(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2012 年 10 月 4 日(04.10.2012)

WIPO | PCT

(10) 国際公開番号
WO 2012/133290 A1

(51) 国際特許分類:
H04L 12/28 (2006.01) *G06F 13/00* (2006.01)
G06F 9/50 (2006.01) *H04L 12/56* (2006.01)

(21) 国際出願番号: PCT/JP2012/057727

(22) 国際出願日: 2012 年 3 月 26 日(26.03.2012)

(25) 国際出願の言語: 日本語

(26) 国際公開の言語: 日本語

(30) 優先権データ:
特願 2011-080955 2011 年 3 月 31 日(31.03.2011) JP

(71) 出願人 (米国を除く全ての指定国について): 日本電気株式会社 (NEC CORPORATION) [JP/JP]; 〒1088001 東京都港区芝五丁目７番１号 Tokyo (JP).

(72) 発明者; および
(75) 発明者/出願人 (米国についてのみ): 上野　洋史 (UENO Hiroshi) [JP/JP]; 〒1088001 東京都港区芝五丁目７番１号　日本電気株式会社内 Tokyo (JP).

(74) 代理人: 工藤　実(KUDOH Minoru); 〒1400013 東京都品川区南大井六丁目２４番１０号カドヤビル６階 Tokyo (JP).

(81) 指定国 (表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

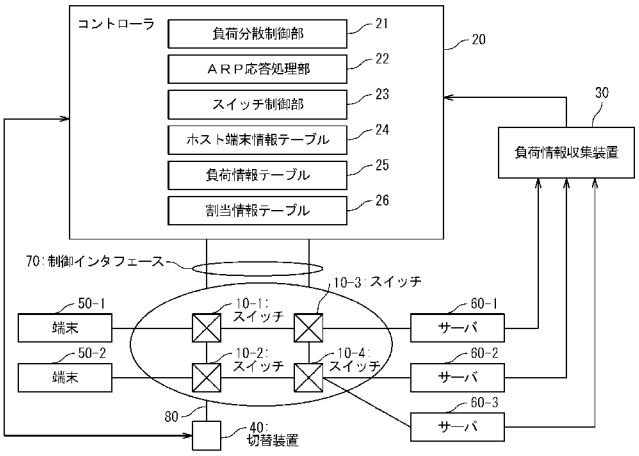
(84) 指定国 (表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR),

[続葉有]

(54) **Title**: COMPUTER SYSTEM, AND COMMUNICATION METHOD

(54) 発明の名称 : コンピュータシステム、及び通信方法

[図1]



20 CONTROLLER
21 LOAD DISTRIBUTION CONTROL UNIT
22 ARP RESPONSE PROCESSING UNIT
23 SWITCH CONTROL UNIT
24 HOST TERMINAL INFORMATION TABLE
25 LOAD INFORMATION TABLE
26 ALLOCATION INFORMATION TABLE
30 LOAD INFORMATION COLLECTING APPARATUS
70 CONTROL INTERFACE
50-1, 50-2 TERMINALS
10-1, 10-2, 10-3, 10-4 SWITCHES
60-1, 60-2, 60-3 SERVERS
40 SWITCHING DEVICE

(57) **Abstract**: In a computer system according to the present invention, a controller uses an ARP response packet, the transmission source of which is the MAC address of a physical server corresponding to a target VIP address included in an ARP request packet from a terminal, to representatively respond to the ARP request packet from the terminal, thereby notifying the terminal of the MAC address of the physical server that is an access destination of the terminal. Also, the controller sets, to a switch, a flow entry that is transferred from the terminal and that specifies to transfer a received packet, the destination MAC address of which is the aforementioned MAC address, to a communication path leading to the aforementioned physical server. In this way, it is made possible to arbitrarily set a server, which is the access destination of a terminal, in a high-speed network environment, while suppressing the increase in cost.

(57) 要約: 本発明によるコンピュータシステムにおいて、コントローラは、端末からのＡＲＰ要求パケットに含まれるターゲットＶＩＰアドレスに対応する物理サーバのＭＡＣアドレスを送信元としたＡＲＰ応答パケットによって、端末からのＡＲＰ要求パケットに対して代理応答することにより、端末のアクセス先となる物理サーバのＭＡＣアドレスを端末に通知する。又、コントローラは、端末から転送され、当該ＭＡＣアドレスを宛先ＭＡＣアドレスとした受信パケットを当該物理サーバに至る通信経路に転送することを規定したフローエントリを、スイッチに設定する。これにより、コスト上昇を抑制しながら、高速化ネットワーク環境下において、端末のアクセス先となるサーバを任意に設定することが可能となる。

WO 2012/133290 A1

OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

添付公開書類:

— 国際調査報告（条約第 21 条(3)）

明　細　書

発明の名称 ： コンピュータシステム、及び通信方法

技術分野

[0001] 本発明は、コンピュータシステム、及びコンピュータシステムにおける通信方法に関し、特にコンピュータシステムにおけるサーバ負荷分散方法に関する。

背景技術

[0002] データセンターなどに構築される大規模システムでは、スケーラビリティを確保するために、サーバ負荷分散装置が利用されている。サーバ負荷分散装置では、仮想ＩＰ（Ｖｉｒｔｕａｌ　Ｉｎｔｅｒｎｅｔ　Ｐｒｏｔｃｏｌ；以下、ＶＩＰと称す）アドレスに対して、実際の振り分け先の宛先サーバを決定し、その宛先にパケットが届くようにパケットヘッダを付け替えるという処理が行われる。例えば、負荷分散装置は、宛先のＶＩＰアドレスを変えずに宛先のＭＡＣアドレスのみを書き換えるＭＡＴ（Ｍｅｄｉａ　Ａｃｃｅｓｓ　Ｃｏｎｔｒｏｌ　Ｔｒａｎｓｌａｔｉｏｎ）処理や、パケットのＩＰアドレスを、実際の宛先のＩＰアドレスに書き換えるＮＡＴ（Ｎｅｔｗｏｒｋ　Ａｄｄｒｅｓｓ　Ｔｒａｎｓｌａｔｉｏｎ）処理を行う。

[0003] 一例として、特開２００８－６０７４７に、レイヤ２の負荷分散を実現するために、ＭＡＴ処理による負荷分散処理方法が記載されている（特許文献１参照）。特許文献１では、負荷分散装置がＡＲＰパケットに対して代理ＡＲＰ応答するとともに、サーバのＭＡＣアドレスを学習し、宛先ＩＰアドレスに応じて、パケットに含まれる負荷分散装置のＭＡＣアドレスをサーバのＭＡＣアドレスに変換する。これにより、ネットワークからのアクセス先サーバを、負荷分散装置によって一元的に変更するとともに、サーバからネットワーク宛のアクセス先も負荷分散装置によって一元的に変更することが可能となる。

[0004] 一方、ネットワークの高速化に伴って、ホストのネットワークインタフェ

ースにおいて１０Ｇｂｐｓの利用が進み、ネットワークスイッチにおいても１０Ｇｂｐｓポートを多数備えたスイッチの利用が一般的になっている。

[0005] しかしながら、高速処理環境では、ヘッダ検索のためのＣＡＭ（Ｃｏｎｔｅｎｔ Ａｄｄｒｅｓｓａｂｌｅ Ｍｅｍｏｒｙ）の高速化が困難であるため、ＭＡＴ処理やＮＡＴ処理の高速化も難しくなる。ＭＡＴ処理を高速化するためには、例えばパケットヘッダを加工するための専用ＡＳＩＣを利用することが考えられるが、この場合、デバイスコストが増大する恐れがある。

[0006] 一般に、Ｌ２スイッチは、固定長の宛先ＭＡＣアドレスのみに基づいて転送を行い、ヘッダの変換を行わない。Ｌ２スイッチはネットワーク構築に不可欠であり、製品価格そのものがスケールメリットで削減されるため、高速ネットワーク環境ではＬ２スイッチのような単純な機能の利活用が望まれる。

[0007] Ｌ２スイッチを用いてサーバに対する負荷分散を行う負荷分散装置が、例えば特開２００９－２５３５７８に記載されている（特許文献２参照）。特許文献２に記載の負荷分散装置は、クライアントからのＡＲＰ要求に対してＡＲＰ応答があった複数のサーバのうち、各サーバとクライアントとの接続状況に基づいて選択したサーバからのＡＲＰ応答をクライアントに転送する。

[0008] 特許文献２では、クライアントは実サーバのＭＡＣアドレスを直接指定してパケットを送信することができるため、ネットワークの高速化の阻害要因となるＭＡＣアドレスの変換処理が不要となる。しかしながら、特許文献２の方法では、クライアントとサーバとの間におけるＡＲＰ中継のために、ネットワーク内にパケット処理装置を挿入しなくてはならず、当該パケット処理装置が、高速処理環境における処理ボトルネックとなることが予想される。

[0009] 一方、コンピュータネットワークにおいて、各スイッチの転送動作等を外部のコントローラによって一元的に制御する技術（オープンフロー）が、ＯｐｅｎＦｌｏｗ Ｃｏｎｓｏｒｔｉｕｍによって提案されている（非特許文

献１参照）。オープンフローに関するシステムの一例が、Ｎｉｃｋ　ＭｃＫｅｏｗｎほか７名による“ＯｐｅｎＦｌｏｗ：　Ｅｎａｂｌｉｎｇ　Ｉｎｎｏｖａｔｉｏｎ　ｉｎ　Ｃａｍｐ　ｕｓ　Ｎｅｔｗｏｒｋｓ”に記載されている（非特許文献２参照）。

[0010] オープンフローは、通信をエンドツーエンドのフローとして捉え、フロー単位で経路制御を行うものである。転送ノードとして機能するオープンフロースイッチは、オープンフロープロトコルに従ってオープンフローコントローラにより追加・更新されるフローテーブルに従って動作する。フローテーブルには、パケットを特定するパケットマッチングルール（以下、ルールと称す）と、「特定のポートに出力する」、「廃棄する」、「ヘッダを書き換える」といった受信パケットに対する動作を規定したアクションの組がフローエントリとして登録される。オープンフロースイッチは、受信パケットのヘッダ情報が、自身に登録されたフローエントリのルールにマッチする場合、当該フローエントリに記載されたアクションに従って受信パケットの処理を行う。又、受信パケットのヘッダ情報とマッチするルール（フローエントリ）がない場合は、オープンフローコントローラに当該パケット（以下、ファーストパケットと称す）を通知する。

[0011] ファーストパケットを受け取ったオープンフローコントローラは、当該パケット内に含まれている送信元ホストや送信先ホストという情報を元に経路を決定し、経路上の全てのオープンフロースイッチに対して、パケットの転送先を規定するフローエントリの設定を指示する（フローテーブル更新指示を発行）。経路上のオープンフロースイッチは、フローテーブル更新指示に応じて、自身で管理しているフローテーブルを更新する。その後、オープンフロースイッチは、更新したフローテーブルに従い、パケットの転送を開始することで、オープンフローコントローラが決定した経路を経由して、宛先のホストへパケットが到達するようになる。

[0012] オープンフロープロトコルを利用したネットワークでは、コントローラ（制御系）とスイッチ（パケット転送系）が分離されるため、既存のスイッチ

動作への機能拡張が容易に実現できる。しかし、既存のオープンフロースイッチは、ヘッダ参照やヘッダ変換を伴うスイッチ処理を行うため、ネットワークを高速化する場合、上述と同様に専用の高速デバイスが必要となりコストの上昇を招く恐れがある。

先行技術文献

特許文献

[0013] 特許文献1：特開２００８－６０７４７

特許文献2：特開２００９－２５３５７８

非特許文献

[0014] 非特許文献1：ＯｐｅｎＦｌｏｗ Ｓｗｉｔｃｈ Ｓｐｅｃｉｆｉｃａｔｉｏｎ Ｖｅｒｓｉｏｎ １．０．０ (Ｗｉｒｅ Ｐｒｏｔｏｃｏｌ ０ｘ０１) Ｄｅｃｅｍｂｅｒ ３１， ２００９

非特許文献2：Ｎｉｃｋ ＭｃＫｅｏｗｎほか７名、“ＯｐｅｎＦｌｏｗ：Ｅｎａｂｌｉｎｇ Ｉｎｎｏｖａｔｉｏｎ ｉｎ Ｃａｍｐ ｕｓ Ｎｅｔｗｏｒｋｓ”、［ｏｎｌｉｎｅ］、［平成２２年１０月２２日検索］、インターネット＜ＵＲＬ：ｈｔｔｐ：／／ｗｗｗ．ｏｐｅｎｆｌｏｗｓｗｉｔｃｈ．ｏｒｇ／／ｄｏｃｕｍｅｎｔｓ／ｏｐｅｎｆｌｏｗ－ｗｐ－ｌａｔｅｓｔ．ｐｄｆ＞

発明の概要

[0015] 以上のことから、本発明の目的は、コスト上昇を抑制しながら、高速化ネットワーク環境下において、端末のアクセス先となるサーバを任意に設定することにある。

[0016] 本発明の他の目的は、高速化ネットワーク環境下において負荷分散を実現するコンピュータシステムのコスト上昇を抑制することにある。

[0017] 本発明の更に他の目的は、高速ネットワーク環境下において、高価なネットワークデバイスを用いなくてもスケーラビリティの高いサーバ負荷分散処理を実現することにある。

[0018] 本発明の更に他の目的は、アドレス変換が必要なレイヤ４の負荷分散を高速ネットワーク環境下で実現することにある。

[0019] 本発明による一の態様に係るコンピュータシステムは、コントローラと、コントローラによって設定されたフローエントリに適合する受信パケットに対し、フローエントリで規定された中継動作を行うスイッチとを具備する。コントローラは、ＶＩＰ（Ｖｉｒｔｕａｌ Ｉｎｔｅｒｎｅｔ Ｐｒｏｔｃｏｌ）アドレスと物理サーバとを対応付けて記録されたテーブルを参照して、端末からのＡＲＰ要求パケットに含まれるターゲットＶＩＰアドレスに対応する物理サーバを、前記端末のアクセス先となる物理サーバとして選択する。又、コントローラは、選択した物理サーバのＭＡＣ（Ｍｅｄｉａ Ａｃｃｅｓｓ Ｃｏｎｔｒｏｌ）アドレスを送信元ＭＡＣアドレスとしたＡＲＰ応答パケットによって、ＡＲＰ要求パケットに対して代理応答する。これにより、端末のアクセス先となる物理サーバのＭＡＣアドレスは端末に通知される。更にコントローラは、端末から転送され、当該ＭＡＣアドレスを宛先ＭＡＣアドレスとした受信パケットを当該物理サーバに至る通信経路に転送することを規定したフローエントリを、スイッチに設定する。

[0020] 又、本発明による他の態様に係るコンピュータシステムは、コントローラと、コントローラによって設定されたフローエントリに適合する受信パケットに対し、フローエントリで規定された中継動作を行うスイッチと、切替装置とを具備する。スイッチは、コントローラによって設定された第１フローエントリに従い、第１サーバの第１ＭＡＣアドレスを宛先ＭＡＣアドレスとしたパケットを、第１サーバに至る通信経路へ転送する。コントローラは、第１サーバの第１ＭＡＣアドレスを宛先ＭＡＣアドレスとしたパケットを切替装置に転送することを規定した第２フローエントリを、第１フローエントリに替えてスイッチに設定する。スイッチは、第２フローエントリに従い、第１ＭＡＣアドレスを宛先ＭＡＣアドレスとしたパケットを、切替装置へ転送する。切替装置は、スイッチからの受信パケットの宛先ＭＡＣアドレスを、第１ＭＡＣアドレスから、第１サーバとは異なる第２サーバの第２ＭＡＣ

アドレスに変更して、第２サーバに至る通信経路へ転送する。

[0021] 上述のコントローラは、通信制御プログラムをコンピュータによって実行することで、実現されることが好ましい。

[0022] 又、上述の切替装置は、通信制御プログラムをコンピュータによって実行することで実現されることが好ましい。

[0023] 本発明による他の態様における通信方法は、コントローラによって設定されたフローエントリに適合する受信パケットに対し、前記フローエントリで規定された中継動作を行うスイッチを備えるコンピュータシステムにおいて実行される通信方法である。本発明による通信方法は、コントローラが、ＶＩＰ（Ｖｉｒｔｕａｌ Ｉｎｔｅｒｎｅｔ Ｐｒｏｔｃｏｌ）アドレスと物理サーバとを対応付けて記録されたテーブルを参照して、端末からのＡＲＰ要求パケットに含まれるターゲットＶＩＰアドレスに対応する物理サーバを、前記端末のアクセス先となる物理サーバとして選択するステップと、コントローラが、当該物理サーバのＭＡＣ（Ｍｅｄｉａ Ａｃｃｅｓｓ Ｃｏｎｔｒｏｌ）アドレスを、送信元ＭＡＣアドレスとしたＡＲＰ応答パケットによって、ＡＲＰ要求パケットに対して代理応答することにより、端末のアクセス先となる物理サーバのＭＡＣアドレスを端末に通知するステップと、端末から転送され、当該ＭＡＣアドレスを宛先ＭＡＣアドレスとした受信パケットを物理サーバに至る通信経路に転送することを規定したフローエントリを、コントローラがスイッチに設定するステップとを具備する。

[0024] 又、本発明による他の態様における通信方法は、コントローラによって設定されたフローエントリに適合する受信パケットに対し、フローエントリで規定された中継動作を行うスイッチを備えるコンピュータシステムにおいて実行される通信方法である。本発明による通信方法は、スイッチが、コントローラによって設定された第１フローエントリに従い、第１サーバの第１ＭＡＣアドレスを宛先ＭＡＣアドレスとしたパケットを、第１サーバに至る通信経路へ転送するステップと、コントローラが、第１サーバの第１ＭＡＣアドレスを宛先ＭＡＣアドレスとしたパケットを切替装置に転送することを規

定した第２フローエントリを、第１フローエントリに替えて前記スイッチに設定するステップと、スイッチが、第２フローエントリに従い、第１ＭＡＣアドレスを宛先ＭＡＣアドレスとしたパケットを、切替装置へ転送するステップと、切替装置が、スイッチからの受信パケットの宛先ＭＡＣアドレスを、第１ＭＡＣアドレスから、第１サーバとは異なる第２サーバの第２ＭＡＣアドレスに変更して、第２サーバに至る通信経路へ転送するステップを具備する。

