Cisco Nexus 9000 数据中心运营商

指南

2015 年 3 月

目录

1. 简介

1.1 概述

Cisco Nexus® 9000 系列交换机产品使任何规模的客户都能享受到下一代数据中心交换功能的优势。本白皮书主要面向从未使用过 Cisco® Nexus 9000，但希望了解如何将 Cisco Nexus 9000 轻松部署在其数据中心的商业客户。

本白皮书重点介绍了 Cisco Nexus 9000 的优势，概述了适合中小型客户部署的几种设计，讨论了与现有网络的集成，并介绍了经过思科验证的 Nexus 9000 拓扑，同时也列举了一些配置示例。

这些极具特色的设计既适用于初次使用 Cisco Nexus 9000 交换机的组织，也适用于正在横向扩展数据中心的现有客户组织或成长型组织。上述配置示例将逻辑设计转变成易于使用的有形模板，对于 IT 员工人数不多或日益增加的组织来说，这些模板可简化部署和运营。

（可选）读者可以参阅本文档末尾的附录，从初学者的角度了解如何开始使用思科 NX-OS 的许多强大的可编程功能。本白皮书还提供了许多重要的链接，便于读者进一步阅读有关协议、解决方案和设计的讨论。对 Cisco Nexus 9000 可编程功能的全面讨论不属于本文档的讨论范围。

1.2 免责声明

如需获得有关软件版本、支持的最大配置规模和设备规格的最新信息，请随时参考思科网站：http://www.cisco.com/go/nexus9000。

1.3 为什么选择 Cisco Nexus 9000 系列交换机

Cisco Nexus 9000 系列交换机（图 1）是中小型数据中心的理想之选，可提供五项主要优势：价格、性能、端口密度、可编程性和能效。

图 1. Cisco Nexus 9000 系列交换机产品

Cisco Nexus 9000 系列交换机采取商用增强型交换机设计方法，可实现成本效益。Cisco Nexus 9000 交换机的强大功能源自于思科自主开发的芯片及商用芯片专用集成电路 (ASIC) Trident II（有时缩写为 T2）。T2 ASIC 还可以提升能效。此外，Cisco Nexus 9000 系列交换机在 10 GE 和 40 GE 每端口密度价格方面领先于行业。Cisco Nexus 9000 采用经济高效的设计方法，并且具有丰富的功能集，这使得它非常适合用于商业数据中心。

Cisco Nexus 9000 的许可方式得到极大的简化。在撰写本文时，已经提供的许可证有两种：企业服务包许可证可以启用动态路由协议和虚拟可扩展局域网 (VXLAN) 支持；数据中心网络管理器 (DCNM) 许可证用于为整个数据中心网络提供单一平台 GUI 管理工具。今后随着新功能的推出，可能会提供其他许可证。

最后，Cisco Nexus 9000 提供强大的可编程功能，可支持充分采用 NX-API、Python、Chef 和 Puppet 等工具的新兴网络模型（包括自动化以及开发和运营 [DevOps] 模型）。

对于中小型商业客户，Cisco Nexus 9000 系列交换机产品是 1 GE 到 10 GE 迁移和 10 GE 到 40 GE 迁移的最佳平台，并且是替代数据中心内过时的 Cisco Catalyst® 交换机的理想产品。Cisco Nexus 9000 可以轻松与现有网络集成。本白皮书将介绍一种由两台 Cisco Nexus 9000 系列交换机组成的小型设计，并提供随着数据中心的发展横向扩展数据中心的路径，重点介绍了接入/汇聚设计以及主干/枝叶设计。

1.4 关于 Cisco Nexus 9000 系列交换机

Cisco Nexus 9000 系列包括较大的 Cisco Nexus 9500 系列模块化交换机和较小的 Cisco Nexus 9300 系列固定配置交换机。本白皮书后面部分将详细讨论产品配置。

思科为 Cisco Nexus 9000 系列交换机提供了两种操作模式。客户可以使用思科 NX-OS 软件在标准的 Cisco Nexus 交换机环境中部署 Cisco Nexus 9000 系列。或者，客户可以使用硬件就绪的思科以应用为中心的基础设施 (ACI)，以充分利用一种自动化、基于策略的系统管理方法。

除了传统的 NX-OS 功能（如虚拟 PortChannel [vPC]、服务中软件升级（ISSU - 未来）、加电自动调配 [POAP] 和 Cisco Nexus 2000 系列交换矩阵扩展器支持）外，Cisco Nexus 9000 上运行的单映像 NX-OS 还推出了多项重要的新功能：

- 智能的思科 NX-OS API (NX-API) 为管理员提供了在 HTTP/HTTPS 上通过远程过程调用（JSON 或 XML）来管理交换机的方法，而非直接访问思科 NX-OS 命令行。
- Linux Shell 访问使得可以通过 Linux Shell 脚本对交换机进行配置，有助于实现多个交换机配置的自动化，以及确保多个交换机之间的一致性。
- 通过冷修补和热修补方法保持持续运行，冷修补和热修补在常规的维护版本之间或在最终维护版本与维护终止版本之间以不具中断性的方式（对于热补丁而言）提供补丁。
- 硬件中以全线速提供的 VXLAN 桥接和路由可促进并加速虚拟和物理服务器之间的通信。VXLAN 提供与 VLAN 相同的第 2 层以太网服务，但是具有更高的灵活性且大规模提供这些服务。

有关升级的详细信息，请参阅《Cisco Nexus 9000 系列 NX-OS 软件升级和降级指南》。

本白皮书将重点介绍基本设计、集成的各项功能，例如 vPC、VXLAN、接入层设备连接和第 4-7 层服务插入。如需了解 Cisco Nexus 9000 系列交换机上的 NX-OS 的高级可编程功能，请参阅附录。其他功能不属于本白皮书的讨论范围。

2. ACI 就绪性

2.1 什么是 ACI?

采用思科以应用为中心的基础设施 (ACI) 的未来网络将是一种能够以支持快速应用变更的方式进行部署、监控和管理的网络。ACI 通过降低复杂性并采用可自动调配和管理资源的通用策略框架来实现这一目标。

思科 ACI 旨在解决以下业务问题：因主要专注于技术网络调配和变更管理问题而导致应用部署缓慢，解决方法是支持快速部署应用以满足不断变化的业务需求。思科 ACI 提供了一种集成的方法：提供以应用为中心的端到端可视性（从软件重叠一直到物理交换基础设施），同时它可以加快并优化第 4-7 层服务插入，以构建一个为网络提供应用语言的系统。

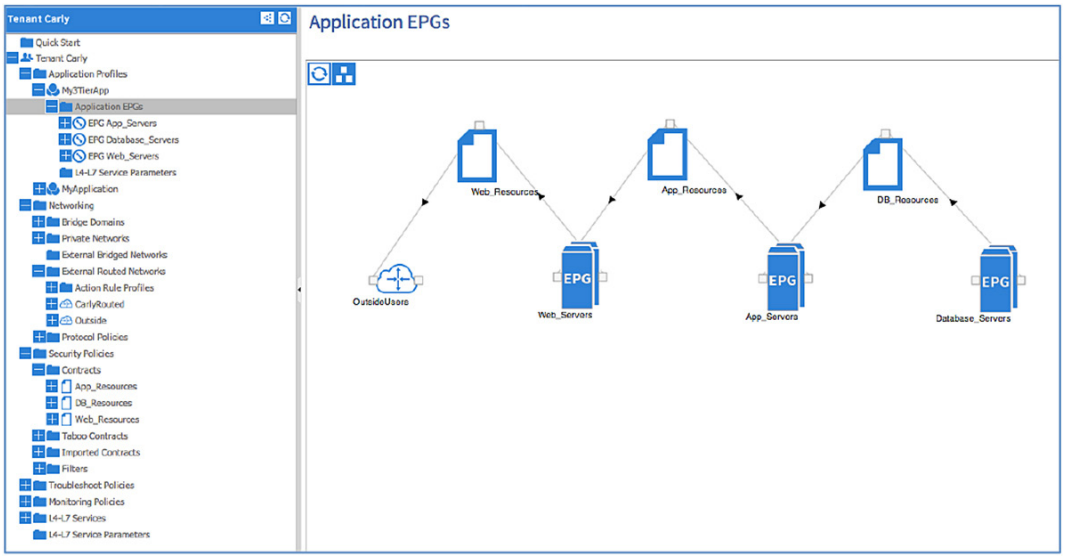
ACI 通过允许根据业务层应用需求使网络自动化和配置网络，来提供自动化、可编程性和集中调配功能。

ACI 可在网络和第 4-7 层基础设施中更快地统一部署应用，并在应用层提供可视性和管理能力。使用先进的遥感勘测功能可查看网络运行状况，简化的第二天运营也为应用本身的故障排除创造了机会。ACI 多样化、开放式的生态系统设计为可插入到任何上层的管理或协调系统，吸引了众多的开发者。利用思科和第三方的第 4-7 层虚拟和物理服务设备的集成和自动化功能，您可以使用单个工具管理整个应用环境。

借助 ACI 模式，客户可以采用策略的形式根据应用需求部署网络，无需改变当前网络限制的复杂性。此外，ACI 有助于确保安全性和性能，同时保持全面了解虚拟和物理资源的应用状况。

图 2 显示了如何从 ACI GUI 界面为三层应用定义网络通信。按照应用的需求定义网络，通过在应用配置文件内定义一组策略或合同确定允许谁与谁通信以及允许的通信内容是什么，而不是在多台交换机、路由器和应用上配置一行又一行的命令行界面 (CLI) 代码。

图 2. 三层应用策略示例



2.2 将 Nexus 9000 NX-OS 模式转换为 ACI 模式

本白皮书重点介绍 NX-OS（单机）模式下的 Cisco Nexus 9000 系列交换机。但是，Cisco Nexus 9000 硬件是 ACI 就绪型硬件。Cisco Nexus 9300 交换机和许多 Cisco Nexus 9500 线卡可以转换为 ACI 模式。

Cisco Nexus 9000 交换机是思科 ACI 架构的基础，并提供网络交换矩阵。在 ACI 模式下运行的 Cisco Nexus 9000 交换机使用新的操作系统。这些交换机与名为思科应用策略基础设施控制器 (APIC) 的集中式控制器及其开放式 API 结合使用。APIC 可统一对 ACI 交换矩阵进行自动化操作、遥感勘测和管理，帮助支持对数据中心采取应用策略模型方法。

在 Cisco Nexus 9000 上从单机 NX-OS 模式转换为 ACI 模式不属于本白皮书的讨论范围。

有关 Cisco Nexus 9000 系列交换机的 ACI 模式的详细信息，请参阅思科 ACI 网站。

3. 数据中心设计演进

本部分介绍了推动数据中心设计的关键考虑事项，以及 Cisco Nexus 9000 系列交换机如何处理这些考虑事项。

灵活的工作负载部署

虚拟机可以部署在数据中心内的任何位置，无需考虑机架的物理边界。在初始部署后，可能出于优化、整合或其他原因（这可能包括迁移到其他数据中心或公共云）移动虚拟机。该解决方案应提供机制以允许无缝地迁移虚拟机。所需的功能使用 VXLAN 和分布式任播网关实现。

数据中心内的东-西流量

数据中心流量模式正在不断变化。现在，更多的流量在服务器之间通过接入层按东-西方向流动，因为服务器彼此之间需要通信和使用数据中心的服务。这一转变主要由数据中心的整合、集群应用（例如 Hadoop）的演进、虚拟桌面和多租户推动。传统的三层数据中心设计不是最优的，因为东-西流量通常被迫上行通过核心层或汇聚层，采用次优路径。

东-西流量的需求由充分利用枝叶-主干架构的两层平面数据中心设计处理，在接入层的第一跳路由器实现最具体的路由。交换主机路由以帮助确保最具体的路由往返服务器和主机。通过检测虚拟机连接并用信号将新位置发送至网络其余部分来支持虚拟机移动性，以便至虚拟机的路由继续保持最优。

Cisco Nexus 9000 既可以用作行尾式或架顶式接入层交换机，也可以作为汇聚或核心交换机部署在传统的两层或三层分层网络设计或现代的枝叶-主干架构中。本白皮书将讨论接入/汇聚和枝叶/主干设计。

多租户以及第 2 层和第 3 层流量的分段

大型数据中心(尤其是运营商数据中心）需要托管多个客户和租户，他们将具有重叠的私有 IP 地址空间。数据中心网络必须允许来自共享网络基础设施上的多个客户的流量共存，并在这些流量之间提供隔离，除非特定的路由策略允许。流量分段通过以下方法实现：a) 使用 VXLAN 封装，其中 VNI（VXLAN 网络标识符）作为网段标识符；b) 通过虚拟路由转发 (VRF) 以多路分用重叠的私有 IP 地址空间。

消除或减少数据中心的第 2 层泛洪

在数据中心，第 2 层中的未知单播流量和广播流量泛洪，地址解析协议 (ARP) 和 Ipv6 邻居请求是广播流量的最重要部分。虽然 VXLAN 可以启用分布式交换域，这些域使得可以将虚拟机 (VM) 部署在数据中心内的任何位置，但是这样做的代价是必须向现在遍布数据中心的分布式交换域广播流量。

因此，必须为 VXLAN 实施补充一种将减少广播流量的机制。所需功能的实现方法：通过边界网关协议以太网 VPN (BGP EVPN) 分发 MAC 可达性信息，以优化与未知第 2 层单播流量相关的泛洪。通过 BGP EVPN 分发必要的信息并将其缓存在接入交换机，实现优化（减少与 ARP 和 IPv6 邻居请求相关联的广播）。地址请求可以在本地进行响应，无需发送广播。

多站点数据中心设计

大多数运营商数据中心以及大型企业数据中心有多个站点，以支持诸如地理覆盖范围、灾难恢复之类的需求。数据中心租户要求能够在不同的站点构建其基础设施，以及在私有和公共云上运营。

4. 数据中心运营演进

受下列一些需求推动，数据中心运营过去几年来迅速发展：

- **虚拟化 -** 所有流行的虚拟化平台（ESXi、OpenStack 和 HyperV）均内置了虚拟网络功能。物理网络元素必须跨这些虚拟网络元素进行联接和优化。
- **计算、网络和存储元素的融合 -** 用户希望通过集成的协调器实施完整的调配使用案例（即创建租户、创建租户网络、调配虚拟机、将虚拟机绑定到租户网络、创建 vDisk、将 vDisk 绑定到虚拟机）。为了实现此目的，所有基础设施元素需要提供 API 以允许协调器调配和监控设备。
- **DevOps -** 由于数据中心的规模已扩大，数据中心管理越来越多地通过编程框架（如 Puppet 和 Chef）进行。这些框架都具有一个通过在目标设备本地运行的“代理”控制目标设备的“主框架”。Puppet 和 Chef 使用户可以通过清单/菜谱（一组可重复使用的配置或管理任务）定义其意图，并允许将菜谱部署在由“代理”执行的众多设备上。

- **动态应用调配 -** 数据中心基础设施正在从支持相对静态的工作负载的环境过渡到高度动态的环境，在前一种环境中，工作负载局限于特定的基础设施孤岛，而在后一种环境中，任何工作负载都可以调配到任何位置，并且可以根据应用需求按需进行扩展。在调配、扩展和迁移应用时，需要无缝地强制执行其相应的基础设施要求、IP 寻址、VLAN 和策略。
- **软件定义网络 (SDN) -** SDN 将控制平面与网络内的数据平面分离，使得可以集中管理网络的智能和状态，同时将底层物理网络的复杂性抽象化。行业已统一采用诸如 OpenFlow 之类的协议。用户使用 SDN 将使网络可以动态适应各种应用（如 Hadoop、视频传输等）的应用需求。

Cisco Nexus 9000 旨在满足不断发展的数据中心的运营需求。有关各项功能的详细讨论，请参考本文档的第 6 部分。

- **可编程性和自动化**
  - **NX-API：**智能的思科 NX-OS API (NX-API) 为管理员提供了在 HTTP/HTTPS 上通过远程过程调用（JSON 或 XML）来管理交换机的方法，替代了直接访问思科 NX-OS 命令行的方式。
  - **与 Puppet 和 Chef 相集成：**Cisco Nexus 9000 交换机为代理提供了与 Puppet 和 Chef 相集成的方法，并提供了允许自动配置和管理 Cisco Nexus 9000 系列交换机的菜谱。菜谱在部署到 Cisco Nexus 9000 系列交换机后，会转换为网络配置设置以及用于收集统计数据和分析信息的命令。
  - **OpenFlow：**思科通过其 OpenFlow 插件（安装在 NX-OS 支持的设备上）和思科 ONE 企业网络控制器（安装在服务器上）支持 OpenFlow，并使用 OpenFlow 界面管理设备。ONE 控制器反过来向应用提供 Java 抽象以将网络抽象化和管理网络。
  - **Cisco OnePK™：**OnePK 是一个易于使用的工具包，用于开发、自动化、快速服务创建等各种用途。它支持 C、Java 和 Python，并与 PyCharm、PyDev、Eclipse、IDLE、NetBeans 等集成。借助其丰富的 API 集，您可以轻松访问网络内的宝贵数据并执行功能。示例包括自定义路由逻辑；创建基于流量的服务（例如服务质量 [QoS]）；根据不断变化的网络状况（例如带宽）调整应用；自动执行跨多个设备的工作流程；并为管理应用提供新信息。
  - **Guest Shell：**从思科 NX-OS 软件版本 6.1(2)I3(1) 开始，Cisco Nexus 9000 系列设备支持访问名为“Guest Shell”的分离执行空间。在 Guest Shell 内，网络管理员被授予 Bash 访问权限，并且可使用熟悉的 Linux 命令管理交换机。Guest Shell 环境：
    - 可访问网络，包括 NX-OS 软件已知的所有 VRF
    - 具有托管 Cisco Nexus 9000 Bootflash 所需的读取和写入访问权限
    - 能够执行 Cisco Nexus 9000 CLI
    - 可访问 Cisco onePK API
    - 能够开发、安装和运行 Python 脚本
    - 能够安装和运行 64 位 Linux 应用
    - 具有在系统重新加载或切换期间持续存在的根文件系统
- **与协调器和管理平台的集成**
  - **Cisco UCS® Director：**使用 Cisco UCS Director 可轻松管理思科融合基础设施平台、vBlock 和 FlexPod，并且它对有关思科 UCS 服务器、Nexus 交换机和存储阵列的使用案例提供端到端的调配和监控。
  - **Cisco Prime™ 数据中心网络管理器 (DCNM)** 是一个非常强大的工具，用于思科数据中心计算、网络和存储基础设施的集中数据中心监控、管理和自动化。DCNM 的基本版本免费提供，更高级的功能需要许可证。DCNM 允许集中管理所有 Cisco Nexus 交换机、思科 UCS 和思科 MDS 设备。

- **OpenStack：**OpenStack 是领先的开源云管理平台。Cisco Nexus 9000 系列包括对 OpenStack 的 Neutron 的插件支持。Cisco Nexus 插件接受 OpenStack Networking API 调用，并直接配置 Cisco Nexus 交换机以及 Open vSwitch (OVS)，后者在虚拟机监控程序上运行。Cisco Nexus 插件不仅在物理和虚拟网络上配置 VLAN，而且还智能地分配 VLAN Id，在不再需要 VLAN ID 时取消调配它，并在可能的情况下将其重新分配给新租户。系统将配置 VLAN，以便在属于同一租户网络的不同虚拟化（计算）主机上运行的虚拟机通过物理网络透明地进行通信。此外，系统对从计算主机到物理网络的连接进行中继，以便只允许流量来自虚拟交换机在主机上配置的 VLAN。有关详细信息，请访问 Cisco OpenStack。

- **基于策略的网络可解决动态应用调配：**Nexus 9000 在单机和 ACI 模式下运行。ACI 模式提供功能丰富、基于策略的网络结构和一个框架，以整体定义应用的基础设施需求以及调配和管理基础设施。ACI 支持的一些结构包括租户的定义、应用、终端组 (EPG)、合同与策略以及 EPG 与网络交换矩阵的关联。有关详细信息，请参阅《思科以应用为中心的基础设施设计指南》：http://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-731960.html。

5. 思科 NX-OS 软件的主要功能

适用于 Cisco Nexus 9000 系列交换机的思科 NX-OS 软件以两种模式运行：

- 单机的思科 NX-OS 部署
- 思科 ACI 部署

Cisco Nexus 9000 系列使用增强版的思科 NX-OS 软件，该版本具有单一的二进制镜像，可支持此系列中的每款交换机，从而简化映象管理。

思科 NX-OS 旨在满足各种客户的需求，包括中端市场、大型企业、运营商和一系列特定的行业。思科 NX-OS 使客户可以在数据中心为局域网和 SAN 创建稳定、标准的交换环境。思科 NX-OS 基于高度安全、稳定、标准的 Linux 内核，长期提供模块化、可持续的基础。思科 NX-OS 旨在统一并简化数据中心，为思科统一数据中心提供网络软件基础。

它的突出特点包括：

- **模块化 -** 思科 NX-OS 在设备内的控制平面与数据转发平面之间以及在软件组件之间提供隔离，因此一个平面或流程内的故障不会导致其他平面或流程中断。大多数系统功能、特性和服务处于隔离状态，因此在发生故障时，它们可以独立启动和重新启动，而其他服务则继续运行。大多数系统服务可以执行有状态重新启动，这使该服务可以对其他服务透明地恢复运行。

- **弹性 -** 思科 NX-OS 采用全新设计，可为要求最严苛的网络环境提供持续、可预测且高弹性的运行。通过精细的流程模块化、自动故障隔离和包容以及紧密集成的硬件弹性功能，思科 NX-OS 提供了高度可靠的操作系统来确保运行连续性。

  思科 NX-OS 弹性包括多种功能。思科服务中软件升级 (ISSU) 提供问题修复、功能增强甚至完整的操作系统升级，无需中断设备的运行。每个流程的模块化使客户可以重新启动各个流程或更新各个流程，而不会中断设备上的其他服务。思科 NX-OS 还允许将控制平面与数据平面分离；数据平面事件无法阻止控制命令的执行，有助于确保正常运行。

- **效率 -** 思科 NX-OS 包含大量传统和高级功能，可简化实施和持续运营。监控工具、分析器和集群技术集成到思科 NX-OS。这些功能提供可简化运营和提升效率的单一管理点。

  拥有跨数据中心网络的所有主要组件的单个网络软件平台将创建一个可预测、一致的环境，使配置网络、诊断问题和实施解决方案更加容易。

思科数据中心网络管理器 (DCNM) 是可以处理所有思科 NX-OS 设备的集中式管理器，允许集中管理在设备级别执行的所有监控和分析，并提供高水平的总体控制。此外，思科 NX-OS 提供在 Cisco IOS® 软件率先推出的同一行业标准命令行环境，使从 Cisco IOS 软件过渡到思科 NX-OS 软件变得容易。

- **虚拟化 -** 思科 NX-OS 旨在提供交换机级别的虚拟化。使用思科 NX-OS，可以在多个逻辑设备中虚拟化交换机，每个设备独立运行。在多租户环境以及出于监管原因必须严格分离的环境中，设备分区尤其有用。思科 NX-OS 提供 VLAN 和 VSAN，并且还支持新技术（例如 VXLAN），有助于实现网络隔离。整合到思科 NX-OS 中的技术提供网络与虚拟化服务器环境之间的紧密集成，从而实现简化数据中心资源的管理和调配。

有关思科 NX-OS 的其他信息，请参阅 Cisco Nexus 9500 和 9300 系列交换机 NX-OS 软件产品手册。

6. 数据中心设计考虑事项

6.1 传统数据中心设计

传统数据中心采用包含核心、汇聚和接入层的三层架构构建（图 3），或者采用汇聚层和核心层合并在一层的两层紧缩核心构建（图 4）。此架构可适应北-南流量模式，在此模式中，客户端数据来自广域网或互联网，将由数据中心的服务器处理，然后再从数据中心推出。此模式常用于网络服务等应用，因为在网络服务中，大多数通信在外部客户端与内部服务器之间进行。南北流量模式允许硬件超订用，因为大多数流量通过带宽较低的广域网或互联网瓶颈流入和流出。

在本文档中，经典网络是指许多数据中心环境中通常部署的典型三层架构。它具有不同的核心层、汇聚层和接入层，这些层一起为任何数据中心设计提供了基础。表 1 概述了典型的三层设计的每一层及其功能。

表 1. 经典的三层数据中心设计

层	说明
核心	此层为所有进出数据中心的流量提供高速数据包交换背板。核心层提供与多个汇聚模块的连接，并提供无单点故障 (SPOF) 的弹性第 3 层路由交换矩阵。核心层运行内部路由协议，例如开放最短路径优先 (OSPF) 或边界网关协议(BGP)，并且在数据中心内所有连接的网段之间平衡流量负载。
汇聚	此层提供重要的功能，例如服务模块集成、第 2 层域定义和转发以及网关冗余。服务器到服务器的多层流量流经汇聚层并且可以使用各种服务，例如防火墙和服务器负载均衡，以优化和保护应用。此层为所有北向和南向流量提供第 2 层和第 3 层分界点，并处理数据中心内的大多数东向和西向流量。
接入	此层是服务器以物理方式连接到网络的点。服务器组件由不同类型的服务器组成： • 具有集成交换机的刀片服务器 • 采用直通布线的刀片服务器 • 集群的服务器 • 可能包括大型机 接入层网络基础设施还包括各种模块化交换机和集成的刀片服务器交换机。交换机提供第 2 层和第 3 层拓扑，满足各种服务器广播域和管理要求。在现代数据中心，该层进一步划分成一个使用基于虚拟机监控程序的网络的虚拟接入层，此内容不属于本文档的讨论范围。

