

Use o recurso "Slow Peer" do BGP para resolver problemas lentos do peer

Contents

[Introduction](#)

[Informações de Apoio](#)

[Atualizar grupos](#)

[Problema](#)

[Solução](#)

[Detecção](#)

[Identificação de peer lento](#)

[Movimento](#)

[Movimento sem recurso de peer lento](#)

[Movimento de peer lento e estático](#)

[Movimento dinâmico de peer lento](#)

[Recuperação](#)

[Limpe o status do peer lento](#)

Introduction

Este documento descreve como resolver um problema de peer lento com o uso do *recurso* BGP (Border Gateway Protocol) de *peer lento*, que identifica um peer lento em um grupo de atualização BGP e pode mover o peer lento para fora do grupo de atualização permanente ou temporariamente.

Informações de Apoio

Esta seção fornece uma visão geral do recurso de peer lento e do uso de grupos de atualização.

Atualizar grupos

O recurso de peer lento é usado em grupos de atualização. Um grupo de atualização é um método dinâmico que é usado para agrupar peers BGP com a mesma política de saída. O benefício dos grupos de atualização é que a política de grupo é usada para formatar mensagens uma vez, e depois elas são replicadas e transmitidas aos outros membros do grupo. Esse método é mais eficiente do que a necessidade de formatar atualizações BGP para cada peer separadamente.

Quando esse método é implementado, se a política de saída for alterada, os grupos de peer

mudam por grupo de atualização. Os grupos de atualização são formados por família de endereços (AF).

Aqui está um exemplo de dois peers BGP em diferentes grupos de atualização para unicast AF IPv4, mas com o mesmo grupo de atualização para AF VPNv4:

```
R2#show ip bgp update-group
```

```
BGP version 4 update-group 1, external, Address Family: IPv4 Unicast  
Has 1 member (* indicates the members currently being sent updates):  
 10.1.3.4
```

```
BGP version 4 update-group 2, external, Address Family: IPv4 Unicast  
Has 1 member (* indicates the members currently being sent updates):  
 10.1.2.3
```

```
R2#show ip bgp vpnv4 all update-group
```

```
BGP version 4 update-group 1, external, Address Family: VPNv4 Unicast  
Has 2 members (* indicates the members currently being sent updates):  
 10.1.2.3          10.1.3.4
```

O grupo de atualização torna-se mais eficiente à medida que o número de peers BGP incluídos no grupo de atualização aumenta. Normalmente, os peers BGP internos (iBGP) têm a mesma política de saída. Para o iBGP, um refletor de rota (RR) pode ter muitos peers iBGP; portanto, ele terá grandes grupos de atualização. Os roteadores Provider Edge (PE) podem ter muitos peers BGP externos (eBGP) em direção aos roteadores Customer Edge (CE) em um Virtual/Routing Forwarding (VRF). Os roteadores PE também podem ter grandes grupos de atualização para os pares com roteadores CE nas interfaces VRF.

Problema

Um peer lento é um peer que não pode acompanhar a taxa na qual o roteador gera mensagens de atualização de BGP durante um período prolongado (em ordem de minutos) em um grupo de atualização. O motivo para isso pode ser problemas de rede persistentes. Os motivos da rede podem ser a perda de pacotes e/ou links carregados, ou problemas de throughput com as sessões BGP. Além disso, um peer BGP pode estar muito carregado em termos de CPU e não pode atender à conexão TCP na velocidade necessária.

Os peers lentos afetam a convergência de BGP do grupo de atualização completo. Se um peer de BGP estiver lento, isso fará com que todo o grupo de atualização fique lento. O resultado é que os outros membros do grupo de atualização também terão uma convergência mais lenta. Por isso, a questão deve ser resolvida.

Você pode identificar o peer lento e movê-lo para fora do grupo de atualização. Para concluir esta tarefa, você pode alterar a política de saída para esse peer BGP; no entanto, esta é uma tarefa manual. Primeiro, você deve identificar o peer que está lento e depois movê-lo para fora do grupo de atualização. O recurso de peer lento pode fazer isso automaticamente, para que não seja necessária nenhuma intervenção do usuário.

