(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2005-141739
(P2005-141739A)
(43)公開日　平成17年6月2日(2005.6.2)

(51)Int.Cl.[7]	F I			テーマコード（参考）
G06F 13/14	G06F	13/14	330E	5B014
G06F 13/36	G06F	13/36	530A	5B061

審査請求 未請求　請求項の数 23　ＯＬ　（全 9 頁）

(21)出願番号　特願2004-311154 (P2004-311154)
(22)出願日　平成16年10月26日 (2004.10.26)
(31)優先権主張番号　10/702,832
(32)優先日　平成15年11月6日 (2003.11.6)
(33)優先権主張国　米国(US)

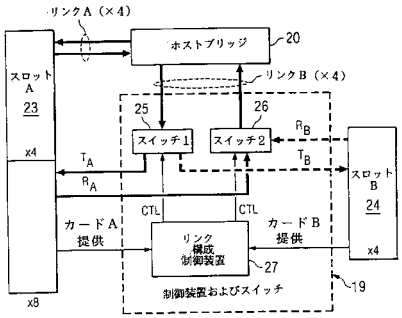
（特許庁注：以下のものは登録商標）
１．ペンティアム

(71)出願人　500391866
デル・プロダクツ・エル・ピー
Ｄｅｌｌ　Ｐｒｏｄｕｃｔｓ，　Ｌ．Ｐ．
アメリカ合衆国、テキサス州７８６８２、ラウンド・ロック、ワン・デル・ウェイ（番地なし）
(74)代理人　100058479
弁理士　鈴江　武彦
(74)代理人　100091351
弁理士　河野　哲
(74)代理人　100088683
弁理士　中村　誠
(74)代理人　100108855
弁理士　蔵田　昌俊

最終頁に続く

(54)【発明の名称】ＰＣＩエクスプレスリンクのダイナミック再構成

(57)【要約】
【課題】本発明は、ＰＣＩエクスプレスレーンのダイナミックな再構成を可能にしてＰＣＩエクスプレスバスの帯域幅の制限を克服することを目的としている。
【解決手段】情報処理システムは、中央処理装置と、そのプログラムを記憶しているメモリと、入力／出力エンドポイントをシステムへ接続し、スイッチ構造とホストブリッジからエンドポイントへのリンクとを有するＰＣＩエクスプレスバスと、ＣＰＵ、メモリ、およびバスを接続するホストブリッジ20と、ＰＣＩエクスプレスバスのリンクを再構成し、１以上のエンドポイントの状態を検出するための制御装置27と、制御装置27からの信号に応答して、情報処理システムが動作している期間中に、そのリンクの全てまたは一部を１つのエンドポイントから別のエンドポイントへ切換えるように動作可能な少なくとも１つのリンクに関連するスイッチ25，26とを有するリンク再構成回路とを具備している
【選択図】　図２



【特許請求の範囲】
【請求項１】
リンクがバス上のエンドポイントへ経路を設定される情報処理システムのＰＣＩエクスプレスバスのリンクを再構成する方法において、
１以上のエンドポイントの状態を検出し、
前記検出ステップの結果に基づいて、１つのエンドポイントから別のエンドポイントへリンクの全てまたは一部を切換えるステップを含んでいる方法。
【請求項２】
検出ステップはエンドポイントがポピュレートされるか否かを検出することにより行われる請求項１記載の方法。
【請求項３】
切換えステップはポピュレートされていないエンドポイントから１以上のポピュレートされているエンドポイントへリンクを切換えることにより行われる請求項２記載の方法。
【請求項４】
検出ステップはポピュレートされたエンドポイントに設置される装置の帯域幅要件を検出することにより行われる請求項１記載の方法。
【請求項５】
切換えステップはＰＣＩエクスプレスバス切換え構造に対して外部のスイッチにより行われる請求項１記載の方法。
【請求項６】
情報処理システムはオペレーティングシステムを有し、検出ステップはオペレーティングシステムに対して外部の回路を使用して行われる請求項１記載の方法。
【請求項７】
情報処理システムはオペレーティングシステムを有し、検出ステップはオペレーティングシステムを使用して行われる請求項１記載の方法。
【請求項８】
切換えステップは２以上の他のエンドポイントへの切換えにより行われる請求項１記載の方法。
【請求項９】
情報処理システムは検出および切換えステップ期間中は動作している請求項１記載の方法。
【請求項１０】
リンクがバス上のエンドポイントへ経路を設定されている情報処理システムのＰＣＩエクスプレスバスのリンクを再構成する回路において、
１以上のエンドポイントの状態を検出する制御装置と、
制御装置からの信号に応答して、情報処理システムが動作している期間中に、１つのエンドポイントから別のエンドポイントへそのリンクの全てまたは一部を切換えるように動作可能な少なくとも１つのリンクに関連するスイッチとを具備している回路。
【請求項１１】
制御装置はエンドポイントがポピュレートされるか否かを検出する請求項１０記載の回路。
【請求項１２】
スイッチはポピュレートされていないエンドポイントから１以上のポピュレートされているエンドポイントへリンクを切換えるように動作可能である請求項１１記載の回路。
【請求項１３】
制御装置はポピュレートされたエンドポイントに設置された装置の帯域幅要求を検出する請求項１０記載の回路。
【請求項１４】
スイッチはＰＣＩエクスプレスバスのスイッチ構造に対して外部のものである請求項１０記載の回路。

