

Solucionar problemas de rendimiento de routers C8000v

Contenido

[Introducción](#)

[Componentes Utilizados](#)

[Resolución general de problemas](#)

[Desbordamientos](#)

[Caídas de funciones](#)

[Gotas](#)

[Hipervisores](#)

[VMware ESXi](#)

[AWS](#)

[Colas Multi-TX](#)

[Métricas excedidas](#)

[Microsoft Azure](#)

[Redes aceleradas](#)

[Azure y fragmentación](#)

[Tipos de instancia admitidos para Microsoft Azure](#)

[Recursos adicionales](#)

Introducción

Este documento describe cómo resolver problemas de rendimiento en routers empresariales C8000v en nubes públicas y escenarios de ESXi.

Componentes Utilizados

La información de este documento se basa en estos componentes de hardware y software:

- C8000v con versión 17.12
- ESXi Versión 7.0 U3

La información que contiene este documento se creó a partir de los dispositivos en un ambiente de laboratorio específico. Todos los dispositivos que se utilizan en este documento se pusieron en funcionamiento con una configuración verificada (predeterminada). Si tiene una red en vivo, asegúrese de entender el posible impacto de cualquier comando.

Resolución general de problemas

Aunque puede alojar su C8000v en diferentes entornos, todavía puede seguir algunos pasos de solución de problemas que son idénticos independientemente de dónde esté alojado el C8000v.

Comencemos con los aspectos básicos. Lo primero que debe asegurarse es si el dispositivo está alcanzando sus límites de capacidad o no. Para ello, puede comenzar por la verificación de estos dos resultados:

1. `show platform hardware qfp active datapath util summary` - Este comando le brinda la información completa de la entrada/salida que el C8000v está recibiendo y transmitiendo desde cada puerto. Debe centrar su atención en el porcentaje de carga de procesamiento. Si se encuentra en una situación en la que alcanza el 100%, significa que está alcanzando el límite de capacidad

```
----- show platform hardware qfp active datapath utilization summary -----  
  
CPP 0:                5 secs          1 min          5 min          60 min  
Input:   Total (pps)      93119          92938          65941          65131  
         (bps)      997875976     1000204000     708234904     699462016  
Output:  Total (pps)      93119          92949          65944          65131  
         (bps)      1052264704     1054733128     746744264     737395744  
Processing: Load (pct)      14             14             10             10
```

2. `show platform hardware qfp active datapath infrastructure sw-cio` - Considere este comando como una versión más detallada del anterior. Proporciona más detalles sobre los núcleos individuales, incluidos los núcleos de E/S y criptografía que no forman parte del número de utilización de QFP. Es muy útil en un escenario donde se desea ver si un núcleo de plano de datos específico está causando un cuello de botella.

7	Gi4	4:	1993	0	0	0	0	0	0	0	0	0	0
8	Gi5	4:	2009	0	0	0	0	0	0	0	0	0	0
9	Gi6	4:	2015	0	0	0	0	0	0	0	0	0	0
10	Gi7	4:	2002	0	0	0	0	0	0	0	0	0	0
11	vpg0	400:	490	0	0	0	0	0	0	0	0	0	0

Core Utilization over preceding 107352.2729 seconds

ID:	0	1	2	3	4	5	6	7	8	9	10	11
% PP:	2.98	2.01	1.81	1.67	1.60	1.53	1.35	1.30	1.25	1.19	2.19	1.19
% RX:	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
% TM:	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
% IDLE:	97.02	97.99	98.19	98.33	98.40	98.47	98.65	98.70	98.75	98.81	97.81	98.81

Ahora, ha determinado si está alcanzando el límite de la plataforma o no. El siguiente paso sería comprobar si hay caídas. Están inherentemente conectados a problemas de rendimiento. Existen tres tipos de caídas que puede considerar dependiendo de dónde se estén produciendo.

- Desbordamientos: Este tipo de caída de paquetes ocurre en el extremo Rx. Se producen porque se ha superado la capacidad de procesamiento de uno o varios núcleos.
- Caídas de funciones: Este tipo de descarte de paquetes se produce en el PPE. Están relacionados con funciones del router como una ACL o QoS.
- Gotas de cola: Este tipo de caída de paquetes ocurre en el extremo Tx. Ocurren debido a la congestión en las memorias intermedias Tx.

Para identificar qué caídas está experimentando, puede utilizar los siguientes resultados:

- show platform hardware qfp active drop state clear
- show interface
- show policy map interface

Usted verifica cómo identificar a qué caídas se enfrenta y cómo mitigarlas. Sin embargo, el mayor enfoque en este artículo va a ser las caídas que se conocen como Taildrops, ya que son particularmente difíciles de resolver en routers virtuales.

Desbordamientos

Una caída de desbordamiento en Cisco IOS XE ocurre cuando la interfaz de red recibe los paquetes más rápido de lo que puede procesarlos o almacenarlos en su buffer. Específicamente, las memorias intermedias internas de las interfaces (cola FIFO) se llenan porque la velocidad de datos entrantes excede la capacidad del hardware para manejarlo. Como resultado, los nuevos paquetes entrantes no se pueden almacenar y se descartan, lo que incrementa el contador de desbordamiento. Básicamente, se trata de una pérdida de paquetes causada por la saturación temporal de la interfaz.

