

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4898781号
(P4898781)

(45) 発行日 平成24年3月21日(2012.3.21)

(24) 登録日 平成24年1月6日(2012.1.6)

(51) Int.Cl.

F I

H04L 12/28 (2006.01)

H04L 12/28 200Z

G06F 13/00 (2006.01)

G06F 13/00 353A

請求項の数 11 (全 25 頁)

(21) 出願番号 特願2008-503470 (P2008-503470)
 (86) (22) 出願日 平成18年3月15日(2006.3.15)
 (65) 公表番号 特表2008-535342 (P2008-535342A)
 (43) 公表日 平成20年8月28日(2008.8.28)
 (86) 国際出願番号 PCT/EP2006/060733
 (87) 国際公開番号 W02006/103168
 (87) 国際公開日 平成18年10月5日(2006.10.5)
 審査請求日 平成20年12月22日(2008.12.22)
 (31) 優先権主張番号 11/097,051
 (32) 優先日 平成17年4月1日(2005.4.1)
 (33) 優先権主張国 米国 (US)

(73) 特許権者 390009531
 インターナショナル・ビジネス・マシー
 ズ・コーポレーション
 INTERNATIONAL BUSIN
 ESS MACHINES CORPO
 RATION
 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
 (74) 代理人 100108501
 弁理士 上野 剛史
 (74) 代理人 100112690
 弁理士 太佐 種一
 (74) 代理人 100091568
 弁理士 市位 嘉宏

最終頁に続く

(54) 【発明の名称】 オペレーティング・システム・パーティションのためのネットワーク通信

(57) 【特許請求の範囲】

【請求項 1】

複数のオペレーティング・システム・パーティションを動作させる1つのコンピュータ・システム上でネットワーク通信を提供する方法であって、

前記コンピュータ・システムにおいて論理スイッチでデータ・パケットを受信するステップであって、前記データ・パケットはパケット制御記述子を含み、前記パケット制御記述子は標準ステータス、追い出しステータス、又は再循環ステータスを含むパケット・ステータスを含み、前記追い出しステータスは物理ポートから自動的に送出されるように変更されており、前記再循環ステータスは、データ・パケットを前記1つのコンピュータ・システム内の1つのオペレーティング・システム・パーティションから前記1つのコンピュータ・システム内の他のオペレーティング・システム・パーティションに内部で送る(以下、「ラップする」という)ように変更されている、前記受信するステップと、

前記データ・パケット内の情報がラップ経路を示すときに、前記コンピュータ・システムの1つのオペレーティング・システム・パーティションから前記コンピュータ・システムの他のオペレーティング・システム・パーティションへ前記コンピュータ・システム内のハードウェア経路を通して前記データ・パケットをラップするステップと、

前記データ・パケット内の情報が非ラップ経路を示すときに、オペレーティング・システム・パーティションから受信された前記データ・パケットを外部ネットワークに結合された物理ポートへ、或いは前記外部ネットワークから受信されたデータ・パケットをオペレーティング・システム・パーティションへ送るステップと

10

20

を含む、前記方法。

【請求項 2】

前記コンピュータ・システムは前記外部ネットワークに接続された少なくとも 1 つの物理ポートを含み、

各オペレーティング・システム・パーティションはデータを送信し受信するための 1 つ以上の論理ポートを有し、

前記論理ポートと、関連する物理ポートとは、前記データ・パケットを前記論理ポート及び前記関連する物理ポートを通して送るために前記論理スイッチにより選択されることゝるの仮想ポートとして動作するようになっている、請求項 1 に記載の方法。

【請求項 3】

前記データ・パケット内の前記情報は、宛先アドレスとパケット制御記述子とを含む、請求項 1 又は 2 に記載の方法。

【請求項 4】

前記データ・パケットはユニキャスト・パケット、マルチキャスト・パケット、及びブロードキャスト・パケットのうちの 1 つである、請求項 1 ~ 3 のいずれか一項に記載の方法。

【請求項 5】

前記データ・パケットがユニキャスト・パケットである場合に、前記データ・パケットを内部オペレーティング・システム・パーティションへ送るか或いは前記外部ネットワーク上の宛先へ送るかを判定するために前記データ・パケットの前記宛先アドレスが調べられる、請求項 4 に記載の方法。

【請求項 6】

前記データ・パケットがマルチキャスト・パケット又はブロードキャスト・パケットである場合に、前記データ・パケットは前記コンピュータ・システム上に実装されたマルチキャスト/ブロードキャスト・マネージャに送られる、請求項 4 に記載の方法。

【請求項 7】

前記マルチキャスト/ブロードキャスト・マネージャは、前記データ・パケットについての追加宛先のために前記データ・パケットを複製し、前記データ・パケットと複製されたデータ・パケットとについて、当該データ・パケットと当該複製されたデータ・パケットが前記マルチキャスト/ブロードキャスト・マネージャにより処理されたことを示すように前記パケット制御記述子を変更し、前記データ・パケットと前記複製されたデータ・パケットとを、当該データ・パケットと当該複製されたデータ・パケットが前記コンピュータ・システムと前記外部ネットワークとの適切な宛先に送られるよう出力する、請求項 6 に記載の方法。

【請求項 8】

前記マルチキャスト/ブロードキャスト・マネージャにより受信されたデータ・パケットが前記コンピュータ・システムのオペレーティング・システム・パーティションに送られるべきである場合に、前記マルチキャスト/ブロードキャスト・マネージャは、前記マルチキャスト/ブロードキャスト・マネージャにより受信されたデータ・パケットが送られる特定のオペレーティング・システム・パーティションを示すタグを前記データ・パケットに付け加える、請求項 7 に記載の方法。

【請求項 9】

前記受信されたデータ・パケットがマルチキャスト・パケット又はブロードキャスト・パケットであり、且つ、前記パケット制御既述子が前記受信されたデータ・パケットが前記マルチキャスト/ブロードキャスト・マネージャにより指定された追い出しパケットであることを示す場合に、前記データ・パケットは前記外部ネットワーク上の宛先に送られる、請求項 6 に記載の方法。

【請求項 10】

前記受信されたデータ・パケットがマルチキャスト・パケット又はブロードキャスト・パケットであり、且つ、前記パケット制御既述子が前記受信されたデータ・パケットが前

10

20

30

40

50

記マルチキャスト/ブロードキャスト・マネージャにより指定された再循環パケットであることを示す場合に、前記データ・パケットは、前記データ・パケットに関連付けられ且つ記マルチキャスト/ブロードキャスト・マネージャにより提供されるタグにより示される前記コンピュータ・システムのオペレーティング・システム・パーティションに送られる、請求項 6 に記載の方法。

【請求項 11】

前記標準ステータスは、当該標準ステータスを含むデータ・パケットが改変されていないことを示し、

前記追い出しステータスは、当該追い出しステータスを含むデータ・パケットが、前記物理ポートから自動的に送出されるように前記標準ステータスを含むデータ・パケットから改変されているデータ・パケットであることを示し、

10

前記再循環ステータスは、当該再循環ステータスを含むデータ・パケットが、前記ラップするように前記標準ステータスを含むデータ・パケットから改変されているデータ・パケットであることを示し、

前記非ラップ経路は前記 1 つのコンピュータ・システム内の 1 つのオペレーティング・システム・パーティションと外部ネットワークとの間でデータ・パケットが伝達される経路である、請求項 1 ~ 10 のいずれか一項に記載の方法。

【発明の詳細な説明】

【技術分野】

【0001】

20

関連出願との相互参照

本出願は、下記の同時係属中の米国特許出願と関連している。

【0002】

本願と同じ日に出願されて本発明の譲受人に譲渡された“サーバ環境においてオフロードをネットワーク化するためのホスト・イーサネット(登録商標)・アダプタ(Host Ethernet Adapter for Networking Offload in Server Environment)”と題された米国特許出願第 11/97,608 号(代理人事件整理番号 RPS 920050059US1/3485P)。

【0003】

本願と同じ日に出願されて本発明の譲受人に譲渡された、“数個のイーサネット(登録商標)・ポートと、簡易化されたフレーム・バイ・フレーム上部構造により処理されるラップ伝送されるフローとに配慮するための方法及びシステム(Method and System for Accommodating Several Ethernet Ports and a Wrap Transmitted Flow Handled by a Simplified Frame-By-Frame Upper Structure)”と題された米国特許出願第 11/96,363 号(代理人事件整理番号 RPS 920050060US1/3486P)。

30

【0004】

本願と同じ日に出願されて本発明の譲受人に譲渡された、“ネットワーク接続テーブルを提供するための方法及び装置(Method and Apparatus for Providing a Network Connection Table)”と題された米国特許出願第 11/96,571 号(代理人事件整理番号 RPS 920050061US1/3487P)。

40

【0005】

本願と同じ日に出願されて本発明の譲受人に譲渡された、“ホスト・イーサネット(登録商標)・アダプタのための設定可能なポート(Configurable Ports for a Host Ethernet Adapter)”と題された米国特許出願第 11/97,652 号(代理人事件整理番号 RPS 920050073US1/3502P)。

【0006】

50

本願と同じ日に出願されて本発明の譲受人に譲渡された、“ホスト・イーサネット（登録商標）・アダプタ（HEA）においてチェックサムを分解し、フィルタリングし、計算するためのシステム及び方法（System and Method for Parsing, Filtering, and Computing the Checksum in a Host Ethernet Adapter（HEA））”と題された米国特許出願第 11 / 96 , 365 号（代理人事件整理番号 RPS 920050074US1 / 3503P）。

【0007】

本願と同じ日に出願されて本発明の譲受人に譲渡された、“ホスト・イーサネット（登録商標）・アダプタ（HEA）において待ち時間を短縮する方法のためのシステム及び方法（System and Method for a Method for Reducing Latency in a Host Ethernet Adapter（HEA））”と題された米国特許出願第 11 / 96 , 353 号（代理人事件整理番号 RPS 920050075US1 / 3504P）。

【0008】

本願と同じ日に出願されて本発明の譲受人に譲渡された、“ネットワーク伝送のためのブラインド・チェックサム及び訂正のための方法及び装置（Method and Apparatus for Blind Checksum and Correction for Network Transmissions）”と題された米国特許出願第 11 / 97 , 055 号（代理人事件整理番号 RPS 920050076US1 / 3505P）。

【0009】

本願と同じ日に出願されて本発明の譲受人に譲渡された、“パケット・ヘッダ・ルックアップを実行する方法及びシステム（Method and System for Performing a Packet Header Lookup）”と題された米国特許出願第 11 / 96 , 362 号（代理人事件整理番号 RPS 920050082US1 / 3512P）。

【0010】

本願と同じ日に出願されて本発明の譲受人に譲渡された、“ホスト・イーサネット（登録商標）・アダプタ（HEA）においてブラインド・チェックサムを計算するシステム及び方法（System and Method for Computing a Blind Checksum in a Host Ethernet Adapter（HEA））”と題された米国特許出願第 11 / 97 , 430 号（代理人事件整理番号 RPS 920050089US1 / 3516P）。

【0011】

技術分野

本発明は、コンピュータ・ネットワークにおける通信に関し、特にオペレーティング・システム・パーティションとコンピュータ・ネットワークとの間の通信に関する。

【背景技術】

【0012】

コンピュータ・システムは、イーサネット（登録商標）のような公知標準規格を用いてネットワーク接続を確立し使用することによってネットワークを介して互いに通信する。通例、システム上で動作する複数のエンティティ又はプロセスはネットワークを介して各々独立に通信することができる。システムとの複数の接続を介して効率的に通信することを可能にすることを補助するために、ネットワーク・ポートがシステムの入力及び出力のために設けられる。例えば、システムは、ネットワークへの1つ以上の物理的ポートを含むネットワーク・アダプタ・インターフェース・カードを含むことができる。

