

Comprensión de las redes virtuales NFVIS: OVS, DPDK y SR-IOV

Contenido

[Introducción](#)

[Componentes Utilizados](#)

[Descripción general de las redes en NFVIS](#)

[Plataforma ENCS54XX](#)

[Catalyst 8200 uCPE](#)

[Catalyst 8300 uCPE 1N20](#)

[Tecnologías de virtualización de redes](#)

[vSwitch abierto \(OVS\)](#)

[Puentes OVS](#)

[Déficits de cambio de contexto](#)

[Kit de desarrollo de plano de datos \(DPDK\)](#)

[Copia de datos](#)

[Paso a través de PCIe](#)

[Virtualización de E/S de raíz única \(SR-IOV\)](#)

[Funciones físicas \(PF\)](#)

[Funciones virtuales \(VF\)](#)

[Controladores recomendados para la aceleración de SR-IOV en hardware compatible con NFVIS](#)

[Casos prácticos de DPDK y SR-IOV](#)

[Preferencia DPDK](#)

[Preferencia de SR-IOV](#)

[Configuración](#)

[Habilitación de DPDK](#)

[Crear una red nueva y asociarla a un nuevo puente de OVS](#)

[Conexión de VNF](#)

[Artículos y documentación relacionados](#)

Introducción

Este documento describe el esquema de red virtual que la plataforma NFVIS proporciona para comunicar VNF en redes empresariales y de servicios.

Componentes Utilizados

La información de este documento se basa en estos componentes de hardware y software:

- ENCS5412 con NFVIS 4.7.1-FC4
- c8300 uCPE 1N20 que ejecuta NFVIS 4.12.1-FC2

La información que contiene este documento se creó a partir de los dispositivos en un ambiente de laboratorio específico. Todos los dispositivos que se utilizan en este documento se pusieron en funcionamiento con una configuración verificada (predeterminada). Si tiene una red en vivo, asegúrese de entender el posible impacto de cualquier comando.

Descripción general de las redes en NFVIS

Una red de gestión interna (int-mgmt-net) y un puente (int-mgmt-br) se utilizan internamente para la supervisión de VNF, asignando direcciones IP de gestión de la subred 10.20.0.0/24.

Plataforma ENCS54XX

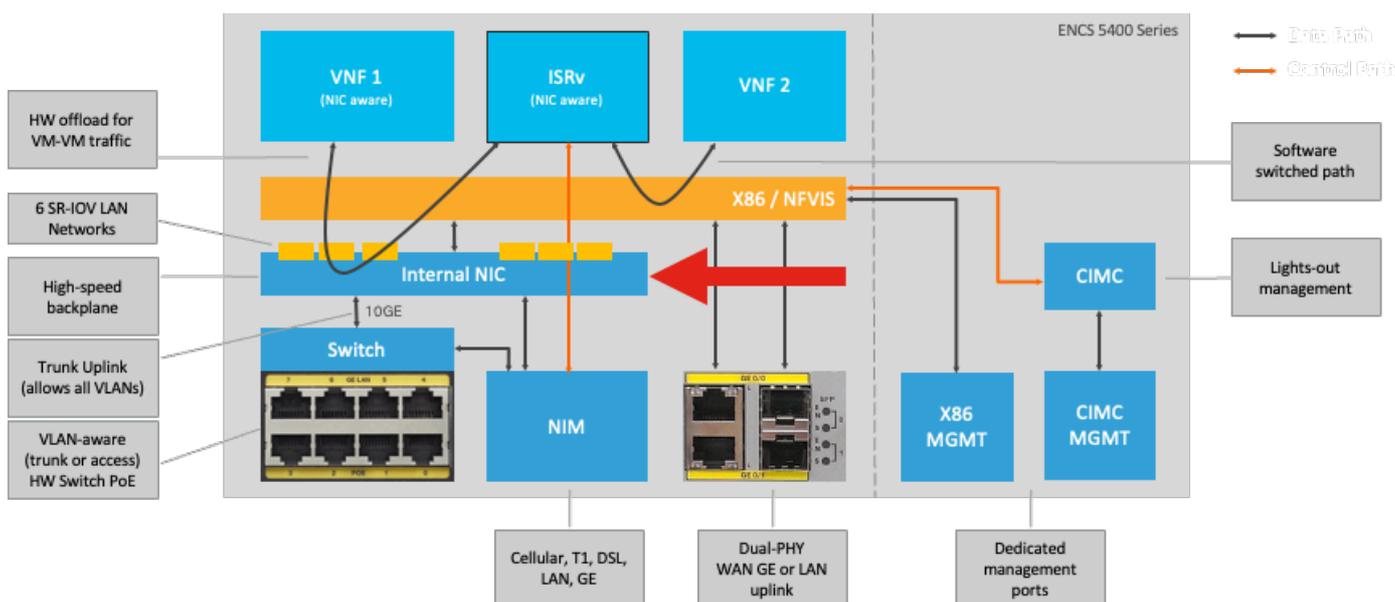


Figura 1. Conexiones internas del switch de hardware y NIC de enlace ascendente WAN/LAN

Catalyst 8200 uCPE

- Se puede acceder a NFVIS de forma predeterminada a través del puerto WAN o del puerto LAN GE0/2 para la gestión.
- La red WAN (wan-net y wan2-net) y el puente WAN (wan-br y wan2-br) están configurados para activar DHCP de forma predeterminada. GE0-0 está asociado al puente WAN y GE0-1 al puente WAN2 de forma predeterminada.
- La dirección IP de administración 192.168.1.1 en Catalyst 8200 UCPE es accesible a través de GE0-2.
- GE0-2 está asociado al puente LAN.
- Se crea una red de gestión interna (int-mgmt-net) y un puente (int-mgmt-br) que se utilizan internamente para supervisar el sistema.

