



(21) 申请号 202110658145.0

(22) 申请日 2016.07.21

(65) 同一申请的已公布的文献号  
申请公布号 CN 113315682 A

(43) 申请公布日 2021.08.27

(30) 优先权数据  
62/195,478 2015.07.22 US

(62) 分案原申请数据  
201680051441.1 2016.07.21

(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·W·霍博古德 S·J·布朗  
J·H·考威

(74) 专利代理机构 中国贸促会专利商标事务所  
有限公司 11038  
专利代理师 刘前红

(51) Int.Cl.

H04L 43/0852 (2022.01)

H04L 43/10 (2022.01)

H04L 43/106 (2022.01)

H04L 43/16 (2022.01)

H04L 41/0677 (2022.01)

H04L 61/4511 (2022.01)

G06F 11/30 (2006.01)

(56) 对比文件

CN 101505230 A, 2009.08.12

CN 1989745 A, 2007.06.27

US 2007250625 A1, 2007.10.25

US 2006143496 A1, 2006.06.29

WO 0198916 A1, 2001.12.27

郭琳; 张大方; 黎文伟; 谢鲲. 基于行为模型的IP Forwarding异常检测方法. 计算机应用. 2006, (03), 第53-55页.

审查员 包文韬

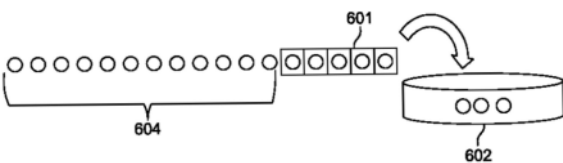
权利要求书3页 说明书17页 附图18页

(54) 发明名称

生成信息传输性能警告的方法、系统和装置

(57) 摘要

本发明涉及生成信息传输性能警告的方法、系统和装置。提供了一种警告系统和方法，以识别和表征高频的全球和本地跟踪路由数据中的实时信息传输异常。该系统包括活动的网络传感器和/或收集器设备，其收集与到计算机网络中的不同点的传输相关联的跟踪路由数据，并将跟踪路由数据提供给主服务器。通过将探测数据分组发送到位于目标设备本地和/或全球的多个计算目标设备，跟踪路由数据由活动的网络传感器获得。主服务器根据接收到的跟踪路由数据确定一个或多个异常，并依据类型、严重性、位置、受影响的域名系统 (DNS) 服务器和/或互联网服务提供商 (ISP) 来表征异常。



1. 一种包括指令的非瞬态计算机可读介质,所述指令当由一个或多个硬件处理器执行时引起执行包括以下的操作:

估计与计算机网络中的第一点和计算机网络中的第二点之间的网络流量相关联的基线延迟状态;

检测与计算机网络中的第一点和计算机网络中的第二点之间的所述网络流量相关联的当前延迟状态,所述当前延迟状态相对于所述基线延迟状态有偏差;

识别与相对于所述基线延迟状态有偏差的所述当前延迟状态对应的异常;

识别包括以下至少之一的一个或多个网络设备:

引起了所述异常的第一设备;

发送了因所述异常而延迟的网络流量的第二设备;和

接收了因所述异常而延迟的网络流量的第三设备;以及向识别出的一个或多个网络设备发送与所述异常对应的警告。

2. 如权利要求1所述的计算机可读介质,其中所述警告还包括给所述一个或多个网络设备的指令,该指令使所述一个或多个网络设备解决所述当前延迟状态相对于所述基线延迟状态的偏差。

3. 如权利要求2所述的计算机可读介质,其中:

所述一个或多个网络设备包括域名系统(DNS)服务器;并且

给所述DNS服务器以解决所述偏差的指令包括从(a)将所述网络流量转向受所述异常影响的第一设备的第一查询解析策略变为(b)将所述网络流量转向不受所述异常影响的第二设备的第二查询解析策略的指令。

4. 如权利要求3所述的计算机可读介质,其中:

第一查询解析策略将所述网络流量转向第一地理区域、第一组DNS查询源或第一组互联网协议(IP)地址中的一个或多个;以及

第二查询解析策略将所述网络流量转向不与所述当前延迟状态相对于所述基线延迟状态的偏差相关联的第二地理区域、第二组DNS查询源或第二组互联网协议(IP)地址中的一个或多个。

5. 如权利要求2所述的计算机可读介质,其中:

所述一个或多个网络设备包括网络路由器;并且

给所述网络路由器的指令包括从经由受所述异常影响的第一网络路由转向网络流量的第一路由策略变为从经由不同于所述第一网络路由的第二网络路由转向网络流量的第二路由策略的指令,其中所述第二网络路由不受所述异常影响。

6. 如权利要求1所述的计算机可读介质,其中所述异常包括一组受影响的互联网协议(IP)地址、受影响的地理区域和错误类型中的一个或多个。

7. 如权利要求1所述的计算机可读介质,其中所述一个或多个网络设备包括负载平衡器、网络优化设备和配置管理系统中的一个或多个。

8. 如权利要求1所述的计算机可读介质,其中所述警告包括与所述当前延迟状态相关联的一个或多个属性。

9. 一种包括指令的非瞬态计算机可读介质,所述指令当由一个或多个硬件处理器执行时引起执行包括以下的操作:

估计与(a)计算机网络中的第一节点和计算机网络中的第二节点之间的第一路由或(b)特定地理区域相关联的基线延迟状态;

计算与(a)计算机网络中的第一节点和计算机网络中的第二节点之间的第一路由或(b)特定地理区域相关联的当前延迟状态;

响应于确定所述当前延迟状态相对于所述基线延迟状态有偏差,识别异常;

响应于确定(a)所述异常与包括引起了所述异常的第一网络设备、发送了因所述异常而延迟的网络流量的第二网络设备和接收了因所述异常而延迟的网络流量的第三网络设备的多个网络设备中的一个或多个相关联以及(b)所述多个网络设备中的第四网络设备满足(i)不引起所述异常以及(ii)不受所述异常影响中的至少一项;

从包括所述第一网络设备、所述第二网络设备、所述第三网络设备和所述第四网络设备中的一个或多个的一组设备中选择所述第四网络设备;

将与所述异常对应的警告发送到所述第一网络设备、所述第二网络设备和所述第三网络设备中识别出的一个或多个;以及

将流量转向与所述第四网络设备对应的IP地址。

10.如权利要求9所述的计算机可读介质,其中所述选择和转向操作由DNS服务器执行。

11.如权利要求9所述的计算机可读介质,其中所述选择和转向操作由网络路由器执行。

12.如权利要求9所述的计算机可读介质,其中所述异常包括一组受影响的互联网协议(IP)地址、受影响的地理区域和错误类型中的一个或多个。

13.一种系统,包括:

包含硬件处理器的至少一个设备;

所述系统被配置为执行包括以下的操作:

估计与计算机网络中的第一点和计算机网络中的第二点之间的网络流量相关联的基线延迟状态;

检测与计算机网络中的第一点和计算机网络中的第二点之间的所述网络流量相关联的当前延迟状态,所述当前延迟状态相对于所述基线延迟状态有偏差;

识别与相对于所述基线延迟状态有偏差的所述当前延迟状态对应的异常;

识别包括以下至少之一的一个或多个网络设备:

引起了所述异常的第一设备;

发送了因所述异常而延迟的网络流量的第二设备;和

接收了因所述异常而延迟的网络流量的第三设备;以及

向识别出的一个或多个网络设备发送与所述异常对应的警告。

14.如权利要求13所述的系统,其中所述警告还包括给所述一个或多个网络设备的指令,该指令使所述一个或多个网络设备解决所述当前延迟状态相对于所述基线延迟状态的偏差。

15.如权利要求14所述的系统,其中:

所述一个或多个网络设备包括域名系统(DNS)服务器;并且

给所述DNS服务器以解决所述偏差的指令包括从(a)将所述网络流量转向受所述异常影响的第一设备的第一查询解析策略变为(b)将所述网络流量转向不受所述异常影响的第

二设备的第二查询解析策略的指令。

16. 如权利要求15所述的系统, 其中:

第一查询解析策略将所述网络流量转向第一地理区域、第一组DNS查询源或第一组互联网协议 (IP) 地址中的一个或多个; 以及

第二查询解析策略将所述网络流量转向不与所述当前延迟状态相对于所述基线延迟状态的偏差相关联的第二地理区域、第二组DNS查询源或第二组互联网协议 (IP) 地址中的一个或多个。

17. 如权利要求14所述的系统, 其中:

所述一个或多个网络设备包括网络路由器; 并且

给所述网络路由器的指令包括从经由受所述异常影响的第一网络路由转向网络流量的第一路由策略变为从经由不同于所述第一网络路由的第二网络路由转向网络流量的第二路由策略的指令, 其中所述第二网络路由不受所述异常影响。

18. 如权利要求13所述的系统, 其中所述异常包括一组受影响的互联网协议 (IP) 地址、受影响的地理区域和错误类型中的一个或多个。

19. 如权利要求13所述的系统, 其中所述一个或多个网络设备包括负载均衡器、网络优化设备和配置管理系统中的一个或多个。

20. 如权利要求13所述的系统, 其中所述警告包括与所述当前延迟状态相关联的一个或多个属性。

## 生成信息传输性能警告的方法、系统和装置

[0001] 本申请是申请号为201680051441.1、申请日为2016年7月21日、发明名称为“生成信息传输性能警告的方法、系统和装置”的发明专利申请的分案申请。

[0002] 相关申请的交叉引用

[0003] 本申请援引35U.S.C.§119(e)要求于2015年7月22日提交的标题为“Methods, Systems, and Apparatus to Generate Information Transmission Performance Alerts”的美国申请62/195478的优先权益。上面引用的申请通过引用整体上并入本文。

### 背景技术

[0004] 互联网是使用分组交换协议的互联网协议(IP)族进行通信的互连计算机网络的全球系统。连接到互联网的设备通过交换寻址到特定IP地址(通常表示为多位数字(例如,对于IPv4是32位,对于IPv6是128位))的信息分组来相互通信。

[0005] 路由器、交换机、桥接器和其它设备基于分组的IP地址将每个分组朝着其目的地转发。一些设备(诸如路由器)可以基于到分组的目的地不同路径上的流量来路由分组,例如以避免拥塞或平衡跨可用路径的流量负载。路由器也可以基于分组所携带的数据的类型来路由分组。例如,路由器可以在其它分组之前发送携带容易受延迟(latency)影响的数据(诸如语音或视频数据)的分组,以避免不必要地增加延迟。

[0006] 连接到互联网的每个路由器都有一个或多个自己的IP地址,并构成互联网上的众多节点之一。因为互联网使用分组交换并且具有许多节点,因此通过动态地路由受影响区域周围的流量,可以承受由于装备故障、自然灾害、政治事件等引起的关键节点的损失。尽管如此,节点的暂时丢失甚至也可能通过减少连接性或增加延迟而中断互联网服务。互联网服务的中断可能使得内容提供商向受影响区域中的以及可能超出受影响区域的消费者可靠地交付内容更加昂贵。

### 发明内容

[0007] 发明人已经认识到,由节点故障等造成的互联网服务的中断有时会导致与正常网络性能模式的偏差。这些偏差(也称为异常)可以使用全球分布的跟踪路由(traceroute)数据收集器系统获取的跟踪路由数据来检测。如果系统检测到异常,那么它可以发出关于异常的性能警告。在一些情况下,系统可以表征异常并且基于异常的特点来确定负责造成异常的互联网服务中断的存在、位置和/或类型。

[0008] 在一些实现中,主服务器从第一网络传感器或跟踪路由数据收集器接收指示网络传感器与目标计算设备之间的传输延迟的跟踪路由数据。主服务器可以通过确定相对于目标计算设备的预期延迟与当前延迟之间的差是否显著和/或超出预定义阈值来识别异常指示。类似地,主服务器可以接收由第二网络传感器或跟踪路由数据收集器发送的跟踪路由数据分组。主服务器可以通过确定相对于目标计算设备的预期延迟与当前延迟之间的差是否显著和/或超出预定义阈值来识别异常指示。

[0009] 主服务器可以利用从跟踪路由数据收集器收集的数据导出的一个或多个异常指

示来计算相干性值 (coherency value) 或置信度值 (confidence value), 并且进一步至少基于从第一和第二跟踪路由数据收集器计算出的异常指示来进行网络性能异常推断。

[0010] 在一些实现中, 主服务器可以估计与计算机网络中的第一点和计算机网络中的第二点之间的信息传输相关联的基线延迟状态。这种基线延迟状态表示存储在基线库 (baseline reservoir) 中的离散延迟值的分布。另外, 主服务器可以估计与计算机网络中的第一点和计算机网络中的第二点之间的信息传输相关联的当前延迟状态。这种当前延迟状态表示存储在最近样本队列中的离散延迟值的分布。随后, 主服务器可以通过执行所估计的基线延迟状态与所估计的当前延迟状态的比较来识别第一点与第二点之间的异常的存在。

[0011] 在一些实现中, 主服务器从众多网络传感器或跟踪路由数据收集器接收跟踪路由数据。跟踪路由数据分组包含众多跟踪路由数据收集器中的每一个跟踪路由数据收集器与一组目标计算设备之间的传输延迟数据。传输延迟数据可以表示在不同时间段观察到的传输延迟。例如, 跟踪路由数据的第一集合可以表示从时间T1至时间T2观察到的传输延迟, 而跟踪路由数据的第二集合可以表示在稍后 (并且可能重叠) 的时间段 (例如从T3至T4) 观察到的传输延迟。

[0012] 主服务器可以基于跟踪路由数据的第一集合来计算每个跟踪路由数据收集器与目标计算设备之间的预期延迟值。此外, 主服务器可以基于跟踪路由数据的第二集合计算每个跟踪路由数据收集器与目标计算设备之间的当前延迟值。主服务器还可以计算每个跟踪路由数据收集器相对于计算目标设备或一组计算目标设备的当前延迟值与预期延迟值之间的差。此后, 主服务器可以推断与一个或多个跟踪路由数据收集器和一个目标计算设备和/或一组目标计算设备相关联的异常。

[0013] 关于异常的信息可以以各种不同的方式被使用, 以改进计算机网络的性能。例如, 主服务器可以发出关于异常类型、严重性、位置等的自动通知。这种通知的接收方 (包括用户、路由器和服务器) 可以使用关于异常的这个信息来适当地路由流量, 例如, 以减少拥塞、避免受连接性问题影响的区域、负载平衡、减少延迟, 等等。例如, 如果域名系统 (DNS) 服务器接收到异常通知, 那么DNS服务器可以基于异常来响应DNS请求。类似地, 如果异常通知指示异常影响第一互联网服务提供商 (ISP) 但不影响第二ISP, 那么路由器可以经由第二ISP路由流量, 以避免异常。

