



(12) 发明专利

(10) 授权公告号 CN 108476146 B

(45) 授权公告日 2022. 02. 25

(21) 申请号 201680057593.2
(22) 申请日 2016.09.06
(65) 同一申请的已公布的文献号
 申请公布号 CN 108476146 A
(43) 申请公布日 2018.08.31
(30) 优先权数据
 62/214,814 2015.09.04 US
(85) PCT国际申请进入国家阶段日
 2018.04.03
(86) PCT国际申请的申请数据
 PCT/US2016/050429 2016.09.06
(87) PCT国际申请的公布数据
 W02017/041107 EN 2017.03.09
(73) 专利权人 动态网络服务股份有限公司
 地址 美国新罕布什尔
(72) 发明人 A·波帕库 P·斯坦霍帕

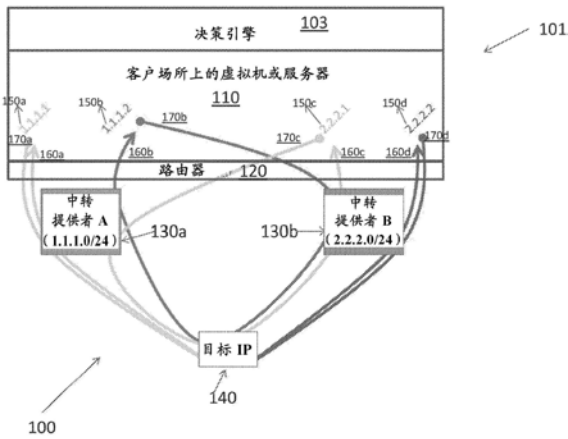
A·梅赫拉 S·胡尔绍夫
B·普雷莫雷 D·马多里
S·M·布拉齐纳 C·巴克尔
R·吉布森
(74) 专利代理机构 中国贸促会专利商标事务所
 有限公司 11038
 代理人 鲍进
(51) Int.Cl.
 H04L 45/121 (2022.01)
 H04L 45/00 (2022.01)
 H04L 43/0864 (2022.01)
(56) 对比文件
 CN 102457933 A, 2012.05.16
 CN 101411141 A, 2009.04.15
 US 2010272010 A1, 2010.10.28
 US 2014064100 A1, 2014.03.06
 审查员 任冠桥
 权利要求书2页 说明书22页 附图14页

(54) 发明名称

使用实时用户监控数据进行实时流量引导的方法和装置

(57) 摘要

常规互联网路由使用诸如边界网关协议 (BGP) 之类的路由协议来处理。但是,简单的BGP不考虑等待时间、分组丢失或成本。为了解决这个问题,实现了智能路由系统,其快速且以经济高效的方式路由流量。在一种做法中,智能路由系统测量、比较和分析客户场所与一个或多个端点之间的往返等待时间和其它度量。基于这些测量,为每个端点选择最优的入站和出站中转提供者。其它智能路由系统收集并分析真实用户监控 (RUM) 数据,以基于客户端的IP地址和内容起源的IP地址预测用于向特定客户端提供数据的不同内容起源的等待时间性能,其中所述IP地址按照性能进行排名。通过基于与IP地址相关联的性能来解析域名系统 (DNS) 查询,这些排名用于沿着较低等待时间路径引导流量。



1. 一种在多个中转提供者之间将来自客户场所的流量指引到互联网协议 (IP) 地址的方法, 所述方法包括:

测量在第一路径上从客户场所处的收集器设备到所述IP地址的第一等待时间, 其中所述第一路径包括所述多个中转提供者中的第一中转提供者以及耦合到所述第一中转提供者的第一边界路由器;

测量在第二路径上从所述收集器设备到所述IP地址的第二等待时间, 其中所述第二路径包括所述多个中转提供者中的第二中转提供者、第二边界路由器以及所述第一边界路由器;

执行所述第一等待时间与所述第二等待时间的比较;

基于所述第一等待时间与所述第二等待时间的比较来选择所述第二中转提供者; 以及

将来自所述客户场所的流量指引到所述IP地址, 其中将所述流量指引到所述IP地址包括经由所述第一边界路由器和所述第二边界路由器将所述流量指引到所述第二中转提供者。

2. 如权利要求1所述的方法, 其中测量第一等待时间包括:

经由第一中转提供者向所述IP地址发送回声请求; 以及

经由第二中转提供者接收回声应答。

3. 如权利要求1所述的方法, 其中测量第一等待时间包括:

经由耦合到所述第一中转提供者的边界路由器的第一接口向所述IP地址发送回声请求; 以及

经由所述边界路由器的第二接口接收回声应答。

4. 如权利要求1所述的方法, 其中执行所述第一等待时间与所述第二等待时间的比较包括:

比较所述第一中转提供者的入站等待时间与所述第二中转提供者的入站等待时间; 以及

将所述第一中转提供者的出站等待时间与所述第二中转提供者的出站等待时间进行比较。

5. 如权利要求1所述的方法, 其中选择所述第二中转提供者还包括:

基于所述第一中转提供者的成本和所述第二中转提供者的成本选择所述第一中转提供者和所述第二中转提供者中的一个。

6. 如权利要求1所述的方法, 其中选择所述第二中转提供者还包括:

基于所述第一中转提供者的分组丢失和所述第二中转提供者的分组丢失选择所述第一中转提供者和所述第二中转提供者中的一个。

7. 如权利要求1所述的方法, 其中将所述流量指引到所述IP地址包括:

将分组的主机前缀与边界网关协议 (BGP) 社区属性相关联; 以及

基于所述主机前缀的BGP社区属性将所述分组指引到所述第一中转提供者和所述第二中转提供者中的一个。

8. 如权利要求1所述的方法, 其中将所述流量指引到所述IP地址包括:

将去往所述IP地址的流量的下一跳设置为耦合到所述第一中转提供者和所述第二中转提供者中的一个的边界路由器。

9.一种在多个中转提供者之间将来自客户场所的流量指引到互联网协议(IP)地址的方法,所述方法包括:

测量在第一路径上从客户场所处的收集器设备到所述IP地址的第一等待时间,其中所述第一路径包括所述多个中转提供者中的第一中转提供者以及耦合到所述第一中转提供者和第二中转提供者的边界路由器上的第一接口;

测量在第二路径上从所述收集器设备到所述IP地址的第二等待时间,其中所述第二路径包括所述多个中转提供者中的第二中转提供者以及所述边界路由器上的第二接口;

执行所述第一等待时间与所述第二等待时间的比较;

基于所述第一等待时间与所述第二等待时间的比较来选择所述第二中转提供者;以及

经由所述第二中转提供者将来自所述客户场所的流量指引到所述IP地址,其中将所述流量指引到所述IP地址包括经由所述边界路由器上的第二接口将所述流量指引到所述第二中转提供者。

10.一种电子设备,包括一个或多个处理器,其被配置为执行如权利要求1-9中任一项所述的方法。

11.一种在多个中转提供者之间将来自客户场所的流量指引到互联网协议(IP)地址的装置,包括用于执行如权利要求1-9中任一项所述的方法的步骤的组件。

12.至少一种其中存储有指令的非临时性计算机可读存储介质,所述指令在被一个或多个处理器执行时,使得所述一个或多个处理器执行根据权利要求1-9中任一项所述的方法。

使用实时用户监控数据进行实时流量引导的方法和装置

[0001] 对(一个或多个)相关申请的交叉引用

[0002] 本申请依据35U.S.C.§119(e)要求于2015年9月4日提交的标题为“Methods and Apparatus for Real-time Traffic Steering and Real-User Monitoring Data”的美国申请No.62/214,814的优先权益,其公开内容通过引用整体上并入本文。

背景技术

[0003] 互联网上的路径选择主要依赖于边界网关协议(BGP)。BGP是一种标准化的可扩展协议,其允许在跨互联网的自治系统之间进行通信。BGP使用确定性规则集合在端点之间选择路径,并被分类为距离向量路由协议。

[0004] 在选择一对端点之间的路径时,BGP不考虑等待时间、分组丢失和成本。因此,在一些电路中,等待时间会低得令人无法接受。在其它情况下,可以存在高水平的分组丢失,这会导致通信低效并实际上增加等待时间。最后,BGP对中转提供商(transit provider)可能关心的商业考虑(包括成本)以及优选的中转提供者视而不见。

[0005] 目前,存在一些优化BGP路由中的第一跳的做法。特定于路由器的解决方案包括Cisco的IP-SLA以及Juniper的RPM解决方案。其它做法或者是特定于硬件的(Notion)或者驻留在中转提供者的层面(Internap)来全局地评估流量流并相应地修改BGP表。

[0006] 目前还没有整体或全面的解决方案,其相对独立于路由器并且可以集成无限数量的输入来帮助确定路径选择。而且,目前没有解决方案在目标服务器上没有资产的情况下考虑或控制返回路径。

[0007] 域名系统(DNS)可以从基于文本的域名映射互联网协议(IP)空间中的逻辑端点/地址。当需要到达目标时,有一些方法可以基于源地理位置选择目标。

[0008] 但是,由于互联网的变幻莫测,来自不同源的目标可用性可以各不相同:服务器可以不堪重负,路径可以拥塞或不可用。因此,必须有方法来实时地确定针对给定用户或市场的最佳目标位置。

发明内容

[0009] 本技术解决与用于在互联网上路由的路径选择和目标选择确定相关联的问题,以及其它问题。这种技术的实施例包括用于在多个中转提供者之间将来自客户场所的流量指引到互联网协议(IP)地址的方法和装置。该系统包括在客户场所处的收集器设备,该收集器设备测量经由多个中转提供者中的第一中转提供者到该IP地址的第一等待时间,并且测量从该收集器设备经由多个中转提供者中的第二中转提供者到该IP地址的第二等待时间。耦合到收集器设备或由收集器设备实现的决策引擎执行第一等待时间与第二等待时间的比较,并基于该比较来选择第一中转提供者或第二中转提供者。路由器、交换机或其它设备经由选定的中转提供者将来自客户场所的流量指引到该IP地址。

[0010] 在一些情况下,收集器设备通过经由第一中转提供者向IP地址发送回声请求并且经由第二中转提供者接收回声应答来测量第一等待时间。收集器设备还可以通过经由耦合

到第一中转提供者的边界路由器的第一接口向IP地址发送回声请求并经由该边界路由器的第二接口接收回声应答来测量第一等待时间。收集器设备可以经由耦合到第一中转提供者的第一边界路由器测量第一等待时间,并且经由耦合到第一边界路由器和第二中转提供者的第二边界路由器测量第二等待时间。在这种情况下,路由器可以经由第一边界路由器和第二边界路由器将流量指引到第二中转提供者。并且收集器设备可以经由耦合到第一中转提供者和第二中转提供者的边界路由器上的第一接口测量第一等待时间,并且经由该边界路由器上的第二接口测量第二等待时间。在这种情况下,路由器可以经由第一边界路由器和第二边界路由器将流量指引到第二中转提供者。

[0011] 决策引擎可以通过比较第一中转提供者的进站等待时间与第二中转提供者的进站等待时间来执行第一等待时间与第二等待时间的比较。决策引擎还可以将第一中转提供者的出站等待时间与第二中转提供者的出站等待时间进行比较。

[0012] 决策引擎还可以基于第一中转提供者的成本和第二中转提供者的成本来选择第一中转提供者或第二中转提供者。决策引擎还可以基于第一中转提供者的分组丢失和第二中转提供者的分组丢失来选择第一中转提供者或第二中转提供者。

[0013] 路由器、交换机或其它路由设备可以通过将分组的主机前缀与边界网关协议(BGP)社区属性相关联并且基于BGP社区属性将该分组指引到选定的中转提供者来将流量指引到IP地址。路由器、交换机或其它设备还可以通过将用于去往该IP地址的流量的下一跳设置为耦合到第一中转提供者和第二中转提供者中的一个的边界路由器来将流量指引到该IP地址。

[0014] 本技术的其它实施例包括用于测量目标IP地址与客户场所之间的等待时间的方法和系统,其中客户场所包含通知第一前缀并连接到第一中转提供者的第一路由器和通知第二前缀并连接到第二中转提供者的第二路由器。客户场所处的收集服务器经由第二路由器和第二中转提供者将来自具有第一前缀的第一IP地址的第一回声请求发送到目标IP地址。第一回声请求包括指示第一回声请求何时由第一IP地址发送的第一时间戳。收集服务器经由第一中转提供者和第一中转提供者从目标IP地址接收第一回声应答。第一回声应答包括指示第一回声应答何时由目标IP地址发送的第一时间戳应答。耦合到收集服务器或由收集服务器实现的决策引擎基于第一时间戳和第一时间戳应答来确定第一往返等待时间。

[0015] 在一些情况下,收集服务器还经由第一路由器和第一中转提供者从客户场所处的第二IP地址向目标IP地址发送第二回声请求。这个第二回声请求包括指示第二回声请求何时由具有第二前缀的第二IP地址发送的第二时间戳。收集服务器经由第二中转提供者和第二中转提供者接收来自目标IP地址的第二回声应答。第二回声应答包括指示第二回声应答何时由目标IP地址发送的第二时间戳应答。

[0016] 在这些情况下,决策引擎基于第二时间戳和第二时间戳应答来确定第二往返等待时间。它还可以执行第一等待时间与第二等待时间的比较,并且基于第一等待时间与第二等待时间的比较来选择第一中转提供者或第二中转提供者。耦合到决策引擎的路由器、交换机或其它路由设备经由选定的中转提供者将来自客户场所的流量指引到IP地址。

[0017] 本技术的其它实施例包括用于响应域名系统(DNS)请求的系统和方法。示例方法包括从递归解析器接收DNS请求,以解析具有存储在多个内容起源中的每一个内容起源处的内容的域,该域可以提供来自内容递送网络(CDN)或云提供者的内容。响应于对DNS请求,

权威DNS服务器或其它处理器从IP地址层次结构中选择IP地址集合。这个IP地址集合包含递归解析器的IP地址和至少预定数量的样本。权威DNS服务器基于(i) 该IP地址集合和(ii) 多个内容起源中的与该IP地址集合相关联的内容起源的排名来从多个内容起源中选择内容起源。权威DNS服务器将所选内容起源的IP地址发送到递归解析器。

[0018] 本技术的又一个实施例包括实时监控由客户端进行的下载的方法。在一个示例中,由客户端执行的软件(可执行代码)使客户端识别多个递归解析器中用于解析针对内容递送网络(CDN)的域名系统(DNS)请求的递归解析器的IP地址。客户端测量用于解析DNS请求的解析时间,响应于DNS请求而识别由递归解析器返回的CDN的统一资源定位符(URL),并测量用于从内容起源向客户端下载内容的下载时间。客户端可以响应于来自客户端的对内容的请求而下载并执行这个软件。

[0019] 客户端可以测量识别递归解析器的IP地址,测量DNS解析时间,识别URL,并且测量用于从多个CDN中的每一个CDN进行下载的下载时间。客户端或另一个处理器(例如,权威DNS服务器)可以针对多个CDN执行解析时间和下载时间的比较。例如,可以基于用于每个递归解析器的解析时间和下载时间为那个递归解析器进行CDN排名。递归解析器可以使用这个比较来对后续的DNS请求做出响应。

[0020] 应当认识到的是,前述概念和以下更详细讨论的附加概念的所有组合(假设这些概念并非相互不一致)被预期是本文公开的发明性主题的一部分。特别地,在本公开结尾处出现的要求保护的主题的所有组合被认为是本文公开的发明性主题的一部分。还应当认识到的是,本文明确采用的术语也可以出现在通过引用并入本文的任何公开中,如果符合与本文公开的特定概念最一致的的含义的话。

附图说明

[0021] 本领域技术人员将理解的是,附图主要是用于说明的目的,而并非旨在限制本文所述的发明性主题的范围。图示不一定按比例绘制;在一些情况下,本文公开的发明性主题的各个方面可以在图示中被夸大或放大,以便于对不同特征的理解。在附图中,相同的标号一般指相同的特征(例如,功能上相似和/或结构上相似的元件)。

