

# 瞭解NFVIS虛擬網路：OVS、DPDK和SR-IOV

## 目錄

---

### [簡介](#)

### [採用元件](#)

### [NFVIS中的網路概述](#)

[ENCS54XX平台](#)

[Catalyst 8200 uCPE](#)

[Catalyst 8300 uCPE 1N20](#)

### [網路虛擬化技術](#)

[開放式vSwitch \(OVS\)](#)

[OVS網橋](#)

[環境切換缺陷](#)

[資料平面開發套件\(DPDK\)](#)

[資料複製](#)

[PCIe傳輸](#)

[單一根I/O虛擬化\(SR-IOV\)](#)

[物理功能\(PF\)](#)

[虛擬功能\(VE\)](#)

[在支援NFVIS的硬體上建議的SR-IOV加速驅動程式](#)

[DPDK和SR-IOV的使用案例](#)

[DPDK偏好設定](#)

[SR-IOV偏好設定](#)

### [組態](#)

[啟用DPDK](#)

[建立新網路並將其與新的OVS網橋關聯](#)

[連線VNE](#)

### [相關文章和文檔](#)

---

## 簡介

本文檔介紹NFVIS平台為企業和服務網路中通訊VNF而提供的虛擬網路方案。

## 採用元件

本檔案中的資訊是根據以下硬體和軟體元件而定：

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

本文中的資訊是根據特定實驗室環境內的裝置所建立。文中使用到的所有裝置皆從已清除（預設）的組態來啟動。如果您的網路運作中，請確保您瞭解任何指令可能造成的影響。

# NFVIS中的網路概述

內部管理網路(int-mgmt-net)和網橋(int-mgmt-br)在內部用於VNF監控，從10.20.0.0/24子網分配管理IP地址。

## ENC54XX平台

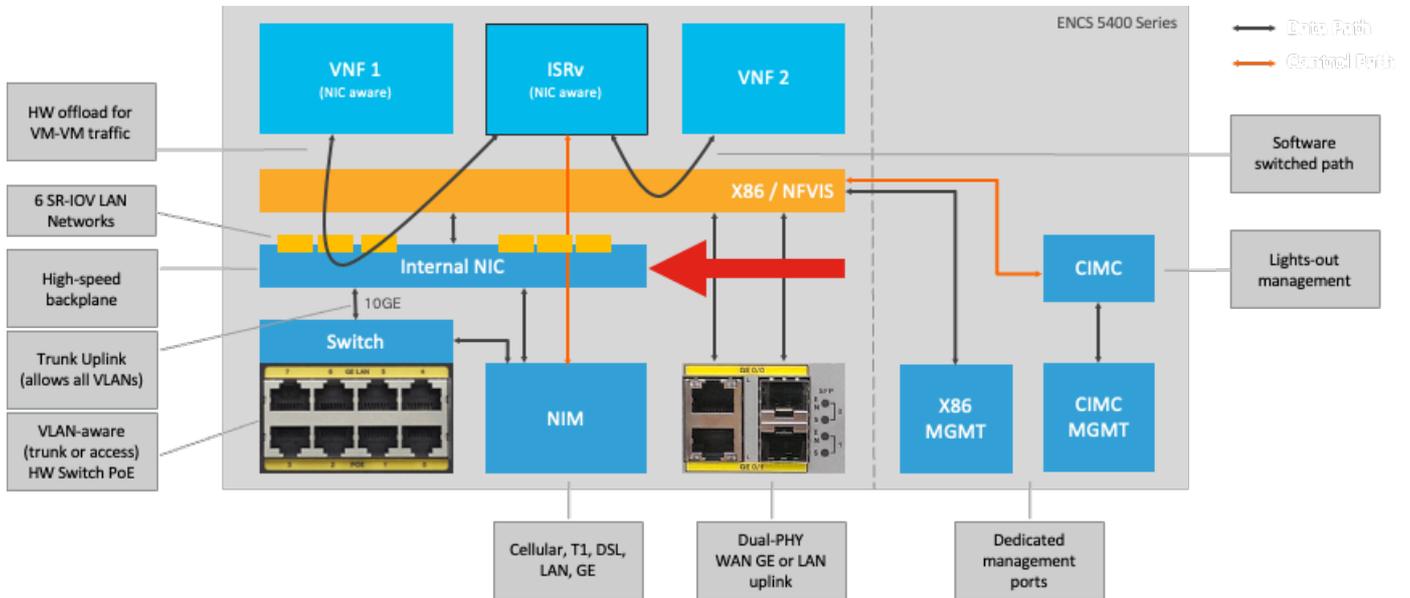


圖1.硬體交換機和WAN/LAN上行鏈路NIC內部連線

## Catalyst 8200 uCPE

- 預設情況下，可以透過WAN埠或GE0/2 LAN埠訪問NFVIS以進行管理。
- 預設情況下，WAN網路 ( wan-net和wan2-net ) 和WAN網橋 ( wan-br和wan2-br ) 設定為啟用DHCP。預設情況下，GE0-0與WAN網橋關聯，而GE0-1與WAN2網橋關聯。
- Catalyst 8200 UCPE上的管理IP地址192.168.1.1可以透過GE0-2訪問。
- GE0-2與LAN網橋關聯。
- 建立內部管理網路(int-mgmt-net)和網橋(int-mgmt-br)，並在內部用於系統監控。

