

Solucionar problemas de alta utilização da CPU causados pelo scanner BGP ou pelo processo do roteador

Contents

[Introduction](#)

[Prerequisites](#)

[Requirements](#)

[Componentes Utilizados](#)

[Conventions](#)

[Informações de Apoio](#)

[Entender os processos de BGP](#)

[CPU elevada devido ao scanner de BGP](#)

[CPU alta devido ao processo do roteador BGP](#)

[Melhorias no desempenho](#)

[Conexões de Fila para Par TCP](#)

[Grupos de paridade BGP](#)

[MTU de caminho e o comando ip tcp path-mtu-discovery](#)

[Aumentar filas de entrada de interface](#)

[Melhorias adicionais no Cisco IOS](#)

[Procedimentos de Troubleshooting](#)

[Informações Relacionadas](#)

Introduction

Este documento descreve como solucionar problemas de causa de leituras altas de CPU quando o scanner ou roteador BGP é usado.

Prerequisites

Requirements

Este documento requer conhecimento de como interpretar o comando `show process cpu`.

Componentes Utilizados

As informações contidas neste documento são baseadas no software IOS® da Cisco versão 12.0.

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. Se a rede estiver ativa, certifique-se de que você entenda o impacto potencial de qualquer comando.

Conventions

Consulte as [Convenções de Dicas Técnicas da Cisco para obter mais informações sobre convenções de documentos](#).

Informações de Apoio

Este documento descreve situações em que um roteador Cisco IOS pode experimentar alta utilização da CPU devido ao processo do roteador Border Gateway Protocol (BGP) ou ao processo do scanner BGP, como mostrado na saída do comando **show process cpu**. A duração da condição elevada de CPU varia conforme o número de condições, particularmente o tamanho da tabela de roteamento à Internet e o número de rotas que um determinado roteador mantém em seu roteamento e tabelas de BGP. O comando **show process cpu** mostra a média de utilização da CPU durante os últimos cinco segundos, o último minuto e os últimos cinco minutos. Os números de utilização da CPU não oferecem uma indicação linear real da utilização no que diz respeito à carga oferecida.

Estes são alguns dos principais motivos:

- Numa rede do mundo real, a CPU tem de gerenciar várias funções de manutenção do sistema, por exemplo, o gerenciamento de rede.
- O CPU tem que processar atualizações de roteamento periódico e acionados por eventos.
- Há outras operações internas de sobrecarga do sistema, como polling de disponibilidade de recursos, que não são proporcionais à carga de tráfego.

Você também pode usar o comando **show processes cpu** para obter alguma indicação de atividade da CPU.

Note: Para obter mais informações sobre os comandos show, consulte a [Referência de Comandos do Cisco IOS Configuration Fundamentals](#)

Entender os processos de BGP

Um processo do Cisco IOS, em geral, consiste em segmentos individuais e dados associados que executam tarefas, como manutenção do sistema, comutação de pacotes e implementação de protocolos de roteamento. Vários processos do Cisco IOS executados no roteador permitem que o BGP seja executado. Usar o comando **show process cpu | incluir o comando BGP** para ver a quantidade de utilização da CPU devido aos processos BGP.

Esta tabela lista a função dos processos BGP e mostra que cada processo é executado em momentos diferentes, e esses momentos dependem das tarefas que são tratadas. Como os processos do scanner BGP e do roteador BGP são responsáveis por uma grande quantidade de cálculos, você pode ver alta utilização da CPU devido a qualquer um desses processos. As próximas seções discutem esses processos com mais detalhes.

Nome do processo	Descrição	Intervalo
BGP aberto	Realiza o estabelecimento de peer de BGP.	Na inicialização, quando uma conexão TCP com um par BGP estabelecida.
BGP I/O	Gerencia pacotes BGP que estão	À medida que pacotes de contr

Scanner BGP	na fila e são processados, como ATUALIZAÇÕES e KEEPALIVES. Passa pela tabela de BGP e confirma a alcançabilidade dos próximos saltos. O scanner BGP também verifica o anúncio condicional para determinar se o BGP anuncia prefixos de condição e executa a redução de rota. Em um ambiente VPN MPLS, o scanner BGP importa e exporta rotas para uma instância específica de roteamento e encaminhamento (VRF) de VPN.	BGP são recebidos. Após um minuto.
Roteador BGP	Calcula o melhor caminho BGP e processa qualquer rotatividade de rotas . Também envia e recebe rotas, estabelece pares e interage com o Routing Information Base (RIB).	Uma vez por segundo e quando peer de BGP é adicionado, removido ou configurado por software.

