ETSI GS NFV-INF 004 V1.1.1 (2015-01)

Network Functions Virtualisation (NFV); Infrastructure; Hypervisor 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-INF004

Keywords
interface, 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

Important notice

The present document can be downloaded from:
http://www.etsi.org

The present document may be made available in electronic versions and/or in print. The content of any electronic and/or print versions of the present document shall not be modified without the prior written authorization of ETSI. In case of any existing or perceived difference in contents between such versions and/or in print, the only prevailing document is the print of the Portable Document Format (PDF) version kept on a specific network drive within ETSI Secretariat.

Users of the present document should be aware that the document may be subject to revision or change of status. Information on the current status of this and other ETSI documents is available at
http://portal.etsi.org/tb/status/status.asp

If you find errors in the present document, please send your comment to one of the following services:
http://portal.etsi.org/chaircor/ETSI_support.asp

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 Documents	Document #
Overview	GS NFV INF 001
Architecture of Compute Domain	GS NFV INF 003
Architecture of Hypervisor Domain	GS NFV INF 004
Architecture of Infrastructure Network Domain	GS NFV INF 005
Scalability	GS NFV INF 006
Interfaces and Abstraction	GS NFV INF 007
Test Access	GS NFV INF 009

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 the architecture of the Hypervisor Domain of the NFV Infrastructure which supports deployment and execution of virtual appliances. The present document will primarily focus on the use of hypervisor for virtualisation, due to time and resource constraints, However, the hypervisor requirements are similar if not the same for implementing linux containers or other methods for virtualisation.

NOTE: From WikiArch: "Linux Containers (LXC) are an operating system-level virtualisation method for running multiple isolated server installs (containers) on a single control host. LXC does not provide a virtual machine, but rather provides a virtual environment that has its own process and network space. It is similar to a chroot, but offers much more isolation".

There needs to be further research w.r.t to Linux Containers, including developing the ecosystem.

As well as presenting a general overview description of the NFV Infrastructure, the present document sets the NFV infrastructure and all the documents which describe it in the context of all the documents of the NFV. It also describes how the documents which describe the NFV infrastructure relate to each other.

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-INF 001 (V1.1.1): "Network Functions Virtualisation (NFV); Infrastructure Overview".

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 004: "Network Functions Virtualisation (NFV); Virtualisation Requirements".

[i.2] IETF RFC 4133: "Entity MIB (Version 3)".

[i.3] IEEE 802.1D™: "IEEE Standard for Local and Metropolitan Area Networks -- Media access control (MAC) Bridges".

[i.4] IEEE 802.1Q™ MIB: "IEEE Standard for Local and Metropolitan Area Networks, Management Information Base".

[i.5] IETF draft-ietf-opsawg-vmm-mib-00: "Management Information Base for Virtual Machines Controlled by a Hypervisor".

NOTE: Available at http://tools.ietf.org/html/draft-ietf-opsawg-vmm-mib-00.

[i.6] IEEE 1588™: "Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems".

[i.7] IEEE 802.11™: "Wireless LANS IEEE Standard for Information technology - Telecommunications and information exchange between systems Local and metropolitan area networks - Specific requirements Part 11: Wireless LAN".

[i.8] IEEE 802.3ad™: "Link Aggregation".

[i.9] IEEE 802.3™ MIB: "Link Aggregation, Management Information Base".

[i.10] Hotlink: http://www.virtualizationpractice.com/hotlink-supervisor-vcenter-for-hyper-v-kvm-and-xenserver-15369/.

[i.11] Systems Management Architecture for Server Hardware (SMASH).

NOTE: Available at http://www.dmtf.org/standards/smash.

[i.12] NFVINF(13)VM_019_Data_plane_performance.

[i.13] http://pubs.vmware.com/vsphere-51/index.jsp?topic=%2Fcom.vmware.vsphere.networking.doc%2FGUID-E8E8D7B2-FE67-4B4F-921F-C3D6D7223869.html

[i.14] http://msdn.microsoft.com/en-gb/library/windows/hardware/hh440249(v=vs.85).aspx

[i.15] http://www.vcritical.com/2013/01/sr-iov-and-vmware-vmotion/

[i.16] ETSI GS NFV-INF 010: "Network Functions Virtualisation (NFV); Service Quality Metrics".

