Considerações de Escala RR BGP e Monitoramento de KPI

Contents

Introdução

Este documento descreve os principais contribuidores para a escala máxima que um Border Gateway Protocol (BGP) Route-Refletors (RR) pode alcançar e orientação sobre o monitoramento de desempenho BGP RR.

Seleção de plataforma de hardware/software

Um BGP RR de alta escala normalmente não está no caminho de encaminhamento de pacotes que transportam serviços fornecidos por um Provedor de Serviços de Internet. Portanto, os requisitos de hardware para um BGP RR e roteadores que estão predominantemente encaminhando pacotes no caminho de dados são diferentes. Os roteadores padrão são construídos com um poderoso elemento de encaminhamento de caminho de dados e um elemento de caminho de controle comparativamente moderado. Um BGP RR executa todas as suas tarefas em um plano de controle.
Dentro da família de produtos Cisco IOS® XR, você pode escolher entre 3 tipos de plataformas HW/SW para uma função BGP RR:

Roteador Cisco IOS XR físico	Dispositivo Cisco IOS XRv 9000	Roteador Cisco IOS XRv 9000 (também conhecido como XRv9k)
• Capacidade do plano de controle moderada (geralmente entre 2 e 6 núcleos de CPU alocados para RP XR VM) • Capacidade não utilizada do caminho de dados	• Alta capacidade de plano de controle (em dispositivos com base no Cisco UCS M5, 36 núcleos de CPU são dedicados a RP XR VM) • Divisão igual entre capacidade de caminho de dados e caminho de controle. • Imagem XRv9k executada em barebone para desempenho máximo	• Capacidade personalizável do plano de controle • Divisão igual entre a potência do caminho de dados e do caminho de controle ao usar a imagem BGP RR. • Uma camada adicional de virtualização afeta o desempenho.

A partir desta redação, o XRv9k Appliance é a escolha ideal de plataforma para BGP RR porque fornece a maior capacidade de plano de controle com desempenho máximo.

Considerações Sobre Escalabilidade E Desempenho

A escala suportada de entidades de plano de dados é relativamente fácil de expressar porque o desempenho do elemento de caminho de dados raramente depende da escala. Por exemplo, uma pesquisa de TCAM leva o mesmo tempo, independentemente do número de entradas de TCAM ativas.
A escala suportada de entidades de plano de controle é frequentemente muito mais complexa porque a escala e o desempenho estão interconectados. Considere um BGP RR com rotas de 1M. O trabalho que um processo BGP deve realizar para manter esta tabela BGP depende de:

1. Quantos peers BGP estão ativos?
2. Quais famílias de endereços estão ativas?
3. Como eles são distribuídos em grupos de atualização?
4. A complexidade das RPLs (Políticas de Rota)
5. Frequência de atualizações (atualizações recebidas e também enviadas - intervalo de anúncio).
6. MSS de TCP, MTU de Interface/Caminho - o ajuste ajudará no melhor desempenho
7. Se for RP duplo, o NSR está habilitado
8. Qualquer par lento conhecido, que não esteja em um grupo de atualização separado
9. valor de retardo de gatilho Nexthop

Número de pares BGP

O número de peers BGP é geralmente a primeira e, infelizmente, muitas vezes a única coisa que vem à mente quando se considera a escala BGP. Embora a escala de BGP suportada não possa ser representada sem mencionar o número de peers de BGP, não é o fator mais importante. Muitos outros aspectos são igualmente relevantes.

Famílias de Endereços

O tipo de família de endereços (AF) é um fator importante nas considerações de desempenho do BGP porque em implantações típicas ele afeta o tamanho de uma única rota. O número de rotas IPv4 que podem ser empacotadas em um único segmento TCP é significativamente maior que o número de rotas VPNv4. Portanto, para a mesma escala de alterações de tabela de BGP, um RR de BGP IPv4 tem menos trabalho a fazer em comparação a um RR de BGP VPNv4. Obviamente, em implantações onde um número significativo de comunidades é adicionado a cada rota, a diferença entre AFs se torna menos significativa, mas o tamanho de uma única rota é ainda maior e requer consideração.

Número De Grupos De Atualização

O processo BGP prepara uma única atualização para todos os membros do mesmo grupo de atualização. Em seguida, o processo TCP divide os dados de atualização em um número necessário de segmentos TCP (dependendo do TCP MSS) para cada membro do grupo de atualização. Você pode ver os grupos de atualização ativos e seus membros usando o show bgp update-group comando. Você pode influenciar quais e quantos peers são membros de um grupo de atualização criando uma política de saída comum para um grupo de peers que você deseja que esteja no mesmo grupo de atualização. Uma única atualização enviada pelo BGP RR para um grande número de clientes BGP RR pode disparar uma rajada de TCP ACKs que pode ser descartada no componente Local Packet Transport Service (LPTS) dos roteadores Cisco IOS XR.

