ETSI GS NFV-EVE 003 V1.1.1 (2016-01)

Network Functions Virtualisation (NFV); Ecosystem; Report on NFVI Node Physical Architecture Guidelines for Multi-Vendor Environment

Disclaimer

Reference
DGS/NFV-EVE003

Keywords
architecture, NFV, NFVI

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 (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.

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 guidelines for NFVI Node physical architecture. It is limited to the hardware resources - compute, storage, and network - needed to construct and support the functions of an NFVI Node. This includes physical components needed to house and interconnect nodes.

The present document also provides some examples on "building" specific NVFI Node configurations and addresses related issues such as reliability and energy efficiency.

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.

Not applicable.

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] IEEE 802.3ae™ Standard for Information technology: "Telecommunications and information exchange between systems - Local and metropolitan area networks, - Specific requirements Part 3: Carrier Sense Multiple Access with Collision Detection (CSMA/CD) Access Method and Physical Layer Specifications, Amendment 1: Media Access Control (MAC) Parameters, Physical Layers, and Management Parameters for 10 Gb/s Operation".

[i.3] Introducing data centre fabric, the next-generation Facebook™ data center network.

NOTE: Available at https://code.facebook.com/posts/360346274145943/introducing-data-center-fabric-the-next-generation-facebook-data-center-network/.

[i.4] ETSI EN 300 019-1-3: "Environmental Engineering (EE); Environmental conditions and environmental tests for telecommunications equipment; Part 1-3: Classification of environmental conditions; Stationary use at weatherprotected locations".

[i.5] ETSI EN 300 753: "Environmental Engineering (EE); Acoustic noise emitted by telecommunications equipment".

[i.6] NEBS GR-63: "NEBS Requirements: Physical Protection".

[i.7] ASHRAE: "Thermal Guidelines for Data Processing Environment", 3rd edition, 2012.

[i.8] ETSI GS NFV 002 (V1.2.1): "Network Functions Virtualisation (NFV); Architectural Framework".

[i.9] ETSI GS NFV-INF 003 (V1.1.1): "Network Functions Virtualisation (NFV); Infrastructure; Compute Domain".

[i.10] ETSI GS NFV-MAN 001 (V1.1.1): "Network Functions Virtualisation (NFV); Management and Orchestration".

[i.11] EIA/ECA-310, Revision E, December 1, 2015: "Electronic Components Industry Association (ECIA)".

[i.12] ETSI ETS 300 119-4: "Equipment Engineering (EE); European telecommunication standard for equipment practice; Part 4: Engineering requirements for subracks in miscellaneous racks and cabinets".

[i.13] ETSI GS NFV-REL 003 (V0.3.0) (08-2015): "Network Functions Virtualisation (NFV); Reliability; Report on Models and Features for E2E Reliability".

NOTE: Available at https://docbox.etsi.org/ISG/NFV/Open/Drafts/REL003_E2E_reliability_models/NFV-REL003v030.zip.

[i.14] Final Report: "Virtualised Mobile Network with Integrated DPI, ETSI ISG NFV, October 31st 2014".

[i.15] "Refactoring Telco Functions, the Opportunity for OCP in telco SDN and NFV Architecture", Tom Anschutz, March 9, 2015.

[i.16] IETF RFC 7075: "Diameter Base Protocol".

3 Definitions and abbreviations

3.1 Definitions

For the purposes of the present document, the terms and definitions given in ETSI GS NFV-INF 003 [i.9], ETSI GS NFV-INF 001 [i.1] apply.

3.2 Abbreviations

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

AC	Alternating Current
AES	Advanced Encryption Standard
AES-NI	Advanced Encryption Standard New Instructions
API	Application Program Interface
ASHRAE	American Society of Heating, Refrigerating, and Air-conditioning Engineers
CPE	Customer Premise Equipment
CPU	Central Processing Unit
CRC	Cyclic Redundancy Check
DC	Direct Current
DC-DC	Direct Current-Direct Current
DMA	Direct Memory Access
DSC	Diameter Signaling Controller
E2E	End to End
ECC	Error Correcting Code
EMI	Electromagnetic Interference
EPC	Evolved Packet Core
ER	Extended Reach
FRU	Field Replaceable Unit
FW	FirmWare

GPON	Gigabit Passive Optical Network
GR	Generic Requirements
HDD	Hard Disk Drive
ISG	Industry Specification Group
IT	Information Technology
JBOD	Just a Bunch Of Disks
KVM	Kernel-based Virtualisation Machine
LAN	Local Area Network
LR	Long Reach
LW	Long Wavelength
MANO	Management and Orchestration
MME	Mobility Management Entity
NAS	Network Attached Storage
NEBS	Network Equipment Building System
NFV	Network Functions Virtualisation
NFVI	NFV Infrastructure
NFVI-PoP	NFV Infrastructure Point of Presence
NUMA	Non-Uniform Memory Access
OCP	Open Compute Project
OLT	Optical Line Terminator
OS	Operating System
PCI	Peripheral Component Interconnect
PGW	Packet Data Network Gateway
SDDC	Single Device Data Correction
SGW	Serving Gateway
SONET	Synchronous Optical Networking
SR	Special Report
SR-IOV	Single Root Input/Output Virtualisation
SW	Short Wavelength
ToR	Top of Rack
UMA	Uniform Memory Access
UPS	Uninterruptible Power Supply
VDC	Volts of Direct Current
VLAN	Virtual Local Area Network
VM	Virtual Machine
VNF	Virtualised Network Function
WAN	Wide Area Network

4 Principles for development of physical components

4.1 Introduction

Virtualised Network Functions (VNFs) have to reside and operate on physical hardware. The telecommunications industry is moving away from specialized, sophisticated, and possibly proprietary hardware; instead the goal is to move towards commercially available off-the-shelf products in terms of processors, disks, racks, and other physical elements.

The goal of the present document is to provide guidance for an ecosystem of generic and commonly available sets of physical products and components for the industry.

The guidelines will be beneficial to telecommunication equipment providers and other vendors as well as service providers in the development and acquisition of desired components for building NFVI Nodes.

The focus of the present document is limited to the description of the physical hardware - compute, storage, and network domains shown in figure 1.

The present document draws upon available hardware principles (e.g. Open Compute Project) as necessary. Other topics covered include the following:

- Node functions (compute, storage, network);
- Physical components (racks, frames, processors, etc.);

- Power and cooling issues - guidelines on potential relationships between power delivery, heat build-up and heat dissipation will be provided as appropriate;
- Interconnection methods.

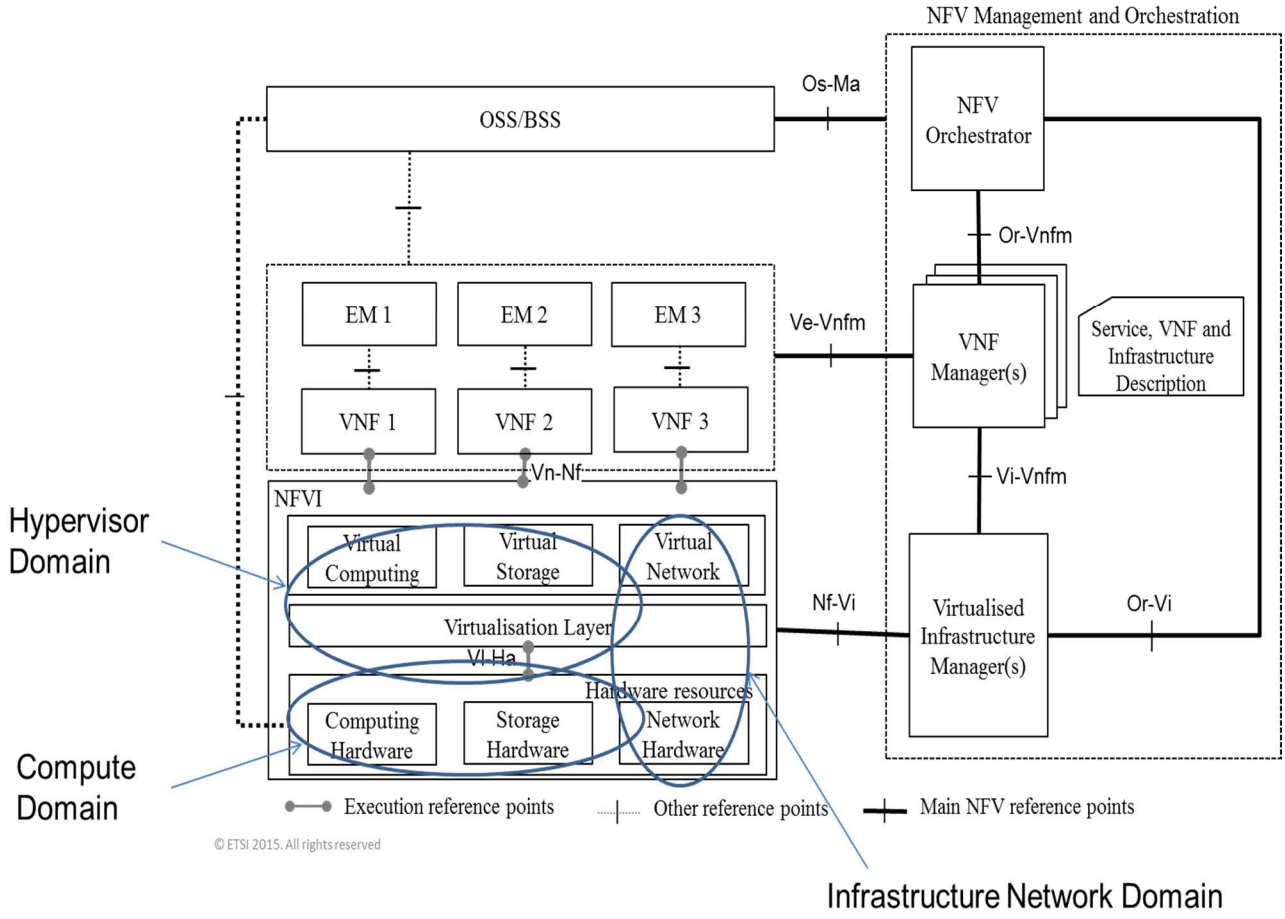

Figure 1: NFV architectural framework and identification of NFVI domains

Clause 4.2 provides general principles and goals for the NFVI Node physical architecture. Clause 4.3 provides key criteria for NFVI Nodes. Additional clauses outline common architectural practices with evaluation of their applicability to the NFVI Node physical architecture.

4.2 General principles

4.2.1 Observations

These general observations are included as guidance when considering architectural implementations for NFVI Nodes. Unlike software elements of NFV, NFVI has unique characteristics due to its physical nature. General observations include:

- The end goals for the platform dictate the architectural choices. Without alignment around goals, it is difficult to determine a cohesive architecture;
- Infrastructure, by nature, is physical. Infrastructure has size and weight. It consumes power and generates heat and noise. These are all interrelated and need to be balanced to meet equipment needs. Ambient temperature, power delivery and acoustic limits are derived from building and equipment practices;
- Infrastructure nodes live within the equipment practices in which they are deployed. They interface with other mechanical, electrical and software elements. These interfaces and behaviours need to be taken into account.

4.2.2 High level goals for NFVI Nodes

The following high level goals are desired for whatever NFVI Nodes are developed. These goals are consistent with a vision to enable a robust software ecosystem on commercially available hardware infrastructure.

High level goals for NFVI Nodes are:

- Multi-vendor, multi-customer, commercially available off-the-shelf ecosystem. Specific products do not need to be developed for each customer. Suppliers are free to innovate within the ecosystem;
- Economical scalability, addressing a wide range of application sizes and capacities;
- Appropriate features: solutions take into account concerns regarding space, power, cooling, and maintenance;
- Able to address multiple layers of the network: including transport, switching/routing and applications.

4.2.3 Other solution values

The following additional solution values are desired:

- Manageable. Application as well as field replaceable units may be managed. Management method integrates with NFV management interfaces;
- Resilient. Failure of components within the solution is detectable in order to support failover to another resource. Fail-over is handled in such a way to reduce the possibility of outages. Resiliency objectives (e.g. 5 nines) may be specified by service providers;
- Efficient (power, space, etc.);
- Interoperable. Equipment from one vendor interoperates with equipment from other vendors;
- Backward compatible. New equipment works with older equipment;
- Future proofed. Current-generation equipment works with future-generation equipment.

