

(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8月6日(06.08.2015)



(10) 国際公開番号
WO 2015/114754 A1

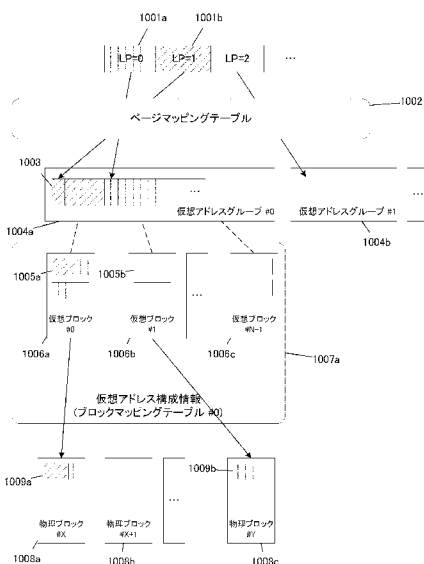
- (51) 国際特許分類:
G06F 12/02 (2006.01) *G06F 3/08* (2006.01)
G06F 3/06 (2006.01) *G06F 12/00* (2006.01)
- (21) 国際出願番号: PCT/JP2014/051949
- (22) 国際出願日: 2014年1月29日(29.01.2014)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (71) 出願人: 株式会社日立製作所 (HITACHI, LTD.)
[JP/JP]; 〒1008280 東京都千代田区丸の内一丁目
6番6号 Tokyo (JP).
- (72) 発明者: 河村 篤志 (KAWAMURA, Atsushi); 〒
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). 小川 純司
(OGAWA, Junji); 〒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 (PAT-
ENT 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,

[続葉有]

(54) Title: STORAGE DEVICE

(54) 発明の名称: ストレージ装置

図10



1002 Page mapping table
1004a, 1004b Virtual address group
1006a, 1006b, 1006c Virtual block
1007a Virtual address configuration information (block
mapping table #0)
1008a, 1008b, 1008c Physical block

(57) Abstract: The present invention provides a storage device employing, as a storage medium, a semiconductor device that has nonvolatile storage properties and that requires an erase operation thereon prior to a write operation thereon, wherein said storage device manages a logical storage space to be provided for a host device, by dividing the logical storage space into individual logical pages, and also manages a virtual address space, which is a linear address space having mapped thereto a plurality of physical blocks of the semiconductor device. The storage device manages the mapping between the individual logical pages and the storage regions of the semiconductor device by use of both a page mapping table, which manages the mapping between the logical pages and addresses in the virtual address space, and virtual address configuration information, which manages the mapping between regions in the virtual address space and the physical blocks.

(57) 要約: 本発明は、不揮発性を持ち、書き込みのために消去を必要とする半導体デバイスを記憶媒体として採用したストレージ装置において、上位装置に対して提供する論理記憶空間を論理ページ単位に分割して管理するとともに、半導体デバイスが有する複数の物理ブロックが対応付けられた線形なアドレス空間である仮想アドレス空間を管理している。ストレージ装置は、論理ページと仮想アドレス空間上のアドレスとの対応関係を管理するページマッピングテーブルと、仮想アドレス空間上の領域と物理ブロックとの対応関係を管理する仮想アドレス構成情報とを用いることで、各論理ページと半導体デバイスの記憶領域の対応関係を管理する。

WO 2015/114754 A1



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))

明 細 書

発明の名称：ストレージ装置

技術分野

[0001] 本発明は、不揮発性半導体メモリを用いたストレージ装置とその制御方法に関する。

背景技術

[0002] 不揮発性半導体メモリは、磁気記憶デバイスに比べて省電力・高性能であり、高価格であった。このような不揮発性半導体メモリは、例えばNAND型のフラッシュメモリである。近年は半導体技術の進歩に伴って低価格化が進み、HDDに替わるメインストリームの記憶デバイスとして注目されてきている。

[0003] フラッシュメモリにおいて、ブロックは、データの消去単位であり、ページは、データを読み書きする単位の記憶領域である。今後特に断りがない限り、単にブロック／ページと記述した場合はフラッシュメモリのものを指す。後述のように1つのブロック内に複数のページが設けられる。またフラッシュメモリは、特性上、記憶しているデータを直接書き換えることができない。つまり、フラッシュメモリは、記憶しているデータを書き換える場合、記憶している有効なデータを他のブロックに退避させる。次に、フラッシュメモリは、記憶しているデータをブロック単位で消去する。そして、フラッシュメモリは、消去したブロックにデータを書き込む。このように、フラッシュメモリにおけるデータの書き換えは、ブロックごとのデータの消去を伴う。しかし、フラッシュメモリの1ブロック分のデータ消去にかかる時間は、1ページ分のデータを書き込むのに要する時間と比べて約1桁長い。従って、1ページ分のデータの書き換えのために毎回1ブロック分のデータ消去を行なうようにすると、フラッシュメモリのデータ書き換え性能が低下する。このような問題に対応するため、フラッシュメモリからデータを消去する時間を隠蔽するアルゴリズムを用いてデータをフラッシュメモリへ書き込むこと

が知られている。

[0004] 通常、フラッシュメモリに対するデータの書換え動作は、未使用領域へデータを追記する方式で行う。しかし、データの書き換え量が増えると、フラッシュメモリ内の未使用領域が少なくなるため、フラッシュメモリに書き込まれている不要なデータを消去して記憶領域を再利用可能な状態にする必要が生ずる。そこで、旧データを含むブロック内の、有効なデータのみを未使用領域にコピーし、コピー元のブロックを消去して再利用可能な状態にするブロック再生処理が知られている。以下、これをリクラメーションと呼ぶ。場合によってはR Cと省略して表記する。このリクラメーションは、無効データが多くなったブロックを対象に実行することで効率よく空き領域を作成可能である。

[0005] フラッシュメモリは、使用とともに特性が劣化する。劣化は、セルの消去回数に相関があるとされ、一般的にSSDベンダはデータ保持保証可能な、最大の消去回数を提示している。従って、特定のブロックにデータの書換えが集中し、消去回数が増大して使用不可能となった場合、最悪、他の領域が健全な状態にも関わらず、一部のブロックが書き換え限界に達したために、装置容量を満足できなくなる問題が発生する。そのため、劣化が特定ブロックに集中しないように、各ブロックの消去回数を平準化する技術が知られている。

[0006] またフラッシュメモリは、一度書き込んだページでも、時間とともに読み出しエラー率が増加する。このように保持しておくだけでも発生するエラーは、リテンションエラーと呼ばれる。これを回避するために、書き込んでから一定時間経過したページを、別ページにコピーする処理が知られている。この処理は、リフレッシュと呼ばれる。場合によってはR Fと省略して表記する。リフレッシュにおいても、既に述べた平準化の問題や性能影響を考慮する必要がある。

[0007] 以上で述べたデータ消去時間の隠蔽とデータ消去回数の平準化のため、フラッシュストレージは、データ書込みの際に、論理アドレスを物理アドレス

へ変換する論理物理アドレス変換処理を行う。フラッシュメモリモジュールは、1つ以上のフラッシュメモリチップと、フラッシュメモリチップに対するデータの読み書きを制御するフラッシュメモリコントローラとを有する。このフラッシュメモリコントローラが、論理物理アドレス変換処理を行なう。さらにフラッシュメモリコントローラは、論理物理アドレス変換処理のための情報を、論理物理アドレス変換テーブルとして保存する。以下、論理物理アドレス変換処理を論物変換、論理物理アドレス変換テーブルを論物変換テーブルと呼ぶことがある。

[0008] 論物変換は、フラッシュメモリを効率的に使用するために重要な役割を有している。低い自由度の論物変換を用いる場合、論物変換テーブルのサイズを抑えられる代わりに、リクラメーションが頻発するなどにより性能が低下する。一方、高い自由度の論物変換を用いる場合、論物変換テーブルのサイズが膨大になるとともに、制御コストが大幅に増大する。

[0009] 上記論物変換方式は、目的とする動作に応じて様々な方式が提案されているが、その中でも広く知られているものとして、上位装置に提供するアドレス空間をフラッシュメモリのページサイズで区切って形成される各論理単位を、フラッシュメモリの各ページに対応付ける方式がある。これは、ページサイズが2KB（キロバイト）、4KB、8KBといった、ホストOSやファイルシステム又はアプリケーションプログラムからのI/Oパターンによくマッチした格納単位であり最適化がしやすいこと、論理単位とページを1対1対応させることで管理を簡易化できること、また、論理単位のページ跨がりを抑止することで、複数ページリードの頻度を減らすなどの利点がある。この論理単位を論理ページ、フラッシュメモリのページを、これと区別するために物理ページと呼ぶことがある。

[0010] 一方で、論理ページサイズと物理ページサイズが一致しないケースが考えられる。まずは、圧縮などのデータ変換処理により、最終的なサイズが変動する場合である。フラッシュメモリのビットコスト低減は、HDDの代替を目指す上で大きな目標であり、データ圧縮のフラッシュストレージへの適用

は広く行われている。

[0011] 続いて考えられるのは、物理ページ内のECC等のデータ保護情報の割合が増加することで、格納できるデータ量が減少する場合である。従来は、物理ページの余剰領域にECCを格納することで、データ格納領域については2の冪乗サイズを維持していたが、微細化に伴い高いエラー訂正能力が必要となっており、要求されるECCサイズも拡大している。

[0012] また、フラッシュメモリの特性は、全ての物理ページについて均一ではなく、先天的な差もあれば、物理的なロケーションによる違いも存在する。従って、全ての物理ページに対して同じECCを使用するのではなく、状況に合わせて最適な格納方式を採ることも考えられる。その場合、論理ページと物理ページのサイズ不一致はより顕著となる。

[0013] 論理・物理のページサイズが異なる場合でも、論理ページと物理ページの1対1対応付けを行おうとすると、まず論理ページサイズは物理ページサイズより小さい必要がある。また、複数の論理ページを物理ページに格納できないため、場合によっては無効なデータで物理ページをパディングする必要が出てくるため、格納効率が低下する。

[0014] 一方で、論理ページが物理ページを跨がることを許容する場合、パディングを最小化し、格納効率を上げることができるが、より複雑な管理方式が必要となる。例えば特許文献1では、複数の物理ページに跨がった論理ページをサブブロックに分割し、それぞれの分割された論理ページの格納情報を管理することで、1つの論理ページを複数の物理ページに跨がって格納される態様を許容している。

先行技術文献

特許文献

[0015] 特許文献1：米国特許第7529905号明細書

発明の概要

発明が解決しようとする課題

[0016] フラッシュメモリにおける物理ページの様に、書き込み単位が存在するデバイスにおいて、論理ページと物理ページのサイズを一致させることは利点が多い。しかし、圧縮などのデータ変換、デバイスの物理特性に由来するECCサイズの拡大、などの理由により、論理ページサイズと物理ページサイズが不一致となることを考慮する必要がある。論理ページの物理ページ跨がりを許容しなかった場合、物理ページ中に使用されない領域が増えて格納効率が低下する。一方で跨がりを許容する場合、何らかの形で跨がり情報を保持する必要がある。特許文献1に開示されているように、論理ページ毎に跨がり情報を保持する方法が考えられるが、その場合、それぞれの論理ページ毎に2以上のアドレス情報が必要となり、管理情報サイズの増大を招く。

課題を解決するための手段

[0017] 本発明の一実施形態に係るフラッシュメモリストレージ装置は、コントローラと複数のフラッシュメモリチップを備える。コントローラは、上位装置に対して提供する論理記憶空間を、所定サイズの領域である論理ページ単位に分割して管理するとともに、複数のフラッシュメモリチップの有する複数の物理ブロックが対応付けられた線形なアドレス空間である仮想アドレス空間を有している。上位装置から論理ページに対する書き込みがあった場合には、前記仮想アドレス空間上の未使用領域を選択し、この選択した前記仮想アドレス空間上の未使用領域のアドレスと論理ページとの対応関係をページマッピングテーブルに記録する。仮想アドレス空間から領域を選択する際、未使用領域であれば任意の領域を選択することができ、物理ページ境界を意識する必要はない。

発明の効果

[0018] 本発明によれば、書き込み単位を有する記憶デバイスにおいて、書き込み単位とは異なる論理管理単位データを、書き込み単位境界を跨いで格納することが可能になり、格納効率が向上する。また、論理管理単位データを、書き込み単位境界を跨いで格納しても、管理情報のサイズを増大させることなく、論理管理単位データの記憶デバイス上格納位置の情報を管理するこ

とができる。

図面の簡単な説明

[0019] [図1] 図1はストレージシステムの構成図を示す。

[図2] 図2はFMPKGの構成図を示す。

[図3] 図3はフラッシュメモリストレージのメインメモリの内容を示す。

[図4] 図4はフラッシュメモリのブロックとページの関係を示す。

[図5] 図5は物理ページ内の一般的なデータ格納構成を示す。

[図6] 図6は従来のフラッシュストレージで行われている、ページベースの論物変換処理の概要を示す。

[図7] 図7はページベースの論物変換テーブルを示す。

[図8] 図8は論理ページサイズと物理ページサイズが不一致である場合の格納態様を示す。

[図9] 図9は1論理ページが複数物理ページに跨がる場合の格納態様を示す。

[図10] 図10は本発明の実施例1に係るFMPKGで行われる、論理物理変換の概念を示した図である。

[図11] 図11はページマッピングテーブルの構成を示す。

[図12] 図12は仮想アドレス構成情報の構成を示す。

[図13] 図13は本発明の実施例1に係るFMPKGにおける、正引き処理のフローを示す。

[図14] 図14は本発明の実施例1に係るFMPKGにおけるライト処理のフローを示す。

[図15] 図15は本発明の実施例1に係るFMPKGにおける、新規グループ構築処理のフローを示す。

[図16] 図16は本発明の実施例1に係るFMPKGにおける、RC/RF処理のフローを示す。

[図17] 図17は、本発明の実施例2におけるページマッピングテーブルの構成を示す。

[図18] 図18は本発明の実施例2に係るFMPKGにおける、正引き処理の

フローを示す。

[図19]図 19 は本発明の実施例 2 に係る F M P K G におけるライト処理のフローを示す。

[図20]図 20 は本発明の実施例 2 に係る F M P K G における、R C / R F 処理のフローを示す。

[図21]図 21 は、物理ページサイズが異なるブロックの存在するデバイスで仮想アドレスグループを構成する例を示す。

[図22]図 22 は、本発明の実施例 3 に係る F M P K G における仮想アドレス構成情報の構成を示す。

[図23]図 23 は本発明の実施例 3 に係る F M P K G における、正引き処理のフローを示す。

[図24]図 24 は本発明の実施例 3 に係る F M P K G における、新規グループ構築処理のフローを示す。

発明を実施するための形態

[0020] 以下、本発明のいくつかの実施例を、図面を用いて説明する。なお、これらの実施例は本発明を実現するための一例に過ぎず、本発明の技術的範囲を限定するものではない。

[0021] なお、以後の説明では「a a a テーブル」、「a a a リスト」等の表現にて本発明の情報を表記している箇所があるが、これら情報はテーブル、キュー等のデータ構造以外で表現されていてもよい。そのため、データ構造に依存しないことを示すために「a a a テーブル」、「a a a キュー」、「a a a リスト」等について「a a a 情報」と呼ぶことがある。

[0022] 以後の説明では「プログラム」を主語として説明を行う場合もあるが、プログラムはプロセッサによって実行されることで定められた処理をメモリ及び通信ポートを用いながら行うため、これは実際にはプロセッサが動作主体となって所定の処理を実行していることを表している。また、プログラムの一部または全ては専用ハードウェアによって実現されてもよい。

[0023] また、各図において共通の構成物については、同一の参照番号を付して説

明する。また、共通の構成物に関し、各構成物を識別する場合には、999a、999bなど、数字の最後に英文字を付すか、又は、#1、#2等の個別の番号を付して説明する。ただし、必要に応じて英文字又は番号を省略して説明する場合がある。

[0024] また、本明細書に記載の各種プログラムは、プログラム配布サーバや、計算機が読み取り可能な記憶メディアによって提供されるようにしてもよい。

[0025] 本発明は、以下の実施例に限定されない。以下の実施例における不揮発性メモリは、FM (Flash Memory) である。本実施例におけるFMは、ブロック単位で消去が行われ、ページ単位でリード及びライトが行われる種類のFM、典型的にはNAND型のフラッシュメモリである。しかし、FMは、NAND型に代えて他種のフラッシュメモリ（例えばNOR型）でも良い。また、FMに代えて、他種の不揮発性メモリ、例えば、磁気抵抗メモリであるMRAM (Magnetoresistive Random Access Memory) や、抵抗変化型メモリであるReRAM (Resistance Random Access Memory)、強誘電体メモリであるFeRAM (Ferroelectric Random Access Memory) 等の半導体メモリや、相変化メモリが使用されても良い。

実施例 1

[0026] 図1は、本発明の実施例1に係る計算機システムの構成を示す。

[0027] 計算機システムは、ストレージシステム（ストレージ装置）101と、1つ以上のホスト計算機（図中では「ホスト」と略記する）103と、管理端末104とを有する。ホスト計算機103a、103bのそれぞれは、SAN (Storage Area Network) 105を介して、ストレージシステム101に接続される。

[0028] ストレージシステム101は、ストレージコントローラ102と、不揮発メモリ記憶装置である複数のフラッシュメモリストレージ装置112とを有する。なお、図面及び以下の説明では、フラッシュメモリストレージ装置を

FMPKG (Flash Memory Package) と呼ぶことがある。また、本実施例では、ストレージコントローラ 102 は 1 つであるが、複数のストレージコントローラ 102 で冗長構成されても良い。

[0029] ストレージコントローラ 102 は、CPU (Central Processing Unit) 108 と、メモリ 109 と、複数のホストインターフェイス (以下、「ホスト I/F」と略記する) 107 と複数のドライブインターフェイス (以下、「ドライブ I/F」と略記する) 110 と、保守インターフェイス (以下、「保守 I/F」と略記する) 106 とを有する。ストレージコントローラ 102 内の各部は、バスを介して接続される。メモリ 109 は、ストレージシステム 101 を制御するためのプログラムを格納する領域及びデータを一時的に格納するキャッシュメモリとしての領域を有する。CPU 108 は、メモリ 109 に格納されたプログラムに従ってストレージシステム 101 を制御する。

[0030] ホスト I/F 107 は、計算機 103 との通信を行うためのインターフェイスである。ドライブ I/F 110 は、ストレージコントローラ 102 と FMPKG 112 との通信を行うためのインターフェイスである。保守 I/F 106 は、管理端末 104 と接続され管理端末 104 との通信を行うためのインターフェイスである。

[0031] なお、管理者は、管理端末 104 からストレージコントローラ 102 の管理やメンテナンスを行う。しかし、管理端末 104 は必須要素ではなく、ストレージコントローラ 102 の管理やメンテナンスなどを、例えば、ホスト計算機 103 から行うようにすることも可能である。