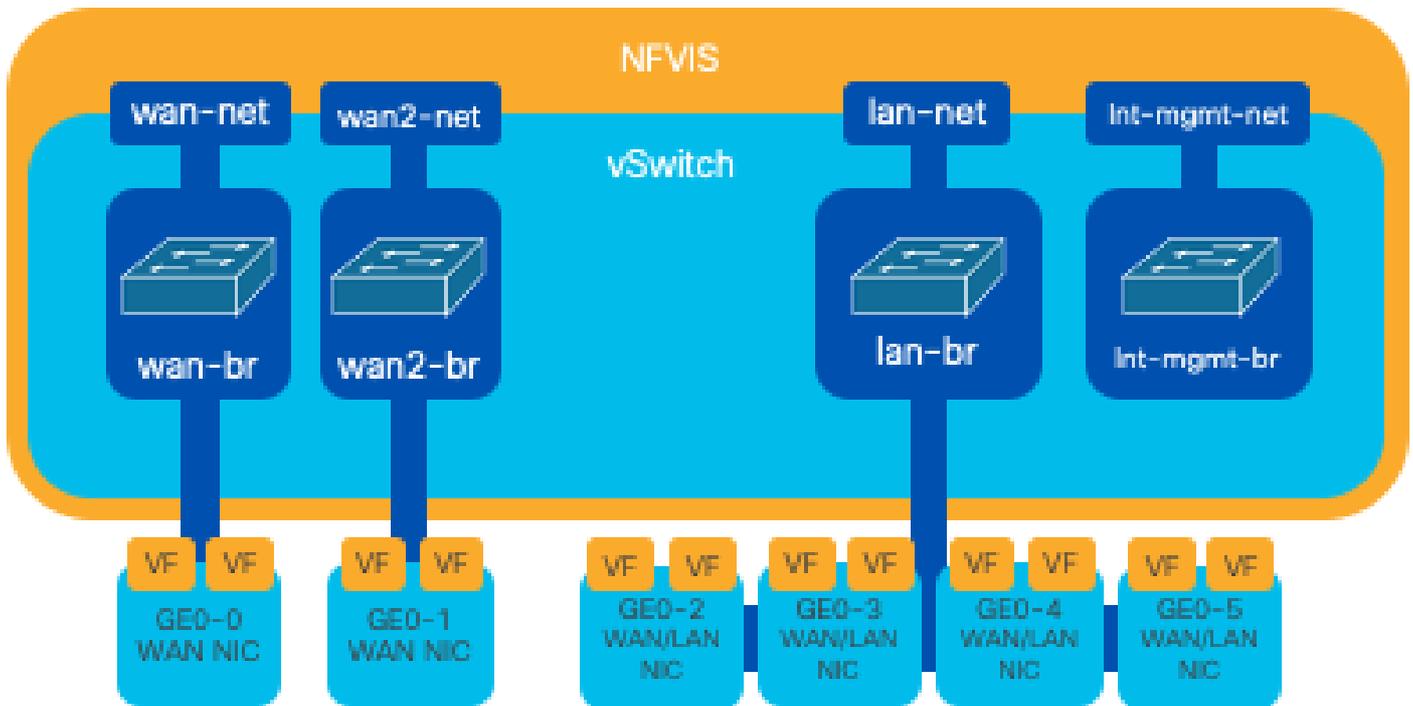


圖2. 分配給8200 NIC的內部橋接和虛擬交換機

## Catalyst 8300 uCPE 1N20

1. 預設情況下，可以透過FPGE ( 前面板Gigabit乙太網 ) WAN埠或透過GE0-2 LAN埠訪問NFVIS以進行管理
2. 預設情況下，WAN網路(wan-net)和WAN網橋(wan-br)設定為啟用DHCP。預設情況下，GE0-0與WAN網橋關聯
3. 預設情況下會建立WAN網路(wan2-net)和WAN網橋(wan2-br)，但不會與任何物理埠關聯
4. GE0-2與LAN網橋關聯，所有其他埠不與OVS關聯
5. C8300-uCPE上的管理IP 192.168.1.1可透過GE0-2訪問
6. 建立內部管理網路(int-mgmt-net)和網橋(int-mgmt-br)，並在內部用於系統監控。

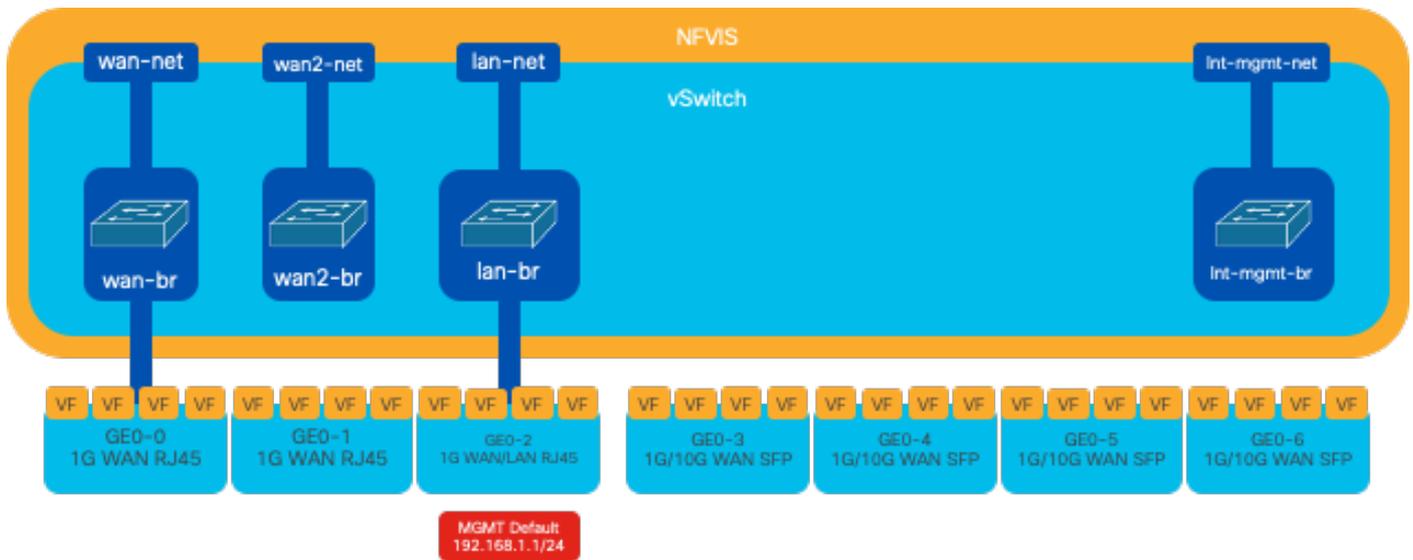


圖3. 分配給8300 NIC的內部橋接和虛擬交換機

## 網路虛擬化技術

### 開放式vSwitch (OVS)

開放式vSwitch (OVS)是一種開源多層虛擬交換機，旨在透過程式設計擴展實現網路自動化，同時支援標準管理介面和協定，例如NetFlow、sFlow、IPFIX、RSPAN、CLI、LACP和802.1ag。它廣泛用於大型虛擬化環境，特別是使用Hypervisor來管理虛擬機器(VM)之間的網路流量。它允許建立透過NFVIS介面直接管理的複雜網路拓撲和策略，為網路功能虛擬化提供多功能環境。