【0013】

或るコンピュータ・システムは、複数のオペレーティング・システム（OS）が同時に独立に動作することを許す。例えば、IBM社の或るサーバ・システムは各オペレーティ

10

20

30

40

50

ング・システムをサーバ上のそれ自身のパーティションに設けることによって複数のオペレーティング・システムを動作させる。例えば、リナックス・パーティション、AIX（IBM社の商標）パーティション、異なるユニックス・パーティションが1つのサーバ上で同時に動作することができる。これらのOSパーティションの全てがネットワークを介して通信することを可能にするために、種々の方法が使用され得る。1つの方法では、サーバのI/Oホスティング・パーティションが、システム上で動作している全てのパーティションの通信ニーズを処理しなければならない。これはOSパーティションに対して低速の通信を生じさせる。なぜならば、複数のパーティションが、ネットワークを介するこれらのI/O通信において単一のI/Oホスティング・パーティションを通して妨げられ、例えば各パーティションは、単一のネットワーク・アダプタに接続されたパケット・フォワーダを含む単一のI/Oホスティング・パーティションを通してそのデータを送受信しなければならないからである。

10

【0014】

他の構成では、高速通信を可能にするために、同じ機械上で各OSパーティションにネットワーク・アダプタと物理的ネットワーク・ポートとを提供することができる。例えば、5つの異なるOSパーティションの各々に、物理的ポートと共にそれ専用のネットワーク・アダプタが設けられる。しかし、特に多数のOSパーティションが実装されるとき、その様な構成の費用は過大になり得る。

【0015】

システム上で動作する複数のOSパーティションを持つためには、システムにより受信されシステムにより送られる情報が全ての適切なネットワーク宛先に経路指定されること、及び、パーティションが同じサーバ上で動作している他のパーティションと情報をやり取りできることが必要である。他のOSパーティションは、データ・パケットについての唯一の宛先として指定され得、或いはパケットについての数個の宛先のグループに含まれ得る。例えば、イーサネット（登録商標）・ネットワーク上のブロードキャスト（BC）ネットワーク機能及びマルチキャスト（MC）ネットワーク機能は、特定のシステム又はパーティションにより送られた情報を複数の宛先が受信することを可能にする。もし1つのパーティションがネットワーク上の全ての利用可能な宛先にブロードキャスト情報を送れば、システム上の他のOSパーティションの全てはそのブロードキャスト情報を受信しなければならない。

20

30

【0016】

現存するシステムでは、幾つかの利用可能な方法のいずれかをを用いて種々のOSパーティションへの、またそれらからのネットワーク・トラフィックが提供される。1つの方法ではサーバ・システムの外側のルータを通してレイヤ3経路指定が行われ、全てのデータ・パケットがシステムの外へ送られ、また逆にシステムの適切なOSパーティションに送られる。しかし、これは、より長くかかる外部経路に起因して非効率的であり得るし、また、レイヤ2プロトコル・サポートのような所望の機能の全てをサポートしないかもしれない。他の1つの方法では、完全なイーサネット（登録商標）・スイッチをネットワーク・アダプタに埋め込むことができる。そのスイッチは、適切な宛先にパケットを提供するために必要に応じて種々のパーティションを切り替える。しかし、この方法は、イーサネット（登録商標）・スイッチがアダプタ上の現存する回路の外側にあるために高価であり、また、非標準的な特注の特徴及び機能を必要とすることがある。

40

【発明の開示】

【発明が解決しようとする課題】

【0017】

サーバ・システムについての他のネットワーク通信問題は、データ・パケットのキューイングに係る。慣例上、オペレーティング・システムから送出され又はそれにより受け取られるデータ・パケットを格納するために単一のキューが設けられ、そのオペレーティング・システムのカーネルのスタックがそのキュー中のパケットを処理する。他の或るシステムでは、他のタイプのキューイングを用いることができる。しかし、現存するキュー

50

ーイング方法は、複数のパーティションを有するシステムのニーズや、ユーザ・スペース・キュー及びOSカーネル・キューの両方を持つ可能性に対処しない。

【0018】

従って、必要なのは、限られた数の物理的ポート、設定可能な回線速度、リソースを割り当てる効率的な方法、及び種々のパケット・キューを有するコンピュータ・システムでの複数のOSパーティションの使用を含む、オペレーティング・システム・パーティション及びコンピュータ・ネットワークと効率的に且つ安価に通信するための装置及び方法である。本発明は、このようなニーズに対処する。

【課題を解決するための手段】

【0019】

本出願の発明は、コンピュータ・ネットワークにおける、またオペレーティング・システム・パーティションとコンピュータ・ネットワークとの間における、通信に関する。本発明の1つの側面において、コンピュータ・システムに設けられた複数のオペレーティング・システム・パーティションにネットワーク通信を分配するための装置はネットワークに結合された少なくとも1つの物理的ポートを含み、その物理的ポートはネットワークとコンピュータ・システムとの間の通信を可能にする。その少なくとも1つの物理的ポートに関連する論理ポートが含まれ、それらの論理ポートの各々はオペレーティング・システム・パーティションの1つと関連付けられる。論理ポートの各々は、関連する物理的ポートと関連するオペレーティング・システム・パーティションとの間での情報の伝達を可能にし、システムのネットワーク・リソースのコンフィギュレーション能力を可能にする。

【0020】

本発明の他の側面において、コンピュータ・システムに設けられた複数のオペレーティング・システム・パーティションにネットワーク通信を分配する方法は、オペレーティング・システム・パーティションと、そのオペレーティング・システム・パーティションに関連する少なくとも1つの論理ポートとの間で情報を伝達することを含む。複数のオペレーティング・システム・パーティションの各々は、関連する少なくとも1つの論理ポートと通信することができる。この方法は、そのオペレーティング・システム・パーティションと関連する少なくとも1つの論理ポートと、その少なくとも1つの論理ポートと関連する少なくとも1つの物理的ポートとの間で情報を伝達することをも含み、物理的ポートはネットワークに結合される。

【0021】

本発明の他の側面において、コンピュータ・システムにおいてネットワーク通信を分配する方法は、コンピュータ・システム内の論理スイッチでデータ・パケットを受け取ることを含み、コンピュータ・システムは複数のオペレーティング・システム・パーティションを動作させる。そのデータ・パケットは、そのパケット内の情報がラップ経路を示すときには、コンピュータ・システムの1つのオペレーティング・システム・パーティションからコンピュータ・システム内のハードウェア経路を通してコンピュータ・システムの他のオペレーティング・システム・パーティションにラップされる。データ・パケット内の情報が非ラップ経路を示すときには、データ・パケットは、もしオペレーティング・システム・パーティションから受け取られたならば、外部ネットワークに結合された物理的ポートに送られ、或いは、もし外部ネットワークから受け取られたならば、オペレーティング・システム・パーティションに送られる。

【0022】

本発明の他の側面において、コンピュータ・システムに設けられた複数のオペレーティング・システム・パーティションにネットワーク通信を分配するための装置は、外部ネットワークに結合された少なくとも1つの物理的ポートを含み、この物理的ポートは外部ネットワークとコンピュータ・システムとの間の通信を可能にする。含まれる論理ポートは、各々、オペレーティング・システム・パーティションの1つと関連付けられる。コンピュータ・システム内に実装された論理スイッチは、物理的ポート及び論理ポートを選択し、オペレーティング・システム・パーティションと物理的ポートとへの情報の伝達を可能

10

20

30

40

50

にする。

【 0 0 2 3 】

本発明の他の側面において、コンピュータ・システムにネットワーク通信を提供するための装置は、ネットワークに結合された少なくとも1つの物理的ポートを含み、この物理的ポートはネットワークとコンピュータ・システムとの間の通信を可能にする。コンピュータ・システムの少なくとも1つのユーザ・スペース内の少なくとも1つのプロセスから送られ又はそれにより受け取られるべきデータ・パケットを格納するために複数の接続キューが包含され、各接続キューは少なくとも1つのプロセスのうちの1つとネットワークとの間の異なるネットワーク接続と関連付けられる。コンピュータ・システムにより実装されたオペレーティング・システム・パーティションのカーネルから送られ又はそれにより受信されるデータ・パケットを格納するためにデフォルト・オペレーティング・システム・キューも包含される。

10

【 0 0 2 4 】

本発明の他の側面において、コンピュータ・システムのためにネットワーク通信を提供する方法はネットワーク・アダプタでデータ・パケットを受け取ることを含む。データ・パケットは接続キューに格納され、接続キューは、関連するネットワーク接続を介してコンピュータ・システムの少なくとも1つのユーザ・スペース内の少なくとも1つのプロセスから送られ又はそれにより受け取られるデータ・パケットを格納する。データ・パケットは、もしそのデータ・パケットのためにネットワーク接続が見出され得なければデフォルト・オペレーティング・システム・キューに格納され、デフォルト・オペレーティング・システム・キューは、コンピュータ・システムにより実装されたオペレーティング・システム・パーティションのカーネルから送られ又はそれにより受信されるデータ・パケットを格納する。

20

【 0 0 2 5 】

本発明は、サーバ・システムに設けられた複数のオペレーティング・システム（OS）パーティションへの高速で効率的な通信を可能にする。ソフトウェアにとってはトランスペアレントな、本発明のネットワーク・アダプタを用いて、全てのパーティションにアダプタを供給することを必要とせず、また標準的アダプタ・ハードウェアを用いて、各パーティションがそれ自身の専用イーサネット（登録商標）・アダプタを持っているのと同じような性能を達成することができる。本発明の論理スイッチは、ユニキャスト・パケット、ブロードキャスト・パケット、及びマルチキャスト・パケットを含む全てのタイプのデータ・パケットをネットワーク及びOSパーティションに効率的に提供することを可能にし、本発明のパケット・キューは、例えば複数のOSパーティションを有するシステムにおいて多様なキューイングをサポートすることを可能にする。本発明は、イーサネット（登録商標）又は類似のネットワークをバーチャル化し、リソースを割り当て、回線速度及び他のリソースを設定する効率的な方法を提供する。

30

【発明を実施するための最良の形態】

【 0 0 2 6 】

本発明は、コンピュータ・ネットワークにおける通信に関し、特にオペレーティング・システム・パーティションとコンピュータ・ネットワークとの間の通信に関する。以下の記述は、当業者が本発明品を製造し使用することを可能にするべく提示され、特許出願及びその要件と関連して提供される。本書に記載されている好ましい実施態様と一般原理と特徴とに対する種々の改変は、当業者にとっては明白であろう。従って、本発明は、示された実施態様に限定されるべく意図されてはならず、ここに記載されている原理及び特徴と矛盾しない最も広い範囲が与えられるべきである。

40

【 0 0 2 7 】

本発明は、主として、特定のインプリメンテーションで提供されたシステムに関して記載される。しかし、当業者は、この方法及びシステムが實際上他のインプリメンテーションでも働くであろうことを容易に認めるであろう。例えば、本発明で使用し得るシステム・アーキテクチャ及びネットワーク構成はいろいろな形をとることができる。本発明は、

50

一定のステップを有する特定の方法の文脈においても記載される。しかし、本方法及びシステムは、本発明と矛盾しない異なるステップ又は追加のステップ或いはその両方を有する他の方法に関しても有効に働く。

【 0 0 2 8 】

本発明の特徴をより具体的に記載するために、以下の検討と関連させて図 1 ないし 1 0 を参照されたい。本発明は、IEEE 802.3 イーサネット（登録商標）・ネットワーク・プロトコル及び伝送制御プロトコル/インターネット・プロトコル（TCP/IP）クライアント・サーバ・システムと関連して記載される。他のプロトコル（ユーザ・データグラム・プロトコル（UDP）など）及び構成も他の実施態様において使用され得る（例えば、UDP 5 タプルをキューにマッピングすることができる）。

