

(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
2010年10月7日(07.10.2010)

PCT

(10) 国際公開番号
WO 2010/113262 A1

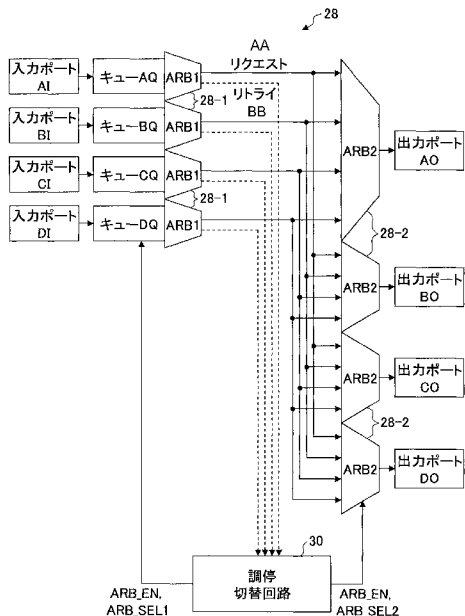
- (51) 国際特許分類:
G06F 15/173 (2006.01)
- (21) 国際出願番号: PCT/JP2009/056671
- (22) 国際出願日: 2009年3月31日(31.03.2009)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (71) 出願人(米国を除く全ての指定国について): 富士通株式会社(FUJITSU LIMITED) [JP/JP]; 〒2118588 神奈川県川崎市中原区上小田中4丁目1番1号 Kanagawa (JP).
- (72) 発明者; および
- (75) 発明者/出願人(米国についてのみ): 畑井田 誠(HATAIDA, Makoto) [JP/JP]; 〒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). 植木 俊和(Ueki, Toshikazu) [JP/JP]; 〒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). 木下 貴行(KINOSHITA, Takayuki) [JP/JP]; 〒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). 岩見 義和(IWAMI, Yoshikazu) [JP/JP]; 〒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). 小佐野 秀和(OSANO, Hidekazu) [JP/JP]; 〒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: 伊東 忠彦(ITO, Tadahiko); 〒1506032 東京都渋谷区恵比寿4丁目20番3号 恵比寿ガーデンプレイスタワー32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, BR, BW, BY, BZ, CA, CH, 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, KM, 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, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

[続葉有]

- (54) Title: ARBITRATION METHOD, ARBITRATION CIRCUIT, AND DEVICE USING THE ARBITRATION CIRCUIT
- (54) 発明の名称: 調停方法、調停回路、及び調停回路を備えた装置

[図3]



AI, BI, CI, DI Input port
 AQ, BQ, CQ, DQ Queue
 AO, BO, CO, DO Output Port
 30 Arbitration switching circuit
 AA Request
 BB Retry

(57) Abstract: When each arbitration method performs routing control to physically transfer the data from a plurality of input ports which logically have a plurality of virtual channels to any output port of a plurality of output ports, the arbitration method includes a first arbitration process which can select only one channel in one input port at any time and arbitrates between each input port channel in accordance with any arbitration algorithm other than a time-division algorithm, and a second arbitration process which arbitrates between the plurality of input ports in accordance with any arbitration algorithm. The arbitration algorithms used in the first and second arbitration processes switch to a constant period time-division algorithm in response to a trigger.

(57) 要約: 調停方法は、各々が論理的には複数の仮想的なチャンネルを有する物理的には複数の入力ポートからのデータを複数の出力ポートのうち任意の出力ポートへ転送するための経路制御を行う際に、任意の時点では1つの入力ポートで1つのチャンネルのみ選択可能であり、各入力ポートのチャンネル間の調停を時分割アルゴリズム以外の任意の調停アルゴリズムに従って行う第1の調停処理と、前記複数の入力ポート間の調停を前記任意の調停アルゴリズムに従って行う第2の調停処理を含む。前記第1及び第2の調停処理で用いる調停アルゴリズムは、トリガにตอบสนองして一定期間時分割アルゴリズムに切り替える。

WO 2010/113262 A1



(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, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), ヨーロッパ (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, SE, SI, SK, TR), OAPI (BF, BJ, CF,

CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

添付公開書類:

— 国際調査報告 (条約第 21 条(3))

明 細 書

調停方法、調停回路、及び調停回路を備えた装置

技術分野

[0001] 本発明は、調停方法、調停回路、及び調停回路を備えた装置に係り、特に並列計算機等に適した調停方法、調停回路、及び調停回路を備えた装置に関する。

背景技術

[0002] 図1は、並列計算機の一例を示すブロック図である。並列計算機1は、図1に示す如く接続された複数のシステムボード(SB: System board) 12、クロスバースイッチ14、及び複数の入出力ボード(IOB: Input Output Board) 15を有する。各SB12は、複数の中央処理装置(CPU: Central Processing Unit) 11及び複数のメモリ13を有する。各IOB15は、複数の入力部及び複数の出力部(又は、入出力インタフェース)を有する。クロスバースイッチ14は、一方のSB12と接続される入力ポートAI及び出力ポートAO、他方のSB12と接続される入力ポートBI及び出力ポートBO、一方のIOB15と接続される入力ポートCI及び出力ポートCI、及び他方のIOB15と接続される入力ポートDI及び出力ポートDOを有する。

[0003] クロスバースイッチ14は、上記の如き複数の入力ポート及び複数の出力ポートを有し、SB12やIOB15等の任意のノードからのデータとしてのパケットを他のノードへ転送するためのルーティング(経路制御)を行う。デッドロックを避けるために、クロスバースイッチ14は1つのポートについて複数の仮想的なチャンネルを持つ構成を有する。つまり、クロスバースイッチ14は物理的には複数のポートを有し、各ポートは論理的には複数のチャンネル(即ち、仮想的なチャンネル)を有するが、任意の時点では1つのポートで1つのチャンネルのみ選択可能である。

[0004] クロスバースイッチ14において、複数のポート及び複数のチャンネルからのパケットを調停する調停回路は、全てのポート及び全てのチャンネルを平等に扱うことが理想的である。図2は、このような理想的な調停回路を示す図である。図2において、調停回路17は、入力ポートAIからのチャンネルC0, C1の入力、入力ポートBIからのチャンネルC0, C1の入力、入力ポートCIからのチャンネルC0, C1の入力、及び入力ポートDIか

らのチャンネルC0, C1の入力に対して調停処理を行い、調停処理の結果に基づいて1つの入力ポートからのルーティング要求(経路制御要求)を任意の出力ポートへ出力する。

- [0005] しかし、実際の回路設計では、図2のような構成を有する調停回路17を作成することは物理的に難しく、信号遅延を考慮した設計とすることも難しい。このため、図3のような2段構成の調停回路が提案されている。
- [0006] 図3は、調停回路の一例を示す図である。図3において、キュー調停回路18-1は、各入力ポートAI, BI, CI, DI毎のキューAQ, BQ, CQ, DQからパケットを選択し、ポート間調停回路18-2は、複数の入力ポートAI, BI, CI, DIから1つのポートを選択する。キューAQ, BQ, CQ, DQは、クロスバースイッチ14内の対応するバッファ(図示せず)に保持されており、図3ではキューAQ, BQ, CQ, DQのブロックがこれらのバッファに相当する。このように、調停回路18は、2種類の調停回路18-1, 18-2を有する2段構成である。
- [0007] 図4は、キュー調停回路18-1の構成を示す図である。図4に示すように、各チャンネルC0, C1毎のキューの集まりに対して、平等な調停となるような調停アルゴリズムが用いられる。図4は、調停アルゴリズムとしてLRU(Least Recently Used)アルゴリズムが用いられる場合を示す。これにより、セクタ181は、LRUアルゴリズム部180がLRUアルゴリズムに基づいて求めた演算結果に応じてチャンネルC0, C1のうち一方のキューを選択出力する。
- [0008] ポート間調停回路18-2の図示は省略するが、チャンネルC0, C1毎のキューの集まりの代わりに、入力ポートAI, BI, CI, DI毎のキューの集まりに対してLRU等のキュー調停回路18-1と同じ調停アルゴリズムを用いて調停を行う。資源枯渇等のために、ポート間調停回路18-1においてキュー調停回路18-1からの経路制御要求(ルーティングリクエスト)が長時間採用されずにストール状態が続く場合には、デッドロックを防止するために一旦上記要求をキャンセルして別の要求を出し直すようなリトライ制御を行う。
- [0009] 上記の如き調停回路18は、全体としては平等に調停を行っているように見えるが、あるパケットに着目して見たときに、キュー調停回路18-1からポート間調停回路18

−2へ一向に出力されない状態のキューが発生する可能性があり、このような状態をライブロックと言う。しかし、調停アルゴリズムにLRUアルゴリズムではなく時分割アルゴリズムを用いたのでは、ライブロックは防止できても、調停時間が長くなり、並列計算機1の性能を低下させてしまう。

特許文献1:特開2001−22711号公報

発明の開示

発明が解決しようとする課題

[0010] 従来の調停方法では、ライブロックを防止しようとするすると調停時間が長くなってしまいう問題があり、その結果並列計算機の性能を低下させてしまいう問題もあった。

[0011] そこで、本発明は、調停時間の増加を抑制してライブロックを防止することのできる調停方法、調停回路、及び調停回路を備えた装置を提供することを目的とする。

課題を解決するための手段

[0012] 本発明の一観点によれば、各々が論理的には複数の仮想的なチャンネルを有する物理的には複数の入力ポートからのデータを複数の出力ポートのうち任意の出力ポートへ転送するための経路制御を行う際に、任意の時点では1つの入力ポートで1つのチャンネルのみ選択可能であり、各入力ポートのチャンネル間の調停を時分割アルゴリズム以外の任意の調停アルゴリズムに従って行う第1の調停処理と、前記複数の入力ポート間の調停を前記任意の調停アルゴリズムに従って行う第2の調停処理と、トリガに応答して一定期間前記第1及び第2の調停処理で用いる調停アルゴリズムを時分割アルゴリズムに切り替える処理を含む調停方法が提供される。