[0022] 图1A示出了一种系统,该系统监控和测量沿着由不同中转提供者提供的、客户场所与目标互联网协议(IP)地址之间的入站和出站路径的等待时间。

[0023] 图1B是图1A中所示的不同入站和出站路径的等待时间的矩阵比较。

[0024] 图2示出了经由通过基于等待时间测量选择中转提供者而将流量引导(steering)到端点的处理。

[0025] 图3示出了测量经由不同中转提供者在一对IP地址之间的往返等待时间的处理。

[0026] 图4示出了基于一对IP地址之间的往返等待时间来选择入站和出站中转提供者的处理。

[0027] 图5A图示了一个网络,其中将路由图(route-map)应用于边界路由器接口,以监控收集器与若干目标IP地址之间的等待时间。

[0028] 图5B图示了手动地引导在收集器与若干目标IP地址之间经由静态路由的网络中的流量。

[0029] 图5C图示了通过基于测得的中转提供者的等待时间将BGP社区属性与不同的中转

提供者动态地关联来引导流量。

[0030] 图6是绘出用于创建路由图并将路由图应用于边界路由器接口以进行等待时间和分组丢失监控的处理的流程图。

[0031] 图7是绘出进行基于策略的静态路由以基于等待时间测量来引导流量的流程图。

[0032] 图8是绘出进行自动覆盖以基于等待时间测量来引导流量的流程图。

[0033] 图9图示了用于收集实时用户监控 (RUM) 数据并基于RUM数据和类型标记IP集 (TYLIPS) 数据引导流量的系统。

[0034] 图10和11示出了用于收集和聚合RUM数据的处理。

[0035] 图12示出了根据图10和11中所示的处理对RUM数据进行聚合而形成的直方图。

[0036] 图13示出了用于生成如图12中所示的直方图的处理。

[0037] 图14A和14B示出了图示用于选择TYLIPS数据以解析DNS请求的不同层次结构的Venn图。

[0038] 图15示出了基于RUM数据的客户端IP地址的示例性排名。

[0039] 图16示出了基于RUM数据引导流量的处理。

具体实施方式

[0040] 固定端点之间基于等待时间、分组丢失和成本的智能路由

[0041] 分组根据边界网关协议 (BGP) 在互联网上被路由, BGP使用规则和表来确定分组到给定目的地的路由上的“下一跳”。遗憾的是, 简单的BGP在路由分组时没有考虑等待时间、分组丢失或成本。这对于期望低等待时间的应用 (诸如针对以每次印象 (per-impression) 为基础出售的互联网广告库存的实时出价) 会导致次优路由。利用实时出价, 广告买方可以对通过用户访问发布者网站而被触发的印象进行出价。如果买方赢得了拍卖, 那么买方的广告会立即显示在发布者的网站上。因为拍卖是响应于用户访问发布者的网站时而实时发生的, 所以它持续很短时间, 因此快速出价就势在必行。如果买方的出价直到拍卖结束之后 (例如, 由于从买方向拍卖方传输出价的等待时间) 才到达拍卖方, 那么买方肯定也失去拍卖。

[0042] 幸运的是, 本文公开的智能路由系统和方法可以测量两个端点 (例如, 买方的IP地址和拍卖方的IP地址) 之间的等待时间并选择那些端点之间的最低等待时间路由。在很多情况下, 这些路由由不同的中转提供者提供, 并且决策引擎或智能路由经由提供具有最低可用等待时间的路由的中转提供者将分组发送到目的地IP地址。在一些情况下, 决策引擎可以基于用户可配置的、等待时间和其它因素 (诸如每个中转提供者的成本和分组丢失) 的加权组合来选择中转提供者。这种中转提供者的智能选择有效地覆盖了BGP策略。由于中转提供者的相对等待时间、分组丢失和成本改变, 因此决策引擎通过选择在给定时间的“最佳”路径来相应地重新路由流量。

[0043] 虽然其它做法优化了出站路径, 但它们通常或者特定于路由器或者在出站目标方面被一般化。本文公开的做法可以被实现为独立于路由器并针对具体的目标或端点而量身定制的软件层。与常规做法不同, 本发明性过程还可以考虑变化的入站流和倒数第二个入站步骤的优化。

[0044] 智能路由系统

[0045] 图1A是系统100的示意图,系统100监控和测量客户场所101(例如,实时出价处理中的买方)处的虚拟机或服务器110与一个或多个目标IP地址140(例如,拍卖方或发布者站点)之间的等待时间和分组丢失。系统沿着入站路径160a-160d(统称为流入或入站路径160)和出站路径170a-170d(统称为流出或出站路径170)监控这些参数,并相应地在不同路径之间路由流量。在这个示例中,服务器110耦合到路由器120,路由器120进而经由中转提供者130a和130b(统称为中转提供者130)将流量路由到目标IP地址140。系统100还包括耦合到虚拟机/服务器110或由虚拟机/服务器110实现的决策引擎103。决策引擎103基于由虚拟机110测量得的中转提供者130的等待时间和分组丢失性能来选择用于在客户场所101与目标IP地址140之间承载流量的中转提供者130。

[0046] 客户场所101包括被指派有多个始发IP地址150a-150d(统称为始发IP地址150)的一个或多个虚拟机或服务器110。这些始发IP地址150中的每一个用于测量由中转提供者130提供的唯一入站/出站路径对的等待时间和分组丢失。每个中转提供者130提供入站路径160和出站路径170两者,因此,如果存在将客户场所101连接到目标IP地址140的N个中转提供者130,那么向虚拟机110指派 N^2 个不同的始发IP地址150。换句话说,所选择的始发IP地址150表示客户场所101与目标IP地址140之间经由中转提供者130的唯一往返路径的总数。

[0047] 如果客户场所101足够大,那么它可以通告一个或多个唯一的主机前缀。在IPv4系统中,每个唯一的路由前缀可以是“/24”,即,路由前缀具有为网络前缀分配的24位和为主机寻址保留的8位。/24是BGP将路由到的最小前缀。如图1A中所示,来自始发IP地址150的与唯一/24对应的分组通过相应的中转提供者130被信道传输。用于每个中转提供者130的唯一/24通过识别最小前缀BGP将经由那个中转提供者130路由来确定的。在这个示例中,始发IP地址150a和150b(1.1.1.1和1.1.1.2)与在中转提供者130a后面通告的/24(1.1.1.0/24)对应,并且始发IP地址150c和150d(2.2.2.1和2.2.2.2)与在中转提供者130b后面通告的/24(2.2.2.0/24)对应。

[0048] 路由器120根据它存储在其内部存储器上的BGP表通过中转提供者130将流量路由到目标IP地址140。如上所述,这些BGP表用于将始发IP地址150指派给虚拟机110,用于探测每个中转提供者130的时变等待时间和分组丢失。并且如下所述,取决于目标IP 140的等待时间、分组丢失和成本的加权组合,这些BGP表可以由决策引擎130覆盖或修改,以强制流量通过给定的中转提供者130。

[0049] 入站路径160和出站路径170经由中转提供者130连接起始IP地址150和目标IP地址140。取决于分组源自的始发IP地址150,沿着入站路径160行进的分组通过中转提供者130被路由。在这个示例中,入站路径160a和160b上的分组通过中转提供者130a被路由,而入站路径160c和160d上的分组通过中转提供者130b被路由。通过中转路径130a信道传输连接始发IP地址150a和目标IP地址140的出站路径170a上的分组。并且通过中转提供者130b信道传输连接始发IP地址150b和目标IP 140的出站路径170b上的分组。类似地,分别通过中转提供者130a和130b信道传输分别从IP地址150c和150d连接到目标IP地址140的出站路径170c和170d。

[0050] 决策引擎103比较中转提供者130的等待时间、分组丢失和/或成本,并基于这些比较实现智能路由,如下面更详细描述。决策引擎103修改或覆盖BGP策略,以强制流量通过

提供到目标IP地址140的最快、最成本有效的路由的中转提供者130。决策引擎103可以被实现为独立于路由器的软件层并且针对目标IP地址140量身定制。在一些实施例中,决策引擎103可以包括在服务器110中。在其它实施例中,决策引擎103可以包括在路由器120中。它也可以在另一个处理器上实现。

[0051] 图1B图示了表示与经由图1A中所示的中转提供者130发送和接收分组相关联的等待时间的矩阵105。矩阵105列出通过中转提供者130到目标IP 140的进站等待时间115和出站等待时间125。可以将这些等待时间115、125求和并进行比较,以识别客户场所101与目标IP地址140之间的最低等待时间往返路径。

[0052] 在这个示例中,矩阵105包括经由中转提供者130a和130b到目标IP 140的出站路径170和来自目标IP 140的进站路径160的每个组合的往返等待时间测量。由于存在两个中转提供者130,因此矩阵105中的路径总数为 $2^2=4$ 。在这个示例中,矩阵105示出了用于每个始发IP地址150的流入和流出等待时间,这些等待时间唯一地映射到不同的进站和出站路径,如上面解释和图1A中所示的。比较矩阵105中的往返等待时间显示客户场所101与目标IP地址140之间经由中转提供者130b的分组传输给出最低等待时间(12ms)。基于这个等待时间测量,到中转提供者130b的进站最后一跳由BGP设置,并且到中转提供者130b的出站第一跳由基于策略的路由来控制。

[0053] 虽然图1B中的矩阵105仅示出等待时间测量,但本领域的普通技术人员将容易理解的是,除了等待时间测量之外或代替等待时间测量,还可以包括其它测量。例如,矩阵还可以包括分组丢失、成本或者等待时间、分组丢失和成本的加权组合。类似地,矩阵105可以被扩展到任何数量的中转提供者和目标IP地址。

[0054] 在端点之间测量等待时间和引导

[0055] 图2是示出用于基于等待时间测量来经由选定的中转提供者路由流量的处理200的流程图。这个处理200可以使用图1A中所示的系统100或者任何其它合适的系统或网络(包括图5A-5C中所示的那些)来实现。处理200涉及针对单个目标IP地址测量经由不同中转提供者的等待时间。例如,给定两个中转提供者-中转提供者1和中转提供者2,在步骤210中测量经由每个中转提供者到目标IP的等待时间。在步骤220中,比较测得的等待时间,以选择到目标IP地址的具有最小等待时间的中转提供者。在一些情况下,外部度量(诸如初始成本和分组丢失)也被考虑到比较中(230)。在比较(220)之后,在步骤240中选择用于目标IP地址的适当的中转提供者。经由选定的中转提供者来引导流量(250)。出站第一跳通过使用路由图的基于策略的路由或通过使用与中转提供者相关联的BGP属性进行引导来控制。

[0056] 虽然图2仅示出了两个中转提供者和单个目标IP地址,但是本领域的普通技术人员将容易认识到的是,这个处理可以应用于多于两个中转提供者和多于一个IP地址。例如,它可以并行地应用于若干IP地址中的每一个,可能用于到每个IP地址的不同数量的中转提供者(例如,两个中转提供者到第一IP地址,五个提供者到第二IP地址,等等)。单个中转提供者可以服务于多个IP地址,但不需要服务于每一个IP地址。

[0057] 通过Ping端点(目标/目的地IP地址)来测量等待时间

[0058] 图3是图示用于例如使用图1A中所示的虚拟机110来收集进站和出站等待时间数据的处理300的流程图。在客户场所310上的虚拟机或裸金属(bare metal)上建立收集器(收集服务器)。收集器通过ping在由客户提供的目标IP地址列表上的每个目标IP地址来收

集等待时间数据。如上所述,这些ping可以源自不同的源IP地址,以探测不同的前向和返回路径。

[0059] Ping如下发生。在步骤320中,收集器经由被监控的每个中转提供者向每个目标IP地址发送互联网控制消息协议(ICMP)回声请求分组。例如,给定具有两个中转提供者TP1和TP2的网络,收集器经由TP1从第一源IP地址并经由TP2从第二源IP地址向第一目标IP地址发送重复的回声请求分组。在步骤330中,目标IP地址通过发送回声应答分组来响应回声请求分组。每个回声应答分组被路由到对应的回声请求分组的表面(apparent)源IP地址。如上面所解释的,如果应用于边界路由器接口的路由图强制对应的回声请求经由通告不同前缀的中转提供者到达目标IP地址,那么表面源IP地址可以与实际源IP地址不同。

[0060] 在步骤340中,收集器基于回声应答分组中的时间戳(例如,回声请求时间和回声应答时间)来确定与往返测量相关联的等待时间。如本领域中所熟知的,每个回声请求消息可以包括指示发送时间的时间戳和序列号。同样,每个回声应答消息可以包括发送时间和序列号。每个回声应答还包括对应的回声请求消息的时间戳和序列号。回声请求与由回声应答中的时间戳所指示的回声应答发送时间之间的差异指示等待时间,其可以被记录并存储在数据库或其它存储器存储装置中。

[0061] 收集器还可以基于每个中转提供者的回声请求与回声应答之比或回声应答的时间线来评估分组丢失。例如,如果收集器经由第一中转提供者发送了十个回声请求,但只接收到八个回声应答,那么收集器可以确定第一中转提供者具有20%的分组丢失率。分组丢失可以在中转提供者之间相互关联,以评估进站和出站路径的分组丢失,很像经由不同中转提供者的往返等待时间测量可以用于评估进站和出站路径的等待时间。

[0062] 基于度量比较来确定最佳路径

[0063] 图4是图示由图1A中所示的决策引擎103实现的、用于比较连接客户场所和目标IP地址的不同中转提供者的各种度量的处理400的流程图。在步骤410中,决策引擎例如响应于用户输入而选择目标IP地址。在步骤420a和420b中,决策引擎识别将客户场所连接到步骤410中选择的目標IP地址的中转提供者。在步骤430a-430d中,决策引擎通过在步骤420中识别出的中转提供者来确定到所选目标IP的进站和出站度量。进站和出站度量包括但不限于等待时间、成本和分组丢失。等待时间和分组丢失可以通过如上面关于图3所述的那样ping目标IP地址来确定。成本信息可以由客户提供。

[0064] 在步骤440中,决策引擎准备进站和出站路径的每个组合的矩阵(例如,如图1B中所示的那样)或其它表示。例如,如果中转提供者的数量是N,那么通过组合获得的路径总数为 N^2 。如果N非常大,那么决策引擎可以对等待时间、分组丢失和/或其组合进行采样,以便为了更快计算而减少测量和/或组合的数量。在步骤450中,决策引擎把外部度量考虑在内,诸如与经由每个中转提供者路由分组相关联的成本。通过把外部度量考虑在内,在步骤460中,对于每个进站/出站中转提供者组合,进站和出站度量被整合在优先度得分的矩阵或其它表示(等待时间、成本和/或分组丢失的加权组合)中。在步骤470,决策引擎基于矩阵中的优先度得分的比较来识别最佳路径。

[0065] 经由多个边界路由器监控并引导流量

[0066] 图5A-5C绘出了测量客户场所虚拟机510与一个或多个目标/目的地IP地址540a和540b(统称为目标IP地址540)之间的等待时间和分组丢失的网络500。包括存储目标IP地址

的列表以进行监控和引导的收集器52的虚拟机510耦合到聚合路由器590a和590b(统称为聚合路由器590)。添加使用耦合到聚合路由器590的边界路由器520a和520b(统称为520)的另一层路由接口。边界路由器520将流量路由到其它边界路由器520和中转提供者530a-530n(统称为530)。

[0067] 在操作中,如上文和下文更详细描述的那样,收集器512监控始发或源IP地址IP1-IPN与列出的目标IP地址540之间的传输的等待时间和分组丢失。再次,收集器512使用客户提供的面向公众的IP地址作为用于收集等待时间和分组丢失数据的源点。目标IP地址540的列表可以通过向客户询问不同的IP地址来建立。包括在客户虚拟机510中的客户收集器512允许以周期性的方式进行从客户端点或目标IP地址的连接。

[0068] 虚拟机510通过使用路由聚合方法生成具体路由的聚合路由器590a(SW1)和590b(SW2)路由流量。该方法通过用单一且一般性路由替换多条路由来组织网络500。这减少了路由器的数量,并最小化了与路由协议相关的开销。耦合到聚合路由器590的边界路由器520a(BR1)和520b(BR2)被部署成执行入口和出口过滤。边界路由器520在骨干网络和一个或多个开放最短路径优先(OSPF)区域之间建立连接。

[0069] 聚合路由器590基于由客户场所虚拟机510和收集器512进行的等待时间和分组丢失测量通过适当的中转提供者530将流量路由到(一个或多个)目标IP地址540和从(一个或多个)目标IP地址540路由的所述流量。如上文和下文中描述的,通过分析和比较中转提供者130的诸如等待时间、成本和分组丢失之类的被监控的度量来选择用于具体目标IP地址540的适当中转提供者530。路由可以使用路由图或通过如下所述基于分析的度量将BGP属性与中转提供者相关联来实现。

