

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4884525号
(P4884525)

(45) 発行日 平成24年2月29日(2012.2.29)

(24) 登録日 平成23年12月16日(2011.12.16)

(51) Int.Cl. F I
 HO 4 L 12/46 (2006.01) HO 4 L 12/46 V
 HO 4 L 12/56 (2006.01) HO 4 L 12/56 I O O Z

請求項の数 3 (全 25 頁)

(21) 出願番号	特願2009-503814 (P2009-503814)	(73) 特許権者	000005223 富士通株式会社 神奈川県川崎市中原区上小田中4丁目1番1号
(86) (22) 出願日	平成19年3月13日(2007.3.13)	(74) 代理人	100070150 弁理士 伊東 忠彦
(86) 国際出願番号	PCT/JP2007/054952	(72) 発明者	西村 和人 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内
(87) 国際公開番号	W02008/111173	審査官	中木 努
(87) 国際公開日	平成20年9月18日(2008.9.18)		
審査請求日	平成21年9月7日(2009.9.7)		

最終頁に続く

(54) 【発明の名称】 通信経路制御方法、通信装置及び通信システム

(57) 【特許請求の範囲】

【請求項1】

識別情報を付して通信網に向かって送信された送信パケットを受信した各ノードが当該送信パケットを受信した経路を当該送信パケットが有する識別情報について学習することにより以後当該識別情報に係るパケットについて当該学習に係る経路を適用し、同じ識別情報を有する送信パケットが複数受信された場合には最初に受信された送信パケットについてのみ学習を行う通信経路制御方法であって、

前記通信網に設けられた各ノードにおいて、送信パケットに係る所定の条件に応じて当該送信パケットの送出タイミングを制御する段階を有し、

前記送信パケットの送出タイミングの制御は当該送信パケットに対する所定の遅延の付与であり、前記所定の遅延は当該送信パケットに係る設定帯域に応じて、当該設定帯域が広いほど小さい遅延である、通信経路制御方法。

10

【請求項2】

識別情報を付して送信された送信パケットを受信した際に受信した経路を当該送信パケットが有する前記識別情報について学習する学習手段と、

当該識別情報に係るパケットについて当該学習に係る経路を適用する経路決定手段とを有し、

前記学習手段は同じ識別情報を有する送信パケットが複数受信された場合には最初に受信された送信パケットについてのみ学習を行い、

更に当該送信パケットを送出する際、当該送信パケットに係る所定の条件に応じて当該

20

送信パケットの送出タイミングを制御する送出態様制御手段を有し、

前記送出態様制御手段による前記送信パケットの送出タイミングの制御は当該送信パケットに対する所定の遅延の付与であり、前記所定の遅延は当該送信パケットに係る設定帯域に応じて、当該設定帯域が広いほど小さい遅延である、通信装置。

【請求項 3】

複数の通信装置が通信網によって相互に接続された通信システムであって、

前記複数の通信装置の各々は、識別情報を付して送信された送信パケットを受信した際に受信した経路を当該送信パケットが有する前記識別情報について学習する学習手段と、当該識別情報に係るパケットについて当該学習に係る経路を適用する経路決定手段とを有し、

前記学習手段は同じ識別情報を有する送信パケットが複数受信された場合には最初に受信された送信パケットについてのみ学習を行い、

更に当該送信パケットを送出する際、当該送信パケットに係る所定の条件に応じて当該送信パケットの送出タイミングを制御する送出態様制御手段を有し、

前記送出態様制御手段による前記送信パケットの送出タイミングの制御は当該送信パケットに対する所定の遅延の付与であり、前記所定の遅延は当該送信パケットに係る設定帯域に応じて、当該設定帯域が広いほど小さい遅延である、通信システム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は通信経路制御方法、通信装置及び通信システムに関する。

【背景技術】

【0002】

イーサネット網等のパケット通信網におけるパケットフォワーディングにおいて通信経路を決定する方法として、いわゆる先着ポート学習方式がある。

【0003】

この方式は例えば特許文献 1、2 に開示されている。

【0004】

この方式では、通信経路が未確定のデータパケット（すなわち未学習パケット）をいわゆるフラッディングの手法により当該通信網に対し発信し、これを受けた各ノードにおいて最も早く到着したデータパケットについて、その発信元の MAC アドレス等を学習（すなわちアドレス学習）し、それ以外の後着のデータパケットを廃棄する。このようにして学習された先着データパケットの経路が以後当該データパケットの発信元の MAC アドレスに係るデータパケットの通信経路として適用される。

【特許文献 1】特開平 3 - 162035 号

【特許文献 2】特開平 6 - 37764 号

【特許文献 3】特開 2003 - 157198 号

【発明の開示】

【発明が解決しようとする課題】

【0005】

この方式によればデータパケットの先着性のみにより自動的に経路が決定されるため、人為的な経路操作やリンクの帯域を考慮した経路操作を行うことができなかった。

【0006】

本発明は上記先着ポート学習方式によるパケットフォワーディングにおいて、人為的な経路操作やリンクの帯域を考慮した経路操作を行えるようにすることによって効果的な負荷分散を実現し得る構成を提供することを目的とする。

【課題を解決するための手段】

【0007】

本発明によれば、通信網に設けられた各ノードにおいて、受信データパケットを送出する際、当該データパケットに係る所定の条件に応じて当該データパケットの送出態様を制

10

20

30

40

50

御するようにした。

【0008】

ここで上記データパケットの送出態様の制御とは、当該データパケットに対し適宜遅延を付加したり、或いは当該データパケットを廃棄したりする処理を含む。

【0009】

このように送信元からフラッディングされたデータパケットの送出態様がこれを受信した各ノードにおいて制御されることにより、それに応じてその後当該パケットを受信するノードにおいて当該データパケットを受信するタイミングが適宜変化される。あるいはフラッディングされたデータパケットを受信した各ノードの内部において、到着タイミングが適宜変化される。

10

【0010】

このようにして、フラッディングにより様々な経路を通じて転送されるデータパケットの到着時間が適宜制御される。その結果特定の経路を通じて転送されるデータパケットが宛先に最先に到着しないように制御することが可能となり、もってそれ以外の経路を通じて転送されるデータパケットが最先に到着することとなり、当該経路が以降当該データパケットのMACアドレス等を有するデータパケットの通信経路として採用されるようになる。したがって人為的な経路操作が可能となる。

【発明の効果】

【0011】

このように本発明によれば先着ポート学習方式によるパケットフォワーディングにおいて人為的な経路操作が可能となるため、リンクの帯域を考慮した効果的な負荷分散が可能となる。

20

【図面の簡単な説明】

【0012】

【図1】従来技術の問題点を説明するための図(その1)である。

【図2】従来技術の問題点を説明するための図(その2)である。

【図3】先着ポート学習方式について説明するための図(その1)である。

【図4】先着ポート学習方式について説明するための図(その2)である。

【図5】先着ポート学習方式における問題点を説明するための図(その1)である。

【図6】先着ポート学習方式における問題点を説明するための図(その2)である。

30

【図7】先着ポート学習方式における問題点を説明するための図(その3)である。

【図8】本発明の第1実施例による方式を適用した通信装置の構成を説明するための機能ブロック図である。

【図9】図8に示されるフィルタリング/遅延処理部の機能の例を説明するための図(その1)である。

【図10】図8に示されるフィルタリング/遅延処理部の機能の他の例を説明するための図(その2)である。

【図11】図8に示されるフィルタリング/遅延処理部の機能であって図9, 図10, 図14及び図15に示された機能を実現するための構成として適用可能な遅延挿入部の機能を説明するための図である。

40

【図12】図8に示されるフィルタリング/遅延処理部の機能の更に他の例を説明するための図である。

【図13】図8に示されるフィルタリング/遅延処理部の機能の更に他の例を説明するための図である。

【図14】図8に示されるフィルタリング/遅延処理部の機能の更に他の例を説明するための図である。

【図15】図8に示されるフィルタリング/遅延処理部の機能の更に他の例を説明するための図である。

【図16A】図8に示される通信装置の図9及び図14に示された機能による具体的な動作例を説明するための図(その1)である。

50

【図16B】図8に示される通信装置の図9及び図14に示された機能による具体的な動作例を説明するための図(その2)である。

【図16C】図8に示される通信装置の図9及び図14に示された機能による具体的な動作例を説明するための図(その3)である。

【図17A】図8に示される通信装置の図10及び図15に示された機能による具体的な動作例を説明するための図(その1)である。

【図17B】図8に示される通信装置の図10及び図15に示された機能による具体的な動作例を説明するための図(その2)である。

【図17C】図8に示される通信装置の図10及び図15に示された機能による具体的な動作例を説明するための図(その3)である。

10

【符号の説明】

【0013】

- 1 パケット送受信部
- 2 フィルタリング/遅延処理部
- 3 パケット処理部
- 4 MACテーブル(アドレステーブル)
- 5 同一性情報テーブル

【発明を実施するための最良の形態】

【0014】

本発明の実施例はパケット通信網上の各ノードにおいて、例えばあるポートに係る特定の packets に対し意図的に遅延を挿入したり或いはこれを廃棄することにより、当該ポートを先着候補から除外されやすくすることによって経路の操作を行う等の構成を設ける。又この場合リンクの速度に鑑みて自動的に遅延を挿入するように構成することにより、より大きな帯域の経路を選択され易くし、もって効果的な負荷分散を図ることを可能にする。

20

【0015】

ここで本発明の実施例の構成の理解を容易にする目的で以下に関連する技術分野についての説明を行う。

【0016】

近年広域イーサネットサービスの普及に伴い、イーサネットのMAN, WAN領域への進出がめざましい。レイヤ2スイッチ(L2SW)はルータと比べて安価であり、上位のプロトコルを選ばない等の方式的なメリットも大きく、今後さらに普及するものと考えられていれる。しかしながらL2SWは元々LANの技術でありWAN領域での使用を前提としていないため、以下の如くの方式的な問題点を内包する。

