

(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2月5日(05.02.2015)



(10) 国際公開番号
WO 2015/015589 A1

- (51) 国際特許分類:
G06F 11/34 (2006.01) G06F 12/00 (2006.01)
- (21) 国際出願番号: PCT/JP2013/070698
- (22) 国際出願日: 2013年7月31日(31.07.2013)
- (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: 富士通株式会社 (FUJITSU LIMITED) [JP/JP]; 〒2118588 神奈川県川崎市中原区上小田中4丁目1番1号 Kanagawa (JP).
- (72) 発明者: ▲高▼見浩之 (TAKAMI, Hiroyuki); 〒2118588 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内 Kanagawa (JP).
- (74) 代理人: 大菅義之 (OSUGA, Yoshiyuki); 〒1020084 東京都千代田区二番町8番地20 二番町ビル3F 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, 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, 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.

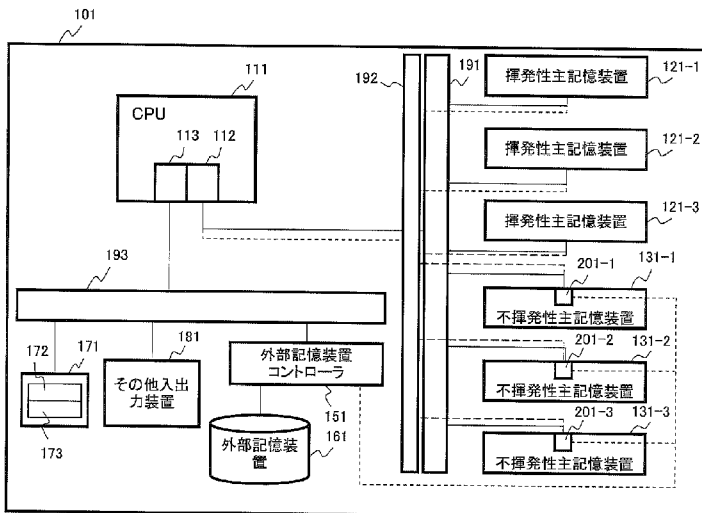
- (84) 指定国 (表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), ユーロピア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, 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: INFORMATION PROCESSING DEVICE, MEMORY DUMP METHOD, AND MEMORY DUMP PROGRAM

(54) 発明の名称: 情報処理装置、メモリダンプ方法、およびメモリダンププログラム



121-1, 121-2, 121-3 Main volatile storage device
 131-1, 131-2, 131-3 Main non-volatile storage device
 151 External storage device controller
 161 External storage device
 181 Other input/output device

(57) Abstract: An information processing device is provided with: a processing unit that runs an operating system; a main non-volatile storage device that can be directly accessed by the processing unit and that comprises a controller; and an external storage device that cannot be directly accessed by the processing unit. The information processing device is characterized in that when the processing unit detects an error in the operating system, the processing unit resets except for the main non-volatile storage device and restarts the operating system, and the controller writes the data of the main non-volatile storage device to the external storage device.

(57) 要約: オペレーティングシステムを実行する処理部と、前記処理部が直接アクセス可能かつコントローラを有する不揮発性主記憶装置と、前記処理部が直接アクセス可能でない外部記憶装置と、を備え、前記処理部が前記オペレーティングシステムのエラーを検出したとき、前記処理部は、前記不揮発性主記憶装置以外をリセットして、前記オペレーティングシステムを再起動し、前記コントローラは、前記不揮発性主記憶装置のデータ

を前記外部記憶装置に書き込むことを特徴とする情報処理装置。

WO 2015/015589 A1

明 細 書

発明の名称：

情報処理装置、メモリダンプ方法、およびメモリダンププログラム

技術分野

[0001] 本発明は、情報処理装置、メモリダンプ方法、およびメモリダンププログラムに関する。

背景技術

[0002] 近年、UNIX（登録商標）サーバ、IAサーバが基幹システムに導入されるようになり、UNIX（登録商標）サーバ、IAサーバの高可用性が重要視されている。一般的に、システムに致命的なエラー（クラッシュ事象）が発生した場合はシステムを緊急停止（パニック）させて、その原因を調査するためにメモリダンプをディスクに保存している。

[0003] システムを停止している間は、システムを使用できないので、速やかにシステムを再起動して業務停止時間を短縮することが重要な要件となる。

[0004] しかし、近年では、実装メモリの容量がテラバイト（TB）オーダのサーバが登場し、このようなシステムでは、メモリダンプを採取するのに非常に時間がかかり、速やかにシステムを再起動することができなくなっている。

先行技術文献

特許文献

[0005] 特許文献1：特開2004-102395号公報

特許文献2：特開平4-182748号公報

特許文献3：特開平10-333944号公報

発明の概要

発明が解決しようとする課題

[0006] 本発明の課題は、メモリダンプ処理による業務停止時間を短縮することである。

課題を解決するための手段

- [0007] 実施の形態の情報処理装置は、処理部と、不揮発性主記憶装置と、外部記憶装置と、を備える。
- [0008] 前記処理部は、オペレーティングシステムを実行する。
前記不揮発性主記憶装置は、前記処理部が直接アクセス可能且つコントローラを有する。
前記外部記憶装置は、前記処理部が直接アクセス可能でない。
- [0009] 前記処理部は、前記オペレーティングシステムのエラーを検出したとき、前記不揮発性主記憶装置以外をリセットして、前記オペレーティングシステムを再起動する。
- [0010] 前記コントローラは、前記不揮発性主記憶装置のデータを前記外部記憶装置に書き込む。

発明の効果

- [0011] 実施の形態の情報処理装置によれば、メモリダンプ処理による業務停止時間を短縮することができる。

図面の簡単な説明

- [0012] [図1]第1の実施の形態に係る情報処理装置の構成図である。
[図2]BIOS設定情報の例を示す図である。
[図3]第1の実施の形態に係る不揮発性主記憶装置の構成および接続バスを示す図である。
[図4]記憶素子管理テーブルの例を示す図である。
[図5]EEPROMに記録されている情報を示す図である。
[図6]第1の実施の形態に係るメモリダンプ処理のフローチャートである。
[図7]第1の実施の形態に係るBGダンプ処理の詳細なフローチャートである。
。
[図8]第1の実施の形態に係る割り込みダンプ処理の詳細なフローチャートである。
[図9]第1の実施の形態に係る割り込みダンプ処理を示す図である。

[図10]優先度による段階的なメモリダンプを示す図である。

[図11]第1の実施の形態に係るクラッシュ事象発生直前およびBGダンプフラグ設定直後の記憶素子管理テーブルの状態を示す図である。

[図12]第2の実施の形態に係る不揮発性主記憶装置の状態を示す図である。

[図13]拡張記憶素子管理テーブルの例である。

[図14]第2の実施の形態に係る割り込みダンプ処理の詳細なフローチャートである。

[図15]割り当て変更前と割り当て変更後の拡張記憶素子管理テーブルを示す図である。

[図16]第3の実施の形態に係るクラッシュ事象発生直前およびBGダンプフラグ設定直後の記憶素子管理テーブルの状態を示す図である。

[図17]記憶素子管理テーブルのフラグの状態遷移図である。

[図18A]第4の実施の形態に係るメモリダンプ処理のフローチャートである。

[図18B]第4の実施の形態に係るメモリダンプ処理のフローチャートである。

[図19]第4の実施の形態に係る各処理時のデータの位置を示す図である。

[図20]コンピュータシステムの構成図である。

発明を実施するための形態

[0013] 以下、図面を参照しながら実施の形態について説明する。

図1は、第1の実施の形態に係る情報処理装置の構成図である。

[0014] 情報処理装置101は、例えば、パーソナルコンピュータ(PC)やサーバ等のコンピュータである。

[0015] 情報処理装置101は、Central Processing Unit(CPU)111、揮発性主記憶装置121-i(i=1~3)、不揮発性主記憶装置131-i、外部記憶装置コントローラ151、外部記憶装置161、Basic Input/Output System(BIOS)格納域171、およびその他入出力装置181を備える。

[0016] CPU111は、各種処理を実行するプロセッサ(処理部)である。

CPU111は、BIOS格納域171に格納されているBIOS172

を読み出して実行する。また、CPU 111は、揮発性主記憶装置121または不揮発性主記憶装置141に格納されているプログラムを読み出して実行する。

[0017] 情報処理装置101には、Operating System (OS) がインストールされており、CPU 111は、外部記憶装置161からOSを揮発性主記憶装置121および／または不揮発性主記憶装置141に読み出して実行する。

[0018] CPU 111は、メモリコントローラ112およびI/Oコントローラ113を有する。

[0019] メモリコントローラ112は、揮発性主記憶装置121および不揮発性主記憶装置131に対する読み書きの制御を行う。メモリコントローラ112は、メモリバス191を介して揮発性主記憶装置121および不揮発性主記憶装置131と接続している。揮発性主記憶装置121および不揮発性主記憶装置131のデータは、メモリバス192を介して送信される。

[0020] また、メモリコントローラ112は、Inter-Integrated Circuit (I²C) バス192を介して揮発性主記憶装置121および不揮発性主記憶装置131と接続している。メモリコントローラ112は、I²Cのマスタコントローラを兼ねている。

[0021] I/Oコントローラ113は、入出力装置の制御を行う。I/Oコントローラ113は、Peripheral Component Interconnect (PCI) バス193を介して、BIOS格納域171、外部記憶装置コントローラ151、およびその他入出力装置181と接続している。

[0022] 揮発性主記憶装置121は、揮発性の記憶素子を用いたデータを格納する記憶装置である。揮発性主記憶装置121は、例えば、Dynamic Random Access Memory (DRAM) である。

揮発性主記憶装置121は、CPU 111から直接アクセス可能である。

[0023] 不揮発性主記憶装置131は、不揮発性の記憶素子を用いたデータを格納する記憶装置である。不揮発性主記憶装置131は、例えば、NAND型フラッシュメモリやMagnetoresistive Random Access Memory (MRAM) 等である。

[0024] 不揮発性主記憶装置 131 は、電力が供給されない場合でも、データを保持することができる。

不揮発性主記憶装置 131 は、CPU 111 から直接アクセス可能である。

[0025] 外部記憶装置コントローラ 151 は、外部記憶装置 161 を制御する。外部記憶装置コントローラ 151 は、I/O コントローラ 113、外部記憶装置 161、および NonVolatile Memory (NVM) コントローラ 201 と接続している。

[0026] 外部記憶装置 161 は、データを格納する不揮発性の記憶装置である。

外部記憶装置 161 は、CPU 111 から直接アクセスできない。よって、CPU 111 が外部記憶装置 161 のデータを読み出す場合、該データは、揮発性主記憶装置 121 または不揮発性主記憶装置 131 に転送され、揮発性主記憶装置 121 または不揮発性主記憶装置 131 から読み出される。

[0027] 外部記憶装置 161 は、例えば、ハードディスクドライブ、Solid State Drive (SSD)、または磁気テープ装置等である。外部記憶装置 161 は、揮発性主記憶装置 121 および不揮発性主記憶装置 131 よりもアクセス速度が遅い。

[0028] 外部記憶装置 161 は、補助記憶装置または二次記憶装置とも呼ばれる。

BIOS 格納域 171 は、BIOS Flash Firmware (以下、BIOS と表記する) 172 および BIOS 設定情報 173 を格納する。BIOS 格納域 171 は、例えば、不揮発性のメモリである。

[0029] BIOS 172 は、ハードウェアの初期化やブートローダーの起動等を行うプログラムである。また、BIOS 172 は、ダンプ処理や NVM コントローラの設定等を行う。

[0030] BIOS 172 は、例えば、Unified Extensible Firmware Interface (UEFI) や拡張ファームウェア等である。

[0031] BIOS 172 は、CPU 111 により読み出されて実行される。

BIOS 設定情報 173 は、メモリダンプ処理時に使用する設定値やメモ

リスロットに搭載される主記憶装置（揮発性主記憶装置 1 2 1 および不揮発性主記憶装置 1 3 1）の特性などが記載される情報である。

[0032] 尚、BIOS設定情報 1 7 3の詳細については後述する。

その他入出力装置 1 8 1は、データの入力および／または出力を行う装置である。その他入出力装置 1 8 1は、例えば、Universal Serial Bus (USB) デバイスやPeripheral Component Interconnect (PCI) カード等である。

[0033] 図 2 は、BIOS設定情報の例を示す図である。

BIOS設定情報 1 7 3は、Element、Attribute1、Attribute2、およびValueが対応付けられて記載されている。

[0034] Elementは、対象の装置や設定項目を示す。図 2 において、Elementとして、ダンプ先外部記憶装置、メモリ手動交換、メモリスロット 1～6 が記載されている。

[0035] Attribute1およびAttribute2は、装置の特性を示す。図 2 において、ダンプ先外部記憶装置のAttribute1として、「共有」が記載されている。これは、外部記憶装置が複数の主記憶装置のダンプ先として設定されていることを示す。

[0036] また、図 2 において、メモリスロット 1～3のAttribute1として「揮発性」、メモリスロット 4～6のAttribute1として「不揮発性」が記述されている。これは、メモリスロット 1～3に搭載されている主記憶装置が揮発性、メモリスロット 4～6に搭載されている主記憶装置が不揮発性であることを示す。

[0037] また、図 2 において、メモリスロット 4のAttribute2として「Block (対象外)」、メモリスロット 5、6のAttribute2として「Byte (対象)」が記述されている。これは、メモリスロット 4に搭載されている主記憶装置に対してブロックアクセスを行い、ダンプ対象でないことを示し、メモリスロット 5、6に搭載されている主記憶装置に対してバイトアクセスを行い、ダンプ対象であることを示す。

[0038] 不揮発性主記憶装置 131 へのアクセス方法は以下 2 種類ある。

(1) ブロックアクセス：OS からブロックストレージデバイスとして見える

(2) バイトアクセス：OS から物理メモリ空間として見える

[0039] Value は、設定されている値である。

図 2 において、メモリ手動交換の Value として、N が記載されている。これは、不揮発性主記憶装置 131 の手動交換を行わないことを示す。尚、Y が記載されている場合は、不揮発性主記憶装置 131 の手動交換を行うことを示す。尚、「メモリ手動交換」は、後述の第 4 の実施の形態で利用される。

[0040] 図 2 において、ダンプ先外部記憶装置の Value として、Ctrl=0, Device=2 が記載されている。Ctrl=0, Device=2 は、ダンプ先の外部記憶装置 161 を識別する情報である。

[0041] また、図 2 において、メモリスロット 5 の Value として「Pri=High」、メモリスロット 6 の Value として「Pri=Low」が記載されている。「Pri」は優先度を示し、メモリスロット 5 に搭載されている主記憶装置の優先度が High、メモリスロット 6 に搭載されている主記憶装置の優先度が Low であることを示している。