[0013] 本発明の一観点によれば、各々が論理的には複数の仮想的なチャンネルを有する物理的には複数の入力ポートからのデータを複数の出力ポートのうち任意の出力ポートへ転送するための経路制御を行う際に、任意の時点では1つの入力ポートで1つのチャンネルのみ選択可能であり、各入力ポートのチャンネル間の調停を時分割アルゴリズム以外の任意の調停アルゴリズムに従って行う第1の調停回路と、前記複数の入力ポート間の調停を前記任意の調停アルゴリズムに従って行う第2の調停回路と、トリガに応答して一定期間前記第1及び第2の調停処理で用いる調停アルゴリズムを時

分割アルゴリズムに切り替える調停切替回路を備えた調停回路が提供される。

- [0014] 本発明の一観点によれば、複数のプロセッサ及びメモリを有するシステムと、複数の入出力インタフェースを有する入出力システムを接続する装置であつて、各々が論理的には複数の仮想的なチャンネルを有する物理的には複数の入力ポートと、複数の出力ポートと、前記複数の入力ポートからのデータを前記複数の出力ポートのうち任意の出力ポートへ転送するための経路制御を行う際に、任意の時点では1つの入力ポートで1つのチャンネルのみ選択可能であり、各入力ポートのチャンネル間の調停を時分割アルゴリズム以外の任意の調停アルゴリズムに従って行う第1の調停回路と、前記複数の入力ポート間の調停を前記任意の調停アルゴリズムに従って行う第2の調停回路と、トリガに応答して一定期間前記第1及び第2の調停処理で用いる調停アルゴリズムを時分割アルゴリズムに切り替える調停切替回路を備えた装置が提供される。

発明の効果

- [0015] 開示の調停方法、調停回路、及び調停回路を備えた装置によれば、調停時間の増加を抑制してライブロックを防止することが可能となる。

図面の簡単な説明

- [0016] [図1]並列計算機の一例を示すブロック図である。
[図2]理想的な調停回路を示す図である。
[図3]調停回路の一例を示す図である。
[図4]キュー調停回路の構成を示す図である。
[図5]第1実施例における調停回路の一例を示す図である。
[図6]調停切替回路の構成を示す図である。
[図7]キュー調停回路の構成を示す図である。
[図8]入力ポートAIのキューAQの状態の一例を示す図である。
[図9]LRUアルゴリズムに基づくクロック毎の調停結果を示す図である。
[図10]時分割アルゴリズムに基づくクロック毎の調停結果を示す図である。
[図11]第2実施例における調停回路の一例を示す図である。
[図12]キュー調停回路の構成を示す図である。

[図13]調停切替回路の構成を示す図である。

[図14]第3実施例における調停回路の一例を示す図である。

[図15]調停切替回路の構成を示す図である。

符号の説明

- [0017] 1 並列計算機
- 11 CPU
- 12 SB
- 13 メモリ
- 14 クロスバースイッチ
- 15 ICB
- 28, 28A, 28B 調停回路
- 28-1, 28-1A キュー調停回路
- 28-2 ポート調停回路
- 30, 30A, 30B 調停切替回路

発明を実施するための最良の形態

[0018] 開示の調停方法、調停回路、及び調停回路を備えた装置によれば、各々が論理的には複数の仮想的なチャンネルを有する物理的には複数の入力ポートからのデータを複数の出力ポートのうち任意の出力ポートへ転送するためのルーティング(経路制御)を行う際に、任意の時点では1つの入力ポートで1つのチャンネルのみ選択可能であり、第1の調停処理により各入力ポートのチャンネル間の調停を時分割アルゴリズム以外の任意の調停アルゴリズムに従って行い、第2の調停処理により複数の入力ポート間の調停を前記任意の調停アルゴリズムに従って行う。又、トリガに応答して、一定期間前記第1及び第2の調停処理で用いる調停アルゴリズムを時分割アルゴリズムに切り替える。

[0019] これにより、調停時間の増加を抑制してライブロックを防止することが可能となる。

[0020] 以下に、開示の調停方法、調停回路、及び調停回路を備えた装置の各実施例を、図5以降と共に説明する。

実施例

- [0021] 図5は、第1実施例における調停回路の一例を示す図である。図5に示す調停回路28は、例えば図1に示す並列計算機1のクロスバースイッチ14内に設けられているものとする。図5中、図1と同一部分には同一符号を付し、その説明は省略する。
- [0022] 図5に示す調停回路28は、キュー調停回路28-1、ポート調停回路28-2、及び調停切替回路30を有する。キュー調停回路28-1は、各入力ポートAI, BI, CI, DI毎のキューAQ, BQ, CQ, DQからデータとしてのパケットを選択し、ポート間調停回路28-2は、複数の入力ポートAI, BI, CI, DIから1つのポートを選択する。このように、調停回路28は、2種類の調停回路28-1, 28-2を有する2段構成である。
- [0023] キュー調停回路28-1は、調停切替回路30により指示された第1又は第2の調停アルゴリズムに基づいたキューの調停を行う。同様に、ポート調停回路28-2は、調停切替回路30により指示された第1又は第2の調停アルゴリズムに基づいたポートの調停を行う。
- [0024] キュー調停回路28-1からポート調停回路28-2への調停リクエストがある一定期間採用(又は、選択)されない場合、キュー調停回路28-1はこの調停リクエストを一旦キャンセルして別の調停リクエストを再度要求する。この調停リクエストのキャンセルを通知するのが図5に示すリトライである。
- [0025] 調停切替回路30は、各入力ポートAI, BI, CI, DIにおけるリトライの頻度を監視し、リトライの頻度が閾値に達すると調停回路28-1, 28-2の調停アルゴリズムを第1の調停アルゴリズムから第2の調停アルゴリズムに切り替える指示を出力する。調停切替回路30は、キュー調停回路28-1に調停切替信号ARB__EN及び調停アルゴリズムを指示する選択信号ARB__SEL1を供給し、ポート調停回路28-2に調停切替信号ARB__EN及び調停アルゴリズムを指示する選択信号ARB__SEL2を供給する。各入力ポートAI, BI, CI, DIのリトライの頻度が閾値に達すると、調停切替回路30は調停アルゴリズムの切り替えを許可する調停切替信号ARB__ENをアサートすることで調停回路28-1, 28-2の調停アルゴリズムを選択信号ARB__SEL1, ARB__SEL2が指示する調停アルゴリズムに切り替えるよう指示する。
- [0026] 第1及び第2の調停アルゴリズムは、互いに異なる調停アルゴリズムであり、本実施例では第1の調停アルゴリズムがLRUアルゴリズムであり、第2の調停アルゴリズムが

時分割アルゴリズムである。尚、第1の調停アルゴリズムはLRUアルゴリズムに限定されるものではなく、例えばラウンドロビンで選択を行うラウンドロビンアルゴリズム、固定の優先順位で選択を行う固定プライオリティアルゴリズム等を用いても良いことは言うまでもない。

[0027] 図6は、調停切替回路30の構成を示す図である。調停切替回路30は、図6に示す如く接続されたフリーランカウンタ31、頻度監視カウンタ32、オア回路33、34、一定時間計測カウンタ35、及びデコーダ36を有する。この例では、カウンタ31、32、35は、夫々10ビットカウンタ、4ビットカウンタ、10ビットカウンタである。

[0028] フリーランカウンタ31のカウント値が最大値(例えば1024サイクル)に達すると、フリーランカウンタ31からのクリア信号が頻度監視カウンタ32のクリア端子clrに入力されて頻度監視カウンタ32がクリアされる。各入力ポートAI, BI, CI, DIに対するリトライ指示は、オア回路33を介して頻度監視カウンタ32のカウントアップ端子upに入力され、頻度監視カウンタ32は各入力ポートAI, BI, CI, DIに対するリトライ指示の回数をカウントさせる。カウントしたリトライ指示の頻度がある閾値(例えば15回/1024サイクル)に達すると、頻度監視カウンタ32はオア回路34を介して一定時間計測カウンタ35のカウントアップ端子upに信号を入力して一定時間計測カウンタ35のカウントをスタートさせる。一定時間計測カウンタ35が一定時間をカウントする間(即ち、カウント値が一定値を超えるまで)、出力する調停切替信号ARB_ENをアサートする。調停切替信号ARB_ENがアサートされている期間は調停アルゴリズムがLRUアルゴリズムから時分割アルゴリズムに切り替えられる。一定時間計測カウンタ35が一定時間をカウントすると(カウント値が一定値を超えると)、カウント値が0となり調停切替信号ARB_ENがデアサートされるので、調停アルゴリズムは元のLRUアルゴリズムに戻される。調停切替信号ARB_ENは、オア回路34にフィードバックされる。一方、調停切替信号ARB_ENは、デコーダ36に入力され、デコーダ36からは選択信号ARB_SEL1, ARB_SEL2が出力される。

[0029] 図7は、キュー調停回路28-1の構成を示す図である。ポート調停回路28-2は、キュー調停回路28-1と同じ構成の回路で実現できるので、ポート調停回路28-2の構成の図示及び説明は省略する。ただし、図7の調停回路をポート調停回路28-

2として使用する場合には、選択信号ARB_SEL1の代わりに選択信号ARB_SEL2が入力されることは言うまでもない。

- [0030] キュー調停回路28-1は、図7に示す如く接続されたLRUアルゴリズム部281、セクタ282、283、アンド回路284、285、オア回路286、及びカウンタ287を有する。セクタ282には、LRUアルゴリズム部281に従った信号及び選択信号ARB_SEL1が入力され、調停切替信号ARB_ENにより出力が選択される。セクタ283には、例えば入力ポートAIにおけるキューAQが入力され、セクタ282の出力信号により出力が選択される。各アンド回路284、285には、セクタ283の出力信号及び出力ポートビジー信号が入力される。アンド回路284からはルーティングリクエスト(経路制御要求)が出力される。
- [0031] アンド回路285からはストール信号が出力されてオア回路286及びカウンタ287のカウントアップ端子upに入力される。カウンタ287の出力信号はLRUアルゴリズム部281及びオア回路286に入力されると共に、この場合は入力ポートAIに対するリトライ指示として調停切替回路30に入力される。オア回路286の出力信号はカウンタ287のクリア端子clrに入力されるので、カウンタ287のカウント値が一定値に達するとカウンタ287はクリアされる。カウンタ287は、例えば4ビットカウンタである。
- [0032] これにより、キュー調停回路28-1は、通常時はLRUアルゴリズム部281によるLRUアルゴリズムに従って調停を行うが、アサートされた調停切替信号ARB_EN信号によって時分割アルゴリズムに従った調停に切り替えられる。この調停切替信号ARB_ENが一定時間後にデアサートされると、調停アルゴリズムは時分割アルゴリズムからLRUアルゴリズムに戻される。
- [0033] 次に、本実施例においてどのようにしてライブロックが防止されるかを、図8～図10と共に説明する。クロスバースイッチ14は物理的には複数のポートを有し、各ポートは論理的には複数のチャンネル(即ち、仮想的なチャンネル)を有するが、任意の時点では1つのポートで1つのチャンネルのみ選択可能である。
- [0034] 図8は、入力ポートAIのキューAQの状態の一例を示す図である。図9は、LRUアルゴリズムに基づくクロック毎の調停結果を示す図であり、図10は、時分割アルゴリズムに基づくクロック毎の調停結果を示す図である。図9及び図10は一例として、出力