【請求項１５】
情報処理システムはホストブリッジを有し、制御装置はホストブリッジに集積されている請求項１０記載の回路。
【請求項１６】
制御装置からの信号はスイッチに対して直接的である請求項１０記載の回路。
【請求項１７】
制御装置からの信号は情報処理システムのオペレーティングシステムを通過する請求項１０記載の回路。
【請求項１８】
中央処理装置と、
この中央処理装置によって実行可能なプログラムを記憶するメモリと、
入力／出力エンドポイントをシステムへ接続し、スイッチ構造とホストブリッジからエンドポイントへのリンクとを有するＰＣＩエクスプレスバスと、
ＣＰＵ、メモリ、およびバスを接続するホストブリッジと、
ＰＣＩエクスプレスバスのリンクを再構成し、１以上のエンドポイントの状態を検出するための制御装置と、
制御装置からの信号に応答して、情報処理システムが動作中である期間中に、そのリンクの全てまたは一部を１つのエンドポイントから別のエンドポイントへ切換えるように動作可能な少なくとも１つのリンクに関連するスイッチとを有するリンク再構成回路とを具備している情報処理システム。
【請求項１９】
前記制御装置はエンドポイントがポピュレートされるか否かを検出する請求項１８記載のシステム。
【請求項２０】
前記スイッチはポピュレートされていないエンドポイントから１以上のポピュレートされているエンドポイントへリンクを切換えるように動作可能である請求項１９記載のシステム。
【請求項２１】
前記制御装置はポピュレートされたエンドポイントに設置されている装置の帯域幅要件を検出する請求項１８記載のシステム。
【請求項２２】
前記スイッチはＰＣＩエクスプレスバスのスイッチ構造に対して外部に位置している請求項１８記載のシステム。
【請求項２３】
前記制御装置はホストブリッジに集積されている請求項１８記載のシステム。
【発明の詳細な説明】
【技術分野】
【０００１】
本発明はコンピュータシステム、特にコンピュータシステムのバス接続に関する。
【背景技術】
【０００２】
プロセッサ、チップセット、キャッシュ、メモリ、拡張カード、記憶装置を含むコンピュータのコンポーネントは相互に１以上の“バス”によって通信する。“バス”は通常のコンピュータ用語では、情報が２以上の装置間で流れるチャンネルである。バスは通常、アクセス点、または装置がバスに接続できる場所を有する。一度接続されると、バス上の装置は他の装置へ情報を送信し、また、そこから受信できる。
【０００３】
今日のパーソナルコンピュータは少なくとも４つのバスを有する傾向がある。各バスはさらにある程度プロセッサから除去され、それぞれはその上のレベルへ接続される。
【０００４】

プロセッサバスは最高レベルのバスであり、プロセッサとの情報の送受信のためにチップセットにより使用される。キャッシュバス（時には、バックサイドバスと呼ばれる）はシステムキャッシュのアクセスのために使用される。メモリバスはメモリサブシステムをチップセットとプロセッサへ接続する。多くのシステムでは、プロセッサおよびメモリバスは同一であり、集合的にフロントサイドバスまたはシステムバスと呼ばれている。

【０００５】

ローカルＩ／Ｏ（入力／出力）バスは周辺機器を、メモリ、チップセット、プロセッサへ接続している。ビデオカード、ディスク記憶装置、ネットワークインターフェースカードは通常このバスを使用する。２つの最も普通のローカルＩ／ＯバスはＶＥＳＡローカルバス（ＶＬＢ）と周辺機器コンポーネント相互接続（ＰＣＩ）バスである。産業標準アーキテクチャ（ＩＳＡ）Ｉ／Ｏバスはまたマウス、モデム、低速度の音響およびネットワーク化装置のような低速度の周辺機器でも使用されることができる。

