

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2006-108954
(P2006-108954A)

(43) 公開日 平成18年4月20日(2006.4.20)

(51) Int. Cl. F I テーマコード (参考)
H O 4 L 12/56 (2006.01) H O 4 L 12/56 1 O O Z 5 K O 3 O

審査請求 未請求 請求項の数 12 O L (全 21 頁)

(21) 出願番号 特願2004-290892 (P2004-290892)
(22) 出願日 平成16年10月4日 (2004.10.4)

(特許庁注：以下のものは登録商標)
1. E T H E R N E T

(71) 出願人 000005108
株式会社日立製作所
東京都千代田区丸の内一丁目6番6号
(74) 代理人 100075096
弁理士 作田 康夫
(74) 代理人 100100310
弁理士 井上 学
(72) 発明者 赤羽 真一
東京都国分寺市東恋ヶ窪一丁目280番地
株式会社日立製作所中央研究所内
(72) 発明者 須貝 和雄
神奈川県秦野市堀山下1番地 株式会社日立製作所 I P ネットワーク事業部内
Fターム(参考) 5K030 GA12 HA08 HD03 HD09 JA10
KA05 MD02

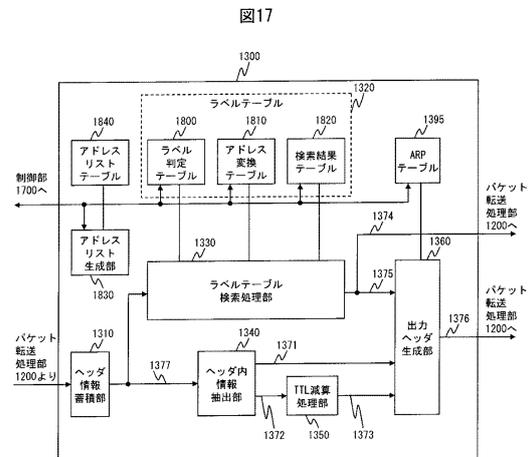
(54) 【発明の名称】 検索テーブル高速切替方式およびパケット転送装置

(57) 【要約】

【課題】 IPおよびMPLSネットワーク内のルータが収容する回線に障害が発生した場合に、高速に、出力先の回線を迂回路用の回線に切替え、パケットの出力ヘッダ情報を迂回用のヘッダ情報に切替えることを可能にすること。

【解決手段】 2つの出力回線を収容するパケット転送装置において、複数の入力ヘッダ情報毎の出力情報を記憶する手段と、入力ヘッダ情報に対応付けられた障害時の迂回路となるべき出力回線の出力情報を記憶する手段と、出力回線の障害時に、迂回路となるべき出力回線に対応付けられた出力情報のアドレスを書き込む出力情報アドレス書替処理部を有する。

【選択図】 図17



【特許請求の範囲】**【請求項 1】**

第一及び第二の出力回線とメモリと上記第一の出力回線または上記第二の出力回線のうち、使用すべき回線を判定する判定部を有するパケット転送装置であって、

上記メモリには、

受信パケットのヘッダ情報の少なくとも一部に対応付けられた、上記第一の出力回線の識別子、第一の新たなヘッダ情報、上記第二の出力回線の識別子、および第二の新たなヘッダ情報と、

上記第一の出力回線および上記第一の新たなヘッダ情報の記憶位置を示す第一のアドレスと、

上記第二の出力回線および上記第二の新たなヘッダ情報の記憶位置を示す第二のアドレスと、が保持されており、

上記判定部で上記第一の出力回線を使用すべきと判定された場合には、上記受信パケットのヘッダ情報を、上記第一のアドレスに基づいて上記メモリから読み出された上記第一の新たなヘッダ情報に書き換えて、該受信パケットを上記第一の出力回線から送信し、

上記判定部で上記第二の出力回線を使用すべきと判定された場合には、上記受信パケットのヘッダ情報を、上記第二のアドレスに基づいて上記メモリから読み出された上記第二の新たなヘッダ情報に書き換えて、該受信パケットを上記第二の出力回線から送信するパケット転送装置。

【請求項 2】

請求項 1 に記載のパケット転送装置であって、

上記受信パケットのヘッダ情報は、該受信パケットが使用する論理的な回線を識別するラベルを含むことを特徴とするパケット転送装置。

【請求項 3】

請求項 2 に記載のパケット転送装置であって、

上記受信パケットのヘッダ情報は、MPLSのShimヘッダであることを特徴とするパケット転送装置。

【請求項 4】

請求項 1 に記載のパケット転送装置であって、

上記受信パケットのヘッダ情報は、IPヘッダであることを特徴とするパケット転送装置。

【請求項 5】

請求項 1 に記載のパケット転送装置であって、

上記受信パケットのヘッダ情報は、Ethernetヘッダであることを特徴とするパケット転送装置。

【請求項 6】

請求項 1 乃至 5 のいずれかに記載のパケット転送装置であって、

上記メモリにはさらに、

上記受信パケットのヘッダ情報の少なくとも一部に対応付けられた第一及び第二の転送処理優先度と、

上記第一の転送処理優先度の記憶位置を示す第三のアドレスと、

上記第二の転送処理優先度の記憶位置を示す第四のアドレスと、が保持されており、

上記判定部で上記第一の出力回線を使用すべきと判定された場合には、上記第三のアドレスに基づいて上記メモリから読み出された上記第一の転送処理優先度に従って上記受信パケットを転送処理し、

上記判定部で上記第二の出力回線を使用すべきと判定された場合には、上記第四のアドレスに基づいて上記メモリから読み出された上記第二の転送処理優先度に従って上記受信パケットを転送処理することを特徴とするパケット転送装置。

【請求項 7】

第一及び第二の出力回線とメモリと上記第一の出力回線または上記第二の出力回線のうち、使用すべき回線を判定する判定部を有するパケット転送装置であって、

10

20

30

40

50

上記メモリには、

受信パケットのヘッダ情報の少なくとも一部に対応付けられた、上記第一の出力回線の識別子、第一の新たなヘッダ情報、上記第二の出力回線の識別子、および第二の新たなヘッダ情報が保持されており、

上記判定部で上記第一の出力回線を使用すべきと判定された場合には、上記受信パケットのヘッダ情報を、上記第一の新たなヘッダ情報に書き換えて、該受信パケットを上記第一の出力回線から送信し、

上記判定部で上記第二の出力回線を使用すべきと判定された場合には、上記受信パケットのヘッダ情報を、上記第二の新たなヘッダ情報に書き換えて、該受信パケットを上記第二の出力回線から送信するパケット転送装置。

10

【請求項 8】

請求項 7 に記載のパケット転送装置であって、

上記受信パケットのヘッダ情報は、該受信パケットが使用する論理的な回線を識別するラベルを含むことを特徴とするパケット転送装置。

【請求項 9】

請求項 8 に記載のパケット転送装置であって、

上記受信パケットのヘッダ情報は、MPLS の Shim ヘッダであることを特徴とするパケット転送装置。

【請求項 10】

請求項 7 に記載のパケット転送装置であって、

上記受信パケットのヘッダ情報は、IP ヘッダであることを特徴とするパケット転送装置。

20

【請求項 11】

請求項 7 に記載のパケット転送装置であって、

上記受信パケットのヘッダ情報は、Ethernet ヘッダであることを特徴とするパケット転送装置。

【請求項 12】

請求項 7 乃至 11 のいずれかに記載のパケット転送装置であって、

上記メモリにはさらに、

上記受信パケットのヘッダ情報の少なくとも一部に対応付けられた第一及び第二の転送処理優先度が保持されており、

30

上記判定部で上記第一の出力回線を使用すべきと判定された場合には、上記第一の転送処理優先度に従って上記受信パケットを転送処理し、

上記判定部で上記第二の出力回線を使用すべきと判定された場合には、上記第二の転送処理優先度に従って上記受信パケットを転送処理することを特徴とするパケット転送装置。

【発明の詳細な説明】

【技術分野】

【0001】

本発明はパケット転送方法およびその装置に関し、検索テーブルを高速に切替える方式およびその装置に関する。

【背景技術】

40

【0002】

Internet Protocol (IP) ネットワークはその低コスト性から、電話網に代わる重要な社会インフラの一つになりつつある。それに伴い、従来の電話網で実現されていたネットワーク障害発生時における高速復旧等の高信頼性の実現、低遅延時間や低廃棄率等の通信品質(QoS: Quality of Service)の保証が要求されている。

ネットワーク障害の回復方式としては、例えば、Multi Protocol Label Switching(MPLS)を用いた高速パス切替え技術がある(従来技術1)(MPLSに関しては非特許文献1、2参照。MPLSを用いた高速パス切替え技術に関しては非特許文献3参照。)

まず、MPLSの概要について説明する。MPLSは、ルータがパケットに付与されたラベルの値に基づきパケットの転送先を決定するパケット転送方式である。

50

図1を用いてMPLSを用いたパケット転送の概要を説明する。

図1のMPLSネットワーク100は、ネットワークの出入り口に位置するエッジルータER1(141)、ER2(142)、ER3(143)、ER4(144)、MPLSネットワーク100内に位置するコアルータCR1(151)、CR2(152)、CR3(153)から構成されている。MPLSネットワーク100は、第2層の転送プロトコル(例えばEthernet)等を用いて構成される。(Ethernetは登録商標)以下では、第2層の転送プロトコルとしてEthernetを用いる場合を例にして説明する。