ポートDOの出力ポートビジー信号、リトライ指示の有無、入力ポートAI, BI, CI、及び調停結果を示す。出力ポートビジー信号は、「1」の場合にその出力ポートがビジーであることを示す。又、リトライ指示の有無は、「1」の場合にリトライ指示が有ることを示す。又、A1x, A2xは入力ポートAIにおけるチャンネルC0, C1側のパケット、B1x, B2xは入力ポートBIにおける仮想的なチャンネルC0, C1側のパケット、C1x, C2xは入力ポートCIにおけるチャンネルC0, C1側のパケットを示す。

[0035] 説明の便宜上、例えば図8のように、入力ポートAIのキューAQに全て出力ポートDO向けのチャンネルC0側のパケットA10, A11, A12及びチャンネルC1側のパケットA20, A21, A22がクロスバースイッチ14内のバッファに滞留しているとする。図8は入力ポートAIのみについて示すが、入力ポートBI, CIについて以下の如き調停が行われることは言うまでもない。

[0036] これら各入力ポートAI, BI, CIから出力ポートDOへの調停は、通常時はLRUアルゴリズムに従って図9のように調停される。図9は、サイクル単位の調停を示し、出力ポートビジー信号が「1」であり出力ポートDOがビジーであることを示すサイクルでは、出力ポートDOが資源ビジーのために調停を行うことはできない。又、リトライ指示の有無が「1」でありリトライ指示が有ることを示すサイクルでは、各入力ポートAI, BI, CIのルーティングリクエストを一旦キャンセルしてから再度調停を行う。従って、キャンセルされたパケットはLRUアルゴリズムによれば最低順位を有することになる。図9の例では、出力ポートDOが恒常的に資源ビジーとなっており、一定間隔で1個ずつしかパケットを調停できない状態に陥っている。このような場合、入力ポートAIで見ると、図8のキューAQのチャンネルC0側のパケットA1xのみが処理されており、キューAQのチャンネルC1側のパケットA2xはクロスバースイッチ14内のバッファに滞留したままの状態に陥っている。更に、後続のパケットがチャンネルC0側のパケットA1xのキューに追加投入されると、チャンネルC1側のパケットA2xのキューがライブロックしてしまう。

[0037] 一方、図9と同じような状況下で、図10に示すような、時分割の調停アルゴリズムを用いた場合は、ある一定時間に同一入力ポート及び同一チャンネルのキューを固定的に調停し、更にこのような調停を全入力ポートの全チャンネルに平等に割り当てる。このため、一部のチャンネルのライブロックは発生しないことがわかる。

- [0038] 尚、単に時分割アルゴリズムのみによる調停を行うだけでは並列計算機1の性能が低下してしまうが、ライブロックに陥っていると懸念される場合にのみ調停アルゴリズムをLRUアルゴリズムから時分割アルゴリズムに切り替えることで、並列計算機1の性能低下を抑制しつつライブロックを防止することが可能となる。
- [0039] 上記第1実施例では、調停アルゴリズムの切り替えをリトライ指示の頻度が閾値に達したことをトリガとして行っているが、調停アルゴリズムの切り替えはリトライ指示の頻度以外をトリガとしても良いことは言うまでもない。
- [0040] 図11は、第2実施例における調停回路の一例を示す図である。図11に示す調停回路28Aは、例えば図1に示す並列計算機1のクロスバースイッチ14内に設けられているものとする。図11中、図5と同一部分には同一符号を付し、その説明は省略する。
- [0041] 本実施例では、各入力ポートAI, BI, CI, DI毎のロック情報が対応するキュー調停回路28-1Aから調停切替回路30Aに供給される点が上記第1実施例と異なる。
- [0042] 図12は、キュー調停回路28-1Aの構成を示す図である。図12中、図7と同一部分には同一符号を付し、その説明は省略する。図12において、各アンド回路301, 302には、セレクタ282の出力信号、出力ポートビジイー信号、及び入力ポートAIにおけるチャンネルC0側のパケットA1x(ルーティングリクエスト)が入力される。各アンド回路311, 312には、セレクタ282の出力信号、出力ポートビジイー信号、及び入力ポートAIにおけるチャンネルC1側のパケットA2x(ルーティングリクエスト)が入力される。
- [0043] 図12において、アンド回路301の出力信号は、ハイレベルであると入力ポートAIにおけるチャンネルC0側のパケットA1xが調停されなかったことを示し、カウンタ303のカウントアップ端子upに入力される。アンド回路302の出力信号は、ハイレベルであると入力ポートAIにおけるチャンネルC0側のパケットA1xが調停されたことを示し、オア回路304を介してカウンタ303のクリア端子clrに入力される。カウンタ303は、例えば4ビットカウンタである。カウンタ303の出力信号は、入力ポートAIにおけるチャンネルC0側のパケットA1xをロックするポートA1ロック信号として調停切替回路30Aへ出力されると共に、オア回路304にフィードバックされる。カウンタ303のカウント値が一定値に達すると、カウンタ303の出力信号(ポートA1ロック信号)は、入力ポートAIにおけ

るチャンネルC0側のパケットA1xがデッドロックした状態であることを示す。

[0044] アンド回路311の出力信号は、ハイレベルであると入力ポートAIにおけるチャンネルC1側のパケットA2xが調停されなかったことを示し、カウンタ313のカウントタップ端子upに入力される。アンド回路312の出力信号は、ハイレベルであると入力ポートAIにおけるチャンネルC1側のパケットA2xが調停されたことを示し、オア回路314を介してカウンタ313のクリア端子clrに入力される。カウンタ313は、例えば4ビットカウンタである。カウンタ313の出力信号は、入力ポートAIにおけるチャンネルC1側のパケットA2xをロックするポートA2ロック信号として調停切替回路30Aへ出力されると共に、オア回路314にフィードバックされる。カウンタ313のカウント値が一定値に達すると、カウンタ313の出力信号(ポートA2ロック信号)は、入力ポートAIにおけるチャンネルC1側のパケットA2xがデッドロックした状態であることを示す。

[0045] 図13は、調停切替回路30Aの構成を示す図である。図13中、図6と同一部分には同一符号を付し、その説明は省略する。

[0046] 図13において、調停切替回路30Aは、オア回路330、34、一定時間計測カウンタ35、及びデコーダ36を有する。オア回路330には、各キュー調停回路28-1Aで検出されたロック情報、即ち、入力ポートAIに対するポートA1、A2ロック信号、入力ポートBIに対するポートB1、B2ロック信号、入力ポートCIに対するポートC1、C2ロック信号、及び入力ポートDIに対するポートD1、D2ロック信号が入力される。オア回路330の出力信号は、オア回路34の一方の入力に供給される。

[0047] 各入力ポートAI、BI、CI、DIからのロック情報が入力されると、一定時間計測カウンタ35が一定時間をカウントする間(即ち、カウント値が一定値を超えるまで)、出力する調停切替信号ARB_ENをアサートする。調停切替信号ARB_ENがアサートされている期間は調停アルゴリズムがLRUアルゴリズムから時分割アルゴリズムに切り替えられる。一定時間計測カウンタ35が一定時間をカウントすると(カウント値が一定値を超えると)、カウント値が0となり調停切替信号ARB_ENがデアサートされるので、調停アルゴリズムは元のLRUアルゴリズムに戻される。

[0048] これにより、本実施例では、調停アルゴリズムの切り替えを少なくとも1つの入力ポートにおけるデッドロックをトリガとして行うことができる。

- [0049] 図14は、第3実施例における調停回路の一例を示す図である。図14に示す調停回路28Bは、例えば図1に示す並列計算機1のクロスバースイッチ14内に設けられているものとする。図14中、図5と同一部分には同一符号を付し、その説明は省略する。
- [0050] 本実施例では、各入力ポートAI, BI, CI, DIのキューAQ, BQ, CQ, DQが空になったことを示すエンプティ信号EMPTY__A, EMPTY__B, EMPTY__C, EMPTY__Dがバッファから調停切替回路30Bに供給されると共に、調停切替回路30Bが調停アルゴリズムをLRUアルゴリズムから時分割アルゴリズムに切り替えている期間はパケットの投入を抑えるインヒビット信号INHを各入力ポートAI, BI, CI, DIに供給する点が上記第1実施例と異なる。
- [0051] 図15は、調停切替回路30Bの構成を示す図である。図15中、図6と同一部分には同一符号を付し、その説明は省略する。
- [0052] 図15において、アンド回路401には、各入力ポートAI, BI, CI, DIのキューAQ, BQ, CQ, DQが空になったことを示すエンプティ信号EMPTY__A, EMPTY__B, EMPTY__C, EMPTY__Dがバッファから入力される。アンド回路401の出力信号は、一定時間計測カウンタ35のクリア端子clrに入力される。これにより、各キューAQ, BQ, CQ, DQからのエンプティ信号EMPTY__A, EMPTY__B, EMPTY__C, EMPTY__Dが全てアサートされると、一定時間計測カウンタ35のカウンタ値は一定時間の計測を終了する前であっても0にクリアされる。又、一定時間計測カウンタ35が出力する調停切替信号ARB__ENは、バッファ402を介してインヒビット信号INHとして各入力ポートAI, BI, CI, DIに入力される。
- [0053] これにより、調停切替信号ARB__ENがアサートされている期間は、アサートされたインヒビット信号INHが各入力ポートAI, BI, CI, DIに供給されるので、各入力ポートAI, BI, CI, DIへのパケットの投入が抑えられる。一方、調停切替信号ARB__ENがアサートされている期間は調停アルゴリズムがLRUアルゴリズムから時分割アルゴリズムに切り替えられているが、各入力ポートAI, BI, CI, DIのキューAQ, BQ, CQ, DQが全て空になりエンプティ信号EMPTY__A, EMPTY__B, EMPTY__C, EMPTY__Dが全てアサートされると、一定時間計測カウンタ35のカウンタ値がその

