



(12) 发明专利

(10) 授权公告号 CN 1610331 B

(45) 授权公告日 2010.05.26

(21) 申请号 200410054404.5

55-86 段、图 1, 6.

(22) 申请日 2004.07.20

US 2003/0169689 A1, 2003.09.11, 说明书第 20-51 段、图 1-7.

(30) 优先权数据

2003-364044 2003.10.24 JP

审查员 汤广强

(73) 专利权人 株式会社日立制作所

地址 日本东京都

(72) 发明人 柘植宗俊 续木彰人

(74) 专利代理机构 永新专利商标代理有限公司

72002

代理人 胡建新

(51) Int. Cl.

H04L 12/56 (2006.01)

H04L 12/28 (2006.01)

H04L 29/06 (2006.01)

(56) 对比文件

CN 1404268 A, 2003.03.19, 全文.

US 2003/0162499 A1, 2003.08.28, 说明书第

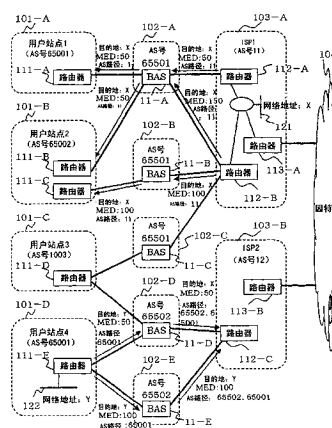
权利要求书 3 页 说明书 22 页 附图 10 页

(54) 发明名称

进行 AS 间的路由控制的通信装置及其路由控制方法

(57) 摘要

本发明提供一种进行 AS 间的路由控制的通信装置及其路由控制方法,在存根 AS 方和上游 AS 方之间可以实现策略路由控制,该策略路由控制包括能够只在相邻 AS 间使用的控制。在将自身的 AS 号通知和包含自身 AS 号的属性信息通知向存根 AS 方发出时,不使用自身的 AS 号或者取而代之使用上游 AS 的 AS 号。除此之外,在从一方 AS 方将所接收到的路由信息中继到另一方 AS 方时,通常将不通过 AS 加以中继的属性信息废弃,进行中继。



1. 一种通信装置,将端(stub)网群和赋予了自主系统AS号码的上游网连接,该端网群由赋予了指定的AS(Autonomous System)号码的一个以上的端网构成,其特征为:

该通信装置被赋予自己的AS号码,具有:

对该自己的AS号码进行存储的单元;

接口,用来向配置于上述端网或上述上游网中的其他通信装置发送接收遵守指定的协议的AS间传送控制用路径控制信息以及通知自己的AS号码的自身AS号码通知信息;

通知单元,经由上述接口向上述其他通信装置通知上述路径控制信息及上述自身AS号码通知信息;

处理器,用来进行上述路径控制信息的处理;

存储单元,存储用于进行上述路径控制的路径控制软件,

上述处理器在对位于上述端网侧的上述其他通信装置发送上述自身AS号码通知信息时,通知上述上游网的AS号码来取代自己的AS号码,

在将从上述上游网侧发送的路径控制信息向上述端网侧中继时,对于与数据包传送的优先路径有关的信息之中、上述路径控制协议规定的通过AS不被中继的信息,不予以废弃,而进行上述路径控制信息的处理。

2. 一种通信装置,将端网群和赋予了自主系统AS号码的上游网连接,该端网群由赋予了指定的AS(Autonomous System)号码的一个以上的端网构成,其特征为:

该通信装置被赋予自己的AS号码,具有:

对该自己的AS号码进行存储的单元;

接口,用来向配置于上述端网或上述上游网中的其他通信装置发送接收遵守指定的协议的AS间传送控制用路径控制信息以及通知自己的AS号码的自身AS号码通知信息;

通知单元,用来经由上述接口向上述其他通信装置通知上述路径控制信息及上述自身AS号码通知信息;

处理器,用来进行上述路径控制信息的处理;

存储单元,存储用于进行上述路径控制的路径控制软件,

上述处理器在对位于上述端网侧的上述其他通信装置发送上述自身AS号码通知信息时,通知上述上游网的AS号码来取代自己的AS号码,

在将从上述端网侧发送的路径控制信息向上述上游网侧中继时,对于与数据包传送的优先路径有关的信息之中、上述路径控制协议的规定的经过AS不被中继的信息,不予以废弃,而进行上述路径控制信息的处理。

3. 根据权利要求1或2记载的通信装置,其特征为:

还具有存储向上述端网发送的路径控制信息中不包含的AS号码的集合的存储单元,

上述处理器在将从上述上游网侧发送的路径控制信息向上述端网侧中继时,对于存储于路径控制信息内的、该路径控制信息通过的AS的历史信息,即使在该历史信息中包含自己的AS号码,也不废弃该路径控制信息,且从该历史信息中删除属于上述AS号码的集合的全部AS号码,并且不将自己的AS号码追加到该历史信息中,来进行上述路径控制信息的处理。

4. 根据权利要求3记载的通信装置,其特征为:

采用自己的AS号码代替上述AS号码的集合,来进行上述路径控制信息的处理。

5. 根据权利要求 3 记载的通信装置,其特征为:

上述处理器在对从上述端网侧发送的路径控制信息进行处理时,对于存储于路径控制信息内的、该路径控制信息通过的 AS 的历史信息,即使该历史信息中包含自己的 AS 号码,也不废弃该路径控制信息,来进行上述路径控制信息的处理。

6. 根据权利要求 5 记载的通信装置,其特征为:

上述处理器在对从上述端网侧发送的路径控制信息进行处理时,对于存储于路径控制信息内的、该路径控制信息通过的 AS 的历史信息,如果该历史信息中包含上述上游网的 AS 号码,则废弃该路径控制信息,来进行上述路径控制信息的处理。

7. 根据权利要求 3 记载的通信装置,其特征为:

上述处理器在对从上述端网侧发送的路径控制信息进行处理时,对于存储于路径控制信息内的、该路径控制信息通过的 AS 的历史信息,从该历史信息中将属于上述 AS 号码的集合的全部 AS 号码删除,来进行上述路径控制信息的处理。

8. 根据权利要求 3 记载的通信装置,其特征为:

上述路径控制协议使用 BGP。

9. 根据权利要求 8 记载的通信装置,其特征为:

作为上述路径控制协议的规定的通过 AS 不被中继的、与数据包传送的优先路径有关的信息,使用多出口标示的属性即 MED 属性。

10. 根据权利要求 3 记载的通信装置,其特征为:

上述路径控制协议使用 BGP;作为路径控制信息通过的 AS 的历史信息,使用 AS 路径属性。

11. 一种通信方法,在经由第 1 路由器连接了端网群和上游网的网络中使用,并且采用了自主系统 AS 间路径控制协议,该端网群由被赋予了 AS 号码的一个以上端网构成,该上游网被赋予了与上述 AS 号码的任一个都不同的 AS 号码,其特征为:

从上述第 1 路由器向位于上述端网内的第 2 路由器发送将上述上游网的 AS 号码作为自己的 AS 号码进行通知的信息,

从上述第 1 路由器向位于上述上游网内的第 3 路由器发送将指定的 AS 号码作为自己的 AS 号码进行通知的信息,该指定的 AS 号码与上述上游网的 AS 号码及上述端网群的 AS 号码的任一个都不同,

通过上述第 1 路由器来接收从上述第 3 路由器发送的路径控制信息,并按接收时的原样附加下述信息,来向上述第 2 路由器中继;该下述信息是上述第 1 路由器接收到的路径控制信息中含有的、与数据包传送的优先路径有关的信息中、路径控制协议规定的经过 AS 不被中继的信息。

12. 一种通信方法,在经由第 1 路由器连接了端网群和上游网的网络中使用,并且采用了自主系统 AS 间路径控制协议,该端网群由被赋予了 AS 号码的一个以上端网构成,该上游网被赋予了与上述 AS 号码的任一个都不同的 AS 号码,其特征为:

从上述第 1 路由器向位于上述端网内的第 2 路由器发送将上述上游网的 AS 号码作为自己的 AS 号码进行通知的信息,

从上述第 1 路由器向位于上述上游网内的第 3 路由器发送将指定的 AS 号码作为自己的 AS 号码进行通知的信息,该指定的 AS 号码与上述上游网的 AS 号码及上述端网群的 AS

号码的任一个都不同，

通过上述第 1 路由器来接收从上述第 2 路由器发送的路径控制信息，并按接收时的原样附加下述信息，来向上述第 3 路由器中继；该下述信息是上述第 1 路由器接收到的路径控制信息中含有的、与数据包传送的优先路径有关的信息中、路径控制协议的规定的经过 AS 不被中继的信息。

13. 根据权利要求 11 或 12 记载的通信方法，其特征为：

在上述第 1 路由器将从上述第 3 路由器接收到的路径控制信息向上述第 2 路由器进行中继时，即使存储于该路径控制信息内的、该路径控制信息通过的 AS 的历史信息中，包含向上述第 3 路由器通知了作为自己的 AS 号码的 AS 号码也进行中继，并且在存储于被中继信息内的 AS 的历史信息中，不包含向上述第 3 路由器通知了作为自己的 AS 号码的 AS 号码。

14. 根据权利要求 11 或 12 记载的通信方法，其特征为：

AS 间路径控制协议是 BGP。

15. 根据权利要求 14 记载的通信方法，其特征为：

路径控制协议的规定的通过 AS 不被中继的、与数据包传送的优先路径有关的信息，是多出口表示属性即 MED 属性。

16. 根据权利要求 13 记载的通信方法，其特征为：

AS 间路径控制协议是 BGP，路径控制信息通过的 AS 的历史信息是 AS 路径属性。

进行 AS 间的路由控制的通信装置及其路由控制方法

技术领域

[0001] 本发明涉及一种在通过路由器连接一个以上的存根 (stub) AS 和作为转接 AS 或存根 AS 的一个上游 AS 的通信网络中,作为路由器配置于网络内部的通信装置及其路由控制方法。

背景技术

[0002] 以日本·韩国·欧美为中心,取代以往的拨号式因特网连接,而以 ADSL(Asymmetric Digital Subscriber Line)为首的廉价固定连接高速线路(俗称的宽带线路)正在面向个人快速普及。许多 ISP(Internet Service Provider)面向个人提供有使用固定连接高速线路的因特网连接服务。另外,ISP 和通信公司已开始出现,该 ISP 和通信公司采用将与家庭用固定连接高速线路通用的基础设施应用于企业的小规模网点和面向 SOHO(Small Office/Home Office)企业的接入线路中的形式,提供对于用户来说廉价且易于导入的面向企业的接入线路。

[0003] 就以前的拨号式连接而言,RAS(Remote Access Server)被置于 ISP 网内,该 RAS 是进行用户连接认证和计费信息收集的通信装置。RAS 和用户之间只是通过 PSTN(Public Switched Telephone Networks) 电话线路或者利用在 IP 层次认为相同的隧道协议,进行连接。因而,ISP 和用户是与网络层有关被直接连接的状态。在此,网络层指的是,OSI(Open System Interconnection) 参考模型中的网路层,例如在 TCP/IP(Transmission Control Protocol/Internet Protocol) 协议体系的场合下,意味着 IP 层。

[0004] 图 4 表示的是,采用以往面向企业的接入线路的用户网和 ISP 网之间的网络结构、以及该结构的以往的 BGP 路由控制方法。

[0005] 以往的专用线、ATM 线路及帧中继线路之类面向企业的线路,可以当作从网络层也就是 IP 层所直接连接的线路。为此,在通过这些线路来连接企业用户网 101 和 ISP 网 103 的场合下,由于用户网 101 和 ISP 网 103 成为由直接线路所连结的相邻 AS,因而可以从一方 AS 向另一方 AS 利用各通常的 BGP 路由器建立通常的 EBGp 配对连接。而且,可以在其配对连接的基础上很容易地进行使用各种各样属性信息的策略路由控制,该属性信息如同 MED 属性那样,包含只通过相邻 AS 就可利用的属性信息。

[0006] 但是,有关 ADSL 和 FTTH(Fiber To The Home)之类的固定连接高速线路,除线路的带宽与拨号式连接相比大的多之外,就 ADSL 来说,需要在线路台站分别设置被称为 DSLAM(Digital Subscriber Line Access Multiplexer)的用户线路集线装置,该线路台站在狭窄的地域存在多个。为此,在日本国内除去一部分例外,接受用户线路的接入线路运营商和向用户提供因特网连接服务的 ISP 已变成不同的运营商,接入线路运营商向多个 ISP 提供 ADSL 和 FTTH 的接入线路。接入线路运营商采取下述形式,即将通过用户线路集线装置所连接的多条用户线路向被称为 BAS(Broadband Access Server)的一种路由器进行连接,因此取代 ISP 而进行用户认证和流通量集中,将其流通量传输给各 ISP。这是为了使用户认证和带宽利用有效化。

[0007] 另一方面,将按照其网络管理者的意向对在网络内通过何种路由收发数据包进行控制的方法,称为策略路由控制。作为策略路由控制的示例来说,列举出「想从哪个 ISP 获取通信数据包」「向哪条线路发送通信数据包」之类的控制。将在同一个策略下所管理范围的网路称为 AS(Autonomous System)。在 AS 中,存在从分配机关接受 AS 号分配的全局 AS 和在所管理的网路内可独自进行分配的专用 AS。在不同 AS 间进行路由控制所需的网路控制协议,被称为 EGP(Exterior Gateway Protocol)。在 EGP 之中被广泛实施的协议内,有 BGP(Border Gateway Protocol)。在 BGP 中,路由信息不断通过的路由器自动或者根据管理者的设定将各种各样的属性信息添加到该路由信息中,通告给相邻 AS。据此,可以跨越不同策略的 AS 间,统一实现策略路由控制。

[0008] AS 可以根据和其它 AS 之间数据包传输方式的不同,分成转接 AS 和存根 AS。存根 AS 在与相邻 AS 之间只进行以自 AS 为发送源的数据包发送和以自 AS 为目的地的数据包接收,不从相邻 AS 接收发给其它 AS 的数据包。另一方面,转接 AS 不仅仅进行与自 AS 相关的数据包收发,还经由自 AS 将从某个相邻 AS 到达并且发给其它 AS 的数据包中继到其它相邻 AS。作为转接 AS 的代表例,可以列举出 ISP 网,作为存根 AS 的代表例,可以列举出一般的企业网(或者其各网站点)。

[0009] BGP 在各转接 AS 之间的路由控制过程中,作为实质性标准的 AS 间路由控制协议已被广泛使用。另一方面,在转接 AS 和存根 AS 或者存根 AS 和存根 AS 之间的路由控制过程中,根据两个 AS 间网路结构的不同,也有时不对路由控制使用 BGP。例如,如同普通的个人用户与 ISP 进行连接的形式那样,通过单条线路与单个转接 AS 所连接的存根 AS,一般情况下不对路由控制使用 BGP。其另一方面,在使用多重初始地址连接以及 IP-VPN(Virtual Private Network) 的场合下,一般情况下在置于存根 AS 中的边缘路由器和转接 AS 的边缘路由器之间进行利用 BGP 的路由控制,该多重初始地址连接通过多条线路与转接 AS 连接,该 IP-VPN 是通过 ISP 网建立多个站点间虚拟专用线路的技术之一。

[0010] 在 BGP,在作为路由信息直接交换对方的路由器之间预先建立利用 TCP 的虚拟连接,通过该 TCP 连接来交换路由信息。将该 TCP 连接称为配对(pair)连接,将从进行配对连接的路由器之间一方的角度看上去的另一方路由器称为配对路由器。在配对连接中有 EBG(External BGP) 连接和 IBGP(Internal BGP) 连接的 2 种,该 EBG 连接是路由器属于相互不同 AS 的 AS 间配对连接,该 IBGP 连接是属于相互相同 AS 的 AS 内配对连接。EBG 连接是在不同的 AS 间进行路由信息交换所使用的,IBGP 连接是将通过 EBG 连接所得到的路由信息在 AS 内的全部 BGP 路由器上加以共享所使用的。由任一种形式配对连接所得到的路由信息都被使用于 AS 间的策略路由控制,在 AS 内的路由控制中通常不使用。

[0011] 为了在路由器间建立 EBG 连接,各配对路由器必须在 IP 层次上进行直接连接。原因是,由于 AS 间连接的 IP 路由控制通常只利用 BGP 来进行,因而在 EBG 连接前的时刻路由器不能得知直接路由之外的 AS 外路由。其结论为,如果不清楚 IP 层中到配对路由器的到达路由,当然既不可能在和配对路由器之间进行下述 TCP 连接,也不能进行配对连接,该 TCP 连接是在 IP 的一个上层进行动作的。另一方面,对于 IBGP 连接而言,各配对路由器也可以不在 IP 层次进行直接连接。原因是,一般情况下在 AS 内与 BGP 并行被称为 IGP(Interior Gateway Protocol) 的在 AS 内进行路由控制所用的路由控制协议另行进行动作,可以通过该 IGP 知道直接路由以外的 AS 内路由。

