Resolución de problemas de reenvío intrafabric de ACI: caídas intermitentes

Contenido

Introducción

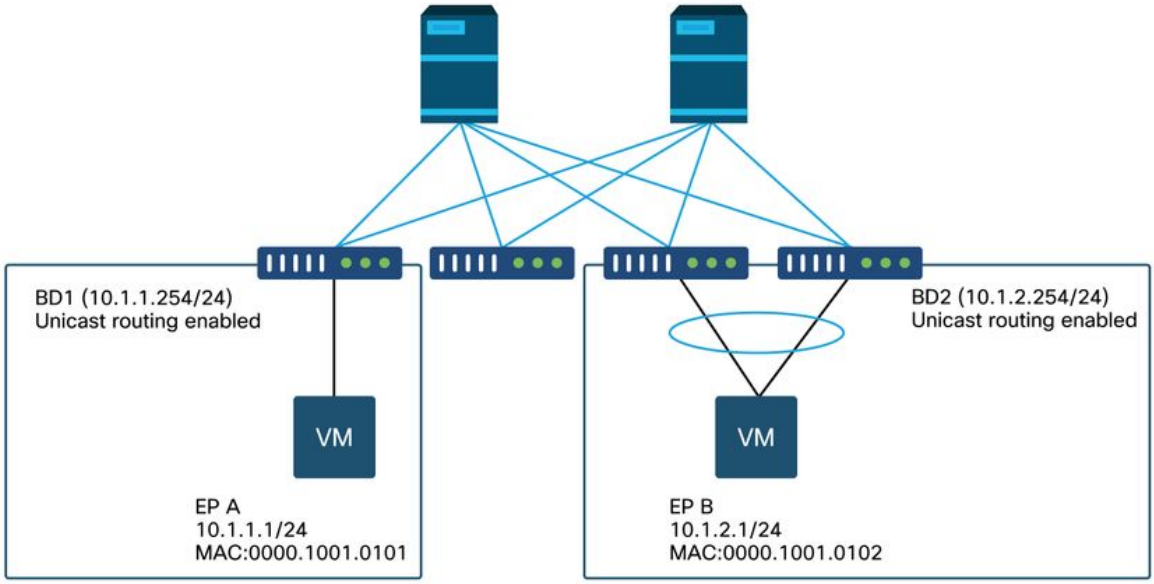
Este documento describe los pasos para solucionar problemas de caídas intermitentes en ACI.

Antecedentes

El material de este documento se extrajo del libro Troubleshooting Cisco Application Centric Infrastructure, Second Edition, específicamente del capítulo **Reenvío Intra-Fabric - Caídas intermitentes**.

Resolución de problemas de reenvío de ACI dentro del fabric: caídas intermitentes

Ejemplo de topología



En este ejemplo, el ping de EP A (10.1.1.1) a EP B (10.1.2.1) está experimentando las caídas intermitentes.

```
[EP-A ~]$ ping 10.1.2.1 -c 10
PING 10.1.2.1 (10.1.2.1) 56(84) bytes of data.
64 bytes from 10.1.2.1: icmp_seq=1 ttl=231 time=142 ms
64 bytes from 10.1.2.1: icmp_seq=2 ttl=231 time=141 ms
                                  <-- missing icmp_seq=3

64 bytes from 10.1.2.1: icmp_seq=4 ttl=231 time=141 ms
64 bytes from 10.1.2.1: icmp_seq=5 ttl=231 time=141 ms
64 bytes from 10.1.2.1: icmp_seq=6 ttl=231 time=141 ms
                                   <-- missing icmp_seq=7

64 bytes from 10.1.2.1: icmp_seq=8 ttl=231 time=176 ms
64 bytes from 10.1.2.1: icmp_seq=9 ttl=231 time=141 ms
64 bytes from 10.1.2.1: icmp_seq=10 ttl=231 time=141 ms

--- 10.1.2.1 ping statistics ---
10 packets transmitted, 8 received, 20% packet loss, time 9012ms
```

Flujo de resolución de problemas

1. Determine qué dirección está causando las caídas intermitentes

Realizar una captura de paquetes (tcpdump, Wireshark, etc.) en el host de destino (EP B). Para ICMP, concéntrese en el número de secuencia para ver que los paquetes perdidos intermitentemente se observan en el EP B.

```
[admin@EP-B ~]$ tcpdump -ni eth0 icmp
11:32:26.540957 IP 10.1.1.1 > 10.1.2.1: ICMP echo request, id 3569, seq 1, length 64
11:32:26.681981 IP 10.1.2.1 > 10.1.1.1: ICMP echo reply, id 3569, seq 1, length 64
11:32:27.542175 IP 10.1.1.1 > 10.1.2.1: ICMP echo request, id 3569, seq 2, length 64
11:32:27.683078 IP 10.1.2.1 > 10.1.1.1: ICMP echo reply, id 3569, seq 2, length 64
11:32:28.543173 IP 10.1.1.1 > 10.1.2.1: ICMP echo request, id 3569, seq 3, length 64   <---
11:32:28.683851 IP 10.1.2.1 > 10.1.1.1: ICMP echo reply, id 3569, seq 3, length 64     <---
11:32:29.544931 IP 10.1.1.1 > 10.1.2.1: ICMP echo request, id 3569, seq 4, length 64
11:32:29.685783 IP 10.1.2.1 > 10.1.1.1: ICMP echo reply, id 3569, seq 4, length 64
11:32:30.546860 IP 10.1.1.1 > 10.1.2.1: ICMP echo request, id 3569, seq 5, length 64
...
```

- Patrón 1: todos los paquetes se observan en la captura de paquetes EP B.

Las caídas deben ser en respuesta de eco ICMP (EP B a EP A).

- Patrón 2 - Las caídas intermitentes se observan en la captura de paquetes EP B.