[0032] 上記の計算機のシステムでは、ホスト計算機 103 と FMPKG 112 とが、ストレージコントローラ 102 を介して接続されているが、例えば、ストレージコントローラ 102 を省き、ホスト計算機 103 と FMPKG 112 とが直接接続されても良い。

[0033] 以下、FMPKG 112 の構成を説明する。なお、各 FMPKG 112 の基本的な構成は同じである。

[0034] 図2は、FMPKG112の構成を示す。

[0035] FMPKG112は、フラッシュメモリ制御装置201と1つ以上のフラッシュメモリチップ（FMチップとも呼ばれる。また図中では「FM Chip」と表記）209a～hにより構成される。また、先に述べたとおり、FMチップに代えて他種の不揮発性メモリ素子を用いることも可能である。

[0036] フラッシュメモリ制御装置201は、図1におけるストレージコントローラ102あるいはホスト計算機103のように、FMPKG112に対してデータアクセス要求を発行する上位装置（以下、ストレージコントローラ102またはホスト計算機103を、「上位装置102」と呼ぶ）とFMPKG112とを接続するためのストレージインターフェイス202、ASIC203、バッファ204、非常時に給電を行うバッテリー205、CPU206、メインメモリ207、フラッシュメモリとの接続を行うフラッシュメモリインターフェイス208により構成される。

[0037] CPU206は、フラッシュメモリ制御装置201の全体を制御するプロセッサであり、メインメモリ207に格納されたマイクロプログラムに基づき動作する。例えば、受領したI/Oに対して、メインメモリ207に格納してある論物変換テーブルを参照してフラッシュメモリチップ209にデータを読み書きする。また、フラッシュメモリの使用状況に応じて、リクレーションやウェアレベリングを実施する。

[0038] ASIC203は、CPU206によって制御され、フラッシュメモリチップ209に対してデータの読み書きを実施する。また、フラッシュメモリチップ209に対するものと同様にストレージインターフェイス202を通じて上位装置との通信を制御する。

[0039] 本図では、ASIC203に対してCPU206は外付けとなっているが、ASIC203は、CPU206や、ストレージインターフェイス202、フラッシュメモリインターフェイス208を含む1つのLSIで構成されていてもよい。

[0040] メインメモリ207とバッファ204は、より高速なアクセスが可能な、

揮発性の記憶領域であり、例えばDRAMである。メインメモリ207はCPU206が直接制御に用いるワークスペースであり、バッファ204より短いレイテンシを提供する。一方、バッファ204は、ユーザデータのバッファ用や、メインメモリ207に格納しきれない大サイズテーブルを格納する。

[0041] 本図では、メインメモリ207とバッファ204は分離した構成物として
いるが、これらを1つの記憶素子で実現してもよい。

[0042] ストレージインターフェイス202は、上位装置とフラッシュメモリ制御
装置201を接続し、通信を行う。このインターフェイスには、例えばS
ATAやSAS、FCなどのストレージインターフェイスが採用されてもよい
し、PCI-Expressなどのインターフェイスが採用されてもよい。

[0043] フラッシュメモリ制御装置201の内部アーキテクチャは必ずしも図の通
りである必要は無く、それぞれの機能を1つないしは複数のデバイスによ
って代用してもよい。

[0044] 図3は、FMPKG112のメインメモリ207に格納される情報の一例
を示す。

[0045] メインメモリ207は、例えば、オペレーティングシステム303、フラ
ッシュストレージ制御プログラム302、データ転送制御部制御プログラム
301、入出力制御部制御プログラム304、論物変換プログラム306及
び論物変換情報305を記憶する。

[0046] オペレーティングシステム303は、CPU206が各プログラムを実行
する際のスケジューリング等の基本処理を行うプログラムである。

[0047] フラッシュストレージ制御プログラム302は、フラッシュメモリ制御装
置201が上位装置へ提供するボリューム（あるいはボリュームの記憶空間
）の管理や、バッファメモリの管理など、フラッシュメモリ制御装置201
がストレージデバイスとして動作するための制御に用いるプログラムである
。

[0048] データ転送制御部制御プログラム301は、ASIC203の制御に用い

るプログラムである。

[0049] 入出力制御部制御プログラム304は、ストレージインターフェイス202、フラッシュメモリインターフェイス208の制御に用いるプログラムである。

[0050] 論物変換プログラム306は、上位装置102から発行された入出力要求（I/O要求）に含まれている論理アドレスを、フラッシュメモリ上の物理位置である物理アドレスに変換するプログラムである。なお、本実施例で言う「論理アドレス」は、例えば、LBA(Logical Block Address)で良い。また「物理アドレス」とは、例えばFMPKG112内の物理ページ及び／または物理ブロックを一意に識別するアドレス（ページ番号、ブロック番号）のことである。

[0051] 論物変換情報305は、論物変換プログラム306が動作する際に使用する変換のための情報である。

[0052] 図4は、フラッシュメモリチップ209の内部構成を示す。

[0053] フラッシュメモリチップ209は、内部に複数のダイ(DIE)401a～b、フラッシュメモリ制御装置201から発行されたフラッシュメモリI/Oコマンドの対象データを一時的に格納するページバッファ402a～bを持つ。そして各ダイ401は、実際の記憶領域である、1つ以上（たとえば128個、256個など）の物理ブロック403a～cを有する。

[0054] 物理ブロック403は、フラッシュメモリにおける消去単位である。物理ブロック403は、1つ以上の物理ページ404a～cにより構成される。物理ページはフラッシュメモリにおける書き込み・読み込み単位である。

[0055] 書き込みとは、消去済みの物理ページに対するデータの書き込みであり、読み込みとは物理ページに書き込まれたデータを読み取ることである。原則として、書き込みが行われた物理ページに再度書き込みを行うことはできず、物理ページに再度書き込みを行う場合には、一度、当該物理ページを含む物理ブロック403単位で消去を実施する必要がある。消去処理により、物理ブロック403に含まれる物理ページ404に格納されたデータは消去さ

れ、再び書き込みが可能となる。1つの物理ページのみを消去することは出来ない。書き込み・読み込み・消去などのコマンドは、フラッシュメモリチップ209とフラッシュメモリ制御装置201を接続するフラッシュメモリインターフェイス208から受領する。

[0056] 図5は物理ページ404内の一般的なデータ格納構成を示す。

[0057] Code Word (CW) 501は、データ502と、それを保護するECC 503の1セットであり、エラー訂正単位である。物理ページには1つ以上のCWが存在する。物理ページは一般的に、2の冪乗のバイト数分の領域に加えて余剰の領域を保持しており、この余剰領域にECCやメタ情報を格納する。ここではメタ情報は特に無いものとして記述するが、存在したとしても特に問題無い。ECCサイズは、チップに要求される信頼性レベルに応じてサイズが決まるため、1物理ページ内に必ずしも2の冪乗分のデータ領域サイズが確保できないこともある。以降、物理ページサイズといった場合、元のFMの物理ページのサイズからECCやメタ情報を格納する領域を含まない、ユーザデータを格納する領域のサイズを指す。

[0058] 図6は、従来からあるフラッシュストレージで行われている、ページベースの論物変換処理の概要を示している。図中の要素600は、フラッシュストレージが上位装置に対して提供するボリュームの記憶空間を表しており、上位装置がフラッシュストレージの記憶領域にアクセスする際の最小アクセス単位である各セクタ（図中の要素601）に、0番から始まるアドレスが付されており、そのアドレスのことを論理ブロックアドレス（LBA）と呼ぶ。一般に、1セクタのサイズは512バイトであるが、その他の大きさ（たとえば520バイト）でもよい。またフラッシュストレージは、フラッシュストレージが上位装置に対して提供する記憶空間を先頭から所定サイズ（たとえば図6に示した例では16セクタ）ごとに分割し、分割された各領域を「論理ページ」（図中の要素602）として管理している。また、図6において、ブロックX（603）中のページY（604）は、物理ページを表している。

[0059] 従来からあるフラッシュストレージでは、全ての論理ページ602、全ての物理ページ604のサイズは固定サイズで（図6の場合、16セクタ、つまり8KB）、また論理ページ602と物理ページ604のサイズは等しい。従って、1つの論理ページ602は任意の1つの物理ページ604に一意に対応づけられ、その対応関係についての情報を論物変換テーブル701にて管理する。ある論理ページ602への書き込みを行う場合、新規に未使用の物理ページ604が取得され、当該論理ページ602の更新データをその未使用物理ページ604へ書き込み、対応付けを更新する。論理ページ602からの参照が外れた旧物理ページ604は、無効なページとして登録され、リクラメーションによって回収・消去され、空きページとして再利用される。

[0060] 図7は、ページベース論物変換のための論物変換テーブル701の構成の一例を示している。このテーブルは従来からあるフラッシュストレージが通常有するテーブルである。

[0061] 論物変換テーブル701は、各論理ページ602に対応する物理ページ604のアドレスを管理するテーブルで、各論理ページ602の論理ページ番号（702）と、当該論理ページ番号702の論理ページ602が対応付けられている物理ページ604のアドレス（格納アドレス）703を保持している。たとえばフラッシュストレージが上位装置に対して、N論理ページ分の大きさの記憶空間を提供している場合、N個のエントリが存在する。論理ページ602の格納先が、異なる物理ページ604に変更された場合、このテーブルの格納アドレス703の情報が更新される。

[0062] 次に、図8、図9を用いて、論理ページサイズが物理ページサイズと異なる場合の問題点について説明する。図8のCase1は、論理ページサイズが物理ページサイズの1/2の場合の例を示しており、Case2は、論理ページサイズが物理ページサイズの1/2よりも大きい例を示している。

[0063] 図8のCase1のように、論理ページサイズが物理ページサイズの1/2の場合、あるいは1/4、1/8のように、2の冪乗分の1のサイズであ

れば、1 物理ページ中に複数の論理ページがちょうど格納される。そのため、論理ページサイズが物理ページサイズの2の冪乗分の1のサイズである場合は、物理ページを論理ページサイズと等サイズの領域（仮にこれを「スロット」と呼ぶ）に分割して管理し、図7を用いて説明した論物管理テーブルでは、各論理ページの論理ページ番号と、物理ページ上のスロットとの対応関係を管理するように変更するだけで対応可能である。しかしながら図8のCase 2のように、論理ページサイズが物理ページサイズの2の冪乗分の1のサイズと等しくない場合は、必ず物理ページ上に、1つの論理ページが収納しきれない領域（無効領域）ができる。論理ページを2つ以上の物理ページに対応づけない格納形態を採用する場合、それらの領域は無効とするしかない。この無効領域が大きいほど、格納効率は低下し、ビットコストの悪化につながる（ビットコストが高くなる）。

[0064] 図9は、1つの論理ページが複数の物理ページに跨がって格納される場合の、データ格納構成を示す。

[0065] 先に説明したとおり、論理ページサイズと物理ページサイズが異なる場合、物理ページ上に1論理ページを収納しきれない無効領域ができる。物理ページに無効領域ができないようにするために、1つの論理ページを複数の物理ページに跨って格納される態様を許す方法が考えられるが、そのようにすると、以下のような問題がある。論理ページは、1つの物理ページ内に収まる場合（図9の論理ページ901の場合）もあれば、2つ以上の物理ページに跨がって格納される場合（図9の論理ページ902の場合）もある。論理ページに対する物理ページのアドレスを管理する場合、論理ページ（LP=0）901は単一物理ページに収まっているため、論理ページ901について、当該論理ページの格納されている物理ページのアドレス及び当該物理ページ内オフセット位置の情報を持てば良い。一方で論理ページ（LP=1）902は2つの物理ページ903a、903bに跨がっているため、論理ページに対する物理ページのアドレスを管理する場合、論理ページ（LP=1）902の前半（たとえば論理ページの0バイト目からaバイト目まで）は

物理ページ 903 a の一部 (x バイト目から終端まで) に格納され、また論理ページ (LP = 1) 902 の後半 (たとえば論理ページの a + 1 バイト目から終端まで) は物理ページ 903 b の一部 (0 バイト目から y バイト目まで) に格納される、という情報を管理する必要がある。つまり 1 論理ページの情報を管理するために、論理ページと物理ページの対応を管理するための情報 (論物変換テーブル) に 2 エントリを登録する必要があるため、論物変換テーブルのサイズが大きくなる。

[0066] 続いて、本発明の実施例 1 に係る FMPKG 112 における、論理ページと物理ページの対応関係の管理方法について説明する。なお、本発明の実施例 1 に係る FMPKG 112 において、各論理ページのサイズは等しく、また各物理ページのサイズも等しい。ただし論理ページのサイズと物理ページのサイズは一致していない (論理ページのサイズと物理ページのサイズが一致している場合であっても、本発明の実施例に記載の FMPKG 112 は動作可能であるが、以下では論理ページのサイズと物理ページのサイズは一致していないという環境であるという前提で説明する)。

[0067] 図 10 は、FMPKG 112 が実施する論理物理変換の概念を表している。FMPKG 112 の実施する論理物理変換は、図 6、7 を用いて説明したページベース論物変換の概念に、仮想アドレス空間の概念を組み込んだものである。なお、図 10 では図示を省略しているが、図 6 と同様、FMPKG 112 は上位装置に対してアドレス 0 から始まる論理ブロックアドレス (LBA) でアドレッシングされた一次元の論理記憶空間 (図 6 の 600 相当物) を提供し、この一次元の論理空間を先頭から所定サイズ (たとえば 16 セクタ) ごとに分割し、分割された各領域を論理ページとして管理している。つまり一次元の論理記憶空間の論理ブロックアドレス (LBA) の 0 番地から 15 番地までの領域は、論理ページ番号 (LP) が 0 の論理ページ 1001 a に対応付けられ、以下順に、一次元の論理記憶空間の LBA が 16 番地から 31 番地の領域が、論理ページ番号 (LP) が 1 の論理ページ 1001 b に、一次元の論理記憶空間の LBA が 32 番地から 47 番地の領域が、論

理ページ番号（LP）が2の論理ページに対応付けられている。なお、論理ページのサイズは任意であるが、後述するようにFMPKG112は論理ページ単位で更新を行う（論理ページ未満のサイズのデータのライト要求が来た場合でも、当該データを含む論理ページ全体を書き換える）ため、上位装置、たとえばホスト計算機103で稼働するOSやファイルシステムやアプリケーションプログラムが発行するI/O要求で指定されるデータサイズ（たとえば4KBや8KB）等を選択するのが好ましい。

[0068] 本発明の実施例1に係るFMPKG112では、論理ページと物理ページとの対応を直接管理するわけではなく、各論理ページ（1001a、1001b、…）は、FMPKG112内でのみ用いられる1次元のアドレス空間である、仮想アドレス空間1003上のアドレスに対応付けられている。この論理ページと仮想アドレス空間1003上アドレスとの対応関係についての情報は、後述するページマッピングテーブル1002（図11）を用いて管理される。仮想アドレス空間1003は、物理ページ境界を意識させない1次元の線形なアドレス空間であり、FMPKG112はこの仮想アドレス空間1003を構成する情報を、後述する仮想アドレス構成情報1007（図12）を用いて管理している。なお、仮想アドレス空間1003が、物理ページ境界を意識させないアドレス空間である、とは、論理ページを仮想アドレス空間1003に対応付ける場合、論理ページが物理ページを跨って格納されないように対応付ける等の考慮をせずに、単純に、仮想アドレス空間上の任意の空き領域（論理ページがまだ対応付けられていない仮想アドレス空間上の領域）に論理ページを対応付ければよいことを意味する。

[0069] 仮想アドレス空間1003はFMPKG112内に1つだけ定義される態様に限定されるわけではなく、複数存在し得る。仮想アドレス空間が複数存在する場合、本明細書では、それぞれの仮想アドレス空間を仮想アドレスグループ1004と呼ぶ（以下、仮想アドレスグループのことを単にグループと呼ぶこともある）。また、それぞれのグループに1つの仮想アドレス構成情報1007が対応づけられている。グループが複数存在する場合、各論理

ページはどのグループのアドレスに対応づけられてもよい。

[0070] 仮想アドレス空間 1003 は、FM の物理ブロック 1008 と同サイズに分割され、各分割された領域は、「仮想ブロック 1006」と呼ばれる。各仮想ブロック 1006 には、仮想アドレスグループ 1004 内で一意な識別番号（仮想ブロック番号と呼ぶ）が付されている。仮想アドレスグループ 1004 内の先頭の仮想ブロック（図 10 の 1006 a）には、0 番の仮想ブロック番号が付されており、以下、2 番目以降の領域には順に、1 番、2 番…の仮想ブロック番号が付されている。

[0071] また、それぞれの仮想ブロック 1006 は、FM 上の物理ブロック 1008（図 4 の 403 に相当する）へ 1 : 1 にマップされる。このマッピング情報は仮想アドレス構成情報 1007 に管理される。これにより、仮想アドレス空間 1003 には、複数の物理ブロックがマッピングされることになる。

[0072] ここからわかるとおり、たとえば図 10 の物理ページ 1009 a と 1009 b が同一の仮想アドレス空間の構成物であるという情報は、ページ毎ではなく、ブロック単位で持つことになる。これにより、論理ページ毎に物理ページ間の関係を保持する必要は無くなるため、管理情報を大幅に減らすことができる。ただし、1 論理ページのデータがグループを跨がって格納されることは許容されないため、グループの最終アドレス付近では、バウンダリを合わせるためにユーザデータを格納しない領域が発生する。これらの領域は、無効な領域としても良いし、何かの管理情報を格納する領域として利用しても良い。

[0073] 仮想アドレス空間を構成する要素として、上では物理ブロックを前提としたが、物理ページより大きな物理アドレス構成要素であれば、管理情報を減らすという目標は実現可能である。ただし、消去単位である物理ブロックと対応づけることは、管理上のメリットが大きい。例えば、ある物理ブロックを消去する際に、これが複数のグループから共有されている場合、それぞれのグループについてコピー処理を実施する必要がある。他にも、従来物理ブロックを単位に実施してきた処理を、グループ単位に拡張する場合、物理ブ

ロックはグループ間で共有されるべきではない。以降、グループ構成要素としては、物理ブロックを想定して記述する。

[0074] 図11は、実施例1に係るFMPKG112が管理するページマッピングテーブル1002の構成を表している。ページマッピングテーブル1002は、1つのFMPKG112に一つ定義されるテーブルで、論理ページと仮想アドレス空間1003上のアドレス（以下、仮想アドレス空間上のアドレスのことを「仮想アドレス」とよぶ）との対応付けを管理するものである。またページマッピングテーブル1002は、メインメモリ207の論物変換情報305を構成する1つの情報である。論理ページ毎に設定される仮想アドレスは、仮想アドレスグループ番号（グループ#）1102とグループ内のオフセットアドレス1103とで構成される。オフセットアドレス1103には、グループ#1102で特定されるグループの先頭セクタのアドレスを0とした相対アドレスが格納される。各論理ページが初期状態、つまり論理ページにデータが書き込まれていない場合、グループ#1102及びオフセットアドレス1103には値が格納されていない。論理ページに対するデータの書き込み要求を上位装置から受け付けると、FMPKG112は当該論理ページに対応付けるべき仮想アドレス空間上の未書き込み領域アドレス（グループ番号及びグループ内オフセットアドレス）を決定し、グループ#1102及びオフセットアドレス1103に値を格納する。また、論理ページ上のデータが更新（上書き）される時にも同様に、仮想アドレス空間上の未書き込み領域のアドレス（グループ番号及びグループ内オフセットアドレス）を決定し、グループ#1102及びオフセットアドレス1103に値を格納する。

