦＭＣ）２００７、およびデータ転送を相互に行う内部スイッチ（Ｓ／Ｗ）２００４を備えている。インタフェースコントローラ２００１、プロセッサ２００３、データ処理アシストユニット２００６、ＦＭコントローラ２００７は内部バスを介して内部スイッチ２００４に接続される。また内部メモリ２００５は、メモリバス２００９を介して内部スイッチ２００

４に接続されている。

[0031] インタフェースコントローラ２００１は、ストレージコントローラ１００が備えるＣＰＵ１２０から、キャッシュＳＳＤ２００に対する各種コマンドを受け付ける、あるいはデータ転送を行うためのものである。インタフェースコントローラ２００１は、ストレージ装置１０内のストレージコントローラ１００が備えるＡＳＩＣ１４０と接続される。

[0032] ＣＰＵ２００３は、内部スイッチ２００４を介してＳＳＤコントローラ２０１の各部位と接続され、内部メモリ２００５に記録されたプログラム及び管理情報（後述する内部キャッシュ管理テーブル２００５２など）を用いて、ＳＳＤコントローラ２０１全体を制御する。

[0033] 内部メモリ２００５には具体的には、ビットコストはＦＭ２０１１より高いものの、高速でかつＦＭ２０１１より長寿命である、ＤＲＡＭなどの揮発性メモリが用いられる。内部メモリ２００５上には、ＳＳＤコントローラ２０１でのデータ転送処理途中のデータを一時的に格納するために用いられる内部キャッシュ領域２００５１、内部キャッシュ領域２００５１に格納されたデータの情報を管理する内部キャッシュ管理テーブル２００５２、ＦＭ２０１１の格納情報を管理する論理物理変換テーブル２００５３、ＦＭ２０１１の物理ブロックに有効な情報が格納されているか否かなどの情報を格納するブロック情報２００５４が存在する。なお、これらの領域やテーブルについては、１つの内部メモリ２００５上に格納してもよいし、複数の内部メモリ２００５に分散して格納してもよい。以降、内部キャッシュ領域２００５１については、単に内部キャッシュ２００５１と呼ぶこともある。また、本発明の実施例に係るキャッシュＳＳＤでは、内部キャッシュ２００５１のサイズ（容量）は、複数のＦＭ２０１１の合計容量よりも小さいものとするが、内部キャッシュ２００５１のサイズがＦＭ２０１１の合計容量よりも小さくない場合であっても、本発明は有効である。

[0034] ＦＭコントローラ２００７は、複数バス（例えば１６）２０１０によってＦＭ２０１１と接続する。各バス２０１０には複数（例えば２）のＦＭ２０

１１が接続される。

[0035] また詳細は後述するが、本発明の実施例に係るキャッシュＳＳＤ２００は、複数のＦＭ２０１１の有する記憶領域から構成される論理的な記憶空間を、キャッシュＳＳＤ２００が接続されるストレージコントローラ１００に対して提供する。ストレージコントローラ１００がキャッシュＳＳＤ２００にアクセス（リード、ライト等）する際、たとえばキャッシュＳＳＤ２００にデータの書き込み（ライト）を行う場合、論理的な記憶空間上の一次元の論理アドレスを指定したライトコマンドをキャッシュＳＳＤ２００に対して発行することで、ライト対象データをキャッシュＳＳＤ２００が提供する論理的記憶空間上に（仮想的に）書き込むことができる。なお、ライトコマンドとともにキャッシュＳＳＤ２００に到来したデータは、一旦内部キャッシュ領域２００５１に格納され、その後、論理的な記憶空間に対応付けられたＦＭ２０１１上の記憶領域へと移動（内部ディステージ）される。これらの制御は、ＣＰＵ２００３が内部メモリ２００５に記録されたプログラム及び管理情報を用いて実行することで実現される。

[0036] キャッシュＳＳＤ２００は、キャッシュＳＳＤ２００内に格納するデータの圧縮／伸長やＰａｒｉｔｙ生成、暗号化／復号化などの、データ変換処理及び逆変換処理を行う機能を備える。データ処理アシストユニット２００６は、これらのデータ変換処理／逆変換処理、たとえばデータ圧縮／伸長処理を行う機能を有するハードウェアである。データ処理アシストユニット２００６は、ＣＰＵ２００３からの指示に従って、インタフェースコントローラ２００１から到来してＦＭ２０１１に書き込まれるデータに対するデータ変換処理（圧縮、暗号化など）を行い、またはＦＭ２０１１からインタフェースコントローラ２００１へ送出されるデータに対する、データ逆変換処理（圧縮データの伸長、暗号化データの復号化など）を行う。なお、これらのデータ変換／逆変換処理を行う機能をキャッシュＳＳＤ２００に実装する方法として、データ処理アシストユニット（Ａｓｓｉｓｔ）２００６のようなハードウェアを設ける方法のほか、ＣＰＵ２００３にデータの圧縮／伸長やＰ

ａｒｉｔｙ生成、暗号化／復号化などのデータ処理を行うプログラムを実行させることで、データ変換／逆変換を行う方法をとってもよい。

[0037] 以上説明した、ＳＳＤコントローラ２０１の各部位は、ＡＳＩＣやＦＰＧＡ（Ｆｉｅｌｄ Ｐｒｏｇｒａｍｍａｂｌｅ Ｇａｔｅ Ａｒｒａｙ）として、一つの半導体素子内で構成してもよいし、複数の個別専用ＩＣ（Ｉｎｔｅｇｒａｔｅｄ Ｃｉｒｃｕｉｔ）を相互に接続した構成であってもよい。また、キャッシュＳＳＤ２００に内部メモリ２００５用のキャパシタやバッテリを備え、電断時などに揮発する内部メモリ２００５に格納されたデータや管理情報をＦＭ２０１１に退避する構成としてもよい。

[0038] なお、本実施例に係るキャッシュＳＳＤ２００では、図３に示すように、主たる記憶媒体としてＦＭを使用しているが、キャッシュＳＳＤ２００に搭載する記憶媒体はＦＭに限定されるものではなく、Ｐｈａｓｅ Ｃｈａｎｇｅ ＲＡＭやＲｅｓｉｓｔａｎｃｅ ＲＡＭなどの不揮発メモリであってもよい。また、ＦＭ２０１１の一部または全部を揮発性のＲＡＭ（ＤＲＡＭ等）とする構成であってもよい。

[0039] 続いて図４を用いて、内部キャッシュ管理テーブル２００５２について説明する。

[0040] 先に述べたとおり、ストレージコントローラ１００はキャッシュＳＳＤ２００が提供する論理的な記憶空間に対してデータの書き込み（ライト）を行う場合、ライト先として論理的記憶空間上のアドレス（論理アドレス）を指定したライトコマンドをキャッシュＳＳＤ２００に対して発行する。ライトコマンドとともにキャッシュＳＳＤ２００に到来したデータは、一旦内部キャッシュ領域２００５１に格納（キャッシュ）される。キャッシュＳＳＤ２００のＣＰＵ２００３は、ライトデータのライト先である論理アドレスと、当該ライトデータが一旦格納される内部キャッシュ２００５１上の位置とを対応付けて、内部キャッシュ管理テーブル２００５２に記録して管理する。

[0041] ストレージコントローラ１００のキャッシュ領域１３０２と同様、キャッシュＳＳＤ２００のＣＰＵ２００３は、内部キャッシュ領域２００５１を、

所定サイズ（たとえば４ＫＢあるいは８ＫＢ）の固定領域（内部キャッシュセグメントと呼ぶ）に分割して管理している。これは、内部キャッシュ領域２００５１の割り当て、解放の単位であるとともに、データ処理アシストユニット２００６による、データ処理（圧縮や暗号化）を行う際の単位である。またＣＰＵ２００３は各内部キャッシュセグメントに、キャッシュＳＳＤ２００内で一意な識別番号（内部キャッシュセグメントＩＤと呼ぶ。ただし「セグメントＩＤ」と略記されることもある）を付して管理する。

[0042] 内部キャッシュ管理テーブル２００５２には、内部キャッシュセグメントのセグメントＩＤ２００５２１で特定される各内部キャッシュセグメントについて、当該内部キャッシュセグメントに格納されているデータの本来の書き込み先（先に述べた論理的な記憶空間上の位置（論理アドレス）であって、ストレージコントローラ１００から指定された書き込み先アドレス）である論理アドレス２００５２２、当該内部キャッシュセグメントのデータ処理（圧縮等）後のデータサイズを表すデータサイズ２００５２３、内部キャッシュ状態２００５２４、ＦＭへのディステージ時に参照する直近アクセス時刻２００５２５、ＦＭへの内部ディステージ優先度２００５２６（ディステージ優先度と略記されることもある）が格納されている。

[0043] データサイズ２００５２３は、キャッシュＳＳＤ２００にデータが格納される時に、データ処理アシストユニット２００６によるデータ処理を行った結果、データサイズが変動する場合に用いられる。たとえばデータ処理として圧縮を行った場合、多くの場合データサイズが小さくなる。データサイズ２００５２３には、データ処理（たとえば圧縮）後のデータサイズの情報が格納される。たとえば図４の例で、内部キャッシュセグメントＩＤ２００５２１が０ｘ１の行（２００５２－０）に格納されている情報を例にとると、内部キャッシュセグメントＩＤ２００５２１が０ｘ１である４ＫＢの内部キャッシュ領域２００５１が確保されているが、当該内部キャッシュ領域にはデータ処理後のデータが格納されており、そのサイズは、データサイズ２００５２３に格納されている内容を参照すると、１ＫＢであることがわかる。

[0044] なお、キャッシュＳＳＤ２００がデータ処理アシストユニット２００６によるデータ処理として、たとえばデータ圧縮を行って圧縮後のデータを内部キャッシュ領域２００５１あるいはＦＭ２０１１に格納する場合であっても、キャッシュＳＳＤ２００がストレージコントローラ１００に提供している論理的な記憶空間に影響はなく、ストレージコントローラ１００からは、あたかも圧縮前のデータ（非圧縮データ）を読み書きしているようにみえている。

[0045] 内部キャッシュ状態２００５２４は、先に説明したキャッシュ管理テーブル１３０１のキャッシュ状態１３０１４と同様のもので、内部キャッシュ状態２００５２４には「ｃｌｅａｎ」または「ｄｉｒｔｙ」のいずれかの情報が格納される。

[0046] 直近アクセス時刻２００５２５も、先に説明したキャッシュ管理テーブル１３０１の直近アクセス時刻１３０１５と同様のもので、直近アクセス時刻２００５２５に対応する内部キャッシュセグメントに最後にアクセスがあった時刻の情報が格納される。キャッシュ領域１３０２と同様、内部キャッシュ領域２００５１からＦＭ２０１１へとデータを移動（ディステージ）する際に、移動対象データの選択ポリシにはＬＲＵアルゴリズムが用いられるが、ポリシは必ずしもＬＲＵアルゴリズムに限定されない。ＬＲＵアルゴリズムを用いない場合、直近アクセス時刻２００５２５に代えて、そのディステージポリシに応じた情報を内部キャッシュ管理テーブル２００５２に格納する。

[0047] 続いて内部ディステージ優先度２００５２６について説明する。本発明の実施例に係るキャッシュＳＳＤ２００では、ストレージコントローラ１００から送信される、後述するライト要求コマンド３０００に含まれるデータ情報３００５を用いて、ＳＳＤコントローラ２０１が、そのライトデータがすぐにキャッシュＳＳＤ２００から追い出されるか（あるいはストレージコントローラ１００が当該ライトデータに対して再アクセスする可能性が低い）、それとも保持される（ストレージコントローラ１００から再アクセスされ

る）かを予測し、内部ディステージ優先度２００５２６を決定し、内部キャッシュ管理テーブル２００５２の更新時に設定する。短寿命なＦＭ２０１１のライトによる劣化を抑止するため、キャッシュＳＳＤ２００からすぐ追い出されると予測されるデータについては、内部ディステージ優先度２００５２６を低と設定して、できるだけＦＭへライトせずに内部キャッシュ２００５１で留める。予測通りであれば、ストレージコントローラ１００による最終記憶媒体１９０へのディステージの際にはデータがキャッシュＳＳＤ２００の内部キャッシュ２００５１上にある。キャッシュＳＳＤ２００からのリードは内部キャッシュ２００５１でヒットとなり、続く無効化要求でキャッシュＳＳＤ２００からデータが無効化されるため、ＦＭへのライトが発生しない。この一連の処理の詳細については後述する。

[0048] ＦＭへのディステージ優先度２００５２６は、内部キャッシュ２００５１からデータを追い出す際に、直近アクセス時刻２００５２５と合わせて、内部キャッシュ２００５１からＦＭへディステージするべきデータの選択に用いられる。ディステージ優先度２００５２６には、高（Ｈｉｇｈ）または低（Ｌｏｗ）のいずれかの情報が格納される（また、高（Ｈｉｇｈ）の情報を格納する代わりに「１」を格納する、あるいは低（Ｌｏｗ）の情報を格納する代わりに「０」を格納するようにしてもよい）。なお、以降では、ＳＳＤコントローラ２０１による、内部キャッシュ２００５１からＦＭへのデータディステージのことを、内部ディステージと呼ぶ。内部ディステージする必要がある時、ディステージ優先度２００５２６の高いデータの中から、ＬＲＵアルゴリズムにより内部ディステージするデータを選択する。ディステージ優先度２００５２６の低いデータについては、ディステージ優先度２００５２６の高いデータがない時にのみ、内部ディステージの対象として選ばれる。

[0049] ただし、内部ディステージポリシは上記のポリシに限定されない。たとえば直近アクセス時刻２００５２５が一定時間以上前であった場合には、ディステージ優先度２００５２６の高いデータに加えてディステージ優先度２０

０５２６の低いデータも内部ディステージ候補とする、などのポリシを採用しても良い。また、ディステージ優先度２００５２６に、高（Ｈｉｇｈ）と低（Ｌｏｗ）の２段階の情報を格納する態様のほか、複数段階の情報を格納するようにしてもよい。

[0050] 本発明の実施例１に係るストレージ装置１０の場合、ストレージコントローラ１００のＣＰＵ１２０が各ライトデータについて、シーケンシャルアクセスデータかランダムアクセスデータかを判定し（判定の方法には、ホスト計算機４０から連続領域に対するアクセス要求が到来しているか否かを判定する等の、公知の手法を用いることができる）、キャッシュＳＳＤ２００へライト要求する際に、ライトコマンド中に（具体的には後述するデータ情報３００５に）、「ランダムアクセスデータ」または「シーケンシャルアクセスデータ」のいずれかの情報を格納してＳＳＤコントローラ２０１へ送信する。シーケンシャルアクセスの場合、ライトデータに対する再アクセスが行われる可能性が低い（つまりキャッシュヒットする可能性が低い）。そのため、キャッシュＤＲＡＭ１３０２およびキャッシュＳＳＤ２００に格納されたシーケンシャルアクセスデータは、ヒットすることなく最終記憶媒体１９０にディステージされる。そのため、ＳＳＤコントローラ２０１は、シーケンシャルアクセスデータであると通知されたデータはキャッシュヒットする可能性が低いと予測し、そのデータの内部キャッシュ２００５１への格納時に、内部ディステージ優先度２００５２６を「低」と設定する。一方、ランダムアクセスデータの場合には内部ディステージ優先度２００５２６を「高」と設定し、シーケンシャルアクセスのデータに比べて、内部ディステージされやすくする（ＦＭへライトされやすくする）。これにより、相対的にシーケンシャルアクセスデータを、できるだけ内部キャッシュ２００５１に留めて、内部ディステージされにくくしている。

[0051] 続いて図５を用いて、キャッシュＳＳＤ２００の論理物理アドレス変換／更新（論物変換／論物更新）について説明する。前述したとおり、ＦＭ２０１１は同物理アドレスに対して上書きができない。そのため、キャッシュＳ

ＳＤ２００はライトコマンドを受け付けると、ライトコマンドで指定されている論理アドレスに対して新たな（未書き込みの）ＦＭ２０１１の物理アドレスを割り当て、当該割り当てられた物理アドレスに対してライトを行う。そのために、ストレージコントローラ１００から指定される論理アドレス２００５３１と、当該論理アドレス２００５３１に対して割り当てられたＦＭ２０１１上物理アドレス２００５３２のマッピングを、内部メモリ２００５上の論理物理アドレス変換テーブル２００５３に格納して、マッピングの管理を行う。ＳＳＤコントローラ２０１がリード／ライトを受けた際に、ＣＰＵ２００３が論理物理アドレス変換テーブル２００５３の参照、更新を行う。

[0052] ＳＳＤコントローラ２０１がリード要求を受けた場合、ＣＰＵ２００３は、論理物理アドレス変換テーブル２００５３を参照して、リード要求のあった論理アドレス２００５３１に対応する物理アドレス２００５３２を取得する。この処理を論物変換と呼ぶ。その後、ＣＰＵ２００３は、その物理アドレス２００５３２を用いて、該当するＦＭ２０１１に対してリード要求を行い、ＦＭ２０１１からデータをリードして、インタフェースコントローラ２００１を介してストレージコントローラ１００へリードデータを転送する。

[0053] ＳＳＤコントローラ２０１がライト要求を受けた場合、ＣＰＵ２００３はライトデータを格納するための、ＦＭ２０１１上の未書き込み領域を確保し、確保された領域のあるＦＭ２０１１にデータをライトする。その後、ＣＰＵ２００３は、論理物理アドレス変換テーブル２００５３を参照して、ライト要求のあった論理アドレス２００５３１に対応する物理アドレス２００５３２を、新たにデータを格納した領域の物理アドレスへと更新する。この処理を論物更新と呼ぶ。

[0054] なお、図４の説明の際に述べたとおり、キャッシュＳＳＤ２００では、データ処理アシストユニット２００６によるデータ処理（圧縮や暗号化）を行う際に、４ＫＢ単位でデータ処理を行うため、図５の論理物理アドレス変換テーブル２００５３における論理アドレスの管理も４ＫＢ単位で行う。

[0055] 次に、図６を用いてストレージコントローラ１００のＣＰＵ１２０からキャッシュＳＳＤ２００へ発行されるライトコマンド（ライト要求コマンド）について説明する。

[0056] ライト要求コマンド３０００には、コマンドを識別するコマンド番号（コマンドＮｏ）３００１、ライトであることを示す依頼内容３００２、ライト先の論理空間上先頭アドレス３００３、ライトデータのデータサイズ３００４、ＳＳＤコントローラ２０１がＦＭへのディステージ優先度を設定するためのヒントとなるデータ情報３００５が含まれる。依頼内容３００２には、コマンドの種類を表す情報（ライトであることを示す情報）や、ライト対象データが現在格納されている領域のアドレス（たとえばキャッシュ領域１３０２のアドレス）のほか、データ処理アシストユニット２００６によって、データ圧縮や暗号化などを行うか否かを指定する情報が含まれる。データ情報３００５には先に述べたとおり、「ランダムアクセスデータ」または「シーケンシャルアクセスデータ」のいずれかの情報が格納される。

[0057] キャッシュＳＳＤ２００によるライト処理が終了（正常終了または失敗）した後、ＳＳＤコントローラ２０１はライト要求コマンド３０００に対する応答コマンド４０００をストレージコントローラ１００のＣＰＵ１２０へ送信する。応答コマンド４０００は、コマンドを識別するコマンド番号４００１、完了または失敗などの情報を示す結果情報４００２から成る。コマンド番号４００１には、ライト要求コマンド３０００に含まれているコマンド番号３００１と同内容の情報が格納される。これにより、ストレージコントローラ１００のＣＰＵ１２０は、自身が発行したライト要求コマンド３０００によるライト要求が正常に終了したのか失敗したのかを判別できる。

[0058] 続けて、図７を用いてストレージコントローラ１００のＣＰＵ１２０からキャッシュＳＳＤ２００へのリードまたはデータ無効化要求するためのコマンドについて説明する。ストレージコントローラ１００のＣＰＵ１２０は、ＳＳＤコントローラ２０１にリードまたはデータ無効化要求コマンド５０００を送信する。リードまたはデータ無効化要求コマンド５０００は、コマン

ドを識別するコマンド番号５００１、リードまたはデータ無効化であることを示す依頼内容５００２、リードまたは無効化先の先頭アドレス５００３、リードまたは無効化要求サイズ５００４から成る。依頼内容５００２には、コマンドの種類を表す情報（リード要求であること、あるいは無効化要求であることを示す情報）のほか、データ処理アシストユニット２００６によって、リードデータを伸長あるいは復号化してストレージコントローラ１００に返送するか否かを指定する情報が含まれる。

[0059] キャッシュＳＳＤ２００が要求に対する処理を完了または失敗した後、応答コマンド６０００がＳＳＤコントローラ２０１からストレージコントローラ１００のＣＰＵ１２０へ送信される。応答コマンド６０００は、ライト要求コマンド３０００に対する応答コマンド４０００と同様、コマンドを識別するコマンド番号６００１、完了または失敗などの情報を示す結果情報６００２から成る。

[0060] 続いて図８を用いて、ホスト計算機４０からのライト要求を受け取ったときにストレージ装置１０で行われる処理の流れを説明する。ここでは、ストレージコントローラ１００のＣＰＵ１２０によるキャッシュ先選択について、キャッシュＤＲＡＭ１３０２を１次キャッシュとし、キャッシュＳＳＤ２００を２次キャッシュとする例について説明する。また、以降では断りのない限り、ホスト計算機４０からのリード、ライト要求で指定される領域が、ホスト計算機４０がアクセスする最終記憶媒体１９０（あるいは１以上の最終記憶媒体１９０から構成されるボリューム）の記憶空間上の４ＫＢ境界に一致している場合の例（これはデータ処理アシストユニット２００６によるデータ処理の最小単位が４ＫＢのため）について説明する。また、図８の処理は特に断りのない限り、ストレージコントローラ１００のＣＰＵ１２０が実行する処理である。