Este tipo de caídas de paquetes ocurre en el extremo Rx. Se producen porque se ha superado la capacidad de procesamiento de uno o más núcleos y el subproceso Rx no puede distribuir los paquetes entrantes al subproceso PP relevante y los búferes de entrada ya están llenos. Para hacer una analogía simple, puede pensarlo como una cola en un contador de caja que se llena

Port Instance Bindings:

```
ID Port IOS Port WRKR 2
1 rc10 rc10 Rx+Tx
2 ipc ipc Rx+Tx
3 vxe_punti vxe_puntif Rx+Tx
4 Gi1 GigabitEthernet1 Rx+Tx
5 Gi2 GigabitEthernet2 Rx+Tx <<< in this case, WRKR2 is the thread responsible for both Rx and Tx
```

Luego, puede analizar la utilización del subproceso específico responsable del tráfico Rx de esta interfaz y su número de créditos.

En un escenario donde Gig2 está observando problemas de rendimiento debido a sobrecostos, puede esperar que el PP#2 esté constantemente utilizado (Inactivo = 0%), y créditos bajos/cero para la interfaz Gig2:

```
#show platform hardware qfp active datapath infrastructure sw-cio
Credits Usage:
```

```
ID Port Wght Global WRKR0 WRKR1 WRKR2 Total
1 rc10 16: 487 0 0 25 512
1 rc10 32: 496 0 0 16 512
2 ipc 1: 490 0 0 21 511
3 vxe_punti 4: 459 0 0 53 512
4 Gi1 4: 477 0 0 35 512
5 Gi2 4: 474 0 0 38 512 <<< low/zero credits for interface Gig2:
```

```
Core Utilization over preceding 1.0047 seconds
```

```
-----
ID: 0 1 2
% PP: 0.77 0.00 0.00
% RX: 0.00 0.00 0.44
% TM: 0.00 0.00 5.63
% IDLE: 99.23 99.72 93.93 <<< the core ID relevant in this case would be PP#2
```

Caídas de funciones

Los paquetes son manejados por cualquier subproceso de plano de datos disponible y se distribuyen estrictamente en función de la disponibilidad de los núcleos QFP a través de la función Rx del software (x86) - Distribución basada en carga (LBD). Los paquetes que llegan al PPE se pueden descartar con una razón de descarte QFP específica, que se puede comprobar usando esta salida:

```
#show drops
----- show platform hardware qfp active statistics drop detail -----
```

```
Last clearing of QFP drops statistics : never
```

```
-----
ID Global Drop Stats Packets Octets
```

```

-----
319 BFDoffload          403          31434
139 Disabled            105          7487
 61 Icmp                135          5994
 94 Ipv4NoAdj           1            193
 33 Ipv6NoRoute        2426         135856
215 UnconfiguredIpv4Fia 1937573      353562196
216 UnconfiguredIpv6Fia 8046173      1057866418
-----

```

```

----- show platform hardware qfp active interface all statistics drop_summary -----
-----

```

Drop Stats Summary:

note: 1) these drop stats are only updated when PAL reads the interface stats.

2) the interface stats include the subinterface

Interface	Rx Pkts	Tx Pkts
GigabitEthernet1	9980371	0
GigabitEthernet2	4012	0

Las razones de las caídas son diversas y generalmente se explican por sí mismas. Para investigar más a fondo, se puede utilizar un [seguimiento de paquetes](#).

Gotas

Como se mencionó anteriormente, las caídas de cola ocurren cuando el dispositivo está tratando de transmitir paquetes, pero las memorias intermedias de transmisión están llenas.

En esta subsección, va a ver qué resultados puede examinar cuando se enfrente a este tipo de situación. Qué valores puede ver en ellos significan y qué puede hacer para mitigar el problema.

En primer lugar, debe saber cómo identificarlos. Una de estas maneras es simplemente mirar la interfaz show. Esté atento a cualquier aumento de las gotas de salida:

```

GigabitEthernet2 is up, line protocol is up
Hardware is vNIC, address is 0050.56ad.c777 (bia 0050.56ad.c777)
Description: Connected-To-ASR Cloud Gateway
Internet address is 10.6.255.81/29
MTU 1500 bytes, BW 10000000 Kbit/sec, DLY 10 usec,
reliability 255/255, txload 2/255, rxload 3/255
Encapsulation ARPA, loopback not set
Keepalive set (10 sec)
Full Duplex, 10000Mbps, link type is force-up, media type is Virtual
output flow-control is unsupported, input flow-control is unsupported
ARP type: ARPA, ARP Timeout 04:00:00
Last input 00:00:00, output 00:00:00, output hang never
Last clearing of "show interface" counters 03:16:21
Input queue: 0/375/0/0 (size/max/drops/flushes); Total output drops: 7982350 <<<<<<<<
Queueing strategy: fifo
Output queue: 0/40 (size/max)
5 minute input rate 150449000 bits/sec, 20461 packets/sec

```

5 minute output rate 89116000 bits/sec, 18976 packets/sec

Este comando es particularmente útil para entender si está experimentando congestión o no:

