使用ASR 1000配置重疊傳輸虛擬化

目錄

簡介

本文檔介紹ASR1000和Catalyst 8300/8500系列路由器上支援的重疊傳輸虛擬化(OTV)網路拓撲。

必要條件

需求

本文件沒有特定需求。

採用元件

本文中的資訊係根據以下軟體和硬體版本：

- ASR1000、IOS® XE 16.10.1a版及更高版本

- Catalyst 8300、IOS® XE版本17.5.1a及更高版本
- Catalyst 8500、IOS® XE版本17.6.1a及更高版本

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。

背景資訊

自Cisco IOS® XE版本3.5起，ASR1000支援OTV。Catalyst 8300系列路由器開始支援IOS® XE17.5.1a，Catalyst 8500系列路由開始支援IOS® XE版本17.6.1a。

OTV透過傳輸網路中基於MAC地址的路由和IP封裝轉發(MAC-in-IP)在遠端網路站點之間提供第2層連線，為需要第2層鄰接的應用（如群集和虛擬化）提供支援。OTV使用重疊控制平面協定在重疊網路中學習和傳播MAC路由資訊。OTV控制平面協定使用中間系統到中間系統(IS-IS)消息來建立與遠端站點的鄰接關係並向遠端站點傳送MAC路由更新。透過自動發現遠端OTV裝置，OTV與重疊網路上的遠端站點建立第2層鄰接。

OTV在第2層擴展的優勢包括：

- 無MPLS要求
- 網狀架構沒有複雜的多重協定乙太網路標籤交換(EoMPLS)組態
- 第2層擴展無需複雜的虛擬專用LAN服務(VPLS)部署
- 本地生成樹隔離
    - 無需明確配置網橋資料協定單元(BPDU)過濾器
    - 預設隔離給定資料中心的生成樹問題
- 本地未知單播泛洪隔離
    - 未知的單播MAC資料包未轉發
    - 允許支援每MAC未知單播轉發
- 使用OTV ARP快取最佳化地址解析協定(ARP)
    - 減少不必要的廣域網流量
- 簡化第一躍點備援通訊協定(FHRP)隔離的布建
- 簡化站點的增加
- 簡化的冗餘配置
- 在需要臨時服務時，能夠使用「裝置部署」進行遷移

需求

設計OTV部署時，需要記住的主要規則是後續專案。 如果這些規則得到遵守，設計和部署就會得到簡化。

- 對於所有已配置的OTV重疊介面，一個且只有一個介面可用於傳輸OTV封裝的流量（稱為加入介面）
- 一個且僅一個介面可用於配置OTV站點VLAN的資料中心L2服務例項和所有已配置OTV覆蓋介面的資料中心之間擴展的VLAN
- 埠通道可用於介面冗餘和與VSS或VPC交換機的連線，並且支援作為「唯一一個」介面進行連線。
- 所有OTV路由器都必須可透過加入介面聯絡

- 必須在指向資料中心的OTV路由器上配置生成樹
- 必須配置IGMP監聽和查詢才能正確轉發資料中心組播流量
- 給定資料中心可以配置1個或2個OTV路由器。使用兩台路由器時，它們會根據VLAN編號以奇數/偶數方式分配VLAN轉發。 資料中心中的每台OTV路由器都充當另一台路由器的備份。
- 多宿主對必須使用相同的OTV站點識別符號進行配置
- ASR1000/Catalyst 8300/Catalyst 8500和Nexus 7000可以參與同一個OTV網路
    - Nexus 7000不支援OTV分段或加密，因此這些功能不能用於「混合」部署。

某些背對背連線設計受支援，不符合上述規則。 雖然支援這些組態，但不建議使用。 有關這些操作的詳細資訊，請參閱後面的「特殊情況單播拓撲」部分。

目前的OTV軟體在設定OTV的加入和L2存取介面時，具有「一個且僅一個」介面限制，這一點再怎麼強調都不為過。連線埠通道介面可用來做為備援。 支援埠通道到VPC中的Nexus 7000的連線。還支援與單個交換機的基本埠通道連線。

OTV實施型別

OTV需要單個加入介面和單個L2介面。 每個OTV路由器只能支援其中一個。 OTV還要求配置站點VLAN，以便多宿主OTV路由器能夠透過本地網路相互通訊。 即使是單宿主OTV路由器也必須配置OTV站點VLAN。 此外，每個站點或資料中心必須配置唯一的站點識別符號。 雙宿主OTV路由器必須使用相同的站點識別符號，並且能夠透過同一個VLAN通訊。

後續配置提供了OTV所需的基本配置。但是，它並不完整，因為必須增加單播或組播核心配置。這些將在本文檔後面的部分中詳細介紹。

```
otv site bridge-domain 100
otv site-identifier 0000.0000.1111
!
interface Overlay1
  no ip address
  otv join-interface GigabitEthernet0/0/0
  service instance 99 ethernet
    encapsulation dot1q 99
    bridge-domain 99
  !
  service instance 90 ethernet
    encapsulation dot1q 90
    bridge-domain 90
!
interface GigabitEthernet1/0/1
  no ip address
  negotiation auto
  service instance 100 ethernet
    encapsulation dot1q 100
    bridge-domain 100
  !
  service instance 99 ethernet
    encapsulation dot1q 99
    bridge-domain 99
  !
  service instance 98 ethernet
    encapsulation dot1q 98 second-dot1q 1098
```

```
    rewrite ingress tag trans 2-to-1 dot1q 90 symmetric
    bridge-domain 90
```

服務例項配置用於所有帶OTV的L2介面配置。

L2介面上的每個服務例項都必須與特定的單標籤或雙標籤封裝相關聯。

反過來，這些服務例項中的每一個都必須與一個橋接域相關聯。

該橋接域用於重疊介面上配置的服務例項。

網橋域是將Overlay服務例項連結到L2介面服務例項的粘合劑。

重疊介面上的流量封裝必須與L2介面上重寫入口後的流量封裝匹配。

在示例中，在Gig1/0/1服務例項99上進入的流量具有單個99的802.1Q VLAN和網橋域99。在重疊介面上使用網橋域99的相應服務例項也配置為99的單個802.1Q VLAN。此情況最為簡單。

在示例中，進入Gig1/0/1服務例項98的流量具有兩個802.1Q VLAN，分別為99和1098，橋接域為90。在重疊介面上橋接域為90的對應服務例項配置為單個802.1Q VLAN，即90。顯然，這些並不相同。 rewrite ingress命令可確保標籤在流量透過入口介面時正確轉換。 進入L2介面的流量將98/1098 802.1Q VLAN替換為單個90的VLAN。symmetric關鍵字可確保從L2介面流出的流量將單個802.1Q VLAN 90替換為98/1098。

具有多個802.1Q VLAN且由OTV擴展的任何服務例項都必須使用rewrite ingress命令。 OTV封裝僅支援一個VLAN識別符號。 因此，必須將L2介面上的任何雙VLAN配置重寫為重疊介面服務例項上的單個標籤。 這就排除了對不明確VLAN配置的支援。

有關標籤重寫的詳細資訊，請參閱此文檔：https://www.cisco.com/c/en/us/td/docs/ios-xml/ios/cether/command/ce-cr-book/ce-m1.html

在本示例中，OTV站點網橋域為100。

- OTV站點橋接域僅在L2介面上配置。
- OTV站點橋接域絕不能在重疊介面上配置，因為這會導致OTV部署不穩定。
- OTV站點VLAN必須僅連線到OTV路由器，並且不能傳輸任何其他資料中心/使用者流量。
- OTV站點VLAN必須與OTV擴展VLAN位於同一物理介面上。

多重首頁

資料中心可與單台OTV主機或最多2台主機連線，以實現冗餘，也稱為多宿。 Multihome用於恢復能力和負載均衡。當一個站點中存在多個邊緣裝置，並且兩個裝置都參與同一個重疊網路時，該站點將被視為多宿主。OTV Multihome根據VLAN編號以奇數/偶數方式將屬於同一站點的兩台OTV路由器中的VLAN拆分。一台邊緣裝置被選舉為所有奇數VLAN的AED，而另一台OTV路由器被選舉為所有偶數VLAN的AED。每個AED也是另一台路由器上處於活動狀態的VLAN的備用裝置。如果其中一個AED中的鏈路或節點發生故障，備用AED對於所有VLAN都變為活動狀態。

如果兩個ASR1000在同一資料中心中連線以進行Multihome，則兩個ASR1000之間無需專用鏈路。

OTV使用透過內部介面傳播的OTV站點VLAN和透過加入介面進行通訊來確定哪些路由器負責奇偶的VLAN。

ASR1000和Nexus 7000不能在同一資料中心內混合使用，且兩台路由器上均配置了OTV作為另一台路由器的備份。相配的平台（ASR1000或Nexus 7000）支援給定資料中心的多主環境。您可以在一個資料中心安裝ASR1000，在另一個資料中心安裝Nexus 7000。這兩個平台之間的互用性已經過測試並受到支援。 有些資料中心可以是多宿的，而有些是單宿的。

多宿主ASR1000路由器對必須運行相同版本的Cisco IOS® XE軟體。

如果使用Multihome，強烈建議必須在OTV路由器上啟用生成樹，因為這將使OTV路由器發出拓撲更改通知(TCN)，從而導致相鄰的L2交換機裝置（以及生成樹中的其他交換機）將它們的老化計時器從預設設定為15秒。 當多宿主對之間存在故障或恢復時，這可以極大地提高速度收斂。透過將子行增加到全局配置中，可以為所有已配置的服務例項（連線到OTV或其他）啟用生成樹。

