Dépannage des problèmes de base du protocole BGP

Table des matières

Introduction

Ce document décrit comment dépanner les problèmes les plus courants avec le Border Gateway Protocol (BGP) et fournit des solutions et des directives de base.

Conditions préalables

Exigences

Aucune condition préalable spécifique n'est requise pour ce document. La connaissance de base du protocole BGP est utile, vous pouvez vous référer au Guide de configuration BGP pour plus d'informations.

Composants utilisés

Ce document n'est pas limité à des versions logicielles et matérielles spécifiques, mais les commandes sont applicables à Cisco IOS® et Cisco IOS® XE.

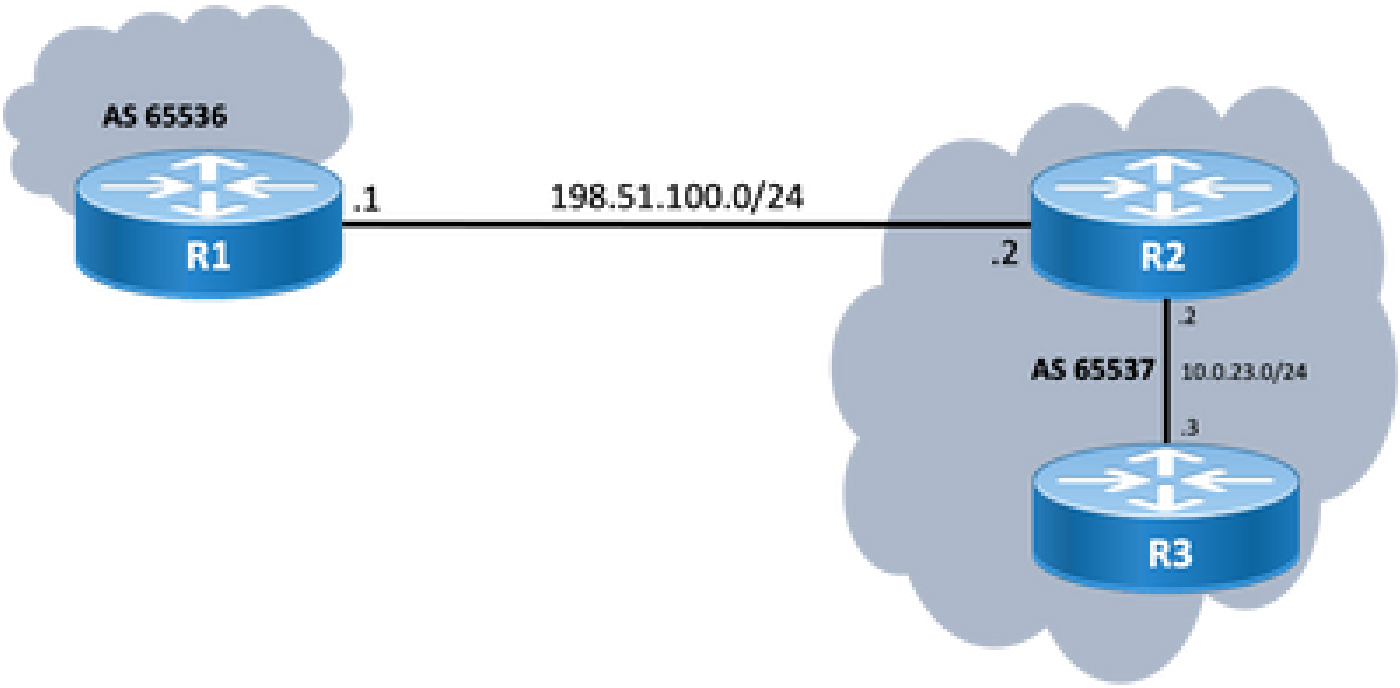
The information in this document was created from the devices in a specific lab environment. All of the devices used in this document started with a cleared (default) configuration. Si votre réseau est en ligne, assurez-vous de bien comprendre l’incidence possible des commandes.

Informations générales

Ce document décrit un guide de base pour dépanner les problèmes les plus courants dans le protocole BGP (Border Gateway Protocol), fournit des actions correctives, des commandes/débogages utiles pour détecter la cause racine des problèmes et les meilleures pratiques pour éviter les problèmes potentiels. Gardez à l'esprit que toutes les variables et tous les scénarios possibles ne peuvent pas être pris en compte et qu'une analyse plus approfondie pourrait être requise par le TAC Cisco.

Topologie

Utilisez ce schéma de topologie comme référence pour les résultats fournis dans ce document.



Scénarios et problèmes

Contiguïté descendante

Si une session BGP est arrêtée et ne s'ouvre pas, émettez la commande show ip bgp all summary command. Vous trouverez ici l'état actuel de la session :

- Si l'état de la session n'est pas up, il peut varier entre IDLE et ACTIVE (dépend du processus de la machine à états finis).
- Si la session est active, vous voyez le nombre de préfixes reçus.

<#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
```

Aucune connectivité

La première condition à remplir est la connectivité entre les deux homologues afin que la session TCP sur le port 179 puisse être établie. Soit ils sont directement connectés, soit ils ne le sont pas. Une simple requête ping est utile dans ce cas. Si l'appairage est établi entre les interfaces de bouclage, une requête ping de bouclage à bouclage doit être effectuée. Si un test ping est effectué sans bouclage spécifique comme interface source, l'adresse IP de l'interface physique sortante est utilisée comme adresse IP source du paquet au lieu de l'adresse IP de bouclage du routeur.

Si la requête ping échoue, tenez compte des causes suivantes :

- Aucun homologue de route connecté ou aucune route : show ip route peer_IP_address peut être utilisé.
- Problème de couche 1 : l'interface physique, le SFP (connecteur), le câble ou un problème externe (transport et fournisseur, le cas échéant) doit être pris en compte.
- Vérifiez tout pare-feu ou toute liste d'accès pouvant bloquer la connexion.

Si la requête ping aboutit, tenez compte des points suivants :

Problèmes liés à la configuration

- Adresse IP incorrecte ou AS configuré : adresse IP incorrecte , aucun message de ce type

n'est affiché, mais assurez-vous que la configuration est correcte. Pour les AS incorrects, vous devez voir un message comme avec le `show logging` erasecat4000_flash:.

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

Vérifiez la configuration BGP aux deux extrémités pour corriger les numéros de système autonome ou l'adresse IP homologue.

- Identificateur de routeur en double:

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

Vérifiez l'identificateur BGP aux deux extrémités via `show ip bgp all summary` et corrigez le problème en double. Ceci peut être réalisé manuellement avec la commande globale `bgp router-id X.X.X.X` sous bgp router configuration. Il est recommandé de définir manuellement l'ID de routeur sur un numéro unique.

- Source BGP et durée de vie :

La plupart des sessions iBGP sont configurées sur les interfaces de bouclage accessibles via un IGP. Cette interface de bouclage doit être explicitement définie comme source. Pour ce faire, utilisez la commande `neighbor ip-address update-source interface-id` .

Pour l'homologue eBGP, les interfaces connectées directement sont généralement utilisées pour l'appairage, et il y a une vérification pour que Cisco IOS/Cisco IOS XE remplisse cet objectif, ou il le fait même pas essayer d'établir une session. Si eBGP est essayé de bouclage en bouclage sur des routeurs connectés directement, cette vérification peut être désactivée pour un voisin spécifique aux deux extrémités via `neighbor ip-address disable-connected-check` .

Cependant, s'il y a plusieurs sauts entre les homologues eBGP, un nombre de sauts correct est requis, assurez-vous que **neighbor ip-address ebgp-multihop [hop-count]**est configuré avec le nombre de sauts correct afin que la session puisse être établie.

Si le nombre de sauts n'est pas spécifié, la valeur TTL par défaut pour les sessions iBGP est 255, tandis que la valeur TTL par défaut pour les sessions eBGP est 1.

Problèmes de session TCP

Une action utile pour tester le port 179 est une connexion Telnet manuelle d’un homologue à l’autre :

<#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]
```

L'ouverture/la fermeture de la connexion ou le refus de la connexion par l'hôte distant indique que les paquets atteignent l'extrémité distante, puis assurez-vous qu'il n'y a aucun problème avec le plan de contrôle à l'extrémité distante. Sinon, si une destination est inaccessible, vérifiez tout pare-feu ou toute liste d'accès qui peut bloquer le port TCP 179, ou les paquets BGP, ou toute perte de paquets sur le chemin.

En cas de problème d'authentification, les messages suivants s'affichent :

```
%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
```

Vérifiez les méthodes d'authentification, le mot de passe et la configuration associée, et pour plus de détails sur le dépannage, référez-vous à Exemple de configuration d'authentification MD5 entre homologues BGP.

