전송 MPLS 트래픽에 대한 잘못된 L2 헤더 트러블슈팅

목차

소개

이 문서에서는 NCS4K 노드(6.5.26)의 L2 헤더 손상으로 인해 MPLS 전달 경로가 끊어진 경우 문제를 해결하는 방법에 대해 설명합니다.

문제

이 시나리오에서는 TE(Traffic Engineering) 터널이 작동하지만 MPLS(Multiprotocol Label Switching) ping이 MPLS 터널을 통해 작동하지 않습니다.

```
tunnel-te5180 10.38.101.62 Up Up default #ping mpls traffic-eng tunnel-te 5180 Thu Jan 5 21:30:29.245 UTC Sending 5, 100-byte MPLS Echos to tunnel-te5180, timeout is 2 seconds, send interval is 0 msec: Codes: '!' - success, 'Q' - request not sent, '.' - timeout, 'L' - labeled output interface, 'B' - unlabeled output interface, 'D' - DS Map mismatch, 'F' - no FEC mapping, 'f' - FEC mismatch, 'M' - malformed request, 'm' - unsupported tlvs, 'N' - no rx label, 'P' - no rx intf label prot, 'p' - premature termination of LSP, 'R' - transit router, 'I' - unknown upstream index, 'X' - unknown return code, 'x' - return code 0 Type escape sequence to abort. ..... Success rate is 0 percent (0/5)
```

MPLS traceroute에서 다음과 같은 홉이 표시됩니다.

```
#traceroute mpls traffic-eng tunnel-te 5180 Thu Jan 5 21:30:49.405 UTC Tracing MPLS TE Label Switched Path on tunnel-te5180, timeout is 2 seconds Codes: '!' - success, 'Q' - request not sent, '.' - timeout, 'L' - labeled output interface, 'B' - unlabeled output interface, 'D' - DS Map mismatch, 'F' - no FEC mapping, 'f' - FEC mismatch, 'M' - malformed request, 'm' - unsupported tlvs, 'N' - no rx label, 'P' - no rx intf label prot, 'p' - premature termination of LSP, 'R' - transit router, 'I' - unknown upstream index, 'X' - unknown return code, 'x' - return code 0 Type escape sequence to abort. 0 172.16.61.78 MRU 9582 [Labels: 27769 Exp: 0] L 1 172.16.61.79 MRU 9582 [Labels: 28136 Exp: 0] 7 ms . 2 * . 3 * . 4 * . 5 *^C
```

MPLS 터널을 선택하면 ERO(Explicit Router)에 대해 두 개의 홉이 더 표시됩니다.

```
#show mpls traffic-eng tunnels 5180 Thu Jan 5 21:31:11.958 UTC Name: tunnel-te5180 Destination: 10.38.96.1 Ifhandle:0x80002c4 Signalled-Name: MIVLPAMI-0112003A_t5180 Status: Admin: up Oper: up Path: valid Signalling: connected path option 10, type dynamic (Basis for Setup, path weight 3000) Accumulative metrics: TE 3000 IGP 30 Delay 900000 Path-option attribute: eline-any Number of affinity constraints: 1 Include bit map : 0x2 Include ext bit map : Length: 256 bits Value : 0x::2 Include affinity name : eline(1) G-PID: 0x0800 (derived from egress interface properties) Bandwidth Requested: 7 kbps CT0 Creation Time: Thu Nov 10 22:17:55 2022 (7w6d ago) Config Parameters: Bandwidth: 0 kbps (CT0) Priority: 5 5 Affinity: 0x0/0xffff Metric Type: TE (interface) Path Selection: Tiebreaker: Min-fill (default) Hop-limit: disabled Cost-limit: disabled Delay-limit: disabled Path-invalidation timeout: 10000 msec (default), Action: Tear
```

```
(default) AutoRoute: enabled LockDown: disabled Policy class: not set Forward class: 0 (not enabled) Forwarding-Adjacency: disabled Autoroute Destinations: 0 Loadshare: 0 equal loadshares Auto-bw: enabled Last BW Applied: 7 kbps CT0 BW Applications: 29 Last Application Trigger: Periodic Application Bandwidth Min/Max: 0-4294967295 kbps Application Frequency: 60 min Jitter: 0s Time Left: 46m 48s Collection Frequency: 5 min Samples Collected: 2 Next: 1m 3s Highest BW: 0 kbps Underflow BW: 0 kbps Adjustment Threshold: 10% 10 kbps Overflow Detection disabled Underflow Detection disabled Resignal Last-bandwidth Disabled Auto-Capacity: Disabled: Fast Reroute: Enabled, Protection Desired: Any Path Protection: Not Enabled BFD Fast Detection: Disabled Reoptimization after affinity failure: Enabled Soft Preemption: Disabled History: Tunnel has been up for: 7w6d (since Thu Nov 10 22:17:55 UTC 2022) Current LSP: Uptime: 15:09:12 (since Thu Jan 05 06:22:00 UTC 2023) Reopt. LSP: Last Failure: LSP not signalled, identical to the [CURRENT] LSP Date/Time: Thu Jan 05 19:03:33 UTC 2023 [02:27:39 ago] Prior LSP: ID: 32 Path Option: 10 Removal Trigger: reoptimization completed Path info (IS-IS 1 level-2): Node hop count: 3 Hop0: 172.16.61.79 Hop1: 172.16.57.244 Hop2: 172.16.6.59 Hop3: 10.38.96.1
```

