

# NFVIS virtuele netwerken begrijpen: OVS, DPDK en SR-IOV

## Inhoud

---

[Inleiding](#)

[Gebruikte componenten](#)

[Overzicht van netwerken in NFVIS](#)

[ENCS E540X-platform](#)

[Catalyst 8200 CPE switch](#)

[Catalyst 3800 CPE-1N20 switch](#)

[Virtualisatietechnologieën voor netwerken](#)

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

[OVS-bruggen](#)

[Tekorten bij contextswitching](#)

[DPDK \(Data Plane Development Kit\)](#)

[Gegevens kopiëren](#)

[PCIe-doorgifte](#)

[I/O-virtualisatie met één sleuf \(SR-IOV\)](#)

[Fysieke functies \(PF's\)](#)

[Virtuele functies \(VLAN's\)](#)

[Aanbevolen stuurprogramma's voor SR-IOV versnelling op voor NFVIS geschikte hardware](#)

[Use Cases voor DPDK en SR-IOV](#)

[DPDK-voorkeur](#)

[SR-IOV voorkeur](#)

[Configuratie](#)

[DPDK inschakelen](#)

[Een nieuw netwerk maken en aan een nieuwe OVS-brug koppelen](#)

[VPN's aansluiten](#)

[Verwante artikelen en documentatie](#)

---

## Inleiding

In dit document wordt het virtuele netwerkschema beschreven dat het NFVIS-platform biedt voor communicatie tussen VNF's in ondernemings- en servicenetwerken.

## Gebruikte componenten

De informatie in dit document is gebaseerd op deze hardware- en softwarecomponenten:

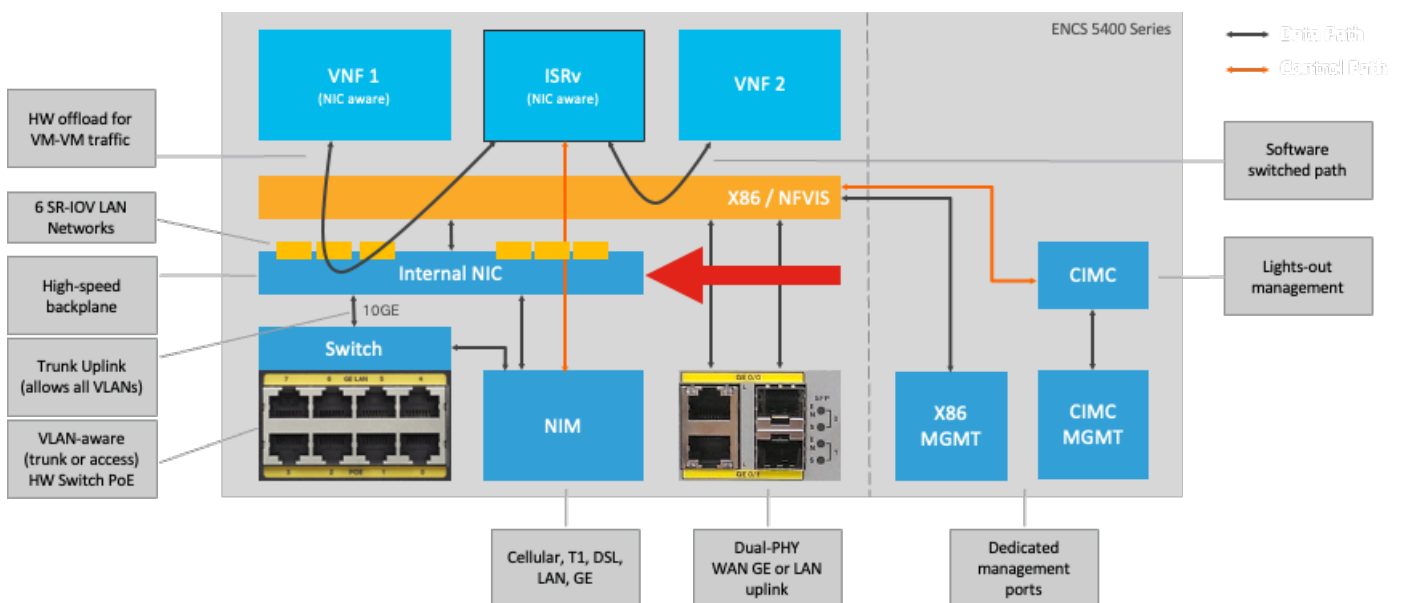
- ENCS 5412 met NFVIS 4.7.1-FC4
- c8300 CPE 1N20 met NFVIS 4.12.1-FC2

De informatie in dit document is gebaseerd op de apparaten in een specifieke laboratoriumomgeving. Alle apparaten die in dit document worden beschreven, hadden een opgeschoonde (standaard)configuratie. Als uw netwerk live is, moet u zorgen dat u de potentiële impact van elke opdracht begrijpt.

## Overzicht van netwerken in NFVIS

Intern wordt een intern beheernetwerk (int-mgmt-net) en een bridge (int-mgmt-br) gebruikt voor VNF-bewaking, waarbij IP-adressen voor beheerdoeleinden worden toegewezen aan het 10.20.0.0/24-submenu.