[0061] なお、ストレージコントローラ１００のＣＰＵ１２０がキャッシュＳＳＤ２００に発行するライト要求コマンド３０００、リード要求コマンド５０００と区別するため、以下ではホスト計算機４０から受信するライト要求のこ

とを、ホストライト要求と呼ぶこともあり、またホスト計算機４０から受信するリード要求のことを、ホストリード要求と呼ぶこともある。

[0062] ストレージコントローラ１００のＣＰＵ１２０は、ホスト計算機４０からホストライト要求を受けると（Ｓ１０）、キャッシュ管理テーブル１３０１を用いてキャッシュＤＲＡＭ１３０２のヒット判定を行う（Ｓ２０）。キャッシュＤＲＡＭ１３０２にヒットした、つまりキャッシュＤＲＡＭ１３０２上にホストライト要求で指定されたデータを格納するためのキャッシュセグメントが確保済みの場合（Ｓ２０：ヒット）、キャッシュＤＲＡＭ１３０２上の該当セグメントに対してライトデータの上書きを行う（Ｓ６０）。キャッシュＤＲＡＭ１３０２でミスした場合（Ｓ２０：ミス）、キャッシュＤＲＡＭ１３０２に空き（未使用）のセグメントがあるかを確認する（Ｓ３０）。空きがない場合（Ｓ３０：Ｎｏ）、空きセグメントを確保するため、ホストライト要求で指定されたデータを格納するために必要となるサイズ以上のデータを、キャッシュＤＲＡＭ１３０２から２次キャッシュであるキャッシュＳＳＤ２００へとディステージする（Ｓ４０）。このディステージ処理については後述する。

[0063] 空きがある場合（Ｓ３０：Ｙｅｓ）またはデータディステージ後（Ｓ４０の実行後）、キャッシュＤＲＡＭ１３０２に該当する新たなセグメントを確保してキャッシュＤＲＡＭ１３０２にライトデータを格納する（Ｓ５０）。Ｓ５０において、セグメントを確保すると、キャッシュ管理テーブル１３０１の内容を更新する。

[0064] キャッシュＤＲＡＭ１３０２にデータ格納後、キャッシュ管理テーブル１３０１を用いてキャッシュＳＳＤ２００にヒットしているか（ホストライト要求によってライトされるライト対象データを格納するためのキャッシュセグメントが確保済みか）を確認する（Ｓ７０）。ヒットしている場合（Ｓ７０：ヒット）には、ストレージコントローラ１００はキャッシュＳＳＤ２００に対して該当データの無効化を要求する（Ｓ８０）。なぜなら、ヒットしている場合とは、キャッシュＳＳＤ２００に、ホストライト要求によってラ

イトされるライト対象データの更新前データが格納されていることを意味するが、更新前データは不要なデータであるため無効化処理を行う。キャッシュＳＳＤ２００でのデータ無効化処理については後述する。

[0065] 無効化要求が完了、またはミス（Ｓ７０：ミス）で無効化の必要がない場合、ストレージコントローラ１００はホスト計算機４０へライト完了を通知する（Ｓ９０）。また、キャッシュ管理テーブル１３０１を用いたキャッシュＳＳＤ２００のヒット判定（Ｓ７０）については、Ｓ２０におけるキャッシュＤＲＡＭ１３０２のヒット判定の際にキャッシュＳＳＤ２００のヒット判定を行うようにしてもよい。

[0066] 図９を用いて、キャッシュＤＲＡＭ１３０２の空きを確保するために行われる、キャッシュＳＳＤ２００へのディステージ（Ｓ４０）処理について説明する。この処理も、ストレージコントローラ１００のＣＰＵ１２０によって実行される。

[0067] まずストレージコントローラ１００のＣＰＵ１２０は、キャッシュＤＲＡＭ１３０２からキャッシュＳＳＤ２００へディステージすべきデータを選択する（Ｓ４１）。この選択はすでに述べたように、キャッシュ管理テーブル１３０１を用いて、直近アクセス時刻（最終アクセス時刻）が最も古いデータから順に選択する。また、このディステージ処理は、ホストライト要求で指定されたデータを格納するために必要な領域を確保するための処理であるため、Ｓ４１では、ディステージすべきデータを複数個（少なくともホストライト要求で指定されたデータと同量またはそれより大きい量）選択することもある。

[0068] 続いてキャッシュＳＳＤ２００へディステージされるデータについて、キャッシュ管理テーブル１３０１を用いて、キャッシュＳＳＤ２００のヒット判定、つまりディステージ対象のデータを格納するためのセグメントがキャッシュＳＳＤ２００に確保済みか確認する処理を行う（Ｓ４２）。ミスである場合（Ｓ４２：ミス）、キャッシュＳＳＤ２００に空きセグメントがあるかを確認する（Ｓ４３）。空きがない場合（Ｓ４３：Ｎｏ）、キャッシュＳ

ＳＤ２００内のデータを最終記憶媒体１９０へとディステージする（Ｓ４４）。このキャッシュＳＳＤ２００から最終記憶媒体１９０へのディステージ処理については後述する。このディステージ処理後、またはキャッシュＳＳＤ２００に空きがあった場合（Ｓ４３：Ｙｅｓ）、キャッシュＳＳＤ２００に該当するキャッシュセグメントを新たに確保し（Ｓ４５）、キャッシュＳＳＤ２００に対してライト要求コマンド３０００を送信することで、キャッシュＤＲＡＭ１３０２からキャッシュＳＳＤ２００にディステージするデータのライトを行う（Ｓ４６）。このライト要求コマンドを送信する際、データ情報３００５に、「ランダムアクセスデータ」または「シーケンシャルアクセスデータ」のいずれかの情報を格納して送信する。Ｓ４７で、ライト要求の完了結果である応答コマンド４０００をキャッシュＳＳＤ２００から受領し、キャッシュＤＲＡＭ１３０２からのディステージ処理を完了する。

[0069] 図１０を用いて、キャッシュＳＳＤ２００の空きを確保するために行われる、最終記憶媒体１９０へのディステージ処理（Ｓ４４）の流れについて説明する。

[0070] ストレージコントローラ１００のＣＰＵ１２０は、キャッシュ管理テーブル１３０１を用いて、すでに述べたようにキャッシュＳＳＤ２００からディステージするデータをＬＲＵなどで選択する（Ｓ４４１）。続いてＣＰＵ１２０は、選択されたデータについて、キャッシュＳＳＤ２００からデータをリードし（Ｓ４４２、Ｓ４４３、Ｓ４４４）、リードしたデータを最終記憶媒体１９０へライトする（Ｓ４４５、Ｓ４４６、Ｓ４４７）。その後、データは最終記憶媒体１９０に格納されたため、ＣＰＵ１２０はキャッシュＳＳＤ２００上のデータを無効化し（Ｓ４４８、Ｓ４４９、Ｓ４４１０）、最後に無効化したキャッシュＳＳＤ２００上のセグメントについての情報をキャッシュ管理テーブル１３０１から削除することにより、セグメントの解放処理を行う（Ｓ４４１１）。キャッシュＳＳＤ２００からのリード（Ｓ４４３）およびキャッシュＳＳＤ２００上データの無効化要求（Ｓ４４９）については、図７に記載のコマンドを用いるが、この処理の詳細については後述す

る。

[0071] 図１１を用いて、キャッシュＳＳＤ２００で実行されるリード処理（Ｓ４４３）を説明する。

[0072] ＳＳＤコントローラ２０１のＣＰＵ２００３は、リード要求コマンド５０００を受領すると（Ｓ４４３１）、内部キャッシュ管理テーブル２００５２を用いて、内部キャッシュ２００５１のヒット判定を行う（Ｓ４４３２）。内部キャッシュ２００５１にヒットした場合（Ｓ４４３２：ヒット）、ＣＰＵ２００３は内部キャッシュ２００５１からデータをリードする（Ｓ４４３５）。内部キャッシュ２００５１でミスした場合（Ｓ４４３２：ミス）、ＣＰＵ２００３は前述の論物変換を行って（Ｓ４４３３）、ＦＭ２０１１からデータをリードする（Ｓ４４３４）。その後ＣＰＵ２００３は必要に応じてリードしたデータにキャッシュＳＳＤ２００のデータ処理機能を適用し（Ｓ４４３６）、ストレージコントローラ１００へデータを転送し（Ｓ４４３７）、ＣＰＵ１２０へ完了通知する（Ｓ４４３８）。Ｓ４４３６の処理では、ＣＰＵ２００３はデータ処理アシストユニット２００６を用いてデータ処理を行う。

[0073] 図１１の例では、キャッシュＳＳＤ２００の内部キャッシュ２００５１にリードデータをキャッシュしない。しかし、内部キャッシュ２００５１にリード時のデータをキャッシュしてもよい。その場合、ＦＭからデータをリードする前に、内部キャッシュ２００５１の空きを確認し、空きがなければ内部ディステージを行って空きを確保する。その後、空きセグメントを新たに確保し、ＦＭからリードするデータをストレージコントローラ１００へ転送するとともに、内部キャッシュ２００５１に格納する。また、本発明の実施例１に係るストレージ装置１０では、ストレージコントローラ１００が最終記憶媒体１９０にディステージするために、キャッシュＳＳＤ２００からデータを読み出す際には、データ変換／逆変換処理は行わせない。たとえば、キャッシュＳＳＤ２００に圧縮データが格納されている場合、圧縮されたままのデータを読み出して最終記憶媒体１９０にディステージするので、スト

レージコントローラ１００が最終記憶媒体１９０にディステージするために、キャッシュＳＳＤ２００に対してリード要求コマンド５０００を発行する場合、リード要求コマンドの依頼内容５００２には、リードデータの変換／逆変換は行わない旨が指定されている（そのため、上で説明したＳ４４３６の処理は行われない）。逆に、ストレージコントローラ１００がホスト計算機４０へ返送するためのリードデータをキャッシュＳＳＤ２００から読み出す場合（かつ、キャッシュＳＳＤ２００に、データ処理アシストユニット２００６によって変換処理が施されたデータが格納されている場合）には、データ逆変換（伸長や復号化等）を行ったデータを読み出す必要がある。そのため、ストレージコントローラ１００のＣＰＵ１２０がキャッシュＳＳＤ２００に対して発行するリード要求コマンド５０００の依頼内容５００２には、リードデータの逆変換を行う旨が指定されている（そのため、上で説明したＳ４４３６の処理が行われる）。

[0074] 続いて、図１２を用いて、キャッシュＳＳＤ２００がデータ無効化要求を受信した時に行われる処理（Ｓ４４９）の流れを説明する。

[0075] ＳＳＤコントローラ２０１のＣＰＵ２００３は、データ無効化要求コマンド５０００を受領すると（Ｓ４４９１）、内部キャッシュ管理テーブル２００５２を用いて、内部キャッシュ２００５１のヒット判定を行う（Ｓ４４９２）。内部キャッシュにヒットした場合（Ｓ４４９２：ヒット）、ＣＰＵ２００３は内部キャッシュ２００５１の該当セグメントを解放する（Ｓ４４９３）。セグメント解放後または内部キャッシュ２００５１でミスした場合（Ｓ４４９２：ミス）、ＣＰＵ２００３は前述の論物変換を行って（Ｓ４４９４）、論理物理アドレス変換テーブル２００５３の該当テーブルエントリを確認する（Ｓ４４９５）。これは、内部キャッシュ２００５１およびＦＭ２０１１の両方にデータが格納されていることがあり、その両方を無効化する必要があるため内部キャッシュヒット時にも論物変換を行う。ＦＭ２０１１にデータが格納されている場合は、該当テーブルエントリは有効であるため（Ｓ４４９５：有効）、ＣＰＵ２００３は該当物理アドレスのデータが無効

である旨をブロック情報２００５４に記録し（Ｓ４４９６）、さらに論理物理アドレス変換テーブル２００５３から、論理物理アドレス変換テーブル２００５３に記録されている該当物理アドレスの情報を削除する（Ｓ４４９７）。その後ＣＰＵ２００３は、ストレージコントローラ１００のＣＰＵ１２０へ完了通知を行い（Ｓ４４９８）、処理を終了する。

[0076] 続いて図１３を用いて、ストレージコントローラ１００からライト要求コマンド（Ｓ４６）を受信したときに、キャッシュＳＳＤ２００が実行する処理について説明する。

[0077] ＳＳＤコントローラ２０１のＣＰＵ２００３は、ライト要求コマンド３０００を受領すると（Ｓ４６１）、コマンド３０００の依頼内容３００２（たとえば圧縮の指示、暗号化の指示など）に従って、必要に応じてデータ処理アシストユニット２００６を用いて、ライトデータにデータ変換処理を施す（Ｓ４６２）。なお、本発明の実施例１に係るストレージ装置１０では、ストレージコントローラ１００がホスト計算機４０から受信したライトデータをキャッシュＳＳＤ２００に格納する場合には、キャッシュＳＳＤ２００にデータ変換を行わせる。そのため、ストレージコントローラ１００がホスト計算機４０から受信したライトデータをキャッシュＳＳＤ２００に格納するために、ライト要求コマンド３０００を発行する場合には、当該ライト要求コマンドの依頼内容３００２にはデータの変換を行う旨が指定されている。一方、最終記憶媒体１９０に格納されていたデータをキャッシュＳＳＤ２００に格納（ステージング）する場合には、キャッシュＳＳＤ２００にデータ変換／逆変換処理は行わせない。そのため、この場合にストレージコントローラ１００のＣＰＵ１２０がキャッシュＳＳＤ２００に発行するライト要求コマンドの依頼内容３００２には、データの変換／逆変換は行わない旨が指定されている（そのため、上で説明したＳ４６２の処理は行われない）。

[0078] 続いてＣＰＵ２００３は内部キャッシュ管理テーブル２００５２を用いて、内部キャッシュ２００５１のヒット判定を行う（Ｓ４６３）。内部キャッシュにヒットした場合（Ｓ４６３：ヒット）、ＣＰＵ２００３は内部キャッ

シュの該当セグメントにデータを上書きする（Ｓ４６７）。内部キャッシュ２００５１でミスした場合（Ｓ４６３：ミス）、ＣＰＵ２００３は内部キャッシュ２００５１に空きの（未使用の）セグメントがあるか確認する（Ｓ４６４）。空きがない場合（Ｓ４６４：Ｎｏ）、空きセグメントを確保するため、ＣＰＵ２００３は内部キャッシュ２００５１上のデータをＦＭ２０１１に内部ディステージする（Ｓ４６５）。内部ディステージ処理については後述する。なお、Ｓ４６４の判定において、未使用のセグメントがなくなった場合に内部ディステージ（Ｓ４６５）処理を実施する態様に限定されるわけではない。たとえば、未使用の領域サイズ（セグメント数）が不足した場合（ストレージコントローラ１００からライト要求のあったデータのサイズよりも未使用の領域のサイズが小さい場合、あるいは未使用の領域のサイズが所定の閾値未満になった場合など）に内部ディステージ（Ｓ４６５）処理を行うようにしてもよい。

[0079] 内部ディステージ後、または内部キャッシュ２００５１に空きがある場合（Ｓ４６４：Ｙｅｓ）、ＣＰＵ２００３は内部キャッシュ２００５１のセグメントを新たに確保し（Ｓ４６６）、内部キャッシュ２００５１にデータをライトする（Ｓ４６７）。なお、Ｓ４６７では、ＣＰＵ２００３はＳ４６１で受信したライト要求コマンド３００のデータ情報３００５に含まれている「ランダムアクセスデータ」または「シーケンシャルアクセスデータ」の情報に基づいて、ライトデータの格納されるキャッシュセグメントに対応する、内部キャッシュ管理テーブル２００５２のディステージ優先度２００５２６の欄に、「高」または「低」を格納する。その後ＣＰＵ２００３は、ストレージコントローラ１００のＣＰＵ１２０に応答コマンド４０００を送信することでライト完了通知を行う（Ｓ４６８）。

[0080] 続いて図１４を用いて、キャッシュＳＳＤ２００で行われる内部ディステージ処理（Ｓ４６５）について説明する。

[0081] 内部キャッシュ２００５１からデータをＦＭ２０１１へ内部ディステージする必要がある場合、ＳＳＤコントローラ２０１のＣＰＵ２００３は、内部

ディステージするデータを選択する（Ｓ４６４１）。先に述べた通り、内部ディステージ対象データの選択は、一例として、内部キャッシュ管理テーブル２００５２の直近アクセス時間２００５２５とディステージ優先度２００５２６を用いて、ディステージ優先度２００５２６が「高」であるデータの中から直近アクセス時間２００５２５の最も古いものを選択することによって行う。そしてディステージ優先度２００５２６が「高」であるデータが存在しない場合には、ディステージ優先度２００５２６が「低」のデータの中から直近アクセス時間２００５２５の最も古いものを選択する。ただし、これ以外のデータ選択方法を採用してもよい。

[0082] 内部ディステージするデータを選択後、ＣＰＵ２００３は内部ディステージ先となるＦＭ２０１１の空き（未使用）領域（物理アドレス）を探す（Ｓ４６４２）。内部ディステージ対象データを格納可能な空き領域（物理アドレス）が複数存在する場合、ＣＰＵ２００３はＦＭ２０１１の寿命等を考慮して適切な物理アドレスを選択する。そしてＣＰＵ２００３は、選択した書き先物理アドレスに対応するＦＭ２０１１へのライト要求を行い（Ｓ４６４３）、書き先物理アドレスのデータが有効である旨をブロック情報２００５４に書き込む（Ｓ４６４４）。さらにＣＰＵ２００３は、要求された論理アドレスに書き先物理アドレスに対応付けるために論理物理アドレス変換テーブル２００５３を更新する（Ｓ４６４５）。以上でデータをＦＭ２０１１に格納したので、ＣＰＵ２００３は内部キャッシュ管理テーブル２００５２を更新して内部キャッシュ２００５１の該当セグメントを解放し（Ｓ４６４６）、内部ディステージ処理を終了する。

[0083] 次に、図１５を用いて、ホスト装置４０からのリード要求（ホストリード要求）を受信した時に、ストレージ装置１０が行う処理について説明する。

[0084] ストレージコントローラ１００のＣＰＵ１２０は、ホスト計算機４０からホストリード要求を受領すると（Ｓ１００）、キャッシュ管理テーブル１３０１を用いて、キャッシュヒット判定を行う（Ｓ１１０、Ｓ１２０）。キャッシュＤＲＡＭ１３０２にヒットした場合（Ｓ１１０：ヒット）、ＣＰＵ１

２０はキャッシュＤＲＡＭ１３０２からデータをリードする（Ｓ１９０）。キャッシュＤＲＡＭ１３０２にヒットせず（Ｓ１１０：ミス）、キャッシュＳＳＤ２００にヒットした場合（Ｓ１２０：ヒット）、ＣＰＵ１２０はキャッシュＳＳＤ２００へリード要求コマンドを発行する（Ｓ１４０）。キャッシュＳＳＤ２００で行われるリード処理は、Ｓ４４３で行われる処理と同様で、データが読み出される過程で、キャッシュＳＳＤ２００はデータにキャッシュＳＳＤ２００内部のデータ処理機能を適用し、データに対する処理（データ伸長やデータ復号化）を行う。そして処理が施されたデータがストレージコントローラ１００のバッファ領域１３０３にに転送される。

[0085] キャッシュＳＳＤ２００でもミスした場合（Ｓ１２０：ミス）、ＣＰＵ１２０は最終記憶媒体１９０からデータをリードし、メモリ１３０、ＡＳＩＣ１４０、あるいはＩＦ　ＣＴＬ１５０内の一時記憶領域にデータを格納する（Ｓ１３０）。さらにＣＰＵ１２０は、キャッシュＳＳＤ２００に空きセグメントが存在するか確認し（Ｓ１５０）、空きがないなら（Ｓ１５０：Ｎｏ）、キャッシュＳＳＤ２００から最終記憶媒体１９０へディステージ処理を行う（Ｓ１６０）。なお、このキャッシュＳＳＤ２００の空き確認とディステージ処理（Ｓ１５０、Ｓ１６０）については、最終記憶媒体１９０からのリード（Ｓ１３０）の前に行うようにしてもよい。ディステージ処理（Ｓ１６０）は、Ｓ４４と同様である。ディステージ後、またはキャッシュＳＳＤ２００に空きがある場合（Ｓ１５０：Ｙｅｓ）、ＣＰＵ１２０はキャッシュＳＳＤ２００に該当するキャッシュセグメントを新たに確保し、最終記憶媒体１９０からリードしたデータをライトする要求をキャッシュＳＳＤ２００に発行する（Ｓ１７０）。キャッシュＳＳＤ２００へのライト処理はＳ４６と同様である。続いてＳ１７０でＣＰＵ１２０は、キャッシュＳＳＤ２００に書き込んだデータをリードする要求を、キャッシュＳＳＤ２００に発行する（Ｓ１８０）。このリード要求を受け付けたキャッシュＳＳＤ２００ではＳ４４３と同様、データに対する処理（データ伸長やデータ復号化）を行い、処理が施されたデータがストレージコントローラ１００のバッファ領域１

３０３にに転送される。最後に、ストレージコントローラ１００はいずれかの媒体からリードしたデータをホスト計算機４０へ転送し（Ｓ２００）、処理を終了する。

[0086] 上で述べたように、いずれのキャッシュでもミスし、最終記憶媒体１９０からリードした場合（Ｓ１３０）、ＣＰＵ１２０は最終記憶媒体１９０からリードしたデータを一旦キャッシュＳＳＤ２００にライトし、その後キャッシュＳＳＤ２００にライトしたデータをリードする。最終記憶媒体１９０上のデータは、ライト時にキャッシュＳＳＤ２００のデータ処理機能が適用されて、圧縮や暗号化等の処理がなされているので、一度キャッシュＳＳＤ２００に格納し、再度データ処理機能を適用してデータ伸長あるいは復号化を行うことでホスト計算機４０が扱えるデータにする必要があるためである。そのため、ホスト計算機４０に送信するデータを、最終記憶媒体１９０から一旦キャッシュＳＳＤ２００にライトする必要があるが、内部キャッシュ２００５１にヒットする、或いは内部キャッシュ２００５１の空き領域に書き込める可能性が高いため、ＦＭ２０１１へのライトが発生する可能性は低く、リードレイテンシも小さい。

