

Comprensión de las Entradas de Catalyst 6500 S2T CEF

Contenido

[Introducción](#)

[Prerequisites](#)

[Requirements](#)

[Componentes Utilizados](#)

[Antecedentes](#)

[Diagrama de la red](#)

[Identificación de Entradas CEF en Motores de Reenvío Distribuidos](#)

[Eliminar entradas CEF](#)

[Agregar una entrada CEF](#)

[Agregar y Eliminar Entradas para Tablas de Ruteo VRF](#)

Introducción

Este documento describe cómo Cisco Catalyst 6500 con Supervisor Sup2T programa las entradas CEF (Cisco Express Forwarding) configuradas en el software Cisco IOS en el hardware de las tarjetas de línea utilizado para lograr el reenvío de paquetes.

Prerequisites

Requirements

Cisco recomienda que tenga conocimiento sobre estos temas:

- Reenvío express de Cisco (CEF)
- Cisco Catalyst 6500 Series Switches
- Tarjeta de reenvío distribuido (DFC) de Cisco

Componentes Utilizados

La información que contiene este documento se basa en estas versiones de software y hardware.

- Tarjeta de línea Cisco Catalyst 6500 WS-X6848-GE-TX (con DFC4).
- Cisco Catalyst 6500 con Supervisor 2T en IOS versión 15.2.1SY5

The information in this document was created from the devices in a specific lab environment. All of the devices used in this document started with a cleared (default) configuration. Si tiene una red en vivo, asegúrese de entender el posible impacto de cualquier comando.

Antecedentes

La mayoría de los switches multicapa de Cisco utilizan CEF como mecanismo de switching de Capa 3. Es imperativo que los ingenieros de redes entiendan cómo funciona CEF para resolver las interrupciones de red, la pérdida de paquetes o los escenarios de demora de paquetes a diario.

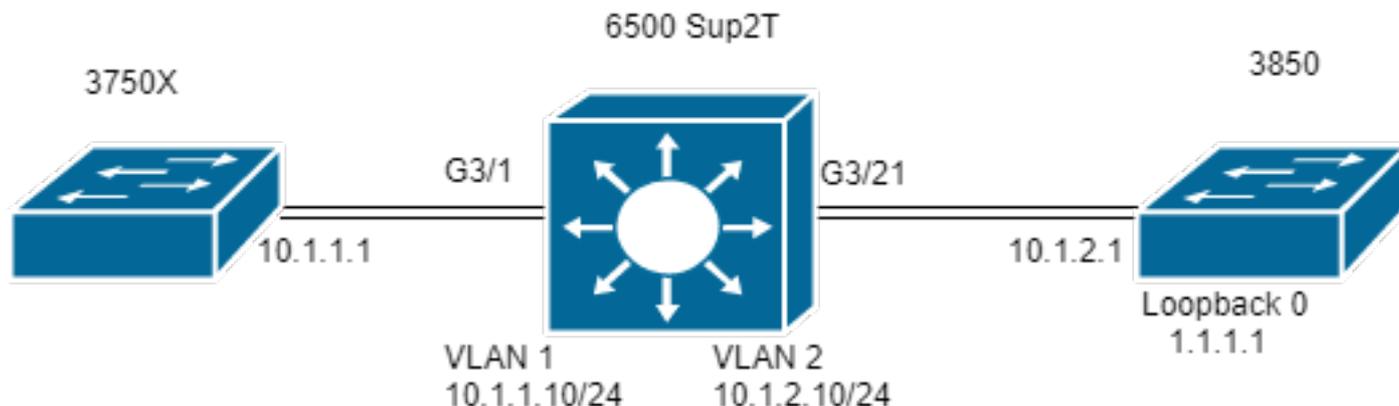
El supervisor del motor supervisor 2T en modo independiente o, como VSS, actualmente lo implementan muchas redes empresariales como switch principal, agrega prácticamente todos los demás dispositivos de routing o switching. Esto también significa que reenvía la mayor parte del tráfico dentro y entre dominios para entregar con éxito los paquetes a sus destinos. Para lograr esto, el motor supervisor 2T debe tener información de ruteo adecuada aprendida estática o dinámicamente a través de los protocolos de ruteo.

