

了解NFVIS虚拟网络：OVS、DPDK和SR-IOV

目录

[简介](#)

[使用的组件](#)

[NFVIS中的网络概述](#)

[ENCS54XX平台](#)

[Catalyst 8200 uCPE](#)

[Catalyst 8300 uCPE 1N20](#)

[网络虚拟化技术](#)

[开放式vSwitch \(OVS\)](#)

[OVS网桥](#)

[环境交换缺陷](#)

[数据平面开发套件\(DPDK\)](#)

[数据复制](#)

[PCIe直通](#)

[单根I/O虚拟化\(SR-IOV\)](#)

[物理功能\(PF\)](#)

[虚拟功能\(VE\)](#)

[在支持NFVIS的硬件上加速SR-IOV的建议驱动程序](#)

[DPDK和SR-IOV的使用案例](#)

[DPDK首选项](#)

[SR-IOV首选项](#)

[配置](#)

[启用DPDK](#)

[创建新网络并将其关联到新的OVS网桥](#)

[连接VNE](#)

[相关文章和文档](#)

简介

本文档介绍NFVIS平台为通信企业和服务网络中的VNF提供的虚拟网络方案。

使用的组件

本文档中的信息基于以下硬件和软件组件：

- 运行NFVIS 4.7.1-FC4的ENCS5412
- 运行NFVIS 4.12.1-FC2的c8300 uCPE 1N20

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。

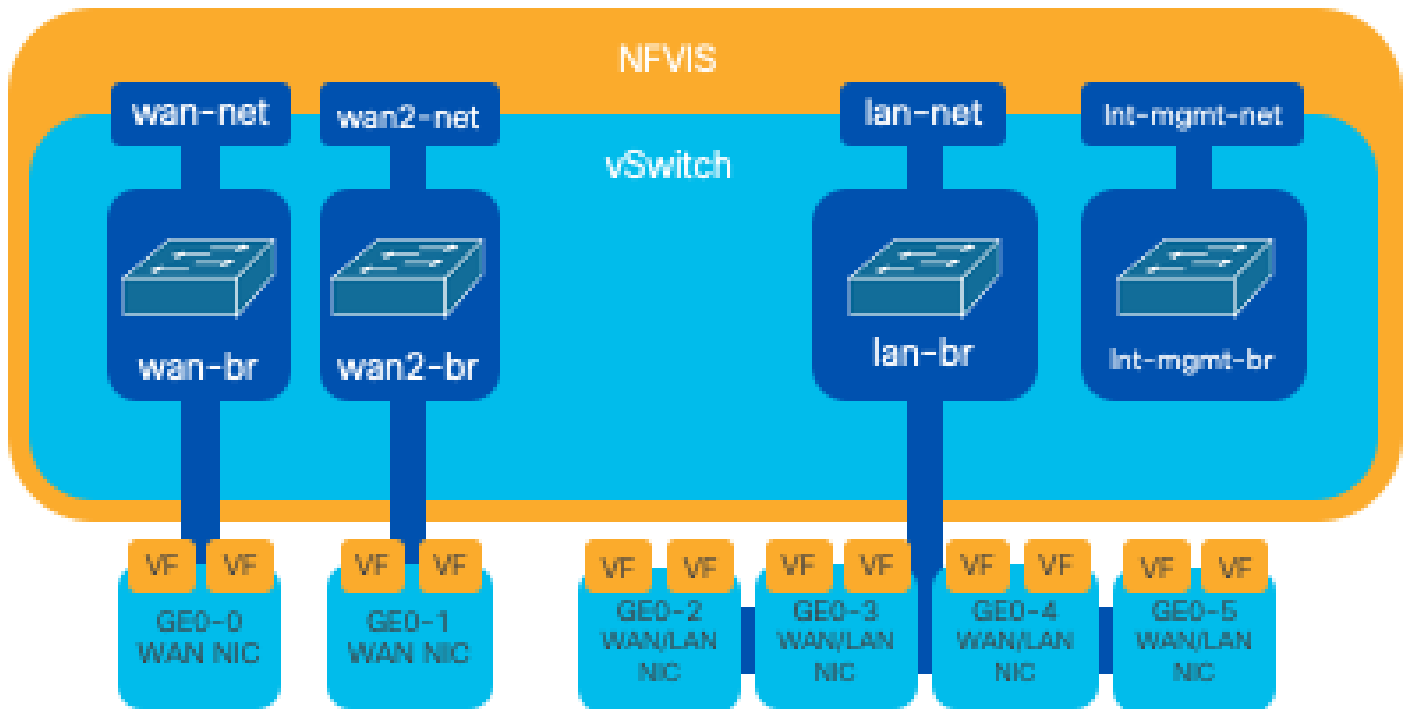


图 2. 分配给8200 NIC的内部桥接和虚拟交换机

Catalyst 8300 uCPE 1N20

1. 默认情况下，可以通过FPGE（前面板千兆以太网）WAN端口或通过GEO-2 LAN端口访问NFVIS以进行管理
2. 默认情况下，WAN网络(wan-net)和WAN网桥(wan-br)设置为启用DHCP。默认情况下，GEO-0与WAN网桥关联
3. 默认情况下会创建WAN网络(wan2-net)和WAN网桥(wan2-br)，但不会将其与任何物理端口相关联
4. GEO-2与LAN网桥关联，所有其他端口不与OVS关联
5. 可以通过GEO-2访问C8300-uCPE上的管理IP 192.168.1.1
6. 创建内部管理网络(int-mgmt-net)和网桥(int-mgmt-br)，并在内部用于系统监控。

