# ETSI GS NFV-EVE 001 V3.1.1 (2017-07)


# Network Functions Virtualisation (NFV);
# Virtualisation Technologies;
# Hypervisor Domain Requirements specification;
# Release 3

***Disclaimer***

The present 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-EVE001

Keywords
network, NFV, virtualisation

***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

## Essential patents

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 (https://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.

## Trademarks

The present document may include trademarks and/or tradenames which are asserted and/or registered by their owners. ETSI claims no ownership of these except for any which are indicated as being the property of ETSI, and conveys no right to use or reproduce any trademark and/or tradename. Mention of those trademarks in the present document does not constitute an endorsement by ETSI of products, services or organizations associated with those trademarks.

# Foreword

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

# Modal verbs terminology

In the present document "**shall**", "**shall not**", "**should**", "**should not**", "**may**", "**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 provides requirements for the hypervisor domain as it pertains to an operator's network. It focuses on gaps between Network Functions Virtualisation (NFV) use cases and the industry state of art at the time of publication. Therefore requirements that are deemed to be supported by most hypervisor solutions at the time of publication are not repeated in the present document.

# 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 reference 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: "Network Functions Virtualisation (NFV); Terminology for Main Concepts in NFV".

[2] ETSI GS NFV-IFA 002: "Network Functions Virtualisation (NFV); Acceleration Technologies; VNF interfaces specification".

[3] ETSI GS NFV-IFA 004: "Network Functions Virtualisation (NFV); Acceleration Technologies; Management aspects specification".

[4] ETSI GS NFV-IFA 018: "Network Functions Virtualisation (NFV); Acceleration Technologies; Network Acceleration Interface Specification".

[5] ETSI GS NFV-IFA 019: "Network Functions Virtualisation (NFV); Acceleration Technologies; Acceleration Resource Management Interface Specification".

[6] ETSI GS NFV-SEC 012: "Network Functions Virtualisation (NFV); Security; System architecture specification for execution of sensitive NFV components".

[7] Unified Extensible Firmware Interface Forum: "Advanced Configuration and Power Interface Specification (ACPI)".

NOTE: Available at http://www.uefi.org/specifications.

[8] ETSI GS NFV-TST 008: "Network Functions Virtualisation (NFV); Testing; NFVI Compute and Network Metrics Specification".

## 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 004: "Network Functions Virtualisation (NFV); Infrastructure; Hypervisor Domain".

# 3 Definitions and abbreviations

## 3.1 Definitions

For the purposes of the present document, the terms and definitions given in ETSI GS NFV 003 [1] and the following apply:

**hypervisor:** piece of software which partitions the underlying physical resources and creates Virtual Machines, and isolates the VMs from each other

NOTE: The hypervisor is a piece of software running either directly on top of the hardware (bare metal hypervisor) or running on top of a hosting operating system (hosted hypervisor). The abstraction of resources comprises all those entities inside a computer/server which are accessible, like processor, memory/storage, NICs.

**hypervisor domain:** general area for focus which includes hypervisors

NOTE: Other elements such as a vswitch can be implemented in the hypervisor domain.

## 3.2 Abbreviations

For the purposes of the present document, the abbreviations given in ETSI GS NFV 003 [1] and the following apply:

| | |
|---|---|
| ACPI | Advanced Configuration and Power Interface |
| CD | Compute Domain |
| DOPFR | Dynamic Optimization of Packet Flow Routing |
| EPD | Extended Para-virtualised Device |
| eSwitch | embedded Switch |
| HD | Hypervisor Domain |
| IND | Infrastructure Networking Domain |
| INF | Infrastructure |
| MWAIT | Monitor Wait |

NOTE: x86 Instruction Set Reference.

| | |
|---|---|
| NFVI | NFV Infrastructure |
| NIC | Network Interface Controllers |
| RDMA | Remote Direct Memory Access |
| RDTSC | Read Time-Stamp Counter |

NOTE: x86 Instruction Set Reference.

RDTSCP Read Time-Stamp Counter and Processor identifier

NOTE: x86 Instruction Set Reference.

| | |
|---|---|
| VIM | Virtualisation Infrastructure Manager |
| VM | Virtual Machine |
| VN | Virtual network |
| VNFC | Virtualised Network Function Component |
| vRouter | virtual Router |
| vSwitch | virtual Switch |

# 4 Hypervisor Domain Overview

Figure 4-1 describes the logical placement of the hypervisor domain, which is circled in dotted lines, within the NFV architectural framework. The hypervisor domain abstracts the hardware resources from the compute domain. The hypervisor domain overlaps with the network domain in that it can include virtual switch (vSwitch) and virtual router (vRouter) software elements and abstract network-related hardware resources provided by the compute domain such as network interface controllers (NIC) and embedded switches (eSwitch).

