Considérations sur l'échelle BGP RR et surveillance KPI

Table des matières

Introduction

Ce document décrit les principaux contributeurs à l'échelle maximale qu'un réflecteur de route (RR) Border Gateway Protocol (BGP) peut atteindre et des conseils sur la surveillance des performances BGP RR.

Sélection de la plate-forme matérielle/logicielle

Un RR BGP à grande échelle n'est généralement pas dans le chemin de transmission des paquets transportant des services fournis par un fournisseur d'accès Internet. Par conséquent, les exigences matérielles pour un RR BGP et les routeurs qui transfèrent principalement des paquets dans le chemin de données sont différentes. Les routeurs standard sont construits avec un élément de transmission de chemin de données puissant et un élément de chemin de contrôle relativement modéré. Un RR BGP effectue toutes ses tâches dans un plan de contrôle.
Dans la gamme de produits Cisco IOS® XR, vous pouvez choisir entre 3 types de plates-formes matérielles/logicielles pour un rôle BGP RR :

Routeur physique Cisco IOS XR	Appliance Cisco IOS XRv 9000	Routeur Cisco IOS XRv 9000 (alias XRv9k)
• Capacité modérée du plan de contrôle (généralement entre 2 et 6 coeurs de CPU alloués à la machine virtuelle RP XR) • Capacité de chemin de données inutilisée	• Capacité élevée du plan de contrôle (sur l'appliance Cisco UCS M5, 36 coeurs de CPU sont dédiés à la machine virtuelle RP XR) • Partage égal entre la capacité du chemin de données et celle du chemin de contrôle. • L'image XRv9k s'exécute sur le réseau sans système d'exploitation pour des performances maximales	• Capacité du plan de contrôle personnalisable • Partage égal entre la puissance du chemin de données et du chemin de contrôle lors de l'utilisation d'une image BGP RR. • Une couche supplémentaire de virtualisation a un impact sur les performances.

Au moment de la rédaction de ce document, l'appliance XRv9k est le choix de plate-forme optimal pour le RR BGP, car il fournit la capacité de plan de contrôle la plus élevée avec des performances maximales.

Considérations D'Évolutivité Et De Performances

L'échelle prise en charge des entités de plan de données est relativement facile à exprimer, car les performances de l'élément de chemin de données dépendent rarement de l'échelle. Par exemple, une recherche TCAM prend le même temps quel que soit le nombre d'entrées TCAM actives.
L'échelle prise en charge des entités de plan de contrôle est souvent beaucoup plus complexe car l'échelle et les performances sont interconnectées. Considérez un RR BGP avec des routes de 1 M. Le travail qu'un processus BGP doit effectuer pour maintenir cette table BGP dépend de :

1. Combien d'homologues BGP sont actifs ?
2. Quelles sont les familles d'adresses actives ?
3. Comment sont-ils répartis en groupes de mise à jour ?
4. La complexité des RPL (Route Policies)
5. Fréquence des mises à jour (mises à jour entrantes et mises à jour sortantes - intervalle d'annonce).
6. TCP MSS, MTU interface/chemin : le réglage de ce paramètre améliore les performances
7. En cas de RP double, NSR est-il activé ?
8. Tous les homologues lents connus, qui ne sont pas dans un groupe de mise à jour distinct
9. Valeur du délai de déclenchement du tronçon suivant

Nombre d'homologues BGP

Le nombre d'homologues BGP est généralement le premier et malheureusement, souvent la seule chose qui vient à l'esprit quand on considère l'échelle BGP. Bien que l'échelle BGP prise en charge ne puisse pas être représentée sans mentionner le nombre d'homologues BGP, ce n'est pas le facteur le plus important. De nombreux autres aspects sont également pertinents.

Familles d'adresses

Le type de famille d'adresses (AF) est un facteur important dans les considérations de performances BGP car dans les déploiements typiques, il a un impact sur la taille d'une route unique. Le nombre de routes IPv4 pouvant être regroupées dans un seul segment TCP est nettement supérieur au nombre de routes VPNv4. Par conséquent, pour la même échelle de changements de table BGP, un RR BGP IPv4 a moins de travail à faire comparé à un RR BGP VPNv4. Évidemment, dans les déploiements où un nombre important de communautés est ajouté à chaque route, la différence entre les AF devient moins importante, mais la taille d'une seule route est alors encore plus grande et nécessite une attention particulière.

