



互联网跨域端到端质量监测及故障定位方案

颜永明¹, 陈兵², 许文杰¹

(1. 中国电信股份有限公司上海分公司, 上海 200085;
2. 上海市信息网络有限公司, 上海 200081)

摘要: 随着互联网市场的迅猛发展, 互联网内容提供商对网络质量提出了更高的要求。大型互联网内容提供商因业务需要, 应用遍布于各地数据中心, 对互联网跨域访问质量有很高的要求。建立互联网跨域端到端质量监测系统, 快速定位域外网络故障, 对互联网运营商及服务提供商具有较大的意义。分析了常用网络监控技术和等价路径中的散列(Hash)算法, 提出利用散列算法实现网络路径全遍历, 监测互联网跨域端到端质量, 最后对故障定位方案给出了建议。

关键词: 跨域; 端到端质量; 散列; 路径全遍历

中图分类号: TN915.41

文献标识码: A

doi: 10.11959/j.issn.1000-0801.2018239

Internet cross-domain end-to-end quality monitoring and trouble location scheme

YAN Yongming¹, CHEN Bing², XU Wenjie¹

1. Shanghai Branch of China Telecom Co., Ltd., Shanghai 200085, China
2. Shanghai Information Network Co., Ltd., Shanghai 200081, China

Abstract: With the rapid development of internet, the internet content providers need much better quality of internet. They usually need to deploy their applications in the internet data centers all over the country due to their business requirements. Users usually cross domain to access in their data centers. So internet content providers have high level of requirements for cross-domain internet quality. It is very important for internet service providers to establish a monitoring system which could monitor end-to-end cross-domain internet quality and locate the troubles of network out of autonomous systems rapidly. The commonly used network monitoring technology and the hash algorithm in equal-cost multipath routing were analyzed. A solution of using Hash algorithm to realize full traversal of network paths and monitor end-to-end cross-domain internet quality, was presented. Some suggestions were also given for trouble location scheme.

Key words: cross domain, end-to-end quality, Hash, full traversal of network paths

1 引言

在各类移动端和桌面端应用, 如游戏、社交

应用、购物以及搜索引擎等需求的推动下, 互联网企业对网络带宽的需求持续增长, 数据中心经多年发展, 建设规模不断扩大。同时 ICP (internet



content provider, 互联网内容提供商) 对网络访问质量提出了更高的要求。云计算数据中心更因业务跨机房调度、配置部署等特性, 对网络高可靠性、稳定性、网络时延、分组丢失、抖动等质量问题更为敏感^[1]。网络异常对这些业务的用户体验造成极大影响。因此, 本文针对因自治域问题而无法直接采集到网元、链路状态的互联网跨域场景, 重点研究了端到端质量监测及故障定位方案。

2 互联网跨域端到端质量监测场景需求和问题

目前, 大型互联网公司大多在全国租用或自建数据中心, 网络游戏、社交应用、在线购物及搜索等各类业务分布式部署, 访问这些服务的用户来自全国各地, 且数量庞大。以手机端游戏业务为例, 玩家来自全国各地, 而服务器则可能集中部署于北京、上海或者广州等主要城市的数据中心, 玩家使用手机终端经本地运营商接入网络、城域网、骨干网, 再到业务所在数据中心城域网, 数据中心服务器经返程完成完整的数据交互。此类业务特性决定了 ICP 非常关注跨域端到端网络质量。从运营商收到的日常申告来看, ICP 主要反映从数据中心到某个/某些省市的访问质量劣化, 影响业务, 并经常要求确认相应省间、区域间 (例如全国往华东区域) 的网络是否存在问题。事实上, 只要端到端路径中某个段落的网络出现异常, 就有可能导致分组丢失、拥塞、时延增大、数据重传, 造成游戏卡顿, 降低用户体验^[2], 因此, 亟需建立一套端到端网络质量监控系统来实时监控网络质量, 快速定位故障点, 以便 ISP (internet service provider, 互联网服务提供商) /ICP 及时排障、迂回, 确保业务质量保持稳定。

