ETSI GS NFV-INF 005 V1.1.1 (2014-12)

Network Functions Virtualisation (NFV);
Infrastructure;
Network Domain

Disclaimer

This document has been produced and approved by the Network Functions Virtualisation (NFV) ETSI Industry Specification Group (ISG) and represents the views of those members who participated in this ISG.
It does not necessarily represent the views of the entire ETSI membership.

Reference
DGS/NFV-INF005

Keywords
network, NFV

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - NAF 742 C
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° 7803/88

Contents

Intellectual Property Rights

IPRs essential or potentially essential to the present document may have been declared to ETSI. The information pertaining to these essential IPRs, if any, is publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available from the ETSI Secretariat. Latest updates are available on the ETSI Web server (http://ipr.etsi.org).

Pursuant to the ETSI IPR Policy, no investigation, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Foreword

This Group Specification (GS) has been produced by ETSI Industry Specification Group (ISG) Network Functions Virtualisation (NFV).

Infrastructure Architecture Document		Document #
Overview		GS NFV INF 001
Illustrative Use Cases for the NFV Infrastructure		GS NFV INF 002
Architecture of the Infrastructure Domains	Compute Domain	GS NFV INF 003
	Hypervisor Domain	GS NFV INF 004
	Infrastructure Network Domain	GS NFV INF 005
Architectural Methodology	Interfaces and Abstraction	GS NFV INF 007
Service Quality Metrics		GS NFV INF 010

Modal verbs terminology

In the present document "**shall**", "**shall not**", "**should**", "**should not**", "**may**", "**may not**", "**need**", "**need not**", "**will**", "**will not**", "**can**" and "**cannot**" are to be interpreted as described in clause 3.2 of the ETSI Drafting Rules (Verbal forms for the expression of provisions).

"**must**" and "**must not**" are **NOT** allowed in ETSI deliverables except when used in direct citation.

1 Scope

The present document presents an architectural description of the Infrastructure Network domain of the infrastructure which supports virtualised network functions. It sets out the scope of the infrastructure domain acknowledging the potential for overlap between infrastructure domains, and between the infrastructure and the virtualised network functions. Its also sets out the nature of interfaces needed between infrastructure domains and within the infrastructure network domain.

The present document does not provide any detailed specification but makes reference to specifications developed by other bodies and to potential specifications, which, in the opinion of the NFV ISG could be usefully developed by an appropriate standards developing organisation (SDO).

2 References

2.1 Normative references

References are either specific (identified by date of publication and/or edition number or version number) or non-specific. For specific references, only the cited version applies. For non-specific references, the latest version of the referenced document (including any amendments) applies.

Referenced documents which are not found to be publicly available in the expected location might be found at http://docbox.etsi.org/Reference.

NOTE: While any hyperlinks included in this clause were valid at the time of publication, ETSI cannot guarantee their long term validity.

The following referenced documents are necessary for the application of the present document.

[1] ETSI GS NFV 003 (V1.1.1): "Network Functions Virtualisation (NFV); Terminology for Main Concepts in NFV".

[2] ETSI GS NFV 002 (V1.1.1): "Network Functions Virtualisation (NFV); Architectural Framework".

[3] ETSI GS NFV 001 (V1.1.1): "Network Functions Virtualisation (NFV); Use Cases".

[4] ETSI GS NFV-MAN 001 (V1.1.1): "Network Functions Virtualisation (NFV); Management and Orchestration".

2.2 Informative references

References are either specific (identified by date of publication and/or edition number or version number) or non-specific. For specific references, only the cited version applies. For non-specific references, the latest version of the referenced document (including any amendments) applies.

NOTE: While any hyperlinks included in this clause were valid at the time of publication, ETSI cannot guarantee their long term validity.

The following referenced documents are not necessary for the application of the present document but they assist the user with regard to a particular subject area.

[i.1] ETSI GS NFV-INF 001 (V1.1.1): "Network Functions Virtualisation (NFV); Infrastructure Overview".

[i.2] ETSI GS NFV-INF 003 (V1.1.1): "Network Functions Virtualisation (NFV); Infrastructure; Compute Domain".

[i.3] ETSI GS NFV-INF 004 (V1.1.1): "Network Functions Virtualisation (NFV); Infrastructure; Hypervisor Domain".

[i.4] IEEE Std 802.1Q™ (2012): "IEEE Standard for Local and metropolitan area networks -- Media Access Control (MAC) Bridges and Virtual Bridges".

[i.5] MEF 6.1 (2008-04): "MEF Technical Specification; MEF 6.1; Ethernet Services Definitions - Phase 2".

[i.6] draft-davie-stt-04 (work in progress): "A Stateless Transport Tunneling Protocol for Network Virtualization (STT)".

[i.7] draft-mahalingam-dutt-dcops-vxlan-06 (work in progress and experimental): "VXLAN: A Framework for Overlaying Virtualized Layer 2 Networks over Layer 3 Networks".

[i.8] draft-sridharan-virtualization-nvgre-03 (work in progress): "NVGRE: Network Virtualization using Generic Routing Encapsulation".

[i.9] IETF RFC 2784 (2000-03): "Generic Routing Encapsulation (GRE)".

[i.10] IETF RFC 1702 (1994-10): "Generic Routing Encapsulation over IPv4 networks".

[i.11] IETF RFC 3985 (2005-03): "Pseudo Wire Emulation Edge-to-Edge (PWE3) Architecture".

[i.12] IETF RFC 4448 (2006-04): "Encapsulation Methods for Transport of Ethernet over MPLS Networks".

[i.13] IETF RFC 4761 (2007-01): "Virtual Private LAN Service (VPLS) Using BGP for Auto-Discovery and Signaling".

[i.14] IETF RFC 4762 (2007-01): "Virtual Private LAN Service (VPLS) Using Label Distribution Protocol (LDP) Signaling".

[i.15] IEEE Std 802.3™ (2012): "Ethernet working group".

[i.16] IETF RFC 4364 (2006-02): "BGP/MPLS IP Virtual Private Networks (VPNs)".

[i.17] IETF RFC 2661 (1999-08): "Layer Two Tunneling Protocol "L2TP"".

[i.18] IETF RFC 6439 (2011-11): "Routing Bridges (RBridges): Appointed Forwarders".

[i.19] IEEE Std 802.1Qbp™ (2013): "IEEE Standard for Local and metropolitan area networks - Media Access Control (MAC) Bridges and Virtual Bridged Local Area Networks - Amendment: Equal Cost Multiple Paths (ECMP)".

[i.20] IEEE Std 802.1AX™ (2014): "IEEE Standard for Local and metropolitan area networks -- Link Aggregation".

[i.21] IETF RFC 6325 (2011-07): "Routing Bridges (RBridges): Base Protocol Specification".

[i.22] IETF RFC 6327 (2011-07): "Routing Bridges (RBridges): Adjacency".

[i.23] IEEE Std 802.1ag™ (2007): "IEEE Standard for Local and Metropolitan Area Networks Virtual Bridged Local Area Networks Amendment 5: Connectivity Fault Management".

[i.24] IEEE Std 802.1AB™ (2009): "IEEE Standard for Local and Metropolitan Area Networks -- Station and Media Access Control Connectivity Discovery".

[i.25] IEEE Std 802.1Qbg™ (2012): "IEEE Standard for Local and metropolitan area networks -- Media Access Control (MAC) Bridges and Virtual Bridged Local Area Networks -- Amendment 21: Edge Virtual Bridging".

[i.26] IEEE Std 802.1Qbb™ (2011): "IEEE Standard for Local and metropolitan area networks -- Media Access Control (MAC) Bridges and Virtual Bridged Local Area Networks -- Amendment 17: Priority-based Flow Control".

[i.27] IEEE Std 802.1Qaz™ (2011): "IEEE Standard for Local and metropolitan area networks -- Media Access Control (MAC) Bridges and Virtual Bridged Local Area Networks -- Amendment 18: Enhanced Transmission Selection for Bandwidth Sharing Between Traffic Classes".

[i.28] IEEE Std 802.1AX™ (2008): "IEEE Standard for Local and metropolitan area networks -- Link Aggregation".

[i.29] IEEE Std 802.1AS™ (2011): "IEEE Standard for Local and Metropolitan Area Networks - Timing and Synchronization for Time-Sensitive Applications in Bridged Local Area Networks".

[i.30] IEEE Std 802.1Qau™ (2010): "IEEE Standard for Local and Metropolitan Area Networks -- Virtual Bridged Local Area Networks -- Amendment 13: Congestion Notification".

[i.31] IETF STD 62 (2002): "STD 62 (RFC 3417) Transport Mappings for the Simple Network Management Protocol (SNMP)"; "STD 62 (RFC 3416) Version 2 of the Protocol Operations for the Simple Network Management Protocol (SNMP)"; "STD 62 (RFC 3415) View-based Access Control Model (VACM) for the Simple Network Management Protocol (SNMP)"; "STD 62 (RFC 3414) User-based Security Model (USM) for version 3 of the Simple Network Management Protocol (SNMPv3)"; "STD 62 (RFC 3413) Simple Network Management Protocol (SNMP) Applications"; "STD 62 (RFC 3412) Message Processing and Dispatching for the Simple Network Management Protocol (SNMP)"; "STD 62 (RFC3411) An Architecture for Describing Simple Network Management Protocol (SNMP) Management Frameworks".

[i.32] IETF RFC 6241 (2011-06): "Network Configuration Protocol (NETCONF)".

[i.33] IETF RFC 3954 (2004-10): "Cisco Systems NetFlow Services Export Version 9".

[i.34] MEF 17 (2007-04): "MEF Technical Specification; MEF 17; Service OAM Requirements & Framework - Phase 1".

[i.35] MEF 30.1 (2013-04): "MEF Technical Specification; MEF 30.1; Service OAM Fault Management Implementation Agreement: Phase 2".

[i.36] MEF.35 (2012-04): "MEF Technical Specification; MEF 35; Service OAM Performance Monitoring Implementation Agreement".

[i.37] IEEE Std 802.1BR™ (2012): "IEEE Standard for Local and metropolitan area networks -- Virtual Bridged Local Area Networks--Bridge Port Extension".

[i.38] draft-ietf-nvo3-security-requirements-02: "Security Requirements of NVO3".

[i.39] IETF RFC 4031 (2005-04): "Service Requirements for Layer 3 Provider Provisioned Virtual Private Networks (PPVPNs)".

[i.40] IETF RFC 4110 (2005-07): "A Framework for Layer 3 Provider-Provisioned Virtual Private Networks (PPVPNs)".

[i.41] IETF RFC 4271 (2006-01): "A Border Gateway Protocol 4 (BGP-4)".

[i.42] IETF RFC 4760 (2007-01): "Multiprotocol Extensions for BGP-4".

[i.43] draft-ietf-l3vpn-end-system-02 (work in progress): "End-system support for BGP-signaled IP/VPNs".

[i.44] IETF RFC 4664 (2006-09): "Framework for Layer 2 Virtual Private Networks (L2VPNs)".

[i.45] IETF RFC 4665 (2006-09): "Service Requirements for Layer 2 Provider-Provisioned Virtual Private Networks".

[i.46] draft-ietf-l2vpn-evpn-req-06 (work in progress): "Requirements for Ethernet VPN (EVPN)".

[i.47] draft-ietf-opsawg-oam-overview-16 (work in progress): "An Overview of Operations, Administration, and Maintenance (OAM) Tools".

[i.48] ETSI GS NFV-SWA 001 (V1.1.1): "Network Functions Virtualisation (NFV); Virtual Network Function Architecture".

[i.49] IETF RFC 4655 (2006-08): "A Path Computation Element (PCE)-Based Architecture".

[i.50] ETSI GS NFV-PER 001 (V1.1.1): "Network Functions Virtualisation (NFV); NFV Performance & Portability Best Practises".

[i.51] ETSI GS NFV-REL 001 (V1.1.1): "Network Functions Virtualisation (NFV); Resiliency Requirements".

[i.52] IETF RFC 792 (1981-09): "Internet Control Message Protocol".

[i.53] IETF RFC 4443 (2006-03): "Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification".

[i.54] IETF RFC 2151 (1997-06): "A Primer On Internet and TCP/IP Tools and Utilities".

[i.55] IETF RFC 5880 (2010-06): "Bidirectional Forwarding Detection (BFD)".

[i.56] IETF RFC 5881 (2010-06): "Bidirectional Forwarding Detection (BFD)for IPv4 and IPv6 (Single Hop)".

3 Definitions and abbreviations

3.1 Definitions

For the purposes of the present document, the following terms and definitions apply:

policy group: grouping of nodes (VNFCI(s), external entities, infrastructure components, etc.) in an NFV environment that share a common policy

NOTE: That policy is usually, but not limited to, a security or traffic isolation model. Other possible uses of a policy group could include common traffic forwarding class, policy based routing, etc.

security group: security group is a subset of Policy Groups that are only concerned with traffic isolation

NOTE: An example of a traffic isolation policy group might be that all the VNFCI(s) deployed to provide a load-balancing function as part of some service function can receive TCP traffic from any external source addressed to port 80 or port 443, and can communicate with other VNFCI(s) deployed as part of the same service using TCP addressed to port 80 or port 443, and ICMP PING protocols.

virtual network: See ETSI GS NFV-INF 001 [i.1].

3.2 Abbreviations

For the purposes of the present document, the following abbreviations apply:

AMI	Advanced Metering Infrastructure
API	Application Programming Interface
ARP/ND	Address Resolution Protocol/ Neighbor Discovery
BEB	Backbone Edge Bridge
BFD	Bidirectional Forwarding Detection
BGP	Border Gateway Protocol
BSS	Business Support System
CD	Compute Domain
CFM	Connectivity Fault Management
CIR	Committed Information Rate
CPU	Central Processing Unit
C-VID	Customer VLAN Identifier
DHCP	Dynamic Host Configuration Protocol
D-LAG	Distributed Link Aggregation
E-BGP	External Border Gateway Protocol
ECMP	Equal-Cost Multi-Path
EIR	Excess Information Rate

EVPN	Ethernet Virtual Private Network
FIB	Forwarding Information Base
ForCES	Forwarding and Control Element Separation
GRE	Generic Routing Encapsulation
HD	Hypervisor Domain
HTTP	Hypertext Transfer Protocol
HW	Hardware
I-BGP	Internal Border Gateway Protocol
ICMP	Internet Control Message Protocol
ID	Identifier
IETF	Internet Engineering Task Force (http://www.ietf.org/)
IG	Interworking Gateway
IND	Infrastructure Network Domain
IP	Internet Protocol
IPFIX	Internet Protocol Flow Information Export
IS-IS	Intermediate System to Intermediate System
LAG	Link Aggregation Group
LAN	Local Area Network
LDP	Label Distribution Protocol
LLDP	Link Layer Discovery Protocol
MAC	Media Access Control
MANO	Management and Orchestration
MEF	Metro Ethernet Forum (http://metroethernetforum.org/)
MEP	Maintenance association End Point
MIB	Management Information Base
MIP	Maintenance domain Intermediate Point
MP-BGP	Multiprotocol Border Gateway Protocol
MPLS	Multi-Protocol Label Switching
MSTP	Multiple Spanning Tree Protocol [i.4]
NAT	Network Address Translation
NF	Network Function [1]
NFCI	Network Function Component Instance
NFV	Network Functions Virtualisation
NFVI	Network Functions Virtualisation Infrastructure [1]
NFVI-PoP	Network Functions Virtualisation Infrastructure Point of Presence [i.1]
NFVO	Network Functions Virtualisation Orchestrator
NI	Network Intensive
NIC	Network Interface Card
N-PoP	Network Point of Presence [1]
NVE	Network Virtualisation Edge
NVGRE	Network Virtualisation using Generic Routing Encpasulation
OA&M	Operations, Administration and Maintenance
OAM	Operations, Administration and Maintenance
ONF	Open Networking Foundation (https://www.opennetworking.org/)
OS	Operating System
OSPF	Open Shortest Path First
OSS	Operations Support System
OTN	Optical Transport Network
PBB	Provider Backbone Bridge
PBB-TE	Provider Backbone Bridge Traffic Engineering
PCE	Path Computation Element
PE	Provider Edge
PNF	Physical Network Function [1]
PPP	Point-to-Point Protocol
PPTP	Point-to-Point Tunneling Protocol
QoS	Quality of Service
RFC	Request for Comments
ROCE	Remote Direct Memory Access (RDMA) over Converged Ethernet
RSTP	Rapid Spanning Tree Protocol
RSVP	Resource Reservation Protocol
SDH	Synchronous Digital Hierarchy
SDL	Software Development Lifecycle

SDN	Software-Defined Networking
SDO	Standards Development Organization
SID	Service Instance Identifier
SLA	Service Level Agreement [1]
SNMP	Simple Network Management Protocol
SPB	Shortest Path Bridging
SPBM	SPB-MAC
SPBV	SPB-VID
STP	Spanning Tree Protocol
STT	Stateless Transport Tunneling
S-VID	Service VLAN Identifer
TCP	Transmission Control Protocol
TE	Traffic Engineering
TEP	Tunnel End Point
TOR	Top Of Rack
TORS	Top-Of-Rack Switch
TRILL	Transparent Interconnection of Lots of Links (http://datatracker.ietf.org/wg/trill/)
UDP	Stateless Transport Tunneling
UNI	User Network Interface
VDP	Virtual Station Interface (VSI) Discovery and Configuration Protocol
VEB	Virtual Ethernet Bridging
VEPA	Virtual Ethernet Port Aggregator
VID	VLAN Identifier
VIM	Virtualisation Infrastructure Manager
VLAN	Virtual LAN
VM	Virtual Machine [1]
VN	Virtual Network
VNF	Virtualised Network Function [1]
VNFC	Virtual Network Function Component [i.1]
VNFCI	Virtual Network Function Component Instance
VNI	VxLAN Network Identifier
VNIC	Virtual Network Interface Card
VNID	Virtual Network Interface Device
VNP	Virtual Network Protocol
VPLS	Virtual Private LAN Service
VPN	Virtual Private Network
VPRN	Virtual Private Routed Network
VRF	Virtual Routing and Forwarding
VSID	Virtual Subnet Identifier
VTN	Virtual Tenant Network
VXLAN	Virtual eXtensible LAN
WAN	Wide Area Network
WIM	WAN Infrastructure Manager
XML	Extensible Markup Language

4 Domain Overview

Figure 1 illustrates the four domains described in [i.1], their relationship with each other and their relationship to other domains outside the infrastructure. The figure also sets out the primary interfaces.

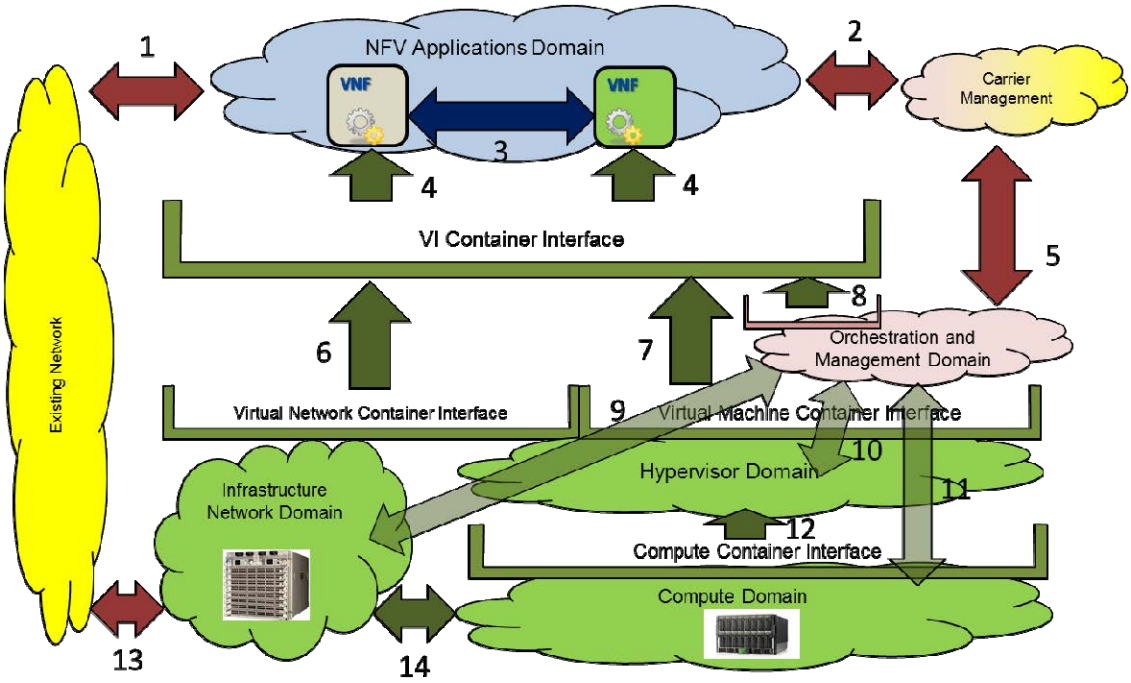

Figure 1: General Domain Architecture and Associated Interfaces

Figure 2 [i.1] gives a high level overview of the three domains within the NFVI and shows how the domains realise the primary interfaces of the NFV overall architectural framework.

