

(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
2016年9月22日(22.09.2016)



(10) 国際公開番号

WO 2016/147351 A1

- (51) 国際特許分類:
G06F 12/00 (2006.01) G06F 12/02 (2006.01)
G06F 3/06 (2006.01)
- (21) 国際出願番号: PCT/JP2015/058100
- (22) 国際出願日: 2015年3月18日(18.03.2015)
- (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: 仁藤 拓実(NITO, Takumi); 〒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). 内垣 洋 (UCHI-
GAITO, Hiroshi); 〒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: 特許業務法人ウィルフォート国際特許
事務所 (WILLFORT INTERNATIONAL PATENT
FIRM); 〒1030016 東京都中央区日本橋小網町1
9-7 日本橋TCビル 1階 Tokyo (JP).

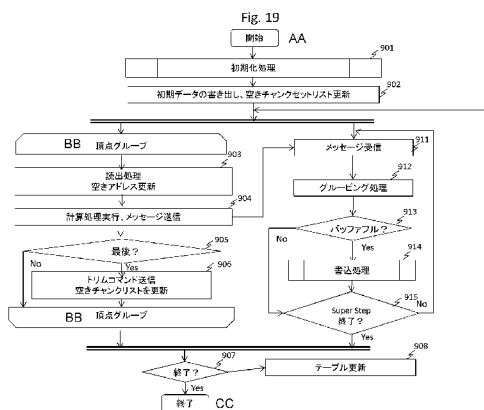
- (81) 指定国 (表示のない限り、全ての種類の国内保
護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA,
BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN,
CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES,
FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN,
IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR,
LS, LU, LY, MA, MD, ME, 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, ST, 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: COMPUTER SYSTEM, METHOD, AND HOST COMPUTER

(54) 発明の名称: 計算機システム、方法、及びホスト計算機



- 901 Initialization process
902 Store initial data and update free chunk set list
903 Perform read process and update free address list
904 Perform calculation process and transmit message
905 Last group?
906 Transmit trim command and update free chunk list
907 Completed?
908 Update table
911 Receive message
912 Perform grouping process
913 Is buffer full?
914 Write process
915 Super-step completed?
AA Start
BB Apex group
CC End

(57) Abstract: According to the present invention, a storage device stores limit information about the size of a first limited logical region, which is a part of a logical storage region of a nonvolatile memory, associates a first physical storage region, which is a part of a physical storage region of the nonvolatile memory, with the first limited logical region on the basis of the limit information, and transmits the limit information to a host computer. The host computer performs a first process, thereby generating first processed data. The host computer calculates a range of first logical addresses indicating the first limited logical region, on the basis of the limit information, and transmits, to the storage device, the first processed data and a first write command specifying a particular first logical address within the range of first logical addresses. After the completion of the first process, the host computer transmits a read command specifying the particular first logical address to the storage device, and further transmits, to the storage device, a specific command indicating that the range of first logical addresses is no longer necessary.

(57) 要約: ストレージ装置は、不揮発性メモリの論理記憶領域の一部である第1限定論理領域のサイズに関する限定情報を記憶し、限定情報に基づいて、不揮発性メモリの物理記憶領域の一部である第1物理記憶領域を、第1限定論理領域に関連付け、限定情報を、ホスト計算機に送信する。ホスト計算機は、第1処理を実行し、第1処理により第1処理データを生成する。ホスト計算機は、限定情報に基づき、第1限定論理領域を示す第1論理アドレス範囲を算出し、第1論理アドレス範囲内の第1論理アドレスを指定した第1ライトコマンドと第1処理データをストレージ装置へ送信する。ホスト計算機は、第1処理の後、第1論理アドレスを指定した第1論理アドレス範囲が不要であることを示す特定コマンド

リードコマンドをストレージ装置へ送信し、
をストレージ装置へ送信する。

明 細 書

発明の名称： 計算機システム、方法、及びホスト計算機

技術分野

[0001] 本発明は、不揮発性半導体メモリを用いた計算機システムに関する。

背景技術

[0002] 近年、HDD (Hard Disk Drive) に代えて又は加えて、不揮発性メモリを記憶媒体とするストレージ装置が用いられる。このような不揮発性メモリは、例えばNAND型のフラッシュメモリである。不揮発性メモリを記憶媒体とするストレージ装置は、複数の小型ディスクドライブを有するストレージ装置に比べて、省電力及びアクセス時間等に優れていると考えられている。

[0003] フラッシュメモリは、メモリチップ内に複数のブロックが含まれる。ブロックはデータを一括消去する単位の記憶領域であり、ページはデータを読み書きする単位の記憶領域である。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。記憶しているデータを書き換える場合、フラッシュメモリは、記憶している有効なデータを他のブロックに退避させた後、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換えは、ブロックごとのデータの消去を伴う。データの消去により、メモリチップは消耗する。このため、ブロック毎に消去可能な回数が規定されている。

[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、データの書換えが進むと、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。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、これをガーベージコレクションと呼ぶ。場合によってはGCと省略して表記する。ガーベージコレクションは、無効データが多く含まれるブロックを対象に実行される。

[0005] フラッシュメモリを用いたストレージ装置において、コントローラが、データのライト先の論理アドレスに従う論理ページに割り当てられている論理チャンクに割り当てられている物理領域を特定し、その特定した物理領域にデータを書き込む技術が知られている（特許文献1）。

[0006] 一方、インターネット環境の発展により飛躍的に増加するオブジェクトを効率的に処理する技術として、例えば、グラフ処理が知られている（特許文献2）。グラフ処理において処理されるオブジェクトは年々増加している。

先行技術文献

特許文献

[0007] 特許文献1：国際公開WO2013/046464号公報

特許文献2：特開2004-31884号公報

発明の概要

発明が解決しようとする課題

[0008] 上述のように、ストレージ装置が処理するデータは、年々増加している。不揮発性メモリ含むストレージ装置において、膨大なデータ処理を行った場合、無秩序にデータを書き込めばデータの書換え動作が増加しGCが多発し、不揮発メモリの寿命が低下するおそれがある。

課題を解決するための手段

[0009] かかる課題を解決するために本発明に係る計算機システムは、ストレージ装置と、ストレージ装置に接続されるホスト計算機とを備える。ストレージ装置は、不揮発性メモリと、不揮発性メモリ及びホスト計算機に接続されるコントローラと、を含む。ホスト計算機は、記憶デバイスと、記憶デバイス及びストレージ装置に接続されるプロセッサと、を含む。コントローラは、

不揮発性メモリの論理記憶領域の一部である第1限定論理領域のサイズに関する限定情報を記憶し、限定情報に基づいて、不揮発性メモリの物理記憶領域の一部である第1物理記憶領域を、第1限定論理領域に関連付ける。コントローラは、限定情報を、ホスト計算機に送信する。プロセッサは、第1処理を実行し、第1処理により第1処理データを生成する。プロセッサは、限定情報に基づき、第1限定論理領域を示す第1論理アドレス範囲を算出する。プロセッサは、第1論理アドレス範囲内の第1論理アドレスを指定した第1ライトコマンドと第1処理データとをストレージ装置へ送信する。コントローラは、第1ライトコマンドに応じて、第1物理記憶領域内の第1物理アドレスを、第1論理アドレスに割当て、第1物理アドレスにより示された領域に第1処理データを書き込む。プロセッサは、第1処理の後、第1論理アドレスを指定したリードコマンドをストレージ装置へ送信する。コントローラは、リードコマンドに応じて、第1物理アドレスにより示された領域から第1処理データを読み出す。プロセッサは、第1論理アドレス範囲が不要であることを示す特定コマンドをストレージ装置へ送信する。コントローラは、特定コマンドの受信の後に、第1処理データを消去する。

発明の効果

[0010] データの書込範囲を制御することで、不揮発性半導体メモリの寿命および性能の低下を防止できる。

図面の簡単な説明

[0011] [図1]グラフの概念を説明する図である。

[図2]B S P グラフ処理を説明する図である。

[図3]本実施例に係る計算機システムの構成を示す。

[図4]主記憶122に格納される情報の一例を示す。

[図5]グラフ処理に関するプログラムの構成を示す図である。

[図6]不揮発性メモリデバイス101内部の接続状態を示す図である。

[図7]論理ページとチャンクの関係を示す図である。

[図8]コントローラ112による論物変換を説明するための図である。

[図9]チャンク、チャンクセット及びチャンクグループの関係を説明する図である。

[図10]頂点 1 1 0 1 のグルーピング処理を説明する図である。

[図11]スーパーステップにおける頂点グループの処理の順序を説明する図である。

[図12]頂点セットを説明する図である。

[図13]アドレステーブルの一例である。

[図14]チャンクセット割当テーブル 2 2 5 の一例である。

[図15]空きチャンクセットリスト 2 2 6 の一例である。

[図16]チャンクセット空きリストテーブル 2 2 7 の一例である。

[図17]メッセージ 1 1 0 2 のグルーピング処理及び書込処理を説明する図である。

[図18]メッセージ 1 1 0 2 の読出処理を説明する図である。

[図19]本実施例にかかる書込範囲制御処理のフローチャートを説明する。

[図20]初期化処理のフローチャートである。

[図21]書込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

[0012] 以下の説明では、「k k k リスト」、「k k k テーブル」、「k k k 情報」の表現にて情報を説明することがあるが、情報は、どのようなデータ構造で表現されていてもよい。データ構造に依存しないことを示すため「k k k リスト」、「k k k テーブル」、「k k k 情報」は、いずれの呼び方にも置換できる。さらに、各情報の内容を説明する際に、「ID」、「番号」、「識別子」という表現を用いるが、これらに代えて又は加えて他種の識別情報が使用されてよい。

[0013] また、以下の説明では、「プログラム」を主語として処理を説明する場合があるが、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、適宜に記憶資源（例えば、メモリ）及び／又は通信インターフェイスデバイス（例えば、通信ポート）を用いながら行うため、処理の主語が

プロセッサとされてもよい。逆に、プロセッサが主語となっている処理は、1以上のプログラムを実行することにより行われると解釈することができる。

[0014] 以下、ホスト計算機と不揮発性メモリデバイスとを有する計算機システムを例に取り、実施例を説明する。この計算機システムは、ホスト計算機が、データの演算方法としてBSP (Bulk Synchronous Parallel) グラフ処理を行う。

[0015] 図1は、グラフ処理の概念を説明する図である。グラフは、頂点と、頂点間のつながりを表す辺とを示す概念である。BSPグラフ処理を実行するにあたり、処理は頂点ごとに行われ、辺に沿ってメッセージを送る。例えば、頂点V1と頂点V2及びV3を結ぶ辺は、頂点V1が頂点V2及びV3へメッセージを通知することを示す。同様に、頂点V2は頂点V4にメッセージを通知し、頂点V3は頂点V2にメッセージを通知し、頂点V4は頂点V1及びV3にメッセージを送る。

[0016] 図2は、BSPグラフ処理を説明する図である。BSPグラフ処理においては、全ての頂点1101が概念上並列に処理される。頂点1101には、頂点1101を識別するための頂点IDが含まれる。例えば、V1～V4が頂点IDである。各頂点1101の処理に際して出力されるデータがメッセージ1102である。メッセージ1102には、出力元の頂点1101に係性のある1の頂点1101を示す宛先が付される。宛先の頂点（以下、宛先頂点）は、メッセージ1102を入力し、処理を行う。最初の頂点1101の処理の開始から最後の頂点1101の処理の完了までをスーパーステップという。スーパーステップ毎に同期がとられる。

[0017] 例えば、スーパーステップSにおいて、（1）頂点が、自身を宛先とするメッセージを入力し（read messages）、頂点が、処理を実行して頂点の状態を更新する（modify state）。（2）頂点が、次の宛先の頂点に対するメッセージを出力する（send message）。（3）宛先の頂点が、出力されたメッセージを受け取る（Recv messages）。（4）全ての頂点がメッセージを

受け取ったときに、スーパーステップSが完了し（同期；Synchronization）、次のスーパーステップS 1に移行する。

[0018] B S P グラフ処理を行う計算機がメッセージを不揮発性メモリへ格納する場合、メッセージのサイズは不揮発性メモリのブロックサイズより小さいため、不揮発性メモリに不要となったメッセージが無秩序に格納されていれば、ブロック内のデータの断片化が多発し、G C が頻繁に行われる可能性がある。G C の際には、消去ブロック内の有効データを他のブロックに移動しなければならない。従って、断片化がおきると余分な書き換えが発生してしまう。つまり、ブロックの消去回数が増加し、不揮発性記憶媒体の寿命を縮めることとなる。

[0019] 本実施例の計算機システムは、複数のメッセージをまとめて特定の物理記憶領域へ書き込み、複数のメッセージをまとめて消去することにより、G C を減らすことで、ブロックの消去回数を低減でき、不揮発性メモリの寿命を縮めることを防止することができる。さらに、G C による不揮発性メモリのアクセス性能低下を防止できる。

[0020] 図3は、本実施例に係る計算機システムの構成を示す。

[0021] 計算機システムは、不揮発性メモリデバイス101とホスト計算機（以下、ホスト）102が通信ネットワークを介して接続されたシステムである。不揮発性メモリデバイス101及びホスト102は、不揮発性メモリデバイス101内での通信プロトコルとホスト102内での通信プロトコルと同じ通信プロトコル（例えばP C I E x p r e s s（登録商標））で通信されてもよいし異なる通信プロトコル（例えばF i b r e C h a n n e l）で通信されてもよい。なお、計算機システムは、不揮発性メモリデバイス101とホスト102が1つの筐体に収納されたコンバインドシステムであってもよい。

[0022] 不揮発性メモリデバイス101は、例えば、1つ以上のメモリチップ111を有する記憶メディアドライブ（例えばS S D ; Solid State Drive）である。また、不揮発性メモリデバイス101は、複数の不揮発性メモリデバ

イス101を有する大規模メモリシステムであってもよい。本実施例では、不揮発性メモリデバイス101は、1つ以上のメモリチップ111、コントローラ112、チャンク情報レジスタ114及びホストI/F118を有するメディアドライブであるとして説明する。これらの要素は、例えばバス（例えばPCI Expressバス）等で相互に接続されていてよい。

[0023] メモリチップ111は、本実施例では、NAND型のフラッシュメモリである。従って、メモリチップ111は、複数のブロックで構成されており、各ブロックは、複数の物理ページで構成されており、物理ページ毎にデータが読み書きされ、ブロック毎にデータが消去される。つまり、物理ページがデータの書込単位であり、ブロックがデータの消去単位である。メモリチップ111は、NAND型のフラッシュメモリに代えて、他種の不揮発性メモリ（例えば、MRAM（Magnetoresistive random access memory）、ReRAM（resistance random access memory）、又は、FeRAM（Ferroelectric random access memory））でもよい。