[0070] 监控等待时间和分组丢失

[0071] 在图5A中,收集器512经由中转提供者530监控到不同目标IP地址的等待时间和分组丢失。其通过使用基于策略的路由以匹配在收集器512上配置的每个IP地址并将针对每个IP地址的分组路由出其相关联的中转提供者530来完成这件事。路由图被应用于边界路由器520上的每个入口接口522a和522b。(在图5A中,入口接口522a和522b也被标记为gig 1/1、gig 1/2和gig 1/3。)这些路由图使用访问列表来匹配来自收集服务器512的IP地址,并设置用于将分组中转到适当的中转提供者530或相邻边界路由器520的下一跳。应用于每个面向聚合路由器590的边界路由器接口的路由图覆盖所有IP地址,并且强制出站回声请求分组或者到耦合到对应边界路由器530的中转提供者或者到邻居边界路由器530。通过强制回声请求分组离开不同的中转提供者530,收集器512可以沿着它与目标IP地址540之间可用的入站和出站路径的每个组合来测量等待时间。

[0072] 图6和下面的伪代码图示了处理600的示例,该处理配置并应用路由图以监控第一IP地址(例如,在客户场所)与第二IP地址(例如,目标IP地址)之间的等待时间和分组丢失。在步骤610中,客户在收集器512上为每个目标IP地址设置访问:

	边界路由器 520a 配置	边界路由器 520b 配置
[0073]	访问列表 1 允许<IP_1> 访问列表 2 允许<IP_2> 访问列表 3 允许<IP_3> 访问列表 n 允许<IP_N>	访问列表 1 允许<IP_1> 访问列表 2 允许<IP_2> 访问列表 3 允许<IP_3> 访问列表 n 允许<IP_N>

[0074] 在步骤620中,客户创建路由图 (COLLECTOR_SW),以在来自聚合交换机590的分组中寻找每个目的地IP地址:

	边界路由器 520a 配置	边界路由器 520b 配置
[0075]	路由图 COLLECTOR_SW 允许 10 匹配 ip 地址 1 设置 ip 下一跳<T1_IP> ! 路由图 COLLECTOR_SW 允许 20 匹配 ip 地址 2 设置 ip 下一跳<T2_IP> ! 路由图 COLLECTOR_SW 允许 30 匹配 ip 地址 3 设置 ip 下一跳<BR2_P2P>	路由图 COLLECTOR_SW 允许 10 匹配 ip 地址 1 设置 ip 下一跳<BR1_P2P> ! 路由图 COLLECTOR_SW 允许 20 匹配 ip 地址 2 设置 ip 下一跳<BR1_P2P> ! 路由图 COLLECTOR_SW 允许 30 匹配 ip 地址 3 设置 ip 下一跳<T3_IP>
[0076]	! 路由图 COLLECTOR_SW 允许 40 匹配 ip 地址 4 设置 ip 下一跳<BR2_P2P> !	! 路由图 COLLECTOR_SW 允许 40 匹配 ip 地址 4 设置 ip 下一跳<TN_IP> !

[0077] 在步骤630中,顾客为边界路由器530之间的交叉链接创建路由图 (COLLECTOR_BR1和COLLECTOR_BR2):

[0078]	边界路由器 520a 配置 路由图 COLLECTOR_BR1 允许 10 匹配 ip 地址 1 设置 ip 下一跳<T1_IP> ! 路由图 COLLECTOR_BR1 允许 20 匹配 ip 地址 2 设置 ip 下一跳<T2_IP> !	边界路由器 520b 配置 路由图 COLLECTOR_BR2 允许 10 匹配 ip 地址 3 设置 ip 下一跳<T3_IP> ! 路由图 COLLECTOR_BR2 允许 20 匹配 ip 地址 N 设置 ip 下一跳<TN_IP> !
--------	--	--

[0079] 在步骤640中,客户将COLLECTOR_SW路由图应用于面向聚合交换机590的每个接口522(gig 1/1和gig 1/2),并将面向边界路由器520的接口522(gig 1/3)与COLLECTOR_BR1和COLLECTOR_BR2路由图交叉链接:

[0080]	边界路由器 520a 配置 接口 gig 1/1 ip 策略路由图 COLLECTOR_BR1 !	边界路由器 520b 配置 接口 gig 1/1 ip 策略路由图 COLLECTOR_BR2 !
[0081]	接口 gig 1/2 ip 策略路由图 COLLECTOR_SW ! 接口 gig 1/3 ip 策略路由图 COLLECTOR_SW !	接口 gig 1/2 ip 策略路由图 COLLECTOR_SW ! 接口 gig 1/3 ip 策略路由图 COLLECTOR_SW !

[0082] 如果每个中转提供者530如上面关于图1A所述的那样通告不同的/24,那么收集服务器512可以测试N个流出乘N个流入。如上所述,源自属于由中转提供者530a通告的前缀的源IP地址的分组(例如,用于测量等待时间的回声请求分组)通过每个中转提供者530被信道传输到其目的地IP地址。但是,来自目的地IP地址的返回分组(例如,响应于回声请求的回声应答分组)通过中转提供者530a中转回源IP地址。测量给定源和目的地IP地址之间的每个流入/流出组合产生流入和流出等待时间性能的矩阵(例如,如图1B中所示)。比较等待时间测量使得有可能在给定时间选择到给定目的地IP地址的流入和流出的“最佳”(例如,最低等待时间)组合。

[0083] 用于确定到端点的路径的最佳中转决策引擎

[0084] 虚拟机510、收集服务器512或单独的处理器可以实现执行以下伪代码以确定用于将流量路由到给定目标IP地址的期望路径的决策引擎(未示出):

[0085] 对于每个目标IP地址t以规律的间隔向应用程序接口(API)轮询数据:

- [0086] 获取针对t的所有最近的测量(按时间或按计数)
- [0087] 对于每个中转提供者P;
- [0088] 基于使用P的等待时间、分组丢失和客户配置的成本来为t
- [0089] 计算优先度得分
- [0090] 决定是否为t推荐路由交换机
- [0091] 在决定是否推荐交换机时,决策引擎检查哪个中转提供者530具有针对目标IP地址的最高优先度得分,并将那个得分与当前用于那个目标IP地址的中转提供者530的优先度得分进行比较。决策引擎通过或者维护路由状态(例如,最后一次推荐)的记录或者通过向客户的系统查询其路由配置来“知道”当前的提供者。决策引擎可以对其推荐交换中转提供者530的频率强加限制,以防止中转提供者530之间的不期望的振荡。类似地,收集器512可以调节其等待时间和分组测量的频率,以防止不利地影响路由性能(例如,由于具有太多回声请求和回声应答而造成拥塞)。
- [0092] 静态路由(手动流量引导)
- [0093] 图5B示出了网络500如何经由根据决策引擎的推荐设置的静态路由手动地引导流量。基于优先度得分为每个端点确定适当的中转提供者530,该得分是客户配置的等待时间、成本和分组丢失的加权组合。客户服务器510通过为每个端点应用静态路由来覆盖BGP策略。例如,如果中转提供者530c为到端点A的流量提供具有最低优先度得分的路径,那么BGP策略被覆盖,并且针对端点A的流量通过中转提供者530c被引导。类似地,如果中转提供者530a为到端点B的流量提供具有最低优先度得分的路径,那么BGP策略被覆盖,并且针对端点B的流量通过中转提供者530a被引导。
- [0094] 图7和下面的伪代码图示了用于(手动地)将流量引导到两个目标端点的处理700。在步骤710中,决策引擎生成到由收集器512监控的每个端点(目标IP地址540)的静态路由。在步骤720中,决策引擎将针对端点A和B的静态路由应用于每个边界路由器520:

边界路由器 520a 配置	边界路由器 520b 配置
Ip 路由<endpoint_a> 255.255.255.255 <T1_IP>	ip 路由<endpoint_a> 255.255.255.255 <BR1_P2P>
Ip 路由<endpoint_b> 255.255.255.255 <BR2_IP>	ip 路由<endpoint_B> 255.255.255.255 <T3_IP>

- [0095] 对于指向适当的中转提供者530或边界路由器520的每个端点,可以重复这个处理700(步骤730和740)。要注意的是,静态路由可以以任何次序或甚至同时被应用于用于端点540的边界路由器520。

- [0096] 一旦被应用,静态路由就如下工作。如果去往端点A的分组出现在边界路由器520a处,那么边界路由器520a经由中转提供者530a路由分组。类似地,如果去往端点B的分组出现在边界路由器520b处,那么边界路由器520b经由中转提供者530c将分组路由到端点B。但是,如果去往端点B的分组出现在边界路由器520a处,那么边界路由器520a将分组路由到边界路由器520b,边界路由器520b进而经由中转提供者530c将分组路由到端点B。类似地,如果去往端点A的分组出现在边界路由器520b处,那么边界路由器520b将分组路由到边界路

由器520a,边界路由器520a进而经由中转提供者530a将分组路由到端点A。

[0098] 通过将BGP属性与中转提供者相关联来引导流量

[0099] 如上所述的静态路由不能很好地扩展并且在配置和错误管理方面可能是笨拙的。幸运的是,可以在客户位置处的收集服务器512上使用BGP守护进程来完成动态引导。这个BGP守护进程使得收集服务器512能够或者直接或者经由一个或多个路由反射器585a和585b(统称为路由反射器585)与客户路由器530对等,如图5C中所示。

[0100] 图8和以下的伪代码图示了用于使用BGP社区属性经由特定中转提供者动态地配置路由器以引导针对给定IP地址的流量的处理800。每个中转提供者530与不同的BGP社区属性相关联。每个主机前缀可以通过改变BGP“社区”而与给定的中转提供者530相关联。一旦形成这种关联,边界路由器520就将经由相关联的中转提供者530,例如按照以下伪代码,引导具有给定BGP“社区”属性的流量:

[0101]

示例 IOS 路由图	
	ip 前缀列表 Slash32 seq 5 允许 0.0.0.0/0 ge 32
	路由图 TRAFFIC_STEERING 允许 10
	匹配 ip 地址前缀列表 Slash32
	匹配社区<T1_COMMUNITY>
	设置 ip 下一跳<T1_P2P>
	设置社区无导出添加(no-export additive)
	路由图 TRAFFIC_STEERING 允许 20
	匹配 ip 地址前缀列表 Slash32
	匹配社区<T2_COMMUNITY>
	设置 ip 下一跳<T2_P2P>
	设置社区无导出添加
	路由图 TRAFFIC_STEERING 允许 30
	匹配 ip 地址前缀列表 Slash32
	匹配社区<T3_COMMUNITY>
	设置 ip 下一跳<T3_P2P>
	设置社区无导出添加

[0102] 处理800可以通过利用与指派给每个中转提供者530的唯一BGP社区属性匹配的策略配置边界路由器520来实现(步骤810)。一旦与给定的BGP社区匹配,每个边界路由器520就改变其朝着相关联的中转提供者530的对应的下一跳(步骤820)。如果另一个中转提供者530开始提供到同一端点的较低等待时间,那么BGP守护进程通过改变用于那个端点的BGP社区属性的值来改变边界路由器520和中转提供者530之间的关联。例如,BGP守护进程可以对BGP会话应用以下路由图:

	示例 IOS 路由图
[0103]	路由器 bgp ABCD <snip> 邻居<BGP_DAEMON> 路由图 TRAFFIC_STEERING in <snip>

[0104] 用于解析DNS查询的实时用户监控

[0105] 以最佳方式引导互联网数据流量的另一种做法是通过基于各种性能度量(包括等待时间和成本)来解析域名系统(DNS)查询。这是通过收集和分析真实用户监控(RUM)数据以预测从不同内容起源向客户端提供内容以及用给定的递归解析器查询给定权威服务器的性能来实现的。可以使用预测的性能来将进行DNS查询的客户端引导到给定的内容起源,该内容起源可以提供来自内容递送网络(CDN)或云提供者的内容,和/或将递归解析器引导到给定的权威服务器。流量是基于内容起源相对于客户端IP地址的预测性能来引导的。更具体而言,一个或多个服务器收集用于不同的RUM数据,在IP级别处理RUM数据,聚合跨IP地址集合的RUM数据,并创建用于快速访问聚合数据的数据库。然后该数据库用于以最佳方式引导流量。

[0106] 如本领域技术人员所理解的,DNS是用于计算机、服务或连接到互联网或私有网络的任何资源的分层分布式命名系统。它将各种信息与指派给每个参与实体的域名相关联。DNS还将更容易记忆的域名翻译为用于定位和识别利用底层网络协议的计算机服务和设备的数字互联网协议(IP)地址。

[0107] 权威DNS服务器(也称为权威名称服务器或权威)对关于域名到数字IP地址的映射的查询并且还对其它资源记录(RR)(诸如邮件交换(MX)记录)的请求做出响应。为了对这些查询做出响应,每个权威都有其自己的DNS记录的DNS数据库。存储在DNS数据库中的常用记录类型包括用于对应域的IP地址(A和AAAA)、简单邮件传输协议(SMTP)MX记录和名称服务器(NS)记录。DNS数据库还可以存储用于其它类型数据的记录(包括可用于认证DNS记录的域名别名(CNAME)和DNS安全扩展(DNSSEC)记录)。

[0108] 为了向互联网添加新域,基本的DNS标准要求域所有者或注册者从注册员(registrar)处购买域名,并指定用于回答针对新域的查询的权威DNS服务器的名称。注册人从权威DNS提供者(诸如Dynamic Network Service公司,Manchester,NH)获得权威DNS服务,并利用权威DNS提供者为其域名(或更精确地说,区)配置记录。当最终用户的机器尝试访问新的域名时,它请求递归DNS服务器(也称为递归服务器、递归解析器或递归)检索新域的DNS记录,最常见的是A或AAAA(IPv4或IPv6地址)。这些DNS记录包括提供最终用户请求的内容或其它信息的内容起源的IP地址。递归服务器定位由权威DNS提供者维护的权威DNS服务器(也称为权威服务器或简称为权威),然后向权威DNS服务器查询DNS记录。递归DNS服务器将权威DNS服务器的答案返回给最终用户的机器,并且还可以根据其生存时间(TTL)高速缓存答案。然后最终用户的机器尝试使用由权威DNS服务器提供的DNS记录来访问域。

[0109] 常规的递归和权威DNS服务器不考虑与用于新域的内容起源的IP地址与最终用户的机器(客户端)之间的分组传输相关联的等待时间。相反,权威服务器只是根据由新域的

运营商设置的(静态)策略来提供IP地址。因此,常规的名称服务器不能将客户端引导至提供最低等待时间的内容起源。

[0110] RUM数据监控系统

[0111] 图9示出了监控RUM数据并基于等待时间和其它因素使用RUM数据来解析DNS请求的系统900。系统900包括经由互联网901或另一分组交换网络与内容起源910a-910c(统称为内容起源910)及权威服务器920a和递归解析器920b(统称为权威DNS服务器递归解析器920)通信的客户端930。系统900还包括存储RUM数据的RUM数据库940和存储类型标记IP集(TYLIPS)的另一个数据库950。

[0112] 在操作中,作为访问存储在内容起源910之一上的内容的处理的一部分,客户端930请求向权威服务器920a发送DNS请求931。权威服务器920a通过基于客户端的IP地址和内容起源性能的性能排名选择具有期望性能的内容起源来对这个请求做出响应,用于向具有与客户端在同一子网上或在相同地理区域中的IP地址的其它客户端(未示出)提供数据。这种排名可以专门针对权威服务器920a定制。权威服务器920a向客户端930提供所选内容起源(例如,内容起源910a)的IP地址或主机名921。客户端930从所选的内容起源910下载内容911,从而生成用于解析未来DNS请求的附加性能数据。来自客户端930的相关请求经由递归解析器920b发送。

[0113] 收集RUM数据

[0114] 图10是图示用于使用图9中所示的系统900来创建RUM数据的数据库以引导流量的处理1000的流程图。在步骤1010中,互联网性能数据作为真实用户监控(RUM)记录的流被收集。数据是从多个内容提供者或内容起源请求或下载数据的用户或客户端收集的。在步骤1020中,通过几乎同时测量来自各种内容起源的相同数据样本的下载时间来处理每个RUM记录。处理允许直接比较每对内容起源。

[0115] 在步骤1030中,IP地址被分组为类型标记IP集(TYLIPS)。TYLIPS是共享共同特征的IP地址集合,诸如同一个国家、源自相同的提供商,等等。相对内容起源性能的直方图为每个IP地址累积。然后为TYLIPS组合并累积这些直方图。

