在IOS XE EVPN/VXLAN中配置DHCP

目錄

簡介

本檔案介紹不同情況下乙太網路VPN(EVPN)虛擬可擴充區域網路(VXLAN)的動態主機設定通訊協定(DHCP)組態，以及Win2012和Win2016 DHCP伺服器的特定方面。

必要條件

需求

思科建議您瞭解EVPN/VXLAN和DHCP。

採用元件

本文中的資訊係根據以下軟體和硬體版本：

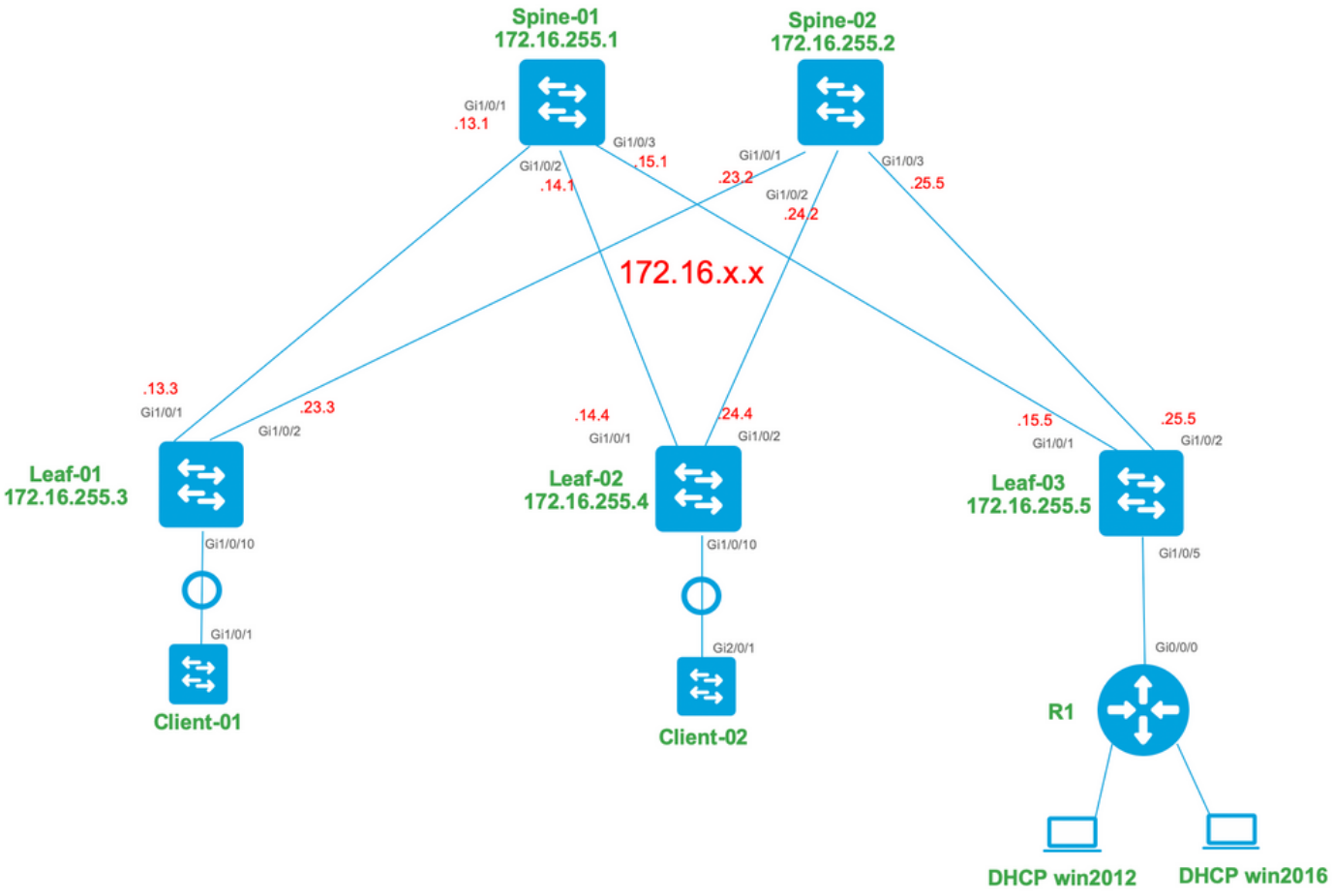
- C9300
- C9400
- C9500
- C9600
- MSFT Windows Server 2012 R2
- MSFT Windows Server 2016

- Cisco IOS XE 16.9.x或更高版本提供的功能

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。

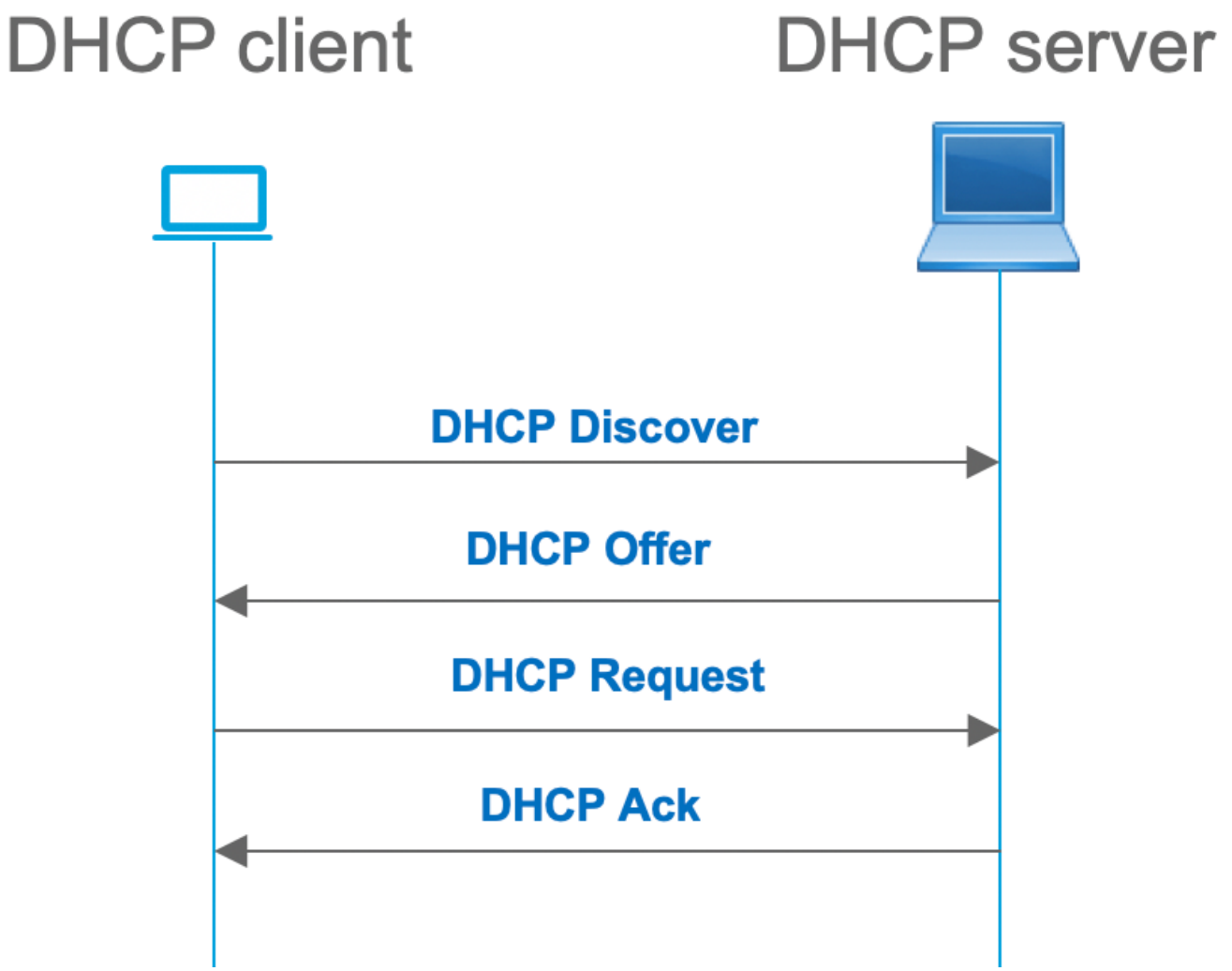
設定

網路圖表



組態

現在，讓我們檢查DHCP客戶端和伺服器之間的消息流。共有四個階段：

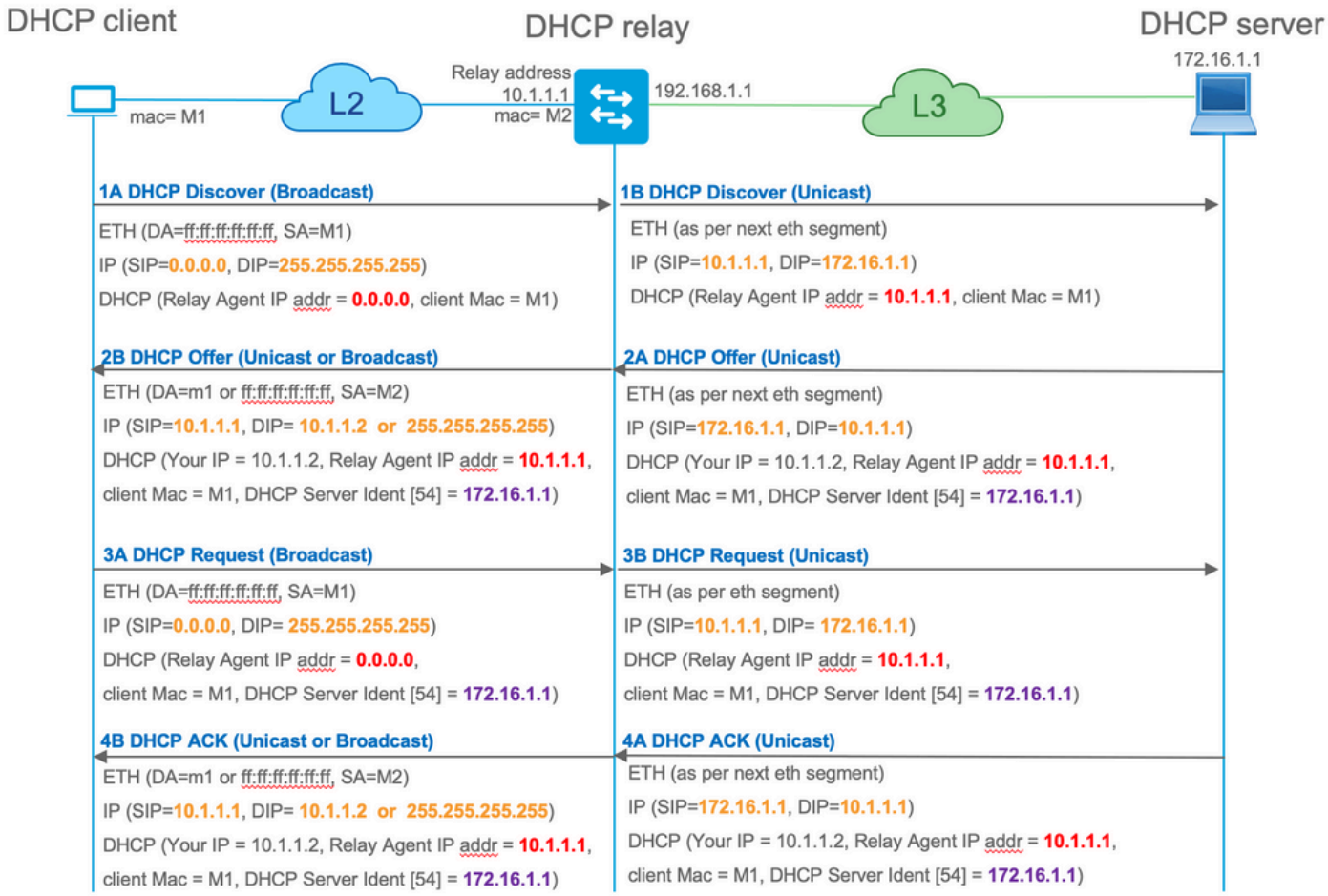


這適用於客戶端和伺服器位於同一子網的情況，但通常情況並非如此。在大多數情況下，DHCP伺服器與客戶端不在同一個子網中，必須通過第3層路由路徑而不是第2層到達。在這種情況下，需要DHCP中繼功能。DHCP中繼（交換機或路由器）功能將廣播轉換為可路由的udp封裝單播，並將其傳送到DHCP伺服器。它如今已成為網路中廣泛使用的配置。

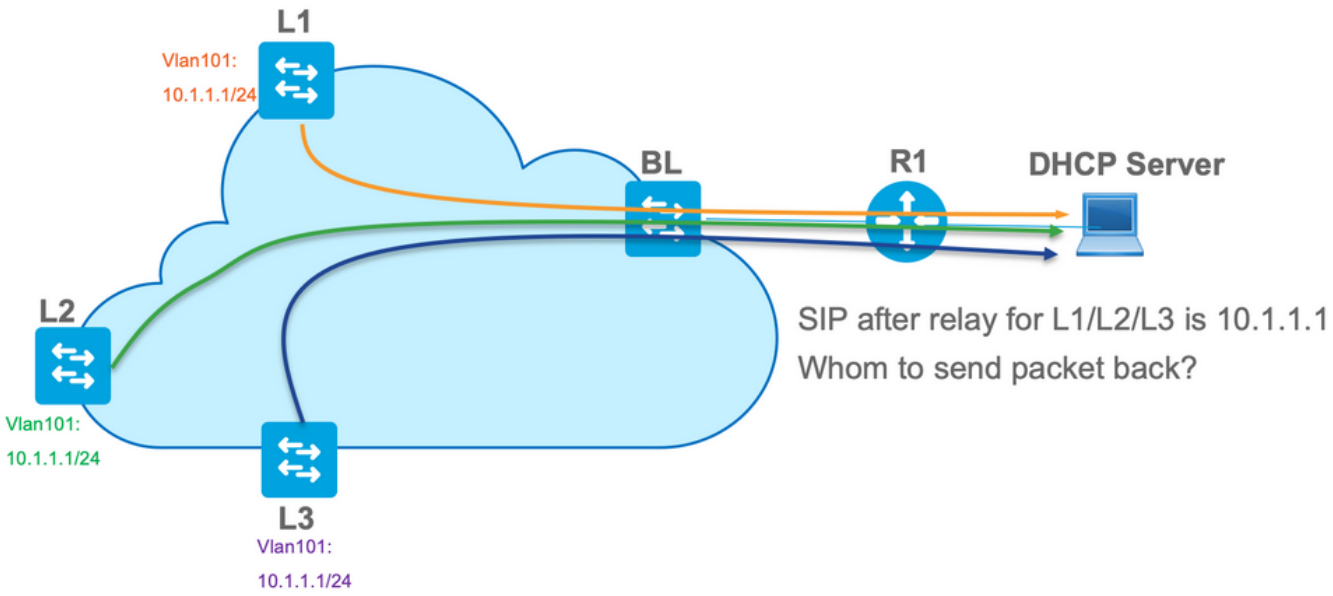
DHCP和EVPN/VXLAN交換矩陣面臨的挑戰：

通常，DHCP伺服器通過L3網路連線到EVPN交換矩陣。這意味著您必須使用DHCP中繼功能將第2層DHCP廣播資料包轉換為第3層單播可路由資料包。

使用DHCP中繼功能時，客戶端、中繼和伺服器之間的DHCP呼叫流的工作方式與以下類似：



中繼後，資料包的源IP是中繼IP。但是，這會在VXLAN/EVPN部署中產生問題，因為由於使用了分散式任播GW(DAG)，通常的源IP不是唯一的。 由於所有VTEP SVI源IP都相同，因此這會導致將來自DHCP伺服器的應答資料包轉發到最近的枝葉。



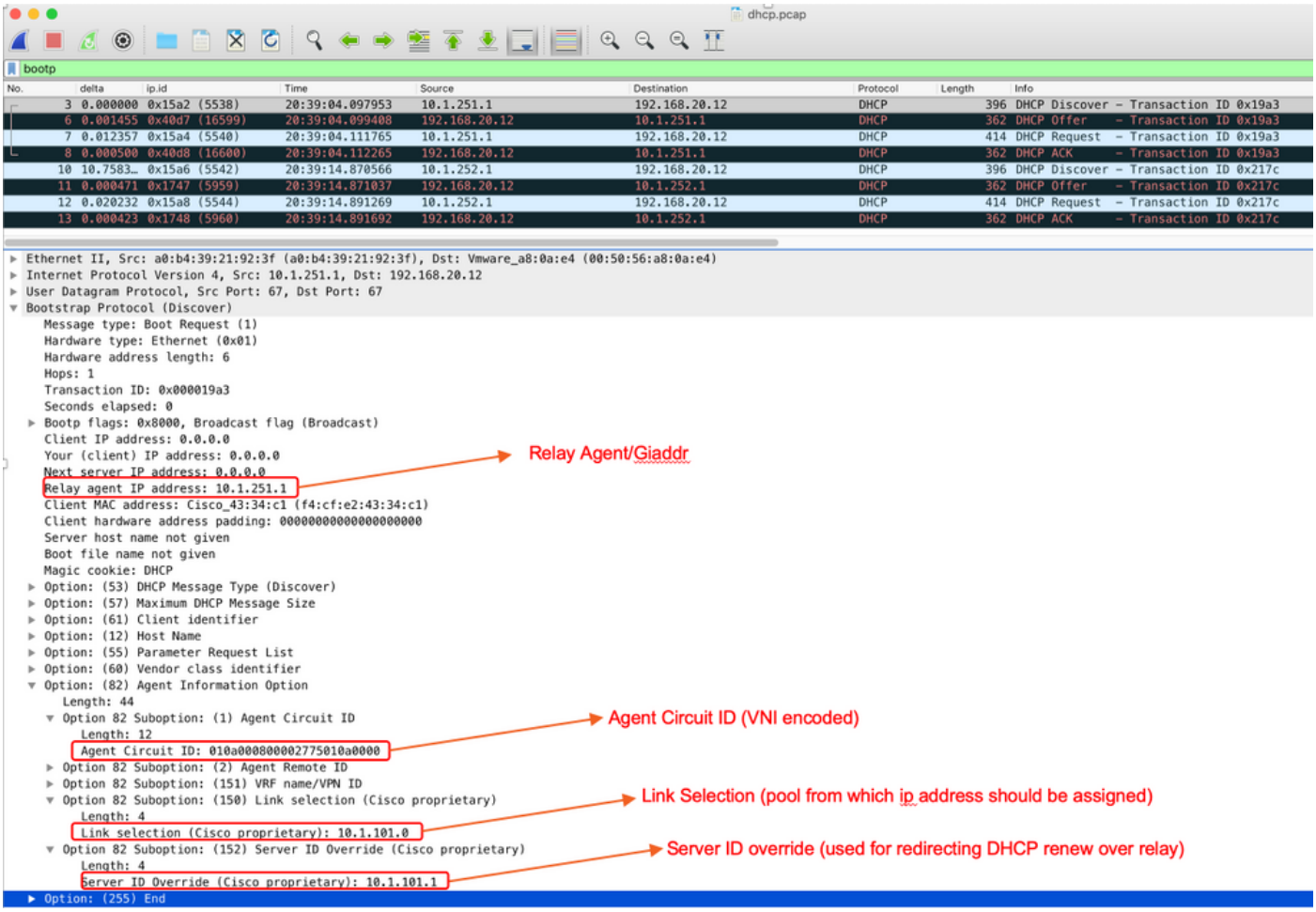
為了解決非唯一源問題，您必須能夠對每個枝葉的中繼DHCP資料包使用唯一的IP地址。另一個問題與GIADDR替換有關。在DHCP伺服器上，您必須選擇正確的池來分配IP地址。從地址池中完成，地址池包含網關IP地址(giaddr)。 對於EVPN交換矩陣，它必須是SVI的IP地址，但在中繼之後，giaddr將被中繼IP地址替換，在本例中該地址是唯一的環回。

如何通知DHCP伺服器，它必須使用哪些池？

為了解決這個問題，使用了選項82。主要有以下幾個重要的子選項：

- 1 — 代**理電路ID**。在VXLAN/EVPN的情況下，此子選項會傳輸VNI ID
- 5 —（或150表示思科專有）。 DHCP資料包來自的實際子網的**Link selection**子選項
- 11 -(或152（思科專有）。 具有DHCP伺服器地址的**Server Identifier Override**子選項
- 151 - **VRF名稱/VPN ID**。此子選項具有VRF名稱/VPN ID

在從DHCP中繼到DHCP伺服器的資料包捕獲中，您可以看到DHCP資料包中存在的這些不同選項，如下圖所示。



交換機配置：

- 選項82具有選擇正確的DHCP池並將資料包從伺服器返回到正確的枝葉所需的所有必要資訊。
- 這僅當DHCP伺服器可以處理選項82資訊時有效，儘管並非所有伺服器都完全支援它（例如win2012 r2）。

```
ip dhcp relay information option vpn        <<<  adds the VRF name/VPN ID to the option 82
ip dhcp relay information option            <<< enables option 82
!
ip dhcp snooping vlan 101-102,201-202
ip dhcp snooping
!
vlan configuration 101
member evpn-instance 101 vni 10101
!
interface Loopback101
```

```
 vrf forwarding green
 ip address 10.1.251.1 255.255.255.255
!
interface Vlan101
 vrf forwarding green
ip dhcp relay source-interface Loopback101    <<< DHCP relay source is unique Loopback
 ip address 10.1.101.1 255.255.255.0
 ip helper-address 192.168.20.12               <<< 192.168.20.12 - DHCP server
```

伺服器配置

Win2012 R2配置選項1 — 每個VNI/SVI每個VTEP的唯一中繼IP

Win2012的主要問題是選項82不完全支援，因此「鏈路選擇」子選項（5或Cisco專有 — 150）不能用於選擇DHCP伺服器上的正確池。

要解決此類問題，可以使用以下方法：

- 必須建立中繼IP地址的作用域，否則DHCP找不到與DHCP GIADDR匹配的池並忽略資料包。必須將完整的IP範圍從DHCP中排除，以防止從中繼IP池進行分配。我們稱這個池為RELAY_POOL
- 必須建立要分配的IP範圍的作用域。我們稱此池IP_POOL
- 必須建立超級作用域，並且必須同時包括RELAY_POOL和IP_POOL

我們來看看DHCP資料包在伺服器上是如何處理的。

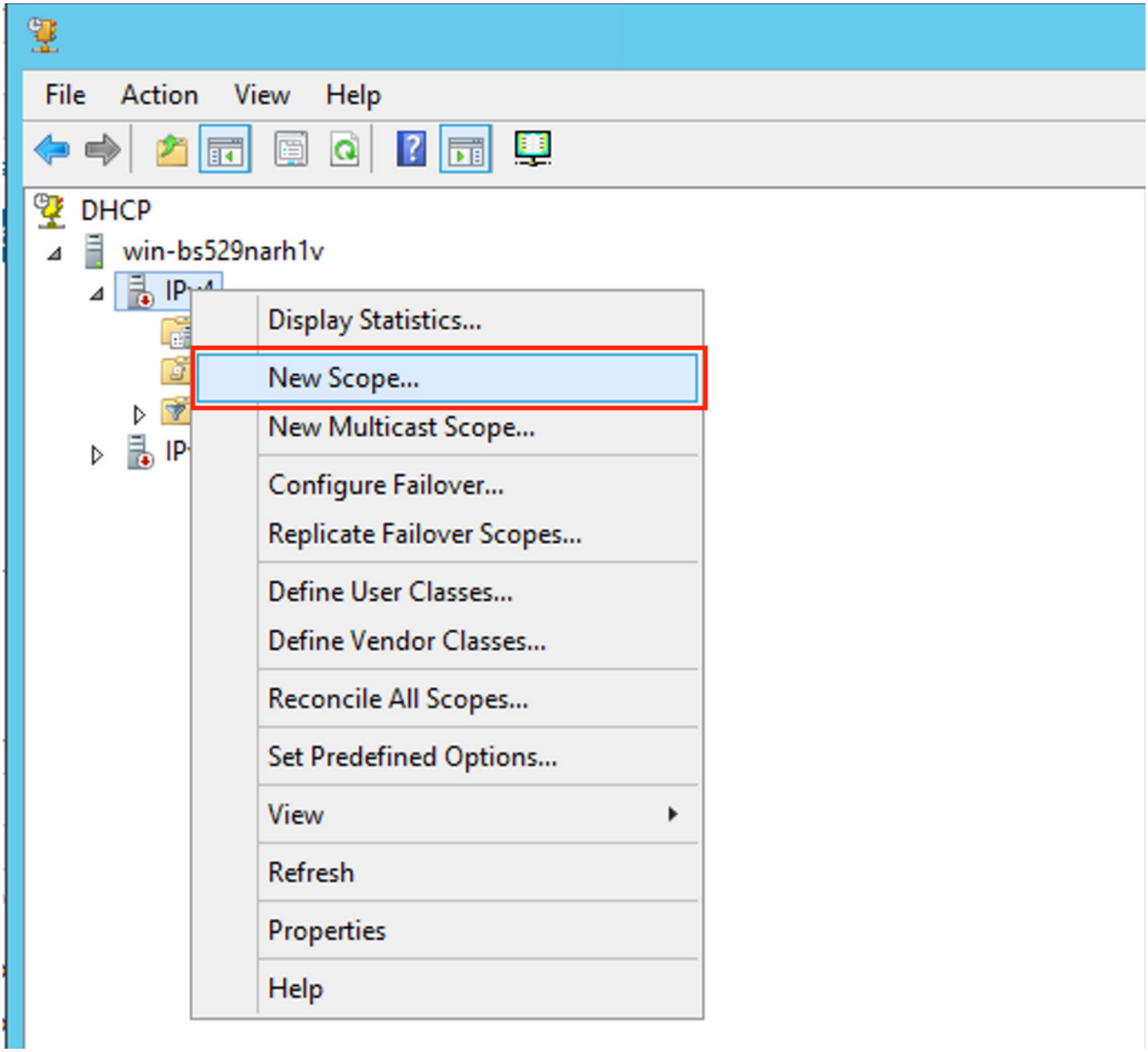
1. 伺服器收到DHCP資料包。
2. 基於GIADDR。在相應的超級作用域中選擇相應的池RELAY_POOL。
3. 由於RELAY_POOL中沒有可用的IP地址（是否記得已排除整個作用域？），因此它在同一超級作用域中回退到IP_POOL。
4. 地址從相應的超級池分配，並傳送回中繼。

此方法的一大缺點是您必須每個vtep的每個VLAN/VNI具有唯一的環回，因為DHCP池是根據中繼地址選擇的。

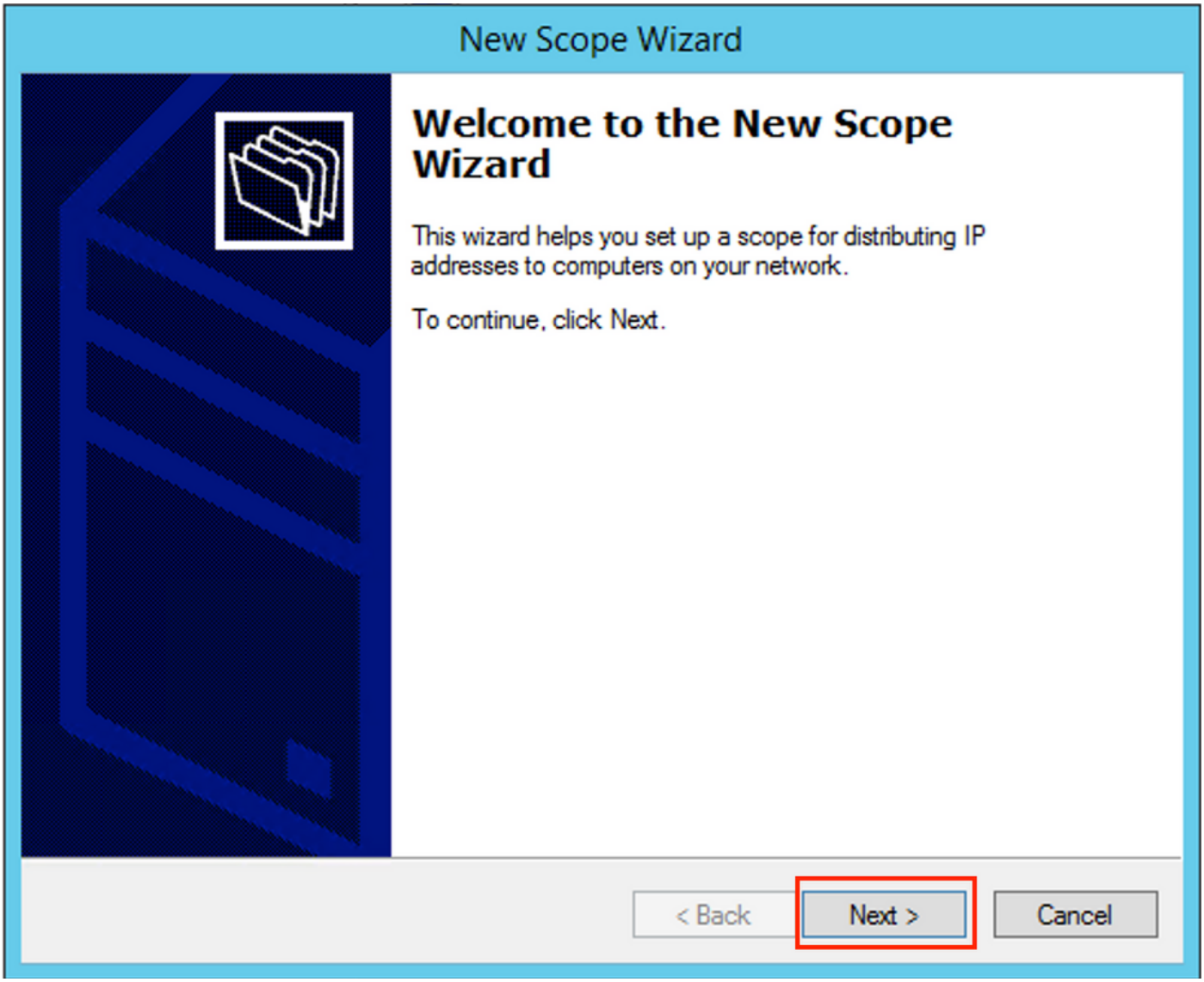
該選項使我們能夠使用大的IP範圍作為中繼IP地址。

選項1：有關如何配置win2012 r2的逐步說明。

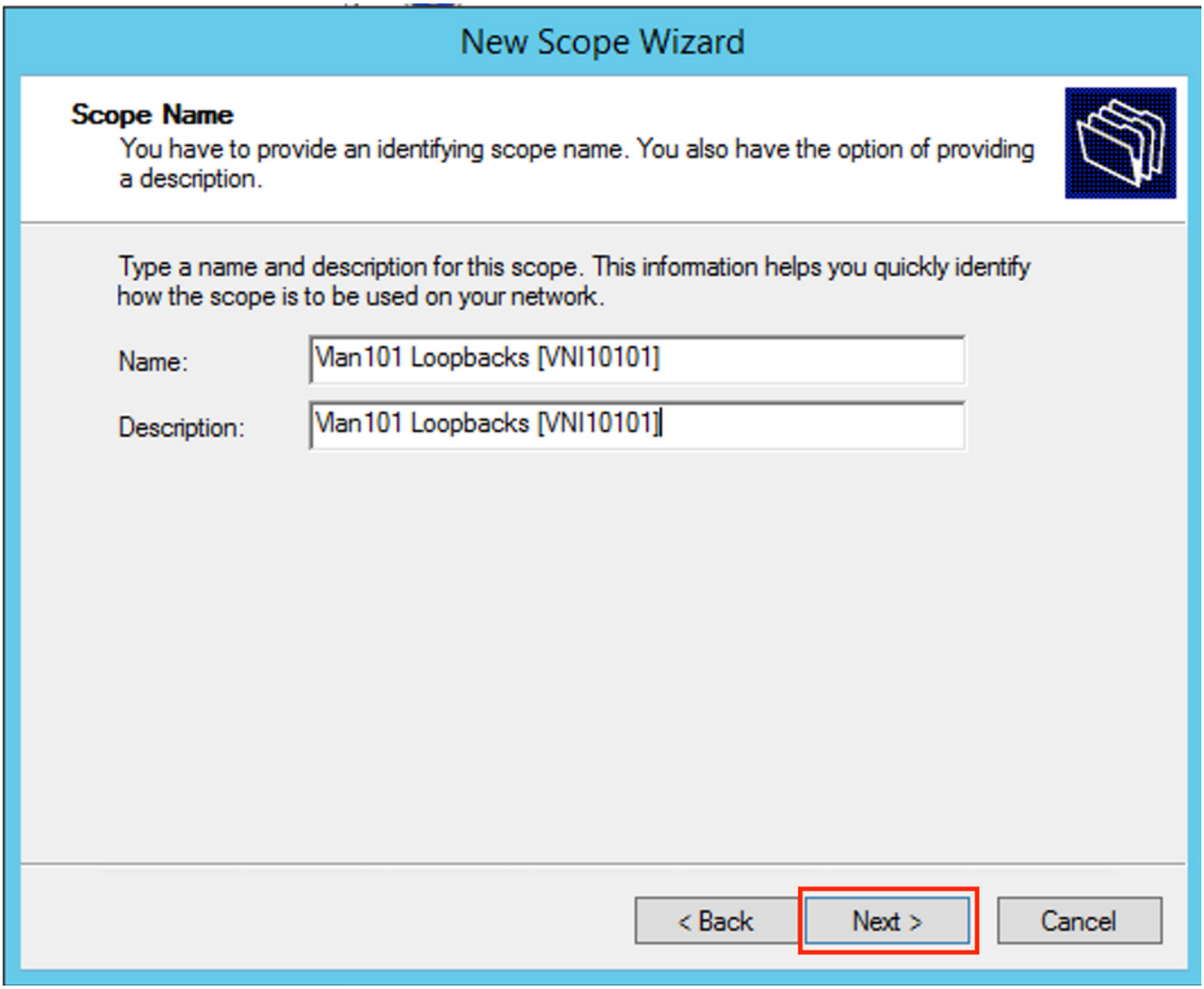
為中繼地址建立DHCP作用域。按一下右鍵並選擇**New Scope**，如下圖所示。



