

排除Cisco Nexus 9000上損壞的乙太網資料包故障

目錄

[簡介](#)

[背景資訊](#)

[交換器如何處理封包](#)

[流量通過N9K時，使用標籤的VLAN修改填充](#)

[解決方案](#)

簡介

本檔案介紹如何在Cisco Nexus 9000上排除損壞的乙太網路封包的故障，因為填充資訊已損毀或格式不正確。

背景資訊

無論是否存在任何VLAN標籤，乙太網幀的最小大小都是64位元組。

最小乙太網負載大小為：

- 如果VLAN標籤不存在，則為46位元組。
- 如果存在VLAN標籤，則為42位元組。

您可以驗證以下事實：

- 在Wikipedia上，部分**負載**：https://en.wikipedia.org/wiki/Ethernet_frame
- 在IEEE 802.3標準

(http://people.ee.duke.edu/~mbrooke/EE164.02/Spring_2004/group_2/index_files/8023.pdf)中，第39頁3.1.1節中定義了MAC幀格式（無VLAN），第43頁3.5節定義了標籤MAC幀的元素。

無論是否存在VLAN標頭，乙太網資料包的最小大小都是64位元組。允許伺服器傳送包含VLAN的64位元組資料包，您應該接受並正確處理該資料包。

附註：此行為由Catalyst 4500x正確處理，而不是Nexus 9k。

交換器如何處理封包

步驟1.接收有效的64位元組乙太網幀。

步驟2.刪除幀校驗序列(FCS)，使資料包變長60位元組。

步驟3.刪除VLAN標籤，使資料包變長56位元組。

步驟4.新增填充以使資料包的長度達到60位元組。

步驟5.新增FCS，使資料包的長度為64個位元組。

封包經過直通交換器時，不應修改填充。

流量通過N9K時，使用標籤的VLAN修改填充

資料包不是用零填充，而是用垃圾字元填充，在大多數情況下，這樣做沒有影響，因為校驗和沒有修改，因此沒有人使用這些資料。但是，如果客戶有特殊用途並且需要重新計算校驗和，則這些垃圾資料會導致末尾的校驗和損壞（其他裝置，如NAT/負載平衡器也可能發現問題）。

裝置是N9K 93120TX（最初是在9372TX上檢測到的），版本是最新的NXOS 7.0(3)I2(2a)。

使用帶有直接連線到N9K（無任何型別的虛擬化）的Linux主機（1000base-T連結）。

使用以下設定：

```
interface Ethernet1/59
    switchport mode trunk
!
interface Ethernet1/60
    switchport mode trunk
```

linux configurations:

```
inet 10.2.1.1/24 brd 10.2.1.255 scope global eth1 <= native vlan
inet 10.1.1.1/24 brd 10.1.1.255 scope global eth1.100 <= tagged vlan 100
```

或

只要連線Windows主機並傳送已標籤的幀，它們就會觸發問題。此外，請確保網路介面卡(NIC)能夠標籤封包。

交換機將非零填充新增到通過的幀中。

例如：主機 — [中繼] N9K [中繼] — 主機

您可以使用netcat來傳送和接收封包。

如圖所示，它會傳送交換器上的Side（已標籤VLAN 100）連線埠e1/59。

```
6: eth1.100@eth1: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 9000 qdisc noqueue state UP group default
link/ether 44:a8:42:2c:5f:c4 brd ff:ff:ff:ff:ff:ff
inet 10.1.1.1/24 brd 10.1.1.255 scope global eth1.100
valid_lft forever preferred_lft forever
inet6 fe80::46a8:42ff:fe2c:5fc4/64 scope link
valid_lft forever preferred_lft forever
```

```
root@s35-c2-0:~# nc 10.1.1.2 3002 -u
a
^C
root@s35-c2-0:~#
```

它接收交換機上的端 (VLAN 100標籤) 埠e1/60 , 如下圖所示 :

```
7: eth1.100@eth1: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 9000 qdisc noqueue state UP group default
link/ether 44:a8:42:2c:63:d1 brd ff:ff:ff:ff:ff:ff
inet 10.1.1.2/24 brd 10.1.1.255 scope global eth1.100
valid_lft forever preferred_lft forever
inet6 fe80::46a8:42ff:fe2c:63d1/64 scope link
valid_lft forever preferred_lft forever
```

```
root@s35-c2:~# nc -l -u -p 3002
a
^C
root@s35-c2:~#
```