[0012] 进行 EBGp 连接的各配对路由器可以在 IP 层次上进行直接连接,也可以不必在数据链路以下的层次上进行直接连接。也就是说,各配对路由器也可以采用 MPLS(Multi Protocol Label Switching)、PPP(Point to Point Protocol)及 L2TP(Layer 2 Tunneling Protocol)等与 IP 层有关透过性的隧道协议,进行连接。例如,在特开 2002-368788 号公报中公示出下述技术,即在多个 AS 通过 MPLS 网相互连接的状况下属于各个 AS 的各路由器采用 EBGp 连接方式来进行策略路由控制时,将与 BGP 属性信息对应的 MPLS 标志添加到接收数据包中,经由 MPLS 向其它 AS 传输,以此使 MPLS 网内的数据包传输反映出 BGP 的策略路由控制。

[0013] 专利文献 1

[0014] 特开 2002-368788 号公报

[0015] 就采用专用线、ATM 及帧中继等的面向企业的线路的 AS 间连接来说,连结两个 AS 的线路中途,例如企业网的边缘路由器和 ISP 的边缘路由器之间,在网络层次上进行直接连接。因而,这些 AS 如企业网的 AS 和 ISP 所属的 AS 可以作为完全相邻的 AS 加以运用,能够相互在 EBGp 连接的基础上进行 BGP 路由通告。也就是说,可以使用全部的策略路由控制,该策略路由控制也包括只在相邻 AS 间可使用的策略。另外,即使在以前的拨号式连接中,由于 RAS 置于 ISP 中,因而用户和 ISP 间在网络层次上进行直接连接。因而,暂且不论有没有实用上进行的意义,只要 RAS 具备 BGP 路由器的功能,则也可以给拨号用户分配与 ISP 不同的 AS 号,在用户和 ISP 之间进行 EBGp 连接。

[0016] 但是,在构建 ADSL 和 FTTH 等网络中,很多时候用户和 ISP 网不在网络层次进行直接连接结构的网络被构建。就拥有 BAS 的接入线路运营商和 ISP 运营商来说,则是因为很多时候企业运营商不同。在这种状况下接入线路运营商的 BAS 和 ISP 网之间进行 BGP 路由控制时,一般来说给 BAS 和 ISP 网的各自分配不同的 AS 号。通过分配不同的 AS 号,而可以在接入线路运营商和 ISP 网之间维持与路由控制有关的管理独立性,除此之外因为 BAS 和 ISP 网之间成为 EBGp 连接,所以此处也可以进行策略路由控制。

[0017] 在此,考虑 ADSL 和 FTTH 等的用户想在自身的用户网和 ISP 网之间进行利用 BGP 的路由控制的情形。这种应用形式,在用户不是个人和家庭而是小规模企业和企业各地网点之类企业用户网的场合下被充分考虑。

[0018] 这种场合下,因为用户想在和 ISP 之间利用 BGP 进行策略路由控制的缘由,所以当然给用户网分配与 ISP 网不同的 AS 号。除此之外,由于在用户网的边缘路由器和 ISP 网方的边缘路由器之间配置有 BAS 也就是路由器,因而不能在用户网和 ISP 网之间直接建立通常的 EBGp 连接。因而,在以往网络结构的 ADSL 和 FTTH 的基础上,为了采用通常的 EBGp 连接方式,在用户和 ISP 之间进行利用 BGP 的策略路由控制,而必须给用户网、BAS 及 ISP 网的各自分配不同的 AS 号,在用户网和 BAS 以及 BAS 和 ISP 网的各自之间建立 EBGp 连接。但是,就这种网络结构来说,因为不通过 AS 的属性信息,不在用户网和 ISP 网之间进行传送,所以不能使用策略路由控制之中只在相邻 AS 间可使用的路由控制。

[0019] 另一方面,由于采用下述 EBGp 多转发连接功能在用户网的边缘路由器和 ISP 网的边缘路由器之间直接建立 EBGp 多转发连接,因而也能够实现用户网和 ISP 网之间的策略路由控制,该 EBGp 多转发连接功能是在网络层次不相邻的路由器间进行 EBGp 连接的功能。但是,EBGp 多转发连接,存在下述等之类的与网络管理・运营的复杂性和保密性有关的问题,

[0020] 1) 需要附加设定,包括向用户网的边缘路由器和 ISP 网方的边缘路由器双方静态地设定到相互的路由器的路由等,

[0021] 2) 因为 ISP 网方边缘路由器的冗余配置和 IP 地址变更等,很多时候若对 BAS 和 ISP 网之间的网络结构进行变更,则其影响波及到用户网边缘路由器的设定,所以 ISP 方难以进行 BAS 和 ISP 网间的通信路由设定变更,若一旦对 BAS 和 ISP 网间的网络结构进行变更,则用户网方的边缘路由器也必须对设定加以变更,

[0022] 3) ISP 网方和用户网方双方的边缘路由器由于不能判别经由 BAS 所接收到的 BGP 数据包发送源 IP 地址是否是被冒充的地址,因而有成为 DoS(Denial of Service) 攻击等目标的危险,该攻击是由冒充发送源 IP 地址的配对连接请求而引起的。

发明内容

[0023] 本发明的目的为,提供一种上述通信装置及其路由控制方法,在下述通信网络中,在存根 AS 方和上游 AS 方之间不用 EBGp 多转发连接,而可以实现包括可以只在相邻 AS 间适用的控制的策略路由控制,该通信网络以经由 BAS 来连接用户网和 ISP 网的结构为典型示例,其结构为通过作为路由器进行动作的通信装置来连接一个以上的存根 AS 和作为转接 AS 或存根 AS 的一个上游 AS。

[0024] 在本发明中,由于通信装置中所进行的网络层次路由控制协议处理,对于存根 AS 方来说如同其通信装置是否属于上游 AS 那样进行动作,因而使通信装置所属的 AS 从存根 AS 方看来成为透明,解决上述课题。为此,上述通信装置在对存根 AS 方的配对路由器发出由路由控制协议处理所进行的自身 AS 号通知和包含自身 AS 号的属性信息通告时,不使用自身的 AS 号或者取而代之使用上游 AS 的 AS 号。除此之外,在将从一方 AS 的配对路由器所接收的路由信息中继到另一方 AS 的配对路由器时,一般情况下不将通过 AS 不被中继的属性信息废弃,而进行中继。

[0025] 下面,采用下述示例来说明发明的概要,该示例为使用本发明的通过 BAS 来连接用户网和 ISP 网的例子。BAS 如果从作为 EBGp 配对路由器的用户网边缘路由器接收到附加属性信息后的路由信息,则与通常的 BGP 路由器相同,将其路由信息使用于 BAS 自身的路由表中,此后将该路由信息中继到作为另一方 EBGp 配对路由器的 ISP 网边缘路由器。但是此时,在通常路由器的场合下,由其路由器自身对不越过 AS 进行传播的属性(MED 属性)进行解释,在向 ISP 网边缘路由器中继路由信息时,不附加其属性,而在本实施示例的 BAS 中,则将这些属性按原状态中继给 ISP 网边缘路由器。对于其它越过 AS 进行传播的属性(AS 路径属性、下一跳属性等),与通常的路由器相同地进行处理。

[0026] 另外,本发明的 BAS 如果从 ISP 网边缘路由器接收到附加属性信息后的路由信息,则与通常的 BGP 路由器相同,将其路由信息使用于 BAS 自身的路由表中,此后将该路由信息中继到用户网边缘路由器。此时,本实施示例的 BAS 与通常的路由器相同,也包括不越过 AS 进行传播的属性(MED 属性等)由自身来解释属性信息,而在向用户网边缘路由器中继路由信息时,将原本不越过 AS 进行传播的属性之中的若干个按原状态附加到路由信息中。对于 AS 路径属性(路由信息所不断通过的 AS 按最近通过的顺序被记录的属性),即使在 AS 路径中包含 BAS 自身的 AS 号也不将路由信息废弃,而从 AS 路径整体将其 AS 号(或是预先设定于 BAS 中的 AS 号集合内所包含的 AS 号)全部删除,不在该 AS 路径前面附加 BAS 自身的

AS 号,而将加工后的 AS 路径属性附加到路由信息中。对于其它越过 AS 进行传播的属性,也取代包含 BAS 自身的 AS 号而包含 ISP 网的 AS 号。通过上面的处理,将所生成的带属性信息的路由信息中继到用户网的边缘路由器。

[0027] 除此之外,有关 BGP 的配对连接,虽然在连接确立时将自路由器的 AS 号通知给配对路由器,但是此时 BAS 对用户网边缘路由器通知 ISP 网的 AS 号。这如同 MED 属性那样,是为了使通常只在相同从相邻 AS 所通告的路由信息之间进行比较的属性有效工作,以及为了避免下述状况,该状况包括因在用户站点方显示专用 AS 号而使用户站点方的路由器设定变得复杂,和在用户站点方给路由控制引起混乱。还有, BAS 将自 AS 号向 ISP 网边缘路由器通知。

[0028] 发明效果

[0029] 这样,由于中间隔者本实施示例的 BAS,在用户网边缘路由器和 BAS 之间以及 BAS 和 ISP 网边缘路由器之间双方进行本发明方式的 BGP 路由控制,因而尽管中间隔着具有与用户网和 ISP 网都不同独立 AS 号的 BAS,也可以在 ISP 网和 BAS 之间、ISP 网和用户网之间双方的区间上实现策略路由控制,该策略路由控制采用原来只能在直接相邻的 AS 间进行交换的 MED 属性等属性信息。而且,此时在用户网边缘路由器和 ISP 网边缘路由器的双方上不需要特殊的功能和设定,而可以按照通常的 BGP 设定来使用通常的 BGP 路由器。尤其是对于用户网的边缘路由器,可以通过与隔着 BAS 而和 ISP 网直接连接的情况相同的功能和设定,来实现几乎相同的策略路由控制。

[0030] 还有,虽然本发明将利用 BAS 的网络结构作为典型的实施示例做出假设,但是实施本发明的 IP 路由器所需要的功能,只是采用通常 BGP 协议的路由控制功能、和本发明的向 BGP 处理部的追加功能。为此,本发明不限于 BAS,而可以应用于与本实施示例 BAS 相同的、在网络结构中使用的全部 BGP 对应 IP 路由器。

附图说明

[0031] 图 1 表示的是实施示例 1 的网络结构以及本发明的 BGP 路由控制方法。

[0032] 图 2 表示的是可以满足本实施示例 BAS 功能的内部处理体系结构一个示例。

[0033] 图 3 表示的是在实施示例 1 的一部分上采用本发明 BGP 路由控制方法的配对连接确立处理时的开放报文。

[0034] 图 4 表示的是采用以往面向企业接入线路的用户网和 ISP 网之间的网络结构以及其结构上的以往 BGP 路由控制方法。

[0035] 图 5 表示的是按照与实施示例 1 相同的网络结构将不具备本发明 BGP 路由控制方法的以往 BGP 路由器作为 BAS 加以使用时的 BGP 路由控制动作。

[0036] 图 6 表示的是按照与实施示例 1 相同的网络结构在用户网和 ISP 网之间设建立 EBGp 多转发连接来进行 BGP 路由控制并不使 BAS 加入 BGP 路由控制时 EBGp 多转发连接的状态。

[0037] 图 7 表示的是 BGP 数据包的数据结构。

[0038] 图 8 表示的是 BGP 开放报文的数据结构。

[0039] 图 9 表示的是 BGP 更新报文的数据结构。

[0040] 图 10 是在实施示例 1 的一部分上从 ISP 网方本实施示例的 BAS 接收 BGP 路由控

制报文将其中继到用户网方时的收发顺序图。

[0041] 图 11 是在实施示例 1 的一部分上从用户网方本实施示例的 BAS 接收 BGP 路由控制报文将其中继到 ISP 网方时的收发顺序图。

[0042] 图 12 表示的是实施示例 2 的网络结构以及本发明的 BGP 路由控制方法。

具体实施方式

[0043] 下面,有关本发明的实施示例,采用附图予以说明。

[0044] 实施示例 1

[0045] 图 1 表示的是实施示例 1 的网络结构及本发明的 BGP 路由控制方法。

[0046] 在本实施示例中,本实施示例的 BAS11-A、11-B、11-C、11-D 及 11-E 如图 1 所示,从数个线路接口通过通信线路(或是在网络层上认为相同的隧道线路),与 ISP 网边缘路由器 112-A、112-B 及 112-C 相连接。另外, BAS11-A、11-B、11-C、11-D 及 11-E 如图 1 所示,从其它若干个线路接口通过通信线路(或是在网络层上认为相同的隧道线路),与用户网边缘路由器 111-A、111-B、111-C、111-D 及 111-E 相连接。

[0047] 还有,在实施本发明时,这些通信线路的连接方式也可以不是本实施示例附图所示的那样, BAS 可以采用插入 ISP 网和(单个或多个)用户网之间的形式来连接线路。但是,在 1 台 BAS 与多个 ISP 网边缘路由器相连接的情况下,那些 ISP 网边缘路由器必须全部属于相同的 AS。也就是说,连接 1 台 BAS 的 ISP 网的 AS 是单一的。一般情况下,如果路由器所属的 AS 相同,则路由器所属的 ISP 也是相同的,因此 1 台 BAS 与单个 ISP 网连接。在本实施示例中设为, ISP 网 1103-A、ISP 网 2 103-B 都设立 AS 边界使之与 ISP 的网络边界相一致。而且, BAS11-A、11-B 及 11-C 对于 ISP 来说只与 ISP 网 1 103-A 连接,同样 BAS11-D 及 11-E 只与 ISP 网 2 103-B 连接。

[0048] BAS11-A、11-B、11-C、11-D 及 11-E 通过这些连接线路,在用户网 101-A、101-B、101-C、101-D 以及 ISP 网 103-A、103-B 之间对 IP 数据包流通量进行中继。也就是说,对于从 ISP 网发向用户网的数据包而言,ISP 网边缘路由器 112-A、112-B 及 112-C 的下一跳(next hop)成为 BAS11-A、11-B、11-C、11-D 及 11-E 的任一个, BAS 的下一跳成为用户网边缘路由器 111-A、111-B、111-C、111-D 及 111-E 的任一个。相反,对于从用户网发向 ISP 网的数据包而言,用户网边缘路由器 111-A、111-B、111-C、111-D 及 111-E 的下一跳成为 BAS11-A、11-B、11-C、11-D 及 11-E 的任一个, BAS 的下一跳成为 ISP 网边缘路由器 112-A、112-B 及 112-C 的任一个。还有,下一跳指的是,某个路由器准备将作为目的地具有某个 IP 地址的 IP 数据包传达到其目的地时该数据包的传输目标路由器。该下一跳根据各路由器和其它路由器进行交换动态得到的路由信息以及在其路由器自身中静态设定的路由信息来决定。

[0049] 但是,与通常的 IP 路由器不同, BAS11-A、11-B、11-C、11-D 及 11-E 既不从用户网 101-A、101-B、101-C 及 101-D 向其它用户网 101-A、101-B、101-C 及 101-D 中继流通量(这种中继是经由 ISP 网来进行的),也不进行将下述流通量中继到与 ISP 网之间的其它连接线路的处理(这种中继设为在 ISP 网内进行),该流通量是从与 ISP 网 103-A、103-B 之间的连接线路所到达的。另外,设为用户网 101-A、101-B、101-C 及 101-D 是存根 AS。也就是说,这些用户网将经由该用户网从用户网外(和 BAS 之间的连接线路等)所到达的数据包中继到用户网外,即不进行作为转接 AS 的动作。

[0050] 用户网 101-A、101-B、101-C、101-D 以及 ISP 网 103-A、103-B 各自具有相互不同的 AS 号。关于用户网和 ISP 的 AS 号是否应当是专用 AS 号,本发明没有特别的限制,可以按照通常的 BGP 应用方法加以决定。例如,在本实施示例的场合下设为,因 ISP 网 103-A、103-B 是与因特网所连接的转接 AS,所以具有全局(global)的 AS 号 11、12;因用户网 3 101-C 与多个 ISP 网进行多重初始地址连接,所以具有全局的 AS 号 1003;因其它用户网 101-A、101-B、101-D 只与单个 ISP 网相连接,所以具有从 ISP 所分配的专用 AS 号 65001、65002、65001。