[0024] コントローラ112は、CPU（Central Processing Unit）を含んだモジュール（例えば、CPUそれ自体、又は、CPUとDMA等の専用ハードウェアとを含んだモジュール）である。コントローラ112は、ホスト102からのリクエストに応答して処理を行う。コントローラ112は、不揮発性メモリデバイス101内でのデータ転送を制御する。

[0025] チャンク情報レジスタ114は、不揮発性メモリデバイス101内の図示しないメモリに設けられた記憶領域である。メモリは、揮発メモリでも不揮発性メモリでもよい。メモリは、例えばDRAM（Dynamic Random Access Memory）である。チャンク情報レジスタには、後述のチャンク情報が格納される。

[0026] ホストI/F113は、ホスト102に接続されるインターフェイスデバ

イスである。

[0027] ホスト102は、CPU121、主記憶122及び不揮発性メモリ1/F123を有する。CPU121は、ホスト102の動作を制御する。主記憶122は、ホスト102が有するメモリである。主記憶122は、例えばDRAMであるが、DRAM122に代えて他種のメモリが採用されてもよい。不揮発性メモリ1/F123は、不揮発性メモリデバイス101のホスト1/F113に接続されるインターフェイスデバイスである。ホスト102は、CPU121で実行されるプログラムを記憶した不揮発性メモリ（以下、プログラムメモリ）を有していてもよい。プログラムメモリからプログラムが主記憶122にロードされ、主記憶122にロードされたプログラムがCPU121により実行されてよい。プログラムメモリが無く、CPU121で実行されるプログラムが、不揮発性メモリデバイス101のメモリチップ111から主記憶122にロードされてもよい。なお、ホスト102には、ユーザが後述のユーザコード等のデータを入力するための図示しない入力装置、及びユーザに処理の結果を表示するための図示しない出力装置が接続されてよい。

[0028] 図4は、ホスト102の主記憶122に格納される情報の一例を示す。主記憶122は、ユーザコードプログラム221と、グラフフレームワーク（グラフFW）プログラム222と、現アドレステーブル223と、前アドレステーブル224と、チャンクセット割当テーブル225と、空きチャンクセットリスト226と、チャンクセット空きリストテーブル227と、管理ミドルウェア228を格納する。これらの情報は、前述の通り、プログラムメモリに記憶されてもよいし、他所から主記憶122に一時的に読み出されて使用されてもよい。また、主記憶122には、グラフ処理に必要なデータを読み出すためのワーキングエリア229が確保される。

[0029] ユーザコードプログラム221は、ユーザが入力するユーザコードにより、グラフ処理の内容を指定するプログラムである。グラフFWプログラム222は、グラフ処理を実行するプログラムである。管理ミドルウェア228

は、これらのプログラム 221、222 を制御するプログラムであり、例えば、OS である。

[0030] 例えば、ユーザコードプログラム 221 は、グラフフレームワークプログラム 222 に対して、ユーザからの入力に従った論理アドレスのサイズを指定する。グラフフレームワークプログラム 222 は、指定された論理アドレスサイズを管理ミドルウェア 228 に要求する。管理ミドルウェア 228 は、要求に含まれる論理アドレスのサイズに基づき、利用可能なメモリチップ 111 の記憶領域を確保して、利用可能な論理アドレスの範囲（開始論理アドレス及び終了論理アドレス）を、グラフフレームワークプログラム 222 に通知する。各情報 223～227 の説明は、後述する。

[0031] 図 5 は、グラフ処理に関するプログラムの構成を示す。

[0032] ユーザコードプログラム 221 及びグラフフレームワークプログラム 222 が、管理ミドルウェア 228 上で実行される。ユーザコードプログラム 221 には、各頂点計算処理部 301 が含まれる。グラフフレームワークプログラム 222 には、計算処理順制御部 311、計算処理部 312、データ通信処理部 313、書込処理部 314、及び読出処理部 315 が含まれる。書込処理部 314 には、グルーピング処理部 321 及びアドレス管理部 322 が含まれる。これら各部は、例えばプログラムモジュールである。

[0033] 各頂点計算処理部 301 は、ユーザコードにより指定され、各頂点の処理を示す。計算処理順制御部 311 は、計算処理部 312 における計算処理の順序を制御する。計算処理部 312 は、各頂点計算処理部 301 の処理を実行する。データ通信処理部 313 は、312、314 間のメッセージの通信を制御する。書込処理部 314 は、データの不揮発性メモリへの書き込みを制御する。グルーピング処理部 321 は、データをグルーピングする。アドレス管理部 322 は、論理アドレスを管理する。読出処理部 315 は、不揮発性メモリからのデータの読み出しを制御する。

[0034] 図 6 は、不揮発性メモリデバイス 101 内部の接続状態を示す図である。

[0035] コントローラ 112 は、複数の DMA (Direct Memory Access) 115

を有する。DMA 115は、データ転送を制御するハードウェア回路である。各DMA 115が、バス116を介して複数のメモリチップ111に接続される。以下の説明及び図面では、コントローラ112内の複数のDMA 115に対しDMA ID（#）を付して説明する場合がある。図示例では、コントローラ112は、DMA #0～DMA #15の16個のDMAを有する。また、以下の説明及び図面では、各DMAに接続された複数のメモリチップ111に対してCE ID（#）を付して説明する場合がある。図示例では、各DMAに16個のメモリチップ111、CE #0～CE #15が接続されている。なお、コントローラ112は、DMA 115を有さなくてもよい。この場合、コントローラ112が不揮発性メモリデバイス101内のデータ転送を制御する。

[0036] 図7は、論理ページとチャンクの関係を示す図である。

[0037] コントローラ112が、チャンク920と、複数のチャンク920を含むチャンクグループ960とを設定する。チャンクグループ960は、複数のチャンク920の行列である。本実施例においてコントローラ112は、各メモリチップ111を一つのチャンク920に割り当てる。これにより、チャンクグループ960の横（行）方向に並ぶチャンクの数（横チャンク数）が、DMA 115の数（16個）に対応する。チャンクグループ960の縦（列）方向に並ぶチャンクの数（縦チャンク数）が、1つのDMA 115にバスを介して接続されるメモリチップ111の数（CE数、16個）に対応する。

[0038] コントローラ112は、チャンクグループ960内の各行をチャンクセットとして定義する（図9参照）。なお、縦チャンク数は、チャンクセット数と同意であり、横チャンク数は、各チャンクセット内のチャンク数と同意である。

[0039] コントローラ112は、各チャンク920に複数の論理ページ922を設定する。コントローラ112は、各論理ページに論理アドレスを対応づける。なお、本実施例では、コントローラ112は、論理ページ922を各メモ

リチップ111内の物理ページと同一サイズで管理することとするが、異なるサイズで管理してもよい。

[0040] コントローラ112は、各チャンク920に複数の論理ブロック925を割り当てる。本実施例では、論理ブロック925を各メモリチップ111内の物理ブロック930と同一サイズで管理することとするが、異なるサイズで管理してもよい。

[0041] ここで、コントローラ112が設定する論理ページとチャンクの対応づけの具体例を示す。

[0042] 例えば、コントローラ112は、チャンクグループ960内のチャンク920に連続するチャンク番号を割り振る。例えば、コントローラ112は、1行目の各チャンク920をチャンク#00～#15とし、2行目の各チャンク920をチャンク#16～#31とするように、チャンクグループ960内のチャンクにチャンク#00～#255を割り振る。

[0043] そして、コントローラ112は、チャンクグループ960内の各チャンク920に論理ブロック925を割り当てる。例えば、コントローラ112は、1行目の第1列のチャンク（CHUNK#00）に1つの論理ブロック（B#00）を割当て、次に1行目の第2列のチャンク（CHUNK#01）に1つの論理ブロック（B#01）を割り当て、1行目の最終列のチャンク（CHUNK#015）に1つの論理ブロック（B#015）を割り当てた後、2行目の第1列のチャンク（CHUNK#016）に1つの論理ブロックを割当てる。そして、コントローラ112は、最終（16）行の最終（16）列のチャンク（CHUNK#225）まで、1つずつ論理ブロックを割り当てた後、再び当該チャンクグループの先頭のチャンク（CHUNK#00）から、2つめの論理ブロックを割り当てていくというように、チャンクグループ960内の全てのチャンク920に複数の論理ブロックをそれぞれ割り当てる。つまり、コントローラ112は、チャンクグループ960内の全てのチャンク920に対して、まず1つ目の論理ブロックをチャンク番号順に割り当て、次いで2つ目の論理ブロックをチャンク番号順に割り当ててい

くというように、各チャンクの空きがなくなるまで論理ブロックを割り当てる。

[0044] コントローラ 1 1 2 は、論理ブロック 9 2 5 に論理ページ 9 2 2 を割り当てる。ここで、チャンクインタリーブ L P 数は、各チャンク 9 2 0 に連続して割り当てられる論理ページ 9 2 2 の数である。この図においては、チャンクインタリーブ L P 数は、2 である。例えば、コントローラ 1 1 2 は、論理ページ番号順に、当該チャンクグループ内の先頭のチャンク内の 1 つ目の論理ブロック (B # 0 0) に 2 つの論理ページ # 0、# 1 を割り当て、1 行目の第 2 列のチャンク内の 1 つ目の論理ブロック (B # 0 1) に論理ページ # 2、# 3 を割り当て、1 行目の第 1 5 列のチャンク内の 1 つ目の論理ブロック (B # 1 5) に論理ページ # 3 0、3 1 を割り当てた後、1 行目の第 1 列のチャンク内の 1 つ目の論理ブロック (B # 0 0) の続く論理ページ # 3 1、3 2 を割り当てる。このようにして、コントローラ 1 1 2 は、1 行目の各チャンクの 1 つ目の論理ブロック (B # 0 0 ~ 1 5) に全て論理ページを割り当てた後、2 行目の各列の 1 つ目の論理ブロックに、同じように論理ページを割り当てる。そして、コントローラ 1 1 2 は、最終行の各列の 1 つ目の論理ブロックに、同じように論理ページを割り当てた後、1 行目の各列の 2 つ目の論理ブロックに、同じように論理ページを割り当てる。このようにして、コントローラ 1 1 2 は、チャンクグループ 9 6 0 内の各チャンク 9 2 0 に予め決められたチャンクグループ内 L P 数まで論理ページ 9 2 2 を割り当てる。チャンクグループ内 L P 数は、例えば、各チャンクに論理ページ数を均一に配置するために、チャンクグループ内に配置可能な論理ページ番号をチャンクグループ内のチャンク数で割った値 (整数値) に、縦チャンク数と横チャンク数を掛けた値であってもよい。

[0045] コントローラ 1 1 2 が、論理アドレスから、それに対応するチャンクを特定する方法を以下に示す。

[0046] コントローラ 1 1 2 は、チャンク情報として、例えば、論理ページのページサイズ、チャンクグループ 9 6 0 内の L P 数、論理ブロック 9 2 5 内 L P

数、チャンクグループ960内の横チャンク数（チャンクセット数、DMA数）、チャンクグループ960内の縦チャンク数（チャンクセット内チャンク数、1つのDMAに接続されるチャンク数）、チャンクインタリーブLP数を記憶する。

[0047] コントローラ112は、論理アドレス範囲から、各論理アドレスに対応する、論理ページ番号（＃）、チャンクグループ番号（＃）、チャンクグループ内論理ページ番号（＃）、チャンクグループ内論理ページ番号（＃）、横チャンク番号（＃）及び縦チャンク番号（＃）を算出する。これらの値の算出方法は、以下の通りである。

（１）論理ページ番号（＃）＝論理アドレス／ページサイズ

（２）チャンクグループ番号（＃）＝論理ページ番号（＃）／チャンクグループ内LP数

（３）チャンクグループ内LP番号（＃）＝LP番号％チャンクグループ内LP数

（４）縦チャンク番号（＃）＝（チャンクグループ内LP数／（論理ブロック内LP数＊横チャンク数））％チャンクグループ内縦チャンク数

（５）横チャンク番号（＃）＝（チャンク内LP番号／チャンクインタリーブLP数）％チャンクグループ内横チャンク数

なお、コントローラ112は、縦チャンク番号（＃）及び横チャンク番号（＃）に基づき、各チャンクをチャンクグループ内で一意に特定できるチャンク番号（＃）を算出してもよい。

[0048] コントローラ112は、上記計算式を用いて論理アドレスからチャンクを特定することができる。

[0049] また、コントローラ112は、チャンク情報をホスト102に送信する。ホスト102は、コントローラ112から送信されたチャンク情報を用いて、各チャンクセットに対応する論理アドレスの範囲を決定することができる。

[0050] 図8は、コントローラ112による論物変換を説明するための図である。

メモリチップ１１１は、複数の物理ブロックを含む。各物理ブロックは、複数の物理ページを含む。以下の説明及び図面においては、論理ページをＬページ、物理ページをＰページと記載する場合がある。

[0051] ホスト１０２は、チャンク情報に基づいて論理アドレス９１１を決定し、論理アドレス９１１を指定して不揮発メモリ１０１にコマンドを送信する。コントローラ１１２は、ホスト１０２から論理アドレスを指定したライトコマンドを受信した場合、その論理アドレスに対応する論理ページ及びチャンクを特定し、そのチャンクに対応するメモリチップ１１１の中から空き物理ブロックを選択する。そしてコントローラ１１２は、選択した物理ブロック９３０内の先頭から空き物理ページ９３３を当該論理ページ９２２に割当てる。これにより、ライトデータの格納場所が決定する。例えば、ホスト１０２からＬＢＡ０×００を指定したライトコマンドを受信した場合、コントローラ１１２は、ＬＢＡ０×００に対応するＬページ＃０とチャンク＃０を特定し、チャンク＃０に対応するメモリチップ＃０内の空きブロック＃００を選択し、ブロック＃００内のＰページ＃００を選択し、Ｌページ＃００にＰページ＃００を割り当て、Ｐページ＃０にライトコマンドに従うライトデータを書き込む。

[0052] なお、本実施例では、コントローラ１１２は、メモリチップ１１１と、チャンク９２０とを１対１で割り当てている。しかし、チャンク９２０と複数の物理ブロック９３０とが割り当て可能であれば、この態様に限らない。

[0053] また、コントローラ１１２は、消去単位である物理ブロック９３０毎にデータを消去する。この場合、コントローラ１１２は、空き物理ブロック数が予め設定された閾値を下回ったか否かを判定し、下回った場合に、最も無効ページ率が高い物理ブロック内の有効ページを他の物理ブロックにコピーするＧＣを実行する。コントローラ１１２は、ＧＣにより、全て無効ページとなった物理ブロック内のデータを消去することにより、その物理ブロック内の物理ページをデータの格納が可能な空きページとする。

