(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号

特許第4909384号
(P4909384)

(45)発行日　平成24年4月4日(2012.4.4)　(24)登録日　平成24年1月20日(2012.1.20)

(51)Int.Cl.		F I		
H04L 29/10	(2006.01)	H04L	13/00	309Z
G06F 13/42	(2006.01)	G06F	13/42	350C
G06F 13/38	(2006.01)	G06F	13/38	320A
H04L 12/56	(2006.01)	H04L	12/56	300Z

請求項の数 19　外国語出願　(全 15 頁)

(21)出願番号　特願2009-165061 (P2009-165061)
(22)出願日　平成21年7月13日 (2009.7.13)
(65)公開番号　特開2010-88106 (P2010-88106A)
(43)公開日　平成22年4月15日 (2010.4.15)
審査請求日　平成21年7月14日 (2009.7.14)
(31)優先権主張番号　12/218,410
(32)優先日　平成20年7月15日 (2008.7.15)
(33)優先権主張国　米国 (US)

(73)特許権者　591003943
インテル・コーポレーション
アメリカ合衆国　９５０５２　カリフォルニア州・サンタクララ・ミッション　カレッジ　ブーレバード・２２００
(74)代理人　110000877
龍華国際特許業務法人
(72)発明者　ハリマン、デービット　ジェイ.
アメリカ合衆国　９５０５２　カリフォルニア州・サンタクララ・ミッション　カレッジ　ブーレバード・２２００　インテル・コーポレーション内

審査官　森谷　哲朗

最終頁に続く

(54)【発明の名称】プロトコルスタックのタイミングの管理

(57)【特許請求の範囲】
【請求項１】
第１のプロトコルにしたがってデータを処理する第１のプロトコルスタックであって、前記第１のプロトコルスタックをトンネリングインターコネクトにインターフェースするインターフェースロジックを有する前記第１のプロトコルスタックと、
前記第１のプロトコルスタックをリンクに結合する前記トンネリングインターコネクトと
を備え、
前記インターフェースロジックは、前記トンネリングインターコネクトによるトンネリングによって発生するタイミング遅延に少なくとも部分的に基づいて前記第１のプロトコルスタックの少なくとも１つのタイマを制御する
装置。
【請求項２】
前記インターフェースロジックは、前記タイミング遅延に対応付けられているタイミング遅延情報を、前記第１のプロトコルスタックの少なくとも１つのスタックロジックの、タイミングについての制約を示すタイミング要件に対応付ける
請求項１に記載の装置。
【請求項３】
前記インターフェースロジックは、前記タイミング遅延情報、及び前記タイミング遅延情報に対応付けられている前記タイミング要件に少なくとも部分的に基づいて、前記第１

のプロトコルスタックの前記少なくとも１つのタイマを制御することによって、前記第１のプロトコルスタックのタイミングを変更するか否かを判断する
請求項２に記載の装置。
【請求項４】
前記インターフェースロジックは、前記タイミング遅延情報を前記タイミング要件に動的に対応付けして、前記第１のプロトコルスタックは、前記トンネリングインターコネクトの共通物理層または別の物理層に動的に結合される
請求項２に記載の装置。
【請求項５】
前記インターフェースロジックは、前記第１のプロトコルスタックの第１のスタックロジックが、前記第１のプロトコルの、タイミングについての制約を示すリンクタイミング要件を満たすべく、前記第１のスタックロジックに第１のクロック信号を供給する第１のクロックを、予め定められた時間にわたってディセーブルする
請求項１に記載の装置。
【請求項６】
前記トンネリングインターコネクトは、前記トンネリングインターコネクトのプロトコルによって、前記第１のプロトコルのパケットを前記リンクにトンネリングする
請求項５に記載の装置。
【請求項７】
前記リンクは、前記第１のプロトコルスタックと第２のプロトコルスタックとの間で共有する統合型インターコネクトであり、前記第１のプロトコルスタックは、ＰＣＩｅ（Ｐｅｒｉｐｈｅｒａｌ Ｃｏｍｐｏｎｅｎｔ Ｉｎｔｅｒｃｏｎｎｅｃｔ Ｅｘｐｒｅｓｓ（登録商標））スタックである
請求項１から６のいずれか１項に記載の装置。
【請求項８】
前記トンネリングインターコネクトは、前記第１のプロトコルスタックに第１のスロットおよび第２のスロットを割り当てて、前記第２のプロトコルスタックに第３のスロットを割り当てる
請求項７に記載の装置。
【請求項９】
前記リンクに結合されており、前記トンネリングインターコネクトによりトンネリングされたパケットを受信する受信機
をさらに備え、
前記受信機は、前記リンクに結合されているインターフェースロジックによって、割り当てられた前記第１のスロットおよび前記第２のスロットを利用する
請求項８に記載の装置。
【請求項１０】
トンネリングインターコネクトに結合されている第１のプロトコルスタックのインターフェースロジックにおいて通信を受信する段階と、
前記通信の通信種類に基づいてテーブルにアクセスして、前記通信種類に対応付けられている、前記トンネリングインターコネクトによるトンネリングによって発生するタイミング遅延を示すタイミング遅延情報を取得する段階と、
前記タイミング遅延情報が示す遅延に対応するべく、前記通信の通信種類について前記第１のプロトコルスタックの少なくとも１つのスタックロジックのタイミングを制御するべきか否かを決定する段階と、
前記決定する段階において変更するべきと決定した場合に、前記第１のプロトコルスタックの前記少なくとも１つのスタックロジックのタイミングを調整して、前記遅延に対応する段階と、
調整された前記タイミングを用いて前記第１のプロトコルスタックにおいて前記通信を処理する段階と

