

Entender as entradas de CEF do Catalyst 6500 S2T

Contents

[Introduction](#)

[Prerequisites](#)

[Requirements](#)

[Componentes Utilizados](#)

[Informações de Apoio](#)

[Diagrama de Rede](#)

[Identificar entradas CEF sobre mecanismos de encaminhamento distribuído](#)

[Excluir entradas CEF](#)

[Adicionar uma entrada CEF](#)

[Adicionar e excluir entradas para tabelas de roteamento VRF](#)

Introduction

Este documento descreve como o Cisco Catalyst 6500 com Supervisor Sup2T programa as entradas CEF (Cisco Express Forwarding) configuradas no software Cisco IOS no hardware de placas de linha usado para alcançar o encaminhamento de pacotes.

Prerequisites

Requirements

A Cisco recomenda que você tenha conhecimento destes tópicos:

- Cisco Express Forwarding (CEF)
- Cisco Catalyst 6500 Series Switches
- Cisco Distributed Forwarding Card (DFC)

Componentes Utilizados

As informações neste documento são baseadas nas seguintes versões de hardware e software:

- Placa de linha Cisco Catalyst 6500 WS-X6848-GE-TX (com DFC4).
- Cisco Catalyst 6500 com Supervisor 2T no IOS versão 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. Se a rede estiver ativa, certifique-se de que você entenda o impacto potencial de qualquer comando.

Informações de Apoio

O CEF como um mecanismo de comutação de Camada 3 é usado pela maioria dos switches multicamada da Cisco. É imperativo que os engenheiros de rede entendam como o CEF funciona para solucionar problemas de interrupções da rede, perda de pacotes ou cenários de atraso de pacotes diariamente.

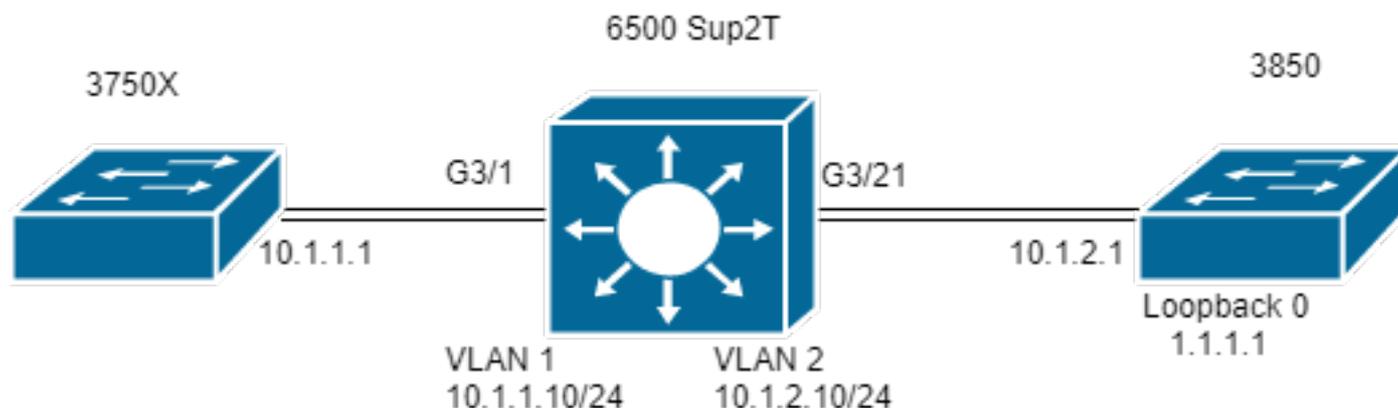
O supervisor Sup2T em modo autônomo ou como o VSS é atualmente implantado por muitas redes corporativas como um switch central, agrega praticamente todos os outros dispositivos de roteamento ou switching. Isso também significa que encaminha a maioria do tráfego intra e inter-dominios para entregar com sucesso os pacotes aos seus destinos. Para isso, o Sup2T deve ter informações de roteamento corretas aprendidas estática ou dinamicamente por meio de protocolos de roteamento.

Em um chassi modular, podem existir vários mecanismos de encaminhamento além dos do supervisor. Determinadas placas de linha (especialmente as de nova geração, como C6800-32P10G) já incluem seu próprio mecanismo de encaminhamento para melhorar o desempenho de comutação de pacotes, a pesquisa de entradas CEF é executada localmente e faz com que os recursos sejam melhor distribuídos para o tráfego que ingressa em diferentes placas de linha. Elas são conhecidas como DFCs (Distributed Forwarding Cards).