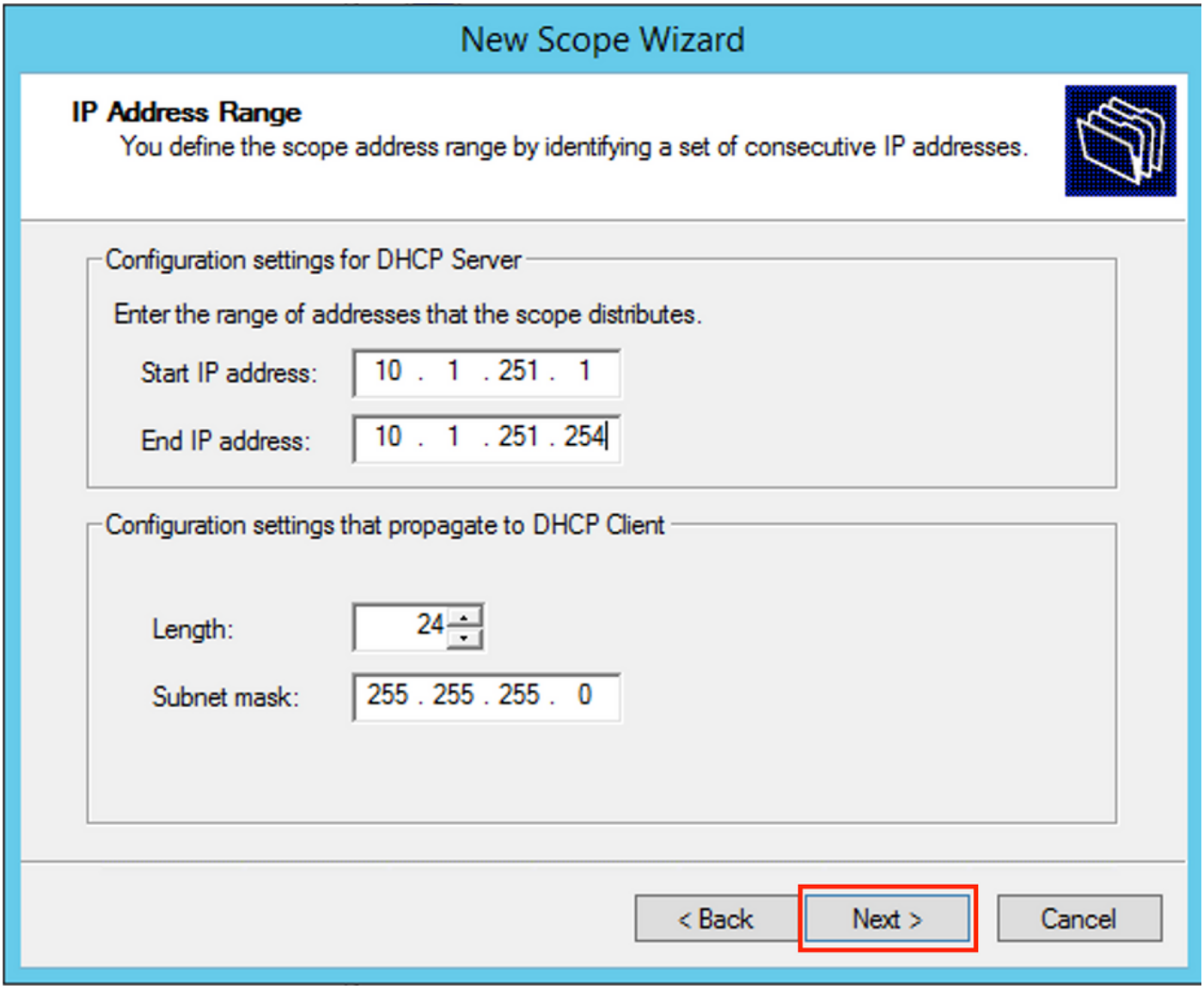
選擇**Next**，如下圖所示。



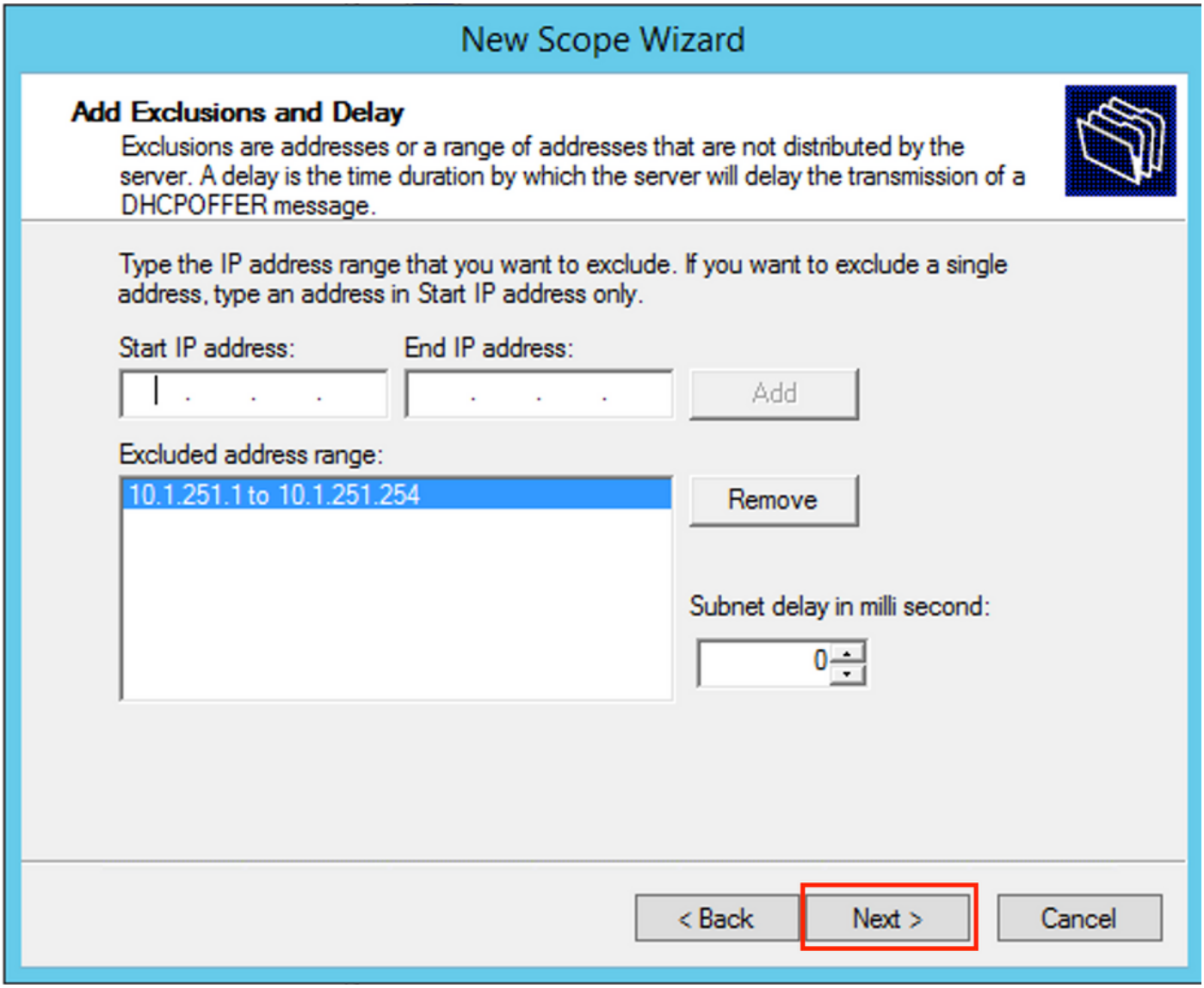
填寫一個有意義的名稱和描述，然後選擇**下一步**，如下圖所示。



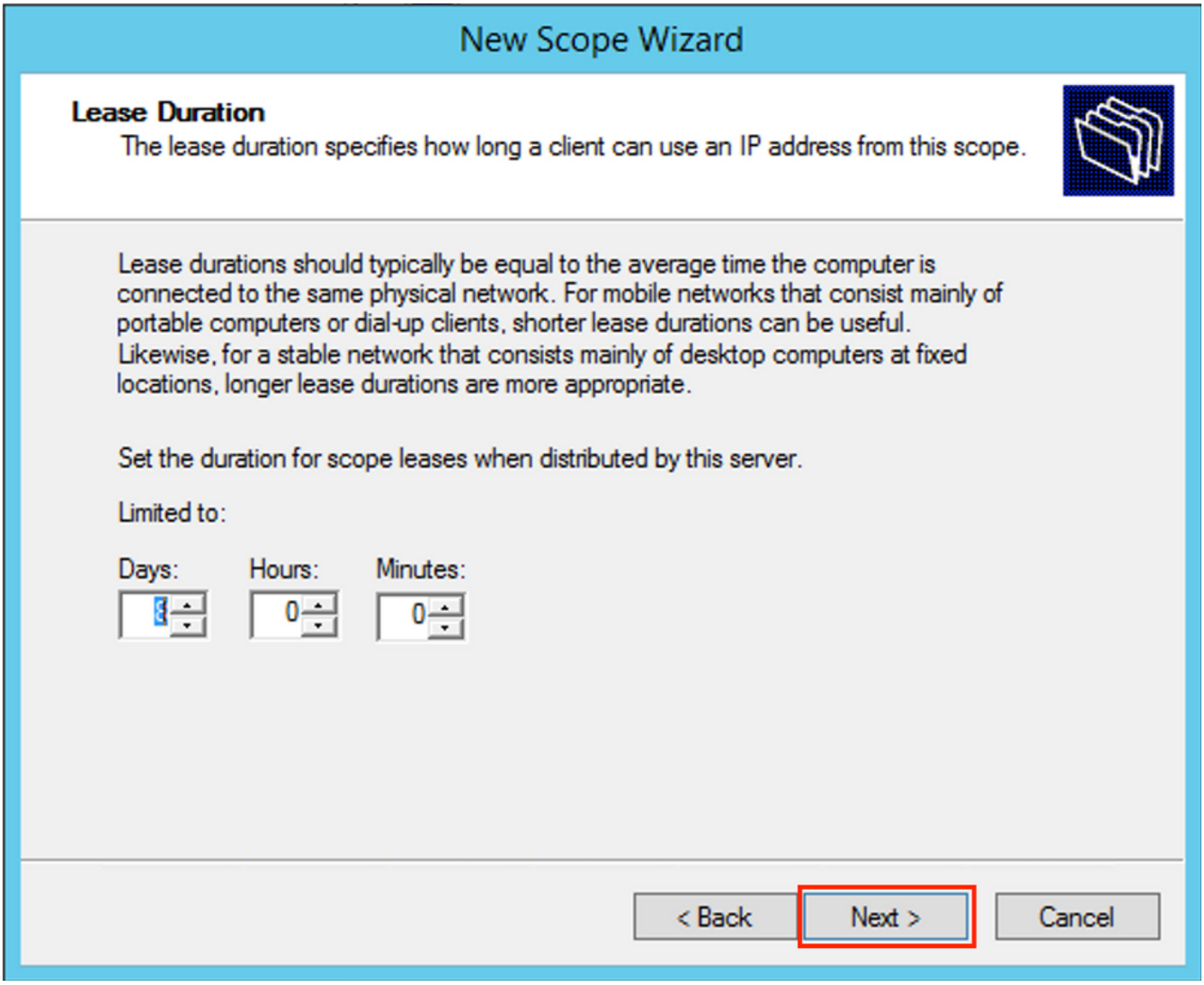
填寫中繼池的IP地址資訊。在此範例中，網路遮罩為/24，但如圖所示，它可以大於或小於（視網路大小而定）。



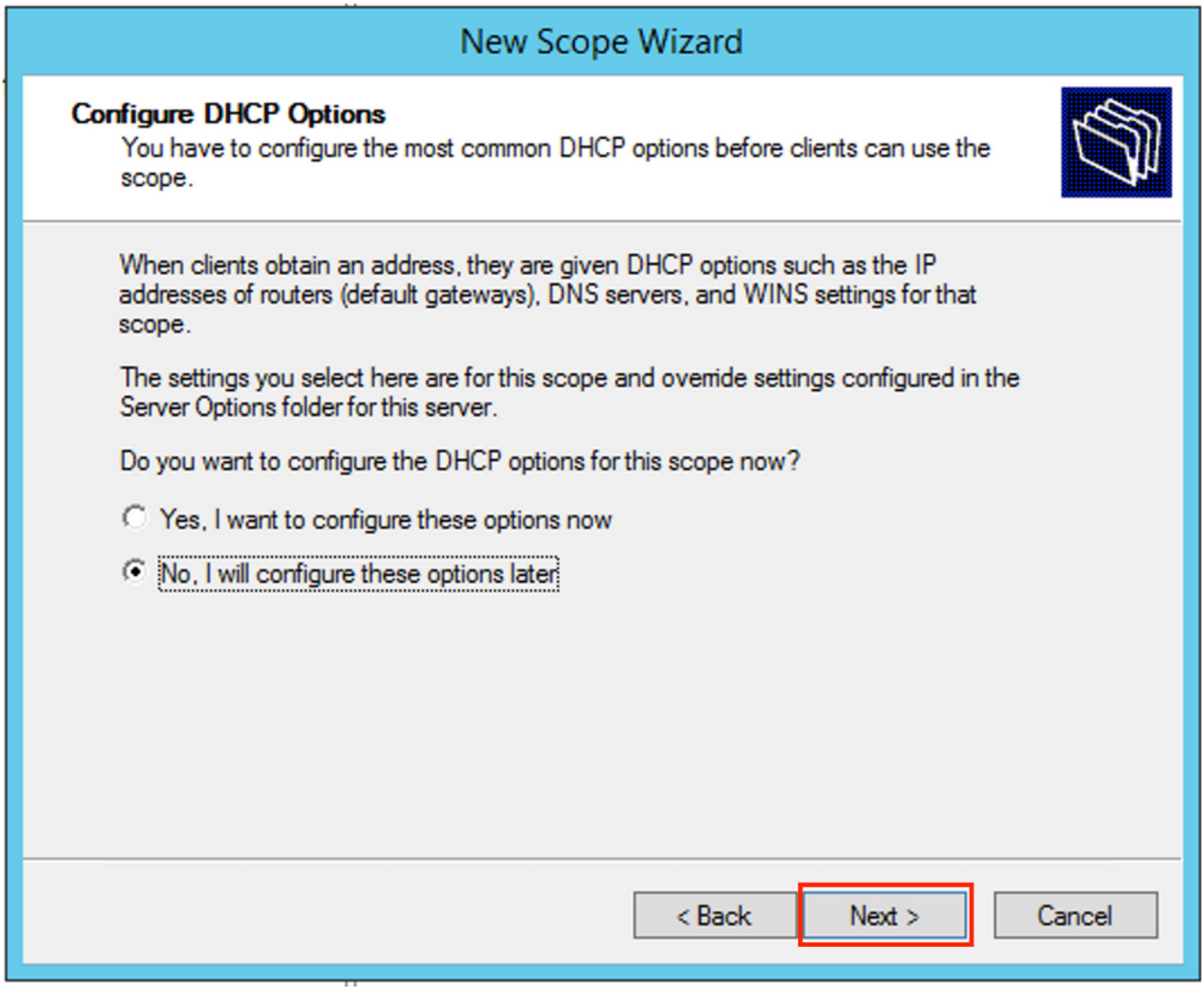
從池中排除所有範圍。這很重要，否則可以從此池分配IP地址。



配置租用時間（預設為8天），如下圖所示。



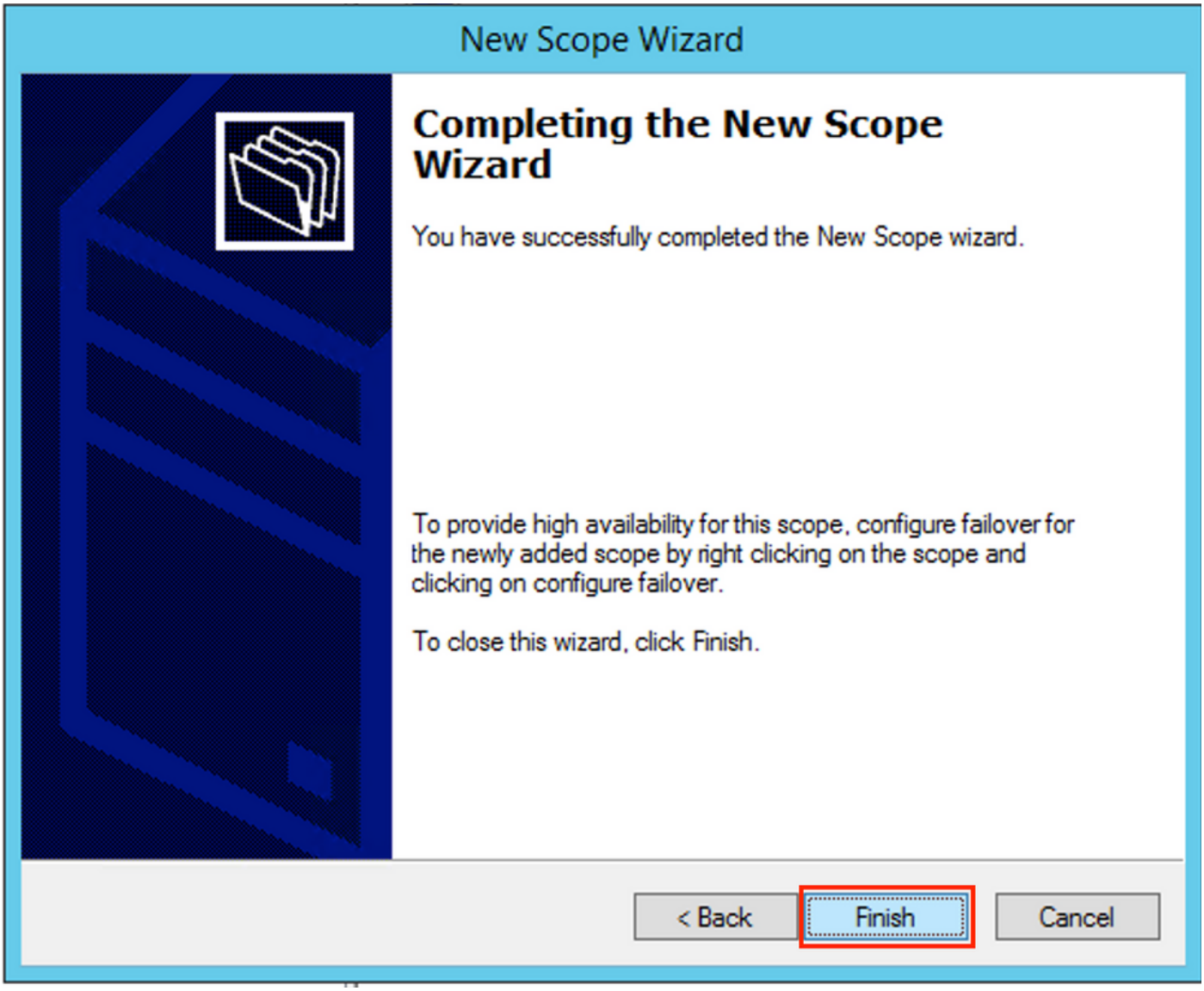
您可以配置DHCP選項引數，如DNS/WINS（在此示例中跳過）。


New Scope Wizard
Configure DHCP Options
You have to configure the most common DHCP options before clients can use the scope.
When clients obtain an address, they are given DHCP options such as the IP addresses of routers (default gateways), DNS servers, and WINS settings for that scope.
The settings you select here are for this scope and override settings configured in the Server Options folder for this server.
Do you want to configure the DHCP options for this scope now?
Yes, I want to configure these options now
No, I will configure these options later
< Back
Next >
Cancel

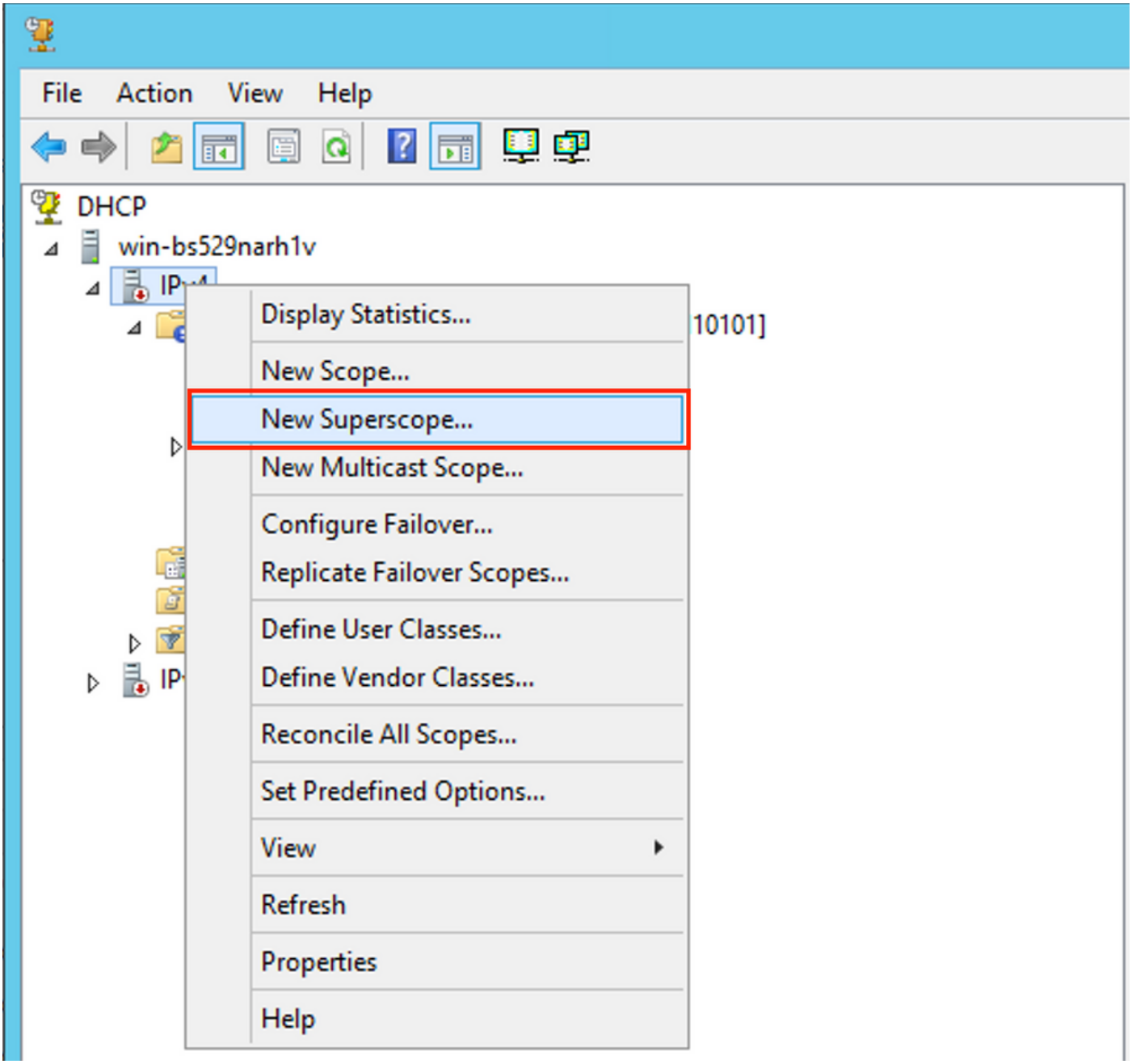
啟用作用域，如下圖所示。


New Scope Wizard
Activate Scope
Clients can obtain address leases only if a scope is activated.
Do you want to activate this scope now?
Yes, I want to activate this scope now
No, I will activate this scope later
< Back
Next >
Cancel

