

# 请使用BGP“慢对等体”功能解决缓慢的对等体问题

## 目录

[简介](#)

[背景信息](#)

[更新组](#)

[问题](#)

[解决方案](#)

[检测](#)

[减慢对等体识别](#)

[移动](#)

[没有缓慢的对等体功能的移动](#)

[静态缓慢的对等体移动](#)

[动态请减慢对等体移动](#)

[恢复](#)

[清除缓慢的对等体状态](#)

## 简介

本文描述如何解决与使用的一缓慢的对等体问题边界网关协议(BGP)缓慢的对等体功能，在BGP更新组中识别一慢对等体，并且能永久或临时地移动慢对等体在更新组外面。

## 背景信息

此部分提供缓慢的对等体功能的概述和使用更新组。

## 更新组

缓慢的对等体功能用于更新组。更新组是使用为了分组有同样出局策略的BGP对等体的一个动态方法。更新组的好处是组策略用于为了一次格式化消息，他们然后复制并且传送到组的其他组员。此方法比需要是更有效的分开格式化每对等体的BGP更新。

当此方法实现，如果出局策略变更，对等组每更新组更改。更新组每个地址家族(AF)被建立。

这是两个BGP对等体示例用不同的更新组AF IPv4单播的，但是有AF Vpvn4的同一更新组的：

```
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
```

更新组变为更有效的作为在更新组增加包括的编号BGP对等体。一般，内部BGP (iBGP)对等体有同一出局策略。对于iBGP，路由反射器(RR)能有许多iBGP对等体;因此，它将有更大更新组。服务商边缘路由器能有往用户边缘(CE)路由器的许多外部BGP (EBGP)对等体一个虚拟/路由转发的(VRF)。PE路由器能有大更新组对等互连的用VRF接口的CE路由器。

## 问题

一慢对等体是不能跟上速率路由器很久生成BGP更新消息时间的对等体(按分钟的顺序)在更新组中。对此的原因可以是不变网络问题。网络原因能是包丢失和已加载链路或者吞吐量问题与BGP会话。并且，BGP对等体也许大量地装载根据CPU，并且不能服务TCP连接以需要的速度。

慢对等体影响完整更新组的BGP收敛。如果一个BGP对等体慢，它造成整个更新组减速。结果是其他更新组成员将有缓慢的聚合。为此，问题应该是解决的。

您能识别慢对等体和移动它在更新组外面。为了完成此任务，您能更改该BGP对等体的出局策略;然而，这是一手工的任务。您必须首先识别慢的对等体，然后移动它在更新组外面。缓慢的对等体功能能自动地执行此，因此用户干涉没有要求。

## 解决方案

有三部分对缓慢的对等体功能：

- 慢对等体的检测
- 慢对等体的移动到一缓慢的更新组里
- 移动被恢复的对等体回到其原始更新组)的恢复慢对等体(这些进程在跟随的部分的更详细的资料描述。

## 检测

缓慢的对等体功能在更新组中检测慢对等体。每更新组有一个高速缓冲存储队列，被格式化的BGP更新在发射前临时地存储。

这是这样更新组缓存示例：

```
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	

缓存的大小动态地计算并且取决于：

- 对等体数量在更新组中
- 已安装系统内存
- 对等体种类在更新组中
- AF种类

等候发射被格式化的BGP更新的数量能在一更新组中构件，当一对等体(慢一个)时一样迅速不确认BGP消息象其他成员。当缓存限制达到时，组没有排队的配额新建的消息。新的消息不可以被格式化，直到减少缓存(直到一些消息由慢对等体确认)。这禁止BGP对等体，并且不允许它传送新建的信息(更新或让步)到组的更加快速的组员。因此，这在更新组中减速所有对等体收敛。

为了缓慢的对等体功能能识别一慢对等体，它是指BGP更新时间戳和对等体TCP参数。

默认情况下缓慢的对等体检测禁用。为了启动缓慢的对等体检测，请使用这些方法之一：

- 启用BGP进程的功能(能从AF/VRF配置)：

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

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

**Note:** 阈值能排列在120和3,600秒之间，并且默认值是300秒。

- 启用功能每对等体：

```
neighbor {<nbr-addr>/<peer-grp-name>} slow-peer detection [threshold < seconds >]
```

```
[no] neighbor {<nbr-addr>/<peer-grp-name>} slow-peer detection
```

- 通过对等体策略模板启用功能：

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

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

当一慢对等体检测时，系统消息类似于此打印：

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

您能输入这些显示命令为了查看慢对等体：

- 慢的show ip bgp summary
- 慢的show ip bgp neighbors
- show ip bgp慢更新组的摘要

这是示例show命令输出，当使用时缓慢的关键字：

```
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
```

如输出所显示，对等体10.1.6.7是AF IPv4单播的一慢对等体。另一AFs不显示其中任一慢对等体。

为了验证当前检测计时器运行和其值，是否输入此命令：

```
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
```

如示例输出所显示，检测计时器开始。当更新组缓存全双工，检测计时器启动。

在本例中，您能看到一慢对等体检测，但是只搬出更新组，在慢对等体检测计时器超时后：

