排除Catalyst 9000 DHCP中繼代理上的慢速或間歇性DHCP故障

目錄

簡介

本文說明如何解決作為DHCP中繼代理的Catalyst 9000系列交換機上的慢速動態主機配置協定(DHCP)地址分配或間歇性DHCP地址分配故障。

必要條件

需求

思科建議您瞭解以下主題：

- DHCP和DHCP中繼代理
- 網際網路控制訊息通訊協定 (ICMP)
- 控制階段管制(CoPP)

採用元件

本文中的資訊係根據以下軟體和硬體版本：

- Catalyst 9000 系列交換器
- Cisco IOS XE®版本16.x和17.x

本文中的資訊是根據特定實驗室環境內的裝置所建立。文中使用到的所有裝置皆從已清除（預設）的組態來啟動。如果您的網路運作中，請確保您瞭解任何指令可能造成的影響。

相關產品

本文件也適用於以下硬體和軟體版本：

- 採用Cisco IOS XE® 16.x的Catalyst 3650/3850系列交換器

背景資訊

控制階段管制(CoPP)功能可保護CPU免受不必要的流量和拒絕服務(DoS)攻擊，從而提高您裝置上的安全性。它還可以保護控制流量和管理流量，使其免受其他高優先順序流量高導致的流量丟棄的影響。

您的裝置通常分為三個操作平面，每個平面都有其自己的目標：

- 資料平面，用於轉發資料包。
- 控制平面，用於正確路由資料。
- 管理平面，用於管理網路元素。

您可以使用CoPP保護大多數CPU繫結的流量，並確保路由穩定性、可達性和資料包傳送。最重要的是，您可以使用CoPP保護CPU免受DoS攻擊。

CoPP使用模組化QoS命令列介面(MQC)和CPU隊列來實現這些目標。根據特定標準將不同型別的控制平面流量組合在一起，並分配給CPU隊列。您可以通過配置硬體中的專用策略來管理這些CPU隊列。例如，您可以修改特定CPU隊列（流量型別）的監察器速率，或者為特定型別的流量禁用監察器。

雖然策略器在硬體中進行了配置，但CoPP不會影響CPU效能或資料平面的效能。但是由於它限制了通向CPU的資料包數量，CPU負載會受到控制。這意味著等待來自硬體的資料包的服務可以看到更受控的輸入資料包速率（該速率可由使用者配置）。

問題

在路由介面或SVI上配置**ip helper-address** 命令時，Catalyst 9000交換機被配置為DHCP中繼代理。配置幫助程式地址的介面通常是下遊客戶端的預設網關。為使交換機向其客戶端提供成功的DHCP中繼服務，它必須能夠處理入站DHCP發現消息。這需要交換機接收DHCP發現並將此資料包推送到其CPU進行處理。在接收並處理DHCP發現後，中繼代理會建立一個新的單播資料包，該資料包源自DHCP發現接收所在的介面，目的地為**ip helper-address**配置中定義的IP地址。建立資料包後，硬體將轉發資料包並傳送到DHCP伺服器，然後對其進行處理，最後將其傳送回中繼代理，以便客戶端的DHCP過程可以繼續。

遇到的常見問題是中繼代理的DHCP事務資料包由於受到特定ICMP方案（如ICMP重定向或ICMP目標無法到達消息）的影響而意外受到傳送到CPU的流量影響。此行為可能表現為客戶端無法及時從DHCP獲取IP地址，甚至無法完成DHCP分配失敗。在某些場景中，可能只在每天的某些時間觀察到行為，例如網路負載完全最大時的峰值工作時間。

如背景部分所述，Catalyst 9000系列交換機附帶在裝置上配置和啟用的預設CoPP策略。此CoPP策略用作服務品質(QoS)策略，該策略位於前面板埠上接收的流量路徑中，並且發往裝置CPU。其速率根據流量型別和策略中配置的預定義閾值來限制流量。一些預設情況下分類和速率受限的流量示例包括路由控制資料包（通常標有DSCP CS6）、拓撲控制資料包(STP BPDU)和低延遲資料包(BFD)。應優先處理這些資料包，因為能夠對這些資料包進行可靠處理，可以確保網路環境穩定。