3 传统端到端质量监测和故障定位方式

传统网络质量监控通过部署在数据中心的网

络探针向目的被监控 IP 地址进行单向 ping 和 trace 测试, 以发现、定位故障^[3]。这一方式存在如下缺陷:

- 探针部署位置固定, 无法拟合用户数据分组转发路径;
- 由于是网络探针单向做 ping 和 trace, 当监测到故障时无法判断是 ICMP (internet control message protocol, Internet 控制报文协议) 数据分组去程方向路径异常还是回程路径异常;
- 目前数据网络都具有冗余保护措施, 网络设备都是多上联架构, 每条路径都有 2 个以上的冗余节点, 整个网络路径呈现复杂的 ECMP (equal-cost multipath, 等价多路径) 路由, 传统的单向监测的探针 IP 地址和目的 IP 地址相对固定, 导致监测路径也是固定的, 由于无法遍历每条路径, 若监测数据分组正好走在状态正常的路径, 就无法监测其他可能发生质量劣化的等价链路, 大概率与实际用户业务数据分组流经的路径不一致, 对网络质量监测和故障定位造成严重干扰^[4-5]。

4 互联网跨域场景下网络路径全遍历的质量监测方案

4.1 端到端网测监控系统技术分析

有多种方法可实现支持目前主流网络协议场景的网络质量监控, 分为被动式和主动式网络质量监测。

被动式网络质量监控, 如网元告警系统等是运营商常用的监控方式。该监控室采集各种信息例如设备告警、端口流量等, 然后把采集到的信息传送到相关服务器进行筛选、分析、告警和存储等处理。但考虑到用此种方式采集到的数据主要是监控自治域内的网络设备、链路告警, 存在很大的局限性, 且由于维护权限问题, 很难做到跨域的端到端网络质量监控。

采用主动式网络质量监控,在网络部署探针发起监测,对网络上的一些目的地址,例如实际用户访问的互联网网站,也可以通过数据中心间的网络探针做双向监控,根据应用场景灵活部署监控任务。

IP 在当前网络中使用最广泛,ICMP 作为 IP 的一个子协议具有一定的控制能力和网络监控能力,经常用于网络监控和链路故障定位。

4.1.1 ping 测试

ping 测试通过发送 ICMP 数据分组给目标主机,要求目标主机返回应答消息。ICMP 是 TCP/IP (transmission control protocol/internet protocol, 传输控制协议/互联网互联网协议) 协议族的一个子协议,属于网络层协议,广泛使用于检测网络连通性、路由可用性、目的主机是否可达等。遇到数据分组无法到达目的网络、网络带宽速率低、时延高等情况发生时,可发送 ICMP 消息,测试当前网络状态。

ping 测试将 ICMP 数据分组发出,目的 IP 地址对收到的数据进行分组后检查目的地址,如和自身 IP 地址相符则接受,并且把数据分组中相关信息交给 ICMP 处理,封装应答数据分组后回传给源地址。当 ping 测路径异常时,返程应答数据分组里包含路由不可用、目的主机不可达等错误信息。在日常网络运维中经常使用 ping 测试用于检测网络连通性。

4.1.2 trace 工具

trace 发出 TTL (time to live, 存活时间) 值追踪到达目的主机所经过的节点,从节点收到 ICMP 检测的应答数据,用来检测数据发送源到目的节点之间所经节点。信息在网络中传输会经过服务器、交换机、路由器等网络设备,这些设备通常都配置 IP 地址,trace 可以让运维人员知道数据端到端传输所经过的路径。trace 可以测量数据分组从发送到目的地址再返回源地址所需要的时间,其测试报告包含节点 IP 地址、每个节点所花