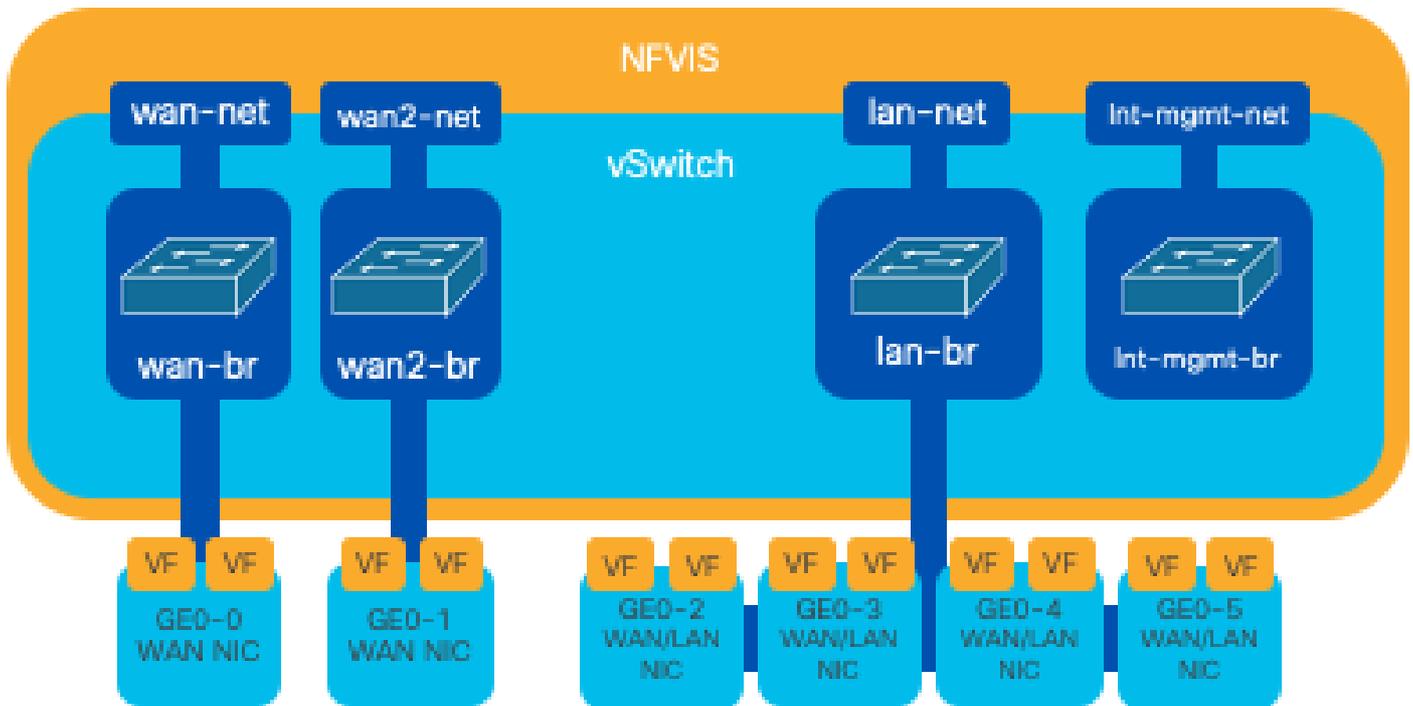


Figura 2 Puentes internos y switches virtuales asignados a las tarjetas NIC 8200

Catalyst 8300 uCPE 1N20

1. Se puede acceder a NFVIS de forma predeterminada a través de los puertos WAN FPGE (Gigabit Ethernet del panel frontal) o a través del puerto LAN GE0-2 para administración
2. La red WAN (wan-net) y un puente WAN (wan-br) están configurados de forma predeterminada para activar DHCP. GE0-0 está asociado de forma predeterminada al puente WAN
3. La red WAN (wan2-net) y un puente WAN (wan2-br) se crean de forma predeterminada, pero no están asociados a ningún puerto físico
4. GE0-2 está asociado con el puente LAN, todos los demás puertos no están asociados con OVS
5. La IP de administración 192.168.1.1 en C8300-uCPE es accesible a través de GE0-2
6. Se crea una red de gestión interna (int-mgmt-net) y un puente (int-mgmt-br) que se utilizan internamente para la supervisión del sistema.

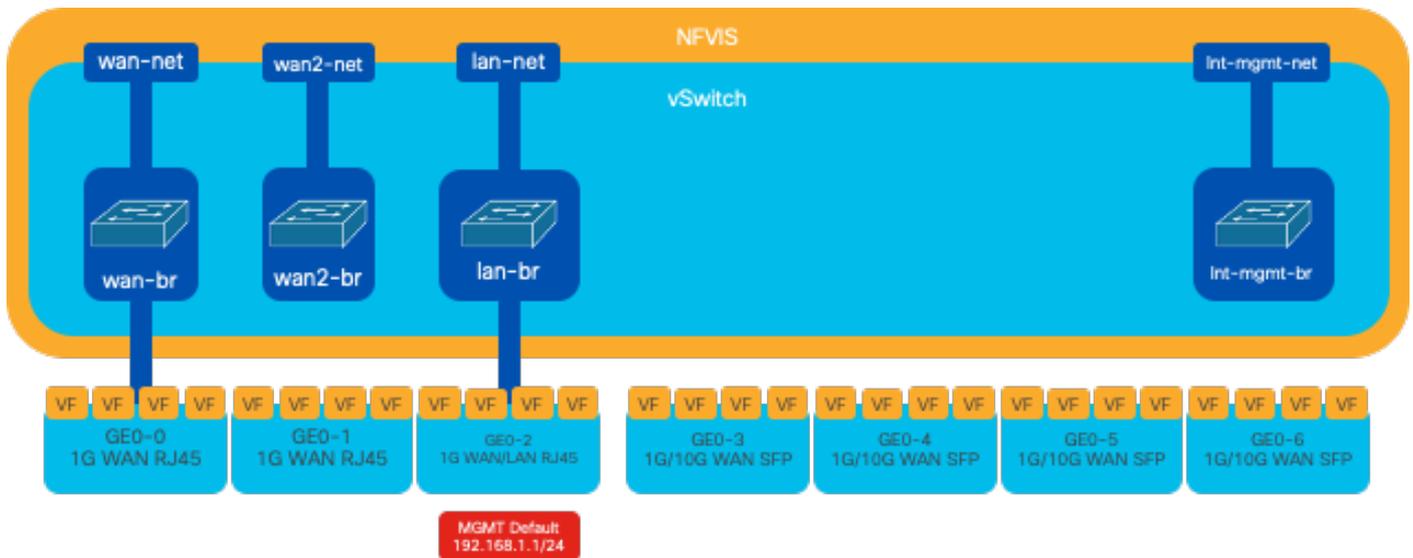


Figura 3. Puentes internos y switches virtuales asignados a las tarjetas NIC 8300

Tecnologías de virtualización de redes

vSwitch abierto (OVS)

Open vSwitch (OVS) es un switch virtual multicapa de código abierto diseñado para habilitar la automatización de la red mediante extensiones programáticas, a la vez que proporciona compatibilidad con interfaces y protocolos de gestión estándar, como NetFlow, sFlow, IPFIX, RSPAN, CLI, LACP y 802.1ag. Se utiliza ampliamente en grandes entornos virtualizados, especialmente con hipervisores, para gestionar el tráfico de red entre máquinas virtuales (VM). Permite la creación de políticas y topologías de red sofisticadas gestionadas directamente a través de la interfaz NFVIS, lo que proporciona un entorno versátil para la virtualización de las funciones de red.