【発明の開示】

【発明が解決しようとする課題】

【０００６】

現世代のＰＣＩバスはＰＣＩエクスプレスバスとして知られている。このバスは高帯域幅シリアルバスであり、既存のＰＣＩ装置とのソフトウェアの競合性を維持する。

【０００７】

本発明の目的は、ＰＣＩエクスプレスレーンのダイナミックな再構成を可能にしてＰＣＩエクスプレスバスの帯域幅の制限を克服することである。

【課題を解決するための手段】

【０００８】

本発明の１特徴はＰＣＩエクスプレスバスのリンクを再構成する方法である。エンドポイントがポピュレートされているか否か、エンドポイントが必要とする帯域幅の量のようなバスのエンドポイントの状態が検出される。この検出に基づいて、未使用の帯域幅を有している全てまたは一部のリンクが別のエンドポイントに切換えられることができる。

【０００９】

例えば、ポピュレートされていないエンドポイントへ経路を設定されているリンクの全てのレーンはポピュレートされたエンドポイントへ再度経路を設定されることができる。またさらに別の例では、そのリンクにより提供される帯域幅よりも少ない帯域幅を必要とするエンドポイントへ経路を設定されたリンクの１以上のレーンはさらに多くの帯域幅を必要とするエンドポイントへ切換えられてもよい。

【００１０】

本発明の利点はＰＣＩエクスプレスバスの帯域幅の制限を克服することを助けることである。ＰＣＩエクスプレスレーンのダイナミックな再構成は未使用の帯域幅がバスの他の装置へ切換えられることを許容する。

【発明を実施するための最良の形態】

【００１１】

本発明の実施形態およびその利点のさらに完全な理解は添付図面を伴った以下の説明を参照することにより得られる。図面の中で同一の参照符号は類似の特性を示す。

図１は、本発明にしたがった情報処理システムの種々の内部素子を示している。以下説明するように、システム100 はＰＣＩエクスプレスバス17と、バスの１以上のリンク17bをダイナミックに再構成する付加的な回路19を有している。ＰＣＩエクスプレスバス17は周辺機器のコンポーネントを接続する一般的な方法で使用されるが、エンドポイント17cの状態が検出され、そのエンドポイントの帯域幅がそのエンドポイントで必要とされないならば再度経路を設定されるように強化されている。

【００１２】

図１の実施形態では、システム100 は典型的なパーソナルコンピュータシステムであるが、サーバ、ワークステーション、または埋設されたシステムのような幾つかの他のタイプの情報処理システムであってもよい。この説明のためには、情報処理システムはビジネ

ス、科学、制御またはその他の目的のために情報、インテリジェンスまたはデータの任意の形態を計算、分類、処理、送信、受信、検索、発信、切換え、記憶、表示、マニフェスト、検出、記録、再生、処理または使用するように動作する任意の手段または手段の集合を含むことができる。例えば、情報処理システムはパーソナルコンピュータ、ネットワーク記憶装置、または任意の他の適切な装置であってもよく、寸法、形状、性能、機能、価格において変化することができる。情報処理システムはランダムアクセスメモリ（ＲＡＭ）、中央処理装置（ＣＰＵ）のような１以上の処理リソース、ハードウェアまたはソフトウェア制御論理装置、ＲＯＭ、および／またはその他のタイプの不揮発性メモリを含むことができる。情報処理システムの付加的なコンポーネントは１以上のディスクドライブ、外部装置と通信するための１以上のネットワークポート、キーボード、マウス、およびビデオディスプレイのような種々の入力および出力（Ｉ／Ｏ）装置を含むことができる。情報処理システムはまた種々のハードウェアコンポーネント間で通信を送信するように動作可能な１以上のバスを含むことができる。

【００１３】

ＣＰＵ10は任意の中央処理装置であってもよい。典型的なＣＰＵ10の例はインテル社から入手可能なペンティアムファミリのプロセッサである。本発明の目的に対しては、ＣＰＵ10は少なくともＢＩＯＳ（基本入力／出力システム）プログラミングを有するオペレーティングシステムを実行するようにプログラムされている。

