

(19) 日本国特許庁(JP)

(12) **公開特許公報(A)**

(11) 特許出願公開番号

特開2009-199428

(P2009-199428A)

(43) 公開日 平成21年9月3日(2009.9.3)

(51) Int.Cl.

G O 6 F 3/06 (2006.01)

F l

G06F 3/06 301B

G06F 3/06 301S

テーマコード (参考)

5 B 0 6 5

審査請求 未請求 請求項の数 10 O L (全 24 頁)

(21) 出願番号 特願2008-41480 (P2008-41480)

(22) 出願日 平成20年2月22日 (2008. 2. 22)

(71) 出願人 000005108

株式会社日立製作所

東京都千代田区丸の内一丁目6番6号

(74) 代理人 100093861

弁理士 大賀 眞司

(72) 発明者 大野 正智

神奈川県横浜市戸塚区吉田町292番地

株式会社日立アドバンスデジタル内

(72) 発明者 福田 秀明

神奈川県小田原市中里322番2号 株式

会社日立製作所RAIDシステム事業部内

(72) 発明者 五十嵐 康弘

神奈川県横浜市戸塚区吉田町292番地

株式会社日立アドバンスデジタル内

Fターム(参考) 5B065 BA01 CA02 CE12 CE14

(54) 【発明の名称】 ストレージ装置及びアクセス命令送信方法

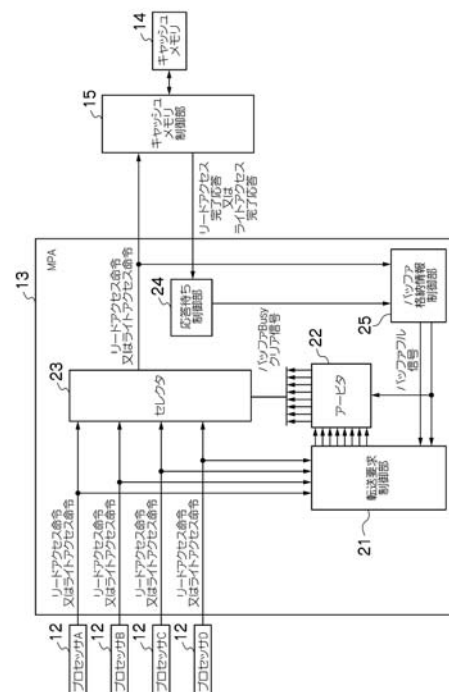
(57) 【要約】

【課題】本発明は、効率的にデータ転送を実施してスループットを向上させ得るストレージ装置を提案する。

【解決手段】ホスト装置から送信されるデータを格納するストレージ装置であって、前記ホスト装置から送信される前記データの送受信に関するアクセス要求に基づいて、キャッシュメモリ制御部に対して前記データの送受信に関するアクセス命令を送信する複数のプロセッサと、前記複数のプロセッサにより送信される前記アクセス命令を排他的に前記キャッシュメモリ制御部に送信するアクセス命令送信部とを備え、前記アクセス命令送信部は、応答が必要なアクセス命令を格納する複数の格納部を備え、すべての前記格納部に前記応答が必要なアクセス命令が格納されたときに、応答の必要がないアクセス命令のみを前記キャッシュメモリ制御部に送信する。

【選択図】 図2

图2



【特許請求の範囲】

【請求項 1】

ホスト装置から送信されるデータを格納するストレージ装置であって、

前記ホスト装置から送信される前記データの送受信に関するアクセス要求に基づいて、キャッシュメモリ制御部に対して前記データの送受信に関するアクセス命令を送信する複数のプロセッサと、

前記複数のプロセッサにより送信される前記アクセス命令を排他的に前記キャッシュメモリ制御部に送信するアクセス命令送信部と

を備え、

前記アクセス命令送信部は、

応答が必要なアクセス命令を格納する複数の格納部

を備え、

すべての前記格納部に前記応答が必要なアクセス命令が格納されたときに、応答の必要がないアクセス命令のみを前記キャッシュメモリ制御部に送信する

ことを特徴とするストレージ装置。

【請求項 2】

前記アクセス命令送信部は、

すべての前記格納部に前記応答が必要なアクセス命令が格納されたときに、ラウンドロビンによって優先順位順に、かつ排他的に、前記応答の必要がないアクセス命令をキャッシュメモリ制御部に送信する

ことを特徴とする請求項 1 に記載のストレージ装置。

【請求項 3】

前記アクセス命令送信部は、

すべての前記格納部に前記応答が必要なアクセス命令が格納されていないときに、ラウンドロビンによって優先順位順に、かつ排他的に、前記応答が必要なアクセス命令及び前記応答の必要がないアクセス命令をキャッシュメモリ制御部に送信する

ことを特徴とする請求項 3 に記載のストレージ装置。

【請求項 4】

前記アクセス命令送信部は、

前記キャッシュメモリ制御部から前記応答が必要なアクセス命令に対応するアクセス完了応答を受信すると、当該応答が必要なアクセス命令を前記格納部から削除する

ことを特徴とする請求項 3 に記載のストレージ装置。

【請求項 5】

前記アクセス命令送信部は、

すべての前記格納部に前記応答が必要なアクセス命令が格納されていたときには、前記格納部からの前記アクセス命令の削除により、前記応答が必要なアクセス命令及び前記応答の必要がないアクセス命令のキャッシュメモリ制御部への送信を再開する

ことを特徴とする請求項 4 に記載のストレージ装置。

【請求項 6】

ホスト装置から送信されるデータを格納するストレージ装置のアクセス命令送信方法であって、

複数のプロセッサが、前記ホスト装置から送信される前記データの送受信に関するアクセス要求に基づいて、キャッシュメモリ制御部に対して前記データの送受信に関するアクセス命令を送信する第 1 のステップと、

アクセス命令送信部が、前記複数のプロセッサにより送信される前記アクセス命令を排他的に前記キャッシュメモリ制御部に送信する第 2 のステップと

を備え、

前記第 2 のステップでは、

応答が必要なアクセス命令を格納する複数の格納部のすべての前記格納部に前記応答が必要なアクセス命令が格納されたときに、応答の必要がないアクセス命令のみを前記キャ

10

20

30

40

50

ッシュメモリ制御部に送信する

ことを特徴とするアクセス命令送信方法。

【請求項 7】

前記第 2 のステップでは、

すべての前記格納部に前記応答が必要なアクセス命令が格納されたときに、ラウンドロビンによって優先順位順に、かつ排他的に、前記応答の必要がないアクセス命令をキャッシュメモリ制御部に送信する

ことを特徴とする請求項 6 に記載のアクセス命令送信方法。

【請求項 8】

前記第 2 のステップでは、

すべての前記格納部に前記応答が必要なアクセス命令が格納されていないときに、ラウンドロビンによって優先順位順に、かつ排他的に、前記応答が必要なアクセス命令及び前記応答の必要がないアクセス命令をキャッシュメモリ制御部に送信する

ことを特徴とする請求項 7 に記載のアクセス命令送信方法。

【請求項 9】

前記第 2 のステップでは、

前記キャッシュメモリ制御部から前記応答が必要なアクセス命令に対応するアクセス完了応答を受信すると、当該応答が必要なアクセス命令を前記格納部から削除する

ことを特徴とする請求項 8 に記載のアクセス命令送信方法。

【請求項 10】

前記第 2 のステップでは、

すべての前記格納部に前記応答が必要なアクセス命令が格納されていたときには、前記格納部からの前記アクセス命令の削除により、前記応答が必要なアクセス命令及び前記応答の必要がないアクセス命令のキャッシュメモリ制御部への送信を再開する

ことを特徴とする請求項 9 に記載のアクセス命令送信方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、ストレージ装置及びアクセス命令送信方法に関し、例えば、複数のプロセッサからのアクセスを調停して排他的にアクセスさせるアービタを有するストレージ装置に適用して好適なものである。

【背景技術】

【0002】

従来、バスに接続された複数のバスマスタ（プロセッサ）がバスを共有して動作する場合、バスマスタは、まずバスの使用权を要求し、バス権を獲得した後、バスを使用していた。バスの使用权は、バスアービタ（アービタ）が管理し、バスアービタは、バスマスタからのバス使用权要求に対し、適切な優先順位でバス使用权を与えるよう構成されている。

【0003】

このような技術として、バスの使用効率を高めるために、スプリットトランザクションをサポートするバスを持つシステムにおいてバススレーブ側の受信バッファの段数に応じてバスマスタ内の DMA コマンドバッファの段数を最適化することでバスの有効利用を図るようにした例がある（例えば、特許文献 1 参照）。

【特許文献 1】特開平 10-011387 号公報

【発明の開示】

【発明が解決しようとする課題】

【0004】

ところで、複数のプロセッサが共通バスを通じて共通資源メモリにライトアクセスしようとした場合、アービタは、ライトアクセスを調停して排他し、排他処理を実施した後に、共通資源メモリのメモリ制御回路にライトアクセス命令を転送する。

10

20

30

40

50

【0005】

このとき、ライトアクセスが応答を待つライトアクセス（応答待ちライトアクセス）の仕様である場合、アービタは、応答を待つライト命令（応答待ちライト命令）の転送した分を、メモリ制御回路からの応答があるまで、バッファに記憶しておかなければならない。

【0006】

この結果、アービタは、ある一定以上の応答待ちライト命令を受け付けることができず、この間に応答を待つ必要がないライト命令（応答不要ライトアクセス命令）を受け付けることができるとしても、ライトアクセス命令を調停して排他しているため、当該応答不要ライトアクセス命令を転送することができず、効率的にデータ転送することができないという問題がある。