を備える方法。
【請求項１１】
前記テーブルは、不揮発性メモリに格納されており、前記トンネリングインターコネクトと前記第１のプロトコルスタックとの間のマッピングを含む第１の部分と、前記トンネリングインターコネクトと前記トンネリングインターコネクトに結合されている第２のプロトコルスタックとの間のマッピングを含む第２の部分とを有する
請求項１０に記載の方法。
【請求項１２】
前記第１のプロトコルスタックと前記第２のプロトコルスタックとの間で前記トンネリングインターコネクトを共有する段階と、
前記第１のプロトコルスタックがあるスロットにおいて通信対象の情報を持たない場合に、前記第２のプロトコルスタックに前記第１のプロトコルスタックの前記スロットを与える段階と
をさらに備える、請求項１１に記載の方法。
【請求項１３】
前記少なくとも１つのスタックロジックの前記タイミングは、前記少なくとも１つのスタックロジックに結合されているクロックをオフにすることによって調整される
請求項１０から１２のいずれか１項に記載の方法。
【請求項１４】
前記遅延に少なくとも部分的に基づいて、第２のスタックロジックのクロックを遅延させる段階
をさらに備える、請求項１３に記載の方法。
【請求項１５】
前記クロックを遅延して、予め定められた時間が経過するまで、前記トンネリングインターコネクトによって時間要件が満たされない場合にトリガされる、受信機からの肯定応答を受領していない旨を指し示すエラー信号が発行されないようにする段階
をさらに備える、請求項１４に記載の方法。
【請求項１６】
リンクに結合されている物理層、および、前記物理層に結合されているプロトコルスタックを有する送信機と、
前記リンクを介して前記送信機に結合されており、第１のプロトコルにしたがってデータを処理する第１のプロトコルスタックを有する受信機と、
前記受信機に結合されているダイナミックランダムアクセスメモリ（ＤＲＡＭ）と
を備え、
前記第１のプロトコルスタックは、トンネリング物理層を介して、前記リンクに前記第１のプロトコルスタックをインターフェースする第１のインターフェースロジックを含み、前記第１のインターフェースロジックは、前記トンネリング物理層によるトンネリングによって発生するタイミング遅延に少なくとも部分的に基づいて、前記第１のプロトコルスタックの少なくとも１つの第１のスタックロジックのタイミングを変更する
システム。
【請求項１７】
前記受信機はさらに、第２のプロトコルにしたがってデータを処理する第２のプロトコルスタックを有し、前記第２のプロトコルスタックは、前記タイミング遅延に少なくとも部分的に基づいて前記第２のプロトコルスタックの少なくとも１つの第２のスタックロジックのタイミングを変更する第２のインターフェースロジックを含む
請求項１６に記載のシステム。
【請求項１８】
前記トンネリング物理層は、前記第１のプロトコルスタックまたは前記第２のプロトコルスタックを選択して前記送信機からのパケットを受信させるコントローラを含む
請求項１７に記載のシステム。

【請求項１９】
前記第１のインターフェースロジックは、前記トンネリング物理層を介して前記送信機から受信したパケットの通信種類に基づいてテーブルにアクセスして、前記通信種類に対応付けられている、タイミング遅延情報を取得し、前記タイミング遅延情報が示す遅延に対応するべく、前記通信種類について前記第１のプロトコルスタックの前記少なくとも第１のスタックロジックのタイミングを変更するべきか否かを決定する
請求項１６から１８のいずれか１項に記載のシステム。
【発明の詳細な説明】
【背景技術】
【０００１】
コンピュータプラットフォームは通常、さまざまなインターコネクトを用いて互いに結合された半導体素子を多数備えている。このようなインターコネクトまたはリンクは、プロトコルが異なっていることが多く、リンク上で実行される通信は、リンクが異なると、実行速度が異なり、準拠するプロトコルも異なる。一部のシステムでは、入出力（Ｉ／Ｏ）プロトコルの通信は、別のインターコネクトにおいてトンネリングされ得る。トンネリングは一般的に、第１のプロトコルに準拠した通信を受け取り、第２のプロトコルにしたがって動作するインターコネクトを介して当該通信を提供することを含む。例えば、第１のプロトコルのパケットに第２のプロトコルのヘッダを適用して、当該パケットをインターコネクトを介して送信することによって、第１のプロトコルのパケットがトンネリングされる。多くの場合、このようなプロトコルトンネリングは非常に上位のレベルで実行され、これら２つのプロトコルが有するソフトウェア抽象化が同じである一方、プロトコル間で共有されているハードウェアはない。このように、ソフトウェア互換性、性能、製品化までの時間において、上述したようなトンネリングによる利点は非常に小さい。
【図面の簡単な説明】
【０００２】
【図１】本発明の一実施形態に係る、共有物理層を介したリンクへのプロトコルスタックの接続を示すブロック図である。
【０００３】
【図２】本発明の別の実施形態に係る、共有物理層に結合されている複数の通信スタックを備えるシステムを示すブロック図である。
【０００４】
【図３】本発明の一実施形態に係る方法を説明するためのフローチャートである。
【０００５】
【図４】本発明の別の実施形態に係る、プロトコルスタックのインターフェースを動作させる方法を説明するためのフローチャートである。
【０００６】
【図５】本発明の一実施形態に係るシステムを示すブロック図である。
【発明を実施するための形態】
【０００７】
さまざまな実施形態によると、１以上の既存のＩＯプロトコルが、比較的低位のレベルで、別のインターコネクトにおいてトンネリングされ得る。本明細書では、「別のインターコネクト」を「トンネリングインターコネクト」と呼ぶことにする。一実施形態によると、このようなインターコネクトの一例として統合型（ｃｏｎｖｅｒｇｅｄ）ＩＯ（ＣＩＯ）が挙げられる。ＣＩＯは、ＰＣＩ Ｅｘｐｒｅｓｓ（登録商標）仕様書の基本仕様書バージョン２．０（２００７年１月１７日発行）に従った（以降では、ＰＣＩｅ（登録商標）仕様書と呼ぶ）、ＰＣＩｅ（Ｐｅｒｉｐｈｅｒａｌ Ｃｏｍｐｏｎｅｎｔ Ｉｎｔｅｒｃｏｎｎｅｃｔ Ｅｘｐｒｅｓｓ）プロトコルの通信、または、同様のプロトコルの通信をトンネリングするべく用いることができる。ＣＩＯの場合、ＰＣＩｅハードウェアスタックの多くは直接実装されるので、ソフトウェア互換性、性能、および製品化までの時間において、利点が得られる。すなわち、低位のレベルのトンネリングにおいて、トンネ

リングされるプロトコルスタックは大半が実装されている。これとは対照的に、高位のレベルのトンネリングの場合は、ソフトウェアアーキテクチャは保存されるが、トンネリングされるプロトコルからのパケット、符号化、または有線用プロトコルメカニズムを必ずしも用いるわけではない。このように低位のトンネリングによって、ＰＣＩｅプロトコルスタックのパケットを、例えば、トンネリングパケットにＣＩＯヘッダを適合させることによって、ＣＩＯインターコネクトを介してトンネリングすることができる。このように送信されたトンネリングパケットを受信機が受信すると、受信機のＣＩＯプロトコルスタックはヘッダを復号化して、受信機の対応ＰＣＩｅプロトコルスタックにＰＣＩｅパケットを渡すことができる。