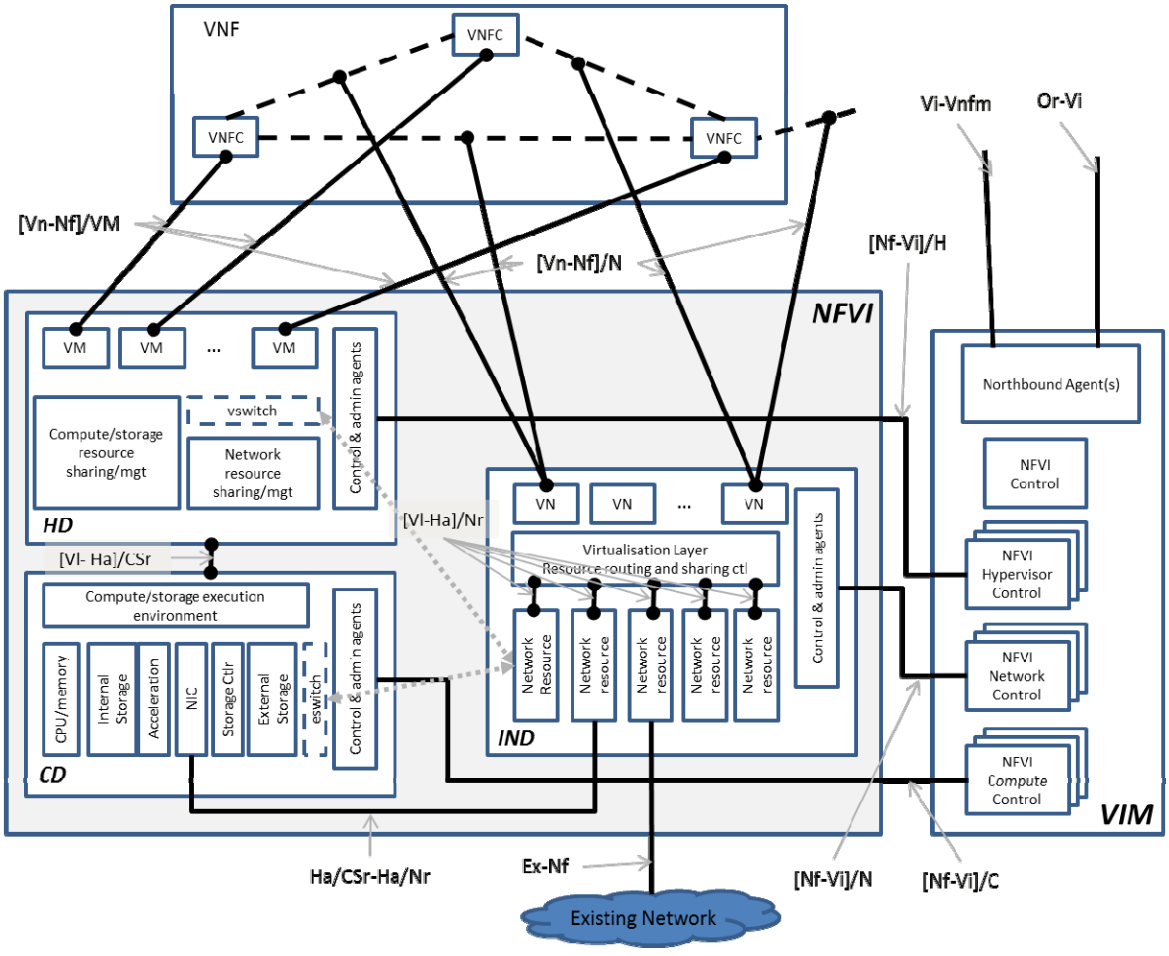

Figure 2: High Level Overview of the NFVI Domains and Interfaces

The general domain architecture of figure 2 is reduced to a reference point architecture (figure 3) showing only the Network Domain and aligning these reference points with the NFV E2E Architecture (ETSI GS NFV 002 [2]). The Network Domain reference architecture has five reference points catalogued in table 1.

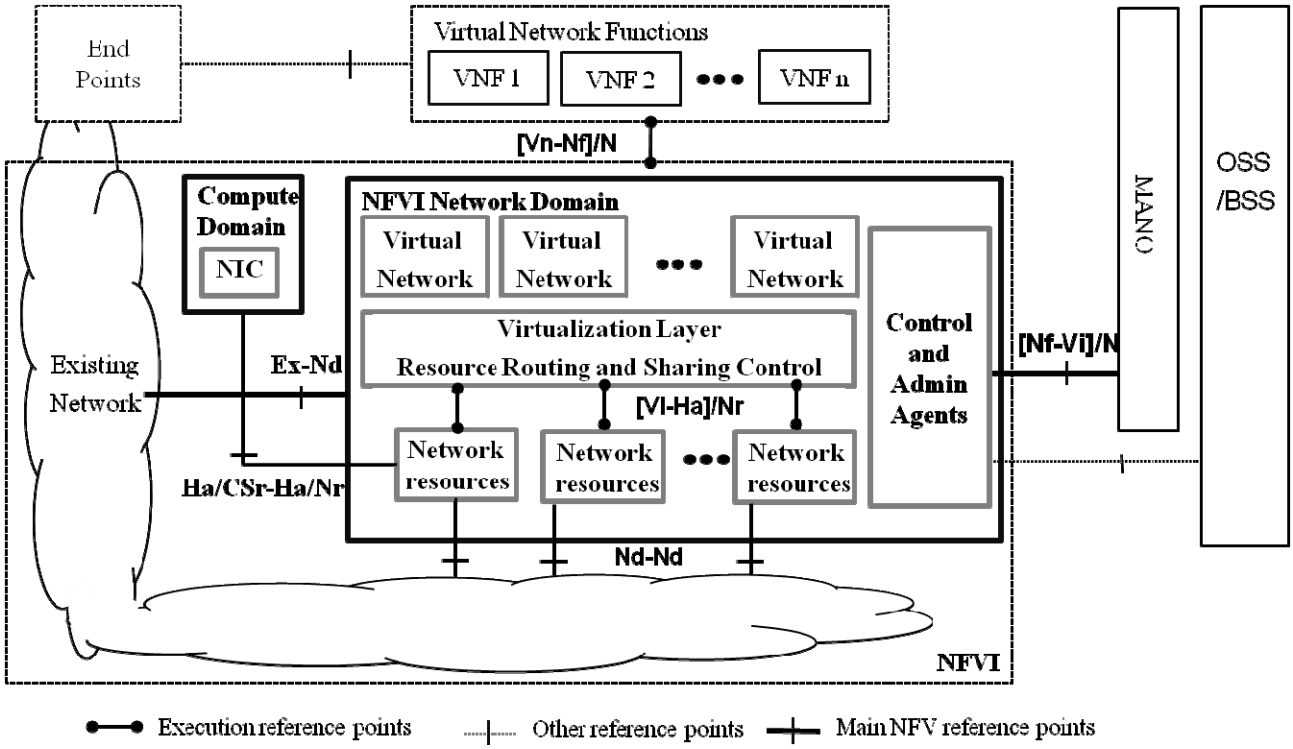

Figure 3: Network Domain Reference Point Architecture

The inter-domain and intra-domain interfaces are summarized in table 1. The inter-domain interfaces are described in more detail in clause 5. The functional blocks consist of the virtual networks, the virtualisation layer resource routing and sharing control, the network resources and the control & admin agents. They are described in clause 6. The interfaces internal to the domain are described in more detail in clause 7.

Table 1: Catalogue of Inter-domain Interfaces related to the Infrastructure Network Domain

Reference Point	Description
[Vn-Nf]/N	This reference point is the virtual network (VN) container interface carrying communication between VNFC instances. Note that a single VN can support communication between more than a single pairing of VNFC instances (eg an E-LAN VN). It is the reference point over which the services of the network domain are delivered. These services may be either IP forwarding services or Ethernet private line/LAN/TREE services provided by the infrastructure. The reference point is providing services at two layers: IP forwarding services across the [Vn-Nf]/N/L3 reference point and Ethernet services, e.g. E-LINE, E-LAN, E-TREE, across the [Vn-Nf]/N/L2 reference point.
[Nf-Vi]/N	This is the reference point between the management and orchestration agents in the infrastructure network domain and the management and orchestration functions in the virtual infrastructure management (VIM). It is the part of the Nf-Vi interface relevant to the infrastructure network domain.
[Vl-Ha]/Nr	The reference point between the virtualisation layer and the network resources.
Ex-Nd	The reference point between the infrastructure network domain and external networks.
Nd-Nd	The reference point between NFVI-PoPs used to extend the virtualisation layer of a single Network Operator's NFVI over multiple geographically separated sites.
Ha/Csr-Ha/Nr	This is the reference point between the infrastructure network domain and the servers/storage of the compute domain.

Table 2 describes various aspects of the infrastructure network and related requirements that it needs to address.

Table 2: Infrastructure Network Requirements

Generic Aspects	Requirements
Address Space and Traffic Isolation	For layer 2 services, the infrastructure network shall provide traffic and address space isolation (see note) between virtual networks. It shall support a large number of virtual networks (some VNFs may require their own virtual network(s)). For layer 3 services, the infrastructure network shall provide traffic isolation between virtual networks. Some use cases require address isolation, and if this requirement maps to isolation of the infrastructure IP address space, then the infrastructure shall support address space isolation. This may be achieved using various techniques: • An encapsulation method to provide overlay networks (L2 or L3 service). • The use of forwarding table partitioning mechanisms (L2 service). • By applying policy control within the infrastructure network (L3 service). The technique shall provide sufficient information for unambiguous mapping of given packet to its associated virtual network. The technique shall support IP traffic and may support multiple L3 protocols. The technique can be applied by the server (vSwitch or vRouter) or external switch/router (tier 1, tier 2 or tier 3). Where encapsulation is employed and there is a requirement for separate service and infrastructure addresses, there shall be a mechanism to resolve service addresses to infrastructure addresses.
Address management	The virtual network(s) shall ensure address uniqueness within a given virtual network. The solution shall be able to translate between overlapping address spaces and/or public/private addresses.
Scalability	The infrastructure network shall be able to support a large number of servers each running many VMs. The infrastructure network may span multiple N-PoPs.
Flexibility	The infrastructure network shall be able to support live VM migration within a data center. It should also aim to support live VM migration between data centers.
Reliability and Availability	The infrastructure network should provide the ability to request different service levels with measurable reliability and availability metrics, e.g. percentage of time the network is available. The infrastructure network shall provide a set of OAM processes to verify reliability, availability and integrity of the infrastructure network layer. The infrastructure network should provide mechanisms to mitigate congestive loss of data frames for applications that require it, for example: FCoE, ROCE, etc.
Network Utilization	The infrastructure network should be able to utilize breadth of connectivity where it exists to maximize network utilization. It should also support multicast/broadcast VNF traffic efficiently . Tradeoffs between performance, network resource utilization and other resources utilization is expected. For example, shutting down network resources to reduce power utilization may increase latency between VNFs.
Performance	Requirements vary greatly based on the network functions performance requirements. The network should allow the infrastructure connectivity services to specify the following performance related parameters: • Maximum overhead (bits required for the network virtualisation technique, per packet or percentage of traffic). • Maximum delay. • Maximum delay variation. • Throughput (CIR, CIR+EIR and packets per second). • Maximum packet loss allowable.
Security	The infrastructure network for NFV should consider both internal and external threats that could compromise the security of the infrastructure. Internal threats are usually from authorized internal personnel who may misbehave to cause damage to the infrastructure. Internal threats should be addressed by rigorous operational procedures. External threats are from outsiders who may gain access to the infrastructure, e.g. by exploiting design and/or implementation vulnerabilities. Once gaining access, an adversary could further escalate its privileges and install backdoor software to maintain long-term control. To address external threats, security should be considered during the whole development process, e.g. by following a Software Development Lifecycle (SDL) process. In addition, infrastructure devices might need to go through a security certification process (e.g. Common Criteria) to gain assurance of their security levels.
NOTE: Traffic isolation refers to privacy and non-leakage. QoS is for further study.	

5 External Interfaces of the Domain

5.1 [Vn-Nf]/N

5.1.1 Nature of the Interface

The Vn-Nf/N interfaces shall provide transparent network services to VNFs. This interface is used to interconnect the following:

- VNFCIs to other VNFCIs within within the same or another VNF.
- VNFCIs to storage.
- VNFs to PNFs and external endpoints.

VNFs use the network services provided by the Vn-Nf/N interface to participate in the VNF Forwarding Graphs as defined in ETSI GS NFV 002 [2]. The Vn-Nf/N interface provides the container interface that provides access to the network services.

Transparent infrastructure network connectivity services present VNFs with a virtual network exhibiting the same properties as a physical network implementing the same connectivity service. The services are transparent in the sense that VNFs are unaware of how the services are provided.

The virtual network(s) may form a discrete subset of the NFVI, or it may be comprised of one or more overlay networks that exist on top of the actual NFV infrastructure.

The VNFCI to VNFCI and VNFCI to storage connectivity services may be provided using a virtual network that implements one of the services described in the following sub-clauses. A VNF may require one or more virtual networks to interconnect the various VNFCIs.

NOTE: The provision of connectivity between a VNF and a PNF or an endpoint requires the use of the Ex-Nd interface described in clause 5.3.1. The provision of connectivity between VNFs located in different NFVI-PoPs requires the use of the Nd-Nd interface described in clause 5.3.2. This is not visible to the VNFs.

The Vn-Nf/N Virtual Network Container interface can take several forms depending on the VNF connectivity requirements and whether the infrastructure network is offered as a service to a provider of the network service.

In the case where each VNFCI port, as defined in ETSI GS NFV-SWA 001 [i.48], has a one-to-one correspondence with a VNIC, as is illustrated in figure 4 it is more difficult for a provider of the network service to make use of resources owned by a distinct NFV Infrastructure provider because this method requires either of the following approach, neither of which is desirable:

- The provider of the network service has to control the virtual networks (e.g. VLAN, VXLAN) to connect two VNICs, while the virtual networks are owned by the infrastructure provider.
- To implement their own virtual networks within the VNFCIs.

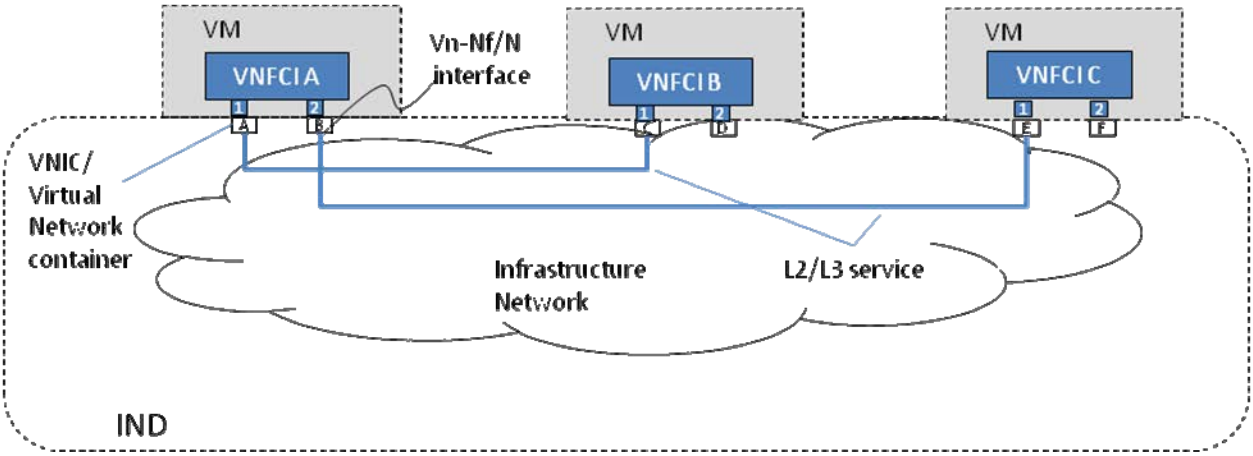

Figure 4: Vn-Nf/N per VNFCI port

In order to support all use cases, the following requirements should be addressed:

- It shall be possible to connect multiple ports within a single VNFCI to a single VNIC.
- The VNF shall remain unaware of the how the infrastructure network enforces the VNF Forwarding Graph and network connectivity between VNFs (and PNFs) of a Network Service, and between the VNFC instances of a VNF.
- It shall be possible to delegate VNF Forwarding Graph decisions to the provider of the network service. This addresses the NFVIaaS use case in ETSI GS NFV 001 [3].
- It shall be possible to determine to which VNFCI port(s) to deliver each packet received on a VNIC.

5.1.1.1 [Vn-Nf]/N/L2 Service

For an L2 service, the VNFCI is free to encapsulate any network layer protocol that is supported over Ethernet. The service will provide connectivity of the Ethernet data frames. Two implementation variants are possible:

1) Ethernet services based on an Ethernet switching network infrastructure; or

2) Layer 2 VPN services based on an IP network.

Both implementation variants are presented to the VNFCI as a capability to exchange Ethernet data frames.

5.1.1.1.1 [Vn-Nf]/N/L2 Service Definition

Table 3

Service Specification		Illustrative Parameters
Control Operations	Establishment	Request • list of vNICs • (pairwise) bandwidth requirement • (pairwise) delay requirement • resiliency requirement Return • instance id • list of vNICs • (pairwise) bandwidth • (pairwise) delay • Resiliency
	Modification	Request • instance id • change to list of vNICs • change to (pairwise) bandwidth requirement • change to (pairwise) delay requirement • change to resiliency requirement Return • instance id • list of vNICs • (pairwise) bandwidth • (pairwise) delay • Resiliency
	Removal	Request • instance id Return • success
Instance Interfaces	vNIC end point	VLANiD/MAC address Physical location
Operational status		OAM parameters
Performance stats	Establishment	Virtual network provisioning latency Virtual network diversity compliance Virtual network provisioning reliability
	Operation	Packet loss Packet delay Packet delay variation Delivered throughput Network outage
Instance Functionality	Forwarding and transport	MAC forwarding

5.1.1.1.2 [Vn-Nf]/N/L2 VPN Service

The Layer 2 VPN solution based on BGP Ethernet VPNs (EVPN) [i.46] provides a virtual Layer 2 bridged connectivity between the VNFCIs. The infrastructure domain could be MPLS technology (providing the benefits of fast-reroute, resiliency, etc.) or IP technology. In an EVPN, the MAC learning between the VNFCI and the domain edge occurs the same as for a Layer 2 bridge. As the VNFCI is unaware that the L2 service is provided using the Layer 2 VPN solution, there is no extra complexity for the VNFCI. Within the network domain, the MAC learning occurs not in the data plane (as with traditional Ethernet bridging) but in the control plane. This offers greater control than with traditional Ethernet bridging, such as restricting who learns what, and the ability to apply policies. The use of a BGP control plane provides greater network scalability than traditional Ethernet bridging and the ability to preserve the virtualisation or isolation of groups of interacting agents (hosts, servers, virtual machines) from each other.

The L2 VPN solution provides the service presented in clause 5.1.1.1.1, with a different set of parameters for the Instance Interfaces:

Instance Interfaces	vNIC end point	MAC address, VLAN ID, B-MAC

5.1.1.1.3 [Vn-Nf]/N/L2 OAM Protocols

The following protocols may be carried inside the Services layer:

- MEF 17 Service OAM [i.34].
- MEF 30.1 Service OAM Phase 2 [i.35].
- MEF 35 Service OAM Performance Monitoring Implementation Agreement [i.36].
- Connectivity Fault Management IEEE Std 802.1ag [i.23].
- Congestion Notification IEEE Std 802.1Qau [i.30].

NOTE: This set of protocols runs at the Virtual L2 Service layer. They are tunnelled over the NFVI core, just like the user data, however will not interact with the core management. It is preferred to never translate these.

5.1.1.2 [Vn-Nf]/N/L3 Service

If the VNFCI communicates using IP-only (together with supporting protocols such as ARP/ND and DHCP), either a Layer 2 or Layer 3 service may be used depending on the service granularity required. Both L2 and L3 services operate by exchanging Ethernet data frames with a VNFCI at a virtual NIC interface as the physical layer interface. For a Layer 3 service, an IP network infrastructure is used and supports IP packet service granularity. The L3 service options presented here provide a very scalable and flexible service solution for the infrastructure domain, especially suited for large scale deployments.

There are two distinct approaches to L3 services, with different characteristics as it relates to address space isolation. However, from the point of view of a VNFCI, there is no detectable difference between these two L3 services. Both services are presented to the VNFCI as a capability to exchange either IPv4 or IPv6 data packets.