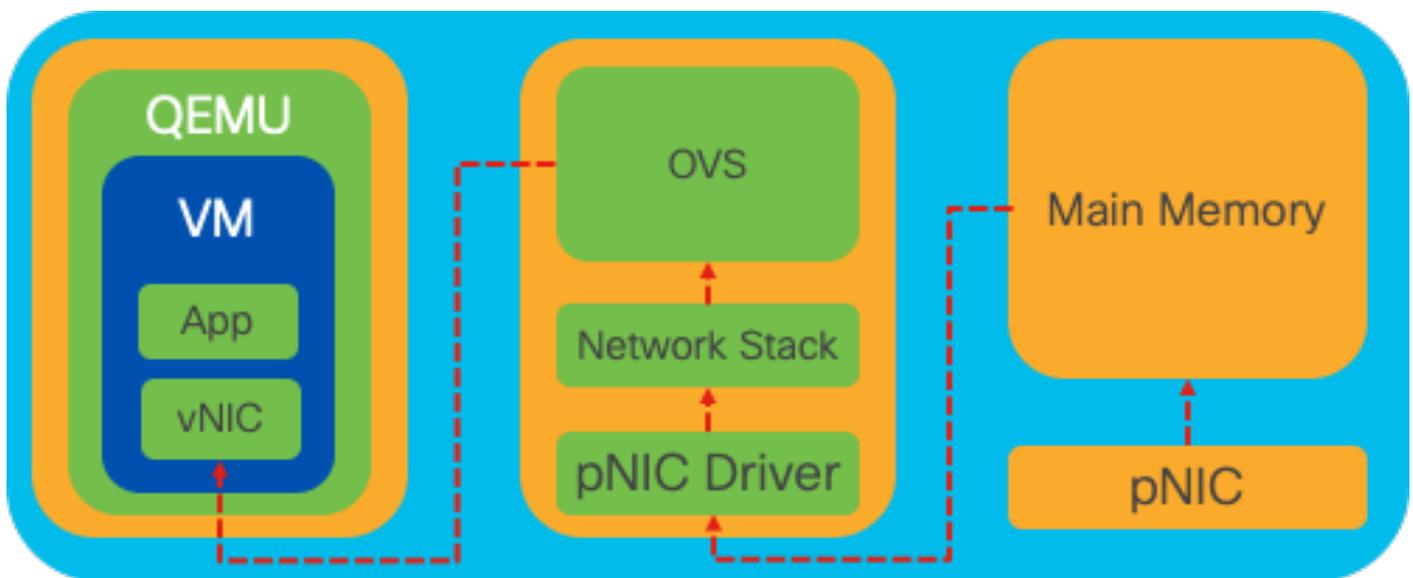


圖4.Linux核心中的OVS配置

### OVS網橋

它使用虛擬網路網橋和流規則在主機之間轉發資料包。其行為與物理交換機類似，只是進行了虛擬化。

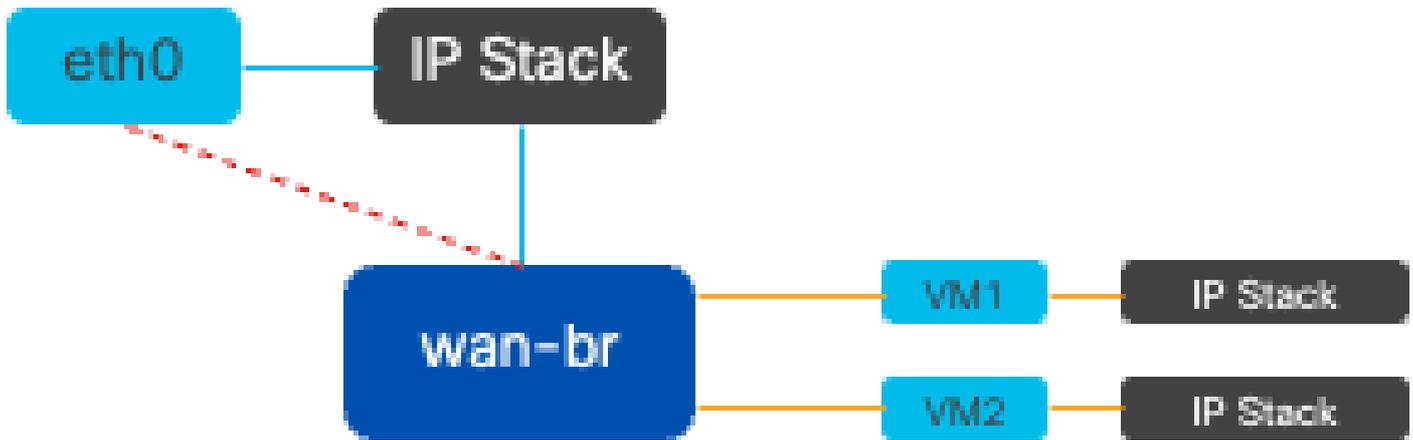


圖5.連線到wan-br網橋的2個VM或VNF的示例實施

### 環境切換缺陷

當網路封包到達網路介面卡(NIC)時，會觸發中斷，即傳送訊號給處理器，表示處理器需要立即處理。CPU暫停其當前任務以處理中斷，該過程稱為中斷處理。在此階段，CPU在作業系統核心的控制下，將資料包從NIC讀入記憶體，並根據資料包的目的地和目的確定後續步驟。目標是快速處理資料包或將資料包路由到其預期應用，從而最大程度地減少延遲並最大程度地提高吞吐量。

上下文切換是CPU從一個環境(上下文)中執行任務切換到另一個環境(上下文)中的過程。當在使用者模式和核心模式之間移動時，這一點尤其重要：

- 使用者模式：這是大多數應用程式執行的受限處理模式。使用者模式中的應用程式無法直接存取硬體或參考記憶體，而且必須與作業系統核心通訊，才能執行這些作業。
- 核心模式：此模式向作業系統授予對硬體和所有記憶體的完全訪問許可權。核心可以執行任何CPU指令並引用任何記憶體地址。執行諸如管理硬體裝置、記憶體和執行系統呼叫等任務時需要核心模式。

當應用程式需要執行需要核心層級許可權的作業(例如讀取網路封包)時，會發生內容切換。CPU從使用者模式轉換到核心模式以執行操作。一旦完成，另一個情景交換機將CPU返回到使用者模式以繼續執行該應用程式。此交換過程對於維護系統穩定性和安全性至關重要，但會引入可能影響效能的開銷。

OVS主要在作業系統使用者空間中運行，隨著資料吞吐量的增加，OVS可能成為瓶頸。這是因為需要更多的上下文切換才能讓CPU移動到核心模式來處理資料包，從而降低效能。這種限制在高資料包速率或準確計時至關重要的環境中尤其明顯。為了解決這些效能限制並滿足現代高速網路的需求，開發了DPDK(資料平面開發套件)和SR-IOV(單根I/O虛擬化)等技術。

### 資料平面開發套件(DPDK)

DPDK是一組程式庫和驅動程式，可加速各種CPU架構上的封包處理工作負載。透過繞過傳統核心網路堆疊(避免上下文交換)，DPDK可以顯著增加資料平面吞吐量和減少延遲。這對於需要低延遲通訊的高吞吐量VNF特別有用，使NFVIS成為效能敏感型網路功能的理想平台。

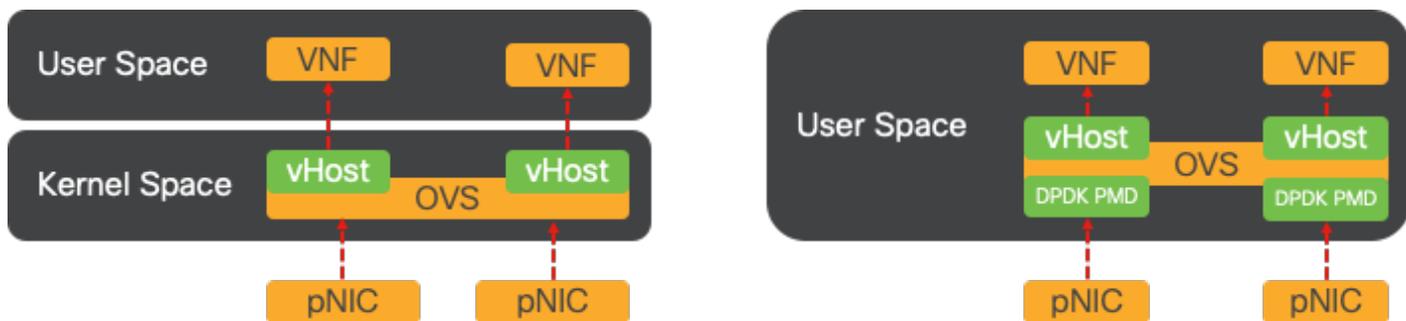


圖6.傳統OVS ( 左側 ) 和DPDK OVS ( 右側 ) 情景交換最佳化

NFVIS 3.10.1 (ENCS)和3.12.2 (OVS)中對DPDK的支援已啟動。

- 接近SRIOV的服務鏈吞吐量，優於非DPDK OVS。
- VNF需要Virtio驅動程式。
- 支援的平台：
- ENCS 3.10.1以上。
- UCSE、UCS-C、CSP5K 3.12.1以上版本。
- 自4.12.1起支援的埠通道的DPDK。
- 資料包/流量捕獲：DPDK不支援。
- PNIC上的SPAN流量：DPDK不支援。
- 啟用OVS-DPDK後，無法將其作為單獨功能停用。停用DPDK的唯一方法就是執行工廠重設。