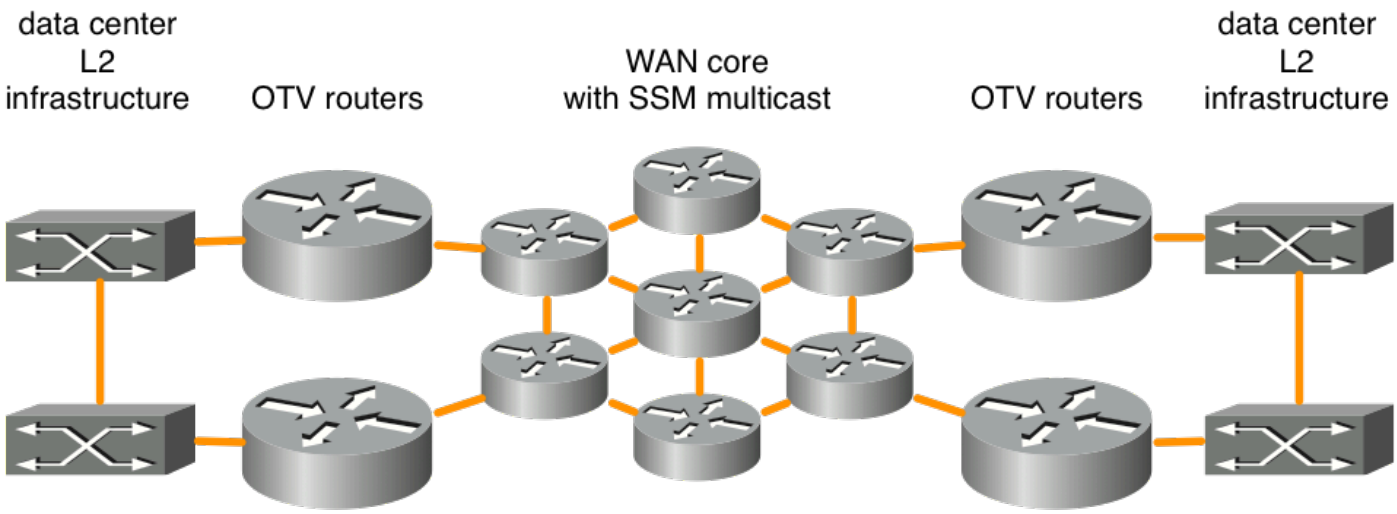
```
spanning-tree mode [ pvst | rapid-pvst | mst ]
```

不需要針對每個vlan或每個服務例項進行特定的配置。

多點傳送核心

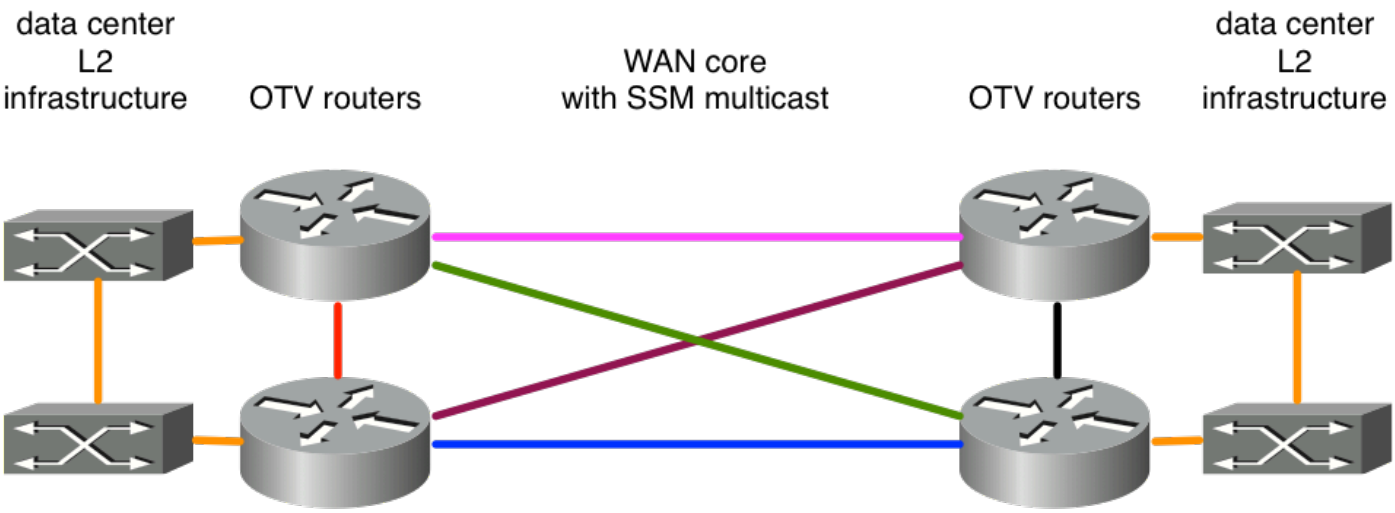
組播網路需要廣域網中的全網狀連線。所有OTV路由器都需要透過加入介面連線在一起。

圖1.支援的組播網路拓撲



下圖顯示了透過全網狀核心連線的兩個資料中心的示例。 源特定組播(SSM)協定獨立組播(PIM)在OTV路由器和WAN核心路由器之間運行。 只要存在全網狀連線，就可以支援任意數量的核心路由器。 對於廣域網核心的OTV連線，沒有明確的最大延遲要求。

圖2.不支援的組播網路拓撲



例如，由於ASR1000/OTV期望在單個加入介面上接收來自所有對等體的組播消息，這將導致OTV部署不穩定。假設將粉紅色和藍色中的東-西鏈路配置為連線介面。 當粉紅色鏈路發生故障時，路由器將無法再在該介面上接收OTV更新。 透過綠色或紫色鏈路的備用路徑是不可接受的，因為已明確配置加入介面。 必須在該介面上接收更新。 目前不支援使用回送介面作為加入介面。

如果使用者不擁有自己的骨幹，他們必須確保服務提供商的核心支援組播，並且服務提供商可以響應網際網路組管理協定(IGMP)查詢消息。ASR1000上的OTV充當組播主機（轉發IGMP加入消息），而不是作為到核心WAN組播拓撲的組播路由器。

OTV路由器之間的傳輸網路必須支援提供商組播組的PIM稀疏模式（任意源組播[ASM]）和交付組的SSM。

組播核心需要在重疊介面上為控制組和用於轉發資料的一系列資料組播組進行某些特定配置。

```
ip multicast-routing distributed
ip pim ssm default
!
interface Port-channel60
 encapsulation dot1Q 30
 ip address 10.0.0.1 255.255.255.0
 ip pim passive
 ip igmp version 3
!
interface Overlay99
 no ip address
 otv control-group 239.1.1.1
 otv data-group 232.192.1.0/24
 otv join-interface Port-ch60
```

組播OTV部署要求將加入介面配置為PIM被動介面。 可以根據需要為不同版本配置IGMP。 重疊介面必須配置控制組和資料組。 控制組是用於OTV管理的單個組播組。 資料組是用於在資料中心之間傳輸使用者資料的一系列組播地址。 如果資料組不在232.0.0.0/8 IP空間中，則必須將附加命令「ip pim ssm range」配置為包括OTV所需的範圍。

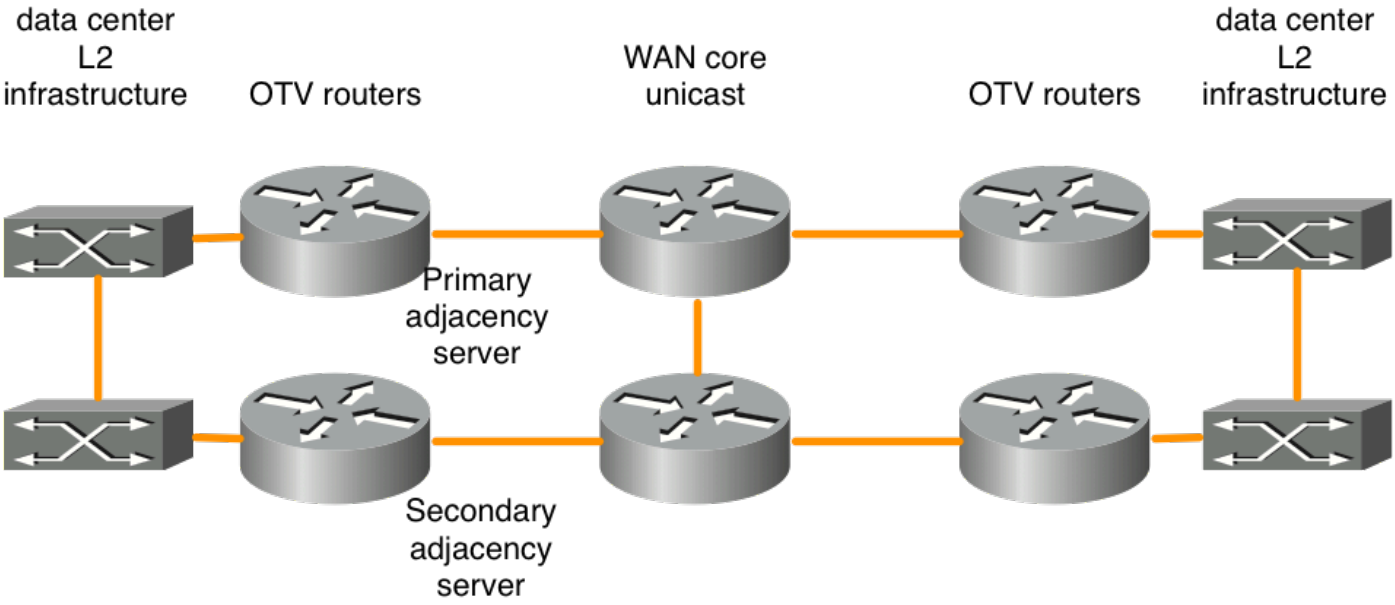
OTV路由器之間的傳輸網路必須支援提供商組播組的PIM稀疏模式（任意源組播[ASM]）和交付組的

源特定組播(SSM)。

帶有鄰接伺服器的單播核心

Cisco IOS® XE 3.9增加了對於具有單播核心的OTV的支援。 所有ASR1000平台和未來版本的Cisco IOS® XE 3.9都繼續支援OTV的單播和組播核心。

圖3.單播網路拓撲



OTV鄰接伺服器功能在OTV邊緣之間啟用僅單播傳輸。 配置了鄰接伺服器角色的OTV路由器會保留所有已知OTV路由器的清單。它們會將該清單提供給所有已註冊的OTV路由器，以便它們擁有必須接收複製的廣播和組播流量的裝置的清單。

僅單播傳輸上的OTV控制平面的工作方式與具有組播核心的OTV完全相同，不同之處在於在單播核心網路中，每個OTV邊緣裝置需要為每個控制平面資料包建立多個副本，並將它們單播到同一邏輯重疊中的每個遠端邊緣裝置。

按照同樣的思路，來自資料中心的任何組播流量都在本地OTV路由器上複製，並且多個副本將傳送到每個遠端資料中心。 與依靠廣域網核心完成複製相比，此方法效率較低，但不需要配置和管理核心組播網路。 如果資料中心組播流量非常少，或者資料中心位置非常少（四個或更少），OTV轉發單播核心通常是最佳選擇。 整體而言，單播模型的操作簡化使得單播核心部署選項在僅需要4個或更少資料中心之間進行LAN擴展連線的情況下更優先。 建議至少配置兩台鄰接伺服器，一台為主伺服器，一台為備份伺服器。 沒有用於主用/主用鄰接伺服器配置的選項。

必須相應地配置OTV路由器，以正確辨識並註冊到相應的鄰接伺服器。

	主要鄰接伺服器	輔助鄰接伺服器	其他OTV路由器
OTV加入介面IP地址	10.0.0.1	10.2.2.24	其他IP地址

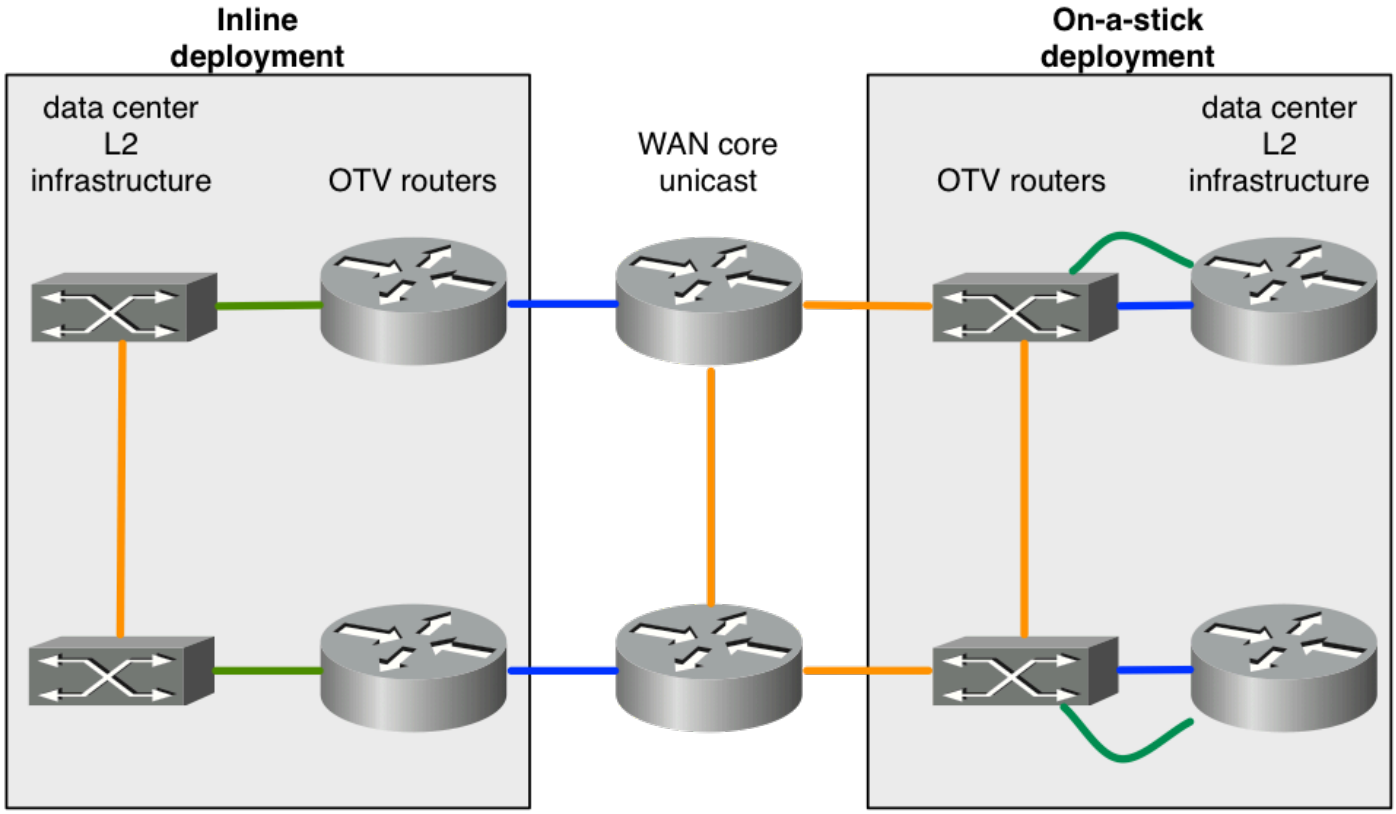
	主要鄰接伺服器	輔助鄰接伺服器	其他OTV路由器
組態	介面重疊1 otv adjacency-server unicast-only	介面重疊1 otv adjacency-server unicast-only otv use-adjacency-server 10.0.0.1 unicast-only	介面重疊1 otv use-adjacency-server 10.0.0.1 10.2.2.24 unicast-only

某些背對背連線設計受單播OTV轉發支援，這些設計不符合「全網狀」規則。 雖然支援這些組態，但不建議使用。 當資料中心透過暗光纖連線時，這種型別的部署最為常見。 有關此配置選項的詳細資訊，請參閱後面的「特殊情況單播拓撲」部分。

OTV on a Stick與Inline

在您的資料中心部署OTV有兩種模式：單臂和內聯。在前面介紹的設計方案中，OTV路由器位於資料中心和服務提供商核心網路之間。但是，最好增加OTV路由器作為不在所有流量傳輸路徑中的裝置。有時，要求是不更改當前拓撲以透過當前裝置連線到服務提供商（例如，使用Catalyst 6000交換機或Nexus交換機硬體進行棕地部署，但不支援OTV）。因此，最好在ASR1000上將OTV部署為單臂作為OTV裝置。

圖4.內聯拓撲與單臂拓撲



該圖展示可以屬於同一重疊的兩種部署模型。 連線到OTV路由器的綠色鏈路被配置為L2接入介面以接受VLAN流量。 連線到OTV路由器的藍色鏈路是承載OTV封裝VLAN流量的加入介面。

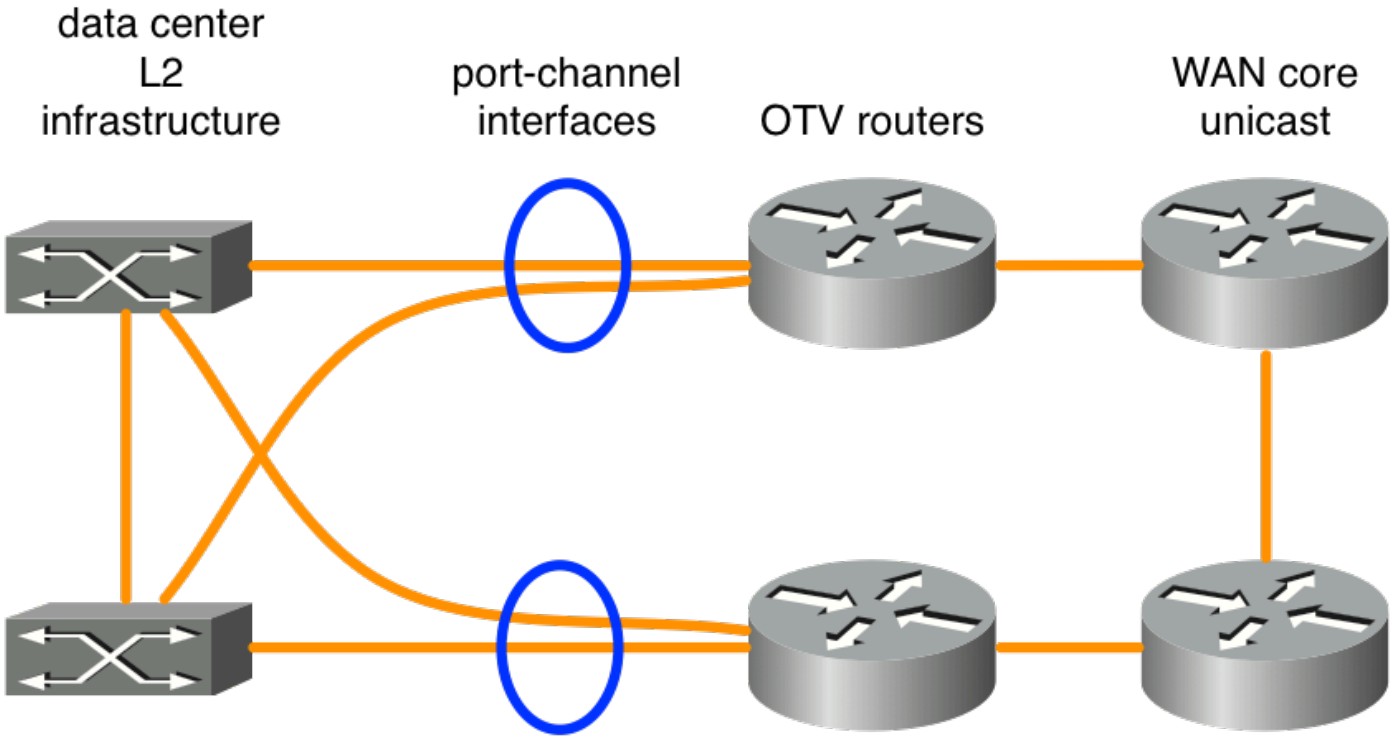
可能需要配置OTV不支援的功能。例如，不能在同一台裝置上配置OTV和MPLS。因此，使用

ASR1000/OTV並在位於OTV路由器前面的路由器上配置MPLS是一個不錯的選擇。

第2層和第3層的埠通道

用於ASR1000的Cisco IOS® XE 3.10代碼增加了支援第2層和第3層埠通道配置與OTV。第2層埠通道可用作內部介面。 Port-channel必須包含最多4個物理介面。第3層埠通道可用作連線介面。

圖5.用於L2連線的埠通道



圖中顯示了一個典型的埠通道方案，其中包含VSS （Catalyst 6000系列）或VPC （Nexus 7000系列）中的兩台交換機。 這種設計透過雙OTV路由器和資料中心基礎設施的雙連線提供冗餘。 如果在與OTV路由器相鄰的L2交換裝置上使用VSS或VPC，則除了基本的埠通道配置之外，不需要對OTV進行特殊配置。

預設閘道

根據定義，OTV會在多個位置建立相同的L3子網。 這需要在將L3流量路由到擴展VLAN或從擴展VLAN路由出時考慮一些特殊事項。 L3路由可以在OTV路由器上配置，也可以在連線到擴展VLAN的其他裝置上配置。 此外，在每個案例中，可部署第一躍點備援通訊協定(FHRP)，例如熱待命備援通訊協定(HSRP)或虛擬路由器備援通訊協定(VRRP)，以提供備援。 HSRP可以在指定資料中心的本地運行，也可以在資料中心之間擴展（非典型）。

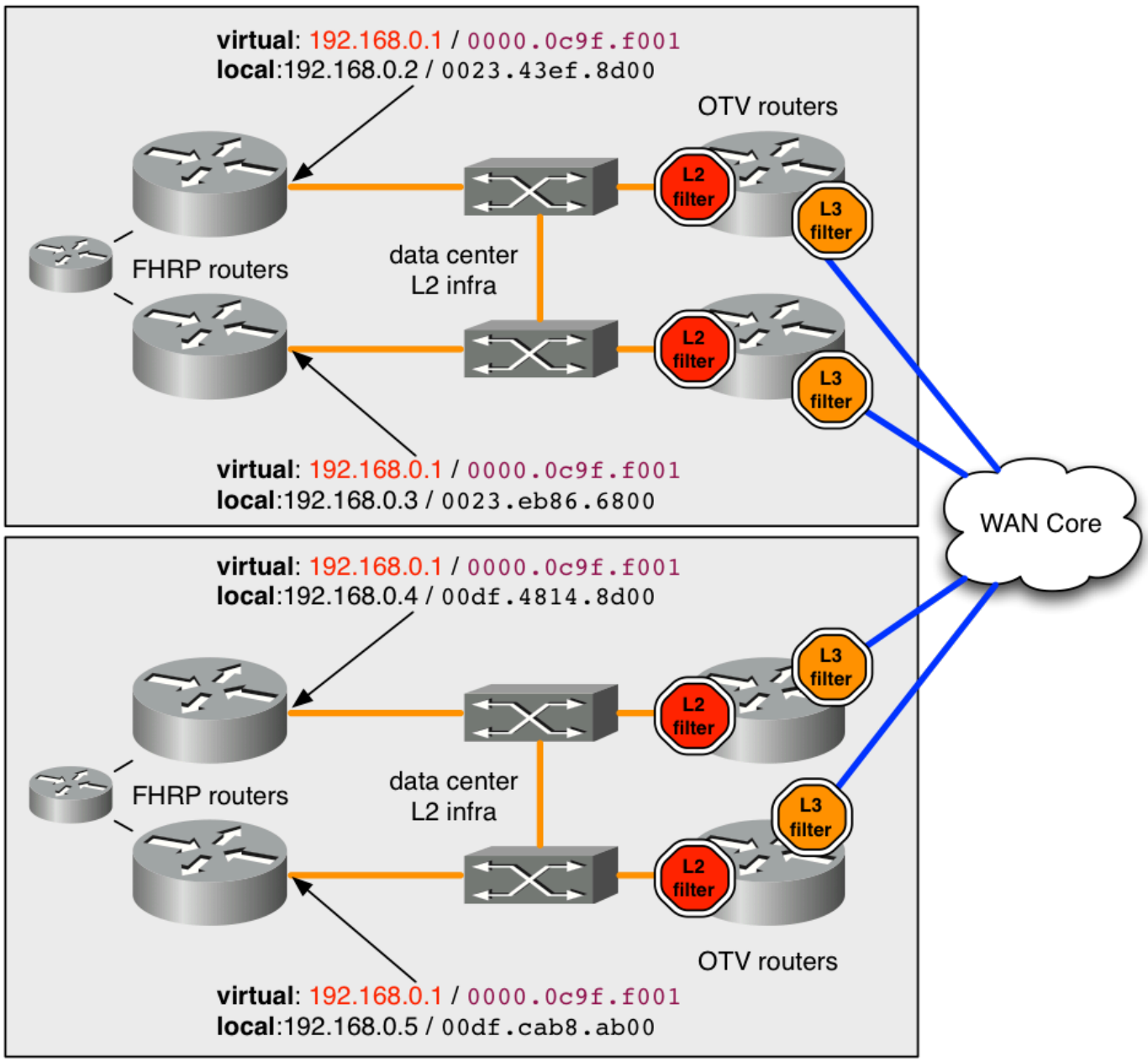
使用FHRP的OTV部署的最佳實踐是在每個資料中心運行FHRP的本地例項。 這些FHRP例項使用相同的虛擬MAC地址和IP地址，因此當虛擬機器(VM)在資料中心之間移動時，它們具有不間斷連線。 如果在資料中心之間更改預設路由器的MAC地址，則虛擬機器將無法通過子網進行通訊，直到虛擬機器的預設網關ARP條目超時。

若要正確部署具有OTV的FHRP，必須考慮必須過濾哪些第2層和第3層流量並從OTV中隔離。在L2層級，必須這樣才能防止OTV看到多個位置的FHRP使用的同一L2虛擬MAC。 L3級別需要過濾

器，以將HSRP和VRRP通告隔離到每個資料中心，從而將活動/偵聽/備用選擇定位到每個資料中心。

預設情況下，啟用OTV時會啟用FHRP過濾器。 如果設計要求在資料中心之間擴展FHRP，則可以停用該功能。 虛擬MAC地址的L2過濾預設處於未啟用狀態，必須手動配置。

圖6. 建議的FHRP部署示例



在示例中，虛擬MAC地址0000.0c9f.f001用於IP地址192.168.0.1，該地址承載擴展VLAN以連線子網。 在兩個資料中心使用相同的虛擬MAC和IP，主機在資料中心之間傳輸時，可無縫連線子網。

為了將MAC地址0000.0c9f.f001隱藏在多個位置的OTV中，必須在為VLAN提供服務的每個OTV路由器上為VLAN部署入口L2過濾器（圖中的紅色標籤）。透過ACL過濾在L2服務例項上為入口配置的過濾器ACL，源自MAC的所有資料包都會在ASR1000上的OTV進程看到它們之前被丟棄。 因此，OTV從不瞭解MAC，也不向遠端資料中心通告MAC。

此處提供了用於捕獲所有已知/預設FHRP虛擬MAC流量的建議配置。

```
mac access-list extended otv_filter_fhrp
 deny 0000.0c07.ac00 0000.0000.00ff any
 deny 0000.0c9f.f000 0000.0000.0fff any
 deny 0007.b400.0000 0000.0000.00ff any
 deny 0000.5e00.0100 0000.0000.00ff any
 permit any any
```

此ACL匹配與HSRP版本1和2、網關負載均衡協定(GLBP)和VRRP相關的已知MAC地址空間（按該順序）。如果虛擬MAC配置為使用非基於FHRP組號的非標準值，則必須將其明確增加到ACL示例。 必須將ACL增加到L2服務例項（如下所示）。

```
interface Port-channel10
  description *** OTV internal interface ***
  no ip address
  no negotiation auto
!
  service instance 800 ethernet
    encapsulation dot1q 800
    mac access-group otv_filter_fhrp in
    bridge-domain 800
```

此外，還需要管理L3級別的FHRP主機之間的通訊。 圖中的單個擴展子網中配置了4台FHRP路由器。 如果沒有某種程度的L3過濾器，則所有四台路由器將會看到對方並選擇一個活動裝置，並且有3台處於各種備用狀態。 因此，一個資料中心將有兩個本地備用FHRP路由器，但由於前面討論的L2過濾器而沒有與遠端活動路由器的L2連線。

理想的結果是在每個資料中心擁有一個主用FHRP路由器和一台備用FHRP路由器。 前面討論的入口L2過濾器不會捕獲此選舉流量，因為選舉過程使用路由器的實際IP和MAC地址。 預設情況下，後續ACL在重疊介面上作為出口應用。 重疊介面的出口將是流向廣域網核心的流量。 ACL不會顯示在運行配置中，但可以使用「show ip access-list」進行觀察。 它根據UDP埠號過濾掉FHRP選舉流量。

```
Extended IP access list otv_fhrp_filter_acl
    10 deny udp any any eq 1985 3222
    20 deny 112 any any
    30 permit ip any
```

停用此過濾器的唯一原因是您希望某個VLAN上的所有FHRP路由器都參與相同的活動狀態選擇。若要停用此過濾器，請在重疊介面上設定「no otv filter-fhrp」。

未知的單點傳播流量

預設情況下，OTV路由器從LAN收到的發往遠端OTV位置未知的MAC地址的單播流量將被丟棄。此流量稱為未知單播。 此丟棄操作用於限制WAN上廣播流量消耗的頻寬量的WAN核心。 一般預

期是，LAN上的所有主機都會發出足夠多的廣播流量（ARP、協定廣播等），這些流量始終會被OTV路由器看到、通告，因此是「已知」的。

某些應用程式會利用靜默式主機。 在正常的交換基礎設施中，這不是問題，因為LAN上未知單播MAC地址的L2廣播允許靜默主機檢視流量。 但是，在OTV環境中，OTV路由器會阻止資料中心之間的流量。

為了彌補這一點，Cisco IOS® XE整合了稱為選擇性單播轉發的功能。XE 3.10.6、XE3.13.3、XE 3.14.1、XE3.15以及之後的所有版本都支援選擇性單播轉發。

可透過在重疊介面上為每個MAC地址增加一個命令來配置它。 舉例來說：

```
interface Overlay1
  service instance 100 ethernet
    encapsulation dot1q 100
    otv mac flood 0000.0000.0001
    bridge-domain 100
```

在本例中，所有發往0000.0001.0001的流量都必須泛洪到所有使用VLAN 100的遠端OTV路由器。這可透過後續命令進行觀察：

```
<#root>

OTV_router_1#

show otv route


Codes: BD - Bridge-Domain, AD - Admin-Distance,SI - Service Instance, * - Backup Route
OTV Unicast MAC Routing Table for Overlay99
 Inst VLAN BD     MAC Address    AD    Owner  Next Hops(s)
-----------------------------------------------------------
 0    100  100    0000.0000.0001 20    OTV    Flood
```

如果在遠端站點獲知了該MAC地址，則必須將優先於泛洪條目的條目增加到轉發表中。

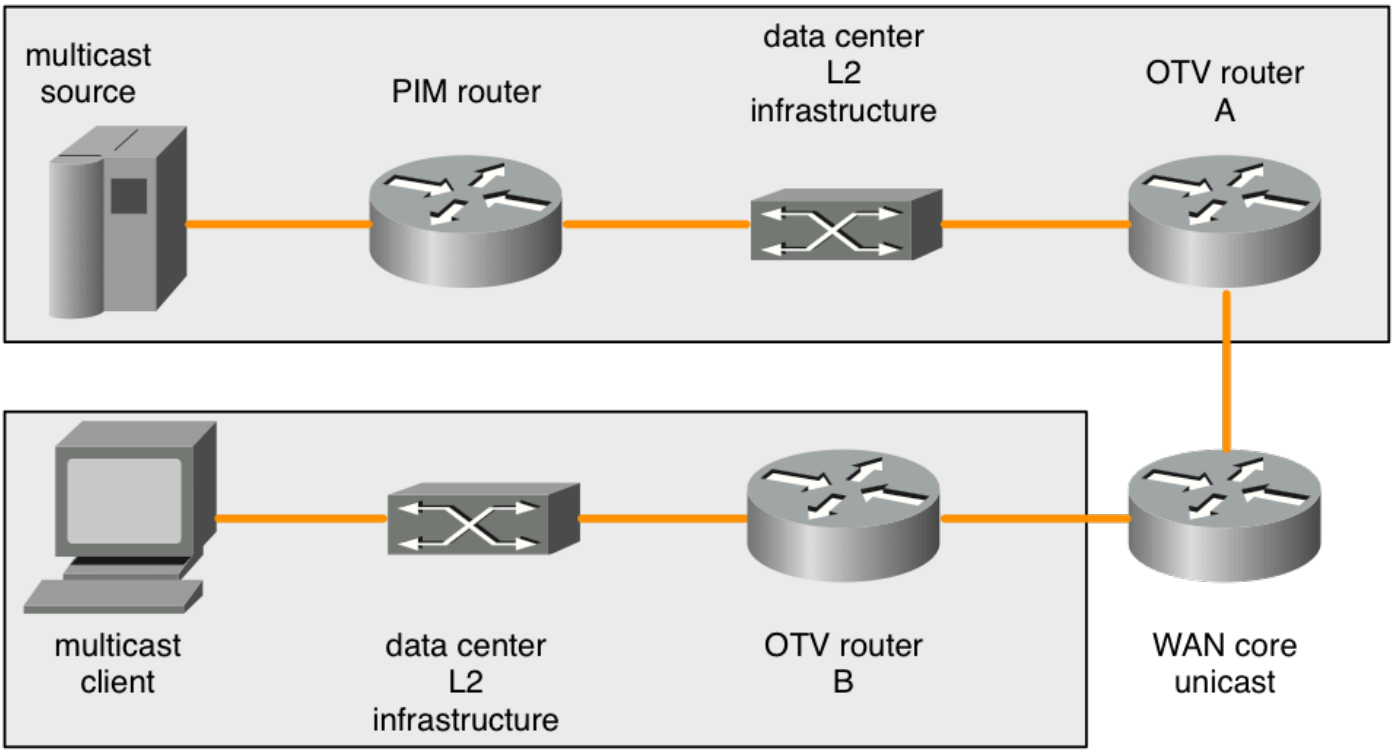
```
<#root>

OTV_router_1#

show otv route


Codes: BD - Bridge-Domain, AD - Admin-Distance,SI - Service Instance, * - Backup Route
OTV Unicast MAC Routing Table for Overlay99
 Inst VLAN BD     MAC Address    AD    Owner  Next Hops(s)
-----------------------------------------------------------
 0    100  100    0000.0000.0001 20    OTV    Flood
 0    100  100    0000.0000.0001 50    ISIS   OTV_router_3
```

通常，必須在具有該VLAN的所有OTV路由器上配置給定MAC地址的泛洪條目。

遠端組播源

ASR1000認為OTV路由器不會轉發從LAN接收的多播IGMP加入請求。後續圖表詳細說明了可能導致此問題的拓撲。

圖7.遠端組播源



當組播客戶端傳送組播IGMP加入時，ASR1000（OTV路由器B）會觀察並通告對組播組的興趣。遠端OTV路由器（OTV路由器A）必須將任何流量轉發到他們在本地L2廣播域中看到的該組播組。但是，當從客戶端的OTV路由器（OTV路由器B）向廣播對組播組的興趣時，遠端ASR1000（OTV路由器A）不會重新生成組播IGMP加入請求。

當組播源與OTV路由器位於同一個L2廣播域時，則不存在問題。 OTV路由器必須配置為IGMP查詢器。 這會顯示在L2廣播域中存在的任何組播流量中。 但是，只有PIM加入請求會導致PIM路由器將組播源從不同的L2廣播域轉發到OTV路由器所在的L2廣播域。

遠端IGMP加入要求未轉送或重新產生。OTV路由器也不是PIM路由器。 因此，當遠端客戶端感興趣時，組播源不直接位於OTV路由器的L2廣播域中的拓撲無法進入PIM路由器轉發源流量。

此問題有兩個解決方法。

首先，可以在連線到OTV路由器（OTV路由器A）的L2廣播域上部署本地IGMP客戶端。該IGMP客戶端必須訂閱遠端客戶端可以訂閱的任何組播組。 這將導致PIM路由器將組播流量轉發到與OTV路由器A相鄰的廣播域。 然後，IGMP查詢將提取任何組播流量，並透過重疊傳送。

另一種解決方案是為遠端客戶端可能訂閱的任何組配置「ip igmp static-join」。 這也會導致PIM路由器將組播流量轉發到與OTV路由器A相鄰的廣播域。

此限制是已知的，並且是設計規範的一部分。 目前不認為是Bug，而是支援拓撲的限制。

QoS注意事項

在ASR1000上，預設情況下，增加的OTV報頭中的TOS值從L2資料包的802.1p位複製。 如果L2資料包未標籤，則使用零值。

Nexus 7000在5.2.1及更高版本的軟體中具有不同的預設行為。 如果期望的行為是將內部資料包TOS值複製到外部，則可透過附加QoS配置來實現此目的。 這與較新的Nexus 7000軟體的行為相同。

將L2資料包L3 TOS值複製到OTV資料包最外部報頭的配置是後續：

```
class-map dscp-af11
 match dscp af11
!
class-map dscp-af21
 match dscp af21
!
class-map qos11
 match qos-group 11
!
class-map qos21
 match qos-group 21
!
policy-map in-mark
 class dscp-af11
  set qos-group 11
 class dscp-af21
  set qos-group 21
!
policy-map out-mark
 class qos11
  set dscp af11
 class qos21
  set dscp af21
!
interface Gig0/0/0
 ! L2 interface
 service instance 100 ethernet
  encapsulation dot1q 100
  service-policy in-mark
  bridge-domain 100
!
interface Gig0/0/1
 ! OTV join interface
 service-policy out-mark
```

提供的配置必須與入口上各種DSCP值的流量匹配。 本地有效的qos組標籤用於在路由器傳輸期間內部標籤該流量。 在出口介面，qos組匹配，然後最外部的TOS位元組相應地更新。

WAN MTU注意事項/分段

OTV實際上使用GRE報頭透過WAN傳輸L2流量。此GRE報頭的大小為42位元組。 在理想的網路部署中，WAN鏈路的最大傳輸單元(MTU)必須至少比OTV預期處理的最大資料包大42位元組。

如果L2介面的MTU為1500位元組，則加入介面的MTU必須為1542位元組或更多。 如果L2介面的MTU為2000位元組，但預期只處理最大為1500位元組的資料包，則1542位元組的WAN MTU就足夠了，但是2000標準相加42比較理想。

```
interface GigabitEthernet0/0/0
  mtu 1600
!
interface Overlay 1
  otv join-interface GigabitEthernet0/0/0
!
interface GigabitEthernet0/0/1
  mtu 1500
  service instance 100 ethernet
    encapsulation dot1q 100
    bridge-domain 100
  service instance 101 ethernet
    encapsulation dot1q 101
    bridge-domain 101
```

一些服務提供商無法為其WAN電路提供更大的MTU值。 如果出現這種情況，ASR1000可以對OTV傳輸的資料執行分段。 Nexus 7000不具備此功能。 不支援在ASR1000上啟用了分段的混合ASR1000和Nexus 7000 OTV網路。

OTV分段的配置為：

```
otv fragmentation join-interface GigabitEthernet0/0/0
!
interface Overlay 1
  otv join-interface GigabitEthernet0/0/0
```

務必在Overlay interface join-interface命令之前配置全局級命令。 如果首先配置重疊介面的otv join-interface命令，請從重疊介面刪除otv join-interface命令，配置otv fragmentation join-interface命令，然後再次配置重疊介面的otv join-interface命令。

未啟用OTV分段時，傳送封裝的L2資料的所有OTV封包都會設定DF位元，如此一來，它們在傳輸過程中就不會分段。 新增分段命令後，DF位元會設定為0。OTV路由器本身可以對封包進行分段，且可在其他路由器傳送時對其進行分段。

ASR1000平台上可用的資料包重組緩衝區數量有限，因此資料包必須截斷的片段越少才能更好地傳輸。 如果存在問題，這將提高效率並降低整個廣域網的整體頻寬消耗。 啟用OTV分段具有效能影響。 如果存在分段，並且預期處理超過1Gb/sec的OTV流量，則必須進一步調查OTV效能。

特殊情況單播拓撲

OTV的現場部署通常在兩個資料中心的OTV路由器之間採用直接背對背光纖連線。

對於單宿主拓撲，這使OTV和非OTV流量共用加入介面的標準部署成為可能。 此設定不需要特殊考量，因此本節不適用。

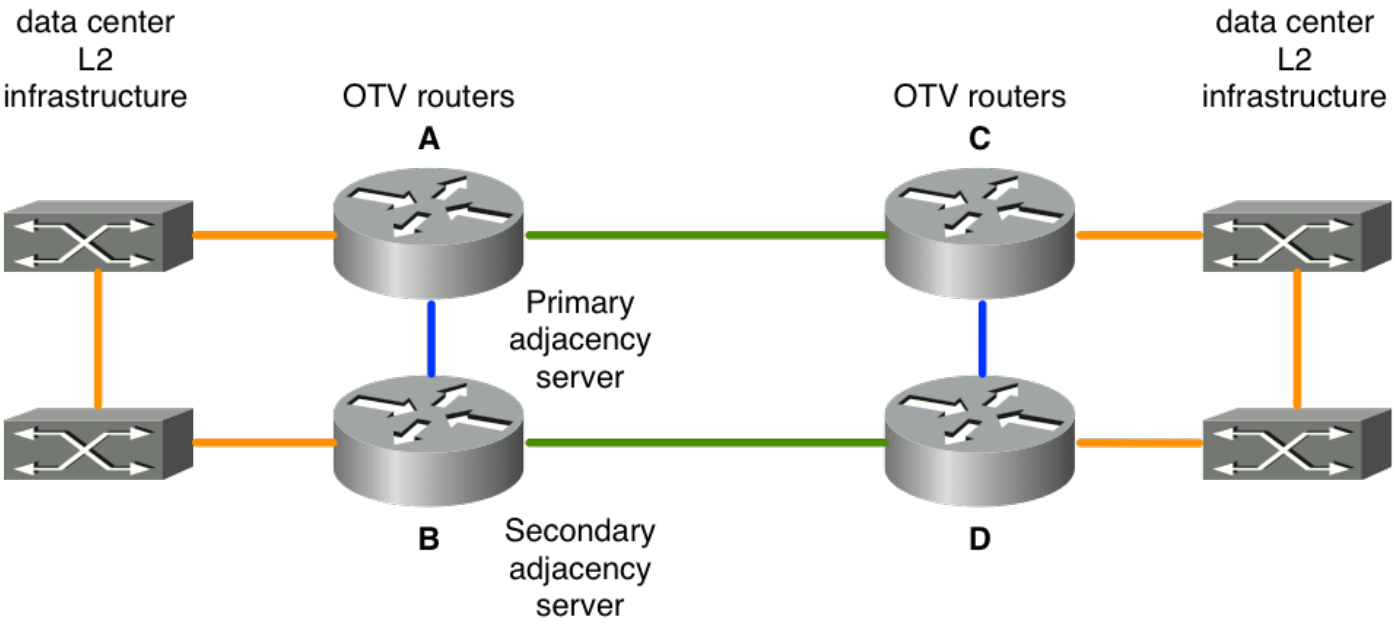
但是，如果部署在兩個資料中心中使用多宿主OTV路由器，則有一些特殊注意事項。 需要其他配置。

如果涉及兩個以上的資料中心，則此特殊配置不適用。

對於使用單或多宿主OTV路由器的兩個以上資料中心的場景，必須使用標準單播或組播OTV部署。

沒有其他支援的替代方案。

圖8. 特殊案例單播



在顯示的拓撲中，綠色鏈路是兩個資料中心之間的暗光纖鏈路。 這些暗光纖直接連線到OTV路由器。 OTV路由器之間的藍色鏈路用於在綠色鏈路出現故障時重新路由非OTV流量。 如果頂部綠色鏈路發生故障（A到C），使用最頂層OTV路由器作為其預設路由的非OTV流量將透過南北藍色鏈路（A到B和C到D）路由到底部OTV路由器對（B到D）之間仍然運行的綠色鏈路。

這種基本的流量重路由對OTV流量不起作用，因為OTV配置將物理介面指定為加入介面。 如果OTV路由器A上的「綠色介面」斷開，則OTV流量不能從備用介面OTV路由器B獲得。此外，由於未通過WAN核心實現完全連線，因此在出現故障時無法通知所有OTV路由器。 為了解決此問題，使用了雙向轉發檢測(BFD)和嵌入式事件管理器(EEM)指令碼。

BFD必須監控東西OTV路由器對（A/C和B/D）之間的WAN鏈路。如果與遠端路由器的連線丟失，則透過東西OTV路由器對（A/C和B/D）上的EEM指令碼關閉OTV覆蓋介面。 這會導致成對的多宿主路由器承擔所有VLAN的轉發任務。 當BFD檢測到鏈路已恢復時，EEM指令碼會觸發以重新啟用覆蓋介面。

使用BFD檢測鏈路故障非常重要。 這是因為重疊介面的「故障」端及其東西對都需要關閉。 根據服務提供商提供的連線型別而定，一條物理鏈路可能斷開（OTV路由器A上的綠色介面），而對應的東西對路由器的介面可能保持運行（OTV路由器C上的綠色介面）。BFD檢測任一介面出現故障或傳輸中的任何其它問題，並立即同時通知兩對。 這同樣適用於需要通知路由器恢復鏈路的情況。

此部署的配置與其他任何部署相同，並增加了後續項：

- WAN介面上的BFD配置
- 後續EEM指令碼
- 匹配偶數/奇數VLAN分佈的OTV ISIS身份

OTV加入介面上的BFD配置不在本文檔的討論範圍之內。 有關如何在ASR1000上配置BFD的資訊，請訪問：

https://www.cisco.com/c/en/us/td/docs/ios-xml/ios/iproute_bfd/configuration/xe-3s/irb-xe-3s-book.html

一旦連線介面對（圖中的綠色鏈路）之間的BFD故障檢測工作正常，就必須部署EEM指令碼。 EEM指令碼必須針對特定路由器進行定製，以修改正確的重疊介面，並可能監控日誌中更精確的字串，以瞭解BFD故障和恢復。

```
event manager environment _OverlayInt Overlay1
!
event manager applet WatchBFDdown
description "Monitors BFD status, if it goes down, bring OVERLAY int down"
event syslog pattern "BFD peer down notified" period 1
action 1.0 cli command "enable"
action 2.0 cli command "config t"
action 2.1 syslog msg "EEM: WatchBFDdown will shut int $_OverlayInt"
action 3.0 cli command "interface $_OverlayInt"
action 4.0 cli command "shutdown"
action 5.0 syslog msg "EEM WatchBFDdown COMPLETE ..."
!
event manager applet WatchBFDup
description "Monitors BFD status, if it goes up, bring OVERLAY int up"
event syslog pattern "new adjacency" period 1
action 1.0 cli command "enable"
action 2.0 cli command "config t"
action 2.1 syslog msg "EEM: WatchBFDup bringing up int $_OverlayInt"
action 3.0 cli command "interface $_OverlayInt"
action 4.0 cli command "no shutdown"
action 5.0 syslog msg "EEM WatchBFDup COMPLETE ..."
!
```

這種型別的部署還要求東-西路由器對（A/C和B/D）在轉發奇數和偶數vlan時匹配。

例如，A和C必須轉發偶數VLAN，而B和D在穩態標稱操作中轉發奇數VLAN。

奇/偶分佈由OTV序號確定，該序號可透過「show otv site」命令進行觀察。

兩個站點路由器之間的序號基於OTV ISIS網路ID確定。

```
OTV_router_A#show otv site
Site Adjacency Information (Site Bridge-Domain: 99)
Overlay99 Site-Local Adjacencies (Count: 2)
  Hostname       System ID      Last Change Ordinal    AED Enabled Status
* OTV_router_A   0021.D8D4.F200 19:32:02    0          site       overlay
  OTV_router_B   0026.CB0C.E200 19:32:46    1          site       overlay
```

必須在所有OTV路由器上配置OTV ISIS網路識別符號。 在配置識別符號時必須小心，以使所有OTV路由器仍然可以相互辨識。

<#root>

```
OTV router A:
otv isis Site
 net
49
.
0001
.
0001
.
0001
.
000a
.
00
```

```
OTV router B:
otv isis Site
 net
49
.
0001
.
0001
.
0001
.
000b
```

.

00

OTV router C:
otv isis Site
 net

49

.

0001

.

0001

.

0001

.

000c

.

00

OTV router

D:
otv isis Site
 net

49

.

0001

.

0001

.

0001

.

000d

.

00

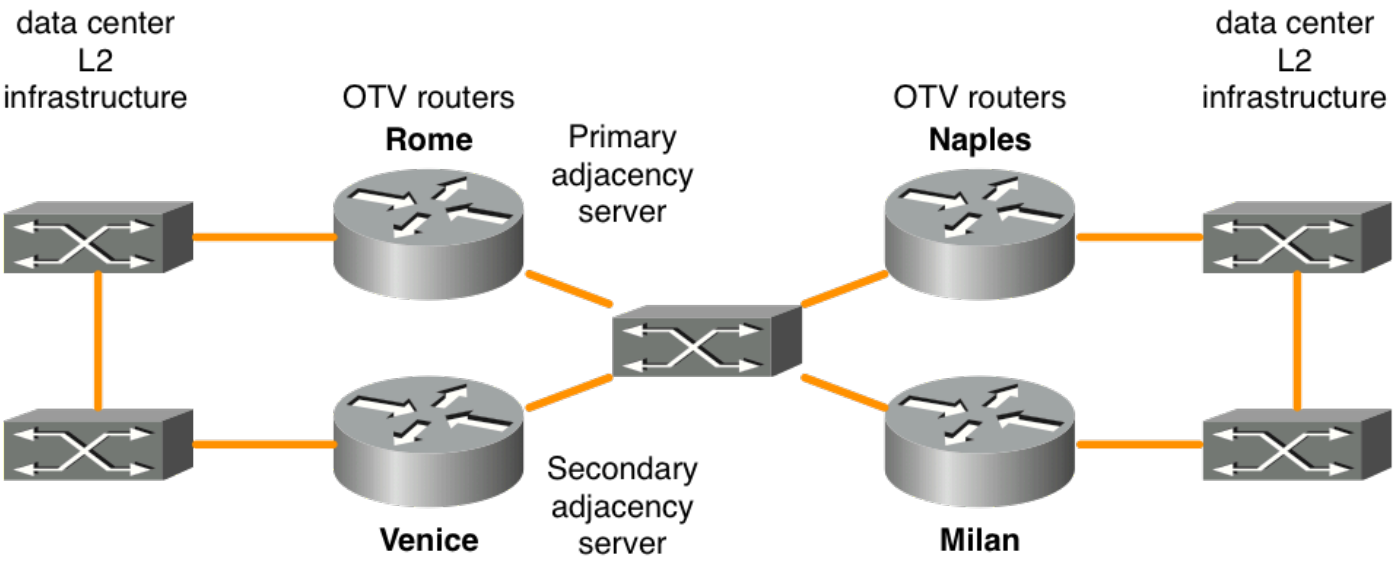
在參與重疊的所有OTV路由器上，識別符號的黑色部分必須匹配。 可以修改識別符號的紅色部分

。 站點上的最低網路識別符號得到序號0，然後轉發偶數編號的VLAN。 站點的最高網路識別符號得到序數1並轉發奇數VLAN。

組態範例

單點傳播

圖9.單播配置示例



Rome配置：

```
!
hostname Rome
!
ip igmp snooping querier version 3
ip igmp snooping querier
!
otv site bridge-domain 99
!
otv site-identifier 0000.0000.0001
!
spanning-tree mode pvst
!
interface Overlay99
 no ip address
 otv join-interface GigabitEthernet1/0/0
 otv adjacency-server unicast-only
 service instance 100 ethernet
  encapsulation dot1q 100
  bridge-domain 100
 !
 service instance 101 ethernet
  encapsulation dot1q 101
  bridge-domain 101
 !
interface GigabitEthernet1/0/0
 ip address 172.16.0.1 255.255.255.0
 negotiation auto
```

```
 cdp enable
!
interface GigabitEthernet1/0/1
 no ip address
 negotiation auto
 cdp enable
 service instance 99 ethernet
  encapsulation dot1q 99
  bridge-domain 99
 !
 service instance 100 ethernet
  encapsulation dot1q 100
  bridge-domain 100
 !
 service instance 101 ethernet
  encapsulation dot1q 101
  bridge-domain 101
 !
```

威尼斯配置：

```
!
hostname Venice
!
ip igmp snooping querier version 3
ip igmp snooping querier
!
otv site bridge-domain 99
!
otv site-identifier 0000.0000.0001
!
spanning-tree mode pvst
!
interface Overlay99
 no ip address
 otv join-interface GigabitEthernet0/0/0
 otv adjacency-server unicast-only
 otv use-adjacency-server 172.16.0.1 unicast-only
 service instance 100 ethernet
  encapsulation dot1q 100
  bridge-domain 100
 !
 service instance 101 ethernet
  encapsulation dot1q 101
  bridge-domain 101
 !
!
interface GigabitEthernet0/0/0
 ip address 172.16.0.2 255.255.255.0
 negotiation auto
 cdp enable
!
interface GigabitEthernet0/0/1
 no ip address
 negotiation auto
 cdp enable
 service instance 99 ethernet
  encapsulation dot1q 99
```

```
  bridge-domain 99
 !
 service instance 100 ethernet
  encapsulation dot1q 100
  bridge-domain 100
 !
 service instance 101 ethernet
  encapsulation dot1q 101
  bridge-domain 101
 !
```

那不勒斯配置：

```
!
hostname Naples
!
ip igmp snooping querier version 3
ip igmp snooping querier
!
otv site bridge-domain 99
!
otv site-identifier 0000.0000.0002
!
spanning-tree mode pvst
!
interface Overlay99
 no ip address
 otv join-interface GigabitEthernet0/0/0
 otv use-adjacency-server 172.16.0.1 172.16.0.2 unicast-only
 service instance 100 ethernet
  encapsulation dot1q 100
  bridge-domain 100
 !
 service instance 101 ethernet
  encapsulation dot1q 101
  bridge-domain 101
 !
!
interface GigabitEthernet0/0/0
 ip address 172.16.0.3 255.255.255.0
 negotiation auto
 cdp enable
!
interface GigabitEthernet0/0/1
 no ip address
 negotiation auto
 cdp enable
 service instance 99 ethernet
  encapsulation dot1q 99
  bridge-domain 99
 !
 service instance 100 ethernet
  encapsulation dot1q 100
  bridge-domain 100
 !
 service instance 101 ethernet
  encapsulation dot1q 101
  bridge-domain 101
```

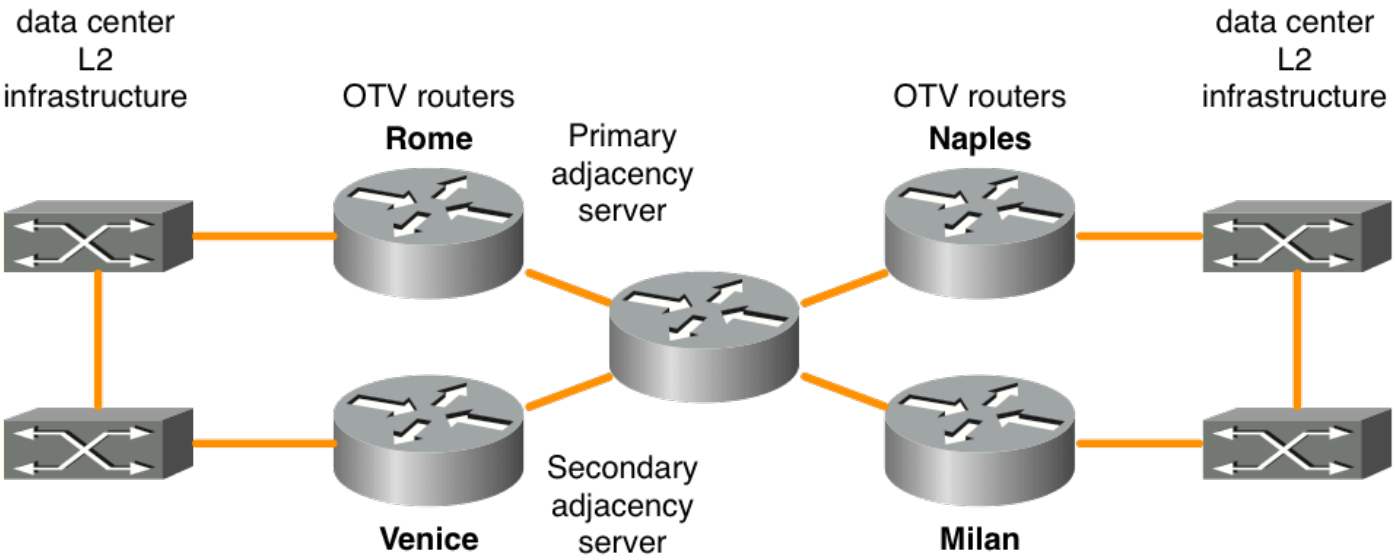
```
 !
!
```

米蘭配置：

```
!
hostname Milan
!
ip igmp snooping querier version 3
ip igmp snooping querier
!
otv site bridge-domain 99
!
otv site-identifier 0000.0000.0002
!
spanning-tree mode pvst
!
interface Overlay99
 no ip address
 otv join-interface GigabitEthernet0/0/0
 otv use-adjacency-server 172.16.0.1 172.16.0.2 unicast-only
 service instance 100 ethernet
  encapsulation dot1q 100
  bridge-domain 100
 !
 service instance 101 ethernet
  encapsulation dot1q 101
  bridge-domain 101
 !
!
interface GigabitEthernet0/0/0
 ip address 172.16.0.4 255.255.255.0
 negotiation auto
 cdp enable
!
interface GigabitEthernet0/0/1
 no ip address
 negotiation auto
 cdp enable
 service instance 99 ethernet
  encapsulation dot1q 99
  bridge-domain 99
 !
 service instance 100 ethernet
  encapsulation dot1q 100
  bridge-domain 100
 !
 service instance 101 ethernet
  encapsulation dot1q 101
  bridge-domain 101
 !
!
```

多點傳播

圖10.組播配置示例



Rome配置：

```
!
hostname Rome
!
ip multicast-routing distributed
!
ip igmp snooping querier version 3
ip igmp snooping querier
!
otv site bridge-domain 99
!
otv site-identifier 0000.0000.0001
!
spanning-tree mode pvst
!
interface Overlay99
 no ip address
 otv join-interface GigabitEthernet1/0/0
 otv control-group 239.0.0.1
 otv data-group 238.1.2.0/24
!
 service instance 100 ethernet
  encapsulation dot1q 100
  bridge-domain 100
 !
 service instance 101 ethernet
  encapsulation dot1q 101
  bridge-domain 101
 !
!
interface GigabitEthernet1/0/0
 ip address 192.168.0.1 255.255.255.0
 ip pim passive
 ip igmp version 3
 negotiation auto
 cdp enable
!
interface GigabitEthernet1/0/1
```

```
 no ip address
 negotiation auto
 cdp enable
!
 service instance 99 ethernet
  encapsulation dot1q 99
  bridge-domain 99
 !
 service instance 100 ethernet
  encapsulation dot1q 100
  bridge-domain 100
 !
 service instance 101 ethernet
  encapsulation dot1q 101
  bridge-domain 101
 !
```

威尼斯配置：

```
!
hostname Venice
!
ip multicast-routing distributed
!
ip igmp snooping querier version 3
ip igmp snooping querier
!
otv site bridge-domain 99
!
otv site-identifier 0000.0000.0001
!
spanning-tree mode pvst
!
interface Overlay99
 no ip address
 otv join-interface GigabitEthernet0/0/0
 otv control-group 239.0.0.1
 otv data-group 238.1.2.0/24
 !
 service instance 100 ethernet
  encapsulation dot1q 100
  bridge-domain 100
 !
 service instance 101 ethernet
  encapsulation dot1q 101
  bridge-domain 101
 !
!
interface GigabitEthernet0/0/0
 ip address 172.17.0.1 255.255.255.0
 ip pim passive
 ip igmp version 3
 negotiation auto
 cdp enable
!
interface GigabitEthernet0/0/1
 no ip address
 negotiation auto
```

```
 cdp enable
!
 service instance 99 ethernet
  encapsulation dot1q 99
  bridge-domain 99
 !
 service instance 100 ethernet
  encapsulation dot1q 100
  bridge-domain 100
 !
 service instance 101 ethernet
  encapsulation dot1q 101
  bridge-domain 101
 !
```

那不勒斯配置：

```
!
hostname Naples
!
ip multicast-routing distributed
!
ip igmp snooping querier version 3
ip igmp snooping querier
!
otv site bridge-domain 99
!
otv site-identifier 0000.0000.0002
!
spanning-tree mode pvst
!
interface Overlay99
 no ip address
 otv join-interface GigabitEthernet0/0/0
 otv control-group 239.0.0.1
 otv data-group 238.1.2.0/24
!
 service instance 100 ethernet
  encapsulation dot1q 100
  bridge-domain 100
 !
 service instance 101 ethernet
  encapsulation dot1q 101
  bridge-domain 101
 !
!
interface GigabitEthernet0/0/0
 ip address 172.18.0.1 255.255.255.0
 ip pim passive
 ip igmp version 3
 negotiation auto
 cdp enable
!
interface GigabitEthernet0/0/1
 no ip address
 negotiation auto
 cdp enable
 service instance 99 ethernet
```

```
  encapsulation dot1q 99
  bridge-domain 99
 !
 service instance 100 ethernet
  encapsulation dot1q 100
  bridge-domain 100
 !
 service instance 101 ethernet
  encapsulation dot1q 101
  bridge-domain 101
 !
!
```

米蘭配置：

```
!
hostname Milan
!
ip multicast-routing distributed
!
ip igmp snooping querier version 3
ip igmp snooping querier
!
otv site bridge-domain 99
!
otv site-identifier 0000.0000.0002
!
spanning-tree mode pvst
!
interface Overlay99
 no ip address
 otv join-interface GigabitEthernet0/0/0
otv control-group 239.0.0.1
 otv data-group 238.1.2.0/24
!
  service instance 100 ethernet
  encapsulation dot1q 100
  bridge-domain 100
 !
 service instance 101 ethernet
  encapsulation dot1q 101
  bridge-domain 101
 !
!
interface GigabitEthernet0/0/0
 ip address 172.19.0.1 255.255.255.0
 ip pim passive
 ip igmp version 3
 negotiation auto
 cdp enable
!
interface GigabitEthernet0/0/1
 no ip address
 negotiation auto
 cdp enable
 service instance 99 ethernet
  encapsulation dot1q 99
  bridge-domain 99
```

```
 !
 service instance 100 ethernet
  encapsulation dot1q 100
  bridge-domain 100
 !
 service instance 101 ethernet
  encapsulation dot1q 101
  bridge-domain 101
 !
!
```

常見問題

問：是否支援與OTV結合使用專用VLAN？

A)是，OTV不需要特殊配置。在專用VLAN配置中，確保連線OTV L2介面的交換機埠配置在混合模式下。

問：IPSEC加密是否支援OTV？

A)是，支援連線介面上的加密對映配置。 OTV無需特殊配置即可支援加密。 但是，加密配置會增加額外開銷，並且必須透過WAN MTU與LAN MTU的增加來補償此開銷。如果無法做到這一點，則必須進行OTV分段。 OTV效能限制為IPSEC硬體效能。

問：MACSEC是否支援OTV？

答：是的，ASR1001-X包括內建介面的MACSEC支援。 OTV可與在LAN和/或WAN介面上配置的MACSEC配合使用。 OTV效能限制為MACSEC硬體效能。

問：環回介面能否用作加入介面？

A)否，只有乙太網、Portchannel或POS介面可用作OTV加入介面。 OTV Loopback join interface已在規劃圖中，但目前尚未計畫發行版本。

問：通道介面能否用作加入介面？

答：否，不支援將GRE隧道、DMVPN隧道或任何其他型別的隧道作為加入介面。 只有乙太網、Portchannel或POS介面可用作OTV加入介面。

問：不同的重疊介面能否使用不同的L2和/或連線介面？

A)所有重疊介面必須指向同一個加入介面。 所有重疊必須連結至相同的實體介面，才能連線至資料中心的L2連線。

Q) OTV站點VLAN是否可以與OTV擴展VLAN位於不同的物理介面上？

A) OTV站點VLAN和擴展VLAN必須位於同一物理介面上。

問：OTV需要什麼功能集？

A) OTV需要高級IP服務(AIS)或高級企業服務(AES)。

問：固定配置平台上的OTV是否需要單獨的許可證？

A)否，只要ASR1000在配置了高級服務或創新引導級別的情況下運行，即可使用OTV。

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）。