[0014] 应当认识到, 前述概念和下面更详细讨论的附加概念的所有组合 (假设这些概念不相互矛盾) 被认为是本文公开的发明性主题的一部分。特别地, 出现在本公开结尾处的所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, 本文明确采用的、也可能出现在通过引用并入的任何公开中的术语应该被赋予与本文公开的特定概念最一致的含义。

## 附图说明

[0015] 本领域技术人员将理解的是, 附图主要是用于说明性目的, 并非旨在限制本文描述的发明性主题的范围。附图不一定按比例绘制; 在一些情况下, 本文公开的发明性主题的各个方面可能在附图中被夸大或放大以便于理解不同的特征。在附图中, 相同的标号一般指代相同的特征 (例如, 功能上相似和/或结构上相似的元件)。

- [0016] 图1例示了包括主服务器、路由设备、跟踪路由收集器设备和目标计算设备的性能警告系统的示例。
- [0017] 图2例示了包括硬件和软件模块的主服务器的示例的框图。
- [0018] 图3例示了包括硬件和软件模块的跟踪路由收集器设备的示例的框图。
- [0019] 图4例示了通过递增TTL来确定特定跟踪路由(在图1中用黑色箭头示出)的处理。
- [0020] 图5A例示了由收集器针对目标计算设备获得的跟踪路由数据的示例。
- [0021] 图5B-5D例示了数据平滑或规范化处理的示例。
- [0022] 图6A-6C例示了包括新的跟踪路由数据、最近的跟踪路由数据收集和基线跟踪路由数据收集的跟踪路由数据替换策略。
- [0023] 图7例示了包括目标计算设备、收集器和路由设备的全球分布系统的子集。
- [0024] 图8绘出了根据本发明一个实施例的用于检测计算机网络中的性能异常的方法。
- [0025] 图9绘出了根据本发明另一个实施例的用于检测计算机网络中的异常的方法。
- [0026] 图10绘出了根据本发明另一个实施例的用于检测包括多个收集器的计算机网络中的性能异常的方法。
- [0027] 图11A是从一组三个不同的(但地理上接近的)跟踪路由收集器对于由位于伊拉克的ScopeSky通信和互联网有限公司(AS50597)控制的一群IP地址测得的往返延迟的图。
- [0028] 图11B是从一组三个类似的收集器对于由位于中国的Pacnet控制的特定网络前缀(222.126.128.0/23)中的IP地址测得的往返延迟的图。
- [0029] 图11C是从一组三个类似的收集器对于由位于南非的TENET(AS2018)控制的地址测得的往返延迟的图。
- [0030] 图11D是从一组三个类似的收集器对于由位于罗马尼亚的布加勒斯特的RCS&RDS控制的特定网络前缀(213.154.100.0/23)中的IP地址测得的往返延迟的图。
- [0031] 图12例示了由警告系统生成的警告消息内包括的内容的示例。
- [0032] 当结合附图时,通过下面阐述的具体实施方式,本技术的特征和优点将变得更加清楚。

## 具体实施方式

[0033] 下面是对在本地和全球范围内生成信息传输性能警告的发明性系统、方法和装置的相关的各种概念以及实施例的更详细描述。应当认识到,以上介绍的和下面更详细讨论的各种概念可以以许多方式中的任何方式来实现,因为所公开的概念不限于任何特定的实现方式。具体实现和应用的示例主要是为了说明性目的而提供。

### [0034] 介绍

[0035] 互联网性能的一个度量是与在互联网上的两个点之间发送分组相关联的时延或延迟。总延迟取决于分组在其源和目的地之间传播所花费的时间。延迟具有由分开两个点的物理距离和光速的乘积给出的下限。在实践中,总延迟取决于分组在两个点之间沿着光纤传播所花费的时间。总延迟还取决于中间设备(诸如路由器、交换机和桥接器)在分组的源和目的地之间路由分组所花费的时间。在一些情况下,特定的一对点之间可以存在许多路径,并且每条路径可以具有不同的延迟。由于流量拥塞、装备故障等等,每条路径的延迟也可以波动。在一些情况下,这些波动可以由互联网使用中的周期性变化(例如,与晚上的

流传输视频相关联的每日波动)造成。在其它情况下,这些波动可以表示由装备故障等造成的互联网流量模式的异常或无规律偏差。

[0036] 可以使用跟踪路由实用程序(traceroute utility)来识别和跟踪互联网流量中的波动,以测量全球分布的收集器的网络(或简称为收集器网络)中的每个收集器的延迟。如本领域技术人员所理解的并且在下面更详细解释的,跟踪路由是记录源计算机(例如,收集器)和指定的目的地计算机之间通过互联网的路由的实用程序。跟踪路由数据包括沿着到目的地的路由的路由器和其它设备的地址以及与沿着到目的地的路由的每一跳相关联的延迟这两项。

[0037] 全球分布的收集器网络收集连接到互联网的网络和设备的至少一部分的跟踪路由数据。总的来说,收集器网络可以在典型的一天中监视大约2000000个独特的地址。收集器网络可以针对一些设备(诸如位于高流量节点处的路由器)比针对其它设备更频繁地收集跟踪路由数据。此外,收集器网络中的不同收集器可以收集针对相同目的地、不同目的地、到相同目的地的不同路由和/或几条路由共用的跳的路由跟踪数据。

[0038] 跟踪路由数据可以被设想为包含以大致每天的频率搜集的对于互联网上的一些点的测量结果的数据流,并且对于互联网上的其它(更常遍历的)点的测量结果可能每秒被多次测量。为了进一步解释,每个跟踪路由可以被分解成[收集器,响应器,延迟]元组的集合。引起来自跳W、X、Y和Z的成功跳响应的来自收集器A的踪迹(例如,A→W(10ms)→X(30ms)→Y(45ms)→Z(55ms))变为:

[0039] [A,W,10ms]

[0040] [A,X,30ms]

[0041] [A,Y,45ms]

[0042] [A,Z,55ms]

[0043] 当这些元组从踪迹分解步骤中出现时,它们不再具有包含在它们当中的路径信息。此外,不同的[收集器,响应器]对可以以不同的速率出现在数据流中,因为常被遍历的路由器(如W、X和Y)倾向于被重复地观察/测量,而端点(Z)可能仅偶尔(例如,每天一次)被看到。每个[收集器,响应器]对都具有自己的最近样本队列和基线库。队列和库的尺寸可以基于每个对独有的观察频率来调整。

[0044] 收集器网络根据跟踪路由数据计算互联网上的许多IP地址的延迟数据,并且将延迟数据积累为基线延迟数据。(跟踪路由数据包含在一些情况下可能被这个系统忽略的路径信息。)这种基线延迟数据可以包括由收集器网络中的每个收集器手机的与每个被监视的IP地址相关联的延迟或延迟状态的分布。因为收集器网络持续地收集跟踪路由数据,因此它还检测由于互联网上节点和连接的数量和位置的变化而引起的延迟变化。换句话说,沿着被跟踪的路由的每一跳响应可以被视为跟踪路由收集器和与作出响应的特定跳相关联的节点之间的从头开始的延迟测量结果。结果,由收集器网络收集的最近跟踪路由数据表示互联网的当前状态(或者更精确地说,数据被收集时互联网的状态)。随着收集器网络收集更新后的最近跟踪路由数据,它将先前收集的跟踪路由数据添加到基线延迟数据,因此基线延迟数据跟踪互联网连接随时间的演变。结果,基线延迟状态可以描述互联网相对于地理区域、目标计算设备、一组互连的路由设备和/或一个或多个目标计算设备的正常/常见次序。



[0045] 作为收集器网络的一部分或可操作地耦接到收集器网络的警告系统计算基线延迟数据与更近收集的延迟数据之间的相似性。例如,警告系统可以将互联网上的特定点对之间的分组传输的当前延迟状态(最近测得的离散延迟值的分布)与同一点对之间的基线延迟状态(离散延迟值的基线分布)进行比较。警告系统可以通过计算一个或多个相似性度量来执行这种比较,以确定当前延迟状态是否表示互联网异常。警告系统还可以确定特定于异常的信息,包括但不限于受影响的地理区域、替代的未受影响的流量路由以及异常的分类。

[0046] 如果比较显示当前延迟状态是异常,那么警告系统可以发出指示异常、受影响的路由、替代的未受影响的路由、异常类型等的性能警告。性能警告可以在观察到如从一组收集器(也被认为是一组相关IP/收集器元组)看到的、属于特定位置的特定网络的一群IP地址上的异常之后生成。这有助于防止由于单个网络节点的不当行为或单个收集器处的网络性能问题而引起的假阳性(false-positives)。一旦生成警告,它就可以被发送到已经表示对受影响的网络或地理区域感兴趣的客户,或者发送到分析师以供进一步调查。

[0047] 为了确保当前延迟状态与基线延迟状态之间的适当比较,收集器系统可以使用相同的获取处理(或处理集)来收集并确定延迟状态和底层的离散延迟值。使用一致的延迟收集和确定处理减少或防止了由于测量结果噪声的变化引起的误报出现在比较中。此外,对于当前延迟状态和基线延迟状态的延迟值分布可以被选择为足够大以捕获动态,或足够小以表现由于异常的路由行为引起的延迟变化。

[0048] 用于检测计算机网络流量异常的系统

[0049] 图1例示了检测计算机网络(诸如互联网)中的性能异常的性能警告系统的示例。图1中所示的性能警告系统包括性能警告服务器110(也称为主服务器110),其耦接到性能警告数据库112、一个或多个客户端114以及跟踪路由收集器120的网络。虽然系统100仅示出一个性能警告服务器110和数据库112,但是系统100可以包括和/或使用多个同步的主服务器110和数据库112。当使用多个主服务器110时,主服务器110可以被同步,以用于处理可以分布在多个数据库112上的数据。相应地,数据库112可以被同步,并且因此可以使用有线和/或无线通信协议和/或技术进行通信。

[0050] 跟踪路由收集器120是驻留在它们各自的提供商的数据中心内的真实或虚拟机器,每个属于自治系统(AS) 130或路由域。在操作中,跟踪路由收集器120测量与到它们自己的AS 130内和其它AS 130内的路由器140、目标计算设备150和边界网关协议(BGP)路由器160(也称为边界路由器160)的路由相关联的延迟。

[0051] AS 130可以被认为是计算设备150的邮政编码,即,每个AS 130可以被描绘为基于ISP并且在范围上不一定是地理的互联网邻居。在每个AS 130内,存在实现AS 130的路由策略并维持到邻近的AS 130中的BGP路由器160的物理连接的边界网关协议(BGP)路由器160(也称为边界路由器160)和其它路由器140。在提交时,全球互联网上的AS的数量超过51000。

[0052] 更正式地,AS 130是具有单个明确定义的由公用网络管理员(或管理员组)代表单个管理实体(诸如大学、商业企业、商业部门等)控制的路由策略的IP网络的连接组。AS 130中的给定IP网络内的节点共享相同的网络前缀,从而采用该前缀内的各个IP地址用于互联网连接。大多数自治系统130包括多个网络前缀。AS 130可以使用作为用于在TCP/IP网络中

执行域间路由的外部网关协议(EGP)的BGP来与其它AS 130共享路由信息。

[0053] 可以通过建立从边界路由器160到其BGP对等体之一的连接来在AS 130内或在AS 130之间共享路由信息以便交换BGP更新。如本领域技术人员所理解的,在边界路由器260之间交换数据的过程被称为“对等操作(peering)”。在对等操作会话中,两个网络直接连接并交换数据。内部BGP对等操作会话涉及直接连接单个AS 130内的边界路由器160。外部BGP对等操作会话涉及将邻近的AS 130中的边界路由器160直接相互连接。

[0054] 跟踪路由收集器设备120沿着到具体目标(其可以是连接到互联网的任何设备)的路径在路由器140和边界路由器160之间探测路由并发送跟踪路由数据。跟踪路由收集器设备120可以将收集到的消息数据转发到(一个或多个)主服务器110,主服务器110将数据存储或存档在一个或多个数据库112中,如图1中所示。(一个或多个)主服务器110处理来自(一个或多个)数据库112的跟踪路由数据并向客户端114提供用于访问、操纵和/或以其它方式显示处理后的数据的工具,这些处理后的数据可以用于监视网络路径、检测性能异常、并基于网络异常做出路由决定。该数据允许客户端114监视来自多个有利点(收集器设备120)的网络路径。

[0055] 数据库112还可以包括其它数据,诸如全球路由状态、路由注册数据、域名系统(DNS)数据、互联网网络数据和/或与跟踪路由数据相关或从跟踪路由数据导出的其它数据。这些数据可以用于聚合、关联和聚类所观察到的延迟异常,并支持人类操作员或分析人员尝试理解数据。

[0056] 主服务器和跟踪路由收集器设备

[0057] 图2例示了包括处理器218的主服务器110的示例的框图,其中处理器218耦接到用户接口212、通信接口219和存储可执行指令216的存储器214。这些可执行指令216定义了分析引擎217,分析引擎217当由处理器218实现时分析跟踪路由数据,以识别互联网流量中的性能异常。

[0058] 处理器218可以包括一个或多个高速数据处理单元以执行用于执行用户和/或系统生成的请求的程序组件。通常,这些高速数据处理单元结合各种专用处理单元,诸如但不限于:集成系统(总线)控制器、存储器管理控制单元、浮点单元、以及甚至专用处理子单元,比如图形处理单元、数字信号处理单元等。此外,处理器218可以包括内部快速访问可寻址存储器,并且能够映射和寻址超出处理器本身的存储器;内部存储器可以包括但不限于:快速寄存器、各种级别的高速缓冲存储器(例如,级别1、2、3等)、RAM、ROM等。处理器218可以通过使用可经由指令地址访问的存储器地址空间来访问存储器214和可执行指令216,其中处理器218可以构造和解码所述指令地址以允许其访问通向具有存储器状态和/或可执行指令的特定存储器地址空间的电路路径。

[0059] 照惯例但不一定地,以适配卡的形式,通信接口219可以接受、连接到多个接口适配器和/或与其通信,诸如是但不限于:输入输出(I/O)接口、存储接口、网络接口等。例如,通信接口219中包括的网络接口可以用于发送和接收来自图3中的跟踪路由收集器设备120的信息。