10

【0007】

本発明は以上の点を考慮してなされたもので、効率的にデータ転送を実施してスループットを向上させ得るストレージ装置及びアクセス命令送信方法を提案するものである。

【課題を解決するための手段】

【0008】

かかる課題を解決するために本発明においては、ホスト装置から送信されるデータを格納するストレージ装置であって、前記ホスト装置から送信される前記データの送受信に関するアクセス要求に基づいて、キャッシュメモリ制御部に対して前記データの送受信に関するアクセス命令を送信する複数のプロセッサと、前記複数のプロセッサにより送信される前記アクセス命令を排他的に前記キャッシュメモリ制御部に送信するアクセス命令送信部とを備え、前記アクセス命令送信部は、応答が必要なアクセス命令を格納する複数の格納部を備え、すべての前記格納部に前記応答が必要なアクセス命令が格納されたときに、応答の必要がないアクセス命令のみを前記キャッシュメモリ制御部に送信する。

20

【0009】

また、本発明においては、ホスト装置から送信されるデータを格納するストレージ装置のアクセス命令送信方法であって、複数のプロセッサが、前記ホスト装置から送信される前記データの送受信に関するアクセス要求に基づいて、キャッシュメモリ制御部に対して前記データの送受信に関するアクセス命令を送信する第1のステップと、アクセス命令送信部が、前記複数のプロセッサにより送信される前記アクセス命令を排他的に前記キャッシュメモリ制御部に送信する第2のステップとを備え、前記第2のステップでは、応答が必要なアクセス命令を格納する複数の格納部のすべての前記格納部に前記応答が必要なアクセス命令が格納されたときに、応答の必要がないアクセス命令のみを前記キャッシュメモリ制御部に送信する。

30

【0010】

従って、すべての格納部に応答が必要なアクセス命令が格納された場合に、応答の必要がないアクセス命令の転送が停止するのを未然かつ有効に防止し、当該応答の必要がないアクセス命令に対応するデータのみを転送することができる。

【発明の効果】

【0011】

本発明によれば、効率的にデータ転送を実施してスループットを向上させ得るストレージ装置及びアクセス命令送信方法を実現できる。

40

【発明を実施するための最良の形態】

【0012】

以下、本発明の一実施の形態を、図面を参照して詳細に説明する。なお、これにより本発明が限定されるものではない。

【0013】

図1は、本実施の形態によるストレージシステム1の構成を示している。ストレージシステム1は、ホスト装置2及びストレージ装置3がSAN（Storage Area Network）4を介して接続されることにより構成されている。

50

【0014】

ホスト装置2は、CPU (Central Processing Unit)、メモリ等の情報処理資源 (図示せず) を備えたコンピュータ装置であり、例えば、パーソナルコンピュータや、ワークステーション、メインフレーム等から構成される。また、ホスト装置2は、SAN4と接続するためのホストバスアダプタ (FC HBA) (図示せず) を備えて構成される。さらに、ホスト装置2は、キーボード、スイッチやポインティングデバイス、マイクロフォン等の情報入力装置 (図示せず) と、モニタディスプレイやスピーカ等の情報出力装置 (図示せず) とを備えて構成される。

【0015】

SAN4は、ホスト装置2が提供する記憶資源におけるデータの管理単位であるブロックを単位として、ホスト装置2とストレージ装置3との間で、コマンドやデータ等の授受を行う。この場合、ホスト装置2とストレージ装置3との間で行われる通信プロトコルは、ファイバチャネルプロトコルである。

10

【0016】

なお、ホスト装置2とストレージ装置3との間は、必ずしもSAN4を介して接続する必要はなく、LAN等を介して接続するようにしてもよい。例えば、LANを介して接続する場合には、例えばTCP/IP (Transmission Control Protocol/Internet Protocol) に従ってコマンドやデータ等の授受を行う。また、LANを介して接続する場合には、ホストバスアダプタに代えて、例えば、LAN対応のネットワークカード等を適用することができる。

20

【0017】

ストレージ装置3は、データを記憶するための複数のハードディスクドライブ (HDD : Hard Disk Drive) 5 及び当該ハードディスクドライブ5に対するデータの入出力を制御するストレージコントローラ6を備えて構成される。

【0018】

ストレージコントローラ6は、複数のチャネル制御部11、複数のプロセッサ12が各々接続された複数のMPA (Microprocessor Adapter) 13、キャッシュメモリ14が各々接続された複数のキャッシュメモリ制御部15、エクスペンダ (Expander) 16が各々接続された複数のHDD制御部17及び管理装置18が相互結合部19を介して接続されることにより構成されている。また、ストレージコントローラ6は、ハードディスクドライブ5とエクスペンダ16を介して接続されている。

30

【0019】

チャネル制御部11は、例えば、ホスト装置2から受信した要求 (例えば、リードアクセス要求やライトアクセス要求) をプロセッサ13に通知して、プロセッサ12からの命令 (例えば、リードアクセス (読み出し) 命令やライトアクセス (書き込み) 命令) により、キャッシュメモリ制御部15及びHDD制御部17を介して、ホスト装置2とキャッシュメモリ14との間のデータの転送を行う。

【0020】

プロセッサ12は、ストレージコントローラ6全体を制御し、例えば、チャネル制御部11から通知された要求を解釈して、チャネル制御部11、キャッシュメモリ制御部15及びHDD制御部17に命令を通知する。また、プロセッサ12は、ハードディスクドライブ5に対して、RAID (Redundant Arrays of Independent Disks) 制御を行うことにより、ストレージ装置3の信頼性、可用性及び性能を向上させることができる。

40

【0021】

この場合、プロセッサ12は、ハードディスクドライブ5をRAID方式で運用する。プロセッサ12は、1又は複数のハードディスクドライブ5により提供される物理的な記憶領域 (RAIDグループ) 上に、1又は複数の論理的なボリューム (以下、これを論理ボリュームと呼ぶ) を設定する。そして、データは、この論理ボリューム内に所定の大きさのブロック (以下、これを論理ブロックと呼ぶ) 単位で記憶される。

【0022】

50

各論理ボリュームには、それぞれ固有の識別子（以下、これをL U（Logical Unit number）と呼ぶ）が付与される。本実施の形態の場合、データの入出力は、このL Uと、各論理ブロックにそれぞれ付与されるその論理ブロックに固有の番号（LBA :Logical Block Address）とを組み合わせたものをアドレスとして、当該アドレスを指定して行われる。

【0023】

また、プロセッサ12には、例えば、キャッシュメモリ14の管理情報やストレージ装置3の構成情報等の情報を格納するためのメモリが内蔵されている。また、プロセッサ12のメモリには、種々の制御プログラムや管理テーブルが格納されている。

【0024】

M P A 1 3 は、複数のプロセッサ12から転送されるリードアクセス命令又はライトアクセス命令を排他的にキャッシュメモリ14のキャッシュメモリ制御部15に転送する。

【0025】

キャッシュメモリ制御部15は、キャッシュメモリ14を制御し、チャネル制御部11又はH D D制御部17から転送されたデータをキャッシュメモリ14に格納する。キャッシュメモリ14は、ハードディスクドライブ5に格納されるべきデータやホスト装置2に転送すべきデータを一時的に格納するためのメモリである。

【0026】

H D D制御部17は、ハードディスクドライブ5を制御し、プロセッサ12からの命令により、エクスパンダ16を介して、ハードディスクドライブ5とキャッシュメモリ14との間のデータの転送を行う。また、H D D制御部17は、ハードディスクドライブ5に対して、R A I D（Redundant Arrays of Independent Disks）制御を行うことにより、ストレージ装置3の信頼性、可用性及び性能を向上させることができる。

【0027】

管理装置18は、ストレージ装置3全体の動作を操作する管理端末であり、例えば、ノート型のパーソナルコンピュータから構成される。管理装置18は、オペレータの操作に応じて各種処理の指令を行う。オペレータは、例えば、管理装置18を操作してストレージ装置3の各種状態を管理装置18の表示部に表示させることにより、ストレージ装置3の状態を確認することができる。

【0028】

図2は、M P A 1 3の構成を示している。M P A 1 3は、転送要求制御部21、アービタ22、セクタ23、応答待ち制御部24及びバッファ格納情報制御部24を備えて構成される。なお、この場合、M P A 1 3は、プロセッサA（12）～プロセッサD（12）に接続されている。

【0029】

転送要求制御部21は、プロセッサA～プロセッサDから転送されたリードアクセス命令又はライトアクセス命令を解析して、応答待ちありの命令であるか否かを判断し、応答待ちアクセス要求又は応答待ちなしアクセス要求をアービタ22に転送する。また、転送要求制御部21は、バッファ格納情報制御部24からバッファフル信号（後述）が送信されている場合には、その間、応答待ちアクセス要求の転送を停止し、応答待ちなしアクセス要求のみをアービタ22に転送する。

【0030】

アービタ22は、転送要求制御部21から転送された応答待ちアクセス要求又は応答待ちなしアクセス要求に基づいて、ラウンドロビンによって優先順位順に、かつ排他的に、キャッシュメモリ制御部15に転送するリードアクセス命令又はライトアクセス命令を、セクタ23に指示する。また、アービタ22は、バッファ格納情報制御部24からバッファフル信号（後述）が送信されている場合には、その間、応答待ちなしアクセス要求のみに基づいて、ラウンドロビンによって優先順位順に、かつ排他的に、キャッシュメモリ制御部15に転送するリードアクセス命令又はライトアクセス命令を、セクタ23に指示する。

【0031】

10

20

30

40

50

セレクタ 2 3 は、アービタ 2 2 からの転送指示に基づいて、プロセッサ A～プロセッサ D から転送されたリードアクセス命令又はライトアクセス命令をキャッシュメモリ制御部 1 5 に転送する。

【0032】

応答待ち制御部 2 4 は、キャッシュメモリ制御部 1 5 から転送されたリードアクセス完了応答又はライトアクセス完了応答を解析し、完了応答のあったバッファ格納情報制御部 2 5 内の格納部 6 2 のフラグ 6 2 B (後述) を「1」から「0」に変更するバッファ BUSY クリア信号をバッファ格納情報制御部 2 5 に送信する。なお、この場合、バッファ格納情報制御部 2 5 内の格納部 6 2 のバッファ 6 2 A (後述) にリードアクセス命令又はライトアクセス命令が格納されている場合には、フラグ 6 2 B が「1」となり、リードアクセス命令又はライトアクセス命令が格納されていない場合には、フラグ 6 2 B が「0」となる。