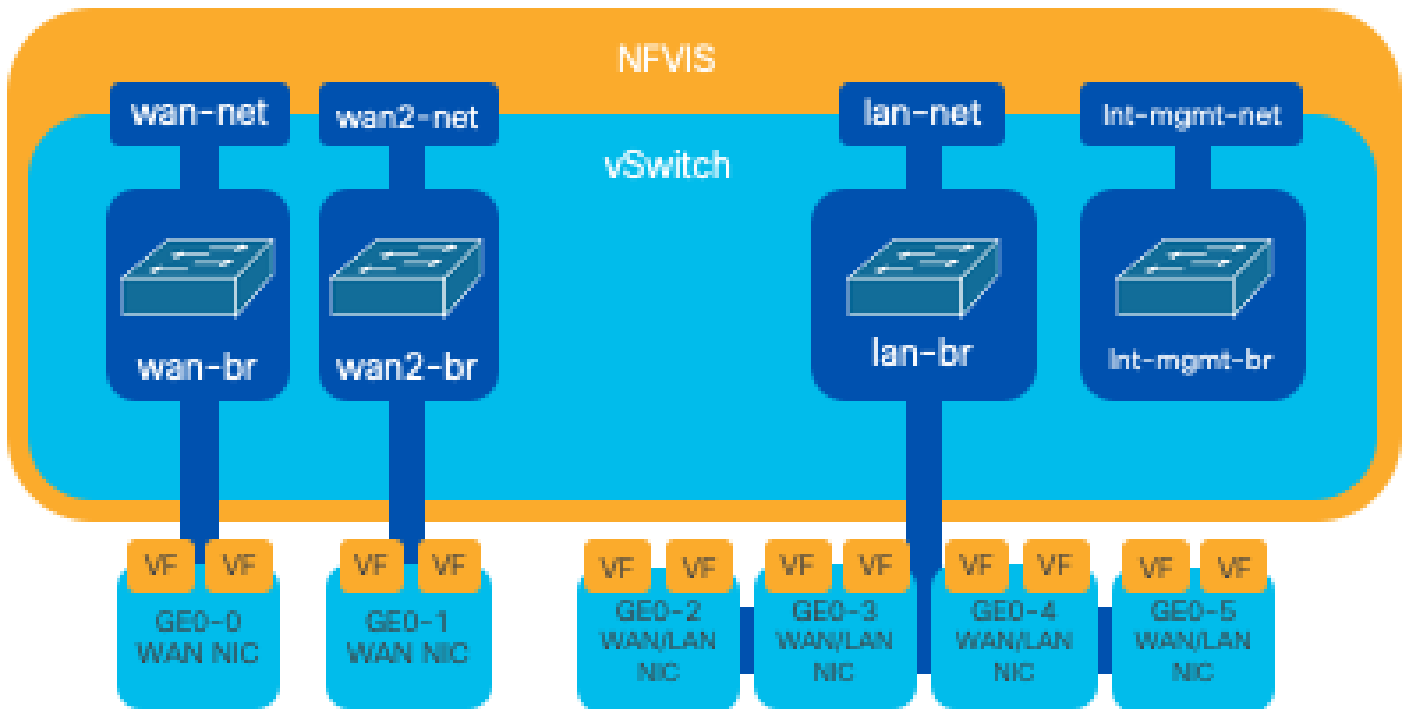
### ENCS E540X-platform



Afbeelding 1. Interne verbindingen voor hardware-Switch en WAN/LAN-uplink-NIC's

### Catalyst 8200 CPE switch

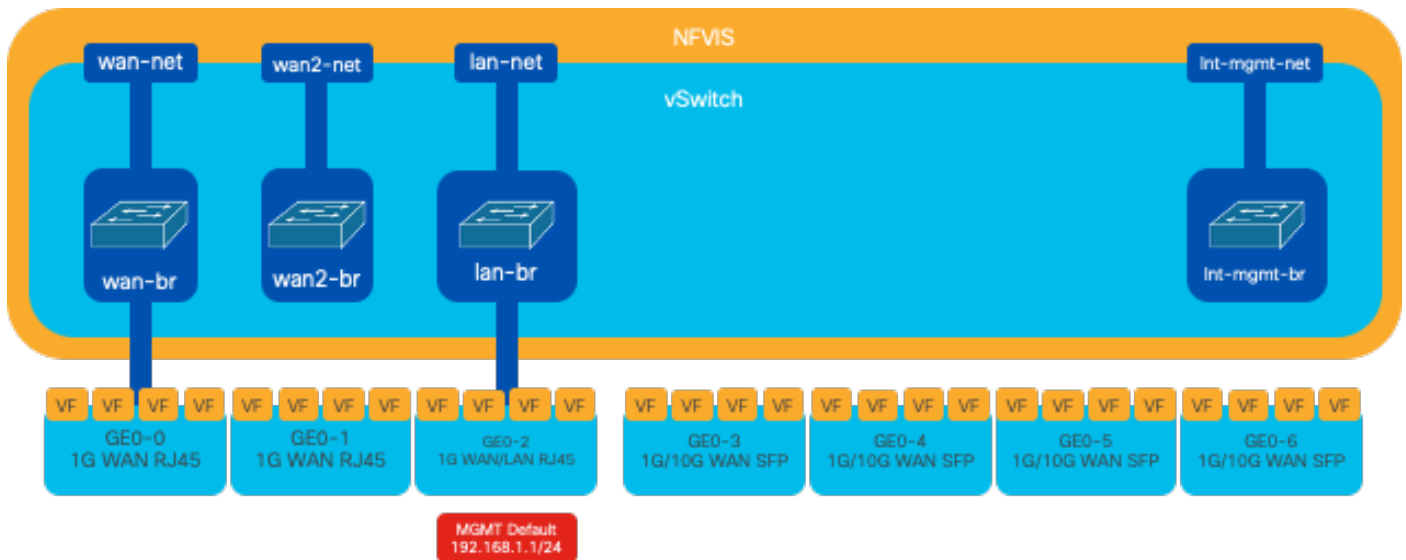
- NFVIS is standaard toegankelijk via de WAN-poort of GE0/2 LAN-poort voor beheer.
- WAN-netwerk (wan-net en wan2-net) en WAN-bridge (wan-br en wan2-br) zijn ingesteld om DHCP standaard in te schakelen. GE0-0 is standaard gekoppeld aan WAN-brug en GE0-1 aan WAN2-brug.
- Het IP-beheeradres 192.168.1.1 op Catalyst 8200 UPE is toegankelijk via GE0-2.
- GE0-2 is gekoppeld aan LAN-brug.
- Er wordt een intern beheernetwerk (int-mgmt-net) en een bridge (int-mgmt-br) gecreëerd en intern gebruikt voor systeembewaking.



Afbeelding 2. Interne overbrugging en virtuele switches toegewezen aan de 8200 NIC's

### Catalyst 3800 CPE-1N20 switch

1. NFVIS is standaard toegankelijk via de FPGE (Front Panel Gigabit Ethernet) WAN-poorten via de GE0-2 LAN-poort voor beheer
2. WAN-netwerk (WAN-net) en een WAN-brug (WAN-br) worden standaard ingesteld om DHCP mogelijk te maken. GE0-0 is standaard gekoppeld aan WAN-brug
3. WAN-netwerk (WAN2-net) en een WAN-brug (WAN2-br) worden standaard gemaakt, maar niet gekoppeld aan fysieke poorten
4. GE0-2 is gekoppeld aan LAN-brug; alle andere poorten zijn niet gekoppeld aan OVS
5. IP-beheer 192.168.1.1 op C8300-CPE is toegankelijk via GE0-2
6. Er wordt een intern beheernetwerk (int-mgmt-net) en een brug (int-mgmt-br) gecreëerd en intern gebruikt voor systeembewaking.

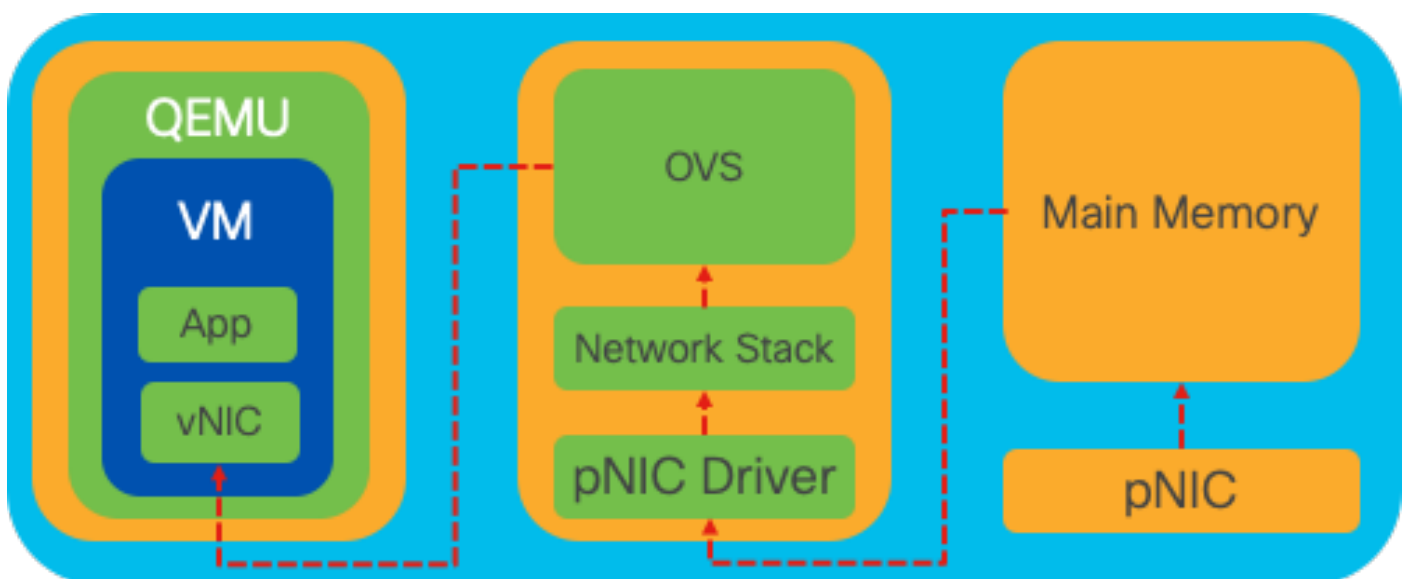


Afbeelding 3. Interne overbrugging en virtuele switches toegewezen aan de 8300 NIC's

