

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第5001966号
(P5001966)

(45) 発行日 平成24年8月15日(2012. 8. 15)

(24) 登録日 平成24年5月25日(2012. 5. 25)

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

請求項の数 10 (全 20 頁)

(21) 出願番号	特願2009-40564 (P2009-40564)	(73) 特許権者	000004226
(22) 出願日	平成21年2月24日(2009. 2. 24)		日本電信電話株式会社
(65) 公開番号	特開2010-199800 (P2010-199800A)		東京都千代田区大手町二丁目3番1号
(43) 公開日	平成22年9月9日(2010. 9. 9)	(74) 代理人	100064414
審査請求日	平成23年2月22日(2011. 2. 22)		弁理士 磯野 道造
		(74) 代理人	100127720
			弁理士 大石 恵
		(74) 代理人	100162374
			弁理士 中村 新二
		(72) 発明者	クリステル ベルサー
			東京都千代田区大手町二丁目3番1号 日
			本電信電話株式会社内
		(72) 発明者	増田 暁生
			東京都千代田区大手町二丁目3番1号 日
			本電信電話株式会社内

最終頁に続く

(54) 【発明の名称】 経路情報管理方法およびその管理システム

(57) 【特許請求の範囲】

【請求項 1】

I P (Internet Protocol) ネットワークを構成する A S (Autonomous System) 内で、他の A S から該他の A S のアドレスの集合であるプレフィックスを含む経路情報を収集しルーティングを行う経路情報管理システムにおいて用いられる経路情報管理方法であって、

前記経路情報管理システムは、前記他の A S のルータに直結する複数の境界ルータと、前記境界ルータによって取得された前記他の A S の経路情報を収集して、その収集した前記経路情報に基づいてデータの転送経路を決定する複数の経路管理装置とを備え、

前記境界ルータは、取得した前記経路情報に含まれる前記プレフィックスを用いて所定の演算を行い、その演算結果に基づいて該プレフィックスに対応付けられた経路管理装置に前記経路情報を送信し、

前記経路管理装置は、受信した該経路情報を記憶し、前記境界ルータから前記データの転送経路を問い合わせる問い合わせ情報を受けたときに、前記記憶した経路情報に基づいて転送経路を決定し、その決定した転送経路を含む経路決定情報を返信し、

前記問い合わせをした境界ルータは、該経路決定情報に基づいて前記データのルーティングを行う

ことを特徴とする経路情報管理方法。

【請求項 2】

前記経路情報管理システムは、さらに、自 A S 内に前記境界ルータとは別に中継ルータ

を備え、

前記境界ルータまたは中継ルータは、データを受信したとき、そのデータに含まれるプレフィックスを抽出し該プレフィックスを用いた前記所定の演算を行い、そのプレフィックスに対応付けられた経路管理装置に前記データの転送経路を問い合わせる問い合わせ情報を送信し、前記問い合わせ情報に対する応答として前記経路管理装置から送信される前記経路決定情報を受信して、その受信した前記経路決定情報に基づいて前記データのルーティングを行う

ことを特徴とする請求項 1 に記載の経路情報管理方法。

【請求項 3】

前記境界ルータおよび中継ルータは、処理部と前記経路管理装置を識別する経路管理装置 ID を記憶する記憶部とを備え、

前記処理部が、前記プレフィックスの値を所定のハッシュ関数に適用してハッシュ値 k を算出し、前記ハッシュ値 k と記憶部に記憶してある前記経路管理装置 ID とを比較して、前記ハッシュ値 k 以上の範囲において最小の値となる前記経路管理装置 ID を算出することによって、前記プレフィックスを用いた前記所定の演算を実行する

ことを特徴とする請求項 1 または請求項 2 に記載の経路情報管理方法。

【請求項 4】

前記経路管理装置は、処理部と記憶部とを備え、

前記記憶部に記憶される前記経路情報には該経路情報を受信した境界ルータの識別情報およびその受信に使用されたプロトコル種別と、前記境界ルータおよびルータ間の経路に付与されたコスト情報とを含み、さらに、前記記憶部に AS 内のトポロジ情報を記憶し、前記処理部が、

前記問い合わせ情報を受け付けて前記データの転送経路を決定するとき、

(1) 前記問い合わせ情報の送信元が境界ルータの場合には、

(1-1) 当該境界ルータが e B G P で受信した経路情報を検索し、この検索した経路情報を用いて転送経路を選択する手順を実行し、

当該手順によって複数の転送経路が選択された場合には、

(1-2) さらに、前記トポロジ情報を参照して、該問い合わせ情報の送信元の境界ルータのトポロジ上の位置を特定し、前記転送経路の候補となる経路に付与されるコストの最も小さい経路を転送経路として選択する手順を実行して、転送経路を決定し、

(2) 前記問い合わせ情報の送信元が中継ルータの場合には、前記トポロジ情報を参照して、該問い合わせ情報の送信元の境界ルータのトポロジ上の位置を特定し、前記転送経路の候補となる経路に付与されるコストの最も小さい経路を転送経路として選択する手順を実行して、転送経路を決定する

ことを特徴とする請求項 2 に記載の経路情報管理方法。

【請求項 5】

前記経路管理装置は、少なくとも 2 つの転送経路を含む経路決定情報を前記境界ルータまたは前記中継ルータに送信し、

該経路決定情報を受信した前記境界ルータまたは前記中継ルータは、ネクストホップとの通信が切断したときに、送信されてきた該経路決定情報に含まれる別の転送経路に切り替える

ことを特徴とする請求項 4 に記載の経路情報管理方法。

【請求項 6】

IP (Internet Protocol) ネットワークを構成する AS (Autonomous System) 内で、他の AS から該他の AS のアドレスの集合であるプレフィックスを含む経路情報を収集しルーティングを行う経路情報管理システムであって、

前記経路情報管理システムは、前記他の AS のルータに直結する複数の境界ルータと、前記境界ルータによって取得された前記他の AS の経路情報を収集して、その収集した前記経路情報に基づいてデータの転送経路を決定する複数の経路管理装置とを備え、

前記境界ルータは、取得した前記経路情報に含まれる前記プレフィックスを用いて所定

の演算を行い、その演算結果に基づいて該プレフィックスに対応付けられた経路管理装置に前記経路情報を送信し、

前記経路管理装置は、受信した該経路情報を記憶し、前記境界ルータから前記データの転送経路を問い合わせる問い合わせ情報を受けたときに、前記記憶した経路情報に基づいて転送経路を決定し、その決定した転送経路を含む経路決定情報を返信し、

前記問い合わせをした境界ルータは、該経路決定情報に基づいて前記データのルーティングを行う

ことを特徴とする経路情報管理システム。

【請求項 7】

前記経路情報管理システムは、さらに、自 A S 内に前記境界ルータとは別に中継ルータを備え、

前記境界ルータまたは中継ルータは、データを受信したとき、そのデータに含まれるプレフィックスを抽出し該プレフィックスを用いた前記所定の演算を行い、そのプレフィックスに対応付けられた経路管理装置に前記データの転送経路を問い合わせる問い合わせ情報を送信し、前記問い合わせ情報に対する応答として前記経路管理装置から送信される前記経路決定情報を受信して、その受信した前記経路決定情報に基づいて前記データのルーティングを行う

ことを特徴とする請求項 6 に記載の経路情報管理システム。

【請求項 8】

前記境界ルータおよび中継ルータは、処理部と前記経路管理装置を識別する経路管理装置 ID を記憶する記憶部とを備え、

前記処理部が、前記プレフィックスの値を所定のハッシュ関数に適用してハッシュ値 k を算出し、前記ハッシュ値 k と記憶部に記憶してある前記経路管理装置 ID とを比較して、前記ハッシュ値 k 以上の範囲において最小の値となる前記経路管理装置 ID を算出することによって、前記プレフィックスを用いた前記所定の演算を実行する

ことを特徴とする請求項 6 または請求項 7 に記載の経路情報管理システム。

【請求項 9】

前記経路管理装置は、処理部と記憶部とを備え、

前記記憶部に記憶される前記経路情報には該経路情報を受信した境界ルータの識別情報およびその受信に使用されたプロトコル種別と、前記境界ルータおよびルータ間の経路に付与されたコスト情報とを含み、さらに、前記記憶部に A S 内のトポロジ情報を記憶し、前記処理部が、

前記問い合わせ情報を受け付けて前記データの転送経路を決定するとき、

(1) 前記問い合わせ情報の送信元が境界ルータの場合には、

(1 - 1) 当該境界ルータが e B G P で受信した経路情報を検索し、この検索した経路情報を用いて転送経路を選択する手順を実行し、

当該手順によって複数の転送経路が選択された場合には、

(1 - 2) さらに、前記トポロジ情報を参照して、該問い合わせ情報の送信元の境界ルータのトポロジ上の位置を特定し、前記転送経路の候補となる経路に付与されるコストの最も小さい経路を転送経路として選択する手順を実行して、転送経路を決定し、