[0060] 用户界面显示器212可以包括具有接受来自视频接口的信号的接口(例如,DVI电路系统和电缆)的基于阴极射线管(CRT)或液晶显示器(LCD)的监视器。可替代地,用户界面显示器212可以包括触摸屏和/或其它内容显示设备。视频接口合成由存储在存储器314中

并由处理器218执行的可执行指令216生成的信息。可执行指令216包括具有用于处理和分析从一个或多个跟踪路由收集器设备220获得的数据的一组指令的分析引擎模块217。用户界面显示器212可以包括如由操作系统和/或操作环境提供的、利用操作系统和/或操作环境提供的和/或在操作系统和/或操作环境之上提供的常规图形用户界面,操作系统和/或操作环境诸如是Apple OS、Windows OS、Linux、基于Unix的OS等。用户界面显示器212可以通过文本和/或图形设施允许程序组件和/或系统设施的显示、执行、交互、操纵和/或操作。用户界面显示器212提供用户可以通过其影响、与之交互和/或操作计算机系统的设施。用户界面显示器212可以与包括它本身的组件集合中的其它组件和/或这类的设施通信。用户界面显示器212可以包含、传送、生成、获得和/或提供程序组件、系统、用户和/或数据通信、请求和/或响应。

[0061] 图3例示了包括硬件和软件模块的跟踪路由收集器设备120的示例的框图。跟踪路由收集器设备120包括具有类似于先前描述的通信接口219的特点的通信接口322。处理器324也可以具有与先前描述的处理器218类似的特点。存储器326包括可执行指令328和踪迹探测器模块329,踪迹探测器模块329具有从一个或多个目标计算设备(例如,图1中的目标计算设备150a、150b和150c)收集跟踪路由数据的指令。

[0062] 跟踪路由数据和数据收集

[0063] 图1和4例示了跟踪路由数据系统的工作原理。为了执行跟踪路由,跟踪路由收集器120a使用互联网控制消息协议(ICMP)将第一分组发送到目的地计算机(150a)。路由跟踪收集器120a还为第一分组指定等于1的跳数限制值(称为“生存时间”(TTL))。当第一路由器140a接收到第一分组时,它递减TTL(从1到0)。在处理TTL=0的分组时,第一路由器向跟踪路由收集器120a返回“超时”消息401a,而不是将第一分组转发到沿着到目的地计算机150a的路径的下一个路由器。这使得跟踪路由收集器120a能够确定与在到目标计算机150a的路径上的到第一路由器140a的跳相关联的延迟。跟踪路由收集器120a然后向目标计算机150a发送具有TTL=2的第二分组。第二路由器160a返回另一个超时消息,等等。后续的分组(包含TTL=3至TTL=7)引起来自路由器160b、160c、140b、160d和160e的超时消息。当目的地计算机150a接收到具有TTL=8的最终分组时,它将“回声应答(Echo Reply)”消息返回给跟踪路由收集器120a,从而使跟踪路由收集器120a能够测量最后一跳的延迟。

[0064] 通过在每次发送分组时增加TTL并且监视来自中间路由器的“超过TTL”响应401a、401b、401c等,跟踪路由收集器设备120a发现到目的地计算机150a的路径上的相继的跳。所收集的“超过TTL”响应和“ICMP回声应答”402的次序和定时被跟踪路由收集器设备120a用来建立由ICMP分组在到目标设备150a的途中遍历的路由器的列表。沿着路径的每一跳都具有通过从响应(超过TTL或回声应答)的到达时间减去ICMP回声请求的发送时间计算的往返延迟。

[0065] 图5A例示了由图1中的跟踪路由收集器120获得的跟踪路由数据的示例。图5A中的跟踪路由数据示出了数据分组从传感器或路由跟踪收集器设备到目标计算设备所取的路径。跟踪路由数据包括跟踪路由中的每个设备的标识符,包括跟踪路由收集器设备120的标识符和/或IP地址501。在502a和502b中包含的IP地址可以表示作为全球或本地计算机网络的一部分的路由器。跟踪路由数据还包括时间503a和503b,时间503a和503b分别表示路由跟踪收集器设备501从路由器502a和502b获得响应所花费的往返时间。类似地,时间503c表

示跟踪路由收集器设备120从由标识符504标识的目标计算设备获得ICMP回声应答所花费的时间。

[0066] 在一些情况下,由一个或多个跟踪路由收集器设备获得的跟踪路由数据可以被主服务器110处理,以按如下所示的数据结构生成中间人类可读格式:

```
[0067] tr_base_fields=[('dcv',str),#data version('ts',int),#timestamp of
start of trace('protocol',str),#[I]CMP,[U]DP,[T]CP('port',int),('collector_
ip',str),('collector_external_ip',str),('collector_name',str),('target_ip',
str),('halt_reason',str),#[S]uccess,[L]oop,[U]nreachable,[G]ap('halt_data',
int),#additional information for failed trace('hoprecords',T5HopList)]
```

[0068] 数据结构tr\_base\_fields可以包括:表示当前内部数据格式(例如,T1或T5)的数据版本dcv字段;指定跟踪路由开始时的时间的ts字段;指定一个或多个互联网协议(例如,互联网控制消息协议(ICMP)、用户数据报协议(UDP)、传输控制协议(TCP)等协议)的protocol字段;指定与跟踪路由数据的传输相关联的端口号或标识符的port字段;指定跟踪路由收集器设备IP地址的collector\_ip字段;指定收集器设备外部IP地址的collector\_external\_ip字段;指定收集器设备名称的collector\_name字段;指定目标设备IP地址的target\_ip字段;指定暂停原因(例如,成功、循环、不可达或间隙)的halt\_reason字段;以及表示在跟踪路由数据传输期间记录的跳的hoprecords列表。

[0069] hoprecords列表可以具有一个或多个元素,其中每个元素表示在从跟踪路由收集器到目标设备的跟踪路由中的类型路由设备或者目标设备。路由设备可以是调节IP网络之间的数据流量的网关路由器。此外,路由设备可以在具有规定分组如何能够到达目标设备的静态或动态规则集合的IP网络内将数据分组从一个接口指引到另一个接口。下面提供了表示hoprecords列表元素的数据结构:

```
[0070] ex=[h.ip,#hop IP address h.rtt,#RTT in milliseconds h.dist,#TTL
distance on the packet that incurred the reply h.replyttl]#TTL of the reply
packet from the target mpls-tags#numeric MPLS tags are appended as comma-
separated values
```

[0071] 数据结构ex可以包括:表示路由设备或目标设备的IP地址h.ip;表示响应于由跟踪路由收集设备(例如,图4中的120c)发送的跟踪路由传输而从接收方设备(例如,图4中的401a-401c和402)接收到确认传输所花费的时间的往返延迟时间h.rtt;指定跟踪路由传输分组在被认为是不成功传输之前可以耗费的跳数限制的生存时间(TTL)距离h.dist;指定对跟踪路由传输的确认在被认为是不成功传输之前可以耗费的跳数的TTL距离h.replyttl;以及指定例如标签交换路径以预先确定在私有和/或公共网络内采取的路由的一个或多个多协议标签交换(MPLS)标签mpls\_tags。

[0072] 下面提供了所收集的如在tr\_base\_fields和ex数据结构所示地组织的跟踪路由数据的示例。

```
[0073] dcv=T5;ts=1419460683;protocol=ICMP;port=0;collector_ip=
162.243.114.207
```

```
[0074] collector_extcrnal_ip=162.243.114.207;
```

```
[0075] collector_name=vps01.nyc3;target_ip=188.241.40.144;
```

[0076] halt\_reasone=Success;halt\_data=11;

[0077] hop\_records={h.ip=192.241.164.241;h.rtt=0.445ms;h.replyttl=63  
h.ip=62.115.44.241;h.rtt=0.468ms h.replyttl=62}

[0078] 图5B-5D例示了可以应用于包括延迟值作为5A中所示的数据的跟踪路由数据的分布的数据平滑或规范化处理的示例。在一些实现中,主服务器110可以准备由一个或多个收集器接收的、要进行处理的数据。数据的准备可以包括拒绝离群值和/或跟踪路由数据内被认为是噪声的其它值。例如,高于和/或低于某个百分点(例如,第10、15、20或25个百分点)的延迟数据可以从分布中省略。此外,主服务器110可以从根据跟踪路由数据获得的离散延迟值的分布(图5C)导出核密度图(图5B)。对于由主服务器接收的每个分布,可以通过例如应用具有 $\sigma^2$ 的核宽度的高斯和(sum-of-Gaussians)技术来估计核密度图。此后,主服务器110可以例如通过根据群体尺寸测量一个或多个核分布估计的高度或者通过从核分布估计重新采样新的分布来在具有相等群体尺寸的共享基础上计算每个分布。核的带宽是可以根据具体情况进行配置的自由参数。类似地,子区间、仓或桶是覆盖一定范围的数据的自由参数,其也可以被相应地配置。图5D例示了与图背景上所示的直方图上的离散数据对应的重叠的核密度估计。

[0079] 积累和更新基线延迟状态

[0080] 图6A-6C例示了用于基于新的跟踪路由数据、最近的跟踪路由数据和基线跟踪路由数据来获取、积累和更新延迟状态的跟踪路由数据替换策略。主服务器110接收包括由一个或多个跟踪路由收集器设备120收集的跟踪路由记录的数据流。每个跟踪路由记录表示在给定跟踪路由收集器设备120与计算机网络中的另一个节点之间的一个或多个跳。如上所示,每个跳记录包括跳的往返时间(延迟)。每个跳记录还可以包括时间戳、跳距离和应答TTL,如hop\_record数据结构中所示。

[0081] 主服务器101基于tr\_base\_fields数据结构中的时间戳ts来对数据流中的跟踪路由记录进行排序。它还将时间戳排序的跟踪路由记录的跳记录分解成一组时序的数据子流,其中每个子流604包括单个收集器与单个接收方IP地址之间的样本。基于在tr\_base\_fields数据结构中描述的收集器和在测量结果的每个hop\_record的h.ip字段中的接收方IP地址,将这些样本指引到子流604。

[0082] 如图6A中所示,样本可以暂时保留在最近样本队列(RSQ)601中的存储器114中。在主服务器110的一些情况中,接收到的样本可以被推送到RSQ 601的尾部。RSQ 601可以包括预配置的保留容量。图6B示出了每当RSQ 601的保留容量被耗尽时,主服务器110就可以将最老的样本从RSQ 601的头部移到基线库(BR)602中。

[0083] 如果期望的话,可以调整RSQ 601的保留容量:一般而言,较小的RSQ保留容量允许系统更快速地对异常做出响应,但倾向于使系统对假阳性更敏感。在一些情况下,每个[收集器,响应器]对具有基于收集器从响应器接收样本的观察频率的RSQ保留容量。如果收集器频繁地从响应器接收样本(例如,因为响应器是互联网阻塞点处的路由器),那么系统可以将RSQ保留容量设置为保持接收到的样本相对短的时间。它也可以只保持接收到的样本的一小部分(例如,每第十个样本)。并且它可以保持更大数量的样本。但是,如果收集器相对不频繁地从响应器接收样本(例如,因为响应器是互联网终端处的客户端设备),那么系统可以将RSQ保留容量设置为保持接收到的样本相对长的时间和/或在RSQ601中共保持较

少的样本。例如,保持少于10个样本的RSQ601可能在仅有少数偏差样本的情况下生成大量的假阳性,而保持>200个样本的RSQ601可能只对于最常观察到的[收集器,响应器]对反应迟缓。

[0084] 类似于RSQ 601的保留容量,可以预先配置BR 602的保留容量,以保留预定数量的样本。可以基于收集器从响应器接收样本的观察频率为每个[收集器,响应器]对设置BR保留容量,就像RSQ保留容量可以被调整那样。典型地,BR 602应当保持对应的RSQ 601的大约5-10倍的样本。

[0085] 如图6C中所示,一旦BR 602达到其保留容量,就可以通过从BR 602中移除对应数量的样本603并将其替换为根据一个或多个替换策略(例如,转盘替换策略等)从RSQ 601取得的样本来将一个或多个样本添加到BR 602。

[0086] 在一些情况下,替换策略(也被称为驱逐策略)可以指定从BR602中随机选择样本进行永久驱逐。在这种情况下,每个样本可以具有相同的驱逐概率。而且,每个样本可以具有相当于由函数 $[1 - (1/s)]^m$ 定义的、保留在BR 602中的概率,其中s是保留在BR602中的样本的数量并且m是放在BR 602中的附加样本的数量。

[0087] 主服务器110可以基于存储在BR 602中的离散延迟值来确定与两个或更多个设备之间的信息传输相关联的传输基线延迟状态。可以通过在保留在BR 602中的样本上计算一个或多个描述性统计量(例如,离散延迟值的分布的离散度的度量和/或集中趋势的度量)来确定基线延迟状态。类似地,主服务器110可以基于存储在RSQ 601中的离散延迟值来确定与两个或更多个设备之间的信息传输相关联的当前延迟状态。当前延迟状态可以类似地通过在保留在BR 602中的样本上计算一个或多个描述性统计量(包括但不限于集中趋势的度量和/或离散度的度量)来确定。

[0088] 在一些情况下,包含在BR 602和RSQ 601中的分布可以经受由主服务器110执行的一个或多个采样策略和/或规范化策略。例如,主服务器110可以应用如下策略:通过指定在从(一个或多个)延迟值分布中得出推断时低于或高于阈值的样本不应当纳入考虑,该策略省略离群值样本。此外,在执行任何统计分析和/或测试之前,BR 602和RSQ 601中的分布可以通过任何合适的操作被平滑和/或规范化。合适的平滑操作包括但不限于针对每个分布确定核密度、在具有相等群体尺寸的共享基础上重新计算每个分布等。

[0089] 主服务器110可以通过分析基线延迟状态与当前延迟状态之间的差异来识别一个或多个信息传输异常。(下面描述的图11A-11D显示了来自基线库和当前延迟状态这两者的样本)。这种分析可以包括离散比较,以确定基本延迟状态与当前延迟状态从相同的群体得出的可能性,例如对数似然分析、Pearson的卡方(Chi-squared)测试等离散比较分析。在其它情况下,分析可以包括基于例如Kolmogorov-Smirnov分析的连续比较以及对两个分布之间的不止一个差异敏感的类似统计分析。