## 資料複製

傳統網路方法通常要求資料在到達VM記憶體中的目標之前進行多次複製。例如，必須將資料包從NIC複製到核心空間，然後複製到使用者空間，以便由虛擬交換機 ( 如OVS ) 進行處理，最後複製到虛擬機器記憶體。儘管DPDK透過繞過核心網路堆疊提供了效能改進，但每次複製操作都會導致延遲並增加CPU使用率。

這些開銷包括記憶體副本以及在將資料包轉發到虛擬機器之前處理使用者空間中的資料包所需的處理時間。PCIe Passthrough和SR-IOV透過允許物理網路裝置 ( 如NIC ) 直接在多個VM之間共用，而不像傳統虛擬化方法那樣涉及主機作業系統，從而解決這些瓶頸。

## PCIe傳輸

此策略包括繞過虛擬機器監控程式，允許虛擬網路功能(VNF)直接訪問網路介面卡(NIC)，從而幾乎實現最大吞吐量。此方法稱為PCI直通，它允許完整NIC專門用於訪客作業系統，而無需虛擬機器監控程式的干預。在此設定中，虛擬機器就像直接連線到NIC一樣。例如，如果提供兩個NIC卡，則每個網路卡都可以專門分配給不同的VNF，從而提供直接訪問。

但是，此方法有一個缺點：如果只有兩個NIC可供兩個單獨的VNF單獨使用，則任何其他VNF ( 例如第三個 ) 將沒有NIC訪問許可權，因為缺少專屬NIC可供使用。另一種解決方案是使用單一I/O虛擬化(SR-IOV)。

## 單一I/O虛擬化(SR-IOV)

是允許單個物理PCI裝置(如網路介面卡(NIC))顯示為多個獨立虛擬裝置的規範。該技術提供對物理網路裝置的直接VM訪問，從而減少開銷並提高I/O效能。其工作方式是將單個PCIe裝置劃分為多個虛擬片，每個虛擬片可分配給不同的VM或VNF，有效地解決了網路卡數量有限所帶來的限制。這些虛擬磁碟片段(稱為虛擬功能(VF))允許多個VNF之間共用NIC資源。物理功能(PF)是指促進SR-IOV功能的實際物理元件。

透過利用SR-IOV，NFVIS可以將專用NIC資源分配給特定的VNF，透過促進直接記憶體訪問(DMA)網路資料包直接進入各自的VM記憶體來確保高效能和低延遲。此方法可將CPU僅處理封包的情況降至最低，因此可降低CPU使用量。這對於需要保證頻寬或具有嚴格效能要求的應用程式特別有用。

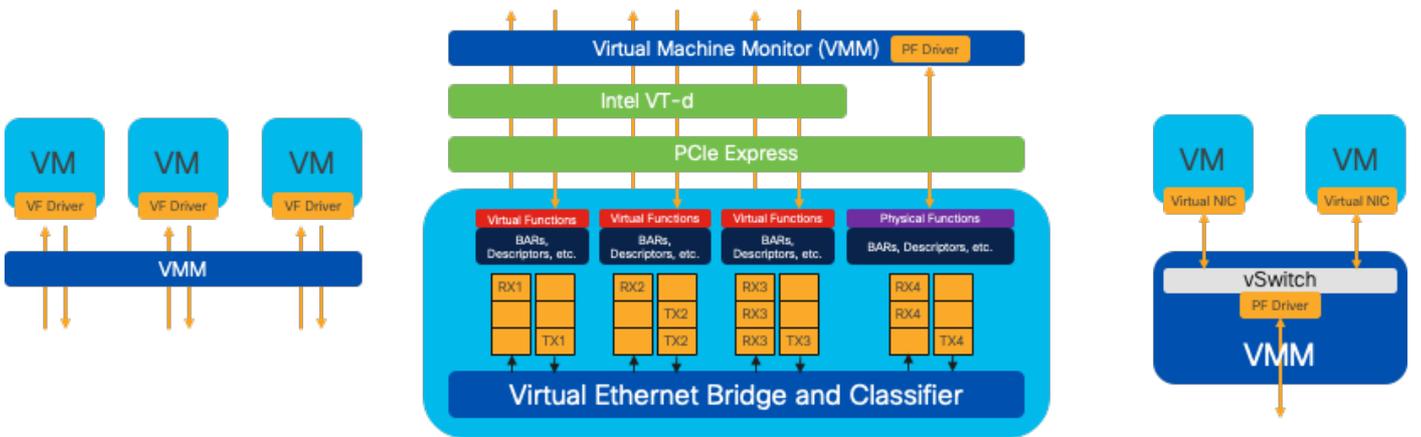


圖7.NFVIS SR-IOV PCIe資源透過硬體功能分離

### 物理功能(PF)

它們是功能齊全的PCIe功能，指的是專門構建的硬體盒，提供特定的網路功能；這些是功能齊全的PCIe功能，可以像任何其他PCIe裝置一樣被發現、管理和操作。物理功能包括可用於配置和控制PCIe裝置的SR-IOV功能。

### 虛擬功能(VF)

它們以最少配置資源（輕量）實現精簡功能，僅側重於將I/O處理為簡單的PCIe功能。每個虛擬功能都源自一個物理功能。裝置硬體限制虛擬功能的可能數量。一個乙太網埠（物理裝置）可以對應於多個虛擬功能，然後這些虛擬功能可以分配給不同的虛擬機器。

### 在支援NFVIS的硬體上建議的SR-IOV加速驅動程式

平台	NIC	NIC驅動程式
ENCS 54XX	背板交換器	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偏好設定

特別是在網路流量主要流向東西方（意味著它位於同一台伺服器內）的情況下，DPDK的效能優於SR-IOV。其基本原理很簡單：當流量在伺服器內進行內部管理而不需要訪問NIC時，SR-IOV不會帶來任何好處。事實上，SR-IOV可能會因不必要地擴展流量路徑並消耗NIC資源而可能導致效率低下。因此，對於內部伺服器流量管理，使用DPDK是更高效率的選擇。