## Virtualisatietechnologieën voor netwerken

### Open vSwitch (OVS)

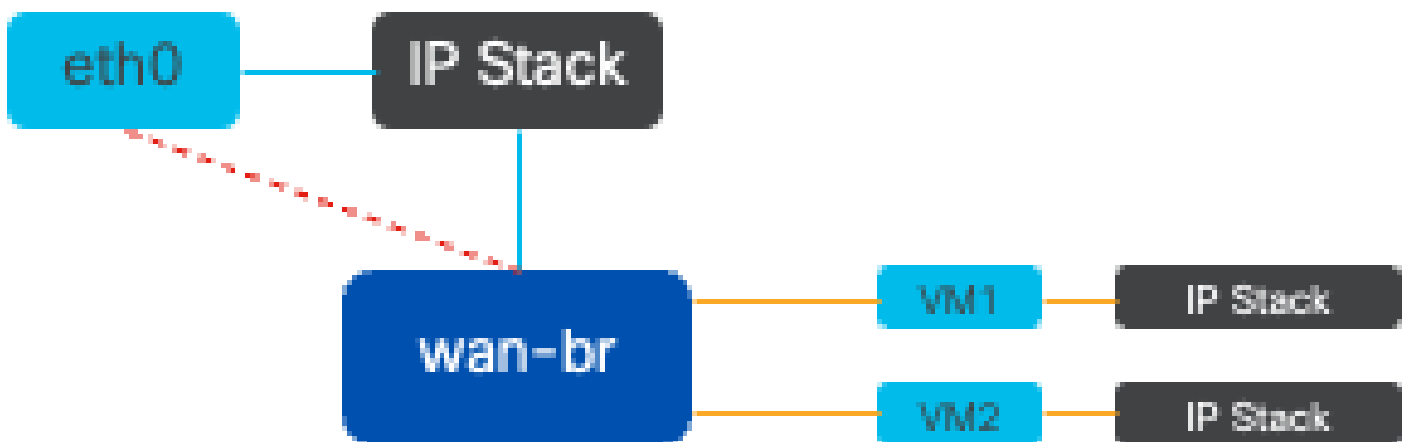
Open vSwitch (OVS) is een open-source, meerlaagse virtuele switch die is ontworpen om netwerkautomatisering mogelijk te maken via programmatische extensies, terwijl ondersteuning wordt geboden voor standaardbeheerinterfaces en -protocollen, zoals NetFlow, sFlow, IPFIX, RSPAN, CLI, LACP en 802.1ag. Het wordt veel gebruikt in grote gevirtualiseerde omgevingen, met name met hypervisors om netwerkverkeer tussen virtuele machines (VM's) te beheren. Het maakt de creatie mogelijk van geavanceerde netwerktopologieën en -beleid dat direct wordt beheerd via de NFVIS-interface, waardoor een veelzijdige omgeving voor netwerkfunctievirtualisatie wordt geboden.



Afbeelding 4. OVS configuratie binnen de Linux kernel

## OVS-bruggen

Het gebruikt virtuele netwerkbruggen en stroomregels om pakketten tussen hosts door te sturen. Het gedraagt zich als een fysische switch, alleen gevirtualiseerd.



Afbeelding 5. Voorbeeld van implementatie van 2 VM's of VNF's die aan de wan-br-brug zijn gekoppeld

### Tekorten bij contextswitching

Wanneer een netwerkpakket bij een netwerkinterfacekaart (NIC) aankomt, brengt het een onderbreking teweeg, een signaal aan de processor erop wijst die dat het onmiddellijke aandacht vergt. De CPU onderbreekt zijn huidige taken om de interrupt te verwerken, een proces dat bekend staat als interrupt processing. Tijdens deze fase leest de CPU, onder de controle van de kernel van het besturingssysteem, het pakket van de NIC in het geheugen en bepaalt de volgende stappen op basis van de pakketbestemming en het doel. Het doel is het pakket snel te verwerken of te routeren naar de beoogde toepassing, waardoor de latentie wordt geminimaliseerd en de doorvoersnelheid wordt gemaximaliseerd.

Context switching is het proces waarbij een CPU switches taken in een omgeving (context) naar een andere uitvoeren. Dit is met name van belang bij het overschakelen tussen de gebruikersmodus en de kernelmodus:

- Gebruikersmodus: Dit is een beperkte verwerkingsmodus waarin de meeste toepassingen worden uitgevoerd. Toepassingen in de gebruikersmodus hebben geen directe toegang tot hardware of referentiegeheugen en moeten communiceren met de kernel van het besturingssysteem om deze bewerkingen uit te voeren.
- Kernel Mode: Deze modus geeft het besturingssysteem volledige toegang tot de hardware en alle geheugen. De kernel kan elke CPU-instructie uitvoeren en verwijst naar elk geheugenadres. De kernmodus is vereist voor het uitvoeren van taken zoals het beheren van hardware-apparaten, geheugen en het uitvoeren van systeemaanroepen.

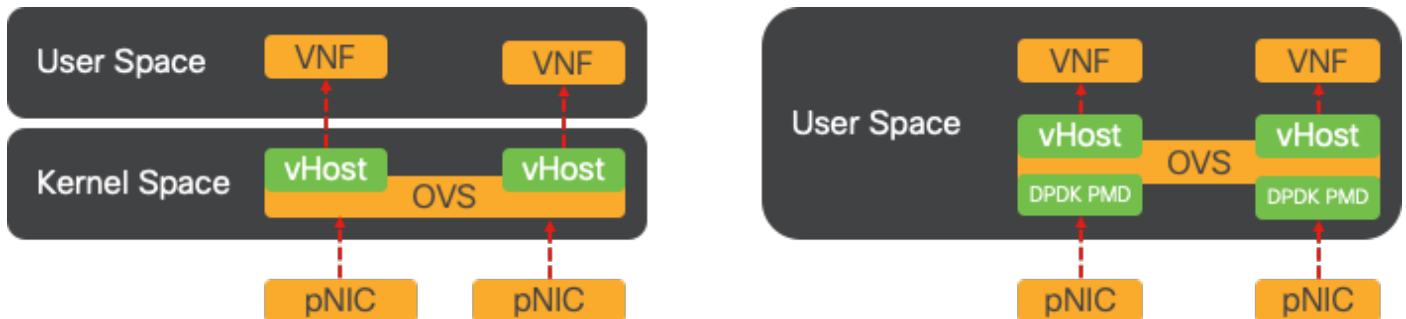
Wanneer een toepassing een handeling moet uitvoeren die rechten op kernelniveau vereist (zoals het lezen van een netwerkpakket), treedt er een context switch op. De CPU overgangen van de

gebruikersmodus naar de kernelmodus om de bewerking uit te voeren. Als deze optie is voltooid, zal een andere context switch de CPU naar de gebruikersmodus terugsturen om door te gaan met het uitvoeren van de toepassing. Dit schakelproces is kritiek voor het handhaven van systeemstabiliteit en veiligheid maar introduceert overheadkosten die prestaties kunnen beïnvloeden.