時点で0にクリアされる。一定時間計測カウンタ35のカウンタ値が0にクリアされると、調停切替信号ARB_ENがデアサートされるので、調停アルゴリズムは元のLRUアルゴリズムに戻される。

[0054] これにより、本実施例では、調停アルゴリズムの切り替えをリトライ指示の頻度が閾値に達したことをトリガとして行い、且つ、各キューQA, QB, QC, QD内のパケットが全て空になったことをトリガとして元の調停アルゴリズムに戻すことができる。

産業上の利用可能性

[0055] 開示の調停方法、調停回路、及び調停回路を備えた装置は、クロスバースイッチや、クロスバースイッチを有する並列計算機等調停処理を要する各種装置に適用可能である。

[0056] 以上、開示の調停方法、調停回路、及び調停回路を備えた装置を実施例により説明したが、本発明は上記実施例に限定されるものではなく、本発明の範囲内で種々の変形及び改良が可能であることは言うまでもない。

請求の範囲

- [1] 各々が論理的には複数の仮想的なチャンネルを有する物理的には複数の入力ポートからのデータを複数の出力ポートのうち任意の出力ポートへ転送するための経路制御を行う際に、任意の時点では1つの入力ポートで1つのチャンネルのみ選択可能であり、各入力ポートのチャンネル間の調停を時分割アルゴリズム以外の任意の調停アルゴリズムに従って行う第1の調停処理と、
- 前記複数の入力ポート間の調停を前記任意の調停アルゴリズムに従って行う第2の調停処理と、
- トリガに応答して一定期間前記第1及び第2の調停処理で用いる調停アルゴリズムを時分割アルゴリズムに切り替える処理を含む、調停方法。
- [2] 前記第2の調停処理において任意の入力ポートからの経路制御要求が一定期間採用されない場合に前記任意の入力ポートに対して前記任意の入力ポートからの経路制御要求を一旦キャンセルするリトライ指示を出力して前記第1の調停処理で再度調停を行い、前記リトライ指示の頻度が閾値に達したことを前記トリガとする、請求項1記載の調停方法。
- [3] 前記複数の入力ポートにおける各キュー内のデータが全て空になったことをトリガとして前記調停アルゴリズムを前記任意の調停アルゴリズムに戻す、請求項2記載の調停方法。
- [4] 前記第1の調停処理において少なくとも1つの入力ポートにおけるデッドロックが検出されたことを前記トリガとする、請求項1記載の調停方法。
- [5] 前記任意の調停アルゴリズムは、LRU (Least Recently Used) アルゴリズム、ラウンドロビンアルゴリズム、及び固定プライオリティアルゴリズムからなるグループから選択された1つのアルゴリズムである、請求項1乃至4のいずれか1項記載の調停方法。
- [6] 各々が論理的には複数の仮想的なチャンネルを有する物理的には複数の入力ポートからのデータを複数の出力ポートのうち任意の出力ポートへ転送するための経路制御要求を行う際に、任意の時点では1つの入力ポートで1つのチャンネルのみ選択可能であり、各入力ポートのチャンネル間の調停を時分割アルゴリズム以外の任意の調停アルゴリズムに従って行う第1の調停回路と、

前記複数の入力ポート間の調停を前記任意の調停アルゴリズムに従って行う第2の調停回路と、

トリガに応答して一定期間前記第1及び第2の調停処理で用いる調停アルゴリズムを時分割アルゴリズムに切り替える調停切替回路を備えた、調停回路。

- [7] 前記第2の調停回路において任意の入力ポートからの経路制御要求が一定期間採用されない場合に前記任意の入力ポートに対して前記任意の入力ポートからの経路制御要求を一旦キャンセルするリトライ指示が出力されて前記第1の調停回路は再度調停を行い、

前記調停切替回路は、前記リトライ指示の頻度が閾値に達したことを前記トリガとする、請求項6記載の調停回路。

- [8] 前記調停切替回路は、前記複数の入力ポートにおける各キュー内のデータが全て空になったことをトリガとして前記調停アルゴリズムを前記任意の調停アルゴリズムに戻す、請求項7記載の調停回路。

- [9] 前記調停切替回路は、前記第1の調停回路において少なくとも1つの入力ポートにおけるデッドロックが検出されたことを前記トリガとする、請求項6記載の調停回路。

- [10] 前記任意の調停アルゴリズムは、LRU (Least Recently Used) アルゴリズム、ラウンドロビンアルゴリズム、及び固定プライオリティアルゴリズムからなるグループから選択された1つのアルゴリズムである、請求項6乃至9のいずれか1項記載の調停回路。

- [11] 複数のプロセッサ及びメモリを有するシステムと、複数の入出力インタフェースを有する入出力システムを接続する装置であって、

各々が論理的には複数の仮想的なチャンネルを有する物理的には複数の入力ポートと、

複数の出力ポートと、

前記複数の入力ポートからのデータを前記複数の出力ポートのうち任意の出力ポートへ転送するための経路制御を行う際に、任意の時点では1つの入力ポートで1つのチャンネルのみ選択可能であり、各入力ポートのチャンネル間の調停を時分割アルゴリズム以外の任意の調停アルゴリズムに従って行う第1の調停回路と、

前記複数の入力ポート間の調停を前記任意の調停アルゴリズムに従って行う第2の

調停回路と、

トリガに応答して一定期間前記第1及び第2の調停処理で用いる調停アルゴリズムを時分割アルゴリズムに切り替える調停切替回路を備えた、装置。

- [12] 前記第2の調停回路において任意の入力ポートからの経路制御要求が一定期間採用されない場合に前記任意の入力ポートに対して前記任意の入力ポートからの経路制御要求を一旦キャンセルするリトライ指示が出力されて前記第1の調停回路は再度調停を行い、

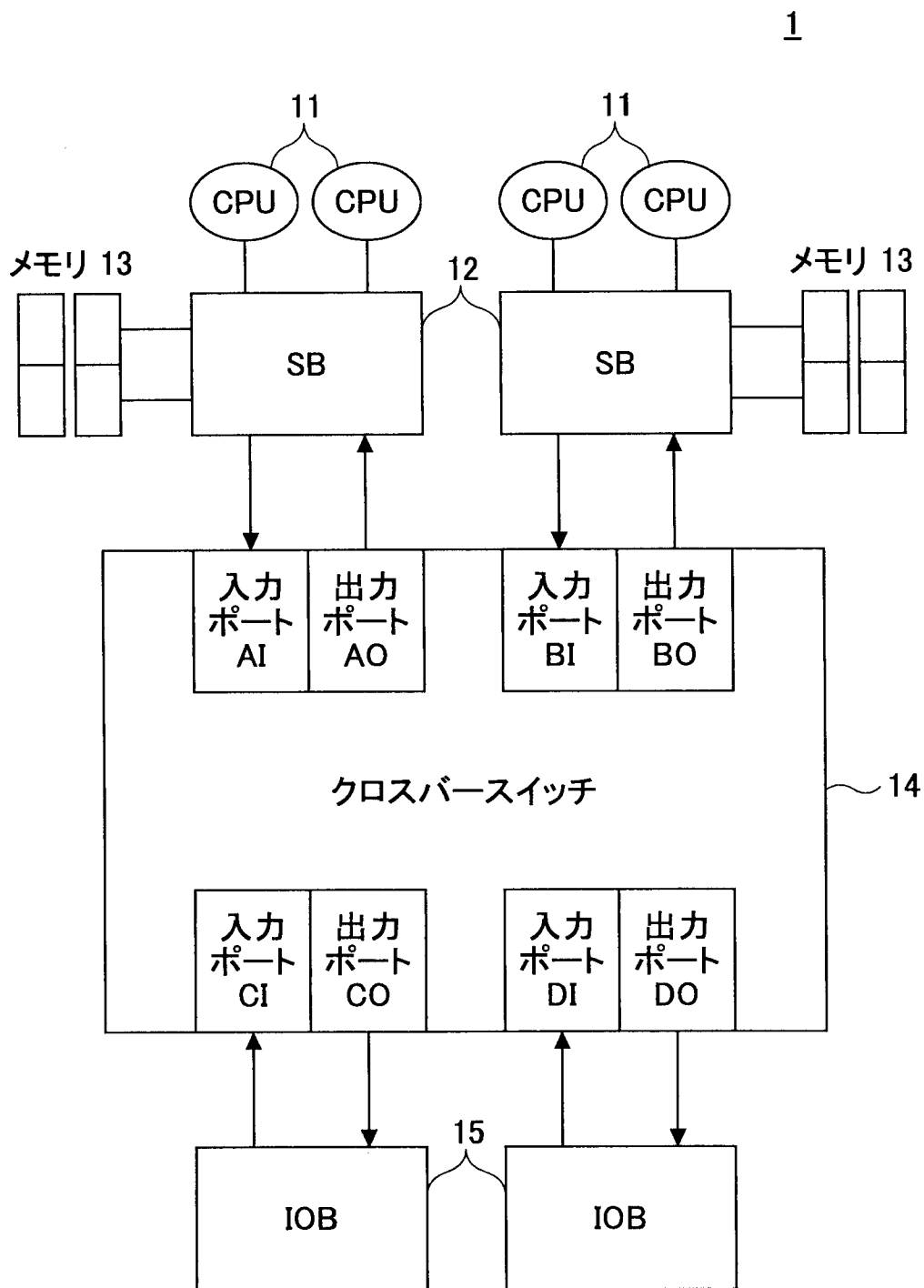
前記調停切替回路は、前記リトライ指示の頻度が閾値に達したことを前記トリガとする、請求項11記載の装置。