4.3 Key criteria

4.3.1 Space

Space criteria relate to the physical height, width, depth and volume that the equipment occupies. Space-related criteria are important to comprehend because NVFI Nodes will be deployed within facilities where space-related constraints exist. The following are key space-related criteria:

- Rack footprint and height compatible with data center and central office;
- Efficient use of the available volume;
- Flexible module widths, heights, etc.;
- Approximately 1 rack unit module height, or multiples thereof (allows use of commercially available off-the-shelf components).

4.3.2 Power

Power criteria are related to the type and amount of power that is supplied to the NFVI Nodes as well as limitations on how much power the equipment draws. Key criteria are:

- High power density within the constraints of climate and cooling scheme;
- Flexibility to interface to various direct current (including high voltage direct current) and alternating current configurations and topologies;
- Capability to support installations in both data centers and central offices, depending on configuration;

- Rack level uninterruptable power supply option with appropriate backup time;
- Maximum rack power consistent with facilities infrastructure practices;
- Maximum node power depends on node density. Full rack of nodes consume no more than the maximum rack power;
- Support power control at NFVI Node as well as individual compute/storage/network node levels;
- Avoidance of single Points of Failures at power system;
- Support power redundancy including N+M (M<N) and N+N.

4.3.3 Cooling

Cooling criteria are related to the capability to remove heat from the NFVI Node. Since dissipation of power by the NFVI Node generates heat, power consumption and cooling capabilities need to be matched. Key criteria are:

- Cooling matches the power needs in the central office and data center environments;
- Air filter option is desirable;
- Roadmap to support for liquid cooling;
- Front-to-back airflow is desirable;
- Placing temperature sensitive parts (e.g. optical module, HDD) at air intake position is desirable;
- Maintenance at cold aisle is desirable.

4.3.4 Physical interconnect

System interconnect criteria provide guidance on how elements of the NFV Node infrastructure are connected together. Key criteria are:

- Common interconnection methods (e.g. Ethernet) for all nodes;
- Capacity scalable to order of Terabits/s per rack unit;
- Modular capacity can grow as installation needs demand;
- Support high bandwidth and low latency switching to meet the resource pooling requirements;
- Support isolation of north-south data flow and east-west data flow;
- Support isolation of service data flow and management data flow.

4.3.5 Management

Infrastructure management applies to how each module or physical subcomponent (power supplies, fans, nodes) is managed. NFVI management fits within the overall framework of NFV management.

4.3.6 Climatic

Since NFVI equipment may be deployed in both central office and datacenter environment, compliance with the climatic standards of central office and datacenter environment (e.g. ETSI EN 300-019-1-3 [i.4], NEBS GR-63 [i.6], ASHRAE Thermal Guidelines for Data Processing Environment [i.7], etc.) is desirable.

4.3.7 Acoustic

For the hearing protection of employees working in high noise emission environment, compliance with noise emission standards (e.g. ETSI EN 300 753 [i.5], NEBS GR-63 [i.6]) is desired. For central office and datacenter deployment, the NFVI equipment adherent to acoustic emission limits is strongly desired and likely will be mandated by operators.

Capability of operating at the NEBS acoustic noise limits is recommended. Capability of operating at the ETSI acoustic noise limit might be needed based on operator demand.

4.4 Open Compute Project

The Open Compute Project is an online community whose mission is to create and deliver efficiently functioning servers, storage capabilities and data centre designs. Open Compute is one example of an open ecosystem of commercially available hardware that is being deployed today. While Open Compute was not architected specifically for application to NFVI Nodes, understanding of its key elements is instructive. Throughout the rest of the present document, Open Compute may be used as an architecture for illustrative purposes.

5 Overview of node functions

5.1 Introduction

As discussed in the ETSI NFV architectural framework [i.8], the physical hardware resources include computing, storage and network resources that provide processing, storage and connectivity to VNFs through the virtualisation layer (e.g. hypervisor).

5.2 Compute node

A compute node is an element used in the compute domain of the NFVI (see figure 1 in clause 4.1). The NFV architectural framework [i.8] states that the hardware resources for the compute node are assumed to be comprised of commercially available products as opposed to purpose-built hardware. The computing resources are commonly pooled.

5.3 Storage node

A storage node is an element used in the compute domain of the NFVI (see figure 1 in clause 4.1). The NFV architectural framework [i.8] states that the storage resources can be differentiated between shared Network Attached Storage (NAS) and storage that resides on the server itself. The storage resources are also commonly pooled.

5.4 Network node

A network node is an element used in the infrastructure network domain of the NFVI (see figure 1 in clause 4.1). The NFV architectural framework [i.8] states that the network resources are comprised of switching functions, e.g. routers, and wired or wireless links. Also, network resources can span different domains. However, the NFV architectural framework [i.8] differentiates only the following two types of networks:

- NFVI-PoP network: the network that interconnects the computing and storage resources contained in an NFVI-PoP. It also includes specific switching and routing devices to enable external connectivity;
- Transport network: the network that:
  - interconnects NFVI-PoPs;
  - connects NFVI-PoPs to other networks owned by the same or different network operator; and
  - connects NFVI-PoPs to other network appliances or terminals not contained within the NFVI-PoPs.

In an NFVI-PoP, a network node can furthermore be categorized as follows:

- A network node which hosts compute and storage nodes, and connects to other network nodes (e.g. Top of Rack (ToR) switch, access switch, leaf switch, etc.);
- A network node which interconnects other network nodes (e.g. aggregation switch, spine switch, etc.);
- A network node which connects to transport network (e.g. gateway router).

5.5 NFVI Node

Figure 2, which is equivalent to figure 22 in ETSI GS NFV-INF 003 [i.9], gives an illustrative example of how an NFVI Node might be constructed from compute, storage, network and gateway nodes.

NOTE: Per definitions in ETSI GS NFV-INF 003 [i.9], compute nodes are expected to have internal storage (local disk) that may be necessary to support the processor(s) in the node. By contrast, a storage node refers to external large scale and non-volatile storage of information.

In this example, a compute node might be implemented as a server blade, a network node might be implemented as a Top of Rack Ethernet switch, a storage node might be implemented as a network attached storage device and a gateway node might be implemented by some optical interface device.

Figure 2 shows a logical view of an NFVI Node without physical aspects (e.g. deployment). The alignment and mapping with physical hardware components/racks is discussed in clause 6.

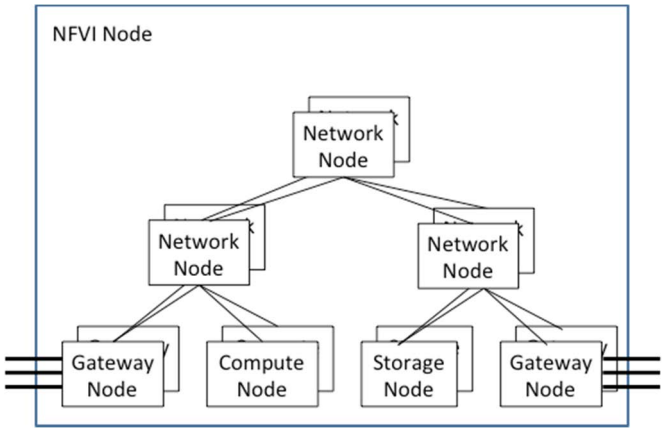

Figure 2: Example NFVI Node implementation including compute nodes and storage Nodes (equivalent to figure 22 of ETSI GS NFV-INF 003 [i.9])

6 Physical components

6.1 Commercial products

NFV aims to leverage standard IT virtualisation technology to consolidate many network equipment types onto industry standard high-volume commercial servers, switches and storage, which could be located in data centres, Network Nodes and in the end user premises [i.1].

Network operators desire to "mix & match" hardware from different vendors, without incurring significant integration costs and avoiding lock-in and eliminating the need for application-specific hardware. The skills base across the industry for operating standard high volume IT servers is much larger and less fragmented than for today's telecom-specific network equipment. Although the hardware deployed for NFVI may not be identical to that deployed in IT applications (due to environmental, regulatory, reliability, and other differences), leveraging IT best practices, technologies, and scale is desired [i.1].

The IT ecosystem is comprised of several main hardware component types: compute nodes (servers), networking nodes (switches and routers), and storage nodes (JBOD and other network attached storage devices). This ecosystem works because there is de-facto agreement within the industry regarding what functionality each of these components provide, what interfaces (both physical and logical) are provided, and what environment they operate in.

IT needs have evolved over time with the emergence of hyper-scale and cloud computing. As a result, the commercial products ecosystem has also continued to expand, embracing new concepts such as Open Compute Project and other rack-scale architectures. It is expected that the wide-scale deployment of NFV will continue to broaden the ecosystem with commercial products that meet the specific needs of NFVI.

6.2 Racks

6.2.1 Introduction

Equipment racks (also referred to as frames) provide a mechanically stable space into which compute, networking, storage and other nodes may be installed. Clauses 6.2.2 to 6.2.8 outline rack considerations related to NFVI Nodes. A simple rack might consist of two upright posts to which other infrastructure components can be affixed. More sophisticated racks could include additional features such as integrated cooling, integrated power, and seismic hardening. For racks with managed resources such as cooling and integrated power, it is common for the rack to include a management controller to configure, monitor and control these resources.

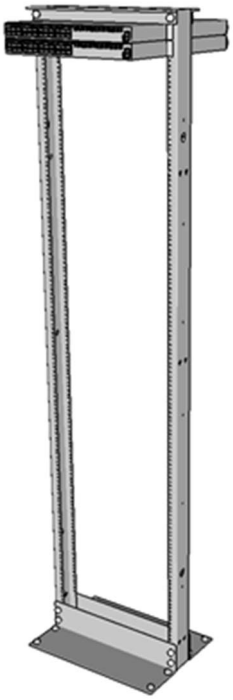

Figure 3: Simple rack with network top-of-rack switches

Clauses 6.2.2 to 6.2.8 outline rack considerations related to NFVI Nodes. Additional clauses within the present document cover power, cooling and hardware management.

6.2.2 Relationship of racks to NFVI Nodes

6.2.2.1 Introduction

NFVI Nodes are physical devices that are deployed and managed as individual entities. The purpose of NFVI Nodes is to provide the NFVI functions (hardware and software) required to support execution of VNFs [i.1]. NFVI Nodes are comprised of compute, network and storage nodes as dictated by the application needs of each particular VNF.

In a similar way, equipment racks are physical devices that can contain compute, storage, and networking resources. Although racks can be used to build NFVI hardware infrastructure, the relationship is not necessarily one-to-one. This clause examines relationship between equipment racks and NFVI Nodes.

6.2.2.2 Geographic location

Geographically speaking, each NFVI Node is sited within an NFVI-PoP [i.1]. Necessarily, its constituent compute nodes, network nodes and storage nodes (also called resource nodes in the present document) are also located within the same NFVI-PoP. The exact locations for each of these resource nodes within the NFVI-PoP, however, is determined by a Management and Orchestration (MANO) mechanism or some other provisioning mechanism so that the precise "location" of an NFVI Node is difficult to define. To complicate matters, the "location" of the NFVI's resource nodes may change over time as VNFs are moved from one resource node to another.

Racks, on the other hand, are physically deployed within the NFVI-PoP at a fixed location (row and aisle) and in most cases will not be moved. The fixed geographic nature of the rack provides a stable reference point within the NFVI-PoP and may be important when directing service/maintenance personnel to particular equipment.

The present document envisions that NFVI Nodes will be mapped onto one or more equipment racks, each containing compute, network, and storage nodes. The particular mapping of an NFVI Node onto physical resources might have implications on performance or reliability based on geographic locality within the NFVI-PoP. As an example, two compute nodes that are placed in geographically distant locations within the NFVI-PoP may traverse through additional switching layers, whereas, if they are collocated within the same rack, the switching latency incurred would be lessened.