The service definition for the L3 service is identical to that presented in clause 5.1.1.1.1, apart from the following differences:

Instance Interfaces	vNIC end point	IP address, MAC address

Instance Functionality	Forwarding and transport	IP forwarding of IPv4 or IPv6 data packets only

5.1.1.2.1 [Vn-Nf]/N/L3 VPN Service

An L3 VPN service based on BGP IP VPN makes use of an overlay approach to provide an arbitrary number of logically separate virtual networks, each of which has its own independent range of IP addresses. The ranges of IP addresses used by different virtual networks may overlap, i.e. the L3 VPN service supports address space isolation. Note, there is no overlay visible to the VNFCI, this service does not introduce any extra complexity for the VNFCI. [i.43] describes a network virtualisation solution that provides an IP service to end-system virtual interfaces. The solution decouples the control plane and the forwarding functionality to enable the forwarding functionality to be implemented flexibly in multiple devices.

5.1.1.2.2 [Vn-Nf]/N/L3 Infrastructure based virtual networks Service

The L3 infrastructure-based service does not make use of an overlay approach. Instead, IP packets are transported between VNFCIs using the L3 infrastructure network directly. VNFCIs may be assigned IP addresses out of one or more non-overlapping private IP address spaces associated with the NFV infrastructure, or they may be assigned public IP addresses, or they may be assigned some mix of both private and public IP addresses. The L3 infrastructure-based service does not support address space isolation. However it does support traffic isolation between defined security groups.

5.1.1.2.3 [Vn-Nf]/N/L3 OAM Protocols

The following protocols may be carried inside the services layer:

- IPv4/IPv6 Ping using ICMP ([i.52], [i.53]).
- IPv4/IPv6 TraceRoute (described in [i.54]).
- BFD for IPv4/IPv6 ([i.55], [i.56]).

These are a (non-exhaustive) list of IP OAM tools available. A more comprehensive (but not complete) list can be found in [i.47].

5.1.2 Specifications in Current Widespread Use

5.1.2.1 MEF Specifications for L2 services

MEF [i.5] defined Ethernet Services including E-Line, E-LAN and E-Tree. E-Line provides a point-to-point service, E-LAN provides a multipoint-to-multipoint service and E-TREE provides a rooted multipoint service between UNI endpoints. Similar services are provided by the infrastructure network between the VNFCIs, storage and other NFs (VNFs and PNFs). Figure 5 shows how the MEF services can be applied to VNFCIs within a VNF or between VNFs.

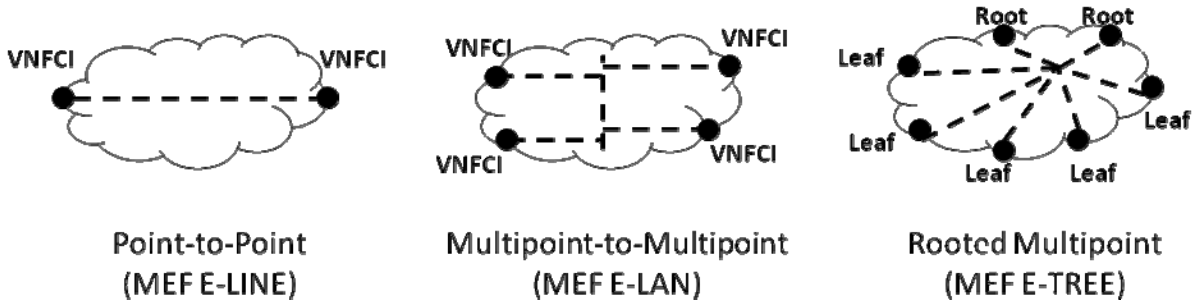

Figure 5: E-line, E-LAN and E-Tree Service

5.1.2.2 IETF Specifications for L3 services

IETF RFC 4031 [i.39] specifies requirements for Layer 3 Provider Provisioned VPN services. IETF RFC 4110 [i.40] specifies a framework for Provider Provisioned L3VPNs. IETF RFC 4271 [i.41] describes BGP and IETF RFC 4760 [i.42] defines multi-protocol extensions (MP-BGP). IETF RFC 4364 [i.16] uses these extensions to describes a method for using MP-BGP to provision L3VPNs over an IP backbone. Draft-ietf-l3vpn-end-system-02 [i.43] describes a solution in which the control plane in RFC 4364 is used to provide a Virtual Network Service between end systems.

IETF RFC 4665 [i.45] specifies requirements for L2VPNs and IETF RFC 4664 [i.44] provides a framework them.

5.2 [NF-Vi]/N

5.2.1 Nature of the Interface

The interface to the VIM shall provide the ability to request infrastructure connectivity services as described in clause 5.1. One of the complexities of this interface is that infrastructure connectivity services may require orchestration and management of both infrastructure network resources (e.g. Ethernet switches and/or routers), compute resources (e.g. NICs) and hypervisor resources (e.g. vSwitches or vRouters) in order to provide the requested infrastructure connectivity services to the VNF, the latter two being covered in other documents [i.2] and [i.3].

It is part of the role of the VIM to create compatible configurations across the domains in order to construct the infrastructure network services, further details of which are covered in [4] (clause 7.1.5). Moreover, the scope of a particular infrastructure network service is likely to go beyond that on any one VIM . The NFVO provides co-ordination of configuration across all the VIMs associated with an infrastructure network service and will also co-ordinate with existing infrastructure which are not managed by a VIM (these are normally managed by an existing OSS). This is illustrated in figure 5.4 of [4].

Figure 6 illustrates the scope of Interface 9 between the Management and Orchestration and the Infrastructure Network domain where the interface defined in this clause corresponds to the arrow between the NaaS Platform Plugins and the Infrastructure Network Management Plane.

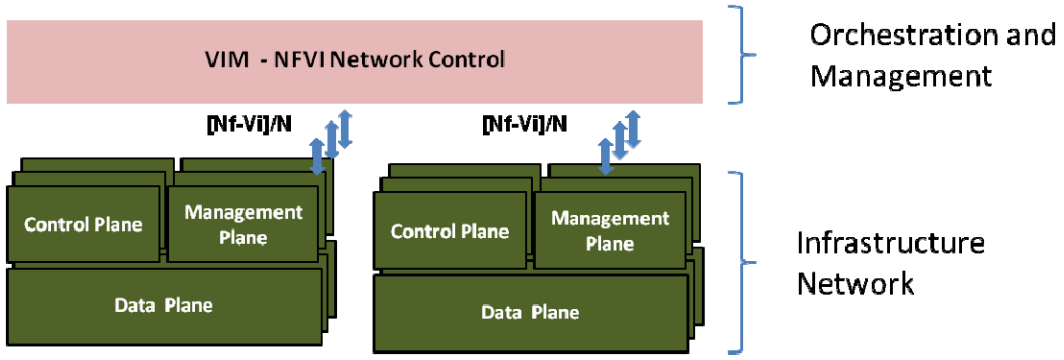

Figure 6: Orchestration and Management within the Infrastructure Network Domain

There are several approaches to infrastructure network management and control. In some approaches, the VIM interfaces a centralized network controller that controls the network equipment through a standard interface, for example an SDN controller using OpenFlow interface to control OpenFlow enabled switches. In other approaches, the VIM may be interfacing the equipment directly. In other approaches, a single VIM may interface a combination of network controllers and/or networking equipment.

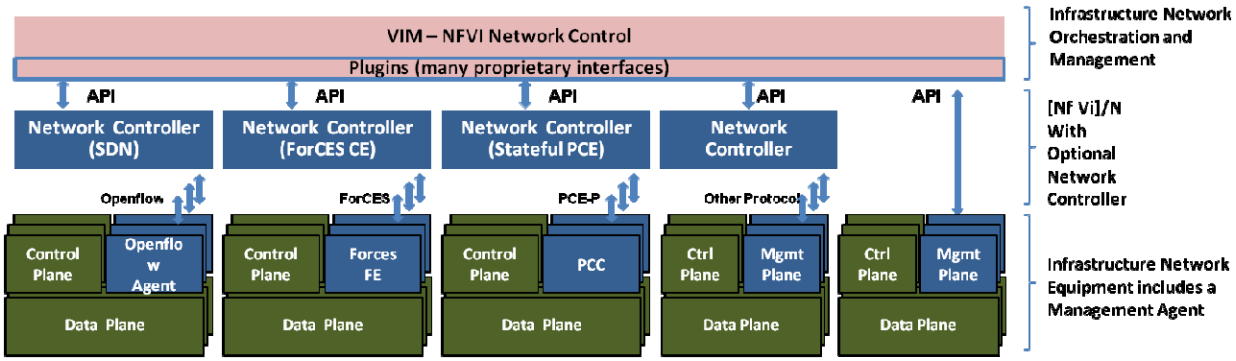

Figure 7: Infrastructure Network Management Plane

Figure 7 illustrates example management plane components. The Network Controller centralizes some or all of the control and management functionality and may provide an abstract view of its network domain to the NFVI Network Control functions of the VIM. Note that some of these components are centralizing functionality that was traditionally implemented as distributed control plane; the remaining distributed control plane aspects in those instances may be quite thin, or even disappear entirely. Interworking between different networking technologies requires co-existence of multiple types of network controllers and multiple technology support in the NFVI Network Control functions of the VIM.

5.3 Ex-Nf

5.3.1 Ex-Nd

The L2 service described in clause 5.1.1.1 and the L3 VPN service described in clause 5.1.1.2.1 both make use of overlay networking techniques. A gateway function may be required to provide an encapsulation/de-encapsulation function between the overlay and the external network domain.

The L3 infrastructure-based service described in clause 5.1.1.2.2 does not make use of overlay networking techniques. VNFCIs that are connected by an L3 infrastructure-based service can communicate directly with physical network functions without the need for a gateway function to perform encapsulation/decapsulation.

NOTE: A NAT function may be required between an L3 infrastructure-based network domain and a network domain to which physical network functions are connected if these respective domains have been assigned IP address spaces that require network address translation between them.

The following clauses address the gateway function needed to enable communication between physical network functions and VNFCIs that are connected by means of virtual network services based on overlays, i.e. the L2 network service or the L3 VPN network service, and do not apply to the L3 infrastructure-based service case.

5.3.1.1 Nature of the Interface

This interface allows VNFs to connect to PNFs, other VNFs and external endpoints. The VNF is not aware how this connectivity is provided, as shown in figure 8.

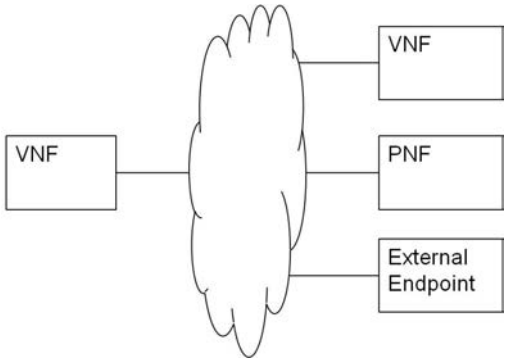

Figure 8: VNF connectivity

At the interface between the infrastructure network and the existing network, a mechanism shall exist to process the traffic received on the existing network (i.e. non-virtualised), map the traffic to the correct infrastructure connectivity service, potentially encapsulate the packet with an additional header corresponding to the identified infrastructure connectivity service and forward the packet to its destination. In the reverse direction, the reverse steps receive the packet for an infrastructure connectivity service (i.e. virtualised), decapsulate and process the packet and forward on the existing network.

In its simplest form, the incoming traffic may be classified based on the physical or virtual port on which it is received at the edge of the NFV infrastructure, i.e. the gateway device between the infrastructure network domain and the existing network. In more complicated forms, it may require packet inspection, for example to steer traffic based on VLAN or other packet characteristics, including potentially higher layer inspection.

EXAMPLE: In figure 9, Gateway B may be forwarding the traffic to the VNF B1, i.e. put it on the virtual network for VNF B1, if it received on interface B1 and forward traffic from VNF B1 on interface B1 while Gateway A may inspect the incoming packets on interfaces A1 and A2 and based on some protocol and field match, forward the packet to the corresponding VNF (A1, A2 or A3).

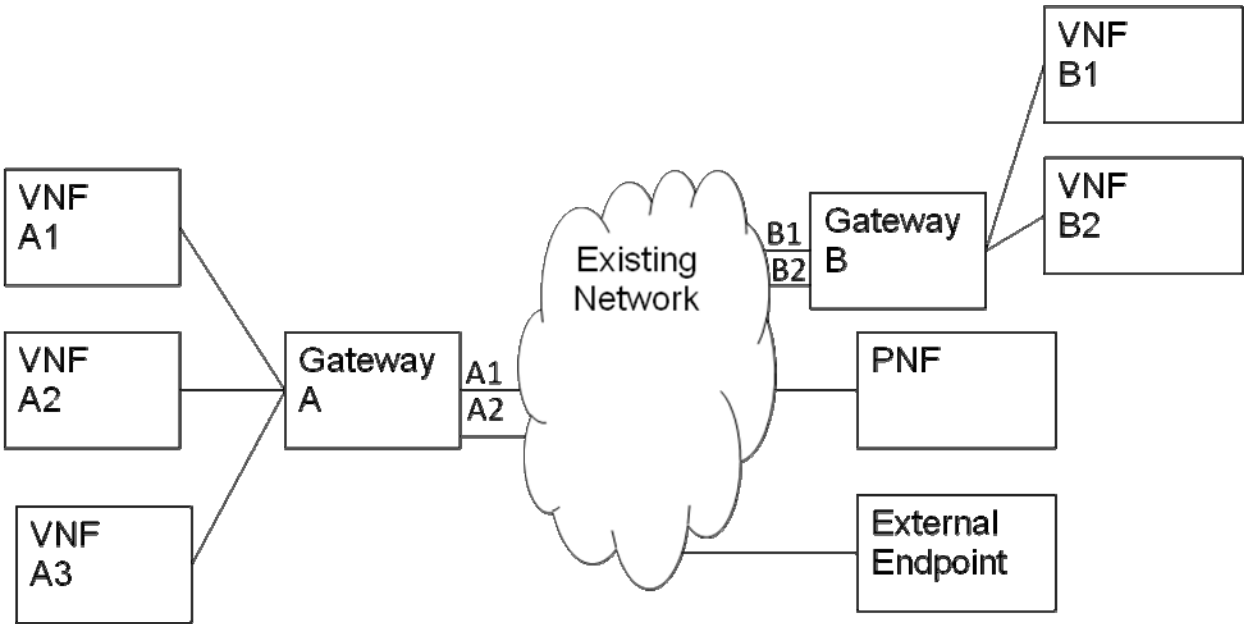

Figure 9: NFV Infrastructure Network Gateway Devices

The interface consists of the protocols that are exposed between the NFV infrastructure network gateway and the existing network and varies depending on the type of NFV infrastructure gateway device and its traffic steering capabilities. As the infrastructure network provides only L2 and L3 services, only steering based on L3 or below is configured through the Nf-Vi interface. Any steering based on L4-7 is configured outside the Nf-Vi interface.

As a VNF may consist of multiple VNFCIs but the VNF may not expose its internal design outside the NFVI, the externally visible interfaces of the VNF may not correspond exactly to the VNFCI interfaces. A number of options are provided by the infrastructure network building blocks of connectivity services and gateway functions. A few examples are provided below.

In its simplest case, the same connectivity service, e.g. E-LINE service, is created between the gateway and a single VNFCI of a VNF. That is reflected in figure 10. This could be the case if an application delivery controller is provided as a PNF but is not limited to that special case. Note that the VNF may include other VNFCIs but for external connectivity to other NFs, a single VNFCI is responsible for that interface. Note that this is merely an example and other similar examples include two VNFs with E-LINE service to the same gateway.

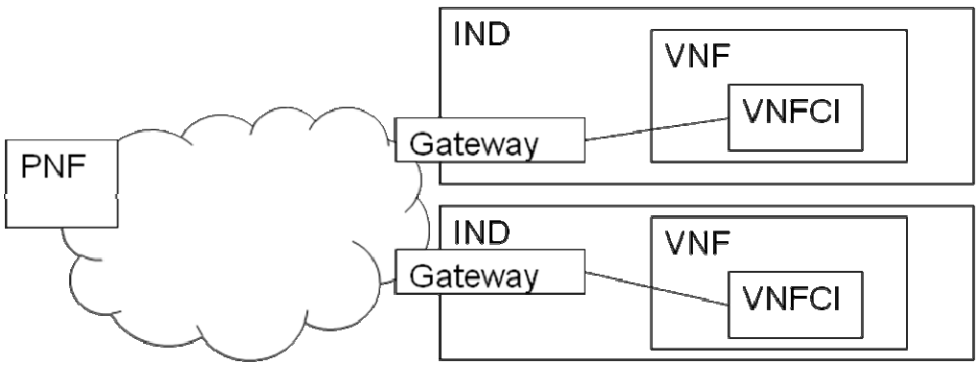

Figure 10: 1:1 Correspondence between VNF External Interfaces and VNFCI Interfaces

A more complex VNF could expose a single interface to other NFs but multiple VNFCI interfaces to the gateway, the gateway providing application delivery controller functions, for example load balancing. In this case, several E-LINE services may be requested; one between each VNFCI and the gateway. Other services that may be requested include E-LAN and E-TREE depending on the service characteristics needed by the VNF. In these cases, the gateway is responsible to steer the traffic based on some protocol fields as visible on the external network. This is illustrated in figures 11 and 12.

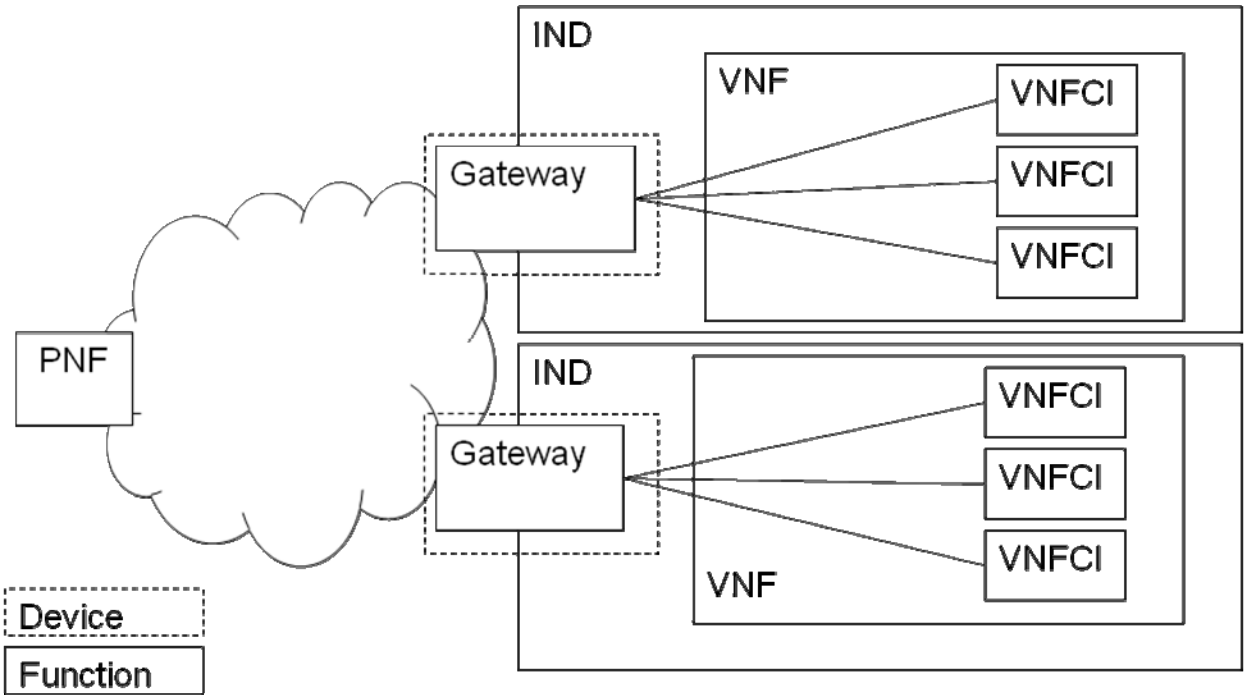

Figure 11: Ex-Nd with Native Gateway for 1:N Correspondence between VNF External Interfaces and VNFCI Interfaces

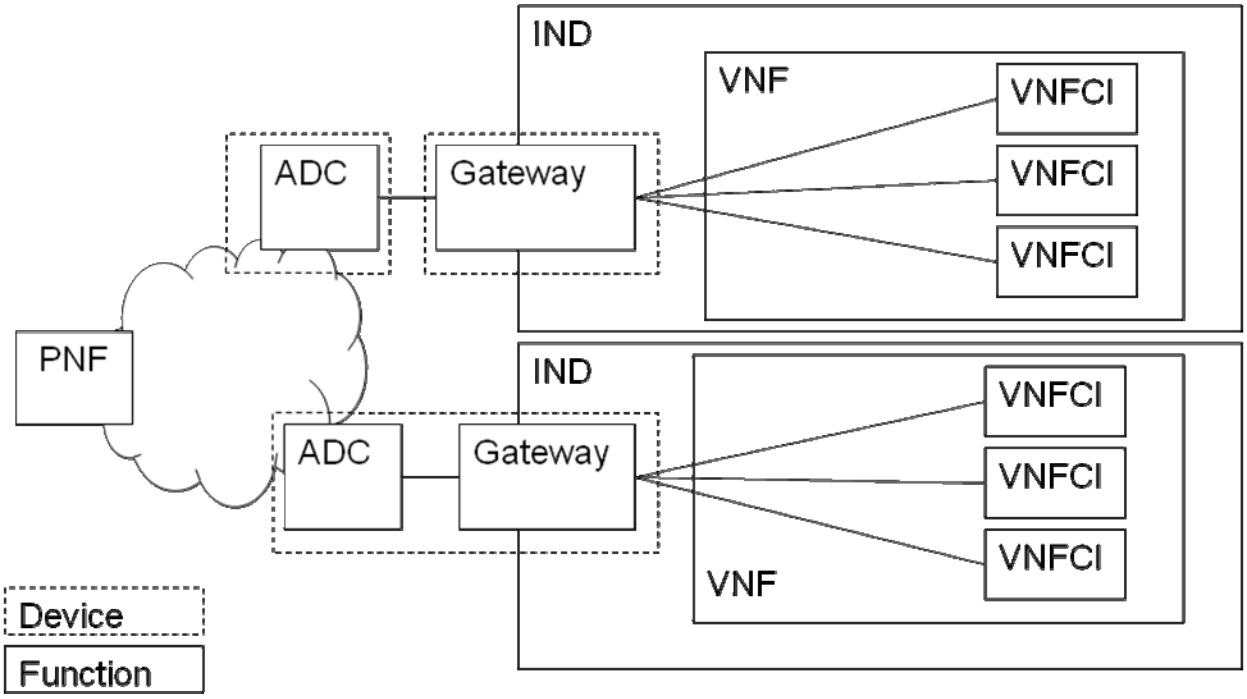

Figure 12: Ex-Nd with Extended Gateway for 1:N Correspondence between VNF External Interfaces and VNFCI Interfaces

A VNF may also expose more than one interface externally regardless of the number of VNFCIs that it contains. In this case, the gateway maps the traffic from VNF to external network based on the virtual network or source VNFCI. In this case, multiple E-LINE connectivity services between a VNFCI and gateway(s) may be used as illustrated in figure 13. Note that other options based on E-LAN and E-TREE services are also possible. In all cases, it is expected that the VNFCI would select the VNIC that corresponds to the externally visible interface, i.e. VNIC 1 for external interface 1. As the only virtual networks exposed to the VNFs are L2 and L3, there is no need to steer traffic towards the external network based on higher layers.