CPU elevada devido ao scanner de BGP

Pode-se esperar alta utilização da CPU devido ao processo de scanner BGP por curtos períodos em um roteador que transporta uma grande tabela de roteamento da Internet. Uma vez a cada minuto, o scanner BGP passa pela tabela RIB do BGP e executa importantes tarefas de manutenção. Essas tarefas incluem o exame do próximo salto mencionado na tabela BGP do roteador e verifica se os dispositivos do próximo salto podem ser alcançados. Portanto, uma tabela BGP grande leva um tempo equivalente para ser examinada e validada.

Como o processo do scanner BGP passa por toda a tabela BGP, a duração da condição de CPU alta varia com o número de vizinhos e o número de rotas aprendidas por vizinho. Use os comando `show ip bgp summary` e `show ip route summary` para capturar essas informações. O processo do scanner BGP passa pela tabela BGP para atualizar qualquer estrutura de dados e passa pela tabela de roteamento para fins de redistribuição de rota. (Neste contexto, a tabela de roteamento também é conhecida como base de informações de roteamento (RIB), que o roteador gera quando você executa o comando `show ip route`). Ambas as tabelas são armazenadas separadamente na memória do roteador e podem ser grandes e consumir ciclos de CPU.

O próximo exemplo de saída do comando `debug ip bgp updates` captura uma execução do scanner BGP:

```
router#
2d17h: BGP: scanning routing tables
2d17h: BGP: 10.0.0.0 computing updates, neighbor version 8,
      table version 9, starting at 0.0.0.0
2d17h: BGP: 10.0.0.0 update run completed, ran for 0ms, neighbor
      version 8, start version 9, throttled to 9, check point net 0.0.0.0
2d17h: BGP: 10.1.0.0 computing updates, neighbor version 8,
      table version 9, starting at 0.0.0.0
2d17h: BGP: 10.1.0.0 update run completed, ran for 4ms, neighbor
      version 8, start version 9, throttled to 9, check point net 0.0.0.0
router#
```

Enquanto o scanner BGP é executado, os processos de baixa prioridade devem esperar por mais tempo para acessar a CPU. Um processo de baixa prioridade controla os pacotes ICMP (protocolo de mensagem de controle da Internet), como pings. Os pacotes destinados ao roteador ou originados dele podem experimentar uma latência maior do que a esperada, já que o processo ICMP deve esperar atrás do scanner BGP. O ciclo é que o scanner BGP é executado por algum tempo e se suspende, e então o ICMP é executado. Por outro lado, os pings enviados através de um roteador devem ser comutados através do Cisco Express Forwarding (CEF) e não devem experimentar nenhuma latência adicional. Ao solucionar problemas de picos periódicos de latência, compare os tempos de encaminhamento de pacotes encaminhados através de um roteador com os pacotes processados diretamente pela CPU no roteador.

Note: Os comandos ping que especificam opções IP, como rota de registro, também exigem que a CPU os processe diretamente e isso pode causar atrasos de encaminhamento mais longos.

Usar o comando **show process | include bgp scanner** para ver a prioridade da CPU. O valor de **Lsi** na próxima saída de exemplo usa **L** para se referir a um processo de baixa prioridade.

```
6513#show processes | include BGP Scanner
172 Lsi 407A1BFC      29144      29130      1000 8384/9000  0 BGP Scanner
```

CPU alta devido ao processo do roteador BGP

O processo do roteador BGP é executado aproximadamente uma vez por segundo para verificar o trabalho. A convergência de BGP define a duração entre o momento em que o primeiro peer de BGP é estabelecido e o ponto em que o BGP é convergido. Para garantir os menores tempos de convergência possíveis, o roteador BGP consome todos os ciclos de CPU livres. Entretanto, após a sua inicialização, ele abre mão da CPU (ou suspende) imediatamente.

