

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4809824号
(P4809824)

(45) 発行日 平成23年11月9日(2011.11.9)

(24) 登録日 平成23年8月26日(2011.8.26)

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

請求項の数 11 (全 19 頁)

(21) 出願番号	特願2007-320015 (P2007-320015)	(73) 特許権者	000004226
(22) 出願日	平成19年12月11日(2007.12.11)		日本電信電話株式会社
(65) 公開番号	特開2009-147457 (P2009-147457A)		東京都千代田区大手町二丁目3番1号
(43) 公開日	平成21年7月2日(2009.7.2)	(74) 代理人	100064414
審査請求日	平成22年1月14日(2010.1.14)		弁理士 磯野 道造
		(74) 代理人	100127720
			弁理士 大石 恵
		(72) 発明者	ベルサー・クリステル
			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	武田 知典
			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】 BGPセッションの設計方法、セッション設計装置およびBGPセッション設計プログラム

(57) 【特許請求の範囲】

【請求項1】

A S (Autonomous System) 内においてルータ同士を接続する B G P (Border Gateway Protocol) セッションの設計方法であって、

前記 A S 内のルータそれぞれにおける、パケットの宛先となる I P (Internet Protocol) アドレス空間を示すプレフィックスごとに、そのプレフィックスへパケットを送信するときのネクストホップを示した B G P ルートの集合である経路情報を記録した経路情報データベースと、前記 A S 内の各ルータ間における i B G P (internal B G P) セッションの設定状態を示す i B G P セッショントポロジ情報とを記憶する記憶部を備え、前記 B G P セッションを設定するルータを計算するセッション設計装置が、

前記経路情報データベースおよび前記 i B G P セッショントポロジ情報を参照して、前記経路情報に、前記ネクストホップが冗長化されていないプレフィックス P r e q を 1 以上持つルータの中から、 i B G P セッション追加対象ルータを特定するセッション追加対象ルータ特定処理を実行することを特徴とする B G P セッションの設計方法。

【請求項2】

前記セッション追加対象ルータ特定処理は、

前記セッション設計装置が、

前記経路情報データベースを参照して、前記ルータのうち、前記経路情報に、前記ネクストホップが冗長化されていないプレフィックス P r e q を 1 以上持つルータを特定し、この特定したルータの中から任意のルータを第 1 のルータとして選択するステップと、

前記 i B G P セッション トポロジ情報を参照して、前記選択した第 1 のルータとの間で前記 i B G P セッションを設定していない第 2 のルータを特定するステップと、

前記経路情報データベースを参照して、前記特定した第 2 のルータの中で、前記プレフィックス P r e q のうち、少なくともいずれか 1 つのプレフィックスの B G P ルートを前記経路情報に持ち、かつ、そのプレフィックスが、前記第 1 のルータに設定されているプレフィックス P r e q に対するネクストホップとは異なるネクストホップが設定されたルータを、前記第 1 のルータに対する i B G P セッション追加対象ルータとして特定するステップとを含むことを特徴とする請求項 1 に記載の B G P セッションの設計方法。

【請求項 3】

前記セッション設計装置は、

前記プレフィックス P r e q に対するネクストホップとは異なるネクストホップが設定されたルータが複数あったとき、これらのルータのうち、前記ネクストホップが異なるプレフィックスの数が最も多いルータを、前記 i B G P セッション追加対象ルータとすることを特徴とする請求項 2 に記載の B G P セッションの設計方法。

【請求項 4】

前記セッション設計装置は、

前記記憶部に、前記 A S 内の各ルータが e B G P (external B G P) セッションにより、外部の A S から受信している B G P ルートを示した e B G P ルート情報を記憶し、

前記 e B G P ルート情報を参照して、自身の A S 内で前記外部の A S との境界に設置される境界ルータのうち、当該プレフィックスの e B G P ルートを受信しているルータが 1 つのみであるプレフィックスがあったとき、当該プレフィックス以外のプレフィックスを、前記プレフィックス P r e q とすることを特徴とする請求項 2 または請求項 3 に記載の B G P セッションの設計方法。

【請求項 5】

前記 A S 内のルータのうち、前記境界ルータは、B G P 拡張として、前記 B G P ルートを広告するとき、自身のルータにおけるベストルートに加え、前記外部の A S から広告された e B G P ルートを、前記 i B G P セッションを設定した他のルータに広告する external best が設定されたルータであり、

前記セッション設計装置は、

前記第 2 のルータを特定するステップにおいて、前記 i B G P セッション トポロジ情報および前記 e B G P ルート情報を参照して、前記第 1 のルータとの間で前記 i B G P セッションを設定しておらず、かつ、前記プレフィックス P r e q を前記 e B G P ルートとして受信しているルータを前記第 2 のルータとして特定することを特徴とする請求項 4 に記載の B G P セッションの設計方法。

【請求項 6】

前記セッション設計装置は、

前記 e B G P ルート情報を参照して、自身の A S 内の前記境界ルータのうち、当該プレフィックスの e B G P ルートを受信しているルータが 1 つのみであるプレフィックスがあったとき、前記自身の A S 内のルータのうち、当該プレフィックスの e B G P ルートを受信しているルータ以外のルータの中から、新たに前記プレフィックスの e B G P ルートを受信するルータを選択することを特徴とする請求項 4 または請求項 5 に記載の B G P セッションの設計方法。

【請求項 7】

前記セッション設計装置は、

前記セッション追加対象ルータ特定処理の対象とするルータおよびプレフィックス P r e q を示した設定情報の入力を受け付け、前記設定情報に示されるルータおよびプレフィックス P r e q を対象に、前記セッション追加対象ルータ特定処理を実行することを特徴とする請求項 1 ないし請求項 6 のいずれか 1 項に記載の B G P セッションの設計方法。

【請求項 8】

前記セッション設計装置は、

10

20

30

40

50

前記セッション追加対象ルータ特定処理により特定された、前記第1のルータとこの第1のルータに対する前記iBGPセッション追加対象ルータそれぞれとの間に前記iBGPセッションを設定するよう指示するステップを実行することを特徴とする請求項2ないし請求項7のいずれか1項に記載のBGPセッションの設計方法。

【請求項9】

前記セッション設計装置は、

前記iBGPセッションを設定するよう指示したルータそれぞれから、前記経路情報を取得し、この取得した経路情報において前記プレフィックスPrefixに対し、ネクストホップが冗長化されていないルータがあったとき、このルータに対し再度iBGPセッションの設定を指示することを特徴とする請求項8に記載のBGPセッションの設計方法。

10

【請求項10】

AS (Autonomous System) 内においてBGP (Border Gateway Protocol) セッションの設定を行うルータを計算するセッション設計装置であって、

前記AS内のルータそれぞれにおける、パケットの宛先となるIP (Internet Protocol) アドレス空間を示すプレフィックスごとに、そのプレフィックスへパケットを送信するときのネクストホップを示したBGPルートの集合である経路情報を記録した経路情報データベースと、前記AS内の各ルータ間におけるiBGP (internal BGP) セッションの設定状態を示すiBGPセッショントポロジ情報とを記憶する記憶部と、

前記経路情報データベースおよび前記iBGPセッショントポロジ情報を参照して、前記経路情報に、前記ネクストホップが冗長化されていないプレフィックスPrefixを1以上持つルータの中から、iBGPセッション追加対象ルータを特定するセッション追加対象ルータ特定部とを備えることを特徴とするセッション設計装置。

20

【請求項11】

請求項1ないし請求項9のいずれか1項に記載のBGPセッションの設計方法を、コンピュータであるセッション設計装置に実行させることを特徴とするBGPセッション設計プログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、複数のAS (自律システム、Autonomous System) からなるネットワークにおける、BGP (Border Gateway Protocol) セッションの設計技術に関する。

30

【背景技術】

【0002】

インターネットは、複数のAS (またはドメイン) から構成されている。このAS間でBGPルート (経路情報) を交換する必要があるが、このときに利用される代表的プロトコルは、BGPである。BGPは、e (external) BGPとi (internal) BGPとがあり、eBGPは、隣接しているASとの間でBGPルートを交換するとき用いられるプロトコルである。iBGPは、AS内でBGPルートを配布 (交換) するとき用いられるプロトコルである。