30

【0017】

イーサネットの標準において唯一ネットワーク冗長方式として規定されているものとしてスパンニングプロトコル(以下STPと略称する)がある。このプロトコルはネットワーク上に物理的にループが存在する場合、論理的にポートをブロックすることにより論理的なツリーネットワークを構築することを可能にする。

【0018】

しかしながらこのSTPを用いた場合、データパケット(以下単にパケットと称する)は論理的なツリー上しか通ることができないため、例えば図1に示される如く、パケットは論理的なブロック(ブロッキングポート)により遮られ最適な経路を通れず、またこの場合ノードB-C間のリンクが全く使用されないため効果的な負荷分散がなし得ない。

40

【0019】

このような問題を解決するためMSTPと称されるSTPの拡張方式が考えられている。この方式ではネットワーク上に複数のツリー(STI)を構築し、VLAN毎に使用するツリーを変更することにより、負荷分散を図る。

【0020】

しかしながらこの方式はVLANとSTIとのマッピングによるため、例えば図2に示

50

される如くノード A , B、C の 3 拠点に同じ V L A N 1 に属するユーザが存在するような場合、どのツリーを使っても最適な経路が通れないようなパターンが生ずる。すなわち S T 1 を使うと端末 1 と端末 2 との間の経路が遠回りとなり、S T 2 を使うと端末 2 と端末 3 との間の経路が遠回りとなる。

【 0 0 2 1 】

また S T P には B P D U (ブリッジ・プロトコル・データ・ユニット) の生成や解釈などのソフトウェア処理が要される。その結果ノードの C P U 負荷が大きくなる。又 M S T P の場合複数の S T I 情報を処理しなければならないためさらに負荷が高くなる。又 C P U に障害が発生した場合にプロトコル動作がうまく働かず、ループが形成されてしまうという可能性もある。

10

【 0 0 2 2 】

また S T P の B P D U にはメッセージ・エイジというフィールドが設けられ、一定ノード以上経由した B P D U は廃棄される構成とされている。S T P の標準ではいくら拡張してもホップ数は 2 0 ホップが限度であるため、それ以上の規模のネットワークを構築できないというスケラビリティの問題も内在している。

【 0 0 2 3 】

上記先着ポート学習方式はこの様な問題点を解決する方法として考えられた。

【 0 0 2 4 】

この先着ポート学習方式によれば、図 3 に示す如く、(i) 宛先アドレス (すなわち宛先識別子、以下同様。具体的には M A C アドレス等) が未学習の packets がノードに到着すると、そのノードは到着ポート以外のポートに当該 packets をフラッディングする。

20

【 0 0 2 5 】

(i i) あるノードにおいて同一 packets が複数のポートから到着した場合、最も早く到着した packets についてのみについてフォーワーディング処理を行う。すなわち当該ノードにおいて当該 packets が未学習の場合、同一の packets が到着していないポートに対して当該 packets のフラッディングを行う。また、このとき最も早く到着したポート (先着ポート) をその packets の送信元アドレス (すなわち送信元識別子、以下同様。具体的には M A C アドレス等) とともに学習する。学習はアドレステーブル (すなわち M A C テーブル等、図 3 参照。以下同様) を使用する。

【 0 0 2 6 】

図 3 の場合、送信元アドレス (S A) が x、宛先アドレス (D A) が y の packets が送信元の端末 1 (M A C アドレス : x) からフラッディングされ、これがノード A 経由でポート 3 から、及びノード B ~ ノード C 経由でポート 2 からノード C に到着している。

30

【 0 0 2 7 】

この場合ノード B ~ ノード C 経由のものが先着であったと仮定すると、図示の如く、先着ポートであるポート 2 が当該 packets の送信元アドレス x について学習される。

【 0 0 2 8 】

又ノード B では送信元アドレス x の未学習 packets がポート 1 から先着しているため、当該送信元アドレス x についてポート 1 が学習されている (ノード B につき、アドレステーブルの図示を省略) 。

40

【 0 0 2 9 】

(i i i) それ以外の後着 packets は全て廃棄する。

【 0 0 3 0 】

図 3 の例の場合、ノード A 経由の後着 packets はノード C において廃棄される。

【 0 0 3 1 】

またその後図 4 に示す如く、端末 2 (M A C アドレス : y) から送信元アドレス y、宛先アドレス x の packets が送出されると、これを受けたノード C では当該宛先アドレス x が既に上記アドレステーブルにおいて学習済みのため、通常のイーサネットスイッチと同様に、当該アドレステーブルにおいて上記の如く当該宛先アドレス x について学習されたポート 2 に対し当該 packets を送出する。その結果当該 packets は当該ポート 2 に接

50

続されている経路を通りノード B に転送される。

【 0 0 3 2 】

又ノード B では上記の如く当該受信パケットの宛先アドレス x が既に自己のアドレステーブルにおいて学習済みのため、当該アドレステーブルにおいて上記の如く当該宛先アドレス x について学習されたポート 1 に対し当該パケットを送出する。その結果当該パケットは当該ポート 1 に接続されている端末 1、すなわち当該パケットの宛先に転送される。

【 0 0 3 3 】

このように先着ポート学習方式によれば、先着パケットが通過した経路が（対応するポートにより）各ノードで学習されることにより、その後当該パケットの送信元アドレスを宛先アドレスに有するパケットについて同上学習結果にしたがって同一の経路が適用される。

10

【 0 0 3 4 】

ここでこの先着ポート学習方式ではパケットが先着したポートが最適ポートと見なされ自動的に経路が決められる。すなわちプロトコルなどによるソフトウェア処理を介さずに最適な経路が選択される。

【 0 0 3 5 】

しかしながら実際には人為的な経路の操作を行うことにより、より効果的な負荷分散がなし得るような場合もある。

【 0 0 3 6 】

例えば、図 5 に示されるように LAN 1 と LAN 2 との間が 3 台のスイッチノード B、ノード A、ノード C により接続されているネットワークがあり、VLAN ID = 1 と VLAN ID = 2 の 2 種のトラフィックが流れていると仮定する。

20

【 0 0 3 7 】

このような場合、LAN 2 へに対してフラディングされたパケットは、上記いずれの VLAN のトラフィックにおいてもホップ数の少ない下側の経路（すなわち下側のリンク、以下同様）、すなわちノード A を経由しない、ノード B から直接ノード C に至る経路を通ったパケットがノード C に先着し、もって該当するポート 2 のみが学習されるようになると考えられる。

【 0 0 3 8 】

その結果ノード C のアドレステーブル（図 5 参照）では全て上記ポート 2 が送出ポートとして学習されるため、その後の LAN 2 から LAN 1 へのユニキャストトラフィックは図 6 に示す如く、常に上記下側の経路のみを使用するようになる。すなわちトラフィックは当該ノード C とノード B とを直接接続する経路に集中してしまい、帯域の有効利用が図れない。

30

【 0 0 3 9 】

またこの先着ポート学習方式ではリンクの帯域は特に考慮されない。

【 0 0 4 0 】

すなわちギガビットイーサネット（以下単に G b E と略称する）のパケット読出し速度（すなわち帯域幅）はファストイーサネット（以下単に F E と略称する）の 10 倍であるため、基本的には G b E のリンクを通る経路を通るパケットの方が先着になりやすいという特性がある。しかしながら図 7 に示す如く G b E のリンクにノードが存在したり、或いはたまたま G b E のリンクに一時的な輻輳が発生したりする場合、F E のリンクを通るパケットばかりが先着するような事態も生じ得る。そのような場合細い側の帯域 F E のリンクにトラフィックが集中し、適切な負荷分散が図れなくなる。

40

【 0 0 4 1 】

本発明の実施例はこのような問題点を解決するために特に有効である。

【 0 0 4 2 】

ここで先着ポート学習方式において O S P F（オープン・ショーテスト・パス・ファースト）や上記 S T P のようなコントロールプレーンを設けリンクにコストという概念を導入することで意図的な経路の操作（トラフィックエンジニアリング）を可能にするという方法

50

も考えられる。しかしながらこのような方法はコントロールプレーンを不要にするという先着ポート学習方式の基本概念と相反することになり好ましくない。すなわち特にコントロールプレーンを要さないで経路操作が行える方法が望ましい。

【 0 0 4 3 】

上記の如く先着ポート学習方式とはフラディングパケットが最も早く到着したポートが最適なポートとして選択される方式である。したがってこの方式によればフラディングパケットの到着が遅いポートは最適経路として選択されない。

【 0 0 4 4 】

そうすると本来最先に到着するポートについて何らかの方法でパケットを遅く到着させる、あるいは到着しないようにすることにより、先着ポートを意図的に変更でき、もって経路の操作が可能になる。

10

【 0 0 4 5 】

例えば、前述の図5の場合において、下側の経路を通るVLAN ID = 2のフラディングパケットに対し遅延を挿入することで、当該経路を通るパケット、すなわちポート2から到着するパケットよりノードAを経由するパケット、すなわちポート3から到着するパケットの方が最先にノードCに到着するようにすることができる。その結果ノードCでは、VLAN ID = 1のトラヒックについてポート2に係る経路を最適経路と判断するが、VLAN ID = 2のトラヒックについては上記の如くの遅延の挿入の結果としてポート3に係る経路を最適経路と判断するようになる。よって負荷分散を図り得る。

20

【 0 0 4 6 】

このような制御を可能にするための構成を含む本発明の実施例による通信装置の装置構成を図8に示す。

【 0 0 4 7 】

この通信装置10は上記各ノードとして機能する装置である。

【 0 0 4 8 】

同通信装置10は複数のポート#1, #2, ..., #Nを持っており、各ポートはパケット送受信部1に接続されている。

【 0 0 4 9 】

パケット送受信部1はパケットのカプセリング等送受信に関する処理を行う機能ブロックであり、主信号のデータパケットはパケット送受信部1を経由してパケット処理部3へと送信される。