There is a direct reference point to the Virtualisation Infrastructure Manager (VIM) from the hypervisor domain.

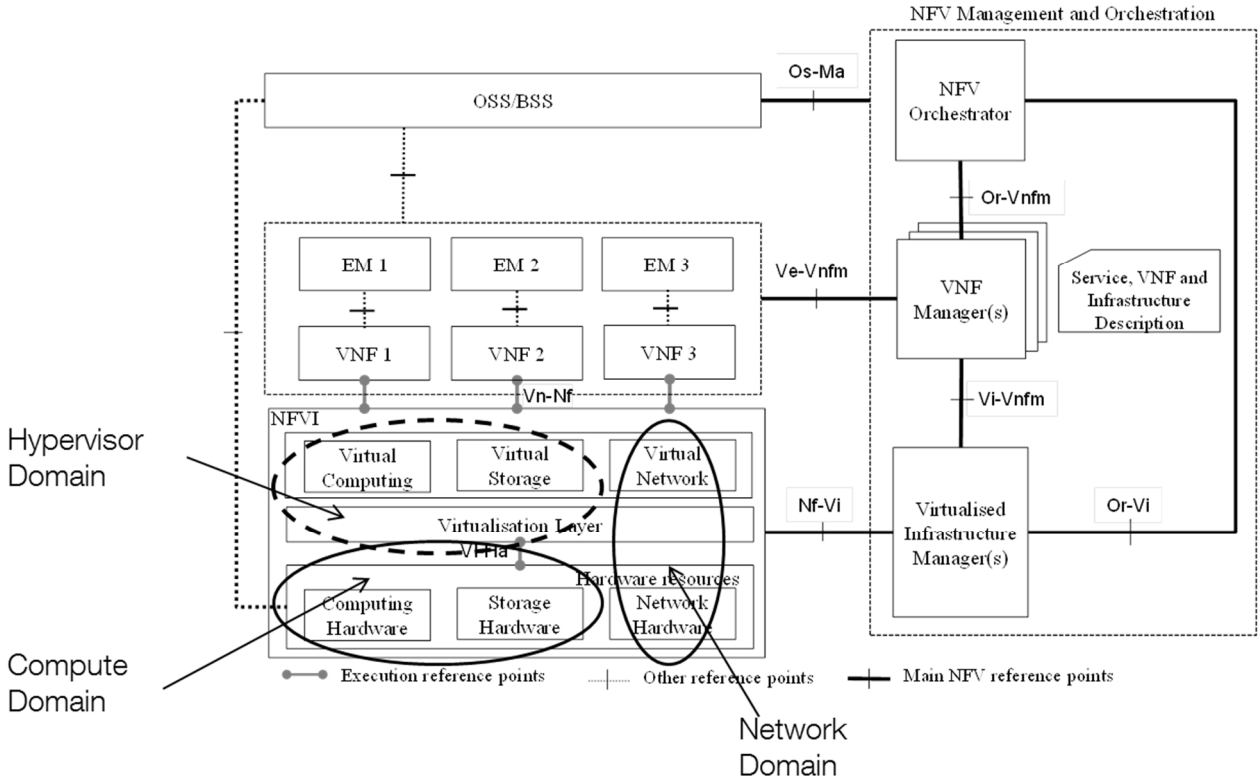


**Figure 4-1: Logical placement of the hypervisor domain**

The hypervisor domain itself is a software environment which abstracts hardware and implements services, such as starting a Virtual Machine (VM), terminating a VM, acting on policies, scaling, live migration, and high availability.

Figure 4-2 provides a more detailed view of the hypervisor domain and its reference points to the other elements of the architectural framework.