(2) 前記問い合わせ情報の送信元が中継ルータの場合には、前記トポロジ情報を参照して、該問い合わせ情報の送信元の境界ルータのトポロジ上の位置を特定し、前記転送経路の候補となる経路に付与されるコストの最も小さい経路を転送経路として選択する手順を実行して、転送経路を決定する

ことを特徴とする請求項 7 に記載の経路情報管理システム。

【請求項 10】

前記経路管理装置は、少なくとも 2 つの転送経路を含む経路決定情報を前記境界ルータまたは前記中継ルータに送信し、

該経路決定情報を受信した前記境界ルータまたは前記中継ルータは、ネクストホップとの通信が切断したときに、送信されてきた該経路決定情報に含まれる別の転送経路に切り

替える

ことを特徴とする請求項 9 に記載の経路情報管理システム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、IP (Internet Protocol) ネットワークにおいて、AS (自律システム; Autonomous System) 内で、別の AS (一つの経路選択プロトコルによって経路を管理しているネットワークの範囲のこと) の経路情報を管理する技術に関する。

【背景技術】

【0002】

インターネット等の IP ネットワークは、異なる組織によって運営される多数の AS が接続されて形成されている。AS 内にはルータが配置され、データを送信する場合、どの経路でそのデータを伝達するかが、ルーティングプロトコルによって決定される。ルーティングプロトコルは、大きく分けて、AS 内用の IGP (Interior Gateway Protocol) と AS 間用の BGP (Border Gateway Protocol) とに分けられる。IGP はデータを AS 内の宛先に送信する場合に用いられ、BGP は異なる AS 間でデータを送信する場合に用いられる。そして、IGP および BGP は、データを送信するルータ間に隣接関係 (ピア) を結んで、互いの経路情報を交換し、交換した経路情報に基づいて、データを所定の宛先に伝達する経路を決定する。BGP では、異なる AS 間で用いる BGP を eBGP (external BGP) と呼び、AS 内で用いる BGP を iBGP (internal BGP) と呼んで区別している。なお、iBGP は、BGP の情報を、他の AS から別の AS に転送する場合に用いられる。

【0003】

IGP や BGP で交換される経路情報は、プレフィックス (prefix) とパス属性情報とを含んで形成される。プレフィックスは、ネットワークアドレスとそのネットワークアドレスのネットマスクによって表記される宛先のアドレスの集合である。また、パス属性情報は経路情報に付加される当該経路の属性情報であり、経路を選択する際に用いられるパラメータである。

【0004】

ここで、BGP に基づく AS 内の経路情報管理の従来技術について説明する。

例えば、図 12 に示すように、AS 1a 内に、複数のルータ R1a, R2a, R3a, …, Rna が配置され、ルータ R1a ~ Rna が相互に接続されているものとする。いま、プレフィックス情報 P1a を含む経路情報が AS 2a から AS 1a 内のルータ R1a に eBGP を用いて広告されたとき、ルータ R1a は、iBGP を用いて、AS 1a 内の他のルータ R2a ~ Rna に、「ルータ R1a にデータを送れば、プレフィックス P1 に届けられる旨」の経路情報を広告する。ただし、iBGP を用いてルータ R1a ~ Rna 間で広告する場合には、IGP によってルータ R1a ~ Rna 間が通信可能に接続されている必要がある。このようにして、各ルータ R1a ~ Rna は、広告された経路情報を記憶して、その記憶した経路情報に基づいて、データの転送経路を決めている。ただし、iBGP では、一般的に AS 1a 内のルータ R1a ~ Rna 間をフルメッシュにピアを結んで運用される。そのため、ルータ数が増加傾向にある AS においては、ルータ記憶容量を大きなものに更改したり、経路計算のために演算速度の早い高性能なものに更改したり、また、回線を通信帯域の広いものに更改したり、つまり設備の更改を行う必要が発生するという問題がある。

【0005】

そこで、数百ルータによって構成されるような AS の場合には、ルートリフレクタ (RR) (非特許文献 1 参照) または AS フェデレーション (非特許文献 2 参照) が用いられている。ルートリフレクタ (RR) は、広告された経路情報をピアを結んだルートリフレクタクライアントに転送する。しかし、ルートリフレクタクライアント同士はピアを

10

20

30

40

50

確立しないのでよいため、AS内に確立すべきピア数をフルメッシュの場合より減少することが可能となる。また、ASコンフェデレーションは、AS内にいくつかのサブASを構成することによって、ピア数を減少することが可能である。そのため、ルータの記憶容量の増加や演算速度の高速化や回線の広帯域化を抑制することが可能となる。

【先行技術文献】

【非特許文献】

【0006】

【非特許文献1】T.Bates, E.Chen, and R.Chandra, "BGP Route Reflection: An Alternative to Full Mesh Internal BGP (IBGP)", RFC4456, IETF, 2006年4月

【非特許文献2】P.Traina, D.McPherson, and J.Scudder, "Autonomous System Confederations for BGP", RFC5065, IETF, 2007年8月

10

【発明の概要】

【発明が解決しようとする課題】

【0007】

しかしながら、非特許文献1に記載のルートリフレクタ(RR)を用いる方法、および非特許文献2に記載のASコンフェデレーションを用いる方法では、iBGPトポロジがフルメッシュでなくなり、また情報を交換する際にTCPによって張られるiBGPセッションにおいてプレフィックスごとに1つの経路しか広告されない。そのため、AS内の各ルータが経路情報の一部しか記憶しないことに起因する問題が発生する。すなわち、AS内の各ルータは、経路選択を行う際に、経路情報の一部に基づいて経路計算を行ってしまうという問題がある。さらに、その一部の経路情報から選択された経路情報が、IGPによってAS内のルータに広告され、また、iBGPによっても広告される。このような状態において、IGPおよびiBGPによって広告される経路情報が異なるタイミングでルータに届けられた場合には、iBGPピアのルータ間で異なる経路情報を保持してしまい、広告しあうことによって、経路決定がいつまでも収束しないという問題がある。

20

【0008】

また、ルートリフレクタを用いる方法では、ルートリフレクタが、パケットの転送元ルータからではなく、自身から見た最適経路を選択して各ルータに通達するので、iBGPトポロジとIGPトポロジとが乖離してしまう。この場合、パケットが境界ルータへ向かって転送されていくIGP経路上のルータにおいて意図しない境界ルータが選択されてしまうことが発生する。これによって、AS内でループが発生する問題や、AS間経路制御のポリシーにおいて一般的に用いられる外部AS宛のパケットはできるだけ早く他のASに送出するルーティングができない問題、すなわち、ホットポテトルーティングができない問題が発生する。

30

【0009】

そこで、本発明の課題は、前記した問題を解決するために、AS内のルータ数の増加に対して設備の大幅な更改を抑制でき、経路情報を全て把握できるようにしてAS内の経路情報を統一する技術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】

【0010】

前記課題を解決するために、請求項1に記載の発明は、IP(Internet Protocol)ネットワークを構成するAS(Autonomous System)内で、他のASから該他のASのアドレスの集合であるプレフィックスを含む経路情報を収集しルーティングを行う経路情報管理システムにおいて用いられる経路情報管理方法であって、前記経路情報管理システムが、前記他のASのルータに直結する複数の境界ルータと、前記境界ルータによって取得された前記他のASの経路情報を収集して、その収集した前記経路情報に基づいてデータの転送経路を決定する複数の経路管理装置とを備え、前記境界ルータが、取得した前記経路情報に含まれる前記プレフィックスを用いて所定の演算を行い、その演算結果に基づいて該プレフィックスに対応付けられた経路管理装置に前記経路情報を送信し、前記経路管理装置が、受信した該経路情報を記憶し、前記境界ルータから前記データの転送経路を問

40

50

合わせる問い合わせ情報を受けたときに、前記記憶した経路情報に基づいて転送経路を決定し、その決定した転送経路を含む経路決定情報を返信し、前記問い合わせをした境界ルータは、該経路決定情報に基づいて前記データのルーティングを行うことを特徴とする。