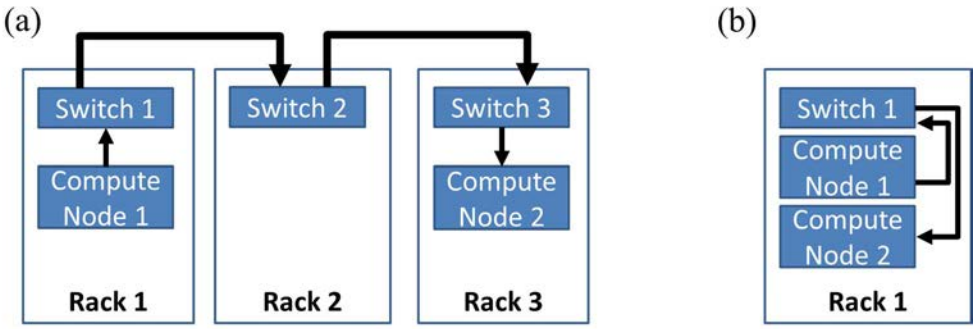

Figure 4: (a) geographically disperse compute nodes (b) compute nodes co-located within a single rack

It is recommended that the provisioning mechanism comprehend these geographic interdependencies.

6.2.2.3 1:N NFVI Node to rack mapping

1:N mapping allows mapping of one NFVI Node exclusively onto multiple racks. This is shown in figure 5. In the case where N is 1, the single NFVI Node is mapped onto a single rack. This special case is referred to as 1:1 mapping.

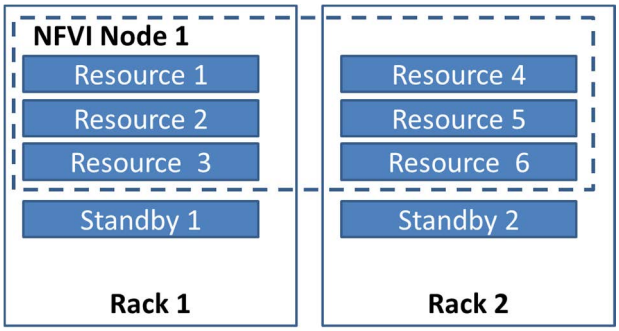

Figure 5: 1:N Mapping of NFVI Node to racks

1:N mapping benefits from allowing NFVI Nodes of larger scale than 1:1 mapping. Redundancy of rack-level shared resources is necessary and some resource inefficiency may exist since standby elements are also present in the rack and cannot be shared with other NFVI Nodes.

6.2.2.4 N:1 NFVI Node to rack mapping

N:1 NFVI Node to rack mapping allows for multiple NFVI Nodes to be mapped to a single rack of equipment. In this case, each of the NFVI Nodes shares the rack level infrastructure (e.g. power, cooling, hardware management), and can possibly share standby resources within the rack. This is shown in figure 6.

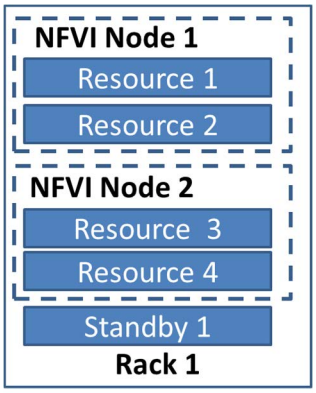

Figure 6: N:1 mapping of NFVI Nodes to rack

This configuration benefits from possibly higher resource efficiency since standby resources can be utilized by multiple different NFVI Nodes, however, additional importance is placed upon rack reliability since failure of shared rack-level resources can cause failure of multiple NFVI Nodes.

Also of note, because rack-level-resources (if any) are now shared between multiple NFVI Nodes, it may be necessary for the rack-level hardware manager to comprehend the relationship between each resource within the rack and the NFVI Node to which it is mapped.

6.2.2.5 N:M NFVI Node to rack mapping

The N:M NFVI Node to rack mapping is the most general of all the mappings. Any number of NFVI Nodes may be mapped onto any number of equipment racks as shown in figure 7.

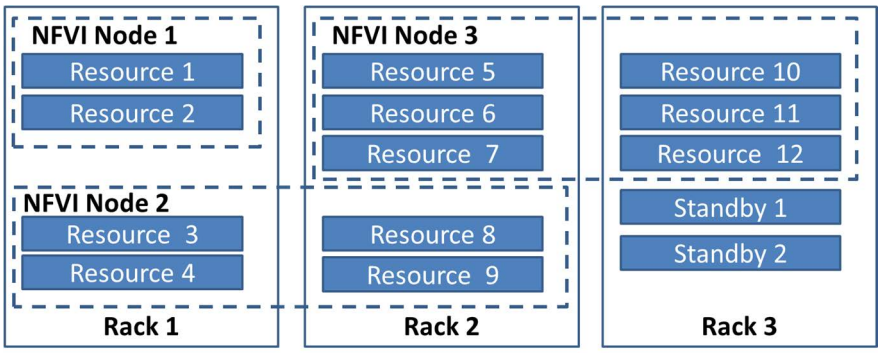

Figure 7: N:M NFVI Node to rack mapping

N:M NFVI Node to rack mapping allows for the greatest resource efficiency and immunity from rack-level failure events but places the largest burden on coordination between rack-level hardware management and VNF provisioning functions.

Because of the flexibility of this model, the present document recommends N:M NFVI Node to rack mapping be supported by NFVI.

6.2.3 Industry equipment practices

6.2.3.1 Mechanical dimensions

Most racks today are based on the Electronic Industries Alliance EIA-310 specification [i.11]. Compliant racks accept equipment that is up to 17,72 inches (450 mm) wide, with front panels of as much as 19 inches wide. The mounting hole pattern on the rack posts repeat every 1.75" (44,45 mm). This distance is defined as one rack unit (1RU) and most rack-mount equipment comes in integer multiples of this height. Because of the width, EIA-310 racks are commonly referred to as "19 inch racks".

19 inch racks come in a variety of heights, with the most common configurations able to accommodate either 42U or 43U of equipment. It should be noted that not all rack heights are acceptable for all facilities or installation procedures. For instance, some installations may require shorter rack heights in order to allow for ease of transport through doorways.

19 inch racks also support a variety of depths of equipment. The two post rack shown in figure 3 does not limit the depth of equipment that can be installed. Enclosed racks may impose restrictions. Typical maximum equipment depths are 600 mm or 800 mm. It should be noted that ETSI ETS 300 119-4 [i.12] defines an alternate rack standard that defines a maximum rack depth of 600 mm for a rack width of approximately 21 inches (535 mm). Rack depth is important in facility design in order to plan for sufficient space for cooling and cabling between adjacent aisles of equipment.

An alternate approach to rack design comes from the Open Compute Project. Open Compute Project's Open Rack defines a rack with an outside dimension of 600 mm and a depth of 1 067 mm. Open Rack supports equipment that is up to 537 mm wide 800 mm deep, and integer multiples of 48 mm in height.

Table 1: Rack standards mechanical dimensions

Rack Standard	Width	Depth	Equipment Size
EIA-310 (19" Rack)	600 mm (Typical)	600 mm, 800 mm typical	17,72" (450 mm) wide Unspecified depth, Multiples of 1,75 (44,45 mm) high
ETSI ETS 300 119-4 [i.12]	600 mm	600 mm	500 mm wide 545 mm deep, Multiples of 25 mm high
Open Rack	600 mm	852,6 mm	538 mm wide, 800 mm deep Multiples of 48 mm high

6.2.3.2 Weight considerations

Two main issues need to be considered with regard to weight. The first consideration is how much weight a rack can safely support. The second consideration is how much weight surface area that the installation facility can withstand. The total equipment within the rack can be no more than the smaller of the two constraints.

Many of today's racks are capable of housing more equipment than typical facility flooring can support. Typical full height racks today support between 907 kg (2 000 lbs) and 1 360 kg (3 000 lbs). Typical raised floor systems can support 600 kg to 1 200 kg per square meter. It is the responsibility of the facility architect to ensure that maximum loading capacity of the rack and flooring is not exceeded.

6.2.3.3 Safety considerations

Safety is a broad topic that encompasses many practices related to the prevention of injury and damage to property. Main rack safety practices can be grouped into the following classifications.

1) Proper labelling and warning information;

2) Protection from electrical discharge or shock;

3) Protection from seismic events;

4) Protection from fire/burns;

5) Protection from tipping/falling;

6) Protection from hazardous chemicals;

7) Protection from cuts.

Most countries have safety regulating bodies and requirements. It is recommended that rack equipment for NFVI be developed with compliance to these safety regulations in mind.

6.2.3.4 Electromagnetic interference considerations

All electronic devices radiate electromagnetic waves into the surrounding environment. The intensity of the emissions is related to a variety of factors including: printed circuit board practices, grounding techniques, signal edge rates and clock frequencies. If emissions are severe enough they can interfere with the normal functioning of other electronic devices.

Since the primary function of equipment racks is to house the other NFVI Node infrastructure components, they are not expected generate significant amounts of electromagnetic interference (EMI). Industry best practices dictate that radiated emissions generally be contained close to the source, however, racks that enclose the equipment that they hold may form a part of the overall EMI mitigation strategy.

Most countries have regulations regarding electromagnetic interference. It is recommended that rack equipment for NFVI be developed with compliance to these EMI regulations in mind.

6.2.4 Installation of compute/network/storage nodes

It is desirable to make installation and removal of compute, storage and network nodes as safe and fool proof as possible. Some industry best practices include:

- Limiting the compute/storage/networking node equipment weight so that it can be installed by no more than two service personnel;
- Use of standard tools or support tool-less installation;
- Support of cable maintenance mechanisms to keep nodes from being "tied in" by cables;
- Proper labelling;
- Reduction of the number of cables that need to be connected to each removable device.

It is expected that NFVI Node rack hardware will incorporate reasonable ease-of-use features.

6.2.5 Rack-level management considerations

As previously noted, racks introduce a purely physical dimension into NFVI equipment. Racks reside in a fixed physical position within an installation. In addition, racks contain physical devices (compute/storage/networking nodes, etc.) that contribute to the overall NFVI equipment While the physical location of these devices need not be known to the VNFs that reside on them, it may be useful to provide geographic location information to aid service personnel in repair and maintenance of the equipment.

6.2.6 Volumetric efficiency considerations

When related to NFVI rack equipment, volumetric efficiency is the amount of useful VNF work that can be accomplished within a specific amount of space. Volumetric efficiency of equipment might be important when facility space is limited since more efficient equipment will be able to accomplish the same amount of work in a smaller space.

Many factors influence volumetric efficiency. Some of them are listed hereafter:

- Size of compute/networking/storage nodes;
- Size of non-compute/networking/storage elements (e.g. power) within the rack;
- Performance of compute/networking/storage nodes relative to the VNFs that are executing on them;
- Rack dimensions;
- Cooling solution size;

- Cable allowances (contributes to depth of rack);
- Air plenum space between racks.

Most of these parameters are interdependent and need to be optimized as a whole. For instance, higher volumetric efficiency can be achieved by making the compute/storage/and networking nodes smaller; however, at some point processing performance, storage capacity and network ports will be limited due to space constraints. Furthermore, a denser node definition might necessitate higher a performance cooling solution - leading to larger fans or air plenums.

In installations where total rack power is limited, densely populated racks may not be possible due to overall power limitations.

6.2.7 Open Compute example

The Open Compute Project is an online community whose mission is to create and deliver efficiently functioning servers, storage capabilities and data centre designs. Open Compute is one example of an open ecosystem of commercial hardware that is being deployed today. This clause examines the key elements of the Open Compute Project Open Rack.

Open Rack v2.0 from Open Compute Project™ defines a rack that is nominally 2 110 mm high, 600 mm wide and 1 067 mm in depth. The rack accommodates sixteen 96 mm high payload trays, three standard IT switches, and two power supplies. Servers can be inserted without tools and are held in place by spring latches.

NOTE: Open Compute Project™ is the trade name of a collaborative open source project from the Open Compute Project Foundation. This information is given for the convenience of users of the present document and does not constitute an endorsement by ETSI of the product named. Equivalent products may be used if they can be shown to lead to the same results.

The base rack configuration supports 500 kg but can be configured with additional cross-members in order to support up to 1 400 kg of equipment. Additional options allow for GR63 Zone 2 and Zone 4 seismic protection.