[0051] 另外,虽然 ISP 网 103-A、103-B 通过上游用边缘路由器 113-A、113-B 连接到因特网 104,但是本发明不管有没有连接都可以使用。还有,在本实施示例中「上游」指的是,从 11-A ~ 11-E 的各 BAS 看来更接近核心网络的方面。「上游网」意味着,从 BAS 看来成为上游方的网络。如果通过图 1 来说,就是 ISP 网 103-A、103-B 为上游网,该 ISP 网 103-A、103-B 是从 BAS 看来更接近因特网 104 的网络。相反,用户网 101A ~ D 成为「下游网」。在存在上游用边缘路由器 113-A、113-B 并且由此采用 BGP 等 AS 间路由控制协议向因特网 104 的那种 ISP 外通告路由信息的场合下,在通常的 BGP 应用过程中,是在将其路由信息中所包含的专用 AS 号全部删除之后进行通告的。

[0052] BAS11-A、11-B、11-C、11-D 及 11-E 被作为与用户网和 ISP 网都不同的 AS 来运用。在本实施示例中,由于各 BAS 只与对每 1 台 BAS 特定的 ISP 网相连接,因而从所连接的 ISP 进行分配,将所连接的 ISP 与用户网不重复的专用 AS 号作为自 AS 号加以使用。还有,为了实施本发明,在向 BAS 所分配的 AS 号与用户网和 ISP 网都不重复并且包含向 BAS 所分配 AS 号的路由信息按原状态流向与 ISP 网相比更上游的场合下,需要使其 AS 号与通过因特网所连接的其它任何 AS 都不重复。在通常的运用过程中,由于完全没有使 BAS 的 AS 号流向与 ISP 网相比更上游的必要性,因而在本实施示例中设为,给 BAS 分配专用 AS 号,在 ISP 向下述其它 AS 通告路由信息时,将其路由信息中含有的专用 AS 号删除进行通告,该其它 AS 是因特网和其它 ISP 等专用 AS 号的分配不在自 ISP 的管理之下加以进行的 AS。还有,虽然取代专用 AS 号而取得与 ISP 网、用户网及其它任何 AS 都不重复的全局 AS 号,并给予 BAS,这在理论上也是可能的,但是没有实际应用上的价值。

[0053] 而且,在本实施示例中 BAS11-A、11-B、11-C、11-D 及 11-E,通过图 1 所示的通信线路中一方末端为 BAS 的全部线路,与线路连接对象的用户网边缘路由器 111-A、111-B、111-C、111-D、111-E 以及 ISP 网边缘路由器 112-A、112-B、112-C,进行 EBG 连接。例如, BAS11-A 与 ISP 网边缘路由器 112-A、112-B 以及用户网边缘路由器 111-A、111-B 进行 EBG 连接。还有,对于该 EBG 连接的初始确立动作,本实施示例的 BAS 也需要进行与通常的 BGP 路由器不同的动作,对此将在下面予以说明。

[0054] 还有,在本实施示例中,将 BAS11-A、11-B、11-C、11-D 及 11-E 分别作为物理上不同的设备,也就是物理路由器加以处理。但是,本发明如果将它们作为按照 IP 路由控制的观点看独立的路由器进行处理,则可以使用,因此 11-A、11-B、11-C、11-D 及 11-E 的各自也可以是收存于相同 BAS 设备内的不同虚拟路由器。

[0055] 上面,在图 1 所示的网络结构之下说明本发明 BGP 路由控制方法的动作。

[0056] 首先,作为本发明的前提条件,使 BAS 所使用的 AS 号成为处于同一 ISP 之下的所有 BAS 共用的 AS 号。但是,该 AS 号需要与处于同一 ISP 之下 BAS 以外所使用的 AS 号(在

使用专用 AS 号的场合下,是用户网等在 ISP 管理之下所分配的其它 AS 号)不重复。若在本实施示例的场合下,则使处于 ISP 网 1 103-A 之下 BAS11-A、11-B、11-C 所属的 AS 102-A、102-B、102-C 全部成为相同的专用 AS 号 65501,同样使处于 ISP 网 2 103-B 之下 BAS11-D、11-E 所属的 AS102-D、102-E 全部成为相同的专用 AS 号 65502。这是为了从附属于路由信息中的 AS 路径属性删除 BAS 自身的 AS 号,向用户网边缘路由器 111-A、111-B、111-C、111-D 及 111-E 进行通告,以此可以在用户网方易于利用 MED 属性,以及为了做出是否全部是从同一相邻 AS 所收到路由信息的假设,向 ISP 网边缘路由器 112-A、112-B、112-C 进行通告,以此可以在 ISP 网方易于利用 MED 属性。

[0057] 该 MED 属性是给 BGP 的路由信息所附加的属性信息一种,是在二个相邻 AS 间通过多条线路连接的场合下,在从一方 AS 向另一方 AS 对 IP 数据包传输时应优先使用的线路进行指定时所采用的。从同一相邻 AS 接收到与相同 IP 地址前缀有关并且 MED 属性值不同的多个路由信息的 BGP 路由器,如果应比 MED 属性更为优先的各种路由选择条件是相同的,则选择 MED 属性值最小的路由。例如,AS1 和 AS2 通过线路 1、线路 2 的 2 条线路相连接,在线路 1 上将 MED 属性的值设为 50 并且在线路 2 上设为 100,使与同一目的 IP 地址前缀有关的路由信息从 AS1 流向 AS2,这种场合下,AS2 的路由器通过 MED 属性值小的线路 1,将 IP 数据包中继到 AS1,该 IP 数据包作为目的地具有与其 IP 地址前缀匹配的 IP 地址。

[0058] 由于 MED 属性值可以由路由信息发送源 AS 的管理者自由决定,因而一般情况下即便对发送源 AS 不同的路由信息中附属的各 MED 属性值进行比较,也没有意义(除去该各发送源 AS 有关 MED 属性值的设定已作出规定的情形)。为此,通常的 BGP 路由器在对从不同相邻 AS 所接收到的各路由信息进行比较,进行路由选择时,不使用 MED 属性。另外,在路由信息通过某个 AS 向其它 AS 进行中继时, MED 属性被删除。这种场合下,该通过 AS 应当按照需要,将 MED 属性值重新确定成适当的值。

[0059] 还有,处于不同 ISP 之下的 BAS,在限于使用 ISP 专用 AS 号的范围内,相互相同的专用 AS 号和不同的专用 AS 号都可以。例如,若在本实施示例的场合下,则 AS 102-A、102-B、102-C 和 AS102-D、102-E 的双方都可以使用相同的专用 AS 号 65501。

[0060] 另外,也可以取代使用处于同一 ISP 之下的全部 BAS 中共用的 AS 号,而使用在 BAS 的若干个中相互不同的 AS 号,并且能够给各 BAS 设定在全部 BAS 中所使用的 AS 号集合。这种场合下,在从向用户通告的路由信息中附属的 AS 路径属性删除 AS 号的处理之时,将后一个 AS 号集合中含有的所有 AS 号从 AS 路径属性删除。在向 ISP 网通告路由时,假设是 BAS 自身的 AS 号而对前一个 AS 号进行处理。这样一来,例如如果给 ADSL 线路接受的 BAS 和光纤线路接受的 BAS 分别分配相互不同的专用 AS 号,优先对从后一个专用 AS 所通告的路由信息进行处理,则与 MED 属性无关,而可以使通过光纤线路的路由总是比通过 ADSL 线路的路由更为优先。该 MED 属性是从用户网所通告的路由信息中附属的。但是,在本实施示例中只要预先没有特别通知则采用下述方法,该方法使用在全部 BAS 中共用的专用 AS 号。

[0061] 下面,有关本发明中从用户网向 ISP 网的路由信息通告方法,予以说明。例如考虑到下述情形,即用户网 4 101-D 内的边缘路由器 111-E 通过与 BAS11-D、11-E 两个 BAS 之间的 EBGp 配对连接,将与到下述网络 122 的到达路由有关的路由信息最终通告给 ISP2103-B 内的边缘路由器 112-C,该网络 122 具有网内的 IP 地址前缀 Y。但是,设为用户网 4 101-D 因某种理由(例如,为了网络 112 的用户特别想确保下行方向的带宽等)而想要通告目的

地 Y 的路由信息,以便经由 BAS11-D 的路由被优先选择。

[0062] 这种场合下,用户网边缘路由器 111-E 可以在与 BAS11-D 之间的 EBG 连接上使 MED 属性值小(此处设为 50)的路由信息流过,在与 BAS11-E 之间的 EBG 连接上使 MED 属性值大(此处设为 100)的路由信息流过。

[0063] 收到这些路由信息的 BAS11-D、11-E 首先按照需要,进行路由信息是否未进行循环的检查。此时,就通常 BGP 中路由信息接收时的处理来说,要是在路由信息中所附加的 AS 路径属性内包含自 AS 号,则认为该路由信息已进行循环将其废弃,而本实施示例的 BAS 取代自 AS 号转而对是否包含 ISP 的 AS 号进行检查。还有,使用 ISP 的 AS 号之路由信息循环检查由于通常也由 ISP 网边缘路由器 112-C 来进行,因而不一定需要 BAS11-D、11-E 加以进行。为此,也可以设为本实施示例的 BAS 完全不进行从用户网边缘路由器所接收到的路由信息循环检查。还有,该路由信息循环指的是,再次由自路由器接收以前自路由器接收并中继到其它路由器的路由信息。此时,若自路由器进行与上次接收时相同的处理,则相同的路由信息被多次发送,存在路由信息无限增加的可能。为此,一般情况下进行路由信息循环的检测处理,在被检测出的场合下将该路由信息废弃。

[0064] 如果判明未进行路由信息循环检查或者作为进行结果路由信息未进行循环,则在自身的路由表中反映出该路由信息,此后对该路由信息的 AS 路径属性和下一跳属性等进行与通常 BGP 路由器相同的处理,来制作用于向 ISP 网边缘路由器 112-C 进行通告的路由信息。然后,虽然在通常的 BGP 路由通告处理过程中,不将原来路由信息中所附加的 MED 属性值附加到新制作出的路由信息中,但是在本发明的路由通告处理过程中,则将其附加。有关 AS 路径属性,与通常的 BGP 路由器相同,可以将自 AS 号追加到 AS 路径的前面,只附加到新制作出的路由信息中。但是,在用户网存在对 AS 路径属性插入无效专用 AS 号,不断通告路由信息的可能性的场合下,为了避免 ISP 网方的路由控制混乱,也可以将从 AS 路径整体只消除 BAS 自身 AS 号(或者,如果预先在 BAS 内设定出在同一 ISP 内所使用全部 BAS 的 AS 号集合,则是该 AS 号集合中含有的全部 AS 号)的处理,作为将自 AS 号追加到 AS 路径前面之前的处理,加以进行。通过上面的处理,将所生成的带属性信息的路由信息向 ISP 网边缘路由器 112-C 进行通告。

[0065] 这样一来,向 ISP 网边缘路由器 112-C 从 BAS11-D、11-E 的各自到达两个路由信息,该两个路由信息其通告源的相邻 AS 是相同的 65502、AS 路径属性的长度相同、MED 属性值不同并且将目的地设为 Y。然后,ISP 网边缘路由器 112-C 根据通常的 BGP 路由选择处理,将这些路由信息之中 MED 属性值小且从 BAS11-D 到达的路由信息作为适当路由予以选择。其结果为,将与 IP 地址前缀 Y 匹配的 IP 地址作为目的地的数据包只要在 BAS11-D 及通过此的路由上不发生故障,就从 ISP 网边缘路由器 112-C 经由 BAS11-D 到达用户网边缘路由器 111-E。

[0066] 还有,如果 BAS11-D 不是 BAS 而是通常的 BGP 路由器,则除非未在该路由器中做出将路由通告加以屏蔽的设定,从用户网边缘路由器 111-E 所获取到的路由信息就不只是作为 EBG 连接配对路由器的 ISP 网边缘路由器 112-C,也同样向作为 EBG 连接配对路由器的用户网边缘路由器 111-D 进行通告。但是,在 11-D 是本实施示例 BAS 的场合下,由于从用户网只经由 BAS 向其它用户网中继流量,是不进行的通告前提条件,因而不会有 BAS11-D 将从用户网边缘路由器 111-E 所获取到的路由信息向用户网边缘路由器 111-D 进行通告的

处理。

[0067] 下面,有关本发明中从 ISP 网向用户网的路由信息通告方法,予以说明。例如考虑到下述情形,即 ISP 用户网 1 103-A 内的边缘路由器 112-A、112-B 通过与 BAS11-A、11-B 两个 BAS 之间的 EBGP 连接,将与到下述网络 121 的到达路由有关的路由信息最终向用户网 1 101-A 内的边缘路由器 111-A 和用户网 2 101-B 内的边缘路由器 111-B、111-C 进行通告,上述网络具有 ISP 网内的 IP 地址前缀 X。但是,设为 ISP1 103-A 因某种理由(例如,根据 ISP 网内部的网络拓扑,ISP 网边缘路由器 112-A 比 ISP 网边缘路由器 112-B 更为接近网络 121,并且 BAS11-A 和 ISP 边缘路由器 112-B 之间的线路是备份用低速线路等),而想要通告目的地 X 的路由信息,使之按照经由 BAS11-A 及 ISP 网边缘路由器 112-A、经由 BAS11-B 及 ISP 网边缘路由器 112-B、经由 BAS11-A 及 ISP 网边缘路由器 112-B 的顺序来优先选择路由。

[0068] 这种场合下,ISP 网边缘路由器 112-A 在与 BAS11-A 之间的 EBGP 连接上使 MED 属性值小(此处设为 50)的路由信息流过,ISP 网边缘路由器 112-B 可以在与 BAS11-A 之间的 EBGP 连接上使 MED 属性值大(此处设为 150)的路由信息流过,在与 BAS11-B 之间的 EBGP 连接上使具有中间 MED 属性值(此处设为 100)的路由信息流过。

[0069] 获取到这些路由信息的 BAS11-A、11-B,虽然在通常 BGP 中的路由信息接收时的处理过程中,首先进行路由信息是否未进行循环的检查,而就本实施示例的 BAS 而言则不进行该检查。然后,首先在自身的路由表中反映该路由信息。此时,由于 BAS11-A 分别从 ISP 网边缘路由器 112-A、112-B 获取到两个路由信息,该两个路由信息其通告源的相邻 AS 是相同的 11、AS 路径属性的长度相同、MED 属性值不同并且将目的地设为 X,因而根据通常的 BGP 路由选择处理,将这些路由信息之中 MED 属性值小并且从 ISP 网边缘路由器 112-A 到达的路由信息,作为适当路由予以选择。

[0070] 然后, BAS11-A、11-B 在将所选择出的路由信息反映到自身的路由表中之后,对其路由信息的下一跳属性等进行与通常 BGP 路由器相同的处理,来制作用于向用户网边缘路由器 111-A、111-B、111-C 进行通告的路由信息。然后,虽然在通常的 BGP 路由通告处理过程中不将原来路由信息中所附加的 MED 属性值附加到新制作出的路由信息中,但是就本发明的路由通告处理而言则将其附加。除此之外,有关 AS 路径属性,虽然在通常的 NGP 路由通告处理过程中只是向 AS 路径的前面附加自 AS 号,但是就本发明的路由通告处理而言,则从 AS 路径整体只将 BAS 自身的 AS 号(或者说,如果预先在 BAS 内设定在同一个 ISP 内所使用全部 BAS 的 AS 号集合,则是该 AS 号集合中含有的全部 AS 号)全部删除,不在该 AS 路径的前面附加 BAS 自身的 AS 号而附加到新制作出的路由信息中。通过上面的处理,将所生成的带属性信息的路由信息向用户网边缘路由器 111-A、111-B、111-C 进行通告。

[0071] 这样一来,从 BAS11-A 或 11-B 向各个用户网边缘路由器 111-A、111-B、111-C 分别到达一个将目的地设为 X 的路由信息,使得通告源的相邻 AS 全部是 11,也就是相邻 AS 是 ISP。此时,由于 BAS11-A、11-B 的 AS 号是 65501,因而根据通常的 BGP 处理方法,在采用该 AS 号确立出和用户网边缘路由器 111-A、111-B、111-C 之间的配对连接的场合下, BAS 在配对连接确立时所告知的 AS 号和路由信息中含有的 AS 号产生矛盾。但是,通过由本实施示例的 BAS 来使用下述的配对连接确立处理,该问题被解决。

