

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第4491397号
(P4491397)

(45) 発行日 平成22年6月30日(2010.6.30)

(24) 登録日 平成22年4月9日(2010.4.9)

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

請求項の数 8 (全 15 頁)

(21) 出願番号	特願2005-294273 (P2005-294273)	(73) 特許権者	504411166 アラクサラネットワークス株式会社 神奈川県川崎市幸区鹿島田890
(22) 出願日	平成17年10月7日(2005.10.7)	(74) 代理人	100100310 弁理士 井上 学
(65) 公開番号	特開2007-104513 (P2007-104513A)	(72) 発明者	渡邊 林音 神奈川県川崎市幸区鹿島田890 アラク サラネットワークス株式会社内
(43) 公開日	平成19年4月19日(2007.4.19)	(72) 発明者	矢崎 武己 神奈川県川崎市幸区鹿島田890 アラク サラネットワークス株式会社内
審査請求日	平成19年11月9日(2007.11.9)	審査官	吉田 隆之

最終頁に続く

(54) 【発明の名称】 トラフィック迂回機能を備えるパケット転送装置。

(57) 【特許請求の範囲】

【請求項1】

パケット転送装置であって、
複数の物理回線を束ねて論理回線として扱う送受信部と、
前記送受信部が出力するパケットを迂回させる予備の送受信部と、
出力ラベルと、前記論理回線の識別子とを対応付けて記憶する第1の記憶部と、
前記論理回線の識別子と、前記論理回線が含む前記物理回線の識別子を対応付けて記憶する第2の記憶部と、
前記第1の記憶部を参照し、前記第1の送受信部で受信したパケットのヘッダ情報に基づいて前記出力ラベルと前記論理回線を検索する出力ラベル検索部と、
前記第2の記憶部を参照し、前記ラベル検索部で検索された前記論理回線が含む前記物理回線を検索する物理回線検索部と、
前記送受信部で受信した第1の受信パケットのヘッダを基に検索された第1の物理回線が障害が発生していた場合、前記論理回線の中の前記第1の物理回線の送受信を停止させ、前記第1のパケットを前記予備の送受信部から送出する宛先判定部とを有することを特徴とするパケット転送装置。

【請求項2】

請求項1記載のパケット転送装置であって、
前記物理回線の障害を検出する障害検出部を有し、
前記障害検出部は、前記物理回線で障害を検知すると、前記第2の記憶部の前記物理回

線の識別子を障害発生を示す値に更新することを特徴とするパケット転送装置。

【請求項 3】

請求項 2 記載のパケット転送装置であって、
前記論理回線は、複数の物理回線のリンクアグリゲーションで構成されていることを特徴とするパケット転送装置。

【請求項 4】

請求項 3 記載のパケット転送装置であって、
前記ラベルは、MPLS のラベルであることを特徴とするパケット転送装置。

【請求項 5】

パケット転送装置であって、
複数の物理回線を束ねて論理回線として扱う送受信部と、
前記送受信部が出力するパケットを迂回させる予備の送受信部と、
出力ラベルと、前記論理回線の識別子と、閾値を対応付けて記憶する第 1 の記憶部と、
前記論理回線の識別子と、前記論理回線が含む前記物理回線の識別子に対応付けて記憶する第 2 の記憶部と、

前記第 1 の記憶部を参照し、前記第 1 の送受信部で受信したパケットのヘッダ情報に基づいて前記出力ラベルと前記論理回線を検索する出力ラベル検索部と、

前記第 2 の記憶部を参照し、前記ラベル検索部で検索された前記論理回線が含む前記物理回線を検索する物理回線検索部と、

前記送受信部で受信した第 1 の受信パケットのヘッダを基に検索された前記論理回線が含む前記物理回線の数が閾値を下回った場合、前記第 1 のパケットを前記予備の送受信部から送出する宛先判定部とを有することを特徴とするパケット転送装置。

【請求項 6】

請求項 5 記載のパケット転送装置であって、
前記物理回線の障害を検出する障害検出部を有し、
前記障害検出部は、前記物理回線で障害を検知すると、前記第 2 の記憶部の前記物理回線の識別子を障害発生を示す値に更新することを特徴とするパケット転送装置。

【請求項 7】

請求項 6 記載のパケット転送装置であって、
前記論理回線は、複数の物理回線のリンクアグリゲーションで構成されていることを特徴とするパケット転送装置。

【請求項 8】

請求項 7 記載のパケット転送装置であって、
前記ラベルは、MPLS のラベルであ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、パケット転送網におけるトラフィック保護機能を備えるパケット転送装置に関する。

【背景技術】

【0002】

ネットワークを構成するノードやその物理回線の障害時に、通信の継続を実現する高可用化技術が開発されている。本技術の一例として、例えば、MPLS(Multi Protocol Label Switching)ネットワークの通信遮断を防止するファストリルート技術がある。非特許文献 1 には、パスであるLSP(Label Switching Path) (保護対象LSPと呼ぶ)を流れるトラフィックの保護方式として1対1バックアップ(one-to-one backup)方式とファシリティバックアップ(facility backup)方式が記載されている。1対1バックアップ方式では、保護対象LSPに1つの予備LSPが設定され、ファシリティバックアップ方式では、複数の保護対象LSPに1つの予備LSPが設定される。保護対象LSPが経由するノードや物理回線に障害が発生した時には、本LSPを流れるトラフィックは予備LSPへと迂回され、障害発生時の通信

10

20

30

40

50

継続が実現される。

【 0 0 0 3 】

高可用性技術の別の例としては、例えば、非特許文献 2 記載のリンクアグリゲーション技術がある。本技術では、複数の物理回線を論理的な一つの回線（論理回線と呼ぶ）として扱い、複数の物理回線にフレームを振り分けて通信する。論理回線の一つの物理回線が故障しても少なくとも一つの物理回線は正常であるため、論理回線で接続されているスイッチ間の通信が途切れることはない。そのため、スイッチ間をリンクアグリゲーションの論理回線で接続することで、一物理回線が故障した時の通信継続が実現される。

【 0 0 0 4 】

【非特許文献 1】 IETF Internet Draft draft-ietf-mpls-rsvp-lsp-fastreroute-07.txt
“Fast Reroute Extensions to RSVP-TE for LSP Tunnels”