30

【 0 0 5 0 】

パケット処理部3は前記先着ポート学習方式によるパケットのフォワーディング処理と同一性判定処理とを行う。

【 0 0 5 1 】

フォワーディング処理は通常のイーサネットスイッチのものと同様とされ、ここでは上記の如くのアドレステーブル4を参照し、現在処理中のパケットの宛先アドレスのレコードアドレステーブルから得られた場合、そのレコードに記録されたポートに当該パケットを送出するようにパケット送受信部1に指示を送る。他方該当するレコードがアドレステーブル4から得られない場合、同パケットをフラディングするようにパケット送受信部1に指示を送る。

40

【 0 0 5 2 】

上記同一性判定処理とは、受け取ったパケットが既に受信したことのあるパケットかどうかを判定する処理である。すなわち現在処理中のパケットが先着パケットか後着パケットかを判定する。

【 0 0 5 3 】

具体的には、パケット処理部3ではパケットが到着すると同一性情報テーブル5にアクセスし、そのパケットに該当する情報が既に存在するか否かを調べる。存在しなければ、そのパケットを先着パケットと見なし、その到着ポートを先着ポートとしてアドレステ-

50

ブル4に書き込む。その場合さらに、パケットの同一性を判断するための情報(例えばH a s hによる縮退情報など)を同一性情報テーブル5に書き込む。

【0054】

他方同一性情報テーブル5において既に到着パケットに該当する情報が存在する場合、これは後着パケットと判定され、廃棄される。

【0055】

上述の動作は先着ポート学習方式の一般的な動作である。

【0056】

本発明の実施例では上記パケット送受信部1とパケット処理部3との間にフィルタリング/遅延処理部2という機能ブロックを設ける。

10

【0057】

この機能ブロックはパケットを送受信する際、当該パケットに対しフィルタリング処理を実施したり遅延を挿入したりするための機能ブロックである。このブロックはユーザインタフェースを持ち、ユーザによるパラメータの設定が可能とされる。すなわちユーザによって設定された値に従って、フィルタリング処理または遅延の挿入を実施する。

【0058】

以下に図9～図15とともに、このフィルタリング/遅延処理部2が有する機能の例について説明する。

【0059】

ここで本発明の実施例による通信装置10のフィルタリング/遅延処理部2は、図9、図10、図12、図13、図14及び図15の各々とともに説明する計5種の機能を有し、或いはそのうちの一又は複数の機能のみを備える構成を有する。ここで複数の機能を有する構成の場合には上記ユーザインタフェースを介して当該複数の機能のうちのいずれかの機能を選択可能とする構成を設ければよい。

20

【0060】

図9は図8とともに上述のフィルタリング/遅延処理部2が有する機能の例として、パケット処理部3からパケット送受信部1へ送出されるパケット、すなわち当該通信装置10から外部へ送出されるパケットに対し遅延を挿入する機能を説明するための図である。

【0061】

この構成では特定のポートを介して外部に送出されるパケットに遅延を挿入することにより、当該ポートに通信経路を介して接続された対向する通信装置に当該パケットが到着するタイミングを遅らせることが可能である。

30

【0062】

その結果当該対向する通信装置における当該受信に係るポートが当該対向する通信装置において先着ポートとなる可能性を効果的に減らし得る。その結果これらの対向する通信装置の当該ポート間の経路がその後のユニキャストのトラヒックにおいて使用される経路として選ばれる可能性を効果的に減らし得る。したがって経路操作が果たされる。

【0063】

この場合遅延の対象となるトラヒックは上記ユーザインタフェースを介してユーザによって設定される。この場合の設定情報として、例えば「ポート全体」、「ポート+VLAN」或いは「ポート+QoSクラス」等が可能である。

40

【0064】

ここで「ポート全体」とは、指定されたポートへ送出されるパケットに対し常に所定の遅延を挿入する設定であり、「ポート+VLAN」とは、指定されたポートへ送出されるパケットのうち指定されたVLANに係るパケットに対し常に所定の遅延を挿入する設定であり、「ポート+QoSクラス」とは、指定されたポートへ送出されるパケットのうち指定されたQoSクラスのパケットに対し常に所定の遅延時間分の遅延を挿入する設定である。

【0065】

ここで例えば「ポート+VLAN」又は「ポート+QoSクラス」の設定の場合、指定

50

された当該通信装置 10 のポートが接続された対向する通信装置のポートが、指定されたトラヒックについては先着ポートとして選ばれ難くなる。

【0066】

例えばポート #1 と #2 とが対向する通信装置に接続され、ポート #1 の VLAN #1 と、ポート #2 の VLAN #2 とを遅延挿入対象とした場合を想定する。

【0067】

このような場合、これら相対向する通信装置間の通信において、VLAN #1 のトラヒックについては、遅延が挿入されない側のポート #2 に接続された対向する通信装置のポートが選ばれる可能性が高まり、同様に VLAN #2 のトラヒックについては、遅延が挿入されない側のポート #1 に接続された対向する通信装置のポートが選ばれる可能性が高まる。

10

【0068】

このような機能により効果的に各経路のトラヒック負荷を分散させることが可能になる。

【0069】

なおこの例の場合、フィルタリング/遅延処理部 2 は、パケット送受信部 1 からパケット処理部 3 へ渡されるパケットについては、特に何の処理も行わない。

【0070】

また図 10 は、図 8 に示す本発明の実施例による通信装置が有する機能の例として、図 9 の機能とは逆に、パケット送受信部 1 からパケット処理部 3 へ送られるパケット、すなわち当該通信装置が受信したパケットに対し遅延を挿入する機能を説明するための図である。

20

【0071】

この構成ではフィルタリング/遅延処理部 2 が、当該通信装置 10 の特定のポートを介して外部から受信されたパケットに遅延を挿入する。その結果当該通信装置 10 のパケット処理部 3 では当該遅延の挿入によりタイミングが遅らされた上で当該パケットを受信する。

【0072】

ここでパケット処理部 3 では先着ポート学習方式による動作においてパケットの受信タイミングを基準としてポート間のパケット到着の先後を判断するため、上記遅延の挿入の結果その受信タイミングが遅れ、もって前記特定のポートが先着ポートと判断される可能性が減少する。このようにして当該ポートに接続された経路がその後のユニキャストのトラヒックにおいて使用される経路として選ばれる可能性を効果的に減らすことが可能である。したがって経路操作が果たされる。

30

【0073】

この場合に遅延を挿入する対象となるパケットはユーザにより上記ユーザインタフェースを介して設定される。この場合の設定情報としても上記の図 9 の場合同様、例えば、「ポート全体」、「ポート + VLAN」、「ポート + QoS クラス」などが可能である。この場合、例えば指定されたポートが指定されたトラヒックについて先着ポートとして選ばれ難くなる。

40

【0074】

例えば上記同様、ポート #1 と #2 とが対向する通信装置に接続されている場合であって、ポート #1 の VLAN #1 と、ポート #2 の VLAN #2 とを遅延挿入対象とした場合を想定する。

【0075】

このような場合、これら相対向する通信装置間の通信において、VLAN #1 のトラヒックについて、遅延が挿入されない側のポート #2 が選ばれる可能性が高まり、同様に VLAN #2 のトラヒックについて、遅延が挿入されない側のポート #1 が選ばれる可能性が高まる。その結果その後の当該通信装置 10 からのユニキャストのトラヒックでは、VLAN #1 のトラヒックについてはポート #2 が最適経路として選択され、VLAN #2

50

のトラヒックについてはポート # 1 が最適経路として選択され易くすることが可能となる。

【 0 0 7 6 】

このような機能により効果的に各経路のトラヒック負荷を分散させることが可能になる。

【 0 0 7 7 】

なおこの例の場合、パケット処理部 3 からパケット送受信部 1 へ渡されるパケットについては、特に何の処理も行われない。

【 0 0 7 8 】

次に図 1 1 とともに、図 9 , 図 1 0 とともに上述のフィルタリング / 遅延処理部 2 によってパケットに遅延を与える機能を実現するための具体的な構成として可能な遅延挿入部の構成例について説明する。

【 0 0 7 9 】

尚ここで述べる遅延挿入部 2 0 0 の構成はあくまで一例に過ぎず、本発明の実施例による通信装置 1 0 におけるフィルタリング / 遅延処理部 2 に適用される遅延挿入部の具体的な構成はこれに限定されることはない。

【 0 0 8 0 】

図 1 1 (a) に当該遅延挿入部 2 0 0 のブロック構成を示す。

【 0 0 8 1 】

遅延挿入部 2 0 0 は周知のパケット待ち合わせ用の F I F O (ファースト・イン・ファースト・アウト)バッファに時刻を管理する部分を追加した構成を採る。

【 0 0 8 2 】

遅延挿入部 2 0 0 はパケットを格納する物理メモリ 2 0 1、当該パケットの情報を管理する F I F O 管理部 2 0 2、時刻を管理するクロック 2 0 3 及びパケットの格納時間が所定の遅延時間を超えたか否かを管理する遅延制御部 2 0 4 を有する。

【 0 0 8 3 】

また F I F O 管理部 2 0 2 は図 1 1 (b) に示す構成を有する F I F O 情報テーブルを有する。

【 0 0 8 4 】

同テーブルの「アドレス」項にはパケットが格納されている物理メモリ 2 0 1 のアドレスが記述され、その「書込み時刻」項にはパケットを物理メモリ 2 0 1 に書き込んだ時刻が記述される。又ここには F I F O の先頭を示す先頭ポインタが存在する。なお上記時刻は標準的な時刻(時 / 分 / 秒)である必要はなく、当該通信装置 1 0 内のクロック 2 0 3 が生成する装置内時刻でもよい。

【 0 0 8 5 】

次に遅延挿入部 2 0 0 の具体的な動作について説明する。

【 0 0 8 6 】

まず、全ての動作に先立ち、上記の如くパケットに付与する所定の遅延時間が遅延制御部 2 0 4 に設定される。この設定はオペレータが上記ユーザインタフェースを介し手動で行う構成(前述の図 9 , 図 1 0 の機能の例の場合)が可能であり、或いはソフトウェアが自動的に行う構成(後述する図 1 4 , 図 1 5 の機能の例の場合)も可能である。

