Solucionar problemas básicos do Border Gateway Protocol

Contents

Introduction

Este documento descreve como solucionar os problemas mais comuns com o Border Gateway Protocol (BGP) e fornece soluções e diretrizes básicas.

Prerequisites

Requirements

Não existem requisitos específicos para este documento. O conhecimento básico do protocolo BGP é útil; você pode consultar o Guia de Configuração BGP para obter mais informações.

Componentes Utilizados

Este documento não está restrito a versões específicas de software e hardware, mas os comandos são aplicáveis para o Cisco IOS® e o Cisco IOS-XE®.

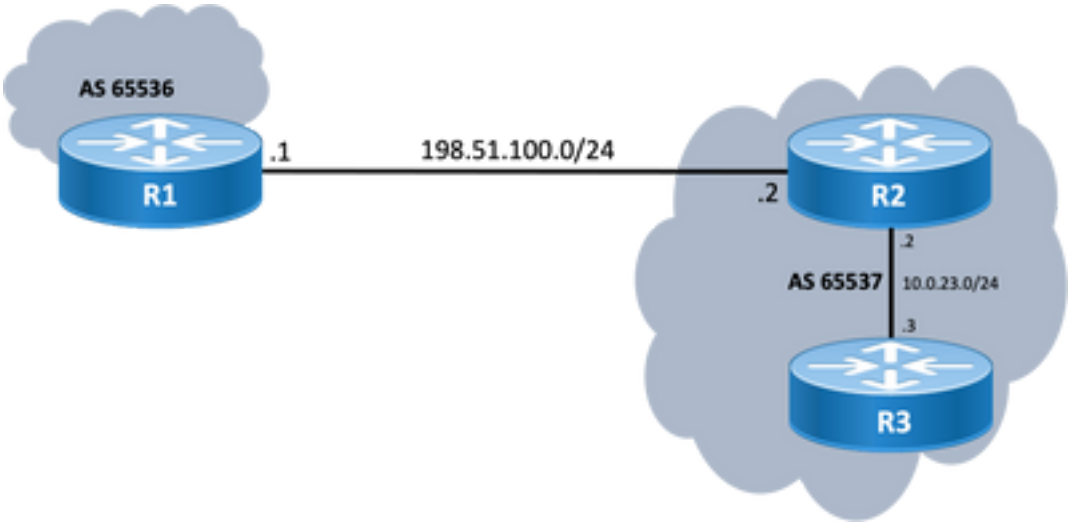
The information in this document was created from the devices in a specific lab environment. All of the devices used in this document started with a cleared (default) configuration. Se a rede estiver ativa, certifique-se de que você entenda o impacto potencial de qualquer comando.

Informações de Apoio

Este documento descreve um guia básico para solucionar os problemas mais comuns no Border Gateway Protocol (BGP), fornece ações corretivas, comandos/depurações úteis para detectar a causa raiz dos problemas e práticas recomendadas para evitar possíveis problemas. Tenha em mente que todas as variáveis e cenários possíveis não podem ser considerados e uma análise mais profunda poderia ser exigida pelo TAC da Cisco.

Topologia

Use este diagrama de topologia como referência para as saídas fornecidas neste documento.



Cenários e problemas

Adjacência para Baixo

Se uma sessão BGP estiver inativa e não for ativada, execute o comando show ip bgp all summary command. Aqui você pode encontrar o status atual da sessão:

- Se a sessão não estiver ativa, o estado pode variar entre IDLE e ATIVE (depende do processo da Máquina de Estado Finito).
- Se a sessão estiver ativa, você verá o número de prefixos recebidos.

```
R2#show ip bgp all summary
For address family: IPv4 Unicast
BGP router identifier 198.51.100.2, local AS number 65537
BGP table version is 19, main routing table version 19
18 network entries using 4464 bytes of memory
```

```
18 path entries using 2448 bytes of memory
1/1 BGP path/bestpath attribute entries using 296 bytes of memory
0 BGP route-map cache entries using 0 bytes of memory
0 BGP filter-list cache entries using 0 bytes of memory
BGP using 7208 total bytes of memory
BGP activity 18/0 prefixes, 18/0 paths, scan interval 60 secs
18 networks peaked at 11:21:00 Jun 30 2022 CST (00:01:35.450 ago)

Neighbor        V           AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
10.0.23.3       4        65537       6       5       19    0    0 00:01:34       18
198.51.100.1    4        65536       0       0        1    0    0 never    Idle
```