3 Definitions and abbreviations

3.1 Definitions

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

application VMs: VM not utilizing an OS

hypervisor: virtualisation environment running on a host

NOTE: The virtualisation environment includes the tools, BIOS, firmware, Operating Systems (OS) and drivers.

portability: See ETSI GS NFV-INF 001 [1].

standard: is de-jure, de-facto or open standard that fulfils the requirement

NOTE: This assumption should be applied to all requirements through the entire document.

3.2 Abbreviations

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

API	Application Programming Interface
BIOS	Basic Input Output System

BSD	Berkeley Software Distribution
CIM	Centralized Interference Mitigation
CLI	Comand Line Interface
CMS	Call Management System
CPU	Compute Processing Unit
DMA	Direct Memory Access
Dpdk	Data plane development kit
EE	Electrical Engineering
FFS	For Further Specification
GUI	Graphical User Interface
HA	High Availability
IETF	Internet Engineering Task Force
INF	Infrastructure
IO	Input Output
IP	Internet Protocol
IPMI	Intelligent Platform Interface
JVM	Java Virtual Machine
KVM	Kernel Virtual Machine
LAN	Local Area Networks
LLC	Lower Level Cache
LLDP	Link Layer Discovery Protocol
LXC	Linux Containers
MAC	Media Access Controller
MANO	Management and Orchastration
MIB	Management Information Base
NF	Network Function
NFVi	Network Function Virtualisation Infrastructure
NFVINF	Network Function Virtualisation Infrastructure
NIC	Network Interface Card
NOC	Network Operator Council
NUMA	Non Uniform Memory Access
OAM	Operations and Maintenance
OS	Operating System
OSS	Operations Systems and Software
OVF	Open Virtual Framework
PAE	Physical Address Extension
PCI	Peripheral Component Interconnect
RAID	Redundant Array of Independent Disks
RAM	Random Access Memory
RAS	Row Address Strobe
RELAV	Reliability and Resiliency Work Group
RFC	Request For Comment
SCSI	Small Computer System Interface
SDO	Standards Development Organizations
SEC	Security Working Group
SLA	Service Level Agreement
SNMP	Signalling Network Management Protocol
SR-IOV	Single Root I/O Virtualisation
SWA	Software Architecture Work group
TCP	Transport Control Protocol
TLB	Translation Lookaside Buffer
UDP	User Datagram Protocol
UUID	Universally Unique Identifier
VIM	Virtualisation Infrastructure Manager
VM	Virtual Machine
VN	Virtual network
VNF	Virtual Network Function
VNFC	Virtual Network Function Component
VNFI	Virtual Network Function Interface
VSCSI	Virtual Small Computer System Interface
vSwitch	virtual Switch
VT	Virtualisation

WG	Working Group
XAPI	eXtended Application Programming Interface

4 Domain Overview

Popek and Goldberg paper 'Formal Requirements for Third Generation Architectures': set the definition of hypervisors in 1974.

- Equivalence: the hypervisor provides an environment for programs which is essentially identical to the original machine.
- Resource control: the hypervisor is in complete control of system resources.
- Efficiency: programs run on this (virtualised) environment show at worst only minor decreases in speed.

Equivalence

The environment provided by a hypervisor is functionally equivalent to the original machine environment. This implies that the same operating systems, tools and application software can be used in the virtual environment. This does not preclude para-virtualisation and other optimization techniques which may require operating systems, tools and application changes.

Resource Control

The hypervisor domain mediates the resources of the computer domain to the virtual machines of the software appliances. Hypervisors as developed for public and enterprise cloud requirements place great value on the abstraction they provide from the actual hardware such that they can achieve very high levels of portability of virtual machines.

In essence, the hypervisor can emulate every piece of the hardware platform even in some cases, completely emulating a CPU instruction set such that the VM believes it is running on a completely different CPU architecture from the actual CPU on which it is running. Such emulation, however, has a significant performance cost. The number of actual CPU cycles needed to emulate virtual CPU cycle can be large.

Efficiency

Even when not emulating a complete hardware architecture, there can still be aspects of emulation which cause a significant performance hit. Typically, computer architectures provide means to offload these aspects to hardware, as so called virtualisation extensions, the set of operations that are offloaded and how they are offloaded varies between different hardware architectures and hypervisors as innovation improves virtualisation performance.

EXAMPLE: Intel VT and ARM virtualisation extensions minimise the performance impact of virtualisation by offloading to hardware certain frequently performed operations.

There can be many virtual machines running on the same host machine. The VMs on the same host may want to communicate between each other and there will be a need to switch between the VMs.