【０００８】

しかし、統合型インターコネクトに対するこのようなアプローチによって、抽象化という、より高位で行われるプロトコルのトンネリングとは対照的に、低位トンネリングに起因する問題が生じる。つまり、プロトコルのタイミングについて制約が潜在的に課されていることが多い。このような制約は、インターコネクトプロトコルのトンネリングされていない元々のインスタンス化では満足させられていたことが自明であるが、インターコネクトプロトコルをトンネリングする場合には、トンネリングに用いられるインターコネクトによる遅延が発生するために、管理するのがより困難になり得る。このような遅延は、トンネリングインターコネクト自身によって生じる場合もあり、または、その他のトンネリングされたプロトコルによるトラフィックによって生じる場合もある。

【０００９】

本発明の実施形態は、トンネリングインターコネクトを介してトンネリングを実行する場合に、トンネリングされるプロトコルの明示的および暗示的なタイマを管理するメカニズムを提供する。本明細書で説明する実施形態ではＣＩＯを介してトンネリングされるＰＣＩｅプロトコルを一例として用いるが、本発明の範囲はこれに限定されるものではなく、同様の原理は、その他のトンネリングインターコネクト、および、有線インターコネクトおよび無線インターコネクトを始めとする、トンネリングに用いられるその他のインターコネクトにも応用され得ると理解されたい。

【００１０】

インターコネクトに関する、明示的および暗示的な、タイミング要件は、大きく分けて２つのカテゴリーに分類され得る。本明細書では、リンクタイミング要件および掛け時計タイミング要件と呼ぶ。リンクタイミング要件は、リンクプロトコル等の低位レベルに対応付けられており、リンク動作を円滑化して、検証に関する問題を最低限に抑えることを目的として通常設けられている。掛け時計タイミング要件は、高位レベルにおいて観察可能なイベント、例えば、オペレーティングシステム（ＯＳ）およびアプリケーションソフトウェアに対応付けられている。リンクタイミング要件は、プロトコルトンネリングによって生じる遅延に直接影響を受ける可能性があり、本発明の実施形態によって問題解決を図る要件である。通常、リンクタイミング要件は、約１０マイクロ秒（μｓ）未満のオーダーであり得るが、掛け時計タイミング要件は、約１０マイクロ秒（μｓ）よりも大きい。掛け時計タイミング要件は、プロトコルのトンネリングによって生じる比較的短時間（例えば、マイクロ秒）の遅延に影響されない程度に十分長い時間値（例えば、ミリ秒（ｍｓ））に対応付けられていることが多いので、プロトコルトンネリングには基本的に影響されない。また、このような要件は、特定のインターコネクトプロトコルを伝送するべく用いられるハードウェアメカニズムに関わらず（本来のもの、または、トンネリング用）等しく望ましい、アプリケーションソフトウェアにおいてユーザに見える動作停止（ｓｔａｌｌ）の発生を防ぐというような、特性に対応付けられている。

【００１１】

以下に記載する表１は、ＰＣＩｅに対応付けられている多数のタイミング要件を一覧にしたものであり、各要件と本開示との関係を示す。「説明」部分の記載は、ＰＣＩ　Ｅｘｐｒｅｓｓ（登録商標）仕様書から引用していることに留意されたい。

【表１】

説明	種類	注
肯定応答/非肯定応答(Ack/Nak) 送信および再生タイマ	リンク	絶対的な要件- トンネリングインターコネクトによって時間要件が満たされない場合、(スプリアス)エラーがトリガされる
リンク状態ゼロスタンバイ(L0s) 呼び出しポリシー： 「ポートは送信レーンを、所定のアイドル条件(以下参照)が7μs以下の期間にわたって満たされた場合には、L0s状態に遷移させなければならない	リンク	リンクが使用されていない場合には、リンク電力管理をトリガする。これは、実際にどれが利用されたかではなく、トンネリングインターコネクト割り当てに従って時間をカウントする場合の一例である。
リンク状態1(L1)エントリ交渉- 「上流の構成要素が、アイドル状態を4シンボル時間以下として、このDLLPを繰り返し送る」	リンク	暗示的なタイミング要件: この場合は、PCIeスタックに対してトラフィックがどのように見えるかを管理して、挿入された遅延をマスクする
フロー制御の更新	リンク	ガイドラインであって要件ではない
PCI電力管理(PM)およびアクティブ状態電力管理(ASPM)： 「L1を終了すると、下流の構成要素が、フロー制御更新データリンク層パケット、データリンク層パケット(DLLP)を、L1終了から1μs内に開始された全てのイネーブルされた仮想チャネル(VC)およびフロー制御(FC)型に対して、送信することが推奨される	リンク	ガイドラインであって要件ではない
L0s/L1終了レイテンシ	掛け時計	これらのタイミングパラメータは、PCIeを介したトラフィックに対する影響を決定できるように設けられたものであり、基本的なPCIeリンク動作ではない
電力管理イベント(PME)- 「100ms(+50%/-5%)後に、要求元エージェントのPME_Statusビットがクリアされていない場合、PMEサービスタイムアウトメカニズムが切れて、PME要求元エージェントをトリガして、一時的に失われたPM_PMEメッセージを再送させる」	掛け時計	PMEが完全に失われてしまわないように設けられている要件である； スプリアストリガを最低限に抑えるように特定の時間を選択すると同時に、PMEが比較的タイミング良く処理されるように十分に短く設定される
出された要求に対する受諾の制限(10us)	掛け時計	ファブリックの混雑によって生じるプラットフォームの観察可能な遅延を制限するためのもの
フロー制御の最低更新頻度(30us)および更新FCPタイマ-200us	掛け時計	フロー制御パケットの損失によって生じる停止を制限するためのもの