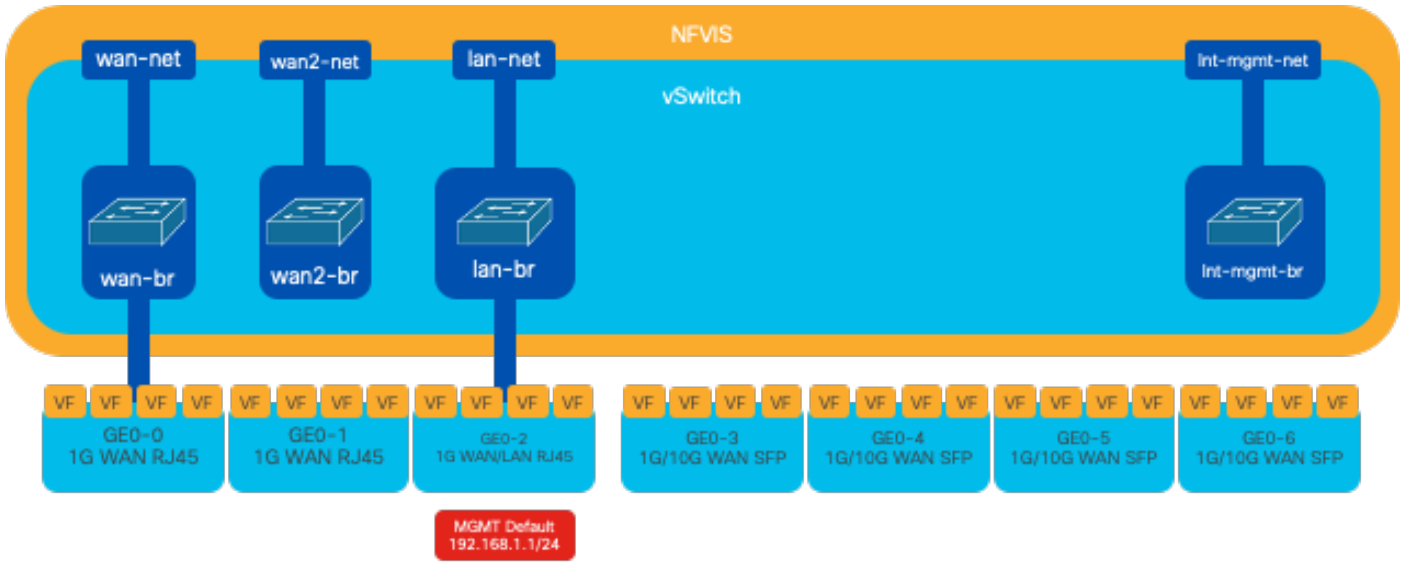


图 3. 分配给8300 NIC的内部桥接和虚拟交换机

网络虚拟化技术

开放式vSwitch (OVS)

开放式vSwitch (OVS)是一种开源的多层虚拟交换机，旨在通过编程扩展实现网络自动化，同时支持标准管理接口和协议，例如NetFlow、sFlow、IPFIX、RSPAN、CLI、LACP和802.1ag。它广泛用于大型虚拟化环境，特别是虚拟机监控程序来管理虚拟机(VM)之间的网络流量。它允许创建通过NFVIS接口直接管理的复杂网络拓扑和策略，为网络功能虚拟化提供多功能环境。