En un chasis modular, pueden existir varios motores de reenvío además de los del supervisor. Ciertas tarjetas de línea (especialmente las de nueva generación como C6800-32P10G) ya incluyen su propio motor de reenvío para mejorar el rendimiento de conmutación de paquetes, la búsqueda de entradas CEF se ejecuta localmente y hace que los recursos se distribuyan mejor para el tráfico que ingresa en diferentes tarjetas de línea. Éstas se conocen como Tarjetas de reenvío distribuidas (DFC).

Estas entradas de CEF compartidas en todos los motores de reenvío pueden no estar asignadas en HW por varias razones, desde una condición de defecto de software, agotamiento de recursos a condiciones de CPU altas y evita que el switch tenga tiempo suficiente para actualizar todas las entradas, esto puede causar una serie de eventos indeseables.

Diagrama de la red

Red:



```
Switch#show module 3
```

```
----- Mod Ports Card Type Model Serial No. --- --
-----
10/100/1000mb Ethernet WS-X6848-GE-TX SAL2003X5AH -----
----- 3 Distributed Forwarding Card WS-F6K-DFC4-A SAL2003X5AH 1.4 Ok
```

Identificación de Entradas CEF en Motores de Reenvío Distribuidos

En el diagrama, un switch 6506 independiente tiene un Supervisor 2T instalado, así como una tarjeta de línea WS-6848-GE-TX con un DFC en la ranura 3. El host 3750X, que está conectado a la tarjeta de línea a través del puerto G3/1, envía tráfico a la dirección 1.1.1.1 del Loopback 0 de 3850.

Para esto, 3750X tiene una ruta estática a la dirección IP 1.1.1.1 al siguiente salto 10.1.1.10 que es la SVI de VLAN 1 en el switch Sup2T. El switch Sup2T necesita rutear este tráfico al switch 3850 basado en una entrada de ruta estática para IP 1.1.1.1/32 a través del salto siguiente 10.1.2.1 que es la interfaz 3850 conectada al Sup2T en VLAN 2.

```
MXC.CALO.3750X#show ip route | inc 1.1.1.1
S 1.1.1.1 [1/0] via 10.1.1.10
```

```
MXC.CALO.Sup2T#show ip route | inc 1.1.1.1
S 1.1.1.1 [1/0] via 10.1.2.1
```

```
CALO.MXC.3850#show ip route | inc 1.1.1.1
C 1.1.1.1 is directly connected, Loopback1
```

Tenga en cuenta que, en aras de la simplicidad, los switches 3750X y 3850 están conectados al 6500 a través de la misma tarjeta de línea. Esto significa que el tráfico también se busca de forma local y local.

Un paquete ingresa el switch Sup2T a través de Gi3/1, finalmente alcanza el motor de reenvío (ya que es un DFC). El motor de reenvío analiza el campo de dirección IP de destino en este paquete y una búsqueda sobre las entradas CEF programadas para la mejor coincidencia (Máscara más larga).

Dado que se trata de una tarjeta DFC, significa que tiene sus propias entradas CEF y para verificarlas, es necesario que se conecte a la tarjeta de línea con

conexión de comandos [dec] o que se adjunte el switch [1-2] mod [dec] para VSS.

Ahora, debe estar en la indicación DFC, el comando **show platform hardware cef** o **show platform hardware cef vpn 0** devuelven todas las entradas CEF programadas para la tabla de ruteo general (VPN 0/ No VRF).

Dado que el objetivo es el prefijo 1.1.1.1/32 usted utiliza el comando **show platform hardware cef vpn 0 lookup 1.1.1.1**. El comando devuelve la mejor coincidencia para el prefijo 1.1.1.1 y el que utiliza para reenviar realmente el tráfico:

```
MXC.CALO.Sup2T#attach 3
Trying Switch ...
Entering CONSOLE for Switch
Type "^C^C^C" to end this session
```

```
MXC.CALO.Sup2T-dfc3#show platform hardware cef vpn 0
Codes: decap - Decapsulation, + - Push Label
Index Prefix Adjacency
32 0.0.0.0/32 receive
33 255.255.255.255/32 receive
34 10.1.85.254/32 glean
35 10.1.85.5/32 receive
36 10.1.86.5/32 receive
[snip...]
```

```
MXC.CALO.Sup2T-dfc3#show platform hardware cef vpn 0 lookup 1.1.1.1
Codes: decap - Decapsulation, + - Push Label
Index Prefix Adjacency
262 1.1.1.1/32 V12 ,0c11.678b.f6f7
```

La entrada CEF está ahí, se programó como resultado de nuestra entrada estática programada en el software IOS a través del comando **ip route 1.1.1.1 255.255.255.255 10.1.2.1**.