- [13] 前記調停切替回路は、前記複数の入力ポートにおける各キュー内のデータが全て空になったことをトリガとして前記調停アルゴリズムを前記任意の調停アルゴリズムに戻す、請求項12記載の装置。

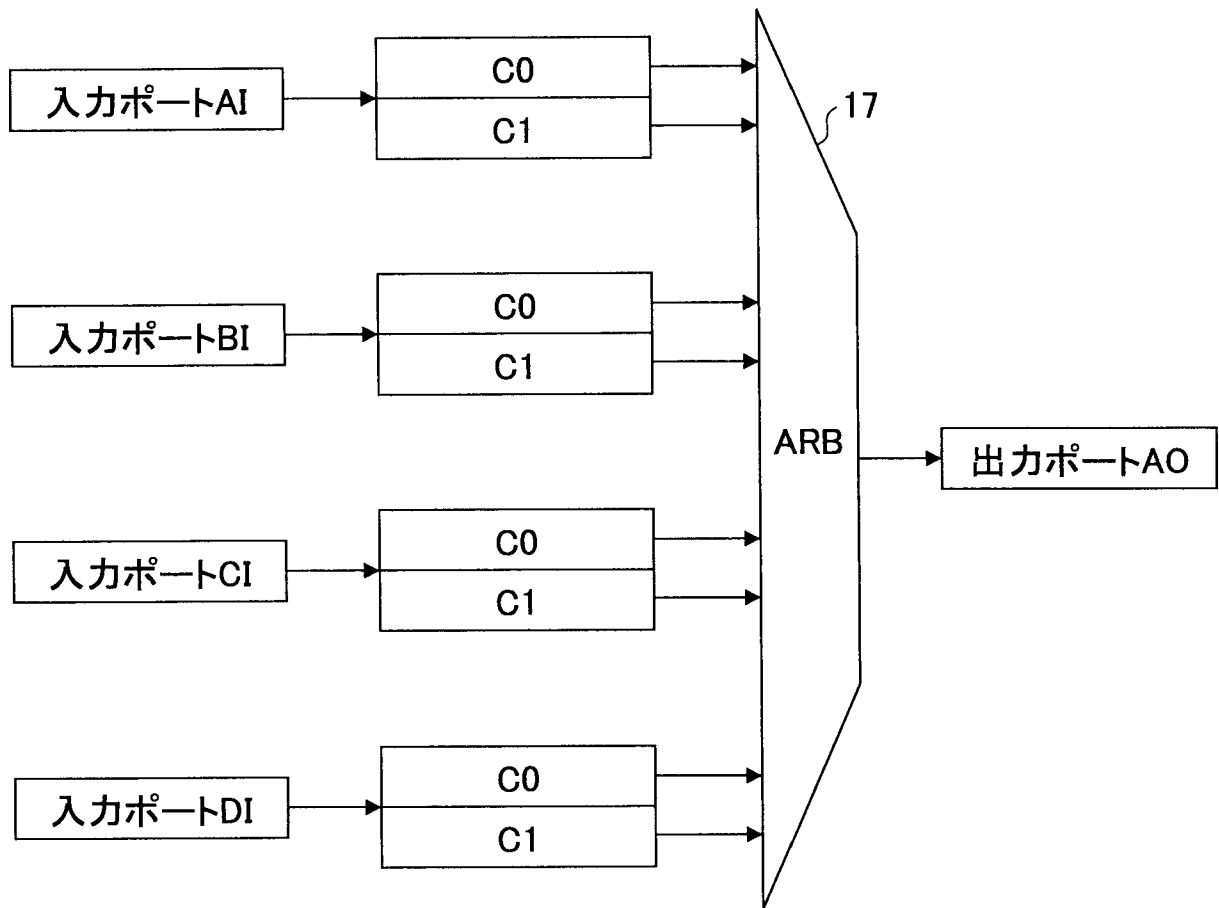
- [14] 前記調停切替回路は、前記第1の調停回路において少なくとも1つの入力ポートにおけるデッドロックが検出されたことを前記トリガとする、請求項11記載の装置。

- [15] 前記任意の調停アルゴリズムは、LRU (Least Recently Used) アルゴリズム、ラウンドロビンアルゴリズム、及び固定プライオリティアルゴリズムからなるグループから選択された1つのアルゴリズムである、請求項11乃至14のいずれか1項記載の装置。

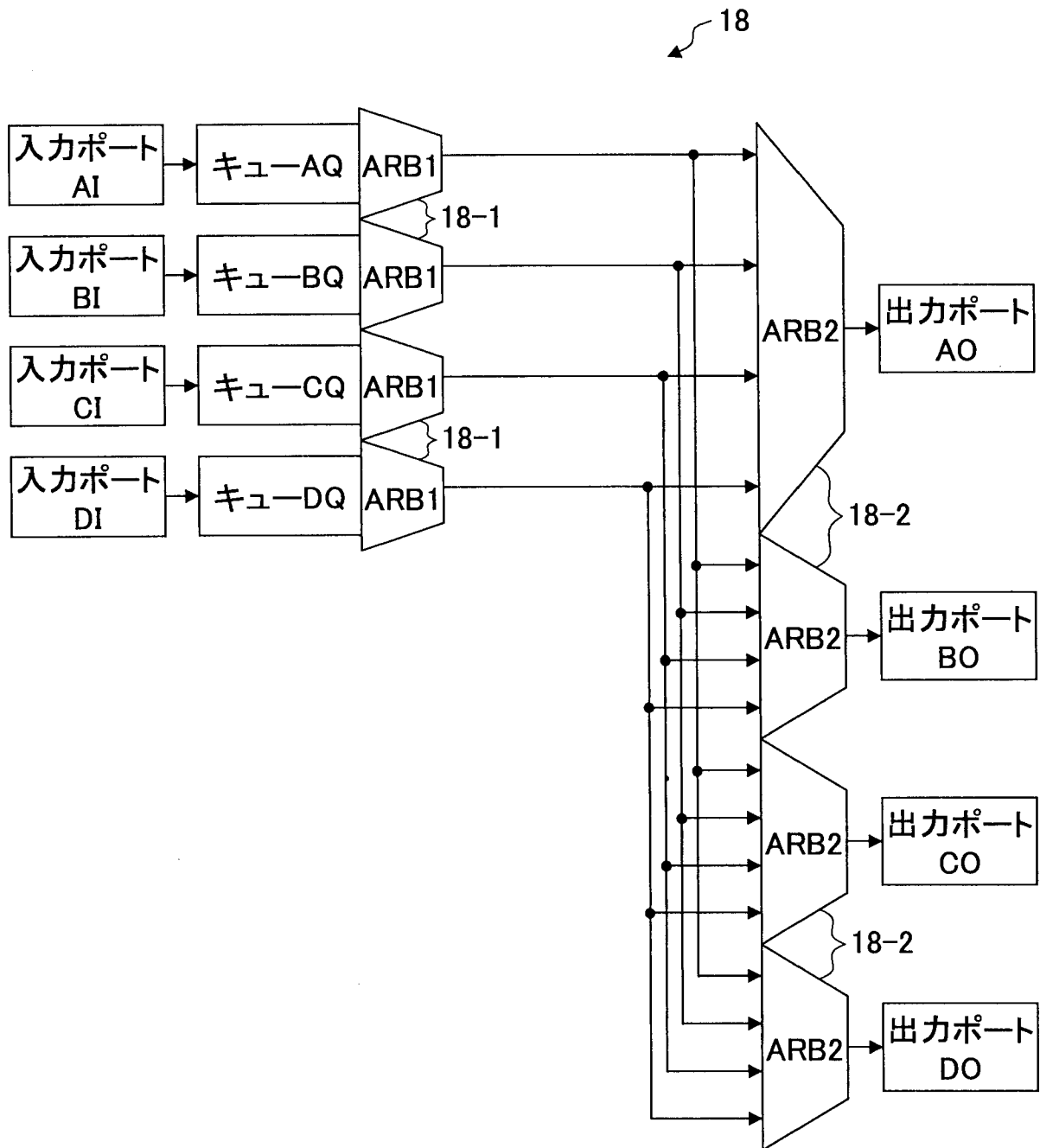
[図1]



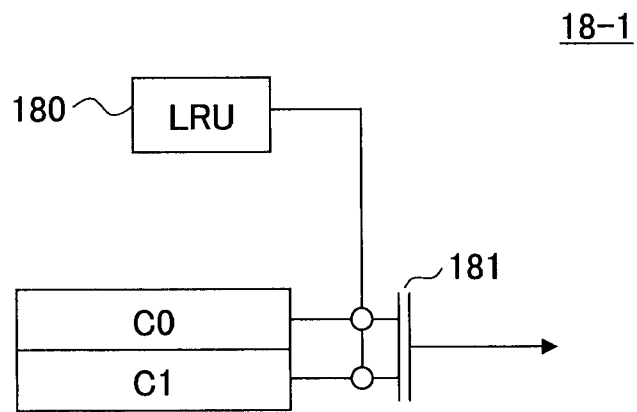
[図2]



