解决企业网络中的路由器问题

目录

简介

本文档介绍如何使用Cisco路由器识别、排除和解决企业网络中的延迟问题。

先决条件

要求

本文档没有特定的前提条件或要求。

使用的组件

本文档不限于特定的软件版本和硬件类型，但命令适用于Cisco IOS® XE路由器，例如ASR 1000、ISR 4000和Catalyst 8000系列。

本文档中的信息都是基于特定实验室环境中的设备编写的。本文档中使用的所有设备最初均采用原始（默认）配置。如果您的网络处于活动状态，请确保您了解所有命令的潜在影响。

背景信息

本文档介绍了了解、隔离和排除一般延迟问题的基本指南，并提供了用于检测根本原因和最佳实践的有用命令/调试。请记住，不能考虑所有可能的变量和方案，而更深入的分析取决于具体情况。

延迟定义

一般而言，在引用存储和转发设备的严格定义（在RFC 1242上）时，延迟是输入帧的最后一位到达输入端口时开始到输出端口上看到输出帧的第一位时结束的时间间隔。

网络延迟可能只是指通过网络传输数据的延迟。对于实际问题，此定义只是一个起点；您需要定义针对每个具体案例所讨论的延迟问题，虽然看上去显而易见的是，解决问题所需要的第一步是定义问题，而且这变得非常重要。

延迟使用情况

许多应用要求实时通信和业务运营的低延迟；随着每天硬件和软件改进，更多应用可用于任务关键型计算、在线会议应用、流传输等；同样，网络流量持续增长，优化网络设计和提高设备性能的需求也不断增加。

除了提供更好的用户体验和满足延迟敏感型应用的最低要求外，有效地识别并减少网络上的延迟问题可以节省大量时间和网络上非常有价值的资源。

处理延迟问题

此类问题的难点在于必须考虑变量数量以及不能出现单点故障。因此，延迟的定义成为解决延迟问题的重要关键，下面是您必须考虑的一些方面才能对问题进行有用的说明。

1.期望和检测

区分所需延迟、预期或基线工作延迟和当前延迟非常重要。根据网络设计、提供商或设备，有时您可能无法实现所需的延迟，在正常情况下测量实际延迟是一个不错的过程，但您需要对测量方法保持一致，以避免产生误导性数字；IP SLA和网络分析工具可在这方面提供帮助。

通过ICMP或ping来识别应用甚至IP SLA延迟的最常用和基本工具之一：

<#root>

```
Router#
ping
 198.51.100.1
Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 198.51.100.1, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5),
round-trip min/avg/max
```

=

2/109/541 ms

除了检查可达性之外，ping还指示从源到目的地的往返时间(RTT)；最小值(2)、平均值(109)和最大值(541)（以毫秒为单位）。这意味着从路由器收到来自设备目标的应答时开始发出请求的持续时间。但是，它不会显示多少跳或更深入的信息，但它是一种检测问题的简单、快速的方法。

2.隔离

与ping一样，traceroute可用作隔离的起点，它发现跳数和RTT每跳：

<#root>

Router#

traceroute

 198.51.100.1
Type escape sequence to abort.
Tracing the route to 198.51.100.1
VRF info: (vrf in name/id, vrf out name/id)
  1 10.0.3.1 5 msec 6 msec 1 msec
  2 10.0.1.1 1 msec 1 msec 1 msec
  3 10.60.60.1 1 msec 1 msec 1 msec
  4 10.90.0.2

362 msec 362 msec 362 msec

<<<< you can see the RTT of the three probes only on both hops

  5 10.90.1.2

363 msec 363 msec 183 msec

  6 10.90.7.7 3 msec 2 msec 2 msec

Traceroute通过发送生存时间(TTL)为1的数据包来运行。第一跳发回ICMP错误消息，指示由于TTL已过期且测量了RTT而无法转发数据包，然后使用TTL 2重新发送第二个数据包，第二跳返回TTL已过期。此过程会一直持续到到达目的地为止。