【0011】

また、請求項6に記載の発明は、IP (Internet Protocol) ネットワークを構成するAS (Autonomous System) 内で、他のASから該他のASのアドレスの集合であるプレフィックスを含む経路情報を収集しルーティングを行う経路情報管理システムであって、前記経路情報管理システムが、前記他のASのルータに直結する複数の境界ルータと、前記境界ルータによって取得された前記他のASの経路情報を収集して、その収集した前記経路情報に基づいてデータの転送経路を決定する複数の経路管理装置とを備え、前記境界ルータが、取得した前記経路情報に含まれる前記プレフィックスを用いて所定の演算を行い、その演算結果に基づいて該プレフィックスに対応付けられた経路管理装置に前記経路情報を送信し、前記経路管理装置が、受信した該経路情報を記憶し、前記境界ルータから前記データの転送経路を問い合わせる問い合わせ情報を受けたときに、前記記憶した経路情報に基づいて転送経路を決定し、その決定した転送経路を含む経路決定情報を返信し、前記問い合わせをした境界ルータが、該経路決定情報に基づいて前記データのルーティングを行うことを特徴とする。

【0012】

このような構成によれば、AS内の複数の経路管理装置は、プレフィックスごとに経路情報を分担して収集することができる。そのため、従来のルータでは、一部の経路情報しか収集できなかったのに対して、経路管理装置は、収集したプレフィックスに係る経路情報については全てを管理し、その経路情報に基づいてAS内の全てのルータに共通の転送経路を決定することが可能となる。つまり、経路決定がいつまでも収束しないという問題およびループの発生を防ぐことができる。また、複数の経路管理装置を備えることによって、経路管理装置一台あたりの分担する経路情報の量を抑制することができ、AS内のルータ数の増加に対して設備の大幅な更改を抑制することができる。

【0013】

請求項2に記載の発明は、請求項1に記載の経路情報管理方法において、前記経路情報管理システムが、さらに、自AS内に前記境界ルータとは別に中継ルータを備え、前記境界ルータまたは中継ルータが、データを受信したとき、そのデータに含まれるプレフィックスを抽出し該プレフィックスを用いた前記所定の演算を行い、そのプレフィックスに対応付けられた経路管理装置に前記データの転送経路を問い合わせる問い合わせ情報を送信し、前記問い合わせ情報に対する応答として前記経路管理装置から送信される前記経路決定情報を受信して、その受信した前記経路決定情報に基づいて前記データのルーティングを行うことを特徴とする。

【0014】

また、請求項7に記載の発明は、請求項6に記載の経路情報管理システムにおいて、前記経路情報管理システムが、さらに、自AS内に前記境界ルータとは別に中継ルータを備え、前記境界ルータまたは中継ルータが、データを受信したとき、そのデータに含まれるプレフィックスを抽出し該プレフィックスを用いた前記所定の演算を行い、そのプレフィックスに対応付けられた経路管理装置に前記データの転送経路を問い合わせる問い合わせ情報を送信し、前記問い合わせ情報に対する応答として前記経路管理装置から送信される前記経路決定情報を受信して、その受信した前記経路決定情報に基づいて前記データのルーティングを行うことを特徴とする。

【0015】

このような構成によれば、AS内の境界ルータまたは中継ルータが、データに含まれるプレフィックスに対応付けられた経路管理装置に問い合わせ情報を送信し、その問い合わせ情報に対する応答として、経路決定情報を受信できる。そのため、AS内の全ての境界ルータおよび中継ルータにとって共通の転送経路を決定することが可能となる。つまり、経路決定がいつまでも収束しないという問題およびループの発生を防ぐことができる。

【 0 0 1 6 】

請求項 3 に記載の発明は、請求項 1 または請求項 2 に記載の経路情報管理方法において、前記境界ルータおよび中継ルータが、処理部と前記経路管理装置を識別する経路管理装置 ID を記憶する記憶部とを備え、前記処理部が、前記プレフィックスの値を所定のハッシュ関数に適用してハッシュ値 k を算出し、前記ハッシュ値 k と記憶部に記憶してある前記経路管理装置 ID とを比較して、前記ハッシュ値 k 以上の範囲において最小の値となる前記経路管理装置 ID を算出することによって、前記プレフィックスを用いた前記所定の演算を実行することを特徴とする。

【 0 0 1 7 】

また、請求項 8 に記載の発明は、請求項 6 または請求項 7 に記載の経路情報管理システムにおいて、前記境界ルータおよび中継ルータが、処理部と前記経路管理装置を識別する経路管理装置 ID を記憶する記憶部とを備え、前記処理部が、前記プレフィックスの値を所定のハッシュ関数に適用してハッシュ値 k を算出し、前記ハッシュ値 k と記憶部に記憶してある前記経路管理装置 ID とを比較して、前記ハッシュ値 k 以上の範囲において最小の値となる前記経路管理装置 ID を算出することによって、前記プレフィックスを用いた前記所定の演算を実行することを特徴とする。

【 0 0 1 8 】

このような構成によれば、AS 内の経路管理装置に ID を割り振るだけで、経路管理装置が分担するプレフィックスを決定することができ、複数の経路管理装置を容易に設定することが可能となる。

【 0 0 1 9 】

請求項 4 に記載の発明は、請求項 2 に記載の経路情報管理方法において、前記経路管理装置が、処理部と記憶部とを備え、前記記憶部に記憶される前記経路情報には該経路情報を受信した境界ルータの識別情報およびその受信に使用されたプロトコル種別と、前記境界ルータおよびルータ間の経路に付与されたコスト情報とを含み、さらに、前記記憶部に AS 内のトポロジ情報を記憶し、前記処理部が、前記問い合わせ情報を受け付けて前記データの転送経路を決定するとき、(1) 前記問い合わせ情報の送信元が境界ルータの場合には、(1-1) 当該境界ルータが e B G P で受信した経路情報を検索し、この検索した経路情報を用いて転送経路を選択する手順を実行し、当該手順によって複数の転送経路が選択された場合には、(1-2) さらに、前記トポロジ情報を参照して、該問い合わせ情報の送信元の境界ルータのトポロジ上の位置を特定し、前記転送経路の候補となる経路に付与されるコストの最も小さい経路を転送経路として選択する手順を実行して、転送経路を決定し、(2) 前記問い合わせ情報の送信元が中継ルータの場合には、前記トポロジ情報を参照して、該問い合わせ情報の送信元の境界ルータのトポロジ上の位置を特定し、前記転送経路の候補となる経路に付与されるコストの最も小さい経路を転送経路として選択する手順を実行して、転送経路を決定することを特徴とする。

【 0 0 2 0 】

また、請求項 9 に記載の発明は、請求項 7 に記載の経路情報管理システムにおいて、前記経路管理装置が、処理部と記憶部とを備え、前記記憶部に記憶される前記経路情報には該経路情報を受信した境界ルータの識別情報およびその受信に使用されたプロトコル種別と、前記境界ルータおよびルータ間の経路に付与されたコスト情報とを含み、さらに、前記記憶部に AS 内のトポロジ情報を記憶し、前記処理部が、前記問い合わせ情報を受け付けて前記データの転送経路を決定するとき、(1) 前記問い合わせ情報の送信元が境界ルータの場合には、(1-1) 当該境界ルータが e B G P で受信した経路情報を検索し、この検索した経路情報を用いて転送経路を選択する手順を実行し、当該手順によって複数の転送経路が選択された場合には、(1-2) さらに、前記トポロジ情報を参照して、該問い合わせ情報の送信元の境界ルータのトポロジ上の位置を特定し、前記転送経路の候補となる経路に付与されるコストの最も小さい経路を転送経路として選択する手順を実行して、転送経路を決定し、(2) 前記問い合わせ情報の送信元が中継ルータの場合には、前記トポロジ情報を参照して、該問い合わせ情報の送信元の境界ルータのトポロジ上の位置を

特定し、前記転送経路の候補となる経路に付与されるコストの最も小さい経路を転送経路として選択する手順を実行して、転送経路を決定することを特徴とする。

【0021】

このような構成によれば、従来はルータが、自分を起点として経路選択を行っていたのに対して、経路管理装置は、他のルータを起点として、トポロジ上のルータの位置を考慮した転送経路を決定することが可能となる。

【0022】

請求項5に記載の発明は、請求項4に記載の経路情報管理方法において、前記経路管理装置が、少なくとも2つの転送経路を含む経路決定情報を前記境界ルータまたは前記中継ルータに送信し、該経路決定情報を受信した前記境界ルータまたは前記中継ルータが、ネクストホップとの通信が切断したときに、送信されてきた該経路決定情報に含まれる別の転送経路に切り替えることを特徴とする。

10

【0023】

また、請求項10に記載の発明は、請求項9に記載の経路情報管理システムにおいて、前記経路管理装置が、少なくとも2つの転送経路を含む経路決定情報を前記境界ルータまたは前記中継ルータに送信し、該経路決定情報を受信した前記境界ルータまたは前記中継ルータが、ネクストホップとの通信が切断したときに、送信されてきた該経路決定情報に含まれる別の転送経路に切り替えることを特徴とする。