[0054] 上記のように、複数の論理ページ９２２を含むチャンク９２０を設定する

ことにより、異なるチャンク 920 に書き込まれたライトデータは、必ず異なる物理ブロック 930 に格納されることとなる。前述の通り、チャンク情報がコントローラ 112 からホスト 102 に通知される。このため、ホスト 102 が、ライトデータの格納先のチャンクセットを選択できる。例えば、ホスト 102 は、予め設定された条件に基づき不要となるデータを 1 つのチャンクセットに纏めて格納することができる。具体的には、第 1 のタイミングで不要となり無効化される複数のデータを、第 1 チャンクセットに対応する複数の物理ブロック 930 に纏めて格納し、纏めて消去し、それらを空きブロックにすることができる。これにより GC を削減することができる。

[0055] また、例えば、ホスト 102 は、異なる条件に基づき不要となるデータを別のチャンクセットに格納することができる。第 1 のタイミングで不要となり、無効化されるデータを第 1 のチャンクセットに格納し、第 2 のタイミングで不要となり、無効化されるデータを第 2 チャンクセットに格納することができる。従って、第 1 のタイミングで不要となり無効化されるデータと、第 2 のタイミングで不要となり無効化されるデータとが、同じ物理ブロックに格納されるのを防止できる。

[0056] またこの場合、行方向に連続するチャンクに、チャンクインタリーブ LP 数の連続する論理ページを配置することで、コントローラ 112 は、チャンクセット内に割り当てられた各メモリチップ 111 にライトデータをチャンクインタリーブ LP 数ずつ分散して格納できる。これにより、メモリチップ 111 間の劣化の偏りを防ぐことができる。

[0057] 図 9 は、チャンク、チャンクセット及びチャンクグループの関係を説明する図である。

[0058] ホスト 102 は、チャンクグループは、複数のチャンクセットを含む。

[0059] 予め設定された条件に基づきホスト 102 により無効化されるデータが、同じ無効化単位に格納される。本実施例では、無効化単位は、データの消去単位である物理ブロックの複数の割り当てられたチャンクセットである。

[0060] 本実施例では、無効化単位に格納されるデータは、例えば B S P グラフ処

理におけるメッセージ群である。各メッセージのサイズはブロックサイズより小さい。BSPグラフ処理では、各スーパーステップで多くの数の頂点が処理される。このため、頂点から出力されるメッセージの数も多くなる。もし、各メッセージを無秩序に不揮発性メモリデバイス101に書き込めば、データの断片化が多発する原因となる。これに対し、本実施例の不揮発性メモリデバイス101は、複数の消去単位を割り当てた無効化単位を設定することで、データの断片化を防止し、GCの回数を減らすことができる。なお、無効化単位であるチャンクセットには、一つのスーパーステップで一つの頂点セットを宛先としたメッセージの集合が格納される。

[0061] また、前述のようにDMA115（バス116）の数に合わせてチャンクセット（縦チャンク）を設定することで、DMA115に接続される各メモリチップに並列にアクセスでき、入出力のバンド幅を確保することができる。なお、チャンクセットは、利用状況に合わせてユーザが変更可能である。ユーザは、各DMA115（バス116）に接続されるメモリチップ111の一部をチャンクセットと設定してもよいし、1又は複数のDMA115（バス116）に接続される複数のチャンク920を纏めてチャンクセットを設定してもよい。

[0062] 図10は、頂点1101のグルーピング処理を説明する図である。頂点のグルーピング処理は、グルーピング処理部321により実行される。グルーピング処理部321は、各BSPグラフ処理において実行される全頂点1101を、頂点グループ1103に振り分ける。頂点グループ1103は、頂点1101の読み出しに必要な主記憶122のワーキングエリア229のサイズによって定義される。換言すると、一つのスーパーステップで一つの頂点グループに入力されるメッセージの合計サイズは、ワーキングエリア229のサイズ以下である。

[0063] ホスト102は、ワーキングエリア229に読み出される頂点のデータ及びメッセージのサイズによって、頂点グループに属する頂点の数を決定する。各頂点1101の振分方法は、例えば、各頂点1101の頂点IDを頂点

グループ数 (g) で割ったときの余りの値を頂点グループ I D とする。なお、振り分け方法は、上記の方法に限られない。例えば、頂点 I D 順に頂点グループに振り分ける方法でもよい。

[0064] なお、各頂点計算処理部 3 0 1 が各頂点を定義し、計算処理順制御部 3 1 1 の指示により計算処理部 3 1 2 が、各頂点の処理を実行する。各頂点は、複数の C P U 1 2 1 によって並列に処理されてもよいし、C P U 1 2 1 内の複数のコアによって並列に処理されてもよい。また、各頂点は、スーパーステップ毎に一括りに処理されるのであれば、複数の C P U 1 2 1 (又はコア) により厳密に並列に処理されなくてもよい。

[0065] 図 1 1 は、スーパーステップにおける頂点グループの処理の順序を説明する図である。

計算処理順制御部 3 1 1 は、スーパーステップにおける各頂点 1 1 0 1 の処理を頂点グループ 1 1 0 3 毎に行うように制御する。このように頂点 1 1 0 1 をグルーピングし、頂点グループ 1 1 0 3 毎に処理を行うことで、頂点 1 1 0 1 及びメッセージ 1 1 0 2 の読み出しサイズを、主記憶 1 2 2 のワーキングエリア 2 2 9 のサイズ以下にすることができる。つまり、ワーキングエリア 2 2 9 のサイズ以下となるように頂点グループ 1 1 0 3 内の頂点 1 1 0 1 の数を決定できる。計算処理順制御部 3 1 1 は、決定した数の頂点 1 1 0 1 を主記憶 1 2 2 のワーキングエリア 2 2 9 に記憶する。

[0066] なお、計算処理順制御部 3 1 1 は、頂点グループの処理順序を、各スーパーステップの処理前に決定してもよいし、グラフ処理の実行前に決定してもよい。また、グラフ処理で処理される頂点の数、頂点グループに属する頂点の数、頂点グループの数は、ホスト 1 0 2 のユーザコードプログラムに対し、ユーザが設定してよい。

[0067] 図 1 2 は、頂点セットを説明する図である。

ホスト 1 0 2 は、予め設定された条件を満たした場合、複数の頂点グループ 1 1 0 3 を含めた頂点セットを定義する。例えば、1 つの頂点グループは 2 つのチャンクセットを用いるので、計算処理順制御部 3 1 1 は、チャンクセ

ットの数と頂点グループの数の2倍よりも少ない場合、チャンクセットが不足するので、頂点セットを定義する。本実施例では、頂点セットが定義される場合を説明する。

[0068] ホスト102は、1つの頂点セットの各頂点を宛先とするメッセージ1102の格納先として1つのチャンクセットを割り当てる。1つの頂点セットが1つの頂点グループであってもいい。チャンクセットの数と頂点グループの数の2倍よりも少なく、1つの頂点グループに2つのチャンクセットを割り当てられない場合に、計算処理制御部311は、1つの頂点セットに複数の頂点グループを含める。

[0069] これにより、1つのスーパーステップの間に、1つの頂点セットに入力されるメッセージを格納するための読出用チャンクセットと、その頂点セットの各頂点を宛先として出力されるメッセージを格納するための書込用チャンクセットとの2つのチャンクセットを用意することができる。

[0070] なお、前述の通り、頂点グループに含まれる頂点の数は、ワーキングエリア229のサイズに合わせて設定される。このため、頂点グループの数がチャンクセットの数の2倍を超えた場合には、頂点セットを定義することで、すべての頂点グループをチャンクセットに割り当てることができる。

[0071] 図13は、アドレステーブルの一例である。

本実施例では、ホスト102は、現アドレステーブル223及び前アドレステーブル224の2つのアドレステーブルを有する。現アドレステーブル223は、BSPグラフ処理において、実行中のスーパーステップ（現スーパーステップ）における各頂点グループ401と、各頂点グループ1103内の頂点1101で処理され、出力されるメッセージ1102を格納する論理ページ922の論理アドレス一覧であるアドレスリスト402とを対応づけるテーブルである。例えば、アドレスリスト402は、各頂点グループ1103に割り当てられたチャンクセットに含まれる論理ページ922を示す論理アドレスが設定される。

[0072] 前アドレステーブル224は、BSPグラフ処理において、現スーパース

トップの1つ前のスーパーステップ（前スーパーステップ）において出力されたメッセージの宛先の頂点グループ1103と、各頂点グループ1103内の各頂点1101で処理されるメッセージ1102の論理アドレスとを対応づけるテーブルである。これらのアドレステーブル222、223は、同様の構成である。以下、これらアドレステーブルの共通構成を説明する。アドレステーブルは、頂点グループ1103毎のエントリを有する。各エントリは、頂点グループ1103を識別するGroup ID401と、頂点グループ1103内の各頂点1101を宛先とするメッセージ1102が格納される論理ページ922の論理アドレスの一覧であるアドレスリスト402とを有する。

[0073] 図14は、チャンクセット割当テーブル225の一例である。

チャンクセット割当テーブル225は、複数の頂点グループからなる頂点セットと複数のチャンクからなるチャンクセットとを対応づけるテーブルである。チャンクセット割当テーブル225は、頂点セット毎のエントリを有する。各エントリは、頂点セットの識別子である頂点セット番号（#）501と、チャンクセットの識別子であるチャンクセット番号（#）502とを有する。

[0074] 図15は、空きチャンクセットリスト226の一例である。

空きチャンクセットリスト226は、頂点セットに割り当てられていないチャンクセットのリストである。

[0075] 図16は、チャンクセット空きリストテーブル227の一例である。

チャンクセット空きリストテーブル227は、チャンクセットと、チャンクセット内の論理ページ922の論理アドレスとを対応づけるテーブルである。チャンクセット空きリストテーブル227は、チャンクセット毎のエントリを有する。各エントリは、チャンクセットの識別子であるチャンクセット番号（#）701と、いずれのチャンクセットにも割り当てられていない論理ページを示す論理アドレスの一覧である空きアドレスリスト702とを有する。

[0076] 図17は、メッセージ1102のグルーピング処理及び書込処理を説明する図である。

メッセージ1102のグルーピング処理及び書込処理は、書込処理部314により実行される。頂点1101の処理に際し入出力されるメッセージ1102には、宛先となる頂点（宛先頂点）を識別するための宛先頂点IDが含まれている。書込処理部314は、メッセージに含まれる宛先頂点IDからグループIDと、グループ内のローカル頂点IDを算出する。例えば、書込処理部314は、宛先頂点IDを頂点グループ数で割ることで、グループIDとローカル頂点IDとを算出する。宛先頂点IDからグループID及びローカル頂点IDを算出する方法は、前述の方法に限られない。

[0077] 主記憶122は、頂点グループ1103のメッセージ1102を格納するためのバッファ1502を有する。例えば、バッファ1502は、主記憶122のワーキングエリア229に設けられる。バッファ1502は、グループIDで示される頂点グループ毎に設けられる。例えば、Gr. 2が頂点グループのグループIDである。メッセージのグルーピング処理において、書込処理部314は、算出したグループIDで指定されるバッファ1502にメッセージ1102を格納する。

[0078] バッファ1502が一杯になった場合、書込処理部314は、書込処理を実行する。書込処理において、書込処理部314は、バッファ1502内のメッセージ1102を不揮発性メモリデバイス101に送信する。具体的には、例えば、書込処理部314は、現アドレステーブル223を参照し、算出したグループID401に対応するアドレスリスト402に含まれる論理アドレスを指定した書込指示を、不揮発性メモリデバイス101に送信する。不揮発性メモリデバイス101のコントローラ112は、チャック情報を参照し、指定された論理アドレスが示す論理ページに対応する物理ページにメッセージ1102を格納する。不揮発性メモリデバイス101に書き込みが行われた後、書き込み処理部314は書き込みを行ったバッファ1502の中身を空にする。

[0079] なお、前アドレステーブル 224 は、宛先頂点に対するメッセージを読み出すためのテーブルである。

[0080] メッセージのグルーピング処理により、メッセージ 1102 を宛先頂点グループ毎に纏めることができる。前述の通り、各アドレステーブル 223 には、頂点グループ 1103 と、これに対応するチャンクセット内の論理ページを示す論理アドレスとが対応づけられている。このため、頂点グループ 1103 内の各頂点を宛先とするメッセージ 1102 を、その頂点グループ 1103 を含む頂点セットに割り当てられたチャンクセットに格納することができる。これにより、無効化単位であるチャンクセットに、頂点セットのメッセージを纏めて格納できる。

[0081] 従って、複数の頂点セット内の頂点宛のメッセージが 1 つの物理ブロック内に混在して格納されることを防ぎ、1 つの頂点セット内の全頂点宛のメッセージを消去することができる。

[0082] 図 18 は、メッセージ 1102 の読出処理を説明する図である。

メッセージの読出処理は、読出処理部 315 により実行される。読出処理部 315 は、頂点の実行に必要なメッセージを主記憶 122 のワーキングエリア 229 に読み出す。具体的には、例えば、読出処理部 315 は、前アドレステーブル 224 を参照し、頂点グループ 1103 に対応する論理アドレスを取得する。

[0083] 読出処理部 315 は、取得した論理アドレスを指定したリードコマンドを不揮発性メモリデバイス 101 に送信する。例えば、頂点グループ #2 の各頂点 1101 を実行に必要なメッセージ 1102 を主記憶 122 に読み出す場合、読出処理部 315 は、前アドレステーブル 224 を参照し、論理アドレス A～E を指定したリードコマンドを不揮発性メモリデバイス 101 に送信する。

[0084] 頂点セット内の全ての頂点グループ 1103 の読出処理が終了した後、計算処理順制御部 311 は、全アドレステーブル 224 を参照し、その頂点セットに含まれる全ての頂点グループ 1103 に対応する論理アドレスを指定

したトリムコマンドを不揮発性メモリデバイス101に送信する。トリムコマンドは、不要となったデータが格納された論理アドレスを通知するコマンドである。本実施例のホスト102は、1つのスーパーステップにおける頂点セット毎に不要となったメッセージの論理アドレスを、トリムコマンドにより不揮発性メモリデバイス101へ通知する。

[0085] リードコマンドを受信したコントローラ112は、チャンク情報に基づき、指定された論理アドレスに対応する物理ページに格納されたメッセージ1102を読み出しホスト102に送信する。そして、トリムコマンドを受信したコントローラ112は、指定された論理アドレスの対象物理ページとの割当を解除し、対象物理ページを無効化する。