Complexidade de RPLs (políticas de rota)

A complexidade das políticas de rota usadas pelo BGP impacta o desempenho do processo BGP. Cada rota recebida ou enviada deve ser avaliada em relação à política de rota configurada. Uma política muito longa exige que muitos ciclos de CPU sejam gastos nessa ação. Uma política de rota que inclui uma expressão regular é especialmente pesada no processamento. Uma expressão regular ajuda a expressar a política de rota em um número menor de linhas, mas requer mais ciclos de CPU durante o processamento do que a política de rota equivalente que não usa expressão regular.

Frequência De Atualizações

A frequência das atualizações tem um efeito importante na escala de BGP. O número de atualizações é frequentemente difícil de prever. Você pode influenciar a frequência de atualizações usando o comando "**`advertisement-interval`**", que define o intervalo mínimo entre o envio de atualizações de roteamento BGP. O valor padrão para peers iBGP é 0 segundo e 30 para peers eBGP é 30 segundos.

MSS de TCP e MTU de Interface/Caminho

Dividir uma atualização em muitos segmentos TCP pode sobrecarregar bastante os recursos do processo TCP em um ambiente de alta escala e alta frequência de atualização. Um MTU de caminho maior e um TCP MSS maior são melhores para o desempenho do BGP e do TCP.

NSR em roteadores RP duplos

O NSR é um excelente recurso para redundância, mas tem um impacto no desempenho do BGP. Nos roteadores Cisco IOS XR, ambos os RPs estão recebendo simultaneamente todas as atualizações de BGP diretamente da NPU na placa de linha de entrada, o que significa que o RP Ativo não precisa gastar tempo replicando a atualização para o RP em standby. No entanto, cada atualização gerada pelo RP Ativo deve ser enviada para o RP em Espera e daí para o correspondente BGP. Isso permite que o RP em standby esteja sempre atualizado nos números de sequência e de confirmação, mas tem um impacto no desempenho geral do BGP. É por isso que é recomendado que um BGP RR seja um roteador de RP único.

Pares lentos

Um peer lento pode retardar as atualizações em relação a todos os membros do grupo de atualização porque o processo BGP deve manter a atualização em sua memória até que todos os peers a confirmem. Se você souber que alguns pares são muito mais lentos (por exemplo, roteadores em uma parte antiga da rede), separe-os antecipadamente em um grupo de atualização. Por padrão, o Cisco IOS XR relata um peer lento através de uma mensagem de syslog. Você pode criar peers lentos estáticos (que nunca compartilham o grupo de atualização com outros) ou ajustar o comportamento dinâmico do peer lento usando o comando de configuração BGPslow-peer no modo de configuração global ou por vizinho. Uma boa leitura a mais sobre isso pode ser encontrada em Troubleshoot Slow BGP Convergence Due to Suboptimal Route Policies on IOS-XR no portal Cisco xrdocs.io.

Nexthop trigger-delay

Se vários próximos saltos do BGP mudarem em um curto intervalo de tempo e o valor crítico de retardo de gatilho de próximo salto de zero for configurado em uma família de endereços (AF) com um número alto de rotas, um passo completo do AF deve ser executado em cada evento de mudança de próximo salto. Caminhos repetidos desse AF aumentam o tempo de convergência em famílias de endereços com valores críticos mais baixos de retardo de gatilho de próximo salto. Você pode ver os valores de trigger-delay do próximo salto executando o comando "show bgp all nexthops".

Exemplo de escala de RR de BGP multidimensional validada

Os resultados da escala multidimensional, especialmente para as características do plano de controle, são altamente dependentes do ambiente de teste específico. Os resultados do teste podem variar significativamente se alguns dos parâmetros forem alterados.

Parâmetro	Valor	Valor

Platform	Dispositivo XRv9k (baseado em UCS M5)	ASR9902
versão IOS XR	7.5.2 + SMU de guarda-chuva para ID de bug Cisco CSCwf09600 . (Os componentes deste SMU de guarda-chuva estão integrados no Cisco IOS XR versão 7.9.2 e posterior)	7.11.2
Pares	VPNv4 eBGP: 2500 iBGP de VPNv4: 1700	iBGP de VPNv4: 2000
Rotas BGP	Por sessão: 200 Total: 400 mil Caminhos por rota: 1	Por sessão: 750 VPNv4: 1,36M VPNv6: 150 mil IPv4: 950 mil IPv6: 200 mil Total: ~2,6 milhões Caminhos por rota: 1
Rotas IGP	10 k (ISIS)	10 k (ISIS)
Grupos de atualização do BGP	1	1
Temporizadores BGP	padrão	padrão
LPTS BGP-known policer rate	50,000	25,000
configuração tcp num-thread	16 16	16 16

