(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7098711号
(P7098711)

(45)発行日　令和4年7月11日(2022.7.11)

(24)登録日　令和4年7月1日(2022.7.1)

(51)国際特許分類

		FI	
H04L	*45/021 (2022.01)*	H04L	45/021
H04L	*67/00 (2022.01)*	H04L	67/00
H04L	*69/00 (2022.01)*	H04L	69/00

請求項の数　10　（全20頁）

(21)出願番号　特願2020-500079(P2020-500079)
(86)(22)出願日　平成30年7月4日(2018.7.4)
(65)公表番号　特表2020-526968(P2020-526968A)
(43)公表日　令和2年8月31日(2020.8.31)
(86)国際出願番号　PCT/IB2018/054942
(87)国際公開番号　WO2019/012375
(87)国際公開日　平成31年1月17日(2019.1.17)
審査請求日　令和2年12月22日(2020.12.22)
(31)優先権主張番号　15/647,561
(32)優先日　平成29年7月12日(2017.7.12)
(33)優先権主張国・地域又は機関
米国(US)
(31)優先権主張番号　15/840,600
(32)優先日　平成29年12月13日(2017.12.13)
最終頁に続く

(73)特許権者　390009531
インターナショナル・ビジネス・マシーンズ・コーポレーション
ＩＮＴＥＲＮＡＴＩＯＮＡＬ　ＢＵＳＩＮＥＳＳ　ＭＡＣＨＩＮＥＳ　ＣＯＲＰＯＲＡＴＩＯＮ
アメリカ合衆国１０５０４　ニューヨーク州　アーモンク　ニュー　オーチャード　ロード
New Orchard Road, Armonk, New York 10504, United States of America
(74)代理人　100112690
弁理士　太佐 種一
最終頁に続く

(54)【発明の名称】　遠隔ノード発見、ならびに通信チャネル確認および通信チャネル接続のためのプロセッサによって実施される方法、コンピュータ・システム、およびコンピュータ・プログラム

(57)【特許請求の範囲】
【請求項１】
通信チャネル実施のための、遠隔ノード発見、および複数のキュー・ペアを結び付けることのためのプロセッサによって実施される方法であって、
複数のノードの間で１つまたは複数のポイント・ツー・ポイント・リンクを接続すべく複数のハードウェアを準備するステップと、
発見プロセスを通じて前記１つまたは複数のポイント・ツー・ポイント・リンク内で前記複数のノードを発見するステップと、
前記複数のノード間に複数の管理用トラフィック・レーンを作成するステップであって、前記複数の管理用トラフィック・レーンは、チャネルに依存しない複数の開始プロトコルを利用する、ステップと、
前記複数のノード間の複数の接続を検証するステップと、
前記複数の接続上で１つまたは複数の補助的トラフィック・フロー・インターフェースを開始するステップと、
前記複数の接続のための複数のバッファをネゴシエートし、かつ接続すること、および
前記複数のノードの間で複数のノード記述子を交換するステップ
を含む方法。
【請求項２】
デバイス・ドライバ層内で実行される、請求項１に記載の方法。
【請求項３】

複数のノードの間で１つまたは複数のポイント・ツー・ポイント・リンクを接続すべく複数のハードウェアを準備するステップが、
前記複数のハードウェアから複数のバッファ・リソースを前もって確保するステップをさらに含む、請求項１に記載の方法。
【請求項４】
前記複数のハードウェアから前記複数のバッファ・リソースを前もって前記確保するステップが、単一ルート入力／出力仮想化インターフェースを使用して行われる、請求項３に記載の方法。
【請求項５】
前記複数のハードウェアからの複数のバッファ・リソースの確保が、１回、行われる、請求項１に記載の方法。
【請求項６】
チャネル初期設定アービタ機能を利用して、前記複数の接続内の複数のチャネルのすべてが初期設定されていることを確実にすることをさらに含む、請求項１に記載の方法。
【請求項７】
前記複数のバッファをネゴシエートし、かつ接続することが、４ステップのキュー・ペア接続プロセスを含む、請求項１に記載の方法。
【請求項８】
請求項１～７の何れか１項に記載の方法の各ステップをコンピュータ・ハードウェアによる手段として構成した、コンピュータ・システム。
【請求項９】
請求項１～７の何れか１項に記載の方法の各ステップをコンピュータに実行させる、コンピュータ・プログラム。
【請求項１０】
請求項９に記載の前記コンピュータ・プログラムをコンピュータ可読記憶媒体に記録した、記憶媒体。
【発明の詳細な説明】
【技術分野】
【０００１】
本発明は、一般に、コンピューティングの分野に関し、より詳細には、コンピュータ・ネットワーキングの分野に関する。
【背景技術】
【０００２】
コンピュータ・ネットワーキングは、コンピューティング・デバイスがデータを交換することを可能にするデータ・リンクを有する２つ以上のコンピューティング・デバイスの接続を含む。コンピュータ・ネットワークは、多くのなかでもとりわけ、インターネットに対するアクセス、クラウド・ストレージ、プリンタ、ファックス・マシン、電子メール、インスタント・メッセージング、電話サービス、およびデジタル・メディアなどの膨大な数のアプリケーションおよびサービスの基礎の役割をする。ラップトップおよびセル電話などのパーソナル・コンピューティング・デバイスの広範な普及とともに、コンピュータ・ネットワークは、いたるところに存在し、かつ現代の生活の極めて重要な要素となっている。このため、コンピュータ・ネットワーキング技術の維持および改良の進歩は、現代の社会に行き渡る利益をもたらす潜在力を有する。
【発明の概要】
【発明が解決しようとする課題】
【０００３】
したがって発明が解決しようとする課題は、遠隔ノード発見、ならびに通信チャネル確認および通信チャネル接続のための方法、システム、およびそのプログラムを提供することである。
【課題を解決するための手段】

【０００４】
一態様によれば、通信チャネル実施のための、遠隔ノード発見、および複数のキュー・ペアを結び付けることのためのプロセッサによって実施される方法が提供され、方法は、複数のノードの間で１つまたは複数のポイント・ツー・ポイント・リンクを接続すべく複数のハードウェアを準備すること、発見プロセスを通じてその１つまたは複数のポイント・ツー・ポイント・リンク内でその複数のノードを発見すること、その複数のノード間に複数の管理トラフィック・レーンを作成すること、その複数のノード間の複数の接続を検証すること、その複数の接続上で１つまたは複数の補助的トラフィック・フロー・インターフェースを開始すること、その複数の接続のための複数のバッファをネゴシエートし、かつ接続すること、およびその複数のノードの間で複数のノード記述子を交換することを含む。

【０００５】
別の態様によれば、通信チャネル実施のための、遠隔ノード発見、および複数のキュー・ペアを結び付けることのためのコンピュータ・システムが提供され、コンピュータ・システムは、１つまたは複数のプロセッサと、１つまたは複数のコンピュータ可読メモリと、１つまたは複数のコンピュータ可読有形記憶媒体と、その１つまたは複数のメモリのうちの少なくとも１つを介してその１つまたは複数のプロセッサのうちの少なくとも１つによって実行されるようにその１つまたは複数の有形記憶媒体のうちの少なくとも１つの上に記憶されたプログラム命令とを含み、コンピュータ・システムは、複数のノードの間で１つまたは複数のポイント・ツー・ポイント・リンクを接続すべく複数のハードウェアを準備すること、発見プロセスを通じてその１つまたは複数のポイント・ツー・ポイント・リンク内でその複数のノードを発見すること、その複数のノード間に複数の管理トラフィック・レーンを作成すること、その複数のノード間の複数の接続を検証すること、その複数の接続上で１つまたは複数の補助的トラフィック・フロー・インターフェースを開始すること、その複数の接続のための複数のバッファをネゴシエートし、かつ接続すること、およびその複数のノードの間で複数のノード記述子を交換することを含む方法を実行することができる。

【０００６】
別の態様によれば、通信チャネル実施のための、遠隔ノード発見、および複数のキュー・ペアを結び付けることのためのコンピュータ・プログラム製品が提供され、コンピュータ・プログラム製品は、１つまたは複数のコンピュータ可読有形記憶媒体と、その１つまたは複数のコンピュータ可読有形記憶媒体のうちの少なくとも１つの上に記憶されたプログラム命令とを含み、それらのプログラム命令は、プロセッサによって実行可能であり、それらのプログラム命令は、複数のノードの間で１つまたは複数のポイント・ツー・ポイント・リンクを接続すべく複数のハードウェアを準備すること、発見プロセスを通じてその１つまたは複数のポイント・ツー・ポイント・リンク内でその複数のノードを発見すること、その複数のノード間に複数の管理トラフィック・レーンを作成すること、その複数のノード間の複数の接続を検証すること、その複数の接続上で１つまたは複数の補助的トラフィック・フロー・インターフェースを開始すること、その複数の接続のための複数のバッファをネゴシエートし、かつ接続すること、およびその複数のノードの間で複数のノード記述子を交換することを含む方法を実行するようにプロセッサに指示する。

【０００７】
一実施形態によれば、コンピュータ・ネットワーキングのための方法、コンピュータ・システム、およびコンピュータ・プログラム製品が提供される。本発明は、システム間でポイント・ツー・ポイント・リンクを接続すべくハードウェアを準備すること、発見プロセスを通じてそれらのリンクの他方の側の遠隔システムを発見すること、管理トラフィック・レーンを作成すること、ノード間の接続を検証すること、補助的トラフィック・フロー・インターフェースを開始すること、バッファをネゴシエートし、かつ接続すること、およびノード記述子を交換することを含む、１つまたは複数の通信チャネルを実施する際に、遠隔ノードアドレス指定可能性を発見すべく媒体アクセス制御（ＭＡＣ）マルチキャス

