

Comprendre les réseaux virtuels NFVIS : OVS, DPDK et SR-IOV

Table des matières

[Introduction](#)

[Composants utilisés](#)

[Présentation des réseaux dans NFVIS](#)

[Plate-forme ENCS54XX](#)

[Catalyst 8200 uCPE](#)

[Catalyst 8300 uCPE 1N20](#)

[Technologies de virtualisation du réseau](#)

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

[Ponts OVS](#)

[Déficits de la commutation de contexte](#)

[Kit de développement du plan de données \(DPDK\)](#)

[Copie de données](#)

[Passthrough PCIe](#)

[Virtualisation d'E/S racine unique \(SR-IOV\)](#)

[Fonctions physiques \(PF\)](#)

[Fonctions virtuelles \(VF\)](#)

[Pilotes recommandés pour l'accélération SR-IOV sur le matériel compatible NFVIS](#)

[Exemples d'utilisation pour DPDK et SR-IOV](#)

[Préférence DPDK](#)

[Préférence SR-IOV](#)

[Configuration](#)

[Activation de DPDK](#)

[Créer un nouveau réseau et l'associer à un nouveau pont OVS](#)

[Connexion de VNF](#)

[Articles et documentation associés](#)

Introduction

Ce document décrit le schéma de mise en réseau virtuel que la plate-forme NFVIS fournit pour la communication de VNF dans les réseaux d'entreprise et de service.

Composants utilisés

Les informations contenues dans ce document sont basées sur les composants matériels et logiciels suivants :

- ENCS5412 exécutant NFVIS 4.7.1-FC4

- c8300 uCPE 1N20 exécutant NFVIS 4.12.1-FC2

The information in this document was created from the devices in a specific lab environment. All of the devices used in this document started with a cleared (default) configuration. Si votre réseau est en ligne, assurez-vous de bien comprendre l'incidence possible des commandes.

Présentation des réseaux dans NFVIS

Un réseau de gestion interne (int-mgmt-net) et un pont (int-mgmt-br) sont utilisés en interne pour la surveillance VNF, en attribuant des adresses IP de gestion à partir du sous-réseau 10.20.0.0/24.

Plate-forme ENCS54XX

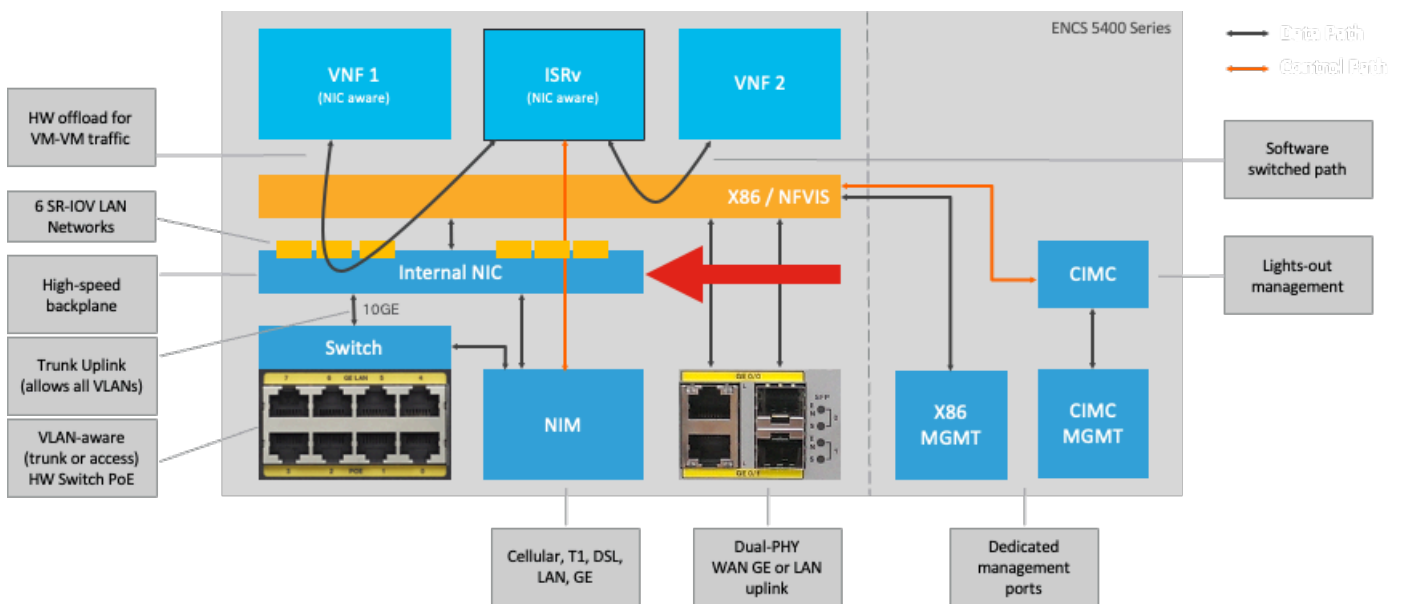


Figure 1. Commutateur matériel et connexions internes des cartes réseau de liaison ascendante WAN/LAN

Catalyst 8200 uCPE

- NFVIS est accessible par défaut via le port WAN ou le port LAN GE0/2 pour la gestion.
- Le réseau WAN (wan-net et wan2-net) et le pont WAN (wan-br et wan2-br) sont configurés pour activer DHCP par défaut. GE0-0 est associé par défaut au pont WAN et GE0-1 au pont WAN2.
- L'adresse IP de gestion 192.168.1.1 sur l'UCPE Catalyst 8200 est accessible via GE0-2.
- GE0-2 est associé au pont LAN.
- Un réseau de gestion interne (int-mgmt-net) et un pont (int-mgmt-br) sont créés et utilisés en interne pour la surveillance du système.