[0025] 本発明によれば、コスト上昇を抑制しながら高速化ネットワーク環境下において、端末のアクセス先となるサーバを任意に設定することができる。

[0026] 又、コスト上昇を抑制しながら高速化ネットワーク環境下におけるサーバ負荷分散処理を実現する。

[0027] 更に、高速ネットワーク環境下において、高価なネットワークデバイスを用いなくてもスケーラビリティの高いサーバ負荷分散処理を実現する。

[0028] 更に、アドレス変換が必要なレイヤ４の負荷分散を高速ネットワーク環境下で実現する。

図面の簡単な説明

[0029] 上記発明の目的、効果、特徴は、添付される図面と連携して実施の形態の記述から、より明らかになる。

[図1]図１は、本発明によるコンピュータシステムの第１の実施の形態における構成を示す図である。

[図2]図２は、本発明に係るホスト端末情報テーブルの構造の一例を示す図である。

[図3]図３は、本発明に係る負荷情報テーブルの構造の一例を示す図である。

[図4]図４は、本発明に係る割当情報テーブルの構造の一例を示す図である。

[図5]図５は、本発明に係るオープンフロースイッチの構成の一例を示す図である。

[図6]図６は、本発明に係る切替装置の構成の一例を示す図である。

[図7]図７は、切替情報テーブルの構造の一例を示す図である。

[図8A]図８Ａは、ＡＲＰ代理応答する場合の負荷分散動作の一例を示すシーケンス図である。

[図8B]図８Ｂは、ＡＲＰ代理応答する場合の負荷分散動作の一例を示すシーケンス図である。

[図9]図９は、本発明によるコントローラにおけるＡＲＰ応答処理の動作の一例を示すフロー図である。

[図10A]図１０Ａは、コントローラにおいてＡＲＰ代理応答を行わない場合のコンピュータシステムの動作を示すシーケン図である。

[図10B]図１０Ｂは、コントローラにおいてＡＲＰ代理応答を行わない場合のコンピュータシステムの動作を示すシーケン図である。

[図11]図１１は、本発明によるコンピュータシステムにおける宛先サーバ切替動作の一例を示すシーケン図である。

[図12]図１２は、本発明によるコントローラ２０におけるサーバ削除処理（宛先サーバ切替処理）の動作の一例を示すフロー図である。

[図13]図１３は、本発明によるコンピュータ装置において、セッションを維持しながら宛先サーバを切替える動作の一例を示すシーケンス図である。

[図14]図１４は、本発明によるコントローラにおける切替処理の動作の一例を示すフロー図である。

[図15]図１５は、本発明による切替装置におけるパケット制御処理の動作の一例を示すフロー図である。

[図16]図１６は、本発明による切替装置４０におけるフローエントリのエキスパイアに応じた切替動作を示すフロー図である。

発明を実施するための形態

[0030] 以下、添付図面を参照しながら本発明の実施の形態を説明する。図面において同一、又は類似の参照符号は、同一、類似、又は等価な構成要素を示す。又、同様な構成が複数ある場合は、符号に追番を付して説明し、複数の構成を区別せずに説明する場合は、追番のない符号を付して説明する。

[0031] （概要）

本発明によるコンピュータシステムは、オープンフロー技術を利用して通信経路の構築及びパケットデータの転送制御を行う。本発明によるコンピュータシステムでは、オープンフローコントローラ（以下、コントローラと称す）が、複数の物理サーバのそれぞれにおける負荷状況に応じて、端末のアクセス先となる物理サーバを決定する。この際、コントローラは、オープンフロースイッチ（以下、スイッチと称す）から端末からのＡＲＰ要求を収集し、ＡＲＰ要求に対して代理応答することで、端末のアクセス先となる物理サーバのＭＡＣアドレスを端末に通知する。又、コントローラは、端末と物理サーバとを端点とした経路上のオープンフロースイッチに、端末から物理サーバ宛のパケットを中継するためのフローエントリを設定する。このように、本発明によるシステムでは、レイヤ２スイッチを利用したまま、アクセス先となる物理サーバを決定するため、高速化ネットワークにおける負荷分散を低コストで実現することが可能となる。

[0032] （コンピュータシステムの構成）

図１から図７を参照して、本発明によるコンピュータシステムの構成の一例を説明する。図２は、本発明によるコンピュータシステムの構成の一例を示す図である。

[0033] 図１を参照して、本発明によるコンピュータシステムは、複数のスイッチ１０－１～１０－４、コントローラ２０、負荷情報収集装置３０、切替装置４０、複数の端末５０－１、５０－２、及び複数のサーバ６０－１～６０－３を具備する。ここで端末５０－１、２、サーバ６０－１～６０－３は、同一のＥｔｈｅｒｎｅｔ　ＬＡＮに接続されたコンピュータ装置に例示される。

[0034] コントローラ２０は、セキュアネットワークである制御インタフェース７０を介して、複数のスイッチ１０－１～１０－４のそれぞれに接続される。コントローラ２０は、オープンフロープロトコルに従って、複数のスイッチ１０－１～１０－４のそれぞれにフローエントリを設定することで、複数のスイッチ１０－１～１０－４のいずれかを介した複数の端末５０－１、５０

－２と複数のサーバ６０－１～６０－３との間の通信を制御する。

[0035] 詳細には、コントローラ２０は、オープンフロー技術により、システム内におけるパケット転送に係る通信経路の構築及びパケット転送処理を制御する。ここで、オープンフロー技術とは、コントローラ２０が、ルーティングポリシー（フローエントリ：フロー＋アクション）に従い、マルチレイヤ及びフロー単位の経路情報を通信経路上のスイッチ１０に設定し、経路制御やノード制御を行う技術を示す（詳細は、非特許文献１を参照）。これにより、経路制御機能がルータやスイッチから分離され、コントローラによる集中制御によって最適なルーティング、トラフィック管理が可能となる。オープンフロー技術が適用されるスイッチ１０は、従来のルータやスイッチのようにパケットやフレームの単位ではなく、ＥＮＤ２ＥＮＤのフローとして通信を取り扱う。

[0036] 図１を参照して、本発明に係るコントローラ２０の構成の詳細を説明する。コントローラ２０は、ＣＰＵ及び記憶装置を備えるコンピュータによって実現されることが好適である。コントローラ２０では、図示しないＣＰＵが記憶装置に格納されたプログラムを実行することで、図１に示す負荷分散制御部２１、ＡＲＰ応答処理部２２、スイッチ制御部２３の各機能が実現される。又、コントローラ２０は、図示しない記憶装置に記録されたホスト端末情報テーブル２４、負荷情報テーブル２５、割当情報テーブル２６を備える。

[0037] 図２は、ホスト端末情報テーブル２４の構造の一例を示す図である。図２を参照して、ホスト端末情報テーブル２４には、端末５０やサーバ６０を識別するホスト端末ＩＤ２４が、端末５０やサーバ６０に設定されたＩＰアドレス２４２及びＭＡＣアドレス２４３に対応付けられて記録される。すなわち、ホスト端末情報テーブル２４には、端末５０やサーバ６０毎に設定されたＩＰアドレスやＭＡＣアドレスが記録される。

[0038] 負荷分散制御部２１は、サーバ６０－１～６０－３のそれぞれから負荷情報収集装置３０を介して負荷情報を収集し、負荷情報テーブル２５に記録及

び管理する。図３は、負荷情報テーブル２５の構造の一例を示す図である。負荷情報テーブル２５には、サーバ６０を特定するサーバ情報２５１（例えばサーバ６０のホスト端末ＩＤ）に、当該サーバに設定された仮想ＩＰアドレス２５２（以下、ＶＩＰ２５２と称す）と負荷情報２５３とが対応付けられて記録される。負荷情報２５３は、ＶＩＰ２５２によって特定されるアプリケーションの負荷やＶＩＰ２５２が割り当てられたサーバの負荷（サーバ負荷）を示す。又、サーバ情報２５１は、ホスト端末情報テーブル２４におけるホスト端末ＩＤ２４１に紐付けられる。これにより、負荷分散制御部２１は、負荷情報テーブル２５を参照することで、ＶＩＰアドレスとサーバ６０の実アドレス（ＩＰアドレス２４２、ＭＡＣアドレス２４３）との関係と、サーバ６０の負荷情報２５３を管理することが可能となる。

[0039] 負荷情報２５３は、負荷情報収集装置３０によって定期的又は任意の時期に各サーバから収集され、サーバ情報２５１やＶＩＰ２５２とともに負荷分散制御部２１に送信される。例えば、負荷情報収集装置３０は、ＳＮＭＰ（Ｓｉｍｐｌｅ Ｎｅｔｗｏｒｋ Ｍａｎａｇｅｍｅｎｔ Ｐｒｏｔｏｃｏｌ）、ＣＬＩ（ｃｏｍｍａｎｄ ｌｉｎｅ ｕｓｅｒ ｉｎｔｅｒｆａｃｅ）等のＡＰＩを介してアプリケーション負荷を収集する方法や、サーバ内におかれた負荷収集エージェントから情報を収集する方法などにより負荷情報２５３を収集する。負荷分散制御部２１は、負荷情報収集装置３０から定期的又は任意の時期に、負荷情報２５３を負荷情報収集装置３０から取得し、負荷情報テーブル２５を更新する。この際、負荷分散制御部２１は、能動的又は受動的に負荷情報２５３を負荷情報収集装置３０から取得する。

[0040] 又、負荷分散制御部２１は、端末５０と、当該端末５０のアクセス先として割り当てたサーバ６０とを対応付けた情報を割当情報テーブル２６に記録及び管理する。図４は、割当情報テーブル２６の構造の一例を示す図である。割当情報テーブル２６には、宛先情報２６１、端末情報２６２、割当サーバ情報２６３が対応付けられて記録される。宛先情報２６１は、端末５０のアクセス先のアプリケーションや仮想サーバを特定する情報であり、例えば

仮想マシンに設定されたＶＩＰアドレスである。あるいは宛先情報２６１は、端末５０のアクセス先のサーバを特定するＩＰアドレスでもよい。端末情報２６２は、端末５０を特定する情報であり、例えば、端末５０を識別するホスト端末ＩＤである。割当サーバ情報２６３は、現在、端末５０のアクセス先として割り当てられているサーバ６０（物理サーバ）を特定する情報であり、例えば、サーバ６０を識別するホスト端末ＩＤである。割当サーバ情報２６３が、ホスト端末情報テーブル２４のホスト端末ＩＤ２４１に紐付けられることで、端末５０のアクセス先として割り当てられたサーバ６０のＭＡＣアドレスを特定することが可能となる。旧宛先サーバ情報２６４は、“アクセス先の切替前”に端末５０のアクセス先として割り当てられていたサーバ６０（物理サーバ）を特定する情報である。負荷分散制御部２１は、端末５０のアクセス先サーバを変更すると、割当サーバ情報２６３の内容を旧宛先サーバ情報２６４として記録し、新たに端末５０の宛先として割当てたサーバのＭＡＣアドレスを割当サーバ情報２６３として更新する。

[0041] ＡＲＰ応答処理部２２は、制御インタフェース７０を介してスイッチ１０から通知されるＡＲＰ要求パケットに対して代理ＡＲＰ応答を実行する。詳細には、ＡＲＰ応答処理部２２は、ＡＲＰ要求パケットに含まれるターゲットアドレスに対応するホスト端末（例えばサーバ６０）のＭＡＣアドレスを送信元としたＡＲＰ応答パケットを要求元の端末５０に送信する。この際、ターゲットアドレスがサーバ６０のＩＰアドレスの場合、ＡＲＰ応答処理部２２は、ホスト端末情報テーブル２４を参照して、ターゲットアドレスに一致するＩＰアドレス２４２が設定されたサーバ６０をＡＲＰ要求先として特定するとともに、そのサーバ６０のＭＡＣアドレス２４３をＡＲＰ応答パケットの送信元ＭＡＣアドレスとして設定する。一方、ターゲットアドレスがＶＩＰアドレスの場合、ＡＲＰ応答処理部２２は、負荷情報テーブル２５を参照して、ターゲットアドレスに一致するＶＩＰ２５２に対応付けられたサーバ情報２５１からＡＲＰ要求先のサーバ６０を特定し、ホスト端末情報テーブル２４において当該サーバ情報２５１に紐付けられたホスト端末ＩＤ２

４に対応するＭＡＣアドレス２４３をＡＲＰ応答パケットの送信元ＭＡＣアドレスとして設定する。

[0042] 以上のような構成により、本発明に係るコントローラ２０は、サーバ６０毎に設定されたＭＡＣアドレスやＶＩＰアドレスを管理するとともに、ＡＲＰ要求のターゲットアドレスに対応するサーバ６０のＭＡＣアドレスを送信元としたＡＲＰ応答パケットを当該サーバ６０に替わって代理応答する。

[0043] スイッチ制御部２３は、オープンフロープロトコルに従ったスイッチ１０に対してフローエントリ（ルール＋アクション）の設定又は削除を行う。これにより、スイッチ１０は、受信パケットのヘッダ情報に応じたルールに対応するアクション（例えばパケットデータの中継や破棄）を実行する。

[0044] フローエントリに設定されるルールには、例えば、ＴＣＰ／ＩＰ（Ｔｒａｎｓｍｉｓｓｉｏｎ Ｃｏｎｔｒｏｌ Ｐｒｏｔｏｃｏｌ／Ｉｎｔｅｒｎｅｔ Ｐｒｏｔｃｏｌ）のパケットデータにおけるヘッダ情報に含まれる、ＯＳＩ（Ｏｐｅｎ Ｓｙｓｔｅｍｓ Ｉｎｔｅｒｃｏｎｎｅｃｔｉｏｎ）参照モデルのレイヤ１からレイヤ４のアドレスや識別子の組み合わせが規定される。例えば、レイヤ１の物理ポート、レイヤ２のＭＡＣアドレス、ＶＬＡＮタグ（ＶＬＡＮ ｉｄ）、レイヤ３のＩＰアドレス、レイヤ４のポート番号のそれぞれの組み合わせがルールとして設定される。尚、ＶＬＡＮタグには、優先順位（ＶＬＡＮ Ｐｒｉｏｒｉｔｙ）が付与されていても良い。

[0045] ここで、スイッチ制御部２３によってルールに設定されるポート番号等の識別子やアドレス等は、所定の範囲で設定されても構わない。又、宛先や送信元のアドレス等を区別してルールとして設定されることが好ましい。例えば、ＭＡＣ宛先アドレスの範囲や、接続先のアプリケーションを特定する宛先ポート番号の範囲、接続元のアプリケーションを特定する送信元ポート番号の範囲がルールとして設定される。更に、データ転送プロトコルを特定する識別子をルールとして設定してもよい。

[0046] フローエントリに設定されるアクションには、例えばＴＣＰ／ＩＰのパケットデータを処理する方法が規定される。例えば、受信パケットデータを中

継するか否かを示す情報や、中継する場合はその送信先が設定される。又、アクションとして、パケットデータの複製や、破棄することを指示する情報が設定されてもよい。

[0047] スイッチ制御部２３は、スイッチ１０からのファーストパケットの通知やＡＲＰ要求パケットの通知に応じてスイッチ１０に対するフローエントリの設定を行う。スイッチ制御部２３における通信経路の設定、フローエントリの設定及び管理は、非特許文献１に記載のオープンフロープロトコルに準拠して行われるため、ここでの詳細な説明は省略する。ただし、当該プロトコルで規定されていない動作については後述する。

[0048] スイッチ１０は、コントローラ２０によって設定（更新）されたフローエントリに従って受信パケットの処理方法（アクション）を決定する。受信パケットに対するアクションとして、例えば、他のスイッチ１０への転送や破棄がある。

[0049] 図５は、本発明に係るスイッチ１０の構成の一例を示す図である。図５を参照して、スイッチ１０は、スイッチ処理部１１、フローテーブル１２、スイッチ制御部１３を備える。スイッチ処理部１１は、回線１４－１～１４－Ｎを収容し、収容した回線を介して転送されるパケットに対するスイッチ機能を実現する。スイッチ制御部１３は、制御インタフェース７０を介してコントローラ２０に接続され、コントローラ２０からの指示に応じてフローテーブル１２へのフローエントリの設定を制御する。スイッチ処理部１１は、フローテーブル１２に設定されたフローエントリに従って、受信パケットに対する処理を決定する。詳細には、スイッチ処理部１１は、受信パケットのヘッダ情報に適合するルールがフローテーブル１２に設定されている場合、当該ルールを含むフローエントリに規定されたアクションを実行する。一方、受信パケットのヘッダ情報に適合するルールがフローテーブル１２に設定されていない場合、当該パケットデータをファーストパケットとして、スイッチ制御部１３に通知する。スイッチ制御部１３は、ファーストパケットを、制御インタフェース７０を介してコントローラ２０に送信する。

[0050] 又、スイッチ処理部１１は、受信パケットがＡＲＰ要求である場合、受信パケットをスイッチ制御部１３に通知する。スイッチ制御部１３は、通知された受信パケット（ＡＲＰ要求パケット）を、制御インタフェース７０を介してコントローラ２０に送信する。

[0051] 切替装置４０は、セキュアネットワークである制御インタフェース８０を介して、複数のスイッチ１０－１～１０－４のそれぞれに接続される。切替装置４０は、ＴＣＰやＵＤＰなどのレイヤ４の粒度でトラヒックを処理できるスイッチやサーバで実現できる。図６は、本発明に係る切替装置４０の構成の一例を示す図である。切替装置４０は、ＣＰＵ及び記憶装置を備えるコンピュータによって実現されることが好適である。切替装置４０では、図示しないＣＰＵが記憶装置に格納されたプログラムを実行することで、図６に示すスイッチ処理部４１、スイッチ制御部４３、切替制御部４４の各機能が実現される。又、切替装置４０の記憶装置（図示なし）には、フローテーブル４２と切替情報テーブル４５が記録される。

[0052] スイッチ処理部４１は、制御インタフェース８０を介して複数のスイッチ１０－１～１０－４で構成されるスイッチネットワークに接続され、制御インタフェース８０を介して転送されるパケットに対するスイッチ機能を実現する。スイッチ制御部４３は、切替情報テーブル４５に基づいてフローテーブル４２へのフローエントリの設定を制御する。切替情報テーブル４５は、切替制御部４４によって記録、管理される。図７は、切替情報テーブル４５の構造の一例を示す図である。図７を参照して、切替制御部４４は、コントローラ２０から送信された宛先情報４５１、端末情報４５２、新宛先サーバ情報４５３、旧宛先サーバ情報４５４を対応付けて記録する。宛先情報４５１は、端末５０のアクセス先のアプリケーションや仮想サーバを特定する情報であり、例えば仮想マシンに設定されたＶＩＰアドレスである。端末情報２６２は、端末５０を特定する情報であり、例えば、端末５０のＭＡＣアドレスである。新宛先サーバ情報４５３は、新たに、端末５０のアクセス先として割り当てられたサーバ６０（物理サーバ）を特定する情報であり、例え

ば、サーバ６０のＭＡＣアドレスである。旧宛先サーバ情報４５４は、“アクセス先の切替前”に端末５０のアクセス先として割り当てられていたサーバ６０（物理サーバ）を特定する情報であり、例えばサーバ６０のＭＡＣアドレスである。

[0053] スイッチ処理部４１は、受信パケットを解析し、解析結果をスイッチ制御部４３に通知する。スイッチ制御部４３は通知された解析結果に基づいて、フローテーブル４２に設定されたフローエントリのアクションを変更する。例えば、スイッチ処理部４１は、受信パケットがＳＹＮフラグを含むＴＣＰパケットであるかや、受信パケットに適合するフローエントリがあるかを判定し、判定結果をスイッチ制御部４３に通知する。スイッチ制御部４３は、受信パケットがＳＹＮフラグを含むＴＣＰパケットである場合や、受信パケットがＴＣＰパケットではない場合、切替情報テーブル４５を参照して、受信パケットに適合するフローエントリのアクションで規定した宛先を、新宛先サーバ情報４５３で示されるＭＡＣアドレスに変更する。あるいは、スイッチ制御部４３は、受信パケットがＳＹＮフラグを含むＴＣＰパケットではなく、当該パケットに適合するフローエントリが設定されている場合、当該フローエントリに規定されたアクションを実行する。

[0054] 以上のような構成により、本発明によるコンピュータシステムでは、サーバの負荷に応じて端末のアクセス先の変更をコントローラ２０によって一元制御するとともに、切替装置４０を利用することで、セッションを維持しながら端末５０のアクセス先のサーバ６０を他のサーバに切り替えることが可能となる。

[0055] （動作）

以下、図８Ａから図１６を参照して、本発明によるコンピュータシステムの動作の詳細を説明する。以下では、（１）ＡＲＰ代理応答する場合の負荷分散動作、（２）ＡＲＰ代理応答しない場合の動作、（３）宛先サーバ切替動作、（４）セッションを維持しながら宛先サーバを切替える動作、及び（５）フローエントリの削除に応じた切替終了動作の５つの動作について説明