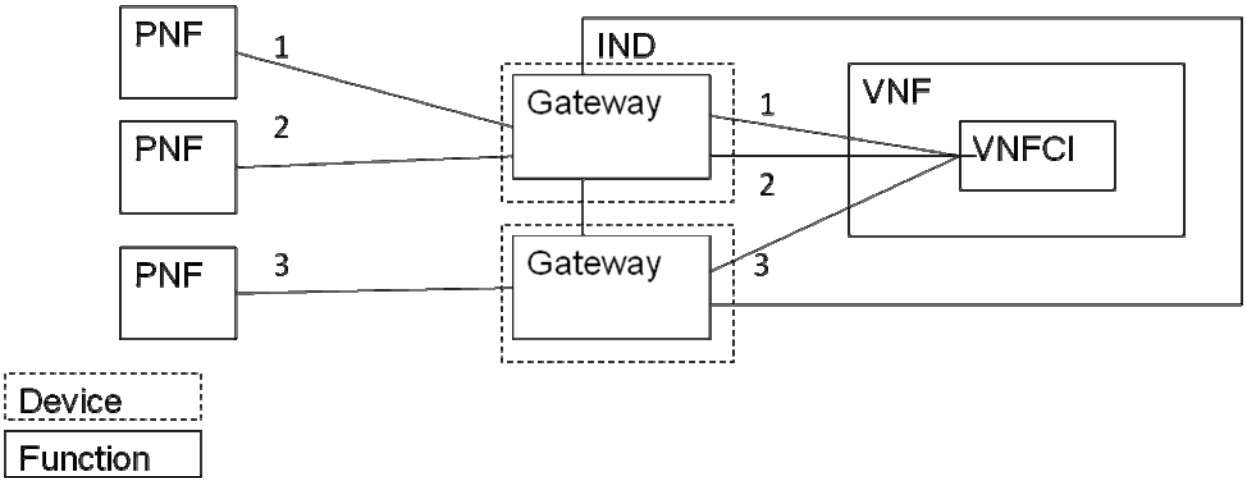

Figure 13: Ex-Nd Native Gateway for M:1 Correspondence between VNF External Interfaces and VNFCI Interfaces

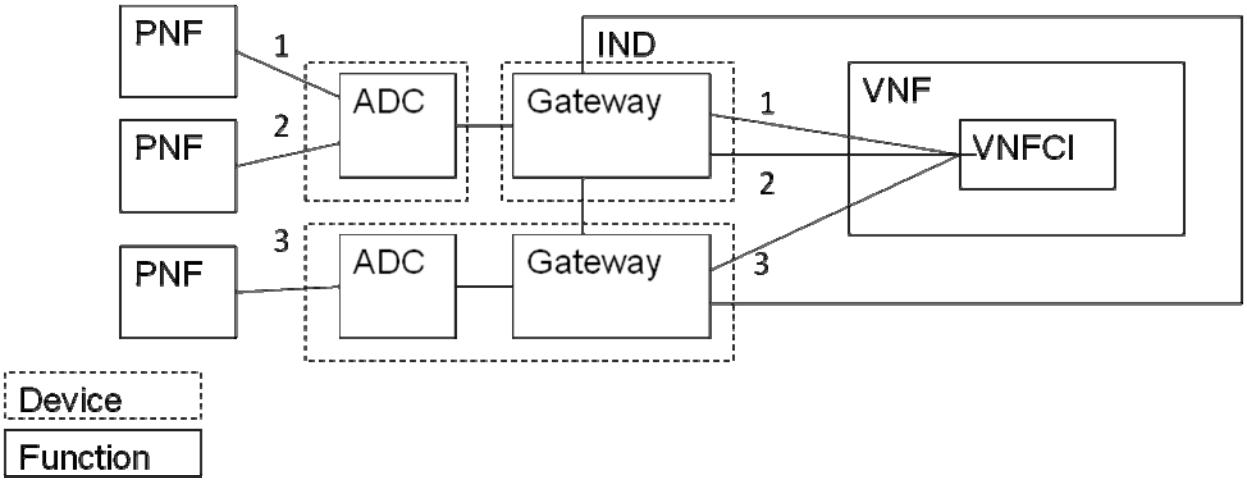

Figure 14: Ex-Nd Extended Gateway for M:1 Correspondence between VNF External Interfaces and VNFCI Interfaces

Note that if an application delivery controller offering load balancing functions and other traffic steering applications is included in the infrastructure, the functions that it provides that are based on layers 4-7 are considered outside the infrastructure network domain, i.e. they are not included in the virtual network service definition and the configuration is done outside the Nf-Vi reference point.

5.3.1.2 Specifications in Current Widespread Use

The use cases documented in ETSI GS NFV 002 [2] provide several contexts for VNFs to be deployed including both access and core network deployments. The dataplane throughput is a general concern for the NFV architecture (see e.g. ETSI GS NFV-PER 001 [i.50]). The throughput capacity of a virtualised network function deployment on an NFVI Node is a fundamental consideration as it is for network elements implementing physical network functions. In core network deployments the I/O of the NFVI node is typically symmetrical, but in access deployments there are typically asymmetries in both the number and line speeds of the interfaces facing the edge of the network compared to those facing the core.

The I/O asymmetry in access deployments is associated with the fan-in or aggregation of lower speed services. Consider the case of a simple VNF forwarding graph with one VNF implemented as a single VNFCI deployed at a NFVI Node for an access application. Figure 15 illustrates this basic fan-in application where multiple inputs are aggregated by a single VNF. This use case would be characterized by a larger number of lower speed interfaces facing towards the edge of the Network Operator's network and a smaller number of higher speed interfaces facing towards the core of the Network Operator's network. An assumption of this use case is that the single VNFCI has sufficient capacity to aggregate all of the input traffic.

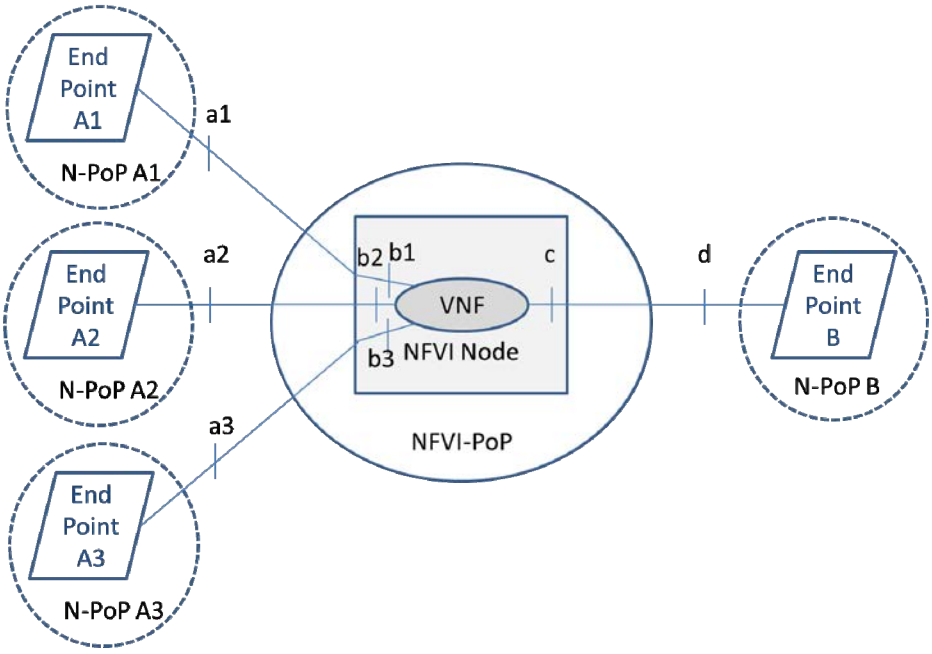

Figure 15: Ingress fan-in at an NFVI Node

Any VNFCI has some capacity limits, as does a VNF with a single VNFCI. If the processing load of a VNFCI is insufficient to fill the high speed links towards the core, then multiple VNFs (VNFCIs) may need to be deployed. figure 16 illustrates this scenario.

In both cases, end-end service graph requirements for service availability may require the traffic paths to be steered to different I/O interfaces of the NFVI Node in response to various failures (e.g. link failures, I/O interface failures). Reliability requirements are discussed in greater detail in ETSI GS NFV-REL 001 [i.51].

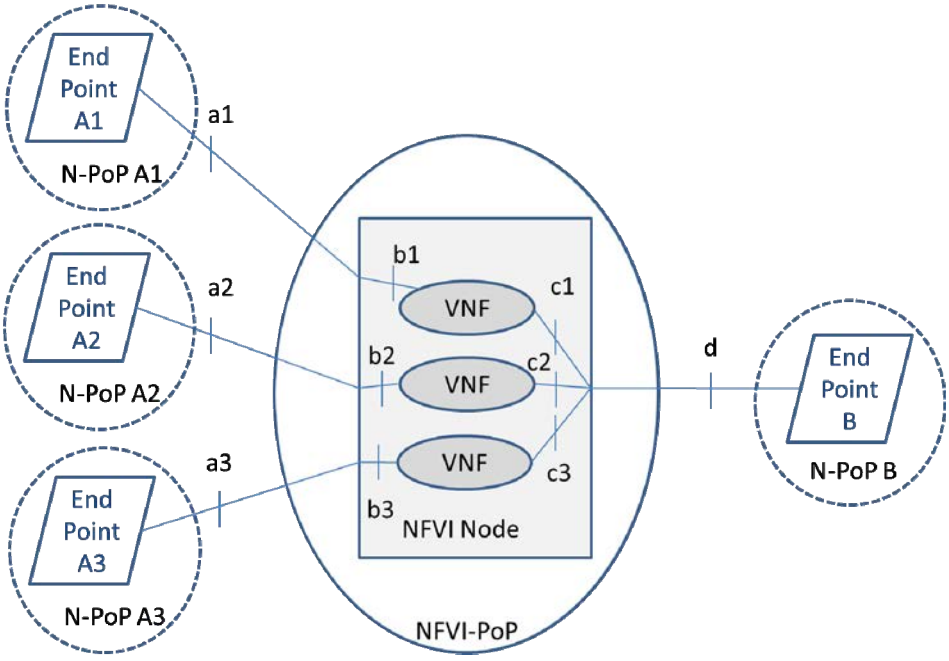

Figure 16: Egress fan-in at the NFVI Node

5.3.2 Nd-Nd

In terms of end to end service abstraction, the infrastructure network virtualisation layer is required to extend to different NFVI-PoPs. As referenced in figure 3, Use case #1: network function virtualisation infrastructure as a service (ETSI GS NFV 001 [3]), a service provider has to extend its NFVI across other service providers who offer NFVIaaS. Nd-Nd interface helps hide underlying network infrastructure details for inter-NFVI-PoP connectivity for seamless connectivity services between VNFs in different PoPs regardless of the connectivity service provided (i.e. L2 or L3).

5.3.2.1 Nature of the Interface

The interface consists of the protocols that are exposed between the NFVI-PoPs. The details of how the inter-NFVI-PoP connectivity services are provided are outside the scope of the present document.

On-demand creation of network connectivity among NFVI-PoPs is essential to meet the dynamic nature of traffic flows produced by VNFs. In cases where excessive traffic flows happen unexpectedly between any two NFVI-PoPs, traditional network management may not be able to deal with it efficiently and advanced network control and orchestration should be introduced, including carrier grade SDN based WAN Infrastructure Manager (WIM), which is outside the scope of the present document.

If new data plane network technologies are adopted in this interface, it may be necessary to extend protocols to provide a full set of connectivity services across domains. At the edge of each NFVI-PoP, a gateway provides the Nd-Nd data plane interface. This interface includes connectivity service virtualisation layers proposed in clause 7.1 over carrier networks such as LANs, MANs, and WANs:

- L2 overlay models.
- L3 models.

A gateway at the Nd-Nd interface includes a number of tools and protocols including centralized control protocols, cross-domain orchestration and routing. The following list is particularly important for the Nd-Nd interface:

- Planning tools e.g. Path Computation Element (PCE) [i.49], Traffic Engineering(TE) and on-demand bandwidth API.
- OAM protocols for measurement and monitoring, e.g. latency, bandwidth utilization.
- Protocols for traffic isolation, mutual authentication and authorization between different NFVI-PoPs.

5.3.2.2 Specifications in Current Widespread Use

In existing carrier networks, operators largely rely on OSS/NMS to provision networks services.

6 Functional Blocks within the Domain

6.1 Virtual Networks

A Virtual network is the network construct that provides network connectivity to one or more VNFs that are hosted on the NFVI. It is currently envisioned that a virtual network could be manifested by either utilizing the underlying NFVI network fabric, or instantiating an L2 or L3-based overlay network. A NFVI could make use of one or both mechanisms for its virtual networking requirements.

6.1.1 Infrastructure based Virtual Networks

An infrastructure-based virtual network is one that utilizes the native networking functions of the NFVI compute and networking components. They ensure traffic isolation by strictly partitioning the address space, but do not provide address space isolation. Where required, these virtual networks can coexist within the same data center as overlay or partitioned networks (where address space isolation is provided). Infrastructure-based virtual networks do not require the operator to manage a network overlay, however they are restricted to cases where the VNFs do not use overlapping addresses.

An example of a L3 infrastructure-based network is as follows:

- Each VNF may be assigned its own unique IP address (out of potentially multiple IP address ranges) which does not overlap with any other address of elements within the NFVI.
- Logical partitioning of the VNFs into their virtual networks is achieved by managing Access Control Lists in the L3 forwarding function in each compute node. The management of these ACLs can be managed by a centralized manager which sits under the network fabric control, e.g. Neutron.
- The L3 forwarding between VNFs and the physical fabric can then be handled by the L3 FIB running on the hosting compute node.
- Control plane solutions, such as BGP, can be used to advertise reachability of the VNFs to other compute hosts.

6.1.2 Layered Virtual Networks

A layered virtual network instantiates one or more private topologies on the underlying NFVI network fabric either by utilizing tunnels to interconnect endpoints in the private network, or by forming a virtual partition of the network and resources. It can support the same capabilities as an infrastructure-based virtual network, and also support overlapping address spaces, but requires the management of multiple overlay networks and the underlay infrastructure. In cases where non-overlapping addresses cannot be assured, NAT function may be required to provide the required connectivity.

6.2 Virtualisation Layer Options

Virtualisation layers are only necessary in the layer-based virtual network model. The infrastructure-based virtual network model does not make use of virtualisation layer(s), nor their component parts.

The primary purpose of the purpose of the virtualisation layer(s) is to provide virtual networks to interconnect the VNFs over the network resources, and to define service graphs that transit VNFs. There are at least two common approaches to virtualisation: virtual overlays and virtual partitioning. The chosen approach and network resources capabilities have an impact on the choice of control and data plane, and also have scalability, management, and networking capability tradeoffs.

6.2.1 Virtual Overlays

Figure 17 illustrates a virtual overlay methodology for creating virtual networks.

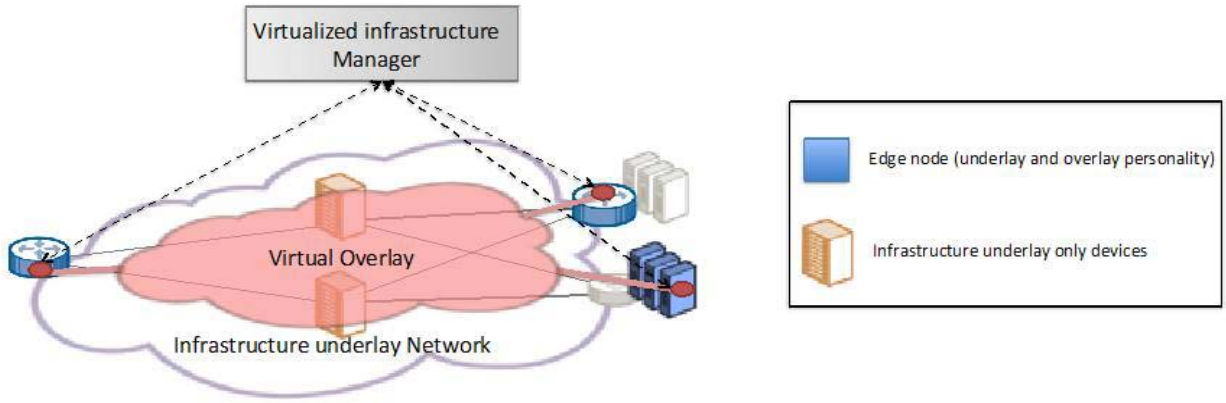

Figure 17: Virtual partitioning using virtual overlays

An overlay network is a network that is built onto another network, which is called the underlay network. In the NFV context the overlay networks are the virtual networks used by the VNFs and the underlay network consists of the infrastructure network resources. These overlay networks are normally created by edge nodes which have a dual personality, participating in both the creation of the virtual networks and also acting as infrastructure network resources. In contrast the core nodes of the infrastructure network only participate in the infrastructure network and have no overlay awareness. In an NFV environment the edge nodes can be virtual switches, TORS and gateways. Typically, each edge node will support multiple virtual network which consist of a per virtual network forwarding tables, which may be L2 or L3 depending on whether the overlay is layer 2 or layer 3 network, along with per virtual network virtual links. These virtual links are normally created by encapsulating the customer's traffic with a packet associated with the underlay network along with a virtual network identifier. This is used by the destination edge device to identify which virtual network the packet belongs too. This type of solution maximizes the separation between the virtual network and the infrastructure network resources, meaning the overlay orchestration and control plane are largely distinct from the underlay network and its control plane. This increases flexibility, allows sharing of network resources and achieving multipath resiliency but it is important to note that the SLAs of the virtual networks or overlays are dependent on the availability and functionality of the underlay or infrastructure network resources. This is a very common way to build virtualised network in both L2 and L3 environments and solutions such as VXLAN, MPLS, PBB all utilize this type of technology.

6.2.2 Virtual Partitioning

A second common approach to building virtual networks is where the virtual network partitions are directly integrated into infrastructure network on an end-to-end basis. This is illustrated in figure 18.