También puede verificar que esta entrada obtiene resultados y el tráfico se reenvía con esta entrada a través de los comandos **show platform hardware cef 1.1.1.1 detail** que devuelve una entrada de adyacencia:

```
MXC.CALO.Sup2T-dfc3#show platform hardware cef 1.1.1.1 detail
Codes: M - mask entry, V - value entry, A - adjacency index, NR- no_route bit
LS - load sharing count, RI - router_ip bit, DF: default bit
CP - copy_to_cpu bit, AS: dest_AS_number, DGTv - dgt_valid bit
DGT: dgt/others value
```

```
Format:IPV4 (valid class vpn prefix)
M(262 ): 1 F 2FFF 255.255.255.255
V(262 ): 1 0 0 1.1.1.1
(A:114689, LS:0, NR:0, RI:0, DF:0 CP:0 DGTv:1, DGT:0)
```

Finalmente, la entrada de adyacencia muestra cómo se reescribe el paquete y si el tráfico es reescrito por esta entrada de adyacencia:

```
MXC.CALO.Sup2T-dfc3#show platform hardware cef adjacencies entry 114689 detail
```

```
RIT fields: The entry has a Layer2 Format
```

```
-----
|decr_ttl = YES | pipe_ttl = 0 | utos = 0
|-----|-----|-----
|l2_fwd = 0 | rmac = 0 | ccc = L3_REWRITE
|-----|-----|-----
|rm_null_lbl = YES| rm_last_lbl = YES| pv = 0
|-----|-----|-----
|add_shim_hdr= NO | rec_findex = N/A | rec_shim_op = N/A
|-----|-----|-----
|rec_dti_type = N/A | rec_data = N/A
|-----|-----|-----
```

```

|modify_smac = YES| modify_dmac = YES| egress_mcast = NO
|_____|_____
|ip_to_mac = NO
|_____
|dest_mac = 0c11.678b.f6f7 | src_mac = d8b1.902c.9680
|_____|_____
|
Statistics: Packets = 642
Bytes = 75756 <<<<

```

Los **dest_mac** y **src_mac** son los valores de interés principal, que indican los nuevos encabezados L2 que se escriben para este paquete. La dirección MAC de destino 0c11.678b.f6f7 es 10.1.2.1 que es el 3850 (salto siguiente para alcanzar 1.1.1.1):

```

MXC.CALO.Sup2T#show ip arp 10.1.2.1
Protocol Address Age (min) Hardware Addr Type Interface
Internet 10.1.2.1 30 0c11.678b.f6f7 ARPA Vlan2

```

Además, el campo **Statistics** muestra que el tráfico realmente llega a esta entrada de adyacencia y los encabezados L2 se reescriben en consecuencia.

Eliminar entradas CEF

La eliminación de entradas CEF puede ayudarnos a eliminar cualquier entrada que se pueda programar erróneamente (por ejemplo, a una entrada de adyacencia incorrecta) o incluso con fines de formación. También proporciona una manera de modificar una trayectoria de ruteo.

Para eliminar una entrada CEF, debe comprender que las entradas CEF se programan secuencialmente y que se asigna un índice de hardware, por ejemplo:

```

MXC.CALO.Sup2T-dfc3#show platform hardware cef vpn 0

```

Códigos: decap - Decapsulation, + - Push Label

```

MXC.CALO.Sup2T-dfc3#show platform hardware cef vpn 0
...
Index Prefix Adjacency 259 10.1.2.255/32 receive 260 10.1.1.1/32 V11 ,a0ec.f930.3f40 261
10.1.2.1/32 V12 ,0c11.678b.f6f7 262 1.1.1.1/32 V12 ,0c11.678b.f6f7 <<<< Our CEF entry of
interest has a HW index of 262.
...