する。

[0056] （１）ＡＲＰ代理応答する場合の負荷分散動作

図８Ａ、図８Ｂ及び図９を参照して、本発明によるコンピュータシステムのＡＲＰ代理応答する場合の負荷分散動作を説明する。ここでは、端末５０－１が、サーバ６０－１～６０－２に対して割り当てられているＶＩＰアドレス宛に通信を開始するものとして説明する。図８Ａ及び図８Ｂは、ＡＲＰ代理応答する場合の負荷分散動作の一例を示すシーケンス図である。図９は、本発明によるコントローラ２０におけるＡＲＰ応答処理の動作の一例を示すフロー図である。

[0057] 本動作に先立って、コントローラ２０は、スイッチ１０－１においてＡＲＰパケットを受信した場合にはコントローラに送信するフローエントリをスイッチ１０－１のフローテーブル１２に記述しておく。例えば、端末５０－１に接続するインタフェース（ポート番号）と、ＡＲＰ要求を示す情報としてイーサネットのタイプフィールドが０ｘ０８０６であるマッチ条件をルールとし、コントローラ２０への転送をアクションとするフローテエントリがスイッチ１０－１に設定される。

[0058] 先ず、端末５０－１は、アクセス先のＶＩＰアドレスに対するレイヤ２のアドレスであるＭＡＣアドレスを取得するため、ＡＲＰ要求パケットをブロードキャスト送信する（ステップＳ１０１）。ここで送信されるＡＲＰ要求パケットは、宛先がブロードキャストアドレス、送信元が端末５０－１のＭＡＣアドレスであり、ペイロー途中に解決したいターゲットであるＶＩＰアドレスが示されている。

[0059] 端末５０－１に隣接するスイッチ１０－１は、端末５０－１からのパケットを受信すると、フローテーブル１２を参照し該当エントリを検索し、当該パケットに適合する（マッチする）フローエントリが設定されているかを判定する（ステップＳ１０２）。ここで、受信パケットは、端末５０－１に接続するインタフェースを介したＡＲＰ要求パケットであるため、スイッチ１０－１は、当該パケットに適合するフローエントリに規定されたアクション

“受信したＡＲＰ要求パケットをコントローラ２０に通知する”を実行する（ステップＳ１０２Ｎｏ、Ｓ１０３Ｙｅｓ、Ｓ１０４）。コントローラ２０は、スイッチ１０－１から通知されたＡＲＰ要求パケットに対して、ＡＲＰ応答処理を実行する（ステップＳ１０５）。

[0060] 図９を参照して、ステップＳ１０５におけるＡＲＰ応答処理の詳細を説明する。コントローラ２０のＡＲＰ応答処理部２２は、通知されたＡＲＰ要求パケットのターゲットアドレスを参照し、当該要求がＶＩＰアドレスに対するＡＲＰ要求か否かを判定する（ステップＳ１１）。ここでは、ＡＲＰ要求パケットがＶＩＰアドレス宛のパケットであるため、ＡＲＰ応答処理部２２は、負荷分散制御部２１から当該ＶＩＰアドレスに対する実サーバ情報（ＭＡＣアドレス）を取得する（ステップＳ１１Ｙｅｓ、Ｓ１２）。

[0061] 詳細には、負荷分散制御部２１は、負荷情報テーブル２５を参照して、受信パケットの宛先ＶＩＰアドレスが割り当てられたサーバ（サーバ情報２５１）を特定し、特定したサーバ情報２５１に対応するＭＡＣアドレス２４３をホスト端末情報テーブル２４から抽出し、ＡＲＰ応答処理部２２に渡す。この際、宛先ＶＩＰアドレスが割り当てられたサーバが複数ある場合、負荷分散制御部２１は、複数のサーバ（サーバ情報２５１）のそれぞれに対応する負荷情報２５３に基づいて、アクセス先として割り当てるサーバを決定する。例えば、負荷分散制御部２１は、宛先ＶＩＰアドレスが割り当てられた複数のサーバのうち、最も小さい負荷のサーバ６０－１を、端末５０－１のアクセス先として割り当て、当該サーバのＭＡＣアドレスをＡＲＰ応答処理部２２に渡す。負荷分散制御部２１は、ＡＲＰ要求パケットのターゲットＶＩＰアドレスにサーバを割り当てると、この割り当てた結果を割当情報テーブル２６に記録する。詳細には、負荷分散制御部２１は、ＡＲＰ要求パケットのターゲットＶＩＰアドレスを宛先情報２６１、ＡＲＰ要求パケットの送信元アドレス（例えば送信元ＭＡＣアドレス）に対応するホスト端末ＩＤを端末情報２６２、当該ＶＩＰアドレスに割り当てたサーバのホスト端末ＩＤを割当サーバ情報２６３とするとともに、それぞれを対応付けて割当情報テ

ーブル２６に記録する。

[0062] あるいは、負荷情報２５３に応じてＶＩＰアドレスに割当られたサーバが、予め割当情報テーブル２６に記録されている場合、負荷分散制御部２１は、割当情報テーブル２６から、ターゲットＶＩＰアドレスに対応する割当サーバ情報２６３を抽出し、割当サーバ情報２６３に対応するＭＡＣアドレス２４３をＡＲＰ応答処理部２２に渡す。尚、ステップＳ１２において、負荷分散制御部２１は、割当情報テーブル２６を参照した後、予めＶＩＰアドレスに割当てられたサーバが記録されていない場合、負荷情報テーブル２５に基づいてターゲットＶＩＰアドレスに割当てるサーバを決定することが好ましい。

[0063] 一方、ステップＳ１１において、受信パケットがＶＩＰアドレスに対するＡＲＰ要求パケットでない場合（例えば宛先ＩＰアドレスに対するＡＲＰ要求パケットの場合）、ＡＲＰ応答処理部２２は、代理応答するか否かを判定する（ステップＳ１４）。例えば、コントローラ２０に、ネットワークに接続している端末の情報が予め設定されている場合、無駄なトラヒックを防ぐため、受信したＡＲＰ要求パケットをフラッディング（ブロードキャスト）しなくても代理で応答を返すことができる。ステップＳ１４においてＡＲＰ応答処理部２２は、所定の端末からのＡＲＰ要求に対して代理応答することが設定されているか否かを判定し、設定されている場合、当該設定に基づいてＡＲＰ応答を行う（ステップＳ１４Ｙｅｓ、Ｓ１５、Ｓ１３）。ここで、ＡＲＰ要求に対する代理応答をする場合、ＡＲＰ応答処理部２２は、ファーストパケットの宛先ＩＰアドレスに対応するＭＡＣアドレス２４３をホスト端末情報テーブル２４から取得する（ステップＳ１５）。負荷分散制御部２１は、ＡＲＰ要求パケットのターゲットＩＰアドレスにサーバが割り当てられると、この割り当てられた結果を割当情報テーブル２６に記録することが好ましい。詳細には、負荷分散制御部２１は、ＡＲＰ要求パケットのターゲットＩＰアドレスを宛先情報２６１、ＡＲＰ要求パケットの送信元アドレス（例えば送信元ＭＡＣアドレス）に対応するホスト端末ＩＤを端末情報２６

２、当該ＩＰアドレスに割り当てたサーバのホスト端末ＩＤを割当サーバ情報２６３とするとともに、それぞれを対応付けて割当情報テーブル２６に記録する。

[0064] ＡＲＰ応答処理部２２は、負荷分散制御部２１又はホスト端末情報テーブル２４から取得したＭＡＣアドレスを送信元ＭＡＣアドレスとし、ＡＲＰ要求パケットの送信元ＭＡＣアドレスを宛先ＭＡＣアドレスとしたＡＲＰ応答パケットを、要求元の端末５０－１に送信する（ステップＳ１３、図８Ｂのステップ１０６）。ここでは、サーバ６０－１のＭＡＣアドレスを送信元ＭＡＣアドレスとし、端末５０－１のＭＡＣアドレスを宛先ＭＡＣアドレスとしたＡＲＰ応答パケットが送信される。

[0065] 図８Ａ及び図８Ｂを参照して、コントローラ２０から送信されたＡＲＰ応答パケットは、制御インタフェース７０を介して、ＡＲＰ要求パケットの通知元のスイッチ１０－１に送信される（ステップＳ１０６）。この際、コントローラ２０は、当該ＡＲＰ応答パケットを端末５０－１に転送するためのフローエントリを、ＡＲＰ応答パケットの送信に先立ってスイッチ１０－１に設定した後、ＡＲＰ応答パケットをスイッチ１０－１に送信する。スイッチ１０－１は、フローテーブルに従い、受信したＡＲＰ応答パケットを端末５０－１に転送する（ステップＳ１０７）。

[0066] 端末５０－１は、受信したＡＲＰ応答パケットの送信元ＭＡＣアドレスを、宛先ＶＩＰアドレスに割り当てられたＭＡＣアドレスとしてＡＲＰテーブルに設定する（ステップＳ１０８）。ここでは、宛先ＶＩＰアドレスのサーバ６０－１のＭＡＣアドレスが割り当てられる。ステップＳ１０８以降、端末５０－１は、ＶＩＰアドレス宛のパケットの宛先ＭＡＣアドレスをサーバ６０－１のＭＡＣアドレスとして、ＶＩＰアドレス宛の通信を行うこととなる（ステップＳ１０９～Ｓ１１５）。詳細には、端末５０－１は、サーバ６０－１のＭＡＣアドレスを宛先としたＶＩＰアドレス宛のパケットを送信する（ステップＳ１０９）。スイッチ１０－１は受信パケットがフローテーブル内のエントリに適合するか否かを判定し、適合しない場合、ファーストパ

ケットとしてコントローラ２０に通知する（ステップＳ１１０Ｙｅｓ、Ｓ１１１）。コントローラ２０は、通知されたファーストパケットを解析し、フロー制御処理を行う（ステップＳ１１２）。詳細には、コントローラ２０は、ファーストパケットの送信元ＭＡＣアドレス及び宛先ＭＡＣアドレスから通信経路の端点となる端末５０－１及びサーバ６０－１を特定し、端点間を結ぶ通信経路を算出する。続いてコントローラ２０は、パケットが通信経路を転送するためのフローエントリを決定する。コントローラ２０は、通信経路上のスイッチ１０にフローエントリの設定を指示する（ステップＳ１１３）。スイッチ１０－１は、フローエントリの設定指示に応じて、自身のフローテーブル１２に、当該フローエントリを設定し、ステップＳ１０９において受信したパケットを当該フローエントリに従った転送先に転送する（ステップＳ１１５）。例えば、“端末５０－１のＭＡＣアドレスを送信元アドレスとし、サーバ６０－１のＭＡＣアドレスを送信元アドレスとしたＶＩＰアドレス宛のパケットを、サーバ６０－１に転送する”ことが規定されたフローエントリが、スイッチ１０－１に設定される。

[0067] ステップＳ１１４以降、端末５０－１のＭＡＣアドレスを送信元アドレスとし、サーバ６０－１のＭＡＣアドレスを送信元アドレスとしたＶＩＰアドレス宛のパケットは、スイッチ１０－１を介してサーバ６０－１に転送されることとなる。尚、図８Ａ及び図８Ｂに示す一例では、スイッチ１０－１からサーバ６０－１にパケットが転送されているが、フローエントリに、サーバ６０－１に至る通信経路上の他のスイッチ１０が転送先として規定されている場合は、他のスイッチ１０を介してサーバ６０－１に転送されることは言うまでもない。

[0068] 又、スイッチ１０に設定されるフローエントリのルールとしては、少なくとも送信元ＭＡＣアドレス及び宛先ＭＡＣアドレスが規定されていればよく、宛先ＩＰアドレスは規定されていなくても良い。これにより、パケットは通常のＬ２スイッチネットワークで転送可能となるため、スイッチ１０のもつＬ２スイッチ機能を利用することで、サーバ６０－１との間で通信が可能

になる。この場合、スイッチ１０はＬ２転送を行うため、パケットはその宛先アドレスのみに基づいて転送される。すなわち、ネットワーク内でパケットヘッダの書き換えが伴われないため、高速かつ安価なスイッチを利用してシステムを構成することが可能である。

[0069] 以上のように、本発明によるコンピュータシステムでは、端末５０が実施したＶＩＰアドレスに対するＡＲＰ要求をコントローラ２０でキャプチャし、ＶＩＰアドレスに割当てたサーバ（のＭＡＣアドレス）を送信元としてＡＲＰ代理応答を行う。又、コントローラ２０によってＶＩＰアドレスに割り当てたＭＡＣアドレス宛に受信パケットを転送するフローエントリが、スイッチ１０に設定される。これにより、スイッチ１０はＭＡＣアドレスを変換することなくＶＩＰアドレスに対応する宛先サーバ（ＭＡＣアドレス）に受信パケットを転送することができる。従って、本発明によれば、高速で且つ安価なＬ２ネットワークにおいて、端末５０のアクセス先をコントローラ２０において任意に設定することができる。又、本発明によるコントローラ２０は、サーバ負荷に応じてアクセス先のサーバを変更するため、高速で且つ安価なＬ２ネットワークにおいて、サーバ負荷分散を実現することができる。

[0070] （２）ＡＲＰ代理応答しない場合の動作

図９に示すステップＳ１４において、ＡＲＰ要求に対する代理応答を行わない場合の動作を、図１０Ａ及び図１０Ｂを参照して説明する。図１０Ａ及び図１０Ｂは、コントローラにおいてＡＲＰ代理応答を行わない場合のコンピュータシステムの動作を示すシーケン図である。

[0071] 図８Ａ及び図８Ｂで示す動作と同様に、端末５０－１から送信されたＶＩＰアドレスをターゲットアドレスとしたＡＲＰ要求パケットがスイッチ１０－１からコントローラ２０に通知される（ステップＳ１０１からＳ１０４）。コントローラ２０は、スイッチ１０－１から通知されたＡＲＰ要求パケットに対して、ＡＲＰ応答処理を実行する（ステップＳ１０５）。

[0072] 図９を参照して、コントローラ２０において、ＡＲＰ要求に対する代理応

答は行わない場合のＡＲＰ応答処理（ステップＳ１０５）を説明する。コントローラ２０のＡＲＰ応答処理部２２は、通知されたＡＲＰ要求パケットのターゲットアドレスを参照し、当該要求がＶＩＰアドレスに対するＡＲＰ要求か否かを判定する（ステップＳ１１）。ここでは、受信パケットがＶＩＰアドレスに対するＡＲＰ要求パケットでないため（例えばＡＲＰ要求パケットはＩＰアドレス宛のパケットであるため）、ＡＲＰ応答処理部２２は、代理応答するか否かを判定する（ステップＳ１４）。ここで、ＡＲＰ要求に対する代理応答をしない場合（ＡＲＰ要求元端末に対して代理応答することが、コントローラ２０に設定されていない場合）、ＡＲＰ応答処理部２２は、ＡＲＰパケットを、通知元のスイッチ１０に返却し、当該ＡＲＰパケットブロードキャストさせるためのＡＲＰ指示を行う（ステップＳ１６、図１０Ａのステップ Ｓ２０１）。

[0073] 図１０Ａ及び図１０Ｂを参照して、ＡＲＰ要求に対する代理応答をしない場合、コントローラ２０は、ＡＲＰ要求パケットをブロードキャストするように当該ＡＲＰ要求パケットの通知元のスイッチ１０に指示する（ステップＳ２０１）。ここでコントローラ２０は、スイッチ１０－１に対し、ＡＲＰ要求パケットをブロードキャストさせるためのフローエントリの設定を指示するとともに、通知されたＡＲＰ要求パケットを返却する。スイッチ１０－１は、フローエントリの設定指示に応じてフローテーブル１２を更新し、当該フローエントリに従い、受信したＡＲＰ応答パケットをブロードキャストする（ステップＳ２０２、Ｓ２０３）。ここでは、ＡＲＰ要求パケットは、サーバ６０－１、６０－２に転送される。尚、図１０Ａ及び図１０Ｂに示す一例では、スイッチ１０－１からサーバ６０－１、６０－２にＡＲＰ要求パケットが転送されているが、フローエントリに、他のスイッチ１０が転送先として設定されている場合は、他のスイッチ１０を介してサーバ６０－１、６０－２に転送されることは言うまでもない。

[0074] ステップＳ２０３以降は、通常の端末によるＡＲＰ手順と同様に宛先ＭＡＣの取得動作が行われる（ステップＳ２０４～Ｓ２０６）。図１０Ａ及び図

１０Ｂに示す一例では、サーバ６０－２からＡＲＰ応答パケットが端末５０－１に送信され、端末５０－１は、ＡＲＰ応答パケットの送信元ＭＡＣアドレスを宛先ＩＰアドレスのＭＡＣアドレスとしてＡＲＰテーブルに設定する（ステップＳ２０４～Ｓ２０６）。

[0075] ステップＳ２０６以降、図８Ａ及び図８Ｂに示す動作と同様に、端末５０－１のＭＡＣアドレスを送信元アドレスとし、サーバ６０－２のＭＡＣアドレスを送信元アドレスとしたＩＰアドレス宛のパケットは、スイッチ１０－１を介してサーバ６０－２に転送されることとなる（ステップＳ１０９～Ｓ１１５）。

[0076] （３）宛先サーバ切替動作

図１１及び図１２を参照して、本発明によるコンピュータシステムにおける宛先サーバ切替動作を説明する。ここでは、端末５０－１のアクセス先のサーバ６０－１を削除し、アクセス先をサーバ６０－２に切り替える場合の動作を一例に説明する。図１１は、本発明によるコンピュータシステムにおける宛先サーバ切替動作の一例を示すシーケン図である。図１２は、本発明によるコントローラ２０におけるサーバ削除処理（宛先サーバ切替処理）の動作の一例を示すフロー図である。

[0077] 図１１を参照して、端末５０－１は、スイッチ１０－１を介してサーバ６０－１に接続している（ステップＳ３０１、Ｓ３０２）。このとき、例えばサーバ６０－１に障害が発生した場合、負荷情報収集装置３０により、サーバ６０－１の削除がコントローラ２０に通知される（図示なし）。コントローラ２０は、サーバ削除の通知に応じてサーバ削除処理（宛先サーバ切替処理）を実行する（ステップＳ３０３）。

[0078] 図１２を参照して、ステップＳ３０３におけるサーバ削除処理の詳細を説明する。負荷分散制御部２１は、負荷情報収集装置３０からサーバの削除情報を取得すると、割当情報テーブル２６を参照して、該当サーバのＭＡＣアドレスを宛先として割当てている端末５０を特定する（ステップＳ２１、２２）。詳細には、負荷分散制御部２１は、削除するサーバのホスト端末ＩＤ

に一致する割当サーバ情報２６３によって割当情報テーブル２６を検索し、対応する宛先情報２６１及び端末情報２６２を特定する（ステップＳ２３）。ここではＶＩＰアドレス及び端末５０－１が特定されるものとする。

[0079] 続いてＡＲＰ応答処理部２２は、負荷分散制御部２１から、特定された宛先情報２６１及び端末情報２６２に割当てる実サーバ情報（ＭＡＣアドレス）を取得する。詳細には、負荷分散制御部２１は、負荷情報テーブル２５を参照して、ステップＳ２２において特定された宛先情報２６１（例えば宛先ＶＩＰアドレス）が割り当てられたサーバ（サーバ情報２５１）を特定し、特定したサーバ情報２５１に対応するＭＡＣアドレス２４３をホスト端末情報テーブル２４から抽出し、ＡＲＰ応答処理部２２に渡す。この際、ＶＩＰアドレスが割り当てられたサーバが複数ある場合、負荷分散制御部２１は、複数のサーバ（サーバ情報２５１）のそれぞれに対応する負荷情報２５３に基づいて、新たなアクセス先として割り当てるサーバを決定する。例えば、負荷分散制御部２１は、ＶＩＰアドレスが割り当てられた複数のサーバのうち、削除されたサーバ６０－１を除く最も小さい負荷のサーバ６０－２を、端末５０－１のアクセス先として割り当て、当該サーバのＭＡＣアドレスをＡＲＰ応答処理部２２に渡す。負荷分散制御部２１は、ＶＩＰアドレスに対応する新たな宛先としてサーバ６０－２を割り当てると、この割り当てた結果を割当情報テーブル２６に記録する。このとき、負荷分散制御部２１は、サーバの割当変更前の割当サーバ情報２６３を旧宛先サーバ情報２６４として記録し、新たにＶＩＰアドレスに割当てるサーバのホスト端末ＩＤを割当サーバ情報２６３として上書きする。例えば、負荷分散制御部２１は、ステップＳ２２で特定されたＶＩＰアドレスを宛先情報２６１、ステップＳ２２において特定された端末５０－１に対応するホスト端末ＩＤを端末情報２６２、当該ＶＩＰアドレスに割り当てたサーバ６０－２のホスト端末ＩＤを割当サーバ情報２６３とするとともに、それぞれを対応付けて割当情報テーブル２６に記録する。