图 3. 传统的三层设计

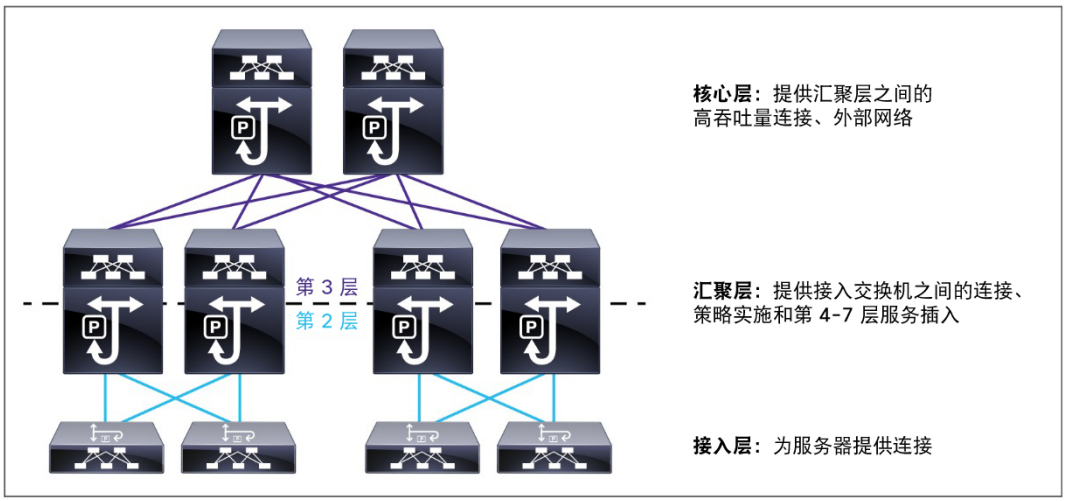
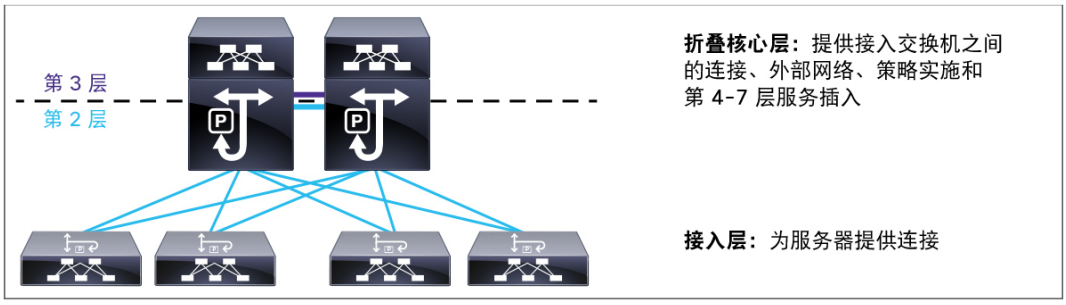


图 4. 传统的两层中型接入-汇聚设计



客户还可以通过 Pod 之间的第 3 层连接将小型紧缩设计复制到其他机架或建筑（图 5）。

图 5. 第 2 层接入、Pod 之间的第 3 层连接

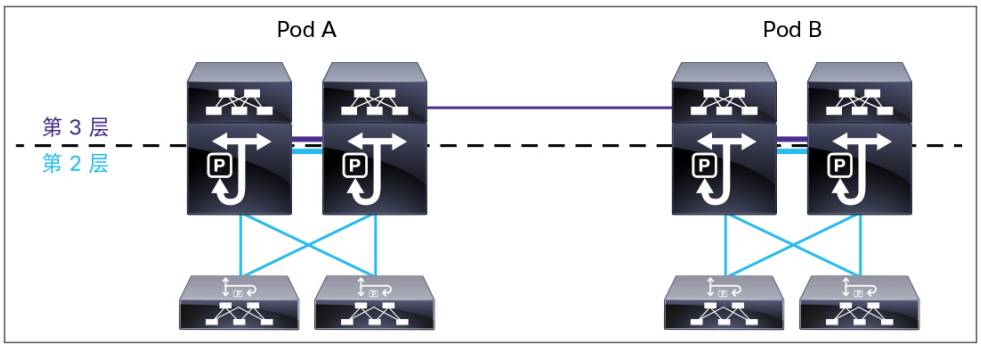


表 2 列出了在做出有关第 2 层-第 3 层边界的决定时考虑的一些因素。

表 2. 第 2 层和第 3 层的边界

考虑事项	位于核心层的第 3 层	位于接入层的第 3 层
多路径	每个 VLAN 一个活动路径，因为接入与核心交换机之间采用生成树模式	通过动态路由协议在接入与核心交换机之间进行等价多路径分流
生成树	第 2 层链路更多，因此需要阻止的环路和链路更多	没有生成树协议 (STP) 在接入层的北向运行
第 2 层可达性	针对工作负载移动性和集群应用提供更高的第 2 层可达性	第 2 层邻接关系限于连接到同一接入层交换机的设备
收敛时间	生成树的收敛速度通常比动态路由协议慢	动态路由协议的收敛速度通常比生成树快

6.2 枝叶-主干架构

主干-枝叶拓扑基于 Clos 网络架构。Clos 一词源自贝尔实验室的 Charles Clos，他在 1953 年发表了一篇论文，其中介绍了一种通过多路径、无阻塞、多级的网络拓扑交换电话呼叫的数学理论。

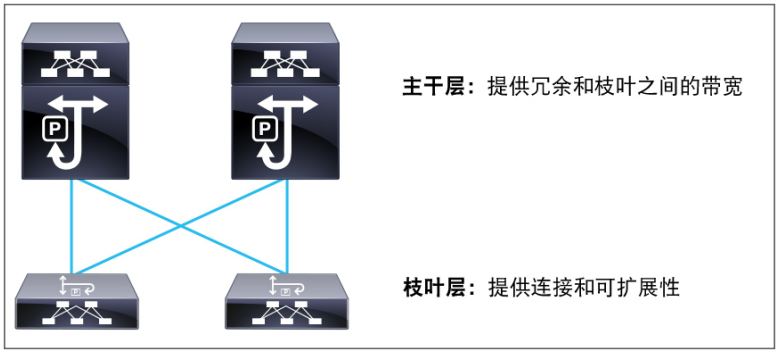
现在，Clos 的原始设计理念已应用于现代的主干-枝叶拓扑。主干-枝叶架构通常部署为两层：主干（例如汇聚层）和枝叶（例如接入层）。主干-枝叶拓扑提供高带宽、低延迟、无阻塞的服务器到服务器连接。

枝叶（汇聚）交换机使设备可以访问交换矩阵（由主干和枝叶交换机组成的网络），通常部署在机架顶部。一般情况下，设备连接到枝叶交换机。设备可以包括服务器、第 4-7 层服务（防火墙和负载均衡器）以及广域网或互联网路由器。枝叶交换机不连接到其它枝叶交换机（除非在 NX-OS 单机模式下运行 vPC）。但是，每个枝叶应连接到全网状网中的每个主干。枝叶上的某些端口将用于终端设备（通常为 10 GB），某些端口用于主干连接（通常为 40 GB）。

主干（汇聚）交换机用于连接到所有枝叶交换机，并且通常部署在行尾或行间。主干交换机不连接到其它主干交换机。主干用作枝叶交换机的骨干互联。一般来说，主干仅连接到枝叶，但是在将 Cisco Nexus 9000 交换机集成到现有环境时，将其他交换机、服务或设备连接到主干也是完全可以接受的。

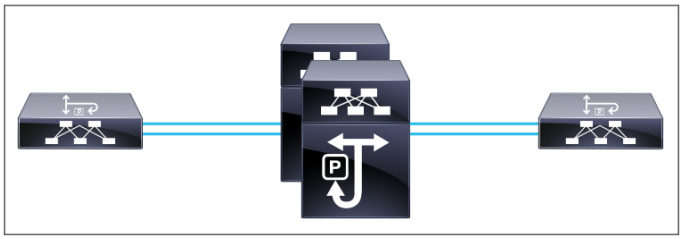
连接到交换矩阵的所有设备距离所有其他设备的跳数相同。这使服务器之间的延迟时间可预测并且提供高带宽。图 6 中的图显示了一种简单的两层设计。

图 6. 两层设计和连接



思考主干-枝叶架构的另一种方法是将主干看作中央骨干，所有枝叶像星星一样从主干分叉出去。图 7 描绘了这种逻辑表示，它在一个替代的可视化映射配置中使用相同的组件。

图 7. 两层设计的逻辑表示



6.3 生成树支持

Cisco Nexus 9000 系列支持两种生成树模式：增强型快速每 VLAN 生成树 (PVST+)（这是默认模式）和多生成树 (MST)。

快速 PVST+ 协议是按 VLAN 实施的 IEEE 802.1w 标准，即快速生成树协议 (RSTP)。快速 PVST+ 与 IEEE 802.1Q VLAN 标准互操作。其中 IEEE 802.1Q VLAN 标准要求为所有 VLAN 分配一个 STP 实例，而不是对每个 VLAN 都分配。默认情况下，快速 PVST+ 在设备上的默认 VLAN (VLAN1) 及新创建的所有 VLAN 上已启用。快速 PVST+ 与运行传统 IEEE 802.1D STP 的设备互操作。RSTP 是对原始 STP 标准 802.1D 的改进，它允许更快速地收敛。

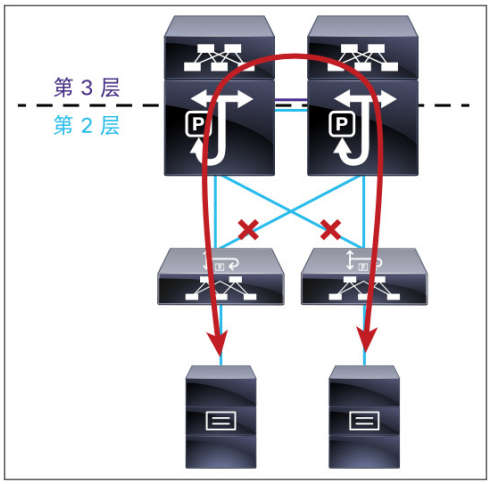
MST 将多个 VLAN 映射到生成树实例，每个实例具有独立于其他生成树实例的生成树拓扑。此架构为数据流量提供多个转发路径，启用负载均衡，并减少支持大量 VLAN 所需的 STP 实例的数量。MST 会提高网络的容错能力，因为某个实例中的故障不会影响其他实例。MST 模式通过显式握手提供快速收敛，因为每个 MST 实例使用 IEEE 802.1w 标准。MST 模式可改善生成树操作，并保持与原始的 802.1D 生成树和增强型快速每 VLAN 生成树（快速 PVST+）的向后兼容。

6.4 第 2 层与第 3 层影响比较

数据中心流量模式正在不断变化。现在，更多的流量在服务器之间通过接入层按东-西方向流动，因为服务器彼此之间需要通信和使用数据中心的服务。超订用的硬件现在不足以处理东-西的 10 GB 至 10 GB 通信。此外，东-西流量通常被迫上行通过核心层或汇聚层，采用次优路径。

生成树是传统的三层数据中心设计的另一种障碍。需要使用生成树阻止环路在以太网网络中泛滥，以便帧不会被无休止地转发。阻止环路意味着阻止链路，只保留一个活动路径（每个 VLAN）。被阻止的链路会严重影响可用带宽和超订用。这种情况还会迫使流量采用次优路径，因为生成树可能会阻止更理想的路径（图 8）。

图 8. 不同 Pod 中的服务器之间因生成树阻止链路而采用的次优路径



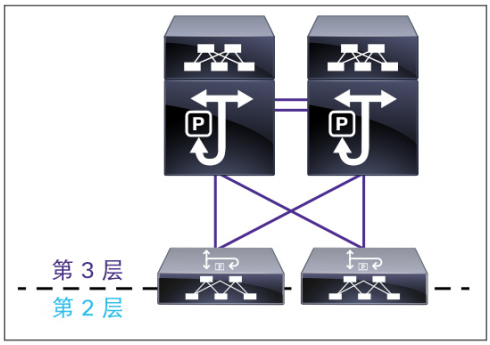
解决这些问题的方法可以包括：

A. 升级硬件以支持 40 GB 或 100 GB 接口

B. 将链路绑定到端口通道以显示为生成树的一条逻辑链路

C. 或将第 2 层/第 3 层边界下移至接入层以限制生成树的可达性（图 9）。在两层之间使用动态路由协议使所有链路成为活动链路，并且可以快速进行重新收敛和等价多路径分流 (ECMP)。

图 9. 两层的路由接入层设计



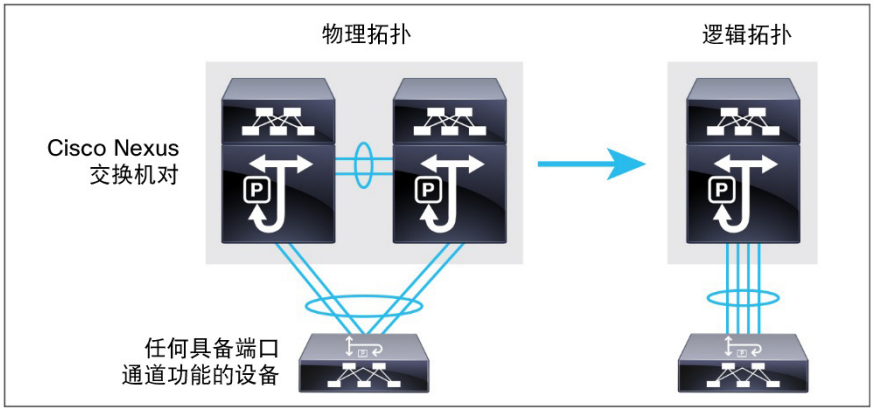
在传统以太网网络中，将第 3 层路由移至接入层时的权衡限制了第 2 层可达性。虚拟机工作负载移动性等应用和某些集群软件要求源服务器与目标服务器之间在第 2 层邻接。在接入层路由时，只有连接到同一接入交换机且相同 VLAN 向下中继的服务器在第 2 层邻接。此缺点通过使用 VXLAN 解决，后续部分将对此进行讨论。

6.5 虚拟端口通道 (vPC)

在过去几年，许多客户寻找各种方法来跨越生成树的限制。随着思科虚拟端口通道 (vPC) 的出现，通往基于 Cisco Nexus 的现代数据中心之路的第一步在 2008 年迈出。vPC 使设备可以使用单个逻辑端口通道接口连接到两台不同的 Cisco Nexus 物理交换机（图 10）。

在 vPC 之前，端口通道通常必须在单台物理交换机上终止。vPC 为设备提供双活动转发路径。由于两台 Cisco Nexus 交换机之间的特殊对等关系，生成树看不到任何环路，使所有链路都处于活动状态。对于连接的设备，若端口通道接口连接显示为正常，则不需要特殊配置。行业标准术语称为“多机箱 EtherChannel”；Cisco Nexus 特定的实施称为“vPC”。

图 10. vPC 物理拓扑与逻辑拓扑比较



在生成树以太网网络上部署的 vPC 是抑制被阻止链路数量非常有效的方式，因而可提高可用的带宽。Cisco Nexus 9000 交换机上的 vPC 解决方案非常适合商业客户以及对当前带宽、超订用和第 2 层可达性要求不满意的客户。

图 11 和 12 描绘了使用 vPC 的中小型传统商业拓扑的两个示例。这些设计使第 2 层/第 3 层边界位于汇聚层，以实现更大的第 2 层可达性，但是所有链路都处于活动状态，因为生成树看不见任何要阻止的链路。

图 11. vPC 的传统一层紧缩设计

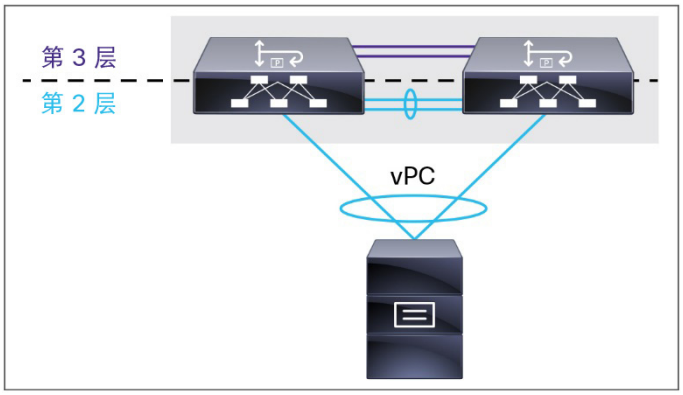
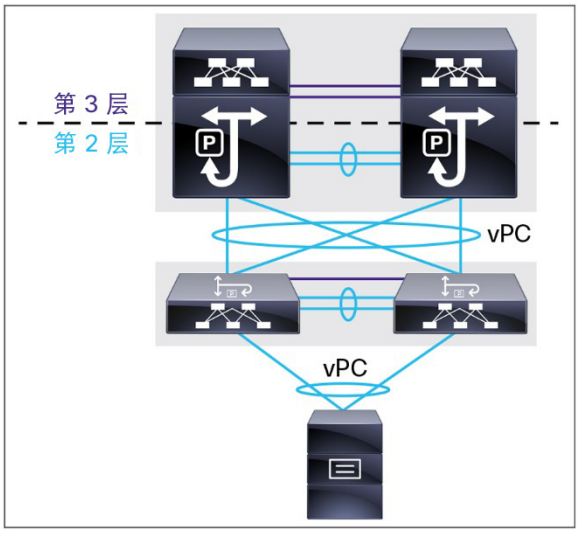


图 12. vPC 的传统两层设计



6.6 重叠

重叠技术有助于将网络的第 2 层边界扩展到大型数据中心以及不同的数据中心。本部分讨论一些重要的重叠技术，包括以下内容：

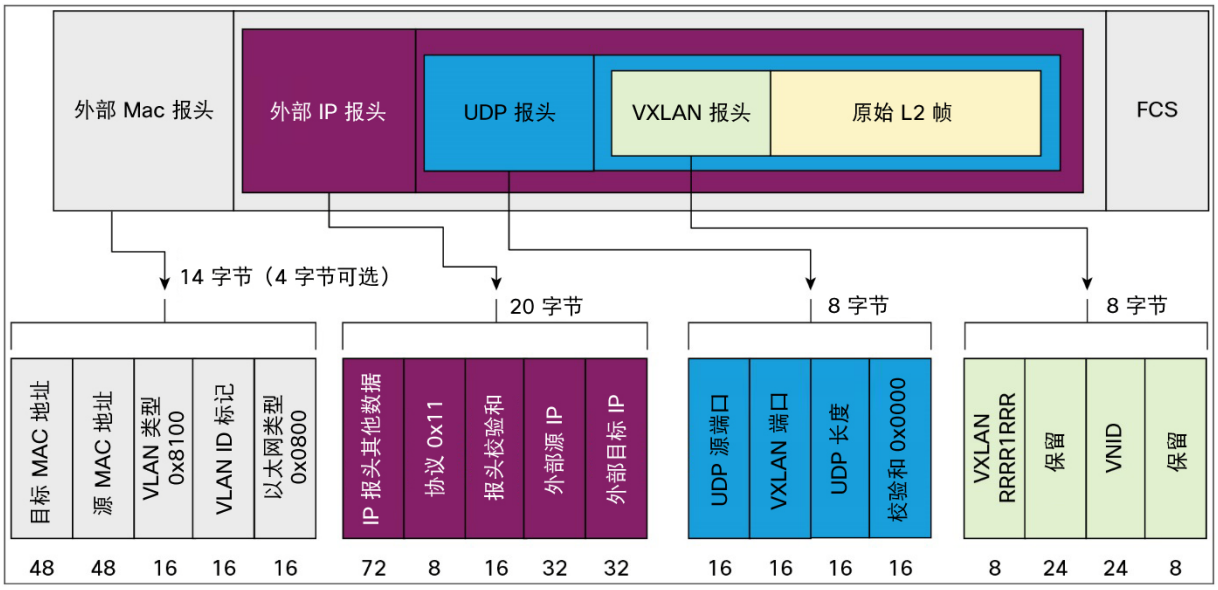
- VXLAN 概述
- 采用 BGP EVPN 作为控制平面的 VXLAN
- 采用 BGP 控制平面的 VXLAN 数据中心互联

有关重叠的更多一般信息，请阅读数据中心重叠技术白皮书。

6.6.1 虚拟可扩展局域网 (VXLAN)

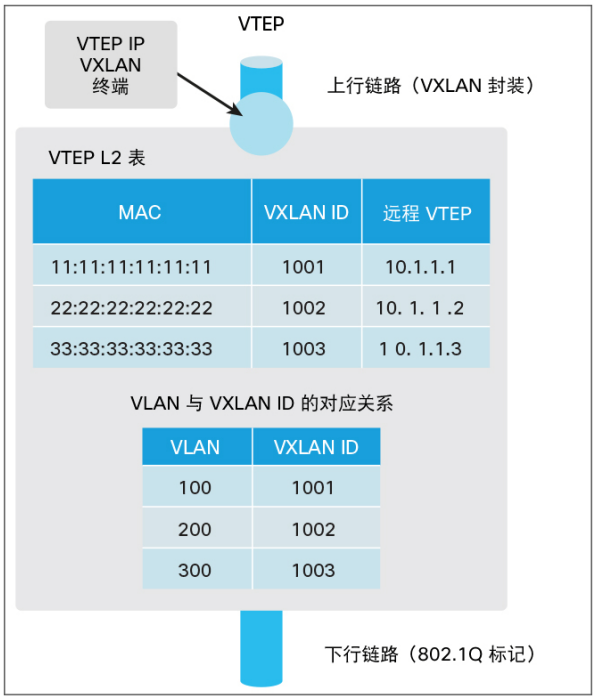
VXLAN 是在第 3 层网络上的第 2 层重叠方案。它使用一种 IP/用户数据报协议 (UDP) 封装，以便运营商或核心网络不需要了解 VXLAN 提供的任何其他服务。24 位 VXLAN 网段 ID 或 VXLAN 网络标识符 (VNI) 包括在封装中，为流量隔离和分段最多提供 1600 万个 VXLAN 网段，相比之下，可由 VLAN 实现的网段为 4000 个。每个这些网段均表示一个唯一的第 2 层广播域，并且可以采用唯一地标识给定租户的地址空间或子网的方式进行管理（图 13）。

图 13. VXLAN 帧格式



VXLAN 可以视为无状态的隧道机制，其中每个帧在 VXLAN 隧道终端 (VTEP) 根据一组规则封装或解封。VTEP 有两个逻辑接口：上行链路和下行链路（图 14）。

图 14. VTEP 逻辑接口

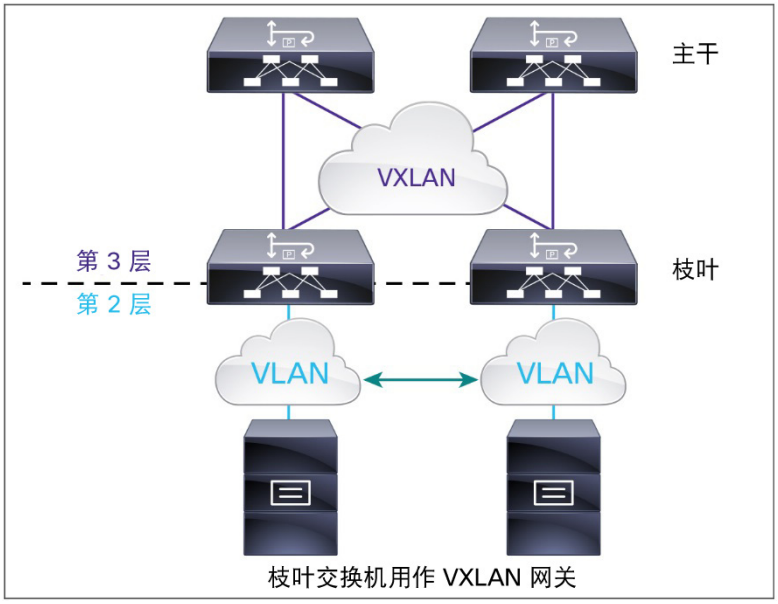


VXLAN 草案标准不要求为发现或学习使用控制协议。它为控制平面来源学习（推送模式）和基于中央目录的查找（拉取模式）提供建议。在撰写本文时，大多数实施依靠泛洪和学习机制来学习终端主机的可达性信息。在此模式中，VXLAN 在与原始 VTEP 相同的网段上建立至所有 VTEP 的一点对多点隧道，以便在交换矩阵中转发未知和多目标的流量。此转发通过关联每个网段的组播组完成，因此它要求底层交换矩阵支持 IP 组播路由。

Cisco Nexus 9000 系列交换机跨数据中心内的 IP 传输网络提供第 2 层连接扩展，并且在 VXLAN 与非 VXLAN 基础设施之间轻松集成。本文档后面的 9.3.2 部分提供了详细的配置，用于使用组播/互联网组管理协议 (IGMP) 方法及 BGP EVPN 控制平面配置 VXLAN 交换矩阵。

除了在这些交换机上发起和终止的简单重叠外，Cisco Nexus 9000 交换机还可以用作基于硬件的 VXLAN 网关。VXLAN 越来越普遍地用于虚拟机监控程序中的虚拟网络，它适用于虚拟机间的通信，而非仅限于交换机间的通信。但是，许多设备无法支持 VXLAN，例如传统的虚拟机监控程序、物理服务器以及服务设备，如防火墙、负载均衡器和存储设备。这些设备需要继续驻留在传统的 VLAN 网段。VXLAN 网段中的虚拟机有时需要访问由传统 VLAN 网段中的设备提供的服务，这种情况并不少见。用作 VXLAN 网关的 Cisco Nexus 9000 可以提供必要的转换，如图 15 所示。

图 15. Cisco Nexus 9000 枝叶交换机将 VLAN 转换为 VXLAN



6.6.2 VXLAN 的 BGP EVPN 控制平面

EVPN 重叠草案指定了对基于 BGP 多协议标签转换 (MPLS) 的 EVPN 解决方案的调整，使其可以应用为采用 VXLAN 封装的网络虚拟化重叠，其中：

- BGP MPLS EVPN 中描述的运营商边缘 (PE) 节点作用相当于 VTEP 或网络虚拟化边缘 (NVE) 设备
- 使用 BGP 分发 VTEP 终端信息
- VTEP 通过 BGP 为远程 MAC 地址使用控制平面学习和分发，而不是数据平面学习
- 广播、未知单播和组播数据流量使用共享组播树或通过入口复制方式发送
- BGP 路由反射器 (RR) 用于将 VTEP 之间的全网状 BGP 会话减少为 VTEP 与 RR 之间的单个 BGP 会话
- 路由过滤和受限制的路由分发用于帮助确保给定重叠的控制平面流量仅分发给位于该重叠实例中的 VTEP
- 主机 MAC 移动性机制可帮助确保重叠实例中的所有 VTEP 了解与 MAC 相关联的特定 VTEP
- 虚拟网络标识符在重叠内具有全局唯一性