Nombre De Groupes De Mise À Jour

Le processus BGP prépare une seule mise à jour pour tous les membres du même groupe de mise à jour. Le processus TCP fractionne ensuite les données de mise à jour en un nombre requis de segments TCP (en fonction de TCP MSS) vers chaque membre du groupe de mise à jour. Vous pouvez afficher les groupes de mise à jour actifs et leurs membres à l'aide de la commande`show bgp update-group` . Vous pouvez déterminer quels homologues sont membres d'un groupe de mise à jour et combien d'entre eux sont membres en créant une stratégie sortante commune pour un groupe d'homologues dont vous souhaitez faire partie du même groupe de mise à jour. Une seule mise à jour envoyée par le RR BGP à un grand nombre de clients RR BGP peut déclencher une rafale d'ACK TCP qui peuvent être abandonnés dans le composant LPTS (Local Packet Transport Service) des routeurs Cisco IOS XR.

Complexité des RPL (politiques de routage)

La complexité des politiques de routage utilisées par BGP a un impact sur les performances du processus BGP. Chaque route reçue ou envoyée doit être évaluée par rapport à la stratégie de route configurée. Une stratégie très longue nécessite de nombreux cycles de CPU à utiliser pour cette action. Une politique de routage qui inclut une expression régulière est particulièrement lourde à traiter. Une expression régulière vous aide à exprimer la stratégie de routage dans un nombre de lignes inférieur, mais nécessite plus de cycles CPU lors du traitement que la stratégie de routage équivalente qui n'utilise pas d'expression régulière.

Fréquence Des Mises À Jour

La fréquence des mises à jour a un impact important sur l'échelle BGP. Le nombre de mises à jour est souvent difficile à prévoir. Vous pouvez influencer la fréquence des mises à jour en utilisant la commande « **`advertisement-interval`** », qui définit l'intervalle minimum entre l'envoi des mises à jour de routage BGP. La valeur par défaut vers les homologues iBGP est 0 seconde et 30 vers les homologues eBGP est 30 secondes.

TCP MSS et MTU interface/chemin

Le fractionnement d'une mise à jour en de nombreux segments TCP peut exercer une forte pression sur les ressources du processus TCP dans un environnement à fréquence de mise à jour élevée et à grande échelle. Un MTU de chemin plus grand et un MSS TCP plus grand sont meilleurs pour les performances BGP et TCP.

NSR sur les routeurs à double RP

NSR est une fonctionnalité très utile pour la redondance, mais elle a un impact sur les performances BGP. Sur les routeurs Cisco IOS XR, les deux RP reçoivent simultanément chaque mise à jour BGP directement du NPU sur la carte de ligne d'entrée, ce qui signifie que le RP actif n'a pas à passer du temps à répliquer la mise à jour sur le RP de secours. Cependant, chaque mise à jour générée par le RP actif doit être envoyée au RP de secours et de là à l'homologue BGP. Cela permet au RP de secours d'être toujours à jour sur la séquence et les numéros d'accusé de réception, mais a un impact sur les performances BGP globales. C'est pourquoi il est recommandé qu'un RR BGP soit un routeur RP simple.

Homologues lents

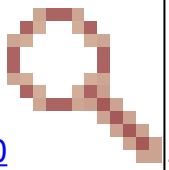
Un homologue lent peut ralentir les mises à jour vers tous les membres du groupe de mise à jour parce que le processus BGP doit garder la mise à jour dans sa mémoire jusqu'à ce que tous les homologues l'aient reconnue. Si vous savez que certains homologues sont beaucoup plus lents (par exemple, les routeurs d'une partie héritée du réseau), séparez-les d'emblée en un groupe de mise à jour. Par défaut, Cisco IOS XR signale un homologue lent via un message syslog. Vous pouvez créer des homologues lents statiques (qui ne partagent jamais le groupe de mise à jour avec d'autres) ou affiner le comportement dynamique des homologues lents à l'aide de la commande de configuration BGPslow-peer en mode de configuration globale ou par voisin. Pour plus d'informations à ce sujet, consultez Troubleshoot Slow BGP Convergence Due to Suboptimal Route Policies on IOS-XR sur le portail Cisco xrdocs.io.

Délai de déclenchement Nexthop

Si plusieurs sauts suivants BGP changent dans un court intervalle de temps et que la valeur critique de délai de déclenchement de saut suivant de zéro est configurée dans une famille d'adresses (AF) avec un nombre élevé de routes, une marche complète de l'AF doit être exécutée sur chaque événement de changement de saut suivant. Les marches répétées de cet AF augmentent le temps de convergence dans les familles d'adresses avec des valeurs de délai de déclenchement de tronçon suivant critiques inférieures. Vous pouvez voir les valeurs de délai de déclenchement du tronçon suivant en exécutant la commande « show bgp all next-thops ».