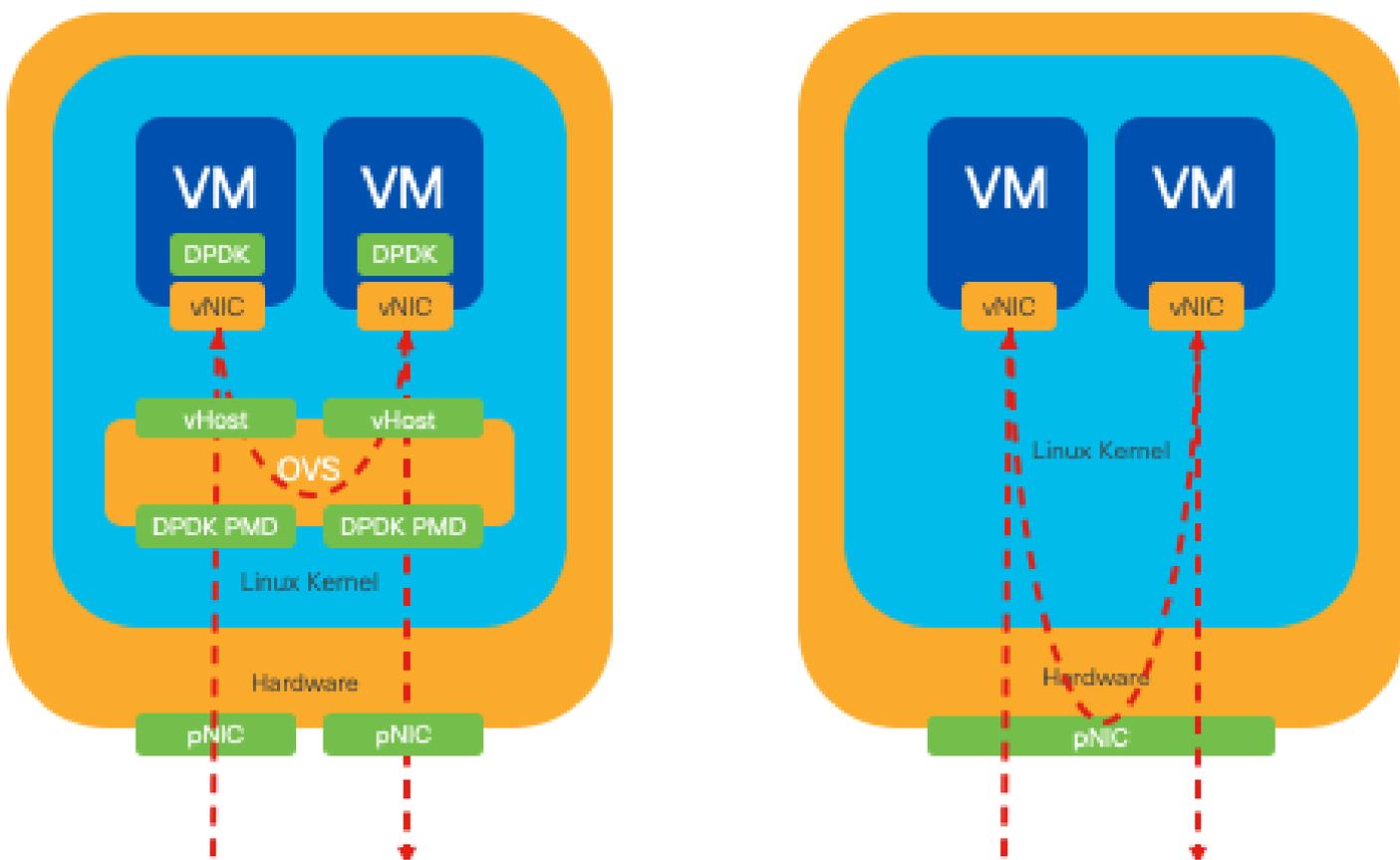


圖8. DPDK和SR-IOV資料包遍歷東到西流量

## SR-IOV偏好設定

在網路流量從北向南，甚至從東向西流動，但特別是在伺服器之間流動的情況下，使用SR-IOV比DPDK更有優勢。對於伺服器到伺服器的通訊尤其如此。由於此類流量不可避免地必須透過NIC，因此選擇使用DPDK增強的OVS可能會不必要地帶來額外的複雜性和潛在的效能限制。因此，在這些情況下，SR-IOV成為首選方案，為處理伺服器間流量提供直接而有效的路徑。