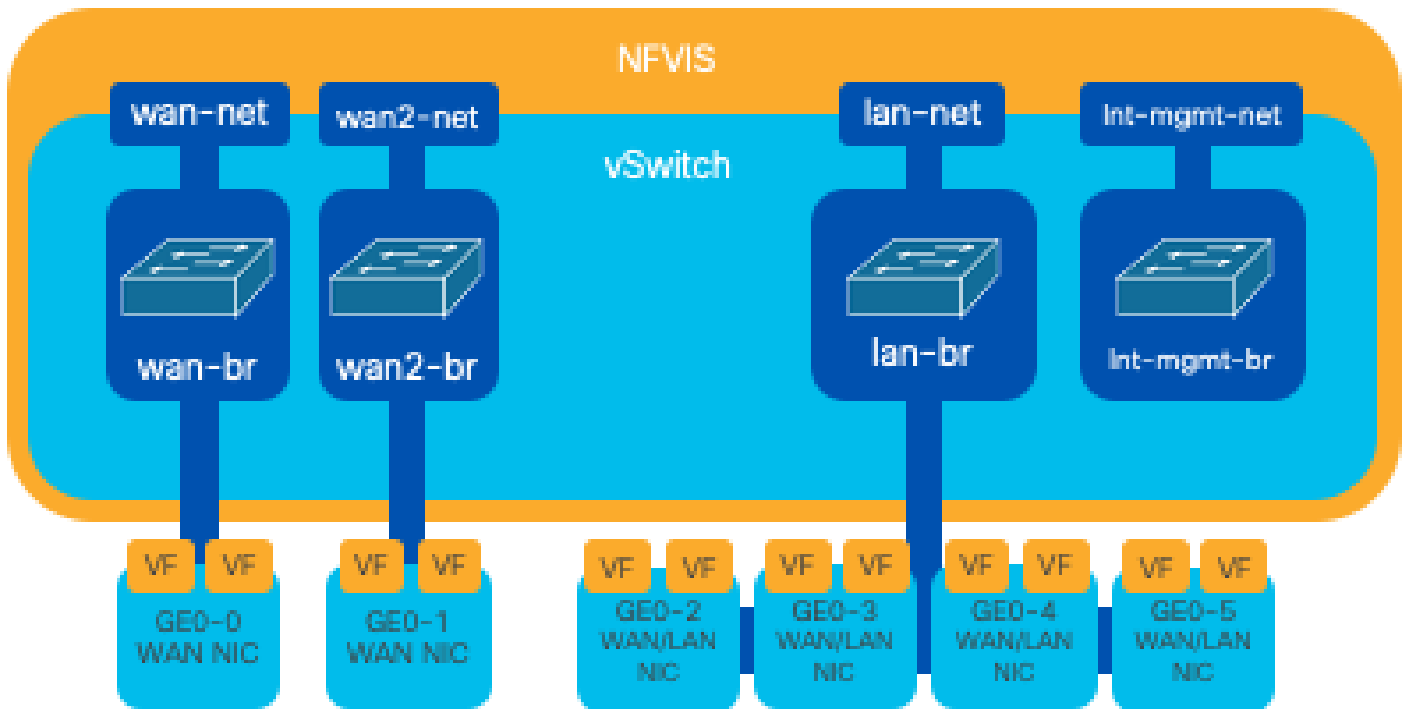


Figure 2. Pontage interne et commutateurs virtuels attribués aux cartes réseau 8200

Catalyst 8300 uCPE 1N20

1. NFVIS est accessible par défaut via les ports WAN FPGE (Front Panel Gigabit Ethernet) ou via le port LAN GE0-2 pour la gestion
2. Le réseau WAN (wan-net) et un pont WAN (wan-br) sont définis par défaut pour activer DHCP. GE0-0 est par défaut associé au pont WAN
3. Le réseau WAN (wan2-net) et un pont WAN (wan2-br) sont créés par défaut, mais ne sont associés à aucun port physique
4. GE0-2 est associé au pont LAN, tous les autres ports ne sont pas associés à OVS
5. L'adresse IP de gestion 192.168.1.1 sur C8300-uCPE est accessible via GE0-2
6. Un réseau de gestion interne (int-mgmt-net) et un pont (int-mgmt-br) sont créés et utilisés en interne pour la surveillance du système.

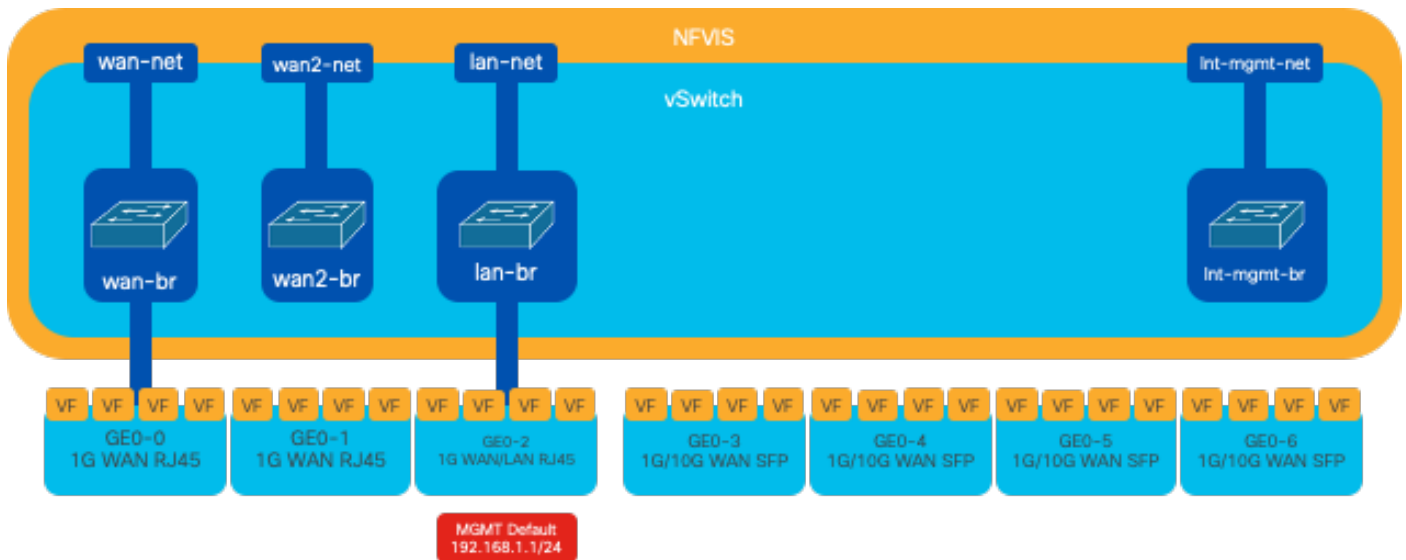


Figure 3. Pontage interne et commutateurs virtuels attribués aux cartes réseau 8300

Technologies de virtualisation du réseau

Open vSwitch (OVS)

Open vSwitch (OVS) est un commutateur virtuel multicouche open source conçu pour permettre l'automatisation du réseau par le biais d'extensions programmatiques, tout en prenant en charge les interfaces et protocoles de gestion standard, tels que NetFlow, sFlow, IPFIX, RSPAN, CLI, LACP et 802.1ag. Il est largement utilisé dans les grands environnements virtualisés, en particulier avec les hyperviseurs pour gérer le trafic réseau entre les machines virtuelles (VM). Il permet la création de topologies et de politiques réseau sophistiquées directement gérées via l'interface NFVIS, offrant ainsi un environnement polyvalent pour la virtualisation des fonctions réseau.