[0072] 上面的结果为,来自用户网边缘路由器 111-A 的将与 IP 地址前缀 X 匹配的 IP 地

址设为目的地的数据包,只要在 BAS11-A 和 ISP 网边缘路由器 112-A 之间的线路以及 ISP 网边缘路由器 112-A 其本身上不发生故障,就经由 BAS11-A 到达 ISP 边缘路由器 112-A。对于用户网边缘路由器 111-B、111-C 而言,如果在这些路由器间进行 IBGP 连接,则一方从 BAS 所接收到的路由信息被通告给另一方,那些路由信息被实施通常的 BGP 路由选择处理,最后选择 MED 属性值较小并且从 BAS11-A 所通告的路由。其结果为,来自用户网 2 101-B 的将与 IP 地址前缀 X 匹配的 IP 地址设为目的地的数据包,只要在路由上的线路或 BAS 或者路由器的任一个中不发生故障,就从用户网边缘路由器 111-B 经由 BAS11-A 到达 ISP 网边缘路由器 112-A。

[0073] 还有,如果 BAS11-A 不是 BAS 而是通常的 BGP 路由器,则除非未在该路由器中做出将路由通告予以屏蔽的设定,从 ISP 网边缘路由器 112-A 所获取到的路由信息就不只是作为配对路由器的用户网边缘路由器 111-A、111-B,也同样向作为配对路由器的 ISP 网边缘路由器 112-B 进行通告。但是,在 BAS11-A 是本实施示例 BAS 的场合下,由于在从 ISP 经由 BAS 将流通量中继到朝向同一 ISP 的其它线路的处理是不进行通告的前提条件,因而不会出现 BAS11-A 将从 ISP 网边缘路由器 112-A 所获取到的路由信息向 ISP 网边缘路由器 112-B 进行通告的处理。这对于 BAS11-A 从 ISP 网边缘路由器 112-B 所获取到的路由是相同的。

[0074] 如上所述,由于将根据本发明 BGP 路由控制方式的 BAS 使用于本实施示例中,因而可以中间隔着具有独立 AS 号的 BAS,在用户网和 ISP 网之间使用策略路由控制,该策略路由使用了原本不隔着 AS 进行通告的 MED 属性。此时所使用的用户网和 ISP 网的边缘路由器可以是已做出标准设定的通常 BGP 路由器。

[0075] 图 2 表示的是可以满足本实施示例 BAS 功能的内部处理体系结构一个示例。

[0076] 还有,在本发明中,作为本实施示例 BAS 功能的前提条件,规定出不直接中继从用户网到用户网的流通量,以及不中继从 ISP 网到 ISP 网的流通量。另外,在本发明中如同此前所述那样,按照与通常 BGP 路由控制处理之间的差分形式规定出本实施示例的 BAS 应进行的处理内容。但是,只要满足上面的功能,则本实施示例的 BAS 内部体系结构可以是任何形式,不需要一定是按照附图的体系结构。

[0077] 在图 2 体系结构示例的场合下,本实施示例 BAS11 的内部由控制处理部 201、下游接口部 202、上游接口部 203 及底板 204 的 4 种内部硬件模块构成。还有,附图中用实线框住的要件意味着硬件,用虚线框住的要件意味着软件或者表格等的的数据。

[0078] BAS11 在其内部具有一个以上的下游接口部 202。下游接口部 202 具有一个以上接口,用来连接到用户网边缘路由器 111 的线路。下游接口部 202 作为其内部硬件具有数据包传输处理用处理机 221 及传输处理用存储器 222。在传输处理用存储器 222 中,存储有程序 225 和下游接口用路由表 227。在下游接口用路由表 227 中存储对应表,用来在从用户网边缘路由器 111 接收到数据包时,按照其数据包的目的 IP 地址来决定作为其数据包传输目标的 ISP 网边缘路由器 112 以及其被连接的上游接口部 203(但是,在目的 IP 地址是 BAS 自身地址的场合下,将传输目标设为控制处理部 201)。还有,下游接口用路由表 227 由 BAS11 的路由控制部 201 随时进行重写。程序 225 是 BAS11 启动时从控制处理部 201 的辅助存储装置 212 所传输的程序代码。数据包传输处理用处理机 221 通过执行程序 225,采用下游接口用路由表 227,来决定从用户网边缘路由器 111 所接收到的数据包适当的传输目标,通过底板 204 向该传输目标传输数据包。另外,数据包传输用处理机 221 通过执行程

序 225, 将从底板 204 所接收到的数据包发送到适当的用户网边缘路由器 111。据此, 向下游网方的通信装置通知路由控制报文和自 AS 号通知报文。

[0079] 另外, BAS11 在其内部具有一个以上的上游接口部 203。上游接口部 203 具有一个以上接口, 用来连接到 ISP 网边缘路由器 112 的线路。上游接口部 203 作为其内部硬件具有数据包传输处理用处理机 231 及传输处理用存储器 232。在传输处理用存储器 232 中, 存储有程序 235 和上游接口用路由表 237。上游接口用路由表格 237 的存储内容和数据包传输处理用处理机 231 的处理内容除去上游和下游以及用户网和 ISP 网进行转换的内容之外, 与下游接口用路由表 227 及数据包传输处理用处理机 221 完全相同。也就是说, 通过数据包传输处理用处理机 231, 向上游网方的通信装置通知路由控制报文和自 AS 号通知报文。

[0080] BAS11 通过其内部所具备的底板 204, 将控制处理部和上游接口部、控制处理部和下游接口部以及上游接口部和下游接口部相互连接。

[0081] BAS11 的内部所具备的控制处理部 201 用来进行以路由控制协议为开始的发给自己或自发 IP 数据包处理以及 BAS11 装置整体管理的处理等。控制处理部 201 作为其内部硬件具有控制处理用处理机 211、辅助存储装置 212 及主存储装置 213。在辅助存储装置 212 中存储有程序代码 215 和初始设定数据 217。程序代码 215 由控制处理部用程序代码和接口部用程序代码构成, 并且全都在 BAS11 启动时向各自的存储器 (主存储装置 213 及传输处理用存储器 222、232) 进行传输, 由各自的处理机 (控制处理用处理机 211 及数据包传输处理用处理机 221、231) 加以进行。在初始设定数据 217 中存储有由管理者设定于 BAS11 内的各种信息, 用来在 BAS11 启动时及管理者指定时由控制处理用处理机 211 来读入, 按照设定数据的内容向 BAS11 内的各部进行适当设定。在初始设定数据 217 中所存储的设定数据内, 包括 BAS11 自身的 AS 号、ISP 网的 AS 号、各接口的 IP 地址设定、各配对路由器的 IP 地址以及静态路由设定等。

[0082] 向主存储装置 213 展开由控制处理用处理机 211 所进行的控制处理部用程序代码, 可以分成 OS・设备驱动程序 251、路由控制处理管理进程 (process) 252、BGP 路由控制进程 255、其它路由控制处理进程组 254 以及其它控制处理进程组 253。OS・设备驱动程序 251 用来将进程管理、输入输出处理及 TCP/IP 协议处理等各种基本处理功能提供给各进程。其它控制处理进程组 253 用来进行与路由控制以外有关的处理如装置管理和隧道协议等的处理。其它路由控制处理进程组 254 用来进行与 BGP 之外的各种路由控制协议有关的处理。路由控制处理管理进程 252 用来对从 BGP 路由控制进程 255 和其它路由控制处理进程组 254 所得到的路由信息及由管理者做出的静态路由设定信息进行汇总, 生成上游・下游接口用路由表 227、237 的数据, 向上游・下游接口部 202、203 进行传输。

[0083] BGP 路由控制进程 255 用来使标准的 BGP 处理进程, 根据本发明反映出下述两种处理, 一是应对用户网边缘路由器 111 进行的开放报文处理, 二是应对从用户网边缘路由器 111 及 ISP 网边缘路由器 112 所获取到的 BGP 路由信息进行的属性信息处理。也就是说, 该 BGP 路由控制进程 255 是具备下述功能的软件进程, 该功能为根据本发明的 AS 间路由控制方法, 发出对配对路由器并采用开放报文的自身 AS 号通知以及在和配对路由器之间的路由控制报文相互通知。还有, 虽然 BGP 是在 TCP 的基础上进行动作的协议, 但是 BGP 路由控制进程只具备与 BGP 其本身有关的处理功能, TCP 之下的协议处理是具备有 TCP/IP 协议堆栈的 OS・设备驱动程序 251 来进行的。

[0084] BGP 路由控制进程 255 在其内部具有上行流通量用 BGP 路由表 262 以及下行流通量用 BGP 路由表 263。BGP 路由控制进程 255 根据从 ISP 网边缘路由器 112 所接收到的路由控制报文,进行上行流通量用 BGP 路由表 262 的更新,将与其更新部分对应的路由控制报文通告给用户网边缘路由器 111。另外,上行流通量用 BGP 路由表 262 的更新部分通过路由控制处理管理进程 252 也反映到下游接口用路由表 227 中。另外,BGP 路由控制进程 255 根据从用户网边缘路由器 111 所接收到的路由控制报文,进行下行流通量用 BGP 路由表 263 的更新,将与其更新部分对应的路由控制报文通告给 ISP 网边缘路由器 112。另外,下行流通量用 BGP 路由表 263 的更新部分通过路由控制处理管理进程 252 也反映到上游接口用路由表 237 中。

[0085] 本实施示例的 BAS11 若从各用户网边缘路由器 111 获取到发给自己的 BGP 路由控制报文,则该报文通过下游接口部 202、底板 204 及控制处理部 201 的 OS・设备驱动程序 251,最后传递到 BGP 路由控制进程 255。BGP 路由控制进程 255 根据本发明的 AS 间路由控制方法对该报文进行处理,更新下行流通量用 BGP 路由表 263,通过路由控制处理管理进程 252 使上游接口用路由表 237 也反映其更新内容。另外,BGP 路由控制进程 255 生成用于将路由信息的更新部分向 ISP 网边缘路由器 112 进行通告的路由控制报文,通过 OS・设备驱动程序 251、底板 204 及上游接口部 203,向 ISP 网边缘路由器 112 进行通告。

[0086] 在从各 ISP 网边缘路由器 112 获取到发给自己的 BGP 路由控制报文的场合下,也通过同样的流程(但是,上行和下行、上游和下游以及 ISP 网和用户网分别相反)来处理报文,进行上行流通量用 BGP 路由表 262 的更新、下游接口用路由表 227 的更新以及向各用户网边缘路由器 111 的路由通告之类的处理。

[0087] 图 3 表示的是在实施示例 1 一部分上采用本发明 BGP 路由控制方法的配对连接确立处理时的开放报文。

[0088] 本实施示例的 BAS 所进行的配对连接确立处理除去开放报文中含有自 AS 号的通知之外,与根据通常 BGP 处理方法的配对连接确立处理完全相同。当然,用户网边缘路由器及 ISP 网边缘路由器所进行的配对连接确立处理也包括开放报文中含有的自 AS 号通知,可以与根据通常 BGP 处理方法的配对连接确立处理完全相同。在此,虽然只是有关本实施示例中的 BAS11-A 进行说明,但是即使对于本实施示例中的其它 BAS,也同样使用下面的处理。

[0089] BAS11-A 与通常的 BGP 处理方法相同,将作为自 AS 号的 65501 插入 BGP 开放报文(message)的自 AS 号字段(field)中,从 BAS 向 ISP 网边缘路由器 112-A、112-B 进行发送(报文 302-A、302-B)。另一方面,对于用户网边缘路由器 111-A、111-B 而言则与通常的 BGP 处理方法不同,将作为 ISP 的 AS 号之 11 插入 BGP 开放报文的自 AS 号字段中,从 BAS 向用户网边缘路由器 111-A、111-B 进行发送(报文 301-A、301-B)。

[0090] 如上所述,ISP 网边缘路由器 112-A、112-B 及用户网边缘路由器 111-A、111-B 如同通常的 BGP 处理方法那样,将各路由器所属的 AS 号插入 BGP 开放报文的自 AS 号字段中,向 BAS 进行发送(报文 303-A、303-B、304-A、304-B)。

[0091] 由于采用配对连接确立处理,该配对连接确立处理使用了上面的开放报文,因而即使按照如图 1 所说明的方法来进行路由信息的收发,配对连接确立时所接收到的 AS 号和所接收到路由信息的 AS 路径属性中含有的 AS 号,从用户网看来也不产生矛盾。加之,由于

可以对用户网方完全隐藏 BAS 的 AS 号,因而可以从用户网看见在和 ISP 网的 AS 号之间直接进行 EBG 配对连接。

[0092] 下面,采用图 5~图 6,来说明不给 BAS 增加功能而进行 BGP 路由控制时的问题所在,同时说明本发明的优越性。

[0093] 图 5 表示的是,将以往的 BGP 路由器作为 BAS 使用,来构成与实施示例 1 相同的网络时 BGP 路由控制的动作。

[0094] 在以往面向个人的 ADSL/FTTH 存取网络中,在使 BAS 具备与本发明无关的以往 BGP 路由控制方法,并且不使用 EBG 多转发连接功能而对利用 BGP 路由控制的企业用户网加以连接的场合下,将成为图 5 的网络结构。

[0095] 给以往的 BGP 相应 BAS511-A、BAS511-B、BAS511-C、BAS511-D 及 BAS511-E 分配专用 AS 号,以使所连接 ISP 网 103-A、103-B 的相同的各 BAS 不为同一个 AS 号。例如,由于 BAS511-A、BAS511-B、BAS511-C 其所连接的 ISP 网是同一个 ISP 网 1 103-A,因而分配各自相互不同的 AS 号 65501、65502、65503。原因是,若给和同一 ISP 网所连接的多个 BAS 分配相同的 AS 号,则在经由 ISP 网将一方 BAS 从用户网所接收到的路由信息中继到另一方 BAS 时,在该路由信息的 AS 路径属性中已经包含与自身 AS 号相同的 AS 号,因此被当作路由信息的循环,致使该路由信息将被废弃。就本实施示例的 BAS 而言,由于不进行路由信息的废弃而从 AS 路径属性删除自身的 AS 号,因而避免了该问题。

[0096] 在此,与图 1 的情况相同考虑到下述情形,即从 ISP 网边缘路由器 112-A、112-B 经由 BAS511-A、511-B,向用户网边缘路由器 111-A、111-B、111-C 通告与到地址前缀 X 网络的到达路由有关的路由信息。此时,与图 1 的情况相同,设为,在从 ISP 网边缘路由器 112-A 向 BAS511-A 所通告的路由信息中,附加值为 50 的 MED 属性,在从 ISP 网边缘路由器 112-B 向 BAS511-A 所通告的路由信息中,附加值为 150 的 MED 属性,在从 ISP 网边缘路由器 112-B 向 BAS511-B 所通告的路由信息中,附加值为 100 的 MED 属性。BAS511-A、511-B 虽然接收这些 MED 属性,但是这些 BAS 是采用以往 BGP 路由控制方法的 BGP 路由器,并且具有与用户网 101-A、101-B 不同的 AS 号,因此不将 MED 属性中继到用户网边缘路由器 111-A、111-B、111-C。其结果为,ISP 网所通告路由信息的 MED 属性在和 BAS 之间是有效的,而在和用户网之间却变得无效。也就是说,在 BAS511-A 向 ISP 网传输下述 IP 数据包时,使用 MED 属性值较小的到 ISP 网边缘路由器 112-A 的线路,上述 IP 数据包将与地址前缀 X 匹配的 IP 地址作为目的地。但是,在经由 BAS511-A、511-B 的任一个,向 ISP 网 1 103-A 传输从用户网 2 101-B 内所发送的相同 IP 数据包时,作为下述资料来说 MED 属性完全不起作用,该资料用来判断用户网边缘路由器 111-B、111-C 经由哪方的 BAS 进行发送是适当的。

[0097] 假设增加功能以使 BAS511-A、511-B 中继 MED 属性,从 BAS511-A 到达用户网边缘路由器 111-B 的路由信息以及从 BAS511-B 到达用户网边缘路由器 111-C 的路由信息,从用户网 2 看来也被当作各自不同的从相邻 AS 到达的路由信息。为此,在用户网边缘路由器 111-B、111-C 中,只要没有做出特别设定,即在从不同的相邻 AS 到达的路由信息之间比较 MED 属性,则在这些路由器比较双方的路由信息来选择适当路由时, MED 属性就不被利用。就本实施示例的 BAS 来说,由于中继 MED 属性并且对用户网表明 BAS 是否属于 ISP 网的 AS,因而避免了上面的问题。