费时间、分组丢失率等信息。

需要注意的是,根据散列算法数据分组从源地址到目的地址走的路径可能会随着每次测试而改变,使用固定源、目的 IP 地址测试分组测试网络状态时,ping、trace 报告显示路径正常,只能代表这次 ping、trace 测试分组恰巧经过了正常路径,但如果用户业务数据分组经过的是等价链路中的劣化路径,则测试结果与用户感知不一致。此外,每一 trace 返回数据分组因源地址被修改为响应设备地址,五元组散列结果差异会造成回程路径的不同,因此测试结果无法真实反映实际情况。

4.2 散列算法实现全路径覆盖监控方案

数据分组在等价路由转发时,网络设备会将每个连接的所有数据分组都发往多个链路中的其中一条,数据分组建立连接和转发链路是通过网络设备散列算法来完成的,散列算法是根据散列因子来计算路径的^[6],比如,包含源 IP 地址、源端口、目的 IP 地址、目的端口和传输层协议的五元组。

发起监控 ICMP 分组时的散列因子和用户业务数据流建立连接时的散列因子不同,用户数据流恰好走故障链路而影响业务,而监控流经过的链路则有可能是正常路径。日常运维中用户业务有时延、分组丢失等现象,而 ICMP 监控报告则显示正常,此时就有可能是在等价链路中监控数据分组是通过链路 A 传输,而用户业务数据分组是从链路 B 转发。这就要求网络质量监控系统首先要满足监控路径覆盖全链路,实现遍历性。

4.2.1 单向监测

部署在数据中心的探针发起对外部公网互联网网站的 ICMP 监控。探针支持多 IP 地址配置,选取多个互联网网站作为目的 IP 地址进行监测。探针发起对不同目的互联网网站的 trace,探针切换不同源地址继续对不同互联网 IP 地址发起测



试，部署在不同网段的其他探针也持续重复以上监控。经过上述监控步骤可以得到数据中心到互联网的不同访问路径，画出全路径网络监控拓扑结构，如图 1 所示，可以看出左边探针发起对互联网网站地址的 ICMP 监控，当探针用不同源地址或者多个探针发起监控时，会经过多个不同中间节点，满足监控链路的遍历性要求。逐渐提高目的互联网网站地址的数量，查看中间路径层数和节点数是否会明显增加，如果调整到无明显增加时，可以得到覆盖整个中间路径的网络监控拓扑。

单向监测的特点是每个探针进行一对多监测，目的 IP 地址可以在探针最大性能范围内找海量公网存活 IP 地址，根据散列算法，足够多的源-目的 IP 地址因子组合可以实现监测覆盖全路径的效果，用少量探针实现监测遍历性，可以减少部署

在异地数据中心的探针数量，控制成本。而缺点是单向监测只能从全局宏观上发现某一方向的网络质量异常情况，定性给出网络变化情况，无法精确定位故障点。

4.2.2 双向监测

用户网络服务是用户发送请求到数据中心，再从数据中心反馈内容，是双向的，因此，网络设备在两个方向分开进行散列算法。如果说用户到数据中心的的方向是路径 A，那么数据中心分组返回时走的路径可能是 B，而 ping 和 trace 报告都是单向监控，如果需要监控双向路径，就要在正反方向分别相互做 ping 和 trace 监控。单向监控的局限性决定，单向监控只能用来做告警发现和绘制全路径网络监控拓扑图，要实现故障点定位功能，需要结合探针双向监控。

