

瞭解整個ACI交換矩陣的PXE引導並對其進行故障排除

目錄

- [預啟動執行環境 \(PXE, 有時發音為pixie\)](#)
- [將PXE客戶端重定向到其他TFTP伺服器](#)
- [重要故障排除說明](#)
- [可能用於PXE引導的DHCP選項](#)
- [客戶拓撲](#)

預啟動執行環境(PXE, 有時發音為精靈)

PXE或網路引導允許BIOS通過網路下載和引導作業系統(OS)。在Cisco NX-OS路由器中，這相當於啟動載入器[loader>]，也就是Kickstart。

在PXE引導中，PXE客戶端的BIOS檢視配置的引導順序，如果配置了PXE引導，則確定硬體是否支援PXE引導。如果是，它會嘗試通過DHCP從網路卡獲取IP地址。在DHCP請求中，PXE客戶端請求某些DHCP選項，例如67 [檔名]。在DHCP應答中，DHCP伺服器使用請求的選項進行應答。然後，PXE客戶端下載在DHCP響應中指定的檔名[網路載入程式(NBP)]，並運行該檔名以使作業系統載入到伺服器上。

將PXE客戶端重定向到其他TFTP伺服器

預設情況下，PXE客戶端將嘗試TFTP從DHCP伺服器的DHCP響應中下載檔名，除非它從DHCP響應中收到選項66、next-server或選項60/43，將其重定向到其他IP地址。bootp和PXE中使用下一個伺服器地址來指定將不同的伺服器用於DHCP和TFTP。

next-server (選項12) 和bootfile (選項0) DHCP選項通常指定初始引導載入程式的TFTP伺服器和檔名。客戶端可以選擇使用Bootfile在DHCPDISCOVER消息中請求特定型別的引導檔案。DHCP伺服器在DHCPOFFER中使用它來完全指定引導檔案目錄路徑和檔名。

還必須為檔名指定選項67。bootfile與選項67之間的區別在於DHCP響應中找到的檔名的位置。使用保留欄位時，響應在響應資料包的Options部分不編號，而是按特定順序提供。

例如，伺服器主機名和引導檔案顯示在下圖中的編號選項之前。

```

▶ Internet Protocol Version 4, Src: 10.1.120.91 (10.1.120.91), Dst: 10.1.208.93 (10.1.208.93)
▶ User Datagram Protocol, Src Port: 67 (67), Dst Port: 67 (67)
▼ Bootstrap Protocol (Offer)
  Message type: Boot Reply (2)
  Hardware type: Ethernet (0x01)
  Hardware address length: 6
  Hops: 0
  Transaction ID: 0xb6c37a01
  Seconds elapsed: 0
  ▶ Bootp flags: 0x8000, Broadcast flag (Broadcast)
  Client IP address: 0.0.0.0 (0.0.0.0)
  Your (client) IP address: 10.250.118.251 (10.250.118.251)
  Next server IP address: 0.0.0.0 (0.0.0.0)
  Relay agent IP address: 10.250.118.1 (10.250.118.1)
  Client MAC address: Cisco_c3:7a:01 (00:25:b5:c3:7a:01)
  Client hardware address padding: 00000000000000000000
  Server host name not given
  Boot file name: /pxelinux.0
  Magic cookie: DHCP
  ▼ Option: (53) DHCP Message Type (Offer)
    Length: 1
    DHCP: Offer (2)
  ▼ Option: (54) DHCP Server Identifier
    Length: 4
    DHCP Server Identifier: 10.250.120.2 (10.250.120.2)
  ▼ Option: (51) IP Address Lease Time
    Length: 4
    IP Address Lease Time: (85537s) 23 hours, 45 minutes, 37 seconds
  ▼ Option: (58) Renewal Time Value
    Length: 4
    Renewal Time Value: (42768s) 11 hours, 52 minutes, 48 seconds

