

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5012526号
(P5012526)

(45) 発行日 平成24年8月29日(2012.8.29)

(24) 登録日 平成24年6月15日(2012.6.15)

(51) Int.Cl. F I
H O 4 L 12/56 (2006.01) H O 4 L 12/56 Z

請求項の数 5 (全 10 頁)

<p>(21) 出願番号 特願2008-8254 (P2008-8254) (22) 出願日 平成20年1月17日(2008.1.17) (65) 公開番号 特開2009-171325 (P2009-171325A) (43) 公開日 平成21年7月30日(2009.7.30) 審査請求日 平成22年12月3日(2010.12.3)</p>	<p>(73) 特許権者 000004237 日本電気株式会社 東京都港区芝五丁目7番1号 (74) 代理人 100080816 弁理士 加藤 朝道 (72) 発明者 ピラウォン ミナサイ 東京都港区芝五丁目7番1号 日本電気株式会社内 審査官 永井 啓司</p>
---	---

最終頁に続く

(54) 【発明の名称】 パケット転送装置及び転送方法

(57) 【特許請求の範囲】

【請求項1】

キャッシュと転送部と制御部とを備えたパケット転送装置であって、
 前記キャッシュはフォワーディング情報を記憶し、
 前記転送部は前記フォワーディング情報を参照して受信したパケットの転送先の決定を試み、該パケットの転送先を決定することができたか否かに関わらず該パケットを前記制御部へ送信し、該パケットの転送先を決定することができなかつた場合には前記制御部に該パケットの転送先の決定を依頼し、
 前記制御部は前記転送部からパケットを受信し、前記転送部から該パケットの転送先の決定を依頼された場合には該パケットの転送先を決定し、該転送先へ該パケットを送信した後に該転送先をフォワーディング情報として前記キャッシュに記憶させ、それ以外の場合には前記転送部によって決定された転送先へ該パケットを送信する、パケット転送装置。

【請求項2】

前記転送部は、パケットの転送先を決定することができたか否かを示すフラグを該パケットの内部ヘッダへ追加し、
 前記制御部は、前記フラグを参照して前記パケットの転送先を決定する必要があるか否かを判定する、請求項1に記載のパケット転送装置。

【請求項3】

請求項1又は2に記載のパケット転送装置を備えるルータ。

【請求項 4】

転送部がパケットを受信する工程と、
 前記転送部がキャッシュに記憶されたフォワーディング情報を参照して前記パケットの転送先の決定を試みる工程と、を含み、
 前記転送部が前記パケットの転送先を決定することができた場合には、前記転送部が前記パケットを制御部へ送信する工程と、
 前記制御部が前記転送部によって決定された転送先へ前記パケットを送信する工程と、を含み、
 前記転送部が前記パケットの転送先を決定することができなかつた場合には、前記転送部が前記パケットを前記制御部へ送信するとともに前記制御部に前記パケットの転送先の決定を依頼する工程と、
 前記制御部が前記パケットの転送先を決定し、該転送先へ前記パケットを送信した後に該転送先をフォワーディング情報として前記キャッシュに記憶させる工程と、を含むパケット転送方法。

10

【請求項 5】

前記転送部が前記パケットの転送先を決定することができたか否かを示すフラグを前記パケットの内部ヘッダへ追加する工程と、
 前記制御部が前記フラグを参照して該パケットの転送先を決定する必要があるか否かを判定する工程と、を含む、請求項 4 に記載のパケット転送方法。

20

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、パケット転送装置及び転送方法に関し、特に、パケットの順序を転送の前後において保持するパケット転送装置及び転送方法に関する。

【背景技術】

【0002】

パケットの転送を行うルータは、正しい宛先へパケットを転送することを目的とする。しかし、近年、ルータによって処理されるパケットには、文字や静止画等のデータのみならず、リアルタイム性が必要とされる音声や動画等のストリーミングデータが多く含まれる。ルータは、ストリーミングデータのパケットを到着した順番に処理し、転送する必要がある。パケットの順序が逆転した場合、リアルタイム性が要求されるストリーミングにおいて、パケットのロスが生じるとともに、再生エラーが発生するからである。

30

【0003】

転送部と制御部とが分離されたアーキテクチャに基づくパケット転送装置を備えたフロールータは、到着したパケットを、それぞれ個別のパケット処理時間と個別の処理経路とによって処理する。したがって、かかる処理において、パケットの順序の逆転（すなわち、パケットの追い越し）が生じうる。