【００１４】

ホストブリッジ11（しばしばノースブリッジと呼ばれる）はＣＰＵ10をエンドポイント12と、メモリ13と、ＰＣＩエクスプレスバス17とに接続するチップ（またはチップセットの一部）である。ホストブリッジ11に接続されるエンドポイント12のタイプはアプリケーションにしたがう。例えばシステム100 がデスクトップコンピュータであるならば、エンドポイント12は典型的にグラフィックアダプタ、（シリアルＡＴＡリンクを介する）ＨＤＤ、（ＵＳＢリンクを介する）ローカルＩ／Ｏである。サーバに対しては、エンドポイント12は典型的にＧｂＥ（ギガビットのイーサネット（Ｒ））と、ＩＢＥ装置と、付加的なブリッジ装置である。

【００１５】

ＣＰＵ10とホストブリッジ11間の通信はフロントサイドバス14を介している。

【００１６】

ＰＣＩエクスプレスバス17はスイッチ構造17a とリンク17b とを具備しており、それによって多数のＰＣＩエンドポイント18が接続されることができる。スイッチ構造17a はホストブリッジ11からリンク17b へファンアウトを提供し、リンクスケーリングを行う。

【００１７】

“リンクスケーリング”はＰＣＩエクスプレスバス17の利用可能な帯域幅が割当てられることを意味し、それによってそれぞれＰＣＩエクスプレスアーキテクチャ標準に適合するサイズをそれぞれ有する予め定められた数のリンク17b は物理的にエンドポイント18へ経路を設定される。各リンク17b は１以上のレーンを備えている。単一のレーン（ｘ１幅を有するとして呼ばれる）を有するリンクは２つの低電圧差動対を有し、２つの装置の間で二重の単向シリアル接続である。２つの装置の間のデータ伝送は両方向で同時である。スケール可能な性能は広いリンク幅（ｘ１、ｘ２、ｘ４、ｘ８、ｘ１６、ｘ３２）により実現可能である。リンクは対称的にスケールされ、各方向に同一数のレーンを有する。

【００１８】

ＰＣＩエンドポイント18はカードスロットまたはその他の接続機構を使用して物理的に接続される周辺機器装置またはチップであってもよい。ＰＣＩエクスプレスバス17に接続される特定のエンドポイント18はシステム100 のアプリケーションのタイプにしたがう。デスクトップコンピュータシステムでは、典型的なＰＣＩエンドポイント18の例はモバイルドッキングアダプタ、イーサネット（Ｒ）アダプタ、装置における他の付加物である。サーバプラットフォームでは、エンドポイント18はギガビットのイーサネット（Ｒ）接続、Ｉ／Ｏおよびクラスタ相互接続のための付加的なスイッチング能力である。通信プラッ

トフォームでは、エンドポイント18はラインカードであってもよい。

【００１９】

通常のＰＣＩエキスプレスバス17では、スイッチ構造17a は別々のコンポーネントとして、またはホストブリッジ11を含むコンポーネントの一部として構成された論理素子である。以下説明するように、本発明では、ＰＣＩエクスプレスバス17は付加的なスイッチングおよび制御回路19と共に動作する。この回路19はエンドポイント18の状態を検出し、１つのエンドポイントから別のエンドポイントへリンクを切換えることができる。

【００２０】

図２はシステム100 の部分図であり、本発明にしたがってＰＣＩエクスプレスリンク17b のダイナミックな再構成を示している。各リンク17b は２つの信号対、即ち送信対と受信対として示されている。送信対はＴ信号として示され、受信対はＲ信号として示されている。

【００２１】

スロット23および24はカードタイプのエンドポイント18を接続するように設計されている。２つのスロットしか示されていないが、リンクの所望のスケーリング（ｘ１、ｘ４等）にしたがって任意の数のスロット構造が可能である。スロット23および24は典型的にシステム100 のコンピュータシャーシ内の物理的位置を表しており、ここで種々のＩ／Ｏ装置に対するカードがインストールされることができる。他の実施形態では、システム100 はスロット接続に加えてまたはその代わりに１以上のチップ接続を有することもできる。一般的に、用語“エンドポイント接続”はチップ、カード、または任意の他のタイプのエンドポイントを集合的に指すために使用される。

【００２２】

図２の例では、スロット23はｘ４リンク幅（リンクＡ）で構成される。スロット24はｘ４リンク幅（リンクＢ）で構成される。

【００２３】

再構成はスイッチ25と26およびリンク構成制御装置27を使用して実現される。図２は１例であり、リンク、スロット、スイッチの数の変化と、種々のリンク幅により、多数の異なるバリエーションのスイッチングおよび制御回路が可能であることを理解すべきである。