[0090] 图7例示了图1中所示的设备的子集,包括目标计算设备150a和150y;跟踪路由收集器120a-120c;AS边界路由设备160a-160f;以及内部路由器140a-140c。它示出了主服务器110如何可以基于在两个或更多个设备之间确定的异常来确定替代的信息传输路由。例如,计算机设备150y具有向目标设备150a发送信息的两条路径。第一路径包括边界路由器160c和内部路由器140b。第二路径包括内部路由器140c和边界路由器160f。两条路径都包括边界路由器160a、160b和160d。收集器设备/网络传感器120连接到主服务器110并且可以

以预先配置的间隔向主服务器110发送跟踪路由数据。

[0091] 主服务器110可以基于由收集器设备120a提供的跟踪路由数据来推断不同的猜测。例如,针对关于跟踪路由数据通过第一路径140a-160a-160b-160c-140b-160d-160e到目标设备150a的传输而从收集器设备120a接收的数据,可以如上面所提到的那样确定识别异常的猜测。并行地,主服务器110可以针对关于跟踪路由数据通过第二路径140a-160a-160b-140c-160f-160d-160e到目标设备150a的传输而从收集器设备120b接收的数据确定不存在异常。在这种情况下,主服务器110可以向设备150x警告关于异常以及关于经由不异常执行的第二路径(即,140a-160a-160b-140c-160f-160d-160e)向目标设备150a发送信息的可能性。在一些情况下,在考虑附加的传输信息(例如,当第一路径非不成功时比较第一路径与第二路径的延迟、确定到第一路径的传输在路径内的某个点是否被认为不可达以及类似的条件)之后推荐替代路径。

[0092] 在一些情况下,主服务器110可以基于跟踪路由数据将异常缩小到具体的链路和/或地理区域。例如,假定在第一路径中识别出异常,那么,如果通过第三路径160c-160b-160a-140a-150y在收集器设备120c与目标设备150y之间没有检测到异常,那么主服务器110可以推断异常存在于部分路径160c-140b-160d中的某处,从而从受影响路由中排除连接160a-160b和160d-160e的链路。通过分析经部分重叠路径从其它收集器设备接收到的跟踪路由数据,这种猜测可以被重新确认、忽略和/或进一步缩小到具体链路或地理区域。如果从设备120得出的推断是矛盾的或不和谐的,那么主服务器110可以通过向由跟踪路由数据收集器设备120接收的数据指派相干性值来协调推断。在一些情况下,相干性值可以是与每个收集器对应的可靠性权重和/或其它可靠性度量的函数。

[0093] 检测性能异常

[0094] 图8绘出了用于检测计算机网络(诸如互联网)中的性能异常的方法800。在步骤801中,处理器(例如,主服务器110)接收表示目标设备(诸如计算机网络上的计算机或路由器)与计算机网络中的或耦接到计算机网络的一个或多个收集器(例如,收集器120a)之间的传输延迟的延迟数据,诸如跟踪路由数据或ping数据。延迟数据可以包括来自一个或多个跟踪路由数据分组的测量结果,其中每个跟踪路由数据分组可以被组织成包含上面关于tr\_base\_fields数据结构描述的字段。例如,这些字段可以指示与到计算机网络内的不同目的地(节点)的跳或跳数相关联的延迟。如上所述,每一跳都可能在途中遇到另一个目标。

[0095] 延迟数据还可以包括由一个或多个ping分组(或ping)引出的延迟测量结果,其中ping可以直接以给定的IP地址为目标。每个ping分组在单个收集器与计算机网络内的具体目标目的地之间被收集。每个ping分组测量收集器与目的地之间在一个时刻的单个往返时间(延迟)。

[0096] 在步骤802中,延迟数据被组织或布置成时序的流,其中每个流仅表示在单个收集器与单个目标设备之间观察到的样本。在路由跟踪数据的情况下,这些流是通过路由跟踪测量结果进行排序和分解而获得的。跟踪路由测量结果的每个跳记录可以被理解为发起跟踪路由的收集器与该跳IP地址之间的从头开始的往返时间(延迟)测量结果。在ping数据的情况下,这些流是简单地通过对收集器与目标之间的不同ping测量结果进行排序和整理而获得的。

[0097] 在一些情况下,处理器可以基于目标设备在计算机网络内的位置以及去往、来自



或经过目标设备的流量的量来调节延迟收集参数(804)。例如,在将延迟数据存储在缓冲器(图6中的最近样本队列601)之前,处理器可以减小收集速率(804a)或抽取延迟数据(804c)。例如,处理器可以抽取与往返于频繁运送的路由器的往返行程相关联的延迟数据,以便过滤延迟中的高频波动并确保缓冲器保持足够的延迟数据,以获得对于当前和基线延迟状态的有意义的估计。处理器还可以调节用于保持延迟数据的缓冲区(最近样本队列601和基线库602)的尺寸。对于典型的目标,最近样本队列保持25个样本,并且基线库保持175个样本。取决于由(一个或多个)收集器测量至目标的延迟的频率,这些数字可以增加或减小。(要注意的是,步骤804可以在图8中所绘出的处理流程内的不同点执行)。

[0098] 处理器将输入的延迟数据存储在最近样本队列中(806)。处理器还将较老的延迟数据从最近样本队列移动到基线库(808)。例如,处理器可以在每个测量间隔处用来自最近样本队列的最老延迟样本替换基线库中的最老延迟样本。可替代地,处理器可以每个测量间隔一次地用来自最近样本队列的延迟样本替换基线库中的随机延迟样本。

[0099] 处理器根据基线库中的延迟样本估计基线延迟状态(810),并根据最近样本队列中的延迟样本估计当前延迟状态(812)。它比较当前延迟状态与基线延迟状态(814),并且如果该比较指示异常网络行为,那么识别出异常的存在(816)。处理器还可以尝试识别异常的类型(818),例如路径改变、拥塞、连接性故障等及其在网络和/或特定地理区域内的位置。例如,处理器可以识别与异常相关联和/或受该异常影响的一个或多个IP地址、地理区域或者IP地址和地理区域这两者。

[0100] 处理器可以使用任何(一种或多种)合适的比较技术来估计并比较当前延迟状态与基线延迟状态。例如,处理器可以生成并规范化当前延迟分布和基线延迟分布,然后比较经规范化的分布,以确定当前延迟性能相对于历史延迟性能是异常的概率。处理器还可以计算一个或多个描述性统计量805,包括例如离散延迟值的分布的集中趋势的度量和/或离散度的度量。此后,主服务器110可以计算指示当前延迟值与预期的延迟值之间的偏差对应于网络性能异常807的概率的值。如果偏差在统计上显著,那么它可以是异常的征兆。在实践中, $p \leq 0.00001$ 就足以认为单个收集器/响应器对是异常的。这些收集器/响应器对可以被聚类,以找到遭遇中断的网络或地理区域。

[0101] 处理器还可以基于类型、受影响区域的(一个或多个)IP地址、受影响的地理位置以及异常的定时、先前确定的响应策略、来自用户的指令等来对异常820做出响应。例如,处理器可以利用关于异常的类型、位置和定时的信息向一个或多个用户或其它设备自动发送异常警告(820a)。这些设备包括但不限于DNS服务器、路由器、负载平衡器、网络优化设备及配置管理系统。

[0102] 例如,处理器可以将异常警告发送到DNS服务器,DNS服务器部分地根据计算机网络中是否存在任何异常而解析DNS查询。如本领域技术人员容易理解的,DNS服务器是将域名翻译成可用于通过计算机网络路由分组的IP地址的服务器。DNS服务器有各种类型,包括权威DNS服务器、高速缓存DNS服务器及转发DNS服务器。在无异常的情况下,DNS服务器可以根据第一预定策略来解析DNS查询,该第一预定策略可以被设置为实现期望的延迟、负载平衡、降低的成本、避免或优选经由特定互联网服务提供商(ISP)路由流量,等等。DNS服务器可以切换到用于响应于来自处理器的异常警告而应答DNS查询的另一个预定策略。在一些情况下,DNS服务器可以针对异常的不同的类型、位置和严重性具有不同的策略。每个策略



可以预定或设置,或者响应于异常而修改。

[0103] 换言之,如果异常影响DNS服务器根据第一预定策略解析针对特定域名的DNS查询的能力,那么DNS服务器可以切换到以不同方式处理针对那个特定域名的查询的第二预定策略。例如,如果DNS服务器在第一预定策略下响应于给定查询而将流量转向到第一IP地址并且异常影响去往和来自该IP地址的互联网流量,那么DNS服务器可以在第二预定策略下响应于相同的查询而将流量转向到第二IP地址。同样,DNS服务器可以基于目的地的地理位置、DNS查询源和受异常影响的区域以及基于与在对应的IP地址之间路由的流量相关联的延迟来使流量转向。

[0104] 处理器还可以自动将异常警告发送到强制执行考虑异常的路由策略的可配置网络路由器(例如,在客户站点处)。如果路由器具有多条到互联网的可用路径(例如,经由不同的ISP),那么它可以基于预定策略(包括负载平衡、最低成本或最佳性能)经由对应接口将出站流量指引到路径(ISP)。策略还可以解释一个或多个异常的存在以及不同类型的异常。结果,如果路由器接收到指示经由给定ISP的流量可能经历更高延迟或其它问题的异常警告,那么路由器可以经由替代ISP发送流量,直到异常得到解决。如果路由器与收集器120中的一个或多个(图1)并置,那么,通过经由每个ISP接口对相同的目标进行采样、识别异常并对可配置网络路由器上的策略引擎提出程序化建议,收集器可以通过每个ISP接口从路由器的角度查看互联网。

[0105] 图9绘出了用于检测计算机网络中的异常的另一种方法900。在步骤902中,主服务器110或其它处理器从多个收集器收集延迟数据。这种延迟数据可以包括从每个收集器到分布在整个计算机网络中的一个或多个其它设备的跟踪路由分组,如上面更详细描述。

[0106] 在步骤904中,主服务器110基于可用数据确定和收集器与计算机网络中的每个目标之间的传输相关联的延迟状态。延迟状态可以包括或表示延迟分布,并且可以使用上面关于图8中的步骤810和812描述的技术来估计。例如,主服务器110可以计算由当前延迟状态表示的延迟值的运行均值、中值或模式,或者将峰值拟合到由当前延迟状态表示的延迟值的分布。当前延迟状态可以描述每个收集器/目标对之间最近的、接近实时的传输状态。主服务器110可以成对地执行这些计算,例如,在每个收集器与第一目标、第二目标之间等等,直到已经计算出每个可能的组合为止。(在一些情况下,对于没有数据或没有连接的收集器和目标来说,估计延迟状态可以是不现实或不可能的)。

[0107] 主服务器110还为每个收集器/目标对确定预期的延迟状态。在一些情况下,主服务器110可以根据历史延迟数据来计算预期的延迟状态,如上面关于图8中的步骤810所描述的。在其它情况下,主服务器110可以根据用户指定的数据确定预期的延迟状态-即,用户可以设置期望的延迟或延迟状态,并且主服务器110可以使用这个用户提供的信息来确定预期的延迟状态。主服务器110还可以基于由用户提供的信息来生成预测的或期望的延迟状态。

[0108] 在步骤908中,主服务器110比较预期的延迟状态和当前延迟状态。作为比较的一部分,主服务器110可以拒绝高于第一阈值或低于第二阈值的离散延迟值;构建预期延迟状态的核密度估计;和/或构建当前延迟状态的核密度估计。该比较可以提供足够的信息来推断影响与一个或多个目标设备的通信的网络传输异常的存在(步骤910)。在这种情况下,主服务器110可以例如通过IP地址、地理位置或其它标识符来识别受异常影响的目标设备(步

骤912)以及计算机网络的未受影响的部分(步骤914)。主服务器生成并发送类似于上面关于图8描述的警告消息(步骤916)。接收到该警告消息的DNS服务器、路由器和用户可以以减轻异常的有害影响的方式使用关于异常的信息来解析DNS查询、选择ISP等,如上面所述的。

[0109] 图10绘出了用于检测计算机网络(例如,图1中所示的网络100)中的性能异常的又一种方法1000。在步骤1002中,收集器接收跟踪路由数据分组并且将跟踪路由数据分组中编码的延迟信息提供给主服务器110。在步骤1004和1006中,主服务器110基于收集到的延迟信息和预期的延迟状态来确定实际延迟状态,例如使用如上所述的历史延迟信息和/或用户输入。

[0110] 在一些实现中,主服务器100例如通过生成分组将按给定延迟经过特定路径的概率来比较实际延迟状态与预期的延迟状态。它使用这个概率和其它因素(诸如与收集器设备相关联的可靠性值和延迟数据的质量)来确定性能网络异常猜测(1004a)。猜测可以被解释为许多或至少两个前因之一,这些前因一起足以生成关于一类网络性能异常的完整推断。主服务器110计算相干性值(1004b),以协调从自其它收集器接收到的跟踪路由数据导出的猜测之间的差异。如果相干性值高于预定的异常阈值,那么主服务器110可以生成指示网络性能异常的推断1010。主服务器110可以通过使用一种或多种人工智能技术分析跟踪路由数据来对异常做出响应(步骤1012),以诊断或分类异常类型(1012a)。如上所述,它还可以发送异常警告(1012b)、基于异常来路由分组(1012c),和/或基于异常来解析DNS查询(1012d)。

[0111] 性能异常警告

[0112] 在一些情况下,主服务器可以检测不同类型的异常,并通过主服务器的用户界面显示器212提供图形和/或文本输出。图11A-11D例示了由警告系统检测到的一组现实世界异常,不同的符号表示不同收集器的测量结果。图11A-11C看起来是异常的路径变化,而图11D看起来是异常的拥塞事件。

[0113] 更具体而言,图11A是从一组三个不同的(但地理上接近的)跟踪路由收集器对于由位于伊拉克的ScopeSky通信和互联网有限公司(AS50597)控制的一群IP地址测得的往返延迟的图。该图示出了典型的双峰分布(在大约80ms和130ms处具有峰)与降级的、较慢的单峰分布(具有在160-170ms周围略微变化的峰)之间的过渡。

[0114] 图11B是从一组三个相似的收集器(以不同颜色显示)对于由位于中国的Pacnet控制的特定网络前缀(222.126.128.0/23)中的IP地址测得的往返延迟的图。该图示出了三个分开的模式,每个模式间隔大约80ms(在310ms、390ms和470ms处)。随着上层模式的下降,性能在6月13日得到改进,并且仅与最快的结果间隔10-20ms。