表１は、一部を例示することを目的とするものであって、ＰＣＩｅにおけるタイミング関連の要件を全て完全に網羅したものではないことに留意されたい。

【００１２】
リンクタイミング要件は、ＰＣＩｅスタック自身によって「測定」されるので、ＰＣＩｅスタックの時間に対する概念が変わると、このようなタイミング要件を感知する方法も変化し得る。このような構成を実現するべく、時間変更を実現するためのメカニズムが、スタックタイミングをいつ、そして、どのように変更するかを決定するためのハードウェア、ソフトウェアまたはファームウェアと共に、提供されるとしてよい。
【００１３】
ＰＣＩｅスタックの時間の概念を変更するメカニズムは、複数の方法で実現することができる。例えば、ＰＣＩｅスタックロジックにおけるさまざまな要素に対するクロックをゲーティングまたはオフに制御して、対象となるロジックの時間を止めることによって、実現することができる。尚、この方法によれば、利用されていないＰＣＩｅスタックロジックの電力消費を低減するという効果もさらに得られる。ほかの実施形態によると、時間をカウントすべき場合を指し示す明示的な制御信号を、ＰＣＩｅスタックロジックに追加することができる。スタック全体を１つの単位として制御することは一般的に十分ではなく、スタックのサブ構成要素は、プロトコルメカニズムが異なれば、異なるロジックブロックに対する影響も異なるので、半独立制御が実行され得る。同様に、リンクタイミング要件を遵守するべく、全ての通信についてタイミング変更が必要となるわけではない。一実施形態によると、制御ロジックを用いてＰＣＩｅスタックの時間の概念をいつおよびどのように調整するかを決定するとしてよく、当該ロジックはＰＣＩｅスタックの一部であってよい。
【００１４】
図１は、ＰＣＩｅスタック（およびその他のトンネリングされたプロトコル）が共有トンネリングリンクにインターフェースされる様子を示すブロック図である。一実施形態によると、共有トンネリングリンクは、ＣＩＯリンクであってよい。図１に示すように、システム１０は、第１のスタック２０ａと、第２のスタック２０ｂ（プロトコルスタック２０と総称する）とを備える。一実施形態によると、第１のプロトコルスタック２０ａは、ＰＣＩｅスタックであってよく、第２のプロトコルスタック２０ｂは、ユニバーサルシリアルバス（ＵＳＢ）、ディスプレイインターコネクト、またはその他の同様のプロトコルスタックであってよい。図示の便宜上、ＰＣＩｅプロトコルスタックの詳細のみを示す。具体的には、プロトコルスタック２０ａは、トランザクション層２２と、データリンク層２４と、インターフェースまたはガスケット層２６とを有する。インターフェースまたはガスケット層２６は、ＰＣＩｅプロトコルとトンネリングプロトコルとの間のインターフェースとして機能する。このようなインターフェースロジックの動作の詳細は後述する。
【００１５】
図１に示すように、統合型ＩＯ層は、第１および第２のプロトコルスタック２０と、リンク７０との間に結合されるとしてよい。リンク７０は、一実施形態によると、光リンク、電気リンク、またはその他の同様のリンクであってよい。図１に示すように、ＣＩＯプロトコルスタックは、ＣＩＯプロトコルトランスポート層３０と、物理層の論理ブロック４０と、物理層の電気ブロック５０と、物理層の光ブロック６０とを有するとしてよい。このように、ブロック４０から６０は、物理層と通信している複数のプロトコルによって共有されて、これらの複数のプロトコルの情報をリンク７０によってトンネリングする共通物理層として機能する。
【００１６】
図２は、共有物理層に結合されている複数の通信スタックを備えるシステムを示す図である。具体的には、図２では、ＰＣＩｅ送信（Ｔｘ）および受信（Ｒｘ）スタック２０ａに加えて、複数のその他の送信および受信スタック２０ｂ－２０ｄが設けられるとしてよい。同図に示すように、一対のマルチプレクサ３５ａおよび３５ｂ（マルチプレクサ３５と総称する）が、これらのスタックと共有物理層４０－６０との間に、結合されているとしてよい。マルチプレクサ３５は、プロトコルトランスポート層制御３０の制御下で、動作させられるとしてよい。図２に示すように、ＣＩＯプロトコルトランスポート（ＰＴ）

層３０は、ＰＣＩｅおよびその他のプロトコルをトンネリングするための、マルチプレクシングメカニズム（マルチプレクサ３５ａおよび３５ｂに基づく）および制御メカニズムを実装する。ＰＴ層制御３０は、送信機の抽象化と、送信機とは別個である受信機の操作を実装する。このような構造はこの説明の残りの部分でも用いるが、本発明の実施形態は、例えば、同時に送信機および受信機を抽象化することによって、または、単一の双方向接続を用いることによって、送信機および受信機を別々に制御するその他の種類のインターコネクトに応用され得る。

【００１７】

インターコネクトのタイミング制御を実装する方法は、実施形態によって異なるとしてよい。例えば、一部の実装によると、動的遅延バインディング（ｄｙｎａｍｉｃ　ｌａｔｅ　ｂｉｎｄｉｎｇ）が生じるとしてよく、その結果、そのようなインターフェースロジックが、結合されるべきトンネリングインターコネクトを動的に決定でき、トンネリングインターコネクトに対応するべくプロトコルのタイミング要件を動的に制御する。ほかの実施形態によると、設計者は、システム開発時において、１以上のプロトコルスタックによって使用されるべきトンネリングインターコネクトを決定するとしてよく、このため、当該トンネリングインターコネクトによって影響され得るリンクタイミング要件が、システム設計時において決定され得る。このため、トンネリングインターコネクトとプロトコルスタックとの間で、例えばインターフェースロジックにおいて、プロトコルスタックのタイミングを制御するロジックが組み込まれ得る。プロトコルスタックのタイミングの制御は、例えば、トンネリングインターコネクトによって発生したさらなる遅延に対応するべくプロトコルスタックのタイミングの概念を変更することによって、実現する。

【００１８】

図３は、リンクタイミング要件を処理するための上述した方法の実装例、つまり、プロトコルスタックが共通物理層または別の物理層に動的に結合され得るように、インターフェースロジックによって実装され得る動的遅延バインディングの実装例を示す図である。具体的に説明すると、図３は、例えば、プロトコルスタック（所与のプロトコルの標準スタックであってよい）と、さまざまなプロトコルのパケットをトンネリングすることができる統合型インターコネクトのような共通物理層との間の通信のためのプロトコルスタックのインターフェースロジックで実装され得る方法１００を説明するためのフローチャートである。図３に示すように、方法１００は、トンネリングインターコネクトのタイミング遅延情報を取得することによって開始されるとしてよい（ブロック１１０）。この情報を取得する方法としては、さまざまな方法が実装され得る。例えば、一実施形態によると、共有物理層が、インターフェースロジックに対して所定の遅延情報の一覧を提供するとしてよい。これに代えて、インターフェースロジックは、共通物理層との間で発生しているパケット通信を分析して、タイミング遅延情報を決定するとしてもよい。より一般的には、一部の実施形態では所定の方法でタイミング情報を取得して、その他の実装でこのタイミング情報を動的に算出するとしてよい。それぞれについていくつかバリエーションが考えられ、例えば、人間が予め決定する場合と機械が予め決定する場合とがあってよく、または、算出される場合には、一度チェックするとしてもよいし、もしくは、定期的にチェックを繰り返すとしてもよい。尚、このような情報については例としてさまざまなものが考えられ、通信の性質および関連するロジックの実体に応じて、通信の種類によって発生する遅延も異なる。