还可以调整 VXLAN 的 EVPN 重叠解决方案，以应用为采用 VXLAN 的网络虚拟化重叠，以便进行第 3 层流量分段。第 3 层 VXLAN 的调整类似于第 2 层 VXLAN，以下方面除外：

- VTEP 通过 IP 地址（而不是 MAC 地址）的 BGP 使用控制平面学习和分发
- 虚拟路由和转发实例映射到 VNI
- VXLAN 报头中的内部目标 MAC 地址不属于主机，但是属于执行 VXLAN 负载路由的接收 VTEP。使用 BGP 属性及 EVPN 路由分发此 MAC 地址

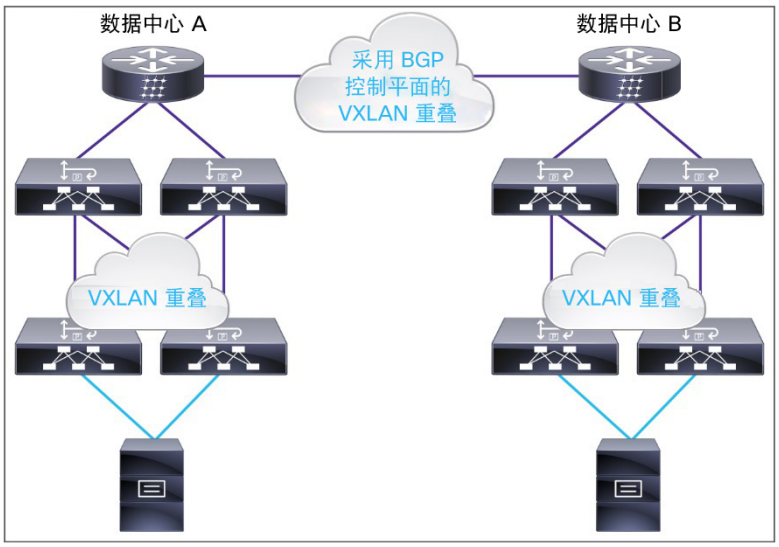
请注意，由于 IP 主机具有关联的 MAC 地址，因此将支持第 2 层 VXLAN 和第 3 层 VXLAN 重叠共存。此外，第 2 层 VXLAN 重叠还将用于促进非基于 Ip（仅限第 2 层）的主机之间的通信。

6.6.3 采用 BGP 控制平面的 VXLAN 数据中心互联 (DCI)

Cisco Nexus 9000 的 BGP EVPN 控制平面功能可以用于将第 2 层域的范围扩展到数据中心 Pod、域和站点。

图 16 显示了正在使用的两个重叠网络：思科 ASR 1000 上的重叠传输虚拟化 (OTV)（用于数据中心之间的连接），以及数据中心内的 VXLAN。两者均提供第 2 层可达性和扩展。Cisco Nexus 7000 系列交换机上也提供 OTV。

图 16. 虚拟重叠网络为应用提供动态可达性



7. 与现有网络集成

在将数据中心迁移到 Cisco Nexus 9000 系列交换机时，您不仅需要考虑与现有传统服务器和设备的兼容性，还需要考虑 Cisco Nexus 9000 系列的下一代功能，这包括：

- 添加了用于实现扩展的 BGP 控制平面功能的 VXLAN
- 交换矩阵扩展器 (FEX)
- vPC
- 10/40 Gbps 连接
- 可编程性

Cisco Nexus 9000 系列交换机具有出色的性能和全面的功能集，可以部署在多种场景中的通用平台，包括以下场景：

- 分层的接入-汇聚-核心设计
- 枝叶和主干架构
- 紧凑型汇聚层解决方案

Cisco Nexus 9000 系列交换机提供全面的思科 NX-OS 软件数据中心交换功能集。表 2 列出了当前的外形规格。访问 http://www.cisco.com/go/nexus9000，了解 Cisco Nexus 9000 产品组合的最新更新。

表 3. Cisco Nexus 9000 系列交换机

设备型号	线卡和扩展模块	说明	部署
Cisco Nexus 9500 模块化交换机	N9K-X9636PQ	36 个端口的 40 Gbps 增强型四通道小型封装热插拔 (QSFP+)	行尾式 (EoR)、行间式 (MoR)、汇聚层和核心
	N9K-X9564TX	48 个端口的 1/10GBASE-T 和 4 个端口的 40 Gbps QSFP+	
	N9K-X9564PX	48 个端口的 1/10 Gbps SFP+ 和 4 个端口的 40 Gbps QSFP+	
Cisco Nexus 9396PX 交换机	N9K-C9396PX	带 48 个端口的 1/10 Gbps SFP+ 的 Cisco Nexus 9300 平台	架顶式 (ToR)、EoR、MoR、聚合层和核心
Cisco Nexus 93128TX 交换机	N9K-C93128TX	带 96 个端口的 1/10GBASE-T 的 Cisco Nexus 9300 平台	ToR、EoR、MoR、汇聚层和核心

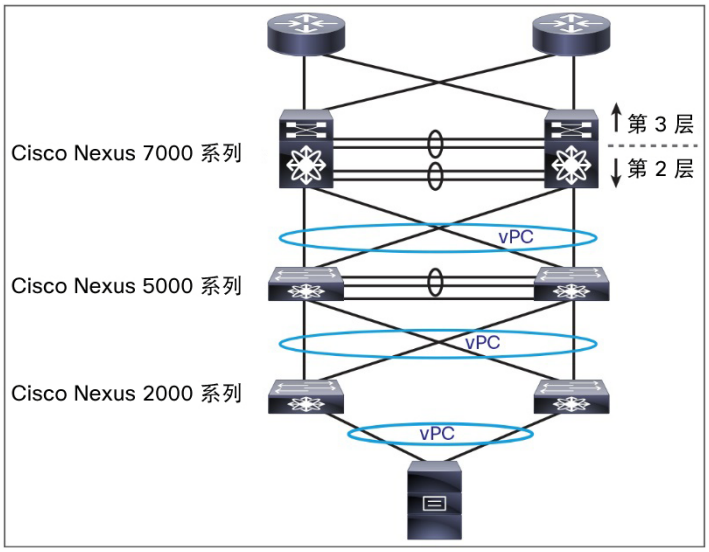
7.1 采用 vPC 的 Pod 设计

vPC 可使物理上连接到两台不同 Cisco Nexus 9000 系列交换机的链路对第三个设备显示为单个端口通道。vPC 可实现第 2 层多路径处理，因此您可通过增加带宽、在节点间支持多条并行路径，以及在有备选路径的条件下平衡流量负载，来创建冗余。

vPC 设计与 vPC 设计指南所述保持相同，例外之处在于 Cisco Nexus 9000 系列不支持 vPC 双活 FEX 或两层 vPC (eVPC)。有关详细信息，请参阅 vPC 设计和最佳实践指南：http://www.cisco.com/en/US/docs/switches/datacenter/sw/design/vpc_design/vpc_best_practices_design_guide.pdf。

图 17 显示了采用 Cisco Nexus 交换机和 vPC 的下一代数据中心。Cisco Nexus 7000 系列交换机与 Cisco Nexus 5000 系列交换机之间具有 vPC，Cisco Nexus 5000 系列交换机与 Cisco Nexus 2000系列 FEX 之间具有双宿主 vPC，服务器与 Cisco Nexus 2000 系列 FEX 之间具有双宿主 vPC。

图 17. Cisco Nexus 7000 系列位于核心层的 vPC 设计考虑事项



在 vPC 拓扑中，汇聚与接入层之间的所有链路均在转发，并且是 vPC 的一部分。

千兆以太网连接利用了后续部分中介绍的 FEX 概念。生成树协议不在 Cisco Nexus 9000 系列交换机与 Cisco Nexus 2000 系列 FEX 之间运行。相反，专有技术使交换机与交换矩阵扩展器之间的拓扑不存在环路。将 vPC 添加到接入层中的 Cisco Nexus 9000 系列交换机可将额外的负载从服务器分配到交换矩阵扩展器，再分配到 Cisco Nexus 9000 系列交换机。

现有的 Cisco Nexus 7000 系列交换机可以替换为 Cisco Nexus 9500 平台交换机，但有一个例外：Cisco Nexus 9000 系列交换机不支持 vPC 双活或两层 vPC (eVPC) 设计。网络拓扑和设计的其余部分没有变化。图 18 显示了新拓扑。图 19 显示了 Cisco Nexus 9500 平台之间的对等连接。

图 18. Cisco Nexus 9500 平台位于核心层的 vPC 设计

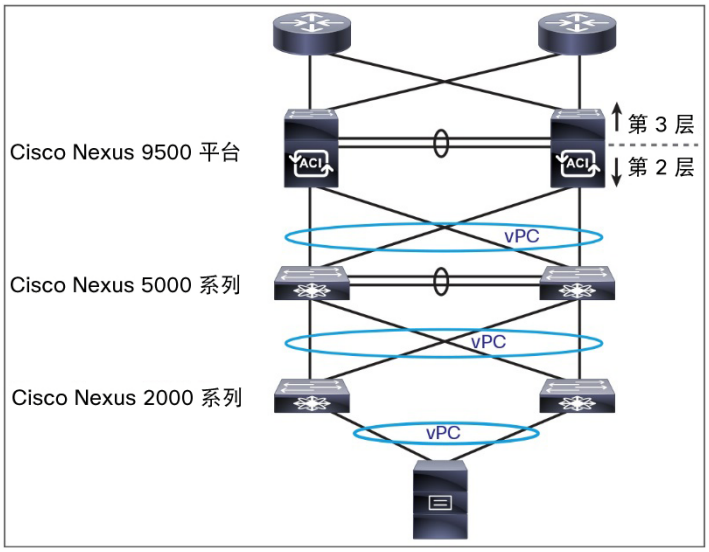
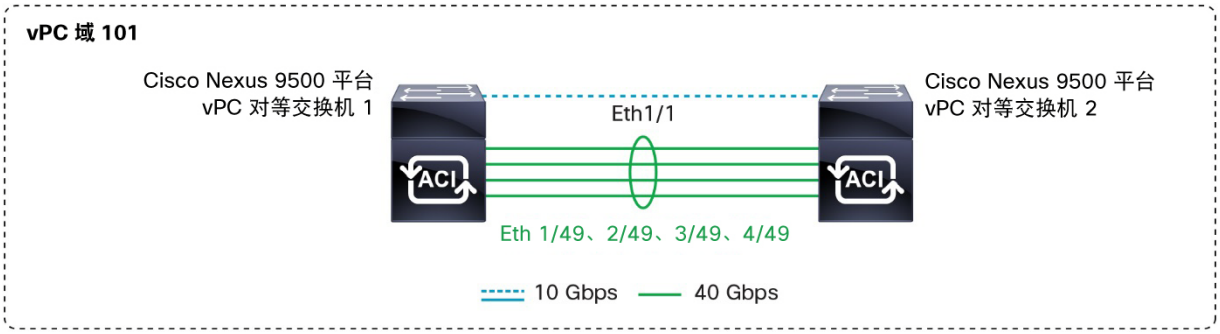


图 19. Cisco Nexus 9500 平台之间的对等连接



7.2 交换矩阵扩展器支持

为使用现有硬件、提高接入端口密度和提高 1 千兆以太网 (GE) 端口可用性，可以在单宿主的直通配置中，将 Cisco Nexus 2000 系列交换矩阵扩展器连接到 Cisco Nexus 9300 平台。最多 16 个交换矩阵扩展器可以连接到单个 Cisco Nexus 9300 系列交换机如果配置在 vPC 中，连接到交换矩阵扩展器的主机上行链路可以是主用/备用或主用/主用链路。

当前支持以下 Cisco Nexus 2000 系列交换矩阵扩展器：

- N2224TP
- N2248TP
- N2248TP-E
- N2232TM
- N2232PP
- B22HP

图 20. Cisco Nexus 2000 系列交换矩阵扩展器

如需了解最新的功能支持，请参阅 Cisco Nexus 9000 软件版本说明。

还支持交换矩阵扩展器收发器 (FET)，以在 Cisco Nexus 2000 交换矩阵扩展器与其父交换机 Cisco Nexus 9300 之间提供经济高效的连接解决方案 (FET-10 Gb)。

有关 FET-10 GB 收发器的更多信息，请参阅 Nexus 2000 系列交换矩阵扩展器产品手册。

图 21 和 22 中显示了支持的 Cisco Nexus 9000 至 Nexus 2000 交换矩阵扩展器拓扑。对于其他 Cisco Nexus 平台，可将思科交换矩阵扩展器技术看作父交换机 Cisco Nexus 9000 的逻辑远程线卡。每个思科 FEX 连接到一台父交换机。服务器应双宿至两个不同的交换矩阵扩展器。服务器上行链路可以在主用/备用网络接口卡 (NIC) 组合中，也可以在 vPC 中（如果在 vPC 域中设置父交换机 Cisco Nexus 9000）。

图 21. 支持的 Cisco Nexus 9000 系列交换机和交换矩阵扩展器设计（对于主用/备用服务器）

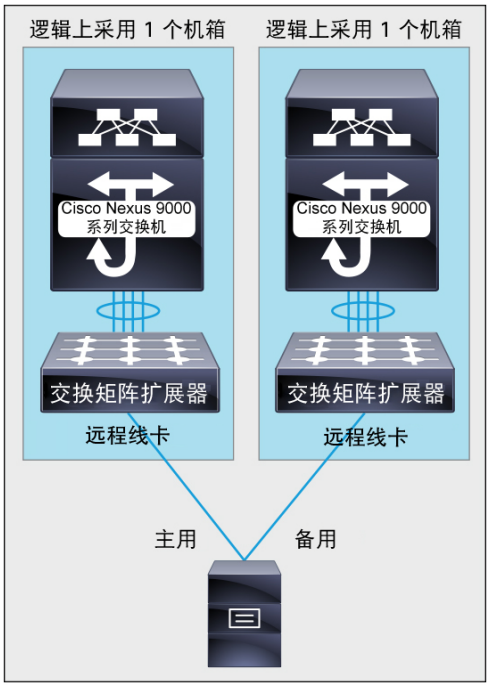
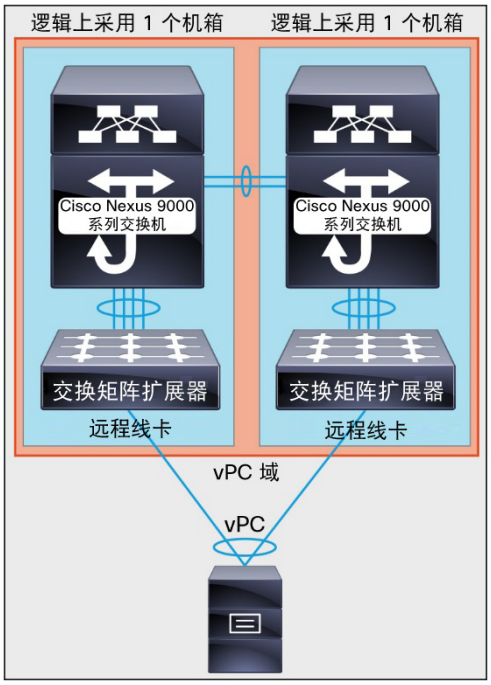


图 22. 支持的 Cisco Nexus 9000 系列交换机和交换矩阵扩展器设计（对于服务器 vPC）



有关详细信息，请参阅 Cisco Nexus 9000 系列交换机的 Cisco Nexus 2000 系列 NX-OS 交换矩阵扩展器配置指南 6.0 版。

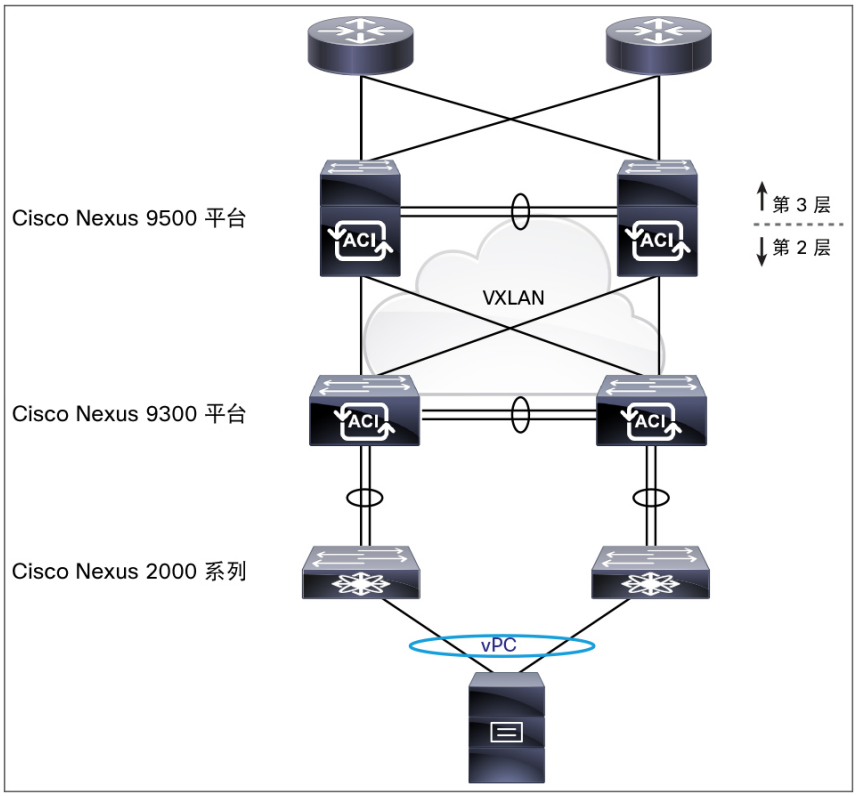
7.3 采用 VXLAN 的 POD 设计

Cisco Nexus 9500 平台在第 3 层网络上使用第 2 层重叠方案 VXLAN。可在基于虚拟机监控程序的虚拟交换机上实施 VXLAN 以允许可扩展的虚拟机部署，也可在物理交换机上实施 VXLAN 以将 VXLAN 网段桥接回 VLAN 网段。

VXLAN 将第 2 层网段 ID 字段扩展到 24 位，可能最多允许 1600 万个唯一的第 2 层网段（相比之下，在同一网络上可通过 VLAN 实现 4000 个网段）。每个这些网段均表示一个唯一的第 2 层广播域，并且可以采用唯一地标识给定租户的地址空间或子网的方式进行管理。请注意，核心层和接入层交换机必须是Cisco Nexus 9000系列交换机实施的VXLAN。

在图 23 中，位于核心的 Cisco Nexus 9500 平台提供第 2 层和第 3 层连接。Cisco Nexus 9500 和 9300 平台通过 40 Gbps 链路连接，并在平台之间使用 VXLAN。现有的 FEX 交换机使用链路汇聚控制协议 (LACP) 端口通道单宿至每台 Cisco Nexus 9300 平台交换机。终端服务器是双宿至两台 Cisco Nexus 2000 系列 FEX 的 vPC。

图 23. Cisco Nexus 9000 系列的 VXLAN 设计

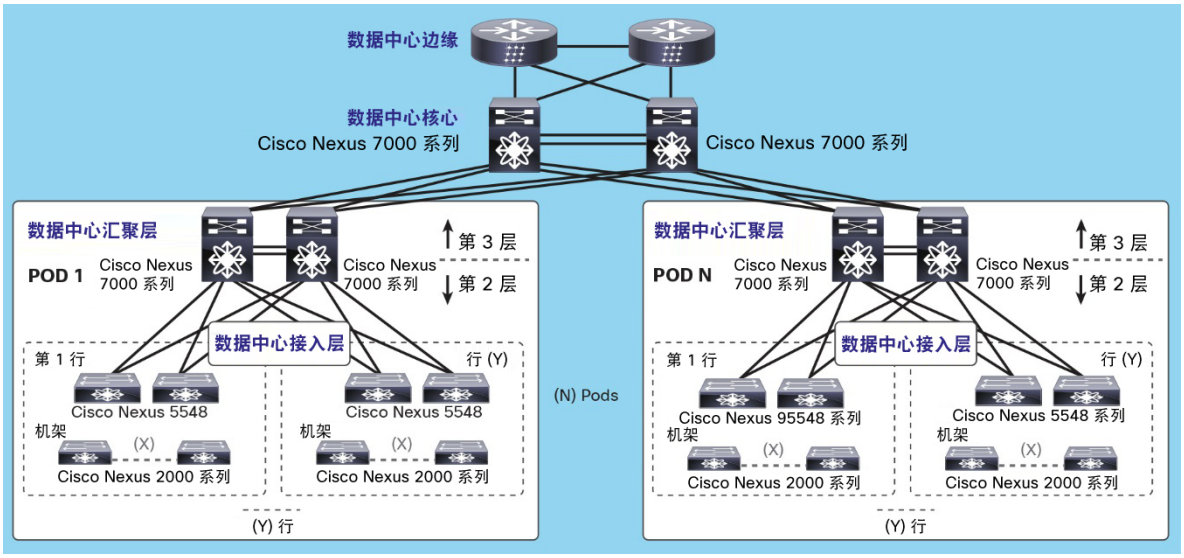


7.4 接入 1/10 千兆以太网服务器的传统三层架构

在典型的数据中心设计中，汇聚层需要高水平的灵活性、可扩展性和功能集成，因为汇聚设备构成第 3 层和第 2 层边界，该边界需要路由和交换功能。接入层连接定义了总转发能力、端口密度和第 2 层域灵活性。

图 24 描绘了位于核心和汇聚层的 Cisco Nexus 7000 系列交换机，在该设计中，一对数据中心核心交换机通常使用 10 千兆以太网第 3 层接口互联多个汇聚模块。

图 24. 经典的三层设计



选项 1：位于核心和汇聚层的 Cisco Nexus 9500 平台

在此设计中，Cisco Nexus 9500 平台（图 24）取代了位于核心和汇聚层的 Cisco Nexus 7000 系列。

Cisco Nexus 9508 8 插槽交换机是具有以下功能的下一代高密度模块化交换机：

- 现代操作系统
- 高密度（40/100 Gbps 汇聚）
- 低能耗

Cisco Nexus 9500 平台使用 Broadcom Trident-2 专用集成电路 (ASIC) 与 Insieme ASIC 的独特组合，以提供更快的部署速度、增强的数据包缓冲能力和全面的功能集。

Cisco Nexus 9508 机箱是一款具有 13 个机架单元 (13RU)、8 个插槽的模块化机箱，采用自前而后气流，非常适合大型数据中心部署。Cisco Nexus 9500 平台最多支持 3456 个 10 千兆以太网端口和 864 个 40 千兆以太网端口，并且每个机架系统可以实现 30 Tbps 的交换矩阵吞吐量。

Cisco Nexus 9508 包括以下通用设备：

- 两个半插槽管理引擎
- 四个冗余电源
- 三个交换矩阵（可升级到六个）
- 三个热插拔风扇托架

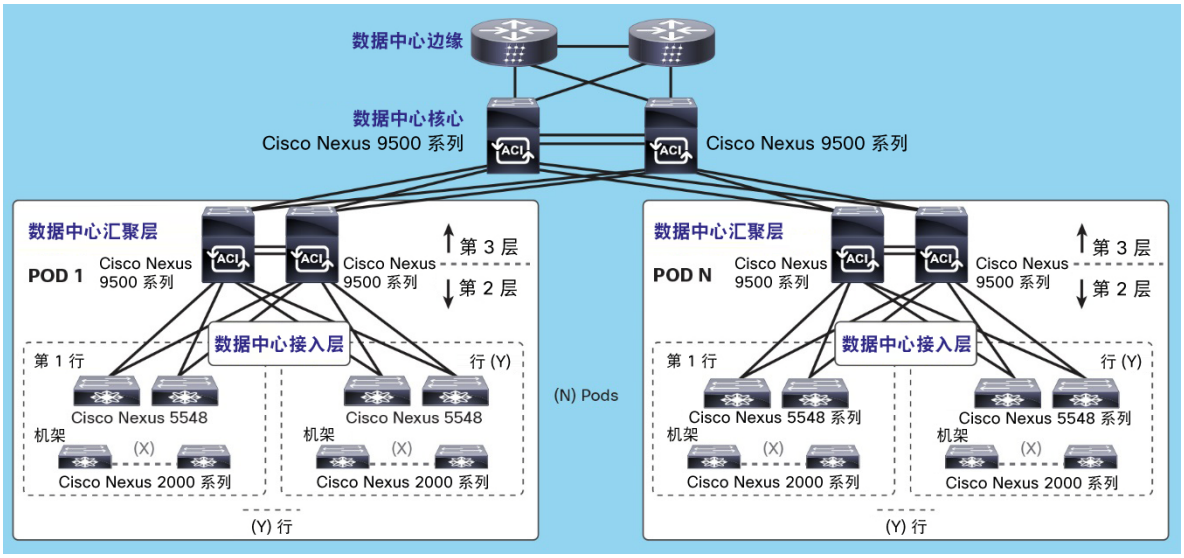
可通过机箱的后部操作风扇托架和交换矩阵模块。机箱有专用于 I/O 模块的八个水平插槽。

Cisco Nexus 9508 交换机可以完全填充 10 千兆、40 千兆和（未来）100 千兆以太网模块，不存在任何带宽或插槽限制。所有八个 I/O 插槽均支持在线插拔所有线卡。

选项 2：位于核心层的 Cisco Nexus 9500 平台和位于汇聚层的 Cisco Nexus 9300 平台

根据数据中心的增长，可以采用 Cisco Nexus 9500 平台在核心层和汇聚层（图 25）以及 Cisco Nexus 9500 平台在核心层、Cisco Nexus 9300 平台在汇聚层的组合，以实现更好的可扩展性（图 26）。

