Troubleshoot Flapping BGP Routes (Recursive Routing Failure)

Contents

Introduction

This document describes how to troubleshoot flapping Border Gateway Protocol (BGP) routes caused by recursive routing failure.

Prerequisites

Requirements

There are no specific requirements for this document.

Components Used

This document is not restricted to specific software and hardware versions.

The information in this document was created from the devices in a specific lab environment. All of the devices used in this document started with a cleared (default) configuration. If your network is live, ensure that you understand the potential impact of any command.
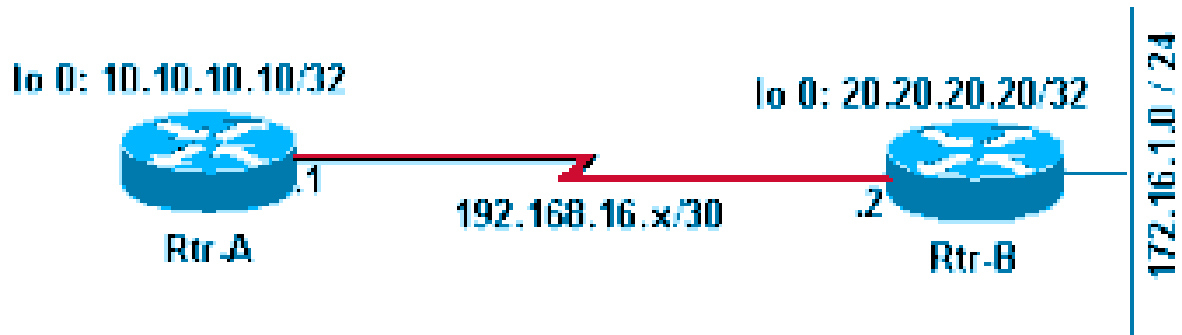
Background Information

This document describes how to troubleshoot flapping Border Gateway Protocol (BGP) routes caused by recursive routing failure.

Common symptoms of recursive routing failure in BGP are:

- Constant deletion and reinsertion of BGP routes into the routing table.

- Loss of connectivity towards destinations learned through BGP.

Refer to this network diagram as you use this document:

Network Diagram

Refer to these configurations as you use this document:

Rtr-A

```
<#root>

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 10.20.20.20 remote-as 2
 neighbor 10.20.20.20 ebgp-multihop 2
 neighbor 10.20.20.20 update-source Loopback0
!

ip route 10.20.20.0 255.255.255.0 192.168.16.2
```

Rtr-B

```
hostname RTR-B

!
interface Loopback0
 ip address 10.20.20.20 255.255.255.255
!
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 10.20.20.20 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
!
```

Conventions

Refer to Cisco Technical Tips Conventions for more information on document conventions.

Problem

Symptoms

These two symptoms are observed with recursive routing failure:

- The continuous flapping of BGP-learned routes in the IP routing table.

  Observe the routing table continuously for a couple of minutes in order to see the flapping.

  ```
  <#root>

  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

         10.0.0.0/8 is variably subnetted, 2 subnets, 2 masks
  B         10.20.20.20/32 [20/0] via 10.20.20.20, 00:00:35
  S         10.20.20.0/24 [1/0] via 192.168.16.2
         172.16.0.0/24 is subnetted, 1 subnets
  B

  172.16.1.0 [20/0] via 10.20.20.20, 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
```

📝 **Note**: It is helpful to use the **show ip route | include , 00:00** command in order to observe flapping routes when you deal with large routing tables.

After you wait for approximately one minute, the **show ip route** command results change to this:

```
<#root>

RTR-A#

show ip route

[..]

Gateway of last resort is not set

     10.0.0.0/24 is subnetted, 1 subnets
S       10.20.20.0 [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
```

📝 **Note**: The BGP routes are missing in the previous routing table.

- When the BGP routes are present in the routing table, connectivity to those networks fails.

  In order to observe this, when the routing table of the Rtr-A has BGP-learned route 172.16.1.0/24 in its routing table, a ping to valid host 172.16.1.1 fails.

  ```
  <#root>

  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 10.20.20.20 00:00:16 ago
    Routing Descriptor Blocks:

  * 10.20.20.20, from 10.20.20.20, 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#
```

Recursive Routing Failure

On Rtr-A, observe the route towards the BGP peer 10.20.20.20. The route flaps between the two next hops consistently every minute or so.

<#root>

RTR-A#

show ip route 10.20.20.20

Routing entry for

10.20.20.20/32

Known via "bgp 1"

, distance 20, metric 0
  Tag 2, type external
  Last update from 10.20.20.20 00:00:35 ago
  Routing Descriptor Blocks:
  *

10.20.20.20

, from 10.20.20.20, 00:00:35 ago
      Route metric is 0, traffic share count is 1
      AS Hops 1

The route towards the BGP peer IP address is learned through BGP itself; thus it creates a recursive routing failure.

After approximately a minute, the route changes to:

<#root>

RTR-A#

show ip route 10.20.20.20

Routing entry for 10.20.20.0/24

Known via "static"

, distance 1, metric 0
  Routing Descriptor Blocks:

```
* 192.168.16.2

        Route metric is 0, traffic share count is 1
```

Causes of Recursive Routing Failure

These steps describe the cause of recursive routing failures:

1. Refer to the configuration of Rtr-A. In this configuration, a static route 10.20.20.0/24 is configured to point to the directly connected next-hop 192.168.16.2. With this static route, a BGP session with peer Rtr-B 10.20.20.20 is established.

2. Rtr-B announces BGP routes 172.16.1.0/24 and 10.20.20.20/32 to Rtr-A with its loopback IP address 10.20.20.20 as the next-hop.

3. Rtr-A receives BGP routes announced by Rtr-B and tries to install the 10.20.20.20/32. This is more specific than 10.20.20.0/24, which is already configured in Rtr-A as a static route. Because the longest matching route is preferred, 10.20.20.20/32 is preferred over 10.20.20.0/24. Refer to Route Selection in Cisco Routers for more information. The installed route 10.20.20.20/32 has next-hop of 10.20.20.20 (Rtr-B peering IP address) in the routing table. This leads to recursive routing failure since the route towards 10.20.20.20/32 has a next-hop of itself.

   In order to understand the reason behind why recursive routing fails in this particular situation, you need to understand how the routing algorithm works. For any nondirectly-connected route in the routing table whose next hop IP address is not a directly-connected interface of the router, the algorithm looks recursively into the routing table until it finds a directly-connected interface to which it can forward the packets.

   In this particular situation, Rtr-A learns a route to the nondirectly-connected network 10.20.20.20/32 with a nondirectly-connected next hop of 10.20.20.20 (itself). The routing algorithm runs into a recursive routing loop failure because it is unable to find any directly-connected interface to which to send packets destined for 10.20.20.20/32.

4. The router detects that this nondirectly-connected route 10.20.20.20/32 has a recursive routing failure and withdraws 10.20.20.20/32 from the routing table. Consequently, all BGP-learned routes with the next hop IP address 10.20.20.20 are also withdrawn from routing table.

5. The whole process repeats from Step 1. You can confirm this if you issue the **debug ip routing** command.

   > **Note**: Before you run any **debug** command, run the **debug** command against an access control list (ACL) for a specific network in order to limit the output of debug. In this example, configure an ACL in order to limit the debug output.

   ```
   <#root>

   RTR-A(config)#

   access-list 1 permit 10.20.20.20

   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 10.20.20.20/32 via 10.20.20.20, bgp metric [20/0]
00:29:50: RT: add 172.16.1.0/24 via 10.20.20.20, bgp metric [20/0]

00:30:45: RT: recursion error routing 10.20.20.20 - probable routing loop

00:30:45: RT: recursion error routing 10.20.20.20 - probable routing loop
00:30:45: RT: recursion error routing 10.20.20.20 - probable routing loop
00:30:46: RT: recursion error routing 10.20.20.20 - probable routing loop
00:30:46: RT: recursion error routing 10.20.20.20 - probable routing loop
00:30:48: RT: recursion error routing 10.20.20.20 - probable routing loop
00:30:48: RT: recursion error routing 10.20.20.20 - probable routing loop
00:30:50: RT: del 10.20.20.20/32 via 10.20.20.20, bgp metric [20/0]
00:30:50: RT: delete subnet route to 10.20.20.20/32
00:30:50: RT: del 172.16.1.0/24 via 10.20.20.20, bgp metric [20/0]
00:30:50: RT: delete subnet route to 172.16.1.0/24
```

6. If the route recursion fails continuously, then this error message appears:

```
%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
```

This is due to the TCP retransmissions occurring on MPLS enabled network. If a BGP keepalive message has failed once and is sent to BGP Peer because the transport link is down, the neighbor BGP Peer does not accept any further keepalive packets even though TCP retransmits the failed message through the backup path, and it eventually leads to BGP peer down with holdtime expiration. This issue is seen only when MPLS is configured on Catalyst6500 or Cisco7600. This information is included in Cisco bug ID CSCsj89544 .

Note: Only registered Cisco users can access internal bug information and other tools.

Solution

The solution(s) to this problem are explained in these details.

Add a specific static route in Rtr-A for the BGP peer IP address (10.20.20.20 in this case).

```
<#root>

RTR-A#

configure terminal


Enter configuration commands, one per line.  End with CNTL/Z.
RTR-A(config)#

ip route 10.20.20.20 255.255.255.255 192.168.16.2
```

The configuration of a static route for prefix 10.20.20.20/32 ensures that a dynamically-learned BGP route 10.20.20.20/32 does not get installed in the routing table and thus avoids the recursive routing loop situation. Refer to Route Selection in Cisco Routers for more information.

✎ **Note**: When EBGP peers are configured to reach each other with default routes, the BGP neighborhood does not appear. This is done in order to avoid route flapping and routing loops.

A ping to 172.16.1.1 confirms the solution.

```
<#root>

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
```

Route Dampening

Route dampening is a BGP feature designed to minimize the propagation of flapping routes across an internetwork. The values the ISP recommended are the defaults on Cisco IOS® and you only need to configure this command in order to enable it.

```
<#root>

router bgp <AS number>


bgp dampening
```

The**bgp dampening**command sets default values for the dampening parameters such as Halftime= 15 minutes, reuse = 750, Suppress = 2000 and Max Suppress Time= 60. These values are user configurable, but Cisco recommends that they remain unchanged.

Related Information

- [What Does the "#%BGP-3-INSUFCHUNKS: Insufficient chunk pools for aspath" Error Message Mean?](#)
- [Troubleshoot BGP Neighborship Connection Issues](#)
- [IP Routing/BGP Support](#)
- [Cisco Technical Support & Downloads](#)