Dépanner les failles BGP entre les commutateurs Ultra Packet Core et Nexus en raison d'une configuration incorrecte

Table des matières

Introduction

Ce document décrit la solution aux failles du protocole BGP (Border Gateway Protocol) entre Cisco Ultra Packet Core (UPC) et le commutateur Nexus 9000 configuré avec une connexion BGP redondante.

Problème

Les défaillances BGP sont déclenchées lorsque l'une des interfaces redondantes entre le coeur de réseau Cisco Ultra Packet Core et le commutateur Nexus est défaillante.

Conditions

Le noeud UPC (Ultra Packet Core) est connecté à Nexus Leaf A et Leaf B sur des ports distincts. Les homologues IPv6 BGP sont établis et les routes par défaut sont installées sur le noeud UPC. La Figure 1 présente le schéma de réseau de haut niveau avec un chemin redondant vers les commutateurs Leaf.

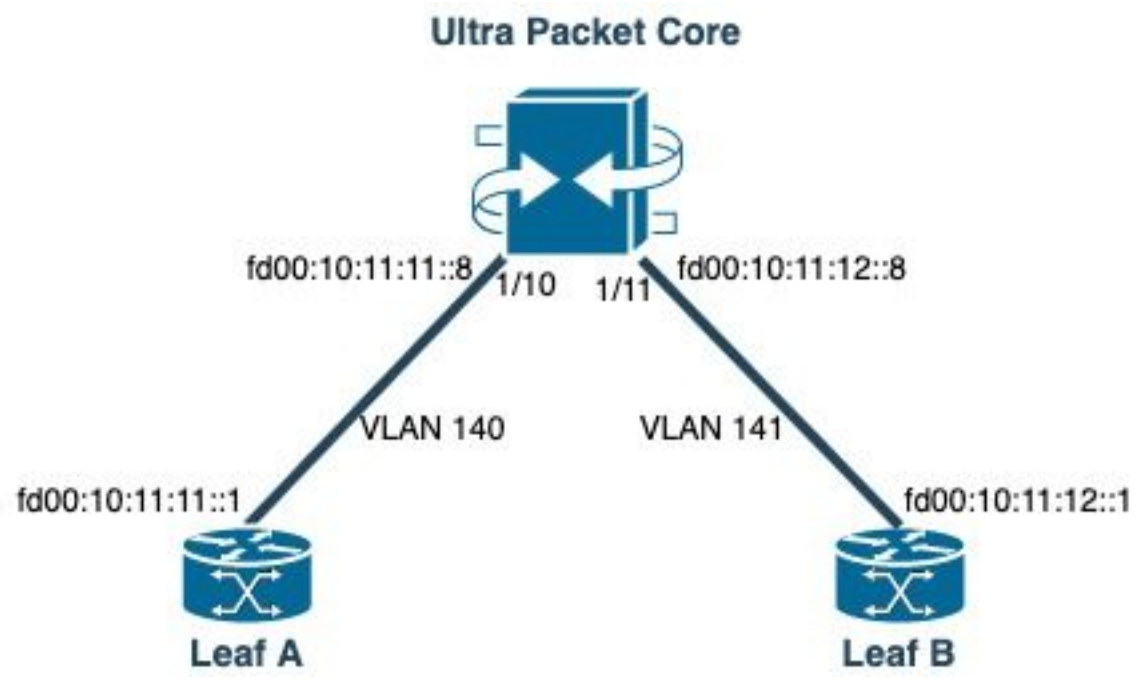

Figure 1 : Schéma de réseau

Configuration

Configuration des ports UPC avec liaison VLAN et interface :

```
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
```

Configuration d'interface UPC avec adresses 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
```

Configuration BGP UPC :

```
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
```

Configuration du commutateur 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
```

Analyse

Initialement, une communication BGP normale entre l'une des interfaces UPC (fd00:10:11:12::8) et le commutateur Nexus (fd00:10:11:12::1 appartient à vlan141) est observée qui inclut des

messages 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
```

En cas de défaillance de l'interface Leaf-B vers UPC, un comportement incorrect est observé dans les journaux où une nouvelle tentative de connexion BGP est initiée par l'UPC (source: fd00:10:11:12::8) vers le Leaf-A sur l'interface fd00:10:11:11::1, qui appartient à un VLAN différent, 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
```

Un tel message SYN BGP non valide envoyé sur une interface incorrecte entraîne l'arrêt du BGP. Lorsque le Nexus annonce sa propre route connectée et que l'UPC obtient une route pour l'interface qui était en panne sur BGP, alors l'UPC tente une connexion via une autre interface avec une adresse IP sortante différente/incorrecte.

Solution

En raison de la configuration mentionnée dans la section Condition de cet article, puisque UPC reçoit les informations de route connectée des deux Leaf des deux interfaces, quand l'une des interfaces est désactivée, UPC tente de communiquer avec ce Leaf par l'intermédiaire de l'autre interface.

Pour éviter qu'UPC n'envoie les messages d'établissement de connexion BGP à partir d'une interface incorrecte, voici les modifications de configuration à prendre en compte :

1. Dans la configuration UPC, ajoutez update-source pour le voisin. Cette configuration empêche la connexion BGP à partir d'une autre interface, si l'interface principale est désactivée. Par exemple, lorsque saegw_vlan140_1/10 (fd00:10:11:11::1/64) est en panne, le noeud ne peut pas utiliser l'interface sortante saegw_vlan141_1/11 pour l'homologue BGP fd00:10:11:11::8. Voici un exemple de configuration :

   ```
   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. Dans la configuration Nexus, bloquez les préfixes des mauvaises interfaces.
   Par exemple, nous refusons les routes pour le leaf redondant sur le voisin fd00:10:11:11::1

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

3. Dans le commutateur Nexus, l'appairage EBGP du VTEP à un noeud externe sur VXLAN doit se trouver dans un VRF de locataire et doit utiliser le update-source d'un loopback (appairage sur VXLAN) comme recommandé dans le Guide de configuration Cisco Nexus 9000

À propos de cette traduction
Cisco a traduit ce document en traduction automatisée vérifiée par une personne dans le cadre d'un service mondial permettant à nos utilisateurs d'obtenir le contenu d'assistance dans leur propre langue.

Il convient cependant de noter que même la meilleure traduction automatisée ne sera pas aussi précise que celle fournie par un traducteur professionnel.