【 0 0 8 7 】

次に遅延挿入部 2 0 0 における、パケット入力時の動作について説明する。パケットが入力されると F I F O 管理部 2 0 2 が当該パケット本体を物理メモリ 2 0 1 に書き込む。そして、書き込んだ物理メモリ 2 0 1 のアドレスと、クロック 2 0 3 から得た現在時刻とをそれぞれ図 1 1 (b) の F I F O 情報テーブルの空いている部分の先頭に書き込む。

【 0 0 8 8 】

次にパケット出力時の動作について説明する。遅延制御部 2 0 4 は、F I F O 情報テーブルの先頭に登録されている格納パケットが格納された時刻を F I F O 管理部 2 0 2 から入手し、クロック 2 0 3 から得られる現在時刻と常に比較する。そして現在時刻と先頭パ

10

20

30

40

50

ケットの格納時刻との差分が上記設定された遅延時間を超えた場合、遅延制御部 204 は F I F O 管理部 201 に対し出力命令を出す。

【0089】

F I F O 管理部 201 はこの出力命令を受けると、F I F O 情報テーブルの先頭に登録されているケットを物理メモリ 201 から取り出して遅延挿入部 200 から出力する。その後、出力したケットのエントリを F I F O 情報テーブルから削除し、先頭のポインタを一つ後ろのエントリへ移動する。

【0090】

尚フィルタリング/遅延処理部 2 において上述した遅延挿入部 200 の F I F O 情報テーブルを当該通信装置 10 のポートごとに設け、遅延制御部 204 では遅延時間の設定および格納時間と設定遅延時間との比較動作をこれらの F I F O 情報テーブルごとに順次個別に実行するように構成することで、ポートごとに挿入すべき遅延の遅延時間を個別に設定可能となる。この場合、遅延を挿入する必要のないポートに係るケットについては物理メモリ 201 への格納を省略する構成とすればよい。

10

【0091】

このような構成の遅延挿入部 200 によれば、フィルタリング/遅延処理部 2 にケットが入力された後フィルタリング/遅延処理部 2 から出力される迄の時間を所定の遅延時間として設定可能であり、図 9 , 図 10 とともに上述の如く、ポートごとに、受信されたケット(図 10 の場合)或いは送出されるケット(図 9 の場合)に対し所定の遅延時間の遅延を与えることが可能となる。

20

【0092】

又上記「ポート+VLAN」或いは「ポート+QoSクラス」の設定の場合、該当するVLANに係るケット或いは該当するQoSクラスに属するケット以外のケットについては物理メモリ 201 への格納を省略する構成とすればよい。

【0093】

又一のポートについて、各ケットのVLAN或いはQoSクラスごとに異なる遅延時間を設定する場合(後述する図 15 の例)、物理メモリ 201 に格納するケットに対する遅延時間をVLANごと或いはQoSクラスごとに管理する必要がある。これを実現するため、例えば図 11 (b) の F I F O 情報テーブルを設定されたVLAN或いはQoSクラスごとに設け、遅延制御部 204 では遅延時間の設定およびケットの格納時間と設定遅延時間との比較動作をこれらの F I F O 情報テーブルごとに順次個別に実行するように構成すればよい。又同じVLANについて上限帯域を細かく設定するような場合(図 17 A、図 17 B、図 17 C の例の場合)も同様に、更に各VLANについて、その設定された上限帯域ごとに F I F O 情報テーブルを設け、遅延制御部 204 では遅延時間の設定およびケットの格納時間と設定遅延時間との比較動作をこれらの F I F O 情報テーブルごとに順次個別に実行するように構成すればよい。

30

【0094】

次に通信装置 10 のフィルタリング/遅延処理部 2 の機能の例として、図 12 に示される如く、ケット処理部 3 からケット送受信部 1 へ送出されるケット、すなわちこれから外部に送出されるケットに対してフィルタリングを行う機能について説明する。

40

【0095】

この場合、ユーザから指定されたポートに係る特定の異トラヒックがフィルタリング/遅延処理部 2 において廃棄され、外部に送出されないように構成する。

【0096】

その結果該当するポートに接続された経路は指定されたトラヒックについては対向する通信装置において完全に先着ポート学習の対象から外される。その結果、当該ポートに接続された経路は指定されたトラヒックについては学習されることが無く、もって指定されたトラヒックについての当該経路の適用が意図的に排除される。よって経路操作が実現される。

【0097】

50

なおこの場合、パケット送受信部 1 から処理部 3 へ渡されるパケットについては、フィルタリング / 遅延処理部 2 は特に何の処理も行わない。

【 0 0 9 8 】

また上記図 1 2 の例とは逆に、図 1 3 に示される如くパケット送受信部 1 からパケット処理部 3 へ送られるパケット、すなわち外部から受信されたパケットを廃棄する動作を行わせることも可能である。

【 0 0 9 9 】

図 1 3 の例ではユーザから指定された特定のトラヒックはフィルタリング / 遅延処理部 2 にて廃棄され、パケット処理部 3 によって受信されることがなくなる。これにより、該当するポートは指定されたトラヒックについては完全に先着ポート学習の対象から外される。

10

【 0 1 0 0 】

この場合も図 1 2 の場合同様、当該ポートに接続された経路は指定されたトラヒックについては当該通信装置 1 0 において学習されることがなく、もって当該経路の指定されたトラヒックについての適用が意図的に排除される。よって経路操作が実現される。

【 0 1 0 1 】

なおこの場合、パケット処理部 3 からパケット送受信部 1 へ渡されるパケットについては、フィルタリング / 遅延処理部 2 は特に何の処理も行わない。

【 0 1 0 2 】

ここで本発明の実施例による通信装置 1 0 は、このように経路を操作する動作としてフィルタリングまたは遅延処理を実施する動作は、上記図 9、図 1 0、図 1 2、図 1 3 の例の如くユーザの手動による設定に基づく場合に限られない。

20

【 0 1 0 3 】

すなわち通信装置 1 0 が自律的に遅延を挿入するようにする構成も可能である。

【 0 1 0 4 】

その具体例として、物理ポートの速度に応じて自動的に遅延時間を決定する方法が考えられる。

【 0 1 0 5 】

この場合の本発明の実施例による通信装置 1 0 のフィルタリング / 遅延処理部 2 の機能の例につき、図 1 4 とともに説明する。

30

【 0 1 0 6 】

この場合図 1 4 に示す如く、フィルタリング / 遅延処理部 2 はパケット送受信部 1 から、各物理ポート # 1、# 2、...、# N から受信されるパケットの受信速度を示す情報を収集し、その情報に応じて自動的に遅延時間を決定し、その決定された遅延を挿入するように構成される。

【 0 1 0 7 】

具体的な挿入すべき遅延時間の決定方法として、通信速度の遅いポートについてより多くの遅延時間の遅延を挿入するようにする。その結果当該ポートに接続された対向する通信装置のポートが先着ポートとして選ばれ難くなり、もって当該経路が最適経路と判定される可能性が減る。もって経路操作が果たされる。

40

【 0 1 0 8 】

図 1 4 の例の場合、ポート # 1 の通信速度は図示の如く 1 0 G b / s、ポート # 2 の通信速度は 1 0 0 G b / s、ポート # 3 の通信速度は 1 G b / s、...、ポート # N の通信速度は 1 0 G b / s であるため、これらの計 4 つのポートの中では、ポートごとに挿入する遅延時間は、ポート # 3 について最も長く、ポート # 2 について最も短くする。その結果通信速度が最も遅いポート # 3 はその後のユニキャストのトラヒックにおいて採用され難くなり、逆に通信速度が最も早いポート # 2 が採用され易くなる。

【 0 1 0 9 】

なお、図 1 4 は、図 9 の機能の例に対応する送信時遅延挿入方式、すなわち送信するパケットに遅延を挿入する方式を用いた場合を表しているが、同様の構成にて図 1 0 の機能

50

の例に対応する受信時遅延挿入方式、すなわち受信されたパケットに遅延を挿入する方式に置き換えることも可能である。その場合の具体的な挿入すべき遅延時間の決定方法としても上記同様、通信速度の遅いポートについてより長い遅延時間の遅延を挿入することにより、当該ポートが先着ポートとなり難くするように構成する。

【0110】

更に他の例として、物理ポートではなく論理ポートの速度に応じて自動的に遅延を挿入する構成が考えられる。

【0111】

ここで論理ポートとは、一の物理ポートについて複数種設定可能な仮想的なポートを意味する。具体的には、イーサネットインタフェースの場合、Tag VLANポートにおける各VLAN等がこれに該当する。

10

【0112】

Tag VLANポートの使い方としては、当該ポートの通過が許されるパケットのVLANを特定するVLAN値のみを設定し特に帯域を設定しないような使い方の他に、VLAN毎に論理的なキューを与え読出し帯域を設定するような使い方も行われている。後者の様な使用法の場合、一つの物理的なリンクの中に複数の細い論理的なリンクが存在することになる。

【0113】

したがって図14とともに上述の例で述べた如くの、物理的なリンク（以下単に物理リンクとも称する）の帯域に応じて決定された遅延時間の遅延を挿入するという構成を、そのまま論理的なリンク（以下単に論理リンクとも称する）に対する構成に置き換えることが可能である。この場合も図14の例の場合同様、帯域の細い論理リンクほど、より長い遅延時間の遅延を挿入する構成を採る。

20

【0114】

なお、上記読出し帯域とは、最低限読出しが保証される帯域或いは読出し上限帯域を意味する。

【0115】

この論理ポートの速度に応じて遅延を挿入する機能につき、図15とともに説明する。

【0116】

この例においてパケット送受信部1に対し、ユーザによって各論理リンクの帯域が設定される。フィルタリング/遅延処理部2は、パケット送受信部1からこのように設定された各論理リンクの速度情報を収集し、それに依りて自動的に挿入すべき遅延時間を決定する。