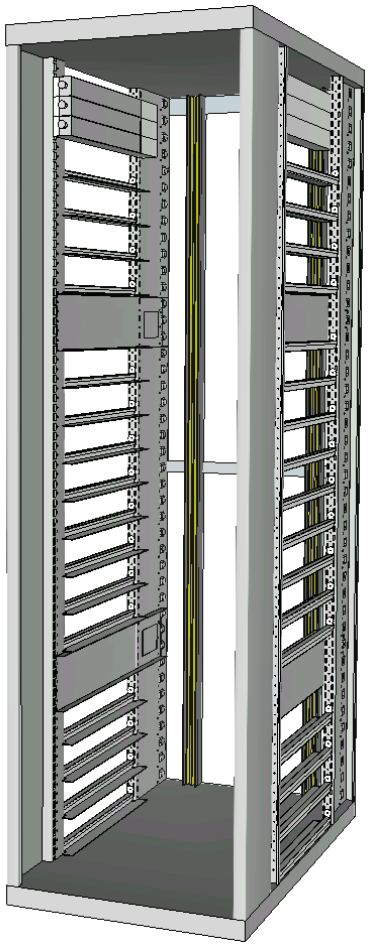

Figure 8: An illustrative example of an Open Compute Open Rack

6.2.8 Recommendations

The present document makes the following recommendations related to NFVI Node hardware:

1) Racks support a variety of NFVI Node scales;

2) Racks support a mapping of an arbitrary number of NFVI Nodes onto one or more physical racks of equipment (N:M mapping);

3) Racks support geographical addressing in order to facilitate ease of maintenance and possible optimization based on locality of compute/storage/networking resources;

4) Racks support 600 mm width options in order to fit within existing facilities;

5) Racks support a range of weight capacities consistent with common facility load-bearing capabilities;

6) Racks support relevant regional and global safety requirements for telecommunications and/or data centre equipment;

7) Racks support relevant regional and global radiated emissions requirements for telecommunications and/or data centre equipment;

8) Racks offer compelling volumetric density options for installations that desire it.

Some open questions and areas for further investigation for NFVI racks include:

1) What is the typical load (weight) rating needed for NFVI racks?

2) What is the maximum allowable depth (including cabling) for NFVI racks?

6.3 Processors

6.3.1 Introduction

Central Processing Unit (CPU) is an essential part of the NFVI Node, and is the core of data processing. It is expected to be an integral part of all compute, storage and networking nodes, thus its performance, reliability, compatibility can have a big influence on the whole performance of the NFVI Node. Clauses 6.3.2 to 6.3.6 outline processor considerations related to NFVI Nodes.

6.3.2 Instruction set

General purpose instructions are of variable length, and comprise optional prefixes, the opcode, and optional modifiers, address offsets, and immediate data. In addition to general purpose instructions, there are a number of instructions that have been introduced over different processor generations to accelerate specific important operations. An example would be Advanced Encryption Standard New Instructions (AES-NI). This is a set of instructions to greatly improve performance for encryption/decryption using AES. Cyclic Redundancy Check (CRC) is another such function.

Some instructions are used to enable new operating modes, such as the virtual machine (VM) extensions used to support virtualisation.

6.3.3 Multi-core support

CPUs now contain multiple cores, with each core capable of virtualisation to support many VMs. Multi-core processors take advantage of a fundamental relationship between power and frequency. By incorporating multiple cores, each core is able to run at a lower frequency, reducing the processor power consumption for a given workload on a core or thread.

When discussing multi-core, it is important to distinguish the multiple ways to implement multi-core architectures. They are:

- Multi-Socket: having 2 or more physical processors in different sockets. One of the most common is a Dual Processor implementation, comprising two separate processor packages, linked together over a bus;
- Multi-Core: having multiple cores within a single socket. As silicon technology progresses, more cores can fit within one package;
- Hyper-Threading: Hyper-Threading provides two or more execution pipelines on a single core. Thus from a software point of view there are two or more distinct logical cores.

All of the above technologies can be combined within a single platform, hence one could have hyper-threading on a multi-core dual socket platform. From software's point of view, each core (logical or physical) appears the same.

Besides having multiple cores in one CPU, the NFVI Node can have multiple CPUs and the single blade/motherboard can have multiple sockets. The same processor, blade, motherboard can be used for multiple different applications and the applications can change dynamically enabling reuse of the processor platform.

6.3.4 Operating system & hypervisor (virtualisation) support

Every CPU is expected to support all types of Operating Systems. Virtualisation environments include open source environments such as Xen™ and Kernel-based Virtualisation Machine (KVM), as well as commercially supported environments also need support from CPU. To enable service agility, the CPU system is expected to support widely used Operating Systems (OSs), hypervisors and other third party software.

NOTE: Xen™ is a trade name of Citrix Systems Inc. This information is given for the convenience of users of the present document and does not constitute an endorsement by ETSI of the product named. Equivalent products may be used if they can be shown to lead to the same results.

Besides basic support for virtualisation, some CPUs contain virtualisation features that can further boost the performance and reliability of the virtualisation system. For example, the CPU system can include the Single Root Input/Output Virtualisation (SR-IOV) capability which is the Peripheral Component Interconnect (PCI) standard for how peripheral devices are supported in a virtualised environment. This allows a single peripheral device (e.g. Network Interface Component) to be partitioned into multiple virtual functions. It is recommended for the CPUs to support these features.

6.3.5 Registers, cache & memory architecture

System software is expected to be enabled to create multiple Direct Memory Access (DMA) protection domains. Each protection domain is an isolated environment containing a subset of the host physical memory. Depending on the software usage model, a DMA protection domain may represent memory allocated to a VM, or the DMA memory allocated by a guest-OS driver running in a VM or as part of the hypervisor itself.

The CPU architecture enables system software to assign one or more Input/Output (I/O) devices to a protection domain. DMA isolation is achieved by restricting access to a protection domain's physical memory from I/O devices not assigned to it by using address-translation tables. This provides the necessary isolation to assure separation between each VM's computer resources.

Some operations require the use of specific registers for different purposes, but many can be used as the software wishes. As software developers use higher level languages, the compiler is responsible for making efficient use of the available registers, including:

- General registers;
- Segment registers;
- Index and pointers;
- Indicators.

Information on these and similar registers is generally available in the software developer's manuals. Different instruction sets may cause a re-compile when moving from one architecture to another. As new generation of processors, with new instructions, registers, etc. become available, it is expected that the legacy software can run on the new generation of processors.

The type of cache available depends on which specific processor is chosen. Most processors have L1/L2 cache, depending on the particular processor there may be a L3 cache as well. The various classes CPUs have wide ranges of available cache, depending on whether it is for a low power embedded kind of processor all the way up to a high performance server. Not only the amount of cache, but how it is shared (or not) among the various cores varies widely. Various protocols are used to ensure cache coherency is maintained. Memory support is similarly varied amongst different CPU implementations. One aspect to be aware of is the Uniform Memory Access (UMA) vs Non-Uniform Memory Access (NUMA). In UMA, all processors access shared memory through a common bus (or another type of interconnect). By contrast a dual socket implementation typically would have memory physically attached to each socket, hence it is a NUMA architecture. Most operating systems allocate memory in a NUMA aware manner, thus minimizing any additional latency for accessing remote memory.

Some high availability technologies are used to guarantee the reliability of data stored in memory module. For example, Error Correcting Code (ECC) can detect and correct common data corruption. Other examples include Single Device Data Correction (SDDC), device tagging, patrol scrubbing, failure isolation, mirroring, etc. In scenarios that have a high requirement of data reliability, memory modules supporting these technologies are recommended.

6.3.6 Processor recommendations

The present document makes the following recommendations with regard to processors:

- Multi-core processors are preferred;
- NFVI Node equipment supports multiple processors;
- Processors support widely used operating systems, hypervisors (virtual environments) and third-party software;
- Processors support virtualisation features that can boost the performance and reliability of the virtualisation system, such as SR-IOV;
- Processors are reusable for most if not all types of applications;
- Processors support backward and forward compatibility;
- DMA is enabled to provide the necessary isolation of each VM's compute resources - this is one use case;

- NUMA is supported to minimize additional latency for accessing remote memory;
- Memory modules support high availability technologies to guarantee the reliability of stored data, such as ECC, SDDC, etc.

6.4 Power

6.4.1 Introduction

Power is needed for normal operation of any electronic equipment. NFVI nodes receive power from facility or installation sources and distribute it to the various sub-systems within the node. Clauses 6.4.2 to 6.4.7 outline power considerations related to NFVI Nodes.

6.4.2 Typical elements of power distribution

6.4.2.1 Introduction

For illustrative purposes, figure 9 shows a typical facility power distribution system. A wide range of variation exists between different types of facilities. However the general subsystems and flow remain similar.

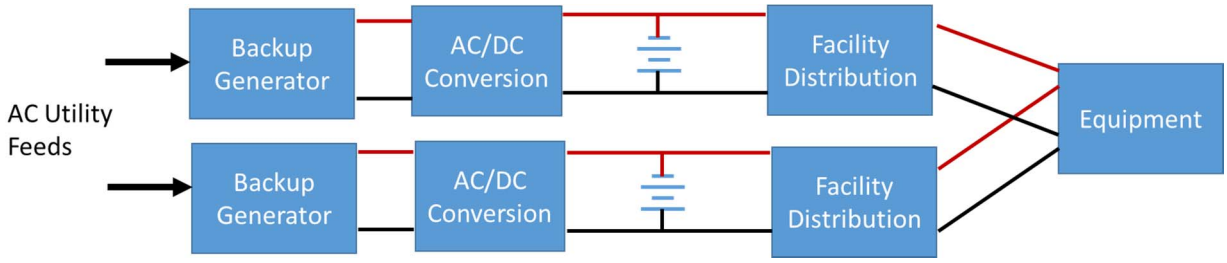

Figure 9: Typical facility-level power distribution architecture (simplified)

AC power entering the facility gets converted to facility power and is used to power the equipment as well as charge large backup batteries. Backup petrol-powered generators may also be present in order to protect the facility from outage when utility power fails.

Recent innovations such as Open Compute Project's Data Centre specification focus on optimization specific to the data centre environment which might also be applicable to NFVI. In particular, higher voltage AC and DC power is supplied to the equipment racks in order to limit distribution losses, and the centralized UPS batteries are replaced with more efficient cabinet-level battery backup.

The present document focuses on power recommendations specific to NFVI Nodes and therefore limits discussion to the equipment rather than the facility power architecture. It is important to note, however, that the NFVI Node power architecture is a small part of the overall facility power architecture and needs to be designed in such a way to provide efficient, safe, and reliable power within the larger context.

6.4.2.2 Facility power

6.4.2.2.1 Contex

Facility power is delivered to each equipment rack in order to power the devices within the rack. In most cases, redundant power feeds are supplied in order to mitigate loss of function within the rack if one of the power feeds fails. Today, multiple different facility power feed types are used and rack equipment may require multiple different types of power to operate.

6.4.2.2.2 -48 VDC power

-48 DC power is commonly used in telecommunications data centres in order to supply power to networking equipment. It is expected that some of the facilities in which NFVI Nodes is deployed will deliver -48 VDC power. The present document recommends that NFVI Nodes support -48 VDC input feeds as an option.

It should be noted that AC powered equipment operating in racks powered exclusively by -48 VDC power will need AC power inversion (conversion of DC power to AC power) in order to operate. This need can be eliminated if the facility also supplies AC power to the rack.

6.4.2.2.3 AC power

Many data centres use AC power facility feeds. There are regional differences in both voltage and frequency for these feeds. The present document recommends that NFVI Nodes support AC feeds as an option.

It should be noted that DC powered equipment operating in racks powered exclusively by AC power will need a DC power supply (conversion of DC power to AC power) in order to operate. This need can be eliminated if the facility also supplies the appropriate DC voltage to the rack.

6.4.2.2.4 High voltage DC power

Some datacentre architectures, such as the Open Compute Project Data Center, distribute high voltage DC power to the equipment. High voltage DC offers some advantages of limiting distribution losses within a facility. Many different high voltage DC power options are being contemplated. The present document recommends that NFVI Nodes support high-voltage DC feeds as an option.

DC powered equipment operating in racks will need a DC-DC power supply (conversion of high voltage DC power to lower voltage DC power) in order to operate.

It should be noted that AC powered equipment operating in racks powered exclusively by high voltage DC power will need AC power inversion (conversion of DC power to AC power) in order to operate. This need can be eliminated if the facility also supplies the appropriate DC voltage to the rack.

6.4.2.3 Voltage conversion

Voltage conversion is the process of converting one voltage type (AC or DC) to another, or one voltage level (high voltage DC, DC) to another. Almost all electronic equipment performs some form of voltage conversion to supply the proper type of power to its components; however, since voltage conversion always introduces losses, best practice dictates avoiding extra voltage conversion steps whenever possible.