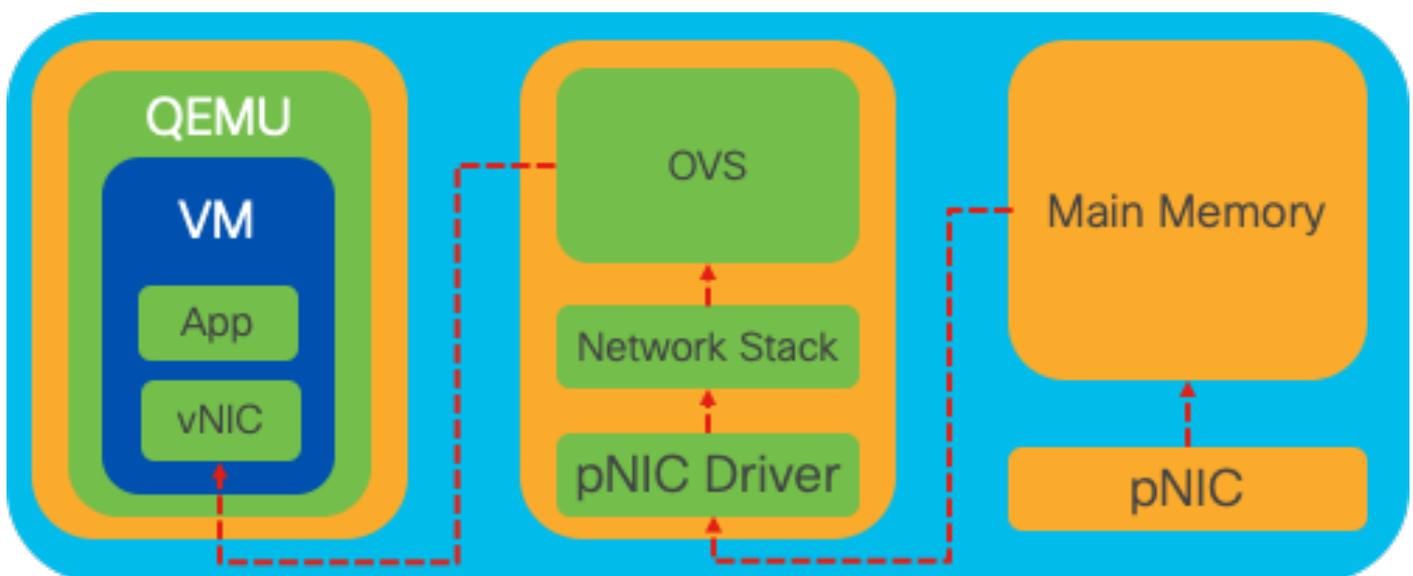


Figura 4 Configuración de OVS dentro del kernel de Linux

Puentes OVS

Utiliza puentes de red virtuales y reglas de flujo para reenviar paquetes entre hosts. Se comporta como un switch físico, solo virtualizado.

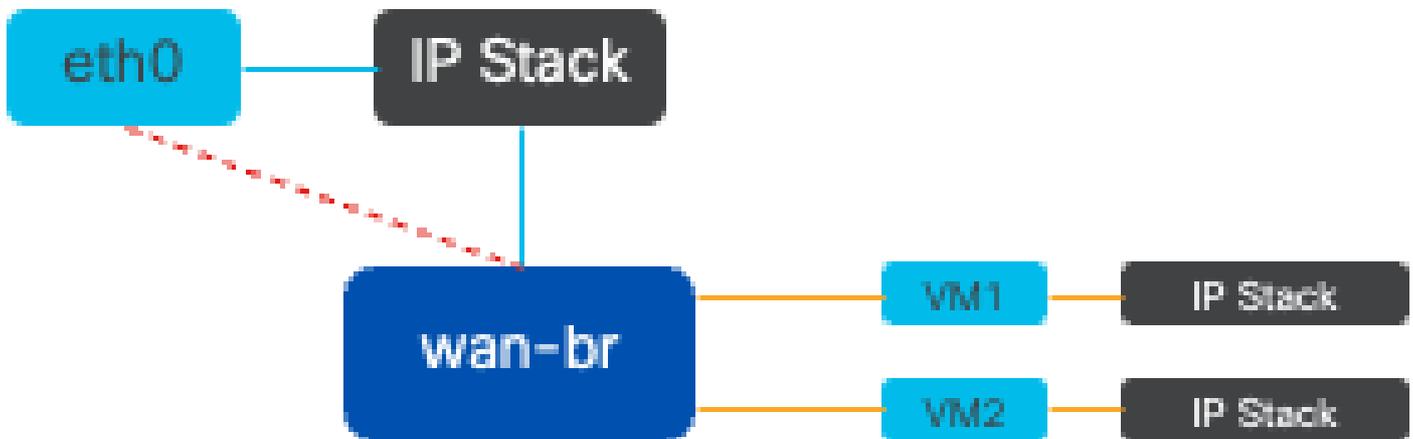


Figura 5. Ejemplo de implementación de 2 VM o VNF conectadas al puente wan-br

Déficits de cambio de contexto

Cuando un paquete de red llega a una tarjeta de interfaz de red (NIC), activa una interrupción, una señal al procesador que indica que necesita atención inmediata. La CPU hace una pausa en sus tareas actuales para gestionar la interrupción, un proceso conocido como procesamiento de interrupciones. Durante esta fase, la CPU, bajo el control del kernel del sistema operativo, lee el paquete de la NIC en la memoria y decide los siguientes pasos en función del destino y propósito del paquete. El objetivo es procesar o enrutar rápidamente el paquete a la aplicación deseada, minimizando la latencia y maximizando el rendimiento.

El cambio de contexto es el proceso por el cual una CPU pasa de ejecutar tareas en un entorno (contexto) a otro. Esto es particularmente relevante cuando se mueve entre el modo de usuario y el modo kernel:

- **Modo de usuario:** es un modo de procesamiento restringido en el que se ejecutan la mayoría de las aplicaciones. Las aplicaciones en modo de usuario no tienen acceso directo al hardware o a la memoria de referencia y deben comunicarse con el kernel del sistema operativo para realizar estas operaciones.
- **Modo kernel:** este modo otorga al sistema operativo acceso completo al hardware y a toda la memoria. El núcleo puede ejecutar cualquier instrucción de CPU y hacer referencia a cualquier dirección de memoria. El modo kernel es necesario para realizar tareas como administrar dispositivos de hardware, memoria y ejecutar llamadas del sistema.

Cuando una aplicación necesita realizar una operación que requiere privilegios a nivel del núcleo (como leer un paquete de red), se produce un cambio de contexto. La CPU pasa del modo de usuario al modo kernel para ejecutar la operación. Una vez completada, otro modificador de contexto devuelve la CPU al modo de usuario para continuar ejecutando la aplicación. Este

proceso de switching es fundamental para mantener la estabilidad y la seguridad del sistema, pero introduce una sobrecarga que puede afectar al rendimiento.

OVS se ejecuta principalmente en el espacio de usuarios del sistema operativo, lo que puede convertirse en un cuello de botella a medida que aumenta el rendimiento de los datos. Esto se debe a que se necesitan más switches de contexto para que la CPU pase al modo kernel para procesar los paquetes, lo que ralentiza el rendimiento. Esta limitación es particularmente notable en entornos con velocidades de paquetes altas o cuando el tiempo preciso es crucial. Para abordar estas limitaciones de rendimiento y satisfacer las demandas de las redes modernas de alta velocidad, se desarrollaron tecnologías como DPDK (kit de desarrollo de plano de datos) y SR-IOV (virtualización de E/S de raíz única).