10

【 0 0 2 9 】

図 1 は、本発明に用いるのに適するシステム 1 0 のブロック図である。システム 1 0 は、サーバ、メインフレーム、デスクトップ・クライアント・コンピュータ、ワークステーション、又は他のコンピュータ又は電子装置のようなコンピュータ・システムである。ここに記載される実施態様では、システム 1 0 は一般的にサーバ・システムと称される。良く知られているように、システム 1 0 には 1 つ以上のマイクロプロセッサ、メモリ（RAM 又は ROM 或いはその両方）、及び他のコンポーネントが設けられる。

【 0 0 3 0 】

記載される実施態様では、サーバ・システム 1 0 はネットワーク接続 1 4 を介して他の種々のコンピュータ・システム 1 2 と通信することができる。コンピュータ・システム 1 2 は、サーバ、クライアント・マシン、又は他のコンピュータ又は電子装置であり得る。例えば、サーバ・システム 1 0 は、イーサネット（登録商標）、TCP/IP のようなネットワークング・プロトコル、又は他のタイプのプロトコル、或いはその両方を用いてコンピュータ・システム 1 2 と通信することができる。

20

【 0 0 3 1 】

ここに記載される実施態様のいくつかについて、サーバ・システム 1 0 は数個のオペレーティング・システム（OS）パーティション 1 6 を含む。各パーティション 1 6 でオペレーティング・システムは独立に動作することができる。例えば、ユニックス、リナックス、IBM 社の AIX（IBM 社の商標）、マイクロソフト社のウィンドウズ（登録商標）、アップルコンピュータ社のマック OS、又はその他のようなオペレーティング・システムが各々パーティション 1 6 で動作することができ、一定タイプの OS の複数のインスタンスが別々のパーティション 1 6 で動作することができる。

30

【 0 0 3 2 】

OS パーティション 1 6 の各々は、ネットワーク 1 4 を介して情報を送受信することができ、ネットワーク上のデータ・パケットのソース又は宛先或いはその両方として行動することができる。例えば、パーティション 1 6 で動作する OS は、例えばウェブ・ページ・データをコンピュータ・システム 1 2 に提供するウェブ・サーバ・アプリケーションなどの、“ユーザ・スペース”内のサーバ・アプリケーション又はプロセスを提供することができ、又は、1 つ以上のコンピュータ・システム 1 2 により提出されたデータベース・データ又はプロセス・データを提供し、或いはそれら両方を提供する。

40

【 0 0 3 3 】

各々のアクティブなパーティション 1 6 は、ネットワーク・アダプタ 2 0 を通してネットワーク 1 4 にデータを送り、ネットワーク 1 4 からデータを受け取る。アダプタ 2 0 は、コンピュータ・システム 1 2 からネットワークを介して受け取られたデータ・パケットの適切な OS パーティション 1 6 への提供を処理する。また、ネットワーク・アダプタ 2 0 は、OS パーティション 1 6 からパケットを受け取って、その宛先がネットワーク 1 4 上のコンピュータ・システム 1 2 であっても他の OS パーティション 1 6 であっても、その情報の適切な宛先への送信を制御する。ネットワーク・アダプタは、入来データストリームの適切なデータを別々の OS パーティションに送る多重分離機能を提供する。ネットワーク・アダプタ 2 0 は、1 つ以上の物理的ポート 2 2 をネットワーク 1 4 へのハードウ

50

エア接続として提供するハードウェア・インターフェース・アダプタであり得る。アダプタ20は、以下で詳しく記載される本発明の機能を実行するハードウェア論理、マイクロプロセッサ、メモリ(RAM又はROM或いはその両方)、他の適切なコンポーネントのいずれか1つ以上を含むことができる。アダプタ20は、サーバ・システムに差し込まれる取り外し可能な別個のカードとして設けられ得、或いは他のサーバ・システム・コンポーネントに組み込まれ得る。他の実施態様では、アダプタ20の機能はソフトウェアで、或いはハードウェアとソフトウェアとの組み合わせで実行され得る。

【0034】

この明細書の全体にわたって記載される本発明の特徴と共に使用され得るシステム・コンポーネントの例が、2005年4月1日出願された、“サーバ環境におけるネットワーク化オフロードのためのホスト・イーサネット(登録商標)・アダプタ(Host Ethernet Adapter for a Networking Offload in a Server Environment)”と題された関連する多重分離特許出願第11/97,608号(RPS920050059US1/3485)に同時係属記載されて見出され得、その全体が参照により本書に組み込まれる。

10

【0035】

図2は、図1に示されているアダプタ20により使用され実装されるように提供された本発明の論理ポート・システム50の線図である。図1に関して記載されたようにシステム10においてOSパーティション16が実装される。図2の例では、利用可能な32個のOSパーティションがある。

20

【0036】

2グループ54及び56の論理ポート58が図2に示されている。本発明の論理ポート58は、システム10のOSパーティション16の、物理的ポート22へのマッピングである。図2の例では、5個の物理的ポート22a、22b、22c、22d、22eがネットワーク・アダプタ20上で利用可能である。これらの物理的ポートは、好ましくは、それらの各々にどれだけの量の帯域幅を割り当てられるかに関して設定可能である。図2の例では、4つの物理的ポート22a-dに夫々1ギガビット/秒が割り当てられ、物理的ポート22eに10ギガビット/秒が割り当てられている。他の実施態様では他の割り当てを用いることができ、好ましくは、エンド・ユーザは、ネットワーク・アダプタ20の物理的限界内で希望に応じて帯域幅割り当てを設定することができる。

30

【0037】

論理ポート58の各々は、関連付けられたOSパーティションがその特定の論理ポートを“所有する”ように、物理的ポート22をOSパーティション16の1つにマッピングする。図示されている例では、論理ポート58aはパーティション16a内のOS2を物理的ポート22aにマッピングし、論理ポート58bはパーティション16b内のOS4を物理的ポート22aにマッピングする、などである。図示されているように、複数の論理ポート58が単一のOSパーティション16を1つ又は複数の物理的ポートにマッピングすることができる。合計3個の論理ポート58がOS4を合計4個の物理的ポート22にマッピングして示されている。各物理的ポートと、その物理的ポートに割り当てられた論理ポートとは、図3に関して以下で記載される論理スイッチにより選択される。各論理ポートは、論理スイッチ上に1つの独自のポートを持つことができる。

40

【0038】

各OSパーティション16において動作するOSの観点からは、そのOSは、利用可能な物理的ポート22の排他的使用権を得る。OSパーティションが論理ポートの仮想化及び構成に気づかないように、すなわち、各OSパーティションがそれに割り当てられた論理ポートを物理的ポートと見るように、論理ポート58はOSパーティションにとってはトランスペアレントであるようにセットアップされる(論理ポートはOSにとっては正確に物理的ポート構造のようには見えない)。OSは、例えば、(図9に関して記載されるように)各論理ポートに関連するキューを見、各論理ポートに関連するリソースと影響を及ぼし合う。

50

【 0 0 3 9 】

論理ポート 5 8 は、ポート・グループ 5 4 及び 5 6 のようなポート・グループに組織され得る。各ポート・グループ内の論理ポートは、自分たちが参照することのできる物理的ポートの特定のセットに割り当てられ、それらの物理的ポートはそのポート・グループ内にあるとも考えられる。例えば、ポート・グループ 5 4 内の論理ポート 5 8 は、物理的ポート 2 2 e にではなくて物理的ポート 2 2 a - d に割り当てられ得る。同様に、ポート・グループ 5 6 内の論理ポート 5 8 は、物理的ポート 2 2 a - d にではなくて物理的ポート 2 2 e に割り当てられ得る。

【 0 0 4 0 】

各ポート・グループ 5 4 又は 5 6 は利用可能な N 個の論理ポートを持つように指定され、それは、システムが設定されるときにニーズが要求するとおりにシステム・マネージャが割り当てることのできるリソース・プールである。各論理ポートは、そのポート・グループ内のどの物理的ポートにも割り当てられ得る。これは、ユーザ、システム管理機能、又はプログラムが、ポート・グループ内の物理的ポートへの帯域幅の分配、パーティション・アクセスの分配など、各ポート・グループ内でリソースを希望通りに設定することを可能にする。システム制約条件（メモリなど）が各ポート・グループの論理ポートの数を制限し得る。例えば、図 2 の例では、各ポート・グループは 1 3 論理ポートに制限され得る。

【 0 0 4 1 】

ポート・グループ 5 4 又は 5 6 の使用は、帯域幅などの、ほぼ同じ一定量のリソース使用を各ポート・グループのために割り当ててを可能にする。例えば、4 つの 1 ギガビット物理的ポートを有することのできるポート・グループ 5 4 は、1 つの 1 0 ギガビット物理的ポートを有するポート・グループ 5 6 とほぼ同じ帯域幅リソースに制限され得る。（例えば、5 つ以上の 1 ギガビット・ポートなど、より多くの物理的ポートのために必要な I / O ピンの数などのリソース割り当てを左右する他の要素も作用し始めることができる。）物理的ポートにおけるボトルネックを減少させるために、各々のより狭い帯域幅の物理的ポート（例えば、1 ギガビット）におけるネットワーク・スループットは、好ましくは、より大きな帯域幅のポート（例えば、1 0 ギガビット）におけるスループットと同様には共有されない。図 2 の例では、各 1 ギガビット・ポート 2 2 a - d は、0.5 帯域幅使用を 4 つの物理的ポート 2 2 a - d の全てに分散させるために 5 つに及ぶ 0.5 パーティションを扱うように割り当てられる。対照的に、単一の 1 0 ギガビット物理的ポート 2 2 e は、ポート・グループ 5 6 の 1 3 個の論理ポート 5 8 の全てに帯域幅を分配する。

【 0 0 4 2 】

入来パケットを論理ポートに割り当てることができるように、各論理ポート 5 8 に識別値又は名称を割り当てることができる。例えば、一実施態様は、各論理ポート 5 8 に、論理ポートを識別するために異なる媒体アクセス制御（MAC）アドレスを割り当てることができる。MAC アドレスは、ネットワーク内のノードを一意に識別するハードウェア・アドレスである。物理的ポートに到着した各パケットは、そのパケット内の宛先（DA）MAC アドレスと一致する MAC アドレスを有する論理ポートに割り当てられる。

【 0 0 4 3 】

各論理ポートは送信及び受信動作のためにデフォルト論理ポート・キュー及びキューイング・プロセスにも関連付けられ得る。このことは、以下で図 9 及び 1 0 に関して詳しく記載される。

【 0 0 4 4 】

各論理ポート 5 8 は、好ましくは、通常各物理的ポート 2 2 と関連する、関連するリソース及び構成パラメータの自分自身のセットを有する。これらのリソース及び構成パラメータは、各論理ポートがあたかも物理的ポートであるかのように機能を実行することを可能にする。例えば、各論理ポート 5 8 は管理情報ベース（MIB）カウンタの自分自身のセットを有することができ、それは、その論理ポートにより送信及び受信されたバイト及びパケットの数、その論理ポートで検出された各タイプのエラーの数、などを数える。各