30

【0117】

なお、図15は、図9の機能の例に対応する送信時遅延挿入方式、すなわち送信するパケットに遅延を挿入する方式を用いた場合を表しているが、同様の構成にて図10の機能の例に対応する受信時遅延挿入方式、すなわち受信されたパケットに遅延を挿入する方式に置き換えることも可能である。その場合の具体的な挿入すべき遅延時間の決定方法としても上記同様、通信速度の遅いリンクについてより長い遅延時間の遅延を挿入することにより、当該リンクがその後のユニキャストのトラヒックにおいて採用され難くすることが可能である。

40

【0118】

図15の例の場合図示の如く、ポート#1の通信速度はVLAN1のトラヒックについて90Gb/s、VLAN2のトラヒックについて40Gb/s、ポート#2の通信速度はVLAN2のトラヒックについて30Gb/s、VLAN5のトラヒックについて40Gb/s、VLAN8のトラヒックについて30Gb/s、...、ポート#Nの通信速度はVLAN9のトラヒックについて10Gb/sであるため、これらの計5種の論理ポートVLAN1、VLAN2、VLAN5、VLAN8、VLAN9の中では、挿入する遅延時間は最も狭い帯域のVLAN9について最も長くし、最も広い帯域のVLAN1について最も短くする。

50

【0119】

次に、図16A、図16B、図16C、図17A、図17B、図17Cとともに、図9、10、12、13、14及び15とともに上述の本発明の実施例による通信装置10のフィルタリング/遅延処理部2のそれぞれの機能に係る、通信装置10の具体的な動作例につき説明する。

【0120】

図16A、図16B、図16Cは図9及び14とともに上述した動作に係る動作例を説明するための図である。

【0121】

ここでは図16A、図16B、図16Cに示す如く3個のノード、すなわちノードB、ノードA及びノードCを有するネットワーク形態を想定する。このネットワークは以下の構成および設定を有する。

- 1) ノードB、Cのポート2、3、4はそれぞれTag VLANポートであり、それぞれVLAN = 1、2が設定されている。
- 2) ノードAを通過する際にフラッディングパケットが被る遅延はaである。
- 3) ノードBのポート3につき、VLAN = 2のトラヒックに遅延bを設定する。
- 4) GbE、すなわちポート2、3に接続された経路には自動的な遅延挿入は行わないが、FE、すなわちポート4に接続された経路には自動的に遅延cを挿入する。

【0122】

このとき、それぞれ遅延a、b、cの遅延時間の長短関係は、 $a < b < c$ とする。このような設定の下では、図16Aに示す如く、ノードBのポート3にVLAN = 2について遅延bが、ポート4に遅延cが、ノードCのポート4に遅延cがそれぞれ設定される。

【0123】

このようなネットワークにおいて、LAN1からLAN2に対し、VLAN = 1ならびにVLAN = 2の未学習パケットをそれぞれ送出したとする。このときの様子を図16Bに示す。

【0124】

これらのパケットはノードBにおいて学習されていないので、ノードBからは、到着ポート以外の全てのブロードキャストドメインに対し当該パケットがフラッディングされる。

【0125】

まず、VLAN = 1のパケットについては、ノードCにおいてポート2から到着するパケットは上記設定2)により、ノードAによる遅延aを受ける。このため「伝送路遅延にこの遅延aが加えられた遅延時間」後にノードCに到着する。

【0126】

他方ポート3からノードCに到着するVLAN = 1のパケットは特に遅延を受けない。上記設定3)により、ポート3に対してはVLAN = 2のトラヒックに対し遅延bが付与されるが、VLAN = 1のパケットに対しては遅延が付与されないからである。その結果この場合は「伝送路遅延のみ」でノードCに到着する。

【0127】

ポート4からノードCに到着するパケットは「伝送路遅延および遅延c」の経過後にノードCに到着する。上記設定4)によりこの経路を通るパケットについては一律に遅延cが付与されるからである。

【0128】

すなわち経路ごとのノードCへの到着時間の遅延は、ポート2からのものが「経路遅延 + a」、ポート3からのものが「経路遅延のみ」、ポート4からのものが「経路遅延 + c」となり、経路遅延による遅延時間がこれら3つの経路間で差がないと仮定すると、「経路遅延のみ」のポート3からのものが最も遅延が少ない。したがって特に輻輳が無い環境化においては、VLAN = 1のパケットに関してはノードCではポート3が先着ポートとして学習される。

【 0 1 2 9 】

一方、VLAN = 2のパケットについては、ポート2, ポート4からのものはVLAN = 1の場合と同じであるが、ポート3から到着するパケットは、「伝送路遅延 + b」後に到着する。

【 0 1 3 0 】

したがってこの場合、ポート2からのものが遅延「経路遅延 + a」、ポート3からものが「経路遅延 + b」、ポート4からのものが「経路遅延 + c」となり、上記同様経路遅延による遅延時間がこれら3つの経路間で差がないと仮定すると、上記 $a < b < c$ の関係からポート2からのものが最も遅延時間が少ない。したがってVLAN = 2のパケットに関してはポート2が先着ポートとして学習される。

10

【 0 1 3 1 】

このようにVLAN = 1のパケットについてはポート3の経路が最適経路と判定され、VLAN = 2のパケットについてはポート2の経路が最適経路と判定される。その結果その後のLAN2からLAN1へのユニキャストのトラヒックは図16Cに示す如く、VLAN = 1のパケットはポート3の経路を通り、VLAN = 2のパケットはポート2の経路を通るように経路制御がなされる。このようにVLAN = 1と2とでトラヒックが別の経路を通ることになり、効果的な負荷分散が可能となる。

【 0 1 3 2 】

また、帯域の小さな(すなわちFEの)ポート4は自動遅延挿入によって常に先着ポートとなる可能性が小さく、もってユニキャストのトラヒックに使用される経路として選択され難くなる。その意味でも効果的な負荷分散が図れる。

20

【 0 1 3 3 】

次に図17A、図17B, 図17Cとともに、図10及び15とともに上述した機能例に係る動作例について説明する。

【 0 1 3 4 】

この場合、図17Aに示す如く、2つのノード、すなわちノードA, Bのネットワーク形態を想定する。このネットワークでは、以下の構成とされ、設定がなされるものとする。

- 1) ノードB, C間のリンクは全てGbEによるものであり、その伝送遅延は全て等しい。
- 2) ノードB, Cのそれぞれのポート2, 3, 4は、それぞれTag VLANポートであり、それぞれについてVLAN = 1, 2, 3が設定されている。
- 3) ノードB - C間のポート2, 3, 4にはそれぞれ論理リンクが設定されており、以下の上限帯域が設定されている(設定帯域は対称)。

30

ポート2 : VLAN1 = 100 [Mb / s] , VLAN2 = 500 [Mb / s] , VLAN3 = 800 [Mb / s]

ポート3 : VLAN1 = 400 [Mb / s] , VLAN2 = 300 [Mb / s] , VLAN3 = 200 [Mb / s]

40

ポート4 : VLAN1 = 200 [Mb / s] , VLAN2 = 400 [Mb / s] , VLAN3 = 100 [Mb / s]

これら論理ポートに対して自動的に設定される遅延時間は、 $1 [ms] \times 1000 [Mb / s] / (論理帯域)$ とする。

【 0 1 3 5 】

ノードBとCのポート2のVLAN2に対しては、オペレータが人為的に5 [ms] の遅延を挿入する。

【 0 1 3 6 】

50

このような構成および設定の下では、図 17B に示す如くに各ポートにおいてそれぞれ遅延が設定される。

【 0 1 3 7 】

以下にポートごとに上記論理帯域に基づいて自動的に設定される遅延および人為的に設定される遅延を示す。

ポート 2 : V L A N 1 = 1 0 [m s] , V L A N 2 = 2 [m s] + 5 [m s] , V L A N 3 = 1 . 2 5 [m s]

ポート 3 : V L A N 1 = 2 . 5 [m s] , V L A N 2 = 3 . 3 3 [m s] , V L A N 3 = 5 [m s] 10

ポート 4 : V L A N 1 = 5 [m s] , V L A N 2 = 2 . 5 [m s] , V L A N 3 = 1 0 [m s]

このうち論理帯域に基づいて自動的に設定される遅延時間は、上記の如く各論理ポートに自動的に設定される遅延時間が $1 [m s] \times 1 0 0 0 [M b / s] / (\text{論理帯域})$ であり、ポート 2 では

V L A N 1 = 1 0 0 [M b / s] であるため $1 0 0 0 / 1 0 0 = 1 0 [m s]$ 20

V L A N 2 = 5 0 0 [M b / s] であるため $1 0 0 0 / 5 0 0 = 2 [m s]$ 、
V L A N 3 = 8 0 0 [M b / s] では $1 0 0 0 / 8 0 0 = 1 . 2 5 [m s]$

として得られる。

【 0 1 3 8 】

又ポート 3 では

V L A N 1 = 4 0 0 [M b / s] であるため $1 0 0 0 / 4 0 0 = 2 . 5 [m s]$ 、
V L A N 2 = 3 0 0 [M b / s] であるため $1 0 0 0 / 3 0 0 = 3 . 3 3 [m s]$ 30

V L A N 3 = 2 0 0 [M b / s] では $1 0 0 0 / 2 0 0 = 5 [m s]$

として得られる。

【 0 1 3 9 】

又ポート 4 では

V L A N 1 = 2 0 0 [M b / s] であるため $1 0 0 0 / 2 0 0 = 5 [m s]$ 、
V L A N 2 = 4 0 0 [M b / s] であるため $1 0 0 0 / 4 0 0 = 2 . 5 [m s]$ 、
V L A N 3 = 1 0 0 [M b / s] では $1 0 0 0 / 1 0 0 = 1 0 [m s]$ 40

として得られる。

【 0 1 4 0 】

なお、本動作例の如く、自動的に設定された遅延にさらに人為的に遅延を付加することが可能である。この手法は自動設定だけでは適切に負荷分散がとれない場合等に有効である。

【 0 1 4 1 】

本動作例の場合 V L A N 2 については自動設定のみでは先着ポートがポート 2 となる可能性が高く、V L A N 3 についてもポート 2 が先着となる可能性が高い。よって、意図的に V L A N 2 においてポート 2 が選ばれ難くする目的でポート 2 では V L A N 2 について 50