【0003】

各エッジルータはMPLSネットワークを介して相互接続する複数のネットワークNA(161)、NB(162)、NC(163)、ND(164)を収容する。各エッジルータは、前記収容するネットワークからパケットを受信し、受信パケットをMPLSカプセルヘッダでカプセル化し、MPLSネットワーク100内に転送する。図1では、表記を簡単にするため、前記MPLSカプセルヘッダの内容のうち、ラベルのみを図示している。例えば、エッジルータER1はネットワークNAから受信したパケットP1(171)をラベルL11(111)を含むMPLSカプセルヘッダでカプセル化し、CR1へ転送する。この際、出力回線181、ラベルL11等のMPLSカプセルヘッダは、パケットP1の第3層の転送プロトコル(例えばIP)等のヘッダ内の情報を用いて決定される。以下では、第3層の転送プロトコルとしてIPを用いる場合を例にして説明する。

10

【0004】

コアルータCR1、CR2は、MPLSカプセルヘッダを用いて出力回線と出力用の新しいカプセルヘッダを決定する。例えば、パケットP1を受信したCR1は、パケットP1に付与されているラベルL11の値を用いてCR1内に保持しているラベルテーブルを検索し、出力回線1020-1、出力用のラベルL12(112)、次に転送すべきルータCR2の回線1020-1のインターフェースに割り当てられている次転送先IPアドレス(以下、Next Hop IP)であるIP152-1を決定する。CR1は、パケットP1に付与されていたラベルL11を出力用のラベルL12に置換える。また、CR1は、Next Hop IP(IP152-1)からCR2の回線1020-1のインターフェースに付与されているMedia Access Control(以下、MAC)アドレス(MAC152-1)を決定してEthernetフレームの宛先MACアドレスとして付与する。その後、CR1はパケットP1を出力回線1020-1へ送信する。以下、受信パケットに付与されているラベルを入力ラベル、送信時にルータが新たに付与するラベルを出力ラベルと呼ぶ。また、CR2も同様にして、入力ラベルL12が付与されたパケットP1の出力回線182、出力ラベルL13(113)、Next Hop IP(IP143)を決定し、入力ラベルL12を出力ラベルL13で置き換え、Next Hop IP(IP143)から宛先MACアドレス(MAC143)を決定してEthernetフレームに付与し、出力回線182へパケットを送信する。

20

30

【0005】

エッジルータER3は、ラベルL13が付与されたパケットP1を受信すると、ラベルL13のみ、あるいはラベルL13とパケットP1の第3層の転送プロトコル(本例ではIP)ヘッダ内の情報を用いて、転送先(図1ではネットワークNC)と、出力回線183と、Next Hop IPを決定し、MPLSカプセルヘッダを除去して回線183へパケットP1を送信する。パケットP2の転送も前記と同様に行われる。

【0006】

上述のように、パケットP1は、ルータER1とCR1を結ぶ経路、CR1とCR2とを結ぶ経路、およびCR2とER3とを結ぶ経路をラベルL11、L12、L13を用いて転送される。

40

また、パケットP2は、ルータER2とCR1とを結ぶ経路、CR1とCR2とを結ぶ経路、およびCR2とER4とを結ぶ経路をラベルL21、L22、L23を用いて転送される。これらの経路をLabel Switched Path(LSP)と呼ぶ。LSPは、ラベル分配プロトコル、あるいはネットワーク管理者の手動設定により設定される。ラベル分配プロトコルにはLabel Distribution Protocol(LDP)、Constraint-based Routing Label Distribution Protocol(CR-LDP)、Extensions to Resource Reservation Protocol for LSP Tunnels(RSVP-TE)等がある(LDPに関しては非特許文献4参照。CR-LDPに関しては非特許文献5参照。RSVP-TEに関しては非特許文献6参照。)

【0007】

MPLSでは、LSPを明示的に設定することが可能であるため、トラフィック毎の要求通信

50

品質に基づいた経路の提供や、ネットワーク内の経路の負荷分散を行うトラフィックエンジニアリングが可能になるという利点がある。

【 0 0 0 8 】

次に、図2、図3を用いて、非特許文献3に記載されているMPLSを用いた高速パス切替え技術のうち、Facility Backupと呼ばれる技術について説明する。

この高速パス切替え技術では、同一回線上に設定されている複数のLSPに対し、前記回線に障害が発生した場合の切替え用LSPを1つだけ設定する。以下、障害時の切替え用LSPを予備LSPと呼び、正常時のLSPを現用LSPと呼ぶ。また、その高速パス切替え技術では、予備LSPの終点ルータでの出力ラベルと出力回線を決定するためのラベルテーブルのエントリ数を節約するため、2つ以上のラベルを付与するラベルスタック技術が用いられる。

10

【 0 0 0 9 】

図2、図3は図1のMPLSネットワーク100のみを抜粋したものであり、ネットワークNA、NB、NC、NDは省略されている。

図2は、回線1020-1が正常な場合の packets 転送動作を示す。また、図3は、回線1020-1に障害が発生した場合(図3の回線1020-1上に「X」と表す)の packets 転送動作を示す。

【 0 0 1 0 】

図2において、回線1020-1が正常な場合に packets P1が転送されるLSPをLSP1、 packets P2が転送されるLSPをLSP2とする。転送に用いられるラベルは図1と同様である。LSP1で用いられるラベルは、ER1-CR1間のラベルL11(111)、CR1-CR2間のラベルL12(112)、CR2-ER3間のラベルL13(113)となる。また、LSP2で用いられるラベルは、ER2-CR1間のラベルL21(121)、CR1-CR2間のラベルL22(122)、CR2-ER4間のラベルL23(123)となる。

20

【 0 0 1 1 】

図3では、回線1020-1上に設定されているLSP1用およびLSP2用の予備LSP3が、CR1-CR3間の回線1020-2上とCR3-CR2間の回線184上とに設定されている。CR1-CR3間で使用するラベルをL31(131)、CR3-CR2間で使用するラベルをL32(132)とする。

【 0 0 1 2 】

図3を用いて回線1020-1に障害が発生した場合の packets P1の転送方法を説明する。図3のコアルータCR1は、回線1020-1に障害が発生したことを検知すると、ラベルテーブル内の入力ラベルL11に対する出力ラベルをL12一つから、L12、L31の二つに変更する。また、出力回線を回線1020-1から回線1020-2に変更する。さらに、Next Hop IPを、CR2のIP152-1からCR3のIP153に変更する。 packets P1を受信したルータCR1は、入力ラベルL11から、出力ラベルL12、L31と、出力回線1020-2と、Next Hop IP IP153を決定する。その後、 packets P1から入力ラベルL11を除去し、ラベルスタック機能を用いて前記2つの出力ラベルを付与する。この際、予備LSPで用いるラベルL31を、現用LSP1で用いるラベルL12の上位(データとして先に転送される位置)に付与する。

30

【 0 0 1 3 】

また、Next Hop IP (IP153)からCR3の回線1020-2のインターフェースに付与されているMACアドレス(MAC153)を決定してEthernetフレームの宛先MACアドレスとして付与する。その後、CR1は packets P1を回線1020-2へ送信する。前記 packets P1を受信したCR3は、 packets P1に付与されている2つのラベルのうち、上位に付与されているラベルL31のみを用いて、出力回線184と、出力ラベルL32と、Next Hop IP IP152-2を決定する。CR3は、 packets P1の入力ラベルのうち上位に付与されているラベルL31のみを出力L32で置換え、Next Hop IP(IP152-2)からCR2の回線184のインターフェースに付与されているMACアドレス(MAC152-2)を決定してEthernetフレームの宛先アドレスとして付与する。その後、CR3は packets P1を回線184へ送信する。CR3から送信された packets P1を受信したCR2は、 packets P1に付与されている2つのラベルのうち、上位に付与されているラベルL32の値から、自分が予備LSP3の終点であることを認識し、ラベルL32の下位に付与されているラベルL12を用いて出力回線と出力ラベルを決定すべきであることを判断する。

40

【 0 0 1 4 】

その後、CR2はラベルL12を用いて出力回線182、出力用のラベルL13、Next Hop IP IP14

50

3を決定する。CR2は、パケットP1に付与されているラベルL32を除去し、ラベルL12を出力ラベルL13で置換える。さらに、Next Hop IP (IP143)からER3の回線182のインターフェースに付与されているMACアドレス(MAC143)を決定してEthernetフレームの宛先アドレスとして付与する。その後、CR2はパケットP1を回線182へ送信する。

【0015】

パケットP2の転送方法も、パケットP1の場合と同様である。ただし、CR1で付与される2つの出力ラベルのうち、上位に付与されるラベルL31はP1の場合と同一だが、もう一方のラベルがL22である点が異なる。また、CR2での出力ラベルが、ラベルL23である点が異なる。以上説明したように、現用LSP1で転送されていたパケットP1も、現用LSP2で転送されていたパケットP2も、LSP3では同様に転送される。