경로를 따라 첫 번째 홉으로 이동하면 이 터널에 대한 올바른 MPLS LFIB(Label Forwarding Information Base) 항목이 표시됩니다.

```
#show mpls forwarding labels 27769 Fri Jan 6 06:13:04.220 UTC Local Outgoing Prefix Outgoing Next Hop Bytes Label Label or ID Interface Switched ------ ----------- ------------------ ------ ------ --------------- ------------ 27769 28136 TE: 5180 Hu0/10/0/11/2.4001 172.16.57.244 0 28136 TE: 5180 tt60409 point2point 0 (!)
```

이 노드의 이그레스 인터페이스는 이 MAC 주소를 사용하므로, L2 프레임의 L2 헤더에서 소스(SRC) MAC으로 사용할 수 있습니다.

```
#show interfaces hundredGigE 0/10/0/11/2.4001 Fri Jan 6 06:14:45.773 UTC HundredGigE0/10/0/11/2.4001 is up, line protocol is up Interface state transitions: 79 Hardware is VLAN sub-interface(s), address is 0c11.67c8.2041 Description: To HundredGigE1/3/0/10/2.PHLAPALO-12121302A:CID:I1001/GE100/PHLAPALO/SLTNPAST Internet address is 172.16.57.245 MTU 9600 bytes, BW 100000000 Kbit (Max: 100000000 Kbit) reliability 255/255, txload 0/255, rxload 0/255 Encapsulation 802.1Q Virtual LAN, VLAN Id 4001, loopback not set, Last link flapped 1w6d ARP type ARPA, ARP timeout 04:00:00 Last input 00:00:00, output 00:00:00 Last clearing of "show interface" counters never 5 minute input rate 64000 bits/sec, 62 packets/sec 5 minute output rate 2198000 bits/sec, 699 packets/sec 4529877895 packets input, 2267795435148 bytes, 6 total input drops 0 drops for unrecognized upper-level protocol Received 124 broadcast packets, 0 multicast packets 3926978895 packets output, 1611587340639 bytes, 0 total output drops Output 0 broadcast packets, 0 multicast packets
```

그러나 **show captured packets ingress location <active LC VM>**의 네이버 측에서는 올바른 MPLS 레이블이 표시되지만 완전히 잘못된 L2 SRC and Destination(DST) MAC 주소가 표시됩니다.

```
[200] Jan 6 06:10:12.449, len: 103, hits: 1, buffhdr type: 1 i/p i/f: HundredGigE1/3/0/10/2 punt reason: DROP_PACKET Ingress Headers: port_ifh: 0x8001ae4, sub_ifh: 0x0, bundle_ifh: 0x0 logical_port: 0x6c1, pd_pkt_type: 3 punt_reason: DROP_PACKET (0) payload_offset: 21, l3_offset: 21 FTMH: pkt_size: 0x7e, tc: 0, tm_act_type: 0, ssp: 0x981 PPH: pph_fwd_code: CPU Trap (7), fwd_hdr_offset: 0 inlif: 0x0, vrf: 0x0, rif: 0x0 FHEI: trap_code: Rx_UNKNOWN_PACKET (63), trap_qual: 193 [ether dst: 0000.0000.0000 src: 0c11.67c8.2000 type/len: 0x8847] [MPLS label: 28136, exp 0x6, eos 0, ttl 255]
```