O tempo de convergência é uma medição direta do tempo que o roteador BGP gasta na CPU, não do tempo total. Este procedimento mostra uma condição de CPU alta durante a convergência do BGP e troca prefixos BGP com dois peers BGP externos (eBGP).

1. Capture uma linha de base para utilização normal da CPU antes de iniciar o teste.

```
router#show process cpu
CPU utilization for five seconds: 0%/0%; one minute: 4%; five minutes: 5%
```

2. Quando o teste começa, o CPU atinge 100% de uso. O comando **show process cpu** mostra que a condição de alta utilização da CPU é causada pelo roteador BGP, indicado por 139 (o ID de processo do Cisco IOS para o roteador BGP) na próxima saída.

```
router#show process cpu
CPU utilization for five seconds: 100%/0%; one minute: 99%; five minutes: 81%
```

```
!--- Output omitted. 139 6795740 1020252 6660 88.34% 91.63% 74.01% 0 BGP Router
```

3. Você pode monitorar e capturar várias saídas dos comandos **show ip bgp summary** e **show process cpu** neste momento. O comando **show ip bgp summary** captura o estado dos vizinhos BGP.

```
router#show ip bgp summary
Neighbor    V    AS  MsgRcvd  MsgSent    TblVer  InQ  OutQ  Up/Down  State/PfxRcd
10.0.0.0    4   64512  309453   157389     19981    0   253  22:06:44  111633
10.1.0.0    4   65101  188934    1047      40081    41    0  00:07:51  58430
```

4. Quando o roteador conclui a troca de prefixo com seus peers BGP, as taxas de utilização da CPU retornam aos níveis normais. As médias computadas de um minuto e cinco minutos

também podem se acalmar e podem mostrar níveis superiores ao normal por um período maior que a taxa de cinco segundos.

```
router#show process cpu
```

```
  CPU utilization for five seconds: 3%/0%; one minute: 82%; five minutes: 91%
```

5. Use a saída capturada dos comandos show anteriores para calcular o tempo de convergência do BGP. Em particular, use a coluna **Up/Down** do comando **show ip bgp summary** e compare os tempos de início e parada da condição de CPU alta.

Geralmente, a convergência de BGP pode levar vários minutos quando uma grande tabela de roteamento de Internet. é trocado

Note: A alta utilização da CPU no dispositivo também pode ocorrer devido à instabilidade da tabela BGP. Isso acontece se o roteador recebe duas cópias da tabela de roteamento, uma do peering EBGP com o ISP e a outra do peering IBGP na rede. A causa raiz disso é a quantidade de memória no dispositivo. A Cisco recomenda um mínimo de 1 Gig de RAM para uma única cópia da tabela de roteamento da Internet. Para contornar essa instabilidade, aumente a RAM no dispositivo ou filtre os prefixos para que a tabela BGP e a memória ocupada por ela sejam aliviadas.

Melhorias no desempenho

À medida que o número de rotas na tabela de roteamento da Internet cresce, o mesmo ocorre com o tempo que leva para o BGP convergir. Em geral, a convergência é definida como o processo em que todas as tabelas de roteamento são colocadas em um estado de consistência. O BGP é considerado convergido quando estas condições são verdadeiras:

- Todas as rotas foram aceitas.
- Todas as rotas foram instaladas na tabela de roteamento.
- A versão da tabela para todos os peers é igual à versão da tabela de BGP.
- InQ e OutQ para todos os peers é zero.

Esta seção descreve algumas melhorias de desempenho do Cisco IOS para reduzir o tempo de convergência de BGP, que resultam em uma redução de uma condição de CPU alta devido a um processo BGP.

Conexões de Fila para Par TCP

O BGP agora enfileira dados agressivamente do BGP OutQ para o soquete TCP para cada peer até que os OutQs tenham sido completamente drenados. Como o BGP faz envios em uma taxa mais rápida, o BGP converge mais rapidamente.