【００１９】

図３に戻って、ブロック１２０に進む。ブロック１２０では、タイミング遅延情報が、第１のプロトコルスタックのタイミング要件に対してマッピングされるとしてよい。一例を挙げると、プロトコルスタックのタイミング要件は、上記の表１において記載したように、リンク層通信によって異なるとしてよい。そして、ひし形１３０に進み、マッピングに基づいて、第１のプロトコルスタックのタイミングの概念を変更する必要があるか否かを決定するとしてよい。すなわち、共通物理層内に存在し得るレイテンシのために、当該プロトコルスタックの所定のロジックに対応付けられている１以上のタイマを、例えば、

加速、減速、ディセーブル等することによって、制御することができる。このようなタイミングの概念の変更が必要ない場合には、ブロック１３５に進み、プロトコルスタックの標準タイミングを用いてデータを送信および／または受信するとしてよい。

【００２０】

図３を参照しつつさらに説明すると、タイミングの概念を変更すべきであると決定される場合、ブロック１４０に進み、少なくとも１つのスタックロジックのタイミングを制御して、第１のプロトコルスタックのタイミングを変更するとしてよい。上述したように、このようなタイミングの変更は、タイマを制御すること、所定の間隔をカウントするように（またはしないように）ロジックを制御すること等に基づいて実現され得る。このようなタイミングの制御が実行された後、変更後のプロトコルスタックタイミングを用いて、所望のデータを送信／受信するとしてよい（ブロック１５０）。さらに図３に示すように、続いて、通信、つまり、所定のトランザクションが完了したか否かを判断するとしてよい（ひし形１６０）。完了していれば、当該方法を終了するとしてよい。完了していない場合は、ブロック１４０および１５０を繰り返すべく戻る。図３に図示する実施形態における具体的な実装に基づいて説明したが、本発明の範囲は、これに限定されない。

【００２１】

例えば、その他の実装例においては、所定のプロトコルスタックは、遅延が既知である既知のトンネリングインターコネクトを介してトンネリングされるように、システム設計が固定されるとしてもよい。したがって、システム設計時には、当該トンネリングインターコネクトに関して発生する可能性のあるいずれの遅延についても対応するべく、必要に応じて、さまざまなプロトコルトランザクションのタイミングの制御を実行するロジックが実装され得る。上記の表１は、リンク層タイミング要件の例を示している。

【００２２】

図４は、本発明の別の実施形態に係るプロトコルスタックのインターフェースを動作させる方法を示すフローチャートである。図４に示すように、方法２００は、静的設計パラメータに基づいて、必要に応じて、プロトコルスタックのタイミングの概念を変更することができるインターフェースロジックによって実装され得る。図４に示すように、方法２００は、トンネリングインターコネクトに対する通信またはトンネリングインターコネクトからの通信を受信することによって開始され得る（ブロック２０５）。このため、当該通信は、プロトコルスタックのインターフェースロジックにおいて、出力方向または入力方向で、受信される。インターフェースロジックではさまざまな種類の通信が処理されるとしてよい。例えば、肯定的応答（ＡＣＫ）等のさまざまなプロトコルパケット、電力管理、フロー制御等に用いられる制御パケット、データパケットの送信および受信であってよい。

【００２３】

パケットの種類に基づいて、所定の通信の種類が、変更後タイミングに影響されるか否かを、インターフェースロジックにおいて決定するとしてよい（ひし形２１０）。例えば、インターフェースロジックは、所定の表を含むとしてもよいし、所定の表に対応付けられているとしてもよい（不揮発性メモリに存在するとしてよい）。この表は、トランザクションの種類と、その種類の通信についてプロトコルスタックのタイミングの概念を変更すべきか否かとを特定しており、さらに、適用可能な遅延の指定、タイミングを変更するべくインターフェースロジックが適用すべき制御方策の種類を示す命令またはその他の識別子を記載する。尚、この表は複数の部分に分割されているとしてよく、それぞれの部分は、所定のスタックに対応付けられており、それぞれの部分が、あるスタック－トンネリングインターコネクト関係専用のマッピングを提供している。

【００２４】

さらに図４を参照しつつ説明すると、変更が必要ない場合、標準のプロトコルスタックタイミングを用いて通信を処理するとしてよく、標準のプロトコルスタックタイミングを用いてデータを送信／受信してよい（ブロック２２０）。タイミングの概念を変更すべきであると決定される場合、ブロック２３０に進み、少なくとも１つのスタックロジックの

タイミングを制御して、タイミングを変更するとしてよい。そして、変更後のプロトコルスタックタイミングを用いて所望のデータを送信／受信するとしてよい（ブロック２４０）。さらに図４に示すように、通信、つまり、所定のトランザクションが完了したか否かを決定するとしてよい（ひし形２６０）。完了している場合、当該方法を終了するとしてよい。完了していない場合、ブロック２３０および２４０を繰り返すべく戻る。リンクタイミング要件を処理する静的制御はこのように実現され得る。

【００２５】

上記の図３および図４に示すように、所定の通信の種類についてタイミング制御を変更することができる一方、その他の通信の種類については、変更なしで、通常のプロトコルスタックタイミングに従って処理が進められるとしてよい。以下の説明では、リンクタイミング要件に対応するべく、プロトコルスタックのタイミングを変更し得る場合の例をいくつか紹介する。

【００２６】

一実施形態によると、ＰＴ層制御３０は、ＰＣＩｅに割り当てられる送信機「スロット」を提供し得る。尚、この送信機「スロット」は、送信すべきＰＣＩｅトラフィックが存在しない場合、その他の種類のトラフィックにも利用可能である。このため、第１のプロトコルスタックに割り当てられたスロットは、第１のプロトコルスタックが何も送信しない場合には、別のプロトコルスタックによって用いられるとしてもよい。同様に、受信機においても、ＰＣＩｅトラフィックが受信されるはずだが、ほかの構成要素が送信すべきＰＣＩｅトラフィックを有していないか、または、別の種類のより優先度が高いトラフィックを有しているかという理由で、受信機がその期間中ＰＣＩｅトラフィックを一切受信しないという期間があり得る。