Infrastructure Domains

Figure 1 illustrates the four domains of the NFV architecture, their relationship with each other and their relationship to other domains outside the infrastructure. The figure also sets out the primary interfaces. Hypervisor for the present document entails tools, kernel, host.

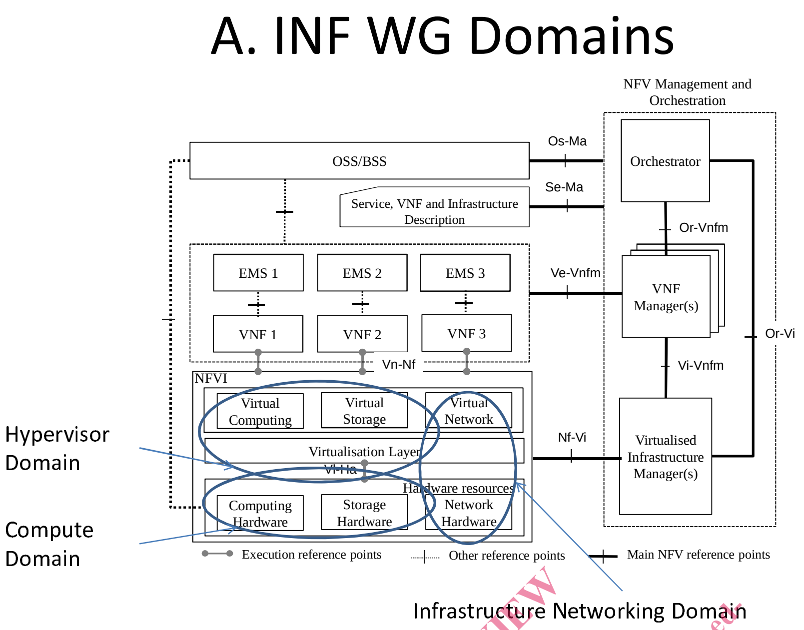

Figure 1: General Domain Architecture and Associated Interfaces

The NFV Infrastructure (NFVI) architecture is primarily concerned with describing the Compute, Hypervisor and Infrastructure domains, and their associated interfaces.

The present document is primarily focused on describing the hypervisor domain, which comprise the hypervisor which:

- provides sufficient abstract of the hardware to provide portability of software appliances;
- allocates the compute domain resources to the software appliance virtual machines;
- provides a management interface to the orchestration and management system which allows for the loading andmonitoring of virtual machines.

Figure 2 depicts the NFV reference architectural framework. Table 1 gives description and definition to the interfaces.