オペレータが5 [ms] の遅延を挿入している。

【0142】

このようなネットワークにおいてLAN1からLAN2に対しVLAN = 1, 2, 3の未学習パケットをそれぞれ送出したとする。このときの様子を図17Cに示す。

【0143】

これらのパケットはノードBにおいて学習されていないため、ノードBからは、到着ポート以外の全てのブロードキャストドメインに対して当該パケットがフラッディングされる。

【0144】

まず、VLAN = 1については、ポート2からは10 [ms]後に、ポート3からは2.5 [ms]後に、ポート4からは5 [ms]後にそれぞれノードCに到着する。よってポート3からの到着が2.5 [ms]後で最も早い。したがってノードCでは通常はポート3が先着ポートとして学習される。

10

【0145】

VLAN = 2については、ポート2からは2 + 5 = 7 [ms]後に、ポート3からは3.33 [ms]後に、ポート4からは2.5 [ms]後にそれぞれノードCに到着する。よってポート4からの到着が2.5 [ms]後で最も早い。したがってノードCでは通常はポート4が先着ポートとして学習される。

【0146】

VLAN = 3については、ポート2からは1.25 [ms]後に、ポート3からは5 [ms]後に、ポート4からは10 [ms]後にそれぞれノードCに到着する。よってポート3からの到着が2.5 [ms]後で最も早い。したがってノードCでは通常はポート2が先着ポートとして学習される。

20

【0147】

すなわちノードCではVLAN = 1, 2, 3についてそれぞれポート3, 4, 2が先着ポートとして学習される。

【0148】

これら先着ポートとして学習されたポートがそのままその後のLAN2からLAN1へのユニキャストのトラヒックの際に使用されるため、VLAN = 1, 2, 3でそれぞれ別の経路を通ることになり、効果的な負荷分散が図れる。

30

【0149】

以上説明したように、本発明の実施例では、パケットネットワークにおいて、発信ノードは主信号パケットに情報を付加することなく、そのパケットを受け取ったポート以外のブロードキャストドメインにパケットをフラッディングし、各ノードにおいて到着の最も早かったポート(先着ポート)を、そのパケットの送信元識別子と共にフォワーディングテーブル(すなわちアドレステーブル)に学習する最適経路学習方法(先着ポート学習方式)において、各ノードにおいて、指定されたポートに送信または受信される特定のトラヒックに対して遅延またはフィルタリング処理を実施することにより、選択される経路を任意に操作可能な構成とした(図8の構成)。

【0150】

その結果、コントロールプレーンを持たないという先着ポート学習方式の特性を失うことなく経路の操作が可能となり、トラヒックの負荷分散を図ることが可能になる。

40

【0151】

また上記遅延またはフィルタリング処理として、フラッディングパケット送出時に遅延を挿入するように構成することが可能である(図9の構成例)。

【0152】

その結果、送信側ノードにおける処理により、特定のトラヒックに対し受信側のノードにおいて特定のポートが選択され難くすることが可能となる。

【0153】

この遅延挿入という方法によれば、フィルタリング(すなわち廃棄)の場合とは異なり

50

、一定遅延後には該当するパケットが対向ノードに到着する。その結果、仮に他の経路に障害が発生し当該他の経路に係るフラッディングパケットが目的端末に到着しないような場合においては、当該遅延挿入経路に係るポートが先着ポートとして選択される余地がある。したがって従来のプロトコルのコスト設定と同等の機能を実現することができる。

【0154】

又上記遅延またはフィルタリング処理として、パケット受信時に遅延を挿入するように構成することも可能である（図10の構成例）。

【0155】

その結果、受信側ノードにおいて、特定のトラヒックに対して特定のポートが選択され難くすることが可能となる。この場合も遅延の挿入による方法であるため、上記同様従来のプロトコルのコストを高く設定した場合と同等の効果が得られる。

10

【0156】

又上記遅延またはフィルタリング処理として、フラッディングパケット送信時にパケットを廃棄する（すなわちフィルタリングする）ようにすることも可能である（図12の構成例）。

【0157】

その場合、送信側ノードにおける処理により、特定のトラヒックに対して受信側ノードにおいて特定のポートが完全に選択されないようにすることが可能となる。

【0158】

このフィルタリングという方法は、指定ポートを先着ポートの候補から完全に除外してしまうため、上記の如く他の経路に障害が発生して当該他の経路に係るフラッディングパケットが届かないような場合においても上記指定ポートが選ばれることがない。この方法によれば、上記遅延挿入方式において不可欠な遅延時間のチューニング等の処理が不要となる。すなわち、より簡易に構成にて経路制御を実現し得る。

20

【0159】

又上記遅延またはフィルタリング処理として、パケット受信時にパケットを廃棄するように構成することも可能である（図13の構成例）。

【0160】

その結果、受信側ノードにおいて、特定のトラヒックに対し指定したポートが完全に選択されないようにすることが可能となる。この場合も上記の場合同様、より簡易な構成で経路制御を実現し得る。

30

【0161】

又上記遅延またはフィルタリング処理として、各ポートの物理速度に応じて自動的に遅延時間を決定するように構成することが可能である（図14の構成例）。

【0162】

この場合には帯域の小さな物理リンクを自動的に選ばれ難くすることが可能となる。その結果より効率的な負荷分散が可能となる。

【0163】

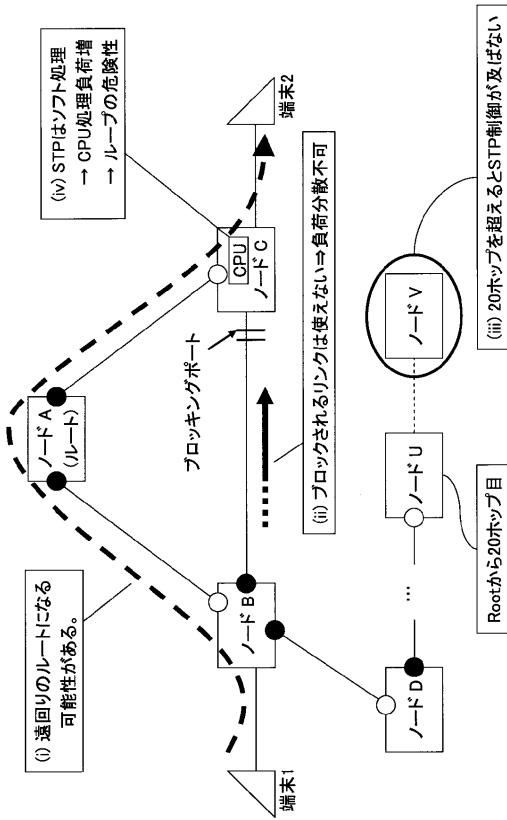
又上記遅延またはフィルタリング処理として、各ポートの論理速度に応じて自動的に遅延時間を設定するように構成することも可能である（図15の構成例）。

40

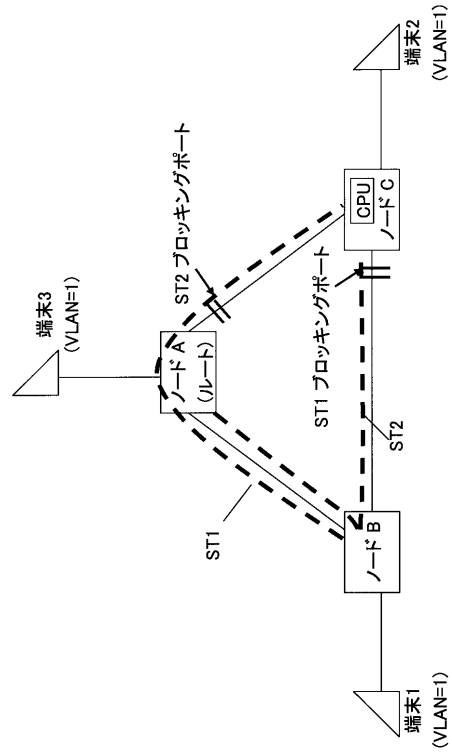
【0164】

この場合、上記ポートの物理速度に応じて自動的に遅延時間を決定するように構成の考え方を論理ポートにまで拡張し、帯域の小さな論理リンクを自動的に選ばれ難くすることが可能となる。これにより、より効率的な負荷分散が可能となる。