[0116] 对于每个TYLIPS,在步骤1040中使用来自最近RUM数据的定时、失败和其它信息来计算性能得分。这些得分用于对属于那个TYLIPS的IP地址的内容起源进行排名。内容起源及其相关联的TYLIP排名存储在TYLIPS数据库中。这个数据库提供对聚合数据的快速访问并用于以最佳方式引导流量。

[0117] 例如,可以从多个内容提供者或内容起源请求并下载数据的用户或客户端收集互联网性能数据。这些数据被组织成记录,并且每个RUM记录可以与从用户到多个内容起源的下载相关联,并且包含以下一个或多个:(1)客户的互联网协议(IP)地址,(2)用于DNS解析的一个或多个递归解析器的(一个或多个)IP地址,(3)与内容起源唯一地相关联的标识符,例如统一资源标识符(URI),以及(4)与下载相关联的时间数据。内容起源可以是例如内容递送网络(CDN)(诸如Akamai、Level 3)、云提供商(诸如Digital Ocean、Amazon)或内容发布者的私有数据中心。与下载相关联的时间数据可以包括下载速度的各种中间测量(诸如域名服务器(DNS)解析的时间、建立连接的时间、到第一个字节的时间、到最后一个字节的时间、从特定的内容起源下载数据样本的总持续时间,等等)。

[0118] 图11图示了用于收集真实用户监控(RUM)记录的流的另一个处理1100。在步骤

1110中,部署诸如JavaScript之类的代码,以下载RUM数据。在一些实施例中,代码是经由内容起源的网页部署的。在其它实施例中,代码被部署在内容起源的数据中心中。当客户端访问内容起源的网页时,所部署的代码被下载到客户端的浏览器上(步骤1120)。在步骤1130中,客户端执行这个代码,以收集RUM数据。在步骤1140中,客户端向RUM数据库发送RUM数据,包括但不限于客户端和内容起源IP地址、DNS请求解析时间、内容起源的URL以及每次下载的下载时间。这个客户端可以重复数据收集和传输,以用于来自不同内容起源的下载、来自相同内容起源的多次下载或者在同时和/或顺序测量中快速相继地进行这二者(步骤1150),针对不同的IP地址比较DNS请求解析时间和下载时间产生与客户端、递归解析器和/或内容起源的不同组合相关联的排名(步骤1160)。这些排名可以用于对进一步的DNS请求做出响应,例如,以减少总时间或与内容起源识别和内容下载处理的一个或多个步骤相关联的时间(步骤1170)。

[0119] 参考以下非限制性示例说明用于收集RUM记录的代码的操作。客户经由web浏览器访问美国专利整页图像<http://patft.uspto.gov/netathtml/PTO/patimg.htm>的网页。该网页提供由客户端的浏览器下载的JavaScript代码。在这个示例中,当客户端从网页下载美国专利的整页图像时,JavaScript代码在web浏览器上执行并收集RUM数据。JavaScript代码可以收集(1)访问美国专利整页图像的网页的客户端的IP地址,(2)客户端使用的DNS递归解析器的IP地址,(3)用于数据样本的内容起源的URI,以及(4)下载数据样本的处理的各种中间时间。

[0120] RUM数据处理

[0121] 在每个RUM记录中,几乎同时测量来自各种起源的同一数据样本的下载时间。这允许每对内容起源的直接性能比较。对于每对内容起源,计算相同数据样本的对应下载时间之间的差异,例如,DNS解析时间和/或连接时间之间的差异。这些差异在一段时间内累积。这些差异的直方图针对每对起源和每种类型的定时计算。定时的一些示例是DNS解析或下载时间。

[0122] 例如,具有地理定位于Houston,Texas的IP地址并使用具有地理定位于Dallas,Texas的IP地址的递归解析器的客户端可以看到分别来自OID1的10ms和40ms的DNS解析和下载时间,以及来自OID2的15ms和50ms的DNS解析和下载时间。在这个示例中,对于对(OID1,OID2),DNS解析时间的差异为-5ms,下载时间的差异为-10ms。这些随时间累积的差异可以用于创建由客户端或递归IP或这些IP的特征(例如,地理位置、互联网服务提供商)、定时类型(例如,DNS解析、下载时间)和起源对联合索引的直方图。

[0123] 图12示出了用于Houston,Texas的所有客户端IP的示例直方图,其来自在24小时内收集的两个不同内容递送网络(CDN)(在这里是Akamai和Fastly)中的内容起源的DNS解析时间的差异。起源对之间差异的直方图用于比较一个起源相对于另一个起源的性能。在一些情况下,包括图12,通过计数在每个差异直方图中出现多少正值和负值来实现这种比较。例如,在用于一对起源(OID1,OID2)的差异直方图上,负值表示用于OID1的定时低于用于OID2的定时时的情况,因此OID1的性能更好。正值表示在OID2的性能更好时的相反情况,候。如果差异为零,那么OID1和OID2的性能是完全相同的。如果定时的绝对差低于给定阈值,例如,如果定时的绝对差小于20ms,那么性能也可以被认为是等同的。对于每一个起源,可以将对照队列中的其它起源进行的性能的正向(head-to-head)比较求平均,以产生表示

特定起源的整体性能的单个得分。为了相对于其它起源对每个起源进行评分,使用特定于起源的其它测量(诸如中值等待时间、等待时间随时间的变化性和/或稳定性、失败率,等等)。

[0124] 类型标记IP集(TYLIPS)数据聚合

[0125] RUM数据在共享一个或多个共同特征的IP地址组上被聚合。这些集合被称为类型标记IP集,或“TYLIPS”。TYLIPS的一些示例是:法国(类型=国家),波士顿(类型=城市),AS 174(类型=互联网服务提供商(ISP))和(巴黎,AS5511)(类型=城市-提供商)。术语“TYLIPS”也可以与术语“IP特征”互换使用。单个IP是尺寸为1的TYLIPS。因为TYLIPS是IP地址集,所以它们允许使用数学集运算并具有集合的所有特性。例如,两个TYLIPS可以相交,或者一个TYLIPS可以包含在另一个更大的TYLIPS中。

[0126] 图13是图示用于聚合RUM数据的处理的流程图。在步骤1310中,几乎同时测量每对起源的下载时间。这些差异在一段时间内累积。在步骤1320中,针对累积时间段计算每对起源的差值并生成直方图。

[0127] 这些相对起源性能的直方图或者可以建立在“每个客户端IP地址”的基础上,或者可以建立在“每个递归解析器IP地址”的基础上。针对每个IP地址累积直方图,并将其组合并累积以生成IP地址组的直方图(步骤1330)。每组IP地址是根据来自共享一个或多个共同特征的组中的所有IP地址来选择的。共同特征可以是例如地理位置(来自该组的IP地址被地理定位到特定的城市、国家或大洲)、源自特定互联网服务提供商(ISP)或通过其中转,或共同组织中的成员资格。共同特征也可以是联合特征,诸如地理位置和ISP(例如,源自位于赫尔辛基的TeliaSonera的IP地址)。

[0128] 在步骤1340中,比较同一TYLIPS内的一个起源相对于另一个起源的性能。TYLIPS基于其性能和总得分进行排名(步骤1350)。响应于后续的DNS请求,递归解析器可以使用TYLIP排名来选择用于特定客户端的特定内容起源(步骤1360)。

[0129] 在使用RUM数据对互联网流量进行最佳引导的情况下,一些IP地址有可能具有不完整或质量不合适的数据。如果完整的数据可用,那么可以使用与那个IP地址相关联的数据来为每个IP地址做决定。为了在实践中解决数据质量或数据稀疏的问题,RUM数据被聚合到TYLIPS中并定义TYLIPS的层次结构。给定的IP包含在该层次结构中的若干相继的更大TYLIPS中。TYLIPS基于其特异度、可用数据的量、所收集的数据的质量以及其它相似准则进行排名。当用于IP的数据不可用时,对其有足够质量的足够数据可用的最具体的TYLIPS(例如最小的)被使用。底层的假设是,给定IP的性能简档与最具体的TYLIPS中IP的性能简档相似。

[0130] 图14A和14B示出了TYLIPS的不同层次结构。在图14A中,基于TYLIPS集的层次结构,对于IP1,选择TYLIPS“网络前缀”。对于IP2,所选择的TYLIPS是“管理部门”中的“BGPASN”,例如,美国州中的ISP。而对于IP3,所选择的TYLIPS是“国家”。在图14B中,IP地址88.221.8.1属于以下TYLIPS层次结构,其可以被或可以不被预定义:前缀88.221.8.0/22, AS5511,城市Madrid,省Comunidad de Madrid,国家西班牙,地区西欧,以及大洲欧洲。如果请求来自这个IP地址,那么选择具有足够数据的最具体的TYLIPS。

[0131] 一些TYLIPS包含在其它TYLIPS中,例如省中的城市,国家中的省,而一些TYLIPS只相交,例如前缀和国家。对于另一个IP地址,比如说90.84.255.1,可用的最具体的数据可能

只在国家级别,并且其TYLIPS层次结构是国家西班牙,地区西欧,大洲欧洲。

[0132] 从层次结构中,可以从可用的TYLIPS中选择最具体的TYLIPS。最具体的TYLIPS是通过选择其性能数据与IP地址的简档最佳匹配的TYLIPS来选择的。此外,最具体的TYLIPS是基于足够的可用数据来选择的。

[0133] 对TYLIPS进行排名

[0134] 如上面关于收集RUM数据所描述的,当客户端访问内容起源时,客户端下载并执行部署在该内容起源中的图像或代码。这种代码或图像记录RUM数据(诸如下载时间、进行测量的时间、客户端的IP的位置,等等)。测量集合被称为从一个客户端IP到所有内容起源的定时测量集合,其中代码或图像作为信标被托管。这些信标被传送到数据处理服务器以处理数据。

[0135] 内容起源排名是从RUM数据计算的。数据在一个时间间隔内聚合。例如,数据在24小时内聚合。对于时间间隔中的每个信标,计算每对内容起源之间的定时差异。

[0136] 例如,如果在信标中对于3个内容起源测量以下时间:

[0137] • CDN_A:60ms,CDN_B:100ms,CDN_C:40ms,CDN_D:200ms那么按对的差异是:

[0138] • CDN_A-B:-40ms,CDN_A-C:20ms,CDN_A-D:-140ms

[0139] • CDN_B-C:60ms,CDN_B-D:-100ms

[0140] • CDN_C-D:-160ms

[0141] 计算在时间间隔内的每对内容起源的时间差的分布。这允许计算平均差或识别一个内容起源比另一个更好的时间百分比。

[0142] 可以从按对的内容起源统计信息中计算若干排名。例如,公开了基于“更好的时间百分比”的排名的说明。对于给定的内容起源,识别当与其竞争对手相比时内容起源更好的时间百分比。然后将这些百分比求平均,以计算得分。这些得分被用于对内容起源进行排名,并将具有相似得分的内容起源分组到各级中。以示例来最好地说明这一点。对于四个内容起源的按对比较,以下百分比示出何时一个内容起源比另一个更好:

[0143] • A-B 52%-48%,A-C 75%-25%,A-D 95%-5%

[0144] • B-C 70%-30%,C-D 90%-10%

[0145] • C-D 60%-40%

[0146] 以表格形式,这些排名是:

[0147]

内容起源	A	B	C	D
A	-	52%	75%	95%
B	48%	-	70%	90%
C	25%	30%	-	60%
D	5%	10%	40%	-

[0148] 在这个示例中,内容起源A是最好的,但它与内容起源B几乎相同,但是明显优于内容起源C,并且比内容起源D好得多,内容起源D是所有当中最糟糕的。

[0149] 与内容起源B、C和D相比,内容起源A分别更好52%、75%和95%的时间。百分比的平均值是74%。内容起源B、C和D的平均百分比分别是69.33%、38.33%和18.33%。使用这些平均值,内容起源排名如下:

[0150]

CDN	得分	排名
-----	----	----

A	74.00%	1
B	69.33%	2
C	38.33%	3
D	18.33%	4

[0151] 图15图示了对于在2016年3月2日0:00UTC和2016年3月3日0:00UTC之间的数据针对日本的客户端IP地址的排名的示例性示例。排名与得分一起在1510中示出。百分比矩阵1520是在行上具有每个内容起源更好的百分比的矩阵。例如, Fastly比Edgecast更好大约54%的时间。并非所有百分比加起来都是100%。这是因为测量的百分比是完全相同的, 即, 两个内容起源具有完全相同的性能。具有相似性能的内容起源被分成各等级。例如, 在等级A中, 一对内容起源之间的最大得分差异小于10。

[0152] 为了将内容起源分成各等级, 得分以降序考虑。最初考虑用于实例A或0的最高等级。对于具有等级的给定得分, 按次序计算与前一个和下一个得分的间隙。此外, 还计算该得分与同一等级内的最高得分之间的间隙。如果得分差异严格地大于二并且如果: (1) 当前得分与下一个得分之间的间隙大于10%, 即, 下一个得分小于当前得分的90%或, (2) 下一个得分与最高得分之间的间隙大于10% (下一个得分小于最高得分的90%) 并且下一个得分与当前得分之间的间隙大于前一个得分与当前得分之间的间隙的三倍, 那么前进到下一个等级。

[0153] 例如, 在50和40之间中断, 因为它们之间的差异是10, 这大于2, 并且40小于50的90% (即, 45)。类似地, 对于得分60、57、55、53、51、50、46, 在50和46之间中断, 因为 $50 - 46 - 4 > 2$, $46 < (0.9 \times 60) = 54$ 并且 $50 - 46 = 4 > 3 \times (51 - 50) = 1$ 。

[0154] 如果当前等级与下一个等级之间的差异大于20%, 那么如下将等级前进多于一级: (1) 对于10%与20%之间的差异, 前进一个等级, 例如A到B; (2) 对于20%与40%之间的差异, 前进两个等级, 例如A到C; (3) 对于40%与80%之间的差异, 前进三个等级, 例如A到D; 以及(4) 对于超过80%的差异, 前进到F。

[0155] 因此, 相似的得分得到相同的等级。等级中断在具有显著差异(10%)的得分之间应用。等级中断使得等级内的最高分和最低分并没有太大差异。

[0156] 使用在TYLIPS上聚合的RUM数据引导流量

[0157] 图16图示了用于使用在一个或多个TYLIPS上聚合的RUM数据来引导流量的处理1005。客户端930向递归解析器920b发送DNS请求, 以将域解析为IP地址, 并且递归解析器将DNS请求转发到权威DNS服务器920a (步骤1610)。权威DNS服务器和递归解析器920向TYLIPS数据库950提供客户端的DNS递归解析器的IP地址和客户端的IP地址 (如果可用的话) (步骤1620), TYLIPS数据库950检索与客户端的DNS递归解析器的IP地址和客户端的IP地址相关联的最具体的可用TYLIPS。TYLIPS数据库950还检索用于最具体的可用TYLIPS的内容起源的对应排名, 并基于排名选择 (一个或多个) 内容起源。TYLIPS数据库950将选定的内容起源1630提供给权威DNS服务器920a和递归解析器920b, 其通过将 (一个或多个) 最优内容起源的IP地址发送到客户端930来响应客户端的DNS请求。

[0158] 基于RUM数据, 递归解析器的IP地址可以与客户端的IP地址的TYLIPS相关联。如果在RUM数据中没有观察到递归解析器的IP地址, 那么可以使用递归解析器的IP地址来查找TYLIPS的层次结构。客户端子网数据 (诸如客户端的前缀) 也可以用于查找TYLIPS的层次结

构。

[0159] 递归解析器IP地址可以与属于由递归解析器IP地址表示的客户端IP地址的IP特征的层次结构或TYLIPS相关联。例如,递归解析器可以代表美国东海岸许多城市中的客户端向权威DNS服务器进行查询。在其它情况下,递归解析器IP地址的一个或多个特征可以被直接用作引导的基础。例如,递归解析器可以被地理定位到波士顿。无论如何,选择与来自递归解析器的查询相关联的TYLIPS集合,并且这个选定的TYLIPS集合用于引导决定。

[0160] 以地理定位在Somerville,Massachusetts并源自Comcast的递归解析器为例。如果有足够的RUM记录与这个递归解析器的IP地址相关联,以区分多个起源候选并选择适当的起源,那么可以使用与递归解析器的IP地址专门相关联的排名。否则,可以使用与具有足够的RUM记录的具有较低特异性的区域(例如,城市Somerville,州Massachusetts或者地区新英格兰)相关联的起源排名,以形成区分的基础。常常期望使用具有足够高质量数据的最具体的特征或联合特征,以允许明确推荐正确的起源。换句话说,选择对其具有足够数据的最具体的TYLIPS,并且对于最具体的TYLIPS推荐具有最佳得分的内容起源。通过为期望数量的RUM数据提供阈值并分析当前数量是否大于该阈值,数据可以被足够量化。