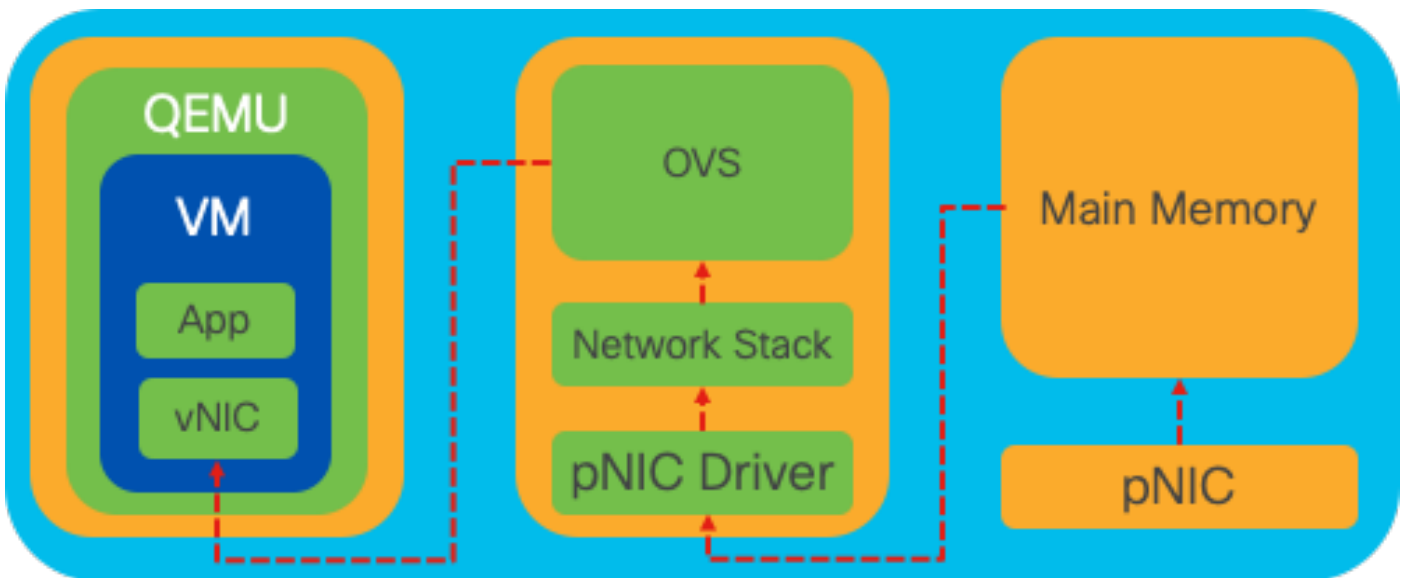


图 4.Linux内核中的OVS配置

OVS网桥

它使用虚拟网络网桥和流规则在主机之间转发数据包。它的行为类似于物理交换机，只是虚拟化的。

。

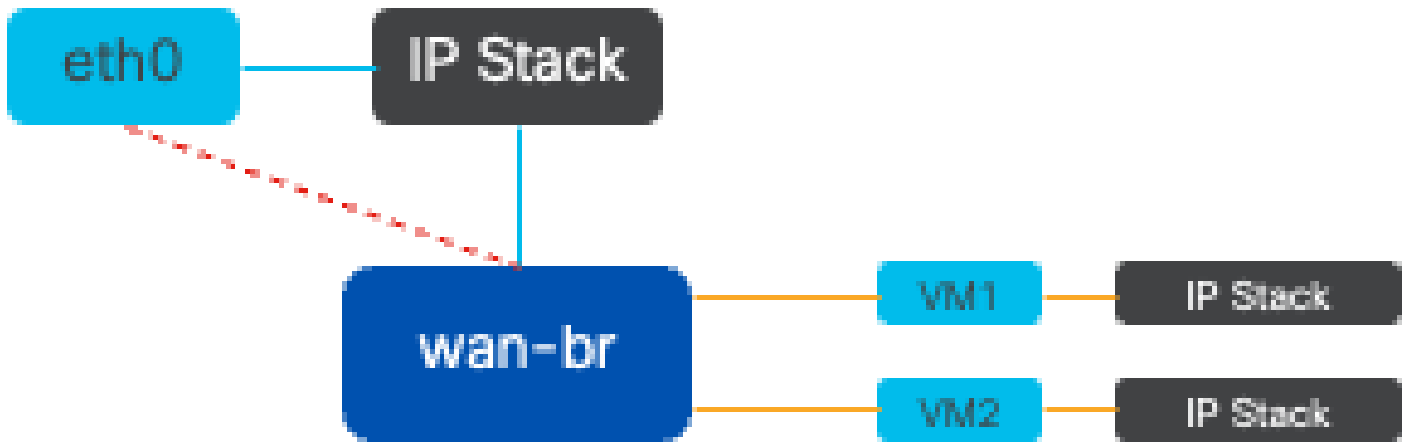


图 5. 连接到wan-br网桥的2个VM或VNF的示例实施

环境交换缺陷

当网络数据包到达网络接口卡(NIC)时，它会触发中断，即发信号给处理器，表明它需要立即处理。CPU暂停其当前任务以处理中断，该过程称为中断处理。在此阶段，CPU在操作系统内核的控制下，将数据包从网卡读入内存，并根据数据包的目的地和用途决定后续步骤。目标是快速处理数据包或将数据包路由到其预期应用，从而最大程度地减少延迟并最大化吞吐量。

情景交换是CPU从一个环境（情景）中的任务切换到另一个环境的过程。在用户模式和内核模式之间切换时，这一点尤其重要：

- 用户模式：这是限制性的处理模式，大多数应用程序在此模式下运行。用户模式下的应用程序不能直接访问硬件或参考内存，并且必须与操作系统内核通信才能执行这些操作。
- 内核模式：此模式授予操作系统对硬件和所有内存的完全访问权限。内核可以执行任何CPU指令并引用任何内存地址。执行管理硬件设备、内存和执行系统调用等任务需要内核模式。

当应用程序需要执行需要内核级权限的操作（例如读取网络数据包）时，就会发生上下文切换。CPU从用户模式转换到内核模式以执行操作。完成后，另一个情景交换机将CPU返回到用户模式以继续执行应用。此交换过程对于维护系统的稳定性和安全性至关重要，但会引入可能会影响性能的开销。

OVS主要运行于操作系统的用户空间中，随着数据吞吐量的增加，它可能会成为瓶颈。这是因为CPU需要更多的上下文切换才能转到内核模式来处理数据包，从而降低了性能。这种限制在高数据包速率或精确计时至关重要的环境中尤为明显。为了解决这些性能限制并满足现代高速网络的需求，开发了DPDK（数据平面开发套件）和SR-IOV（单根I/O虚拟化）等技术。

数据平面开发套件(DPDK)

DPDK是一组库和驱动程序，旨在加速各种CPU架构上的数据包处理工作负载。通过绕过传统内核网络堆栈（避免情景交换），DPDK可以显著提高数据平面吞吐量并减少延迟。这对于需要低延迟通信的高吞吐量VNF特别有益，使NFVIS成为性能敏感型网络功能的理想平台。

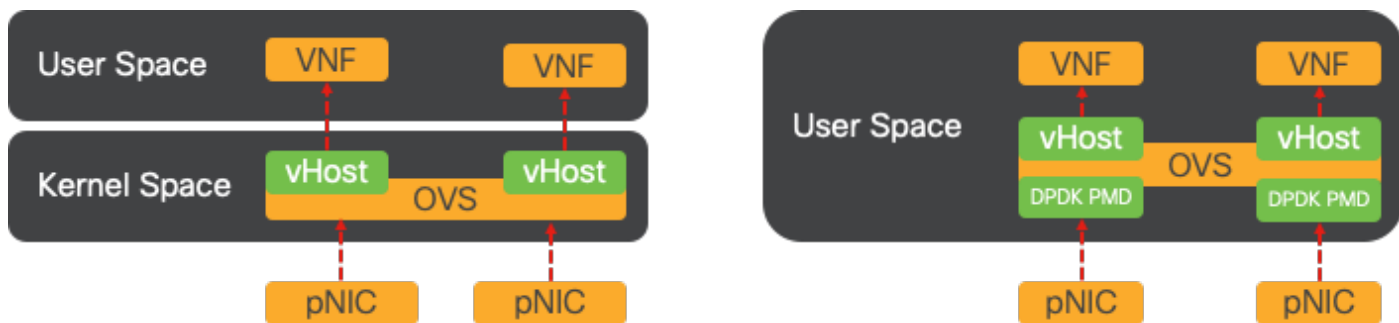


图 6.传统OVS (左侧) 和DPDK OVS (右侧) 环境交换优化

对OVS的DPDK的支持从NFVIS 3.10.1 (对于ENCs) 和3.12.2 (对于其他平台) 开始。

- SRIOV附近的服务链吞吐量，优于非DPDK OVS。
- VNF需要虚拟驱动程序。
- 支持的平台:
- 从ENCs 3.10.1开始。
- UCSE、UCS-C、CSP5K 3.12.1及以上版本。
- 自4.12.1起支持的端口通道的DPDK。
- 数据包/流量捕获：DPDK中不支持。
- PNIC上的SPAN流量：DPDK中不支持。
- 启用OVS-DPDK后，无法将其作为单独功能禁用。禁用DPDK的唯一方法是重置出厂设置。

数据复制

传统网络方法通常要求数据在到达VM内存中的目的地之前多次复制。例如，必须将数据包从NIC复制到内核空间，然后复制到用户空间，以便由虚拟交换机（如OVS）进行处理，最后复制到VM内存。尽管DPDK通过绕过内核网络堆栈提供了性能改进，但每次复制操作都会导致延迟并增加CPU使用率。

这些开销包括内存副本以及在将数据包转发到VM之前在用户空间处理数据包所需的处理时间。PCIe Passthrough和SR-IOV通过允许物理网络设备（如NIC）在多个虚拟机之间直接共享，而不像传统虚拟化方法那样涉及主机操作系统，解决了这些瓶颈。