[0087] なお、たとえばデータ処理機能によりデータ暗号化・復号化を行う場合、データ暗号化・復号化のために必要な暗号鍵などの情報については、ライト時からキャッシュＳＳＤ２００が保持し続けてもよいし、ストレージコントローラ１００が管理してもよい。ストレージコントローラ１００が管理する場合、キャッシュＳＳＤ２００へのライト要求コマンド３０００、リード要求コマンド５０００によって必要な情報をキャッシュＳＳＤ２００に通知する。あるいはライト要求コマンド３０００やリード要求コマンド５０００とは異なる、別のコマンドを発行することで、必要な情報をキャッシュＳＳＤ２００に通知するようにしてもよい。

[0088] また、上では、ホスト計算機４０からのホストリード要求、ホストライト要求で指定される領域が、４ＫＢ境界に一致している場合の例について説明したが、ホスト計算機４０からのリード、ライト要求で指定される領域が４

ＫＢ境界に一致しておらず、かつキャッシュＤＲＡＭあるいはキャッシュＳＳＤにアクセス対象のデータが格納されていない（キャッシュミス）の場合には、最終記憶媒体１９０からアクセス対象の範囲を含む、４ＫＢ境界に一致したサイズのデータをキャッシュＤＲＡＭあるいはキャッシュＳＳＤに読み出して、キャッシュされたデータに対して、リード、ライト処理を行えばよい。

実施例 2

[0089] 実施例２に係るストレージ装置１０及びキャッシュＳＳＤ２００のハードウェア構成は、実施例１で説明したものと同じであるため、ここでは説明しない。また、実施例２に係るストレージ装置１０及びキャッシュＳＳＤ２００で実施される大半の処理も、実施例１で説明したものと同じであるため、ここでは、実施例１と異なる点（具体的には、実施例１における、キャッシュＳＳＤ２００の内部ディステージ処理（図１３：Ｓ４６５、図１４）、そしてストレージコントローラ１００のＣＰＵ１２０がキャッシュＳＳＤ２００にライト要求を発行する際の処理（図９：Ｓ４６）が、実施例２では異なる処理になる）を中心に説明する。

[0090] 実施例１に係るストレージ装置１０及びキャッシュＳＳＤ２００では、ディステージ優先度２００５２６を高く設定したデータが優先的にＦＭ２０１１にライトされる。これにより、ディステージ優先度２００５２６が低いデータが極力ＦＭ２０１１にライトされないようにし、ＦＭ２０１１の寿命劣化を最小限に抑える。ただし、ディステージ優先度２００５２６が高いデータがキャッシュＳＳＤ２００の内部キャッシュ２００５１に存在しない場合には、ディステージ優先度２００５２６が低いデータもＦＭ２０１１にライトされる。一方、実施例２に係るストレージ装置１０及びキャッシュＳＳＤ２００では、ディステージ優先度２００５２６が高いデータのみをＦＭ２０１１にライトするように制御する。

[0091] しかし、そのままでは内部ディステージ優先度２００５２６が低いデータが内部キャッシュ２００５１を占有してしまうため、ストレージコントロー

ラ１００がキャッシュＳＳＤ２００から最終記憶媒体１９０へデータをディステージする必要がある。そのためにキャッシュＳＳＤ２００からストレージコントローラ１００のＣＰＵ１２０へディステージが必要であることを通知する。この通知は、キャッシュＳＳＤ２００がライト要求を受領したタイミングで行うが、それ以外の契機で通知するようにしてもよい。例えば、ＳＳＤコントローラ２００が内部キャッシュ２００５１の空き容量を定期的に監視し、内部キャッシュ２００５１の空き容量が小さくなった際にストレージコントローラ１００のＣＰＵ１２０にディステージが必要であると通知してもよい。

[0092] 図１６は、実施例２におけるキャッシュＳＳＤ２００で行われる内部ディステージ処理のフローの一例である。

[0093] ＳＳＤコントローラ２０１のＣＰＵ２００３は、内部キャッシュ２００５１内のデータについて、内部ディステージ優先度２００５２６が高いデータの有無を確認し（Ｓ４６４１１）、そのようなデータがない場合（Ｓ４６４１１：Ｎｏ）、ＳＳＤコントローラ２０１から最終記憶媒体１９０へのディステージが必要である旨、及び内部キャッシュ２００５１内のデータのうち、最終アクセス時刻（内部キャッシュ管理テーブル２００５２で管理しているデータの最終アクセス時刻２００５２５）が最も古いデータ（キャッシュセグメントに格納されているデータ）の情報を、ライト要求コマンド３０００に対する応答コマンド４０００の結果情報４００２に含めて、ストレージコントローラ１００のＣＰＵ１２０へ通知し（Ｓ４６４１２）、キャッシュＳＳＤ２０の内部ディステージ処理は終了する。なお、詳細は後述するが、ストレージコントローラ１００のＣＰＵ１２０は、ＳＳＤコントローラ２０１から返送された応答コマンド４０００の結果情報４００２に、最終記憶媒体１９０へのディステージが必要である旨の情報が含まれていた場合、キャッシュＳＳＤ２００内のデータを最終記憶媒体１９０へディステージする処理を実施する。

[0094] 一方、内部ディステージ優先度２００５２６が高いデータがある場合（Ｓ

４６４１１：Ｙｅｓ）、ＣＰＵ２００３は実施例１における内部ディステージ処理（Ｓ４６４１～Ｓ４６４６）と同様の処理を行う。

[0095] 続いて図１７を用いて、ストレージコントローラ１００のＣＰＵ１２０がキャッシュＳＳＤ２００へライト要求を発行したときに、ストレージコントローラ１００のＣＰＵ１２０の実行する処理の流れを説明する。

[0096] 実施例１と同様に、ストレージコントローラ１００のＣＰＵ１２０がキャッシュＳＳＤ２００へライト要求コマンド３０００を送信することでライト要求（Ｓ４６）した後、キャッシュＳＳＤ２００からライト要求コマンド３０００に対する応答コマンド４０００でライト完了通知を受領する（Ｓ４７）。本実施例では、ＣＰＵ１２０はその結果情報４００２からライト完了またはディステージが必要なのかを確認する（Ｓ４８）。結果情報４００２に、最終記憶媒体１９０へのディステージが必要である旨の情報が含まれていた場合（Ｓ４８：ディステージ要）、ＣＰＵ１２０はキャッシュＳＳＤ２００からデータを最終記憶媒体１９０へディステージする（Ｓ４９）。このディステージ処理は実施例１のＳ４４（つまり図１０に記載の処理）とほぼ同様であるが、ＣＰＵ１２０はディステージするデータとして、キャッシュＳＳＤ２００から通知された、内部キャッシュ２００５１内のデータのうち最終アクセス時刻が最も古いデータを選択するようにする。これは、内部キャッシュ２００５１のデータをディステージしなければ、キャッシュＳＳＤ２００に新たにデータを書き込めないからである。ディステージ後、ＣＰＵ１２０は再度キャッシュＳＳＤ２００へライト要求を行い、ライトが完了するまで上記の処理を繰り返す。

[0097] 実施例２に係るストレージ装置１０及びキャッシュＳＳＤ２００では、ディステージ優先度が低いデータのＦＭ２０１１への書き込みが完全に抑止される。そのため、キャッシュＳＳＤの記憶媒体（ＦＭ）に書き込まれるデータ量を削減することができ、ＦＭ２０１１の寿命を延ばすことができる。

実施例 3

[0098] 実施例１または２に係るストレージ装置１０では、ストレージコントロー

ラ１００上のメモリ１３０（具体的にはメモリ１３０内のキャッシュＤＲＡＭ１３０２）を１次キャッシュとし、キャッシュＳＳＤ２００を２次キャッシュとして用いていたが、実施例３に係るストレージ装置１０では、１次キャッシュを用いない。つまりキャッシュＳＳＤ２００のみをストレージ装置１０のキャッシュとして用いる点が、実施例１または２に係るストレージ装置１０と異なる。

[0099] 実施例３に係るストレージ装置１０の構成は、実施例１で説明したもの（図１に記載のもの）とほぼ同様であるため、図示は省略する。ただし実施例３に係るストレージ装置１０の場合、メモリ１３０内にメモリキャッシュ領域１３０２が存在しない点が、実施例１に係るストレージ装置１０と異なる。その他の点は実施例１で説明したものと同じである。

[0100] 続いて実施例３に係るストレージ装置１０が管理する、キャッシュ管理テーブル１３０１について説明する。実施例３に係るキャッシュ管理テーブル１３０１に含まれる情報は、実施例１におけるものと同じであるので図示は省略する。ただし、実施例１におけるキャッシュ管理テーブル１３０１には、キャッシュＤＲＡＭ１３０２のキャッシュセグメントについての情報（図２：１３１０）が格納されていたが、実施例３に係るストレージ装置１０では、キャッシュとして用いる記憶領域は、キャッシュＳＳＤ２００だけであるため、キャッシュＳＳＤのキャッシュセグメントについての情報（図２：１３１１）だけが格納され、キャッシュＤＲＡＭ１３０２のキャッシュセグメントについての情報（図２：１３１０）は格納されない。

[0101] なお、実施例３に係るキャッシュＳＳＤ２００の構成、及びキャッシュＳＳＤ２００が内部メモリ２００５で管理する情報（内部キャッシュ管理テーブル、論理物理アドレス変換テーブル２００５３、ブロック情報２００５４）は、実施例１に係るキャッシュＳＳＤ２００と同一であるため、ここでは説明を行わない。

[0102] 続いて図１８を用いて、実施例３に係るストレージ装置１０が、ホスト計算機４０からライト要求を受信した時の処理の流れを説明する。この処理は

、一般的なディスクキャッシュ付きストレージ装置が行う処理と同様の処理である。

[0103] ストレージコントローラ１００のＣＰＵ１２０は、ホスト計算機４０からライト要求を受信する（Ｓ１０）。この処理は図８を用いて説明した処理Ｓ１０と同じである。続いてストレージコントローラ１００は、キャッシュ管理テーブル１３０１を用いてキャッシュＳＳＤ２００のヒット判定を行う（Ｓ２０’）。キャッシュＳＳＤ２００にヒットした、つまりキャッシュＳＳＤ２００上に、ライト要求で指定されたデータを格納するためのキャッシュセグメントが確保済みの場合（Ｓ２０’：ヒット）、キャッシュＳＳＤ２００上の該当セグメントに対してライトデータの上書きを行う（Ｓ６０’）。キャッシュＳＳＤ２００でミスした場合（Ｓ２０’：ミス）、キャッシュＳＳＤ２００に空き（未使用）のセグメントがあるかを確認する（Ｓ３０’）。空きがない場合（Ｓ３０’：Ｎｏ）、空きセグメントを確保するため、キャッシュＳＳＤ２００から最終記憶媒体１９０へデータのディステージを行う（Ｓ４０’）。このディステージ処理は、実施例１で説明した図１０の処理（図９：Ｓ４４の処理）と同じ処理である。

[0104] 空きがある場合（Ｓ３０’：Ｙｅｓ）またはデータディステージ後（Ｓ４０’の実行後）、キャッシュＳＳＤ２００に新たなセグメントを確保してキャッシュＳＳＤ２００にライトデータを格納する（Ｓ５０’）。この処理は、実施例１で説明した、図９のＳ４５、Ｓ４６と同じ処理である。キャッシュＳＳＤ２００にデータ格納後、ストレージコントローラ１００はホスト計算機４０へライト完了を通知し（Ｓ９０）、処理を終了する。

[0105] 次に、図１９を用いて、ホスト装置４０からのリード要求を受信した時のストレージ装置１０の行う処理について説明する。この処理は、図１５を用いて説明した、実施例１に係るストレージ装置１０が実施するリード処理から、キャッシュＤＲＡＭ１３０２に対する処理を取り除いたものである。

[0106] 最初にストレージコントローラ１００のＣＰＵ１２０は、ホスト計算機４０からリード要求を受領すると（Ｓ１００）、キャッシュ管理テーブル１３

０１を用いて、キャッシュヒット判定を行う（Ｓ１２０）。キャッシュＳＳＤ２００にヒットした場合（Ｓ１２０：ヒット）、ＣＰＵ１２０はキャッシュＳＳＤ２００へリード要求する（Ｓ１４０）。キャッシュＳＳＤ２００のリード処理は、実施例１で説明したＳ４４３の処理（つまり図１１の処理）と同様である。キャッシュＳＳＤ２００でキャッシュミスの場合（Ｓ１２０：ミス）、ＣＰＵ１２０は最終記憶媒体１９０からデータをリードして、メモリ１３０、ＡＳＩＣ１４０、あるいはＩＦ　ＣＴＬ１５０内の一時記憶領域に格納する（Ｓ１３０）。

[0107] さらにＣＰＵ１２０は、キャッシュＳＳＤ２００に空きセグメントが存在するか確認し（Ｓ１５０）、空きがない場合（Ｓ１５０：Ｎｏ）、キャッシュＳＳＤ２００から最終記憶媒体１９０へデータのディステージ処理を行う（Ｓ１６０）。このキャッシュＳＳＤ２００の空き確認とディステージ処理（Ｓ１５０、Ｓ１６０）については、最終記憶媒体１９０からのリード（Ｓ１３０）の前に行うようにしてもよい。ディステージ処理（Ｓ１６０）は、Ｓ４４と同様である。ディステージ後、またはキャッシュＳＳＤ２００に空きがある場合（Ｓ１５０：Ｙｅｓ）、ＣＰＵ１２０はキャッシュセグメントを新たに確保し、最終記憶媒体１９０からリードしたデータをキャッシュＳＳＤ２００にライトする（Ｓ１７０）。キャッシュＳＳＤ２００へのライト処理はＳ４６と同様である。このライト処理の過程で、データにキャッシュＳＳＤ２００内部のデータ処理機能が適用される。続いて、Ｓ１７０でＣＰＵ１２０はキャッシュＳＳＤ２００に書き込んだデータをリードする要求を、キャッシュＳＳＤ２００に発行する（Ｓ１８０）。このリード要求を受け付けたキャッシュＳＳＤ２００ではＳ４４３と同様、データに対する処理（データ伸長やデータ復号化）を行い、処理が施されたデータがストレージコントローラ１００のバッファ領域１３０３に転送される。最後にＣＰＵ１２０は、いずれかの媒体からリードしたデータをホスト計算機４０へ転送し（Ｓ２００）、処理を終了する。

[0108] なお、実施例３に係るキャッシュＳＳＤ２００でも、実施例１と同様に、

内部キャッシュ領域２００５１に空きセグメントが存在しない場合、図１４に記載された内部ディステージ処理が行われる。内部ディステージ処理では実施例１と同様、ディステージ優先度２００５２６の高いデータの中から、ＬＲＵアルゴリズムにより内部ディステージするデータを選択することにより、シーケンシャルアクセスデータのように、再アクセスされる可能性が低いデータをＦＭ２０１０に格納（内部ディステージ）されにくくしている。

[0109] また、実施例３に係るストレージ装置１０においても、実施例２において説明した内部ディステージ処理（図１６の処理）を採用してもよい。そうすると、実施例３に係るストレージ装置１０は、実施例２に係るストレージ装置１０と同様に、ディステージ優先度２００５２６が高いデータのみをＦＭ２０１１にライトするように動作する。

実施例 4

[0110] 続いて実施例４の説明を行う。実施例４に係るストレージ装置及びキャッシュＳＳＤのハードウェア構成は、実施例１または２で説明したものと同じであるため、ここでは説明しない。実施例１または２に係るストレージ装置１０では、ストレージコントローラ１００上のメモリ１３０（具体的にはメモリ１３０内のキャッシュＤＲＡＭ１３０２）を１次キャッシュとし、キャッシュＳＳＤ２００を２次キャッシュとして用いていたが、実施例４では、ストレージコントローラ１００のＣＰＵがデータ特性（例えば、ライトデータの最終的な格納先）に応じて、キャッシュＳＳＤ２００又はキャッシュＤＲＡＭ１３０２のいずれか一方を選択してデータをキャッシュする。ライトデータの最終的な格納先については、ホストライト要求に含まれるライト対象データのアドレス情報（ＬＢＡ等）を参照することで判断できる。

[0111] 最終記憶媒体１９０が、例えばＳＳＤを始めとする高速な媒体である場合、アクセス速度が同等であるキャッシュＳＳＤ２００にキャッシュしても、アクセス性能向上の効果が小さいため、ストレージコントローラ１００のメモリ１３０上のキャッシュＤＲＡＭ１３０２にのみ格納する。一方、最終記憶媒体１９０がＨＤＤのような、ＳＳＤよりも低速な媒体である場合、キャ

ッシュＳＳＤ２００にのみキャッシュする。この場合では、キャッシュＤＲＡＭ１３０２に格納されたデータについては、キャッシュＳＳＤ２００が備えるデータ処理機能を用いることができないため、ストレージコントローラ１００のＣＰＵ１２０または別ハードウェアで同様の処理を行う。

[0112] キャッシュＳＳＤ２００にキャッシュされたデータの扱いは実施例３等と同等となるので詳細は省略する。

[0113] 実施例４では、最終記憶媒体の種類などの第一のデータ特性に基づき、キャッシュＤＲＡＭ１３０２又はキャッシュＳＳＤ２００のいずれか一方にライトデータをキャッシュする。キャッシュＳＳＤ２００に格納されたデータに関して、キャッシュＳＳＤ２００が、ストレージコントローラ１００のＣＰＵ１２０から受信した第二のデータ特性（例えば、ライトアクセスの種類（シーケンシャル／ランダム））に基づき、内部キャッシュからＦＭへのディステージの優先度を決定する。

[0114] 図２０を用いて、ホスト装置４０からのライト要求を受信した時のストレージ装置１０の行う処理について説明する。

[0115] ストレージコントローラ１００のＣＰＵ１２０は、ホスト計算機４０からライト要求を受信する（Ｓ１０）。この処理は図８を用いて説明した処理Ｓ１０と同じである。続いてＣＰＵ１２０は、ライト要求コマンドのＬＢＡを確認して、ライトデータの最終格納先がＳＳＤのような高速な媒体かＨＤＤのような低速な媒体かを判定する（Ｓ１１）。最終格納先がＳＳＤであれば（Ｓ１１：ＳＳＤ）、ライトデータをキャッシュＤＲＡＭ１３０２へ格納する（Ｓ１３）。このキャッシュＤＲＡＭ１３０２への格納処理については、図８のＳ７０、Ｓ８０を除いた処理となる。なぜなら、キャッシュＤＲＡＭ１３０２とキャッシュＳＳＤ２００は同じＬＢＡのデータを保持しないため、キャッシュＤＲＡＭ１３０２上のデータがキャッシュＳＳＤ２００上にはないからである。

[0116] 一方、最終格納先がＨＤＤであれば（Ｓ１１：ＨＤＤ）、ライトデータをキャッシュＳＳＤ２００に格納する（Ｓ１２）。キャッシュＳＳＤ２００へ

の格納処理は、図９のＳ４２以降と同様である。

[0117] 実施例４に係るストレージ装置によれば、キャッシュＳＳＤ２００には、ＨＤＤのように、ＳＳＤよりもアクセス性能の低い記憶媒体が用いられた最終記憶媒体１９０に格納されるデータのみがキャッシュされ、ＳＳＤ等のようにキャッシュＳＳＤ２００と同等のアクセス性能の記憶媒体が用いられた最終記憶媒体１９０に格納されるデータはキャッシュされない。そのため、キャッシュＳＳＤ２００に書き込まれるデータ量を削減することができ、キャッシュＳＳＤ２００の記憶媒体（ＦＭ２０１１）の寿命を延ばすことができる。

実施例 5

[0118] 続いて実施例５の説明を行う。実施例５に係るストレージ装置及びキャッシュＳＳＤのハードウェア構成は、実施例１または２で説明したものと同じである。実施例５に係るストレージ装置及びキャッシュＳＳＤ２００では、キャッシュＳＳＤ２００のＦＭ２０１１の劣化が進み、データを保持し続けることが難しくなった場合に、キャッシュＳＳＤ２００が内部ディステージを行わずに、実施例２と同様にストレージコントローラ１００へディステージを要求して、データを最終記憶媒体１９０へディステージする。このとき、ＦＭ２０１１に格納されているデータのうち、キャッシュ状態１３０１４が「ｄｉｒｔｙ」のデータについても、ストレージコントローラ１００へディステージ要求する。これにより、キャッシュＳＳＤ２００がキャッシュ可能な容量は小さくなるが、ＦＭ２０１１の劣化によってデータが消失してしまうことを防止しつつ、ライトデータまたは最終記憶媒体１９０からリードされたデータに対して、キャッシュＳＳＤ２００の持つデータ変換等の機能を引き続き適用できる。

[0119] 図２１を用いて、キャッシュＳＳＤ２００のＦＭ２０１１が寿命に到達した場合の動作を説明する。図２１は、キャッシュＳＳＤ２００がストレージコントローラ１００からライト要求コマンドを受信したときに、キャッシュＳＳＤ２００で実行される処理の流れを説明する図であり、多くの処理は実

施例１の図１３と共通している。

[0120] キャッシュＳＳＤ２００は、実施例１と同様に、ストレージコントローラ１００のＣＰＵ１２０からライト要求を受け取ると（Ｓ４６１）、データ処理を行い（Ｓ４６２）、内部キャッシュ管理テーブル２００５２を用いて、内部キャッシュ２００５１のヒット判定を行う（Ｓ４６３）。内部キャッシュにヒットした場合（Ｓ４６３：ヒット）、内部キャッシュの該当セグメントに上書きライトする（Ｓ４６７）。