Si la session TCP ne s'ouvre pas, vous pouvez utiliser les commandes suivantes pour l'isolation :

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

Rebonds de contiguïté

Si la session est active et inactive, recherchez `show log` et vous pouvez voir quelques scénarios.

Volet D'Interface

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

Comme l'indique le message, la raison de cette panne est la situation d'interface inactive, recherchez tout problème physique sur le port/SFP, le câble ou les déconnexions.

Expiration du minuteur de conservation

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

C'est une situation très courante ; cela signifie que le routeur n'a pas reçu ou traité de message de test d'activité ou de message de mise à jour avant l'expiration du minuteur de mise en attente. Le périphérique envoie un message de notification et ferme la session. Les raisons les plus communes de ce problème sont énumérées ici :

- Problèmes d'interface : recherchez les erreurs d'entrée, les pertes de file d'attente d'entrée ou les problèmes physiques sur les interfaces connectées des deux homologues ; `show interface` peut être utilisé à cette fin.
- Perte de paquets en transit : parfois, les paquets Hello peuvent être abandonnés en transit, la meilleure façon de s'assurer que c'est une capture de paquets au niveau de l'interface.
    - Vous pouvez utiliser la capture de paquets intégrée sur les périphériques Cisco IOS et Cisco IOS XE.
    - Si des paquets sont vus au niveau de l'interface, vous devez vous assurer qu'ils atteignent le plan de contrôle, EPC sur le plan de contrôle, ou `debug bgp [vrf name] ipv4 unicast keepalives` est utile.
- CPU élevé : une condition CPU élevée peut provoquer des chutes sur le plan de contrôle, `show processes cpu [sorted|history]` est utile pour identifier le problème. En fonction de la plate-forme, vous pouvez trouver l'étape suivante de dépannage avec le document Référence CPU
- Problèmes de stratégie CoPP : la méthodologie de dépannage varie pour chaque plate-forme et n'est pas couverte par ce document.

- MTU mismatch : s'il y a des différences de MTU dans le chemin, et si les messages ICMP sont bloqués dans le chemin de la source à la destination, PMTUD ne fonctionne pas et peut entraîner un battement de session. Les mises à jour sont envoyées avec la valeur MSS négociée et un bit DF défini. Si un périphérique sur le chemin ou même la destination n'est pas en mesure d'accepter les paquets avec un MTU plus élevé, il renvoie un message d'erreur ICMP au haut-parleur BGP. Le routeur de destination attend que le keepalive BGP ou le paquet de mise à jour BGP mette à jour son minuteur de mise hors service.
    - Vous pouvez vérifier le MSS négocié avec `show ip bgp neighbors ip_address`.

Un test Ping sur un voisin spécifique avec df set peut vous montrer si ce MTU est valide le long du

chemin :

<#root>

ping 198.51.100.2 size

max_seg_size

df

Si des problèmes de MTU sont détectés, un examen précis de la configuration doit être effectué pour s'assurer que les valeurs de MTU sont cohérentes sur l'ensemble du réseau.

> Remarque : Pour plus d'informations sur MTU, référez-vous à BGP Neighbor Flaps with MTU Troubleshooting .

Questions relatives à 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 000
```

L'identificateur de famille d'adresses (AFI) est une extension de capacité ajoutée par le protocole BGP multiprotocole (MP-BGP). Il établit une corrélation avec un protocole réseau spécifique, tel qu'IPv4, IPv6 et similaire, et une granularité supplémentaire via un identificateur de famille d'adresses suivantes (SAFI), tel que la monodiffusion et la multidiffusion. MBGP réalise cette séparation par les attributs de chemin BGP (PA) MP_REACH_NLRI et MP_UNREACH_NLRI. Ces attributs sont transportés à l'intérieur des messages de mise à jour BGP et sont utilisés pour transporter des informations d'accessibilité du réseau pour différentes familles d'adresses.

Le message vous donne les numéros de ces AFI/SAFI enregistrés par l'IANA :

- Numéros de famille d'adresses IANA
- Paramètres des identificateurs de famille d'adresses suivants (SAFI)

- Vérifiez la configuration BGP pour les familles d'adresses destinées aux deux côtés pour corriger les familles d'adresses indésirables.
- Utilisation neighbor ip-address dont-capability-negotiate aux deux extrémités. Pour plus d'informations, référez-vous à Capacités non prises en charge causant un dysfonctionnement d'homologue BGP.

Installation et sélection du chemin

Pour une meilleure explication sur le fonctionnement de BGP, et pour sélectionner le meilleur chemin, référez-vous à Algorithme de sélection du meilleur chemin BGP.

Saut suivant

Pour qu'une route soit installée dans notre table de routage, le saut suivant doit être accessible, sinon, même si le préfixe est sur notre table BGP Loc-RIB, il n'entre pas dans RIB. En tant que règle d'évitement de boucle, sur Cisco IOS/Cisco IOS XE, iBGP ne modifie pas l'attribut de saut suivant et laisse AS_PATH seul tandis qu'eBGP réécrit le saut suivant et ajoute son AS_PATH en préfixe.

Vous pouvez vérifier le saut suivant avec show ip bgp [prefix]. Il vous donne le prochain saut et le mot inaccessible. Dans l'exemple, il s'agit d'un préfixe annoncé par R1 via eBGP à R2 et appris par R3 via une connexion iBGP de R2.

<#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
```