10

【0033】

バッファ格納情報制御部 2 5 は、プロセッサ A～プロセッサ D から転送されたリードアクセス命令又はライトアクセス命令を当該バッファ格納情報制御部 2 5 内の格納部 6 2 のバッファ 6 2 A に格納してフラグ 6 2 B を「0」から「1」に変更する。そして、バッファ格納情報制御部 2 5 は、すべてのバッファ 6 2 A にリードアクセス命令又はライトアクセス命令が格納され、すべてのフラグ 6 2 B が「1」に変更されたときに、変更されている間、バッファ 6 2 A にリードアクセス命令又はライトアクセス命令を格納することができない旨を通知するバッファフル信号を転送要求制御部 21 及びアービタ 2 2 に送信する。

20

【0034】

図 3 は、転送要求制御部 2 1 の構成を示している。転送要求制御部 2 1 は、転送制御シーケンサ 3 1 A～3 1 D、データ解析シーケンサ 3 2 A～3 2 D、AND 回路 3 3 A～3 3 D、AND 回路 3 4 A～3 4 D 及びインバータ 3 5 A～3 5 D を備えて構成される。

【0035】

転送制御シーケンサ 3 1 A は、プロセッサ A からリードアクセス命令又はライトアクセス命令を受信すると、当該リードアクセス命令又はライトアクセス命令の転送を要求する転送要求リクエストを AND 回路 3 3 A 及び AND 回路 3 4 A に送信する。

【0036】

データ解析シーケンサ 3 2 A は、プロセッサ A からリードアクセス命令又はライトアクセス命令を受信すると、当該リードアクセス命令又はライトアクセス命令内のコマンドを解析し、応答ありのリードアクセス命令又はライトアクセス命令であるか、応答なしのリードアクセス命令又はライトアクセス命令であるかをチェックする。

30

【0037】

そして、データ解析シーケンサ 3 2 A は、応答ありのリードアクセス命令又はライトアクセス命令である場合には、応答待ちありである旨の信号を AND 回路 3 4 A に送信し、応答なしのリードアクセス命令又はライトアクセス命令である場合には、応答待ちなしである旨の信号を AND 回路 3 3 A に送信する。

【0038】

AND 回路 3 3 A は、転送制御シーケンサ 3 1 A から転送要求リクエストを受信すると共に、データ解析シーケンサ 3 2 A から応答待ちなしである旨の信号を受信すると、応答待ちなしアクセス要求を応答あり／応答なしラウンドロビンシーケンサ 4 1 及び応答なしラウンドロビンシーケンサ 4 2 に転送する。

40

【0039】

AND 回路 3 4 A は、転送制御シーケンサ 3 1 A から転送要求リクエストを受信し、データ解析シーケンサ 3 2 A から応答待ちありである旨の信号を受信すると共に、停止回路 3 5 A からバッファレディ信号 (後述) を受信すると、応答待ちアクセス要求をアービタ 2 2 の応答あり／応答なしラウンドロビンシーケンサ 4 1 に転送する。

【0040】

50

インバータ 3 5 A は、リードアクセス命令又はライトアクセス命令をまだバッファ 6 2 A に格納することができる旨を通知するバッファレディ信号を A N D 回路 3 4 A に送信し、バッファ格納情報制御部 2 5 の A N D 回路 6 3 (後述) からバッファフル信号が受信されている間、バッファレディ信号の送信を停止する。これにより、転送要求制御部 2 1 は、リードアクセス命令又はライトアクセス命令をバッファ 6 2 A に格納することができない場合には、応答待ちアクセス要求の転送を停止することができる。

【0041】

転送制御シーケンサ 3 1 B ~ 3 1 D、データ解析シーケンサ 3 2 B ~ 3 2 D、A N D 回路 3 3 B ~ 3 3 D、A N D 回路 3 4 B ~ 3 4 D 及びインバータ 3 5 B ~ 3 5 D は、プロセス B ~ D からリードアクセス命令又はライトアクセス命令を受信すると、上述の転送制御シーケンサ 3 1 A、データ解析シーケンサ 3 2 A、A N D 回路 3 3 A、A N D 回路 3 4 A 及びインバータ 3 5 A と同様の処理を行う。

【0042】

図 4 は、アービタ 2 2 の構成を示している。アービタ 2 2 は、応答あり／応答なしラウンドロビンシーケンサ 4 1、応答なしラウンドロビンシーケンサ 4 2 及びセクタ 4 3 を備えて構成される。

【0043】

応答あり／応答なしラウンドロビンシーケンサ 4 1 は、転送要求制御部 2 1 の A N D 回路 3 3 A ~ 3 3 D から転送される応答待ちなしアクセス要求及び A N D 回路 3 4 A ~ 3 4 D から転送される応答待ちアクセス要求により、ラウンドロビンによって優先順位順に、かつ排他的に、キャッシュメモリ制御部 1 5 に転送するリードアクセス命令又はライトアクセス命令を決定し、当該転送指示をセクタ 4 3 に転送する。

【0044】

応答なしラウンドロビンシーケンサ 4 2 は、転送要求制御部 2 1 の A N D 回路 3 3 A ~ 3 3 D から転送される応答待ちなしアクセス要求により、ラウンドロビンによって優先順位順に、かつ排他的に、キャッシュメモリ制御部 1 5 に転送するリードアクセス命令又はライトアクセス命令を決定し、転送指示をセクタ 4 3 に転送する。

【0045】

セクタ 4 3 は、バッファ格納情報制御部 2 5 の A N D 回路 6 3 (後述) からバッファフル信号を受信している間、応答なしラウンドロビンシーケンサ 4 2 から転送される転送指示をセクタ 2 3 に転送し、当該 A N D 回路 6 3 からバッファフル信号の受信が停止すると、応答あり／応答なしラウンドロビンシーケンサ 4 1 から転送される転送指示をセクタ 2 3 に転送する。これにより、アービタ 2 2 は、リードアクセス命令又はライトアクセス命令をバッファ 6 2 A に格納することができない場合には、応答待ちアクセス要求の転送を再開することができると共に、応答あり／応答なしラウンドロビンシーケンサ 4 1 からの転送指示の転送を停止することができる。

【0046】

図 5 は、応答待ち制御部 2 4 の構成を示している。応答待ち制御部 2 4 は、データ解析シーケンサ 5 1 及び応答制御シーケンサ 5 2 を備えて構成される。

【0047】

データ解析シーケンサ 5 1 は、キャッシュメモリ制御部 1 5 からリードアクセス完了応答又はライトアクセス完了応答を受信すると、当該リードアクセス完了応答又はライトアクセス完了応答内のコマンドを解析し、応答ありのリードアクセス完了応答又はライトアクセス完了応答であるか、応答なしのリードアクセス完了応答又はライトアクセス完了応答であるかをチェックする。

【0048】

そして、データ解析シーケンサ 5 1 は、応答ありのリードアクセス完了応答又はライトアクセス完了応答である場合には、応答待ちありである旨の信号を応答制御シーケンサ 5 2 に送信する。

【0049】

10

20

30

40

50

応答制御シーケンサ 5 2 は、データ解析シーケンサ 5 1 から送信される応答待ちありである旨の信号を受信すると、当該応答待ちありである旨の信号のリードアクセス完了応答又はライトアクセス完了応答に対応するリードアクセス命令又はライトアクセス命令が格納されている格納部 6 2 にバッファ B U S Y クリア信号を送信する。

【0050】

図 6 は、バッファ格納情報制御部 2 5 の構成を示している。バッファ格納情報制御部 2 5 は、転送制御シーケンサ 6 1、格納部 6 2 並びに A N D 回路 6 3 を備えて構成される。また、バッファ格納情報制御部 2 5 は、「0」～「n」の「n+1」個の格納部 6 2 を備えている。

【0051】

転送制御シーケンサ 6 1 は、プロセッサ A ～プロセッサ D からセクタ 2 3 を介して転送されたリードアクセス命令又はライトアクセス命令を受信すると、当該リードアクセス命令又はライトアクセス命令を空いている格納部 6 2 のバッファ 6 2 A に転送すると共に、リードアクセス命令又はライトアクセス命令が格納された格納部 6 2 のフラグ 6 2 B を「0」から「1」に変更するバッファ B U S Y セット信号をフラグ 6 2 B に送信する。

【0052】

格納部 6 2 は、リードアクセス命令又はライトアクセス命令を格納するバッファ 6 2 A 及びバッファ 6 2 A にリードアクセス命令又はライトアクセス命令が格納されているか否かを判断するためのフラグ 6 2 B を備えて構成される。格納部 6 2 は、転送制御シーケンサ 6 1 からリードアクセス命令又はライトアクセス命令を受信すると、当該リードアクセス命令又はライトアクセス命令を指定されたバッファ 6 2 A に格納し、転送制御シーケンサ 6 1 からバッファ B U S Y セット信号を受信すると、リードアクセス命令又はライトアクセス命令が格納された格納部 6 2 のフラグ 6 2 B を「0」から「1」に変更する。

【0053】

また、格納部 6 2 は、リードアクセス命令又はライトアクセス命令が格納された格納部 6 2 のフラグ 6 2 B を「0」から「1」に変更すると、「1」に変更されている間、リードアクセス命令又はライトアクセス命令がバッファ 6 2 A に格納された旨の信号であるバッファ B U S Y 信号を A N D 回路 6 3 に送信する。

【0054】

一方、格納部 6 2 は、応答待ち制御部 2 4 の応答制御シーケンサ 5 2 からバッファ B U S Y クリア信号が送信されると、リードアクセス命令又はライトアクセス命令が格納された格納部 6 2 のフラグ 6 2 B を「1」から「0」に変更し、バッファ 6 2 A に格納されているリードアクセス命令又はライトアクセス命令を削除する。