```

重要故障排除說明

如果PXE引導客戶端沒有在DHCP請求資料包中的選項55引數請求清單中請求特定DHCP選項，則即使在DHCP伺服器上配置了DHCP選項，該DHCP選項也不會在DHCP伺服器響應資料包中。獲得監聽器追蹤軌跡以驗證所請求和給定的選項非常重要。

附註：如果為DHCP伺服器的DHCP ip幫助地址（DHCP中繼）配置了ACI，則可以在第一跳枝葉交換機上使用tcpdump。由於枝葉交換機必須充當DHCP中繼，因此CPU必須處理DHCP資料包並將其轉發到DHCP伺服器。您可以使用管理員帳戶登入並使用命令

```
tcpdump -i any -w /tmp/packet.pcap "埠67"
```

捕獲主機和DHCP伺服器之間的DHCP資料包。如果枝葉交換機沒有任何OOB IP地址，您可以通過使用admin使用者名稱登入apic並使用

```
scp admin@<switch TEP>:/tmp/packet.pcap /home/admin
```

然後，可以使用GUI IP地址將檔案從APIC傳輸。

根據PXE伺服器支援的內容，可以有許多不同的檔名。Linux伺服器最常見的啟動載入程式檔名稱為"/pxelinux.0"，通常位於/tftpboot目錄中。

可能用於PXE引導的DHCP選項

名稱	選項#	組態
引導檔案	0	Bootfile <file>
DEFAULT_ROUTER	3	Default-router 10.250.118.1
網路	11	網路10.250.118.0 255.255.255.0
NEXT_SERVER	12	Next-server 10.45.67.8
使用匹配選項60傳送	43	選件43 hex 06010708070000010 afbd015
選項43的匹配字串	60	選項60 「PXEClient」
單個TFTP伺服器	66	選項66 ip 10.240.221.21
PXE檔名	67	選項67 「 /pxelinux.0」
TFTP伺服器清單	150	選項150 10.240.208.21

如果DHCP伺服器將其選項60字串與客戶端傳送的選項60字串相匹配，則選項43將返回給PXE客戶端。選項43有幾個升序子選項（選項6和選項8等）：

子選項6定義為

06:子選項6

01:1位元組長

07:使用PXE_BOOT_SERVERS清單，禁用組播和廣播發現（僅單播）

子選項8定義為

08:子選項8

07:08選項的大小（位元組）

0000:PXE伺服器型別0（PXE引導伺服器）

01:pxe伺服器數量

0afbd015:10.251.208.21

在DHCP配置中，該選項如下所示

選項43 hex 06010708070000010afbd015

根據<https://support.microsoft.com/en-us/kb/259670>，當DHCP伺服器設定了以下選項時：

60 =客戶端識別符號（設定為「PXEClient」）

66 =引導伺服器主機名或IP地址

67 =引導檔名

並且，當DHCP伺服器的初始DHCP提供包含這些引導選項時，會嘗試從PXE客戶端連線到DHCP伺服器上的埠4011，並發生錯誤「代理DHCP服務未回覆埠4011」。

附註：Microsoft不支援在DHCP伺服器上使用這些選項來重定向PXE客戶端。

客戶拓撲

在此示例中，使用者有2個不同的ACI交換矩陣。在ACI交換矩陣#1配置中，PXE客戶端通過DHCP運行，並從交換矩陣內的Microsoft伺服器啟動。在其他ACI交換矩陣中，PXE客戶端通過L3Out從外部IOS-XE 4507路由器上DHCP關閉，但使用同一Microsoft伺服器下載檔案，因此DHCP伺服器沒有充當TFTP伺服器。

客戶最終採用的解決方案是遷移到linux DHCP伺服器。分析DHCP資料包跟蹤時，客戶似乎錯誤配置了選項43字串，而IOS-XE DHCP伺服器從未正常工作。

要使其在ACI中正常工作，請執行以下操作：

1. 必須配置DHCP中繼以中繼DHCP客戶端資料包
2. 合約必須允許PXE客戶端所在的L3Out和EPG之間的DHCP和TFTP資料包

使用命令，您可能看到TFTP資料包在枝葉交換機上的合約丟棄

```
calo2-leaf2# show logging ip access-list internal packet-log | grep <client ip address>
```