Grupos de paridade BGP

Embora ajudem a simplificar a configuração do BGP, os grupos de peer do BGP também podem melhorar a escalabilidade. Todos os membros de grupo de peer devem compartilhar uma política de saída comum. Assim, os mesmos pacotes de atualização podem ser enviados para cada membro do grupo, e isso reduz o número de ciclos de CPU que o BGP requer para anunciar rotas aos pares. Em outras palavras, com grupos de peer, o BGP vai para a tabela BGP apenas no líder do grupo de peer, filtra os prefixos através das políticas de saída e gera atualizações que envia para o líder do grupo de peer. Por sua vez, o líder replica as atualizações para membros do grupo com os quais está sincronizado. Sem grupos de peer, o BGP deve percorrer a tabela para cada peer, filtrar prefixos através de políticas de saída e gerar atualizações que são enviadas

apenas para o peer.

MTU de caminho e o comando `ip tcp path-mtu-discovery`

Todas as sessões TCP são limitadas por um limite no número de bytes que podem ser transportados em um único pacote. Esse limite, conhecido como Tamanho máximo de segmento (MSS), é de 536 bytes por padrão. Em outras palavras, o TCP divide os pacotes em uma fila de transmissão em blocos de 536 bytes antes de passar os pacotes para a camada IP. Usar o comando `show ip bgp neighbors | include max data` para exibir o MSS de peers BGP:

```
Router#show ip bgp neighbors | include max data
Datagrams (max data segment is 536 bytes):
Datagrams (max data segment is 536 bytes):
Datagrams (max data segment is 536 bytes):
Datagrams (max data segment is 536 bytes):
```

A vantagem de um MSS de 536 bytes é a pouca probabilidade de que os pacotes sejam fragmentados em um dispositivo de IP ao longo do caminho para o destino já que a maioria dos links usa um MTU de, no mínimo, 1500 bytes. A desvantagem é que os pacotes menores aumentam a quantidade de largura de banda usada para transportar a sobrecarga. Como o BGP cria uma conexão TCP para todos os peers, um MSS de 536 bytes afeta os tempos de convergência do BGP.

A solução é ativar o recurso Path MTU (PMTU) com o comando `ip tcp path-mtu-discovery`. Você pode usar esse recurso para determinar dinamicamente o tamanho do valor do MSS e, enquanto isso, não criar pacotes que precisem ser fragmentados. O PMTU permite que o TCP determine o menor tamanho de MTU entre todos os links em uma sessão TCP. Em seguida, o TCP usa esse valor de MTU, menos espaço para os cabeçalhos IP e TCP, como o MSS da sessão. Se uma sessão TCP atravessa apenas segmentos Ethernet, o MSS é 1460 bytes. Se ele atravessar apenas segmentos de Pacote sobre SONET (POS), o MSS será de 4.430 bytes. O aumento no MSS de 536 para 1460 ou 4430 bytes reduz a sobrecarga de TCP/IP, o que ajuda a convergência de BGP mais rápida.

Depois de habilitar a PMTU, use novamente o comando `show ip bgp neighbors | include max data` para ver o valor de MSS por peer:

```
Router#show ip bgp neighbors | include max data
Datagrams (max data segment is 1460 bytes):
Datagrams (max data segment is 1460 bytes):
Datagrams (max data segment is 1460 bytes):
Datagrams (max data segment is 1460 bytes):
```

Aumentar filas de entrada de interface

Se o BGP anunciar milhares de rotas para muitos peers, o TCP deve transmitir milhares de pacotes em uma curta duração. Os peers BGP recebem esses pacotes e enviam confirmações TCP para o alto-falante BGP que anuncia, o que faz com que o alto-falante BGP receba uma inundação de TCP ACKs em um curto período de tempo. Se os ACKs chegarem a uma taxa que é muito alta para o processador de rota, os pacotes voltam para as filas de interface de entrada. Por padrão, as interfaces de roteador usam um tamanho de fila de entrada de 75 pacotes. Além disso, pacotes de controle especiais como ATUALIZAÇÕES DE BGP usam uma fila especial com SPD (Descarte seletivo de pacotes). Essa fila especial mantém 100 pacotes. Quando ocorre a convergência de BGP, os ACKs de TCP podem preencher rapidamente os 175 pontos de buffer