【0004】

図 6 は、転送部と制御部が分離されたアーキテクチャに基づく従来のパケット転送装置のブロック図である。図 6 において、ルータ内部のパケット転送装置 110 は、転送部 111、制御部 112 及びキャッシュ 113 を備える。

40

【0005】

制御部 112 は、最初に到着した（先頭）パケットの転送先を決定するために、パケットヘッダの宛先 IP アドレスを参照して、パケットの転送先のネクストホップ及び出力インタフェースを計算し、パケットを転送する。このとき、制御部 112 は、主記憶装置（非図示）に格納されたフォワーディング情報を参照して検索を行うため、長い検索時間を必要とする。また、制御部 112 は、計算したネクストホップ及び出力インタフェースを、フォワーディング情報として、キャッシュ 113 のフォワーディングテーブルへ登録する。

【0006】

50

キャッシュ 113 は、記録用の高速メモリであって、登録されたフォワーディング情報を一時的に記憶（保持）する。

【0007】

転送部 111 は、到着したパケットのパケットヘッダの宛先 IP アドレスを参照するとともに、キャッシュ 113 に記憶されているフォワーディングテーブルを参照し、所望のフォワーディング情報が記憶されている場合には、短時間かつ高速に転送先を決定し、パケットを転送する。

【0008】

全てのパケットの転送先のフォワーディング情報を計算するのではなく、キャッシュ 113 に保存されたフォワーディング情報を利用することによって、転送部 111 は短時間でパケットの転送先を判断することができ、高速にパケットを転送することができる。

10

【0009】

なお、特許文献 1 において、パケットのフォワーディング性能を向上させつつ、装置内におけるパケット転送の負荷を軽減させたネットワーク中継装置が開示されている。

【0010】

【特許文献 1】特開 2001 - 094568 号公報

【発明の開示】

【発明が解決しようとする課題】

【0011】

以下の分析は、本発明者によってなされたものである。図 6 に示すように、転送部 111 と制御部 112 とが分離したパケット転送装置 110 では、宛先のネクストホップ及び出力インタフェースへパケットを転送する場合、次のような問題が生じる。

20

【0012】

キャッシュ 113 にフォワーディング情報がまだ登録されていない最初の（先頭）パケットが転送部 111 に到着した場合、転送部 111 はパケットヘッダの宛先 IP アドレス情報に基づいて、キャッシュ 113 に保存されているフォワーディングテーブルを参照する。しかし、最初のパケットであるため、キャッシュ 113 にフォワーディング情報（ネクストホップ及び出力インタフェース）がまだ登録されていないため、転送部 111 は制御部 112 へパケットを転送する。制御部 112 はパケットヘッダの宛先 IP アドレスを利用し、パケットの転送先のネクストホップ及び出力インタフェースを計算し、パケットを転送する。制御部 112 は計算したフォワーディング情報をキャッシュ 113 のフォワーディングテーブルに登録する。

30

【0013】

キャッシュ 113 にフォワーディング情報が登録された後においては、同じフローに含まれるパケットに関し、転送部 111 はキャッシュ 113 のフォワーディングテーブルを参照することによって転送先を決定し、転送処理を行うことができる。

【0014】

このとき、制御部 112 によって長い時間を要して転送先を検索（計算）して転送される先頭パケットと、転送部 111 によってキャッシュ 113 に登録されたフォワーディング情報に基づいて短時間で転送される後続パケットとの間では処理の経路及び処理の時間が異なる。したがって、これらのパケットの間で転送前後においてパケットの順序が逆転するという問題がある。なお、特許文献 1 に開示されたネットワーク中継装置においても、同様の問題が生じる。

40

【0015】

そこで、パケット転送の前後におけるパケットの順序を保持するパケット転送装置、方法等を提供することが課題となる。

【課題を解決するための手段】

【0016】

本発明の第 1 の視点に係るパケット転送装置は、
キャッシュと転送部と制御部とを備えたパケット転送装置であって、

50

前記キャッシュはフォワーディング情報を記憶し、

前記転送部は前記フォワーディング情報を参照して受信したパケットの転送先の決定を試み、該パケットの転送先を決定することができたか否かに関わらず該パケットを前記制御部へ送信し、該パケットの転送先を決定することができなかつた場合には前記制御部に該パケットの転送先の決定を依頼し、