10

【0016】

また、図3において、予備LSPの終点であるCR2は、上位のラベルL32を用いて、自分が予備LSP3の終点であり、L32の下位に付与されているラベル(L12あるいはL22)を用いて出力回線と出力ラベルを決定すべきであると判定している。この場合、CR2では、ラベルテーブルの検索方式によってはラベルテーブルを2度検索する必要があり、処理時間がかかる。この問題を回避する方式として、予備LSPの終点から一つ手前のルータで、予め上位に付与されたラベルを除去する方式(Penultimate Hop Popping; 以下、PHP方式)がある。図4を用いて本方式を説明する。

【0017】

図4において、ラベルの付与方式、パケットの転送方式は図3で説明した方式とほぼ同様である。ただし、予備LSP3の終点ルータCR2の一つ手前のルータCR3でのラベル付与方式が異なる。CR2は、受信パケットP1、P2の上位に付与されているラベルL31を用いて、自分が予備LSP3の終点から一つ手前のルータであることを認識する。ラベルL31の下位に付与されているラベルL12、L22は、予備LSP3の終点ルータCR2においては現用LSPの転送に用いられているラベルと同一ラベルである。したがって、CR2はこのラベルを使用すれば現用LSPと同一の出力回線、出力ラベルを決定することができる。このため、CR3は上位に付与されているラベルL31の除去のみを行い、パケットを出力回線184へ送信する。前記パケットを受信するルータCR2は、一つだけ付与されているラベルL12、L22を用いて、出力回線、出力ラベルを決定する。以上説明したPHP方式を用いることにより、CR2におけるラベルテーブルの検索は一回行うだけでよく、処理時間は現用LSPを用いた転送処理と同様になる。

20

30

【0018】

また、MPLSではなくAsynchronous Transfer Mode (以下、ATM)技術を用いてネットワークを構成する際、データ転送を行うATM交換機において障害時の高速切替を行う技術がある(従来技術2)(特許文献1参照)。従来技術2は、ATM交換機に予め通常経路の変換用ヘッダ情報を格納する現用ルーティングテーブルと、予備経路の変換用ヘッダ情報を格納する予備用ルーティングテーブルを所持する。さらに、障害が発生した回線、および障害が発生した仮想回線識別子(Virtual Path Identifier、以下VPI)をビットマップとして記憶するメモリを所持する。ATM交換機は、データを53Byteのセルという単位で処理する。従来技術2のATM交換機は、セル到着時に、前記障害回線および障害VPIが記憶されているメモリを参照して、出力回線の障害の有無を検査する。出力回線に障害が無い場合は、現用ルーティングテーブルに格納されているヘッダ情報をセルのヘッダ情報変換に用いる。出力回線に障害が有る場合は、予備用ルーティングテーブルに格納されているヘッダ情報をセルのヘッダ情報変換に用いる。ヘッダ情報変換後、従来技術2のATM交換機はセルを次に転送すべきATM交換機へ送信する。

40

【0019】

【特許文献1】特開平8-24220号公報

【0020】

【非特許文献1】E. Rosen, A. Viswanathan, R. Callon, "Multiprotocol Label Switching Architecture", RFC3031, Internet Engineering Task Force, January, 2001

50

【非特許文献2】E. Rosen, T. Tappan, G. Fedorkow, Y. Rekhter, D. Farinacci, T. Li, A. Conta, "MPLS Label Stack Encoding", RFC3032, Internet Engineering Task Force, January, 2001

【非特許文献3】P. Pan, G. Swallow, A. Atlas, "Fast Reroute Extensions to RSVP-TE for LSP Tunnels", draft-ietf-mpls-rsvp-lsp-fastreroute-04.txt, Internet draft, Internet Engineering Task Force, 2004

【非特許文献4】L. Andersson, "LDP Specification", RFC3036, Internet Engineering Task Force, January 2001

【非特許文献5】B. Jamoussi, "Constraint-Based LSP Setup using LDP", RFC3212, Internet Engineering Task Force, January 2002

【非特許文献6】D. Awduche, L. Berger, T. Li, V. Srinivasan, G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC3209, Internet Engineering Task Force, December 2001

10

【発明の開示】

【発明が解決しようとする課題】

【0021】

前記の従来技術1では、同一回線内に設定されている複数のLSPに対し、一つの共通LSPを設定するため、複数LSPに対し一対一に予備LSPを設定する方式に比べて、MPLSネットワーク内のルータが管理するLSP数を少なくすることができる。たとえば、ある回線内に設定されているLSP数をNとすると、従来技術1では、(N+1)のLSPを管理すればよいのに比べ、一対一に予備LSPを設定する場合は、2NのLSPを管理する必要がある。

20

【0022】

しかし、前記の従来技術1では、ルータにおいて現用LSPから予備LSPへ切替える際の具体的な方法に関しては記述されていない。以下で、ルータにおいて現用LSPから予備LSPへ切替える際の課題について説明する。

【0023】

図5に、図2で示した、回線1020-1が正常な場合のCR1のラベルテーブルの内容の1例を示す。また、図6に、図3で示した、回線1020-1に障害が発生した場合のCR1のラベルテーブルの内容の一例を示す。ラベルテーブルは、入力ラベルに対応する出力ラベル、出力回線、Next Hop IPの組合せ(以下、ラベルテーブルエントリと呼ぶ)が複数個設定される。図5および図6では、ラベルテーブルエントリはLE1、LE2の2つ設定されている。

30

図2では、回線1020-1上に設定されるLSPは、パケットに付与されるラベルが、ラベルL11、ラベルL12、ラベルL13と変換されるLSP(以下、LSP1と呼ぶ)と、ラベルL21、ラベルL22、ラベルL23と変換されるLSP(以下、LSP2と呼ぶ)の2つである。この際、CR1が保持する、入力ラベルから出力ラベル、出力回線、Next Hop IPを決定するためのラベルテーブルのエントリは、図5に示す通り、ラベルL11用エントリ(LE1)とラベルL12用エントリ(LE2)の2エントリである。

【0024】

回線1020-1に障害が発生した場合、図5のLE1、LE2の出力ラベル、出力回線、Next Hop IPを、図6のLE1、LE2の通り書替える必要がある。図5、図6のLE1を例にして、書替え処理を具体的に説明する。回線1020-1に障害が発生した場合、CR1は、図5のLE1の出力ラベルを、L21からL21とL31に書替える。また、図5のLE1の出力回線を1020-1から1020-2に書替える。また、図5のLE1のNext Hop IPを、IP152-1からIP253に書替える。前記の書替え処理を、障害が発生した回線1020-1を出力回線として設定する全てのラベルテーブルエントリ(本例では、LE1とLE2の2エントリ)について行う必要がある。

40

【0025】

以上、図5、図6を用いて説明したように、ある回線に障害が発生した場合、ルータは、障害が発生した回線を出力回線として設定してある全てのラベルテーブルエントリの出力ラベル、出力回線、Next Hop IPを書替える必要がある。大規模ネットワークでは、同一回線上に設定されるLSP数が多くなり、ラベルテーブルエントリが多くなる。従って、回

50

線に障害が発生した場合、多数のラベルテーブルエントリの書替え時間が長くなる。書替え処理が終了していないエントリに対応するパケットは、障害が発生した回線に出力されて廃棄されるため、ラベルテーブルエントリの書替え時間が長いと、廃棄されるデータ量が多くなる。従って、障害時には、高速にラベルテーブルエントリを書き替える方式が必要になる。

【0026】

また、前記の従来技術2では、ATM交換機において、高速にルーティングテーブルを切替えることができる。しかし、今後インターネットを始めとするネットワークを構築する際に用いられるIPおよびMPLSへの対応が記載されていない。具体的には、IPおよびMPLSでは、入力ラベルに対応する情報として、出力ラベル、出力回線の他に、Next Hop IPが必要

10

【0027】

本発明の第一の目的は、ネットワーク障害時に、高速に検索テーブルを切替えることができる手段を提供することである。本発明の第二の目的は、今後のネットワーク構築の主要プロトコルであるIP、MPLSに対して、前記高速に検索テーブルを切替える手段を提供することである。

【課題を解決するための手段】**【0028】**

前記課題を解決するために、本発明のパケット転送装置は、複数の入出力回線を収容したスイッチ手段と、パケットのスイッチングに用いられる複数のラベル毎に、正常時の出力回線と障害時に迂回路となるべき出力回線とに対応付けられた複数のヘッダ情報およびヘッダ情報に対応する別の情報を記憶するヘッダ情報記憶手段とを有することを特徴とする。

20

【0029】

さらに、本発明のパケット転送装置は、収容する入出力回線毎に正常か異常かを記憶する回線状態記憶手段と、前記各入力回線から受信したパケット毎に、前記回線状態記憶手段から読み出された、前記パケットを出力すべき出力回線の状態に応じて、前記ヘッダ情報記憶手段から読み出されたヘッダ情報のうちの少なくとも一つを選択する選択手段と、前記選択手段によって選択されたヘッダ情報に基づいて入力パケットのヘッダ情報の内容を書き替えるヘッダ書替手段とを有することを特徴とする。