[0121] 内部キャッシュ２００５１でミスした場合（Ｓ４６３：ミス）、内部キャッシュ２００５１に空きの（未使用の）セグメントがあるか確認する（Ｓ４６４）。内部キャッシュ２００５１に空きがある場合（Ｓ４６４：Ｙｅｓ）、実施例１と同様に新たに内部キャッシュセグメントを確保し、データを内部キャッシュへライトする。一方、内部キャッシュ２００５１に空きのセグメントがない場合（Ｓ４６４：Ｎｏ）、ＦＭ２０１１が寿命に到達したかを判定する（Ｓ４６９）。ここでの判定方法として例えば、ＳＳＤコントローラ２０１のＣＰＵ２００３が、上位装置からの要求とは無関係にデータを定期的にリードして、データに付与されているＣＲＣなどの保証コードを用いてエラー率を確認し、エラー率が予め設定してある閾値を越えた場合に寿命に到達したと判定する、という方法があり得る。

[0122] ＦＭ２０１１が寿命に到達していない場合（Ｓ４６９：Ｎｏ）、他の実施例で説明したライト処理と同様に内部ディステージを行って（Ｓ４６５）、空きセグメントを確保する。空きセグメントを確保するための内部ディステージを行うことができない場合、すなわちＦＭ２０１１が寿命に到達した場合（Ｓ４６９：Ｙｅｓ）、内部ディステージを行うとデータを消失する危険がある。そこで、実施例２と同様に、キャッシュＳＳＤ２００からストレージコントローラ１００のＣＰＵ１２０へディステージが必要であることを要求する（Ｓ４６４１２）。ただし、それ以外の契機でディステージが必要であることをＣＰＵ１２０へ通知するようにしてもよい。実施例１と同様に、ＳＳＤコントローラ２００が内部キャッシュ２００５１の空き容量を定期的

に監視し、内部キャッシュ２００５１の空き容量が小さくなった際にストレージコントローラ１００のＣＰＵ１２０にディステージが必要であると通知してもよい。さらに、キャッシュＳＳＤ２００はＦＭ２０１１が寿命に到達したことをストレージコントローラ１００に通知してもよい。

[0123] キャッシュＳＳＤ２００のＦＭ２０１１が寿命に到達した旨の通知を受けたストレージコントローラ１００のＣＰＵ１２０は、すでにキャッシュＳＳＤ２００に格納してあるデータについて、キャッシュ状態１３０１４が「ｄｉｒｔｙ」であるデータを最終記憶媒体１９０へディステージする。また、キャッシュ状態１３０１４が「ｃｌｅａｎ」であるデータについては、キャッシュＳＳＤ２００に格納したままでは読み出せない可能性があるため、それらのデータについては、キャッシュ管理テーブル１３０１を更新し、キャッシュＳＳＤ２００に格納されていないものとして扱う。具体的には、キャッシュの管理情報を、キャッシュＳＳＤ２００に記録したデータが消失したものとして更新する。以降、キャッシュＳＳＤ２００に格納されるデータは、内部キャッシュ２００５１に格納できないものについては、上述の通り、ストレージコントローラ１００にディステージが必要であることを通知し、ストレージコントローラ１００にディステージしてもらうため、寿命に到達したＦＭ２０１１にデータが格納されることはない。

[0124] 本発明の実施例５に係るストレージ装置及びキャッシュＳＳＤ２００によれば、ストレージコントローラ１００のＣＰＵ１２０が、キャッシュＳＳＤ２００からＦＭ２０１１の寿命に基づく情報を取得し、キャッシュＳＳＤ２００のＤＲＡＭなどからなる内部メモリ２００５のデータを最終記憶媒体１９０へディステージすることで、信頼性を高めることを可能とする。

実施例 6

[0125] 続いて実施例６について説明する。実施例６に係るストレージ装置及びキャッシュＳＳＤのハードウェア構成も、実施例１または２で説明したものと同様である。実施例６は、ストレージ装置が、複数の最終記憶媒体１９０をＲＡＩＤグループとして管理し、ホスト計算機４０からのライトデータから

ＲＡＩＤ　Ｐａｒｉｔｙ（以下、「Ｐａｒｉｔｙ」と略記することもある）を生成し、ライトデータとともにＰａｒｉｔｙを最終記憶媒体１９０に格納する構成を対象とする。

[0126] また実施例６に係るストレージ装置１０では、キャッシュＳＳＤ２００がストレージコントローラ１００からの指示に基づいて、キャッシュＳＳＤ２００に格納されたデータ（ホスト計算機４０から書き込まれたライトデータ）からＲＡＩＤ　Ｐａｒｉｔｙを生成する。この構成によれば、ストレージコントローラ１００からＲＡＩＤ　Ｐａｒｉｔｙ生成に係る負荷をオフロードすることができ、ストレージ装置１０全体の性能を向上させる事が可能となる。そのため、実施例６に係るキャッシュＳＳＤは、Ｐａｒｉｔｙを生成する機能を備える。実施例１においても述べたとおり、この機能はデータ処理アシストユニット２００６に備えられていてもよいし、あるいはデータ処理アシストユニット２００６とは異なる、Ｐａｒｉｔｙ生成用のハードウェアとしてキャッシュＳＳＤ２００に実装されていてもよい。あるいは、ＣＰＵ２００３がＰａｒｉｔｙ生成を実行する構成であってもよい。

[0127] 実施例６に係るストレージ装置では、ストレージコントローラ１００はホスト計算機４０からのライトデータをキャッシュＳＳＤ２００に格納した後、ＳＳＤに対しＰａｒｉｔｙ生成コマンドを発行する。

[0128] このＰａｒｉｔｙ生成コマンドは少なくとも、Ｐａｒｉｔｙ生成対象とするデータの格納位置を指定するＬＢＡ（これは実施例１で説明した、キャッシュＳＳＤがストレージコントローラに提供している論理的な記憶空間上のアドレスである）を複数個、生成されるＰａｒｉｔｙの格納先を指定するＬＢＡを１または２個（ＲＡＩＤ５であれば１つ、ＲＡＩＤ６であれば２つ）、そしてデータ長の情報を有している。

[0129] このＰａｒｉｔｙ生成コマンドをストレージコントローラ１００のＣＰＵ１２０から受領したキャッシュＳＳＤ２００は、コマンドにて指定されたＬＢＡに対応付けられたデータを、ＦＭ２０１１または内部キャッシュ２００５１より取得し、Ｐａｒｉｔｙ演算を実施する。そして取得されたデータを

用いてＰａｒｉｔｙを生成し、内部キャッシュ２００５１に格納し、Ｐａｒｉｔｙ生成コマンドにて指定されたＬＢＡに対応付けて管理する（具体的には、内部キャッシュ管理テーブル２００５２を用いて管理しておく）。

[0130] ストレージ装置は一般に、キャッシュの空き容量が低下し、新たなデータをキャッシュに載せる事ができない場合に、アクセス頻度の低いデータをキャッシュからＨＤＤなどの最終記憶媒体に転送するディステージ処理を行い、ディステージによって最終記憶媒体に記録されたデータをキャッシュから削除することで空き領域を生成する。またストレージ装置は一般に、ディステージされるべきと決定されたデータ（ディステージ対象データ）に対してＰａｒｉｔｙを生成するため、生成されたＰａｒｉｔｙもディステージ対象データと同様に、生成後、短期間で最終記憶媒体に転送され、キャッシュから消去される。

[0131] このことから、本発明の実施例６に係るキャッシュＳＳＤでも、上で説明したＰａｒｉｔｙ生成により生成されたＰａｒｉｔｙは、短期間でディステージされることが期待される。そのためキャッシュＳＳＤ２００は、生成されたＰａｒｉｔｙを内部キャッシュ２００５１に格納する際、当該Ｐａｒｉｔｙが内部ディステージ優先度の低いデータであるとして管理する（内部キャッシュ管理テーブル２００５２の、内部ディステージ優先度２００５２６を「Ｌｏｗ」として記録する）。

[0132] これにより、ストレージコントローラがＰａｒｉｔｙを最終記憶媒体１９０にディステージし、キャッシュＳＳＤ２００からのデータ（Ｐａｒｉｔｙ）消去が指示されるまでの期間、Ｐａｒｉｔｙは優先的にＤＲＡＭ２００５（内部キャッシュ２００５１）に格納され、ＦＭ２０１１に書き込まれる（内部ディステージされる）確率を低下させることで、ＦＭ２０１１へのライト量を減らし、ＦＭの劣化を軽減できる。

[0133] 以上で、本発明の実施例に係るストレージ装置及びキャッシュＳＳＤの説明を終了する。本発明の実施例に係るストレージ装置では、ＤＲＡＭ等の高速の記憶媒体から成る内部キャッシュと、ＦＭのような低価格で大容量の記

憶媒体を主たる記憶領域とするキャッシュ装置（キャッシュＳＳＤ）を備える。ＦＭは、データ消去回数（書き換え回数）に制限があるという欠点があるため、高頻度の書き替えが行われると、寿命が短くなってしまう。

[0134] 本発明の実施例に係るストレージ装置では、キャッシュ装置がデータの属性・特性に基づいて、キャッシュ対象データの内部ディステージの優先度を制御する。一例として、ストレージコントローラ等の上位装置がキャッシュ装置にデータを格納（キャッシュ）する際に、データの属性・特性に関する情報（たとえばシーケンシャルアクセス、ランダムアクセスである等）をキャッシュ装置に通知することにより、キャッシュ装置はデータの属性・特性を得る。これにより、たとえばシーケンシャルライトデータのように、上位装置から再アクセスされることなく追い出されてしまうデータ（キャッシュする効果が小さいデータ）が、内部キャッシュからＦＭにデータ移動（内部ディステージ）されることを抑制できるので、ＦＭへのライト頻度を低下させ、キャッシュ装置の寿命（耐用年数）を延ばすことが可能である。

[0135] また、内部キャッシュからＦＭへのデータ移動（内部ディステージ）の優先度とは異なる基準（ＬＲＵなど）に基づき、データをキャッシュ装置から最終記憶媒体へディステージすることにより、内部キャッシュからＦＭを介さずに最終記憶媒体へディステージすることを可能とし、ＦＭへのライト頻度を抑制している。

[0136] また、ＦＭのような書き替え回数に制限のある記憶媒体を用いたキャッシュ装置を長寿命化させる場合、キャッシュする効果が小さいデータをキャッシュ装置に格納しないように制御する方法も考えられるが、そのように制御すると、本発明の実施例に係るキャッシュ装置（キャッシュＳＳＤ）のように、ライトデータの圧縮や暗号化などを行うデータ処理機能を有している場合、キャッシュする効果が小さいデータに対して、キャッシュ装置のデータ処理機能を適用できない。本発明の実施例に係るストレージ装置及びキャッシュ装置では、キャッシュする効果が小さいデータを極力ＦＭに格納しないように制御できるため、キャッシュする効果が小さいデータに対してもキャ

ッシュ装置のデータ処理機能を適用でき、かつキャッシュ装置の長寿命化が可能である。

[0137] なお、実施例６で説明したとおり、本発明はストレージコントローラがデータの属性を指定し、その属性に基づいて、キャッシュ装置が内部ディステージの優先度を変更する例に限定されるものではない。実施例６に係るキャッシュＳＳＤのように、キャッシュ装置がデータの種別によって自律的にディステージ優先度を変更する態様であっても、ＦＭへのライト頻度を低下させ、キャッシュ装置の寿命（耐用年数）を延ばすという効果を得ることが可能である。

符号の説明

[0138] １０：ストレージ装置

２０：ネットワーク

４０：ホスト計算機

１００：ストレージコントローラ

１１０：Ｉ／Ｆ　ＣＴＬ

１２０：ＣＰＵ

１３０：メモリ

１４０：ＡＳＩＣ

１５０：Ｉ／Ｆ　ＣＴＬ

１６０：専用内部バス

１７０：内部バス

１８０：メモリバス

１９０：最終記憶媒体

２００：キャッシュＳＳＤ

２００１：Ｉ／Ｆ　ＣＴＬ

２００３：ＣＰＵ

２００４：内部スイッチ

２００５：内部メモリ

２００６：データ処理アシストユニット（Ａｓｓｉｓｔ）

２００７：ＦＭＣ

２０１１：ＦＭ

請求の範囲

[請求項1] １以上の最終記憶装置と、ホスト計算機及び前記最終記憶装置とが接続されるストレージコントローラを有するストレージ装置であって、

前記ストレージコントローラは、少なくともプロセッサと、前記ホスト計算機からのライトデータを一時的に格納するキャッシュ装置を有し、

前記キャッシュ装置は、書き替え回数に制限のある不揮発記憶媒体と、前記不揮発記憶媒体に比べて高速なアクセスが可能な揮発性記憶媒体から成る内部キャッシュを有し、

前記ストレージコントローラは、前記キャッシュ装置に前記ホスト計算機からのライトデータを格納する際、前記ライトデータの格納を指示するためのライト要求であって、前記ライトデータの特性に関する情報が含まれている前記ライト要求を前記キャッシュ装置に発行し、

前記キャッシュ装置は、前記ライト要求を受信すると、前記ライト要求に伴うライトデータを前記内部キャッシュに格納するとともに、前記ライトデータの特性に関する情報から前記ライトデータの内部ディステージ優先度を決定し、前記内部ディステージ優先度を前記ライトデータに対応付けて記憶し、

前記キャッシュ装置は前記内部キャッシュに格納された前記ライトデータのうち、前記ライトデータに対応付けられた前記内部ディステージ優先度に基づき、前記不揮発記憶媒体へと移動する、

ことを特徴とする、ストレージ装置。

[請求項2] 前記ストレージコントローラは、前記ライトデータの特性に関する情報として、ライトデータがシーケンシャルアクセスデータであるかランダムアクセスデータであるかを示す情報を、前記ライト要求に含ませて、前記キャッシュ装置に発行し、

前記キャッシュ装置は、前記ライト要求を受信すると、前記ライト要求に伴うライトデータを前記内部キャッシュに格納するとともに、前記ライトデータの特性に関する情報がシーケンシャルアクセスデータであることを示す情報であった場合、前記ライト要求に伴うライトデータの前記内部ディステージ優先度は低いと決定し、前記ライトデータの特性に関する情報がランダムアクセスデータであることを示す情報であった場合、前記ライト要求に伴うライトデータの前記内部ディステージ優先度は高いと決定することを特徴とする、請求項１に記載のストレージ装置。

[請求項3] 前記キャッシュ装置は、前記ライトデータに対して最後にアクセスがあった時刻を前記ライトデータに対応付けて記憶しており、

前記内部キャッシュの未使用領域の量が所定の閾値未満になった時、前記内部キャッシュに格納された前記ライトデータの中に、前記ライトデータに対応付けられた前記内部ディステージ優先度の高いライトデータが存在しなかった場合、前記ライトデータに対応付けられた前記内部ディステージ優先度が低い前記ライトデータの中から、前記ライトデータに最後にアクセスがあった時刻が最も古いデータを前記不揮発記憶媒体へと移動することを特徴とする、請求項２に記載のストレージ装置。

[請求項4] 前記キャッシュ装置は、前記ライトデータに対して最後にアクセスがあった時刻を前記ライトデータに対応付けて記憶しており、

前記キャッシュ装置は、前記内部キャッシュの未使用領域の量が所定の閾値未満になった時、前記内部キャッシュに格納された前記ライトデータの中に、前記ライトデータに対応付けられた前記内部ディステージ優先度の高いライトデータが存在しなかった場合、前記ストレージコントローラに、前記ライトデータの前記最終記憶装置へのディステージが必要である旨を通知し、

前記ストレージコントローラは、前記通知に応じて、前記キャッシ

ュ装置内の内部キャッシュに格納された前記ライトデータの中から、前記ライトデータに最後にアクセスがあった時刻が最も古いデータを読み出して、前記最終記憶装置へ書き込むことを特徴とする、請求項２に記載のストレージ装置。

[請求項5] 前記キャッシュ装置は、前記前記ストレージコントローラからライトデータを受信すると、前記ライトデータに対して変換処理を施して、前記キャッシュ装置に格納し、

前記ストレージコントローラが前記キャッシュ装置に格納された前記変換処理が施されたデータを前記最終記憶装置へディステージする際には、前記変換処理が施されたデータを読み出して、前記最終記憶装置に格納することを特徴とする、

請求項１に記載のストレージ装置。

[請求項6] 前記ストレージコントローラは、前記ホスト計算機からデータリード要求を受け付けると、

前記最終記憶装置から前記変換処理が施されたデータを読み出して、前記キャッシュ装置に格納し、

前記キャッシュ装置に対し、前記キャッシュ装置に格納された前記変換処理が施されたデータを読み出すリード要求であって、データ逆変換を行う旨が指定された前記リード要求を発行し、

前記キャッシュ装置は、前記リード要求を受信すると、前記変換処理が施されたデータを逆変換して前記ストレージコントローラに送信することを特徴とする、

請求項５に記載のストレージ装置。

[請求項7] 前記変換処理は、データの圧縮処理であって、前記変換処理が施されたデータを逆変換する処理は、圧縮データの伸長処理であることを特徴とする、

請求項６に記載のストレージ装置。

[請求項8] 前記変換処理は、データの暗号化処理であって、前記変換処理が施

されたデータを逆変換する処理は、暗号化データの復号化処理であることを特徴とする、

請求項６に記載のストレージ装置。

[請求項9] 前記ストレージ装置はさらに、前記ホスト計算機からのライトデータを一時的に格納するためのメモリを有し、

前記ストレージコントローラは、前記ホスト計算機からホストライト要求を受信すると、前記ホストライト要求で対象とされるライト対象データを格納するための空き領域が前記メモリに存在するか判定し、

前記空き領域が前記メモリに存在しない場合、前記メモリ内に格納されているデータのうち、最後にアクセスがあった時刻が最も古いデータを前記キャッシュ装置へ移動した後で空き領域を確保し、

前記確保された空き領域に前記ライト対象データを格納する、

ことを特徴とする、請求項１に記載のストレージ装置。

[請求項10] 前記ストレージコントローラは、前記確保された空き領域に前記ライト対象データを格納した後、前記ライト対象データの更新前データが前記キャッシュ装置に格納されているか判定し、

前記ライト対象データの更新前データが前記キャッシュ装置に格納されている場合、前記キャッシュ装置に前記更新前データの無効化を要求することを特徴とする、

請求項９に記載のストレージ装置。

[請求項11] １以上の最終記憶装置と、ホスト計算機及び前記最終記憶装置とが接続されるストレージコントローラを有するストレージ装置の制御方法であって、

前記ストレージコントローラは、少なくともプロセッサと、前記ホスト計算機からのライトデータを一時的に格納するキャッシュ装置を有し、

前記キャッシュ装置は、書き替え回数に制限のある不揮発記憶媒体

と、前記不揮発記憶媒体に比べて高速なアクセスが可能な揮発性記憶媒体から成る内部キャッシュを有するキャッシュ装置であって、

前記方法は、

前記ストレージコントローラが前記キャッシュ装置に前記ホスト計算機からのライトデータを格納する際、前記ライトデータの格納を指示するためのライト要求であって、前記ライトデータの特性に関する情報が含まれている前記ライト要求を前記キャッシュ装置に発行し、

前記キャッシュ装置は前記ライト要求を受信すると、

前記内部キャッシュに、前記ライト要求に伴うライトデータを格納するための未使用の領域が存在するか判定し、

前記未使用の領域が存在しない場合、前記内部キャッシュに格納されたデータのうち、前記データに対応付けられた内部ディステージ優先度の高いデータを、前記不揮発記憶媒体へと移動した後、前記ライト要求に伴うライトデータを格納するための前記未使用の領域を確保して、

前記ライト要求に伴うライトデータを前記確保された前記未使用の領域に格納するとともに、前記ライトデータの特性に関する情報から前記ライトデータの内部ディステージ優先度を決定し、前記内部ディステージ優先度を前記ライトデータに対応付けて記憶する、

処理を実行することを特徴とする、ストレージ装置の制御方法。

[請求項12] 前記ストレージコントローラは、前記ライトデータの特性に関する情報として、ライトデータがシーケンシャルアクセスデータであるかランダムアクセスデータであるかを示す情報を、前記ライト要求に含ませて、前記キャッシュ装置に発行し、

前記キャッシュ装置は、前記ライト要求に伴うライトデータを前記確保された前記未使用の領域に格納する際、前記ライトデータの特性に関する情報がシーケンシャルアクセスデータであることを示す情報であった場合、前記ライト要求に伴うライトデータの前記内部ディス

テージ優先度は低いと決定し、前記ライトデータの特性に関する情報がランダムアクセスデータであることを示す情報であった場合、前記ライト要求に伴うライトデータの前記内部ディステージ優先度は高いと決定することを特徴とする、

請求項１１に記載のストレージ装置の制御方法。

[請求項13] 前記キャッシュ装置は、前記ライトデータに対して最後にアクセスがあった時刻を前記ライトデータに対応付けて記憶しており、

前記内部キャッシュに、前記ライト要求に伴うライトデータを格納するための未使用の領域が存在せず、かつ前記内部キャッシュに格納された前記ライトデータの中に、前記ライトデータに対応付けられた前記内部ディステージ優先度の高いライトデータが存在しなかった場合、

前記キャッシュ装置は、前記ライトデータに対応付けられた前記内部ディステージ優先度が低い前記ライトデータの中から、前記ライトデータに最後にアクセスがあった時刻が最も古いデータを前記不揮発記憶媒体へと移動する処理を行うことを特徴とする、

請求項１２に記載のストレージ装置の制御方法。

[請求項14] 前記キャッシュ装置は、前記ライトデータに対して最後にアクセスがあった時刻を前記ライトデータに対応付けて記憶しており、

前記キャッシュ装置は、前記内部キャッシュに、前記ライト要求に伴うライトデータを格納するための未使用の領域が存在せず、かつ前記内部キャッシュに格納された前記ライトデータの中に、前記ライトデータに対応付けられた前記内部ディステージ優先度の高いライトデータが存在しなかった場合、