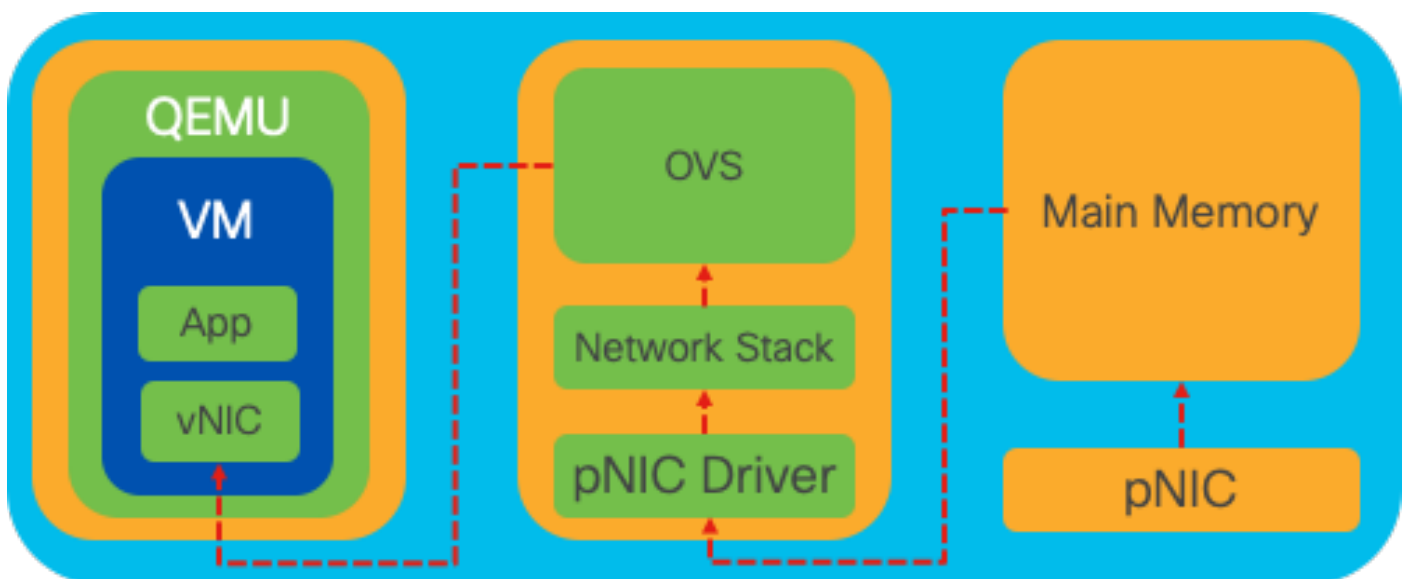


Figure 4. Configuration OVS dans le noyau Linux

Ponts OVS

Il utilise des ponts réseau virtuels et des règles de flux pour transférer des paquets entre les hôtes. Il se comporte comme un commutateur physique, mais virtualisé.

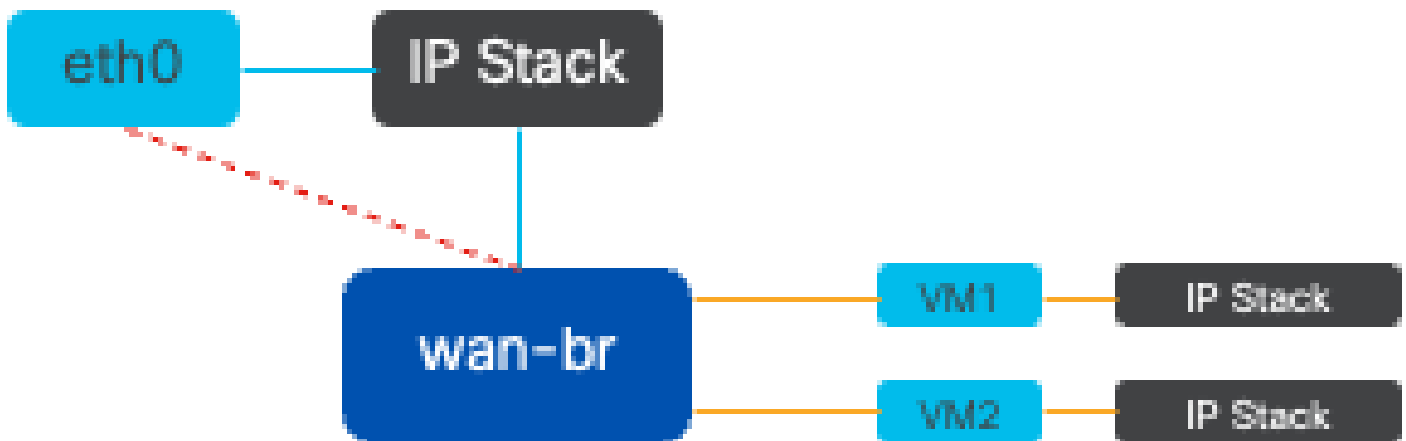


Figure 5. Exemple de mise en oeuvre de 2 machines virtuelles ou VNF reliées au pont wan-br

Déficits de la commutation de contexte

Lorsqu'un paquet réseau arrive sur une carte réseau, il déclenche une interruption, un signal envoyé au processeur indiquant qu'il a besoin d'une attention immédiate. Le processeur interrompt ses tâches en cours pour gérer l'interruption, un processus appelé traitement des interruptions. Au cours de cette phase, le processeur, sous le contrôle du noyau du système d'exploitation, lit le paquet de la carte réseau dans la mémoire et décide des étapes suivantes en fonction de la destination et du but du paquet. L'objectif est de traiter ou d'acheminer rapidement le paquet vers l'application voulue, en minimisant la latence et en optimisant le débit.

La commutation de contexte est le processus par lequel un processeur passe de l'exécution de tâches dans un environnement (contexte) à un autre. Ceci est particulièrement pertinent lorsque vous passez du mode utilisateur au mode noyau :

- Mode utilisateur : il s'agit d'un mode de traitement restreint dans lequel la plupart des applications s'exécutent. Les applications en mode utilisateur n'ont pas d'accès direct au matériel ou à la mémoire de référence et doivent communiquer avec le noyau du système d'exploitation pour effectuer ces opérations.
- Kernel Mode : ce mode accorde au système d'exploitation un accès complet au matériel et à toute la mémoire. Le noyau peut exécuter n'importe quelle instruction CPU et référencer n'importe quelle adresse mémoire. Le mode noyau est requis pour effectuer des tâches telles que la gestion des périphériques matériels, de la mémoire et l'exécution des appels système.

Lorsqu'une application doit effectuer une opération nécessitant des privilèges de niveau noyau (comme la lecture d'un paquet réseau), un changement de contexte se produit. Le processeur passe du mode utilisateur au mode noyau pour exécuter l'opération. Une fois terminé, un autre commutateur de contexte remet le processeur en mode utilisateur pour continuer à exécuter l'application. Ce processus de commutation est essentiel au maintien de la stabilité et de la sécurité du système, mais il introduit une surcharge pouvant affecter les performances.

OVS s'exécute principalement dans l'espace utilisateur du système d'exploitation, ce qui peut devenir un goulot d'étranglement à mesure que le débit de données augmente. En effet, davantage de commutateurs de contexte sont nécessaires pour que le processeur passe en mode noyau pour traiter les paquets, ce qui ralentit les performances. Cette limitation est particulièrement visible dans les environnements à haut débit de paquets ou lorsque la synchronisation précise est cruciale. Pour répondre à ces limitations de performances et aux exigences des réseaux modernes à haut débit, des technologies telles que DPDK (Data Plane Development Kit) et SR-IOV (Single Root I/O Virtualization) ont été développées.

Kit de développement du plan de données (DPDK)

DPDK est un ensemble de bibliothèques et de pilotes conçus pour accélérer les charges de travail de traitement des paquets sur une large gamme d'architectures de CPU. En contournant la pile réseau traditionnelle du noyau (évitant la commutation de contexte), DPDK peut augmenter considérablement le débit du plan de données et réduire la latence. Cela est particulièrement avantageux pour les VNF à haut débit qui nécessitent une communication à faible latence, ce qui fait de NFVIS une plate-forme idéale pour les fonctions réseau sensibles aux performances.