Kit de desarrollo de plano de datos (DPDK)

DPDK es un conjunto de bibliotecas y controladores diseñados para acelerar las cargas de trabajo de procesamiento de paquetes en una amplia gama de arquitecturas de CPU. Al omitir la pila de red del núcleo tradicional (evitando el cambio de contexto), DPDK puede aumentar significativamente el rendimiento del plano de datos y reducir la latencia. Esto resulta especialmente beneficioso para las VNF de alto rendimiento que requieren una comunicación de baja latencia, lo que convierte a NFVIS en una plataforma ideal para las funciones de red sensibles al rendimiento.

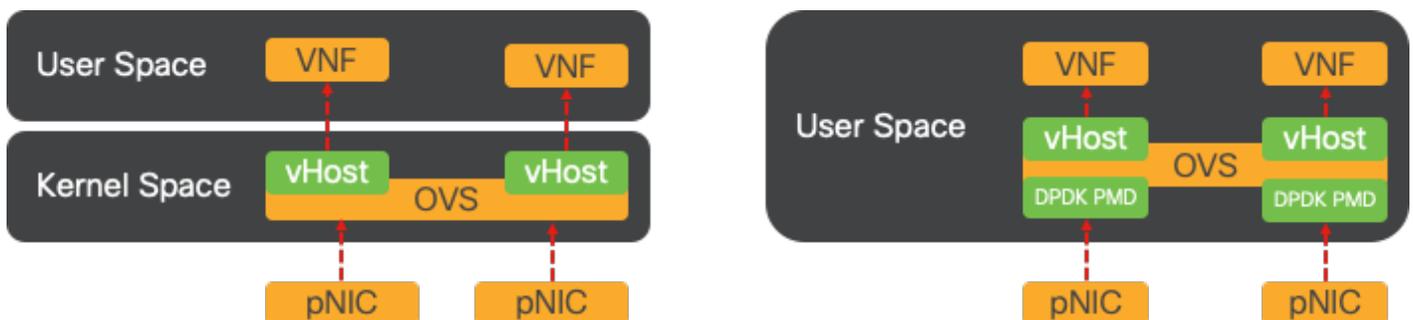


Figura 6. Optimizaciones de switching de contexto de OVS (izquierda) tradicional y de OVS (derecha) DPDK

La compatibilidad con DPDK para OVS se inició en NFVIS 3.10.1 para ENCS y 3.12.2 para otras plataformas.

- Rendimiento de la cadena de servicios cerca de SRIOV, mejor que los OVS no DPDK.
- Se requiere controlador Virtio para VNF.
- Plataformas Soportadas:
- ENCS 3.10.1 en adelante.
- UCSE, UCS-C, CSP5K 3.12.1 en adelante.
- DPDK para canales de puerto admitidos desde 4.12.1.
- Captura de paquetes/tráfico: no compatible con DPDK.
- Tráfico de extensión en PNIC: no compatible con DPDK.
- Después de activar OVS-DPDK, no se puede desactivar como una función individual. La única manera de inhabilitar DPDK sería un restablecimiento de fábrica.

Copia de datos

Los enfoques de red tradicionales a menudo requieren que los datos se copien varias veces antes de llegar a su destino en la memoria de la VM. Por ejemplo, un paquete se debe copiar de la NIC al espacio del núcleo, luego al espacio del usuario para ser procesado por un switch virtual (como OVS) y, finalmente, a la memoria de la VM. Cada operación de copia conlleva un retraso y aumenta el uso de la CPU a pesar de las mejoras de rendimiento que ofrece DPDK al omitir la pila de red de núcleos.

Estas sobrecargas incluyen copias de memoria y el tiempo de procesamiento necesario para manejar los paquetes en el espacio de usuario antes de que se puedan reenviar a la VM. PCIe Passthrough (Paso a través de PCIe) y SR-IOV afrontan estos cuellos de botella permitiendo que un dispositivo de red físico (como una NIC) se comparta directamente entre varias VM sin necesidad de involucrar al sistema operativo del host en la misma medida que los métodos de virtualización tradicionales.

Paso a través de PCIe

La estrategia consiste en omitir el hipervisor para permitir el acceso directo de las funciones de red virtual (VNF) a una tarjeta de interfaz de red (NIC), con lo que se logra un rendimiento casi máximo. Este enfoque se conoce como paso a través de PCI, que permite que una NIC completa se dedique a un sistema operativo invitado sin la intervención de un hipervisor. En esta configuración, la máquina virtual funciona como si estuviera conectada directamente a la NIC. Por ejemplo, con dos tarjetas NIC disponibles, cada una se puede asignar exclusivamente a diferentes VNF, proporcionándoles acceso directo.

Sin embargo, este método tiene un inconveniente: si solo hay dos NIC disponibles y las utilizan exclusivamente dos VNF independientes, cualquier VNF adicional, como una tercera, se quedaría sin acceso a NIC debido a la falta de una NIC dedicada disponible para ella. Una solución alternativa consiste en utilizar la virtualización de E/S de raíz única (SR-IOV).

Virtualización de E/S de raíz única (SR-IOV)

Es una especificación que permite que un único dispositivo PCI físico, como una tarjeta de interfaz de red (NIC), aparezca como varios dispositivos virtuales independientes. Esta tecnología proporciona acceso directo de VM a los dispositivos de red físicos, lo que reduce la sobrecarga y mejora el rendimiento de E/S. Funciona dividiendo un único dispositivo PCIe en varios segmentos virtuales, cada uno asignable a diferentes VM o VNF, resolviendo de forma eficaz la limitación causada por un número finito de NIC. Estos segmentos virtuales, conocidos como funciones virtuales (VF), permiten compartir recursos de NIC entre varias VNF. La función física (PF) hace referencia al componente físico real que facilita las capacidades de SR-IOV.

Al aprovechar SR-IOV, NFVIS puede asignar recursos de NIC dedicados a VNF específicos, lo que garantiza un alto rendimiento y una baja latencia al facilitar el acceso directo a memoria (DMA) de los paquetes de red directamente en la memoria de VM correspondiente. Este enfoque minimiza la participación de la CPU para que solo procese los paquetes, reduciendo así el uso de la CPU. Esto resulta especialmente útil para aplicaciones que requieren un ancho de banda

garantizado o que tienen estrictos requisitos de rendimiento.

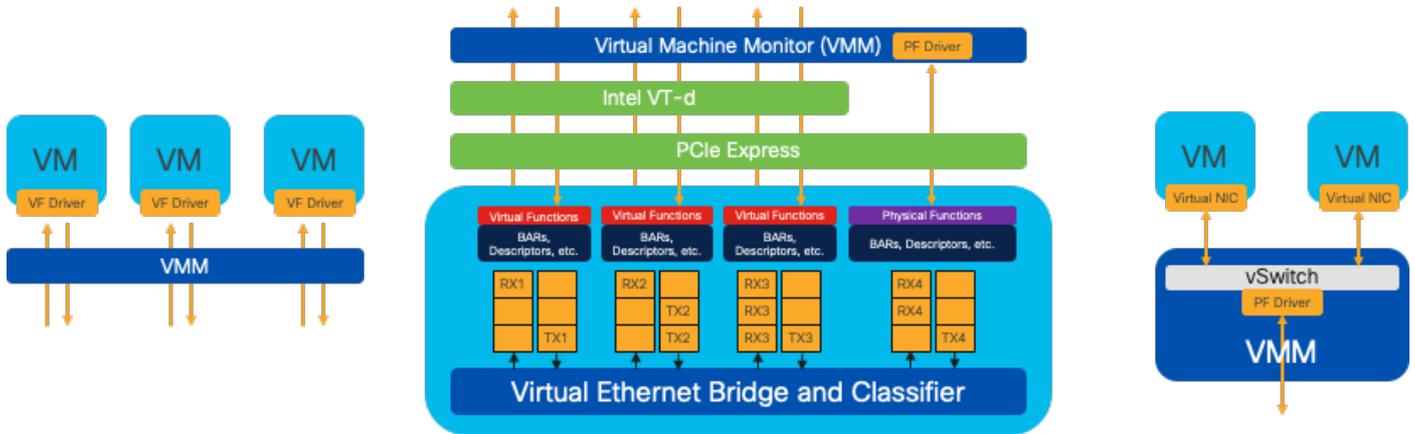


Figura 7 Separación de recursos PCIe de NFVIS SR-IOV mediante funciones de hardware