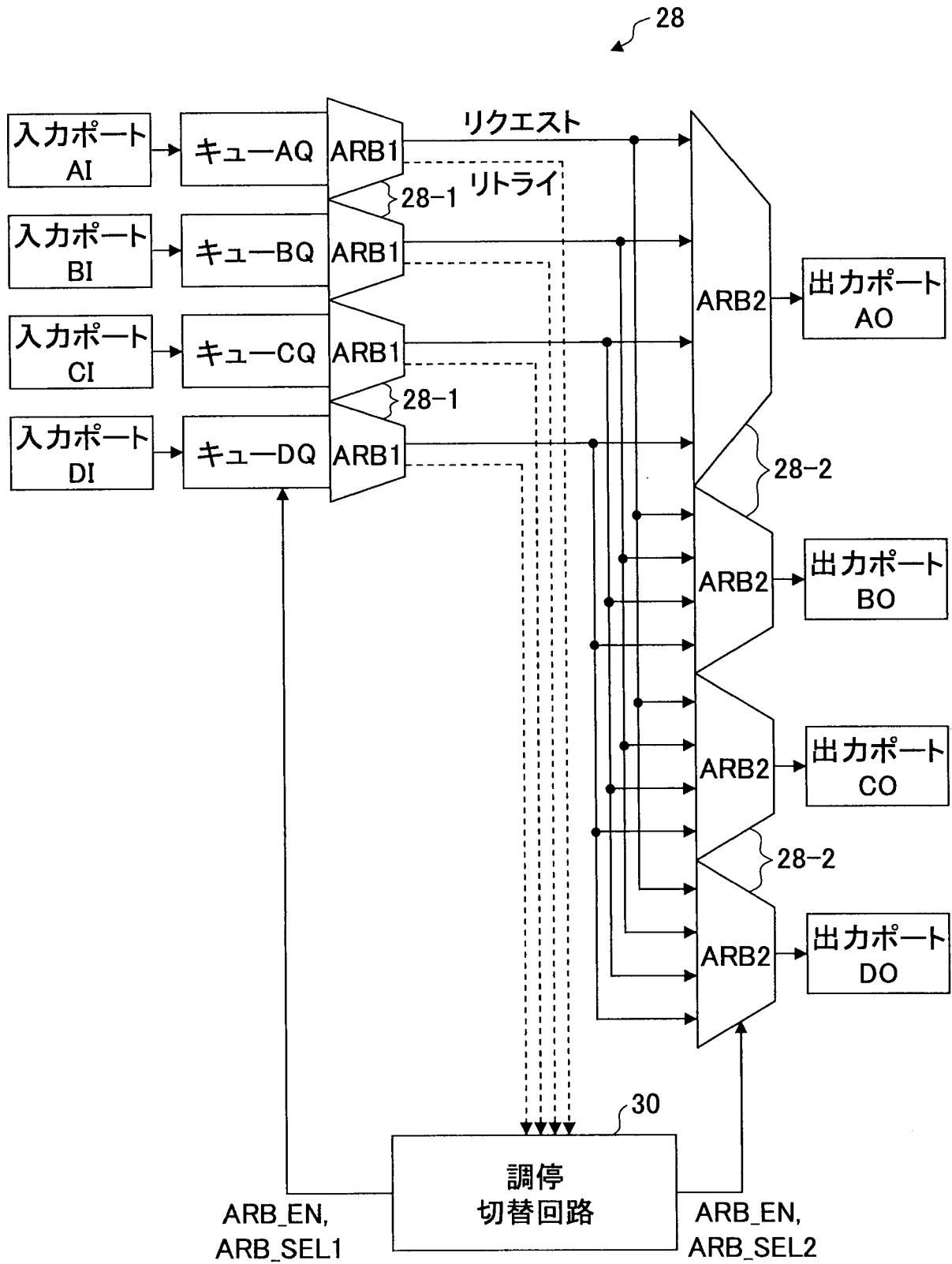
[図3]



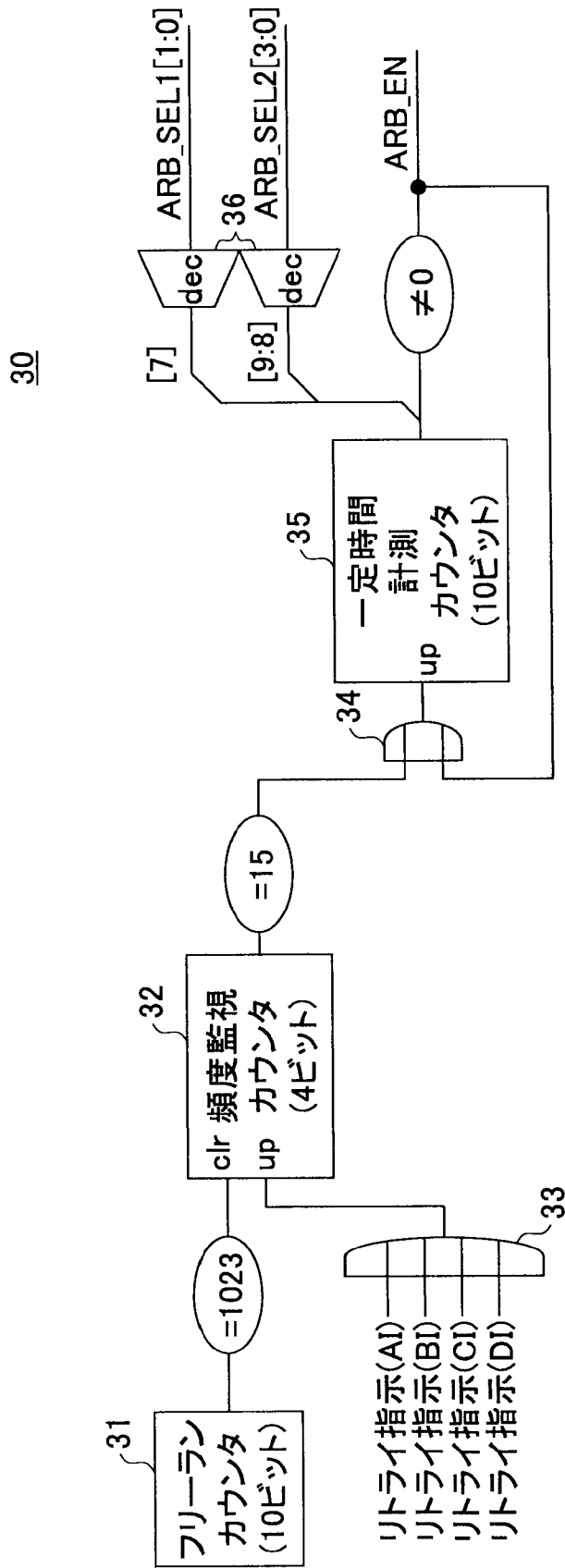
[図4]



[図5]

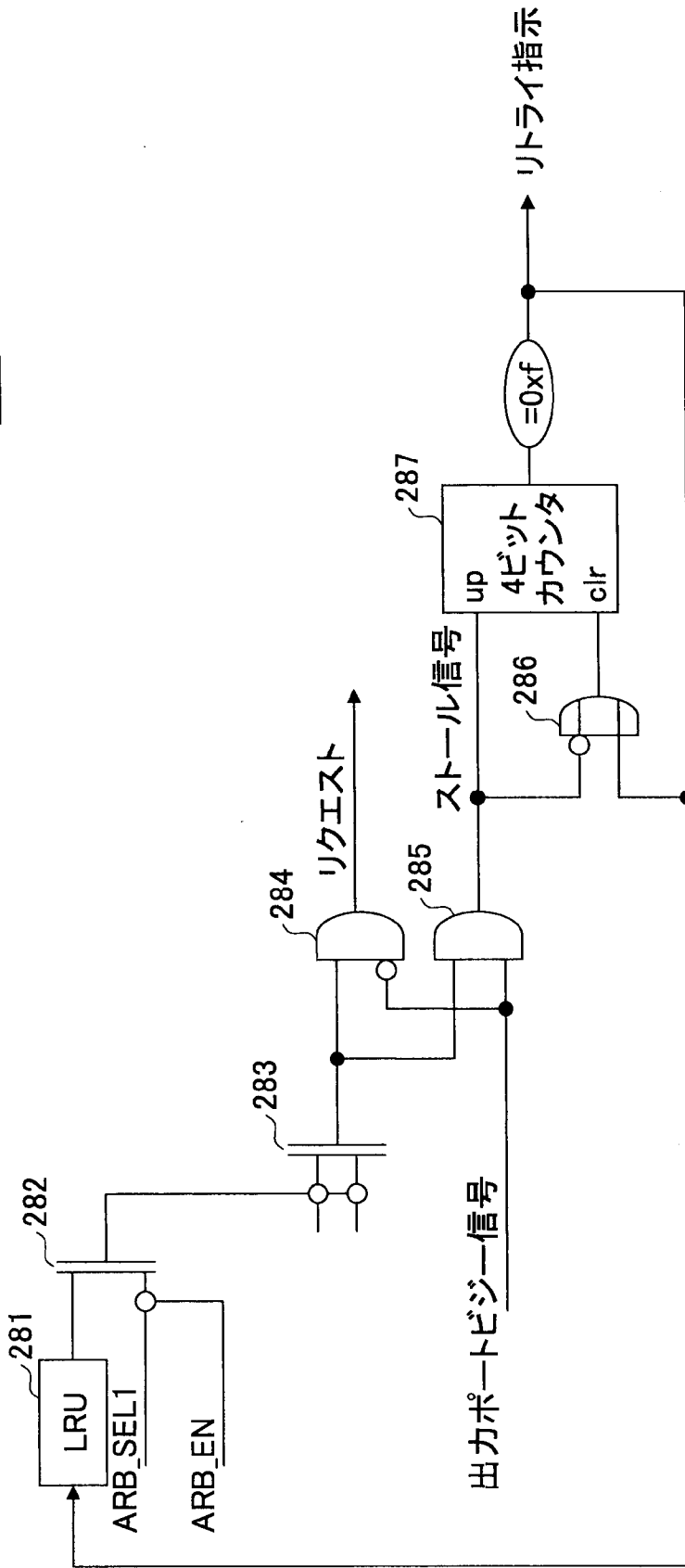


[図6]

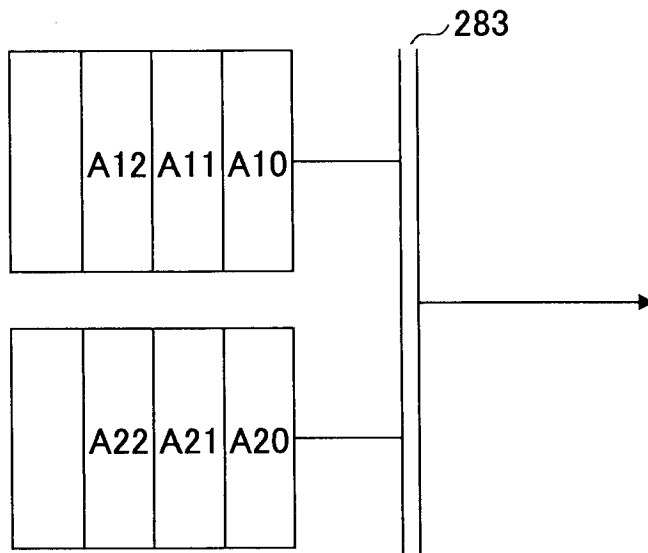


[図7]