前記キャッシュ装置は、前記ストレージコントローラに、前記ライトデータの前記最終記憶装置へのディステージが必要である旨を通知し、

前記ストレージコントローラは、前記通知に応じて、前記キャッシ

ュ装置内の内部キャッシュに格納された前記ライトデータの中から、前記ライトデータに最後にアクセスがあった時刻が最も古いデータを読み出して、前記最終記憶装置へ書き込む処理を行うことを特徴とする、

請求項１２に記載のストレージ装置の制御方法。

[図1]

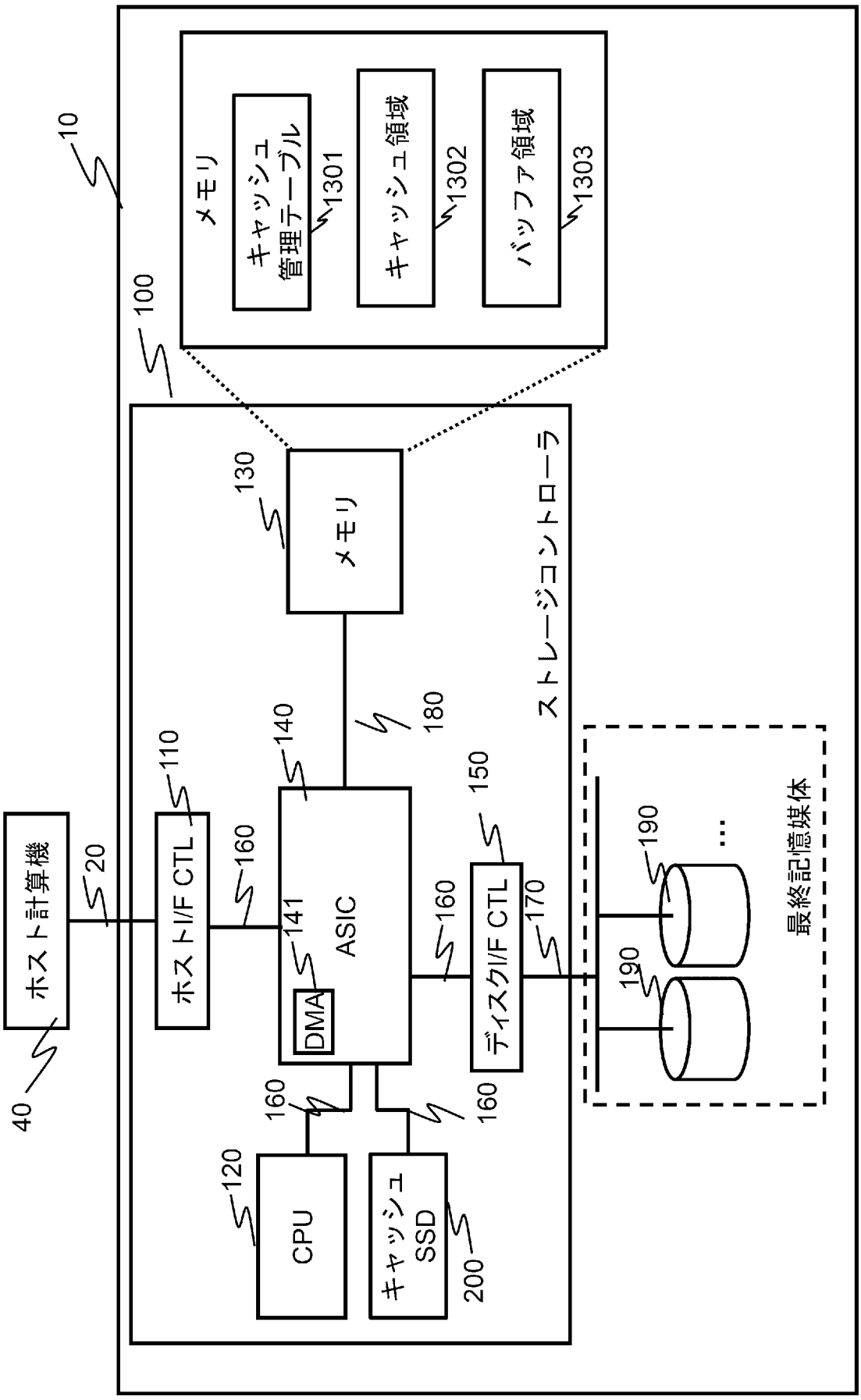

10
100
メモリ
キャッシュ管理テーブル
1301
キャッシュ領域
1302
バッファ領域
1303
ストレージコントローラ
130
メモリ
180
140
ASIC
141
DMA
110
ホストI/F CTL
160
150
ディスクI/F CTL
170
160
160
160
ホスト計算機
20
40
120
CPU
キャッシュSSD
200
190
190
...
最終記憶媒体

図1

[図2]

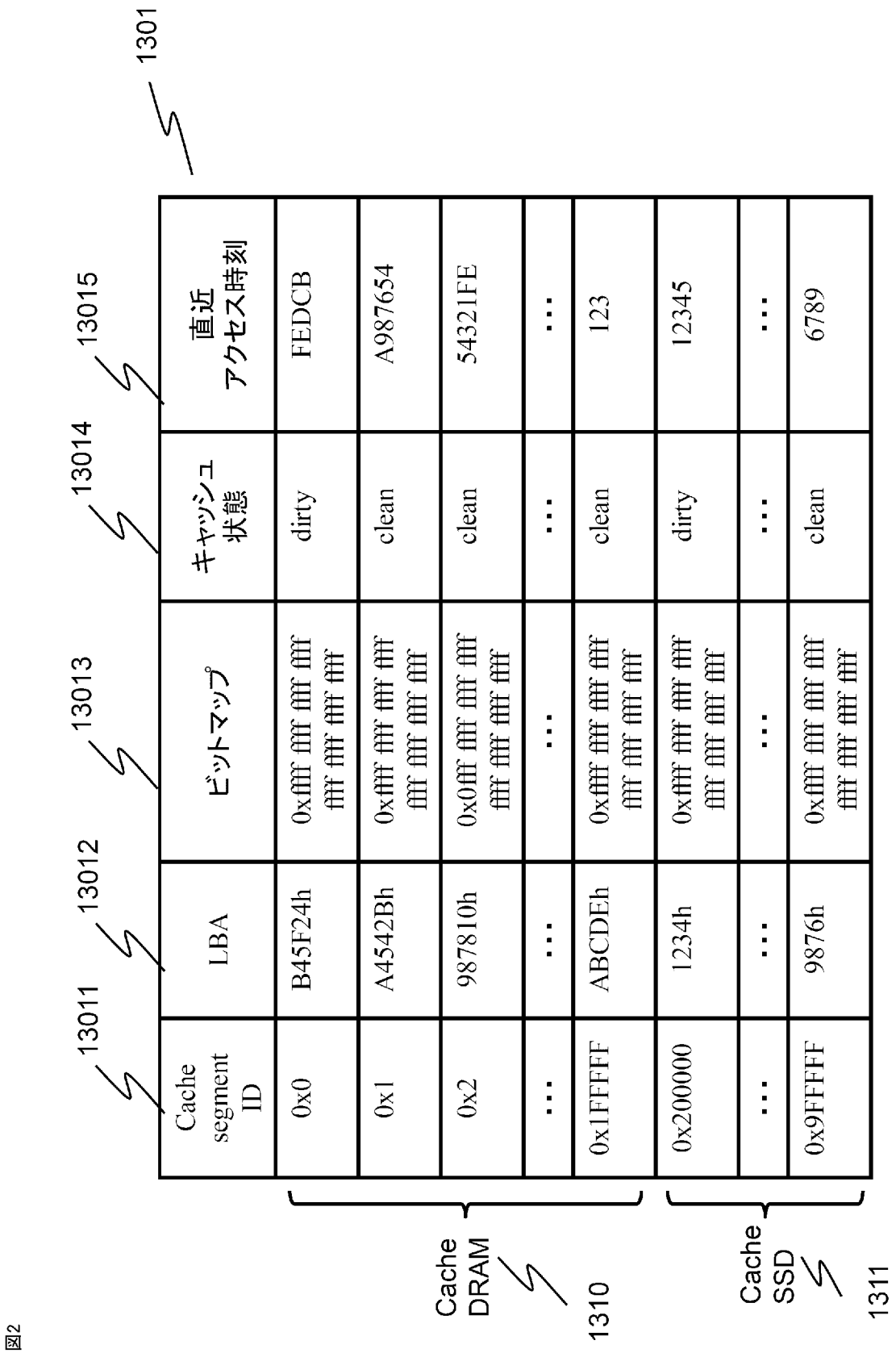

1301
13011
13012
13013
13014
13015
Cache segment ID
LBA
ビットマップ
キャッシュ状態
直近アクセス時刻
0x0
B45F24h
0xffff ffff ffff ffff ffff ffff ffff ffff
dirty
FEDCB
0x1
A4542Bh
0xffff ffff ffff ffff ffff ffff ffff ffff
clean
A987654
0x2
987810h
0x0fff ffff ffff ffff ffff ffff ffff ffff
clean
54321FE
...
...
...
...
...
0x1FFFFF
ABCDEh
0xffff ffff ffff ffff ffff ffff ffff ffff
clean
123
0x200000
1234h
0xffff ffff ffff ffff ffff ffff ffff ffff
dirty
12345
...
...
...
...
...
0x9FFFFF
9876h
0xffff ffff ffff ffff ffff ffff ffff ffff
clean
6789
Cache DRAM
1310
Cache SSD
1311

図2

[図3]

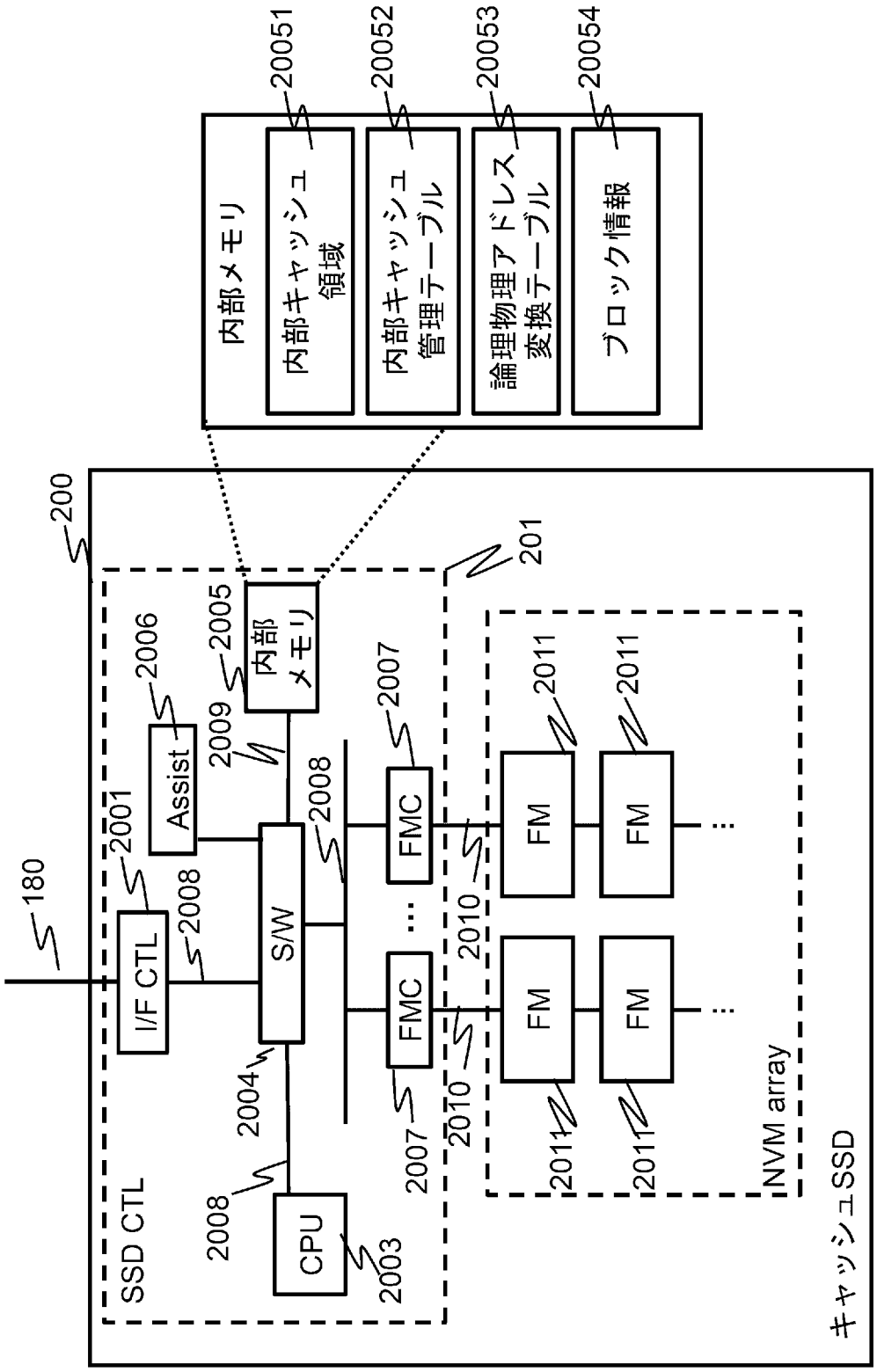

20051
20052
20053
20054
内部メモリ
内部キャッシュ領域
内部キャッシュ管理テーブル
論理物理アドレス変換テーブル
ブロック情報
200
201
180
SSD CTL
I/F CTL
Assist
内部メモリ
S/W
CPU
FMC
FM
NVM array
キャッシュSSD
2001
2003
2004
2005
2006
2007
2008
2009
2010
2011

図3

[図4]

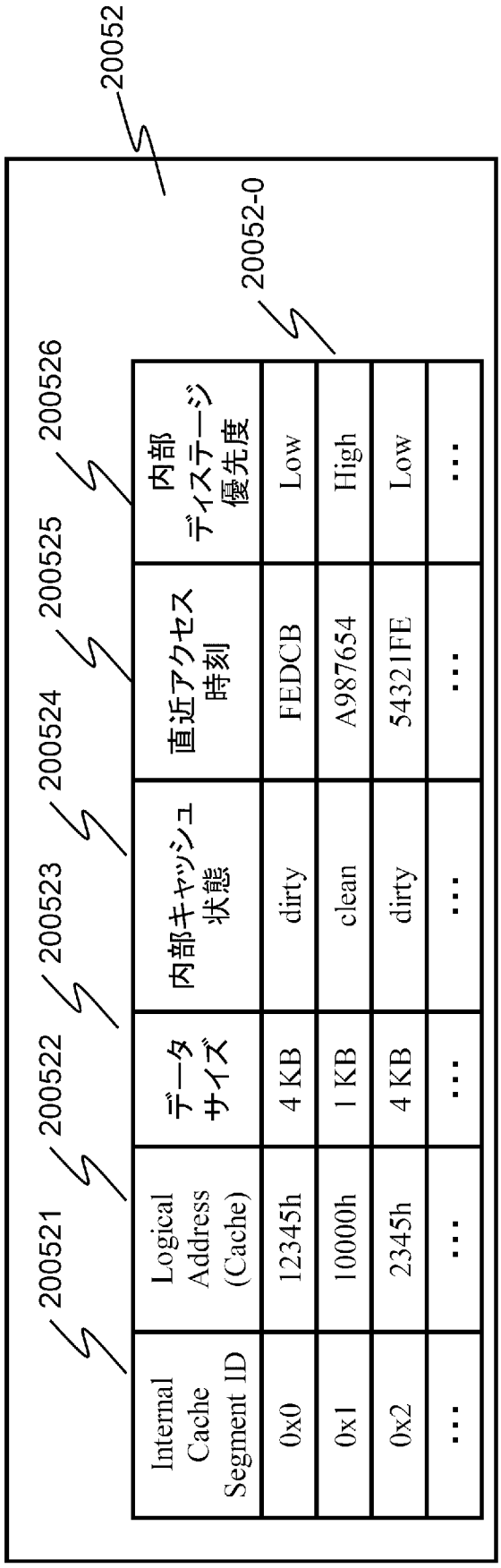

20052
20052-0
200521
200522
200523
200524
200525
200526
Internal Cache Segment ID
Logical Address (Cache)
データサイズ
内部キャッシュ状態
直近アクセス時刻
内部ディステージ優先度
0x0
12345h
4 KB
dirty
FEDCB
Low
0x1
10000h
1 KB
clean
A987654
High
0x2
2345h
4 KB
dirty
54321FE
Low
...
...
...
...
...
...

図4

[図5]

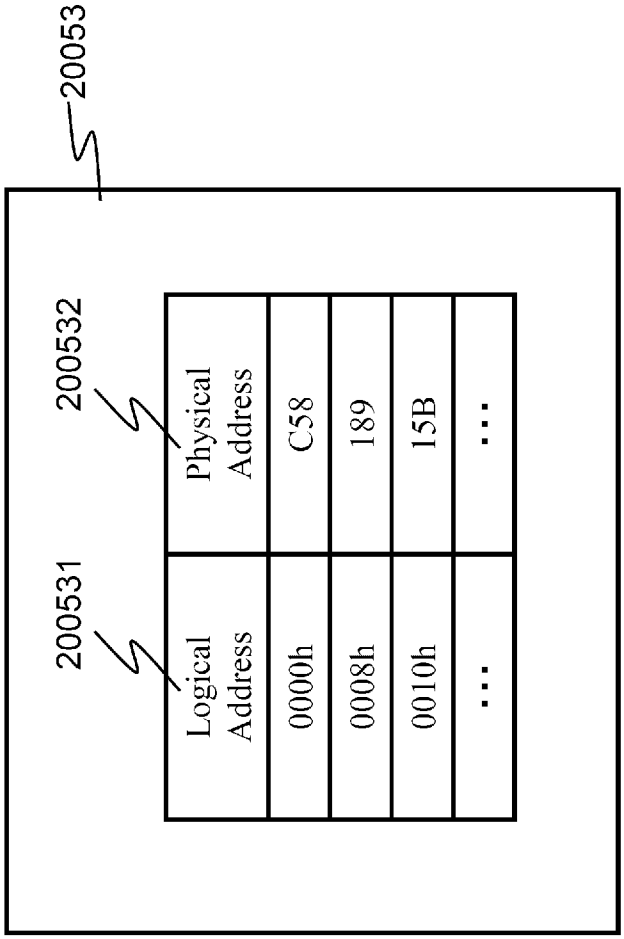

20053
200532
200531
Logical Address
Physical Address
0000h
0008h
0010h
...
C58
189
15B
...