【0003】

40

BGPルートは、パケットの宛先となるアドレスであるプレフィックス (IP (Internet Protocol) アドレス、サブネットアドレス等)、そのプレフィックス宛にパケットを転送するときの直近の境界ルータのアドレスであるネクストホップ、BGPに関する各種属性情報等を含んで構成される。なお、ここでの境界ルータとはドメイン (AS) の境界にあるルータである。

【0004】

ここで、AS間を繋ぐ経路の障害発生時に、通信を速やかに復旧させるためには、AS内の各ルータが、隣接するASへのルート (経路) を複数持つことが好ましい。つまり、AS内の各ルータが同じプレフィックスに対し、異なるルータをネクストホップとするBGPルート (ネクストホップの冗長性を持つルート情報) を複数保持するのが好ましい。

50

【 0 0 0 5 】

このことを図 6 および図 7 を用いて説明した図 6 および図 7 は本願発明の比較例となる技術を説明した図である。ここでは、図 6 に示すように、A S - A と、A S - B とを含むネットワークを例に説明する。この A S - A は、ルータ 2 0 (2 0 A , 2 0 B) を備え、A S - B は、ルータ 2 0 (2 0 C , 2 0 D , 2 0 E , 2 0 F) を備える。ルータ 2 0 A ~ 2 0 D は、境界ルータであり、ルータ 2 0 A , 2 0 C 間、ルータ 2 0 B , 2 0 D 間にはそれぞれ e B G P セッションが確立され、ルータ 2 0 C , 2 0 D はこの e B G P セッションによりプレフィックス P の B G P ルート (e B G P ルート) を受信し、経路情報 2 0 1 C , 2 0 1 D に記録する。また、ルータ 2 0 E , 2 0 F は R R (ルートリフレクタ) であり、他のルータ 2 0 から広告された B G P ルートを転送する。ここで、A S - B のルータ 2 0 C , 2 0 E 間、ルータ 2 0 E , 2 0 F 間、ルータ 2 0 F , 2 0 D 間には i B G P セッションが確立され、お互いに B G P ルートを交換しているものとする。

【 0 0 0 6 】

ここで、各ルータ 2 0 C ~ 2 0 F は i B G P セッションを用いて、B G P ルートを広告し、経路情報 2 0 1 (2 0 1 C , 2 0 1 D , 2 0 1 E , 2 0 1 F) に示すような経路情報を登録する。なお、経路情報 2 0 1 2 0 1 において、プレフィックスに「 > 」が付いたルートはベストルートである。ルータ 2 0 は同じプレフィックスの B G P ルートが複数あるとき、この「 > 」のついたベストルートを選択する。

【 0 0 0 7 】

このようなネットワーク構成において、ルータ 2 0 C , 2 0 A 間の経路に障害が発生すると、ルータ 2 0 C の経路情報 2 0 1 C には、プレフィックス P に対するネクストホップはルータ 2 0 A しか設定されていないので、ルータ 2 0 C からプレフィックス P へのパケットを A S - A へ転送することができない。つまり、ルータ 2 0 C 経由での通信が中断することになる。

【 0 0 0 8 】

ここで、図 7 の破線矢印で示すように、ルータ 2 0 C , 2 0 D 間に i B G P セッションがあれば、ルータ 2 0 C の経路情報 2 0 1 C に、プレフィックス P のネクストホップはルータ 2 0 B という B G P ルートが登録される。これにより、ルータ 2 0 A , 2 0 C 間の経路に障害が発生した場合でも、ルータ 2 0 C において経路の切り替えを行うことで、プレフィックス P 宛のパケットはルータ 2 0 C ルータ 2 0 D、ルータ 2 0 D ルータ 2 0 B という経路を辿って、プレフィックス P へ到達することができる。このように、各ルータ 2 0 にネクストホップの冗長性を持つ B G P ルートを配布することは、A S 間の経路の耐障害性を考えたときに非常に重要な問題である。ここで、ネクストホップの冗長性を高めるため、A S 内の各ルータ 2 0 間にフルメッシュで i B G P セッションを確立することも考えられるが、各ルータ 2 0 の i B G P ピア (i B G P セッションで隣接するノード) の数が膨大になり、現実的ではない。つまり、R R (ルートリフレクタ) を用いて、i B G P ピア の数を低減しつつ、A S 内のルータ 2 0 のネットワーク構成を考える方が現実的である。

【 0 0 0 9 】

また、ルータ 2 0 に external best と呼ばれる B G P 拡張を用いることで、境界ルータであるルータ 2 0 は、ベストルートに加え、外部 A S から受信した外部 A S への B G P ルートも i B G P セッションにより広告することができる (非特許文献 1)。これにより、A S 内の各ルータ 2 0 は、異なるネクストホップを持つ B G P ルートを受信しやすくなる。つまり、A S 内の各ルータ 2 0 の経路情報 2 0 1 においてネクストホップを冗長化しやすくなる。

【 非特許文献 1 】 Juniper Networks, Inc., " Configuring BGP routing advertising routes: bgp advertise best external to internal, " [online], [2007年10月27日検索]、インターネット、<URL: <http://www.juniper.net/techpubs/software/erx/junose71/swconfig bgp mpls/html/bgp config10.html>>

【 発明の開示 】

【発明が解決しようとする課題】

【0010】

しかし、前記した技術のいずれもが、AS内にどのようにiBGPセッションを設定すれば、各ルータがネクストホップの冗長性を持つようにBGPルートの広告をできるかについては考慮されていなかった。また、前記したexternal bestによりAS内の各ルータが、eBGPルートを受け取りやすくしたとしても、そのことにより、AS内の各ルータが確実にネクストホップの冗長性を持つBGPルートを持つことになるとは限らない。

【0011】

本発明は、前記した問題を解決し、AS内の各ルータが、ネクストホップの冗長性のあるBGPルートを持つようなBGPセッションの設計を行うことを目的とする。

10

【課題を解決するための手段】

【0012】

請求項1に記載の発明は、AS(AutonomousSystem)内においてルータ同士を接続するBGP(Border Gateway Protocol)セッションの設計方法であって、前記AS内のルータそれぞれにおける、パケットの宛先となるIP(Internet Protocol)アドレス空間を示すプレフィックスごとに、そのプレフィックスへパケットを送信するときのネクストホップを示したBGPルートの集合である経路情報を記録した経路情報データベースと、前記AS内の各ルータ間におけるiBGP(internalBGP)セッションの設定状態を示すiBGPセッショントポロジ情報とを記憶する記憶部を備え、前記BGPセッションを設定するルータを計算するセッション設計装置が、前記経路情報データベースおよび前記iBGPセッショントポロジ情報を参照して、前記経路情報に、前記ネクストホップが冗長化されていないプレフィックスPreqを1以上持つルータの中から、iBGPセッション追加対象ルータを特定するセッション追加対象ルータ特定処理を実行することを特徴とする。

20

【0013】

請求項2に記載の発明は、請求項1に記載のBGPセッションの設計方法において、前記セッション追加対象ルータ特定処理は、前記セッション設計装置が、前記経路情報データベースを参照して、前記ルータのうち、前記経路情報に、前記ネクストホップが冗長化されていないプレフィックスPreqを1以上持つルータを特定し、この特定したルータの中から任意のルータを第1のルータとして選択するステップと、前記iBGPセッショントポロジ情報を参照して、前記選択した第1のルータとの間で前記iBGPセッションを設定していない第2のルータを特定するステップと、前記経路情報データベースを参照して、前記特定した第2のルータの中で、前記プレフィックスPreqのうち、少なくともいずれか1つのプレフィックスのBGPルートを前記経路情報に持ち、かつ、そのプレフィックスが、前記第1のルータに設定されているプレフィックスPreqに対するネクストホップとは異なるネクストホップが設定されたルータを、前記第1のルータに対するiBGPセッション追加対象ルータとして特定するステップとを含むことを特徴とする。