PCIe直通

该策略包括绕过虚拟机监控程序，以允许虚拟网络功能(VNF)直接访问网络接口卡(NIC)，从而实现接近最大吞吐量的目标。此方法称为PCI直通，它允许完整NIC专门用于访客操作系统，而不需要虚拟机监控程序的干预。在此设置中，虚拟机的运行方式与其直接连接到网卡。例如，由于有两个NIC卡可用，每个卡可以专门分配给不同的VNF，从而提供直接访问。

但是，此方法有一个缺点：如果只有两个NIC可供两个单独的VNF独占使用，则任何其他VNF（如第三个VNF）将没有NIC访问权限，因为缺少可供其使用的专用NIC。替代解决方案包括使用单根I/O虚拟化(SR-IOV)。

单根I/O虚拟化(SR-IOV)

是允许单个物理PCI设备（如网络接口卡[NIC]）显示为多个独立虚拟设备的规范。该技术提供对物理网络设备的直接VM访问，从而降低开销并提高I/O性能。其工作原理是将单个PCIe设备划分为多个虚拟片，每个虚拟片可分配给不同的VM或VNF，从而有效地解决了有限数量的NIC导致的限制。这些虚拟片称为虚拟功能(VF)，允许在多个VNF之间共享NIC资源。物理功能(PF)是指促进SR-IOV功能的实际物理组件。

通过利用SR-IOV，NFVIS可以将专用NIC资源分配给特定VNF，通过促进网络数据包的直接内存访问(DMA)直接进入相应VM内存，确保高性能和低延迟。此方法将CPU仅处理数据包的情况减至最少，从而降低CPU使用率。这对于需要保证带宽或具有严格性能要求的应用程序尤其有用。

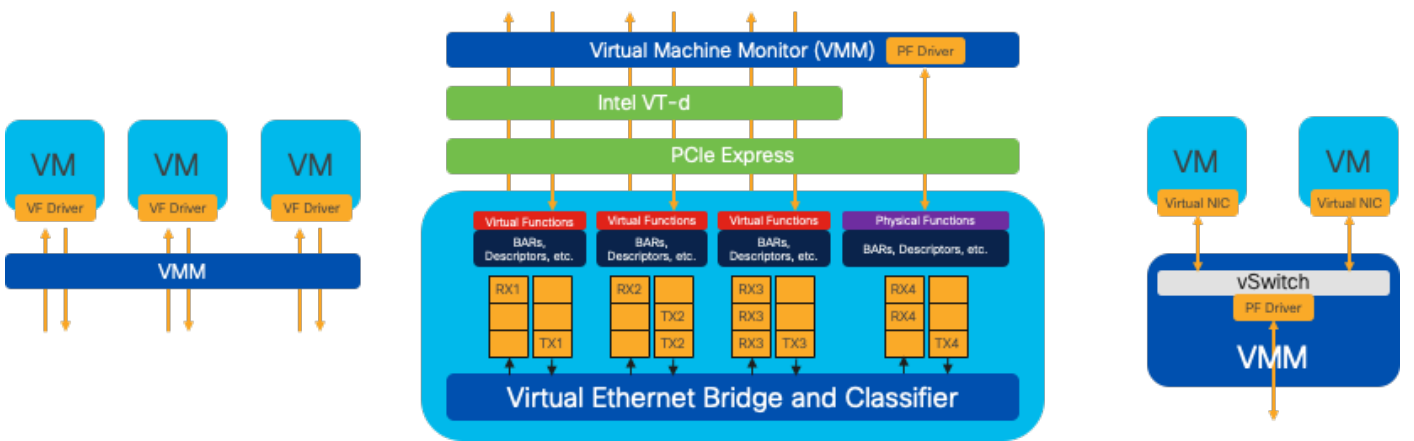


图 7.NFVIS SR-IOV PCIe资源通过硬件功能分离

物理功能(PF)

它们是功能齐全的PCIe功能，是指提供特定网络功能的专用硬件盒；它们是功能齐全的PCIe功能，可以像任何其他PCIe设备一样发现、管理和操作。物理功能包括可用于配置和控制PCIe设备的SR-IOV功能。

虚拟功能(VF)

它们是精简型功能，配置资源最少（轻量），仅侧重于将I/O处理为简单的PCIe功能。每个虚拟功能都源自一个物理功能。设备硬件限制了虚拟功能的可能数量。一个以太网端口（物理设备）可与多个虚拟功能对应，然后这些虚拟功能可分配给不同的虚拟机。

在支持NFVIS的硬件上加速SR-IOV的建议驱动程序

Platform	网卡	NIC驱动程序
ENCS 54XX	背板交换机	i40e
ENCS 54XX	GE0-0和GE0-1	IGB
Catalyst 8200 uCPE	GE0-0和GE0-1	ixgbe
Catalyst 8200 uCPE	GE0-2和GE0-5	IGB

DPDK和SR-IOV的使用案例

DPDK首选项

特别是在网络流量主要以东-西流量（意味着它位于同一服务器内）的情况下，DPDK的性能优于SR-IOV。原理很简单：当流量在服务器内部管理而不需要访问网卡时，SR-IOV不会提供任何好处。事实上，SR-IOV可能会通过不必要地扩展流量路径和消耗NIC资源而导致效率低下。因此，对于内部服务器流量管理，使用DPDK是更有效的选择。