Las caídas deben ser en eco ICMP (EP A a EP B).

2. Verifique si otro protocolo con la misma IP de origen/destino tiene el mismo problema

Si es posible, intente probar la conectividad entre los dos terminales usando un protocolo diferente permitido por el contrato entre ellos (como ssh, telnet, http,..)

- Patrón 1 - Otros protocolos tienen la misma caída intermitente.

El problema podría estar en la inestabilidad del terminal o en la colocación en cola/almacenamiento en búfer, como se muestra a continuación.

- Patrón 2 - Solo el ICMP tiene la caída intermitente.

Las tablas de reenvío (como la tabla de terminales) no deben presentar ningún problema, ya que el reenvío se basa en MAC e IP. La colocación en cola/almacenamiento en búfer tampoco debería ser la razón, ya que esto afectaría a otros protocolos. La única razón por la que ACI tomaría una decisión de reenvío diferente basada en el protocolo sería el caso práctico de PBR.

Una posibilidad es que uno de los nodos de la columna tiene un problema. Cuando un protocolo es diferente, el paquete con el mismo origen y destino podría ser balanceado de carga a otro puerto de link ascendente/fabric (es decir, otra columna) por la hoja de ingreso.

Los contadores atómicos se pueden utilizar para garantizar que los paquetes no se descarten en los nodos de columna y lleguen a la hoja de salida. En caso de que los paquetes no alcancen la hoja de egreso, verifique el ELAM en la hoja de ingreso para ver qué puerto de fabric envían los paquetes. Para aislar el problema en una columna específica, se pueden cerrar los enlaces ascendentes de hoja para forzar el tráfico hacia otra columna.

3. Compruebe si está relacionado con un problema de aprendizaje de terminal

ACI utiliza una tabla de terminales para reenviar paquetes de un terminal a otro terminal. Un problema de disponibilidad intermitente puede ser causado por la inestabilidad del terminal porque la información inapropiada del terminal hará que el paquete se envíe a un destino incorrecto o que el contrato se descarte como clasificado en el EPG incorrecto. Incluso si el destino se supone que es un L3Out en lugar de un grupo de terminales, asegúrese de que la IP no se aprenda como un punto final en el mismo VRF a través de cualquier switch de hoja.

Consulte la subsección "Inestabilidad de terminales" de esta sección para obtener más información sobre cómo solucionar problemas de inestabilidad de terminales.

4. Compruebe si está relacionado con problemas de almacenamiento en búfer cambiando la frecuencia del tráfico

Aumente o disminuya el intervalo de ping para ver si cambia el ratio de descarte. La diferencia de intervalo debe ser lo suficientemente grande.

En Linux, la opción '-i' se puede utilizar para cambiar el intervalo (s):

```
[EP-A ~]$ ping 10.1.2.1 -c 10 -i 5     -- Increase it to 5 sec
[EP-A ~]$ ping 10.1.2.1 -c 10 -i 0.2   -- Decrease it to 0.2 msec
```

Si la proporción de caídas aumenta cuando disminuye el intervalo, probablemente esté relacionada con la colocación en cola o el almacenamiento en buffer de los terminales o switches.

La proporción de caídas a considerar es (número de caídas/total de paquetes enviados) en lugar de (número de caídas/tiempo).

En tal escenario, verifique lo siguiente.

1. Verifique si algún contador de caídas en las interfaces del switch está aumentando junto con el ping. Consulte la sección "Descartes de interfaz" en el capítulo "Reenvío dentro del fabric" para obtener más información.
2. Verifique si el contador Rx está aumentando junto con los paquetes en el punto final de destino. Si el contador Rx aumenta con el mismo número que los paquetes transmitidos, es probable que los paquetes se descarten en el terminal mismo. Esto podría deberse al almacenamiento en búfer del terminal en la pila TCP/IP.

Por ejemplo, si se envían 100000 pings con el intervalo más corto posible, se puede observar el contador Rx en el punto final a medida que aumenta 100000.

```
[EP-B ~]$ ifconfig eth0
eth0: flags=4163<UP,BROADCAST,RUNNING,MULTICAST>  mtu 1500
      inet 10.1.2.1  netmask 255.255.255.0  broadcast 10.1.2.255
      ether 00:00:10:01:01:02  txqueuelen 1000  (Ethernet)
      RX packets 101105  bytes 1829041
      RX errors 0  dropped 18926930  overruns 0  frame 0
      TX packets 2057  bytes 926192
      TX errors 0  dropped 0 overruns 0  carrier 0  collisions 0
```

5. Compruebe si ACI está enviando los paquetes o si el destino está recibiendo los paquetes

Realice una captura SPAN en el puerto de salida del switch de hoja para eliminar el fabric ACI de la ruta de solución de problemas.

Los contadores Rx en el destino también pueden ser útiles para eliminar los switches de red completos de la trayectoria de troubleshooting como se muestra en los pasos anteriores para el almacenamiento en memoria intermedia.

Inestabilidad de terminal

En esta sección se explica cómo comprobar si hay inestabilidad de terminales. Puede encontrar más información en estos documentos:

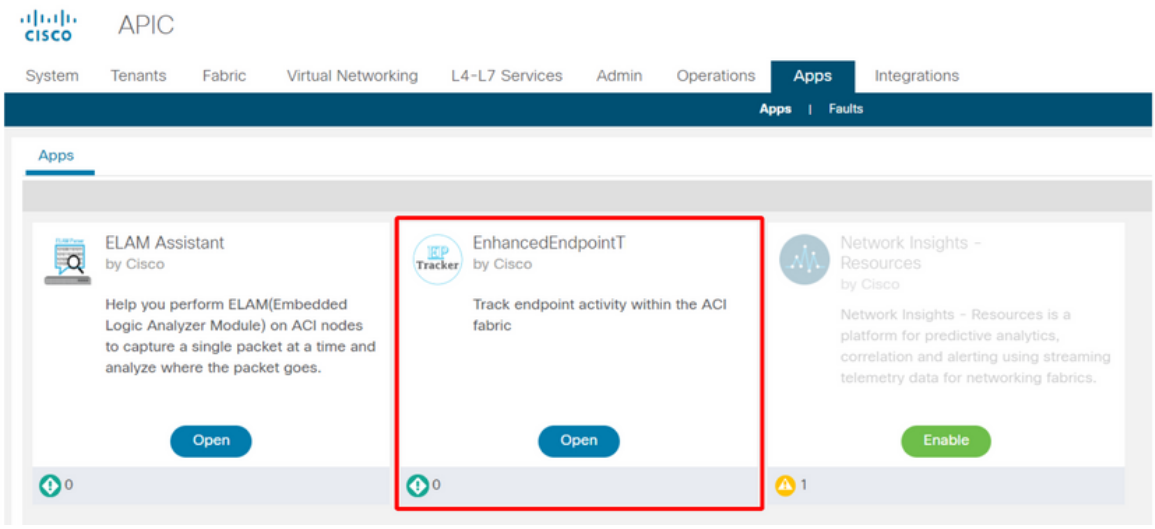
- "Informe técnico sobre aprendizaje de terminales de fabric de ACI" en www.cisco.com
- "Solución de problemas de ACI BRKACI-2641 de Cisco Live: terminales" en www.ciscolive.com

Cuando ACI aprende la misma dirección MAC o IP en varias ubicaciones, tendrá la misma apariencia de haberse movido el terminal. Esto también puede ser causado por un dispositivo de simulación o una configuración incorrecta. Este comportamiento se denomina inestabilidad de punto final. En tal escenario, el tráfico hacia el punto final en movimiento/inestable (dirección MAC para tráfico en puente, dirección IP para tráfico ruteado) fallará intermitentemente.

El método más eficaz para detectar el inestabilidad de los terminales es utilizar el rastreador de terminales mejorado. Esta aplicación se puede ejecutar como una aplicación de ACI AppCenter o como una aplicación independiente en un servidor externo en caso de que necesite administrar un fabric mucho más grande.

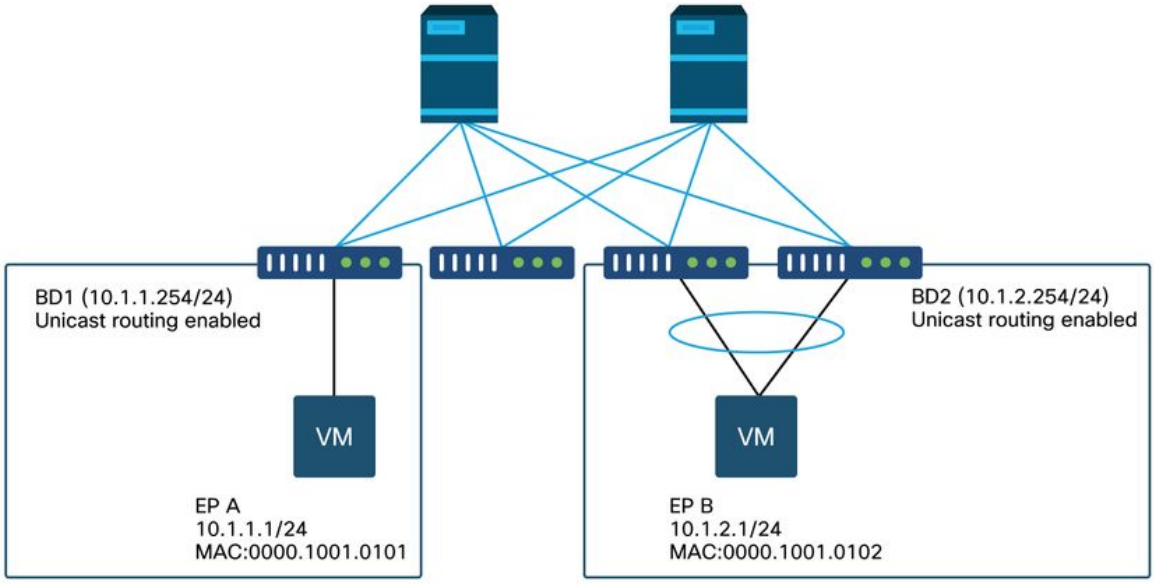
Seguimiento mejorado de terminales

¡ADVERTENCIA DE DEPRECIACIÓN! Esta guía fue escrita en 4.2; desde entonces, la aplicación Enhanced Endpoint Tracker ha quedado obsoleta en favor de la funcionalidad de Nexus Dashboard Insights. Para obtener más información, consulte Cisco bug ID CSCvz59365 .