30

【0014】

請求項10に記載の発明は、AS(AutonomousSystem)内においてBGP(Border Gateway Protocol)セッションの設定を行うルータを計算するセッション設計装置であって、前記AS内のルータそれぞれにおける、パケットの宛先となるIP(Internet Protocol)アドレス空間を示すプレフィックスごとに、そのプレフィックスへパケットを送信するときのネクストホップを示したBGPルートの集合である経路情報を記録した経路情報データベースと、前記AS内の各ルータ間におけるiBGP(internalBGP)セッションの設定状態を示すiBGPセッショントポロジ情報とを記憶する記憶部と、前記経路情報データベースおよび前記iBGPセッショントポロジ情報を参照して、前記経路情報に、前記ネクストホップが冗長化されていないプレフィックスPreqを1以上持つルータの中から、iBGPセッション追加対象ルータを特定するセッション追加対象ルータ特定部とを備えることを特徴とする。

40

【0015】

50

このようにすることで、セッション設計装置は、A S内の各ルータがネクストホップの冗長性のあるB G Pルートを受信できるようなi B G Pセッションの設計を行うことができる。なお、このように予めネクストホップの冗長性のあるB G Pルートを受信できるようなルータを、i B G Pセッションを追加するルータとすることで、A S内に無駄なi B G Pセッションを設定することなく、A S内の各ルータがネクストホップの冗長性のあるB G Pルートを受信できるようになる。つまり、A S内の各ルータのi B G Pセッション数を低減できる。

【0016】

請求項3に記載の発明は、請求項2に記載のB G Pセッションの設計方法において、前記セッション設計装置が、前記プレフィックスP r e qに対するネクストホップとは異なるネクストホップが設定されたルータが複数あったとき、これらのルータのうち、前記ネクストホップが異なるプレフィックスの数が最も多いルータを、前記i B G Pセッション追加対象ルータとすることを特徴とする。

【0017】

このようにすることで、セッション設計装置は、A S内の各ルータにおけるi B G Pセッション数をより低減することができる。つまり、セッション設計装置は、セッション追加対象ルータとして、i B G Pセッションを1本追加することで、ネクストホップが異なるB G Pルートをより多数受信できるルータを選択するので、ルータがネクストホップの冗長化のために設定するi B G Pセッション数を低減できる。

【0018】

請求項4に記載の発明は、請求項2または請求項3に記載のB G Pセッションの設計方法において、前記セッション設計装置が、前記記憶部に、前記A S内の各ルータがe B G P (external B G P)セッションにより、外部のA Sから受信しているB G Pルートを示したe B G Pルート情報を記憶し、前記e B G Pルート情報を参照して、自身のA S内で前記外部のA Sとの境界に設置される境界ルータのうち、当該プレフィックスのe B G Pルートを受信しているルータが1つのみであるプレフィックスがあったとき、当該プレフィックス以外のプレフィックスを、前記プレフィックスP r e qとすることを特徴とする。

【0019】

このようにすることで、セッション設計装置は、セッション追加対象ルータ特定処理を効率よく行うことができる。つまり、自身のA S内の境界ルータのうち、当該プレフィックスのe B G Pルートを受信しているルータが1つのみであるプレフィックスがあれば、A S内にどのようにi B G Pセッションを設定しても、そのプレフィックスに対しネクストホップは1つのみなので、冗長化できない。よって、セッション設計装置は、このプレフィックスを除外すれば、効率よくセッション追加対象ルータ特定処理を実行することができる。

【0020】

請求項5に記載の発明は、請求項4に記載のB G Pセッションの設計方法において、前記A S内のルータのうち、前記境界ルータは、B G P拡張として、前記B G Pルートを広告するとき、自身のルータにおけるベストルートに加え、前記外部のA Sから広告されたe B G Pルートを、前記i B G Pセッションを設定した他のルータに広告するexternal b estが設定されたルータであり、前記セッション設計装置は、前記第2のルータを特定するステップにおいて、前記i B G Pセッショントポロジ情報および前記e B G Pルート情報を参照して、前記第1のルータとの間で前記i B G Pセッションを設定しておらず、かつ、前記プレフィックスP r e qを前記e B G Pルートとして受信しているルータを前記第2のルータとして特定することを特徴とする。

【0021】

このようにすることで、A S内のルータは、境界ルータからi B G Pセッションにより、e B G Pルート(外部A Sへ接続するB G Pルート)を受信しやすくなる。つまり、このような設定がされた境界ルータとの間でi B G Pセッションを設定したルータは、この

境界ルータから、この境界ルータのベストルート以外の e B G P ルートも受信できるので、ネクストホップの冗長化を行いやすくなる。

【 0 0 2 2 】

請求項 6 に記載の発明は、請求項 4 または請求項 5 に記載の B G P セッションの設計方法において、前記セッション設計装置は、前記 e B G P ルート情報を参照して、自身の A S 内の前記境界ルータのうち、当該プレフィックスの e B G P ルートを受信しているルータが 1 つのみであるプレフィックスがあったとき、前記自身の A S 内のルータのうち、当該プレフィックスの e B G P ルートを受信しているルータ以外のルータの中から、新たに前記プレフィックスの e B G P ルートを受信するルータを選択することを特徴とする。

【 0 0 2 3 】

このようにすることで、セッション設計装置は、e B G P セッションの追加により、A S 内の各ルータにネクストホップの冗長性を持たせることができる。

【 0 0 2 4 】

請求項 7 に記載の発明は、請求項 1 ないし請求項 6 のいずれか 1 項に記載の B G P セッションの設計方法において、前記セッション設計装置が、前記セッション追加対象ルータ特定処理の対象とするルータおよびプレフィックス P r e q を示した設定情報の入力を受け付け、前記設定情報に示されるルータおよびプレフィックス P r e q を対象に、前記セッション追加対象ルータ特定処理を実行することを特徴とする。

【 0 0 2 5 】

このようにすることで、セッション設計装置は、このセッション設計装置の利用者が所望するプレフィックス、ルータを対象として、セッション追加対象ルータ特定処理を行うことができる。

【 0 0 2 6 】

請求項 8 に記載の発明は、請求項 2 ないし請求項 7 のいずれか 1 項に記載の B G P セッションの設計方法において、前記セッション設計装置が、前記セッション追加対象ルータ特定処理により特定された、前記第 1 のルータとこの第 1 のルータに対する前記 i B G P セッション追加対象ルータそれぞれとの間に前記 i B G P セッションを設定するよう指示するステップを実行することを特徴とする。

【 0 0 2 7 】

このようにすることで、セッション設計装置は、第 1 のルータとこの第 1 のルータに対する i B G P セッション追加対象ルータとの間に i B G P セッションを設定できる。

【 0 0 2 8 】

請求項 9 に記載の発明は、請求項 8 に記載の B G P セッションの設計方法において、前記セッション設計装置が、前記 i B G P セッションを設定するよう指示したルータそれぞれから、前記経路情報を取得し、この取得した経路情報において前記プレフィックス P r e q に対し、ネクストホップが冗長化されていないルータがあったとき、このルータに対し再度 i B G P セッションの設定を指示することを特徴とする。

【 0 0 2 9 】

このようにすることで、セッション設定装置は、特定したセッション追加対象ルータに対し、i B G P セッションを設定することができる。

【 0 0 3 0 】

請求項 1 1 に記載の発明は、請求項 1 ないし請求項 9 のいずれか 1 項に記載の B G P セッションの設計方法を、コンピュータであるセッション設定装置に実行させることを特徴とする B G P セッション設計プログラムである。

【 0 0 3 1 】

このようなプログラムによれば、請求項 1 ないし請求項 9 のいずれか 1 項に記載の B G P セッションの設計方法を、コンピュータであるセッション設定装置に実行させることができる。

【 発明の効果 】

【 0 0 3 2 】

10

20

30

40

50

本発明によれば、セッション設計装置は、AS内の各ルータに、ネクストホップの冗長性が高い経路情報を持たせるようなBGPセッションの設計を行うことができる。これにより、各ルータは障害発生時においても即座に経路の切り替えを行えることができるので、経路障害による通信の中断時間を短くすることができる。