ト手続きを使用し、かつキュー・ペアのセットを接続すべくＩｎｆｉｎｉＢａｎｄ（Ｒ）（ＩｎｆｉｎｉＢａｎｄ（Ｒ）およびすべてのＩｎｆｉｎｉＢａｎｄ（Ｒ）ベースの商標およびロゴは、システムＩ／Ｏ株式会社、またはシステムＩ／Ｏ株式会社の関連会社、あるいはその両方の商標または登録商標である）プロトコルを使用する方法を含んでよい。

【０００８】

次に、本発明の好ましい実施形態が、単に例として、以下の図面を参照して説明される。

【図面の簡単な説明】

【０００９】

【図１】少なくとも１つの実施形態による例示的なネットワーク化されたコンピュータ環境を示す図である。

【図２】少なくとも１つの実施形態によるノード発見およびアービトレーション・プロセスを示す動作フローチャートである。

【図３】少なくとも１つの実施形態によるリンク初期設定手続きを示す機能ブロック図である。

【図４】少なくとも１つの実施形態による図１に示されるコンピュータおよびサーバの内部構成要素および外部構成要素を示すブロック図である。

【図５】本発明の実施形態によるクラウド・コンピューティング環境を示す図である。

【図６】本発明の実施形態による抽象モデル層を示す図である。

【発明を実施するための形態】

【００１０】

図面の様々な特徴は、これらの例示が、当業者が、詳細な説明と併せて本発明の好ましい実施形態を理解するのを容易にする際の簡明さを目的とするので、一律の縮尺に従ってはいない。

【００１１】

特許請求される構造および方法の詳細な実施形態が本明細書において開示されるが、開示される実施形態は、様々な形態において実施されてよい特許請求される構造および方法を例示するに過ぎないものと理解されることが可能である。しかし、本発明は、異なる多くの形態で実施されてよく、本明細書に記載される例示的な実施形態に限定されるものと解釈されるべきではない。この説明において、よく知られた特徴および技法の詳細は、提示される実施形態を不必要に不明瞭にするのを回避すべく省略されることが可能である。

【００１２】

本発明の実施形態は、コンピューティングの分野に関し、より詳細には、コンピュータ・ネットワーキングに関する。説明される以下の例示的な実施形態は、１つまたは複数の通信チャネルを実施する際に、とりわけ、遠隔ノードアドレス指定可能性を発見すべく媒体アクセス制御（ＭＡＣ）マルチキャスト手続きを使用し、かつキュー・ペアを接続すべくＩｎｆｉｎｉＢａｎｄ（Ｒ）（ＩｎｆｉｎｉＢａｎｄ（Ｒ）およびすべてのＩｎｆｉｎｉＢａｎｄ（Ｒ）ベースの商標およびロゴは、システムＩ／Ｏ株式会社、またはシステムＩ／Ｏ株式会社の関連会社、あるいはその両方の商標または登録商標である）プロトコルを使用するシステム、方法、およびプログラム製品を提供する。したがって、本実施形態は、切り換えられる環境を許容するノード発見の帯域内方法を提供することによってコンピュータ・ネットワーキングの技術分野を改良する能力を有する。

【００１３】

前述したとおり、コンピュータ・ネットワーキングは、コンピューティング・デバイスがデータを交換することを可能にするデータ・リンクを有する２つ以上のコンピューティング・デバイスの接続を含む。コンピュータ・ネットワークは、多くのなかでもとりわけ、インターネットに対するアクセス、クラウド・ストレージ、プリンタ、ファックス・マシン、電子メール、インスタント・メッセージング、電話サービス、およびデジタル・メディアなどの膨大な数のアプリケーションおよびサービスの基礎の役割をする。ラップトップおよびセル電話などのパーソナル・コンピューティング・デバイスの広範な普及とともに、コンピュータ・ネットワークは、いたるところに存在し、かつ現代の生活の極めて重

要な要素となっている。このため、コンピュータ・ネットワーキング技術の維持および改良の進歩は、現代の社会に行き渡る利益をもたらす潜在力を有する。

【００１４】

多くのネットワーク接続は、市販のオペレーティング・システムに組み込まれたプロトコルの標準のセットを利用する。これらのプロトコルは、汎用であり、ネットワーク接続の普及のための基礎である。しかし、そのような汎用性は、ネットワーク・パフォーマンスを犠牲にしている。ネットワーク潜時などの測定基準に左右されるシステムは、汎用ではなく、特定のアプリケーションおよびシステムの動作環境のためにカスタマイズされたカスタム通信プロトコルを開発することを選択することが可能である。したがって、集中型サーバの必要なしに、ポイント・ツー・ポイント・リンクにおけるノード発見を使用し、かつ動的プロトコルを使用して作成されたチャネルの使用によってではなく、初期の発見を介して確立された固定のチャネルに基づいてデータを転送する方法の必要性が存在する。

【００１５】

一実施形態によれば、マルチキャスト・プロトコルを使用して遠隔アドレス指定可能性を帯域内で発見し、次に、遠隔チャネル定義が確認され、かつ初期設定された直接に接続された管理経路を確立することによって、２つのシステムの間で１つまたは複数のチャネルのセットが確立される。

【００１６】

そのような実施形態において、ＩｎｆｉｎｉＢａｎｄ（Ｒ）（ＩｎｆｉｎｉＢａｎｄ（Ｒ）およびすべてのＩｎｆｉｎｉＢａｎｄ（Ｒ）ベースの商標およびロゴは、システムＩ／Ｏ株式会社、またはシステムＩ／Ｏ株式会社の関連会社、あるいはその両方の商標または登録商標である）などのトランスポート層プロトコルをサポートする通信アダプタを使用して、イーサネット（Ｒ）などの物理層プロトコルを使用して２つのサーバ・システムの間で物理接続がまず、確立される。その接続上で、遠隔システム上の接続されたポートのアドレスを包含するマルチキャスト・メッセージを受信し、かつローカル・システム上の接続されたポートのアドレスを包含するマルチキャスト・メッセージを送信するローカル・マルチキャスト・キュー・ペア（ＱＰ）をセット・アップすることによって、遠隔システムのアドレス指定可能性が獲得される。

【００１７】

遠隔ポート・アドレスが獲得されたとき、そのアドレスは、遠隔システムを相手に直接に接続された管理ＱＰを確立するパラメータとして使用される。管理ＱＰは、チャネルのローカル定義および遠隔定義などの情報が送られ、かつチャネルを含むＱＰの数などのパラメータがネゴシエートされる経路として使用される。

【００１８】

合致するローカルと遠隔の定義を有する各チャネルに関して、管理ＱＰ経路は、補助的ＱＰ、および構成バッファＱＰのセットを定義し、かつ接続するのに使用される。補助的ＱＰは、チャネル特有の非データ・メッセージ（メッセージ中止コマンドなどの）のために使用され、バッファＱＰは、チャネルに関するデータ・メッセージを送受信するのに使用される。

【００１９】

本発明は、統合の任意の可能な技術的詳細レベルにおけるシステム、方法、またはコンピュータ・プログラム製品、あるいはその組合せであってよい。コンピュータ・プログラム製品は、プロセッサに本発明の態様を実行させるためのコンピュータ可読プログラム命令を有する１つの（または複数の）コンピュータ可読記憶媒体を含んでよい。

【００２０】

コンピュータ可読記憶媒体は、命令実行デバイスによって使用されるように命令を保持すること、および記憶することができる有形デバイスであることが可能である。コンピュータ可読記憶媒体は、例えば、電子ストレージ・デバイス、磁気ストレージ・デバイス、光ストレージ・デバイス、電磁ストレージ・デバイス、半導体ストレージ・デバイス、または以上の任意の適切な組合せであってよいが、これらには限定されない。コンピュータ可

読記憶媒体のより具体的な例の網羅的でないリストは、以下、すなわち、ポータブル・コンピュータ・ディスケット、ハード・ディスク、ランダム・アクセス・メモリ（ＲＡＭ）、読取り専用メモリ（ＲＯＭ）、消去可能なプログラマブル読取り専用メモリ（ＥＰＲＯＭもしくはフラッシュ・メモリ）、スタティック・ランダム・アクセス・メモリ（ＳＲＡＭ）、ポータブル・コンパクト・ディスク読取り専用メモリ（ＣＤ－ＲＯＭ）、デジタル・バーサタイル・ディスク（ＤＶＤ）、メモリ・スティック、フロッピ・ディスク、パンチカード、もしくは命令が記録されている溝の中の隆起構造などの機械的に符号化されたデバイス、および以上の任意の適切な組合せを含む。本明細書において使用されるコンピュータ可読記憶媒体は、電波もしくは他の自由に伝播する電磁波、導波路もしくは他の伝送媒体を伝播する電磁波（例えば、光ファイバ・ケーブルを通過する光パルス）、または配線を通して伝送される電気信号などの、一過性の信号自体であると解釈されるべきではない。

【００２１】

本明細書において説明されるコンピュータ可読プログラム命令は、コンピュータ可読記憶媒体からそれぞれのコンピューティング／処理デバイスにダウンロードされること、またはネットワーク、例えば、インターネット、ローカル・エリア・ネットワーク、ワイド・エリア・ネットワーク、または無線ネットワーク、あるいはその組合せを介して外部コンピュータもしくは外部ストレージ・デバイスにダウンロードされることが可能である。ネットワークは、銅伝送ケーブル、伝送光ファイバ、無線伝送、ルータ、ファイアウォール、スイッチ、ゲートウェイ・コンピュータ、またはエッジ・サーバ、あるいはその組合せを含んでよい。各コンピューティング／処理デバイスにおけるネットワーク・アダプタ・カードまたはネットワーク・インターフェースが、ネットワークからコンピュータ可読プログラム命令を受信し、そのコンピュータ可読プログラム命令を、それぞれのコンピューティング／処理デバイス内のコンピュータ可読記憶媒体に記憶するために転送する。