【0055】

また、格納部 6 2 は、リードアクセス命令又はライトアクセス命令が格納された格納部 6 2 のフラグ 6 2 B を「1」から「0」に変更すると、バッファ B U S Y 信号の A N D 回路 6 3 への送信を停止する。

【0056】

A N D 回路 6 3 は、「0」～「n」のすべての格納部 6 2 からバッファ B U S Y 信号を受信すると、当該すべてのバッファ B U S Y 信号が受信されている間、バッファフル信号を転送要求制御部 2 1 及びアービタ 2 2 に送信する。また、A N D 回路 6 3 は、すべての格納部 6 2 からバッファ B U S Y 信号を受信していないときには、バッファフル信号の送信を停止する。

【0057】

これにより、バッファ格納情報制御部 2 5 は、リードアクセス命令又はライトアクセス命令をバッファ 6 2 A に格納することができるようになった場合には、応答あり／応答なしラウンドロビンシーケンサ 4 1 から転送する転送指示を再開することができる。

【0058】

次に、本実施の形態におけるストレージシステム 1 のストレージコントローラ 6 によるアクセス命令転送処理について説明する。

10

20

30

40

50

【0059】

図7は、このストレージシステム1のストレージコントローラ6のアクセス命令転送処理に関する、ストレージコントローラ6のMPA13の具体的な処理手順を示したフローチャートの一例である。

【0060】

MPA13は、プロセッサA～D(12)からリードアクセス命令又はライトアクセス命令を受信すると、図7に示すアクセス命令転送処理手順RT1に従って、データ解析シーケンサ32A～32Dにより当該リードアクセス命令又はライトアクセス命令内のコマンドを解析する(SP1)。

【0061】

続いて、MPA13は、データ解析シーケンサ32A～32Dにより解析したリードアクセス命令又はライトアクセス命令内のコマンドが応答ありのリードアクセス命令又はライトアクセス命令であるか否かをチェックする(SP2)。

【0062】

そして、MPA13は、解析したリードアクセス命令又はライトアクセス命令内のコマンドが応答ありのリードアクセス命令又はライトアクセス命令でない場合(SP2:NO)には、ステップSP4に進む。

【0063】

これに対して、MPA13は、解析したリードアクセス命令又はライトアクセス命令内のコマンドが応答ありのリードアクセス命令又はライトアクセス命令である場合(SP2:YES)には、インバータ35A～35Dにより格納部62のバッファ62Aがフルであるか否か、すなわち、インバータ35A～35Dにバッファフル信号が送信されているか否かをチェックする(SP3)。

【0064】

そして、MPA13は、格納部62のバッファ62Aがフルである場合(SP3:YES)には、格納部62のバッファ62Aがフルでなくなるのを待機モードで待ち受ける(SP3)。

【0065】

これに対して、MPA13は、格納部62のバッファ62Aがフルである格納部62のバッファ62Aがフルでない場合(SP3:NO)には、AND回路33A～33D又はAND回路34A～34Dにより応答待ちアクセス要求又は応答なしアクセス要求をアービタ22の応答あり／応答なしラウンドロビンシーケンサ41及び又は応答なしラウンドロビンシーケンサ42に転送する(SP4)。

【0066】

続いて、MPA13は、応答あり／応答なしラウンドロビンシーケンサ41及び又は応答なしラウンドロビンシーケンサ42によりラウンドロビンによって優先順位順に、かつ排他的に、キャッシュメモリ制御部15に転送するリードアクセス命令又はライトアクセス命令を決定し、セクタ43により転送指示をセクタ23に転送する(SP5)。

【0067】

続いて、MPA13は、転送指示が応答ありのリードアクセス命令又はライトアクセス命令である場合には、セクタ23により転送された応答ありのリードアクセス命令又はライトアクセス命令を、転送制御シーケンサ61により格納部62のバッファ62Aに格納して、フラグ62Bを「0」から「1」に変更する(SP6)。

【0068】

続いて、MPA13は、セクタ23により転送指示のリードアクセス命令又はライトアクセス命令をキャッシュメモリ制御部15に転送する(SP7)。

【0069】

やがて、MPA13は、この後、この図7に示すアクセス命令転送処理手順RT1を終了する(SP8)。

【0070】

10

20

30

40

50

次に、本実施の形態におけるストレージシステム 1 のストレージコントローラ 6 によるアクセス完了応答転送処理について説明する。

【0071】

図 8 は、このストレージシステム 1 のストレージコントローラ 6 のアクセス完了応答転送処理に関する、ストレージコントローラ 6 の M P A 1 3 の具体的な処理手順を示したフローチャートの一例である。

【0072】

M P A 1 3 は、キャッシュメモリ制御部 1 5 からリードアクセス完了応答又はライトアクセス完了応答を受信すると、図 8 に示すアクセス完了応答転送処理手順 R T 2 に従って、データ解析シーケンサ 5 1 により当該リードアクセス完了応答又はライトアクセス完了応答内のコマンドを解析する (S P 1 1) 。

10

【0073】

続いて、M P A 1 3 は、データ解析シーケンサ 5 1 により解析したリードアクセス完了応答又はライトアクセス完了応答内のコマンドが応答ありのリードアクセス完了応答又はライトアクセス完了応答であるか否かをチェックする (S P 1 2) 。

【0074】

そして、M P A 1 3 は、解析したリードアクセス完了応答又はライトアクセス完了応答内のコマンドが応答ありのリードアクセス完了応答又はライトアクセス完了応答でない場合 (S P 1 2 : N O) には、ステップ S P 1 4 に進む。

【0075】

20

これに対して、M P A 1 3 は、解析したリードアクセス完了応答又はライトアクセス完了応答内のコマンドが応答ありのリードアクセス完了応答又はライトアクセス完了応答である場合 (S P 1 2 : Y E S) には、格納部 6 2 によりフラグ 6 2 B を「1」から「0」に変更して、当該リードアクセス完了応答又はライトアクセス完了応答に対応するリードアクセス命令又はライトアクセス命令を、格納されているバッファ 6 2 A から削除する (S P 1 3) 。

【0076】

続いて、M P A 1 3 は、A N D 回路 6 3 により格納部 6 2 のバッファ 6 2 A がフルであったか否か、すなわち、A N D 回路 6 3 にすべての格納部 6 2 のバッファ B U S Y 信号が送信されていたか否かをチェックする (S P 1 4) 。

30

【0077】

そして、M P A 1 3 は、格納部 6 2 のバッファ 6 2 A がフルでなかった場合 (S P 1 4 : N O) には、この後、この図 8 に示すアクセス完了応答転送処理手順 R T 2 を終了する (S P 1 6) 。

【0078】

これに対して、M P A 1 3 は、格納部 6 2 のバッファ 6 2 A がフルであった場合 (S P 1 4 : Y E S) には、A N D 回路 6 3 により送信されているバッファフル信号の送信を停止し (S P 1 5) 、この後、この図 8 に示すアクセス完了応答転送処理手順 R T 2 を終了する (S P 1 6) 。

【0079】

40

なお、ストレージシステム 1 では、アクセス命令又はアクセス完了応答に基づいて、チャンネル制御部 1 1 及びキャッシュメモリ制御部 1 5 間において、リードアクセス命令又はライトアクセス命令のデータの転送が行われるようになされている。

【0080】

図 9 は、他の実施の形態によるストレージシステム 1 の構成を示している。本実施の形態におけるストレージシステム 1 は、バッファ格納情報制御部 2 5 がバッファフル信号を転送要求制御部 2 1 に送信し、転送要求制御部 2 1 が応答待ちアクセス要求の転送を停止した場合について述べたが、本発明はこれに限らず、図 9 のように、バッファ格納情報制御部 2 5 がバッファフル信号をアービタ 2 2 にのみ送信し、アービタにより応答待ちアクセス要求の転送を停止するようにしても良く、この他種々の形態に適用することができる

50

。

【0081】

図10は、他の実施の形態によるストレージシステム1における転送要求制御部71の構成を示している。転送要求制御部71は、例えば、インバータ35A～35Dを備えていないことを除いて、本実施の形態による転送要求制御部21と同様に構成される。

【0082】

図11は、他の実施の形態によるストレージシステム1におけるアービタ72の構成を示している。アービタ72は、応答あり／応答なしラウンドロビンシーケンサ73を備えて構成される。応答あり／応答なしラウンドロビンシーケンサ73は、バッファフル信号の送信の有無に基づいて、転送要求制御部71のAND回路33A～33Dから転送される応答待ちなしアクセス要求及びAND回路34A～34Dから転送される応答待ちアクセス要求により、ラウンドロビンによって優先順位順に、かつ排他的に、キャッシュメモリ制御部15に転送するリードアクセス命令又はライトアクセス命令を決定し、当該転送指示をセクタ23に転送する。

10

【0083】

このようにして、ストレージシステム1では、プロセッサA～プロセッサDから転送されたリードアクセス命令又はライトアクセス命令を当該バッファ格納情報制御部25内の格納部62のバッファ62Aに格納し、すべてのバッファ62Aにリードアクセス命令又はライトアクセス命令が格納されたときに、応答待ちなしのリードアクセス命令又はライトアクセス命令のみをキャッシュメモリ制御部15に転送する。

20

【0084】

従って、すべてのバッファ62Aに応答待ちありのリードアクセス命令又はライトアクセス命令が格納された場合に、応答待ちなしのリードアクセス命令又はライトアクセス命令の転送が停止するのを未然かつ有効に防止し、当該応答待ちなしのリードアクセス命令又はライトアクセス命令に対応するデータのみを転送することができる。

【0085】

なお、本実施の形態においては、リードアクセス命令又はライトアクセス命令を送信した場合について述べたが、本発明はこれに限らず、例えば、この他種々の他のアクセス命令についても適用することができる。