【0024】

このような構成によれば、ルータは、複数の転送経路を含む経路決定情報を取得するので、ネクストホップとの通信が切断した際に残りの経路に容易に切り替えることが可能となる。

20

【発明の効果】

【0025】

本発明によれば、AS内のルータ数の増加に対して設備の大幅な更改を抑制でき、経路情報を全て把握できるようにしてAS内の経路情報を統一する技術を提供することができる。

【図面の簡単な説明】

【0026】

【図1】本実施形態の概要を説明する図である。

30

【図2】本実施形態における経路情報管理システムの構成例を示す図である。

【図3】ルートサーバ決定方法の例を示す図である。

【図4】ルートサーバの追加方法の例を示す図である。

【図5】ルートサーバの機能ブロックの例を示す図である。

【図6】経路表の例を示す図である。

【図7】ルータの機能ブロックの例を示す図である。

【図8】ルートサーバの経路情報収集手順の例を示す図である。

【図9】ルータの経路決定情報取得手順の例を示す図である。

【図10】ルートサーバにおける経路決定手順処理の例を示す図である。

【図11】経路決定手順処理の具体例を示す図である。

40

【図12】従来のAS内の経路情報管理を示す図である。

【発明を実施するための形態】

【0027】

次に、本発明を実施するための形態（以降「本実施形態」と称す）について、適宜図面を参照しながら詳細に説明する。

【0028】

本実施形態の概要

初めに、本実施形態の概要について、図1を用いて説明する。

図1に示すように、経路情報管理システム30は、AS1内に、経路情報の収集およびルーティングの管理を行う複数のルートサーバ（経路管理装置）10D、10Eおよび複

50

数のルータ 20 (20 H , 20 I , 20 J , 20 K , 20 L 等) を備えている。そして、各ルートサーバ 10 D , 10 E は、ルータ 20 と通信可能に接続される。また、各ルータ 20 間は、実線 (直線) で示されるように通信可能に接続されている。なお、図 1 は一例であり、ルートサーバ 10 D , 10 E の台数、ルータ 20 の台数はこれに限られない。また、ルータ 20 間の接続は、図 1 に示した状態に限られない。

【 0 0 2 9 】

図 1 の破線に示すように、AS 1 内のルータ 20 H , 20 I が、AS 2 からプレフィックス P 1 を含む経路情報を e B G P によって受信したものとする。そして、受信したプレフィックス P 1 を含む経路情報は、ルートサーバ 10 D に収集される。また、AS 1 内のルータ 20 J , 20 K が、AS 3 からプレフィックス P 9 を含む経路情報を e B G P によって受信したものとする。そして、受信したプレフィックス P 9 を含む経路情報は、ルートサーバ 10 E に収集される。すなわち、ルートサーバ 10 D , 10 E は、それぞれ、異なるプレフィックスに係る経路情報を収集する。なお、どのルートサーバ 10 D , 10 E が、どのプレフィックス P 1 , P 9 を含む経路情報を収集するかを決定するルートサーバ決定方法については、後記する。

【 0 0 3 0 】

このようにして、各ルートサーバ 10 D , 10 E は、プレフィックス P 1 , P 9 ごとの全ての経路情報を収集できるので、全ての経路情報に基づいて転送経路の決定を行うことが可能となる。つまり、前記した従来のルータに見られるような一部の経路情報しか取得できず、その一部の経路情報に基づいて転送経路の決定を行うというようなことは起きない。また、ルートサーバ 10 D , 10 E がプレフィックス P ごとの経路情報を分散して担当するので、AS 内のルータ台数が増加しても、ルートサーバ一台あたりの記憶容量の増加や処理負荷の増加を抑制することが可能となる。また、ルートサーバが複数備えられるため、経路情報の送受信が分散されて、回線の通信帯域の増加を抑制することも可能となる。つまり、AS 内のルータ数の増加に対して設備の大幅な更改を抑制することができる。

【 0 0 3 1 】

次に、各ルータ 20 は、パケットを受信したとき、ルートサーバ 10 D , 10 E に問い合わせ、ルートサーバ 10 D , 10 E によって決定された経路情報を取得する。各ルータ 20 が、どのルートサーバ 10 D , 10 E に問い合わせればよいかについては、後記するルートサーバ決定方法に基づいて、決定される。

例えば、図 1 に示すように、ルータ 20 L にプレフィックス P 1 を宛先とするパケットが到着した (図 1 の (1) 参照) とき、ルータ 20 L は、後記するルートサーバ決定方法に基づいて、プレフィックス P 1 に係る経路情報を取得しているルートサーバ 10 D へ、問い合わせを行う (図 1 の (2) 参照) 。そして、ルートサーバ 10 D は、後記する経路決定手順に基づいて、ルータ 20 L のネクストホップがルータ 20 H であるとする転送経路を決定し、その決定した結果 (経路決定情報) を返信する (図 1 の (3) 参照) 。そして、ルータ 20 L は、受信した経路決定情報に基づいて、パケットをルーティングする。このようにして、プレフィックス P 1 については、ルートサーバ 10 D が転送経路を決定し、問い合わせてきたルータ 20 にその転送経路に係る経路情報 (経路決定情報) を返信することによって、AS 1 内の経路情報を統一することができる。

【 0 0 3 2 】

また、ルートサーバ 10 D , 10 E が最適経路を決定する場合に、ルータの位置を考慮した後記する経路決定手順を用いる。さらに、その経路決定手順において、最適経路のみを選択するのではなく、次善の経路を選択するようにする。そして、最適経路および次善の経路に係る経路情報 (経路決定情報) を、i B G P の拡張機能を用いて、問い合わせてきたルータ 20 に返信する。このことによって、ルータ 20 は、ネクストホップのルータとの通信が切断した際に、新しい経路として次善の経路に直ちに切り替えることができる。

【 0 0 3 3 】

10

20

30

40

50

経路情報管理システム 30 の構成例

次に、本実施形態における経路情報管理システムの構成例を、図 2 を用いて説明する。図 2 に示すように、AS 1 内の経路情報管理システム 30 は、3 つのルートサーバ 10 (10 A, 10 B, 10 C) と 7 つのルータ 20 (20 A, 20 B, 20 C, 20 D, 20 E, 20 F, 20 G) とを含んで構成される。なお、図 2 に示したルータ 20 は、他の AS から経路情報を e B G P で受信する境界ルータのみであって、AS 1 内の他のルータの記載を省略している。また、図 2 では、実線矢印や破線矢印によって、情報の流れを示している。したがって、図 2 では、ルートサーバ 10 とルータ 20 との接続表記を省略しているが、実線矢印で示した箇所は i B G P による通信を可能とし、破線矢印で示した箇所は e B G P による通信を可能とする。なお、ルータ 20 は、他の AS のルータと直結されている。

【 0 0 3 4 】

図 2 に示すように、ルートサーバ 10 は、ルートサーバ 10 A, 10 B, 10 C ごとに、ルートサーバ 10 を識別する経路管理装置 I D (I D) = 5, 10, 15 が付与されている。なお、経路管理装置 I D は、AS 1 のネットワークを管理しているオペレータによって付与される。そして、各ルートサーバ 10 は、ルータ 20 に対して、自身がルートサーバ 10 であることおよびその経路管理装置 I D を、I G P を用いて広告する。

【 0 0 3 5 】

ルートサーバ 10 は、経路情報を、プレフィックスごとに分担して管理している。図 2 の例では、説明を簡単にするために、各ルートサーバ 10 A, 10 B, 10 C は、それぞれプレフィックス P 1 ~ P 5, P 6 ~ P 10, P 11 ~ P 15 を含む経路情報を管理するものとする。なお、本実施形態では、各ルートサーバ 10 によって管理されるプレフィックス P は、後記するルートサーバ決定方法を用いて決定される。

【 0 0 3 6 】

図 2 に示すように、ルータ 20 は、e B G P によって、外部のプレフィックス P を含む経路情報を受け取ると、I B G P を用いて、該当するプレフィックス P を管理するルートサーバ 10 に、その経路情報を広告する。なお、ルータ 20 は、後記するルートサーバ決定方法を用いて、広告すべきルートサーバ 10 を特定する。図 2 では、プレフィックス P 1, P 2 を含む経路情報はルートサーバ 10 A に広告され、プレフィックス P 6, P 9 を含む経路情報はルートサーバ 10 B に広告され、プレフィックス P 12 を含む経路情報はルートサーバ 10 C に広告される。これによって、各ルートサーバ 10 は、自身の担当するプレフィックス P に係る経路情報の全てを収集することが可能となる。