28-1



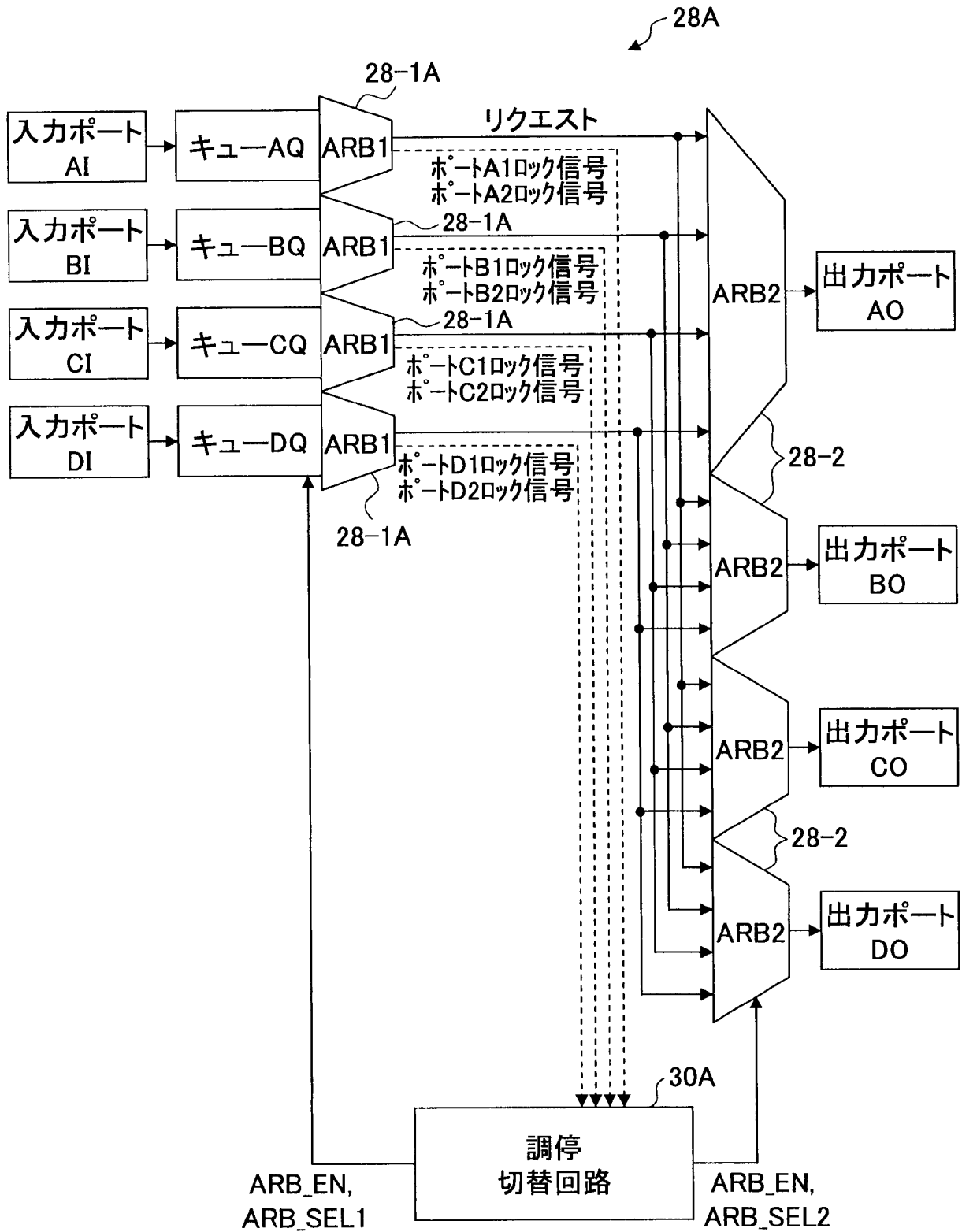
[図8]

AQ

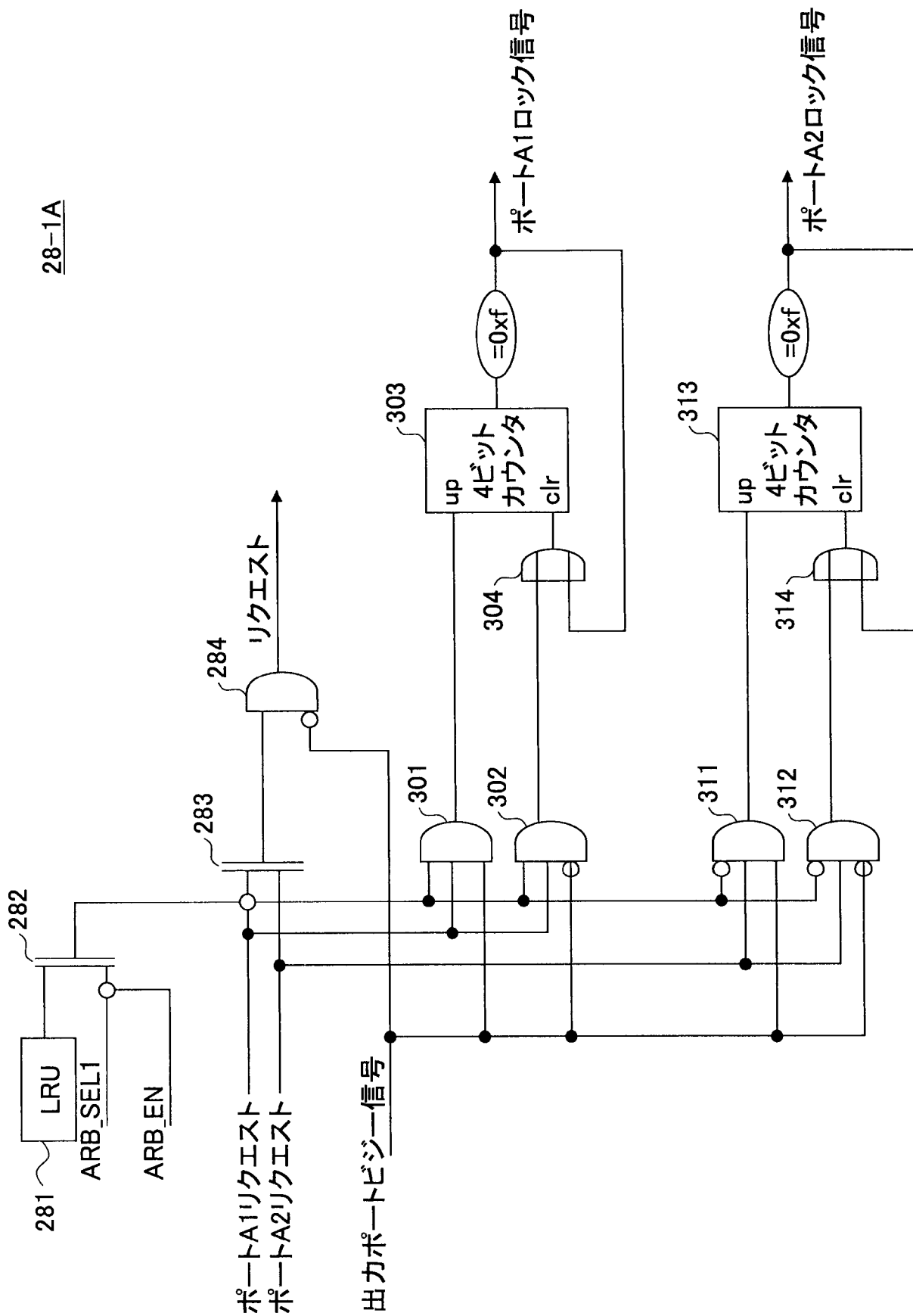
[図9]

出力ポートDO ビット信号	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
リトライ指示の 有無		1		1		1		1		1		1		1		1		1		1		1		1	
入力ポートAI	A10	A20	A20	A11	A11	A20	A20	A20	A11	A20	A20	A20	A20	A12	A12	A12	A20	A20	A20	A12	A20	A20	A13	A13	A20
入力ポートBI	B10	B10	B10	B20	B10	B10	B21	B21	B10	B10	B10	B10	B21	B10	B10	B10	B22	B22	B22	B10	B10	B10	B10	B10	B23
入力ポートCI	C10	C10	C10	C20	C20	C10	C20	C20	C11	C11	C11	C20	C20	C20	C11	C20	C20	C12	C12	C12	C20	C20	C20	C12	
調停結果	A1			B2		C1		A1		B2		C1		A1					A1				B2		C1

[図11]

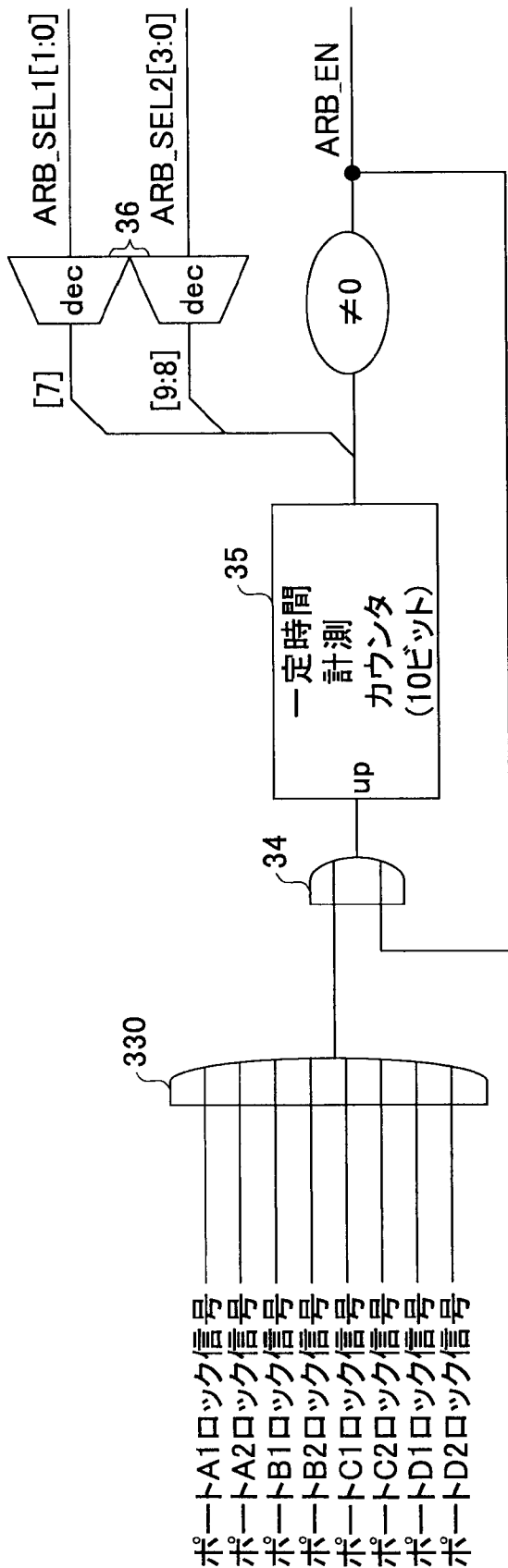


[図12]