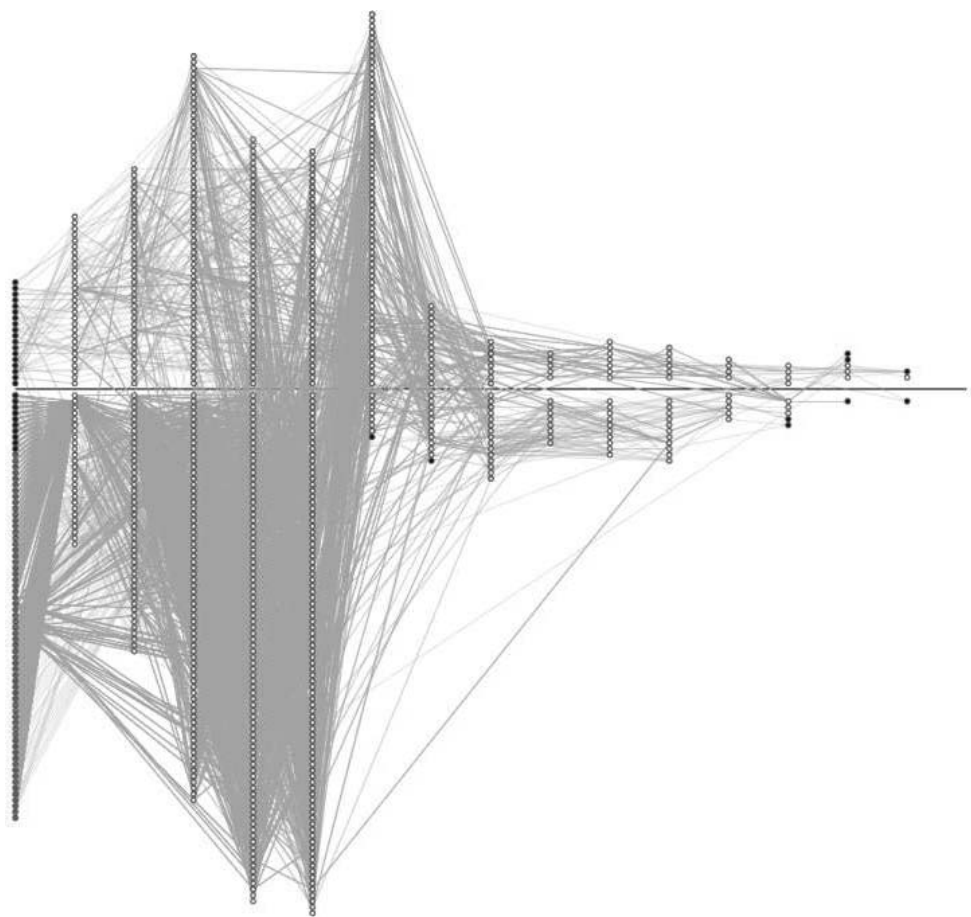


图 1 数据中心到公网网络监控拓扑结构

在日常网络运维中会碰到以下情况，见表 1。

如表 1 所示，正向 trace 报告中第 10 跳已到用户侧网络且显示分组丢失出现在最后一跳。第 10 跳分组丢失率为 0，依据排障经验可以判断之前显示的分组丢失率为中间节点设备对 ICMP 分组做的保护机制，对正常业务数据转发无影响，故障点在最后一跳用户设备。

但是从表 2 反向 trace 报告可看出，分组丢失出现在第 8 跳节点，之后分组丢失率一直延续，故障点在第 8 跳省际链路中间节点。最后的排查结果是上海到北京的一根链路故障导致的拥塞引起分组丢失。业务请求和数据反馈走的双向路径会不同，且故障往往发生在其中一条链路上，在实际日常运维中会经常碰到结合单向和反向监测

报告判断出故障点的情况。

在需要监控路径的多个数据中心部署多个探针，探针支持多 IP 地址配置，数据中心中的多个探针与另外数据中心探针进行相互 ICMP 监控。不同数据中心探针间形成 full-mesh 监控，如图 2 所示。

探针源地址数和探针部署量，可以根据单向监控方式绘制的全路径网络监控拓扑图作调整。探针间 full-mesh 监控结果可得出数据中心间全路径覆盖监控，且可以准确定位路径中的故障节点，如图 3 所示。