論理ポートは、各論理ポートのために独立の判定基準を用いる、パケット・タイプに基づくパケットのタイプ、又は仮想ローカル・エリア・ネットワーク（VLAN）タイプなどのフィルタリング、又はマルチキャスト・アドレスのフィルタリング、或いはそれらのフィルタリングの両方のようなフィルタリング・プロセスを各論理ポートが独立に実行することを可能にする自分自身のフィルタ・テーブルを有することができる。他のフィルタリング構成は、イーサネット（登録商標）のための“SNAP”フレーム、“JUMBO”（大型サイズ）フレーム、又はパケットのための他のフレーム・フォーマットをフィルタリングするための判定基準を含む。構成はVLAN抽出構成を含むことができ、それは入来パケットのフレームからのVLANタグの抽出を可能にする。更に、構成は、パケットの通過を許したり、エラーが発見された場合に許さなかったりするというような、チェックサム・エラーを発見したときにとられ得る処置（例えば、IP又はTCPプロトコルについて）のためにセットされ得る。また、構成リソースは、パー・コネクション・キューイング（図9に関連して以下で記載される）など、本発明の他の特徴のためのイネーブルメント・セッティングを含むことができる。

10

【0045】

論理ポート58は、単一のネットワーク・アダプタ20上に設けられ得る物理的ポート22の単一のセットを複数のOSパーティションが共有することを可能にする。本発明では、潜在的に大きな数のOSパーティション（例えば、25）が、全て、遥かに少数の物理的ポート（例えば、5）を使用するときにネットワークへのアクセスを持つことができる。他の利点は、論理ビューやレベル・ビューに影響を及ぼさずに、グループのポートを物理的レベルで例えば1×10G又は4×1G帯域幅にセットできるようなポート・グループのコンフィギュレーション能力である。従って、ポートのOSパーティション・ビューは、物理的ポート帯域幅割り当てが変化したとき、変化しなくても良い。更に、より大きな効率及び速度を目的として帯域幅を広げるように分配され得る論理ポートをOSパーティションに割り当てることができ、すなわち、各ポート・グループはほぼ同じ総帯域幅に制限され、各OSパーティションのために専用のネットワーク・アダプタ及び物理的ポートを設けることを必要とせず、高速通信を可能にする。

20

【0046】

図3は、論理スイッチを含む本発明の通信システム100のブロック図である。OSパーティション16がサーバ・システム10に設けられて、論理ポート58によってポート・グループ102にマッピングされている。物理的ポート22もポート・グループ102に含まれている。

30

【0047】

論理ポート58は、論理スイッチ104によって、関連する物理的ポート22に接続される。これらは、ネットワーク・アダプタ20において好ましくはハードウェアで実装される論理スイッチであって、OSパーティション16により送信及び受信されるパケットのための経路を決定する。論理ポートと物理的ポートとは全て論理スイッチに接続された仮想ポートになり、論理スイッチはこれらの仮想ポートのいずれにも切り替わってそのポートを通してデータを送り或いはそのポートからデータを受け取ることができる。

【0048】

図3に示されているように、論理スイッチ104は、ネットワーク・アダプタの各物理的ポート22のために設けられて、その物理的ポートを、その物理的ポート22と関連する論理ポート58に結合させる。物理的ポート105は論理スイッチ108に関連付けられ、これは論理ポート109a-c及びパーティション112、114、及び116に接続される。同様に、物理的ポート106は論理スイッチ110に関連付けられ、これは論理ポート111a-c及びパーティション118、120、及び122に接続される。

40

【0049】

ネットワーク14と物理的ポートとの間でインターフェースする外部スイッチ124を物理的ポート22に結合させることができる。このスイッチは、通常はネットワーク・アダプタ20及びシステム10の外にあるハードウェア・コンポーネントである。例えば、

50

外部スイッチ 124 は、入来するパケットを適切な物理的ポートに経路指定することができ、（例えば、それはポートへの宛先 MAC アドレスのテーブルを維持する）、全ての物理的ポートから出力されるパケットをネットワーク 14 に経路指定することができる。外部スイッチ 124 は、レイヤ 2 又はレイヤ 3 外部スイッチであり得、或いは両方のレイヤを処理することができる。

【0050】

論理スイッチ 104 は、入来するパケットを適切な OS パーティション 16 に送り、また OS パーティション 16 からの送出パケットを関連する物理的ポートに送ることができるだけでなく、1つの OS パーティション 16 からの適切なパケットを同じサーバ・システム 10 内の 1つ以上の他の OS パーティション 16 に送ることもできる。パケットが 1つの OS パーティションから他の OS パーティションに内部で送られるべきときには、それは本書では“ラップ”経路と称され、パケットがサーバ内でラップされることを示す。“非ラップ”経路は、パケットが OS パーティションと外部ネットワーク 14 との間で伝達される場合の経路である。論理スイッチ 104 は、経路 128 で示されているようにサーバ・システム 10 と外部ネットワークとの間で、経路 130 で示されているように同じ論理スイッチ 104 に存する 2つの OS パーティションの間で、また、経路 132 により示されているように異なる論理スイッチに存する 2つの OS パーティション間で、パケットを経路指定することができる。経路 132 は、2つの論理スイッチ 18 の間の経路が適切な物理的ポート 22 を通り得、また外部スイッチ 124 を通り得ることを示している。

【0051】

図 4 は、サーバ・システム上で複数の OS パーティションに、また複数の OS パーティションから、パケットを送る方法を示す流れ図である。このプロセスは、OS パーティション 16 から、物理的ポート 22 を通してネットワーク 14 から、或いはマルチキャスト/ブロードキャスト・マネージャ（後述する）から情報のパケットを得るときのネットワーク・アダプタ 20 の見地からのものである。プロセス 200 は、パケットがネットワークに送信されることと、パケットがネットワークから受信されることを記述する。このプロセスを 2つの異なるプロセス、すなわち 1つはネットワーク 14 へのパケットの送信のためのプロセス、1つはネットワークから、またラップ経路から、パケットを受信するためのプロセス、に分離し得ることを当業者は理解するであろう。この方法は、好ましくは論理回路及びゲート、プロセッサなどのようなハードウェアで実行されるが、ソフトウェアでも、或いはソフトウェアとハードウェアとの組み合わせでも実行され得る。

【0052】

本方法は 202 から始まり、ステップ 204 で、パケットがネットワーク・アダプタ 20 において得られる。1つのシナリオでは、パケットがネットワーク 14 を横断して或いは他の OS パーティション 16 へ送信されるべく OS パーティション 16 又はサーバ・システムのマルチキャスト/ブロードキャスト・マネージャ（後述する）から得られる。他のシナリオでは、パケットはコンピュータ・システム 12 又は他の装置からネットワーク 14 からネットワーク・アダプタの物理的ポート 22 を通して得られる（受信される）。

【0053】

ステップ 206 で、プロセスは、受信されたパケットで得られたパケット制御記述子を、そのパケットのステータスを判定するために、チェックする。記述される実施態様では、パケット制御記述子は、そのパケットについての異なるステータスを各々示す 3つの可能な値の 1つを持つことができる。1つの値は、パケットが“標準の”ステータスを有することを示す、すなわち、それは本発明で用いられるように改変されていないパケットである。このタイプのパケットは、普通は、OS パーティション 16 により送出されるか、又はネットワーク 14 を介して送信/受信される。第 2の値は、パケットが“追い出し（force out）”ステータスを有することを示す、すなわち、それは本発明で用いられるように改変され（図 5 に詳しく記載されているように）、ネットワーク・アダプタの物理的ポート 22 から自動的に送出されるパケットである。第 3の値はパケットが“再

循環 (r e c i r c u l a t e) ” ステータスを有することを示す、すなわち、それはサーバ・システム 10 内で自動的に OS パーティション 16 にラップされるように改変されている (図 5 に詳しく記載されているように) パケットである。

【 0 0 5 4 】

ステップ 208 で、プロセスは、パケットが標準ステータスを有するか非標準ステータスを有するかを判定する。もしそれが標準パケットであれば、プロセスはステップ 210 に続き、ここでプロセスはそのパケットがユニキャスト (UC) パケットであるか、それともマルチキャスト (MC) / ブロードキャスト (BC) パケットであるかを調べる。良く知られているように、ユニキャスト・パケットは単一の指定された宛先を有し、マルチキャスト・パケットは複数の指定された宛先を有し、ブロードキャスト・パケットは全ての使用可能な宛先に送られなければならない。もしそれが UC パケットであれば、プロセスはステップ 212 に続き、ここでプロセスは、そのパケットの宛先 (DA) MAC アドレスがそのパケットがサーバ・システム 10 の OS パーティション 16 のような内部宛先に送られるべきであることを示しているのか、それとも物理的ポート 22 を介して外のネットワークに送信されるべきであることを示しているのかをチェックする。パケットの宛先は、この MAC アドレスから判定され得る。システム 10 の各論理ポート 58 には MAC アドレスが割り当てられており、物理的ポート 22 を介してネットワークからパケットが受信されると、得られたパケットの DA MAC は、そのパケットを受信した物理的ポートと関連する論理スイッチの各論理ポートの各 MAC と比較される。

【 0 0 5 5 】

パケットが内部宛先に送られるべきであれば、プロセスはステップ 214 に続き、ここでパケットは、例えば、そのパケットにおいて見出された宛先アドレス MAC を有する論理ポートに存する OS パーティション 16 に内部で送られる。上で記載されたように、宛先 MAC アドレスは、所望の OS パーティションにマッピングされている同じ MAC アドレスを有する論理ポート 58 を見出すために使用される。パケットがラップ経路上にあるときには、すなわちそれが OS パーティション又は MC / BC マネージャから得られて別の OS パーティションに送られるのであれば、そのパケットは、あたかもそのパケットがネットワーク及び物理的ポートからのラインで受信されたかのようにネットワーク・アダプタ 20 の受信側多重分離論理に注入される。その後、プロセスは 218 で完了する。

【 0 0 5 6 】

パケットがステップ 212 で内部で送られないのであれば、すなわちそのパケットがネットワークで送信されるべきであれば、ステップ 216 において、そのパケットはネットワーク 14 を介して宛先に送られるべく適切な物理的ポート 22 から出力される。上で記載されたように、パケットは、ステップ 204 でそのパケットがそこから得られたところの OS パーティションの論理ポートを含むポート・グループと関連する物理的ポート 22 を通して送出される。その後、プロセスは 218 で完了する。

【 0 0 5 7 】

パケットがステップ 210 でマルチキャスト又はブロードキャスト (MC / BC) パケットであることが見出されたならば、プロセスはステップ 220 に続き、ここでパケットはネットワーク・アダプタ 20 からサーバ・システム 10 の内部 MC / BC マネージャに送られる。もし OS パーティション 16 がステップ 204 でそのパケットを提供したのであれば、そのパケットは、あたかもそのパケットがネットワークからのラインで受信されたかのようにラップ経路でネットワーク・アダプタ 20 の受信側多重分離論理へ送られ、MC / BC マネージャへの多重分離フローに従う。MC / BC マネージャは、サーバ・システム 10 に設けられたソフトウェア・プロセスであり得て、ハードウェアを使用するより低速であっても良いけれども、MC 又は BC パケットは通常はネットワーク・トラフィック全体の小さなパーセンテージであるので、これはシステム性能を顕著には低下させない。パケットを受信したときの MC / BC マネージャのプロセスが以下で図 5 と関連して記載される。

【 0 0 5 8 】

ステップ208で、受信されたパケットが標準ステータスを持っていないことが見出されたならば、そのパケットはMC/BCマネージャから得られた。MC/BCマネージャは、ステップ220の前の反復でそれに送られたパケットのステータスを、そのパケット制御記述子を変更することによって、改変しているか、或いは図5に関して以下で記載されるように非標準ステータスを有する複製パケットを作っている。プロセス200は、受信された非標準パケットが“追い出し”又は“再循環”のステータスを有するか判定するためにステップ222でパケット制御記述子をチェックする。もしパケットが追い出しパケットであれば、プロセスはステップ224に続き、ここでそのパケットは、そのマルチキャスト/ブロードキャストMACアドレスにより示される宛先に送信されるべく物理的ポートを介してネットワーク14へ送出される、すなわちこのパケットはMC/BCマネージャにより外部に送信されるように指定されている。その後、プロセスは218で完了する。