Voltage conversion is likely to occur in two main locations within NFVI Nodes as shown in figure 10.

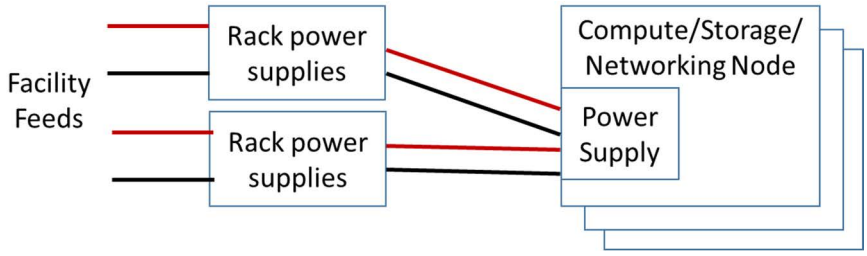

Figure 10: Representative NFVI Node power distribution showing power conversion stages

Facility feeds entering the rack may need to be converted to the voltage level and type distributed within the rack. This is the case when one voltage level DC (e.g. -48 VDC) is converted to another (e.g. 12 VDC) or when AC input feeds are converted to DC. Resource nodes typically have dedicated power supplies in order to generate the specific voltage levels required by their circuitry.

It is necessary that rack power supplies be appropriately sized to provide all the power needed by equipment within the rack.

In order to offer highly reliable operation of the system, power supplies typically incorporate voltage (and sometimes power) sensing. If a voltage fault occurs, it is recommended that NFVI Node hardware be capable of detecting the fault and notifying the hardware management system.

6.4.2.4 Backup power

In traditional facilities, backup power is provided by large facility batteries. In newer installations, such as those based on the Open Compute Project Data Centre specification, backup power is supported by cabinet-based batteries. It is desirable for NFVI Node hardware to support both models.

Battery systems typically have sensors to detect the charge level and utilization of the backup power. If backup power is incorporated into the NFVI Node rack, it is recommended that these sensors be present and capable of integrating with facility hardware management system.

6.4.2.5 In-rack power distribution

In-rack power distribution is responsible for distributing power from the in-rack power supplies or facility feeds to each of the resource nodes within the rack. Other than meeting relevant safety and industry safety regulations, no special needs are envisioned.

6.4.3 Power redundancy models

6.4.3.1 Redundant rack feeds

Redundant power feeds are simply more than one set of power feeds from the facility power distribution source. Redundant power feeds protect the rack from possibility of outage should a single facility power feed fail. The present document recommends that all NFVI Nodes support multiple power feeds.

6.4.3.2 Redundant rack power conversion

Redundant rack power conversion allows for multiple power supplies sharing the responsibility of providing power to the rack. If any one of the power supplies fails, the others have sufficient capacity to keep the rack running. A typical, redundant power contains multiple power supply modules that can be hot-swapped individually when in need of service. A representative rack power supply is shown in figure 11.

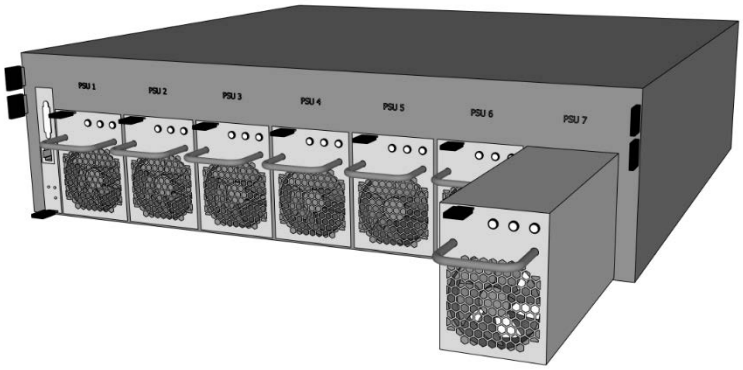

Figure 11: Representative rack power supply unit with redundant hot-swappable modules

Since the loss of the rack power supply could create a rack-level outage, the present document recommends redundant rack-level power supplies when rack power supplies are required.

6.4.3.3 Redundant rack power distribution

Power distribution is responsible for distributing the power from the in-rack power supplies to the individual resource nodes. Certain failure scenarios, such as a short circuit condition, could render an entire in-rack distribution feed inoperable. If only one in-rack feed is present, then a rack-level outage will occur. Redundant power distribution is desirable to protect from situations such as these.

6.4.3.4 Redundant compute/storage/network node power

Some networking equipment supports multiple redundant power supplies at the resource node. This could be particularly useful when loss of the resource node constitutes a significant level of service outage. Because of the additional complexity associated with redundant power supplies at the compute/storage/networking node, the present document recommends that redundant power supplies at the resource node be optional.

6.4.4 Power safety considerations

Safety is a broad topic that encompasses many practices related to the prevention of injury and damage to property. Main power practices can be grouped into the following classifications:

1) Proper labelling and warning information;

2) Protection from electrical discharge or shock;

3) Protection from short circuit;

4) Protection from fire/burns;

5) Protection from cuts.

Most countries have safety regulating bodies and requirements. It is recommended that power equipment for NFVI be developed with compliance to these safety regulations in mind.

6.4.5 Power efficiency

6.4.5.1 Introduction

Power efficiency is related to the amount of power that performs useful work (at the resource nodes) versus the amount of power that is used in overhead (e.g. cooling, voltage conversion). Clauses 6.4.5.2 to 6.4.5.5 examine power efficiency considerations as they relate to NFVI Node equipment.

6.4.5.2 Power conversion

Power conversion losses result every time power conversion is required from one format to another (e.g. AC to DC) or form one voltage to another. Because of these unnecessary power conversion stages, the present document recommends avoiding unnecessary power conversion stages when possible. Additionally, power supplies can be found that operate at a variety of different efficiencies. The present document recommends selection of power supply efficiencies to meet overall system design goals.

6.4.5.3 Compute, storage and networking Efficiency

Compute, storage, and networking resources consume power in order to perform work. With a push toward commercial servers, switches and storage, it is important to note that not all solutions will be equally efficient at performing the same amount of work. In particular, certain forms of accelerators will be more efficient than implementations using purely general purpose processing. The present document recommends consideration of a mix of commercial servers and commercial hardware acceleration solutions to improve power efficiency performance.

6.4.5.4 Power management

Power management encompasses a variety of technologies to limit the power consumed by electronic hardware when not running at full capacity. The present document recommends that resource nodes implement all appropriate power management options.

6.4.5.5 Redundancy models

Whenever redundant resource nodes are deployed, the possibility of inefficiency exists since the redundant node may draw same power while providing no useful work to the system. This impact can be mitigated by deploying redundancy schemes that minimize the number of standby nodes, and by ensuring that standby nodes consume as little power as possible.

6.4.6 Example from the Open Compute Project Open Rack

Open Rack from the Open Compute Project supports 3-phase facility AC or DC feeds in a variety of formats. All power supplies offer redundancy modes, and internal rack 12V power distribution is accomplished by redundant bus-bars.

6.4.7 Power Recommendations

The present document makes the following recommendations related to NFVI Node power:

- NFVI Node equipment supports -48 VDC input feeds as an option;
- NFVI Node equipment supports AC input feeds as an option;
- NFVI Node equipment supports High-voltage DC feeds as an option;
- NFVI Node is capable of detecting power faults and notifying the hardware management system;
- NFVI Node is capable of detecting in-rack backup power state (when present) and is capable of reporting the state to facility hardware management system;

- NFVI Node equipment supports multiple facility power feeds;
- NFVI Node equipment supports redundant rack-level power supplies when rack power supplies are required;
- Redundant power supplies at the compute/storage/network node are optional;
- Power equipment for NFVI be developed with compliance to safety regulations in mind;
- NFVI Node hardware avoids unnecessary power conversion stages when possible;
- Selection of power supply efficiencies meet overall system design goals;
- Consider a mix of commercial servers and commercial hardware acceleration solutions to improve power efficiency performance;
- Compute/storage/networking nodes implement all appropriate power management options.

6.5 Interconnections

6.5.1 Ethernet

The Open Compute Networking Project in OCP develops specifications for hardware and software of top-of-rack switches. It also includes spine switches in the scope as a future target. The project has already published their accepted or contributed switches, where 10/ 40 Gigabit Ethernet interfaces are commonly used.

The generic term of the 10 Gigabit Ethernet interconnection refers to the current IEEE 802.3ae specification [i.2]. The family is configured by 10GBASE-LX4, 10GBASE-SR, 10GBASE-LR, 10GBASE-ER, 10GBASE-SW and 10GBASE-LW. The former and latter parts of the interconnection names have different meanings. The former part, "10GBASE", shows the data rate of the interconnection. The latter part is divided into two meanings. The first capital of 'S', 'L', and 'E' indicates a type of optical media categorized with the operating distance, and those mean "Short Wavelength Serial", "Long Wavelength Serial" and "Extra Long Wavelength Serial" respectively.

Table 2: Types of 10GBase interconnections (1)

Former capital	Meanings
S	Short Wavelength Serial
L	Long Wavelength Serial
E	Extra Long Wavelength Serial

The latter term of 'X', 'R', and 'W' means specific encoding methods. The capital of 'X' and 'R' shows the encoding methods of "8B/10B" and "64B/66B" for LAN interface respectively. The 10GBASE-W is specified for WAN interface. The reference [i.2] describes that the data-rate and format are compatible with the SONET STS-192c transmission format.

Table 3: Types of 10GBase interconnections (2)

Latter capital	Meanings
X	encoding methods of "8B/10B"
R	encoding methods of "64B/66B"
W	WAN interface compatible with the SONET STS-192c

Table 4 shows an overview of the 10 Gigabit Ethernet interconnections. "Medium" shows the type of each interconnection. "Media type" shows the type of cables to be used for the interconnections. "Connector" shows the physical types of attachment to NFVI Nodes. "Operating distance" shows the maximum distance of the interconnection. Appropriate interconnections in table 4 are picked up and discussed for intra domain (between compute, storage, and network domains), intra NFVI Node (physical domain to hypervisor) and inter NFVI Node in clauses 6.5.2 to 6.5.6.

Table 4: Types of interconnections

Interconnection	Medium (fiber/copper/etc.)	Media type (single/multi/etc.)	Operating distance
10GBASE-LX4 (see note)	Optical	Single mode fiber	10 km
		Multi-mode fiver	300 m
10GBASE-SR	Optical	Multi-mode fiber	300 m
10GBASE-LR	Optical	Single mode fiber	10 km
10GBASE-ER	Optical	Single mode fiber	30 km/40 km
10GBASE-SW	Optical	Multi-mode fiber	300 m
10GBASE-LW	Optical	Single mode fiber	10 km
NOTE: LX4 uses four wavelengths encoded on the same fiber.			

6.5.2 Intra domain (between compute, storage and network domains)

A physical infrastructure within an NFVI-PoP should continuously be scalable and adaptable to network service needs. The requirement is carefully considered in current data centres [i.3]. A set of servers, storages, switches and other hardware resources are, for example, installed in a hardware rack, and inter-connected with leaf-spine network topology. As shown in figure 12, the leaf switches on the top of the rack are inter-connected to spine switches with fully-meshed manner.

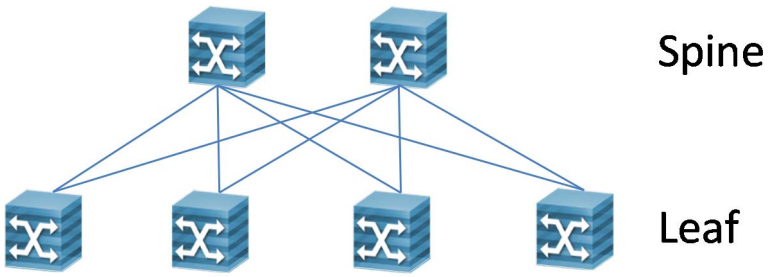

Figure 12: An overview of a leaf-spine architecture

An interconnection in an intra domain refers to a physical interconnection for Ha/Csr-Ha/Nr between compute, storage and network domains. The issue has been discussed in ETSI GS NFV-INF 001 [i.1], which shows interconnections among multiple hardware components when the hardware resources are, for instance, installed in an OpenRack standardized in OCP: two switches, three 10-OpenUnit and three power shelves are installed in the OpenRack. Compute and storage nodes are installed as components in the 10-OpenUnit innovation zone. Optical media is used to inter-connect the physical resources to the ToR switches.