【００２７】

「ＰＣＩｅ期間」という考え方をＰＣＩｅスタックに正確に伝える上で、受信時刻および送信時刻はある程度別個に考えることができる。表１に記載した、Ｌ０ｓ呼び出しポリシーおよび「Ｌ１が終了すると・・・」という要件のような、一部の場合によると、時刻が測定されるのは、１つの観点のみからである（このような場合は、送信機の観点である）。

【００２８】

しかし、Ａｃｋ／Ｎａｋプロトコルの場合は、受信機および送信機の観点が両方とも考慮される必要がある。物理ＰＣＩｅポートに基づいて送信パイプラインを伝播する際に特定のレイテンシを仮定して決定された、トランザクション層パケット（ＴＬＰ）が送信された、ＰＣＩｅ送信機の観点での時間は、ＣＩＯ送信パイプラインの遅延が異なっている場合には、不正確である場合がある。ほかの構成要素（つまり、受信機）は、自身のＰＣＩｅスタックが共有リンク上においてＰＣＩｅ時間を割り当てられた場合にのみ応答が可能である。この割り当ては、自身の（受信機の）時間の概念を調整する必要がある旨を意味するものとして、受信機には感知される。ＰＣＩｅスタックは、送信パイプラインの遅延を５０ナノ秒（ｎｓ）と予測しているが、ＣＩＯリンクの送信パイプライン遅延は７０ｎｓだと仮定する。この場合、差分に対処するべく、２０ｎｓだけ、送信機の時間の概念を（この遅延を認識しているかどうかに応じて決まるプロトコルの側面について）停止または調整する必要がある。このようにすることで、送信機は、受信機（共有物理層によって遅延され得る）からのＡＣＫ信号について適切な長さの時間だけ待機して、不適切にエラー信号が発行されないようにする。

【００２９】

受信機の場合は、ほかの構成要素の送信機がＰＣＩｅに対して割り当てた時間（利用した時間ではない）を考慮しなければならない。一部のケースでは、受信機に直接通知されるが、その他のケースでは、ほかの構成要素の受信機が時間の概念をどれだけ進めるべきかを指し示すメッセージ等のトンネリングプロトコルメカニズムが、トンネリングされたプロトコルのそれぞれについて、設けられるとしてよい。例えば、ＰＣＩｅ送信機に２つの１００ｎｓスロットが割り当てられるが、送信すべきＰＣＩｅトラフィックが少ないた

めに、そのうち１つのみが送信機によって利用される場合、受信機は２００ｎｓを考慮しなければならない。このように、ほかの構成要素が送信に利用可能なスロットを利用せずに、タイミングルールを守らない場合、ルールを守らなかったことは受信機において認識可能である。これは、（割り当てられたものではなく）使用された送信スロットのみを考慮する場合には、この限りでない。

【００３０】

尚、所定のプロトコルについては多岐にわたる最適化が可能であるとしてよい。例えば、既知の帯域幅トラフィックは、実際に与えられているリンク抽象化には関係なく、カウンタメカニズムを用いて考慮するとしてよい。プロトコルの受信割り当ておよび送信割り当てが等しいことが保証されている場合、片方のみ（例えば、送信機）が、他方（受信機）の時間の概念が一致すべきであると理解していると考えてよい。

【００３１】

先述したように、本発明の実施形態は、ＣＩＯまたはＰＣＩ等の具体的な内容に左右されるものでは一切なく、ディスプレイ、ＵＳＢ、ネットワーク等のトンネリングされるその他のプロトコルに適用され得る。本発明の実施形態はさらに、その他のトンネリングプロトコル／環境、例えば、有線または無線ＵＳＢインターコネクトを介したトンネリングＰＣＩｅにも適用される。

【００３２】

本発明の一実施形態に応じてトンネリングを実行することによって、より汎用度が高いハードウェアから成る一般的なハードウェア群によって、より多くの個別のＩＯアプリケーションで要件を満たすことができる。例えば、プラットフォームは１２個のＵＳＢポート、８個のＰＣＩｅポート、およびさまざまな特定用途向けのポート（例えば、ディスプレイ）を備えるとしてよい。トンネリングによって、これらのポートは、例えば１６個の複合型ポートから成るポート群にまとめることができ、各ポートは、従前のポートのいずれ（１または複数）としても利用することができる。

【００３３】

本発明の実施形態は、多くの異なる種類のシステムで実装することができる。図５は、シリアルリンクであるトンネリングインターコネクトによってコントローラハブに結合された複数のデバイスを備える、本発明の一実施形態に係るシステムを示すブロック図である。システム３００は、コントローラハブ３１５に結合されている、プロセッサ３０５と、システムメモリ３１０とを備える。プロセッサ３０５は、マイクロプロセッサ、ホストプロセッサ、埋め込み型プロセッサ、コプロセッサ等のプロセッサを含む任意の処理素子を含む。プロセッサ３０５は、フロントサイドバス（ＦＳＢ）３０６を介して、コントローラハブ３１５に結合されている。一実施形態によると、ＦＳＢ３０６は、シリアル方式のポイントツーポイント（ＰｔＰ）インターコネクトである。

【００３４】

システムメモリ３１０は、ランダムアクセスメモリ（ＲＡＭ）、不揮発性（ＮＶ）メモリ等の、システム３００が備えるデバイスがアクセス可能なメモリなど、任意のメモリデバイスを含む。システムメモリ３１０は、メモリインターフェース３１６を介して、コントローラハブ３１５に結合される。

【００３５】

一実施形態によると、コントローラハブ３１５は、ＰＣＩｅインターコネクト階層構造において、ルートハブまたはルートコントローラである。コントローラハブ３１５の例を挙げると、チップセット、メモリコントローラハブ（ＭＣＨ）、ノースブリッジ、入出力コントローラハブ（ＩＣＨ）、サウスブリッジ、およびルートコントローラ／ハブがある。尚、コントローラハブ３１５は、シリアルリンク３１９を介して、スイッチ／ブリッジ３２０に結合される。入出力モジュール３１７および３２１は、インターフェース／ポート３１７および３２１と呼ばれることもあるが、コントローラハブ３１５とスイッチ３２０との間の通信を実現するべく層状プロトコルスタックを含む／実装する。一実施形態によると、スイッチ３２０には複数のデバイスを結合することができる。