图 25. 基于 Cisco Nexus 9500 平台的设计



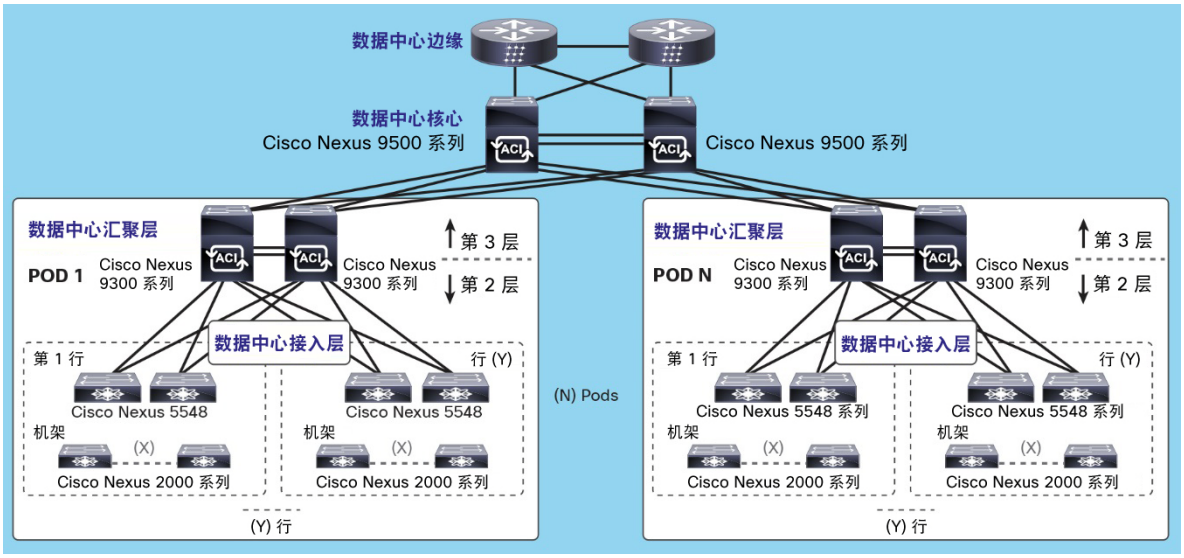
Cisco Nexus 9300 平台目前可采用两种固定配置：

- Cisco Nexus 9396PX：2 个 RU，具有 48 个 10 Gbps 的端口和 12 个 40 Gbps 的端口
- Cisco Nexus 93128TX：3 个 RU，具有 96 个 1/10 Gbps 的端口和 8 个 40 Gbps 的端口

在这两种选项中，可以将位于核心和汇聚层的现有 Cisco Nexus 7000 系列交换机换成 Cisco Nexus 9508 交换机，同时保留现有的布线连接。

目前，此设计不支持以太网光纤通道 (FCoE)。

图 26. 基于 Cisco Nexus 9500 和 9300 平台的设计



7.5 传统的思科统一计算系统和刀片服务器接入

在多层数据中心设计中，您可以使用 Cisco Nexus 9500 平台替换核心层的 Cisco Nexus 7000 系列交换机，或者使用 Cisco Nexus 9500 平台替换核心层或使用 Cisco Nexus 9300 平台替换接入层。您还可以将现有的思科统一计算系统™ (Cisco UCS®) 和刀片服务器接入层连接到 Insieme 硬件（图 27 和 28）。

图 27. 使用 Cisco Nexus 7000 和 5000 系列以及交换矩阵扩展器的经典设计

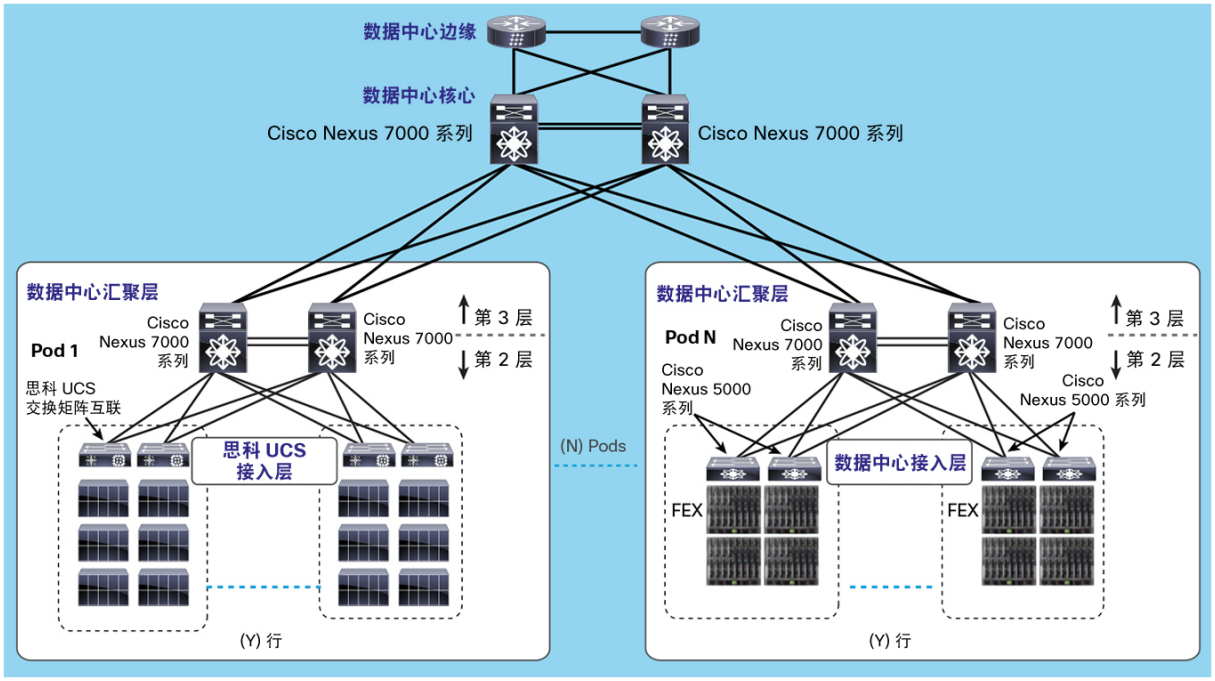
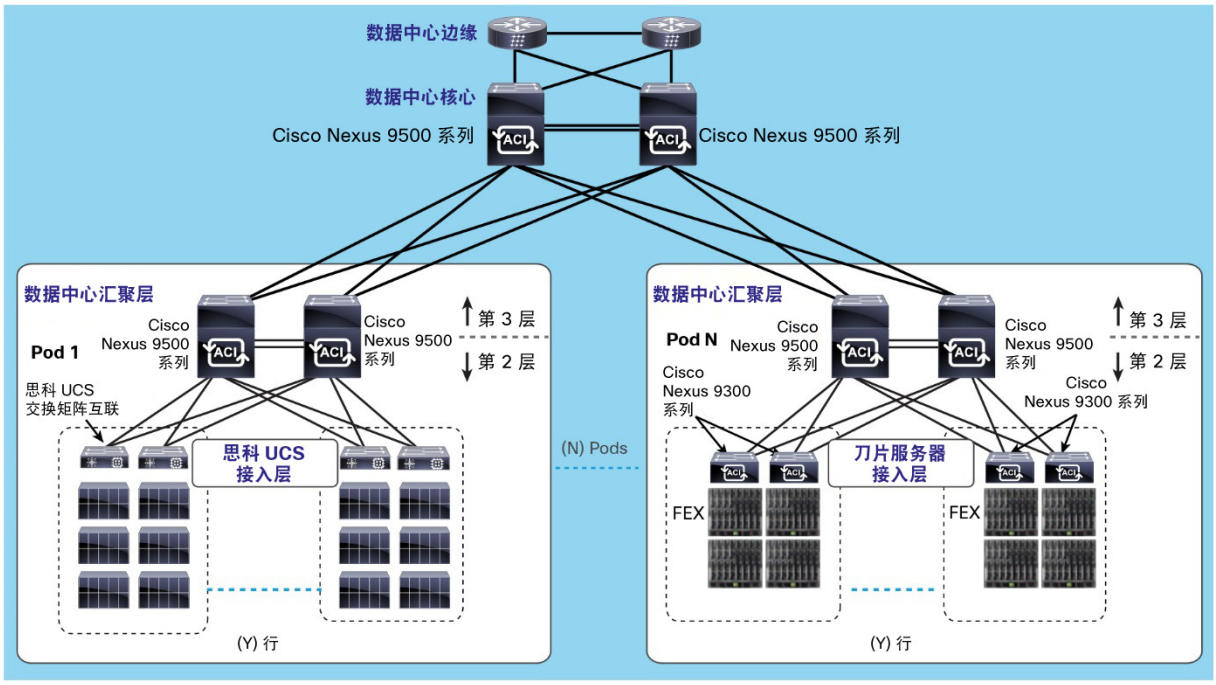


图 28. 使用 Cisco Nexus 9500 和 9300 系列以及交换矩阵扩展器的经典设计



8. 集成第 4-7 层服务

Cisco Nexus 9000 系列交换机可集成到任何供应商服务设备。第 9 部分概述了各种拓扑，其中 Cisco Nexus 9000 连接到：

- 路由模式下的思科 ASA 防火墙，ASA 充当连接到被屏蔽子网的主机的默认网关
- 单臂模式下的 F5 网络负载均衡器

可以在路由或透明模式下连接防火墙。它们需要满足典型数据中心场景的以下条件：

- 一个外部用户访问 Web 服务器
- 内部用户访问其它网络中的内部主机
- 内部用户访问外部 Web 服务器

有关连接思科 ASA 防火墙的详细信息，请访问 http://www.cisco.com/c/en/us/td/docs/security/asa/asa93/configuration/general/asa-general-cli/intro-fw.html。

9. Cisco Nexus 9000 数据中心拓扑设计和配置

此部分将使用实际的设备介绍 Cisco Nexus 9000 设计的一个示例，这些设备包括 Cisco Nexus 9508 和 Nexus 9396 交换机、VMware ESXi 服务器和独立的裸机服务器以及思科 ASA 防火墙。已在思科实验室中测试此拓扑，以验证和演示 Cisco Nexus 9000 解决方案。

本部分的目的是提供一个网络设计和配置示例，包括本白皮书前面部分讨论的功能。请务必参考思科配置指南，获取最新的信息。

本部分使用的术语和交换机名称引用“枝叶”和“主干”，但相似的配置可以应用于接入-汇聚设计。

9.1 硬件和软件规格

- 两台 Cisco Nexus N9K-C9508 交换机（8 个插槽），每台交换机中具有以下组件：
  - 一个 N9K-X9636PQ（36 端口 QSFP 40 千兆）以太网模块
  - 六个 N9K-C9508-FM 交换矩阵模块
  - 两个 N9K-SUP-A 管理引擎模块
  - 两个 N9K-SC-A 系统控制器
- 两台 Cisco Nexus 9396PX 交换机（48 端口 SFP+ 1/10 千兆），每台交换机中具有以下扩展模块：
  - 一个 N9K-M12PQ（12 端口 QSFP 40 千兆）以太网模块
- 三台思科 UCS C 系列服务器（对于较小型的部署，可以替换为思科 UCS 快捷版服务器）。有关服务器选项的详细信息，请访问思科统一计算系统快捷版。
- 一台思科 ASA 5510 防火墙设备
- 一台 F5 网络 BIG-IP 负载均衡器设备

本设计中使用的 Cisco Nexus 9000 交换机运行思科 NX-OS 软件版本 6.1(2)I2(3)。使用的 VMware ESXi 服务器运行 ESXi 5.1.0 版本 799733。思科 ASA 运行 8.4(7) 版本。F5 运行 11.4.1 内部版本 625.0 修补程序 HF1。

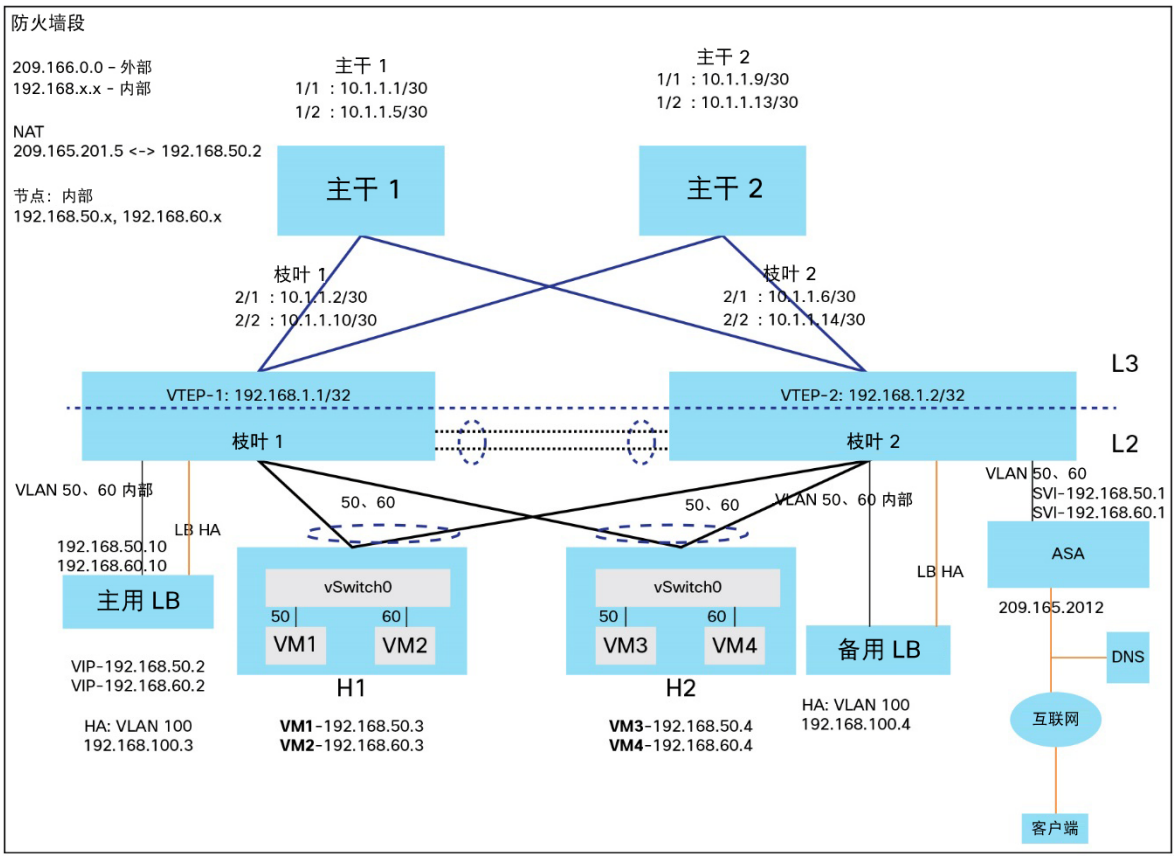
9.2 基于枝叶-主干的数据中心

9.2.1 拓扑

图 29 显示了基本网络设置。它包括：

- 应用 - Cisco.app1.com：Ex - 外部 IP- 209.165.201.5
- 此应用的外部 Web 服务器在负载均衡器（IP 为 192.168.50.2）和应用服务器（IP 为 192.168.60.2）的内部区域中运行
- Web 服务器节点在以下 IP 地址运行：192.168.50.5、192.168.50.6
- 应用服务器节点在以下 IP 地址运行：192.168.60.5、192.168.60.6

图 29. 枝叶-主干架构的第 4-7 层服务拓扑



在这个简单拓扑中，Web 服务器虚拟机连接到 VLAN 50，应用服务器虚拟机连接到 VLAN 60，IP 寻址方案分别为 192.168.50.x 和 192.168.60.x。服务器虚拟机网络被称为内部。服务器虚拟机连接到 vPC 模式下的枝叶交换机。

枝叶交换机端口接口连接到防火墙和负载均衡设备。枝叶路由端口接口连接到数据中心内的每个主干接口。使用每个枝叶中的环回接口创建 VTEP。

服务器虚拟机的默认网关在思科 ASA 5510 防火墙设备上运行。在上面的拓扑中，VLAN 50 和 VLAN 60 内部网络的 IP 地址是 192.168.50.1 和 192.168.60.1。防火墙内部接口针对每种流量划分为多个子接口，正确的 VLAN 与子接口关联。此外，还有一个访问列表，用于指示哪些区域和设备可以相互通信。

连接到单臂模式下的 F5 负载均衡器，在与内部网络的负载均衡器相同的网络中连接负载均衡器网络。

图 30 显示枝叶-主干架构 VXLAN 拓扑。

图 30. 第 4-7 层服务的 VXLAN 拓扑的主干-枝叶架构

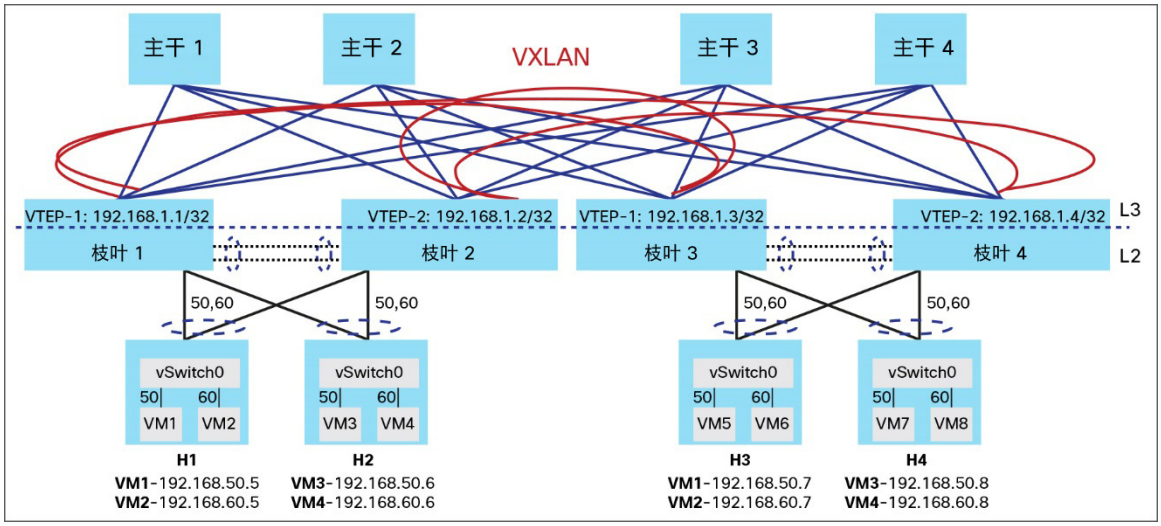
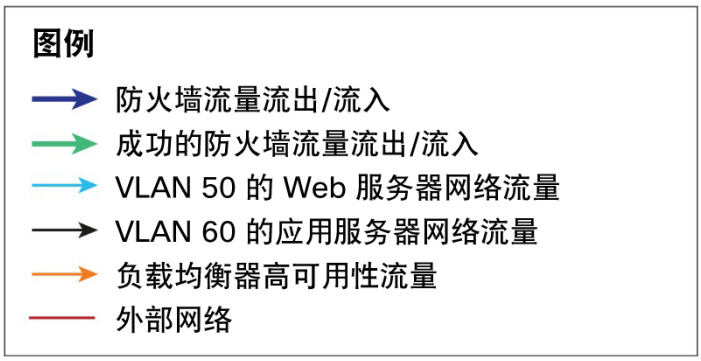


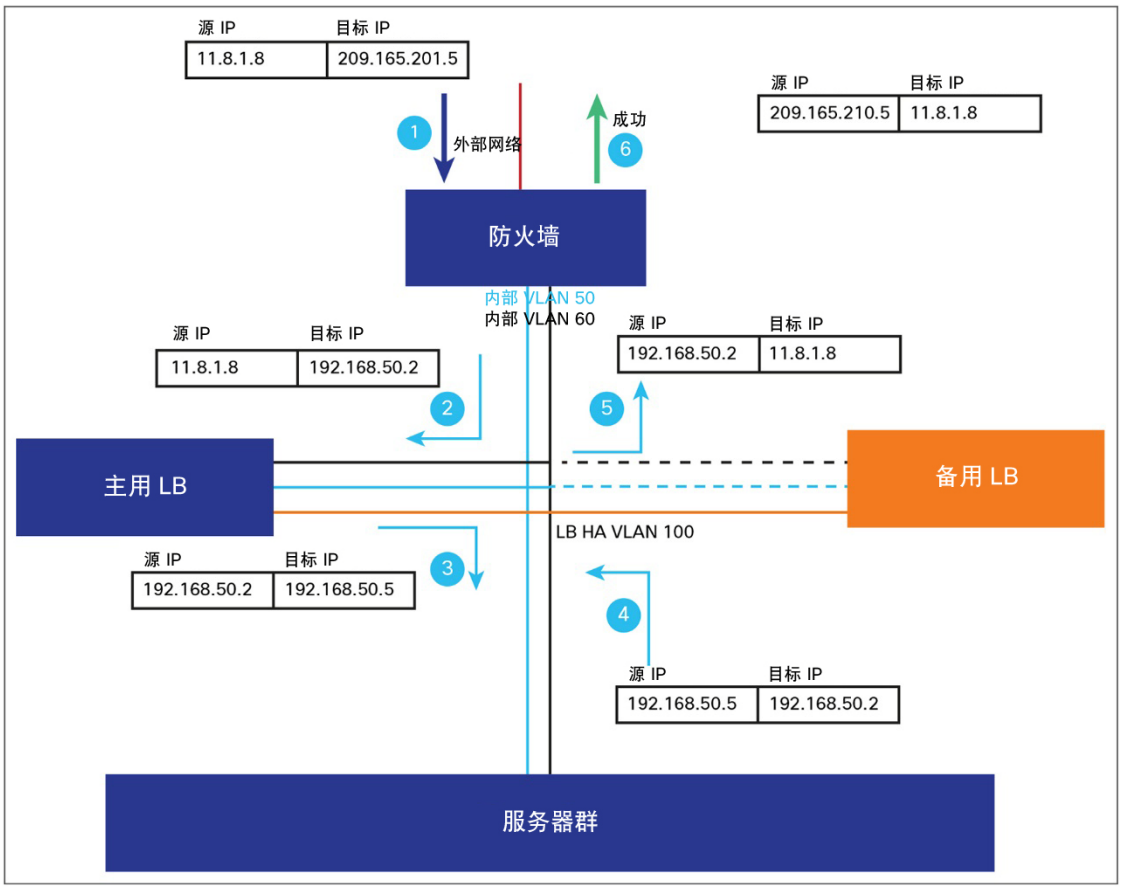
图 31 和图 32 说明了北-南流量和东-西流量的逻辑模型。

图例



在图 31 中，一个外部用户访问 Web 服务器。

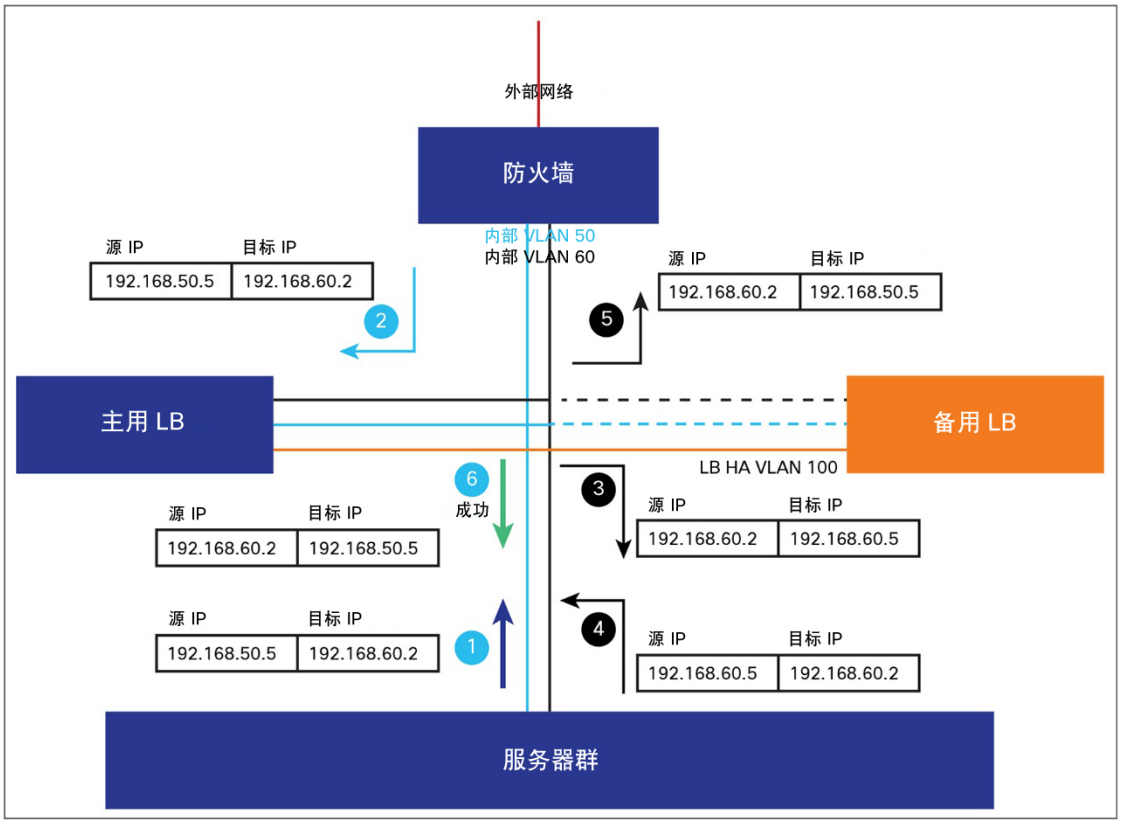
图 31. 北-南流量



1. 外部网络上的用户 (11.8.1.8) 使用全局目的地址 209.165.201.5 请求访问网页，该地址的网络位于防火墙的外部接口子网上。
2. 防火墙接收数据包，由于是新会话，因此安全设备会根据安全策略条款（访问列表、过滤器、AAA）验证数据包是否获得允许。防火墙将目的地址 (209.165.201.5) 转换为负载均衡器中存在的 Web 服务器的 VIP 地址 (192.168.50.2)。然后，防火墙将会话条目添加到快速路径，并转发到 Web 服务器子接口网络。
3. 数据包进入负载均衡器接口，该接口根据服务池状况为数据包提供服务。负载均衡器根据服务池状况，使用负载均衡器 VIP 地址的 IP 地址 (192.168.50.2) 和相应服务节点 IP 地址 (192.168.50.5) 的目标 NAT 进行源网络地址转换 (NAT)。然后，负载均衡器使用负载均衡器中创建的会话条目通过 Web 服务器接口转发数据包。
4. 节点 (192.168.50.5) 从负载平衡器回应请求。
5. 负载均衡器从建立的会话使用目的地址的客户端 IP 地址 (11.8.1.8) 执行反向 NAT。
6. 当数据包到达防火墙时，数据包会绕过许多与新连接相关联的查找，因为之前已建立该连接。防火墙执行反向 NAT，方法是使用位于防火墙外部网络的全局 IP 地址 (209.165.201.5) 转换源服务器 IP 地址 (192.168.50.2)。防火墙随后将数据包转发到外部客户端 (11.8.1.8)。