[0161] 例如,内容提供者(例如,纽约时报)可以向权威DNS提供者付费,以将用户引导至www.nytimes.com,以到在多个内容起源(例如,位于欧洲和北美的起源)中的正确内容起源。权威DNS提供者收集RUM记录(例如,使用上面讨论的方法),基于用户位置比较www.nytimes.com的性能,并提供针对内容起源的推荐。例如,权威DNS提供者可以推荐欧洲,而不是北美。

[0162] 用于给定区域的权威DNS提供者提供实时流量引导,使得由查询递归DNS解析器表示的一个或多个用户被映射到针对所请求的内容的起源。可以基于低等待时间、高可用性、稳定性和其它相似的属性来选择起源。换句话说,权威DNS提供者可以将请求域的递归解析器引导到最期望的位置,以获得用于那个域的数据。

[0163] 因此,公开了一种从递归IP和/或客户端IP计算TYLIPS的层次结构的处理。对于每个TYLIPS,对具有内容起源的预先计算的得分和排名的数据库的访问都是可用的。选择最佳TYLIPS,例如,对其具有足够数据的最具体的TYLIPS。最佳TYLIPS的得分和排名被用于匹配由递归IP请求的域与具有(一个或多个)最佳得分的(一个或多个)内容起源。

[0164] 本发明的实施例可以被用于实时地引导流量,或者以基于过去的性能实现期望的性能水平这样一种方式配置DNS。使用本发明的方法,可以生成将地理区域匹配到期望的CDN的地图。例如,即使用户当前没有从那个内容起源购买CDN服务,也可以为源自Massachusetts的查询推荐适当的内容起源。因此,本发明的实施例可以被用于评估新的CDN服务购买。

[0165] 结论

[0166] 虽然各种发明性实施例已经在本文中进行了描述和图示,但本领域普通技术人员将容易想到用于执行本文所述的功能和/或获得本文所述的结果和/或优点中的一个或多个的各种其它装置和/或结构,并且这些变化和/或修改当中每一个都被认为在本文所述的发明性实施例的范围内。更一般而言,本领域技术人员将容易认识到的是,本文所述的所有参数、维度、材料和配置都意在是示例性的,并且实际的参数、维度、材料和配置将取决于本发明性教导所用于的具体应用(一个或多个)。本领域技术人员将认识到,或者能够仅仅使

用常规实验来确定,本文所描述的具体发明性实施例的许多等价物。因此,应当理解的是,前述实施例仅仅是作为示例给出的,并且,在所附权利要求及其等价物的范围内,发明性实施例可以以不同于具体描述并要求保护的那样的方式来实现。本公开的发明性实施例涉及本文所述的每个单独的特征、系统、制品、材料、套件、和/或方法。此外,如果这种特征、系统、制品、材料、套件、和/或方法并非相互不一致,那么两个或更多个这种特征、系统、制品、材料、套件、和/或方法的任意组合包括在本公开的发明性范围内。

[0167] 上述实施例可以以多种方式中的任何一种来实现。例如,设计并产生本文公开的技术的实施例可以使用硬件、软件或其组合来实现。当用软件实现时,软件代码可以在无论在单个计算机中提供或者在多个计算机之间分布的任何合适的处理器或处理器集合上执行。

[0168] 另外,应当认识到的是,计算机可以以多种形式中的任何一种来体现,诸如机架式计算机、台式计算机、膝上型计算机或平板计算机。此外,计算机可以嵌入在一般不被视为计算机但具有合适处理能力的设备(其包括个人数字助理(PDA)、智能电话或任何其它合适的便携式或固定电子设备)中。

[0169] 而且,计算机可以具有一个或多个输入和输出设备。除此之外,这些设备可以用于呈现用户接口。可以用于提供用户接口的输出设备的示例包括用于输出的视觉呈现的打印机或显示屏或用于输出的可听呈现的扬声器和其它声音生成设备。可以用于用户接口的输入设备的示例包括键盘以及定点设备(诸如鼠标、触摸板和数字化平板)。作为另一个示例,计算机可以通过语音识别或其它可听格式接收输入信息。

[0170] 这种计算机可以通过任何合适形式的一个或多个网络互连,包括局域网或广域网(诸如企业网络 and 智能网络(EN)或互联网)。这种网络可以基于任何合适的技术并且可以根据任何合适的协议来操作,并且可以包括无线网络、有线网络或光纤网络。

[0171] 本文概述的(例如,设计和制造上面公开的技术的)各种方法或处理可以被编码为可在采用各种操作系统或平台中的任何一种的一个或多个处理器上执行的软件。此外,这种软件可以使用多种合适的编程语言和/或编程或脚本工具中的任何一种来编写,并且还可以被编译为在框架或虚拟机上执行的可执行机器语言代码或中间代码。

[0172] 在这方面,各个发明性概念可以体现为利用一个或多个程序编码的计算机可读存储介质(或多个计算机可读存储介质)(例如,计算机存储器、一个或多个软盘、紧凑盘、光盘、磁带、闪速存储器、现场可编程门阵列或其它半导体器件中的电路构造、或者其它非瞬态介质或有形计算机存储介质),当所述一个或多个程序在一个或多个计算机或其它处理器上执行时,执行实现以上讨论的本发明的各个实施例的方法。计算机可读介质(一个或多个)可以是便携式的,使得存储在其上的程序(一个或多个)可以被加载到一个或多个不同的计算机或其它处理器上,以实现如上讨论的本发明的各个方面。

[0173] 本文使用的术语“程序”或“软件”在一般意义上指任何类型的计算机代码或处理器可执行指令集,该计算机代码或处理器可执行指令集可以用来对计算机或其它处理器编程以实现如以上讨论的实施例的各个方面。此外,应当认识到的是,根据一方面,当被执行时执行本发明的方法的一个或多个计算机程序无需驻存在单个计算机或处理器上,而是可以以模块方式分布在多个不同的计算机或处理器当中以实现本发明的各个方面。

[0174] 计算机可执行指令可以具有许多形式,诸如由一个或多个计算机或其它设备执行

的程序模块。一般而言,程序模块包括执行特定任务或实现特定抽象数据类型的例程、程序、对象、组件、数据结构等。通常,在各个实施例中可以如所期望地组合或分布程序模块的功能。

[0175] 而且,数据结构可以以任何合适形式存储在计算机可读介质中。为了简化说明,数据结构可以被示为具有通过数据结构中的位置相关的字段。这种关系同样可以通过以下方式来实现-在计算机可读介质中为具有用以传达字段之间的关系的的位置的字段指派存储空间。但是,可以使用任何合适的机制来建立数据结构的字段中的信息之间的关系,包括通过使用指针、标签或建立数据元素之间关系的其它机制。

[0176] 而且,各种发明性概念可以被体现为已经提供了其示例的一个或多个方法。作为该方法的一部分执行的动作可以以任何合适的方式进行排序。因此,实施例可以被构造为其中动作是以不同于图示的次序执行的,其可以包括同时执行一些动作,尽管其在说明性实施例中所示为顺序的动作。

[0177] 如本文所定义和使用的,所有的定义都应该被理解为优先于字典的定义控制通过引用被结合的文档中的定义和/或所定义术语的普通含义。

[0178] 如本文在说明书和在权利要求中所使用的,除非明确地指示为相反,否则不定冠词“一”和“一个”应当被理解为意在“至少一个”。

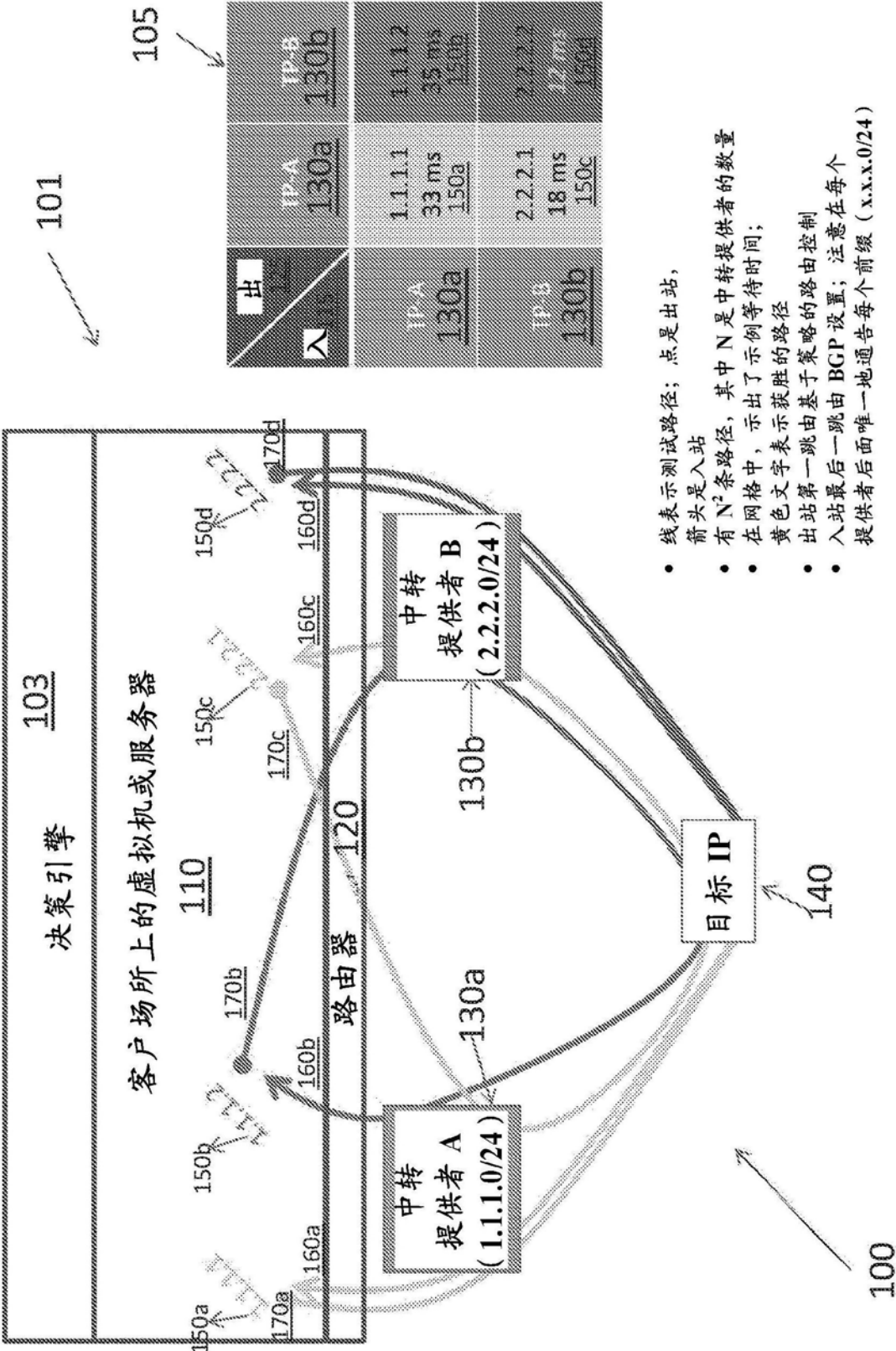
[0179] 如本文在说明书和在权利要求中所使用的,短语“和/或”应当被理解为意在如此连接的元素的“任一个或两者”,即,在一些情况下结合存在并且在另一些情况下分离存在的元素。利用“和/或”列出的多个元素应该以同样的方式进行解释,即,如此结合的元素的“一个或多个”。除由“和/或”子句具体识别的元素之外,其它元素也可以可选地存在,无论是与具体识别的那些元素有关还是无关。因此,作为非限制性的示例,当结合诸如“包括”的开放式语言使用时,对“A和/或B”的引用可以,在一个实施例中,仅仅指A(可选地包括除B之外的元素);在另一个实施例中,仅仅指B(可选地包括除A之外的元素);在还有的另一个实施例中,指A和B两者(可选地包括其它元素);等等。

[0180] 如本文在说明书和权利要求中所使用的,“或”应当被理解为具有与如上定义的“和/或”相同的含义。例如,当分隔列表中的项时,“或”或“和/或”应当被解释为包括性的,即,包括多个元素或元素列表当中至少一个,但也包括一个以上,以及可选地,包括附加的未列出的项目。只有明确地指示相反的术语,诸如“仅…之一”或者“确切地…之一”,或者在权利要求中使用“由…组成”,将指包括多个元素或元素列表当中确切一个元素。一般而言,如本文所使用的,当前面有排他性术语(诸如“任意一个”、“…之一”、“仅…之一”或者“确切地…之一”)时,如本文中使用的术语“或”将只被解释为指示排他性的替代(即“一个或另一个但不是两者”)。当在权利要求书中使用时,“基本上由…组成”将具有其如在专利法领域中使用的普通含义。

[0181] 如本文在说明书和权利要求中所使用的,短语“至少一个”在引用一个或多个元素的列表时应当理解为意味着从元素列表中的元素中的任何一个或多个元素中选择的至少一个元素,但是未必包括在元素列表内具体列举的每一个元素中的至少一个元素并且未排除元素列表中的元素的任意组合。这一定义也允许可以可选地存在除了在短语“至少一个”所引用的元素列表内具体识别的元素之外的、无论是与具体识别的那些元素有关还是无关的元素。因此,作为非限制性示例,“A和B中的至少一个”(或等价地,“A或B中的至少一个”,

或等价地,“A和/或B中的至少一个”)可以,在一个实施例中,指至少一个,可选地包括多于一个A,其中不存在B(并且可选地包括除B之外的元素);在另一个实施例中,指至少一个,可选地包括多于一个B,其中不存在A(并且可选地包括除A之外的元素);在还有的另一个实施例中,指至少一个,可选地包括多于一个A,和至少一个,可选地包括多于一个B(并且可选地包括其它元素);等等。

[0182] 在权利要求中,以及在上述说明书中,所有过渡短语(诸如“包含”、“包括”、“携带”、“具有”、“含有”、“涉及”、“持有”、“有”等)被理解为是开放式的,即,意味着包括但不限于。只有过渡短语“由…组成”和“基本上由…组成”才将分别是封闭或半封闭过渡短语,如在美国专利局专利审查程序手册2111.03节中所阐述的。



- 线表示测试路径；点是出站，箭头是入站
- 有 N^2 条路径，其中 N 是中转提供者的数量
- 在表格中，示出了示例等待时间；黄色文字表示获胜的路径
- 出站最后一跳由策略的路由控制
- 入站最后一跳由 BGP 设置；注意在每个提供者后面唯一地通告每个前缀 (x.x.x.0/24)