[0075] なお、本発明の実施例においては、オフセットアドレス1103の単位はセクタとしているが、その他の単位を用いても良い。たとえば1論理ページのサイズを1単位としたアドレスをオフセットアドレス1103に格納するようにしても良い。また、先に述べたとおり、論理ページがグループを跨がることは許容されないため、1つの論理ページ（論理ページ1101で特定

される論理ページ）が、複数のグループに対応付けられることはない。そのため、ページマッピングテーブル１００２には、FMPKG１１２で管理する論理ページの数と同数の行（エントリ）のみが存在する。

[0076] 図１２は、実施例１に係るFMPKG１１２が管理する仮想アドレス構成情報１００７の構成を表している。この情報もページマッピングテーブル１００２と同様、メインメモリ２０７の論物変換情報３０５を構成する１つの情報である。

[0077] 仮想アドレス構成情報１００７は、仮想アドレスグループ（グループ１２０１）毎に、次の書き込み位置を保持するポインタである次書き込み位置１２０２（なお、本発明の実施例では、単位はセクタとするが、その他の単位を用いてもよい）と、対応するグループの仮想アドレス空間を構成する複数の物理ブロックの物理ブロック番号（物理ブロック＃）１２０４を保持する。本発明の実施例に係るFMPKG１１２では、FMPKG内の各物理ブロックについて、FMPKG内で一意な識別番号を付して管理しており、この識別番号を物理ブロック番号と呼ぶ。物理ブロック番号１２０４の欄には、この物理ブロック番号が格納される。また物理ブロック番号１２０４は、１グループ内にN個が登録され、各グループの先頭に登録されている物理ブロック番号は、０番の仮想ブロック番号（１２０３）に対応付けられており、それ以降の各物理ブロックは順に、１番、２番…の仮想ブロック番号（１２０３）に対応付けられている。仮想アドレスグループを構成する物理ブロック数Nの大きさによって、仮想アドレスグループ内のアドレス空間の大きさが定まる。

[0078] 次書き込み位置１２０２は、対応する仮想アドレスグループ１２０１の先頭未書き込み位置を表す情報である。詳細は後述するが、本発明の実施例に係るFMPKGでは、上位装置１０２からのライトデータを仮想アドレスグループ（に対応付けられた物理ブロック・物理ページ）に書き込む際、グループ内の未書き込み領域の先頭から順に、追記書きの要領で書き込む。またFMPKGが、物理ブロック（内の物理ページ）にデータを書き込む際にも

、物理ブロック中の未書き込みの先頭ページから順に書き込んでいく（たとえば当該物理ブロックに全くデータが書き込まれていない場合、先頭ページはページ番号 0 番のページであり、以下 1 番、2 番…のページへと、順番に書き込みを行う）。そして物理ブロック内全物理ページにデータが書き込まれるまでは、次の物理ブロックの物理ページへのデータ書き込みは行わない。そのため、次書き込み位置 1 2 0 2 より前のアドレスで特定される仮想アドレス空間には、すでにデータが格納され論理ページにマッピングされた領域であることを意味し、次書き込み位置 1 2 0 2 以降のアドレスで特定される空間は、未使用領域、つまりデータがまだ書き込まれておらず、論理ページにもマッピングされていない領域であることを表している。初期状態では次書き込み位置 1 2 0 2 の値は 0 で、ライト処理によってたとえば n セクタの値が書き込まれると、次書き込み位置 1 2 0 2 に値 n が加算される。仮想アドレスグループの終端物理ページ（の近く）までデータが書き込まれると、それ以上当該仮想アドレスグループにデータを書き込むことができないため、次書き込み位置 1 2 0 2 には、データが書き込めないことを表す値 (N/A) を格納する。

[0079] また、仮想アドレス構成情報 1 0 0 7 の中には、次書き込み位置 1 2 0 2 に、0 以上の値（つまり N/A でない値が格納されている）が格納されていて、書き込み可能になっているグループが 1 つだけ存在する。上位装置からのライトデータを格納する際には、原則としてそのグループに対してデータを書き込む。以下ではそのグループのことを、「オープンされたグループ」または「オープングループ」と呼ぶ。そして、オープンされたグループにデータを書き込む領域がなくなった時点で、新たにグループを構築する。新たにグループを構築する際に、グループを構成する物理ブロック（複数）を選択し、物理ブロック番号 1 2 0 4 の欄に物理ブロック番号を格納する。ただし、オープングループを複数設けておき、FMPKG 1 1 2 がデータを物理ページに書き込む際に任意のオープングループを選択して書き込むようにしてもよい。

[0080] また、本発明の実施例 1 における仮想アドレス構成情報 1007 では、グループごとに無効ページ数 1205、最終書き込み時刻 1206 の情報を保持している。無効ページ数 1205 は、グループ内の物理ページのうち、データの書き換えが発生したために無効になったページの数で格納され、また最終書き込み時刻 1206 には、当該グループに最後にデータが書き込まれた時刻が格納される。これらの情報は、リクレーション（本明細書では RC と略記されることもある）あるいはリフレッシュ（本明細書では、RF と略記されることもある）処理で用いられる。本発明の実施例 1 に係る FMPKG112 では、これら情報をグループごとに管理しているが、物理ブロック毎など、その他の単位で管理してもよい。また、RC、RF 処理などの用途のために、無効ページ数 1205、最終書き込み時刻 1206 以外の情報を管理するようにしても良い。

[0081] 続いて図 13 を用いて、本発明の実施例 1 に係る FMPKG112 が実行する、論理アドレスから物理アドレスへの変換処理（正引き処理 s1301）のフローを説明する。正引き処理 s1301 は、CPU206 が実行しているフラッシュストレージ制御プログラム 302 が、たとえば上位装置からリード要求等のアクセス要求を受信した時に、当該アクセス要求でアクセス対象となるデータの物理格納位置（物理ページ）を算出するために実行される。この場合、フラッシュストレージ制御プログラム 302 が論物変換プログラム 306 を呼び出して、CPU206 に正引き処理 s1301 を実行させ、論物変換プログラム 306 は正引き処理 s1301 の処理の結果、算出された物理アドレス情報（物理ブロック番号と、当該物理ブロック番号で特定される物理ブロック内物理ページ番号の組）を要求元であるフラッシュストレージ制御プログラム 302 に返却する。以下では、リード要求で指定されるアクセス対象データの先頭位置は論理ページ境界と一致しており、かつアクセス対象データのサイズが 1 論理ページサイズと等しい場合、つまり 1 論理ページ分の領域の物理位置の算出処理について説明するが、それ以外の場合でも同様の処理を行うことで、アクセス対象データの物理格納位置の算

出ができる。

[0082] まずCPU 206は、処理対象のコマンドに含まれる、アクセス対象データのアドレス（LBA）の情報から、アクセス対象となる論理ページの論理ページ番号を算出する（s 1302）。これはアクセス対象データのLBAを論理ページサイズで除算することで算出される。続いてCPU 206は、ページマッピングテーブル1002を参照して、論理ページ1101がs 1302で算出した論理ページ番号に一致する行（エントリ）を特定し（s 1303）、当該行のグループ番号1102と、グループ内のオフセットアドレス1103を取得する（s 1304）。

[0083] 続いてCPU 206は、s 1304で取得したオフセットアドレス1103から、仮想ブロック番号と、仮想ブロック内のオフセットアドレス（仮想ブロックの先頭位置のアドレスを0とした相対アドレス）を計算する（s 1305）。なお、s 1305では、s 1304で取得したオフセットアドレス1103を物理ブロックサイズ（1物理ページのサイズ×1物理ブロック内ページ数）で除算し、このときの商を仮想ブロック番号とし、剰余を仮想ブロック内のオフセットアドレスとする。そしてs 1306では、CPU 206は仮想アドレス構成情報1007を参照し、グループ1201がs 1304で取得したグループ番号と等しく、かつ仮想ブロック番号1203がs 1305で算出した仮想ブロック番号と等しい行を参照し、対応する物理ブロック番号（物理ブロック#1204）を取得する。s 1306で取得した物理ブロック番号で特定される物理ブロックの中に、アクセス対象データが存在することが分かる。

[0084] s 1307ではCPU 206は、s 1305で算出した仮想ブロック内オフセットアドレスを物理ページサイズで除算することで、s 1306で取得した物理ブロック番号で特定される物理ブロック内の相対ページ番号（物理ブロック内先頭ページの番号を0としたページ番号）を割り出し、上位装置からのアクセス要求に含まれるアドレスが対応する物理ページ（アクセス対象データの格納されている先頭の物理ページ）の物理ページ番号を算出する

。また、本発明の実施例に係る F M P K G 1 1 2 の場合、1つの論理ページが複数の物理ページに跨って格納される場合があるので、s 1 3 0 5 で算出した仮想ブロック内オフセットアドレスに「論理ページサイズ-1」を加算した値を物理ページサイズで除算し、アクセス対象データの格納されている終端物理ページの物理ページ番号を算出し、算出された先頭物理ページ及び終端物理ページの物理アドレス（物理ブロック番号と物理ページ番号の組）を要求元に返却することで、処理を終了する（ただし先頭物理ページと終端物理ページの物理番号が同一の場合には、1つの論理ページが1物理ページに収まっていることを意味するので、その場合には1つの物理ページ番号のみが要求元に返却される）。物理アドレスを受け取った要求元は、例えば要求元が行っていた処理がリード処理だった場合には、受け取った物理アドレスに該当する物理ページからデータを読み出し、上位装置に返却する等の処理を行う。

[0085] 続いて、F M P K G 1 1 2 が上位装置 1 0 2 からのライトデータを、F M チップの物理ページに格納する処理（ライト処理）の流れについて説明する。仮に、ライトデータを論理ページ単位で書き込むことを想定した場合、ライト処理の流れの概略は以下の通りになる。F M P K G 1 1 2 が論理ページに対するライト要求及びライトデータを受信すると、仮想アドレス構成情報 1 0 0 7 を用いて未書き込み領域（まだ論理ページが対応付けられていない仮想アドレス空間上の領域）のアドレス（グループ番号及びグループ内オフセットアドレス）を決定し、当該決定された領域のアドレス情報をページマッピングテーブルのグループ 1 1 0 2 及びオフセットアドレス 1 1 0 3 の欄に登録する。そして決定された未書き込み領域のアドレスに対応する1または複数の物理ページを特定し、その特定された物理ページに対してライトデータを格納すればよい。

[0086] 本発明の実施例に係る F M P K G 1 1 2 では、基本的には上記の流れで処理を行えば、論理ページ単位でライトデータを物理ページに格納することは可能である。ただし F M チップへのデータ書き込み最小単位はページ（物理

ページ)である、という制約があるため、上記の流れの処理では物理ページに無効領域が発生することがある。実際のライト処理では、物理ページに無効領域が発生しないように、複数の論理ページのデータをFMPKG112内のバッファ204に蓄積してから、まとめて1または複数の物理ページに書き込む。そのため上記の説明とはやや異なった順で処理が行われる。

[0087] 図14を用いて、FMPKG112が上位装置102からライト要求を受信した時に実行する、ライト処理s1401の流れを説明する。この処理は、CPU206がメインメモリ207内のフラッシュストレージ制御プログラム302や論物変換プログラム306を実行することによって実現される。なお、以下では主に、上位装置102から、ライト対象データの先頭アドレスと終端アドレスが論理ページのページ境界に合致するデータのライト要求を受信した場合について説明する。

[0088] ライトコマンド処理s1401は、上位装置からのライト要求に対する処理である。まずCPU206は、ライトコマンドをs1402で取得する。なお、s1402で取得するライトコマンドには、ライトデータの格納位置(ライトアドレス。通常はLBAが指定されている)及びライトデータ長(セクタ数)が含まれている。

[0089] s1403では、CPU206はバッファ204にライトデータを格納する。その際、ライトアドレス及びライトデータ長から、ライトデータの格納される論理ページの論理ページ番号を算出し(なお、ライトデータ長が複数ページ分の長さの場合、変換によって導出される論理ページ番号は複数ある)、バッファにライトデータを、この変換処理により導出された論理ページ番号とともに格納する。このようにライトデータを論理ページ番号と対応付けて格納しておくことで、バッファ上ライトデータがFMに書き込まれる前に、上位装置からリード・ライト要求が到来したときには、バッファ204に格納されたデータに対するリード、ライトを行うことで、上位装置からの要求を処理することが出来る。バッファ204にライトデータを格納した後、CPU206は上位装置へ完了通知を行う(s1404)。上位装置への

完了通知は、FMにデータを格納する前に行ってよい。

[0090] 続いて、s 1 4 0 5 で、CPU 2 0 6 はバッファ 2 0 4 上に一定以上の量のデータが存在するか確認し、存在した場合には、バッファ 2 0 4 上のデータをFMの物理ページへ書き込む処理を行う。本発明のFMPKG 1 1 2 では、論理ページが複数の物理ページに跨がって格納されることを許容しているが、一方でFMの最小書き込み単位は、物理ページである。そのため、論理ページ単位、あるいは物理ページ単位で、FMへ書き込みを行おうとすると、論理ページの全体または断片が物理ページの一部にのみ書き込まれ、物理ページのそれ以外の領域が無効領域になる事態が発生する。無効領域が極力発生しないようにするために、本発明の実施例における書き込み処理では、ある程度多くの量のデータをまとめてFMに書き込むようにする。好ましくは、物理ページサイズと論理ページサイズの公倍数に等しいサイズのデータがバッファに蓄積された時点で、FMに書き込みを行うとよい（たとえば論理ページサイズが8KB、物理ページサイズ[先に述べたとおり、ここでいう「物理ページサイズ」とは、ECC等の付加情報を除く、上位装置から到来したデータを書き込む領域のサイズである]が7KBの場合、8と7の最小公倍数である56KBあるいは56KBの倍数に等しい量のデータがバッファに蓄積された時点で、FMへの書き込みを行うと、無効領域が発生しないため、記憶領域を有効に活用できる書き込みが実現される）。

[0091] s 1 4 0 6 では、CPU 2 0 6 はオープングループのグループ番号を取得する。オープングループとは、先に述べたとおり、書き込み可能になっているグループのことを意味する。続いてs 1 4 0 7 でCPU 2 0 6 は、取得されたオープングループ内の、次書き込み位置 1 2 0 2 及び今回FMに書き込むデータ量を元にして、書き込み先物理ブロックの物理ブロック番号 1 2 0 4 及び物理ブロック内ページ番号を算出する。物理ブロック番号 1 2 0 4 及び物理ブロック内ページ番号の算出方法であるが、次書き込み位置 1 2 0 2 を「1 物理ブロック内物理ページ数×物理ページサイズ（物理ページサイズの単位はセクタ数）」で除算することで、データ書き込み対象となる先頭物

理ブロックの仮想ブロック番号が算出でき、また仮想アドレス構成情報 1007 を参照することで、算出された仮想ブロック番号に対応する物理ブロック番号 1204 を特定できる。また次書き込み位置 1202 を「1 物理ブロック内物理ページ数×物理ページサイズ」で除算したときの剰余を求めることで、今回の書き込み先となる物理ブロック内ページの先頭ページ番号（先頭物理ページ番号）を取得できる。さらに今回 FM に書き込むデータ量が複数物理ページ分にわたる場合、必要とされる物理ページ数分のページ番号を連番で取得する（物理ページを 5 ページ分取得する必要がある、かつ今回の書き込み先となる物理ブロック内ページの先頭物理ページ番号が 3 の場合、ページ番号 3、4、5、6、7 のページ番号を取得する）。

[0092] s 1408 では CPU 206 は、バッファ上のデータを、s 1407 で取得した物理アドレス（書き込み先物理ブロックの物理ブロック番号 1204 及び物理ブロック内ページ番号）で特定される物理ページに順次書き込む。この時、無効領域が発生しないように、バッファ上の複数論理ページ分のデータを連続して物理ページに書き込む（たとえば図 9 のように）。なお、物理ページに格納された各データについて、当該データの格納された物理ページ位置に対応する仮想アドレス空間上のアドレス（グループ番号、グループ内オフセットアドレス）を算出し（データの格納された物理ページの物理ブロック番号、物理ページ番号、物理ページ内オフセット位置と、仮想アドレス構成情報 1007 とから算出可能である）、それを各データの論理ページ番号（バッファ上に記憶されている）と対応付けて一時的に記憶しておく。これは s 1409 の管理情報の更新で必要となる。その後 s 1409 で、CPU 206 はページマッピングテーブル 1002 等の管理情報の更新を行う。この際、s 1408 で一時的に記憶しておいた情報（各データの記憶された仮想アドレス空間上のアドレス、及び当該データの論理ページ番号の情報）を用いて、今回物理ページへの書き込み対象となった論理ページ（1101）に対応するグループ # 1102、オフセットアドレス 1103 を更新する。さらに仮想アドレス構成情報 1007 の次書き込み位置 1202 を更新

する（次書き込み位置 1 2 0 2 に、今回書き込まれたデータ量（セクタ数）を加算する）。

[0093] s 1 4 0 8 でのフラッシュチップ内物理ページへの書き込みは、上位装置からのライト要求に対して非同期で実施されること、また一定量のデータをまとめて物理ページへ書き込むことを想定している。これは、物理ページが書き込み単位であり、かつ論理ページが物理ページを跨がることを想定した場合、効率よく、かつ管理を楽にするために有効である。さらに、バッファが充分大きく、複数物理ブロック分のデータを蓄積できる場合、複数物理ブロック分のデータが蓄積されるまでバッファにデータを格納し、複数の物理ブロックに対して並列にデータを書き込むように制御し、アクセス性能をより向上させる方法もあり得る。この場合、仮想アドレス構成情報 1 0 0 7 において、グループ 1 2 0 1 内に登録される複数の物理ブロックを、それぞれ異なる FM チップ 2 0 9 から選択するとよい。たとえばグループ構築処理（後述）において、仮想ブロック # 1 2 0 3 が 0 番、1 番、2 番、3 番に対応付けられる物理ブロックがそれぞれ、図 2 の FM チップ 2 0 9 a、2 0 9 b、2 0 9 e、2 0 9 f 内の物理ブロックになるように、仮想アドレス構成情報 1 0 0 7 を設定し、4 物理ブロック分のデータをまとめて FM チップに格納する際には、FM チップ 2 0 9 a、2 0 9 b、2 0 9 e、2 0 9 f に対して、並列にライト要求を発行すると、単一の FM チップ 2 0 9 に対してライト要求を発行する場合に比べて、性能が向上する。