前記制御部は前記転送部からパケットを受信し、前記転送部から該パケットの転送先の決定を依頼された場合には該パケットの転送先を決定し、該転送先へ該パケットを送信した後に該転送先をフォワーディング情報として前記キャッシュに記憶させ、それ以外の場合には前記転送部によって決定された転送先へ該パケットを送信する。

【0017】

本発明の第1の展開形態のパケット転送装置は、

前記転送部は、パケットの転送先を決定することができたか否かを示すフラグを該パケットの内部ヘッダへ追加するように構成され、

前記制御部は、前記フラグを参照して前記パケットの転送先を決定する必要があるか否かを判定するように構成されることが好ましい。

【0018】

本発明の第2の展開形態のルータは、上記のパケット転送装置を備えることが好ましい。

【0019】

本発明の第2の視点に係るパケット転送方法は、

転送部がパケットを受信する工程と、

前記転送部がキャッシュに記憶されたフォワーディング情報を参照して前記パケットの転送先の決定を試みる工程と、を含み、

前記転送部が前記パケットの転送先を決定することができた場合には、前記転送部が前記パケットを制御部へ送信する工程と、

前記制御部が前記転送部によって決定された転送先へ前記パケットを送信する工程と、を含み、

前記転送部が前記パケットの転送先を決定することができなかつた場合には、前記転送部が前記パケットを前記制御部へ送信するとともに前記制御部に前記パケットの転送先の決定を依頼する工程と、

前記制御部が前記パケットの転送先を決定し、該転送先へ前記パケットを送信した後に該転送先をフォワーディング情報として前記キャッシュに記憶させる工程と、を含む。

【0020】

本発明の第3の展開形態のパケット転送方法は、

前記転送部が前記パケットの転送先を決定することができたか否かを示すフラグを前記パケットの内部ヘッダへ追加する工程と、

前記制御部が前記フラグを参照して該パケットの転送先を決定する必要があるか否かを判定する工程と、を含むことが好ましい。

【発明の効果】

【0021】

本発明のパケット転送装置によって、パケット転送の前後におけるパケットの順序を保持することができる。転送部は、キャッシュに記憶されたフォワーディング情報に基づいてパケットの転送先を決定することができたか否かに関わらず、パケットを制御部に送信するため、同じフローに含まれる全てのパケットは転送部と制御部とによって逐次処理されるからである。

【0022】

さらに、本発明の転送装置によって、高速にパケットの転送を行うことができる。制御部は転送部によって、パケットに挿入されたフラグを参照して、パケットの転送先を自身が決定する必要があるか否かを判定することができるからである。

【発明を実施するための最良の形態】

10

20

30

40

50

【 0 0 2 3 】

本発明の実施形態に係るパケット転送装置について図面を参照して説明する。図 1 は、本実施形態のパケット転送装置の構成を示すブロック図である。図 1 を参照すると、パケット転送装置 1 0 は、転送部 1 1 と制御部 1 2 とキャッシュ 1 3 とを備える。

【 0 0 2 4 】

キャッシュ 1 3 はフォーディング情報を記憶する。転送部 1 1 はフォーディング情報を参照して受信したパケットの転送先の決定を試み、パケットの転送先を決定することができたか否かに関わらずパケットを制御部 1 2 へ送信し、パケットの転送先を決定することができなかった場合には制御部 1 2 にパケットの転送先の決定を依頼する。

【 0 0 2 5 】

制御部 1 2 は転送部 1 1 からパケットを受信し、転送部 1 1 からパケットの転送先の決定を依頼された場合にはパケットの転送先を決定し、その転送先へパケットを送信するとともにその転送先をフォーディング情報としてキャッシュ 1 3 へ登録し、それ以外の場合には転送部 1 1 によって決定された転送先へパケットを送信する。

【 0 0 2 6 】

また、転送部 1 1 は、パケットの転送先を決定することができたか否かを示すフラグをパケットの内部ヘッダへ追加し、制御部 1 2 は、フラグを参照してパケットの転送先を決定する必要があるか否かを判定することが好ましい。

【 0 0 2 7 】

さらに、ルータにおいて、上記パケット転送装置を備えることが好ましい。

【実施例 1】

【 0 0 2 8 】

本発明の第 1 の実施例に係るパケット転送装置について、図面を参照して説明する。図 2 はルータの内部に備えたパケット転送装置に格納されたレイヤー 3 ・フォーディングテーブルの一例を示す。パケットを転送する際、ルータの内部に備えたパケット転送装置は、宛先の IP アドレスに基づいて、ネクストホップ IP アドレス（ネクストホップ）を計算し、ネクストホップへの出力インタフェースを決定する。また、パケット転送装置は、フォーディングテーブルを検索し、検索結果のネクストホップ及び出力インタフェースへパケット送信する。

【 0 0 2 9 】

図 3 は本実施形態に係るパケット転送装置のブロック図である。パケット転送装置 2 0 は、転送部 2 1 と制御部 2 2 とキャッシュ 2 3 とを備える。パケット転送装置 2 0 は、キャッシュ 2 3 に一時的に保存しているフォーディング情報に基づいて、入力インタフェース 3 0 から入力したパケットを処理し、宛先の出力インタフェース 4 0 へ出力する。

【 0 0 3 0 】

転送部 2 1 は、全てのパケットについて、キャッシュ 2 3 に保存しているフォーディングテーブルを参照し、宛先アドレスに相当するフォーディング情報（ネクストホップ及び出力インタフェース）が含まれるか否かを判定する。

【 0 0 3 1 】

キャッシュ 2 3 にフォーディング情報が含まれない場合には、転送部 2 1 から制御部 2 2 へ転送先のネクストホップ及び出力インタフェースの計算を依頼する。制御部 2 2 は、パケットヘッダに含まれる IP アドレスに基づいて、主記憶装置 5 0 上に格納されたフォーディングテーブルを検索する。制御部 2 2 は、検索したフォーディング情報に基づいてパケットを送信する。

【 0 0 3 2 】

一方、キャッシュ 2 3 にフォーディング情報が存在する場合、転送部 2 1 は、そのフォーディング情報に基づいて、転送先のネクストホップ及び出力インタフェースを短時間で決定することができる。この場合にも、パケットは転送部 2 1 から制御部 2 2 へ転送されるものの、制御部 2 2 はパケットの転送先を計算することなく、パケットを送信する。

【 0 0 3 3 】

以上のように、本実施形態に係るパケット転送装置 2 0 においては、転送部 2 1 と制御部 2 2 とによるパイプライン処理に基づいてパケット転送が行われる。すなわち、同じフローのパケットは、順次転送先が決定されて送信されるため、パケットの順序が逆転することはない。

【 実施例 2 】

【 0 0 3 4 】

本発明の第 2 の実施例に係るパケット転送装置について図面を参照して説明する。図 3 は、ルータ内部におけるパケット転送装置 2 0 を示す。転送部 2 1 は、短時間で転送先を決定し、高速にパケット転送を行う。制御部 2 2 は、主記憶装置 5 0 に格納されたフォーワーディング情報を最長一致検索 (Longest Prefix Match、LPM) し、転送先を決定してパケットを転送する。キャッシュ 2 3 は高速メモリであって、制御部 2 2 によって検索されたフォーワーディング情報を格納する。転送部 2 1 はキャッシュ 2 3 に登録されたフォーワーディング情報を参照する。

10

【 0 0 3 5 】

図 4 は、転送部 2 1 によって作成された内部ヘッダの詳細を示す。転送部 2 1 は、内部ヘッダ “ 1 ” を追加し、制御部 2 2 へ転送する。制御部 2 2 は、内部ヘッダが “ 1 ” である場合には、宛先のネクストホップ及び出力インタフェースの計算を転送部 2 1 から依頼されたものと判断する。一方、転送部 2 1 が内部ヘッダ “ 0 ” を追加した場合には、制御部 2 2 は、宛先のネクストホップ及び出力インタフェースの計算が不要であるものと判断する。この場合には、パケットは制御部 2 2 を単に通過するのみである。

20

【 0 0 3 6 】

最初のパケットが入力された場合、本実施例のパケット転送装置 2 0 における転送部 2 1 は、ネクストホップへの出力インタフェースを決定するため、パケットヘッダの宛先の IP アドレスと、キャッシュ 2 3 に格納されたフォーワーディングテーブルと、を参照する。しかし、最初のパケットであるため、キャッシュ 2 3 には、フォーワーディング情報がまだ保存されていない。転送部 2 1 は、パケットの転送の宛先が未決定であるため、パケットの内部ヘッダに “ 1 ” を追加し、制御部 2 2 に対して宛先のネクストホップ及び出力インタフェースの計算を依頼する。