- `show platform hardware qfp active datapath infrastructure` - HQF significa `estructura jerárquica de almacenamiento en cola`. Se trata de una función que permite la gestión de la calidad del servicio (QoS) en diferentes niveles (físico, lógico y de clase) mediante la interfaz de línea de comandos (MQC) de QoS modular. Muestra los costes actuales de RX y TX. Cuando la cola TX está llena, como muestra el resultado (1959 completo)

```
pmd b1689fc0 device Gi1
RX: pkts 5663120 bytes 1621226335 return 0 badlen 0
Out-of-credits: Hi 0 Lo 0
pkts/burst 1 cycl/pkt 1565 ext_cycl/pkt 1173
Total ring read 12112962299, empty 12107695202
TX: pkts 8047873582 bytes 11241140363740
pri-0: pkts 8047873582 bytes 11241140363740
pkts/send 3
Total: pkts/send 3 cycl/pkt 452
send 2013612969 sendnow 1810842
forced 2013274797 poll 724781 thd_poll 0
blocked 2197451 retries 7401 mbuf alloc err 0
TX Queue 0: full 1959 current index 0 hiwater 224
```

El resultado sugiere que el hardware subyacente no está al día con el envío de paquetes. Para depurar la interfaz subyacente, es posible que deba buscar fuera del C8000v y en el entorno subyacente en el que se está ejecutando el C8000v para ver si se han notificado errores adicionales en las interfaces físicas subyacentes.

Para verificar el entorno, hay un paso que puede realizar antes de verificar en qué hipervisor se está ejecutando el router C8000v. Esto es para verificar la salida del comando `show controller`. Sin embargo, usted puede encontrarse perdido en lo que cada contador significa o dónde mirar.

En primer lugar, un detalle importante que debe tener en cuenta al ver este resultado es que la información se obtiene principalmente de las propias vNIC. Cada controlador NIC tiene un conjunto específico de contadores que utilizan, que pueden variar según el controlador, naturalmente. Diferentes hipervisores tienen algún tipo de efecto en lo que se presenta también. Algunos contadores, como los contadores mbuf, son estadísticas del controlador DPDK. Estos pueden variar para diferentes controladores DPDK. El recuento real lo realiza generalmente el hipervisor en la capa de virtualización.

```
GigabitEthernet2 - Gi2 is mapped to UIO on VXE
rx_good_packets 1590
tx_good_packets 1402515
rx_good_bytes 202860
tx_good_bytes 1857203911
rx_missed_errors 0
```

```
rx_errors 0
tx_errors 0
rx_mbuf_allocation_errors 0
rx_q0_packets 1590
rx_q0_bytes 202860
rx_q0_errors 0
tx_q0_packets 1402515
tx_q0_bytes 1857203911
rx_q0_drop_total 0
rx_q0_drop_err 0
rx_q0_drop_fcs 0
rx_q0_rx_buf_alloc_failure 0
tx_q0_drop_total 976999540797
tx_q0_drop_too_many_segs 0
tx_q0_drop_tso 0
tx_q0_tx_ring_full 30901211518
```

Tómese un minuto aquí para aprender a interpretar y leer estos contadores:

1. Si ve subX, significa que es una sub-interfaz - una división lógica de la interfaz principal. El sub0 es generalmente el primario/predeterminado. Estos se utilizan a menudo cuando hay varias VLAN involucradas.
2. Luego, tiene "rx = recepción" y "tx = transmisión".
3. Finalmente, q0 se refiere a la cola primero/predeterminado utilizada por esa interfaz

Aunque no hay una descripción para cada contador, el artículo describe algunos de ellos, que pueden ser relevantes para su solución de problemas:

- "RX_MISSED_ERRORS:" aparece cuando el búfer de NIC (Rx FIFO) se llena en exceso. Esta condición conduce a caídas y a un aumento de la latencia. Una posible solución a esto es aumentar el búfer NIC (lo cual es imposible en nuestro caso) o cambiar el controlador NIC.
- "tx_q0_drop_total" y "tx_q0_tx_ring_full": Estos pueden indicarle que el host está descartando paquetes y que el C8000v está experimentando descartes de cola en el C8000v porque el host está presionando hacia atrás el C8000v

En el resultado anterior, no vemos ningún "rx_missing_errors". Sin embargo, como nos estamos centrando en las caídas de cola, vemos "tx_q0_drop_total" y "tx_q0_tx_ring_full". Con esto, podemos concluir que existe una congestión causada por el hardware subyacente del host.

Como se ha mencionado anteriormente, cada hipervisor tiene algún tipo de efecto en lo que se presenta. El artículo se centra en esto en la siguiente sección, ya que trata las diferencias entre los diferentes hipervisores en los que se puede alojar el C8000v. También puede encontrar las diferentes recomendaciones para tratar de mitigar este tipo de problema en cada una de ellas.

Hipervisores

Un hipervisor es una capa de software que permite que varios sistemas operativos (denominados máquinas virtuales o VM) se ejecuten en un único host de hardware físico mediante la administración y asignación de los recursos de hardware, como la CPU, la memoria y el

almacenamiento, a cada VM. Garantiza que estas máquinas virtuales funcionen de forma independiente sin interferir unas con otras.

