# 某局点S5560X-EI几乎所有端口流量被打满问题案例

二层转发 张文宁 2020-02-28 发表

## 组网及说明

/

## 问题描述

现场两台S5560X-EI两台堆叠作为接入设备，核心是S7500E，客户反馈只有该接入设备下的服务器业务卡顿。

查看该接入设备发现很多端口出方向带宽都打满了：

```
行   字符串
1281 Last 300 second output: 80225 packets/sec 121358857 bytes/sec 98%
1325 Last 300 second output: 80225 packets/sec 121358785 bytes/sec 98%
1369 Last 300 second output: 80233 packets/sec 121359387 bytes/sec 98%
1413 Last 300 second output: 80249 packets/sec 121370932 bytes/sec 98%
1457 Last 300 second output: 80249 packets/sec 121369008 bytes/sec 98%
1501 Last 300 second output: 81370 packets/sec 121469524 bytes/sec 98%
1545 Last 300 second output: 0 packets/sec 0 bytes/sec -%
1589 Last 300 second output: 80228 packets/sec 121360559 bytes/sec 98%
1633 Last 300 second output: 0 packets/sec 0 bytes/sec -%
1677 Last 300 second output: 80258 packets/sec 121381725 bytes/sec 98%
1721 Last 300 second output: 0 packets/sec 0 bytes/sec -%
1765 Last 300 second output: 80451 packets/sec 121414641 bytes/sec 98%
1809 Last 300 second output: 80225 packets/sec 121359025 bytes/sec 98%
1853 Last 300 second output: 80225 packets/sec 121358875 bytes/sec 98%
1897 Last 300 second output: 80225 packets/sec 121359013 bytes/sec 98%
1941 Last 300 second output: 80225 packets/sec 121358942 bytes/sec 98%
1985 Last 300 second output: 0 packets/sec 0 bytes/sec -%
2029 Last 300 second output: 80224 packets/sec 121358580 bytes/sec 98%
2073 Last 300 second output: 80225 packets/sec 121359064 bytes/sec 98%
2117 Last 300 second output: 0 packets/sec 0 bytes/sec -%
2161 Last 300 second output: 80224 packets/sec 121358717 bytes/sec 98%
2205 Last 300 second output: 0 packets/sec 0 bytes/sec -%
2249 Last 300 second output: 80224 packets/sec 121358798 bytes/sec 98%
2293 Last 300 second output: 80224 packets/sec 121358644 bytes/sec 98%
2337 Last 300 second output: 80241 packets/sec 121373411 bytes/sec 98%
2381 Last 300 second output: 80241 packets/sec 121373363 bytes/sec 98%
2425 Last 300 second output: 80241 packets/sec 121373393 bytes/sec 98%
2469 Last 300 second output: 80391 packets/sec 121404229 bytes/sec 98%
2513 Last 300 second output: 80373 packets/sec 121406005 bytes/sec 98%
2557 Last 300 second output: 80275 packets/sec 121386980 bytes/sec 98%
2601 Last 300 second output: 0 packets/sec 0 bytes/sec -%
2645 Last 300 second output: 80276 packets/sec 121394452 bytes/sec 98%
2689 Last 300 second output: 0 packets/sec 0 bytes/sec -%
2733 Last 300 second output: 80247 packets/sec 121374547 bytes/sec 98%
2777 Last 300 second output: 0 packets/sec 0 bytes/sec -%
2821 Last 300 second output: 80311 packets/sec 121378600 bytes/sec 98%
```