在本例中，现在您可以缩小到两台特定主机，然后可以开始隔离该主机。

尽管这些命令非常有用，可以轻松识别问题，但它们不会考虑其他变量，例如协议、数据包标记和大小（虽然您可以将其设置为第二步）、不同的IP源、多个因素中的目的地。

说延迟是一个非常宽泛的概念，您通常只看到应用程序、浏览、呼叫或特定任务上的症状。首先要限制的事情之一是了解影响并更详细地定义问题，回答接下来的问题和元素可以帮助进行这种划分：

- 延迟是否只影响特定类型的流量或应用？示例：仅UDP、TCP、ICMP...
- 如果是，此流量是否具有唯一标识符？示例：特定QoS标记、仅确定的数据包大小、IP选项……
- 有多少用户或站点受到影响？示例：只有一个特定子网、一台或两台终端主机、连接到一个或多个设备的整个站点……
- 是否确定了特定的时间戳？示例：这是否仅在高峰时间、任何时间模式或完全随机……
- 设计方面。示例：流量通过特定设备（可能有许多设备，但只连接到一个提供商），流量进行负载均衡，但影响一个路径……

还有其他许多注意事项，但是完成不同的答案（甚至可以对其进行回答的测试）可以有效隔离并限制故障排除的进行范围。例如，只有一个应用（相同类型的流量）在所有分支上受到影响，该应用通过不同的提供商，并在高峰时段在同一数据中心结束。在这种情况下，您不会开始检查所有分支机构中的所有接入交换机，而是集中精力收集有关数据中心的更多信息，然后在该侧进一步检查，

您在网络上可以使用的监控工具和一些自动化功能在很大程度上也依赖于这种隔离，这真的取决于您拥有的资源和独特的情况。

常见原因故障排除

一旦限制故障排除的范围，就可以开始检查特定原因，例如，在提供的traceroute示例中，可以隔离到两个不同的跳，然后缩小到可能的原因。

平台相关

高 CPU 利用率

其中一个常见原因可能是设备的CPU使用率较高，导致延迟处理所有数据包。对于路由器而言，最有用、最基本的路由器检查命令是

路由器的整体性能：

<#root>

Router#

show platform resources

```
**State Acronym: H - Healthy, W - Warning, C - Critical
Resource                Usage                Max            Warning         Critical        State
-------------------------------------------------------------------------------------------------
RP0 (ok, active)                                                                               H

Control Processor       1.15%                100%           80%             90%             H

DRAM                    3631MB(23%)          15476MB        88%             93%             H
```

```
  bootflash              11729MB(46%)          25237MB         88%             93%             H
  harddisk               1121MB(0%)            225279MB        88%             93%             H
ESP0(ok, active)                                                                                 H
 QFP                                                                                             H
  TCAM                   8cells(0%)            131072cells     65%             85%             H
  DRAM                   359563KB(1%)          2097152OKB      85%             95%             H
  IRAM                   16597KB(12%)          131072KB        85%             95%             H

CPU Utilization        0.00%                 100%            90%             95%             H


Crypto Utilization     0.00%                 100%            90%             95%             H

  Pkt Buf Mem (0)        1152KB(0%)            164864KB        85%             95%             H
  Pkt Buf CBlk (0)       14544KB(1%)           986112KB        85%             95%             H
```

同时查看内存和CPU利用率很有用，它在控制平面和数据平面(QFP)上与每个平面的阈值相同。内存本身不会产生延迟问题，但是，如果控制平面没有更多DRAM内存，则思科快速转发(CEF)被禁用并导致高CPU使用率（这会产生延迟），这就是保持数字处于正常状态的重要原因。有关内存故障排除的基本指南不在讨论范围之内，但请参考“相关信息”部分中的有用链接。