[0080] ＡＲＰ応答処理部２２は、負荷分散制御部２１において新たに割当てられ

たサーバのＭＡＣアドレスを送信元ＭＡＣアドレスとするＧＡＲＰ（Ｇｒａｔｕｉｏｕｓ　ＡＲＰ）パケットを、ステップＳ２２で特定された端末５０に発光する（ステップＳ２４、図１１のステップＳ３０４）。ここで、ＧＡＲＰパケットには、端末５０－１とその宛先ＶＩＰアドレスに対して新たに割当てられたサーバ６０－２のＭＡＣアドレスが送信元ＭＡＣアドレスとして設定されている。図１１を参照して、ＧＡＲＰパケットはスイッチ１０－１を介して端末５０－１に送信される（ステップＳ３０４、Ｓ３０５）。

[0081] 端末５０－１は、受信したＧＡＲＰパケットの送信元ＭＡＣアドレスを、宛先ＶＩＰアドレスに割り当てられたＭＡＣアドレスとしてＡＲＰテーブルを更新する（ステップＳ３０６）。ここでは、宛先ＶＩＰアドレスにサーバ６０－２のＭＡＣアドレスが割り当てられる。ステップＳ３０６以降、端末５０－１は、図１０Ａ及び図１０Ｂのステップ１０９～Ｓ１１５と同様に、ＶＩＰアドレス宛のパケットの宛先ＭＡＣアドレスをサーバ６０－２のＭＡＣアドレスとして、ＶＩＰアドレス宛の通信を行うこととなる（ステップＳ３０７～Ｓ３１３）。

[0082] 以上のように、本発明によれば、宛先サーバが削除されても、コントローラ２０による一元制御により、負荷を考慮した宛先サーバの切替が可能となる。

[0083] （４）セッションを維持しながら宛先サーバを切替える動作

図１３から図１５を参照して、本発明によるコンピュータシステムのセッションを維持しながら宛先サーバを切替える動作について説明する。ここでは、端末５０－１とサーバ６０－１との間の通信を、端末５０－１とサーバ６０－２との間の通信に切り替える場合について説明する。図１３は、本発明によるコンピュータ装置において、セッションを維持しながら宛先サーバを切替える動作の一例を示すシーケンス図である。図１４は、本発明によるコントローラ２０における切替処理の動作の一例を示すフロー図である。図１５は、本発明による切替装置４０におけるパケット制御処理の動作の一例を示すフロー図である。

[0084] 図１３を参照して、端末５０－１は、スイッチ１０－１を介してサーバ６０－１に接続している（ステップＳ４０１、Ｓ４０２）。このとき、コントローラ２０は、例えばサーバ６０－１への負荷が増大すると、端末５０－１とサーバ６０－１との間の通信を、負荷の小さな他のサーバ６０との通信に切り替えるための切替処理を実行する（ステップＳ４０３）。

[0085] 図１４を参照して、ステップＳ４０３における切替処理の詳細を説明する。負荷分散制御部２１は、負荷情報テーブル２５及び割当情報テーブル２６を参照して、端末の宛先ＶＩＰアドレスに割当てられたサーバのうち、負荷が増大したサーバを、負荷の小さな他のサーバに切り替える（ステップＳ３１）。詳細には、負荷分散制御部２１は、端末の宛先ＶＩＰアドレスに割当てられたサーバの負荷が、所定の閾値を超えた場合、あるいは、当該宛先ＶＩＰアドレスをもつ他のサーバの負荷が所定の閾値よりも低下した場合、現在ＶＩＰアドレスに割当てている宛先サーバを、負荷の小さな他のサーバに切り替える。負荷分散制御部２１は、サーバの割当変更前の割当サーバ情報２６３を旧宛先サーバ情報２６４として記録し、新たに宛先サーバとして割当てるサーバのホスト端末ＩＤを割当サーバ情報２６３として上書きする。例えば、負荷分散制御部２１は、負荷が閾値を越えたサーバ６０－１のホスト端末ＩＤを旧宛先サーバ情報２６４とするとともに、負荷が小さいサーバ６０－２のホスト端末ＩＤを割当サーバ情報２６３として宛先情報２６１（ＶＩＰアドレス）及び端末情報２６２（端末５０－１）に対応付けて記録する。尚、負荷分散制御部２１による宛先サーバの切替動作は、定期的に行われても良いし、任意の時期（例えば、負荷情報テーブル２５の更新時（負荷情報収集装置３０からの負荷情報の通知時））に行われても良い。

[0086] 負荷分散制御部２１は、宛先サーバを切替えると、切替前後の宛先サーバの情報（ＭＡＣアドレス）を、切替情報として切替装置４０に通知する（ステップＳ３２、図１３のＳ４０４）。詳細には、ステップＳ３１において更新された割当サーバ情報２６３及び旧宛先サーバ情報２６４に対応するＭＡＣアドレス２４３と、これらに対応付けられた端末情報２６２に対応するＭ

ＡＣアドレス及び宛先情報２６１（ＶＩＰアドレス）が、切替情報として切替装置４０に送信される。ここでは、端末５０－１からＶＩＰアドレス宛のトラヒックに割当てられた、切替前のサーバ６０－１のＭＡＣアドレスと切替後のサーバ６０－２のＭＡＣアドレスとの対応情報が、切替情報として切替装置４０に通知される。

[0087] 図１３を参照して、切替装置４０の切替制御部４４は通知された切替情報を、切替情報テーブル４５に記録する（ステップＳ４０５）。ここでは、切替情報に含まれる宛先情報２６１（ＶＩＰアドレス）、端末情報２６２に対応するＭＡＣアドレス、割当サーバ情報２６３に対応するＭＡＣアドレス及び旧宛先サーバ情報２６４に対応するＭＡＣアドレス２４３がそれぞれ、宛先情報４５１、端末情報４５２、新宛先サーバ情報４５３及び旧宛先サーバ情報４５４として切替情報テーブル４５に記録される。

[0088] 切替装置４０に切替情報を通知すると、コントローラ２０は、切替前のサーバ６０宛の通信経路を切替装置４０に振り向けるようにフローエントリの設定（変更）を、当該通信経路上のスイッチ１０に対して指示する（リダイレクト設定指示、ステップＳ４０６、図１４のＳ３３）。詳細には、コントローラ２０のスイッチ制御部２３は、リダイレクト設定の対象となる端末５０と切替前のサーバ６０との間の通信経路上において端末５０に隣接するスイッチ１０のフローエントリを変更する。ここでは、端末５０－１に隣接するスイッチ１０－１のフローエントリの変更を指示する。

[0089] スイッチ１０－１は、フローエントリの設定指示に応じて、自身のフローテーブル１２に、当該フローエントリを設定する（ステップＳ４０７）。ここでは、“端末５０－１のＭＡＣアドレスを送信元ＭＡＣとし、サーバ６０－１のＭＡＣアドレスを宛先ＭＡＣアドレスとしたパケットを、切替装置４０に転送する”と規定されたフローエントリが設定される。

[0090] ステップＳ４０７以降、端末５０－１のＭＡＣアドレスを送信元アドレスとし、サーバ６０－１のＭＡＣアドレスを送信元アドレスとしたＶＩＰアドレス宛のパケットは、スイッチ１０－１を介してサーバ６０－１に転送され

ることとなる。（ステップＳ４０８、Ｓ４０９）。詳細には、端末５０－１は、サーバ６０－１のＭＡＣアドレスを宛先としたＶＩＰアドレス宛のパケットを送信する（ステップＳ４０８）。スイッチ１０－１は受信パケットをフローエントリに従った切替装置４０に転送する（ステップＳ４０９）。これにより、端末５０－１から宛先ＭＡＣが６０－１宛のトラヒックのリダイレクトが実現される。

[0091] 切替装置４０は、パケットを受信すると、パケット制御処理により受信パケットを解析し、解析結果に応じてフローエントリの変更やフローエントリに従った処理を行う（ステップＳ４１０）。

[0092] 図１５を参照して、ステップＳ４１３におけるパケット制御処理の一例を説明する。切替装置４０では、ＴＣＰのセッションを意識して、新規ＴＣＰコネクションから順次、その宛先サーバを切替先のサーバに切替える。宛先サーバの切替方法については、切替装置４０内でパケット内の宛先ＭＡＣアドレスを切替先のサーバのＭＡＣアドレスに変換することで実現される。

[0093] 先ず、切替装置４０は、受信パケットがＴＣＰパケットか否かを判定する（ステップＳ４１）。受信パケットがＴＣＰパケットではない場合、コネクションの概念がないため、スイッチ処理部４１は、受信パケットの宛先を新宛先サーバ情報４５３に対応するサーバ６０とし、“受信パケットの宛先ＭＡＣアドレスを当該サーバ６０のＭＡＣアドレスに変換して転送する”ことを規定したフローエントリを、フローテーブル４２に登録する（ステップＳ４１Ｎｏ、Ｓ４３）。詳細には、ステップＳ４３においてスイッチ制御部４３は、切替情報テーブル４５から、受信パケットの送信元ＭＡＣアドレス及び宛先ＶＩＰアドレスに対応する新宛先サーバ情報４５３を特定する。続いてスイッチ制御部４３は、受信パケットの送信元ＭＡＣアドレス及び宛先ＶＩＰアドレスをルールとし、パケットの宛先ＭＡＣアドレスを、特定した新宛先サーバ情報４５３（ＭＡＣアドレス）に変換して転送することをアクションとしたフローエントリを、フローテーブル４２に登録する。

[0094] 一方、ステップＳ４１において、受信パケットがＴＣＰパケットである場

合、切替装置４０は受信パケットのＳＹＮフラグビットを確認する（ステップＳ４１Ｙｅｓ、Ｓ４２）。ここで、受信パケットがＳＹＮパケットである場合、当該パケットはＴＣＰの新規コネクションを示すため、ステップＳ４３に移行し、このコネクションについて新規サーバ宛（新宛先サーバ情報４５３）に宛先が変更される（ステップＳ４２Ｙｅｓ）。

[0095] ステップＳ４３においてフローテーブル４２が変更されると、スイッチ処理部４１は、変更されたフローテーブルに規定されたアクションを実行する（ステップＳ５０）。ここでスイッチ処理部４１は、受信パケットの宛先ＭＡＣアドレスを、コントローラ２０によって新たに割り当てたサーバのＭＡＣアドレス（新宛先サーバ情報４５３）に変換して転送する。

[0096] 一方、受信パケットがＳＹＮパケットではなく、ＴＣＰ終了判定条件を満たしていない場合、スイッチ処理部４１は、受信パケットに適合するフローエントリに規定されたアクションに従って、受信パケットを転送する（ステップＳ４２Ｎｏ、Ｓ４４Ｙｅｓ、Ｓ４５、Ｓ４８Ｎｏ、Ｓ５０）。ここで、ＴＣＰ終了判定条件を満たすとは、例えばＦＩＮパケット及びＦＩＮ－ＡＣＫパケットが双方向に送信される場合を示す。

[0097] しかし、ＳＹＮパケットではない受信パケットに適合するフローエントリが、フローテーブル４２に設定されていない場合、スイッチ処理部４１は、受信パケットの宛先を旧宛先サーバ情報４５４に対応するサーバとし、“受信パケットの宛先ＭＡＣアドレスを当該サーバのＭＡＣアドレスに変換して転送する”ことを規定したフローエントリをフローテーブル４２に登録する（ステップＳ４２Ｎｏ、Ｓ４４Ｎｏ、Ｓ４６、Ｓ４７）。詳細には、スイッチ制御部４３は、切替情報テーブル４５から、受信パケットの送信元ＭＡＣアドレス及び宛先ＶＩＰアドレスに対応する旧宛先サーバ情報４５４を特定する。続いてスイッチ制御部４３は、受信パケットの送信元ＭＡＣアドレス及び宛先ＶＩＰアドレスをルールとし、パケットの宛先ＭＡＣアドレスを、特定した旧宛先サーバ情報４５４（ＭＡＣアドレス）に変換して転送することをアクションとしたフローエントリを、フローテーブル４２に登録する。

フローエントリの登録後、ＴＣＰ終了判定条件を満たしていない場合、スイッチ処理部４１は、ステップＳ４７において登録したフローエントリに規定されたアクションに従って、受信パケットを転送する（ステップＳ５０）。

[0098] 又、受信パケットがＴＣＰパケットであって、ＦＩＮパケット及びＦＩＮ－ＡＣＫパケットが双方向に送信されるような、終了判定条件を満たしている場合、コネクションがクローズされたとして明示的にフローテーブルのエントリを削除する（ここでは、明示的に削除宣言されるだけで、実際の削除はアクション実行後に行われる）（ステップＳ４８Ｙｅｓ、Ｓ４９）。

[0099] 以上のように、切替装置４０は、受信パケットの宛先サーバが旧サーバか切替後の新サーバかを判断し、判断結果に応じたフローエントリに従った宛先に受信パケットを転送する。本発明では、切替装置４０によって端末５０から転送されたパケットの宛先ＭＡＣアドレスを変更することで、セッションを維持しつつ宛先サーバを他のサーバに切り替えることができる。図１４に示す一例では、切替装置４０は、受信パケットに応じて設定されたフローエントリ、又は受信パケットに適合するフローエントリに従って、端末５０－１からサーバ６０－１のＭＡＣアドレス宛のパケットを、サーバ６０－２のＭＡＣアドレス宛に変換して転送する（ステップＳ４１０、Ｓ４１１）。本発明では、スイッチ１０にＭＡＣアドレスの変換機能（セッションレイヤの処理）を持たせずに、サービス中のコネクションを維持しながら端末５０のアクセス先サーバを変更できるため、スイッチ１０としてＬ２スイッチを利用することができる。従って、本発明によれば、コスト上昇を抑制しながら高速化ネットワーク環境下における負荷分散を実現することが可能となる。又、代理ＡＲＰ機能はスイッチ１０と独立したコントローラ２０によって行われるため、スイッチネットワークの規模によらずに負荷分散が可能となる。このため、高価なネットワークデバイスを用いなくてもスケーラビリティの高いサーバ負荷分散処理が実現される。

[0100] 更に、ステップＳ４０３においてコントローラ２０は、端末－サーバ間通信のプロトコルに応じて、宛先サーバの割当変更時において切替装置を利用

するか否かを決定してもよい。例えば。端末－サーバ間の通信がＴＣＰ通信である場合、サーバ割当変更時に、切替装置４０を通過させることで、セッションを切らずに徐々にサーバ割当の変更を行う。一方、ＨＴＴＰ通信の場合は、切替装置４０を通過させずサーバ割当を変更し、セッションの中断を許す。あるいは、コントローラ２０は、ＨＴＴＰプロトコルでも特定の重要なアプリケーションの場合は切替装置を通過させて、サービスを途切れないようにする、といった制御をおこなってもよい。更に、コントローラ２０は、送信元ＩＰアドレスに応じて、切替装置に分岐させるかどうかを決めることで、プロトコル単位のみならずユーザ単位でも、サーバ割当の変更時にセッションを維持するか否かを設定することが可能となる。

[0101] （５）フローエントリの削除に応じた切替終了動作

図１６は、本発明による切替装置４０におけるフローエントリのエキスパイアに応じた切替動作を示すフロー図である。図１６を参照して、本発明による切替装置４０におけるフローエントリのエキスパイアに応じた切替動作を説明する。ここでは、切替装置４０のフローテーブル４２に登録されるフローエントリは、所定の時間の経過後や所定の条件を満たした場合、フローテーブル４２から削除される。この場合の切替装置４０の動作を説明する。

[0102] ＴＣＰコネクションの終了条件を満たす場合や、タイムアウトによってフローエントリがエキスパイアされると、切替制御部４４は、フローテーブル４２に記録されたフローエントリのうち、切替前のサーバのＭＡＣアドレス（旧宛先サーバ情報４５４）を宛先ＭＡＣアドレスとするフローエントリの有無を確認する（ステップＳ６０、Ｓ６１、Ｓ６２）。ここで、フローテーブル４２に旧宛先サーバ情報４５４を宛先ＭＡＣアドレスとするフローエントリがある場合、切替制御部４４は、新たなフローエントリのエキスパイアが発生するまで、待機する（ステップＳ６２Ｎｏ）。一方、旧宛先サーバ情報４５４を宛先ＭＡＣアドレスとするフローエントリが無くなった場合、切替制御部４４は、切替処理の終了をコントローラ２０に通知する（ステップＳ６２Ｙｅｓ、Ｓ６３）。例えば、切替制御部４４は、フローエントリを削

除する都度、旧宛先サーバ情報４５４を宛先ＭＡＣアドレスとするフローエントリをカウントし、当該カウント値が０になった時点、あるいは切替処理から一定時間が経過した時点で、切替終了として、コントローラに通知する。

[0103] コントローラ２０は切替終了通知をうけると端末５０に対して、図１１又は図１２と同様にＧＡＲＰパケットによってＶＩＰアドレスに対する物理サーバ（ＭＡＣアドレス）を通知する（図示なし）。例えば、コントローラ２０は切替終了通知に応じて端末５０－１に対して、ＧＡＲＰパケットによってＶＩＰアドレスに対する物理サーバが、サーバ６０－２（のＭＡＣアドレス）となったことを通知する。端末５０－１はＶＩＰアドレスに対するＡＲＰキャッシュエントリをサーバ６０－２のＭＡＣとして、以降のＶＩＰアドレス宛通信をサーバ６０－２のＭＡＣアドレス宛に出力する。更に、コントローラ２０は端末５０－１から切替装置４０に向けてリダイレクトしていた経路用のフローエントリを削除するとともに、切替装置のリソースを開放する（図示なし）。

[0104] 以上のように、本発明によるコンピュータシステムによれば、スイッチにおいてヘッダ変換を行うことなくアクセス先のサーバを変更することができるため、低コストで高速なＬ２スイッチによって負荷分散システムを構築することが可能となる。又、本発明では、オープンフローネットワークを利用して、サーバ負荷やサーバの削除に応じてフローの設定やアクセス先サーバの変更を実現しているため、スケーラビリティの高いサーバ負荷分散処理が可能となる。

[0105] 以上、本発明の実施の形態を詳述してきたが、具体的な構成は上記実施の形態に限られるものではなく、本発明の要旨を逸脱しない範囲の変更があっても本発明に含まれる。例えば、スイッチ１０、端末５０、切替装置４０、サーバ６０の数は上述の例に限らないことは言うまでもない。又、本発明によるコンピュータシステムが複数設けられ、それぞれのコントローラ２０の連携により、複数のシステム間で経路制御（負荷分散制御）されても構わな

い。更に、上述の例では、サーバ負荷やアプリケーション負荷に応じて宛先サーバの変更が行われたが、これに限らず、ネットワーク負荷（輻輳状況）やベンダによる設定によって任意に宛先サーバの変更が行われても良い。例えば、サーバの保守時に、宛先サーバの変更が行われる。更に、切替装置４０は、コントローラ２０とは別の装置で実現されても、同じ装置で実現されてもどちらでも良い。切替装置４０がコントローラ２０と同じ装置で実現される場合、スイッチ１０の機能（パケット中継機能）も適用されることとなる。すなわち、切替装置４０はコントローラ２０とスイッチ１０とを組み合わせることでも実現できる。