En el contexto de Cisco Catalyst 8000V (C8000v), el hipervisor es la plataforma que aloja la máquina virtual C8000v. ¿Cómo puede averiguar qué hipervisor aloja su C8000v? Hay un resultado bastante útil que nos da esa información. Además, también puede comprobar a qué tipo de recursos tiene acceso nuestro router virtual:

```
C8000v#show platform software system all
Processor Details
=====
Number of Processors : 8
Processor : 1 - 8
vendor_id : GenuineIntel
cpu MHz : 2593.906
cache size : 36608 KB
Crypto Supported : Yes
model name : Intel(R) Xeon(R) Platinum 8272CL CPU @ 2.60GHz

Memory Details
=====
Physical Memory : 32817356KB

VNIC Details
=====
Name Mac Address Driver Name Status Platform MTU
GigabitEthernet1 0022.480d.7a05 net_netvsc UP 1500
GigabitEthernet2 6045.bd69.83a0 net_netvsc UP 1500
GigabitEthernet3 6045.bd69.8042 net_netvsc UP 1500

Hypervisor Details
=====
Hypervisor: AZURE
Manufacturer: Microsoft Corporation
Product Name: Virtual Machine
Serial Number: 0000-0002-0201-5310-5478-4052-71
UUID: 8b06091c-f1d3-974c-85a5-a78dfb551bf2
Image Variant: None
```

VMware ESXi

ESXi es un hipervisor de tipo 1 desarrollado por VMware que se instala directamente en servidores físicos para habilitar la virtualización. Permite ejecutar varias máquinas virtuales (VM) en un único servidor físico mediante la abstracción de los recursos de hardware y su asignación a cada VM. El router C8000v es una de esas VM.

Puede empezar por repasar un escenario común en el que se está produciendo una congestión. Esto se puede confirmar al verificar el contador tx_q0_tx_ring_full:

Ejemplo:

```
----- show platform software vnic-if interface-mapping -----
```

```
-----  
Interface Name Driver Name Mac Addr  
-----
```

```
GigabitEthernet3 net_vmxnet3 <-- 0050.5606.2239  
GigabitEthernet2 net_vmxnet3 0050.5606.2238  
GigabitEthernet1 net_vmxnet3 0050.5606.2237  
-----
```

```
GigabitEthernet3 - Gi3 is mapped to UIO on VXE
```

```
rx_good_packets 99850846  
tx_good_packets 24276286  
rx_good_bytes 78571263015  
tx_good_bytes 14353154897  
rx_missed_errors 0  
rx_errors 0  
tx_errors 0  
rx_mbuf_allocation_errors 0  
rx_q0packets 99850846  
rx_q0bytes 78571263015  
rx_q0errors 0  
tx_q0packets 24276286  
tx_q0bytes 14353154897  
rx_q0_drop_total 0  
rx_q0_drop_err 0  
rx_q0_drop_fcs 0  
rx_q0_rx_buf_alloc_failure 0  
tx_q0_drop_total 160945155  
tx_q0_drop_too_many_segs 0  
tx_q0_drop_tso 0  
tx_q0_tx_ring_full 5283588 <-----
```

Esta congestión se produce cuando el C8000V intenta enviar paquetes a través de la interfaz VMXNET3. Sin embargo, el anillo de búfer ya está lleno de paquetes, que terminan en retrasos o caídas.

En estas condiciones, estas caídas se producen en el hipervisor, como hemos mencionado anteriormente. Si se cumplen todas las recomendaciones, se recomienda consultar con el soporte de VMware para saber qué está sucediendo en la NIC.

A continuación, se ofrecen algunas sugerencias sobre cómo mejorar el rendimiento:

- Utilice un vSwitch dedicado y un enlace ascendente para obtener un rendimiento óptimo
- Al asignar el C8000V a un vSwitch dedicado respaldado por su propio enlace ascendente físico, podemos aislar su tráfico de vecinos ruidosos y evitar cuellos de botella de recursos compartidos.

Hay algunos comandos que vale la pena examinar desde el lado de ESXi. Por ejemplo, para comprobar la pérdida de paquetes desde la interfaz ESXi, podemos hacer lo siguiente:

1. Activar SSH.
2. Conexión a ESXi mediante SSH.

3. Ejecute esxtop.

4. Escriba n.

El comando esxtop puede mostrar los paquetes descartados en el switch virtual si el controlador de red de la máquina virtual se queda sin memoria intermedia Rx. Aunque esxtop muestra los paquetes como descartados en el switch virtual, en realidad se descartan entre el switch virtual y el controlador del sistema operativo invitado.

Busque los paquetes que se están descartando en %DRPTX y %DRPRX:

12:34:43pm up 73 days 16:05, 907 worlds, 9 VMs, 53 vCPUs; CPU load average: 0.42, 0.42, 0.42

```
PORT-ID USED-BY TEAM-PNIC DNAME PKTTX/s MbTX/s PSZTX PKTRX/s MbRX/s PSZRX %DRPTX %DRPRX
67108870 Management n/a vSwitch-to-9200 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00
67108872 Shadow of vmnic1 n/a vSwitch-to-9200 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00
67108876 vmk1 vmnic1 vSwitch-to-9200 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00
67108890 2101719:c8kv-gw-mgmt vmnic1 vSwitch-to-9200 76724.83 792.35 1353.00 16180.39 9.30 75.00 0.00 0.00
100663305 Management n/a vSwitch-to-Cisc 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00
100663307 Shadow of vmnic0 n/a vSwitch-to-Cisc 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00
100663309 vmk0 vmnic0 vSwitch-to-Cisc 3.64 0.01 280.00 3.29 0.00 80.00 0.00 0.00
100663310 2100707:gsoaresc-On_Prem vmnic0 vSwitch-to-Cisc 0.00 0.00 0.00 2.43 0.00 60.00 0.00 0.00
100663311 2100993:cats-vmanage void vSwitch-to-Cisc 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00
100663312 2100993:cats-vmanage void vSwitch-to-Cisc 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00
100663313 2100993:cats-vmanage vmnic0 vSwitch-to-Cisc 5.38 0.01 212.00 9.71 0.01 141.00 0.00 0.00
100663314 2101341:cats-vsmart void vSwitch-to-Cisc 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00
100663315 2101341:cats-vsmart vmnic0 vSwitch-to-Cisc 2.60 0.00 164.00 6.94 0.01 124.00 0.00 0.00
100663316 2101522:cats-vbond vmnic0 vSwitch-to-Cisc 0.00 0.00 0.00 0.00 0.00 0.00 0.00 100.00
100663317 2101522:cats-vbond vmnic0 vSwitch-to-Cisc 0.00 0.00 0.00 0.00 0.00 0.00 0.00 100.00
100663318 2101522:cats-vbond vmnic0 vSwitch-to-Cisc 4.33 0.01 174.00 7.80 0.01 162.00 0.00 0.00
100663319 2101522:cats-vbond vmnic0 vSwitch-to-Cisc 0.00 0.00 0.00 4.16 0.00 90.00 0.00 0.00
100663320 2101547:gdk-backup vmnic0 vSwitch-to-Cisc 0.00 0.00 0.00 3.12 0.00 77.00 0.00 0.00
100663321 2101703:sevvy vmnic0 vSwitch-to-Cisc 0.00 0.00 0.00 3.12 0.00 77.00 0.00 0.00
100663323 2101719:c8kv-gw-mgmt vmnic0 vSwitch-to-Cisc 16180.91 9.09 73.00 76755.87 792.44 1353.00 0.00 0.00
100663324 2137274:telemetry-server vmnic0 vSwitch-to-Cisc 0.00 0.00 0.00 3.12 0.00 77.00 0.00 0.00
100663335 2396721:netlab vmnic0 vSwitch-to-Cisc 0.00 0.00 0.00 3.12 0.00 77.00 0.00 0.00
2214592519 vmnic1 - vSwitch-to-9200 76727.26 792.38 1353.00 16182.64 9.30 75.00 0.00 0.00
2248146954 vmnic0 - vSwitch-to-Cisc 16189.05 9.32 75.00 76736.97 792.38 1353.00 0.00 0.00
```

Este comando enumera todas las NIC configuradas en un host:

```
esxcli network nic list
```

```
Name PCI Device Driver Admin Status Link Status Speed Duplex MAC Address MTU Description
-----
vmnic0 0000:01:00.0 igbn Up Up 1000 Full fc:99:47:49:c5:0a 1500 Intel(R) I350 Gigabit Network Connection
vmnic1 0000:01:00.1 igbn Up Up 1000 Full fc:99:47:49:c5:0b 1500 Intel(R) I350 Gigabit Network Connection
vmnic2 0000:03:00.0 ixgben Up Up 1000 Full a0:36:9f:1c:1f:cc 1500 Intel(R) Ethernet Controller 10 Gigabit Network Connection
vmnic3 0000:03:00.1 ixgben Up Up 1000 Full a0:36:9f:1c:1f:ce 1500 Intel(R) Ethernet Controller 10 Gigabit Network Connection
```

También existe un comando útil para comprobar el estado de la vNIC asignada a una VM concreta.

```
esxcli network vm list
World ID Name Num Ports Networks
-----
2137274 telemetry-server 1 Cisco Backbone 10.50.25.0/24
2101703 sevy 1 Cisco Backbone 10.50.25.0/24
2396721 netlab 1 Cisco Backbone 10.50.25.0/24
2101547 gdk-backup 1 Cisco Backbone 10.50.25.0/24
2101522 cats-vbond 4 VPNO, VPNO, VPNO, VPNO
2101719 c8kv-gw-mgmt 2 c8kv-to-92001, c8kv-to-cisco
2100707 gsoaresc-On_Prem 1 Cisco Backbone 10.50.25.0/24
2100993 cats-vmanage 3 VPNO, VPNO, VPNO
2101341 cats-vsmart 2 VPNO, VPNO
[root@localhost:~]
```

Si observamos c8kv-gw-mgmt , que es una máquina virtual C8000v, hay 2 redes asignadas:

- c8kv-to-92001
- c8kv a cisco

Puede utilizar el ID de mundo para buscar más información sobre esta VM:

```
[root@localhost:~] esxcli network vm port list -w 2101719
Port ID: 67108890
vSwitch: vSwitch-to-9200L
Portgroup: c8kv-to-92001
DVPort ID:
MAC Address: 00:0c:29:31:a6:b6
IP Address: 0.0.0.0
Team Uplink: vmnic1
Uplink Port ID: 2214592519
Active Filters:

Port ID: 100663323
vSwitch: vSwitch-to-Cisco
Portgroup: c8kv-to-cisco
DVPort ID:
MAC Address: 00:0c:29:31:a6:ac
IP Address: 0.0.0.0
Team Uplink: vmnic0 <----
Uplink Port ID: 2248146954
Active Filters:
[root@localhost:~]
```

Una vez que disponga de esta información, podrá identificar a qué red está asignado el vSwitch.