Sem conectividade

O primeiro requisito que deve ser assegurado é a conectividade entre os dois peers para que a sessão TCP na porta 179 possa ser estabelecida, estejam eles diretamente conectados ou não. Um ping simples é útil para essa questão. Se o peering for estabelecido entre interfaces de loopback, um loopback para o ping de loopback deverá ser feito. Se um teste de ping for executado sem loopback específico como a interface de origem, o endereço IP da interface física de saída será usado como o endereço IP de origem do pacote em vez do endereço IP de loopback do roteador.

Se o ping não for bem-sucedido, considere estas causas:

- Nenhum par de rota conectado ou nenhuma rota: show ip route *peer_IP_address* pode ser usado.
- Questão da camada 1: a interface física, o SFP (conector), o cabo ou a questão externa (transporte e provedor, se aplicável) precisam ser considerados.
- Verifique qualquer firewall ou lista de acesso que possa bloquear a conexão.

Se o ping tiver êxito, considere isto:

Problemas de configuração

- Endereço IP errado ou AS configurado: para IP errado endereço, essa mensagem não será exibida, mas verifique se a configuração apropriada foi feita. Para AS errado, você deve ver uma mensagem como na show logging comando.

```
%BGP-3-NOTIFICATION: sent to neighbor 198.51.100.1 passive 2/2 (peer in wrong AS) 2 bytes 1B39
```

Verifique a configuração do BGP em ambas as extremidades para corrigir os números AS ou o endereço IP do peer.

- ID do roteador duplicada:

```
%BGP-3-NOTIFICATION: sent to neighbor 198.51.100.1 passive 2/3 (BGP identifier wrong) 4 bytes
0A0A0A0A
```

Verifique o identificador de BGP em ambas as extremidades via show ip bgp all summary e corrigir o problema duplicado, isso pode ser obtido manualmente com o comando global bgp router-id X.X.X.X na configuração do roteador bgp. Como prática recomendada, verifique se o ID do roteador está definido manualmente como um número exclusivo.

- origem BGP e TTL:

A maioria das sessões de iBGP são configuradas nas interfaces de loopback acessíveis através de um IGP. Essa interface de loopback deve ser explicitamente definida como a origem. Faça

isso com o comando neighbor *ip-address* update-source *interface-id* .

Para o peer do eBGP, as interfaces diretamente conectadas são geralmente usadas para peering e há uma verificação para que o Cisco IOS/Cisco IOS-XE cumpra essa finalidade ou não nem mesmo tentar estabelecer uma sessão. Se o eBGP é tentado de loopback para loopback em roteadores conectados diretamente, esta verificação pode ser desabilitada para um vizinho específico em ambas as extremidades via neighbor *ip-address* disable-connected-check .

No entanto, se houver vários saltos entre os peers do eBGP, uma contagem de saltos apropriada será necessária, assegure-se de que **neighbor *ip-address* ebgp-multihop [*hop-count*]** está configurado com a contagem de saltos correta para que a sessão possa ser estabelecida.

Se a contagem de saltos não for especificada, o valor TTL padrão para sessões iBGP é 255, enquanto o valor TTL padrão para sessões eBGP é 1.

Problemas de sessão TCP

Uma ação útil para testar a porta 179 é um telnet manual de um peer para o outro:

```
R1#telnet 198.51.100.2 179
Trying 198.51.100.2, 179 ... Open

[Connection to 198.51.100.2 closed by foreign host]
```

A opção Aberta/conexão fechada ou Conexão recusada pelo host remoto indica que os pacotes atingem a extremidade remota; em seguida, verifique se não há problemas com o plano de controle na extremidade distante. Caso contrário, se houver um Destino inalcançável, verifique qualquer firewall ou lista de acesso que possa bloquear a porta TCP 179 ou pacotes BGP ou qualquer perda de pacote no caminho.