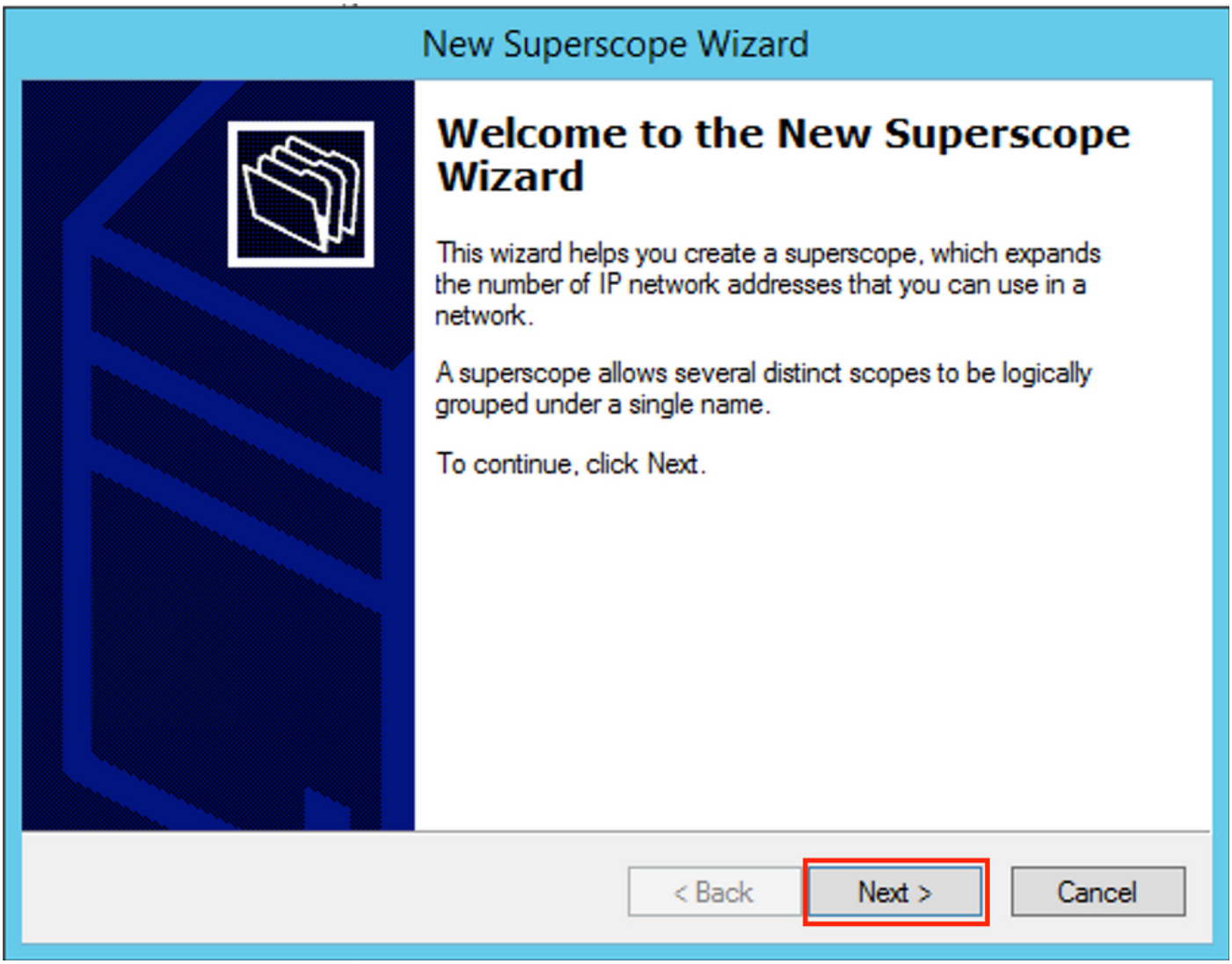
完成配置，如下圖所示。



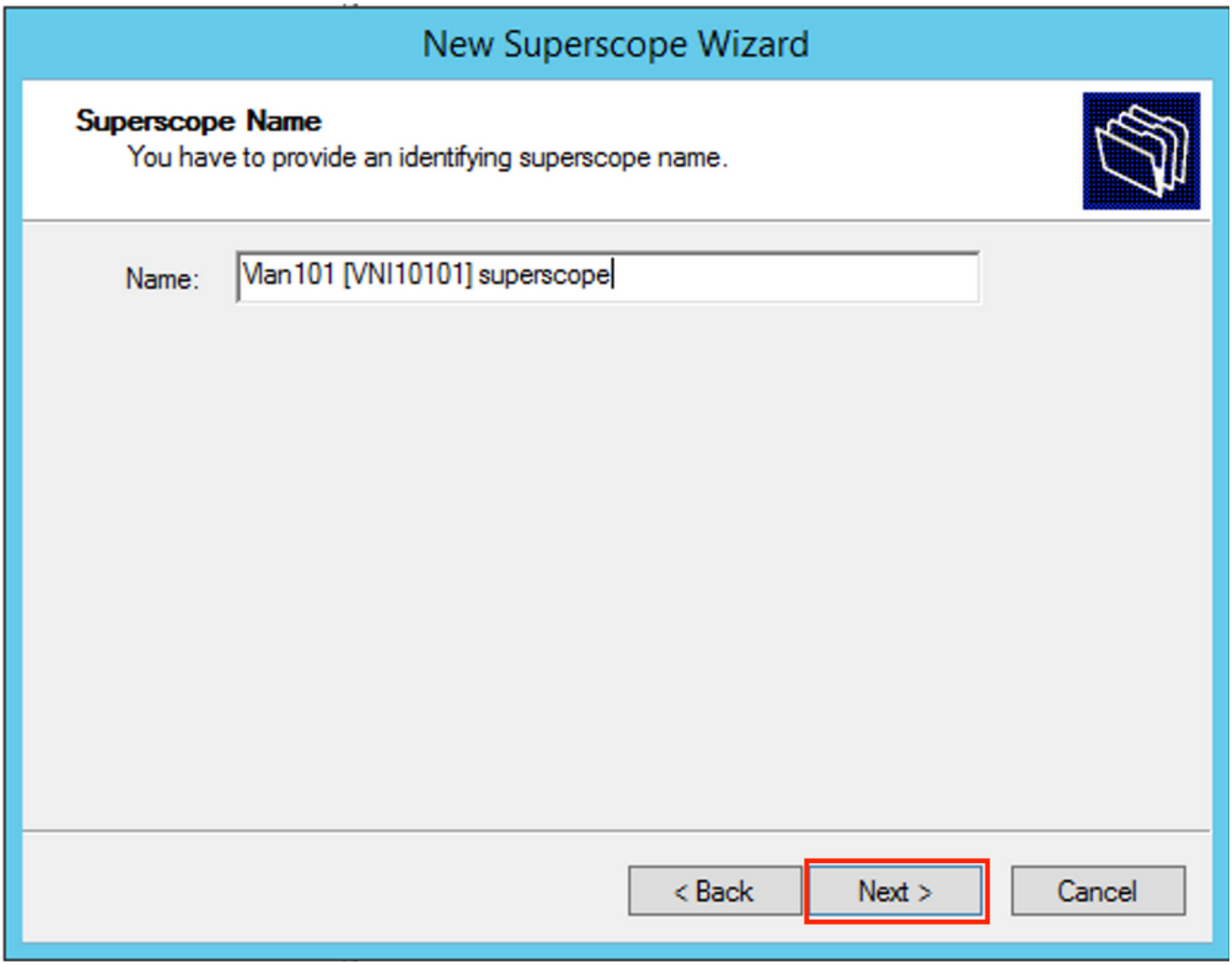
現在建立超級作用域。按一下右鍵並選擇**新建超級作用域**，如下圖所示。



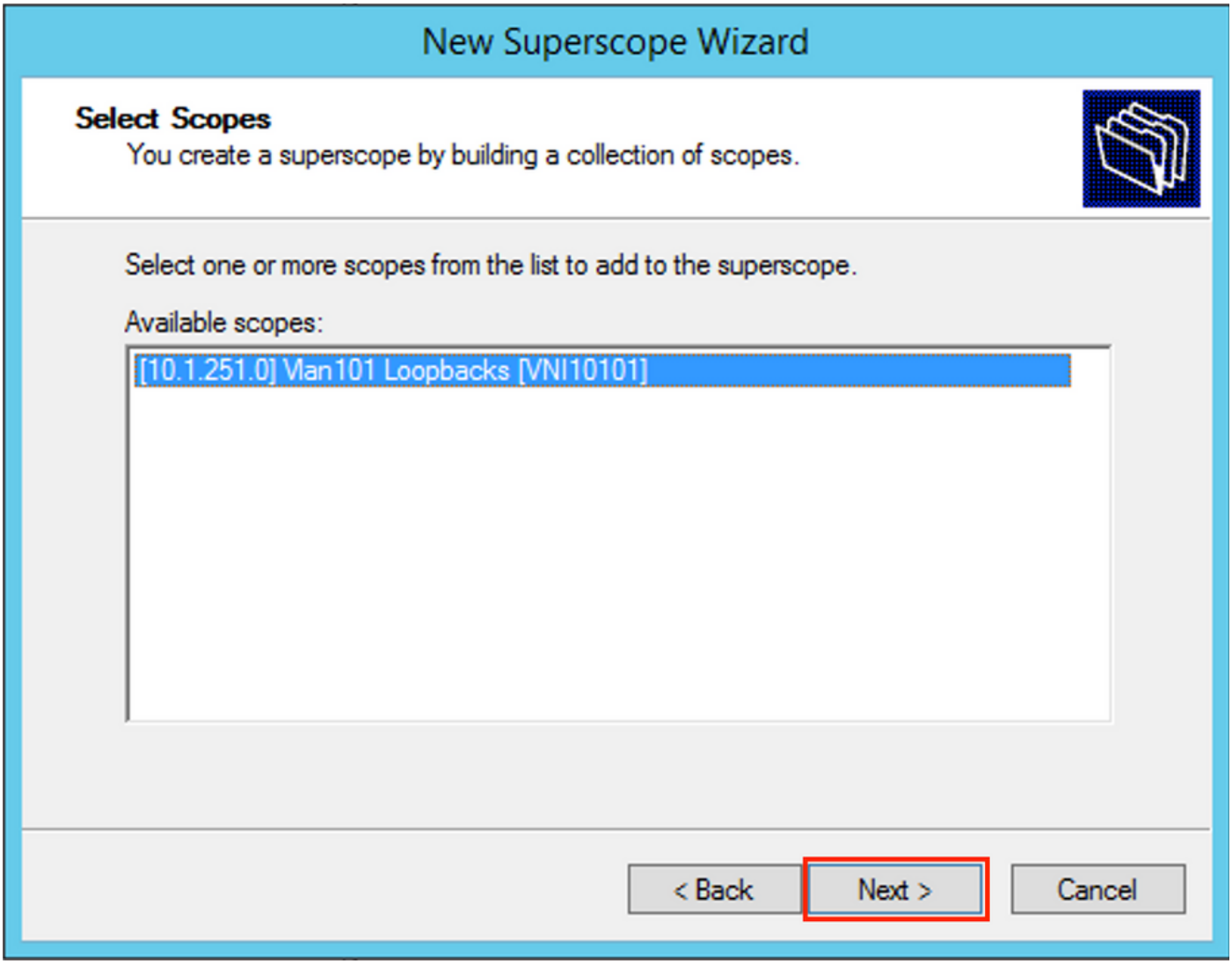
選擇**Next**，如下圖所示。



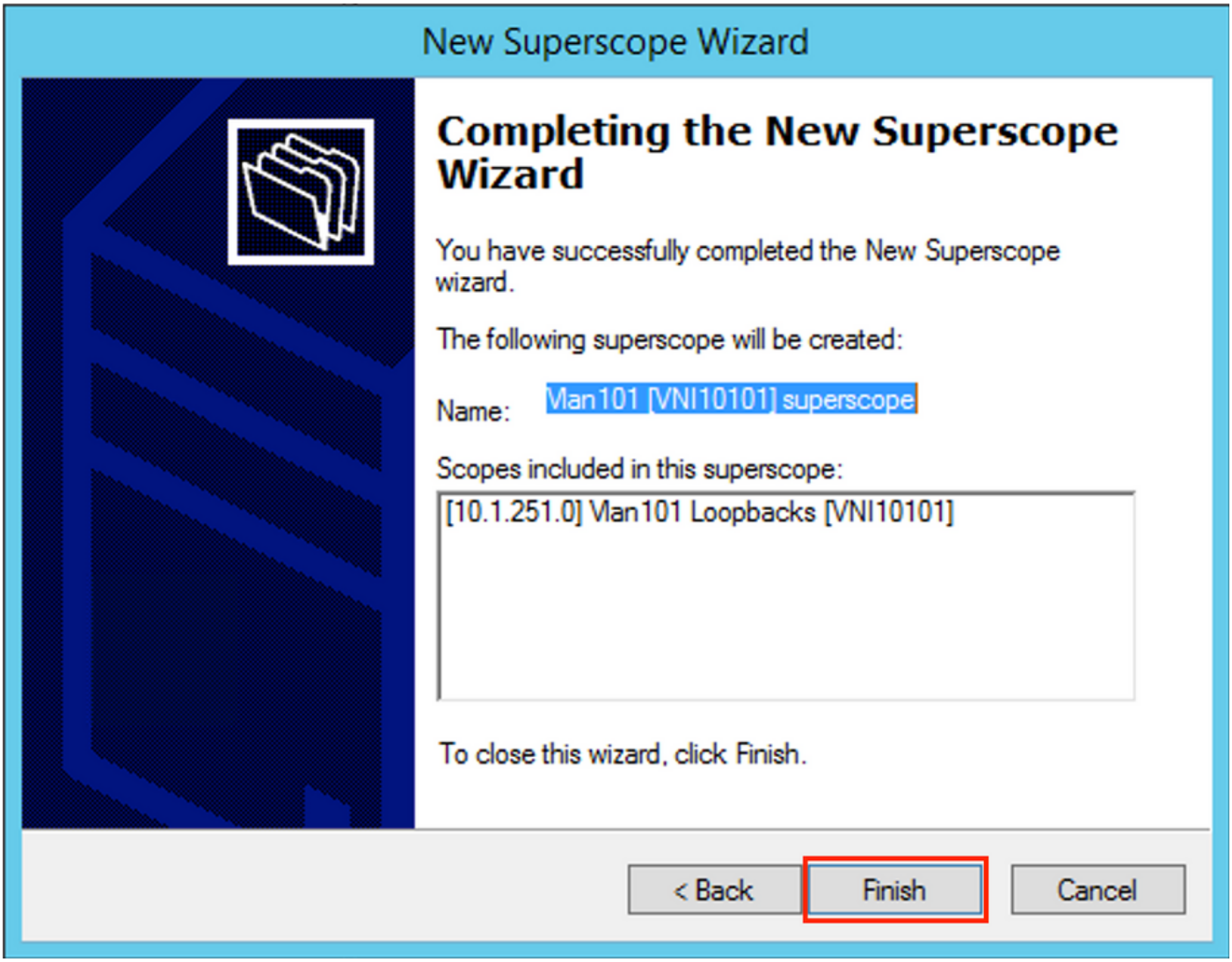
為**超級作用域**選擇一個有意義的名稱，如下圖所示。



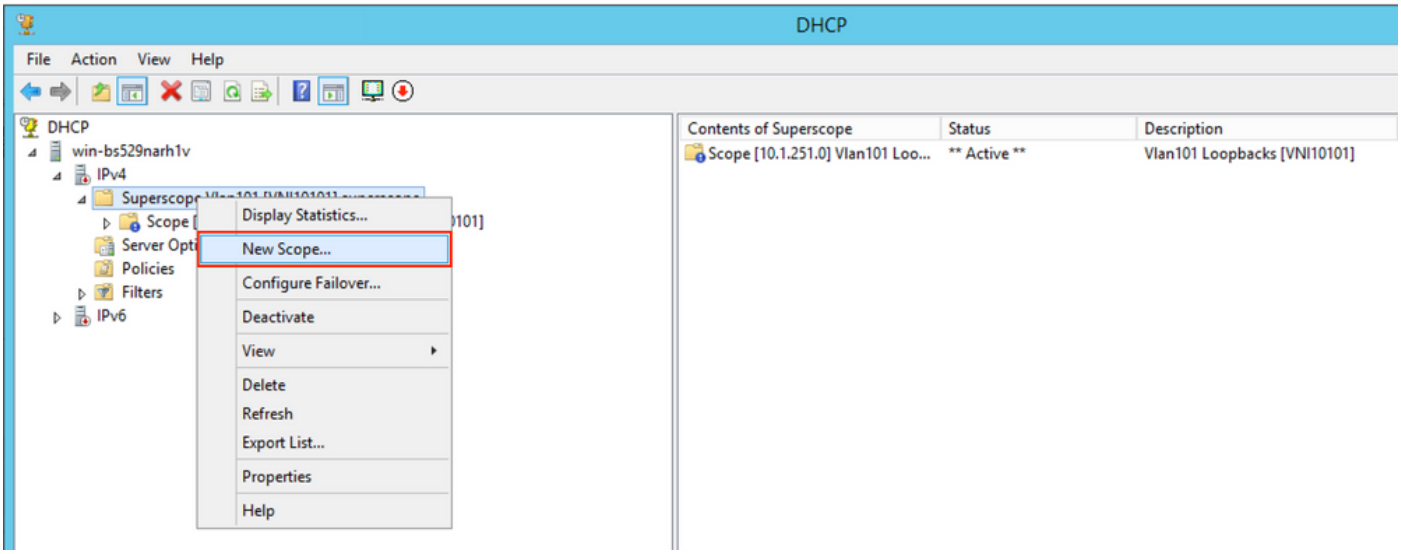
選擇要新增到超級作用域的作用域。


New Superscope Wizard
Select Scopes
You create a superscope by building a collection of scopes.
Select one or more scopes from the list to add to the superscope.
Available scopes:
[10.1.251.0] Vlan101 Loopbacks [VNI10101]
< Back
Next >
Cancel

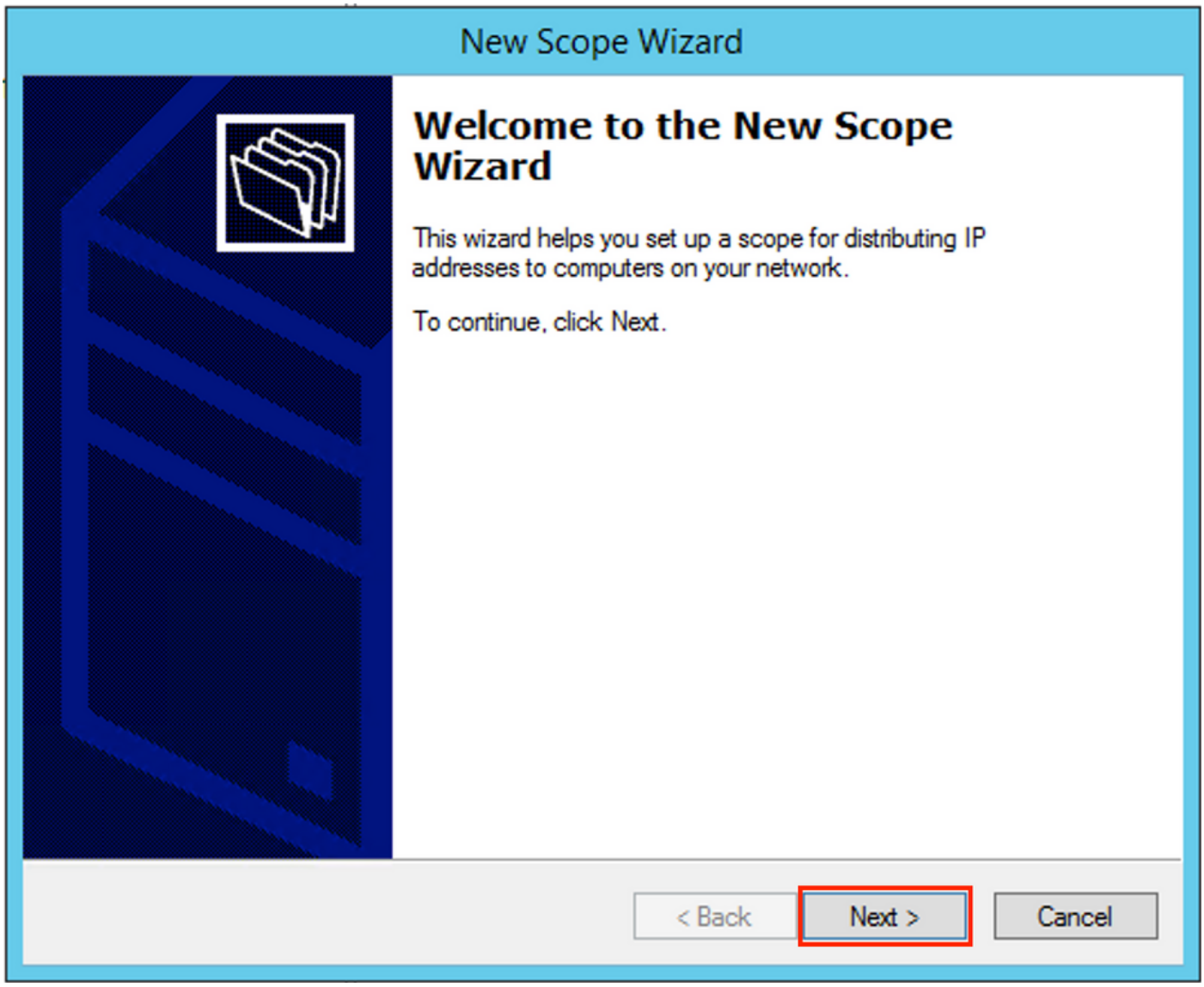
如圖所示完成設定。



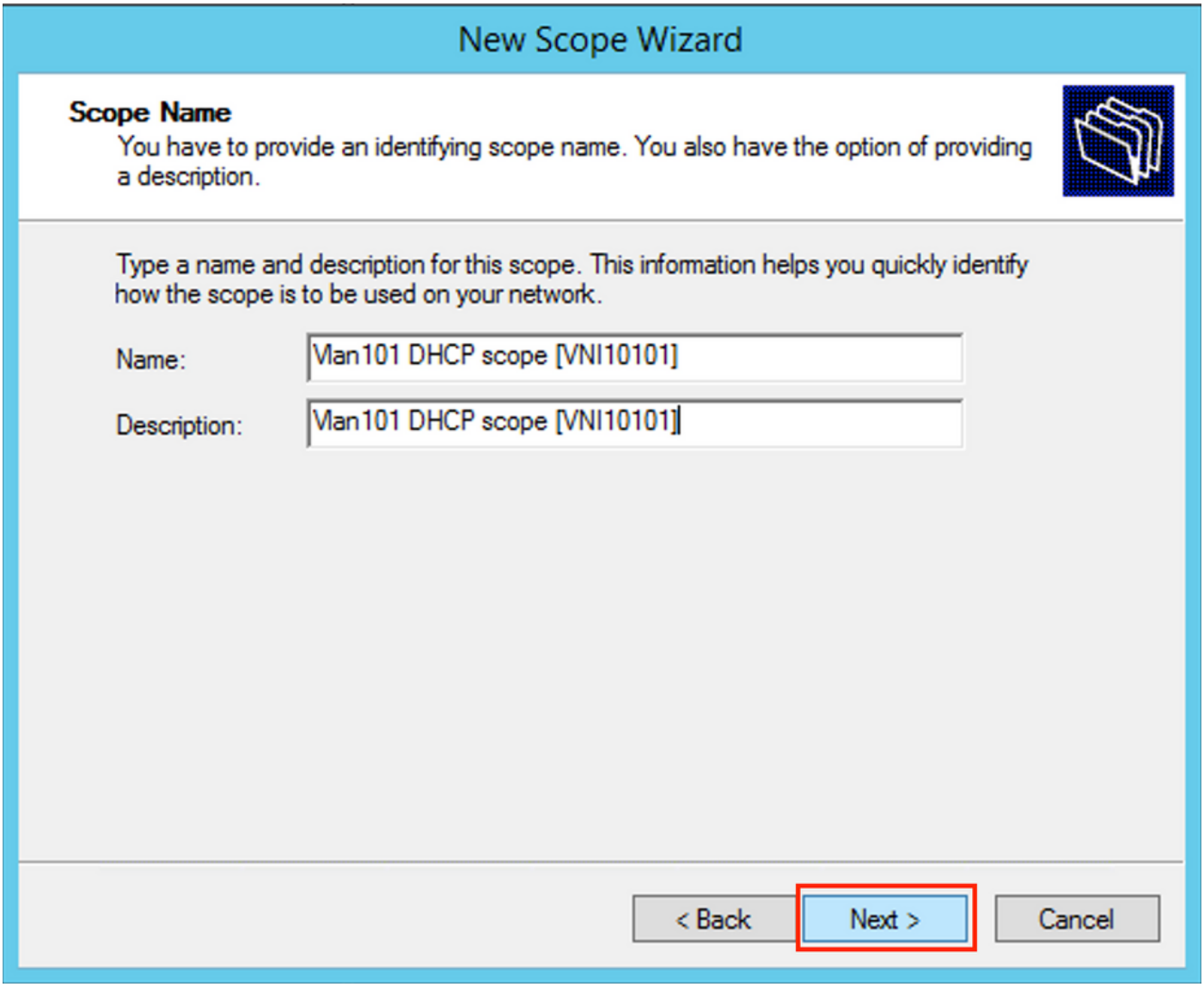
建立從中分配IP地址的DHCP池。按一下右鍵並選擇**新建範圍.....** 如下圖所示。