【発明を実施するための最良の形態】

【0033】

以下、本発明を実施するための最良の形態（以下、実施の形態という）を、図面を参照しながら説明する。まず、図1を用いて本実施の形態のシステムの全体構成および動作概要を説明した図1は、本実施の形態のシステムの全体構成および動作概要を例示した図である。

10

【0034】

図1に示すように、システムは、ASB、ASC、ASDおよびASEを含んで構成される。AS-BにはAS-B内の各ルータ（経路制御装置）20（20C、20D、20E、20F）と、このルータ20C～20FのBGPセッションの制御を行うセッション設計装置100が設置される。このルータ20C～20Fと、セッション設計装置100とはネットワーク経路で接続される。なお、ここでのルータ20（20C、20D、20G～20I）はASBR (Autonomous System Boundary Router、外部のASと接続する境界ルータ)とする。ルータ20E、20Fは、ASBRであるが、RR（ルートリフレクタ）としての機能も備えるものとする。ルータ20は、通信プロトコルとしてBGP（iBGPおよびeBGP）を実装し、他のルータ20との間でBGPルートを交換する。

20

【0035】

ここで、ルータ20Cと、AS-Cのルータ20Hとの間でeBGPセッションが設定（確立）され、ルータ20Cは、このルータ20HからプレフィックスP、Q、RのBGPルート（eBGPルート）の広告を受けている。また、ルータ20Eと、AS-Dのルータ20Gとの間でeBGPセッションが確立され、ルータ20Eは、このルータ20GからプレフィックスPのBGPルート（eBGPルート）の広告を受けている。さらに、ルータ20Fと、AS-Eのルータ20Iとの間でeBGPセッションが確立され、ルータ20Fはこのルータ20IからプレフィックスRのBGPルート（eBGPルート）の広告を受けている。

【0036】

30

このようなネットワーク構成において、AS-Bのルータ20C、20E間、ルータ20E、20F間およびルータ20F、20D間にはそれぞれiBGPセッションが確立される。このiBGPセッションにより、ルータ20C～20DがBGPルートを交換することで、それぞれのルータ20の経路情報201は、経路情報201（201C、201D、201E、201F）に示すようになる。なお、前記したとおり、この経路情報201も、プレフィックスと、ネクストホップとを含んで構成される。プレフィックスとはパケットの宛先となるアドレスであり、ネクストホップは、そのプレフィックス宛にパケットを送信するための次のルータのアドレスである。さらに、プレフィックスに「>」が付いたルートはベストルートである。ルータ20は同じプレフィックスのBGPルートが複数あるとき、この「>」のついたベストルートを選択してパケットの経路制御を行う。

40

【0037】

ここでセッション設計装置100は、各ルータ20C～20Dの経路情報201（201C～201F）を参照して、経路情報201に、ネクストホップが冗長化されていないルータ20を探す。例えば、ルータ20Cの経路情報201Cは、プレフィックスP、Rについてネクストホップが冗長化されていない。また、ルータ20Dの経路情報201Dは、プレフィックスP、Q、Rすべてについてネクストホップが冗長化されていない。さらに、ルータ20Eの経路情報201Eは、プレフィックスPについてネクストホップが冗長化されていない。また、ルータ20Fの経路情報201Fは、プレフィックスP、Qについてネクストホップが冗長化されていない。

【0038】

50

ここで、セッション設計装置 100 は、各ルータ 20 が外部の AS から広告されている e B G P ルートの情報 (e B G P ルート情報) から、ネクストホップを冗長化できるプレフィックスを判断する。例えば、図 1 に示すネットワーク構成の場合、プレフィックス Q , R については、このプレフィックスを外部の AS から受信しているルータ 20 が複数あるが、プレフィックス P については、このプレフィックスを外部の AS から受信しているのはルータ 20 C のみなので、このプレフィックス P についてはネクストホップを冗長化できない。従って、セッション設計装置 100 は、この AS - B が e B G P ルートを受信しているプレフィックス P , Q , R から、プレフィックス P を除いたプレフィックス Q , R についてネクストホップの冗長化を行えばよいと判断する。

【 0039 】

この後、セッション設計装置 100 は、各ルータ 20 に対し、プレフィックス Q , R についてネクストホップの冗長化をするようにルータ 20 間に i B G P セッションを確立する。例えば、図 1 に示すネットワーク構成の場合、セッション設計装置 100 は、ルータ 20 C , 20 D 間と、ルータ 20 C , 20 F 間とに i B G P セッションを追加するよう指示する。このようにすることで、ルータ 20 C は、ルータ 20 F からプレフィックス P についてネクストホップはルータ 20 F という B G P ルート (符号 101 C) を経路情報 201 C に追加できる。また、ルータ 20 F はルータ 20 C から、プレフィックス Q についてネクストホップはルータ 20 C という B G P ルート (符号 101 F) を経路情報 201 F に追加できる。さらに、ルータ 20 D はルータ 20 C からプレフィックス R , Q についてネクストホップはルータ 20 C という B G P ルート (符号 101 D) を経路情報 201 D に追加できる。つまり、AS - B 内のルータ 20 はプレフィックス Q , R についてネクストホップの冗長化をすることができる。

【 0040 】

次に、このようなセッション設計装置 100 の構成を詳細に説明する。図 2 (a) は、図 1 のセッション設計装置の機能ブロック図である。図 2 (b) は、図 2 (a) の経路情報 DB を例示した図である。

【 0041 】

図 2 (a) に示すように、セッション設計装置 100 は、設定情報等の入出力を司る入出力部 11 と、i B G P セッションの追加対象となるルータ 20 を特定する処理部 12 と、処理部 12 が i B G P セッションの追加対象となるルータ 20 の特定等を行うときに参照するデータを記憶する記憶部 13 と、各ルータ 20 から経路情報 201 や e B G P ルート、i B G P トポロジ等を受信する通信部 14 とを備える。

【 0042 】

入出力部 11 は、外部の入力装置や出力装置とのデータ入出力を司る入出力インタフェースから構成される。また、処理部 12 は、このセッション設計装置 100 が備える CPU (Central Processing Unit) によるプログラム実行処理や、専用回路等により実現される。さらに、記憶部 13 は、RAM (Random Access Memory)、ROM (Read Only Memory)、HDD (Hard Disk Drive)、フラッシュメモリ等の記憶媒体から構成される。なお、セッション設計装置 100 をプログラム実行処理により実現する場合、記憶部 13 には、このセッション設計装置 100 の機能を実現するためのプログラムが記録される。通信部 14 は、ルータ 20 と通信可能なネットワークインタフェースカードから構成される。

【 0043 】

< 入出力部 >

入出力部 11 は、外部の入力装置 (例えば、キーボードやマウス) から、このセッション設計装置 100 の設定情報 134 (後記) 等の入力を受け付ける。この設定情報 134 は、例えば、このセッション設計装置 100 においてネクストホップの冗長性を確保するルータ 20 の識別情報や、ネクストホップの冗長性を確保するプレフィックス等を示した情報である。このような設定情報 134 は、このセッション設計装置 100 が一部のルータ 20 のみに限定して前記したネクストホップの冗長性を確保する i B G P セッションの

10

20

30

40

50

確立を指示したり、このセッション設計装置 100 の属する AS が広告を受けている e B G P ルートのうち、一部のプレフィックスのみに限定してネクストホップの冗長性を確保するようにしたりするとき大変便利である。なお、このような設定を行わないときは、設定情報 134 の入力を受け付けないようにしてもよい。

【0044】

< 処理部 >

処理部 12 は、セッション設計装置 100 全体の制御を司り、トポロジ情報収集部 121 と、e B G P ルート収集部 122 と、経路情報収集部 123 と、セッション制御部 124 と、セッション追加対象ルータ特定部 125 とを含んで構成される。