在图 32 中，Web 服务器用户访问应用服务器。

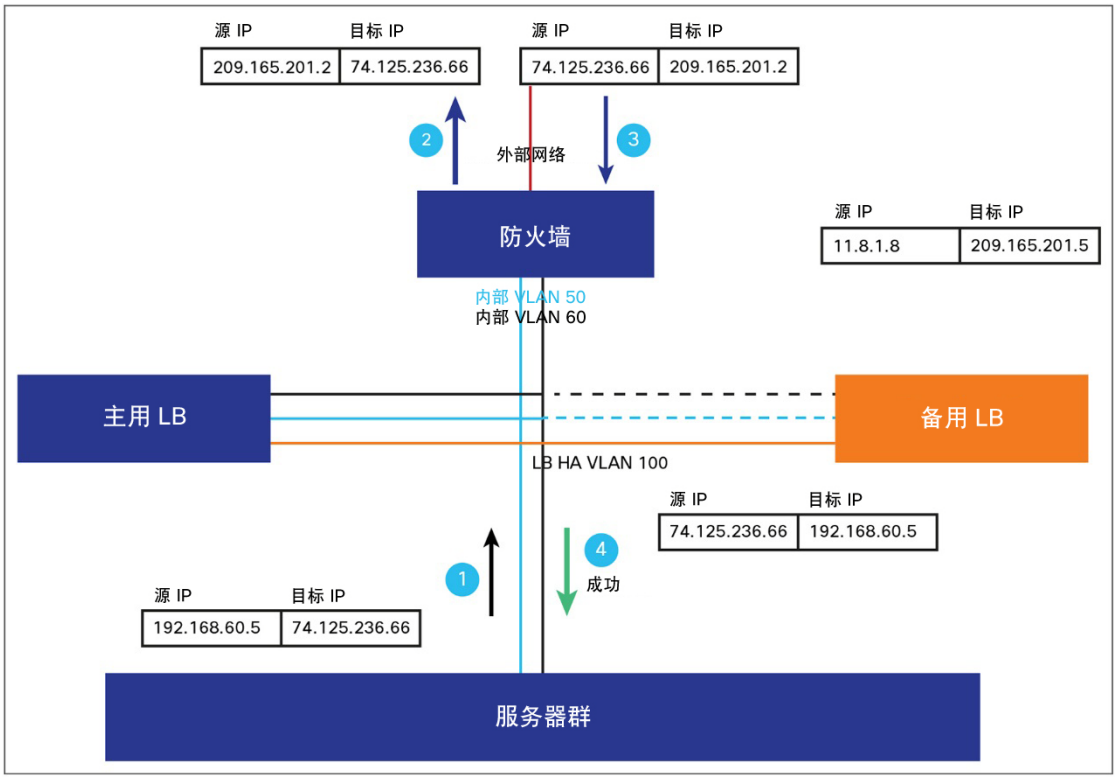
图 32. 东-西流量



1. VLAN 50 的 Web 服务器网络 (192.168.50.5) 上的用户对 VLAN 60 中目的地址为 192.168.60.2 的应用服务器提出请求，该目的地址是应用服务器的 VIP 地址。由于目的地址属于另一个子网，因此数据包到达 VLAN 50 网络的网关，它在防火墙的 Web 服务器子接口上（即 192.168.50.1）。
2. 防火墙接收数据包，由于是新会话，因此安全设备会根据安全策略条款（访问列表、过滤器、AAA）验证数据包是否获得允许。然后，防火墙会记录有关会话已建立的信息，并将数据包从 VLAN 60 的应用服务器子接口转发出去，因为根据 NAT 配置，目的 IP 在 VLAN 60 子接口上。
3. 由于数据包进入负载均衡器应用服务器 VIP，因此与该接口相关联的服务池根据服务池状况为数据包提供服务。负载均衡器使用 VIP 的 IP 地址和节点 IP 地址为 192.168.60.5 的应用服务器服务池的目标 NAT 进行源网络地址转换 (NAT)。然后，负载均衡器使用负载均衡器中创建的会话条目转发数据包。
4. 节点 (192.168.60.5) 从负载平衡器回应请求。
5. 负载均衡器从已建立的会话使用 VIP 192.168.60.2 的源 IP 和请求者 IP 192.168.50.5 的目的地址执行反向 NAT。
6. 当发往其他子网的数据包到达防火墙时，数据包会绕过许多与新连接相关联的查找，因为之前已建立该连接。由于目的 IP 在 Web 服务器网络上，因此数据包放置在 Web 服务器子接口中，然后防火墙将数据包转发给内部用户。

在图 33 中，Web 服务器用户访问外部互联网服务器。同一流程适用于访问外部互联网服务器的应用服务器用户。

图 33. 至外部互联网的流量



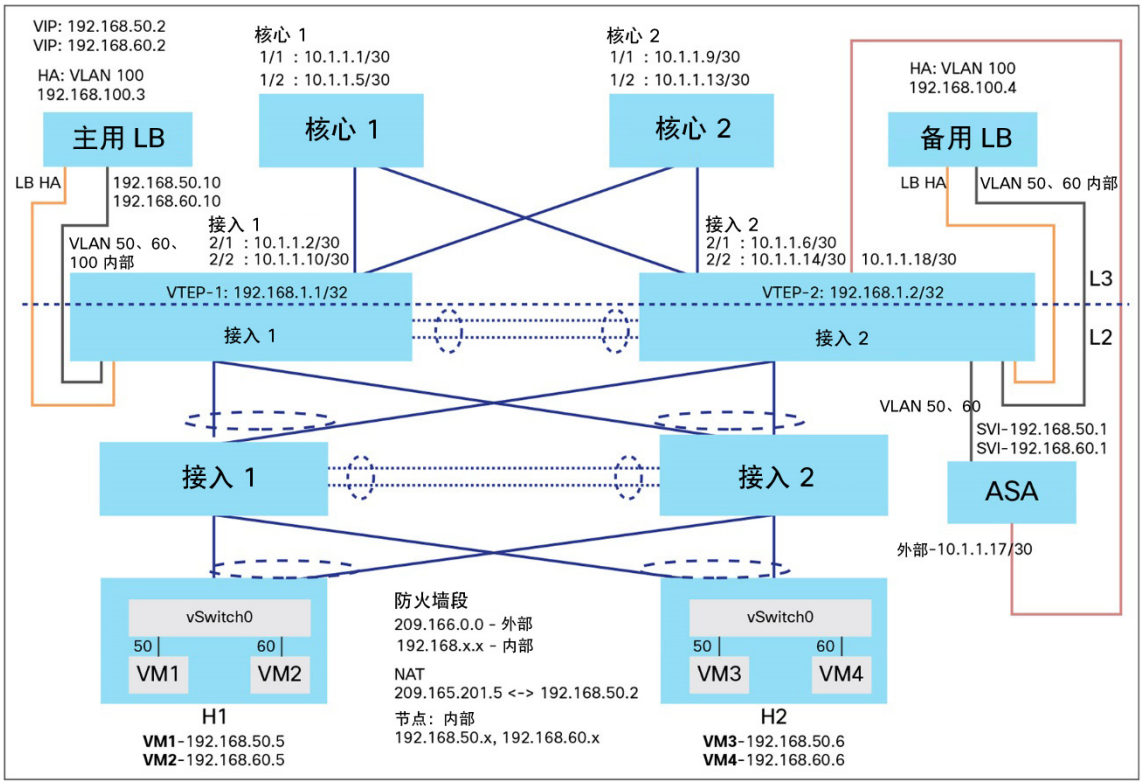
1. 内部网络上的用户从 http://www.google.com (74.125.236.66) 请求访问网页。
2. 防火墙接收数据包，因为其 Web 服务器子接口是 Web 服务器主机的默认网关。由于它是新会话，因此防火墙会根据安全策略条款（访问列表、过滤器、AAA）验证数据包是否获得允许。防火墙将本地源地址 (192.168.60.5) 转换为外部全局地址 209.165.201.2，该地址是外部接口地址。防火墙会记录有关会话已建立的信息，并从外部接口转发数据包。
3. 当 http://www.google.com 响应请求时，数据包会通过防火墙，由于已建立会话，因此数据包会绕过许多与新连接关联的查找。防火墙设备通过将全局目的地址 (209.165.201.2) 转换为本地用户地址 192.168.60.5 来执行 NAT，对于该本地用户地址，已存在一个已建立的现有连接。
4. 数据包从防火墙的 Web 服务器网络发回给 Web 服务器用户，防火墙将数据包转发给 Web 服务器用户。

9.3 传统的数据中心

9.3.1 拓扑

在图 34 概述的简单拓扑中，Web 服务器虚拟机连接到 VLAN 50，应用服务器虚拟机在 VLAN 60 中，IP 寻址方案分别为 192.168.50.x 和 192.168.60.x。服务器虚拟机连接到 vPC 模式下的接入交换机。接入交换机与汇聚交换机之间的连接通过 vPC 进行。

图 34. 传统的三层数据中心架构的第 4-7 层服务拓扑

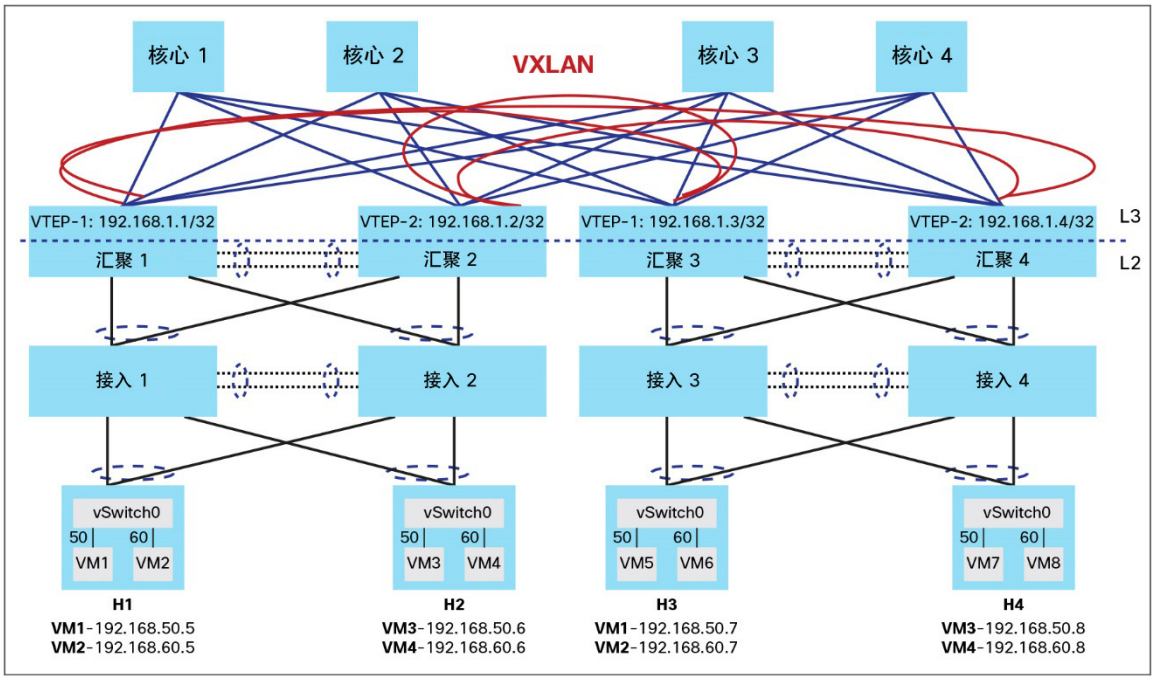


第 4-7 层服务调配发生在层中（图 34）。汇聚设备交换机端口与负载均衡器和防火墙设备连接，并且其路由端口连接到核心设备。使用每个汇聚设备中的环回接口创建 VTEP，以便进行 VXLAN 通信。

服务器虚拟机的默认网关在思科 ASA 5510 防火墙设备上运行。在上面的拓扑中，VLAN 50 和 VLAN 60 内部网络的 IP 地址是 192.168.50.1 和 192.168.60.1。防火墙内部接口针对每种流量划分为多个子接口，正确的 VLAN 与子接口关联。此外，还有一个访问列表，用于指示哪些区域和设备可以相互通信。

传统三层架构中的北-南流量和东-西流量的逻辑模型类似于枝叶-主干架构中的逻辑模型。

图 35. 第 4-7 层服务的 VXLAN 拓扑的传统数据中心架构



10. 管理交换矩阵

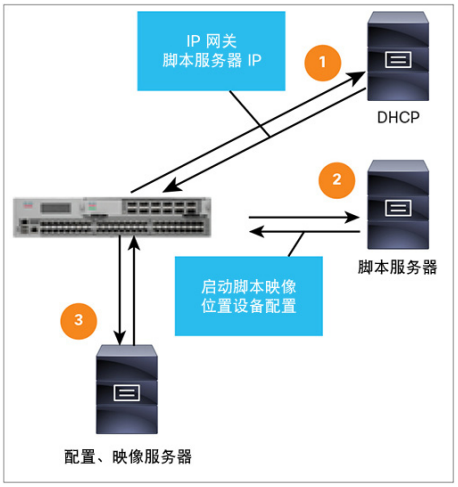
10.1.1 添加交换机和加电自动调配

在首次部署到网络中的 Cisco Nexus 交换机上，加电自动调配 (POAP) 会自动执行安装和升级软件映象以及安装配置文件的流程。

如果具有 POAP 功能的 Cisco Nexus 交换机在启动时未找到启动配置，交换机将进入 POAP 模式，查找域主机配置协议 (DHCP) 服务器，并以其接口 IP 地址、网关和域名称系统 (DNS) 服务器 IP 地址自行启动。交换机还获得简单文件传输协议 (TFTP) 服务器的 IP 地址或 HTTP 服务器的 URL，并下载一个配置脚本，该脚本可以使交换机能够下载并安装合适的软件映象和配置文件。

POAP 可以启用新 Cisco Nexus 9000 系列交换机的无触摸启动和配置，从而减少了执行耗时、容易出错的手动任务以扩展网络容量的需求（图 36）。

图 36. 使用 POAP 自动调配 Cisco Nexus 9000 系列



如果符合以下条件，将启用通电自动调配：

a. 启动配置不存在

B. 网络中存在的 DHCP 服务器响应所启动设备的 DHCP 发现消息

要在 Cisco Nexus 设备上使用 POAP 功能，请执行以下步骤：

1. 配置 DHCP 服务器以便为 Cisco Nexus 9000 交换机分配 IP 地址。图 37 显示了 Ubuntu DHCP 配置示例。

图 37. Ubuntu DHCP 服务器配置示例

```
DHCPARGS="eth0";
ddns-update-style interim;

allow booting;
allow bootp;

ignore client-updates;
set vendorclass = option vendor-class-identifier;

subnet 172.31.216.0 netmask 255.255.252.0 {

        default-lease-time      21600;
        max-lease-time          72000;

        option bootfile-name "poap.py";
        option tftp-server-name "172.31.216.138";

        option domain-name-servers      172.21.157.5,172.21.157.6;
        option routers 172.31.216.1;

        option broadcast-address 172.31.216.255;
        option subnet-mask      255.255.252.0;

        host leaf3 {
                hardware ethernet 88:f0:31:2b:a4:78;
                fixed-address 172.31.216.132;
        }

        host leaf2 {
                hardware ethernet 7c:69:f6:df:e8:a0;
                fixed-address 172.31.216.131;
        }

        host leaf1 {
                hardware ethernet 7c:69:f6:10:5a:b0;
                fixed-address 172.31.216.130;
        }
}
```

2. DHCP 配置中指定的 Poap.py 文件名存在于 https://github.com/datacenter/nexus9000/blob/master/nx-os/poap/poap.py。
3. 配置 TFTP 服务器并将 poap.py 文件放置在默认的 TFTP 服务器目录中。
4. 根据需要更改映像文件名。然后根据设置和需要在 poap.py 脚本中配置文件、IP 地址和凭证。
5. 将映象文件及映象的 .md5 放置在 TFTP 服务器中。可在任何 Linux 服务器上通过命令“md5sum <Image_file> <image_file>.md5”命令生成 Md5。
6. 启动设备。

图 38 显示了使用 POAP 执行 Cisco Nexus 交换机启动流程期间的输出示例。

图 38. 示例输出

```
Abort Auto Provisioning and continue with normal setup ?(yes/no)[n]: 2014 Dec 18 21:02:50 switch %$ VDC-1 %$ %POAP-2-POAP_DHCP
_DISCOVER_START: POAP DHCP Discover phase started
2014 Dec 18 21:02:59 switch %$ VDC-1 %$ %POAP-2-POAP_INFO: Using DHCP, information received over mgmt0 from 172.31.216.138
2014 Dec 18 21:02:59 switch %$ VDC-1 %$ %POAP-2-POAP_INFO: Assigned IP address: 172.31.216.132
2014 Dec 18 21:02:59 switch %$ VDC-1 %$ %POAP-2-POAP_INFO: Netmask: 255.255.252.0
2014 Dec 18 21:02:59 switch %$ VDC-1 %$ %POAP-2-POAP_INFO: DNS Server: 172.21.157.5
2014 Dec 18 21:02:59 switch %$ VDC-1 %$ %POAP-2-POAP_INFO: Default Gateway: 172.31.216.1
2014 Dec 18 21:02:59 switch %$ VDC-1 %$ %POAP-2-POAP_INFO: Script Server: 172.31.216.138
2014 Dec 18 21:02:59 switch %$ VDC-1 %$ %POAP-2-POAP_INFO: Script Name: poap.py
2014 Dec 18 21:03:09 switch %$ VDC-1 %$ %POAP-2-POAP_INFO: The POAP Script download has started
2014 Dec 18 21:03:09 switch %$ VDC-1 %$ %POAP-2-POAP_INFO: The POAP Script is being downloaded from [copy tftp://172.31.216.13
8/poap.py bootflash:scripts/script.sh vrf management ]
2014 Dec 18 21:03:10 switch %$ VDC-1 %$ %POAP-2-POAP_SCRIPT_DOWNLOADED: Successfully downloaded POAP script file
2014 Dec 18 21:03:10 switch %$ VDC-1 %$ %POAP-2-POAP_INFO: Script file size 18941, MD5 checksum f2f412844a905838d81ad0b52d1724
01
2014 Dec 18 21:03:10 switch %$ VDC-1 %$ %POAP-2-POAP_SCRIPT_STARTED_MD5_NOT_VALIDATED: POAP script execution started(MD5 not v
alidated)
2014 Dec 18 21:03:50 switch %$ VDC-1 %$ %POAP-2-POAP_SCRIPT_EXEC_SUCCESS: POAP script execution success
2014 Dec 18 21:03:50 switch %$ VDC-1 %$ %POAP-2-POAP_RELOAD_DEVICE: Reload device
2014 Dec 18 21:03:50 switch %$ VDC-1 %$ %PLATFORM-2-PFM_SYSTEM_RESET: Manual system restart from Command Line Interface
[  161.096020] [1418936639]  writing reset reason 9,
[ ?
CISCO SWITCH Ver7.17
Device detected on 0:6:0 after 0 msecs
Device detected on 0:1:1 after 0 msecs
```

10.1.2 软件升级

要升级接入层而不中断通过 vPC 双宿的主机，请执行以下步骤：

- 升级第一台 vPC 交换机（vPC 主交换机）。在此升级期间，将重新加载该交换机。该交换机重新加载后，服务器或下游交换机会检测与第一台交换机的连接的丢失情况，并开始将流量转发至第二台交换机（VPC 次要交换机）。
- 确认该交换机的升级已成功完成。升级完成后，交换机将恢复 vPC 对等和所有 vPC 链路。
- 升级第二台交换机。如果在第二台交换机上重复相同流程，将导致第二台交换机在升级过程中重新加载。在重新加载期间，第一台（已升级）交换机将转发所有进出服务器的流量。
- 确认第二个交换机的升级已成功完成。此升级结束时，完整的 vPC 对等关系已建立，并且整个接入层已升级。

10.1.3 Guest Shell 容器

从思科 NX-OS 软件版本 6.1(2)I3(1) 开始，Cisco Nexus 9000 系列设备支持访问名为“Guest Shell”的分离执行空间。在 Guest Shell 内，网络管理员被授予 Bash 访问权限，并且可使用熟悉的 Linux 命令管理交换机。Guest Shell 环境：

- 可访问网络，包括思科 NX-OS 软件已知的所有 VRF
- 具有托管 Cisco Nexus 9000 Bootflash 所需的读取和写入访问权限
- 能够执行 Cisco Nexus 9000 命令行界面 (CLI)
- 可访问 Cisco onePK™ API
- 能够开发、安装和运行 Python 脚本
- 能够安装和运行 64 位 Linux 应用
- 具有在系统重新加载或切换期间持续存在的根文件系统

将执行空间与本地主机系统分离，可以自定义 Linux 环境，这样能满足应用的需求，而不会影响主机系统的完整性。Guest shell 文件系统内安装或修改的应用、库或脚本与主机的这些内容分离。

默认情况下，当备用管理引擎使用该管理引擎上提供的 guest shell 软件包过渡到主用角色时，会全新安装 guest shell。借助双管理引擎系统，网络管理员可以利用 guest shell 同步命令，该命令将 guest shell 内容从主用管理引擎同步到备用管理引擎。

有关更多详情，请参阅《Cisco Nexus 9000 系列 NX-OS 可编程性指南 6.x 版》：
http://www.cisco.com/c/en/us/td/docs/switches/datacenter/nexus9000/sw/6-x/programmability/guide/b_Cisco_Nexus_9000_Series_NX-OS_Programmability_Guide.pdf。

11. 虚拟化和云协调

11.1 虚拟机跟踪器

Cisco Nexus 系列交换机提供了虚拟机跟踪器功能，它使交换机可以与最多四个 VMware vCenter 连接进行通信。在创建、删除或移动虚拟机时，虚拟机跟踪器会在服务器的接口上动态配置 VLAN。

目前，VMware vCenter 的 ESX 5.1 和 ESX 5.5 版本支持虚拟机跟踪器功能。它支持每个主机最多 64 个虚拟机，所有 vCenter 最多 350 个主机。它在 MST 模式下最多支持 600 个 VLAN，在 PVRST 模式下最多支持 507 个 VLAN。

虚拟机跟踪器的当前版本依靠思科发现协议将主机与交换机端口相关联。

以下是可以使用虚拟机跟踪器执行的一些操作：

- **启用虚拟机跟踪器 -** 默认情况下，会为交换机的所有接口启用虚拟机跟踪器。您可以根据需要通过使用 [no] vmtracker enable 命令在特定接口上禁用和启用它。
- **创建与 VMware vCenter 的连接**。
- **配置动态 VLAN 连接 -** 默认情况下，虚拟机跟踪器会跟踪来自 VMware vCenter 的所有异步事件，并立即更新交换机端口配置。您还可以根据需要配置一种同步机制，按指定的时间间隔将所有主机、虚拟机和端口组信息自动与 VMware vCenter 同步。
- **启用动态创建 VLAN -** 默认情况下全局启用动态创建和删除 VLAN。在启用动态创建 VLAN 后，如果将虚拟机从一台主机移至另一台主机，并且该虚拟机所需的 VLAN 在该交换机上不存在，则会自动在该交换机上创建所需的 VLAN。您也可以禁用此功能。但是，如果您禁用动态创建 VLAN，则必须手动创建所有必需的 VLAN。

- **VPC 兼容性检查 -** 在 VPC 中，在两个对等体之间 vCenter 连接需要相同。虚拟机跟踪器提供了一个命令来检查两个对等体之间的 vPC 兼容性。

如果使用虚拟机跟踪器，则用户无法执行与交换机端口和 VLAN 相关的任何第 2 层或第 3 层配置，除非是为了更新本地 VLAN。

虚拟机跟踪器配置

```
Leaf1(config)# feature vmtracker
Leaf1(config)# vmtracker connection vCenter_conn1
Leaf1(config)# remote ip address 172.31.216.146 port 80 vrf management
Leaf1(config)# username root password vmware
Leaf1(config)# connect
Leaf1(config)# set interval sync-full-info 120

Leaf2(config)# feature vmtracker
Leaf2(config)# vmtracker connection vCenter_conn1
Leaf2(config)# remote ip address 172.31.216.146 port 80 vrf management
Leaf2(config)# username root password vmware
Leaf2(config)# connect
Leaf2(config)# set interval sync-full-info 120
```

表 4 列出了虚拟机跟踪器配置的相关显示命令。

表 4. 显示命令

show vmtracker status	提供 vCenter 的连接状态
show vmtracker info detail	提供有关与本地接口相关联的虚拟机属性的信息

有关详细信息，请参阅《Cisco Nexus 9000 系列 NX-OS 虚拟机跟踪器配置指南 6.x 版》：
http://www.cisco.com/c/en/us/td/docs/switches/datacenter/nexus9000/sw/6-x/vm_tracker/configuration/guide/b_Cisco_Nexus_9000_Series_NX-OS_Virtual_Machine_Tracker_Configuration_Guide/b_Cisco_Nexus_9000_Series_Virtual_Machine_Tracker_Configuration_Guide_chapter_011.html。

11.2 OpenStack

Cisco Nexus 9000 系列包括对 OpenStack Networking 的 Cisco Nexus 插件 (Neutron) 的支持。该插件使客户可以使用行业的领先网络平台轻松构建其基础设施即服务 (IaaS) 网络，以熟悉的可管理性和可控性提供性能、可扩展性和稳定性。该插件有助于为云网络部署带来操作简便性。OpenStack 构建按需自我服务、多租户的计算基础设施的能力众所周知。但是，跨虚拟和物理基础设施实施 OpenStack VLAN 网络模型可能很困难。

OpenStack Networking 提供了支持插件直接配置网络的可扩展架构。但是，每个网络插件仅启用该插件的目标技术的配置。在 VLAN 的多台主机上运行 OpenStack 集群时，典型的插件将配置虚拟网络或物理网络，而非同时配置两者。