【 0 0 3 7 】

ルートサーバ決定方法

ここで、ルータ 20 におけるルートサーバ決定方法について、図 3 を用いて説明する。

図 3 は、印で示すノードとそれをカバーするハッシュ空間を表している。図 3 中の印の中の数字は、I D 番号 (経路管理装置 I D) を示している。したがって、図 2 に示すルートサーバ 10 A (I D = 5), 10 B (I D = 10), 10 C (I D = 15) はそれぞれ、図 3 の 5, 10, 15 にマッピングされる。

【 0 0 3 8 】

そして、まず、next(k) という関数を導入する。next(k) はハッシュ空間においてハッシュ空間を時計回り (つまり I D を大きくする方向) で次のノードの I D を返す。例えば、next(2)=5、next(6)=10、next(15)=15 となる。すなわち、next(k) は、ハッシュ空間のある値 k 以上の範囲において最小の値となる I D (経路管理装置 I D) を返す。

ここで、k は、プレフィックス P の値をハッシュ関数に代入して算出されたハッシュ値である。ただし、図 3 では、 $0 < k \leq 15$ とする。

【 0 0 3 9 】

そして、図 3 に示すように、ハッシュ空間におけるルートサーバ 10 A の分担範囲は 0 超え 5 以下、ルートサーバ 10 B の分担範囲は 5 超え 10 以下、ルートサーバ 10 C の分担範囲は 10 超え 15 以下となる。このようにして、ルートサーバ 10 の I D を予め決め

ておけば、プレフィックスPに関する経路情報を問い合わせるルートサーバ10を特定することが可能となる。

【0040】

なお、本実施形態では、ルートサーバ10を複数設ける構成としているので、ルートサーバ10の追加や削除を容易に行える必要がある。そこで、まず、ルートサーバ10の追加について、図4を用いて説明する。

図4に示すように、ID=8のルートサーバ10Fが新たに追加設定される場合について説明する。まず、ID=8のルートサーバ10Fは、next(k=8)として、時計回り(つまりIDを大きくする方向)で次のノードのID(図4ではルートサーバ10B(ID=10))に自分の存在をIGPを用いて広告する。そして、ルートサーバ10Fは、ルートサーバ10Bの分担範囲であった $5 < k < 10$ の中から、 $5 < k < 8$ の経路情報を分けてもらう。また、ルートサーバ10Bは、分担範囲を $8 < k < 10$ に変更する。なお、本実施形態では、各ルートサーバ10は、AS1内のルータ20に対して、IGPを用いて、自身がルートサーバ10であることおよび自身のIDを広告する。

【0041】

次に、ID=8のルートサーバ10Fが削除される場合について図4を用いて説明する。削除の場合を想定して、予め、各ルートサーバ10が、少なくとも、反時計回りで次のノードのIDの経路情報を記憶しているようにしておく。すなわち、ルートサーバ10Bが、ID=8のルートサーバ10Fの経路情報を記憶していれば、ID=8のルートサーバ10Fが故障等によって削除されても、 $5 < k < 8$ に係る経路情報が失われることはない。

【0042】

ルートサーバ10の機能

次に、本実施形態の経路情報管理システムに用いられるルートサーバ10の機能について、図5を用いて説明する。図5は、図2のルートサーバ10の機能ブロック図である。なお、図5は、本実施形態において特徴となる、ルートサーバ10の主要な機能のみを示し、iBGPやIGP等に係る機能については記載を省略している。

【0043】

図5に示すように、ルートサーバ10は、経路情報や経路決定情報等の各種データの入出力を司る入出力部11と、経路情報の収集や収集した経路情報に基づいて決定した経路決定情報を生成する処理部12と、記憶部13とを備える。

入出力部11は、ルータ20等とのデータ入出力を行う入出力インタフェースである。入出力部11におけるデータ入出力は、処理部12からの指示に基づいて行われる。

【0044】

処理部12は、ルートサーバ10が備える図示しないCPU(Central Processing Unit)と図示しないメインメモリによって構成されている。そして、このCPUがメインメモリにアプリケーションプログラムを展開して、それを実行することにより種々の演算処理を具現化する。なお、アプリケーションプログラムは、記憶部13に格納されている。

【0045】

また、処理部12は、少なくとも、経路情報収集部121と経路決定部122とを機能として備えている。

経路情報収集部121は、ルータ20から自身のルートサーバ10に広告されてくる経路情報を受信する。なお、ルータ20から広告されてくる経路情報は、予め、ルータ20によって広告先のルートサーバ10を選択されたものとなっている。そして、経路情報収集部121は、受信した経路情報を記憶部13の経路表131に記憶する。

経路決定部122は、ルータ20から、転送経路の問い合わせを受けたときに、記憶部13から経路表131に記憶されている経路情報とトポロジ情報132とを読み出して、予め決めておいた所定の経路決定手順に基づいて、問い合わせをしてきたルータ20を起点とする最適経路情報および次善の経路情報(経路決定情報)を生成する。そして、生成した経路決定情報を、問い合わせをしてきたルータ20に返信する。

10

20

30

40

50

【 0 0 4 6 】

記憶部 1 3 は、少なくとも、経路表 1 3 1 (図 6 参照) およびトポロジ情報 1 3 2 を記憶している。この経路表 1 3 1 は、図 6 に示すように、例えば、プレフィックスとパス属性情報とを含んでいる。パス属性情報は、Local pref、AS path、MED (Multi Exit Disc)、この経路情報を受信したルータ 2 0 を識別する情報および受信に使用されたプロトコル種別 (e B G P / i B G P)、ネクストホップに至る経路に割り当てられたコスト等である。そして、パス属性情報は、プレフィックス P ごとに関連付けられている。また、トポロジ情報 1 3 2 は、A S 内のルータ 2 0 間の接続状態を表す情報である。

【 0 0 4 7 】

パス属性情報のうち、Local prefは、同じA S 内のルータ 2 0 間で通知される属性である。同じプレフィックスに対して複数の経路が存在する場合、Local prefが大きい経路ほど優先される。

AS pathは、経由してきたA S を記録しておくためのものであり、A S を経由するごとにA S 番号が追加される。そして、AS pathに記録されたA S 番号の数が少ない方の経路が優先して用いられる。

MEDは、接続相手のA S との間に複数のピアが存在するときに、どのピアを使うかを相手のA S が指定するときに使われる。ルータ 2 0 は、e B G P によって同じ宛先の経路情報を複数受信すると、MED値が小さい方の経路が優先される。

【 0 0 4 8 】

図 6 に示す「受信したルータの識別情報および受信に使用されたプロトコル種別 (e B G P / i B G P) 」は、問い合わせをしてきたルータ 2 0 の識別情報と一致するか否かの判定に使われる。すなわち、問い合わせをしてきたルータ 2 0 がe B G P で受信した経路情報が、他のルータ 2 0 によって受信された経路情報よりも優先される。なお、識別情報に一致するものがない場合には、i B G P によって受信された経路情報が選択される。ここで、ルータの識別情報とは、例えば、問い合わせを行うときにパケットのヘッダ等に付けられる論理アドレスや物理アドレスや名称等である。

図 6 に示す「ネクストホップに至る経路に割り当てられたコスト」は、ネクストホップに至る経路について、経路を構成するリンクごとに割り当てられたコストを合計した値である。コストの合計が小さい経路が優先される。

【 0 0 4 9 】

ルータ 2 0 の機能

次に、本実施形態における経路情報管理システムに用いられるルータ 2 0 (境界ルータおよびA S 内に配置されるルータ) の機能について、図 7 を用いて説明する。図 7 は、図 2 のルータ 2 0 の機能ブロック図である。なお、図 7 は、本実施形態において特徴となる、ルータ 2 0 の主要な機能のみを示し、他のルータ 2 0 との間のデータ送受信やi B G P やI G P 等に係る機能については記載を省略している。

【 0 0 5 0 】

図 7 に示すように、ルータ 2 0 は、経路情報や経路決定情報等の各種データの入出力を司る入出力部 2 1 と、経路情報の収集や収集した経路情報の広告先の決定等を行う処理部 2 2 と、記憶部 2 3 とを備える。

入出力部 2 1 は、ルートサーバ 1 0 とのデータ入出力を行う入出力インタフェースである。入出力部 2 1 におけるデータ入出力は、処理部 2 2 からの指示に基づいて行われる。

【 0 0 5 1 】

処理部 2 2 は、ルータ 2 0 が備える図示しないC P U (Central Processing Unit) と図示しないメインメモリによって構成されている。そして、このC P U がメインメモリにアプリケーションプログラムを展開して、それを実行することにより種々の演算処理を具現化する。なお、アプリケーションプログラムは、記憶部 2 3 に格納されている。

【 0 0 5 2 】

また、処理部 2 2 は、少なくとも、経路情報収集部 2 2 1、ルートサーバ選択部 2 2 2、経路問い合わせ部 2 2 3、および経路決定情報受信部 2 2 4 を機能として備えている。