【0059】

もしパケットがステップ222で再循環パケットであれば、プロセスはステップ226に続き、ここでそのパケットは、パケットに付随する宛先タグ（このタグは好ましくはパケットとは別である）により示される内部論理ポートへ送られる。このタグとパケット制御既述子とは、内部ラップ経路に沿ってラップされるパケット共に受信側多重分離論理に渡される。宛先タグは、そのパケットが経路指定されるべき内部論理ポート58へのポイントを提供するためにMC/BCマネージャによりそのパケットに付け加えられたので、アダプタ20はこのステップにおいて該パケットのDA MACアドレスを無視して該パケットをそのタグで指定されている論理ポートに存するOSパーティションへ経路指定する。その後、プロセスは218で完了する。本発明における宛先タグの使用は、パケットのオリジナルDA MACアドレスが改変されずにパケットに存続することを可能にし、これによりパケットのDA MACアドレスがそれを必要とする任意の他の後のネットワーク・プロセスにより使用されることを可能にし、本インプリメンテーションの重要な利点である。

【0060】

図5は、サーバ・システム10からの又はその中での配布のためにマルチキャスト又はブロードキャスト・パケットを管理する本発明の方法250を示す流れ図である。この方法は、好ましくは、サーバ・システム10のソフトウェアで実装されてネットワーク・アダプタ20とインターフェースすることのできるMC/BCマネージャ（図7に示されている）により実行される。

【0061】

本方法は252から開始し、ステップ254で、MC/BCパケットがMC/BCマネージャによりネットワーク・アダプタ20から受信される。MC/BCパケットは、図4の方法200のステップ220でアダプタによりマネージャに送信された。マルチキャスト又はブロードキャスト・パケットは、そのパケットについてのこのMC又はBCステータスを示すMACアドレスを有する。ステップ256で、プロセスは、MC/BCマネージャが中に存するシステム10のOSパーティション16にこのパケットがラップされるべきか否か調べる。MC又はBCパケットは複数の宛先に送信されるので、マネージャは全ての宛先をチェックすることができる。この判定において、MC/BCマネージャは、残っているOSパーティションがそのパケットを受信する必要があるのかそれとも初めに受信されたパケットの複製を受信する必要があるのかを調べる。ブロードキャスト・パケットは常にラップされる必要があると共に物理的ポートからネットワーク14上の使用可能な宛先に出力される必要があり、希望に応じてラップ送信又は外部送信が始めに実行され得る（以下で記載されるように、結局は、複製されたパケットが全ての宛先に送信される）。マルチキャスト・パケットは、マルチキャストMACにより提供される指定された宛先に依存して（MC宛先は、特定のMCアドレスについて登録されているパーティションの維持されているデータベースを用いて決定される）、1つ以上の内部OSパーティションにラップされ又は外部宛先に送信されなければならないことがあり或いはその両方が行

われなければならないことがある。

【 0 0 6 2 】

マネージャは、自分が受信したパケットが内部で O S パーティション 1 6 から発せられたものであるならば、常にそのパケットをネットワークに送信されるように送出する。アダプタ 2 0 の送信側論理はそのパケットをマネージャに提供する。マネージャは、パケットのソース M A C をシステム 1 0 の内部 M A C と比較して一致を見出すことによって、それが内部で発せられたことを知り、この場合には、マネージャは常にパケットを物理的ポートから送出し、そのパケットの具体的な外部宛先は後に決定される（例えば外部スイッチにより。マネージャはシステム 1 0 の中にあり、どの具体的な外部宛先がマルチキャスト・パケットを受信するべきかを知らない）。しかし、もしマネージャがそのパケットのソース M A C をシステム 1 0 の内部 M A C のうちの 1 つと一致させなければ、それは、それがネットワーク 1 4 からアダプタ 2 0 の受信側多重分離論理を通して受信されたことを知り、外部ネットワークに送信されるべきパケットを送出しない。

10

【 0 0 6 3 】

パケットのこのコピーについてラップが行われないのであれば、プロセスはステップ 2 5 8 に続き、ここでマネージャはパケットのパケット制御既述子に適切な値を書き込むことによってパケットのステータスを“ 追い出し ” にセットする。その後、プロセスはステップ 2 6 0 に続き、ここでマネージャはパケットをネットワーク・アダプタ 2 0 に送る。ネットワーク・アダプタは、上記のように、方法 2 0 0 のステップ 2 0 4 でパケットを受信し、それを物理的ポート 2 2 から送出する。

20

【 0 0 6 4 】

このパケットについてラップが行われるのであれば、プロセスはステップ 2 6 2 に続き、ここでマネージャはパケットのパケット制御既述子に適切な値を書き込むことによってパケットのステータスを“ 再循環 ” にセットする。更に、マネージャはパケットに付随する再循環タグを付け加え、そのタグは、そのパケットが送られるべき論理ポートの名称を含む。これは、ネットワーク・アダプタ 2 0 が M A C アドレスとは異なる情報を読むことによって再循環されるパケットを論理ポートへ経路指定することを可能にし、従ってそのパケットのオリジナルのマルチキャスト M A C アドレスが改変されずに存続することを可能にする。プロセスはその後ステップ 2 6 0 に続き、ここでマネージャはそのパケット（及びタグ）をネットワーク・アダプタ 2 0 に送る。

30

【 0 0 6 5 】

ステップ 2 5 8 又はステップ 2 6 2 の後にステップ 2 6 4 が実行され、ここでプロセスは送信されるべきパケットの他のコピーがあるか否かを調べる。マルチキャスト及びブロードキャスト・パケットは複数の宛先に送られるので、M C / B C マネージャは、どの宛先が指定されているか判定し、このステップにおいていずれかの宛先にそのパケットの複製がまだ送られていないかを調べる。送信されるべき他のパケットがあるならば、ステップ 2 6 6 で、マネージャは初めに受信されたパケットを複製し、プロセスはステップ 2 5 6 に戻って、その複製されたパケットを、適切な経路指定のためにネットワーク・アダプタ 2 0 に送る前に、適宜に改変する。ステップ 2 6 4 で送信されなければならない複製パケットがそれ以上無ければ、プロセスは 2 6 8 で完了する。記載されている実施態様では、ネットワーク上の外部宛先に送られるべきパケットを外部スイッチ 1 2 4 が複製できるので、M C / B C マネージャは追い出しパケットを複製しなくても良い。

40

【 0 0 6 6 】

他の実施態様では、プロセス 2 5 0 において、記載されたステップが図示の順序とは異なる順序でも実行され得ることに留意するべきである。例えば、マネージャは、受信されたパケットの複製を作り、全てがステップ 2 5 8 又は 2 6 2 或いはその両方を介して準備されるまで全ての複製されたパケットをバッファにおいて維持し、その後それら全てをステップ 2 6 0 と同様のステップと一緒にネットワーク・アダプタに送ることができる。

【 0 0 6 7 】

図 6 は、本発明のシステム 3 0 0 のブロック図であり、図 4 に関して上で記載されたユ

50

ユニキャスト（UC）パケットの流れを示す。ネットワーク・アダプタ20は、イーサネット（登録商標）・アダプタとして示されていて、論理ポート304a-cと物理的ポート306とを参照する論理スイッチ302を含む。OSパーティション308、310、及び312が論理ポート304a-cに結合されている。一例では、パーティション308、310、及び312は、ネットワーク14上のコンピュータ・システム324及び326と同じサブネット上に存在し得る。例えば、パーティション308は11というMACアドレスと9.1.1.1というTCP/IPアドレスとを持つことができ、パーティション310は12というMACと9.1.1.2というTCP/IPアドレスとを持つことができ、パーティション312は13というMACと9.1.1.3というTCP/IPアドレスとを持つことができ、コンピュータ・システム324は31というMACと9.1.1.13というTCP/IPアドレスとを持つことができ、コンピュータ・システム326は41というMACと9.1.1.14というTCP/IPアドレスとを持つことができる。

10

【0068】

OSパーティションにより送信されたパケットをラップする例では、パーティション308は経路318で標準のユニキャスト・パケットを送信し、そのパケットは、ネットワーク・アダプタ20の論理スイッチ302により受信されるとUC宛先MACアドレス・ルックアップ・プロセス212aを受ける。UC MACルックアップは、上で図4のステップ212に記載されている。MACルックアップはパケットがラップされるべきことを示すので、パケットは論理スイッチにより送られ、続いて経路318で宛先へ続き、それは論理ポート304bを介してOSパーティション310である。もしパーティション310がパーティション308という宛先を有する標準のユニキャスト・パケットを発したならば、同様のプロセスが逆方向に行われる。

20

【0069】

別の、非ラップのパケット送信又は受信の例では、パーティション312は経路320で標準UCパケットを送出し、そのパケットは、論理スイッチ302により受信されると、同様にUC宛先MACアドレス・ルックアップ・プロセス212bを受ける。そのMACルックアップはそのパケットが外部に送出されるべきことを示すので、論理スイッチはそのパケットを経路320で物理的ポート306を通し、外部スイッチ322を通して宛先コンピュータ・システム324に送る。もしコンピュータ・システム324が物理的ポート306を通してネットワーク・アダプタ20により受信された標準ユニキャスト・パケットを発したのであれば、同様のプロセスが逆方向に行われて、パケットは論理スイッチでDA MACルックアップを受けて論理ポートを通してその宛先へ、例えば論理ポート304cからパーティション312へ提供される。

30

【0070】

図7は、本発明のシステム350のブロック図であり、図4及び5に関して上で記載されたマルチキャスト/ブロードキャスト（MC/BC）パケットのアウトバウンドの流れ（送信）を示す。ネットワーク・アダプタ20は、イーサネット（登録商標）・アダプタとして示されており、論理ポート304a-cと物理的ポート306とを参照する論理スイッチ302を含む。OSパーティション308、310、及び312が論理ポート304a-cに結合されている。一例では、図6について上で記載されたものと同様に、パーティション308、310、及び312はネットワーク14上のコンピュータ・システム324及び326と同じサブネット上に存在し得る。

40

【0071】

この例では、パケットのアウトバウンドの流れはOSパーティションから他の宛先へ提供される。パーティション308は標準のMC又はBCパケットを経路352に送出し、そのパケットは、上で図4のステップ210において記載されたように、論理スイッチ302により受信されるとMC宛先MACアドレス・チェック210aを受ける。MACチェックはそのパケットがMC又はBCパケットであることを示すので、パケットはサーバ・システム10において実装されているMC/BCマネージャ354に経路352で送ら

50

れる。

【 0 0 7 2 】

図 5 に関して上で記載されたように、MC / BC マネージャ 3 5 4 は、パケットの宛先を判定し、複製パケットを作り、パケット記述子を変更し、再循環タグを付け加える。この例では、マネージャ 3 5 4 はそれが 3 つの宛先を有する MC であること、また 4 つのパケットが、2 つは OS パーティションにラップされ、2 つは外部宛先に送信されるように、送出されるべきことを判定している。マネージャ 3 5 4 は再循環 MC パケット及び再循環タグ 3 5 6 を経路 3 5 8 でアダプタ 2 0 へ送出し、これは、タグ 3 5 6 を読んで論理ポート 3 0 4 c を見出してパケットをパーティション 3 1 2 へラップする。同様に、マネージャ 3 5 4 は再循環 MC パケットと再循環タグ 3 6 0 とを経路 3 6 2 でアダプタ 2 0 へ送出し、これは、タグ 3 6 0 を読んで論理ポート 3 0 4 b を見出してパケットをパーティション 3 1 0 へラップする。マネージャ 3 5 4 は、また、追い出し MC パケットを経路 3 6 4 でアダプタ 2 0 へ送出し、これは、これらのパケットを物理的ポート 3 0 6 から外部スイッチ 3 2 2 を通して送出する。外部スイッチ 3 2 2 は、ネットワーク 1 4 を横断して 1 つが外部コンピュータ・システム 3 2 4 に提供され、1 つが外部コンピュータ・システム 3 2 6 に提供されるように、そのパケットを複製する。