[0086] 読出処理により、頂点グループ1103の頂点の処理に必要なメッセージ1102を纏めて読み出すことができる。また、頂点セット内の全ての頂点グループ1103の読出処理が終了した後、頂点セットに対応するチャンクセット毎に、読み出したメッセージ1102が格納された物理ページを無効化できる。前述の通り、1つの頂点セットが使用するメッセージが、1つのチャンクセットに纏めて格納されている。このため、1つのチャンクセット内のメッセージが纏めて無効化される。従って、GCの回数を低減することができる。

[0087] なお、スーパーステップにおいて、頂点の処理が全て終了すれば、前アドレステーブル224の内容は不要となる。スーパーステップにおける全ての頂点の処理が終了したときに、書込処理部314は、前アドレステーブル224の内容を破棄し、現アドレステーブル223の内容を前アドレステーブル224にコピーする。

[0088] 図19は、本実施例にかかる書込範囲制御処理のフローチャートを説明する。

[0089] この処理は、BSPグラフ処理の開始を契機にホスト102が実行する処理である。このとき、BSPグラフ処理に必要な頂点1101を含む初期データを、不揮発性メモリデバイス101とは別の図示しない記憶デバイス（

例えば、SSD、HDD ; Hard disk driveなど) に格納されている。

[0090] 計算処理順制御部311は、初期化処理を実行する(S901)。初期化処理の詳細は、後述する。初期化処理により、BSPグラフ処理に使用する論理アドレス範囲が設定され、この論理アドレス範囲に基づき利用可能なチャンクグループ960、及びチャンクグループ960内のチャンクセット数が算出される。そして、チャンクセットの論理アドレス範囲が頂点セットに割り当てられ、チャンクセット割当テーブル225が生成される

[0091] 計算処理順制御部311は、BSPグラフ処理に使用する各頂点の初期値である初期データを1または複数のチャンクセットに纏めて格納し、空きチャンクセットリスト226を更新する(S902)。例えば、計算処理順制御部311は、各チャンクセットに対応する論理ページの論理アドレスを算出し、その論理アドレスを指定したリードコマンドを不揮発性メモリデバイス101に送信する。この場合、例えば、計算処理順制御部311は、初期データを頂点グループ1103毎に纏めてチャンクセットに格納してもよい。

[0092] これにより、不揮発メモリ101のコントローラ112は、初期データを受信し、チャンクセットに対応する物理ページ933に初期データを格納する。

[0093] スーパーステップにおいて実行される頂点セット内の各頂点グループに対し、計算処理順制御部311は、頂点グループ処理(S903~S906)を繰り返す。以下頂点グループ処理を説明する。この説明において、頂点グループ処理の対象となる頂点グループを対象頂点グループという。

[0094] 計算処理順制御部311は、前アドレステーブル224を参照し、対象頂点グループの各頂点を宛先頂点とするメッセージを主記憶122に読み出す読出処理を実行する(S903)。読出処理は、図18で説明した通りである。具体的には、例えば、計算処理順制御部311は、前アドレステーブル224を参照し、対象頂点グループの頂点グループID401に対応するアドレスリスト402に含まれる論理アドレスを取得する。そして、計算処理

順制御部 311 は、取得した論理アドレスを指定したリードコマンドを不揮発性メモリデバイス 101 に送信する。計算処理順制御部 311 は、取得した論理アドレスを、チャンクセット空きリストテーブル 227 の空きアドレスリスト 702 に追加する。

[0095] なお、読出処理により、不揮発性メモリデバイス 101 のコントローラ 112 は、チャンク情報に基づきリードコマンドで指定された論理アドレスに対応する物理ページに格納されたメッセージ 1102 をホスト 102 へ送信する。計算処理順制御部 311 は、メッセージ 1102 を受信し、受信したメッセージ 1102 を主記憶 122 に格納する。

[0096] 計算処理順制御部 311 は、計算処理部 312 に処理を引き渡す。計算処理部 312 は、頂点 1101 の処理を実行する (S904)。計算処理部 312 は、各頂点 1101 の処理の実行により発生したメッセージ 1102 をデータ通信処理部 313 に引き渡す。データ通信処理部 313 は、メッセージ 1102 を書込処理部 314 に引き渡す (S911)。

[0097] 計算処理順制御部 311 は、頂点セットにおいて、対象頂点グループが、この処理を実行する最後のグループか否かを判定する (S905)。最後の頂点グループでない場合 (S905; No)、計算処理順制御部 311 は、S903 に処理を戻す。

[0098] 一方、最後の頂点グループの場合 (S905; Yes)、計算処理順制御部 311 は、前アドレステーブル 224 を参照し、各対象頂点グループの頂点グループ ID 401 に対応するアドレスリスト 402 から S903 で頂点グループ毎に読み出したメッセージ 1102 が格納される論理アドレスを取得する。そして、計算処理順制御部 311 は、取得した論理アドレスを指定したトリムコマンドを不揮発性メモリデバイス 101 に送信する。計算処理順制御部 311 は、取得したチャンクセット番号 (#) を、空きチャンクセットリスト 226 に追加し、S907 に処理を進める。

[0099] コントローラ 112 は、トリムコマンドで指定された論理アドレスに対応する物理ページを無効化する。

- [0100] 書込処理部314は、データ通信処理部313を介して、計算処理順制御部311からメッセージを受信する(S911)。
- [0101] 書込処理部314は、受信したメッセージ1102を、宛先頂点毎にグルーピングするメッセージのグルーピング処理を実行する(S912)。なお、グルーピング処理は、図17において説明した通りである。メッセージのグルーピング処理により、頂点グループ毎のバッファ1502に、その頂点グループ頂点を宛先としたメッセージが格納される。
- [0102] 書込処理部314は、ワーキングエリア229内にメッセージ1102で一杯になるバッファ領域1502があるか(バッファフル)否かを判定する(S913)。メッセージ1102で一杯になるバッファ領域1502がある場合(S913; Yes)、書込処理部314は、書込処理を行い(S914)、S915に処理を進める。書込処理の概要は図17において説明した通りである。書込処理の詳細は後述する。
- [0103] 一方、メッセージ1102で一杯になるバッファ領域1502がない場合(S913; No)、書込処理部314は、S915に処理を進める。
- [0104] 書込処理部314は、スーパーステップが終了したか否かを判定する(S915)。スーパーステップが終了していない場合(S915; No)、書込処理部314は、S911に処理を戻す。一方、スーパーステップが終了した場合(S915; Yes)、書込処理部314は、スーパーステップが終了したことを計算処理順制御部311に通知し、待機する。
- [0105] 計算処理順制御部311は、書込処理部314から通知を受信し、かつ、頂点セット内の複数の頂点グループ1103の頂点グループ処理がすべて終了したときに、S907に処理を進める。
- [0106] 計算処理順制御部311は、BSPグラフ処理が終了したか否かを判定する(S907)。BSPグラフ処理が終了していないと判定した場合(S907; No)、計算処理順制御部311は、チャンクセットと頂点グループの割当を外すようチャンクセット割当テーブル225を更新し(S908)、前アドレステーブル224の情報を破棄して、現アドレステーブル223

のデータを、前アドレステーブル 224 にコピーし、S903 に処理を戻す。これにより、次のスーパーステップが開始される。

[0107] 書込範囲制御処理により、メッセージの書込範囲を制御することで、不揮発性メモリデバイス 101 の性能の低下を防止できる。具体的には、頂点の処理に使用された後に不要となるメッセージを、宛先の頂点セット毎にチャンクセットに纏めて格納することで、各頂点セットを宛先とするメッセージ群が格納される物理ブロックの範囲を限定することができる。これにより、BSP グラフ処理などのように非常に多くの頂点の処理を実行する場合であっても、不揮発性メモリデバイス 101 の GC が多発するのを防止できる。従って、不揮発性メモリの寿命の低下および不揮発性メモリのアクセス性能の低下を防止できる。

[0108] また、メッセージ 1102 は、そのメッセージ 1102 の宛先（入力先）の頂点 1101 の処理が終了すれば不要となる。頂点セットを設定することで、このように、あるタイミングで不要となるメッセージ 1102 を、同じチャンクセットに纏めて格納することができる。

[0109] また、頂点セット内の全ての頂点グループの読出処理が終了した後、無効化指示（トリムコマンド）を不揮発性メモリデバイス 101 に送信することで、頂点セットに対応するチャンクセットに割り当てられた物理ブロックを纏めて消去することができる。従って、物理ブロック内のデータの断片化を防止することができ、GC の回数を低減することができ、不揮発性メモリの寿命の低下を防止することができる。

[0110] また、チャンクセットを、ハードウェア構成に基づく複数のチャンクで構成することで、例えば、不揮発性メモリデバイス 101 内の全てのバスにそれぞれ 1 つずつ接続されたメモリチップ 111 に対応する複数のチャンクで構成することで、不揮発性メモリデバイス 101 に対するデータの読出し及び書込みを並列に処理することができる。従って、不揮発性メモリデバイス 101 へのホスト 102 のアクセス性能を向上させることができる。

[0111] また、宛先頂点を頂点グループ 1103 毎に纏めることで、主記憶 122

のワーキングエリア 229 に一度に読み出されるメッセージ 1102 を限定することができる。

[0112] 図 20 は、初期化処理のフローチャートである。

初期化処理は、書込範囲制御処理における S901 の処理である。

[0113] ユーザがユーザコードプログラム 221 を介して、BSP グラフ処理で使
用したい不揮発性メモリデバイス 101 のアドレス範囲が指定され、管理ミ
ドルウェアが、使用する論理アドレス範囲（論理アドレスサイズ、又は開始
論理アドレスー終了論理アドレス）を決定したとき、フレームワークプログ
ラム 222 の計算処理順制御部 311 が起動する（S801）。

[0114] 計算処理順制御部 311 は、決定した論理アドレス範囲、及び、不揮発性
メモリデバイス 101 から受信したチャンク情報に基づき、BSP グラフ処
理に利用可能なチャンク 920 の範囲としてチャンクグループ 960 を決定
する（S802）。例えば、計算処理順制御部 311 は、利用可能なチャン
クグループ 960 のチャンクグループ番号、チャンクグループ 960 内の各
チャンク 920 のチャンク番号（縦チャンク番号及び横チャンク番号）、チ
ャンクグループ 960 内のチャンク数などを算出してもよい。また、計算処
理順制御部 311 は、チャンク情報に基づきから、チャンク内の論理ページ
とメモリチップ内の物理ページとを対応づける論理物理変換情報を生成記憶
してもよい。

[0115] 計算処理順制御部 311 は、S802 の算出結果及びチャンク情報に基づ
き、チャンクセット数を算出する（S803）。例えば、計算処理順制御部
311 は、チャンクグループ 960 内のチャンク数を横チャンク数（不揮発
性メモリデバイス 101 の DMA の数）で割った値をチャンクセット数とし
て算出する。

[0116] 計算処理順制御部 311 は、チャンクセット数に合わせて頂点セット数を
決定する（S804）。頂点セット数の算出方法は、図 12 で説明したとお
りである。計算処理順制御部 311 は、決定した数の頂点セットを作成する
。

[0117] 初期化処理により、不揮発性メモリデバイス101から送信されるチャンク情報に基づき、ホスト102が、チャンクセット数等を算出し、各チャンクセットに対応する頂点セット生成できる。

[0118] 図21は、書込処理のフローチャートである。

書込処理は、書込範囲制御処理におけるS914の処理である。以下、書込対象の頂点グループを、書込頂点グループという。

[0119] 書込処理部314は、書込頂点グループの頂点グループIDから、書込頂点グループが属する頂点セット（書込頂点セット）の頂点セットIDを算出する（S1002）。

[0120] 書込処理部314は、チャンクセット割当テーブル225を参照し、（a）書込頂点セットに対応するチャンクセット（書込チャンクセット）が設定され、かつ、（b）書込チャンクセットに空きアドレスが存在するか否かを判定する。書込処理部314は、（a）かつ（b）の場合、S1007に処理を進める。一方、書込処理部314は、（a）又は（b）の少なくともいずれかの要件を満たさない場合、S1004に処理を進める。

[0121] 書込処理部314は、空きチャンクセットリスト226を参照し、空きチャンクセットが存在するか否かを判定する（S1004）。空きチャンクセットが存在する場合（S1004；Yes）、書込頂点セットにいずれかの空きチャンクセットを設定し、チャンクセット割当テーブルを更新する（S1006）。そして、書込処理部314は、S1007に処理を進める。

[0122] 一方、空きチャンクが存在しない場合（S1004；No）、書込処理部314は、チャンクセット割当テーブル225及びチャンクセット空きテーブル227を参照し、書込頂点セットの頂点セットとは別の頂点セットに対応するチャンクセットであって、かつ空きアドレスがあるチャンクセットを取得する（S1005）。なお、このステップにおいて取得可能なチャンクセットが複数ある場合、書込処理部314は、書込頂点セットのIDに近いIDの頂点セットに対応するチャンクセットを取得するとしてもよい。書込処理部314は、書込頂点セットに取得したチャンクセットを設定し（S1

006)、S1007に処理を進める。

- [0123] 書込処理部314は、チャンクセット空きリストテーブル227を参照し、設定したチャンクセットに対応する空きアドレス（論理アドレス）を取得し、空きアドレスリストを更新する（S1007）。具体的には、書込処理部314は、取得した論理アドレスを空きアドレスリスト226から削除する。
- [0124] 書込処理部314は、書込頂点グループの各頂点を宛先とするメッセージに対応する論理アドレスを、不揮発性メモリデバイス101に送信する（S1008）。
- [0125] 不揮発性メモリデバイス101のコントローラ112は、論理アドレスを受信し、チャンク情報に基づき受信した論理アドレスに対応する物理ページにメッセージを格納する。
- [0126] 書込処理により、書込頂点グループが属する頂点セット（書込頂点セット）に割り当てられたチャンクセットにメッセージを格納することができる。
- [0127] すでに、書込頂点セットがチャンクセットに割り当てられている場合であって、そのチャンクセットに空きがある場合には（a and b）、そのチャンクセットにメッセージを格納できる。
- [0128] また、書込頂点セットがチャンクセットに割り当てられていないか、書込頂点セットがチャンクセットに割り当てられている場合であっても、そのチャンクセットに空きアドレスが割り当てられていない場合には、頂点セットが割り当てられていない別の空きチャンクセットに、書込頂点セットを割り当て、そのチャンクセットにメッセージを格納できる。
- [0129] これらの処理により、適切なチャンクセットにメッセージを格納することができる。
- [0130] 以上、一実施例を説明したが、これは本発明の説明のための例示であって、本発明の範囲をこの実施例にのみ限定する趣旨ではない。本発明は、他の種々の形態でも実施することが可能である。
- [0131] 本実施例では、ホスト102が、1つの頂点セットに1つのチャンクセッ