此命令將在2.0版中更改，以新增deny或permit關鍵字，並變成

```
calo2-leaf2# show logging ip access-list internal packet-log deny | grep <client ip address>
```

在該示例中，PXE客戶端不是請求選項66，而是指定選項60並在選項55引數清單中請求選項43。它還請求下一伺服器（選項12）。注意，它通過選項67請求PXE檔名。

```

▶ Bootp flags: 0x8000, Broadcast flag (Broadcast)
Client IP address: 0.0.0.0 (0.0.0.0)
Your (client) IP address: 0.0.0.0 (0.0.0.0)
Next server IP address: 0.0.0.0 (0.0.0.0)
Relay agent IP address: 10.250.118.1 (10.250.118.1)
Client MAC address: Cisco_c3:7a:01 (00:25:b5:c3:7a:01)
Client hardware address padding: 00000000000000000000
Server host name not given
Boot file name not given
Magic cookie: DHCP
▶ Option: (53) DHCP Message Type (Discover)
▼ Option: (55) Parameter Request List
  Length: 24
  Parameter Request List Item: (1) Subnet Mask
  Parameter Request List Item: (2) Time Offset
  Parameter Request List Item: (3) Router
  Parameter Request List Item: (5) Name Server
  Parameter Request List Item: (6) Domain Name Server
  Parameter Request List Item: (11) Resource Location Server
  Parameter Request List Item: (12) Host Name
  Parameter Request List Item: (13) Boot File Size
  Parameter Request List Item: (15) Domain Name
  Parameter Request List Item: (16) Swap Server
  Parameter Request List Item: (17) Root Path
  Parameter Request List Item: (18) Extensions Path
  Parameter Request List Item: (43) Vendor-Specific Information
  Parameter Request List Item: (54) DHCP Server Identifier
  Parameter Request List Item: (60) Vendor class identifier
  Parameter Request List Item: (67) Bootfile name
  Parameter Request List Item: (128) DOCSIS full security server IP [TODO]
  Parameter Request List Item: (129) PXE - undefined (vendor specific)
  Parameter Request List Item: (130) PXE - undefined (vendor specific)
  Parameter Request List Item: (131) PXE - undefined (vendor specific)
  Parameter Request List Item: (132) PXE - undefined (vendor specific)
  Parameter Request List Item: (133) PXE - undefined (vendor specific)
  Parameter Request List Item: (134) PXE - undefined (vendor specific)
  Parameter Request List Item: (135) PXE - undefined (vendor specific)

```

響應顯示引導檔案和選項67

- ▶ Bootp flags: 0x8000, Broadcast flag (Broadcast)
Client IP address: 0.0.0.0 (0.0.0.0)
Your (client) IP address: 10.250.118.251 (10.250.118.251)
Next server IP address: 0.0.0.0 (0.0.0.0)
Relay agent IP address: 10.250.118.1 (10.250.118.1)
Client MAC address: Cisco_c3:7a:01 (00:25:b5:c3:7a:01)
Client hardware address padding: 0000000000000000000000
Server host name not given
Boot file name: /pxelinux.0
Magic cookie: DHCP
- ▶ Option: (53) DHCP Message Type (Offer)
- ▶ Option: (54) DHCP Server Identifier
- ▶ Option: (51) IP Address Lease Time
- ▶ Option: (58) Renewal Time Value
- ▶ Option: (59) Rebinding Time Value
- ▶ Option: (1) Subnet Mask
- ▶ Option: (6) Domain Name Server
- ▶ Option: (3) Router
- ▼ Option: (67) Bootfile name
 Length: 11
 Bootfile name: /pxelinux.0
- ▶ Option: (82) Agent Information Option
- ▶ Option: (255) End

枝葉交換機上的TCPdump: