M9K设备单核丢包典型案例分析

会话 孔凡安 2023-08-18 发表

组网及说明

不涉及

告警信息

不涉及

问题描述

现场不定时有如下单核丢包日志：

> %Aug 6 11:17:07:280 2023 XX-D1-F13&E13-6U-PTFW-M9010-01&02 SIB/6/SIB_CORE_ATTACK_DROP: -Chassis=1-Slot=6.1; Dropped 121535 packets because of core attack.
> %Aug 6 11:12:07:273 2023 XX-D1-F13&E13-6U-PTFW-M9010-01&02 SIB/6/SIB_CORE_ATTACK_DROP: -Chassis=1-Slot=6.1; Dropped 129312 packets because of core attack.
> %Aug 6 11:07:07:266 2023 XX-D1-F13&E13-6U-PTFW-M9010-01&02 SIB/6/SIB_CORE_ATTACK_DROP: -Chassis=1-Slot=6.1; Dropped 169859 packets because of core attack.
> %Aug 6 11:02:07:259 2023 XX-D1-F13&E13-6U-PTFW-M9010-01&02 SIB/6/SIB_CORE_ATTACK_DROP: -Chassis=1-Slot=6.1; Dropped 218690 packets because of core attack.
> %Aug 6 10:57:07:253 2023 XX-D1-F13&E13-6U-PTFW-M9010-01&02 SIB/6/SIB_CORE_ATTACK_DROP: -Chassis=1-Slot=6.1; Dropped 310252 packets because of core attack.

过程分析

出现上述日志之后，可以通过命令display attack-defense cpu-core flow info chassis X slot Y cpu 1 （X=框号，Y=槽位号）查看对应板卡处理的攻击流量，详细介绍可见链接：
https://zhiliao.h3c.com/Theme/details/217682

> 定位到流量之后，发现是NAS挂载相关的流量：
>
> TimeStamp: 2023-08-06 07:19:03
> CPUID: 17
> SMAC: 74:3a:20:2b:fc:01 DMAC: a2:c9:70:01:00:02
> VlanID: 2000 Interface: Ten-GigabitEthernet1/3/2/5
> SIP: 172.18.18.50 DIP: 172.20.23.123
> Pro: 6
> SPort: 2049 DPort: 1016
> CPUUsage: 69% IfIsolate: false

进一步定位为何该流量把单核打高，该设备为单逻辑板卡，如果流量全部由逻辑处理理论上不会存在CPU高的情况。唯一的可能就是逻辑上没有会话。
通过查看逻辑会话果然没有（read session burst data回显部分全零可以理解为逻辑上没有会话）。

```
[XX-D1-F13&E13-6U-PTFW-M9010-01&02-probe]archer  chassis  1 slot  6 cpu 1 0 0 rd-tbl sess_tbl_id 172.20.23.123 172.18.18.50 1016 2049 6 1 0 2
===>CONSOLE_CMD_READ_TBL
===>input hash key
    ac 14 15 0a ac 12 12 32 00 63 08 01 06 00 00 01
    00 00 00 02
===>hash crc32          :0xa04cd797
===>fold xor(11 bits)   :0x0000015f
===>hash23              :0x00002bf4
===>ddr mod             :0x00000017
===>addr                :0x004cd797
===>read session hash burst data
    55 40 84 00 00 00 00 00 00 00 00 00 00 00 00 00
    00 00 00 00 ef 0f 00 00 00 00 00 00 00 00 00 00
===>read session burst data
    00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
    00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
    00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
    00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
    00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
    00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
    00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
    00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
    00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
    00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
```

进一步检查配置，发现设备开启了DPI功能，因此执行inspect bypass并重置会话之后查看报文是否逻辑处理。（流量无法走硬件快转的场景可以参考技术公告：
https://zhiliao.h3c.com/Theme/details/182583）
遗憾的是，会话重新建立之后，逻辑上还是没有会话。这一下整的就很尴尬，只能继续查看设备配置并进行分析。

解决方法

后续经过定位发现设备开启了会话引流，但是没有配置会话引流支持逻辑快转导致流量全部由CPU处理。增加如下命令后恢复正常。

> session flow-redirect hardware-fast-forwarding命令用来开启会话引流的硬件快速转发功能。
> undo session flow-redirect hardware-fast-forwarding命令用来关闭会话引流的硬件快速转发功能。
> 【命令】
> session flow-redirect hardware-fast-forwarding
> undo session flow-redirect hardware-fast-forwarding
> 【缺省情况】
> 会话引流的硬件快速转发功能处于关闭状态。
> 【视图】
> 系统视图
> 【缺省用户角色】
> network-admin
> mdc-admin
> 【使用指导】
> 非缺省vSystem不支持本命令。
> 开启此功能后，设备会将首报文触发创建的会话表项下发硬件芯片，后续报文将直接匹配硬件芯片中的会话表项进行转发，从而提高设备对报文的转发速度。
> 此功能仅在会话引流功能和硬件快速转发功能均处于开启状态时才能生效。有关硬件快速转发功能的详细介绍，请参见“三层技术-IP业务配置指导”中的“快速转发”。
> 当需要定位硬件芯片是否存在故障时，可以关闭此功能。
> 【举例】
> # 开启会话引流的硬件快速转发功能。
> <Sysname> system-view
> [Sysname] session flow-redirect hardware-fast-forwarding