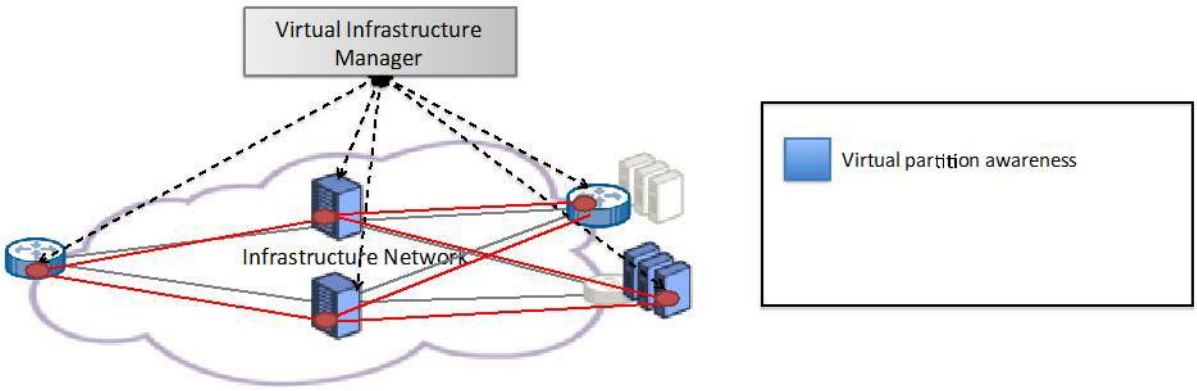

Figure 18: Virtual partitioning in the infrastructure network

In this case discrete virtual topologies are built in both the edge and core nodes of the infrastructure network for each virtual network. This can consist of per virtual network forwarding tables, logical links and even control planes on an end-to-end basis across the infrastructure network. This approach is seen in MSTP VLAN environments, where each virtual network partition is created using a unique MAC table, connected together with a dedicated VLAN with a dedicated "Spanning Tree" instance running on the infrastructure network. A similar scenario could be envisaged in some flow-based environments where flow tables are programmed on an end-to-end basis across the infrastructure network. It is important to note that the SLAs of the virtual networks are dependent on the availability of the network resources and physically constrained by loop-free path alternatives.

6.2.3 Abstract Layering Model

The existence of VNFs that are transparent to the customer layer and the requirement to be able to uniquely address and instrument service graphs suggests the existence of three or more protocol layers or sub-layers discernable within the NFVI virtual network domain. These are:

1) The customer layer which consists of packets originating with and/or terminated by the customer end-system or host.

2) The service layer, which exists exclusively within the NFVI and has end to end significance between the point of ingress to the NFVI and point of egress to the NFVI of the customer traffic.

3) The virtual network layer which interconnects VNFCIs or connects VNFCIs to the NFVI ingress or egress.

4) The infrastructure layer which in an overlay context connects hypervisors together.

The relationship between these layers is purely hierarchical with lower layers encapsulating the higher layers and encapsulated frames being explicitly opaque to underlying layers. Figure 19 depicts this.

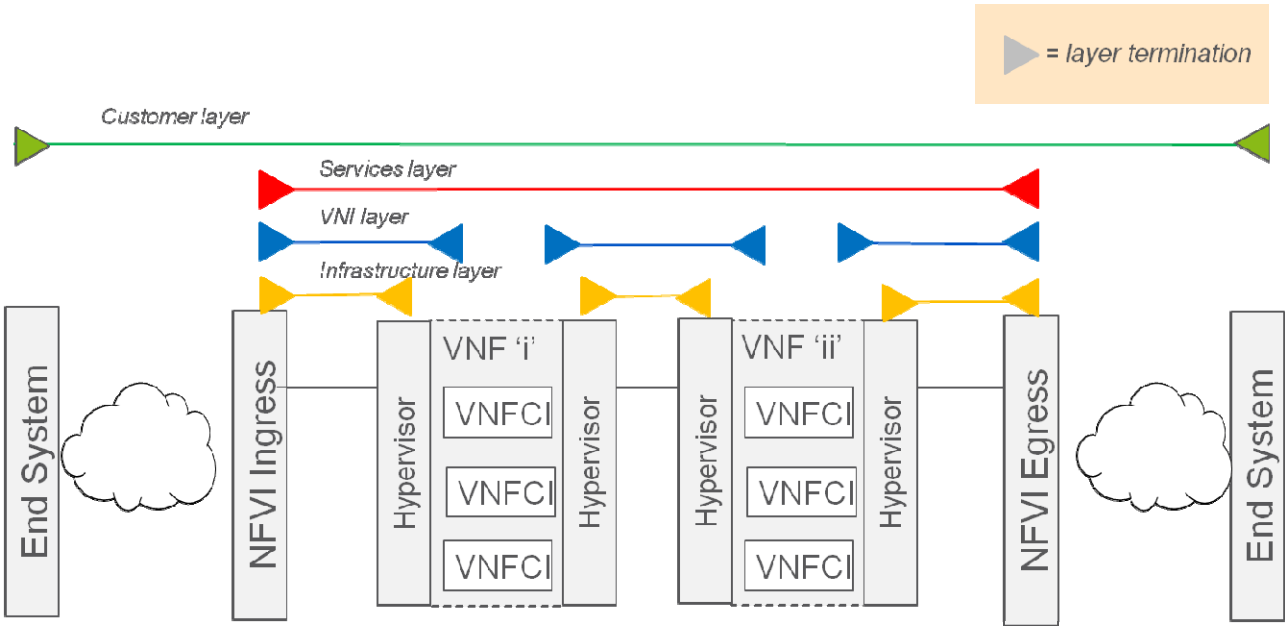

Figure 19: Layering model

6.2.4 Examples

Table 4 gives an overview of explicit encapsulation technologies that add specific information to a packet that is needed for tunnelling. Note that there are also techniques in use, such as for example based on OpenFlow, that provide similar tunnelling without adding information to a packet.

Table 4 provides a comparison of encapsulation methods.

Table 4 highlights some of features of well-known encapsulation mechanisms. The status of the referenced documents is set out in the reference list as some are still work in progress.

Table 4: Comparison of Encapsulation Methods

Encapsulation mechanism	Reference	Network Id (tVNID)	Notes
STT	draft-davie-stt-04 [i.6]	64-bit Context ID	STT is an IP-based encapsulation and utilizes a TCP-like header inside the IP header. It is, however, stateless. STT is particularly useful when some tunnel endpoints are in end-systems, as it utilizes the capabilities of standard network interface cards to improve performance. A 64-bit field in the STT frame header that conveys information about the disposition of the STT frame between the tunnel endpoints. One example use of the Context ID is to direct delivery of the STT frame payload to the appropriate virtual network or virtual machine.
VxLAN	draft-mahalingam-dutt-dcops-vxlan-06 [i.7]	24-bit VXLAN Network Identifier (VNI)	VXLAN addresses Layer 2 and Layer 3 data center infrastructure network in the presence of VMs in a multitenant environment. It uses a VLAN-like encapsulation technique to encapsulate MAC-based layer 2 Ethernet frames within layer 3 UDP packets. VXLAN increases scalability up to 16 million logical networks and allows for layer 2 adjacency across IP networks. Each overlay is termed a VXLAN segment. Only VMs within the same VXLAN segment can communicate with each other.
NVGRE	draft-sridharan-virtualization-nvgre-03 [i.8]	24-bit Virtual Subnet Identifier (VSID)	NVGRE uses Generic Routing Encapsulation (GRE) to tunnel layer 2 packets over layer 3 networks. In NVGRE, every virtual Layer-2 network is associated with a 24-bit VSID carried in an outer header that supports up to 16 million virtual subnets in the same management domain. Each VSID represents a virtual Layer-2 broadcast domain.
IEEE Std 802.1Q-2012	802.1Q-2012 [i.4]	24-bit Backbone Service Instance Identifier (I-SID) 12-bit & 12-bit Customer VLAN ID & Service VLAN ID (C-VID & S-VID) 12-bit VLAN Identifier (VID also called a C-VID)	IEEE Standard for Local and metropolitan area networks—MAC Bridges and Virtual Bridged LANs. IEEE Std 802.1Q-2012 [i.4] supports three types of Ethernet virtual network encapsulations which are Virtual LANs (VLANs, or Customer VLANs), Service VLANs (S-VLANs), and Backbone Service Instances (BSIs). The standard supports arranging these three encapsulations in a hierarchy allowing up to 16,776,959 backbone service instances each carrying up to 4094 Service VLANs (S-VLANs) and each of these S-VLANs carrying up to 4 094 Customer VLANs (C-VLANs or just VLANs). The standard defines: • a system of VLAN and Service Instance tagging and encapsulation for Ethernet frames which allows both backward compatibility and upward scaling from older Ethernet bridge revisions (clause 5, annex A); • link state (SPB), distance vector (RSTP), and software defined (PBB-TE) topology protocols to configure forward of both individual and group addressesed frames (clauses 13, 14, 25, 26, 27, 28, IEEE Std 802.1Qbp-2013) [i.19]; • hierarchical OA&M management which operates across all tags and encapsulations (clauses 19, 20, 21, 22); • element management MIBs for each component (clauses 12, 17); • provisions for a quality of service including: class queuing and class transmission selection, priority queuing, L2 flow control, L2 congestion management, traffic shaping, and network time synchronization (clauses 8, 9, 30, 31, 32, 33, 34, 35, 36, 37, 38); • specifications for L2 vSwitches and protocols for coupling vSwitches to the physical network infrastructure (clauses 40, 41, 42, 43).
MAC in GRE IP in GRE	IETF RFC 2784 [i.9] IETF RFC 1702 [i.10]	n.a.	GRE is a simple, general purpose encapsulation mechanism. [i.4]
MAC in PWE3	IETF RFC 3985 [i.11]	n.a.	Pseudowire Emulation Edge to Edge (PWE3) specifies the encapsulation, transport, control, management, interworking and security of services emulated over IETF-specified PSNs.

Encapsulation mechanism	Reference	Network Id (tVNID)	Notes
VPLS (layer 2 MPLS VPN)	IETF RFC 4448 [i.12] IETF RFC 4761 [i.13] IETF RFC 4762 [i.14]	n.a.	The IETF RFC 4448 [i.12], describes a VPN model to carry Ethernet [i.15] packets over an MPLS network. It enables service providers to offer "emulated" Ethernet services over existing MPLS networks. The L2VPN Working Group produced two separate documents, IETF RFC 4762 [i.14] and IETF RFC 4761 [i.13] that perform similar functions in different manners. Two different approaches have been defined for two distinct applications. Virtual Private LAN Service (VPLS), also known as Transparent LAN Service and Virtual Private Switched Network service offers a Layer 2 Virtual Private Network (VPN); however, in the case of VPLS, the customers in the VPN are connected by a multipoint Ethernet LAN, in contrast to the usual Layer 2 VPNs, which are point-to-point in nature. VPLS offers a "switch in the cloud" style VPLS service. VPLS provides the ability to span VLANs between sites. L2 VPNs are typically used to route voice, video, and AMI traffic between substation and data center locations.
VPRN (layer 3 MPLS VPN)	IETF RFC 4364 [i.16] draft-ietf-l3vpn-end-system-02 [i.43]	n.a.	MPLS VPN is a family of methods for harnessing the power of multiprotocol label switching (MPLS) to create virtual private networks (VPNs). MPLS VPN gives network engineers the flexibility to transport and route several types of network traffic using the technologies of a MPLS backbone. Utilizes layer 3 VRF (VPN/virtual routing and forwarding) to segment routing tables for each "customer" utilizing the service. The customer peers with the service provider router and the two exchange routes, which are placed into a routing table specific to the customer.
L2TP	IETF RFC 2661 [i.17]	n.a.	L2TP facilitates the tunneling of PPP packets across an intervening network in a way that is as transparent as possible to both end-users and applications. The Layer 2 Tunnel Protocol (L2TP) is an emerging IETF standard that combines the best features of two existing tunneling protocols: Cisco's Layer 2 Forwarding (L2F) and Microsoft's Point-to-Point Tunneling Protocol (PPTP). L2TP (Layer 2 Tunneling Protocol) is also known as a pseudowire.
TRILL	IETF RFC 6439 [i.18]	16 bit TRILL header	TRILL creates a cloud with a flat Ethernet address, so that nodes can move around within the cloud and not need to change their IP address. Nodes attached to the cloud perceive the cloud as Ethernet while the packet is traversing the cloud, however it is encapsulated with a TRILL header, which contains a source (ingress RBridge), destination (egress RBridge), and hop count. The addresses in the TRILL header supports 64,000 RBridges. TRILL supports VLANs and multicast.

6.3 Network Resources

At the data plane level, the infrastructure network could employ any of the following technologies: Ethernet switching, TRILL switching, MPLS switching, IP routing, flow based switching or a combination; for example L2 switching at the edge (in the TORS and virtual switches) and L3 switching in the core of the infrastructure network. Clearly the control plane used needs to match the infrastructure networking data plane technology. Some of the distributed control plane options are outlined below:

- Ethernet switching ([i.4], [i.19],[i.20]): SPBV, SPBM, SPB-ECMP, SPB-PCR, RSTP, MSTP, PBB-TE, LAG, D-LAG.
- TRILL switching: IETF RFC 6325 [i.21], IETF RFC 6327 [i.22], IETF RFC 6439 [i.18].
- MPLS switching: Layer 3 protocol + LDP, RSVP, MP-BGP.
- IP switching: OSPF, IS-IS, E-BGP, I-BGP.

In addition the infrastructure control plane could be implemented using a centralized SDN approach. In this case there would be a centralized control element that computes the forwarding tables. This information is then communicated to the infrastructure network elements, which then forms the basis for forwarding traffic. Forwarding decisions could be made based on flow entries or traditional MAC/L3 forwarding entries.

Hybrid operation with SDN and distributed control may be supported using SDN in parallel with and the IEEE distributed protocols. The hybrid control division is supported by IEEE Std 802.1Q [i.4] through allocation of VID and S-VID spaces to the independent control protocols. In addition, it is possible to use SDN at the edge and a distributed control protocol in the core. The coupling between these can be achieved using the IEEE Std 802.1Q [i.4] Virtual Station Interface Discovery and Configuration Protocol (VDP).

6.4 Control & Admin Agents

The Control & Admin Agents include control and OAM functions.

6.4.1 Control plane

The NFV architecture requires a control plane at the infrastructure network level. This control plane provides the fundamental services needed to manage connectivity in the network and control packet forwarding. There are a number of control plane approaches that could be employed in an NFV infrastructure network. The precise choice will depend on the forwarding technology employed, the virtualisation paradigm and whether the control plane is implemented in a distributed fashion, for example OSPF routing or in a logically centralised fashion, for example, following the SDN approach.

6.4.1.1 Control Plane Functions

The control plane provides a wide range of functionality including the following: topology detection, traffic isolation, routing and path convergence.

6.4.1.1.1 Topology and device Detection

The discovery of the physical and logical network components associated with infrastructure network.

Topology detection and particularly topology change detection is typically conducted in a distributed manner by a family of control protocols including LLDP, OSPF, BGP, etc. Note that some of the protocols mainly serve routing purposes, but also provide topology information.

Note that the control plane operating at the network infrastructure level may be used to provide topology or reachability information of the VNFs.

EXAMPLE: IETF RFC 4364 [i.16] describes how BGP can be used to advertise reachability of elements within a client layer, such as VNFs. By flooding this information in the control plane, other NFVI elements can determine where the relevant VNFCIs reside, and route or switch traffic destined for them to the relevant NFVI egress.

6.4.1.1.2 Virtual partitioning

In some scenarios the infrastructure network control plane is also responsible for building the virtual network in which the VNFs operate. This is described in clause 6.2.

6.4.1.1.3 Traffic Isolation

Traffic isolation is a key function of the control plane in data centers that typically does not use specific protocols but is implemented as component of management and control functions.

The traffic isolation refers to how traffic is identified and forwarded to associated virtual networks on NFVI. The control plane should know how traffic is identified in order to restrict the incoming traffic. Unknown traffic should not be associated with and transferred to any virtual network.

The control plane should also know how traffic is associated with virtual networks. Packet header parameters (for example VLAN tags, source and destination addresses, port numbers, protocols or MPLS labels) can be used to classify traffic from a shared network into flows. These flows can then be mapped to one or more virtual networks.

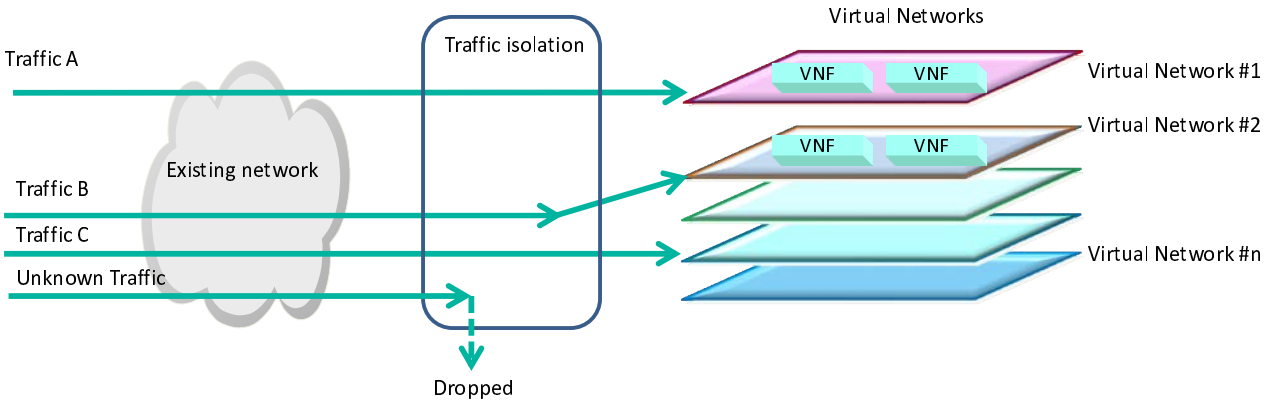

Figure 20: Traffic Isolation

6.4.1.1.4 Reachability

Reachability: Calculation/learning and distribution of reachability information associated with the infrastructure network.

Routing is most commonly conducted in a distributed manner by a family of control protocols including IS-IS, OSPF, BGP, etc.

EXAMPLE: BGP can advertise (and withdraw) reachability to elements within the NFV infrastructure or higher layers.

6.4.1.1.5 Traffic Engineering/Path Computation

Traffic Engineering provides a network operator the ability to control the usage of resources in the network. Examples include: better capacity utilization, preferred use of paths or nodes, diversity, etc. Explicit path computation and configuration are used to support Traffic Engineering in today's networks, though it is limited in that it is dependent on manual intervention or use of a control plane/management system which can support it. For complex networks (multi-layer, large-scale networks), it is very challenging for network operators to utilize Traffic Engineering and to be able to exploit the benefits of its use. SDN and NFV will enable supporting this functionality more efficiently, for either a distributed control plane or a centralized control plane, with a variety of possible eco-systems and open standards.

6.4.1.1.6 Flow Management

As described in clause 5.2.1, the infrastructure network domain can be implemented using a centralized controller and provide network programmability. With flow management, as in OpenFlow, flow tables in an infrastructure network equipment can be managed, e.g. in an OpenFlow enabled switch, to control forwarding behavior.

EXAMPLE: At the point of encapsulation, incoming flows are identified and forwarded to appropriate tunnel end points based on the flow table entries. The flow management adds, updates, and deletes the flow entries, resulting in unicasting, multicasting and broadcasting of the flows.

6.4.1.1.7 Failure detection

Failure detection is the discovery and notification of node and link failures in the infrastructure network.

6.4.1.1.8 Convergence

Convergence is the synchronization of the topology state and reachability recalculation triggered by an event within the infrastructure network (e.g. failure, addition of links).

6.4.1.1.9 Quality of Service

Techniques to support Quality of Service vary based on the technology, e.g. traffic prioritization, flow policing, and more sophisticated techniques using Traffic Engineering.

6.4.1.1.10 Policy

Policy can be based on simple policies (preconfigured) or be more complex and require communication with an external policy management entity.

6.4.1.2 Control Plane Approaches

Traditionally control planes execute in a distributed fashion in packet switching environments, meaning there is a control plane instance per switching/routing instance whether this is a physical or logical packet device. With the advent of SDN, new solutions are emerging where the control plane is removed from the data plane and executes in a central location and the forwarding information is conveyed from the central "brain" to the data plane devices using a protocol such as OpenFlow. It is anticipated that some users will select a distributed control plane approach, some will use a fully centralized approach, while others will use a hybrid approach. This combines centralized and distributed functionality. The different approaches are illustrated in figure 21.

