Risoluzione dei problemi di base del protocollo Border Gateway

Sommario

Introduzione

Questo documento descrive come risolvere i problemi più comuni con il Border Gateway Protocol (BGP) e fornisce soluzioni e linee guida di base.

Prerequisiti

Requisiti

Non sono previsti prerequisiti specifici per questo documento. Per ulteriori informazioni, è possibile fare riferimento alla Guida alla configurazione BGP.

Componenti usati

Il documento può essere consultato per tutte le versioni software o hardware, ma i comandi sono validi per Cisco IOS® e Cisco IOS® XE.

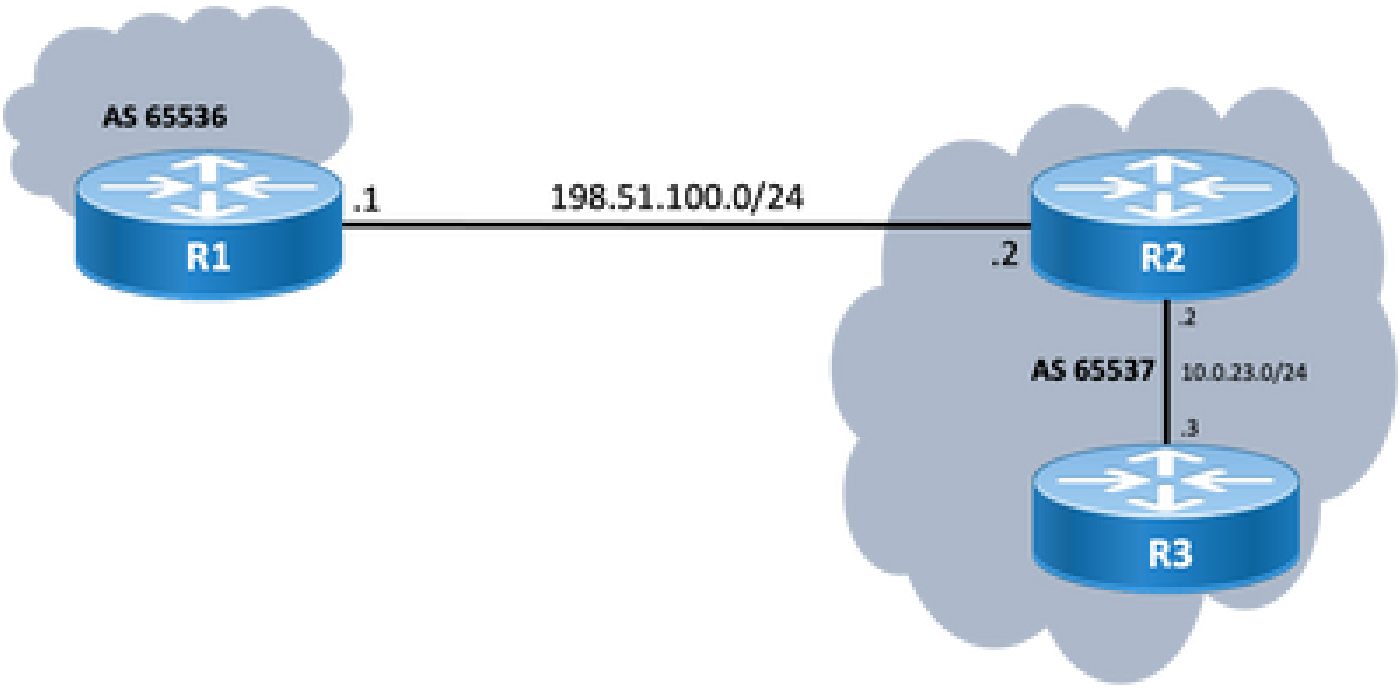
Le informazioni discusse in questo documento fanno riferimento a dispositivi usati in uno specifico ambiente di emulazione. Su tutti i dispositivi menzionati nel documento la configurazione è stata ripristinata ai valori predefiniti. Se la rete è operativa, valutare attentamente eventuali conseguenze derivanti dall'uso dei comandi.

Premesse

Questo documento descrive una guida di base per la risoluzione dei problemi più comuni del Border Gateway Protocol (BGP), fornisce azioni correttive, comandi/debug utili per rilevare la root cause dei problemi e best practice per evitare potenziali problemi. Tenere presente che non è possibile prendere in considerazione tutte le variabili e gli scenari possibili e che Cisco TAC potrebbe richiedere un'analisi più approfondita.