【0086】

また、本実施例においては、図2のMPA13を機能ブロック図として示した場合について述べたが、本発明はこれに限らず、より具体的には、一連の処理を、それぞれの機能を有するハードウェアによって実行することにより、各処理を実施するようにしても良く、MPA13の制御部等が制御プログラムを実行することにより各処理を実行するようにしても良く、この他種々の形態に適用することができる。

30

【産業上の利用可能性】

【0087】

本発明は、複数のプロセッサからのアクセスを調停して排他的にアクセスさせるアービタを有するストレージ装置に広く適用することができる。

【図面の簡単な説明】

40

【0088】

【図1】本実施の形態によるストレージシステムの概略的な構成を示すブロック図である。

。

【図2】MPAの概略的な構成を示すブロック図である。

【図3】転送要求制御部の概略的な構成を示すブロック図である。

【図4】アービタの概略的な構成を示すブロック図である。

【図5】応答待ち制御部の概略的な構成を示すブロック図である。

【図6】バッファ格納情報制御部の概略的な構成を示すブロック図である。

【図7】アクセス命令転送処理手順を示すフローチャートである。

【図8】アクセス完了応答転送処理手順を示すフローチャートである。

50

【図 9】他の実施の形態によるストレージシステムの概略的な構成を示すブロック図である。

【図 10】他の実施の形態による転送要求制御部の概略的な構成を示すブロック図である。

【図 11】他の実施の形態によるアービタの概略的な構成を示すブロック図である。

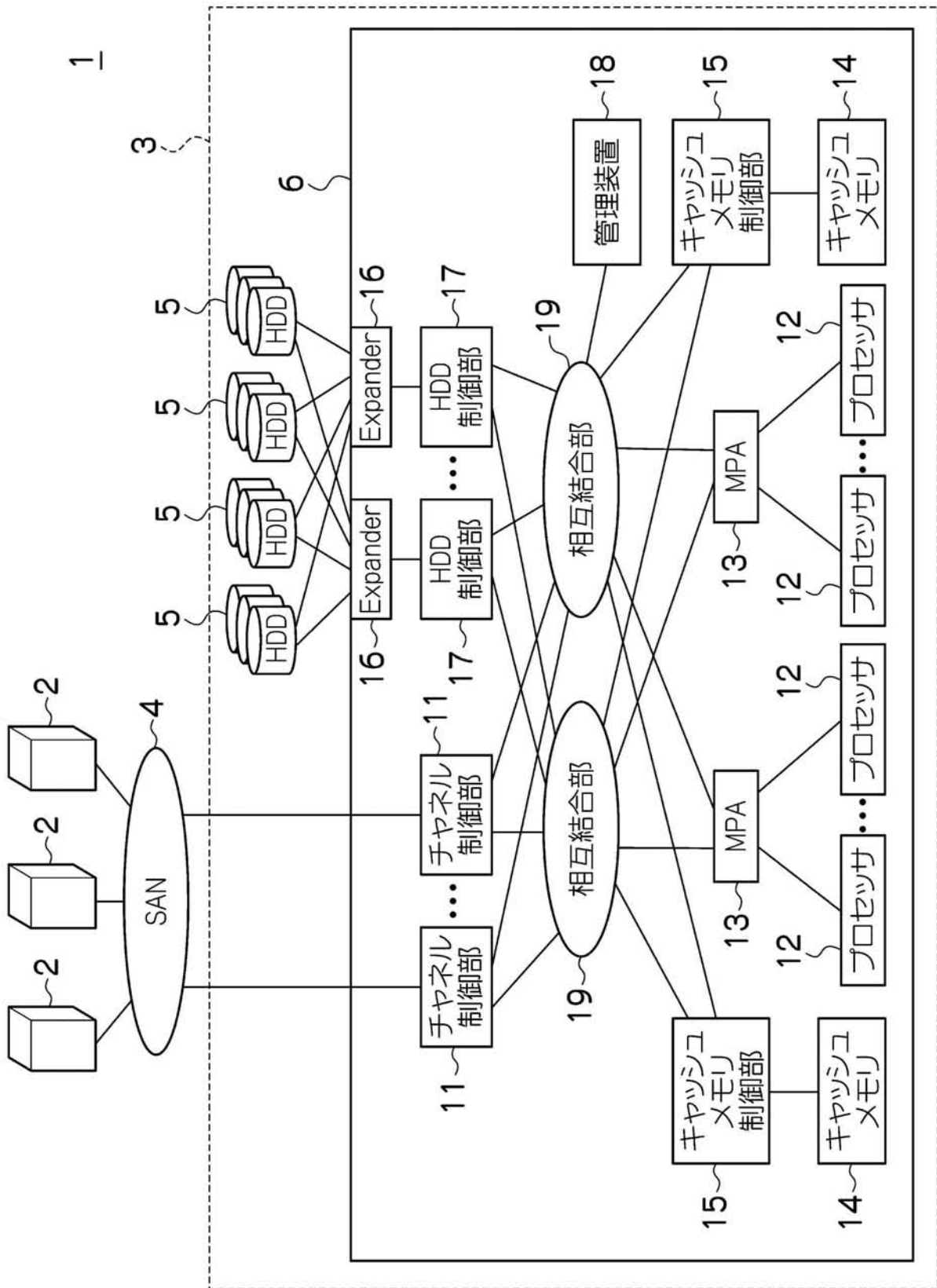
【符号の説明】

【0089】

1 ……ストレージシステム、2 ……ホスト装置、3 ……ストレージ装置、4 ……S A N、
5 ……ハードディスクドライブ、6 ……ストレージコントローラ、11 ……チャンネル制御
部、12 ……プロセッサ、13 ……M P A、14 ……キャッシュメモリ、15 ……キャッ
シュメモリ制御部、17 ……H D D制御部、18 ……管理装置、21、71 ……転送要求
制御部、22、72 ……アービタ、23、43 ……セクタ、24 ……応答待ち制御部、
25 ……バッファ格納情報制御部、31 ……転送要求シーケンサ、32、51 ……データ
解析シーケンサ、33、34、63 ……A N D回路、41 ……応答あり／応答なしラウン
ドロビンシーケンサ、42 ……応答なしラウンドロビンシーケンサ、52 ……応答制御シ
ーケンサ、61 ……転送制御シーケンサ、62 ……格納部、62 A ……バッファ、62 B
……フラグ