【００２２】

本発明の動作を実行するためのコンピュータ可読プログラム命令は、アセンブラ命令、命令セット・アーキテクチャ（ＩＳＡ）命令、マシン命令、マシン依存命令、マイクロコード、ファームウェア命令、状態設定データ、集積回路のための構成データ、またはＳｍａｌｌｔａｌｋ（Ｒ）、Ｃ＋＋、もしくは類似したものなどのオブジェクト指向プログラミング言語、および「Ｃ」プログラミング言語もしくは類似したプログラミング言語などの手続き型プログラミング言語を含め、１つまたは複数のプログラミング言語の任意の組合せで書かれたソース・コードもしくはオブジェクト・コードであってよい。コンピュータ可読プログラム命令は、完全にユーザのコンピュータ上で、部分的にユーザのコンピュータ上で、スタンドアロンのソフトウェア・パッケージとして、部分的にユーザのコンピュータ上で、かつ部分的に遠隔コンピュータ上で、または完全に遠隔コンピュータもしくは遠隔サーバの上で実行されてよい。完全に遠隔コンピュータもしくは遠隔サーバの上で実行されるシナリオにおいて、遠隔コンピュータは、ローカル・エリア・ネットワーク（ＬＡＮ）もしくはワイド・エリア・ネットワーク（ＷＡＮ）を含む、任意のタイプのネットワークを通じてユーザのコンピュータに接続されてよく、または接続は、外部コンピュータに対して行われてよい（例えば、インターネット・サービス・プロバイダを使用してインターネットを通じて）。一部の実施形態において、例えば、プログラマブル論理回路、フィールド・プログラマブル・ゲート・アレイ（ＦＰＧＡ）、またはプログラマブル・ロジック・アレイ（ＰＬＡ）を含む電子回路が、本発明の態様を実行するためにその電子回路を個人設定すべく、コンピュータ可読プログラム命令の状態情報を利用することによってコンピュータ可読プログラム命令を実行してよい。

【００２３】

本発明の態様は、本発明の実施形態による方法、装置（システム）、およびコンピュータ・プログラム製品のフローチャートまたはブロック図、あるいはその両方を参照して本明細書において説明される。フローチャートまたはブロック図、あるいはその両方の各ブロック、ならびにフローチャートまたはブロック図、あるいはその両方におけるブロックの

組合せは、コンピュータ可読プログラム命令によって実装されることが可能であることが理解されよう。

【００２４】

これらのコンピュータ可読プログラム命令は、コンピュータまたは他のプログラマブル・データ処理装置のプロセッサを介して実行される命令が、フローチャートまたはブロック図、あるいはその両方の１つまたは複数のブロックにおいて指定される機能／動作を実施するための手段を作り出すべく、汎用コンピュータ、専用コンピュータ、または他のプログラマブル・データ処理装置のプロセッサに提供されてマシンを作り出すものであってよい。また、これらのコンピュータ可読プログラム命令は、コンピュータ可読記憶媒体に記憶された命令により、フローチャートまたはブロック図、あるいはその両方の１つまたは複数のブロックにおいて指定される機能／動作の態様を実施する命令を含む製品を作り出すべく、コンピュータ可読記憶媒体に記憶されて、コンピュータ、プログラマブル・データ処理装置、または他のデバイス、あるいはその組合せに特定の様態で機能するように指示するものであってもよい。

【００２５】

また、コンピュータ可読プログラム命令は、コンピュータ、他のプログラマブル・データ処理装置、または他のデバイスの上で実行される命令が、フローチャートまたはブロック図、あるいはその両方の１つまたは複数のブロックにおいて指定される機能／動作を実施するように、コンピュータによって実施されるプロセスを作り出すべく、コンピュータ、他のプログラマブル・データ処理装置、または他のデバイスにロードされて、コンピュータ、他のプログラマブル・データ処理装置、または他のデバイスの上で一連の動作ステップを実行させるものであってもよい。

【００２６】

図におけるフローチャートおよびブロック図は、本発明の様々な実施形態によるシステム、方法、およびコンピュータ・プログラム製品の可能な実装形態のアーキテクチャ、機能、および動作を示す。これに関して、フローチャートまたはブロック図における各ブロックは、指定された論理機能を実施するための１つまたは複数の実行可能命令を含む、命令のモジュール、セグメント、または一部分を表すことが可能である。一部の代替の実装形態において、ブロックに記載される機能は、図に記載される順序を外れて生じてよい。例えば、連続して示される２つのブロックが、実際には、実質的に同時に実行されてよく、またはそれらのブロックが、関与する機能に依存して、ときとして、逆の順序で実行されてよい。また、フローチャートまたはブロック図、あるいはその両方の各ブロック、ならびにフローチャートまたはブロック図、あるいはその両方におけるブロックの組合せは、指定された機能もしくは動作を実行する、または専用ハードウェア命令とコンピュータ命令の組合せを実行する専用ハードウェア・ベースのシステムによって実装されることが可能であることにも留意されたい。

【００２７】

説明される以下の例示的な実施形態は、マルチキャスト・プロトコルを使用して遠隔アドレス指定可能性を帯域内で発見し、次に、遠隔チャネル定義が確認され、かつ初期設定された直接に接続された管理経路を確立することによって、２つのコンピューティング・システムの間で１つまたは複数の通信チャネルのセットを確立するシステム、方法、およびプログラム製品を提供する。

【００２８】

図１を参照すると、少なくとも１つの実施形態による例示的なネットワーク化されたコンピューティング環境１００が示される。ネットワーク化されたコンピュータ環境１００は、通信ネットワーク１１２を介して互いに接続されたネットワーク・ノード１０２Ａおよび１０２Ｂを含んでよい。少なくとも１つの実装形態によれば、ネットワーク化されたコンピュータ環境１００は、複数のネットワーク・ノード１０２Ａ、１０２Ｂを含んでよく、そのうちの２つだけが、例示を簡単にするために図示される。

【００２９】

通信ネットワーク１１２は、ワイド・エリア・ネットワーク（ＷＡＮ）、ローカル・エリア・ネットワーク（ＬＡＮ）、遠隔通信ネットワーク、無線ネットワーク、公衆交換ネットワーク、または衛星ネットワーク、あるいはその組合せなどの、様々なタイプの通信ネットワークを含んでよい。通信ネットワーク１１２は、配線、無線通信リンク、または光ファイバ・ケーブルなどの接続を含んでよい。図１は、一実装形態の例示だけを含み、異なる実施形態が実施されてよい環境に関して何ら限定を暗示するものではないことが認識されることが可能である。示される実施形態の多くの変形形態が、設計上および実施上の要件に基づいて行われてよい。

【００３０】

本発明の一実施形態による、ソフトウェア・プログラム１０８Ａならびにノード発見およびアービタ・プログラム１１０Ａをホストすること、および実行すること、ならびに通信ネットワーク１１２を介してネットワーク・ノード１０２Ｂと通信することを可能にされたネットワーク・ノード１０２Ａは、プロセッサ１０４Ａと、データ・ストレージ・デバイス１０６Ａとを含んでよい。ネットワーク・ノード１０２Ａは、例えば、モバイル・デバイス、電話機、携帯情報端末、ネットブック、ラップトップ・コンピュータ、タブレット・コンピュータ、デスクトップ・コンピュータ、あるいはプログラムを実行すること、およびネットワークにアクセスすることができる任意のタイプのコンピューティング・デバイスであってよい。図５を参照して説明されるとおり、ネットワーク・ノード１０２Ａは、内部構成要素５０２ａと、外部構成要素５０４ａとをそれぞれ含んでよい。また、ネットワーク・ノード１０２Ａは、サービスとしてのソフトウェア（ＳａａＳ）、サービスとしてのプラットフォーム（ＰａａＳ）、またはサービスとしてのインフラストラクチャ（ＩａａＳ）などのクラウド・コンピューティング・サービス・モデルにおいて動作してもよい。また、ネットワーク・ノード１０２Ａは、プライベート・クラウド、コミュニティ・クラウド、パブリック・クラウド、またはハイブリッド・クラウドなどのクラウド・コンピューティング展開モデルに配置されてもよい。

【００３１】

本発明の一実施形態による、ソフトウェア・プログラム１０８Ｂならびにノード発見およびアービタ・プログラム１１０Ｂをホストすること、および実行すること、ならびに通信ネットワーク１１２を介してネットワーク・ノード１０２Ａと通信することを可能にされたネットワーク・ノード１０２Ｂは、プロセッサ１０４Ｂと、データ・ストレージ・デバイス１０６Ｂとを含んでよい。ネットワーク・ノード１０２Ｂは、例えば、モバイル・デバイス、電話機、携帯情報端末、ネットブック、ラップトップ・コンピュータ、タブレット・コンピュータ、デスクトップ・コンピュータ、あるいはプログラムを実行すること、およびネットワークにアクセスすることができる任意のタイプのコンピューティング・デバイスであってよい。図５を参照して説明されるとおり、ネットワーク・ノード１０２Ｂは、内部構成要素５０２ａと、外部構成要素５０４ａとをそれぞれ含んでよい。また、ネットワーク・ノード１０２Ｂは、サービスとしてのソフトウェア（ＳａａＳ）、サービスとしてのプラットフォーム（ＰａａＳ）、またはサービスとしてのインフラストラクチャ（ＩａａＳ）などのクラウド・コンピューティング・サービス・モデルにおいて動作してもよい。また、ネットワーク・ノード１０２Ｂは、プライベート・クラウド、コミュニティ・クラウド、パブリック・クラウド、またはハイブリッド・クラウドなどのクラウド・コンピューティング展開モデルに配置されてもよい。