BGP send-buffer-size	padrão	padrão
Resumo dos KPIs (Key Performance Indicators, principais indicadores de desempenho)	• Caso de teste com a maior taxa de pacotes de entrada e saída: ◦ Entrada: 49,4 kpps ◦ Saída: 95 kpps ◦ ==> quedas de LPTS (vigilante a 50kpps) ◦ ==> Não há quedas em clientes NetIO ◦ ==> Tamanho máximo da fila XIPC (BGP): 1362 ◦ ==> Tamanho máximo da fila XIPC (TCP): 1248	• Caso de teste com a maior taxa de pacotes de entrada: ◦ Entrada: 16030 pacotes/s ◦ Saída: 31 pkts/s ◦ ==> Não há quedas nos clientes LPTS nem NetIO ◦ ==> Tamanho máximo da fila XIPC (BGP): 378 ◦ ==> Tamanho máximo da fila XIPC (TCP): 1021 • Caso de teste com maior taxa de pacotes de saída: ◦ Entrada: 12172 pacotes/s ◦ Saída: 23465 pacotes/s ◦ ==> Não há quedas nos clientes LPTS nem NetIO ◦ ==> Tamanho máximo da fila XIPC

		(BGP): 109 ==> Tamanho máximo da fila XIPC (TCP): 1518

Considerações do projeto

Há duas abordagens para o posicionamento do BGP RR na rede:

- Projeto de RR de BGP centralizado/plano.
- Projeto BGP RR distribuído/hierárquico.

Em um design centralizado/plano, todos os clientes BGP RR na rede estabelecem o peering BGP com um conjunto (geralmente um par) de dispositivos BGP RR que mantêm exatamente as mesmas informações. Essa abordagem é simples de implementar e funciona bem em redes de pequena a moderada escala. Qualquer alteração na tabela BGP é propagada rapidamente para todos os clientes BGP RR. À medida que o número de clientes BGP RR cresce, o design pode atingir um limite de escala quando o número de conexões TCP nos dispositivos BGP RR cresce na extensão em que seu desempenho é afetado.

Em um projeto distribuído/hierárquico, a rede é dividida em várias regiões. Todos os roteadores em uma região estabelecem peering BGP com um conjunto (geralmente um par) de dispositivos BGP RR que mantêm exatamente as mesmas informações. Esses dispositivos BGP RR atuam como clientes BGP RR para outro conjunto (geralmente um par) de dispositivos BGP RR. Essa abordagem de projeto permite uma fácil expansão da rede, mantendo o número de conexões TCP em cada BGP RR abaixo de um certo limite.
Outra consideração de design é adaptar o escopo dos destinatários de atualizações de BGP. Dependendo da distribuição de VRF entre clientes BGP RR, vale a pena considerar a distribuição de rota restrita RT. Se todos os clientes BGP RR tiverem interfaces no mesmo VRF, a Distribuição de Rota Restrita RT não traz muitos benefícios. No entanto, se os VRFs forem distribuídos de forma esparsa entre todos os clientes BGP RR, o uso da Distribuição de Rota Restrita RT reduz significativamente$f$ a carga no processo de bgp no BGP RR.

Monitorar KPIs (Key Performance Indicators, principais indicadores de desempenho) do BGP

O monitoramento dos KPIs (Key Performance Indicators, principais indicadores de desempenho) do BGP RR é importante para garantir a operação adequada da rede.

Uma alteração significativa na topologia de rede (por exemplo, um flap de enlace DWDM principal) pode disparar atualizações de roteamento que geram tráfego excessivo para e/ou do BGP RR. Tráfego significativo que atinge o BGP RR normalmente carrega:

- Atualizações de pares BGP.
- ACKs TCP gerados pelos peers BGP, em resposta a atualizações enviadas pelo BGP RR e vice-versa

Esta seção do documento explica o KPI que precisa ser monitorado em um BGP RR típico e também como saber quais dos dois tipos de tráfego BGP significativos estão causando alta taxa de tráfego de plano de controle.

O caminho dos pacotes BGP dentro do roteador pode ser descrito da seguinte maneira:

Punt
Controlador Ethernet -(pacote)-> encaminhador de caminho de dados -(pacote)-> LPTS -(pacote)-> SPP -(pacote) -> NetIO -(pacote)-> TCP -(mensagem)-> BGP
Injetar
BGP -(mensagem) -> TCP -(pacote) -> NetIO -(pacote) -> SPP -(pacote) -> encaminhador de caminho de dados -(pacote) -> controlador Ethernet

Os KPIs podem ser divididos em:

Fundamentos:

- Encaminhador de caminho de dados
- LPTS (configurações de vigilantes de punt de hardware, contadores de aceitação e contadores de queda)
- SPP
- NetIO
- Filas IPC (NetIO <==> TCP <==> BGP)
- Tamanhos BGP InQ/OutQ

Opcional:

- utilização de CPU
- Utilização de memória
- estatísticas de TCP
- desempenho do processo BGP
- convergência de BGP

Monitorar Encaminhador de Caminho de Dados

No XRv9000, o encaminhador de caminho de dados é o DPA (Data Plane Agent), enquanto nas plataformas ASR9000 ele é o NP (Network Processor).