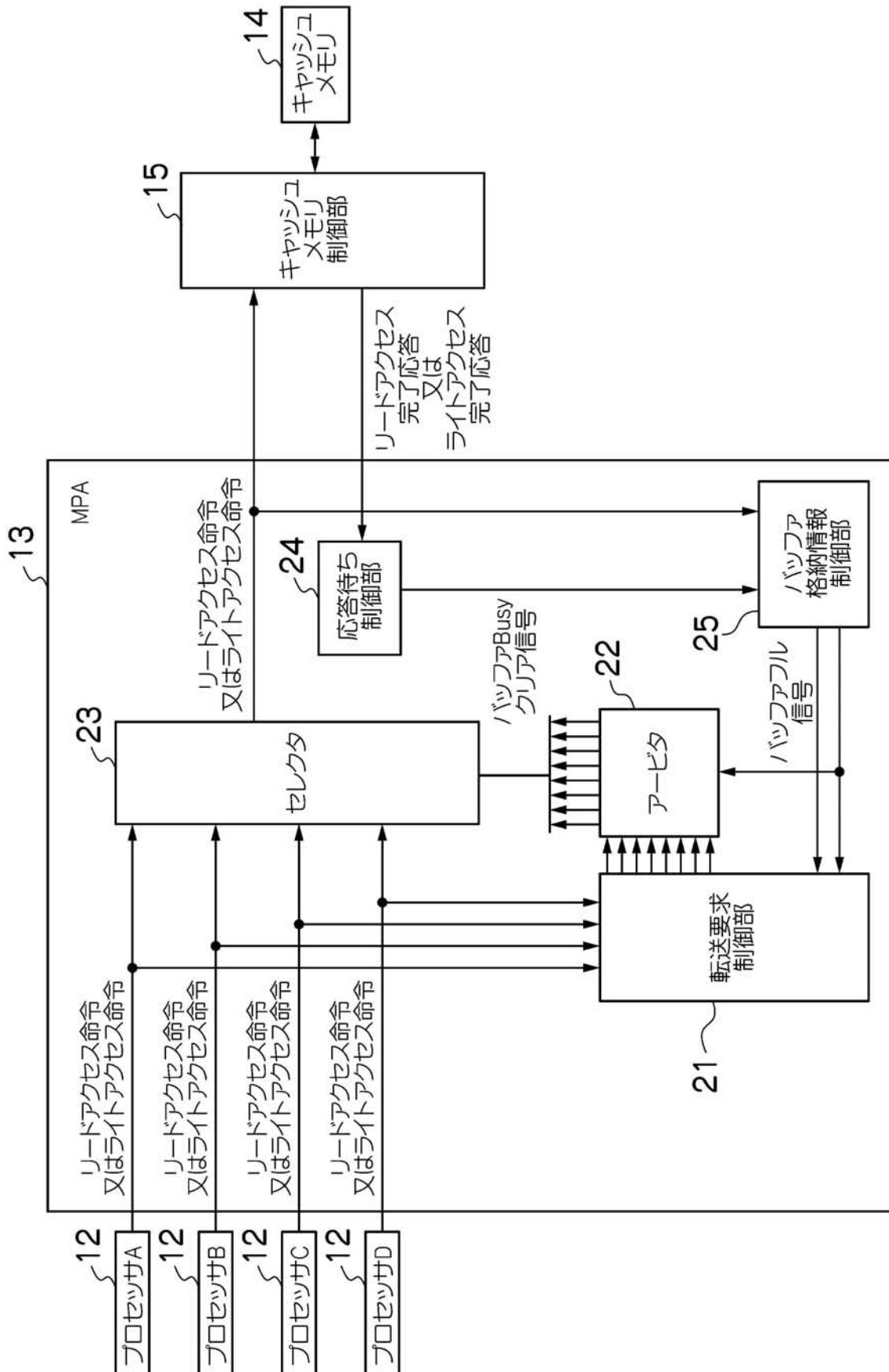
【図 1】

図 1



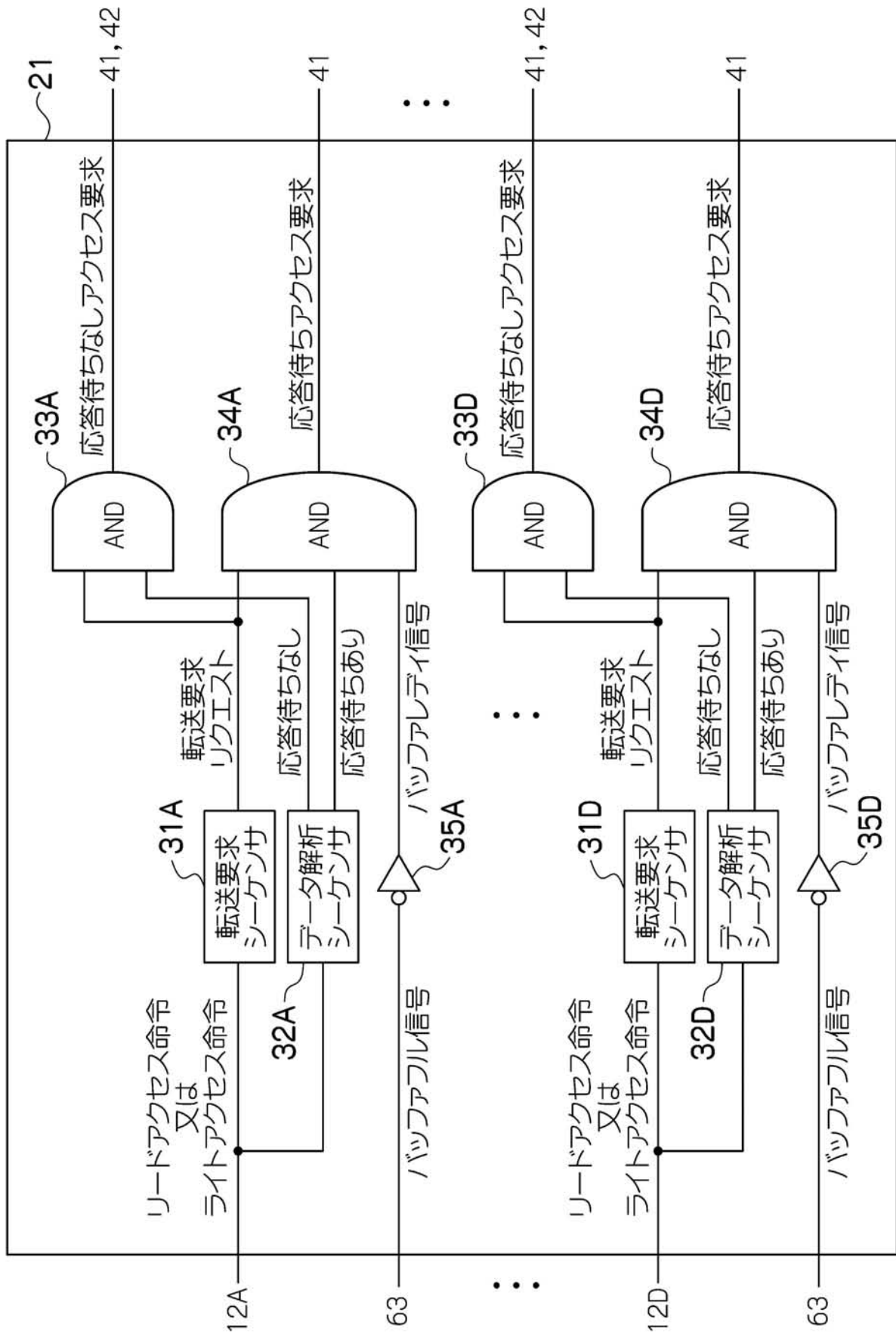
【図2】

図2



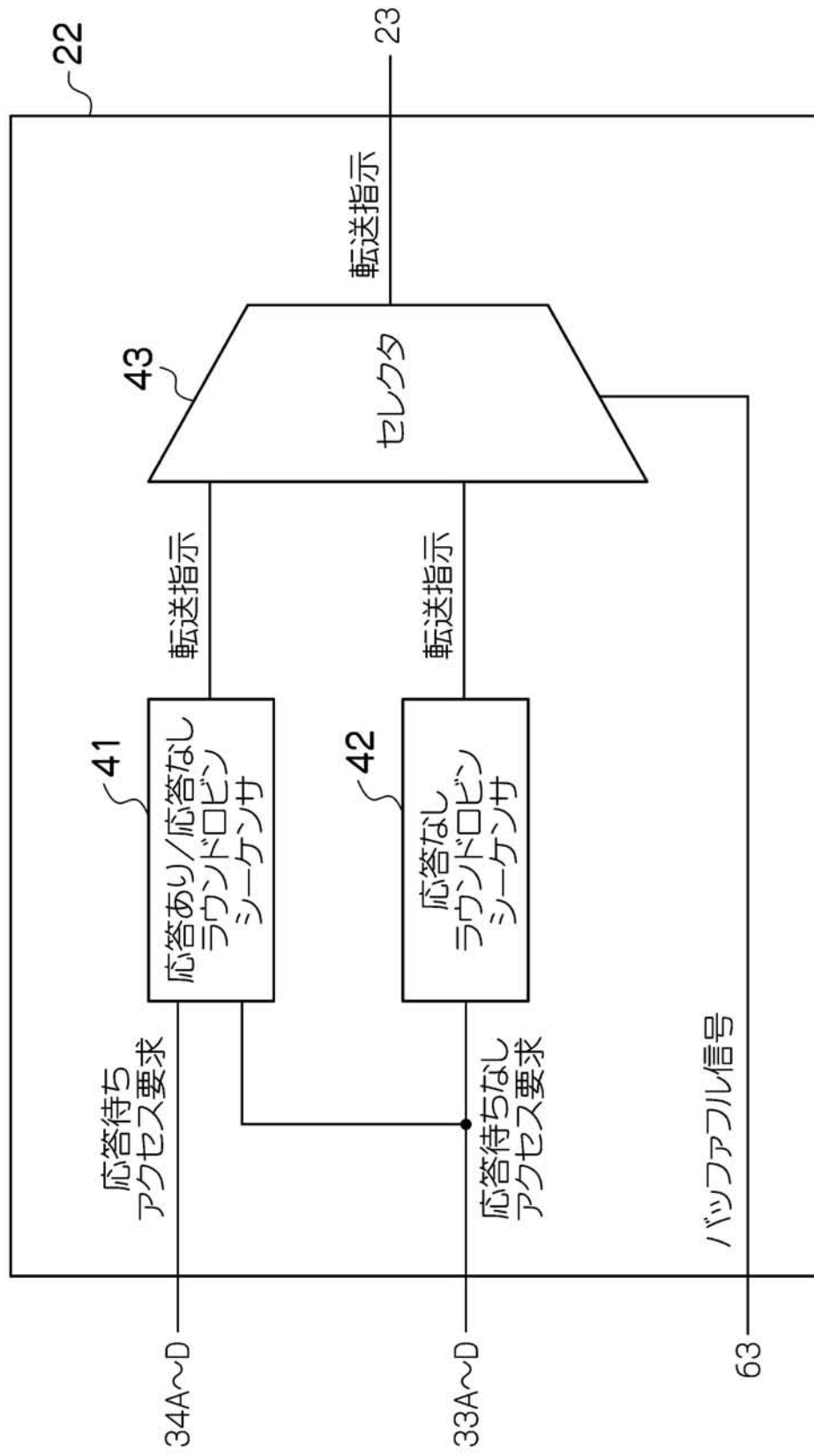
【図 3】

図 3



