

知 CR16K-F插上非法光模块后直连ping延时大案例

光模块 付睿 2015-10-26 发表

某局点新上线CR16K-F一台，经过CR16K-F的业务转发正常(Pc ping防火墙延时正常)，但Pc和16k互ping，延时大，时延不固定，约为300-500ms。

如下图，CR16K-F的0槽位使用CSPEX-1204板卡，2槽位使用CSPC-XP8LB板卡。G0/1/9口连接二层交换机，下联PC，G2/0/8口直连防火墙。



从问题现象来看，经过设备转发正常，芯片故障的可能性较小。查看设备诊断日志，发现有slot 0大量I2C bus告警：

```
%@20536^Aug 30 17:30:10:963 2015 HX-XiAn-CR16 NULL/3/UNKNOWN: -Slot=0; 0x00000000 [685]: I2C bus 1 write failed, retry resume! i2c trace: 1, error code: 1, dev: 0x51, reg: 0x7b
```

//光模块I2C总线提示写错误

```
%@20537^Aug 30 17:30:10:966 2015 HX-XiAn-CR16 DSYSM/3/NP: -Slot=0; 0x0F709006 [2257]: SubCardI2CWrite(1,81,6,123,4) failed!
```

```
%@20538^Aug 30 17:30:11:007 2015 HX-XiAn-CR16 PORT/3/API: -Slot=0; 0x0F9D3000 [3961]: PORT Check Fake Opt, Phony Check, p1=7, p2=1073807361!
```

//Slot 0提示接口校验，发现光模块是非法光模块

查看诊断信息，通过display transceiver diagnosis interface发现Slot 0上确实有非法光模块。其中，g0/1/2口和g0/1/7口显示从模块读取信息失败，而其他非法光模块接口直接提示设备不支持读取该模块信息：

GigabitEthernet0/1/1 transceiver diagnostic information:

The transceiver does not support this function.

GigabitEthernet0/1/2 transceiver diagnostic information:

Reading information from the transceiver failed.

GigabitEthernet0/1/3 transceiver diagnostic information:

The transceiver does not support this function.

GigabitEthernet0/1/4 transceiver diagnostic information:

The transceiver does not support this function.

GigabitEthernet0/1/6 transceiver diagnostic information:

The transceiver does not support this function.

GigabitEthernet0/1/7 transceiver diagnostic information:

Reading information from the transceiver failed.

GigabitEthernet0/1/8 transceiver diagnostic information:

The transceiver does not support this function.

GigabitEthernet0/1/9 transceiver diagnostic information:

Current diagnostic parameters:

Temp.(°C)	Voltage(V)	Bias(mA)	RX power(dBm)	TX power(dBm)
-----------	------------	----------	---------------	---------------

28	3.29	9.57	-5.34	-5.42
----	------	------	-------	-------

Alarm thresholds:

Temp.(°C)	Voltage(V)	Bias(mA)	RX power(dBm)	TX power(dBm)
-----------	------------	----------	---------------	---------------

High	73	3.60	42.49	-3.00	6.00
Low	-3	3.00	1.00	-22.08	-12.50

此时，这些插了非法光模块的接口物理协议都up，看起来转发正常，并且连接PC的g0/1/9口连接的是合法光模块。是否可能是非法光模块导致的问题呢？我们尝试插拔光模块，发现：一旦将设备提示读取光模块信息失败的g0/1/2和g0/1/7口的光模块拔出，PC ping设备延时就恢复正常。插回这两个光模块，故障复现。

这是为什么呢？首先，我们要知道在ping报文的处理流程中，ping报文没有任何优先保证，相比较其他业务流量，ping报文的处理级别是最低的。这是因为，ping报文只是网络管理员和用户的探测报文，甚至可能是恶意攻击报文，设备出于保护设备尽量免于被攻击默认就把ping流量的重要程度和优先级设置为最低，所以ping流量的时延会随着任务调用的不同而出现时延波动的情况。只有这样，更为重要的业务流量尤其是实时业务流量的时延才能得到更好的保证。

现场采用的是光口，有些缺省的后台任务在运行，比如端口监控任务，这个任务会定时对所有在位光模块、PHY通道进行技术指标监控，由于轮询的端口多，内容相对较多，会占用一定比例的CPU调用周期。这个进程是硬件核心监控进程，优先级高，调度实时性要求高。当类似高优先级进程占用CPU周期时，Ping进程的报文处理就会滞后。

由于现场采用了非法光模块，设备无法成功读取光模块信息，因此会不停对光模块端口进行轮询，造成，Ping进程的报文处理一直滞后，表现为直连延时大。而转发报文不用上送cpu，因此不受影响。

更换合法光模块后，故障恢复。

从本则案例我们可以看到，在今后的局点应尽量使用从H3C出产的正版光模块，使用非法光模块或者伪造光模块，虽然可能接口也能up，但却存在潜在隐患，应尽量避免。