[0115] 图11C是从一组三个相似的收集器(以不同颜色显示)对于由位于南非的TENET(AS2018)控制的地址测得的往返延迟的图。该图显示延迟(大部分)逐渐改进,但是存在保持独立的两种模式(间隔大约20ms),因为它们变得快40ms。

[0116] 图11D是从一组三个相似的收集器(以不同颜色显示)对于由位于罗马尼亚的布加勒斯特的RCS&RDS控制的特定网络前缀(213.154.100.0/23)中的IP地址测得的往返延迟的图。该图显示了来自全部三个收集器的大约35ms的一致延迟。从6月12日开始,被测量的地址的子群体开始表现出日间延迟上升。到6月18日,该子群体表现出从基线起多达300ms的日间延迟变化。

[0117] 主服务器110可以通过一种或多种人工智能技术被训练成识别随时间的跟踪路由数据模式,从而允许根据跟踪路由数据样本的观察来对信息传输系统和/或网络100自动建模。人工智能技术可以包括但不限于时间自动机、实时自动机等人工智能技术。一旦系统100被建模,主服务器110就可以操作为跟踪路由数据的分类器来识别异常的类型。可由主服务器110识别的一些异常可以包括模式改变1102、模式整合1104、多模式转移1106、变差的日间拥塞1108等异常。而且,主服务器还可以确定受异常影响的国家1110、受影响国家的区域互联网注册机构1112、受异常影响的DNS标识符1111和/或受异常影响的ISP 1109。

[0118] 性能异常警告

[0119] 如果主服务器110例如通过识别延迟数据中的模式或者当前延迟状态与基线延迟状态的比较来识别性能异常,那么它可以生成并发出如图12中所示的警告消息1201的警告。在一些情况下,主服务器110可以向一个或多个客户端114警告关于识别出的异常。异常警告可以指示受警告影响的地理区域。例如,警告消息可以指示受影响的地理区域是西北华盛顿特区。除了地理区域之外,警告消息还可以示出受影响的路径1205。在本示例中受影响的路径包括图7中所示的设备(140a-160a-160b-160c-160b-160d-160e)。

[0120] 另外,警告消息可以指示不受影响的地理区域1003,例如西南华盛顿特区。在一些情况下,警告消息可以包括建议的替代路径。例如,替代路径可以包括图7中所示的设备140a-160a-160b-140c-160f-160d-160e。可以在警告消息中提供与替代路径对应的估计延迟传输时间1007,其指示通过建议的替代路径发送数据分组的估计时间。

[0121] 结论

[0122] 虽然已经在本文中描述和例示了各个发明性实施例,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同。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同的范围内,发明性实施例可以如不同于具体描述并要求保护的那样来实践。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/或方法不是相互不一致的,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。

[0123] 本发明的上述实施例可以以多种方式中的任何一种来实现。例如,一些实施例可以使用硬件、软件或其组合来实现。当实施例的任何方面至少部分地用软件实现时,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。

[0124] 在这方面,本发明的各个方面可以至少部分地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、或者其它有形计算机存储介质或非瞬态介质),所述一个或多个程序当在一个或多个计算

机或其它处理器上执行时执行实现以上讨论的技术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,以实现如上讨论的技术的各个方面。

[0125] 本文使用的术语“程序”或“软件”在一般性意义上指任何类型的计算机代码或计算机可执行指令集,该计算机代码或计算机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技术的各个方面。此外,应当认识到的是,根据本实施例的一个方面,当被执行时执行本技术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技术的各个方面。

[0126] 计算机可执行指令可以以许多形式被一个或多个计算机或其它设备执行,诸如程序模块。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。典型地,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。

[0127] 此外,本文描述的技术可以体现为已经提供了至少一个示例的方法。作为方法的一部分执行的动作可以以任何合适的方式排序。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,尽管在说明性实施例中示出为顺序的动作。

[0128] 如本文所定义和使用的,所有的定义都应该被理解为在字典定义上控制通过引用并入的文档中的定义和/或所定义术语的普通含义。

[0129] 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”。

[0130] 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“任一者或两者”,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”子句具体标识的元素之外,其它元素也可以可选地存在,无论是否与具体标识的那些元素相关。因此,作为非限制性的示例,当结合诸如“包括”之类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。

[0131] 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的至少一个,但也包括一个以上,以及可选地,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”,或者当在权利要求中使用时的“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,“基本由…组成”将具有其如在专利法领域中使用的普通含义。

[0132] 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个或每一个元素中的至少一个并且不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。

[0133] 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”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。