30

【0030】

前記ヘッダ情報記憶手段に記憶されるヘッダ情報およびヘッダ情報に対応する別の情報として、例えば、Request for Comment(RFC) 3032に記述されているShimヘッダ内のラベル、出力回線番号、次に転送すべきパケット転送装置の回線インターフェースに割り当てられるIPアドレス、などを用いてもよい。

前記のパケット転送装置は、出力回線に異常が発生した場合、前記回線状態記憶手段の対応する回線の状態を正常から異常に書き替えることにより、受信したパケットの出力回線および出力字に付与するヘッダ情報を高速に切替えることが可能になる。

【0031】

さらに、本発明のパケット転送装置は、前記複数の入力ラベル毎に有する出力情報記憶手段を識別するアドレスを前記入力ラベル毎に記憶する出力情報アドレス記憶手段と、前記入力ラベルに対応付けられた障害時の迂回路となるべき出力回線に対応付けられた出力情報が登録されている複数のアドレスと、前記入力ラベルに対応付けられた出力情報アドレス記憶手段のアドレスとを対にして、パケットの正常時の出力回線毎にリスト構造に構成して記憶する出力情報登録アドレスリスト記憶手段と、出力回線の障害時に、前記出力情報登録アドレスリスト記憶手段から障害が発生した出力回線に対応する前記出力情報登録アドレスリストを読み出し、出力情報アドレス記憶手段のアドレスに対して迂回路となるべき出力回線に対応付けられた出力情報のアドレスを書き込む出力情報アドレス書替手段を有してもよい。

40

【0032】

50

前記の packets 転送装置は、出力回線に異常が発生した場合、前記ヘッダ情報登録アドレスリスト記憶手段から前記異常が発生した出力回線に対応付けて記憶されたヘッダ情報登録アドレスリストを読み出し、前記ヘッダ情報アドレスリストに登録されている複数のヘッダ情報アドレス記憶手段のアドレスに対して、迂回路となるべき出力回線に対応付けられたヘッダ情報の登録アドレスを連続して書き込むことにより、受信した packets の出力回線および出力字に付与するヘッダ情報を高速に切替えることが可能になる。

【0033】

本発明の要旨は、これら packets 転送方式、および、それを実現する装置構成である。本発明のその他側面については、発明を実施するための最良の形態および図面により明らかにされる。

10

【発明の効果】

【0034】

本発明の packets 転送装置をネットワークに適用することにより、packets 転送装置が収容する回線に障害が発生した場合、packets の出力回線とヘッダ情報を、迂回路となるべき別の出力回線とそれに対応するヘッダ情報に高速に切替えることが可能となり、障害が発生した回線に出力される packets の数を最小限に抑えることが可能となる。

【実施例1】

【0035】

まず、図7、図8、図9、図10を用いて、MPLSカプセル化された packets のデータフォーマットと、MPLSカプセルヘッダフォーマットについて説明する。

20

図7に、MPLSカプセル化された packets のデータフォーマットを示す。MPLSカプセル化された packets は、PPPやEthernet等の第2層プロトコルのヘッダであるLayer2ヘッダ、MPLSカプセルヘッダ部200、packets データから構成される。MPLSカプセルヘッダ部200は、 n 個(n は1以上の整数)のShimヘッダ、Shimヘッダ1からShimヘッダ n (200-1~200- n)から構成される。データ部220は、例えば、IP等の第3層プロトコルヘッダによりカプセル化されたデータである。

【0036】

図8に、レイヤ2のプロトコルで用いるヘッダの一例として、Ethernet V2フレームヘッダ構造を示す。Ethernet V2フレームヘッダには、フレームの開始を示すプリアンブル/SFD、レイヤ2における宛先の端末のMACアドレス(宛先MACアドレス)、レイヤ2における送信元のMACアドレス(送信元MACアドレス)、レイヤ3のプロトコルを示すタイプが格納されている。

30

【0037】

図9に、レイヤ2のプロトコルで用いるヘッダの一例として、Tag VLAN対応Ethernet V2フレームヘッダ構造を示す。Tag VLAN対応Ethernet V2フレームヘッダには、フレームの開始を示すプリアンブル/SFD、レイヤ2における宛先の端末のMACアドレス(宛先MACアドレス)、レイヤ2における送信元のMACアドレス(送信元MACアドレス)、Tag VLAN対応であることを示すTAGプロトコルID、フレームの転送優先度を示す優先度、レイヤ2におけるルーティング情報を格納していることを示すCanonical Field Indicator (CFI)、VLANの識別子であるVLAN ID、レイヤ3のプロトコルを示すタイプが格納されている。

40

【0038】

図10に、MPLSカプセルヘッダ部を構成するShimヘッダのフォーマットを示す。ShimヘッダのフォーマットはShimヘッダ200-1からShimヘッダ200- n に関して共通のフォーマットである。ルータが出力回線と出力ラベルを決定するために用いるラベルフィールド201、MPLS内ネットワーク内の packets の通信品質を示すExperimental Use (EXP)フィールド202、2つ以上のShimヘッダを付与する(以下、Shimヘッダを付与することを「PUSHする」と記述する)場合の最下位に位置するラベルであることを示すSビットフィールド(Bottom of Stackフィールド)203、ネットワーク内で通過可能なルータ数を示すTime to Live (TTL)フィールド204、から構成される。図10に示すフォーマットは、RFC3032、3ページ、図1に記載されているものと同様である。

50

【 0 0 3 9 】

MPLSカプセル化されたパケットを受信したルータは、ラベルフィールド201の値を検索キーとしてラベルテーブルを検索し、出力回線と1つ以上の出力ラベルと次に転送すべきルータのIPアドレスであるNext Hop IPを決定する。また、ルータはEXPの値に基づきパケットの通信品質を判定し、通信品質を保証するための優先転送制御を行う。また、ルータは入力Shimヘッダ内のTTLフィールド204の値を1だけ減算し、その値を出力Shimヘッダ内のTTLフィールドに設定する。また、ルータは、レイヤ2のプロトコルとしてEthernetを用いる場合、Next Hop IPから対応するMACアドレスを決定し、図8あるいは図9の宛先MACアドレスに書き込む。

【 0 0 4 0 】

次に、本発明のルータの一実施例を図11から図23を用いて説明する。図11は本発明のルータの一構成例である。ルータ1000はパケットが入力するN本の入力回線1010-i(i=1~N)とパケットの受信処理を行うパケット受信処理部1100-iと、パケット転送処理部1200と、パケットの出力回線とヘッダ情報を決定するルーティング処理部1300と、パケットをスイッチングするスイッチ1400と、出力回線毎の優先転送制御を行うパケット送信処理部1500-jと、回線状態を監視する回線状態監視部1600-jと、パケットが出力されるN本の出力回線1020-jと、ルータ全体の制御およびルーティング処理を行う制御部1700より構成される。制御部1700はルーティング処理部内の各テーブルの設定、各回線状態監視部からの障害通知の処理等を行う。

【 0 0 4 1 】

図12は、ルータ1000内部におけるパケットフォーマットの一実施例を示す。ルータ1000内部におけるパケットのフォーマットは、MPLSネットワークにおけるパケットのフォーマットに内部ヘッダ部230が付加されたものになる。前記内部ヘッダ部230は、パケットが入力された回線の識別子である入力回線番号231と、パケットが出力される回線の識別子である出力回線番号232と、パケットの装置内での通信品質を示す装置内QoS情報233から構成される。

【 0 0 4 2 】

ルータ1000の入力回線1010-iよりパケットが入力されると、パケット受信回路1100-iは内部ヘッダ部230を付加し、パケットが入力された入力回線1010-iの回線番号iを入力回線番号231に書き込み、パケット転送処理部1200へパケットを送信する。なお、この時点では出力回線番号232、装置内QoS情報233は無意味な値となっている。

【 0 0 4 3 】

パケット転送処理部1200はパケットを受信するとMPLSカプセルヘッダ部200、データ、内部ヘッダ部230の全てを蓄積し、MPLSカプセルヘッダ情報200と内部ヘッダ情報230を抽出してルーティング処理部1300へ送信する。

【 0 0 4 4 】

ルーティング処理部1300は、前記MPLSカプセルヘッダ情報210と内部ヘッダ情報230を用いてラベルテーブル検索処理を行い、出力回線(例えば回線1020-j)、出力Shimヘッダ、POP数、PUSH数、パケットの装置内の通信品質を示す装置内QoS情報、を決定し、それらの情報をパケット転送処理部1200に通知する。さらに、ルーティング処理部1300は、ラベルテーブル検索処理を行い、Next Hop IPを決定する。その後、Next Hop IPから次に転送すべきルータのMACアドレスを決定し、前記情報をパケット転送処理部1200に通知する。なお、ルーティング処理部1300の構成および動作については後に詳細に述べる。