Another interconnection in an intra domain refers to a physical interconnection between leaf and spine switches. As shown in figure 13, spine switches forwards aggregated traffic among leaf switches. Interconnections to gateways depend on the physical deployment model. Gateways can be inter-connected to the spine switches as shown in figure 13, and can also be inter-connected to intermediate switches as introduced in a current data centre model [i.3]. These physical interconnections are deployed within an infrastructure network domain and cannot be seen from the ETSI NFV architecture.

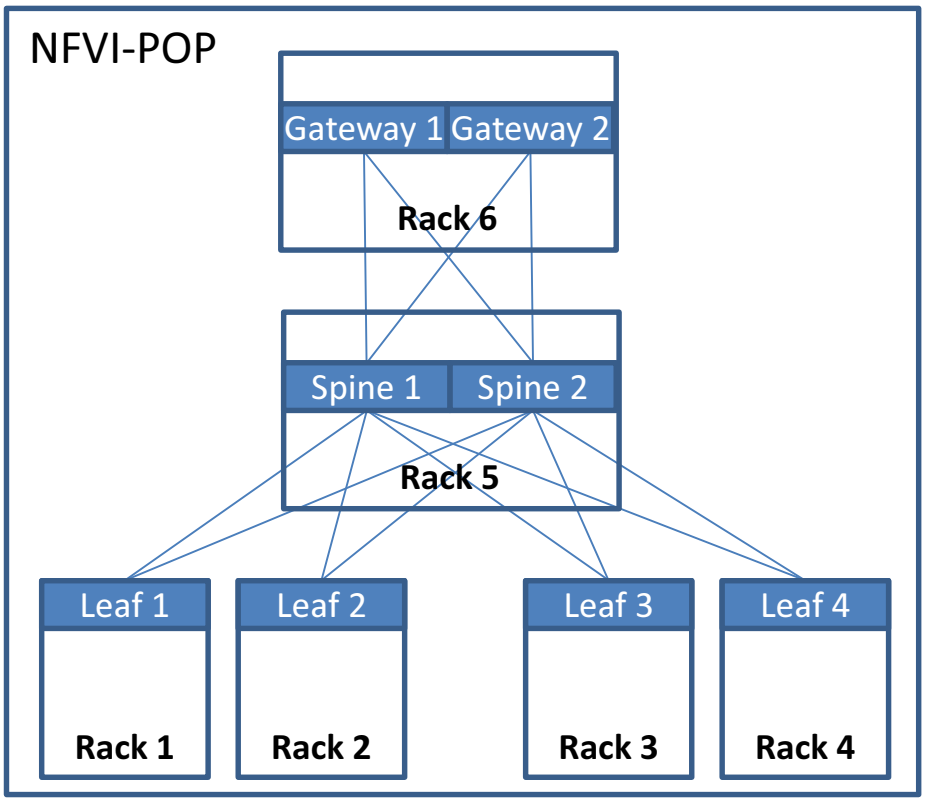

Figure 13: Deployment example

In a data centre network topology [i.3], the servers/leaf switches, spine switches and gateways are grouped respectively, and then installed into racks. Racks hosting the same type of resources are physically put in the same row or the same island. Different grouped resources are put in different rows or islands of racks.

6.5.3 Intra NFVI Node

As shown in figure 14, an NFVI Node is composed of compute node, storage node, network node, gateway, etc. Network node connects compute nodes and storage nodes, and provides network resources for switching/routing. There can be multiple interfaces between network node and other nodes. The interfaces can be classified into two categories: management interface and traffic interface. From the perspective of traffic flow direction, the traffic interface can be further classified into internal (east-west) traffic interface and external (north-south) traffic interface.

In order to ensure the NFVI Node to work properly, there are some recommendations to the interfaces/interconnections:

- QoS: adequate interface bandwidth is needed; high bandwidth with low latency switching is desired;
- Reliability: it is recommended to provide redundant interfaces/interconnections;
- Security: it is recommended for the interfaces/interconnections to support data flow isolation.

The compute, storage, and other resources (for example, hardware acceleration resources) in a NFVI Node form a resource pool, which supports the deployment and execution of VNFs. To ensure E2E service QoS, adequate interface bandwidth and low latency switching are needed. Besides, the capability of flexible scaling interface bandwidth and switching capacity according to service demand is desired to meet the need of different kinds of VNFs.

No single point failure, fast detection of component failure and self-recovery are the ways to achieve hardware high reliability and fulfil the need of service continuity. It is highly recommended to provide redundant interfaces and interconnections (illustrated as dotted box and lines in the figure) to avoid service performance downgrade caused by component failure.

To ensure the security and reliability of the communication system, different kinds of data flow (e.g. management data flow, service traffic) inside the NFVI Node is recommended to be isolated. To prevent external network attacks, isolation of east-west traffic (internal traffic) and north-south traffic (external traffic) is recommended. Data flow isolation can be implemented physically or logically. Normally, physical isolation would provide higher security and reliability than logical isolation.

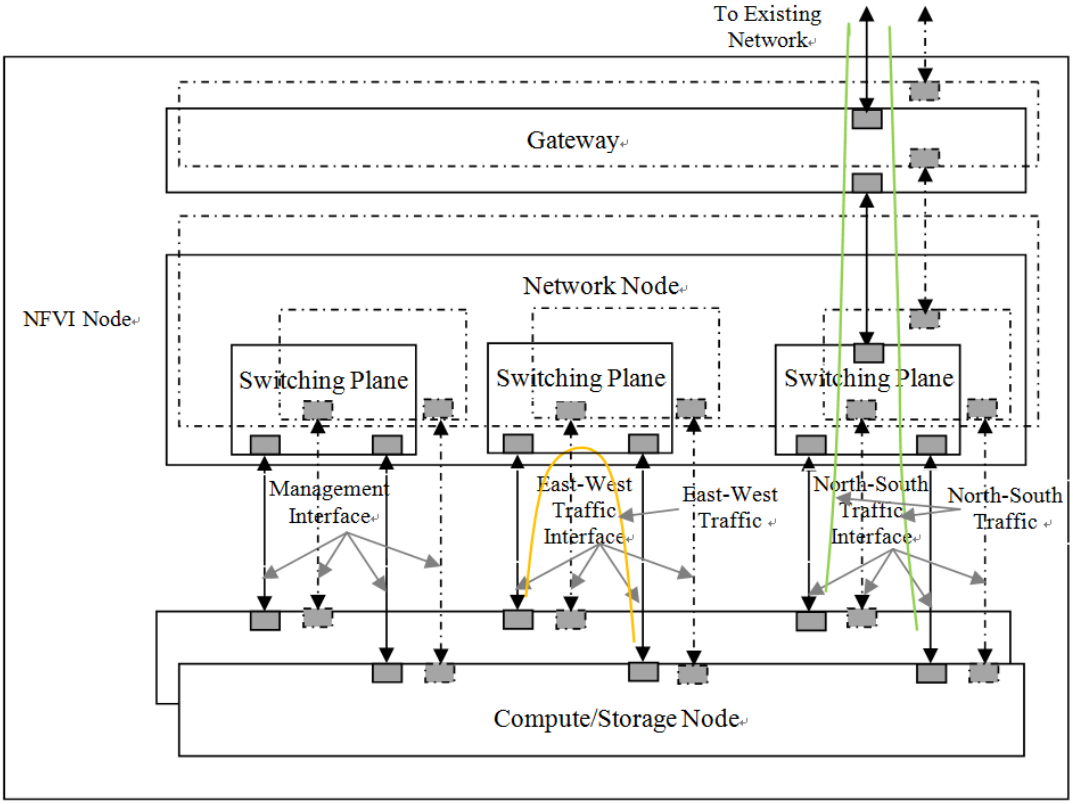

Figure 14: Example of physical interface and interconnection isolation in the NFVI Node

As shown in figure 14, for physical isolation, different types of data flow (management data flow, east-west service traffic and north-south service traffic) are switched in different physical switching planes through different physical interfaces.

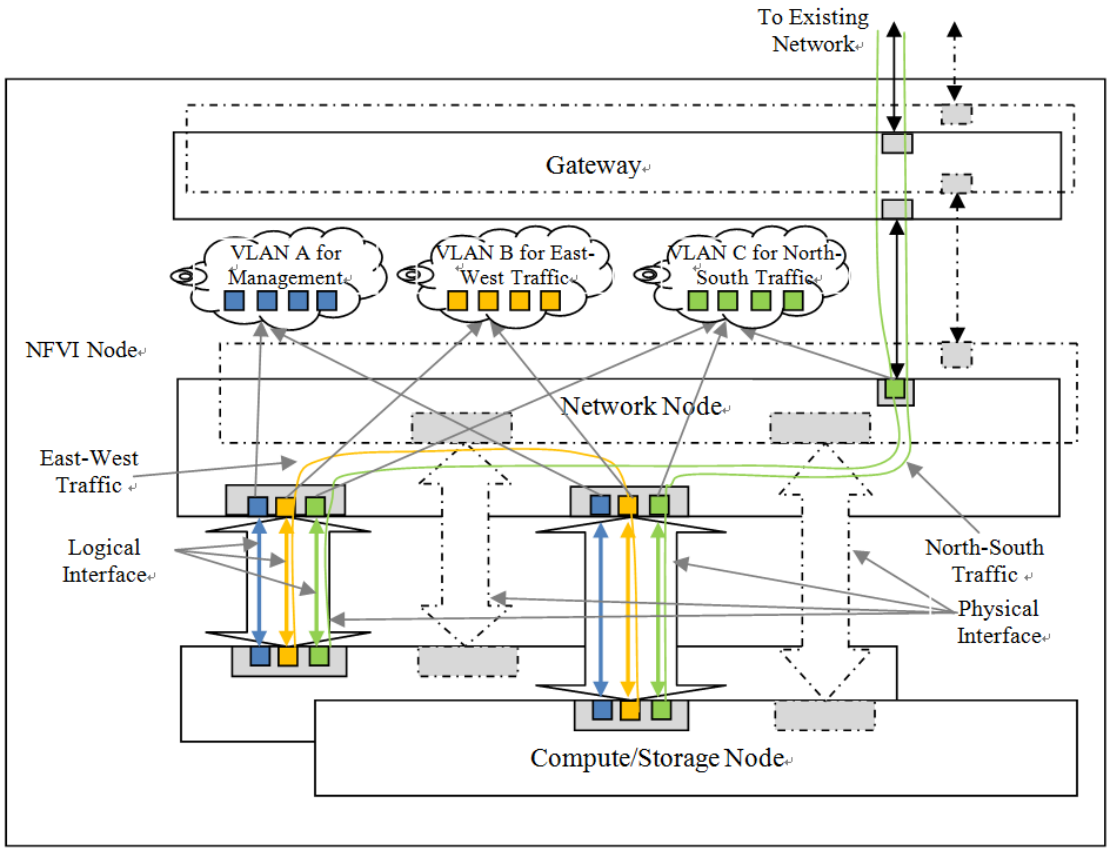

Figure 15: Example of logical interface and interconnection isolation in the NFVI Node

As shown in figure 15, for logical isolation, different types of data flow share physical interface(s) and switching plane(s). Methods like Virtual Local Area Network (VLAN) can be used to create isolated logical interfaces and switching planes to implement data flow isolation.

6.5.4 Inter NFVI Node

In an example shown in figure 16, a pair of NFVI Nodes in the same NFVI-PoP is connected through a physical interconnection. As discussed in clause 6.2.2, the abstract NFVI Node comprises one or multiple hardware resource nodes (e.g. compute node, storage node, network node). In a case where resource nodes of the two NFVI Nodes are deployed in the same rack, the physical interconnection between the NFVI Nodes is configured within the rack. In another case where the resource nodes of those NFVI Nodes are deployed in different racks, the physical interconnection traverses through multiple network nodes and racks.