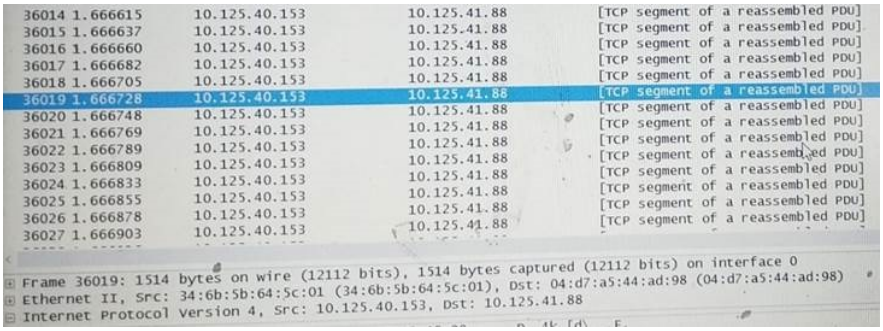
## 过程分析

查看input流量看到，只有一个业务上行口ten2/0/25口inbound方向有9%的带宽，还有个是irf端口1/0/28口有9%的带宽，这两个口都是10G口，所以9%就是千兆端口的90%+。因此怀疑是ten2/0/25上行口收到的流量全部单播泛红了：

```
Ten-GigabitEthernet2/0/25
Current state: UP
Line protocol state: UP
IP packet frame type: Ethernet II, hardware address: 307b-ace1-dc1b
Description: Ten-GigabitEthernet2/0/25 Interface
Bandwidth: 10000000 kbps
Loopback is not set
Media type is optical fiber, Port hardware type is 10G_BASE_SR_SFP
10Gbps-speed mode, full-duplex mode
Link speed type is autonegotiation, link duplex type is autonegotiation
Flow-control is not enabled
Maximum frame length: 10000
Allow jumbo frames to pass
Broadcast max-ratio: 100%
Multicast max-ratio: 100%
Unicast max-ratio: 100%
PVID: 1
MDI type: Automdix
Port link-type: Trunk
 VLAN Passing:   1000, 1041
 VLAN permitted: 1000, 1041
 Trunk port encapsulation: IEEE 802.1q
Port priority: 0
Last link flapping: 11 weeks 4 days 9 hours 14 minutes
Last clearing of counters: Never
 Peak input rate: 127055803 bytes/sec, at 2019-09-04 08:04:58
 Peak output rate: 86467874 bytes/sec, at 2019-08-20 19:01:49
 Last 300 second input: 80094 packets/sec 121627569 bytes/sec 9%
 Last 300 second output: 1341 packets/sec 1422070 bytes/sec 0%
 Input (total):  587829958955 packets, 300874875026319 bytes
  587765685711 unicasts, 10806296 broadcasts, 53466948 multicasts, 0 pauses
 Input (normal):  587829958925 packets, - bytes
  587765685708 unicasts, 10806296 broadcasts, 53466948 multicasts, 0 pauses
 Input:  0 input errors, 0 runts, 0 giants, 0 throttles
  0 CRC, 0 frame, - overruns, 0 aborts
  - ignored, - parity errors
 Output (total): 25771276025 packets, 29141803417940 bytes
  25767216889 unicasts, 51830 broadcasts, 4007306 multicasts, 0 pauses
 Output (normal): 25771276025 packets, - bytes
```

```
1896 Last 300 second input: 0 packets/sec 0 bytes/sec 0%
1940 Last 300 second input: 0 packets/sec 0 bytes/sec 0%
1984 Last 300 second input: 0 packets/sec 0 bytes/sec -%
2028 Last 300 second input: 0 packets/sec 0 bytes/sec 0%
2072 Last 300 second input: 0 packets/sec 0 bytes/sec 0%
2116 Last 300 second input: 0 packets/sec 0 bytes/sec -%
2160 Last 300 second input: 0 packets/sec 0 bytes/sec 0%
2204 Last 300 second input: 0 packets/sec 0 bytes/sec -%
2248 Last 300 second input: 0 packets/sec 0 bytes/sec 0%
2292 Last 300 second input: 0 packets/sec 0 bytes/sec 0%
2336 Last 300 second input: 0 packets/sec 0 bytes/sec 0%
2380 Last 300 second input: 0 packets/sec 2 bytes/sec 0%
2424 Last 300 second input: 0 packets/sec 2 bytes/sec 0%
2468 Last 300 second input: 99 packets/sec 67304 bytes/sec 0%
2512 Last 300 second input: 76 packets/sec 47060 bytes/sec 0%
2556 Last 300 second input: 1087 packets/sec 1177778 bytes/sec 0%
2600 Last 300 second input: 0 packets/sec 0 bytes/sec -%
2644 Last 300 second input: 17 packets/sec 6812 bytes/sec 0%
2688 Last 300 second input: 0 packets/sec 0 bytes/sec -%
2732 Last 300 second input: 15 packets/sec 5031 bytes/sec 0%
2776 Last 300 second input: 0 packets/sec 0 bytes/sec -%
2820 Last 300 second input: 206 packets/sec 184633 bytes/sec 0%
2864 Last 300 second input: 0 packets/sec 0 bytes/sec -%
2908 Last 300 second input: 0 packets/sec 0 bytes/sec -%
2952 Last 300 second input: 0 packets/sec 0 bytes/sec -%
2996 Last 300 second input: 0 packets/sec 0 bytes/sec -%
3040 Last 300 second input: 0 packets/sec 0 bytes/sec -%
3084 Last 300 second input: 0 packets/sec 0 bytes/sec -%
3128 Last 300 second input: 0 packets/sec 0 bytes/sec -%
3172 Last 300 second input: 0 packets/sec 0 bytes/sec -%
3216 Last 300 second input: 0 packets/sec 0 bytes/sec -%
3260 Last 300 second input: 0 packets/sec 0 bytes/sec -%
3304 Last 300 second input: 0 packets/sec 0 bytes/sec -%
3348 Last 300 second input: 0 packets/sec 0 bytes/sec -%
3455 Last 300 second input: 1815 packets/sec 327151 bytes/sec 0%
3498 Last 300 second input: 0 packets/sec 0 bytes/sec -%
3530 Last 300 second input: 75 packets/sec 30543 bytes/sec 0%
3562 Last 300 second input: 80114 packets/sec 122600233 bytes/sec 9%
3606 Last 300 second input: 80094 packets/sec 121627569 bytes/sec 9%
3649 Last 300 second input: 0 packets/sec 0 bytes/sec -%
3681 Last 300 second input: 372 packets/sec 80067 bytes/sec 0%
3713 Last 300 second input: 159 packets/sec 76493 bytes/sec 0%
```

通过现场随意插一个空闲接口放通所有vlan直接接PC抓包发现，确实抓到了大量单播泛洪流量，如下：



| | | | | |
|---|---|---|---|---|
| 36014 | 1.666615 | 10.125.40.153 | 10.125.41.88 | [TCP segment of a reassembled PDU] |
| 36015 | 1.666637 | 10.125.40.153 | 10.125.41.88 | [TCP segment of a reassembled PDU] |
| 36016 | 1.666660 | 10.125.40.153 | 10.125.41.88 | [TCP segment of a reassembled PDU] |
| 36017 | 1.666682 | 10.125.40.153 | 10.125.41.88 | [TCP segment of a reassembled PDU] |
| 36018 | 1.666705 | 10.125.40.153 | 10.125.41.88 | [TCP segment of a reassembled PDU] |
| 36019 | 1.666728 | 10.125.40.153 | 10.125.41.88 | [TCP segment of a reassembled PDU] |
| 36020 | 1.666748 | 10.125.40.153 | 10.125.41.88 | [TCP segment of a reassembled PDU] |
| 36021 | 1.666769 | 10.125.40.153 | 10.125.41.88 | [TCP segment of a reassembled PDU] |
| 36022 | 1.666789 | 10.125.40.153 | 10.125.41.88 | [TCP segment of a reassembled PDU] |
| 36023 | 1.666809 | 10.125.40.153 | 10.125.41.88 | [TCP segment of a reassembled PDU] |
| 36024 | 1.666833 | 10.125.40.153 | 10.125.41.88 | [TCP segment of a reassembled PDU] |
| 36025 | 1.666855 | 10.125.40.153 | 10.125.41.88 | [TCP segment of a reassembled PDU] |
| 36026 | 1.666878 | 10.125.40.153 | 10.125.41.88 | [TCP segment of a reassembled PDU] |
| 36027 | 1.666903 | 10.125.40.153 | 10.125.41.88 | [TCP segment of a reassembled PDU] |