図5

[図6]

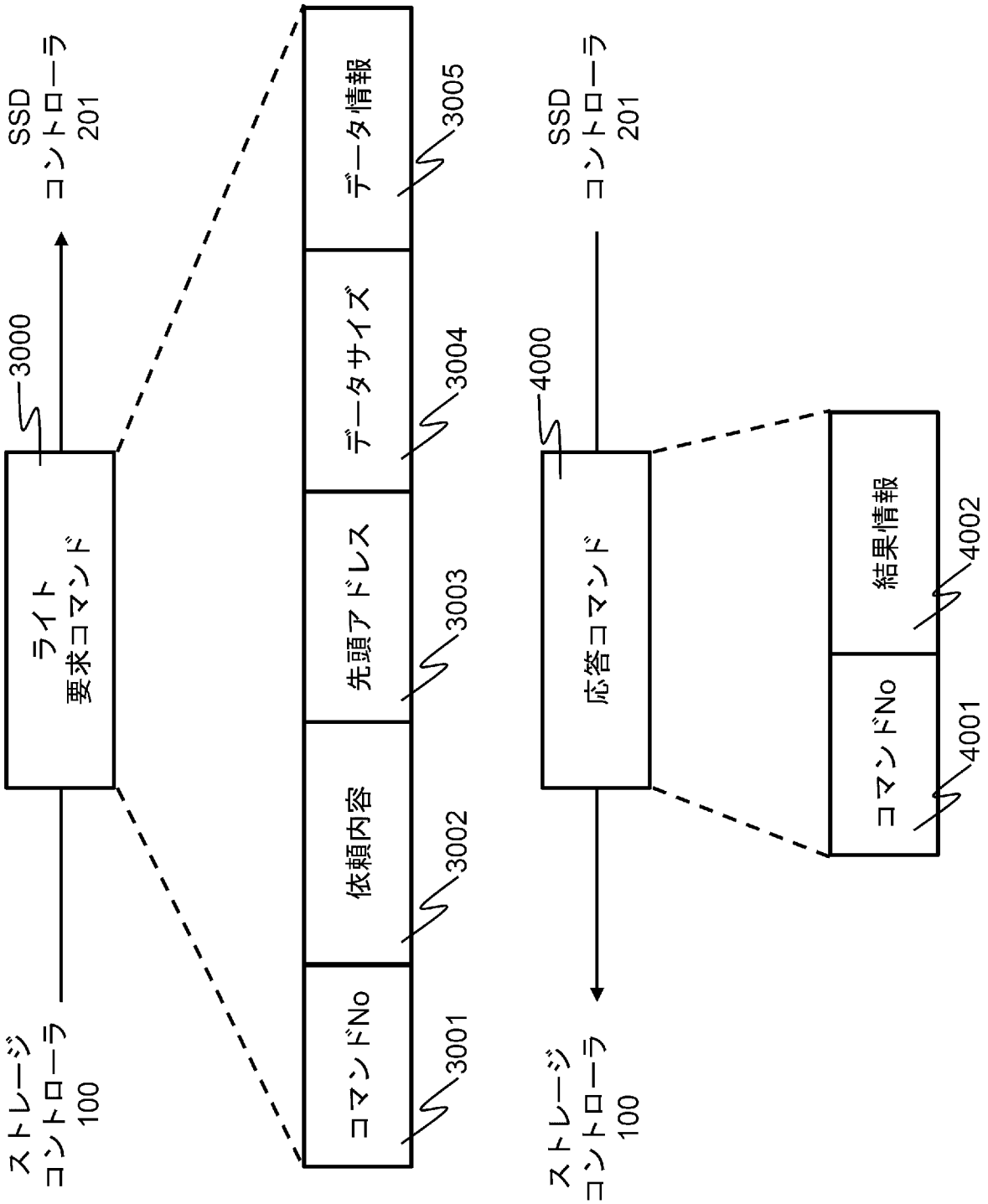

ストレージ
コントローラ
100
SSD
コントローラ
201
3000
ライト
要求コマンド
コマンドNo
3001
依頼内容
3002
先頭アドレス
3003
データサイズ
3004
データ情報
3005
ストレージ
コントローラ
100
SSD
コントローラ
201
4000
応答コマンド
コマンドNo
4001
結果情報
4002

図6

[図7]

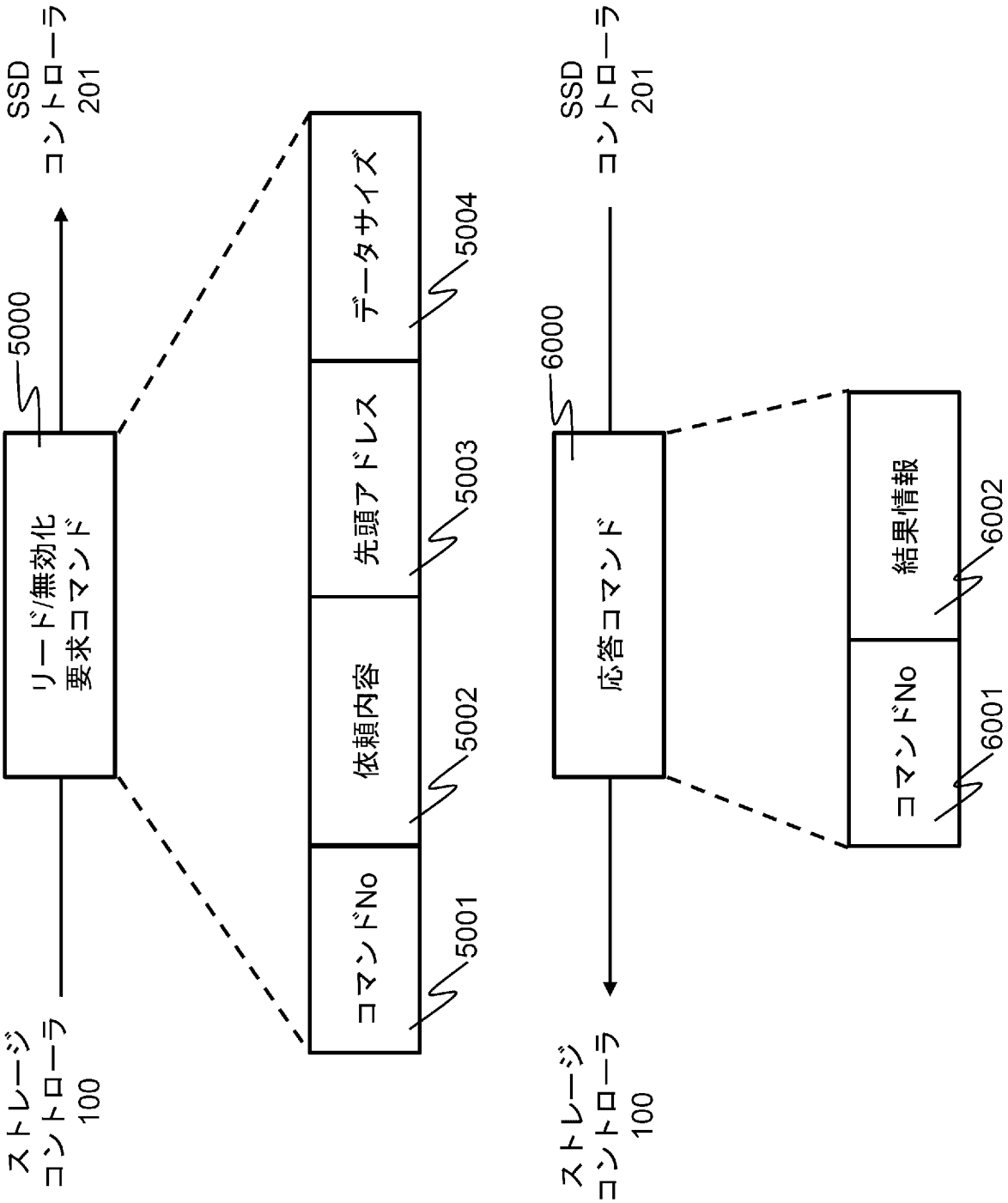

ストレージ
コントローラ
100
SSD
コントローラ
201
5000
リード/無効化
要求コマンド
コマンドNo
依頼内容
先頭アドレス
データサイズ
5001
5002
5003
5004
ストレージ
コントローラ
100
SSD
コントローラ
201
6000
応答コマンド
コマンドNo
結果情報
6001
6002

図7

[図8]

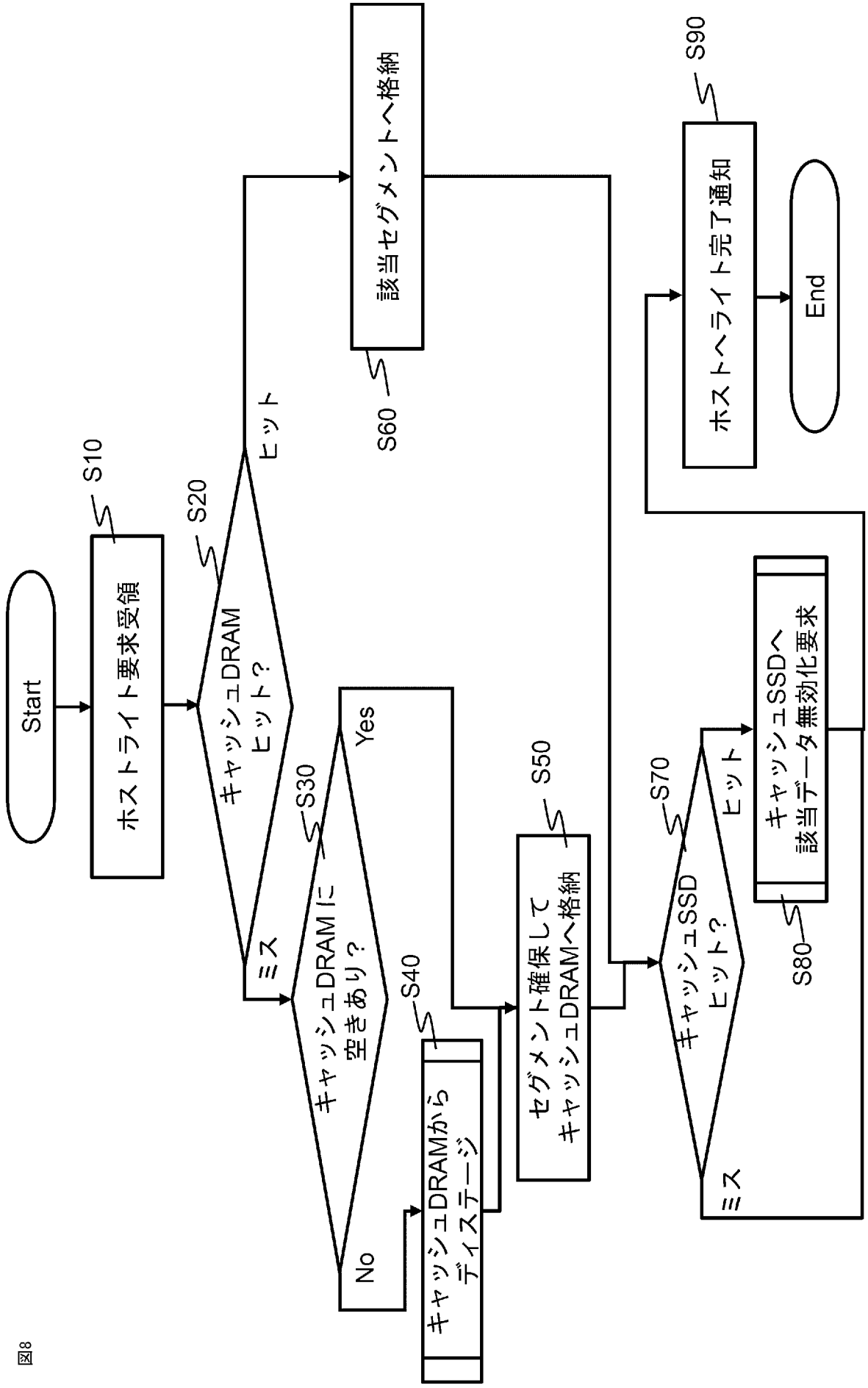

Start
ホストライト要求受領
S10
キャッシュDRAM ヒット？
S20
ヒット
ミス
該当セグメントへ格納
S60
キャッシュDRAM に 空きあり？
S30
Yes
No
キャッシュDRAMから ディステージ
S40
セグメント確保して キャッシュDRAMへ格納
S50
キャッシュSSD ヒット？
S70
ヒット
ミス
キャッシュSSDへ 該当データ無効化要求
S80
ホストへライト完了通知
S90
End

図8

[図9]

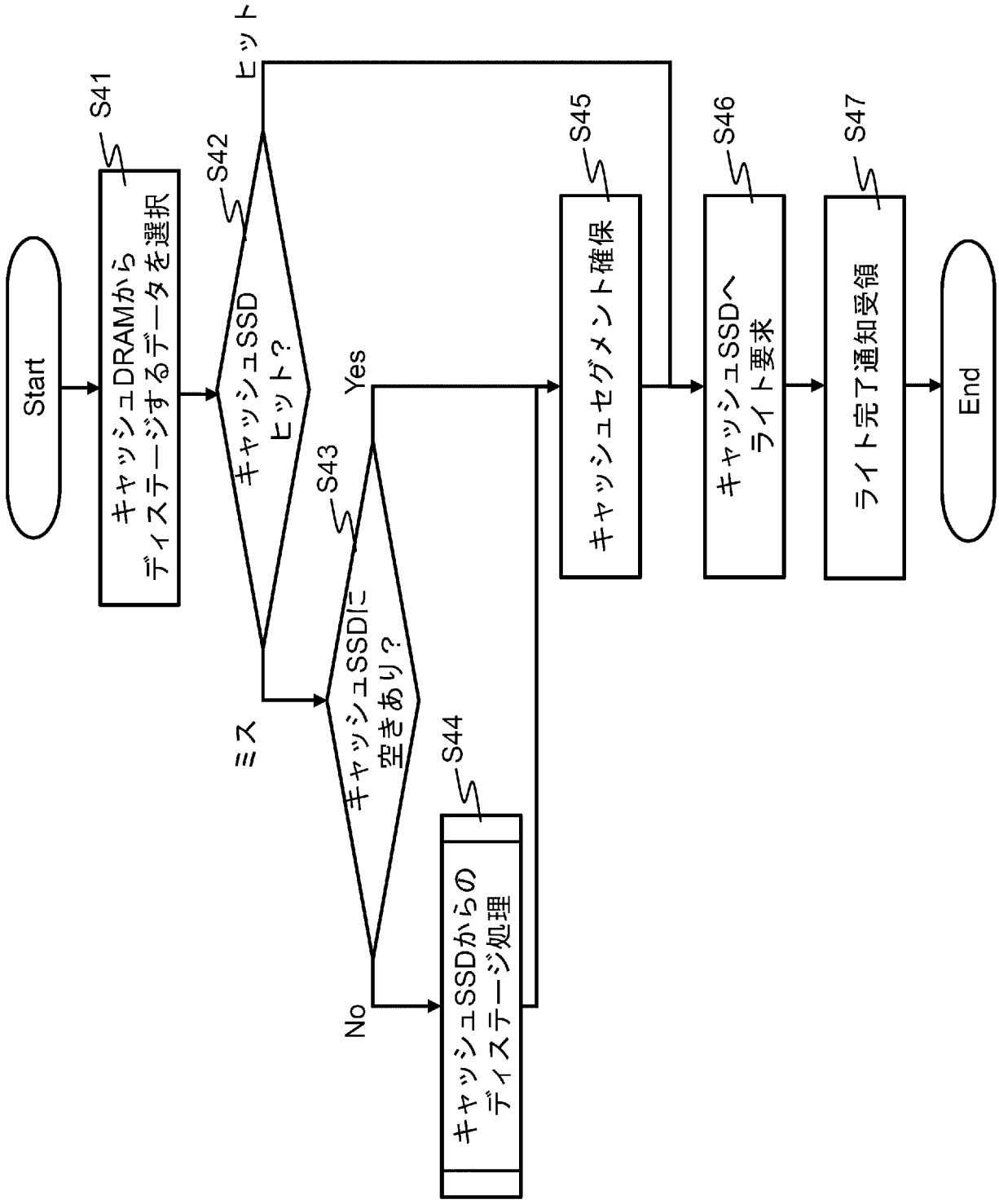

Start
S41
キャッシュDRAMから
ディステージするデータを選択
S42
キャッシュSSD
ヒット?
ヒット
ミス
Yes
No
S43
キャッシュSSDに
空きあり?
S44
キャッシュSSDからの
ディステージ処理
S45
キャッシュセグメント確保
S46
キャッシュSSDへ
ライト要求
S47
ライト完了通知受領
End

図9

[図10]

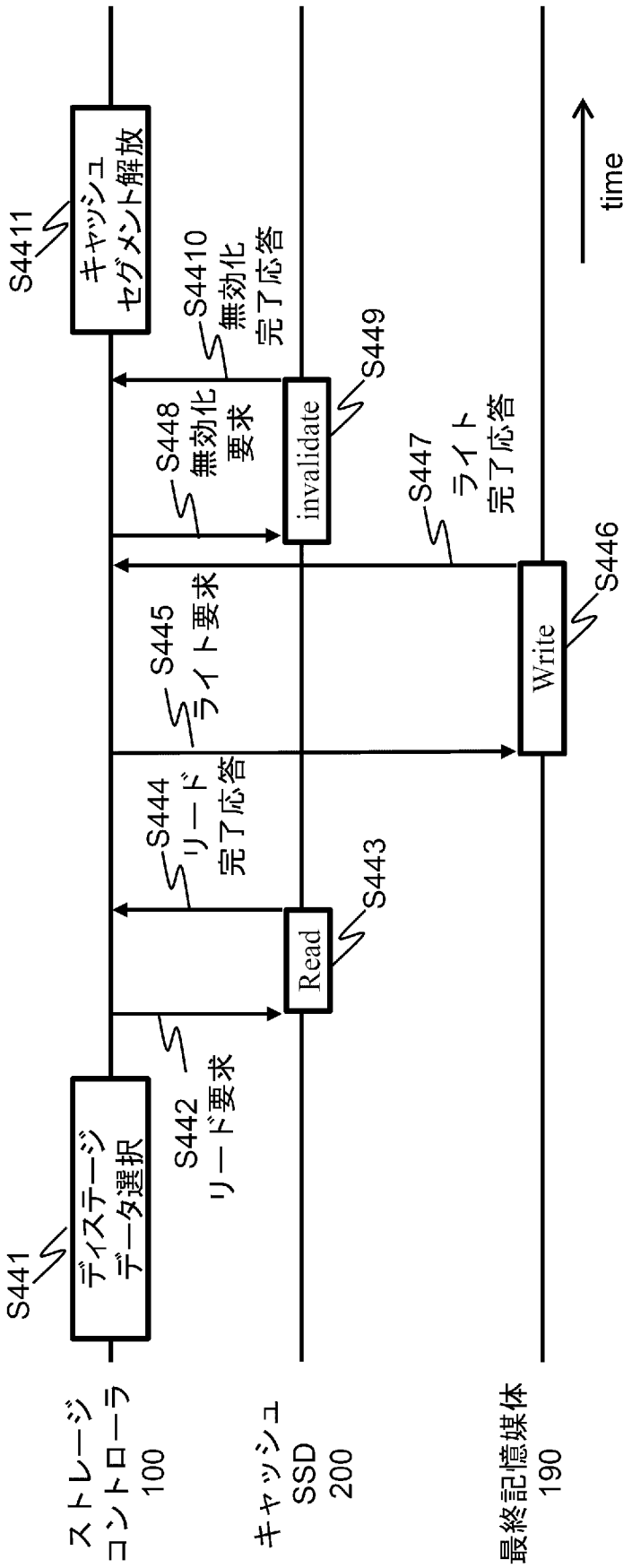

ストレージ
コントローラ
100
キャッシュ
SSD
200
最終記憶媒体
190
time
S441
ディステージ
データ選択
S442
リード要求
S443
Read
S444
リード
完了応答
S445
ライト要求
S446
Write
S447
ライト
完了応答
S448
無効化
要求
S449
invalidate
S4410
無効化
完了応答
S4411
キャッシュ
セグメント解放

図10

[図11]

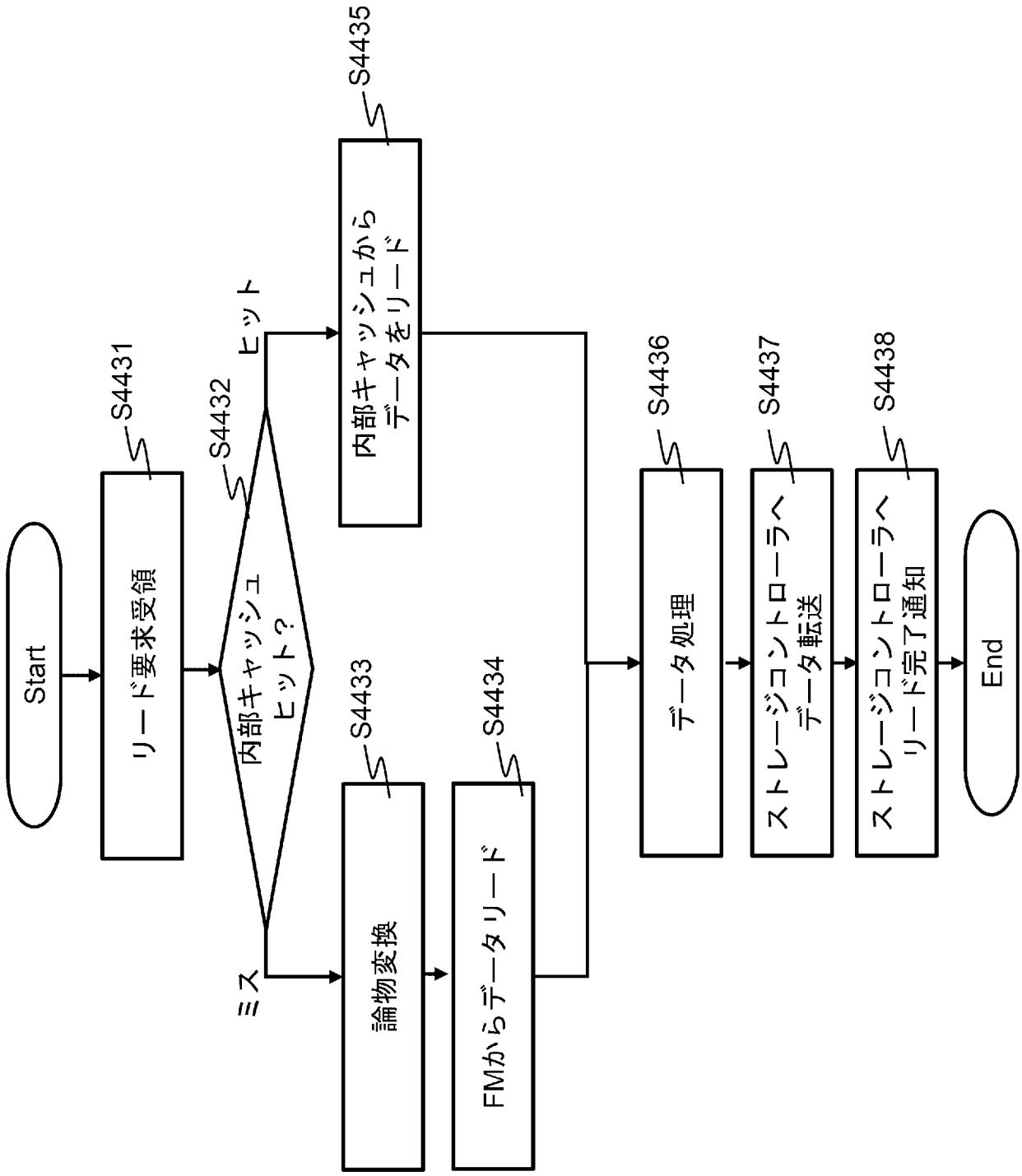

Start
リード要求受領
S4431
内部キャッシュ ヒット?
S4432
ヒット
ミス
内部キャッシュから データをリード
S4435
論物変換
S4433
FMからデータリード
S4434
データ処理
S4436
ストレージコントローラへ データ転送
S4437
ストレージコントローラへ リード完了通知
S4438
End

図11

[図12]