Solução

Há três partes no recurso peer lento:

- Detecção de peer lento
- Movimento do peer lento em um grupo de atualização lento
- Recuperação do peer lento (que move o peer recuperado de volta para seu grupo de atualização original)

Esses processos são descritos com mais detalhes nas seções a seguir.

Detecção

O recurso de peer lento detecta peers lentos em um grupo de atualização. Cada grupo de atualização tem uma fila de cache, onde as atualizações de BGP formatadas são temporariamente armazenadas antes da transmissão.

Aqui está um exemplo desse cache de grupo de atualização:

```
R2#show ip bgp replication
```

Index	Members	Leader	MsgFmt	MsgRepl	Csize	Current Version	Next Version
1	1	10.1.1.1	0	0	0/100	6/0	
2	3	10.1.2.3	2	6	0/1000	6/0	
3	1	10.1.2.6	3	0	0/100	6/0	

O tamanho do cache é calculado dinamicamente e depende de:

- O número de pontos no grupo de atualização
- A memória do sistema instalada
- O tipo de peer no grupo de atualização
- O tipo de AF

O número de atualizações de BGP formatadas que aguardam transmissão pode ser criado em um grupo de atualização quando um peer (o lento) não reconhece as mensagens de BGP tão rapidamente quanto os outros membros. Quando o limite de cache é atingido, o grupo não tem mais cota para colocar novas mensagens em fila. Nenhuma mensagem nova pode ser formatada até que o cache seja reduzido (até que algumas mensagens sejam confirmadas pelos pares lentos). Isso proíbe o peer BGP e não permite que ele envie novas mensagens (atualizações ou retiradas) para os membros mais rápidos do grupo. Assim, isso retarda a convergência de todos os pares no grupo de atualização.

Para que o recurso de peer lento identifique um peer lento, ele se refere aos timestamps de atualização do BGP e aos parâmetros TCP do peer.

Por padrão, a detecção lenta do peer é desativada. Para habilitar a detecção de peer lenta, use um destes métodos:

- Ative o recurso para o processo BGP (pode ser configurado em AF/VRF):

```
bgp slow-peer detection [threshold
```

Note: O valor limite pode variar entre 120 e 3.600 segundos, e o valor padrão é 300 segundos.

- Ative o recurso por peer:

```
neighbor {
```

- Ative o recurso por meio do modelo de política de peer:

```
slow-peer detection [threshold < seconds >]
```

```
[no] slow-peer detection
```

Quando um peer lento é detectado, uma mensagem de syslog semelhante a esta é impressa:

```
%BGP-5-SLOWPEER_DETECT: Neighbor IPv4 Unicast 10.1.6.7 has been detected as a slow peer.
```

Você pode inserir estes comandos **show** para exibir os pares lentos:

- **show ip bgp summary slow**
- **show ip bgp neighbors low**
- **show ip bgp update-group summary slow**

Aqui está um exemplo de saída do comando **show** quando a palavra-chave **lento** é usada:

```
R2#show ip bgp update-group summary slow
Summary for Update-group 1, Address Family IPv4 Unicast
Summary for Update-group 2, Address Family IPv4 Unicast
Summary for Update-group 3, Address Family IPv4 Unicast
Summary for Update-group 4, Address Family IPv4 Unicast
BGP router identifier 10.1.6.2, local AS number 2
BGP table version is 966013, main routing table version 966013
BGP main update table version 966013
50000 network entries using 6050000 bytes of memory
50000 path entries using 2600000 bytes of memory
5001/5000 BGP path/bestpath attribute entries using 700140 bytes of memory
5000 BGP AS-PATH entries using 183632 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 9533772 total bytes of memory
BGP activity 208847/158847 prefixes, 508006/458006 paths, scan interval 60 secs
Neighbor      V      AS MsgRcvd MsgSent  TblVer  InQ  OutQ Up/Down  State/PfxRcd
10.1.6.7      4       7    165   50309      0     0   100 00:10:35      0
```

Como mostrado na saída, o peer **10.1.6.7** é um peer lento para o unicast AF IPv4. Os outros AFs não mostram nenhum par lento.