图1A

图1B

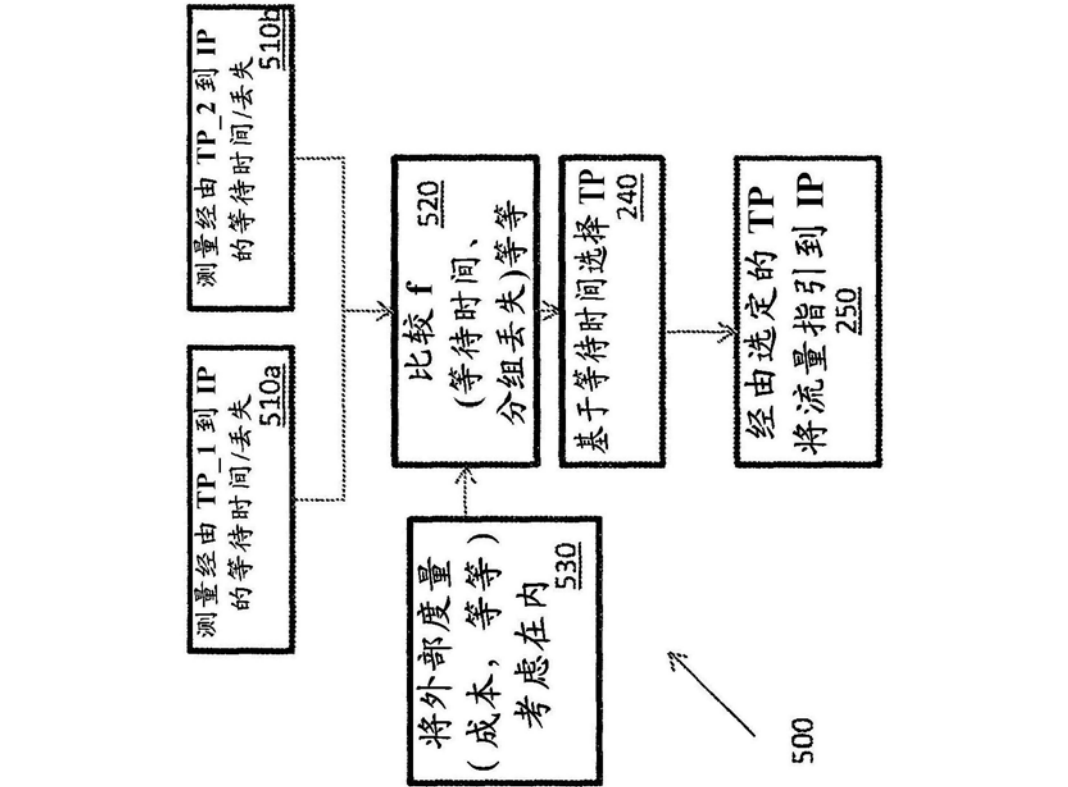


图2

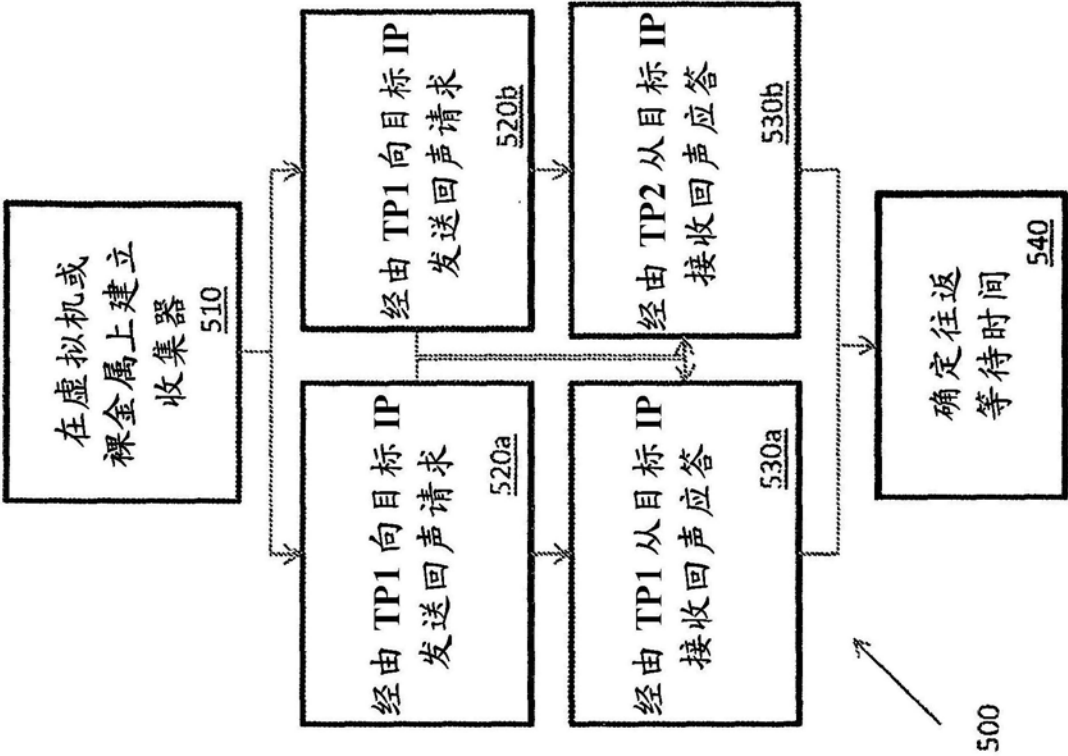


图3

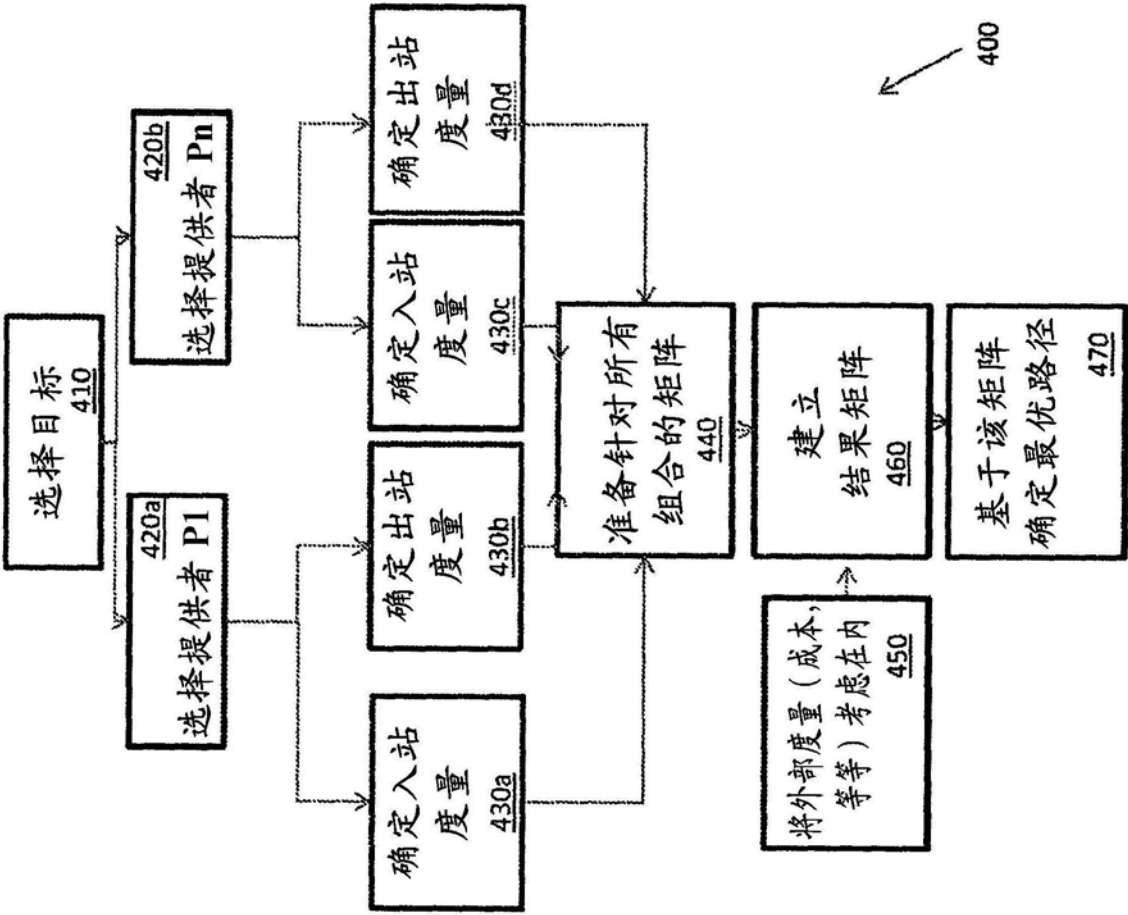


图4

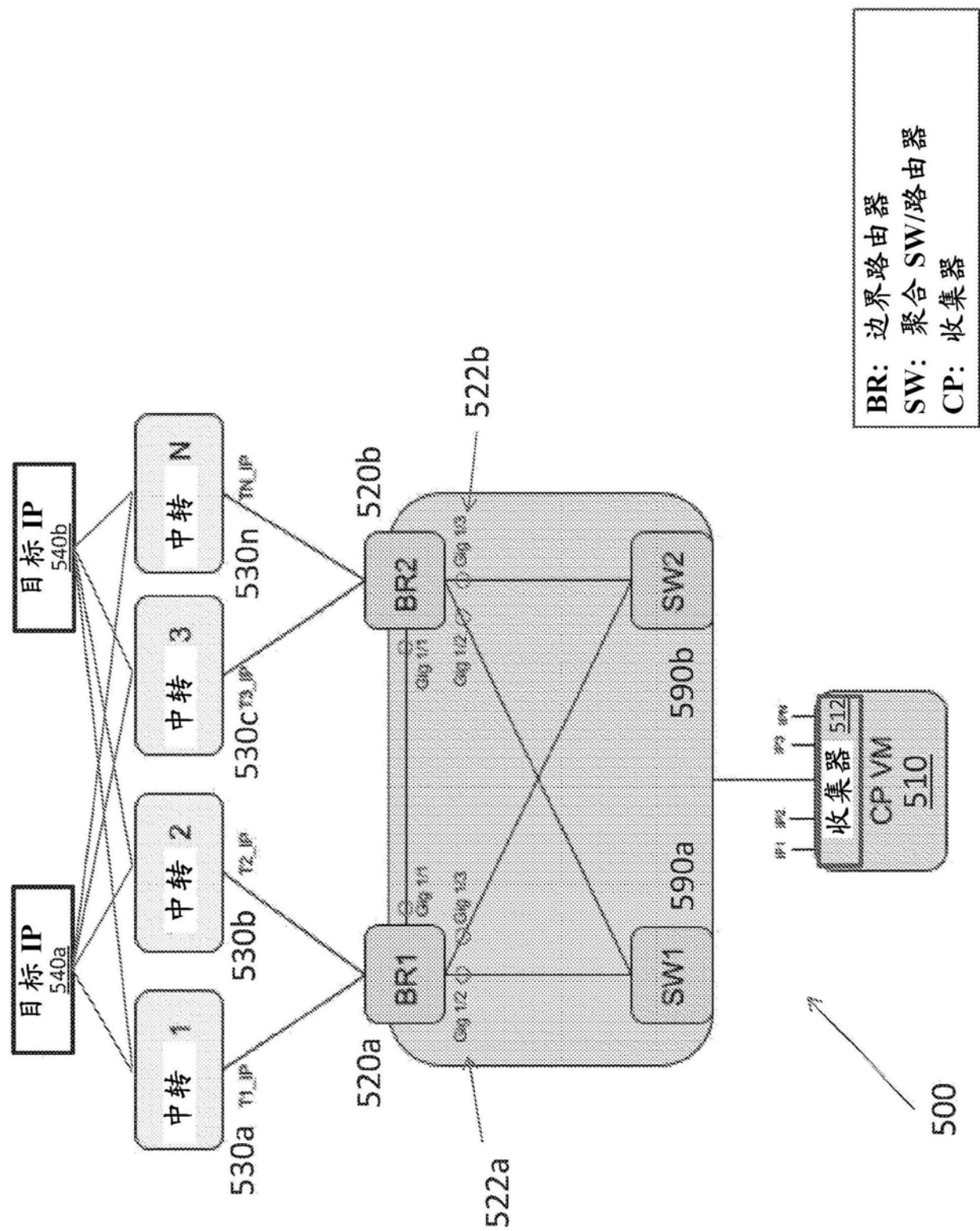


图5A

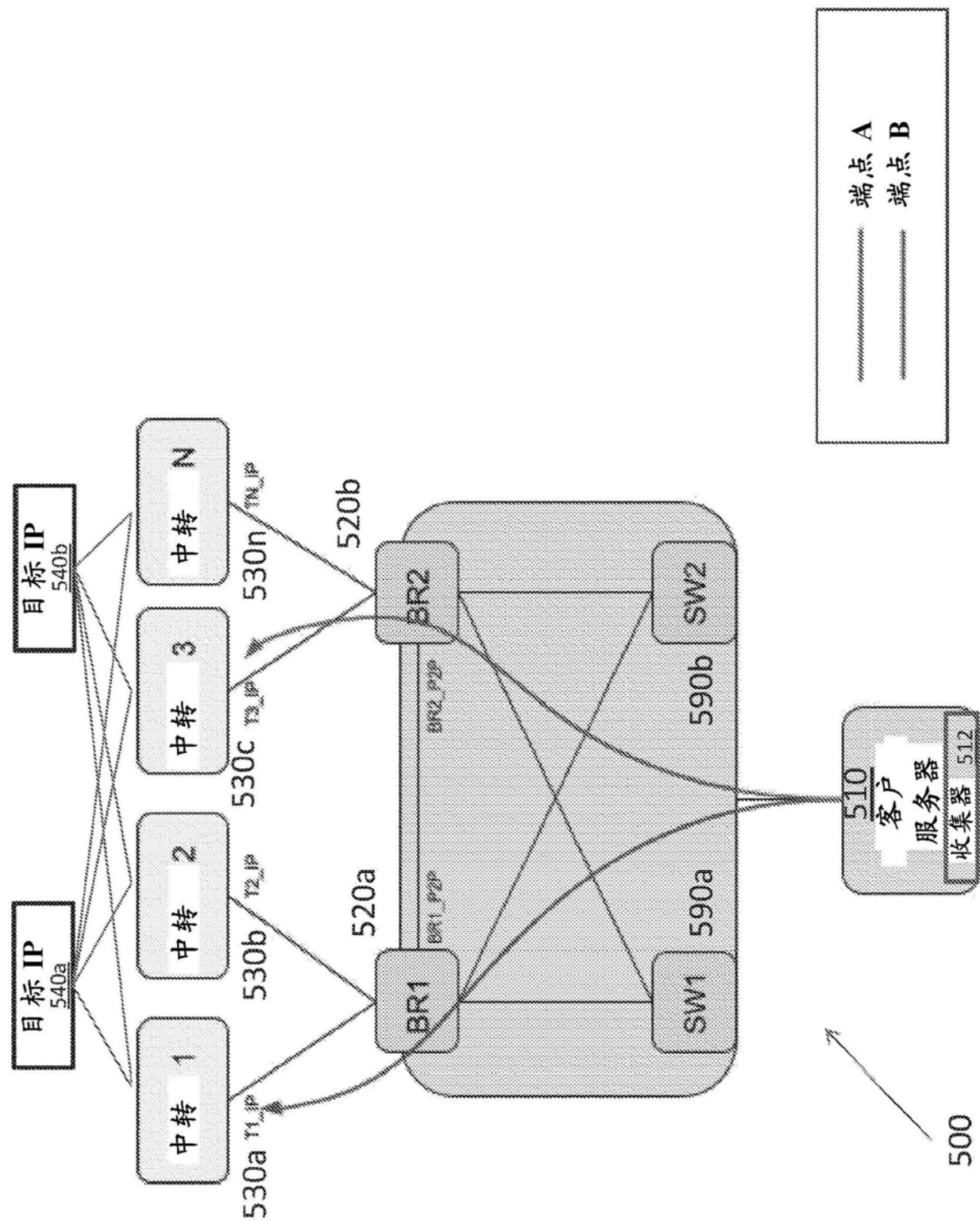


图5B

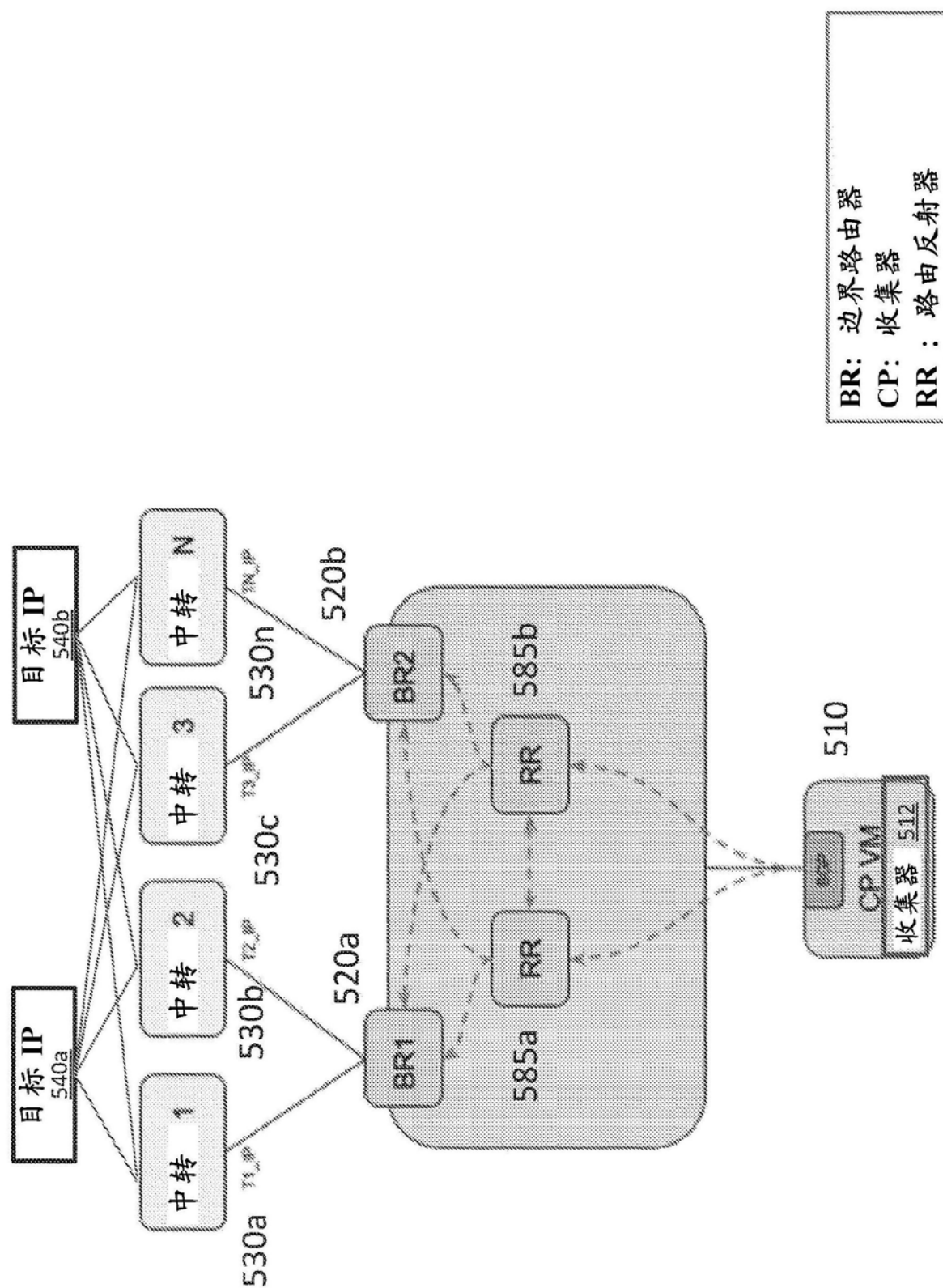


图5C