Exemple d'échelle RR BGP multidimensionnelle validée

Les résultats d'échelle multidimensionnelle, en particulier pour les fonctions du plan de contrôle, dépendent fortement de l'environnement d'essai spécifique. Les résultats des tests peuvent varier considérablement si certains paramètres sont modifiés.

Paramètre	Valeur	Valeur
Plateforme	Appliance XRv9k (basé sur UCS M5)	ASR9902
Version d'IOS XR	7.5.2 + SMU de parapluie pour l'ID de bogue Cisco  CSCwf09600 . (les composants de cette SMU parapluie sont intégrés dans Cisco IOS XR version 7.9.2 et ultérieure)	7.11.2
Pairs	VPNv4 eBGP : 2500 VPNv4 iBGP : 1 700	VPNv4 iBGP : 2000
Routes BGP	Par session : 200 Total : 400 k Chemins par route : 1	Par session : 750 VPNv4 : 1,36 M VPNv6 : 150 k IPv4 : 950 K IPv6 : 200 K Total : ~2,6 millions Chemins par route : 1
Routes IGP	10 000 (ISIS)	10 000 (ISIS)
Groupes de mise à jour BGP	1	1
Minuteurs BGP	manquer à ses obligations	manquer à ses obligations

Débit de contrôle connu BGP LPTS	50,000	25,000
tcp num-thread configuration	16 16	16 16
BGP send-buffer-size	manquer à ses obligations	manquer à ses obligations
Résumé des indicateurs de performance clés (KPI)	• Cas de test avec débit de paquets d'entrée et de sortie le plus élevé : ◦ Entrée : 49,4 kpps ◦ Sortie : 95 kpps ◦ ==> LPTS drops (régulateur à 50 kpps) ◦ ==> Aucune perte dans les clients NetIO ◦ ==> Taille maximale de la file d'attente XIPC (BGP) : 1 362 ◦ ==> Taille maximale de la file d'attente XIPC (TCP) : 1 248	• Cas de test avec le débit de paquets en entrée le plus élevé : ◦ Entrée : 16030 pkts/s ◦ Sortie : 31 pkts/s ◦ ==> Aucune perte dans les clients LPTS ou NetIO ◦ ==> Taille maximale de la file d'attente XIPC (BGP) : 378 ◦ ==> Taille maximale de la file d'attente XIPC (TCP) : 1021 • Cas de test avec le débit de paquets de sortie le plus élevé : ◦ Entrée : 12172 pkts/s ◦ Sortie : 23465 pkts/s ◦ ==> Aucune perte dans les clients LPTS ou

		NetIO ◦ ==> Taille maximale de la file d'attente XIPC (BGP) : 109 ◦ ==> Taille maximale de la file d'attente XIPC (TCP) : 1 518

Considérations conceptuelles

Il existe deux approches pour le placement de RR BGP dans le réseau :

- Conception RR BGP centralisée/plate.
- Conception RR BGP distribuée/hiérarchique.

Dans une conception centralisée/plate, tous les clients BGP RR dans le réseau établissent l'appairage BGP avec un ensemble (généralement une paire) de périphériques BGP RR qui contiennent exactement les mêmes informations. Cette approche est simple à mettre en oeuvre et fonctionne bien dans les réseaux de petite à moyenne échelle. Toute modification de la table BGP est propagée rapidement à tous les clients BGP RR. À mesure que le nombre de clients BGP RR augmente, la conception peut atteindre une limite d'échelle lorsque le nombre de connexions TCP sur les périphériques BGP RR augmente au point où leurs performances sont affectées.

Dans une conception distribuée/hiérarchique, le réseau est divisé en plusieurs régions. Tous les routeurs d'une région établissent l'appairage BGP avec un ensemble (généralement une paire) de périphériques RR BGP qui contiennent exactement les mêmes informations. Ces périphériques RR BGP agissent comme des clients RR BGP vers un autre ensemble (généralement une paire) de périphériques RR BGP. Cette approche de conception permet une extension facile du réseau, tout en maintenant le nombre de connexions TCP sur chaque RR BGP sous une certaine limite.
Une autre considération de conception est la personnalisation de l'étendue des destinataires des mises à jour BGP. En fonction de la distribution VRF parmi les clients BGP RR, il est intéressant de considérer la distribution de route contrainte RT. Si tous les clients RR BGP ont des interfaces dans le même VRF, la distribution de route contrainte RT n'apporte pas beaucoup d'avantages. Cependant, si les VRF sont distribués de manière éparse entre tous les clients RR BGP, l'utilisation de la distribution de route contrainte RT réduit de manière significative*f* la charge sur le processus bgp sur le RR BGP.

Contrôler les indicateurs de performance clés (KPI) BGP

La surveillance des indicateurs de performance clés (KPI) du RR BGP est importante pour garantir le bon fonctionnement du réseau.

Une modification importante de la topologie du réseau (par exemple, un battement de liaison DWDM majeur) peut déclencher des mises à jour de routage qui génèrent un trafic excessif vers et/ou depuis le RR BGP. Le trafic important qui atteint le RR BGP transporte généralement :

- Mises à jour des homologues BGP.
- ACK TCP générés par les homologues BGP, en réponse aux mises à jour envoyées par BGP RR et vice-versa

Cette section du document explique les KPI qui doivent être surveillés sur un RR BGP typique et également comment dire lequel des deux types de trafic BGP significatifs est à l'origine d'un taux de trafic de plan de contrôle élevé.

Le chemin des paquets BGP à l'intérieur du routeur peut être représenté comme suit :

Pointer
Contrôleur Ethernet -(paquet)-> redirecteur de chemin de données -(paquet)-> LPTS -(paquet)-> SPP -(paquet) -> NetIO -(paquet)-> TCP -(message)-> BGP
Injecter
BGP -(message)-> TCP -(paquet)-> NetIO -(paquet)-> SPP -(paquet) -> redirecteur de chemin de données -(paquet)-> contrôleur Ethernet

Les IPC peuvent être divisés en :

Essentials :

- DataPath Forwarder
- LPTS (paramètres des régulateurs de point matériel, acceptez les compteurs et supprimez les compteurs)
- SPP
- NetIO
- Files d'attente IPC (NetIO <==> TCP <==> BGP)
- Tailles BGP InQ/OutQ

Facultatif:

- Utilisation du processeur
- Utilisation de la mémoire

- Statistiques TCP
- Performances du processus BGP
- Convergence BGP

Surveiller Datapath Forwarder

Sur XRv9000, le redirecteur de chemin de données est l'agent de plan de données (DPA), tandis que sur les plates-formes ASR9000, il s'agit du processeur réseau (NP).

Surveillance de l'agent du plan de données XRv9000 (DPA)

Commande utile pour voir la charge et les statistiques du DPA est :

show controllers dpa statistics global

Cette commande affiche tous les compteurs non nuls, qui vous donnent un aperçu du type et du nombre de paquets envoyés des interfaces réseau au processeur RP, injectés du processeur RP vers les interfaces réseau, et le nombre de paquets abandonnés :

<#root>

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

show controllers dpa statistics global

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

Surveillance du processeur réseau ASR9000 (NP)

Les commandes utiles pour afficher la charge et les statistiques de chaque processeur réseau du système sont les suivantes :

show controllers np load all

show controllers np counters all

NP sur ASR9000 dispose d'un ensemble complet de compteurs qui vous indiquent le nombre, le taux et le type de paquets traités et abandonnés.

<#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
```

Moniteur LPTS

Comme un RR BGP standard n'est pas dans le chemin de transmission, tous les paquets reçus sur l'interface réseau sont dirigés vers le plan de contrôle. L'élément de chemin de données sur un RR BGP effectue un petit nombre d'opérations simples avant que les paquets ne soient dirigés vers le plan de contrôle. Étant donné que l'élément de chemin de données ne risque pas d'être un point de congestion, le seul élément de la carte de ligne qui nécessite une surveillance est l'état LPTS.

Veuillez noter que dans le cas de XRv9k, les statistiques matérielles correspondent à la vPP

commande :

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

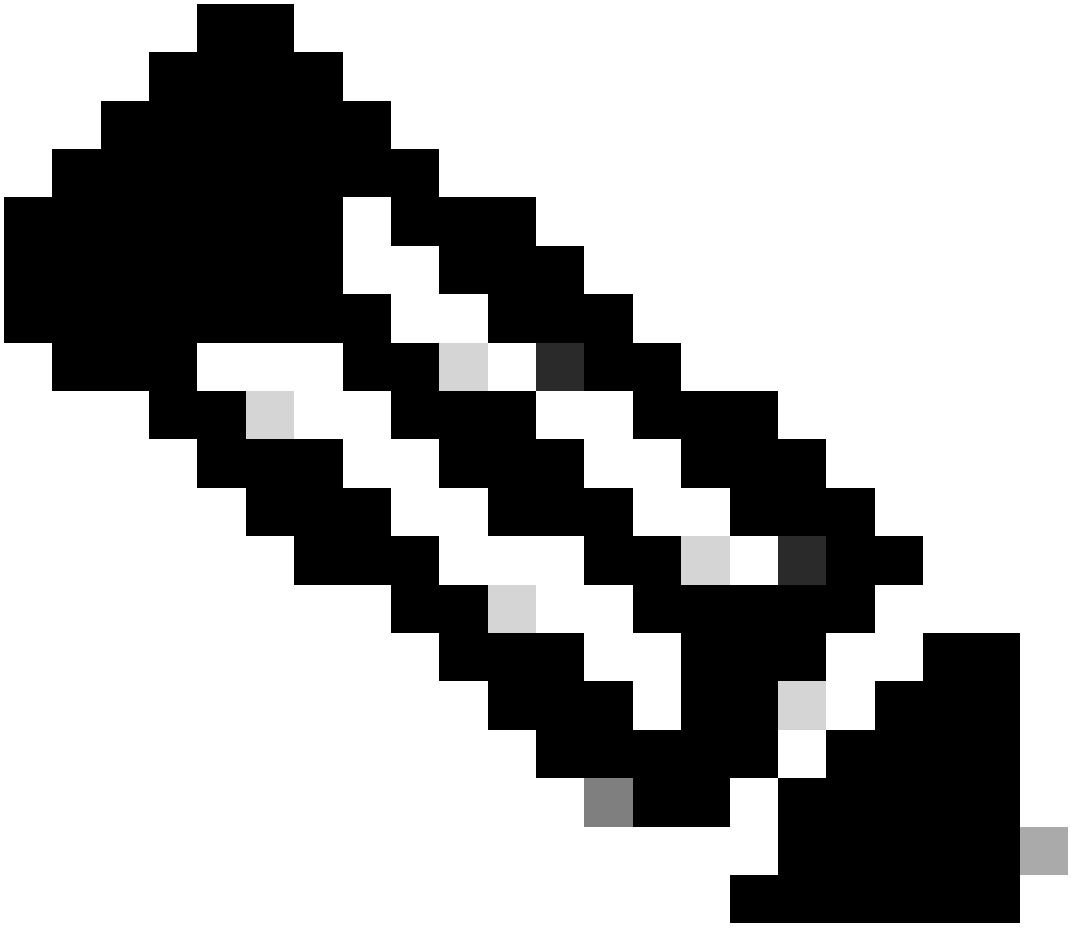
Exemple :

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

Ce qu'il faut rechercher :
Si un saut significatif dans AggDrops par rapport au type de flux connu par BGP est observé, commencez à rechercher les changements de topologie du réseau qui ont déclenché un tel désordre massif du plan de contrôle.

Chemin des données de télémétrie :

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



Remarque : les compteurs de statistiques LPTS peuvent être effacés. Votre système de surveillance doit tenir compte de cette possibilité.

Surveiller SPP

SPP est la première entité du processeur de routage ou du processeur de la carte de ligne qui reçoit le paquet envoyé par le NP ou le DPA via le fabric interne, et le dernier point du traitement du paquet logiciel avant qu'il ne soit transmis au fabric pour injection dans le NP ou le DPA.

Commandes pertinentes pour la surveillance SPP :

show spp node-counters

show spp client

La **show spp node-counters** commande affiche le taux de paquets injectés/perforés et est facile à lire et à comprendre. Pour les sessions BGP, les compteurs appropriés sont sous **client/pun**t et **client/inject** sur le RP actif.

Le **show spp client** est plus riche en résultats et donne une vue plus détaillée du nombre de paquets mis en file d'attente/abandonnés vers les clients, ainsi que du filigrane haut.

<#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) ----------

Surveillance de NetIO

Alors que le régulateur LPTS affiche uniquement le nombre de paquets acceptés ou abandonnés par un régulateur correspondant, au niveau de NetIO, nous pouvons voir le taux de paquets envoyés au CPU RP. Puisque sur un RR BGP typique, la grande majorité des paquets reçus sont des paquets BGP, le taux global de NetIO indique de très près le taux de paquets BGP reçus.

<#root>

Command:

show netio rates

Exemple :

<#root>

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

show netio rates

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

Que faut-il rechercher :

- Si une augmentation significative du taux d'E/S réseau est observée, commencez à rechercher les modifications de la topologie du réseau qui ont déclenché un tel désordre massif du plan de contrôle.

Chemin des données de télémétrie :

- non applicable, car la télémétrie doit diffuser des valeurs de compteur, et non des débits. Le compteur d'acceptation du régulateur LPTS connu par BGP peut être utilisé sur le collecteur de télémétrie pour estimer le taux moyen de paquets BGP reçus d'homologues connus.

Contrôler les files d'attente XIPC

Sur le chemin ponctuel, les paquets reçus par NetIO à partir de LPTS sont transmis à TCP et BGP. Il est important de surveiller ces files d'attente :

1. File d’attente TCP de priorité élevée par laquelle NetIO livre des paquets au TCP
2. File d'attente de contrôle BGP
3. File d'attente de données BGP

Sur le chemin d'injection, les paquets sont créés par TCP et transmis à NetIO. Il est important de surveiller ces files d'attente :

- File d'attente XIPC OutputL

Commandes :

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

Exemples:

NetIO vers TCP, vue du point de vue de 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 vers NetIO, vue du point de vue de NetIO :

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

NetIO vers TCP, vue du point de vue du processus 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 vers 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 ------ ----------
```

File d'attente de données 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
```

File de contrôle 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:
```

Que faut-il rechercher :

- il ne doit pas y avoir de pertes dans les files d'attente correspondantes
- dans les statistiques de file d'attente XIPC La limite supérieure (HWM) ne doit pas dépasser 50 % de la taille de la file d'attente

Pour un meilleur suivi de l'évolution de la valeur de filigrane élevée, vous devez effacer la valeur de filigrane élevée après chaque lecture. Notez que cela n'efface pas seulement le compteur HWM, mais également toutes les statistiques de file d'attente. Le format de la commande pour effacer les statistiques de file d'attente XIPC est : clear xipcq statistics queue-name <queue_name>

Comme le nom de la file d'attente inclut souvent l'ID de processus (PID), le nom de la file d'attente change après le redémarrage du processus. Voici quelques exemples de commandes permettant d'effacer les statistiques de files d'attente correspondantes :

```
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
```

Chemin de télémétrie :

- Il n'existe aucun chemin de capteur de télémétrie pour XIPC.

Surveillance des files d'attente BGP en entrée et en sortie

BGP gère une file d'attente d'entrée et de sortie pour chaque homologue BGP. Les données se trouvent dans InQ quand TCP les a passées à BGP, mais BGP ne les a pas encore traitées. Les données se trouvent dans OutQ tandis que BGP attend que TCP les divise en paquets et les transmette. La taille instantanée de BGP InQ/OutQ fournit une bonne indication du niveau d'occupation du processus BGP.

commande :

show bgp <AFI> <SAFI> summary

Exemple :

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

Ce qu'il faut rechercher :

- La taille de InQ/OutQ doit être égale à zéro lorsque le réseau est stable. Il change rapidement lorsque des mises à jour sont échangées.
- La taille d'InQ/OutQ ne doit pas augmenter de façon monotone au fil du temps.

Chemin de télémétrie :

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

Surveiller les taux de messages BGP

Certains voisins BGP peuvent envoyer continuellement des mises à jour ou des retraits si la topologie du réseau est instable. Le RR BGP doit ensuite répliquer ce changement de table de routage des milliers de fois sur tous ses clients RR. Par conséquent, il est important de surveiller les débits de messages reçus des voisins, afin de suivre les sources d'instabilités.

commande :

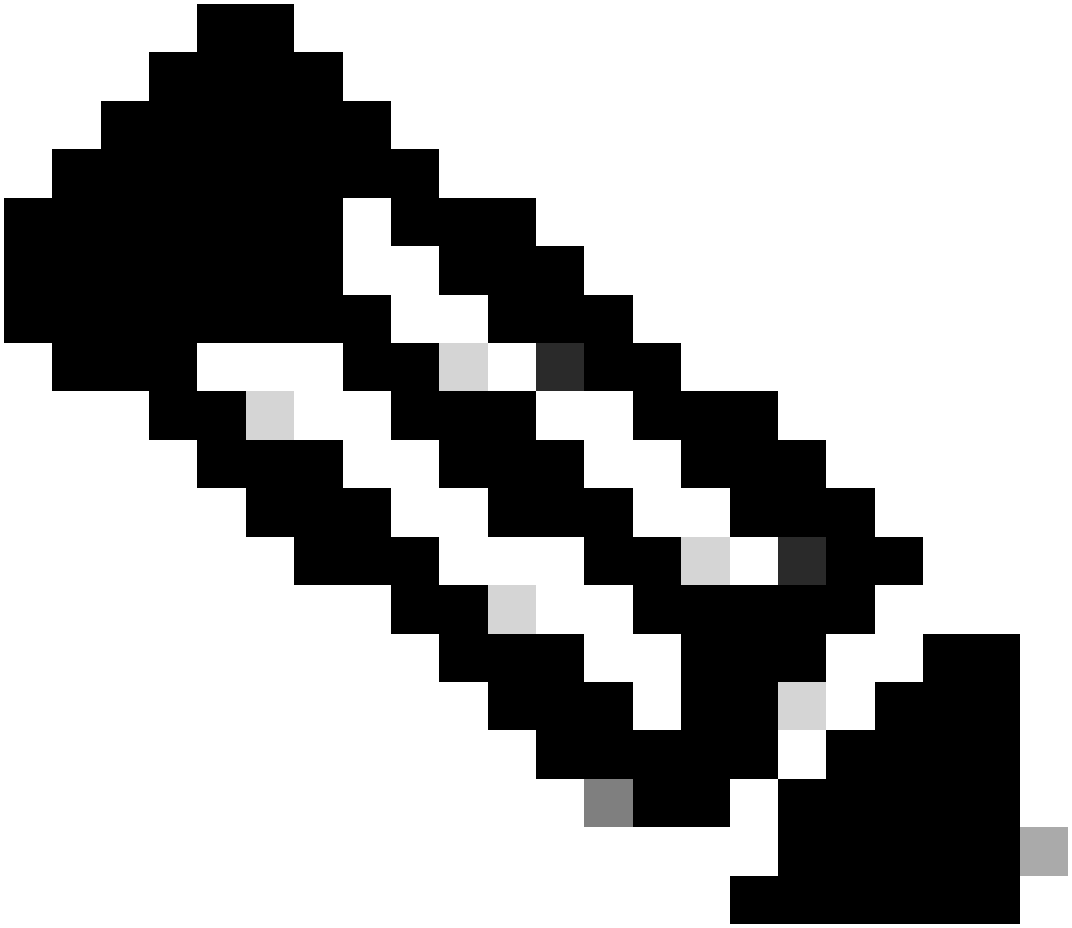
show bgp <AFI> <SAFI> summary

Exemple :

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

Les files d'attente des clients RR ont à peu près la même quantité de MsgSent, mais certains voisins peuvent avoir un nombre de MsgRcvd plus élevé que d'autres. Vous devez capturer plusieurs instantanés de cette commande pour évaluer le taux de messages.

Une fois que vous avez identifié les homologues incriminés, vous pouvez passer en revue d'autres commandes telles que **show bgp neighbor <neighbor> detail** et **show bgp neighbor <neighbor> performance-statistics** ou **show bgp recent-prefixe**s pour essayer de comprendre quels préfixes sont instables et s'il s'agit toujours des mêmes préfixes ou de préfixes différents.



Remarque : les compteurs MsgRcvd et MsgSent sont par voisin mais pas par famille d'adresses. Ainsi, lorsque vous exécutez une commande comme show bgp all all summary , vous voyez les mêmes compteurs par voisin dans les sections pour les différentes familles d'adresses. Ils ne représentent pas le nombre de messages reçus/envoyés de/vers ce voisin pour cette famille d'adresses, mais entre les familles d'adresses.

Surveiller l'utilisation du processeur

L'utilisation du CPU doit être surveillée sur chaque routeur, mais sur un routeur avec un nombre élevé de coeurs de CPU dédiés au plan de contrôle, certaines lectures peuvent être peu intuitives. Sur un RR BGP avec un nombre élevé de coeurs de CPU dédiés au processeur de routage (RP), comme dans le cas de l'appliance XRv9k, les threads actifs s'exécutent sur différents coeurs de CPU, tandis qu'un certain nombre de coeurs de CPU restent inactifs. Par conséquent, certains coeurs de processeur peuvent être très occupés, mais l'utilisation globale du processeur calculée sur tous les coeurs de processeur reste modérée.

Par conséquent, pour surveiller correctement l'utilisation des coeurs de CPU via l'interface de ligne de commande, utilisez la **show processes cpu thread** commande.