如圖所示 , 封包會傳輸。

```
root@s35-c2-0:~# tcpdump -i eth1.100 -nvex
tcpdump: listening on eth1.100, link-type EN10MB (Ethernet), capture size 65535 bytes
10:42:20.953994 44:a8:42:2c:5f:c4 > 44:a8:42:2c:63:d1, ethertype IPv4 (0x0800), length 44: (tos 0x0, ttl 64, id 64283, offset 0, flags [DF], proto UDP (17), length 30)
  10.1.1.1.41675 > 10.1.1.2.3002: UDP, length 2
  0x0000: 4500 001e fb1b 4000 4011 29af 0a01 0101
  0x0010: 0a01 0102 a2cb 0bba 000a 1620 610a
^C
1 packet captured
1 packet received by filter
0 packets dropped by kernel
root@s35-c2-0:~#
```

系統收到封包 , 如下圖所示 :

```
10:43:12.665897 44:a8:42:2c:5f:c4 > 44:a8:42:2c:63:d1, ethertype IPv4 (0x0800), length 60: (tos 0x0, ttl 64, id 64283, offset 0, flags [DF], proto UDP (17), length 30)
  10.1.1.1.41675 > 10.1.1.2.3002: UDP, length 2
  0x0000: 4500 001e fb1b 4000 4011 29af 0a01 0101
  0x0010: 0a01 0102 a2cb 0bba 000a da45 610a 0000
  0x0020: 0000 0000 0000 0000 0000 7562 710e
^C
7 packets captured
7 packets received by filter
0 packets dropped by kernel
root@s35-c2:~#
```

如圖所示 , 錯誤填充突出顯示。

No.	Time	Source	Destination	Protocol	Length	Info
1	0.000000	10.1.1.1	10.1.1.2	UDP	60	Source port: 40849 Destination port: 3002

▶ Frame 1: 60 bytes on wire (480 bits), 60 bytes captured (480 bits)

▼ Ethernet II, Src: Dell_2c:5f:c4 (44:a8:42:2c:5f:c4), Dst: Dell_2c:63:d1 (44:a8:42:2c:63:d1)

- ▶ Destination: Dell_2c:63:d1 (44:a8:42:2c:63:d1)
- ▶ Source: Dell_2c:5f:c4 (44:a8:42:2c:5f:c4)
- Type: IP (0x0800)

Padding: 000000000000000000000000f1b7bc5c

▼ Internet Protocol Version 4, Src: 10.1.1.1 (10.1.1.1), Dst: 10.1.1.2 (10.1.1.2)

- 0100 = Version: 4
- 0101 = Header Length: 20 bytes
- ▶ Differentiated Services Field: 0x00 (DSCP 0x00: Default; ECN: 0x00: Not-ECT (Not ECN-Capable Transport))
- Total Length: 30
- Identification: 0xfb1d (64285)
- ▶ Flags: 0x02 (Don't Fragment)
- Fragment offset: 0
- Time to live: 64
- Protocol: UDP (17)
- ▶ Header checksum: 0x29ad [validation disabled]
- Source: 10.1.1.1 (10.1.1.1)
- Destination: 10.1.1.2 (10.1.1.2)
- [Source GeoIP: Unknown]
- [Destination GeoIP: Unknown]

▼ User Datagram Protocol, Src Port: 40849 (40849), Dst Port: 3002 (3002)

- Source Port: 40849 (40849)
- Destination Port: 3002 (3002)
- Length: 10
- ▼ Checksum: 0xdd7f [validation disabled]
- [Good Checksum: False]
- [Bad Checksum: False]
- [Stream index: 0]

▼ Data (2 bytes)

- Data: 610a
- [Length: 2]

0000	44 a8 42 2c 63 d1 44 a8 42 2c 5f c4 08 00 45 00	D,B,c,D, B,....E.
0010	00 1e fb 1d 40 00 40 11 29 ad 0a 01 01 01 0a 01	...@.@.).....
0020	01 02 9f 91 0b ba 00 0a dd 7f 61 0a 00 00 00 00a.....
0030	00 00 00 00 00 00 00 00 f1 b7 bc 5c\

這也會透過封包分析器顯示 (在另一個封包中，資料與之前的截圖不同，但測試和錯誤相同)，

解決方案

因應措施是在已連線此伺服器的介面上禁用[buffer-boost](#)。

```
C9396PX-1(config)# int et 1/7
C9396PX-1(config-if)# no buffer-boost
```

相關缺陷：

[N9000](#)上採用dot1q報頭L2交換的CSCva46849 60位元組幀