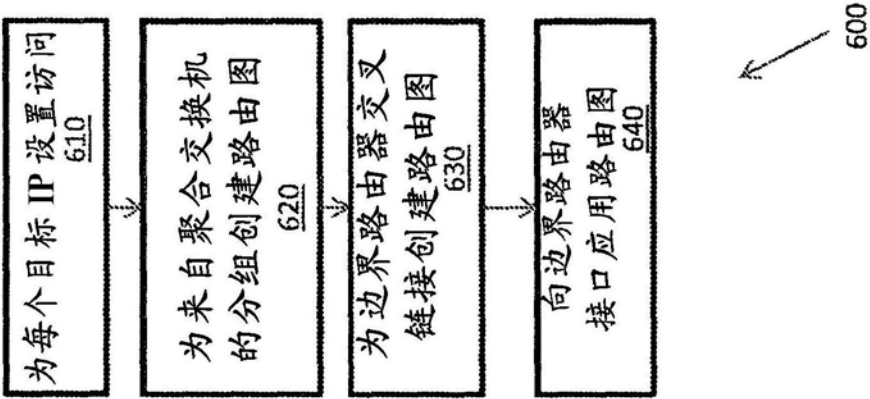


图6

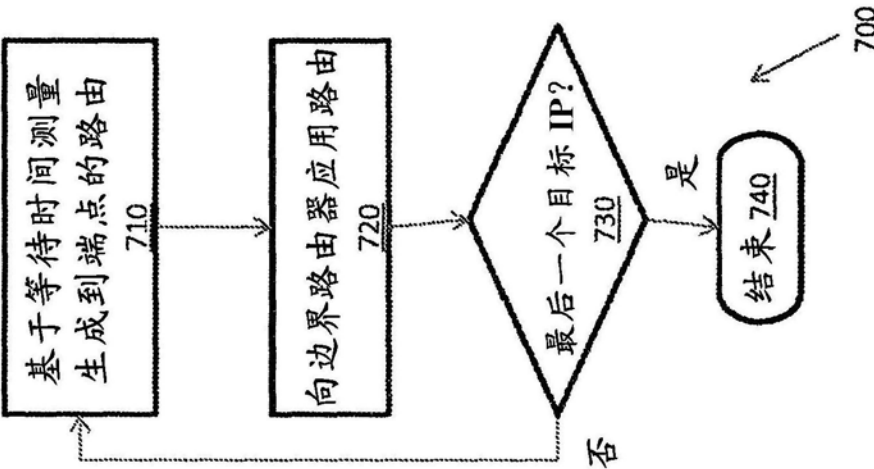


图7

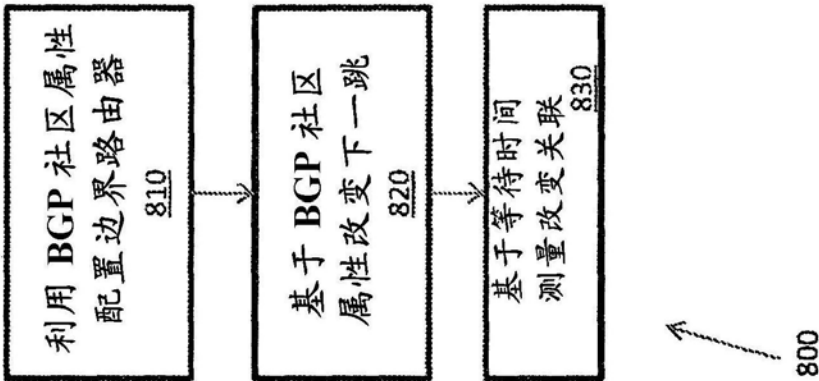


图8

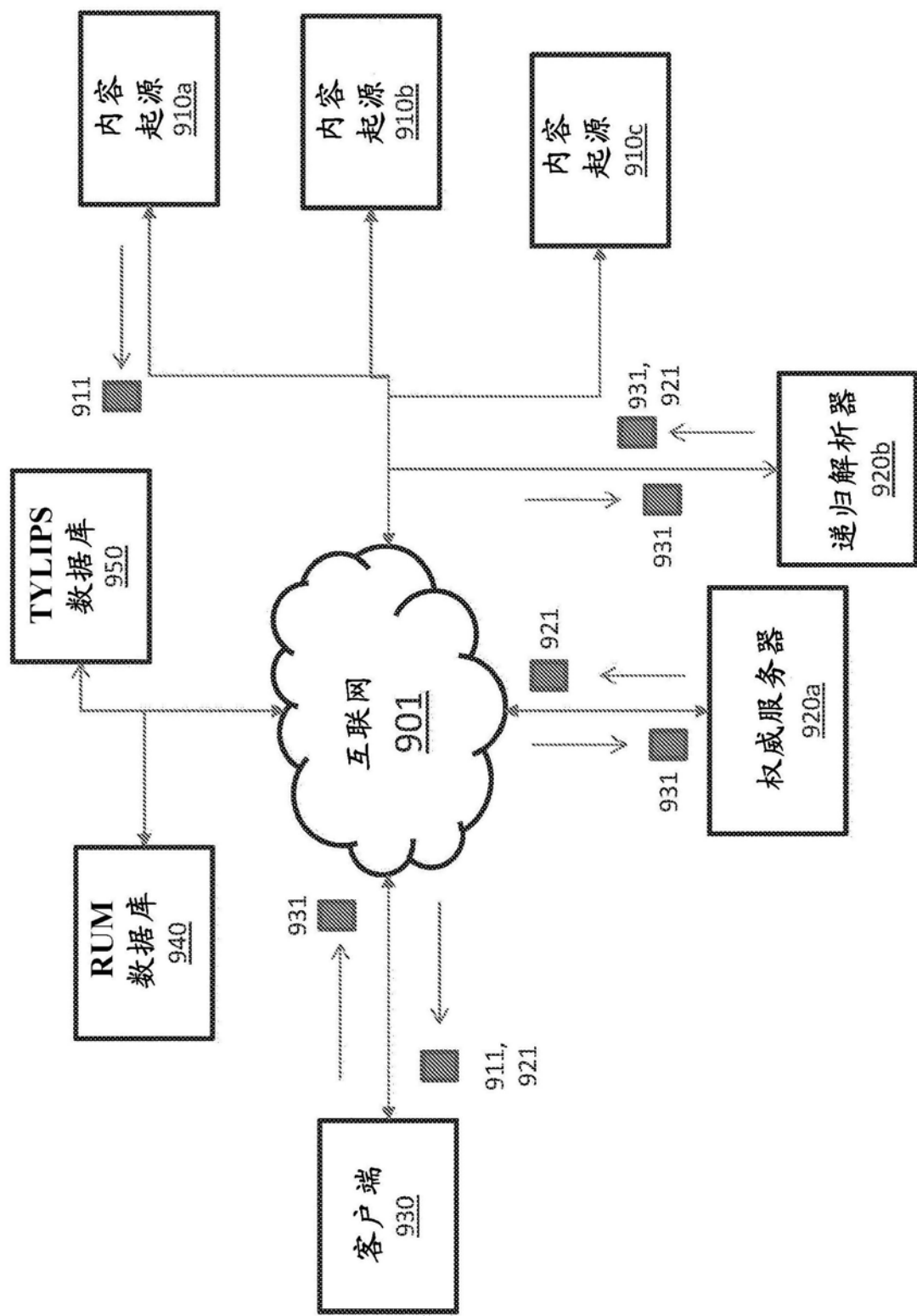


图9

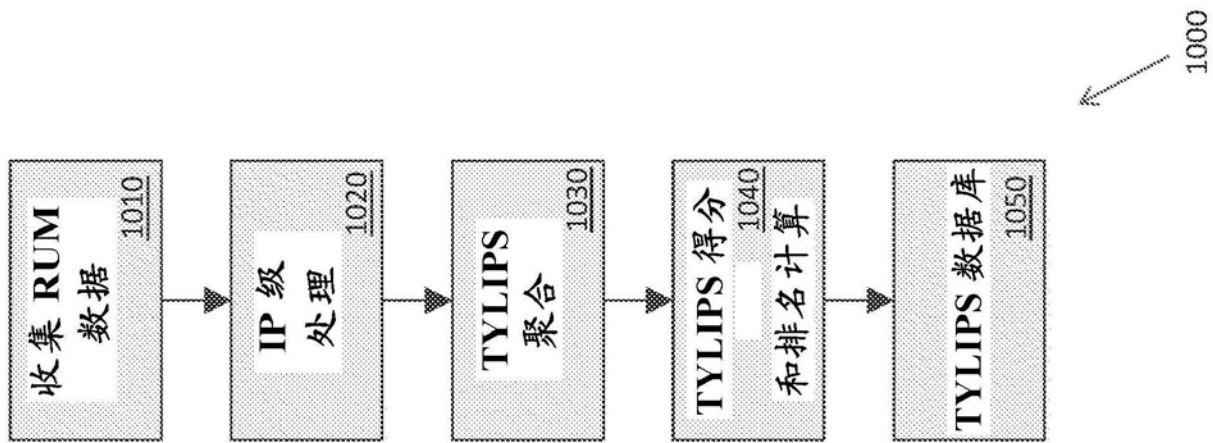


图10

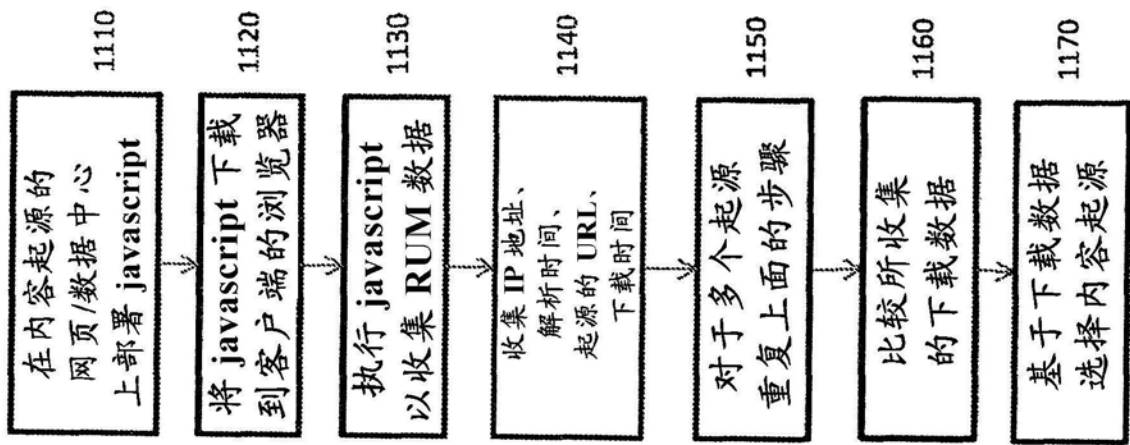


图11

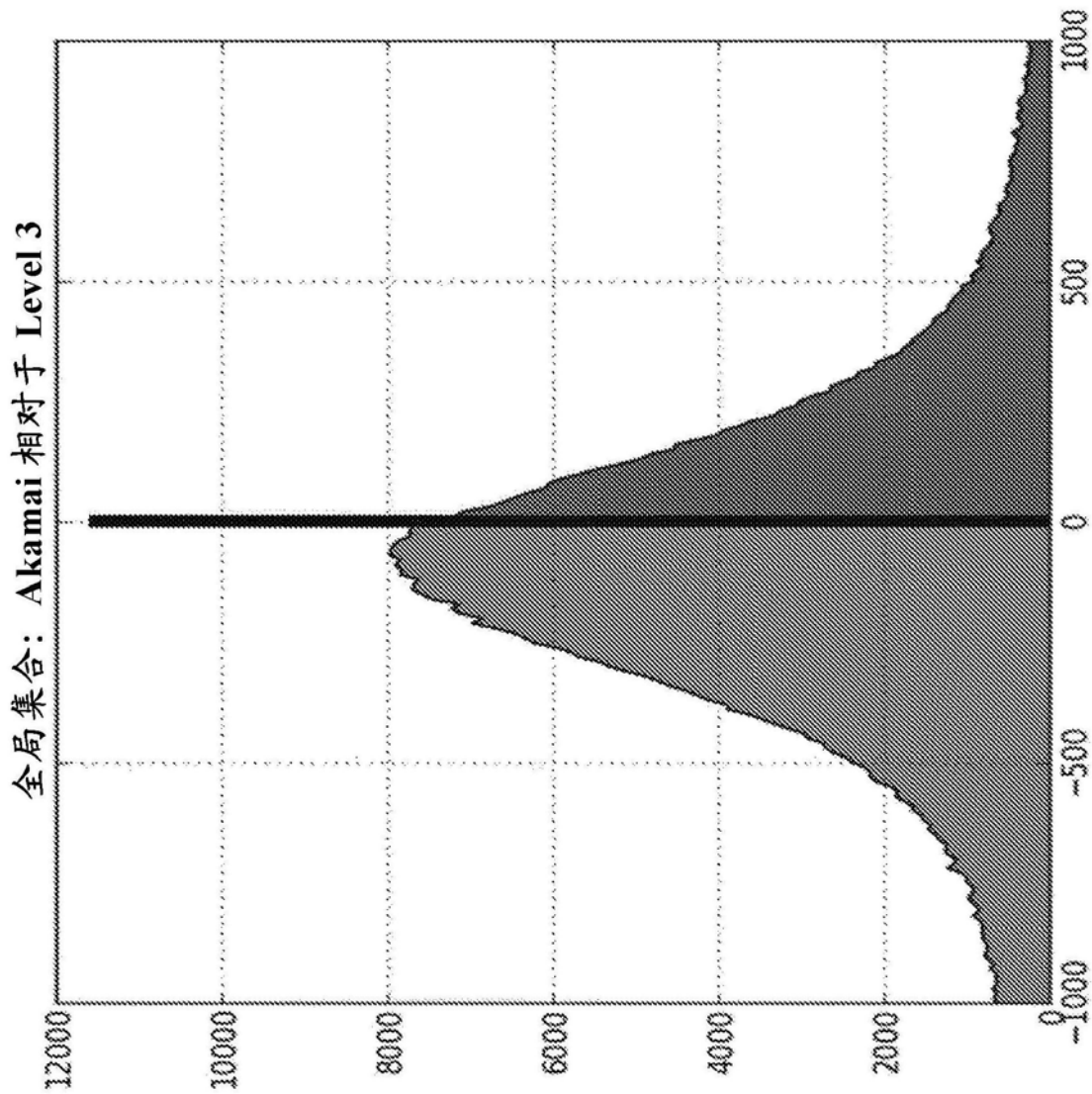