使用**show platform hardware fed switch active qos queue stats internal cpu policer**命令檢視CoPP管制器統計資訊。

ICMP重新導向佇列（佇列6）和BROADCAST佇列（佇列12）共用相同的PlcIdx（0位元管制器索引）。這表示任何需要由裝置CPU處理的廣播流量（例如DHCP發現）都會與目的地為ICMP重定向隊列中裝置CPU的流量共用。這可能會導致前面提到的問題，其中DHCP事務由於ICMP重定向隊列流量耗盡需要由BROADCAST隊列服務的流量，從而導致合法廣播資料包被丟棄。

```
9300-Switch#show platform hardware fed switch active qos queue stats internal cpu policer

CPU Queue Statistics
============================================================================================
(default) (set) Queue Queue
QId PlcIdx Queue Name Enabled Rate Rate Drop(Bytes) Drop(Frames)
--------------------------------------------------------------------------------------------
0 11 DOT1X Auth Yes 1000 1000 0 0
1 1 L2 Control Yes 2000 2000 0 0
2 14 Forus traffic Yes 4000 4000 0 0
3 0 ICMP GEN Yes 600 600 0 0
4 2 Routing Control Yes 5400 5400 0 0
5 14 Forus Address resolution Yes 4000 4000 0 0
6    0      ICMP Redirect               Yes     600       600      0            0  <-- Policer
Index 0
7 16 Inter FED Traffic Yes 2000 2000 0 0
8 4 L2 LVX Cont Pack Yes 1000 1000 0 0
9 19 EWLC Control Yes 13000 13000 0 0
10 16 EWLC Data Yes 2000 2000 0 0
11 13 L2 LVX Data Pack Yes 1000 1000 0 0
12   0      BROADCAST                   Yes     600       600      0            0  <-- Policer
Index 0
13 10 Openflow Yes 200 200 0 0
14 13 Sw forwarding Yes 1000 1000 0 0
15 8 Topology Control Yes 13000 16000 0 0
16 12 Proto Snooping Yes 2000 2000 0 0
17 6 DHCP Snooping Yes 500 500 0 0
18 13 Transit Traffic Yes 1000 1000 0 0
19 10 RPF Failed Yes 250 250 0 0
20 15 MCAST END STATION Yes 2000 2000 0 0
<snip>
```

超過CoPP策略中預設每秒600資料包速率的流量在到達CPU之前被丟棄。

```
9300-Switch#show platform hardware fed switch active qos queue stats internal cpu policer

CPU Queue Statistics
============================================================================================
(default) (set) Queue Queue
QId PlcIdx Queue Name Enabled Rate Rate Drop(Bytes) Drop(Frames)
--------------------------------------------------------------------------------------------
0 11 DOT1X Auth Yes 1000 1000 0 0
1 1 L2 Control Yes 2000 2000 0 0
2 14 Forus traffic Yes 4000 4000 0 0
3 0 ICMP GEN Yes 600 600 0 0
4 2 Routing Control Yes 5400 5400 0 0
5 14 Forus Address resolution Yes 4000 4000 0 0
6    0      ICMP Redirect               Yes     600       600      3063106173577 3925209161
<-- Dropped packets in queue
7 16 Inter FED Traffic Yes 2000 2000 0 0
8 4 L2 LVX Cont Pack Yes 1000 1000 0 0
9 19 EWLC Control Yes 13000 13000 0 0
10 16 EWLC Data Yes 2000 2000 0 0
11 13 L2 LVX Data Pack Yes 1000 1000 0 0
12   0      BROADCAST                   Yes     600       600      1082560387 3133323
<-- Dropped packets in queue
```

```
13 10 Openflow Yes 200 200 0 0
14 13 Sw forwarding Yes 1000 1000 0 0
15 8 Topology Control Yes 13000 16000 0 0
16 12 Proto Snooping Yes 2000 2000 0 0
17 6 DHCP Snooping Yes 500 500 0 0
18 13 Transit Traffic Yes 1000 1000 0 0
19 10 RPF Failed Yes 250 250 0 0
20 15 MCAST END STATION Yes 2000 2000 0 0
<snip>
```