10

【 0 0 7 3 】

図 8 は、本発明のシステム 3 7 0 のブロック図であり、図 4 及び 5 に関して上に記載されたマルチキャスト / ブロードキャスト (MC / BC) パケットのインバウンドの流れ (受信) を示す。ネットワーク・アダプタ 2 0 は、イーサネット (登録商標) ・アダプタとして示され、論理ポート 3 0 4 a - c と物理的ポート 3 0 6 とを参照する論理スイッチ 3 0 2 を含む。OS パーティション 3 0 8 , 3 1 0 、及び 3 1 2 が論理ポート 3 0 4 a - c に結合されている。上記と同様に、一例では、パーティション 3 0 8 , 3 1 0 、及び 3 1 2 はネットワーク 1 4 上のコンピュータ・システム 3 2 4 及び 3 2 6 と同じサブネット上に存在し得る。

20

【 0 0 7 4 】

この例では、コンピュータ・システム 3 2 6 のような外部宛先からのパケットのインバウンドの流れが内部 OS パーティションに提供される。コンピュータ・システム 3 2 6 は標準の MC 又は BC パケットを経路 3 7 2 に送出し、外部スイッチ 3 2 2 は、もし適切ならばそのパケットの 1 コピーをネットワーク 1 4 上の他の宛先に、外部コンピュータ・システム 3 2 4 などに、送る。スイッチ 3 2 2 は他のパケット・コピーを物理的ポート 3 0 6 を通してシステム 1 0 に送り、そのパケットは経路 3 7 4 で論理スイッチへ経路指定される。上で図 4 のステップ 2 1 0 において記載されたように、パケットは、論理スイッチ 3 0 2 によって受信されると、MC 宛先 MAC アドレス・チェック 2 1 0 a を受ける。MAC チェックは、そのパケットが MC 又は BC パケットであることを示すので、そのパケットは経路 3 7 4 で MC / BC マネージャ 3 5 4 に送られる。

30

【 0 0 7 5 】

図 5 に関して上で記載されたように、MC / BC マネージャ 3 5 4 は、パケットの宛先を判定し、複製パケットを作り、パケット記述子を変更し、再循環タグを付け加える。この例では、マネージャ 3 5 4 は、それが 3 つの内部宛先を有する MC パケットであること、そして 3 つのパケットが内部 OS パーティションに送られるべきであることを判定している。マネージャ 3 5 4 は再循環 MC パケット及び再循環タグ 3 7 6 を経路 3 7 8 でアダプタ 2 0 へ送出し、これは、タグ 3 7 8 を読んで論理ポート 3 0 4 c を見出してパケットをパーティション 3 1 2 へラップする (マネージャ 3 5 4 から内部 OS パーティションへの経路は “ ラップ ” であるとも考えられる) 。同様に、マネージャ 3 5 4 は、論理ポート 3 0 4 b 及びパーティション 3 1 0 へラップされる再循環 MC パケット及び再循環タグ 3 8 0 を経路 3 8 2 で送出し、また、論理ポート 3 0 4 a 及びパーティション 3 0 8 へラップされる再循環 MC パケット及び再循環タグ 3 8 4 を経路 3 8 6 で送出する。

40

【 0 0 7 6 】

図 9 は、ネットワーク・パケットをキューイングするための本発明のキューイング・シ

50

ステム４００の線図である。サーバ・システム１０は、ＯＳパーティション・カーネルにより受信されるか又は各ＯＳパーティションのユーザ・スペースの中のアプリケーション／プロセスにより直接受信されるべきパケットを保持し、また、パーティション又はユーザ・スペース・プロセスによって物理的ポート２２へ送信されるべきパケットを保持するためのキューのセットを含むことができる。以下に記載されるキューは、パケットをキューに格納するためにネットワーク・アダプタ２０によりアクセスされ得ると共にパケットをキューから取り出すためにプロトコル・スタック及びデバイス・ドライバによりアクセスされ得るシステム１０のメイン・メモリにおいて実現され得る。

【００７７】

システム４００に含まれる第１タイプのキューは接続キュー４０２である。これらのキューは、サーバ・システムのユーザ・スペース内のプロセス４０４に送られるパケットのためにバッファを提供し、また、プロセッサ４０４から他のプロセス、ＯＳパーティション、又はネットワークの向こうの外部コンピュータ・システムなどの宛先に送られるパケットのためのバッファである。ユーザ・スペースは、アプリケーションがそこで動作するところのプロセスを提供するＯＳパーティション内のソフトウェア・スペースであり、そのアプリケーションはネットワーク接続によりネットワーク１４と通信する。慣例上ＯＳのカーネルで動作するＴＣＰ／ＩＰスタックの一部又は全部は、各ユーザ・スペースに移され、それ自身の接続キューを持つ。これは、多くのＴＣＰ／ＩＰ送信及び受信操作のためにカーネルが迂回され得るので、効率を与える。ユーザ・スペース内のプロセス／アプリケーション４０４への各ネットワーク接続は、パケットを格納するためのそれ自身の接続キュー４０２を有する。例えば、ＴＣＰ／ＩＰパケットが受信されたとき、ＴＣＰ／ＩＰスタック・プロセスは、受信されたパケットが適切な接続キューに置かれて結局は該スタックにより処理され得るように、接続ダイレクト・テーブル内の接続情報を調べて接続キュー番号を見出すことができる。そのキュー内のパケットから処理された情報がユーザ・スペース内のプロセス（例えば、各プロセス内のアプリケーション）に送られる。受信されたパケットは、ネットワーク・アダプタ２０内のハードウェアにより正しいキューに向けられ（“サーバ環境におけるネットワーク化オフロードのためのホスト・イーサネット（登録商標）・アダプタ（Host Ethernet Adapter for a Networking Offload in a Server Environment）”と題された関連特許出願、２００５年４月１日に出願された出願第１１／９７、６０８号（ＲＰＳ ９２００５００５９ ＵＳ １／３４８５に記載されている）、ユーザ・スペース内のＴＣＰスタックはそのキューへのアクセスの管理を引き継ぐ。

【００７８】

一例では、接続キューは、ＩＰパケットの断片化が予期されなくて短待ち時間が期待される永続的接続のために最善に適合し得る。ユーザ・スペース・アプリケーション又はプロセスは、ネットワーク・アダプタ１０により提供される接続キューイング・メカニズムをサポートするユーザ・スペース・ソケット・ライブラリを利用することができる。

【００７９】

システム４００に含まれる第２のタイプのキューはデフォルト論理ポート（ＬＰ）キュー４０６である。これらのキューは、サーバ・システム１０のＯＳパーティション１６のカーネル４０８に送られるパケットのためのバッファを提供する。各デフォルトＬＰキューが所望のＯＳパーティションと関連付けられた単一の論理ポート５８と関連付けられ得るように、各ＯＳパーティション１６はそれ自身のデフォルトＬＰキュー４０６を備えている。従って、各論理ポートは、それ自身のデフォルト・キューを有する論理イーサネット（登録商標）・インターフェースに対応する。通例、デフォルトＬＰキュー内のパケットは、関連するオペレーティング・システムのカーネル４０８により扱われて処理される、すなわち、オペレーティング・システムのカーネル・ネットワーク化スタックはそのオペレーティング・システムのためのデフォルト・キュー４０６を扱う。本発明では、パケットのＭＡＣアドレスのためにパー・コネクション・ルックアップがイネーブルされていなければ（パー・コネクテッド・ルックアップ・イネーブルメント（per - conn

cted lookup enablement) は、システム・マネージャ又はオペレータにより設定され得る論理ポート設定リソースの1つであり得る)、或いは、以下で図10に関して詳しく記載されるようにパケットが再循環させられるマルチキャスト/ブロードキャスト・パケットであるならば、パケットのために接続キュー402が見出され得ないときにはデフォルトLPキュー406がパケットを格納する。

【0080】

図2に関して上で記載されたように論理ポート58に属性が割り当てられ得るのと同様の仕方で、デフォルトLPキュー406(及び、以下で記載される特別デフォルト・キュー410)の各々に属性が割り当てられ得る。例えば、各キューのためのフィルタリングがイネーブル又はディスエーブルされ得る。属性は、容易なパケット管理を可能にするために設定のための同じフォーマットを持つことができる。設定はキュー・タイプ間で同じなので、或るハードウェア論理は共有され得る。

【0081】

システム400に含まれる第3のタイプのキューは特別デフォルト・キュー410である。これらのキューは、他の利用可能なキューに格納されないパケットを格納するために使われる。本発明では、3つの特別デフォルト・キュー、すなわち、ユニキャスト・デフォルト・キュー412と、マルチキャスト・デフォルト・キュー414と、ブロードキャスト・デフォルト・キュー416と、が設けられる。ユニキャスト・デフォルト・キュー412は、パケットのためにデフォルトLPキュー406が見出され得ないときにパケットを格納する。マルチキャスト・デフォルト・キュー414とブロードキャスト・デフォルト・キュー416とは、MACルックアップ・テーブル(ネットワーク・アダプタ20内のハードウェアに置かれる)のMACアドレスの1つと一致しないか又は再循環させられるマルチキャスト・パケット及びブロードキャスト・パケットを夫々格納する。これらは、以下で図10と関連して詳しく記載される。

【0082】

上で列挙された各タイプのキューは、好ましくは、ネットワーク・アダプタ20に設けられている各物理的ポートと関連付けられる。従って、各物理的ポートは、好ましくは、接続キューのセット402と、その物理的ポートのための論理ポート58と関連付けられたデフォルトLPキューのセット406と、特別デフォルト・キューのセット410とを有する。

【0083】

図10は、パケット格納のためのキューを選択する本発明の方法450を示す流れ図である。この方法は、ネットワーク14から又はラップ経路で受信されたパケットのためにネットワーク・アダプタ20においてハードウェアにより実行される。或いは、それはソフトウェアで又はサーバ・システム10の異なるコンポーネントで或いはその両方で実行され得る。

【0084】

本方法は452から開始し、ステップ454で、パケットがネットワーク・アダプタによりネットワーク14から又は内部ラップ経路を介して受信される。ステップ456で、プロセスはそれが再循環させられるパケットであるか否か、すなわち、図4及び5に関して上で記載されたようにMC/BCマネージャ354によりラップ経路へ向けて指定されたパケットであるか否か、調べる。再循環させられるパケットはマネージャによって既に処理されているのでそれらの宛先へ送られ得る。従って、もしそれが再循環させられるパケットであれば、プロセスはステップ470へ続いて、以下で記載されるようにそのパケットをデフォルトLPキューに格納する。もしそれが再循環させられるパケットでなければ、ステップ457においてプロセスはそのパケットがユニキャスト(UC)パケットであるのか或いはマルチキャスト(MC)/ブロードキャスト(BC)パケットであるのかを調べる。これは、例えば、パケットのMACアドレスから判定され得る。もしそれがMC/BCパケットであるならば、ステップ458で、プロセスは、そのパケットを受信した物理的ポートと関連付けられている特別デフォルトMCキュー414又はBCキュー4

16にそのパケットを格納する。その後、プロセスは460で終了する。図5に関して上で記載されたように、デフォルトMCキュー414及びデフォルトBCキュー416内のパケットは、結局は処理を受けるべくMC/BCマネージャ354に供給される。単一のOSパーティションに専用される物理的ポートにパケットが到着したのであれば、MC及びBCキュー内のパケットは中央OSキュー、例えばそのパーティションのためのデフォルトLPキュー406、に供給され得る。