【００３２】

ソフトウェア・プログラム１０８Ａ、１０８Ｂは、ノード発見およびアービタ・プログラム１１０Ａ、１１０Ｂによって確立された接続を活用するいくつかのより高レベルのソフトウェア・プログラムのいずれかであってよい。

【００３３】

本実施形態によれば、ノード発見およびアービタ・プログラム１１０Ａ、１１０Ｂは、データ・チャネル通信の目的で、遠隔ホスト・ノードの身元を発見するランデブ機能を実行すること、およびマルチキャスト・プロトコルを使用することができ、かつ、その後、キ

ュー・ペアのセットを結び付けるプログラムであってよい。ノード発見およびアービタ方法は、図２に関連して後段でさらに詳細に説明される。一実施形態において、ノード発見およびアービタ・プログラム１１０Ａ、１１０Ｂは、コンピューティング・ノード１０２Ａ、１０２Ｂのためのオペレーティング・システムのデバイス・ドライバ層において実行される。

【００３４】

次に、図２を参照すると、動作フローチャートが、遠隔ノード発見およびデータ・チャネル初期設定のための方法を示す。２０２において、ノード発見およびアービタ・プログラム１１０Ａ、１１０Ｂが、システム間のポイント・ツー・ポイント・リンクを接続するようにハードウェアを準備する。このアダプタ初期設定ステップは、以下の３つの段階に分割されてよい。すなわち、段階１が、ノード発見およびアービタ・プログラム１１０Ａ、１１０Ｂが、アダプタを知られている状態にするために、アダプタによって使用される仮想機能上で機能レベル・リセットをトリガすることを伴ってよい。アダプタに関して定義されたそのようなチャネルが存在しない場合、ノード発見およびアービタ・プログラム１１０Ａ、１１０Ｂは、初期設定を中止してよい。次に、ノード発見およびアービタ・プログラム１１０Ａ、１１０Ｂは、ＤＭＡトランザクションおよびＭＳＩ－Ｘ割込みに関してアダプタをイネーブルにしてよい。

【００３５】

アダプタ初期設定の段階２は、ノード発見およびアービタ・プログラム１１０Ａ、１１０Ｂが、アダプタにコマンドを渡すこと、およびアダプタからコマンドを受け取ることに用いられる通信チャネルを確立することを伴ってよい。アダプタ初期設定の段階３は、ノード発見およびアービタ・プログラム１１０Ａ、１１０Ｂが、アダプタに対するコマンドにおいて後に使用されることが可能な極めて重要なハードウェア構成情報に関してアダプタにクエリを行うことを含んでよい。さらに、ノード発見およびアービタ・プログラム１１０Ａ、１１０Ｂは、アダプタ上のイベントについて通知を受けるべく、かつ遠隔直接メモリ・アクセス（ＲＤＭＡ）トランザクションを可能にすべく、イベント・キューおよびメモリ保護テーブル・エントリに登録する。また、ノード発見およびアービタ・プログラム１１０Ａ、１１０Ｂは、この時点で、割当てにおいてＱＰを確保することもし、これらのＱＰが、次に、セットに細分され、ここで、データ・チャネルは、ＱＰのセットから成る。ノード発見およびアービタ・プログラム１１０Ａ、１１０Ｂは、初期設定の単純化をもたらすべく、かつフロー・インターフェースを解体すべく、ＱＰのそれらのセットを順次に割り当てる。アダプタが、通知、ＤＭＡ、割込み、および他のそのようなトラフィックに対して準備万端となると、ノード発見およびアービタ・プログラム１１０Ａ、１１０Ｂは、発見トラフィック、タイミング・トラフィック、および管理トラフィック、ならびにデータ・トラフィックが流れることが可能なフロー・インターフェースを確立してよい。

【００３６】

本発明の一実施形態において、アダプタ初期設定プロセスの段階３中、ノード発見およびアービタ・プログラム１１０Ａ、１１０Ｂは、前もってアダプタからバッファ・リソースを確保してよい。バッファ・プールの確保は、費用のかかる動作であり、この実施形態において、バッファ・プール確保動作は、アダプタ初期設定中、１回だけ実行されてよく、このことが、それらのリソースを敏捷な割当てのために後に利用可能にする。一実施形態において、ノード発見およびアービタ・プログラム１１０Ａ、１１０Ｂは、ネットワーク・アダプタなどのデバイスが、様々なＰＣＩ　Ｅｘｐｒｅｓｓ拡張（商標）ハードウェア機能のなかでそのデバイスのリソースに別々にアクセスすることを可能にするＰＣＩｅである、単一ルート入力／出力仮想化（ｓｉｎｇｌｅ－ｒｏｏｔ　ｉｎ／ｏｕｔ　ｖｉｒｔｕａｌｉｚａｔｉｏｎ：ＳＲＩＯＶ）と呼ばれるインターフェースを使用してバッファ・プール確保を行う。このインターフェースの下で、１つのデバイス・ドライバが、それらのリソースの管理所有者であってよく、ノード発見およびアービタ・プログラム１１０Ａ、１１０Ｂは、ゲストであってよい。

【００３７】

次に、２０４において、ノード発見およびアービタ・プログラム１１０Ａ、１１０Ｂが、発見プロセスを通じてリンクの他方の側の遠隔システムを発見する。ノード／遠隔ホスト発見を実行するために、ノード発見およびアービタ・プログラム１１０Ａ、１１０Ｂは、専有のペイロードと一緒にマルチキャスト宛先ＭＡＣアドレスを使用してよい。ポートが現れ、かつホスト・ファームウェアが構成されているとき、ノード発見およびアービタ・プログラム１１０Ａ、１１０Ｂは、フロー・インターフェースとしても知られるＱＰを確立してよく、マルチキャストＭＡＣアドレスおよびペイロードを有するメッセージを生のイーサネット（Ｒ）・パケットとして周期的に発信してよい。フロー・インターフェースを立ち上げることに加えて、ノード発見およびアービタ・プログラム１１０Ａ、１１０Ｂは、指定されたＭＡＣアドレスを有するトラフィックをいつ転送すべきかを知るために、ハードウェアにおいてＭＡＣアドレスを構成してよい。遠隔ホスト・ファームウェアが作動するようになった後、遠隔ホストは、この指定されたＭＡＣアドレス上でイーサネット（Ｒ）・パケットを受信してよく、かつそのペイロードをメモリに入れてよい。そのペイロードは、リンクの遠隔側についての情報を包含し、接続を確認するのに使用されてよい。リンク情報に加えて、ネットワーク・トラフィックの送信を可能にすべく、ノード発見およびアービタ・プログラム１１０Ａ、１１０Ｂは、リンクのいずれの側が、レベルの高い方の初期設定手続きを開始するかのアービトレーションを行うのに使用される情報を送信してよい。一実施形態において、このアービトレーション情報は、タイムスタンプ・データから成ることが可能である。アービトレーション情報の他に、そのペイロードを介して送信される情報は、イーサネット（Ｒ）・アドレスと、グローバル識別子と、管理ＱＰ番号と、シスプレックス（Ｒ）時間プロトコル（ＳＴＰ）ＱＰと、遠隔直接メモリ・アクセス（ＲＤＭＡ）キー情報と、ホスト・ファームウェア・バージョンとを含んでよい。イーサネット（Ｒ）・アドレスおよびグループ識別子（ＧＩＤ）は、向けられたネットワーク・トラフィックをポイント・ツー・ポイントとネットワーク化された交換環境の両方で送信するための識別情報であることが可能である。ノード発見およびアービタ・プログラム１１０Ａ、１１０Ｂは、ＱＰ番号およびＲＤＭＡキー情報を使用して、パケットが、特定の終点アダプタに着信した後、どこに向かうべきかを翻訳する。最後に、ファームウェア・バージョンは、制御ブロックにおいて設定されてよく、ノード発見およびアービタ・プログラム１１０Ａ、１１０Ｂは、ファームウェア・バージョンを使用して、サポートされる符号機能を識別してよい。

【００３８】

次に、２０６において、ノード発見およびアービタ・プログラム１１０Ａ、１１０Ｂが、管理トラフィック・レーンを作成する。管理トラフィック・レーンは、必要である可能性がある特定のプロトコルがどのようなものであれ、そのプロトコルのために汎用の経路を提供するネットワーク・トラフィックのための通信チャネルであることが可能である。このことは、時間同期プロトコル、またはシスプレックス（Ｒ）・ワイド・ロギング・プロトコルなどの情報収集プロトコルを含むことが可能である。代替の実施形態において、管理レーンは、チャネル特有ではない、ＳＴＰなどの複数の開始プロトコルを利用してよい。

【００３９】

ステップ２０８から２１６は、チャネル初期設定手続きを含む。より早期のステップにおいてポート・レベル／管理フロー・インターフェースのすべてが確立されると、次のステップは、チャネル初期設定手続きである。このチャネル初期設定手続きを実行するエンティティは、チャネル初期設定アービタとして知られ、いずれのチャネルが有効であるか、いずれのチャネルが誤り状態になく、または閾値に入れられているか、およびいずれのチャネルが正しいアーキテクチャ状態にあるかを判定することが可能である。定義された各チャネルに関して、チャネル初期設定アービタは、タイマを開始し、かつ接続の検証、補助的キュー・ペア接続、バッファ接続、およびノード記述子交換のステップを含むチャネル初期設定手続きを始める。