4.3 端到端网络质量监控方案的实现

系统由网络探针和数据处理服务器组成，网络探针用于发起监控，收集监控数据，数据处理

表 1 正向 trace

IP地址	数据分组				ping		
	分组丢失率	次数	最近	平均	最佳	最差	标准偏差
1. 123.1xx.xx.x	0	138	0.6	0.8	0.5	3.3	0.3
2. ???							
3. 123.15x.xx.xx	52.6%	138	2.6	2.5	1.6	5.8	0.7
4. 221.2xx.xx.xxx	0	138	1.8	2.0	1.6	10.0	0.9
5. 221.2xx.x.xxx	0	138	1.7	1.9	1.6	10.7	1.1
6. 201.9x.xx.xx	59.1%	138	4.7	4.5	4.4	4.7	0.1
7. 101.9x.xx.xxx	95.6%	137	32.1	33.3	32.0	39.8	3.1
8. ???							
9. 101.xx.xxx.xxx	29.4%	137	28.5	28.5	28.4	28.7	0.1
10. 101.2.xx.xxx.xx	0	137	29.7	29.6	29.3	30.0	0.1
11. ???							
12. 101.xx.x.xx	0.7%	137	131.1	130.6	127.6	131.4	0.7

表 2 反向 trace

IP地址	数据分组				ping		
	分组丢失率	次数	最近	平均	最佳	最差	标准偏差
1. 101.xx.x.x	0	119	0.7	0.7	0.5	1.5	0.2
2. ???							
3. 101.2xx.xxx.xx	41.2%	119	0.7	0.7	0.6	1.9	0.2
4. 101.9x.xxx.xxx	0	119	1.4	1.7	1.4	8.7	0.9
5. 101.xx.xxx.xxx	0	119	6.3	6.0	1.8	9.6	2.4
6. 202.xxx.xx.xxx	0	119	128.4	129.9	125.6	139.5	1.6
7. 202.xx.xx.xx	0	119	130.5	131.3	127.3	133.9	1.5
8. 221.2xx.x.xxx	44.1%	119	130.7	130.3	125.9	131.1	0.9
9. ???							
10. 12x.xxx.xx.xx	0.8%	119	130.8	130.4	126.9	131.3	0.8
11. ???							
12. 12x.xxx.xx.xx	2.5%	118	130.9	130.5	127.1	131.2	0.87