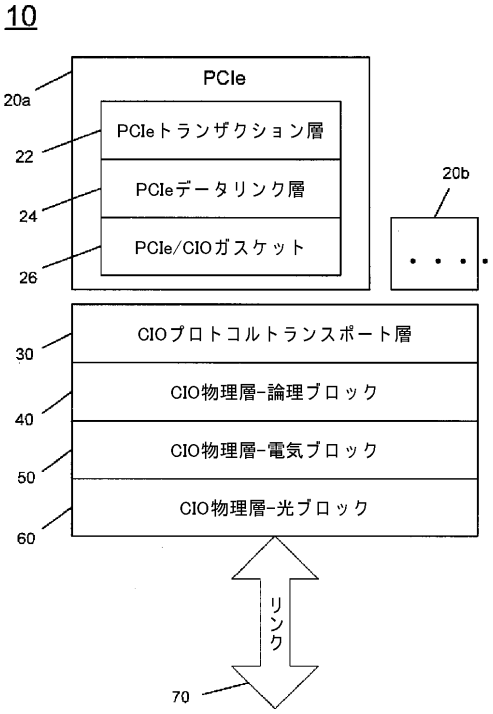
【００３６】
スイッチ３２０は、デバイス３２５からのパケット／メッセージを、上流、つまり、階層構造を上方向に、コントローラハブ３１５に対してルーティングすると共に、下流、つまり、階層構造を下方向に、コントローラハブ３１５から離れるように、デバイス３２５へとパケット／メッセージをルーティングする。ＩＯモジュール３２２および３２６は、スイッチ３２０とデバイス３２５との間の通信を実行するべく、層状プロトコルスタックを実装する。一実施形態によると、ＩＯモジュール３２６は、複数のプロトコルスタック、つまり、スタック３２７および３２８のパケットをトンネリングするためのトンネリング物理層であってよい。デバイス３２５は、電子システムに結合される、任意の内部または外部のデバイスまたは構成要素を含む。例えば、ＩＯデバイス、ネットワークインターフェースコントローラ（ＮＩＣ）、拡張カード、オーディオプロセッサ、ネットワークプロセッサ、ハードドライブ、ストレージデバイス、モニタ、プリンタ、マウス、キーボード、ルータ、ポータブルストレージデバイス、ファイヤワイヤデバイス、ユニバーサルシリアルバス（ＵＳＢ）デバイス、スキャナー、およびその他の入出力デバイスが挙げられる。

【００３７】
コントローラハブ３１５にはさらに、シリアルリンク３３２を介して、グラフィクスアクセラレータ３３０が接続されている。一実施形態によると、グラフィクスアクセラレータ３３０は、ＭＣＨに結合され、ＭＣＨはＩＣＨに結合されている。スイッチ３２０は、そしてＩＯデバイス３２５は、ＩＣＨに結合されている。ＩＯモジュール３３１および３１８もまた、グラフィクスアクセラレータ３３０とコントローラハブ３１５との間において通信を実行するべく、層状プロトコルスタックを実装する。

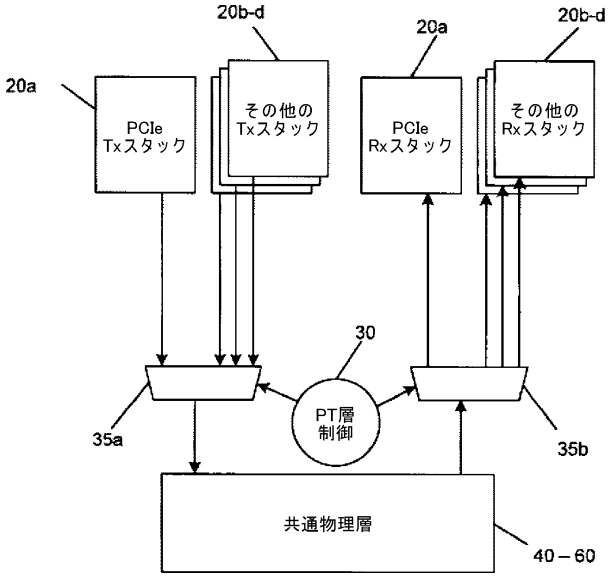
【００３８】
本発明の実施形態は、符号で実装されるとしてよく、命令を格納しており、当該命令を実行するようにシステムをプログラミングするべく用いられる格納媒体に格納されるとしてよい。このような格納媒体は、これらに限定されるものではないが、フロッピーディスク（登録商標）、光ディスク、コンパクトディスクリードオンリーメモリ（ＣＤ－ＲＯＭ）、書き換え可能コンパクトディスク（ＣＤ－ＲＷ）、および光磁気ディスク等の任意の種類のディスク、リードオンリーメモリ（ＲＯＭ）、ダイナミックランダムアクセスメモリ（ＤＲＡＭ）およびスタティックランダムアクセスメモリ（ＳＲＡＭ）のようなランダムアクセスメモリ（ＲＡＭ）、消去可能なプログラム可能リードオンリーメモリ（ＥＰＲＯＭ（登録商標））、フラッシュメモリ、電気的に消去可能なプログラム可能リードオンリーメモリ（ＥＥＰＲＯＭ（登録商標））、磁気カードあるいは光カード等の半導体デバイス、または、電子的に命令を格納するのに適しているその他の任意の種類の媒体を含むとしてよい。

【００３９】
限られた数の実施形態に基づいて本発明を説明してきたが、当業者であれば、多くの変形および変更に想到するであろう。本願の請求項はこのような変形および変更を全て、本発明の真の精神および範囲内にあるものとして、含むものである。

