Border Gateway Protocol(BGP)の基本的な問題のトラブルシューティング

内容

概要

このドキュメントでは、ボーダーゲートウェイプロトコル(BGP)で最も一般的な問題をトラブルシューティングする方法について説明し、基本的なソリューションとガイドラインを提供します。

前提条件

要件

このドキュメントに関しては個別の前提条件はありません。BGPプロトコルに関する基本的な知識が役立ちます。詳細については、『BGPコンフィギュレーションガイド』を参照してください

。

使用するコンポーネント

このドキュメントは、特定のソフトウェアやハードウェアのバージョンに限定されるものではありませんが、Cisco IOS®およびCisco IOS® XEに適用されるコマンドです。

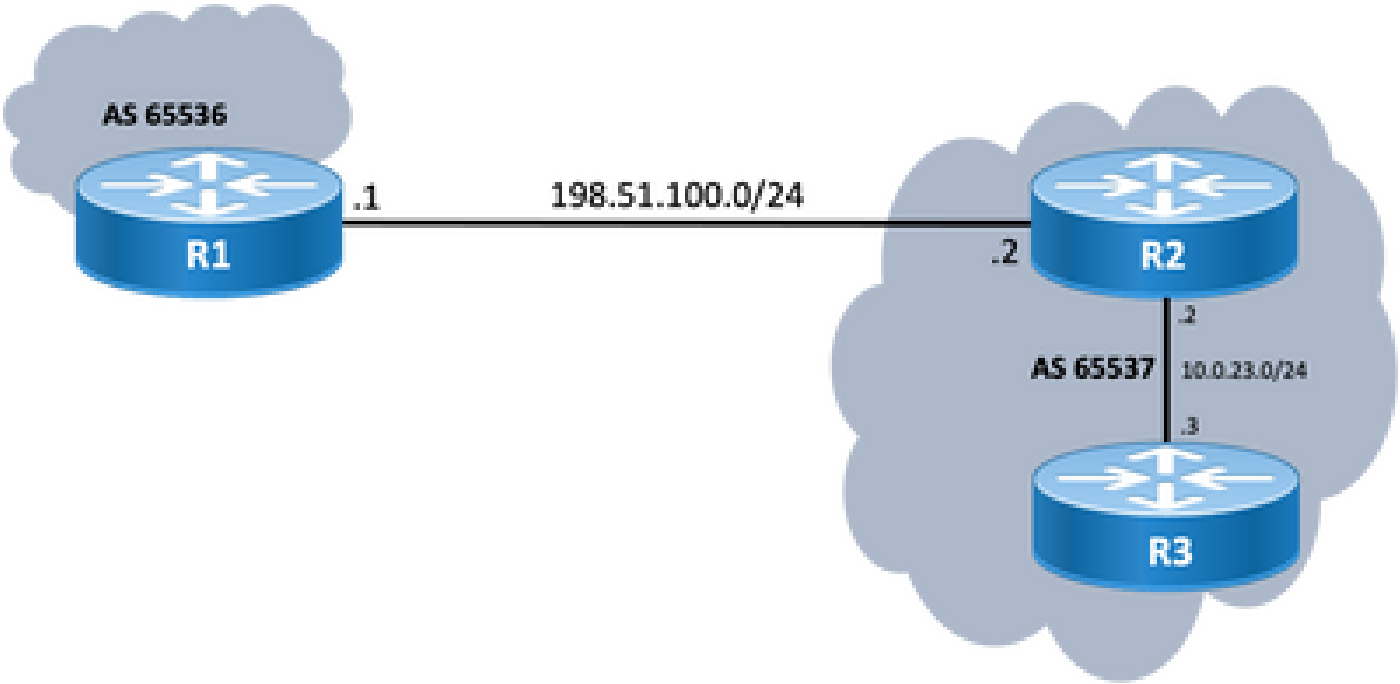
このドキュメントの情報は、特定のラボ環境にあるデバイスに基づいて作成されました。このドキュメントで使用するすべてのデバイスは、クリアな（デフォルト）設定で作業を開始しています。本稼働中のネットワークでは、各コマンドによって起こる可能性がある影響を十分確認してください。