DST Mac 주소는 all-0이고 SRC MAC 주소는 NCS4K 노드의 관리 인터페이스와 일치합니다.

```
#show interfaces mgmtEth 0/rp1/emS/0 Fri Jan 6 06:15:59.141 UTC MgmtEth0/RP1/EMS/0 is down, line protocol is down Interface state transitions: 0 Hardware is Management Ethernet, address is 0c11.67c8.2000 (bia 0c11.67c8.2000) Internet address is 10.230.192.86 MTU 1514 bytes, BW 100000 Kbit (Max: 100000 Kbit) reliability 255/255, txload 0/255, rxload 0/255 Encapsulation ARPA, Full-duplex, 100Mb/s, unknown, link type is autonegotiation loopback not set, ARP type ARPA, ARP timeout 04:00:00 Last input never, output never Last clearing of "show interface" counters never 5 minute input rate 0 bits/sec, 0 packets/sec 5 minute output rate 0 bits/sec, 0 packets/sec 0 packets input, 0 bytes, 0 total input drops 0 drops for unrecognized upper-level protocol
```

```
Received 0 broadcast packets, 0 multicast packets 0 runts, 0 giants, 0 throttles, 0 parity 0
input errors, 0 CRC, 0 frame, 0 overrun, 0 ignored, 0 abort 0 packets output, 0 bytes, 0 total
output drops Output 0 broadcast packets, 0 multicast packets 0 output errors, 0 underruns, 0
applique, 0 resets 0 output buffer failures, 0 output buffers swapped out 0 carrier transitions
```

솔루션

근본 원인은 DDT입니다. **Cisco 버그 ID CSCvz99253** Cisco 버그 ID CSCwa11748 Cisco 버그 ID CSCvz의 중복99253 6.5.32 릴리스에서 고정되어 있습니다.

```
++++++++ Dec 23 23:24:58.214 ofa_ipnhgroup_event 0/LC1 3235382# t4839
TP3147224,dnxsdk_l3_fec_create,enter,trans_id,357633649,npu_id,0,is_modify,1,use_eei_encoding,0,
dest,0x6052,encap_id,0x4000304a,fec_id,0x2001fe98, Dec 23 15:10:16.787 ofa_ipnh_event 0/LC1
210586# t5614
TP2061,client_ipnh_create,grid_res_id_alloc_req_success,trans_id,357386085,encap_id,0x304a,alloc
_sz,2 Dec 23 15:10:16.787 ofa_ipnh_event 0/LC1 286920# t4857
TP9909,dispatch_ipnh,resolve_refhdl_success,ref_l3intf_trans_id,357386081,hdl,0x87a75238 Dec 23
15:10:16.787 ofa_ipnh_event 0/LC1 172418# t4857
TP9913,dummy_block,trans_id,357386085,wait,duration,time,0.1742 Dec 23 15:10:16.787
ofa_ipnh_event 0/LC1 153352# t4857
TP3149344,srv_ipnh_create,entry,trans_id,357386085,npu_mask,0x100000,l3a_mac_addr,00af.1f18.0043
,l3a_intf_id,28,port_id,0 Dec 23 15:10:16.780 ofa_ipnh_event 0/LC1 258284# t5614
TP2061,client_ipnh_create,grid_res_id_alloc_req_success,trans_id,357386077,encap_id,0x3048,alloc
_sz,2 Dec 23 15:10:16.780 ofa_ipnh_event 0/LC1 105614# t4857
TP9909,dispatch_ipnh,resolve_refhdl_success,ref_l3intf_trans_id,357383451,hdl,0x87a75238 Dec 23
15:10:16.780 ofa_ipnh_event 0/LC1 172408# t4857
TP9913,dummy_block,trans_id,357386077,wait,duration,time,0.1947 Dec 23 15:10:16.780
ofa_ipnh_event 0/LC1 286912# t4857
TP3149344,srv_ipnh_create,entry,trans_id,357386077,npu_mask,0x100000,l3a_mac_addr,00af.1f18.0043
,l3a_intf_id,28,port_id,0 ++++++++
```

복구 방법으로 영향을 받는 이그레스 하위 인터페이스를 재구성할 수 있습니다.

이 번역에 관하여

Cisco는 전 세계 사용자에게 다양한 언어로 지원 콘텐츠를 제공하기 위해 기계 번역 기술과 수작업 번역을 병행하여 이 문서를 번역했습니다. 아무리 품질이 높은 기계 번역이라도 전문 번역가의 번역 결과물만큼 정확하지는 않습니다. Cisco Systems, Inc.는 이 같은 번역에 대해 어떠한 책임도 지지 않으며 항상 원본 영문 문서(링크 제공됨)를 참조할 것을 권장합니다.