10

【非特許文献 2】 IEEE802.3ad

【発明の開示】

【発明が解決しようとする課題】

【 0 0 0 5 】

非特許文献 2 を適用したネットワークにおいては、論理回線の一部の物理回線に障害が発生した際には通信の継続は可能であるが、ノードに障害が発生すると通信が継続されないという問題がある。

【 0 0 0 6 】

一方、非特許文献 1 を適用したネットワークにおいては、ノード障害が発生した場合、予備 LSP が障害ノードを経由しないのであれば、予備 LSP を用いて通信は継続される。しかし、両 LSP 設定時には、個々のトラフィックに必要な通信品質（通信遅延など）を満たす経路が、保護対象 LSP 用の経路以外には存在しない場合も考えられる。この場合、ノードや物理回線に障害が発生すると、通信は継続されるが、予備 LSP に迂回されたトラフィックの通信遅延が大きくなる場合や、トラフィックの一部が廃棄される場合がある。

20

【 0 0 0 7 】

そこで、ノード障害時の通信継続を実現しつつ、迂回するトラフィックを最小限に抑えて該トラフィックの通信品質劣化を最小限に抑えるパケット転送装置を提供することを一つの目的とする。

【課題を解決するための手段】

30

【 0 0 0 8 】

上記課題を解決するために、例えば、複数の入力回線と複数の出力回線と、該入力回線より入力するパケットのヘッダ内の情報から、少なくとも一つの出力回線から構成される第一の論理回線と第一の論理回線を構成する第一の出力回線を判定する宛先判定部を備え、前記第一の出力回線に該パケットを送信するパケット転送装置であって、前記宛先判定部が第一の論理回線の出力回線に障害が発生すると、前記第一の出力回線を第二の論理回線を構成する第二の出力回線に変更することを特徴とするパケット転送装置を提供する。

【 0 0 0 9 】

上記以外の本願が解決しようとする課題、その解決手段は、本願の「発明を実施するための最良の形態」の欄および図面で明らかにされる。

40

【発明の効果】

【 0 0 1 0 】

論理回線を構成する一部の物理回線に障害が発生すると、障害回線の帯域分のトラフィックを迂回すれば良く、論理回線の総帯域分のトラフィックを迂回する必要が無い。そのため、迂回するトラフィックを最小限に抑えて該トラフィックの通信品質劣化を最小限に抑えることができる。また、ノード障害が発生すると、論理回線を通るトラフィックを全て迂回することで、通信継続も実現できる。

【発明を実施するための最良の形態】

【 0 0 1 1 】

以下、図 1 から図 10 を用いて本発明に好適な実施形態の例を説明する。ただし、本発

50

明は本実施形態に限定されるものではない。

【 0 0 1 2 】

(1) ネットワークの概要

図 1 0 は、本実施例における網構成の一例を示す。図 1 0 中の IP 網 A から IP 網 E はそれぞれ異なる IP 網である。LER1 から LER4 はラベルエッジルータ (Label Edge Router : LER) であり、IP 網から送られてきた IP パケットに、宛先 IP アドレスや宛先ポート番号などの情報を集約したラベルを付加し、MPLS パケットとして MPLS 網内に送信する機能を備えている。ラベルは MPLS 網内でトラフィックの明示的な経路 (LSP) を識別する情報である。また、MPLS 網内から送られてきた MPLS パケットからラベルを削除し、IP パケットとして IP 網に送信する機能を備えている。R1 から R5 は MPLS コアルータ (MPLS Core Router) であり、MPLS パケットをラベルの値に基づき転送 (ラベルスイッチ) する機能を備えている。R-A から R-E は IP ルータであり、IP パケットを転送する機能を備えている。R1 と R2 は 1Gbps の物理回線 3 本で接続されており、この 3 本の物理回線を集約 (例えば、リンクアグリゲーションによるグループ化) し、3Gbps の帯域幅を持つ 1 本の論理回線 A として扱っている。その他のルータ間は、それぞれ 10Gbps の回線で接続されている。R1 において、論理回線 A を構成する物理回線番号は 1 番から 3 番であり、R4 と接続する物理回線番号は 4 番である。

10

【 0 0 1 3 】

LSP 1、LSP2、LSP3 は、LER1、R1、R2、R3、LER2 を通る LSP であり、それぞれの LSP には 1Gbps の帯域が予約されている。IP 網 B 内のホストと IP 網 E 内のホストが通信を行うためのパスとして、LER2、R4、R5、LER4 を経路とする LSP_A が設定されており、最大 8Gbps の帯域が予約されている。また、R1、R2 を接続する回線の回線障害および、R3 の障害に備え、保護対象 LSP1、LSP2、LSP3 の予備 LSP として LSP_backup が設定されている。それぞれの LSP_backup は R1、R4、R5、R3 を通るパスである。LSP_backup には 3Gbps の帯域が予約されている。

20

【 0 0 1 4 】

IP 網 A 内のホストと IP 網 B、IP 網 C、IP 網 E 内のホストが通信を行うため、R-A は LER1 と、R-C、R-D、R-E は LER3 と接続されている。IP 網 A 内から発信される IP 網 C、IP 網 D、IP 網 E 宛の IP トラフィックは、LER1 で宛先ごとに異なるラベルを付加され、それぞれ、該ラベルに対応した LSP 1、LSP2、LSP3 に沿って転送される。

【 0 0 1 5 】

LSP 1、LSP2、LSP3 は、IP 網 A から IP 網 C、IP 網 D、IP 網 E へ流れるトラフィックの特性を考慮した最適パスである。また、LSP_backup と LSP_A は、R4 と R5 を接続する回線において帯域を共有している。R4 と R5 を接続する回線帯域 10Gbps に対し、LSP_backup の予約帯域と LSP_A の予約帯域は 11Gbps であり、オーバーサブスクライブ (利用可能な帯域幅を越えた帯域予約) となっている。このため、LSP_backup の予約帯域分のトラフィックと LSP_A の予約帯域分のトラフィックが同時に、R4 と R5 を接続する回線帯域に流れ込むと、帯域幅を超えて遅延、輻輳が発生する可能性がある。