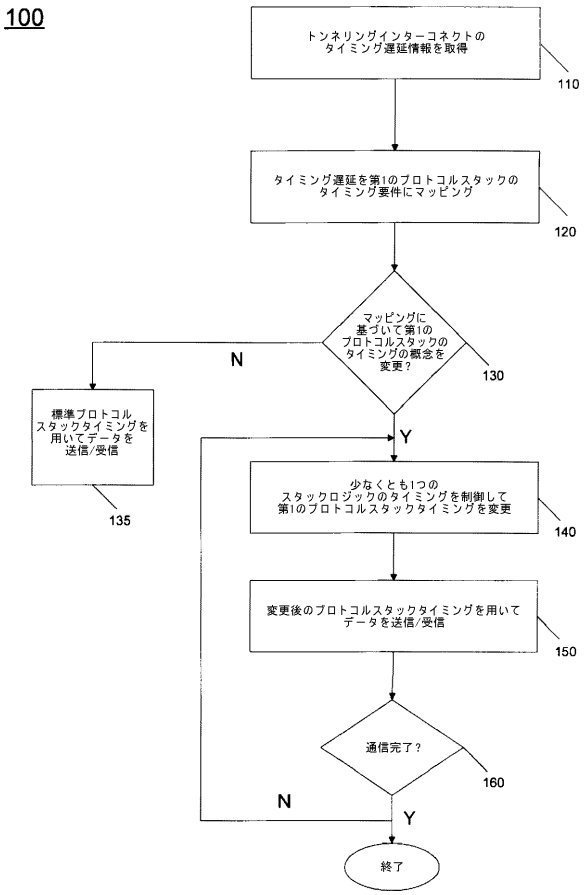
【図１】


10
PCIe
20a
PCIeトランザクション層
22
PCIeデータリンク層
24
PCIe/CIOガスケット
26
20b
CIOプロトコルトランスポート層
30
CIO物理層-論理ブロック
40
CIO物理層-電気ブロック
50
CIO物理層-光ブロック
60
リンク
70

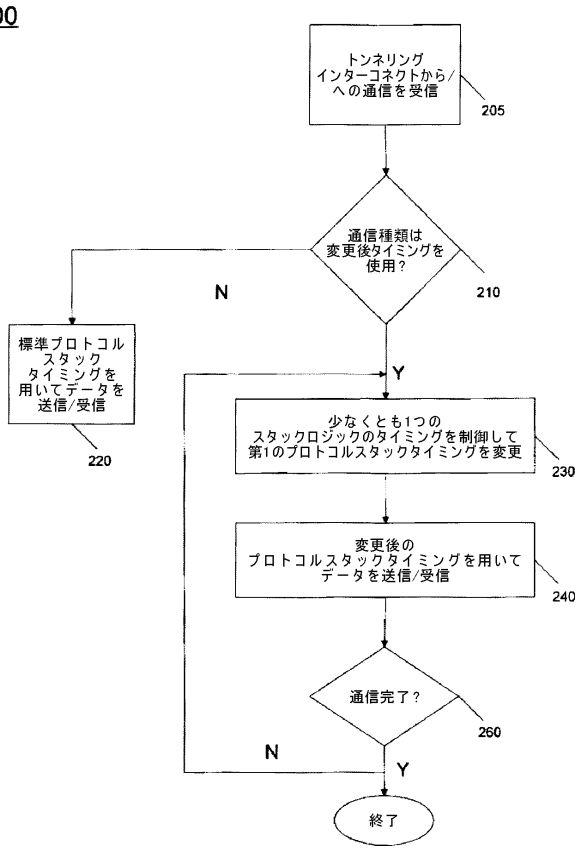
【図２】


20b-d
20a
20b-d
20a
PCIe Txスタック
その他の Txスタック
PCIe Rxスタック
その他の Rxスタック
30
PT層 制御
35a
35b
共通物理層
40－60

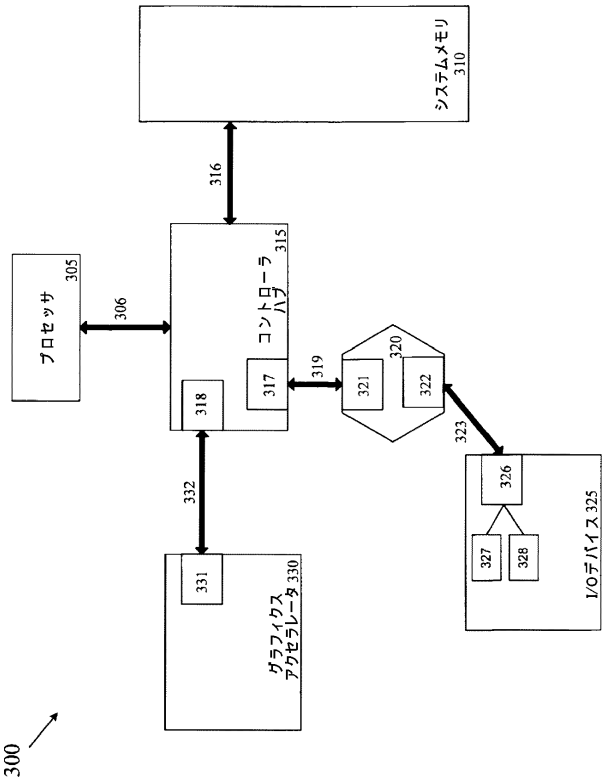
【図３】


100
トンネリングインターコネクトの タイミング遅延情報を取得
110
タイミング遅延を第1のプロトコルスタックの タイミング要件にマッピング
120
マッピングに 基づいて第1の プロトコルスタックの タイミングの概念を 変更？
130
N
Y
標準プロトコル スタックタイミングを 用いてデータを 送信/受信
135
少なくとも1つの スタックロジックのタイミングを制御して 第1のプロトコルスタックタイミングを変更
140
変更後のプロトコルスタックタイミングを用いて データを送信/受信
150
通信完了？
160
N
Y
終了

【図４】


200
トンネリング インターコネクトから/ への通信を受信
205
通信種類は 変更後タイミングを 使用？
210
N
Y
標準プロトコル スタック タイミングを 用いてデータを 送信/受信
220
少なくとも1つの スタックロジックのタイミングを制御して 第1のプロトコルスタックタイミングを変更
230
変更後の プロトコルスタックタイミングを用いて データを送信/受信
240
通信完了？
260
N
Y
終了

【図５】


システムメモリ
310
316
プロセッサ
305
306
コントローラ
ハブ
315
317
318
319
320
321
322
323
326
327
328
I/Oデバイス 325
332
331
グラフィクス
アクセラレータ 330
300

フロントページの続き

(56)参考文献　国際公開第２００５／１１７３５２（WO，Ａ１）
特開２００２－１３５２６４（ＪＰ，Ａ）
特開２００３－２７３９５３（ＪＰ，Ａ）
特開２０００－２５３０７１（ＪＰ，Ａ）
特表２０１０－５００８０７（ＪＰ，Ａ）

(58)調査した分野(Int.Cl.，ＤＢ名)
Ｈ０４Ｌ　２９／１０
Ｇ０６Ｆ　１３／３８
Ｇ０６Ｆ　１３／４２
Ｈ０４Ｌ　１２／５６