Para verificar se o temporizador de detecção está em execução no momento e seu valor, insira este comando:

```
R2#show ip bgp update-group
BGP version 4 update-group 3, external, Address Family: IPv4 Unicast
BGP Update version : 116013/0, messages 164 queue 164, not converged
Private AS number removed from updates to this neighbor
Update messages formatted 5948, replicated 11589
Number of NLRIs in the update sent: max 249, min 1
Minimum time between advertisement runs is 30 seconds
Slow-peer detection timer (expires in 111 seconds)
  Has 3 members (* indicates the members currently being sent updates):
  10.1.4.5          10.1.5.6          10.1.6.7
```

Como mostrado na saída do exemplo, o temporizador de detecção foi iniciado. O temporizador de detecção inicia quando o cache do grupo de atualização está cheio.

Neste exemplo, você pode ver que um peer lento é detectado, mas ele só sai do grupo de atualização depois que o temporizador de detecção de peer lento expira:

```
R2#show ip bgp update-group
&&|
BGP version 4 update-group 3, external, Address Family: IPv4 Unicast
BGP Update version : 516013/566013, messages 357 queue 357, not converged
Private AS number removed from updates to this neighbor
Update messages formatted 27044, replicated 53645
Number of NLRIs in the update sent: max 249, min 0
Minimum time between advertisement runs is 30 seconds
Slow-peer detection timer (expires in 20 seconds)
Has 3 members (* indicates the members currently being sent updates)
(1 dynamically detected as slow):
  *10.1.4.5          *10.1.5.6          10.1.6.7
```

Identificação de peer lenta

Se o recurso de detecção de peer lento não estiver habilitado, você deve identificar o peer lento manualmente. Primeiro, verifique a versão da tabela e a fila de saída dos peers no grupo de atualização:

```
R2#show ip bgp update-group 3 summary
Summary for Update-group 3, Address Family IPv4 Unicast
BGP router identifier 10.1.6.2, local AS number 2
BGP table version is 552583, main routing table version 552583
BGP main update table version 552583
37870 network entries using 4582270 bytes of memory
37870 path entries using 1969240 bytes of memory
5002/3788 BGP path/bestpath attribute entries using 700280 bytes of memory
5001 BGP AS-PATH entries using 183656 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 7435446 total bytes of memory
BGP activity 158847/108847 prefixes, 295876/258006 paths, scan interval 60 secs
Neighbor          V    AS MsgRcvd MsgSent  TblVer  InQ  OutQ  Up/Down  State/PfxRcd
```

10.1.4.5	4	5	77	26840	516013	0	0	01:07:12	0
10.1.5.6	4	6	69	26833	516013	0	0	01:00:30	0
10.1.6.7	4	7	79	26761	516013	0	194	00:45:42	0

Neste exemplo, verifique se a versão da tabela (**TbVer**) dos peers alcançou a versão principal da tabela BGP ou se está sempre atrás. Em segundo lugar, verifique se há um ou mais peers com valores de fila de saída muito altos. É provável que estes sejam os pares lentos.

Quando você visualizar a suspeita de peer BGP lento, considere estas perguntas (em ambos os lados da sessão BGP):

- Há quanto tempo a *última gravação* foi executada?
- Os keepalives estão em aceleração?
- A fila de saída está alta?
- O *SRTT/RTTO* está alto?
- O número de retransmissões aumenta?
- Há algum pacote de retransmissão em fila?
- A janela de *envio TCP* é muito baixa ou zero?