[0042] 図 3 は、第 1 の実施の形態に係る不揮発性主記憶装置の構成および接続バスを示す図である。

[0043] 不揮発性主記憶装置 131-1 は、NVM コントローラ 201-1、Electrically Erasable Programmable Read-Only Memory (EEPROM) 211、外部記憶装置コントローラ 231、コネクタ 241、および記憶素子 251-k (k=1~8)

[0044] NVM コントローラ 201-1 は、バッファ 202、記憶素子管理テーブル 203、および制御ファームウェア 204 を備える。

[0045] NVM コントローラ 201-1 は OS の動作とは無関係に、ダンプ対象のデータを能動的に外部記憶装置 161 にコピーする。コピーは OS からのアクセス同様、最小アクセス単位はページだが、外部記憶装置 161 とのプロ

トコルが一度の転送命令で送信可能なデータ長の範囲で、複数のページが一括して送信されることで効率化を図る。

[0046] バッファ202は、記憶素子251に書き込むデータまたは記憶素子251から読み出したデータを一時的に格納する記憶装置である。

[0047] 記憶素子管理テーブル203は、不揮発性主記憶装置131-1の各ページの状態が記載されたテーブルである。

尚、記憶素子管理テーブル203の詳細については後述する。

[0048] 実施の形態の情報処理装置101では、仮想記憶方式によりOSが主記憶領域を「ページ」と呼ばれる構造体の単位で分割し、それらを管理するための構造として「ページテーブル」が存在するものとする。OSのページテーブルはメモリ空間全体でひとつ保持され、主記憶領域の一部を利用して格納されるため、主記憶装置毎に持つものではない。一方、OSが管理するページテーブルとは別に、個々の不揮発性主記憶装置131のNVMコントローラ201は、ダンプ管理のために独自のページテーブルを保持するものとする。また、各ページのサイズおよび開始位置はOSが管理するものと一致するものとする。

[0049] 制御ファームウェア204は、後述するBackground (BG) ダンプルーチン処理等を行うプログラムである。制御ファームウェア204は、NVMコントローラ201-1により実行される。

[0050] 尚、不揮発性主記憶装置131-2、131-3の構成は、不揮発性主記憶装置131-1の構成と同様であるため説明は省略する。

[0051] EEPROM211は、不揮発性主記憶装置131-1の特性やメモリダンプ処理の設定などの情報を格納する。尚、EEPROM211に格納される情報の詳細については後述する。

[0052] 外部記憶装置コントローラ231は、コネクタ241を介して、外部記憶装置コントローラ151と接続する。この接続はダンプ用途に限らず、不揮発性主記憶装置の故障に備えたバックアップを、OSに負荷をかけずに実現するなどの用途に流用することも可能である。外部記憶装置コントローラ2

31は、外部記憶装置コントローラ151との間でデータの送受信を行う。外部記憶装置コントローラ231と外部記憶装置コントローラ151との間で通信を行うことで、Direct Memory Access (DMA) コントローラを含むI/Oコントローラ113やPCIバス193に影響が生じなくなる。

[0053] 外部記憶装置161との通信プロトコルはSmall Computer System Interface (SCSI) など任意のものでよい。外部記憶装置コントローラ151、231は単純に通信プロトコルをやりとりする機能だけでなく、ブロックまたはファイルシステム上の書き込み位置を理解可能なレベルのものが実装される。

[0054] 記憶素子251は、データを格納する素子である。

図4は、記憶素子管理テーブルの例を示す図である。

[0055] 記憶素子管理テーブル203には、ページ番号とフラグが対応付けられて記載されている。

[0056] ページ番号は、ページに割り当てられた番号である。尚、以下の説明において、ページ番号nのページをページnと表記する場合がある。

[0057] フラグは、ページの状態を示す情報である。フラグとして、空き領域、使用領域 (Dirty)、使用領域 (Sync)、未ダンプ、またはReservedが記載される。

[0058] 空き領域は、データが書き込まれていない空き領域であることを示す。

使用領域 (Dirty) は、データが書き込まれており、且つ該データは外部記憶装置161へコピーされていない、すなわち未ダンプであることを示す。

[0059] 使用領域 (Sync) は、データが書き込まれており、該データは外部記憶装置161へコピー済み、すなわちダンプ済みであることを示す。

[0060] 未ダンプは、ダンプされていないことを示す。また、未ダンプ(1)、未ダンプ(2)のような未ダンプの後ろに付加されている数字は、ダンプの世代を示す。BGダンプ処理中に再度クラッシュ事象が発生した場合も再度ダンプは可能である。この場合、前回のダンプデータの残りと今回のダンプデータの分離を可能とするために、NVMコントローラ201がBGダンプ処

理を実行する初期にすべての使用済みページを「未ダンプ」と設定する処理で、すでに「未ダンプ」となっているページのフラグを見つけた場合「未ダンプ（１）」に書き換えることで、区別を可能とする。この場合、NVMコントローラ２０１は前回と今回のダンプ先のファイル名を区別可能なようにするなどの処理を行う。同様に前々回の未ダンプデータは「未ダンプ（２）」などとインクリメントすることで複数世代の管理も可能であるが、フラグの実装上どれだけ未アサインの番号が存在するかが世代の上限となる。例えば、フラグが４ビット（１６値）で管理され、４値が既存のフラグとしてアサイン済みであれば、追加で最大１２世代分が管理可能である。

[0061] Reservedは、未使用の（すなわち、外部から見たアドレスが割り当てられていない）予備領域であることを示す。

[0062] 図５は、EEPROMに記録されている情報を示す図である。

EEPROM 211には、主記憶モジュールの種類、主記憶モジュールの容量、ECCの有無、温度センサ値、BGダンプフラグ、およびダンプ先外部記憶装置が記録されている。

[0063] 主記憶モジュールの種類は、主記憶装置の種類を示す。詳細には、主記憶モジュールの種類は、主記憶装置が揮発性か不揮発性であるかを示す。図５は、不揮発性主記憶装置 131-1 のEEPROM 211なので、「不揮発性」が記載されている。

[0064] 主記憶モジュールの容量は、不揮発性主記憶装置 131-1 の容量を示す。

ECCの有無は、不揮発性主記憶装置 131-1 のError Check and Correct (ECC) 機能の有無を示す。

[0065] 温度センサ値は、温度センサから取得した不揮発性主記憶装置 131-1 の温度である。

[0066] BGダンプフラグは、BGダンプ処理を行うか否かを示す。BGダンプフラグは０～２のいずれかが設定される。０は無効（すなわち、BGダンプ処理を行わない、若しくはBGダンプ処理完了）、１は有効（すなわち、BG

ダンプ処理を行う)、2は有効且つWait(すなわち、BGダンプ処理を行うために待機)を示す。

[0067] ダンプ先外部記憶装置は、ダンプ先の外部記憶装置を示す。

尚、揮発性主記憶装置121にも揮発性主記憶装置121の特性(主記憶モジュールの種類、主記憶モジュールの容量、ECCの有無、温度センサ値等)が記録されたEEPROMが搭載されている。

[0068] 図6は、第1の実施の形態に係るメモリダンプ処理のフローチャートである。

先ず、情報処理装置101では、OSが実行されているものとする。

[0069] ステップS501において、情報処理装置101のOSでクラッシュ事象、すなわち致命的なエラーが発生する。

[0070] ステップS502において、CPU111は、ダンプ処理を開始する。

ステップS503において、CPU111は、不揮発性主記憶装置131のBGダンプフラグを設定する。詳細には、CPU111は、BIOS設定情報173を参照し、各不揮発性主記憶装置131の優先度(Pri)を確認する。CPU111は、Pri=Highである不揮発性主記憶装置131のBGダンプフラグを1、Pri=Lowである不揮発性主記憶装置131のBGダンプフラグを2に設定する。

[0071] ステップS504において、NVMコントローラ201は、記憶素子管理テーブル203のフラグを設定する。詳細には、NVMコントローラ201は、「使用領域(Dirty)」を「未ダンプ」に設定する。NVMコントローラ201は、「使用領域(Dirty)」以外のフラグを「空き領域」に設定する。

[0072] ステップS505において、CPU111は、不揮発性主記憶装置131以外の装置(CPU111および揮発性主記憶装置121を含む)をリセットする。

[0073] ステップS506において、CPU111は、OSを起動する。

以下、ステップS507~S510とステップS511が並列に実行される。

- [0074] ステップS507において、Pri=Highである不揮発性主記憶装置131のNVMコントローラ201は、BGダンプ処理を実行する。尚、BGダンプルーチン処理の詳細については後述する。
- [0075] ステップS508において、CPU111は、Pri=Highである全ての不揮発性主記憶装置131のEEPROM211内のBGダンプフラグが0であることを検出する。すなわち、Pri=Highである全ての不揮発性主記憶装置131でBGダンプルーチン処理が完了したことを検出する。
- [0076] ステップS509において、CPU111は、Pri=Lowである不揮発性主記憶装置131のEEPROM211内のBGダンプフラグを1に設定する。
- [0077] ステップS510において、Pri=Lowである不揮発性主記憶装置131のNVMコントローラ201は、BGダンプ処理を実行する。尚、BGダンプルーチン処理の詳細については後述する。
- [0078] ステップS511において、各NVMコントローラ201は、割り込みダンプ処理を行う。尚、割り込みダンプ処理については後述する。
- [0079] 尚、上述のCPU111が実行する各ステップは、CPU111がBIOS172を読み出して実行することにより実現される。
- [0080] 図7は、第1の実施の形態に係るBGダンプ処理の詳細なフローチャートである。
- 図7は、図6のステップS507またはステップS510に対応する。
- [0081] ステップS521において、NVMコントローラ201は、以下のステップS522～S524で処理対象とするページを未選択のページの中から1つ選択する。例えば、NVMコントローラ201は、未選択のページの中から、ページ番号が一番小さいページを選択する。これにより、ページ番号の昇順にページが選択される。以下、選択されたページを選択ページと表記する。
- [0082] ステップS522において、NVMコントローラ201は、記憶素子管理テーブル203を参照し、選択ページのフラグが「未ダンプ」であるかチェックする。選択ページのフラグが「未ダンプ」の場合、制御はステップS5

23に進み、「未ダンプ」でない場合、制御はステップS525に進む。

[0083] ステップS523において、NVMコントローラ201は、選択ページのデータを外部記憶装置161にコピー（送信）する。これにより、選択ページのデータは、外部記憶装置161に書き込まれる。

[0084] ステップS524において、NVMコントローラ201は、記憶素子管理テーブル203の選択ページのフラグを「空き領域」に設定する。

[0085] ステップS525において、未選択のページがある場合、制御はステップS521に戻り、未選択のページがない場合（すなわち、フラグが「未ダンプ」であるページが無い場合）、制御はステップS526に進む。

[0086] ステップS526において、NVMコントローラ201は、EEPROM211のBGダンプフラグを0に設定する。

[0087] 図8は、第1の実施の形態に係る割り込みダンプ処理の詳細なフローチャートである。

図8は、図6のステップS511に対応する。

[0088] ステップS531において、NVMコントローラ201は、OSからの書き込み要求を待つ。OSからの書き込み要求を受信した場合、制御はステップS532に進む。また、BGダンプ処理が終了した場合、割り込みダンプ処理は終了する。

[0089] ステップS532において、NVMコントローラ201は、記憶素子管理テーブル203を参照し、書き込み要求による書込み先のページのフラグが「未ダンプ」であるか否かチェックする。書込み先のページのフラグが「未ダンプ」である場合、制御はステップS533に進み、「未ダンプ」でない場合、制御はステップS534に進む。

[0090] ステップS533において、NVMコントローラ201は、書き込み要求データをバッファ202に書き込み、書き込み完了をOSに報告する（Write Back方式）。そして、NVMコントローラ201は、書き込み要求データの書込み先のページを外部記憶装置161にコピー（送信）し、該ページのフラグを「空き領域」に設定する。

- [0091] ステップS534において、NVMコントローラ201は、書き込み要求データを書込み先のページ（記憶素子251）に書き込む。また、データが書き込まれたページのフラグは、「使用領域（Dirty）」となる。そして、制御はステップS531に戻る。
- [0092] 図9は、第1の実施の形態に係る割り込みダンプ処理を示す図である。
図9では、書き込み先のページが未ダンプである場合の処理を示す。
先ず、NVMコントローラ201は、OSからの書き込み要求を受信する（ステップS541）。
- [0093] NVMコントローラ201は、記憶素子管理テーブル203を参照し、書き込み要求による書き込み先のページ（ページ番号=50）のフラグが「未ダンプ」であるか否かチェックする。ここでは、書き込み先のページのフラグが「未ダンプ」であるので、NVMコントローラ201は、書き込み要求データをバッファ202に書き込み、書き込み完了をOSに報告する（ステップS542）。
- [0094] NVMコントローラ201は、書き込み先のページのデータを外部記憶装置161に書き込むためのキューを生成し、順番待ちになっているキューの先頭に割り込ませる。
- [0095] そして、書き込み先のページのデータが外部記憶装置161に書き込まれた後、NVMコントローラ201は、書き込み要求データを書込み先のページ（記憶素子251）に書き込む。
- [0096] 図9において、BGダンプ処理により、ページを順次ダンプするキューが生成されている。生成されたキューは、生成された順に並んでおり、先頭から順次、外部記憶装置コントローラ151に送信される。
- [0097] 図9においては、ページ番号=6~9のページを外部記憶装置161に書き込むためのキューが順番に並んでいる状態で、ページ番号=50のページを外部記憶装置161に書き込むためのキューがページ番号=6のページのキューの前に割り込んだ状態を示している。このように、ページ番号=50のページを外部記憶装置161に書き込むためのキューが並んでいるキュー

に割り込み、先に実行される。

[0098] 割り込みダンプ処理を用いることにより、OSからの初回のデータ書き込みこそ低速であるが、それ以後の同一ページへの読み書きは通常時の不揮発性主記憶装置と同等の速度が得られる。また、ダンプに割り込みが発生したり、複数の不揮発性主記憶装置から非同期にダンプされることで、ダンプファイルは断片化（フラグメンテーション）が発生する。ダンプファイルが不連続となる箇所には開始アドレスとオフセットが合わせて記録されるようにすることで、ダンプの解析時に正しいアドレス空間に復元することが可能である。

[0099] 図10は、優先度による段階的なメモリダンプを示す図である。

上述のように、図6のメモリダンプ処理では、Pri=Highの不揮発性主記憶装置のBGメモリダンプ処理の後に、Pri=Lowの不揮発性主記憶装置のBGメモリダンプ処理が行われている。

