

BGP邻居抖动与MTU故障排除技术说明

目录

[简介](#)

[先决条件](#)

[问题](#)

[解决方案](#)

简介

本文档介绍如何确定内部或外部边界网关协议(BGP)邻居抖动是由最大传输单元(MTU)问题引起的。

先决条件

在完成本文档中的步骤之前，请确保在两台BGP路由器上完成以下任务：

- 检查BGP配置。
- 验证BGP邻居是否可通过互联网控制消息协议(ICMP)到达，并且未观察到丢弃。
- 验证用于对等BGP的已连接接口未超订用，并且没有任何输入/输出丢弃或错误。
- 检查CPU和内存利用率。

问题

BGP邻居形式；但是，在前缀交换时，BGP状态丢弃和日志会生成缺少BGP hello keepalive或其他对等体终止会话。

完成以下步骤以确定MTU是否导致BGP邻居摆动：

1. 使用以下命令检查哪个邻居受到影响，以及两台BGP路由器上的已连接接口。如果对等地址是环回地址，请检查可通过其访问环回的已连接接口。此外，检查两台对等路由器上的BGP OutQ。一致的非零OutQ强烈表明，由于路径中的MTU问题，更新未到达对等体。

```
Router#show ip bgp summ | in InQ|10.10.10.2
Neighbor      V   AS MsgRcvd MsgSent  TblVer  InQ  OutQ Up/Down  State/PfxRcd
10.10.10.2    4   3     64     62      3     0    0  00:00:3  2
```

```
Router#show ip route 10.10.10.2
Routing entry for 10.10.10.0/24
  Known via "connected", distance 0, metric 0 (connected, via interface)
  Routing Descriptor Blocks:
    * directly connected, via GigabitEthernet1/0
      Route metric is 0, traffic share count is 1
```

2. 检查两端的接口MTU:

```
Router#show ip int g1/0 | i MTU
```

```
MTU is 1500 bytes
```

```
Router#
```

3. 确认TCP同意的两个BGP扬声器的最大数据段：

```
Router#show ip bgp neigh 20.20.20.2 | inc segment
```

```
Datagrams (max data segment is 1460 bytes):
```

```
Router#
```

在上例中，1460是正确的，因为TCP报头分配了20个字节，IP报头分配了20个字节。

4. 确认是否启用了BGP使用的path-mtu:

```
Router#show ip bgp neigh 10.10.10.2 | in tcp
```

```
Transport(tcp) path-mtu-discovery is enabled
```

```
Router#
```

5. 对设置了最大接口MTU和DF（不分段）位的BGP对等体执行ping操作：

```
Router#ping 10.10.10.2 size 1500 df
```

```
Type escape sequence to abort.
```

```
Sending 5, 1500-byte ICMP Echos to 10.10.10.2, timeout is 2 seconds:
```

```
Packet sent with the DF bit set
```

```
.....
```

```
Success rate is 0 percent (0/5)
```

6. 减小ICMP大小值，以确定可使用的最大MTU大小：

```
ping 10.10.10.2 size 1300 df
```

解决方案

以下是一些可能的原因：

- 两台路由器上的接口MTU不匹配。
- 两台路由器上的接口MTU匹配，但形成BGP会话的第2层域不匹配。
- 路径MTU发现确定了TCP BGP会话的不正确最大数据分配。
- BGP路径最大传输单元发现(PMTUD)可能因PMTUD ICMP数据包被阻止（防火墙或ACL）而失败

以下是解决MTU问题的可能方法：

1. 两台路由器上的接口MTU应相同；运行show ip int | in MTU命令，以检查当前MTU设置。
2. 如果两台路由器上的接口MTU正确（例如，1500），但DF位设置的ping测试不超过1300，则形成受影响BGP会话的第2层域可能包含不一致的MTU配置。检查每个第2层接口MTU。更正第2层接口MTU以解决此问题。
3. 如果无法检查/更改第2层域，可以将ip tcp mss 全局命令设置为较小的值，如1000，这将强制所有本地发起的TCP最大数据段会话（包括BGP）都设置为1000。有关此命令的详细信息，请参阅《Cisco IOS IP应用服务命令参考》的[ip tcp mss](#)部分。

此外，您可以使用ip tcp adjust-mss命令进一步排除故障；此命令在接口级别配置，并影响所有TCP会话。有关此命令的详细信息，请参阅《Cisco IOS IP应用服务命令参考》的[ip tcp adjust-mss](#)部分。

4. (可选)BGP路径最大传输单元发现(PMTUD)可能无法生成正确的最大数据大小。您可以全局禁用它，也可以按邻居禁用它，以确认这是否是原因。禁用BGP PMTUD时，BGP最大分段大小(MSS)默认为536，如RFC 879中[所定义](#)。

有关如何禁用PMTUD的信息，请参阅[Cisco IOS BGP配置指南的配置BGP对TCP路径MTU每会话发现的支持部分](#)。

有关PMTUD的详细信息，请参阅[什么是PMTUD?](#)