```

Este índice de hardware es el elemento más importante para eliminar una entrada CEF ya que se utiliza como referencia. Sin embargo, para realizar cualquier cambio en él, debe convertirse en un identificador de software. Puede lograr esto con el comando **test platform hardware cef index-conv hw_to_sw [hw index]**

```

MXC.CALO.Sup2T-dfc3#test platform hardware cef index-conv hw_to_sw 262
hw index: 262 ----> sw handle: 101
Ahora que conoce el identificador del software, puede continuar con la eliminación de la entrada CEF con el comando test platform hardware cef v4-delete [sw handle] mask [mask length] vpn [dec]

```

```

MXC.CALO.s2TVSS-sw2-dfc3#test platform hardware cef v4-delete 101 mask 32 vpn 0
test_ipv4_delete: done.

```

Nota: El valor de longitud de la máscara es 32, ya que se trata de una ruta específica del host (1.1.1.1/32)

Ahora, se elimina nuestra entrada CEF:

```

MXC.CALO.Sup2T-dfc3#show platform hard cef vpn 0 1.1.1.1
Codes: decap - Decapsulation, + - Push Label
Index Prefix Adjacency

```

```

MXC.CALO.Sup2T-dfc3#show platform hard cef vpn 0
[snip...]
259 10.1.2.255/32 receive
260 10.1.1.1/32 V11 ,a0ec.f930.3f40
261 10.1.2.1/32 V12 ,0c11.678b.f6f7
288 224.0.0.0/24 receive <<<<<<< Index 262 no longer exists in the CEF entries.
289 10.1.85.0/24 glean

```

Observe que el comando **test platform hardware cef vpn 0** se ejecutó bajo el indicador DFC. De esta manera, la entrada CEF se quitó de la tabla CEF de DFC y NO del supervisor, debe tener mucho cuidado en qué motor de reenvío se quitan las entradas.

Un cambio en el tráfico tiene el riesgo de no tener visibilidad (en el caso de una prueba de laboratorio), esto puede deberse al resultado de otra entrada CEF. Considere siempre coincidir con el más exacto (máscara más larga). En este laboratorio, se produce el siguiente impacto:

```

MXC.CALO.Sup2T-dfc3#show plat hard cef vpn 0 lookup 1.1.1.1
Codes: decap - Decapsulation, + - Push Label
Index Prefix Adjacency
262048 0.0.0.0/0 glean

```

Entonces, ¿qué hace realmente esta entrada con el paquete?:

```

MXC.CALO.Sup2T-dfc3#show platform hardware cef adjacencies entry 262048
RIT fields: The entry has a Recirc. Format
-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
add_shim_hdr = YES |-----|-----|-----|-----|-----|rc_fidrx=0 |
rc_shimop=1 | rc_dti_type=4 | rc_data = 0x10B
|-----|-----|-----|-----|-----|-----|-----|-----|-----|
Statistics: Packets = 2163 Bytes =
255234

```

Taken from a CPU packet capture using Catlayst 6500 NETDR tool. For NETDR capture tool details refer to: [Catalyst 6500 Series Switches Netdr Tool for CPU-Bound Packet Captures](#)

----- dump of incoming inband packet -----

```

l2idb Po1, l3idb V11, routine inband_process_rx_packet, timestamp 01:00:17.841
dbus info: src_vlan 0x1(1), src_indx 0xB40(2880), len 0x82(130)
bpdu 0, index_dir 0, flood 0, dont_lrn 0, dest_indx 0x5FA4(24484), CoS 0
cap1 0, cap2 0
78020800 00018400 0B400100 82000000 1E000464 2E000004 00000010 5FA45BDD
destmac D8.B1.90.2C.96.80, srcmac A0.EC.F9.30.3F.40, shim ethertype CCF0
earl 8 shim header IS present:
version 0, control 64(0x40), lif 1(0x1), mark_enable 1,
feature_index 0, group_id 0(0x0), acos 0(0x0),
ttl 14, dti 4, dti_value 267(0x10B)
10000028 00038080 010B
ethertype 0800
protocol ip: version 0x04, hlen 0x05, tos 0x00, totlen 100, identifier 51573
df 0, mf 0, fo 0, ttl 255, src 10.1.1.1, dst 1.1.1.1
icmp type 8, code 0