【 0 0 4 5 】

パケット転送処理部1200は出力回線1020-jの回線番号jを内部ヘッダ内の出力回線番号232に書き込み、装置内QoS情報を内部ヘッダ内の装置内QoS情報233に書き込む。また、出力Shimヘッダと、POP数、PUSH数を用いて、MPLSカプセルヘッダ部200内のShimヘッダを書き替える。入力時のMPLSカプセルヘッダ部200のデータ長と、出力時のMPLSカプセルヘッダ部のデータ長は、POP数、PUSH数により異なる場合がある。その場合はPOP数、PUSH数に従いMPLSカプセルヘッダ部のデータ長を再調整する。また、パケット転送処理部1200は、次

10

20

30

40

50

に転送すべきルータのMACアドレスをレイヤ2ヘッダ内の宛先MACアドレスに書き込む。その後、パケット転送処理部1200は内部パケットをスイッチ1400に送信する。

【0046】

スイッチ1400は出力回線番号232に従いパケットをスイッチングし、出力回線毎のパケット送信処理部1500-jに送信する。パケット送信処理部1500-jは装置内QoS情報233に従いパケットの優先度に対応した送信制御を行い、内部ヘッダ部230を削除し、出力回線1020-jにパケットを送信する。

以上、本発明のルータの一実施例について、その構成と動作について説明した。

【0047】

次に、図13から図16を用いて、ルーティング処理部1300の一実施例を説明する。

図13に、ルーティング処理部1300の一構成例を示す。ルーティング処理部1300は、ヘッダ情報蓄積部1310と、ラベルテーブル1320と、前記ラベルテーブル1320を検索するラベルテーブル検索処理部1330と、ヘッダ内情報抽出部1340と、TTL減算部1350と、出力ヘッダ生成部1360と、Next Hop IPと次に転送すべきルータのMACアドレスを対応づけて記憶するAddress Resolution Protocol(以下、ARPと呼ぶ)テーブル1395と、現用予備判定部1380と、回線状態テーブル1390とから構成される。

【0048】

ヘッダ情報蓄積部1310は、パケット転送処理部1200から受信したMPLSカプセルヘッダ部200と内部ヘッダ部230を蓄積する。ヘッダ内情報抽出部は、ヘッダ情報蓄積部に蓄積された入力Shimヘッダ内のEXP値とTTL値を抽出し、信号線1371を介してEXP値を出力Shimヘッダ生成部1360へ出力し、信号線1372を介してTTL値をTTL減算部1350に出力する。TTL減算部1350は入力Shimヘッダ内のTTL値を1だけ減算し、信号線1373を介して計算結果を出力ヘッダ生成部1360へ出力する。

【0049】

回線状態テーブル1390は、ルータ1000が所持する全回線に対し、正常か障害発生かを示すビットを保持するテーブルである。図14に回線状態テーブル1390の一構成例を示す。図14は回線番号1391と回線状態1392から構成される。回線状態は、正常の場合"0"、障害発生時の場合"1"に設定するとする。

【0050】

ラベルテーブル1320の一構成例を図15に示す。ラベルテーブル1320の各エントリは、検索キーであるラベル1321およびEXP1322と、それに対応する正常時の検索結果(以下、現用の検索結果と呼ぶ)と、異常時の検索結果(以下、予備の検索結果と呼ぶ)の両方から構成される。前記各検索結果1323-j(j=1, 2)は、それぞれ、現用、予備用の入力ShimヘッダのPOP数1324-j、出力ShimヘッダのPUSH数1325-j、PUSH数分の0個以上の出力ラベル1326-j、出力回線番号1327-j、Next Hop IP 1328-j、装置内QoS情報1329-jから構成される。

【0051】

図13では、本発明のルータ1000を図2、図3、図4のCR1に適用した場合のテーブル設定値を示している。この場合、現用の検索結果1323-1に関しては、入力ラベルL11、L21に対応するPOP数は1、PUSH数は1、出力ラベルはラベル1フィールド(1326-1)内にそれぞれL12、L22の一つのみ設定され、出力回線は"1"、Next Hop IPはIP152-1となる。また、入力EXP値に対応する装置内QoS情報を設定するため、入力ラベルL21に対応するエントリはEXP値がE2とE3の2つ分用意されている。また、予備用の検索結果1323-2に関しては、ラベルL11、L21に対応するPOP数は1、PUSH数は2、出力ラベルはラベル1、ラベル2の二つ設定され、出力回線は"2"、Next Hop IPはIP153となる。

【0052】

ARPテーブル1395の一構成例を図16に示す。ARPテーブル1395の各エントリは、検索キーであるIPアドレス1395-1と、それに対応するMACアドレスから構成される。図16は、図2、図3、図4のルータCR1が保持するARPテーブルの内容の一例を示している。この場合、図16のARPテーブル1395には、CR2の回線1020-1のインターフェースのIPアドレス(IP152-1)に対応して、MACアドレス(MAC152-1)が登録され、また、CR3の回線1020-2のインターフェー

10

20

30

40

50

スのIPアドレス(IP153)に対応して、MACアドレス(MAC153)が登録されている。826, 2461前記ARPテーブル1395の各エントリは、ルータCR1が、隣接ルータCR2、CR3に対して、IP version 4の場合は、RFC826に記載されているAddress Resolution Protocolなど、あるいはIP version 6の場合は、RFC2461に記載されているNeighbor Discovery Protocolなどを用いてCR2、CR3のIPアドレスに対応する隣接ルータCR2、CR3のMACアドレスを問い合わせることにより生成される。

【0053】

次に、ルーティング処理部1300の動作について述べる。まず、ラベルテーブル検索処理部1330が、ヘッダ情報蓄積部1310に蓄積されている情報の内、ラベルテーブル検索に必要な情報を検索キーとしてラベルテーブル1320を検索する。検索の結果、検索キーに対応する現用の検索結果1323-1と予備の検索結果1323-2を得る。ラベルテーブル検索処理部1330は、これら2組の検索結果を現用予備判定部1380に出力する。現用予備判定部1380は、受信した2組の検索結果のうち、現用の出力回線番号と予備の出力回線番号を検索キーとして回線状態テーブル1390を検索し、現用の出力回線と予備の出力回線の状態を読み出す。この結果、現用の出力回線が正常の場合は、現用予備判定部1380は、現用の検索結果1323-1を選択し、前記検索結果のうち、検索キーに対応する入力ShimヘッダのPOP数1324-1、出力ShimヘッダのPUSH数1325-1、PUSH数分の出力ラベル1326-1、出力回線番号1327-1、Next Hop IP 1328-1、装置内QoS情報1329-1を、信号線1374を介してパケット転送処理部1200へ出力する。

【0054】

また、現用予備判定部1380は、POP数、PUSH数、出力ラベル、Next Hop IPを、信号線1375を介して出力ヘッダ生成部1360へ出力する。また、現用の出力回線に障害が発生し、かつ、予備用の出力回線が正常な場合は、予備の検索結果1323-2を選択し、前記現用の出力回線が正常の場合と同様の情報をパケット転送処理部1200と出力ヘッダ生成部1360へ出力する。現用の出力回線、予備の出力回線ともに障害が発生している場合は、パケットを廃棄する指示をパケット転送処理部1200へ通知する。

【0055】

ラベルテーブル検索と並行して、ヘッダ内情報抽出部1340は、信号線1377を介して、ヘッダ情報蓄積部1310に蓄積されている情報のうち、入力Shimヘッダ内のEXP値、TTL値とを抽出し、EXP値を出力ヘッダ生成部1360へ出力し、TTL値をTTL減算処理部1350へ出力する。TTL減算処理部1350はTTL値を1だけ減算し、信号線1373を介して計算結果を出力ヘッダ生成部1360へ転送する。

【0056】

出力ヘッダ生成部1360は、信号線1375を介してラベルテーブル検索処理部1330からPOP数、PUSH数、出力ラベル、Next Hop IPを受信し、信号線1371を介してヘッダ内情報抽出部1340からEXP値を受信し、信号線1373を介してTTL減算処理部1350からTTL値の減算結果を受信すると、出力ヘッダ生成処理を行う。この際、ラベル値としては、現用予備判定部1380から受信したラベル値を用いる。複数の出力ShimヘッダのEXP値としては、ヘッダ内情報抽出部1340から受信した入力Shimヘッダ内のEXP値を用いる。また、TTL値はTTL減算処理部1350で減算された値を用いる。出力ヘッダ生成部1360は、PUSH数分の出力Shimヘッダを生成する。また、出力ヘッダ生成部1360は、Next Hop IPを検索キーにしてARPテーブル1395を検索して次に転送すべきルータのMACアドレスを決定する。その後、出力ヘッダ生成部1360は、前記PUSH数分の出力Shimヘッダと次に転送すべきルータのMACアドレスを、信号線1376を介してパケット転送処理部1200に送信する。

【0057】

以上で説明したルータの一実施例では、パケット転送プロトコルとして、MPLSを用いた例を説明したが、パケット転送プロトコルとして、MPLS以外のプロトコル、例えば、IPを用いてもよい。

以上、本発明のルーティング処理部1300の動作について説明した。以上説明したように、ラベルテーブルエントリに現用の検索結果と予備の検索結果を予め両方用意しておき、さ

10

20

30

40

50

らに回線状態テーブルを保持することにより、回線障害が発生した際、制御部は回線テーブルの障害が発生した回線番号の状態を書替えるだけで、テーブルを切替えることができる。したがって、テーブルの変更時間の短縮を実現することができる。