[0098] 对于从用户网边缘路由器 111-E 经由以往的 BGP 相应 BAS511-D、511-E,向 ISP 网

边缘路由器 112-C 通告与到地址前缀 Y 网络的到达路由有关的路由信息的情形,虽然也因相同的理由而在用户网和 BAS 之间利用 MED 属性,但是在用户网和 ISP 网之间 MED 属性不起作用。就本实施示例的 BAS 而言,由于中继 MED 属性并且给在同一 ISP 网的下游所使用的全部 BAS 分配同一个 AS 号,因而避免了上面的问题。

[0099] 另外,用户网 3 101-C 和多个 ISP 网进行多重初始地址连接。为此,ISP 网 103-A、103-B 需要对用户网 3 101-C 隐藏各 ISP 独自分配利用的专用 AS 号(如果不那样做,则存在对用户网 3 的路由控制引起故障的可能性)。许多路由器具备有下述功能,即在通告路由信息时,从其 AS 路径属性将专用 AS 号全部删除。为此,可以认为以往的 BGP 相应 BAS511-C、511-D 也可以在向用户网边缘路由器 111-D 通告路由信息时,从 AS 路径属性将全部的 AS 号删除进行通告。但是,BAS511-C、511-D 作为自身的 AS 号使用专用 AS 号,并且该号码包含于配对连接确立时从 BAS 向用户网边缘路由器所发送的开放报文中。加之,假使做出从 AS 路径属性将全部专用 AS 号删除的设定,则根据 BGP 路由控制功能安装的不同,存在将其该路由器自身的 AS 号附加到 AS 路径属性中的状况不被避免的可能性。而采用本实施示例的 BAS,即使给 BAS 分配专用 AS 号,也不引起这些问题。

[0100] 图 6 表示的是,按照与实施示例 1 相同的网络结构,在用户网和 ISP 网之间建立 EBGp 多转发连接,来进行 BGP 路由控制,并不使 BAS 加入 BGP 路由控制时 EBGp 多转发连接的方式。

[0101] 如上所述,EBGP 多转发连接功能指的是,在网络层也就是 IP 层次不相邻的路由器间进行 EBGp 连接的功能。可以认为,在以往面向个人的 ADSL/FTTH 存取网中,采用 EBGp 多转发连接功能对利用 BGP 路由控制的企业用户网加以连接的场合下,将成为图 6 的那种网络结构。

[0102] 若采用 EBGp 多转发连接功能,则如同图 6 那样,尽管中间相隔 BAS,但是仍可以在用户网和 ISP 网之间建立 EBGp 多转发连接 621。由于使用该 EBGp 多转发连接 621,因而与通常的 EBGp 连接相同,可以进行相邻 AS 间的策略路由控制。

[0103] BAS611 在用户网边缘路由器 111 和 ISP 网边缘路由器 112 之间,对包含 BGP 路由控制报文且非发给自己的 IP 数据包进行中继。但是,由于 BAS611 不进行 BGP 路由控制,因而没有发送自发 BGP 路由控制报文的处理,并且将全部发给自己的 BGP 路由控制报文废弃。但是,在 BAS611 和 ISP 网边缘路由器 112 之间也想进行采用 BGP 的策略路由控制的场合下,也可以将 BAS611 作为 BGP 路由器,与附图所示的 EBGp 多转发连接 621 不同,在 ISP 网边缘路由器 112 和 BAS611 之间建立 EBGp 连接。在图 6 中,设为在 ISP 网边缘路由器 112 和 BAS611 之间没有 EBGp 连接的状况。

[0104] 这样,如果采用 EBGp 多转发连接功能,则在用户网和 ISP 网间表明很容易使之满足进行 BGP 策略路由控制这样的要求,而在 EBGp 多转发连接功能中存在若干问题。

[0105] 第一个问题是,为了在未直接连接的路由器间进行 EBGp 连接,若不以某种形式预先设定至配对路由器的路由信息,则 EBGp 配对连接其本身不能进行。例如,由于从用户网边缘路由器 111-B 到 BAS611-A 是通过线路直接连接的,因而从用户网边缘路由器 111-B 到 BAS611-A 的到达路由是清楚的。但是,由于与 ISP 网边缘路由器 112-A 未直接连接,因而若不按照静态路由设定等预先给予路由信息,则用户网边缘路由器 111-B 在传送用于 BGP 配对连接确立的 IP 数据包时,不清楚如果向哪条线路传送数据包,该数据包将传递到 ISP 网

边缘路由器 112-A。为了避免该问题,而不得不在用户网边缘路由器 111-B 中预先设定通向 ISP 网边缘路由器 112-A 的静态路由。

[0106] 第二个问题是,若 BAS 和 ISP 网边缘路由器之间的网络结构产生变化,则很多时候用户网边缘路由器的设定也需要进行变更。例如设为,当以前不存在对 BAS611-A 和 ISP 网边缘路由器 112-B 加以连接的线路时,新设置该线路并成为图 6 的状态。此时,虽然 BAS611-A 和 ISP 网 1103-A 之间的线路被冗余化为 2 条线路,但是用户网 1101-A、用户网 2101-B 不能获得这种好处。原因是,用户网和 ISP 网之间经由 BAS611-A 的配对连接只存在一方配对路由器为 ISP 网边缘路由器 112-A 的配对连接 621-A、621-B。为此,如果 ISP 网边缘路由器 112-A 停止,则用户网将经由 BAS611-A 的与 AS 外路由有关的信息全部丢失。为了避免这种情况,需要经由 BAS611-A 在用户网边缘路由器 111-A、111-B 和 ISP 网边缘路由器 112-B 之间新建立 EBGp 多转发连接。对此,进行新的 EBGp 多转发连接所需的设定不只是在 ISP 网边缘路由器 112-B 中,在用户网边缘路由器 111-A、111-B 中也是必要的。

[0107] 第三个问题是,因 IP 数据包发送源地址冒充而引起的保密性问题。一般情况下,IP 数据包标题中包含的发送源 IP 地址的值可以由发送源主机自由设定,并且不给通向目的地主机的到达性带来影响。为此,也可以在发送源 IP 地址的字段中设定虚假的值。但是,就采用直接线路的通常 EBGp 连接来说,由于 BGP 数据包的真的发送源主机被限定为存在于其直接线路上主机之中的任一个,因而可以轻易地认清发送源 IP 地址的冒充。另一方面,就 EBGp 多转发连接来说,由于配对路由器不存在于直接线路上,因而即使发送源 IP 地址被冒充也不能进行判别。例如,在图 6 的场合下考虑到下述情形,即用户网边缘路由器 111-B 将下述 BGP 数据包传送给 ISP 网边缘路由器 112-A,该 BGP 数据包将发送源 IP 地址冒充为用户网边缘路由器 111-A 的地址。此时,接收到该 BGP 数据包的 ISP 网边缘路由器 112-A 不清楚该数据包是否真是从用户网边缘路由器 111-A 到达的,或者是从其它主机冒充发送源 IP 地址所发送的。为此,由于从配对路由器之外的第三个主机发送向 BGP 配对连接所使用 TCP 端口的 TCP 连接或断开数据包,因而存在引起 DDoS 攻击和不能预期配对连接断开的可能性。

[0108] 若采用 EBGp 多转发连接在用户网和 ISP 网之间进行 BGP 路由控制,则发生如上的新问题。另一方面,采用本发明的 BGP 路由控制方法,则不发生这些问题。

[0109] 图 7 表示的是 BGP 数据包的数据结构。

[0110] BGP 数据包指的是,在配对路由器之间收发 BGP 报文时包含有其收发所使用的 TCP 连接控制信息或者在其连接上所收发数据的 IP 数据包。

[0111] 在 IP 标题 701 中,按照 IP 协议标准规定包含 BGP 数据包发送源地址 711 及目的地址 712 等信息。就 BGP 数据包来说,在目的地址 712 中插入配对路由器的 IP 地址。还有,作为 IP 协议而言,一般使用版本 4(Ipv4) 和版本 6(Ipv6) 的 2 种,而本发明使用任一个都可以实施。另外,在实际上使物理线路流过 BGP 数据包时,一般来说在 IP 标题之前附加线路种类所依赖的数据链路层标题,而本发明不依赖于线路种类和数据链路层标题的形式就可以实施。

[0112] 在 IP 标题 701 之后接着 TCP 标题 702。在 TCP 标题 702 中按照 TCP 协议标准规定,包含发送源端口 721 和目的端口 722 等信息。发送源端口 721 是 BGP 数据包发送源的 BGP 路由器使用于配对连接中的 TCP 端口号码。目的端口 722 是 BGP 数据包的目的地也就

是配对路由器使用于配对连接中的 TCP 端口号码。配对连接用的 TCP 连接是通过从一方配对路由器（连接源端口号码为任意）向另一方配对路由器的配对连接受理用端口请求 TCP 连接而建立的。在该配对连接受理用端口中，通常使用作为 BGP 用标准 TCP 端口所规定的第 179 号端口。但是，即便使用此外的号码，本发明也能够实施。

[0113] 在 TCP 数据中包含 0 个以上 BGP 报文的一部分或全部。由于 BGP 报文在 TCP 连接的流上进行流动，因而并不一定在一个 IP 数据包的 TCP 数据 703 中包含一个 BGP 报文。例如，如果也有时在一个 IP 数据包的 TCP 数据 703 中包含多个 BGP 报文，则还有时在一个 IP 数据包的 TCP 数据 703 中只包含一个 BGP 报文的一部分。

[0114] 还有，BGP 路由器为了区别通常的通信数据包和 BGP 数据包，而使用 IP 数据包的目的地址 712 和目的端口 722。下面，作为示例根据图 2 列举出的本实施示例 BAS 内部体系结构进行说明。若接口部 202、203 从外部接口接收到 IP 数据包，则检查其目的地址 712 是否与该路由器自身所具有的 IP 地址相一致，如果相一致则向控制处理部 201 进行传输。如果不一致，则若找出应传送的接口就从其外部接口进行传送，若未找到就将其废弃。接收到发给自己 IP 数据包的控制处理部 201 内 OS・设备驱动程序 251，如果其 IP 数据包是 TCP 的数据包，则检查 BGP 路由控制进程 255 为配对连接用所打开的端口号码和该 IP 数据包的目的端口 722 是否相一致，如果一致则向 BGP 路由控制进程 255 传递该数据包的 TCP 数据 703 部分。

[0115] 图 8 表示的是 BGP 开放报文的数据结构。

[0116] 根据 BGP 规定出若干个 BGP 报文，而开放报文是其中之一。开放报文只在配对连接确立时 TCP 连接刚刚确立后进行发送。但是，本发明不根据通知 AS 号的报文形式和时刻，就可以实施。

[0117] 开放报文按照 BGP 协议标准规定，由 BGP 报文标题 801 和开放报文数据 802 构成。在 BGP 报文标题 801 中包含报文长度 811 和报文类型 812 等信息。在报文长度 811 中存储将 BGP 报文标题 801 和开放报文数据 802 合并后的字节长度。BGP 报文为了流到 TCP 连接的流上，而采用该报文长度 811 来判别从所接收到 TCP 流数据的何处到何处是一个 BGP 报文。在报文类型 812 中存储开放报文及更新报文等为每个报文类型所确定的值。通过查看该报文类型 812，可以判别接于 BGP 报文标题 801 之后的报文数据是开放报文数据还是其它报文数据部分。在开放报文数据 802 中，包含自 AS 号 821 等信息。在自 AS 号 821 中，存储将开放报文发送后的 BGP 路由器所属的 AS 号。根据该自 AS 号 821，BGP 路由器可以对其配对路由器传输自身的 AS 号。

[0118] 图 9 表示的是 BGP 更新报文的数据结构。

[0119] 与开放报文相同，更新报文也是 BGP 报文的一种。更新报文用来向配对路由器通知路由信息的添加・变更・删除，也是包含路由信息的唯一 BGP 报文。为此，在本说明书中一直称为 BGP 路由控制报文的就相当于该 BGP 更新报文。但是，只要属性信息的处理与本发明的前提条件相一致，本发明不依赖于路由控制报文的形式和种类就可以实施。

[0120] 更新报文按照 BGP 协议标准规定，由 BGP 报文标题 801 和更新报文数据 902 构成。BGP 报文标题 801 的形式除去在报文长度 811 中存储字节长度并且在报文类型 812 中存储报文类型值之外，与开放报文的情形完全相同，该字节长度是将 BGP 报文标题 801 和更新报文数据 902 合并后的长度，该报文类型值表示更新报文。更新报文数据 902 由删除路由集

合 921、属性信息集合 922 及添加路由集合 923 三个部分构成。在删除路由集合 921 中,存储 0 个以上成为不能到达路由的目的地址前缀。在属性信息集合 922 中,存储附随添加路由集合 923 的 0 个以上属性信息。一个属性信息由属性类型、属性长度及属性值三个组来表示。在属性类型中存储表示其属性信息类型(AS 路径属性、MED 属性、下一跳属性及其它各种属性)的值。在属性长度中存储属性值的字节长度。在属性值中存储其属性信息的值(AS 路径、MED 属性值等)。在添加路由集合 921 中存储 0 个以上新成为可到达或者已是可到达而属性被更新的路由的目的地址前缀。对添加路由集合 921 中含有的全部添加·更新路由,使用属性信息集合 922 中含有的全部属性信息。

[0121] 还有,在采用 BGP 对 Ipv4 之外与网络层协议(Ipv6 等)有关的路由信息进行通告的场合下,取代删除路由集合 921 和添加路由集合 923,而采用具有同等功能的特殊属性,将不能到达·可到达的目的地址前缀集合包含于属性信息集合 922 中。但是,本发明在按照任一个数据结构通告路由信息的场合下,都可以使用。

[0122] 图 10 是在实施示例 1 中,本实施示例的 BAS 从 ISP 网侧接收 BGP 路由控制报文,将其中继到用户网方时的收发顺序图。

[0123] BAS11-A 在从 ISP 网边缘路由器 112-A、112-B 接收到路由信息时,不进行路由信息循环的检查。另外,在向用户网边缘路由器 111-A、111-B 通告路由信息时,按照原来的值将 MED 属性附加到路由信息中,如果在路由信息的 AS 路径属性中已存在自身的 AS 号,则将其除去,不将自身的 AS 号追加到 AS 路径属性中,通告该路由信息。如果不包括这些动作,则图 10 的收发顺序与采用通常 BGP 路由控制的路由控制报文收发顺序完全相同。还有,在图 10 的顺序开始时刻,设为在用户网边缘路由器 111-A、111-B 和 BAS11-A 之间以及 ISP 网边缘路由器 112-A、112-B 和 BAS11-A 之间的双方已建立 EBGP 配对连接,还未进行任何与地址前缀 X 有关的路由信息通告。但是,在 ISP 网边缘路由器 112-A、112-B 进行路由通告之后进行配对连接确立等场合下,如果有关上述场合之外的路由信息通告也不包括上述属性操作的差异,则可以是与采用通常 BGP 路由控制的收发顺序完全相同的处理。

[0124] 下面,详细说明收发顺序。设从 ISP 网边缘路由器 1 112-A 向 BAS11-A 发送路由控制报文(步骤 1011),该路由控制报文其 MED 属性值为 50、AS 路径属性只是 ISP 网的 AS 号 11 并且与目的地址前缀 X 的路由添加有关。接收到该报文的 BAS11-A 如果是通常的 BGP 路由控制处理,则首先检查该报文的 AS 路径属性中是否包含自身的 AS 号,如果包含则将其路由信息废弃,在本实施示例的 BAS 不进行该检查。然后,与通常的 BGP 路由控制处理相同,和此前从其它配对路由器所接收到目的地 X 的路由信息做比较,决定最佳路由。此处,由于未从其它任何配对路由器再接收到目的地 X 的路由信息,因而本次所接收到的路由信息成为最佳路由。如果新接收到的路由信息是最佳路由,则将由该路由信息而引起的更新反映到 BAS11-A 自身的上行流通量用 BGP 路由表中,将与更新部分相当的路由信息作为路由控制报文发送给用户网边缘路由器 111-A、111-B(步骤 1012-A、1012-B)。此时,如果是通常的 BGP 路由控制处理,则不给所发送的路由信息附加接收时的 MED 属性值,在 AS 路径属性的前面追加自身的 AS 号,而本实施示例的 BAS 按原状态附加接收时的 MED 属性值,不向 AS 路径属性追加自身的 AS 号,取而代之从 AS 路径属性将自身的 AS 号(或是, BAS 内所预先设定的 AS 号集合的要件)全部去除。