Funciones físicas (PF)

Se trata de funciones PCIe completas que hacen referencia a una caja de hardware específica que proporciona funciones de red específicas; se trata de funciones PCIe completas que se pueden detectar, gestionar y manipular como cualquier otro dispositivo PCIe. Entre las funciones físicas se incluyen las capacidades SR-IOV que se pueden utilizar para configurar y controlar un dispositivo PCIe.

Funciones virtuales (VF)

Se trata de funciones simplificadas con recursos de configuración mínimos (ligero), centradas únicamente en el procesamiento de E/S como funciones PCIe sencillas. Cada función virtual se origina a partir de una función física. El hardware del dispositivo limita el posible número de funciones virtuales. Un puerto Ethernet, el dispositivo físico, puede corresponder a numerosas funciones virtuales que, a continuación, se pueden asignar a distintas máquinas virtuales.

Controladores recomendados para la aceleración de SR-IOV en hardware compatible con NFVIS

Platform	NIC(s)	Controlador de NIC
ENCS 54XX	Switch de backplane	i40e
ENCS 54XX	GE0-0 y GE0-1	IGB
Catalyst 8200 uCPE	GE0-0 y GE0-1	ixgbe
Catalyst 8200 uCPE	GE0-2 y GE0-5	IGB

Casos prácticos de DPDK y SR-IOV

Preferencia DPDK

Especialmente en escenarios donde el tráfico de red fluye principalmente de este a oeste (lo que

significa que permanece dentro del mismo servidor), DPDK supera a SR-IOV. La lógica es sencilla: cuando el tráfico se gestiona internamente dentro del servidor sin necesidad de acceder a la NIC, SR-IOV no aporta ninguna ventaja. De hecho, SR-IOV podría generar ineficiencias ampliando innecesariamente la ruta de tráfico y consumiendo recursos de NIC. Por lo tanto, para la gestión del tráfico de servidores internos, aprovechar DPDK es la opción más eficaz.

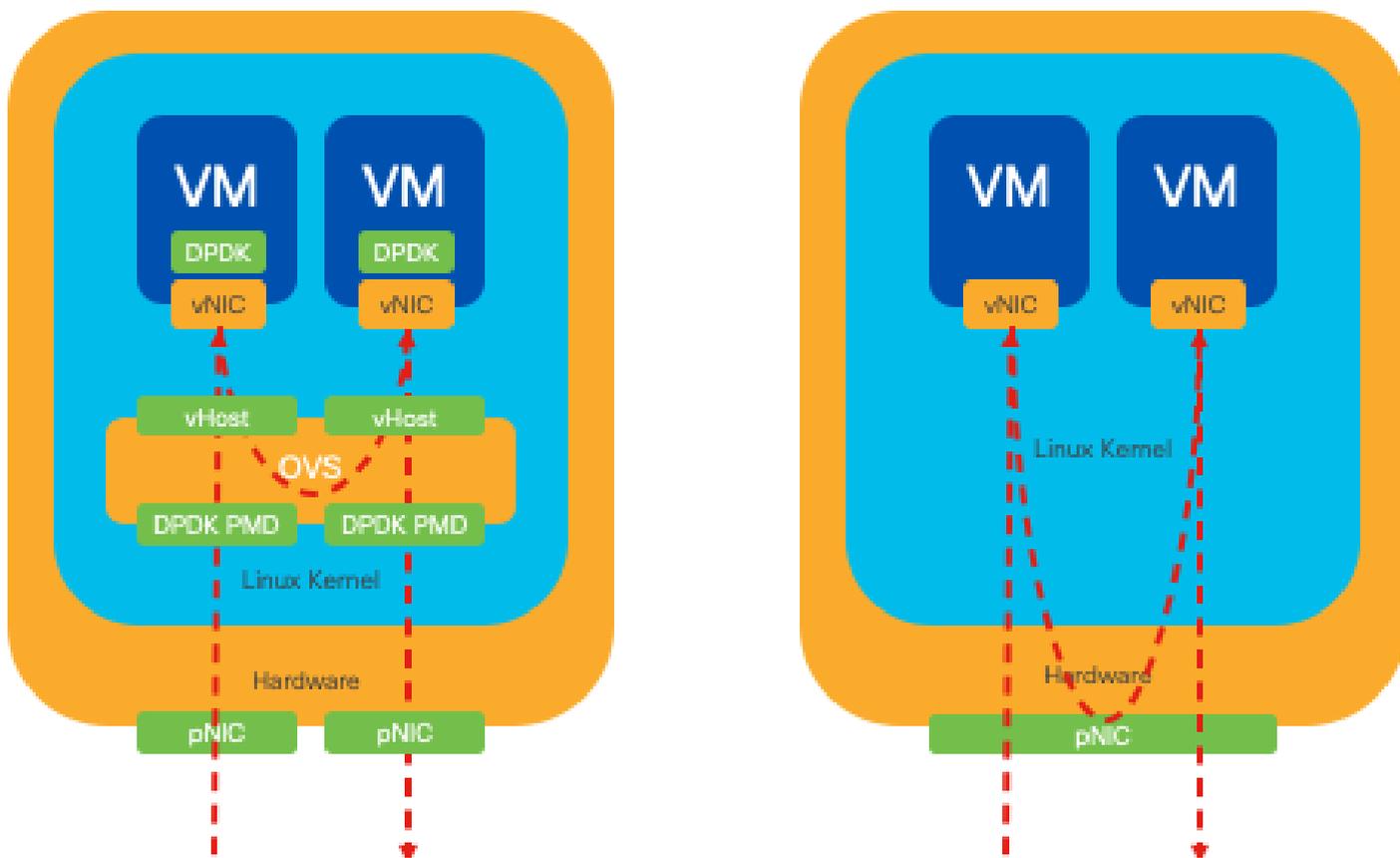


Figura 8 Traversal de paquetes DPDK y SR-IOV en tráfico de este a oeste

Preferencia de SR-IOV

En situaciones en las que el tráfico de red fluye de norte a sur, o incluso de este a oeste pero específicamente entre servidores, el uso de SR-IOV resulta más ventajoso que DPDK. Esto es especialmente cierto en el caso de la comunicación entre servidores. Dado que dicho tráfico tiene que atravesar inevitablemente la NIC, la elección de OVS mejorados con DPDK podría introducir innecesariamente una complejidad adicional y posibles limitaciones de rendimiento. Por lo tanto, SR-IOV surge como la opción preferible en estas circunstancias, ya que ofrece una ruta sencilla y eficaz para gestionar el tráfico entre servidores.