トを割り当て、1つのスーパーステップで1つの頂点セット内の複数の頂点グループを宛先とし、入力されるメッセージを、1つのチャンクセットに書き込む。しかし、この態様に限られない。

[0132] 例えば、ホスト102は、1つの頂点グループに1つのチャンクを割り当て、1つのスーパーステップで1つの頂点グループ1103を宛先として入力されるメッセージ1102を、1つのチャンク920へ書き込んでもよい。この場合、ホスト102は、チャンク情報を用いて、頂点グループ番号に対応するチャンク番号を特定し、チャンク番号に対応する論理アドレス範囲へメッセージを書き込む。これにより、より小さい単位で書込範囲制御処理を実行する事ができる。

[0133] なお、ストレージ装置は、不揮発性メモリデバイス101に対応し、プロセッサは、CPU121に対応し、記憶デバイスは主記憶122に対応する。また、第1限定論理領域及び第2限定論理領域は、それぞれ異なるチャンクセット又はそれぞれ異なるチャンクに対応する。また、限定情報は、チャンク情報に対応し、特定コマンドは、トリムコマンドに対応する。また、処理データはメッセージに対応する。また、限定論理領域情報は、チャンクセット割当テーブル225、空きチャンクセットリスト226、チャンクセット空きリストテーブル227等に対応する。

符号の説明

[0134] 101…不揮発性半導体 102…ホスト計算機 121…CPU 112…コントローラ 111…メモリチップ

請求の範囲

[請求項1]

ストレージ装置と、前記ストレージ装置に接続されるホスト計算機と、を備え、

前記ストレージ装置は、不揮発性メモリと、前記不揮発性メモリ及び前記ホスト計算機に接続されるコントローラと、を含み、

前記ホスト計算機は、記憶デバイスと、前記記憶デバイス及び前記ストレージ装置に接続されるプロセッサと、を含み、

前記コントローラは、前記不揮発性メモリの論理記憶領域の一部である第1限定論理領域のサイズに関する限定情報を記憶し、前記限定情報に基づいて、前記不揮発性メモリの物理記憶領域の一部である第1物理記憶領域を、前記第1限定論理領域に関連付け、

前記コントローラは、前記限定情報を、前記ホスト計算機に送信し、

前記プロセッサは、第1処理を実行し、前記第1処理により第1処理データを生成し、

前記プロセッサは、前記限定情報に基づき、前記第1限定論理領域を示す第1論理アドレス範囲を算出し、

前記プロセッサは、前記第1論理アドレス範囲内の第1論理アドレスを指定した第1ライトコマンドと前記第1処理データとを前記ストレージ装置へ送信し、

前記コントローラは、前記第1ライトコマンドに応じて、前記第1物理記憶領域内の第1物理アドレスを、前記第1論理アドレスに割当て、前記第1物理アドレスにより示された領域に前記第1処理データを書き込み

前記プロセッサは、前記第1処理の後、前記第1論理アドレスを指定したリードコマンドを前記ストレージ装置へ送信し、

前記コントローラは、前記リードコマンドに応じて、前記第1物理アドレスにより示された領域から前記第1処理データを読み出し、

前記プロセッサは、前記第1論理アドレス範囲が不要であることを示す特定コマンドを前記ストレージ装置へ送信し、

前記コントローラは、前記特定コマンドの受信の後に、前記第1処理データを消去する、
計算機システム。

[請求項2]

請求項1に記載の計算機システムであって、

前記限定情報は、さらに前記第1限定論理領域と異なる第2限定論理領域のサイズに関する情報を含み、

前記コントローラは、前記限定情報に基づいて、前記第1物理記憶領域と異なる第2物理記憶領域を、前記第2限定論理領域に関連付け、

前記プロセッサは、前記第1処理データを用いて第2処理を実行し、前記第2処理により第2処理データを生成し、

前記プロセッサは、前記限定情報に基づき、前記第2限定論理領域を示す第2論理アドレス範囲を算出し、

前記プロセッサは、前記第2論理アドレス範囲内の第2論理アドレスを指定した第2ライトコマンドと前記第2処理データとを前記ストレージ装置へ送信し、

前記コントローラは、前記第2ライトコマンドに応じて前記第2物理記憶領域内の第2物理アドレスを、前記第2論理アドレスに割り当て、

前記コントローラは、前記第2物理アドレスより示された領域に前記第2処理データを書き込む
計算機システム。

[請求項3]

請求項2に記載の計算機システムであって、

前記コントローラは、各限定論理領域内に複数の論理ページを割り当て、

前記限定情報は、前記論理ページのサイズ、前記論理記憶領域内の

論理ページ数、前記論理記憶領域内の限定論理領域数、を含む、
計算機システム。

[請求項4] 請求項3に記載の計算機システムであって、
前記不揮発性メモリは、複数のメモリチップを有し、
各メモリチップは、複数のブロックを含み、
各メモリチップは、複数のバスの何れか一つを介して前記コントローラに接続され、
前記第1 限定論理領域及び前記第2 限定領域のそれぞれは、前記複数のメモリチップのうち、前記複数のバスに接続されたメモリチップ群に割り当てられる、
計算機システム。

[請求項5] 請求項4に記載の計算機システムであって、
前記第1 処理は、B S P グラフ処理における第1 スーパーステップであり、
前記第2 処理は、前記B S P グラフ処理における第2 スーパーステップであり、
前記プロセッサは、前記第1 処理において複数の頂点を実行して前記第1 処理データを生成し、前記第2 処理において、前記第1 処理データを用いて前記複数の頂点を実行し、前記第2 処理データを生成し、
前記プロセッサは、前記複数の頂点の一部である第1 頂点グループを前記第1 限定論理領域に対応づけ、
前記プロセッサは、前記第1 処理において生成された前記第1 処理データのうち、前記第1 頂点グループの各頂点を宛先とする第1 グループ処理データを、前記第1 論理アドレス範囲を指定したライトコマンドとともに送信する、
計算機システム。

[請求項6] 請求項5に記載の計算機システムであって、

前記プロセッサは、限定論理領域の数が頂点グループの数の2倍よりも少ない場合に、前記第1頂点グループと、前記複数の頂点のうち前記第1頂点グループ以外の一部である第2頂点グループとを、頂点セットとして定義し、

前記プロセッサは、前記頂点セットを前記第1限定論理領域に対応づけ、

前記プロセッサは、前記第1処理において生成された前記第1処理データのうち、前記頂点セットの各頂点を宛先とする前記第1処理データを、前記第1論理アドレス範囲を指定したライトコマンドとともに送信する、

計算機システム。

[請求項7] 請求項6に記載の計算機システムであって、

前記プロセッサは、前記限定情報に基づき、頂点セットと限定論理領域との対応づけと、処理データが書き込まれていない限定論理領域を示す論理アドレス範囲である空きアドレス範囲と、前記処理データが書き込まれていない限定論理領域との対応づけと、頂点セットに対応づけられていない限定論理領域と、を示す、限定論理領域情報を前記記憶デバイスに記憶し、

前記プロセッサは、前記限定論理領域情報に基づき、前記第1限定論理領域に対応する前記第1論理アドレス範囲を算出する計算機システム。

[請求項8] 請求項7に記載の計算機システムであって、

前記プロセッサは、前記第1グループ処理データを前記記憶デバイスに格納し、

前記プロセッサは、前記記憶デバイス内に前記第1グループ処理データが揃ったときに、前記限定論理領域情報に基づき、前記頂点セットに空きアドレス範囲を有する限定論理領域が対応づけられているかを判定し、

前記プロセッサは、前記頂点セットに空きアドレス範囲を有する限定論理領域が対応づけられている場合に、前記空きアドレス範囲を取得し、

前記プロセッサは、前記空きアドレス範囲を指定したライトコマンドと前記第1頂点処理データとを前記ストレージ装置に送信する、
計算機システム。

[請求項9] 請求項8に記載の計算機システムであって、

前記プロセッサは、前記頂点セットに限定論理領域が対応づけられていないか、又は、前記頂点セットに対応する限定論理領域に空きアドレス範囲が存在しない場合に、前記限定論理領域情報を参照し、頂点セットに対応づけられていない空き限定論理領域が存在するか否かを判定し、

前記プロセッサは、前記空き限定論理領域が存在する場合に、前記空き限定論理領域に頂点セットを対応づけ、

前記プロセッサは、前記空き限定論理領域に対応する論理アドレスを指定したライトコマンドと前記第1グループ処理データとを前記ストレージ装置に送信する
計算機システム。

[請求項10] 請求項9に記載の計算機システムであって、

前記不揮発性メモリは、NAND型のフラッシュメモリである、
計算機システム。

[請求項11] 請求項4に記載の計算機システムであって、

前記プロセッサは、前記第1限定論理領域及び前記第2限定論理領域のそれぞれを、前記複数のメモリチップのうち、前記複数のバスに接続されたメモリチップ群の一部に割り当てる、
計算機システム。

[請求項12] 請求項に記載の計算機システムであって、

前記プロセッサは、前記第1限定論理領域及び前記第2限定論理領

域のそれぞれを、各メモリチップに割り当てる、
計算機システム。

[請求項13] 不揮発性メモリを有するストレージ装置に接続されるホスト計算機
が実行する、前記ストレージ装置への書込範囲を制御する方法であっ
て、

前記ストレージ装置は、前記不揮発性メモリの論理記憶領域の一部
である第1限定論理領域のサイズに関する限定情報を記憶し、前記限
定情報に基づいて、前記不揮発性メモリの物理記憶領域の一部である
第1物理記憶領域を、前記第1限定論理領域に関連付け、前記限定情
報を、前記ホスト計算機に送信し、

前記ホスト計算機は、

第1処理を実行し、前記第1処理により第1処理データを生成し
、

前記限定情報に基づき、前記第1限定論理領域を示す第1論理ア
ドレス範囲を算出し、

前記第1論理アドレス範囲内の第1論理アドレスを指定した第1
ライトコマンドと前記第1処理データとを前記ストレージ装置へ送信
し、

前記第1処理の後、前記第1論理アドレスを指定したリードコマ
ンドを前記ストレージ装置へ送信し、

前記第1論理アドレス範囲が不要であることを示す特定コマンド
を前記ストレージ装置へ送信する
方法。

[請求項14] 不揮発性メモリを有するストレージ装置に接続されるホスト計算機
が実行する、前記ストレージ装置への書込範囲を制御する方法であっ
て、

前記ストレージ装置は、前記不揮発性メモリの論理記憶領域の一部
である第1限定論理領域のサイズに関する限定情報を記憶し、前記限

定情報に基づいて、前記不揮発性メモリの物理記憶領域の一部である第1物理記憶領域を、前記第1限定論理領域に関連付け、前記限定情報を、前記ホスト計算機に送信し、

前記ホスト計算機は、

第1処理を実行し、前記第1処理により第1処理データを生成し、

前記限定情報に基づき、前記第1限定論理領域を示す第1論理アドレス範囲を算出し、

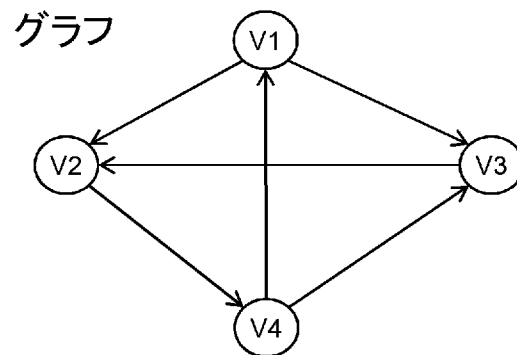
前記第1論理アドレス範囲内の第1論理アドレスを指定した第1ライトコマンドと前記第1処理データとを前記ストレージ装置へ送信し、

前記第1処理の後、前記第1論理アドレスを指定したリードコマンドを前記ストレージ装置へ送信し、

前記第1論理アドレス範囲が不要であることを示す特定コマンドを前記ストレージ装置へ送信するホスト計算機。

[図1]

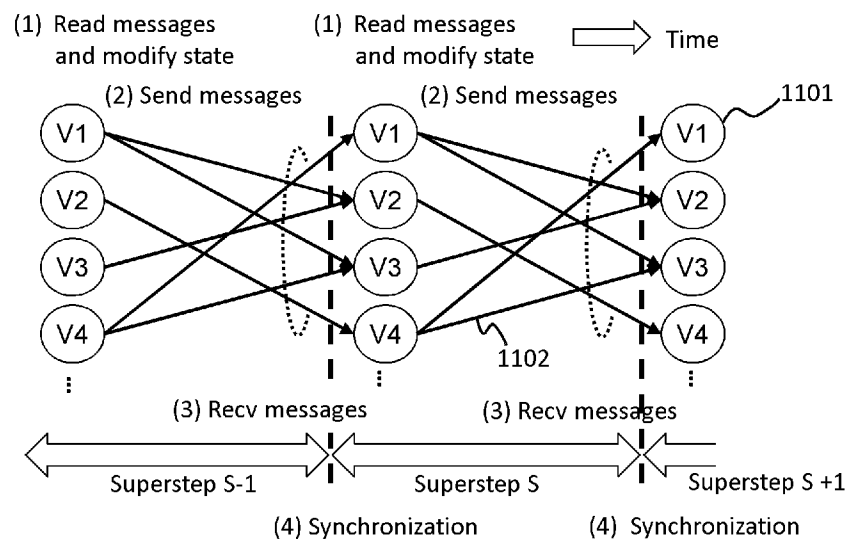
Fig. 1



[図2]