【 0 0 3 7 】

制御部 2 2 は、内部ヘッダにおける識別子 “ 1 ” を参照し、自身が処理を依頼されたものと判断する。このとき、制御部 2 2 は、パケットヘッダに含まれる宛先の IP アドレスに基づいて、パケットの転送先を計算し、宛先のネクストホップ及び出力インタフェースを決定することによって、パケットを転送する。さらに、制御部 2 2 は、計算したネクストホップ及び出力インタフェースをフォーワーディング情報としてキャッシュ 2 3 へ登録する。

30

【 0 0 3 8 】

フォーワーディング情報の登録後において、同じフローに含まれる後続のパケットを転送する場合には、転送部 2 1 は、キャッシュ 2 3 に登録されたフォーワーディング情報を検索し、検索がヒットしたネクストホップ及び出力インタフェースに基づいてパケットを転送する。また、フォーワーディング情報の登録後において、転送部 2 1 は、内部ヘッダへ識別子 “ 0 ” を追加し、制御部 2 2 へ転送する。制御部 2 2 は、内部ヘッダの識別子 “ 0 ” を参照し、転送先の計算が必要でないものと判断し、パケットを通過させる。

40

【 0 0 3 9 】

このように、本実施例においては、フロー内に含まれるパケット毎に、転送部 2 1 と制御部 2 2 とによってパイプライン処理が行われるため、パケットの順序を維持することができる。また、最初に到着した先頭のパケットのみを制御部 2 2 によって処理し、その後到着した後続のパケットは転送部 2 1 によって高速に処理するため、パケットの順序を維持しつつ、高速なパケット転送を実現することができる。

【 0 0 4 0 】

50

次に、本実施例のパケット転送装置の動作について図面を参照して説明する。図5は、本実施例のパケット転送装置の動作のフローチャートである。

【0041】

図3及び図5を参照すると、パケット転送装置20の転送部21は、入力インタフェース30からパケットデータを受信する(ステップS11)。転送部21は、パケットのレイヤー3ヘッダ(又はレイヤー2ヘッダ)の情報を抽出する(ステップS12)。転送部21は、レイヤー3ヘッダに含まれる宛先IPアドレス(又はレイヤー2ヘッダに含まれる宛先MACアドレス)の情報と、キャッシュ23に一時保存しているフォワーディングテーブルとを参照し、転送先のネクストホップ及び出力インタフェースを決定する。

【0042】

キャッシュ23のフォワーディングテーブルにパケットの転送先の情報が未登録である場合、キャッシュ23はヒットしないものと判定される(ステップS13のNo)。この場合、転送部21は、パケットの内部ヘッダへ“1”を追加し(ステップS14)、制御部22へのネクストホップ及び出力インタフェースの計算を依頼する。制御部22は、レイヤー3の宛先IPアドレス情報を利用し、ネクストホップ及び出力インタフェースを計算し(すなわちL3ヘッダの解析を行い)(ステップS17)、出力インタフェースへパケットを転送する(ステップS18)。また、制御部22は、計算したフォワーディング情報をキャッシュ23へ登録する。

【0043】

登録後において、転送部21が同じフローに含まれる後続のパケット群に対してキャッシュ23のフォワーディング情報を参照した場合、キャッシュ23に登録されているフォワーディング情報にヒットする(ステップS13のYes)。したがって、転送部21は、ヒットしたフォワーディング情報に基づいて、後続のパケット群に対するネクストホップ及び出力インタフェースを決定することができる。キャッシュ23に登録されたフォワーディング情報を利用することができる場合、転送部21は、パケットの内部ヘッダへ“0”を追加し(ステップS15)、制御部22へ転送する。制御部22は、内部ヘッダにおける“0”の識別子を参照し、パケットの転送先の計算を行うことなく(すなわちL3ヘッダの解析を行うことなく)(ステップS16)、転送部21によるフォワーディング情報を利用し、そのネクストホップ及び出力インタフェースの情報に基づいてパケットを送信する(ステップS18)。