【0045】

トポロジ情報収集部 121 は、AS 内のルータ 20 のトポロジや、このルータ 20 間における i B G P セッションの確立（設定）状態等を示した情報（i B G P セッション情報）を取得する。そして、この情報を i B G P セッショントポロジ情報 131 として記憶部 13 に記録する。例えば、図 1 に示すネットワークの場合、ルータ 20 C, 20 E 間、ルータ 20 E, 20 F 間およびルータ 20 F, 20 D 間に i B G P セッションが確立されているという情報を i B G P セッショントポロジ情報 131 として記録する。この情報は、後記するセッション追加対象ルータ特定部 125 がどのルータ 20 間に i B G P セッションを追加すればよいかを判断するときに参照される。

【0046】

e B G P ルート収集部 122 は、AS 内の境界ルータであるルータ 20 から、e B G P ルートを収集する。この e B G P ルートは、AS の境界に設置される境界ルータ（AS B R）が外部の AS から受信する B G P ルートである。そして、この e B G P ルートを e B G P ルート情報 132 として記録する。例えば、図 1 に示すネットワークの場合、e B G P ルート収集部 122 は、ルータ 20 C は AS - C のルータ 20 H からプレフィックス P, Q, R の e B G P ルートの広告を受けていることや、ルータ 20 E は、AS - D のルータ 20 G からプレフィックス P の e B G P ルートの広告を受けていることや、ルータ 20 F は、AS - E のルータ 20 I からプレフィックス R の e B G P ルートの広告を受けていること等を e B G P ルート情報 132 に記録する。この e B G P ルート情報 132 は、セッション追加対象ルータ特定部 125 がネクストホップの冗長性を確保できるプレフィックスを判断するときに参照される。

【0047】

経路情報収集部 123 は、AS 内のルータ 20 から経路情報 201 を収集し、経路情報 DB（データベース）133 に記録する。例えば、図 1 に示すネットワークの場合、ルータ 20（20 C ~ 20 F）から、経路情報 201（201 C ~ 201 F）を収集し、経路情報 DB 133 に記録する。

【0048】

セッション制御部 124 は、後記するセッション追加対象ルータ特定部 125 において特定されたルータ 20 間に i B G P セッションの確立を指示する。つまり、セッション制御部 124 は、i B G P セッションの確立を指示するセッション制御情報を通信部 14 経由で出力する。

【0049】

セッション追加対象ルータ特定部 125 は、i B G P セッショントポロジ情報 131、e B G P ルート情報 132、経路情報 DB 133 および設定情報 134 を参照して、AS 内のルータ 20 のうち、どのルータ 20 間に i B G P セッションを追加すれば、AS 内のルータ 20 それぞれにおいてネクストホップが冗長化された経路情報 201 を保持できるかを判断する。このセッション追加対象ルータ特定部 125 の処理手順の詳細はフローチャートを用いて後記する。

【0050】

< 記憶部 >

記憶部 13 は、i B G P セッショントポロジ情報 131 と、e B G P ルート情報 132

10

20

30

40

50

と、経路情報DB133と、設定情報134とを記憶する。

【0051】

iBGPセッショントポロジ情報131は、前記したとおり、このセッション設計装置100がセッション制御を行うルータ20の識別情報、AS内におけるルータ20のトポロジ、iBGPセッションの確立状態等を示した情報である。

【0052】

eBGPルート情報132は、AS内の境界ルータがeBGPセッションにより、外部のASから受信しているeBGPルートを示した情報である。

【0053】

経路情報DB133は、AS内の各ルータ20における経路情報201を記憶する。例えば、図2(b)に示すように、経路情報DB133は、図1のルータ20の識別情報(例えば、「ルータ20C,20D,20E,20F」)ごとに、そのルータ20における経路情報201が記録される。 10

【0054】

図2(a)の説明に戻る。設定情報134は、セッション追加対象ルータ特定部125が処理対象とするルータ20の識別情報や、プレフィックス等を示した情報である。

【0055】

<通信部>

通信部14は、所定のプロトコルを用いて、図1の各ルータ20から経路情報201や、eBGPルート情報132、iBGPセッショントポロジ情報131等の元となる各種情報を受信する。受信した情報は処理部12へ出力する。また、この通信部14は、セッション制御部124から出力されるセッション制御情報を、各ルータ20へ出力する。 20

【0056】

なお、iBGPセッショントポロジ情報131およびeBGPルート情報132は、このセッション設計装置100の利用者(オペレータ)が入力したものでよい。また、経路情報DB133に記憶される各経路情報201は、iBGPセッショントポロジ情報131やeBGPルート情報132に基づいて、セッション設計装置100が計算した情報でもよい。

【0057】

次に、図1および図2を参照しつつ、図3および図4を用いて、図2(a)のセッション設計装置100の処理手順を説明した図3および図4は、図2(a)のセッション設計装置の処理手順を示した図である。 30

【0058】

まず、図2(a)のセッション設計装置100のトポロジ情報収集部121は、AS内の各ルータ20からiBGPセッショントポロジ情報131の元となる情報(iBGPセッション情報等)を収集する(S101)。そして、トポロジ情報収集部121は、収集した情報を、記憶部13のiBGPセッショントポロジ情報131として記録する。また、eBGPルート収集部122は、AS内の各ルータ20(境界ルータ)からeBGPルートを収集する(S102)。そして、eBGPルート収集部122は、収集したeBGPルートを記憶部13のeBGPルート情報132として記録する。さらに、経路情報収集部123は、AS内のルータ20から経路情報201を収集する(S103)。そして、経路情報収集部123は、経路情報DB(データベース)133に記録する。 40

【0059】

この後、セッション設計装置100は入出力部11経由で、ネクストホップの冗長性を必要とするプレフィックスの集合Preqと、冗長性を必要とするルータ20の集合Reqとを示した設定情報134の入力を受け付ける(S104)。入力された設定情報134は、記憶部13に記録される。この設定情報134は、例えば、ネクストホップの冗長性を必要とするプレフィックスPreqはプレフィックスP,Q,Rであり、冗長性を必要とするルータReqはルータ20C,20D,20E,20Fであるという情報である。

。

【 0 0 6 0 】

次に、セッション追加対象ルータ特定部 1 2 5 は、e B G P ルート情報 1 3 2 を参照して、ネクストホップの冗長性を確保できないプレフィックスを特定し、前記した P r e q から取り除く (S 1 0 5)。つまり、セッション追加対象ルータ特定部 1 2 5 は、e B G P ルート情報 1 3 2 を参照して、ネクストホップの冗長性を確保できないプレフィックスがあるか否かを判断し、ネクストホップの冗長性を確保できないプレフィックスがあれば P r e q から取り除く。ここで、ネクストホップの冗長性を確保できないプレフィックスとは、どのようにルータ 2 0 間に i B G P セッションを追加しても、当該プレフィックス宛に異なる複数の経路を確保できないプレフィックスである。例えば、図 1 に示したネットワーク構成の場合、プレフィックス P を e B G P ルートとして受信しているルータはルータ 2 0 C のみである。一方、プレフィックス Q については、このプレフィックスを e B G P ルートとして受信しているルータ 2 0 は複数ある (ルータ 2 0 E およびルータ 2 0 C)。また、プレフィックス R についても、このプレフィックスを e B G P ルートとして受信しているルータ 2 0 は複数ある (ルータ 2 0 C およびルータ 2 0 F)。つまり、プレフィックス Q, R については、ネクストホップの冗長性を確保できるが、プレフィックス P についてネクストホップの冗長性を確保できない。従って、セッション追加対象ルータ特定部 1 2 5 は、プレフィックス P, Q, R から、プレフィックス P を除いたプレフィックス Q, R を、P r e q とする。