图12

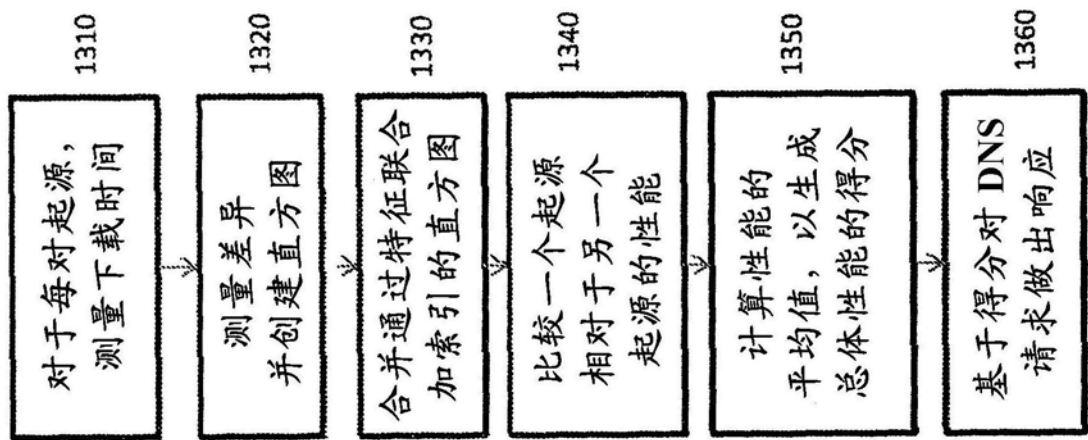


图13

图示处理 - 选择 TYLIPS - 更具体到更不具体

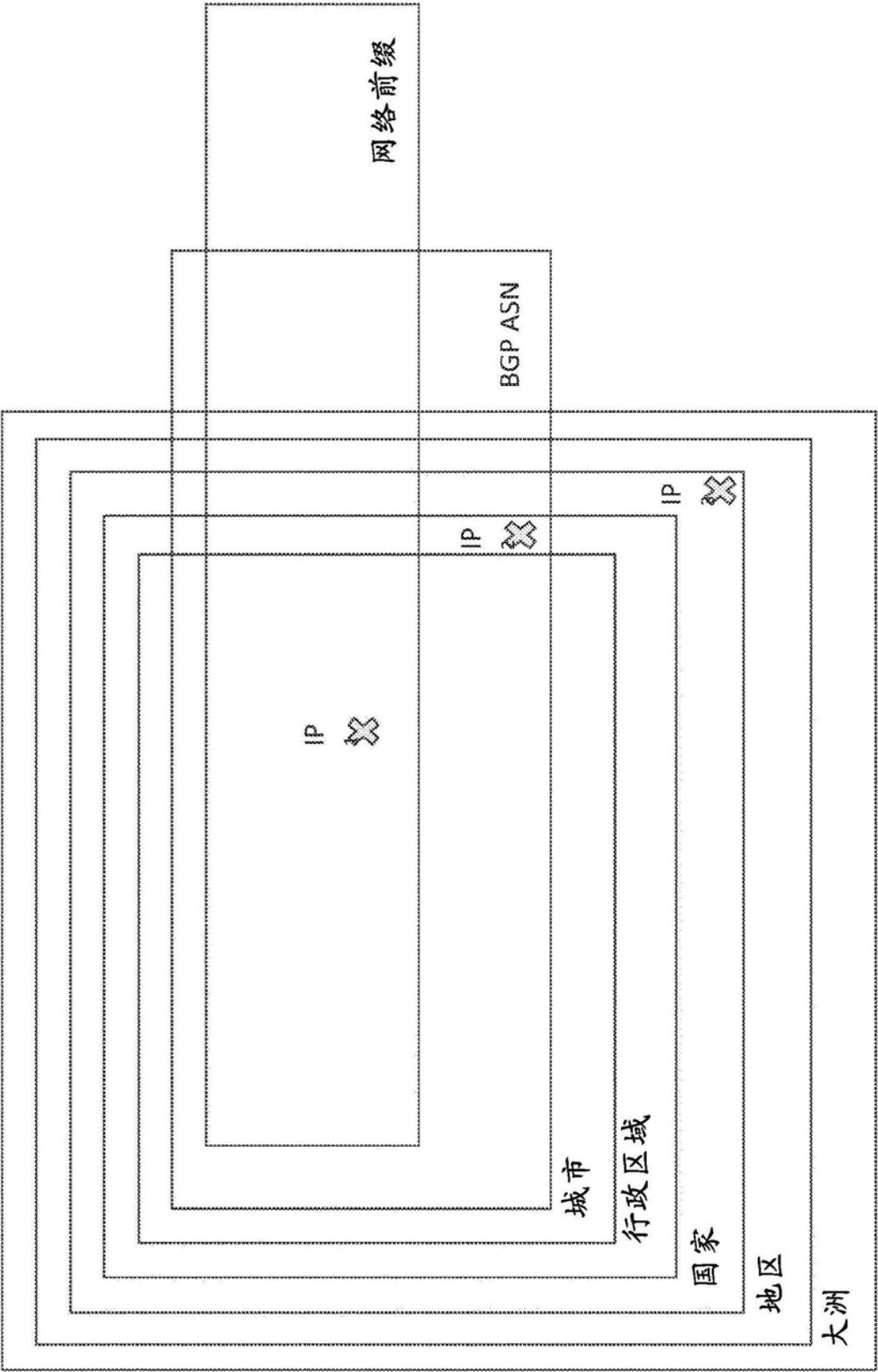


图14A

图示处理 - 选择 TYLIPS

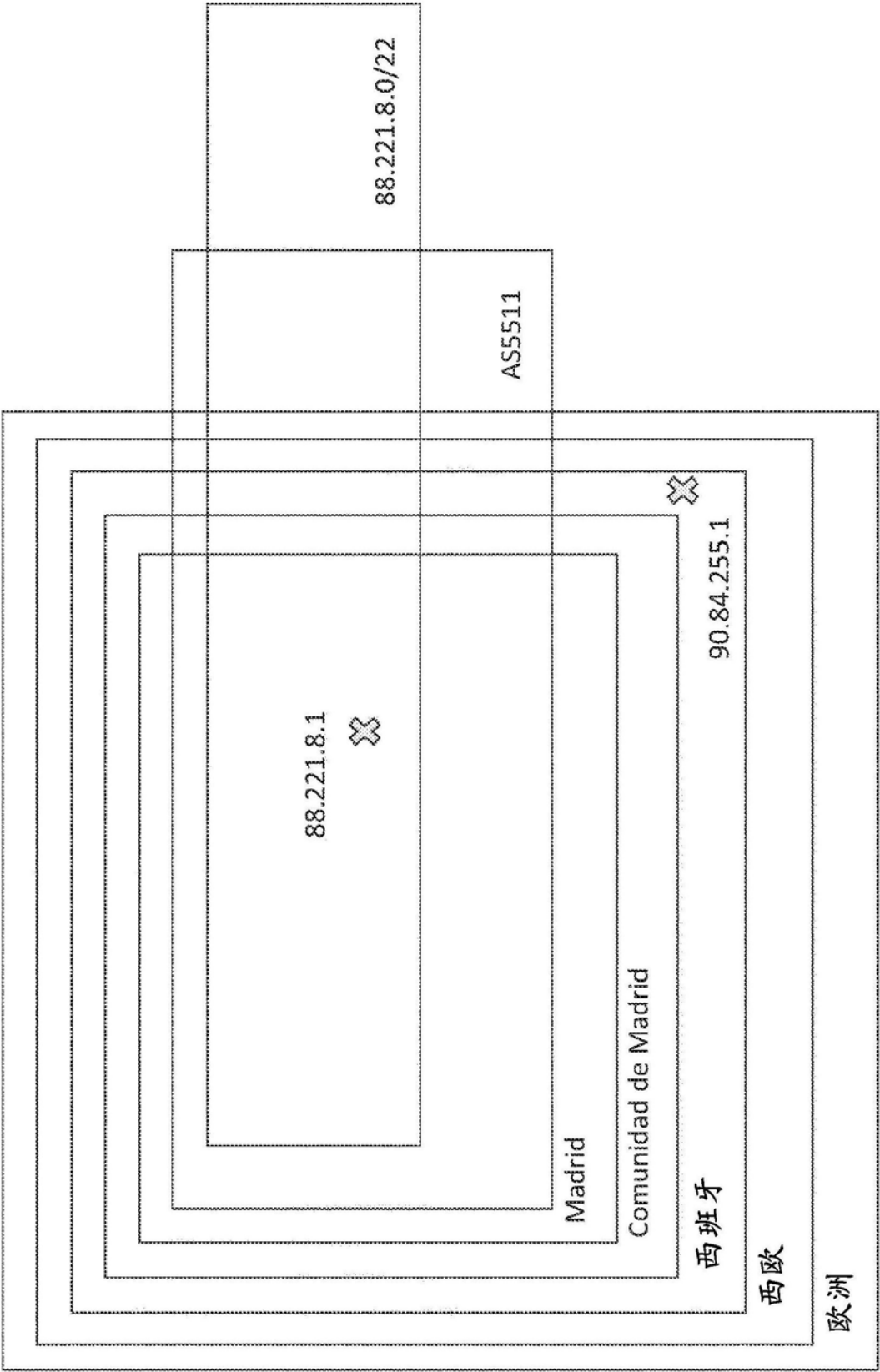


图14B

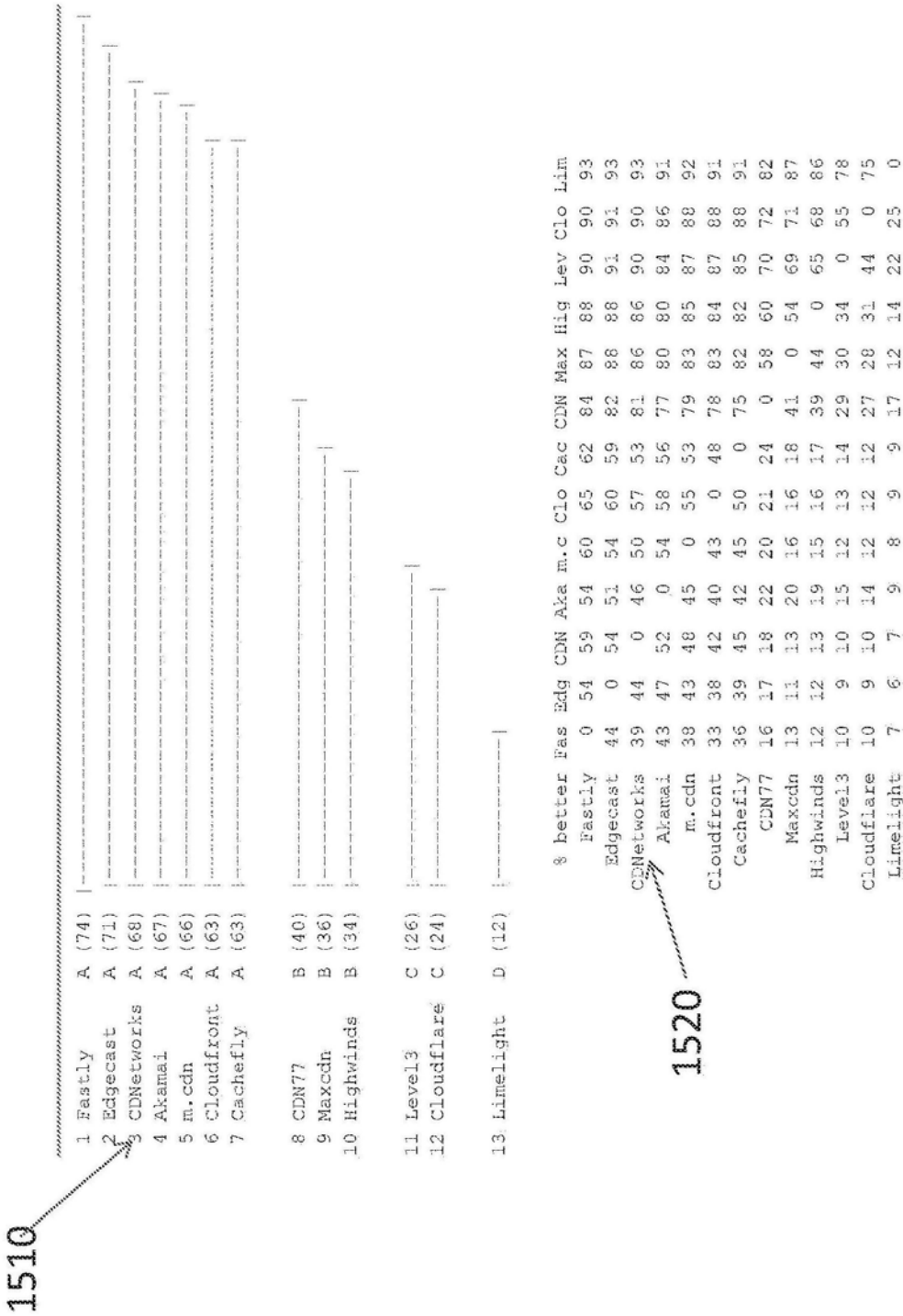


图15

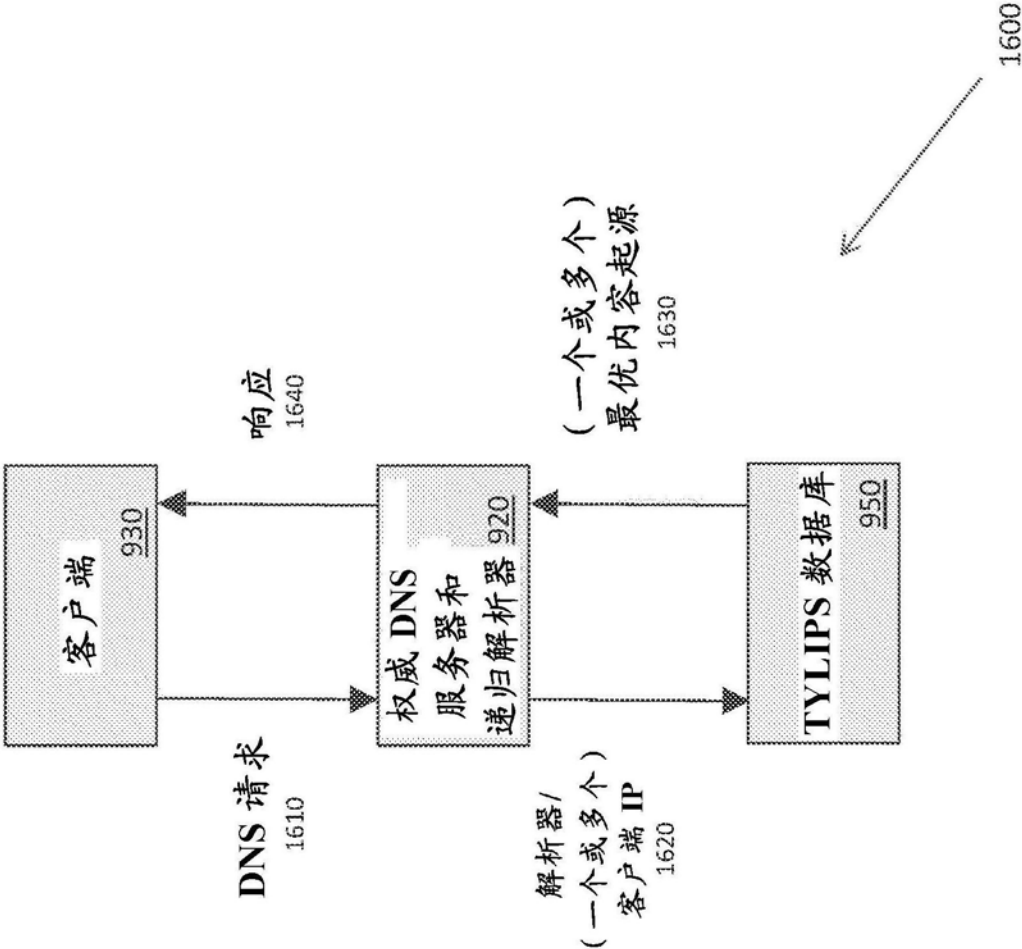


图16