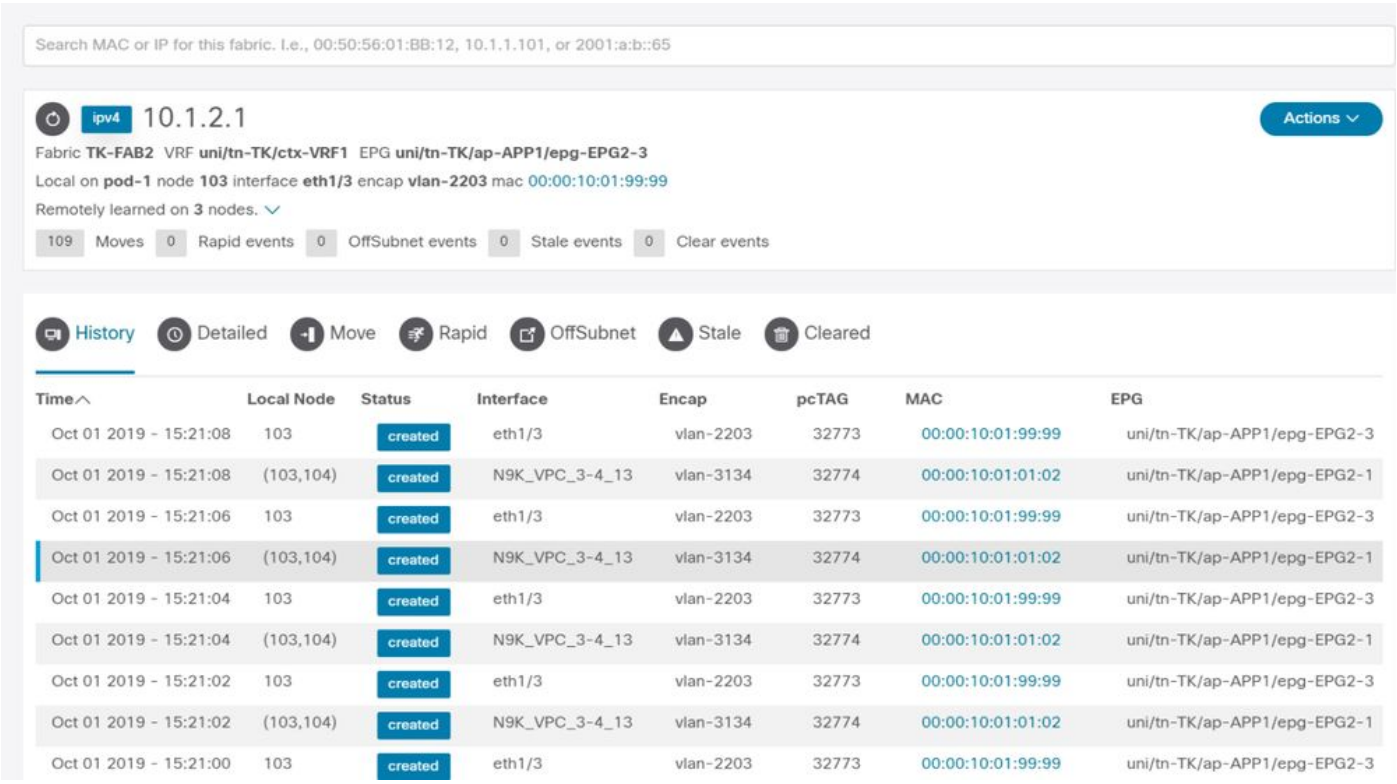
La imagen anterior muestra el rastreador de terminales mejorado en AppCenter. A continuación se muestra un ejemplo de cómo buscar extremos inestables con el Rastreador de extremos mejorados.

Ejemplo de inestabilidad de terminal



En este ejemplo, IP 10.1.2.1 debe pertenecer a EP B con MAC 0000.1001.0102. Sin embargo, un EP X con MAC 0000.1001.9999 también está suministrando tráfico con IP 10.1.2.1 debido a una configuración incorrecta o quizás a una suplantación de IP.

Resultado de Enhanced Endpoint Tracker: movimientos

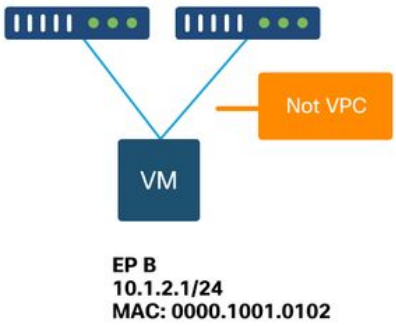


Enhanced Endpoint Tracker muestra cuándo y dónde se aprendió IP 10.1.2.1. Como se muestra en la captura de pantalla anterior, 10.1.2.1 está inestable entre dos terminales con MAC 0000.1001.0102 (esperado) y 0000.1001.9999 (no esperado). Esto causará un problema de alcance hacia IP 10.1.2.1 porque cuando se detecta en la dirección MAC incorrecta, el paquete se envía a un dispositivo incorrecto a través de la interfaz incorrecta. Para resolver esto, tome medidas para evitar que la máquina virtual inesperada obtenga tráfico con una dirección IP

inapropiada.

A continuación se muestra un ejemplo típico de inestabilidad de terminal debido a una configuración inapropiada.

Ejemplo de topología que podría provocar inestabilidad de terminales



Cuando un servidor o una VM se conecta a nodos de hoja de ACI a través de dos interfaces sin una VPC, el servidor necesita utilizar la agrupación NIC activa/en espera. De lo contrario, los paquetes tienen la carga equilibrada en ambos enlaces ascendentes y parecería que los puntos finales están oscilando entre dos interfaces desde la perspectiva del switch de hoja de ACI. En este caso, se requiere el modo de agrupamiento activo/en espera o NIC equivalente, o simplemente utilice un VPC en el lado de ACI.

Caídas de la interfaz

Este capítulo describe cómo verificar los contadores principales relacionados con la caída de la interfaz de ingreso.

Tipos de contadores de caídas de hardware

En los switches Nexus 9000 que se ejecutan en modo ACI, hay tres contadores de hardware principales en ACI para las caídas de interfaz de entrada.

Reenvío

Las principales razones de las caídas son:

- SECURITY_GROUP_DENY: Una baja por falta de contratos para permitir la comunicación.
- VLAN_XLATE_MISS: Una caída debido a una VLAN inapropiada. Por ejemplo, una trama ingresa al entramado con una VLAN 10 802.1Q. Si el switch tiene VLAN 10 en el puerto, inspeccionará el contenido y tomará una decisión de reenvío basada en la MAC de destino. Sin embargo, si VLAN 10 no está permitida en el puerto, lo descartará y lo etiquetará como VLAN_XLATE_MISS.
- ACL_DROP: Una caída debido a SUP-TCAM. SUP-TCAM en switches ACI contiene reglas especiales que se deben aplicar sobre la decisión de reenvío de L2/L3 normal. Las reglas de SUP-TCAM están integradas y el usuario no puede configurarlas. El objetivo de las reglas

SUP-TCAM es principalmente gestionar algunas excepciones o parte del tráfico del plano de control y no está pensado para que lo comprueben o supervisen los usuarios. Cuando un paquete está alcanzando las reglas SUP-TCAM y la regla es descartar el paquete, el paquete descartado se cuenta como ACL_DROP y aumentará el contador de descartes de reenvío.