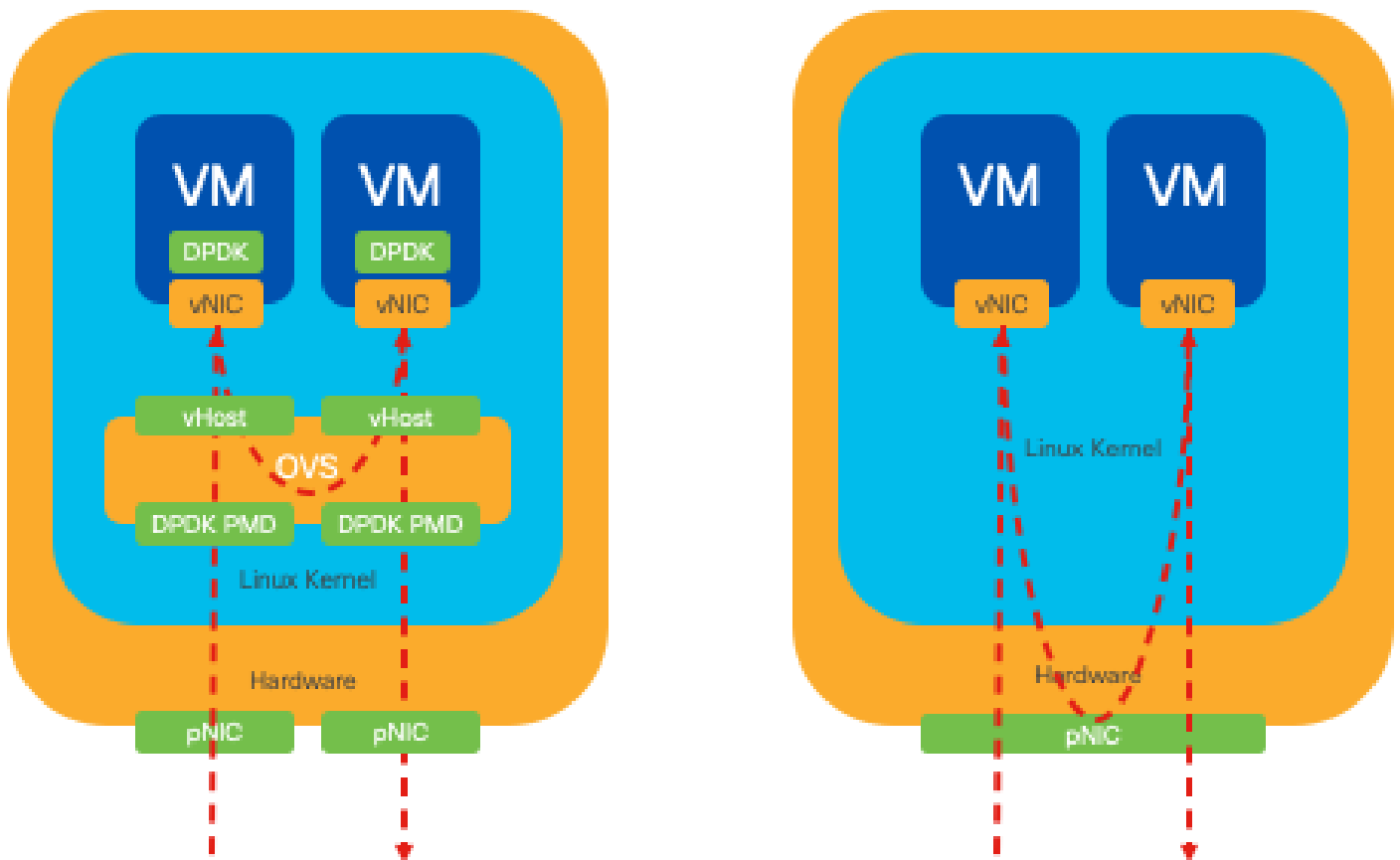


图 8.East-to-West流量中的DPDK和SR-IOV数据包遍历

SR-IOV首选项

在网络流量从北向南、甚至从东向西流动（尤其是在服务器之间）的情况下，使用SR-IOV比DPDK更具优势。对于服务器到服务器通信尤其如此。由于此类流量不可避免地必须流经NIC，因此选择使用DPDK增强的OVS可能会不必要地引入额外的复杂性和潜在的性能限制。因此，SR-IOV逐渐成为这些情况下的首选，为处理服务器间流量提供了直观、高效的途径。