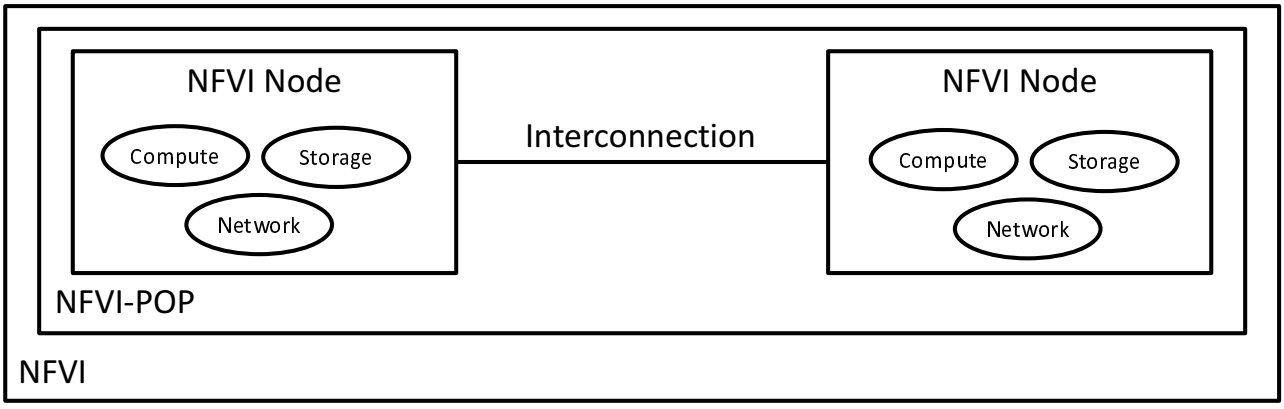

Figure 16: Interconnection between NFVI Nodes in an NFVI-PoP

In another example shown in figure 17, a pair of NFVI Nodes located at different NFVI-PoPs is connected through a transport network. The case is derived from an example where multiple VNFs located at different NFVI-PoPs are interconnected through a transport network [i.10]. In this case, there is no direct interconnection between the NFVI Nodes which execute their VNFs. Instead, gateway nodes deployed at the edge of the NFVI Nodes are connected to a network node in the transport network.

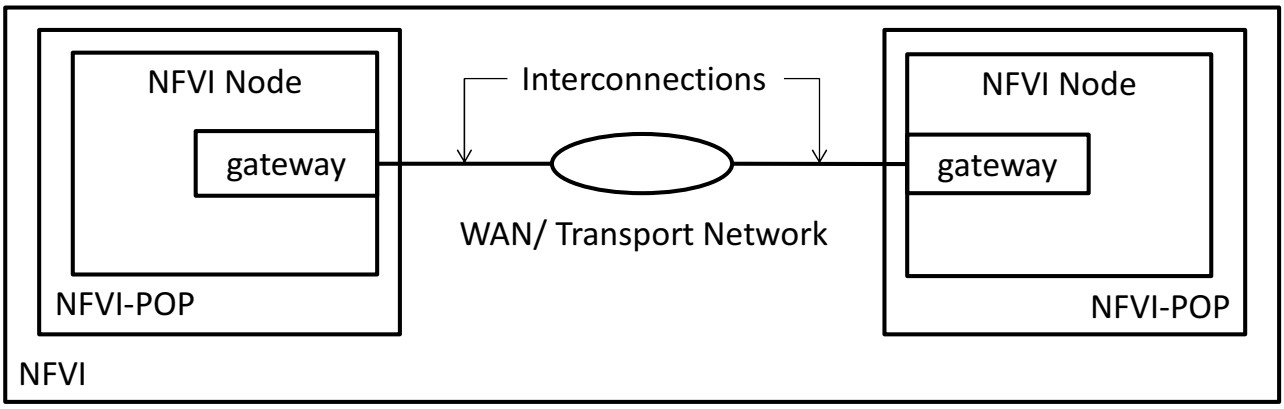

Figure 17: Interconnections between NFVI Nodes through a transport network

6.5.5 Other types of interconnections

At the time of the writing of the present document, Ethernet is the predominant interconnect envisioned for NFVI. Other specialized interconnect technologies exist that may offer advantages for some particular applications. The present document recognizes the existence of these technologies as possible alternative solutions. Discussion of these technologies, their use, and trade-offs is beyond the scope of the present document.

6.5.6 Recommendations

In general distributed system, virtual machines executing a service are deployed as much as close with each other in order to achieve better performance. On the other hand, a set of back-up system for the service is deployed in another separated area in order to achieve reliability. Physical distance for an interconnection is, therefore, an important aspect to keep the balance of performance and reliability.

The maximum length of an interconnection in a rack would be from a single digit meters or a little bit more. But the maximum length of the interconnections between racks depends on the physical footprint size of the NFVI-PoP. Two or three digits meters should be considered. Appropriate types of interconnections should be applied to individual physical interfaces.

In this study, it is recognized that 10 and 40 Gigabit Ethernet interconnections are commonly used in current data center. Therefore, photonic media is preferable to deal with the broadband and aggregated traffic rather than electronic interconnections.

6.6 Cooling

6.6.1 Introduction

Cooling is essential for normal operation of any electronic equipment. As the equipment runs, it dissipates power and generates heat. The NFVI Node cooling is responsible for removing excess heat and keeping components within acceptable temperature limits. Clauses 6.6.2 to 6.6.7 outline cooling considerations related to NFVI Nodes.

6.6.2 Typical elements of cooling

6.6.2.1 Facility and environmental

The first level of the NFVI Node cooling system is the environment in which the equipment is installed. The ambient temperature of the air around the equipment, the relative humidity, air pressure, and amount of direct sunlight the equipment receives all impact the cooling behaviour of the equipment.

6.6.2.2 Rack cooling

When present, rack cooling solutions are responsible for providing cooling to all components within the rack. Rack cooling solutions generally consist of multiple fans (to move cool air into the rack and heated air out). Chiller units may also be installed.

In many cases, racks do not provide cooling. Instead, each piece of equipment installed within the rack is expected to provide its own cooling solution.

6.6.2.3 Chip cooling

Chip cooling is associated with removing heat from individual electronic components. Processors, memories and other devices may generate tens to hundreds of watts in a very small surface area. If excess heat is not removed from the chip, permanent component failure can result. Typically, heat sinks are affixed to the surface of the chips and are used to extract heat from the integrated circuit as shown in figure 18.

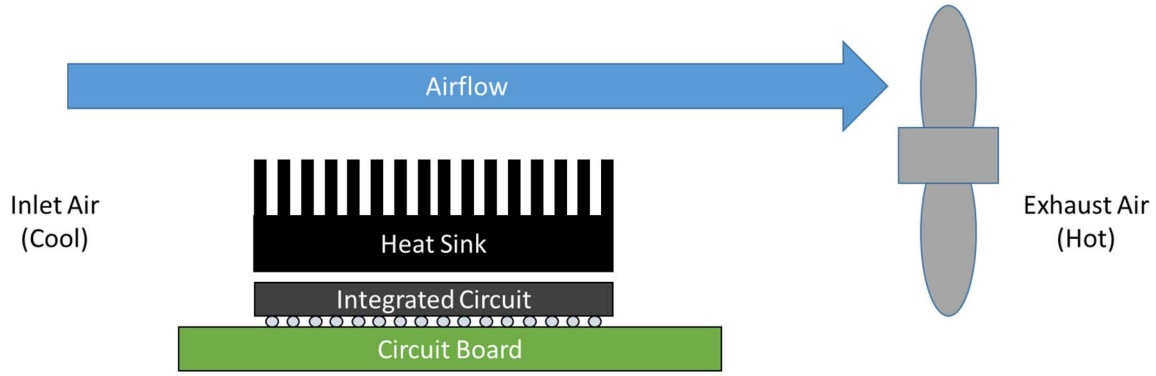

Figure 18: Forced-air cooling showing integrated circuit, fan and heat sink

As shown in figure 18, heat from the chip is transferred to the heat sink. Heat from the heat sink fins is transferred to the surrounding air which is then expelled from the enclosure with the help of fans or blowers.

Because the efficiency of this cooling mechanism depends upon air being forced over the heat sink fins, redundant fans are often deployed to ensure proper airflow even in cases when a single stops working.

It is expected that NFVI compute, storage, and networking node suppliers will provide adequate chip cooling mechanisms for devices within their nodes.

6.6.2.4 Liquid cooling

In some very high-power applications, air may not have sufficient cooling capacity to remove heat from chips. In these cases, liquid may be a more appropriate cooling medium. Liquid cooling solutions also tend to create less ambient noise than air cooling solutions and may offer advantages in reliability since they eliminate the possibility of fan failure.

Liquid cooling solutions vary from cold-plates affixed to components within the rack, to heat exchangers at the rear of rack, or even submersing the entire system in non-conductive coolant. Each of these solutions tends to be more complex than forced-air cooling.

The present document recommends forced air cooling whenever possible for NFVI Node infrastructure.

6.6.2.5 Air filters

In forced-air cooled systems where dust and particulate matter are present, air filters typically form part of the overall cooling solution. Air filters limit accumulation of dust on other system components but need to be serviced at regular intervals in order to maintain the cooling system efficiency. Air filters with internal electrically grounded metallic mesh may also serve as part of the overall strategy to limit radiated emissions from the enclosure.

6.6.3 Cooling reliability

6.6.3.1 Introduction

Hardware redundancy is often employed within the cooling solution in order to prevent component damage or service outages. This clause investigates some common cooling reliability techniques and concepts.

6.6.3.2 Cooling zones

A cooling zone is an area within the NFVI Node that is cooled by the same set of fans such that failure or removal of the fans might impact the equipment within the zone. If fan failure or removal cannot impact the temperature of equipment, it is said to reside in a separate thermal zone.

If all fans providing airflow to a certain thermal zone fail, equipment in that zone needs to be quickly shut down in order to avoid permanent damage. Best practices dictate the presence of redundant fans within the zone to prevent this situation.

6.6.3.3 Fan redundancy

To minimize the impact to equipment when a fan fails, it is customary to include redundant fans within each cooling zone. If any single fan within the cooling zone fails, equipment still receives enough airflow to continue operation. Redundant fans for NFVI Nodes are recommended by the present document.

6.6.3.4 Fan replacement

When a fan fails, it needs to be easily replaceable by service personnel. In order to facilitate this, industry best practices dictate that the fan failure be detectable by the hardware platform management software, and that the fan be serviceable without use of specialized tools. Furthermore, it is desirable that the fan be replaceable without powering down the equipment that it cools.

6.6.4 Cooling safety considerations

The main cooling safety practices can be grouped into the following classifications:

- Proper labelling and warning information;
- Protection from fire/burns;
- Protection from liquid spills (liquid cooling);
- Protection from hazardous chemicals (liquid cooling).

Most countries have safety regulating bodies and requirements. It is recommended that NFVI cooling equipment be developed with compliance to these safety regulations in mind.

6.6.5 Cooling efficiency

Cooling efficiency is the ratio of the cooling capability of the cooling solution versus the energy used to provide the cooling. It is desirable that the cooling system be as efficient as possible.

For forced air cooling, a number of factors impact efficiency including:

- Ambient temperature where the equipment is deployed (lower is better);
- Air density (higher is better);
- Fan size (larger is better);
- Fan speed control algorithm;

Cooling efficiency may present trade-offs against other system attributes, such as reliability and manageability. For instance, small fans located in each compute/storage/network node will have poorer reliability and efficiency than larger fans located within the rack. However, a solution with large rack fans introduces another layer of hardware management complexity in order to ensure that adequate airflow is always provided to each resource node.

The present document recommends that these trade-offs be examined when determining the overall equipment cooling strategy.

6.6.6 Example from Open Compute Project

Open Rack from the Open Compute Project accommodates sized payload trays that are multiples of 48 mm in height. It is expected that each of the payload trays provide their own individual cooling solution. An example implementation is shown in figure 19.