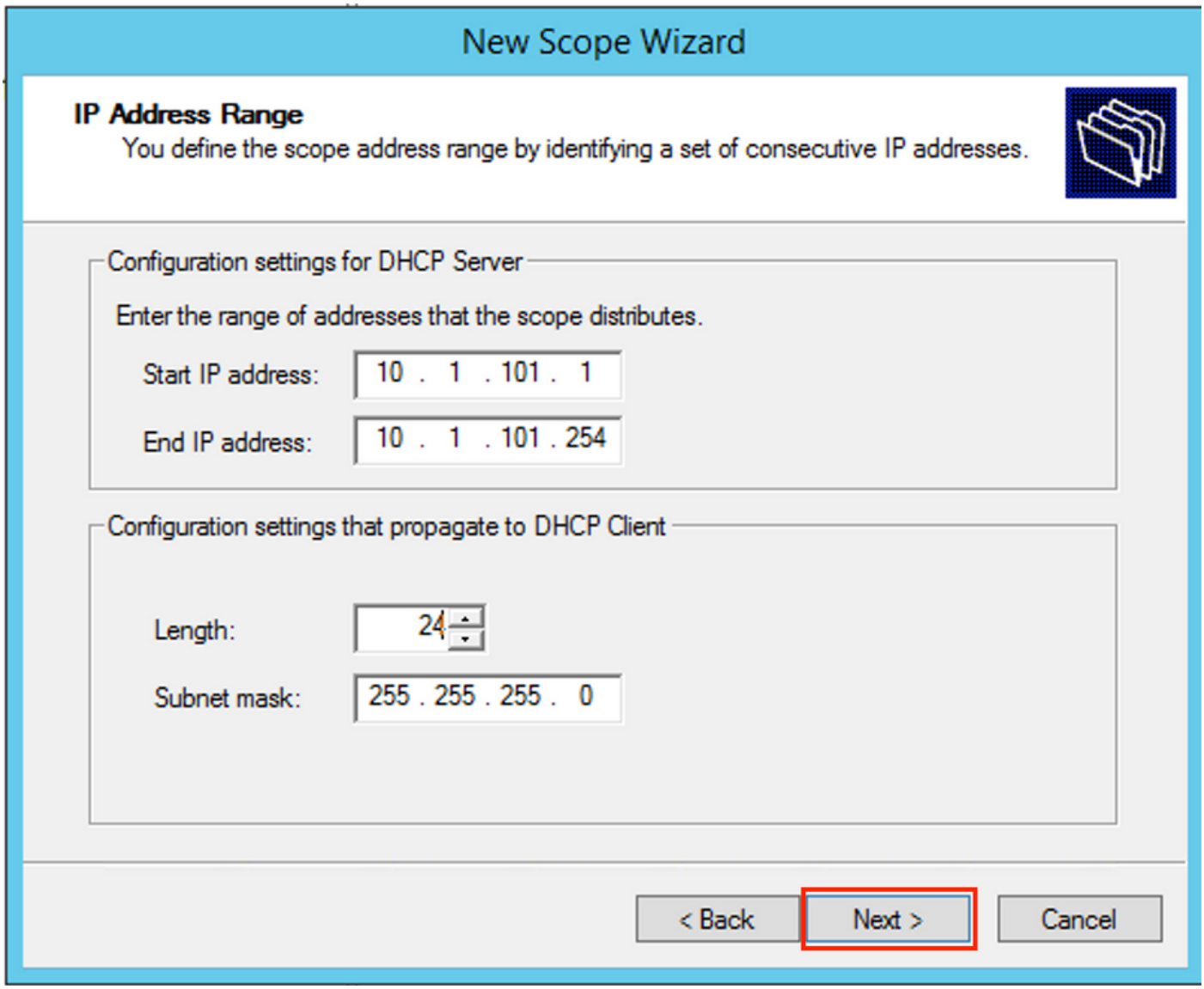
選擇**Next**，如下圖所示。



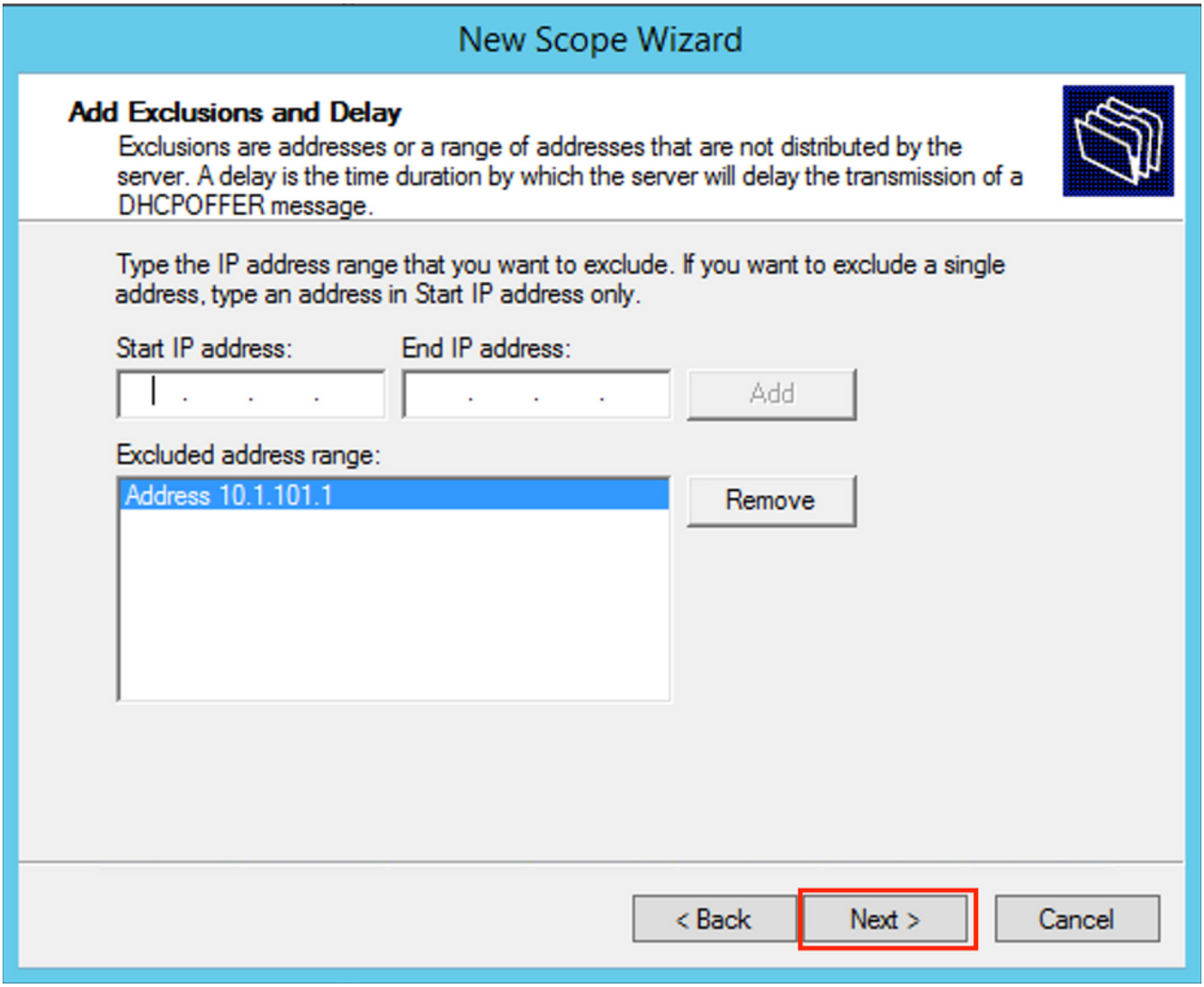
選擇有意義的名稱和描述，如下圖所示。



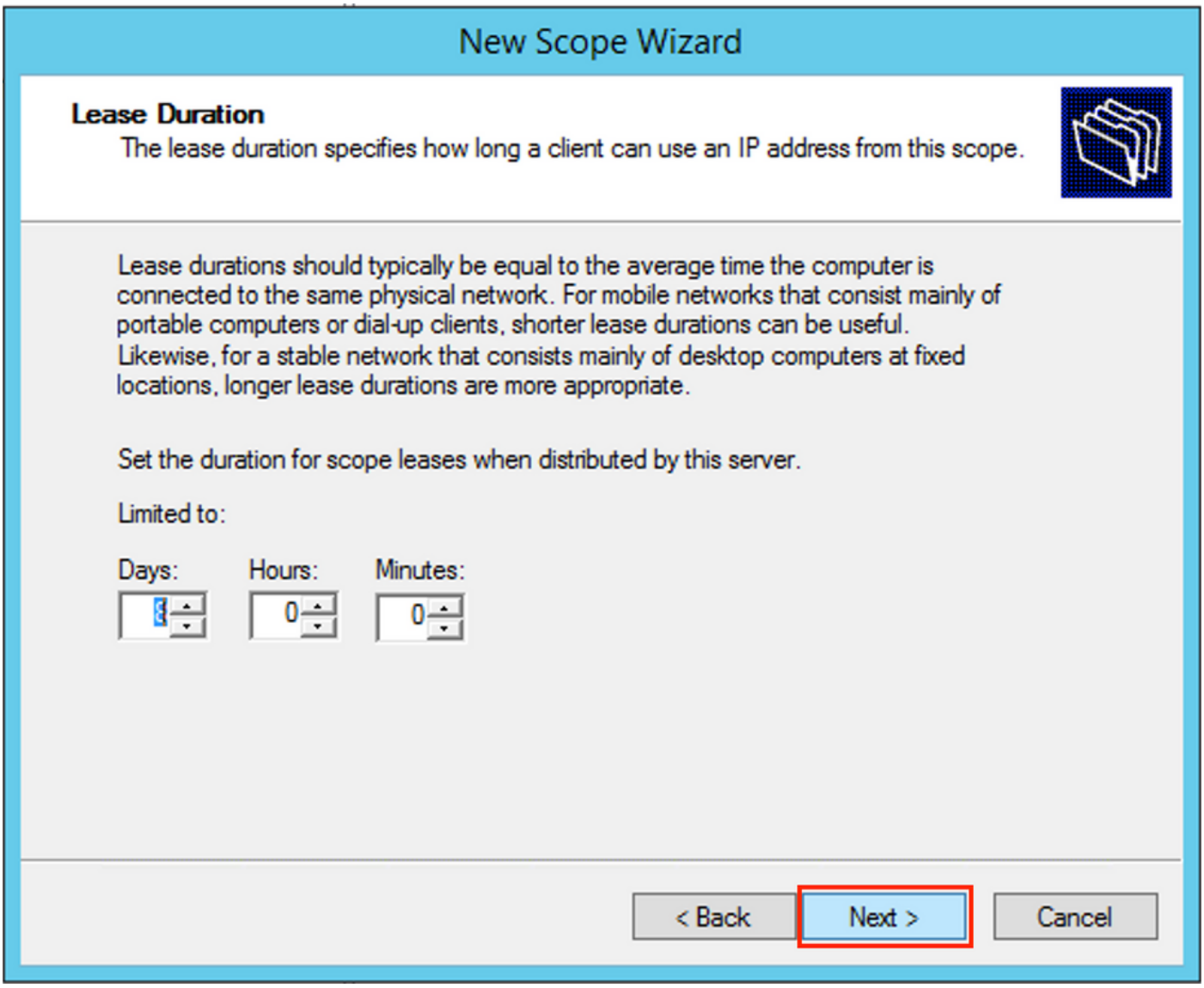
如圖所示，指定要將IP地址分配給客戶端的池的網路和掩碼。



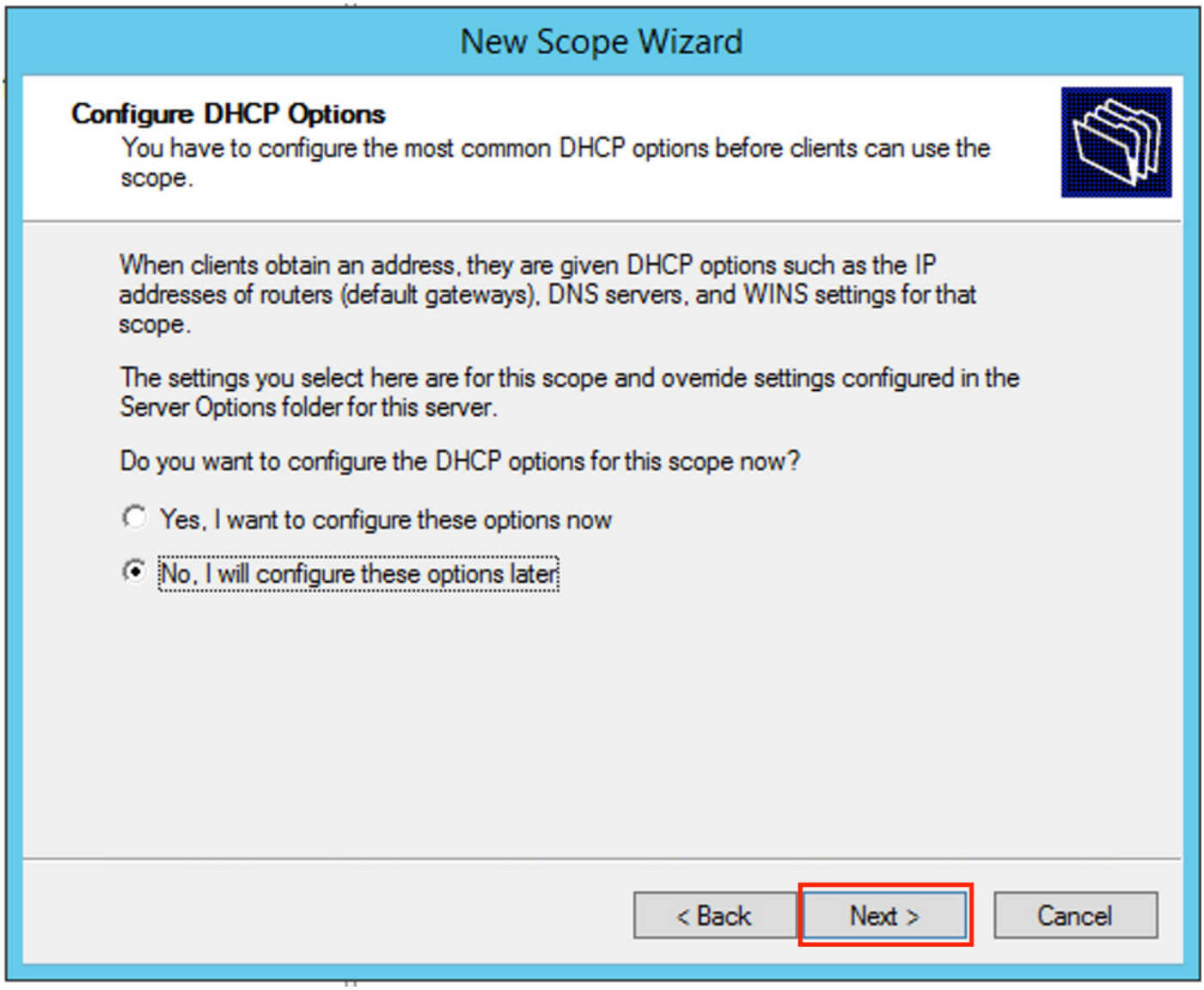
從池中排除預設網關的IP地址（在本例中為10.1.101.1），如下圖所示。



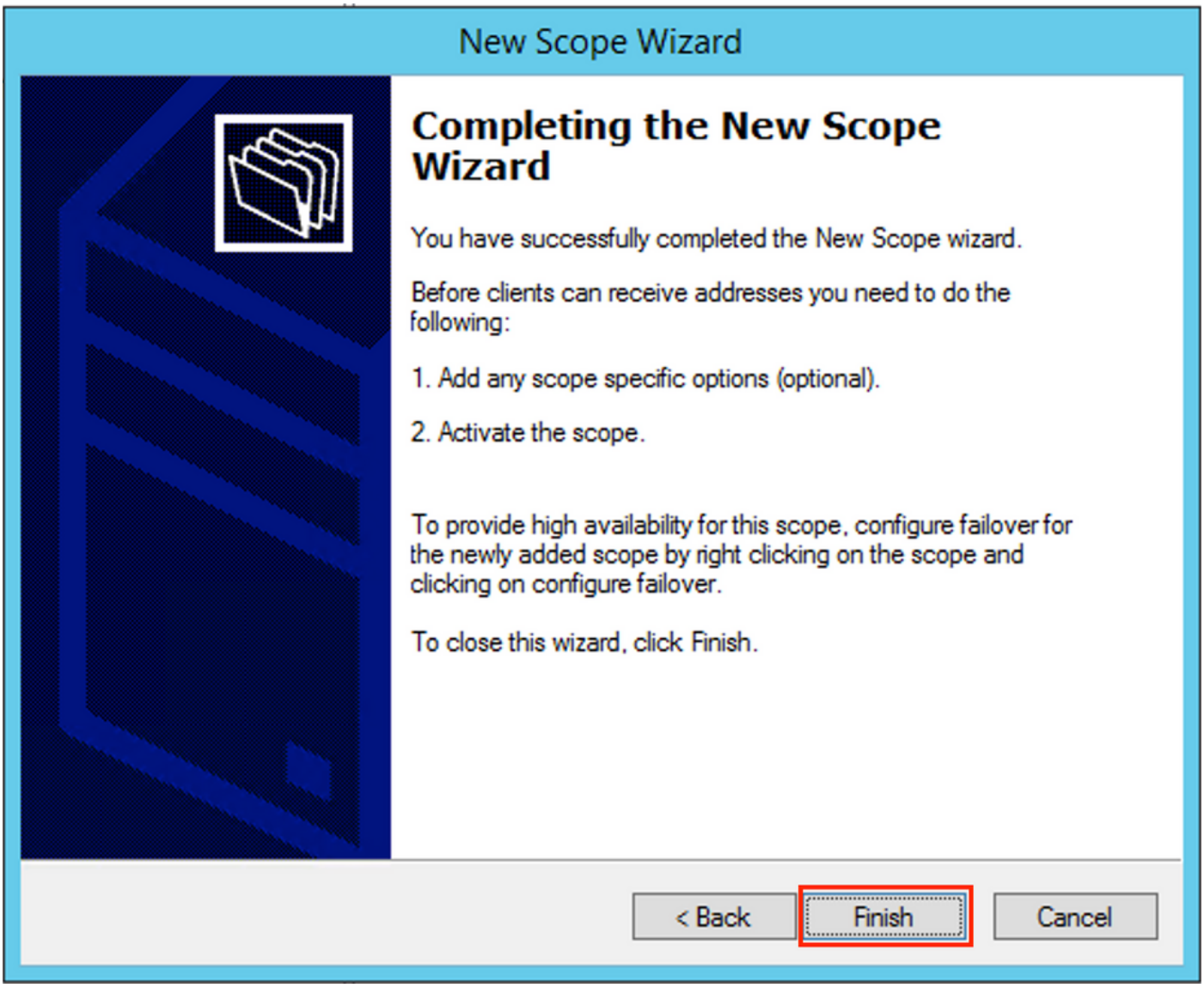
指定租用計時器，如下圖所示。



或者，您可以指定DNS/WINS（在此示例中跳過）。


New Scope Wizard
Configure DHCP Options
You have to configure the most common DHCP options before clients can use the scope.
When clients obtain an address, they are given DHCP options such as the IP addresses of routers (default gateways), DNS servers, and WINS settings for that scope.
The settings you select here are for this scope and override settings configured in the Server Options folder for this server.
Do you want to configure the DHCP options for this scope now?
Yes, I want to configure these options now
No, I will configure these options later
< Back
Next >
Cancel

完成配置，如下圖所示。



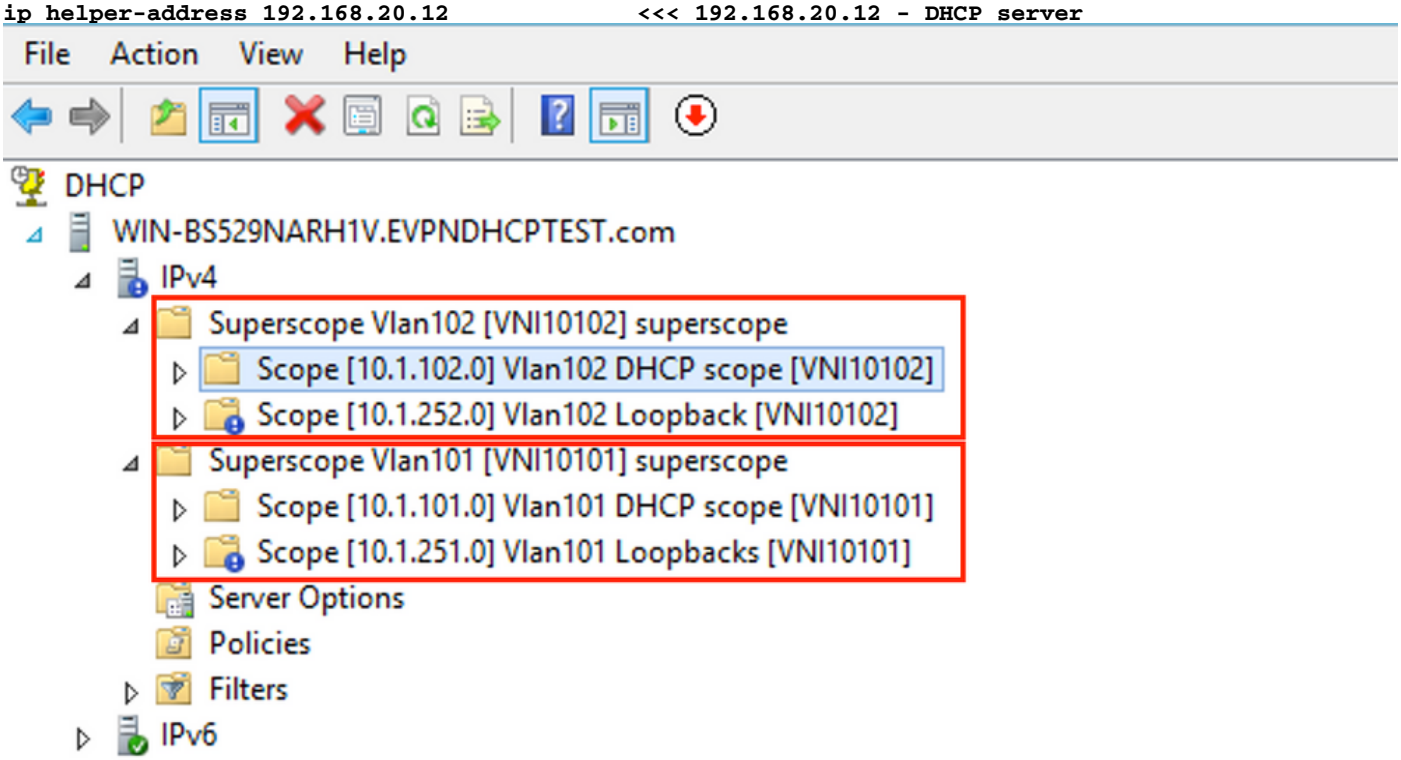
建立池後，必須為池建立策略。

- 在策略中匹配代理電路ID [1]
- 如果您有多個VLAN/VNI，則必須建立具有中繼IP地址子池和每個VLAN/VNI分配的實際IP範圍的超級池
- 此範例使用VNIs 10101和10102

交換機配置：

```
ip dhcp relay information option vpn   <<<  add the VRF name/VPN ID to the option 82
ip dhcp relay information option       <<< enables option 82
!
ip dhcp snooping vlan 101-102,201-202
ip dhcp snooping
!
vlan configuration 101
member evpn-instance 101 vni 10101
!
interface Loopback101
 vrf forwarding green
 ip address 10.1.251.1 255.255.255.255
!
interface Loopback102
 vrf forwarding green
 ip address 10.1.251.2 255.255.255.255
```

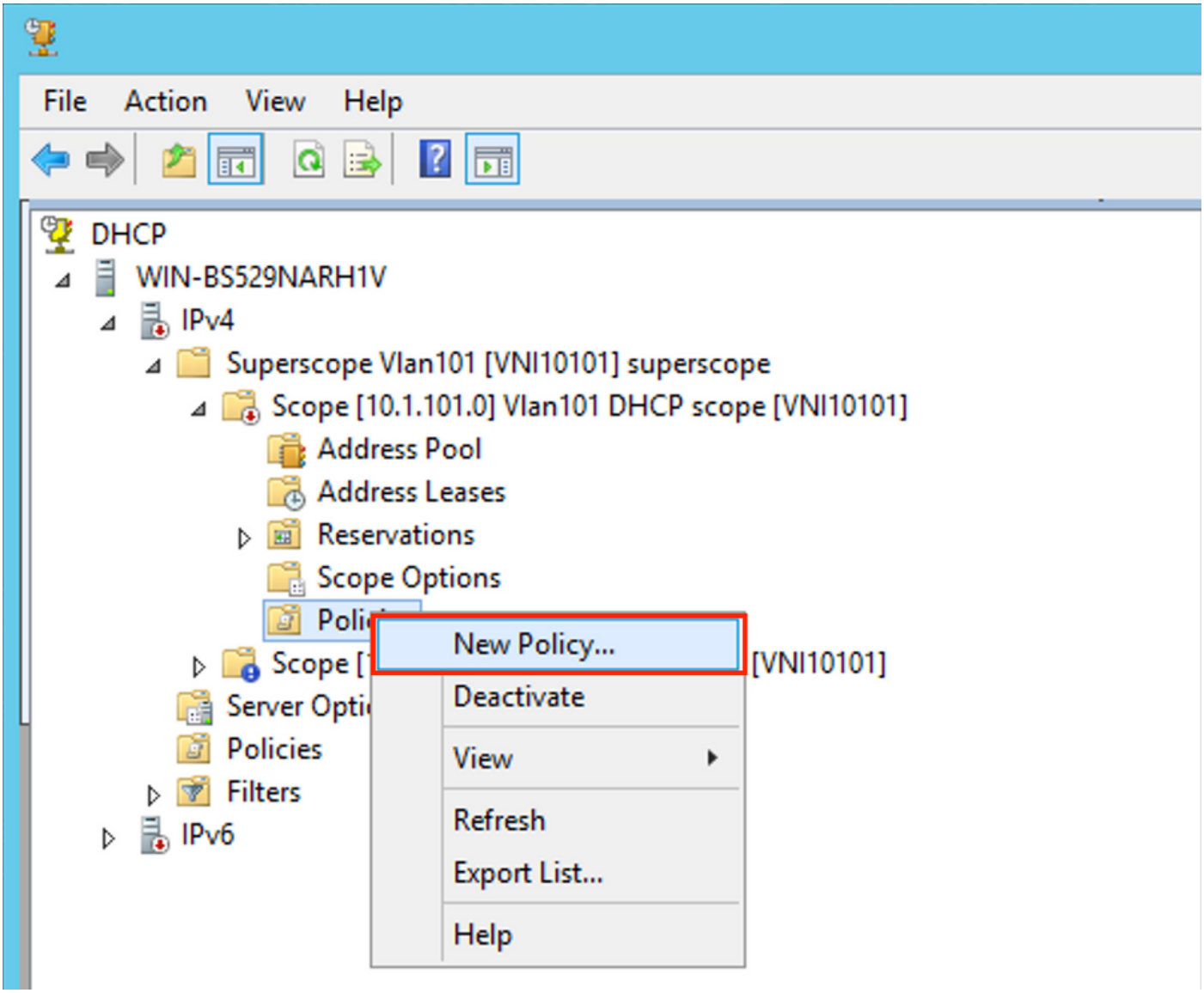
```
!
interface Vlan101
 vrf forwarding green
 ip dhcp relay source-interface Loopback101  <<< DHCP relay source is unique Loopback101
 ip address 10.1.101.1 255.255.255.0
ip helper-address 192.168.20.12              <<< 192.168.20.12 - DHCP server
!
interface Vlan102
 vrf forwarding green
 ip dhcp relay source-interface Loopback102  <<< DHCP relay source is unique Loopback102
 ip address 10.1.101.1 255.255.255.0
ip helper-address 192.168.20.12              <<< 192.168.20.12 - DHCP server
```



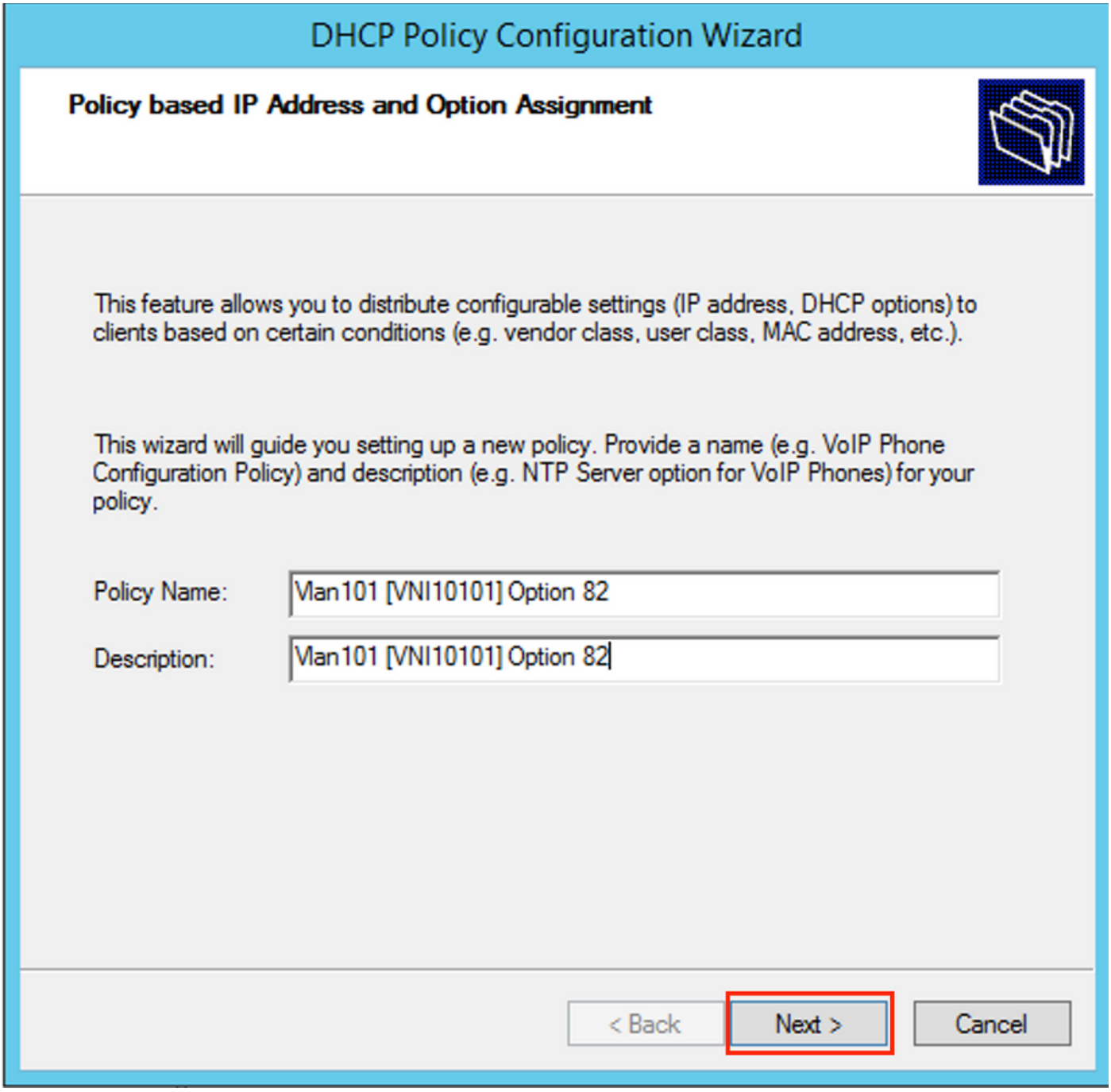
Win2012 R2配置選項2 — 匹配代理電路ID欄位

- 最後一個方法的缺點是唯一的環回的利用率很高，因此另一個選項是匹配代理電路ID欄位。
- 步驟是相同的，但您為範圍選擇新增策略時，不是基於Agent Circuit ID欄位而不是中繼IP。

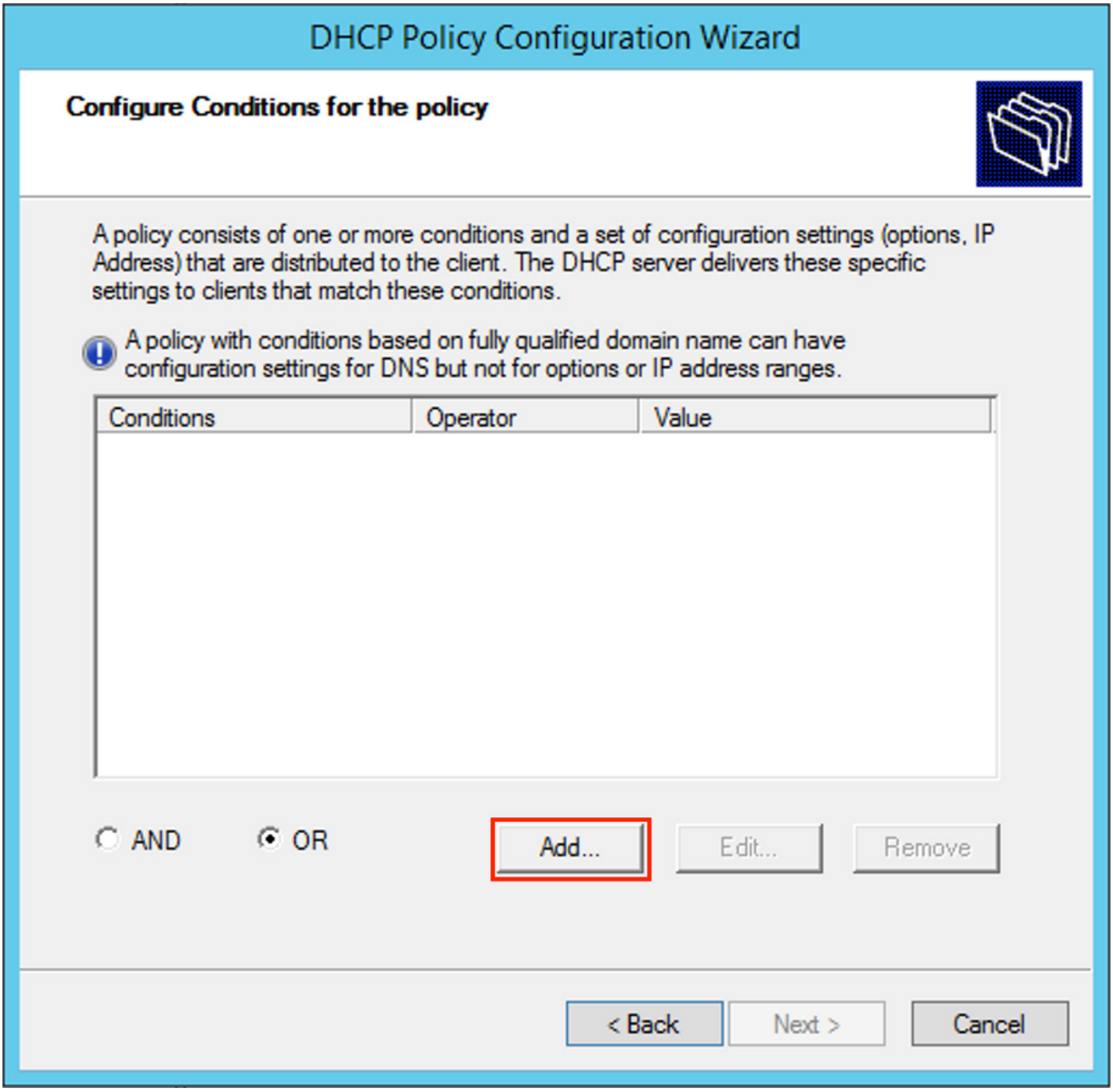
策略建立。按一下右鍵池並選擇**New Policy**，如下圖所示。



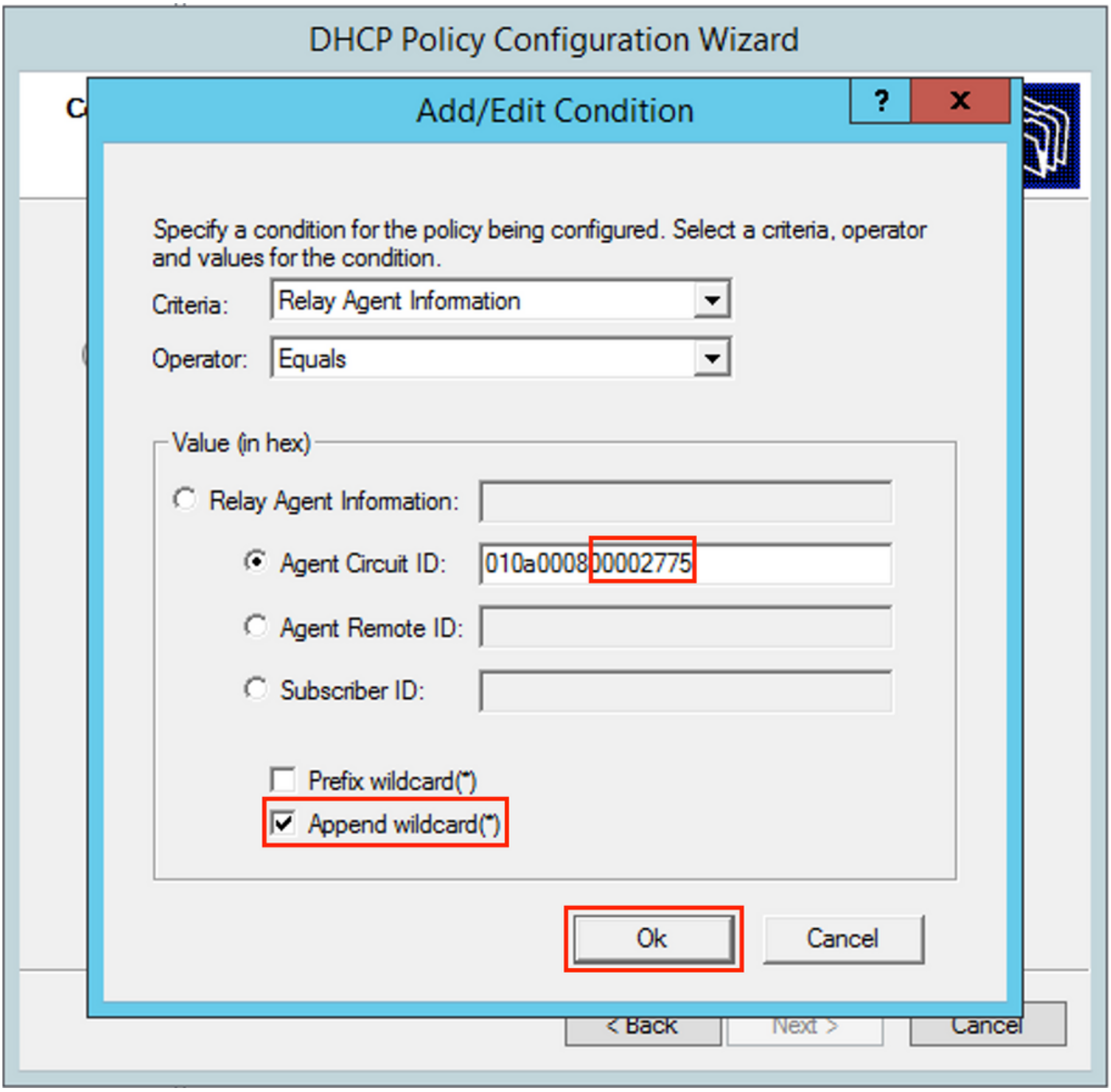
為策略選擇有意義的名稱和說明，如下圖所示。


DHCP Policy Configuration Wizard
Policy based IP Address and Option Assignment
This feature allows you to distribute configurable settings (IP address, DHCP options) to clients based on certain conditions (e.g. vendor class, user class, MAC address, etc.).
This wizard will guide you setting up a new policy. Provide a name (e.g. VoIP Phone Configuration Policy) and description (e.g. NTP Server option for VoIP Phones) for your policy.
Policy Name:
Vlan101 [VNI10101] Option 82
Description:
Vlan101 [VNI10101] Option 82
< Back
Next >
Cancel