【0058】

また、現用の検索結果と予備用の検索結果毎に異なる装置内QoS情報を設定できるため、現用の出力回線にパケットを出力する際のパケットの転送制御と、予備用の出力回線にパケットを出力する際のパケットの転送制御とを異なる転送制御にすることが可能となる。例えば、現用の出力回線の回線帯域に比べて、予備用の出力回線の回線帯域が小さい場合、現用の出力回線では全てのパケットの出力時の優先度を高く設定しておくが、予備用の出力回線では、重要なパケットの出力時の優先度のみ高くしておき、比較的重要なでないパケットの出力時の優先度を低くすることで、出力回線の帯域が小さい予備用の出力回線に切替えた場合でも、重要なパケットは優先して出力することが可能になる。

【実施例2】

【0059】

次に、図17から図21を用いて、ルーティング処理部1300の、図13から図16を用いて説明した実施例とは別の一実施例を説明する。

図17に、ルーティング処理部1300の一構成例を示す。図17のルーティング処理部1300は、ヘッダ情報蓄積部1310と、ラベルテーブル1320と、前記ラベルテーブル1320を検索するラベルテーブル検索処理部1330と、ヘッダ内情報抽出部1340と、TTL減算部1350と、出力ヘッダ生成部1360と、ARPテーブル1395と、アドレスリスト生成部1830と、アドレスリスト

【0060】

図17のラベルテーブル1320は、ラベル判定テーブル1800と、アドレス変換テーブル1810と、検索結果テーブル1820とから構成される。以下では、まず、図2で説明した、CR1の出力回線1020-1が正常な場合のラベルテーブルの例を示す。

図18に、ラベル判定テーブル1800の一構成例を示す。ラベル判定テーブル1800は、検索キーであるラベル1321およびEXP1322と、それに対応するアドレス変換テーブルのアドレス1801から構成される。

【0061】

図19に、アドレス変換テーブル1810の一構成例を示す。アドレス変換テーブル1810は、

【0062】

図20に、検索結果テーブル1820の一構成例を示す。検索結果テーブル1820の各エントリは、各検索結果が登録される位置を示すアドレス1821と、入力ShimヘッダのPOP数1324、出力ShimヘッダのPUSH数1325、PUSH数分の0個以上の出力ラベル1326、出力回線番号1327、Next Hop IP 1328、装置内QoS情報1329から構成される。図19のアドレス変換テーブル1810に登録されている検索結果テーブルアドレス1812は、図20の検索結果テーブル1820の

【0063】

図20では、本発明のルータ1000を図2、図3、図4のCR1に適用した場合のテーブル設定値を示している。本実施例では、検索結果テーブル1820に、現用の検索結果と予備用の検索結果の両方を設定する。図20では、アドレス1、アドレス2、アドレス3に、現用の検索結果が設定され、アドレス4、アドレス5、アドレス6に、予備用の検索結果が設定された例を示している。各検索結果は、アドレス1、アドレス2、アドレス3に設定されている現用の検索結果に関しては、入力ラベルL11、L21に対応するPOP数は1、PUSH数は1、出力ラベルはラベル1フィールド(1326)内にそれぞれL12、L22の一つのみ設定され、出力回線は”1”、Next Hop IPはIP152-1となる。また、入力EXP値に対応する装置内QoS情報を設定する

ため、入力ラベルL21に対応するエントリは、EXP値がE2とE3の2つ分がアドレス2とアドレス3に設定されている。また、アドレス4、アドレス5、アドレス6に設定されている予備用の検索結果に関しては、ラベルL11、L21に対応するPOP数は1、PUSH数は2、出力ラベルはラベル1、ラベル2の二つ設定され、出力回線は「2」、Next Hop IPはIP153となる。入力ラベルL21に対応するエントリは、EXP値がE2とE3の2つ分がアドレス5とアドレス6に設定されている。

【0064】

次に、図21、図22、図23を用いて、アドレスリスト生成部1370とアドレスリストテーブル1740について説明する。

図21に、制御部1700が入力ラベルL11、EXP E1に対する現用の検索結果をラベルテーブル1320内の各テーブルに対して書き込む際に、ルーティング処理部1300へ送信する書き込み情報の一例を示す。図21に示す書き込み情報は、ラベル情報として入力ラベルとEXPが格納され、また、書き込み情報が現用の検索結果か予備用の検索結果かを示す予備識別子と、入力ShimヘッダのPOP数、出力ShimヘッダのPUSH数、PUSH数分の0個以上の出力ラベル、現用の出力回線番号、Next Hop IP、装置内QoS情報が格納されている。また、予備用の出力回線番号のフィールドもあるが、本例の場合、現用の検索結果を書き込む際の書き込み情報のため、前記予備用の出力回線番号のフィールドは無意味な値が設定されている。図21は、現用の検索結果を書き込む際の書き込み情報の例を示しているため、予備識別子は、現用を示す'0'という値が格納されている。

【0065】

図22に、制御部1700が入力ラベルL11、EXP E1に対する予備用の検索結果をラベルテーブル1320内の各テーブルに対して書き込む際に、ルーティング処理部1300へ送信する書き込み情報の一例を示す。格納される情報の種類は、図21で示したものと同様である。ただし、図22は、予備用の検索結果を書き込む際の書き込み情報の例を示しているため、予備識別子は、予備用を示す'1'という値が格納されている。また、現用の出力回線番号には、現用の検索結果として記録されている「1」が格納され、予備用の出力回線番号には、「2」が格納される。

【0066】

図23に、アドレスリストテーブル1840の一構成例を示す。アドレスリストテーブル1840は、現用の出力回線番号1841と、前記現用の出力回線番号に対応付けられ、前記現用の出力回線番号1841を出力回線とする全ての入力ラベル情報1843-k-1(k=1、2、3)(入力ラベル1843-k-1-1とEXP1843-k-1-2)と、前記入力ラベルに対応付けられて記録されるアドレス変換テーブルのアドレス1843-k-2と、予備用の検索結果が記録されるアドレス1843-k-3とを、一つのエントリ1843-kとして、前記エントリを複数記録するエントリリスト1843と、前記記録されるエントリの数が記録されるエントリ数1842から構成される。前記アドレスリスト1843の構成は、各エントリをポインタで連結する構成にしてもよいし、アドレスリストテーブルを構成するメモリの連続するアドレスに、各エントリを連続的に配置する構成にしてもよい。

【0067】

次に、制御部1700が入力ラベルL11、EXP E1に対する現用の検索結果をラベルテーブル1320内の各テーブルに対して書き込む際の、アドレスリスト生成部1830の動作について説明する。アドレスリスト生成部1830は、制御部1700が現用の検索結果をラベルテーブル1320内の各テーブルに対して書き込む際、図21に示した書き込み情報内の、ラベル情報としてラベル判定テーブル1800に書き込まれるラベル「L11」とEXP「E1」と、前記ラベル判定テーブル1800に登録されるラベル情報に対応するアドレス変換テーブル1810の設定アドレス「1」と、前記アドレスに記録される出力回線番号「1」を抽出し、ラベルと、EXPと、アドレス変換テーブル1810の設定アドレスとを1つのエントリとして、出力回線「1」に対応するにアドレスリストテーブルのエントリ1(1843-1)に書き込む。

【0068】

次に、制御部1700が入力ラベルL11、EXP E1に対する予備用の検索結果をラベルテーブ 50

ル1320内の各テーブルに対して書き込む際の、アドレスリスト生成部1830の動作について説明する。アドレスリスト生成部1830は、制御部1700が予備用の検索結果をラベルテーブル1320内の検索結果テーブル1820に対して書き込む際、図22に示した書き込み情報内の、現用出力回線番号「1」に対応するアドレスリストのエントリを全て検索し、前記図22に示した書き込み情報内のラベル情報(ラベル「L11」とEXP「E1」)に一致するエントリを検索する。本例の場合、エントリ1のラベル情報にラベル「L11」とEXP「E1」が記録されているため、前記検索の結果、エントリ1が一致する。その後、アドレスリスト生成部1830は、図22に示したPOP数、PUSH数、出力ラベル情報、予備出力回線番号、Next Hop IP、装置内QoS情報が記録された検索結果テーブル内のアドレス「4」を、前記エントリ1の予備用検索結果記録アドレス1343-1-3に書き込む。

10

【0069】

次に、出力回線「1」に障害が発生した際の、アドレスリスト生成部1830の処理について説明する。出力回線「1」に障害が発生した際には、前記アドレスリストテーブルの出力回線「1」に対応するアドレスリストを読み出し、各エントリ1843-k内のアドレス変換テーブルアドレス1843-k-2に記録されているアドレス変換テーブルのアドレスの内容を、同じく各エントリ1843-k内の予備用検索結果記録アドレス1843-k-3内に登録されている値に書替える。以上の処理を障害が発生した出力回線「1」に対応するアドレスリスト内のエントリ全てに対して行うことにより、回線「1」を出力回線としていたラベルテーブルエントリの検索結果を、現用の検索結果から予備用の検索結果に切替えることが可能になる。