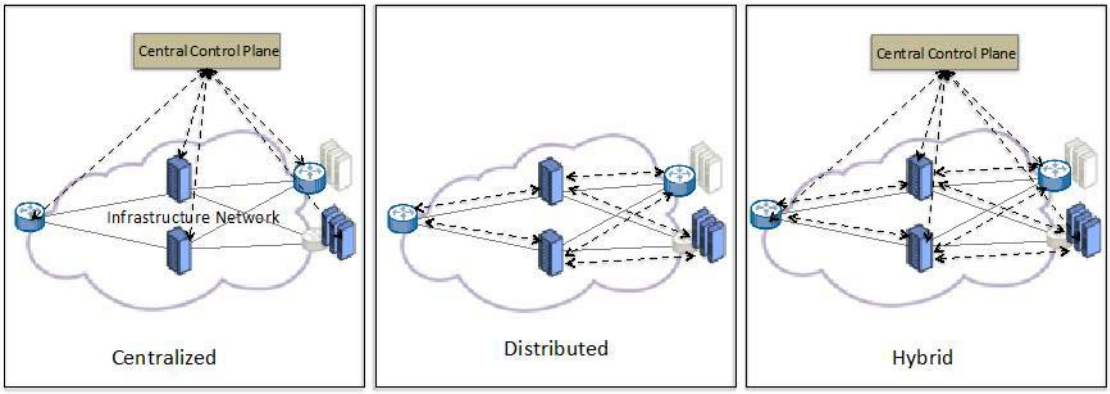

Figure 21: Control Plane Models

6.4.2 North-South OAM Interface

North-South OAM interfaces can be grouped into three different categories: management protocols, management modules/information and data models and flow monitoring protocols.

- The configuration of virtual networks requires interfaces to all relevant components, including network elements in the infrastructure network domain, vSwitches/vRouters in the hypervisor domain and embedded switches/routers in the compute domain. The interfaces to the vSwitches/vRouters and embedded switches/routers are described in their respective domain documents. The relevant interfaces to the infrastructure network domain network elements include: SNMP [i.31].
- NetConf [i.32].

SNMP developed by the IETF is the most widespread management protocol with probably the largest number of data models called MIB modules. NetConf is the designated new management protocol by the IETF. It uses new modern XML encoding and comes with a new modelling language called YANG. As YANG is relatively new, data models are under development. Availability by equipment is rising but still not as omnipresent as SNMP. Its advantage is the simple modelling and the support of different transport protocols it can use. This also simplifies security for the protocol.

Regarding data models there are multiple sources for the above protocols:

- SNMP MIB modules are provided by IETF and vendors.

- NetConf models are provided by IETF and vendors.
- OpenFlow specific NetConf modules are provided by ONF.

SNMP and NetConf both allow implementing vendor specific data models apart from those that are standardized by the IETF. One vendor specific model that is relevant for Orchestration and Management is the OF-CONFIG module based on NetConf and specified by the ONF.

Besides the management protocols and their data models there are also a number of flow monitoring and sampling protocols used in the industry:

- sFlow (http://www.sflow.org/).
- NetFlow [i.33].
- IPFIX (https://tools.ietf.org/wg/ipfix/).

All those protocols do flow sampling and monitoring and report their results to a server instance. sFlow version 4 and 5 and NetFlow version 9 are industry standards. IPFIX is the standardized flow sampling and monitoring solution of the IETF. IPFIX is sometimes also referenced as NetFlow version 10.

6.4.3 East-West OAM Interface

6.4.3.1 OAM and the Abstract layering model

Figure 22 illustrates well understood industry MEP/MIP models mapped onto the abstract layering model presented in clause 6.2. Layer terminations correspond to MEPs and interrogatable intermediate points correspond to MIPs. Common practice is that a MEP of a server layer is co-located with either a MEP or a MIP of a client layer. This permits comprehensive capability to perform fault sectionalisation.

OAM itself is typically of two forms, proactive and reactive. Proactive OAM, e.g. CFM [i.23], is in the form of scheduled and continuous probing of the network to proactively detect faults or degradation of service and is performed end to end for a given layer. Reactive OAM, e.g. Link trace [i.23], Loopback, is usually operator attended and is used for fault sectionalisation in response to an already reported fault condition. Reactive OAM involves interrogation of MIPs to determine which section of a path has failed, and hence the server layer to progressively "drill down" until the location of the root cause of the fault can be ascertained. Refer to figure 22.

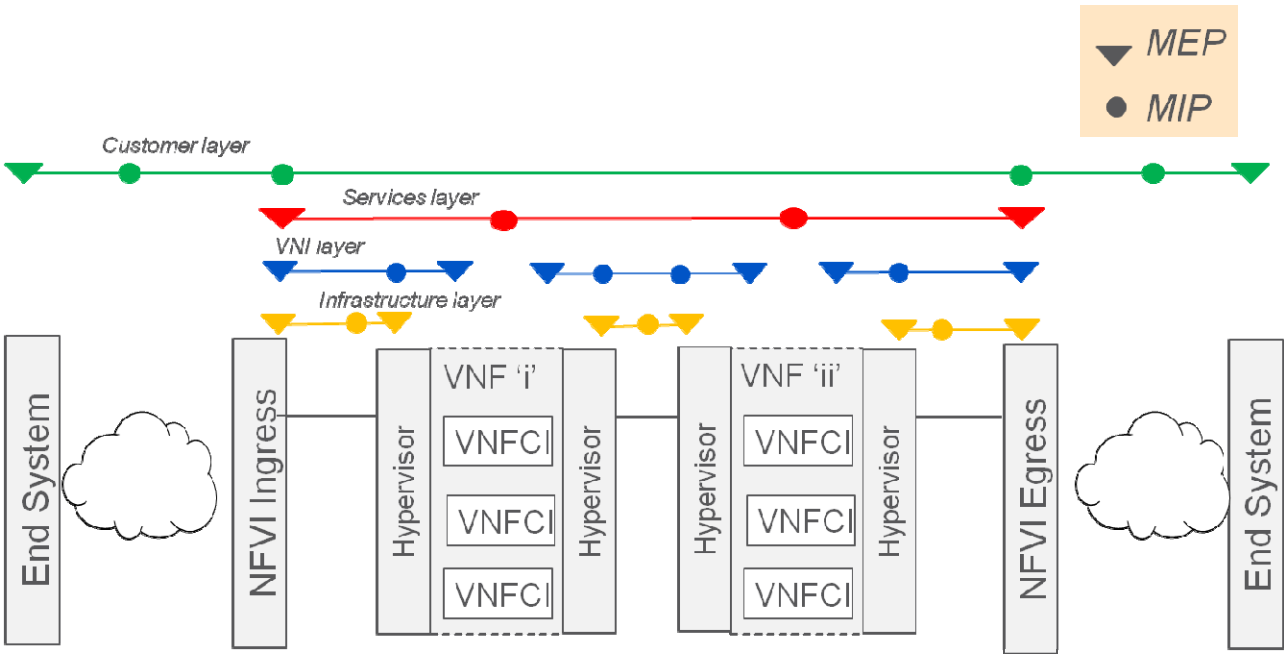

Figure 22: OAM MEP/MIP model

6.4.3.2 Layer 2 OAM Protocols

Within the NFVI layer, a set of IEEE Ethernet OAM protocols may be used to manage the Ethernet links and L2 segments of the NFVI.

- Station and Media Access Control Connectivity Discovery IEEE Std 802.1AB [i.24].
- Edge Virtual Bridging IEEE Std 802.1Qbg [i.25].
- Priority Based Flow Control IEEE Std 802.1Qbb [i.26].
- Enhanced Transmission Selection IEEE Std 802.1Qaz [i.27].
- Link Aggregation IEEE Std 802.1AX [i.28] and IEEE 802.1AX-Rev [i.20] (draft in progress).
- Timing and Synchronization IEEE Std 802.1AS [i.29].
- Connectivity Fault Management IEEE Std 802.1ag [i.23].
- Congestion Notification IEEE Std 802.1Qau [i.30].

This set of protocols runs at the L2 layer. An end-to-end service may span across different types of network technologies. While different OAM protocols can be applied for the different network technologies, the transparency for OAM protocols for an end-to-end service shall be ensured over underlying transport technologies such as MPLS or Optical-related (e.g. SDH, OTN).

6.4.3.3 Layer 3 OAM Protocols

There is a large set of OAM protocols for IP and MPLS, including: IP Ping, IP Traceroute, BFD OAM, MPLS OAM, and Pseudowire OAM. A comprehensive description is provided in [i.47].

NOTE: This set of protocols runs at L3. They will be tunneled over any L2 or L1 boundary (encapsulation), and will not interact with these transport technologies.

6.4.3.4 Layer 3 OAM Protocols

The following protocols may be used both within the services layer, and, in the case of an Infrastructure L3 service, to provide OAM capabilities for the service itself:

- IPv4/IPv6 Ping using ICMP (IETF RFC 792 [i.52], IETF RFC 4443 [i.53]).
- IPv4/IPv6 TraceRoute (described in IETF RFC 2151 [i.54]).
- BFD for IPv4/IPv6 (IETF RFC 5880 [i.55], IETF RFC 5881 [i.56]).

These are a (non-exhaustive) list of IP OAM tools available. A more comprehensive (but not complete) list can be found in [i.47].

7 Interfaces within the Domain

7.1 [Vl-Ha]/Nr

Virtual networks are created by the virtualisation layer resource routing and sharing control that provides abstraction of network resources at the [Vl-Ha]/Nr interface. Figure 23 illustrates ways in which NFVI network resources from the Compute, Hypervisor and Network Domains may work in concert to provide a virtual network in terms of interfaces that correspond to the following reference points:

- [Vn-Nf]/N, the service interface over which VNFs access the virtual network;
- Ha/CSr-Ha/Nr, the access interface between the network resources in the Compute and Hypervisor Domains and NFVI infrastructure network equipment; and

- Ex-Nf, the external interface between the NFVI and the existing network.

While the primary focus of the present document is infrastructure networking, this clause also describes at a high level the functions provided by the Compute and Hypervisor Network Resources and by Infrastructure Network Equipment.

Networking functions provided within the Compute Domain include:

- Physical network interface controllers (NICs), documented in ETSI GS NFV-INF 003 [i.2];
- Virtual Ethernet Bridges (VEB) in NICs;
- Virtual Ethernet Port Aggregation (VEPA) in NICs; and
- Base L3 forwarding capabilities in the underlying kernel.

Networking functions provided within the Hypervisor Domain include:

- Virtual network interface controllers (vNICs);
- Virtual switches (vSwitches), which may be supplemented by a native or virtualised router; and
- Virtual routers (vRouters).

vNICs and vSwitches are documented in ETSI GS NFV-INF 004 [i.3].

Networking functions are also provided in the Infrastructure Network Equipment that makes up the switching fabric.

Transport network functions can potentially utilize a combination of techniques to isolate virtual networks, such as:

- VLANs or other headers to mark traffic;
- Virtual Private Networks (VPNs); and/or
- Overlay tunnels.

Gateway functions provide the interconnection between NFVI-PoPs and the transport networks. They also connect virtual networks to existing network components.

7.1.1 Layer 2 overlay model

Figure 23 illustrates a NFVI-PoP using a simple two level data center switching fabric topology.

NOTE: This is but one example of possible topologies, many of which use more than two levels.

Figure 23 shows different scenarios for interconnecting VNFs with various hosting options, ranging from fully virtualised to bare metal. Note that these are for illustrative purposes only and are not prescriptive. The example here shows an overlay virtualised network providing L2 connectivity. Similar mechanisms apply to other scenarios, such as a L3 overlay operating over a L2 underlay, or an infrastructure-based virtual network.

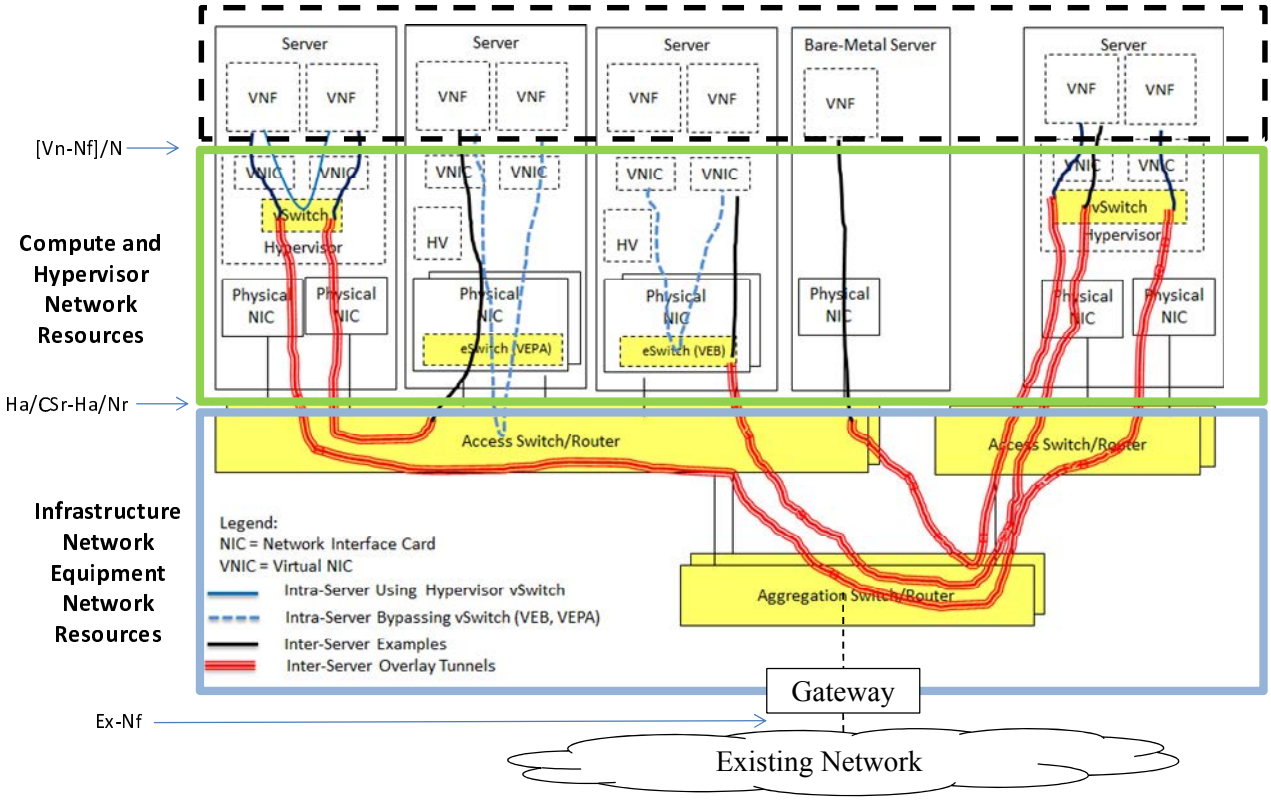

Figure 23: L2 Connectivity Examples - Range of Virtualisation

As shown above, figure 23 illustrates the use of L2 overlay tunneling over a L3 IP underlay network fabric to interconnect VNFs attached to NICs or vNICs within a data center. These tunnels originate and terminate in edge devices, which could be vSwitches or access switches. The underlay fabric is represented by access and aggregation devices. The forwarding tables in these devices may be provided using a traditional distributed control plane, such as one using IS-IS to compute routes, or by a logically centralized SDN controller communicating with fabric devices using a protocol such as OpenFlow. Overlay tunneling technology can also be used to implement virtual networks that span multiple physical sites using tunnel end points (TEPs) in different NFVI-PoPs.

Figure 23 also illustrates how multiple connectivity approaches might be involved in providing the virtual networks. A solid blue line is used to represent the traffic path between two VNFs in the same virtual network that are hosted on guest virtual machines by the same physical server. In this case, intra-server VNF traffic can be directly forwarded by a VEB (vSwitch) implemented in software in the hypervisor. Similarly, the dashed blue lines show variants of intra-server VNF interconnection where the switching is done using a hardware VEB in the Physical NIC, and at the edge hardware switch using VEPA or VN-tags [i.37]. In both of these cases, VMs bypass the hypervisor to directly access the physical NIC.

The black lines illustrate VNF to VNF traffic that crosses server boundaries and needs to be overlay tunneled. Overlay tunnel traffic paths are shown as thick red lines. Two cases are shown, the first case being where the tunnels originate and terminate in the hypervisor software vSwitches, and the second where tunnels originate and terminate in edge hardware switches.

A bare metal example is also shown where the VNF is running directly on the server hardware and not in a virtual machine hosted by a hypervisor. This case may be used to accommodate, for example, software component VNFs that run on operating systems that do not have tunneling capabilities. Such a VNF can participate in the virtual network enabled by overlay tunnels that originate and terminate in edge switches.

7.1.2 Layer 3 models

Figure 24 is similar in scope to figure 23, but instead of utilizing the L2 overlay model, it utilizes the L3 models (L3VPN and Infrastructure). It goes further than figure 23 and shows an NFV infrastructure that spans two NFVI-PoPs, using a intra-NFVI-PoP switching plane and an inter-NFVI-PoP routing plane to interconnect the vRouters and base L3 forwarding engines.

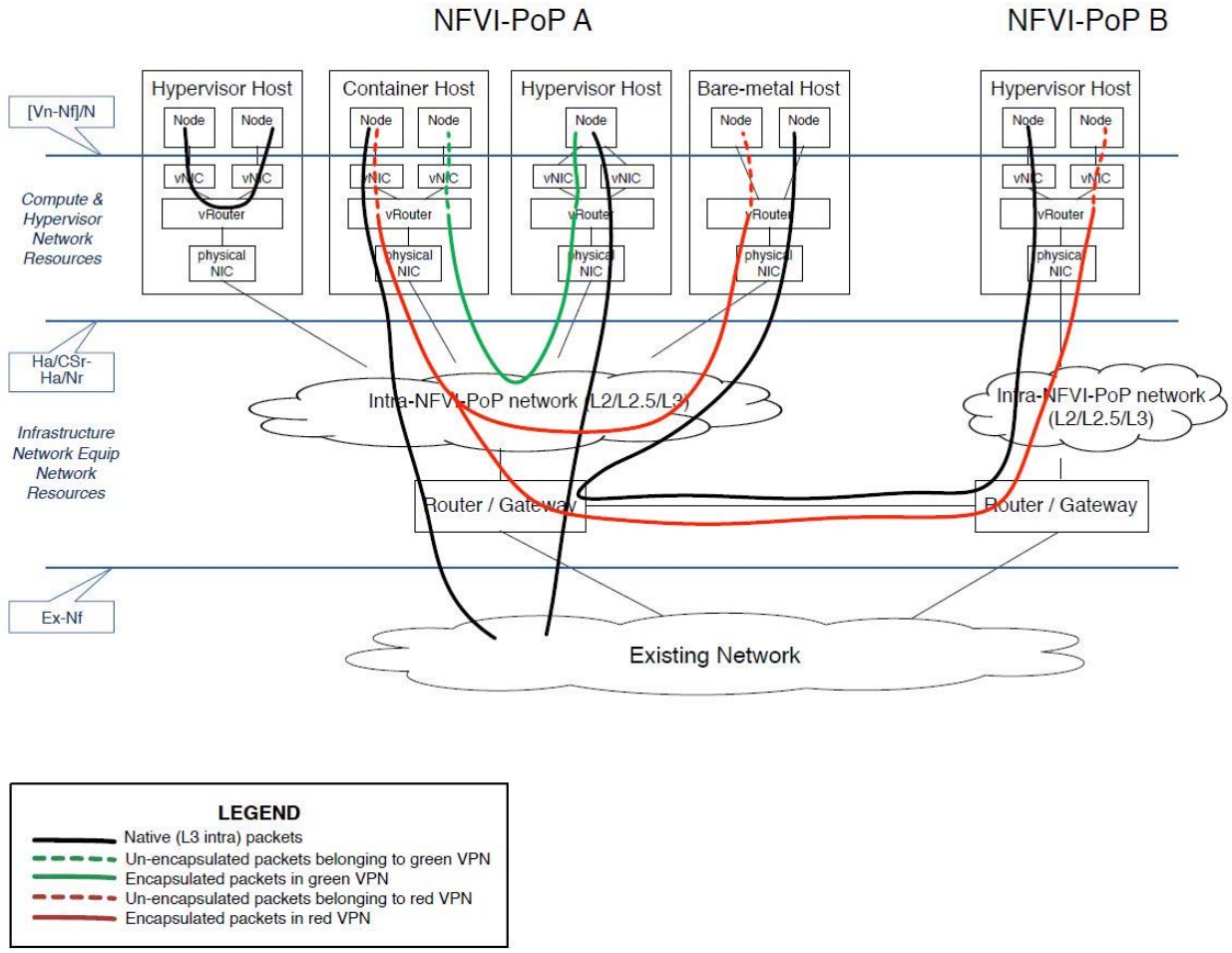

Figure 24: L3 Connectivity Examples - NFVI network of two PoPs

Figure 24 provides an illustrative set of data plane flows that may be found in both L3 VPN and L3 infrastructure NFVI networks. It is not an exhaustive set. In contrast with the L2 overlay model illustrated in figure 23, figure 24 also shows an NFVI that is comprised of multiple sites or PoPs. The hosts represented in the diagram comprise hypervisor compute nodes, container (or non-hypervisor compute nodes) and bare metal devices, which could be bare metal compute hosts, storage appliances, legacy network appliances, etc.