新增新條件，如下圖所示。



輸入正確的電路ID(不要忘記**附加萬用字元**(*)框)，如下圖所示。



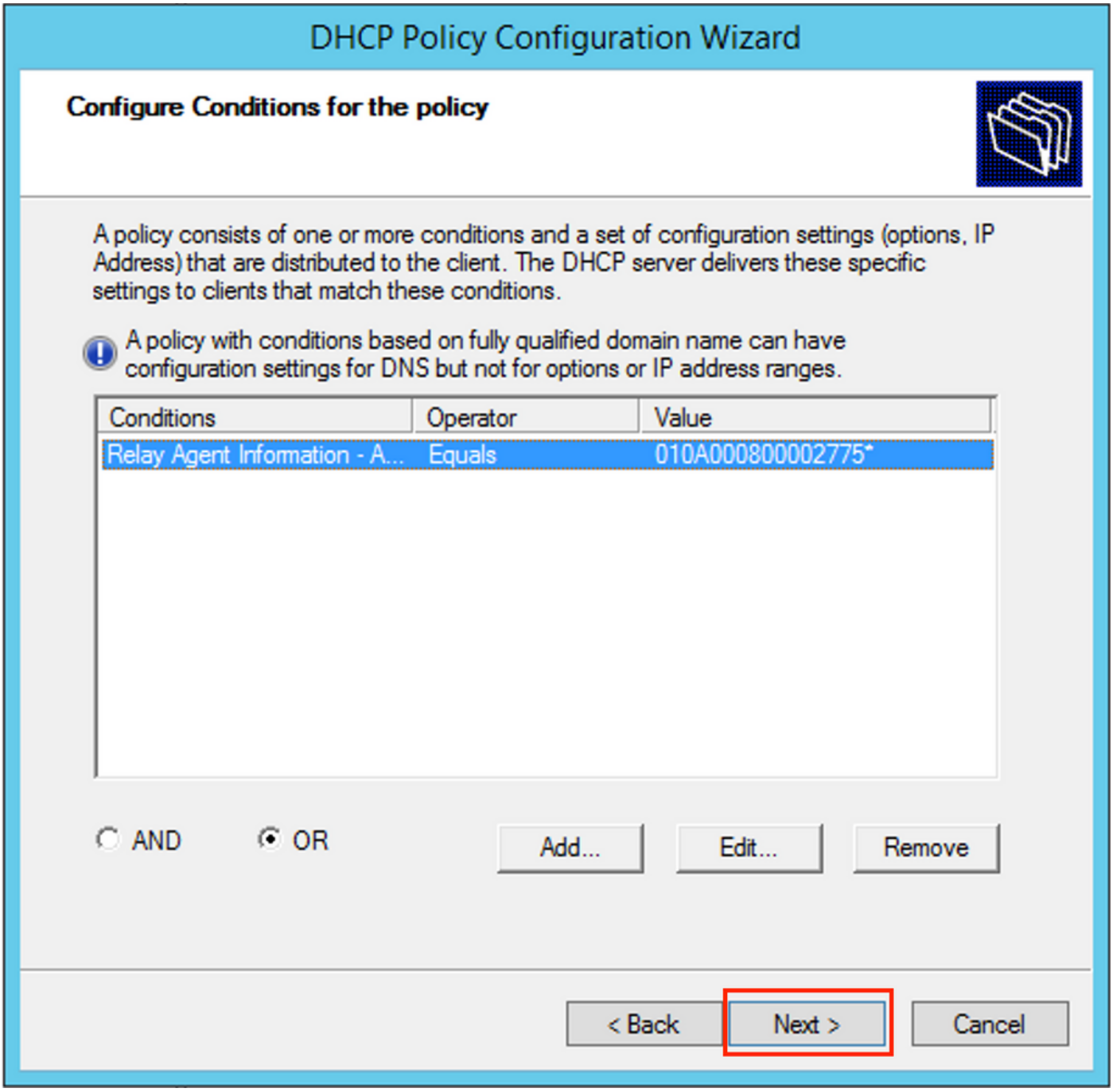
關於為什麼選擇此數字的說明：

在Wireshark中，可以看到代理電路ID等於**010a000800002775010a000**，該值是從這裡派生的(00002775十六進位制= 10101十進位制等於為VLAN 101配置的VNI 10101)。

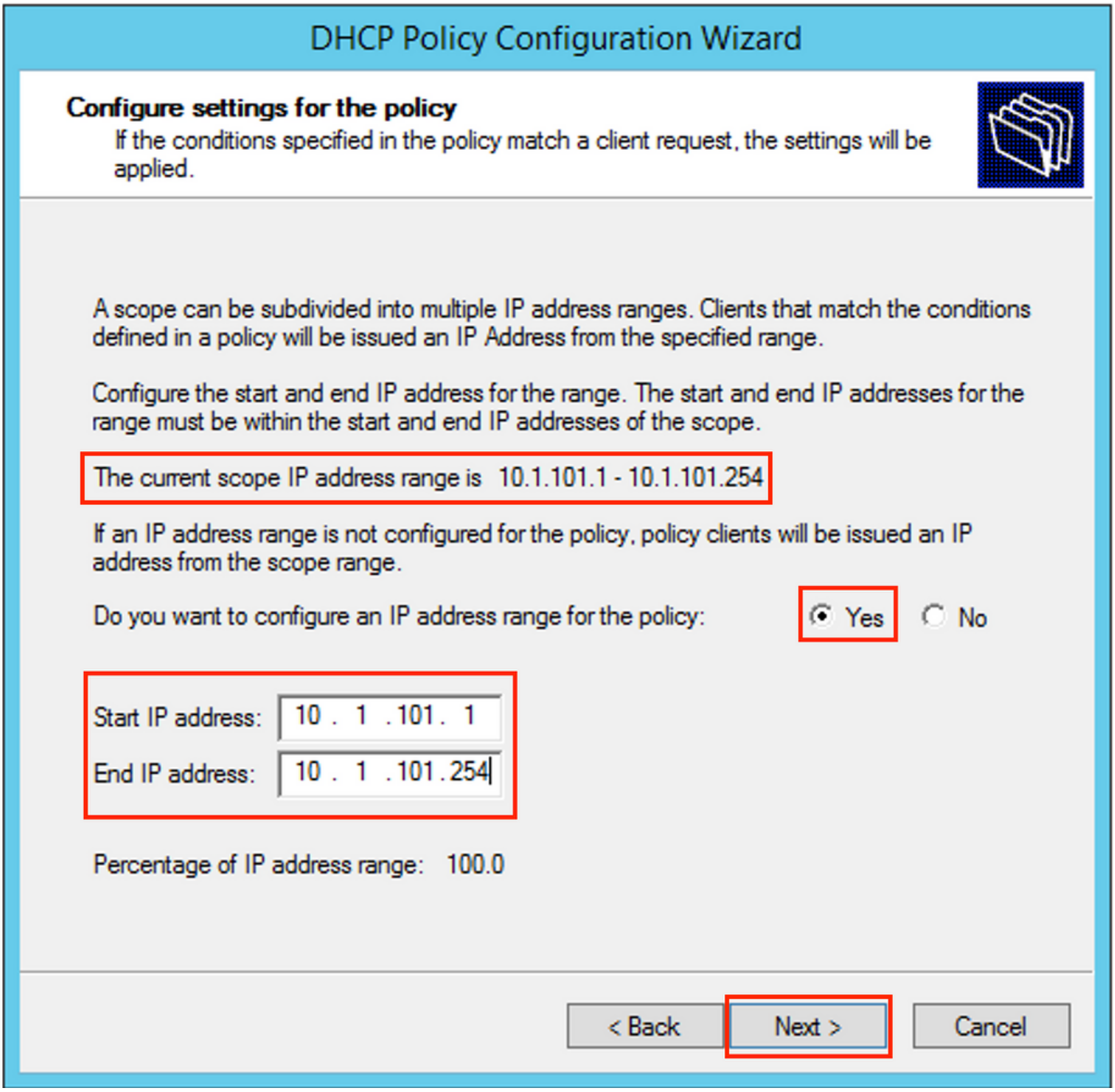
```
▼ Option: (82) Agent Information Option
    Length: 44
  ▼ Option 82 Suboption: (1) Agent Circuit ID
      Length: 12
      Agent Circuit ID: 010a000800002775010a0000
  ▶ Option 82 Suboption: (2) Agent Remote ID
  ▶ Option 82 Suboption: (151) VRF name/VPN ID
  ▼ Option 82 Suboption: (150) Link selection (Cisco proprietary)
      Length: 4
      Link selection (Cisco proprietary): 10.1.101.0
  ▼ Option 82 Suboption: (152) Server ID Override (Cisco proprietary)
      Length: 4
      Server ID Override (Cisco proprietary): 10.1.101.1
```

VXLAN VN的代理電路ID子選項按以下格式編碼：

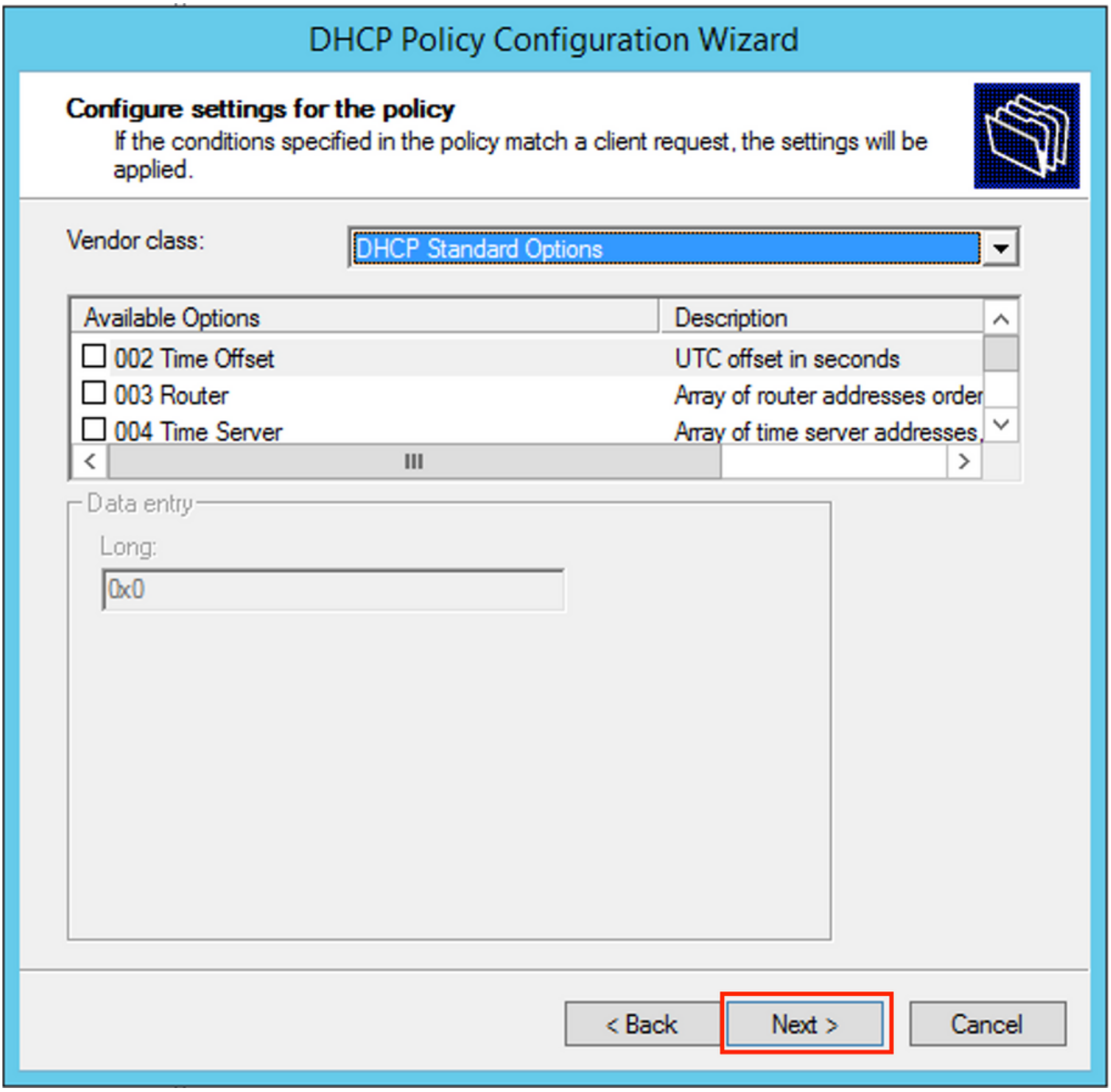
子選項型別	長度	電路ID型別	長度	VNI	mod	連接埠
1位元組	1位元組	1位元組	1位元組	4 位元組	2位元組	2位元組
01	0a	00	08	00002775	*	*



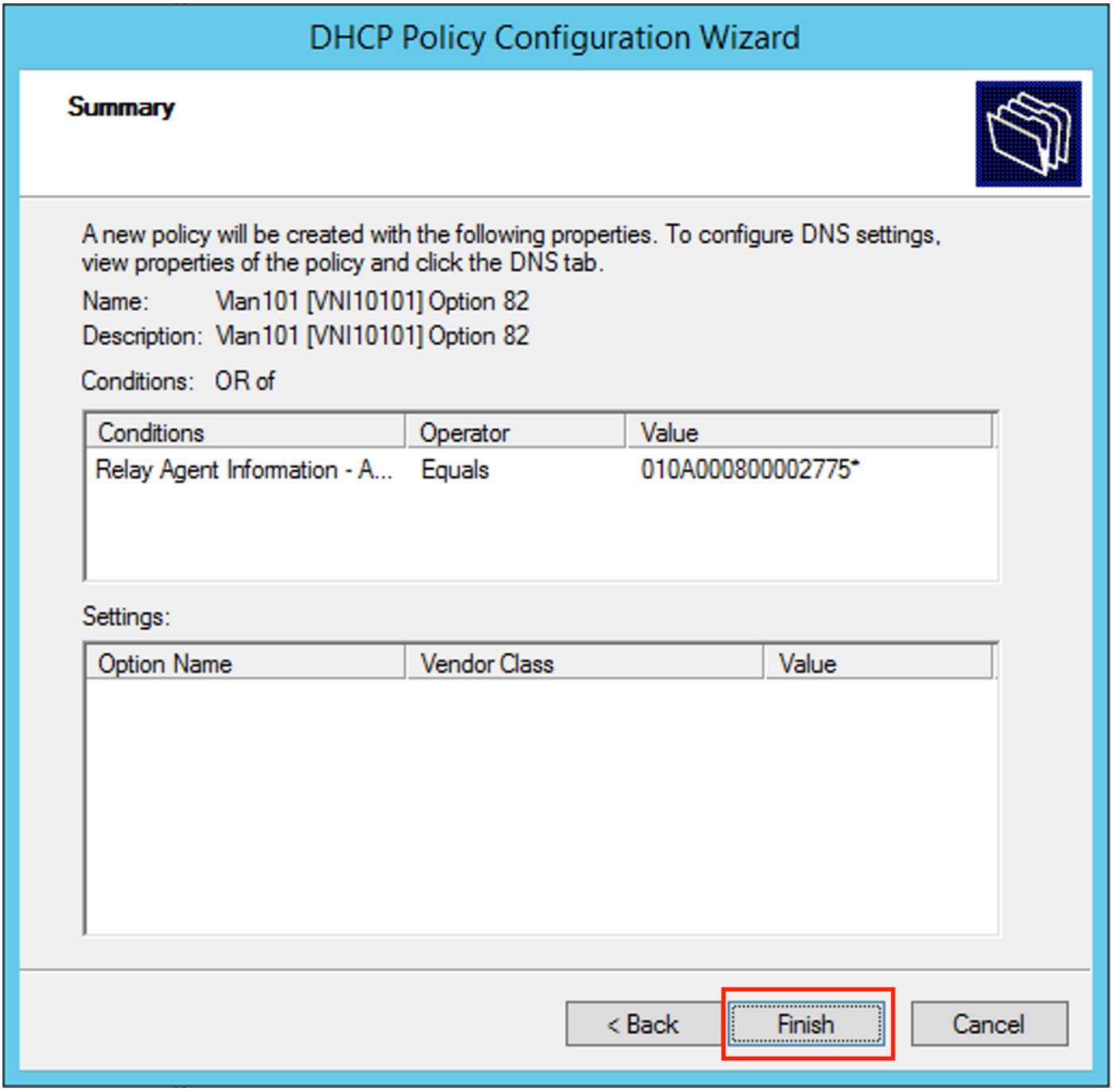
配置從中分配IP地址的IP範圍。如果沒有此配置，則無法**分配當**前範圍。



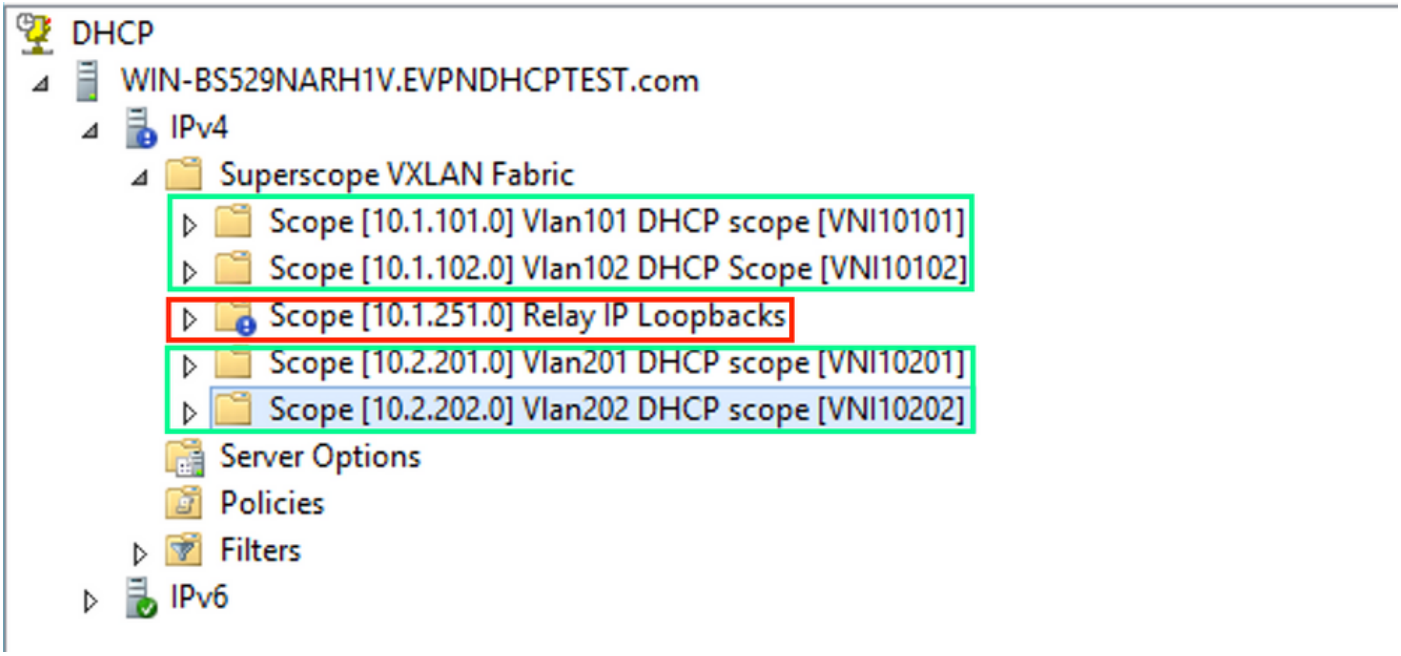
您還可以在此階段選擇標準DHCP選項，如下圖所示。


DHCP Policy Configuration Wizard
Configure settings for the policy
If the conditions specified in the policy match a client request, the settings will be applied.
Vendor class:
DHCP Standard Options
Available Options
Description
002 Time Offset
UTC offset in seconds
003 Router
Array of router addresses order
004 Time Server
Array of time server addresses
Data entry
Long:
0x0
< Back
Next >
Cancel

選擇**完成**，如下圖所示。



必須為其他範圍執行類似的配置，如下圖所示。



在此案例中，每個VTEP只能對SVI的數量使用一個唯一的IP地址，而不能對每個VNI/SVI每個VTEP使用一個唯一的環回。

交換機配置：

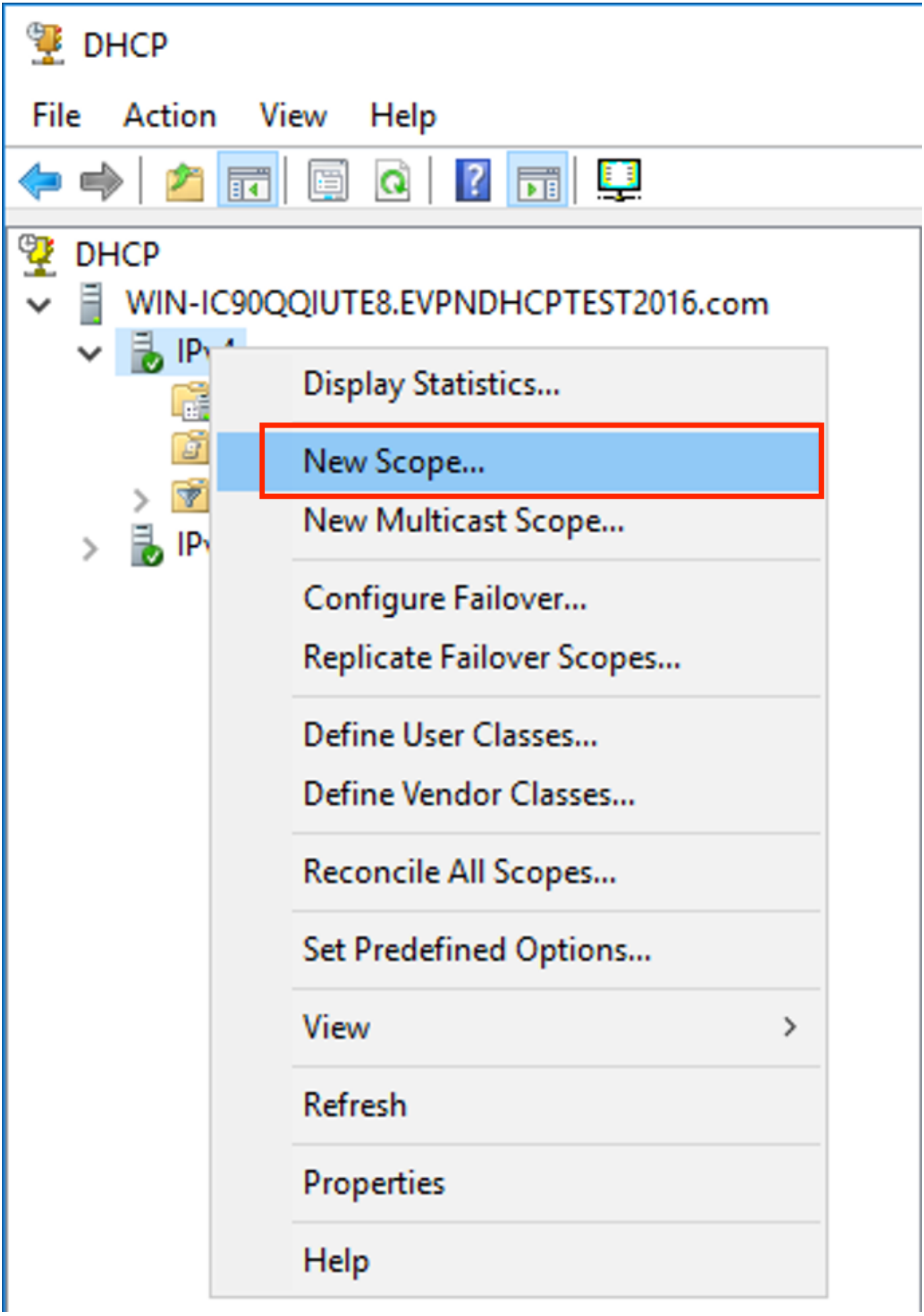
```
ip dhcp relay information option vpn     <<<  adds the VRF name/VPN ID to the option 82
ip dhcp relay information option         <<< enables option 82
!
ip dhcp snooping vlan 101-102,201-202
ip dhcp snooping
!
vlan configuration 101
member evpn-instance 101 vni 10101
!
interface Loopback101
 vrf forwarding green
 ip address 10.1.251.1 255.255.255.255
!
interface Vlan101
 vrf forwarding green
 ip dhcp relay source-interface Loopback101 <<< DHCP relay source
 ip address 10.1.101.1 255.255.255.0
 ip helper-address 192.168.20.12            <<< 192.168.20.12 - DHCP server
!
interface Vlan102
 vrf forwarding green
 ip dhcp relay source-interface Loopback101 <<< DHCP relay source
 ip address 10.1.101.1 255.255.255.0
ip helper-address 192.168.20.12           <<< 192.168.20.12 - DHCP server
```

Windows Server 2016配置

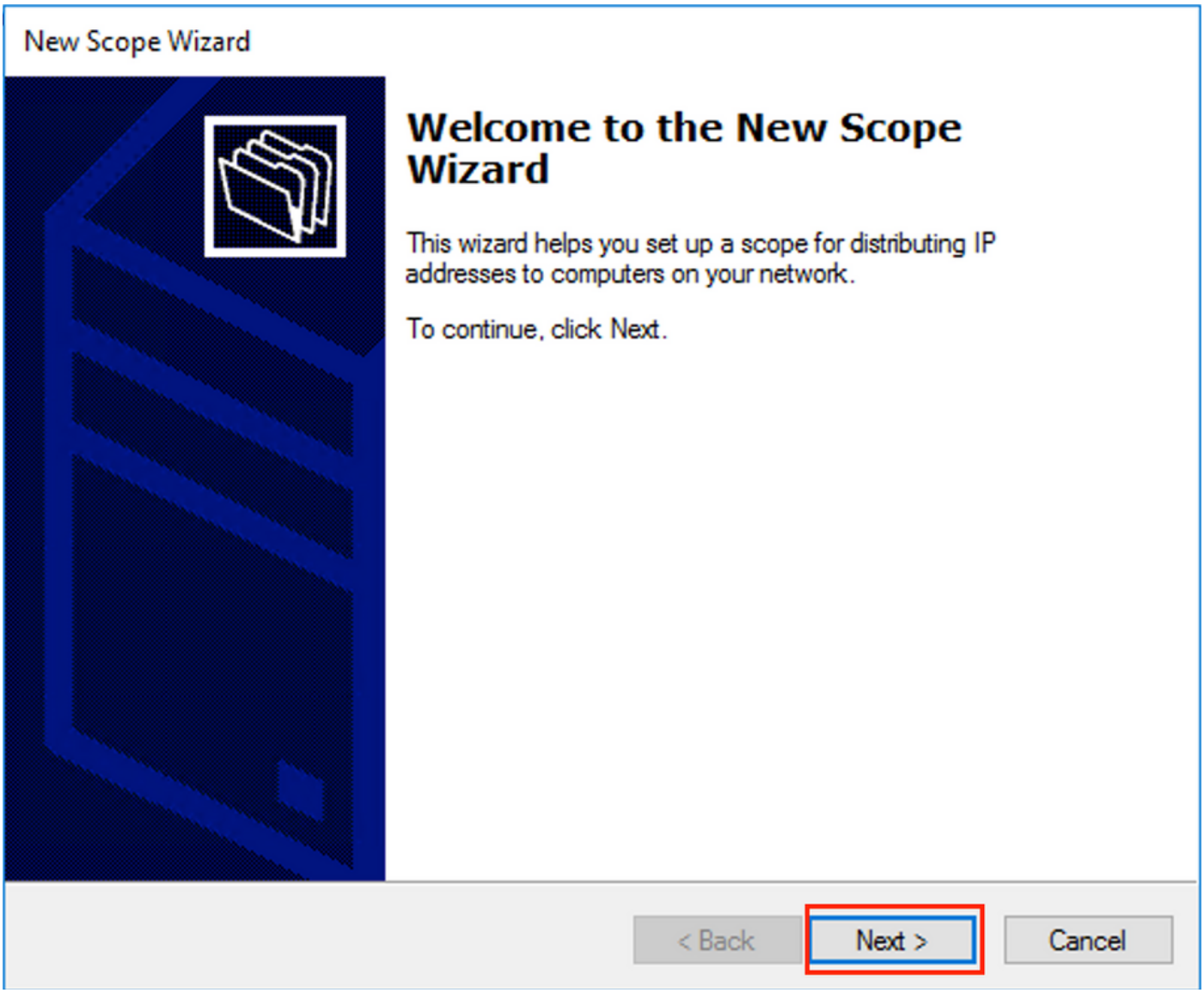
- Windows Server 2016支援選項82子選項5（Cisco專有150）「鏈路選擇」，這意味著您不會使用唯一的中繼IP地址進行池選擇。而是使用「Link selection」子選項，從而顯著簡化配置。
- 最好您仍然有一個中繼IP地址池，否則DHCP資料包與任何作用域不匹配，不會進行處理。