```
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
```

## 减慢对等体识别

如果缓慢的对等体检测功能没有启用，则您必须手工识别慢对等体。首先，请检查表格版本和对等体的输出队列在更新组中：

```

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

```

在本例中，请验证表格版本(TblVer)对等体是否跟上主要BGP表版本或是否总是后边。其次，请检查一个或更多对等体与非常高产队列值。很可能这些是慢对等体。

当您查看怀疑的慢BGP对等体时，请考虑这些问题(在BGP会话的两边)：

- 前多久是为时写入实行？
- Keepalive在节流孔？
- 输出队列是否是高？
- SRTT/RTTO是否是高？
- 编号重新传输增加？
- 有没有任何排队重传数据包？
- TCP发送窗口是否是非常低或零？

示例如下：

```

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&colon; 65, total data bytes: 1261  
Sent: 1463 (retransmit: 29 fastretransmit: 1),with data&colon; 1391, total  
data bytes: 1245129
```

## 移动

此部分关于缓慢的对等体功能在各种情况下描述移动进程。

### 没有缓慢的对等体功能的移动

一慢对等体可以手工移动到一新的更新组里，不用缓慢的对等体功能。

在缓慢的对等体功能是可用的前，您要求识别慢对等体手工然后移动它在更新组外面。这完成与对该BGP对等体出局策略的一更改。此出局策略跟使用的任何其他一定不同，您必须保证慢对等体不移动向当前存在的另一更新组(和请移动问题向该更新组)。您能应用的最好的更改是不影响实际策略的一个。例如，您可能更改最低的路由通告间隔(MRAI)对等体(在特定AF下)。

这是显示一慢对等体的手工的移动的示例，当缓慢的对等体功能不是可用的时：

```
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
```

### 静态缓慢的对等体移动

当静态减慢对等体，为了移动从更新组的一对等体到一新的更新组里，您能配置它。如果有广泛慢对等体，则放置有同一出局策略的静态慢对等体到同一缓慢的更新组里。

为了静态移动一慢对等体，您能用使用这些命令配置它：

- 启用静态对等体移动每个邻居或每对等组：

```
[no] neighbor {<nbr-addr>/<peer-grp-name>} slow-peer split-update-group static
```

- 通过对等体策略模板启用静态对等体移动：

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

## 动态请减慢对等体移动

默认情况下缓慢的对等体移动禁用。为了启用缓慢的对等体移动，您能通过这些方法之一配置它：

- BGP进程的Enable (event)缓慢的对等体移动：

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

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

**Note:**这可以从*address-family/topology/VRF*视图配置。

- Enable (event)缓慢的对等体移动每对等体：

```
neighbor {<nbr-addr>/<peer-grp-name>} slow-peer split-update-group dynamic [permanent]
```

```
[no] neighbor {<nbr-addr>/<peer-grp-name>} slow-peer split-update-group dynamic
```

- Enable (event)缓慢的对等体移动通过对等体策略模板：

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

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

**Note:**永久性关键字表明慢对等体不会自动地恢复。在这种情况下，您能移动被恢复的慢对等体回到其原始更新组通过其中一个[清除命令](#)。

静态慢对等体和动态慢对等体是在同一缓慢的对等体更新组中。在本例中您在一缓慢的更新组中能看到一慢对等体：

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

```
â€”|
```

```
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 NLRIs 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
```

## 恢复

一慢对等体可以被重新组合在其原始更新组下(该匹配出局策略)，一旦被确认它不再是一慢对等体(追上)。当缓慢的对等体更新组聚合，恢复计时器启动。当恢复计时器超时时，慢对等体移动回到定期更新组。

**Note:**为了看到与检测/恢复计时器涉及的行为，请输入**debug ip bgp updates**事件命令。

当一慢对等体移动回到原始更新组(这含义一恢复)，系统消息类似于此打印：

```
%BGP-5-SLOWPEER_RECOVER: Slow peer IPv4 Unicast 10.1.1.6.7 has recovered.
```

为了验证当前恢复计时器运行和值，是否输入此命令：



```
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
```

在本例中，恢复计时器，有值的16秒，表明一可能慢对等体也许移动回到其原始更新组在16秒。

在本例中，您能看到从缓慢的对等体状态恢复的对等体：

```
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
```

## 清除缓慢的对等体状态

缓慢的对等体状态可以用这些命令手工清除：

- 慢的clear ip bgp \*
- clear ip bgp AF {单播|组播} <AS的number>
- clear ip bgp AF {单播|组播}慢对等组的<group name>
- 慢clear ip bgp的<neighbor-address>
- clear bgp AF {单播|组播} \*请减慢
- clear bgp AF {单播|组播} <AS的number>
- clear bgp AF {单播|组播}慢对等组的<group name>
- clear bgp AF {单播|组播}的<neighbor-address>

**Note:**当您使用这些命令时，请用实际地址家族替换AF。

使用使用这些命令，对等体移动回到原始更新组。

输入show ip bgp内部命令为了查看缓慢的对等体检测和移动设置：

```
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:**总之，BGP慢对等体是在BGP更新组中检测一慢对等体并且允许与慢对等体移动的更加快速的BGP收敛在更新组外面的功能。