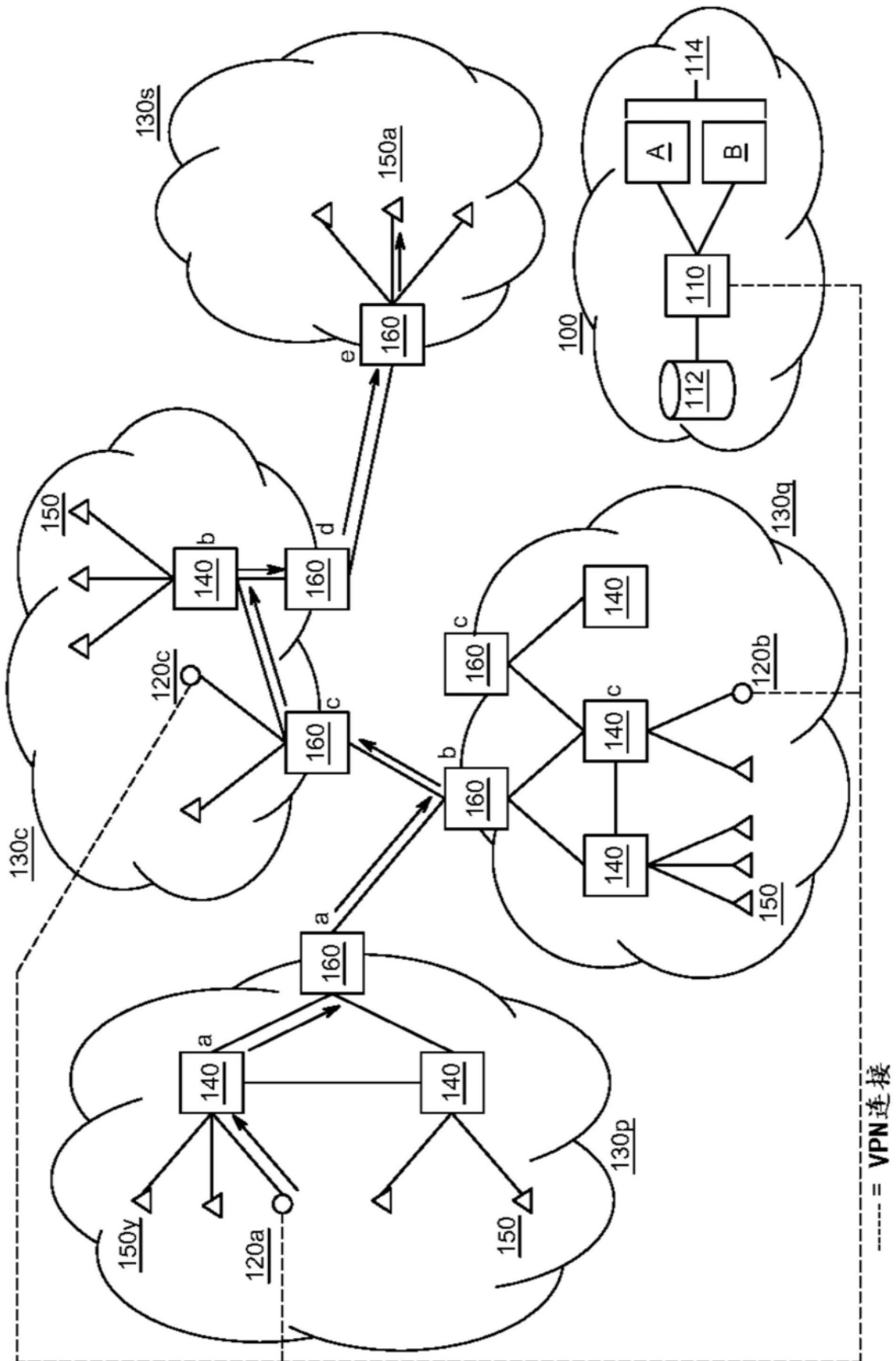


图1

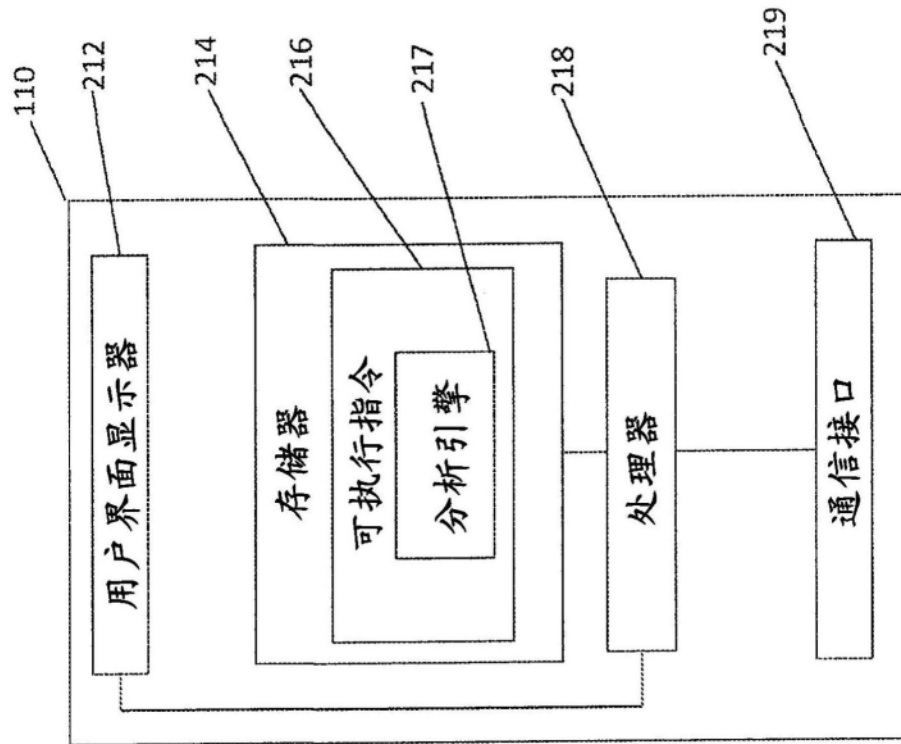


图2

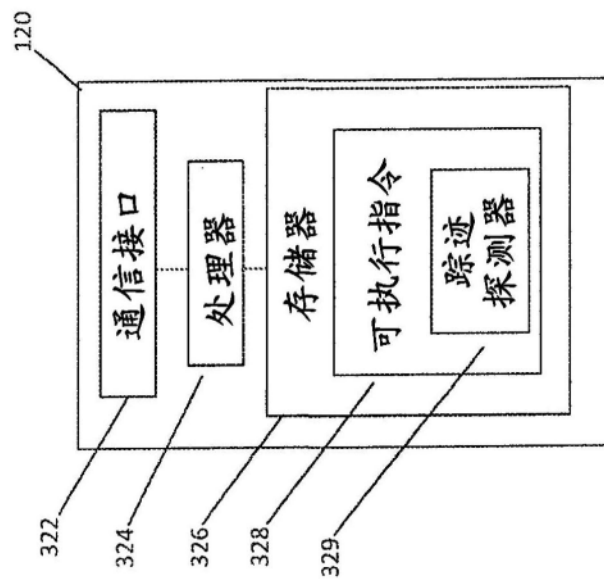


图3

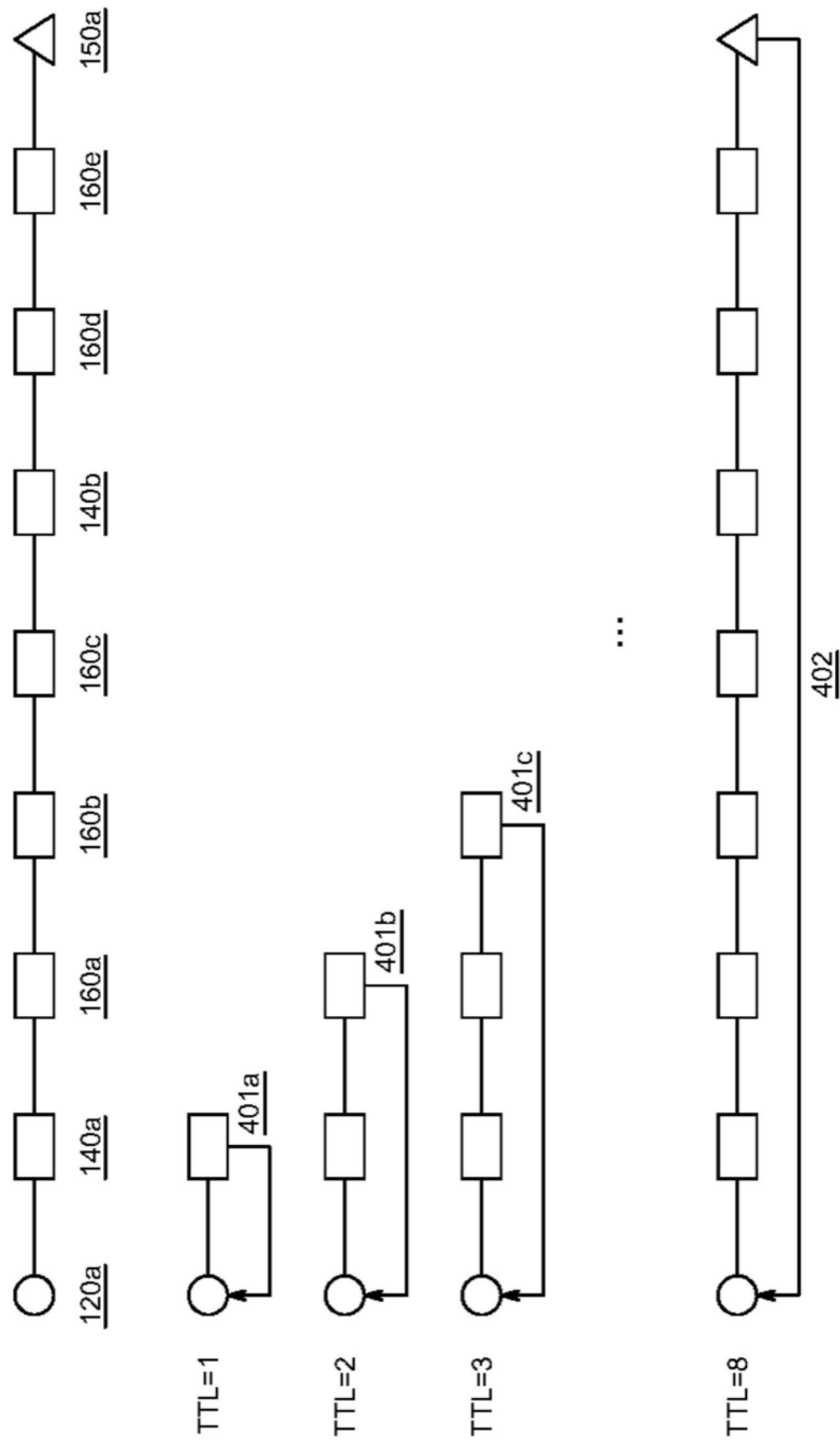


图4



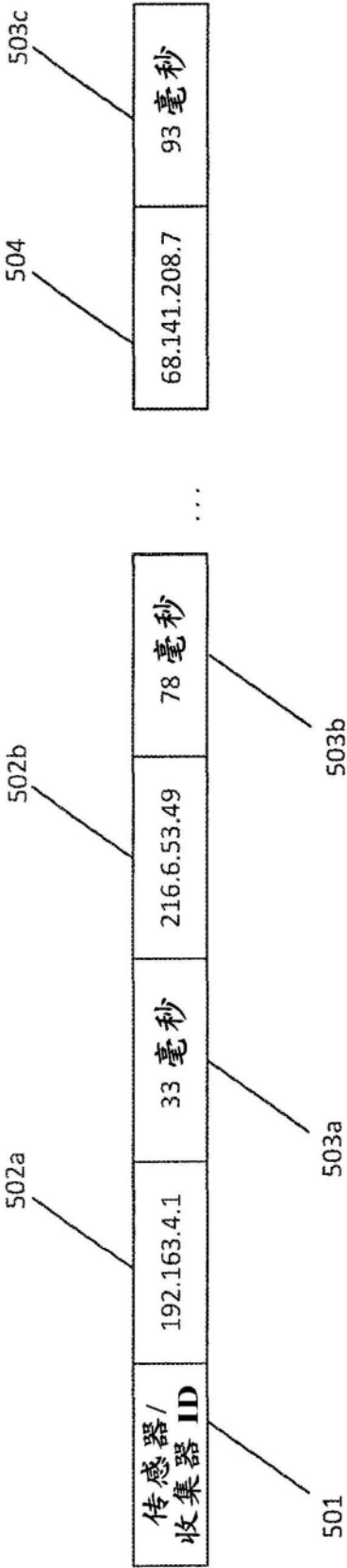


图5A

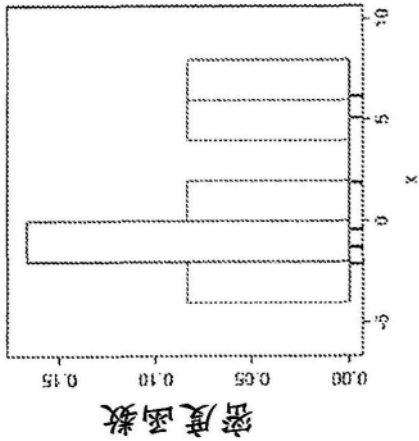


图5B

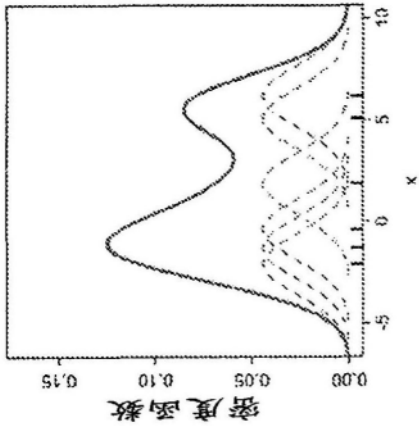