此示例演示了「鏈路選擇」選項的用法。

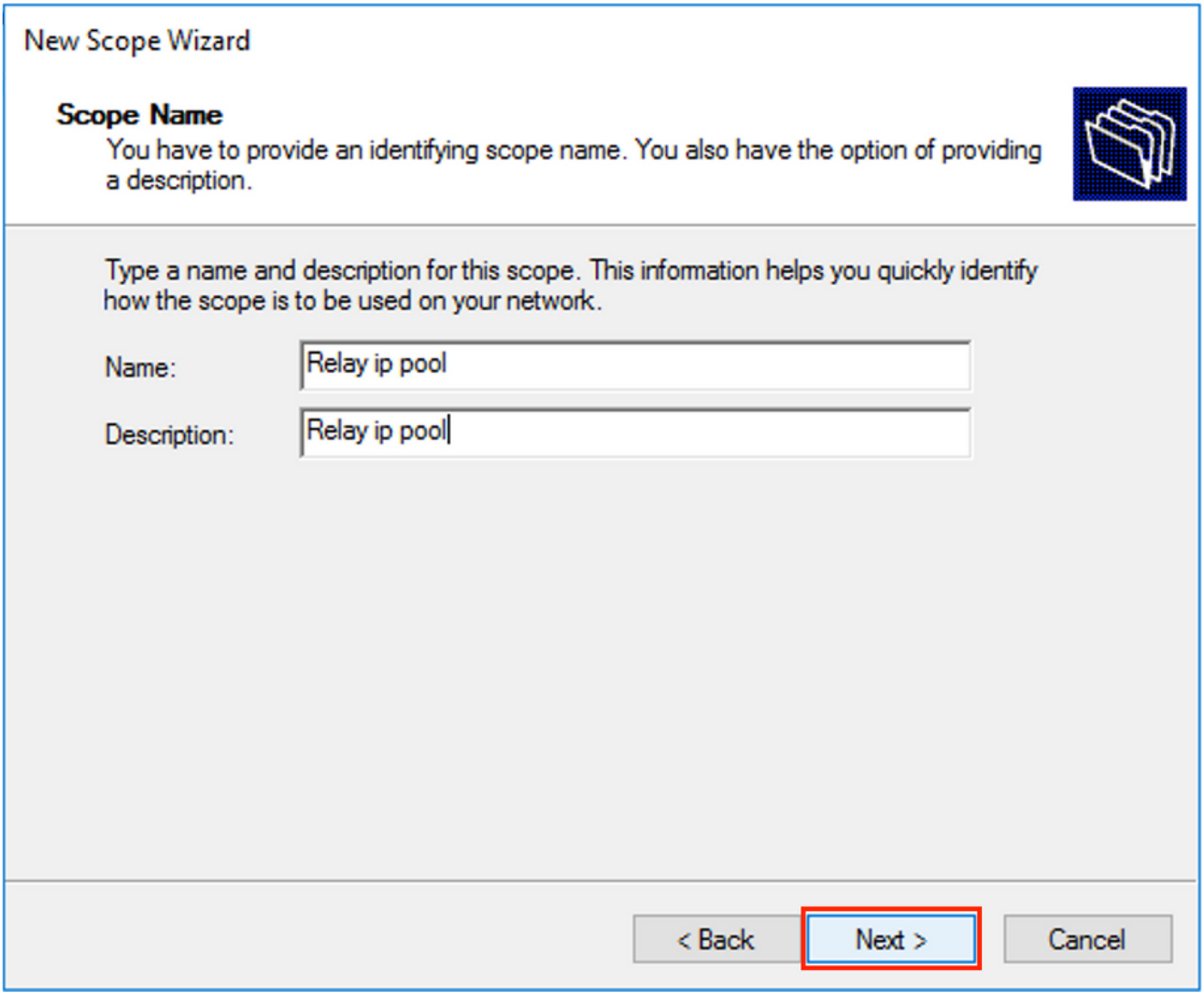
啟動中繼IP地址的IP地址池，如下圖所示。


DHCP
File
Action
View
Help
DHCP
WIN-IC90QQIUTE8.EVPNDHCPTEST2016.com
IPv
IPv
Display Statistics...
New Scope...
New Multicast Scope...
Configure Failover...
Replicate Failover Scopes...
Define User Classes...
Define Vendor Classes...
Reconcile All Scopes...
Set Predefined Options...
View
Refresh
Properties
Help

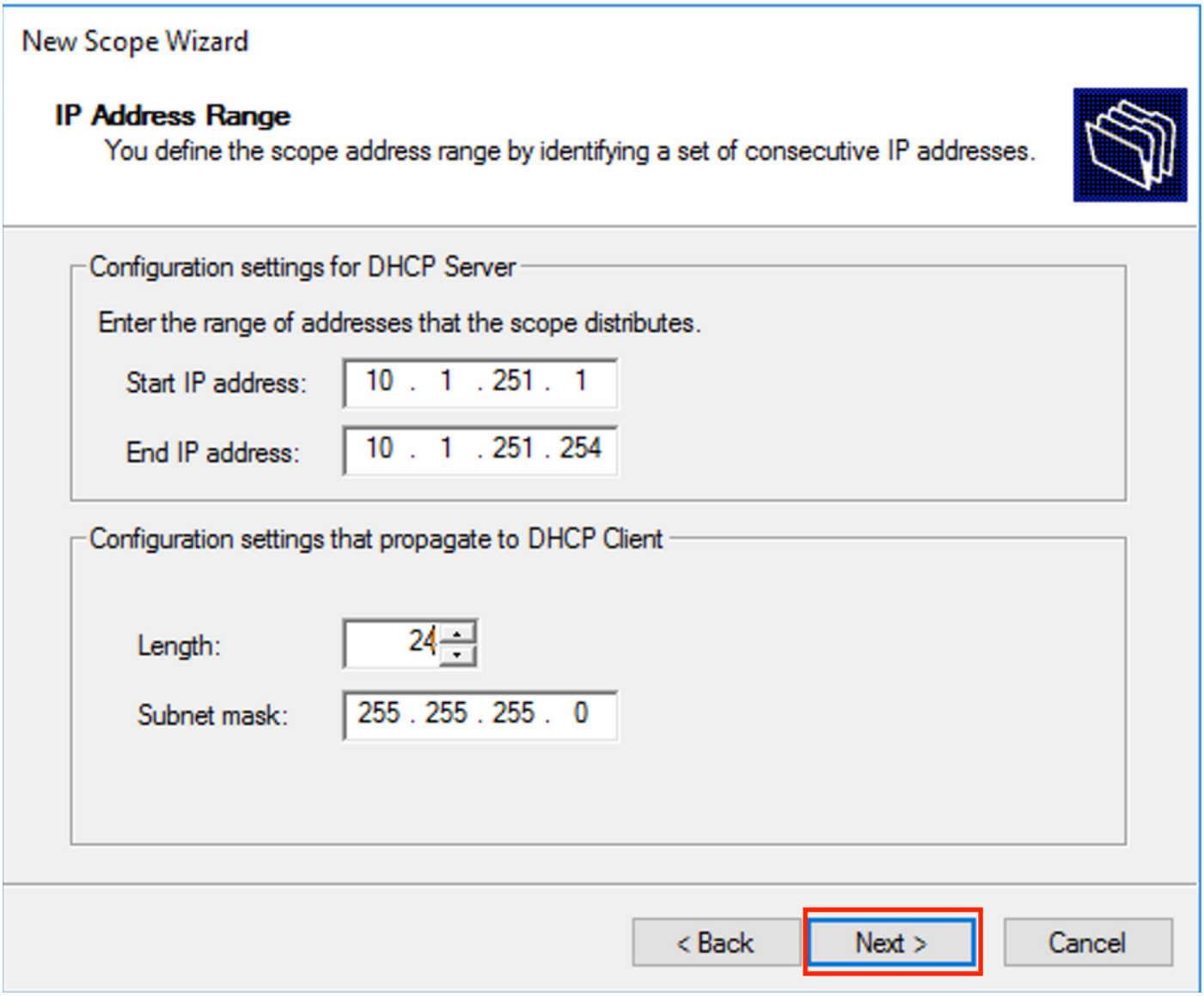
選擇**Next**，如下圖所示。



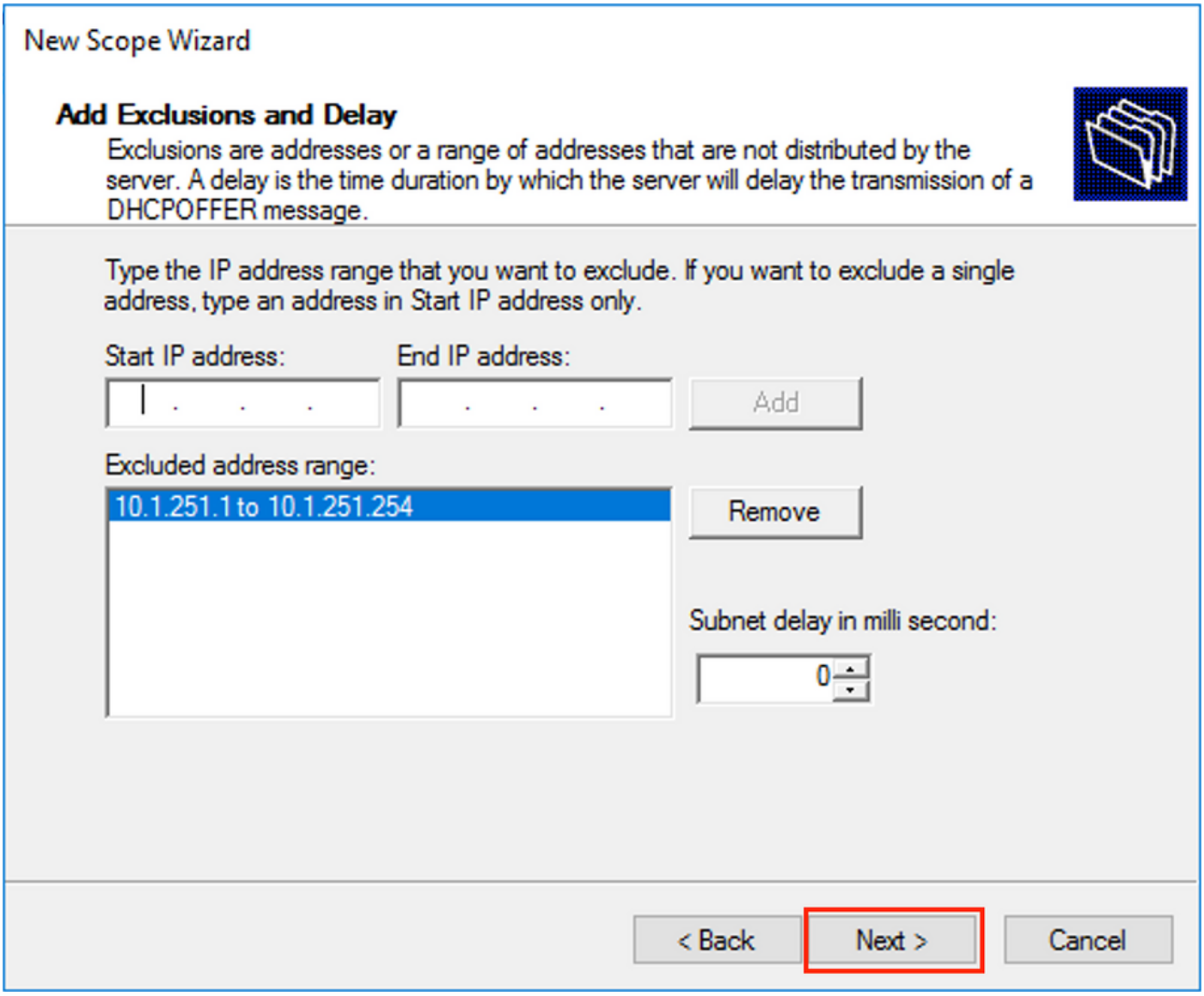
為範圍選擇有意義的名稱和說明，如下圖所示。



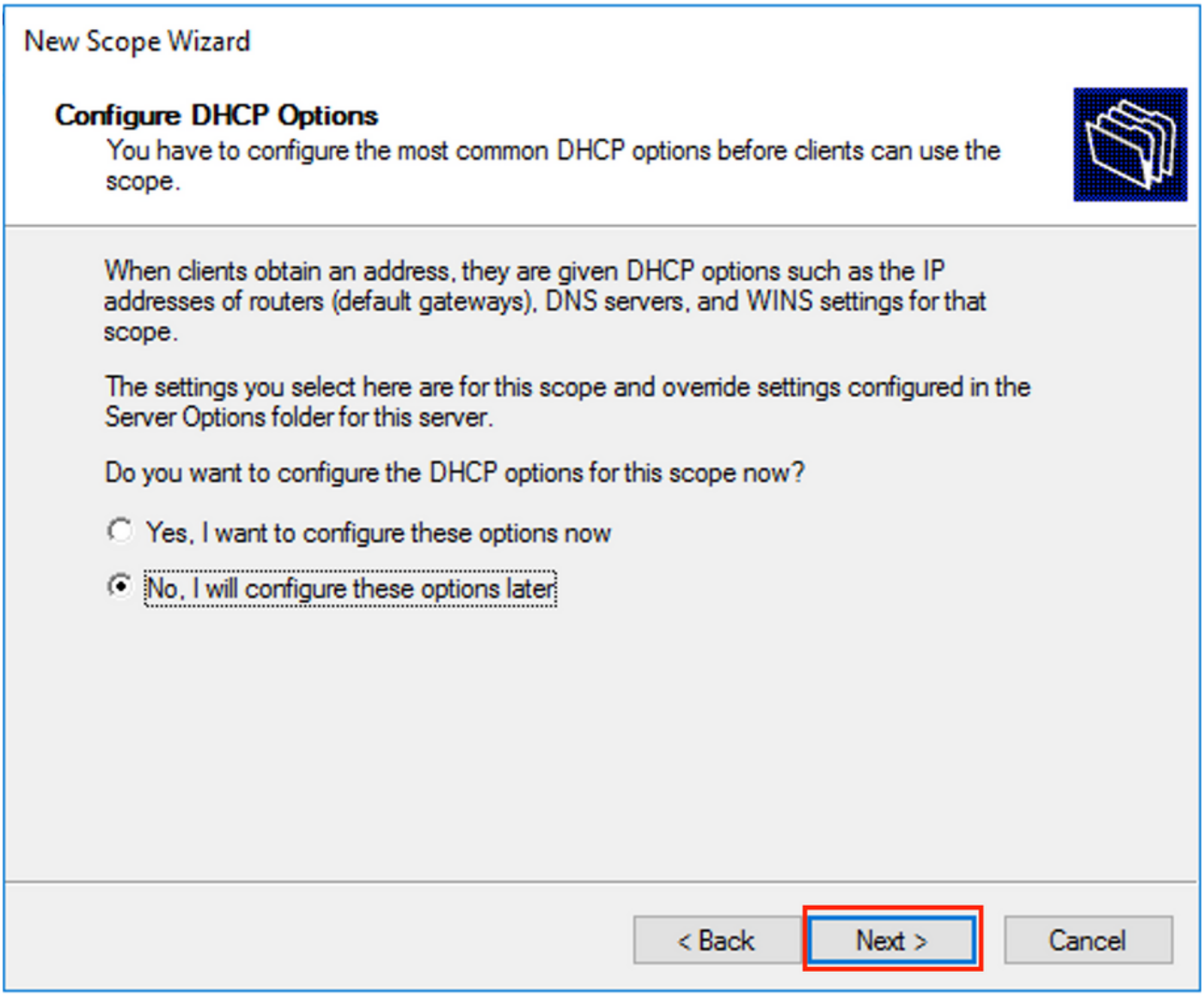
輸入用於IP中繼的IP地址空間，如下圖所示。



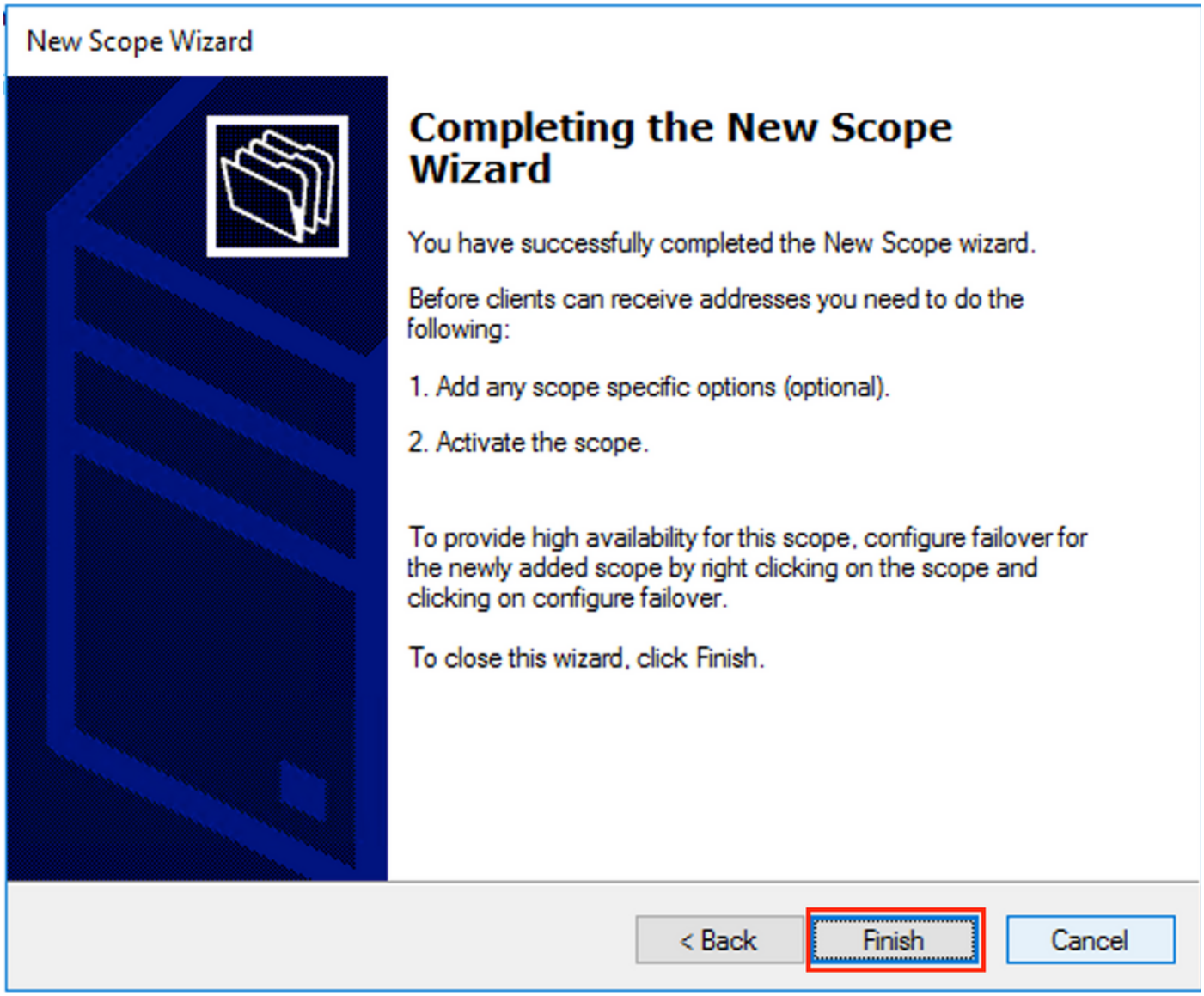
從作用域中排除所有範圍，以阻止從此範圍中進行分配，如下圖所示。



您也可以選擇選項DNS/WINS等引數（在此範例中跳過），如下圖所示。

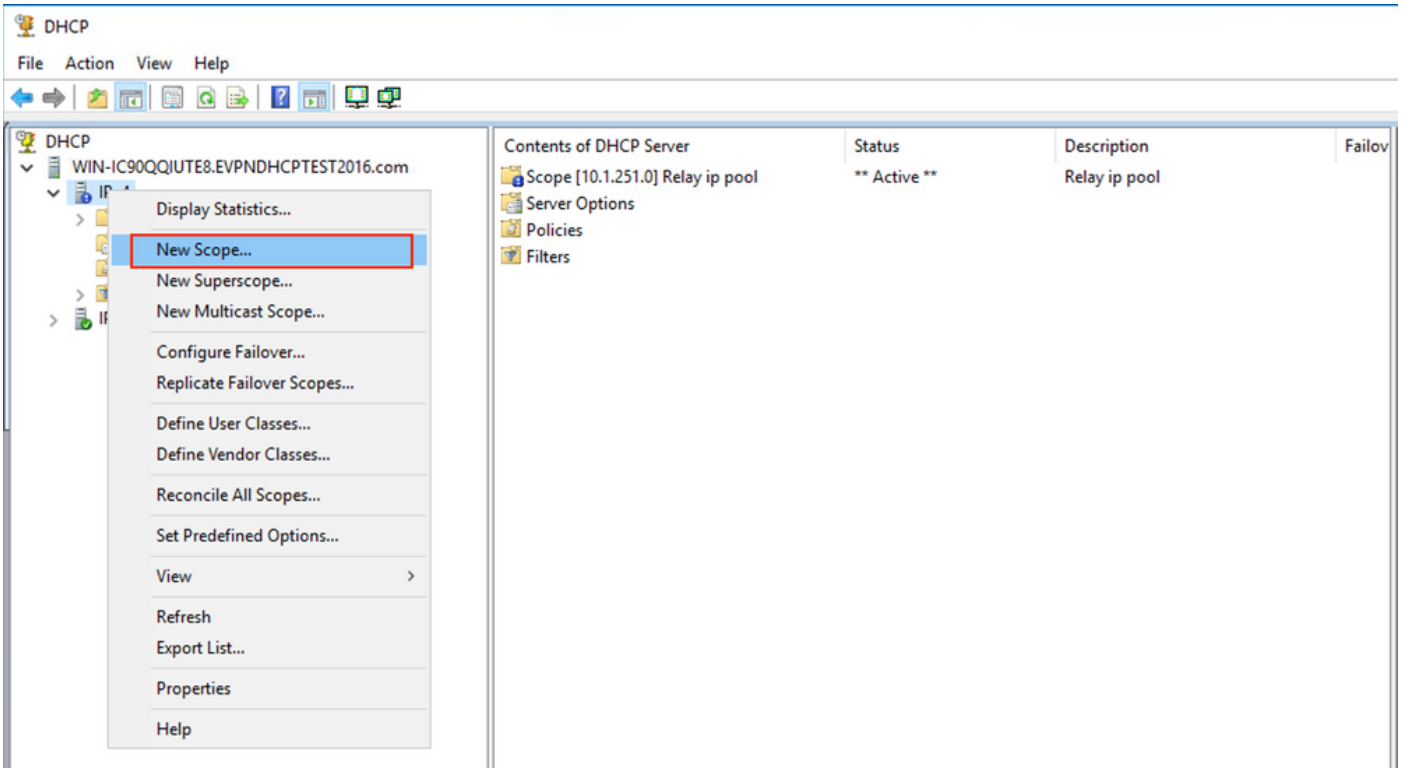

New Scope Wizard
Configure DHCP Options
You have to configure the most common DHCP options before clients can use the scope.
When clients obtain an address, they are given DHCP options such as the IP addresses of routers (default gateways), DNS servers, and WINS settings for that scope.
The settings you select here are for this scope and override settings configured in the Server Options folder for this server.
Do you want to configure the DHCP options for this scope now?
Yes, I want to configure these options now
No, I will configure these options later
< Back
Next >
Cancel

選擇**完成**，如下圖所示。

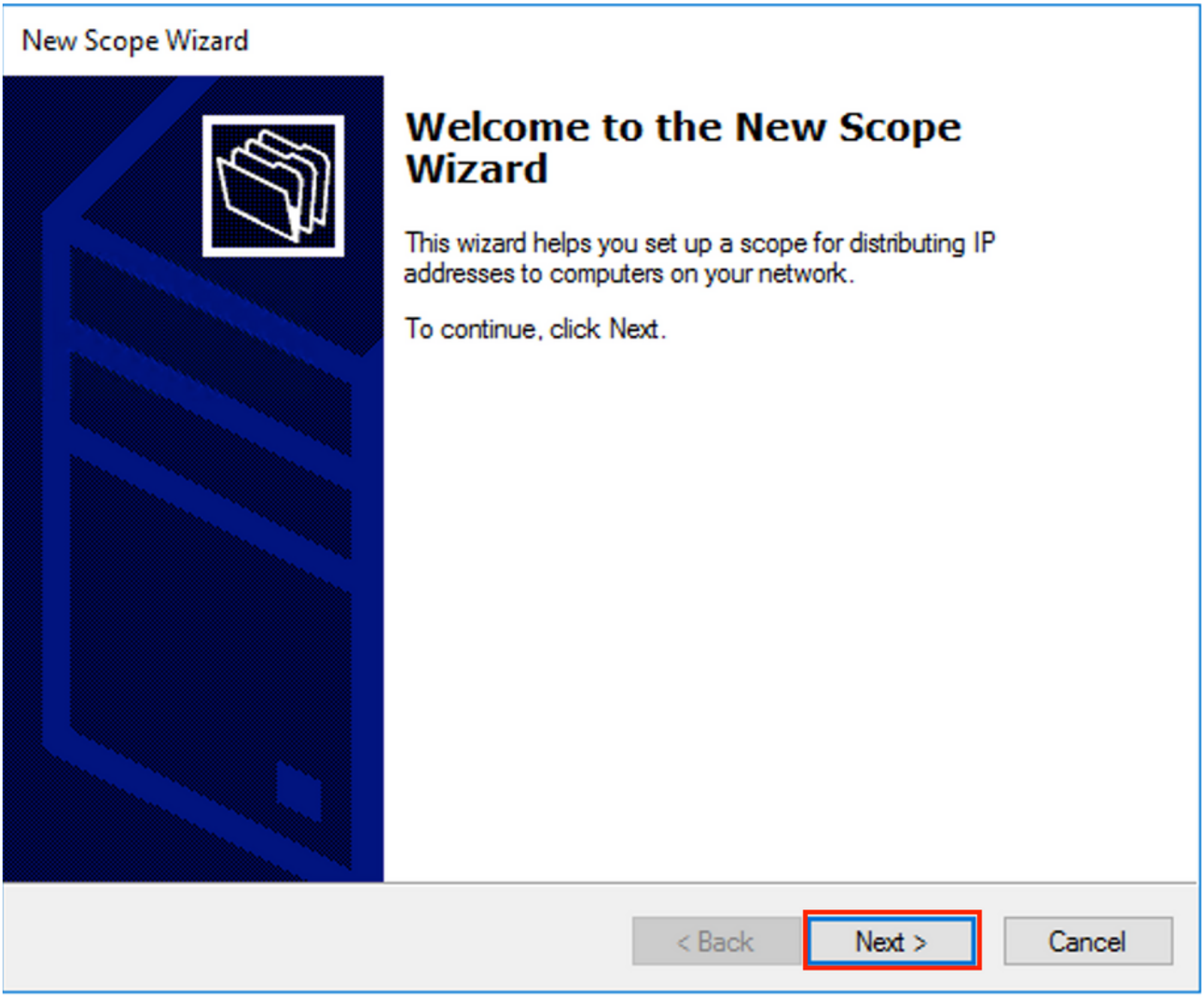


中繼的範圍現已就緒。

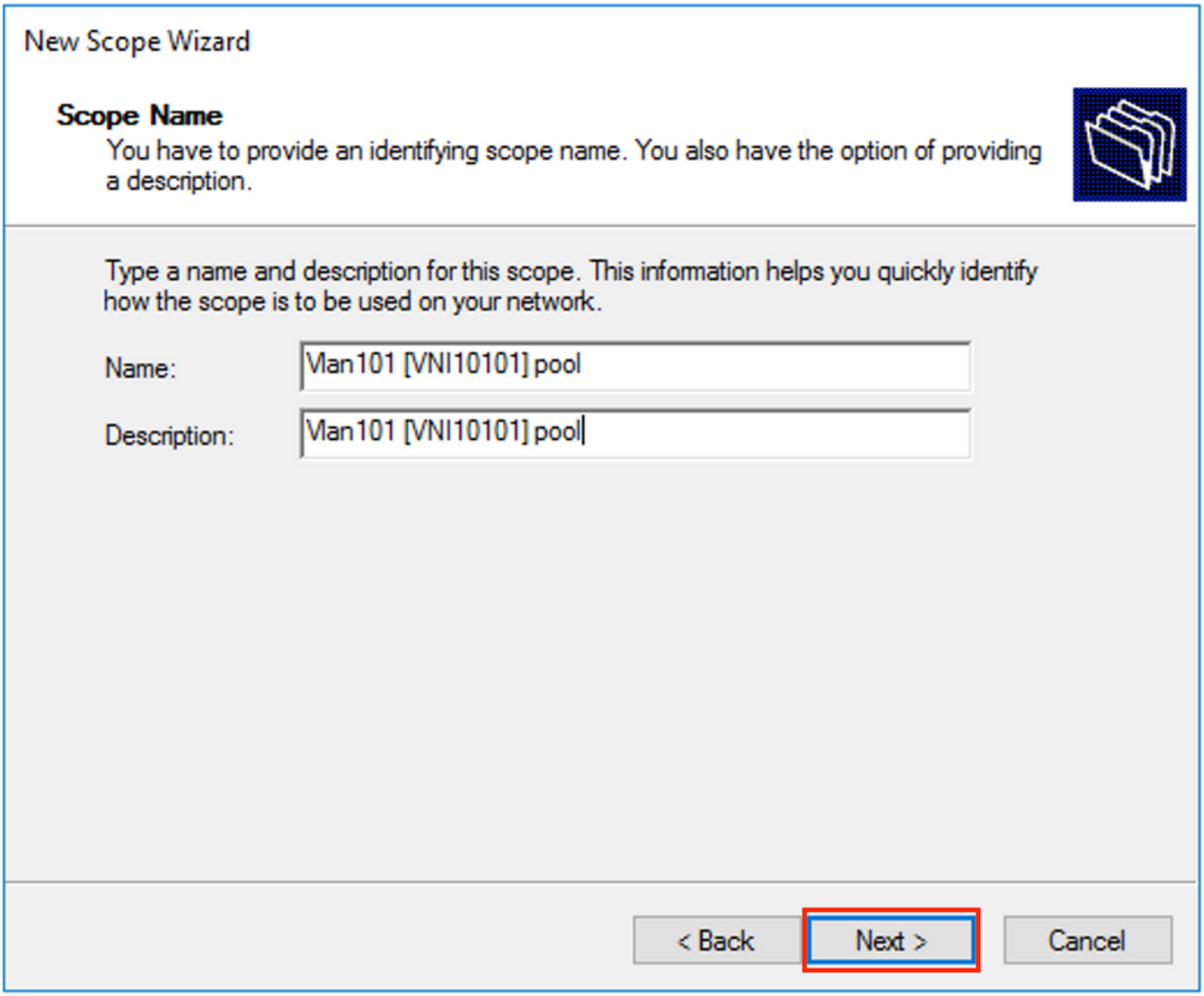
- 接下來，建立客戶端從中獲取IP地址的池。
- 按一下右鍵並選擇**New Scope**，如下圖所示。