20

【0070】

次に、図17のルーティング処理部1300の動作について述べる。まず、ラベルテーブル検索処理部1330が、ヘッダ情報蓄積部1310に蓄積されている情報の内、ラベル判定テーブル1800の検索に必要な情報を検索キーとしてラベル判定テーブル1800を検索する。検索の結果、検索キーに対応するアドレス変換テーブル1810のアドレス1811を決定する。次に、ラベルテーブル検索処理部1330は、アドレス変換テーブル1810の前記決定したアドレス1811に記録されている検索結果テーブルのアドレス1812を読み出す。次に、テーブル検索処理部1330は、前記読み出した検索結果テーブルのアドレス1812に記録されている検索結果である、入力ShimヘッダのPOP数1324、出力ShimヘッダのPUSH数1325、PUSH数分の出力ラベル1326、出力回線番号1327、Next Hop IP 1328、装置内QoS情報1329を決定する。ラベル

30

【0071】

ラベルテーブル検索と並行して、ヘッダ内情報抽出部1340は、信号線1377を介して、ヘッダ情報蓄積部1310に蓄積されている情報のうち、入力Shimヘッダ内のEXP値、TTL値とを抽出し、EXP値は出力ヘッダ生成部1360へ出力し、TTL値をTTL減算処理部1350へ出力する。TTL減算処理部1350はTTL値を1だけ減算し、信号線1373を介して計算結果を出力ヘッダ生成部1360へ転送する。

40

【0072】

出力ヘッダ生成部1360は、信号線1375を介してラベルテーブル検索処理部1330からPOP数、PUSH数、出力ラベル、Next Hop IPを受信し、信号線1371を介してヘッダ内情報抽出部1340からEXP値を受信し、信号線1373を介してTTL減算処理部1350からTTL値の減算結果を受信すると、出力ヘッダ生成処理を行う。この際、ラベル値としては、ラベルテーブル検索処理部1330から受信したラベル値を用いる。複数の出力ShimヘッダのEXP値としては、ヘッダ内情報抽出部1340から受信した入力Shimヘッダ内のEXP値を用いる。また、TTL値はTTL減算処理部1350で減算された値を用いる。出力ヘッダ生成部1360は、PUSH数分の出力Shimヘッダを生成する。また、出力ヘッダ生成部1360は、Next Hop IPを検索キーにしてARPテーブル1395を検索して次に転送すべきルータのMACアドレスを決定する。その後、

50

出力ヘッダ生成部1360は、前記PUSH数分の出力Shimヘッダと次に転送すべきルータのMACアドレスを、信号線1376を介してパケット転送処理部1200に送信する。

【0073】

以上、本発明のルーティング処理部1300に関し、図13で説明した一実施例とは別の実施例についてその動作について説明した。以上説明したように、検索結果テーブル1820には現用、予備用の検索結果を予め用意しておき、出力回線が正常の場合は、アドレス変換テーブル内の検索結果テーブルアドレスに現用の検索結果が記録されている検索結果テーブルのアドレスを記録しておくことで、出力回線が正常の場合の出力ラベル、出力回線、Next Hop IP、装置内QoS情報を決定することができる。また、出力回線に異常が発生した場合には、異常が発生した出力回線毎に予め記録してあるアドレスリストテーブル内の書替えるべきアドレス変換テーブルのアドレスと、書替えるべきアドレス変換テーブルの内容を連続的に読み出し、アドレス変換テーブル内の検索結果テーブルアドレスを現用の検索結果が記録されている検索結果テーブルのアドレスから、予備用の検索結果が記録されている検索結果テーブルのアドレスに書替えることにより、検索結果を現用の検索結果から予備用の検索結果に高速に切替えることができる。また、本発明の図17で説明した一実施例を用いることにより、パケットを受信する際に行う検索結果テーブルからの検索結果情報の読出しデータ量が、現用あるいは予備用のどちらか一方のみで済み、検索結果テーブルを格納するメモリへの読出しアクセス回数を削減することができ、パケットの検索処理、転送処理の高速化が可能である。

10

【産業上の利用可能性】

20

【0074】

本発明は、IPパケットおよびMPLSパケットを転送するルータおよび、Ethernetフレームを転送するスイッチなどのパケット転送装置において、前記パケット転送装置が複数所持する1つの回線に障害が発生した際に、パケットの出力回線およびパケットに付与するヘッダ情報を高速に別の回線と前記別の回線に対応するヘッダ情報に切替える方式として利用可能である。

【図面の簡単な説明】

【0075】

【図1】MPLSのパケット転送概要を説明する図。

【図2】従来のMPLSを用いた高速パス切替え技術を説明する図。

30

【図3】従来のMPLSを用いた高速パス切替え技術を説明する図。

【図4】従来のMPLSを用いた高速パス切替え技術を説明する図。

【図5】図2のルータCR1が所持するラベルテーブルを説明する図。

【図6】図3、図4のルータCR1が所持するラベルテーブルを説明する図。

【図7】MPLSネットワークにおけるパケットのフォーマットの一例を示す図。

【図8】Ethernet V2のヘッダ構造を示す図。

【図9】Tag VLAN対応Ethernet V2のヘッダ構造を示す図。

【図10】MPLSカプセルヘッダであるShimヘッダのフォーマットの一例を示す図。

【図11】本発明のルータの一構成例を示す図。

【図12】本発明のルータ1000内部におけるパケットのフォーマットの一例を示す図。

40

【図13】ルーティング処理部1300の一構成例を示す図。

【図14】回線状態テーブル1390の一構成例を示す図。

【図15】ラベルテーブル1320の一構成例を示す図。

【図16】ARPテーブル1395の一構成例を示す図。

【図17】ルーティング処理部1300の図13とは別の構成例を示す図。

【図18】ラベル判定テーブル1800の一構成例を示す図。

【図19】アドレス変換テーブル1810の一構成例を示す図。

【図20】検索結果テーブル1820の一構成例を示す図。

【図21】現用の検索結果を書き込む際のラベルテーブル書き込み情報の一構成例を示す図。

50

【図22】予備用の検索結果を書き込む際のラベルテーブル書き込み情報の一構成例を示す図。

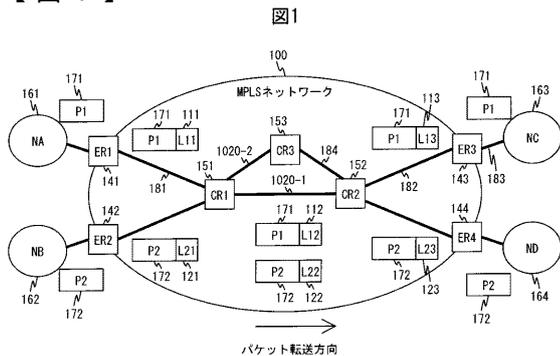
【図23】アドレスリストテーブルの一構成例を示す図。

【符号の説明】

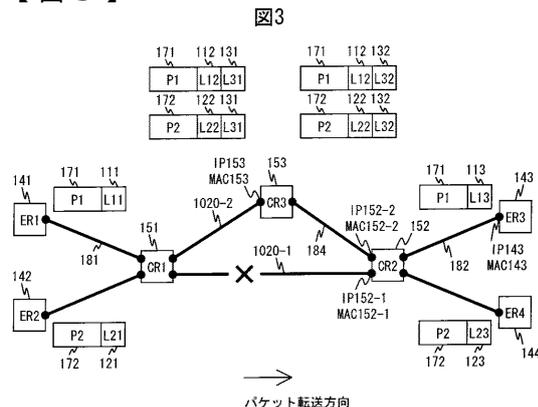
【0076】

100...MPLSネットワーク、141~144...エッジルータ、151~153...コアルータ、1000...本発明のルータ、1100...パケット受信処理部、1200...パケット転送処理部、1300...ルーティング処理部、1310...ヘッダ情報蓄積部、1320...ラベルテーブル、1330...ラベルテーブル検索処理部、1340...ヘッダ内情報抽出部、1360...出力ヘッダ生成部、1380...現用予備判定部、1400...スイッチ、1500...パケット送信処理部、1700...制御部。

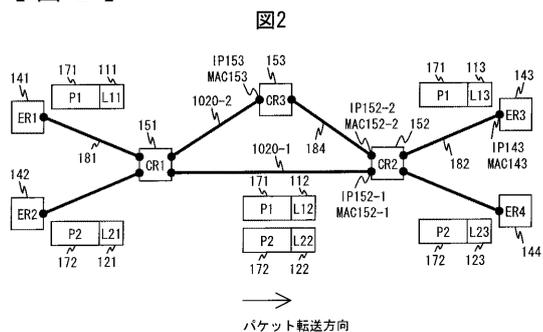
【図1】



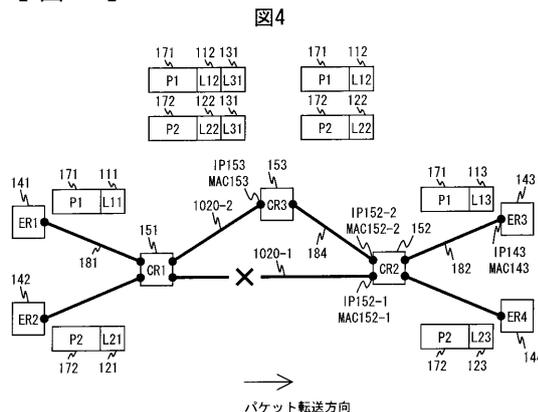
【図3】



【図2】



【図4】



【 図 5 】

図5

入力ラベル	出力ラベル	出力回線	Next Hop IP
L11	L12	1020-1	IP152-1
L12	L22	1020-1	IP152-1

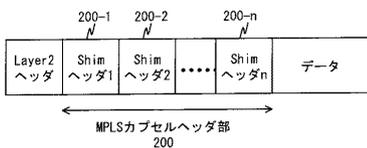
【 図 6 】

図6

入力ラベル	出力ラベル	出力回線	Next Hop IP
L11	L12, L31	1020-2	IP153
L12	L22, L31	1020-2	IP153

【 図 7 】

図7



【 図 8 】

図8

プリアンブル /SFD	宛先MAC アドレス	送信元MAC アドレス	タイプ
-------------	------------	-------------	-----