案例1: ICMP重新導向

對於第一個場景，請考慮以下拓撲：

事件的順序如下：

1. 10.10.10.100上的使用者發起到裝置10.100.100.100（遠端網路）的telnet連線。

2.目的IP位於不同的子網中，因此資料包將傳送到使用者的預設網關10.10.10.15。

3.當Catalyst 9300收到此封包進行路由時，它會將該封包傳送到其CPU以產生ICMP重新導向。

之所以產生ICMP重新導向，是因為，從9300交換器的角度來看，如果筆記型電腦直接將此封包傳送到位於10.10.10.1的路由器，效率會更高，因為無論怎樣，它都是Catalyst 9300的下一躍點，且使用者位於相同的VLAN中。

問題在於，整個流量會在CPU上處理，因為它符合ICMP重新導向條件。如果其他裝置傳送的流量符合ICMP重新導向案例，則有更多流量開始傳送到此佇列中的CPU，這可能會影響BROADCAST佇列，因為它們共用同一個CoPP管制器。

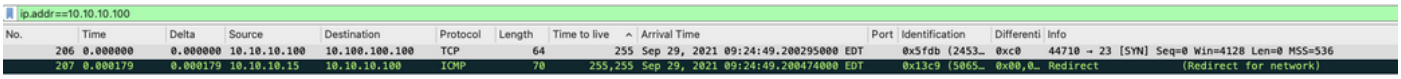
調試ICMP以檢視ICMP重定向系統日誌。

```
9300-Switch#debug ip icmp       <-- enables ICMP debugs
ICMP packet debugging is on
9300-Switch#show logging | inc ICMP
*Sep 29 12:41:33.217: ICMP: echo reply sent, src 10.10.10.15, dst 10.10.10.100, topology BASE,
dscp 0 topoid 0
*Sep 29 12:41:33.218: ICMP: echo reply sent, src 10.10.10.15, dst 10.10.10.100, topology BASE,
dscp 0 topoid 0
*Sep 29 12:41:33.219: ICMP: echo reply sent, src 10.10.10.15, dst 10.10.10.100, topology BASE,
dscp 0 topoid 0
*Sep 29 12:41:33.219: ICMP: echo reply sent, src 10.10.10.15, dst 10.10.10.100, topology BASE,
dscp 0 topoid 0
*Sep 29 12:43:08.127: ICMP: redirect sent to 10.10.10.100 for dest 10.100.100.100, use gw
10.10.10.1
*Sep 29 12:50:09.517: ICMP: redirect sent to 10.10.10.100 for dest 10.100.100.100, use gw
10.10.10.1
*Sep 29 12:50:10.017: ICMP: redirect sent to 10.10.10.100 for dest 10.100.100.100, use gw
```

```
10.10.10.1     <-- ICMP Redirect to use 10.10.10.1 as Gateway
*Sep 29 12:50:14.293: ICMP: redirect sent to 10.10.10.100 for dest 10.100.100.100, use gw
10.10.10.1
*Sep 29 12:50:19.053: ICMP: redirect sent to 10.10.10.100 for dest 10.100.100.100, use gw
10.10.10.1
*Sep 29 12:50:23.797: ICMP: redirect sent to 10.10.10.100 for dest 10.100.100.100, use gw
10.10.10.1
*Sep 29 12:50:28.537: ICMP: redirect sent to 10.10.10.100 for dest 10.100.100.100, use gw
10.10.10.1
*Sep 29 12:50:33.284: ICMP: redirect sent to 10.10.10.100 for dest 10.100.100.100, use gw
10.10.10.1
```