Em caso de problema de autenticação, as mensagens que você pode ver:

```
%TCP-6-BADAUTH: Invalid MD5 digest from 198.51.100.1(179) to 198.51.100.2(20062) tableid - 0
%TCP-6-BADAUTH: No MD5 digest from 198.51.100.1(179) to 198.51.100.2(20062) tableid - 0
```

Verifique os métodos de autenticação, a senha e a configuração relacionada e, para fazer Troubleshooting adicional, consulte Exemplo de Configuração de Autenticação MD5 Entre Pares BGP.

Se a sessão TCP não aparecer, você pode usar os próximos comandos para isolamento:

```
show tcp brief all
show control-plane host open-ports
debug ip tcp transactions
```

Devoluções de adjacências

Se a sessão estiver ativa ou inativa, procure show log e podemos ver alguns cenários.

Flap de Interface

```
%BGP-5-ADJCHANGE: neighbor 198.51.100.2 Down Interface flap
```

Como a mensagem indica, o motivo dessa falha é a situação de inatividade da interface, procure por problemas físicos na porta/SFP, no cabo ou nas desconexões.

Temporizador em Espera Expirado

```
%BGP-3-NOTIFICATION: sent to neighbor 198.51.100.2 4/0 (hold time expired) 0 bytes
```

Éuma situação muito comum; significa que o roteador não recebeu nem processou uma mensagem de keepalive ou qualquer mensagem de atualização antes de o temporizador de espera expirar. O dispositivo envia uma mensagem de notificação e fecha a sessão. Os motivos mais comuns para esse problema estão listados aqui:

  - **Problemas de interface:** Procure erros de entrada, quedas de fila de entrada ou problemas físicos nas interfaces conectadas de ambos os peers; show interface pode ser usado para esse fim.
  - **Perda de pacotes em trânsito:** às vezes, os pacotes Hello podem ser descartados em trânsito, a melhor maneira de garantir que isso seja uma captura de pacotes no nível da interface. Você pode usar o Embedded Packet Capture nos dispositivos Cisco IOS e Cisco IOS-XE. Caso os pacotes sejam vistos no nível da interface, precisamos ter certeza de que eles alcançam o plano de controle, EPC no plano de comando ou debug bgp [vrf name] ipv4 unicast keepalives é útil.
  - **CPU de alto desempenho:** uma condição de CPU de alto nível pode causar quedas no plano de controle, show processes cpu [sorted|history] é útil para identificar o problema. Com base na plataforma, você pode encontrar a próxima etapa para solucionar o problema com o documento Referência da CPU
  - **Problemas de política de CoPP:** a metodologia de identificação e solução de problemas varia para cada plataforma e está fora do escopo deste documento.
- **Incompatibilidade de MTU:** se houver discrepâncias de MTU no caminho e se as mensagens ICMP forem bloqueadas no caminho da origem para o destino, o PMTUD não funcionará e poderá resultar em oscilação de sessão. As atualizações são enviadas com o valor MSS negociado e um bit DF definido. Se um dispositivo no caminho ou mesmo no destino não puder aceitar os pacotes com MTU mais alto, ele enviará uma mensagem de erro ICMP de volta ao alto-falante BGP. O roteador de destino aguarda o keepalive do BGP ou o pacote de atualização do BGP para atualizar seu temporizador de retenção. Você pode verificar o MSS negociado com show ip bgp neighbors *ip_address*.

Um teste de ping para um vizinho específico com o conjunto df pode mostrar se esse MTU é válido ao longo do caminho:

```
ping 198.51.100.2 size max_seg_size df
```

Se forem encontrados problemas de MTU, uma revisão precisa da configuração deverá ser feita para garantir que os valores de MTU sejam consistentes em toda a rede.

> **Nota:** Para obter mais informações sobre MTU, consulte Flaps de Vizinhos BGP com Troubleshooting de MTU .

Problemas AFI/SAFI

```
%BGP-5-ADJCHANGE: neighbor 198.51.100.2 passive Down AFI/SAFI not supported
%BGP-3-NOTIFICATION: received from neighbor 198.51.100.2 active 2/8 (no supported AFI/SAFI) 3
bytes 000000
```

O identificador de família de endereços (AFI) é uma extensão de capacidade adicionada pelo Multi-Protocol BGP (MP-BGP), que se correlaciona a um protocolo de rede específico, como IPv4, IPv6 e similares, e granularidade adicional através de um identificador de família de endereços subsequente (SAFI), como unicast e multicast. O MBGP obtém essa separação por atributos de caminho (PAs) do BGP MP_REACH_NLRI e MP_UNREACH_NLRI. Esses atributos são transportados dentro das mensagens de atualização do BGP e são usados para transportar informações de alcançabilidade de rede para diferentes famílias de endereços.