[0100] このような優先度に基づく段階的なBGメモリダンプ処理の意義について説明する。

OSのメモリ管理の特性上、BIOSが決定した物理メモリアドレスにはアクセス順序や頻度に局所性があることが予想される。特にOSの起動にはカーネルが利用する領域を優先的にダンプする必要がある。このような特性をプロファイルとしてBIOS設定情報173に保持し、不揮発性主記憶装置131の単位で優先順位付けをすることで、カーネルが利用する領域を優先的にダンプすることができる。

[0101] 図10では、不揮発性主記憶装置1～Nがあり、不揮発性主記憶装置1、2にカーネル空間、不揮発性主記憶装置3～Nにユーザプロセス空間が割り当てられるものとする。

[0102] よって、不揮発性主記憶装置1、2の優先度はHigh、不揮発性主記憶装置3～Nの優先度はLowに設定される。

[0103] メモリダンプの第1段階では、不揮発性主記憶装置1、2においてBGダンプフラグ=1（有効）に設定され、不揮発性主記憶装置3～NにおいてB

G ダンプフラグ=2（有効：Wait）に設定される。

- [0104] BIOS 172 を実行する CPU 111 は、EEPROM 211 を監視して、優先度が高い不揮発性主記憶装置 1、2 の BG ダンプフラグ=0（無効）となった時点で、優先度が低い不揮発性主記憶装置 3～N に BG ダンプフラグ=1（有効）を設定する（第 2 段階）。
- [0105] 尚、BG ダンプフラグ=2（有効：Wait）の場合、NVM コントローラ 201 は OS からの書き込み要求に応じて割り込みダンプ処理を実施するときだけダンプを実施する。
- [0106] 図 10 において、ダンプ第二段階の主記憶装置 3～N で飛び飛びに網掛けになっている部分が、割り込みダンプ処理によってダンプ済みの領域を示す。
- [0107] 図 11 は、第 1 の実施の形態に係るクラッシュ事象発生直前および BG ダンプフラグ設定直後の記憶素子管理テーブルの状態を示す図である。
- [0108] 図 11 の上側はクラッシュ事象発生直前の記憶素子管理テーブルの状態を示し、下側は BG ダンプフラグ設定直後の記憶素子管理テーブルの状態を示す。
- [0109] 図 11 では、記憶素子管理テーブル 203 の各レコードを記憶素子の物理的な位置（Column, Row）に対応するように配置して記載している。
- [0110] 図 11 では、クラッシュ事象発生直前において、ページ 1～3、6、X-1 のフラグは「使用領域（Dirty）」であり、ページ 4、5、X のフラグは「空き領域」である。
- [0111] 図 6 で述べたように、BG ダンプフラグが設定される（ステップ S503）と NVM コントローラ 201 は記憶素子管理テーブル 203 のフラグを設定する（ステップ S504）。図 11 の下側に示すように、「使用領域（Dirty）」は、「未ダンプ」に設定される。
- [0112] 第 1 の実施の形態の情報処理装置によれば、メモリダンプ処理による業務停止時間を短縮することができる。
- [0113] 第 1 の実施の形態の情報処理装置によれば、不揮発性主記憶装置から外部

記憶装置へのダンプ処理をOS再起動後のバックグラウンド処理とすることで、業務停止時間を短縮することができる。

[0114] (第2の実施の形態)

第2の実施の形態では、不揮発性記憶装置131において、外部から見たアドレスに対して記憶素子が可変的に割り当てられる場合について説明する。

[0115] 尚、第2の実施の形態の情報処理装置の構成は、特に断りのない限り第1の情報処理装置の構成と同様であるため説明は省略する。

[0116] 第2の実施の形態の不揮発性主記憶装置131は、NAND型フラッシュメモリを用いたSSDで用いられているような、外部から見たアドレスに対して記憶素子が可変的に割り当てられているような方式が実装されているものとする。

[0117] このような不揮発性主記憶装置131は、ある記憶素子が保持するデータを直接更新することができず、別の記憶素子に書き込む、またはひとまとまりの記憶素子の列をイレース後に再度書き込むなどの処理を行う。

[0118] また、記憶素子の寿命が書き込み回数に依存するため、特定の記憶素子に書き込みが偏らないよう、不揮発性主記憶装置131は、外部から見て連続したアドレスであっても、内部的には不連続な領域に書き込むウェアレベリングと呼ばれる機能を持つ。このため、不揮発性主記憶装置131は、外部から見たアドレスと、記憶素子の対応を管理するためのテーブルを持つ。第2の実施の形態では、第1の実施の形態の記憶素子管理テーブル203を拡張することで、外部から見たアドレスと記憶素子との対応を管理するテーブルを実現する。

[0119] 実際には記憶素子1つにつき1つのアドレスを割り当てると、その管理のために膨大な量の管理表が必要となるため、ブロックと呼ばれる数キロバイト程度の単位で対応関係は管理される。

[0120] 図12は、第2の実施の形態に係る不揮発性主記憶装置の状態を示す図である。

第2の実施の形態の不揮発性主記憶装置131は、データの読み書きに用いられる有効記憶素子と、ウェアレベリングに用いられるウェアレベリング用記憶素子を有する。

[0121] 第2の実施の形態の不揮発性主記憶装置131は、外部から見た不揮発性主記憶装置131の容量に対し、実際には数割程度余分にウェアレベリング用の記憶素子を有する。

[0122] 第2の実施の形態の不揮発性主記憶装置131は、記憶素子管理テーブル203の代わりに、記憶素子管理テーブル203を拡張した記憶素子管理テーブルを有する。

[0123] 以下、第1の実施例の記憶素子管理テーブル203を拡張した記憶素子管理テーブルを拡張記憶素子管理テーブル203'と表記する。

[0124] 図13は、拡張記憶素子管理テーブルの例である。

拡張記憶素子管理テーブル203'は、外部から見た物理アドレス範囲、使用記憶素子位置、ページ番号、およびフラグが対応付けられて記載されている。

[0125] 外部から見た物理アドレス範囲は、外部(OSやメモリコントローラ112)から見た不揮発性主記憶装置131の物理アドレスの範囲である。

[0126] 使用記憶素子位置は、不揮発性主記憶装置131の記憶素子251の位置を示す。

ページ番号は、ページに割り当てられた番号である。

[0127] フラグは、ページの状態を示す情報である。

第2の実施の形態において、メモリダンプ処理やBGダンプ処理は第1の実施の形態と同様である。第2の実施の形態では、図8の割り込みダンプ処理の代わりに下記のような割り込みダンプ処理が行われる。

[0128] 図14は、第2の実施の形態に係る割り込みダンプ処理の詳細なフローチャートである。

図14は、図6のステップS511に対応する。

[0129] ステップS601において、NVMコントローラ201は、OSからの書

き込み要求を待つ。OSからの書き込み要求を受信した場合、制御はステップS602に進む。また、BGダンプ処理が終了した場合、割り込みダンプ処理は終了する。

[0130] ステップS602において、NVMコントローラ201は、拡張記憶素子管理テーブル203'を参照し、書き込み先のメモリアドレスを含むページのフラグが「未ダンプ」であるか否かチェックする。書き込み先のメモリアドレスを含むページのフラグが「未ダンプ」である場合、制御はステップS603に進み、「未ダンプ」でない場合、制御はステップS605に進む。

[0131] ステップS603において、NVMコントローラ201は、書き込み先のメモリアドレスを含む外部から見た物理アドレス範囲を無効にする。詳細には、NVMコントローラ201は、拡張記憶素子管理テーブル203'の書き込み先のメモリアドレスを含む外部から見た物理アドレス範囲に「無効」を書き込む。

[0132] ステップS604において、NVMコントローラ201は、無効にした外部から見た物理アドレス範囲を未割り当ての記憶素子に割り当てる。詳細には、NVMコントローラ201は、拡張記憶素子管理テーブル203'の未割り当ての記憶素子の使用記憶素子位置に対応する外部から見た物理メモリアドレス範囲およびページ番号に、それぞれ無効にした外部から見た物理アドレス範囲と無効にしたページのページ番号を書き込む。

[0133] ステップS605において、NVMコントローラ201は、書き込み要求データを新たに割り当てたページに書き込む。また、データが書き込まれたページのフラグは、「使用領域 (Dirty)」となる。そして、制御はステップ601に戻る。

[0134] 図15は、割り当て変更前と割り当て変更後の拡張記憶素子管理テーブルを示す図である。

[0135] 図15の上側は変更前の拡張記憶素子管理テーブル203'、下側は変更後の拡張記憶素子管理テーブル203'を示す。

[0136] 割り当て変更前において、図15の上側の拡張記憶素子管理テーブル20

3' に示すように、外部から見た物理メモリアドレスの範囲0x01010000~0x010101FFFFに対応する使用記憶素子位置は0x1254、ページ番号は20、フラグは未ダンプである。

[0137] また、未割り当ての記憶素子として、記憶素子位置が0x57B9である記憶素子がある。

図15では、書込み要求の書込み先のメモリアドレスが0x01010000~0x010101FFFFの範囲に含まれるものとする。

[0138] 図14で述べたように、書込み先のメモリアドレスを含むページのフラグが「未ダンプ」である場合、NVMコントローラ201は、拡張記憶素子管理テーブル203'の書込み先のメモリアドレスを含む外部から見た物理アドレス範囲に「無効」を書き込む。これにより、図15の下側の拡張記憶素子管理テーブル203'に示すように、無効としたページに対応する外部から見た物理メモリアドレスの範囲は、0x01010000~0x0101FFFF（無効）となっている。

[0139] また、NVMコントローラ201は、拡張記憶素子管理テーブル203'の未割り当ての記憶素子の使用記憶素子位置に対応する外部から見た物理アドレス範囲およびページ番号に、それぞれ無効にした外部から見た物理メモリアドレス範囲と無効にしたページのページ番号を書き込む。NVMコントローラ201は、書込み要求データを新たに割り当てたページに書き込み、フラグを「使用領域 (Dirty)」とする。

[0140] これにより、図15の下側の拡張記憶素子管理テーブル203'に示すように、記憶素子位置が0x57B9に対応する外部から見た物理メモリアドレス範囲は0x01010000~0x0101FFFF、ページ番号は20、フラグは「使用領域 (Dirty)」となる。

[0141] また、BGダンプ処理は割り込みダンプ処理とは非同期に順次実行されるため、もともとのアドレス範囲が後から判別可能なように、アドレスデータそのものは保持される。ダンプ後には当該記憶素子のアドレス範囲及びページ番号は未割当に変更され、ウェアレベリング用にプールされる。

- [0142] 第2の実施の形態の情報処理装置によれば、メモリダンプ処理による業務停止時間を短縮することができる。
- [0143] 第2の実施の形態の情報処理装置によれば、OSからの書き込み処理とダンプ処理とを非同期に実行することで、第1の実施の形態に比べて性能劣化が発生しにくくすることができる。
- [0144] (第3の実施の形態)
- 第3の実施の形態では、NVMコントローラ201がメモリダンプ処理を行っていない通常時に不揮発性記憶装置131のデータを外部記憶装置161にバックアップする場合について説明する。
- [0145] 尚、第3の実施の形態の情報処理装置の構成は、特に断りのない限り第1の情報処理装置の構成と同様であるため説明は省略する。
- [0146] 第3の実施の形態では、記憶素子管理テーブルのフラグとしてさらに「使用領域 (Sync)」を用いる。
- [0147] 通常時のバックアップ処理は、下記の通りである。
- 第3の実施の形態のNVMコントローラ201は、記憶素子管理テーブル203を監視し、フラグが「使用領域 (Dirty)」であるページがあるかチェックする。
- [0148] NVMコントローラ201は、フラグが「使用領域 (Dirty)」であるページを検出すると、該ページのデータを外部記憶装置161に書き込む (バックアップ)。
- [0149] そして、NVMコントローラ201は、該ページのフラグを「使用領域 (Sync)」に設定する。
- [0150] 尚、NVMコントローラ201は、OSからの書き込み要求によりページにデータを書き込んだ場合には、該ページのフラグを「使用領域 (Dirty)」に設定する。
- [0151] NVMコントローラ201は、常に記憶素子管理テーブル203を監視し、フラグが「使用領域 (Dirty)」であるページを減らすようにバックアップを行う。

- [0152] また、第3の実施の形態のメモリダンプ処理は、基本的に第1の実施の形態のメモリダンプ処理（図6）と同様である。
- [0153] 図16は、第3の実施の形態に係るクラッシュ事象発生直前およびBGダンプフラグ設定直後の記憶素子管理テーブルの状態を示す図である。
- [0154] 図16の上側はクラッシュ事象発生直前の記憶素子管理テーブルの状態を示し、下側はBGダンプフラグ設定直後の記憶素子管理テーブルの状態を示す。
- [0155] 図16では、記憶素子管理テーブル203の各レコードを記憶素子の物理的な位置（Column, Row）に対応するように配置して記載している。
- [0156] 図16では、クラッシュ事象発生直前において、ページ1、6のフラグは「使用領域（Dirty）」であり、ページ2、3、X-1のフラグは「使用領域（Sync）」であり、ページ4、5、Xのフラグは「空き領域」である。
- [0157] 上述のように、第3の実施の形態のメモリダンプ処理は、第1の実施の形態のメモリダンプ処理と同様である。BGダンプフラグが設定される（ステップS503）とNVMコントローラ201は記憶素子管理テーブル203のフラグを設定する（ステップS504）。図16の下側に示すように、「使用領域（Dirty）」は「未ダンプ」に設定され、「使用領域（Sync）」は「空き領域」に設定される。
- [0158] 第3の実施の形態では、フラグ＝使用領域（Sync）のページは、外部記憶装置161にバックアップ済みであるので、ダンプする必要が無い。
- [0159] 図11の下側と図16の下側とを比較すると、第3の実施の形態では、フラグが「未ダンプ」であるページが少なくなっている。
- [0160] よって、第3の実施の形態では、第1の実施の形態よりもメモリダンプ処理の時間が短縮できる。
- [0161] 図17は、記憶素子管理テーブルのフラグの状態遷移図である。
初期状態では、S0（フラグ＝空き領域）である。
- [0162] S0の状態において、ページにデータが書き込まれるとS1（フラグ＝使用領域（Dirty））に遷移する。