[0094] 一方で、必ずしも上で述べたように、複数ページをまとめて FM へ格納する方式を採る必要は無く、例えば 1 物理ページ単位でデータを FM に格納する方式を採ってもよい。その場合、FM に格納した物理ページに論理ページが跨がっている場合は、論理ページの一部領域が物理ページに格納されず、バッファに残されることがあるが、バッファに残された論理ページの論理ページ番号、論理ページ内オフセットアドレスの情報をバッファ上に残されたデータとともに管理しておき、他の論理ページの手書き込み処理（図 1 4 の処理）が行われるときに、あわせてバッファに残された論理ページの一部のデ

ータを物理ページに書き込むようにすればよい（もちろん、先にFMに格納された論理ページの前半部と、今回書き込まれる論理ページの後半部が、仮想アドレス空間上で連続するように書き込む必要がある）。

[0095] また、上で述べた方式は、いわゆるライトバック方式である。従って、停電等の障害が発生した場合に、揮発メモリ（バッファ等）上に格納されているライトデータを退避させることができるように、キャパシタやバッテリーなどによる、非常時の給電システムが必要である。

[0096] 最後にs 1 4 1 0で、CPU 2 0 6は今回の書き込みの結果、オープングループの空き領域が枯渇した（次書き込み位置1 2 0 2が、1グループ内物理ページ数を超過した）かどうかを判定し、枯渇した場合は、今回書き込みを行ったグループ（オープングループ）について、仮想アドレス構成情報1 0 0 7の次書き込み位置1 2 0 2の値をN/Aにすることで、今回書き込みを行ったグループを非オープングループに変更し、その後新規グループの構築を行って（s 1 4 1 1）、処理を終了する。s 1 4 1 1の処理については、後述する。

[0097] なお、上の説明では、ライトデータの格納位置が論理ページのページ境界に合致している場合について説明したが、上位装置1 0 2から、論理ページのページ境界に合致しないライト要求、たとえば1論理ページサイズに満たないサイズのデータのライト要求を受け付けたときには、s 1 4 0 2とs 1 4 0 3の間に、FMチップからライト対象データを含む1論理ページ（または複数論理ページ）分のデータをバッファに読み出し、バッファ上に読み出された1または複数論理ページ分のデータに対してライト対象データを上書きし、上書きされた1または複数論理ページ分のデータを、s 1 4 0 4以降の処理を実施することでFMへライトする。つまり、論理ページ境界に合致しないサイズのデータのライト要求を受け付けたときは、当該データを含む論理ページ全体のリードモディファイドライト（read modified write）を行って、論理ページ全体を書き換える。これは、ページマッピングテーブル1 0 0 2で、論理ページ単位で情報を管理しているため

、論理ページ内の一部領域のみの更新であっても、論理ページ全体を書き換える必要があるためである。ただし、この処理は、従来からあるフラッシュストレージでも同様であり、このような処理を行ったからといって、従来からあるフラッシュストレージに比べて性能が悪化するわけではない（同等である）。また論理ページサイズを、上位装置（たとえばホスト計算機103）で稼働するプログラム（OS、ファイルシステム、アプリケーションプログラム）が発行するI/O要求で指定されるデータサイズ（たとえば4KBや8KB）と同サイズにしておけば、論理ページのページ境界に合致していないライト要求が来ることは少ない。

[0098] 続いて図15を用いて、図14のs1411で行われる、新規グループ構築処理s1501のフローを説明する。

[0099] 新規グループ構築処理s1501は、オープングループの空きページがなくなり、次の書き込み先グループが必要となったときに実施される。従って、グループをオープングループにする処理と同義である。まずCPU206は、仮想アドレス構成情報1007の中から、物理ブロックの登録されていない未使用のグループ番号を取得する（s1502）。続いてCPU206は、消去済みの空きブロックリストを取得し、その中から1つのグループを構成するために必要となる数の物理ブロック番号を取得する（s1503, s1504）。最後にs1505で、CPU206はこれらの物理ブロック番号を仮想アドレス構成情報1007に登録し、次書き込み位置1202に値0を設定することにより今回作成した新規グループをオープングループとし、新規グループ構築処理を終了する。

[0100] 空きブロックリストは、消去済み物理ブロックの、物理ブロック番号が登録されているリストである。空きブロックリストは最低限、FMPKG112内に1つ用意されていれば良いが、用途に応じて適切な物理ブロックを選択できるように、複数の空きブロックリストを設け、各空きブロックリストに同種の属性・特性を有する空き物理ブロックの物理ブロック番号を登録しておくようにしても良い。

[0101] 例えば物理ブロックは、その消去回数などの要因によって、徐々に寿命が短くなる（劣化する）。そこで、各物理ブロックについて、消去回数に関する情報などに基づく、劣化度合いを示す指標値を算出して管理し（以下ではこの、消去回数に関する情報等に基づいて算出される指標値を「劣化度」と呼ぶ）、所定範囲の劣化度を有する空きブロックを集めた空きブロックリストを複数設け（たとえば劣化度が0から1までの値（小数）で定義される指標値で、値が小さいほど劣化度合いが少ない（寿命が長い）物理ブロックであるとする）、劣化度0～0.5の値を有する（寿命が長い）空き物理ブロックを管理する空きブロックリストと、劣化度0.5を超える値を有する（寿命が短い）空き物理ブロックを管理する空きブロックリストの2つを設けるなど）、新規グループ構築処理を実行する際に、一つの空きブロックリストを選択し、選択された1つの空きブロックリストから空き物理ブロックを複数選択して（これにより、選択される物理ブロックはいずれも所定範囲内の劣化度を有する物理ブロックになる）、新規構築されるグループに登録するようにしてもよい。空きブロックリストが複数ある場合の選択方法は、上位装置からあらかじめ劣化度の範囲を指定され、FMPKG112は上位装置から指定された劣化度の範囲の物理ブロックを管理する空きブロックリストを選択する方法、あるいはFMPKG112がI/O状態や稼働時間等の情報をもとに自動的に選択する方法などがあり得、いずれの方法をとってもよい。

[0102] また、先に述べたとおり、FMへのデータライト時の並列書き込みを可能にするために、FMチップ209やダイ（DIE）401などの物理的に異なる単位毎に空きブロックリストを設け（つまり、各空きブロックリストでは、同一のFMチップ209、あるいは同一ダイ401内の空き物理ブロックが管理される）、新規グループ構成時にそれぞれの空きブロックリストから順に空き物理ブロックを1つずつ取得してグループに登録するようにしても良い。これにより、1つのグループには、複数のFMチップ209あるいは複数のダイ401内の物理ブロックが登録されることになり、FMへのデ

ータライト時の並列書き込みが可能になる。

[0103] 続いて図16を用いて、本発明の実施例1に係るFMPKG112が実行する、グループ単位リクラメーション(RC)／リフレッシュ(RF)処理s1601のフローを説明する。

[0104] RCとRFは、それぞれ目的は異なるが、ブロック内の有効なデータを他のブロックにコピーしてコピー元のブロックを消去する処理である。本発明の実施例1に係るFMPKG112では、グループ単位でRC／RF処理を実施する。これは、グループ内の1つのブロックのみを消去すると、消去対象ブロックに隣接しているブロックに跨がった論理ページもコピーする必要があるためである。ただし、1つのブロックずつRC／RFを実施するようにしてもよい。

[0105] CPU206はまず、s1602で対象とするグループを取得する。RC／RFの対象を選択するのは、例えばRCに関して言えばコピー量になるべく減るように、仮想アドレス構成情報1007で管理されている無効ページ数1205を参照し、有効データの最も少ないグループ（無効ページ数1205がもっとも大きいグループ）から優先的にグループを選択する。RFでは、たとえば最終書き込み時刻1206を参照し、書き込んでからの時間経過が長いページを優先的に選択する。あるいは各グループについて、信頼性に関する情報を管理するようにし、信頼性に問題のあるグループを選ぶようにしても良い。なお、原則としてここではオープングループは選択しない。

[0106] s1603では、書き込み先のグループが選択される。書き込み先としては、オープングループが選択される。ただし、例えばウェアレベリングを考慮して、I/O頻度に適したグループを選択する、あるいは新規構築するようにしても良い。s1604でCPU206は、RC／RF対象グループ（s1602で選択されたグループ）内に有効データが存在するか判定を行い、存在する場合は、有効データの一部をs1605でバッファ204上にコピーする。バッファ204上にコピーされたデータについて、CPU206は、図14で説明したライト処理中のs1412（s1405～s1411ま

での処理)と同様の処理を行うことで、FMへと書き込む。ただし、s 1 4 1 2の処理を行う場合、s 1 4 0 6では、s 1 6 0 3で選択したグループがオープングループとして取得される。

[0107] CPU 2 0 6は、s 1 6 0 4、s 1 6 0 5、s 1 4 1 2の処理を、RC／RF対象グループ内の有効データが無くなるまで実施する。RC／RF対象グループ内の全ての有効データがコピーされ、RC／RF対象グループ内の有効データが無くなったら、CPU 2 0 6は、RC／RF対象グループに登録されている全物理ブロックについて消去処理を行い、仮想アドレス構成情報 1 0 0 7から、当該グループの情報を削除する(s 1 6 1 0、s 1 6 1 1)。消去されたブロックは空きブロックリストに追加される。

[0108] 以上が、本発明の実施例 1 に係る F M P K G 1 1 2 についての説明である。本発明の実施例 1 に係る F M P K G 1 1 2 の特徴は、論理ページサイズと物理ページサイズとが一致しない場合でも、管理情報の量が過剰に増加しない点にある。論理ページサイズと物理ページサイズとが等しくない場合、図 9 で説明したように、1 つの論理ページが複数の物理ページに跨がって格納される場合が発生する。この場合、1 論理ページについて、論理ページの前半部が格納される物理ページの情報、及び論理ページの後半部が格納される物理ページの情報という、2 つの情報を管理する必要がある。そのため、たとえばほぼすべての論理ページが 2 物理ページに跨って格納される場合、論理ページと物理ページとの対応関係を管理する情報の量は、1 論理ページが 1 物理ページ内に収まる場合に比べて 2 倍必要となる。

[0109] 一方、本発明の実施例 1 に係る F M P K G 1 1 2 の場合、論理ページと物理ページとの対応関係を管理する代わりに、論理ページを、仮想アドレス空間という 1 次元アドレス空間に対応付け、ページマッピングテーブルを用いて論理ページと仮想アドレス空間との対応関係を管理する。ページマッピングテーブルには、1 論理ページの位置を表現する情報は 1 エントリしか格納されないため、ページマッピングテーブルのサイズは、従来のフラッシュストレージ(論理ページと物理ページのサイズが等しい)における論物変換テ

ーブルと同様のサイズしか必要としないため、ほぼすべての論理ページが2物理ページに跨って格納される場合でも、管理情報のサイズが増加しない。

[0110] もちろん、実施例1にて説明したとおり、本発明の実施例1に係るFMPKG112ではページマッピングテーブルに加え、仮想アドレス空間と物理ブロックとの対応関係を管理する仮想アドレス構成情報を管理する必要があるが、仮想アドレス構成情報のサイズは、論物変換テーブルに比べて極めて小さい。なぜならば、論物変換テーブルには、ページ毎の情報が格納されるため、論物変換テーブルのサイズは、ページ数に比例する。一方仮想アドレス構成情報は、物理ブロック毎の情報が格納されるため、物理ブロック数に比例したサイズを有する。そして1物理ブロックには、たとえば128個あるいは256個のページを含んでいる（将来はもっと多くのページが格納される）ので、FMPKG112内の物理ブロック数は、物理ページ数の128分の1あるいは256分の1しかない。そのため、仮想アドレス構成情報のサイズは、論物変換テーブルの1/100以下と、極めて小さなサイズとなるため、仮想アドレス構成情報を持つことは、それほど管理情報を増加させることにはならない。

実施例 2

[0111] 次に、実施例2について説明する。以下の説明では、実施例1との違いを重点的に説明する。従って、実施例1と同じ構成及び処理については、説明を省略または簡略する場合がある。

[0112] 実施例2におけるストレージシステム及びFMPKGの基本的な構成は、実施例1に係るストレージシステム101、FMPKG112とほとんど同じであるため、図示を省略する。ただし、実施例2におけるFMPKGは、上位装置からのライトデータを圧縮する機能及び／または暗号化する機能を備える点が、実施例1に係るFMPKG112と相違する。この機能は、図2のCPU206あるいはASIC203で実施される態様でもよいし、あるいは圧縮または暗号化用の専用ハードウェアが設けられる態様もありえる。また、実施例2に係るFMPKGでは、ライトデータを変換（たとえば圧

縮、暗号化)して格納するが、格納されたデータが読み出される場合には、データが逆変換(圧縮データの伸長、あるいは暗号化されたデータの復号化)されて上位装置に渡されるため、ストレージシステムやホスト計算機からは、実施例1に係るFMPKG112と同じ装置としてみえる。そのため、実施例2に係るストレージシステムの構成は、実施例1におけるものと同じである。

[0113] 実施例1においては、論理ページが物理ページを跨がるという点以外には限定をおいていなかったが、本実施例は、特に論理ページの格納時のデータサイズが変動する場合の処理方式について述べる。これは、データ圧縮や暗号化などのデータ変換が加わるためである。この場合、各論理ページのフラッシュでの格納状況はより複雑となる。

[0114] 図17は、実施例2におけるページマッピングテーブル1700の内容を示す。

[0115] 実施例1と同じく、実施例2に係るFMPKGでは、上位装置に対して一次元の論理記憶空間(図6の600相当物)を提供し、この一次元の論理空間を先頭から所定サイズ(たとえば16セクタ)ごとに分割し、分割された各領域を論理ページとして管理している。実施例1と異なる点は、圧縮・暗号化などのデータ変換を行う点だが、データ変換はこの論理ページごとに行われる。つまり1論理ページ未満のサイズのデータを変換(たとえば圧縮)したり、論理ページ境界を跨いで存在するデータを変換(圧縮)する、ということは行わない。

[0116] ページマッピングテーブル1700が、実施例1のページマッピングテーブル1002と異なるのは、ページマッピングテーブル1700内の各エントリについて、データの長さを表すレングス1705の項目が加えられている点である。各論理ページはデータ変換によりサイズが変動するため、変換後の論理ページのデータのサイズをこのエントリで管理する。レングス1705の項目に格納される数値の単位は、例えばセクタ(512バイトあるいは520バイト)であるが、その他の単位(バイトなど)を用いてもよい。

つまり、ページマッピングテーブル 1700 では、変換（たとえば圧縮）後の論理ページのデータが対応付けられる仮想アドレス空間上の範囲の情報を、論理ページ毎に管理している。

[0117] また、実施例 2 に係る FMPKG でも実施例 1 と同様に、仮想アドレス構成情報を管理するが、その内容は実施例 1 に係る仮想アドレス構成情報 1007 と同じであるため、ここでは説明しない。

[0118] 図 18 は、実施例 2 に係る FMPKG が実行する、論理アドレスから物理アドレスへの変換処理（正引き処理）のフローを示す。実施例 1 における正引き処理と同じく、上位装置からリード要求を受け付けた場合に行われる論理アドレスの物理ページ番号への変換処理で、かつリード要求で指定されるアクセス対象データの先頭位置は論理ページ境界と一致しており、アクセス対象データのサイズが論理ページサイズと等しい場合、つまり 1 論理ページ分の領域の物理位置の算出処理について説明する。

[0119] 実施例 1 の正引きフロー（図 13）と異なるのは、図 13 の s1304 の処理（グループ番号とグループ内オフセットアドレスの取得処理）が s1802 に代わっている点、及び図 13 の s1307 の処理が、s1803 に代わっている点である。s1802 では、FMPKG の CPU は、ページマッピングテーブル 1700 の中から、該当する論理ページのグループ番号 1703、グループ内オフセットアドレス 1704 に加え、レンジ 1705 を取得する。グループ番号とオフセットアドレスにより、当該論理ページが格納されている仮想アドレス空間上の先頭アドレスは取得できるため、s1305～s1306 の処理は実施例 2 でも実施例 1 と同様に行われる。

[0120] そして s1803 では、CPU はアクセス対象データの先頭物理ページ番号及び終端物理ページ番号を算出する。先頭物理ページ番号の算出処理は、実施例 1 のものと同じであるが、アクセス対象データの終端物理ページ番号の算出処理の際には、1 論理ページのサイズを用いることに代えて、レンジ 1705 の情報を用いて終端物理ページ番号を算出する点が、実施例 1 のものと相違する。その他の処理は実施例 1 で説明した処理と同様である。

[0121] 図19は、実施例2におけるFMPKGが上位装置からライト要求を受信した時に実行する処理の流れを示している。図14の説明と同様、ここでも、ライトデータの格納位置が論理ページのページ境界に合致しているケースを中心に説明する。

[0122] 実施例1のライト処理（図14の処理）と異なるのは、まずs1402～s1404の処理によってバッファに格納された、上位装置から受領したデータに対して変換処理が施される（s1902）点である（これにより、バッファに格納されていた変換前のデータは、変換後データに置き換わる）。変換処理は先に述べたとおり、FMPKGのCPUが行うようにしても良いし、あるいはCPUが専用のハードウェアに実施させてもよい。またs1902の後に行われるs1903で、CPUは変換後のサイズでバッファ上のデータ量をチェックし、バッファ上のデータ量（変換後のデータのサイズ）が所定サイズ以上であると判定されるとフラッシュへのライトを開始する、そして管理情報更新（s1409）の際には、レングス1705の情報更新を行う点が、実施例1の図14の処理と異なる。それ以外の点は実施例1と同様である。

[0123] なお、実施例2の変形として、FMPKGが上位装置から受信したライトデータをFMに格納する時にはデータ変換を行わず（図19の処理の実行時にs1902の処理を行わない）、RC/RF処理の際に、圧縮などのデータ変換を行う方法もある。その場合のRC/RF処理の流れを、図20に示す。

[0124] 図20の処理が実施例1の図16の処理と異なる点は、データをバッファにコピーする際（実施例1の図16のs1605）に、図20の処理では、CPU（あるいは専用ハードウェア）がデータ変換を実施し、変換されたデータをバッファに格納する点である（s2002）。それ以外の点は、実施例1のRC/RF処理と同じである。

[0125] 以上で、実施例2に係るFMPKGについての説明を終了する。実施例2に係るFMPKGで管理する管理情報は、実施例1におけるものと同様に、

ページマッピングテーブル 1700 及び仮想アドレス構成情報である。ページマッピングテーブル 1700 は、実施例 1 におけるものと比べると、データの長さ（レングス）の情報を保持する必要があるものの、実施例 1 の場合と同様に、ページマッピングテーブルには、1 論理ページの位置を表現する情報は 1 エントリ分しか格納する必要がないため、管理情報の量はそれほど増加しない。