【0085】

パケットがステップ457で判定されたUCパケットであれば、プロセスはステップ462に続き、ここで接続ルックアップが実行される。(1つのインプリメンテーションでは、たとえ接続キューに格納するとしても論理ポートについてのMACアドレス・ルックアップも実行され得、例えば、6タプル・ルックアップが実行され得、ここで6番目のタプルは論理ポート番号である。これは同じ5タプル接続情報が各論理ポート58でアピールし、また重なり合わないことを可能にする。)記載されている実施態様では、アダプタ20上のハードウェアは、パケット内の接続情報の2タプル、3タプル、又は5タプル・パターンをとり、接続が使用可能にされ又は確立されたときにそのパターンを接続テーブルに格納されている接続情報と比較しようと試みる。その様な方法の一実施態様が“ネットワーク接続テーブルを提供するための方法及び装置(Method and Apparatus for Providing a Network Connection Table)”と題された同時係属特許出願(RPS920050061US1/3487)、すなわち2005年4月1日に出願された出願第11/96,571号に記載されており、その全体が参照により本書に組み込まれる。

【0086】

次のステップ464で、プロセスは接続ルックアップについて一致が見出されたか否か、例えば、接続テーブルにおいて接続情報の一致するパターンが見出されたか否か、調べる。もしそうならば、プロセスは、一致する接続と関連付けられた接続キューにパケットを格納する。例えば、所望の接続キューを参照する番号又はポインタが、一致した接続情報と共に接続テーブルに格納され得る。パケットは、結局は接続キューから処理されてユーザ・スペース内の適切なプロセス404に提供され得る。その後、プロセスは460で終了する。

【0087】

ステップ464でパケットについて一致するパターンが発見されなければ(或いはパケット内のMACアドレスのためにパー・コネクション・キューイングがイネーブルされなければ)、プロセスはステップ468でパケットのためにデフォルトLPキューが発見されたか否かを調べる。上記のように、各論理ポート58は、関連するデフォルトLPキュー406を有する。デフォルト・キューは、論理ポートMACアドレスのリスト又はテーブルにおいてパケット内の宛先MACアドレスについてルックアップを実行することにより判定される(上記のように、各論理ポート58にMACアドレスが割り当てられている)。このステップにおいても論理ポートのMACアドレスがパケット内の宛先MACアドレスと一致したならば、プロセスはステップ470においてパケットをその論理ポートのデフォルトLPキュー406に格納し、プロセスは460で終了する。デフォルトLPキュー406内のパケットは、これらのキューと関連付けられたOSパーティション16のカーネル・ネットワーク化スタックへ経路指定される。

【0088】

パケットのためにデフォルトLPキュー406が発見されなければ、プロセスはステップ472に続いて、(もし特別キュー412がイネーブルされているならば)パケットを受信した物理的ポートと関連付けられている特別デフォルトUCキュー412にそのパケットを格納し、プロセスは460で終了する。パケットはMACアドレス及びシステム10の論理ポート58のいずれにも向けられないかもしれないので、そのパケットは設定された論理ポートMACアドレスのいずれとも一致しなくて、従ってそれはデフォルト・キュー412に格納され得る。例えば、これは、もしシステムが“プロミスカス(promiscuous)”モードで動作しているならば、そのパケットはデフォルト・キュー412に格納され得る。

miscuous)モード”であるか、又は“スニファ－(sniffer)”機能を持っているならば起こり得ることであり、この“スニファ－”機能では、送信された全てのデータ・パケットが、例えば、システム10に向けられたパケットと同様に他の宛先へのパケットも分析されるようにネットワーク活動を監視するために、ネットワーク・アダプタにより受信されて読まれ得る。或いは、これは、もしネットワーク・ブリッジ機能がシステム10によって実装されているならば起こり得ることであり、その場合、ネットワーク14は他のネットワークに接続され、各パケットは、それがネットワーク14内の宛先へ渡されるべきであるのか、それとも他の相互接続されているネットワーク上にあると知られている宛先へ転送されるべきであるのかを判定するために、分析される。

【0089】

10

デフォルトUCキュー412はオプションであるので、他の実施態様では、又はもしキュー412がディスエーブルされていれば、ステップ472は実行されず、むしろ、もしデフォルトLPキューがパケットのために発見されなければ、そのパケットは捨てられる。例えば、もし“プロミスキヤス(promiscuous)モード”がディスエーブルされれば特別キュー412はディスエーブルされるかもしれない。

【0090】

本発明の方法450は、接続ごとのキューイング(queueing per connection)と論理ポートごとのキューイング(queueing per logical port)(OSキュー)という両極端の間の柔軟性を考慮している。本発明では、両方のモデルが、それら自身のキューイングを有する或る接続及びデフォルト論理ポート・キューでキューイングされる或る接続と共に働くことができる。方法450は、複数のOSパーティションも考慮しており、また、マルチキャスト/ブロードキャスト・パケットを含む、パーティション間でのイーサネット(登録商標)のレイヤ2スイッチングをサポートする(例えば、IEEE802.3)。

20

【0091】

本発明は、図示された実施態様に従って記載されたけれども、それらの実施態様の変更があり得ること、それらの変更が本発明の範囲内にあるであろうことを当業者は直ぐに認めるであろう。従って、添付されている請求項の範囲から逸脱すること無く多くの改変が当業者によりなされ得る。

【図面の簡単な説明】

30

【0092】

【図1】本発明に用いるのに適するシステムの例のブロック図である。

【図2】図1に示されているアダプタにより使用されるように設けられ実装される本発明の論理ポート・システムの線図である。

【図3】論理スイッチを含む本発明の通信システムのブロック図である。

【図4】コンピュータ・システム上の複数のオペレーティング・システム・パーティションに、またそれらから、パケットを送る方法を示す流れ図である。

【図5】コンピュータ・システムからの又はその中での配布のためにマルチキャスト・パケット又はブロードキャスト・パケットを管理する本発明の方法を示す流れ図である。

【図6】論理スイッチを含む本発明の通信システムのブロック図である。

40

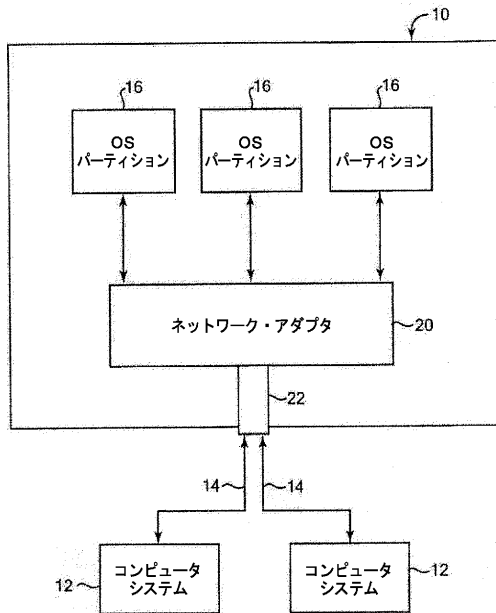
【図7】論理スイッチを含む本発明の通信システムのブロック図である。

【図8】論理スイッチを含む本発明の通信システムのブロック図である。

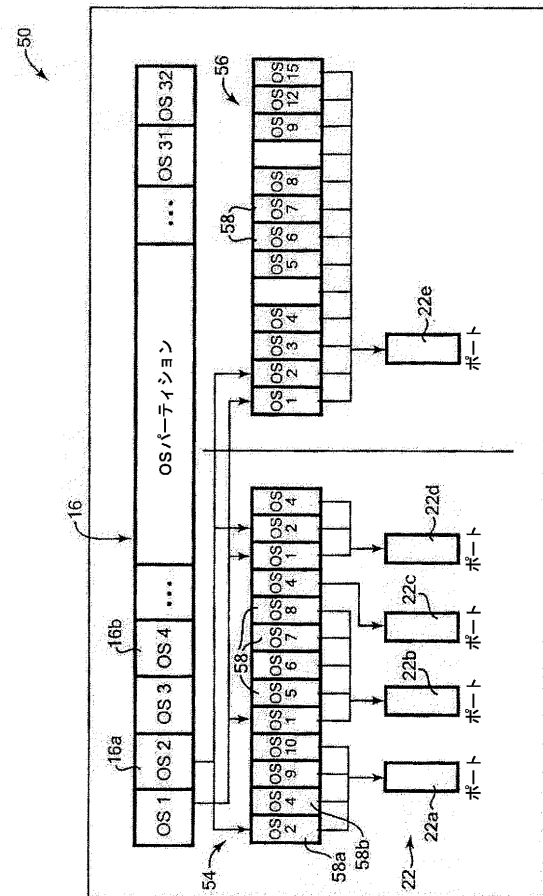
【図9】ネットワーク・パケットをキューイングするための本発明のキューイング・システムの線図である。