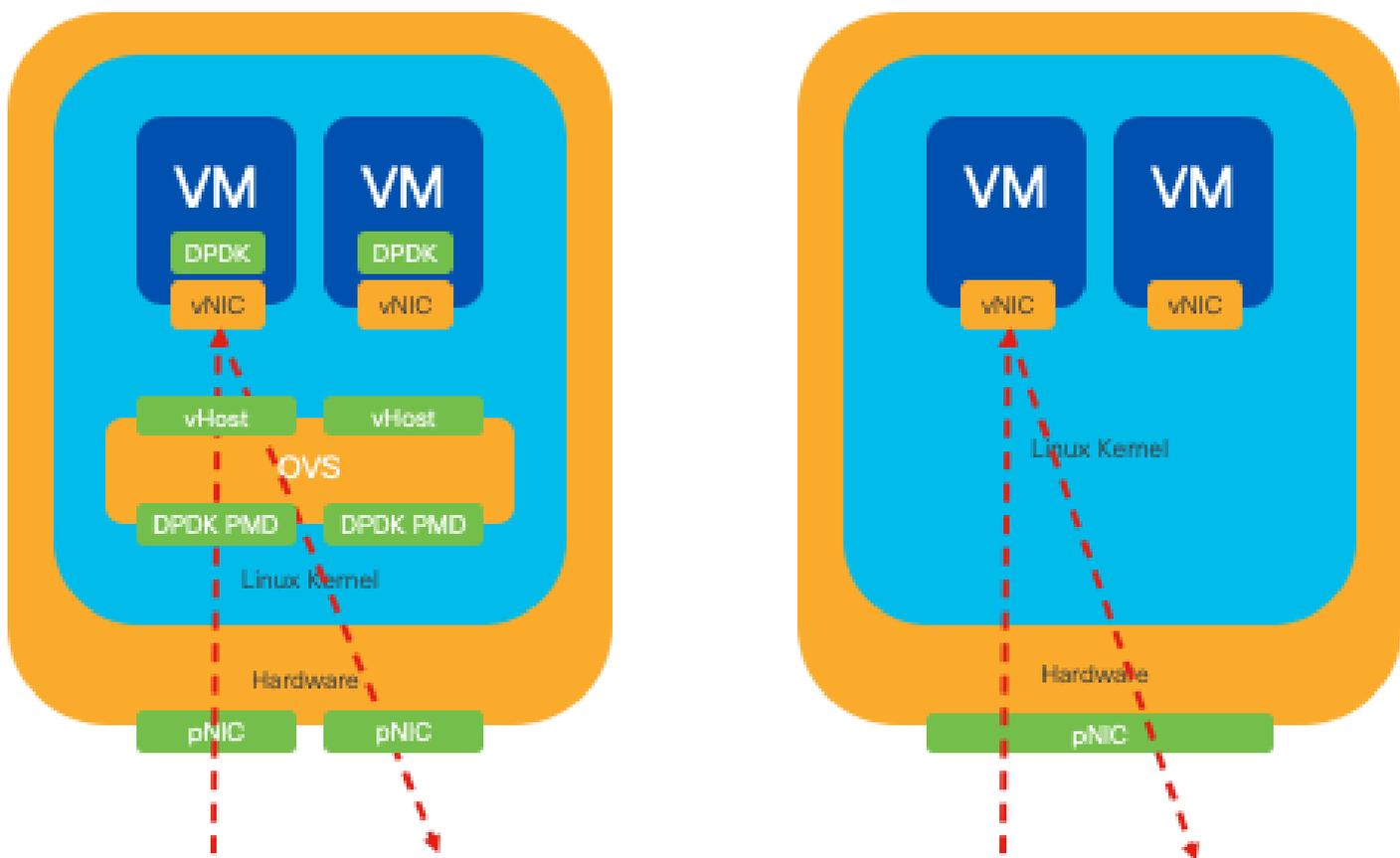
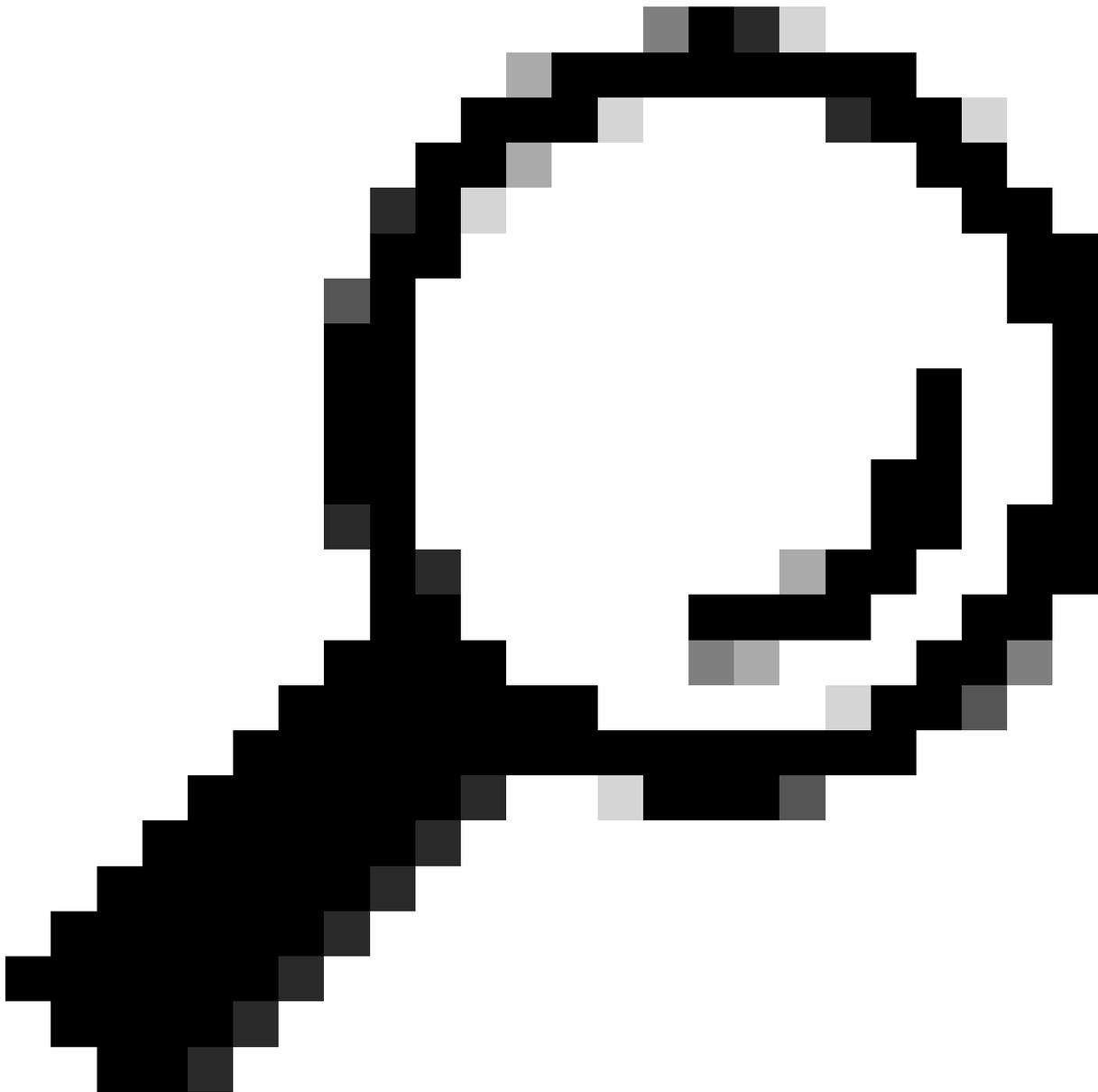