【００２４】

リンク構成制御装置27はスロット23および24のどちらが（使用において）ポピュレートされるかを検出する。ＰＣＩバス17はスロットが“ホットプラグ”および、“ホットスワップ”されることを許容するので、この検出は装置がスロット23または24に取付けられるか取外されるときにはいつでも、制御装置27が迅速にその事象を検出する意味でダイナミックである。

【００２５】

リンク構成制御装置27にはプログラム可能な論理装置により構成され、独立式の論理回路であってもよく、または他のシステム論理装置と集積されてもよい。例えばリンク構成制御装置がホストブリッジ20に集積されることができる。

【００２６】

スロットの状態（ポピュレートされているか、されていない）が変化するならば、制御装置27は信号をスイッチ25と26へ出力する。スイッチ25と26には高速スイッチング装置が構成されることができる。制御装置27のように、スイッチ25と26は制御装置27および／またはホストブリッジ20のような他の回路と集積されることができる。

【００２７】

図２の例では、リンクＢはその送信レーンにスイッチ25を有し、その受信レーンにスイッチ26を有している。スイッチ25と26は両者ともリンクＢをスロット23またはスロット24へ切換えるように動作可能である。リンクＢがスロット23に切換えられるならば、スロット23はｘ８リンクを受ける。リンクＢがスロット24に切換えられるならば、スロット24はｘ４リンクを受ける。スイッチ25および26とスロット23の間の適切な物理的接続が行われ

、それによって代わりのパス間の切換えが可能であることが仮定される。

【００２８】

この例では、スロット23はポピュレートされ、スロット24はポピュレートされていない。この状態は制御装置27により検出され、これは全てのリンクＢをスロット23へ切換えるようにスイッチ25と26を設定する。

【００２９】

図３は本発明の動作の別の例を示している。この例では、スロット33と34の両者がポピュレートされている。システムは３つのｘ４リンクにより構成されている。リンクＡはｘ４リンクであり、スロット33へ経路を設定されている。リンクＢもまたｘ４リンクであり、スロットＢへ経路を設定されている。リンクＣはｘ４リンクであり、スイッチ35と36へ経路を設定され、“切換え可能な”リンクにしている。

【００３０】

制御装置27はスロット33と34の両者がポピュレートされていることを検出するが、スロット33がｘ８リンクを必要とし、スロット34がｘ４リンクだけを必要とすることも検出する。応答において、制御装置27は制御信号をスイッチ35と36へ伝送し、それによってリンクＣはスロット33へ経路を設定され、ｘ８スロットにする。この例では、スロット33および34をポピュレートするカードは（直接的にまたはシステム100 のオペレーティングシステムを介して）制御装置27にそれらの帯域幅の要求を知らせる幾つかの手段であることが仮定される。

【００３１】

図４は第３の例を示し、ここではスイッチはリンクの一部だけが再度経路を設定されるようにエンドポイントへのリンクを再構成するために使用される。図４の例では、スロット43およびスロット44の既存の構造はそれぞれｘ４とｘ８である。しかしながらｘ８のエンドポイントはスロット43に置かれ、ｘ４のエンドポイントはスロット44に置かれている。制御装置27は両スロットの状態および帯域幅の要求を検出し、リンクＢの一部がスロット43へ再度経路を設定されるようにスイッチ45と46を動作する。この例の変形では、スロット44はポピュレートされず、リンクＢはスロット43へのｘ４バスと幾つかの他のエンドポイントへのｘ４パスへ分割されるように切換えられる。

【００３２】

前述の例はこれらが既存のリンク、即ち既に物理的にバス上の種々のエンドポイントへ物理的に経路を設定されているリンクを再度経路設定する意味で“再構成”を実現する。本発明がなければ、ＰＣＩエクスプレスバスはシステム100 の開始において設定されるあらゆるリンク構造にもしたがって動作する。さらに、前述の例の方法および回路はシステム100 が動作のために付勢されながら（スタートアップ期間中）およびオペレーティングシステムが作動中でありながら、状態検出および切換えが行われる意味で“ダイナミック”である。したがって、状態検出はエンドポイントの実時間（現在）状態である。本発明の検出およびスイッチングはＰＣＩエクスプレスバスのスケーリング能力にしたがって先にスケールされているリンクで動作する。スケーリングから生じる静止構造と比較するとき、これはダイナミックな再構成である。