Monitore o DPA (Data Plane Agent) XRv9000

O comando útil para ver a carga e as estatísticas do DPA é:

show controllers dpa statistics global

Esse comando mostra todos os contadores diferentes de zero, que fornecem informações sobre o tipo e o número de pacotes lançados das interfaces de rede para a CPU RP, injetados da CPU RP para as interfaces de rede, e o número de pacotes descartados:

<#root>

RP/0/RP0/CPU0:xrv9k-01#

show controllers dpa statistics global

```
 Index Debug Count ------------------------------------------------------------------------------ 350 TBP
```

Monitore o processador de rede (NP) ASR9000

Comandos úteis para ver a carga e as estatísticas de cada NP no sistema são:

show controllers np load all

show controllers np counters all

NP no ASR9000 tem um rico conjunto de contadores que mostram o número, a taxa e o tipo de pacotes processados e descartados,.

<#root>

RP/0/RSP0/CPU0:ASR9k-B#

show controllers np load all

```
 Node: 0/0/CPU0: ------------------------------------------------------------ Load Packet Rate NP0:
```

<#root>

RP/0/RSP0/CPU0:ASR9k-B#

show controllers np counters all

```
Node: 0/0/CPU0: ---------------------------------------------------------------- Show global stats cou
```

Monitorar LPTS

Como um BGP RR padrão não está no caminho de encaminhamento, todos os pacotes recebidos na interface de rede são apontados para o plano de controle. O elemento de caminho de dados em um BGP RR executa um pequeno número de operações simples antes que os pacotes sejam direcionados para o plano de controle. Como é improvável que o elemento do caminho de dados seja um ponto de congestionamento, o único elemento na placa de linha que precisa de monitoramento são as estatísticas LPTS.

Observe que, no caso de XRv9k, as estatísticas de hardware mapeiam para o vPP

Comando:

```
show lpts pifib hardware police location <location> | inc "Node|flow_type|BGP"
```

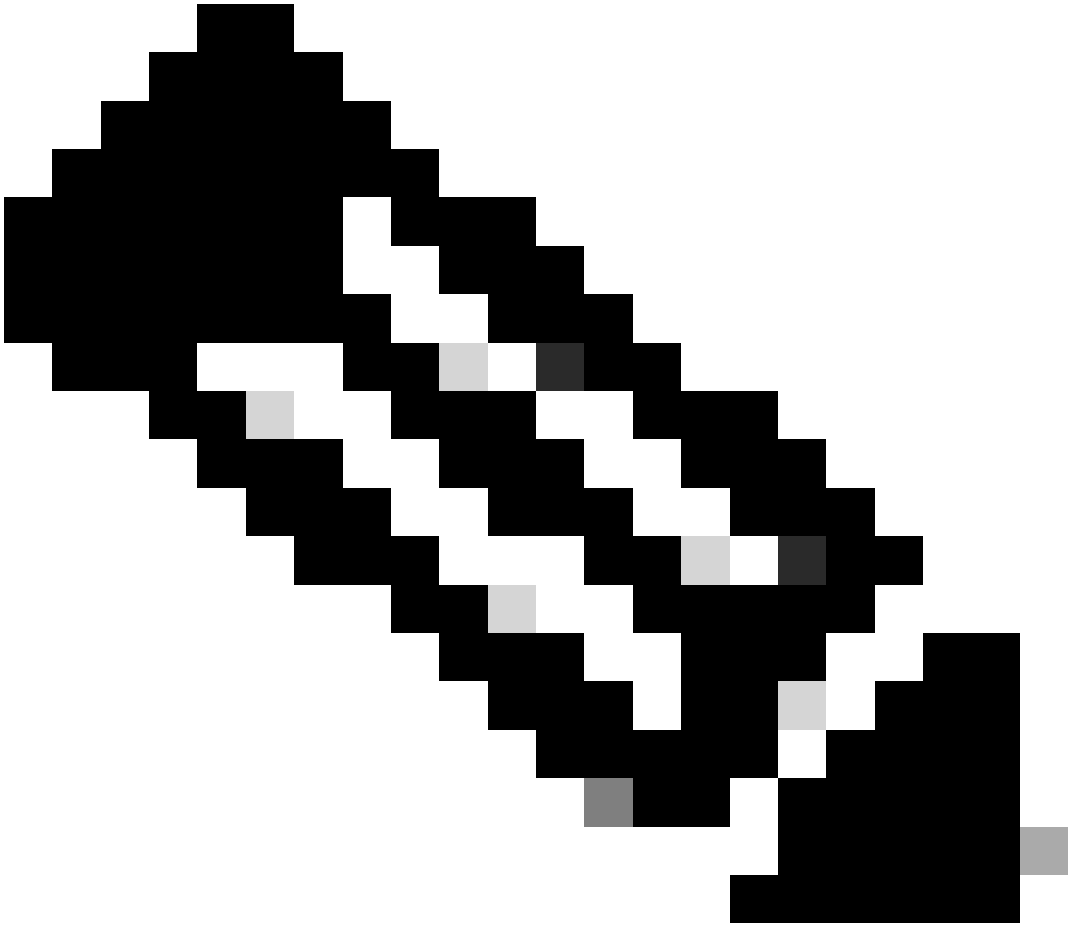
Exemplo:

```
RP/0/RP0/CPU0:xrv9k-01#sh lpts pifib hardware police location 0/0/CPU0 | i "Node|flow_type|BGP" Node 0/0/CPU0: flow_type priority sw_police_id hw
```

O que procurar:
Se for observado um salto significativo em AggDrops em relação ao tipo de fluxo conhecido por BGP, comece a procurar alterações na topologia da rede que acionaram essa rotatividade maciça no plano de controle.