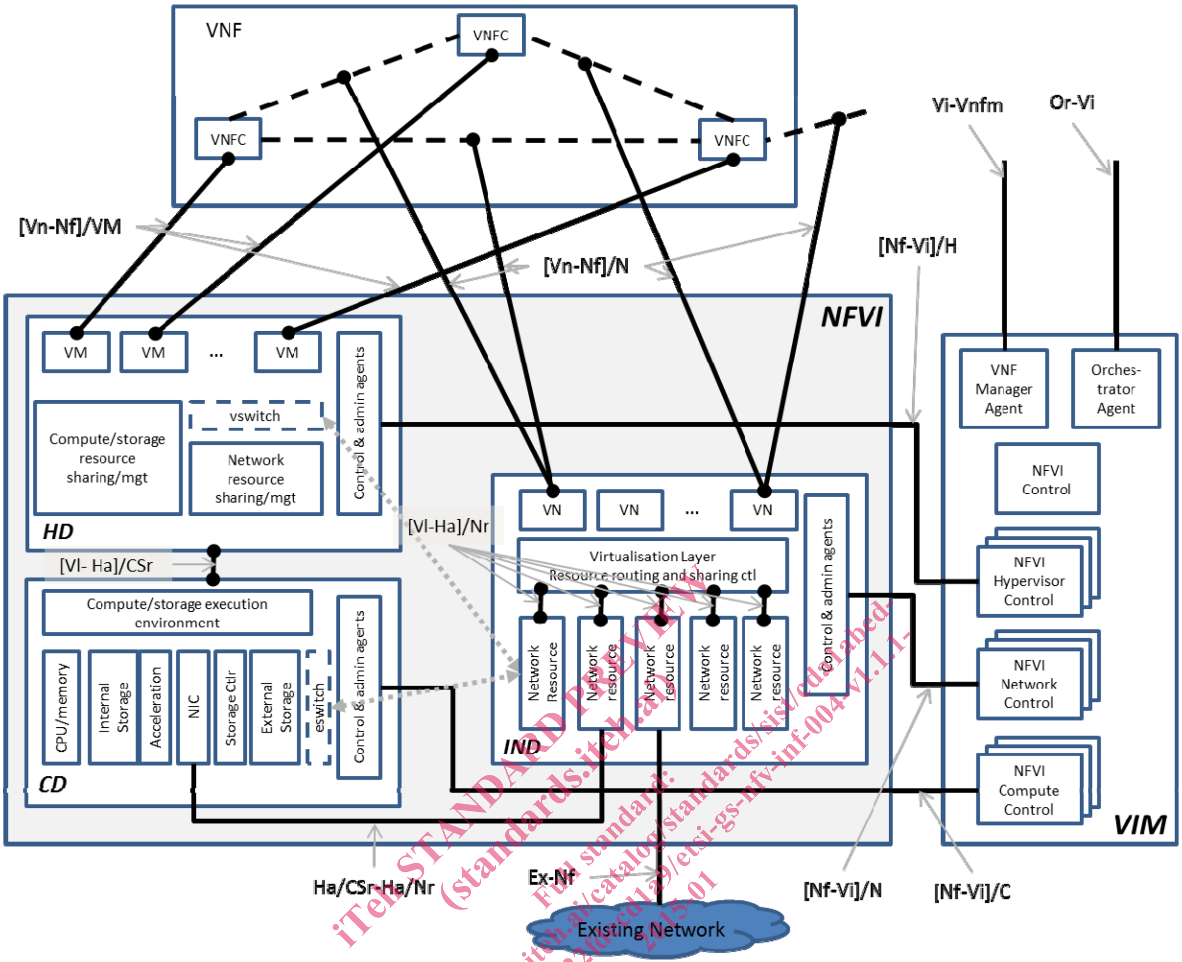

Figure 2: NFV Reference Architectural Framework

Table 1

INF Context	NFV Framework Reference Point	INF Reference Point	Reference Point Type	Description and Comment
External	Vn-Nf	[Vn-Nf]/VM	Execution Environment	This reference point is the virtual machine (VM) container interface which is the execution environment of a single VNFC instance.
		[Vn-Nf]/VN	Execution Environment	This reference point is the virtual network (VN) container interface (eg an E-Line or E-LAN) which 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).
	Nf-Vi	[Nf-Vi]/N	Management, and Orchestration Interface	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.
		[Nf-Vi]/H	Management, and Orchestration Interface	This is the reference point between the management and orchestration agents in hypervisor 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 hypervisor domain.
		[Nf-Vi]/C	Management, and Orchestration Interface	This is the reference point between the management and orchestration agents in compute 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 compute domain.
	Vi-Vnfm		Management, Interface	This is the reference point that allows the VNF Manager to request and/or for the VIM to report the characteristics, availability, and status of infrastructure resources.
	Or-Vi		Orchestration Interface	This is the reference point that allows the Orchestrator to request resources and VNF instantiations and for the VIM to report the characteristics, availability, and status of infrastructure resources.
		Ex-Nf	Traffic Interface	This is the reference point between the infrastructure network domain and any existing and/or non-virtualised network. This reference point also carries an implicit reference point between VNFs and any existing and/or non-virtualised network.
Internal	Vl-Ha	[Vl-Ha]/CSr	Execution Environment	The framework architecture (see figure 2, NFV Reference Architectural Framework) shows a general reference point between the infrastructure 'hardware' and the virtualisation layer. This reference point is the aspect of this framework reference point presented to hypervisors by the servers and storage of the compute domain. It is the execution environment of the server/storage.
		[Vl-Ha]/Nr	Execution Environment	The framework architecture (see figure 2, NFV Reference Architectural Framework) shows a general reference point between the infrastructure 'hardware' and the virtualisation layer. While the infrastructure network has 'hardware', it is often the case that networks are already layered (and therefore virtualised) and that the exact choice of network layering may vary without a direct impact on NFV. The infrastructure architecture treats this aspect of the Vi-Ha reference point as internal to the infrastructure network domain.
		Ha/CSr-Ha/Nr	Traffic Interface	This is the reference point between the infrastructure network domain and the servers/storage of the compute domain.

The present document focuses on the hypervisor domain. Figures 3 to 5 will depict how the hypervisor domain inter works within the Infrastructure (INF).