Cisco Nexus 插件支持同时使用多个插件，以此来解决上述问题。除了标准的 Open vSwitch (OVS) 插件外，典型的部署还运行 Cisco Nexus 插件。Cisco Nexus 插件接受 OpenStack Networking API 调用，并直接配置 Cisco Nexus 交换机以及在虚拟机监控程序上运行的 OVS。Cisco Nexus 插件不仅在物理和虚拟网络上配置 VLAN，而且还智能地分配 VLAN ID，在不再需要 VLAN ID 时将其取消调配，并在可能的情况下将其重新分配给新租户。系统将配置 VLAN，以便在属于同一租户网络的不同虚拟化（计算）主机上运行的虚拟机通过物理网络透明地进行通信。此外，系统对从计算主机到物理网络的连接进行中继，以便只允许流量来自虚拟交换机在主机上配置的 VLAN（图 39）。

表 5 概述了网络管理员遇到的各种挑战以及 OpenStack 如何解决这些挑战。

图 39. 支持 Cisco Nexus 9000 系列交换机的 Cisco OpenStack Neutron 插件

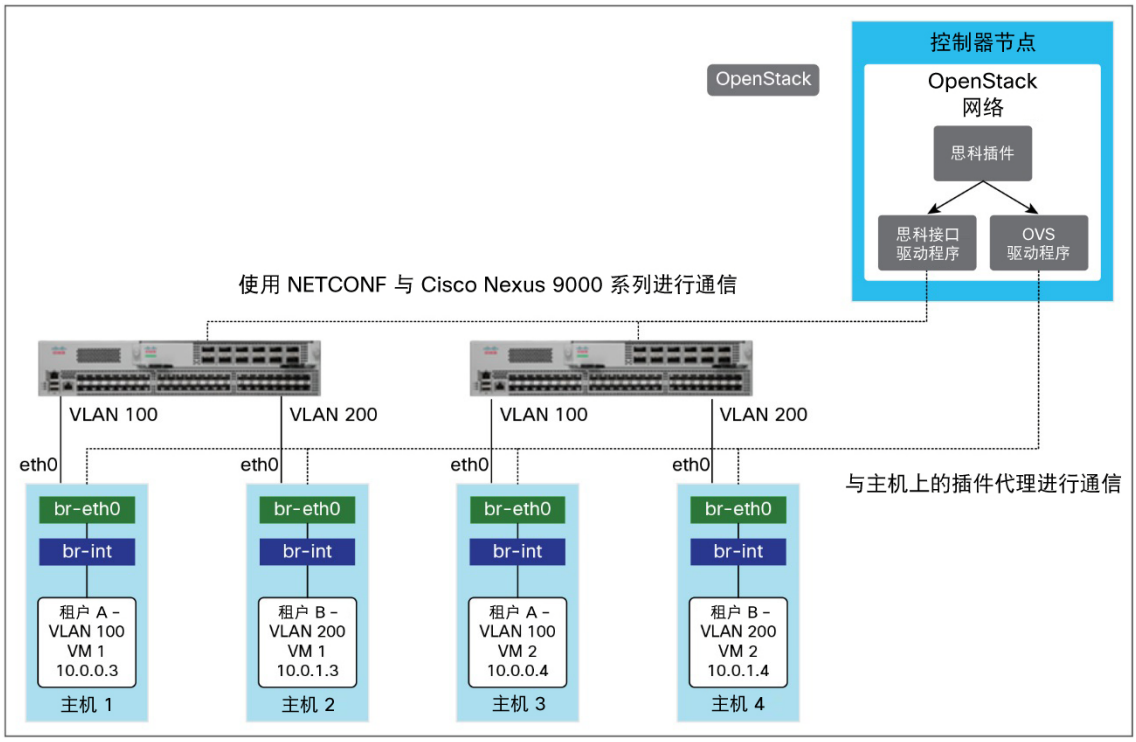


表 5. OpenStack Networking 的 Cisco Nexus 插件

要求	挑战	思科插件解决方法
将租户 VLAN 扩展到虚拟主机	必须在物理和虚拟网络上配置 VLAN。OpenStack 一次仅支持一个插件。操作员必须选择要手动配置网络的哪些部分	接受 OpenStack API 调用并配置物理和虚拟交换机
高效利用有限的 VLAN ID	在每台交换机上静态调配 VLAN ID 会快速消耗所有可用的 VLAN ID，从而限制可扩展性并使网络更易遭受广播风暴攻击	高效利用有限的 VLAN ID，方法是在创建和破坏网络时在交换机上调配和取消调配 VLAN
在架顶式 (ToR) 交换机中轻松配置租户 VLAN	操作员需要在所有物理交换机上静态调配所有可用的 VLAN，这是一个手动且容易出错的过程	在通过 Cisco Nexus 插件驱动程序连接到虚拟化主机的交换机端口上，动态调配租户网络特定的 VLAN
智能分配 VLAN ID	连接到虚拟化主机的交换机端口配置为处理所有 VLAN，因此很快就会达到硬件限制	仅为与主机上配置的网络相对应的 VLAN 配置连接到虚拟化主机的交换机端口，实现准确的端口与 VLAN 关联
大型、多机架的部署的汇聚交换机 VLAN 配置	当计算主机在多个机架中运行时，ToR 交换机需要形成全网状，或者需要手动对汇聚交换机进行中继	支持 Cisco Nexus 2000 系列交换矩阵扩展器以启用大型、多机架的部署，并消除对汇聚交换机 VLAN 配置的需要

11.3 Cisco UCS Director

Cisco UCS Director 是一个非常强大的集中管理和协调工具，它可以让小型商业公司 IT 员工的日常操作变得非常轻松。通过 UCS Director 提供的自动化功能，少量 IT 员工即可加速交付新服务和应用。Cisco UCS Director 将硬件和软件抽象化为可编程任务，并充分利用工作流程设计器，使管理员只需将任务拖放到工作流程中即可提供必要的资源。图 40 描绘了 UCS Director 工作流程示例。

图 40. Cisco UCS Director 工作流程示例

Cisco UCS Director 允许您将通常从 UCS 管理器 GUI 手动执行的许多常见任务自动化。图 41 列出了可使用 UCS Director 自动执行的许多任务。

图 41. 常见的 UCS Director 自动化任务

1. 选择 UCS 服务器	21. 修改 UCS 引导策略 WWPN
2. 创建 UCS 服务器池	22. 创建 VLAN 组
3. 删除 UCS 服务器池	23. 删除 UCS VLAN 组
4. 将服务器添加到 UCS 服务器池	24. 修改 UCS VLAN/VLAN 组组织权限
5. 从 UCS 服务器池中删除服务器	25. 服务器维护
6. 关联 UCS 服务配置文件模板	26. 重新确认服务器插槽
7. 重置 UCS 服务器	27. 添加 VLAN
8. 打开 UCS 服务器电源	28. 删除 UCS 引导策略
9. 关闭 UCS 服务器电源	29. 删除 UCS VLAN
10. 通过模板创建 UCS 服务配置文件	30. 将 VLAN 添加到服务配置文件
11. 创建 UCS 服务配置文件	31. 从服务配置文件中删除 VLAN
12. 选择 UCS 服务配置文件	32. 从服务配置文件中添加 iSCSI vNIC
13. 修改 UCS 服务配置文件引导策略	33. 从服务配置文件中删除 ISCSI vNIC
14. 删除 UCS 服务配置文件	34. 从服务配置文件中添加 vNIC
15. 关联 UCS 服务配置文件	35. 从服务配置文件中删除 vNIC
16. 取消关联 UCS 服务器	36. 创建服务配置文件 iSCSI 引导策略
17. 取消关联 UCS 服务配置文件	37. 将服务配置文件引导策略修改为从 iSCSI 引导
18. 创建 UCS 引导策略	38. 从服务配置文件 vNIC 中删除 VLAN
19. 修改 UCS 引导策略 LUN ID	39. 将 VLAN 添加到 vNIC 模板
20. 克隆 UCS 引导策略	40. 从 vNIC 模板中删除 VLAN

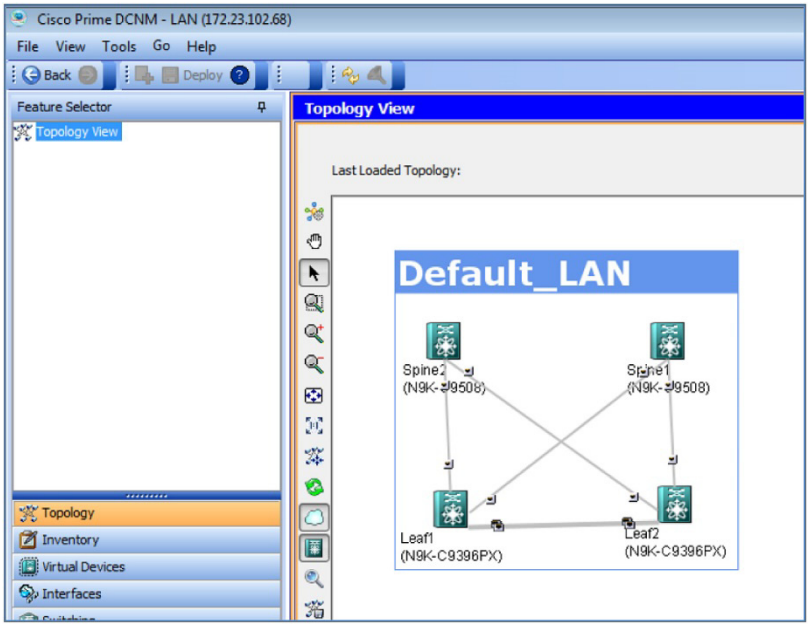
有关详细信息，请阅读 Cisco UCS Director 解决方案概述。

11.4 Cisco Prime 数据中心网络管理器

Cisco Prime 数据中心网络管理器 (DCNM) 是一个非常强大的工具，用于思科数据中心计算、网络和存储基础设施的集中数据中心监控、管理和自动化。DCNM 的基本版本免费提供，更高级的功能需要许可证。DCNM 允许集中管理所有 Cisco Nexus 交换机以及思科 UCS 和 MDS 设备。

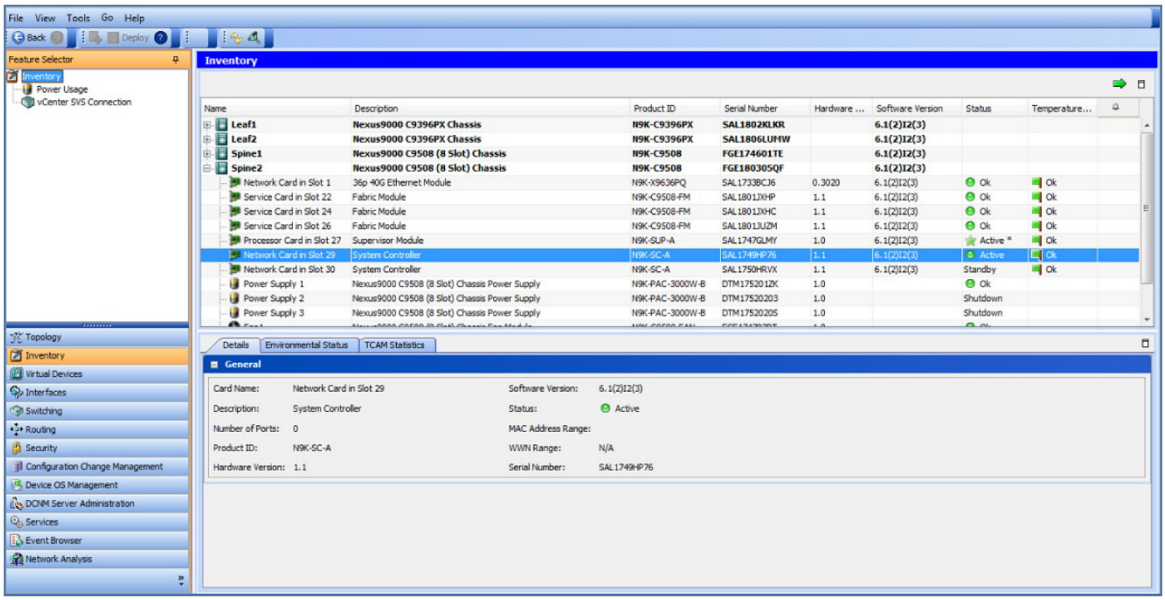
基础设施管理的设计示例中使用了 DCNM-LAN 6.x 版。DCNM 的一项强大功能是能够具有可视的自动更新拓扑图。图 42 显示了实验室拓扑示例。

图 42. DCNM 拓扑图



DCNM 还可用于管理高级功能，例如 vPC、思科 NX-OS 映像管理和资产控制。可视资产示例拍摄自其中一台主干交换机（图 43）。

图 43. DCNM 交换机资产



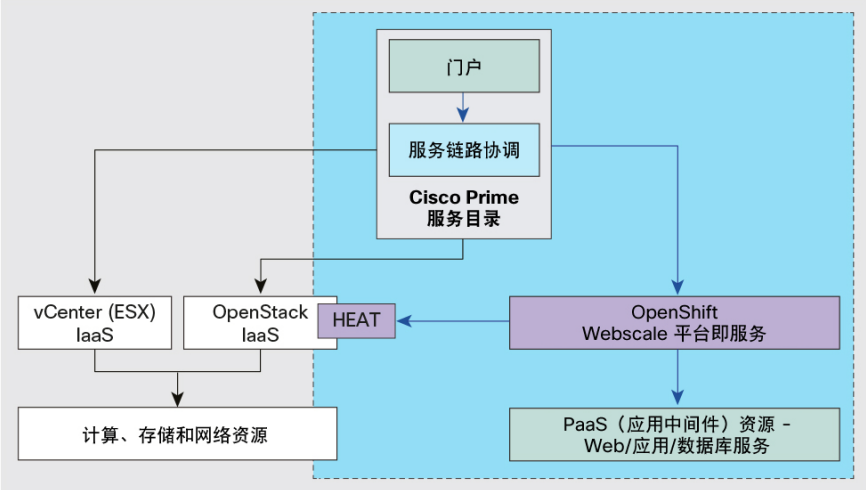
11.5 Cisco Prime 服务目录

随着组织继续扩大自动化，自助服务门户是一个非常重要的组件，用于为消费者使用基础设施或应用提供门店或市场视图。

Cisco Prime 服务目录提供以下主要功能：

- 调配、配置和管理 IaaS、PaaS 及其他 IT 服务的单一平台（图 44）
- IaaS、PaaS 及其他 IT 服务的单点登录 (SSO)；无需登录到多个系统。
- IT 管理员可以用图形方式轻松创建服务目录，以将应用元素与业务策略和监管相结合
- 通过 API 与各种基础设施和各种 IT 子系统集成，以促进信息交流以及调配基础设施和应用。示例包括 OpenStack、Cisco UCS Director、Red Hat Openshift PaaS、vCenter 等
- 创建门店，以便应用开发人员和 IT 项目经理可以通过直观的用户界面浏览和请求可用的应用堆栈

图 44. Cisco Prime 服务目录与 PaaS 和 IaaS 平台的集成示例



12. 自动化和可编程性

12.1 支持传统网络功能

虽然已讨论许多新功能和重要功能，但是有必要强调对 Cisco Nexus 9000 系列交换机提供的其他功能支持，例如服务质量 (QoS)、组播以及简单网络管理协议 (SNMP) 和支持的 MIB。

服务质量

新应用和不断发展的协议正在改变现代数据中心的 QoS 要求。高频率交易大厅应用对延迟非常敏感，而高性能计算的特征通常是突发、多对一的东-西流量。任何数据中心内的存储、语音和视频等应用也需要特殊且不同的处理。

像 Cisco Nexus 系列的其他成员一样，Cisco Nexus 9000 系列交换机通过思科模块化 QoS CLI (MQC) 支持 QoS 配置。配置涉及根据协议或数据包报头标记等信息使用类映射来识别流量，定义如何通过策略映射来处理不同的类，可能采用的方法是标记数据包、排队或调度，然后使用服务策略命令将策略映射应用于接口或整个系统。QoS 在默认情况下处于启用状态，且不需要许可证。

与 Cisco IOS 软件不同，Cisco Nexus 9000 交换机上的思科 NX-OS 使用三种不同的策略映射类型，具体取决于您尝试实施的 QOS 功能。思科 NX-OS QoS 的三种类型及其主要用途包括：

- QoS 类型 - 用于分类、标记和策略管制
- 排队类型 - 用于缓存、排队和调度
- 网络 QoS 类型 - 用于系统范围的设置、拥塞控制和暂停行为

网络-QOS 类型策略映射仅在系统范围内应用，而 QoS 类型和排队类型可以同时应用在单个接口上，一种类型应用于每个入口和出口，每个接口总共四个 QoS 策略（如果需要）。

组播

在启用组播路由的情况下，Cisco Nexus 9300 平台以最多 8000 个组播路由的规模低延迟地提供高性能、线速的第 2 层和第 3 层组播吞吐量。Cisco Nexus 9300 平台通过为第 2 层和第 3 层组播执行基于 IP 的查找来优化组播查找和复制，以避免出现组播 IP 至 MAC 编码中常见的别名设置问题。

Cisco Nexus 9300 平台支持互联网组管理协议 (IGMP) 版本 1、2 和 3，协议无关组播 (PIM) 稀疏模式、任播交汇点和组播源发现协议 (MSDP)。

SNMP 和支持的 MIB

SNMP 提供标准框架和语言来管理网络上的设备，包括 Cisco Nexus 9000 系列交换机。SNMP 管理器使用 SNMP 控制和监控网络设备的活动。SNMP 代理在受管设备中运行，并将数据报告给管理系统。Cisco Nexus 9000 系列交换机上的代理必须配置为与管理器通信。MIB 是 SNMP 代理的受管对象集合。Cisco Nexus 9000 系列交换机支持 SNMP v1、v2c 和 v3。交换机还可以支持基于 IPv6 的 SNMP。

SNMP 会生成关于 Cisco Nexus 9000 系列交换机的通知并发送给管理器，例如在邻近路由器丢失时通知管理器。陷阱通知从 Cisco Nexus 9000 交换机上的代理发送给管理器，管理器未确认消息。用于告知的通知由管理器确认。表 6 提供了默认情况下启用的 SNMP 陷阱。

在 Cisco Nexus 9000 交换机上，思科 NX-OS 支持无状态重新启动，并且也知道虚拟路由和转发 (VRF)。使用 SNMP 不需要许可证。

如需 Cisco Nexus 9000 系列交换机上支持的 MIB 列表，请参阅此列表。

如需获得配置帮助，请参阅《Cisco Nexus 9000 系列 NX-OS 系统管理配置指南 6.0 版》的“配置 SNMP”部分。

表 6. Cisco Nexus 9000 系列交换机上默认情况下启用的 SNMP 陷阱

陷阱类型	说明
generic	: coldStart
generic	: warmStart
entity	: entity_mib_change
entity	: entity_module_status_change
entity	: entity_power_status_change
entity	: entity_module_inserted
entity	: entity_ module_removed
entity	: entity_unrecognised_module
entity	: entity_fan_status_change
entity	: entity_power_out_change
link	: linkDown
link	: linkup
link	: extended-linkDown
link	: extended-linkUp
link	: cieLinkDown
link	: cieLinkUp
link	: delayed-link-state-change
rf	: redundancy_framework
license	: notify-license-expiry
license	: notify-no-license-for-feature
license	: notify-licensefile-missing
license	: notify-license-expirty-warning
upgrade	: UpgradeOpNotifyOnCompletion
upgrade	: UpgradeJobStatusNotify
rmon	: risingAlarm
rmon	: fallingAlarm
rmon	: hcRisingAlarm
rmon	: hcFallingAlarm
entity	: entity_sensor

12.2 通过 NX-API 对 Cisco Nexus 9000 交换机进行编程

Cisco NX-API 允许对 Cisco Nexus 9000 系列平台进行基于 HTTP 的编程访问。此支持由开源 Web 服务器 NX-API 提供。NX-API 通过基于 Web 的 API 提供思科 NX-OS CLI 的配置和管理功能。设备可以设置为以 XML 或 JSON 格式发布 API 调用的输出。此 API 支持在 Cisco Nexus 9000 系列平台上快速开发。

本部分提供了 API 配置，用于实现先前通过 CLI 执行（如上一部分所述）的功能。配置详细信息作为第 14.2.3 节（NX-API 沙盒）的一部分提供。

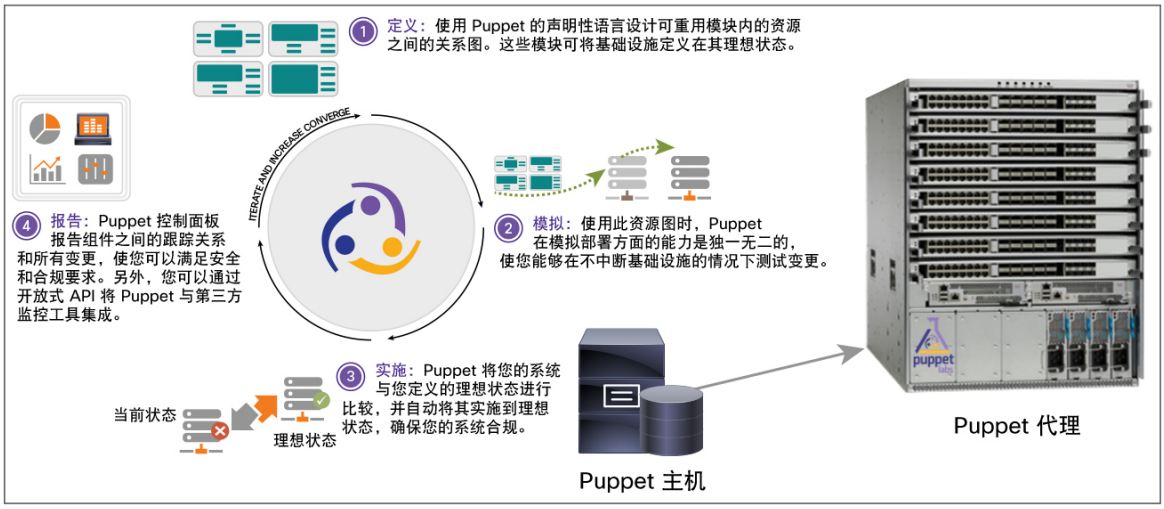
12.3 Chef、Puppet 和 Python 集成

Puppet 和 Chef 是两种流行的、基于意图的基础设施自动化框架。Chef 使用户可以通过菜谱（一组可重复使用的配置或管理任务）定义其意图，并允许将菜谱部署在众多设备上。菜谱在部署到 Cisco Nexus 9000 系列交换机后，会转换为网络配置设置以及用于收集统计数据和分析信息的命令。使用菜谱可以自动配置和管理 Cisco Nexus 9000 系列交换机。

Puppet 提供相似的意图定义结构，其名为清单。清单在部署到 Cisco Nexus 9000 系列交换机后，会转换为网络配置设置以及用于从交换机收集信息的命令。

Puppet 和 Chef 在基础设施自动化和 DevOps 社区中得到广泛部署且备受关注。Cisco Nexus 9000 系列支持 Puppet 和 Chef 框架，Puppet 和 Chef 的客户端集成到交换机上的增强型思科 NX-OS（图 45）。

图 45. 支持在 Cisco Nexus 9000 系列上通过 Puppet 实现自动化



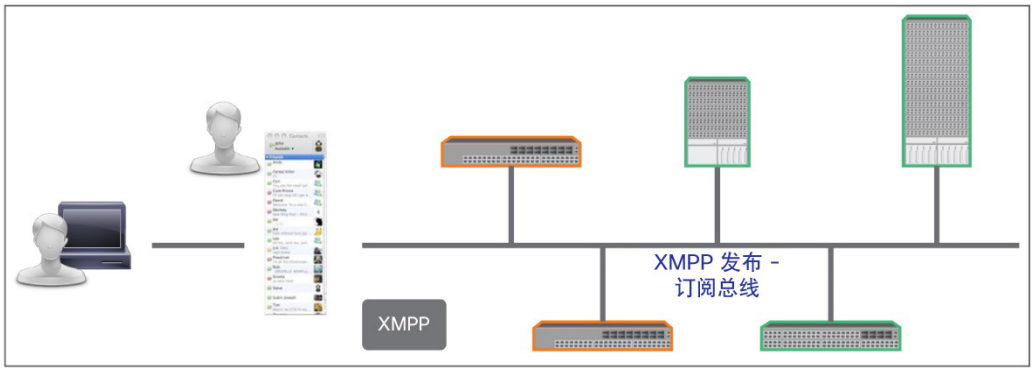
12.4 可扩展消息传送和网真协议支持

Cisco Nexus 9000 系列交换机上的增强型思科 NX-OS 将可扩展消息传送和网真协议 (XMPP) 客户端集成到操作系统。此集成使得可以由 XMPP 启用的聊天客户端管理和配置 Cisco Nexus 9000 系列交换机，这些客户端常用于人际交流。XMPP 支持可能启用多项有用的功能：

- **组配置 -** 将一组 Cisco Nexus 9000 系列设备添加到聊天组，并以组的形式管理一组 Cisco Nexus 9000 系列交换机。此功能有助于将通用配置推送到一组 Cisco Nexus 9000 系列设备，而不是逐个配置设备。
- 单一管理点 - XMPP 服务器可以充当单一管理点。用户通过单个 XMPP 服务器的身份验证，然后获得该服务器上注册的所有设备的访问权限。
- **安全 -** XMPP 接口支持基于角色的访问控制 (RBAC)，并帮助确保用户只能运行他们有权运行的命令。

- **自动化 -** XMPP 是一个基于标准的开放接口。脚本和管理工具可以使用此接口自动管理 Cisco Nexus 9000 系列设备（图 46）。

图 46. 支持在 Cisco Nexus 9000 系列上通过 XMPP 实现自动化



12.5 OpenDayLight 集成和 OpenFlow 支持

Cisco Nexus 9000 系列将支持与思科推崇的开源 OpenDaylight 项目集成（图 47）。OpenDaylight 在某些用户群体中越来越受欢迎，因为它可以满足一些基础设施要求。