- [0163] S 1 の状態において、通常時にバックアップ処理が行われると S 2（フラグ＝使用領域（Sync））に遷移する。また、S 1 の状態において、ダンプフラグが有効になると、S 3（フラグ＝未ダンプ）に遷移する。また、S 1 の状態において、ページを含む不揮発性主記憶装置 1 3 1 が解放されると（メモリ解放）、S 0（フラグ＝未ダンプ）に遷移する。
- [0164] S 2 の状態において、ページにデータが書き込まれると、S 1（フラグ＝使用領域（Dirty））に遷移する。S 2 の状態において、ダンプフラグが有効になると、S 3（フラグ＝未ダンプ）に遷移する。
- [0165] S 3 の状態において、ダンプフラグが有効になると、S 4（フラグ＝未ダンプ（1））に遷移する。S 3 の状態において、メモリダンプ処理が実施されると、S 0（フラグ＝空き領域）に遷移する。
- [0166] 以下、同様に状態 S 4～S（X－1）において、ダンプフラグが有効になると、「未ダンプ」の世代を示す番号が 1 インクリメントされる。尚、状態 S X では、「未ダンプ」の世代を示す番号がインクリメントされない。また、S 4～S X の状態において、メモリダンプ処理が実施されると、S 0（フラグ＝空き領域）に遷移する。
- [0167] 第 3 の実施の形態の情報処理装置によれば、メモリダンプ処理による業務停止時間を短縮することができる。
- [0168] 第 3 の実施の形態の情報処理装置によれば、通常時にバックアップを採取しているため、メモリダンプ処理時にはバックアップとの差分のみをダンプすればよいので、メモリダンプ処理の時間を短縮することができる。
- [0169]（第 4 の実施の形態）
- 第 4 の実施の形態では、第 1～第 3 の実施の形態のいずれかにおいて、さらに不揮発性主記憶装置のすべてまたは一部をダンプ対象外とする処理や揮発性主記憶装置から不揮発性主記憶装置へのコピー等の処理を行う場合について説明する。
- [0170] 尚、第 4 の実施の形態の情報処理装置の構成は、特に断りのない限り第 1 の情報処理装置の構成と同様であるため説明は省略する。

[0171] 図18A、Bは、第4の実施の形態に係るメモリダンプ処理のフローチャートである。

先ず、情報処理装置101では、OSが実行されているものとする。

[0172] ステップS701において、情報処理装置101のOSでクラッシュ事象、すなわち致命的なエラーが発生する。

[0173] ステップS702において、CPU111は、ダンプ処理を開始する。

以下、揮発性主記憶装置ダンプルーチン（ステップS703～S707）と不揮発性主記憶装置フラグ設定処理（ステップS708～S712）が並列に実行される。尚、揮発性主記憶装置ダンプルーチンと不揮発性主記憶装置フラグ設定処理は、直列に実行されても良いし、どちらが先に実行されてもよい。

[0174] ステップS703において、CPU111は、フォアグラウンドの揮発性主記憶装置ダンプルーチン（フォアグラウンドダンプ）を開始する。尚、フォアグラウンドとは、情報処理装置101が専らダンプ処理のために動作しており、業務が稼働していない状態を示す。

[0175] 尚、CPU111は、主記憶装置（揮発性主記憶装置121および不揮発性主記憶装置131）のEEPROMを参照し、主記憶装置が揮発性か不揮発性か判別する。

[0176] ステップS704において、CPU111は、不揮発性主記憶装置131に揮発性主記憶装置121の全データまたは一部のデータをコピー可能な空き領域があるかチェックする。揮発性主記憶装置121のデータをコピー可能な空き領域がある場合、制御はステップS705に進み、揮発性主記憶装置121のデータをコピー可能な空き領域がない場合、制御はステップS706に進む。

[0177] ステップS705において、CPU111は、揮発性主記憶装置121のデータを不揮発性主記憶装置131にコピー（書き込み）する。揮発性主記憶装置121のデータを不揮発性主記憶装置131にコピーするときに、CPU111は、揮発性主記憶装置のももとのメモリアドレスが後から判別可

能なような付加情報（例えば、コピーしたデータの開始アドレスおよびサイズまたはアドレスのオフセットなど）を同時に保存する。

[0178] また、コピーの途中で不揮発性主記憶装置 131 の空き領域がなくなった場合は、制御はステップ S706 に進み、未コピーの揮発性主記憶装置 121 のデータは、外部記憶装置 161 にコピーされる。

[0179] ステップ S706 において、CPU 111 は、揮発性主記憶装置 121 のデータを外部記憶装置 161 にコピー（書き込み）する。

[0180] ステップ S707 において、CPU 111 は、揮発性主記憶装置ダンプリンを終了する。

[0181] 尚、不揮発性主記憶装置 131 の空き領域は、あらかじめ予約されているように実装することも可能だし、揮発性主記憶装置ダンプリンは、クラッシュ事象発生時の状態により変化する空き領域に動的に対応できるように実装されることが望ましい。

[0182] ステップ S708 において、CPU 111 は、不揮発性主記憶装置フラグ設定処理を開始する。

[0183] ステップ S709 において、CPU 111 は、BIOS 設定情報 173 を参照し、不揮発性主記憶装置 131 がブロックアクセスであるかチェックする。不揮発性主記憶装置 131 がブロックアクセスの場合、制御はステップ S710 に進み、不揮発性主記憶装置 131 がブロックアクセスでない場合（バイトアクセスである場合）、制御はステップ S711 に進む。尚、上記チェックは、全ての不揮発性主記憶装置 131 に対して行う。

[0184] ステップ S710 において、CPU 111 は、ブロックアクセスである不揮発性主記憶装置 131 をダンプの対象外とする。

[0185] ステップ S711 において、CPU 111 は、バイトアクセスである不揮発性主記憶装置 131 の BG ダンプフラグを設定する。詳細には、CPU 111 は、BIOS 設定情報 173 を参照し、各不揮発性主記憶装置 131 の優先度 (Pri) を確認する。CPU 111 は、Pri=High である不揮発性主記憶装置 131 の BG ダンプフラグを 1、Pri=Low である不揮発性主記憶装置 13

1のBGダンプフラグを2に設定する。

- [0186] BGダンプフラグが設定されたNVMコントローラ201は、記憶素子管理テーブル203のフラグを設定する。詳細には、NVMコントローラ201は、「使用領域(Dirty)」を「未ダンプ」に設定する。NVMコントローラ201は、「使用領域(Dirty)」以外のフラグを「空き領域」に設定する。
- [0187] ステップS712において、CPU111は、不揮発性主記憶装置フラグ設定処理を終了する。
- [0188] ステップS713において、CPU111は、不揮発性主記憶装置131の手動交換を行うか否か判断する。詳細には、CPU111は、BIOS設定情報173を参照し、「手動交換」がYかNかによって、揮発性主記憶装置131の手動交換を行うか否か判断する。手動交換を行う場合、制御はステップS714に進み、手動交換を行わない場合、制御はステップS717に進む。
- [0189] ステップS714において、CPU111は、情報処理装置101の電源をオフにする。
- [0190] また、CPU111は、ダンプ処理の開始および電源オフのタイミングそれぞれでBIOSから情報処理装置101を管理するBaseboard Management Controller(BMC)に対してイベント通知を行い、BMCは手動交換の準備および着手のタイミングをSimple Network Management Protocol(SNMP)トラップや電子メールなどの方法で保守員に知らせる。
- [0191] ステップS715において、保守員が手動で不揮発性主記憶装置131を交換し、情報処理装置101の電源をオンにする。
- [0192] 保守員はその場またはサポートセンターに別途用意した計算機で取り外した不揮発性主記憶装置のデータをダンプし、外部記憶装置161にダンプされていたデータとマージすることで解析に必要なダンプデータを得る。
- [0193] 尚、セキュリティを考慮して不揮発性主記憶装置131のデータが暗号化されている場合、保守員は情報処理装置101の所有者の了承のもとに解読

用コードを入手する。

- [0194] ステップS 7 1 6において、CPU 1 1 1は、OSを起動する。
ステップS 7 1 7において、CPU 1 1 1は、不揮発性主記憶装置1 3 1のメモリアドレス空間が消えてしまわないよう、CPU 1 1 1と揮発性主記憶装置1 2 1をリセットする。
- [0195] ステップS 7 1 8において、CPU 1 1 1は、OSを起動する。
ステップS 7 1 9において、CPU 1 1 1およびNVMコントローラ2 0 1は、BGダンプ処理を実行する。ステップS 7 1 9では、例えば、図6のステップS 5 0 7～S 5 1 0の処理が実行される。尚、BGダンプ処理は、ステップS 7 1 1においてBGダンプフラグが設定された不揮発性主記憶装置1 3 1で行われる。
- [0196] ステップS 7 2 0において、各NVMコントローラ2 0 1は、割り込みダンプ処理を行う。尚、割り込みダンプ処理は、図8または図14で説明した通りである。
- [0197] 尚、上述のCPU 1 1 1が実行する各ステップは、CPU 1 1 1がBIOS 1 7 2を読み出して実行することにより実現される。
- [0198] 上記に様にして得られたダンプデータは取得タイミングやアドレスがバラバラのファイルとなってしまいうため、別途用意するソフトウェアツールによりマージするか、バラバラのファイルを読み込み可能なダンプの解析ソフトを利用する。
- [0199] 図19は、第4の実施の形態に係る各処理時のデータの位置を示す図である。
図19では、不揮発性主記憶装置1 3 1-1がブロックアクセス、不揮発性主記憶装置1 3 1-1がバイトアクセスとする。また、不揮発性主記憶装置1 3 1-3の記載は省略する。また、図19では、揮発性主記憶装置1 2 1-1を記載し、揮発性主記憶装置1 2 1-2、1 2 1-3は省略する。
- [0200] 先ず、クラッシュ事象発生時には、不揮発性主記憶装置1 3 1-1はデータA、Bを格納している。不揮発性主記憶装置1 3 1-2はデータCを格納

しており、空き領域 (Free) を有する。揮発性主記憶装置 121-1 は、データ D、E を格納している (図 19 のクラッシュ事象発生時の状態)。尚、不揮発性主記憶装置 131-1 のデータ A、B は、バックグラウンドダンプ完了まで保持される。

[0201] 揮発性主記憶装置ダンプルーチン (フォアグラウンドダンプ) が開始され、揮発性主記憶装置 121-1 のデータ D は、不揮発性主記憶装置 131-2 にコピーされる (図 19 のフォアグラウンドダンプ中)。また、データ D の、不揮発性主記憶装置 131-2 へのコピーにより、不揮発性主記憶装置 131-2 の空き領域は無くなる。

[0202] 不揮発性主記憶装置 131-2 の空き領域が無い場合、揮発性主記憶装置 121-1 のデータ E は、外部記憶装置 161 にコピーされる (図 19 のフォアグラウンドダンプ完了)。

[0203] OS が再起動され (ステップ S718)、BG ダンプ処理が実行される (ステップ S719) と、不揮発性主記憶装置 131-2 のデータ C、D は、外部記憶装置 161 にコピーされる。また、再起動された OS のデータ F、G が揮発性主記憶装置 121-1 に書き込まれ、再起動された OS のデータ H が不揮発性主記憶装置 131-2 に書き込まれる (図 19 のバックグラウンドダンプ中)。

[0204] BG ダンプ処理が終了すると、揮発性主記憶装置 121-1 は、データ F、G を格納しており、不揮発性主記憶装置 131-2 は、データ H を格納しており、データ D があった領域は空き領域となっている (図 19 のバックグラウンドダンプ完了)。

[0205] 第 4 の実施の形態の情報処理装置によれば、メモリダンプ処理による業務停止時間を短縮することができる。

[0206] 第 4 の実施の形態の情報処理装置によれば、ブロックアクセスの不揮発性主記憶装置はダンプ不要であるから、これをダンプ対象外となるよう制御することで無駄な処理をなくすことができる。

[0207] 第 4 の実施の形態の情報処理装置によれば、揮発性主記憶装置のデータを

外部記憶装置よりも高速にアクセス可能な不揮発性主記憶の空き領域へコピーすることにより、フォアグラウンドのダンプの所要時間を短縮することができる。

[0208] 第4の実施の形態の情報処理装置によれば、不揮発性主記憶装置から外部記憶装置へのダンプ処理をOS再起動後のバックグラウンド処理とすることで、業務停止時間を短縮することができる。

[0209] 図20は、コンピュータシステムの構成図である。

以上説明したようなメモリダンプ処理は、当然一般的なコンピュータシステムによって実現することが可能である。図20において、実施の形態のメモリダンプ処理を実現するための情報処理装置11は、本体12とメモリ13とによって構成されている。メモリ13は、揮発性主記憶装置121、不揮発性主記憶装置131、外部記憶装置161、BIOS格納域171などの記憶装置に対応する。

[0210] このようなメモリ13に各実施の形態のメモリダンプ処理等を行うプログラムが格納され、そのプログラムが本体12によって実行されることにより、実施の形態の各種処理を実現することが可能となる。

[0211] このようなプログラムは、プログラム提供者側からネットワーク21を介して情報処理装置11にロードされることもできる。またプログラムは、市販され、流通している可搬型記憶媒体31に格納され、そのような可搬型記憶媒体31が情報処理装置11にセットされ、メモリ13に読み出されて本体12によって実行されることも可能である。

[0212] 可搬型記憶媒体31としてはCD-ROM、フレキシブルディスク、光ディスク、光磁気ディスクなど様々な形式の記憶媒体を使用することができ、このような記憶媒体から読み出されたプログラムコード自体が実施の形態の機能を実現する。

請求の範囲

[請求項1]

オペレーティングシステムを実行する処理部と、
前記処理部が直接アクセス可能且つコントローラを有する不揮発性主記憶装置と、
前記処理部が直接アクセス可能でない外部記憶装置と、
を備え、
前記処理部が前記オペレーティングシステムのエラーを検出したとき、
前記処理部は、前記不揮発性主記憶装置以外をリセットして、前記オペレーティングシステムを再起動し、
前記コントローラは、前記不揮発性主記憶装置のデータを前記外部記憶装置に書き込むことを特徴とする情報処理装置。

[請求項2]

前記不揮発性主記憶装置は、複数の不揮発性主記憶装置であり、
前記複数の不揮発性主記憶装置の各不揮発性主記憶装置の優先度が記載された設定情報を格納する情報格納部をさらに備え、
前記処理部は、前記設定情報を参照し、前記優先度の高い順に、前記不揮発性主記憶装置のデータを前記外部記憶装置に書き込む処理を前記コントローラに実行させることを特徴とする請求項1記載の情報処理装置。

[請求項3]

前記コントローラは、
前記オペレーティングシステムから前記不揮発性主記憶装置へ書き込む第1のデータを受信し、
前記第1のデータが書き込まれる領域に存在する第2のデータが前記外部記憶装置に書き込まれていない場合、前記第2のデータを前記外部記憶装置に書き込み、
前記第2のデータを前記外部記憶装置に書き込み後、前記第1のデータを前記不揮発性主記憶装置へ書き込むことを特徴とする請求項1記載の情報処理装置。