Topologia

Utilizzare questo diagramma della topologia come riferimento per gli output forniti in questo documento.



Scenari e problemi

Adiacente verso il basso

Se una sessione BGP non è attiva e non viene visualizzata, eseguire il comando `show ip bgp all`

summary command. Qui è possibile trovare lo stato corrente della sessione:

- Se la sessione non è nello stato attivo, può variare tra IDLE e ACTIVE (dipende dal processo di macchina a stati finiti).
- Se la sessione è attiva, verrà visualizzato il numero di prefissi ricevuti.

<#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
```

Nessuna connettività

Il primo requisito da garantire è la connettività tra entrambi i peer, in modo da poter stabilire una sessione TCP sulla porta 179. È possibile che siano collegati direttamente o meno. In questo caso, è utile un semplice ping. Se il peering viene stabilito tra interfacce di loopback, è necessario eseguire un ping di loopback. Se si esegue un test ping senza un loopback specifico come interfaccia di origine, l'indirizzo IP dell'interfaccia fisica in uscita viene usato come indirizzo IP di origine del pacchetto anziché come indirizzo IP di loopback del router.

Se il ping ha esito negativo, tenere presenti le seguenti cause:

- Nessuna route peer connessa o nessuna route: show ip route peer_IP_address possono essere utilizzate.
- Problema di layer 1: è necessario prendere in considerazione l'interfaccia fisica, l'SFP (connettore), il problema del cavo o esterno (trasporto e provider, se applicabile).
- Controllare eventuali firewall o elenchi degli accessi che possono bloccare la connessione.

Se il ping ha esito positivo, tenere presente quanto segue:

Problemi di configurazione

- Indirizzo IP errato o AS configurato: per IP errato , non viene visualizzato alcun messaggio di questo tipo, ma verificare che venga eseguita la configurazione corretta. Per AS errato, è necessario visualizzare un messaggio simile al seguente show logging

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

Controllare la configurazione BGP su entrambe le estremità per correggere i numeri AS o l'indirizzo IP del peer.

- ID router duplicato:

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

Controllare l'identificatore BGP su entrambi i lati tramite show ip bgp all summary e correggere il problema duplicato. Questa operazione può essere effettuata manualmente con il comando global bgp router-id X.X.X.X in configurazione router bgp. È buona norma verificare che l'ID del router sia impostato manualmente su un numero univoco.

- Origine BGP e TTL:

La maggior parte delle sessioni iBGP è configurata sulle interfacce di loopback raggiungibili tramite IGP. L'interfaccia di loopback deve essere definita in modo esplicito come origine. A tale scopo, eseguire il comando neighbor ip-address update-source interface-id .

Per il peer eBGP, le interfacce con connessione diretta sono in genere utilizzate per il peering e, in alternativa, è possibile verificare che Cisco IOS/Cisco IOS XE soddisfi questo requisito non tentare nemmeno di stabilire una sessione. Se si tenta di eseguire il loopback di eBGP su router connessi direttamente, questo controllo può essere disabilitato per un router adiacente specifico su entrambe le estremità tramite neighbor ip-address disable-connected-check .

Tuttavia, se tra i peer eBGP sono presenti più hop, è necessario un numero di hop corretto. Verificare la **neighbor ip-address ebgp-multihop [hop-count]**sia configurato con il numero di hop corretto in modo da poter stabilire la sessione.

Se non si specifica il numero di hop, il valore TTL predefinito per le sessioni iBGP è 255, mentre il valore TTL predefinito per le sessioni eBGP è 1.

Problemi della sessione TCP

Un'azione utile per verificare la porta 179 è un telnet manuale da un peer all'altro:

<#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'apertura o la connessione chiusa o la connessione rifiutata dall'host remoto indica che i pacchetti raggiungono l'estremità remota, quindi accertarsi che non vi siano problemi con il control plane all'estremità remota. In caso contrario, se la destinazione è irraggiungibile, controllare eventuali firewall o elenchi degli accessi che possano bloccare la porta TCP 179 o i pacchetti BGP o qualsiasi perdita di pacchetti sul percorso.

In caso di problemi di autenticazione, vengono visualizzati i messaggi seguenti:

```
%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
```

Controllare i metodi di autenticazione, la password e la configurazione correlata e per ulteriori informazioni sulla risoluzione dei problemi, fare riferimento all'esempio di configurazione dell'autenticazione MD5 tra peer BGP.

Se la sessione TCP non viene visualizzata, è possibile utilizzare i comandi successivi per l'isolamento:

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

Rimbalzi adiacenti

Se la sessione è attiva e inattiva, cercare `show log` e vedete alcuni scenari.

Interfaccia Flap

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

Come indicato nel messaggio, la causa dell'errore è la mancata disponibilità dell'interfaccia. Cercare eventuali problemi fisici sulla porta/SFP, sui cavi o sulle disconnessioni.

Timer di attesa scaduto

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

Si tratta di una situazione molto comune; indica che il router non ha ricevuto o elaborato un messaggio keepalive o alcun messaggio di aggiornamento prima della scadenza del timer di attesa. Il dispositivo invia un messaggio di notifica e chiude la sessione. I motivi più comuni di questo problema sono elencati di seguito:

- Problemi dell'interfaccia: cercare eventuali errori di input, interruzioni della coda di input o problemi fisici sulle interfacce connesse di entrambi i peer; `show interface` possono essere utilizzati a questo scopo.
- Perdita di pacchetti in transito: a volte, i pacchetti Hello possono essere scartati in transito, il modo migliore per essere certi che si tratti di un'acquisizione di pacchetti a livello di interfaccia.
    - È possibile utilizzare Embedded Packet Capture sui dispositivi Cisco IOS e Cisco IOS XE.
    - Se i pacchetti vengono visualizzati a livello di interfaccia, è necessario verificare che raggiungano il control plane, EPC sul piano di comando, o `debug bgp [vrf name] ipv4 unicast keepalives` è utile.
- CPU alta: una condizione di CPU alta può causare cadute sul control plane, `show processes cpu [sorted|history]` è utile per identificare il problema. A seconda della piattaforma in uso, il passaggio successivo per la risoluzione dei problemi è indicato nel documento di riferimento della CPU
- Problemi relativi alla policy CoPP: la risoluzione dei problemi varia a seconda della piattaforma ed è fuori ambito per questo documento.

- MTU non corrispondente: se il percorso contiene discrepanze MTU e i messaggi ICMP sono bloccati nel percorso tra l'origine e la destinazione, il rilevamento dell'MTU del percorso non funziona e può causare un flap nella sessione. Gli aggiornamenti vengono inviati con il valore MSS negoziato e un bit DF impostato. Se un dispositivo nel percorso o anche la destinazione non è in grado di accettare i pacchetti con MTU superiore, invia un messaggio di errore ICMP all'altoparlante BGP. Il router di destinazione attende che il pacchetto BGP keepalive o il pacchetto di aggiornamento BGP aggiorni il proprio timer di attesa.
    - È possibile selezionare il valore MSS negoziato con `show ip bgp neighbors ip_address`.

Un test Ping su un router adiacente specifico con df impostato può indicare se l'MTU è valida sul

percorso:

<#root>

ping 198.51.100.2 size

max_seg_size

df

Se vengono rilevati problemi di MTU, è necessario rivedere accuratamente la configurazione per garantire che i valori MTU siano coerenti su tutta la rete.

> Nota: per ulteriori informazioni sull'MTU, fare riferimento alla sezione BGP Neighbor Flaps with MTU Troubleshooting.

Problemi 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'identificatore della famiglia di indirizzi (AFI) è un'estensione di funzionalità aggiunta dal protocollo BGP (MP-BGP). Corrisponde a un protocollo di rete specifico, ad esempio IPv4, IPv6 e simili, e a una maggiore granularità tramite un successivo identificatore della famiglia di indirizzi (SAFI, Address-Family Identifier), ad esempio unicast e multicast. MBGP raggiunge questa separazione tramite gli attributi di percorso BGP (PA) MP_REACH_NLRI e MP_UNREACH_NLRI. Questi attributi vengono inseriti nei messaggi di aggiornamento BGP e vengono utilizzati per trasferire le informazioni sulla raggiungibilità della rete per le diverse famiglie di indirizzi.

Il messaggio riporta i numeri di questi AFI/SAFI registrati da IANA:

- Numeri famiglia di indirizzi IANA
- Parametri SAFI (Address Family Identifier) successivi

- Controllare la configurazione BGP per individuare le famiglie di indirizzi da entrambi i lati e correggere eventuali famiglie di indirizzi indesiderate.
- Utilizzo neighbor ip-address dont-capability-negotiate su entrambi i lati. Per ulteriori informazioni, fare riferimento a Funzionalità non supportate che causano problemi di funzionamento del peer BGP.

Installazione e selezione dei percorsi

Per una spiegazione migliore di come funziona BGP e per selezionare il percorso migliore, fare riferimento all'algoritmo di selezione del percorso migliore BGP.

Hop successivo

Affinché un percorso venga installato nella tabella di routing, è necessario che l'hop successivo sia raggiungibile. In caso contrario, anche se il prefisso si trova nella tabella BGP Loc-RIB, non verrà inserito nella tabella RIB. Come regola per evitare i loop, in Cisco IOS/Cisco IOS XE, iBGP non modifica l'attributo next-hop e lascia solo AS_PATH mentre eBGP riscrive l'hop successivo e lo precede.

È possibile controllare l'hop successivo con `show ip bgp [prefix]`. Vi dà l'hop successivo e una parola inaccessibile. Nell'esempio, questo è un prefisso annunciato da R1 tramite eBGP a R2 e appreso da R3 tramite connessione iBGP da 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
```