de entrada e os novos pacotes que chegam devem ser descartados. Em roteadores com 15 ou mais pares BGP e troca da tabela de roteamento de Internet completa, mais de 10.000 quedas por interface por minuto podem ser vistas. Este é um exemplo da saída de um roteador 15 minutos depois da reinicialização:

```
Router#show interface pos 8/0 | include input queue
Output queue 0/40, 0 drops; input queue 0/75, 278637 drops
Router#
```

Se você aumentar a profundidade da fila de entrada da interface (com o comando **hold-queue in**), isso ajudará a reduzir o número de ACKs TCP descartados, o que reduz a quantidade de trabalho que o BGP deve fazer para convergir. Normalmente, um valor de 1000 resolve problemas causados por quedas de fila de entrada.

Note: Use isso conscientemente, pois o incremento da fila de entrada pode adicionar algum atraso.

Melhorias adicionais no Cisco IOS

O Cisco IOS inclui várias otimizações para o código do grupo de pares BGP para melhorar o empacotamento e a replicação de atualizações. Antes de analisar essas melhorias, analise o pacote de atualizações e a replicação com mais detalhes.

Uma atualização de BGP consiste em uma combinação de atributos (como MED = 50 e LOCAL_PREF = 120) e uma lista de prefixos Network Layer Reachability Information (NLRI) que compartilham essa combinação de atributos. Quanto mais prefixos NLRI o BGP puder listar em uma única atualização, mais rápido o BGP poderá convergir, já que a sobrecarga (como cabeçalhos IP, TCP e BGP) é reduzida. **O empacotamento de atualização** se refere ao empacotamento de NLRI em atualizações BGP. Por exemplo, se uma tabela de BGP contém 100.000 rotas com 15.000 combinações de atributos exclusivos, o BGP precisa enviar somente 15.000 atualizações se o NLRI for compactado com 100 por cento de eficiência.

Note: Zero por cento de eficiência de empacotamento significa que o BGP precisaria enviar 100.000 atualizações neste ambiente.

Use o comando `show ip bgp peer-group` para exibir a eficiência das atualizações BGP.

Se um membro do grupo de peer estiver em sincronia, um roteador BGP pega uma mensagem de atualização formatada para o líder do grupo de peer e a replica para o membro. É muito mais eficiente reproduzir uma atualização para um membro de grupo de peer do que reformatar a atualização. Por exemplo, suponha que um grupo de correspondentes tenha 20 membros e todos eles precisem receber 100 mensagens BGP. A replicação de cem por cento significa que um roteador BGP formata as 100 mensagens para o líder do grupo de correspondentes e replica essas mensagens para os outros 19 membros do grupo de correspondentes. Para confirmar os aprimoramentos de replicação, compare o número de mensagens replicadas com o número de mensagens formatadas, conforme mostrado pelo comando **show ip bgp peer-group**. As melhorias fizeram uma grande diferença nos tempos de convergência e permitem que o BGP suporte muitos mais peers.

Por exemplo, use o comando **show ip bgp peer-group** para verificar a eficiência do empacotamento de atualizações e da replicação de atualizações. A próxima saída é de um teste

de convergência com 6 grupos de peer, 20 peers em cada um dos primeiros 5 grupos de peer (peers eBGP) e 100 peers no sexto grupo de peer (peers BGP internos (iBGP)). Além disso, a tabela BGP usada tinha 36.250 combinações de atributos.

O próximo exemplo de saída do comando **show ip bgp peer-group | include replicated** em um roteador que executa o Cisco IOS 12.0(18)S exibe estas informações:

```
Update messages formatted 836500, replicated 1668500
Update messages formatted 1050000, replicated 1455000
Update messages formatted 660500, replicated 1844500
Update messages formatted 656000, replicated 1849000
Update messages formatted 501250, replicated 2003750
```

!-- The first five lines are for eBGP peer groups. Update messages formatted 2476715, replicated 12114785 !-- The last line is for an iBGP peer group.