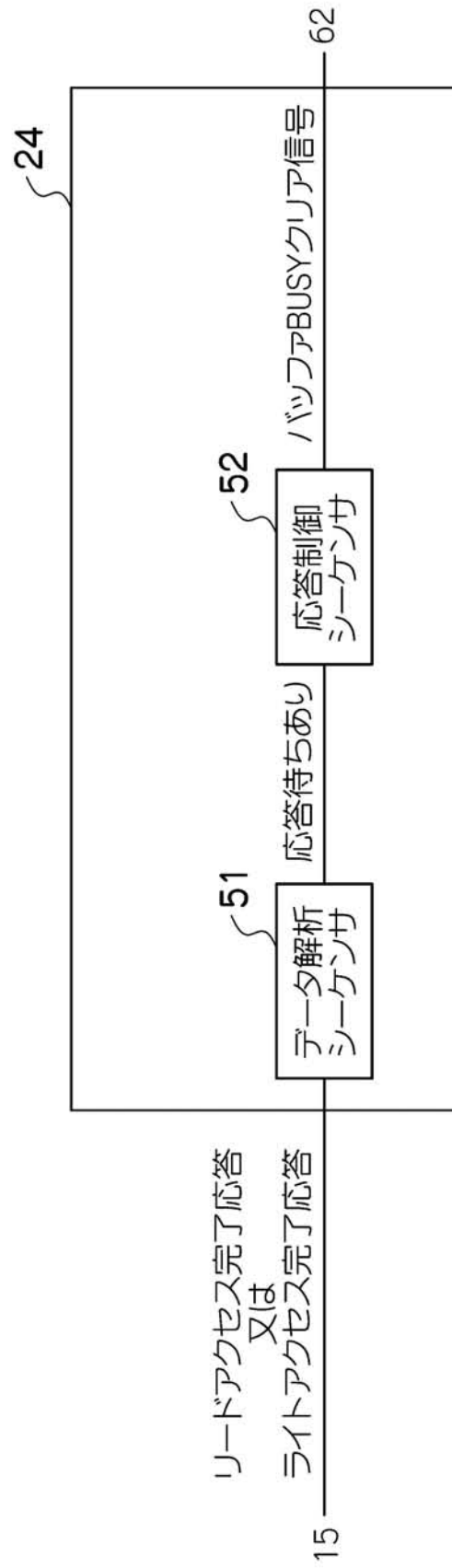
【図4】

図4

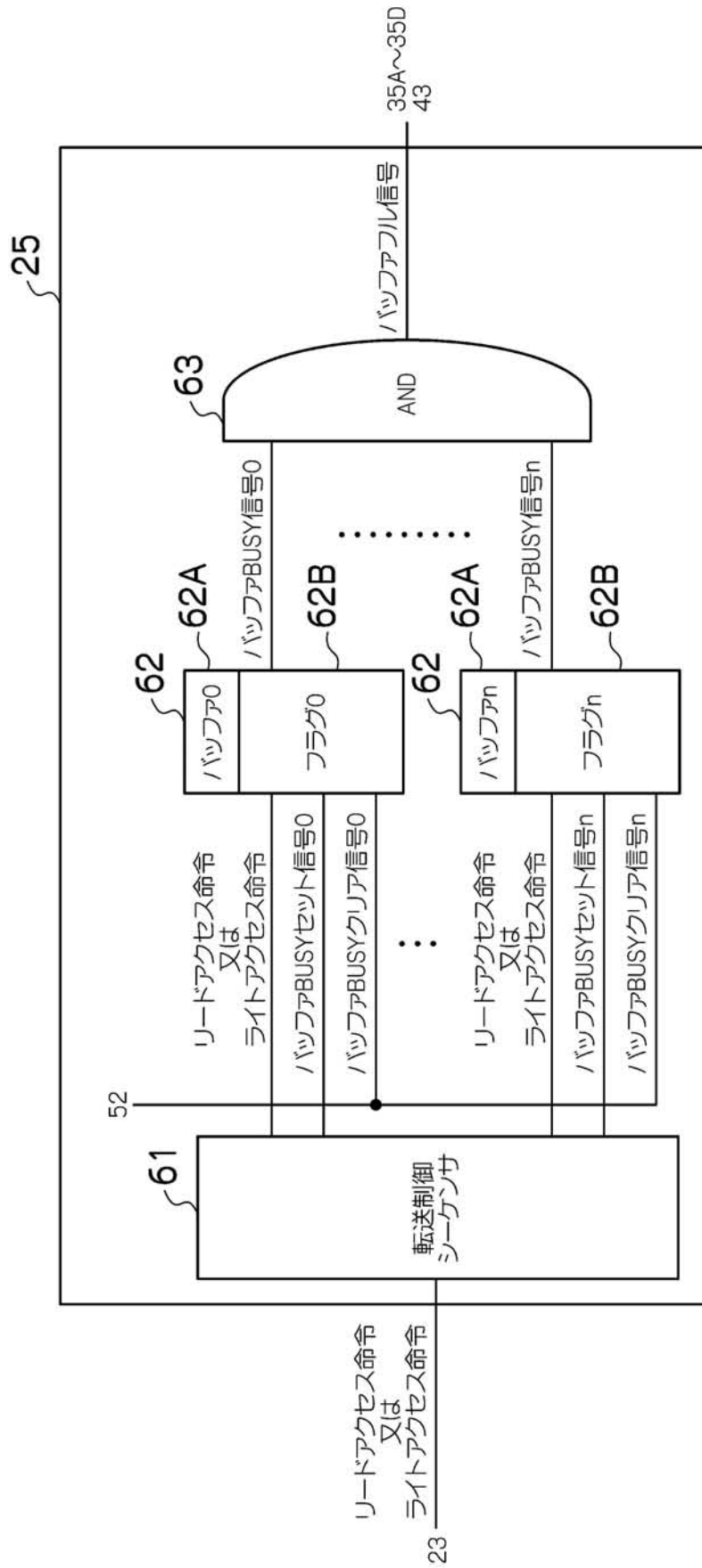


【図5】

図5

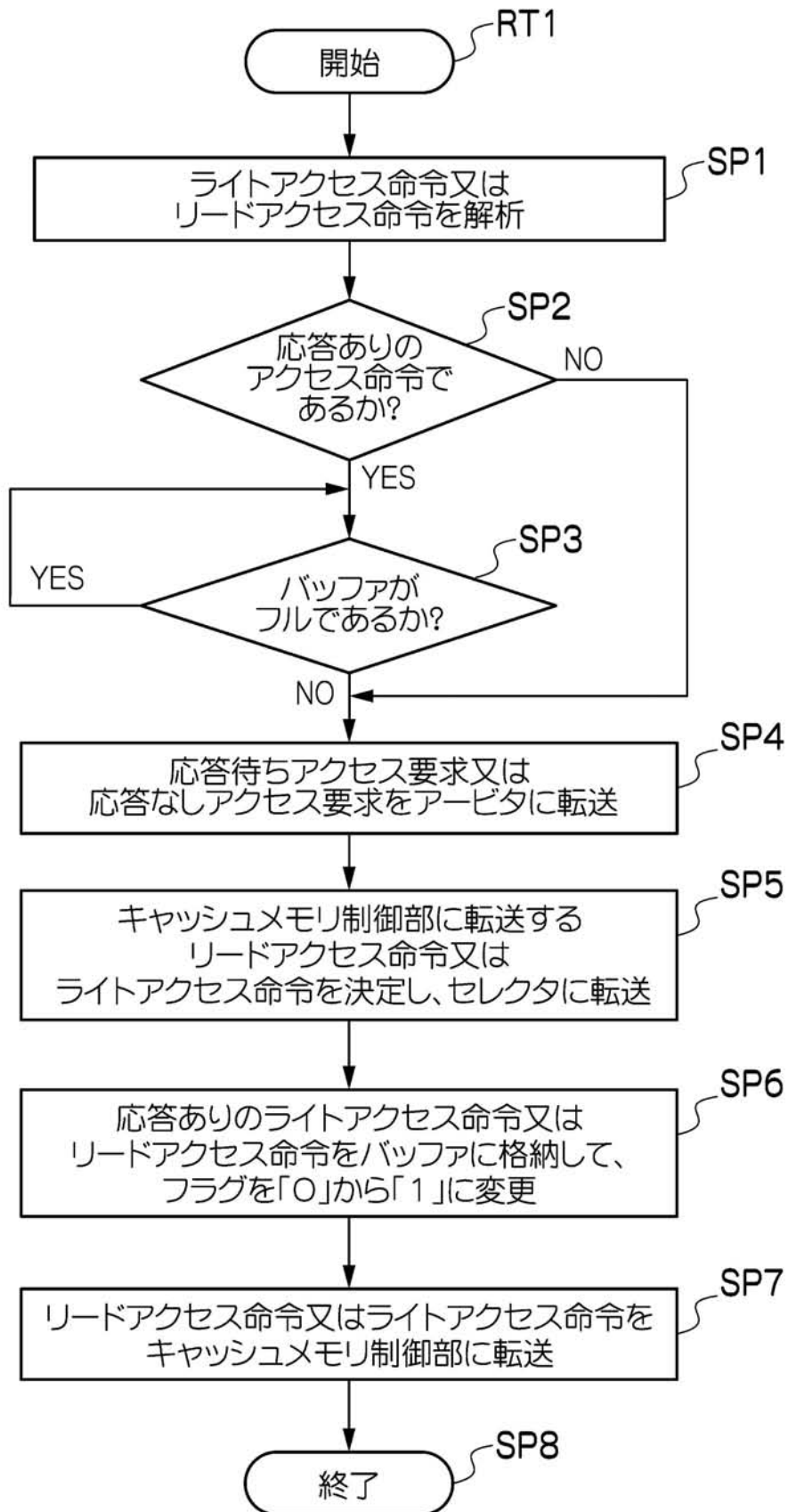


【図 6】
図6