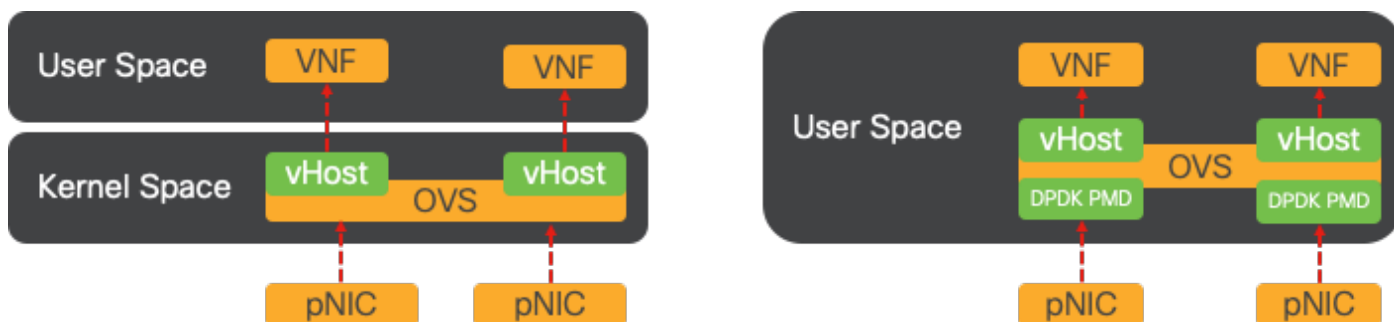


Figure 6. Optimisations de commutation de contexte OVS (côté gauche) et DPDK OVS (côté droit) traditionnelles

La prise en charge de DPDK pour OVS a débuté dans NFVIS 3.10.1 pour ENCS et 3.12.2 pour d'autres plates-formes.

- Débit de la chaîne de services proche de SRIOV, meilleur que celui des OVS non-DPDK.
- Pilote de virtio requis pour VNF.
- Plates-formes prises en charge:
- ENCS 3.10.1 et versions ultérieures.
- UCSE, UCS-C, CSP5K 3.12.1 et versions ultérieures.
- DPDK pour port-channels pris en charge depuis 4.12.1.
- Capture paquet /trafic : non prise en charge dans DPDK.
- Trafic étendu sur pNIC : non pris en charge dans DPDK.
- Une fois qu'OVS-DPDK est activé, il ne peut pas être désactivé en tant que fonctionnalité individuelle. La seule façon de désactiver le DPDK serait une réinitialisation en usine.

Copie de données

Les approches de mise en réseau traditionnelles exigent souvent que les données soient copiées

plusieurs fois avant d'atteindre leur destination dans la mémoire de la machine virtuelle. Par exemple, un paquet doit être copié de la carte réseau vers l'espace noyau, puis vers l'espace utilisateur pour être traité par un commutateur virtuel (comme OVS), et enfin vers la mémoire de la machine virtuelle. Chaque opération de copie entraîne un retard et augmente l'utilisation du CPU malgré les améliorations de performances que le DPDK offre en contournant la pile réseau des noyaux.

Ces frais généraux incluent les copies de mémoire et le temps de traitement nécessaire pour traiter les paquets dans l'espace utilisateur avant qu'ils ne puissent être transférés à la machine virtuelle. PCIe Passthrough et SR-IOV permettent de résoudre ces goulots d'étranglement en permettant à un périphérique réseau physique (tel qu'une carte réseau) d'être partagé directement entre plusieurs machines virtuelles sans impliquer le système d'exploitation hôte dans la même mesure que les méthodes de virtualisation traditionnelles.

Passthrough PCIe

La stratégie consiste à contourner l'hyperviseur pour permettre aux fonctions de réseau virtuel (VNF) d'accéder directement à une carte réseau (NIC), ce qui permet d'atteindre un débit presque maximal. Cette approche est connue sous le nom de PCI passthrough, qui permet à une carte réseau complète d'être dédiée à un système d'exploitation invité sans l'intervention d'un hyperviseur. Dans cette configuration, la machine virtuelle fonctionne comme si elle était directement connectée à la carte réseau. Par exemple, avec deux cartes réseau disponibles, chacune peut être attribuée exclusivement à différents VNF, ce qui leur donne un accès direct.

Cependant, cette méthode présente un inconvénient : si seulement deux cartes réseau sont disponibles et utilisées exclusivement par deux VNF distincts, tout VNF supplémentaire, tel qu'un troisième, serait laissé sans accès à la carte réseau en raison de l'absence d'une carte réseau dédiée disponible pour lui. Une autre solution consiste à utiliser la virtualisation E/S à racine unique (SR-IOV).

Virtualisation d'E/S racine unique (SR-IOV)

Spécification qui permet à un périphérique PCI physique unique, tel qu'une carte réseau, d'apparaître comme plusieurs périphériques virtuels distincts. Cette technologie fournit un accès direct aux machines virtuelles aux périphériques réseau physiques, réduisant ainsi les frais généraux et améliorant les performances d'E/S. Il fonctionne en divisant un seul périphérique PCIe en plusieurs tranches virtuelles, chacune pouvant être attribuée à différentes VM ou VNF, ce qui résout efficacement la limitation causée par un nombre fini de cartes réseau. Ces tranches virtuelles, appelées fonctions virtuelles (VF), permettent le partage des ressources de la carte réseau entre plusieurs VF. La fonction physique (PF) fait référence au composant physique réel qui facilite les capacités SR-IOV.

Grâce à SR-IOV, NFVIS peut allouer des ressources de carte réseau dédiées à des VNF spécifiques, garantissant ainsi des performances élevées et une faible latence en facilitant l'accès direct à la mémoire (DMA) des paquets réseau directement dans la mémoire de la machine virtuelle concernée. Cette approche minimise l'implication du CPU pour traiter simplement les paquets, réduisant ainsi l'utilisation du CPU. Cela est particulièrement utile pour les applications

qui nécessitent une bande passante garantie ou des exigences de performances strictes.

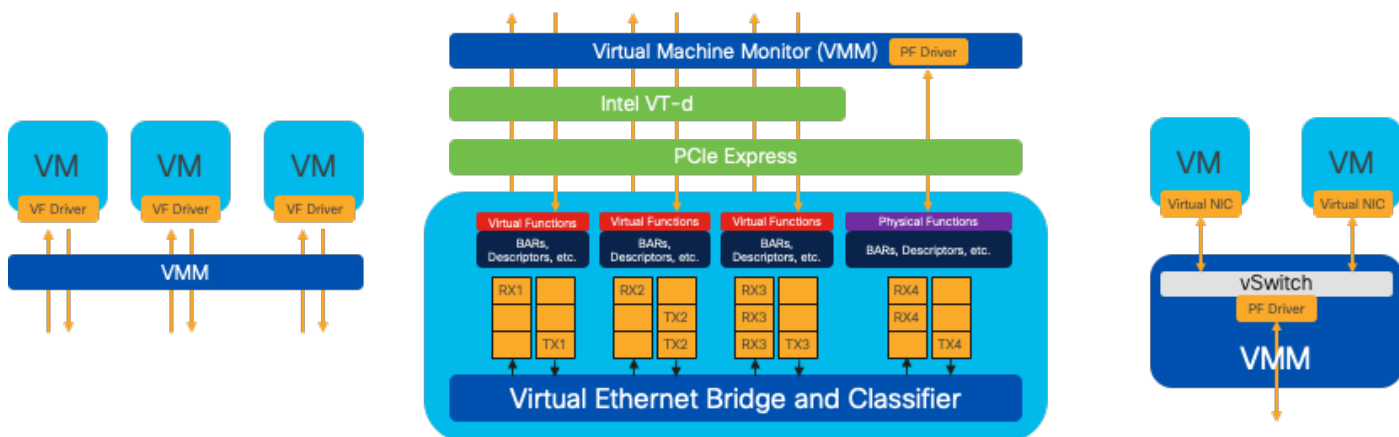


Figure 7. Séparation des ressources NFVIS SR-IOV PCIe par des fonctions matérielles