图5C

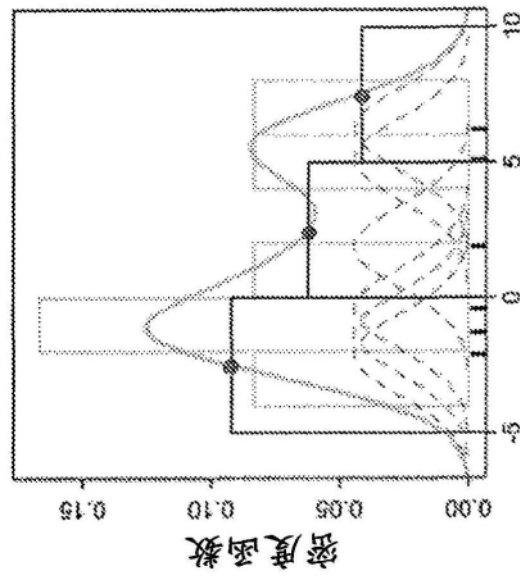


图5D

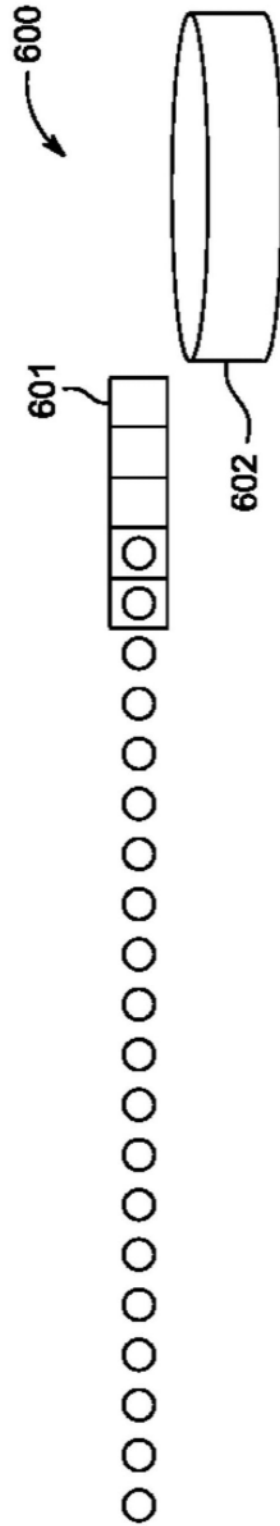


图6A

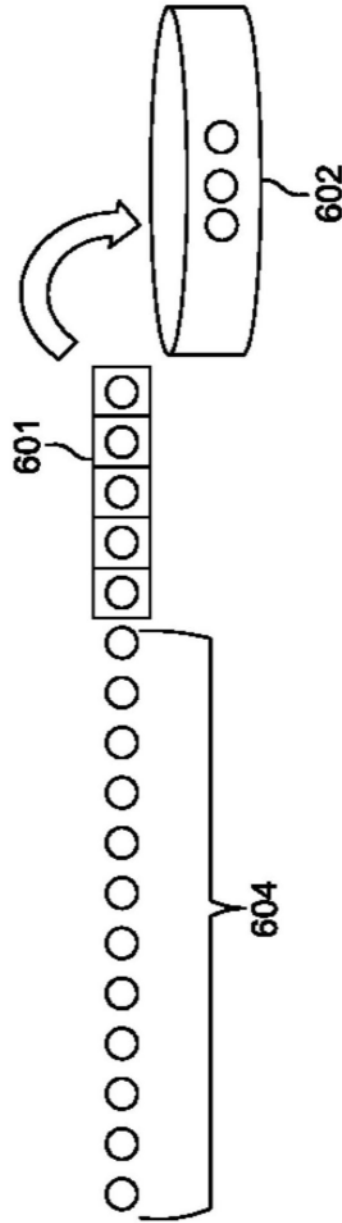


图6B

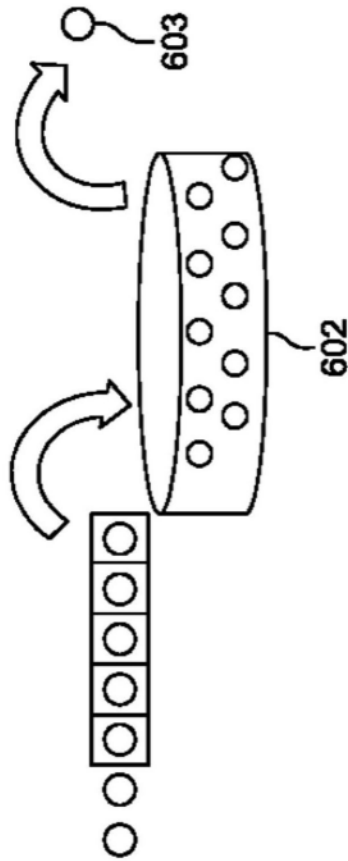


图6C

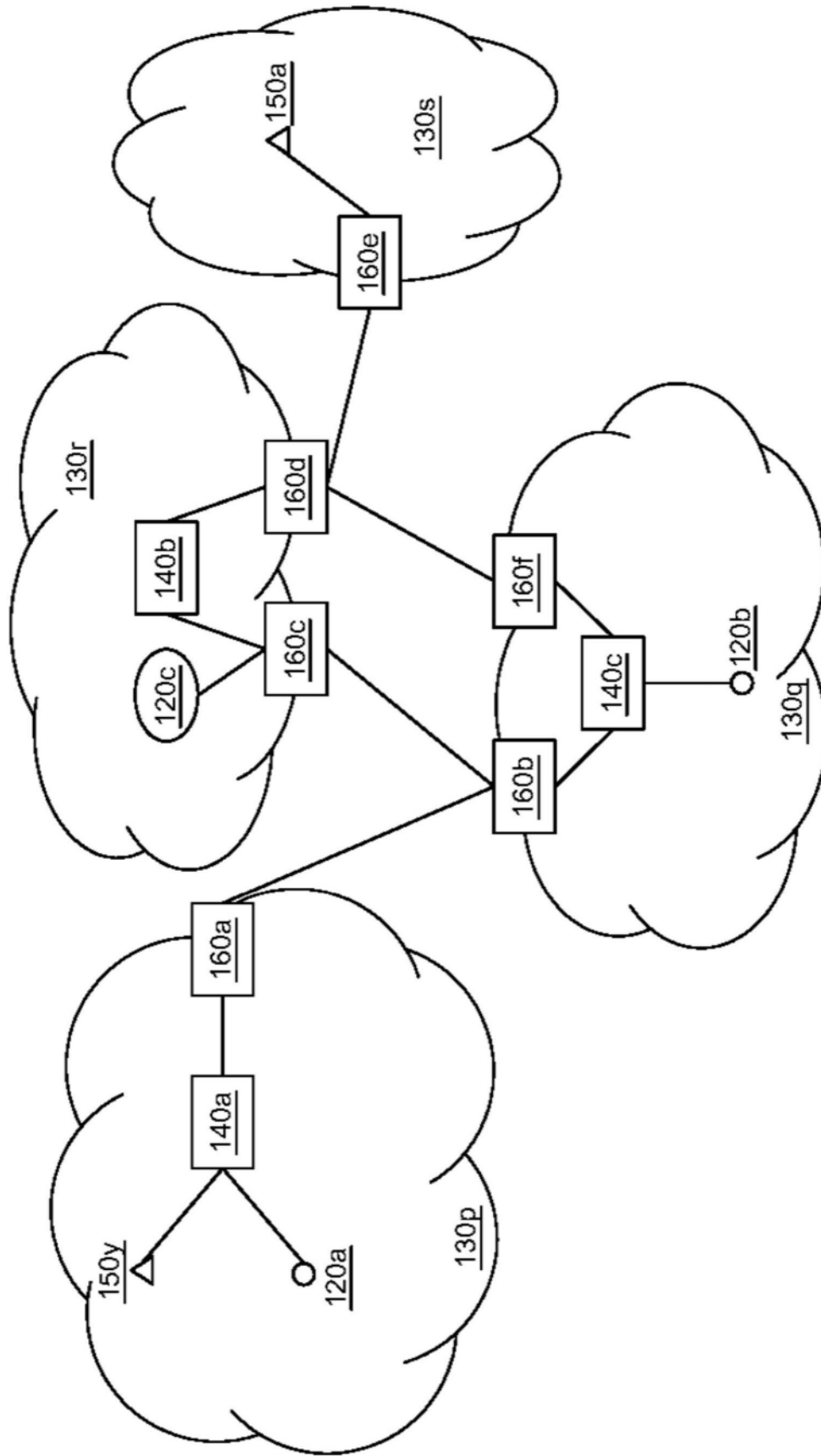


图7

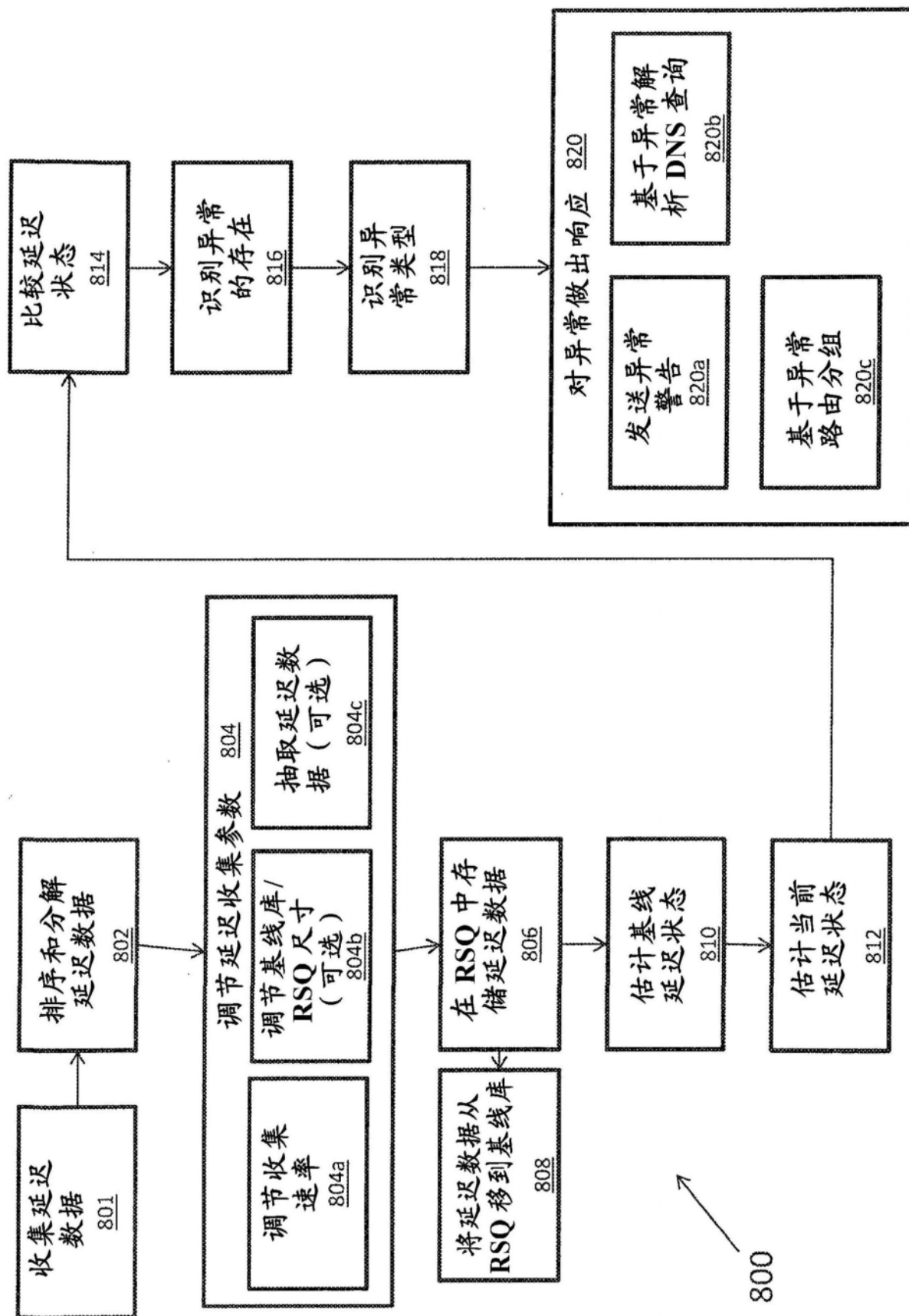


图8