The following reference points are identified:

a) The [Nf-Vi]/H reference point provides the interfaces to the VIM. This is where the requests for hypervisor services occur. Only the VIM shall interact with the hypervisor through these interfaces. Hypervisors shall not implement services autonomously unless within the context of the VIM applied policy.

b) The [Vi-Ha]/Csr reference point provides the interfaces that the hypervisor pulls hardware information from and uses to create virtual hardware components which the VM utilizes.

c) The [Vn-Nf]/VM reference point represents the execution environment provided by the NFVI to a VNF. A VNF is created essentially via one or more VMs. A VM is in essence software running a function, algorithm, application without being aware of the type, model or number of actual physical units 'underneath' the function, algorithm and/or application.

The purpose of the other reference points identified in figure 4-2 are described in table 1 of ETSI GS NFV-INF 004 [i.1].

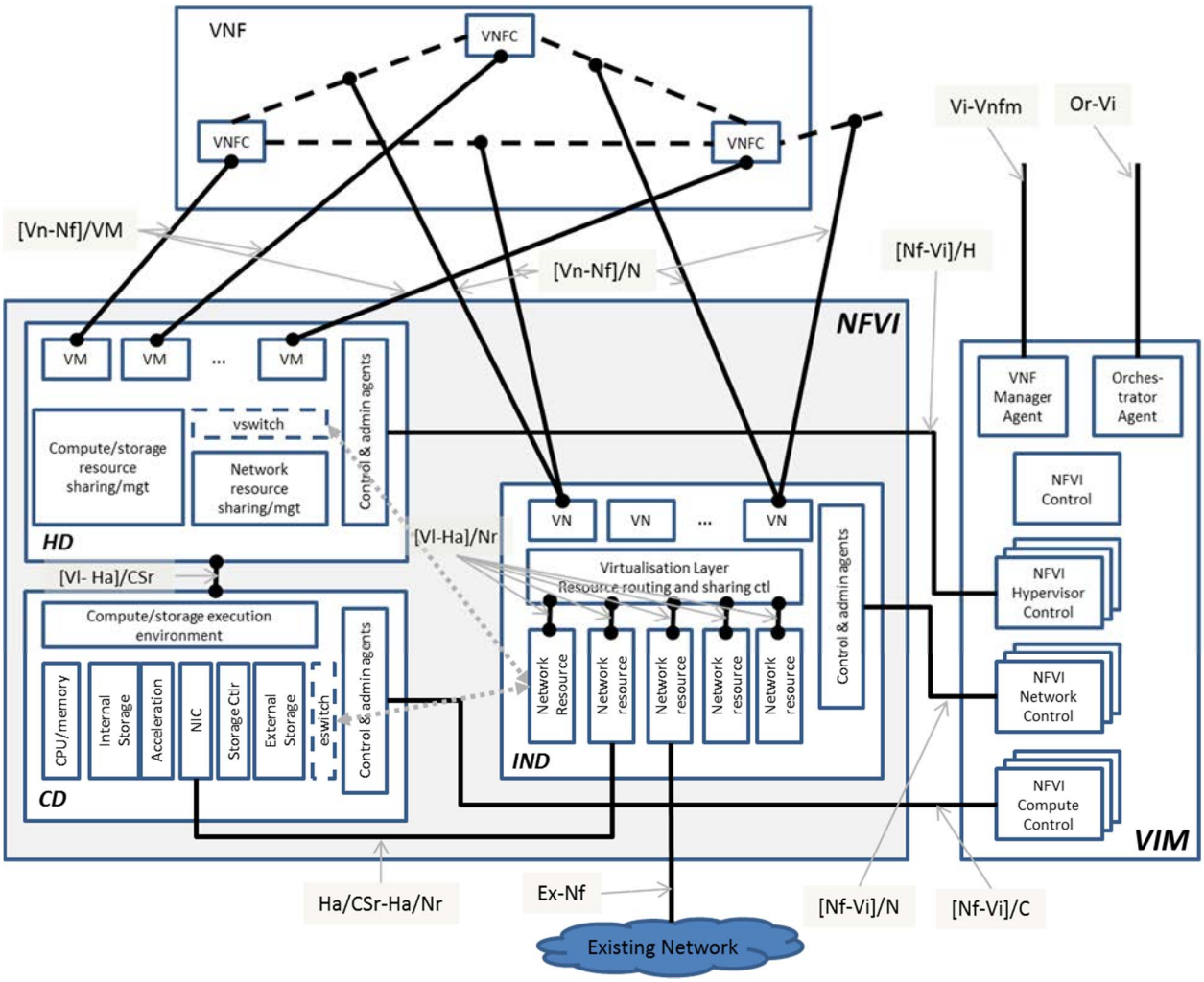


**Figure 4-2: Hypervisor domain reference points**

The general architecture of the hypervisor domain is shown in figure 4-3.