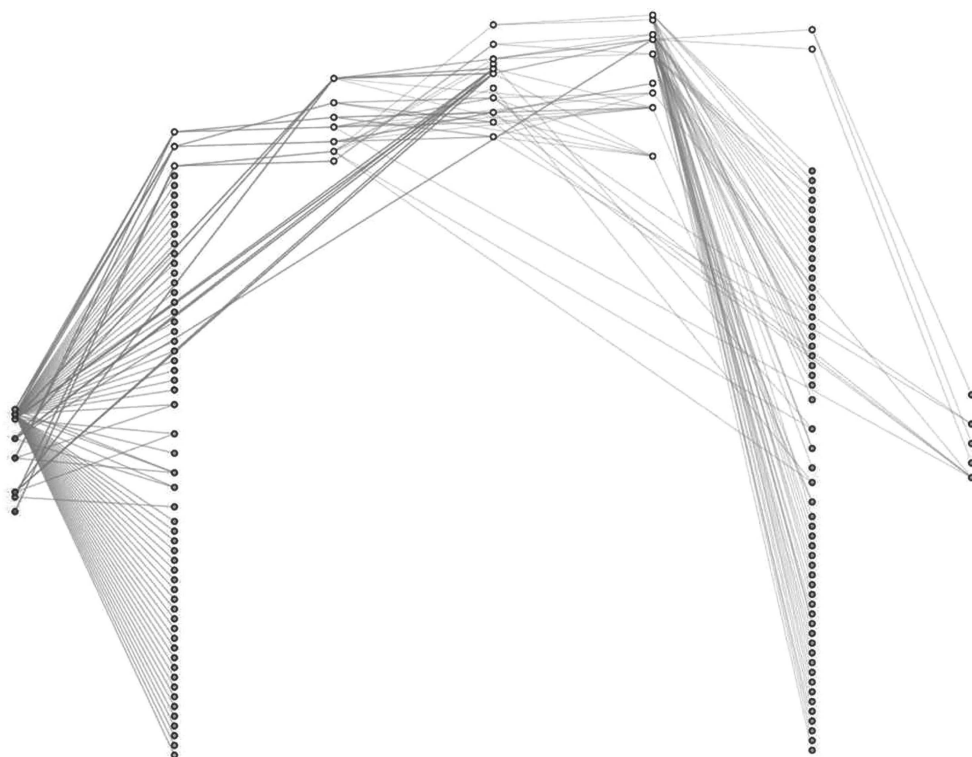


图2 数据中心间形成 full-mesh 监控

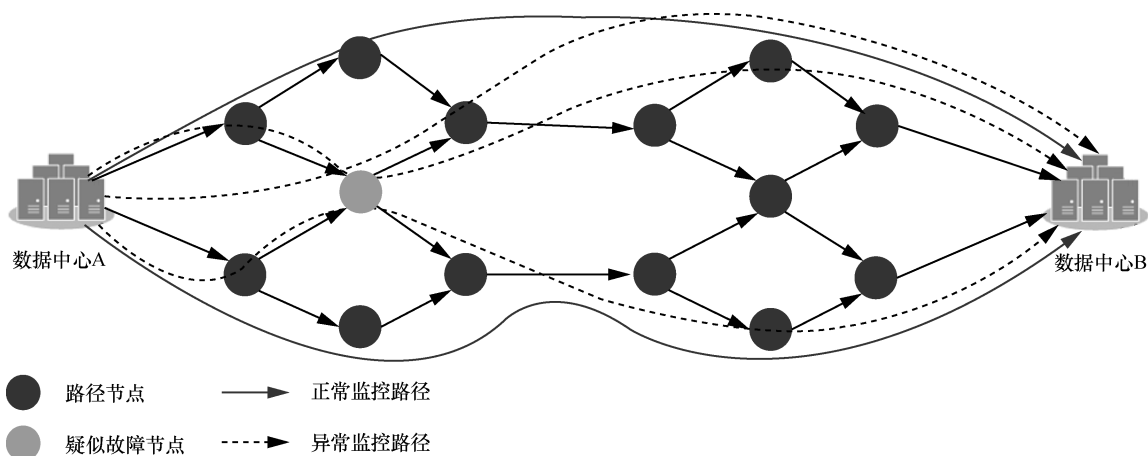


图3 定位故障点

服务器负责对探针进行配置下发和管理，然后对收集到的监控数据进行处理。方案中单向监测和双向监测互为补充。

单向监测选取海量互联网网站地址作为目的IP地址，可以用少量探针达到监测遍历性的效果，用单向 trace 可以画出某方向的全路径监测拓扑，并且根据周期性监测数据动态调整拓扑图，同时，用单向 ping 方式得到某一方向的宏观网络质量。

双向监测在满足遍历性的情况下实现故障点精确定位，通过两端数据中心部署的探针相互 ping/trace，可以做到快速发现故障并准确定位故障点。

从单向和双向监测的历史数据中可以看出网络周期性变化情况，分析计算出各方向、各时间段的网络质量基线，再将网络质量基线应用在日常实时监测中，当实时监测数据偏离质量基线时

触发告警。为了减少探针和网络负荷，在日常监测中单向和双向监测频率可以适当降低，且只进行 ping 测定性网络质量情况。当定性发现异动时再触发高频率的监测，用双向 trace 方式定位故障点。

探针采用虚拟化方式部署在云资源池。虚拟化部署成本低、部署简便，其次在后续维护中无需到达现场，在远端即可登录查看，方便管理。因为探针部署在省外，要确保探针不会失联。虚拟化探针支持配置多个 IP 地址，当数据上传或者管理通道走在故障路径时，只要探针本身状态正常没有僵死，或者链路全断的情况外，数据处理服务器可以通过轮询配置好的多个 IP 地址正常工作。