[0106] 上記の実施の形態の一部又は全部は、以下の付記のようにも記載し得るが、以下には限られない。

[0107] （付記１）

コントローラと、

前記コントローラによって設定されたフローエントリに適合する受信パケットに対し、前記フローエントリで規定された中継動作を行うスイッチと

を具備し、

前記コントローラは、端末からのＡＲＰ要求パケットに対して代理応答することにより、前記端末のアクセス先となる物理サーバのＭＡＣ（Ｍｅｄｉａ Ａｃｃｅｓｓ Ｃｏｎｔｒｏｌ）アドレスを前記端末に通知し、

前記端末から転送され、前記ＭＡＣアドレスを宛先ＭＡＣアドレスとした受信パケットを前記物理サーバに至る通信経路に転送することを規定したフローエントリを、前記スイッチに設定する

コンピュータシステム。

[0108] （付記２）

付記１に記載のコンピュータシステムにおいて、

前記コントローラは、ＶＩＰ（Ｖｉｒｔｕａｌ Ｉｎｔｅｒｎｅｔ Ｐｒｏｔｃｏｌ）アドレスと物理サーバとを対応付けて記録されたテーブルを参照して、前記ＡＲＰ要求パケットに含まれるターゲットＶＩＰアドレスに対

応する物理サーバを、前記端末のアクセス先となる物理サーバとして選択し、前記物理サーバのＭＡＣアドレスをＡＲＰ応答パケットの送信元ＭＡＣアドレスとして、前記端末に送信する

コンピュータシステム。

[0109] （付記３）

付記１又は２に記載のコンピュータシステムにおいて、

前記コントローラは、複数のサーバのそれぞれに対する負荷に応じて、前記端末のアクセス先となる物理サーバのＭＡＣアドレスを選択する

コンピュータシステム。

[0110] （付記４）

付記１から３のいずれか１項に記載のコンピュータシステムにおいて、

前記コントローラは、前記物理サーバと異なる他の物理サーバのＭＡＣアドレスを送信元ＭＡＣアドレスとしたＧＡＲＰパケットを前記端末に送信することで、前記端末のＡＲＰテーブルに設定された宛先ＭＡＣアドレスを、前記物理サーバのＭＡＣアドレスから前記他の物理サーバのＭＡＣアドレスに変更する

コンピュータシステム。

[0111] （付記５）

付記１から４のいずれか１項に記載のコンピュータシステムにおいて、

前記スイッチは、前記コントローラによって設定されたフローエントリに従い、受信パケットがＡＲＰ要求パケットである場合、前記ＡＲＰ要求パケットを前記コントローラに通知する

コンピュータシステム。

[0112] （付記６）

コントローラと、

前記コントローラによって設定されたフローエントリに適合する受信パケットに対し、前記フローエントリで規定された中継動作を行うスイッチと、

切替装置と

を具備し、

前記スイッチは、前記コントローラによって設定された第１フローエントリに従い、第１サーバの第１ＭＡＣアドレスを宛先ＭＡＣアドレスとしたパケットを、前記第１サーバに至る通信経路へ転送し、

前記コントローラは、前記第１サーバの第１ＭＡＣアドレスを宛先ＭＡＣアドレスとしたパケットを前記切替装置に転送することを規定した第２フローエントリを、前記第１フローエントリに替えて前記スイッチに設定し、

前記スイッチは、前記第２フローエントリに従い、前記第１ＭＡＣアドレスを宛先ＭＡＣアドレスとしたパケットを、前記切替装置へ転送し、

前記切替装置は、前記スイッチからの受信パケットの宛先ＭＡＣアドレスを、前記第１ＭＡＣアドレスから、第１サーバとは異なる第２サーバの第２ＭＡＣアドレスに変更して、前記第２サーバに至る通信経路へ転送する

コンピュータシステム。

[0113] （付記７）

付記６に記載のコンピュータシステムにおいて、

前記切替装置は、前記受信パケットが、ＴＣＰ（Ｔｒａｎｓｍｉｓｓｉｏｎ　Ｃｏｎｔｒｏｌ　Ｐｒｏｔｏｃｏｌ）パケットであり、適合するフローエントリが自身のフローテーブルに設定されていない場合、前記スイッチからの受信パケットの宛先ＭＡＣアドレスを、前記第１ＭＡＣアドレスに変更して、前記第１サーバに至る通信経路へ転送する

コンピュータシステム。

[0114] （付記８）

付記６又は７に記載のコンピュータシステムにおいて、

前記切替装置は、受信パケットに基づきＴＣＰ終了判定を行い、ＴＣＰコネクションが終了したと判定すると、自身のフローテーブルに設定され、前記受信パケットに適合するフローエントリに従った宛先に前記受信パケットの転送した後、前記フローエントリを前記フローテーブルから削除する

コンピュータシステム。

[0115] （付記９）

付記６から８のいずれか１項に記載のコンピュータシステムにおいて、

前記切替装置は、自身のフローテーブルからフローエントリが削除されると、切替終了を前記コントローラに通知し、

前記コントローラは、前記切替終了に応じて前記第２フローエントリを前記スイッチから削除する

コンピュータシステム。

[0116] （付記１０）

付記６から９のいずれか１項に記載のコンピュータシステムにおいて、

前記コントローラは、複数のサーバのそれぞれに対する負荷に応じて、前記端末のアクセス先となる前記第２サーバの第２ＭＡＣアドレスを選択して、前記切替装置に通知する

コンピュータシステム。

[0117] （付記１１）

付記１から１０のいずれか１項に記載のコンピュータシステムで利用されるコントローラ。

[0118] （付記１２）

付記６から１０のいずれか１項に記載のコンピュータシステムで利用される切替装置。

[0119] （付記１３）

コンピュータによって実行されることで、付記１１項に記載のコントローラを実現する通信制御プログラム。

[0120] （付記１４）

コンピュータによって実行されることで、付記１２に記載の切替装置を実現する通信制御プログラム。

[0121] （付記１５）

コントローラによって設定されたフローエントリに適合する受信パケットに対し、前記フローエントリで規定された中継動作を行うスイッチを備える

コンピュータシステムにおいて実行される通信方法において、

前記コントローラが、端末からのＡＲＰ要求パケットに対して代理応答することにより、前記端末のアクセス先となる物理サーバのＭＡＣ（Ｍｅｄｉａ　Ａｃｃｅｓｓ　Ｃｏｎｔｒｏｌ）アドレスを前記端末に通知するステップと、

前記端末から転送され、前記ＭＡＣアドレスを宛先ＭＡＣアドレスとした受信パケットを前記物理サーバに至る通信経路に転送することを規定したフローエントリを、前記コントローラが前記スイッチに設定するステップと

を具備する

通信方法。

[0122] （付記１６）

付記１５に記載の通信方法において、

前記コントローラが、ＶＩＰ（Ｖｉｒｔｕａｌ　Ｉｎｔｅｒｎｅｔ　Ｐｒｏｔｃｏｌ）アドレスと物理サーバとを対応付けて記録されたテーブルを参照して、前記ＡＲＰ要求パケットに含まれるターゲットＶＩＰアドレスに対応する物理サーバを、前記端末のアクセス先となる物理サーバとして選択するステップを更に具備し、

前記ＭＡＣアドレスを通知するステップは、前記コントローラが、前記物理サーバのＭＡＣアドレスをＡＲＰ応答パケットの送信元ＭＡＣアドレスとして、前記端末に送信するステップを備える

通信方法。

[0123] （付記１７）

付記１５又は１６に記載の通信方法において、

前記コントローラが、複数のサーバのそれぞれに対する負荷に応じて、前記端末のアクセス先となる物理サーバのＭＡＣアドレスを選択するステップを更に具備する

通信方法。

[0124] （付記１８）

付記１５から１７のいずれか１項に記載の通信方法において、

前記コントローラが、前記物理サーバと異なる他の物理サーバのＭＡＣアドレスを送信元ＭＡＣアドレスとしたＧＡＲＰパケットを前記端末に送信することで、前記端末のＡＲＰテーブルに設定された宛先ＭＡＣアドレスを、前記物理サーバのＭＡＣアドレスから前記他の物理サーバのＭＡＣアドレスに変更するステップを更に具備する

通信方法。

[0125] （付記１９）

付記１５から１８のいずれか１項に記載の通信方法において、

前記スイッチが、前記コントローラによって設定されたフローエントリに従い、受信パケットがＡＲＰ要求パケットである場合、前記ＡＲＰ要求パケットを前記コントローラに通知するステップを更に具備する

通信方法。

[0126] （付記２０）

コントローラによって設定されたフローエントリに適合する受信パケットに対し、前記フローエントリで規定された中継動作を行うスイッチを備えるコンピュータシステムにおいて実行される通信方法において、

前記スイッチが、前記コントローラによって設定された第１フローエントリに従い、第１サーバの第１ＭＡＣアドレスを宛先ＭＡＣアドレスとしたパケットを、前記第１サーバに至る通信経路へ転送するステップと、

前記コントローラが、前記第１サーバの第１ＭＡＣアドレスを宛先ＭＡＣアドレスとしたパケットを切替装置に転送することを規定した第２フローエントリを、前記第１フローエントリに替えて前記スイッチに設定し、

前記スイッチが、前記第２フローエントリに従い、前記第１ＭＡＣアドレスを宛先ＭＡＣアドレスとしたパケットを、前記切替装置へ転送するステップと、

前記切替装置が、前記スイッチからの受信パケットの宛先ＭＡＣアドレスを、前記第１ＭＡＣアドレスから、第１サーバとは異なる第２サーバの第２

ＭＡＣアドレスに変更して、前記第２サーバに至る通信経路へ転送するステップと

を具備する

通信方法。

[0127] （付記２１）

付記２０に記載の通信方法において、

前記受信パケットが、ＴＣＰ（Ｔｒａｎｓｍｉｓｓｉｏｎ Ｃｏｎｔｒｏｌ Ｐｒｏｔｏｃｏｌ）パケットであり、適合するフローエントリが自身のフローテーブルに設定されていない場合、前記切替装置が、前記スイッチからの受信パケットの宛先ＭＡＣアドレスを、前記第１ＭＡＣアドレスに変更して、前記第１サーバに至る通信経路へ転送するステップを更に具備する

通信方法。

[0128] （付記２２）

付記２０又は２１に記載の通信方法において、

前記切替装置が、受信パケットに基づきＴＣＰコネクションが終了したと判定するステップと、

前記切替装置が、前記受信パケットに適合するフローエントリに従った宛先に前記受信パケットの転送した後、前記フローエントリを前記フローテーブルから削除するステップと

を更に具備する

通信方法。

[0129] （付記２３）

付記２０から２２のいずれか１項に記載の通信方法において、

前記切替装置が、自身のフローテーブルからフローエントリが削除されると、切替終了を前記コントローラに通知するステップと、

前記コントローラが、前記切替終了に応じて前記第２フローエントリを前記スイッチから削除するステップと

を更に具備する

通信方法。

[0130] （付記２４）

付記２０から２３のいずれか１項に記載の通信方法において、

前記コントローラが、複数のサーバのそれぞれに対する負荷に応じて、前記端末のアクセス先となる前記第２サーバの第２ＭＡＣアドレスを選択して、前記切替装置に通知するステップを更に具備する

通信方法。

[0131] 尚、本出願は、日本出願番号２０１１－８０９５５に基づき、日本出願番号２０１１－８０９５５における開示内容は引用により本出願に組み込まれる。

請求の範囲

［請求項1］ コントローラと、

前記コントローラによって設定されたフローエントリに適合する受信パケットに対し、前記フローエントリで規定された中継動作を行うスイッチと

を具備し、

前記コントローラは、ＶＩＰ（Ｖｉｒｔｕａｌ　Ｉｎｔｅｒｎｅｔ　Ｐｒｏｔｃｏｌ）アドレスと物理サーバとを対応付けて記録されたテーブルを参照して、端末からのＡＲＰ要求パケットに含まれるターゲットＶＩＰアドレスに対応する物理サーバを、前記端末のアクセス先となる物理サーバとして選択し、前記物理サーバのＭＡＣ（Ｍｅｄｉａ　Ａｃｃｅｓｓ　Ｃｏｎｔｒｏｌ）アドレスを送信元ＭＡＣアドレスとしたＡＲＰ応答パケットによって、前記ＡＲＰ要求パケットに対して代理応答することにより、前記端末のアクセス先となる物理サーバのＭＡＣアドレスを前記端末に通知し、

前記端末から転送され、前記ＭＡＣアドレスを宛先ＭＡＣアドレスとした受信パケットを前記物理サーバに至る通信経路に転送することを規定したフローエントリを、前記スイッチに設定する

コンピュータシステム。

［請求項2］ 請求項１に記載のコンピュータシステムにおいて、

前記コントローラは、複数のサーバのそれぞれに対する負荷に応じて、前記端末のアクセス先となる物理サーバのＭＡＣアドレスを選択する

コンピュータシステム。

［請求項3］ 請求項１又は２に記載のコンピュータシステムにおいて、

前記コントローラは、前記物理サーバと異なる他の物理サーバのＭＡＣアドレスを送信元ＭＡＣアドレスとしたＧＡＲＰパケットを前記端末に送信することで、前記端末のＡＲＰテーブルに設定された宛先

ＭＡＣアドレスを、前記物理サーバのＭＡＣアドレスから前記他の物理サーバのＭＡＣアドレスに変更する

コンピュータシステム。

[請求項4] コントローラと、

前記コントローラによって設定されたフローエントリに適合する受信パケットに対し、前記フローエントリで規定された中継動作を行うスイッチと、

切替装置と

を具備し、

前記スイッチは、前記コントローラによって設定された第１フローエントリに従い、第１サーバの第１ＭＡＣアドレスを宛先ＭＡＣアドレスとしたパケットを、前記第１サーバに至る通信経路へ転送し、

前記コントローラは、前記第１サーバの第１ＭＡＣアドレスを宛先ＭＡＣアドレスとしたパケットを前記切替装置に転送することを規定した第２フローエントリを、前記第１フローエントリに替えて前記スイッチに設定し、

前記スイッチは、前記第２フローエントリに従い、前記第１ＭＡＣアドレスを宛先ＭＡＣアドレスとしたパケットを、前記切替装置へ転送し、

前記切替装置は、前記スイッチからの受信パケットの宛先ＭＡＣアドレスを、前記第１ＭＡＣアドレスから、第１サーバとは異なる第２サーバの第２ＭＡＣアドレスに変更して、前記第２サーバに至る通信経路へ転送する

コンピュータシステム。

[請求項5] 請求項１から４のいずれか１項に記載のコンピュータシステムで利用されるコントローラ。

[請求項6] 請求項３に記載のコンピュータシステムで利用される切替装置。

[請求項7] コンピュータによって実行されることで、請求項５項に記載のコン

トローラを実現する通信制御プログラムが格納された記録媒体。

[請求項8] コンピュータによって実行されることで、請求項６に記載の切替装置を実現する通信制御プログラムが格納された記録媒体。

[請求項9] コントローラによって設定されたフローエントリに適合する受信パケットに対し、前記フローエントリで規定された中継動作を行うスイッチを備えるコンピュータシステムにおいて実行される通信方法において、

前記コントローラが、ＶＩＰ（Ｖｉｒｔｕａｌ　Ｉｎｔｅｒｎｅｔ　Ｐｒｏｔｃｏｌ）アドレスと物理サーバとを対応付けて記録されたテーブルを参照して、端末からのＡＲＰ要求パケットに含まれるターゲットＶＩＰアドレスに対応する物理サーバを、前記端末のアクセス先となる物理サーバとして選択するステップと、

前記コントローラが、前記物理サーバのＭＡＣ（Ｍｅｄｉａ　Ａｃｃｅｓｓ　Ｃｏｎｔｒｏｌ）アドレスを、送信元ＭＡＣアドレスとしたＡＲＰ応答パケットによって、前記ＡＲＰ要求パケットに対して代理応答することにより、前記端末のアクセス先となる物理サーバのＭＡＣアドレスを前記端末に通知するステップと、

前記端末から転送され、前記ＭＡＣアドレスを宛先ＭＡＣアドレスとした受信パケットを前記物理サーバに至る通信経路に転送することを規定したフローエントリを、前記コントローラが前記スイッチに設定するステップと

を具備する

通信方法。

[請求項10] コントローラによって設定されたフローエントリに適合する受信パケットに対し、前記フローエントリで規定された中継動作を行うスイッチを備えるコンピュータシステムにおいて実行される通信方法において、

前記スイッチが、前記コントローラによって設定された第１フロー

エントリに従い、第１サーバの第１ＭＡＣアドレスを宛先ＭＡＣアドレスとしたパケットを、前記第１サーバに至る通信経路へ転送するステップと、

前記コントローラが、前記第１サーバの第１ＭＡＣアドレスを宛先ＭＡＣアドレスとしたパケットを切替装置に転送することを規定した第２フローエントリを、前記第１フローエントリに替えて前記スイッチに設定し、

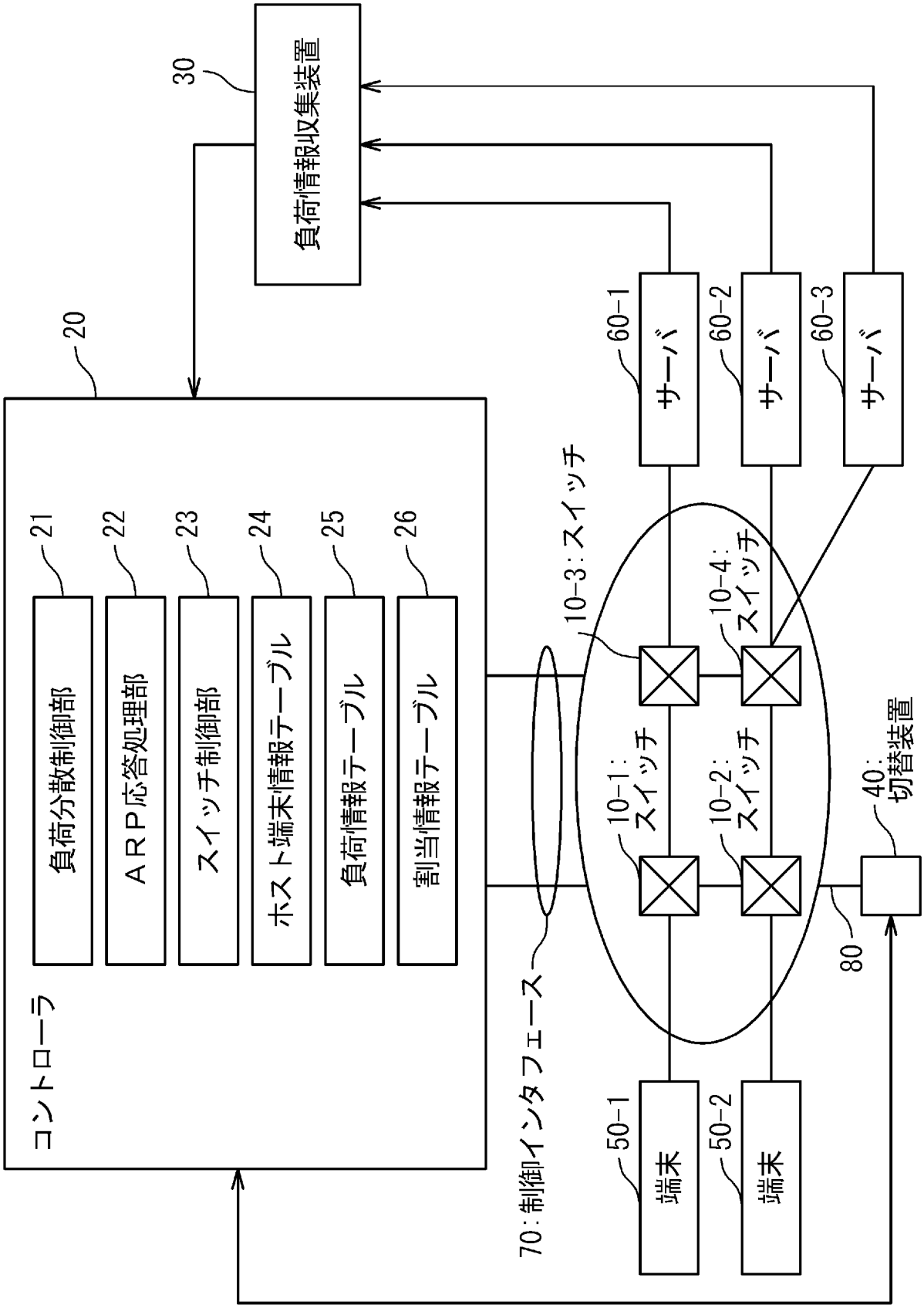
前記スイッチが、前記第２フローエントリに従い、前記第１ＭＡＣアドレスを宛先ＭＡＣアドレスとしたパケットを、前記切替装置へ転送するステップと、