[0125] 下面,设从 ISP 网边缘路由器 2 112-B 向 BAS11-A 发送了路由控制报文(步骤

1021),该路由控制报文的 MED 属性值为 150、AS 路径属性只是 ISP 网的 AS 号 11 并且与目的地址前缀 X 的路由添加有关。接收到该报文的 BAS11-A 与步骤 1011 的情况相同,跳过 AS 路径属性检查处理,然后和从其它配对路由器所接收到的目的地 X 的路由信息做比较,决定最佳路由。此处,由于已从 ISP 网边缘路由器 112-A 接收到 MED 属性值较小的目的地 X 的路由信息,因而本次所接收到的路由信息不成为最佳路由。为此,上行流通量用 BGP 路由表的更新和向用户网边缘路由器 111-A、111-B 的路由信息发送都不进行。

[0126] 下面,设从 ISP 网边缘路由器 1 112-A 向 BAS11-A 发送了与目的地址前缀 X 的路由删除有关的路由控制报文(步骤 1031)。接收到该报文的 BAS11-A 将以前从 ISP 网边缘路由器 1 112-A 所接收到目的地 X 的路由信息废弃,然后对从 ISP 网边缘路由器 1 112-A 之外配对路由器所接收到目的地 X 的各路由信息进行比较,决定最佳路由。此处,在步骤 1021 中从 ISP 网边缘路由器 2 112-B 所接收到的路由信息成为最佳路由。将该更新反映到 BAS11-A 自身的上行流通量用 BGP 路由表中,并将与更新部分相当的路由信息作为路由控制报文发送给用户网边缘路由器 111-A、111-B(步骤 1032-A、1032-B)。此时,与步骤 1012-A、1012-B 的情况相同,按原状态附加接收时的 MED 属性值,不向 AS 路径属性追加自身的 AS 号,取而代之从 AS 路径属性将自身的 AS 号(或是,BAS 内所预先设定的 AS 号集合的要件)全部去除。

[0127] 图 11 是在实施示例 1 的一部分上从用户网方本实施示例的 BAS 接收 BGP 路由控制报文,将其中继到 ISP 网方时的收发顺序图。

[0128] BAS11-D、11-E 在从用户网边缘路由器 111-E 接收到路由信息时,使用 ISP 网的 AS 号来进行路由信息循环的检查,或者全都不进行。另外,在向 ISP 网边缘路由器 112-C 通告路由信息时,按照原来的值将 MED 属性附加到路由信息中,通告其路由信息。如果不包括该动作,则图 11 的收发顺序与采用通常 BGP 路由控制的路由控制报文收发顺序完全相同。还有,在图 11 的顺序开始时刻设为,在用户网边缘路由器 111-E 和 BAS11-D、11-E 之间以及 ISP 网边缘路由器 112-C 和 BAS11-D、11-E 之间的双方已建立 EBGp 配对连接,还未进行任何与地址前缀 Y 有关的路由信息通告。但是,在用户网边缘路由器 111-E 进行路由通告之后进行配对连接确立等场合下,如果有关上述场合以外的路由信息通告也不包括上述属性操作的差异,则可以是与采用通常 BGP 路由控制的收发顺序完全相同的处理。

[0129] 下面,详细说明收发顺序。设从用户网边缘路由器 111-E 向 BAS411-D 发送了路由控制报文(步骤 1111),该路由控制报文其 MED 属性值是 50、AS 路径属性只是用户网的 AS 号 65001 并且与目的地址前缀 Y 的路由添加有关。接收到该报文的 BAS4 11-D 如果是通常的 BGP 路由控制处理,则检查在其报文的 AS 路径属性中是否未包含自身的 AS 号,如果包含则将其路由信息废弃,在本实施示例的 BAS 中,在该检查中使用 ISP 网的 AS 号或者不进行检查。然后,与通常的 BGP 路由控制处理相同,和从其它配对路由器所接收到目的地 Y 的路由信息做比较,决定最佳路由。此处,由于未从其它任何配对路由器再接收到目的地 Y 的路由信息,因而本次所接收到的路由信息成为最佳路由。如果新接收到的路由信息是最佳路由,则将由该路由信息而引起的更新反映到 BAS4 11-D 自身的下行流通量用 BGP 路由表中,将与更新部分相当的路由信息作为路由控制报文,发送给 ISP 网边缘路由器 112-C(步骤 1112)。此时,如果是通常的 BGP 路由控制处理,则不给所发送的路由信息附加接收时的 MED 属性值,而本实施示例的 BAS 却按原状态附加接收时的 MED 属性值。有关 AS 路径属性,

与通常的 BGP 路由控制处理相同,在其前面追加自身的 AS 号。这些 MED 属性和 AS 路径属性等路由控制报文的处理是由处理机 211 来进行的。

[0130] 接着,设从用户网边缘路由器 111-E 向 BAS5 11-E 发送了路由控制报文(步骤 1121),该路由控制报文其 MED 属性值是 100、AS 路径属性只是用户网的 AS 号 65001,并且与目的地址前缀 Y 的路由添加有关。接收到该报文的 BAS5 11-E 进行与步骤 1111 时的 BAS4 11-D 相同的接收处理、最佳路由决定处理以及路由表更新处理,与步骤 1112 相同,将与更新部分相当的路由信息予以发送(步骤 1122)。

[0131] 接着,设从用户网边缘路由器 111-E 向 BAS4 11-D 发送了与目的地址前缀 Y 的路由删除有关的路由控制报文(步骤 1131)。接收到该报文的 BAS4 11-D 将以前从用户网边缘路由器 111-E 所接收到目的地 Y 的路由信息废弃,然后对从用户网边缘路由器 111-E 以外配对路由器所接收到目的地 Y 的各路由信息进行比较,决定最佳路由。此处,由于未从其它任何配对路由器再接收到目的地 Y 的路由信息,因而 BAS4 11-D 失去到目的地 Y 的路由。将该更新也就是到目的地 Y 的路由删除反映到 BAS4 11-D 自身的下行流通量用 BGP 路由表中,将与更新部分相当的路由信息也就是目的地 Y 的路由删除作为路由控制报文,发送给 ISP 网边缘路由器 112-C(步骤 1132)。

[0132] 实施示例 2

[0133] 图 12 表示的是实施示例 2 的网络结构以及本发明的 BGP 路由控制方法。

[0134] 在本实施示例中,企业的主干网 1203 和一个以上支干网 1201-A、1201-B、1201-C,中间相隔本实施示例的 LNS(L2TP Network Server)1212-A、1212-B 被连接。这些网络结构要件具有与实施示例 1 的网络结构要件对应的作用。也就是说,本实施示例的主干网 1203 相当于实施示例 1 的 ISP 网 103。它们的相同之处为都是转接 AS,只是用途各自不同,该转接 AS 用来使支干网和支干网(或者用户网和用户网)之间的流通量(traffic)通过。但是,本实施示例的主干网 1203 即使不是转接 AS 也就是说是存根 AS,本发明也可以使用。在将主干网 1203 设为存根 AS 的场合下,虽然可以实现主干网和支干网之间的通信,但是不能进行支干网和其它支干网之间的通信。本实施示例的支干网 1201 相当于实施示例 1 的用户网 101。它们的相同之处为都是存根 AS,只是用途各自不同,该存根 AS 用来在和主干网(或者 ISP 网)之间只收发自 AS 发或自 AS 收的流通量。本实施示例的 LNS 1212 相当于实施示例 1 的 BAS11。它们的相同之处为都是进行本发明 BGP 路由控制方法的 BGP 路由器,就本实施示例的 LNS 1212 来说,除去支干网方的线路是 L2TP 这种利用隧道协议的数据链路层虚拟线路之外,在功能上都是相同的。

[0135] 主干网 1203、支干网 1201-A、1201-B、1201-C 是作为各自不同的 AS 来构成的,全都具有相互不同的专用 AS 号。另外,与实施示例 1 的 BAS 相同,本实施示例的 LNS 1212-A、1212-B 作为自身的 AS 号全都具有相互相同且与主干网 1203、支干网 1201-A、1201-B、1201-C 的任一个都不同的专用 AS 号。

[0136] 本实施示例的 LNS 1212-A、1212-B 全都与主干网边缘路由器 1213-A、1213-B 的双方相连接。还有,与实施示例 1 相同,连接 1 台 LNS 的主干网边缘路由器必须全部属于相同的 AS。

[0137] 在本实施示例的 LNS 1212-A、1212-B 和支干网 1201-A、1201B、1201-C 之间,设置 LAC(L2TP Access Concentrator)1214-A、1214-B、1214-C。在 LAC 1214 和 LNS 1212 之间

存在 L2TP 网 1204, 是利用 L2TP 进行封装所用的 IP 网, 在 LAC 1214 和 LNS 1212 之间如图所示建立利用 L2TP 的虚拟数据链路层线路。支干网边缘路由器 1211-A、1211-B、1211-C 通过物理或逻辑线路分别与 LAC 1214-A、1214-B、1214-C 相连接, 在该线路上利用 PPP 协议建立有以 1 对 1 方式与 LNS 和 LAC 之间的 L2TP 隧道对应的隧道。据此, 支干网边缘路由器 1211-A、1211-B、1211-C 在 IP 层次上与本实施示例的 LNS 1212-A、1212-B 的各自直接连接。LAC 1214 虽然对于 L2TP 网 1204 内部是作为路由器进行动作的, 但是对于外部则有关 IP 层是透过的, 因此在关于支干网边缘路由器 1211 和 LNS 1212 之间的 BGP 路由控制进行设计时, 可以忽视 LAC 1214 的存在。

[0138] 还有, 当实施本发明时, 在具备本发明 BGP 路由控制方法的 BGP 路由器 (本实施示例的 LAC 和实施示例 1 的 BAS) 和其配对路由器之间, 如果数据链路层上的连接性被保证则也可以使用任何线路。例如, 也可以利用其它隧道协议来连接本实施示例的支干网边缘路由器 1211 和 LNS 1212 之间, 该其它隧道协议是 MPLS 和 IPSec (Security Architecture for Internet Protocol) 等提供数据链路层连接性的协议。

[0139] 在上面的连接线路基础上, 主干网边缘路由器 1213-A、1213-B 全都与本实施示例的 LNS 1212-A、1212-B 双方建立 EBGp 配对连接。支干网边缘路由器 1211-A、1211-B、1211-C 全都与本实施示例的 LNS 1212-A、1212-B 双方建立 EBGp 配对连接。由于本实施示例的 LNS 1212-A、1212-B 进行按照本发明 BGP 路由控制方法的路由控制, 因而可以通过这些配对连接, 在主干网 1203 和支干网 1201-A、1201-B、1201-C 之间进行采用 BGP 的策略路由控制。

[0140] 例如考虑到下述情形, 即采用 BGP 使路由信息流向各支干网, 该路由信息是处于主干网内与到 IP 地址前缀 Z 网络的到达路由有关的信息。但是, 根据主干网 1203 内的网络结构以及主干网 1203 和 LNS 1212 之间的线路带宽等的状况, 设为想按照主干网边缘路由器 1213-A 和 LNS 1212-A 之间的线路、主干网边缘路由器 1213-B 和 LNS 1212-B 之间的线路、主干网边缘路由器 1213-A 和 LNS 1212-B 之间的线路以及主干网边缘路由器 1213-B 和 LNS 1212-A 之间的线路优先次序, 使 IP 数据包通过, 该 IP 数据包具有与地址前缀 Z 匹配的目的地。

[0141] 这种场合下, 可以从主干网边缘路由器 1213-A 向 LNS 1212-A 将 MED 属性值设为 50, 从主干网边缘路由器 1213-B 向 LNS 1212-B 将 MED 属性值设为 100, 从主干网边缘路由器 1213-A 向 LNS 1212-B 将 MED 属性值设为 150, 从主干网边缘路由器 1213-B 向 LNS 1212-A 将 MED 属性值设为 200, 而通告与目的地 Z 有关的 BGP 路由信息。接收到这些路由信息的本实施示例 LNS 1212-A、1212-B 对与目的地 Z 有关的路由信息进行比较来选择最佳路由, 以保持 MED 属性值的原状态将其最佳路由的路由信息向支干网边缘路由器 1211-A、1211-B、1211-C 进行通告。据此, 来自 LNS 1212-A 其 MED 属性值为 50 以及来自 LNS 1212-B 其 MED 属性值为 100 的与目的地 Z 有关的路由信息, 到达支干网边缘路由器 1211-A、1211-B、1211-C 的各自。而且, 从支干网边缘路由器 1211 表明, 本实施示例的 LNS 1212-A、1212-B 属于主干网 1203 的 AS。为此, 支干网边缘路由器 1211 使用 MED 属性来比较双方的路由信息, 以 MED 属性值较小的来自 LNS 1212-A 的路由信息作为最佳路由加以选择。

[0142] 作为结果, 从支干网 1201-A、1201-B、1201-C 发给主干网 1203 内网络 1221 的 IP 数据包流量, 经由 LNS 1212-A 和主干网边缘路由器 1213-A 到达。倘若, 在主干网边缘路

由器 1213-A、LNS 1212-A 及将它们连结的线路任一个中发生故障的场合下,由于 MED 属性值 50 的路由信息被取消, MED 属性值 100 的路由信息成为最佳路由,因而转换成经由 LNS 1212-B 和主干网边缘路由器 1213-B 的路由。