All of the hosts in the diagram show a vRouter component. That could be either an add-on software component, or simply the L3 forwarding capabilities of the underlying operating system. The nodes would normally be VNFCI nodes, but could be other end points in the service infrastructure. The Intra-PoP network can be any infrastructure that allows for the transport of IPv4 and IPv6 packets (e.g. Ethernet, MPLS, IP, or native IP). All other terms are self-explanatory.

Black paths show nodes communicating via the L3 infrastructure option. The node forwards its traffic to its local vRouter, where the traffic is routed to the destination (another node in the same or a different host in the PoP, a node in a different PoP, or some external end point), provided any policy constraints are met.

Red paths show nodes that are members of the Red private topology communicating over a dedicated L3 VPN. The dashed Red paths show the "native" packets sent to or received from the local vRouter by the node. The solid Red paths show the encapsulated L3 VPN traffic carried over the common infrastructure.

Green paths are the same as for the Red paths, except showing another private topology and related dedicated L3 VPN. In the case of one of the nodes in the Green path, there is no dashed Green path. That is because the node itself is the endpoint of the VPN (the node is L3 VPN capable); in this case, the local vRouter forwards the traffic in the same way that a L3 infrastructure packet is forwarded.

Figure 24 also shows that a node can utilize both the L3 infrastructure and L3 VPN modes concurrently. The decision whether to forward via L3 VPN or L3 infrastructure can be made by the vRouter, under the control of the policy framework, or by the node itself (in the case where a node has two vNICs, one for L3 infrastructure traffic and one for the L3 VPN traffic, or in the case where the node is L3 VPN-aware and makes the decision to encapsulate a given packet in an L3 VPN overlay or not).

7.1.3 Specifications in Current Widespread Use

7.1.3.1 Encapsulation Specifications

IETF has published or is in the process of developing specifications for encapsulation protocols. Refer to table 3 in clause 6.3.3.

7.2 Ha/CSr-Ha/Nr

7.2.1 Interface to the NIC

This clause describes the interface between the NIC on the server and the Ethernet Switch or router. This interface has dependencies on the encapsulation location and interworking.

7.2.1.1 Nature of the Interface

This interface consists of the protocols visible between the server NIC and its adjacent Ethernet Switch or router.

8 Modularity and Scalability

The NFV architecture allows for more than one technology to be used in the infrastructure network. This plurality of infrastructure networking solutions creates an interworking requirement.

8.1 Interworking Strategies

There are three basic strategies that can be used to provide an interworking function. These are:

- *Interworking using a Gateway:* An interworking gateway may be placed between the regions with different technologies;
- *Interworking using Multi-Protocol Tunnel Terminations:* The NFVI edge switches (i.e. vSwitch, TORS, etc.) or routers can support multiple protocols; and
- *Interworking in the VNFCI:* All VNFCIs in an NFVI that need to communicate over a plurality of infrastructure networks are connected to each required network.

The NFVI provides a consistent L2 or L3 service to every NF. In an overlay-based NFVI, the service is carried over the NFVI through an encapsulation that supports VNP. A model for the NFVI, illustrating the placement of an Interworking Gateway (IG), is shown in figure 25. Tunnel End Points (TEPs) can be located within a vSwitch, vRouter, TORS or a VNFCI dedicated for that purpose. The network management and orchestration co-ordinate the creation of VNs by provisioning the IGs to connect between two technology regions. If the IGs are multi-protocol, then it may be necessary to provision the required encapsulation for each destination within the VN, if encapsulation (i.e. overlay-based virtual networks) is in use.

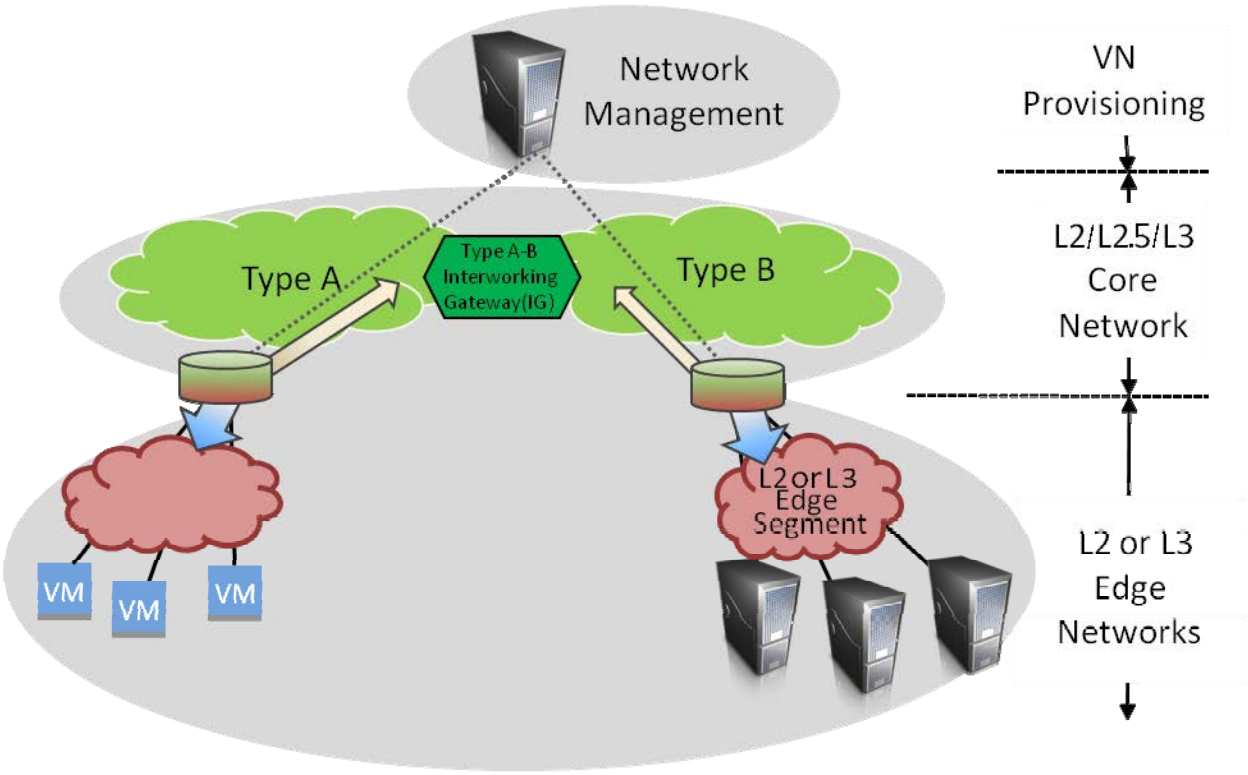

Figure 25: Relationship between TEPs, Gateways and Management in NFVI

Table 5 shows a mapping of a selection of L2 bearing technologies and some of the basic service profiles they can natively support (for example, E-LAN, E-LINE and E-TREE). It is not an exhaustive list, and the capabilities of these technologies may change over time based on implementation and standards development. It is recommended that, for definitive answers as to capability, the relevant standards and/or implementation guides for the services/profiles in question be consulted. Note that, although several L2-bearing technologies could in principle be used to carry an E-TREE service, there are currently no complete standards for some of them. Completed standards may emerge in due course.

Table 5: Mapping of Ethernet Network Service Models to Underlying Technology Capability

		Ethernet Services		
		E-LINE	**E-LAN**	**E-TREE**
L2-bearing technologies	VxLAN	YES	YES	No current standards
	VPLS	YES	YES	YES
	MACinMAC	YES	YES	YES
	TRILL	YES	YES	No current standards
	NVGRE	YES	YES	No current standards

IP services can be natively carried by L3VPN and L3 Infrastructure technologies, which are described in clause 5. Services could also be carried non-natively. If required, IP services could also be carried by L2-bearing technologies, or L2 traffic could be encapsulated and carried by L3-bearing technologies. Note that carrying L2 traffic over L3 bearing infrastructure will require more layers of encapsulation than if L2 traffic is carried natively by L2-bearing technologies across the NFVI.

The rest of clause 8.1 describes the various ways of implementing interworking function.

8.1.1 Interworking Using a Gateway

In the Interworking Gateway case, as illustrated in figure 26, there is an Interworking Gateway (VNF or PNF) between the technology regions. This interworking gateway transforms packets between the type A and B technologies. The transform may operate in the data plane, control plane and (or) management plane. The Interworking Gateway (IG) shall be provisioned along with the other end points (vSwitches, vRouters orTORS) whenever a VN is created or deleted. An IG may then automatically transform any packets that are part of the same VN.

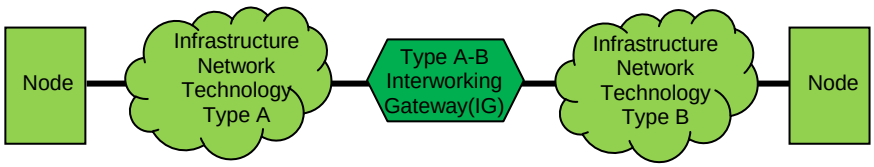

Figure 26: Dedicated Interworking Gateway (IG)

Typically, an IG is placed between the regions of the infrastructure where different technologies exist. An IG may be placed within the NFVI, depending on the boundaries of homogeneous protocol regions. Three common cases can be considered:

- The NFVI has a single protocol within each site, however each site has a different protocols;
- The NFVI has a single protocol within each VN, however different protocols in different VNs; and
- The NFVI has multiple protocols within each VN.

8.1.2 Interworking Using Multi-Protocol Tunnel Terminations

In the multiple protocol case, as illustrated in figure 27, there is no need for an IG. Instead, each edge switch or router (vSwitch, TOR or vRouter) is capable of decoding and (or) encoding multiple protocols, as well as on-ramping/off-ramping from/to the native protocol to the overlay protocol(s). In this case, the infrastructure network technology type can be completely mixed within each VN. This model may be able to operate using a single core switch protocol or may also require that the core switches support multiple protocols. Whether the core needs to be multi-protocol is dependent on whether the edge protocols are all carried over a single sub-protocol. For instance VxLAN, NVGRE and L3VPNs may all be carried over either an Ethernet or IP core. In this example core switches don't need to know the difference between the two protocols since the core is operating at the L2/L3 layers.

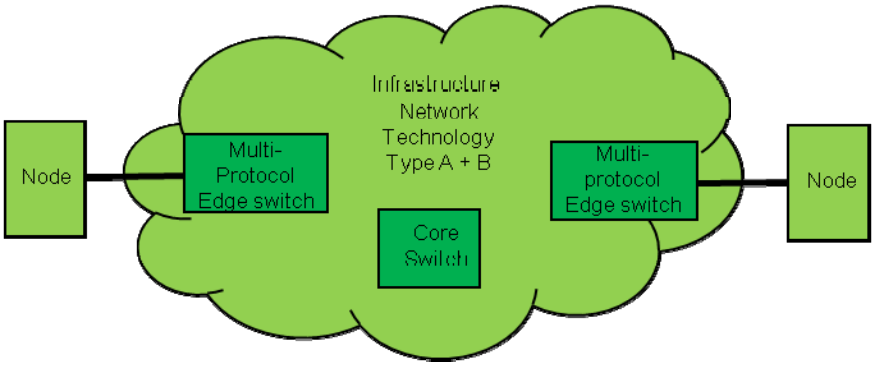

Figure 27: Multi-Protocol Edge Switches

8.1.3 Interworking in the VNFCI

In the case where the VNFCIs themselves perform the interworking function, as seen in figure 24, the VNFCI itself connects to the appropriate network topology or topologies in the Infrastructure. An example might be a VNFCI that communicates mainly over the L3-infrastructure service, but maintains a private network connection to a service-specific management VNFCI.

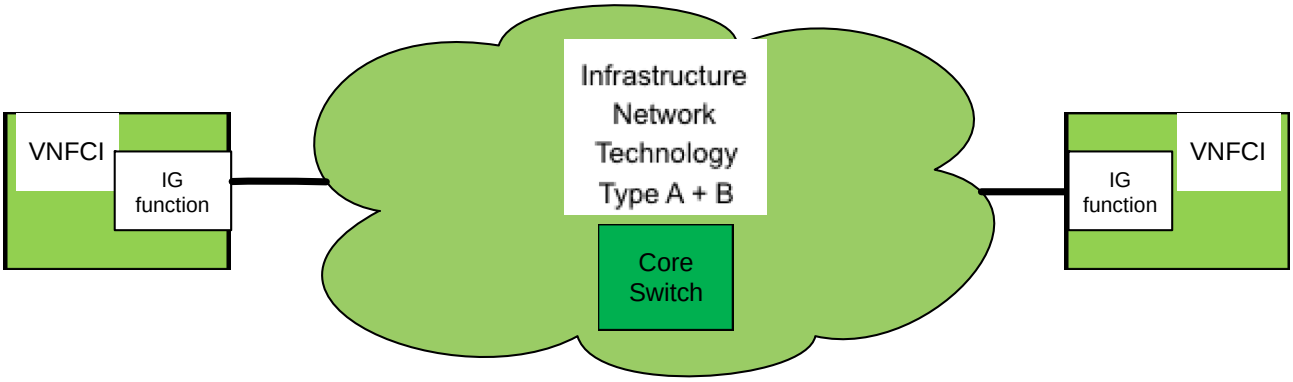

Figure 28: Example of interworking Gateway Function embedded in a VNFCI

8.2 Operations, Administration and Management Interworking

While end-to-end (service layer) OAM will detect failure between VNFCI nodes or between VNFCI nodes and foreign nodes or external end points, they will not, necessarily, identify the failure point within the infrastructure if overlay networking model(s) are used. The L3-infrastructure case, where the service and infrastructure networks are congruent, should be able to use service layer OAM to detect the actual point of failure.

In the case of overlay networks, if the service layer OAM detects a fault, then OAM shall be used on the underlying infrastructure to detect where the fault has actually occurred. The industry has substantial experience with OAM in interworking environments. The same resources already identified in clauses 5.1.1.1.1 and 5.1.1.2.3 will be useful in interworking infrastructure OAM.

8.3 Control Plane Interworking

Interworking control planes may be performed either by an Interworking Function or by supporting multiple control planes in a NFVI network domain with each control plane domain operating as a separate network from the other control domains. The interworking of various control planes is a large body of work that is both well understood in the service provider community, and larger in scope than NFV (e.g. SPBM and STP). Leveraging of the established interworking capabilities already developed and deployed is encouraged.

8.4 Data Plane Interworking

Interworking at the data plane may be performed either by an Interworking Gateway (IG) as depicted in figure 26 or by using a multi-protocol edge switch as depicted in figure 27. In the case of a multi-protocol edge switch, there is no need to perform transformations between encapsulations since the network edge (vSwitch or TORS) simply de-encapsulates whatever type it receives. This therefore concentrates on the transforms required by an IG. Table 4 lists the encapsulation types which may be deployed in the infrastructure network to carry the virtual network service traffic. Ethernet transport mechanisms, such as TRILL and VPLS can provide infrastructure connectivity for any Ethernet or IP based overlays, while L3 routed infrastructure (such as the L3-infrastructure approach above) can provide infrastructure connectivity for IP based overlays (such as L3-VPNs, L3-Infrastructure, GRE, NVGRE, VXLAN, etc.).

The interworking of various L2 and L3 technologies is a large body of work that is both well understood in the service provider community, and larger in scope than NFV. Leveraging of the established interworking capabilities already developed and deployed is encouraged.

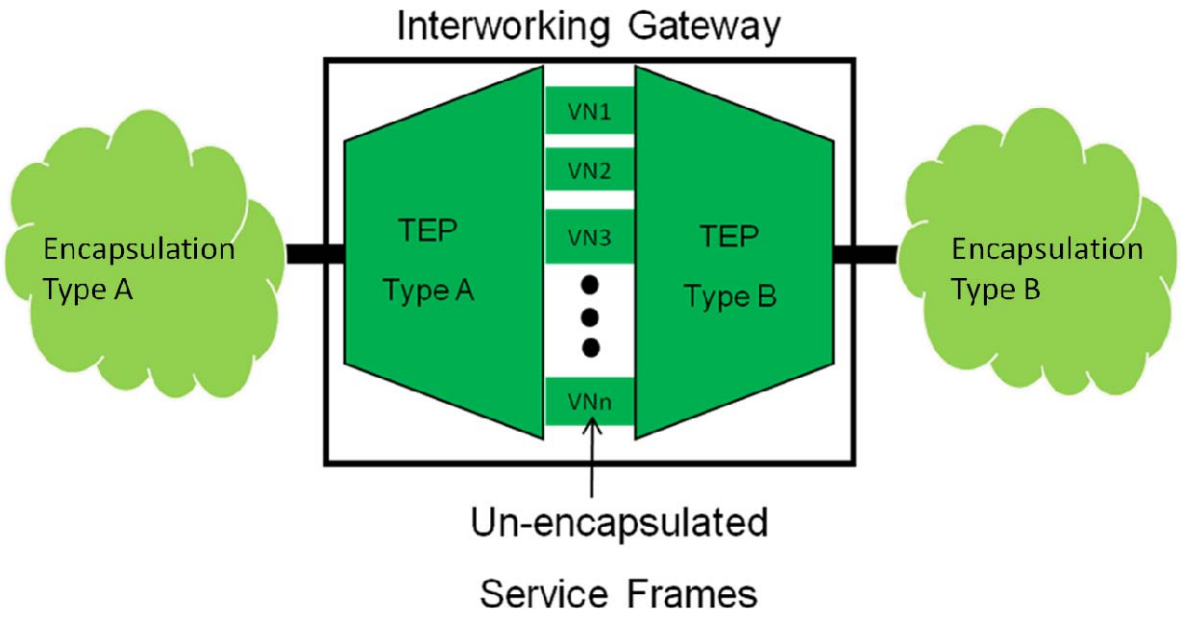

Figure 29: Generalized Interworking Gateway Architecture

A generalized IG relies on the fact that there is a single consistent packet type (IP packet and or Ethernet frame) that is carried to all NFs that attach to a VN. The receiving side of the gateway un-encapsulates each packet or frame, then passes the un-encapsulated service packet or frame over the (logical) interface associated with the specific service network to be re-encapsulated using the new encapsulation.

An example of a generalized IG is a VNF that contains two vSwitches or vRouters, one performing encapsulation/de-encapsulation for encapsulation type A and the other performing encapsulation/de-encapsulation for encapsulation type B. The two vSwitches or vRouters interface to each other via the service network specific interfaces, which act as the relay for frames between the two vRouters/vSwitches. As these vRouter/vSwitches are just more instances of the existing vRouter/vSwitch infrastructure, their OAM support is similar, and uses the same infrastructure. Care shall be taken, however, that security policies are not accidently breached in the shared vRouter/vSwitch. Because the generalized IG completely removes the encapsulation, it is possible to make a transform while preserving the VNID or to translate the VNIDs over the IG. Such a gateway works for any encapsulation type capable of carrying the infrastructure connectivity service traffic.

9 Features of the Domain Affecting Management and Orchestration

The features of the domain affecting management and orchestration are described in clause 5.2.

10 Features of the Domain Affecting Performance

10.1 Dynamic Optimization of Packet Flow Routing

Dynamic optimization of packet flow routing is a mechanism that enables VNFs to optimize packet flow routing and minimize infrastructure resource usage by transferring some of the traffic processing functions from the VNF to an infrastructure network element, including those within the infrastructure network domain (switch/router), the hypervisor domain (vSwitch/vRouter) and the compute domain (eSwitch, eRouter). The main benefits are considerable reduction of the HW dedicated to the VNF, higher throughput and shorter latency.

The dynamic optimization is based on the observation that in most of the network functions (security, NAT, load balancing, etc.) only a small fraction of a flow requires network intensive processing while the remaining part of the flow requires limited processing. By routing only a subset of the packet flow to the VNF for network intensive functions, it decreases the amount of VNF resources necessary and improves performance, making the virtualisation of some network intensive NF possible.

Network intensive applications are applications that involve some static and relatively small CPU code where the data processing is dynamic and typically implemented by embedded systems and network processors.

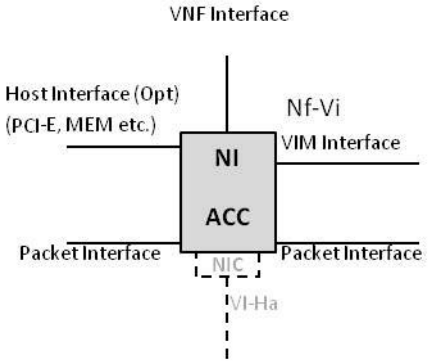

Figure 30: Network Intensive Acceleration

Interfaces to NI acceleration include:

- VIM Interface for acceleration resource management.
- Interface to the VNF (for programming).
- Data packet interface ports.
- Optional Host Interface (for management).
- Optional: NIC Interface to the Infrastructure Network.