Para verificar algunas estadísticas de tráfico de la NIC física asignada al vSwitch, tenemos el

comando this:

```
# esxcli network nic stats get -n <vmnic>
```

Este comando muestra información como paquetes recibidos, bytes recibidos, paquetes descartados y errores recibidos. Esto puede ayudar a identificar si hay alguna caída sucediendo en la NIC.

```
[root@localhost:~] esxcli network nic stats get -n vmnic0
NIC statistics for vmnic0
Packets received: 266984237
Packets sent: 123640666
Bytes received: 166544114308
Bytes sent: 30940114661
Receive packets dropped: 0
Transmit packets dropped: 0
Multicast packets received: 16773454
Broadcast packets received: 36251726
Multicast packets sent: 221108
Broadcast packets sent: 1947649
Total receive errors: 0
Receive length errors: 0
Receive over errors: 0
Receive CRC errors: 0
Receive frame errors: 0
Receive FIFO errors: 0
Receive missed errors: 0
Total transmit errors: 0
Transmit aborted errors: 0
Transmit carrier errors: 0
Transmit FIFO errors: 0
Transmit heartbeat errors: 0
Transmit window errors: 0
```

Hay algunas configuraciones que se deben verificar que pueden mejorar el rendimiento de Cisco Catalyst 8000V que se ejecuta en un entorno ESXi modificando la configuración en el host y la máquina virtual:

- Establezca el hardware virtual: Reserva de CPU establecida en Máximo.
- Reserve toda la memoria de invitado en el hardware virtual: Memoria.
- Seleccione VMware Paravirtual en Hardware virtual: Controlador SCSI.
- Desde el hardware virtual: Adaptador de red: Tipo de adaptador, seleccione SR-IOV para las NIC compatibles
- Establezca la opción General Guest OS Version > VM Options en Other 3.x o posterior Linux (64 bits).
- Establezca la opción Opciones de VM en Sensibilidad de latencia avanzada en Alta.
- En Opciones de VM > Configuración de edición avanzada, agregue "numa.nodeAffinity" al

mismo nodo NUMA que la NIC SRIOV

- Active la configuración de rendimiento del hipervisor.
- Limite la sobrecarga del vSwitch habilitando SR-IOV en las NIC físicas admitidas.
- Configure las vCPU de la VM para que se ejecuten en el mismo nodo NUMA que las NIC físicas.
- Establezca la sensibilidad de latencia de VM en High (Alta).

AWS

El C8000v admite la implementación en AWS al iniciarse como una imagen de máquina de Amazon (AMI) dentro de una nube privada virtual (VPC) de Amazon, lo que permite a los usuarios aprovisionar una sección aislada lógicamente de la nube de AWS para sus recursos de red.

Colas Multi-TX

En un C8000v que se ejecuta en AWS, una función clave es el uso de colas de transmisión múltiple (Multi-TXQs). Estas colas ayudan a reducir la sobrecarga de procesamiento interno y a mejorar la escalabilidad. Disponer de varias colas hace que asignar paquetes entrantes y salientes a la CPU virtual (vCPU) correcta sea más rápido y sencillo.

A diferencia de algunos sistemas en los que las colas RX/TX se asignan por vCPU, en el C8000v, estas colas se asignan por interfaz. Las colas RX (de recepción) y TX (de transmisión) sirven como puntos de conexión entre la aplicación Catalyst 8000V y la infraestructura o hardware de AWS, y administran cómo se envía y recibe el tráfico de red. AWS controla el número y la velocidad de las colas RX/TX disponibles para cada interfaz, dependiendo del tipo de instancia.

Para crear varias colas TX, Catalyst 8000V necesita tener varias interfaces. Cuando se habilitan varias colas TX, el dispositivo mantiene el orden de los flujos de paquetes mediante un método de hashing basado en la 5-tupla del flujo (IP de origen, IP de destino, puerto de origen, puerto de destino y protocolo). Este hashing decide qué cola TX utilizar para cada flujo.

Los usuarios pueden crear varias interfaces en el Catalyst 8000V utilizando la misma tarjeta de interfaz de red (NIC) física conectada a la instancia de AWS. Esto se realiza mediante la configuración de interfaces de loopback o la adición de direcciones IP secundarias.

Con Multi-TXQs, hay varias colas de transmisión para manejar el tráfico saliente. En el ejemplo, hay doce colas TX (numeradas del 0 al 11). Esta configuración le permite supervisar cada cola individualmente para ver si alguna se está llenando.

Si observa el resultado, puede ver que TX Queue 8 tiene un contador "completo" muy alto (56,406,998), lo que significa que su búfer se está llenando con frecuencia. Las otras colas TX muestran cero para el contador "completo", lo que indica que no están congestionadas.

```
Router#show platform hardware qfp active datapath infrastructure sw-cio
```


red uniforme y de alta calidad en diferentes tamaños de instancia. Estos límites ayudan a mantener una red estable para todos los usuarios.

Puede verificar estos límites y las estadísticas relacionadas mediante el comando `show controllers` en su dispositivo. El resultado incluye muchos contadores, pero aquí nos centramos solo en los más importantes para supervisar el rendimiento de la red:

```
c8kv-2#sh control | inc exceed
<snipped>
bw_in_allowance_exceeded 0
bw_out_allowance_exceeded 0
pps_allowance_exceeded 0
conntrack_allowance_exceeded 0
linklocal_allowance_exceeded 0
<snipped>
```

Ahora puede sumergirse y ver a qué se refieren exactamente esos contadores:

- `bw_in_allowance_exceeded`: Número de paquetes en cola o descartados porque el ancho de banda entrante superó el límite de la instancia.
- `bw_out_allowance_exceeded`: Número de paquetes en cola o descartados debido a que el ancho de banda saliente excedió el límite de la instancia.
- `pps_allowance_exceeded`: Número de paquetes en cola o descartados porque el total de paquetes por segundo (PPS) superó el límite de la instancia.
- `conntrack_allowance_exceeded`: Número de conexiones supervisadas que alcanzaron el máximo permitido para el tipo de instancia.
- `linklocal_allowance_exceeded`: Número de paquetes descartados porque el tráfico a los servicios proxy locales (como Amazon DNS, Instance Metadata Service y Time Sync Service) superó el límite de PPS para la interfaz de red. Esto no afecta a los resolvers DNS personalizados.