30

【 0 0 1 6 】

R1 は、論理回線 A を構成する物理回線に障害が起きた場合、LSP1、LSP2 又は LSP3 に流れるトラフィックを LSP_backup に迂回する。論理回線 A を構成する物理回線に障害が発生し、R1 が LSP1 に流れるトラフィックを LSP_backup に迂回する場合を例に挙げる。論理回線 A を構成する物理回線全てがリンクアップしている場合、LSP1、LSP2、LSP3 のトラフィックはそれぞれ、物理回線番号 1、物理回線番号 2、物理回線番号 3 が出力物理回線として選択される。R1 は、物理回線番号 1 の回線に障害が起きたとき、LSP1 に流れるトラフィックのみを LSP_backup に迂回する。

40

【 0 0 1 7 】

仮に、R1 と R2 が 10Gbps の物理回線 1 本で接続されている場合、該物理回線が障害を起こすと、LSP1、LSP2、LSP3 を流れる全てのトラフィックは、LSP_backup に迂回する。R4 と R5 を接続する回線に LSP_A を通る 8Gbps のトラフィックが流れていた場合、遅延、輻輳が生じてしまう。また、R2 に障害が発生したときは、論理回線 A を構成するすべての物理回線

50

で通信不能になるので、LSP1、LSP2、LSP3に流れるトラフィックすべてがLSP_backupに迂回する。これに対し本実施例では、回線障害に応じてトラフィックをLSP_backupに迂回するので、論理回線Aを構成する物理回線全てに障害発生、又はR2に障害発生しない限りR4とR5を接続する回線における輻輳は発生しない。また、LSP2、LSP3を流れるトラフィックには変化がなく、引き続き最適なパスがトラフィックの転送路として選択される。

【0018】

また、R1はLSP毎に迂回優先度を蓄積して、該迂回優先度に基づきトラフィックを迂回する手段を備えている。例えば、LSP1にはVoIPトラフィック、LSP2にはメールトラフィック、LSP3にはファイル転送トラフィックが流れているとする。LSP2のメールトラフィックは遅延や輻輳に対して敏感ではないので、R1の管理者はLSP2の迂回優先度を最大値に設定する。LSP1を流れるVoIPトラフィックは遅延に弱いので、R1の管理者はLSP1の迂回優先度を最小値に設定する。LSP3の迂回優先度はLSP1とLSP3の迂回優先度の中間に設定する。R1は、論理回線Aを構成する物理回線に障害が発生したとき、健全な物理回線数と迂回優先度を比較し、不等式、(物理回線数) < (LSPの迂回優先度)が成り立つLSPに流れるトラフィックから迂回を開始する。上記例では、R1は遅延の影響が少ないメールトラフィックから迂回を開始し、遅延に弱いVoIPトラフィックは継続して最適パスで転送を続ける。

10

【0019】

このように、R1は、論理回線Aを構成する物理回線に障害が発生したとき、該物理回線が出力回線として選択されるトラフィックを予備LSPに迂回する。該トラフィック以外の前記論理回線に流れるトラフィックに変化は生じないため、最適でないパスを通るトラフィックを最小限に抑えることが出来る。また、障害発生した回線を流れるトラフィックのみを予備LSPに迂回するため、迂回先での輻輳発生(パケット廃棄)を抑えることが出来る。

20

(2) ルータの概要

図2は、R1として適用されるルータ200の構成例を示すブロック図である。ルータ200はNの回線部210-i (i=1からN)と、各回線部210が収容する2Nの物理回線201-ij (i=1からN, j=1 or 2)、回線部210-iを結合する一つの packets 中継処理部250と、一つのプロセッサ280から構成される。回線部210-iは、パケットの送受信処理を行うパケット送受信回路230と、回線障害を検出する回線障害検出部240と、本発明固有の宛先判定部100と、ARPテーブル検索部220とから構成される。以下、図8と図5も用いつつルータ200の動作の概要を説明する。

30

【0020】

図8上部は、物理回線201から入出力されるパケットのフォーマットの一例を示す。本フォーマットは、ヘッダ部810とデータ部820から構成される。データ部820はユーザデータ821より構成される。ヘッダ部810は、データリンク層の送信元アドレスである送信先MACアドレス811 (Destination MAC Address: 以下「DMAC」という)と、送信元アドレスである送信元MACアドレス812 (Source MAC Address: 以下「SMAC」という)と、MPLS網内でパケットの転送先を示すラベル813と、送信元IPアドレス(Source IP address)を示すSIP814と、送信先IPアドレス(Destination IP address)DIP815とから構成される。

40

【0021】

また、図8下部にルータ200の内部パケットフォーマットの一例を示す。本フォーマットは図8上部で前述のパケットフォーマットに内部ヘッダ部830が追加される。この内部ヘッダ部830は、パケットを出力する物理回線の番号である出力回線番号831と、該パケットを出力する際パケットに書き込む出力ラベル832と、ラベル813に対するラベル操作833と、該パケットを受信する次のルータ又は端末のIPアドレスである次ホップIPアドレス834 (Next Hop IP Address: 以下「NHIP」という)とから構成される。ここでラベル操作とは、ラベルを一つスタックするプッシュ(PUSH)、ラベルを一つ外すポップ(POP)、ラベル813と出力ラベル832を交換するスワップ(SWAP)の3種類である。

【0022】

図5は、図2のパケット送受信回路230の構成例を示すブロック図である。物理回線201

50

よりパケットを入力すると、内部ヘッダ付加回路510が内部ヘッダ部830を付加し、本パケットをパケットバッファ520に書き込む。パケットヘッダ送信部540Aは、パケットバッファ520内のパケットヘッダ部810をパケットヘッダ情報21として、図2の宛先判定処部100に送信する。このとき、内部ヘッダの各フィールドは意味の無い値となっており、後にヘッダ書き込み回路550Aにより意味のある値が書き込まれる。

【0023】

パケット送受信回路230よりパケットヘッダ情報21を受信した宛先判定部100は、パケットヘッダ情報21内のパケットヘッダ810の少なくとも一つの情報に基づき宛先判定処理を行い、パケット送受信回路230内のヘッダ書き込み回路550Aにパケット出力回線情報22を送信する。パケットヘッダ810内の情報のとして、MPLSコアルータの場合はラベル813であり、入り口エッジルータの場合は送信先IPアドレス(Destination IP address)DIP815や(Destination MAC address)DMAC811などがある。本実施例では、R1をMPLSコアルータとして、パケットヘッダ810内のラベル813に基づく宛先判定処理について説明をする。