Las caídas de reenvío son esencialmente paquetes caídos por una razón conocida válida. Por lo general, se pueden ignorar y no perjudicarán el rendimiento, a diferencia de las caídas reales del tráfico de datos.

Error

Cuando el switch recibe una trama no válida, se descarta como un error. Ejemplos de esto incluyen tramas con errores FCS o CRC. Consulte la sección posterior "CRC — FCS — switching por conexión directa" para obtener más información.

Buffer

Cuando un switch recibe una trama y no hay búferes disponibles para ingreso o egreso, la trama se descartará con 'Buffer'. Esto suele indicar congestión en algún punto de la red. El link que muestra la falla podría estar lleno o el link que contiene el destino está congestionado.

Recopilación de contadores mediante la API

Vale la pena señalar que al aprovechar la API y el modelo de objetos, el usuario puede consultar rápidamente el fabric para todas las instancias de estas caídas (ejecutarlas desde un apic) -

```
# FCS Errors (non-stomped CRC errors)
moquery -c rmonDot3Stats -f 'rmon.Dot3Stats.fCSErrors>="1"' | egrep "dn|fCSErrors"

# FCS + Stomped CRC Errors
moquery -c rmonEtherStats -f 'rmon.EtherStats.cRCAlignErrors>="1"' | egrep "dn|cRCAlignErrors"

# Output Buffer Drops
moquery -c rmonEgrCounters -f 'rmon.EgrCounters.bufferdroppkts>="1"' | egrep "dn|bufferdroppkts"

# Output Errors
moquery -c rmonIfOut -f 'rmon.IfOut.errors>="1"' | egrep "dn|errors"
```

Visualización de estadísticas de caídas en CLI

Si se observan fallas, o hay una necesidad de verificar las caídas de paquetes en las interfaces mediante la CLI, la mejor manera de hacerlo es ver los contadores de la plataforma en el hardware. No todos los contadores se muestran con 'show interface'. Las tres razones principales de la caída solo se pueden ver usando los contadores de la plataforma. Para verlos, realice estos pasos:

Hoja

SSH a la hoja y ejecute estos comandos. Este ejemplo es para Ethernet 1/31.

```
ACI-LEAF# vsh_lc
module-1# show platform internal counters port 31
```

```
Stats for port 31
(note: forward drops includes sup redirected packets too)
IF          LPort                    Input                        Output
                                Packets       Bytes          Packets    Bytes
eth-1/31    31  Total          400719   286628225      2302918    463380330
                Unicast     306610   269471065       453831     40294786
                Multicast        0           0      1849091    423087288
                Flood        56783     8427482            0            0
              Total Drops     37327                        0
                Buffer            0                        0
                Error             0                        0
                Forward      37327
                LB                0
                AFD RED                                    0
...
```

Columna

Se puede comprobar una columna fija (N9K-C9332C y N9K-C9364C) utilizando el mismo método que los interruptores de hoja.

Para una columna modular (N9K-C9504, etc.), la tarjeta de línea se debe conectar a antes de poder ver los contadores de la plataforma. SSH a la columna y ejecute estos comandos. Este ejemplo es para Ethernet 2/1.

```
ACI-SPINE# vsh
ACI-SPINE# attach module 2
module-2# show platform internal counters port 1
Stats for port 1
(note: forward drops include sup redirected packets too)
IF          LPort                    Input                        Output
                                Packets       Bytes          Packets    Bytes
eth-2/1      1  Total        85632884  32811563575     126611414   25868913406
                Unicast   81449096  32273734109     104024872   23037696345
                Multicast  3759719   487617769     22586542   2831217061
                Flood            0           0            0            0
              Total Drops        0                        0
                Buffer           0                        0
                Error            0                        0
                Forward          0
                LB               0
                AFD RED                                    0
...
```

Los contadores de estadísticas de colas se muestran mediante 'show queuing interface'. Este ejemplo es para Ethernet 1/5.