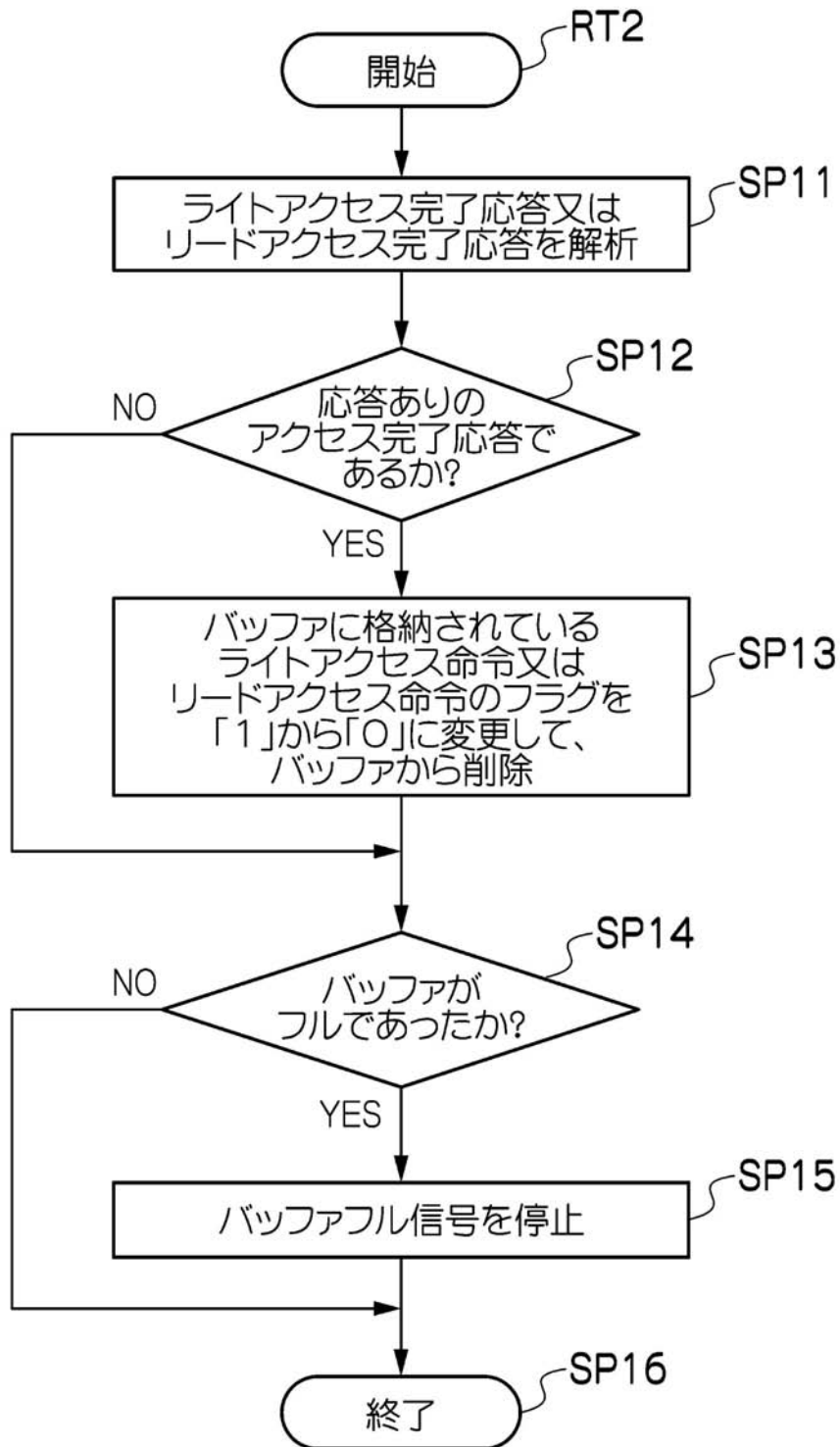
【図 7】

図7



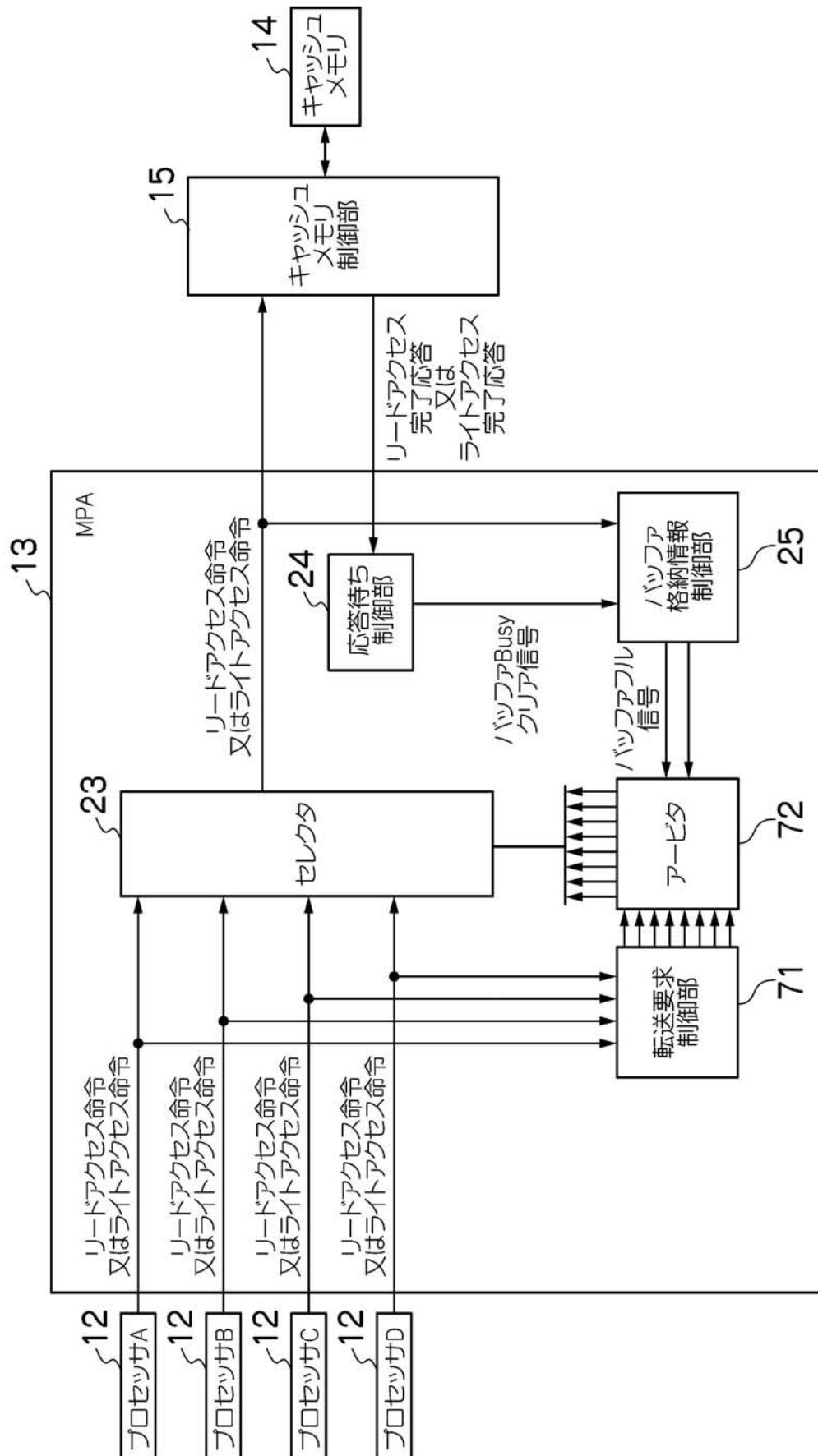
【図 8】

図8



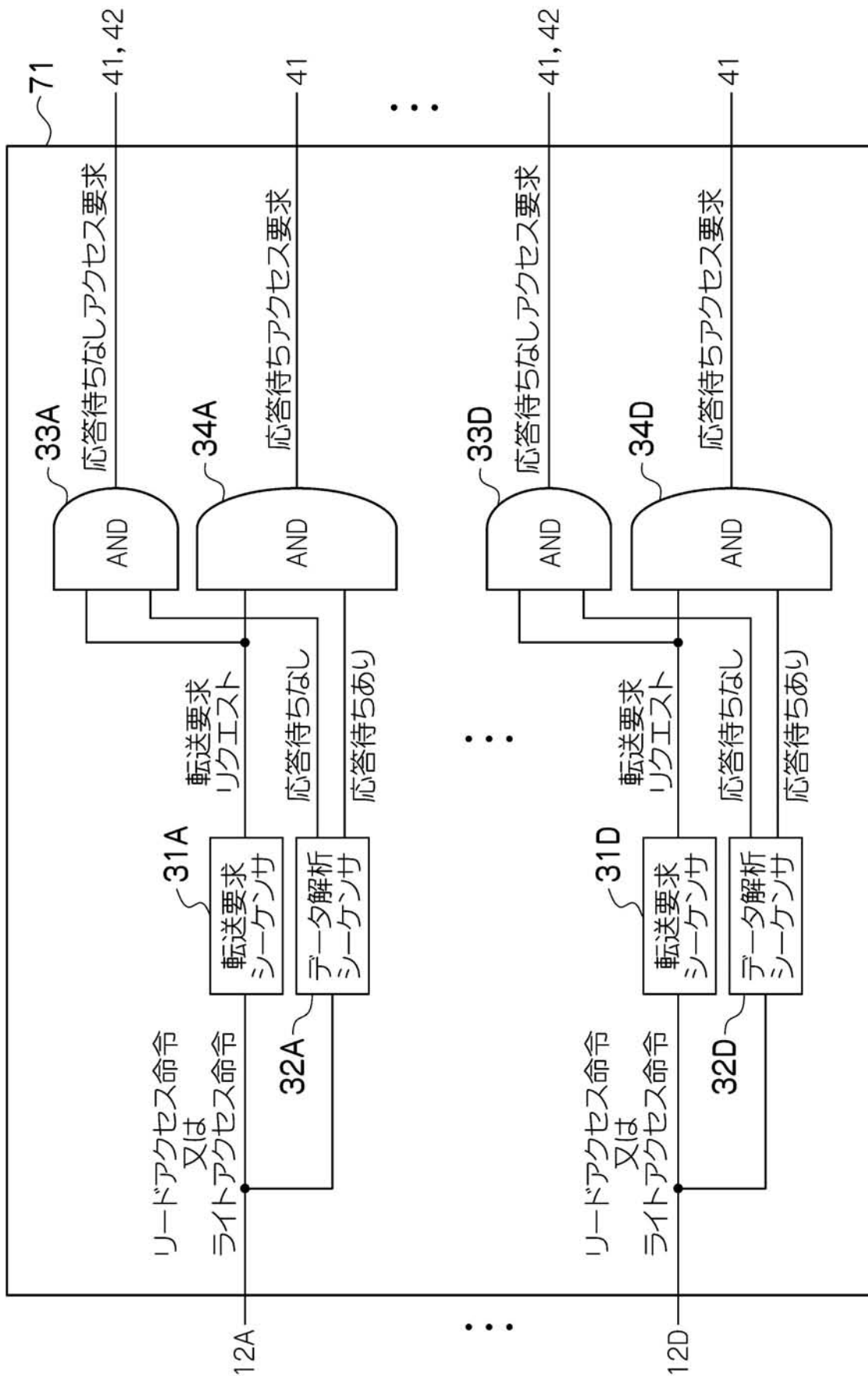
【図 9】

図 9



【図10】

図10



【図 11】

図 11