通过端到端的应用监控，可以很快掌握网络中间的故障点和用户体验，进一步迅速排查网络出现的问题、定位故障点，为减少故障和提高用户体验，起到了非常重要的作用。

4.3.1 了解网络拓扑变化

通过网络探针持续监控各个方向的网络质量和监控路径，可以掌握整个网络拓扑信息。根据历史监控数据，周期性调整拓扑图，比如一周更新一次拓扑图。根据拓扑信息，了解网络变化的趋势，还可以为以后的网络变更及带宽扩容等长期规划提供参考依据，对某些经常发生网络故障的链路和节点及时做优化和整改方案。

4.3.2 制定网络性能基线

长期监控链路质量，根据节点性能，将忙时和闲时、工作日和节假日等不同维度的网络质量监控数据形成动态调整的基线。监控系统把实时收集的监控数据和告警基线相比较分析，之后弹出故障点、故障等级等告警信息，提醒运维人员确认，后续跟进处理。

4.3.3 收敛链路告警数据

数据处理服务器根据链路方向、网络拓扑数据等信息来确定网络探针收集到的监控数据。例如，同一路径方向上的告警数据归为一条告警，

网络拓扑数据中属于相同网络层级或同一网络设备的 IP 地址告警归为相同告警，避免运维人员在海量告警数据中浪费故障处理时间。根据不同监控场景，制定多种收敛算法。

4.3.4 多维度报表

可以根据网络质量监控数据多维度提供网络性能报表，根据不同区域、不同数据中心、不同访问方向、多个监控路径等维度来掌握网络质量信息。还可以根据新的监控场景和业务场景，自定义质量监控数据，以图表和报表方式呈现网络健康状态。

4.4 网络监控拓扑

根据需求该系统可以实现的监控网络拓扑如图 4 所示。

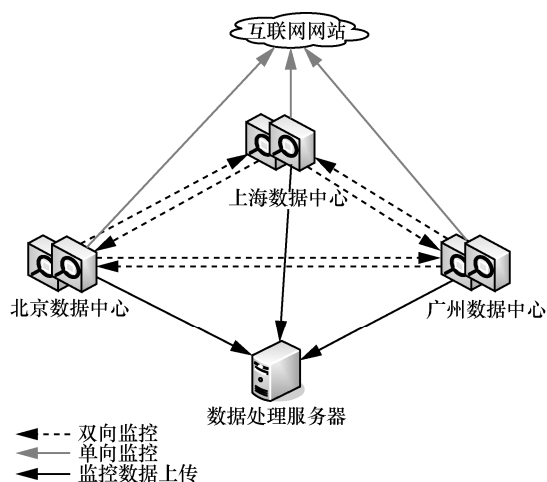


图 4 监控网络拓扑

具体如下：

- 在主要城市数据中心部署网络探针；
- 中间数据处理服务器为部署在各省市数据中心的网络探针下发配置、任务，集中管理部署；
- 部署在各省市数据中心的网络探针相互形成 full-mesh 监控数据流，进行双向监控；
- 部署在各省市数据中心的网络探针也发起对互联网网站 IP 地址的监控数据流，进行单向监控。



4.5 与传统网络质量监测方式对比

数据分组在网络中传递时会经过多个节点，现在网络为了冗余性，节点和节点间用多个等价链路做负载均衡，通常会用散列算法来实现。使用五元组（源地址、目的地址、源端口、目的端口、协议类型）为散列因子，经过散列算法得出的散列值作为区分不同数据流的标志，映射到不同等价链路。在 ICMP 测试中，每次 ping 测试或者 trace 都计算为一次 ICMP 连接，所以每次 ICMP 测试都只会监控多条链路中的一条，在此次 ICMP 测试中其他等价链路不会被监控到。发起监控 ICMP 分组时的散列因子和用户业务数据流建立连接时的散列因子不同，用户数据流恰好走故障链路而影响业务，而监控流经过的链路则有可能是正常路径。