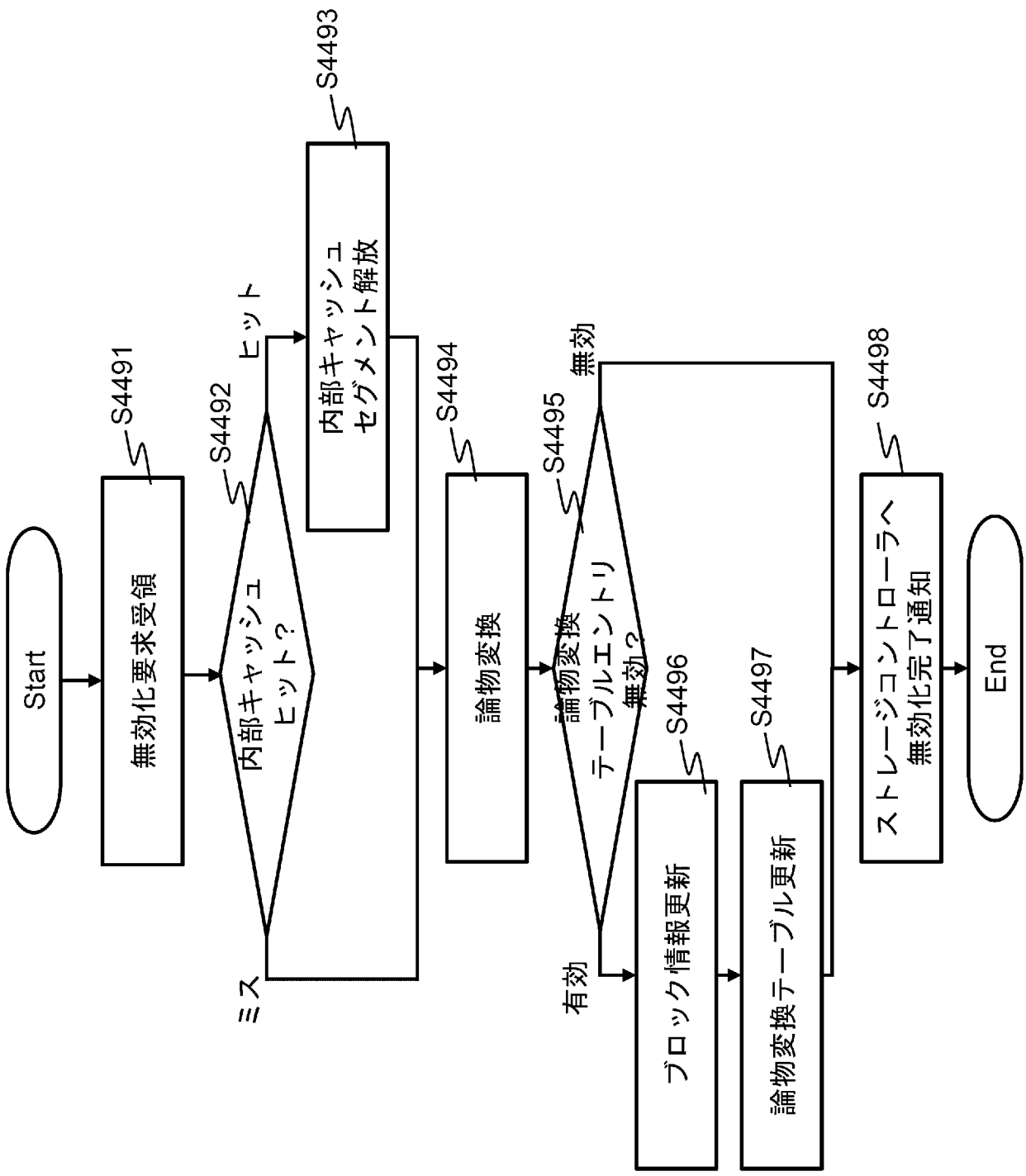

Start
無効化要求受領
S4491
内部キャッシュ
ヒット?
S4492
ヒット
ミス
内部キャッシュ
セグメント解放
S4493
論物変換
S4494
論物変換
テーブルエントリ
無効?
S4495
無効
有効
ブロック情報更新
S4496
論物変換テーブル更新
S4497
ストレージコントローラへ
無効化完了通知
S4498
End

図12

[図13]

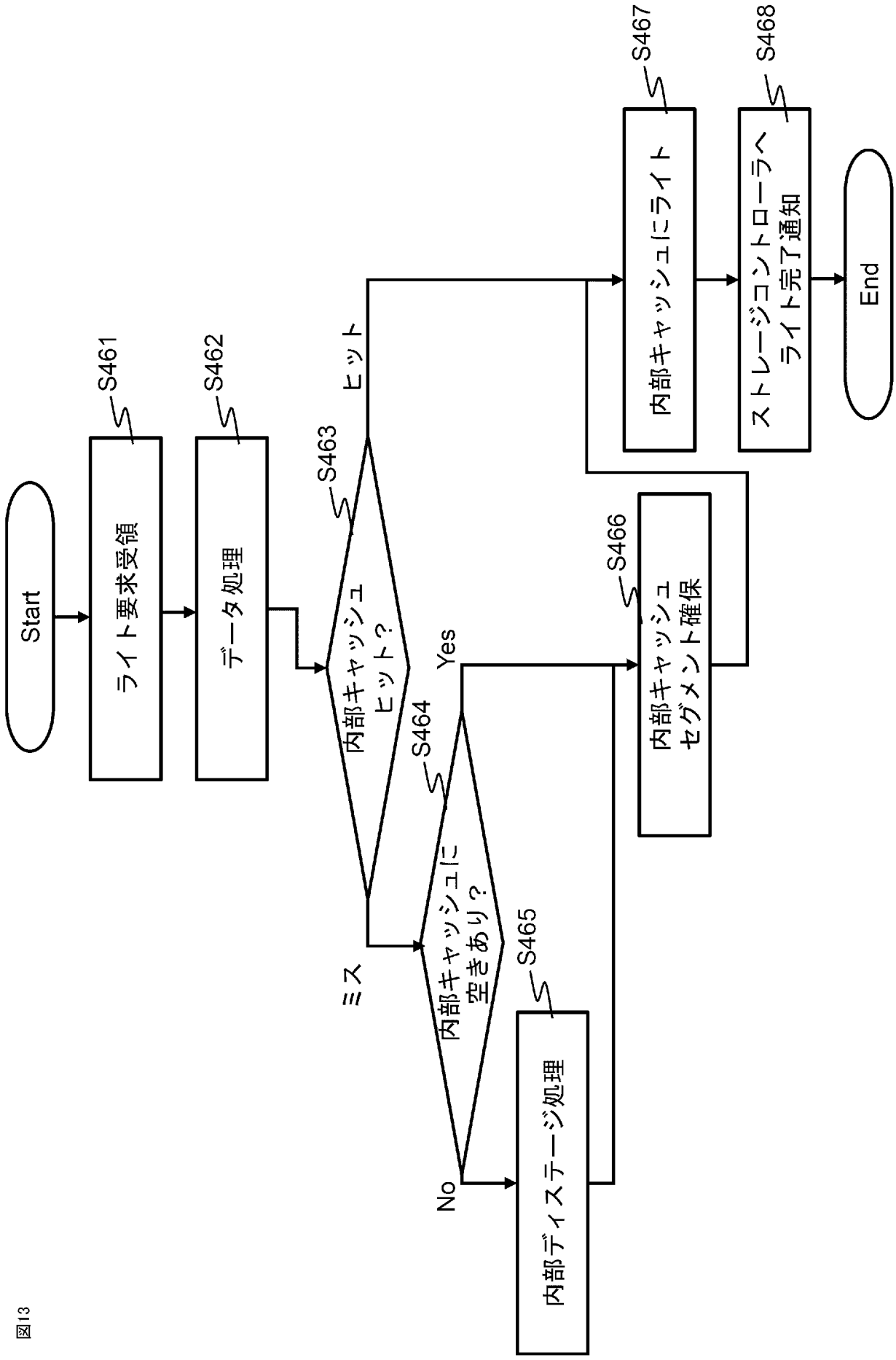

Start
S461
ライト要求受領
S462
データ処理
S463
内部キャッシュ
ヒット?
ヒット
ミス
S464
内部キャッシュに
空きあり?
Yes
No
S465
内部ディステージ処理
S466
内部キャッシュ
セグメント確保
S467
内部キャッシュにライト
S468
ストレージコントローラへ
ライト完了通知
End

図13

[図14]

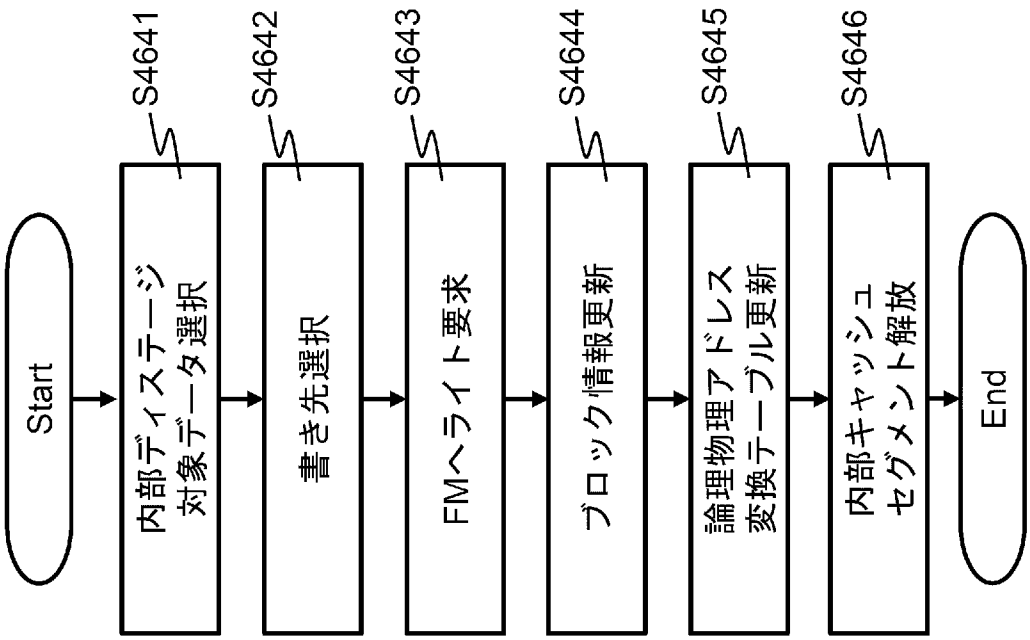

Start
S4641
内部ディスステージ
対象データ選択
S4642
書き先選択
S4643
FMへライト要求
S4644
ブロック情報更新
S4645
論理物理アドレス
変換テーブル更新
S4646
内部キャッシュ
セグメント解放
End

図14

[図15]

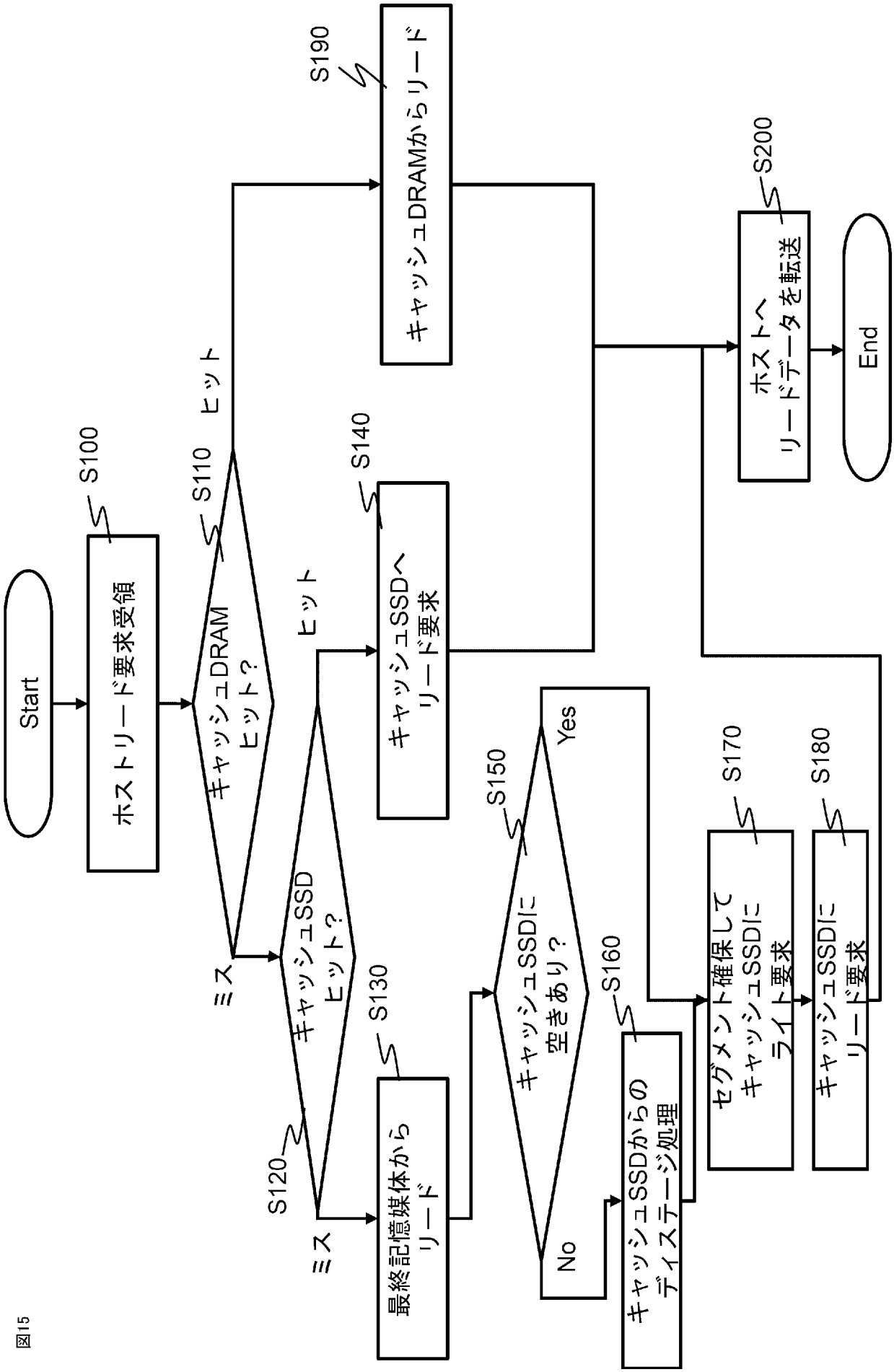

Start
ホストリード要求受領
S100
キャッシュDRAM ヒット?
S110
ヒット
ミス
キャッシュSSD ヒット?
S120
ヒット
ミス
キャッシュSSDへ リード要求
S140
最終記憶媒体から リード
S130
キャッシュSSDに 空きあり?
S150
Yes
No
キャッシュSSDからの ディステージ処理
S160
セグメント確保して キャッシュSSDに ライト要求
S170
キャッシュSSDに リード要求
S180
キャッシュDRAMからリード
S190
ホストへ リードデータを転送
S200
End

図15

[図16]

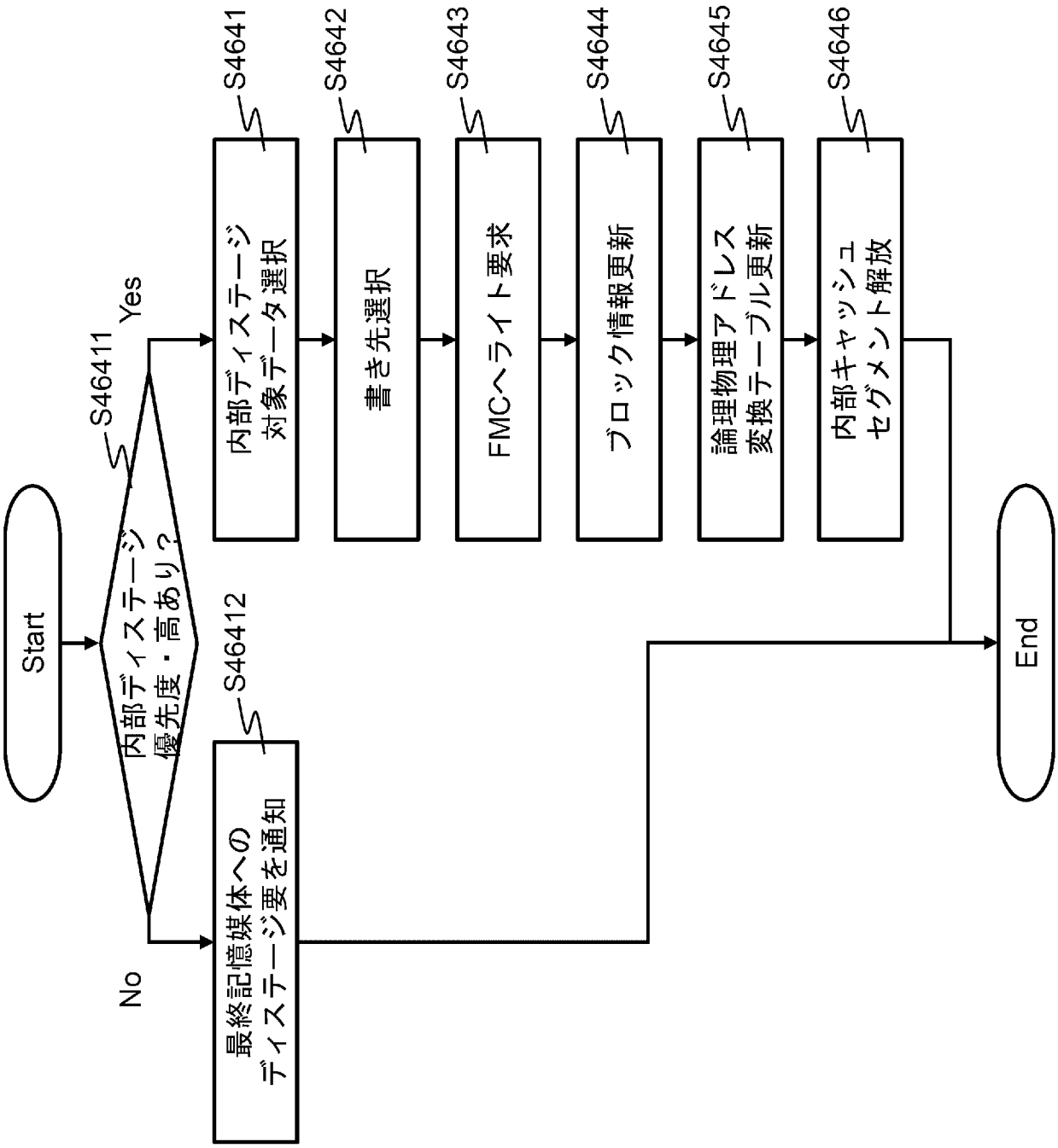

Start
S46411
内部ディスステージ
優先度・高あり？
Yes
No
S46412
最終記憶媒体への
ディスステージ要を通知
S4641
内部ディスステージ
対象データ選択
S4642
書き先選択
S4643
FMCへライト要求
S4644
ブロック情報更新
S4645
論理物理アドレス
変換テーブル更新
S4646
内部キャッシュ
セグメント解放
End

図16

[図17]

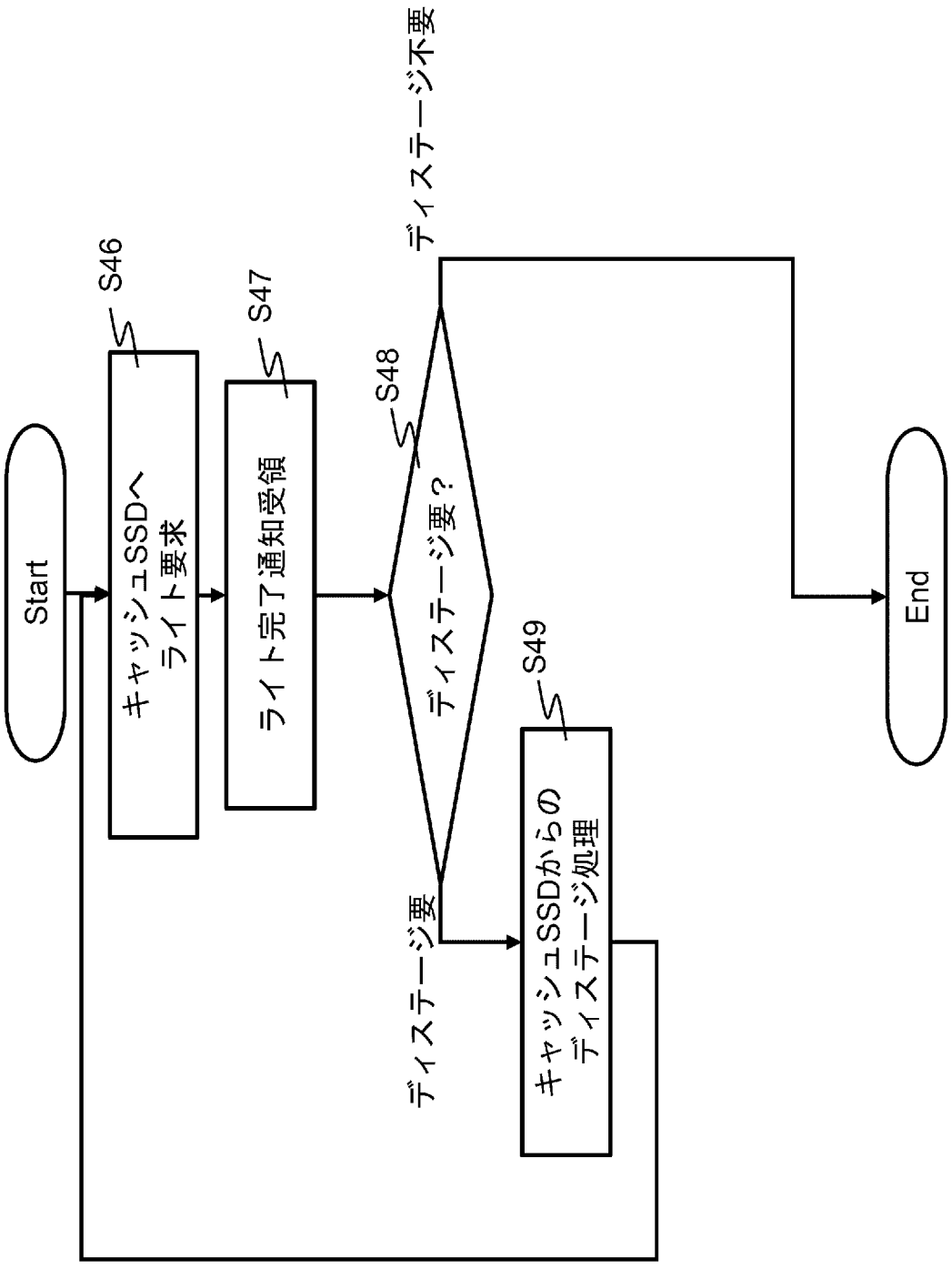

Start
S46
キャッシュSSDへ
ライト要求
S47
ライト完了通知受領
S48
ディステージ要？
ディステージ不要
ディステージ要
S49
キャッシュSSDからの
ディステージ処理
End