経路情報収集部 221 は、他の AS のルータ 20 から e B G P によって広告されてくる経路情報を受信する。そして、経路情報収集部 221 は、他の AS から受信した経路情報を記憶部 23 に記憶する。

ルートサーバ選択部 222 は、経路情報収集部 221 が他の AS のルータ 20 から e B G P によって広告されてくる経路情報を受信したときに、前記したルートサーバ決定方法に基づいて、この経路情報を広告すべきルートサーバ 10 を選択する。そして、ルートサーバ選択部 222 は、他の AS から受信した経路情報を、選択したルートサーバ 10 へ広告する。

【 0 0 5 3 】

経路問い合わせ部 223 は、パケットを受信したとき、そのパケットに含まれている宛先（プレフィックス P）を抽出して、前記したルートサーバ決定方法に基づいて、パケットの転送経路の問い合わせ情報を送信すべきルートサーバ 10 を選択する。そして、経路問い合わせ部 223 は、受信したパケットの転送経路を問い合わせるために、選択したルートサーバ 10 に、宛先（プレフィックス P）に係る情報を含む問い合わせ情報を送信する。なお、問い合わせ情報には、問い合わせ情報を送信するルータ 20 を識別する情報（例えば、IP アドレス等）が含まれている。

【 0 0 5 4 】

経路決定情報受信部 224 は、経路問い合わせ部 223 によって送信した問い合わせ情報に対する回答情報を、ルートサーバ 10 から受信する。回答情報とは、すなわち、経路決定情報（最適経路情報および次善の経路情報）である。経路決定情報受信部 224 は、受信した経路決定情報を記憶部 23 の経路表 231 に記憶する。ただし、経路決定情報受信部 224 は、公知の i B G P セッションを利用して経路決定情報を受信するので、特別な機能を追加される必要はない。

【 0 0 5 5 】

記憶部 23 は、少なくとも、ルートサーバ 10 によって決定された経路決定情報を格納する経路表 231 およびルートサーバ 10 の ID を格納するルートサーバ ID 表 232 を記憶する。なお、ルートサーバ 10 によって決定された経路決定情報がルーティングにおいて参照されるため、記憶部 23 は、広告されてきた経路情報を記憶する必要がなくなる。また、経路表 231 は、経路決定情報以外にも、パケットを転送する方路を決定するための情報を記憶している。また、ルートサーバ ID 表 232 は、ID とルートサーバ 10 とを関連付けた情報を格納しており、前記したルートサーバ決定方法において、ルータ 20 がルートサーバ 10 の ID を参照するとき用いられる。

【 0 0 5 6 】

ルートサーバ 10 の経路情報収集手順

次に、ルートサーバ 10 が、ルータ 20 から経路情報を収集する手順について、図 8 を用いて説明する。

図 8 に示すように、まず、ルータ 20（境界ルータ）が、他の AS から経路情報を取得する（ステップ S701）。次に、該ルータ 20 は、取得した経路情報に含まれるプレフィックス P を抽出して、前記したルートサーバ決定方法に基づいて、広告すべきルートサーバ 10 を選択する（ステップ S702）。そして、該ルータ 20 は、選択したルートサーバ 10 に経路情報を広告する（ステップ S703）。ルートサーバ 10 は、広告された経路情報を受信し、経路表 131（図 5 参照）に記憶する（ステップ S704）。

【 0 0 5 7 】

ルータ 20 の経路決定情報取得手順

次に、ルータ 20 が、ルートサーバ 10 から経路決定情報を取得する手順について、図 9 を用いて説明する。

図 9 に示すように、まず、ルータ 20 がパケットを受信する（ステップ S801）。次に、該ルータ 20 は、受信したパケットに含まれる宛先のプレフィックス P を抽出し、その抽出したプレフィックス P を用いて、前記したルートサーバ決定方法に基づいて、広告すべきルートサーバ 10 を選択する（ステップ S802）。そして、選択したルートサー

サーバ10に転送すべき経路を問い合わせ情報を送信する(ステップS803)。

【0058】

問い合わせを受けたルートサーバ10は、後記する経路決定手順処理を実行する(ステップS804)。そして、ルートサーバ10は、経路決定手順処理に基づいて決定した経路情報を含む経路決定情報を、問い合わせしてきたルータ20に送信する(ステップS805)。

該ルータ20は、経路決定情報を受信する(ステップS806)。そして、受信した経路決定情報に基づいて、パケットを転送する(ステップS807)。

【0059】

経路決定手順処理

次に、ルートサーバ10における経路決定手順処理の詳細について、図10を用いて説明する(適宜図2, 図5, 図6参照)。なお、経路決定手順処理は、図9のステップS804における処理のことである。

まず、ルートサーバ10は、宛先へ向かう経路が一つか否かを判定する(ステップS901)。宛先へ向かう経路が一つでない場合(ステップS901でNo)、ルートサーバ10は、Local Prefの値が最も大きい経路を選択する(ステップS902)。次に、ルートサーバ10は、ステップS902において選択した経路が一つか否かを判定する(ステップS903)。選択した経路が一つでない場合(ステップS903でNo)、ルートサーバ10は、AS pathの値が最も短い経路を選択する(ステップS904)。

【0060】

次に、ルートサーバ10は、ステップS904において選択した経路が一つか否かを判定する(ステップS905)。選択した経路が一つでない場合(ステップS905でNo)、ルートサーバ10は、MEDの値が最も小さい経路を選択する(ステップS906)。次に、ルートサーバ10は、ステップS906において選択した経路が一つか否かを判定する(ステップS907)。選択した経路が一つでない場合(ステップS907でNo)、ルートサーバ10は、eBGPで受信した経路を優先して選択し、該当するものがない場合は、iBGPで受信した経路を選択する(ステップS908)。

【0061】

つまり、ステップS908では、ルートサーバ10の経路決定部122は、経路表131(図6参照)を参照して、問い合わせ情報を送信してきたルータ20について「ルータの識別情報」に一致するものがあるか否かを検索し、「ルータの識別情報」に一致するものがあつた場合には、eBGPで受信した経路情報があるか否かを検索する。そして、eBGPで受信した経路情報がない場合には、iBGPによって受信された経路情報が検索される。すなわち、eBGPで受信した経路が、iBGPで受信された経路よりも優先される。なお、「ルータを識別情報」に一致するものがない場合には、他のルータ20によって受信された経路が選択される。

【0062】

次に、ルートサーバ10は、ステップS908において選択した経路が一つか否かを判定する(ステップS909)。選択した経路が一つでない場合(ステップS909でNo)、ルートサーバ10は、ネクストホップに至る経路に割り当てられたコストの最も小さい経路を選択する(ステップS910)。つまり、ステップS910では、経路決定部122が、トポロジ情報132を参照して、問い合わせをしてきたルータ20または図示しない境界ルータ以外のAS1内のルータのトポロジ上の位置を特定し、転送経路の候補となる経路に付与されるコストの最も小さい経路を転送経路として選択する。次に、ルートサーバ10は、ステップS910において選択した経路が一つか否かを判定する(ステップS911)。このステップS910によっても、トポロジ上のルータの位置が考慮されることになる。そして、選択した経路が一つでない場合(ステップS911でNo)、ルートサーバ10は、いずれか1つの経路を決定または複数の経路を決定する(ステップS912)。

【0063】

10

20

30

40

50

なお、ステップ S 9 0 1、ステップ S 9 0 3、ステップ S 9 0 5、ステップ S 9 0 7、ステップ S 9 0 9、ステップ S 9 1 1 において、選択した経路が 1 つとなった場合（前記各ステップで Yes）、処理はステップ S 9 1 3 へ進む。

そして、ルートサーバ 1 0 は、決定した経路情報を含む経路決定情報をルータ 2 0 に送信する（ステップ S 9 1 3）。なお、ステップ S 9 1 3 の処理は、図 9 に示すステップ S 8 0 5 の処理と同じである。

【 0 0 6 4 】

ただし、ステップ S 9 0 1、ステップ S 9 0 3、ステップ S 9 0 5、ステップ S 9 0 7、ステップ S 9 0 9、ステップ S 9 1 1 において、選択した経路が 1 つとなった場合（前記各ステップで Yes）、次善の経路を選択するためには、既に最適経路として選択された経路を除いて、ステップ S 9 0 1 からの経路決定手順処理を行う必要がある。

【 0 0 6 5 】

具体例

図 1 0 に示す経路決定手順処理の具体例について、図 1 0 および図 1 1 を用いて説明する。

まず、図 1 1 に示すように、経路情報管理システム 3 0 は、AS 1 内に、経路情報の収集およびルーティングの管理を行うルートサーバ 1 0 D および複数のルータ 2 0 を備えている。そして、ルートサーバ 1 0 D は、ルータ 2 0 と通信可能に接続される。また、各ルータ 2 0 間は、実線（直線）で示されるように通信可能に接続されている。そして、実線（直線）の近傍に記されている数値（1, 1 0, 1 0 0）は、その経路のコストを表している。なお、図 1 1 は一例であり、ルートサーバ 1 0 D 以外のルートサーバの記載を省略している。