【００３３】

前述の例では、制御装置27はスロットの状態の検出と、構成スイッチへの制御信号の出力の両者を行う。別の実施形態では、これらの機能の一方または両者はそのＢＩＯＳによるような、システム100 のオペレーティングシステムにより行われる。即ち、ＢＩＯＳはそのＰＣＩエクスプレスバス40のスロットの状態を検出し、および／またはその状態に応答してレーンを切換えるようにプログラムされている。したがって、種々の実施形態では、本発明の検出およびスイッチング機能はハードウェアまたはソフトウェアで制御されることができる。

【００３４】

再構成は本発明の“ダイナミック”検出特性なしでも有効である。換言すると、既存のＰＣＩエクスプレスバスリンクをマニュアルで再度経路を設定することが望まれる状態が

存在する。例えば、ｘ８リンクを必要とするカードは物理的にｘ４リンクを有するシャーシ内のスロットに適合する。ｘ８カードはｘ４カードで切換えられ、それらのリンクは再度経路を設定される。

【図面の簡単な説明】

【００３５】

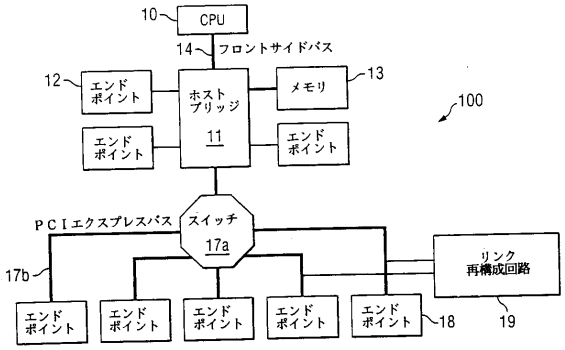
【図１】本発明にしたがった情報処理システムの種々の内部素子を示す図。

【図２】図１のシステムの一部の、リンクを再構成するための構成の第１の例を示すブロック図。

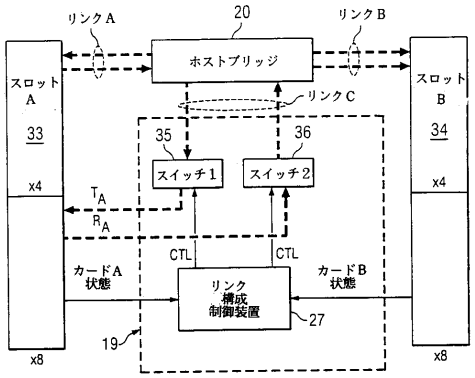
【図３】リンクを再構成する構成の第２の例を示すブロック図。

【図４】リンクを再構成する構成の第３の例を示すブロック図。

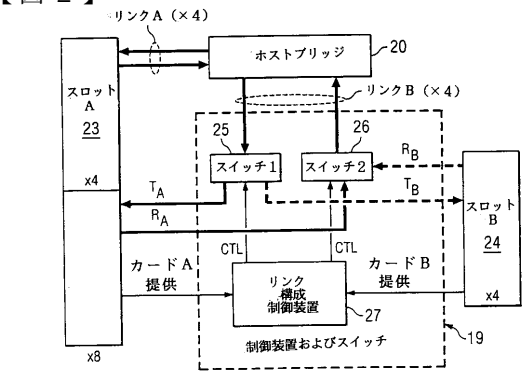
【図１】



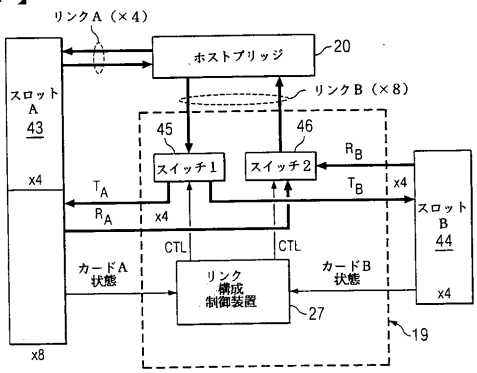
【図３】



【図２】



【図４】



フロントページの続き

(74)代理人 100109830
弁理士 福原 淑弘
(74)代理人 100084618
弁理士 村松 貞男
(74)代理人 100092196
弁理士 橋本 良郎
(72)発明者 マーティン・マッカーフィー
アメリカ合衆国、テキサス州 ７８６４５、ラゴ・ビスタ、パトリオット・ドライブ １９０９
(72)発明者 ルイス・エヌ・キャストロ
アメリカ合衆国、テキサス州 ７８６１３、シーダー・パーク、ウィートン・トレイル ２１１１

Ｆターム(参考) 5B014 HC03 HC07 HC13
5B061 FF01