Risoluzione dei problemi di flap BGP tra Ultra Packet Core e lo switch Nexus a causa di una configurazione errata

Sommario

Introduzione

In questo documento viene descritta la soluzione ai flap Border Gateway Protocol (BGP) tra Cisco Ultra Packet Core (UPC) e lo switch Nexus 9000 configurato con connessione BGP ridondante.

Problema

I flap BGP vengono attivati quando una delle interfacce ridondanti tra Cisco Ultra Packet Core e i flap dello switch Nexus.

Condizioni

Il nodo Ultra Packet Core (UPC) è collegato a Nexus Leaf A e Leaf B su porte separate. I peer IPv6 BGP vengono stabiliti e le route predefinite vengono installate nel nodo UPC. Nella Figura 1 viene mostrato un diagramma di rete di alto livello con un percorso ridondante agli switch foglia.

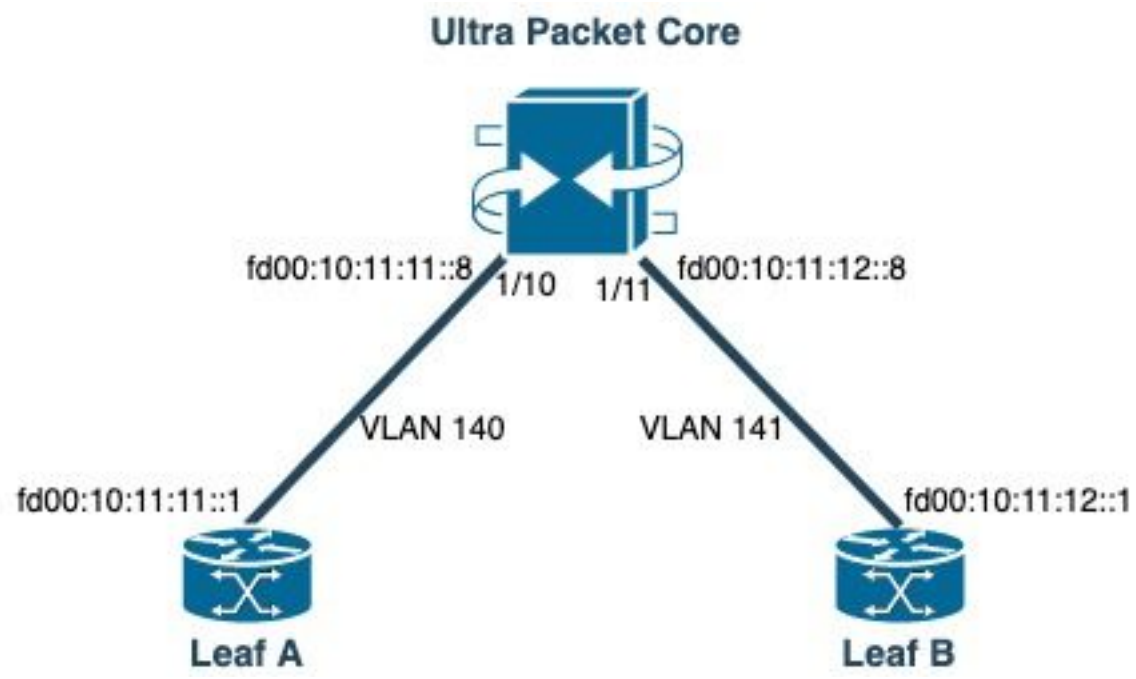

Figura 1: diagramma reticolare

Configurazione

Configurazione della porta UPC con VLAN e binding dell'interfaccia:

```
port ethernet 1/10
    no shutdown
    vlan 140
        no shutdown
        bind interface saegw_vlan140_1/10 saegw
#exit

#exit
port ethernet 1/11
    no shutdown
    vlan 141
        no shutdown
        bind interface saegw_vlan141_1/11 saegw
#exit
#exit
end
```

Configurazione interfaccia UPC con indirizzi IP:

```
interface saegw_vlan140_1/10
  ip address 10.11.11..8 255.255.255.0
  ipv6 address fd00:10:11:11::8/64 secondary
  bfd interval 300 min_rx 300 multiplier 3
#exit
interface saegw_vlan141_1/11
  ip address 10.11.12.8 255.255.255.0
  ipv6 address fd00:10:11:12::8/64 secondary
  bfd interval 300 min_rx 300 multiplier 3
#exit
```

Configurazione UPC BGP:

```
router bgp 25949
  router-id 172.19.20.30
  maximum-paths ebgp 4
  neighbor 10.11.11..1 remote-as 25949
  neighbor 10.11.11..1 fall-over bfd
  neighbor 10.11.12.1 remote-as 25949
  neighbor 10.11.12.1 fall-over bfd
  neighbor fd00:10:11:11::1 remote-as 25949
  neighbor fd00:10:11:12::1 remote-as 25949
  address-family ipv4
    neighbor 10.11.11..1 route-map accept_default in
    neighbor 10.11.11..1 route-map gw-1-OUT out
    neighbor 10.11.12.1 route-map accept_default in
    neighbor 10.11.12.1 route-map gw-1-OUT out
    redistribute connected
#exit
address-family ipv6
  neighbor fd00:10:11:11::1 activate
  neighbor fd00:10:11:11::1 route-map accept_v6_default in
  neighbor fd00:10:11:11::1 route-map allow_service_ips_v6 out
  neighbor fd00:10:11:12::1 activate
  neighbor fd00:10:11:12::1 route-map accept_v6_default in
  neighbor fd00:10:11:12::1 route-map allow_service_ips_v6 out
  redistribute connected
#exit

ipv6 prefix-list name accept_v6_default_routes seq 10 permit ::/0
route-map accept_v6_default permit 10
  match ipv6 address prefix-list accept_v6_default_routes
#exit
```