Fig. 2

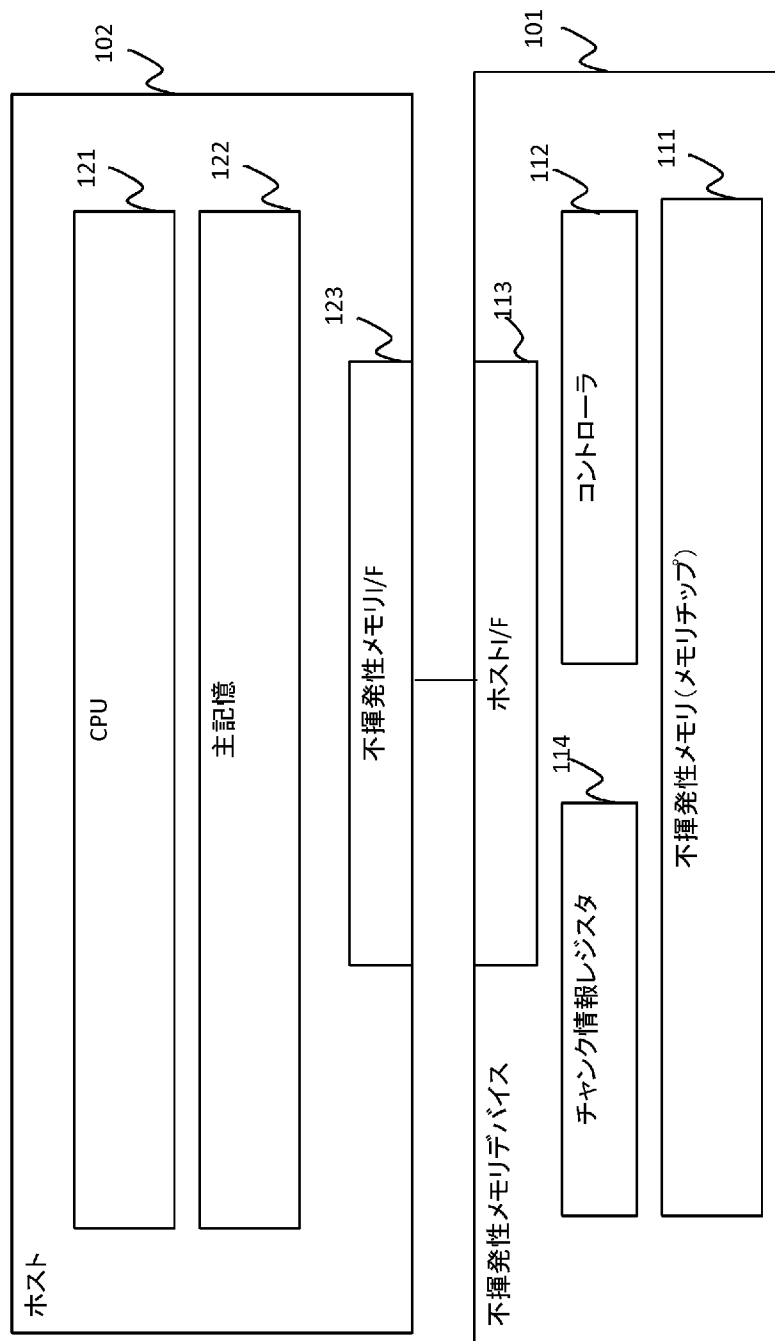
BSPグラフ処理



[図3]

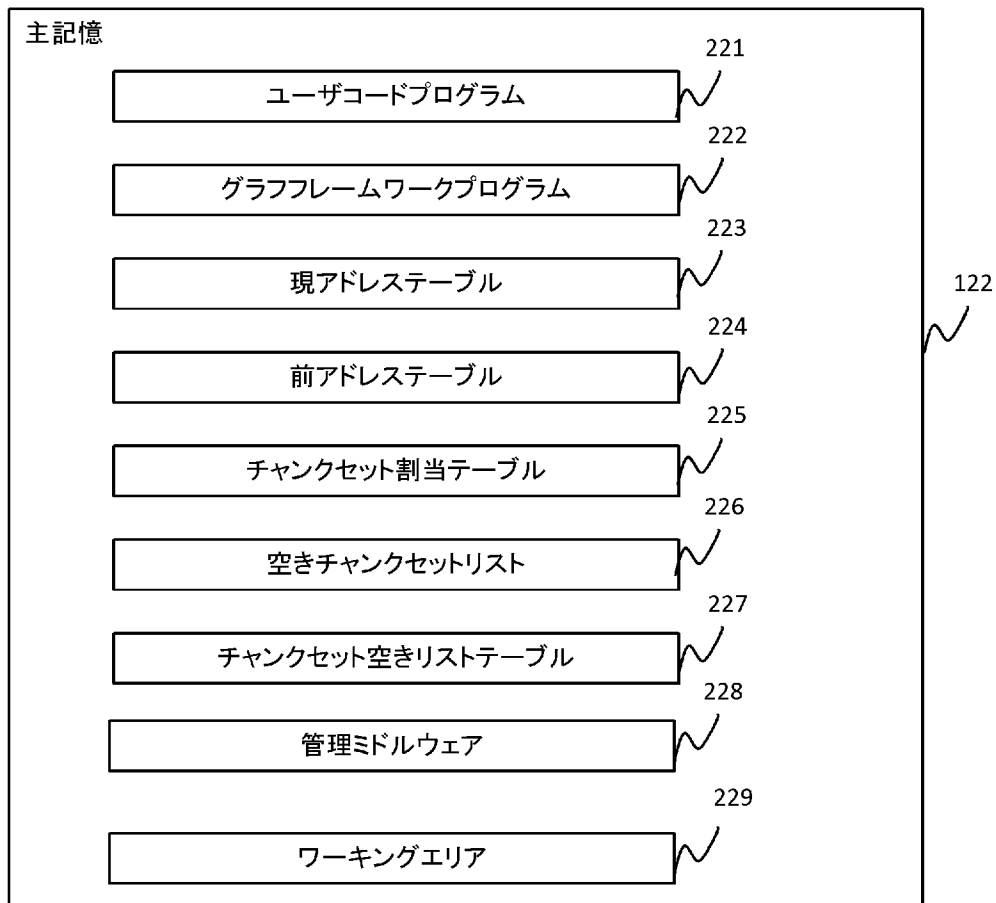
Fig. 3

構成図



[図4]

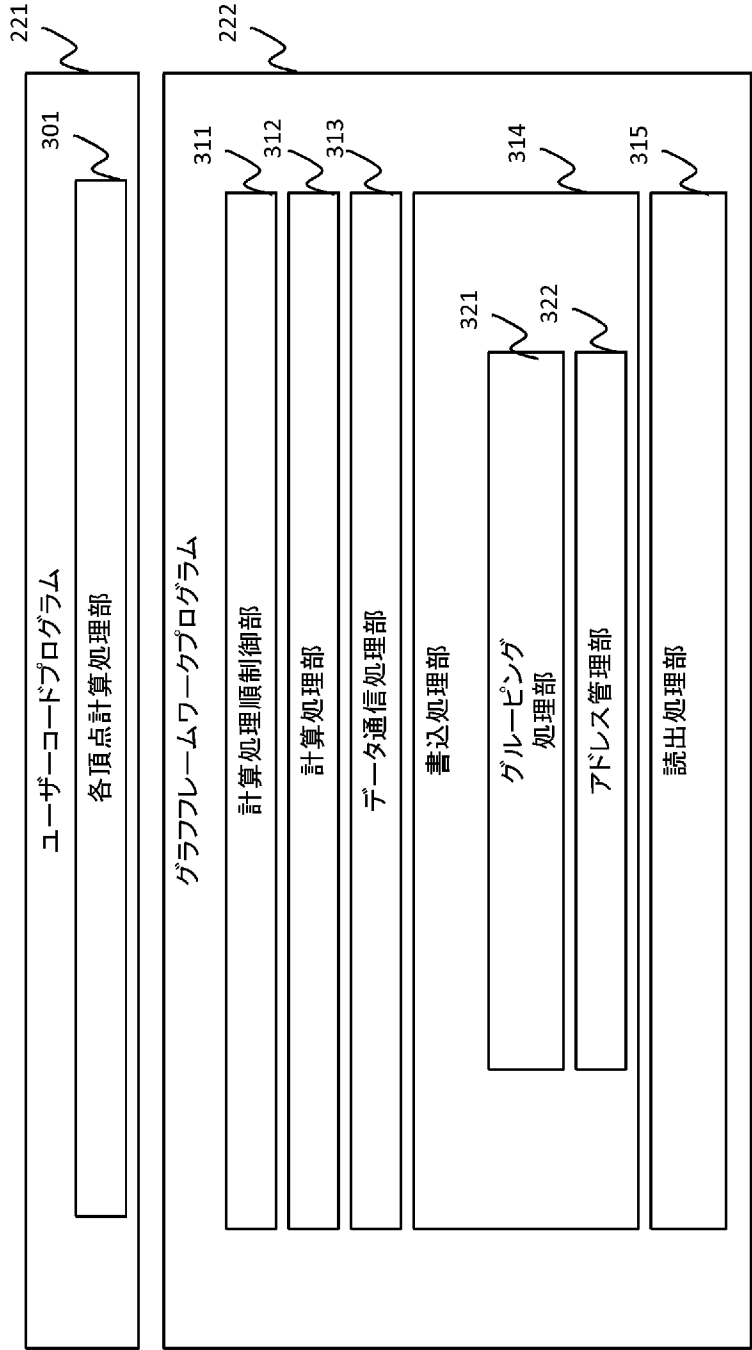
Fig. 4



[図5]

Fig. 5

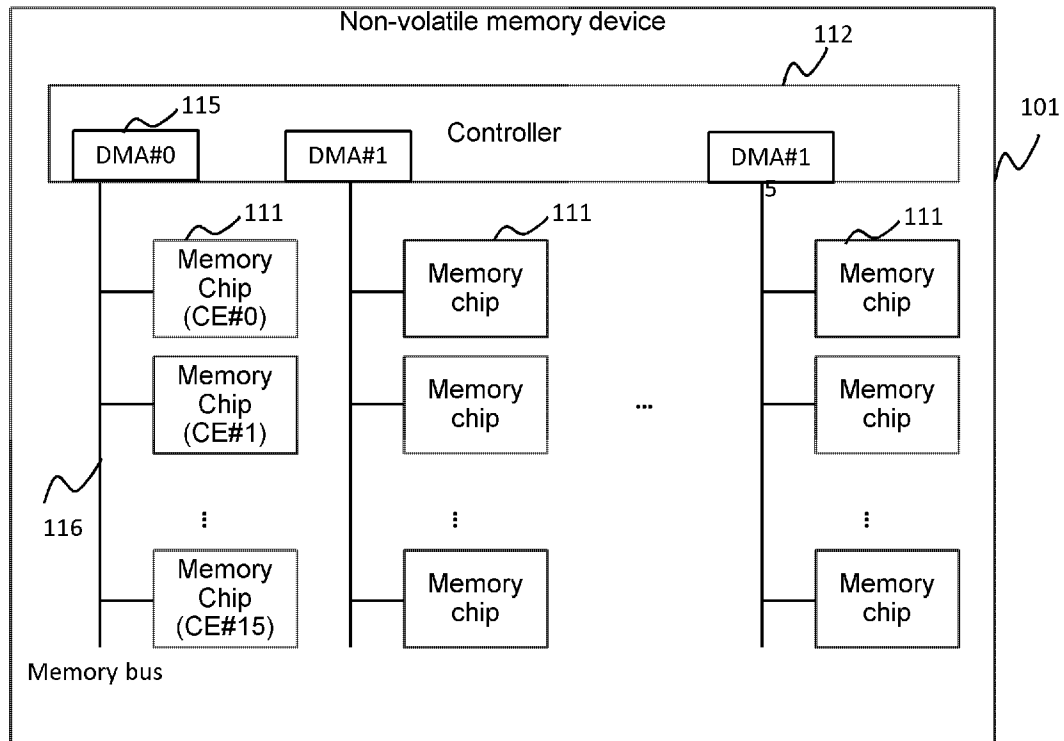
機能図



[図6]

