

BGP の「Slow Peer」機能を使用して遅いピアの問題を解決する

目次

[概要](#)

[背景説明](#)

[アップデート グループ](#)

[問題](#)

[解決策](#)

[検出](#)

[低速ピアの識別](#)

[移動](#)

[低速ピア機能なしの移動](#)

[スタティック低速ピアの移動](#)

[ダイナミック低速ピアの移動](#)

[Recovery](#)

[低速ピア状態のクリア](#)

概要

このドキュメントでは、Border Gateway Protocol (BGP) の低速ピア機能を使用して、低速ピアの問題を解決する方法を説明します。この機能は BGP アップデート グループ内の低速ピアを識別し、永続的または一時的に低速ピアをアップデート グループの外に移動できます。

背景説明

ここでは、低速ピア機能とアップデート グループの使用法の概要を説明します。

アップデート グループ

低速ピア機能はアップデート グループで使用されます。アップデート グループは、同じアウトバウンド ポリシーの BGP ピアをグループ化するために使用されるダイナミックな方式です。アップデート グループの利点は、メッセージを一度フォーマットするためにグループ ポリシーが使用され、その後グループの他のメンバーに複製され、送信されることです。この方式は、各ピアに対して個別に BGP アップデートをフォーマットする必要がある場合に比べて効率的です。

この方式を実装されると、アウトバウンド ポリシーが変更された場合、ピア グループはアップデート グループごとに変更されます。アップデート グループはアドレス ファミリ (AF) ごとに形成されます。

AF IPv4 ユニキャストに対しては異なるアップデートグループに含まれ、AF VPNv4 に対しては同じアップデートグループに含まれる、2つのBGPピアの例を次に示します。

```
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
```

アップデートグループに含まれるBGPピアの数が増えるほど、アップデートグループはより効率的になります。通常、内部BGP (iBGP) ピアは同じ発信ポリシーを持ちます。iBGPの場合、ルートリフレクタ (RR) は多くのiBGPピアを持つことができます。そのため、大きなアップデートグループがあります。プロバイダエッジ (PE) ルータは、1つの仮想/ルーティング転送 (VRF) のカスタムエッジ (CE) ルータに対して、多くの外部BGP (eBGP) ピアが可能です。PE ルータは、VRF インターフェイス上のCE ルータとピアリングに対しても大きなアップデートグループを持つことができます。

問題

低速ピアは、アップデートグループの中で、ルータが長時間 (分単位) に渡ってBGPアップデートメッセージを生成する速度に追従できないピアです。低速ピアの原因は、永続的なネットワークの問題である可能性があります。ネットワークの原因にはパケット損失や負荷のかかったリンク、BGPセッションのスループットの問題が考えられます。また、BGPピアがCPUに関しては高負荷になっているため、必要な速度のTCP接続を提供できない場合もあります。

低速ピアは完全なアップデートグループのBGPコンバージェンスに影響します。1つのBGPピアが低速である場合、全体のアップデートグループが低速になります。その結果、他のアップデートグループメンバーのコンバージェンスも低速になります。こうした理由から、問題を解決する必要があります。

低速ピアを識別し、アップデートグループから移動できます。このタスクを完了するために、該当するBGPピアのアウトバウンドポリシーを変更できます。ただし、これは手動のタスクです。まず、低速なピアを特定し、アップデートグループからピアを移動する必要があります。低速ピア機能はこれを自動で行うことができるので、ユーザの介入は不要です。

解決策

低速ピア機能には次の3つの部分があります。

- 低速ピアの検出
- 低速ピアの低速アップデートグループへの移動
- 低速ピアのリカバリ (リカバリしたピアを元のアップデートグループへ移動)

これらの処理の詳細は、以降の項で説明します。

検出

低速ピア機能はアップデート グループ内の低速ピアを検出します。各アップデート グループにはキャッシュ キューがあり、フォーマットされた BGP アップデートが送信前に一時的に保存されます。