OVS draait voornamelijk in de gebruikersruimte van het besturingssysteem, wat een knelpunt kan worden als de gegevensdoorvoer toeneemt. Dit komt doordat de CPU meer context switches nodig heeft om naar de kernel-modus te kunnen overschakelen om pakketten te verwerken, waardoor de prestaties afnemen. Deze beperking is bijzonder merkbaar in omgevingen met hoge pakquetsnelheden of wanneer nauwkeurige timing van cruciaal belang is. Om deze prestatiebeperkingen aan te pakken en te voldoen aan de eisen van moderne snelle netwerken, zijn technologieën ontwikkeld zoals DPDK (Data Plane Development Kit) en SR-IOV (Single Root I/O Virtualization).

## DPDK (Data Plane Development Kit)

DPDK is een reeks bibliotheken en stuurprogramma's die zijn ontworpen om de pakketverwerkingswerkbelasting op een brede reeks CPU-architecturen te versnellen. Door de traditionele kernelnetwerkstack te omzeilen (contextswitching vermijden), kan DPDK de doorvoersnelheid van het dataplatform aanzienlijk verhogen en latentie verminderen. Dit is met name gunstig voor VNF's met hoge doorvoersnelheid die communicatie met een lage latentie vereisen, waardoor NFVIS een ideaal platform is voor prestatiegevoelige netwerkfuncties.



Afbeelding 6. Optimalisatie van traditionele OVS (linkerkant) en DPDK OVS (rechterkant) context switching

Ondersteuning voor DPDK voor OVS is gestart in NFVIS 3.10.1 voor ENCS en 3.12.2 voor andere platforms.

- Doorvoersnelheid voor serviceketens in de buurt van SRIOV, beter dan non-DPDK OVS.
- Virtueel stuurprogramma vereist voor VNF.
- Ondersteunde platforms:
- ENCS 3.10.1 en verder.
- UCS-C, CSP5K 3.12.1 en verder.
- DPDK voor poortkanalen ondersteund sinds 4.12.1.
- Packet-/verkeersopname: niet ondersteund in DPDK.
- Spanverkeer op PNIC: niet ondersteund in DPDK.

- Nadat OVS-DPDK is ingeschakeld, kan deze niet als afzonderlijke functie worden uitgeschakeld. De enige manier om DPDK uit te schakelen is door een fabrieksreset.

## Gegevens kopiëren

Bij traditionele netwerkbenaderingen moeten gegevens vaak meerdere malen worden gekopieerd voordat ze in het VM-geheugen op hun bestemming komen. Een pakket moet bijvoorbeeld worden gekopieerd van de NIC naar de kernel-ruimte, vervolgens naar de gebruikersruimte voor bewerking door een virtuele switch (zoals OVS) en ten slotte naar het VM-geheugen. Elke kopieerbewerking heeft een vertraging en verhoogt het CPU-gebruik ondanks de prestatieverbeteringen die DPDK biedt door de kernels-netwerkstack te omzeilen.

Deze overheadkosten omvatten geheugenkopieën en de verwerkingstijd die nodig is om pakketten in gebruikersruimte te verwerken voordat ze naar de VM kunnen worden doorgestuurd. PCIe Passthrough en SR-IOV pakken deze knelpunten aan door toe te staan dat een fysiek netwerkapparaat (zoals een NIC) direct wordt gedeeld tussen meerdere VM's zonder dat het host-besturingssysteem in dezelfde mate bij de traditionele virtualisatiemethoden wordt betrokken.

## PCIe-doorgifte

De strategie bestaat uit het omzeilen van de hypervisor om de Virtual Network Functions (VNFs) directe toegang tot een Network Interface Card (NIC) toe te staan, waarbij bijna maximale doorvoersnelheid wordt bereikt. Deze benadering staat bekend als PCI passthrough, waardoor een complete NIC gewijd kan worden aan een gastbesturingssysteem zonder tussenkomst van een hypervisor. In deze opstelling, werkt de virtuele machine alsof zijn direct verbonden met NIC. Bijvoorbeeld, met twee NIC kaarten beschikbaar, kan elke één exclusief worden toegewezen aan verschillende VNFs, die hen directe toegang verlenen.

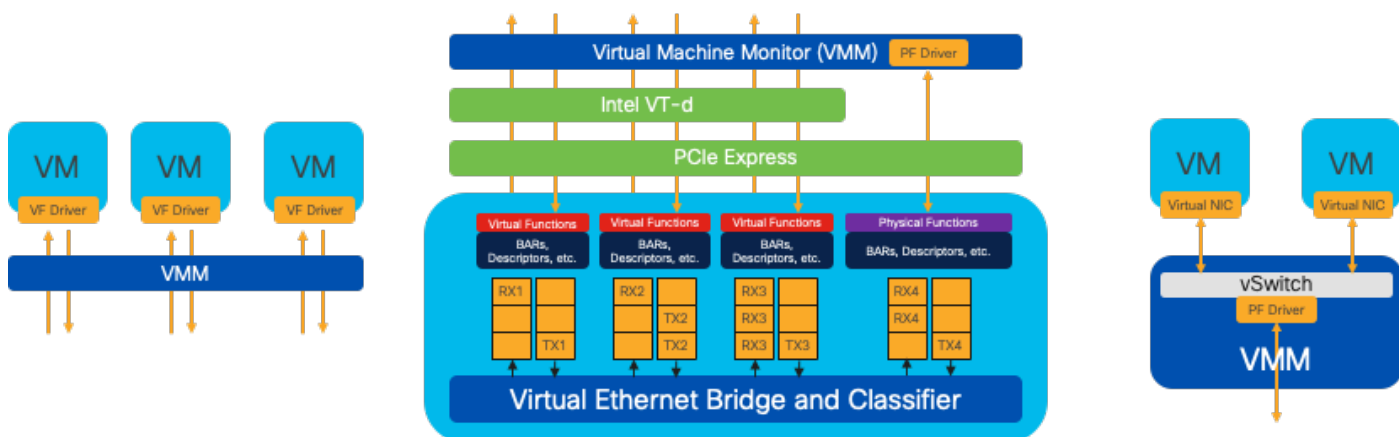
Deze methode heeft echter een nadeel: als slechts twee NIC's beschikbaar zijn en uitsluitend worden gebruikt door twee afzonderlijke VNF's, zouden extra VNF's, zoals een derde, geen NIC-toegang hebben omdat er geen specifieke NIC beschikbaar voor is. Een alternatieve oplossing is het gebruik van Single Root I/O-virtualisatie (SR-IOV).

## I/O-virtualisatie met één sleuf (SR-IOV)

Is een specificatie die het mogelijk maakt dat één fysiek PCI-apparaat, zoals een netwerkinterfacekaart (NIC), wordt weergegeven als meerdere afzonderlijke virtuele apparaten. Deze technologie biedt directe VM-toegang tot fysieke netwerkapparaten, wat de overhead vermindert en de I/O-prestaties verbetert. Het werkt door een enkel PCIe-apparaat te verdelen in meerdere virtuele segmenten, die elk kunnen worden toegewezen aan verschillende VM's of VNF's, waardoor de beperking die wordt veroorzaakt door een eindig aantal NIC's effectief wordt opgelost. Deze virtuele segmenten, bekend als Virtual Functions (VF's), maken gedeelde NIC-bronnen mogelijk tussen meerdere VNF's. De fysieke functie (PF) verwijst naar de fysieke component die de SR-IOV mogelijkheden vergemakkelijkt.