Figura 9. Traversal de paquetes DPDK y SR-IOV en tráfico de norte a sur



Sugerencia: recuerde que es posible mejorar el rendimiento de una configuración basada en SR-IOV integrando SR-IOV con DPDK en una función de red virtual (VNF), excepto en el caso en el que DPDK se utiliza junto con OVS, como se ha descrito anteriormente.

Configuración

Habilitación de DPDK

Para habilitar DPDK desde la GUI, debe navegar hasta Configuration > Virtual Machine > Networking > Networks. Una vez que esté en el menú, haga clic en el interruptor para activar la función

Networks

Networks Information and Configuration

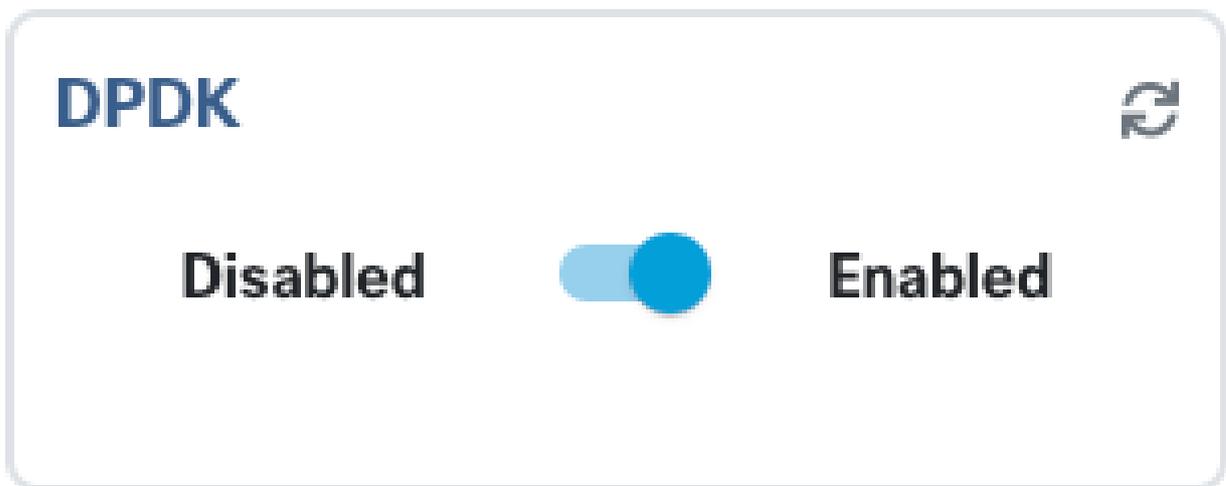
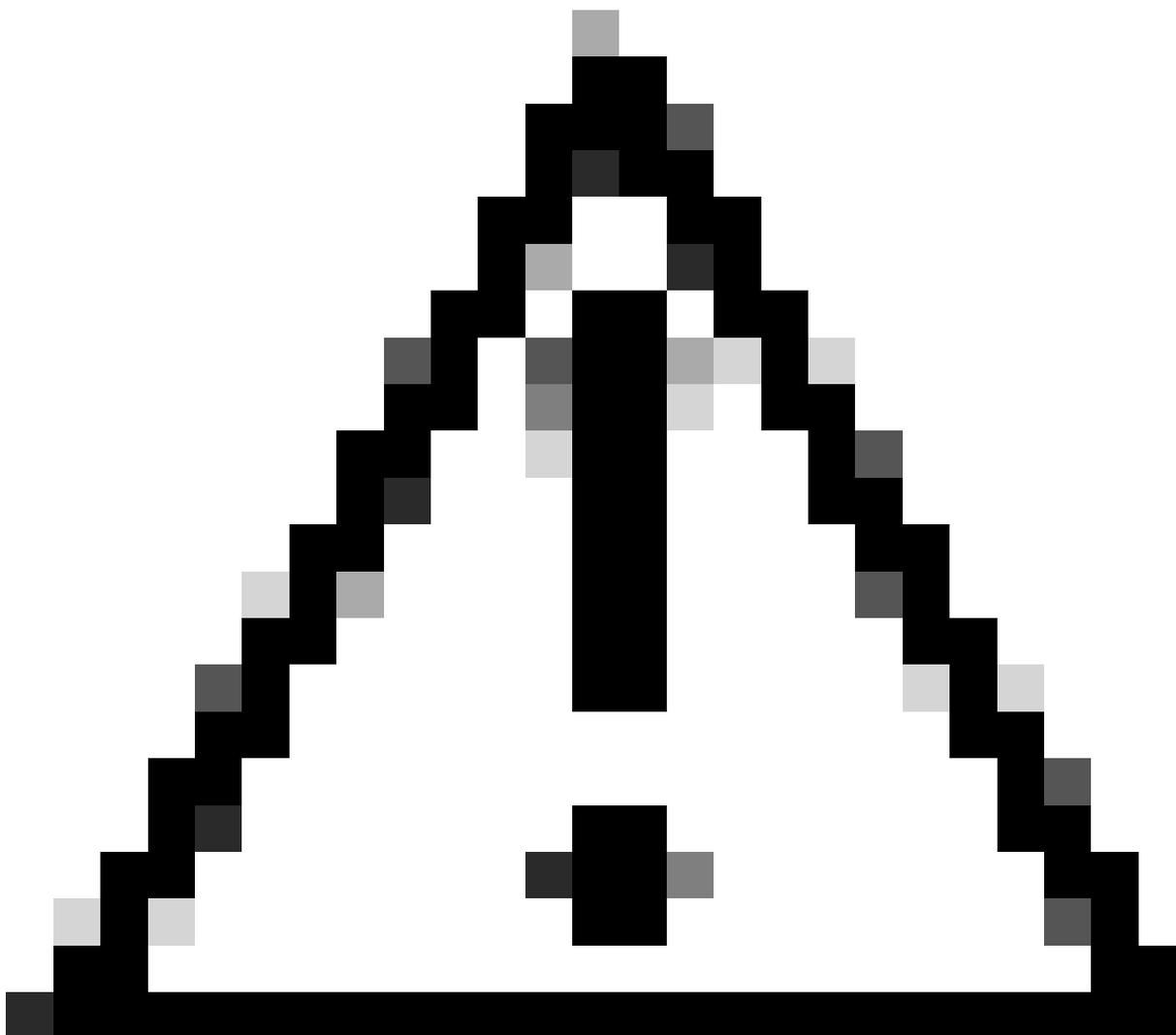


Figura 10. Botón de diapositiva disponible en la interfaz gráfica de usuario para la activación de DPDK

Para la CLI, debe habilitarla desde la configuración global del sistema en el modo de configuración.