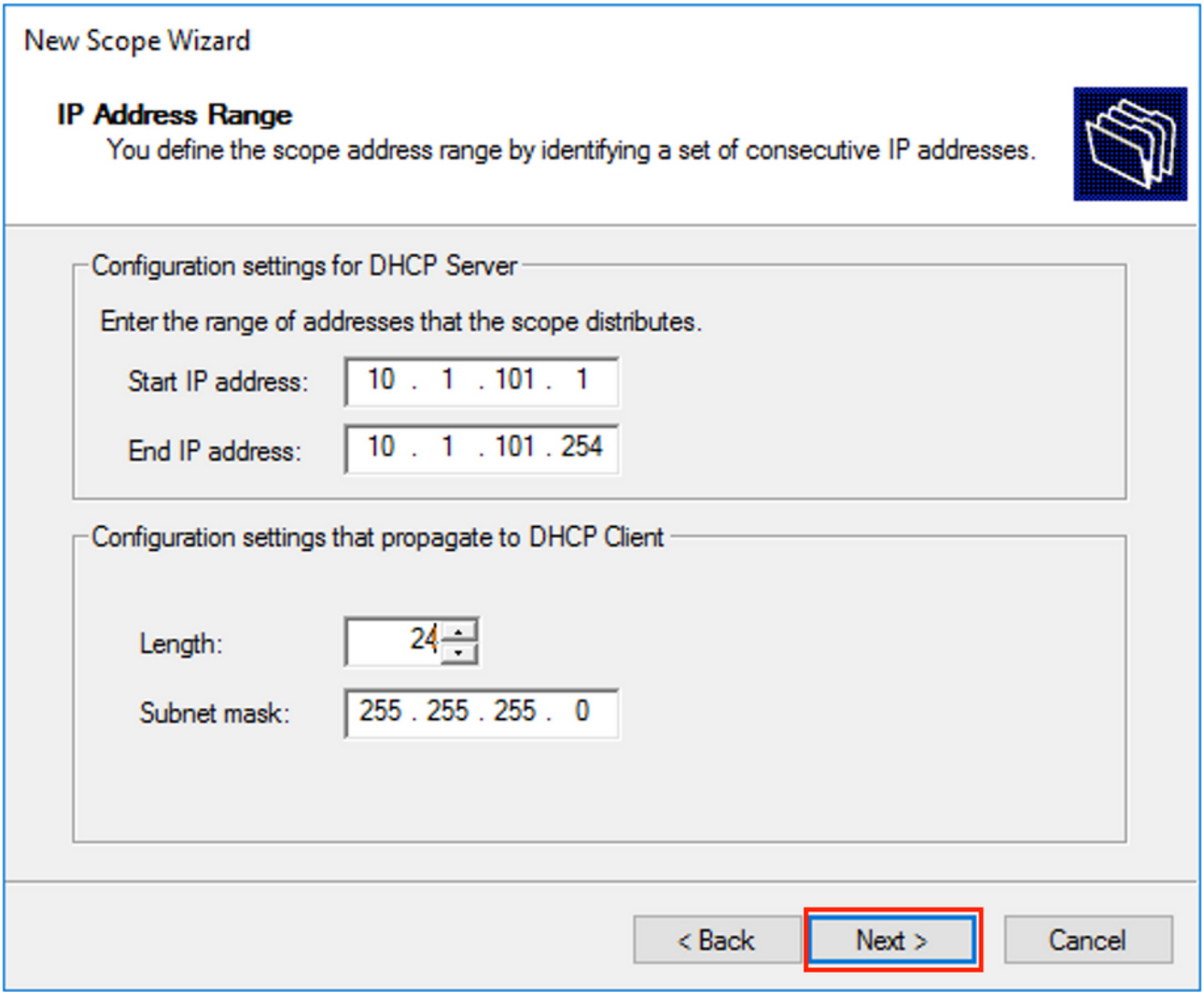
選擇**Next**，如下圖所示。



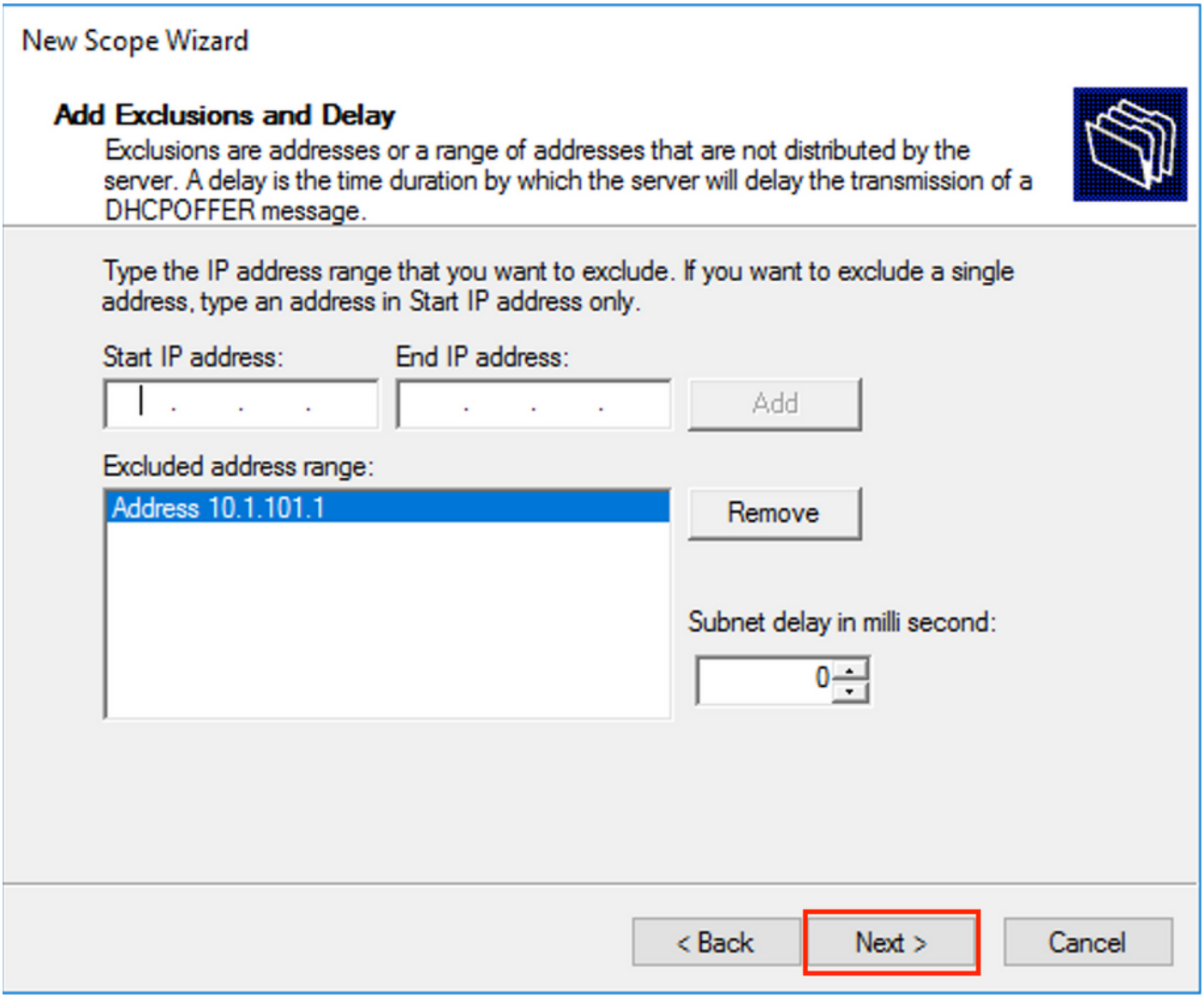
選擇有意義的池名稱和說明，如下圖所示。



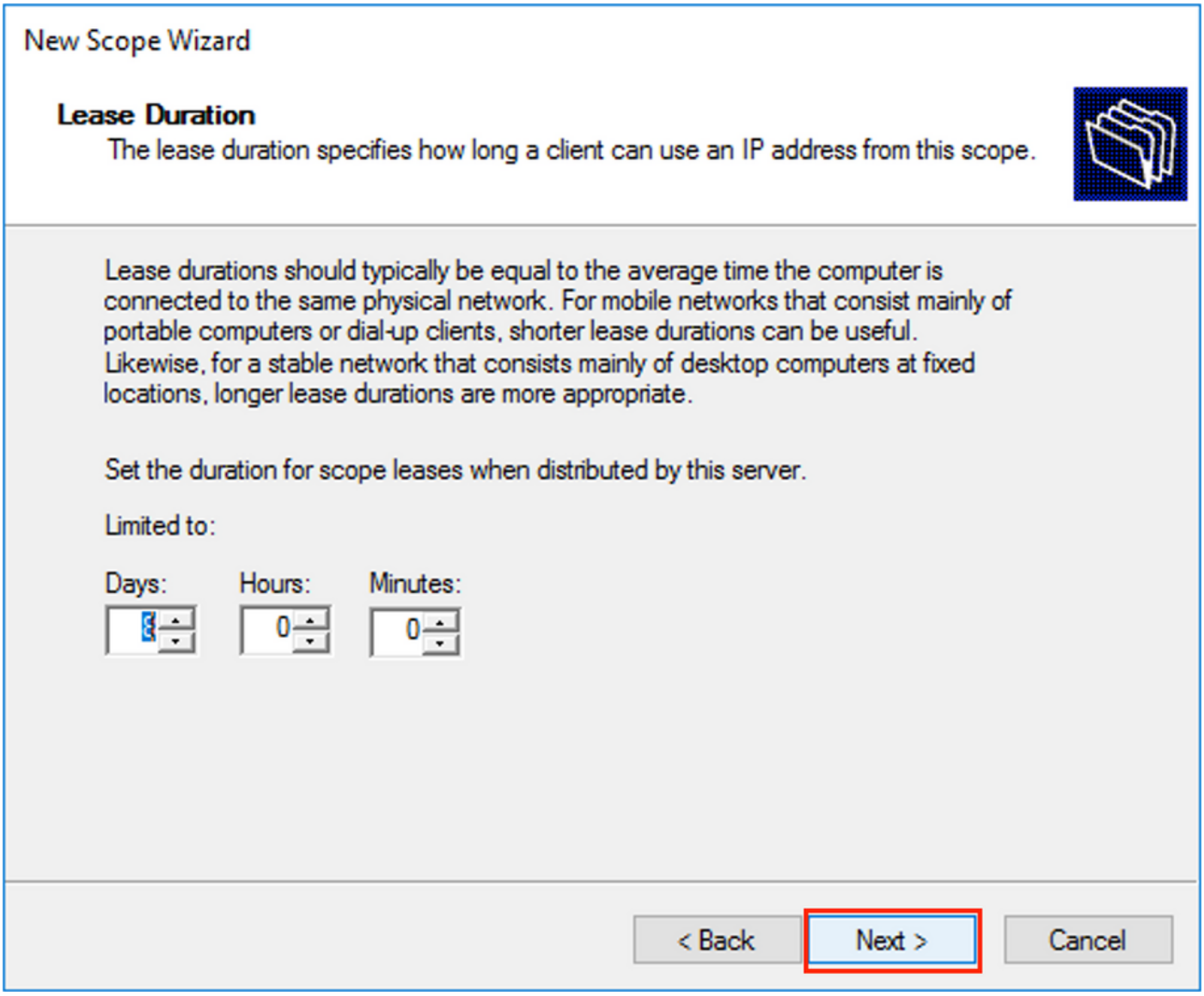
輸入要在vlan101中分配的IP地址空間，如下圖所示。



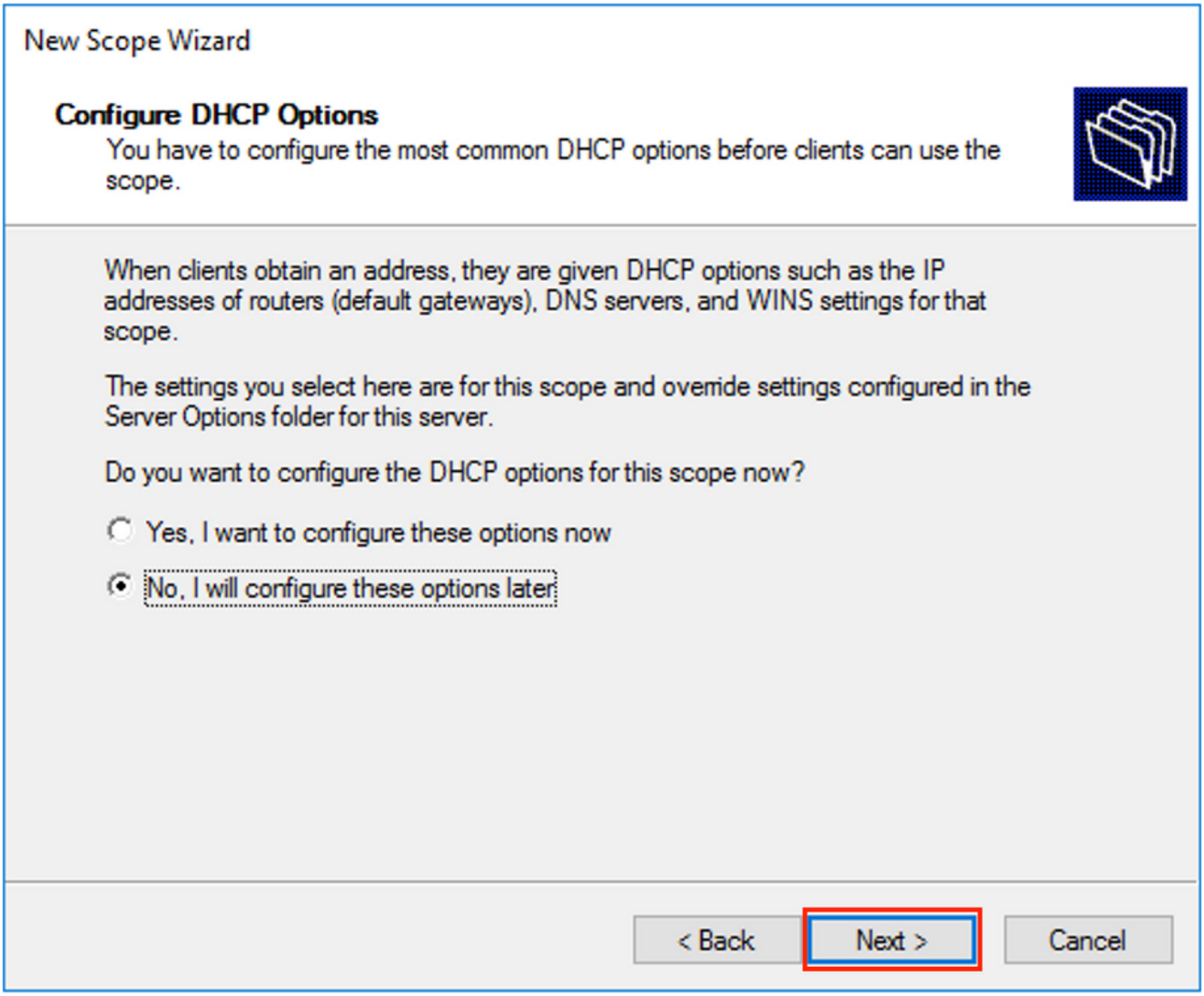
從作用域中排除預設網關IP，如下圖所示。



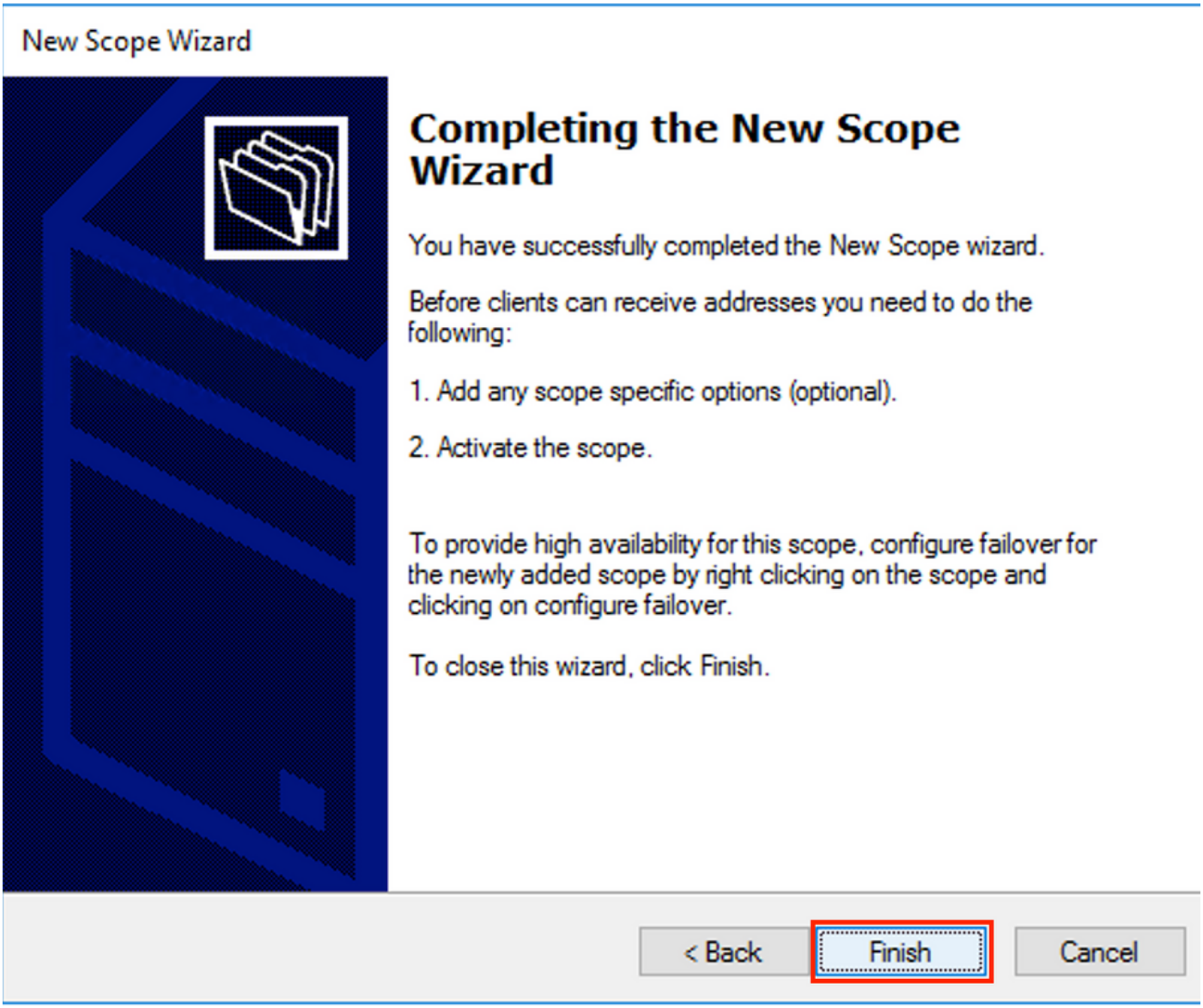
設定租用時間，如下圖所示。



如圖所示，您可以設定DNS/WINS等額外引數（在本範例中跳過）。

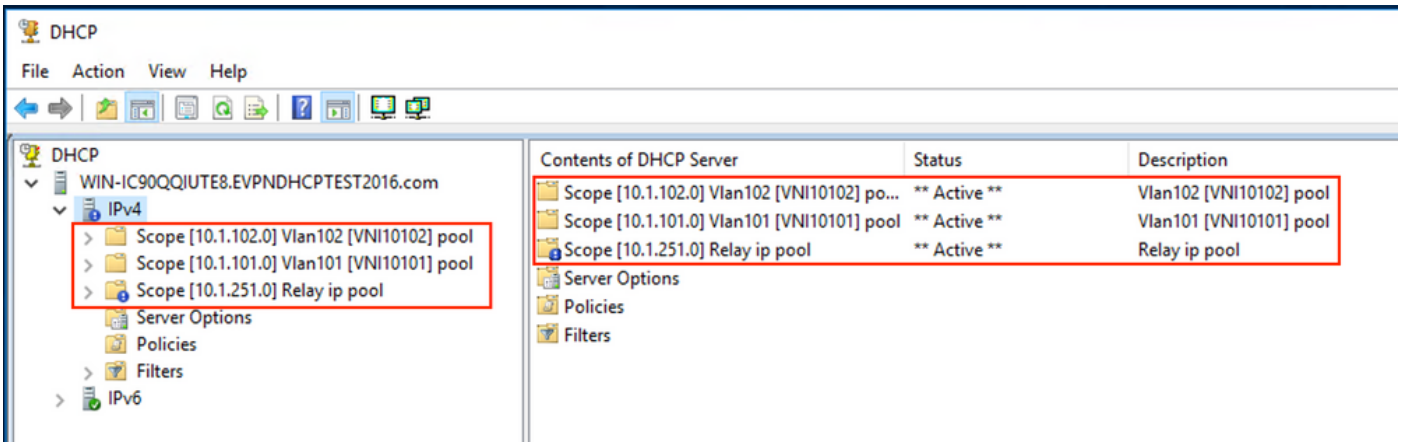


選擇**Finish**以完成設定，如下圖所示。



未配置每個中繼IP地址的池，並且未以十六進位制匹配。池選擇基於子選項**Link selection**。

可以新增新池，無需其他配置，如下圖所示。



Linux DHCP伺服器

檢查Linux上isc-dhcp-server的配置。

- 它支援中繼選項82。在這裡，最重要的選項是鏈路選擇子選項。您仍然可以處理特定欄位的代

理電路ID資訊和十六進位制掩碼/匹配（就像對win2012所做的那樣）。 從實用角度來看，使用82[5]比直接使用Agent電路ID資訊要容易得多。

- 鏈路選擇子選項的配置是在子網定義下完成的。

在本示例中，ISC伺服器用於Ubuntu Linux。

安裝DHCP伺服器：

```
apt-get install isc-dhcp-server
```

要配置DHCP伺服器，請編輯**/etc/dhcp/dhcpd.conf**。（示例中使用了Vim編輯器）

```
vim /etc/dhcp/dhcpd.conf
```

配置片段（省略常規配置）：

```
subnet 10.1.101.0 netmask 255.255.255.0 {

 option agent.link-selection 10.1.101.0;  <<< suboption 82[5] definition

option routers 10.1.101.1;
option subnet-mask 255.255.255.0;

range 10.1.101.16 10.1.101.254;
}

subnet 10.1.102.0 netmask 255.255.255.0 {

option agent.link-selection 10.1.102.0;  <<< suboption 82[5] definition

option routers 10.1.102.1;
option subnet-mask 255.255.255.0;

range 10.1.102.16 10.1.102.254;
}

subnet 10.2.201.0 netmask 255.255.255.0 {

option agent.link-selection 10.2.201.0;  <<< suboption 82[5] definition

option routers 10.2.201.1;
option subnet-mask 255.255.255.0;

range 10.2.201.16 10.2.201.254;
}

subnet 10.2.202.0 netmask 255.255.255.0 {

option agent.link-selection 10.2.202.0;  <<< suboption 82[5] definition

option routers 10.2.202.1;
option subnet-mask 255.255.255.0;

range 10.2.202.16 10.2.202.254;
}
```

交換器組態

此處將檢視一般支援的方案。

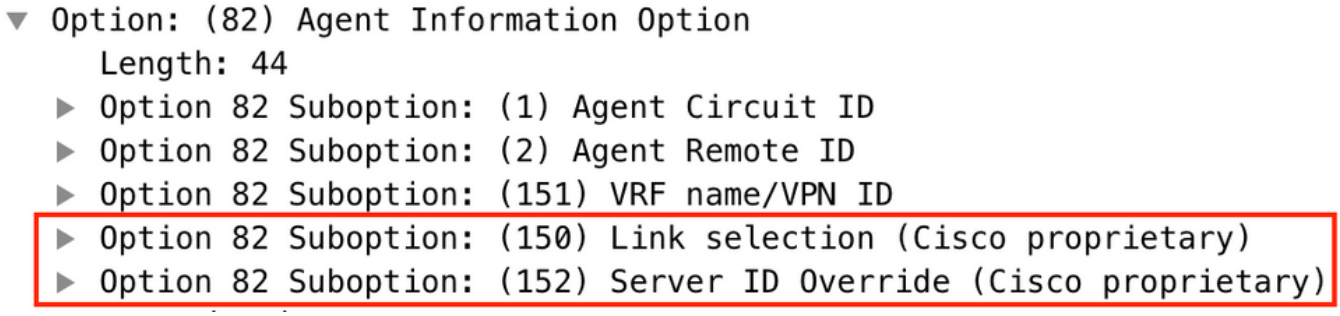
1. DHCP客戶端位於租戶VRF中，DHCP伺服器位於第3層預設VRF中
2. DHCP客戶端位於租戶VRF中，而DHCP伺服器位於同一個租戶VRF中
3. DHCP客戶端位於租戶VRF中，而DHCP伺服器位於不同的租戶VRF中
4. DHCP客戶端位於租戶VRF中，而DHCP伺服器位於非預設非VXLAN VRF中

無論哪種情況，交換機端都需要配置DHCP中繼。

最簡單選項編號2的DHCP配置。

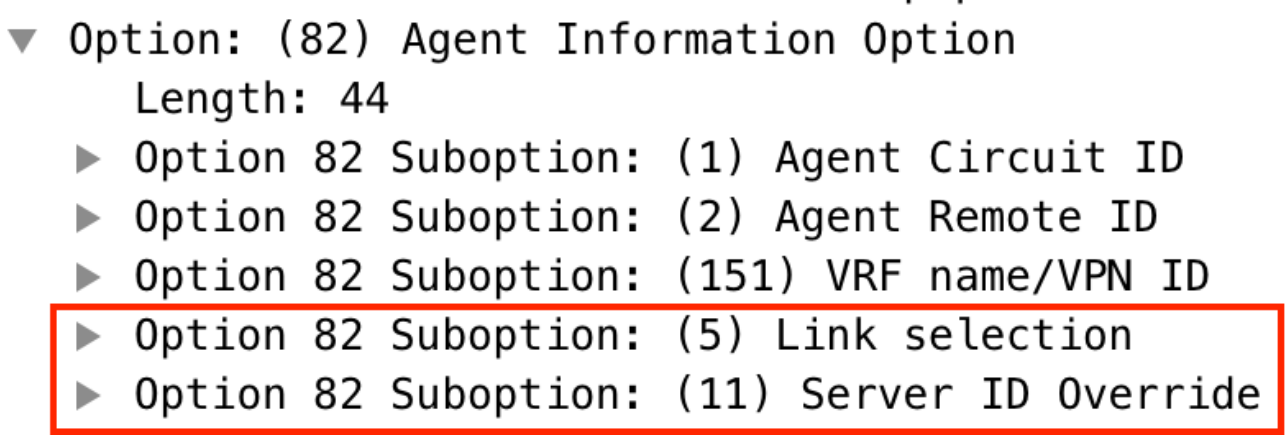
```
ip dhcp relay information option      <<< Enables insertion of option 82 into the packet
ip dhcp relay information option vpn <<< Enables insertion of vpn name/id to the packet - option 82[151]
```

預設情況下，選項82子選項**Link Selection**和**Server ID Override**是思科專有的（分別為150和152）。



如果DHCP伺服器出於任何原因不理解**Cisco專有選項，您可以將其更改為標準選項。**

```
ip dhcp compatibility suboption link-selection standard <<< "Link Selection" suboption
ip dhcp compatibility suboption server-override standard <<< "Server ID Override" suboption
```



必須為必要的VLAN啟用DHCP監聽。

```
ip dhcp snooping vlan 101-102,201-202
ip dhcp snooping
```

您可以使用DHCP-relay source-interface全域性配置。

```
ip dhcp-relay source-interface Loopback101
```

或者，您可以逐個介面進行配置（介面配置將覆蓋全域性配置）。

```
interface Vlan101
vrf forwarding green
ip dhcp relay source-interface Loopback101 <<< DHCP source-interface
ip address 10.1.101.1 255.255.255.0
ip helper-address 192.168.20.20
```

檢查兩個方向是否都存在IP連線b/w中繼IP地址和DHCP伺服器。

```
Leaf-01#ping vrf green 192.168.20.20 source lo101
Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 192.168.20.20, timeout is 2 seconds:
Packet sent with a source address of 10.1.251.1
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 1/1/1 ms
```

在介面配置下，配置DHCP伺服器的地址。此命令可以有3個選項。客戶端和伺服器位於同一個VRF中：

```
interface Vlan101
vrf forwarding green
ip dhcp relay source-interface Loopback101
ip address 10.1.101.1 255.255.255.0
  ip helper-address 192.168.20.20         <<< DHCP server ip address
```

客戶端和伺服器位於不同的VRF中（在本例中，客戶端顯示為綠色，伺服器顯示為紅色）：

```
interface Vlan101
vrf forwarding green
ip dhcp relay source-interface Loopback101
ip address 10.1.101.1 255.255.255.0
ip helper-address vrf red 192.168.20.20   <<< DHCP server is reachable over vrf RED
end
```

VRF中的客戶端和全域性路由表(GRT)中的伺服器：

```
interface Vlan101
vrf forwarding green
ip dhcp relay source-interface Loopback101
ip address 10.1.101.1 255.255.255.0
ip helper-address global 192.168.20.20  <<< DHCP server is reachable over global routing table
end
```