10

【0024】

パケット出力回線情報22は、出力ラベル、ラベル操作、物理回線201に割り振られた回線番号、ネクストホップIPアドレスの少なくとも一部を含む。出力回線が複数の物理回線を集約した論理回線である場合、該物理回線のうち1回線を出力回線として選択する。例えば図10の場合、LER1からラベル20が付加されたパケットを受信すると、次ホップIPアドレスはR2の論理回線Aに割り振られた192.168.0.2、出力ラベル200、出力回線番号は論理回線Aを構成する物理回線に割り振られた番号の中から選択される。出力物理回線番号の選択方法の例として、1から3のいずれかの値を出力するハッシュ関数を用いる方法がある。

20

【0025】

また宛先判定部100は、論理回線を構成する物理回線に障害が発生した場合、保護対象ラベル(保護対象LSPに沿って転送されるパケットに付加されるラベル)が付加されたパケットに予備ラベル(予備LSPに沿って転送されるパケットに付加されるラベル)を付加し、予備LSPに迂回する機能を備えている。

【0026】

回線障害検出部240は、各回線の状態を常時監視している。回線障害検出部240は回線障害を発見すると、障害が発生した回線の物理回線番号を障害発生回線番号情報25として回線部210-1から210-N内の宛先判定部100に送信する。図2中では省略してあるが、回線部210-2から210-Nから全ての回線部210に対し、障害発生回線番号情報25が結線されている。

30

【0027】

宛先判定部100は障害発生回線番号情報25を受信すると、障害発生回線番号情報25が論理回線を構成している物理回線番号であった場合、論理回線を縮退(論理回線を構成する物理回線を減らし、残帯域を減少)させる。宛先判定部100は、入力パケットの出力回線が上記論理回線であり、かつ、上記処理で回線障害と認識した回線が出力回線として選ばれた場合、該パケットに予備ラベルを付加し、予備ラベルに対応する出力回線から出力する。また、宛先判定部100は、保護対象ラベル毎に設定された迂回閾値に基づき、パケット迂回を実行するモードも備える。該モードの詳細については後に説明する。

40

【0028】

パケット送受信回路230のヘッダ書き込み回路550Aはパケット出力回線情報22を受信すると、パケット出力回線情報22内の出力回線番号、出力ラベル、ラベル操作、次ホップIPアドレスをそれぞれ出力回線番号831と出力ラベル832とラベル操作834とNHIP 835の各フィールドに書き込む。パケット読み出し回路560は、パケットバッファ520から蓄積されているパケットを読み出して、パケット中継処理部250へ送信する。

【0029】

パケットを受信したパケット中継処理部250は出力回線番号831に対応する回線部210のパケット送受信回路230に該パケットを送信する。パケット送受信回路230は本パケットをパケットバッファ530に蓄積する。パケットヘッダ送信部540Bは、本パケットの内部ヘッ

50

ダ部830のNHIP834を次ホップIPアドレス情報23としてARPテーブル検索部220に送信する。

【 0 0 3 0 】

ARPテーブル検索部220は次ホップIPアドレス情報23に対応する次ホップルータのMACアドレスを備え、次ホップIPアドレス情報23を受信すると対応するMACアドレスを次ホップMACアドレス情報24としてヘッダ書き込み回路550Bに送信する。ヘッダ書き込み回路550Bは本情報24内のMACアドレス、出力回線番号831に対応する物理回線に割り当てられたMACアドレスをそれぞれ、パケットバッファ530内のパケットのDMAC811、SMAC812に書き込む。該出力物理回線が論理回線の一部である場合、該論理回線に割り振られたSMACがSMAC812に書き込まれる。

【 0 0 3 1 】

続いて、ラベル操作833に従いラベル813の書換えを行う。ラベル操作833がSWAPを示す場合には、ラベル813を出力ラベル833で上書きをする。PUSHの場合はSMAC812とラベル813の間に出力ラベル832を挿入する。POPの場合は、単にラベル813を削除し、SMAC812の直後にデータ821を移動させる。最後に、パケット読み出し回路560は、内部ヘッダ部830を削除し、蓄積したパケットを出力回線番号831に対応する入出力回線201に送信する。

【 0 0 3 2 】

以上に示したように、宛先判定部100は、論理回線を構成する物理回線に障害が発生したとき、該論理回線を縮退させる。入力パケットの出力回線に、上記障害発生した回線が選ばれた場合、該パケットに予備ラベルを付加し、予備ラベルに対応する物理回線又は論理回線から出力することで、出力先を予備ラベルに対応する宛先へ変更することが出来る。複数の保護対象ラベルの出力先が論理回線に設定されているとき、予備LSPへ迂回するトラフィックを障害発生した回線に応じて分割できる。このため、最適でないパスを流れるトラフィックを最小限に抑えることができる。

【 0 0 3 3 】

(3) トラフィック一部迂回方式

図 1 は、図 2 の宛先判定部100の構成例を示すブロック図である。宛先判定部100は、出力ラベル検索部101、出力物理回線番号検索部102、入力ラベル蓄積部103、出力ラベル検索起動部104、迂回判定部105から構成される。

【 0 0 3 4 】

図 3 は、図 1 の出力ラベル検索部の構成例を示すブロック図である。出力ラベル検索部101は、保護対象LSPに対応する出力ラベルテーブル300A、予備LSPに対応する出力ラベルテーブル300B、これらのテーブルの検索制御を行う出力ラベルテーブル制御部310より構成される。出力ラベルテーブル300A、Bの一例を図 6 に示す。出力ラベルテーブル300Aは、入力ラベルに対応する出力ラベル、ラベル操作、NHIP、出力回線番号、迂回閾値、論理回線ビットより構成される。迂回閾値はLSPごとの迂回閾値を表す。迂回閾値の使用法については後に説明する。論理回線ビットは、出力回線が複数の物理回線を束ねた論理回線であることを示す。論理回線ビットが ' 1 ' のとき、出力回線番号は論理回線番号になる。論理回線ビットが ' 0 ' のとき、出力回線番号は入出力物理回線に割り振られた番号になる。出力ラベルテーブル300Bは、出力ラベルテーブル300Aの構成要素より論理回線ビットの列を除いたものに等しい。