端到端网络质量监控方案通过改变监控探针 IP 地址和监控目的 IP 地址，即改变散列算法中散列因子的方式，使监控数据流遍历链路，及时掌握网络拓扑变化，快速定位网络故障，并且通过分析监控数据，动态调整网络性能基线，收敛告警。此方案解决了传统网络质量监控系统遇到的故障发现不及时、故障定位不准确、无法掌握网络拓扑、性能基线设置单一等问题，可使日常运维工作变得更简单、高效。与传统网络质量监测方式比较见表 3。

5 结束语

根据散列算法用足够多的源目 IP 地址组合实现监控路径遍历性，通过单向监测方式绘制监测拓扑发现问题，用双向监测进一步定位故障，使用监控数据制定网络性能基线，收敛告警数据，

建立跨域端到端网络质量监测系统可弥补传统网络质量监控手段的不足。后续可以考虑结合多种不同协议配合 ICMP 做路径监控，如使用 HTTP（hypertext transfer protocol，超文本传输协议）、UDP（user datagram protocol，用户数据报协议）。网络路径中间节点设备会对 HTTP 和 UDP 做转发策略，而针对 ICMP 分组，一些路由器、交换机等设备会为了不影响转发业务数据性能，对 ping 和 trace 协议分组做限制。根据 ICMP、HTTP、UDP 3 种协议各自优点，在不同场景结合使用的话可能会得到更加准确可靠的监测结果。

参考文献：

- [1] 樊自甫, 伍春玲, 王金红. 基于 SDN 架构的数据中心网络路由算法需求分析[J]. 电信科学, 2015, 31(2): 42-51.
FAN Z F, WU C L, WANG J H. Requirements analysis of data center network routing algorithm based on SDN architecture [J]. Telecommunications Science, 2015, 31(2): 42-51.
- [2] 谢海华. 有线 IP 城域网可视化质量监控系统的建设与运用研究[J]. 无线互联科技, 2017(23).
XIE H H. Study on construction and application of visual quality monitoring system for CATV's IP metropolitan area network[J]. Wuxian Hulian Keji, 2017(23).
- [3] 夏刚. 互联网环境的网络质量监测体系研究与实践[J]. 中国金融电脑, 2016(7): 48-51.
XIA G. Research and practice of network quality monitoring system in internet environment[J]. China Financial Computer, 2016(7): 48-51.
- [4] 覃佐曼. 基于 SDN 的数据中心网络多路径负载均衡的研究[D]. 大连: 大连海事大学, 2017.
QIN Z M. Research on multipath load balancing in data center network based on SDN[D]. Dalian: Dalian Maritime University, 2017.
- [5] 安禄. 基于等价多路径的数据中心网络流量优化问题研究[D]. 重庆: 重庆大学, 2014.
AN L. The optimization research of traffic engineering for data

表 3 与传统网络质量监测方式对比

	传统网络质量监控系统	端到端网络质量监控系统
监测方式	单向监测	单向、双向监测相结合
监测遍历性	单向监测且单一源目的地址无法实现监测链路遍历性	单向、双向监测方式相结合，且支持多个源地址和目的地址组合，可以实现监测链路遍历性
定位故障	单向监测方式无法判断是去程方向路径异常还是回程路径出现问题，无法定位故障	双向监测方式可以做到准确定位故障

center networks based on ECMP[D]. Chongqing: Chongqing University, 2014.

- [6] 程光, 龚俭, 丁伟, 等. 面向 IP 流测量的散列算法研究[J]. 软件学报, 2005, 16(5): 652-658.

CHENG G, GONG J, DING W, et al. A hash algorithm for IP flow measurement[J]. Journal of Software, 2005, 16(5): 652-658.

[作者简介]



颜永明 (1978-), 男, 中国电信股份有限公司上海分公司信息网络部综合运营监控中心副经理、高级工程师, 主要研究方向为数据网络、云组网等。



陈兵 (1970-), 男, 上海市信息网络有限公司总经理助理、高级工程师, 主要研究方向为数据通信、大数据挖掘等。



许文杰 (1986-), 男, 中国电信股份有限公司上海分公司信息网络部综合运营监控中心技术工程师、助理工程师, 主要研究方向为数据网络。