A mensagem fornece os números desses AFI/SAFI registrados pela IANA:

- Números da família de endereços IANA
- Parâmetros dos Identificadores da Família de Endereços Subsequentes (SAFI)

- Verifique a configuração do BGP para as famílias de endereços planejadas em ambos os lados para corrigir quaisquer famílias de endereços indesejadas.
- Uso neighbor *ip-address* dont-capability-negotiate em ambas as extremidades. Para obter mais informações, consulte Capacidades não suportadas causam mau funcionamento do correspondente BGP.

Instalação e seleção de caminho

Para obter uma explicação melhor sobre como o BGP funciona e selecionar o melhor caminho, consulte Algoritmo de seleção de melhor caminho BGP.

Próximo Salto

Para que uma rota seja instalada em nossa tabela de roteamento, o próximo salto precisa ser alcançável; caso contrário, mesmo que o prefixo esteja em nossa tabela BGP Loc-RIB, ele não entra em RIB. Como regra de prevenção de loop, no Cisco IOS/Cisco IOS-XE, o iBGP não altera o atributo do próximo salto e deixa o AS_PATH sozinho, enquanto o eBGP regrava o próximo salto e antecede o AS_PATH.

Você pode verificar o próximo salto com show ip bgp [prefix], ele lhe dá o próximo salto e palavra inacessível. No exemplo, esse é um prefixo anunciado por R1 via eBGP para R2 e aprendido por R3 via conexão iBGP de R2.

```
R3#show ip bgp 192.0.2.1
BGP routing table entry for 192.0.2.1/32, version 0
Paths: (1 available, no best path)
  Not advertised to any peer
  Refresh Epoch 1
  65536
    198.51.100.1 (inaccessible) from 10.0.23.2 (10.2.2.2)
      Origin incomplete, metric 0, localpref 100, valid, internal
      rx pathid: 0, tx pathid: 0
      Updated on Jul 1 2022 13:44:19 CST
```

Na saída, o próximo salto é a interface de saída de R1, que não é conhecida por R3. Para corrigir essa situação, você pode anunciar o próximo salto via IGP, rota estática ou usar o comando

neighbor *ip-address next-hop-self* no peer iBGP para modificar o IP do próximo salto (que está diretamente conectado). No exemplo do diagrama, essa configuração precisa estar em R2; o vizinho em direção a R3 (vizinho 10.0.23.3 next-hop-self).

Como resultado, o próximo salto muda (após um clear ip bgp 10.0.23.2 soft) para a interface diretamente conectada (acessível) e o prefixo é instalado.

```
R3#show ip bgp 192.0.2.1
BGP routing table entry for 192.0.2.1/32, version 24
Paths: (1 available, best #1, table default)
  Not advertised to any peer
  Refresh Epoch 1
  65536
    10.0.23.2 from 10.0.23.2 (10.2.2.2)
      Origin incomplete, metric 0, localpref 100, valid, internal, best
      rx pathid: 0, tx pathid: 0x0
      Updated on Jul 1 2022 13:46:53 CST
```

Falha de RIB

Isso acontece quando a rota não pode ser instalada no RIB global, o que resulta em uma falha do RIB, motivo comum é quando o mesmo prefixo já está no RIB para outro protocolo de roteamento com distância administrativa mais baixa, mas o motivo exato de uma falha do RIB é visto com o comando show ip bgp rib-failure. Para uma explicação mais detalhada, você pode consultar o link abaixo:

> **Observação**: você pode identificar e corrigir esse problema conforme explicado em Compreender a falha RIB de BGP e O comando bgp suppress-inative.