```

----- dump of outgoing inband packet -----

```

l2idb NULL, l3idb V12, routine etsec_tx_pak, timestamp 01:03:56.989
dbus info: src_vlan 0x2(2), src_indx 0x380(896), len 0x82(130)
bpdu 0, index_dir 0, flood 0, dont_lrn 0, dest_indx 0x0(0), CoS 0
cap1 0, cap2 0
00020000 0002A800 03800000 82000000 00000000 00000000 00000000 00000000
destmac 0C.11.67.8B.F6.F7, srcmac D8.B1.90.2C.96.80, shim ethertype CCF0
earl 8 shim header IS present:
version 0, control 0(0x0), lif 16391(0x4007), mark_enable 0,
feature_index 0, group_id 0(0x0), acos 0(0x0),
ttl 15, dti 0, dti_value 540674(0x84002)

```

```
000800E0 0003C008 4002
ethertype 0800
protocol ip: version 0x04, hlen 0x05, tos 0x00, totlen 100, identifier 50407
df 0, mf 0, fo 0, ttl 254, src 10.1.1.1, dst 1.1.1.1
icmp type 8, code 0
```

Ahora, todo el tráfico con el destino de 1.1.1.1 que ingresa a través de la tarjeta de línea 3 se recircula con el encabezado shim y se envía a la CPU. A veces, en lugar de esta entrada CEF, se ve otra 0.0.0.0/0 con adyacencia **drop** y hace exactamente lo mismo.

Nota: Evalúe qué entradas CEF se quitan. Debido a esto, puede producirse un uso elevado de la CPU. Normalmente, se configura una ruta predeterminada 0.0.0.0/0 y el tráfico se reenvía en función de ella (y provoca la pérdida de paquetes).

Agregar una entrada CEF

Cuando se agrega una entrada CEF, en la mayoría de los casos resuelve cualquier problema de mala programación que cause pérdida de paquetes, retraso de paquetes o alta utilización de la CPU. Conocimiento de cómo instalar las entradas CEF en el hardware, no solo proporciona la capacidad de corregir una entrada mal programada, sino también de manipular cualquier reenvío de paquetes a través de recirculación del paquete, señalarlo a una interfaz completamente diferente o salto siguiente, reescribir un paquete ruteado como se desee y/o descartarlo, etc. Todo esto, sin una recarga de la caja, quitar y establecer la configuración o cualquier modificación aparente. La entrada CEF la adición se puede realizar sin entrar también en el modo de configuración. (Como también lo hizo con el procedimiento de eliminación de entradas CEF explicado en la sección anterior).

Básicamente, hay dos situaciones aquí, cuando usted tiene una entrada ARP válida al salto siguiente, en este caso 10.1.2.1 y cuando no lo hace (por cualquier razón). La segunda situación obliga a crear una entrada ARP válida (mediante ARP estático):

Paso 1. Hay una entrada ARP en el switch para 10.1.2.1 que es el salto siguiente para 1.1.1.1.

```
MXC.CALO.Sup2T#show ip arp 10.1.2.1
Protocol Address Age (min) Hardware Addr Type Interface
Internet 10.1.2.1 2 0c11.678b.f6f7 ARPA Vlan2
```

```
MXC.CALO.Sup2T#show ip route | inc 1.1.1.1
S 1.1.1.1 [1/0] via 10.1.2.1
Una entrada ARP se programa como una ruta de host ( /32 ) en la tabla CEF:
```

```
MXC.CALO.Sup2T-dfc3#show plat hard cef vpn 0 look 10.1.2.1
Codes: decap - Decapsulation, + - Push Label
Index Prefix Adjacency
53 10.1.2.1/32 V12 ,0c11.678b.f6f7
```

And of course, there is an index for this which again will tell us how a packet should be rewritten to reach 10.1.2.1:

```
MXC.CALO.Sup2T-sw2-dfc3#show plat hard cef vpn 0 10.1.2.1 detail
[snip...]
Format:IPV4 (valid class vpn prefix)
M(53 ): 1 F 2FFF 255.255.255.255
V(53 ): 1 0 0 10.1.2.1
(A:114689, LS:0, NR:0, RI:0, DF:0 CP:0 DGTv:1, DGT:0)
```