背景説明

このドキュメントでは、Border Gateway Protocol（BGP；ボーダーゲートウェイプロトコル）で最も一般的な問題をトラブルシューティングするための基本的なガイド、修正措置、問題の根本原因を検出するための便利なコマンドやデバッグ、潜在的な問題を回避するためのベストプラクティスについて説明します。考えられるすべての変数とシナリオは考慮できず、Cisco TACではより詳細な分析が必要になる可能性があることに注意してください。

トポロジ

このトポロジ図は、このドキュメントで提供する出力の参照用として使用してください。



シナリオと問題

隣接関係ダウン

BGPセッションがダウンしていて起動しない場合は、 show ip bgp all summary command. ここでは、セッションの現在のステータスを確認できます。

- セッションがアップ状態でない場合は、IDLEとACTIVEの間で異なる可能性があります(有限状態マシン(FSM)プロセスによって異なります)。
- セッションがアップしている場合は、受信したプレフィックスの数が表示されます。

<#root>

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
```

接続できない

確認する必要がある最初の要件は、ポート179でTCPセッションを確立できるように両方のピア間を接続することです。直接接続されているかどうかです。この問題には単純なpingが役立ちます。ループバックインターフェイス間でピアリングが確立されている場合は、ループバックからループバックへのpingを実行する必要があります。送信元インターフェイスとして特定のループバックを指定せずにpingテストを実行すると、ルータのループバックIPアドレスではなく、発信物理インターフェイスのIPアドレスがパケットの送信元IPアドレスとして使用されます。

pingが成功しない場合は、次の原因を考慮してください。

- 接続されたルートピアがないか、またはルートがありません： show ip route peer_IP_address を使用できます。
- レイヤ1の問題：物理インターフェイス、SFP（コネクタ）、ケーブルまたは外部の問題（トランスポートおよびプロバイダーが存在する場合）、を考慮する必要があります。
- 接続をブロックする可能性のあるファイアウォールまたはアクセスリストをチェックします。

pingが成功した場合は、次の点を考慮してください。

設定の問題

- 誤ったIPアドレスまたは設定されたAS：誤ったIP用 」というメッセージは表示されませんが、適切な設定が行われていることを確認してください。ASが正しくない場合は、 show logging コマンドを使用して、アップグレードを実行します。

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

両端のBGP設定をチェックして、AS番号またはピアのIPアドレスを修正します。

- ルータ ID の重複:

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

次のコマンドを使用して、両端のBGP識別子をチェックします。 show ip bgp all summary 重複した問題を修正します。これは、グローバルコマンドを使用して手動で実行できます bgp router-id X.X.X.X bgp router configurationの下で設定します。ベストプラクティスとして、ルータIDを手動で一意の番号に設定してください。

- BGPソースとTTL:

ほとんどのiBGPセッションは、IGPを介して到達可能なループバックインターフェイスを介して設定されます。このループバックインターフェイスを送信元として明示的に定義する必要があります。次のコマンドを使用して行います。 neighbor ip-address update-source interface-id を参照。

eBGPピアの場合、直接接続されたインターフェイスは通常ピアリングに使用され、Cisco IOS/Cisco IOS XEがこの目的を満たすかどうかをチェックします セッションの確立を試みることもありません。直接接続されたルータでeBGPをループバックからループバックに試行する場合、このチェックは両端の特定のネイバーに対して無効にすることができます。 neighbor ip-address disable-connected-check を参照。

ただし、eBGPピア間に複数のホップが存在する場合は、適切なホップカウントが必要です。**neighbor ip-address ebgp-multihop [hop-count]**正しいホップカウントで設定されているため、セッションを確立できます。

hop-countを指定しない場合、iBGPセッションのデフォルトのTTL値は255ですが、eBGPセッションのデフォルトのTTL値は1です。

TCPセッションの問題

ポート179をテストする便利なアクションは、ピア間の手動telnetです。

```
<#root>

R1#

telnet 198.51.100.2 179

Trying 198.51.100.2, 179 ... Open

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

open/connection closedまたはconnection refused by remote hostのどちらも、パケットがリモートエンドに到達したことを示します。その後、遠端のコントロールプレーンに問題がないことを確認します。到達不能な宛先がある場合は、TCPポート179、BGPパケット、またはパス上のパケット損失をブロックする可能性のあるファイアウォールまたはアクセスリストを確認します。

認証の問題が発生した場合、次のメッセージが表示されます。

```
%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
```

