設定が誤っているためにUltra Packet CoreとNexusスイッチ間で発生するBGPフラップのトラブルシューティング

内容

[概要](#)
[問題](#)
[条件](#)
[コンフィギュレーション](#)
[分析](#)
[解決方法](#)

概要

このドキュメントでは、Cisco Ultra Packet Core(UPC)と、冗長BGP接続が設定されたNexus 9000スイッチ間のBorder Gateway Protocol(BGP)フラップに対するソリューションについて説明します。

問題

BGPフラップは、Cisco Ultra Packet CoreとNexusスイッチ間の冗長インターフェイスの1つがフラップするとトリガーされます。

条件

Ultra Packet Core(UPC)ノードは、別々のポートでNexusリーフAとリーフBに接続されています。BGP IPv6ピアが確立され、デフォルトルートがUPCノードにインストールされます。図1は、リーフスイッチへの冗長パスを持つ高レベルネットワークダイアグラムを示しています。

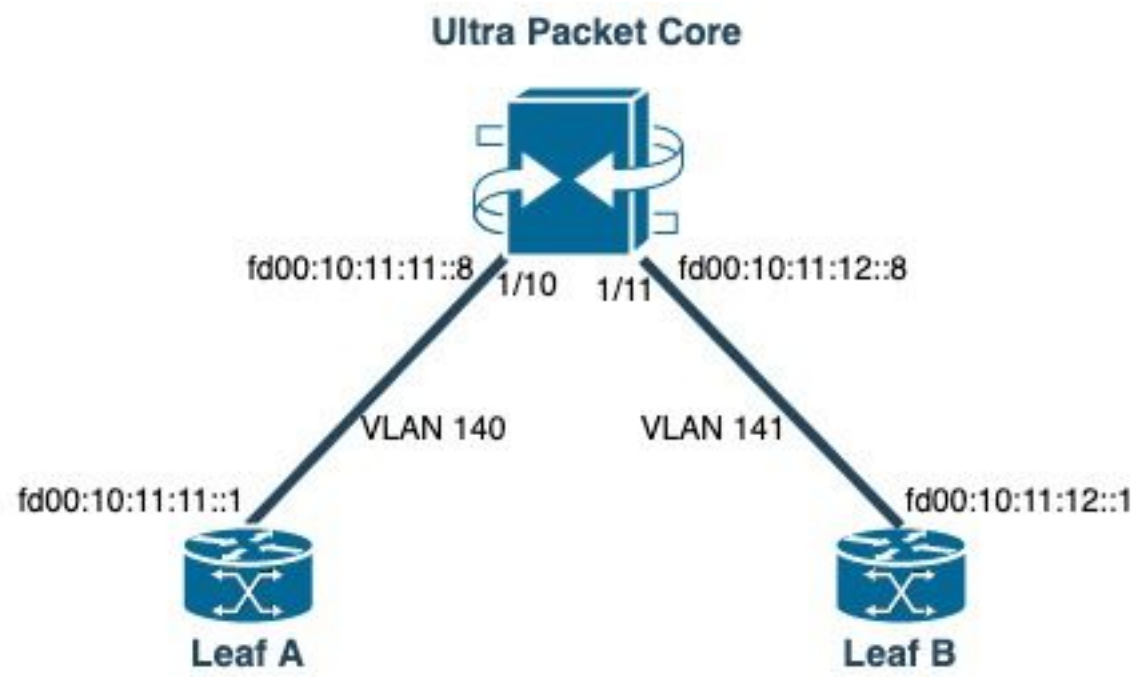

図1：ネットワークダイアグラム

コンフィギュレーション

VLANおよびインターフェイスバインディングを使用したUPCポート設定：

```
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
```

IPアドレスを使用したUPCインターフェイス設定：

```
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設定：

```
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スイッチの設定：

```
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
```

分析

最初に、いずれかのUPCインターフェイス(fd00:10:11:12::8)とNexusスイッチ(fd00:10:11:12::1 belongs to vlan141)の間の通常のBGP通信が観察され、これにはTCP ACKメッセージが含まれま

す。

```
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
```

UPCに対するLeaf-Bインターフェイスで障害が発生すると、UPC（送信元：fd00:10:11:12::8）によって、別のVLAN(vlan140)に属するインターフェイスfd00:10:11:11::1上のLeaf-Aに対して新しいBGP接続が開始されるというログに不正な動作がが示されます。

```
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
```

このような無効なBGP SYNメッセージが誤ったインターフェイスに送信されると、BGPがダウンします。Nexusが自身の接続ルートをアドバタイズし、UPCがBGP経由でダウンしていたインターフェイスのルートを取得すると、UPCは発信IPが異なる/誤っている別のインターフェイスを介して接続を試みます。

解決方法

この記事の「条件」セクションで説明した設定により、UPCは両方のインターフェイスから両方のリーフの接続ルート情報を受信するため、一方のインターフェイスがダウンすると、UPCは他方のインターフェイスを介してそのリーフへの通信を試みます。

UPCが誤ったインターフェイスからBGP接続確立メッセージを送信することを回避するために、考慮すべき設定変更を次に示します。

1. UPC設定で、次のコマンドを追加します。 update-source を設定します。この設定により、メインインターフェイスがダウンしている場合に、別のインターフェイスからのBGP接続が防止されます。たとえば、saegw_vlan140_1/10 (fd00:10:11:11::1/64)がダウンしている場合、ノードは発信インターフェイスsaegw_vlan141_1/11をBGPピアfd00:10:11:11::8に使用できません。
   次に設定例を示します。

   ```
   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. Nexusの設定で、誤ったインターフェイスからのプレフィックスをブロックします。
   たとえば、ネイバーfd00:10:11:11::1上の冗長リーフのルートを拒否します

   ```
   neighbor fd00:10:11:11::1
   update prefix list to deny fd00:10:11:12::8/64
   ```
3. Nexusスイッチでは、VXLANを介したVTEPから外部ノードへのEBGPピアリングがテナントVRF内にあり、 update-source の loopback 『Cisco Nexus 9000コンフィギュレーションガイド』の推奨に従って、インターフェイス(VXLAN経由のピアリング)を設定します。

翻訳について
シスコは世界中のユーザにそれぞれの言語でサポート コンテンツを提供するために、機械と人による翻訳を組み合わせて、本ドキュメントを翻訳しています。ただし、最高度の機械翻訳であっても、専門家による翻訳のような正確性は確保されません。シスコは、これら翻訳の正確性について法的責任を負いません。原典である英語版（リンクからアクセス可能）もあわせて参照することを推奨します。