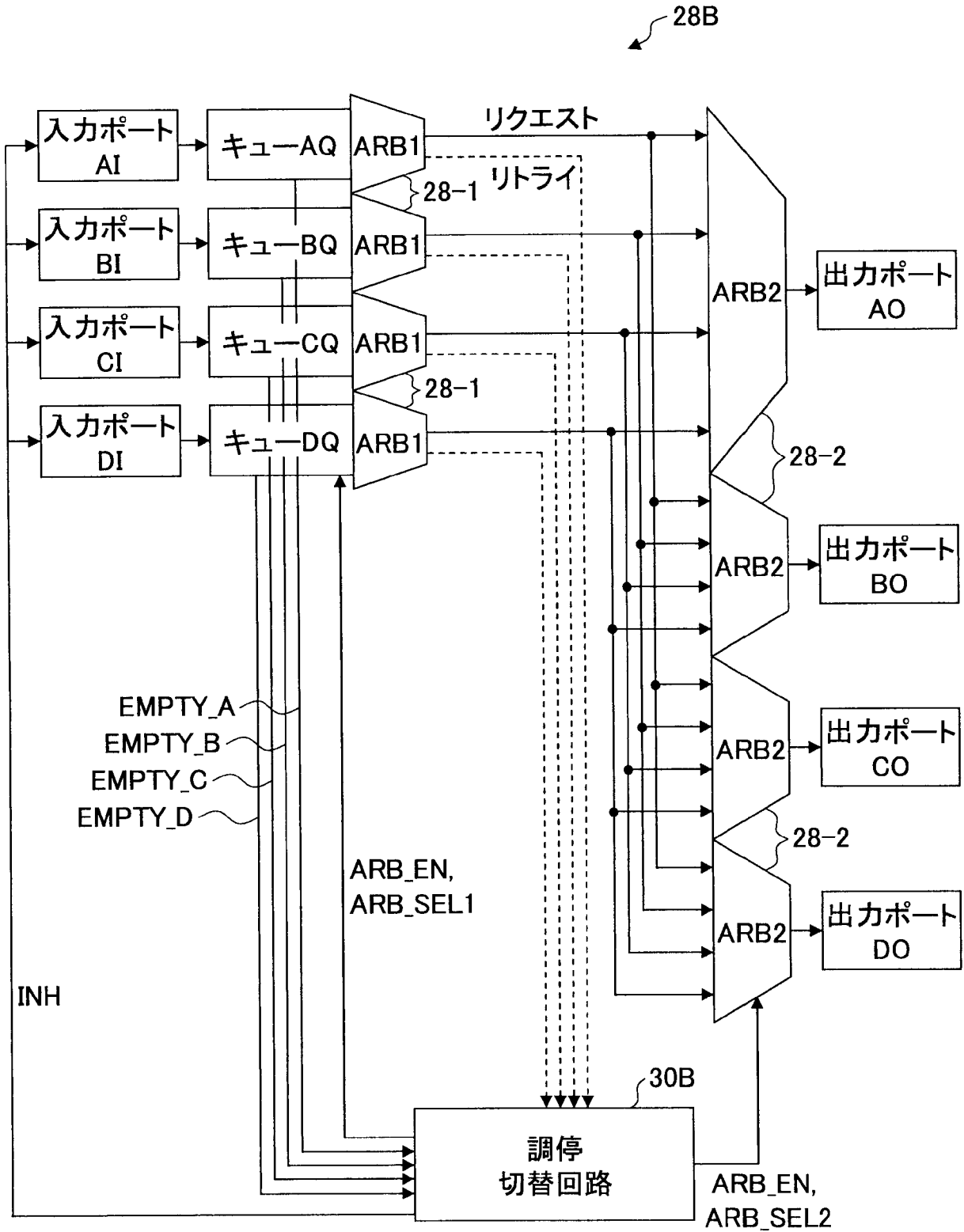


[図13]

30A

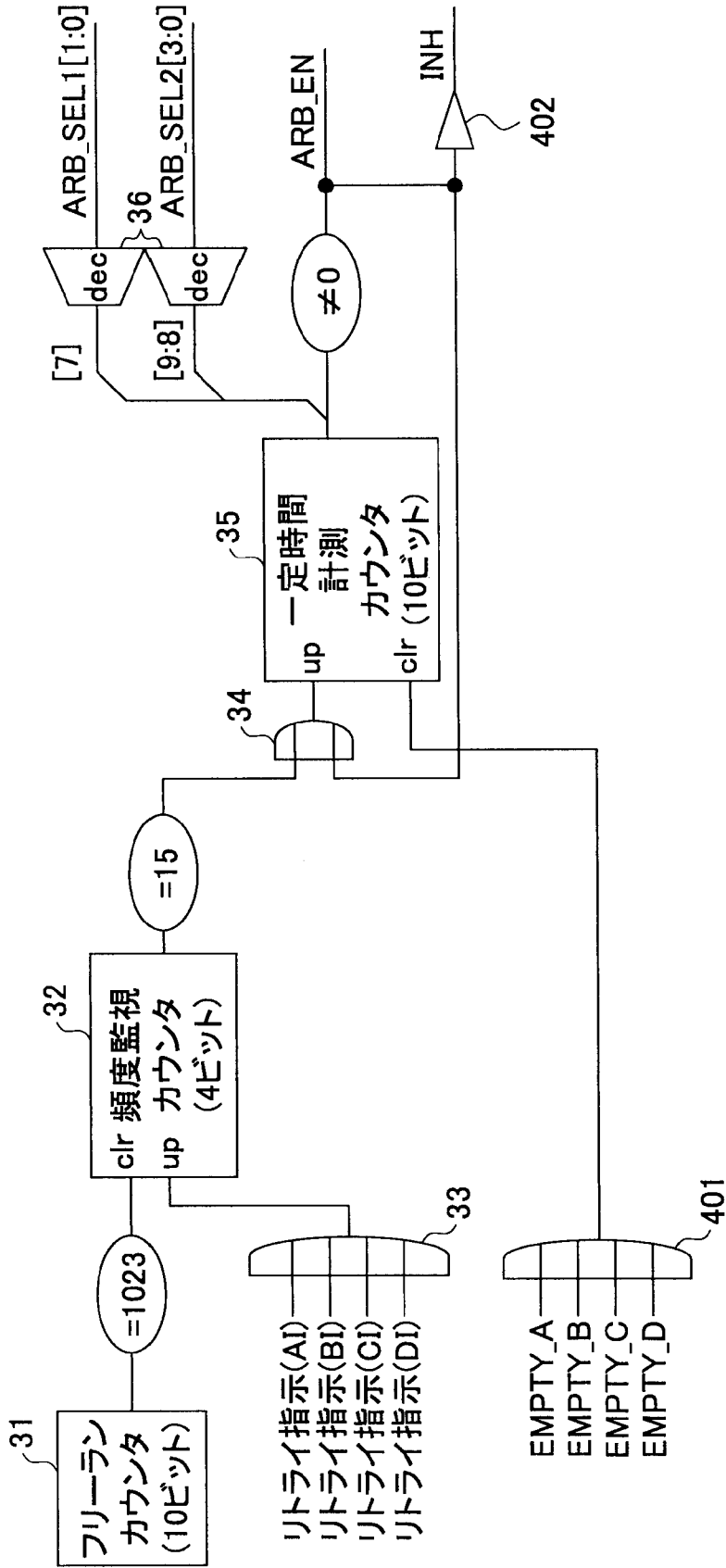


[図14]



[図15]

30B



INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2009/056671

A. CLASSIFICATION OF SUBJECT MATTER
G06F15/173 (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)
G06F15/173

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	1996-2009
Kokai Jitsuyo Shinan Koho	1971-2009	Toroku Jitsuyo Shinan Koho	1994-2009

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.
A	JP 2004-102607 A (NEC Computertechno, Ltd.), 02 April, 2004 (02.04.04), Par. Nos. [0016] to [0057]; Fig. 3 (Family: none)	1-15
A	JP 10-254843 A (Hitachi, Ltd.), 25 September, 1998 (25.09.98), Par. Nos. [0038] to [0068]; Figs. 1 to 2 (Family: none)	1-15

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 19 June, 2009 (19.06.09)	Date of mailing of the international search report 07 July, 2009 (07.07.09)
---	--

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

A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int.Cl. G06F15/173 (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. G06F15/173

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	1971-2009年
日本国実用新案登録公報	1996-2009年
日本国登録実用新案公報	1994-2009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
A	JP 2004-102607 A (エヌイーシーコンピュータテクノ株式会社) 2004.04.02, 段落【0016】 - 【0057】, 第3図 (ファミリーなし)	1-15
A	JP 10-254843 A (株式会社日立製作所) 1998.09.25, 段落【0038】 - 【0068】, 第1-2図 (ファミリーなし)	1-15

☐ 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

19.06.2009

国際調査報告の発送日

07.07.2009

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)

漆原 孝治

5 B

9 3 6 6

電話番号 03-3581-1101 内線 3545