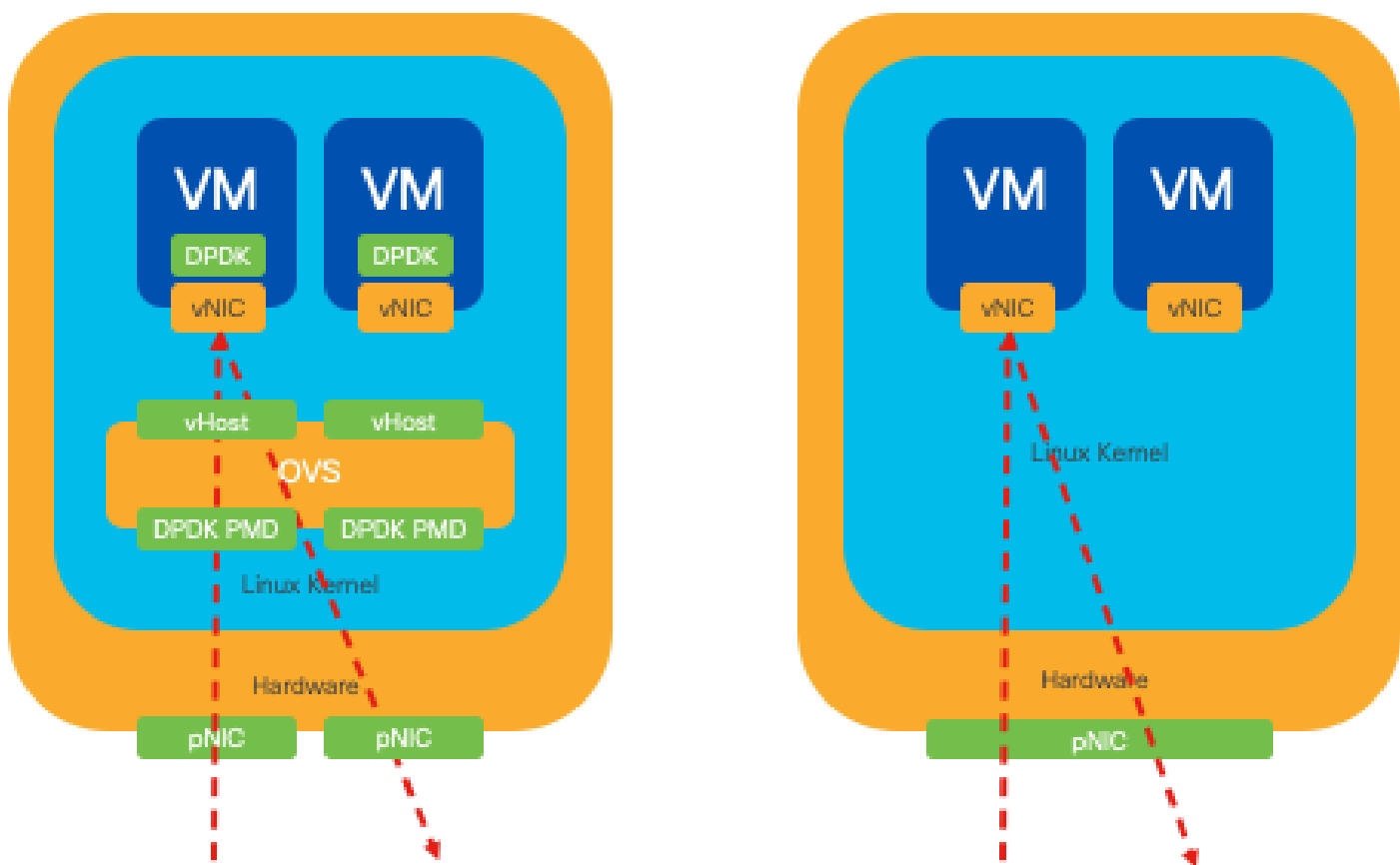
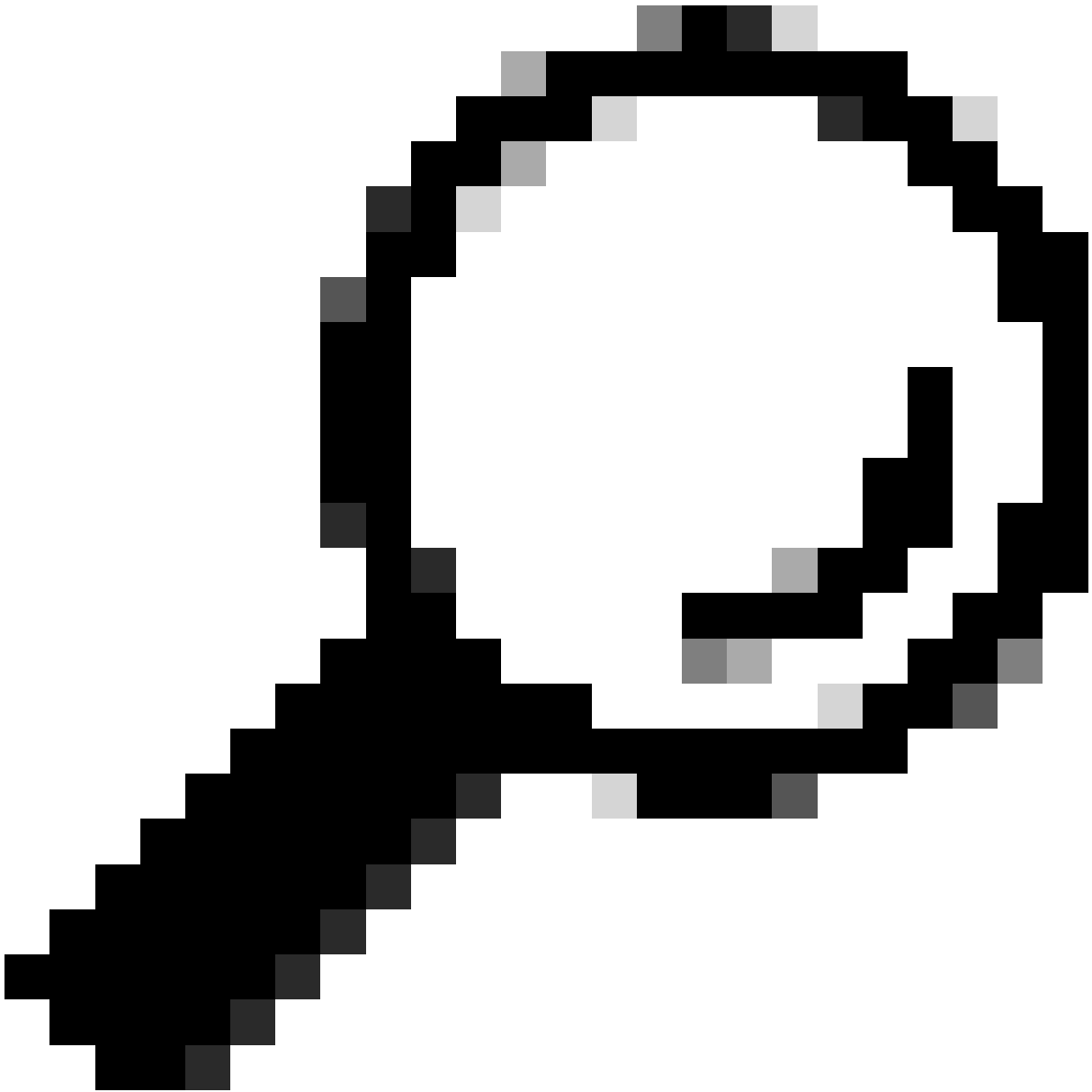


图 9北到南流量中的DPDK和SR-IOV数据包遍历



提示：请记住，通过将SR-IOV与DPDK集成到虚拟网络功能(VNF)中，可以增强基于SR-IOV的设置的性能，但不包括DPDK与之前描述的OVS结合使用的场景。

配置

启用DPDK

要从GUI启用DPDK，您必须导航到Configuration > Virtual Machine > Networking > Networks。进入菜单后，单击交换机激活该功能