【 0 0 6 6 】

いま、AS 1 内のルータ 2 0 H, 2 0 I, 2 0 M, 2 0 N が、プレフィックス P 1 を含む経路情報を e B G P によって受信したものとする。そして、受信したプレフィックス P 1 を含む経路情報は、ルートサーバ 1 0 D に収集される。なお、ルートサーバ 1 0 D に収集されたプレフィックス P 1 を含む経路情報が、いずれも Local pref および AS path が同じで、M E D が設定されていないものとする。

いま、ルータ 2 0 L に宛先がプレフィックス P 1 のパケットが到着したとき、ルータ 2 0 L は、プレフィックス P 1 に係る経路情報を有しているルートサーバ 1 0 D へ、問い合わせを行う。

【 0 0 6 7 】

次に、図 1 0 に移って、経路決定手順処理の経過について説明する。

まず、ステップ S 9 0 1 では、プレフィックス P 1 を宛先とする経路情報は、4 つ（ルータ 2 0 H, 2 0 I, 2 0 M, 2 0 N が受信したもの）ある。したがって、ステップ S 9 0 2 へ進む。ステップ S 9 0 2 では、Local pref の値は 4 つとも同じであるので、全て選択され、ステップ S 9 0 4 へ進む。ステップ S 9 0 4 では、AS path が 4 つとも同じであるので、全て選択され、ステップ S 9 0 6 へ進む。ステップ S 9 0 6 では、4 つとも M E D が設定されていないので、全て選択され、ステップ S 9 0 8 へ進む。

ステップ S 9 0 8 では、ルータ 2 0 L の識別情報を含む経路情報は存在しない（ルータ 2 0 L は、プレフィックス P 1 の経路情報を受信していない）ので、4 つとも選択され、ステップ S 9 1 0 に進む。

【 0 0 6 8 】

そして、ステップ S 9 1 0 では、ネクストホップに至る経路に割り当てられたコストの最も小さい経路が選択される。すなわち、図 1 1 に示される構成では、ルータ 2 0 L からルータ 2 0 H, 2 0 I までのコストの合計は 1 0 2（= 1 + 1 0 0 + 1）となり、ルータ 2 0 L からルータ 2 0 M, 2 0 N までのコストの合計は 2 0 2（= 1 + 1 0 0 + 1 0 0 + 1）となる。すなわち、ステップ S 9 1 0 において、ルータ 2 0 H, 2 0 I が選択され、ステップ S 9 1 2 に進む。そして、ステップ S 9 1 2 では、最適経路および次善の経路が選択される。次に、ステップ S 9 1 3 において、ステップ S 9 1 2 において選択された最

適経路および次善の経路を含む経路決定情報が、ルータ 20L に送信される。

【0069】

図 11 に戻って、ルートサーバ 10D から、経路決定情報として、[プレフィックス：P1、ネクストホップ：ルータ 20H、プレファレンス値：100、・・・]、および [プレフィックス：P1、ネクストホップ：ルータ 20I、プレファレンス値：50、・・・] が、ルータ 20L に送信される。ルータ 20L は、プレファレンス値を比較して、大きい方（プレファレンス値 100）を優先して（最適経路と判定して）、ルーティングを行う。なお、プレファレンス値 50 の経路は、次善の経路と判定され、ルータ 20H が利用できなくなったときに用いられる。

【0070】

以上、本実施形態の経路情報管理システム 30 によれば、AS 内に複数の経路管理装置を備えているため、AS 内のルータ数の増加に対して設備の大幅な更改を抑制でき、プレフィックスごとの経路情報を全て把握できるようにし、AS 内の経路情報を統一し、ネクストホップとの通信が切断した際に新しい経路に容易に切り替える技術を提供することが可能となる。

【0071】

なお、本実施形態は、これらに限定されるものではなく、その趣旨を変えない範囲で実施することができる。

例えば、本実施形態では、ルートサーバ 10 の識別番号（ID）は、AS 1 のネットワークを管理しているオペレータによって付与されるものとしたが、各ルートサーバ間で、分担範囲がほぼ均等あるいは既に付与されている ID の中央となる ID を割り振るというように、ID の決め方を予め定めたルールに基づいて特定のルートサーバが自動的に割り振るようにしても構わない。

【符号の説明】

【0072】

- 1 AS
- 10 ルートサーバ（経路管理装置）
- 12 処理部
- 13 記憶部
- 20 ルータ（境界ルータ、中継ルータ）
- 22 処理部
- 23 記憶部
- 30 経路情報管理システム
- 121 経路情報収集部
- 122 経路決定部
- 131 経路表
- 132 トポロジ情報
- 221 経路情報収集部
- 222 ルートサーバ選択部
- 223 経路問い合わせ部
- 224 経路決定情報受信部
- 231 経路表
- 232 ルートサーバ ID 表

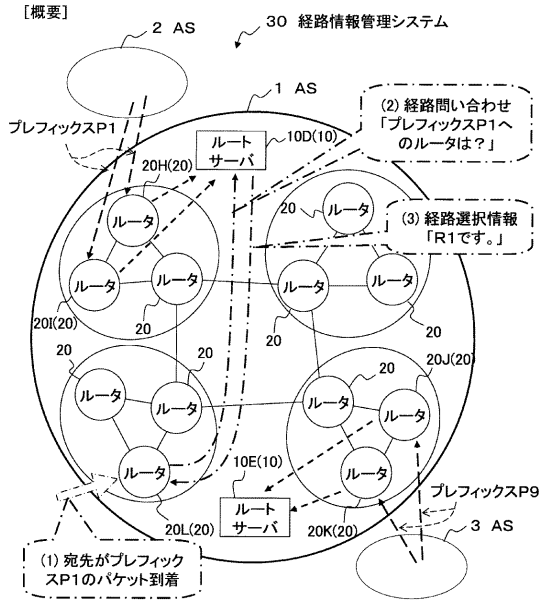
10

20

30

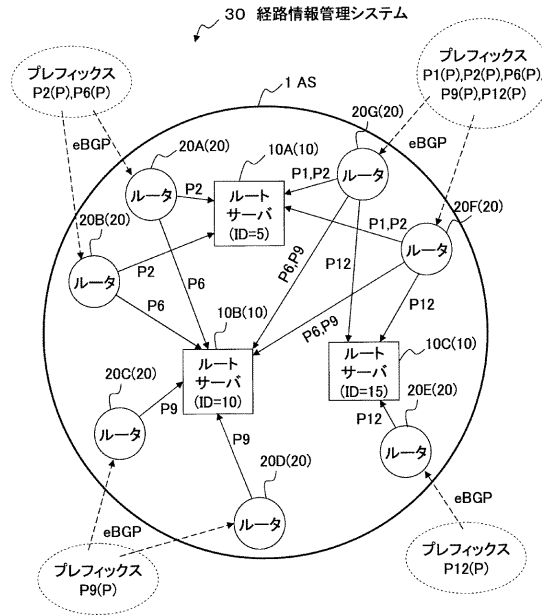
40

【図 1】



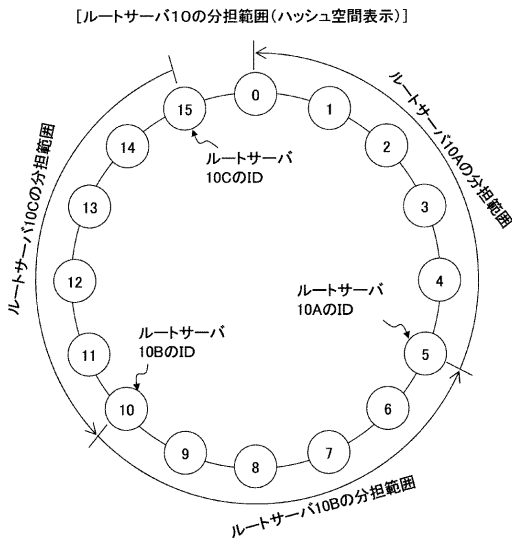
- AS1内に複数のルートサーバ10D, 10Eを備える。
- プレフィックスP1への経路情報はルートサーバ10Dが担当
- プレフィックスP9への経路情報はルートサーバ10Eが担当

