使用ASR 1000配置重叠传输虚拟化

目录

简介

本文档介绍ASR1000和Catalyst 8300/8500系列路由器上支持的重叠传输虚拟化(OTV)网络拓扑。

先决条件

要求

本文档没有任何特定的要求。

使用的组件

本文档中的信息基于以下软件和硬件版本：

- ASR1000、IOS® XE 16.10.1a版及更高版本

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

本文档中的信息都是基于特定实验室环境中的设备编写的。本文档中使用的所有设备最初均采用原始（默认）配置。如果您的网络处于活动状态，请确保您了解所有命令的潜在影响。

背景信息

ASR1000支持Cisco IOS® XE 3.5版开始的OTV。Catalyst 8300系列路由器开始支持IOS® XE17.5.1a，Catalyst 8500系列路由开始支持IOS® XE 17.6.1a版本。

OTV通过传输网络中基于MAC地址的路由和IP封装转发(MAC-in-IP)提供远程网络站点之间的第2层连接，以支持需要第2层邻接的应用，例如集群和虚拟化。OTV使用重叠控制平面协议在整个重叠网络中学习和传播MAC路由信息。OTV控制平面协议使用中间系统到中间系统(IS-IS)消息建立到远程站点的邻接并将MAC路由更新发送到远程站点。OTV通过自动发现远程OTV设备建立与重叠网络上远程站点的第2层邻接。

第2层扩展的OTV的优势包括：

- 无MPLS要求
- 无需针对网状网络进行复杂的多协议标签交换(EoMPLS)配置
- 没有用于第2层扩展的复杂虚拟专用局域网服务(VPLS)部署
- 本地生成树隔离
    - 无需明确配置网桥数据协议单元(BPDU)过滤器
    - 默认隔离特定数据中心的生成树问题
- 本地未知单播泛洪隔离
    - 未知单播MAC数据包未转发
    - 允许支持每MAC未知单播转发
- 使用OTV ARP缓存优化地址解析协议(ARP)
    - 减少不必要的广域网流量
- 简化第一跳冗余协议(FHRP)隔离的调配
- 简化站点的添加
- 简化的冗余配置
- 当需要临时服务时，能够使用“设备加载”进行迁移

要求

设计OTV部署时，需要记住的主要规则是后续项目。 如果这些规则得到遵守，设计和部署就会得到简化。

- 对于所有已配置的OTV重叠接口，一个且只有一个接口可用于传输OTV封装流量，称为加入接口
- 一个且只有一个接口可用于为OTV站点VLAN配置数据中心第2层服务实例，并为所有已配置的OTV重叠接口配置数据中心间扩展的VLAN
- 端口通道可用于接口冗余和与VSS或VPC交换机的连接，并且支持作为“一个且仅一个”接口进行连接。
- 所有OTV路由器必须通过加入接口进行联系

- 必须在指向数据中心的OTV路由器上配置生成树
- 必须配置IGMP监听和查询才能正确转发数据中心组播流量
- 给定数据中心可以配置1台或2台OTV路由器。使用两台路由器时，它们会根据VLAN编号以奇数/偶数方式分发VLAN转发。 数据中心中的每个OTV路由器都充当其他路由器的备份。
- 多宿主对必须使用相同的OTV站点标识符进行配置
- ASR1000/Catalyst 8300/Catalyst 8500和Nexus 7000可以参与同一个OTV网络
    - Nexus 7000不支持OTV分段或加密，因此这些功能不能用于“混合”部署。

某些支持的背对背连接设计不符合所述规则。 虽然这些配置受支持，但不建议使用。 有关这些拓扑的详细信息，请参阅后面的“特殊情况下的单播拓扑”部分。

在为OTV配置加入接口和L2接入接口时，当前OTV软件存在“一个且仅一个”接口限制，这一点再强调也不为过。端口通道接口可用于冗余。 支持端口通道到VPC中的Nexus 7000的连接。 还支持与单个交换机的基本端口通道连接。

OTV实施类型

OTV需要一个加入接口和一个第2层接口。 每个OTV路由器只能支持其中一个。 OTV还要求配置站点VLAN，以便多宿主OTV路由器可以通过本地网络相互通信。 即使单宿主OTV路由器也必须配置OTV站点VLAN。 此外，每个站点或数据中心必须配置唯一的站点标识符。 双宿主OTV路由器必须使用相同的站点标识符，并且能够通过同一VLAN通信。

后续配置提供了OTV所需的基本配置，但由于必须添加单播或组播核心配置，因此该配置并不完整。 这些将在本文档的后续部分中详细说明。

```
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
```

服务实例配置用于所有带OTV的L2接口配置。

L2接口上的每个服务实例都必须与特定的单标记或双标记封装相关联。

反过来，这些服务实例都必须与网桥域相关联。

该网桥域用于重叠接口上配置的服务实例。

网桥域是将重叠服务实例链接到L2接口服务实例的粘合剂。

重叠接口上的流量封装必须与L2接口上重写入口后的流量封装匹配。

在本例中，进入Gig1/0/1服务实例99的流量有一个99的802.1Q VLAN和网桥域99。重叠接口上网桥域99的对应服务实例也配置为一个99的802.1Q VLAN。这种情况最简单。

在本例中，进入Gig1/0/1服务实例98的流量具有两个802.1Q VLAN（99和1098）和网桥域90。重叠接口上网桥域90的对应服务实例被配置为一个90的802.1Q VLAN。很明显，这些不同。 rewrite ingress命令可确保标记在流量通过入口接口时正确转换。 进入L2接口的流量会将98/1098 802.1Q VLAN替换为单个90的VLAN。symmetric关键字可确保从L2接口流出的流量会将90的单个802.1Q VLAN替换为98/1098。

具有多个OTV扩展的802.1Q VLAN的任何服务实例都必须使用rewrite ingress命令。 OTV封装仅支持一个VLAN标识符。 因此，必须将L2接口上的任何双VLAN配置重写为Overlay接口服务实例上的单个标记。 这排除了对不明确VLAN配置的支持。

有关标记重写的详细信息，请参阅本文档：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站点桥接域为100。

- OTV站点网桥域仅在L2接口上配置。
- 永远不能在重叠接口上配置OTV站点网桥域，因为这会导致OTV部署不稳定。
- OTV站点VLAN必须仅连接到OTV路由器，不能传输任何其他数据中心/用户流量。
- OTV站点VLAN必须与OTV扩展VLAN位于同一物理接口上。

Multihome

数据中心可与一台OTV主机或最多2台主机连接，以实现冗余，也称为Multihome。 Multihome用于实现恢复能力和负载均衡。当一个站点中存在多个边缘设备并且它们都参与同一个重叠网络时，该站点被视为多宿主。OTV Multihome根据VLAN编号以奇数/偶数方式在属于同一站点的两个OTV路由器之间拆分VLAN。一个边缘设备被选举为所有奇数VLAN的AED，而另一个OTV路由器被选举为所有偶数VLAN的AED。每个AED也是另一台路由器上处于活动状态的VLAN的备用设备。如果其中一个AED中的链路或节点发生故障，备用AED对所有VLAN变为活动状态。

如果两个ASR1000连接到同一数据中心以进行Multihome，则两个ASR1000之间无需专用链路。

OTV使用通过内部接口传播的OTV站点VLAN和通过加入接口进行通信来确定哪些路由器负责偶数VLAN和奇数VLAN。

不能将ASR1000和Nexus 7000混用在同一数据中心，并且不能将两台路由器上配置的OTV作为另一台路由器的备份。匹配平台（ASR1000或Nexus 7000）支持特定数据中心的多宿主系统。您可以在一个数据中心中使用ASR1000，在另一个数据中心中使用Nexus 7000。这两个平台之间的互操作性已经过测试并受到支持。 有些数据中心可以是多宿主，而其他数据中心是单宿主。

多宿主ASR1000路由器对必须运行相同版本的Cisco IOS® XE软件。

如果使用Multihome，则强烈建议必须在OTV路由器上启用生成树，因为这样会使OTV路由器发出拓扑更改通知(TCN)，导致相邻的L2交换机设备（以及生成树中的其他交换机）将其老化时间计时器从默认值减少到15秒。 当多宿主对之间存在故障或恢复时，这可以大大提高收敛速度。通过将后续行添加到全局配置中，可以为所有已配置的服务实例（连接到OTV或其他）启用生成树。

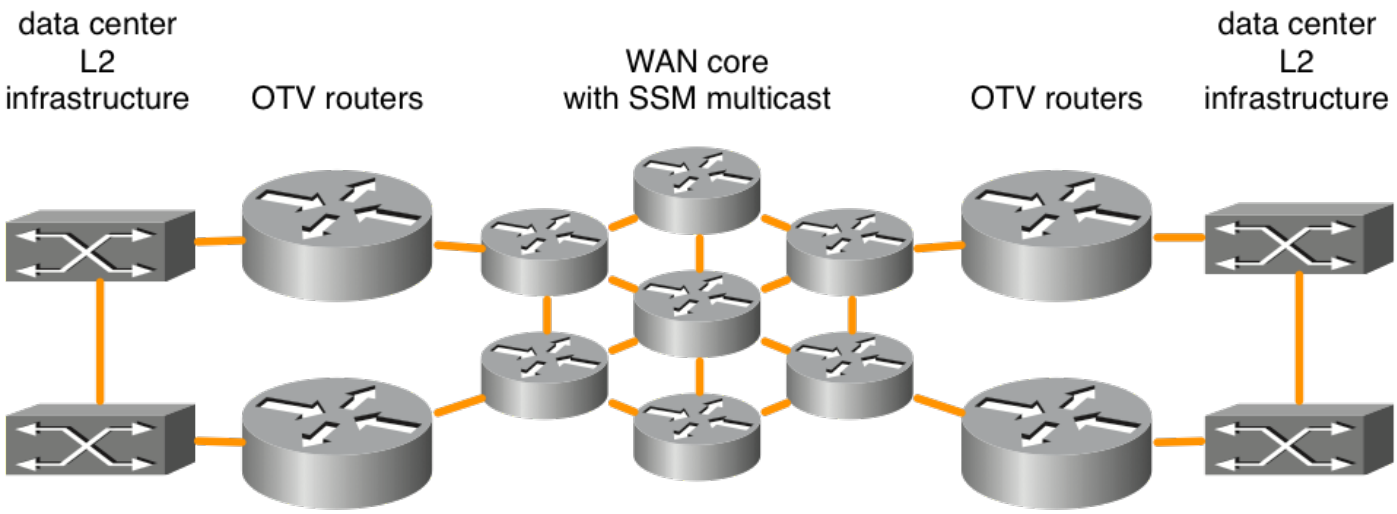
```
spanning-tree mode [ pvst | rapid-pvst | mst ]
```

不需要针对每个vlan或每个服务实例进行特定的配置。

组播核心

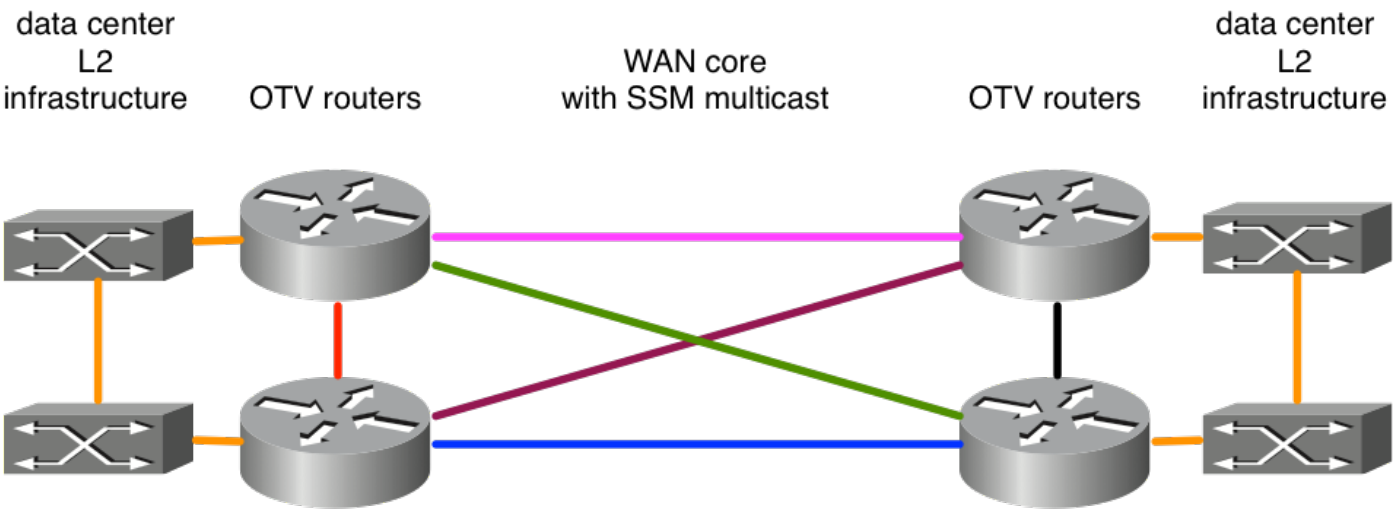
组播网络需要广域网中的全网状连接。所有OTV路由器都需要通过加入接口连接在一起。

图 1.支持的组播网络拓扑



此图显示了通过全网状核心连接的两个数据中心的示例。 源特定组播(SSM)协议独立组播(PIM)在OTV路由器和WAN核心路由器之间运行。 只要具有全网状连接，就可以支持任意数量的核心路由器。 对于跨广域网核心的OTV连接，没有明确的最大延迟要求。

图 2.不支持的组播网络拓扑



由于ASR1000/OTV期望在单个加入接口上接收来自所有其对等体的组播消息（根据示例），这将导致OTV部署不稳定。假设以粉红色和蓝色显示的东-西链路配置为加入接口。 当粉红色链路发生故障时，路由器将无法再在该接口上接收OTV更新。 通过绿色或紫色链路的备用路径是不可接受的，因为已明确配置加入接口。 该接口必须接收更新。 目前不支持使用环回接口作为加入接口。

如果用户没有自己的主干，他们必须确保他们的服务提供商支持其核心的组播，并且服务提供商可以响应互联网组管理协议(IGMP)查询消息。ASR1000上的OTV充当组播主机（转发IGMP加入消息），而不是核心广域网组播拓扑的组播路由器。

OTV路由器之间的传输网络必须支持提供商组播组的PIM稀疏模式（任意源组播[ASM]）和交付组的SSM。

组播核心需要在重叠接口上为控制组以及一系列用于转发数据的数据组播组进行某些特定配置。

```
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
```

组播OTV部署要求将加入接口配置为PIM被动接口。 可以根据需要为不同版本配置IGMP。 重叠接口必须配置控制组和数据组。 控制组是用于OTV管理的单个组播组。 数据组是用于在数据中心之间传输用户数据的一系列组播地址。 如果数据组不在232.0.0.0/8 IP空间中，则必须将其他命令“ip pim ssm range”配置为包含OTV所需的范围。

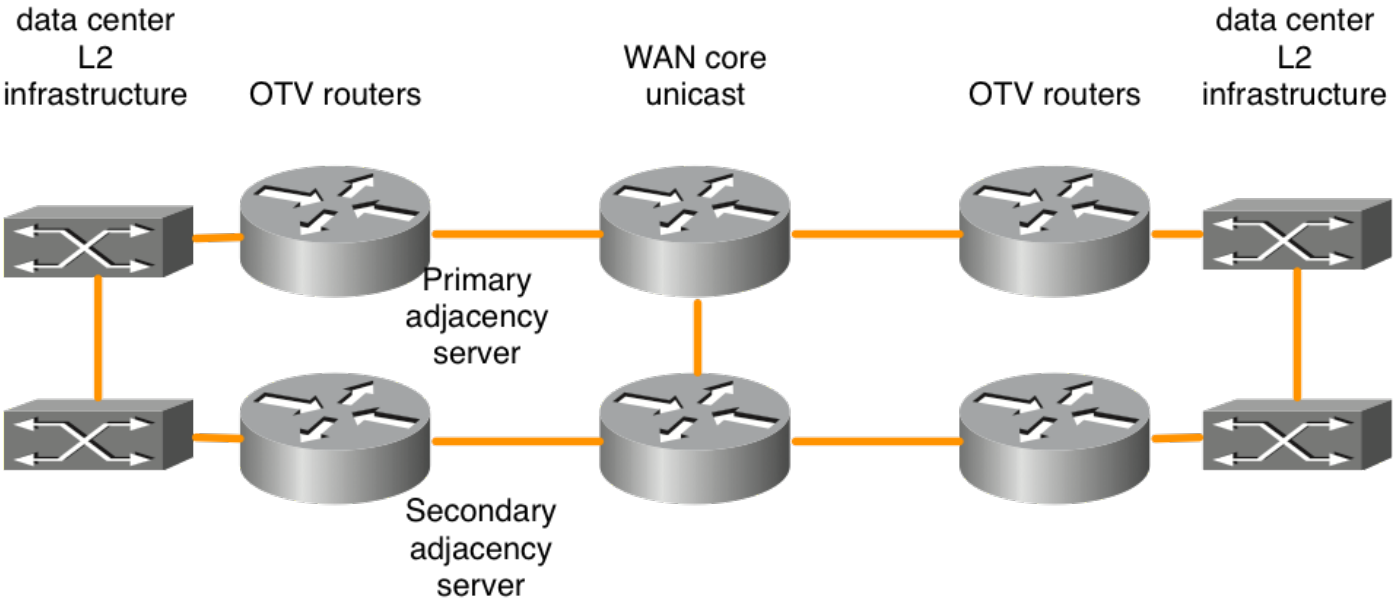
OTV路由器之间的传输网络必须支持提供商组播组的PIM稀疏模式（任意源组播[ASM]）和交付组的

源特定组播(SSM)。

具有邻接服务器的单播核心

Cisco IOS® XE 3.9增加了支持单播核心的OTV。 所有ASR1000平台和来自Cisco IOS® XE 3.9的未来版本继续支持OTV的单播和组播核心。

图 3.单播网络拓扑



OTV邻接服务器功能启用OTV边缘之间的仅单播传输。 配置了邻接服务器角色的OTV路由器会保留所有已知OTV路由器的列表。它们向所有注册的OTV路由器提供该列表，以便它们具有必须接收复制的广播和组播流量的设备的列表。

仅单播传输上的OTV控制平面的工作方式与具有组播核心的OTV完全相同，不同之处在于，在单播核心网络中，每个OTV边缘设备需要创建每个控制平面数据包的多个副本，并将它们单播到同一逻辑重叠中的每个远程边缘设备。

按照同样的思路，来自数据中心的任何组播流量都复制到本地OTV路由器上，并向每个远程数据中心发送多个副本。 虽然这比依赖广域网核心来执行复制要低效，但不需要配置和管理核心组播网络。 如果数据中心组播流量很小或数据中心位置很少（四个或更少），则单播核心通常是OTV转发的最佳选择。 总体而言，仅单播模型的操作简化使单播核心部署选项在仅需要4个或更少数据中心之间建立LAN扩展连接的情况下更适合。 建议至少配置两台邻接服务器，一台主服务器和一台备用服务器。 没有用于主用/主用邻接服务器配置的选项。

必须相应地配置OTV路由器，才能正确识别并注册到相应的邻接服务器。

	主邻接服务器	辅助邻接服务器	其他OTV路由器
OTV加入接口IP地址	10.0.0.1	10.2.2.24	其他IP地址

	主邻接服务器	辅助邻接服务器	其他OTV路由器

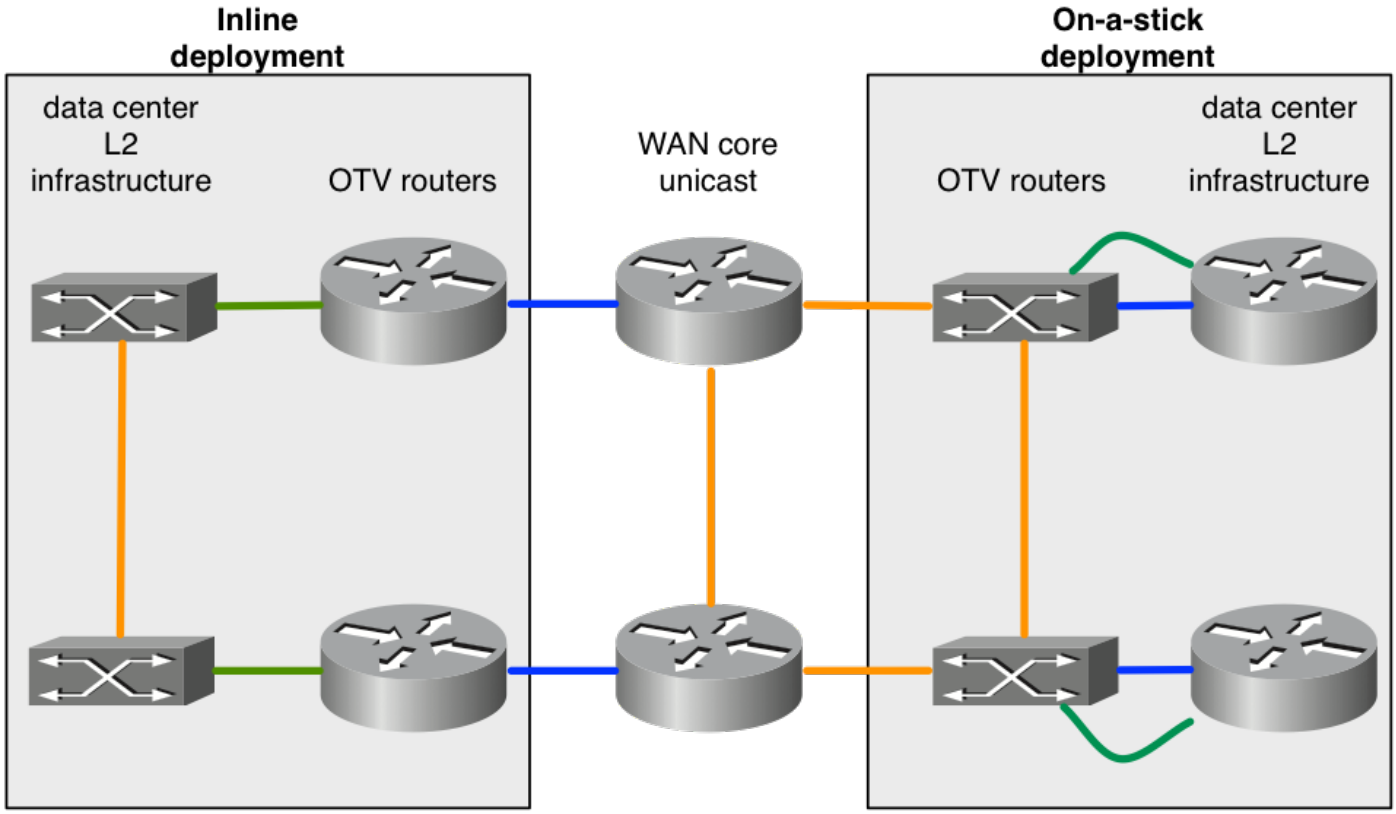
配置	interface Overlay 1
	  otv adjacency-server unicast-only
		interface Overlay 1
		  otv adjacency-server unicast-only
		  otv use-adjacency-server 10.0.0.1 unicast-only
			interface Overlay 1
			  otv use-adjacency-server 10.0.0.1 10.2.2.24 unicast-only

某些背对背连接设计支持单播OTV转发，但不遵循“全网状”规则。 虽然支持这些配置，但不推荐使用这些配置。 当数据中心通过暗光纤连接时，此类部署最为常见。 有关此配置选项的详细信息，请参阅后面的“特殊情况下的单播拓扑”部分。

单臂OTV与内联

在您的数据中心部署OTV有两种模式：单臂模式和嵌入式模式。在前面介绍的设计方案中，OTV路由器在数据中心与服务提供商核心网络之间内联。但是，最好添加OTV路由器作为不在所有流量传输路径中的设备。有时要求不更改当前拓扑以通过当前设备连接到服务提供商（例如，使用Catalyst 6000交换机或Nexus交换机硬件的灰场部署不支持OTV）。因此，最好在ASR1000上将OTV作为单臂设备部署为OTV设备。

图 4.内联拓扑与单臂拓扑



该图演示了可属于同一重叠的两个部署模型。 连接到OTV路由器的绿色链路配置为L2接入接口以接受VLAN流量。 连接到OTV路由器的蓝色链路是承载OTV封装VLAN流量的加入接口。

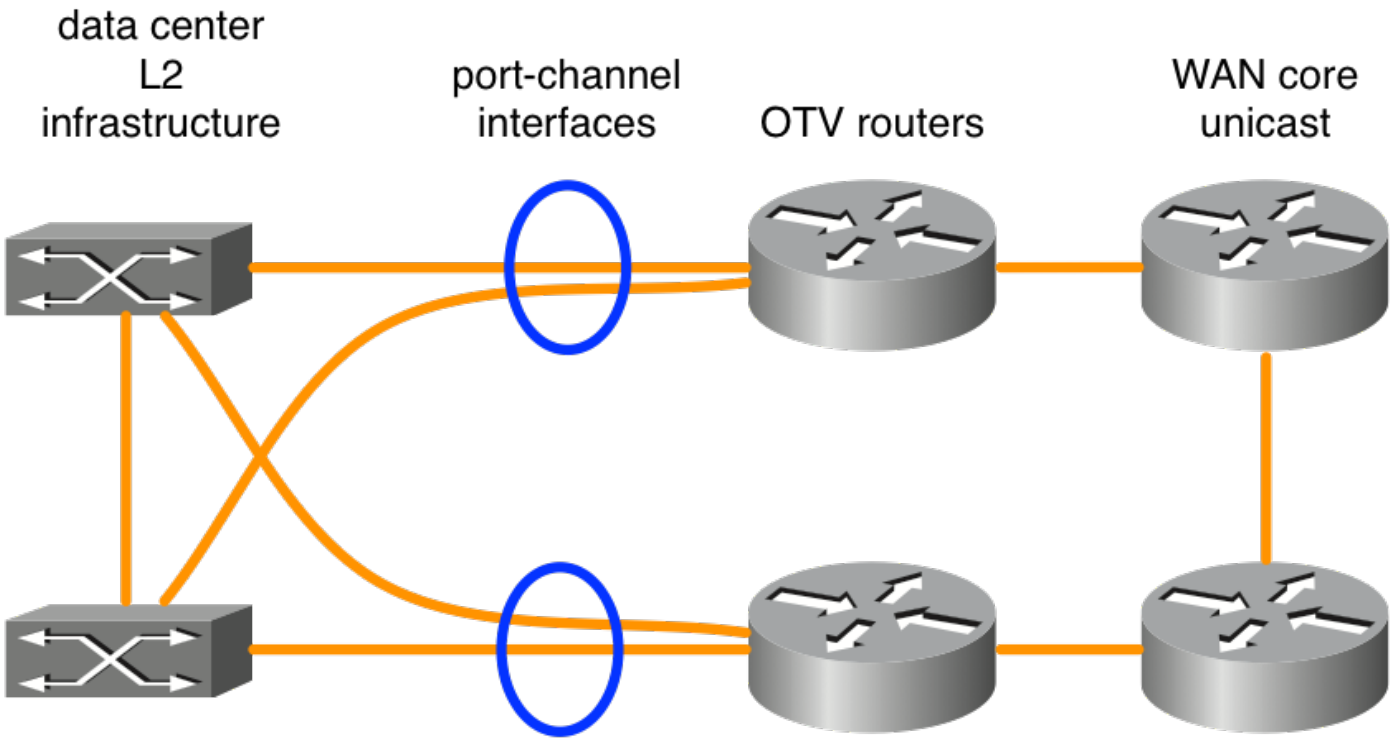
可能需要配置OTV不支持的功能。例如，OTV和MPLS不能配置在同一台设备上。因此，可以选择

单臂使用ASR1000/OTV，并在OTV路由器前面的路由器上配置MPLS。

第2层和第3层的端口通道

用于ASR1000的Cisco IOS® XE 3.10代码增加了对OTV的第2层和第3层端口通道配置的支持。第2层端口通道可用作内部接口。 Port-channel必须包含最多4个物理接口。第3层端口通道可用作加入接口。

图 5.用于L2连接的端口通道



该图显示了一个典型的端口通道方案，包括VSS（Catalyst 6000系列）或VPC（Nexus 7000系列）中的两台交换机。 这种设计通过双OTV路由器和到数据中心基础设施的双连接提供冗余。 如果在与OTV路由器相邻的第2层交换设备上使用VSS或VPC，则无需对OTV进行特殊配置，除非进行基本端口通道配置。

默认网关

根据定义，OTV会在多个位置创建相同的L3子网。 在将L3流量路由到扩展VLAN或从扩展VLAN路由时，需要考虑一些特殊注意事项。 L3路由可以在OTV路由器上配置，也可以在连接到扩展VLAN的其他设备上配置。 此外，在每个场景中，可以部署第一跳冗余协议(FHRP)，例如热备份冗余协议(HSRP)或虚拟路由器冗余协议(VRRP)，以实现冗余。 HSRP可以在本地运行到指定数据中心，也可以在数据中心之间扩展（不典型）。

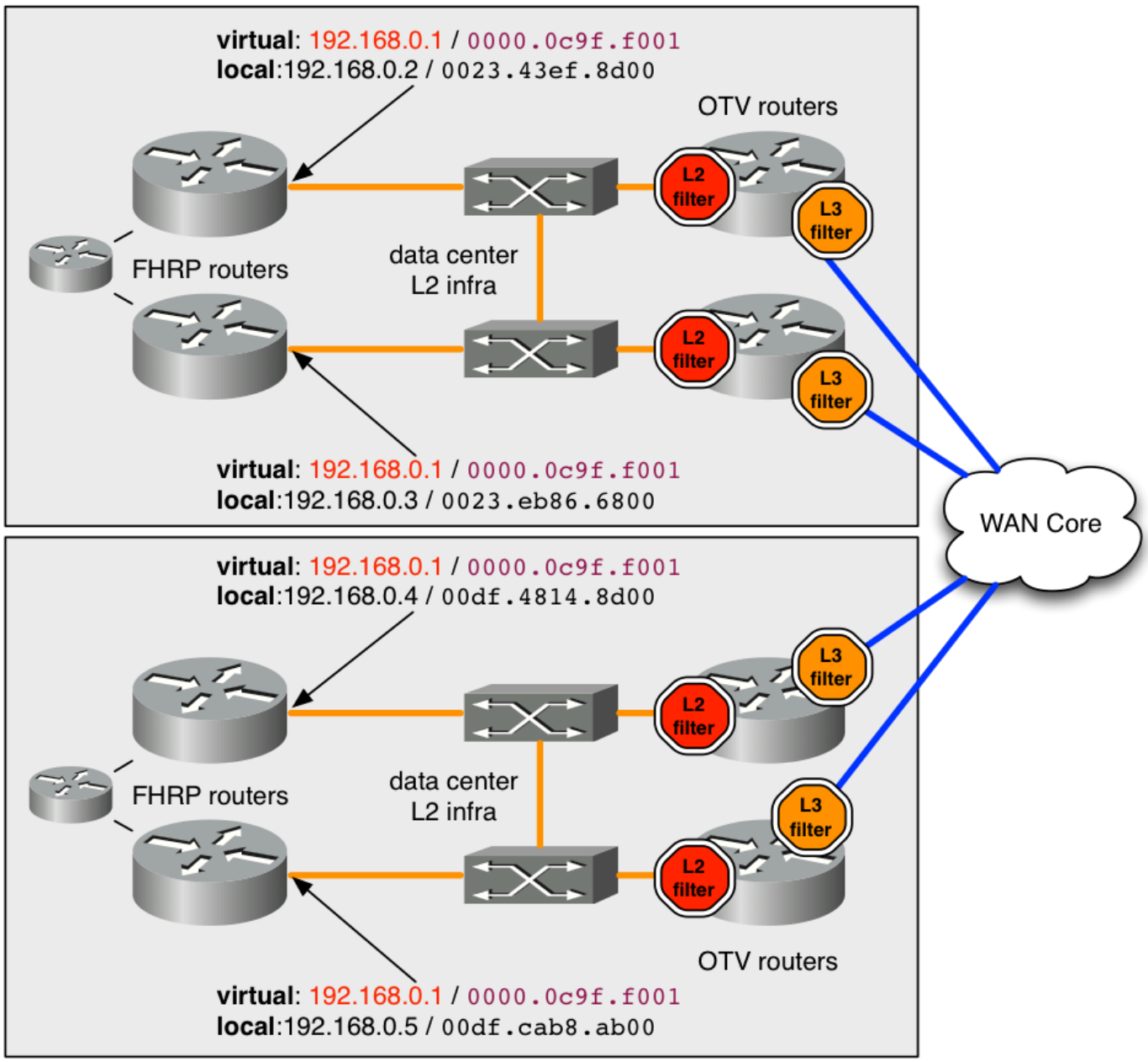
利用FHRP的OTV部署的最佳实践是在每个数据中心运行FHRP的本地实例。 这些FHRP实例使用相同的虚拟MAC地址和IP地址，因此当虚拟机(VM)在数据中心之间移动时，它们具有不间断连接。 如果在数据中心之间更改默认路由器的MAC地址，则虚拟机将无法与子网外通信，直到虚拟机的默认网关ARP条目超时。

要正确部署带OTV的FHRP，必须考虑必须从OTV过滤和隔离哪些L2和L3流量。在L2级别，这是必

要的，以防止OTV在多个位置看到FHRP使用的相同L2虚拟MAC。 在第3层级别需要过滤器，以将HSRP和VRRP通告隔离到每个数据中心，从而使主用/侦听/备用选择本地化到每个数据中心。

默认情况下，启用OTV时启用FHRP过滤器。 如果设计要求在数据中心之间扩展FHRP，则可以禁用该功能。 默认情况下，未启用虚拟MAC地址的L2过滤，必须手动配置。

图 6. 建议的FHRP部署示例



在本例中，虚拟MAC地址0000.0c9f.f001用于IP地址192.168.0.1，该地址在扩展VLAN上托管，用于与子网建立连接。 在两台数据中心使用相同的虚拟MAC和IP地址时，主机在数据中心之间传输时可以无缝连接子网。

为了使MAC地址0000.0c9f.f001在多个位置对OTV隐藏，必须在为VLAN提供服务的每台OTV路由器上为VLAN部署入口L2过滤器（图中的红色标记）。通过ACL过滤在入口的L2服务实例上配置的过滤器ACL，来自该MAC的所有数据包都会被丢弃，然后ASR1000上的OTV进程才能看到它们。 因此，OTV从不了解MAC，也不会将其通告到远程数据中心。

此处提供了捕获所有已知/默认FHRP虚拟MAC流量的建议配置。

```
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匹配与HSRP版本1和2、网关负载均衡协议(GLBP)和VRRP关联的已知MAC地址空间（按此顺序）。如果虚拟MAC配置为使用不基于FHRP组号的非标准值，则必须明确地将其添加到ACL示例。 必须将ACL添加到L2服务实例（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，还必须管理L3级别的FHRP主机之间的通信。 在图中的单个扩展子网上配置了四个FHRP路由器。 如果没有一定程度的L3过滤器，所有四台路由器将会看到对方，会选举出一个活动设备，并且有3台处于各种备用状态。 因此，一个数据中心将有两个本地备用FHRP路由器，但由于前面讨论的L2过滤器原因，没有到远程活动路由器的L2连接。

理想的结果是在每个数据中心拥有一个主用和一个备用FHRP路由器。 前面讨论的入口L2过滤器不会捕获此选举流量，因为选举过程使用路由器的实际IP和MAC地址。 默认情况下，后续ACL在重叠接口上作为出口应用。 重叠接口的出口将是流向广域网核心的流量。 ACL不会显示在运行配置中，但可以使用show ip access-list命令观察它。 它根据UDP端口号过滤FHRP选举流量。

```
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
```

禁用此过滤器的唯一原因是您希望一个VLAN上的所有FHRP路由器参与相同活动状态选举。 要禁用此过滤器，请在重叠接口上配置“no otv filter-fhrp”。

未知单播流量

默认情况下，OTV路由器从LAN接收的发往远程OTV位置未知的MAC地址的单播流量会被丢弃。 此流量称为未知单播。 此丢弃操作针对限制广播流量在WAN上消耗的带宽量的广域网核心。 一般预期是，LAN中的所有主机都会发出足够的广播流量（ARP、协议广播等），这些流量始终会被OTV路由器看到、通告并因此“已知”。

某些应用会利用静默主机。 在正常的交换基础设施中，这不是问题，因为LAN上未知单播MAC地址的L2广播允许静默主机查看流量。 但是，在OTV环境中，OTV路由器会阻止数据中心之间的流量。

为了弥补这一不足，Cisco IOS® XE集成了称为选择性单播转发的功能。XE 3.10.6、XE3.13.3、XE 3.14.1、XE3.15以及之后的所有版本都支持选择性单播转发。

可通过在重叠接口上为每个MAC地址添加一个命令来配置它。 例如：

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

在本例中，所有发往0000.0001.0001的流量都必须泛洪到具有VLAN 100的所有远程OTV路由器。 这可以通过后续命令观察到：

```
<#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
```

如果在远程站点获知该MAC地址，则必须向转发表中添加一个条目，该条目优先于泛洪条目。

```
<#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
```

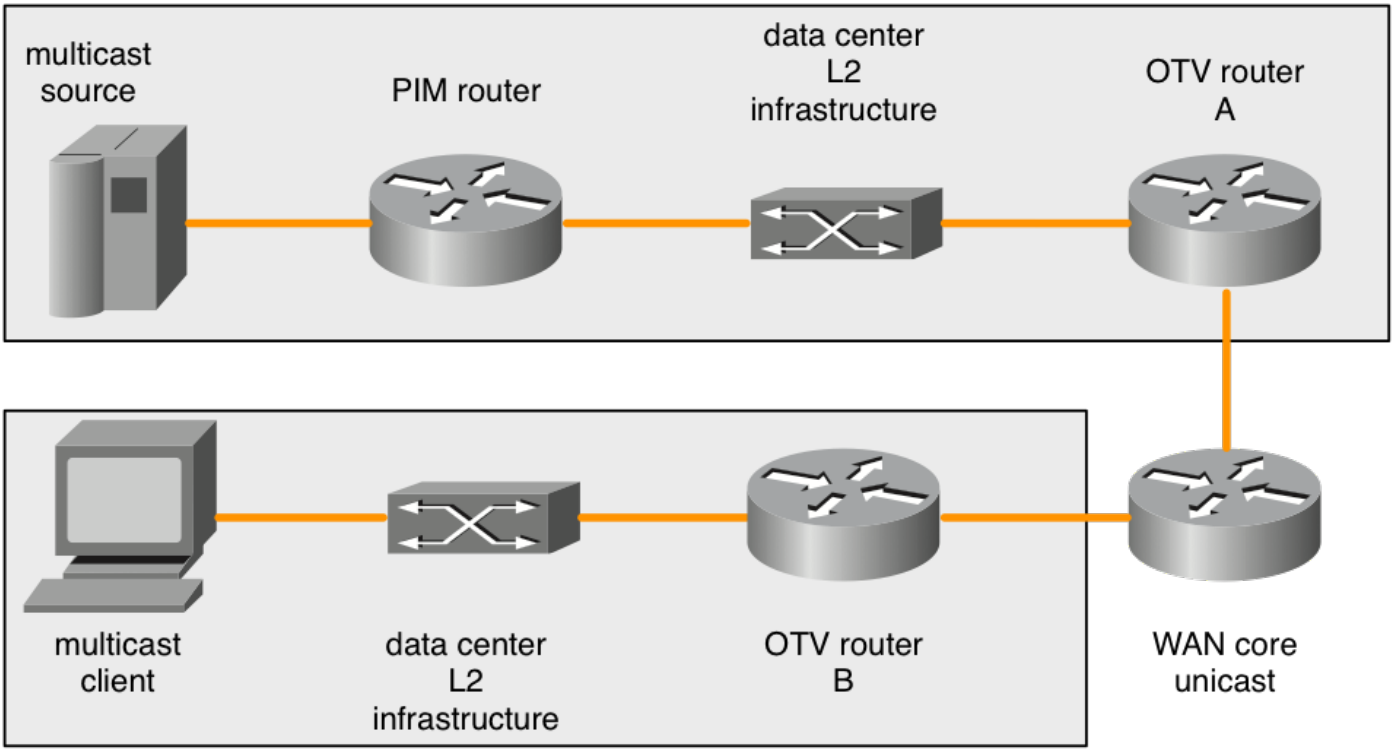
```
0    100  100    0000.0000.0001 50    ISIS   OTV_router_3
```

通常，必须在具有该VLAN的所有OTV路由器上配置给定MAC地址的泛洪条目。

远程组播源

ASR1000认为OTV路由器不会转发从LAN收到的组播IGMP加入请求。后续图表详细说明了可能出现问题的拓扑。

图 7.远程组播源



当组播客户端发送组播IGMP加入时，ASR1000（OTV路由器B）会观察并通告对组播组的兴趣。远程OTV路由器（OTV路由器A）必须将任何流量转发到他们在本地L2广播域中看到的该组播组。但是，当从客户端的OTV路由器（OTV路由器B）向广播对组播组的兴趣时，远程ASR1000（OTV路由器A）不会重新生成组播IGMP加入请求。

当组播源与OTV路由器位于同一L2广播域时，不存在此问题。 必须将OTV路由器配置为IGMP查询器。 这会显示在L2广播域中存在的任何组播流量中。 但是，只有PIM加入请求会导致PIM路由器将来自其他L2广播域的组播源转发到OTV路由器所在的L2广播域。

不会转发或重新生成远程IGMP加入请求。OTV路由器也不是PIM路由器。 因此，当远程客户端感兴趣时，组播源不直接位于带OTV路由器的L2广播域上的拓扑无法从PIM路由器进入以转发源流量。

此问题有两个解决方法。

首先，本地IGMP客户端可以部署在连接到OTV路由器（OTV路由器A）的L2广播域上。该IGMP客户端必须订阅远程客户端可以订阅的任何组播组。 这将导致PIM路由器将组播流量转发到与OTV路

由器A相邻的广播域。 然后，IGMP查询会引入任何组播流量，并通过重叠网络发送。

另一种解决方案是为远程客户端可以订阅的任何组配置“ip igmp static-join”。 这也会导致PIM路由器将组播流量转发到与OTV路由器A相邻的广播域。

此限制是已知的，是设计规范的一部分。 目前不将其视为Bug，而是受支持的拓扑中的限制。

QoS 注意事项

默认情况下，在ASR1000上，添加的OTV报头中的TOS值从L2数据包的802.1p位复制。 如果L2数据包未标记，则使用零值。

Nexus 7000在5.2.1及更高版本的软件中具有不同的默认行为。 如果期望的行为是将内部数据包TOS值复制到外部，则其他QoS配置可实现此目的。 这提供了与较新的Nexus 7000软件相同的行为。

将L2数据包L3 TOS值复制到OTV数据包最外层报头中的配置是后续配置：

```
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
```

提供的配置必须与入口上各种DSCP值的流量匹配。 本地有效的qos组标记用于在通过路由器传输期间在内部标记该流量。 在出口接口，匹配qos组，然后相应地更新最外层的TOS字节。

WAN MTU注意事项/分段

OTV基本上使用GRE报头通过WAN传输L2流量。此GRE报头的大小为42字节。 在理想的网络部署中，WAN链路的最大传输单元(MTU)必须至少比OTV预期处理的最大数据包大42字节。

如果L2接口的MTU为1500字节，则加入接口的MTU必须为1542字节或更多。 如果L2接口的MTU为2000字节，但预计只处理最大为1500字节的数据包，则1542字节的WAN MTU便已足够，但将42字节的标准添加到2000中将是理想的。

```
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
```

一些服务提供商无法为其广域网电路提供更大的MTU值。 如果出现这种情况，ASR1000可以对OTV传输的数据执行分段。 Nexus 7000不具备此功能。 不支持在ASR1000上启用了分段的ASR1000和Nexus 7000 OTV混合网络。

OTV分段的配置为：

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

请务必在Overlay interface join-interface命令之前配置全局级命令。 如果首先配置了Overlay接口的otv join-interface命令，请从Overlay接口删除otv join-interface命令，配置otv fragmentation join-interface命令，然后再次配置Overlay接口的otv join-interface命令。

当OTV分段未启用时，所有承载封装的L2数据的OTV数据包都使用DF位集发送，这样它们在传输时就不会分段。 添加了fragmentation命令后，DF位设置为0。OTV路由器本身可以对数据包进行分段，并且它可在由其他路由器传输时进行分段。

ASR1000平台上可用的数据包重组缓冲区数量有限，因此数据包为了传输而必须截断的片段越少越好。 如果存在问题，这将提高整个广域网的效率并降低整体带宽消耗。 启用OTV分段会对性能产生影响。 如果存在分段，并且预期要处理1Gb/s以上的OTV流量，则必须进一步调查OTV性能。

特殊情况下的单播拓扑

OTV的现场部署通常在两个数据中心的OTV路由器之间采用直接背靠背光纤连接。

对于单宿主拓扑，这使OTV和非OTV流量共享加入接口的标准部署成为可能。 此设置不需要特殊注意事项，因此本部分不适用。

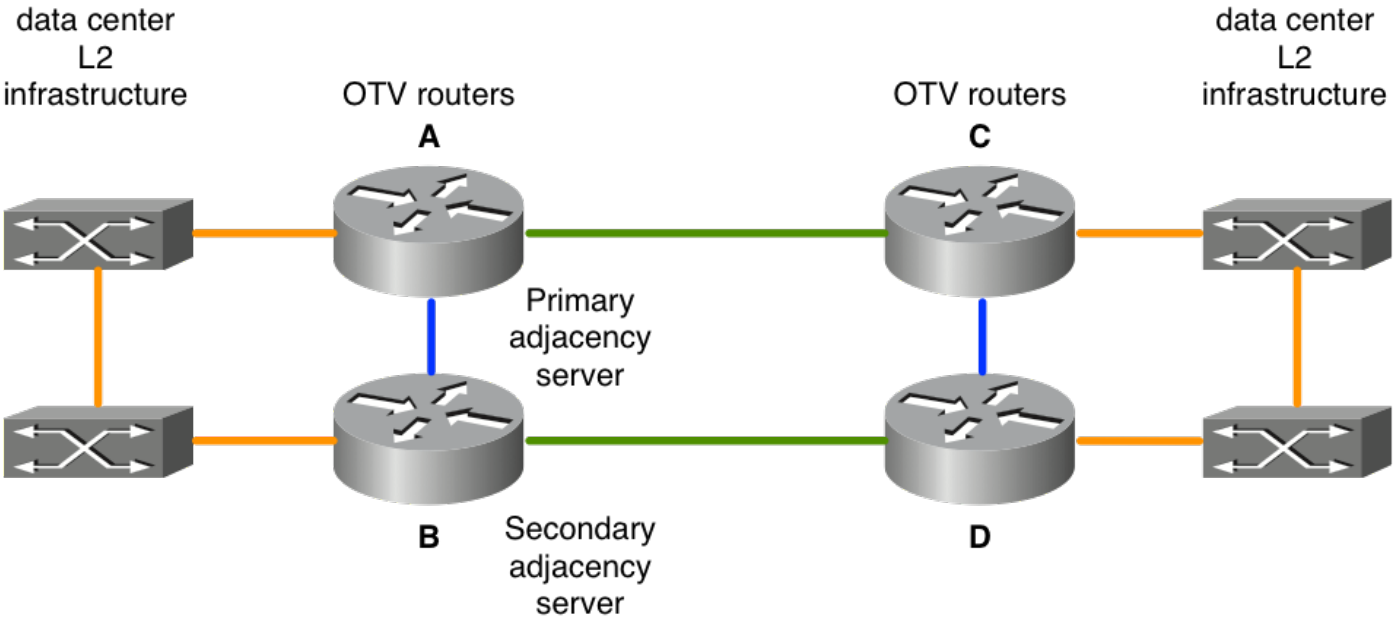
但是，如果部署在两个数据中心中有多宿主OTV路由器，则会有一些特殊的注意事项。 需要进行其他配置。

如果涉及两个以上的数据中心，则此特殊配置不适用。

对于使用单宿主或多宿主OTV路由器的两个以上数据中心的场景，必须使用标准单播或组播OTV部署。

没有其他受支持的替代方案。

图 8. 特殊案例单播



在显示的拓扑中，绿色链路是两个数据中心之间的暗光纤链路。 这些暗光纤直接连接到OTV路由器。 OTV路由器之间的蓝色链路用于在绿色链路出现故障时重新路由非OTV流量。 如果顶部绿色链路发生故障（A到C），则使用最顶部的OTV路由器作为其默认路由的非OTV流量将通过南北蓝色链路（A到B和C到D）路由到底部OTV路由器对（B到D）之间仍然运行的绿色链路。

这种基本的流量重新路由对OTV流量不起作用，因为OTV配置将物理接口指定为加入接口。 如果OTV路由器A上的“绿色接口”断开，则OTV流量不能从备用接口OTV路由器B发出。此外，由于没有通过WAN核心的完全连接，因此当发生故障时，无法通知所有OTV路由器。 为了解决此问题，需要使用双向转发检测(BFD)以及嵌入式事件管理器(EEM)脚本。

BFD必须监控东-西OTV路由器对（A/C和B/D）之间的WAN链路。如果与远程路由器的连接丢失，则OTV重叠接口将通过该东-西OTV路由器对上的EEM脚本关闭。 这会导致成对的多宿主路由器承担所有VLAN的转发工作。 当BFD检测到链路已恢复时，EEM脚本会触发以重新启用重叠接口。

使用BFD检测链路故障非常重要。 这是因为重叠接口的“故障”端及其东-西对都需要关闭。 根据服务提供商提供的连接类型，一个物理链路可以断开（OTV路由器A上的绿色接口），而对应的东-西对路由器的接口可以保持运行（OTV路由器C上的绿色接口）。BFD检测到传输中任一接口出现故障或任何其他问题，并立即同时通知两对。 当需要将恢复链路通知路由器时，也同样如此。

此部署的配置与任何其他部署相同，并添加了后续项目：

- WAN接口上的BFD配置
- 后续EEM脚本
- 匹配偶数/奇数VLAN分配的OTV ISIS身份

OTV加入接口上的BFD配置不在本文档的讨论范围之内。 有关如何在ASR1000上配置BFD的信息，请访问：

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

一旦连接接口对（图中的绿色链接）之间的BFD故障检测正常运行，则必须部署EEM脚本。 EEM脚本必须针对特定路由器进行定制，以修改正确的重叠接口，并可能监控日志中更精确的字符串，以发现BFD故障和恢复。

```
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 ..."
!
```

此类部署还要求东-西路由器对（A/C和B/D）在其奇偶校验vlan的转发中匹配。

例如，A和C必须转发偶数VLAN，而B和D在稳定的标称操作中转发奇数VLAN。

奇/偶分布由OTV序号确定，可通过“show otv site”命令观察该序号。

两个站点路由器之间的序号基于OTV ISIS网络ID确定。

```
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
```

必须在所有OTV路由器上配置OTV ISIS网络标识符。 配置标识符时必须小心，以使所有OTV路由器仍然可以相互识别。

<#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

.

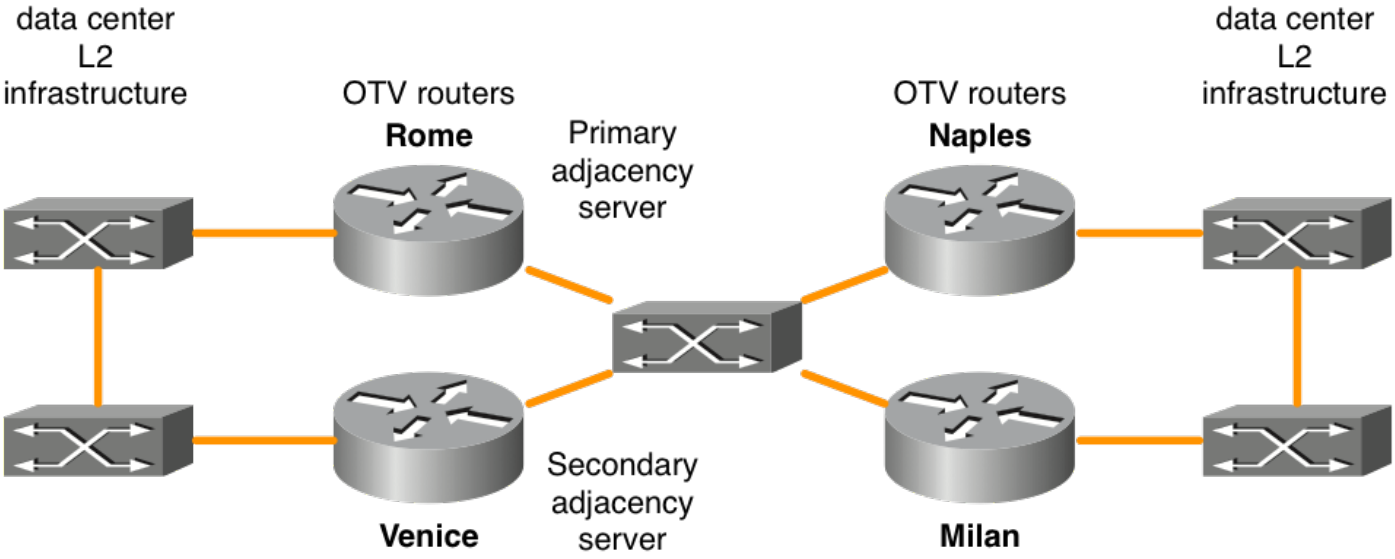
00

在参与重叠的所有OTV路由器上，标识符的黑色部分必须匹配。 可以修改标识符的红色部分。 站点上的最低网络标识符将获得序号0，然后转发偶数VLAN。 站点的最高网络标识符将获得序号1并转发奇数VLAN。

配置示例

单播

图 9单播配置示例



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
 !
```

Venice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
```

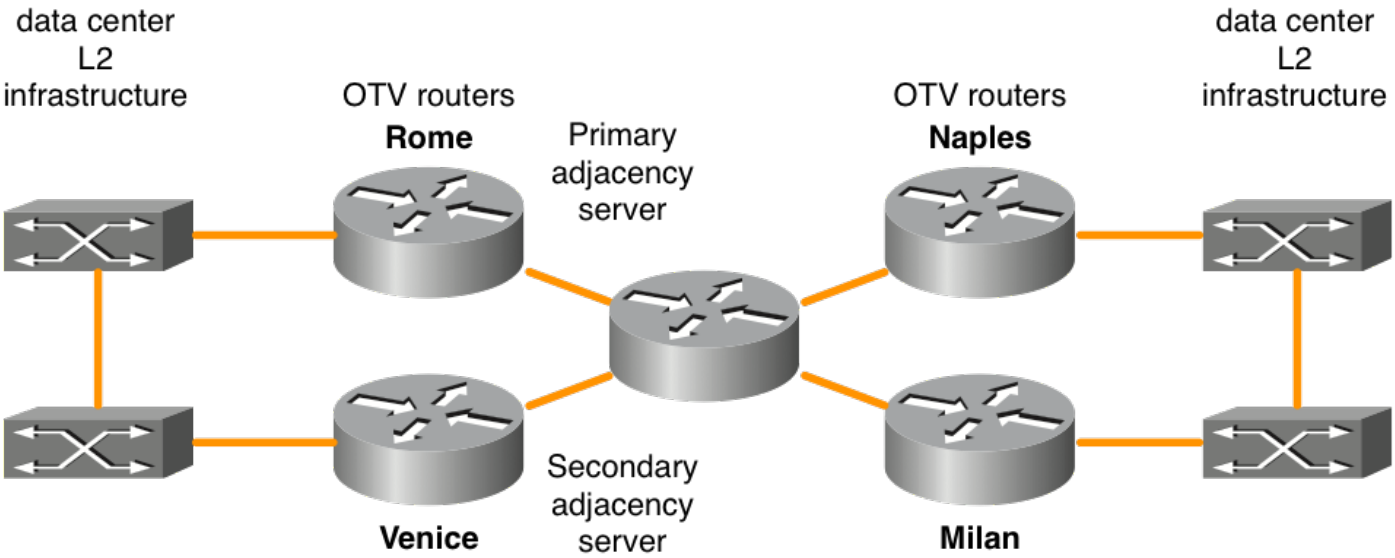
```
  bridge-domain 100
 !
 service instance 101 ethernet
  encapsulation dot1q 101
  bridge-domain 101
 !
!
```

米兰配置：

```
!
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
 !
!
```

组播

图 10.组播配置示例



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
 !
```

Venice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
 !
!
```

米兰配置：

```
!
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
 !
!
```

常见问题解答

问：专用VLAN是否与OTV一起受支持？

A)是，OTV中不需要特殊配置。在专用VLAN配置中，请确保以混合模式配置连接OTV L2接口的交换机端口。

问：IPSEC加密是否支持OTV？

A)是，支持加入接口上的加密映射配置。 OTV无需特殊配置即可支持加密。 但是，加密配置会增加额外开销，并且必须通过WAN MTU比LAN MTU的增加来补偿此开销。如果不能，则必须进行OTV分段。 OTV性能限制为IPSEC硬件的性能。

问：MACSEC是否支持OTV？

答：是的，ASR1001-X包括对内置接口的MACSEC支持。 OTV与LAN和/或WAN接口上配置的MACSEC配合使用。 OTV性能限于MACSEC硬件的性能。

问：环回接口能否用作加入接口？

答：不能，只能将以太网、Portchannel或POS接口用作OTV加入接口。 OTV环回加入接口已在规划图中，但当前未计划在此时发布。

问：是否可以将隧道接口用作加入接口？

答：不支持，GRE隧道、DMVPN隧道或任何其他类型的隧道不支持作为加入接口。 只有以太网、Portchannel或POS接口可用作OTV加入接口。

问：不同的重叠接口能否使用不同的L2和/或加入接口？

A)所有重叠接口必须指向同一个加入接口。 所有重叠必须链接到同一物理接口，以实现到数据中心的L2连接。

问：OTV站点VLAN能否与OTV扩展VLAN位于不同的物理接口上？

A) OTV站点VLAN和扩展VLAN必须在同一物理接口上。

问：OTV需要什么功能集？

A) OTV需要高级IP服务(AIS)或高级企业服务(AES)。

问：固定配置平台上的OTV是否需要单独的许可证？

答：不能。只要ASR1000的运行配置了高级服务或企业引导级别，OTV就可用。

关于此翻译
思科采用人工翻译与机器翻译相结合的方式将此文档翻译成不同语言，希望全球的用户都能通过各自的语言得到支持性的内容。

请注意：即使是最好的机器翻译，其准确度也不及专业翻译人员的水平。

Cisco Systems, Inc. 对于翻译的准确性不承担任何责任，并建议您总是参考英文原始文档（已提供链接）。