そうしたアップデート グループ キャッシュの例を次に示します。

```
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	

キャッシュのサイズは自動的に計算され、次に依存します。

- アップデート グループ内のピアの数
- インストールされているシステム メモリ
- アップデート グループ内のピアのタイプ
- AF のタイプ

送信を待つフォーマットされた BGP アップデートの数は、1つのピア（低速ピア）が他のメンバーほど迅速に BGP メッセージを認識しない場合、1つのアップデート グループ内で増加する可能性があります。キャッシュの制限に達すると、グループに新しいメッセージをキューするためのクォータの割り当てがなくなります。キャッシュが削減されるまで（低速ピアによって一部のメッセージが認識されるまで）、新しいメッセージをフォーマットできません。これにより BGP ピアが妨害され、グループの高速なメンバーに新しいメッセージ（更新または許可）を送信できません。したがって、アップデート グループ内のすべてのピアのコンバージェンスが遅くなります。

低速ピア機能は、低速ピアを特定するために、BGP アップデート タイムスタンプとピア TCP パラメータを参照します。

低速ピア検出はデフォルトで無効になっています。低速ピア検出を有効にするには、次のいずれかの方法を使用します。

- BGP プロセスの機能を有効にします（AF/VRF から設定可能）。

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

[no] bgp slow-peer detection 注: しきい値の範囲は 120 ~ 3600 秒で、デフォルト値は 300 秒です。

- ピア単位で機能を有効にします。

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

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

- ピア ポリシー テンプレートを使用して機能を有効にします。

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

[no] slow-peer detection

低速ピアが検出されると、次のような syslog メッセージが出力されます。

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

次の show コマンドを入力して低速ピアを表示することもできます。

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

slow キーワードを使用した show コマンド出力の例を次に示します。

```
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
```

出力に示されるように、ピア 10.1.6.7 は AF IPv4 ユニキャストの低速ピアです。他の AF では低速ピアが表示されていません。

現在検出タイマーが動作しているかどうか、およびその値を確認するには、次のコマンドを入力します。

```
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
```

出力例が示すように、検出タイマーは開始しています。検出タイマーはアップデートグループ キャッシュが満杯になると開始されます。

この例では、低速ピアが検出されていますが、低速ピア検出タイマーが時間切れになった後で初めて、そのピアがアップデートグループから移動するのを確認できます。

```
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
```

低速ピアの識別

低速ピア検出機能が有効でない場合、手動で低速ピアを識別する必要があります。まず、アップデートグループ内のピアのテーブルバージョンおよび出力キューを確認します。

```
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
```

この例では、ピアのテーブルバージョン (TblVer) がメインの BGP テーブルバージョンに追従しているか、それとも常に遅れているかを確認します。次に、非常に高い出力キュー値のピアがないか確認します。これらは低速ピアである可能性が高くなります。

疑わしい低速 BGP ピアを確認する際、次の点を考慮します (BGP セッションの両側で)。

- 最後の書き込みが実行されたのはいつか。
- スロットルのキープアライブであるか。
- 出力キューは高いか。
- SRTT/RTTO は高いか。
- 再送信の数は増加しているか。
- キューされた再送信パケットはあるか。
- TCP 送信ウィンドウは非常に低いか、またはゼロか。

次に例を示します。

```
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 NLRI's 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, KRRT: 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
```

移動

ここでは、低速ピア機能に関する移動処理について、さまざまなシナリオで説明します。

低速ピア機能なしの移動

低速ピアは、低速ピア機能を使用しないで、手動で新しいアップデートグループに移動できます。

低速ピア機能が使用可能になる前は、低速ピアを特定し、アップデートグループから手動で移動する必要がありました。これは、該当する BGP ピアのアウトバウンドポリシーへの変更により実現されます。このアウトバウンドポリシーは使用される他のポリシーとも異なる必要があります。なぜなら、低速ピアが、現在存在する別のアップデートグループに移動する（問題をそのアップデートグループに移動する）ことがないようにしなければならないからです。適用できる最適な変更は、実際のポリシーに影響を与えないものです。たとえば、ピアの最小ルートアドバタイズメントインターバル（MRAI）を変更できます（特定の AF の下で）。

低速ピア機能が使用できない場合に低速ピアを手動で移動する例を次に示します。