注意：由於規模龐大，建議在啟用ICMP調試之前禁用控制檯日誌記錄和終端監控。

Catalyst 9300 CPU上的嵌入式封包擷取顯示CPU上Telnet連線的初始TCP SYN，以及產生的ICMP重新導向。



ICMP重定向資料包源自Catalyst 9300 VLAN 10介面，目的地為客戶端，其中包含傳送ICMP重定向資料包的原始資料包報頭。

```
▼ Internet Protocol Version 4, Src: 10.10.10.15, Dst: 10.10.10.100
    0100 .... = Version: 4
    .... 0101 = Header Length: 20 bytes (5)
  ▶ Differentiated Services Field: 0x00 (DSCP: CS0, ECN: Not-ECT)
    Total Length: 56
    Identification: 0x13c9 (5065)
  ▶ Flags: 0x0000
    Time to live: 255
    Protocol: ICMP (1)
    Header checksum: 0x7f75 [validation disabled]
    [Header checksum status: Unverified]
    Source: 10.10.10.15
    Destination: 10.10.10.100
▼ Internet Control Message Protocol
    Type: 5 (Redirect)
    Code: 0 (Redirect for network)
    Checksum: 0x2bec [correct]
    [Checksum Status: Good]
    Gateway address: 10.10.10.1
  ▼ Internet Protocol Version 4, Src: 10.10.10.100, Dst: 10.100.100.100
      0100 .... = Version: 4
      .... 0101 = Header Length: 20 bytes (5)
    ▶ Differentiated Services Field: 0xc0 (DSCP: CS6, ECN: Not-ECT)
      Total Length: 44
      Identification: 0x5fdb (24539)
    ▶ Flags: 0x0000
      Time to live: 255
      Protocol: TCP (6)
      Header checksum: 0xd7fa [validation disabled]
      [Header checksum status: Unverified]
      Source: 10.10.10.100
      Destination: 10.100.100.100
  ▶ Transmission Control Protocol, Src Port: 44710, Dst Port: 23
```

解決方案

在此案例中，可能會防止將封包傳送到CPU，這也會停止ICMP重新導向封包的產生。

現代作業系統不使用ICMP重定向消息，因此生成、傳送和處理這些資料包所需的資源並不是網路裝置上CPU資源的有效使用。

或者，將使用者指向使用預設網關10.10.10.1，但此類配置可能因某種原因而存在，並且不在本文檔的討論範圍之內。

只需使用**no ip redirects** CLI禁用ICMP重定向。

```
9300-Switch#conf t
Enter configuration commands, one per line. End with CNTL/Z.
9300-Switch(config)#interface vlan 10
9300-Switch(config-if)#no ip redirects       <-- disable IP redirects
9300-Switch(config-if)#end
```

驗證介面上是否禁用了ICMP重定向。

```
9300-Switch#show ip interface vlan 10
Vlan10 is up, line protocol is up
Internet address is 10.10.10.15/24
Broadcast address is 255.255.255.255
Address determined by setup command
MTU is 1500 bytes
Helper address is not set
Directed broadcast forwarding is disabled
Multicast reserved groups joined: 224.0.0.102
Outgoing Common access list is not set
Outgoing access list is not set
Inbound Common access list is not set
Inbound access list is BLOCK-TELNET
Proxy ARP is disabled
Local Proxy ARP is disabled
Security level is default
Split horizon is enabled
ICMP redirects are never sent       <-- redirects disabled
ICMP unreachables are never sent
ICMP mask replies are never sent
IP fast switching is enabled
IP Flow switching is disabled
IP CEF switching is enabled
IP CEF switching turbo vector
<snip>
```

有關ICMP重定向及其傳送時間的詳細資訊，請訪問以下連結：https://www.cisco.com/c/en/us/support/docs/ip/routing-information-protocol-rip/13714-43.html

案例2:ICMP無法到達

請考慮以下拓撲：10.10.10.100上的使用者發起到10.100.100的Telnet連線。這一次，在VLAN 10 SVI上配置了阻止telnet連線的訪問清單。


```
9300-Switch#show running-config interface vlan 10
Building Configuration..

Current Configuration : 491 bytes
!
interface Vlan10
ip address 10.10.10.15 255.255.255.0
no ip proxy-arp
ip access-group BLOCK-TELNET in              <-- inbound ACL
end
9300-Switch#
9300-Switch#show ip access-list BLOCK-TELNET
Extended IP access list BLOCK-TELNET
10 deny tcp any any eq telnet            <-- block telnet
20 permit ip any any
9300-Switch#
```