認証方式、パスワード、および関連する設定を確認し、さらにトラブルシューティングするには、『BGPピア間のMD5認証の設定例』を参照してください。

TCPセッションが確立されない場合は、次のコマンドを使用して分離できます。

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

隣接関係のバウンス

セッションがアップ状態とダウン状態の場合は、 show log いくつかのシナリオが表示されます

インターフェイスフラップ

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

メッセージが示すように、この障害の原因はインターフェイスのダウン状況です。ポート/SFP、ケーブル、または切断に物理的な問題がないかどうかを確認してください。

ホールドタイマーの期限切れ

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

これは非常に一般的な状況です。ルータがキープアライブメッセージまたは更新メッセージを受信または処理しなかった後、ホールドタイマーが時間切れになったことを意味します。デバイスが通知メッセージを送信し、セッションを閉じます。この問題の最も一般的な原因を次に示します。

- インターフェイスの問題：両方のピアの接続されたインターフェイスで、入力エラー、入力キューのドロップ、または物理的な問題がないか探します。 show interface この目的で使用できます。
- 伝送中のパケット損失：Helloパケットが伝送中にドロップされる場合があります。これは、インターフェイスレベルでのパケットキャプチャを確実にするための最善の方法です。
    - Cisco IOSおよびCisco IOS XEデバイスで組み込みパケットキャプチャ(EPC)を使用できます。
    - パケットがインターフェイスレベルで見られる場合は、パケットがコントロールプレーンのEPCに到達していることを確認する必要があります コントロールプレーン、または debug bgp [vrf name] ipv4 unicast keepalives 便利です。
- CPUの高使用：CPUの高使用状態は、コントロールプレーンでドロップを引き起こす可能性があります。 show processes cpu [sorted|history] 問題を特定するのに役立ちます。プラットフォームに基づいて、トラブルシューティングの次のステップを見つけることができます。詳細については、『CPUリファレンス』ドキュメントを参照してください
- CoPPポリシーの問題：トラブルシューティングの方法はプラットフォームごとに異なるため、このドキュメントでは説明しません。

- MTUの不一致：パス内にMTUの不一致があり、発信元から宛先へのパスでICMPメッセージがブロックされている場合、PMTUDは機能せず、セッションフラップが発生する可能性があります。アップデートは、ネゴシエートされたMSS値とDFビットが設定された状態で送信されます。パス内のデバイスまたは宛先が、より高いMTUのパケットを受け入れることができない場合、ICMPエラーメッセージがBGPスピーカに返されます。宛先ルータは、BGPキープアライブまたはBGPアップデートパケットによってホールドダウンタイマーが更新されるのを待ちます。
    - ネゴシエートされたMSSを確認できます。 show ip bgp neighbors ip_addressを参照。

dfを設定した特定のネイバーに対するpingテストでは、そのようなMTUがパス上で有効かどうか

が示されます。

<#root>

ping 198.51.100.2 size

max_seg_size

df

MTUの問題が見つかった場合は、MTU値がネットワーク全体で一貫していることを確認するために、設定を正確に確認する必要があります。

注：MTUの詳細は、『MTUによるBGPネイバーフラップのトラブルシューティング』を参照してください。

AFI/SAFIの問題

```
%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 000
```

Address-Family Identifier(AFI)は、マルチプロトコルBGP(MP-BGP)によって追加される機能拡張です。このプロトコルは、IPv4、IPv6などの特定のネットワークプロトコルと、ユニキャストやマルチキャストなどのSubsequent Address-Family Identifier(SAFI)による細かい設定に関連しています。MBGPは、BGPパスアトリビュート(PA)MP_REACH_NLRIおよびMP_UNREACH_NLRIによってこの分離を実現します。これらの属性はBGPアップデートメッセージ内で伝送され、さまざまなアドレスファミリのネットワーク到達可能性情報の伝送に使用されます。

メッセージには、IANAによって登録された次のAFI/SAFIの番号が表示されます。

- IANAアドレスファミリ番号
- 後続のアドレスファミリID(SAFI)パラメータ

- 望ましくないアドレスファミリを修正するには、両側で意図されているアドレスファミリのBGP設定を確認します。
- 利用 neighbor ip-address dont-capability-negotiate 両側で行われます詳細は、『サポートされていない機能が原因でBGPピアが誤動作する』を参照してください。

