フラッピング BGP ルート（再帰的ルーティング障害）のトラブルシューティング

目次

概要

この資料では、再帰的なルーティング障害によって引き起こされている Border Gateway Protocol（BGP; ボーダー ゲートウェイ プロトコル）のルートのフラッピングをトラブルシューティングする方法について説明します。

BGP の再帰的ルーティング障害のよく見られる現象は次のとおりです：

- ルーティング テーブルへの BGP ルートの持続する削除と再挿入。
- BGP を通じて学習した宛先への接続の切断。

前提条件

要件

このドキュメントに関する固有の要件はありません。

使用するコンポーネント

このドキュメントは、特定のソフトウェアやハードウェアのバージョンに限定されるものではありません。

背景理論

この文書を使用する際には、次のネットワーク ダイアグラムを参照してください。

この文書を使用する際には、次の設定を参照してください。

Rtr-A

```
hostname RTR-A
!
interface Loopback0
 ip address 10.10.10.10 255.255.255.255
!
interface Serial8/0
 ip address 192.168.16.1 255.255.255.252
!
router bgp 1
 bgp log-neighbor-changes
 neighbor 20.20.20.20 remote-as 2
 neighbor 20.20.20.20 ebgp-multihop 2
 neighbor 20.20.20.20 update-source Loopback0
!
ip route 20.20.20.0 255.255.255.0 192.168.16.2
```

Rtr-B

```
hostname RTR-B

!
interface Loopback0
 ip address 20.20.20.20 255.255.255.255
!
interface Ethernet0/0
 ip address 172.16.1.1 255.255.255.0
!

interface Serial8/0
 ip address 192.168.16.2 255.255.255.252
!
router bgp 2
 no synchronization
 bgp log-neighbor-changes
 network 20.20.20.20 mask 255.255.255.255
 network 172.16.1.0 mask 255.255.255.0
 neighbor 10.10.10.10 remote-as 1
 neighbor 10.10.10.10 ebgp-multihop 2
 neighbor 10.10.10.10 update-source Loopback0
 no auto-summary
!
ip route 10.10.10.0 255.255.255.0 192.168.16.1
!
```

表記法

ドキュメント表記の詳細は、『シスコ テクニカル ティップスの表記法』を参照してください。

問題

症状

これら 2 つの現象は、再帰的ルーティング障害で発生します：

- BGP で学習したルートが、IP ルーティング テーブルの中でフラッピングを繰り返す。フラッピングを表示するには、ルーティング テーブルを継続的に 2、3 分間観察します。RTR-A#**show ip route** Codes: C - connected, S - static, I - IGRP, R - RIP, M - mobile, B - BGP D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2 E1 - OSPF external type 1, E2 - OSPF external type 2, E - EGP i - IS-IS, L1 - ISIS level-1, L2 - ISIS level-2, ia - ISIS inter are * - candidate default, U - per-user static route, o - ODR P - periodic downloaded static route Gateway of last resort is not set 20.0.0.0/8 is variably subnetted, 2 subnets, 2 masks B 20.20.20.20/32 [20/0] via 20.20.20.20, 00:00:35 S 20.20.20.0/24 [1/0] via 192.168.16.2 172.16.0.0/24 is subnetted, 1 subnets B **172.16.1.0 [20/0] via 20.20.20.20, 00:00:35** 10.0.0.0/32 is subnetted, 1 subnets C 10.10.10.10 is directly connected, Loopback0 192.168.16.0/30 is subnetted, 1 subnets C 192.168.16.0 is directly connected, Serial8/0 **注: show ip route | include , 00:00** コマンドは、大きなルーティング テーブルを扱う場合にフラッピング ルートを確認するのに便利です。約 1 分間待機した後、**show ip route コマンド**の出力結果を次のように変更します：RTR-A#**show ip route** [..] Gateway of last resort is not set 20.0.0.0/24 is subnetted, 1 subnets S 20.20.20.0 [1/0] via 192.168.16.2 10.0.0.0/32 is subnetted, 1 subnets C 10.10.10.10 is directly connected, Loopback0 192.168.16.0/30 is subnetted, 1 subnets C 192.168.16.0 is directly connected, Serial8/0 **注: 上記のルーティング テーブルには BGP ルートがありません。**
- BGP ルートがルーティング テーブルにあるとき、これらのネットワークへの接続に失敗する。これを調べるために、Rtr-A のルーティング テーブルに BGP が学習するルート 172.16.1.0/24 がある場合に、有効なホスト 172.16.1.1 への ping が失敗します。RTR-A#**show ip route 172.16.1.0** Routing entry for 172.16.1.0/24 Known via "bgp 1", distance 20, metric 0 Tag 2, type external Last update from 20.20.20.20 00:00:16 ago Routing Descriptor Blocks: ***20.20.20.20, from 20.20.20.20, 00:00:16 ago** Route metric is 0, traffic share count is 1 AS Hops 1 RTR-A#**ping 172.16.1.1** Type escape sequence to abort. Sending 5, 100-byte ICMP Echos to 172.16.1.1, timeout is 2 seconds: ..... Success rate is 0 percent (0/5) RTR-A#