事件的順序如下：

1. 10.10.10.100上的使用者發起到裝置10.100.100的telnet連線。

2.目的地IP位於不同的子網中，因此會將封包傳送到使用者預設閘道。

3.當Catalyst 9300收到此封包時，會根據傳入ACL評估該封包並被封鎖。

4.由於封包遭封鎖，且介面上啟用了IP無法連線，因此封包會被傳送到CPU，以便裝置可以產生ICMP目的地無法連線封包。

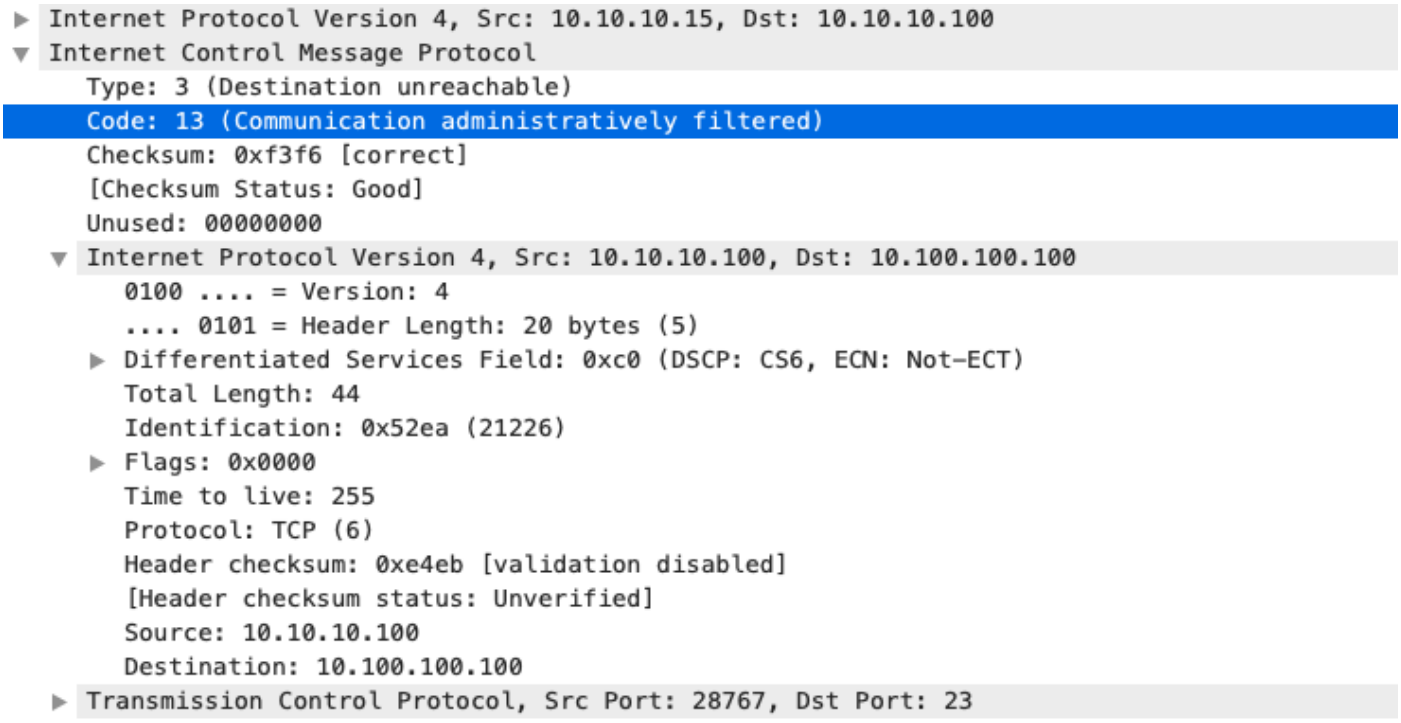
調試ICMP以檢視ICMP目標無法到達的系統日誌。

```
9300-Switch#debug ip icmp                <-- enables ICMP debugs
ICMP packet debugging is on
9300-Switch#show logging | include ICMP
<snip>
*Sep 29 14:01:29.041: ICMP: dst (10.100.100.100) administratively prohibited unreachable sent to
10.10.10.100    <-- packet blocked and ICMP message sent to client
```