Fonctions physiques (PF)

Il s'agit de fonctions PCIe complètes, faisant référence à un boîtier matériel spécifique qui fournit des fonctions réseau spécifiques. Il s'agit de fonctions PCIe complètes qui peuvent être découvertes, gérées et manipulées comme n'importe quel autre périphérique PCIe. Les fonctions physiques incluent les fonctionnalités SR-IOV qui peuvent être utilisées pour configurer et contrôler un périphérique PCIe.

Fonctions virtuelles (VF)

Il s'agit de fonctions rationalisées avec des ressources de configuration minimales (légères), axées uniquement sur le traitement des E/S en tant que fonctions PCIe simples. Chaque fonction virtuelle provient d'une fonction physique. Le matériel du périphérique limite le nombre possible de fonctions virtuelles. Un port Ethernet, le périphérique physique, peut correspondre à de nombreuses fonctions virtuelles, qui peuvent ensuite être allouées à différentes machines virtuelles.

Pilotes recommandés pour l'accélération SR-IOV sur le matériel compatible NFVIS

Plateforme	Carte(s) réseau	Pilote de carte réseau
ENCS 54XX	Commutateur de fond de panier	i40e
ENCS 54XX	GE0-0 et GE0-1	pointe
Catalyst 8200 uCPE	GE0-0 et GE0-1	ixgbe
Catalyst 8200 uCPE	GE0-2 et GE0-5	pointe

Exemples d'utilisation pour DPDK et SR-IOV

Préférence DPDK

En particulier dans les scénarios où le trafic réseau circule principalement d'est en ouest (c'est-à-

dire qu'il reste au sein du même serveur), le DPDK surpasse le SR-IOV. Le raisonnement est simple : lorsque le trafic est géré en interne au sein du serveur sans avoir besoin d'accéder à la carte réseau, SR-IOV n'offre aucun avantage. En fait, la SR-IOV peut potentiellement entraîner des inefficacités en étendant inutilement le chemin de trafic et en consommant des ressources de carte réseau. Par conséquent, pour la gestion du trafic du serveur interne, l'utilisation de DPDK est le choix le plus efficace.

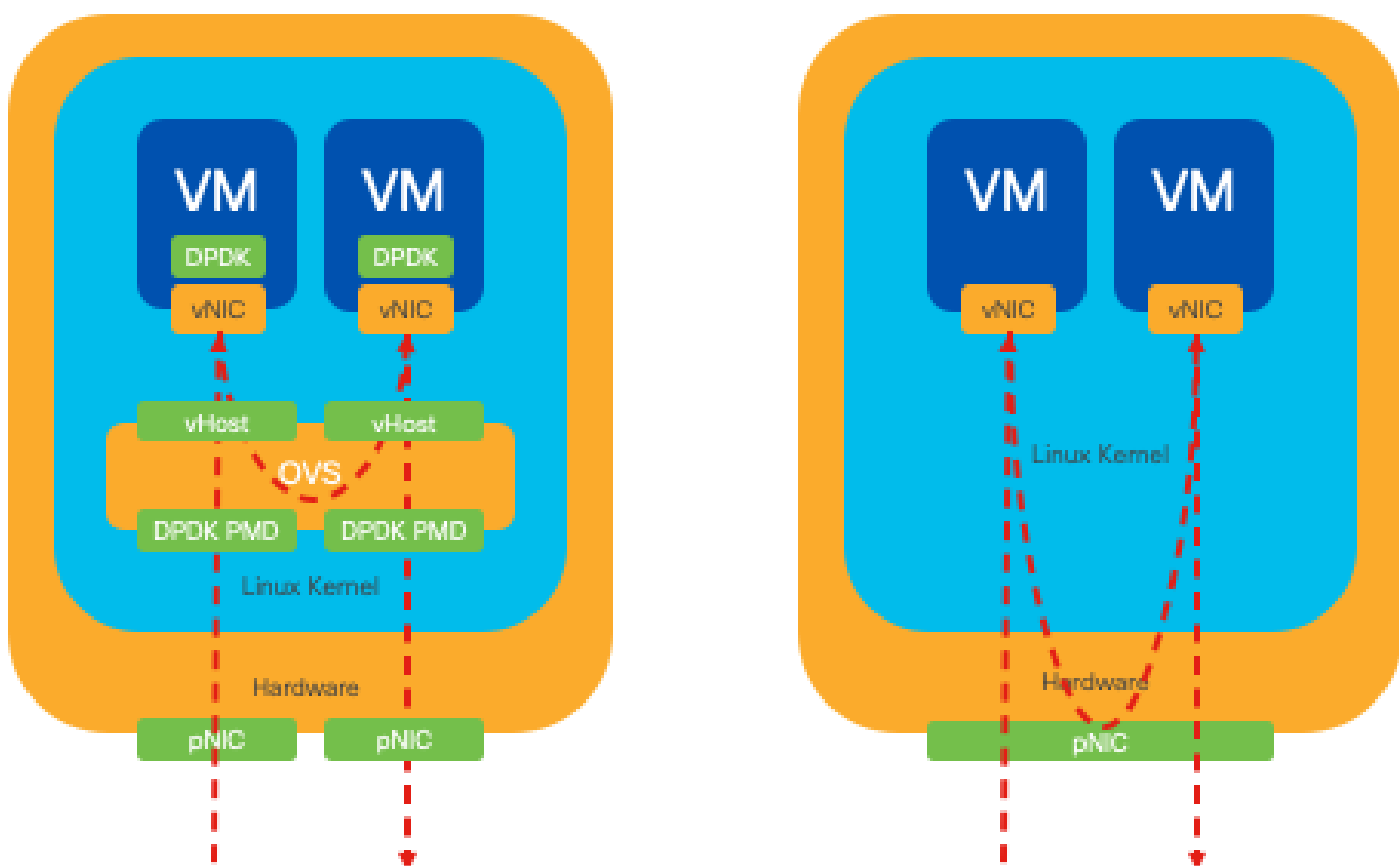


Figure 8. Traversée de paquets DPDK et SR-IOV dans le trafic Est-Ouest

Préférence SR-IOV

Dans les situations où le trafic réseau circule du nord au sud, voire d'est en ouest, mais plus particulièrement entre les serveurs, l'utilisation de SR-IOV s'avère plus avantageuse que le DPDK. Cela est particulièrement vrai pour les communications de serveur à serveur. Étant donné qu'un tel trafic doit inévitablement traverser la carte réseau, le choix d'un système OVS optimisé par DPDK pourrait indubitablement entraîner une complexité supplémentaire et des contraintes de performances potentielles. Par conséquent, SR-IOV apparaît comme le choix préférable dans ces circonstances, offrant un chemin simple et efficace pour gérer le trafic interserveur.