再帰的ルーティング障害

Rtr-A で、BGP ピア 20.20.20.20 へのルートを調べます。 このルートは、次の 2 つのホップ間で、1 分間に 1 回程度フラッピングを繰り返しています。

RTR-A#**show ip route 20.20.20.20** Routing entry for **20.20.20.20/32 Known via "bgp 1"**, distance 20, metric 0 Tag 2, type external Last update from 20.20.20.20 00:00:35 ago Routing Descriptor Blocks: * **20.20.20.20**, from 20.20.20.20, 00:00:35 ago Route metric is 0, traffic share count is 1 AS Hops 1

BGP ピアの IP アドレスへのルートは BGP 自体を介して学習します。 そのため、再帰的ルーティング障害を引き起こします。

約 1 分後、ルートは次のように変更します：

RTR-A#**show ip route 20.20.20.20** Routing entry for 20.20.20.0/24 **Known via "static"**, distance 1, metric 0 Routing Descriptor Blocks: ***192.168.16.2** Route metric is 0, traffic share count is 1

再帰的ルーティング障害の原因

再帰的なルータ障害の原因を、以下で説明します。

1. 「Rtr-A」の設定を参照してください。 この設定では、スタティック ルート 20.20.20.0/24 は、直接接続されたネクスト ホップ 192.168.16.2 を指すように設定されています。 ピア Rtr-B 20.20.20.20 の BGP セッションは、このスタティック ルートを使用して確立されています。

2. Rtr-B では、BGP ルート 172.16.1.0/24 と 20.20.20.20/32 を Rtr-A へアナウンスします。このとき、ループバック IP アドレスである 20.20.20.20 をネクストホップとして使用します。
3. Rtr-A は Rtr-B によりアナウンスされた BGP ルートを受信し、20.20.20.20/32 をインストールしようとします。 これはスタティック ルートとして Rtr-A ですでに設定されている 20.20.20.0/24 よりも詳細です。 最長一致ルートが優先されるため、20.20.20.20/32 は 20.20.20.0/24 に優先されます。 詳細については、『Cisco ルータにおけるルートの選択』を参照してください。 インストールされたルート 20.20.20.20/32 は、ルーティング テーブルに 20.20.20.20（Rtr-B のピア IP アドレス）のネクストホップを持っています。 20.20.20.20/32 へのルートは自分自身のネクストホップであるため、これによって再帰的なルーティング障害が発生します。このような特殊な状況で再帰的ルーティング障害の背後にある原因を理解するためには、ルーティング アルゴリズムがどのように動作するのかを理解する必要があります。 ネクスト ホップの IP アドレスがルータの直接接続インターフェイスではないルーティング テーブル内の非直接接続ルートについては、パケットを転送できる直接接続インターフェイスが見つかるまで、アルゴリズムがルーティング テーブルを再帰的に確認します。このような特定の状況では、Rtr-A は 20.20.20.20 の非直接接続のネクストホップ（自身）を使用して非直接接続されたネットワーク 20.20.20.20/32 へのルートを学習します。 このルーティング アルゴリズムでは、20.20.20.20/32 を宛先としたパケットを送信する直接接続されたインターフェイスを見つけられないため、再帰的なルーティング ループの障害に陥ります。
4. ルータは、この直接接続されていないルート 20.20.20.20/32 が再帰的なルーティング障害になっていることを検出し、ルーティング テーブルから 20.20.20.20/32 を引き出します。 その結果、ネクスト ホップ IP アドレス 20.20.20.20 を使用する BGP から学習したすべてのルートは、ルーティング テーブルから削除されます。
5. プロセス全体がステップ 1 から繰り返されます。 **debug ip routing** コマンドを実行すると、このことを確認できます。**注:** いずれの **debug コマンド**を実行する前でも、デバッグの出力を制限するために、特定のネットワークのアクセス コントロール リスト（ACL）に対する **debug コマンド**を実行します。 この例では、デバッグ出力を制限するように ACL を設定します。
```
RTR-A(config)#access-list 1 permit 20.20.20.20
RTR-A(config)#access-list 1 permit 172.16.1.0
RTR-A(config)#end
RTR-A#debug ip routing 1
IP routing debugging is on for access list 1
00:29:50: RT: add 20.20.20.20/32 via 20.20.20.20, bgp metric [20/0]
00:29:50: RT: add 172.16.1.0/24 via 20.20.20.20, bgp metric [20/0]
00:30:45: RT: recursion error routing 20.20.20.20 - probable routing loop
00:30:45: RT: recursion error routing 20.20.20.20 - probable routing loop
00:30:45: RT: recursion error routing 20.20.20.20 - probable routing loop
00:30:46: RT: recursion error routing 20.20.20.20 - probable routing loop
00:30:46: RT: recursion error routing 20.20.20.20 - probable routing loop
00:30:48: RT: recursion error routing 20.20.20.20 - probable routing loop
00:30:48: RT: recursion error routing 20.20.20.20 - probable routing loop
00:30:50: RT: del 20.20.20.20/32 via 20.20.20.20, bgp metric [20/0]
00:30:50: RT: delete subnet route to 20.20.20.20/32
00:30:50: RT: del 172.16.1.0/24 via 20.20.20.20, bgp metric [20/0]
00:30:50: RT: delete subnet route to 172.16.1.0/24
```
6. ルートの再帰が繰り返し失敗すると、次のエラー メッセージが表示されます：
```
%COMMON_FIB-SP-6-FIB_RECURSION: 10.71.124.25/32 has too many (8) levels of
recursion during setting up switching info
%COMMON_FIB-SP-STDBY-6-FIB_RECURSION: 10.71.124.25/32 has too many (8)
levels of recursion during setting up switching info
```
これは、MPLS 対応するネットワークに TCP の再送信によって発生します。 通信回線がダウンしたために BGP キープアライブ メッセージが一度 BGP ピアへの送信に失敗した場合、TCP がバックアップ パスを介して失敗メッセージを再送信しても、ネイバー BGP ピアはこれ以上のキープアライブ パケットを承認せず、最終的には、ホールドタイムの期限切れで BGP ピアがダウンします。 この問題は、MPLS が Catalyst6500 または Cisco7600 に設定されている場合にのみ発生します。 こ