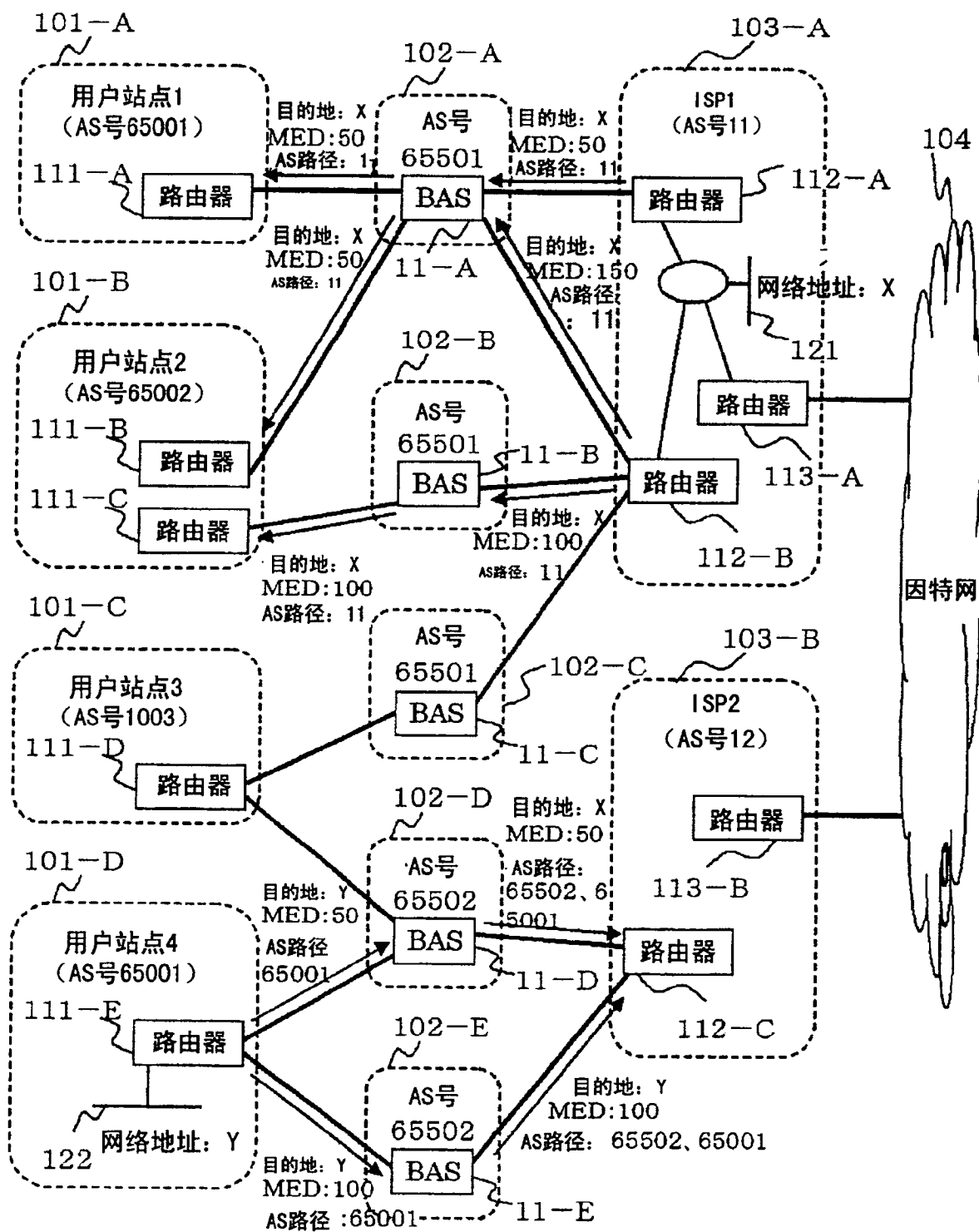


图 1

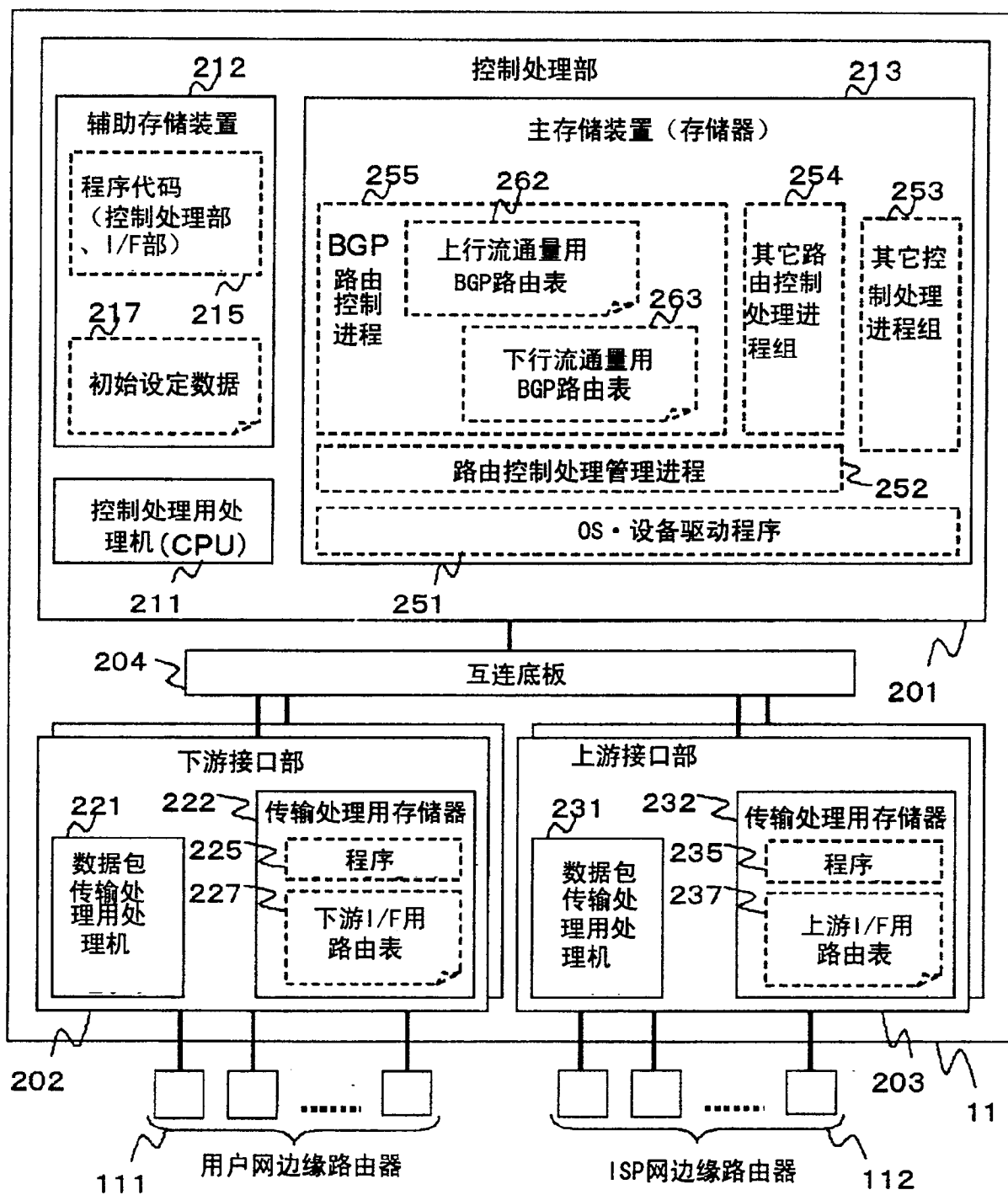


图 2

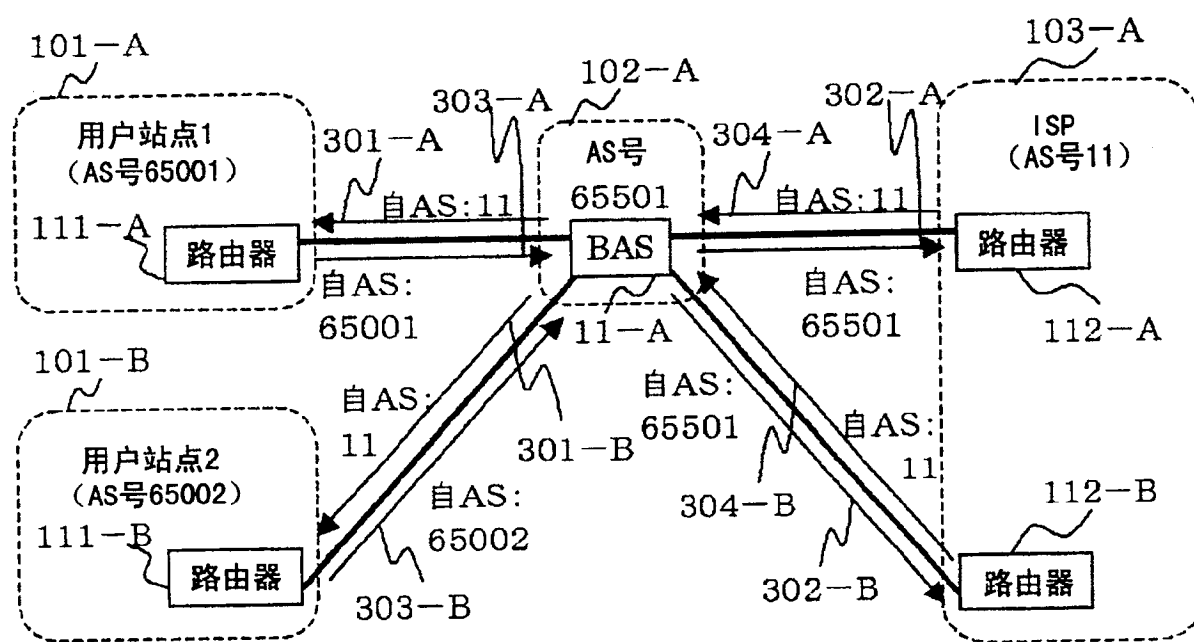


图 3

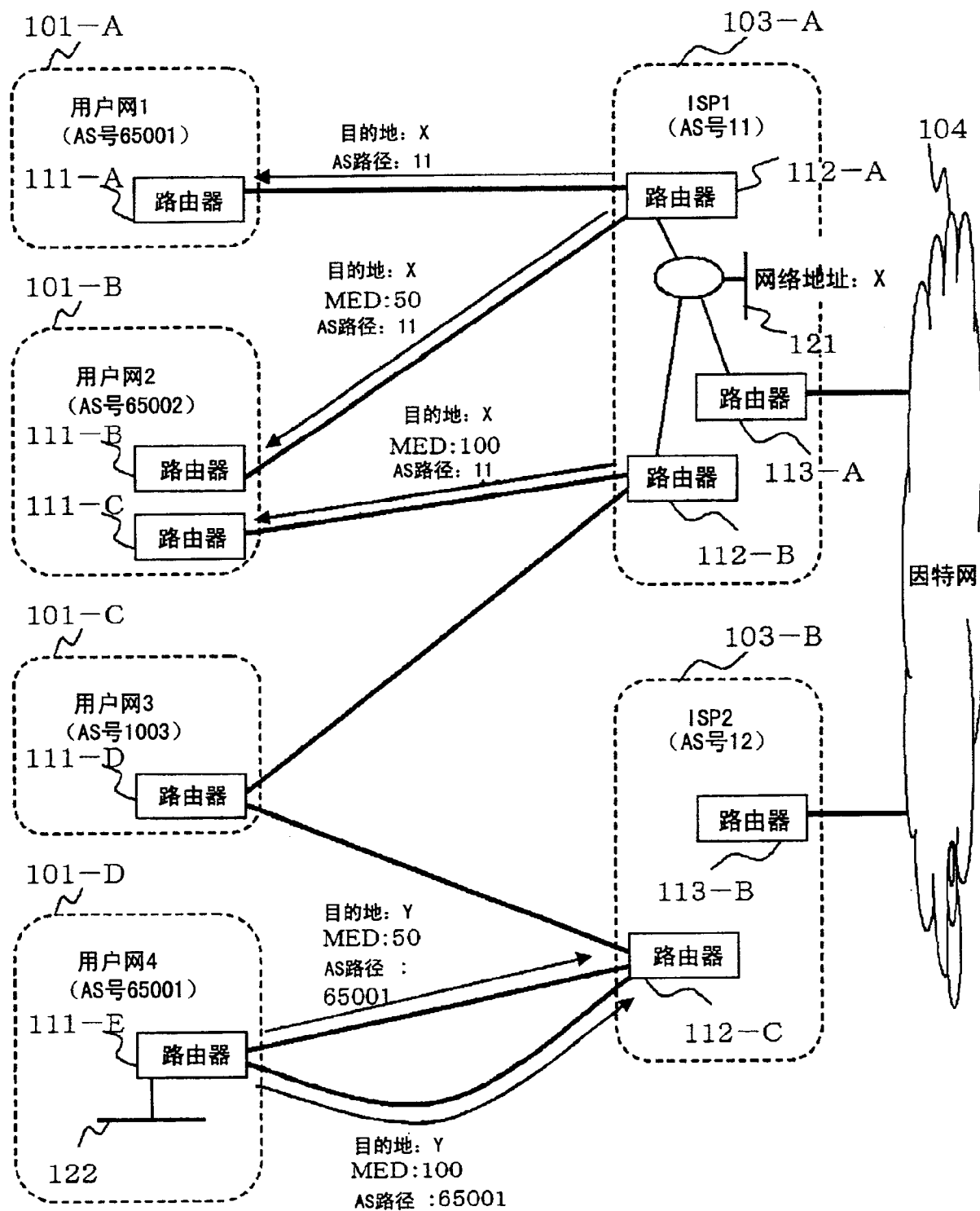


图 4

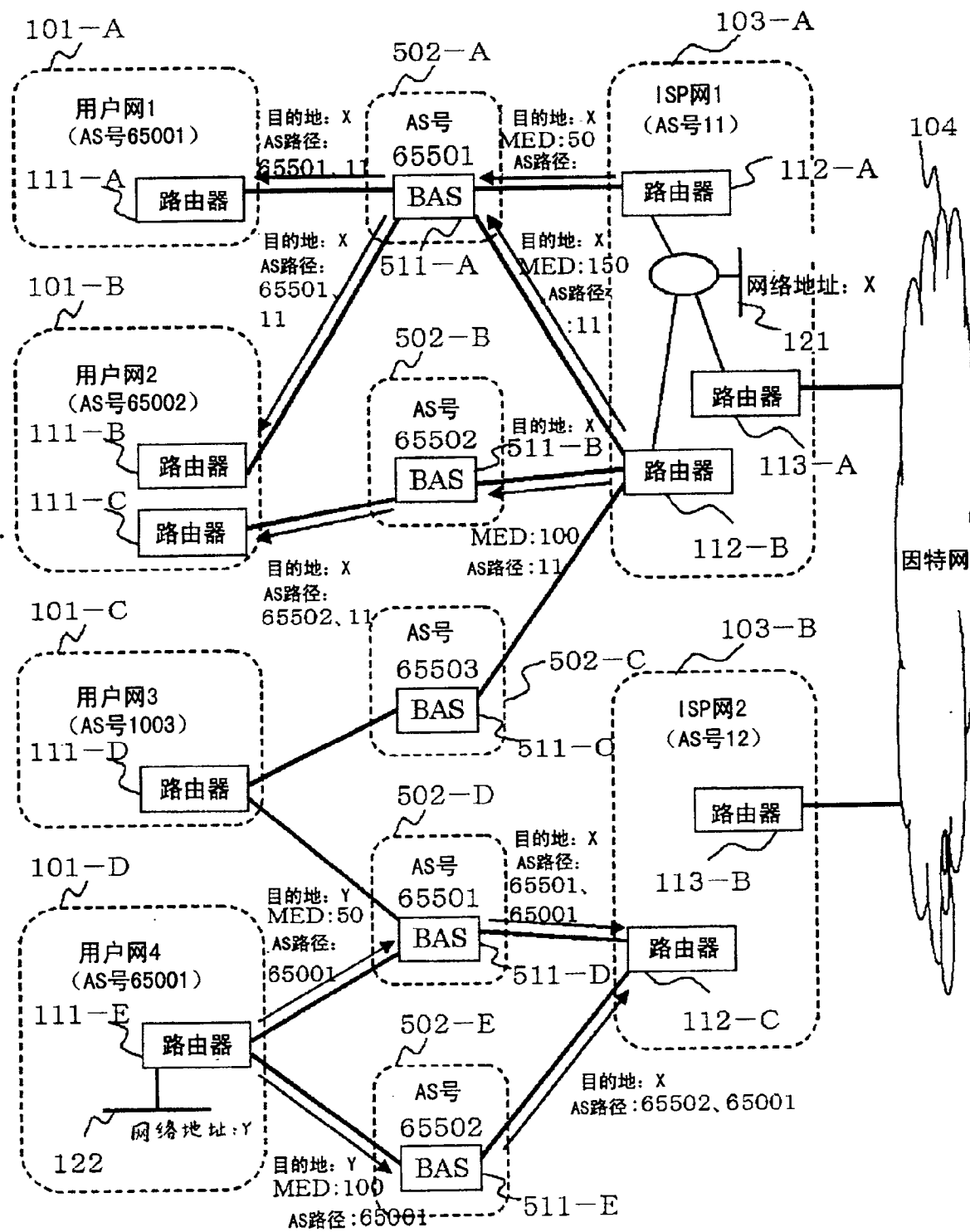


图 5

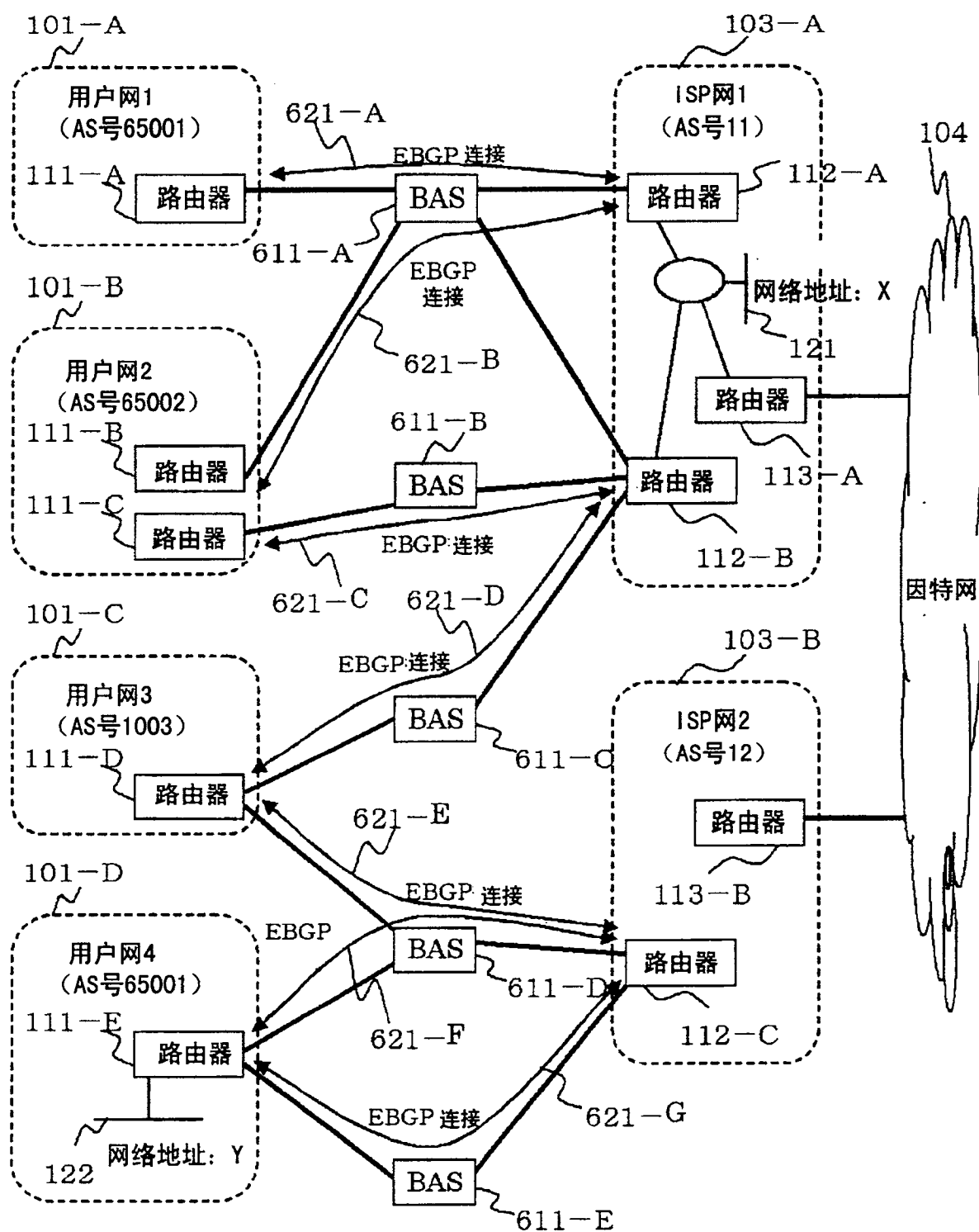


图 6