Configurazione switch Nexus 9000:

```
Interface vlan140
ipv6 address fd00:10:11:11::1/64
no ipv6 redirects

interface vlan141
ipv6 address fd00:10:11:12::1/64
no ipv6 redirects

vrf upc
address-family ipv4 unicast
advertise l2vpn evpn
maximum-paths ibgp 2
address-family ipv6 unicast
advertise l2vpn evpn
maximum-paths ibgp 2
neighbor fd00:10:11:12::5
remote-as 25949
address-family ipv6 unicast
neighbor fd00:10:11:12::6
remote-as 25949
address-family ipv6 unicast
neighbor fd00:10:11:12::8
remote-as 25949
address-family ipv6 unicast
```

Analisi

Inizialmente viene osservata una normale comunicazione BGP tra una delle interfacce UPC (fd00:10:11:12:18:8) e lo switch Nexus (fd00:10:11:12:1 appartiene a vlan141), che include i

messaggi TCP ACK:

```
2023-01-01 01:01:59.000000 fd00:10:11:12::8 -> fd00:10:11:12::1 TCP 35813 > bgp [ACK] Seq=250
Ack=8664 Win=31744 Len=0 TSV=2412344062 TSER=531234647
2023-01-01 01:01:59.000087 fd00:10:11:12::8 -> fd00:10:11:12::1 TCP 35813 > bgp [ACK] Seq=250
Ack=11520 Win=37376 Len=0 TSV=2412344062 TSER=531234647
2023-01-01 01:01:59.000162 fd00:10:11:12::8 -> fd00:10:11:12::1 TCP 35813 > bgp [ACK] Seq=250
Ack=14376 Win=43008 Len=0 TSV=241234062 TSER=531234647
2023-01-01 01:01:59.000281 fd00:10:11:12::8 -> fd00:10:11:12::1 TCP 35813 > bgp [ACK] Seq=250
Ack=17232 Win=49152 Len=0 TSV=2412344062 TSER=531234647
2023-01-01 01:01:59.000936 fd00:10:11:12::8 -> fd00:10:11:12::1 TCP 35813 > bgp [ACK] Seq=250
Ack=20663 Win=48640 Len=0 TSV=2412344063 TSER=531234647
```

In caso di guasto dell'interfaccia Leaf-B verso l'UPC, viene rilevato un comportamento errato nei log in cui un nuovo tentativo di connessione BGP viene avviato dall'UPC (fonte: fd00:10:11:12:8) verso l'interfaccia Leaf-A fd00:10:11:11:10:1, che appartiene a una VLAN diversa, vlan140.

```
2023-01-01 22:36:12.370117 fd00:10:11:12::8 -> fd00:10:11:11::1 TCP 41987 > bgp [SYN] Seq=0
Win=14400 Len=0 MSS=1440 TSV=2412347369 TSER=0 WS=9
```

Il messaggio SYN BGP non valido inviato sull'interfaccia errata causa l'inattività di BGP. Quando Nexus annuncia la propria route connessa e UPC ottiene una route per l'interfaccia non funzionante su BGP, UPC tenta la connessione tramite un'altra interfaccia con un IP in uscita diverso/errato.

Soluzione

A causa della configurazione descritta nella sezione Condizione di questo articolo, dal momento che l'UPC riceve le informazioni sul percorso connesso di entrambe le foglie da entrambe le interfacce, quando una delle interfacce è inattiva, l'UPC tenta di comunicare con quella foglia attraverso l'altra interfaccia.

Per evitare che l'UPC invii i messaggi di connessione BGP dall'interfaccia errata, le modifiche alla configurazione da prendere in considerazione sono le seguenti:

1. Nella configurazione UPC, aggiungere update-source per il vicino. Questa configurazione impedisce la connessione BGP da un'interfaccia diversa, se l'interfaccia principale non è attiva. Ad esempio, quando saegw_vlan140_1/10 (fd00:10:11:11:1/64) è inattivo, il nodo non può utilizzare l'interfaccia in uscita saegw_vlan141_1/11 per il peer BGP fd00:10:11:11:8. Di seguito è riportato un esempio di configurazione:

   ```
   neighbor fd00:10:11:11::1 update-source fd00:10:11:11::8
   neighbor fd00:10:11:12::1 update-source fd00:10:11:12::8
   ```

2. Nella configurazione Nexus, bloccare i prefissi dalle interfacce errate.
   Ad esempio, neghiamo i percorsi per la foglia ridondante sul vicino fd00:10:11:11:1

   ```
   neighbor fd00:10:11:11::1
   update prefix list to deny fd00:10:11:12::8/64
   ```

3. Nello switch Nexus, il peering EBGP dal VTEP a un nodo esterno su VXLAN deve essere in un VRF tenant e deve utilizzare update-source di un loopback (peering su VXLAN) come consigliato nella Guida alla configurazione di Cisco Nexus 9000

Informazioni su questa traduzione

Cisco ha tradotto questo documento utilizzando una combinazione di tecnologie automatiche e umane per offrire ai nostri utenti in tutto il mondo contenuti di supporto nella propria lingua. Si noti che anche la migliore traduzione automatica non sarà mai accurata come quella fornita da un traduttore professionista. Cisco Systems, Inc. non si assume alcuna responsabilità per l’accuratezza di queste traduzioni e consiglia di consultare sempre il documento originale in inglese (disponibile al link fornito).