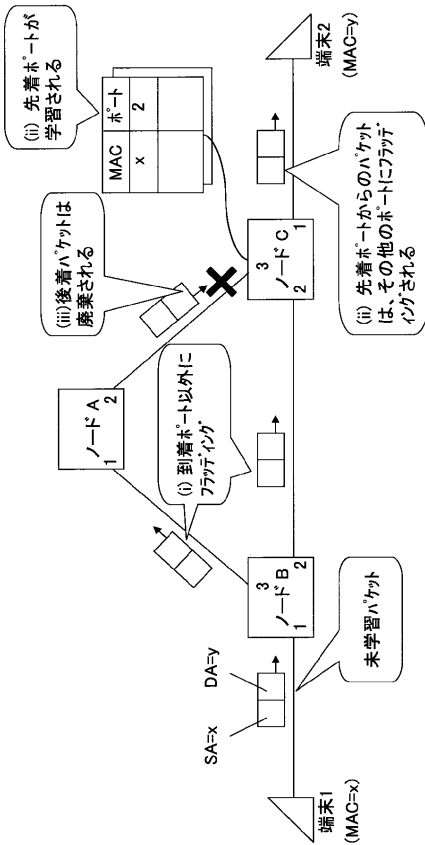
【 図 1 】



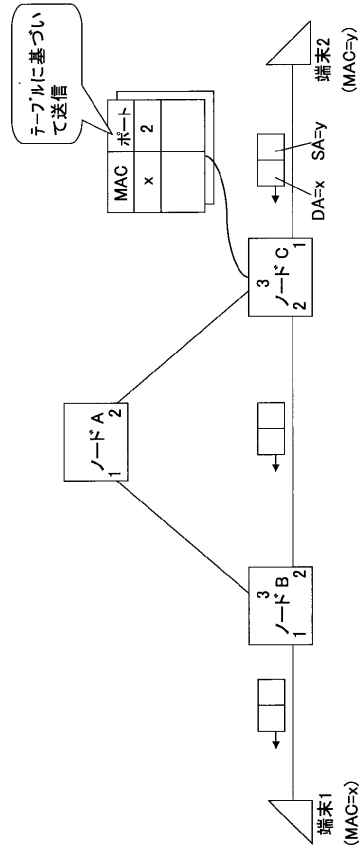
【 図 2 】



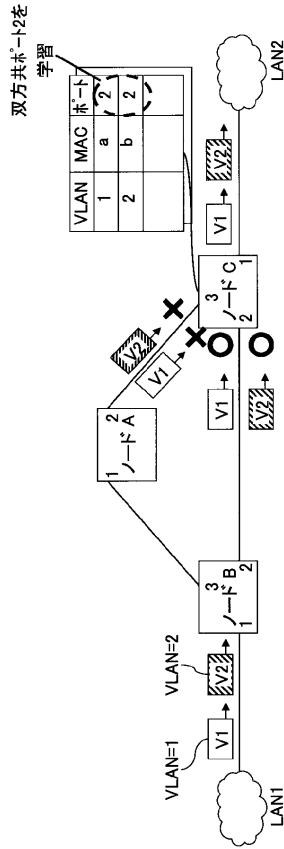
【 図 3 】



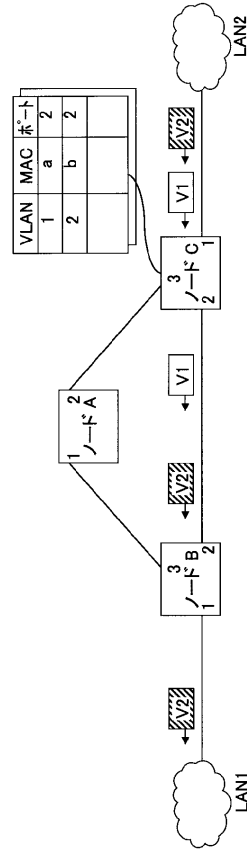
【 図 4 】



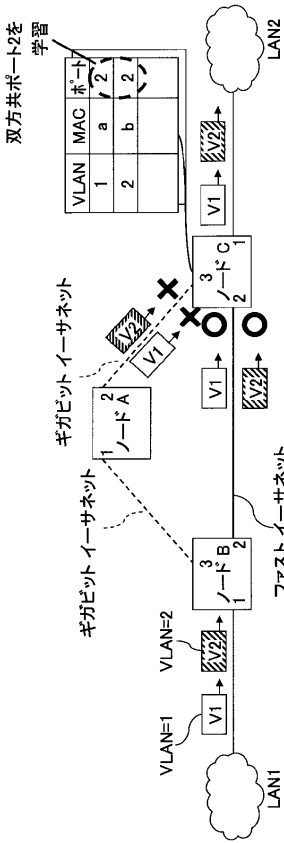
【 図 5 】



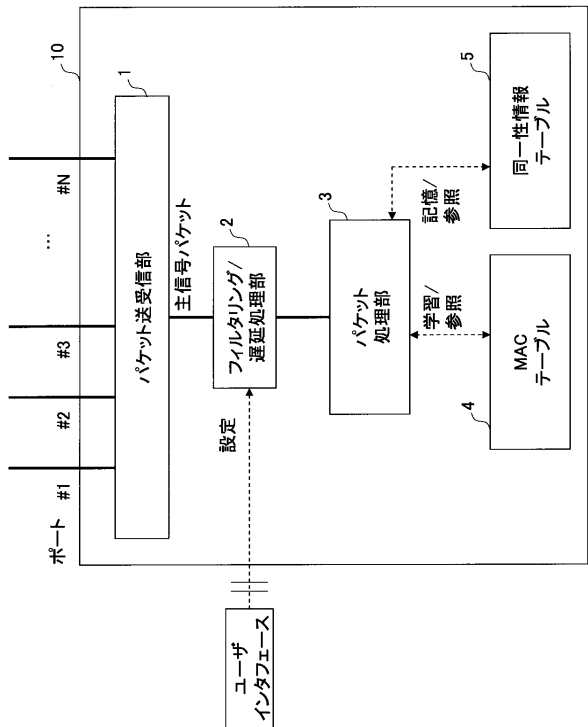
【 図 6 】



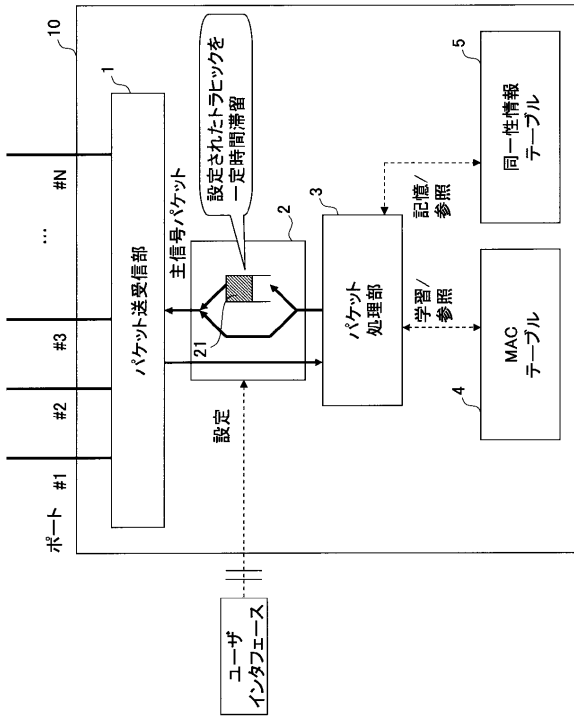
【 図 7 】



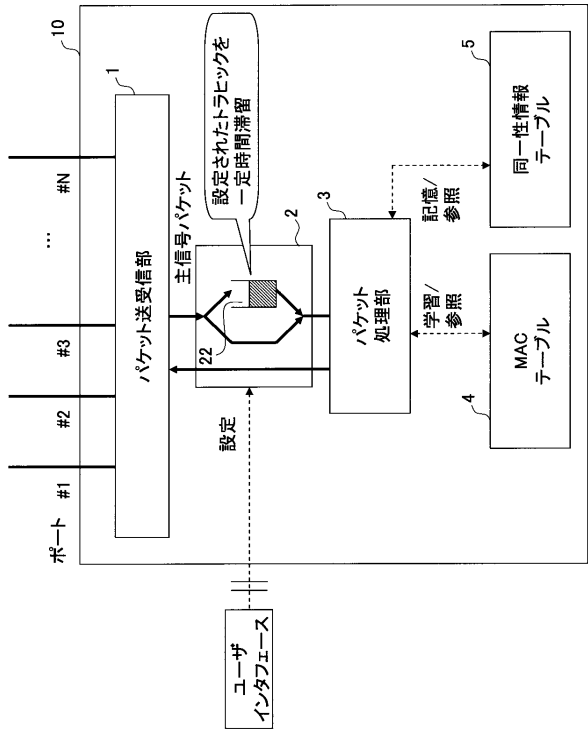
【 図 8 】



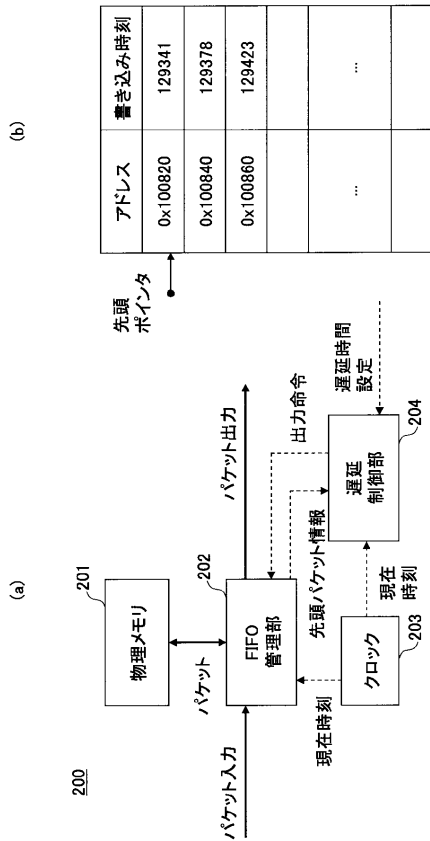
【図9】



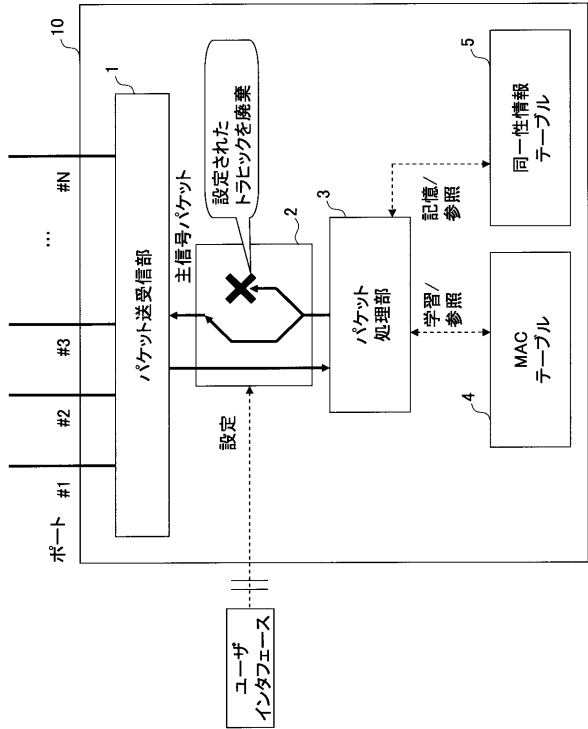
【図10】



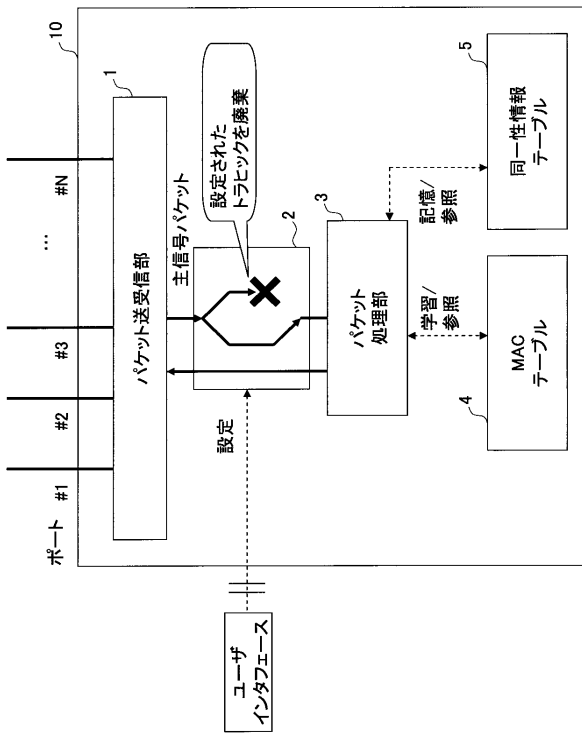
【図11】



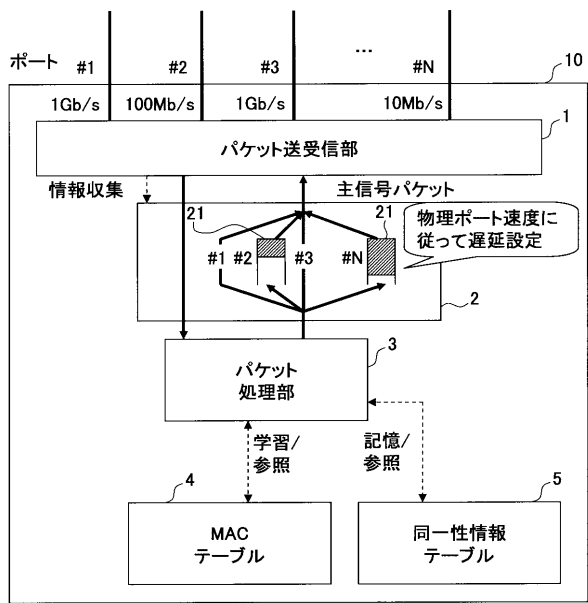