Surveiller les statistiques TCP

Cisco IOS® gère des statistiques détaillées sur chaque session TCP. La commande CLI **show tcp brief** affiche la liste de toutes les sessions TCP existantes. Dans cette sortie récapitulative, pour chaque session TCP, vous pouvez voir les informations suivantes :

- **PCB :** identifiant de session TCP unique.
- **VRF-ID** : ID du VRF dans lequel la session existe.
    - Pour afficher le nom VRF correspondant, exécutez la commande suivante :
    - show cef vrf all summary | utility egrep "^VRF:|Vrfid" | utility egrep -B1 <VRF-ID>

- **Recv-Q** : taille instantanée de la file d'attente de réception. La file d'attente de réception contient les paquets reçus de NetIO. Le processus **tcp** extrait les données d'un paquet et les envoie à l'application correspondante.
- **Send-Q** : taille instantanée de la file d'attente d'envoi. Cette file d'attente contient les données reçues d'une application. Le processus **tcp** fractionne les données en segments TCP (dictés par la taille maximale négociée des segments - TCP MSS), encapsule chaque segment dans un en-tête de couche 3 de la famille d'adresses correspondante (IPv4 ou IPv6) et envoie le paquet à NetIO.
- **Local Address** : adresse IPv4 ou IPv6 locale associée au socket TCP. Les sessions TCP à l'état LISTEN sont généralement liées à l'adresse IP "**any**", qui est représentée par "0.0.0.0" ou "::" dans le cas d'IPv4 ou d'IPv6 respectivement.
- **Adresse étrangère** : adresse IPv4 ou IPv6 distante associée au socket TCP. Les sessions TCP à l'état LISTEN sont généralement liées à l'adresse IP "**any**", qui est représentée par "0.0.0.0" ou "::" dans le cas d'IPv4 ou d'IPv6 respectivement.
- **État** : état de la session TCP. Les états de session TCP possibles sont : LISTEN, SYNSENT, SYNRCVD, ESTAB, LASTACK, CLOSING, CLOSEWAIT, FINWAIT1, FINWAIT2, TIMEWAIT, CLOSED.

Comme le numéro de port BGP bien connu est 179, vous pouvez limiter les sessions TCP affichées à celles qui sont associées à l'application

BGP.

Exemple :

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

Vous pouvez utiliser la valeur PCB affichée pour obtenir les statistiques d'une session TCP particulière. Commandes CLI qui fournissent des informations sur les statistiques de processus TCP :

Global :

show tcp statistics clients location <active_RP>

show tcp statistics summary location <active_RP>

Par PCB :

show tcp brief | i ":179"

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

show tcp statistics pcb <pcb> location <active_RP>

Les commandes de statistiques TCP globales indiquent l'état général des sessions TCP. Outre les statistiques de paquets de données (entrée/sortie), vous pouvez voir par exemple s'il y a des paquets avec des erreurs de somme de contrôle, des paquets mal formés, des paquets abandonnés en raison d'erreurs d'authentification, des paquets dans le désordre, des paquets avec des données après la fenêtre, ce qui vous donne une indication du comportement des homologues TCP.

Dans les commandes par PCB, vous pouvez voir les paramètres importants d'une session TCP, comme MSS, le temps de parcours aller-retour maximum, etc.

Les compteurs pertinents dans le résultat de la commandeshow tcp detail pcb sont :

- **Retrans Timer Starts** : indique le nombre de démarrages du compteur de retransmission.
- **Retrans Timer Wakeups** : indique combien de fois le compteur de retransmission s'est-il écoulé, déclenchant une retransmission du segment TCP.
- **Taille actuelle de la file d'attente d'envoi en octets** : octets sans accusé de réception de l'homologue.
- **Taille actuelle de la file d'attente de réception en octets/paquets** : octets/paquets à lire par l'application (BGP).

- **octets désordonnés** : octets mis en file d'attente dans la file d'attente de sauvegarde en raison d'un trou dans la fenêtre de réception 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

Surveiller l'utilisation de mémoire

La table de routage BGP est stockée dans la mémoire du tas de processus BGP. La table de routage est stockée dans la mémoire du tas de processus RIB.

Commandes utiles pour la surveillance de la mémoire du tas :

show memory summary

show memory summary detail

show memory-top-consumers

show memory heap summary all

Chemin du capteur de télémétrie :

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

FIB stocke les entrées de transfert dans l'espace mémoire partagé.

Commandes utiles pour la surveillance de la mémoire partagée :

show memory summary

show memory summary detail

show shmwin summary

Surveiller les performances du processus BGP

Commande utile qui fournit des données internes sur les performances du processus BGP :

show bgp process performance-statistics

show bgp process performance-statistics detail

Surveiller la convergence BGP

Une autre commande utile est celle qui montre l'état global de la convergence BGP : show bgp convergence

Lorsque le réseau est stable, vous voyez quelque chose comme ceci :

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

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.