如果检测到控制处理器、QFP CPU或加密使用率的CPU使用率较高，您可以使用以下命令：

对于控制平面：

show process cpu sorted

<#root>

```
Router#
show processes cpu sorted
CPU utilization for five seconds:
99%/0%
; one minute: 13%; five minutes: 3%
 PID Runtime(ms)     Invoked      uSecs   5Sec   1Min   5Min TTY Process
  65        1621         638       2540 89.48%  1.82%  0.41%  0 crypto sw pk pro
   9         273          61       4475  1.56%  0.25%  0.05%   0 Check heaps
  51         212          64       3312  0.72%  0.21%  0.05%   0 Exec
 133         128          16       8000  0.60%  0.08%  0.01%   0 DBAL EVENTS
 473          25          12       2083  0.48%  0.04%  0.00%   0 WSMAN Process
  84        1173         353       3322  0.36%  0.07%  0.02%   0 IOSD ipc task
  87          23          12       1916  0.24%  0.02%  0.00%   0 PuntInject Keepa
  78         533         341       1563  0.12%  0.29%  0.07%   0 SAMsgThread
 225          25        1275         19  0.12%  0.00%  0.00%   0 SSS Feature Time
 386           4           4       1000  0.12%  0.00%  0.00%   0 Crypto WUI
 127         204       18810         10  0.12%  0.02%  0.00%   0 L2 LISP Punt Pro
```

如果控制平面CPU过高（由于进程原因，此示例为99%），则需要隔离进程，并视情况继续隔离

（可以为我们传送的数据包，如ARP或控制网络数据包，可以是任何路由协议、组播、NAT、DNS、加密流量或任何服务）。

根据您的流量，这可能会造成进一步处理的问题，如果流量并非发往路由器，您可以重点关注数据平面：

对于数据平面：

show platform hardware qfp active datapath utilization [summary]

<#root>

```
Router#
show platform hardware qfp active datapath utilization
  CPP 0: Subdev 0
5 secs
        1 min        5 min       60 min
Input:  Priority (pps)            0            0            0            0
                 (bps)            0            0            0            0
    Non-Priority (pps)          231          192           68            6
                 (bps)       114616        95392        33920         3008
           Total (pps)          231          192           68            6
                 (bps)       114616        95392        33920         3008
Output: Priority (pps)            0            0            0            0
                 (bps)            0            0            0            0
    Non-Priority (pps)            3            2            2            0
                 (bps)        14896         9048         8968         2368

Total (pps)
         3323         2352          892            0

(bps)
        14896         9048         8968         2368
Processing: Load (pct)

3
            3            3            3

Crypto/IO

Crypto: Load (pct)            0
            0            0            0
        RX: Load (pct)           0            0            0            0
        TX: Load (pct)           1            1            0            0
            Idle (pct)          99           99           99           99
```

如果数据平面很高（通过达到100%的处理负载数字标识），则需要查看通过路由器的流量大小（每秒数据包总数和每秒比特数）和平台的吞吐量性能（您可以在特定数据表上有所了解）。

为了确定此流量是否是预期流量，可以使用数据包捕获(EPC)或任何监控功能（如Netflow）进行进一步分析，一些检查包括：

- 流量是否有效并且预期会通过此路由器？
- 识别异常流量或较高的速率。
- 如果每秒数据包数较高，请查找数据包的大小。确定这是否预期会出现，或者您是否遇到分段问题。

如果所有流量都是预期值，则可能会达到平台限制，然后，通过show running-config查找路由器上运行的功能，作为第二个部分进行分析，主要是在接口上，找出任何不必要的功能并禁用它们，或者平衡流量以释放CPU周期。

但是，如果没有平台限制的指示，另一个用于验证路由器是否在数据包上增加延迟的有用工具是FIA跟踪，您可以看到每个数据包所花费的确切处理时间，以及占用大部分处理时间的功能。完全高CPU故障排查不在本文档的讨论范围之内，但请参阅相关信息部分的链接。

流量相关

MTU和分段

最大传输单位(MTU)是要传输的最大数据包长度，取决于物理链路可以传输的八位组数量。当上层协议将数据提交到底层IP时，IP数据包的长度大于路径MTU，数据包被划分为多个分段。网络中的这种小尺寸导致在某些情况下进行更多处理和不同的处理，因此您必须尽可能避免使用它。

对于NAT或基于区域的防火墙等功能，需要虚拟重组来“拥有整个数据包”、应用所需内容、转发其分段，并丢弃重组后的副本。此过程会增加CPU周期且容易出错。

某些应用程序不依赖分段，检查MTU的最基本测试之一是带无分段选项的ping，并测试不同的数据包大小：ping ip-address df-bit size number。 如果ping不成功，请在发生丢弃时修复路径上的MTU，并导致进一步的问题。