Door gebruik te maken van SR-IOV kan NFVIS speciale NIC-bronnen toewijzen aan specifieke VNF's, waardoor hoge prestaties en lage latentie worden gegarandeerd door Direct Memory

Access (DMA) van netwerkpakketten rechtstreeks in het betreffende VM-geheugen te faciliteren. Deze benadering minimaliseert de CPU betrokkenheid bij het verwerken van de pakketten en verlaagt zo het CPU-gebruik. Dit is met name handig voor toepassingen die een gegarandeerde bandbreedte nodig hebben of waarvoor strikte prestatievereisten gelden.



Afbeelding 7. NFVIS SR-IOV PCIe Resources Scheiding via hardwarefuncties

### Fysieke functies (PF's)

Het zijn volledig-gekenmerkte functies PCIe en verwijzen naar een speciaal gebouwde hardwaredoos die specifieke voorzien van een netwerkfunctie verstrekt; dit zijn volledig gekenmerkte functies PCIe die kunnen worden ontdekt, worden beheerd en, zoals een ander apparaat PCIe worden gemanipuleerd. De fysieke functies omvatten de mogelijkheden van SR-IOV die kunnen worden gebruikt om een apparaat te vormen en te controleren PCIe.

### Virtuele functies (VLAN's)

Ze zijn gestroomlijnde functies met minimale configuratiebronnen (lichtgewicht) en richten zich uitsluitend op het verwerken van I/O als eenvoudige PCIe-functies. Elke virtuele functie is afkomstig van een fysieke functie. De hardware van het apparaat beperkt het mogelijke aantal virtuele functies. Eén Ethernet-poort, het fysieke apparaat, kan corresponderen met talrijke virtuele functies, die vervolgens kunnen worden toegewezen aan verschillende virtuele machines.

### Aanbevolen stuurprogramma's voor SR-IOV versnelling op voor NFVIS geschikte hardware

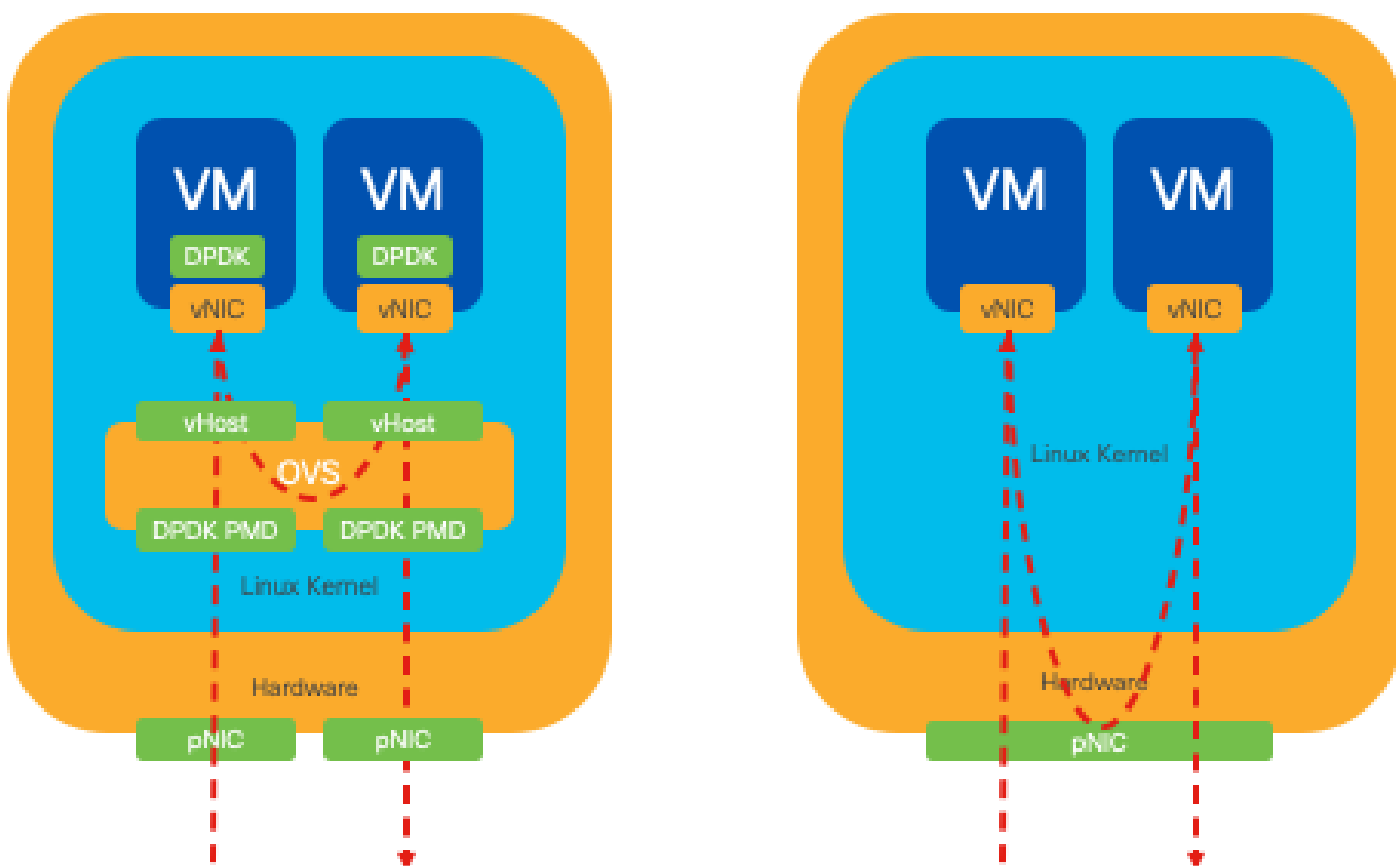
Platform	NIC(s)	NIC-stuurprogramma
ENCS 54XX switch	Backplane Switch	i40e
ENCS 54XX switch	GE0-0 en GE0-1	igb
Catalyst 8200 CPE switch	GE0-0 en GE0-1	ixgbe
Catalyst 8200 CPE switch	GE0-2 en GE0-5	igb