Nell'output, l'hop successivo è l'interfaccia in uscita di R1, che non è nota a R3. Per risolvere questo problema, è possibile pubblicizzare l'hop successivo tramite IGP, l'indirizzamento statico o usare il comando `neighbor ip-address next-hop-self` sul peer iBGP per modificare l'IP dell'hop successivo (connesso direttamente). Nell'esempio di diagramma, questa configurazione deve essere su R2; la porta adiacente verso R3 (adiacente all'hop successivo 10.0.23.3).

Di conseguenza, l'hop successivo cambia (dopo un `clear ip bgp 10.0.23.2 soft`) all'interfaccia collegata direttamente (raggiungibile) e prefisso è installato.

<#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
```

Errore RIB

Questo si verifica quando non è possibile installare il percorso nella NERVATURA GLOBALE, con conseguente errore della NERVATURA. Il motivo più comune è quando lo stesso prefisso è già su RIB per un altro protocollo di routing con distanza amministrativa inferiore, ma la causa esatta di un errore RIB è visibile con il comando show ip bgp rib-failure. Per una spiegazione più dettagliata, consultare questo link:

> Nota: è possibile identificare e risolvere questo problema come spiegato in Comprendere il fallimento del BGP RIB e Il comando bgp suppress-inactive.

Condizione di gara

Il problema più comune è quando l'IGP viene preferito a eBGP nello scenario di ridistribuzione reciproca. Quando una route IGP viene ridistribuita in BGP, viene considerata generata localmente da BGP e per impostazione predefinita riceve un peso di 32768. A tutti i prefissi ricevuti da un peer BGP viene assegnato per impostazione predefinita un peso locale pari a 0. Pertanto, se è necessario confrontare lo stesso prefisso, il prefisso con il peso più alto viene installato nella tabella di routing in base al processo di selezione del miglior percorso BGP ed è per questo che la route IGP viene installata su RIB.

La soluzione di questo problema è impostare un peso maggiore per tutte le route ricevute dal peer BGP in una configurazione bgp del router:

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

Nota: per una spiegazione dettagliata, consultare il documento sull'importanza dell'attributo del percorso di peso BGP negli scenari di failover di rete.

Altri problemi

BGP Slow Peer

Si tratta di un peer che non riesce a tenere il passo con la frequenza con cui il mittente genera messaggi di aggiornamento. Un peer può presentare questo problema per diversi motivi, ad esempio un elevato livello di CPU in uno dei peer, un eccesso di traffico o una perdita di traffico su un collegamento, una risorsa di larghezza di banda e così via.

Nota: per identificare e correggere i problemi dei peer lenti, fare riferimento a Uso della funzione BGP "Slow Peer" per risolvere i problemi dei peer lenti.

Problemi di memoria

Per funzionare correttamente, BGP utilizza la memoria assegnata al processo Cisco IOS per mantenere i prefissi di rete, i percorsi migliori, le policy e tutte le configurazioni correlate. I processi complessivi vengono visualizzati con il comando 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*
```

Il pool di processori è la memoria utilizzata; nell'esempio, circa 2,1 GB. Quindi, è necessario esaminare la colonna Detenzione per identificare il sottoprocesso che contiene la maggior parte di esso. Quindi, è necessario controllare le sessioni BGP in uso, il numero di route ricevute e la configurazione utilizzata.

Passaggi comuni per ridurre la capacità di memoria di BGP:

- Filtro BGP: se non è necessario ricevere una tabella BGP completa, utilizzare i criteri per filtrare le route e installare solo i prefissi necessari.
- Riconfigurazione soft: cercare la riconfigurazione soft dell'indirizzo_ip adiacente in entrata nella configurazione BGP. Questo comando consente di visualizzare tutti i prefissi ricevuti prima di qualsiasi criterio in entrata (Adj-RIB-in). Tuttavia, questa tabella richiede circa la metà della tabella RIB locale BGP corrente per memorizzare queste informazioni, in modo da evitare questa configurazione a meno che non sia obbligatoria o i prefissi attuali siano pochi.