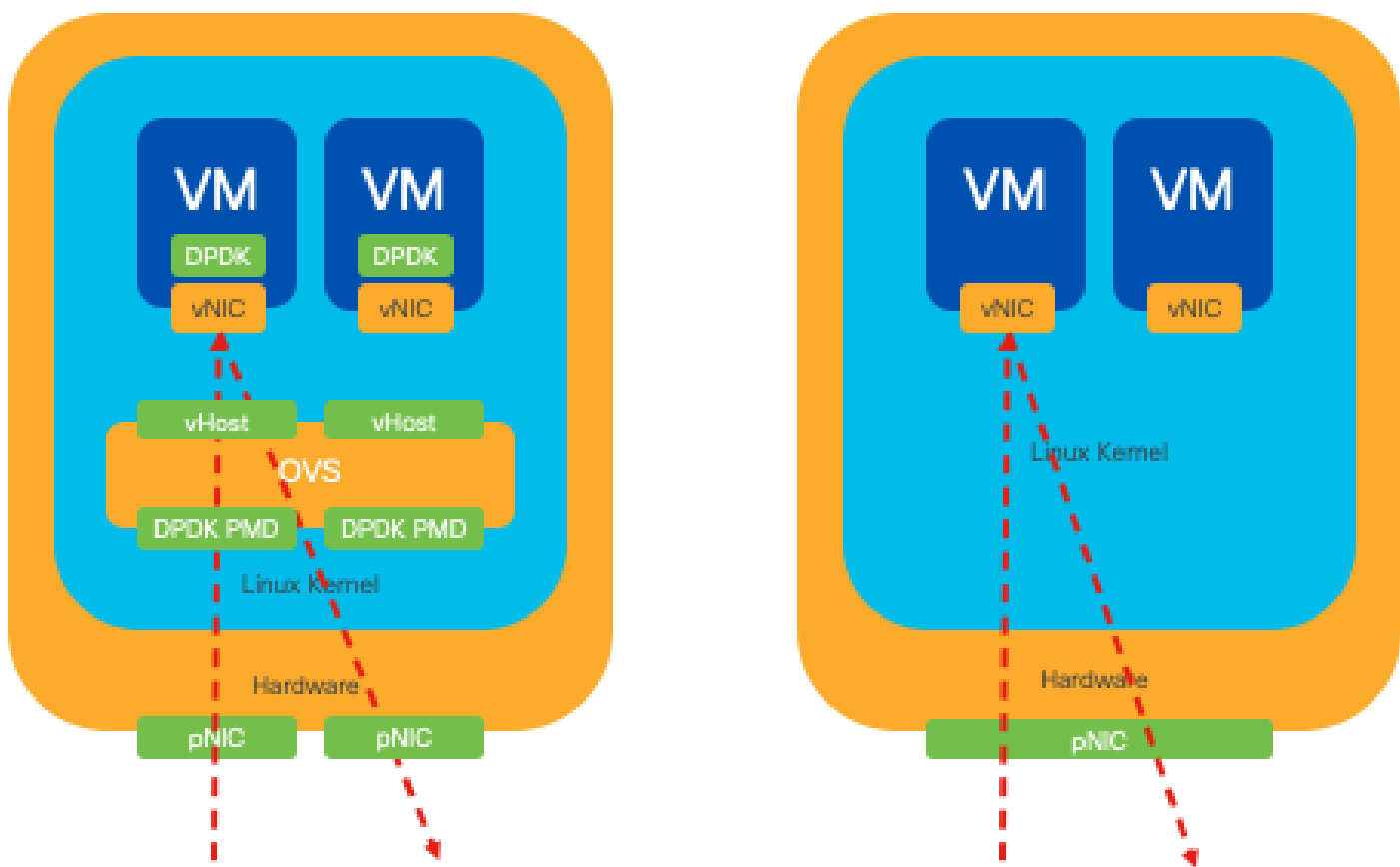
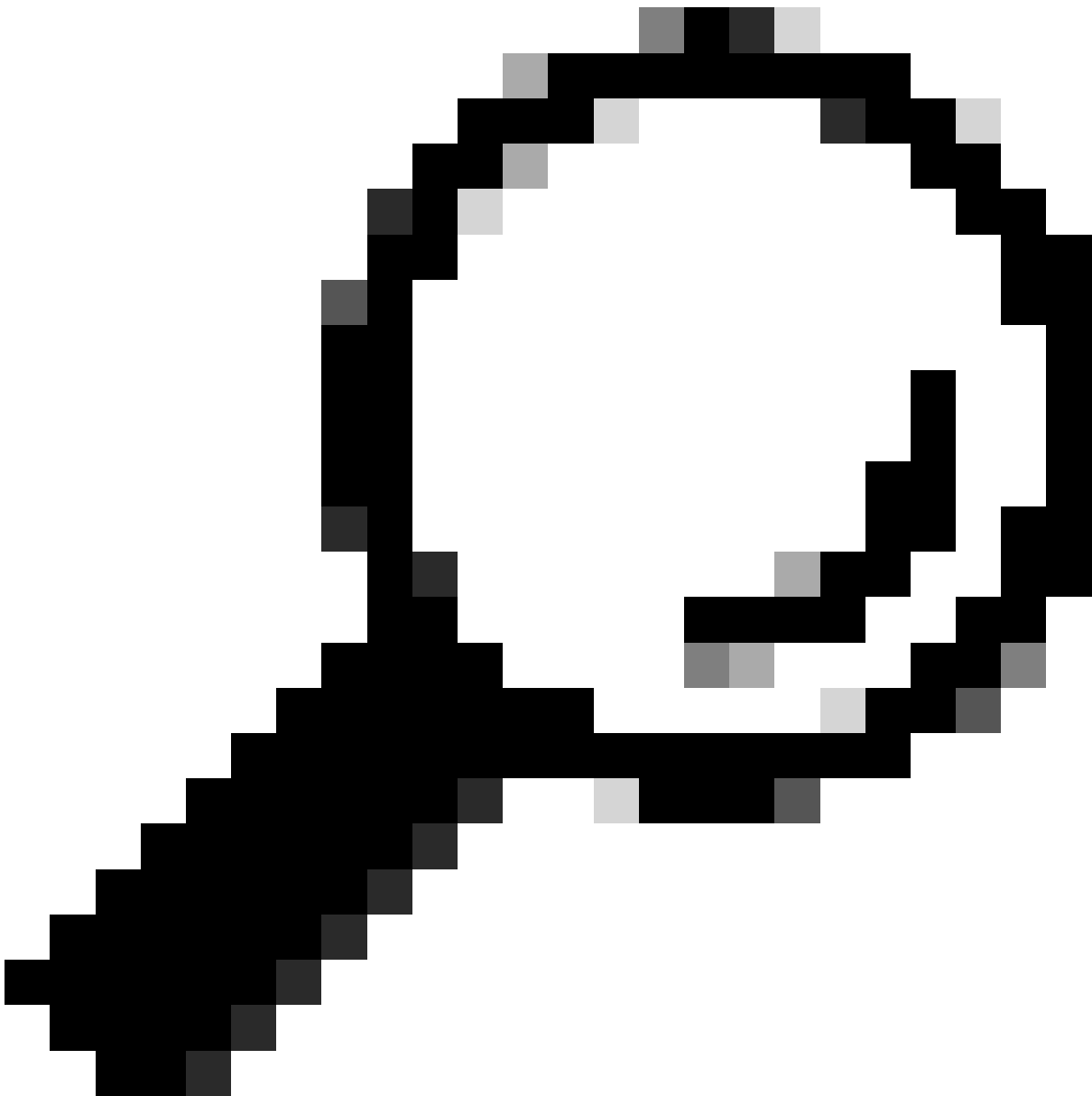