[0126] また、実施例 1 に係る FMPKG では、全ての論理ページサイズが同一であったため、管理情報サイズを小さくするための一つの方法として、論理ページサイズを物理ページサイズと同一にするという方法を採用することが可能である。一方、実施例 2 に係る FMPKG のように、格納データに対してデータ圧縮や暗号化などのデータ変換を行う場合、データ内容に依存してデータサイズが変動する。そのため、全ての論理ページサイズを物理ページサイズと同一にするという方法を採用することができないので、データ格納効率を上げたい場合、1 つの論理ページが複数の物理ページに跨って格納されることを許容する構成にせざるを得ない。上で説明した実施例 2 の方法を用いることで、データ変換によって格納されるデータのサイズが変動する場合であっても、少ない管理情報で、論理ページと物理格納位置の関係を管理することができる。

実施例 3

[0127] 次に、実施例 3 を説明する。以下の説明では、実施例 1 との違いを重点的に説明し、実施例 1 と同じ構成及び処理については、説明を省略または簡略する場合がある。

[0128] 実施例 1、2 においては、個々の物理ページのサイズは固定であったが、本実施例では、物理ページサイズが一定でない場合を述べる。ただし完全に自由なサイズを採用のではなく、1 つの FMPKG 内に、数パタンの物理ページが存在する。これは、例えば物理ページ毎に最適な ECC サイズを適用するような方式が考えられる。

[0129] 図 21 は、物理ページサイズが異なるブロックの存在するデバイスで仮想

アドレスグループ（及び仮想アドレス空間）を構成する例を示したものである。

[0130] デバイス中に、物理ページサイズ特性の異なる２種類以上の物理ブロックがあり、それぞれの種類の物理ブロックを用いて仮想アドレスグループ（グループ）を構成する。ここで「物理ページサイズ特性」とは、１物理ブロック内の物理ページサイズがすべて同じであるという特性、あるいは物理ページサイズの異なる物理ページが同一物理ブロック内に複数種類存在し、それらが特定のパタンで分布している（たとえば各物理ブロックについて、ページ番号０～ n 間の物理ページの物理ページサイズは７ＫＢで、ページ番号（ $n+1$ ）～（ $n+m$ ）の物理ページは、物理ページサイズが８ＫＢである）等のように、物理ブロック内の各物理ページのページサイズの傾向・分布のことを意味する。

[0131] 図２１の例では、ある物理ページサイズ特性を持つ物理ブロックの種別（`type`）を`type A`とし、また`type A`とは異なる物理ページサイズ特性を持つ物理ブロックの種別を`type B`とする。実施例３に係るＦＭＰＫＧでは、`type A`の種別の物理ブロック（２１０２）のみが登録された仮想アドレスグループ２１０１_a、及び`type B`の種別の物理ブロック（２１０３）のみが登録された仮想アドレスグループ２１０１_bを構成している。もちろん、これ以外の仮想アドレスグループを構成してもよい。

[0132] 同一種別のブロックでグループを構成するのは、グループでブロックの特性をそろえることで寿命管理を容易にできる点、またグループ内のオフセットアドレスから当該グループの仮想アドレス空間上に対応付けられている物理ブロック・物理ページを特定する場合、個別のブロックの情報を参照することなく実施できる点に利点がある。

[0133] 図２２は、本実施例における仮想アドレス構成情報２２０１を示す。

[0134] 実施例３における仮想アドレス構成情報２２０１の構成要素は、グループ２２０２～最終書き込み時刻２２０７までは実施例１のものと同一である。

ただし仮想アドレス構成情報 2201 には、グループ毎に、そのグループに対応づけてあるブロックの種別 2208 の情報が保持される点が、実施例 1 の仮想アドレス構成情報 1007 と異なる。FMPKG は種別 2208 に格納されている情報を元にして、グループ内の各物理ページの物理ページサイズを識別する。

[0135] 図 23 は、本発明の実施例 3 に係る FMPKG が実行する、論理アドレスから物理アドレスへの変換処理（正引き処理）のフローを示す。

[0136] 実施例 1 で説明した正引き処理（図 13）と同様に、（FMPKG の）CPU はアクセス対象論理ページ番号を取得し、取得した論理ページ番号を元にページマッピングテーブルを参照し、アクセス対象論理ページ番号に対応づけられたグループ番号とグループ内のオフセットアドレスを取得する（s1302～s1304）。続いて CPU は、仮想アドレス構成情報 2201 を参照し、種別 2206 を取得する（s2302）。この種別 2206 により、CPU は仮想アドレスグループを構成する各物理ブロック内の物理ページ構成（各物理ページのサイズ）を識別し、続く s2303 で、グループ内のオフセットアドレスと物理ページ構成を元に、仮想ブロック番号と仮想ブロック内のオフセットアドレスを算出する。最後に CPU は、s2304 で物理アドレスを取得する。s2304 は、実施例 1（図 13）の s1306、s1307 に相当する処理である。実施例 1 の正引き処理の場合、仮想ブロック内オフセットアドレスを物理ページサイズで除算するだけで物理アドレスを算出可能であったが、実施例 3 の場合、物理ページ構成によっては（たとえば物理ブロック内に異なるサイズの物理ページが存在する構成等）、単純に除算するだけでは物理アドレスは算出できない点が、実施例 1 における処理と異なる（ただし物理ブロック内各物理ページのページサイズの分布パターンが分かっているならば、その情報を用いて物理アドレスを算出することができる）。

[0137] 図 24 は、実施例 3 における新規グループ構築処理のフローを示す。実施例 1 との差異は、本発明の実施例 3 に係る FMPKG は、種別ごとの空きブ

ロックリストを管理していること、そしてブロック取得の時に、CPUは取得対象ブロックの種別を決定し、決定された種別の空きブロックリストを取得する点である（s 2402、s 2403）。たとえば、フラッシュメモリ制御装置201は、常時I/O頻度を監視しており、新規グループ構築処理の際にI/O頻度が高かった場合には比較的信頼性が高い物理ブロックの種別である、種別Aに該当する物理ブロックからグループを構築し、I/O頻度が低い場合には、それほど信頼性が高い物理ブロックではない、種別Bに該当する物理ブロックからグループを構築する、等の選択を行う。これにより、新規グループを構成するブロックは同一種別で構成されることになる。この処理を効率化するために、各ブロックは種別毎に分けられたキューなどで管理されていても良い。また、要求する種別は、例えば種別によって信頼性やデータ格納量に差があるのであれば、格納データの種類に応じて最適なものを選ぶ方式であっても良い。

[0138] 以上が実施例3に係るFMPKGの説明である。たとえば現在研究が進んでいる3次元積層技術を用いたチップ（3D積層チップ）では、同じFMチップ内に特性の異なるページが存在し、それぞれのページで異なるサイズのECCを採用する必要性が出る可能性が高い。そのため、実施例3に係る構成は、3D積層チップのような特性のFMチップを使用したFMPKGを用いる場合に有用である。

符号の説明

[0139] 101：ストレージシステム
102：ストレージコントローラ
103：ホスト計算機
104：管理端末
105：SAN
106：保守インターフェイス
107：ホストインターフェイス
108：CPU

109 : メモリ
110 : ドライブインターフェイス
112 : フラッシュメモリストレージ装置 (FMPKG)
201 : フラッシュメモリ制御装置
202 : ストレージインターフェイス
203 : ASIC
204 : バッファ
205 : バッテリ
206 : CPU
207 : メインメモリ
208 : フラッシュメモリインターフェイス
209 : フラッシュメモリチップ (FMチップ)
401 : ダイ (DIE)
402 : ページバッファ
403 : 物理ブロック
404 : 物理ページ
501 : Code Word (CW)
502 : データ
503 : ECC
1001 : 論理ページ
1002 : ページマッピングテーブル
1003 : 仮想アドレス空間
1004 : 仮想アドレスグループ
1006 : 仮想ブロック
1007 : 仮想アドレス構成情報
1008 : 物理ブロック
1009 : 物理ページ

請求の範囲

- [請求項1] コントローラと複数のフラッシュメモリチップを備えるフラッシュメモリストレージ装置であって、
- 前記コントローラは少なくとも、プロセッサ、揮発性メモリで構成されるバッファ、上位装置に接続するためのストレージインターフェイス、前記複数のフラッシュメモリを前記コントローラに接続するためのフラッシュメモリインターフェイスを備え、
- 前記プロセッサは、前記上位装置に対して提供する論理記憶空間を、所定サイズの領域である論理ページ単位に分割して管理するとともに、前記複数のフラッシュメモリチップの有する複数の物理ブロックが対応付けられた線形なアドレス空間である仮想アドレス空間を管理しており、
- 前記プロセッサはまた、前記論理ページと前記仮想アドレス空間上のアドレスとの対応関係が記録されたページマッピングテーブルと、前記仮想アドレス空間上の領域と前記物理ブロックとの対応関係が記録された仮想アドレス構成情報とを管理することを特徴とする、フラッシュメモリストレージ装置。
- [請求項2] 前記各物理ブロックは、前記フラッシュメモリチップの最小アクセス単位の領域である物理ページを複数有し、
- 前記論理ページのサイズは、前記物理ページのサイズと等しくないことを特徴とする、
- 請求項1に記載のフラッシュメモリストレージ装置。
- [請求項3] 前記プロセッサは、前記上位装置から前記論理ページに対するライト要求を受け付けると、前記仮想アドレス空間上の未使用領域を選択し、前記選択された未使用領域の前記仮想空間上アドレスと前記論理ページとの対応関係をページマッピングテーブルに記録することを特徴とする、請求項1に記載のフラッシュメモリストレージ装置。
- [請求項4] 前記プロセッサは、前記上位装置から前記論理ページに対するリー

ド要求を受け付けると、前記ページマッピングテーブルを参照して前記論理ページに対応する前記仮想空間上アドレスを特定し、前記仮想アドレス構成情報を参照することで、前記仮想アドレス空間上アドレスに対応付けられている物理ブロック内の1または複数の物理ページを特定することを特徴とする、請求項2に記載のフラッシュメモリストレージ装置。

[請求項5] 前記プロセッサは、前記上位装置からのライトデータを受け付けると、前記ライトデータを前記バッファに蓄積し、

前記バッファに1物理ページのサイズ以上のデータが蓄積された時、前記バッファに蓄積されたデータを前記物理ページのサイズ毎に、前記物理ページに格納することを特徴とする、請求項1に記載のフラッシュメモリストレージ装置。

[請求項6] 前記プロセッサは、前記仮想アドレス空間にまだ対応付けられていない物理ブロックを管理する空きブロックリストを複数有し、前記各空きブロックリストには、同種の属性を有する物理ブロックの物理ブロック番号が管理されていることを特徴とする、請求項1に記載のフラッシュメモリストレージ装置。

[請求項7] 前記属性は、物理ブロックの消去回数に関する情報に基づいて算出される指標値である劣化度であって、

第1の範囲の劣化度を有する空き物理ブロックの物理ブロック番号は、前記複数の空きブロックリストの中の第1の空きブロックリストによって管理され、第2の範囲の劣化度を有する空き物理ブロックの物理ブロック番号は、前記複数の空きブロックリストの中の第1の空きブロックリストによって管理されており、

前記プロセッサが、前記仮想アドレス空間上の所定範囲の領域に前記物理ブロックを対応付けるとき、前記第1、第2の空きブロックリストのうち1つの空きブロックリストを選択して、前記選択された空きブロックリストによって管理されている物理ブロックを前記仮想ア

ドレス空間上の所定範囲の領域に対応付けることを特徴とする、
請求項 6 に記載のフラッシュメモリストレージ装置。

[請求項8] 前記属性は、物理ブロックの属するフラッシュメモリチップを特定する属性情報であって、同一の前記属性情報を有する物理ブロックは、同一の前記空きブロックリストによって管理され、

前記プロセッサが、前記仮想アドレス空間上の所定範囲の領域に前記物理ブロックを対応付けるとき、前記複数の空きブロックリストの各々から、空き物理ブロックを 1 つ選択し、前記選択された物理ブロックを前記仮想アドレス空間上の所定範囲の領域に対応付け、

前記プロセッサが、前記仮想アドレス空間上の所定範囲の領域に対応付けられた複数の物理ブロックに対してデータを格納する際、前記複数の物理ブロックに対して並列に書き込み要求を発行することを特徴とする、

請求項 6 に記載のフラッシュメモリストレージ装置。

[請求項9] 前記コントローラは、前記上位装置からのライトデータを前記論理ページごとに変換して格納するよう構成されており、

前記コントローラはまた、前記ページマッピングテーブルに、前記変換された論理ページのデータが対応付けられる前記仮想アドレス空間上の範囲の情報を前記論理ページごとに管理することを特徴とする、

請求項 1 に記載のフラッシュメモリストレージ装置。

[請求項10] 前記コントローラは、前記上位装置からライトデータを受信すると、前記ライトデータを変換してバッファに格納した後、前記バッファに格納された変換済データを前記フラッシュメモリチップに書き込むことを特徴とする、

請求項 9 に記載のフラッシュメモリストレージ装置。

[請求項11] 前記コントローラは、前記上位装置からライトデータを受信すると、前記ライトデータを変換せずバッファに格納した後、前記バッファ

に格納された前記データを前記フラッシュメモリチップに書き込み、前記フラッシュメモリチップに書き込まれたデータを、別のフラッシュメモリチップへコピーするときに、前記フラッシュメモリチップに書き込まれたデータを変換し、前記変換されたデータを前記別のフラッシュメモリチップへコピーすることを特徴とする、請求項9に記載のフラッシュメモリストレージ装置。

[請求項12] 前記フラッシュメモリストレージ装置は、物理ページサイズ特性の異なる2種類以上の物理ブロックを有し、

前記プロセッサは、複数種類の前記仮想アドレス空間を管理しており、

前記それぞれの仮想アドレス空間には、同一の物理ページサイズ特性のみが対応付けられていることを特徴とする、

請求項1に記載のフラッシュメモリストレージ装置。

[請求項13] ストレージコントローラと複数のフラッシュメモリストレージ装置を有するストレージシステムであって、

前記フラッシュメモリストレージ装置は、コントローラと複数のフラッシュメモリチップを備え、

前記コントローラは、前記ストレージコントローラに対して提供する論理記憶空間を、所定サイズの領域である論理ページ単位に分割して管理するとともに、前記複数のフラッシュメモリチップの有する複数の物理ブロックが対応付けられた線形なアドレス空間である仮想アドレス空間を管理しており、

前記コントローラはまた、前記論理ページと前記仮想アドレス空間上のアドレスとの対応関係を管理するページマッピングテーブルと、前記仮想アドレス空間上の領域と前記物理ブロックとの対応関係を管理する仮想アドレス構成情報とを有することを特徴とする、ストレージシステム。

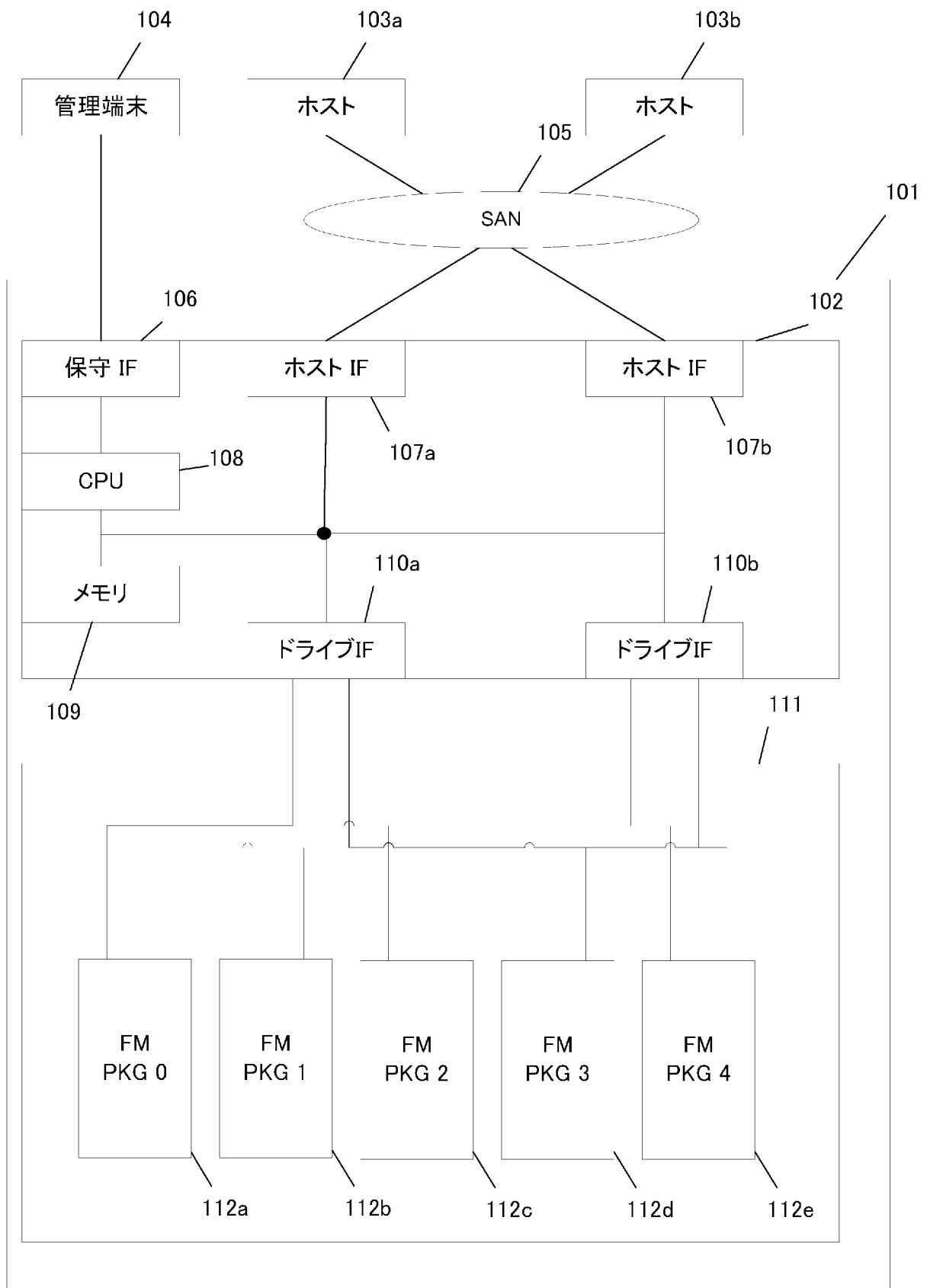
[請求項14] 前記コントローラは、前記ストレージコントローラから前記論理ペ

ージに対するライト要求を受け付けると、前記仮想アドレス空間上の未使用領域を選択し、前記選択された未使用領域の前記仮想空間上アドレスと前記論理ページとの対応関係をページマッピングテーブルに記録することを特徴とする、請求項 13 に記載のストレージシステム。

[請求項15] 前記コントローラは、前記ストレージコントローラから前記論理ページに対するリード要求を受け付けると、前記ページマッピングテーブルを参照して前記論理ページに対応する前記仮想空間上アドレスを特定し、前記仮想アドレス構成情報を参照することで、前記仮想アドレス空間上アドレスに対応付けられている物理ブロック内の 1 または複数の物理ページを特定することを特徴とする、請求項 13 に記載のストレージシステム。