図17

[図18]

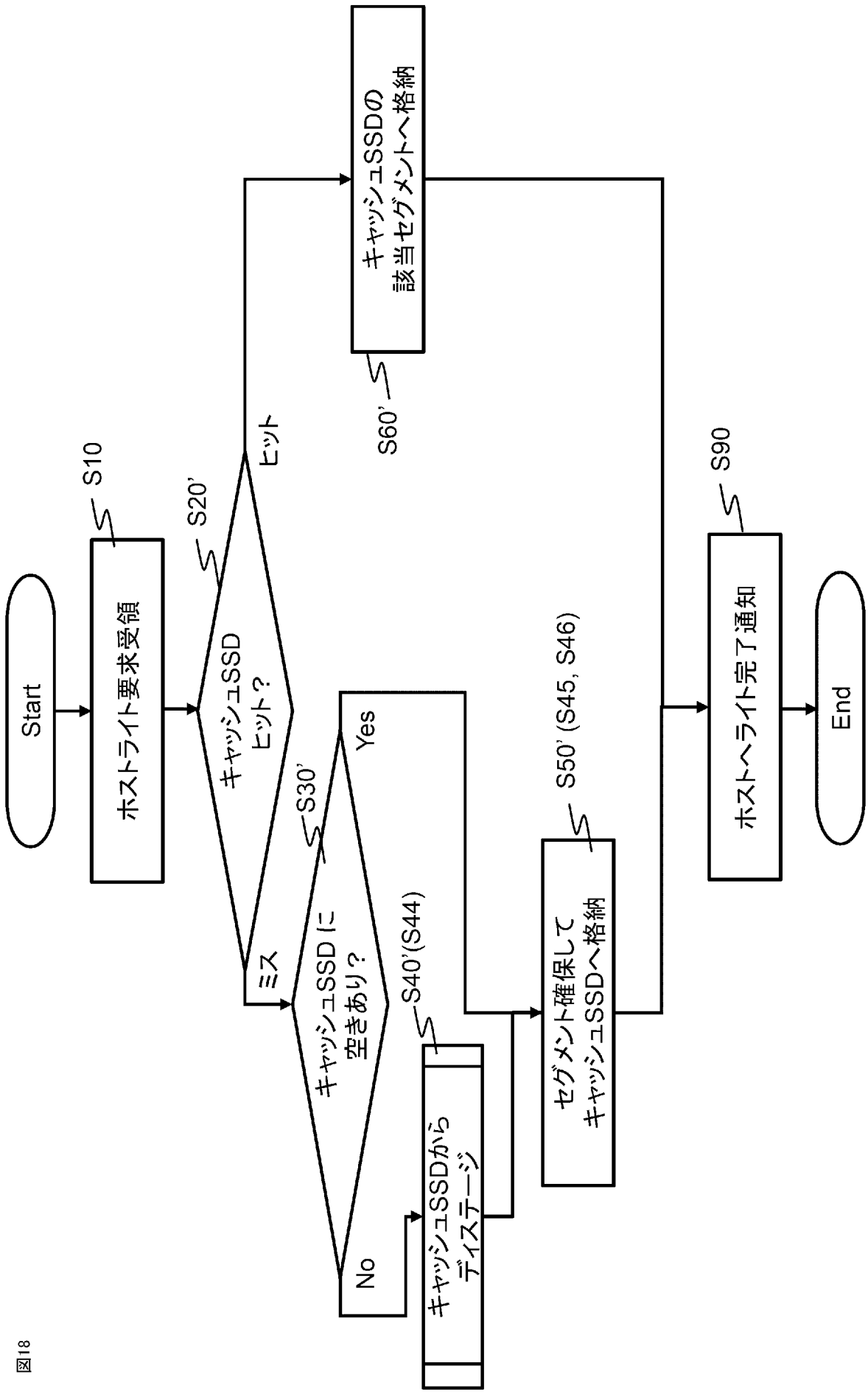

Start
ホストライト要求受領
S10
キャッシュSSD ヒット？
S20'
ヒット
ミス
キャッシュSSDの該当セグメントへ格納
S60'
キャッシュSSD に空きあり？
S30'
Yes
No
キャッシュSSDからディステージ
S40'(S44)
セグメント確保してキャッシュSSDへ格納
S50' (S45, S46)
ホストへライト完了通知
S90
End

図18

[図19]

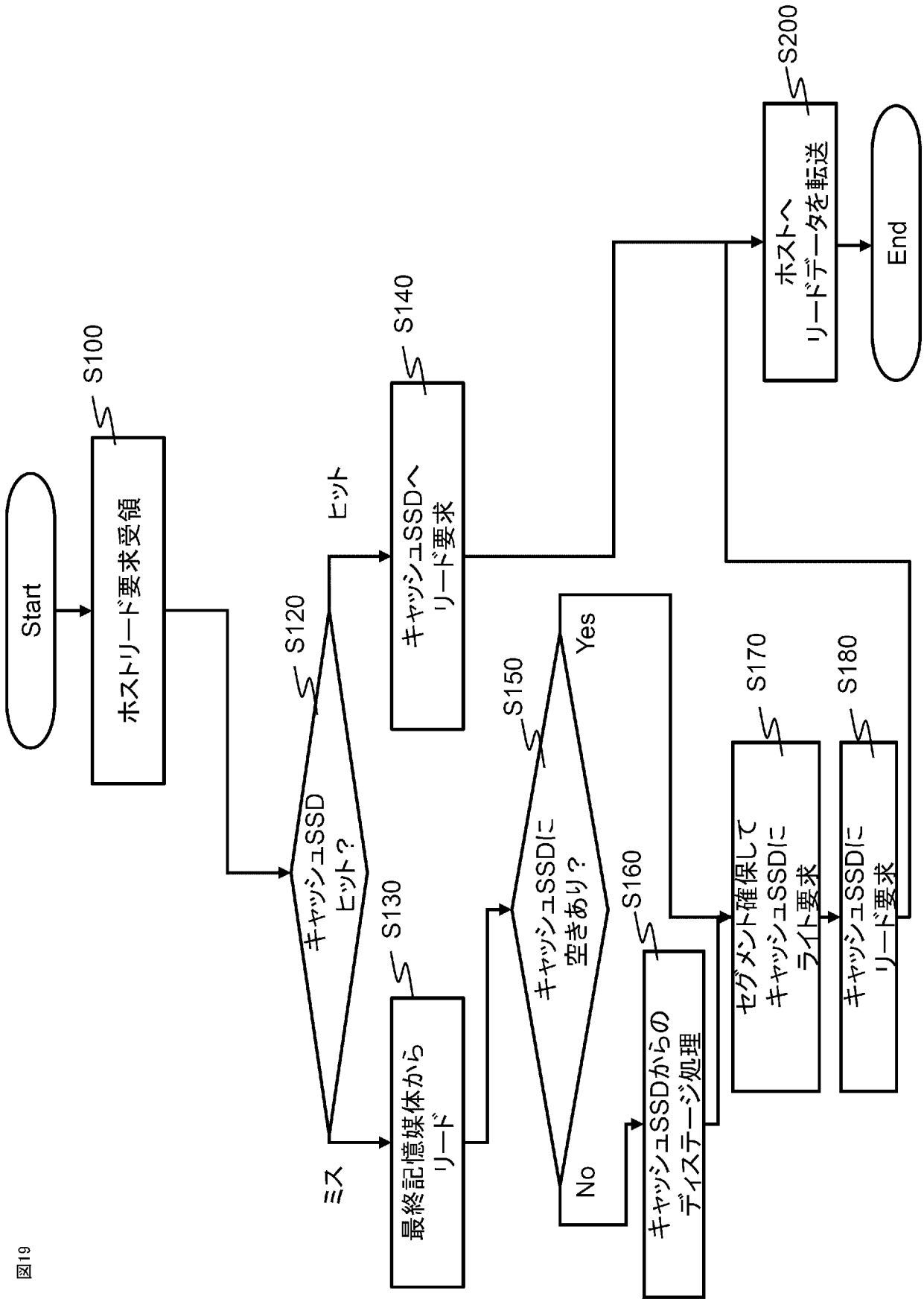

Start
S100
ホストリード要求受領
S120
キャッシュSSD
ヒット？
ヒット
ミス
S130
最終記憶媒体から
リード
S140
キャッシュSSDへ
リード要求
S150
キャッシュSSDに
空きあり？
Yes
No
S160
キャッシュSSDからの
デイステージ処理
S170
セグメント確保して
キャッシュSSDに
ライト要求
S180
キャッシュSSDに
リード要求
S200
ホストへ
リードデータを転送
End

図19

[図20]

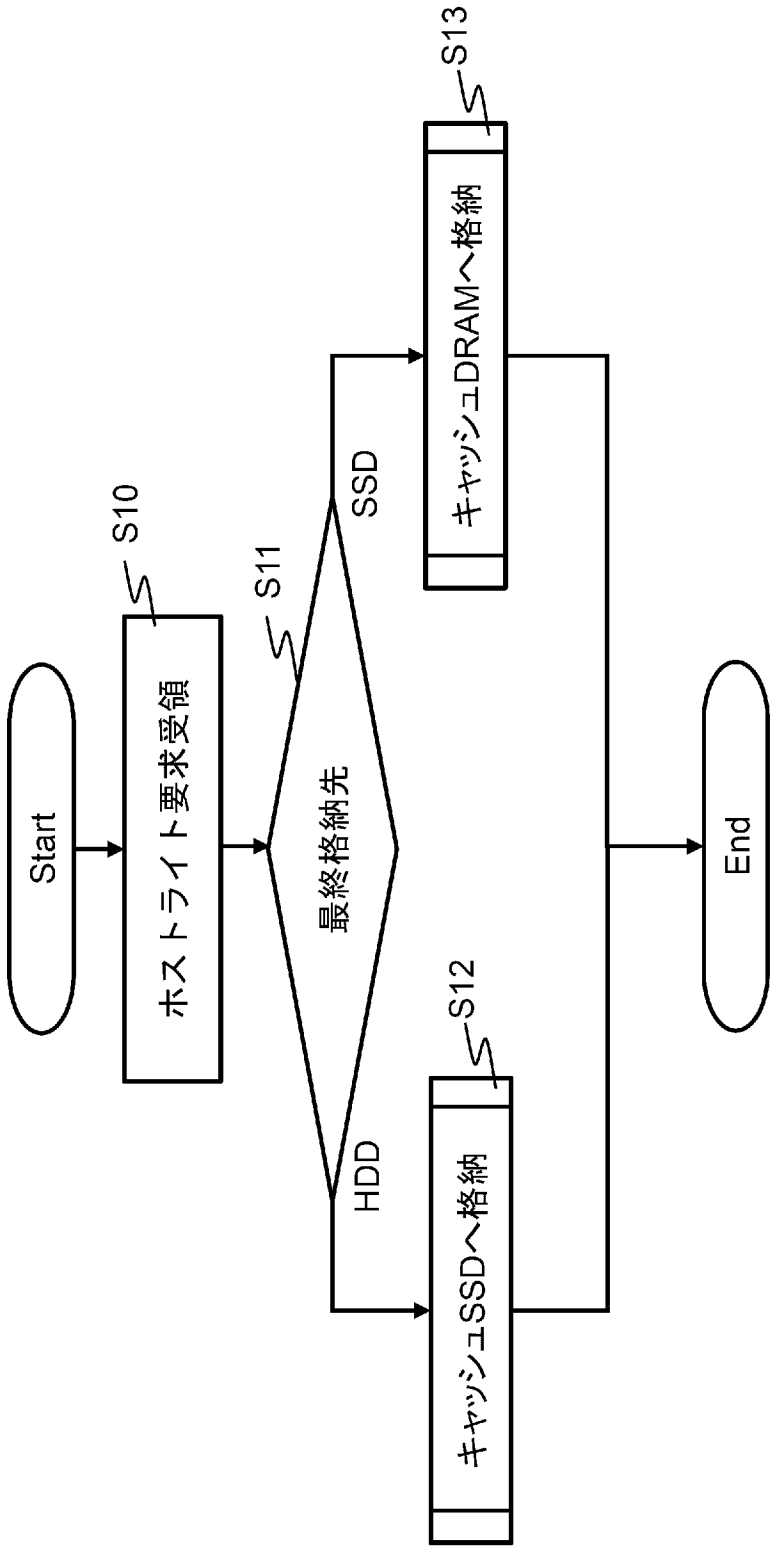

Start
S10
ホストライト要求受領
S11
最終格納先
SSD
HDD
S13
キャッシュDRAMへ格納
S12
キャッシュSSDへ格納
End

図20

[図21]

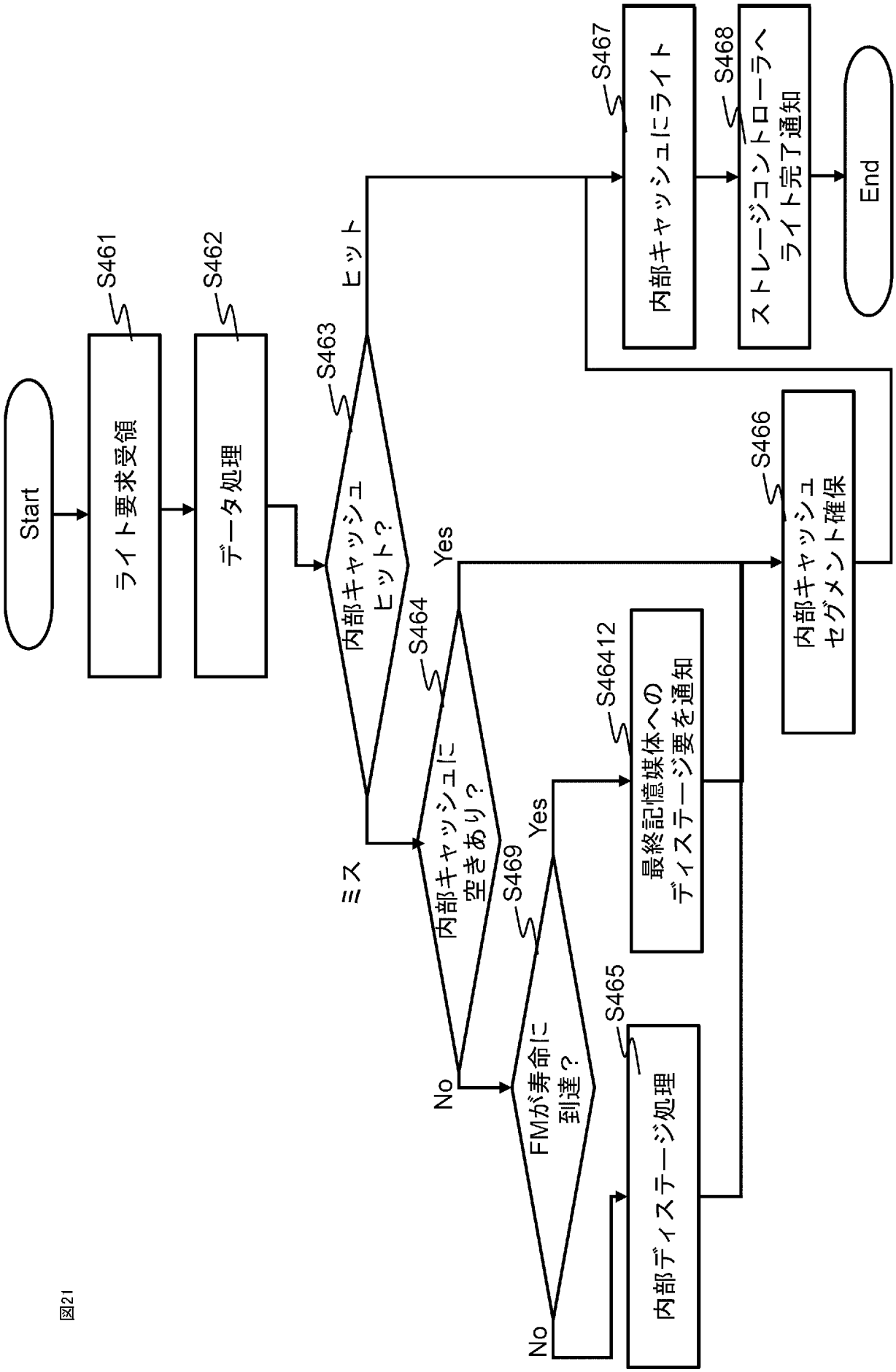

Start
S461
ライト要求受領
S462
データ処理
S463
内部キャッシュ
ヒット？
ヒット
ミス
S464
内部キャッシュに
空きあり？
Yes
No
S469
FMが寿命に
到達？
Yes
No
S46412
最終記憶媒体への
ディステージを通知
S465
内部ディステージ処理
S466
内部キャッシュ
セグメント確保
S467
内部キャッシュにライト
S468
ストレージコントローラへ
ライト完了通知
End

図21

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2014/062714

A. CLASSIFICATION OF SUBJECT MATTER
G06F12/08(2006.01)i, *G06F12/12*(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G06F12/08, G06F12/12

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho	1922-1996	Jitsuyo Shinan Toroku Koho	1996-2014
Kokai Jitsuyo Shinan Koho	1971-2014	Toroku Jitsuyo Shinan Koho	1994-2014

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X Y A	JP 2008-217527 A (Hitachi, Ltd.), 18 September 2008 (18.09.2008), paragraphs [0010] to [0140]; fig. 1 to 2, 7 to 9, 12, 17 to 19 & US 2008/0222359 A1 & US 2011/0271066 A1 & US 2012/0233398 A1 & EP 1967953 A2	1-2,4,9, 11-12,14 5-8 3,10,13
Y A	JP 2008-15918 A (Toshiba Corp.), 24 January 2008 (24.01.2008), paragraphs [0101] to [0105]; fig. 18 (Family: none)	5-8 1-4,9-14
A	JP 2011-204060 A (NEC Corp.), 13 October 2011 (13.10.2011), entire text; all drawings & US 2011/0238908 A1	1-14

☒ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
30 June, 2014 (30.06.14)	08 July, 2014 (08.07.14)
Name and mailing address of the ISA/ Japanese Patent Office Facsimile No.	Authorized officer Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2014/062714

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2009-205335 A (Hitachi, Ltd.), 10 September 2009 (10.09.2009), entire text; all drawings & US 2009/0216945 A1 & US 2011/0296091 A1 & US 2012/0254523 A1	1-14

国際調査報告

国際出願番号　ＰＣＴ／ＪＰ２０１４／０６２７１４

Ａ．　発明の属する分野の分類（国際特許分類（ＩＰＣ））

Int.Cl.　G06F12/08(2006.01)i, G06F12/12(2006.01)i

Ｂ．　調査を行った分野

調査を行った最小限資料（国際特許分類（ＩＰＣ））

Int.Cl.　G06F12/08, G06F12/12

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	１９２２－１９９６年
日本国公開実用新案公報	１９７１－２０１４年
日本国実用新案登録公報	１９９６－２０１４年
日本国登録実用新案公報	１９９４－２０１４年

国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）

Ｃ．　関連すると認められる文献

引用文献のカテゴリー＊	引用文献名　及び一部の箇所が関連するときは、その関連する箇所の表示	関連する請求項の番号
X	JP 2008-217527 A (株式会社日立製作所) 2008.09.18, 段落【0010】-【0140】, 第1-2,7-9,12,17-19図 & US 2008/0222359 A1 & US 2011/0271066 A1 & US 2012/0233398 A1 & EP 1967953 A2	1-2,4,9,11-12,14
Y		5-8
A		3,10,13
Y	JP 2008-15918 A (株式会社東芝) 2008.01.24, 段落【0101】-【0105】, 第18図（ファミリーなし）	5-8
A		1-4,9-14

☑　Ｃ欄の続きにも文献が列挙されている。

☐　パテントファミリーに関する別紙を参照。

＊　引用文献のカテゴリー
「Ａ」特に関連のある文献ではなく、一般的技術水準を示すもの
「Ｅ」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
「Ｌ」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）
「Ｏ」口頭による開示、使用、展示等に言及する文献
「Ｐ」国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献
「Ｔ」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの
「Ｘ」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
「Ｙ」特に関連のある文献であって、当該文献と他の１以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
「＆」同一パテントファミリー文献

国際調査を完了した日	国際調査報告の発送日
３０．０６．２０１４	０８．０７．２０１４

国際調査機関の名称及びあて先
日本国特許庁（ＩＳＡ／ＪＰ）
郵便番号１００－８９１５
東京都千代田区霞が関三丁目４番３号

特許庁審査官（権限のある職員）　５Ｕ　３４５０
野田　佳邦

電話番号　０３－３５８１－１１０１　内線　３５６５

Ｃ（続き）．　関連すると認められる文献

引用文献のカテゴリー＊	引用文献名　及び一部の箇所が関連するときは、その関連する箇所の表示	関連する請求項の番号
A	JP 2011-204060 A (日本電気株式会社) 2011.10.13, 全文, 全図 & US 2011/0238908 A1	1-14
A	JP 2009-205335 A (株式会社日立製作所) 2009.09.10, 全文, 全図 & US 2009/0216945 A1 & US 2011/0296091 A1 & US 2012/0254523 A1	1-14