【図12】



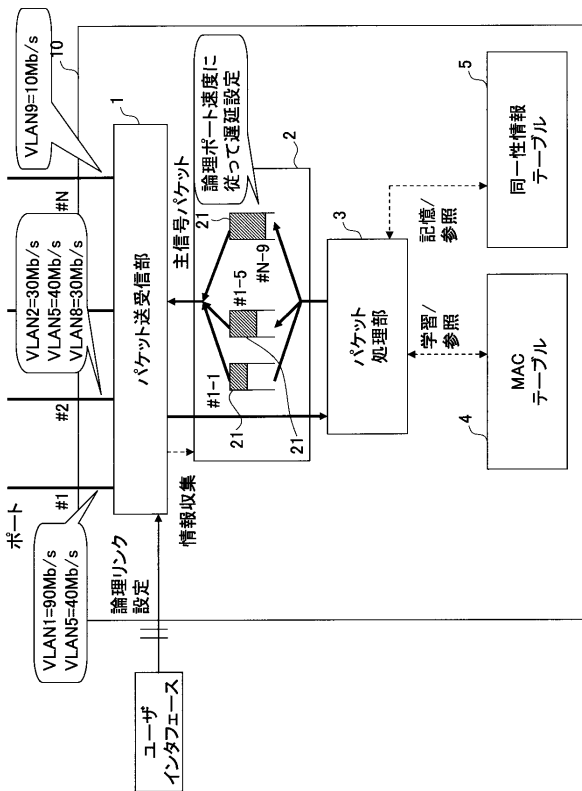
【図13】



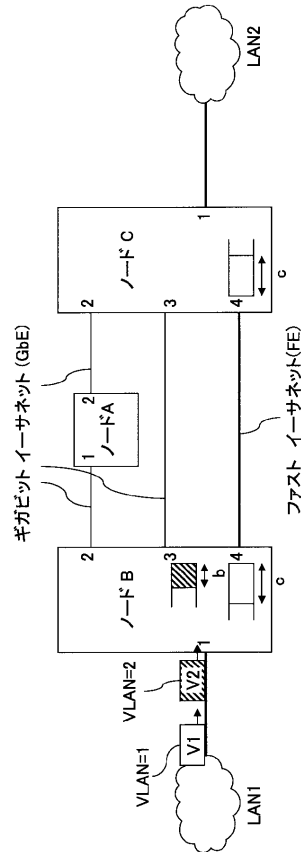
【図14】



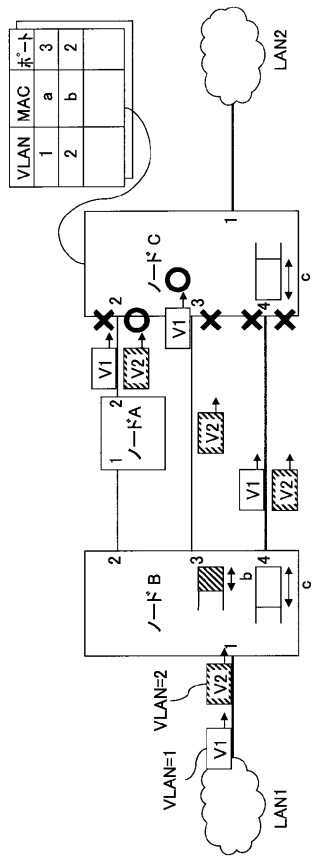
【図15】



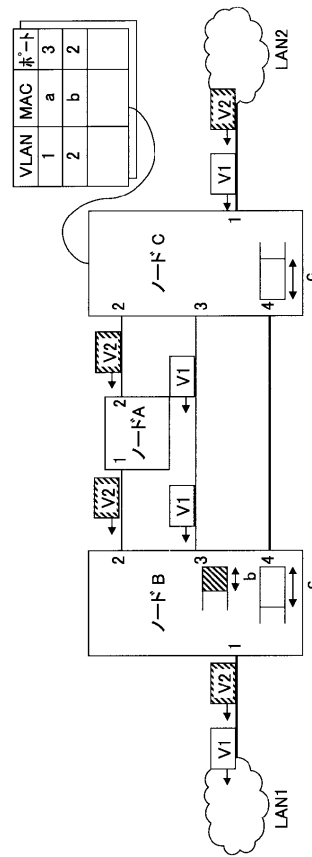
【図16A】



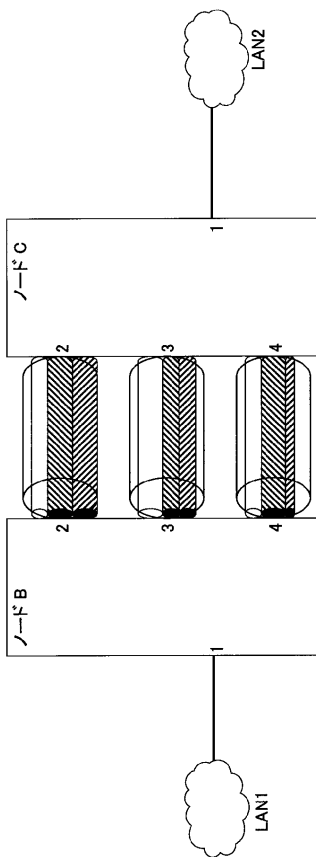
【 図 16 B 】



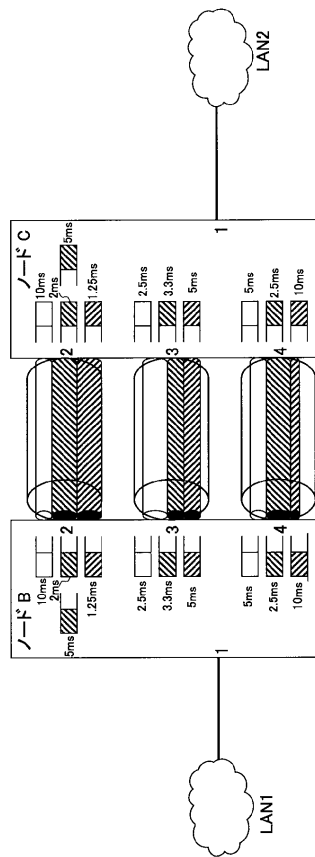
【 図 16 C 】



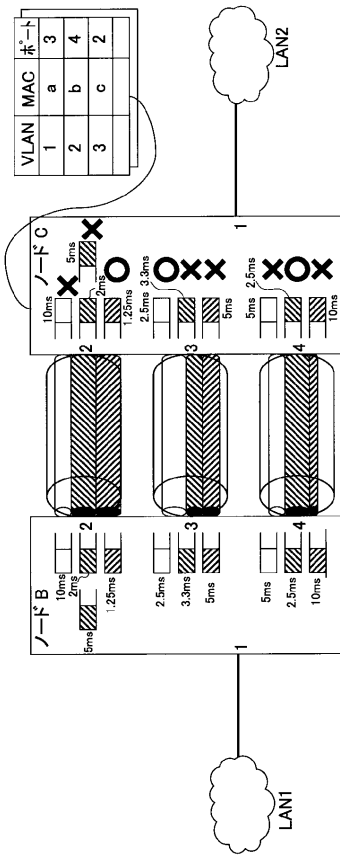
【 図 17 A 】



【 図 17 B 】



【 図 17 C 】



フロントページの続き

- (56)参考文献 特開平06 - 037764 (JP, A)
特開2005 - 303998 (JP, A)
特開平03 - 162035 (JP, A)
特開2005 - 167881 (JP, A)
特開2005 - 136589 (JP, A)

- (58)調査した分野(Int.Cl., DB名)
H04L 12/00-66