Dans la sortie, le saut suivant est l'interface de sortie de R1 qui n'est pas connue par R3. Afin de résoudre cette situation, vous pouvez soit annoncer le saut suivant via IGP, route statique ou utiliser le neighbor ip-address next-hop-self sur l'homologue iBGP pour modifier l'adresse IP du tronçon suivant (qui est directement connectée). Dans un exemple de schéma, cette configuration doit être sur R2 ; le voisin vers R3 (voisin 10.0.23.3 next-hop-self).

Par conséquent, le saut suivant change (après un clear ip bgp 10.0.23.2 soft) à l'interface connectée directement (accessible) et le préfixe est installé.

<#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
```

Défaillance RIB

Cela se produit lorsque la route ne peut pas être installée dans le RIB global, ce qui entraîne une défaillance du RIB. Une raison courante est que le même préfixe est déjà sur RIB pour un autre protocole de routage avec une distance administrative inférieure, mais la raison exacte d'une défaillance RIB est visible avec la commande show ip bgp rib-failure. Pour plus d'explications, vous pouvez consulter ce lien :

> Remarque : vous pouvez identifier et corriger ce problème, comme expliqué dans Comprendre la défaillance RIB BGP et dans La commande bgp suppress-inactive.

Condition De Course

Le problème le plus fréquent est quand IGP est préféré à eBGP sur un scénario de redistribution mutuelle. Lorsqu'une route IGP est redistribuée dans BGP, elle est considérée comme générée localement par BGP et obtient un poids de 32768 par défaut. Tous les préfixes reçus d'un homologue BGP reçoivent un poids local de 0 par défaut. Par conséquent, si le même préfixe doit être comparé, le préfixe ayant le poids le plus élevé est installé dans la table de routage en fonction du processus de sélection du meilleur chemin BGP et c'est pourquoi la route IGP est installée sur RIB.

La solution pour ce problème, est de définir un poids plus élevé pour toutes les routes reçues de l'homologue BGP sous la configuration du routeur bgp :

```
<#root>
neighbor
ip-address
 weight 40000
```

Remarque : Pour une explication détaillée, référez-vous à Comprendre l'importance de l'attribut de chemin de poids BGP dans les scénarios de basculement réseau.

Autres questions

Homologue lent BGP

Il s'agit d'un homologue qui ne peut pas suivre le rythme auquel l'expéditeur génère des messages de mise à jour. Il existe de nombreuses raisons pour qu'un homologue présente ce problème : CPU élevé dans l'un des homologues, trafic excessif ou perte de trafic sur une liaison, ressource de bande passante, entre autres.

Remarque : pour identifier et corriger les problèmes d'homologues lents, référez-vous à Utiliser la fonctionnalité « Homologue lent » BGP pour résoudre les problèmes d'homologues lents.

Problèmes de mémoire

Le protocole BGP utilise la mémoire affectée au processus Cisco IOS pour gérer les préfixes réseau, les meilleurs chemins, les stratégies et toutes les configurations associées afin de fonctionner correctement. Les processus globaux sont affichés avec la commande show processes memory sorted:

<#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*
```

Le pool de processeurs est la mémoire utilisée ; environ 2,1 Go dans l'exemple. Ensuite, vous devez examiner la colonne Mise en attente pour identifier le sous-processus qui en contient la plupart. Ensuite, vous devez vérifier les sessions BGP que vous avez, le nombre de routes reçues et la configuration utilisée.