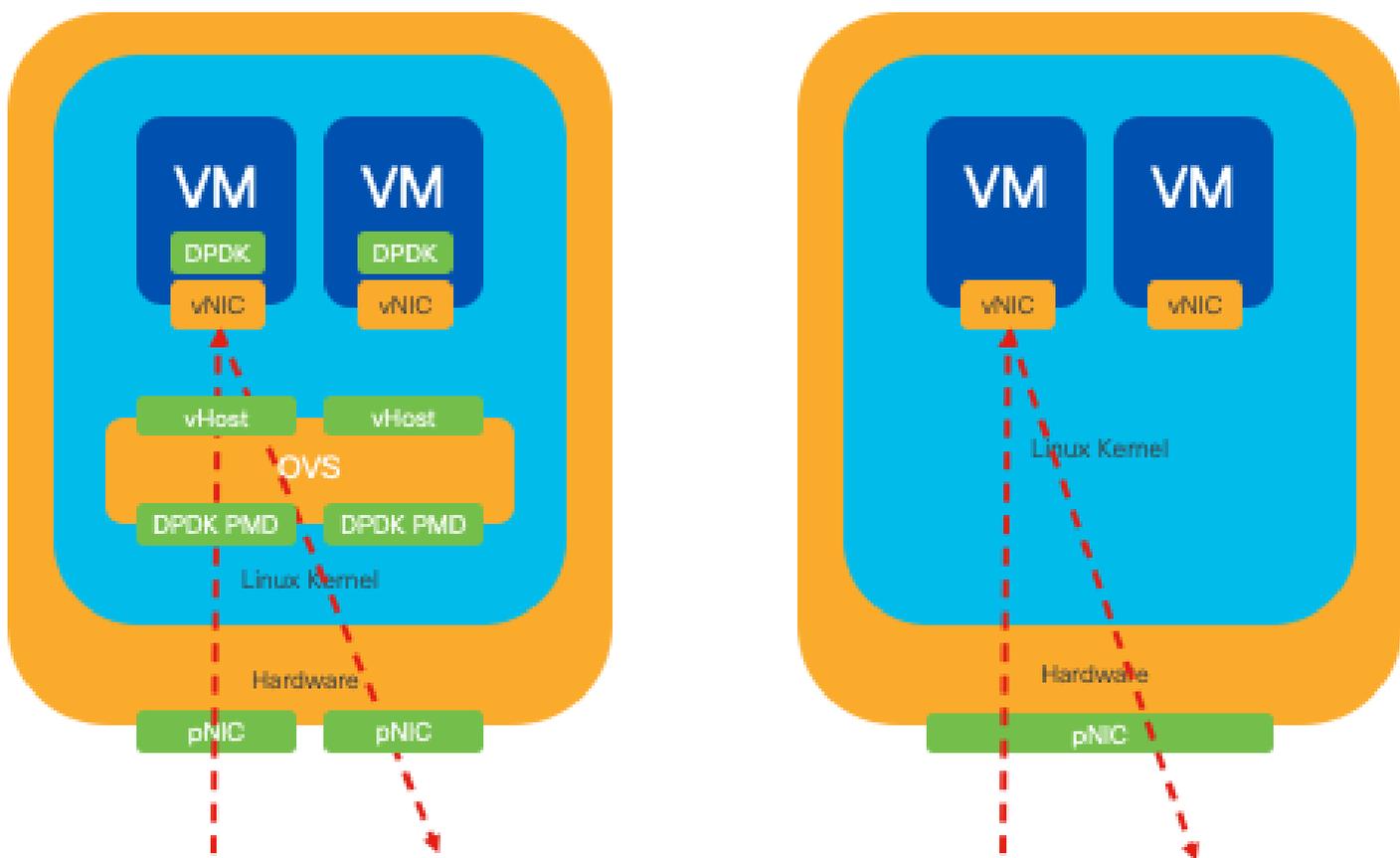
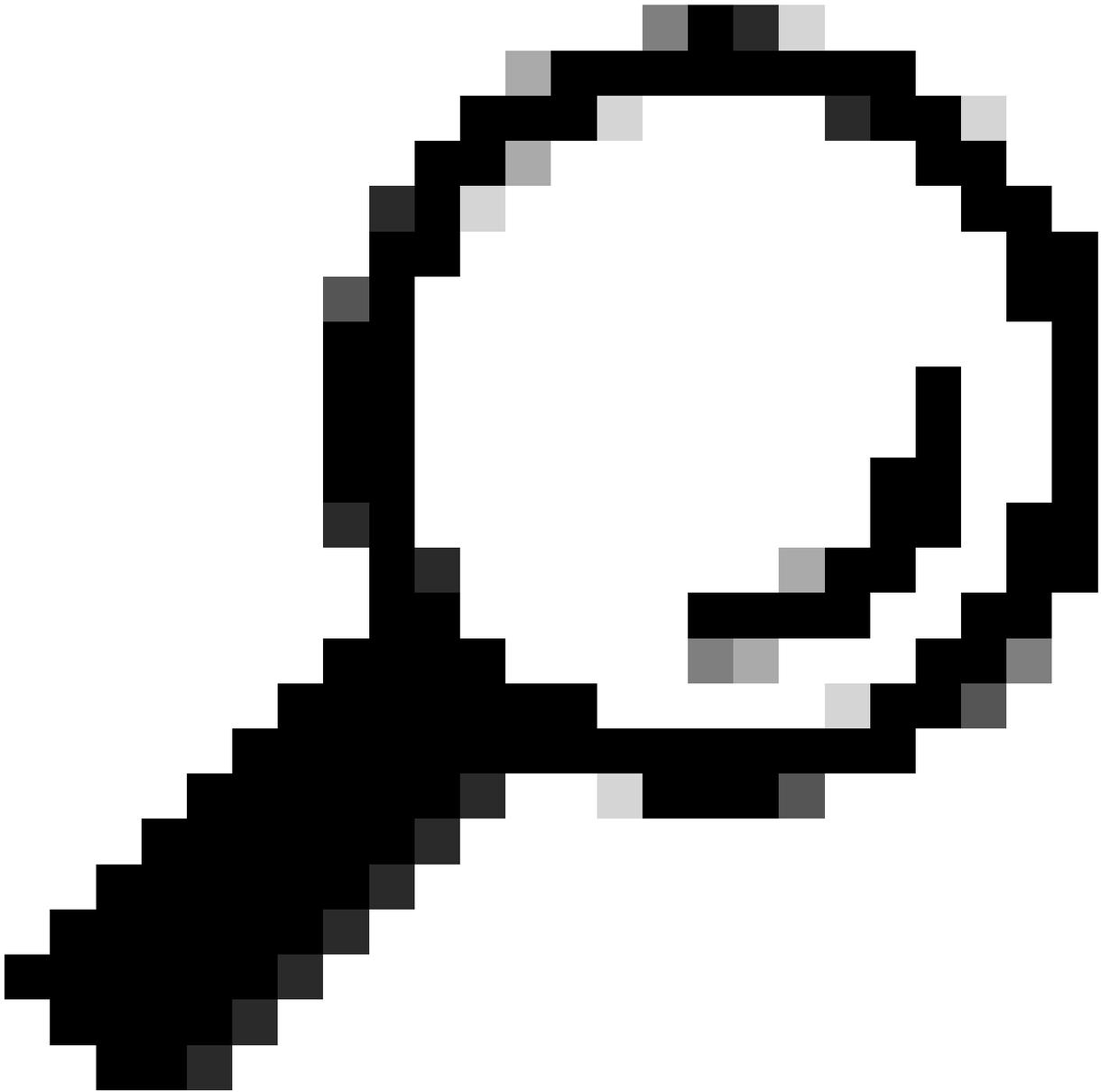


圖9.北向南流量中的DPDK和SR-IOV資料包遍歷



提示：請記住，透過將SR-IOV與DPDK整合到虛擬網路功能(VNF)中，可以增強基於SR-IOV的設定的效能，但不包括DPDK與OVS結合使用的方案（如前所述）。

---

## 組態

### 啟用DPDK

要從GUI啟用DPDK，您必須導航到Configuration > Virtual Machine > Networking > Networks。進入選單後，按一下開關以啟動特徵