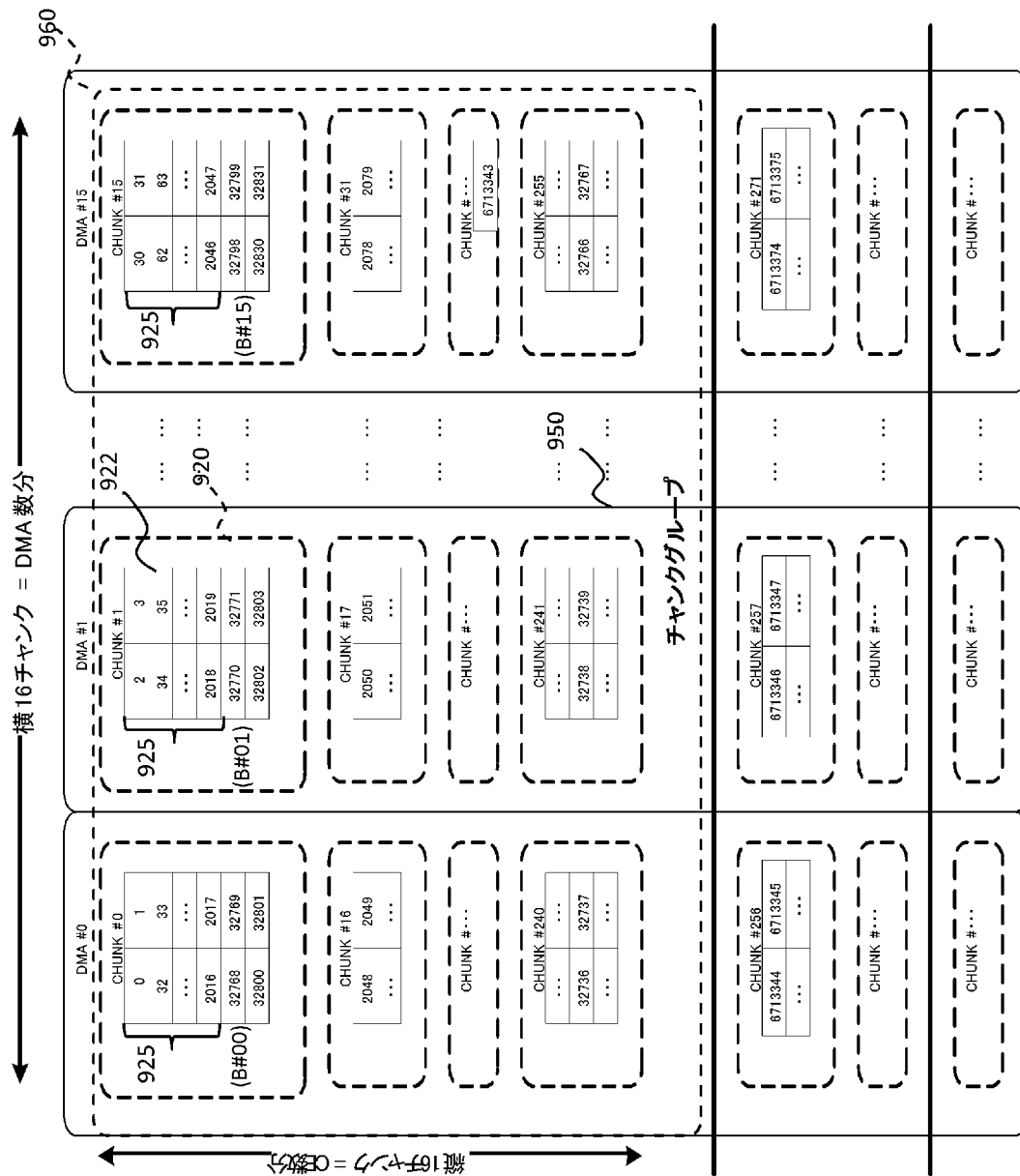
Fig. 6

不揮発性メモリ構成



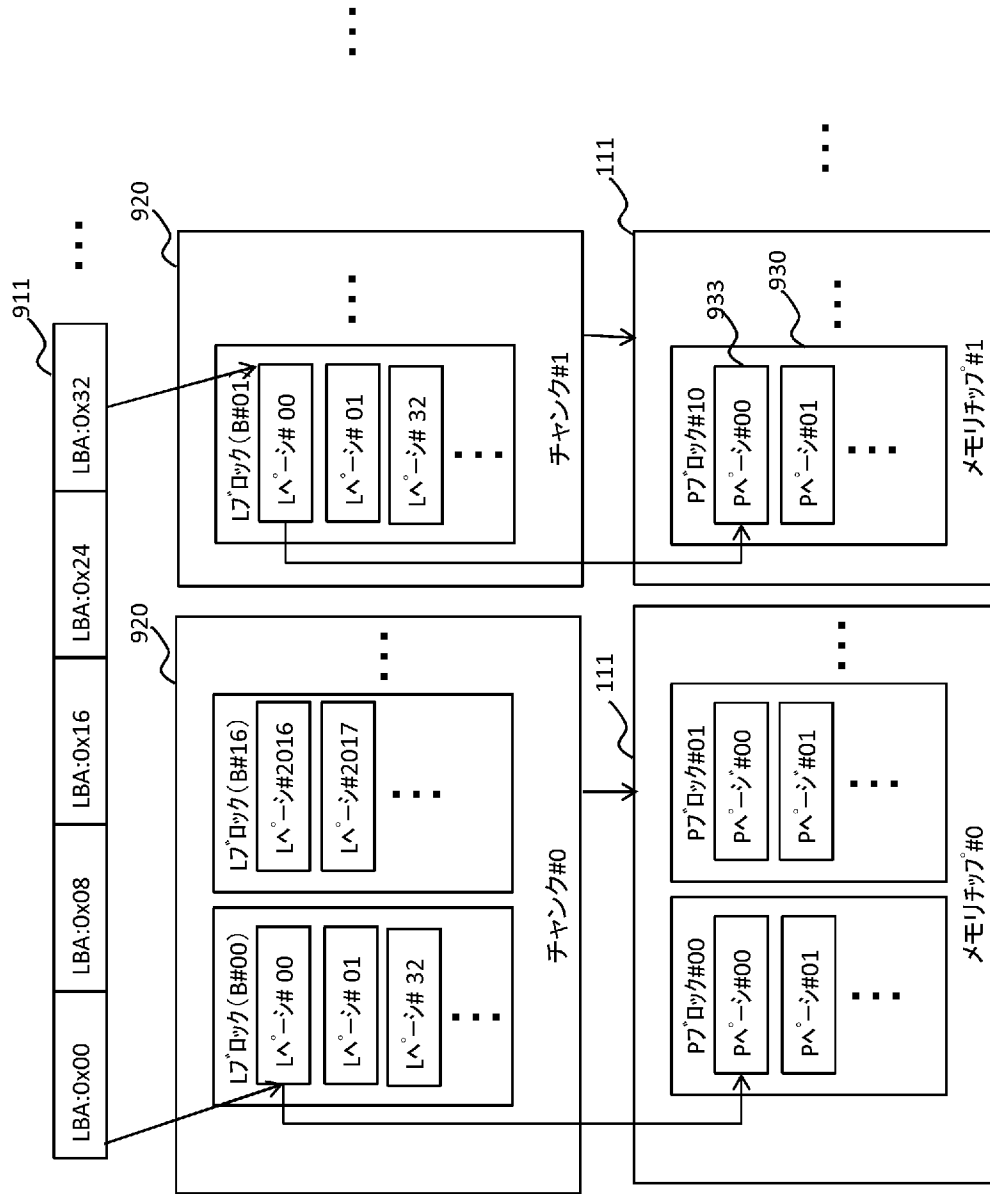
[図7]

Fig. 7



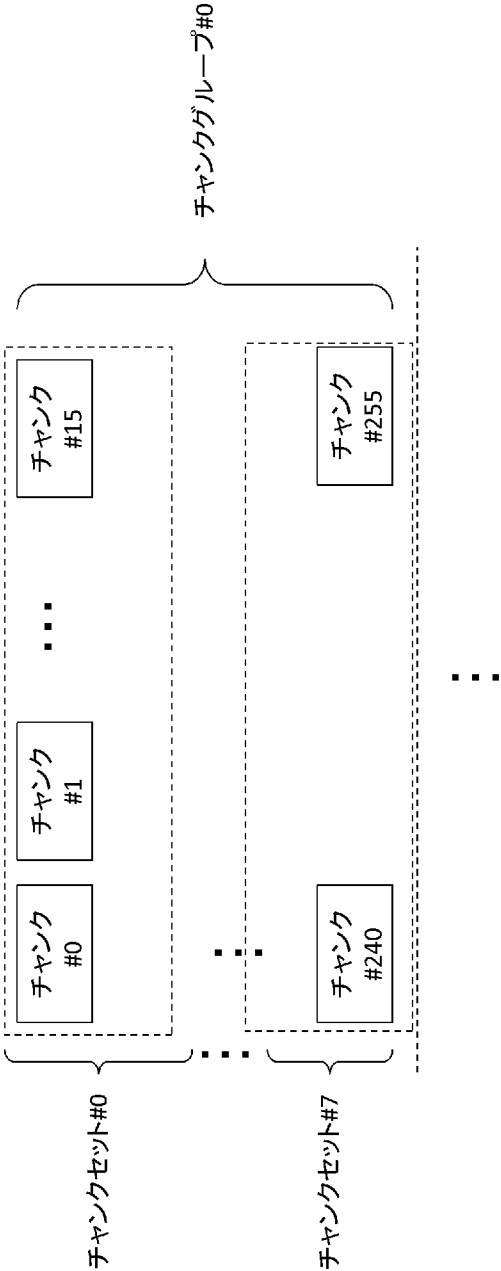
[図8]

Fig. 8



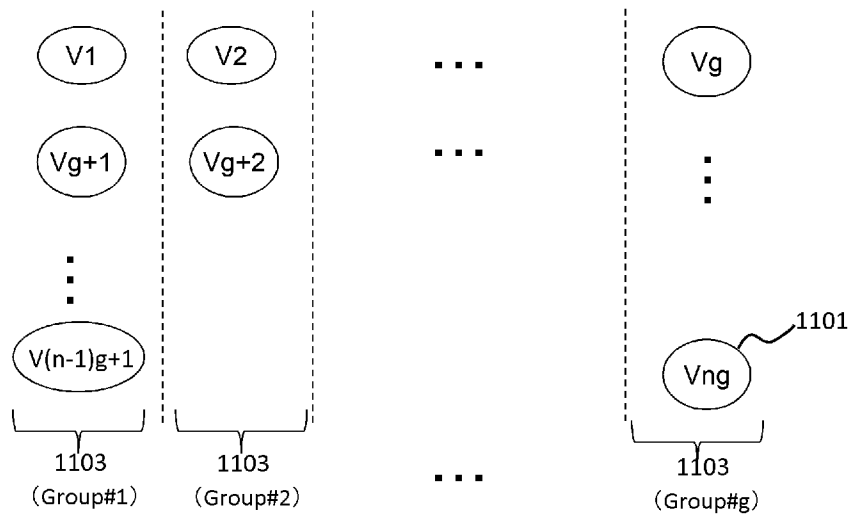
[図9]

Fig. 9



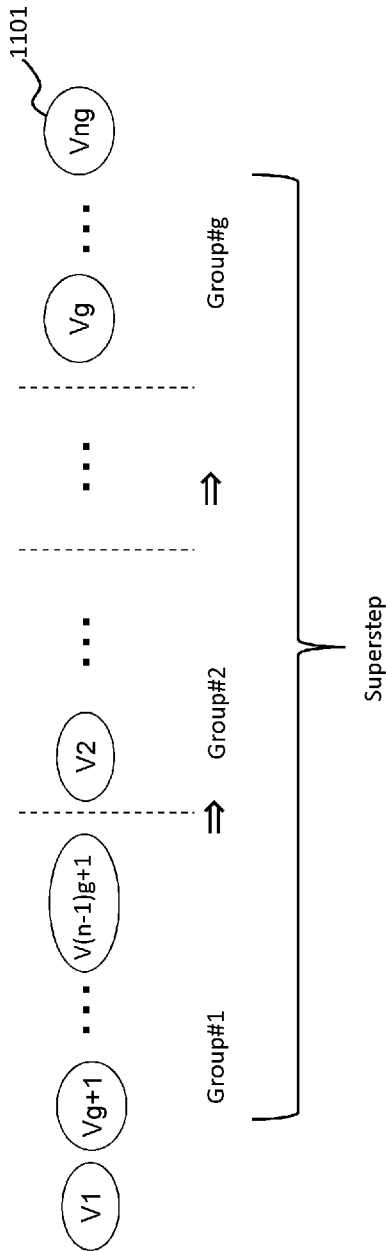
[図10]

Fig. 10



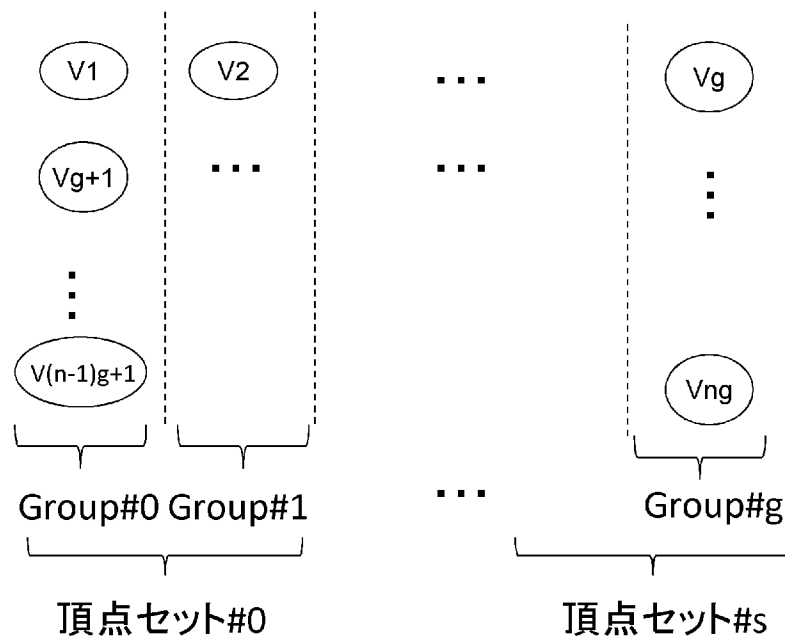
[図11]

Fig. 11



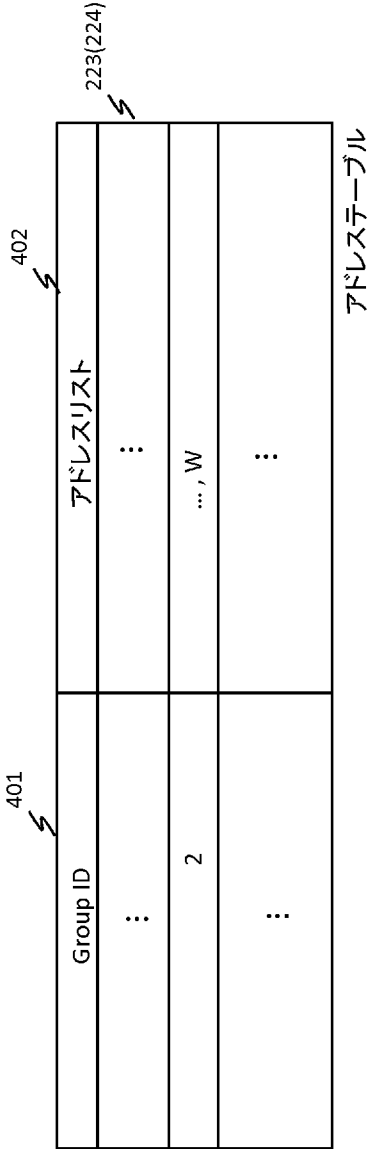
[図12]

Fig. 12



[図13]

Fig. 13



[図14]

Fig. 14

501

502

225

頂点セット#	チャンクセット#
1	
2	
⋮	

チャンクセット割当テーブル

[図15]

Fig. 15

226

空きチャンクセットリスト
1, 2,

[図16]