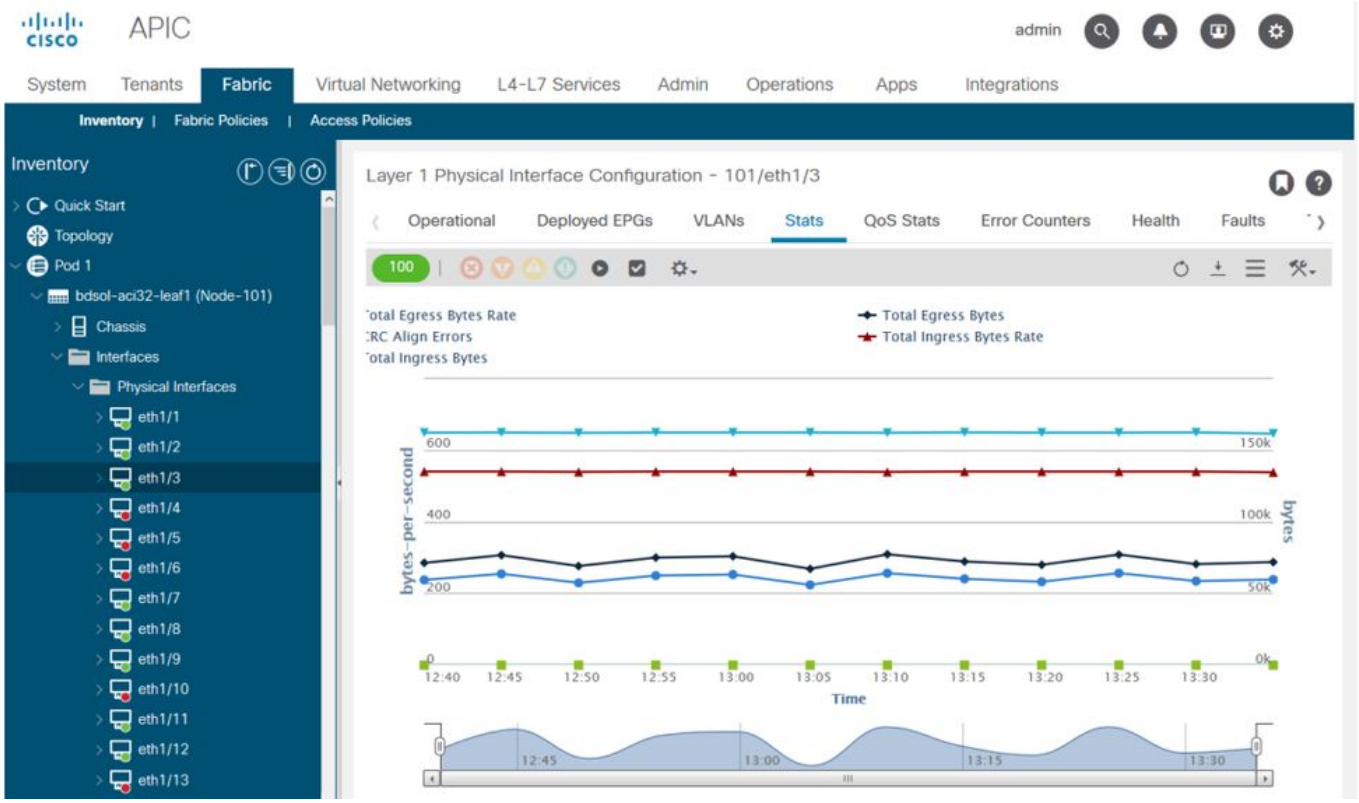
```
ACI-LEAF# show queuing interface ethernet 1/5
================================================================================
               Queuing stats for ethernet 1/5
================================================================================
================================================================================
               Qos Class level1
================================================================================
Rx Admit Pkts :  0                    Tx Admit Pkts :  0
Rx Admit Bytes:  0                    Tx Admit Bytes:  0
Rx Drop Pkts  :  0                    Tx Drop Pkts  :  0
Rx Drop Bytes :  0                    Tx Drop Bytes :  0

================================================================================
```

```
                Qos Class level2
================================================================================
Rx Admit Pkts :  0                      Tx Admit Pkts :  0
Rx Admit Bytes:  0                      Tx Admit Bytes:  0
Rx Drop Pkts  :  0                      Tx Drop Pkts  :  0
Rx Drop Bytes :  0                      Tx Drop Bytes :  0

================================================================================
                Qos Class level3
================================================================================
Rx Admit Pkts :  1756121                Tx Admit Pkts :  904909
Rx Admit Bytes:  186146554              Tx Admit Bytes:  80417455
Rx Drop Pkts  :  0                      Tx Drop Pkts  :  22
Rx Drop Bytes :  0                      Tx Drop Bytes :  3776

...
```

Visualización de estadísticas en la GUI

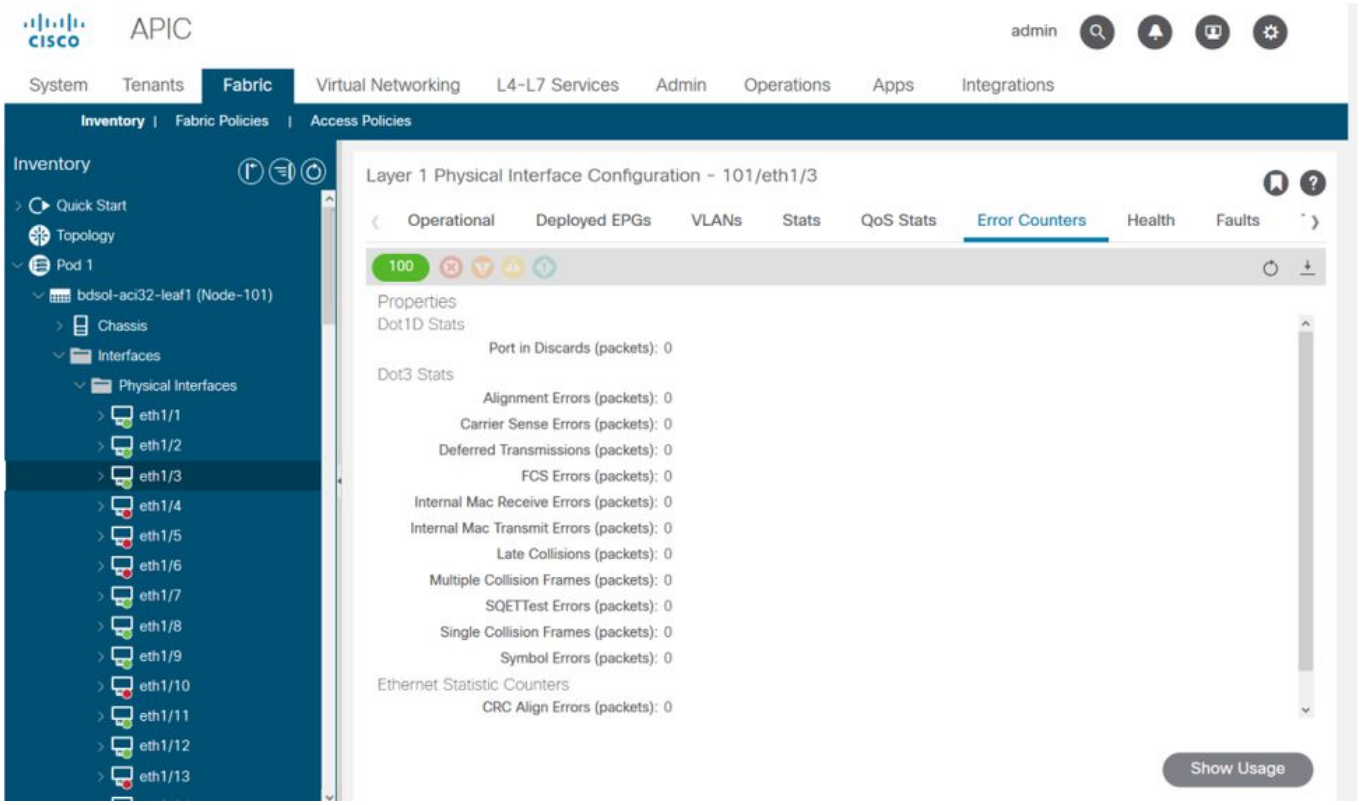
La ubicación es 'Fabric > Inventory > Leaf/Spine > Physical interface > Stats'.