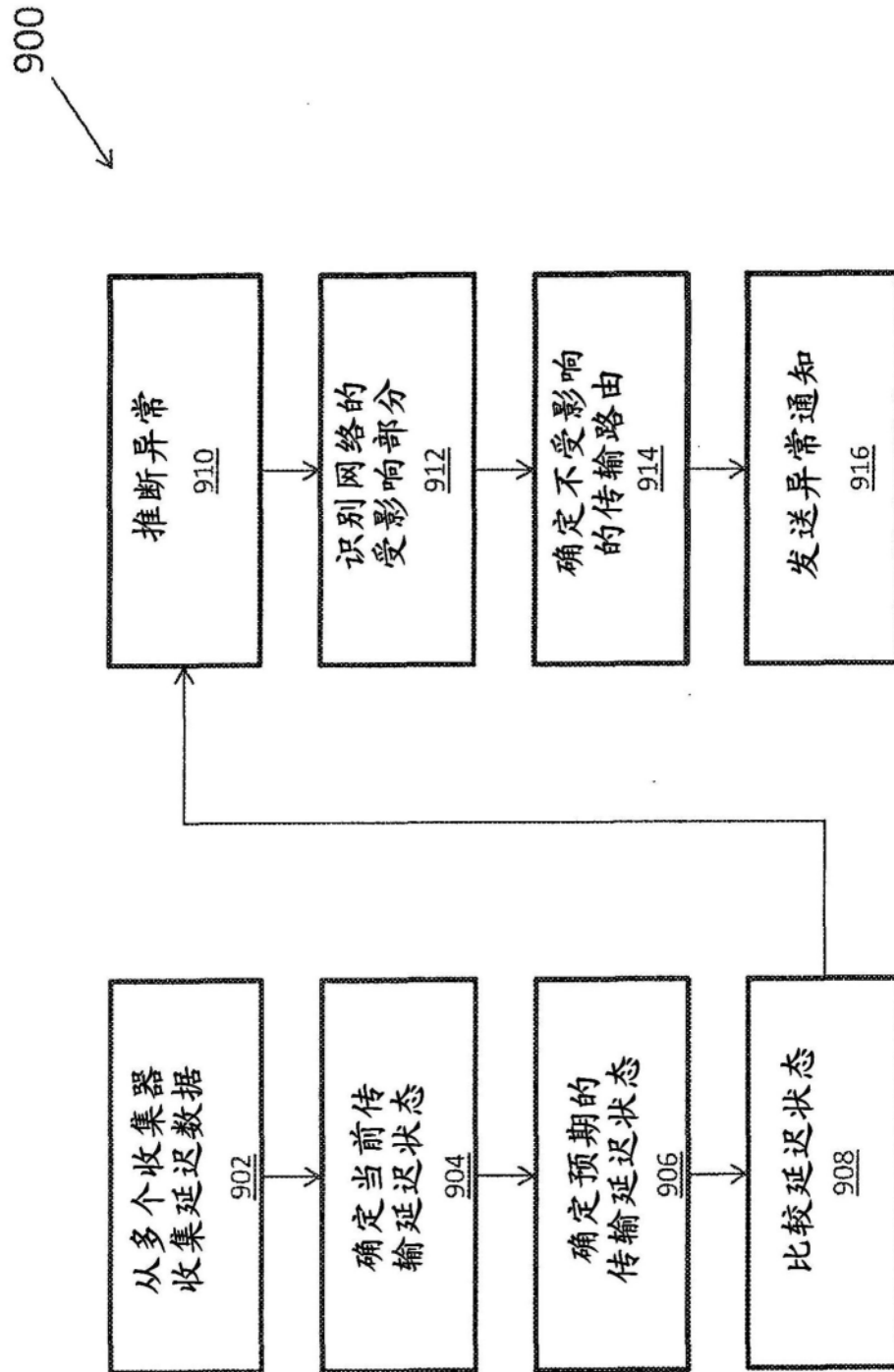


图9

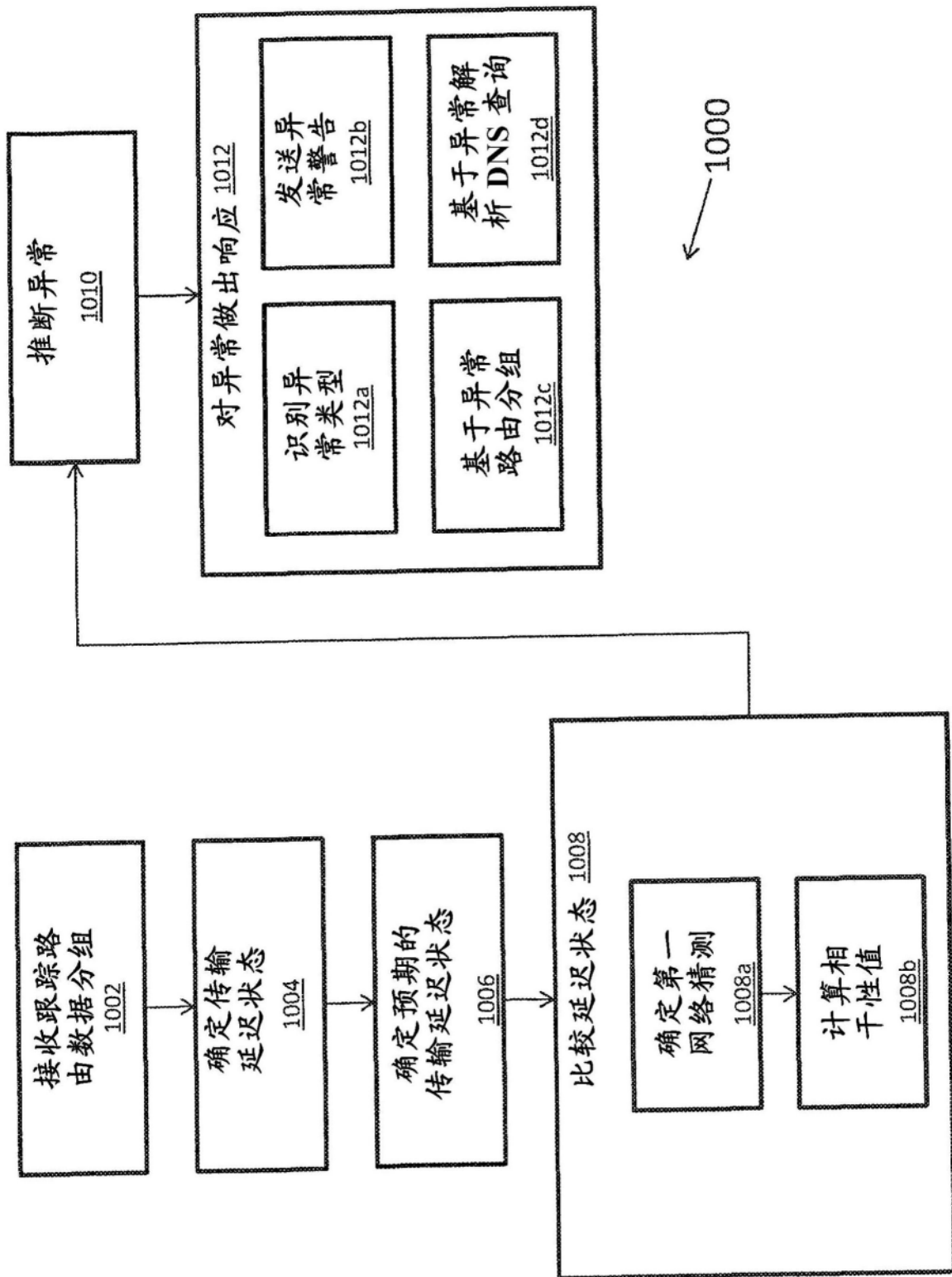


图10

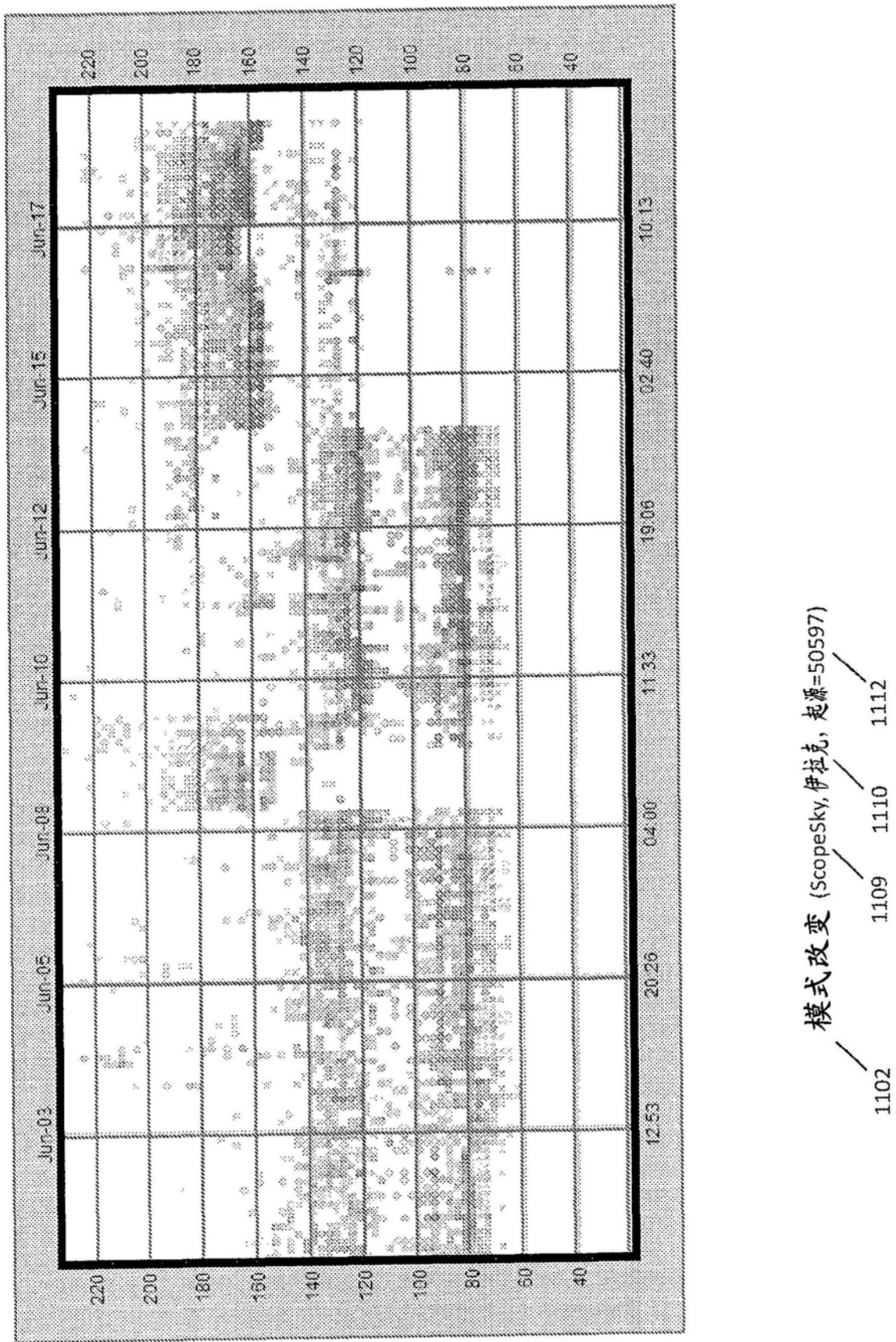


图11A

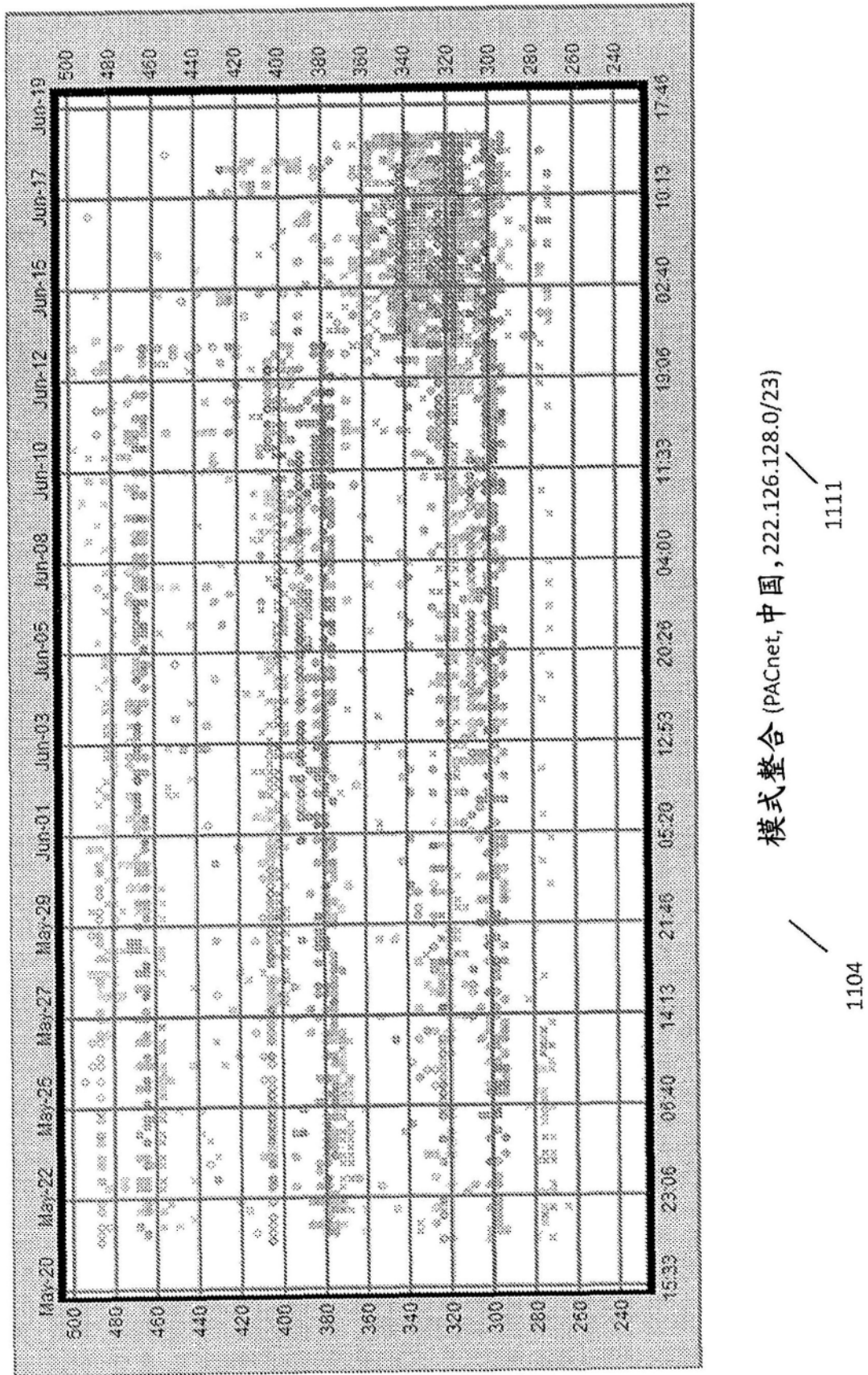


图11B

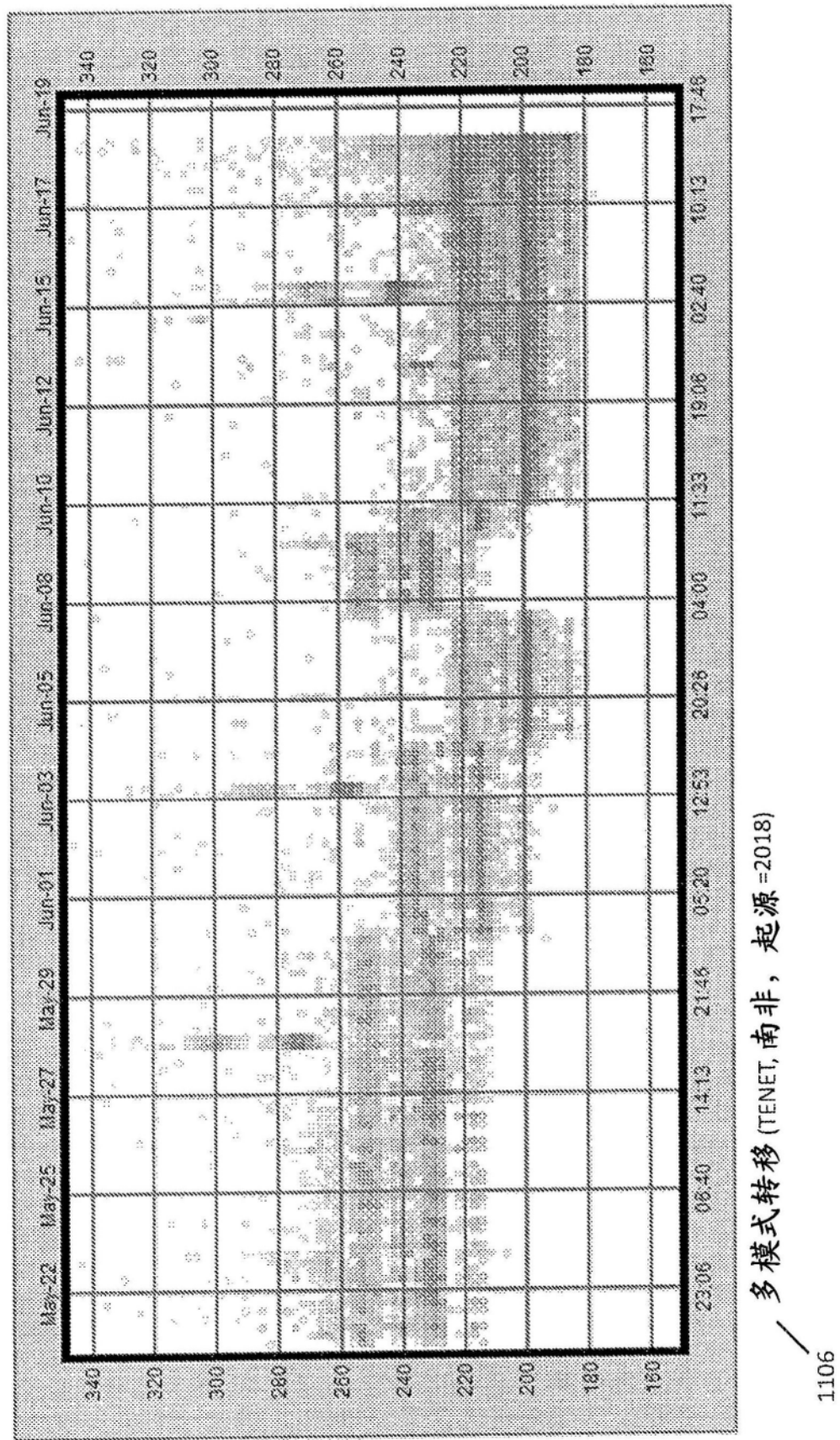


图11C



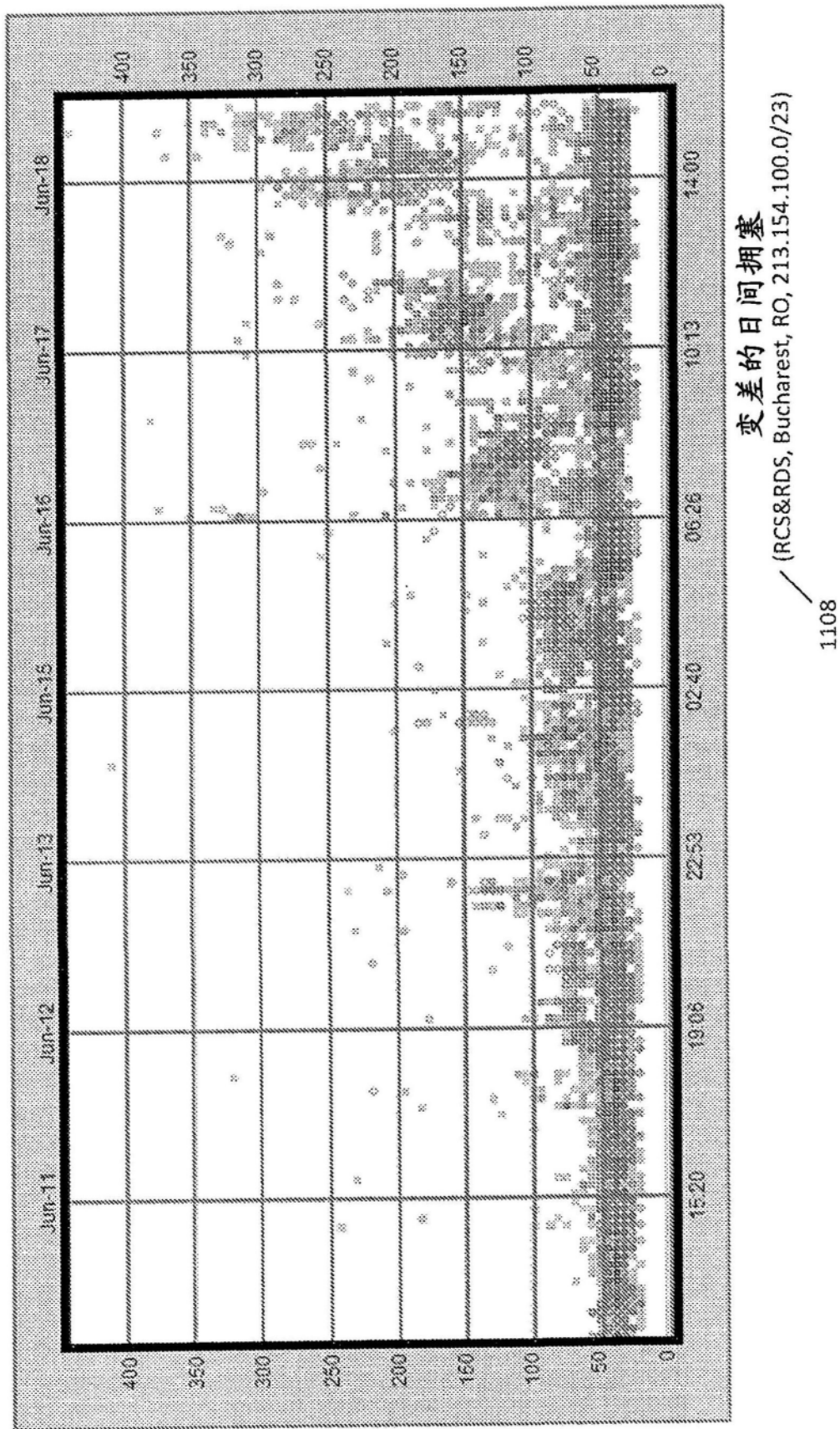


图11D

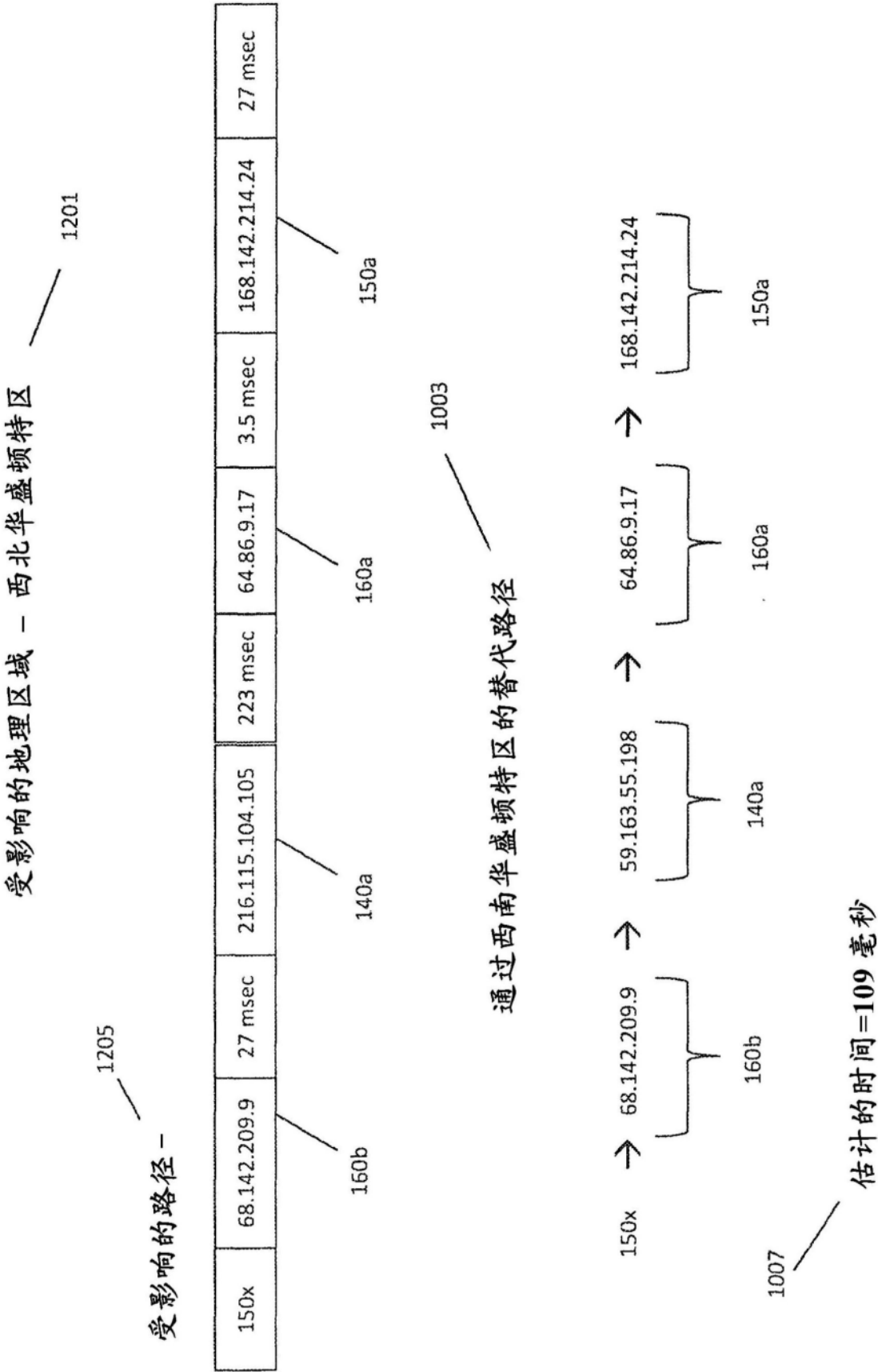


图12