基于策略的路由和带有分段数据包的网络上的等价多路径等功能会造成延迟问题和更多错误，主要表现在高数据速率、导致高组装时间、重复ID和损坏的数据包，如果发现其中一些问题，请尽可能解决此分段。用于检查是否存在分段和任何潜在问题的命令是show ip traffic:

<#root>

```
Router#
```

```
show ip traffic
```

```
IP statistics:
  Rcvd:  9875429 total, 14340254 local destination
         0 format errors, 0 checksum errors, 0 bad hop count
         0 unknown protocol, 0 not a gateway
         0 security failures, 0 bad options, 0 with options
  Opts:  0 end, 0 nop, 0 basic security, 0 loose source route
```

```
         0 timestamp, 0 extended security, 0 record route
         0 stream ID, 0 strict source route, 0 alert, 0 cipso, 0 ump
         0 other, 0 ignored

Frags:


150 reassembled

, 0

timeouts

,

0 could not reassemble

         0

fragmented

, 600

fragments

, 0

could not fragment

         0 invalid hole
  Bcast: 31173 received, 6 sent
  Mcast: 0 received, 0 sent
  Sent:  15742903 generated, 0 forwarded
  Drop:  0 encapsulation failed, 0 unresolved, 0 no adjacency
         0 no route, 0 unicast RPF, 0 forced drop, 0 unsupported-addr
         0 options denied, 0 source IP address zero
<output omitted>
```

在以上输出中，Frags部分上的粗体字是指：

- 重组：重组的数据包数量。
- 超时：每次数据包分段的重组时间到期时。
- 无法重组：无法重组的数据包数量。
- 分段：超过MTU和要分段的主题的数据包数量。
- 分段：数据包被分段到的分块数。
- 无法分段：超过MTU但无法分段的数据包数。

如果使用了分段并且您有超时或无法重组计数器增加，则验证由平台引起问题的一种方法是通过QFP丢弃，使用后面在丢弃部分介绍的相同命令：show platform hardware qfp active statistics drop。查找错误，例如：TcpBadfrag、IpFragErr、FragTailDrop、ReassDrop、ReassFragTooBig、ReassTooManyFrags、ReassTimeout或相关错误。每种情况都有不同的原因，例如没有获取所有分段、重复、CPU拥塞等。同样，用于进一步分析和潜在修复的有用工具可以是FIA跟踪和配置检查。

TCP提供最大分段大小(MSS)机制来解决此问题，但如果发现不正确、未协商的MSS或错误的路径MTU，则可能导致延迟。

由于UDP没有这种分段机制，您可以依靠手动实施PMTD或任何应用层解决方案，因此您可以启用它们（在适用时）发送小于576字节的数据包，这是按照RFC1122发送编号的较小有效MTU，以帮助避免分段。

设计相关的

除了故障排除建议，本节还简要介绍两个可能增加延迟问题的关键组件，它们需要本文档范围之外的广泛讨论和分析。

次优路由

网络中的“次优路由”是指数据包不通过网络中可用的最有效或最短路径传输的情况。相反，这些数据包采用的路由效率较低，可能会导致延迟增加、拥塞或影响网络性能。IGP始终选择最佳路径，这意味着成本较低，但不一定是最便宜的路径或延迟最低的路径（最佳路径可以是带宽较高的路径）。

路由协议问题可能导致次优路由；配置或任何情况，如竞争条件、动态更改（拓扑更改或链路故障）、基于公司策略的流量工程或成本、冗余或故障切换（在特定条件下转到备份路径）等。

诸如traceroute或监控设备之类的工具可帮助识别特定流量的这种情况（如果情况确实如此，并且取决于许多其他因素），满足应用需求并降低延迟可能要求重新设计路由或流量工程。

服务质量 (QoS)

通过配置服务质量(QoS)，您可以优先处理特定类型的流量，而牺牲其他流量类型。如果没有QoS， device 为每个数据包提供尽力而为服务，无论数据包的内容或大小如何。此 device 发送数据包，但不保证可靠性、延迟边界或吞吐量。

如果实施了QoS，则非常有必要确定路由器是标记、重新标记还是仅对数据包进行分类，检查配置并show policy-map [name_of_policy_map | 会话 | interface interface_id]有助于了解受高速率、丢包或错误分类的数据包影响的类。