Figure 9. Traversée de paquets DPDK et SR-IOV dans le trafic Nord-Sud



Conseil : n'oubliez pas qu'il est possible d'améliorer les performances d'une configuration basée sur SR-IOV en intégrant SR-IOV avec DPDK dans une fonction de réseau virtuel (VNF), à l'exception du scénario où DPDK est utilisé en association avec OVS comme décrit précédemment.

Configuration

Activation de DPDK

Pour activer DPDK à partir de l'interface graphique utilisateur, vous devez accéder à Configuration > Virtual Machine > Networking > Networks. Une fois dans le menu, cliquez sur le commutateur pour activer la fonction

Networks

Networks Information and Configuration

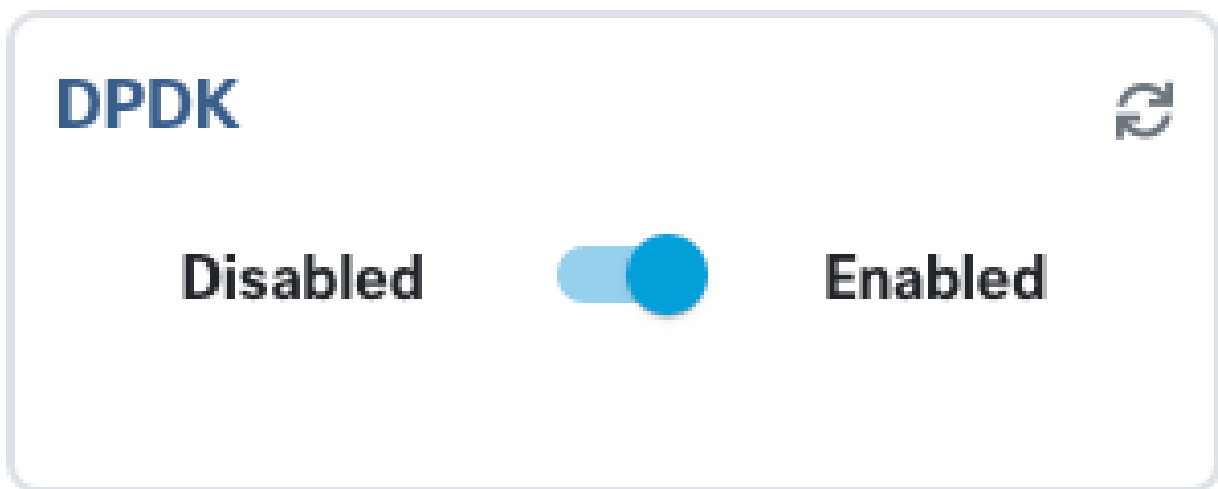
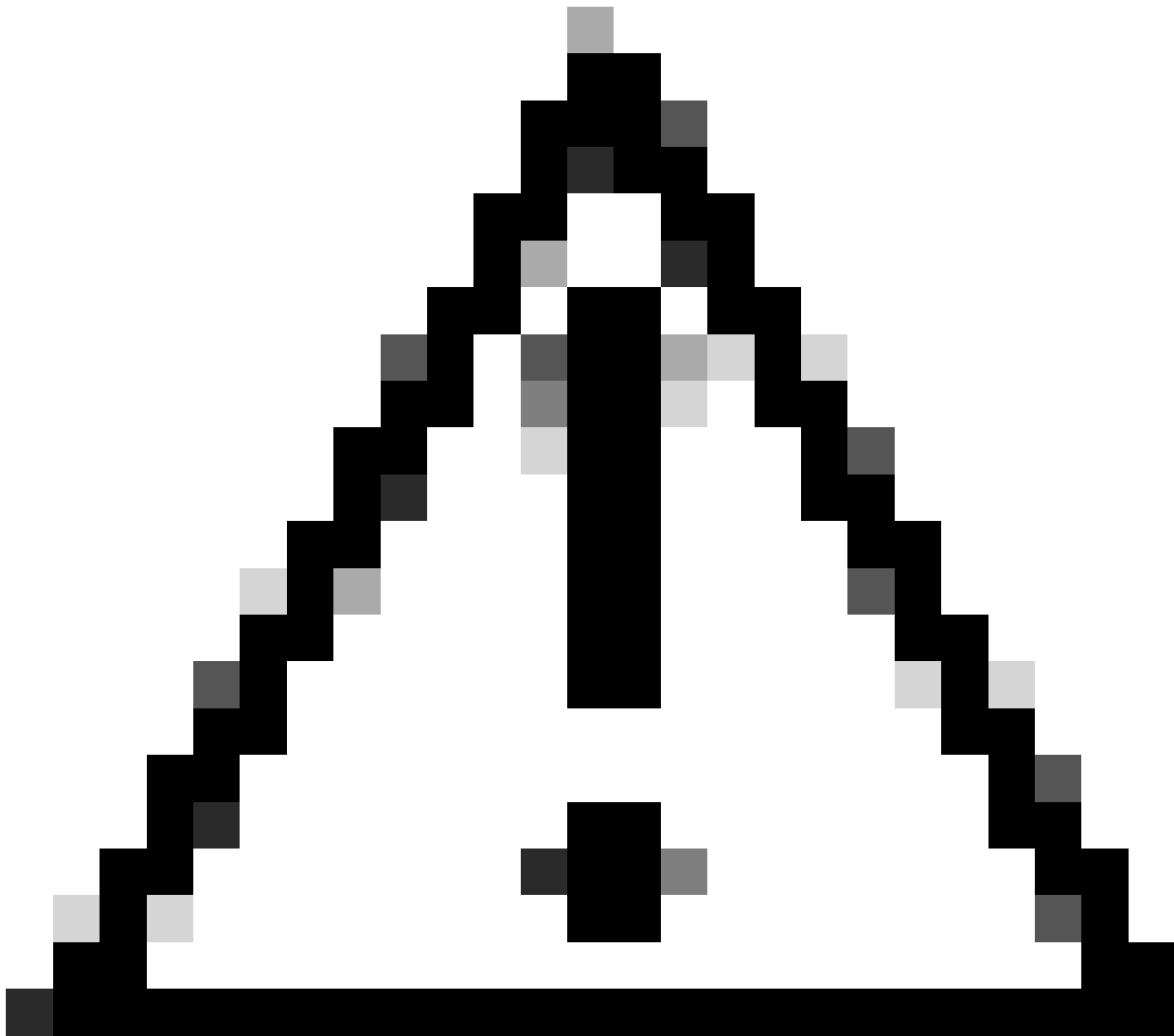


Figure 10. Bouton Diapositive disponible sur l'interface utilisateur graphique pour l'activation DPKD

Pour l'interface de ligne de commande, vous devez l'activer à partir des paramètres système globaux en mode de configuration.