```
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
```

スタティック低速ピアの移動

ピアをアップデートグループから新しいアップデートグループに移動するために、スタティック低速ピアとして設定できます。複数の低速ピアがある場合、同じアウトバウンドポリシーのスタティック低速ピアは、同じ低速アップデートグループに配置されます。

低速ピアを静的に移動するには、次のコマンドを使用して設定できます。

- ネイバー単位またはピアグループ単位のスタティックピアの移動を有効にします。
[no] neighbor {<nbr-addr>/<peer-grp-name>} slow-peer split-update-group static
- ピアポリシーテンプレートによってスタティックピアの移動を有効にします。
[no] slow-peer split-update-group static

ダイナミック低速ピアの移動

低速ピアの移動はデフォルトで無効になっています。低速ピアの移動を有効にするには、次のいずれかの方法を使用して設定できます。

- BGP プロセスの低速ピアの移動を有効にします。
bgp slow-peer split-update-group dynamic [permanent]

[no] bgp slow-peer split-update-group dynamic注: これは [address-family/topology/VRF] ビューから設定できます。
- ピア単位で低速ピアの移動を有効にします。
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
- ピアポリシーテンプレートによって低速ピアの移動を有効にします。
slow-peer split-update-group dynamic [permanent]

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

注: permanent キーワードは低速ピアが自動的にリカバリしないことを示します。この場合、[clear コマンド](#)のいずれかを使用して、リカバリした低速ピアを元のアップデートグループに移動できます。

スタティック低速ピアとダイナミック低速ピアは同じ低速ピアアップデートグループに属します。この例では、1つのアップデートグループ内に1つの低速ピアが確認できます。

```
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 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
```

Recovery

低速ピアは、もはや低速ピアではない（追従している）ことを確認できたら、（アウトバウンドポリシーに一致する）元のアップデートグループの下で、再グループ化できます。リカバリタイマーは、低速ピアアップデートグループが収束すると開始されます。リカバリタイマーが失効すると、低速ピアは通常のアップデートグループへ移動します。

注: 検出/リカバリタイマーに関連する動作を表示するには、`debug ip bgp updates events` コマンドを入力します。

低速ピアが元のアップデートグループに戻ると（これはリカバリを意味します）、次のような syslog メッセージが出力されます。

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

現在リカバリタイマーが動作しているかどうか、およびその値を確認するには、次のコマンドを入力します。

```
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
```

この例では、リカバリタイマーは値が 16 秒で、低速ピアが 16 秒後に元のアップデートグループに戻る可能性があることを示します。

この例では、ピアが低速ピアの状態からリカバリしていることを確認できます。

```
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
```

低速ピア状態のクリア

次のコマンドを使用して、低速ピアの状態を手動でクリアできます。

- `clear ip bgp * slow`
- `clear ip bgp AF {unicast/multicast} <AS 番号> slow`
- `clear ip bgp AF {unicast/multicast} peer-group <グループ名> slow`

- clear ip bgp <ネイバー アドレス> slow
- clear bgp AF {unicast/multicast} * slow
- clear bgp AF {unicast/multicast} <AS 番号> slow
- clear bgp AF {unicast/multicast} peer-group <グループ名> slow
- clear bgp AF {unicast/multicast} <ネイバー アドレス> slow

注: これらのコマンドを使用すると、AF が実際のアドレス ファミリで置き換えられます。

これらのコマンドを使用することにより、ピアは元のアップデート グループへ移動します。

show ip bgp internal コマンドを入力すると、低速ピアの検出と移動の設定が表示されます。

```
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
```

注: 要約すると、BGP 低速ピアは、BGP アップデート グループ内の低速ピアを検出し、低速ピアをアップデート グループの外に移動することで、高速な BGP コンバージェンスを実現する機能です。