Networks

Networks Information and Configuration

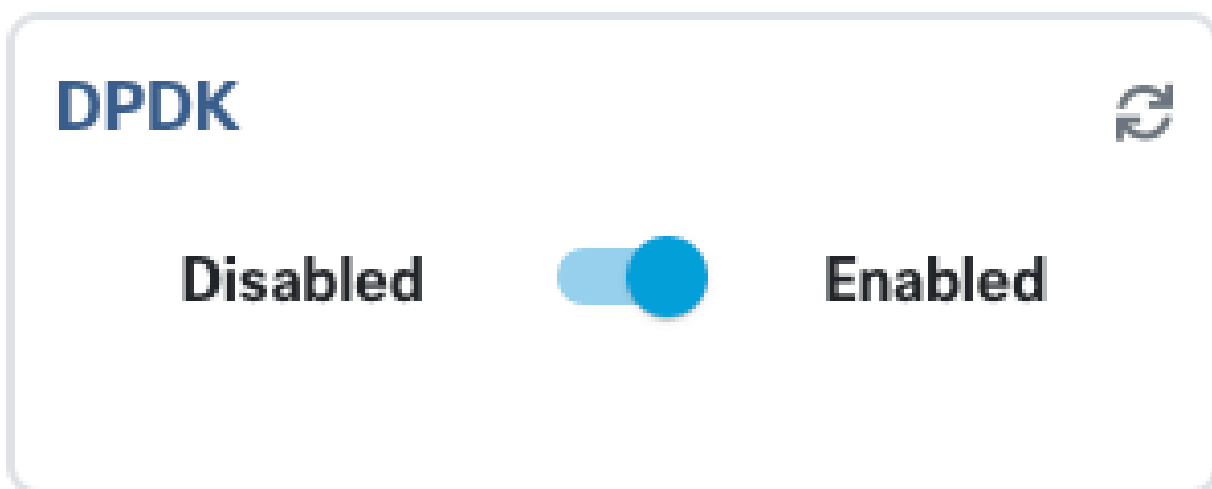
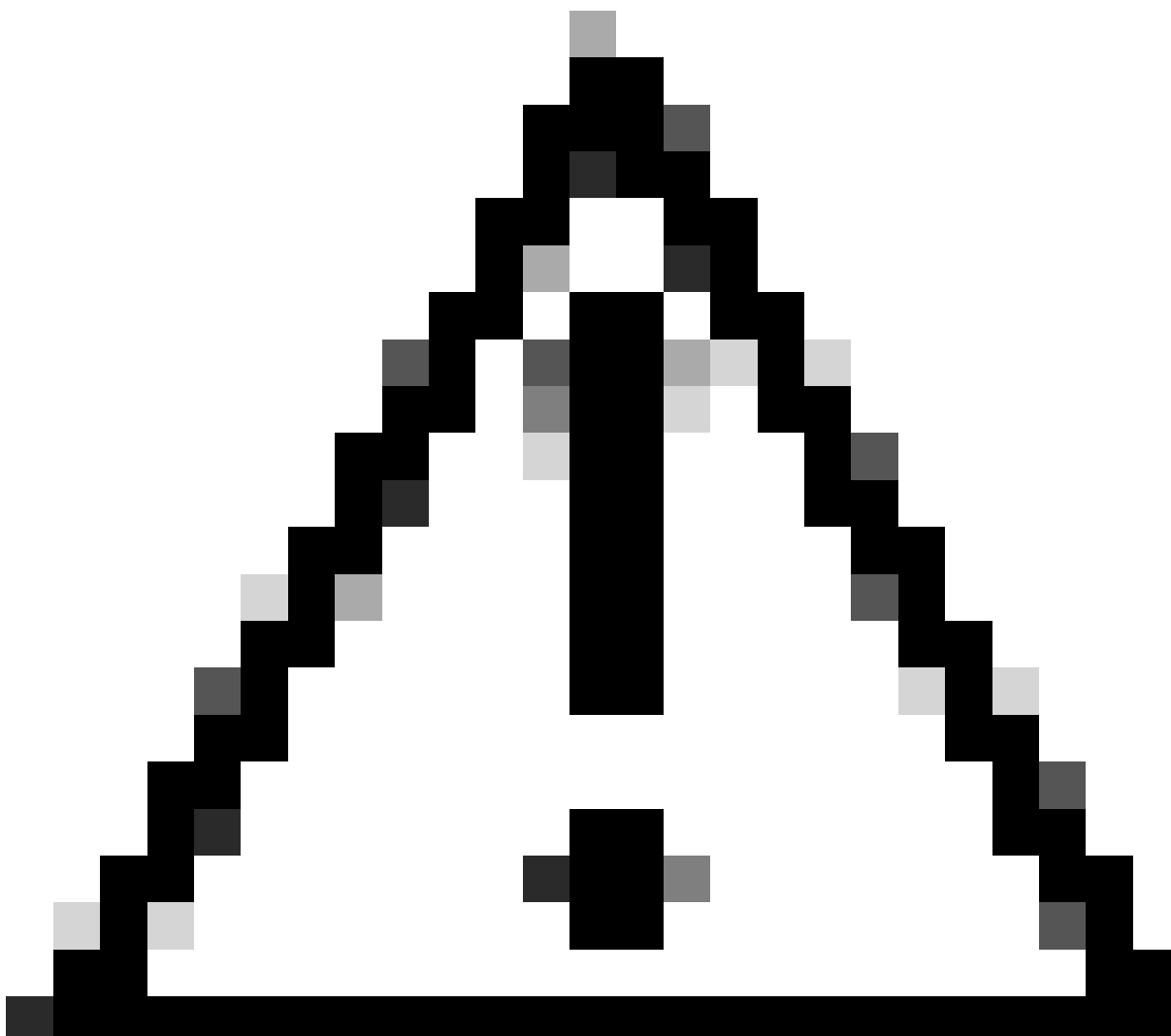


图 10.GUI上提供的幻灯片按钮用于DPDK激活

对于CLI，您必须在配置模式下从全局系统设置中启用它。

```
nfvis(config)# system settings dpdk enable
```



注意：除非从NFVIS执行出厂重置，否则无法禁用DPDK。

创建新网络并将其关联到新的OVS网桥

导航到配置 > 虚拟机 > 网络 > 网络。到达Networks页面后，点击Networks表左上方的加号(+)

#	Network	Mode	Vlan	Vlan-Range	Native Vlan	Bridge	Interface	Action
1	wan-net	trunk				wan-br	GE0-0	
2	wan2-net	trunk				wan2-br	GE0-1	
3	lan-net	trunk				lan-br	GE0-2	