- [請求項4] 前記不揮発性主記憶装置は、外部から見たアドレスに対して未割り当ての領域を有し、
- 前記コントローラは、
- 前記オペレーティングシステムから前記不揮発性主記憶装置へ書き込む第1のデータを受信し、
- 前記第1のデータが書き込まれる第1の領域に存在する第2のデータが前記外部記憶装置に書き込まれていない場合、
- 前記第1の領域に割り当てられている外部から見たアドレスを無効にし、
- 前記未割り当ての領域に前記第1の領域に割り当てられていた外部から見たアドレスを割り当て、
- 前記第1の領域に割り当てられていた外部から見たアドレスを割り当てた前記未割り当ての領域へ前記第1のデータを書き込むことを特徴とする請求項1記載の情報処理装置。
- [請求項5] 前記コントローラは、
- 前記不揮発性主記憶装置の複数の領域の各データが前記不揮発性主記憶装置へ書き込み済みであることを示すフラグが記載された管理テーブルを有し、
- 前記オペレーティングシステムの動作時に、前記管理テーブルを参照し、前記外部記憶装置へ書き込んでいない未書き込みデータを検出し、
- 該未書き込みのデータを前記外部記憶装置へ書き込み、
- 書き込んだデータに対応するフラグを、前記外部記憶装置に書き込み済みであることを示すように設定することを特徴とする請求項1記載の情報処理装置。
- [請求項6] 前記処理部が前記オペレーティングシステムのエラーを検出したとき、
- 前記コントローラは、前記管理テーブルを参照し、前記フラグに基

づいて、前記外部記憶装置へ書き込んでいない前記不揮発性主記憶装置のデータのみを前記外部記憶装置へ書き込むことを特徴とする請求項5記載の情報処理装置。

- [請求項7] 前記情報処理装置は、前記不揮発性主記憶装置がダンプ対象であるか否かを示す設定情報を格納する情報格納部をさらに備え、
- 前記不揮発性主記憶装置は、前記不揮発性主記憶装置のデータを前記外部記憶装置に書き込む処理を実行するか否かを示すダンプフラグを格納する主記憶装置情報格納部をさらに備え、
- 前記処理部は、前記設定情報に基づいて、ダンプ対象である前記不揮発性主記憶装置の前記ダンプフラグを有効に設定し、
- 前記コントローラは、前記ダンプフラグが有効の場合に、前記不揮発性主記憶装置のデータを前記外部記憶装置に書き込むことを特徴とする請求項1記載の情報処理装置。

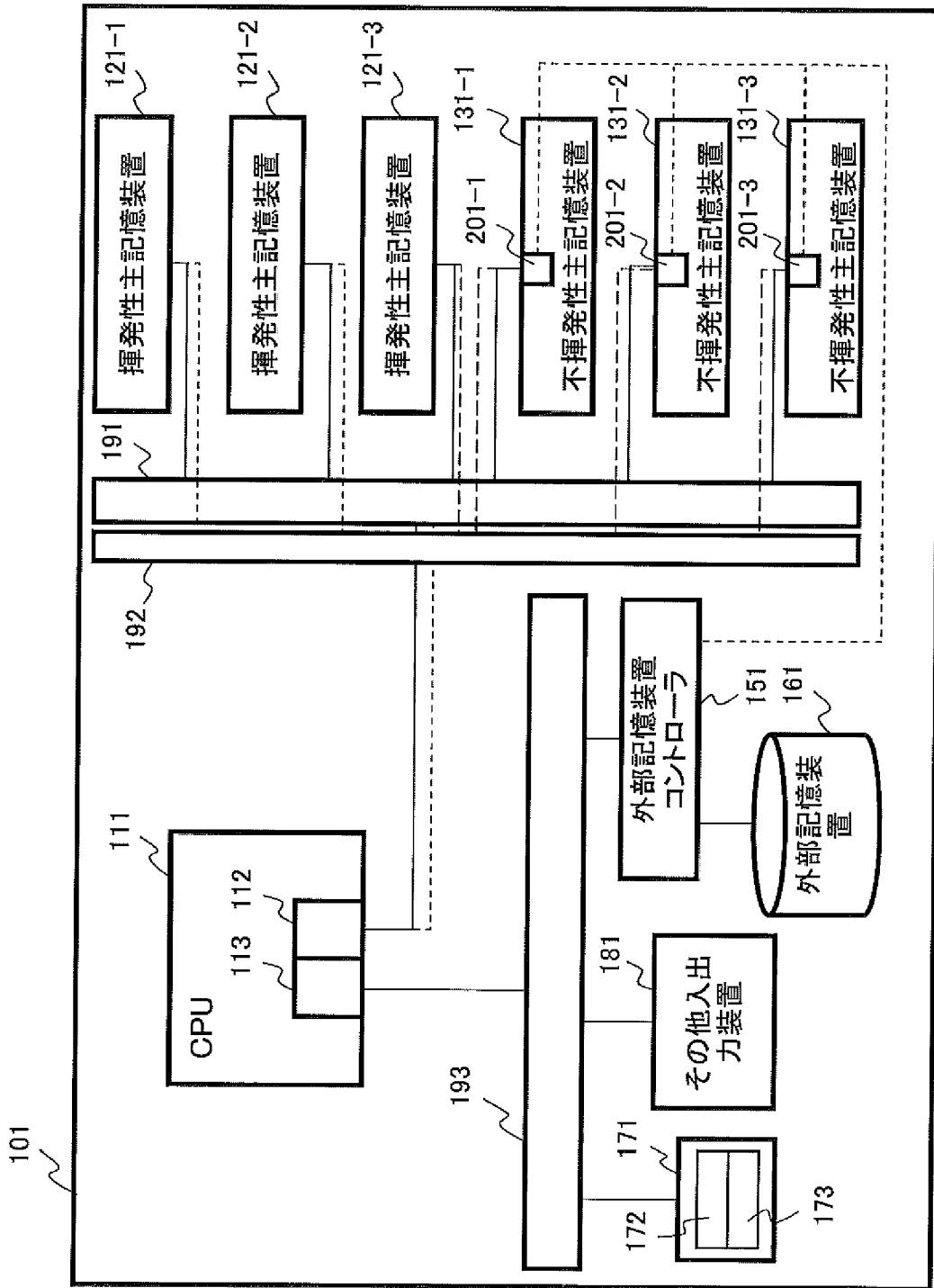
- [請求項8] 前記処理部が直接アクセス可能な揮発性主記憶装置をさらに備え、
- 前記コントローラは、前記不揮発性主記憶装置のデータを前記外部記憶装置に書き込む前に、前記揮発性主記憶装置のデータを前記不揮発性主記憶装置に書き込むことを特徴とする請求項1記載の情報処理装置。

- [請求項9] オペレーティングシステムを実行する処理部と、前記処理部が直接アクセス可能な不揮発性主記憶装置と、前記処理部が直接アクセス可能でない外部記憶装置と、を備える情報処理装置が実行するメモリダンプ方法であって、
- 前記オペレーティングシステムのエラーを検出し、
- 前記不揮発性主記憶装置以外をリセットし、
- 前記オペレーティングシステムを再起動し、
- 前記不揮発性主記憶装置のデータを前記外部記憶装置に書き込む処理を備えるメモリダンプ方法。

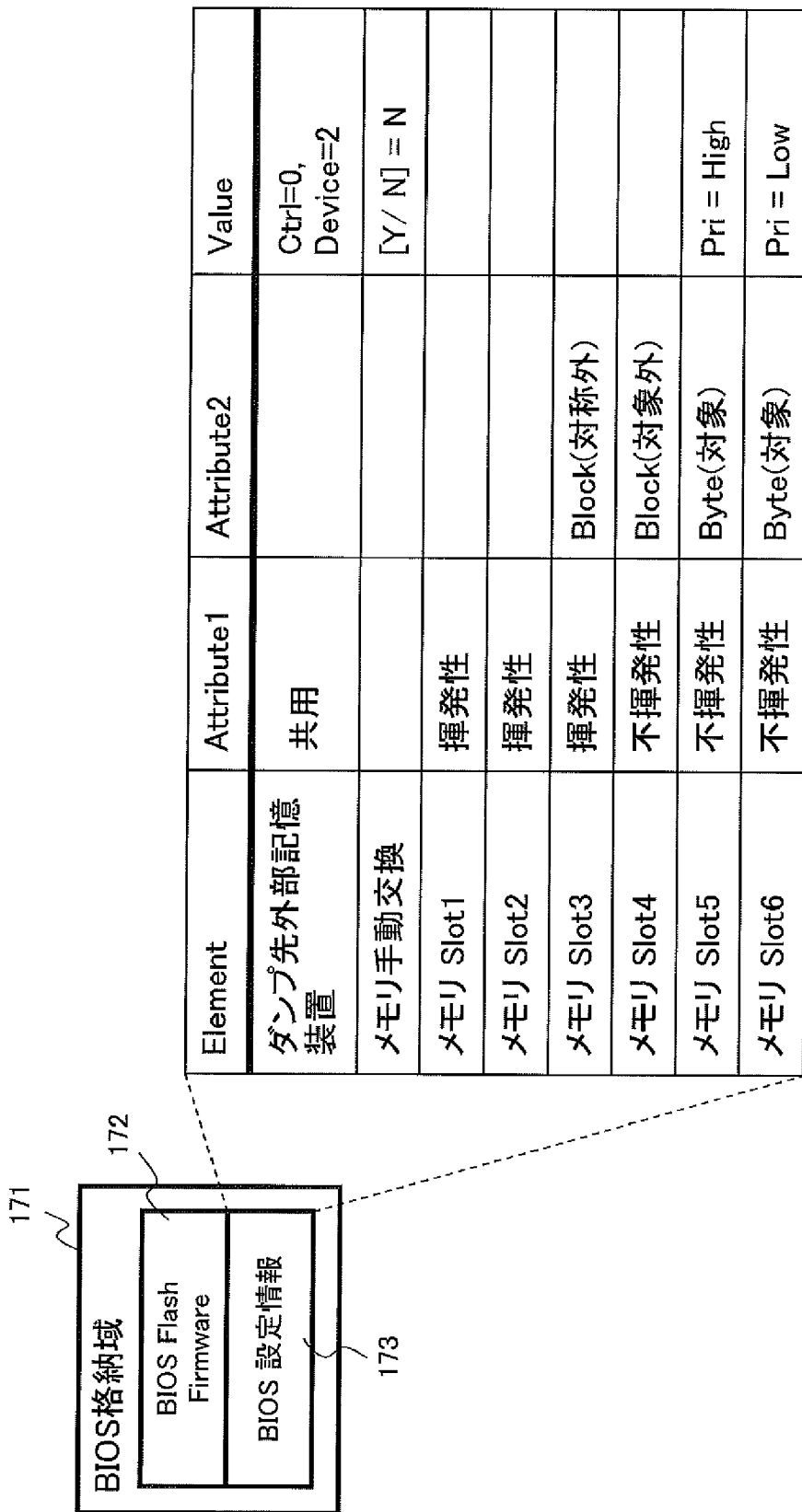
- [請求項10] オペレーティングシステムを実行する処理部と、前記処理部が直接

アクセス可能な不揮発性主記憶装置と、前記処理部が直接アクセス可能でない外部記憶装置と、を備えるコンピュータに、
前記オペレーティングシステムのエラーを検出し、
前記不揮発性主記憶装置以外をリセットし、
前記オペレーティングシステムを再起動し、
前記不揮発性主記憶装置のデータを前記外部記憶装置に書き込む
処理を実行させるメモリダンププログラム。

[図1]



[図2]



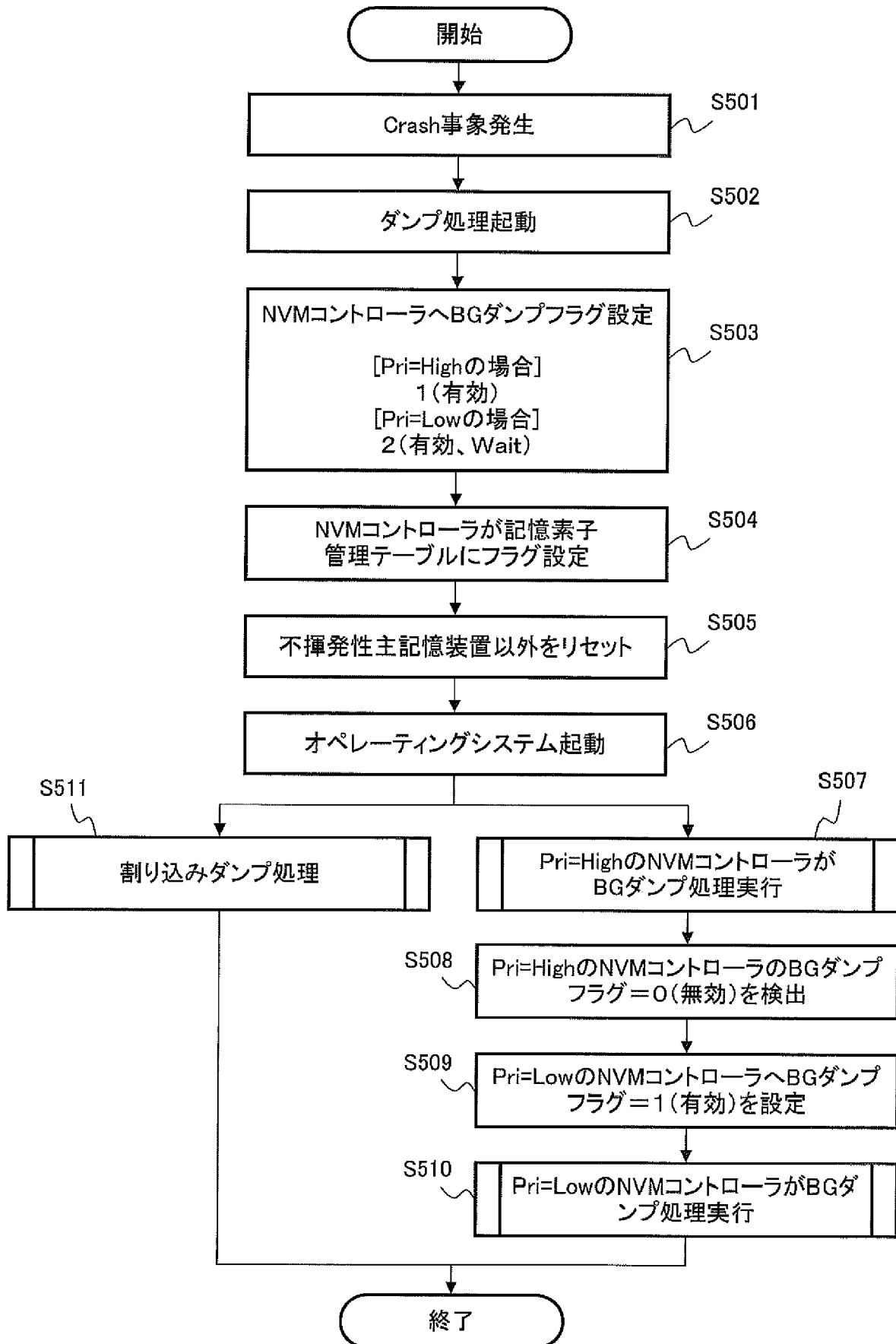
[図4]