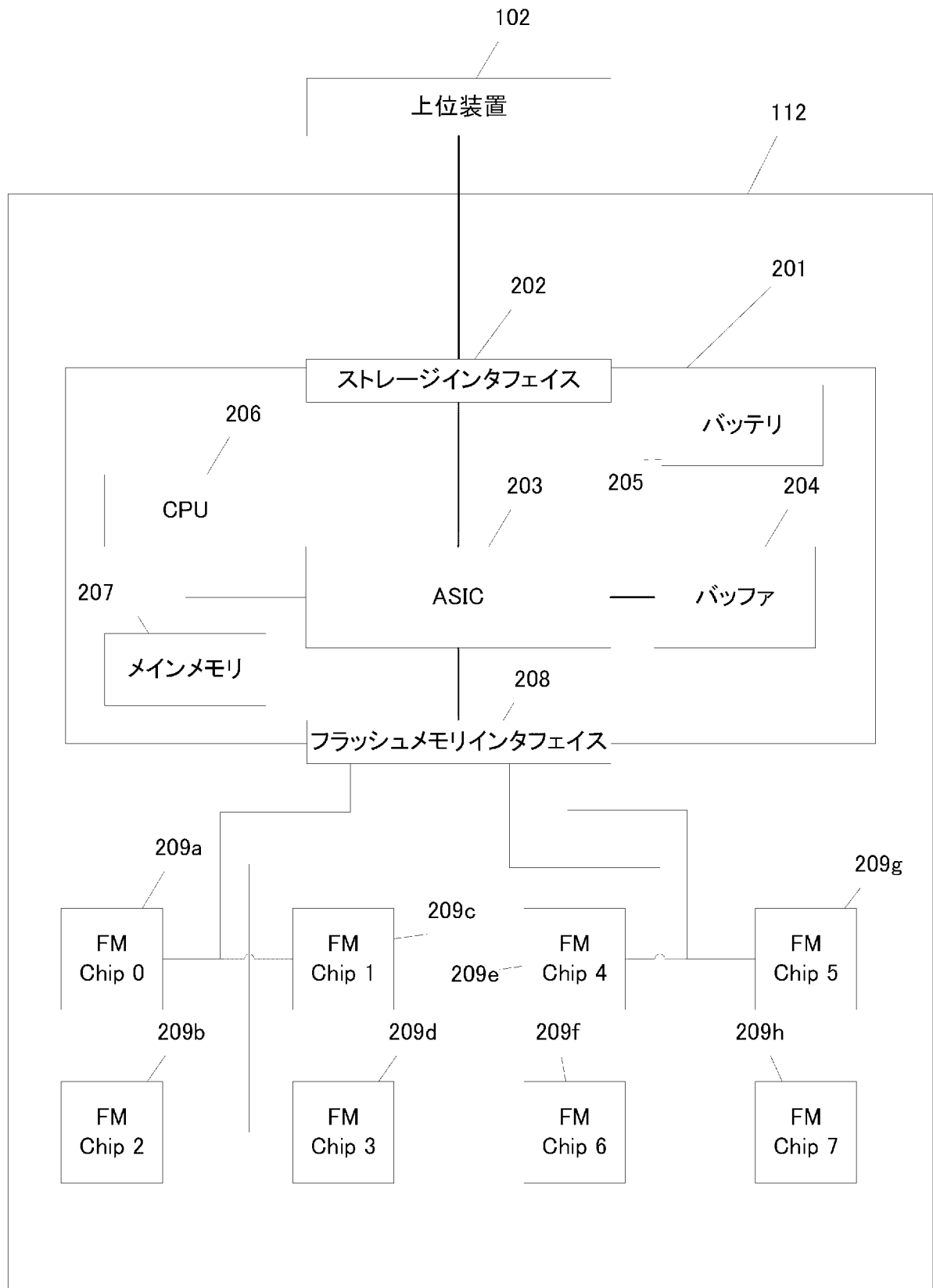
[図1]

図1



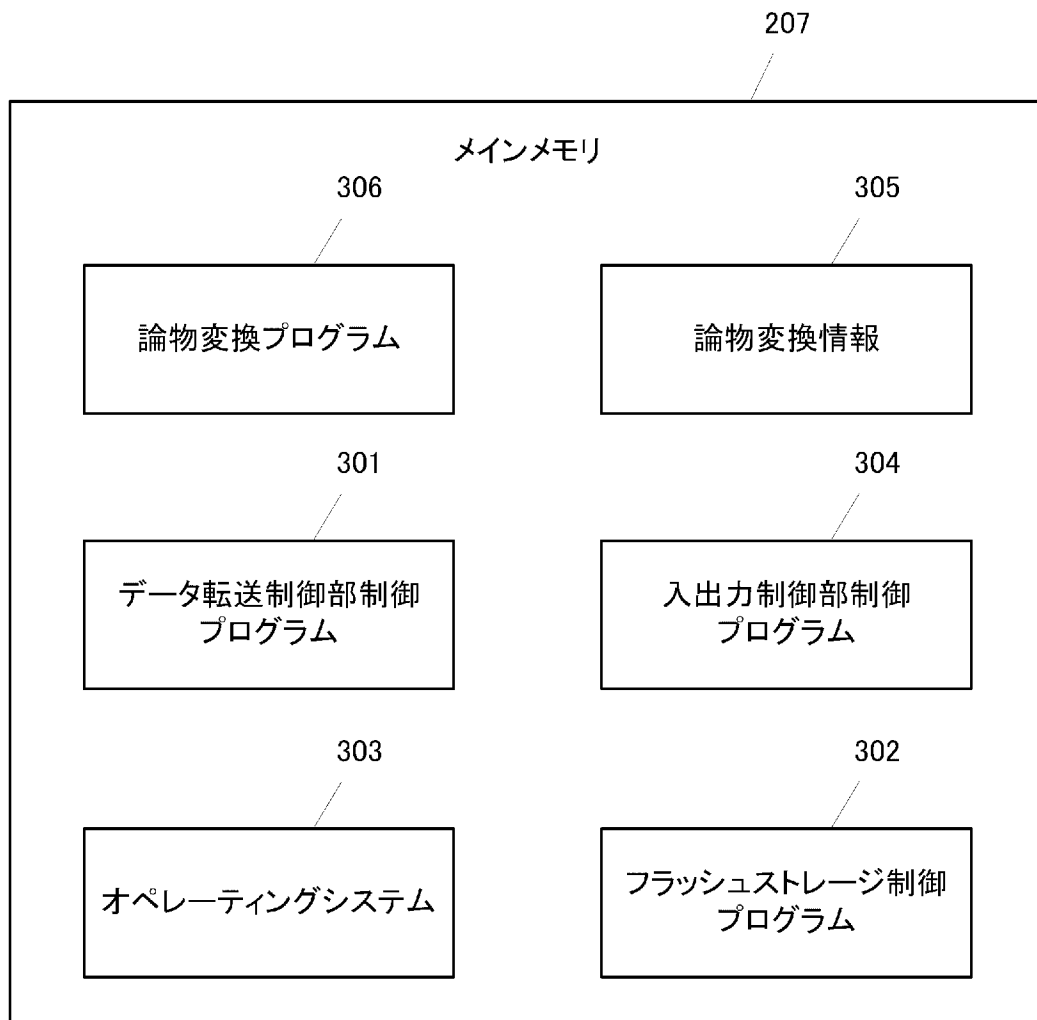
[図2]

図2



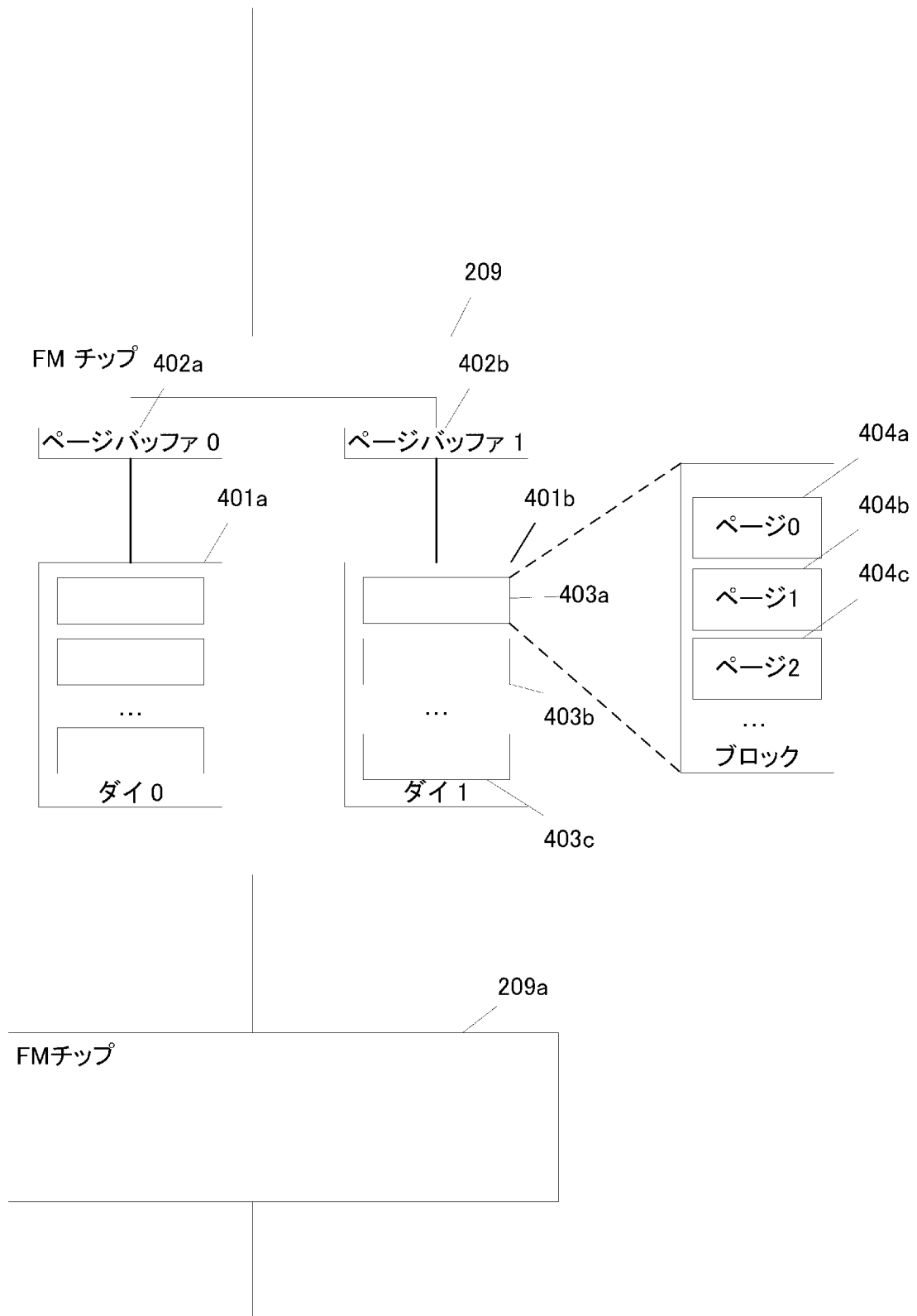
[図3]

図3



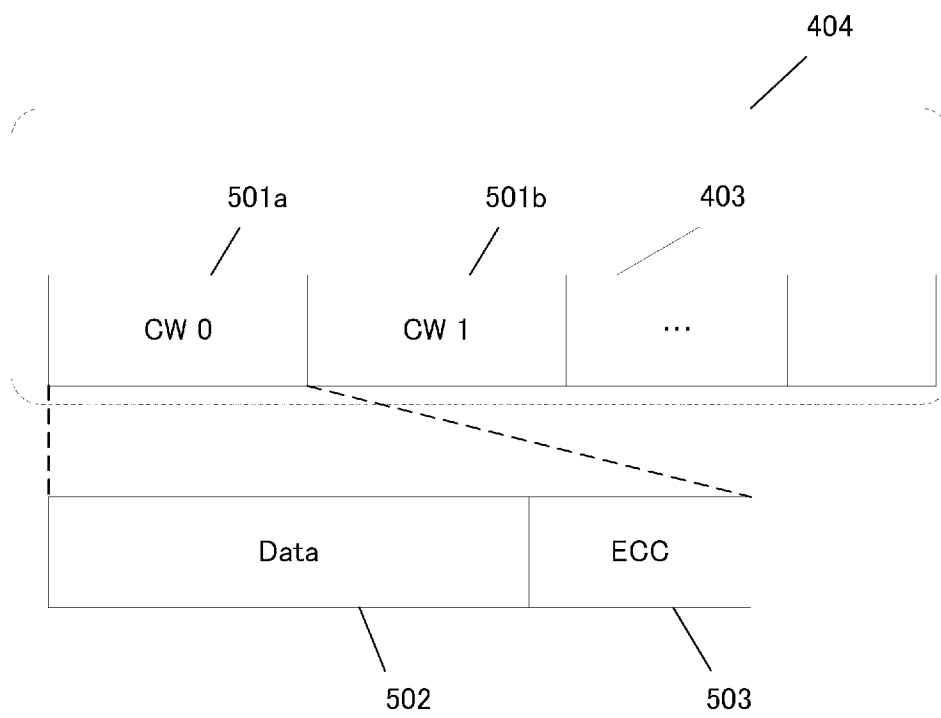
[図4]

図4



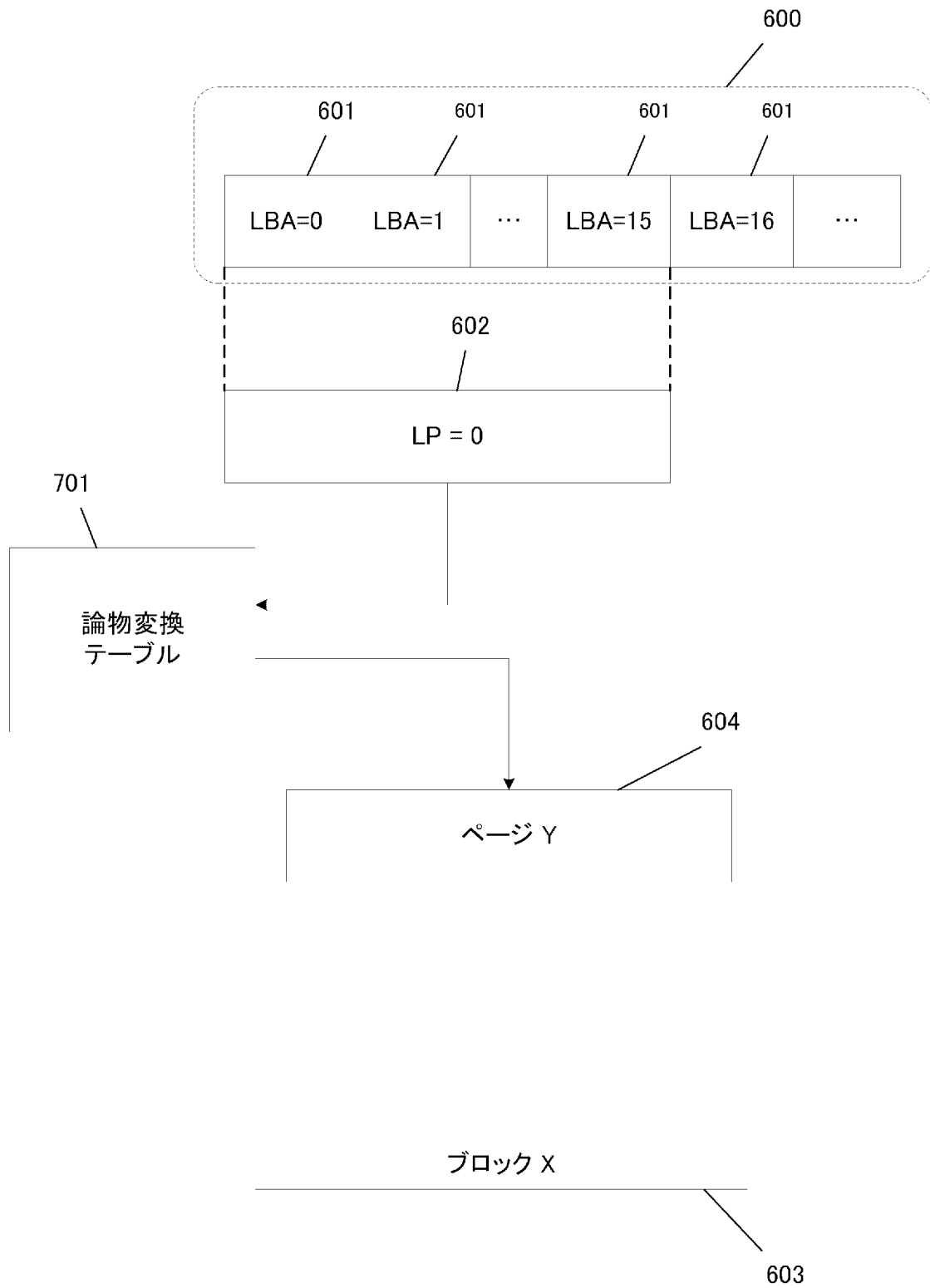
[図5]

図5



[図6]

図6



[図7]

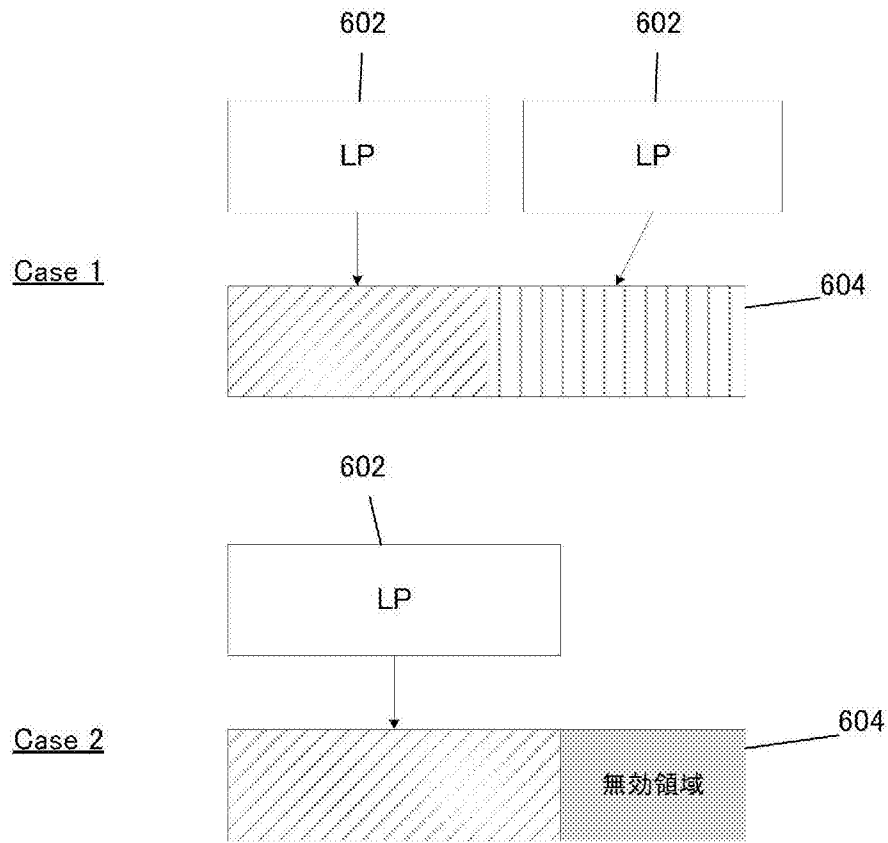
図7

701
↓
論物変換テーブル

702 論理ページ	703 格納アドレス
0	0x00A
1	0x00B
	...