10

【 0 0 6 1 】

次に、セッション追加対象ルータ特定部 1 2 5 は、経路情報 D 1 3 3 を参照して、R r e q の中から、冗長性が必要な P r e q について複数のルートを持たないルータ 2 0 の集合 R を特定する (S 1 0 6)。つまり、セッション追加対象ルータ特定部 1 2 5 は、複数のルートを持たないプレフィックス P r e q を持つがこのプレフィックス 1 以上持つルータ 2 0 を特定する。例えば、P r e q は、プレフィックス Q, R であり、R r e q は、ルータ 2 0 C ~ 2 0 F であるとき、経路情報 D B 1 3 3 におけるルータ 2 0 C ~ 2 0 F の経路情報 2 0 1 (2 0 1 C ~ 2 0 1 F) において、プレフィックス Q, R に関する B G P ルートを複数持たないルータ 2 0 は、ルータ 2 0 C, 2 0 D, 2 0 F であるので、セッション追加対象ルータ特定部 1 2 5 はこのルータ 2 0 C, 2 0 D, 2 0 F を集合 R として特定する。

20

【 0 0 6 2 】

次に、図 4 の説明に移る。セッション追加対象ルータ特定部 1 2 5 は、S 1 0 6 で特定した R (集合 R) の中から、ルータ 2 0 を 1 つ選択する (S 2 0 1)。そして、セッション追加対象ルータ特定部 1 2 5 は、この選択したルータ 2 0 に対する i B G P セッション追加対象ルータを、以下のようにして選択する。

30

【 0 0 6 3 】

すなわち、セッション追加対象ルータ特定部 1 2 5 は、i B G P セッショントポロジ情報 1 3 1 を参照して、S 2 0 1 で選択したルータ 2 0 との間で i B G P セッションを確立していないルータ 2 0 の集合を特定する (S 2 0 2)。例えば、図 1 に示すネットワーク構成において、セッション追加対象ルータ特定部 1 2 5 がルータ 2 0 D を選択したとき、このルータ 2 0 D と i B G P セッションを確立していないルータ 2 0 C, 2 0 E を特定する。

40

【 0 0 6 4 】

そして、セッション追加対象ルータ特定部 1 2 5 は、経路情報 D B 1 3 3 を参照して、S 2 0 2 で特定したルータ 2 0 の集合の中から、冗長性が必要な P r e q を B G P ルートとして受信しているルータ 2 0 の集合を特定する (S 2 0 3)。例えば、セッション追加対象ルータ特定部 1 2 5 が、S 2 0 2 でルータ 2 0 D を選択したとき、このルータ 2 0 D の経路情報 2 0 1 D において、ネクストホップの冗長性が確保されていないプレフィックスはプレフィックス R, Q である。ここで、まだ i B G P セッションを確立していないルータ 2 0 C, 2 0 E は、経路情報 2 0 1 C, 2 0 1 E に示すように、このプレフィックス Q, R を B G P ルートとして受信しているので、セッション追加対象ルータ特定部 1 2 5 は

50

このルータ 20C, 20E を、冗長性が必要な Preq を BGP ルートとして受信しているルータ 20 の集合として特定する。

【0065】

次に、セッション追加対象ルータ特定部 125 は、経路情報 DB 133 を参照して、S203 で特定したルータ 20 の中で、S201 で選択したルータ 20 と Preq のネクストホップが異なる BGP ルートを持つルータ 20 の集合 C を特定する (S204)。つまり、セッション追加対象ルータ特定部 125 は、iBGP セッションの確立により、自身の経路情報 201 に同じプレフィックスに対し異なる複数の BGP ルートが設定できるようなルータ 20 を特定する。

【0066】

なお、ここでいう「ネクストホップが異なる」とは、同じプレフィックスの BGP ルートのうち、このプレフィックスへパケットを送信するとき、自 AS 内において出口となる境界ルータが異なる BGP ルートのことをいう。従って、例えば、経路情報 201D におけるプレフィックス P に対するネクストホップはルータ 20C であり、経路情報 201C におけるプレフィックス P に対するネクストホップは、AS-C のルータ 20H となり、ネクストホップは異なるように見えるが、このルータ 20D から転送されたプレフィックス P 宛のパケットも、ルータ 20C またルータ 20C を最終出口として、プレフィックス P へ転送されるので、これらのネクストホップは実質的に同じである。

【0067】

例えば、S203 で特定したルータ 20 が、図 1 に示すルータ 20D であるとき、セッション追加対象ルータ特定部 125 は、ルータ 20C, 20E のいずれと iBGP セッションを確立しても、プレフィックス R, Q についてネクストホップが異なるものとなる BGP ルートを受信できると判断する。よって、このルータ 20C, 20E を集合 C として特定する。つまり、ルータ 20D は、ルータ 20C と iBGP セッションを確立し、ルータ 20C からプレフィックス Q, R について BGP ルートを受信すれば、プレフィックス Q についてネクストホップがルータ 20C という BGP ルートと、プレフィックス Q についてネクストホップがルータ 20C という BGP ルートが経路情報 201D に追加されることになる。従って、セッション追加対象ルータ特定部 125 は、このルータ 20D を集合 C に含める。

【0068】

また、ルータ 20D は、ルータ 20E と iBGP セッションを確立し、ルータ 20E からプレフィックス Q, R について BGP ルートの広告を受ければ、経路情報 201D にプレフィックス R についてネクストホップがルータ 20C という BGP ルートが追加されることになる。従って、セッション追加対象ルータ特定部 125 は、このルータ 20E を集合 C に含める。なお、ルータ 20 は、iBGP セッションを確立した相手方のルータ 20 に対し、ベストルートのみを広告する。従って、ルータ 20E は、ルータ 20D に対し、プレフィックス Q については、ネクストホップがルータ 20E という BGP ルートを広告することになる。しかし、ルータ 20D の経路情報 201D には既にプレフィックス Q について、ネクストホップがルータ 20E という BGP ルートは経路情報 201D に登録済みであるので、ルータ 20D はルータ 20E と iBGP セッションを確立しても、プレフィックス Q についてはネクストホップの冗長性を確保できないことになる。

【0069】

そして、セッション追加対象ルータ特定部 125 は、C (集合 C) に含まれるルータ 20 が複数あれば (S205 の Yes)、この中から、ネクストホップが異なるプレフィックス (BGP ルート) の数が最も多いルータ 20 を選択する (S206)。例えば、セッション追加対象ルータ特定部 125 がルータ 20D に対し、集合 C としてルータ 20C, 20E を特定したとき、ルータ 20C へ iBGP セッションを確立すると、プレフィックス Q, R の両方についてネクストホップの冗長性を確保できる。一方、ルータ 20E へ iBGP セッションを確立すると、ネクストホップの冗長性を確保することができるのはプレフィックス R のみである。従って、セッション追加対象ルータ特定部 125 は、これら

10

20

30

40

50

のルータ20のうち、ルータ20Cの方を選択する。そして、S208へ進む。なお、セッション追加対象ルータ特定部125は、ネクストホップが異なるプレフィックス(BGPルート)の数が同じルータ20が複数あったときには、これらのルータ20から所定のルールにより、1つのルータ20を選択するようにしてもよい。この所定のルールとは、例えば、このルータ20へのホップ数がより少ないルータ20を選択するルールや、ランダムに選択するルール等が考えられる。なお、このようにセッション追加対象ルータ特定部125が、iBGPセッションの確立を行う相手方のルータ20として、ネクストホップが異なるプレフィックス(BGPルート)の数が最も多いルータ20を選択するようにすることで、少ないiBGPセッション数で、より多くのBGPルートについてネクストホップの冗長化を行うことができる。つまり、AS内の各ルータ20は少ないiBGPセッション数で効率よくネクストホップを冗長化できるので、このiBGPセッションを多数保持することによるルータ20の処理負荷を最低限に抑えることができる。

【0070】

一方、S205において、セッション追加対象ルータ特定部125は、C(集合C)に含まれるルータ20が複数でなければ(S205のNo)、この特定したルータ20を選択し(S207)、S208へ進む。

【0071】

S206またはS207の後、セッション制御部124は、S206またはS207で選択したルータ20と、S201で選択したルータ20との間のiBGPセッションの確立を指示するセッション制御情報を出力する(S208)。例えば、セッション制御部124は、ルータ20D、20CへiBGPセッションの確立を指示するセッション制御情報を出力する。これにより、ルータ20Cから、ルータ20DへプレフィックスR、QのBGPルートが広告されるので、ルータ20Dはネクストホップの冗長性を確保することができる。なお、このようにしてiBGPセッションが追加されたときには、iBGPセッショントポロジ情報131には、この追加されたiBGPセッションの情報が追加される。