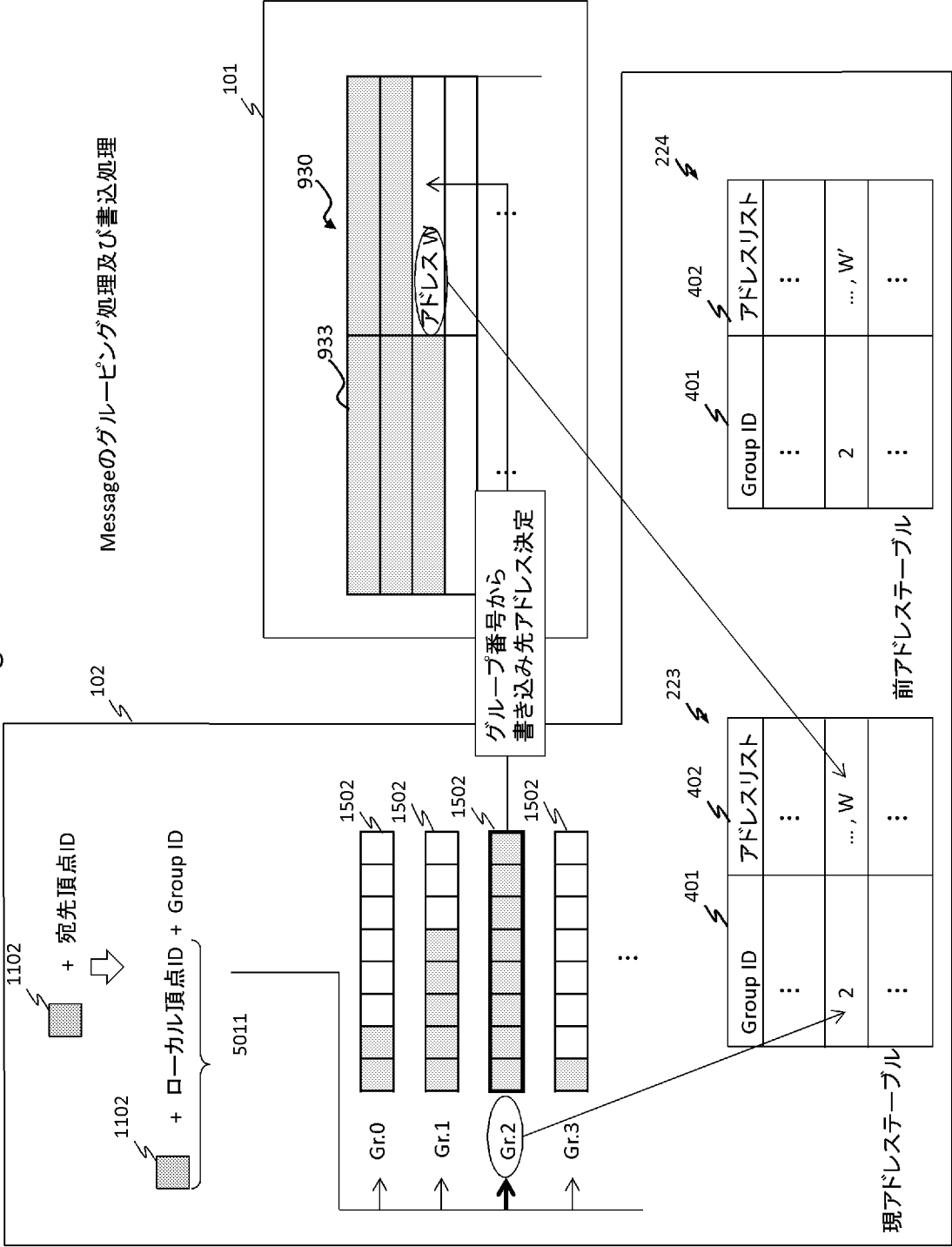
Fig. 16

チャンクセット#	空きアドレスリスト
1	
2	

チャンクセット空きリストテーブル

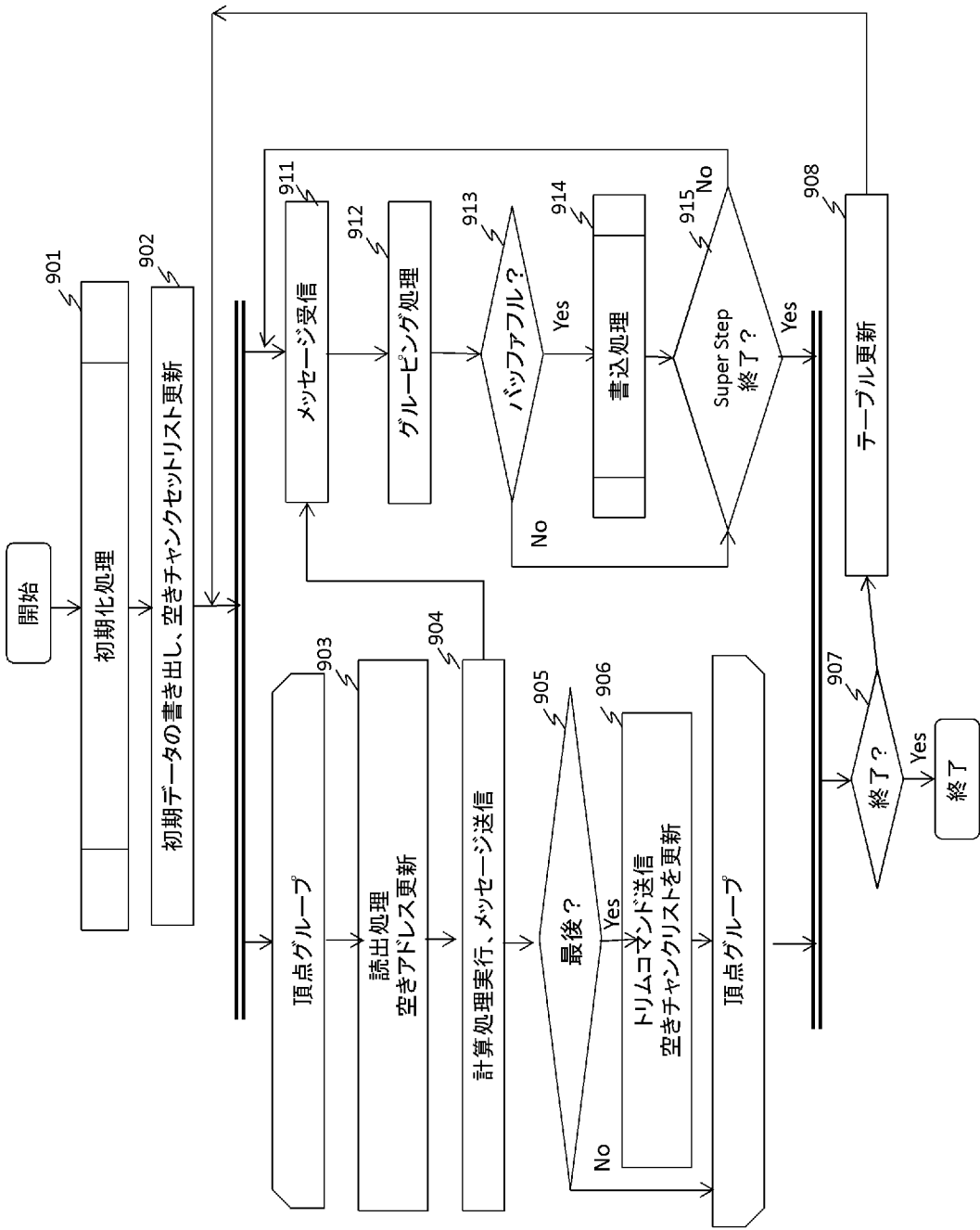
[図17]

Fig. 17



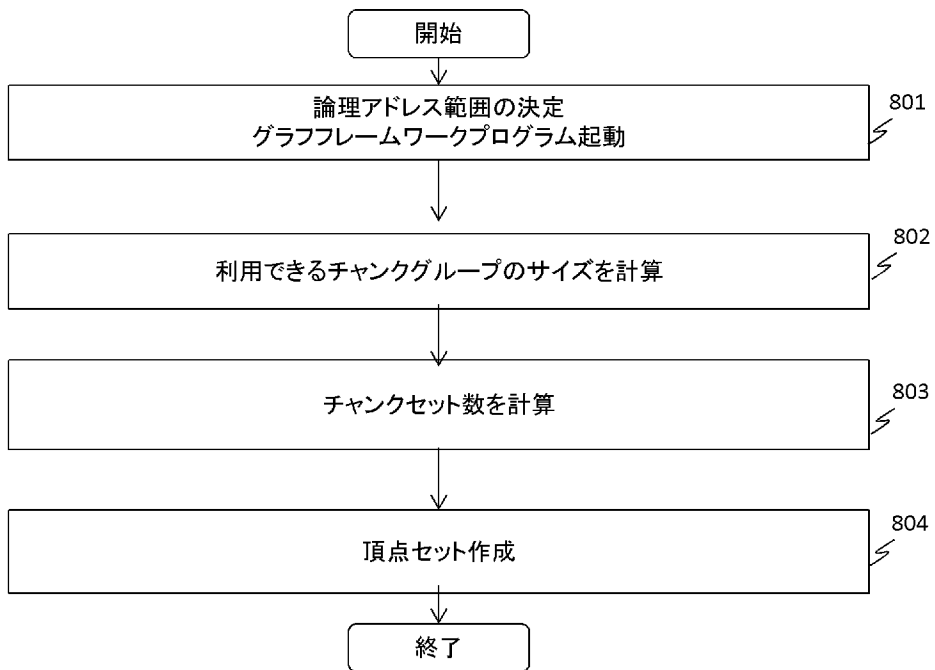
[図19]

Fig. 19



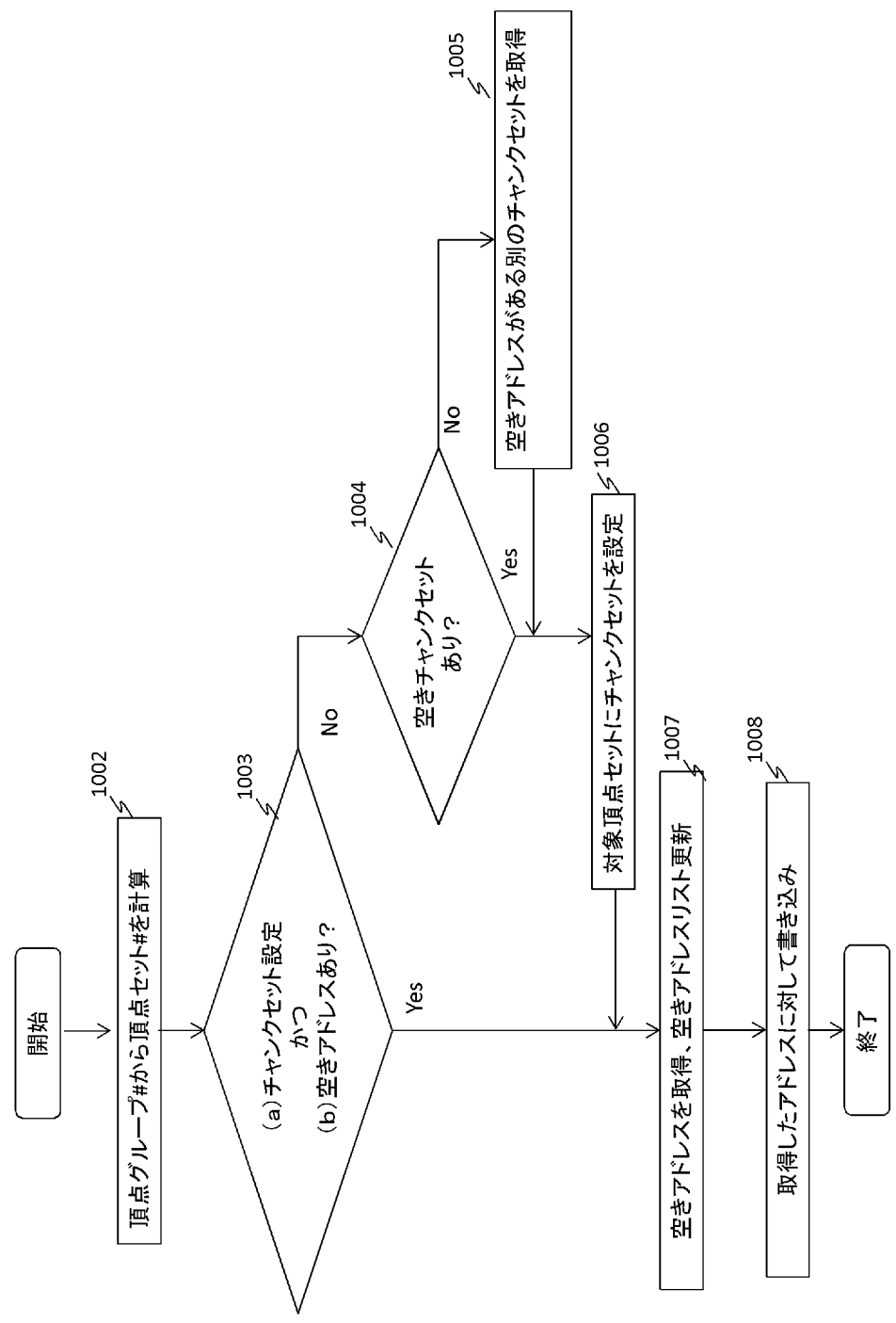
[図20]

Fig. 20



[図21]

Fig. 21



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2015/058100

A. CLASSIFICATION OF SUBJECT MATTER

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

G06F12/00, G06F3/06, G06F12/02

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	1996-2015
Kokai Jitsuyo Shinan Koho	1971-2015	Toroku Jitsuyo Shinan Koho	1994-2015

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

JSTPlus(JDreamIII), IEEE Xplore, THE ACM DIGITAL LIBRARY

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X Y	WO 2015/008358 A1 (Hitachi, Ltd.), 22 January 2015 (22.01.2015), entire text; all drawings; particularly, paragraph [0038]; claims 1 to 3 (Family: none)	1, 13, 14 2-12
Y	WO 2014/196055 A1 (Hitachi, Ltd.), 11 December 2014 (11.12.2014), entire text; all drawings; particularly, paragraphs [0007] to [0009], [0039] (Family: none)	2-12
A	US 6763519 B1 (William F. MCCOLL et al.), 13 July 2004 (13.07.2004), entire text; all drawings (Family: none)	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
18 May 2015 (18.05.15)

Date of mailing of the international search report
26 May 2015 (26.05.15)

Name and mailing address of the ISA/
Japan Patent Office
3-4-3, Kasumigaseki, Chiyoda-ku,
Tokyo 100-8915, Japan

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2015/058100

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	Chen CHEN et al., "The Application of the BSP Model on DataGrid", Services Computing, 2004 (SCC 2004), Proceedings of the 2004 IEEE International Conference on, IEEE, 2004.09.15, pp.471~474	1-14

A. 発明の属する分野の分類（国際特許分類（I P C））

Int.Cl. G06F12/00(2006.01)i, G06F3/06(2006.01)i, G06F12/02(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. G06F12/00, G06F3/06, G06F12/02

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	1 9 2 2 - 1 9 9 6 年
日本国公開実用新案公報	1 9 7 1 - 2 0 1 5 年
日本国実用新案登録公報	1 9 9 6 - 2 0 1 5 年
日本国登録実用新案公報	1 9 9 4 - 2 0 1 5 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）

JSTPlus(JDreamIII), IEEE Xplore, THE ACM DIGITAL LIBRARY

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 Y	WO 2015/008358 A1（株式会社日立製作所）2015.01.22, 全文・全図、特に段落[0038], [請求項 1]-[請求項 3]（ファミリーなし）	1, 13, 14 2-12
Y	WO 2014/196055 A1（株式会社日立製作所）2014.12.11, 全文・全図、特に段落[0007]-[0009], [0039]（ファミリーなし）	2-12
A	US 6763519 B1（William F. MCCOLL ほか 3 名）2004.07.13, 全文・全図（ファミリーなし）	1-14

☒ 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

1 8 . 0 5 . 2 0 1 5

国際調査報告の発送日

2 6 . 0 5 . 2 0 1 5

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）

池田 聡史

5 B

9 4 7 5

電話番号 0 3 - 3 5 8 1 - 1 1 0 1 内線 3 5 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
A	Chen CHEN ほか1名, "The Application of the BSP Model on DataGrid", Services Computing, 2004 (SCC 2004), Proceedings of the 2004 IEEE International Conference on, IEEE, 2004.09.15, pp. 471～474	1-14