前記切替装置が、前記スイッチからの受信パケットの宛先ＭＡＣアドレスを、前記第１ＭＡＣアドレスから、第１サーバとは異なる第２サーバの第２ＭＡＣアドレスに変更して、前記第２サーバに至る通信経路へ転送するステップと

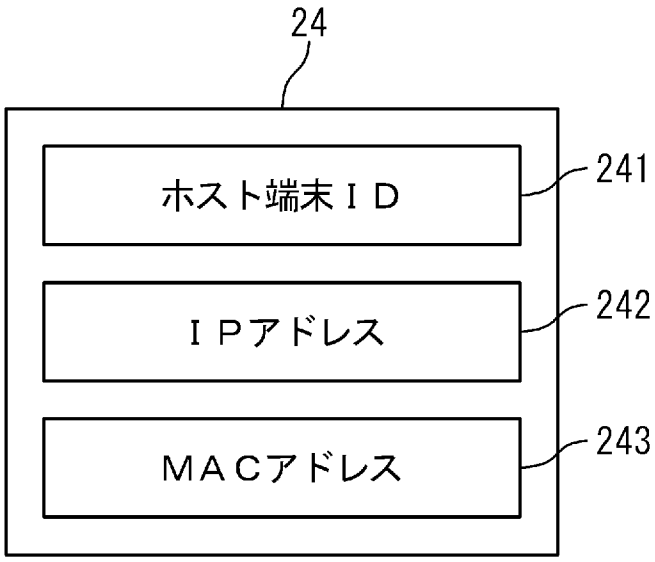
を具備する

通信方法。

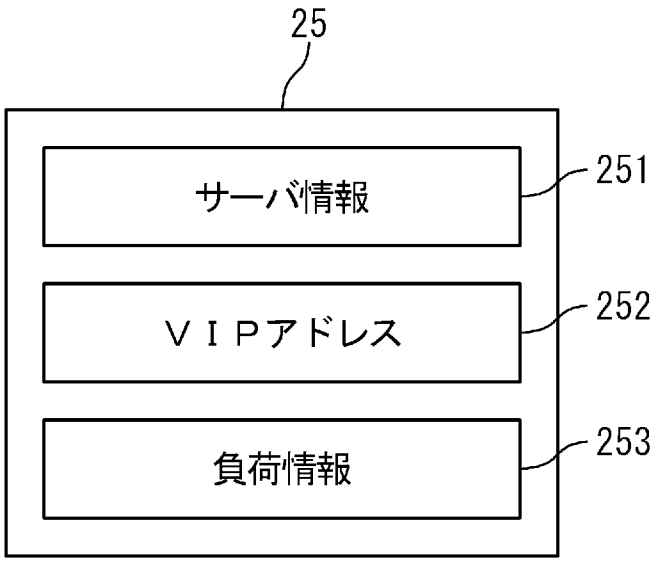
[図1]


30
負荷情報収集装置
20
コントローラ
21
負荷分散制御部
22
ＡＲＰ応答処理部
23
スイッチ制御部
24
ホスト端末情報テーブル
25
負荷情報テーブル
26
割当情報テーブル
70：制御インタフェース
60-1
サーバ
60-2
サーバ
60-3
サーバ
10-3：スイッチ
10-4：スイッチ
10-1：スイッチ
10-2：スイッチ
40：切替装置
80
50-1
端末
50-2
端末

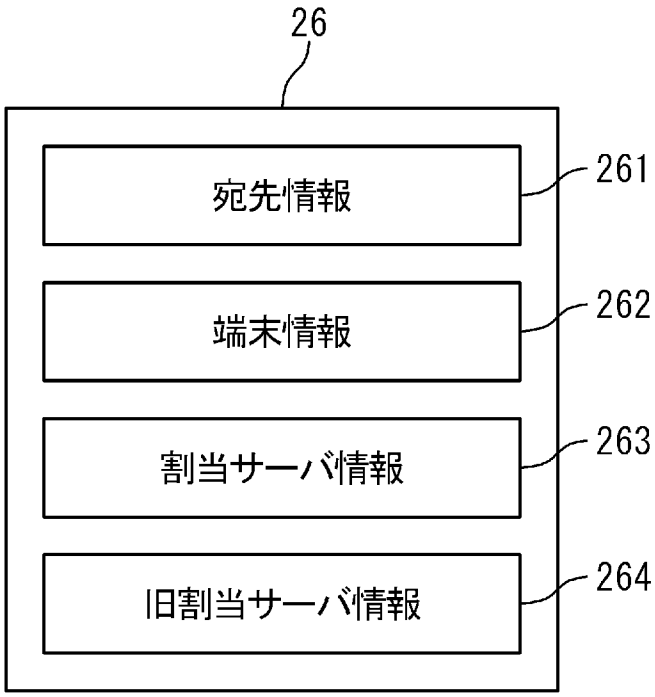
[図2]


24
ホスト端末ＩＤ
241
ＩＰアドレス
242
ＭＡＣアドレス
243

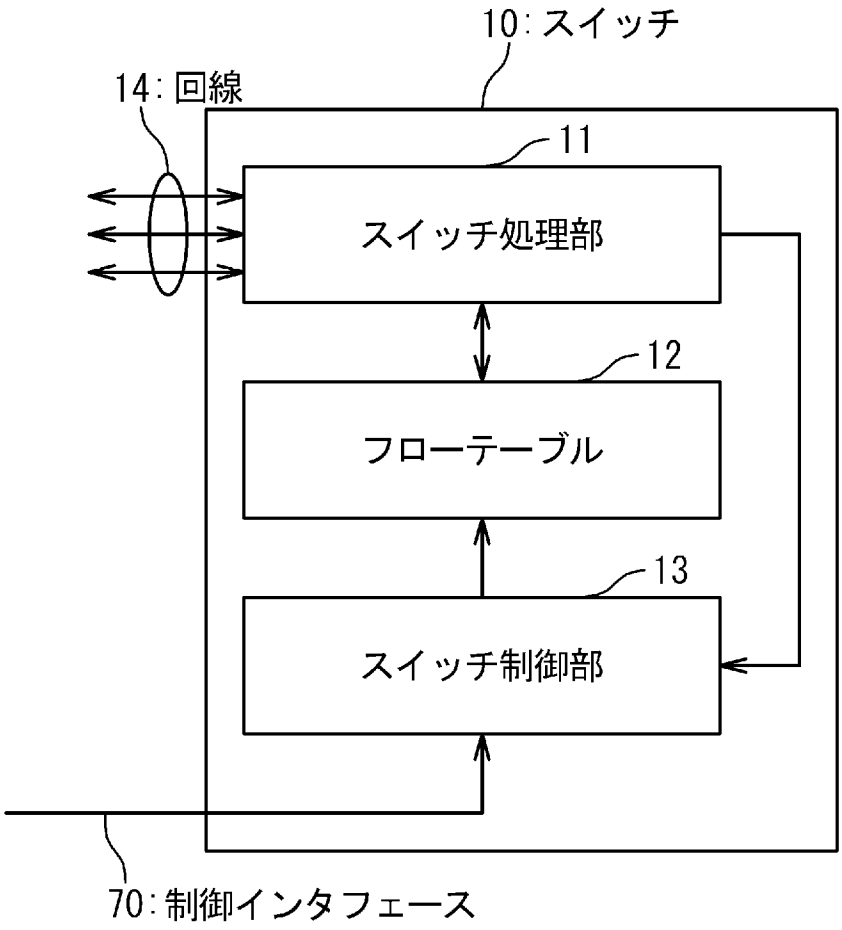
[図3]


25
サーバ情報
251
ＶＩＰアドレス
252
負荷情報
253

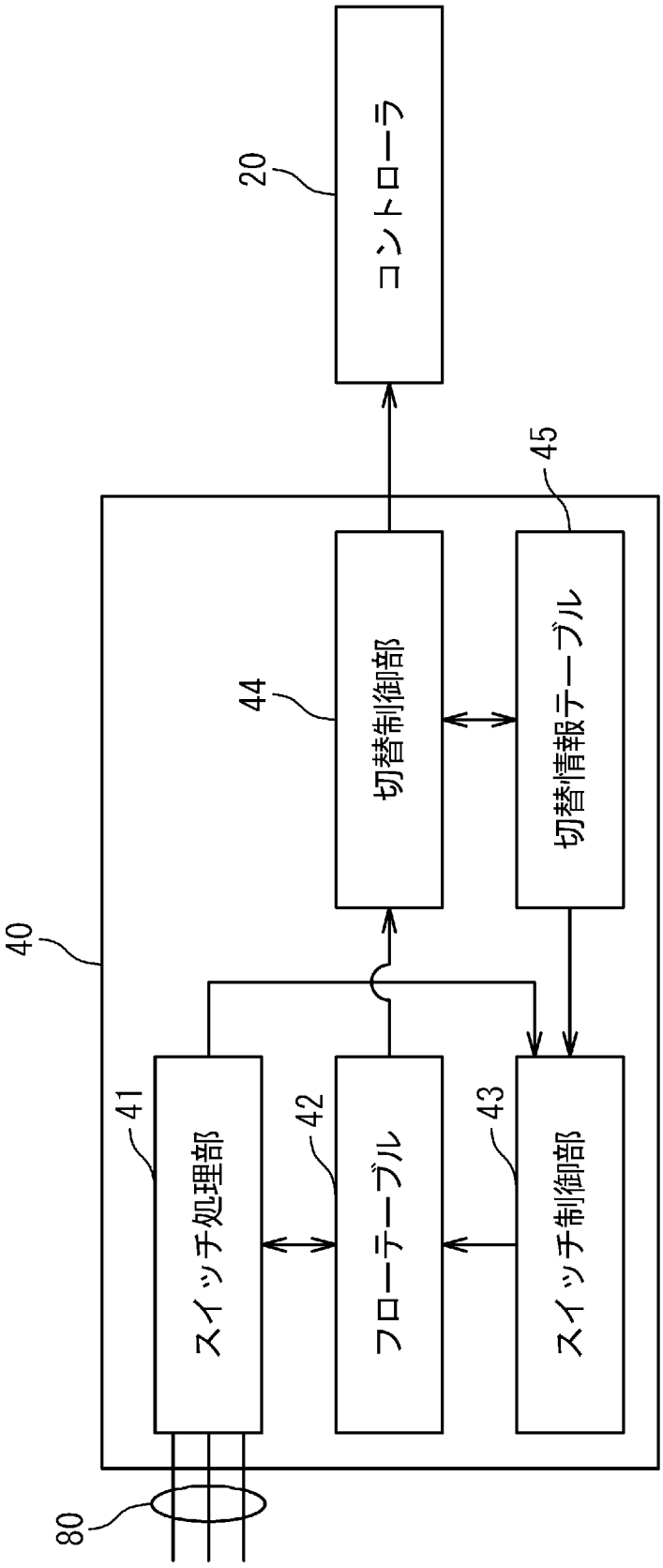
[図4]


26
宛先情報
261
端末情報
262
割当サーバ情報
263
旧割当サーバ情報
264

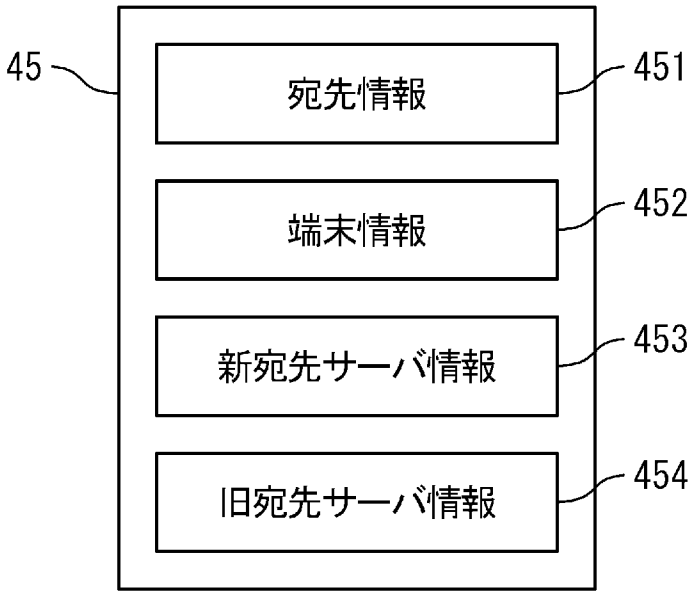
[図5]


10:スイッチ
14:回線
11
スイッチ処理部
12
フローテーブル
13
スイッチ制御部
70:制御インタフェース

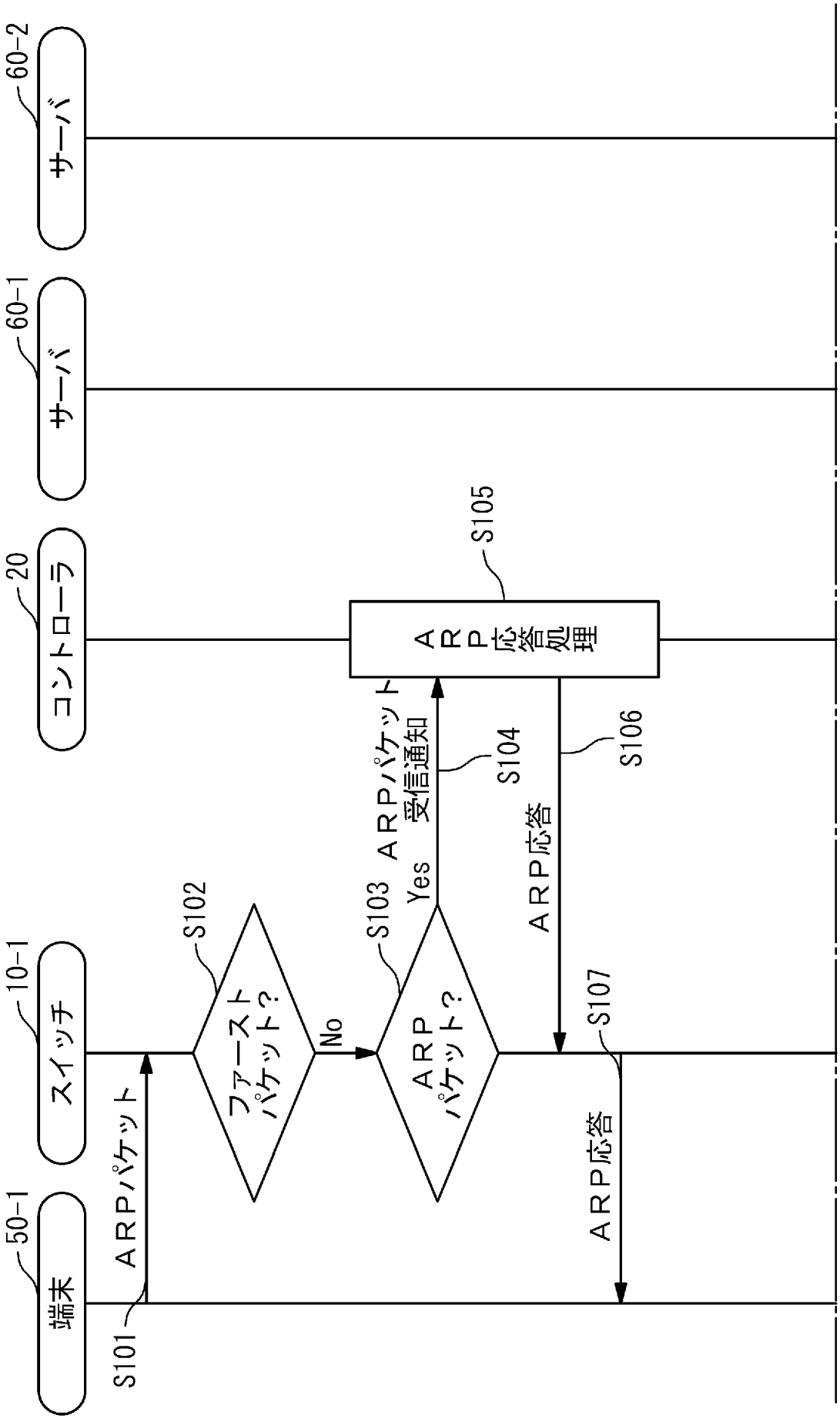
[図6]


20
コントローラ
45
44
切替制御部
切替情報テーブル
40
41
スイッチ処理部
42
フローテーブル
43
スイッチ制御部
80

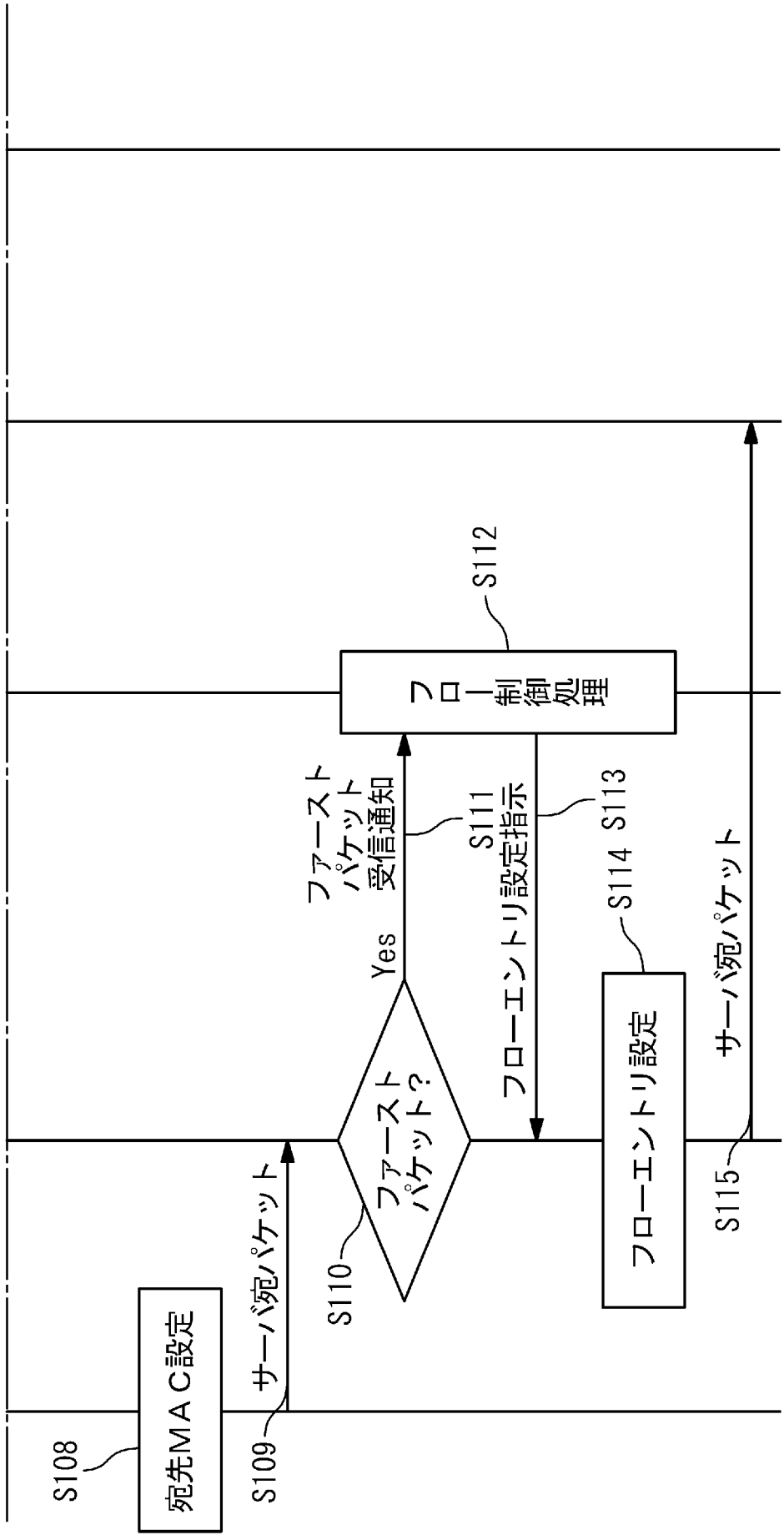
[図7]