Nota: per ulteriori informazioni su come ottimizzare il protocollo BGP, fare riferimento a Configurazione dei router BGP per prestazioni ottimali e consumo di memoria ridotto.

Elevato consumo della CPU

I router utilizzano processi diversi per il funzionamento di BGP. Per verificare che il processo BGP sia la causa di un elevato utilizzo della CPU, utilizzare il show process cpu sorted

<#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
```

Di seguito sono riportati i processi, le cause e i passaggi generali comuni per superare l'elevato utilizzo della CPU dovuto a BGP:

- Router BGP: in esecuzione una volta al secondo per garantire una convergenza più rapida. Si tratta di uno dei temi più importanti. Legge i messaggi di aggiornamento bgp, convalida i prefissi/reti e gli attributi, aggiorna la tabella di prefissi/rete/prefissi AFI/SAFI e la tabella di attributi, esegue il calcolo del miglior percorso tra molte altre attività.
  L'enorme cambiamento del percorso è uno scenario molto comune che porta a questa situazione.
- Scanner BGP: processo a bassa priorità eseguito per impostazione predefinita ogni 60 secondi. Questo processo controlla l'intera tabella BGP per verificare la raggiungibilità dell'hop successivo e aggiorna la tabella BGP di conseguenza, in caso di modifiche per un percorso. Viene eseguito tramite la base di informazioni di routing (RIB, Routing Information Base) a scopo di ridistribuzione.
  Verificare la scalabilità della piattaforma, poiché vengono installati più prefissi e percorsi e viene utilizzato TCAM, sono necessarie più risorse e, in genere, un dispositivo sovraccarico porta a situazioni di questo tipo.

Nota: per ulteriori informazioni su come risolvere i problemi relativi a questi due processi, consultare il documento sulla risoluzione dei problemi relativi alla CPU elevata causata dal processo dello scanner BGP o del router.

- I/O BGP: viene eseguito quando i pacchetti di controllo BGP vengono ricevuti e gestisce l'accodamento e l'elaborazione dei pacchetti BGP. Se i pacchetti ricevuti nella coda BGP sono eccessivi per un lungo periodo o se si verifica un problema con il TCP, il router mostra sintomi di CPU alta a causa del processo di I/O BGP. (In genere, anche il router BGP è ad alta velocità in questa situazione. Esaminare i conteggi dei messaggi per identificare il peer e acquisire i pacchetti per identificare l'origine di questi messaggi.)
- BGP Open: processo utilizzato per stabilire la sessione. Non è un problema di CPU elevato a meno che la sessione non sia bloccata in stato aperto.
- Evento BGP: responsabile dell'elaborazione dell'hop successivo. Cercare i flap dell'hop successivo sui prefissi ricevuti.

Informazioni correlate

- Documentazione e supporto tecnico – Cisco Systems
- Guida alla configurazione BGP
- Esempio di configurazione dell'autenticazione MD5 tra peer BGP
- Embedded Packet Capture
- BGP Neighbor Flaps con MTU Troubleshooting

- Numeri famiglia di indirizzi IANA
- Parametri SAFI (Address Family Identifier) successivi
- Funzionalità non supportate che causano malfunzionamenti del peer BGP
- Algoritmo di selezione del miglior percorso BGP
- Comprendere il fallimento del BGP RIB e il comando bgp suppress-inactive
- Importanza dell'attributo Weight Path del protocollo BGP negli scenari di failover della rete
- Utilizzare la funzione BGP "Slow Peer" per risolvere i problemi dei peer lenti
- Configurazione dei router BGP per prestazioni ottimali e consumo di memoria ridotto
- Risoluzione dei problemi di CPU elevata causati dal processo dello scanner o del router BGP

Informazioni su questa traduzione

Cisco ha tradotto questo documento utilizzando una combinazione di tecnologie automatiche e umane per offrire ai nostri utenti in tutto il mondo contenuti di supporto nella propria lingua. Si noti che anche la migliore traduzione automatica non sarà mai accurata come quella fornita da un traduttore professionista. Cisco Systems, Inc. non si assume alcuna responsabilità per l’accuratezza di queste traduzioni e consiglia di consultare sempre il documento originale in inglese (disponibile al link fornito).