パスのインストールと選択

BGPの動作のしくみについての詳しい説明、およびベストパスの選択については、『BGPでベストパスを選択するアルゴリズム』を参照してください。

ネクストホップ

ルートをルーティングテーブルにインストールするには、ネクストホップが到達可能である必要があります。到達可能でない場合は、プレフィックスがLoc-RIB BGPテーブルにある場合でも、RIBに入りません。 ループ回避ルールとして、Cisco IOS/Cisco IOS XEでは、iBGPはネクストホップのアトリビュートを変更せず、AS_PATHだけを残しますが、eBGPはネクストホップを書き換えてAS_PATHを付加します。

ネクストホップは次のコマンドで確認できます。 show ip bgp [prefix]. ネクストホップとアクセスできないワードが表示されます。この例では、R1がeBGP経由でR2にアナウンスし、R3がR2からのiBGP接続を介して学習したプレフィックスです。

```
<#root>

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
```

出力では、ネクストホップはR3に認識されていないR1の発信インターフェイスです。この状況を解決するには、IGPまたはスタティックルートを使用してネクストホップをアドバタイズするか、 neighbor ip-address next-hop-self コマンドをiBGPピアで発行して、直接接続されているネクストホップIPを変更します。図の例では、この設定はR2、つまりR3へのネイバー（ネイバー10.0.23.3 next-hop-self）上にある必要があります。

その結果、ネクストホップが変更されます(clear ip bgp 10.0.23.2 soft)を直接接続されたインターフェイス（到達可能）に追加し、プレフィクスをインストールします。

```
<#root>

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
```

RIB障害

これは、ルートをグローバルRIBにインストールできない場合に発生し、その結果RIB障害が発生します。一般的な理由は、アドミニストレーティブディスタンスが短い別のルーティングプロトコルの同じプレフィックスがすでにRIBにある場合ですが、RIB障害の正確な理由はshow ip bgp rib-failureコマンドで表示されます。詳細な説明については、次のリンクを参照してください。

注:「BGP RIB-failureについて」および「bgp suppress-inactiveコマンドについて」で説明されているように、このような問題を特定して修正できます。

競合状態

最も一般的な問題は、相互再配布シナリオでeBGPよりもIGPが優先される場合です。IGPルートがBGPに再配布されると、BGPによってローカルに生成されたものと見なされ、デフォルトで32768の重み付けが適用されます。BGPピアから受信したすべてのプレフィックスには、デフォルトで0のローカル重みが割り当てられます。したがって、同じプレフィクスを比較する必要がある場合は、BGPベストパス選択プロセスに基づいて、より大きい重みを持つプレフィクスがルーティングテーブルにインストールされます。これが、IGPルートがRIBにインストールされる理由です。

この問題の解決策は、router bgp設定でBGPピアから受信したすべてのルートに対して、より大きい重みを設定することです。

<#root>

neighbor

ip-address

weight 40000

注：詳細については、『ネットワークフェールオーバーシナリオにおけるBGPウェイトパスアトリビュートの重要性について』を参照してください。

その他の問題

BGP低速ピア

送信者がアップデートメッセージを生成するレートについていけないピアです。ピアでこの問題が発生する理由はさまざまです。たとえば、ピアの1つでCPUの使用率が高くなる、リンク上でのトラフィックの超過や損失、帯域幅リソースなどが考えられます。

> 注：低速ピアの問題を特定して修正するには、『BGPの「低速ピア」機能を使用した低速ピアの問題の解決』を参照してください。

メモリの問題

BGPは、Cisco IOSプロセスに割り当てられたメモリを使用して、ネットワークプレフィックス、ベストパス、ポリシー、およびすべての関連する設定を維持し、正常に動作するようにします。プロセス全体がコマンドで表示されます show processes memory sorted：を入力します。

<#root>

```
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*
```

プロセッサプールは使用されるメモリです。この例では約2.1 GBです。次に、Holding列を見て、その大部分を保持しているサブプロセスを特定します。 次に、所有するBGPセッション、受信さ