Aqui está um exemplo:

```
R2#show ip bgp neighbors 10.1.6.7
BGP neighbor is 10.1.6.7, remote AS 7, external link
Member of peer-group group3 for session parameters
  BGP version 4, remote router ID 10.1.6.7
  BGP state = Established, up for 00:56:09
  Last read 00:00:43, last write 00:00:17, hold time is 180, keepalive interval
  is 60 seconds
Keepalives are temporarily in throttle due to closed TCP window
Neighbor capabilities:
  Route refresh: advertised and received(new)
  Address family IPv4 Unicast:
advertised and received
Message statistics
  InQ depth is 0
  OutQ depth is 0      Partial message pending
                        Sent      Rcvd
Opens:                5         4
Notifications:       0         0
Updates:              29004      0
Keepalives:          0         1426
Route Refresh:        0         0
Total:                30336      1431
Default minimum time between advertisement runs is 30 seconds
For address family: IPv4 Unicast
BGP table version 250001, neighbor version 200001/250001
Output queue size : 410
Index 3, Offset 0, Mask 0x8
3 update-group member
group3 peer-group member
Inbound soft reconfiguration allowed
Private AS number removed from updates to this neighbor
```

Inbound path policy configured

Route map for incoming advertisements is eBGP-in

	Sent	Rcvd
Prefix activity:	----	----
Prefixes Current:	2596	0
Prefixes Total:	102624	0
Implicit Withdraw:	28	0
Explicit Withdraw:	100000	0
Used as bestpath:	n/a	0
Used as multipath:	n/a	0
	Outbound	Inbound
Local Policy Denied Prefixes:	-----	-----
Total:	0	0

Maximum prefixes allowed 20000

Threshold for warning message 80%, restart interval 300 min

Number of NLRIs in the update sent: max 249, min 0

Last detected as dynamic slow peer: never

Dynamic slow peer recovered: never

Oldest update message was formatted: 00:02:24

Address tracking is enabled, the RIB does have a route to 10.1.6.7

Connections established 4; dropped 3

Last reset 00:57:39, due to User reset

Transport(tcp) path-mtu-discovery is enabled

Connection state is ESTAB, I/O status: 1, unread input bytes: 0

Connection is ECN Disabled

Minimum incoming TTL 0, Outgoing TTL 1

Local host: 10.1.6.2, Local port: 20298

Foreign host: 10.1.6.7, Foreign port: 179

Connection tableid (VRF): 0

Enqueued packets for retransmit: 15, input: 0 mis-ordered: 0 (0 bytes)

Event Timers (current time is 0x4A63D14):

Timer	Starts	Wakeups	Next
Retrans	697	29	0x4A6590C
TimeWait	0	0	0x0
AckHold	64	63	0x0
SendWnd	0	0	0x0
KeepAlive	0	0	0x0
GiveUp	0	0	0x0
PmtuAger	128	127	0x4A64CB7
DeadWait	0	0	0x0
Linger	0	0	0x0

iss: 130287252 snduna: 131516888 sndnxt: 131532233 **sndwnd: 16384**

irs: 1184181084 rcvnxt: 1184182346 rcvwnd: 15123 delrcvwnd: 1261

SRTT: 20122 ms, RTTO: 20440 ms, RTV: 318 ms, KRTT: 0 ms

minRTT: 20028 ms, maxRTT: 20796 ms, ACK hold: 200 ms

Status Flags: none

Option Flags: nagle, path mtu capable, higher precedence

Datagrams (max data segment is 1460 bytes):

Rcvd: 922 (out of order: 0), with data: 65, total data bytes: 1261

Sent: 1463 (**retransmit: 29 fastretransmit: 1**), with data: 1391, total data bytes: 1245129

Movimento

Esta seção descreve o processo de movimento em relação ao recurso peer lento em vários cenários.

Movimento sem recurso de peer lento

Um peer lento pode ser movido manualmente para um novo grupo de atualizações *sem* o recurso de peer lento.