Figure 4-3 also illustrates how VMs interact with CPU cores, NICs and vSwitches. It highlights that high performance VMs can be given:

- exclusive access to specific CPU cores;
- exclusive access to allocated memory regions;
- direct access the physical NICs using direct memory mapped polled drivers and user mode instructions requiring no 'context switching';
- direct memory mapped polled drivers for interVM communications using user mode instructions requiring no 'context switching'.

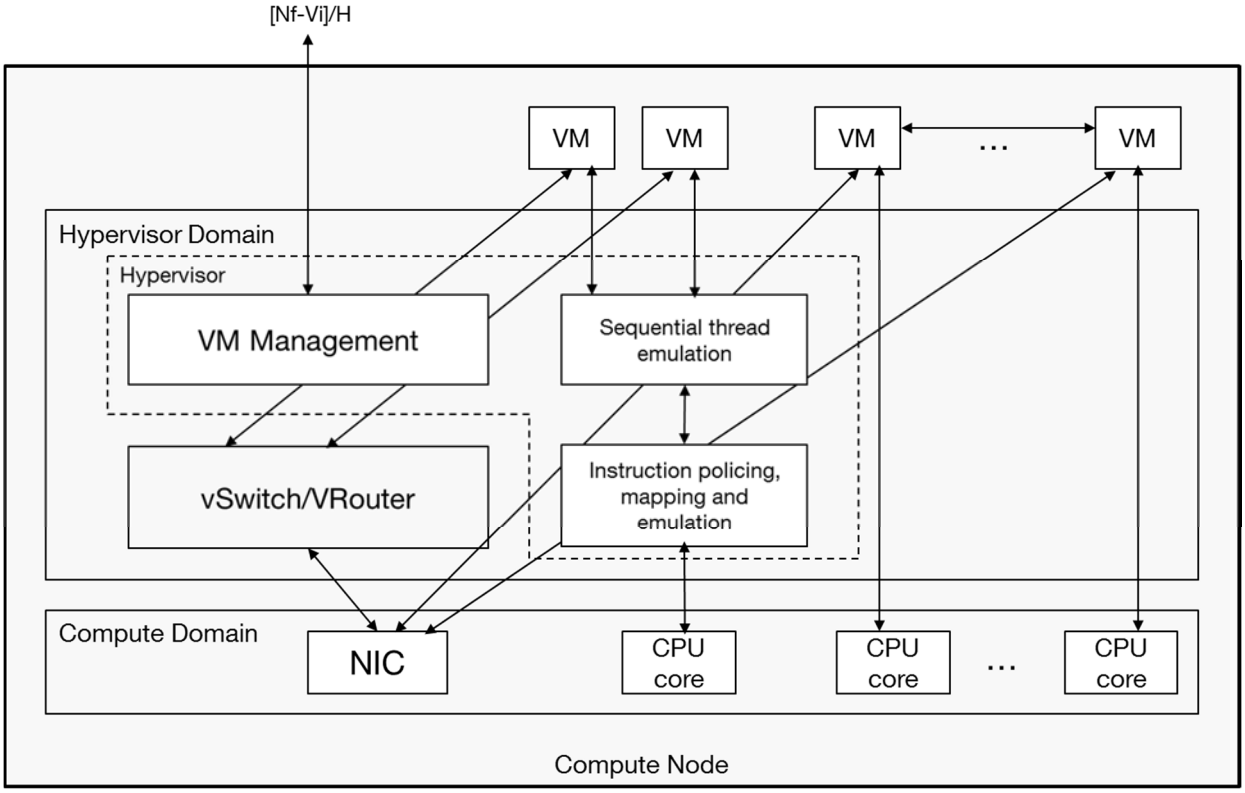


**Figure 4-3: Example hypervisor architecture**

# 5 Requirements of the Hypervisor Domain

## 5.1 Introduction

Some of the requirements on the hypervisor domain are implicitly placing requirements on the underlying hardware. If the hardware does not fulfil such requirements, then the corresponding hypervisor requirements cannot be fulfilled.

## 5.2 General Requirements

Table 5.2-1 specifies general requirements applicable to the hypervisor domain.