实施QoS是一项繁重的任务，需要进行认真分析，并且不在本文档的讨论范围之内，但强烈建议考虑此项，以便优先处理对时间敏感的应用程序，并解决或防止许多延迟和应用程序问题。

其他性能问题

其他情况可能增加您需要检查的慢度、会话重新连接或一般性能不佳，其中一些情况是：

丢弃

与设备处理直接相关的一个问题是丢包，您需要从接口角度检查输入和输出端：

<#root>

```
Router#sh interfaces GigabitEthernet0/0/1
GigabitEthernet0/0/1 is up, line protocol is up
  Hardware is vNIC, address is 0ce0.995d.0000 (bia 0ce0.995d.0000)
```

```
  Internet address is 10.10.1.2/24
  MTU 1500 bytes, BW 1000000 Kbit/sec, DLY 10 usec,
     reliability 255/255, txload 1/255, rxload 1/255
  Encapsulation ARPA, loopback not set
  Keepalive set (10 sec)
  Full Duplex, 1000Mbps, link type is auto, media type is Virtual
  output flow-control is unsupported, input flow-control is unsupported
  ARP type: ARPA, ARP Timeout 04:00:00
  Last input 00:00:19, output 00:08:33, output hang never
  Last clearing of "show interface" counters never


 Input queue: 0/375/6788/0 (size/max/drops/flushes); Total output drops: 18263

  Queueing strategy: fifo
  Output queue: 0/40 (size/max)
  5 minute input rate 114000 bits/sec, 230 packets/sec
  5 minute output rate 0 bits/sec, 0 packets/sec
     193099 packets input, 11978115 bytes, 0 no buffer
     Received 0 broadcasts (0 IP multicasts)
     0 runts, 0 giants, 0 throttles


1572 input errors

,

12 CRC

, 0 frame,

1560 overrun

, 0 ignored
     0 watchdog, 0 multicast, 0 pause input
     142 packets output, 11822 bytes, 0 underruns
     Output 0 broadcasts (0 IP multicasts)
     0 output errors, 0 collisions, 0 interface resets
     23 unknown protocol drops
     0 babbles, 0 late collision, 0 deferred
     0 lost carrier, 0 no carrier, 0 pause output
     0 output buffer failures, 0 output buffers swapped out
Router#
```

在输入端，您有：

- 输入队列丢弃：每个接口都拥有一个输入队列（这是一个可以修改的软件缓冲区），将传入数据包置于等待路由处理器(RP)处理的位置。 如果放置在输入队列上的传入数据包速率超过RP处理数据包的速率，则可以增加丢弃次数。但是，请注意，仅放置控制数据包和“For us”流量，因此，如果在通过流量时发现延迟，即使您有间歇性丢弃，这肯定不是原因。

- 溢出：当由于输入速率超过接收方处理数据的能力，接收方硬件无法将接收的数据包交给硬件缓冲区时，就会发生这种情况。此数字表示路由器的速率和性能有问题，仅捕获此接口的流量并查找流量峰值。常见的解决方法是启用流量控制，但这可能会增加延迟数据包。这也可能是存在瓶颈和超订用的证据。
- CRC：发生的原因在于物理问题，请检查电缆、端口和SFP是否正确连接以及是否正常工作。

在输出端，您有：

- 输出队列丢弃：每个接口都拥有一个输出队列，其中放置了要在接口上发送的传出数据包。有时，RP放置在输出队列上的传出数据包的速率超过了接口发送数据包的速率。如果没有QoS，则这会导致性能问题和延迟问题，否则，由于应用了某些策略，您可能会增加此数量，并建议检查或实施QoS配置，以保护和确保预期流量或关键流量。

最后，QFP上的丢弃与可能导致延迟的高处理直接相关，请通过show platform hardware qfp active statistics drop进行检查：

<#root>

```
Router#
show platform hardware qfp active statistics drop
Last clearing of QFP drops statistics : never

-------------------------------------------------------------------------
Global Drop Stats                          Packets                 Octets
-------------------------------------------------------------------------
Disabled                                         2                    646
Ipv4NoAdj                                   108171                6706602
Ipv6NoRoute                                     10                    560
```