【 0 0 3 5 】

宛先判定部100は、パケットヘッダ情報21を受信すると、入力ラベル813の情報を入力ラベル蓄積部103に蓄積する。次に、出力ラベル検索起動部104が出力ラベル検索部101に検索指示し、入力ラベル蓄積部内103の入力ラベル813を検索キーとして送信する。

【 0 0 3 6 】

出力ラベル検索部101の出力ラベルテーブル制御部310は、出力ラベル検索起動部104からの出力ラベルテーブル検索指示と入力ラベル813を受信すると、入力ラベル813の値を出力ラベルテーブル300A、出力ラベルテーブル300Bのエントリのアドレスとして扱い、該アドレスのエントリを読み出す。このとき、出力ラベルテーブル300A、出力ラベルテーブル300Bの両エントリ情報を読み出す。例えば、入力ラベルの値が20であった場合、出力ラベ

10

20

30

40

50

ルテーブル300A、出力ラベルテーブル300Bのアドレス20のエントリを読み出す。出力ラベルテーブル300A、出力ラベルテーブル300Bから読み出したエントリ情報、有効ビット、出力ラベル、ラベル操作、NHIP、出力回線番号、迂回閾値、論理回線ビットを迂回判定部105に送信する。

【0037】

迂回判定部105は、出力ラベル検索部101から受信したラベルテーブル300A、出力ラベルテーブル300Bのエントリ情報の有効ビットを確認する。有効ビットが‘0’を示す場合、エントリ情報は無効であることを示す。両テーブルのエントリ情報が無効である場合、出力先無しとして、入力パケットは最終的に廃棄される。出力ラベルテーブル300Bのエントリ情報が無効であった場合、入力パケットに対する予備ラベルは存在しないため、迂回パスも存在しないことになる。出力ラベルテーブル300Aのエントリ情報が無効で出力ラベルテーブル300Bのエントリ情報が有効の場合、入力パケットの出力先は出力ラベルテーブル300Bのエントリ情報に従う。つまり迂回パスへ出力するよう判定する。

10

【0038】

出力ラベルテーブル300A又はBのエントリ情報が有効である場合、迂回判定部105は、該エントリ情報の論理回線ビットが‘1’であるか確認する。出力ラベルテーブル300A又はBのエントリ情報の論理回線ビットが‘1’を示している場合、出力回線番号は論理回線であることを示している。迂回判定部105は、出力回線番号が示す論理回線を構成する物理回線の中から出力回線を決定するため、出力回線検索部102に検索指示と出力ラベルテーブル300のエントリ情報の出力回線番号を送信する。迂回判定部105は出力回線検索部105に検索指示を出すとともに、物理回線を選択する情報として、例えば、入力パケットのヘッダ部810を送信する。以上の出力回線検索部105への検索指示は、出力ラベルテーブル300A及びBのどちらか一方、又は、両方のエントリ情報の論理回線ビットが‘1’であった場合に行う。本実施例では、出力ラベルテーブル300Aの検索結果の出力回線が、論理回線の場合について説明を行う。

20

【0039】

次に、図1の出力回線検索部102の構成例を図4に示す。出力物理回線番号検索部102は、物理回線テーブル410と物理回線テーブル制御部400とより構成される。

【0040】

図4の物理回線テーブル410の一例を図7に示す。物理回線テーブル410は、論理回線を構成する物理回線を管理しているテーブルである。物理回線テーブル410は、エントリの有効ビット、論理回線番号、論理回線を構成する物理回線の数を示す物理回線数、論理回線を構成する物理回線の番号が記録されている物理回線番号フィールドより構成される。物理回線番号フィールドの各フィールドには物理回線番号が蓄積されている。物理回線番号フィールドに蓄積されている物理回線番号が‘0’である場合、物理回線に割り振られる番号は1から2Nなので、該出力回線は障害により通信不能であることを示す。

30

【0041】

物理回線テーブル制御部400は、検索指示と入力パケットのヘッダ部810を受信すると、テーブルエントリをアドレスが小さい順に読み出す。読出したエントリの論理回線番号と迂回判定部105から受信した出力回線番号を比較し、最初に一致したエントリを検索結果として採用する。次に該エントリの物理回線番号フィールドの中から、一つの物理回線回線番号を判定する。物理回線回線番号を判定する方法としては、例えば、ハッシュ関数 $H_n(x)$ を用いた方式が考えられる。 $H_n(x)$ は引数 x に対してハッシュ値として1から n の中から一つの整数を出力する。出力回線テーブル制御部400は、物理回線テーブル410の検索により得られた有効な出力回線数である‘ n ’と対応したハッシュ関数 $H_n(x)$ を選択する。例えば、出力回線数‘3’の場合にはハッシュ関数 $H_3(x)$ を選択する。

40

【0042】

本実施例では $H_n(x)$ の引数として、迂回判定部105から受信したパケットヘッダ部810内のラベル813を使用する。出力回線テーブル制御部400は、 $H_n(x)$ 結果が指す前記処理で読み出した物理回線番号フィールドに格納されている物理回線番号を出力回線番号として選

50

折する。出力回線テーブル制御部400は、出力回線番号が決定すると、迂回判定部105に前記処理で読み出した出力回線テーブル410のエントリの物理回線数と、物理回線番号を送信する。

【0043】

迂回判定部105は、物理回線テーブル制御部400から出力回線番号と出力回線数を受信すると、該出力回線番号が‘0’であるか判定する。該出力回線番号が‘0’であった場合、該出力回線番号に該当する物理回線は障害により通信不能であるものとして、パケット出力回線情報22にラベルテーブル300Bの検索結果を採用する。該出力回線番号が‘0’以外の値であった場合、出力ラベル検索部101から受信したラベルテーブル300Aのエントリ情報をパケット出力回線情報22に採用する。迂回判定部105は、上記処理で採用したラベル