图 11.NFVIS GUI中的网络表视图

命名网络并关联到新网桥。VLAN和接口绑定选项取决于网络基础设施的需求。

Add Network

Network *

inter-vnf-net

Mode *

trunk

Vlan

Vlan-Range

Native Vlan

1

Bridge *

Existing Create New

Bridge

inter-vnf-br

Interface

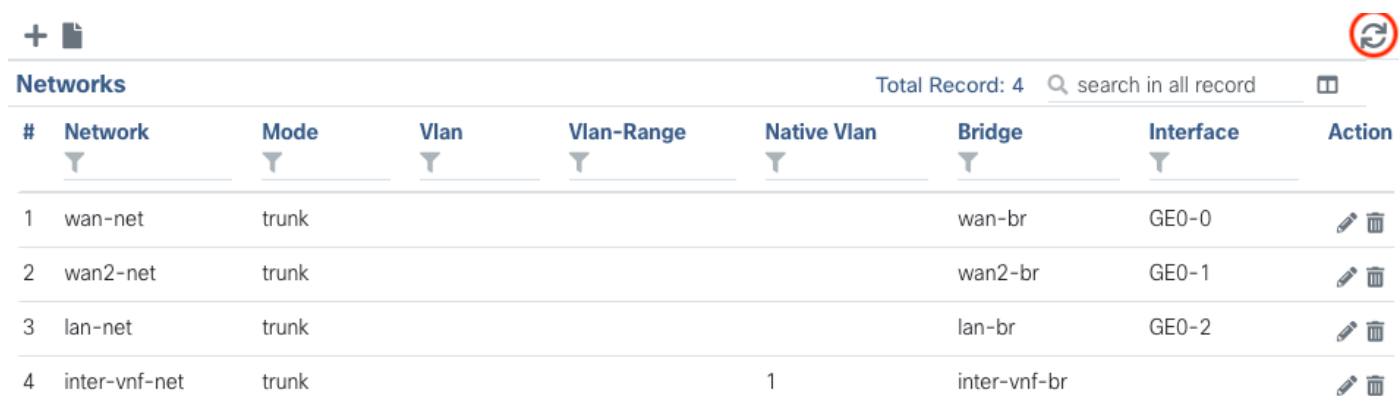
Submit

Cancel

Reset

图 12.在NFVIS GUI中创建虚拟网络的“添加网络”模式

单击submit按钮后，必须能够查看附加到Networks表的新建网络。











#	Network	Mode	Vlan	Vlan-Range	Native Vlan	Bridge	Interface	Action
1	wan-net	trunk				wan-br	GE0-0	 
2	wan2-net	trunk				wan2-br	GE0-1	 
3	lan-net	trunk				lan-br	GE0-2	 
4	inter-vnf-net	trunk			1	inter-vnf-br		 

图 13.NFVIS GUI中的网络表视图，其中“Refresh Icon”（刷新图标）位于右上角（以红色突出显示）

注意：如果在表格中未观察到新网络，请单击右上角的刷新按钮或刷新整个页面。

如果从CLI执行，每个网络和网桥都从配置模式创建，则工作流程与GUI版本相同。

1. 创建新网桥。

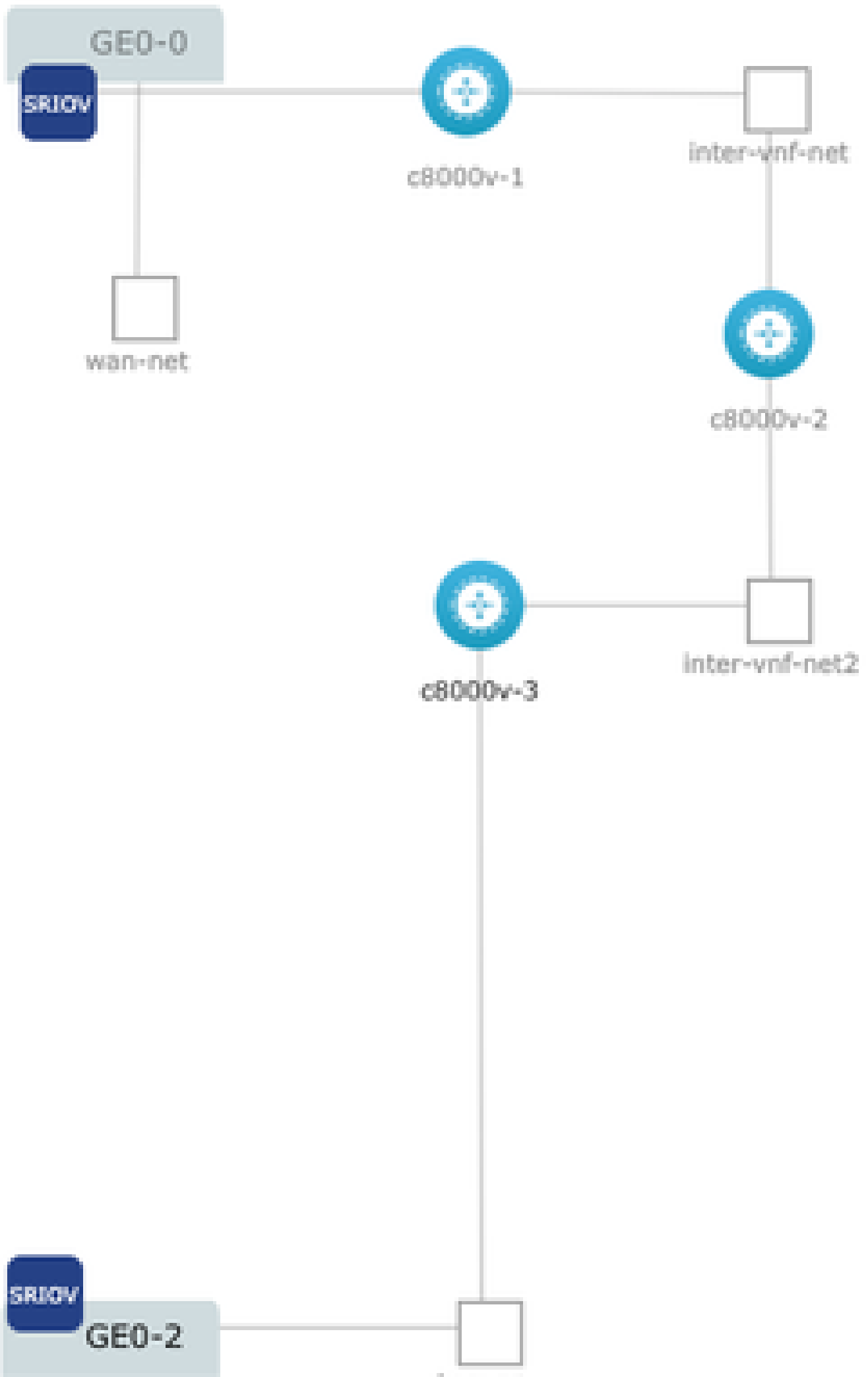
```
nfvis(config)# bridges bridge inter-vnf-br2
nfvis(config-bridge-inter-vnf-br2)# commit
```

2. 创建新网络并将其与之前创建的网桥关联

```
nfvis(config)# networks network inter-vnf-net2 bridge inter-vnf-br2 trunk true native-vlan 1
nfvis(config-network-inter-vnf-net2)# commit
```

连接VNF

要从网络拓扑或单个VFN部署开始，必须导航到配置 > 部署。您可以将VM或容器从选择列表拖到拓扑设计区域，以开始创建您的虚拟化基础设施。



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）。