れたルートの数、および使用されている設定を確認する必要があります。

BGPによるメモリ保持を減らす一般的な手順：

- BGPフィルタリング：完全なBGPテーブルを受信する必要がない場合は、ポリシーを使用してルートをフィルタリングし、必要なプレフィックスだけをインストールします。
- ソフト再構成：BGP設定でneighbor ip_address soft-reconfiguration inboundを探します。このコマンドを使用すると、どの着信ポリシー(Adj-RIB-in)よりも前に受信されたすべてのプレフィックスを確認できます。ただし、この情報を保存するには、このテーブルが現在のBGPローカルRIBテーブルの約半分を必要とします。強制的に必要な場合や現在のプレフィックスが少ない場合を除き、この設定は避けることができます。

注:BGPを最適化する方法の詳細については、『BGPルータを設定して最適なパフォーマンスとメモリ消費の削減を実現する』を参照してください。

CPU の使用率が高い

ルータは、BGPが動作するためのさまざまなプロセスを使用します。BGPプロセスが高いCPU使用率の原因であることを確認するには、 show process cpu sorted コマンドを使用して、アップグレードを実行します。

<#root>

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
```

BGPが原因で発生する高いCPU使用率を克服するための一般的なプロセス、原因、および一般的な手順を次に示します。

- BGPルータ：高速コンバージェンスを保護するために1秒に1回実行されます。これは最も重要なスレッドの1つです。BGP更新メッセージを読み取り、プレフィックス/ネットワークと属性を検証し、AFI/SAFIネットワーク/プレフィックステーブルと属性テーブルごとに更新し、他の多くのタスクの中でベストパス計算を実行します。
  巨大なルートの変更は、この状況につながる非常に一般的なシナリオです。
- BGPスキャナ：デフォルトで60秒ごとに実行される低優先度プロセス。このプロセスは、BGPテーブル全体をチェックしてネクストホップの到達可能性を確認し、パスに変更があった場合はBGPテーブルを適宜更新します。再配布のためにルーティング情報ベース(RIB)を通過する
  プレフィクスとルートのインストールとTCAMの使用が増え、必要なリソースが増え、通常はデバイスの過負荷によりそのような状況が発生するため、プラットフォームの規模を確認します。

注：これら2つのプロセスのトラブルシューティング方法の詳細については、「BGPスキャナまたはルータプロセスが原因で発生するCPUの高使用のトラブルシューティング」を参照してください。

- BGP I/O:BGP制御パケットが受信されると実行され、BGPパケットのキューイングと処理を管理します。BGPキューで長期間にわたって過剰なパケットが受信されたり、TCPに問題がある場合、ルータではBGP I/Oプロセスが原因でCPU使用率が高くなる症状が発生します。(通常、この状況ではBGPルータの使用率も高くなります。メッセージカウントを調べてピアを特定し、パケットをキャプチャしてこれらのメッセージの送信元を特定します)。
- BGP Open：セッション確立時に使用されるプロセス。セッションがOpen状態でスタックしない限り、一般的なCPU高使用率の問題ではありません。
- BGPイベント：ネクストホップ処理を行います。受信したプレフィックスでネクストホップフラップを探します。

関連情報

- テクニカル サポートとドキュメント - Cisco Systems
- BGP 設定ガイド
- BGP ピア間の MD5 認証の設定例
- Embedded Packet Capture
- MTUによるBGPネイバーフラップのトラブルシューティング

- IANAアドレスファミリ番号

- 後続のアドレスファミリID(SAFI)パラメータ

- サポートされていない機能が原因でBGPピアが誤動作する
- BGP でベスト パスを選択するアルゴリズム
- BGP RIB-failureとbgp suppress-inactiveコマンドについて
- ネットワーク フェールオーバー シナリオにおける BGP 重みパス属性の重要性の理解

- BGP の「Slow Peer」機能を使用して遅いピアの問題を解決する
- 最適なパフォーマンスとメモリ消費の削減のためのBGPルータの設定
- BGPスキャナまたはルータプロセスが原因で発生するCPU高使用率のトラブルシューティング

翻訳について
シスコは世界中のユーザにそれぞれの言語でサポート コンテンツを提供するために、機械と人による翻訳を組み合わせて、本ドキュメントを翻訳しています。ただし、最高度の機械翻訳であっても、専門家による翻訳のような正確性は確保されません。シスコは、これら翻訳の正確性について法的責任を負いません。原典である英語版（リンクからアクセス可能）もあわせて参照することを推奨します。