**Table 5.2-1: general requirements**

| Numbering | Requirement description |
|---|---|
| REQ.HYP.GEN.001 | The hypervisor shall support partitioning of the resources of a compute node. |
| REQ.HYP.GEN.002 | The hypervisor should support nested virtualisation. |
| REQ.HYP.GEN.003 | The hypervisor shall support acceleration requirements as specified in ETSI GS NFV-IFA 002 [2]. |
| REQ.HYP.GEN.004 | The hypervisor should support partitioning of the resources. |

## 5.3 Real Time Guest Support

Table 5.3-1 specifies requirements for supporting real time VM execution

**Table 5.3-1: Real time VM execution**

| Numbering | Requirement description |
|---|---|
| REQ.HYP.RT.001 | The hypervisor domain shall be able to support memory bandwidth allocation & control. |
| REQ.HYP.RT.002 | The hypervisor domain shall support non virtualised timing (e.g. rdtsc/rdtscp instructions). |
| REQ.HYP.RT.003 | The hypervisor domain shall support non virtualised synchronization primitives (e.g. monitor/mwait instructions). |
| REQ.HYP.RT.004 | The hypervisor domain should enable the VIM to configure last level cache size allocation. |
| REQ.HYP.RT.005 | The hypervisor domain should enable the VIM to configure cache bandwidth allocation & control. |

## 5.4 vSwitch Resiliency

Table 5.4-1 specifies requirements for increasing vSwitch resiliency.

**Table 5.4-1: vSwitch resiliency requirements**

| | |
|---|---|
| REQ.HYP.VSWR.01 | The hypervisor domain shall support a deployment option where the vswitch functionality is deployed independently from the served hypervisor. |
| REQ.HYP.VSWR.02 | The hypervisor domain shall support a deployment option where the vswitch functionality is hosted in a VM.<br>See note. |
| NOTE: The type of VM hosting this functionality is often referred to as a service VM, a control VM or a stub domain in the industry. | |

## 5.5 NFV Acceleration Support

Table 5.5-1 specifies requirements for enabling fast communication between VNFC instances.

**Table 5.5-1: High-speed communication requirements**

| | |
|---|---|
| REQ.HYP.ACCFC.001 | The hypervisor domain shall support a deployment option where VMs are directly connected with each other via their vNICs (i.e. point-to-point communication), not using any vSwitch, vRouter or eSwitch. |
| REQ.HYP.ACCFC.002 | The hypervisor domain shall support a deployment option where VMs on the same compute node can communicate with each other by sharing memory directly. |
| REQ.HYP.ACCFC.003 | The hypervisor domain shall support a deployment option where VMs can communicate with each other, through distributed memory technology when they reside on different compute nodes (e.g. using RDMA as a cluster technology). |
| REQ.HYP.ACCFC.004 | The hypervisor domain shall support a deployment option where VMs can communicate directly with each other through a serial bus. |

Table 5.5-2 specifies requirements for supporting abstraction of acceleration capabilities.

**Table 5.5-2: Acceleration interfaces requirements**

| | |
|---|---|
| REQ.HYP.ACCIF.001 | The hypervisor domain shall provide the Extended Para-virtualised Device (EPD) backend functionality specified in ETSI GS NFV-IFA 002 [2]. |
| REQ.HYP.ACCIF.002 | The hypervisor domain should provide the EPD backend functionality for all interfaces specified in ETSI GS NFV-IFA 002 [2]. |
| REQ.HYP.ACCIF.003 | The hypervisor domain shall support a deployment option where dedicated switches can be assigned to VNFC instances to enable Dynamic Optimization of Packet Flow Routing (DOPFR) as per ETSI GS NFV-IFA 018 [4]. |

Table 5.5-3 specifies requirements for managing acceleration capabilities.

**Table 5.5-3: Acceleration management requirements**

| | |
|---|---|
| REQ.HYP.ACCMGT.001 | The hypervisor domain shall provide the Acceleration Resource Discovery interface specified in ETSI GS NFV-IFA 019 [5]. |
| REQ.HYP.ACCMGT.002 | The hypervisor domain shall provide the Acceleration Resource Lifecycle Management interface specified in ETSI GS NFV-IFA 019 [5]. |
| REQ.HYP.ACCMGT.003 | The hypervisor domain shall provide the Acceleration Enabler capability as specified in ETSI GS NFV-IFA 004 [3]. |

## 5.6 Security

The hypervisor shall be considered a sensitive component as defined in ETSI GS NFV-SEC 012 [6].