Caminho de dados de telemetria:

```
Cisco-IOS-XR-lpts-pre-ifib-oper:lpts-pifib
```



Observação: os contadores de estatísticas LPTS podem ser limpos. Seu sistema de monitoramento deve levar em conta essa possibilidade.

Monitorar SPP

O SPP é a primeira entidade no processador de rotas ou na CPU da placa de linha que recebe o pacote enviado do NP ou do DPA através da estrutura interna e o último ponto no processamento do pacote de software antes de ser entregue à estrutura para injeção no NP ou no DPA.

Comandos relevantes para a monitoração do SPP:

show spp node-counters

show spp client

O **show spp node-counters** comando mostra a taxa de pacotes apontados/injetados e é fácil de ler e entender. Para sessões BGP, os contadores relevantes estão sob **client/punt** e **client/inject** no RP ativo.

O **show spp client** é mais rico em saída e fornece uma visão mais detalhada sobre o número de pacotes enfileirados/descartados em relação aos clientes, bem como a marca d'água alta.

<#root>

RP/0/RP0/CPU0:xrv9k-01#

show spp node-counters

0/RP0/CPU0:

socket/rx Punted packets: 595305 Punt bulk reads: 6 Punt non-bulk reads: 595293 Management packets: 74

client/inject Injected from client: 140534413 Non-bulk injects: 140534413 ----------

---------- 0/0/CPU0: <. . .>

<#root>

RP/0/RP0/CPU0:xrv9k-01#

show spp client

Sat Apr 20 17:11:40.725 UTC 0/RP0/CPU0: Clients ======= <. . .> netio, JID 254 (pid 4591) ----------

Monitorar NetIO

Enquanto o vigilante LPTS mostra apenas a contagem de pacotes aceitos ou descartados por um vigilante correspondente, no nível NetIO podemos ver a taxa de pacotes apontados para a CPU RP. Como em um BGP RR típico a grande maioria dos pacotes recebidos são pacotes BGP, a taxa NetIO geral indica muito bem a taxa de pacotes BGP recebidos.

<#root>

Command:

show netio rates

Exemplo:

<#root>

RP/0/RP0/CPU0:xrv9k-01#

show netio rates

```
 Netio packet rate for node 0/RP0/CPU0 ---------------------------------- Current rate (updated 0 seco
```

O que procurar:

- Se for observado um salto significativo na taxa de NetIO, comece a procurar alterações na topologia da rede que tenham disparado essa rotatividade maciça no plano de controle.

Caminho de dados de telemetria:

- não aplicável, pois a telemetria deve transmitir valores do contador, não taxas. O contador de aceitação do vigilante LPTS conhecido por BGP pode ser usado no coletor de telemetria para aproximar a taxa média de pacotes BGP recebidos de pares conhecidos.

Monitorar filas XIPC

No caminho de punt, os pacotes recebidos pelo NetIO do LPTS são passados para o TCP e o BGP. É importante monitorar estas filas:

1. Fila de alta prioridade TCP através da qual o NetIO entrega pacotes ao TCP
2. Fila de controle BGP
3. Fila de dados BGP

No caminho de injeção, os pacotes são criados pelo TCP e passados para o NetIO. É importante monitorar estas filas:

- Fila XIPC OutputL

Comandos:

show netio clients show processes bgp | i "Job Id" show xipcq jid <bgp_job_id> show xipcq jid <bgp_job_id> queue-id <n>

Examples:

NetIO para TCP, visão do ponto de vista NetIO:

RP/0/RP0/CPU0:xrv9k-01#show netio clients <. . .> Input Punt XIPC InputQ XIPC PuntQ ClientID Drop/Total Drop/Total Cur/High/Max Cur/High/Max

TCP para NetIO, visão do ponto de vista NetIO:

RP/0/RP0/CPU0:xrv9k-01#show netio clients <. . .> XIPC queues Dropped/Queued Cur/High/Max ---------- Outp

NetIO para TCP, visualização do ponto de vista do processo TCP:

<#root>

RP/0/RP0/CPU0:xrv9k-01#

show processes tcp

```
 | i "Job Id"
 Job Id: 430
RP/0/RP0/CPU0:xrv9k-01#
```

show xipcq jid

```
 430 Mon Apr 17 16:16:11.315 CEST Id Name Size Cur Size Produced Dropped HWM ------ ----------
```

TCP para BGP:

<#root>

RP/0/RP0/CPU0:xrv9k-01#

show processes bgp

```
 | i "Job Id" Job Id: 1078 RP/0/RP0/CPU0:xrv9k-01#
```

show xipcq jid

```
 1078 Mon Apr 17 16:09:33.046 CEST Id Name Size Cur Size Produced Dropped HWM ------ ----------
```

Fila de dados BGP:

<#root>

RP/0/RP0/CPU0:xrv9k-01#

show xipcq jid

```
 1078
```

queue-id 1

```
 XIPC_xipcq_12_0_9854_6506_inst_1_data_toapp
:
 Magic: 12344321 Version: 0 SHM Size: 192392 Owner PID: 9854 Owner JID: 1078 Queue ID: 1 Owner MQ handl
```

Fila de controle BGP:

<#root>

RP/0/RP0/CPU0:xrv9k-01#

show xipcq jid

 1078

queue-id

 2 XIPC_xipcq_12_0_9854_6506_inst_1_ctrl_toapp: Magic: 12344321 Version: 0 SHM Size: 480392 Owner PID:

O que procurar:

- não deve haver quedas nas filas relevantes
- em estatísticas de fila XIPC Marca d'água alta (HWM) não deve exceder 50% do tamanho da fila

Para melhor rastreamento da evolução do valor da marca d'água alta, você deve limpar o valor da marca d'água alta após cada leitura. Observe que isso não limpa apenas o contador HWM, mas também limpa todas as estatísticas da fila. O formato do comando para limpar as estatísticas da fila XIPC é: clear xipcq statistics queue-name <queue_name>

Como o nome da fila geralmente inclui o ID do processo (PID), o nome da fila é alterado após a reinicialização do processo.
Alguns exemplos de comandos para limpar as estatísticas de filas relevantes:

clear xipcq statistics queue-name XIPC_tcp_i0
clear xipcq statistics queue-name XIPC_tcp_i1
clear xipcq statistics queue-name XIPC_xipcq_12_0_9854_6506_inst_1_data_toapp
clear xipcq statistics queue-name XIPC_xipcq_12_0_9854_6506_inst_1_ctrl_toapp

Caminho da telemetria:

- Não há caminhos do sensor de telemetria para XIPC.

Monitorar filas de entrada e saída do BGP

O BGP mantém uma fila de entrada e saída para cada peer de BGP. Os dados ficam no InQ quando o TCP os passou para o BGP, mas o BGP ainda não os processou. Os dados ficam na OutQ enquanto o BGP espera no TCP para dividir os dados em pacotes e transmiti-los. O tamanho instantâneo do BGP InQ/OutQ fornece uma boa indicação de quão ocupado o processo BGP está.

Comando:

show bgp <AFI> <SAFI> summary

Exemplo:

RP/0/RP0/CPU0:xrv9k-01#show bgp all all summary Address Family: VPNv4 Unicast ---------- BGP router identifier 192.168.0.1, local A

O que procurar:

- O tamanho de InQ/OutQ deve ser zero quando a rede estiver estável. Ele muda rapidamente quando as atualizações são trocadas.
- O tamanho de InQ/OutQ não deve aumentar de forma monótona com o tempo.

Caminho da telemetria:

- Cisco-IOS-XR-ipv4-bgp-oper:bgp

Monitorar taxas de mensagens BGP

Alguns vizinhos BGP podem enviar atualizações ou retiradas continuamente se a topologia de rede estiver instável. O BGP RR deve então replicar essa alteração de tabela de roteamento milhares de vezes para todos os seus clientes RR. Portanto, é importante monitorar as taxas de mensagens recebidas de vizinhos, para rastrear fontes de instabilidade.

Comando:

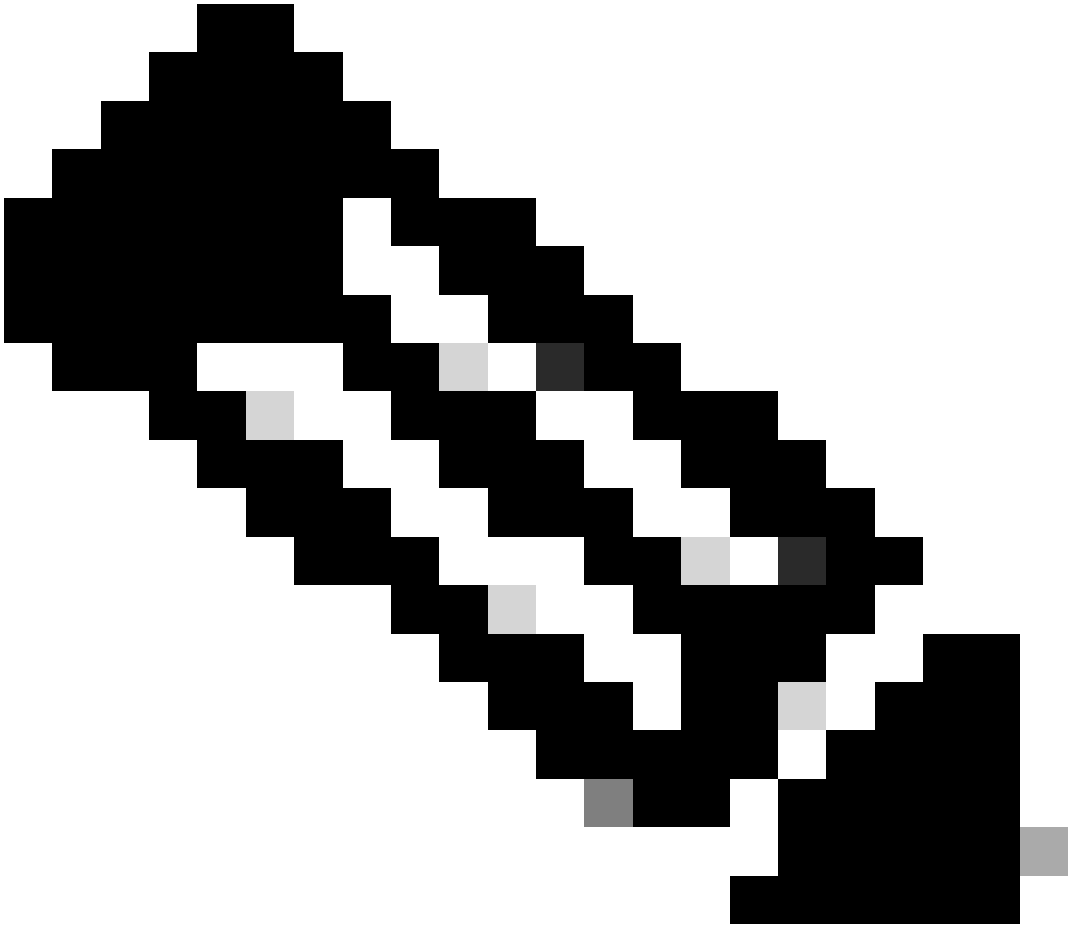
show bgp <AFI> <SAFI> summary

Exemplo:

RP/0/RP0/CPU0:xrv9k-01#show bgp all all summary Address Family: VPNv4 Unicast ---------- BGP router identifier 192.168.0.1, local A

As filas de clientes RR têm aproximadamente a mesma quantidade de MsgSent, mas alguns vizinhos podem ter um número de MsgRcvd maior que outros. Você deve capturar vários instantâneos desse comando para avaliar a taxa de mensagens.

Depois de identificar os peers ofensivos, você pode passar por outros comandos como **show bgp neighbor <neighbor> detail** e **show bgp neighbor <neighbor> performance-statistics** ou **show bgp recent-prefixe**s para tentar entender quais prefixos estão oscilando e se são sempre os mesmos ou diferentes.



Observação: os contadores MsgRcvd e MsgSent são por vizinho, mas não por família de endereços. Assim, ao executar um comando como show bgp all all summary , você vê os mesmos contadores por vizinho nas seções para as várias famílias de endereços. Eles não representam o número de mensagens recebidas/enviadas de/para esse vizinho para essa família de endereços, mas entre famílias de endereços.

Monitorar a utilização da CPU

A utilização da CPU deve ser monitorada em todos os roteadores, mas em um roteador com um alto número de núcleos de CPU dedicados ao plano de controle, algumas leituras podem não ser intuitivas. Em um BGP RR com um alto número de núcleos de CPU dedicados ao processador de roteamento (RP), como no caso do dispositivo XRv9k, threads ativos são executados em diferentes núcleos de CPU, enquanto um número de núcleos de CPU permanece ocioso. Como consequência, alguns núcleos de CPU podem estar muito ocupados, mas a utilização geral de CPU calculada em todos os núcleos de CPU permanece moderada.

Portanto, para o monitoramento adequado da utilização dos núcleos da CPU via CLI, use o **show processes cpu thread** comando.

Monitorar estatísticas de TCP

O Cisco IOS® mantém estatísticas detalhadas sobre cada sessão TCP. O comando CLI **show tcp brief** exibe a lista de todas as sessões TCP existentes. Nesta saída de resumo, para cada sessão TCP você pode ver estas informações:

- **PCB:** identificador exclusivo de sessão TCP.
- **VRF-ID**: o ID do VRF no qual a sessão existe.
  - Para ver o nome VRF correspondente, execute este comando:
  - show cef vrf all summary | utility egrep "^VRF:|Vrfid" | utility egrep -B1 <VRF-ID>
- **Recv-Q**: tamanho instantâneo da fila de recebimento. A fila de recebimento contém os pacotes recebidos do NetIO. O processo **tcp** extrai os dados de um pacote e os envia ao aplicativo correspondente.
- **Send-Q**: tamanho instantâneo da fila de envio. A fila de envio mantém os dados recebidos de um aplicativo. O processo **tcp** divide os dados em segmentos TCP (ditados pelo tamanho máximo negociado do segmento - TCP MSS), encapsula cada segmento em um cabeçalho da camada 3 da família de endereços correspondente (IPv4 ou IPv6) e envia o pacote para NetIO.
- **Endereço local**: endereço IPv4 ou IPv6 local associado ao soquete TCP. As sessões TCP no estado LISTEN são normalmente vinculadas a "**qualquer**" endereço IP, que é representado como "0.0.0.0" ou "::" no caso de IPv4 ou IPv6, respectivamente.
- **Endereço Externo**: endereço IPv4 ou IPv6 remoto associado ao soquete TCP. As sessões TCP no estado LISTEN são normalmente vinculadas a "**qualquer**" endereço IP, que é representado como "0.0.0.0" ou "::" no caso de IPv4 ou IPv6, respectivamente.
- **Estado**: estado da sessão TCP. Os possíveis estados de sessão TCP são: LISTEN, SYNSENT, SYNRCVD, ESTAB, LASTACK, CLOSING, CLOSEWAIT, FINWAIT1, FINWAIT2, TIMEWAIT, CLOSED.