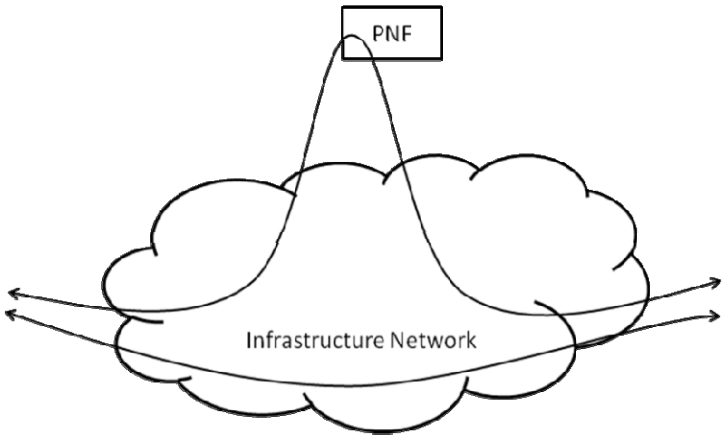

Figure 31: PNF deployment

In today's networks, every new flow that reaches the network (see figure 31) is forwarded to the relevant NFs where it is processed and forwarded to the next NF or an endpoint. In today's implementations, all the packets in the flow are processed by the NF without distinction, creating an unnecessary load on the NF.

In contrast, dynamic optimization of packet flow routing requires an optimization interface (see figure 32) that enables the VNF to optimize packet flow routing for the network intensive NF to the infrastructure network component. It requires the VNF to instruct the infrastructure network to stop forwarding the flow to the VNF and process it directly. The network processing functionality that may be performed by the infrastructure network components includes:

- Forwarding.

- Source IP/Destination IP/Source Port/Destination Port translation.
- TCP sequence number adaptation.
- Performance statistics.
- Qos.
- Etc.

The dynamic optimization of packet flow routing can be requested for the remaining life of the flow or for a given number of bytes. This enables the VNF to request rerouting of packet flows through VNF for a given duration. For example, in order to measure an HTTP server response time, the VNF may process the request and the response HTTP header while instructing the control module to reroute the flow to the network for the duration of the HTTP response object.

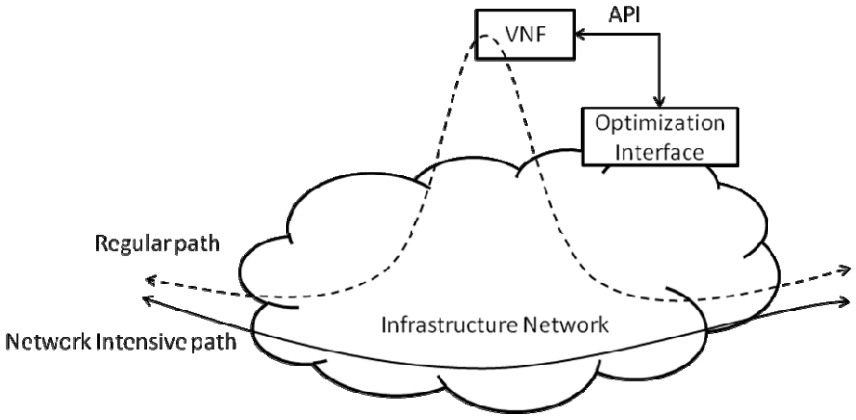

Figure 32: Rerouting the traffic from VNF to the Infrastructure Network

As shown in figure 33, the optimization interface uses the network controller northbound interface to redirect a given flow and the controller calculates the route and updates the flow table of the infrastructure network device along the calculated route. Similarly, the VNF optimization interface can send requests to the infrastructure network controller, which will re-route certain flows to the VNF for further performance evaluation or simply instruct the controller to reroute the flow to the VNF periodically for a limited sampling of the flow. In order to provide dynamic optimization of packet flow routing, the infrastructure network functionality, including its network controller southbound and northbound interfaces may need to be enhanced in order to increase the functionality applied on the fast path such as TCP sequence number adaption, traffic statistics, QoS, etc.

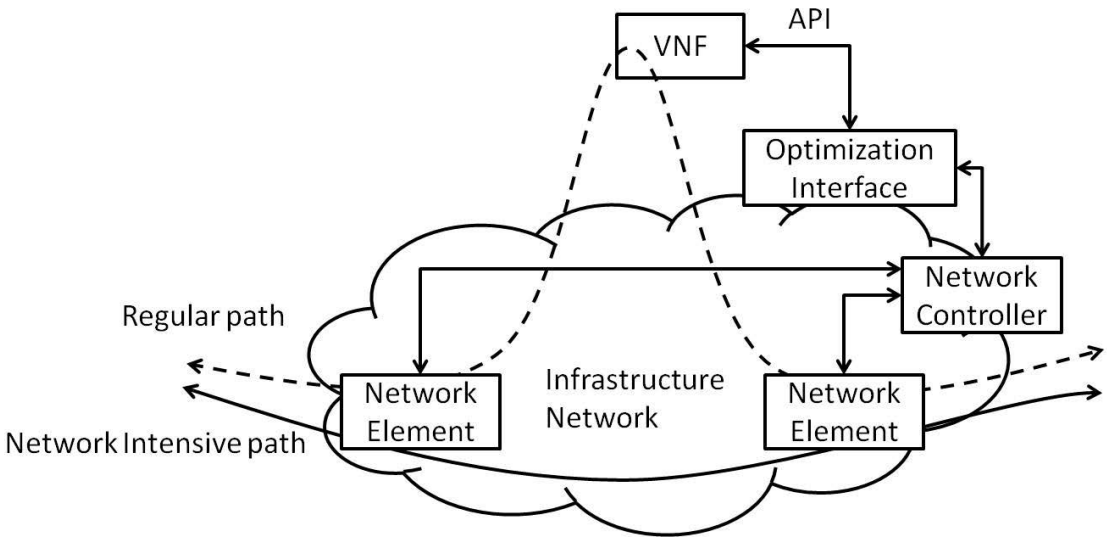

Figure 33: Rerouting flows from VNF to the Infrastructure Network

In addition, this dynamic optimization of packet flow routing mechanism requires the following from the infrastructure network:

- It shall be possible to discover the optimization capabilities of the infrastructure network.
- It shall be possible to establish a virtual network between VNFCIs of a VNF and the infrastructure network elements that provide the offloading function for that VNF. It shall also be possible to create a forwarding graph that ensures that all traffic for the VNF goes through the network elements providing the network processing functions.
- The network elements and associated network controllers that provide the dynamic optimization of the packet flow routing shall provide a mechanism to virtualise the underlying resources to guarantee that the rerouting or packet processing actions initiated by one VNF do not impact the datapath of other VNFs or other infrastructure network functions.

11 Features of the Domain Affecting Reliability

The NFV Infrastructure uses virtual networks to connect Network Function Component Instances (NFCIs).

When parallel paths through the Infrastructure Network have physical diversity, they improve the reliability of communications between the path end-points. Physically diverse paths avoid single points where failure would simultaneously disable all parallel paths. Thus, desirable features of the domain are to offer multiple physically diverse paths, and the ability to respond to virtual network connectivity requests with connectivity that achieves the desired degree of diversity.

When multiple parallel paths are deployed to satisfy connectivity requests, each path will carry traffic that consumes the overall capacity of the Infrastructure resources traversed. Some control or accounting shall be in place to ensure the capacity is not exhausted along the paths, in combination with network design.

In summary, higher reliability can be achieved through both diversity and capacity planning, in addition to other methods involving the selection of individual resources based on their specified failure rates.

12 Features of the Domain Affecting Security

The NFV Infrastructure uses virtual networks to connect Network Function Component Instances (NFCIs). The Virtual Networks maintained by the infrastructure are used to establish a trust relationship between the VN and attached NFCIs. In many cases the ability to access a particular Virtual Network is the first level of security. The type of trust relationship formed by a virtual network is dependent on the type of Virtual Network (e.g. E-LINE, E-LAN or E-TREE), the technology used to implement the virtual network (e.g. VxLAN, L2VPN, SPBM), the specific infrastructure implementation, the specific NFCIs attached, the authentication used for coupling the infrastructure components, etc. For E-LINE and E-LAN type VNs all parties are peers with equal access to each other. For E-TREE type VNs a client server relationship is formed where the root NFCIs can access all NFCIs on the E-TREE while the leaf NFCIs can only access the root NFCIs. An E-TREE enforces a partition between leafs.

Virtual networks are created by the virtualisation layer using either virtual partitioning (e.g. VLAN, VTN, MSTP), virtual overlays (e.g. VxLAN, NVGRE), or a combination of both (e.g. PBB, SPBM). In the case of virtual partitioning, the infrastructure is partitioned under control of a Virtual Infrastructure Manager (VIM). In the case of virtual overlays, the infrastructure does not enforce partitioning; instead, frames are tagged with a virtual network identifier which is enforced at a perimeter composed of Network Virtualisation Edges (NVE). The perimeter is configured under the control of a VIM. In the combined overlay and partitioning cases both a virtual network identifier (i.e. I-SID) and an infrastructure partition are used to enforce the virtual network. L3 Infrastructure-based networks are qualitatively different from those based on virtual partitioning or virtual overlays, both of which involve explicit tagging of packets to identify the VN with which they are associated. The boundaries of L3 Infrastructure-based networks are instead defined by Security Group policies which are configured under the control of a VIM. Each of these partitioning schemes has different vulnerabilities which may be exploited by attackers.

12.1 Threats for Virtual Partitions (VLANs, L2VPN, etc.)

Virtual partition attacks may be classified as: attacks from outside the partition (outside attacks), attacks on the partitioning infrastructure or the Virtual Infrastructure Manager (inside attacks), and attacks from compromised NFCIs systems (VNF attacks).

For virtual partitions, adversaries performing the first type of attack are referred to as outsiders or outside attackers since adversaries do not have to obtain any privilege on the infrastructure supporting the partition or NFCI systems in advance in order to perform this type attack. The second type of attackers are called insiders or inside attackers because they need to get certain privileges in changing the configuration or software of the infrastructure network devices beforehand and initiate the attacks within the infrastructure network. In the third type of attack, an attacker has got certain privileges in changing the configuration or software of an NFCI systems (e.g. hypervisors or virtual machines) and attempts to manipulate the controlled NFCI system to interfere with the normal operations of the VNs.

Virtual partitions have limited vulnerabilities to outside attacks since outsiders can not inspect or inject packets within a Virtual Network Partition from the outside. It is sometimes possible for an attacker to mount a denial of service attack from the outside of a partition if the partition does not protect the infrastructure bandwidth. It may also analyse traffic patterns based on indirect information.

12.2 Threats for Virtual Overlays (VxLAN, NVGRE)

Virtual overlay attacks may be fall into three categories of attacks. These are: attacks from underlying networks (outside attacks), attacks from compromised network virtualisation edge devices (inside attacks), and attacks from compromised NFCIs, guest OS or hypervisors (VNF attacks). This is inline with NVO3 classification [i.38].

For virtual overlays, adversaries performing the first type of attack are referred to as outsiders or outside attackers since adversaries do not have to obtain any privilege on the network virtualisation edge devices or NFCI systems in advance in order to perform this type attack. The second type of attackers are called insiders or inside attackers because they need to get certain privileges in changing the configuration or software of network virtualisation edge devices beforehand and initiate the attacks within the overlay security perimeter. In the third type of attack, an attacker has got certain privileges in changing the configuration or software of an NFCI systems (e.g. hypervisors or virtual machines) and attempts to manipulate the controlled NFCI system to interfere with the normal operations of the virtual overlay.

Virtual overlays are particularly vulnerable to outside attacks, especially when the virtual overlay is extended over a public internet (figure 34). Since the virtual overlay does not enforce a partition between the NVEs which form the security perimeter and other devices on the L3 network, any unsecured device attached to the network can inject or inspect packets on any VN as long as it can determine the VN identifier and the encapsulating IP addresses. In addition, it may be possible for devices attached to disrupt the network or analyse the traffic patterns in the network. When a virtual overlay is extended over a public internet, these attacks may come from anywhere on the public internet. When a virtual overlay is operated on a private internet, outside attacks shall come from unsecured devices connected to the private internet.

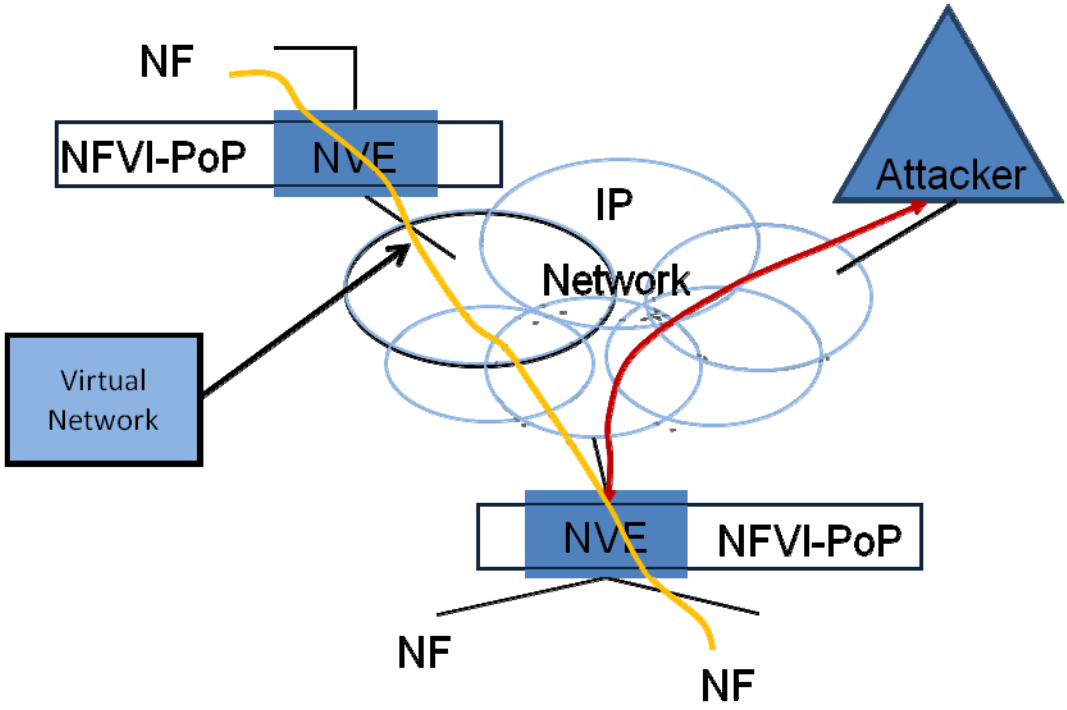

Figure 34: Outside Attacker on a VxLAN/NVGRE Virtual Network

Some of the specific vulnerabilities for virtual overlays are:

1) A variety of man-in-the-middle attacks are possible, providing the opportunity to change the contents of the packets (and associated headers where relevant).

2) Spoofing attacks (related to the above), where malicious packets can be inserted into a stream.

3) Mapping attacks, where an attacker can snoop packets and gain information about the topology of the source and/or destination networks of the packets.

As such, these overlay protocols are not sufficient for many deployments, and tunnels or other cryptographic techniques may be required.

12.3 Threats for Combined Partition/Overlay (PBB, SPBM)

Combined partition/overlay attacks may be classified as: attacks on the partitioning infrastructure (outside attacks); attacks on compromised Backbone Edge Bridge (BEB) devices or the VIM (inside attacks), and attacks on compromised NFCIs systems (NFCI attacks).

For combined partition/overlay, adversaries performing the first type of attack are referred to as outsiders or outside attackers since adversaries do not have to obtain any privilege on the BEB devices or NFCI systems in advance in order to perform this type attack. The second type of attackers are called insiders or inside attackers because they need to get certain privileges in changing the configuration or software of BEB devices beforehand and initiate the attacks within the overlay security perimeter. In the third type of attack, an attacker has gotten certain privileges in changing the configuration or software of an NFCI systems (e.g. hypervisors or virtual machines) and attempts to manipulate the controlled NFCI system to interfere with the normal operations of the VNs.

Combined partition/overlays are more difficult to attack than virtual overlays since attackers can't inspect or inject packets within the partition protecting the BEB perimeter. It is sometimes possible for an attacker to mount a denial of service attack from the outside of a partition if the partition doesn't protect the infrastructure bandwidth.

12.4 Security Model for L3 Infrastructure-based Networks

12.4.1 Security Mechanisms in an L3 Infrastructure-based Network

With VLAN and L2VPN, partitioning takes the form of closed groups of VNFCIs. A collection of VNFCIs that is connected to a given VLAN or L2VPN may communicate freely with one another using any protocol they choose, but may not communicate with any entity outside the VLAN or L2 VPN without transiting through some kind of gateway function.

By contrast, L3 Infrastructure-based virtual networks implement partitioning using Security Groups that apply policies both to the exchange of traffic both within the virtual network, and with entities outside of the virtual network. It is possible to configure Security Groups so as to emulate the "closed groups" capability of VLANs, by defining Security Group policies that permit communications only between and among the set of IP addresses assigned to the VNFCIs within the virtual network. If configured in this way, L3 Infrastructure-based virtual networks exhibit similar characteristics to partitioned virtual networks such as VLANs as it relates to security threats.

However, Security Group policies can be applied at the level of individual VNFCIs, providing more granular control than is possible with VLANs or L2 VPNs. Security Group policies in L3 Infrastructure-based networks can be made more restrictive than VLANs or L2 VPNs, for example by preventing certain VNFCIs from sending traffic to certain other VNFCIs assigned to the same virtual network, or by restricting traffic based on port number or protocol type. In principle, it should be possible to take advantage of such Security Group policies to make L3 Infrastructure-based virtual networks more resistant to security threats than is possible with, for example, VLANs.

Likewise, it is possible to configure Security Group policies that are more permissive, for example allowing certain VNFCIs to communicate directly with an arbitrarily large set of IP addresses that lie outside the boundaries of the virtual network. Such permissive Security Group policies shall be used with care. While there are obvious benefits in enabling VNFCIs to exchange traffic with entities on external networks directly, without the need to transit through a specialized edge function of the type that is always required with partitioned or overlay-based virtual networks, this does open up any VNFCI that has been so configured to various kinds of attack. Security Group policies of this type should therefore only be used with VNFCIs that are explicitly hardened against such attacks or designed to provide some security function in their own right, for example a firewall or a Session Border Controller.

12.4.2 Threat Model in an L3 Infrastructure-based Network

The attack surface in an L3 Infrastructure-based model is broadly centred around three main attack vectors:

- Subversion or bypass of policy control.
- Opposition control of underlying transport.
- (D)DoS of policy control points or forwarding functions.

Subversion or bypass of policy control points can be attempted by *forging* packet header data, such as source-address spoofing, which can be countered by the use of ingress filters and uRPF checking on ingress to the policy controlled network. It can also be attempted through the exploit of *bugs* in the policy enforcement or policy management frameworks. These are corrected by keeping policy as simple as possible, and keeping current on patches and other updates. Regular auditing of policies is also a good practice.

The opposition may attempt to gain access to elements that are directly connected to the transit infrastructure in an L3-infrastructure (or any other network technology) network. If this is accomplished, many forms of attack can be launched, including traffic capture, traffic pattern analysis, and forged traffic injection. The defence against such attacks is the securing and auditing of elements directly connected to the transit infrastructure.

(D)DoS attacks can be mitigated by rate and state limiting of traffic to reasonable levels, such that the volume of attack traffic is kept below the level where it is disruptive, and the source of such traffic can be traced and terminated.

12.5 Security Model for L3 VPN-based Networks

The security model for L3-VPN-based networks, is very similar to the L3-infrastructure based network model, with the addition of partitioning of the VNs by L3-VPN labels. While this can provide yet another level of protection (just as any partitioning may), it complicates policy enforcement and auditing (again, just as in any overlay network) by introducing multiple policy enforcement modes and paths (*i.e.* traffic within the VPN has a different treatment than traffic that has an end-point outside the VPN). There is also a further attack surface, which is an attempt to forge, re-write, or otherwise subvert the VPN ID. However, L3-VPN networks are more secure against such attacks, when compared to other VPN technologies, in that the VPN membership data is transmitted via control plane traffic, rather than just data plane traffic.

Annex A (informative): Authors & contributors

The following people have contributed to the present document:

Rapporteur:

Mrs Evelyne Roch, Huawei Technologies

Other contributors:

Mr Dave Allan, Ericsson
Mr Paul Bottorff, HP
Mr Zvika Bronstein, Huawei
Ms Deborah Brungard, AT&T
Mr Trevor Cooper, Intel
Mr Hiroshi Dempo, NEC
Mr Hans-Jorg Kolbe, NEC
Mr Ramki Krishnan, Brocade
Mr Dirk Kutscher, NEC
Mr Christopher Liljenstolpe, Metaswitch
Mr Al Morton, AT&T
Mr Jurgen Quittek, NEC
Dr Joseph Tardo, Broadcom
Mr Dan Touitou, Huawei Technologies
Mr Ben Wright, Metaswitch
Mr Steven Wright, AT&T
Mr Abraham Yang, ZTE Corporation

History

Document history		
V1.1.1	December 2014	Publication