### Use Cases voor DPDK en SR-IOV



## DPDK-voorkeur

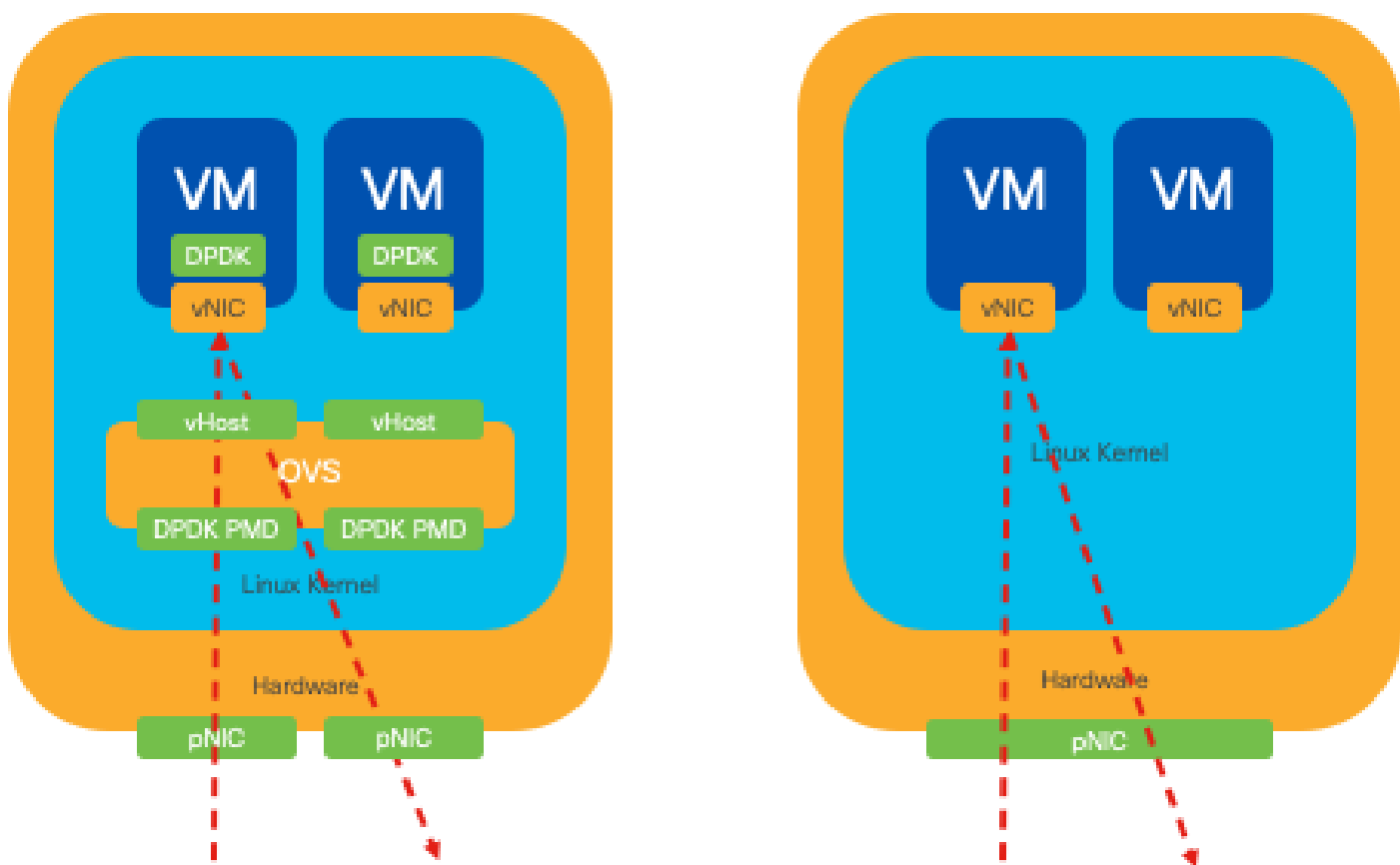
Met name in scenario's waar het netwerkverkeer voornamelijk oost-west stroomt (wat betekent dat het binnen dezelfde server blijft), presteert DPDK SR-IOV beter. De redenering is eenvoudig: als het verkeer intern wordt beheerd binnen de server zonder toegang tot de NIC te hebben, levert SR-IOV geen voordeel op. In feite kan SR-IOV mogelijk tot inefficiënties leiden door het verkeerspad onnodig uit te breiden en NIC-bronnen te gebruiken. Daarom is het gebruik van DPDK voor intern beheer van serververkeer de efficiëntere keuze.



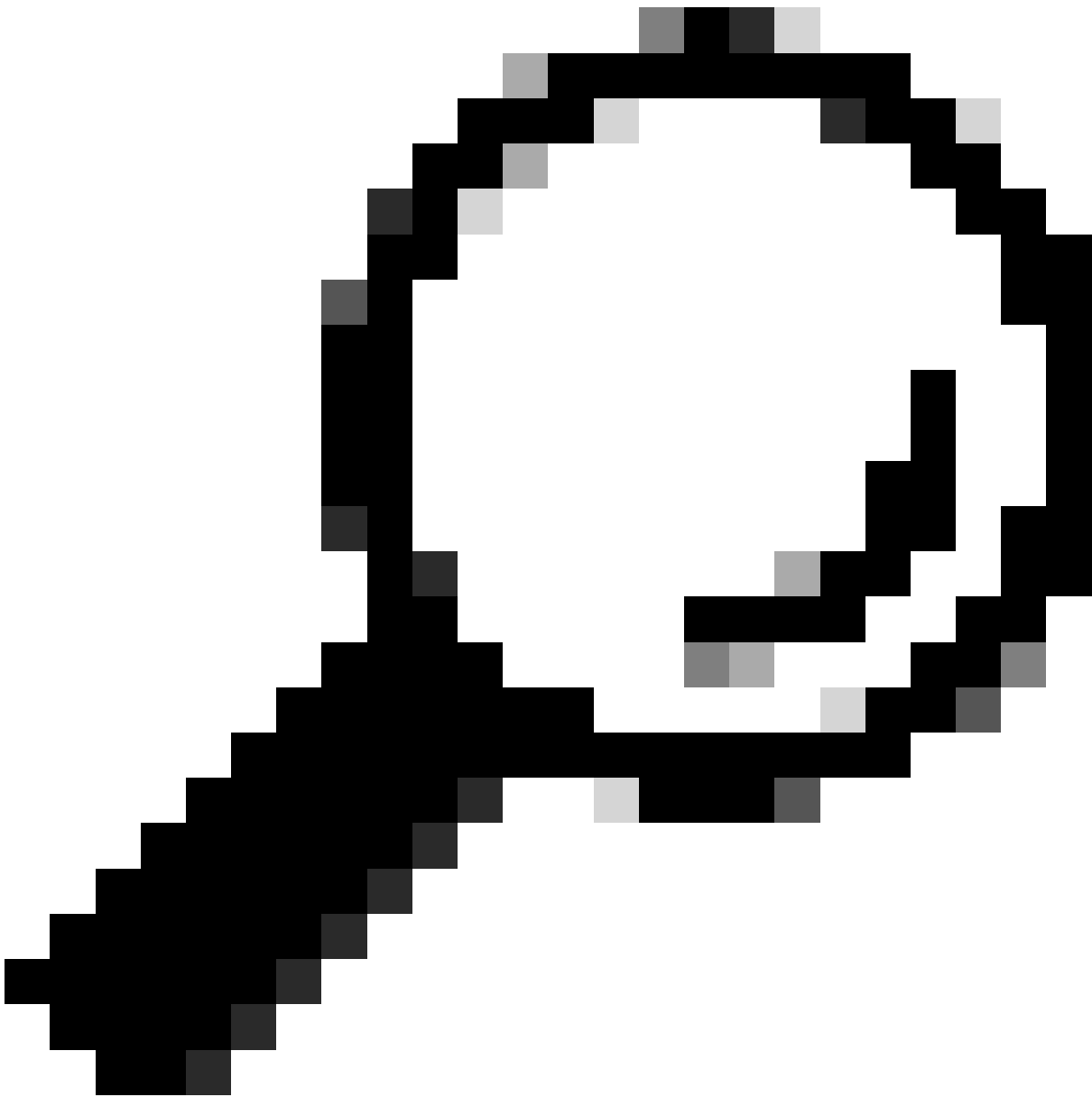
Afbeelding 8. DPDK en SR-IOV pakkettransmissie in oost-naar-westverkeer

## SR-IOV voorkeur

In situaties waarin netwerkverkeer van noord naar zuid, of zelfs van oost naar west, maar specifiek tussen servers stroomt, blijkt het gebruik van SR-IOV voordeliger dan DPDK. Dit geldt met name voor de communicatie van server naar server. Gezien het feit dat dergelijk verkeer onvermijdelijk over de NIC moet lopen, zou het kiezen voor DPDK-verbeterde OVS onnodig extra complexiteit en potentiële prestatiebeperkingen kunnen introduceren. Daarom is SR-IOV in deze omstandigheden de meest geschikte keuze, omdat het een eenvoudige en efficiënte route biedt voor de afhandeling van verkeer tussen servers.