【００４０】

２０８において、ノード発見およびアービタ・プログラム１１０Ａ、１１０Ｂが、ノード

間の接続を検証する。接続検証段階は、ノード発見およびアービタ・プログラム１１０Ａ、１１０Ｂが発見したノードが正しいことを確実にするのに役立つ。この段階はまず、ノード発見およびアービタ・プログラム１１０Ａ、１１０Ｂが、リンク初期設定のノード発見部分においてアービトレーション情報を使用して確立した、リンクの能動側の動作を伴うことが可能であり、まず、ノード発見およびアービタ・プログラム１１０Ａ、１１０Ｂが、チャネルのローカル・ノードの（能動側の）定義を遠隔ノード（受動側）に送信する。チャネル定義は、チャネル・サブシステム識別子（ＣＳＳＩＤ）およびチャネル経路識別子（ＣＨＰＩＤ）として知られる２つの識別子から成ることが可能である。受動側で、能動側のチャネル情報は、受動側のチャネル情報と照合確認される。情報が合致した場合、肯定的な応答が与えられてよく、受動側の情報は、能動側に送信されてよい。情報のいずれかが合致しない場合、誤りを知らせるべく応答符号が能動側に返信されてよく、特定の誤り符号がユーザに送信されてよい。接続検証パケットが、管理ＱＰを使用して送信される。

【００４１】

次に、２１０において、ノード発見およびアービタ・プログラム１１０Ａ、１１０Ｂが、チャネルの補助的トラフィック・フロー・インターフェース（ＡＵＸ　ＱＰ）を開始する。このフロー・インターフェースは、チャネル特有の非データ情報のために使用される。チャネル特有の情報は、そのポート上の他のチャネルに影響を及ぼすことなしにこのチャネルを維持すべく、ハートビートおよび小さい回復アクションを含む。ノード発見およびアービタ・プログラム１１０Ａ、１１０Ｂは、次のステップにおいて説明されるのと同一の接続手続きを使用してＡＵＸ　ＱＰフロー・インターフェースを確立する。ＡＵＸ　ＱＰフロー・インターフェースを確立することに加えて、ノード発見およびアービタ・プログラム１１０Ａ、１１０Ｂは、アダプタから前に確保されているＱＰのセットを選択し、対応する制御メモリ領域を割り当てる。また、ノード発見およびアービタ・プログラム１１０Ａ、１１０Ｂは、何らかの新たなパケットが、そのチャネルに割り当てられたＱＰのうちの１つの上で着信しており、処理される必要があることをホスト・ファームウェアに通知するのに使用される１つの完了キューを確立することもする。

【００４２】

次に、２１２において、ノード発見およびアービタ・プログラム１１０Ａ、１１０Ｂが、バッファをネゴシエートし、かつ接続する。一実施形態において、ノード発見およびアービタ・プログラム１１０Ａ、１１０Ｂは、以下に説明される４ステップ・プロセスを利用する。すなわち、最初に、能動側と受動側の両方のホスト・ファームウェアが、それぞれが接続することを必要とする各種類のＱＰを列挙する。次に、この情報を用いて、いずれかの側のファームウェアが、各ＱＰを１つずつ接続する。使用されるべきＱＰの数をネゴシエートするために、能動側がサポートするＱＰの最大数を包含するメッセージが、能動側から受動側に管理ＱＰを使用して送信される。次に、受動側が、それらの値を受動側が有するものと照合確認し、それらの値が同一である場合、受動側は、それらの値を能動側に戻す。それらの値が同一ではない場合、２つの値のうちの後者が能動側に戻される。次に、能動側が、能動側が受信したＱＰ数を検査し、次に、ネゴシエートされた数のＱＰを１つずつ順次に通って、各側に２段階で各ＱＰを確立する。能動側のＱＰは、リセット状態からセット・アップされ、かつ初期設定され、次に、能動側が、管理ＱＰ上で受動側に接続要求を送信し、ここで、対応するＱＰは、リセット状態から初期設定される。受動側が終わると、受動側は、接続応答を能動側に戻し、能動ＱＰがトラフィックを受信するように準備を整えられ、次に、受動側に接続肯定応答を送信するように、かつ再送するように準備が整えられる。次に、受動側が、受動側のＱＰに関して同一の２ステップを実行し、能動側に「接続完了」メッセージを戻し、ここで、能動側は、次のＱＰ接続を順次に初期設定する。

【００４３】

次に、２１４において、ノード発見およびアービタ・プログラム１１０Ａ、１１０Ｂが、管理ＱＰ上でノード記述子を交換する。このことは、他方の側に特定のチャネル情報を与

え、かつデータ・トラフィックが始まることを可能にする。この情報は、能動側から受動側に行って戻る単純なピンポン・メッセージである。このステップが完了すると、そのチャネルは初期設定され、ノード発見およびアービタ・プログラム１１０Ａ、１１０Ｂが、トラフィックが送られることが可能であること、およびチャネルが公称どおりの性能を示しており、かつチャネルに誤りがないことを確かめるべく試験を行う。ノード発見およびアービタ・プログラム１１０Ａ、１１０Ｂは、チャネルの利用可能性についてソフトウェア・プログラム１０８Ａ、１０８Ｂに通知を行うべく、バックエンドにおいてリンク初期設定プログラムを呼び出してよい。

【００４４】

次に、２１６において、ノード発見およびアービタ・プログラム１１０Ａ、１１０Ｂが、定義されたチャネルのすべてが、チャネル初期設定アービタ機能を使用して初期設定されているかどうかを判定する。ノード発見およびアービタ・プログラム１１０Ａ、１１０Ｂは、定義されたチャネルのリストを横断することによってこの判定を行う。一実施例によれば、方法は、チャネル初期設定アービタ機能が、すべてのチャネルが初期設定されていると判定した（ステップ２１６、「はい」の分岐）場合、終了してよい。チャネル初期設定アービタ機能が、すべてのチャネルが初期設定されているわけではないと判定した（ステップ２１６、「いいえ」の分岐）場合、ノード発見およびアービタ・プログラム１１０Ａ、１１０Ｂは、チャネル初期設定手続きを再開すべく、かつノード間の接続を検証すべくステップ２０８に進んでよい。

【００４５】

図３は、少なくとも１つの実施形態によるリンク初期設定手続きを例示する機能ブロック図である。アダプタ初期設定３０４の実行は、図２のステップ２０２においてさらに説明される。本発明の一実施形態におけるアダプタ初期設定３０４中、ポート初期設定３０６が、図２のステップ２０４および２０６においてさらに説明される。チャネル初期設定３０８が、図２のステップ２０８から２１６においてさらに述べられる。チャネル初期設定３０８中、チャネル初期設定アービタが、いずれのチャネルが有効であるか、いずれのチャネルが誤り状態になく、または閾値に入れられているか、およびいずれのチャネルが正しいアーキテクチャ状態にあるかを判定するのに利用される。チャネル初期設定３０８の後、プロセスは、新たに作成された仮想チャネルをソフトウェア・プログラム１０８に引き渡す。

【００４６】

図２～図３は、一実施例の例示を提供するに過ぎず、異なる実施形態がどのように実施されてよいかに関して何ら限定を暗示するものではないことが認識されてよい。図示される環境の多くの変形が、設計上および実施上の要件に基づいて行われてよい。

【００４７】

図４は、本発明の実施形態による図１に示されるネットワーク・ノード１０２Ａ、１０２Ｂの内部構成要素および外部構成要素のブロック図４００である。図４は、一実施例の例示を提供するに過ぎず、異なる実施形態が実施されてよい環境に関して何ら限定を暗示するものではないことを認識されたい。図示される環境の多くの変形が、設計上および実施上の要件に基づいて行われてよい。

【００４８】

データ処理システム４０２、４０４は、マシン可読プログラム命令を実行することができる任意の電子デバイスを表す。データ処理システム４０２、４０４は、スマート・フォン、コンピュータ・システム、ＰＤＡ、またはその他の電子デバイスを表すことが可能である。データ処理システム４０２、４０４によって表されることが可能なコンピューティング・システム、コンピューティング環境、またはコンピューティング構成、あるいはその組合せの例は、パーソナル・コンピュータ・システム、サーバ・コンピュータ・システム、シン・クライアント、シック・クライアント、ハンドヘルド・デバイスもしくはラップトップ・デバイス、マルチプロセッサ・システム、マイクロプロセッサ・ベースのシステム、ネットワークＰＣ、ミニコンピュータ・システム、ならびに前述のシステムもしくは