Para calcular a taxa de replicação para cada grupo de peer, divida o número de atualizações replicadas pelo número de atualizações formatadas:

$1668500/836500 = 1,99$ $1455000/1050000 = 1,38$ $1844500/660500 = 2,79$ $1849000/656000 = 2,81$ $2003750/501250 = 3,99$ $12114785/2476715 = 4,89$

- Se o BGP foi replicado perfeitamente, os grupos de peer do eBGP têm uma taxa de replicação de 19, pois há 20 peers no grupo de peer. A atualização é formatada para o líder do grupo de correspondentes e depois replicada para os outros 19 correspondentes. Isso fornece uma taxa de replicação ideal de 19. A taxa de replicação ideal para o grupo de peers do iBGP seria 99, já que existem 100 peers.
- Se o BGP empacotou as atualizações perfeitamente, então há apenas 36.250 atualizações formatadas. Você só precisa gerar 36.250 atualizações para cada grupo de peer porque esse é o número de combinações de atributo na tabela BGP. O grupo de peers de iBGP por si só formata aproximadamente 2.500.000 atualizações, enquanto que cada um dos grupos de peers de eBGP gera cerca de 500.000 a 1.000.000 de atualizações.

Em um roteador que executa o Cisco IOS 12.0(19)S, o **show ip bgp peer-group | include replicated** command fornece estas informações:

```
Update messages formatted 36250, replicated 688750
Update messages formatted 36250, replicated 688750
Update messages formatted 36250, replicated 688750
Update messages formatted 36250, replicated 688750
Update messages formatted 36250, replicated 688750
Update messages formatted 36250, replicated 3588750
```

Note: O empacotamento de atualização está ótimo. Exatamente 36.250 atualizações são formatadas para cada grupo de peer. $688750/36250 = 19$ $688750/36250 = 19$ $688750/36250 = 19$ $688750/36250 = 19$ $688750/36250 = 19$ $3588750/36250 = 99$

Note: A replicação de atualização também é perfeita.

Procedimentos de Troubleshooting

Use estes procedimentos para solucionar problemas de alta utilização da CPU devido ao scanner BGP ou roteador BGP:

- Colete informações sobre a topologia do BGP. Determine o número de peers BGP e o número de rotas que são anunciadas por cada peer. A duração da condição de CPU alta é baseada de forma razoável em seu ambiente?
- Determine quando a alta utilização da CPU ocorre. Isso coincide com uma movimentação programada regularmente da tabela BGP?
- A CPU alta seguiu um flap de interface? Você pode usar o comando **show ip bgp dampening flap-statistics** se o ajuste estiver habilitado.
- Emita um ping através do roteador e, em seguida, a partir do roteador. Os ecos ICMP são tratados como um processo de baixa prioridade. O documento [Entendendo os comandos ping e traceroute](#) explica isso em mais detalhes. Assegure que o encaminhamento regular não seja afetado.
- Você precisa garantir que os pacotes possam seguir um caminho de encaminhamento rápido ao verificar se a switching rápida e/ou o CEF estão habilitados nas interfaces de entrada e de saída. Certifique-se de não ver o comando **no ip route-cache cef** na interface ou o comando **no ip cef** na configuração global. Para habilitar o CEF no modo de configuração global, use o comando **ip cef**.
- Verifique a escala da plataforma porque, na maioria dos casos, é devido a um dispositivo sobrecarregado que leva a essas situações. Além disso, garantiu-se de que haja espaço de memória endereçável de conteúdo ternário (TCAM) adequado disponível no roteador.
- Verifique se há memória suficiente no roteador. De acordo com a recomendação, ter um mínimo de 1 GB de DRAM atribuído ao espaço do Cisco IOS por par de BGP que envia a tabela de roteamento de Internet completa. O espaço DRAM mencionado aqui é apenas a memória necessária para o BGP. Outros recursos executados no roteador podem exigir espaço adicional.

Informações Relacionadas

- [Página de Suporte do IP Routing](#)
- [Suporte Técnico - Cisco Systems](#)