

终结者本体有线业务丢包问题处理经验案例

终结者 VLAN 陈铮 2017-12-27 发表

客户反馈有线业务高峰期丢包，重启本体后业务恢复，两周后现象又能复现。

一、

1. 首先排查链路问题，主要通过流量统计确认故障发生时丢包丢在WT1020上；
2. 这个问题目前看来只有这一个局点出现过，而且和设备运行的时间有关，通过对比重启和未重启保留现象的本体发现，有线业务丢包时伴随着cpu某一个转发进程的激增，但是从本体上行口进出方向统计看，丢包和不丢包的两台本体业务流量相差不多；

```
CPU0 states: 78.35% idle, 3.44% user, 17.04% kernel, 1.14% interrupt
CPU1 states: 67.62% idle, 2.49% user, 29.31% kernel, 0.57% interrupt
CPU2 states: 81.20% idle, 1.55% user, 16.86% kernel, 0.38% interrupt
CPU3 states: 94.63% idle, 0.00% user, 5.17% kernel, 0.19% interrupt
CPU4 states: 35.44% idle, 0.00% user, 64.17% kernel, 0.38% interrupt
CPU5 states: 91.01% idle, 0.00% user, 8.41% kernel, 0.57% interrupt
CPU6 states: 89.08% idle, 0.00% user, 10.72% kernel, 0.19% interrupt
CPU7 states: 93.10% idle, 0.00% user, 6.51% kernel, 0.38% interrupt
Memory: 969M total, 644M available, page size 4K
```

JID	PID	PRI	State	FDs	MEM	HH:MM:SS	CPU	Name
89	89	115	R	0	0K	33:08:07	7.58%	[kdrvfd4]
107	107	115	S	0	0K	05:46:43	1.96%	[RipcSendMsg0]
109	109	115	S	0	0K	05:59:17	1.91%	[RipcSendMsg2]
177	177	120	S	62	26740K	05:00:30	1.25%	apmgrlited
108	108	115	S	0	0K	04:57:55	1.04%	[RipcSendMsg1]
90	90	115	R	0	0K	07:29:19	0.94%	[kdrvfd5]
111	111	115	S	0	0K	02:32:01	0.73%	[RipcSendMsg4]
127	127	120	S	14	10880K	03:36:32	0.73%	diagd
110	110	115	S	0	0K	02:47:38	0.68%	[RipcSendMsg3]
91	91	115	R	0	0K	05:22:11	0.59%	[kdrvfd6]
92	92	115	R	0	0K	07:06:25	0.59%	[kdrvfd7]
112	112	115	S	0	0K	03:00:39	0.52%	[RipcSendMsg5]

3. 咨询开发了解到本体设备WT1020虽然是cpu软件转发，和路由器不一样的是本体当前只按照三元组(源mac、源ip、端口号)进行匹配分流，由于有线流量全部是走pppoe认证，来自bras设备的有线流量同源，这也就是为什么本体有多个转发进程只有一个进程非常高的原因，开发给出五元组转发临时版本升级个别本体观察；
4. 观察两周之后问题依然能否复现，现象从单一转发进程高变成了所有转发进程都会很高，说明临时版本能够把pppoe分流到不同转发进程中使得本体对pppoe转发能力有了成倍提升，但是显然pppoe分流不是有线丢包的根因；
5. 设备上似乎无从下手，于是联系代理商在本体上联交换机和下联pc分别抓包，看看本体对数据包做了哪些特殊处理；
6. 如下报文是在学生PC上抓的，可以看出抵达PC的有很多不属于自己的mac，证明了设备是洪泛转发。反查本体mac表项发现学习不全，那么对于学不到mac的接口做未知单播帧转发β-----经过开发确认这部分流量全部按照广播处理，这个数量级应该远远超过我们在接口上统计的广播流量。

No.	Time	Source	Destination	Protocol	Length	Info
483	0.26006800	223.146.243.212	10.64.10.28	ICMP	102	Destination unreachable (Port unreachable)
515	0.27804800	223.146.243.212	10.64.10.28	ICMP	102	Destination unreachable (Port unreachable)
542	0.29546200	223.146.243.212	10.64.10.28	ICMP	102	Destination unreachable (Port unreachable)
590	0.31913300	223.146.243.212	10.64.10.28	ICMP	102	Destination unreachable (Port unreachable)
854	0.47172700	10.64.10.57	10.128.0.3	ICMP	82	Echo (ping) request id=0x0001, seq=15
1010	0.54999900	10.128.0.3	10.64.10.57	ICMP	82	Echo (ping) reply id=0x0001, seq=15
2825	1.48456900	10.64.10.57	10.128.0.3	ICMP	82	Echo (ping) request id=0x0001, seq=15
2989	1.57125600	10.128.0.3	10.64.10.57	ICMP	82	Echo (ping) reply id=0x0001, seq=15
4386	2.32675500	223.146.243.212	10.64.10.28	ICMP	102	Destination unreachable (Port unreachable)
4396	2.34176700	223.146.243.212	10.64.10.28	ICMP	102	Destination unreachable (Port unreachable)
4433	2.36222100	223.146.243.212	10.64.10.28	ICMP	102	Destination unreachable (Port unreachable)
4554	2.43690900	223.146.243.212	10.64.10.28	ICMP	102	Destination unreachable (Port unreachable)
4646	2.48280000	10.64.10.57	10.128.0.3	ICMP	82	Echo (ping) request id=0x0001, seq=15
4846	2.58899700	10.128.0.3	10.64.10.57	ICMP	82	Echo (ping) reply id=0x0001, seq=15
6196	3.48358200	10.64.10.57	10.128.0.3	ICMP	82	Echo (ping) request id=0x0001, seq=15

7. 结合上述分析，我认为前期远程定位排除方向不准确，这个问题的根源应该在mac地址学习到上，由于本体WT1020使用cpu软转，大量复制报文对性能消耗较大，从而导致有线业务丢包。另外，结合前面几次定位过程现象要经过一段时间才会出现，所以推测mac地址学习不全的情况可能也需要时间积累。最后了解现场组网情况后了解到，问题本体上行口还级联了普通AP（该型号其实不允许级联），当sta关联普通AP时，本体从GE口学习到sta的mac地址，当sta漫游到分体时，本体刷新mac-add

，目的端口变成wlan-dbss，这个过程中软件存在bug，老的表项资源无法释放，从而时间积累到一定程度mac表项规格占满，新的mac无法学习时进行未知单播转发。

建议现场修改级联组网，将级联普通AP的接口配置端口隔离。

本体WT1020不允许级联