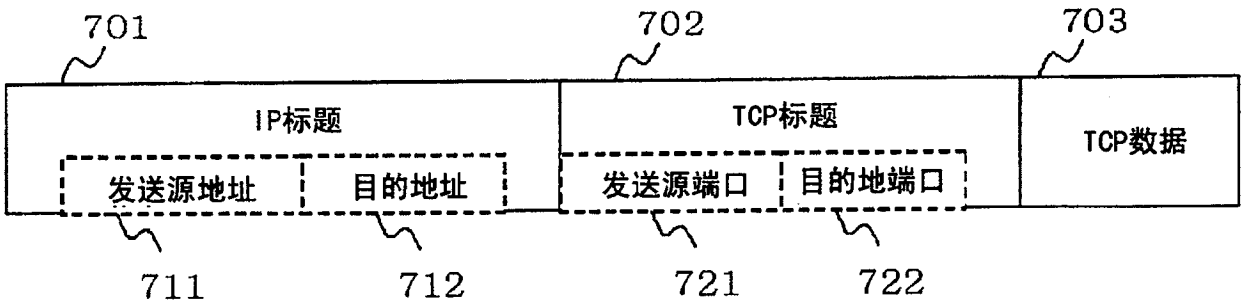


图 7

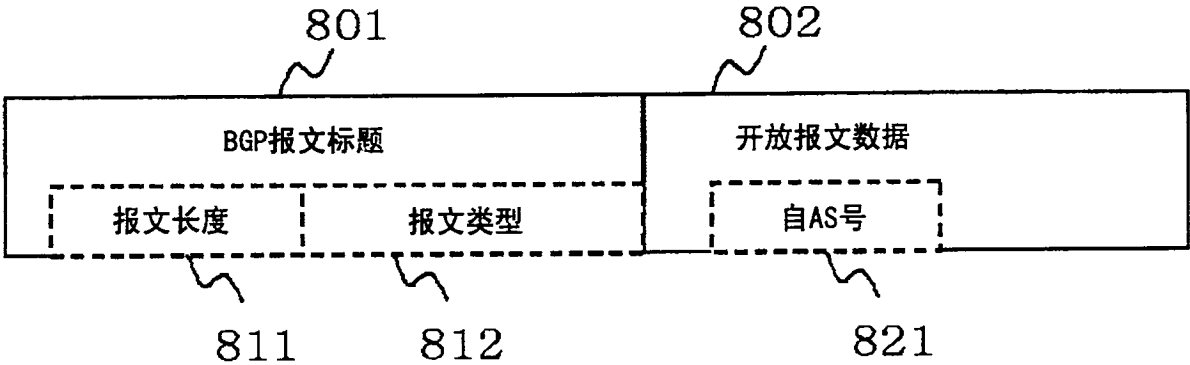


图 8

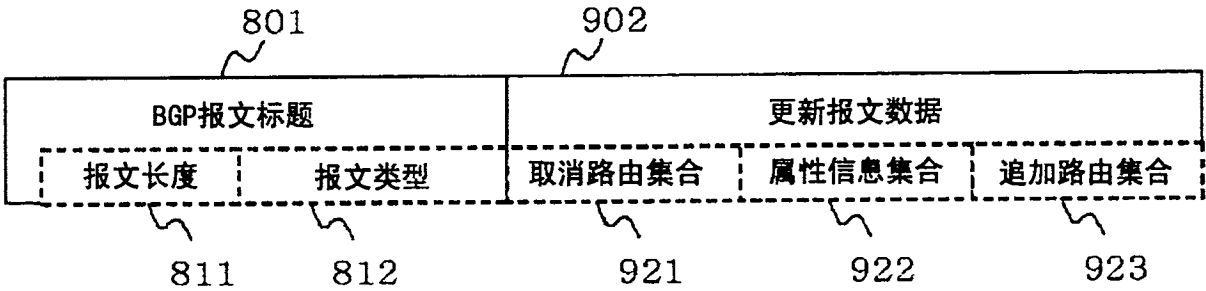


图 9

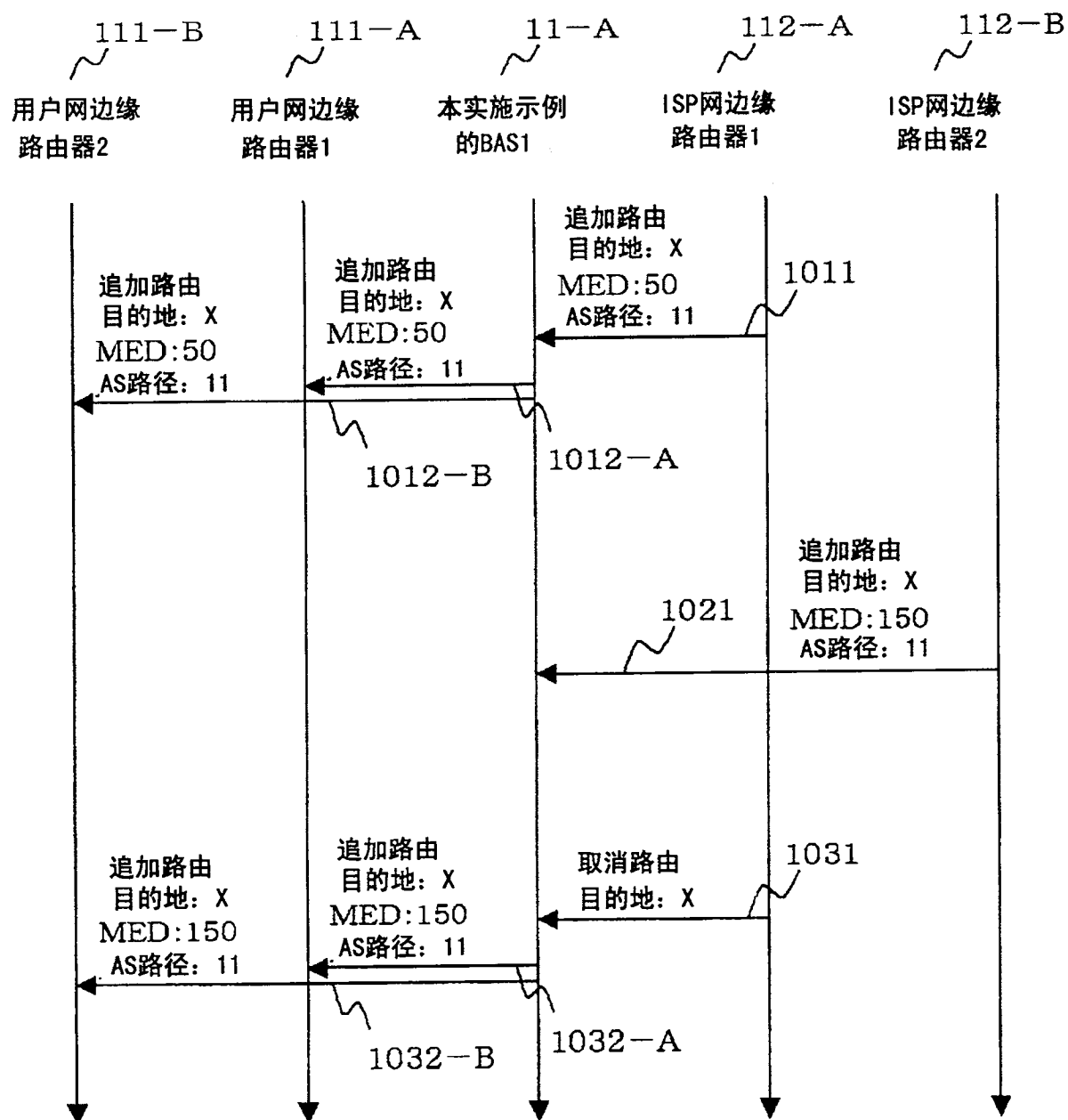


图 10

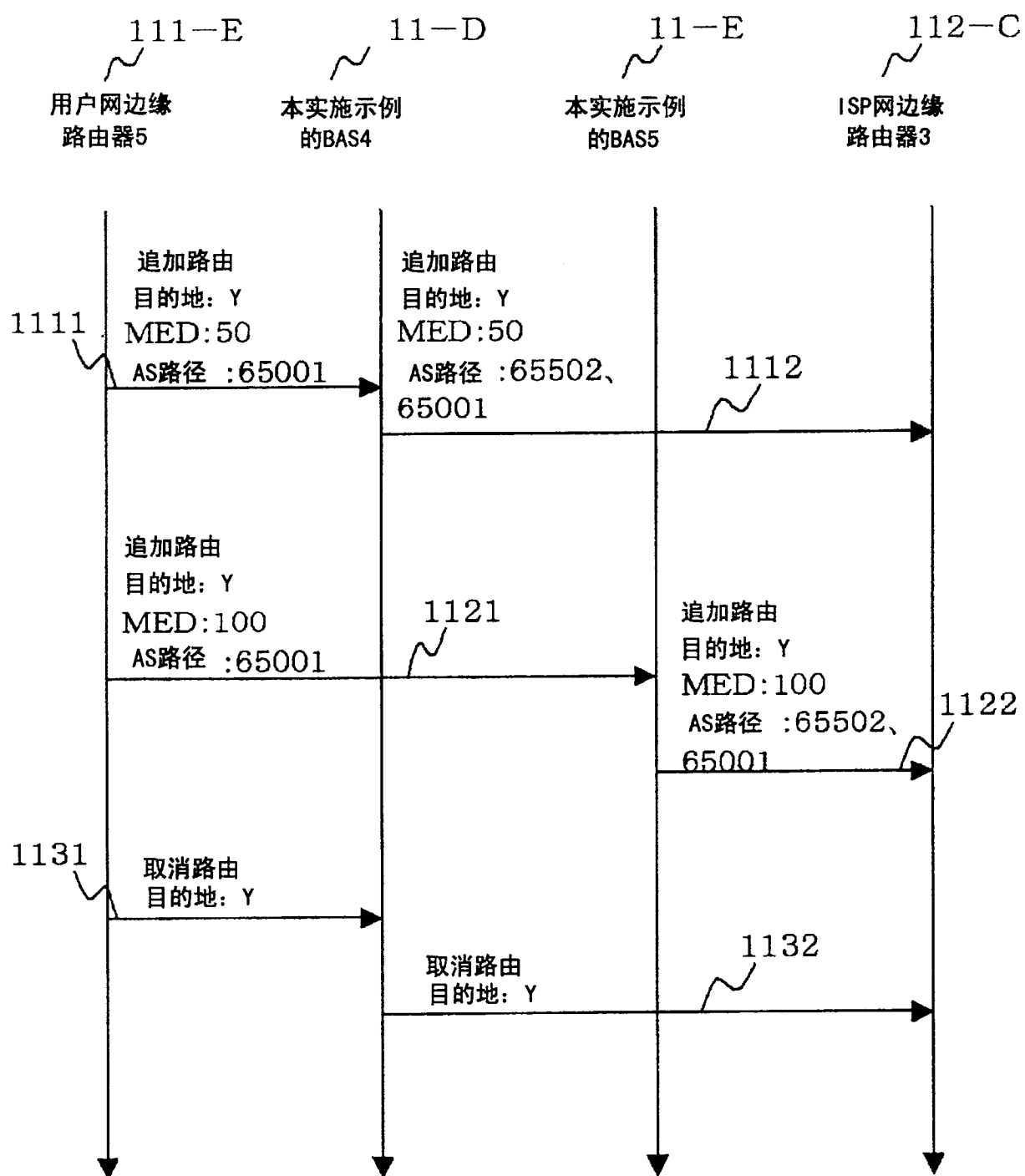


图 11

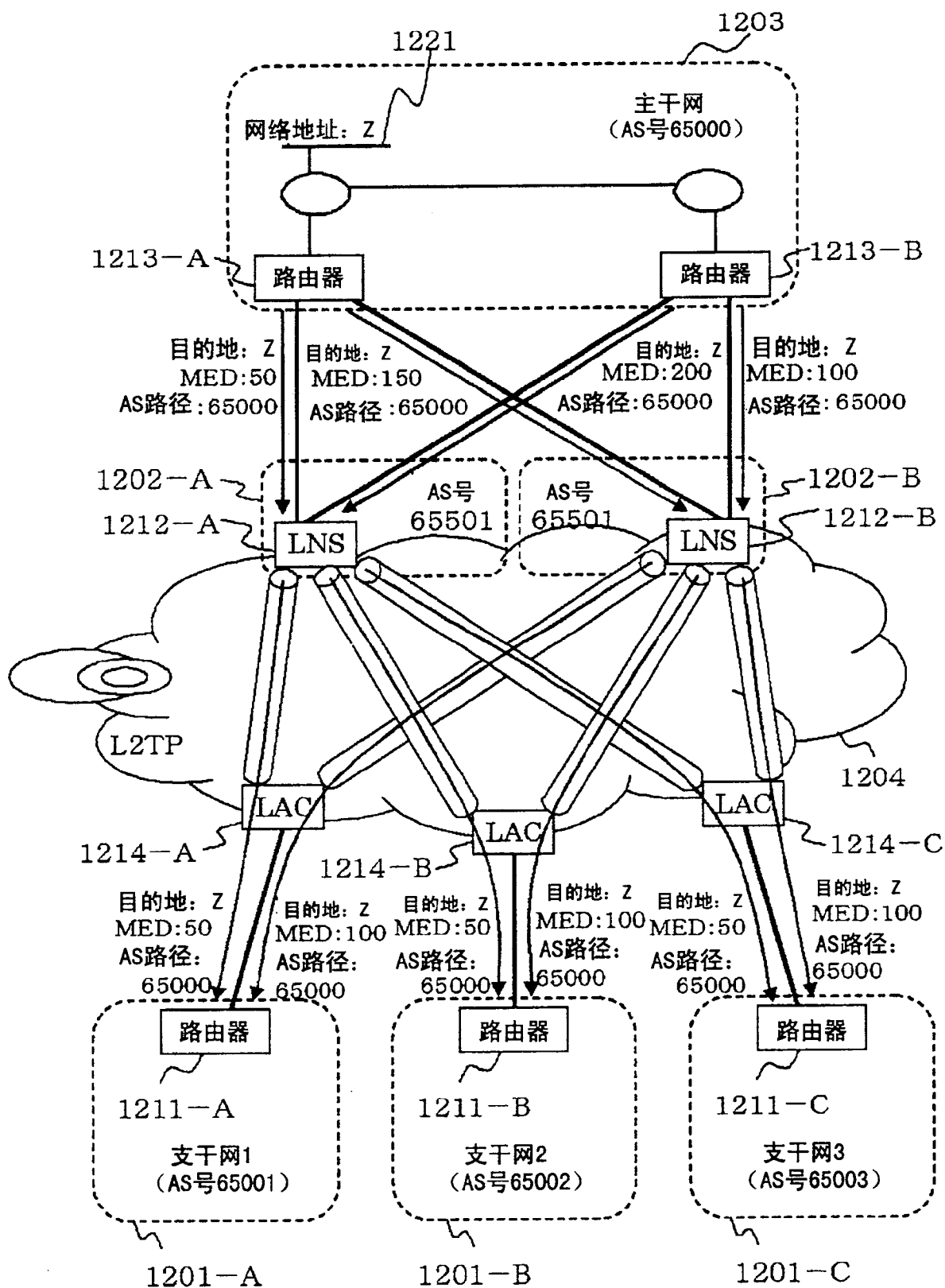


图 12