Condição de corrida

O problema mais comum visto é quando o IGP é preferido sobre o eBGP no cenário de redistribuição mútua. Quando uma rota IGP é redistribuída no BGP, ela é considerada gerada localmente pelo BGP e recebe um peso de 32768 por padrão. Todos os prefixos recebidos de um par BGP recebem um peso local de 0 por padrão. Portanto, se o mesmo prefixo precisar ser comparado, o prefixo com maior peso será instalado na tabela de roteamento com base no processo de seleção do melhor caminho BGP e é por isso que a rota IGP é instalada no RIB.

A solução para esse problema é definir um peso maior para todas as rotas recebidas do peer BGP na configuração de BGP do roteador:

```
neighbor ip-address weight 40000
```

> **Observação**: para obter uma explicação detalhada, consulte Compreender a importância do atributo de caminho de peso BGP em cenários de failover de rede.

Outros problemas

BGP Slow Peer

Éum peer que não consegue acompanhar a taxa em que o remetente gera mensagens de atualização. Há muitas razões para um peer apresentar esse problema: alta utilização de CPU em um dos peers, excesso de tráfego ou perda de tráfego em um link, recurso de largura de banda, entre outros.

> **Observação**: para ajudar a identificar e corrigir problemas de peers lentos, consulte Usar o Recurso "Peer Lento" do BGP para Resolver Problemas de Peer Lento.

Problemas com memória

O BGP usa a memória atribuída ao processo do Cisco IOS para manter os prefixos de rede, os melhores caminhos, as políticas e todas as configurações relacionadas para operar corretamente. Os processos gerais são vistos com o comando show processes memory sorted:

```
R1#show processes memory sorted
Processor Pool Total: 2121414332 Used: 255911152 Free: 1865503180
reserve P Pool Total:     102404 Used:         88 Free:     102316
 lsmpi_io Pool Total:    3149400 Used:    3148568 Free:        832

 PID TTY  Allocated      Freed    Holding    Getbufs    Retbufs Process
   0   0  266231616   81418808  160053760          0          0 *Init*
 662   0   34427640      51720   34751920          0          0 SBC main process
  85   0    9463568          0    8982224          0          0 IOSD ipc task
   0   0   34864888   25213216    8513400    8616279          0 *Dead*
 504   0     696632          0     738576          0          0 QOS_MODULE_MAIN
 518   0     940000       8616     613760          0          0 BGP Router
 228   0     856064     345488     510080          0          0 mDNS
  82   0     547096     118360     417520          0          0 SAMsgThread
   0   0          0          0     395408          0          0 *MallocLite*
```

O pool de processadores é a memória usada; cerca de 2,1 GB no exemplo. Em seguida, devemos examinar a coluna Holding para identificar o subprocesso que contém a maior parte dela. Em seguida, precisamos verificar as sessões de BGP que temos, quantas rotas são recebidas e a configuração usada.

Etapas comuns para reduzir a retenção de memória pelo BGP:

- **Filtragem de BGP:** se não for necessário receber uma tabela de BGP completa, use políticas para filtrar rotas e instalar apenas os prefixos necessários.
- **Reconfiguração suave:** Procure o **vizinho** ***ip_address*** **soft-reconfiguration inbound** ***na configuração BGP; este comando permite que você veja todos os prefixos recebidos antes de qualquer política de entrada (Adj-RIB-in).*** No entanto, essa tabela precisa de cerca de metade da tabela RIB local do BGP atual para armazenar essas informações para que você possa evitar essa configuração, a menos que seja obrigatório ou seus prefixos atuais sejam poucos.

> **Observação**: para obter mais informações sobre como otimizar o BGP, consulte Configurar roteadores BGP para desempenho ideal e consumo de memória reduzido.

Alta utilização da CPU

Os roteadores usam processos diferentes para que o BGP opere. Para verificar se o processo BGP é a causa da alta utilização da CPU, use o comando show process cpu sorted comando.

```
R3#show processes cpu sorted
CPU utilization for five seconds: 0%/0%; one minute: 0%; five minutes: 0%
 PID Runtime(ms)     Invoked      uSecs   5Sec   1Min   5Min TTY Process
 PID Runtime(ms)     Invoked      uSecs   5Sec   1Min   5Min TTY Process
 163          36        1463         24  0.07%  0.00%  0.00%   0 ADJ background
  62          28         132        212  0.07%  0.00%  0.00%   0 Exec
   2          39         294        132  0.00%  0.00%  0.00%   0 Load Meter
   1           0           4          0  0.00%  0.00%  0.00%   0 Chunk Manager
   3          27        1429         18  0.00%  0.00%  0.00%   0 BGP Scheduler
   4           0           1          0  0.00%  0.00%  0.00%   0 RO Notify Timers
  63           4          61         65  0.00%  0.00%  0.00%   0 BGP I/O
  83         924          26      35538  0.00%  0.03%  0.04%   0 BGP Scanner
  96         142       11651         12  0.00%  0.00%  0.00%   0 Tunnel BGP
   7           0           1          0  0.00%  0.00%  0.00%   0 DiscardQ Backgro
```