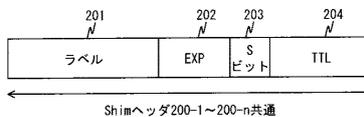
【 図 9 】

図9

プリアンブル /SFD	宛先MAC アドレス	送信元MAC アドレス	TAG プロトコル ID	優先度	C F I	VLAN ID	タイプ
-------------	------------	-------------	--------------	-----	-------	---------	-----

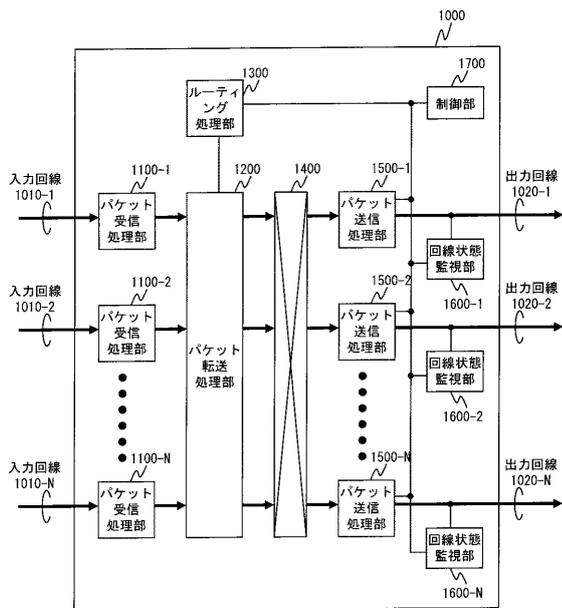
【 図 1 0 】

図10



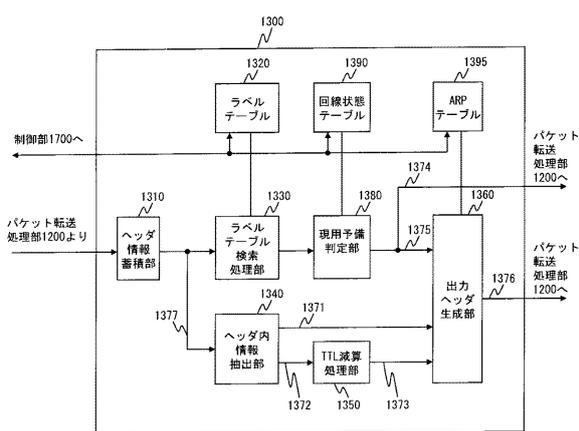
【 図 1 1 】

図11



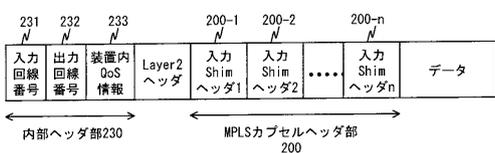
【 図 1 3 】

図13



【 図 1 2 】

図12



【図 14】

図14

回線 番号		回線 状態	
1	1	1	1
2	0	1	0
⋮	⋮	⋮	⋮
N	0	1	0

【図 15】

図15

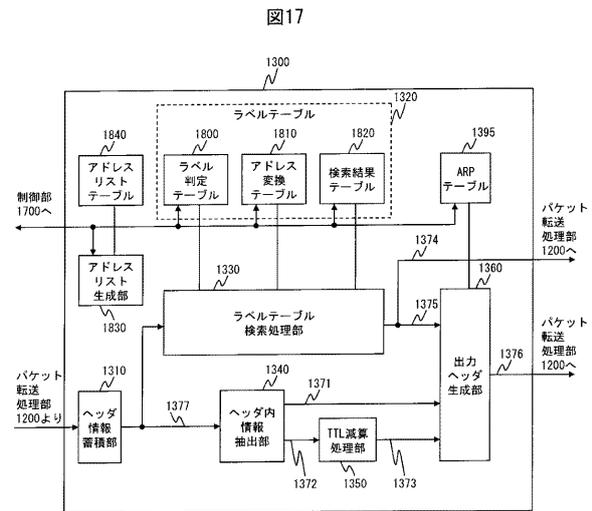
検索キー		検索結果 (現用)										検索結果 (予備)													
入力ラベル 情報		POP 数		PUSH 数		出力ラベル 情報		出力 回線 番号		Next Hop IP		装置 内 QoS 情報		POP 数		PUSH 数		出力ラベル 情報		出力 回線 番号		Next Hop IP		装置 内 QoS 情報	
ラベル	EXP			ラベル	ラベル	...											ラベル	ラベル	...						
L11	E1	1	1	L12	-	-	1	IP 152-1	Q1	1	2	L31	L12	-	2	IP 153	Q1								
L21	E2	1	1	L22	-	-	1	IP 152-1	Q2	1	2	L31	L22	-	2	IP 153	Q2								
L21	E3	1	1	L22	-	-	1	IP 152-1	Q3	1	2	L31	L22	-	2	IP 153	Q3								
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

【図 16】

図16

IP アドレス		MAC アドレス	
IP152-1	MAC152-1	IP153	MAC153
⋮	⋮	⋮	⋮

【図 17】



【図18】

図18

1800		1801	
入カラベル情報		アドレス変換 テーブルアドレス	
ラベル	EXP		
L11	E1		1
L21	E2		2
L21	E3		3
⋮	⋮		⋮
		1321	1322

【図19】

図19

1810	1811	1812
アドレス	検索結果テーブル アドレス	
1		1
2		2
3		3
⋮		⋮

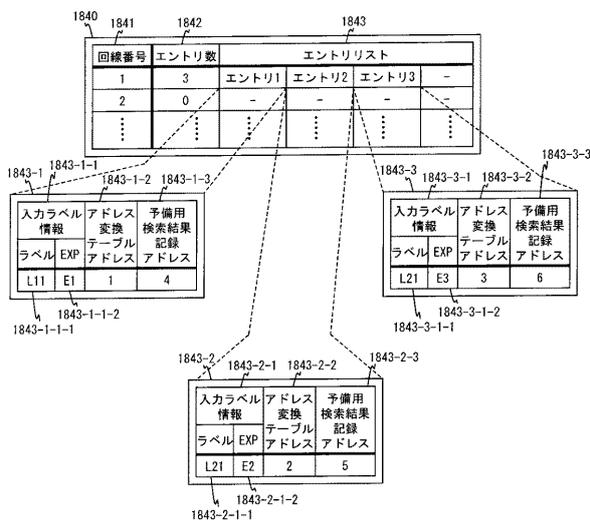
【図20】

図20

1820	1821	1323	1324	1325	1326	1327	1328	1329
検索結果								
アドレス	POP 数	PUSH 数	出カラベル情報			出力回線 番号	Next Hop IP	装置内QoS 情報
			ラベル1	ラベル2	⋮			
1	1	1	L12	-	-	1	IP152-1	Q1
2	1	1	L22	-	-	1	IP152-1	Q2
3	1	1	L22	-	-	1	IP152-1	Q3
4	1	2	L31	L12	-	2	IP153	Q1
5	1	2	L31	L22	-	2	IP153	Q2
6	1	2	L31	L22	-	2	IP153	Q3
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

【図23】

図23



【図21】

図21

入カラベル情報		予備 識別子	POP 数	PUSH 数	出カラベル情報			現用 出力回線 番号	予備 出力回線 番号	Next Hop IP	装置内QoS 情報
ラベル	EXP				ラベル1	ラベル2	⋮				
L11	E1	0	1	1	L12	-	-	1	-	IP152-1	Q1

【図22】

図22

入カラベル情報		予備 識別子	POP 数	PUSH 数	出カラベル情報			現用 出力回線 番号	予備 出力回線 番号	Next Hop IP	装置内QoS 情報
ラベル	EXP				ラベル1	ラベル2	⋮				
L11	E1	1	1	2	L31	L12	-	1	2	IP153	Q1

【手続補正書】

【提出日】平成17年7月1日(2005.7.1)

【手続補正1】

【補正対象書類名】明細書

【補正対象項目名】0019

【補正方法】変更

【補正の内容】

【0019】

【特許文献1】特開平8-242240号公報