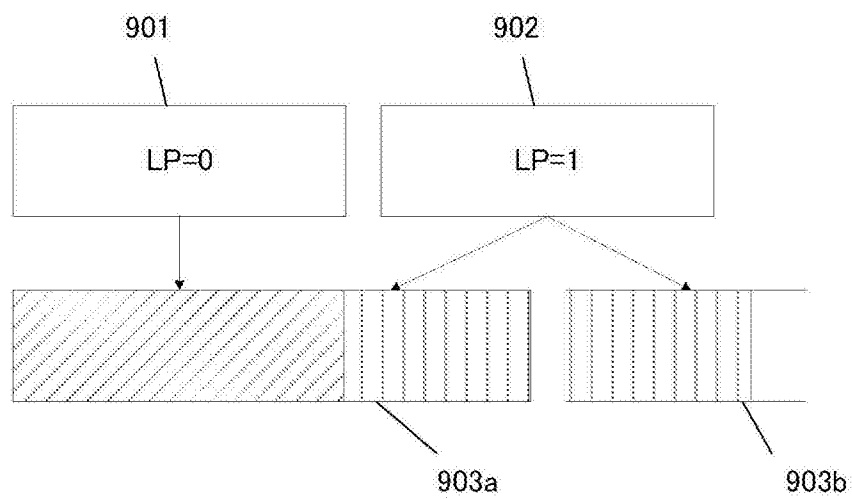
[図8]

図8



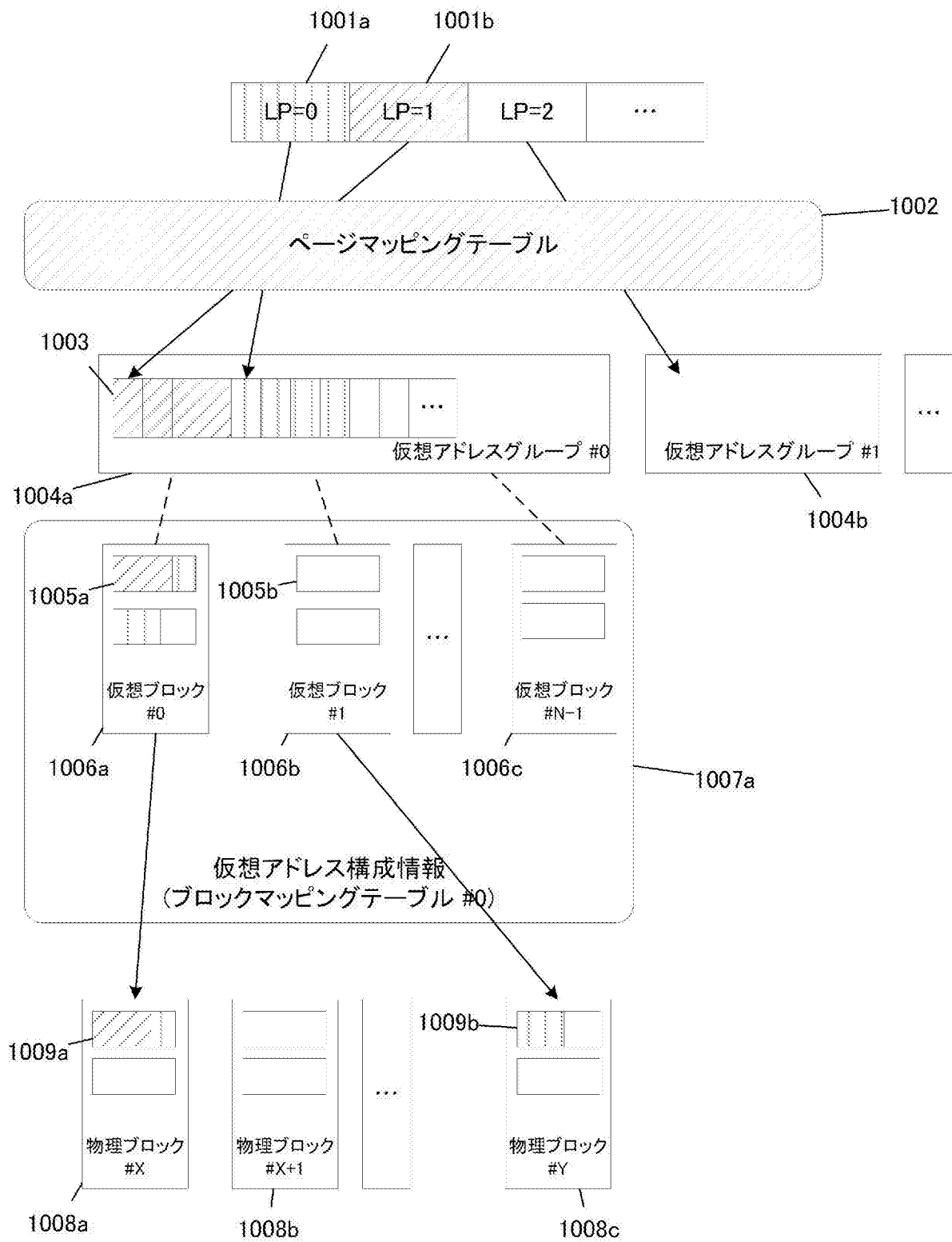
[図9]

図9



[図10]

図10



[図11]

図11

1002

ページマッピングテーブル		
1101 論理ページ	1102 グループ#	1103 オフセットアドレス
0	0	0
1	1	10
...		

[図12]

図12

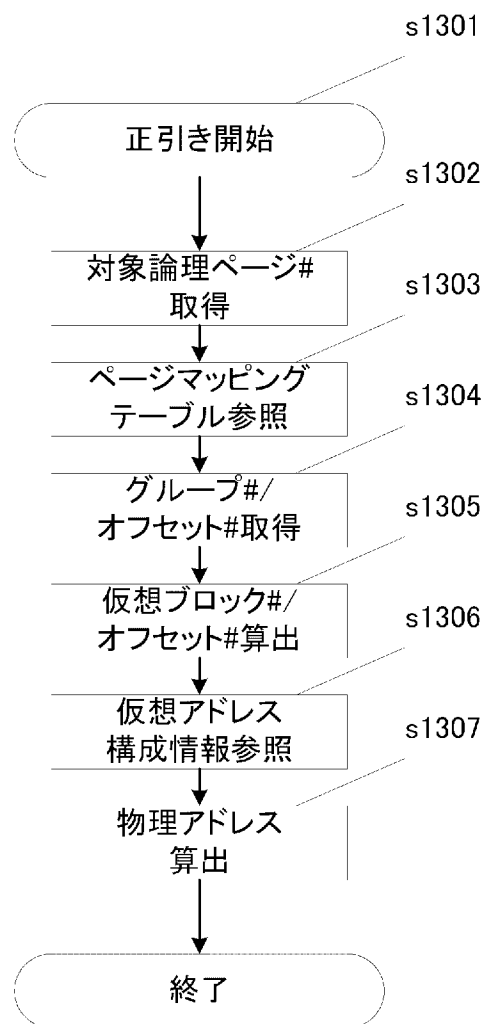
1201 グループ	1202 次書き込み位置	1203 仮想 ブロック#	1204 物理 ブロック#	1205 無効 ページ数	1206 最終書込時刻
0	10	0	0x00A	0	0:00
		1	0x00B		
		N-1	0x00F		
1	N/A	0	0x10A	22	...
		1	0x10B		
		N-1	0x10F		
.....					

仮想アドレス構成情報

1007

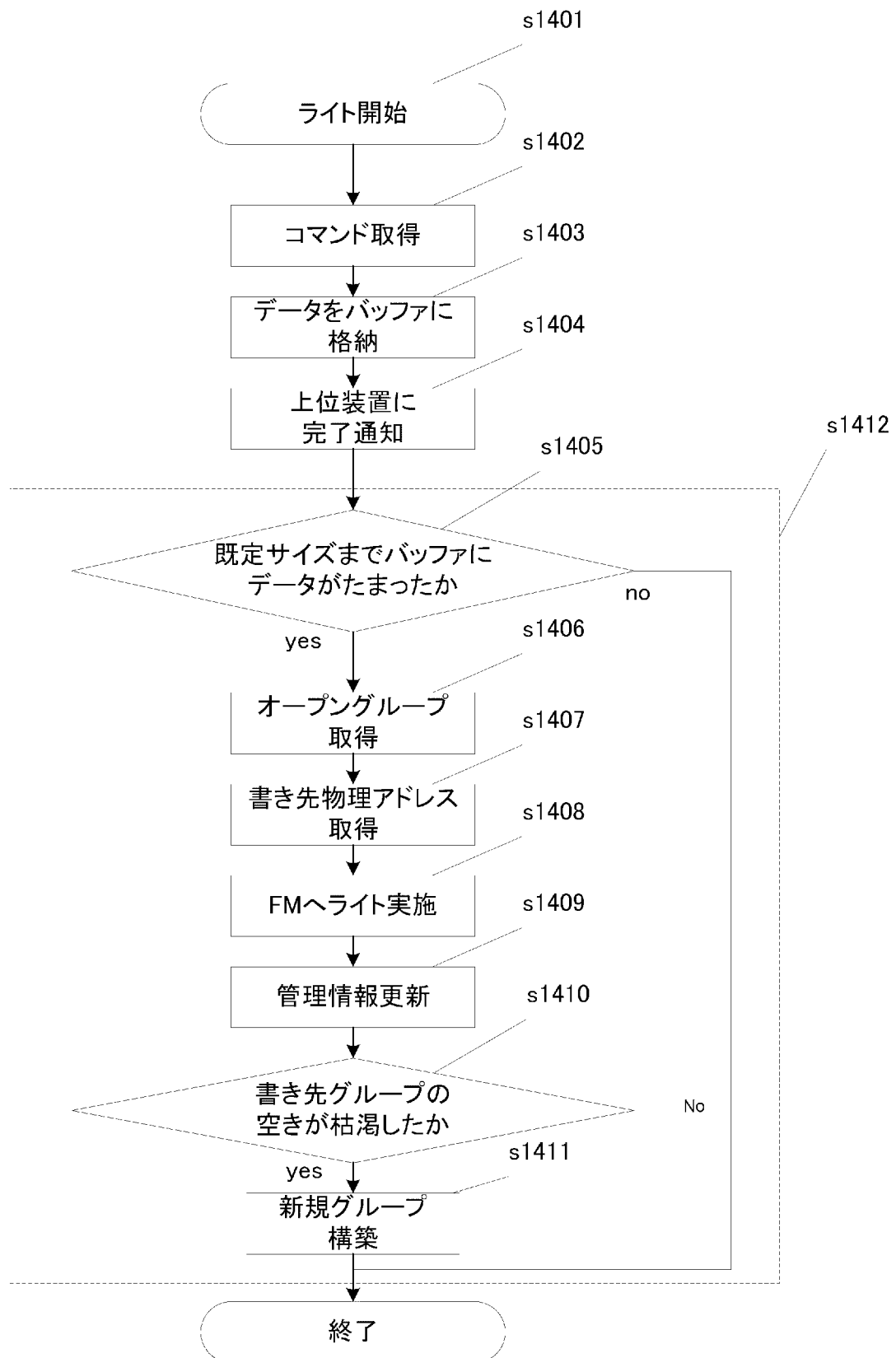
[図13]

図13



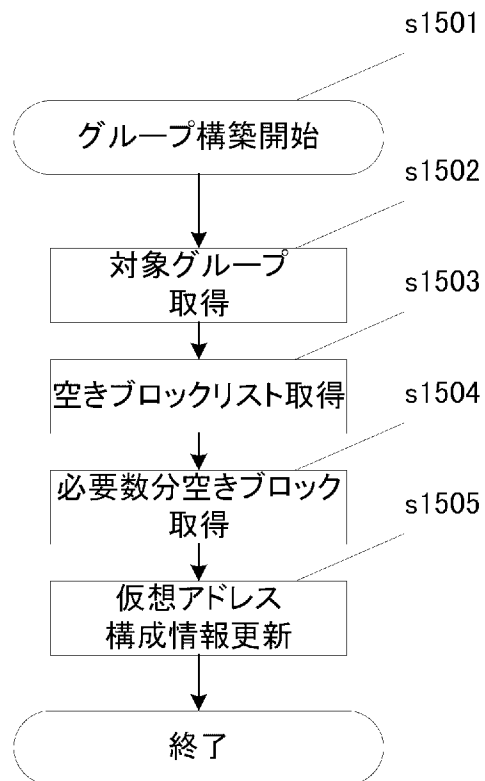
[図14]

図14



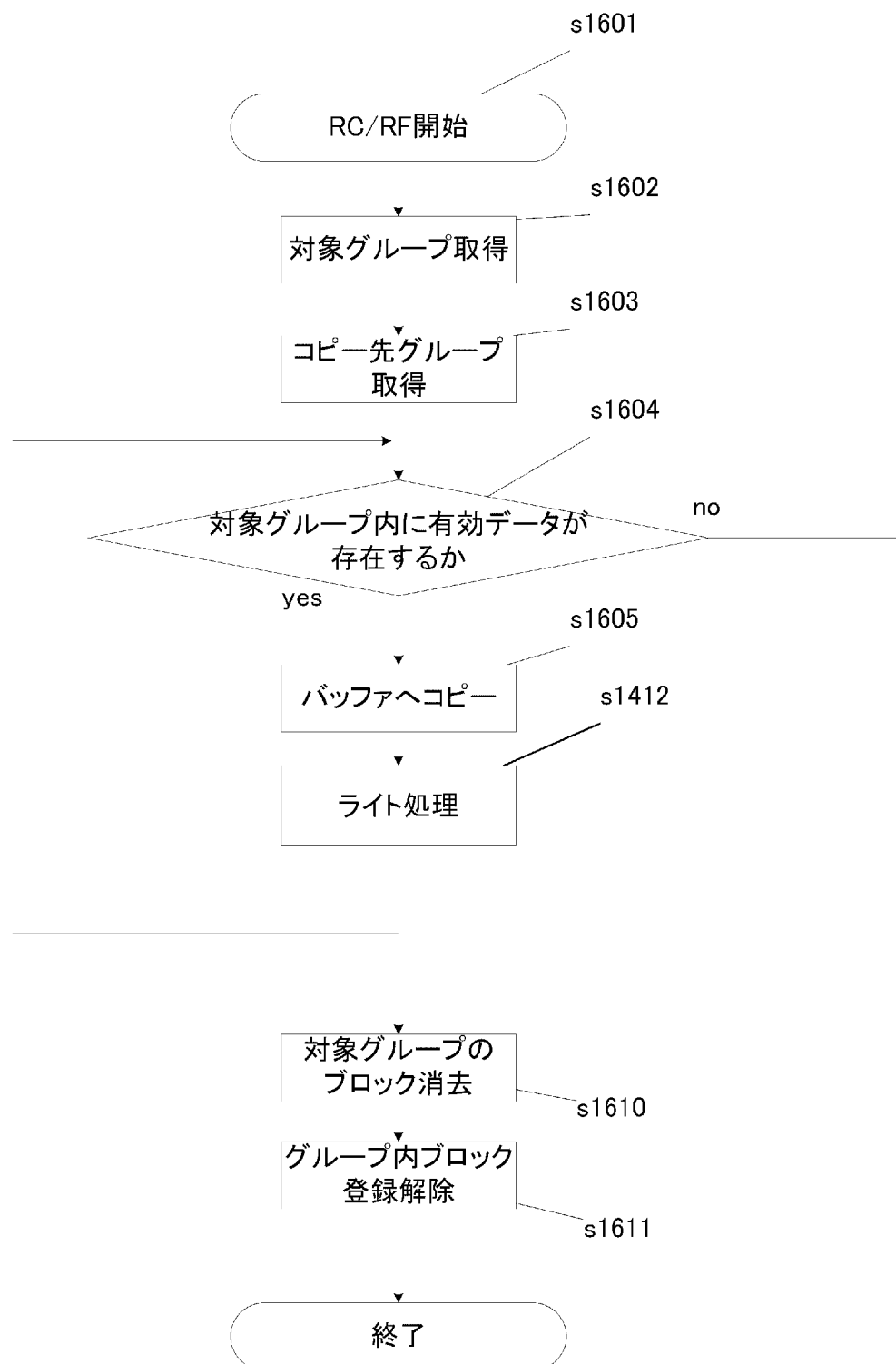
[図15]

図15



[図16]

図16



[図17]

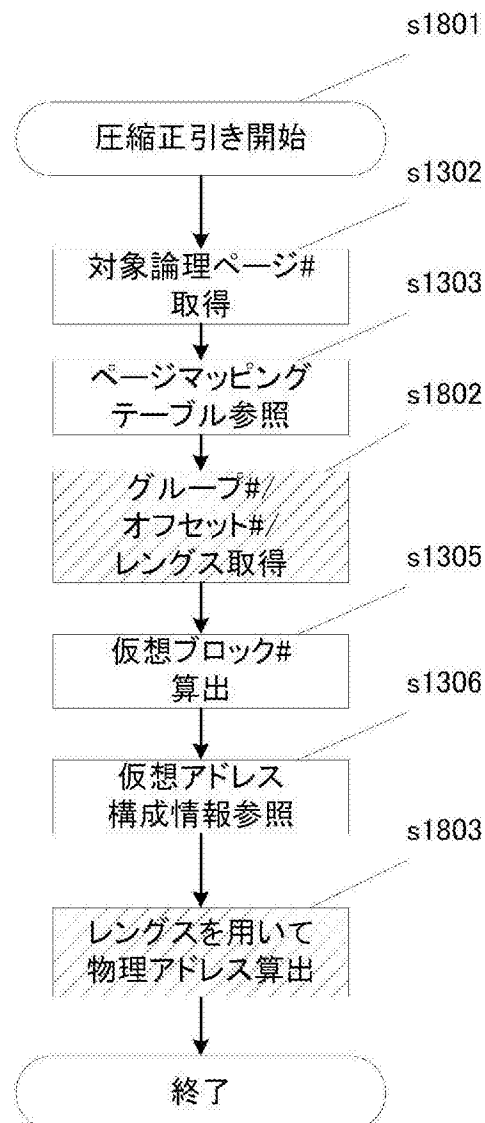
図17

1700

ページマッピングテーブル			
1701 論理ページ	1703 グループ#	1704 オフセットアドレス	1705 レンジス
0	0	0	16
1	1	10	10
...			