【0044】

以上の記載は実施例に基づいて行ったが、本発明は、上記実施例に限定されるものではない。

【産業上の利用可能性】

【0045】

パケットの順序を維持する必要があるストリームデータなどのパケットの転送処理を司るルータにおいて、本発明のパケット転送装置を用いることができる。

【図面の簡単な説明】

【0046】

【図1】本発明の実施形態に係るパケット転送装置のブロック図である。

【図2】キャッシュに格納されたフォワーディングテーブルを示す。

【図3】本発明の第1及び第2の実施例に係るパケット転送装置のブロック図である。

【図4】本発明の第2の実施例におけるパケットのデータフォーマットを示す。

【図5】本発明の第2の実施例に係るパケット転送装置の動作を示すフローチャートである。

【図6】従来のパケット転送装置のブロック図である。

【符号の説明】

【0047】

10、20、110 パケット転送装置

11、21、111 転送部

10

20

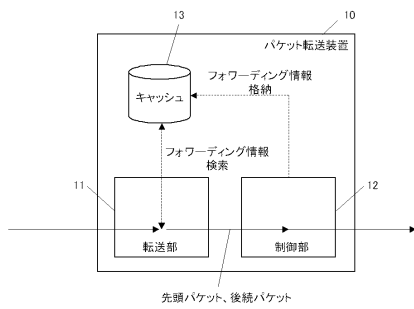
30

40

50

- 1 2、2 2、1 1 2 制御部
- 1 3、2 3、1 1 3 キャッシュ
- 3 0 入力インタフェース
- 4 0 出力インタフェース
- 5 0 主記憶装置

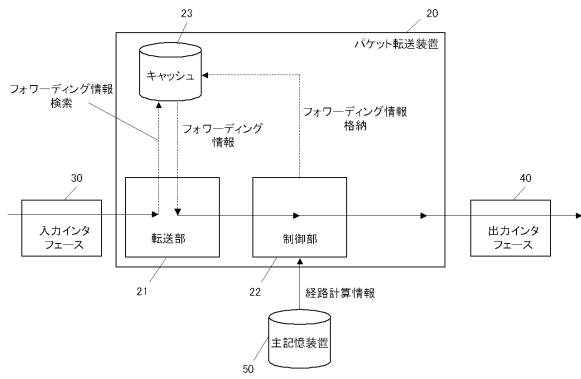
【図 1】



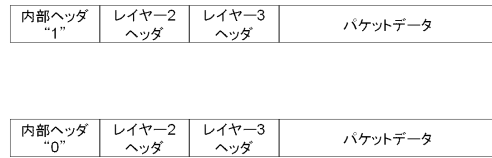
【図 2】

宛先アドレスプレフィックス	ネクストホップ	出力インタフェース
122.82.10/24	185.99.65.123	2
119.110/16	123.11.22.33	3
110.99.88/24	110.77.66.55	1

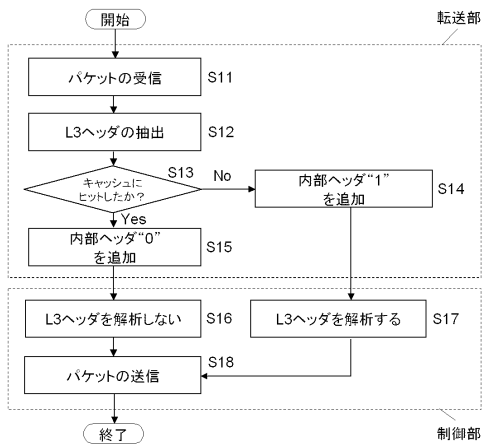
【図3】



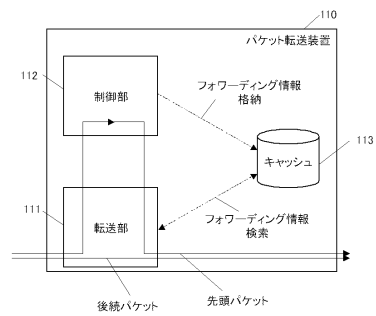
【図4】



【図5】



【図6】



フロントページの続き

(56)参考文献 特開平09 - 036908 (JP, A)
特開2000 - 261477 (JP, A)
特開2007 - 221514 (JP, A)

(58)調査した分野(Int.Cl., DB名)

H04L 12/00 - 12/26、12/50 - 12/66