Antes do recurso de peer lento estar disponível, você era solicitado a identificar o peer lento e depois movê-lo para fora do grupo de atualização manualmente. Isso é concluído com uma alteração na política de saída desse peer BGP. Essa política de saída deve ser diferente de qualquer outra que seja usada, pois você deve garantir que o peer lento não se mova para outro grupo de atualização que existe atualmente (e mova o problema para esse grupo de atualização). A melhor alteração que você pode aplicar é aquela que não afeta a política real. Por exemplo, você pode alterar o MRAI (Minimum Route Advertisement Interval, Intervalo mínimo de anúncio de rota) do peer (sob o AF específico).

Aqui está um exemplo que mostra o movimento manual de um peer lento quando o recurso de peer lento não está disponível:

```
RR1#debug ip bgp groups
BGP groups debugging is on

RR1(config)#router bgp 1
RR1(config-router)#address-family vpnv4
RR1(config-router-af)#neighbor 10.100.1.3 advertisement-interval 3

BGP-DYN(4): 10.100.1.3 cannot join update-group 1 due to an advertisement-interval
mismatch
BGP(4): Scheduling withdraws and update-group membership change for 10.100.1.3
BGP(4): Resetting 10.100.1.3's version for its transition out of update-group 1
BGP-DYN(4): 10.100.1.3 cannot join update-group 1 due to an advertisement-interval
mismatch
BGP-DYN(4): Removing 10.100.1.3 from update-group 1
BGP-DYN(4): 10.100.1.3 cannot join update-group 1 due to an advertisement-interval
mismatch
BGP-DYN(4): Created update-group 0 from neighbor 10.100.1.3
BGP-DYN(4): Adding 10.100.1.3 to update-group 0
```

Movimento de peer lento e estático

Para mover um peer de um grupo de atualização para um novo grupo de atualização, você pode configurá-lo como um peer *estático* lento. Se houver vários peers lentos, os peers lentos estáticos com a mesma política de saída serão colocados no mesmo grupo de atualização *lento*.

Para mover um peer lento estaticamente, você pode configurá-lo com o uso destes comandos:

- Habilitar movimento de peer estático por vizinho ou por grupo de peer:

```
[no] neighbor {
```

- Habilitar movimentação de peer estático por meio de um modelo de política de peer:

```
[no] slow-peer split-update-group static
```


Movimento dinâmico de peer lento

Por padrão, o movimento lento de peer é desativado. Para habilitar o movimento lento de peer, você pode configurá-lo através de um destes métodos:

- Habilite o movimento lento do peer para o processo BGP:

```
bgp slow-peer split-update-group dynamic [permanent]
```

```
[no] bgp slow-peer split-update-group dynamic
```

Note: Isso pode ser configurado na exibição *address-family/topology/VRF*.

- Habilitar movimento lento de peer por peer:

```
neighbor {
```

- Ativar a movimentação lenta de pares através de um modelo de política de pares:

```
slow-peer split-update-group dynamic [permanent]
```

```
[no] slow-peer split-update-group dynamic
```

Note: A palavra-chave **permanente** indica que o peer lento não se recuperará automaticamente. Nesse caso, você pode mover o peer lento recuperado de volta para seu grupo de atualização original por meio de um dos [comandos clear](#).

Os peers lentos estáticos e os peers lentos dinâmicos estão no mesmo grupo de atualização de peer lento. Neste exemplo, você pode ver um peer lento em um grupo de atualização lenta:

```
R2#show ip bgp update-group
```

```
^
```

```
BGP version 4 update-group 4, external, Address Family: IPv4 Unicast
```

```
BGP Update version : 0/566013, messages 100 queue 100, not converged
```

Slow update group

```
Private AS number removed from updates to this neighbor
```

```
Update messages formatted 2497, replicated 0
```

```
Number of NLRI's in the update sent: max 10, min 1
```

```
Minimum time between advertisement runs is 30 seconds
```

```
Has 1 member (* indicates the members currently being sent updates)
```

```
(1 dynamically detected as slow):
```

```
*10.1.1.6.7
```

Recuperação

Um peer lento pode ser agrupado sob seu grupo de atualização original (que corresponde à política de saída) quando for confirmado que ele não é mais um peer lento (ele é capturado). O temporizador de recuperação começa quando o grupo de atualização de peer lento convergiu. Quando o temporizador de recuperação expira, o peer lento é movido de volta para o grupo de atualização regular.