【0072】

この後、セッション追加対象ルータ特定部125は、S106で特定したルータの集合Rのうち未処理のルータ20があれば(S209のYes)、S201へ戻る。一方、S106で特定したルータ20の集合Rすべてについて処理を実行済みであれば(S209のNo)、処理を終了する。

【0073】

なお、前記したセッション制御部124によるiBGPセッションの確立のためのセッション制御情報の出力は、セッション追加対象ルータ特定部125において、iBGPセッションを追加対象となるルータ20をすべて特定した後でもよい。つまり、セッション設計装置100は、複数のルータ20に対しまとめてiBGPセッションの確立を指示するようにしてもよい。

【0074】

また、このセッション設計装置100は、各ルータ20にiBGPセッションの確立指示後、経路情報DB133を参照して、まだネクストホップの冗長性が確保されていない経路情報201を持つルータ20があれば、再度そのルータ20にiBGPセッションの確立を指示するようにしてもよい。

【0075】

さらに、セッション設計装置100は、どのようにルータ20間にiBGPセッションを追加しても、当該プレフィックス宛に異なる複数の経路を確保できないプレフィックスがあったとき(つまり、自身のAS内の境界ルータのうち、当該プレフィックスのeBGPルートを受信しているルータ20が1つのみであるプレフィックスがあったとき)、所定のルータ20間にeBGPセッションを追加することで、ネクストホップの冗長性を確保するようにしてもよい。例えば、セッション設計装置100は、図1に示すルータ20Dと、AS-Cの境界ルータ(例えば、ルータ20H)との間にeBGPセッションを追

加するよう指示してもよい。このようにすることで、ルータ20DもプレフィックスPへのeBGPルートを追加できるので、このルータ20DからBGPルートを広告することで、ネクストホップの冗長性を確保することができる。

【0076】

なお、AS内の各ルータ20はexternal bestというBGP拡張を設定するようにしてもよい。前記したとおり、このexternal bestは、境界ルータが、iBGPセッションを確立したルータ20に対し、ベストルートに加えて、外部ASから受信したeBGPルートも広告する設定である。

【0077】

図5を用いて境界ルータ(ASBR)にexternal bestを設定した場合のBGPルートの広告について説明した図5は、本実施の形態の境界ルータ(ASBR)にexternal bestを設定した場合のBGPルートの広告について説明した図である。ここでも、AS-Bは、ルータ20C~20Fを備え、ルータ20C, 20Dはそれぞれ、AS-Aのルータ20A, 20Bとの間でeBGPセッションを確立し、プレフィックスPのeBGPルートを受信するものとする。また、このようなルータ20C~20Fはそれぞれ、iBGPセッションが確立されたルータ20間でBGPルートの広告を行うことで、経路情報201C~201Fに示すような経路情報を持つことになる。ここでは、ASBRであるルータ20C, 20Dにexternal bestが設定されているものとする。

【0078】

このようなネットワーク構成において、ルータ20C, 20F間と、ルータ20C, 20D間に破線矢印で示すようなiBGPセッションが追加されると、以下ようになる。すなわち、ルータ20Cの経路情報201Cには、プレフィックスPについて、ベストルート以外にも符号301Cに示すBGPルート(ネクストホップがルータ20AのBGPルート)が登録されている。したがって、このルータ20Cは、iBGPセッションを確立したルータ20D, 20Fに対し、このベストルートと、符号301Cに示すBGPルートとを広告する。これより、ルータ20Dの経路情報201Dには、符号301Dに示すBGPルート(ネクストホップがルータ20CのBGPルート)が追加される。また、ルータ20Fの経路情報201Fには、符号301Fに示すBGPルート(ネクストホップがルータ20CのBGPルート)が追加される。これにより、ルータ20D, 20FはプレフィックスPについてネクストホップが冗長化された経路情報201を保持することができる。このように、境界ルータ(ASBR)にexternal bestを設定しておけば、AS内の各ルータ20は境界ルータ4から、ベストルート以外のBGPルートを受信しやすくなり、その結果、各ルータ20の経路情報201に同じプレフィックスについてネクストホップが異なる複数のBGPルートが設定されやすくなる。

【0079】

なお、このように境界ルータにexternal bestを設定したときは、前記した図4のS202において、セッション追加対象ルータ特定部125は、iBGPセッショントポロジ情報131を参照して、S201で選択したルータ20との間でiBGPセッションを確立していない境界ルータの集合を特定する。なお、AS内のルータ20のうち、どのルータ20が境界ルータかは、eBGPルート情報132またはiBGPセッショントポロジ情報131に記される情報を参照して判断するものとする。そして、S203においてセッション追加対象ルータ特定部125は、eBGPルート情報132を参照して、S202で特定した境界ルータの集合の中から、冗長性が必要なPreqをeBGPルートとして受信している境界ルータの集合を特定する。そして、S204において、この特定した境界ルータの集合の中で、iBGPセッションの確立により、S201で選択したルータ20とPreqのネクストホップが異なるものとなるルータ20の集合Cを特定する。S204以降の処理は、前記した処理と同様であるので説明を省略する。

【0080】

本実施の形態に係るセッション設計装置100は、前記したような処理を実行させるGPGセッション制御プログラムによって実現することができ、そのプログラムをコンピュー

タによる読み取り可能な記憶媒体（CD-ROM等）に記憶して提供することが可能である。また、そのプログラムを、インターネット等のネットワークを通して提供することも可能である。