【図 2】



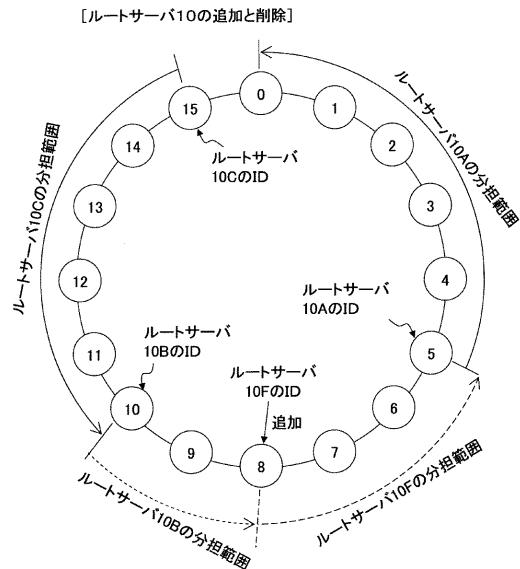
- 例えば、
- ルートサーバ10A: プレフィックスP1, P2を含む経路情報を管理
 - ルートサーバ10B: プレフィックスP6, P9を含む経路情報を管理
 - ルートサーバ10C: プレフィックスP12を含む経路情報を管理

【図 3】



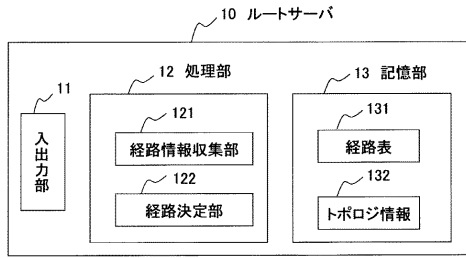
- 内の数字はルートサーバのID番号を表す。
- ルートサーバ10Aの分担範囲: 0超え～5以下
- ルートサーバ10Bの分担範囲: 5超え～10以下
- ルートサーバ10Cの分担範囲: 10超え～15以下

【図 4】

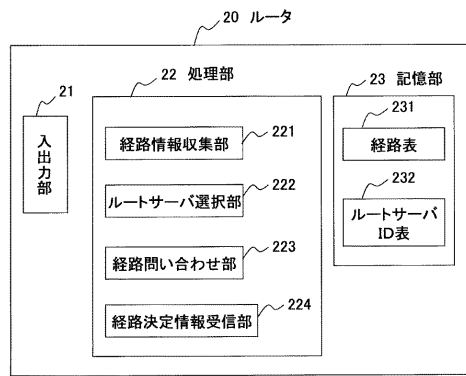


- 内の数字はルートサーバのID番号を表す。
- ルートサーバ10Aの分担範囲: 0超え～5以下
- ルートサーバ10Bの分担範囲: 5超え～10以下
- ルートサーバ10Cの分担範囲: 10超え～15以下
- ルートサーバ10Fの分担範囲: 5超え～8以下

【図5】



【図7】

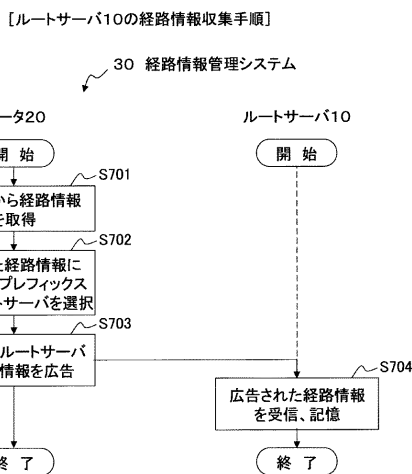


【図6】

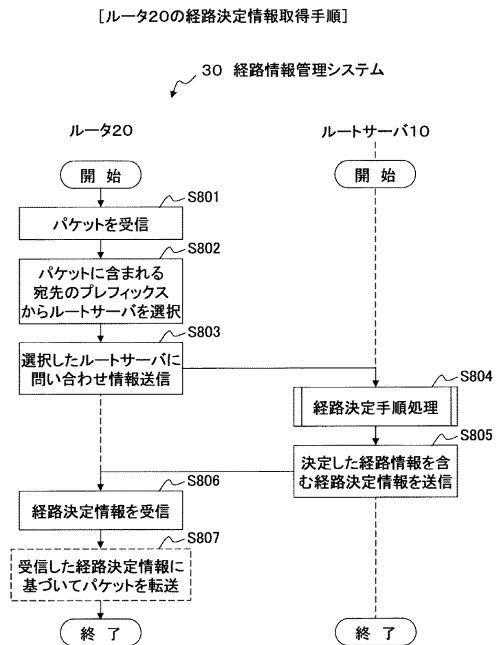
131,231 経路表

パス属性情報	
プロトコル種別	<ul style="list-style-type: none"> Local_pref AS-path MED (Multi_Exit_Disc) 受信したルータの識別情報および受信に使用されたプロトコル種別 (eBGP/iBGP) ネクストホップに至る経路に割り当てられたコスト
プロトコル種別	<ul style="list-style-type: none"> Local_pref AS-path MED (Multi_Exit_Disc) 受信したルータの識別情報および受信に使用されたプロトコル種別 (eBGP/iBGP) ネクストホップに至る経路に割り当てられたコスト
⋮	⋮

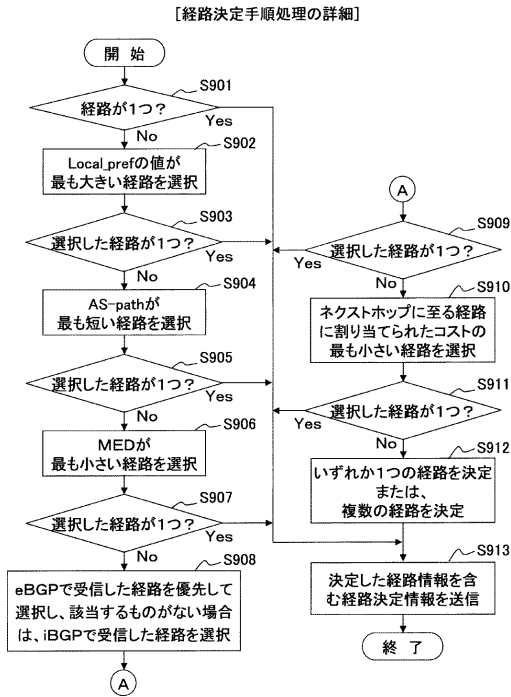
【図8】



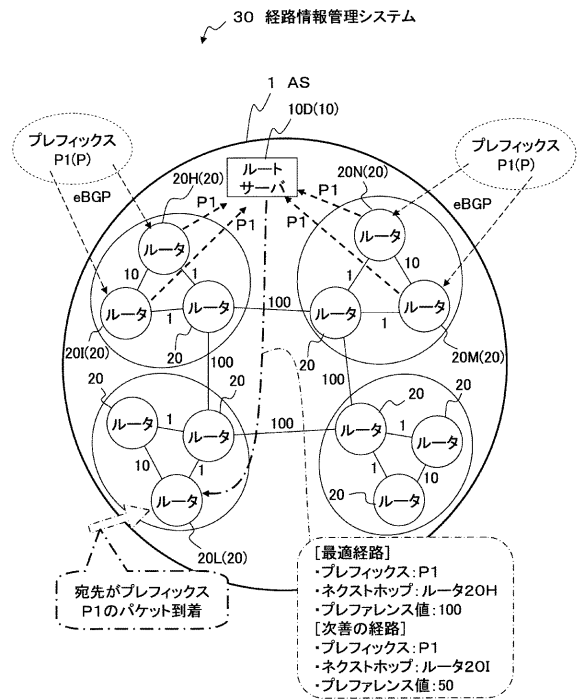
【図9】



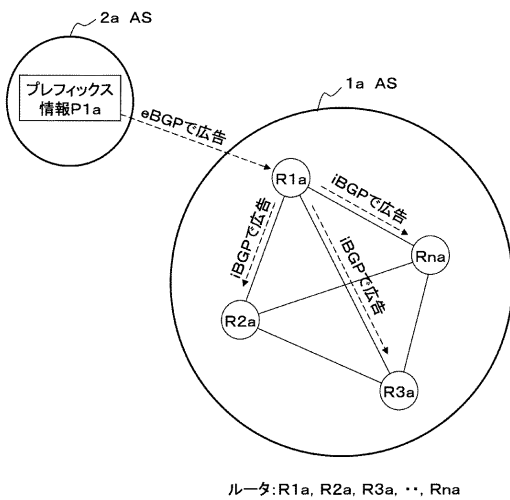
【図 10】



【図 11】



【図 12】



フロントページの続き

審査官 安藤 一道

(56)参考文献 特開2005-318409(JP,A)
特開2008-160385(JP,A)
特開2005-136865(JP,A)
特開2008-011330(JP,A)

(58)調査した分野(Int.Cl., DB名)
H04L 12/56