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



Precaución: DPDK no se puede deshabilitar a menos que se realice un restablecimiento de fábrica desde NFVIS.

Crear una red nueva y asociarla a un nuevo puente de OVS

Vaya a Configuration > Virtual Machine > Networking > Networks. En la página Redes, haga clic en el signo más (+) superior izquierdo de la tabla Redes.

#	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	 

Figura 11. Vista de la tabla de redes desde la GUI de NFVIS

Dé un nombre a la red y asóciela a un nuevo puente. Las opciones de enlace de interfaz y VLAN pueden depender de las necesidades de infraestructura de la red.

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

Figura 12. Modal "Agregar red" para crear redes virtuales en la GUI de NFVIS

Después de hacer clic en el botón submit, debe poder revisar la red recién creada anexada a la tabla Networks.



The screenshot shows the 'Networks' table in the NFVIS GUI. The table has 9 columns: #, Network, Mode, Vlan, Vlan-Range, Native Vlan, Bridge, Interface, and Action. There are 4 records listed. A red refresh icon is highlighted in the top right corner of the table area.

#	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		 

Figura 13. Vista de la tabla de redes de la GUI de NFVIS, donde el icono de actualización se encuentra en la esquina superior derecha (resaltado en rojo)

Nota: si no se observa la nueva red en la tabla, haga clic en el botón de actualización superior derecho o actualice toda la página.

Si se realiza dentro de la CLI, cada red y puente se crea a partir del modo de configuración, el flujo de trabajo es el mismo que la versión de la GUI.

1. Cree el nuevo puente.

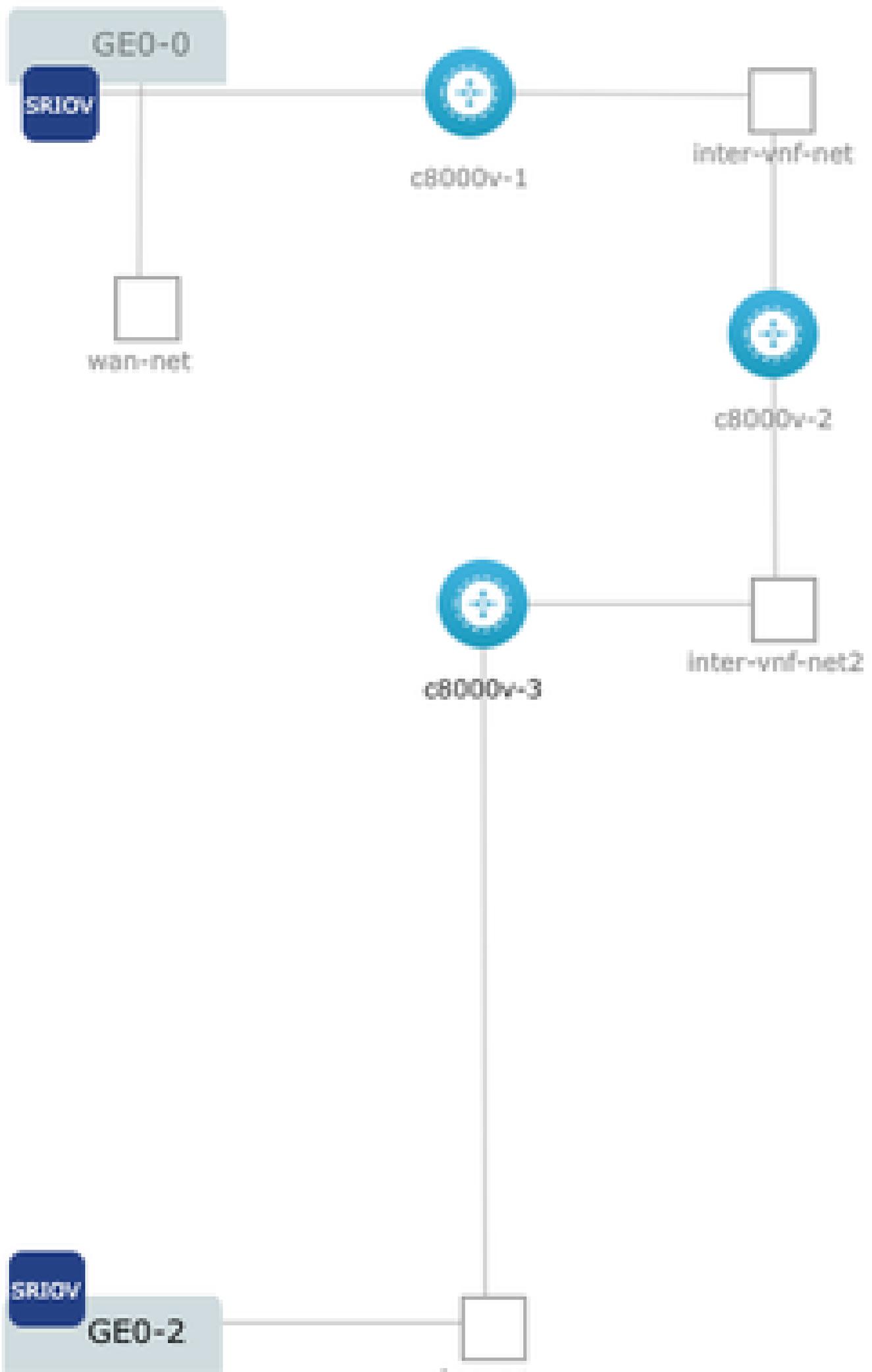
```
nfvis(config)# bridges bridge inter-vnf-br2
nfvis(config-bridge-inter-vnf-br2)# commit
```

2. Cree una nueva red y asóciela al puente creado anteriormente

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

Conexión de VNF

Para comenzar con una topología de red o una implementación de VFN única, debe navegar hasta Configuración > Implementar. Puede arrastrar una VM o un contenedor desde la lista de selección hasta el área de creación de topologías para comenzar a crear su infraestructura virtualizada.



Acerca de esta traducción

Cisco ha traducido este documento combinando la traducción automática y los recursos humanos a fin de ofrecer a nuestros usuarios en todo el mundo contenido en su propio idioma.

Tenga en cuenta que incluso la mejor traducción automática podría no ser tan precisa como la proporcionada por un traductor profesional.

Cisco Systems, Inc. no asume ninguna responsabilidad por la precisión de estas traducciones y recomienda remitirse siempre al documento original escrito en inglés (insertar vínculo URL).