Essas entradas CEF compartilhadas em todos os mecanismos de encaminhamento podem não ser alocadas em HW por vários motivos, desde uma condição de defeito de software, esgotamento de recursos até condições de CPU altas e impede que o switch tenha tempo suficiente para atualizar todas as entradas, o que pode causar uma série de eventos indesejáveis.

Diagrama de Rede

Rede:



```
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
```

Identificar entradas CEF sobre mecanismos de encaminhamento distribuído

No diagrama, um switch 6506 autônomo tem um Supervisor 2T instalado, bem como uma placa de linha WS-6848-GE-TX com um DFC no slot 3. O host 3750X que está conectado à placa de linha através da porta G3/1 envia tráfego para o endereço 1.1.1.1 do Loopback 0 do 3850.

Para isso, o 3750X tem uma rota estática para o endereço IP 1.1.1.1 através do salto seguinte 10.1.1.10, que é a SVI da VLAN 1 no switch Sup2T. O switch Sup2T precisa rotear esse tráfego para o switch 3850 com base em uma entrada de rota estática para IP 1.1.1.1/32 através do salto seguinte 10.1.2.1, que é a interface 3850 conectada ao Sup2T na 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
```

Esteja ciente de que, por uma questão de simplicidade, os switches 3750X e 3850 estão conectados ao 6500 através da mesma placa de linha. Isso significa que o tráfego é pesquisado localmente e encaminhado localmente também.

Um pacote ingressa no switch Sup2T via Gi3/1, eventualmente alcança o mecanismo de encaminhamento (já que este é um DFC). O mecanismo de encaminhamento analisa o campo de endereço IP de destino neste pacote e faz uma pesquisa sobre as entradas CEF programadas para a melhor correspondência (Máscara Mais Longa).

Como esta é uma placa DFC, significa que ela tem suas próprias entradas CEF e para verificá-las, é necessário que nós conectemos à placa de linha com o comando **attach [dec] ou conectemos o switch [1-2] mod [dec]** para VSS.

Agora, você deve estar no prompt DFC, no comando **show platform hardware cef** ou no **show platform hardware cef vpn 0** retornar todas as entradas CEF programadas para a tabela de roteamento geral (VPN 0/ No VRF).

Como o objetivo é o prefixo 1.1.1.1/32, use o comando **show platform hardware cef vpn 0 lookup 1.1.1.1**. O comando retorna a melhor correspondência para o prefixo 1.1.1.1 e a que ele usa para realmente encaminhar o tráfego:

```
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
```

A entrada CEF está lá, ela foi programada como resultado de nossa entrada estática programada no software IOS através do comando **ip route 1.1.1.1 255.255.255 10.1.2.1**.

Você também pode verificar se essa entrada recebe acertos e se o tráfego é encaminhado com essa entrada através dos comandos **show platform hardware cef 1.1.1.1 detail** que retorna uma entrada de adjacência:

```
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, a entrada de adjacência mostra como o pacote é regravado e se o tráfego é regravado por esta entrada de adjacência:

```
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 <<<<

```

O **dest_mac** e **src_mac** são os valores de interesse principal, que indicam os novos cabeçalhos L2 gravados para este pacote. O endereço MAC destino 0c11.678b.f6f7 é 10.1.2.1, que é o 3850 (próximo salto para acessar 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

```

Além disso, o campo **Statistics** mostra que o tráfego realmente atinge essa entrada de adjacência e os cabeçalhos L2 são regravados de acordo.

Excluir entradas CEF

Excluir entradas CEF pode nos ajudar a excluir qualquer entrada que possa ser mal programada (para uma entrada de adjacência errada, por exemplo) ou mesmo para fins de treinamento. Também fornece uma maneira de modificar um caminho de roteamento.

Para excluir uma entrada CEF, você precisa entender que as entradas CEF são programadas sequencialmente e têm um Índice de hardware atribuído, por exemplo:

```

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

```

Códigos: decap - desencapsulamento, + - Rótulo de envio

```

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 é o elemento mais importante para excluir uma entrada CEF, pois é usada como referência. No entanto, para fazer qualquer alteração nele, ele deve ser convertido em um identificador de software. Você pode conseguir isso com o 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
Agora que você conhece o identificador do software, pode continuar com a exclusão da entrada CEF com o 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.

```

Note: O valor do comprimento da máscara é 32, pois essa é uma rota específica do host (1.1.1.1/32)

Agora, nossa entrada CEF é excluída:

```

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 o comando **test platform hardware cef vpn 0** foi executado no prompt DFC. Dessa forma, a entrada CEF foi removida da tabela CEF do DFC e NÃO do supervisor, você deve ter cuidado com qual mecanismo de encaminhamento as entradas são removidas.