デバイスのうちのいずれかを含む分散クラウド・コンピューティング環境を含むが、以上には限定されない。

【００４９】

ネットワーク・ノード１０２Ａ、１０２Ｂは、図４に例示される、内部構成要素４０２ａ、４０２ｂ、および外部構成要素４０４ａ、４０４ｂのそれぞれのセットを含んでよい。内部構成要素４０２のセットのそれぞれは、１つもしくは複数のバス４２６上の１つもしくは複数のプロセッサ４２０、１つもしくは複数のコンピュータ可読ＲＡＭ４２２、および１つもしくは複数のコンピュータ可読ＲＯＭ４２４、ならびに１つもしくは複数のオペレーティング・システム４２８、および１つもしくは複数のコンピュータ可読有形ストレージ・デバイス４３０を含む。ネットワーク・ノード１０２Ａ、１０２Ｂ上の１つもしくは複数のオペレーティング・システム４２８、ソフトウェア・プログラム１０８、ならびにノード発見およびアービタ・プログラム１１０Ａ、１１０Ｂは、それぞれのＲＡＭ４２２（通常、キャッシュ・メモリを含む）のうちの１つまたは複数を介してそれぞれのプロセッサ４２０のうちの１つまたは複数によって実行されるようにそれぞれのコンピュータ可読有形ストレージ・デバイス４３０のうちの１つまたは複数の上に記憶される。図４に例示される実施形態において、コンピュータ可読有形ストレージ・デバイス４３０のそれぞれは、内部ハード・ドライブの磁気ディスク・ストレージ・デバイスである。代替として、コンピュータ可読有形ストレージ・デバイス４３０のそれぞれは、ＲＯＭ４２４、ＥＰＲＯＭ、フラッシュ・メモリなどの半導体ストレージ・デバイス、あるいはコンピュータ・プログラムおよびデジタル情報を記憶することができる他の任意のコンピュータ可読有形ストレージ・デバイスである。

【００５０】

また、内部構成要素４０２ａ、４０２ｂの各セットは、ＣＤ－ＲＯＭ、ＤＶＤ、メモリ・スティック、磁気テープ、磁気ディスク、光ディスク、または半導体ストレージ・デバイスなどの１つまたは複数のポータブルのコンピュータ可読有形ストレージ・デバイス４３８から読取りを行うべく、かつデバイス４３８に書込みを行うべくＲ／Ｗドライブもしくは Ｒ／Ｗインターフェース４３２も含む。ノード発見およびアービタ・プログラム１１０Ａ、１１０Ｂなどのソフトウェア・プログラムは、それぞれのポータブルのコンピュータ可読有形ストレージ・デバイス４３８のうちの１つまたは複数の上に記憶されること、それぞれのＲ／Ｗドライブもしくは Ｒ／Ｗインターフェース４３２を介して読み取られること、およびそれぞれのハード・ドライブ４３０にロードされることが可能である。

【００５１】

また、内部構成要素４０２ａ、４０２ｂの各セットは、ＴＣＰ／ＩＰアダプタ・カード、無線Ｗｉ－Ｆｉインターフェース・カード、または３Ｇもしくは４Ｇの無線インターフェース・カード、または他の有線通信リンクもしくは無線通信リンクなどのネットワーク・アダプタもしくはネットワーク・インターフェース４３６も含む。ネットワーク・ノード１０２Ａ、１０２Ｂにおけるソフトウェア・プログラム１０８、ならびにノード発見およびアービタ・プログラム１１０Ａ、１１０Ｂは、ネットワーク（例えば、インターネット、ローカル・エリア・ネットワーク、または他のワイド・エリア・ネットワーク）およびそれぞれのネットワーク・アダプタもしくはネットワーク・インターフェース５３６を介して外部コンピュータからネットワーク・ノード１０２Ａ、１０２Ｂにダウンロードされることが可能である。ネットワーク・アダプタもしくはネットワーク・インターフェース４３６から、ネットワーク・ノード１０２Ａ、１０２Ｂにおけるソフトウェア・プログラム１０８、ならびにノード発見およびアービタ・プログラム１１０Ａ、１１０Ｂは、それぞれのハード・ドライブ４３０にロードされる。ネットワークは、銅線、光ファイバ、無線伝送、ルータ、ファイアウォール、スイッチ、ゲートウェイ・コンピュータ、またはエッジ・サーバ、あるいはその組合せを含んでよい。

【００５２】

外部構成要素４０４ａ、４０４ｂのセットのそれぞれは、コンピュータ・ディスプレイ・モニタ４４４と、キーボード４４２と、コンピュータ・マウス４３４とを含むことが可能

である。また、外部構成要素４０４ａ、４０４ｂは、タッチ・スクリーン、仮想キーボード、タッチ・パッド、ポインティング・デバイス、およびその他のヒューマン・インターフェース・デバイスを含むことも可能である。また、内部構成要素４０２ａ、４０２ｂのセットのそれぞれは、コンピュータ・ディスプレイ・モニタ４４４、キーボード４４２、およびコンピュータ・マウス４３４とインターフェースするためのデバイス・ドライバ４４０も含む。デバイス・ドライバ４４０、Ｒ／ＷドライブもしくはＲ／Ｗインターフェース４３２、およびネットワーク・アダプタもしくはネットワーク・インターフェース４３６は、ハードウェアと、ソフトウェア（ストレージ・デバイス４３０またはＲＯＭ４２４、あるいはその両方に記憶される）とを含む。

【００５３】

本開示は、クラウド・コンピューティングについての詳細な説明を含むものの、本明細書に記載される教示の実施は、クラウド・コンピューティング環境に限定されないことをあらかじめ理解されたい。むしろ、本発明の実施形態は、現在、知られている、または後に開発される他の任意のタイプのコンピューティング環境と一緒に実施されることが可能である。

【００５４】

クラウド・コンピューティングは、最小限の管理労力、またはサービス・プロバイダとの最小限の対話で迅速にプロビジョニングされ、かつリリースされることが可能である構成可能なコンピューティング・リソース（例えば、ネットワーク、ネットワーク帯域幅、サーバ、処理、メモリ、ストレージ、アプリケーション、仮想マシン、およびサービス）の共有されるプールに対する好都合な、オンデマンドのネットワーク・アクセスを可能にするためのサービス・デリバリのモデルである。このクラウド・モデルは少なくとも５つの特徴と、少なくとも３つのサービス・モデルと、少なくとも４つの展開モデルとを含む。

【００５５】

特徴は、以下のとおりである：

オンデマンド・セルフ・サービス：クラウド・コンシューマが、サービスのプロバイダとの人のやりとりを要求することなしに、必要に応じて自動的に、サーバ時間およびネットワーク・ストレージなどのコンピューティング能力を一方的にプロビジョニングすることが可能である。

ブロード・ネットワーク・アクセス：能力が、ネットワーク上で利用可能であり、かつ異種のシン・クライアント・プラットフォームまたはシック・クライアント・プラットフォーム（例えば、モバイル電話、ラップトップ、およびＰＤＡ）による使用を促進する標準機構を通じてアクセスされる。

リソース・プーリング：プロバイダのコンピューティング・リソースは、様々な物理リソースおよび仮想リソースが、デマンドに応じて動的に割り当てられること、および再割当てされることが行われて、マルチテナント・モデルを使用して複数のコンシューマに役立てられるようにプールされる。コンシューマが、一般に、提供されるリソースの厳密なロケーションをまったく管理しておらず、それについての知識も有さないが、抽象性のより高いレベルにおいてロケーション（例えば、国、州、またはデータセンタ）を指定することができる可能性があるという点で、ロケーション独立の感覚が存在する。

迅速な伸縮性：能力が、急速にスケールアウトするように、かつ急速にスケールインすべく迅速に解放されるように、一部の事例において自動的に、迅速に、かつ伸縮的にプロビジョニングされることが可能である。コンシューマには、プロビジョニングのために利用可能な能力は、しばしば、無制限であるように見え、かつ任意の時点で任意の数量で購入されることが可能である。

測定されるサービス：クラウド・システムが、サービスのタイプ（例えば、ストレージ、処理、帯域幅、およびアクティブ・ユーザ・アカウント）に適切な抽象性の何らかのレベルで測定能力を活用することによって、リソース使用を自動的に制御し、かつ最適化する。リソース使用率が、監視され、制御され、かつ報告されて、利用されるサービスのプロバイダとコンシューマの両方に透明性を提供することが可能である。

【００５６】
サービス・モデルは、以下のとおりである：
サービスとしてのソフトウェア（ＳａａＳ）：コンシューマに提供される能力は、クラウド・インフラストラクチャ上で実行されるプロバイダのアプリケーションを使用することである。アプリケーションは、ウェブ・ブラウザ（例えば、ウェブ・ベースの電子メール）などのシン・クライアント・インターフェースを通じて様々なクライアント・デバイスからアクセス可能である。コンシューマは、限られたユーザ特有のアプリケーション構成設定を可能な例外として、ネットワーク、サーバ、オペレーティング・システム、ストレージ、または個々のアプリケーション能力さえも含め、基礎をなすクラウド・インフラストラクチャを管理することも、制御することもしない。
サービスとしてのプラットフォーム（ＰａａＳ）：コンシューマに提供される能力は、プロバイダによってサポートされるプログラミング言語およびツールを使用して作成された、コンシューマによって作成された、または獲得されたアプリケーションをクラウド・インフラストラクチャ上に展開することである。コンシューマは、ネットワーク、サーバ、オペレーティング・システム、またはストレージを含め、基礎をなすクラウド・インフラストラクチャを管理することも、制御することもしないが、展開されるアプリケーション、および、場合により、アプリケーションをホストする環境構成を管理する。
サービスとしてのインフラストラクチャ（ＩａａＳ）：コンシューマに提供される能力は、コンシューマが、オペレーティング・システムと、アプリケーションとを含むことが可能な任意のソフトウェアを展開すること、および実行することができる場合に、処理、ストレージ、ネットワーク、およびその他の基礎的なコンピューティング・リソースをプロビジョニングすることである。コンシューマは、基礎をなすクラウド・インフラストラクチャを管理することも、制御することもしないが、オペレーティング・システム、ストレージ、展開されるアプリケーション、および、場合により、選定されたネットワーキング構成要素（例えば、ホスト・ファイアウォール）の限られた制御を管理する。
【００５７】
展開モデルは、以下のとおりである：
プライベート・クラウド：クラウド・インフラストラクチャは、専ら組織のために運用される。クラウド・インフラストラクチャは、組織またはサード・パーティによって管理されることが可能であり、かつ敷地内に存在することも、敷地外に存在することも可能である。
コミュニティ・クラウド：クラウド・インフラストラクチャは、いくつかの組織によって共有され、かつ共有される関心（例えば、任務、セキュリティ要件、ポリシー、およびコンプライアンス考慮事項）を有する特定のコミュニティをサポートする。クラウド・インフラストラクチャは、組織またはサード・パーティによって管理されることが可能であり、かつ敷地内に存在することも、敷地外に存在することも可能である。
パブリック・クラウド：クラウド・インフラストラクチャは、一般公衆または大きい業界グループに提供され、かつクラウド・サービスを販売する組織によって所有される。
ハイブリッド・クラウド：クラウド・インフラストラクチャは、固有のエンティティのままであるが、データ移植性およびアプリケーション移植性（例えば、クラウド間の負荷分散のためのクラウド・バースティング）を可能にする標準化された、または専有の技術によって一緒に結び付けられた２つ以上のクラウド（プライベート、コミュニティ、またはパブリック）の合成である。
【００５８】
クラウド・コンピューティング環境は、ステートレスであること、低結合、モジュール性、および意味相互運用性に焦点を合わせて、サービス指向である。クラウド・コンピューティングの中核には、互いに接続されたノードのネットワークを含むインフラストラクチャがある。
【００５９】
次に図５を参照すると、例示的なクラウド・コンピューティング環境５０が、示される。