Aqui estão os processos comuns, as causas e as etapas gerais para superar a alta utilização da CPU devido ao BGP:

- **Roteador BGP:** executa uma vez por segundo para proteger a convergência mais rápida. É um dos segmentos mais importantes, lê as mensagens de atualização do bgp, valida os prefixos/redes e atributos, atualiza a tabela de rede/prefixo por AFI/SAFI e a tabela de atributos, realiza o cálculo do melhor caminho entre muitas outras tarefas.
  Uma rotatividade enorme é um cenário muito comum que leva a essa situação.
- **Scanner BGP:** processo de baixa prioridade que é executado a cada 60 segundos por padrão. Esse processo verifica a tabela de BGP inteira para verificar a acessibilidade do próximo salto e atualiza a tabela de BGP de acordo com o caso haja qualquer alteração para um caminho. Ele é executado através da RIB (Routing Information Base, base de informações de roteamento) para fins de redistribuição.
  Verifique a escala da plataforma, à medida que mais prefixos e rotas são instalados e o TCAM é usado, mais recursos são necessários e, geralmente, um dispositivo sobrecarregado leva a essas situações.

  Observação: para obter mais informações sobre como solucionar problemas desses dois processos, consulte Solucionar problemas de alta CPU causados pelo scanner BGP ou pelo processo do roteador.

- **E/S BGP:** executa quando os pacotes de controle BGP são recebidos e gerencia o enfileiramento e o processamento de pacotes BGP. Se houver pacotes excessivos recebidos na fila de BGP por um longo período, ou se houver um problema com o TCP, o roteador mostra sintomas de alta utilização da CPU devido ao processo de E/S do BGP. (Geralmente, o Roteador BGP também é alto nessa situação. Examine as contagens de mensagens para identificar os pacotes correspondentes e de captura para identificar a origem dessas mensagens.)
- **BGP Open:** processo usado no estabelecimento da sessão. Não é um problema comum de alta utilização da CPU, a menos que a sessão esteja travada no estado Open (Aberto).
- **Evento BGP:** é responsável pelo processamento do próximo salto. Procure oscilações de próximos saltos nos prefixos recebidos.

Informações Relacionadas

- Suporte Técnico e Documentação - Cisco Systems

- Manual de configuração de BGP
- Exemplo de Configuração de Autenticação MD5 Entre Pares BGP
- Captura de pacotes incorporada
- Flaps de vizinhos BGP com Troubleshooting de MTU
- Números da família de endereços IANA
- Parâmetros dos Identificadores da Família de Endereços Subsequentes (SAFI)
- Recursos não suportados causam mau funcionamento do correspondente BGP
- Algoritmo de seleção de melhor caminho BGP
- Entender a falha RIB do BGP e o comando bgp suppress-inative
- Entender a importância do atributo de caminho de peso BGP em cenários de failover de rede
- Usar o recurso "Slow Peer" do BGP para resolver problemas de peer lento
- Configurar roteadores BGP para desempenho ideal e consumo de memória reduzido
- Solucionar problemas de alta utilização da CPU causados pelo scanner BGP ou pelo processo do roteador

Sobre esta tradução
A Cisco traduziu este documento com a ajuda de tecnologias de tradução automática e humana para oferecer conteúdo de suporte aos seus usuários no seu próprio idioma, independentemente da localização.

Observe que mesmo a melhor tradução automática não será tão precisa quanto as realizadas por um tradutor profissional.

A Cisco Systems, Inc. não se responsabiliza pela precisão destas traduções e recomenda que o documento original em inglês (link fornecido) seja sempre consultado.