Fehlerbehebung bei BGP-Flaps zwischen Ultra Packet Core und Nexus-Switch aufgrund falscher Konfiguration

Inhalt

Einleitung

In diesem Dokument wird die Lösung für Border Gateway Protocol (BGP)-Flaps zwischen dem Cisco Ultra Packet Core (UPC)- und Nexus 9000-Switch beschrieben, die mit einer redundanten BGP-Verbindung konfiguriert wurden.

Problem

BGP-Flaps werden ausgelöst, wenn eine der redundanten Schnittstellen zwischen dem Cisco Ultra Packet Core- und Nexus-Switch-Flaps vorhanden ist.

Bedingungen

Der Ultra Packet Core (UPC)-Knoten ist an separaten Ports mit Nexus Leaf A und Leaf B verbunden. Die BGP-IPv6-Peers werden eingerichtet, und die Standardrouten werden auf dem UPC-Knoten installiert. Abbildung 1 zeigt das grobe Netzwerkdiagramm mit dem redundanten Pfad zu Leaf-Switches.

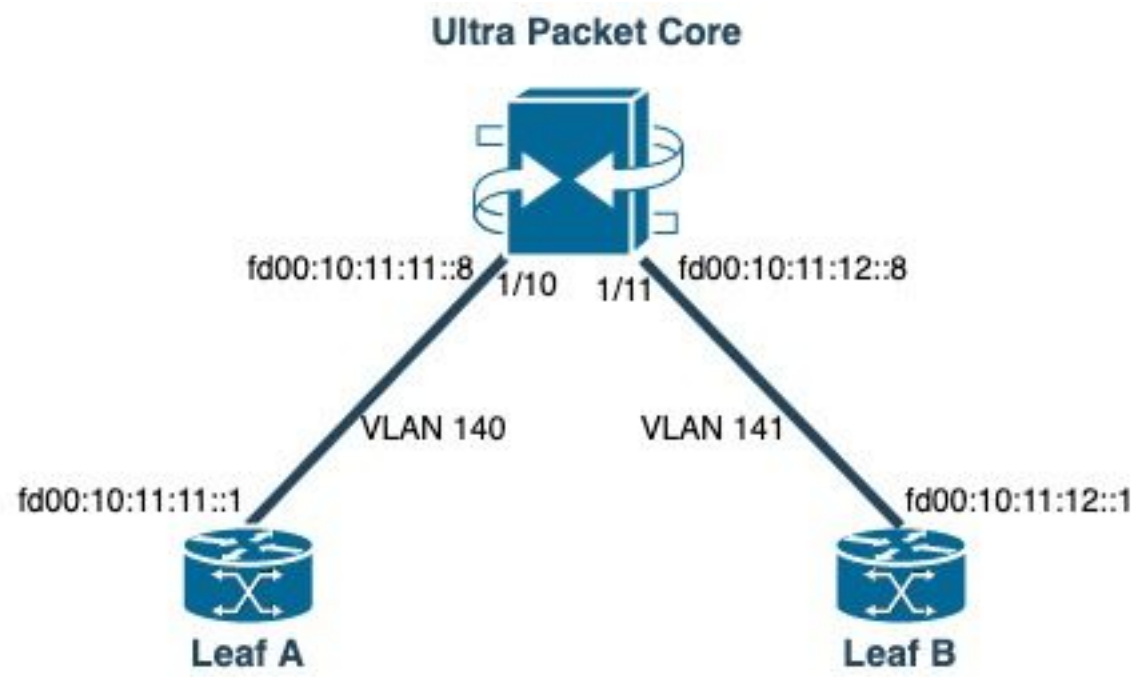

Abbildung 1: Netzwerkdiagramm

Konfiguration

UPC-Port-Konfiguration mit VLAN und Schnittstellenbindung:

```
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
```

Konfiguration der UPC-Schnittstelle mit IP-Adressen:

```
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
```

UPC-BGP-Konfiguration:

```
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
```

Nexus 9000-Switch-Konfiguration:

```
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
```

Analyse

Zunächst wird eine normale BGP-Kommunikation zwischen einer der UPC-Schnittstellen (fd00:10:11:12::8) und dem Nexus-Switch (fd00:10:11:12::1 gehört zu vlan141) beobachtet, die

TCP-ACK-Nachrichten enthält:

```
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
```

Bei Ausfall der Leaf-B-Schnittstelle zu UPC wird ein falsches Verhalten in den Protokollen festgestellt, in denen ein neuer BGP-Verbindungsversuch durch die UPC initiiert wird (Quelle: fd00:10:11:12::8) Richtung Leaf-A auf Schnittstelle fd00:10:11:11::1, die zu einem anderen VLAN gehört, vlll an140.

```
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
```

Eine solche ungültige BGP-SYN-Nachricht, die an die falsche Schnittstelle gesendet wird, führt zu einem BGP-Ausfall. Wenn der Nexus seine eigene verbundene Route meldet und UPC eine Route für die Schnittstelle erhält, die über BGP ausgefallen ist, versucht UPC, eine Verbindung über eine andere Schnittstelle mit einer anderen/falschen ausgehenden IP herzustellen.

Lösung

Aufgrund der Konfiguration im Abschnitt "Condition" dieses Artikels versucht UPC, über die andere Schnittstelle mit diesem Leaf zu kommunizieren, da UPC die verbundenen Routeninformationen beider Leafs von beiden Schnittstellen empfängt, wenn eine der Schnittstellen ausgefallen ist.

Damit UPC die BGP-Verbindungsaufbaumeldungen nicht von der falschen Schnittstelle senden kann, müssen die folgenden Konfigurationsänderungen berücksichtigt werden:

1. Fügen Sie in der UPC-Konfiguration update-source für den Nachbarn. Diese Konfiguration verhindert, dass die BGP-Verbindung von einer anderen Schnittstelle ausgeht, wenn die Hauptschnittstelle ausgefallen ist. Wenn saegw_vlan140_1/10 (fd00:10:11:11::1/64) heruntergefahren ist, kann der Knoten die ausgehende Schnittstelle saegw_vlan141_1/11 für den BGP-Peer fd00:10:11:11 nicht verwenden. 1:8)
   Nachfolgend finden Sie eine Beispielkonfiguration:

   ```
   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. In der Nexus-Konfiguration werden Präfixe für falsche Schnittstellen blockiert. Beispielsweise verweigern wir Routen für das redundante Leaf über den Nachbarn fd00:10:11:11:1.

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

3. Im Nexus-Switch muss sich das EBGP-Peering vom VTEP zu einem externen Knoten über VXLAN in einem Tenant-VRF befinden und den update-source einer loopback Schnittstelle

(Peering über VXLAN) gemäß der Empfehlung im Cisco Nexus 9000 Konfigurationsleitfaden

Informationen zu dieser Übersetzung

Cisco hat dieses Dokument maschinell übersetzen und von einem menschlichen Übersetzer editieren und korrigieren lassen, um unseren Benutzern auf der ganzen Welt Support-Inhalte in ihrer eigenen Sprache zu bieten. Bitte beachten Sie, dass selbst die beste maschinelle Übersetzung nicht so genau ist wie eine von einem professionellen Übersetzer angefertigte. Cisco Systems, Inc. übernimmt keine Haftung für die Richtigkeit dieser Übersetzungen und empfiehlt, immer das englische Originaldokument (siehe bereitgestellter Link) heranzuziehen.