某局点VXLAN组网流量突增问题

VxLAN 云博文 2019-10-10 发表

组网及说明

现场共三个区域网络，服务器区核心连接办公区核心，办公核心连接互联网防火墙，互联网防火墙过WAF到DMZ服务器区与DMZ区共用一套控制器，是一套数据中心强控的解决方案，底层OSPF保障各leaf节点IP互通，上层用vxlan构建二层网络。服务器区网关在服务器核心，DMZ区网关在DMZ核心。办公网为DR2000的方案，为弱控，采用分布式网关。

问题描述

现场下午3点半到4点间出现核心设备上联口流量突然打满的异常情况导致办公区内网30分钟内无法上外网，经内部排查定位为服务器区一台S6800接入流量过大，但没有查到该S6800上哪个接口流量过高时异常就消失了，客户在监控服务器上看到出现异常时该S6800交换机上loopback 1接口地址192.168.136.26与DMZ区接入的S68设备有大量报文传输，并在DMZ区核心上联接口抓取到了异常流量报文。客户想确认一下设备在3点半到4点间是否有产生异常日志或是否可通过后续收集的信息查询到异常流量端口，并想了解一下在DMZ区核心抓取的报文中是否有异常或者是VXLAN封装下的什么类型的数据报文。

问题1：流量突增的原因。

问题2：抓包看报文源IP为何都是一台S6800的Loopback地址。

过程分析

1、查看设备诊断信息，设备本身没问题，接口历史峰值记录中，in方向进入S6800的流量，接口被打满的有TEN-1/0/45、TEN-1/0/46、TEN-1/0/47、TEN-2/0/10、TEN-2/0/12这5个接口。

2、其中TEN-1/0/45、TEN-1/0/46、TEN-1/0/47是堆叠链路，这三个接口流量被占满是跨框的转发的流量。所以异常流量来源应该是TEN-2/0/10、TEN-2/0/12 这两个接口。

Ten-GigabitEthernet2/0/10
Current state: DOWN
Line protocol state: DOWN
Last link flapping: 1 hours 27 minutes 39 seconds
Last clearing of counters: 08:01:46 Mon 07/09/2018
Peak input rate: 1043568650 bytes/sec, at 2019-08-29 07:45:24
Peak output rate: 1053362658 bytes/sec, at 2019-08-29 07:45:24

Ten-GigabitEthernet2/0/12
Current state: DOWN
Line protocol state: DOWN
Last link flapping: 1 hours 27 minutes 39 seconds
Last clearing of counters: 08:01:46 Mon 07/09/2018
Peak input rate: 1053362731 bytes/sec, at 2019-08-29 07:45:24
Peak output rate: 1064177717 bytes/sec, at 2019-08-29 07:47:11

3、目前这两个接口是down的状态，诊断信息的时间是09:28:58.666，接口是收诊断时间1个小时27分钟之前down的，和故障恢复的时间点吻合。

4、设备out方向峰值高峰期的接口很多，都是由TEN-2/0/10、TEN-2/0/12这两个接口发出的广播报文被透传到了其他接口和上行接口。

5、问题1排查结果：这两个接口故障期间直接互联形成了环路导致大量流量在VXLAN隧道中泛洪，关闭这两个接口后故障消失。

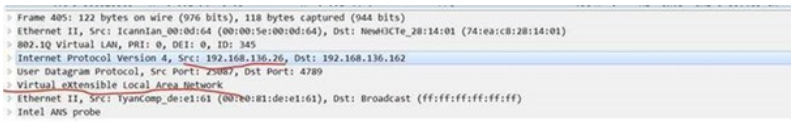
6、现场一线疑惑，这个设备全局开启了STP，两个接口下都配置了边缘端口，形成环路的时候，看接口的STP状态已经变为了Block状态，为何和存在泛洪的流量?

问题3：接口STP被阻塞了，为何还存在环路报文。

Port Ten-GigabitEthernet2/0/12
Role change : DESI->BACK
Time : 2019/08/29 07:34:30
Port priority : 32768.38ad-be97-d284 0 32768.38ad-be97-d284 0
32768.38ad-be97-d284 128.171 128.173
Designated priority : 32768.38ad-be97-d284 0 32768.38ad-be97-d284 0
32768.38ad-be97-d284 128.173 128.173

7、经确认问题3是由于：设备上STP的阻塞是基于PORT和VLAN进行的，接口被阻塞的时候，只能阻塞vlan的流量，对于VXLAN的流量无法阻塞，因此，这两个AC接口的泛洪流量依旧可以正常转发泛洪。现场接口下随然配置了边缘端口，但是没有配置BPDU保护，当出现现场这种情况的时候没法自动关闭其中一个接口。

#
interface Ten-GigabitEthernet2/0/10
 port link-mode bridge
 port link-type trunk
 port trunk permit vlan 1 1000
 stp edged-port
 vtep access port
 #
 service-instance 1000
  encapsulation s-vid 1000
  xconnect vsi SDN_VSI_1000
#
8、对于VXLAN组网中，环路导致泛洪流量的防护建议有如下：
A、基于EVPN的VXLAN组网下，建议关闭VSI的未知广播组播等泛洪。本例中，VXLAN是手工建立的，无法关闭。
B、AC接口可以开启边缘端口加BPDU保护，当出现环路的时候BPDU保护可以直接将环路端口down掉，防止流量泛洪。
9、**问题2看抓包的报文：远端收到的流量通过vxlan泛洪过来的，外层的目的ip是vxlan 隧道的源地址也就是6800的loopback地址，S6800 AC口收到泛洪流量后正常做vxlan封装泛洪出去，设备的正常实现。**
interface Tunnel257 mode vxlan
source 192.168.136.26
destination 192.168.136.143



解决方法

1、设备上STP的阻塞是基于PORT和VLAN进行的，接口被阻塞的时候，只能阻塞端口上VLAN的流量，对于VXLAN的流量无法阻塞。
2、基于EVPN的VXLAN组网下，建议关闭VSI的未知广播组播等泛洪。本例中，VXLAN是手工建立的，无法关闭。
3、AC接口可以开启边缘端口加BPDU保护，当出现环路的时候BPDU保护可以直接将环路端口DOWN掉，防止流量泛洪。