estadísticas de interfaz GUI



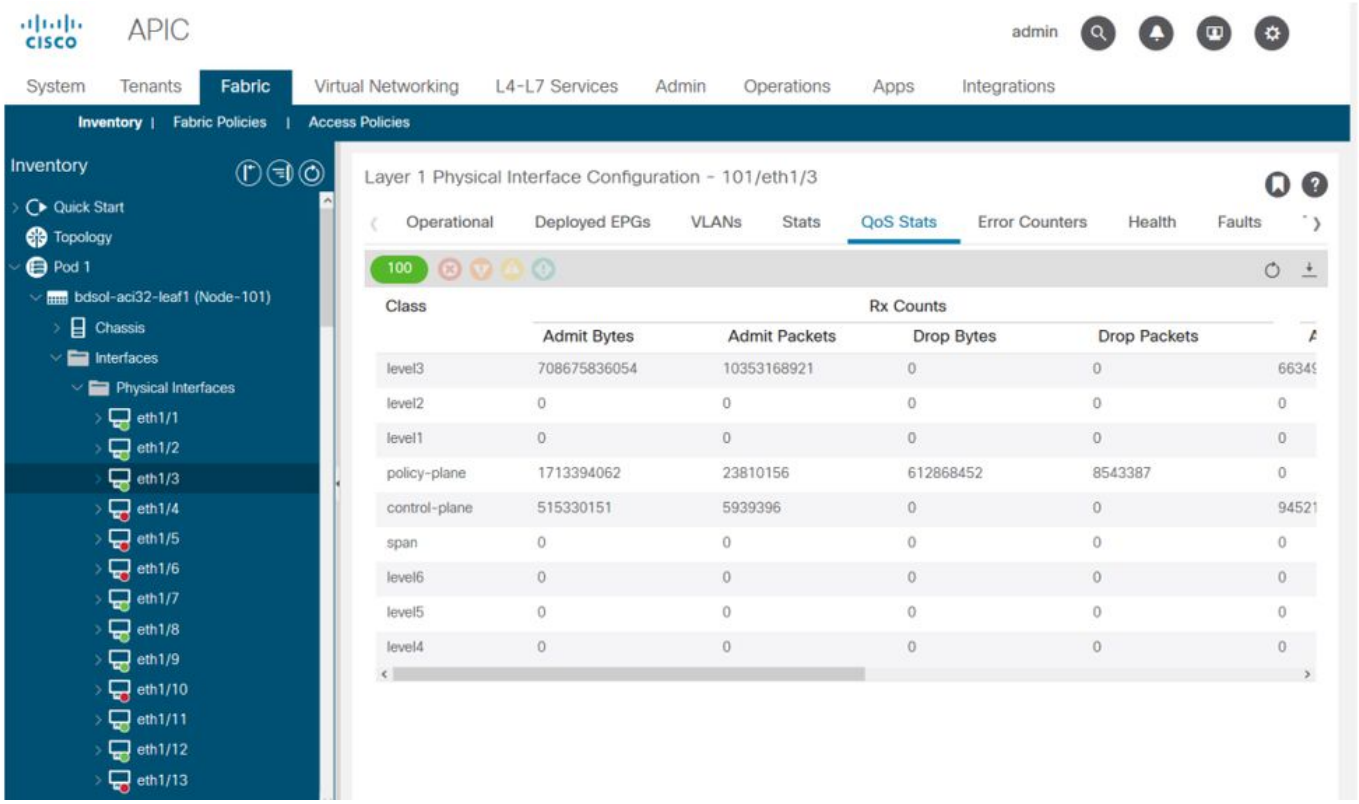
Las estadísticas de error se pueden ver en el mismo lugar:

Errores de interfaz GUI



Por último, la GUI puede mostrar estadísticas de QoS por interfaz:

Contadores QoS de interfaz GUI



CRC — FCS — Switching de conexión directa

¿Qué es la verificación por redundancia cíclica (CRC)?

CRC es una función polinómica en la trama que devuelve un número 4B en Ethernet. Capturará todos los errores de bit simple y un buen porcentaje de errores de bit doble. Por lo tanto, se pretende garantizar que la trama no se dañó en tránsito. Si el contador de errores CRC está aumentando, significa que cuando el hardware ejecutó la función polinómica en la trama, el resultado fue un número 4B que difirió del número 4B encontrado en la trama misma. Las tramas pueden dañarse debido a varias razones, como discordancia dúplex, cableado defectuoso y hardware dañado. Sin embargo, se debe esperar cierto nivel de errores CRC y el estándar permite una tasa de error de hasta 10-12 bits en Ethernet (1 bit de 1012 puede voltearse).