Étapes courantes pour réduire le stockage de mémoire par BGP :

- Filtrage BGP : s'il n'est pas nécessaire de recevoir une table BGP complète, utilisez des stratégies pour filtrer les routes et installer uniquement les préfixes dont vous avez besoin.
- Soft reconfiguration : recherchez neighbor ip_address soft-reconfiguration inbound sous configuration BGP ; cette commande vous permet de voir tous les préfixes reçus avant toute stratégie entrante (Adj-RIB-in). Cependant, cette table a besoin d'environ la moitié de la table RIB locale BGP actuelle pour stocker ces informations afin que vous puissiez éviter cette configuration sauf si elle est obligatoire, ou si vos préfixes actuels sont peu nombreux.

Remarque : Pour plus d'informations sur la façon d'optimiser BGP, référez-vous à Configurer des routeurs BGP pour des performances optimales et une consommation de mémoire réduite.

CPU élevé

Les routeurs utilisent des processus différents pour le fonctionnement du protocole BGP. Pour vérifier que le processus BGP est la cause d'une utilisation CPU élevée, utilisez la show process cpu sorted erasecat4000_flash:.

<#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
```

Voici les processus courants, les causes et les étapes générales pour surmonter l'utilisation élevée du CPU due au BGP :

- Routeur BGP : s'exécute une fois par seconde pour protéger une convergence plus rapide. C'est l'un des fils les plus importants. Il lit les messages de mise à jour bgp, valide les préfixes/réseaux et les attributs, met à jour la table de réseau/préfixe et la table d'attributs par AFI/SAFI, effectue le calcul du meilleur chemin parmi de nombreuses autres tâches. Un grand désordre de route est un scénario très courant qui conduit à cette situation.
- BGP Scanner : processus de faible priorité qui s'exécute toutes les 60 secondes par défaut. Ce processus vérifie l'intégralité de la table BGP pour vérifier l'accessibilité du tronçon suivant et met à jour la table BGP en conséquence, en cas de modification d'un chemin. Il est exécuté via la base d'informations de routage (RIB) à des fins de redistribution. Vérifiez l'évolutivité de la plate-forme, car plus de préfixes et de routes sont installés et plus TCAM est utilisé, plus de ressources sont nécessaires et, généralement, un périphérique surchargé mène à de telles situations.

Remarque : Pour plus d'informations sur la façon de dépanner ces deux processus, référez-vous à Dépanner le CPU élevé causé par le scanner BGP ou le processus du routeur.

- BGP I/O : s'exécute lorsque des paquets de contrôle BGP sont reçus et gère la mise en file d'attente et le traitement des paquets BGP. S'il y a un nombre excessif de paquets reçus dans la file d'attente BGP pendant une longue période, ou s'il y a un problème avec TCP, le routeur montre des symptômes de CPU élevé dû au processus d'E/S BGP. (Généralement, le routeur BGP est également élevé dans cette situation. Examinez le nombre de messages pour identifier l'homologue et capturez les paquets pour identifier la source de ces messages.)
- BGP Open : processus utilisé pour l'établissement d'une session. Pas un problème fréquent de CPU élevé à moins que la session soit bloquée dans l'état Ouvert.
- Événement BGP : est responsable du traitement du tronçon suivant. Recherchez les sauts suivants sur les préfixes reçus.

Informations connexes

- Assistance et documentation techniques - Cisco Systems
- Guide de configuration BGP
- Exemple de configuration d'authentification MD5 entre homologues BGP
- Capture de paquets intégrée

- Volets de voisinage BGP avec dépannage MTU
- Numéros de famille d'adresses IANA
- Paramètres des identificateurs de famille d'adresses suivants (SAFI)
- Des fonctionnalités non prises en charge provoquent un dysfonctionnement des homologues BGP
- Algorithme de sélection de la meilleure route BGP
- Comprendre RIB-failure BGP et la commande bgp suppress-inactive
- Comprendre l’importance de l’attribut de chemin poids du BGP dans les scénarios de basculement réseau
- Utiliser la fonctionnalité « Slow Peer » du protocole BGP pour résoudre les problèmes d'homologue lent
- Configuration des routeurs BGP pour des performances optimales et une consommation de mémoire réduite
- Dépanner le CPU élevé causé par le processus du scanner BGP ou du routeur

À propos de cette traduction
Cisco a traduit ce document en traduction automatisée vérifiée par une personne dans le cadre d'un service mondial permettant à nos utilisateurs d'obtenir le contenu d'assistance dans leur propre langue.

Il convient cependant de noter que même la meilleure traduction automatisée ne sera pas aussi précise que celle fournie par un traducteur professionnel.