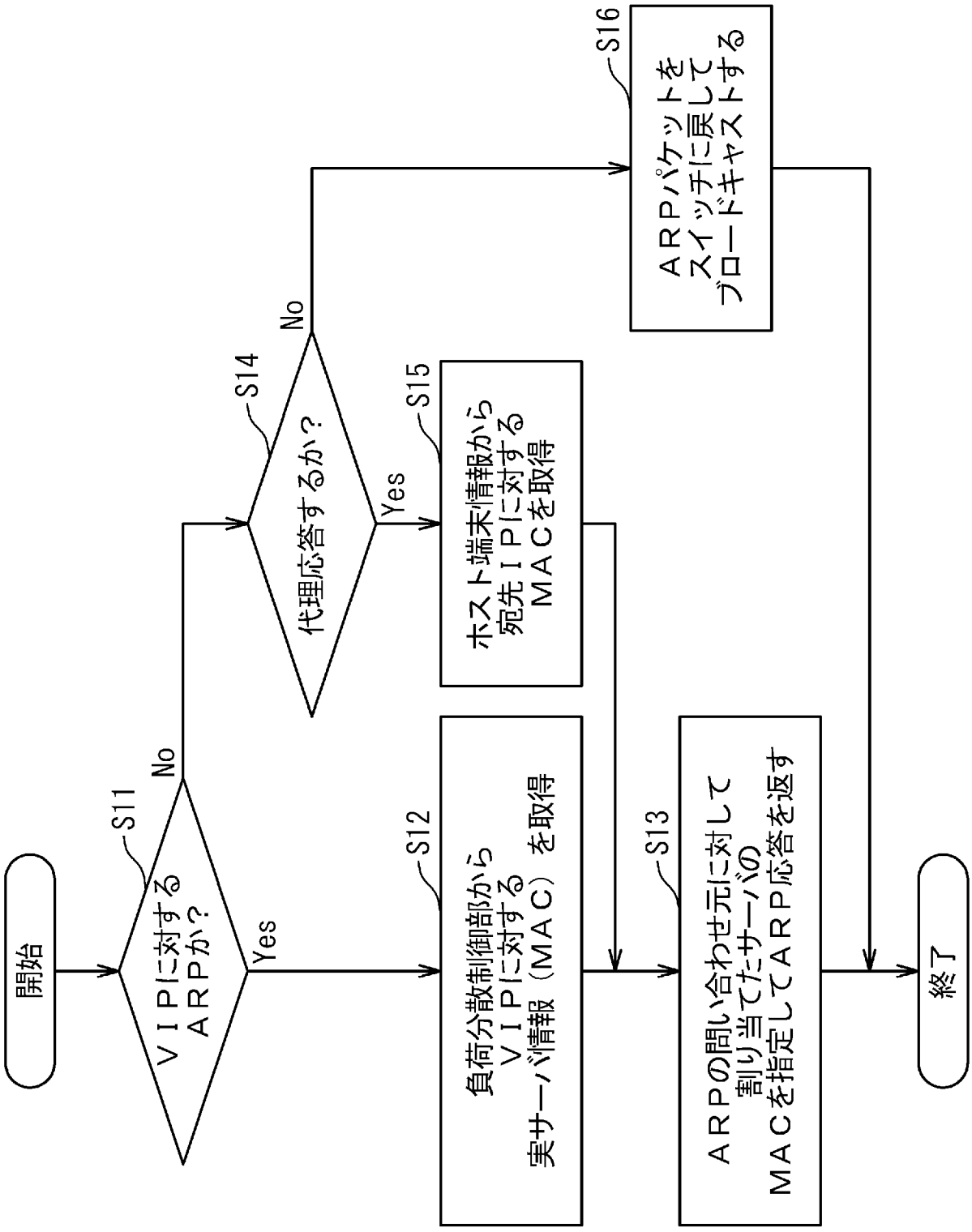
[図8A]


60-2
サーバ
60-1
サーバ
20
コントローラ
S105
ARP応答処理
S104
ARPパケット受信通知
Yes
S106
ARP応答
10-1
スイッチ
S102
ファーストパケット?
No
S103
ARPパケット?
S107
ARP応答
50-1
端末
ARPパケット
S101

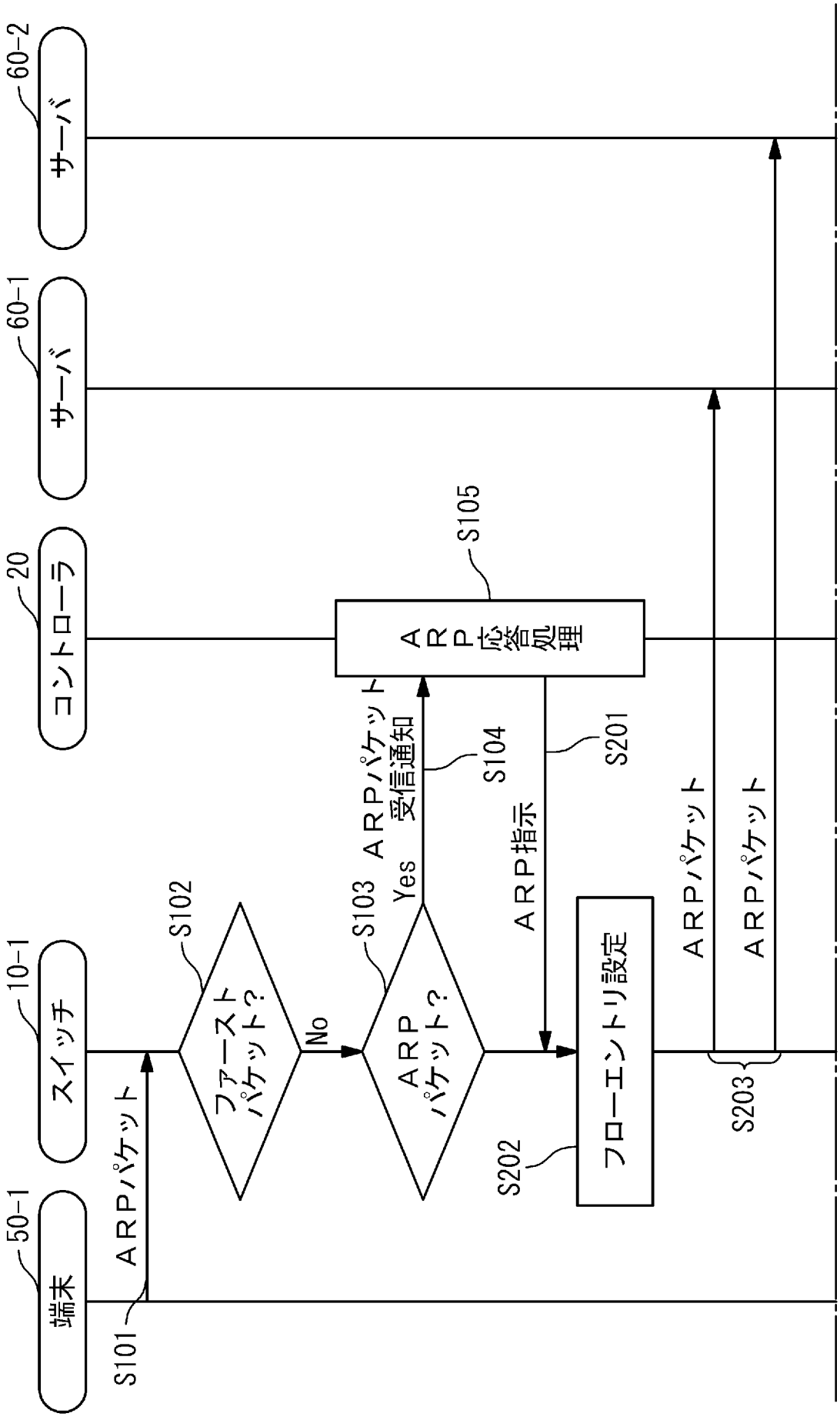
[図8B]


S108
宛先MAC設定
S109
サーバ宛パケット
S110
ファーストパケット?
Yes
ファーストパケット受信通知
S111
フロー制御処理
S112
フローエントリ設定指示
S113
フローエントリ設定
S114
サーバ宛パケット
S115

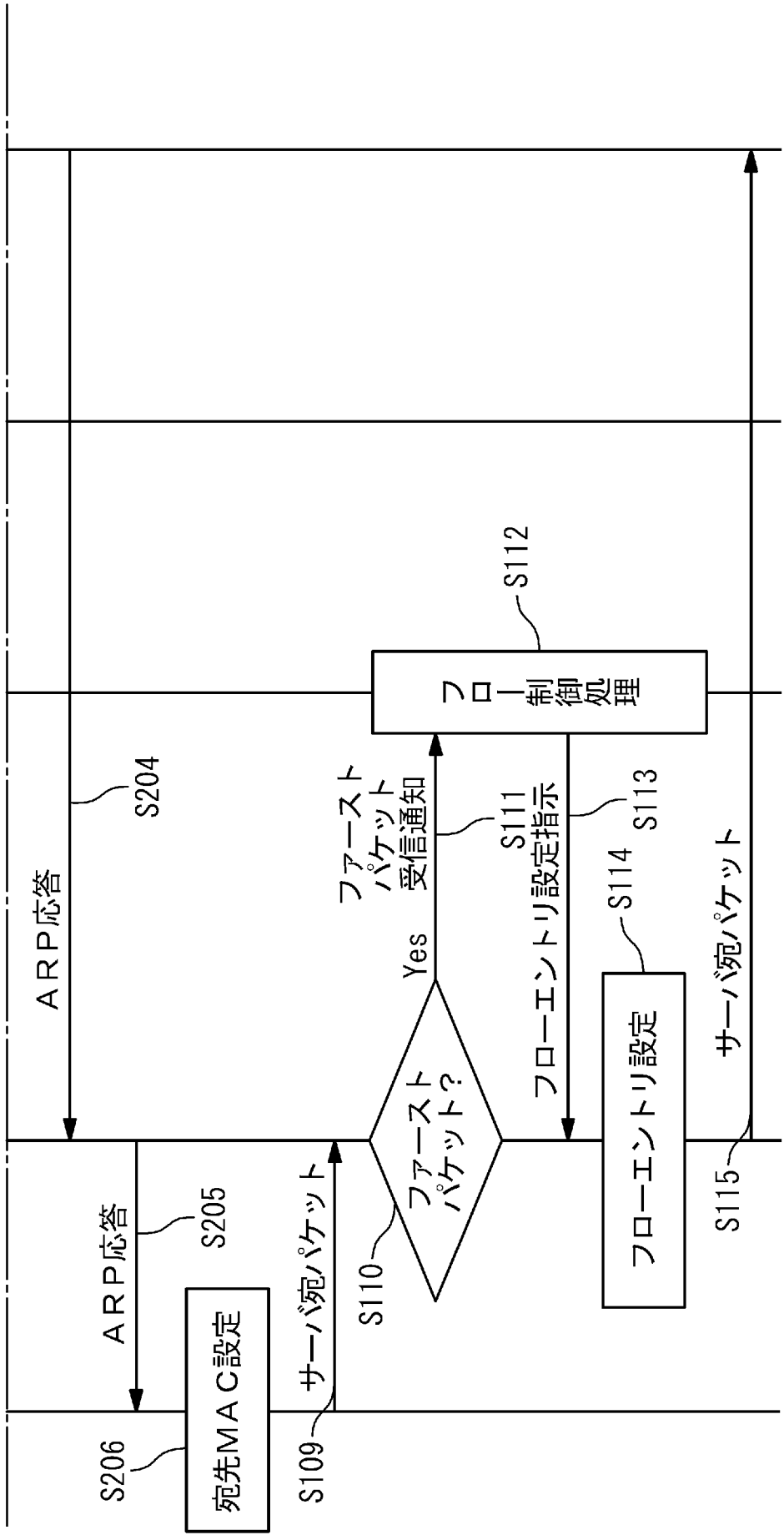
[図9]


開始
S11
VIPに対する
ARPか？
No
Yes
S12
負荷分散制御部から
VIPに対する
実サーバ情報（MAC）を取得
S13
ARPの問い合わせ元に対して
割り当てたサーバの
MACを指定してARP応答を返す
終了
S14
代理応答するか？
No
Yes
S15
ホスト端末情報から
宛先IPに対する
MACを取得
S16
ARPパケットを
スイッチに戻して
ブロードキャストする

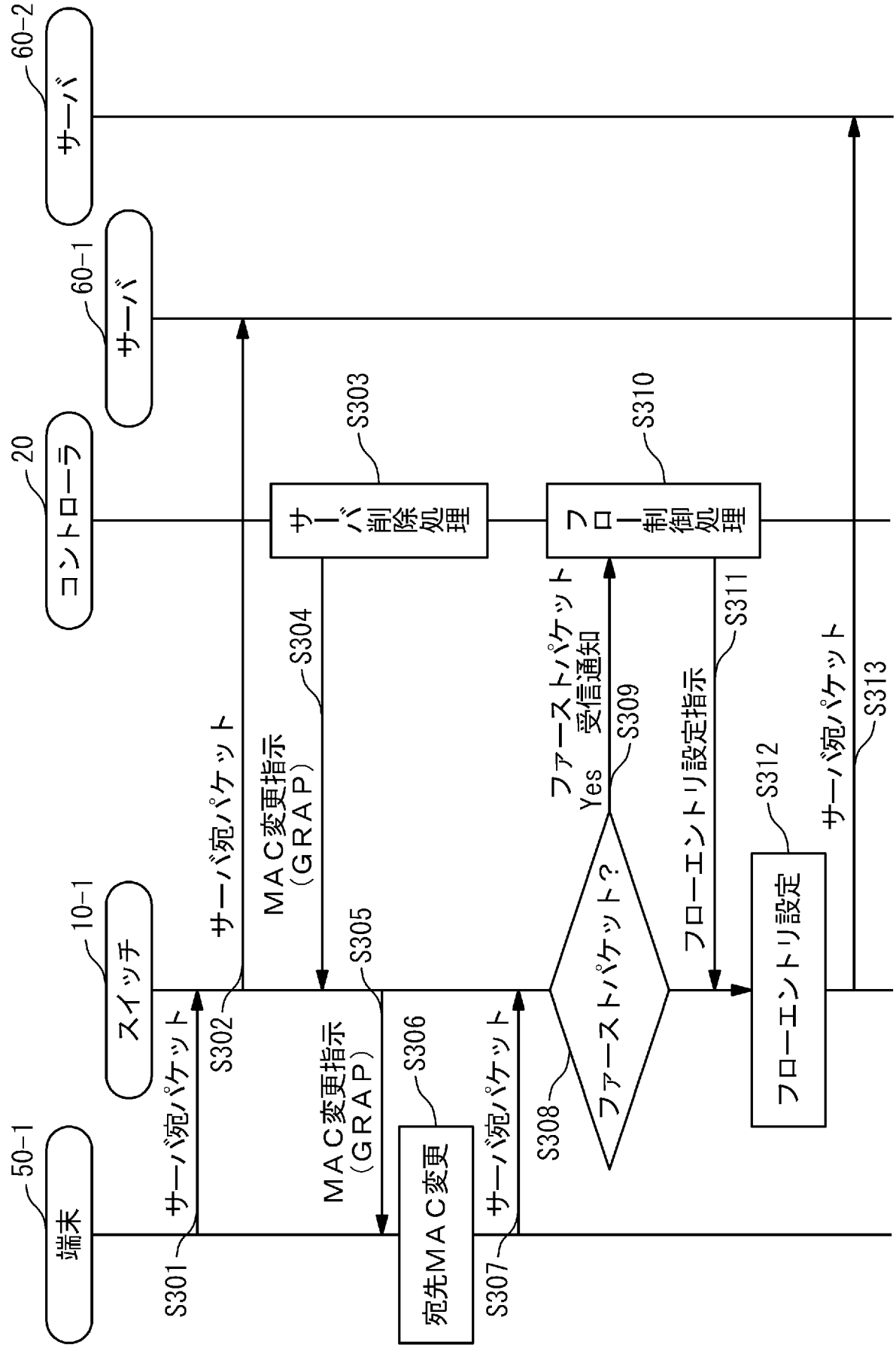
[図10A]


50-1
端末
10-1
スイッチ
20
コントローラ
60-1
サーバ
60-2
サーバ
S101
ARPパケット
S102
ファースト
パケット?
No
S103
ARP
パケット?
Yes
ARPパケット
受信通知
S104
S105
ARP応答処理
ARP指示
S201
S202
フローエントリ設定
ARPパケット
ARPパケット
S203

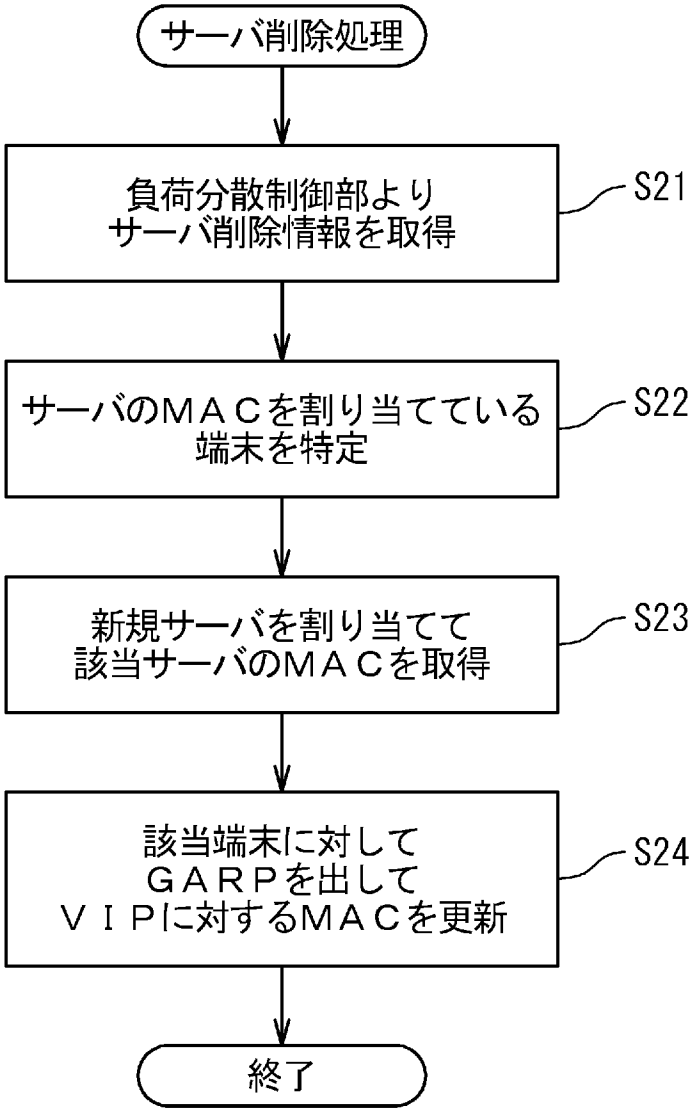
[図10B]