Afbeelding 9. DPDK en SR-IOV pakketverkeer in Noord-naar-Zuid verkeer



Tip: Onthoud dat het mogelijk is om de prestaties van een op SR-IOV gebaseerde setup te verbeteren door SR-IOV te integreren met DPDK in een Virtual Network Function (VNF), met uitzondering van het scenario waarin DPDK wordt gebruikt in combinatie met OVS zoals eerder beschreven.

---

## Configuratie

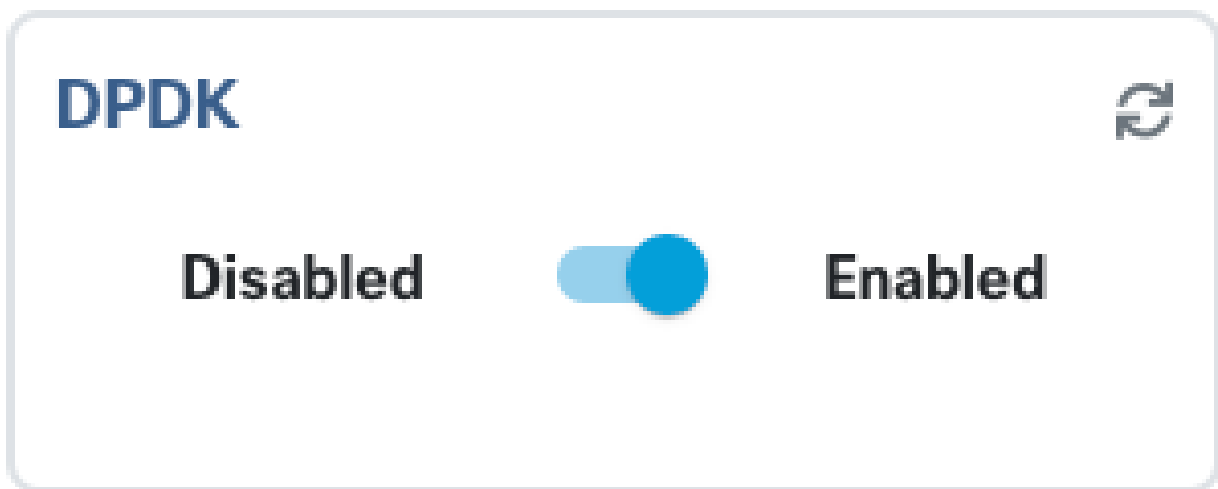
### DPDK inschakelen

Om DPDK in te schakelen vanuit de GUI, moet u naar Configuration > Virtual Machine > Networks > Networks navigeren. Klik zodra u in het menu staat op de switch om de functie te activeren

# Networks

## Networks Information and Configuration

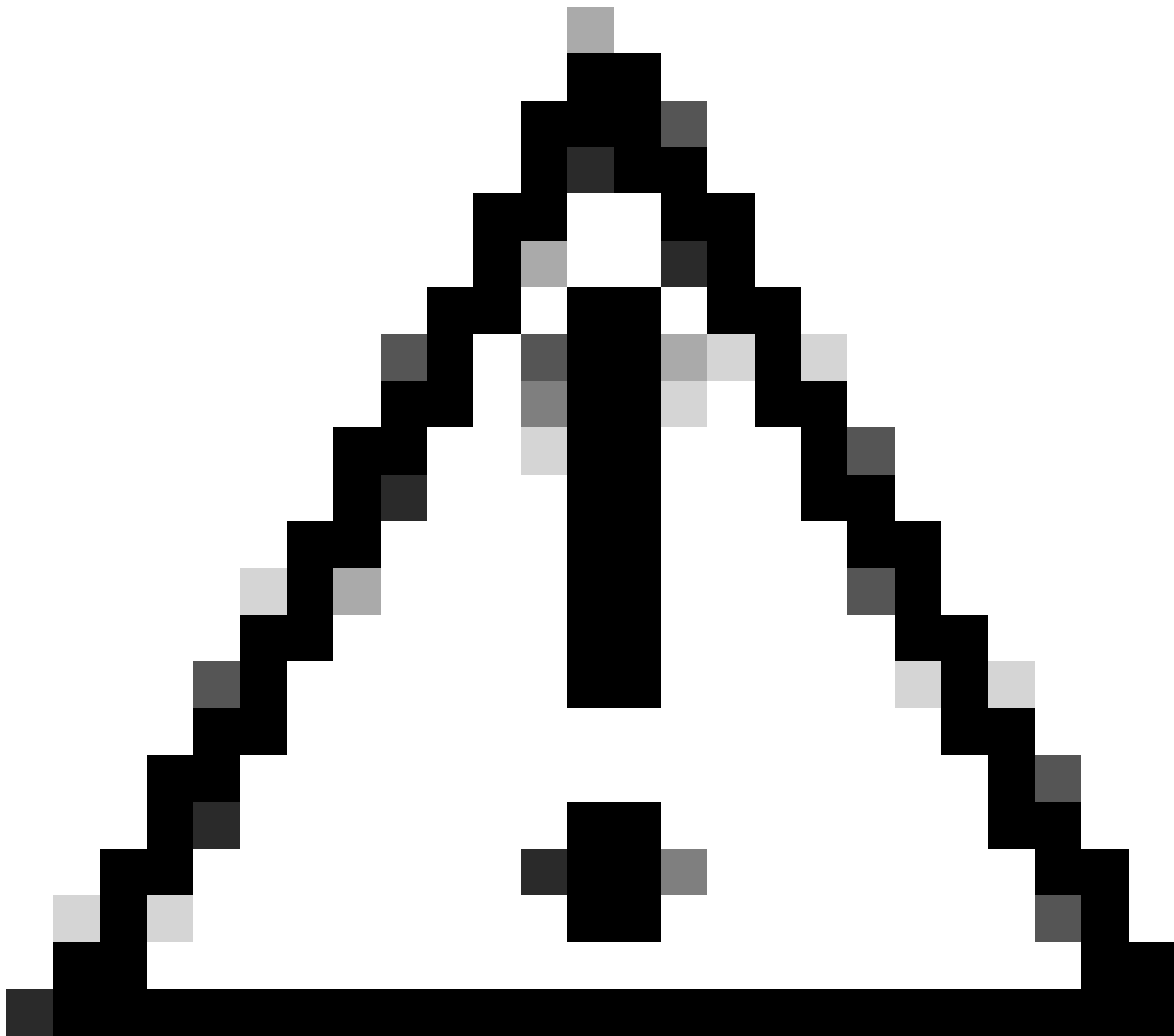
---



Afbeelding 10. Slide-knop beschikbaar op de GUI voor DPDK-activering

Voor de CLI moet u deze in de configuratiemodus inschakelen uit de globale systeeminstellingen.