図示されるとおり、クラウド・コンピューティング環境５０は、例えば、携帯情報端末（ＰＤＡ）もしくはセルラ電話５４Ａ、デスクトップ・コンピュータ５４Ｂ、ラップトップ・コンピュータ５４Ｃ、または自動車コンピュータ・システム５４Ｎ、あるいはその組合せなどの、クラウド・コンシューマによって使用されるローカル・コンピューティング・デバイスが通信相手とすることが可能な、１つまたは複数のクラウド・コンピューティング・ノード１００を含む。ノード１００は、互いに通信することが可能である。ノード１００は、前段で説明されるプライベート・クラウド、コミュニティ・クラウド、パブリック・クラウド、またはハイブリッド・クラウド、あるいはその組合せなどの、１つまたは複数のネットワークにおいて、物理的に、または仮想でグループ化されてよい（図示せず）。このことは、クラウド・コンピューティング環境５０が、インフラストラクチャ、プラットフォーム、またはソフトウェア、あるいはその組合せを、クラウド・コンシューマがローカル・コンピューティング・デバイス上にリソースを保持する必要のないサービスとして提供することを可能にする。図５に示されるコンピューティング・デバイス５４Ａ～Ｎのタイプは、単に例示的であることが意図されること、ならびにコンピューティング・ノード１００およびクラウド・コンピューティング環境５０は、任意のタイプのネットワークまたはネットワーク・アドレス指定可能な接続、あるいはその両方を介して任意のタイプのコンピュータ化されたデバイスと通信することが可能である（例えば、ウェブ・ブラウザを使用して）ことが理解される。

【００６０】

次に図６を参照すると、クラウド・コンピューティング環境５０によって提供される機能抽象化層のセット６００が示される。図６に示される構成要素、層、および機能は、単に例示的であることが意図され、本発明の実施形態は、これらに限定されないことをあらかじめ理解されたい。図示されるとおり、以下の層および対応する機能が提供される。

【００６１】

ハードウェアおよびソフトウェア層６０が、ハードウェア構成要素と、ソフトウェア構成要素とを含む。ハードウェア構成要素の例は、メインフレーム６１、ＲＩＳＣ（縮小命令セット・コンピュータ）アーキテクチャ・ベースのサーバ６２、サーバ６３、ブレード・サーバ６４、ストレージ・デバイス６５、ならびにネットワークおよびネットワーキング構成要素６６を含む。一部の実施形態において、ソフトウェア構成要素は、ネットワーク・アプリケーション・サーバ・ソフトウェア６７と、データベース・ソフトウェア６８とを含む。

【００６２】

仮想化層７０が、仮想エンティティの以下の例、すなわち、仮想サーバ７１、仮想ストレージ７２、仮想プライベート・ネットワークを含む仮想ネットワーク７３、仮想アプリケーションおよびオペレーティング・システム７４、ならびに仮想クライアント７５が提供されることが可能な抽象化層を提供する。

【００６３】

一実施例において、管理層８０が、後段で説明される機能を提供することが可能である。リソース・プロビジョニング８１が、クラウド・コンピューティング環境内でタスクを実行するのに利用されるコンピューティング・リソースおよびその他のリソースの動的調達を提供する。測定および価格設定８２が、クラウド・コンピューティング環境内でリソースが利用されるにつれての費用追跡、およびこれらのリソースの消費に関する料金請求またはインボイス送付を提供する。一実施例において、これらのリソースは、アプリケーション・ソフトウェア・ライセンスを含むことが可能である。セキュリティが、クラウド・コンシューマおよびタスクに関する身元検証、ならびにデータおよびその他のリソースに関する保護を提供する。ユーザ・ポータル８３が、コンシューマおよびシステム管理者にクラウド・コンピューティング環境へのアクセスを提供する。サービス・レベル管理８４が、要求されるサービス・レベルが満たされるようにクラウド・コンピューティング・リソース割当ておよび管理を提供する。サービス・レベル合意（ＳＬＡ）計画および履行８５が、ＳＬＡにより将来の要求が予期されるクラウド・コンピューティング・リソースの

事前準備および調達を提供する。

【００６４】

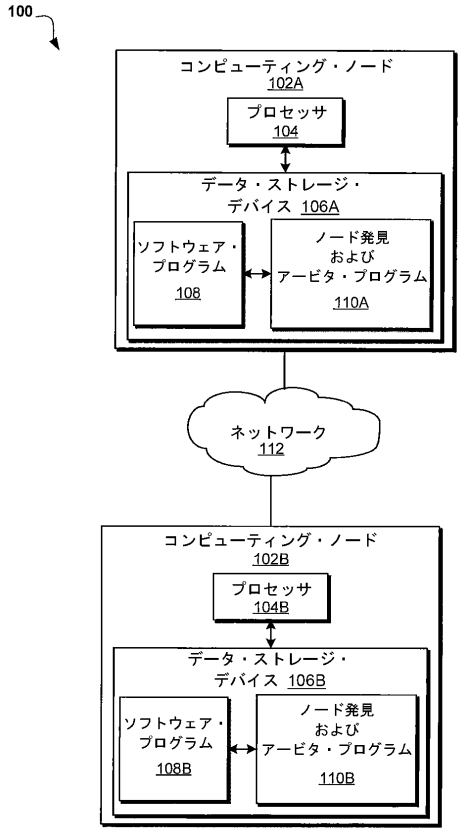
作業負荷層９０が、クラウド・コンピューティング環境が利用されることが可能な機能の例を提供する。この層から提供されることが可能な作業負荷および機能の例は、マッピングおよびナビゲーション９１、ソフトウェア開発およびライフサイクル管理９２、仮想教室教育デリバリ９３、データ・アナリティクス処理９４、トランザクション処理９５、ならびにノード発見およびアービタ・メソッド９６を含む。ノード発見およびアービタ・メソッド９６は、１つまたは複数の通信チャネルを実施する際に、遠隔ノードアドレス指定可能性を発見すべく媒体アクセス制御（ＭＡＣ）マルチキャスト手続きを使用すること、およびキュー・ペアのセットを接続すべくＩｎｆｉｎｉＢａｎｄ（Ｒ）（ＩｎｆｉｎｉＢａｎｄ（Ｒ）およびすべてのＩｎｆｉｎｉＢａｎｄ（Ｒ）ベースの商標およびロゴは、システムＩ／Ｏ株式会社、またはシステムＩ／Ｏ株式会社の関連会社、あるいはその両方の商標または登録商標である）プロトコルを使用することと関係してよい。

【００６５】

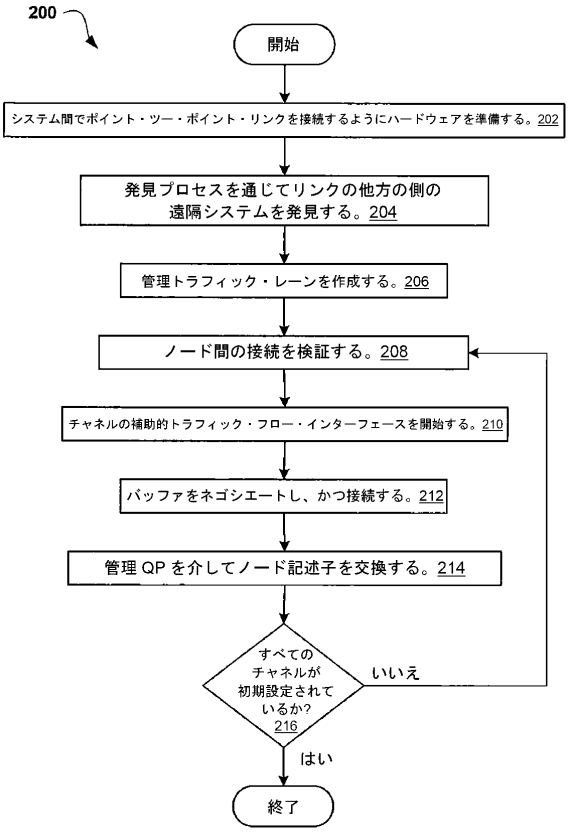
本発明の様々な実施形態の説明は、例示の目的で提示されているが、網羅的であることも、開示される実施形態に限定されることも意図していない。説明される実施形態の範囲を逸脱することなく、多くの変形形態および変更形態が当業者には明白となろう。本明細書において使用される術語は、実施形態の原理、実際的な応用、または市場において見られる技術に優る技術的改良を最もよく説明するように、あるいは当業者が、本明細書において開示される実施形態を理解することを可能にするように選択されている。