ページ番号	フラグ
0	空き領域
1	使用領域(Dirty)
2	使用領域(Sync)
3	未ダンプ
4	未ダンプ(1)
5	未ダンプ(2)
...	...
15	Reserved

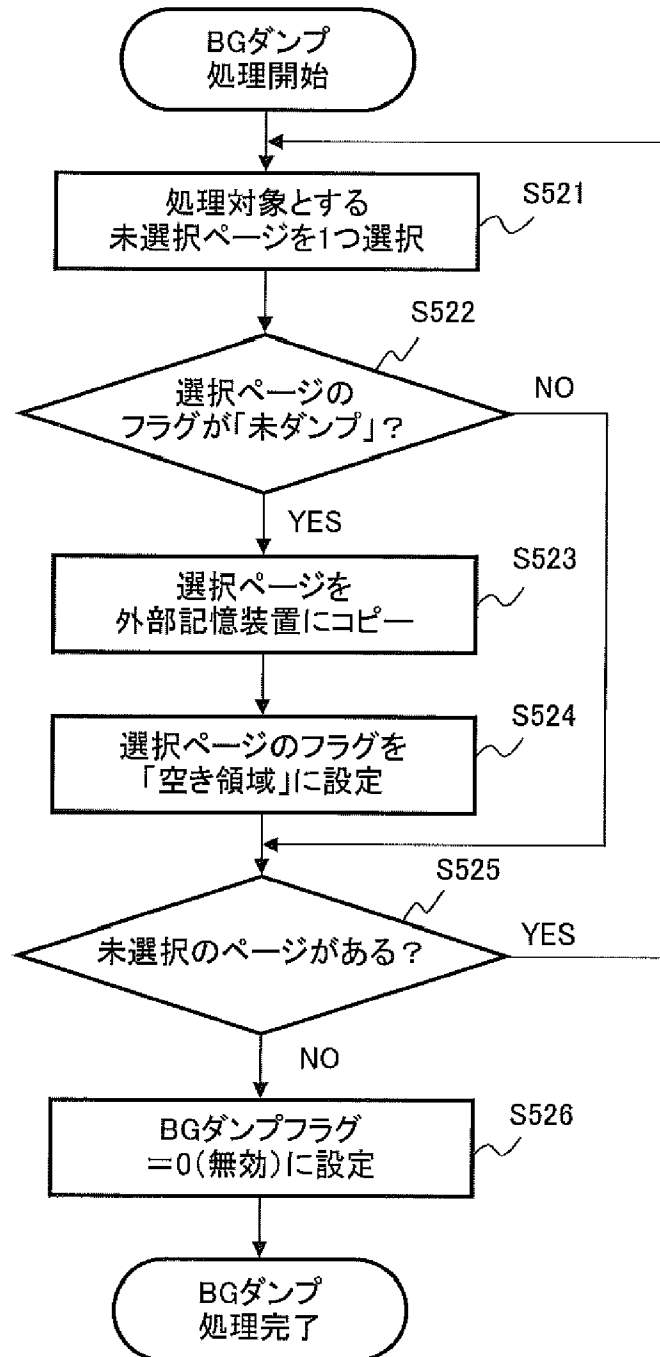
[図5]

主記憶モジュールの種類:不揮発性
主記憶モジュールの容量:100GB
ECC有無:あり
温度センサ値:35°C
...
予備領域
BGダンプフラグ[0-2]=0
ダンプ先外部記憶装置 Ctrl=0, Device=2
自由領域:未使用
...

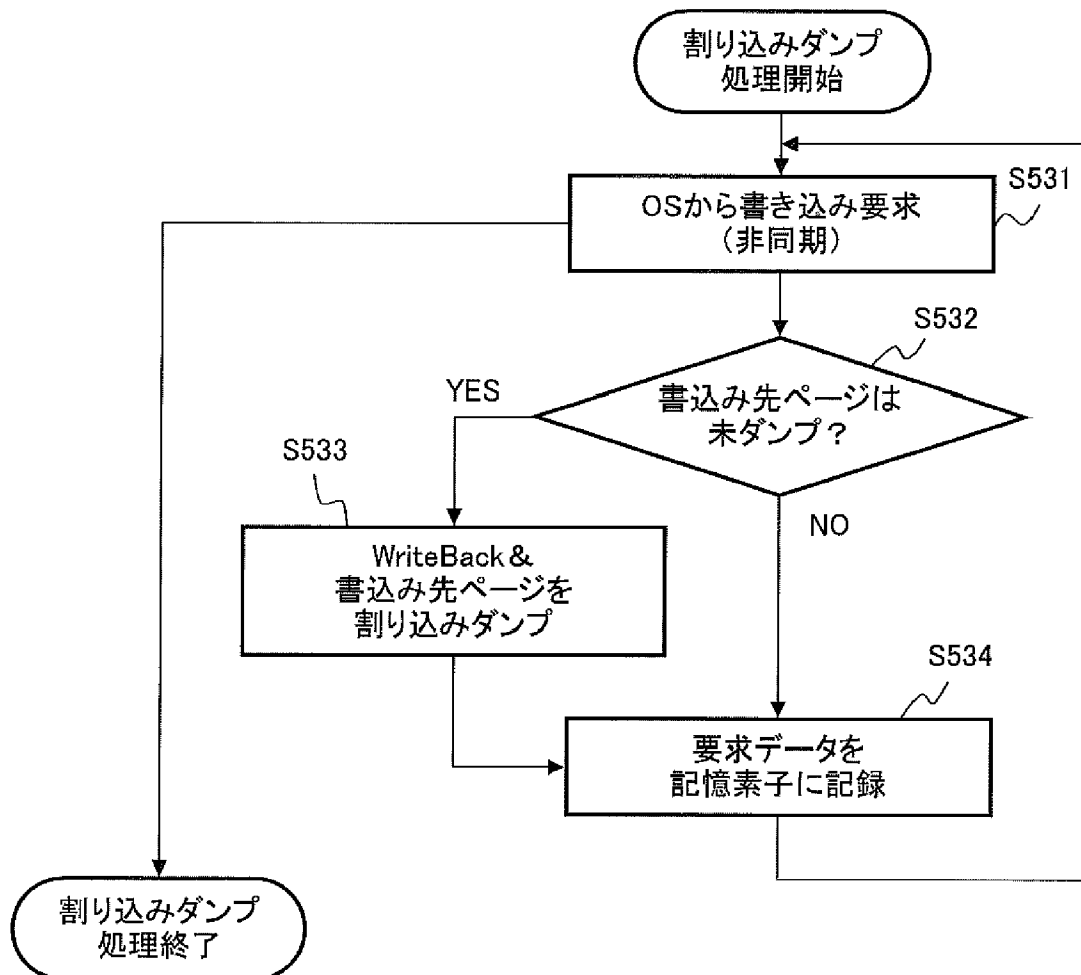
[図6]



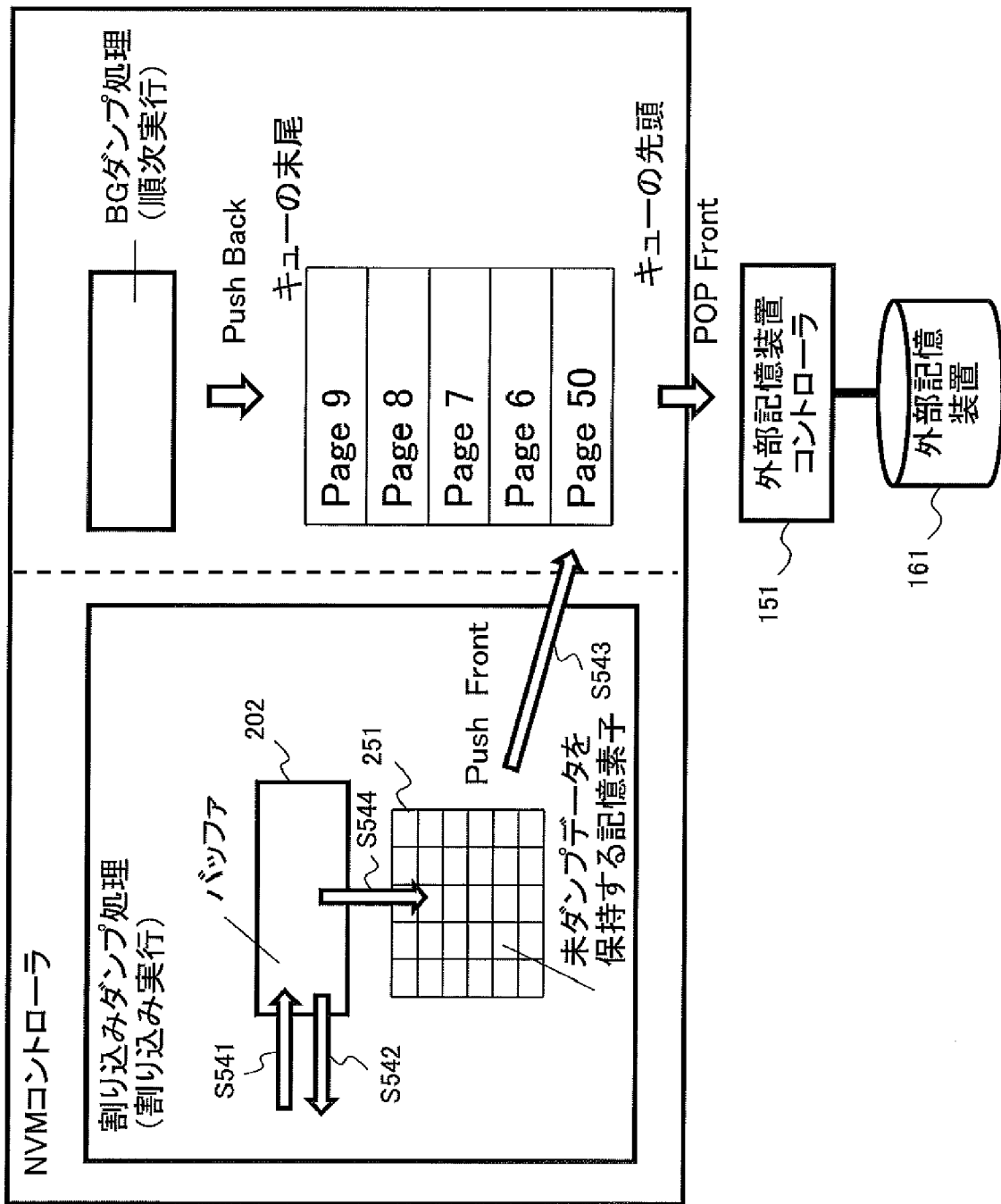
[図7]



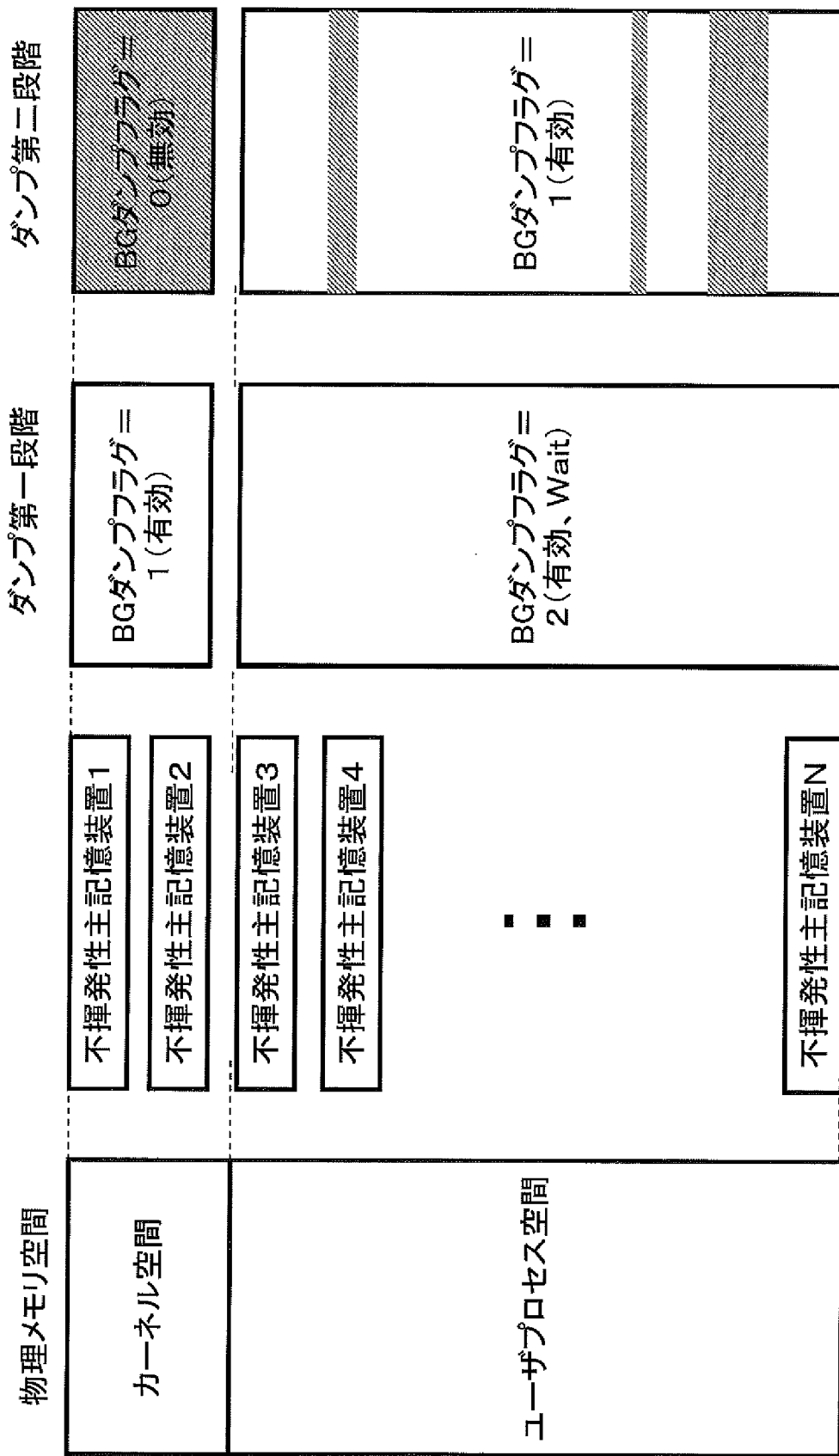
[図8]