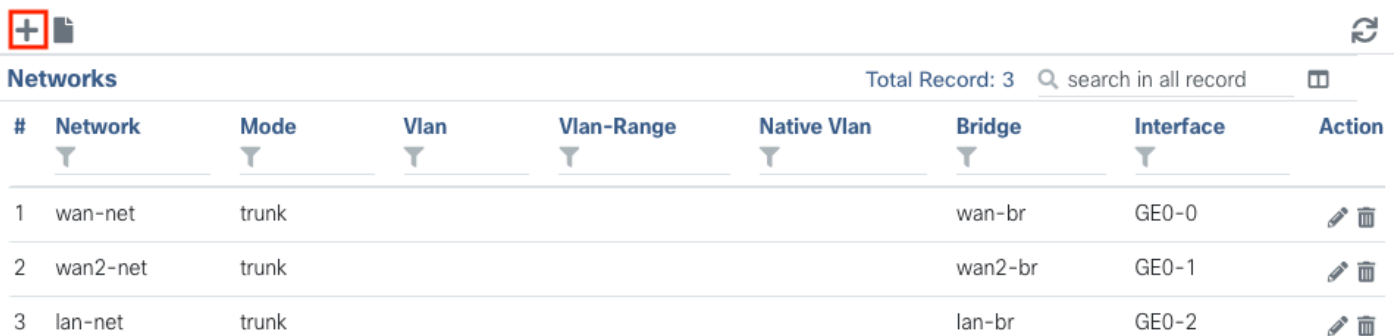
```
nfvis(config)# system settings dpdk enable
```









Let op: DPDK kan niet worden uitgeschakeld tenzij NFVIS een fabrieksreset uitvoert.

## Een nieuw netwerk maken en aan een nieuwe OVS-brug koppelen

Ga naar Configuratie > Virtuele machine > Netwerken > Netwerken. Zodra u op de pagina Netwerken bent, klikt u op links bovenaan het plusteken (+) voor de tabel Netwerken.



#	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	 

Afbeelding 11. Tabelweergave van netwerken vanuit de NFVIS GUI

Geef het netwerk een naam en koppel aan een nieuwe brug. VLAN- en interfacebindingsopties kunnen afhangen van de behoeften van de netwerkinfrastructuur.

# 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

Afbeelding 12. Modal voor "Add Network" waarmee virtuele netwerken kunnen worden gemaakt in de NFVIS GUI

Nadat u op de knop Verzenden hebt geklikt, moet u het nieuwe netwerk kunnen bekijken dat aan de tabel Netwerken is toegevoegd.



The screenshot shows a table titled "Networks" with the following columns: #, Network, Mode, Vlan, Vlan-Range, Native Vlan, Bridge, Interface, and Action. There are four rows of data. A search bar and a refresh icon (circled in red) are visible in the top right corner.

#	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		 

Afbeelding 13. Netwerktafelweergave vanuit de NFVIS GUI, waar het "Vernieuwingspictogram" rechtsboven staat (gemarkeerd in rood)





Opmerking: Als het nieuwe netwerk niet wordt waargenomen in de tabel, klikt u rechtsboven op de knop Vernieuwen of ververs u de gehele pagina.

---

Indien uitgevoerd binnen van CLI, wordt elk netwerk en de brug van configuratiewijze gemaakt, is het werkschema het zelfde als de versie GUI.

1. Maak de nieuwe brug.

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

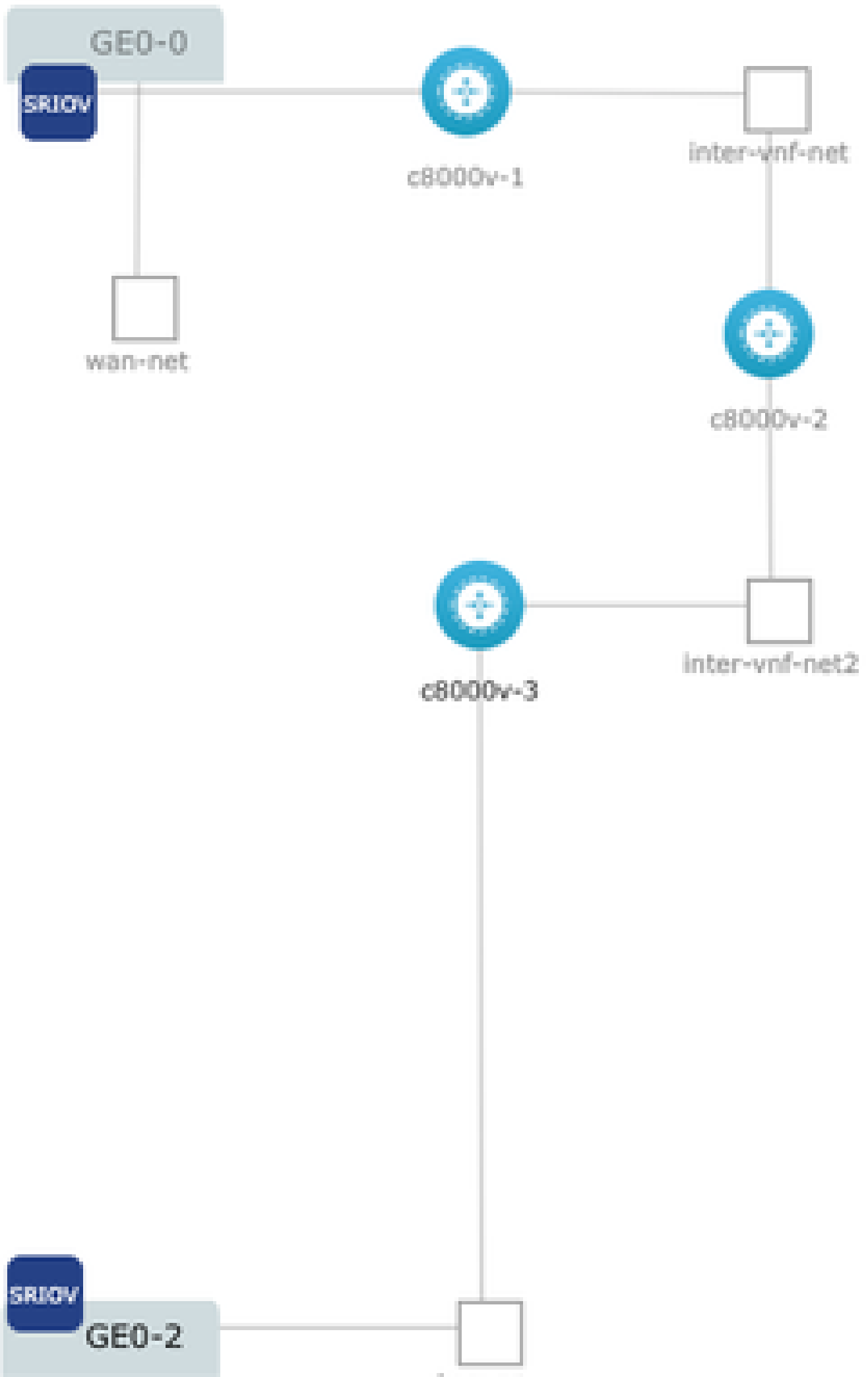
2. Maak een nieuw netwerk en koppel het aan de eerder gemaakte brug

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

```
nfvis(config-network-inter-vnf-net2)# commit
```

## VPN's aansluiten

Om te beginnen met een netwerktopologie of één VFN-implementatie, moet u naar Configuration > Implementation navigeren. U kunt een VM of Container van de selectielijst naar het topologie-ontwerpgebied slepen om te beginnen met het maken van uw gevirtualiseerde infrastructuur.



## Over deze vertaling

Cisco heeft dit document vertaald via een combinatie van machine- en menselijke technologie om onze gebruikers wereldwijd ondersteuningscontent te bieden in hun eigen taal. Houd er rekening mee dat zelfs de beste machinevertaling niet net zo nauwkeurig is als die van een professionele vertaler. Cisco Systems, Inc. is niet aansprakelijk voor de nauwkeurigheid van deze vertalingen en raadt aan altijd het oorspronkelijke Engelstalige document ([link](#)) te raadplegen.