# Networks

## Networks Information and Configuration

---

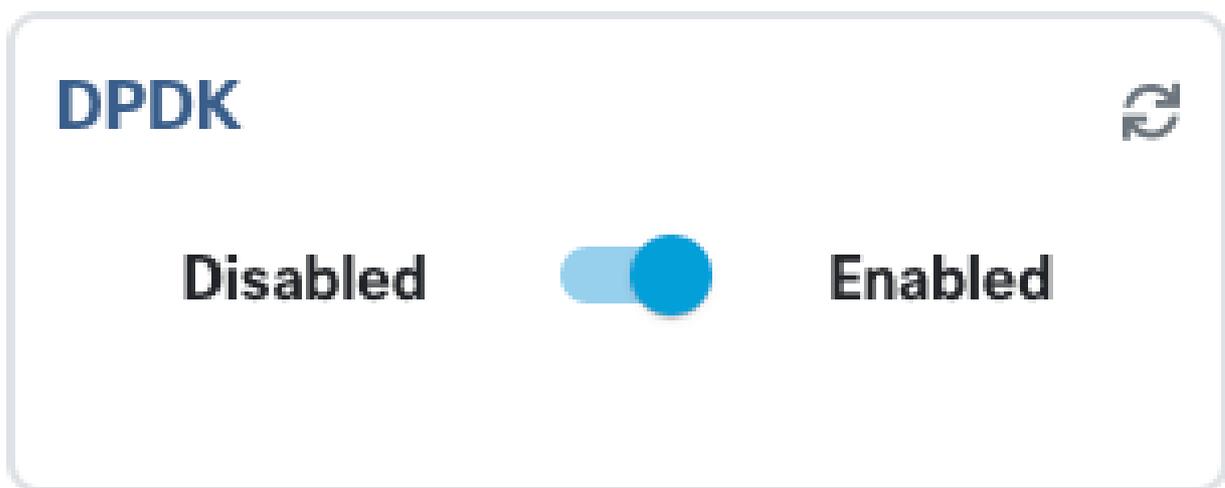
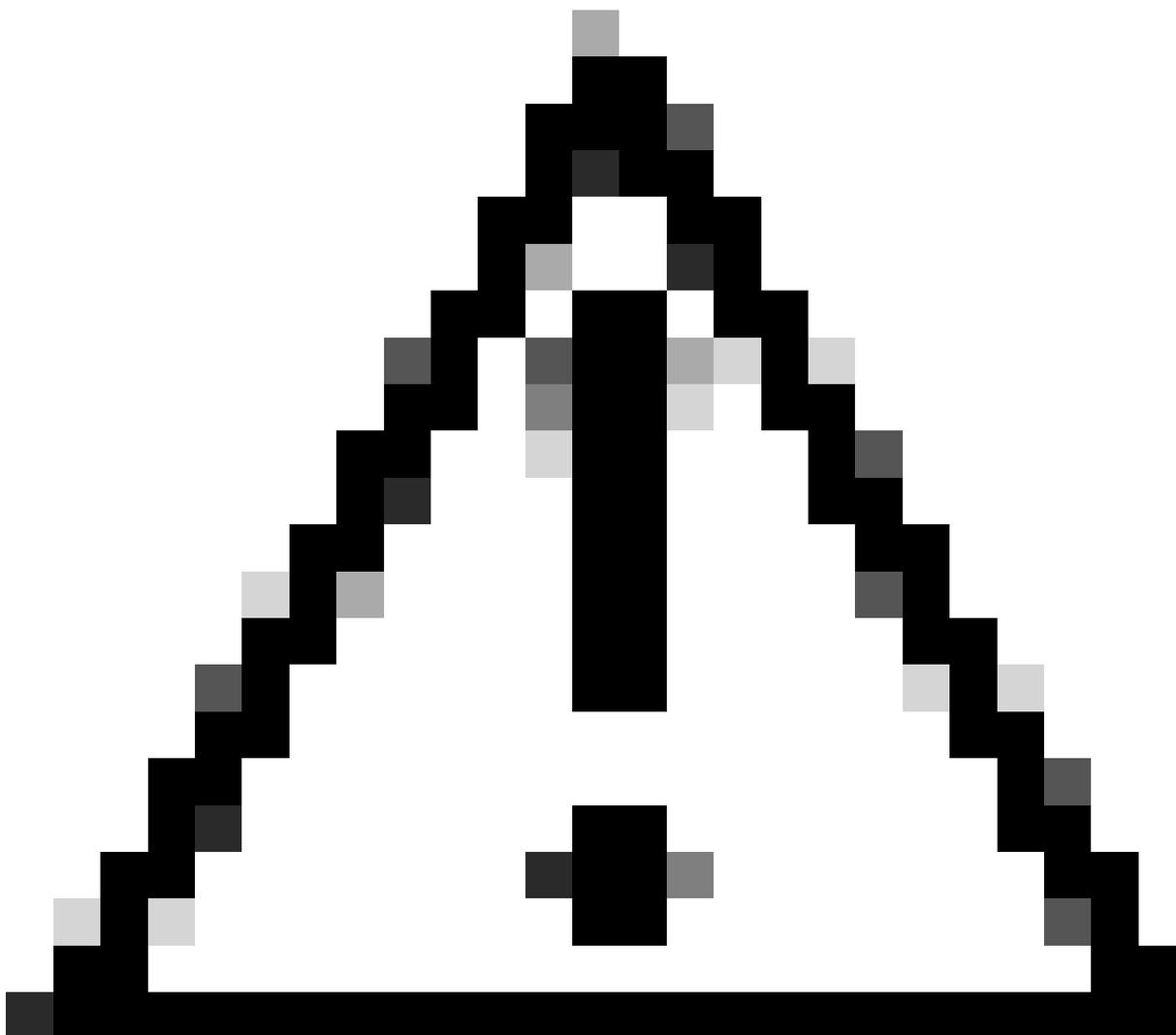


圖10.GUI上提供的幻燈片按鈕，用於DPDK啟用

對於CLI，您必須在配置模式下從全局系統設定啟用它。

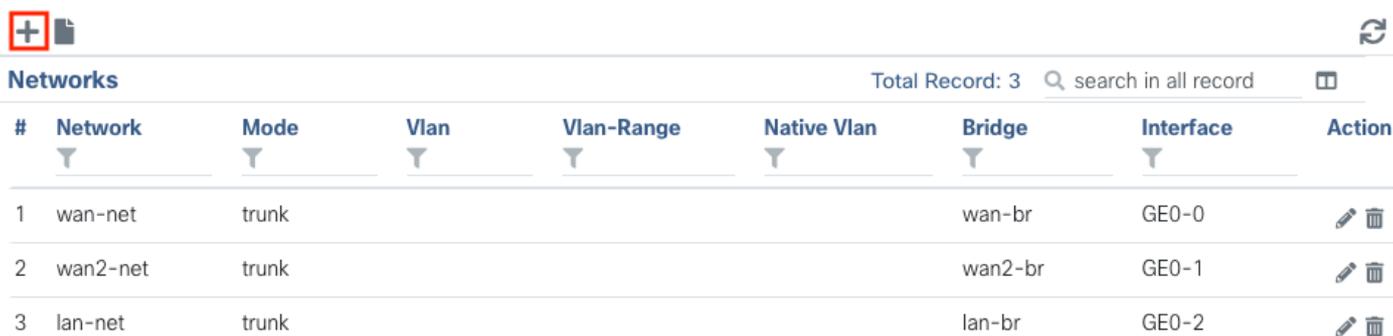
```
nfvis(config)# system settings dpdk enable
```