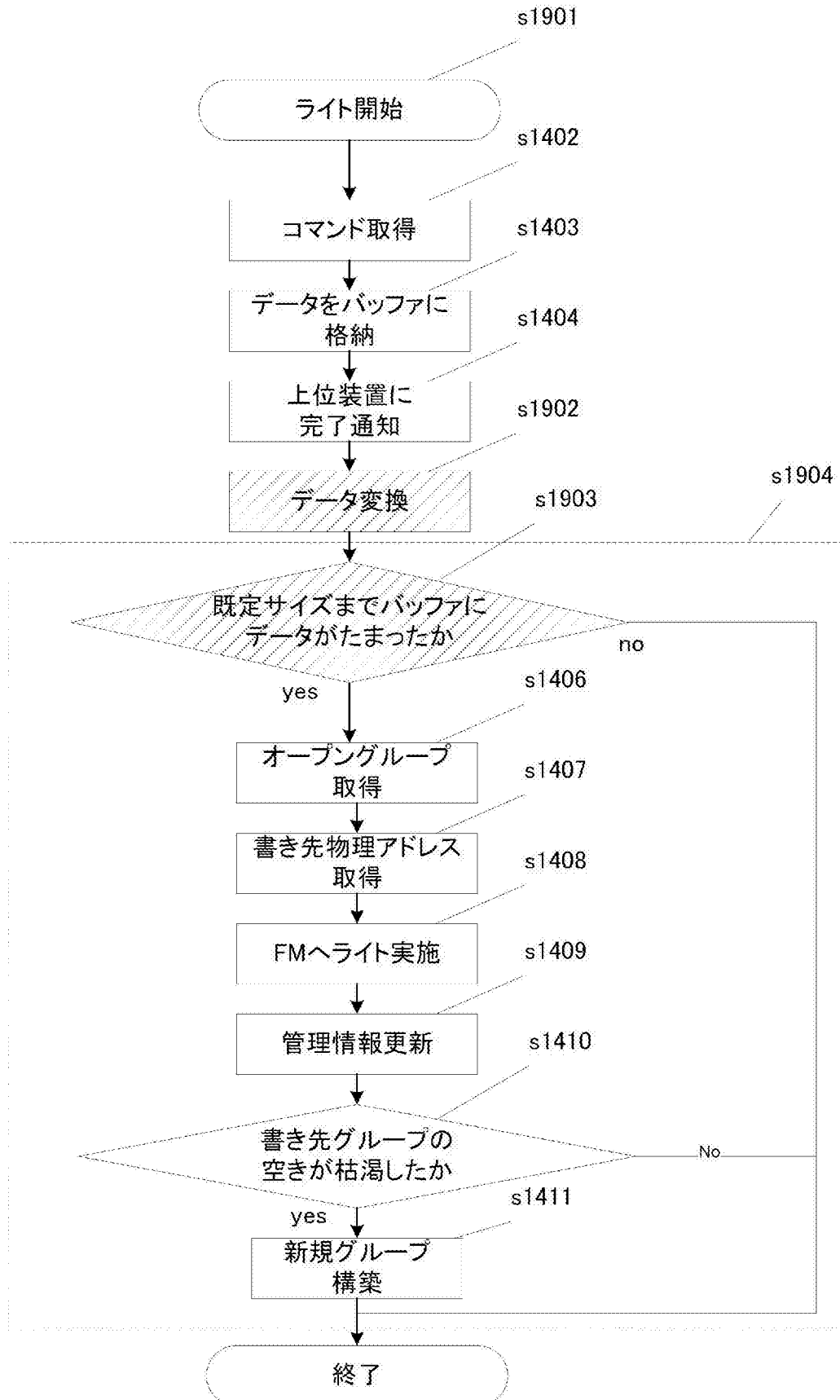
[図18]

図18



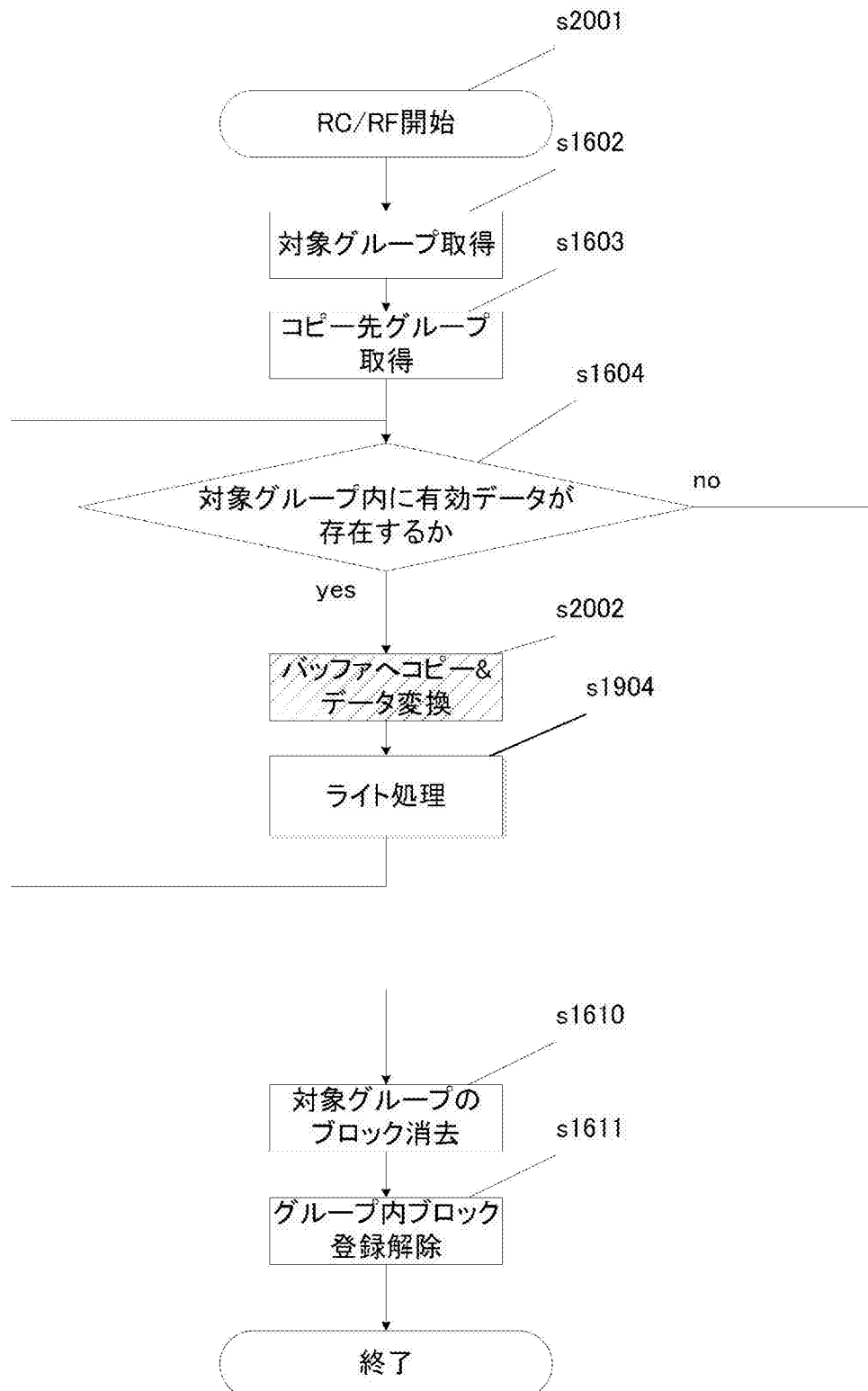
[図19]

図19



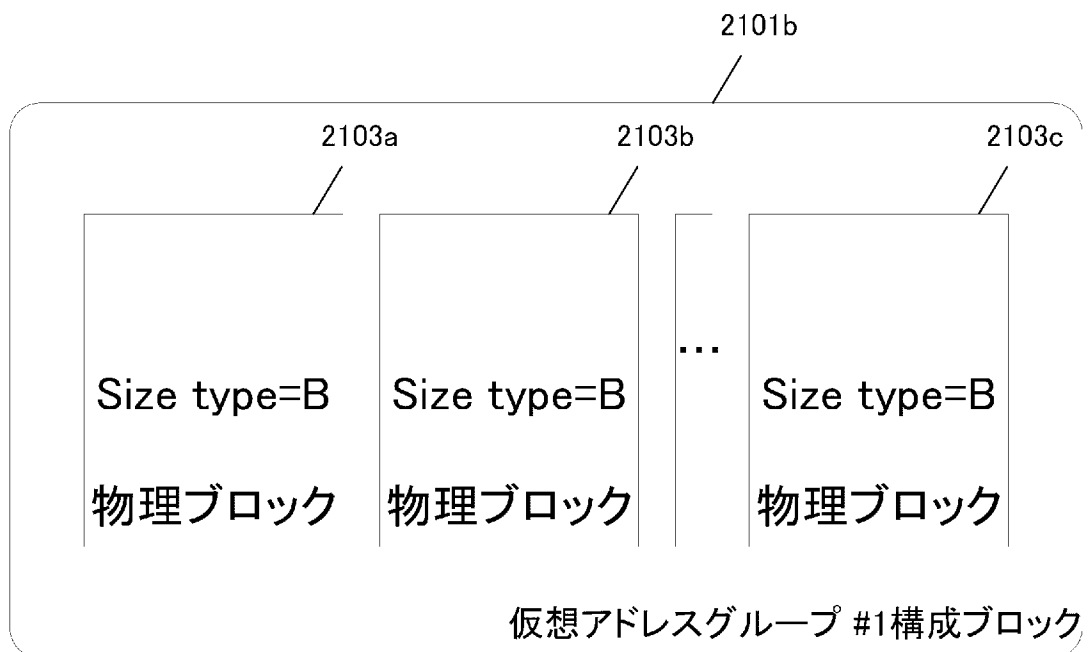
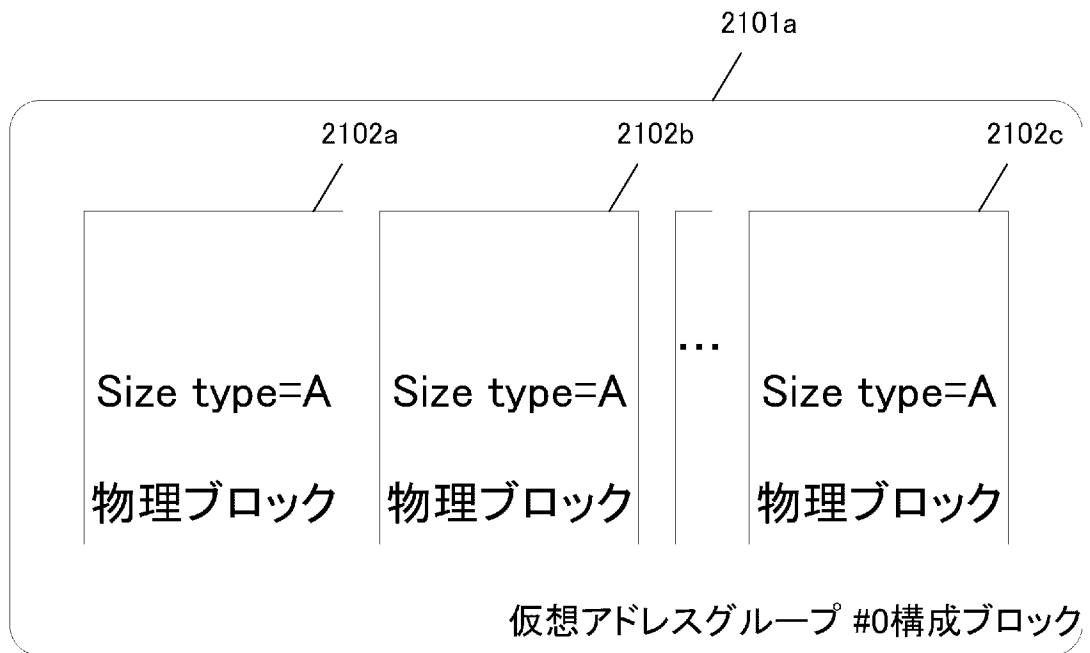
[図20]

図20



[図21]

図21



[図22]

図22

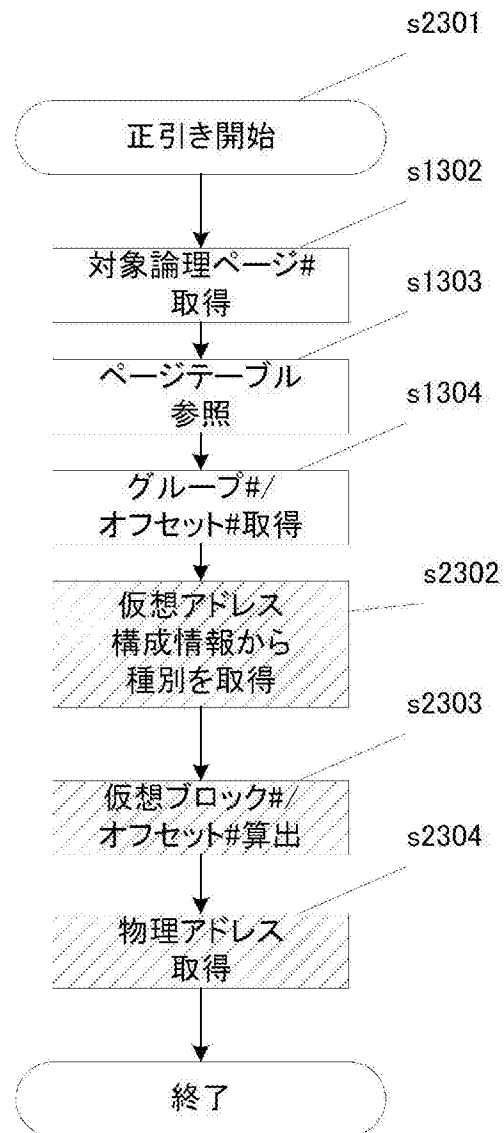
2202 グループ	2203 次書き込み 位置	2204 仮想 ブロック#	2205 物理 ブロック#	2206 無効 ページ数	2207 最終書込 時刻	2208 種別
0	10	0	0x00A	0	0:00	Type A
		1	0x00B			
		N-1	0x00F			
1	N/A	0	0x10A	22	...	Type B
		1	0x10B			
		N-1	0x10F			
.....						

仮想アドレス構成情報

2201

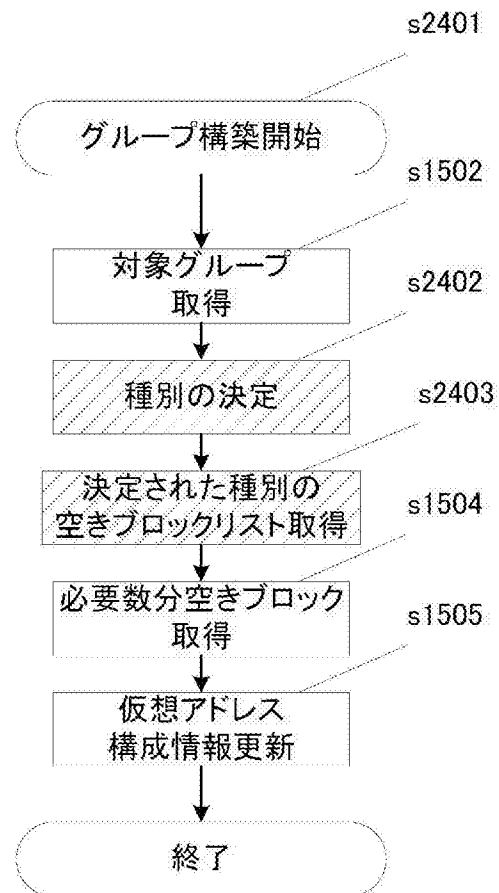
[図23]

図23



[図24]

図24



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2014/051949

A. CLASSIFICATION OF SUBJECT MATTER

G06F12/02(2006.01)i, G06F3/06(2006.01)i, G06F3/08(2006.01)i, G06F12/00(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)

G06F12/02, G06F3/06, G06F3/08, G06F12/00

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 A	JP 2012-99139 A (SanDisk Corp.), 24 May 2012 (24.05.2012), paragraphs [0044] to [0045]; fig. 9A to 9B & US 2005/0144357 A1 & EP 1700204 A & WO 2005/066770 A2 & KR 10-2007-0012630 A & CN 1922571 A & TW 280581 B	1, 3, 5, 13-15 2, 4, 6-12
X A	JP 2012-113699 A (HTC Corp.), 14 June 2012 (14.06.2012), paragraphs [0010] to [0013]; fig. 1B to 3 & US 2012/0137050 A1 & EP 2458494 A1 & CN 102566905 A & TW 201222399 A	1, 3, 5, 13-15 2, 4, 6-12
A	JP 2012-173778 A (Sony Corp.), 10 September 2012 (10.09.2012), entire text; all drawings & US 2012/0215964 A1 & CN 102693185 A	1-15

☐ 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
25 April, 2014 (25.04.14)

Date of mailing of the international search report
13 May, 2014 (13.05.14)

Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

A. 発明の属する分野の分類（国際特許分類（I P C）） Int.Cl. G06F12/02(2006.01)i, G06F3/06(2006.01)i, G06F3/08(2006.01)i, G06F12/00(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（I P C）） Int.Cl. G06F12/02, G06F3/06, G06F3/08, G06F12/00			
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 日本国公開実用新案公報 日本国実用新案登録公報 日本国登録実用新案公報 1 9 2 2－1 9 9 6 年 1 9 7 1－2 0 1 4 年 1 9 9 6－2 0 1 4 年 1 9 9 4－2 0 1 4 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 A	JP 2012-99139 A（サンディスク コーポレーション） 2012.05.24, [0044]-[0045]段落, 図 9A-9B & US 2005/0144357 A1 & EP 1700204 A & WO 2005/066770 A2 & KR 10-2007-0012630 A & CN 1922571 A & TW 280581 B	1, 3, 5, 13-15 2, 4, 6-12	
☒ 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 2 5 . 0 4 . 2 0 1 4		国際調査報告の発送日 1 3 . 0 5 . 2 0 1 4	
国際調査機関の名称及びあて先 日本国特許庁（I S A / J P） 郵便番号 1 0 0 - 8 9 1 5 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） 桜井 茂行 電話番号 0 3 - 3 5 8 1 - 1 1 0 1 内線 3 5 6 5	5 U 2 9 4 5

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
X A	JP 2012-113699 A (宏達国際電子股▲ふん▼有限公司 (HTC CORPORATION)) 2012.06.14, [0010]-[0013]段落, 図 1B-3 & US 2012/0137050 A1 & EP 2458494 A1 & CN 102566905 A & TW 201222399 A	1, 3, 5, 13-15 2, 4, 6-12
A	JP 2012-173778 A (ソニー株式会社) 2012.09.10, 全文, 全図 & US 2012/0215964 A1 & CN 102693185 A	1-15