【図面の簡単な説明】

【0081】

【図1】本実施の形態のシステムの全体構成および動作概要を例示した図である。

【図2】（a）は、図1のセッション設計装置の機能ブロック図であり、（b）は、図2（a）の経路情報DBを例示した図である。

【図3】図2（a）のセッション設計装置の処理手順を示した図である。

【図4】図2（a）のセッション設計装置の処理手順を示した図である。

10

【図5】本実施の形態の境界ルータ（ASBR）にexternal bestを設定した場合のBGPルートの広告について説明した図である。

【図6】本願発明の比較例となる技術を説明した図である。

【図7】本願発明の比較例となる技術を説明した図である。

【符号の説明】

【0082】

11 入出力部

12 処理部

13 記憶部

14 通信部

20

20（20A, 20C, 20D, 20E, 20F, 20G, 20H, 20I） ルータ

100 セッション設計装置

121 トポロジ情報収集部

122 eBGPルート収集部

123 経路情報収集部

124 セッション制御部

125 セッション追加対象ルータ特定部

131 iBGPセッショントポロジ情報

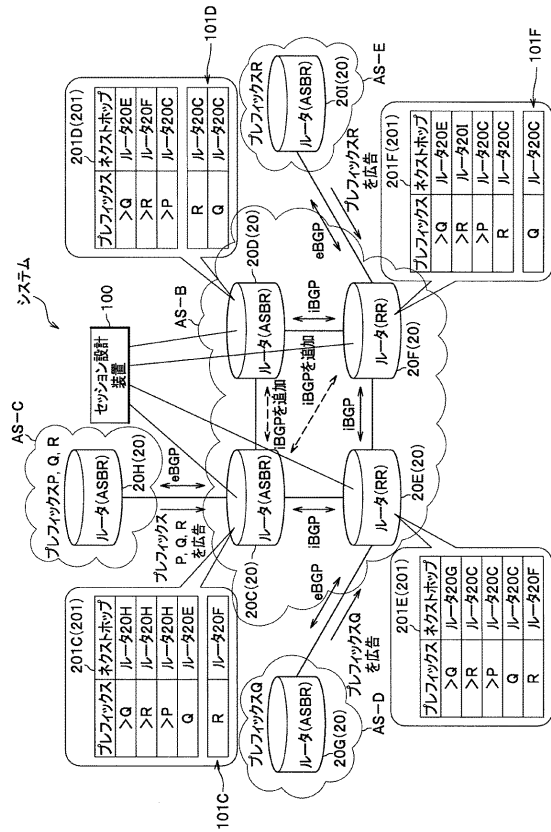
132 eBGPルート情報

133 経路情報DB

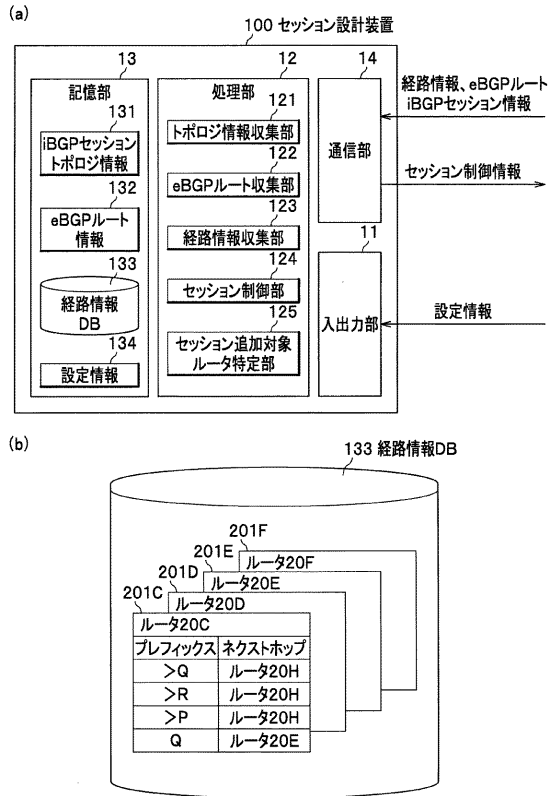
30

134 設定情報

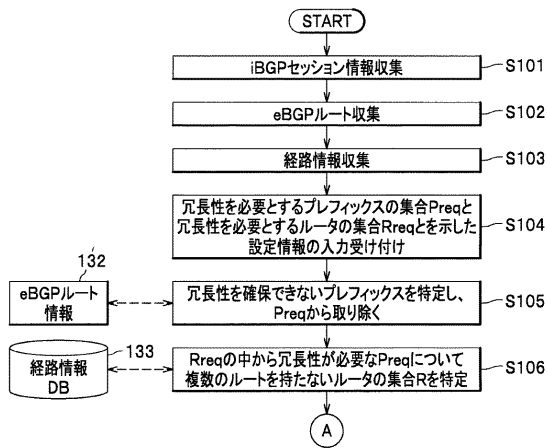
【図1】



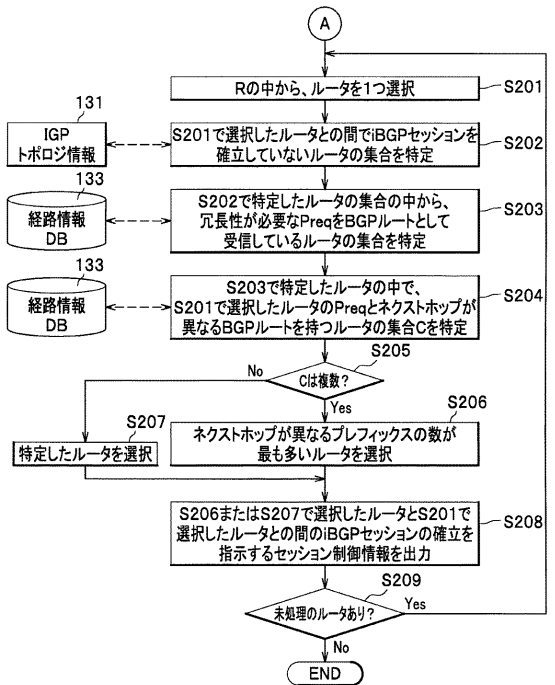
【図2】



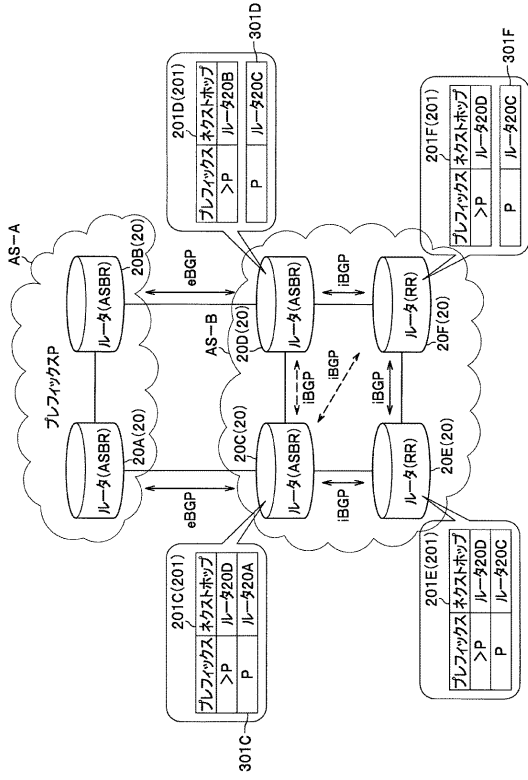
【図3】



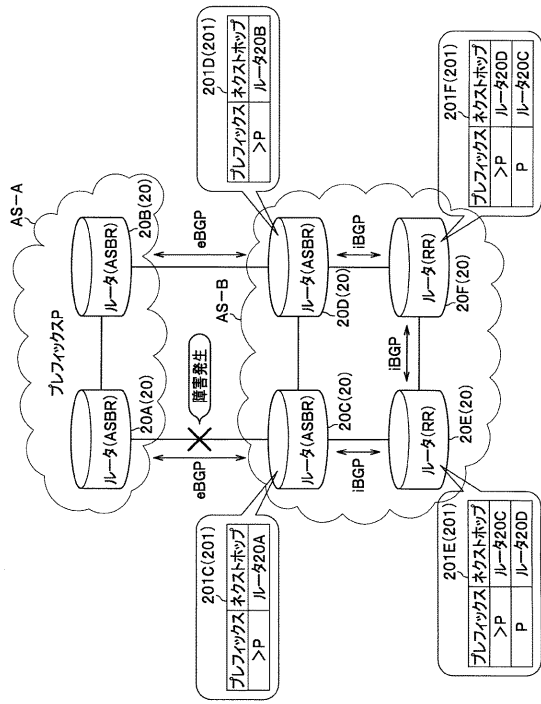
【図4】



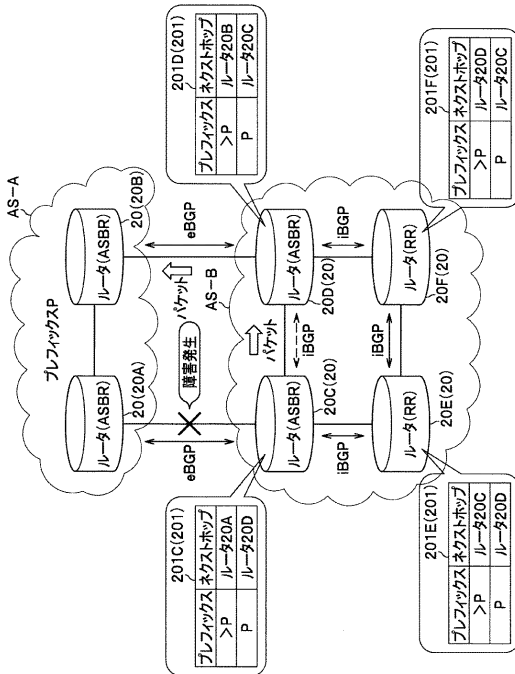
【 図 5 】



【 図 6 】



【 図 7 】



フロントページの続き

(72)発明者 大木 英司

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 塩本 公平

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

審査官 玉木 宏治

(56)参考文献 特開2003-244217(JP,A)

特開2007-180775(JP,A)

特開2005-020600(JP,A)

特表2008-543195(JP,A)

(58)調査した分野(Int.Cl., DB名)

H04L 12/00-66