10

【0044】

以上に宛先判定部100におけるトラフィック迂回方法を説明した。次に、宛先判定部100における物理回線障害発生時の処理について説明を行う。

【0045】

宛先判定部100内の出力回線テーブル制御部400は、障害回線番号情報25を受信すると、該障害回線番号情報25をキーに出力回線テーブル410の検索を行う。出力回線テーブル410の出力回線番号フィールド内に該回線番号を発見すると、該出力回線番号フィールドに蓄積されている出力回線番号を‘0’に書き換える。上記処理は、迂回判定部105での迂回判定処理において、前記障害回線番号情報25に相当する物理回線が、障害により通信不能であると判定するための処理である。例えば、物理回線番号‘1’が通信不能になったとき、図7に示す物理回線テーブル410のエントリ410-1の物理回線番号フィールド‘1’に蓄積されている値を‘0’にする。

20

【0046】

以上に示したように宛先判定部100は、保護対象LSPに相当する出力ラベルテーブル300Aと、予備LSPに相当する出力ラベルテーブル300Bを備える。読み出した出力ラベルテーブル300Aのエントリ情報の出力回線が論理回線のとき、論理回線を構成する物理回線の一つを出力回線として選択する。該物理回線に障害が発生していた場合、出力ラベルテーブル300Bのエントリ情報を入力パケットの宛先とし、パケットの迂回を実現する。このためルータ200は、論理回線Aを構成する物理回線に障害が発生したとき、該物理回線が出力回線として選択されるパケットを、予備ラベルに対応する健全な回線から出力することができる。該パケット以外の前記論理回線に流れるパケットに変化は生じないため、最適でないパスを通るトラフィックを最小限に抑えることができる。また、障害発生により通信不能となった回線を流れるトラフィックのみを予備LSPに迂回するため、迂回先での輻輳発生（パケット廃棄）を抑えることができる。また、ネクストホップのノード障害時は、該ネクストホップノードに接続されている全ての物理回線の物理回線番号が、物理回線テーブル410において‘0’になるため、該論理回線を流れるすべてのトラフィックは予備LSPに迂回する。

30

40

【0047】

ルータ200の管理者は、ルータ200外部の管理端末10から出力ラベルテーブル300の設定を行う。出力ラベルテーブル300A、Bに設定する際に、管理端末10に入力されるコマンドの一例を図9に示す。図9のコマンド901は、入力ラベル20に対するコンフィグを開始することをルータ200の宛先判定部100に伝える。コマンド910は入力ラベル20に対する宛先を設定する。字句911は入力ラベル20に対し出力ラベル200をSWAPすることを宣言している。字句912は入力ラベル20のネクストホップIPアドレスは192.168.0.2であることを宣言している。字句913は出力回線の設定で、論理回線la1が出力回線であることを宣言している。コマンド910を制御端末10から入力すると、プロセッサ280を介してコマンドを受け取った、出力ラベルテーブル310は出力ラベルテーブル300Aの出力ラベルに200、ラベル操作に

50

SWAP、NHIPに192.168.0.2、出力回線番号に‘1’、論理回線ビットに‘1’、有効ビットに‘1’を書き込む。次にコマンド920は、入力ラベル20に対する予備LSPを設定する。字句921は予備LSPを設定開始の字句である。字句922は出力ラベル2000であることを宣言している。字句923は入力ラベル20のネクストホップは192.168.5.2であることを宣言している。字句924は出力回線番号が‘4’であることを宣言している。コマンド920を制御端末10から入力すると、プロセッサ280を介してコマンドを受け取った、出力ラベルテーブル310は出力ラベルテーブル300Bの出力ラベルに2000、ラベル操作にSWAP、NHIPに192.168.5.2、出力回線番号に‘4’、論理回線ビットに‘0’、有効ビットを‘1’を書き込む。

【0048】

(4) 優先度付き一部迂回方式

次に、宛先判定部100における出力ラベルテーブル300Aの各エントリに蓄積する迂回閾値に基づくトラフィック迂回判定の実施例について説明する。該トラフィック迂回法は、論理回線を構成する物理回線に障害が発生したとき、トラフィックの特性に応じてトラフィック迂回の順序を変更できる。例えば、遅延に強いトラフィックから迂回を開始し、遅延に弱いフィックは継続して最適パスで転送を続けることができる。宛先判定部100は、迂回に関するモード蓄積部106を備える。宛先判定部100は、迂回に関する2値のモード情報が蓄積されているモード蓄積部106を備える。該モード情報は、出力回線テーブル制御部400の動作モードを表現している。出力回線テーブル制御部400はモード蓄積部106に蓄積されたモード情報を参照する。出力物理回線番号検索部102と出力回線テーブル制御部400は、モード蓄積部106に蓄積されたモード情報に基づきトラフィック一部迂回方式、または、優先度付き一部迂回方式で動作するかを判断する。

【0049】

まず、物理線検索部102における回線障害発生時の処理を説明する。出力物理回線番号検索部102内の出力回線テーブル制御部400は、障害回線番号情報25を受信すると、該回線番号を検索キーに物理回線テーブル410の検索する。物理回線テーブル410のアドレスの小さい順に各エントリを読み出し、該エントリの物理回線番号フィールドに蓄積している出力回線番号と障害回線番号情報25を比較する。障害回線番号情報25と、物理回線番号フィールドに蓄積している物理回線番号が一致した場合次の処理を行う。アドレスNの出力物理回線番号フィールドMに障害回線番号情報25と一致する物理回線番号が蓄積されていたとする。アドレスNのエントリの物理回線番号数が指す物理回線番号フィールド、つまり、最後の物理回線番号フィールドの物理回線番号を読み出す。続いて、物理回線番号フィールドの最後の値をアドレスNの出力回線番号フィールドMに書きこみ、出力回線番号フィールドの最後の値に‘0’を書き込む。アドレスNの出力回線番号数から‘1’を引く。上記処理を行うことで、論理回線を構成する物理回線を減らし、迂回判定部105において、迂回閾値に基づいたトラフィック迂回処理を行うことができる。