【図面】

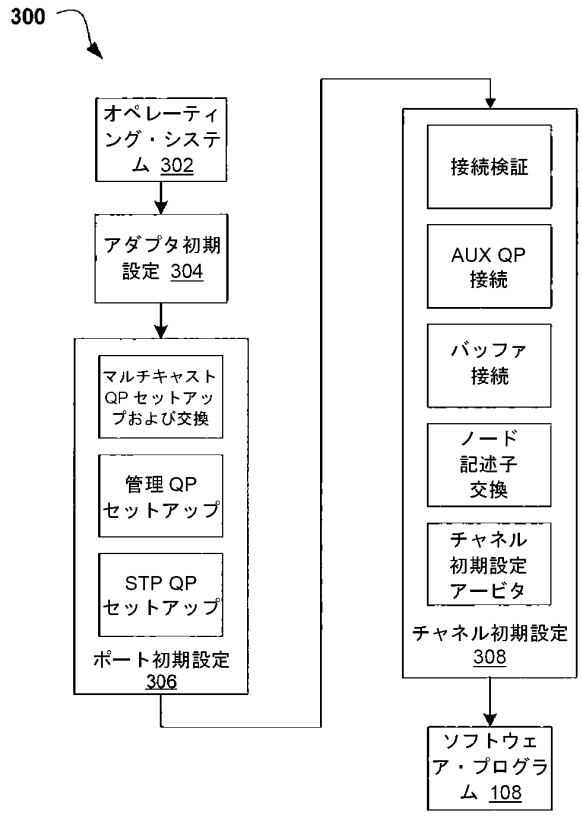
【図１】


100
コンピューティング・ノード
102A
プロセッサ
104
データ・ストレージ・
デバイス 106A
ソフトウェア・
プログラム
108
ノード発見
および
アービタ・プログラム
110A
ネットワーク
112
コンピューティング・ノード
102B
プロセッサ
104B
データ・ストレージ・
デバイス 106B
ソフトウェア・
プログラム
108B
ノード発見
および
アービタ・プログラム
110B

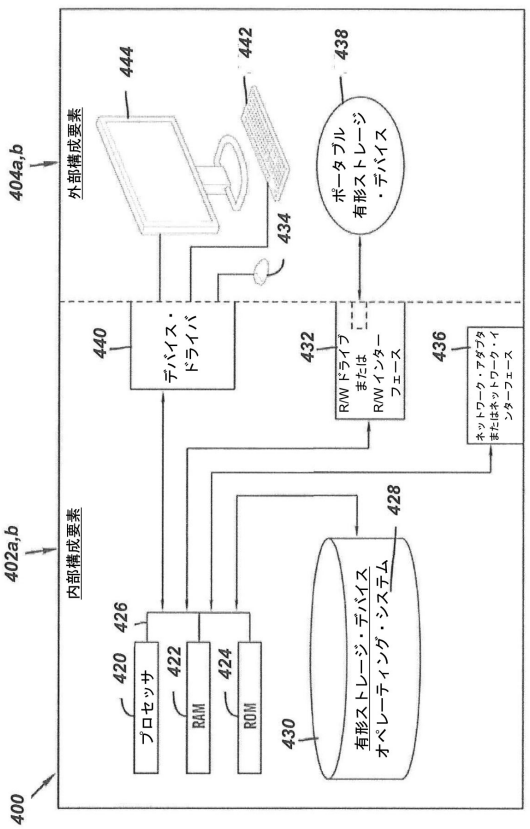
【図２】


200
開始
システム間でポイント・ツー・ポイント・リンクを接続するようにハードウェアを準備する。202
発見プロセスを通じてリンクの他方の側の
遠隔システムを発見する。204
管理トラフィック・レーンを作成する。206
ノード間の接続を検証する。208
チャネルの補助的トラフィック・フロー・インターフェースを開始する。210
バッファをネゴシエートし、かつ接続する。212
管理 QP を介してノード記述子を交換する。214
すべての
チャネルが
初期設定されて
いるか?
216
いいえ
はい
終了

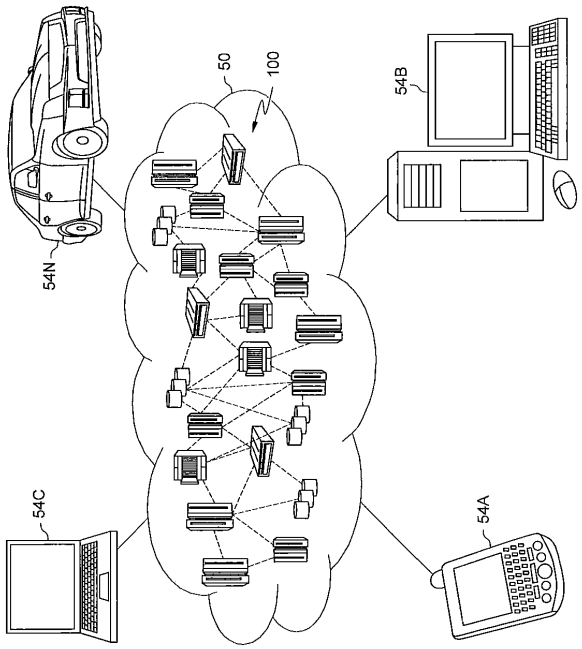
【図３】


300
オペレーティ
ング・システ
ム 302
アダプタ初期
設定 304
マルチキャスト
QP セットアッ
プおよび交換
管理 QP
セットアップ
STP QP
セットアップ
ポート初期設定
306
接続検証
AUX QP
接続
バッファ
接続
ノード
記述子
交換
チャネル
初期設定
アービタ
チャネル初期設定
308
ソフトウェ
ア・プログラ
ム 108

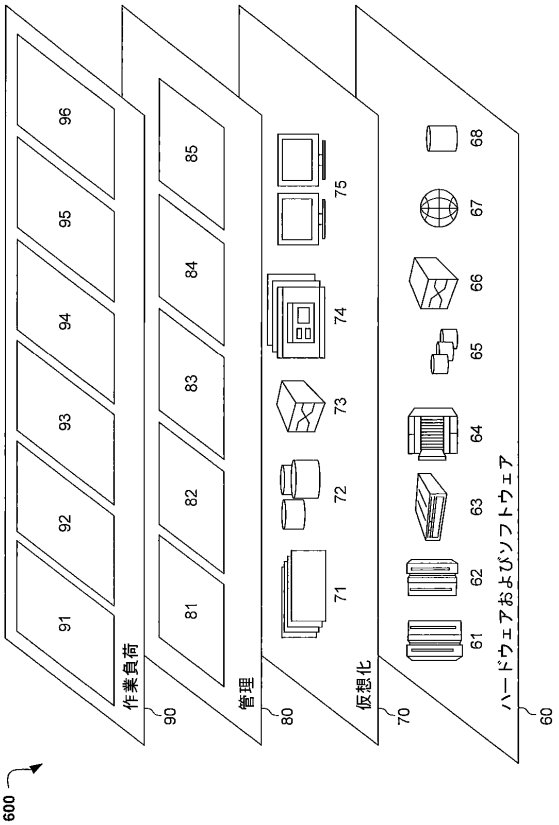
【図４】


400
404a,b
外部構成要素
402a,b
内部構成要素
444
442
438
434
ポータブル
有形ストレージ
・デバイス
440
デバイス・
ドライバ
432
R/W ドライブ
または
R/W インター
フェース
436
ネットワーク・アダプタ
またはネットワーク・イ
ンタフェース
428
426
420
プロセッサ
422
RAM
424
ROM
430
有形ストレージ・デバイス
オペレーティング・システム

【図５】


50
100
54B
54N
54C
54A

【図６】


96
95
94
93
92
91
作業負荷
90
85
84
83
82
81
管理
80
75
74
73
72
71
仮想化
70
68
67
66
65
64
63
62
61
ハードウェアおよびソフトウェア
60
600

フロントページの続き

(33)優先権主張国・地域又は機関
米国(US)
(72)発明者 ライオンズ、マイケル、ピーター
アメリカ合衆国０２１４２－１１２３ マサチューセッツ州ケンブリッジ エムエス・ピー３１８ ビニー・ストリート７５
(72)発明者 スウェッド、ピーター、ケネス
アメリカ合衆国１２６０１－５４００ ニューヨーク州ポキプシー サウス・ロード２４５５
(72)発明者 エリクソン、リチャード
アメリカ合衆国１２６０１ ニューヨーク州ポキプシー サウス・ロード２４５５
(72)発明者 リウ、シアン
アメリカ合衆国１２６０１ ニューヨーク州ポキプシー サウス・ロード２４５５

審査官 中川 幸洋

(56)参考文献
米国特許出願公開第２００９／０２１６８９３（ＵＳ，Ａ１）
米国特許出願公開第２０１７／００８５４３０（ＵＳ，Ａ１）
国際公開第２０１７／０６５９０５（WO，Ａ１）
国際公開第２０１７／０９１４６５（WO，Ａ１）
特開２００７－２４９３５７（ＪＰ，Ａ）

(58)調査した分野 (Int.Cl.，ＤＢ名)
Ｈ０４Ｌ ４５／０２１
Ｈ０４Ｌ ６７／００
Ｈ０４Ｌ ６９／００