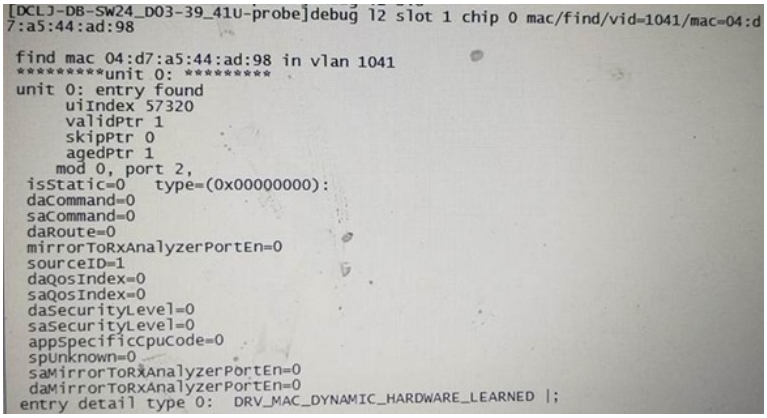
```
Frame 36019: 1514 bytes on wire (12112 bits), 1514 bytes captured (12112 bits) on interface 0
Ethernet II, Src: 34:6b:5b:64:5c:01 (34:6b:5b:64:5c:01), Dst: 04:d7:a5:44:ad:98 (04:d7:a5:44:ad:98)
Internet Protocol Version 4, Src: 10.125.40.153, Dst: 10.125.41.88
```

设备查看该目的mac发现软件表上有：

==========display mac-address==========

| MAC Address | VLAN ID | State | Port/NickName | Aging |
|---|---|---|---|---|
| 04d7-a544-ad98 | 1041 | Learned | GE1/0/3 | Y |

但是查看底层表项发现只有slot 1有，slot 2没有：



```
[DCLJ-DB-SW24_D03-39_41U-probe]debug l2 slot 1 chip 0 mac/find/vid=1041/mac=04:d
7:a5:44:ad:98

find mac 04:d7:a5:44:ad:98 in vlan 1041
*********unit 0: *********
unit 0: entry found
     uiIndex 57320
     validPtr 1
     skipPtr 0
     agedPtr 1
    mod 0, port 2,
 isStatic=0   type=(0x00000000):
 daCommand=0
 saCommand=0
 daRoute=0
 mirrorToRxAnalyzerPortEn=0
 sourceID=1
 daQosIndex=0
 saQosIndex=0
 daSecurityLevel=0
 saSecurityLevel=0
 appSpecificCpuCode=0
 spUnknown=0
 saMirrorToRxAnalyzerPortEn=0
 daMirrorToRxAnalyzerPortEn=0
entry detail type 0:  DRV_MAC_DYNAMIC_HARDWARE_LEARNED |;
```

因为流量都是从ten2/0/25口进来的，而且slot 2底层mac表，所以当做未知单播泛洪了。
设备学习mac有两个途径:1是流量触发芯片学习。 2是软件定时同步到每个芯片（120S）。

经确认现场R1110P06老版本还不支持定时同步mac地址功能，现场设备又配置了本地优先转发，导致服务器上行的流量会优先从Slot1上转发，所以slot2上学不到服务器的mac地址，导致从slot2进来的单播报文，按照未知单播处理而泛洪.

## 解决方法

规避方法：
1.g1/0/3口的服务器发出的流量部分会哈希到slot 2去，触发芯片学习。现场配置后成功学习到mac，泛洪情况立刻消失了.
2.5560X的上行由聚合改为单链路或者5560X与服务器相连单端口也改成聚合口 （上下行都是聚合才是正常的负载分担组网）
解决方法：
升级R1119P12及以后版本彻底解决。