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



Attention : DPDK ne peut pas être désactivé à moins qu'une réinitialisation en usine soit effectuée à partir de NFVIS.

Créer un nouveau réseau et l'associer à un nouveau pont OVS

Accédez à Configuration > Virtual Machine > Networking > Networks. Une fois que vous êtes sur la page Réseaux, cliquez sur le signe plus (+) en haut à gauche pour le tableau Réseaux,








Networks								Total Record: 3	<input type="text" value="search in all record"/>	
#	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	 		

Figure 11. Vue Tableau des réseaux de l'interface graphique NFVIS

Attribuez un nom au réseau et associez-le à un nouveau pont. Les options de liaison de VLAN et d'interface peuvent dépendre des besoins de l'infrastructure réseau.

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

Figure 12. Module « Add Network » pour la création de réseaux virtuels dans l'interface graphique NFVIS

Après avoir cliqué sur le bouton Envoyer, vous devez être en mesure de revoir le réseau nouvellement créé ajouté à la table Réseaux.



The screenshot shows the 'Add Network' module interface. At the top left, there is a '+ [document icon]' button. At the top right, there is a refresh icon (a circular arrow) highlighted with a red circle. Below the header, the text 'Total Record: 4' and a search bar 'search in all record' are visible. The main content is a table with the following columns: #, Network, Mode, Vlan, Vlan-Range, Native Vlan, Bridge, Interface, and Action. The table contains four rows of network data.

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

Figure 13. Vue Tableau des réseaux de l'interface graphique NFVIS, où l'icône Actualiser se trouve dans le coin supérieur droit (surligné en rouge)



Remarque : si le nouveau réseau n'est pas observé sur la table, cliquez sur le bouton d'actualisation en haut à droite ou actualisez la page entière.

S'il est effectué à partir de l'interface de ligne de commande, chaque réseau et pont est créé à partir du mode de configuration, le workflow est identique à la version de l'interface utilisateur graphique.

1. Créez le nouveau pont.

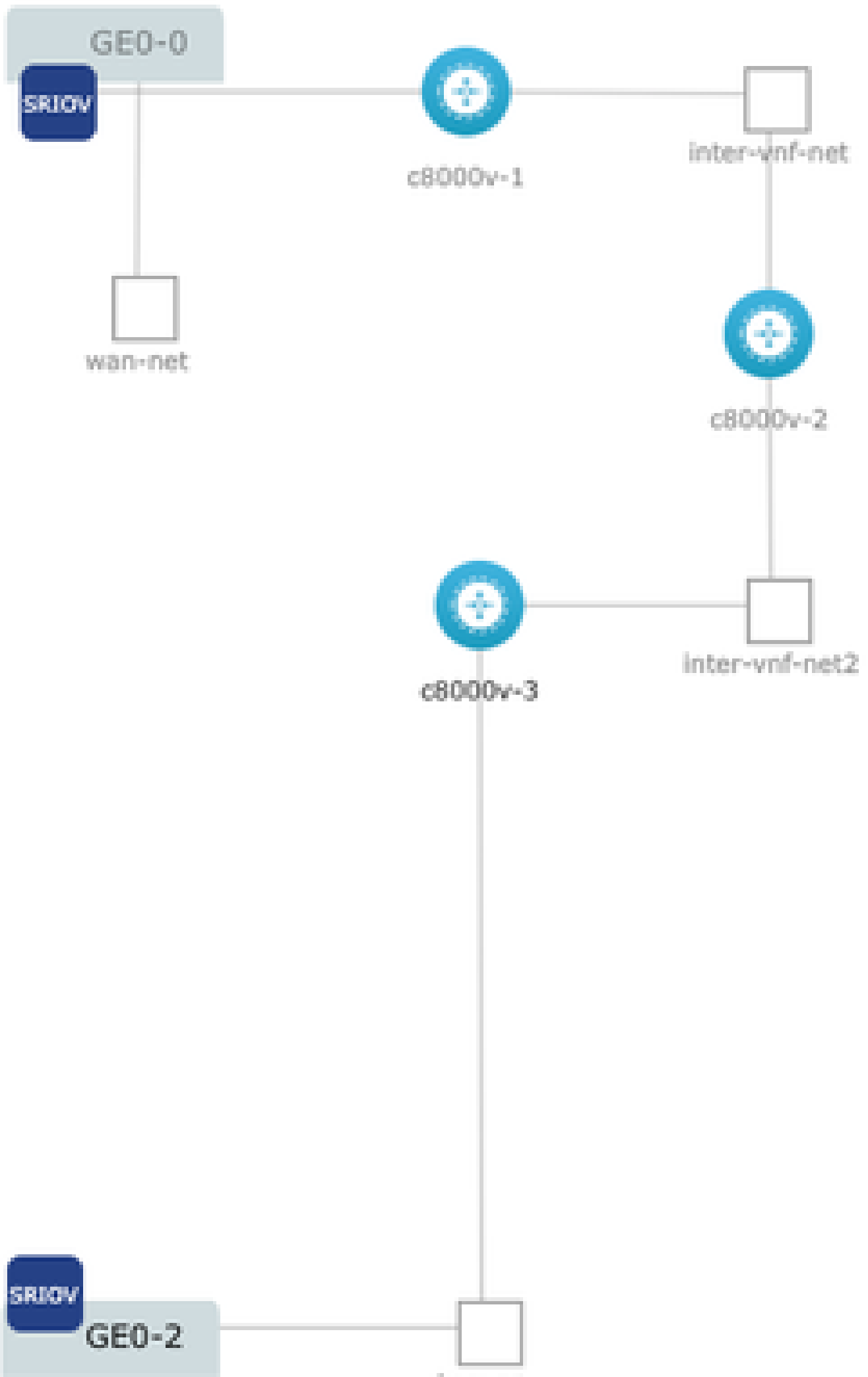
```
nfvis(config)# bridges bridge inter-vnf-br2  
nfvis(config-bridge-inter-vnf-br2)# commit
```

2. Créez un nouveau réseau et associez-le au pont précédemment créé

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

Connexion de VNF

Pour commencer avec une topologie de réseau ou un déploiement VFN unique, vous devez naviguer vers Configuration > Deploy. Vous pouvez faire glisser une machine virtuelle ou un conteneur de la liste de sélection vers la zone d'élaboration de la topologie pour commencer à créer votre infrastructure virtualisée.



À propos de cette traduction

Cisco a traduit ce document en traduction automatisée vérifiée par une personne dans le cadre d'un service mondial permettant à nos utilisateurs d'obtenir le contenu d'assistance dans leur propre langue.

Il convient cependant de noter que même la meilleure traduction automatisée ne sera pas aussi précise que celle fournie par un traducteur professionnel.