注意：由於規模龐大，建議在啟用ICMP調試之前禁用控制檯日誌記錄和終端監控。

Catalyst 9300 CPU上的嵌入式封包擷取顯示CPU上Telnet連線的初始TCP SYN，以及傳送的ICMP目的地無法連線。

ICMP目的地無法連線封包源自於要傳至使用者端的Catalyst 9300 VLAN 10介面，並包含ICMP封包傳送目的地的原始封包標頭。



解決方案

在此案例中，停用已被ACL封鎖的傳出封包的行為，以便產生ICMP目的地無法連線訊息。

Catalyst 9000系列交換器上的路由介面預設啟用IP無法連線功能。

```
9300-Switch#configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
9300-Switch(config)#interface vlan 10
9300-Switch(config-if)#no ip unreachables      <-- disable IP unreachables
```

驗證是否已為介面停用這些功能。

```
9300-Switch#show ip interface vlan 10
Vlan10 is up, line protocol is up
Internet address is 10.10.10.15/24
Broadcast address is 255.255.255.255
Address determined by setup command
MTU is 1500 bytes
Helper address is not set
Directed broadcast forwarding is disabled
Multicast reserved groups joined: 224.0.0.102
Outgoing Common access list is not set
Outgoing access list is not set
Inbound Common access list is not set
Inbound access list is BLOCK-TELNET
Proxy ARP is disabled
Local Proxy ARP is disabled
Security level is default
Split horizon is enabled
ICMP redirects are never sent
ICMP unreachables are never sent     <-- IP unreachables disabled
ICMP mask replies are never sent
IP fast switching is enabled
IP Flow switching is disabled
```

```
IP CEF switching is enabled
IP CEF switching turbo vector
<snip>
```

案例3：超過ICMP TTL

考慮前兩個方案中使用的先前拓撲。這一次，位於10.10.10.100的使用者嘗試訪問網路中已經停用的資源。因此，Catalyst 9300上不再存在用於託管此網路的SVI和VLAN。但是，路由器仍然具有靜態路由，該路由指向Catalyst 9300 VLAN 10介面作為此網路的下一跳。

由於Catalyst 9300不再設定此網路，因此它不會顯示為直接連線，而9300會將它沒有特定路由的任何封包路由到其靜態預設路由（指向10.10.10.1的路由器）。

此行為在使用者嘗試連線到192.168.10.0/24地址空間中的資源時會在網路中引入路由環路。在9300和路由器之間循環該封包，直到TTL到期為止。

1. 使用者嘗試連線到192.168.10/24網路中的資源

2.封包由Catalyst 9300接收，並路由至其下一個躍點的預設路由10.10.10.1，然後遞減TTL 1。

3.路由器收到此資料包，並檢查路由表，發現此網路具有下一跳10.10.10.15的路由。會將TTL遞減1，並將封包路由回到9300。

4. Catalyst 9300收到該資料包，並再次將其路由回10.10.10.1並將TTL減少1。

此程式會重複執行，直到IP TTL達到零。

Catalyst收到IP TTL = 1的封包時，會將封包傳送到CPU並產生ICMP TTL-Exceeded訊息。

ICMP資料包型別為11，其代碼為0（TTL在傳輸過程中過期）。無法通過CLI命令禁用此資料包型別

在此案例中，DHCP流量問題開始發揮作用，因為送回的封包遺漏了接收它們的相同介面，因此會遭受ICMP重新導向。

從使用者傳送的資料包也受到ICMP重定向的影響。在此案例中，DHCP流量很容易從BROADCAST隊列中耗盡。在規模上，由於重定向隊列中投擲的資料包數量，此情況會更糟。

在這裡，通過1000次ping 192.168.10.0/24網路顯示CoPP丟棄，每次ping之間的超時為0秒。9300上的CoPP統計資訊將被清除，並且在傳送ping之前以零位元組丟棄。

```
9300-Switch#clear platform hardware fed switch active qos statistics internal cpu policer
<-- clear CoPP stats

9300-Switch#show platform hardware fed switch active qos queue stats internal cpu policer | i
Redirect|Drop    <-- verify 0 drops
QId PlcIdx Queue Name Enabled Rate Rate Drop(Bytes) Drop(Frames)
```

```
6 0 ICMP Redirect Yes 600 600 0              0                          <-- bytes dropped 0
<snip>
```