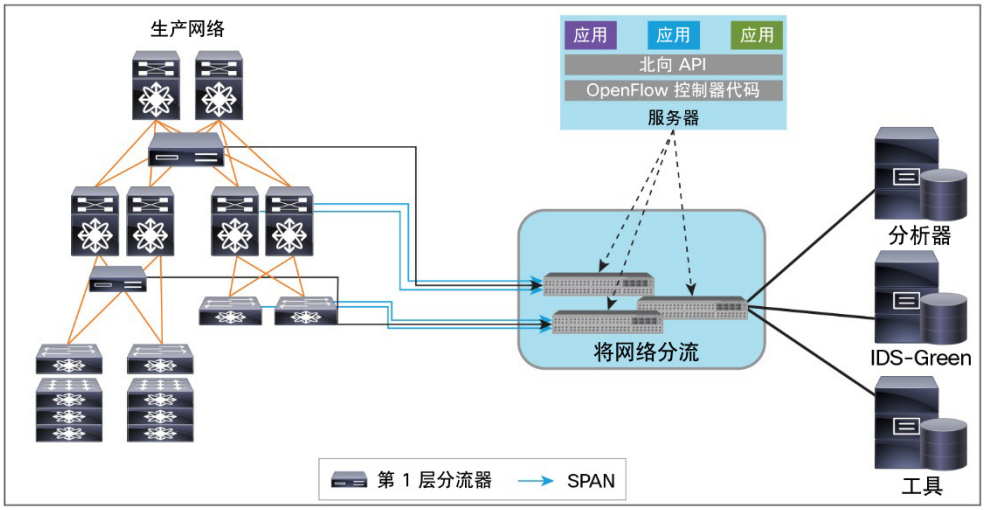
- 操作员希望以经济实惠的方式实时协调和操作集成的虚拟计算、应用和网络资源。
- 应用程序开发人员希望网络使用单个简单的接口。基础详细信息（例如路由器、交换机或拓扑）可能是他们希望抽象化和简化的分心之事。

Cisco Nexus 9000 系列将通过精心发布的全面接口（例如 Cisco onePK）与 OpenDaylight 控制器集成。

图 47. OpenDaylight：未来支持 Cisco Nexus 9000 系列

Cisco Nexus 9000 还将支持 OpenFlow 以帮助启用使用案例，例如网络 Tap 汇聚（图 48）。

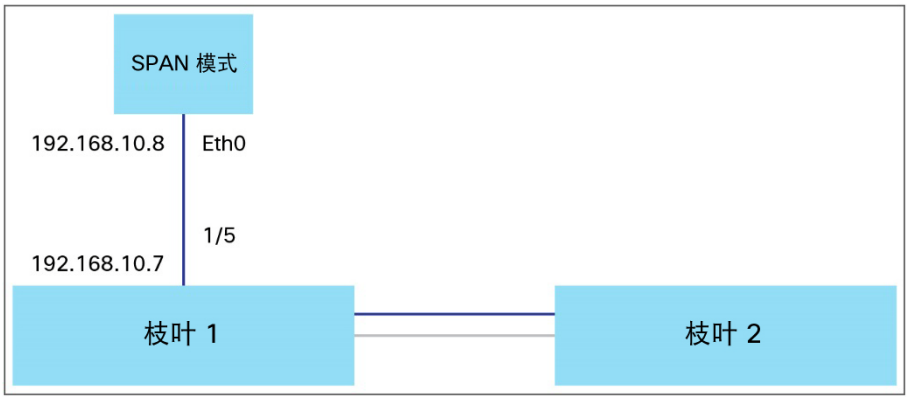
图 48. 在 Cisco Nexus 9000 系列上利用 OpenFlow 支持进行 Tap 汇聚



13. 故障排除

远程封装交换端口分析器 (ERSPAN) 监控一个或多个源端口上的流量，并向另一台交换机上的一个或多个目标端口提供镜像流量（图 49）。流量采用通用路由封装 (GRE) 协议封装，因此可跨第 3 层网络在源交换机与目标交换机之间路由。

图 49. ERSPAN



ERSPAN 复制给定交换机来源的入口和出口，并创建返回到 ERSPAN 目标的 GRE 隧道。这使网络操作员可以将网络监控设备战略性地部署在网络的中心位置。然后，管理员可以非常详细地收集历史流量模式。

ERSPAN 可以对网络上的多台交换机启用远程监控。它将镜像流量从不同交换机的源端口传输到网络分析仪已连接的目标端口。监控源端口的所有接收（入口）、传输（出口）或双向（两者）的数据包。ERSPAN 源包含源端口、源 VLAN 或源 VSAN。当 VLAN 指定为 ERSPAN 源时，该 VLAN 中的所有支持的接口都是 ERSPAN 源。ERSPAN 还可用于数据包监控。

ERSPAN 配置

```
!配置 ERSPAN

 Leaf1# Configure terminal
 Leaf1(config)# Interface Ethernet2/5
 Leaf1(config-if)# ip address 192.168.10.7/24
 Leaf1(config-if)# ip router eigrp 1
 Leaf1(config-if)# no shut

 Leaf1(config)# monitor session 1 type erspan-source
 Leaf1(config)# erspan-id 1
 Leaf1(config)# vrf default
 Leaf1(config)# destination ip 192.168.10.8
 Leaf1(config)# source interface Ethernet2/1 rx
 Leaf1(config)# source interface Ethernet2/2 rx
 Leaf1(config)# source vlan 50,60 rx
 Leaf1(config)# no shut
 Leaf1(config)# monitor erspan origin ip-address 172.31.216.130 global

 Leaf2(config)# monitor session 1 type erspan-source
 Leaf2(config)# erspan-id 1
 Leaf2(config)# vrf default
 Leaf2(config)# destination ip 192.168.10.8
 Leaf2(config)# source interface Ethernet2/1 rx
 Leaf2(config)# source interface Ethernet2/2 rx
 Leaf2(config)# source vlan 50,60,100 rx
 Leaf2(config)# no shut
 Leaf2(config)# monitor erspan origin ip-address 172.31.216.131 global
```

显示命令

show monitor session all	提供有关 ERSPAN 连接的详细信息

Ethanalyzer 是基于 Wireshark（以前称为 Ethereal）的思科 NX-OS 协议分析工具。是捕获和解码数据包的 Wireshark 的命令行版本。Ethanalyzer 是对控制平面和发往交换机 CPU 的流量进行故障排除的实用工具。使用管理接口对影响 mgmt0 接口的数据包进行故障排除。Ethanalyzer 使用与 tcpdump 相同的捕获过滤器语法和 Wireshark 语法的显示过滤器语法。

会传送所有与日志 ACE 相匹配的数据包（具有速率限制）。使用捕获和显示过滤器可仅查看与日志 ACE 相匹配的流量的子集。

```
Leaf2# ethanalyzer local interface inband capture-filter “tcp port 80”
Leaf2# ethanalyzer local interface inband
```

Ethanalyzer 不捕获思科 NX-OS 在硬件中转发的数据流量，但是可以将 ACL 与日志选项配合使用作为一种解决方法。

```
Leaf2(config)# ip access-list acl-cap
Leaf2(config-acl)# permit tcp 192.168.60.5 192.168.60.6 eq 80 log
Leaf2(config-acl)# permit ip any any

Leaf2(config)# interface e1/1
Leaf2(config-if)# ip access-group acl-cap in
Leaf2# ethanalyzer local interface inband capture-filter "tcp port 80"
```

回滚使用户可以获得思科 NX-OS 配置的快照（即用户检查点），然后随时将该配置重新应用于设备，而无需重新加载设备。回滚使任何已获授权的管理员可以应用此检查点配置，而无需对检查点中配置的功能具有专业知识。

思科 NX-OS 会自动创建系统检查点。您可以使用用户检查点或系统检查点执行回滚。

用户可以随时创建当前运行配置的检查点副本。思科 NX-OS 会将此检查点另存为 ASCII 文件，用户可以在未来的某个时间使用该文件将正在运行的配置回滚到检查点配置。用户可以创建多个检查点以保存正在运行的配置的不同版本。

当用户回滚正在运行的配置时，会触发以下回滚类型：

- atomic - 仅在没有发生错误时才实施回滚
- best-effort - 实施回滚并跳过任何错误
- stop-at-first-failure - 实施在发生错误时停止的回滚

默认回滚类型是 atomic。

配置检查点和使用回滚

```
!Configure Checkpoint

 Leaf1# checkpoint stable_Leaf1

 Leaf2# checkpoint stable_Leaf2

 Leaf3# checkpoint stable_Leaf3

!Display content of Check point

 Leaf1# show checkpoint stable

!Display differences between check point stable_Leaf1 and running config
 Leaf1# show diff rollback-patch checkpoint stable running-config

!!Rollback the running config to a user checkpoint stable_Leaf1

 Leaf1# rollback running-config checkpoint stable_Leaf1
```

14. 附录

14.1 产品

14.1.1 Cisco Nexus 9500 产品系列

Cisco Nexus 9500 系列模块化交换机在商业数据中心中通常部署为主干（汇聚或核心交换机）。图 50 列出了在撰写本文时 Cisco Nexus 9500 机箱在 NX-OS 单机模式下提供的线卡（仅限 ACI 模式的线卡已被移除）。

图 50. Cisco Nexus 9500 线卡配置

模块化：Cisco Nexus 9500 线卡类型

线卡	端口	模式	交换矩阵模块	机箱支持
X9600				
X9636PQ	36p QSFP+	NX-OS	6	N9504、N9508
X9500				
X9564PX	48p 1/10G SFP+ 和 4p QSFP +	NX-OS、ACI	3	N9504、N9508、N9516
X9564TX	48 p 1/10G-T 和 4p QSFP+	NX-OS、ACI	3	N9504、N9508、N9516
X9536PQ	36p QSFP+ (1.5:1)	NX-OS、ACI	3	N9504、N9508、N9516
X9400				
X9464PX	48P 1/10G SFP+ 和 4p QSFP +	NX-OS	2	N9504、N9508、N9516
X9464TX	48 p 1/10G-T 和 4p QSFP+	NX-OS	2	N9504、N9508、N9516
X9432PQ	32p QSFP+	NX-OS	4	N9504、N9508、N9516

注：查阅 Cisco Nexus 9500 产品手册，了解最新的产品信息。自撰写本文以来，线卡可用性可能已发生变化。

14.1.2 Cisco Nexus 9300 产品系列

Cisco Nexus 9300 系列固定配置交换机通常部署为枝叶（接入交换机）。某些 9300 交换机通过为额外的端口提供上行链路模块插槽成为半模块化交换机。图 51 列出了在撰写本文时提供的支持 NX-OS 单机模式的 Cisco Nexus 9300 机箱。

图 51. Cisco Nexus 9300 机箱配置

非模块化：Cisco Nexus 9300

线卡	端口	模式	RU	上行链路模块
N9396PX	48p 1/10G SFP+ and 12p QSFP+	NX-OS、ACI	2	是
N9396TX	48p 1/10G-T and 12p QSFP+	NX-OS、ACI	2	是
N93128TX	96p 1/10G-T and 8p QSFP+	NX-OS、ACI	3	是
N93128TX2	96p 1/10G-T and 8p QSFP+	NX-OS、ACI	2	否
N93128PX2	96p 1/10G SFP+ and 8p QSFP+	NX-OS、ACI	2	否
N9372PX	48p 1/10G SFP+ and 6p QSFP+	NX-OS、ACI	1	否
N9372TX	48p 1/10G-T and 6p QSFP+	NX-OS、ACI	1	否
N9332PQ	32p QSFP+	NX-OS、ACI	1	否
N9332PQ2	32p QSFP+	NX-OS、ACI	2	是

注：查阅 Cisco Nexus 9300 产品手册，了解最新的产品信息。自撰写本文以来，交换机可用性可能已发生变化。

14.2 NX-API

14.2.1 关于 NX-API

在 Cisco Nexus 设备上，CLI 仅在设备上运行。NX-API 通过使用 HTTP/HTTPS 在交换机外部提供这些 CLI，从而提高了这些 CLI 的可访问性。对 Cisco Nexus 9000 系列设备上的现有 Cisco Nexus CLI 系统使用此扩展。NX-API 支持显示命令、配置、和 Linux Bash。

14.2.2 使用 NX-API

使用 NX-API 输入 Cisco Nexus 9000 系列设备的命令、命令类型和输出类型，方法是将 CLI 编码到 HTTP/HTTPs POST 的正文中。对请求做出的响应以 XML 或 JSON 输出格式返回。

14.2.3 NX-API 沙盒

NX-API 支持配置、显示命令和 Linux Bash。对于请求和响应，NX-API 支持 XML 和 JSON 格式。NX-API 还可以用于使用 REST 客户端和 Postman 配置 Cisco Nexus 9000 交换机。

访问 NX-API 要求管理员在 Cisco Nexus 9000系列交换机上启用管理接口和 NX-API 功能。

在设备上使用功能管理器 CLI 命令启用 NX-API。默认情况下，NX-API 被禁用。在枝叶和主干节点中启用 NX-API。

如何启用管理接口和 NX-API 功能

```
!Enable Management interface
Leaf1# conf t
Leaf1(config)# interface mgmt 0
Leaf1(config)# ip address 172.31.216.130/24
Leaf1(config)# vrf context management
Leaf1(config)# default gateway 172.31.216.1

!Enable NXAPI feature

Leaf1# conf t
Leaf1(config)# feature nxapi
```

一旦启用管理接口和 NX-API 功能，即可按照紧随其后的步骤所述通过沙盒或 Postman 使用它们。默认情况下，在启用 NX-API 功能时，会启用 http/https 支持。

在使用 NX-API 沙盒时，建议使用 Firefox 浏览器 24.0 版本或更高版本。

1. 打开浏览器并输入 **http(s)://<mgmt-ip>** 以启动 NX-API 沙盒。图 53 和 54 是请求和输出响应的示例。

图 52. 请求 - cli_show

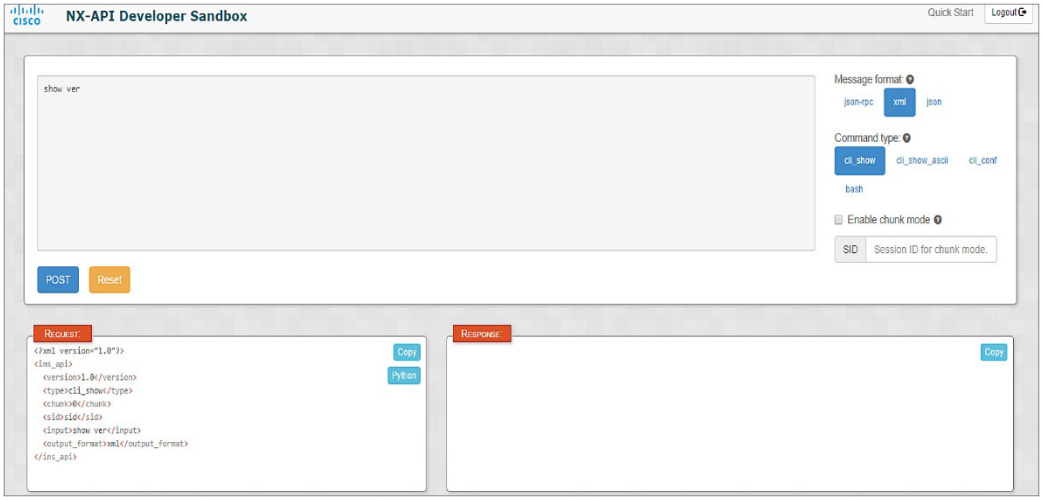
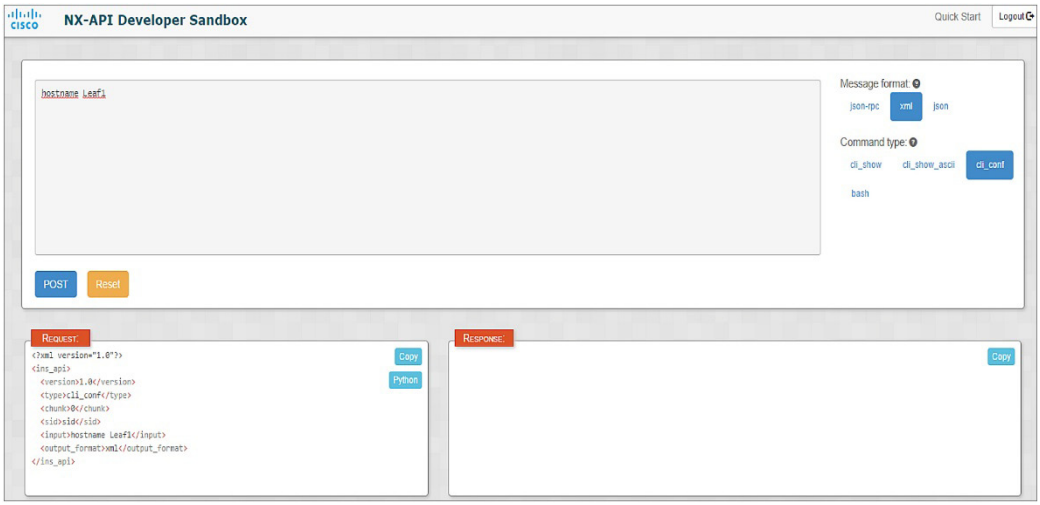


图 53. 输出响应 - cli_config



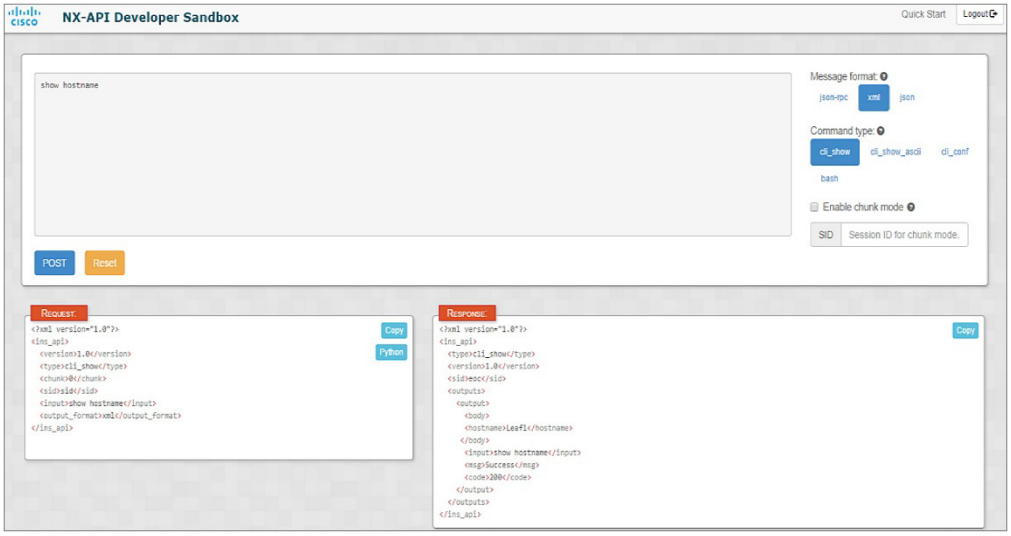
2. 提供要为其生成 XML/JSON 的 CLI 命令。
3. 在顶部窗格中选择 XML 或 JSON 作为消息格式选项。
4. 输入命令类型：cli_show（对于显示命令）和 cli_conf（对于配置命令）。
5. 请求元素的简要说明显示在左下角窗格中。
6. 在请求提交后，输出响应显示在右下角窗格中（图 55）。
7. 如果 CLI 执行成功，响应将包含：

```
...
<msg>Success</msg>
<code>200</code>
...
```

8. 如果 CLI 执行失败，响应将包含：

```
...
<msg>Input CLI command error</msg>
<code>400</code>
...
```

图 54. CLI 执行后的响应

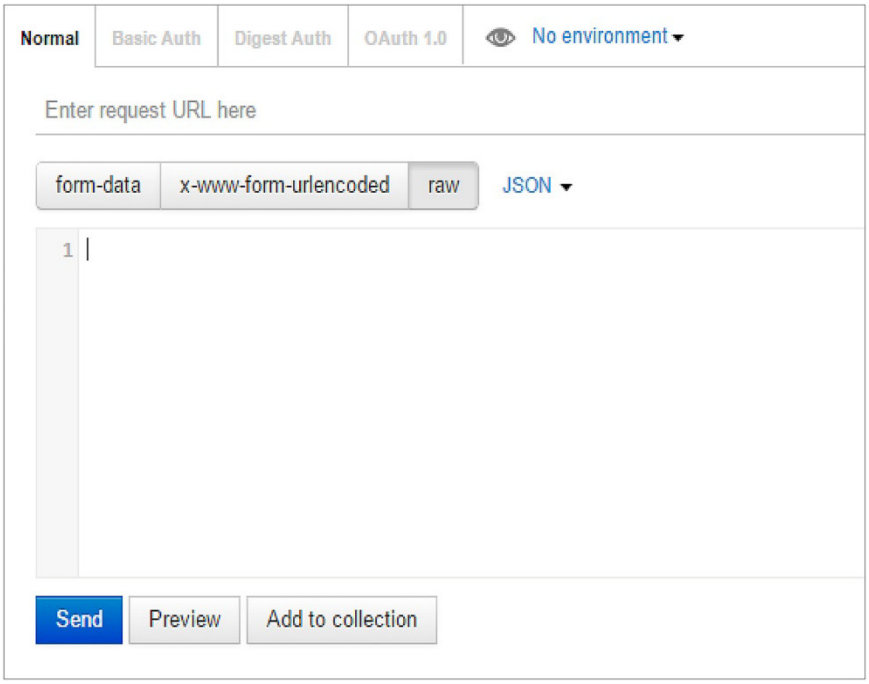


14.2.4 使用 Postman 的 NX-OS 配置

Postman 是作为 Chrome 浏览器扩展提供的 REST 客户端。使用 Postman 可在 Cisco Nexus 设备上执行 NX-API REST 调用。通过编写 API 以远程执行各种配置，Postman 可实现非常好的可重用性。

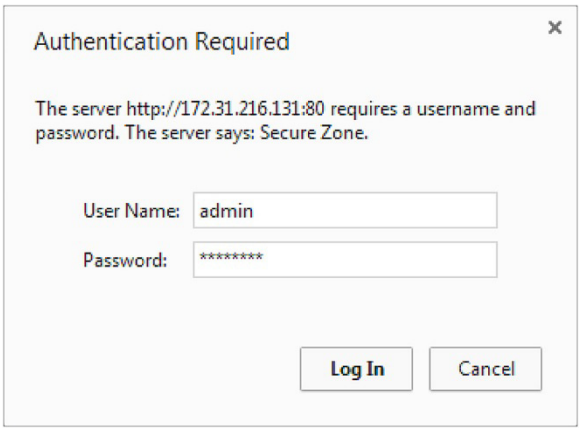
9. 打开网络浏览器并运行 Postman 客户端。空的 Web 用户界面应该与图 56 中的屏幕截图相似。