Wait, wasn't 114689 adj entry the same used for 1.1.1.1?:

```
MXC.CALO.Sup2T-sw2-dfc3#show plat hard cef 1.1.1.1 de
[snip...]
Format:IPV4 (valid class vpn prefix)
M(54 ): 1 F 2FFF 255.255.255.255
V(54 ): 1 0 0 1.1.1.1
(A:114689, LS:0, NR:0, RI:0, DF:0 CP:0 DGTv:1, DGT:0)
```



```
Statistics: Packets = 684
Bytes = 80712
```

```
// Logs in 3850
```

```
CALO.MXC.385024XU#show logging [snip...] *Jan 23 05:59:56.911: ICMP: echo reply sent, src
1.1.1.1, dst 10.1.1.1, topology BASE, dscp 0 topoid 0 *Jan 23 05:59:57.378: ICMP: echo reply
sent, src 1.1.1.1, dst 10.1.1.1, topology BASE, dscp 0 topoid 0 *Jan 23 05:59:57.390: ICMP: echo
reply sent, src 1.1.1.1, dst 10.1.1.1, topology BASE, dscp 0 topoid 0
```

Agregar y Eliminar Entradas para Tablas de Ruteo VRF

A lo largo de la configuración realizada a partir de todos los pasos anteriores, se ha aplicado la cadena **vpn 0** en los comandos **show platform hardware cef**. Incluso si parece completamente innecesario ya que el comando de forma predeterminada devuelve las entradas para la tabla de ruteo general o **vpn 0**, esto se hizo a propósito para siempre tenga en cuenta que las entradas se agregan o eliminan de las instancias de tabla de routing (VRF) específicas, a través del documento que agregó y eliminó la entrada CEF 1.1.1.1/32. Sin embargo, es muy probable que ciertos prefijos existan en diferentes VRF (i. e. 10.x.x.x) y eliminar, agregar o modificar una entrada CEF para un VRF incorrecto puede causar un impacto negativo.

Elimine una entrada CEF con el prefijo 1.1.1.1/32 para VRF **TEST_VRF**. Para obtener una descripción detallada de la adición de entradas CEF, refiérase a la sección **Agregar una Entrada CEF** de este documento.

Para agregar el VRF, cambie las SVI en el switch 6500 al VRF propuesto con el comando **ip vrf forwarding [VRF-NAME]** y finalmente agregue la misma ruta estática en nuestra tabla **TEST_VRF**:

```
MXC.CALO.Sup2T(config)#ip vrf TEST_VRF
MXC.CALO.Sup2T(config-vrf)#int vlan 1
MXC.CALO.Sup2T(config-if)#ip vrf forwarding TEST_VRF
% Interface Vlan1 IPv4 disabled and address(es) removed due to enabling VRF TEST_VRF
MXC.CALO.Sup2T(config-if)#ip add 10.1.1.10 255.255.255.0
MXC.CALO.Sup2T(config-if)#int vlan 2
MXC.CALO.Sup2T(config-if)#ip vrf forwarding TEST_VRF
% Interface Vlan2 IPv4 disabled and address(es) removed due to enabling VRF TEST_VRF
MXC.CALO.Sup2T(config-if)#ip add 10.1.2.10 255.255.255.0
MXC.CALO.Sup2T(config)#ip route vrf TEST_VRF 1.1.1.1 255.255.255.255 10.1.2.1
```

```
MXC.CALO.Sup2T#show ip vrf
Name Default RD Interfaces
TEST_VRF <not set> V11
V12
```

Los VRF también se programan secuencialmente. Este fue el primer VRF en el switch (no se configuró ningún otro VRF anteriormente), por lo que el número **vpn** para esta instancia VRF debe ser 1. Ejecute el comando **show platform hardware cef vpn 1** para verificar que esto sea cierto:

```
MXC.CALO.Sup2T-sw2-dfc3#show plat hard cef vpn 1
Codes: decap - Decapsulation, + - Push Label
Index Prefix Adjacency
34 10.1.1.10/32 receive
35 10.1.1.0/32 receive
36 10.1.1.255/32 receive
38 10.1.2.10/32 receive
43 10.1.2.0/32 receive
44 10.1.2.255/32 receive
53 10.1.2.1/32 V12 ,0c11.678b.f6f7
54 1.1.1.1/32 V12 ,0c11.678b.f6f7
[snip...]
```

However, usually, switches have hundred or thousands of VRFs and just count them in the 'show ip vrf' command output would be quite difficult. In order to know which VPN number is assigned to a