現在，此處會檢視所有選項的典型配置。

DHCP客戶端位於租戶VRF中，DHCP伺服器位於第3層預設VRF中

在這種情況下，GRT中的Lo0是中繼源。DHCP中繼針對某些介面進行了全域性配置+。

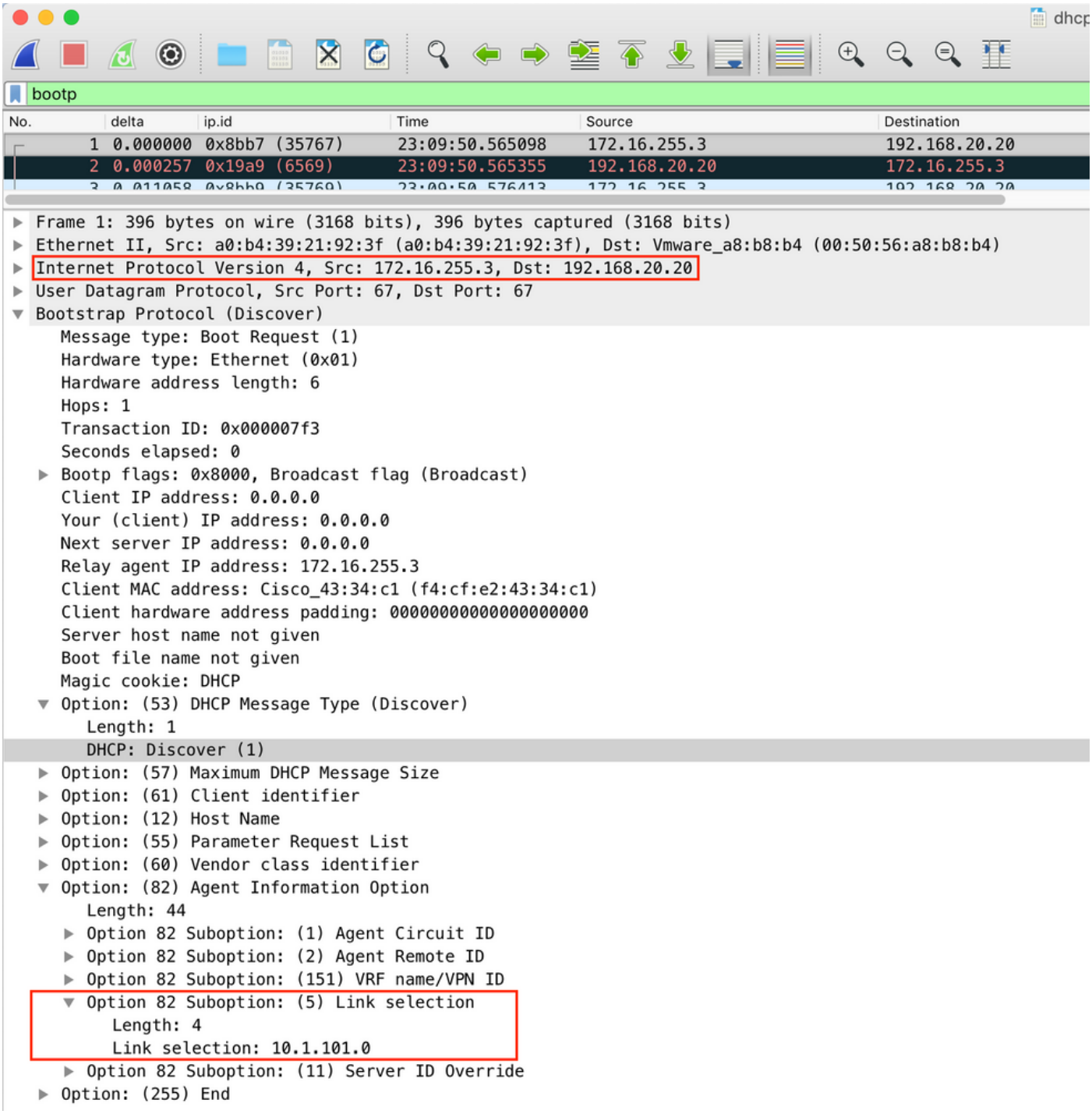
例如，VLAN101命令「IP DHCP relay source-interface Loopback0」丟失，但它使用全域性配置。

```
ip dhcp-relay source-interface Loopback0                 <<< DHCP relay source interface is Lo0
ip dhcp relay information option vpn                     <<< adds the vpn suboption to option 82
ip dhcp relay information option                         <<< enables DHCP option 82
ip dhcp compatibility suboption link-selection standard  <<< switch to standard option 82[5]
ip dhcp compatibility suboption server-override standard <<< switch to standard option 82[11]
ip dhcp snooping vlan 101-102,201-202                    <<< enables dhcp snooping for vlans
ip dhcp snooping                                         <<< enables dhcp snooping globally
!
```

```
interface Loopback0
 ip address 172.16.255.3 255.255.255.255
 ip ospf 1 area 0
!
interface Vlan101
 vrf forwarding green
 ip address 10.1.101.1 255.255.255.0
 ip helper-address global 192.168.20.20           <<< DHCP is reachable over GRT
!
interface Vlan102
 vrf forwarding green
 ip dhcp relay source-interface Loopback0
 ip address 10.1.102.1 255.255.255.0
 ip helper-address global 192.168.20.20           <<< DHCP is reachable over GRT
!
interface Vlan201
vrf forwarding red
ip dhcp relay source-interface Loopback0
ip address 10.2.201.1 255.255.255.0
ip helper-address global 192.168.20.20           <<< DHCP is reachable over GRT
```

因此，使用相同的SRC IP/DST IP但使用不同的子選項通過GRT傳送DHCP中繼資料包。

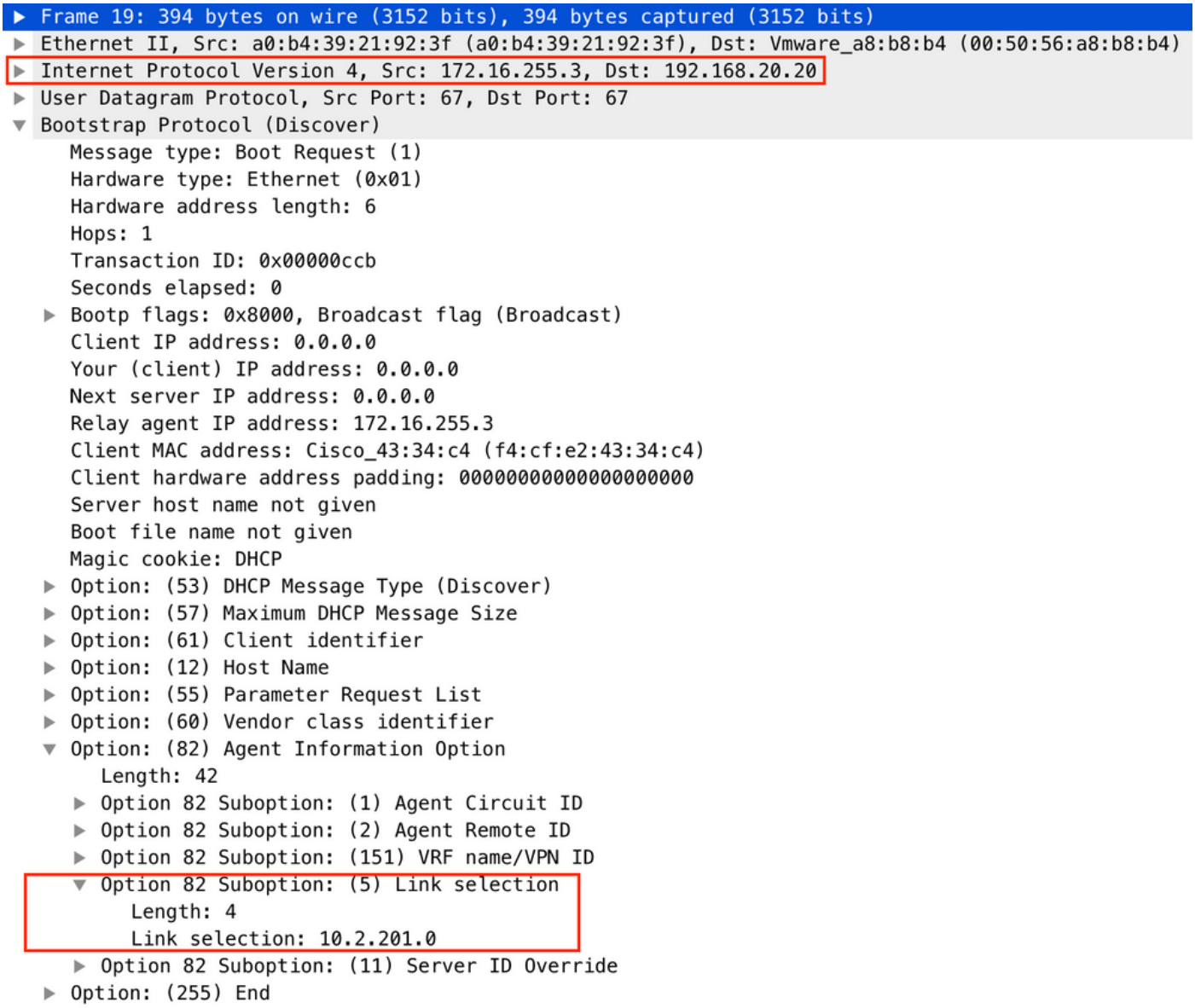
對於vlan101:


dhcp
bootp
No. delta ip.id Time Source Destination
1 0.000000 0x8bb7 (35767) 23:09:50.565098 172.16.255.3 192.168.20.20
2 0.000257 0x19a9 (6569) 23:09:50.565355 192.168.20.20 172.16.255.3
Frame 1: 396 bytes on wire (3168 bits), 396 bytes captured (3168 bits)
Ethernet II, Src: a0:b4:39:21:92:3f (a0:b4:39:21:92:3f), Dst: Vmware_a8:b8:b4 (00:50:56:a8:b8:b4)
Internet Protocol Version 4, Src: 172.16.255.3, Dst: 192.168.20.20
User Datagram Protocol, Src Port: 67, Dst Port: 67
Bootstrap Protocol (Discover)
Message type: Boot Request (1)
Hardware type: Ethernet (0x01)
Hardware address length: 6
Hops: 1
Transaction ID: 0x000007f3
Seconds elapsed: 0
Bootp flags: 0x8000, Broadcast flag (Broadcast)
Client IP address: 0.0.0.0
Your (client) IP address: 0.0.0.0
Next server IP address: 0.0.0.0
Relay agent IP address: 172.16.255.3
Client MAC address: Cisco_43:34:c1 (f4:cf:e2:43:34:c1)
Client hardware address padding: 00000000000000000000
Server host name not given
Boot file name not given
Magic cookie: DHCP
Option: (53) DHCP Message Type (Discover)
Length: 1
DHCP: Discover (1)
Option: (57) Maximum DHCP Message Size
Option: (61) Client identifier
Option: (12) Host Name
Option: (55) Parameter Request List
Option: (60) Vendor class identifier
Option: (82) Agent Information Option
Length: 44
Option 82 Suboption: (1) Agent Circuit ID
Option 82 Suboption: (2) Agent Remote ID
Option 82 Suboption: (151) VRF name/VPN ID
Option 82 Suboption: (5) Link selection
Length: 4
Link selection: 10.1.101.0
Option 82 Suboption: (11) Server ID Override
Option: (255) End

- 對於Vlan102:

```
▶ Frame 8: 396 bytes on wire (3168 bits), 396 bytes captured (3168 bits)
▶ Ethernet II, Src: a0:b4:39:21:92:3f (a0:b4:39:21:92:3f), Dst: Vmware_a8:b8:b4 (00:50:56:a8:b8:b4)
▶ Internet Protocol Version 4, Src: 172.16.255.3, Dst: 192.168.20.20
▶ User Datagram Protocol, Src Port: 67, Dst Port: 67
▼ Bootstrap Protocol (Discover)
    Message type: Boot Request (1)
    Hardware type: Ethernet (0x01)
    Hardware address length: 6
    Hops: 1
    Transaction ID: 0x000007f4
    Seconds elapsed: 0
  ▶ Bootp flags: 0x8000, Broadcast flag (Broadcast)
    Client IP address: 0.0.0.0
    Your (client) IP address: 0.0.0.0
    Next server IP address: 0.0.0.0
    Relay agent IP address: 172.16.255.3
    Client MAC address: Cisco_43:34:c3 (f4:cf:e2:43:34:c3)
    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: (57) Maximum DHCP Message Size
  ▶ Option: (61) Client identifier
  ▶ Option: (12) Host Name
  ▶ Option: (55) Parameter Request List
  ▼ Option: (60) Vendor class identifier
      Length: 8
      Vendor class identifier: ciscopnp
  ▼ Option: (82) Agent Information Option
      Length: 44
    ▶ Option 82 Suboption: (1) Agent Circuit ID
    ▶ Option 82 Suboption: (2) Agent Remote ID
    ▶ Option 82 Suboption: (151) VRF name/VPN ID
    ▼ Option 82 Suboption: (5) Link selection
        Length: 4
        Link selection: 10.1.102.0
    ▶ Option 82 Suboption: (11) Server ID Override
  ▼ Option: (255) End
      Option End: 255
```

對於Vlan201（為vrf紅色，而不是VLAN 101和102的綠色）：



在Spine-01上捕獲的資料包從介面傳輸到枝葉–01:

```
Spine-01#sh mon cap TAC buff br | i DHCP
5401 4.402431 172.16.255.3 b^F^R 192.168.20.20 DHCP 396 DHCP Discover - Transaction ID 0x1feb
5403 4.403134 192.168.20.20 b^F^R 172.16.255.3 DHCP 362 DHCP Offer - Transaction ID 0x1feb
5416 4.418117 172.16.255.3 b^F^R 192.168.20.20 DHCP 414 DHCP Request - Transaction ID 0x1feb
5418 4.418608 192.168.20.20 b^F^R 172.16.255.3 DHCP 362 DHCP ACK - Transaction ID 0x1feb
```

核心層中的DHCP資料包是沒有任何VXLAN封裝的IP:

```
Spine-01#sh mon cap TAC buff det | b Frame 5401:
Frame 5401: 396 bytes on wire (3168 bits), 396 bytes captured (3168 bits) on interface 0
<...skip...>
[Protocols in frame: eth:ethertype:ip:udp:dhcp]
Ethernet II, Src: 10:b3:d5:6a:8f:e4 (10:b3:d5:6a:8f:e4), Dst: 7c:21:0d:92:b2:e4
(7c:21:0d:92:b2:e4)
<...skip...>
Internet Protocol Version 4, Src: 172.16.255.3, Dst: 192.168.20.20
<...skip...>
User Datagram Protocol, Src Port: 67, Dst Port: 67
<...skip...>
Dynamic Host Configuration Protocol (Discover)
<...skip...>
```

此方法的一大優勢是您可以為不同的租戶VRF使用相同的中繼IP地址，而不會在不同的VRF和全域性之間發生路由洩漏。

DHCP客戶端和DHCP伺服器位於同一個租戶VRF中

在這種情況下，租戶VRF中具有中繼IP地址很有意義。

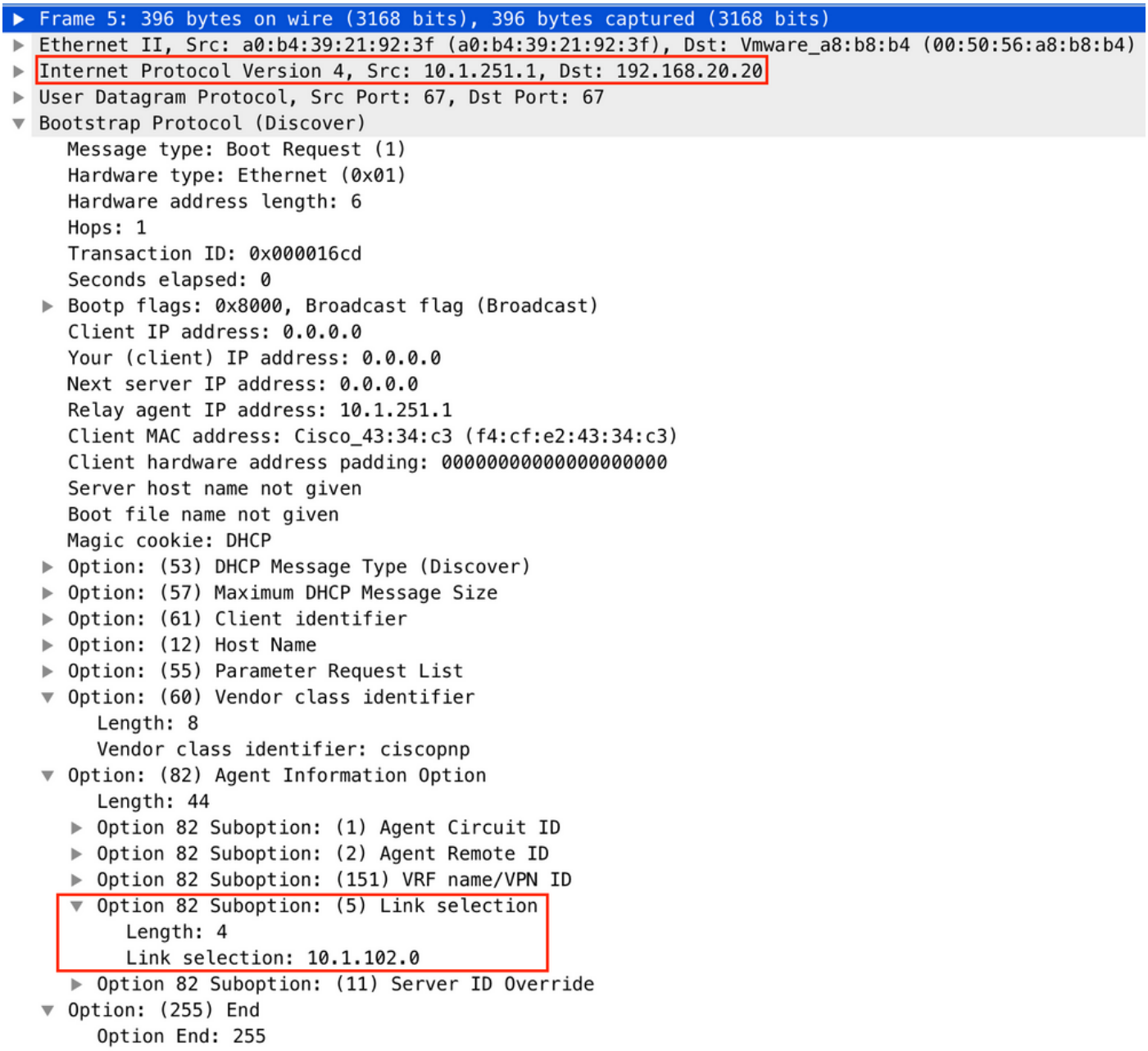
交換機配置：

```
ip dhcp relay information option vpn                     <<< adds the vpn suboption to option 82
ip dhcp relay information option                         <<< enables DHCP option 82
ip dhcp compatibility suboption link-selection standard  <<< switch to standard option 82[5]
ip dhcp compatibility suboption server-override standard <<< switch to standard option 82[11]
ip dhcp snooping vlan 101-102,201-202                    <<< enables dhcp snooping for vlans
ip dhcp snooping                                         <<< enables dhcp snooping globally
!
interface Loopback101
vrf forwarding green
ip address 10.1.251.1 255.255.255.255
!
interface Vlan101
 vrf forwarding green
 ip dhcp relay source-interface Loopback101
 ip address 10.1.101.1 255.255.255.0
 ip helper-address 192.168.20.20                      <<< DHCP is reachable over vrf green
!
interface Vlan102
 vrf forwarding green
 ip dhcp relay source-interface Loopback101
 ip address 10.1.102.1 255.255.255.0
 ip helper-address 192.168.20.20                      <<< DHCP is reachable over vrf green
```

對於vlan101:

```
▶ Frame 1: 396 bytes on wire (3168 bits), 396 bytes captured (3168 bits)
▶ Ethernet II, Src: a0:b4:39:21:92:3f (a0:b4:39:21:92:3f), Dst: Vmware_a8:b8:b4 (00:50:56:a8:b8:b4)
▶ Internet Protocol Version 4, Src: 10.1.251.1, Dst: 192.168.20.20
▶ User Datagram Protocol, Src Port: 67, Dst Port: 67
▼ Bootstrap Protocol (Discover)
    Message type: Boot Request (1)
    Hardware type: Ethernet (0x01)
    Hardware address length: 6
    Hops: 1
    Transaction ID: 0x000016cc
    Seconds elapsed: 0
  ▶ Bootp flags: 0x8000, Broadcast flag (Broadcast)
    Client IP address: 0.0.0.0
    Your (client) IP address: 0.0.0.0
    Next server IP address: 0.0.0.0
    Relay agent IP address: 10.1.251.1
    Client MAC address: Cisco_43:34:c1 (f4:cf:e2:43:34:c1)
    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: (57) Maximum DHCP Message Size
  ▶ Option: (61) Client identifier
  ▶ Option: (12) Host Name
  ▶ Option: (55) Parameter Request List
  ▶ Option: (60) Vendor class identifier
  ▼ Option: (82) Agent Information Option
      Length: 44
    ▶ Option 82 Suboption: (1) Agent Circuit ID
    ▶ Option 82 Suboption: (2) Agent Remote ID
    ▶ Option 82 Suboption: (151) VRF name/VPN ID
    ▼ Option 82 Suboption: (5) Link selection
        Length: 4
        Link selection: 10.1.101.0
    ▶ Option 82 Suboption: (11) Server ID Override
  ▶ Option: (255) End
```

對於vlan102:



從主幹–01到枝葉–01介面的資料包捕獲：

```
Spine-01#sh monitor capture TAC buffer brief | i DHCP
2 4.287466 10.1.251.1 b^F^R 192.168.20.20 DHCP 446 DHCP Discover - Transaction ID 0x1894
3 4.288258 192.168.20.20 b^F^R 10.1.251.1 DHCP 412 DHCP Offer - Transaction ID 0x1894
4 4.307550 10.1.251.1 b^F^R 192.168.20.20 DHCP 464 DHCP Request - Transaction ID 0x1894
5 4.308385 192.168.20.20 b^F^R 10.1.251.1 DHCP 412 DHCP ACK - Transaction ID 0x1894
```

核心中的DHCP資料包具有VXLAN封裝：

```
Frame 2: 446 bytes on wire (3568 bits), 446 bytes captured (3568 bits) on interface 0
<...skip...>
[Protocols in frame: eth:ethertype:ip:udp:vxlan:eth:ethertype:ip:udp:dhcp]
Ethernet II, Src: 10:b3:d5:6a:8f:e4 (10:b3:d5:6a:8f:e4), Dst: 7c:21:0d:92:b2:e4
(7c:21:0d:92:b2:e4)
<...skip...>
Internet Protocol Version 4, Src: 172.16.254.3, Dst: 172.16.254.5 <<< VTEP IP addresses
<...skip...>
User Datagram Protocol, Src Port: 65283, Dst Port: 4789
<...skip...>
```

```
Virtual eXtensible Local Area Network
Flags: 0x0800, VXLAN Network ID (VNI)
0... .... .... .... = GBP Extension: Not defined
.... .... .0.. .... = Don't Learn: False
.... 1... .... .... = VXLAN Network ID (VNI): True
.... .... .... 0... = Policy Applied: False
.000 .000 0.00 .000 = Reserved(R): 0x0000
Group Policy ID: 0
VXLAN Network Identifier (VNI): 50901 <<<<<<<<<<<< L3VNI for VRF green
Reserved: 0
<--- Inner header started --->
Ethernet II, Src: 10:b3:d5:6a:00:00 (10:b3:d5:6a:00:00), Dst: 7c:21:0d:bd:27:48
(7c:21:0d:bd:27:48)
<...skip...>
Internet Protocol Version 4, Src: 10.1.251.1, Dst: 192.168.20.20
<...skip...>
User Datagram Protocol, Src Port: 67, Dst Port: 67
<...skip...>
Dynamic Host Configuration Protocol (Discover)
<...skip...>
```

一個租戶中的DHCP客戶端VRF和另一個租戶VRF中的DHCP伺服器

在本示例中，客戶端為vrf紅色，伺服器為vrf綠色。

有兩種選擇：

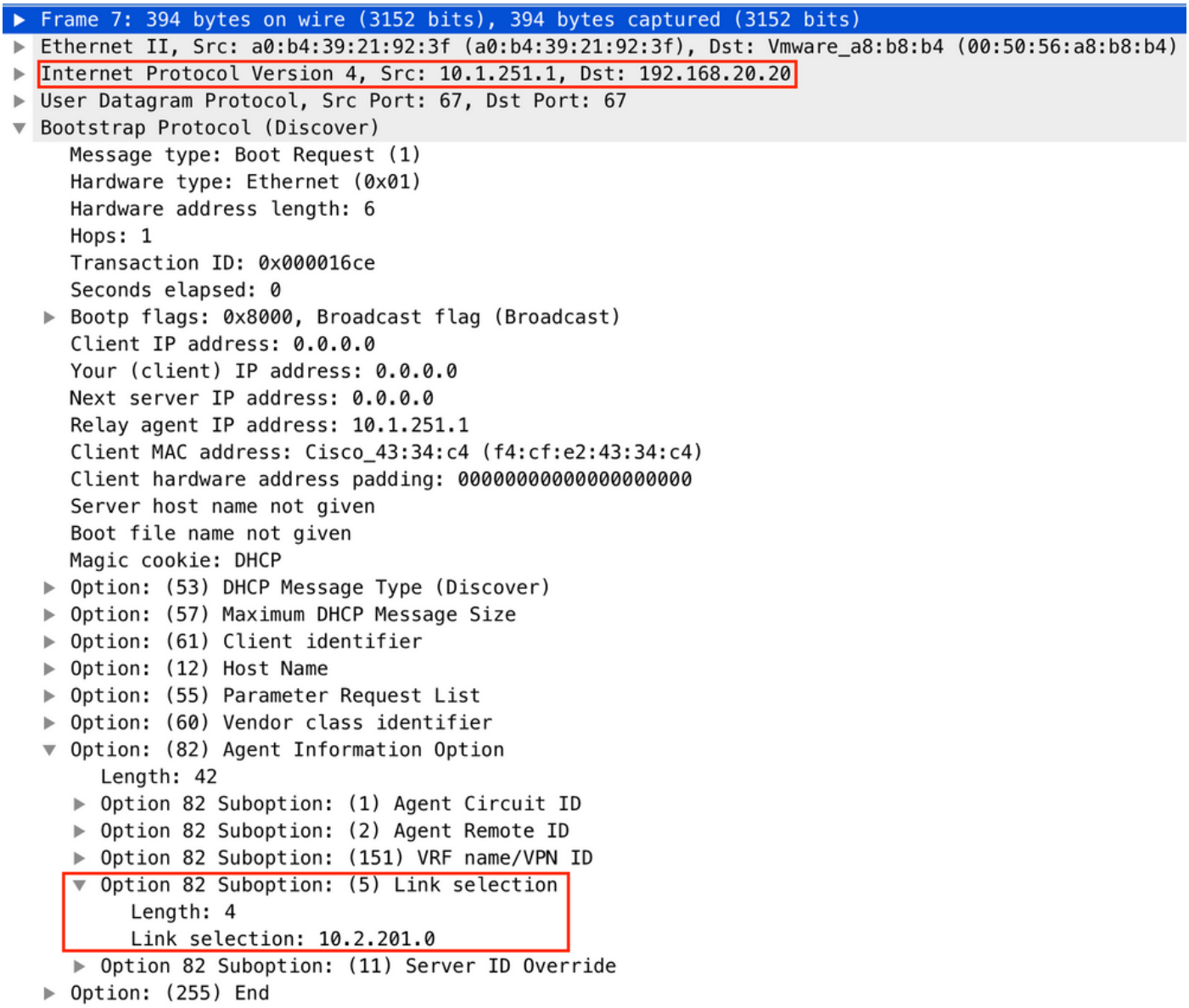
- 保持客戶端vrf中的中繼IP並配置路由洩漏，從而增加複雜性
- 在伺服器vrf中保留中繼IP（類似於您在第一種情況下對GRT所做的操作）

由於支援大量客戶端VRF且不需要路由洩漏，因此選擇第二種方法更簡單。

交換機配置：

```
ip dhcp relay information option vpn                    <<< adds the vpn suboption to option 82
ip dhcp relay information option                        <<< enables DHCP option 82
ip dhcp compatibility suboption link-selection standard  <<< switch to standard option 82[5]
ip dhcp compatibility suboption server-override standard <<< switch to standard option 82[11]
ip dhcp snooping vlan 101-102,201-202                   <<< enables dhcp snooping for vlans
ip dhcp snooping                                        <<< enables dhcp snooping globally
!
interface Loopback101
vrf forwarding green
ip address 10.1.251.1 255.255.255.255
!
interface Vlan201
vrf forwarding red
ip dhcp relay source-interface Loopback101
ip address 10.2.201.1 255.255.255.0
ip helper-address vrf green 192.168.20.20               <<< DHCP is reachable over vrf green
```

對於vlan201:



Spine-01到Leaf-01介面上的資料包捕獲：

```
Spine-01#sh mon cap TAC buff br | i DHCP
2 0.168829 10.1.251.1 b^F^R 192.168.20.20 DHCP 444 DHCP Discover - Transaction ID 0x10db
3 0.169450 192.168.20.20 b^F^R 10.1.251.1 DHCP 410 DHCP Offer - Transaction ID 0x10db
4 0.933121 10.1.251.1 b^F^R 192.168.20.20 DHCP 462 DHCP Request - Transaction ID 0x10db
5 0.933970 192.168.20.20 b^F^R 10.1.251.1 DHCP 410 DHCP ACK - Transaction ID 0x10db
```

在本範例中，核心中的封包採用VXLAN封裝。

```
Frame 2: 446 bytes on wire (3552 bits), 444 bytes captured (3552 bits) on interface 0
<...skip...>
[Protocols in frame: eth:ethertype:ip:udp:vxlan:eth:ethertype:ip:udp:dhcp]
Ethernet II, Src: 10:b3:d5:6a:8f:e4 (10:b3:d5:6a:8f:e4), Dst: 7c:21:0d:92:b2:e4
(7c:21:0d:92:b2:e4)
<...skip...>
Internet Protocol Version 4, Src: 172.16.254.3, Dst: 172.16.254.5  <<< VTEP IP addresses
<...skip...>
User Datagram Protocol, Src Port: 65283, Dst Port: 4789
<...skip...>
Virtual eXtensible Local Area Network
Flags: 0x0800, VXLAN Network ID (VNI)
0... .... .... .... = GBP Extension: Not defined
```

```
.... .... .0.. .... = Don't Learn: False
.... 1... .... .... = VXLAN Network ID (VNI): True
.... .... .... 0... = Policy Applied: False
.000 .000 0.00 .000 = Reserved(R): 0x0000
Group Policy ID: 0
VXLAN Network Identifier (VNI): 50901                                  <<< L3VNI for VRF green
Reserved: 0
<--- Inner header started --->
Ethernet II, Src: 10:b3:d5:6a:00:00 (10:b3:d5:6a:00:00), Dst: 7c:21:0d:bd:27:48
(7c:21:0d:bd:27:48)
<...skip...>
Internet Protocol Version 4, Src: 10.1.251.1, Dst: 192.168.20.20
<...skip...>
User Datagram Protocol, Src Port: 67, Dst Port: 67
<...skip...>
Dynamic Host Configuration Protocol (Discover)
<...skip...>
```

一個租戶中的DHCP客戶端VRF和另一個非VXLAN VRF中的DHCP伺服器

這個案例和上一個非常相似。主要區別在於，封包沒有VXLAN封裝 — 純IP或其他（MPLS/GRE/等），但從組態角度看是相同的。

在本示例中，客戶端為vrf紅色，伺服器為vrf綠色。

有兩種選擇：

- 中繼IP位於客戶端vrf中並配置路由洩漏，從而增加複雜性
- Relay IP在伺服器vrf中（類似於第一種情況下對GRT執行的操作）

由於支援大量客戶端VRF且不需要路由洩漏，因此選擇第二種方法更簡單。

交換機配置：

```
ip dhcp relay information option vpn                      <<< adds the vpn suboption to option 82
ip dhcp relay information option                          <<< enables DHCP option 82
ip dhcp compatibility suboption link-selection standard   <<< switch to standard option 82[5]
ip dhcp compatibility suboption server-override standard <<< switch to standard option 82[11]
ip dhcp snooping vlan 101-102,201-202                     <<< enable dhcp snooping for vlans
ip dhcp snooping                                          <<< enable dhcp snooping globally
!
interface Loopback101
vrf forwarding green
ip address 10.1.251.1 255.255.255.255
!
interface Vlan201
vrf forwarding red
ip dhcp relay source-interface Loopback101
ip address 10.2.201.1 255.255.255.0
ip helper-address vrf green 192.168.20.20               <<< DHCP is reachable over vrf green
```

相關資訊

- RFC 3046
- RFC 3527
- https://docs.microsoft.com
- 技術支援與文件 - Cisco Systems