图 55. 空的 Web 用户界面视图



10. 按照以下步骤将 Postman 用于 NX-API 执行。

I) 在 URL 选项卡中输入 MGMT IP (http://172.31.216.130/ins)。在弹出窗口中提供用户名和密码。

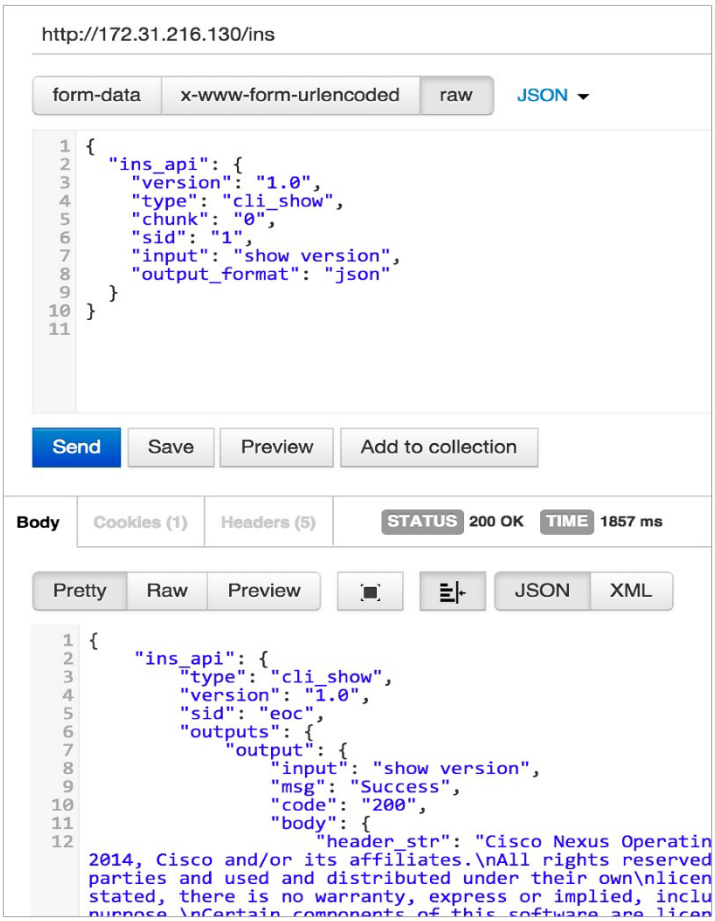


ii) 选择“POST”作为操作类型。

iii) 为原始格式类型选择 ML/JSON。

iv) 请求正文包含 XML/JSON 请求。

V) 按顺序按下以将 API 请求发送到 Cisco Nexus 设备。



14.3 配置

14.3.1 接口和 VLAN 配置

基本 VLAN 和接口配置

```
!Create and (optionally) name VLANs

Leaf1# configure terminal
Leaf1(config)# vlan 50,70,100
Leaf1(config-vlan)# vlan 50
Leaf1(config-vlan)# name Webserver
Leaf1(config-vlan)# vlan 60
Leaf1(config-vlan)# name Appserver



!Configure Layer 3 spine-facing interfaces

!Configure Interface connected to Spine1
Leaf1# configure terminal
Leaf1(config)# interface Ethernet2/1
Leaf1(config-if)# description Connected to Spine1
Leaf1(config-if)# no switchport
Leaf1(config-if)# speed 40000
Leaf1(config-if)# ip address 10.1.1.2/30
Leaf1(config-if)# no shutdown



!Configure Interface connected to Spine2
Leaf1(config)# interface Ethernet2/2
Leaf1(config-if)# description connected to Spine2
Leaf1(config-if)# no switchport
Leaf1(config-if)# speed 40000
Leaf1(config-if)# ip address 10.1.1.10/30
Leaf1(config-if)# no shutdown

!Configure Layer 2 server and service-facing interfaces
!(Optionally) Limit VLANs and use STP best practices

Leaf1(config)# interface Ethernet1/1-2
Leaf1(config-if-range)# switchport
Leaf1(config-if-range)# switchport mode trunk
Leaf1(config-if-range)# switchport trunk allowed vlan 50,60
Leaf1(config-if-range)# no shutdown
```

```
Leaf1(config)# interface Ethernet1/1
Leaf1(config-if)# description to Server1
Leaf1(config-if)# interface Ethernet1/2
Leaf1(config-if)# description to Server2

Leaf1(config-if)# interface Ethernet1/48
Leaf1(config-if)# description to F5 Active LB
Leaf1(config-if)# switchport
Leaf1(config-if)# switchport mode trunk
Leaf1(config-if)# switchport trunk allowed vlan 50,60
Leaf1(config-if)# no shutdown


!Configure Leaf2 Interface connected to Spine1

Leaf2# Configure terminal
Leaf2(config)# interface Ethernet2/1
Leaf2(config-if)# description connected to Spine1
Leaf2(config-if)# no switchport
Leaf2(config-if)# speed 40000
Leaf2(config-if)# ip address 10.1.1.6/30
Leaf2(config-if)# no shutdown

!Configure Interface connected to Spine2
Leaf2(config)# interface Ethernet2/2
Leaf2(config-if)# description to connected to Spine2
Leaf2(config-if)# no switchport
Leaf2(config-if)# speed 40000
Leaf2(config-if)# ip address 10.1.1.14/30
Leaf2(config-if)# no shutdown


!Configure Interface connected to Server1
Leaf2(config)# interface Ethernet1/1
Leaf2(config-if)# description Connected to server1
Leaf2(config-if)# switchport mode trunk 50,60
Leaf2(config-if)# no shutdown



!Configure Interface connected to Server2
Leaf2(config)# interface Ethernet1/2
Leaf2(config-if)# description Connected to server2
Leaf2(config-if)# switchport mode trunk 50,60
```

```
Leaf2(config-if)# no shutdown


!Configure Spine1 Interface connected to Leaf1
Spine1# configure terminal
Spine1(config)# interface Ethernet1/1
Spine1(config-if)# description connected to Leaf1
Spine1(config-if)# speed 40000
Spine1(config-if)# ip address 10.1.1.1/30
Spine1(config-if)# no shutdown

!Configure Interface connected to Leaf2
Spine1(config)# interface Ethernet1/2
Spine1(config-if)# description Connection to Leaf2
```

```
Spine1(config-if)# speed 40000
Spine1(config-if)# ip address 10.1.1.5/30
Spine1(config-if)# no shutdown


!Configure Spine2 Interface connected to Leaf1
Spine2# Configure terminal
Spine2(config)# interface Ethernet1/1
Spine2(config-if)# description Connection to Leaf1
Spine2(config-if)# speed 40000
Spine2(config-if)# ip address 10.1.1.9/30
Spine2(config-if)# no shutdown

!Configure Interface connected to Leaf2
Spine2(config)# interface Ethernet1/2
Spine2(config-if)# description Connection to Leaf2
Spine2(config-if)# speed 40000
Spine2(config-if)# ip address 10.1.1.13/30
Spine2(config-if)# no shutdown
```

14.3.2 路由的配置 - EIGRP

增强型内部网关路由协议 (EIGRP) 配置

```
!Enable the EIGRP feature on Leaf1
!Configure global EIGRP process and (optional) router ID

Leaf1# configure terminal
Leaf1(config)# feature eigrp
```

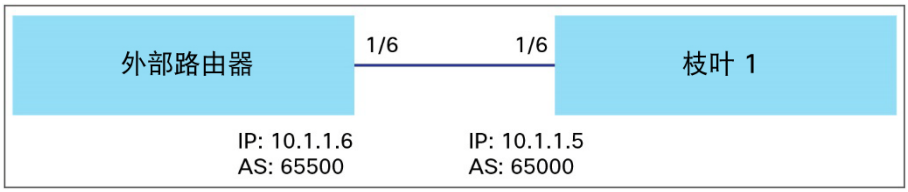
```
Leaf1(config)# router eigrp 1
Leaf1(config-router)# router-id 1.1.1.33

!Enable the EIGRP process on spine-facing interfaces

Leaf1(config-router)# interface Ethernet2/1-2
Leaf1(config-if-range)# ip router eigrp 1


!Enable the EIGRP feature on Leaf2
!Configure global EIGRP process and (optional) router ID

Leaf2# configure terminal
Leaf2(config)# feature eigrp
Leaf2(config)# router eigrp 1
Leaf2(config-router)# router-id 1.1.1.34

!Enable the EIGRP process on spine-facing interfaces

Leaf2(config-router)# interface Ethernet2/1-2
Leaf2(config-if-range)# ip router eigrp 1

!Enable the EIGRP feature on Spine1
!Configure global EIGRP process and (optional) router ID

Spine1# configure terminal
Spine1(config)# feature eigrp
Spine1(config)# router eigrp 1
Spine1(config-router)# router-id 1.1.1.31

!Enable the EIGRP process on Leaf-facing interfaces
Spine1(config-router)# interface Ethernet1/1-3
Spine1(config-if-range)# ip router eigrp 1

!Enable the EIGRP feature on Spine1
!Configure global EIGRP process and (optional) router ID

Spine2# configure terminal
Spine2(config)# feature eigrp
Spine2(config)# router eigrp 1
Spine2(config-router)# router-id 1.1.1.32

!Enable the EIGRP process on Leaf-facing interfaces
```

```
Spine2(config-router)# interface Ethernet1/1-2
Spine2(config-if-range)# ip router eigrp 1
```

14.3.3 DCI 的 BGP 配置

此配置在枝叶与为 DCI/外部连接所连接的外部路由器之间。

配置数据中心内部的 BGP AS 65000，并重新分发路由。




```
Leaf1# Configure terminal
Leaf1(config)# feature bgp
Leaf1(config)# route-map bgp-route permit 65000
Leaf1(config-map)# match ip address prefix-list 172.23.*.*
Leaf1(config-map)# set tag 65500

!Configure global BGP routing process and establish neighbor with External border
router

Leaf1(config)# router bgp 65000
Leaf1(config)# neighbor 192.168.20.8 remote-as 65500
Leaf1(config-router)# address-family ipv4 unicast
Leaf1(config-router-af)# redistribute eigrp bgp-re route-map bgp-route

!Re-distribute BGP routes into EIGRP routing table

Leaf1(config)# router eigrp 1
Leaf1(config-router)# router-id 1.1.1.34
Leaf1(config-router)# redistribute bgp 65000 route-map bgp-route
```

14.3.4 接入层的 vPC 配置

vPC 域和设备配置

```
!Enable feature and create a vPC domain

Leaf1(config)# feature vpc
Leaf1(config)# vpc domain 1

!Set peer keepalive heartbeat source and destination

Leaf1(config-vpc-domain)# )# peer-keepalive destination 172.31.216.131 source
172.31.216.130
```

```
!Automatically and simultaneously enable the following
! best practice commands using the mode auto command: peer-
! gateway, auto-recovery, ip arp synchronize, and ipv6 nd
! synchronize.

Leaf1(config-vpc-domain)# mode auto

!Create port channel interface and attach it to physical interface

Leaf1(config)# feature lacp

!Create the peer-link Port Channel and add to domain

!Configure best practice Spanning Tree features on vPC
!Create a Port Channel and vPC for a server
!The vPC must match on the other peer for the server

Leaf1(config)# interface port-channel10
Leaf1(config-if)# description to Peer
Leaf1(config-if)# switchport mode trunk
Leaf1(config-if)# spanning-tree port type network
Leaf1(config-if)# vpc peer-link
Leaf1(config-if)# no shutdown


!!Please note that spanning tree port type is changed to
"network" port type on vPC peer-link.This will enable
spanning tree Bridge Assurance on vPC peer-link provided
the STP Bridge Assurance (which is enabled by default) is
not disabled.

Leaf1(config)# interface port-channel20
Leaf1(config-if)# description to Sever1
Leaf1(config-if)# switchport mode trunk
Leaf1(config-if)# vpc 20
Leaf1(config-if)# no shutdown

Leaf1(config)# interface port-channel21
Leaf1(config-if)# description to Sever2
Leaf1(config-if)# switchport mode trunk
Leaf1(config-if)# vpc 21
Leaf1(config-if)# no shutdown

Leaf1(config-if)# interface Ethernet1/1
Leaf1(config-if)# switchport mode trunk allowed vlan 50,60
```

```
Leaf1(config-if)# channel-group 20

Leaf1(config-if)# interface Ethernet1/2
Leaf1(config-if)# switchport mode trunk allowed vlan 50,60
Leaf1(config-if)# channel-group 21



Leaf1(config)# interface Ethernet1/33-34
Leaf1(config-if-range)# channel-group 10 mode active



Leaf2

!Enable feature and create a vPC domain

Leaf2(config)# feature vpc
Leaf2(config)# vpc domain 1

!Set peer keepalive heartbeat source and destination

Leaf2(config-vpc-domain)# )# peer-keepalive destination 172.31.216.130 source
172.31.216.131

!Automatically and simultaneously enable the following
! best practice commands using the mode auto command: peer-
! gateway, auto-recovery, ip arp synchronize, and ipv6 nd
! synchronize.

Leaf2(config-vpc-domain)# mode auto

!Create the peer-link Port Channel and add to domain

Leaf2(config-vpc-domain)# feature lacp

Leaf2(config-if)# interface port-channel 10
Leaf2(config-if)# switchport
Leaf2(config-if)# switchport mode trunk
Leaf2(config-if)# switchport trunk allowed vlan 50,60
Leaf2(config-if)# vpc peer-link

Leaf2(config)# interface Ethernet1/33-34
Leaf2(config-if-range)# channel-group 10 mode active


Please note that spanning tree port type is changed to "network" port type on vPC
peer-link.This will enable spanning tree Bridge Assurance on vPC peer-link
provided the STP Bridge Assurance (which is enabled by default) is not disabled.
```

```
Leaf2(config)# no shutdown

!Configure best practice Spanning Tree features on vPC
!Create a Port Channel and vPC for a server
!The vPC #11 must match on the other peer for the server

Leaf2(config)# interface port-channel20
Leaf2(config-if)# description to Sever1
Leaf2(config-if)# switchport mode trunk
Leaf2(config-if)# vpc 20
Leaf2(config-if)# no shutdown

Leaf2(config)# interface port-channel21
Leaf2(config-if)# description to Sever2
Leaf2(config-if)# switchport mode trunk
Leaf2(config-if)# vpc 21
Leaf2(config-if)# no shutdown

Leaf2(config-if)# interface Ethernet1/1
Leaf2(config-if)# switchport mode trunk allowed vlan 50,60
Leaf2(config-if)# channel-group 20

Leaf2(config-if)# interface Ethernet1/2
Leaf2(config-if)# switchport mode trunk allowed vlan 50,60
Leaf2(config-if)# channel-group 21
```

表 7 概述了这些配置的显示命令。

表 7. 显示命令

Show vpc	提供有关 vPC 配置的详细信息和设备中的各种链路的状态
Show vpc consistency-parameters <vlans/global>	提供有关 vPC 类型 1 和类型 2 一致性参数及其对等体的详细信息

14.3.5 组播和 VXLAN

主干协议无关组播 (PIM) 和组播源发现协议 (MSDP) 组播配置

```
Spine1
!Enable PIM and MSDP to turn on multicast routing

Spine1(config)# feature pim
Spine1(config)# feature msdp

!Enable PIM on the spine-facing interfaces
!EIGRP should already be enabled on the interfaces
```

```
Spine1(config)# interface Ethernet1/1-2
Spine1(config-if-range)# ip pim sparse-mode

!Configure the Rendezvous Point IP address for the ASM
! group range 239.0.0.0/8 to be used by VXLAN segments

Spine1(config)# ip pim rp-address 10.2.2.12 group-list 239.0.0.0/8

!Configure MSDP sourced from loopback 1 and the peer spine
! switch's loopback IP address to enable RP redundancy

Spine1(config)# ip pim ssm range 232.0.0.0/8
Spine1(config)# ip msdp originator-id loopback1
Spine1(config)# ip msdp peer 10.2.2.2 connect-source loopback1

!Configure loopbacks to be used as redundant RPs with
! the other spine switch.L0 is assigned a shared RP IP
! used on both spines.L1 has a unique IP address.
!Enable PIM and EIGRP routing on both loopback interfaces.

Spine1(config)# interface loopback0
Spine1(config-if)# ip address 10.2.2.12/32
Spine1(config-if)# ip router eigrp 1
Spine1(config-if)# ip pim sparse-mode

Spine1(config)# interface loopback1
Spine1(config-if)# ip address 10.2.2.1/32
Spine1(config-if)# ip router eigrp 1
Spine1(config-if)# ip pim sparse-mode

Spine2
!Enable PIM and MSDP to turn on multicast routing

Spine2(config)# feature pim
Spine2(config)# feature msdp

!Enable PIM on the spine-facing interfaces
!EIGRP should already be enabled on the interfaces

Spine2(config)# interface Ethernet1/1-2
Spine2(config-if-range)# ip pim sparse-mode

!Configure the Rendezvous Point IP address for the ASM
! group range 239.0.0.0/8 to be used by VXLAN segments
```

```
Spine2(config)# ip pim rp-address 10.2.2.12 group-list 239.0.0.0/8

!Configure MSDP sourced from loopback 1 and the peer spine
! switch's loopback IP address to enable RP redundancy

Spine2(config)# ip pim ssm range 232.0.0.0/8
Spine2(config)# ip msdp originator-id loopback1
Spine2(config)# ip msdp peer 10.2.2.1 connect-source loopback1

Spine2(config)# interface loopback0
Spine2(config-if)# ip address 10.2.2.12/32
Spine2(config-if)# ip router eigrp 1
Spine2(config-if)# ip pim sparse-mode

Spine2(config)# interface loopback1
Spine2(config-if)# ip address 10.2.2.2/32
Spine2(config-if)# ip router eigrp 1
Spine2(config-if)# ip pim sparse-mode
```

枝叶组播配置

```
Leaf1
!Enable and configure PIM on the spine-facing interfaces

Leaf1(config)# feature pim
Leaf1(config)# interface Ethernet2/1-2
Leaf1(config-if-range)# ip pim sparse-mode

!Point to the RP address configured in the spines

Leaf1(config)# ip pim rp-address 10.2.2.12 group-list 239.0.0.0/8
Leaf1(config)# ip pim ssm range 232.0.0.0/8

Leaf2
!Enable and configure PIM on the spine-facing interfaces

Leaf2(config)# feature pim
Leaf2(config)# interface Ethernet2/1-2
Leaf2(config)# ip pim sparse-mode

!Point to the RP address configured in the spines

Leaf2(config)# ip pim rp-address 10.2.2.12 group-list 239.0.0.0/8
Leaf2(config)# ip pim ssm range 232.0.0.0/8
```

VXLAN 配置

```
!Enable VXLAN features

Leaf1(config)# feature nv overlay
Leaf1(config)# feature vn-segment-vlan-based

!Configure loopback used for VXLAN tunnels
!Configure secondary IP address for vPC support

Leaf1(config)# interface loopback1
Leaf1(config-if)# ip address 192.168.1.1/32
Leaf1(config-if)# ip address 192.168.2.1/32 secondary
Leaf1(config-if)# ip router eigrp 1
Leaf1(config-if)# ip pim sparse-mode

!Create NVE interface, using loopback as the source
!Bind VXLAN segments to an ASM multicast group

Leaf1(config-if)# interface nve1
Leaf1(config-if)# source-interface loopback1
Leaf1(config-if)# member vni 5000 mcast-group 239.1.1.50
Leaf1(config-if)# member vni 6000 mcast-group 239.1.1.60

!Map VLANs to VXLAN segments

Leaf2(config-if)# vlan 50
Leaf2(config-vlan)# vn-segment 5000
Leaf2(config-vlan)# vlan 60
Leaf2(config-vlan)# vn-segment 6000
Leaf2(config-vlan)# vlan 100
Leaf2(config-vlan)# vn-segment 10000
```

表 8 列出了这些配置的相关显示命令。

表 8. 显示命令

show nve vni	提供 VNI 的详细信息以及组播地址和状态
show nve interface	提供 VTEP 接口的详细信息
show nve peers	提供 VTEP 对等设备的详细信息
show mac address-table dynamic	提供为设备所知的 MAC 地址详细信息

14.3.6 防火墙 ASA 配置

```
!Create subinterface for Webserver and Appserver using VLAN 50 and 60
!Assign security level 100
!Assign an IP address to act as a gateway for VLAN 50 and 60


ciscoasa(config)# interface Ethernet0/0.50
ciscoasa(config-subif)# description Webserver
ciscoasa(config-subif)# vlan 50
ciscoasa(config-subif)# nameif Webserver
ciscoasa(config-subif)# security-level 100
ciscoasa(config-subif)# ip address 192.168.50.1 255.255.255.0

ciscoasa(config)# interface Ethernet0/0.60
ciscoasa(config-subif)# description Appserver
ciscoasa(config-subif)# vlan 60
ciscoasa(config-subif)# nameif Appserver
ciscoasa(config-subif)# security-level 100
ciscoasa(config-subif)# ip address 192.168.60.1 255.255.255.0


!Create external interface

ciscoasa(config)# interface Ethernet0/1
ciscoasa(config)# nameif outside
ciscoasa(config)# security-level 100
ciscoasa(config)# ip address 209.165.201.2 255.255.255.0

ciscoasa(config)# interface Management0/0
ciscoasa(config)# nameif mgmt
ciscoasa(config)# security-level 100
ciscoasa(config)# ciscoasa(config)# ip address 172.31.216.149 255.255.252.0
ciscoasa(config)# management-only

!Create object group to match web and HTTP traffic

ciscoasa(config)# same-security-traffic permit inter-interface
ciscoasa(config)# same-security-traffic permit intra-interface

!Provide access to traffic from outside to Webserver

ciscoasa(config)# object network Ciscoapp
ciscoasa(config-network-object)# host 209.165.201.5
```

```
ciscoasa(config)# object network Webserver-VIP
ciscoasa(config-network-object)# host 192.168.50.2
ciscoasa(config)# nat (outside,Webserver) source static any destination static
Ciscoapp Webserver-VIP

!Provide access to traffic from inside Webserver/Appserver to outside

ciscoasa(config)# object network Webserver-outside
ciscoasa(config-network-object)# nat (Webserver,outside) dynamic interface
ciscoasa(config)# object network Appserver-outside
ciscoasa(config-network-object)# nat (AppServer,outside) dynamic interface

!Provide access to traffic from Webserver/Appserver

ciscoasa(config)# object network Web-Source
ciscoasa(config-network-object)# subnet 192.168.50.0 255.255.255.0
ciscoasa(config)# object network App-Source
ciscoasa(config-network-object)# subnet 192.168.60.0 255.255.255.0
ciscoasa(config)# nat (Webserver,AppServer) source static Web-Source Web-Source
destination static App-Source App-Source
ciscoasa(config)# nat (AppServer,Webserver) source static App-Source App-Source
destination static Web-Source Web-Source

!Configure access list for HTTP traffic from outside to Webserver

ciscoasa(config)# object service http
ciscoasa(config-service-object)# service 80

ciscoasa(config)# access-list HTTP extended permit object http interface outside
interface Webserver
 ciscoasa(config)# access-group HTTP global
```

有关思科 ASA 5500 防火墙的详细信息，请访问思科 ASA 5500-X 系列下一代防火墙网站。

14.3.7 F5 LTM 负载均衡器配置

负载均衡器配置

```
ltm node 192.168.50.5 {
  address 192.168.50.5
  session monitor-enabled
  state up
}
ltm node 192.168.60.5 {
```

```
    address 192.168.60.5
    session monitor-enabled
    state up
  }
  ltm node 192.168.60.6 {
    address 192.168.60.6
    session monitor-enabled
    state up
  }
  }
  ltm node 192.168.50.6 {
    address 192.168.50.6
    monitor icmp
    session monitor-enabled
    state up
  }

  ltm persistence global-settings { }
  ltm pool App_VIP {
    members {
      192.168.60.5:http {
        address 192.168.60.5
        session monitor-enabled
        state up
      }
      192.168.60.6:http {
        address 192.168.60.6
        session monitor-enabled
        state up
      }
    }
    monitor http and gateway_icmp
  }
  ltm pool WEB_VIP {
    members {
      192.168.50.5:http {
        address 192.168.50.5
        session monitor-enabled
        state up
      }
      192.168.50.6:http {
        address 192.168.50.6
        session monitor-enabled
        state up
```

```
      }
    }
    monitor http and gateway_icmp
  }
  ltm profile fastl4 FASTL4_ROUTE {
    app-service none
    defaults-from fastL4
  }

  ltm traffic-class Traffic_Class {
    classification Telnet
    destination-address 172.16.0.0
    destination-mask 255.255.0.0
    destination-port telnet
    protocol tcp
    source-address 172.16.0.0
    source-mask 255.255.0.0
    source-port telnet
  }

  ltm virtual Virtual_App {
    destination 192.168.60.2:http
    ip-protocol tcp
    mask 255.255.255.255
    pool App_VIP
    profiles {
      tcp { }
    }
    source 0.0.0.0/0
    source-address-translation {
      type automap
    }
    vs-index 12
  }

  ltm virtual Virtual_Web {
    destination 192.168.50.2:http
    ip-protocol tcp
    mask 255.255.255.255
    pool WEB_VIP
    profiles {
      tcp { }
    }
    source 0.0.0.0/0
```

```
    source-address-translation {
      type automap
    }
    vs-index 12
  }

  net interface 1/1.7 {
    if-index 385
    lldp-tlvmap 113264
    mac-address 00:23:e9:9d:f6:10
    media-active 10000SFPCU-FD
    media-max 10000T-FD
    mtu 9198
    serial TED1829B884
    vendor CISCO-TYCO
  }

  net interface 1/1.8 {
    if-index 401
    mac-address 00:23:e9:9d:f6:11
    media-active 10000SFPCU-FD
    media-max 10000T-FD
    mtu 9198
    serial TED1713H0DT
    vendor CISCO-TYCO
  }
  net interface 1/mgmt {
    if-index 145
    mac-address 00:23:e9:9d:f6:09
    media-active 1000T-FD
    media-max 1000T-FD
  }

  net route-domain 0 {
    id 0
    routing-protocol {
      OSPFv2
    }
    vlans {
      VLAN60
      VLAN50
    }
  }
```

```
net self App_IP {
  address 192.168.60.10/24
  traffic-group traffic-group-local-only
  vlan VLAN60
}

net self Web_IP {
  address 192.168.50.10/24
  traffic-group traffic-group-local-only
  vlan VLAN50
}

net self-allow {
  defaults {
    ospf:any
    tcp:domain
    tcp:f5-iquery
    tcp:https
    tcp:snmp
    tcp:ssh
    udp:520
    udp:cap
    udp:domain
    udp:f5-iquery
    udp:snmp
  }
}


net trunk Trunk_Internal {
  bandwidth 10000
  cfg-mbr-count 1
  id 0
  interfaces {
    1/1.8
  }
  mac-address 00:23:e9:9d:f7:e0
  working-mbr-count 1
}

net vlan VLAN50 {
  if-index 832
  interfaces {
    Trunk_Internal {
```

```
        tagged
      }
    }
    tag 50
  }
  net vlan VLAN60 {
    if-index 800
    interfaces {
      Trunk_Internal {
        tagged
      }
    }
    tag 60
  }

  net vlan-group VLAN_Grp_Internal {
    members {
      VLAN50
      VLAN60
    }
  }

  sys management-route default {
    description configured-statically
    gateway 172.31.216.1
    network default
  }
```

14.4 参考

14.4.1 设计指南

- 小型到中型商业数据中心开始使用 Cisco Nexus 9000 系列交换机

14.4.2 Nexus 9000 平台

- Cisco Nexus 9300 平台缓冲区和排队架构
- Cisco Nexus 9300 平台交换机的 VXLAN 设计
- Cisco Nexus 9000 系列交换机的思科 NX-OS 软件增强功能
- Cisco Nexus 9500 系列交换机架构
- Cisco Nexus 9508 交换机电源和性能
- 使用 Cisco Nexus 9000 系列交换机的经典网络设计
- 40 Gbps 双向和并行光纤收发器的光纤布线连接指南
- Cisco Nexus 9000 系列交换机的网络可编程性和自动化
- 采用 Cisco Nexus 9000 系列交换机的简化 40 Gbps 布线部署解决方案
- VXLAN 概述：Cisco Nexus 9000 系列交换机

14.4.3 网络一般信息

- 数据中心重叠技术
- 以应用为中心的基础设施的原理

14.4.4 迁移

- 将数据中心迁移到以应用为中心的基础设施
- 从 Cisco Catalyst 6500 系列交换机迁移到 Cisco Nexus 9000 系列交换机
- 使用 Cisco QSFP BiDi 技术迁移到 40 Gbps 数据中心

14.4.5 分析报告

- Cisco Nexus 9508 电源效率 - Lippis 报告
- Cisco Nexus 9508 交换机性能测试 - Lippis 报告
- Cisco Nexus 9000 可编程网络环境 - Lippis 报告
- Cisco Nexus 9000 系列研究报告 - Lippis 报告
- Nexus 9000 交换系列为何能够提供最高的可用性和可靠性（以平均故障间隔时间 [MTBF] 衡量）- Lippis 报告
- Miercom 报告：Cisco Nexus 9516

美洲总部
Cisco Systems, Inc.
加州圣何西

亚太地区总部
Cisco Systems (USA) Pte.Ltd.
新加坡

欧洲总部
Cisco Systems International BV
荷兰阿姆斯特丹

思科在全球设有 200 多个办事处。 地址、电话号码和传真号码均列在思科网站 **www.cisco.com/go/offices** 中。

美国印刷 C07-733639-00 03/15