S112
フロー制御処理
S204
ARP応答
ファーストパケット受信通知
S111
フローエントリ設定指示
S113
S114
Yes
ファーストパケット?
フローエントリ設定
サーバ宛パケット
S115
S205
ARP応答
S110
サーバ宛パケット
S206
宛先MAC設定
S109

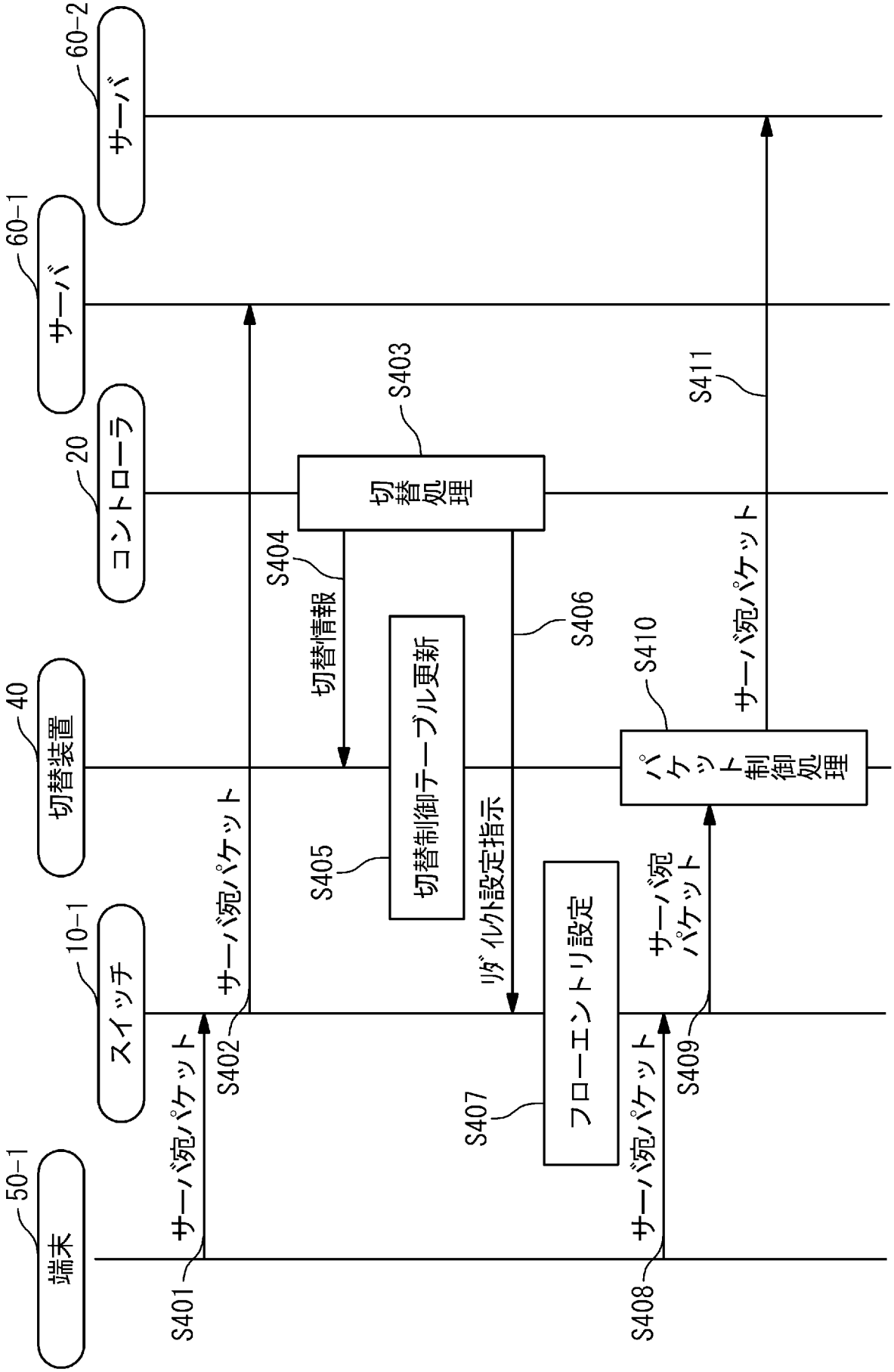
[図11]


60-2
サーバ
60-1
サーバ
20
コントローラ
10-1
スイッチ
50-1
端末
サーバ宛パケット
S301
サーバ宛パケット
S302
MAC変更指示
(GRAP)
S303
サーバ削除処理
S304
MAC変更指示
(GRAP)
S305
S306
宛先MAC変更
サーバ宛パケット
S307
S308
ファーストパケット?
ファーストパケット
Yes 受信通知
S309
S310
フロー制御処理
S311
フローエントリ設定指示
S312
フローエントリ設定
サーバ宛パケット
S313

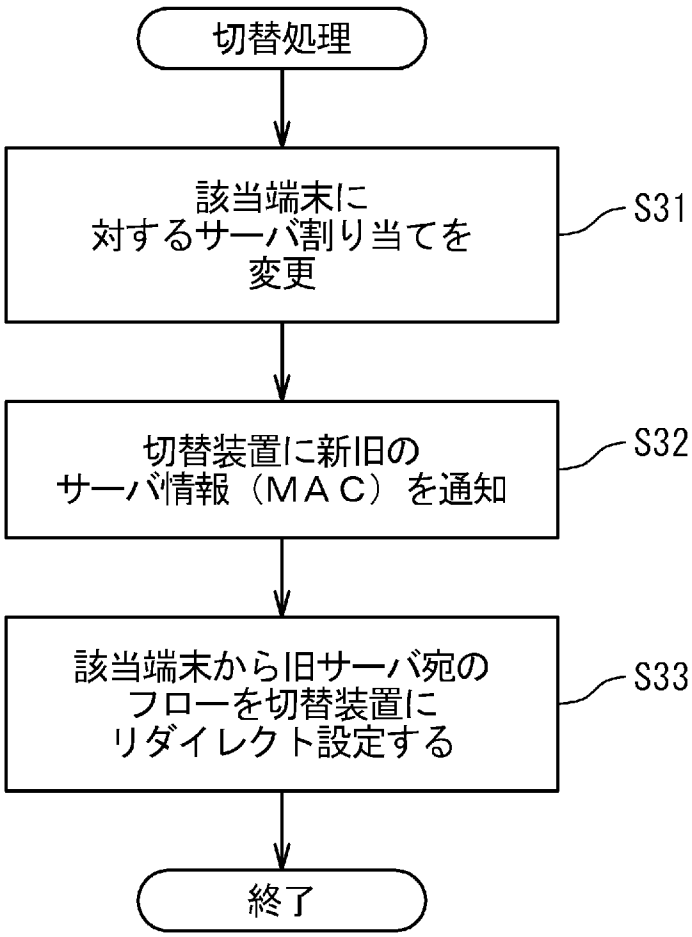
[図12]


サーバ削除処理
負荷分散制御部より
サーバ削除情報を取得
S21
サーバのＭＡＣを割り当てている
端末を特定
S22
新規サーバを割り当てて
該当サーバのＭＡＣを取得
S23
該当端末に対して
ＧＡＲＰを出して
ＶＩＰに対するＭＡＣを更新
S24
終了

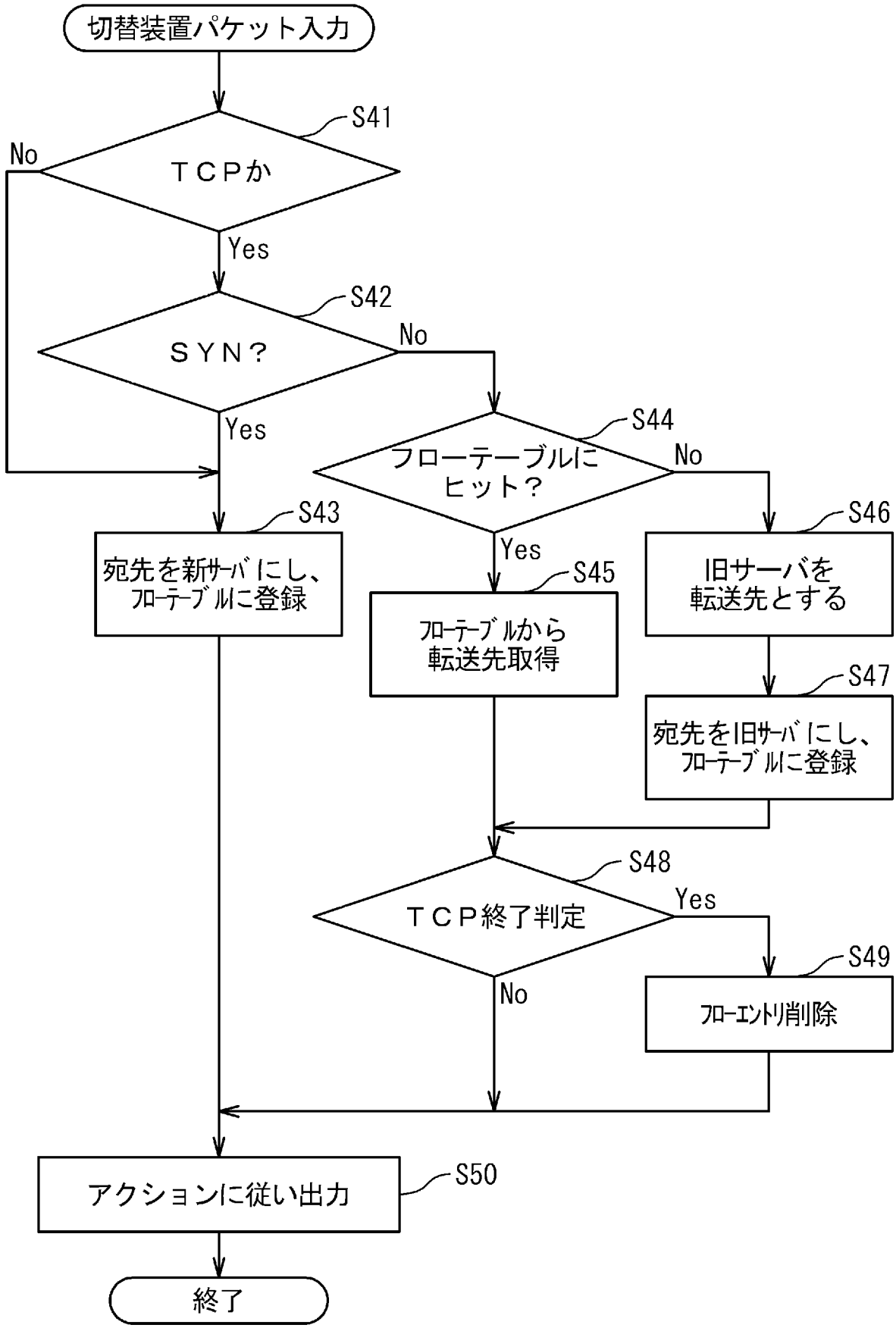
[図13]


50-1
端末
10-1
スイッチ
40
切替装置
20
コントローラ
60-1
サーバ
60-2
サーバ
S401
サーバ宛パケット
S402
サーバ宛パケット
S403
切替処理
S404
切替情報
S405
切替制御テーブル更新
S406
リダイレクト設定指示
S407
フローエントリ設定
S408
サーバ宛パケット
S409
サーバ宛パケット
S410
パケット制御処理
S411
サーバ宛パケット

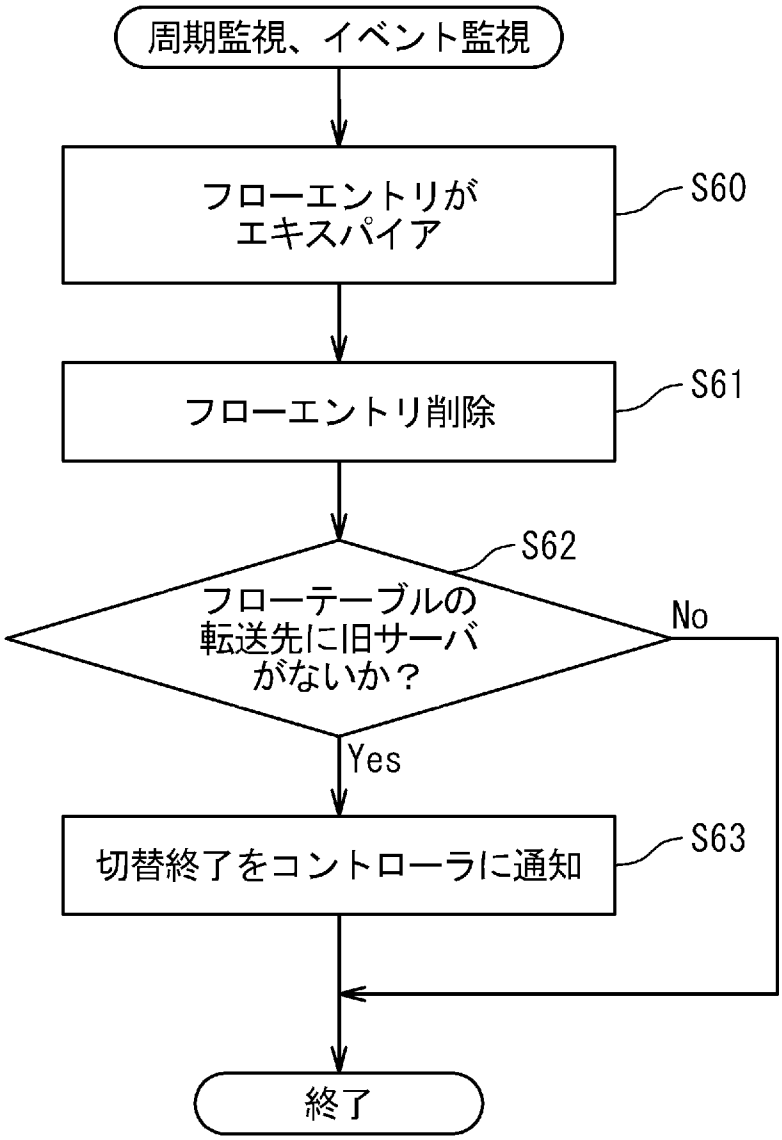
[図14]


切替処理
該当端末に
対するサーバ割り当てを
変更
S31
切替装置に新旧の
サーバ情報（ＭＡＣ）を通知
S32
該当端末から旧サーバ宛の
フローを切替装置に
リダイレクト設定する
S33
終了

[図15]


切替装置パケット入力
S41
TCPか
No
Yes
S42
SYN?
No
Yes
S44
フローテーブルに
ヒット?
No
Yes
S43
宛先を新サーバにし、
フローテーブルに登録
S45
フローテーブルから
転送先取得
S46
旧サーバを
転送先とする
S47
宛先を旧サーバにし、
フローテーブルに登録
S48
TCP終了判定
Yes
No
S49
フローエントリ削除
アクションに従い出力
S50
終了

[図16]


周期監視、イベント監視
フローエントリが
エキスパイア
S60
フローエントリ削除
S61
S62
フローテーブルの
転送先に旧サーバ
がないか？
No
Yes
切替終了をコントローラに通知
S63
終了

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2012/057727

A. CLASSIFICATION OF SUBJECT MATTER
H04L12/28(2006.01)i, *G06F9/50*(2006.01)i, *G06F13/00*(2006.01)i, *H04L12/56*(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
H04L12/28, G06F9/50, G06F13/00, H04L12/56

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho	1922-1996	Jitsuyo Shinan Toroku Koho	1996-2012
Kokai Jitsuyo Shinan Koho	1971-2012	Toroku Jitsuyo Shinan Koho	1994-2012

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	Hiroshi UENO, Masaharu MORIMOTO, Takafumi AIDA, Takahide SUGITA, "A Study on Deployment of Network Appliance Functionalities in Datacenter Network", IEICE Technical Report CPSY2009-36, 13 November 2009 (13.11.2009), vol.109 no.296, pages 7 to 12	1-10
Y	US 2009/0158082 A1 (IBM CORP(AP)), 18 June 2009 (18.06.2009), paragraphs [0052] to [0054] & JP 2011-507426 A & EP 2232828 A & WO 2009/077292 A1 & DE 602008005937 D & CA 2699314 A & KR 10-2010-0097694 A & CN 101884208 A & AT 504153 T	1-10

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
10 April, 2012 (10.04.12)	17 April, 2012 (17.04.12)
Name and mailing address of the ISA/ Japanese Patent Office Facsimile No.	Authorized officer Telephone No.

国際調査報告

国際出願番号　ＰＣＴ／ＪＰ２０１２／０５７７２７

Ａ．　発明の属する分野の分類（国際特許分類（ＩＰＣ））

Int.Cl.　H04L12/28(2006.01)i, G06F9/50(2006.01)i, G06F13/00(2006.01)i, H04L12/56(2006.01)i

Ｂ．　調査を行った分野

調査を行った最小限資料（国際特許分類（ＩＰＣ））

Int.Cl.　H04L12/28, G06F9/50, G06F13/00, H04L12/56

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	１９２２－１９９６年
日本国公開実用新案公報	１９７１－２０１２年
日本国実用新案登録公報	１９９６－２０１２年
日本国登録実用新案公報	１９９４－２０１２年

国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）

Ｃ．　関連すると認められる文献

引用文献のカテゴリー＊	引用文献名　及び一部の箇所が関連するときは、その関連する箇所の表示	関連する請求項の番号
Ｙ	上野洋史・森本昌治・相田貴史・杉田貴英，データセンターネットワークにおけるネットワークアプライアンス機能配備の一検討，電子情報通信学会技術研究報告 CPSY2009-36，2009.11.13，Ｖｏｌ．１０９　Ｎｏ．２９６，pp.7-12	1-10
Ｙ	US 2009/0158082 A1（ＩＢＭ　ＣＯＲＰ（ＡＰ））2009.06.18，【００５２】－【００５４】 & JP 2011-507426 A & EP 2232828 A & WO 2009/077292 A1 & DE 602008005937 D & CA 2699314 A & KR 10-2010-0097694 A & CN 101884208 A & AT 504153 T	1-10

☐　Ｃ欄の続きにも文献が列挙されている。

☐　パテントファミリーに関する別紙を参照。

＊　引用文献のカテゴリー
「Ａ」特に関連のある文献ではなく、一般的技術水準を示すもの
「Ｅ」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
「Ｌ」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）
「Ｏ」口頭による開示、使用、展示等に言及する文献
「Ｐ」国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献
「Ｔ」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの
「Ｘ」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
「Ｙ」特に関連のある文献であって、当該文献と他の１以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
「＆」同一パテントファミリー文献

国際調査を完了した日 １０．０４．２０１２	国際調査報告の発送日 １７．０４．２０１２
国際調査機関の名称及びあて先 日本国特許庁（ＩＳＡ／ＪＰ） 郵便番号１００－８９１５ 東京都千代田区霞が関三丁目４番３号	特許庁審査官（権限のある職員）　５Ｘ　３３６１ 岩田　玲彦 電話番号　０３－３５８１－１１０１　内線　３５９６