注意：除非從NFVIS執行出廠重置，否則無法停用DPDK。

## 建立新網路並將其與新的OVS網橋關聯

導航到配置 > 虛擬機器 > 網路 > 網路。進入「網路」頁面後，按一下「網路」表格左上方的加號 (+)。



#	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	 

圖11.NFVIS GUI中的網路表檢視

命名網路並關聯到新網橋。VLAN和介面繫結選項可以根據網路基礎架構的需求而定。

# 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

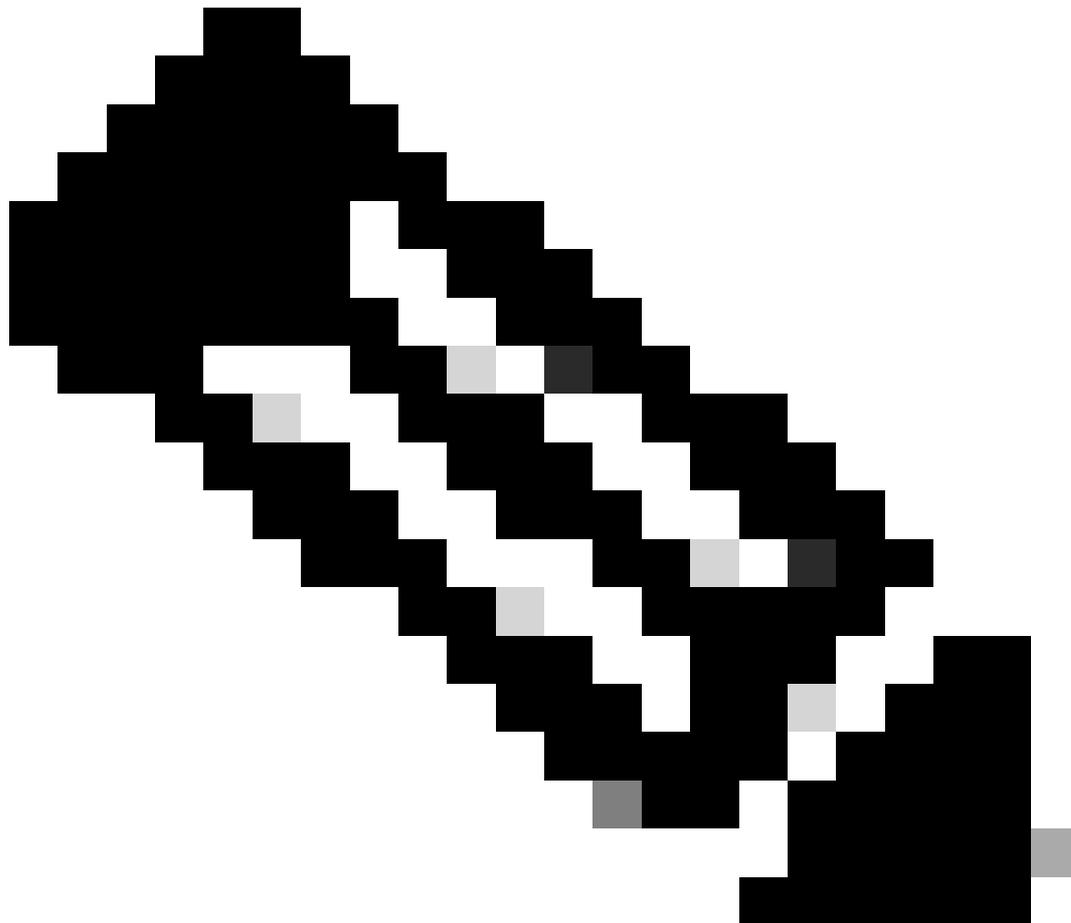
圖12.在NFVIS GUI中建立虛擬網路的「增加網路」模式

按一下submit按鈕後，必須能夠檢視增加到Networks表的新建立的網路。



#	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		 

圖 13.NFVIS GUI中的「Refresh Icon」（刷新圖示）位於右上角的網路表檢視（以紅色突出顯示）



注意：如果在表格中未觀察到新網路，請按一下右上方的重新整理按鈕，或重新整理整個頁面。

如果從CLI執行，每個網路和網橋都從配置模式建立，則工作流程與GUI版本相同。

### 1. 建立新網橋。

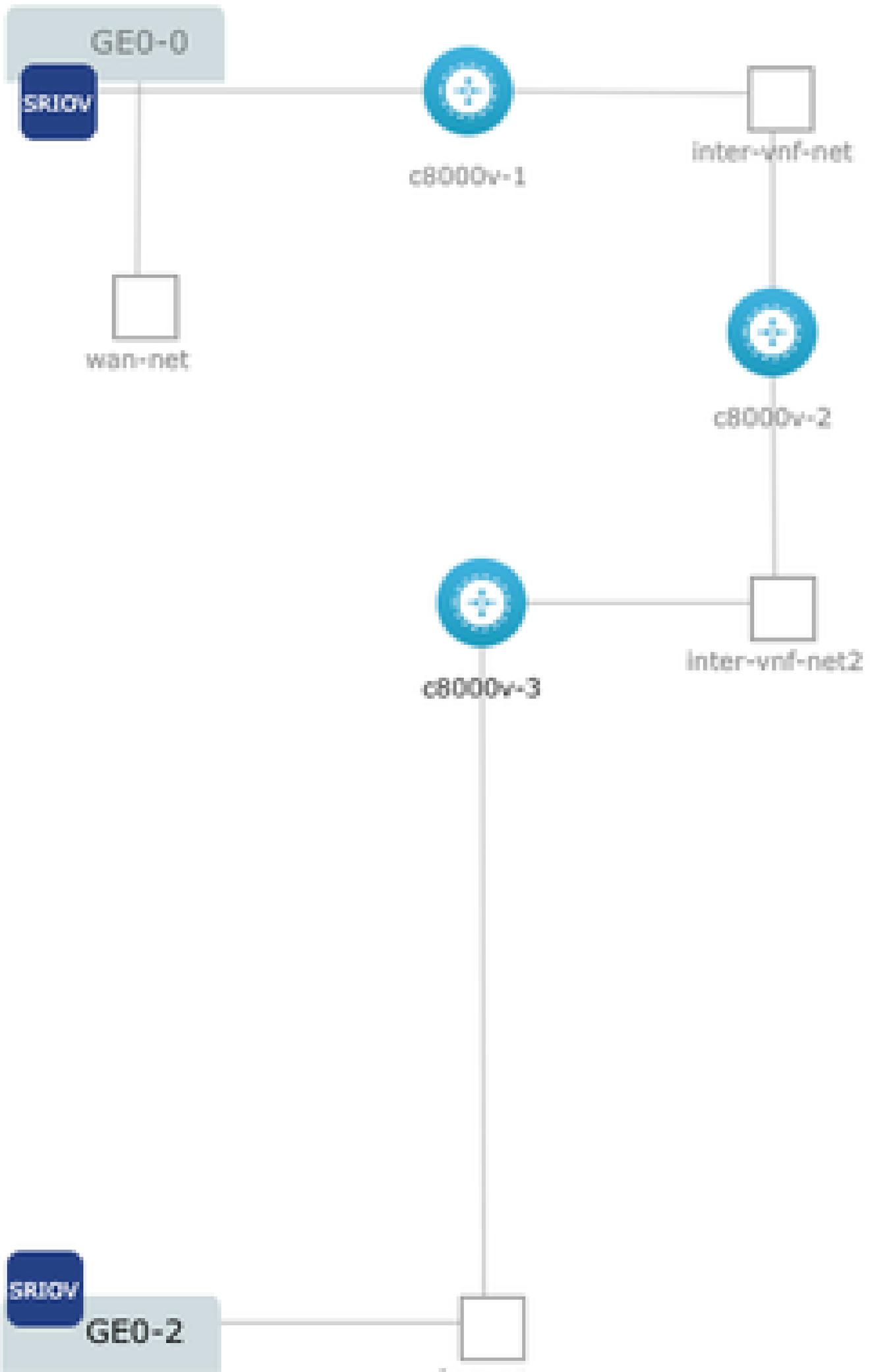
```
nfvis(config)# bridges bridge inter-vnf-br2
nfvis(config-bridge-inter-vnf-br2)# commit
```

### 2. 建立新網路並將其與之前建立的網橋關聯

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

## 連線VNF

要從網路拓撲或單個VFN部署開始，必須導航到Configuration > Deploy。您可以將虛擬機器或容器從選擇清單拖到拓撲構建區域中，以開始建立虛擬化基礎設施。



## 關於此翻譯

思科已使用電腦和人工技術翻譯本文件，讓全世界的使用者能夠以自己的語言理解支援內容。請注意，即使是最佳機器翻譯，也不如專業譯者翻譯的內容準確。Cisco Systems, Inc. 對這些翻譯的準確度概不負責，並建議一律查看原始英文文件（提供連結）。