## 5.7 Energy Efficiency requirements

Table 5.7-1 specifies requirements for enabling energy efficiency in NFV deployments.

**Table 5.7-1: Energy Efficiency requirements**

| | |
|---|---|
| REQ.HYP.EE.001 | The hypervisor shall be configurable with resource utilization thresholds leading to automatic changes of the CPUs' power state. |
| REQ.HYP.EE.002 | The hypervisor shall expose the Advanced Configuration and Power Interface (ACPI) as specified in [7] to the VMs. |
| REQ.HYP.EE.003 | The hypervisor shall be able to apply power management policies provided by the VIM. |
| REQ.HYP.EE.004 | The hypervisor shall be configurable with the time required for a compute resource, to return to a normal operating mode after leaving a specific power-saving mode. |

## 5.8 Performance management requirements

Table 5.8-1 specifies requirements for enabling NFVI performance data to be reported to the VIM.

**Table 5.8-1: Performance management requirements**

| | |
|---|---|
| REQ.HYP.PM.001 | The hypervisor domain shall be able to notify the VIM performance metrics specified in ETSI GS NFV-TST 008 [8]. |
| REQ.HYP.PM.002 | The hypervisor domain shall be able to provide the VIM with performance metrics specified in ETSI GS NFV-TST 008 [8] in response to a query. |
| NOTE: Compute and memory metrics are expected to be reported by the hypervisor. Network metrics are expected to be reported by the vSwitch or vRouter. | |

# Annex A (informative): Authors & contributors

The following people have contributed to this specification:

**Rapporteur:**
Valerie Young, Intel
Bruno Chatras, Orange

**Other contributors:**
François-Frédéric Ozog, 6WIND
Shaoji Ni, Huawei

# Annex B (informative): Bibliography


- ETSI GS NFV-INF 001: "Network Functions Virtualisation (NFV); Infrastructure Overview".
- ETSI GS NFV-INF 003: "Network Functions Virtualisation (NFV); Compute Domain".
- ETSI GS NFV-INF 005: "Network Functions Virtualisation (NFV); Infrastructure; Network Domain".
- ETSI GS NFV-IFA 001: "Network Functions Virtualisation (NFV); Acceleration Technologies; Report on Acceleration Technologies & Use Cases".
- ETSI GS NFV-SWA 001: "Network Functions Virtualisation (NFV); Virtual Network Functions Architecture".
- ETSI GS NFV-EVE 005: "Network Functions Virtualisation (NFV); Ecosystem; Report on SDN Usage in NFV Architectural Framework".

# Annex C (informative): Change History

| Date | Version | Information about changes |
|---|---|---|
| January 12, 2015 | V0.0.1 | ToC added |
| February 20, 2015 | V0.0.2 | Overview and Scope added |
| March 25, 2015 | V0.0.3 | Added NFV INF 004 approved components |
| January 20, 2016 | V0.0.4 | Added approved contributions EVE(15)000253r1 and 4 EVE(15)000409r1 |
| January 21, 2016 | V0.0.5 | Added history information, and disclosure on title page |
| February 11, 2016 | V0.0.6 | Scope update NFVEVE(16)000022 |
| March 22, 2016 | V0.0.7 | Editorial clean-up based on NFVEVE(16)000032 |
| May 5, 2016 | V0.0.8 | Output of NFV#14, incorporating NFVEVE(16)00068, NFVEVE(16)00069R1, NFVEVE(16)00070R1 |
| July 28, 2016 | V0.0.9 | Incorporating NFVEVE(16)000133 (editorial cleanup) agreed at EVE#40 |
| September 9, 2016 | V0.0.10 | Incorporating NFVEVE(16)134R1 and 131R1 agreed at EVE#43 |
| September 22, 2016 | V0.0.11 | Output of NFV#15, incorporating agreed NFVEVE(16)132R1 |
| December 19, 2016 | V0.1.0 | Output of NFV#15, incorporating agreed NFVEVE(16)222 |
| February 27, 2017 | V0.2.0 | Output of NFV#17, incorporating NFVEVE(17)000022 and 023 |
| March 24, 2017 | V0.2.1 | Output of the EVE WG Meeting #56, incorporating NFV(17)000054R2 |

# History

| Document history | | |
|---|---|---|
| V3.1.1 | July 2017 | Publication |
| | | |
| | | |
| | | |
| | | |