[図9]



[図10]



[図11]

クラッシュ事象発生直前の状態

	Column(0~N)	
Row(0~M)	Page1: Flag="使用領域(Dirty)"	Page2: Flag="使用領域(Dirty)"
	Page3: Flag="使用領域(Dirty)"	Page4: Flag="空き領域"
	Page5: Flag="空き領域"	Page6: Flag="使用領域(Dirty)"

	PageX-1: Flag="使用領域(Dirty)"	PageX: Flag="空き領域"



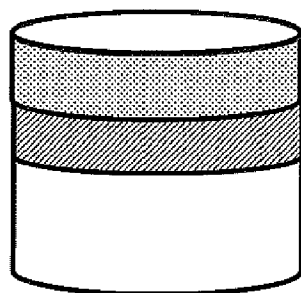
BGダンプフラグが有効にされた直後の状態

	Column(0~N)	
Row(0~M)	Page1: Flag="未ダンプ"	Page2: Flag="未ダンプ"
	Page3: Flag="未ダンプ"	Page4: Flag="空き領域"
	Page5: Flag="空き領域"	Page6: Flag="未ダンプ"

	PageX-1: Flag="未ダンプ"	PageX: Flag="空き領域"

[図12]

131



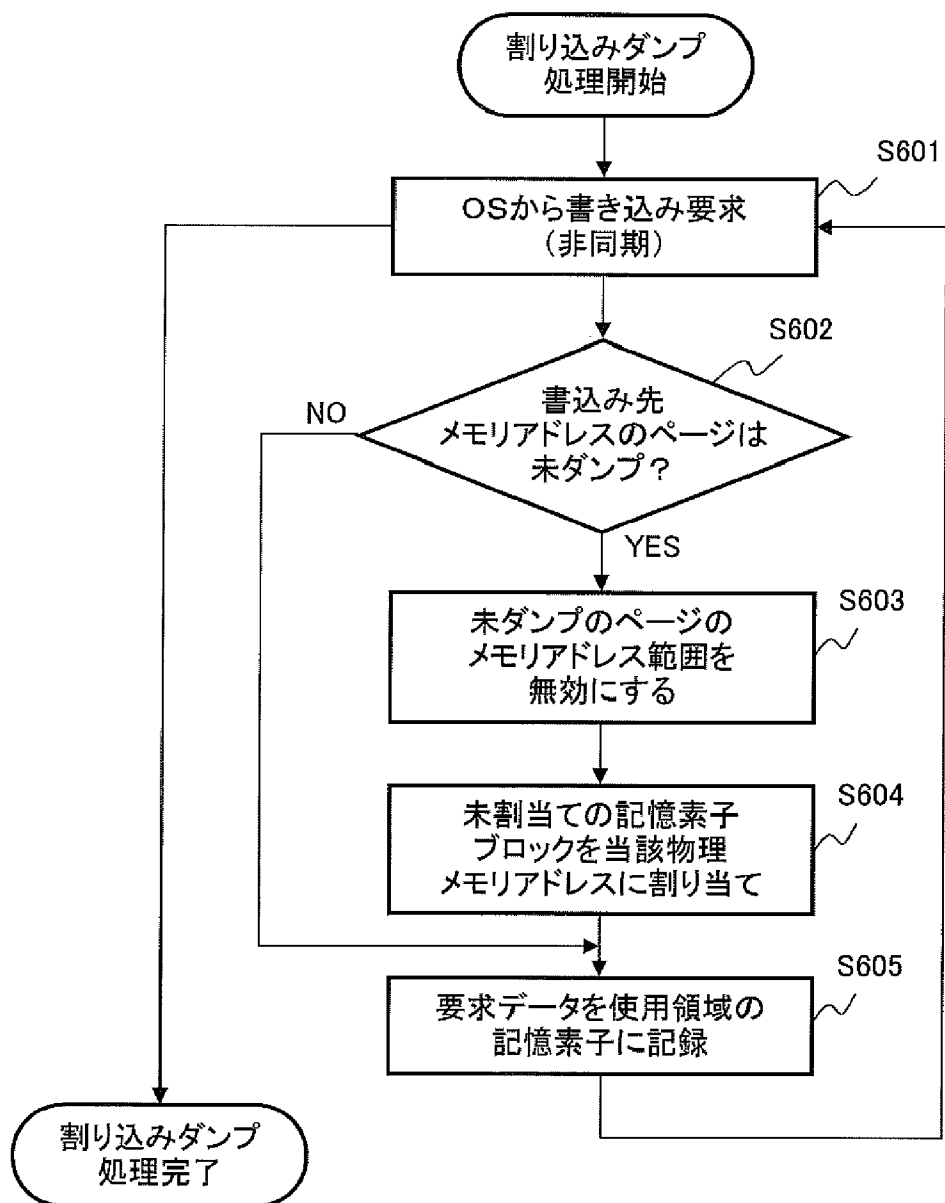
- ウェアレベリング用記憶素子
- 有効記憶素子(空き領域)
- 有効記憶素子(使用領域)

[図13]

203'

外部から見た物理メモリ アドレス範囲	使用記憶素子 位置	ページ 番号	フラグ
0x0100 0000 ~0x0100FFFF	0xABCD	3	使用領域 (Dirty)
0x0101 0000 ~0x0101 FFFF	0x1254	20	未ダンプ
0x0102 0000 ~0x0102 FFFF	0xC05A	35	未ダンプ
...
未割当	0x57B9		
...

[図14]



[図15]

外部から見た物理メモリ アドレス範囲	使用記憶素子 位置	ページ 番号	フラグ
0x0100 0000 ~0x0100FFFF	0xABCD	3	使用領域 (Dirty)
0x0101 0000 ~0x0101 FFFF	0x1254	20	未ダンプ
0x0102 0000 ~0x0102 FFFF	0xC05A	35	未ダンプ
...
未割当	0x57B9		
...



外部から見た物理メモリ アドレス範囲	使用記憶素子 位置	ページ 番号	フラグ
0x0100 0000 ~0x0100FFFF	0xABCD	3	使用領域 (Dirty)
0x0101 0000 ~0x0101 FFFF	0x57B9	20	使用領域 (Dirty)
0x0102 0000 ~0x0102 FFFF	0xC05A	35	未ダンプ
...
0x0101 0000 ~0x0101 FFFF(無効)	0x1254	20	未ダンプ
...

[図16]

クラッシュ事象発生直前の状態

	Column(0~N)	
Row(0~M)	Page1: Flag="使用領域(Dirty)"	Page2: Flag="使用領域(Sync)"
	Page3: Flag="使用領域(Sync)"	Page4: Flag="空き領域"
	Page5: Flag="空き領域"	Page6: Flag="使用領域(Dirty)"

	PageX-1: Flag="使用領域(Sync)"	PageX: Flag="空き領域"

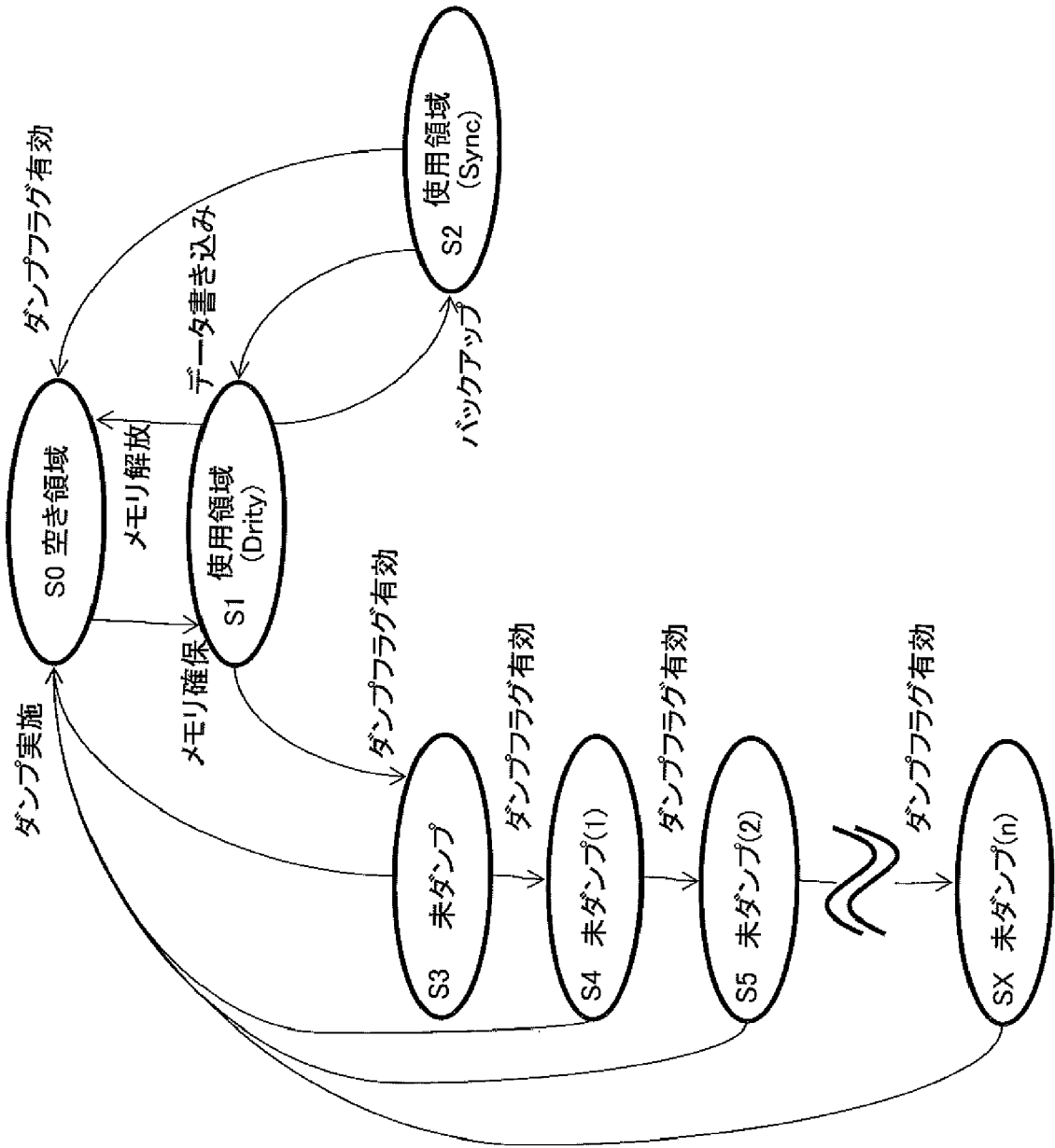


BGダンプフラグが有効にされた直後の状態

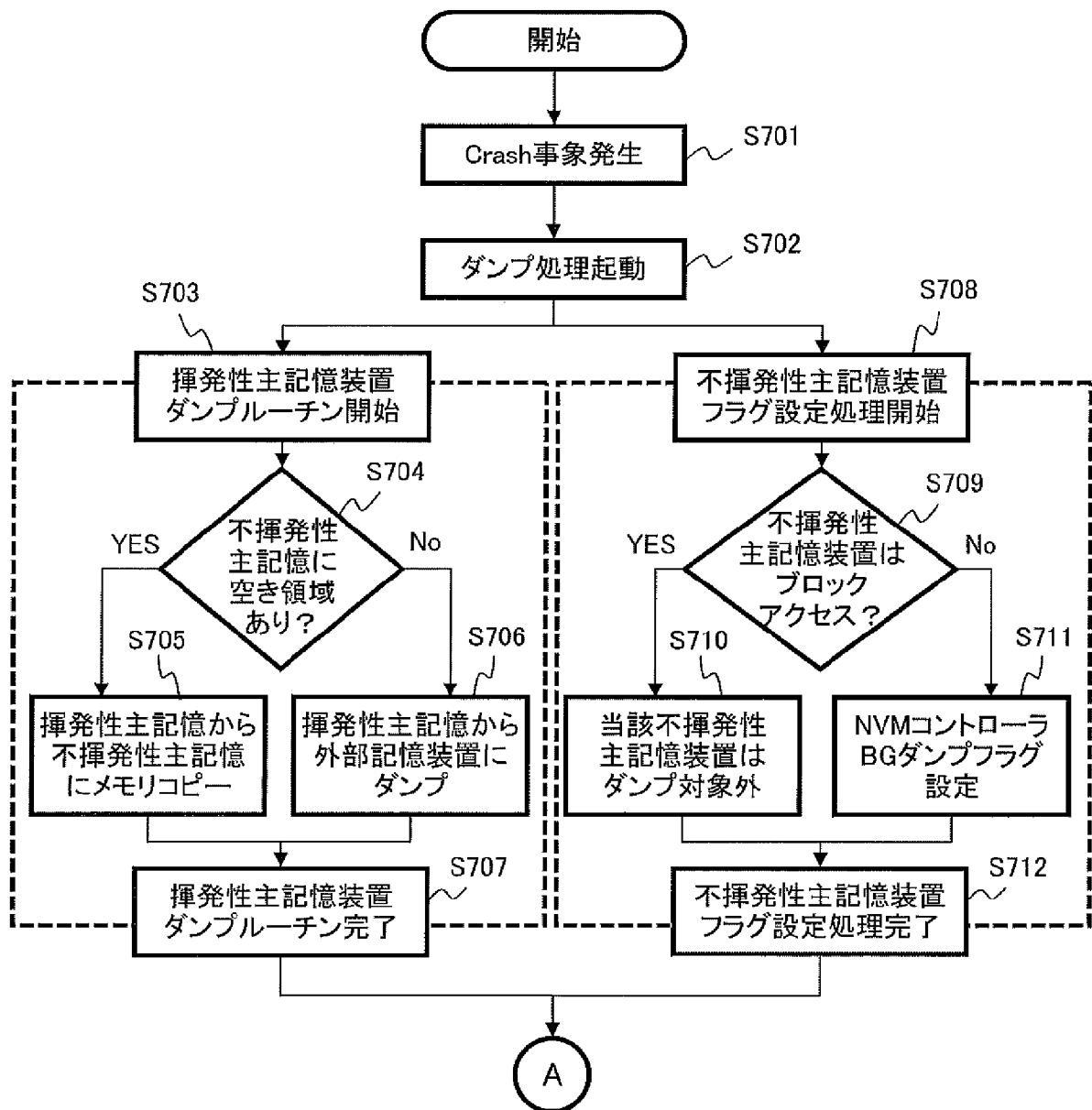
	Column(0~N)	
Row(0~M)	Page1: Flag="未ダンプ"	Page2: Flag="空き領域"
	Page3: Flag="空き領域"	Page4: Flag="空き領域"
	Page5: Flag="空き領域"	Page6: Flag="未ダンプ"