使用者將流量傳送到遠端網路。

```
User#ping 192.168.10.10 timeout 0 rep 1000   <-- User sends 1000 pings
Type escape sequence to abort.
Sending 1000, 100-byte ICMP Echos to 192.168.10.10, timeout is 0 seconds:
.....................................................................
.....................................................................
.....................................................................
.....................................................................
.....................................................................
.....................................................................
.....................................................................
.....................................................................
.....................................................................
.....................................................................
.....................................................................
.....................................................................
.....................................................................
.....................
Success rate is 0 percent (0/1000)
```

ICMP調試顯示由於路由環路導致的重定向和TTL超出系統日誌。

```
9300-Switch#debug ip icmp
ICMP packet deubgging is on
*Sep 29 16:33:22.676: ICMP: redirect sent to 10.10.10.100 for dest 192.168.10.10, use gw
10.10.10.1    <-- redirect sent
*Sep 29 16:33:22.678: ICMP: time exceeded (time to live) sent to 10.10.10.100 (dest was
192.168.10.10), topology BASE, dscp 0 topoid 0    <-- TTL exceeded observed
*Sep 29 16:33:22.678: ICMP: time exceeded (time to live) sent to 10.10.10.100 (dest was
192.168.10.10), topology BASE, dscp 0 topoid 0
*Sep 29 16:33:22.678: ICMP: time exceeded (time to live) sent to 10.10.10.100 (dest was
192.168.10.10), topology BASE, dscp 0 topoid 0
<snip>
```

注意：由於規模龐大，建議在啟用ICMP調試之前禁用控制檯日誌記錄和終端監控。

由於傳送到CPU進行重新導向的流量過多，會看到CoPP捨棄。請注意，這僅適用於單個客戶端。

```
9300-Switch#show platform hardware fed switch active qos queue stats internal cpu policer

CPU Queue Statistics
============================================================================================
(default) (set) Queue Queue
QId PlcIdx Queue Name Enabled Rate Rate Drop(Bytes) Drop(Frames)
--------------------------------------------------------------------------------------------
0 11 DOT1X Auth Yes 1000 1000 0 0
1 1 L2 Control Yes 2000 2000 0 0
2 14 Forus traffic Yes 4000 4000 0 0
3 0 ICMP GEN Yes 600 600 0 0
4 2 Routing Control Yes 5400 5400 0 0
5 14 Forus Address resolution Yes 4000 4000 0 0
6    0      ICMP Redirect               Yes     600       600      15407990     126295    <--
drops in redirect queue
7 16 Inter FED Traffic Yes 2000 2000 0 0
8 4 L2 LVX Cont Pack Yes 1000 1000 0 0
```

```
<snip>
```

解決方案

此案例中的解決方案是停用ICMP重新導向，與案例1相同。路由環路也是個問題，但強度會因資料包被轉發重定向而加劇。

當TTL為1時，也會傳送ICMP TTL-Exceeded封包，但這些封包使用不同的CoPP管制器索引，且不會與BROADCAST共用佇列，因此DHCP流量不會受到影響。

只需使用no ip redirects CLI禁用ICMP重定向。

```
9300-Switch#configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
9300-Switch(config)#interface vlan 10
9300-Switch(config-if)#no ip redirects        <-- disable IP redirects
9300-Switch(config-if)#end
```

相關資訊

- 配置嵌入式資料包捕獲
- 瞭解ICMP重新導向
- 技術支援與文件 - Cisco Systems

關於此翻譯

思科已使用電腦和人工技術翻譯本文件，讓全世界的使用者能夠以自己的語言理解支援內容。請注意，即使是最佳機器翻譯，也不如專業譯者翻譯的內容準確。Cisco Systems, Inc. 對這些翻譯的準確度概不負責，並建議一律查看原始英文文件（提供連結）。