【図10】パケット格納のためにキューを選択する本発明の方法を示す流れ図である。

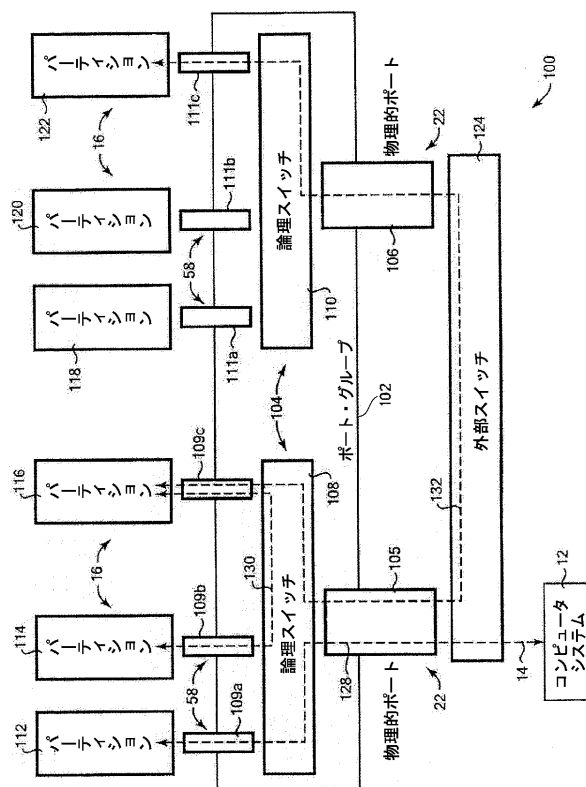
【図 1】



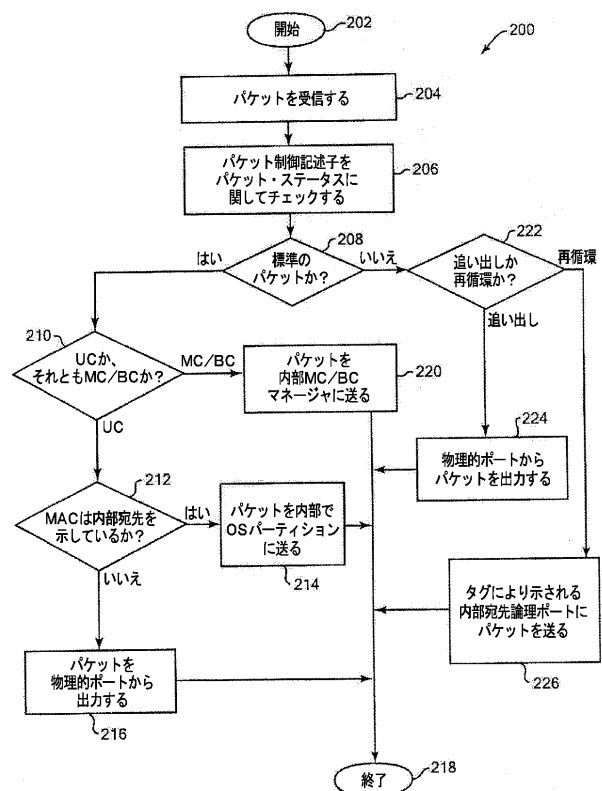
【図 2】



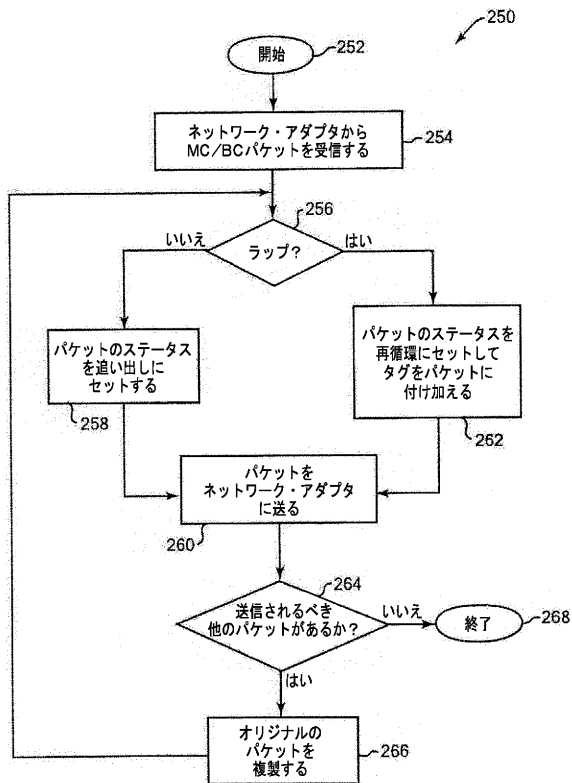
【図 3】



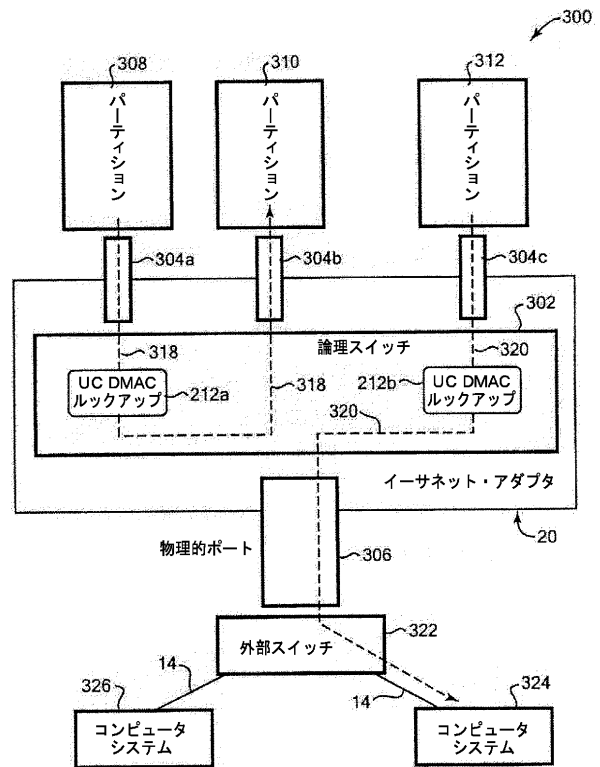
【図 4】



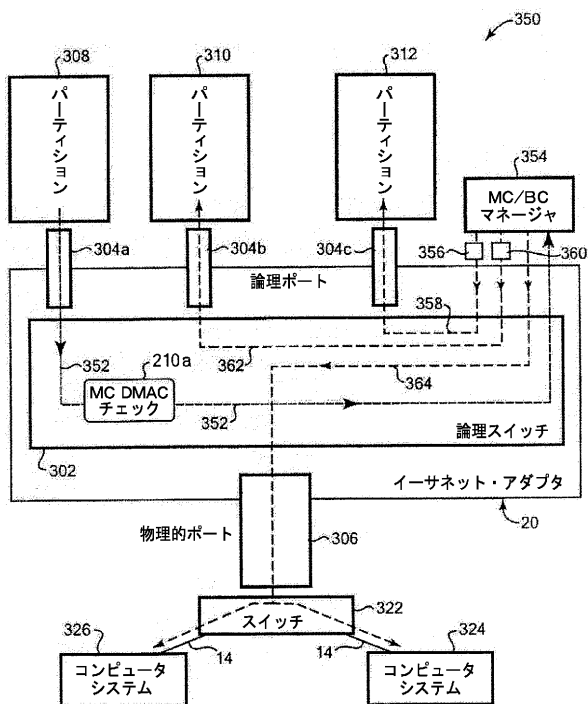
【図 5】



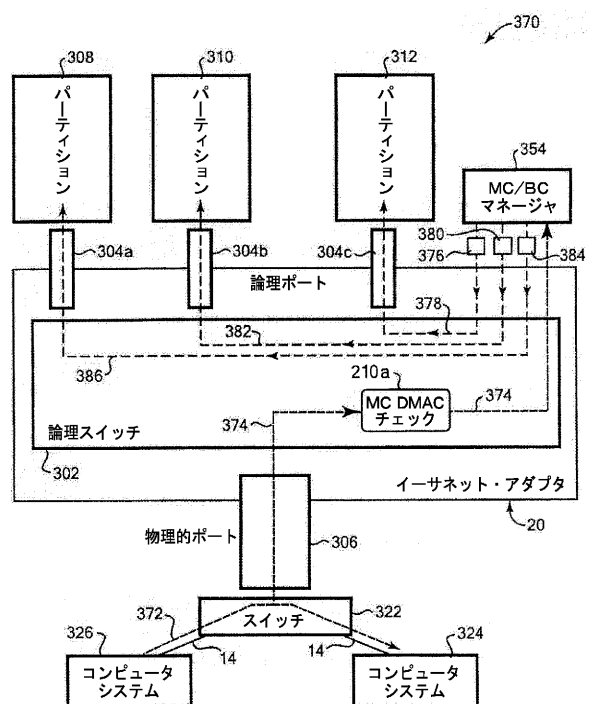
【図 6】



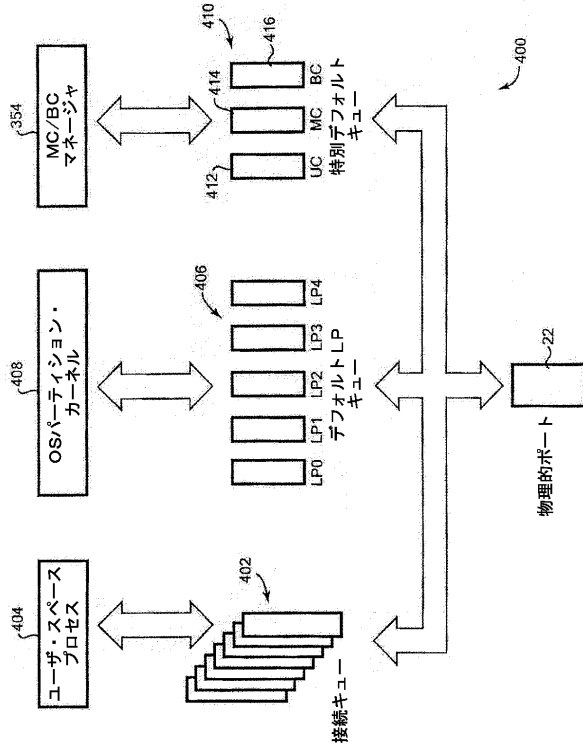
【図 7】



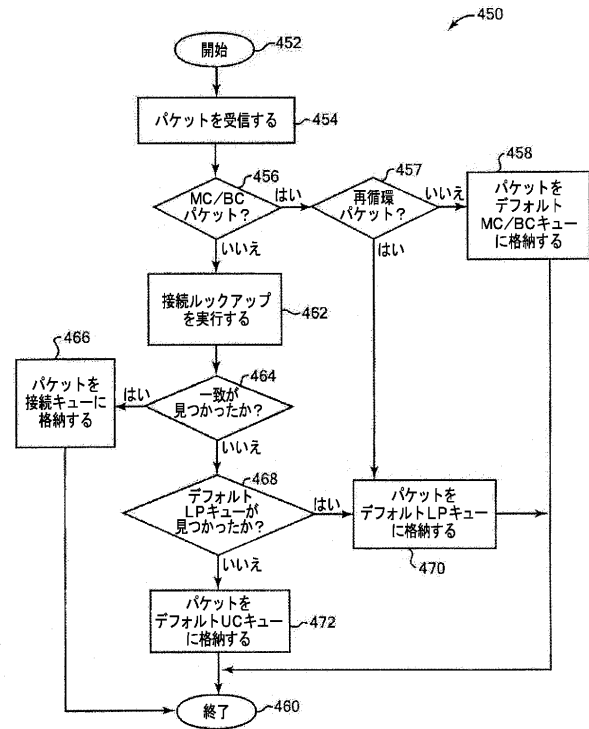
【図 8】



【図 9】



【図 10】



フロントページの続き

- (74)代理人 100086243
弁理士 坂口 博
- (72)発明者 バッソ、クロード
アメリカ合衆国 2 7 6 1 3 ノースカロライナ州 ラーレイ パーシー・コート 7 6 0 4
- (72)発明者 カルヴィニャック、ジーン
アメリカ合衆国 2 7 6 0 6 ノースカロライナ州 ラーレイ アトキンス・ファーム・コート 6
0 0 1
- (72)発明者 チャン、チーチェン
アメリカ合衆国 2 7 5 2 3 ノースカロライナ州 アベックス ウォルデン・メドウ・ドライブ
1 8 0 7
- (72)発明者 デイモン、フィリップ
アメリカ合衆国 2 7 5 1 6 ノースカロライナ州 チャペル・ヒル マウンテン・クリーク・ロー
ド 8 1 0
- (72)発明者 フース、ロナルド、エドワード
アメリカ合衆国 5 5 9 0 1 ミネソタ州 ロチェスター シャレー・ドライブ エヌ・ダブリュ
. 9 6 0
- (72)発明者 キール、ハーヴィー、ジーン
アメリカ合衆国 5 5 9 0 6 ミネソタ州 ロチェスター バックリッジ・ドライブ エヌ・イー
. 1 2 6 8
- (72)発明者 ヴァイドヤナサン、ナタラヤン
アメリカ合衆国 2 7 5 1 0 ノースカロライナ州 カルボロ メイナー・リッジ・ドライブ 2 0
9
- (72)発明者 ヴァーブランケン、ファブリス
フランス共和国 エフ - 0 6 6 1 0 ラ・ゴード ルート・ド・カーニュ 9 1 5 2
- (72)発明者 ヴェリーリ、コリン、ビートン
アメリカ合衆国 2 7 5 0 2 ノースカロライナ州 アベックス ヘイゼルハースト・サークル 7
5 5 3
- (72)発明者 リュッケ、アラン、カイル
アメリカ合衆国 5 5 9 0 6 ミネソタ州 ロチェスター キャシディ・ドライブ エヌ・イー
3 0 2 3

審査官 玉木 宏治

- (56)参考文献 米国特許出願公開第 2 0 0 4 / 0 2 0 2 1 8 9 (U S , A 1)
特開平 0 6 - 0 3 5 7 2 5 (J P , A)
国際公開第 2 0 0 5 / 0 1 8 1 7 5 (W O , A 1)
特開平 0 9 - 0 0 8 8 4 4 (J P , A)

(58)調査した分野(Int.Cl. , D B 名)

H04L 12/00-66

G06F 13/00