【0050】

次に、迂回判定部105におけるトラフィック迂回判定処理について説明する。迂回判定部105は、物理回線テーブル制御部400から物理回線番号と物理回線数を受信すると、該物理回線数と、出力ラベルテーブル制御部310から受信した、出力ラベルテーブル300Aのエントリ情報の迂回閾値との比較を行う。該物理回線数が該迂回閾値を下回っていた(出力回線数<迂回閾値)場合、出力ラベルテーブル300Bのエントリ情報をパケット出力回線情報22としてパケット送受信回路230に送信する。該出力回線数が該迂回閾値と同等以上(出力回線数 ≥ 迂回閾値)の場合、出力ラベルテーブル300Aの検索結果をパケット出力回線情報22としてパケット送受信回路230に送信する。物理回線数テーブル制御部400から受信した物理回線数は、論理回線の残帯域又は残物理回線本数を表現している。出力ラベルテーブル300Aに設定した迂回閾値と、該物理回線数を比較することで、パケット出力回線情報22に、出力ラベルテーブル300A又はBの検索結果を採用するか決定する。例えば、図6の出力ラベルテーブル300の状態のとき、ラベル20が付加されたパケットを受信した場合、迂回閾値は3になる。図7の物理回線テーブル410では物理回線数が3なので(出力回線数 ≥ 迂回閾値)の条件が成立し、出力物理回線は‘1’、‘2’、‘3’のどれかになる。また、同条件で

10

20

30

40

50

ラベル21の packets を受信したときは(出力回線数 < 迂回閾値)が成立するので、出力回線は出力ラベルテーブル300Bのアドレス21エントリの出力回線番号`4`が採用される。

【 0 0 5 1 】

なお、選んだ物理回線の有効ビットを逐次見る方法(一部迂回方式)と、有効な物理回線の数と迂回優先度とを比較する方法(優先度付き一部迂回方式)とは、本実施例ではモードを切り替えることとし同時に実現していないが、例えば物理回線テーブル410のエントリ毎にモードビットを持たせれば同時に実現できる。ただし、この場合、モードが論理回線に括り付けになり、トラフィック(フロー)毎のモードの変更はできない。

(5) ユーザーインターフェース また、迂回閾値をルータ200に設定する際の、管理端末10に入力されるコマンドの一例を図9に示す。図9のコマンド931は、入力ラベル20に対するコンフィグを開始することをルータ200の宛先判定部100に伝える。コマンド940は入力ラベル20に対する宛先を設定する。字句941は入力ラベル20に対し出力ラベル200をSWAPすることを宣言している。字句942は入力ラベル20のネクストホップIPアドレスは192.168.0.2であることを宣言している。字句943は出力回線の設定で、論理回線論理回線1が出力回線であることを宣言している。字句944は迂回閾値が`10`であることを宣言している。コマンド940を制御端末10から入力すると、プロセッサ280を介してコマンドを受け取った、出力ラベルテーブル310は出力ラベルテーブル300Aの出力ラベルに200、ラベル操作にSWAP、NHIPに192.168.0.2、出力回線に論理回線番号`1`、迂回閾値に`3`、論理回線ビットに`1`、有効ビットに`1`を書き込む。次にコマンド950は、入力ラベル20に対する予備LSPを設定する。字句951は予備LSPを設定開始することを宣言している。字句952は出力ラベル200であることを宣言している。字句953は入力ラベル20のネクストホップは192.168.5.2であることを宣言している。字句954は出力回線番号が物理回線番号4番であることを宣言している。コマンド920を制御端末10から入力すると、プロセッサ280を介してコマンドを受け取った、出力ラベルテーブル310は出力ラベルテーブル300Bの出力ラベルに2000、ラベル操作にSWAP、NHIPに192.168.5.2、出力回線番号に`4`、論理回線ビットに`0`、有効ビットを`1`を書き込む。