原因取决于代码，如果此时丢弃了受延迟影响的流量，则FIA跟踪有助于证实或丢弃。

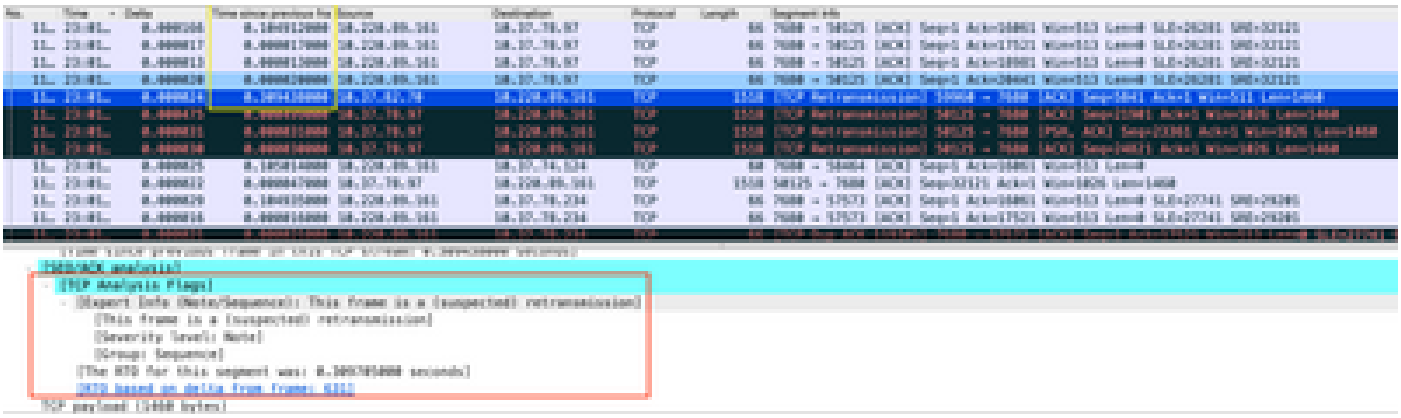
TCP 重新传输

TCP重新传输是症状，或者可能是由于底层问题（如丢包）造成的结果。该问题会导致应用速度缓慢和性能下降。

传输控制协议(TCP)使用重传计时器来确保数据在远程数据接收器没有反馈的情况下传输。此计时器的持续时间称为RTO（重新传输超时）。当重新传输计时器到期时，发送方重新传输尚未被TCP接收方确认的最早数据段，并且RTO增加。

有些重新传输不能完全消除，如果它们很小，则不能反映问题。但是，您可以推断，看到更多重新传输，TCP会话出现更多延迟，需要加以解决。

Wireshark中分析的数据包捕获可以证实此问题，作为下一个示例：



TCP会话捕获

如果存在重新传输，请在路由器入口和出口方向使用相同的捕获方法检查发送和接收的所有数据包。当然，在每一跳上执行此操作可能意味着巨大的努力，因此，需要对TCP的捕获进行详细的分析，查看TTL、同一TCP流上先前帧的时间，以了解延迟或缺乏响应来指导故障排除。

超订用和瓶颈

当所需的资源（带宽）大于实际可用资源时，会发生超订用。上一节中已经介绍了用于识别路由器上是否存在此问题的命令。

因此，当带宽或硬件容量不足导致流量速度减慢时，可能会出现瓶颈。必须确定这种情况是短期发生还是长期存在，以适用解决方案。

没有解决此问题的具体建议，但其中一些选项是平衡流向不同平台的流量、对网络进行分段或者根据当前需求和未来增长分析升级到更强大的设备。

相关信息

- IP SLA ICMP回应操作
- 内存故障排除
- 使用Cisco IOS-XE数据路径数据包跟踪功能进行故障排除
- 排除ASR 1000系列服务路由器上的丢包故障。
- Qos相关信息
- 路由器上的QoS配置
- 思科技术支持和下载

关于此翻译
思科采用人工翻译与机器翻译相结合的方式将此文档翻译成不同语言，希望全球的用户都能通过各自的语言得到支持性的内容。

请注意：即使是最好的机器翻译，其准确度也不及专业翻译人员的水平。

Cisco Systems, Inc. 对于翻译的准确性不承担任何责任，并建议您总是参考英文原始文档（已提供链接）。