Uma alteração no tráfego tem o risco de não ter visibilidade (no caso de um teste de laboratório), isso pode ser devido ao acerto de outra entrada do CEF. Considere sempre corresponder à mais exata (máscara mais longa). Neste laboratório, ele atinge:

```

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

```

Então, o que essa entrada realmente faz com o pacote?:

```

MXC.CALO.Sup2T-dfc3#show platform hardware cef adjacencies entry 262048
RIT fields: The entry has a Recirc. Format
-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
|decr_ttl=NO | 12_fwd=NO | ccc = 6 |
add_shim_hdr = YES |-----|-----|-----|-----|-----|rc_fidx=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
```

Agora, todo o tráfego com destino 1.1.1.1 que entra pela placa de linha 3 é recirculado com cabeçalho de calço e direcionado para a CPU. Às vezes, em vez dessa entrada CEF, outro 0.0.0.0/0 com adjacência **de queda** é visto e faz exatamente a mesma coisa.

Note: Avalie quais entradas CEF foram removidas. Uma alta utilização da CPU pode ser causada por isso. Normalmente, uma rota padrão 0.0.0.0/0 é configurada e o tráfego é encaminhado com base nela (e causa perda de pacotes).

Adicionar uma entrada CEF

Quando uma entrada CEF é adicionada, na maioria dos casos resolve qualquer problema de má programação que cause perda de pacotes, atraso de pacotes ou alta utilização da CPU. Conhecimento de como instalar as entradas CEF no hardware, não só fornece a capacidade de corrigir uma entrada mal programada, como também manipular qualquer encaminhamento de pacotes através de recirculação do pacote, aponte-o para uma interface completamente diferente ou próximo salto, regrave um pacote roteado conforme desejado e/ou descarte-o, etc. Tudo isso, sem recarregar a caixa, remova e defina a configuração ou qualquer alteração aparente. A entrada CEF a adição também pode ser feita sem entrar no modo de configuração. (Como você também fez com o procedimento de remoção de entrada CEF explicado na seção anterior).

Basicamente, há duas situações aqui, quando você tem uma entrada ARP válida para o próximo salto, neste caso 10.1.2.1 e quando não tem (por qualquer motivo). A segunda situação força você a criar uma entrada ARP válida (através de ARP estático):

Etapa 1. Há uma entrada ARP no switch para 10.1.2.1, que é o próximo salto 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
Uma entrada ARP é programada como uma rota de host (/32) na tabela 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
```

Adicionar e excluir entradas para tabelas de roteamento VRF

Durante a configuração feita de todas as etapas anteriores, a string **vpn 0** nos comandos **show platform hardware cef** foi aplicada. Mesmo que pareça completamente desnecessário, uma vez que o comando por padrão retorna as entradas para a tabela de roteamento geral ou **vpn 0**, isso foi feito de propósito para sempre tenha em mente que entradas são adicionadas ou excluídas de instâncias específicas da tabela de roteamento (VRFs), por meio do documento que você adicionou e excluiu a entrada CEF 1.1.1.1/32. No entanto, é muito provável que alguns prefixos existam em VRFs diferentes (i. e. 10.x.x.x) e excluir, adicionar ou modificar uma entrada CEF para um VRF errado pode causar um impacto negativo.

Exclua uma entrada CEF com o prefixo 1.1.1.1/32 para VRF **TEST_VRF**. Para obter uma descrição detalhada da adição de entradas CEF, consulte a seção **Adicionar uma entrada CEF** deste documento.

Para adicionar o VRF, altere os SVIs no switch 6500 para o VRF proposto com o comando **ip vrf forwarding [VRF-NAME]** e, finalmente, adicione a mesma rota estática na nossa tabela **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
```

Os VRFs também são programados sequencialmente. Esse foi o primeiro VRF no switch (nenhum outro VRF foi configurado anteriormente), portanto, o número de vpn para esta instância de VRF deve ser 1. Execute o comando **show platform hardware cef vpn 1** para verificar se isso é verdade:

```
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...]
```