の問題は、Cisco Bug ID CSCsj89544（登録ユーザ専用）に記述されています。

解決策

この問題に対するソリューションの詳細を次に説明します。

BGP ピアの IP アドレス（この場合は 20.20.20.20）用に、Rtr-A に特定のスタティック ルートを追加します。

```
RTR-A#configure terminal Enter configuration commands, one per line. End with CNTL/Z. RTR-A(config)#ip route 20.20.20.20 255.255.255.255 192.168.16.2
```

プレフィックス 20.20.20.20/32 用のスタティック ルートの設定により、ダイナミックに学習された BGP ルート 20.20.20.20/32 がルーティング テーブルにインストールされないことから、再帰的なルーティング ループの状況を回避することができます。 詳細については、『Cisco ルータにおけるルートの選択』を参照してください。

注: EBGP のピアがデフォルト ルートと相互に到達するように設定されている場合、BGP ネイバーシップは表示されません。 これは、ルート フラッピングとルーティング ループを避けるためです。

172.16.1.1 への ping はソリューションを確認します。

```
RTR-A#ping 172.16.1.1 Type escape sequence to abort. Sending 5, 100-byte ICMP Echos to 172.16.1.1, timeout is 2 seconds: !!!!! Success rate is 100 percent (5/5), round-trip min/avg/max = 20/24/40 ms
```

ルート ダンプニング

ルート ダンプニングは、インターネットワーク間でフラッピング ルートの伝搬を最小限に抑えるように設計された BGP 機能です。 ISP が推奨する値は Cisco IOS® のデフォルトであり、この機能を有効にするには、次のコマンドを設定する必要があります。

```
router bgp <AS number>
 bgp dampening
```

bgp dampening コマンドは、ハーフタイム = 15 分、再利用 = 750、サプレス = 2000、最大サプレス時間 = 60 分、などのダンプニング パラメータのデフォルト値を設定します。 これらの値はユーザが設定可能ですが、シスコでは変更しないことを推奨します。

関連情報

- 「#%BGP-3-INSUFCHUNKS: Insufficient chunk pools for aspath」エラー メッセージの意味
- BGP ネイバーがアイドル状態、接続状態、アクティブ状態間でトグルする理由
- BGP に関するサポート ページ
- テクニカルサポートとドキュメント - Cisco Systems