Switching con almacenamiento y retransmisión frente a switching con conexión directa

Tanto los switches de almacenamiento y reenvío como los switches de capa 2 con conexión directa basan sus decisiones de reenvío en la dirección MAC de destino de los paquetes de datos. También aprenden las direcciones MAC mientras examinan los campos MAC de origen (SMAC) de los paquetes a medida que las estaciones se comunican con otros nodos de la red.

Un switch de almacenamiento y reenvío toma una decisión de reenvío en un paquete de datos después de haber recibido la trama completa y comprobado su integridad. Un switch de conexión directa se involucra en el proceso de reenvío poco después de haber examinado la dirección MAC de destino (DMAC) de una trama entrante. Sin embargo, un switch de conexión directa debe esperar hasta que haya visto todo el paquete antes de realizar la verificación CRC. Esto significa que para el momento en que se valida la CRC, el paquete ya ha sido reenviado y no puede ser descartado si falla la verificación.

Tradicionalmente, la mayoría de los dispositivos de red solían funcionar con almacenamiento y retransmisión. Las tecnologías de switching por conexión directa tienden a utilizarse en redes de alta velocidad que exigen un reenvío de baja latencia.

Específicamente, con respecto al hardware ACI de la generación 2 y posteriores, se realiza un switching por conexión directa si la interfaz de ingreso es de mayor velocidad y la interfaz de egreso es de la misma velocidad o de menor velocidad. La conmutación con almacenamiento y retransmisión se realiza si la velocidad de la interfaz de ingreso es menor que la de egreso.

Pisoteo

Los paquetes con un error CRC necesitan descartarse. Si la trama se conmuta en una trayectoria de conexión directa, la validación de CRC ocurre después de que el paquete ya se reenvía. Como tal, la única opción es pisotear la secuencia de verificación de tramas Ethernet (FCS). **Para pisar una trama se configura FCS a un valor conocido que no pasa una verificación CRC**. Debido a esto, una trama incorrecta que falla CRC podría aparecer como CRC en cada interfaz que atraviesa, hasta que alcance un switch de almacenamiento y reenvío que la descartará.

ACI y CRC: buscar interfaces defectuosas

- Si una hoja ve errores CRC en un puerto de link descendente, es principalmente un problema en el SFP de link descendente o con componentes en el dispositivo/red externo.
- Si una columna ve errores CRC, es principalmente un problema en ese puerto local, SFP, fibra o SFP vecino. Los paquetes CRC que fallan de los links descendentes de la hoja no son pisoteados a las columnas. Como si sus encabezados fueran legibles, está encapsulado VXLAN y se calculará el nuevo CRC. Si los encabezados no fueran legibles debido a la

corrupción de tramas, el paquete se descartaría.
- Si una hoja ve errores CRC en links de estructura, puede ser: Un problema en el par SFP/fibra local, la fibra de ingreso de la columna o el par SFP.Un marco pisado que se abre camino a través de la tela.

Detener: Troubleshooting Stomping

- Busque interfaces con errores FCS en el fabric. Dado que FCS se produce localmente en un puerto, lo más probable es que sea la fibra o SFP en cualquiera de los extremos.
- Los errores CRC en la salida 'show interface' reflejan el valor total de FCS+Stomp.\

Vea un ejemplo:

Verifique un puerto con el comando

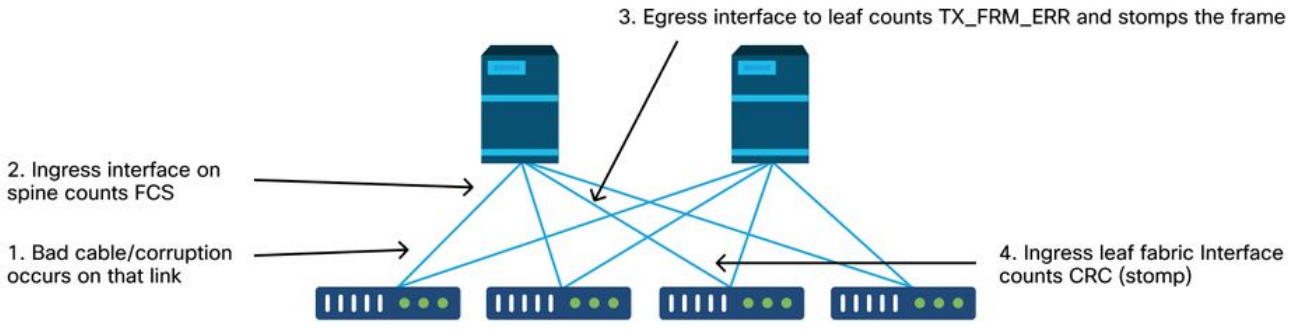
```
vsh_lc: 'show platform internal counter port <X>'
```

En este comando, son importantes 3 valores:

- RX_FCS_ERR: error de FCS.
- RX_CRCERR: trama de error CRC pisoteada recibida.
- TX_FRM_ERROR - Trama de error CRC pisada transmitida.

```
module-1# show platform internal counters port 1 | egrep ERR
       RX_FCS_ERR           0     ---- Real error local between the devices and its direct
neighbor
       RX_CRCERR            0     ---- Stomped frame --- so likely stomped by underlying devices
and generated further down the network
       TX_FRM_ERROR         0     ---- Packet received from another interface that was stomp on
Tx direction
```

escenario de Troubleshooting de CRC Stop



Si un link dañado genera una gran cantidad de tramas dañadas, esas tramas podrían inundarse a todos los demás nodos de hoja y es muy posible encontrar CRC en el ingreso de los enlaces ascendentes de fabric de la mayoría de los nodos de hoja en el fabric. Es probable que todos ellos provengan de un único enlace dañado.

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).