Lo que esto significa para el rendimiento de C8000v:

- Si observa que estos contadores aumentan y experimenta problemas de rendimiento, no siempre significa que el router C8000v sea el problema. En su lugar, a menudo indica que la instancia de AWS que está utilizando ha alcanzado sus límites de capacidad. Puede comprobar las especificaciones de su instancia de AWS para asegurarse de que puede gestionar sus necesidades de tráfico.

Microsoft Azure

En esta sección, explore cómo Microsoft Azure y el router virtual Cisco C8000v se combinan para ofrecer soluciones de red virtuales escalables, seguras y de alto rendimiento en la nube.

Conozca cómo las redes aceleradas (AN) y la fragmentación de paquetes pueden afectar al rendimiento. Además de revisar la importancia de usar un tipo de instancia compatible con Microsoft Azure.

Redes aceleradas

En casos de problemas de rendimiento en los que el C8000v está alojado en la nube de Microsoft Azure. Un aspecto que no puede pasar por alto es si Accelerated Network está o no habilitado. A medida que aumenta considerablemente el rendimiento del router. En pocas palabras, la red acelerada permite la virtualización de E/S de raíz única (SR-IOV) en máquinas virtuales como, por ejemplo, una máquina virtual Cisco Catalyst 8000V. La ruta de red acelerada omite el switch virtual, aumenta la velocidad del tráfico de red, mejora el rendimiento de la red y reduce la latencia y la fluctuación de la red.

Existe una forma muy sencilla de comprobar si la red acelerada está activada. Esto es para verificar la salida de show controllers y verificar si hay o no ciertos contadores:

```
----- show controllers -----
```

```
GigabitEthernet1 - Gi1 is mapped to UIO on VXE
  rx_good_packets 6497723453
  tx_good_packets 14690462024
  rx_good_bytes 2271904425498
  tx_good_bytes 6276731371987
```

```
rx_q0_good_packets 58576251
rx_q0_good_bytes 44254667162
```

```
vf_rx_good_packets 6439147188
vf_tx_good_packets 14690462024
vf_rx_good_bytes 2227649747816
vf_tx_good_bytes 6276731371987
```

Los contadores que está buscando son los que comienzan con vf como vf_rx_good_packets. Si verifica que estos contadores están presentes, puede estar absolutamente seguro de que la red acelerada está habilitada.

Azure y fragmentación

La fragmentación puede tener implicaciones negativas en el rendimiento. Una de las principales razones del efecto sobre el rendimiento es el efecto CPU/memoria de la fragmentación y reensamblado de paquetes. Cuando un dispositivo de red necesita fragmentar un paquete, tiene

que asignar recursos de CPU/memoria para realizar la fragmentación.

Lo mismo sucede cuando se reensambla el paquete. El dispositivo de red debe almacenar todos los fragmentos hasta que se reciban para poder reensamblarlos en el paquete original.

Azure no procesa paquetes fragmentados con Accelerated Networking. Cuando una VM recibe un paquete fragmentado, la trayectoria no acelerada lo procesa. Como resultado, los paquetes fragmentados no aprovechan las ventajas de las redes aceleradas, como una latencia más baja, una fluctuación reducida y paquetes más altos por segundo. Por esta razón, se recomienda evitar la fragmentación si es posible.

Azure, de forma predeterminada, descarta los paquetes fragmentados que llegan a la máquina virtual fuera de servicio, lo que significa que los paquetes no coinciden con la secuencia de transmisión desde el punto final de origen. Este problema puede ocurrir cuando los paquetes viajan a través de Internet u otras WAN grandes.

Tipos de instancia admitidos para Microsoft Azure

Es importante que el C800v utilice un tipo de instancia compatible según los estándares de Cisco. Se pueden encontrar en la [Guía de Instalación y Configuración del Software de Borde Cisco Catalyst 8000V](#).

Esto se debe a que los tipos de instancia de la lista son aquellos en los que se ha probado correctamente el C8KV. Ahora, existe una pregunta válida si el C8000v funciona en un tipo de instancia que no está en la lista. La respuesta probablemente sea sí. Sin embargo, cuando está solucionando problemas tan complejos como los problemas de rendimiento, no desea agregar otro factor desconocido al problema. Solo por ese motivo, Cisco TAC siempre recomienda permanecer en un tipo de instancia compatible.

Recursos adicionales

Un problema de rendimiento solo se puede solucionar realmente cuando está sucediendo en el momento. Sin embargo, esto puede ser difícil de atrapar, ya que puede suceder en cualquier momento dado. Por esa razón, proporcionamos este guion de EEM. Ayuda a capturar salidas importantes en el momento en que los paquetes comienzan a ser descartados y se presentan problemas de rendimiento:

```
ip access-list extended TAC
permit ip host host
```

```
permit ip host
```

```
host
```

```
conf t
event manager applet CONNECTIONLOST1 authorization bypass
event track 100 state down maxrun 500
action 0010 syslog msg "Logging information to file bootflash:SLA-DROPS.txt and bootflash:FIASLA_Decode.txt"
action 0020 cli command "enable"
action 0021 cli command "term length 0"
action 0022 cli command "term exec prompt timestamp"
action 0023 cli command "term exec prompt expand"
action 0095 cli command "show clock | append bootflash:SLA-DROPS.txt"
action 0096 cli command "show platform hardware qfp active statistics drop detail | append bootflash:SLA-DROPS.txt"
action 0097 cli command "show logging | append bootflash:SLA-DROPS.txt"
action 0099 cli command "show interfaces summary | append bootflash:SLA-DROPS.txt"
action 0100 cli command "show interfaces | append bootflash:SLA-DROPS.txt"
action 0101 cli command "show platform hardware qfp active statistics drop clear"
action 0102 cli command "debug platform packet-trace packet 2048 fia-trace"
action 0103 cli command "debug platform packet-trace copy packet both"
action 0104 cli command "debug platform condition ipv4 access-list TAC both"
action 0105 cli command "debug platform condition start"
action 0106 cli command "show platform hardware qfp active data infrastructure sw-cio | append bootflash:SLA-DROPS.txt"
action 0110 wait 60
action 0111 cli command "debug platform condition stop"
action 0112 cli command "show platform packet-trace packet all decode | append bootflash:FIASLA_Decode.txt"
action 0120 cli command "show platform packet-trace statistics | append bootflash:FIASLA_Decode.txt"
action 0121 cli command "show platform packet-trace summary | append bootflash:FIASLA_Decode.txt"
action 0122 cli command "show platform hardware qfp active datapath utilization summary | append bootflash:SLA-DROPS.txt"
action 0123 cli command "show platform hardware qfp active statistics drop detail | append bootflash:SLA-DROPS.txt"
action 0124 cli command "show platform hardware qfp active infrastructure bqs queue output default all | append bootflash:SLA-DROPS.txt"
action 0125 cli command "show platform software status control-processor brief | append bootflash:SLA-DROPS.txt"
action 0126 cli command "show platform hardware qfp active datapath infrastructure sw-pktmem | append bootflash:SLA-DROPS.txt"
action 0127 cli command "show platform hardware qfp active infrastructure punt statistics type per-cause | append bootflash:SLA-DROPS.txt"
action 0128 cli command "show platform hardware qfp active statistics drop | append bootflash:SLA-DROPS.txt"
action 0129 cli command "show platform hardware qfp active infrastructure bqs queue output default all | append bootflash:SLA-DROPS.txt"
action 0130 cli command "show platform hardware qfp active data infrastructure sw-hqf config 0 0 | append bootflash:SLA-DROPS.txt"
action 0131 cli command "show platform hardware qfp active feature lic-bw oversubscription | append bootflash:SLA-DROPS.txt"
action 0132 cli command "show platform hardware qfp active data infrastructure sw-hqf config 0 0 | append bootflash:SLA-DROPS.txt"
action 0133 cli command "show platform hardware qfp active data infrastructure sw-cio | append bootflash:SLA-DROPS.txt"
action 0134 cli command "show platform hardware qfp active data infrastructure sw-hqf sched | append bootflash:SLA-DROPS.txt"
action 0135 cli command "show platform hardware qfp active data infrastructure sw-dist | append bootflash:SLA-DROPS.txt"
action 0136 cli command "show platform hardware qfp active data infrastructure sw-nic | append bootflash:SLA-DROPS.txt"
action 0137 cli command "show platform hardware qfp active data infrastructure sw-pktmem | append bootflash:SLA-DROPS.txt"
action 0138 cli command "show controllers | append bootflash:SLA-DROPS.txt"
action 0139 cli command "show platform hardware qfp active datapath pmd controllers | append bootflash:SLA-DROPS.txt"
action 0140 cli command "show platform hardware qfp active datapath pmd system | append bootflash:SLA-DROPS.txt"
action 0141 cli command "show platform hardware qfp active datapath pmd static-if-config | append bootflash:SLA-DROPS.txt"
action 0150 cli command "clear platform condition all"
action 0151 cli command "clear platform packet-trace statistics"
action 0152 cli command "clear platform packet-trace configuration"
action 0153 cli command "show log | append bootflash:throughput_levelinfoSLA.txt"
action 0154 cli command "show version | append bootflash:throughput_levelinfoSLA.txt"
action 0155 cli command "show platform software system all | append bootflash:throughput_levelinfoSLA.txt"
action 0156 syslog msg "EEM script and FIA trace completed."
action 0180 cli command "conf t"
action 0181 cli command "no event manager applet CONNECTIONLOST1"
```

```
end
```

Acerca de esta traducción

Cisco ha traducido este documento combinando la traducción automática y los recursos humanos a fin de ofrecer a nuestros usuarios en todo el mundo contenido en su propio idioma.

Tenga en cuenta que incluso la mejor traducción automática podría no ser tan precisa como la proporcionada por un traductor profesional.

Cisco Systems, Inc. no asume ninguna responsabilidad por la precisión de estas traducciones y recomienda remitirse siempre al documento original escrito en inglés (insertar vínculo URL).