【 図面の簡単な説明 】

【 0 0 5 2 】

【 図 1 】 図 2 の宛先判定部100の構成例を示すブロック図。

【 図 2 】 図 1 のR1に相当するルータ200の構成例を示すブロック図。

【 図 3 】 図 1 の出力ラベル検索部101の構成例を示すブロック図。

【 図 4 】 図 1 の出力物理回線番号検索部102の構成例を示すブロック図。

【 図 5 】 図 1 のパケット送受信回路230の構成例を示すブロック図。

【 図 6 】 図 3 の出力ラベルテーブル300A、Bの一例を示すテーブル図。

【 図 7 】 図 4 の物理回線テーブル410の一例を示すテーブル図。

【 図 8 】 ルータ200が送受信するパケット、及び、ルータ200内部のパケットフォーマットの一例。

【 図 9 】 出力ラベルテーブル300A、Bを設定するための入力コマンドの一例。

【 図 1 0 】 MPLS網の一例。

【 符号の説明 】

【 0 0 5 3 】

100...宛先判定部

101...出力ラベル検索部

102...物理回線検索部

200...ルータ

230...パケット受信回路

240...回線障害検出部

300...出力ラベルテーブル300

410...物理回線テーブル

21...パケットヘッダ情報

10

20

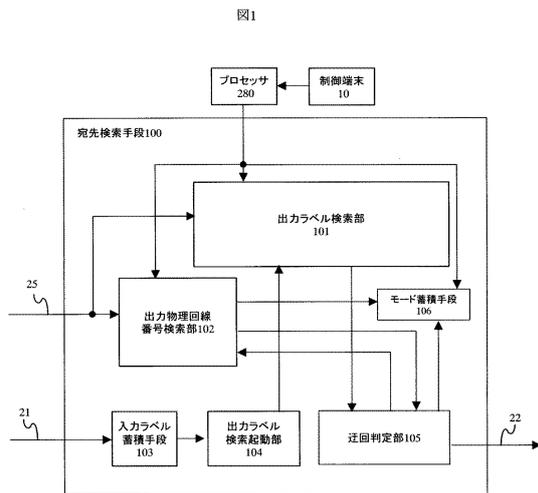
30

40

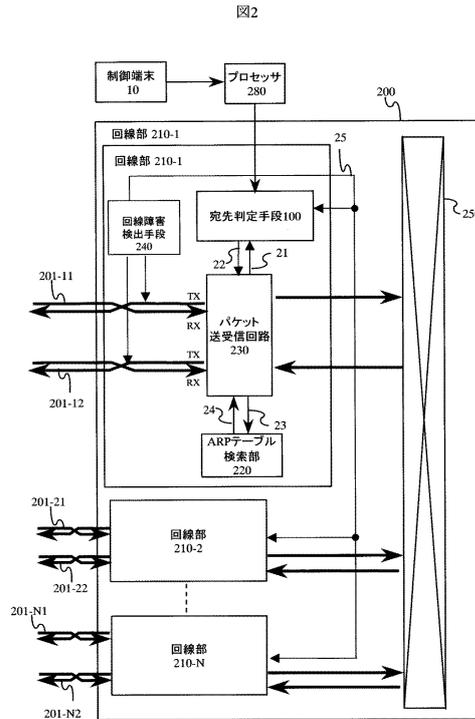
50

- 22... パケット出力回線情報
- 23... 次ホップIPアドレス情報
- 24... MACアドレス情報
- 25... 障害回線番号情報25

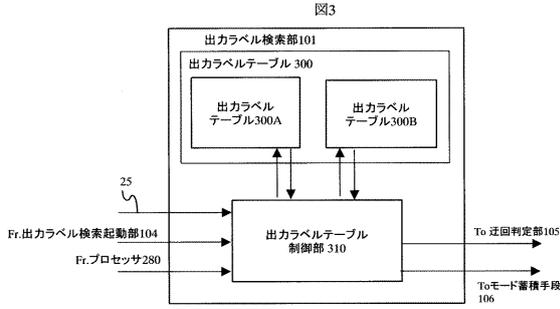
【図1】



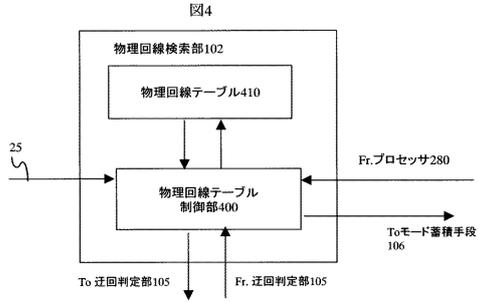
【図2】



【図3】



【図4】



【図6】

図6

出力ラベルテーブル300A

アドレス	有効ビット	出力ラベル	ラベル操作	NHIP	出力回線		論理回線 ビット
					番号	迂回値	
0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
.
20	1	200	SWAP	192.168.0.2	1	3	1
21	1	210	SWAP	192.168.0.2	1	2	1
22	1	220	SWAP	192.168.0.2	1	1	1
.
.

出力ラベルテーブル300B

アドレス	有効ビット	出力ラベル	ラベル操作	NHIP	出力回線		論理回線 ビット
					番号	迂回値	
0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
.
20	1	2000	PUSH	192.168.5.2	4	0	0
21	1	2100	PUSH	192.168.5.2	4	0	0
22	1	2200	PUSH	192.168.5.2	4	0	0
.
.

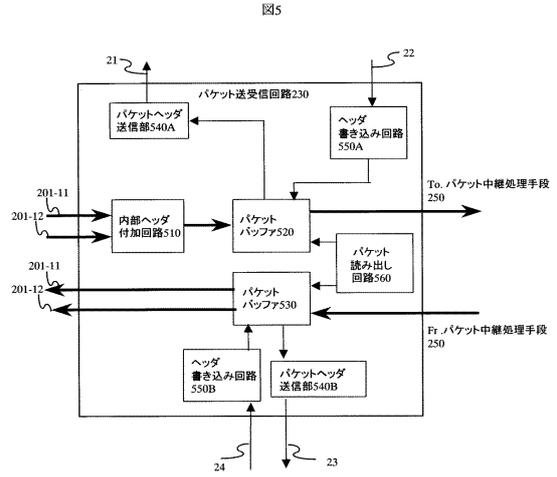
【図7】

図7

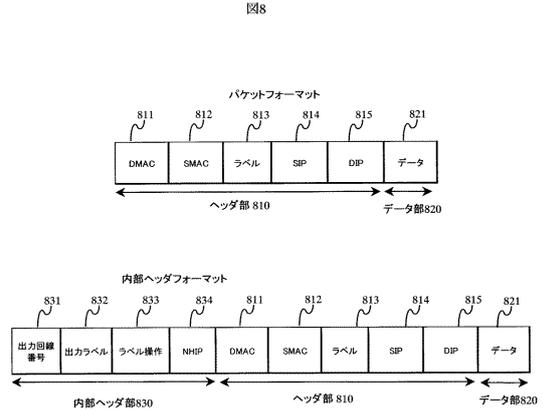
物理回線テーブル410

アドレス	有効ビット	物理回線 番号	物理回線数	物理回線番号フィールド				
				1	2	3	4	5
0	1	1	10	1	2	3	0	0
1	0	2	0	0	0	0	0	0
2	0	3	0	0	0	0	0	0
3	0	4	0	0	0	0	0	0
.
.

【図5】



【図8】



【 図 9 】

図9

```

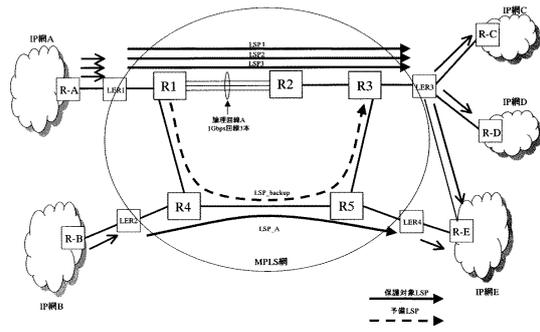
Config# mpls in_label 20
プロンプト      901
Config(in_label20)# out_label 200 swap nhip 192.168.0.2 out_if la1
プロンプト      911      912      913
                  910
Config(in_label20)# backup_lsp out_label 2000 push nhip 192.168.5.2 out_if 4
プロンプト      921      922      923      924
                  920

Config# mpls in_label 20
プロンプト      931
Config(in_label20)# out_label 200 swap nhip 192.168.0.2 out_if la1 threshold 10
プロンプト      941      942      943      944
                  940
Config(in_label20)# backup_lsp out_label 2000 push nhip 192.168.5.2 out_if 4
プロンプト      951      952      953      954
                  950

```

【 図 10 】

図10



フロントページの続き

- (56)参考文献 特開平5 - 191434 (JP, A)
特開昭63 - 94741 (JP, A)
特開2000 - 174818 (JP, A)
特開2003 - 78552 (JP, A)
国際公開第2004 / 086697 (WO, A1)

- (58)調査した分野(Int.Cl., DB名)
H04L 12