Note: Para ver o comportamento relacionado ao temporizador *de detecção/recuperação*, insira o comando **debug ip bgp updates events**.

Quando um peer lento é movido de volta para o grupo de atualização original (isso significa uma recuperação), uma mensagem de syslog semelhante a esta é impressa:

```
%BGP-5-SLOWPEER_RECOVER: Slow peer IPv4 Unicast 10.1.6.7 has recovered.
```

Para verificar se o temporizador de recuperação está em execução no momento e o valor, insira este comando:

```
R2#show ip bgp update-group
BGP version 4 update-group 1, external, Address Family: IPv4 Unicast
BGP Update version : 165973/0, messages 0 queue 0, converged
Route map for outgoing advertisements is dummy
Update messages formatted 0, replicated 0
Number of NLRIs in the update sent: max 0, min 0
Minimum time between advertisement runs is 30 seconds
Slow-peer recovery timer (expires in 16 seconds)
  Has 1 member (* indicates the members currently being sent updates):
  10.1.1.1
```

Neste exemplo, o **temporizador de recuperação**, com um valor de **16 segundos**, indica que um peer possivelmente lento pode voltar ao seu grupo de atualização original em 16 segundos.

Neste exemplo, você pode ver um peer que se recuperou do status de peer lento:

```
R2#show ip bgp neighbor 10.1.6.7
BGP neighbor is 10.1.6.7, remote AS 7, external link
Member of peer-group group3 for session parameters
  BGP version 4, remote router ID 10.1.6.7
  &&|
  3 update-group member
  group3 peer-group member
  &&|
Number of NLRIs in the update sent: max 249, min 0
Last detected as dynamic slow peer: 00:12:49
Dynamic slow peer recovered: 00:01:57
Oldest update message was formatted: 00:00:55
```

Limpe o status do peer lento

O status lento do peer pode ser limpo manualmente com estes comandos:

- **clear ip bgp * low**
- **clear ip bgp AF {unicast/multicast} <número AS> lento**
- **clear ip bgp AF {unicast/multicast} peer-group <group-name> devagar**
- **clear ip bgp <neighbor-address> devagar**

- **clear bgp AF {unicast/multicast} * lenta**
- **clear bgp AF {unicast/multicast} <número AS> lenta**
- **clear bgp AF {unicast/multicast} peer-group <group-name> devagar**
- **clear bgp AF {unicast/multicast} <neighbor-address> lento**

Note: Ao usar esses comandos, substitua **AF** pela família de endereços real.

Com o uso desses comandos, o peer é movido de volta para o grupo de atualização original.

Insira o comando **show ip bgp internal** para exibir as configurações de detecção e movimento lentas de peer:

```
R2#show ip bgp internal
Time left for bestpath timer: 593 secs
Address-family IPv4 Unicast, Mode : RW
  Table Versions : Current 622091, RIB 622091
  Start time : 00:00:01.168      Time elapsed 01:21:56.740
  First Peer up in : 00:00:07      Exited Read-Only in : 00:02:16
  Done with Install in : 00:02:26      Last Update-done in : never
  0 updates expanded
  Attribute list queue size: 0
Slow-peer detection is enabled Threshold is 300 seconds
Slow-peer split-update-group dynamic is enabled
BGP Nexthop scan:-
  penalty: 0, Time since last run: never, Next due in: none
  Max runtime : 0 ms Latest runtime : 0 ms Scan count: 0
BGP General Scan:-
  Max runtime : 14572 ms Latest runtime : 14572 ms Scan count: 78
BGP future scanner version: 79
BGP scanner version: 0
```

Note: Em resumo, o peer lento do BGP é um recurso que detecta um peer lento em um grupo de atualização do BGP e permite uma convergência de BGP mais rápida com o movimento do peer lento fora do grupo de atualização.