F1080 PING大包丢包问题排查思路及解决方法

会话 虚拟分片重组 刘嘉炜 2018-09-27 发表

组网及说明

无

问题描述

现网中两台F1080防火墙做堆叠部署，运行中发现过防火墙PING 1474以上的数据包出现丢包问题，但是PING小包却无丢包。

过程分析

一、设备配置排查

#

interface Reth1

description To H3C交换

ip address 172.15.11.17 255.255.255.252

member interface Ten-GigabitEthernet1/0/25 priority 255

member interface Ten-GigabitEthernet2/0/25 priority 50

#

interface Reth2

description To 华为交换

ip address 172.15.11.22 255.255.255.252

member interface Ten-GigabitEthernet1/0/24 priority 255

member interface Ten-GigabitEthernet2/0/24 priority 50

#

redundancy group zhengtiqiehuan

member interface Reth1

member interface Reth2

node 1

 bind slot 1

 priority 100

 track 1 interface Ten-GigabitEthernet1/0/24

 track 2 interface Ten-GigabitEthernet1/0/25

node 2

 bind slot 2

 priority 50

 track 3 interface Ten-GigabitEthernet2/0/24

 track 4 interface Ten-GigabitEthernet2/0/25

#

session synchronization enable

配置排查安全域、冗余组、会话均没有发现问题，但是在排查到Reth1接口时，发现**入方向存在大量丢包。**

Reth1

Current state: UP

Line protocol state: UP

Description: To NHX_YJ_F5_SW_S10510-2

Bandwidth: 10000000 kbps

Maximum transmission unit: 1500

Internet address: 172.15.11.17/30 (primary)

IP packet frame type: Ethernet II, hardware address: 38ad-8ed3-d138

IPv6 packet frame type: Ethernet II, hardware address: 38ad-8ed3-d138

Physical: Reth, baudrate: 10000000 kbps

Last clearing of counters: Never

Last 300 seconds input rate: 2362274 bytes/sec, 18898192 bits/sec, 9156 packets/sec

Last 300 seconds output rate: 4759888 bytes/sec, 38079104 bits/sec, 9962 packets/sec

Input: 19747326138 packets, 7782643517525 bytes, **7569606 drops**

Output: 23322478643 packets, 17903881878502 bytes, 0 drops

初步怀疑可能与两端设备的报文分片有关，随机排查交换机、防火墙互联接口状态。

防火墙：

Ten-GigabitEthernet1/0/25

Current state: UP

Line protocol state: UP
Description: Ten-GigabitEthernet1/0/25 Interface
Bandwidth: 10000000 kbps
Maximum transmission unit: 1500
Internet protocol processing: Disabled
交换机:
Vlan-interface51
Current state: UP
Line protocol state: UP
Description: to SHX_FW_H3C_F1080
Bandwidth: 10000000 kbps
Maximum transmission unit: 1500
Internet address: 172.15.11.18/30 (primary)
两端接口MTU等信息完全一致，所以此问题应该不是出在报文分片上。
二、抓包分析
在电脑丢包时进行抓包，发现很多丢包都是因为数据分片无法重组导致的丢包。Fragment reassembly time exceeded表示这个包的发送方之前收到了一些分片，但是由于某些原因迟迟无法组装起来。注:可能看到Time-to-live有些人会想到TTL超时，进而错误分析出路由环路的结论，一定要注意括号里面的内容。

4404	142.634614	10.100.0.254	168.168.128.18	ICMP	Echo (ping) reply
4418	143.312280	168.168.128.18	172.15.11.17	ICMP	Time-to-live exceeded (Fragment reassembly time exceeded)
4420	143.436441	168.168.128.18	172.15.11.17	ICMP	Echo (ping) request
4422	143.439785	172.15.11.17	168.168.128.18	ICMP	Echo (ping) reply
4423	143.635414	168.168.128.18	10.100.0.254	ICMP	Echo (ping) request
4424	143.641355	10.100.0.254	168.168.128.18	ICMP	Echo (ping) reply
4439	144.440493	168.168.128.18	172.15.11.17	ICMP	Echo (ping) request
4442	144.639454	168.168.128.18	10.100.0.254	ICMP	Echo (ping) request
4443	144.641928	10.100.0.254	168.168.128.18	ICMP	Echo (ping) reply
4455	145.643490	168.168.128.18	10.100.0.254	ICMP	Echo (ping) request
4456	145.647432	10.100.0.254	168.168.128.18	ICMP	Echo (ping) reply
4473	146.647529	168.168.128.18	10.100.0.254	ICMP	Echo (ping) request
4474	146.649887	10.100.0.254	168.168.128.18	ICMP	Echo (ping) reply
4484	147.651524	168.168.128.18	10.100.0.254	ICMP	Echo (ping) request
4485	147.660961	10.100.0.254	168.168.128.18	ICMP	Echo (ping) reply
4496	148.655608	168.168.128.18	10.100.0.254	ICMP	Echo (ping) request
4497	148.659066	10.100.0.254	168.168.128.18	ICMP	Echo (ping) reply
4508	149.312630	168.168.128.18	172.15.11.17	ICMP	Echo (ping) request
4510	149.318047	172.15.11.17	168.168.128.18	ICMP	Echo (ping) reply
4511	149.659639	168.168.128.18	10.100.0.254	ICMP	Echo (ping) request
4512	149.663310	10.100.0.254	168.168.128.18	ICMP	Echo (ping) reply
4521	150.312505	168.168.128.18	172.15.11.17	ICMP	Time-to-live exceeded (Fragment reassembly time exceeded)

这里想到V7防火墙是多核多线程的转发原理，也就是平时在设备上看到2-15都是属于转发核参与数据转发的。Comware V5\V7早期版本是逐包转发的，也就是数据包
备分片后可能通过不同VCPU处理，进而导致数据过防火墙后无法重组，所以将防火墙改为逐流模式就可以解决此问题。

===============display process cpu slot 1===============
CPU utilization in 5 secs: 7.9%; 1 min: 4.2%; 5 mins: 5.3%

JID	5Sec	1Min	5Min	Name	
171	0.0%	0.0%	0.0%	[kdrvcp0]	
172	0.0%	0.0%	0.0%	[kdrvcp1]	\\控制核
173	0.2%	0.0%	0.0%	[kdrvdp2]	\\转发核
174	0.0%	0.0%	0.0%	[kdrvdp3]	
175	0.0%	0.1%	0.3%	[kdrvdp4]	
176	0.1%	0.0%	0.0%	[kdrvdp5]	
177	0.0%	0.0%	0.1%	[kdrvdp6]	
178	0.2%	0.0%	0.1%	[kdrvdp7]	
179	0.0%	0.0%	0.0%	[kdrvdp8]	
180	0.2%	0.0%	0.1%	[kdrvdp9]	
181	0.4%	0.0%	0.2%	[kdrvdp10]	
182	0.1%	0.0%	0.1%	[kdrvdp11]	
183	0.0%	0.0%	0.0%	[kdrvdp12]	
184	0.1%	0.2%	0.0%	[kdrvdp13]	
185	0.1%	0.0%	0.0%	[kdrvdp14]	
186	0.0%	0.1%	0.1%	[kdrvdp15]	

解决方法

防火墙的逐流模式也称为五元组模式，即匹配报文五元祖后上送同一个VCPU进行处理，从而规避数据分片后报文经过多个VCPU导致重组异常问题。
修改逐流模式的命令命令:
<H3C> system-view
[H3C] forwarding policy per-flow