	PageX-1: Flag="空き領域"	PageX: Flag="空き領域"

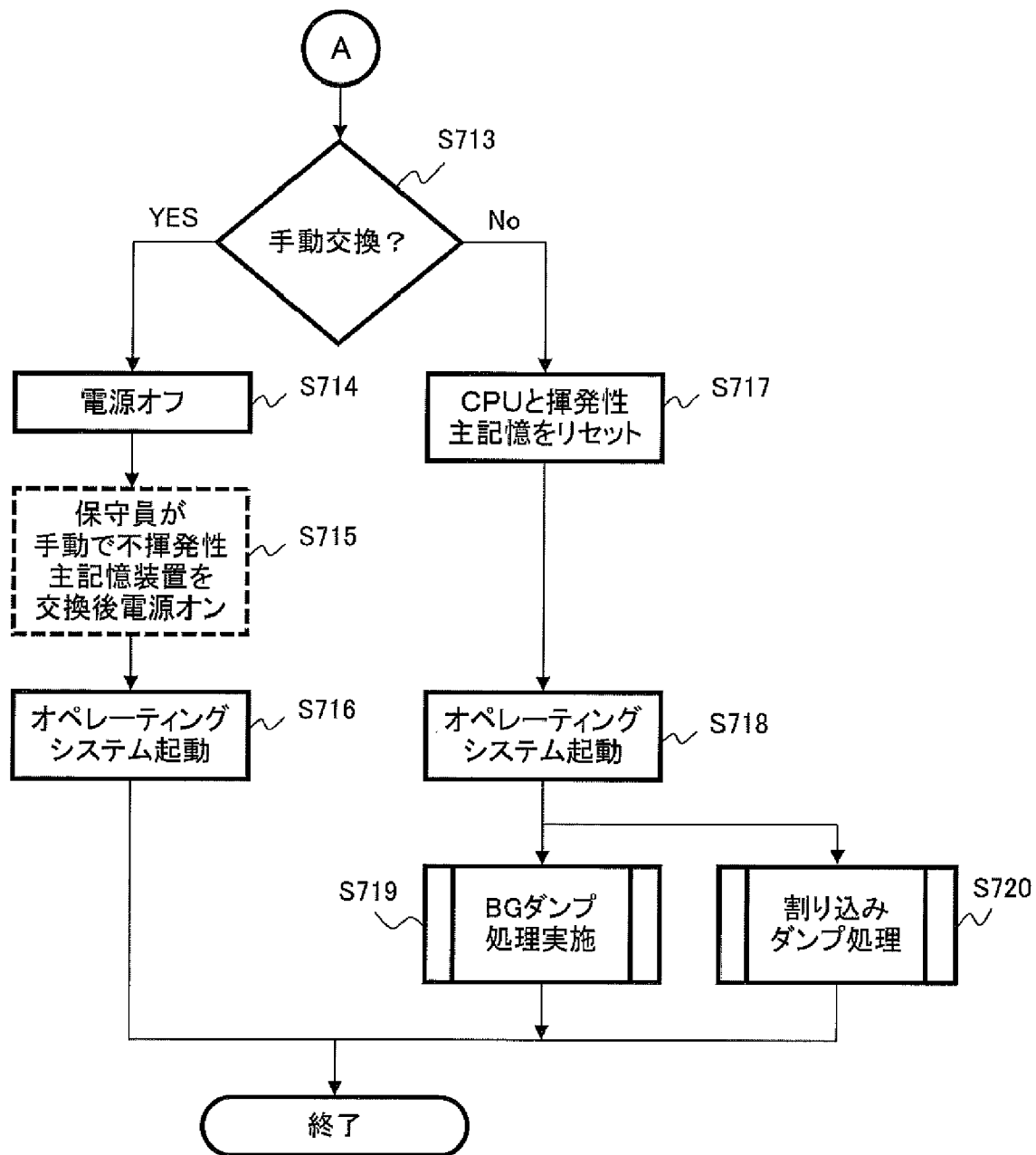
[図17]



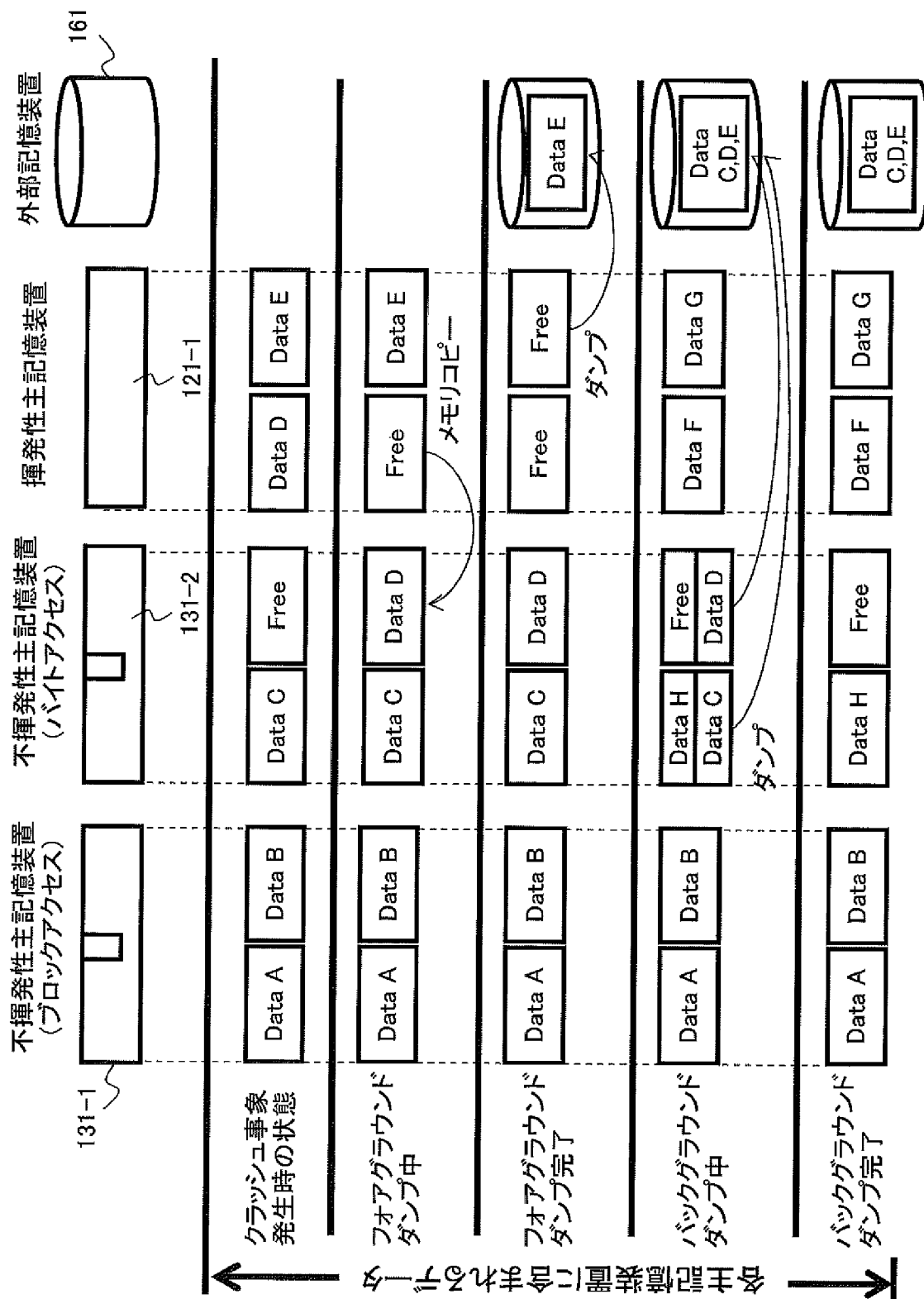
[図18A]



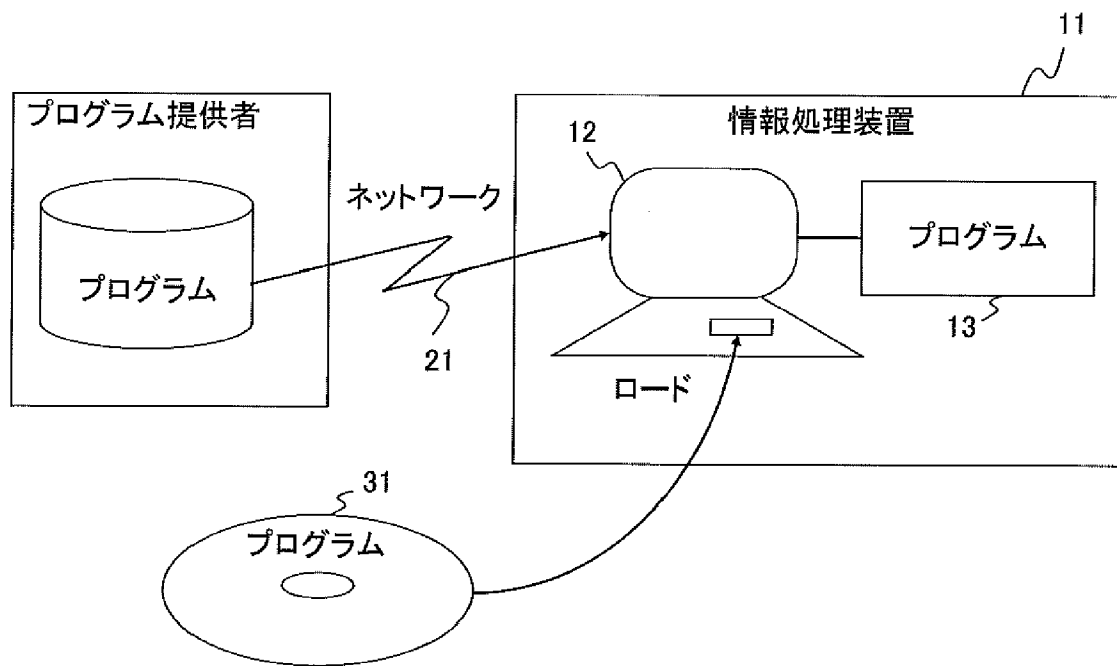
[図18B]



[図19]



[図20]



INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2013/070698

A. CLASSIFICATION OF SUBJECT MATTER
G06F11/34(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)
G06F11/00, G06F11/34, G06F12/00-12/06, G06F12/16, G06F13/16-13/18

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	1996-2013
Kokai Jitsuyo Shinan Koho	1971-2013	Toroku Jitsuyo Shinan Koho	1994-2013

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.
Y A	JP 2011-257987 A (Mitsubishi Electric Corp.), 22 December 2011 (22.12.2011), claim 2; paragraphs [0022] to [0039]; fig. 1, 2 (Family: none)	9, 10 1-8
Y A	US 2005/0204107 A1 (NARAYAN, Ranjani et al.), 15 September 2005 (15.09.2005), paragraphs [0050] to [0083]; fig. 1 to 4 (Family: none)	9, 10 1-8
A	JP 9-259014 A (Meidensha Corp.), 03 October 1997 (03.10.1997), entire text; all drawings (particularly, claim 1; fig. 1) (Family: none)	1-10

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

* Special categories of cited documents:	"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
"A" document defining the general state of the art which is not considered to be of particular relevance	"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
"E" earlier application or patent but published on or after the international filing date	"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
"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)	"&" document member of the same patent family
"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	

Date of the actual completion of the international search 28 August, 2013 (28.08.13)	Date of mailing of the international search report 29 October, 2013 (29.10.13)
---	---

Name and mailing address of the ISA/ Japanese Patent Office	Authorized officer
Facsimile No.	Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2013/070698

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2004-362426 A (Sony Corp.), 24 December 2004 (24.12.2004), paragraphs [0060] to [0081]; fig. 1, 4 to 6 (Family: none)	1-10
A	JP 2009-205254 A (Fujitsu Ltd.), 10 September 2009 (10.09.2009), paragraphs [0079] to [0089]; fig. 8, 9 & US 2009/0216967 A1	2
A	JP 2008-242999 A (Hitachi, Ltd.), 09 October 2008 (09.10.2008), paragraphs [0034] to [0038]; fig. 12 (Family: none)	3
A	JP 5-134906 A (NEC Corp.), 01 June 1993 (01.06.1993), paragraphs [0011] to [0027] (Family: none)	5, 6

A. 発明の属する分野の分類（国際特許分類（IPC）） Int.Cl. G06F11/34(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（IPC）） Int.Cl. G06F11/00, G06F11/34, G06F12/00-12/06, G06F12/16, G06F13/16-13/18		
最小限資料以外の資料で調査を行った分野に含まれるもの 日本国実用新案公報 1922-1996年 日本国公開実用新案公報 1971-2013年 日本国実用新案登録公報 1996-2013年 日本国登録実用新案公報 1994-2013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
Y A	JP 2011-257987 A（三菱電機株式会社）2011.12.22, 【請求項2】，【0022】-【0039】，図1,2（ファミリーなし）	9, 10 1-8
Y A	US 2005/0204107 A1 (NARAYAN, Ranjani et al.) 2005.09.15, [0050]-[0083], 図1-4（ファミリーなし）	9, 10 1-8
A	JP 9-259014 A（株式会社明電舎）1997.10.03, 全文,全図（特に、【請求項1】，図1）（ファミリーなし）	1-10
<input checked="" type="checkbox"/> C欄の続きにも文献が列挙されている。 <input type="checkbox"/> パテントファミリーに関する別紙を参照。		
* 引用文献のカテゴリー 「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 28.08.2013	国際調査報告の発送日 29.10.2013	
国際調査機関の名称及びあて先 日本国特許庁（ISA/J P） 郵便番号100-8915 東京都千代田区霞が関三丁目4番3号	特許庁審査官（権限のある職員） 多賀 実 電話番号 03-3581-1101 内線 3545	5 B 9367

C (続き) . 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
A	JP 2004-362426 A (ソニー株式会社) 2004. 12. 24, 【0060】 - 【0081】 , 図 1, 4-6 (ファミリーなし)	1 - 1 0
A	JP 2009-205254 A (富士通株式会社) 2009. 09. 10, 【0079】 - 【0089】 , 図 8, 9 & US 2009/0216967 A1	2
A	JP 2008-242999 A (株式会社日立製作所) 2008. 10. 09, 【0034】 - 【0038】 , 図 12 (ファミリーなし)	3
A	JP 5-134906 A (日本電気株式会社) 1993. 06. 01, 【0011】 - 【0027】 (ファミリーなし)	5, 6