Como o número de porta BGP bem conhecido é 179, você pode limitar as sessões TCP exibidas àquelas que estão associadas à aplicação BGP.

Exemplo:

RP/0/RSP0/CPU0:ASR9k-B#show tcp brief | include "PCB|:179 " PCB VRF-ID Recv-Q Send-Q Local Address Foreign Address State 0x00007ff7d403bd

Você pode usar o valor PCB exibido para obter as estatísticas de uma sessão TCP específica. Comandos CLI que fornecem informações sobre as estatísticas do processo TCP:

Global:

show tcp statistics clients location <active_RP>

show tcp statistics summary location <active_RP>

Por PCB:

show tcp brief | i ":179"

show tcp detail pcb <pcb> location 0/RP0/CPU0

show tcp statistics pcb <pcb> location <active_RP>

Os comandos de estatísticas TCP globais mostram a integridade geral das sessões TCP. Além das estatísticas de pacote de dados (entrada/saída), você pode ver, por exemplo, se há pacotes com erros de checksum, pacotes malformados, pacotes descartados devido a erros de autenticação, pacotes fora de ordem, pacotes com dados após a janela, o que dá uma indicação do comportamento dos pares TCP.

Nos comandos per-PCB, você pode ver parâmetros importantes de uma sessão TCP, como MSS, tempo máximo de ida e volta, etc.

Os contadores relevantes na saída do show tcp detail pcb comando são:

- **Temporizador de Retransmissão Iniciado**: indica quantas vezes o temporizador de retransmissão foi iniciado.
- **Ativações do temporizador de retransmissão**: indica quantas vezes o temporizador de retransmissão se esgotou, disparando uma retransmissão do segmento TCP.
- **Tamanho atual da fila de envio em bytes**: bytes não confirmados do par.
- **Tamanho da fila de recebimento atual em bytes/pacotes**: bytes/pacotes ainda a serem lidos pelo aplicativo (BGP).
- **bytes mal ordenados**: bytes que são enfileirados na fila de salvamento devido a um buraco na janela de recebimento TCP.

<#root>

RP/0/RSP0/CPU0:ASR9k-B#

show tcp detail pcb 0x4a4400e4

========== Connection state is ESTAB, I/O status:

Current send queue size in bytes: 0 (max 16384)

Current receive queue size in bytes: 0 (max 65535)

mis-ordered: 0 bytes

Current receive queue size in packets: 0 (max 60)

Timer Starts Wakeups Next(msec)

Retrans 2795 0 0

SendWnd 1341 0 0 TimeWait 0 0 0 AckHold 274 2 0 KeepAlive 333 1 299983 PmtuAger 0 0 0 GiveUp 0 0 0 Thro
SRTT: 162 ms, RTTO: 415 ms, RTV: 253 ms, KRTT: 0 ms
minRTT: 0 ms, maxRTT: 247 ms ACK hold time: 200 ms, Keepalive time: 300 sec, SYN waittime: 30 sec Giveu

Monitorar Utilização de Memória

A tabela de rotas BGP é armazenada na memória heap do processo BGP. A tabela de roteamento é armazenada na memória heap do processo RIB.

Comandos úteis para o monitoramento da memória de heap:

show memory summary

show memory summary detail

show memory-top-consumers

show memory heap summary all

Caminho do sensor de telemetria:

Cisco-IOS-XR-nto-misc-oper:memory-summary/nodes/node/detail

A FIB armazena entradas de encaminhamento no espaço de memória compartilhada.

Comandos úteis para o monitoramento de memória compartilhada:

show memory summary

show memory summary detail

show shmwin summary

Monitorar o desempenho do processo BGP

Comando útil que fornece dados internos sobre o desempenho do processo BGP:

show bgp process performance-statistics

show bgp process performance-statistics detail

Monitorar a convergência do BGP

Outro comando útil é aquele que mostra o status geral da convergência de BGP: show bgp convergence

Quando a rede estiver estável, você verá algo como isto:

RP/0/RP0/CPU0:ASR9k-B#show bgp convergence Mon Dec 18 13:55:47.976 UTC Converged. All received routes in RIB, all neighbors updated. All neig

Sobre esta tradução
A Cisco traduziu este documento com a ajuda de tecnologias de tradução automática e humana para oferecer conteúdo de suporte aos seus usuários no seu próprio idioma, independentemente da localização.

Observe que mesmo a melhor tradução automática não será tão precisa quanto as realizadas por um tradutor profissional.

A Cisco Systems, Inc. não se responsabiliza pela precisão destas traduções e recomenda que o documento original em inglês (link fornecido) seja sempre consultado.