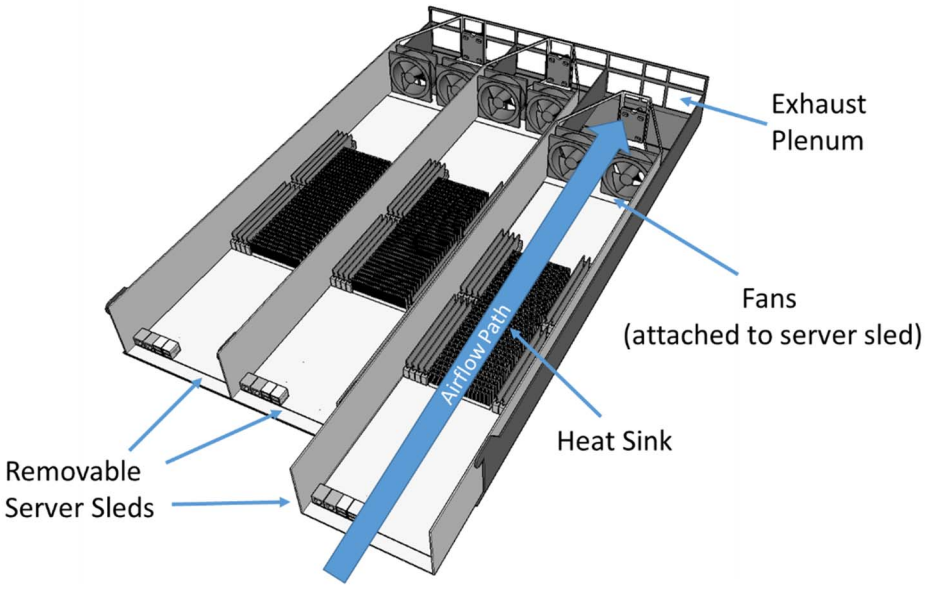

Figure 19: Simplified representative Open Compute servers showing cooling components and airflow path

In this example, three server sleds can be inserted into a single enclosure tray. Each server includes heat sinks on high power devices as well as two 40 mm fans to move air across the server motherboard. In the event of a fan failure, the server can be powered down and the failing fan replaced.

6.6.7 Cooling recommendations

The present document makes the following recommendations with regard to NFVI Node cooling:

- Air forced air cooling is preferred;
- Cooling system supports a variety of installation environments including datacenter central office;
- Air filter options are available for installations that require them;
- Adequate fault protection is present in the cooling solution in order to prevent equipment failure and ensure reliability;
- Fan replacement is possible without specialized tools;
- Relevant global, regional and industry regulations are met.

6.7 Hardware platform management

6.7.1 Introduction

Hardware platform management is a set of operations required to access information about a collection of physical resources, and to perform very basic system management (such as setting fan speeds, taking a system in/out of service, reading manufacturer information, installing firmware, etc.). A NFVINode is expected to provide software APIs to perform a basic level of hardware platform management.

The principles of NFV enabling software ought to apply to the very low level functions described in this clause. It is recommended that specifications and implementations be open, i.e. not proprietary. Re-use of existing technology ought to be considered. Simplicity is essential.

This clause aims to set out recommendations for a minimum set of hardware platform management functions, without going into the detail of implementation.

Hardware platform management will support NFVI failure notification as specified in ETSI GS NFV-REL 003 [i.13].

6.7.2 Typical hardware elements managed via software API

6.7.2.1 Environmental sensors and controls

Commercial hardware platforms typically contain multiple sensors that are essential to monitor a node's physical environment, for example: temperature, voltage, airflow. Controls include, for example, CPU watchdog timer, payload input voltage. The hardware platform management function would provide an API to query the value of these sensors, and issue commands to lower level managers to manipulate controls.

6.7.2.2 Boot/Power

It is desired that the hardware platform management function includes a software interface to control power on/off commands to individual payload elements, including hot swap controls (if present).

6.7.2.3 Cooling/Fans

It is desired that the hardware platform management function includes a software interface to report on the current state of the cooling system (whether fans or some other mechanism) and vary this performance based on environmental factors (such as CPU temperature, ambient temperature, available airflow).

6.7.2.4 Network status

It is desired that the hardware platform management function includes a software interface to report the status of physical connectivity ports. This should include both active (i.e. receive an alert) and passive (poll/query) modes.

6.7.2.5 Inventory data repository

It is desired that the hardware platform management function includes a software interface to query FRU data, such as manufacturer, model, serial number, firmware revision. The FRU data may be used by the node management function to auto-discover a node's contents and, therefore, its capabilities. The inventory data repository is often referred to as 'FRU data'.

6.7.2.6 Firmware upgrade

It is desired that the hardware platform Management function includes the software interfaces to carry out a remote firmware (FW) Upgrade including determining current FW revision, downloading or acquiring a new FW image, placing the node, or a subset of the node, into an appropriate state (powered-on but unavailable for VNF functions), replacing the FW, checking FW integrity, re-booting (as well as security functions to prohibit unauthorized and malicious update attempts).

6.7.2.7 Event logging and diagnostics

It is desired that low level hardware platform errors, events and notifications are logged, and made available to a higher level management function.

6.7.2.8 Alarm management

It is desired that the hardware platform management function provides a software interface to manage alarms which are raised by lower level management functions and drivers. In many cases, this could overlap with event logging and diagnostics. It is recommended that alarm management include a software interface to manipulate any visible/audible alarm mechanisms available on the node.

6.7.3 Hardware platform management features

6.7.3.1 General

It is desired that the NFVI Node presents a single interface for accessing management functions.

It is desired that any application or function which implements this interface will feature redundancy so as not to become a single point of failure for the system.

It is not required for an NFVI Node to implement a distinct intra-node network for management purposes, although this could be implemented if the application use case required.

6.7.3.2 System management

The system management function is an application which aggregates hardware platform management from multiple individual NFVI Nodes. As such, it is outside the scope of the present document and is included for information purposes only (as the consumer of hardware platform management functions).

6.7.3.3 Node management

Node management is a function of a single 'Rack' network element – generally, a collection of compute, networking and/or storage resources within a single enclosure and having connectivity between them. In this clause, Ethernet connectivity is assumed. The node management function is responsible for collating all lower-level FRU hardware management information and commands. The node management function provides a single point of contact for the entire NFVI Node.

It is desired that the node management function supports the following:

1) FRU discovery - automatically detecting all hardware FRU elements in the system and populating a database with that information;

2) Accepting requests for hardware platform information from a higher level management entity (via an Ethernet interface to that external entity);

3) Querying FRU hardware platform information from lower level FRU hardware;

4) Maintaining status information regarding lower level FRU hardware;

5) Responding to higher level management entity requests;

6) Issuing unsolicited status information to a higher level management entity, in response to an error condition.

It is desired that the node management function runs on a general purpose processor and not rely on (for example) a bespoke shelf management controller.

To enable high availability, the node manager ought to support both hardware and software redundancy. If it is implemented as an application running within a FRU, the application should have redundant instances, running on separate FRU instances with no single point of failure.

6.7.3.4 Power and fan management

It is desired that fan (hardware) management be enabled via a software API to read and control current Fan speed.

It is desired that power management be enabled via a software API to monitor voltage and current available from the in-node power supply and the voltage/current available at each FRU.

NOTE: The power and fan hardware management function are not necessarily implemented as a single function.

6.7.3.5 Network management interface

The network interface hardware status is a special case for payload management, being used to monitor the status of network equipment specifically. Management at this level is limited to those functions which allow a higher level management entity to determine the status of network ports.

A typical set of functions (made available via the software API) would be:

- querying a network port to determine status, alerting a network port status change (link up/link down/link failed).

6.7.3.6 Payload management interface

6.7.3.6.1 Introduction

It is desired that hardware management of each payload be made available via a software API implemented on each payload type. The minimum requirements for each payload type are very similar, but the implementation may be very different.

6.7.3.6.2 Compute payload management interface

A compute payload hardware management function would typically include:

1) Access to FRU information;
2) Power on/off controls (including Hot swap);
3) Firmware upgrade;
4) Access to sensors and controls within that payload;
5) Event logging.

6.7.3.6.3 Storage payload management interface

A storage payload hardware management function would typically include:

1) Access to FRU information;
2) Power on/off controls (including hot swap);
3) Access to sensors and controls within that payload;
4) Event logging.

6.7.4 Recommendations

Further study will be required to produce a comprehensive set of recommendations for hardware platform management. Topics for specific consideration include:

1) The availability of open standards for hardware platform management in existing and legacy hardware platform form factors, and the applicability of those standards to NFVI Nodes;
2) A recommended hierarchy and logical structure of hardware management functions (e.g. splitting out compute, storage and network node requirements). This hierarchy may be implicit in the way that an NFVI Node is architected, with hardware management functions cascading down within the system, but requires further investigation to ensure all recommended NFVI Node structures are catered for;
3) The specific functions to be recommended as the minimum set of hardware platform management APIs;
4) The availability of Open Source implementations for APIs and drivers recommended in this clause;
5) Commercially available implementations and any implications this may cause on adoption (such as IPR ownership).

7 NFVI Node examples

7.1 Introduction

NFVI hardware is comprised of network nodes, gateway nodes, compute nodes and storage nodes that are combined in order to provide the resources needed by the VNFs. As much as possible, these nodes are implemented by commercial hardware components. Previous clauses in the present document outlined features and guidelines for the hardware, including interconnect, power, and cooling recommendations. This clause provides a brief discussion on how the various physical hardware components can be combined to support various VNFs. It is the explicit desire that commercial compute, storage, network and gateway node hardware, available from multiple vendors, can be combined together to support various types of NFVI Nodes.

While there are many different types of VNFs, the present document envisions a common architecture can support most, if not all, of them. A simplified representation of this architecture is shown in figure 20. In this architecture, a plurality of network and gateway nodes provides physical connectivity between the various other nodes. Although VNFs may create and reconfigure virtual links as required, the physical connectivity between the nodes is expected to be reconfigured manually and therefore remains relatively static. This relationship of static hardware configuration versus dynamic virtual configuration is a key consideration for the deployment of NFVI Node hardware.

The compute and storage domain demonstrates similar flexibility. Here, a plurality of compute and storage nodes provides resource functionality to the VNFs on an as-needed basis. When VNFs require more compute/storage capacity, additional resource nodes may be logically allocated to the application. As with the infrastructure network, however, these resources are already physically present within the system and deployment of new physical resources will remain fairly static compared to logical allocation to VNFs.

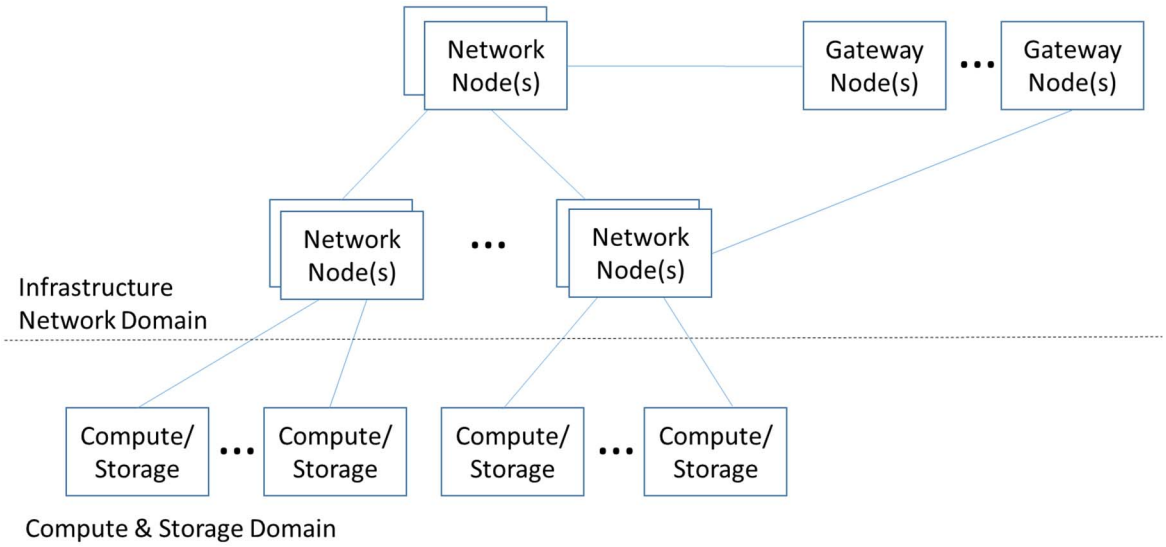

Figure 20: Simplified canonical hardware architecture for NFVI

A common NFVI Node hardware architecture is important because it facilitates the use of common components, simplifying sparing and improving economies of scale.

7.2 Virtual mobile network

The ETSI ISG NFV virtualised mobile network proof-of-concept [i.14] demonstrates how these common hardware elements and architecture can be utilized to implement the key functions of an Evolved Packet Core (EPC). Typical elements of an EPC and their relationship with the rest of the network is shown in figure 21. The virtualised mobile network proof of concept includes eNodeB, mobility management entity (MME), Serving Gateway (SGW) and packet data network gateway (PGW) functions well as a Diameter signalling controller (DSC) virtual network function [i.16].

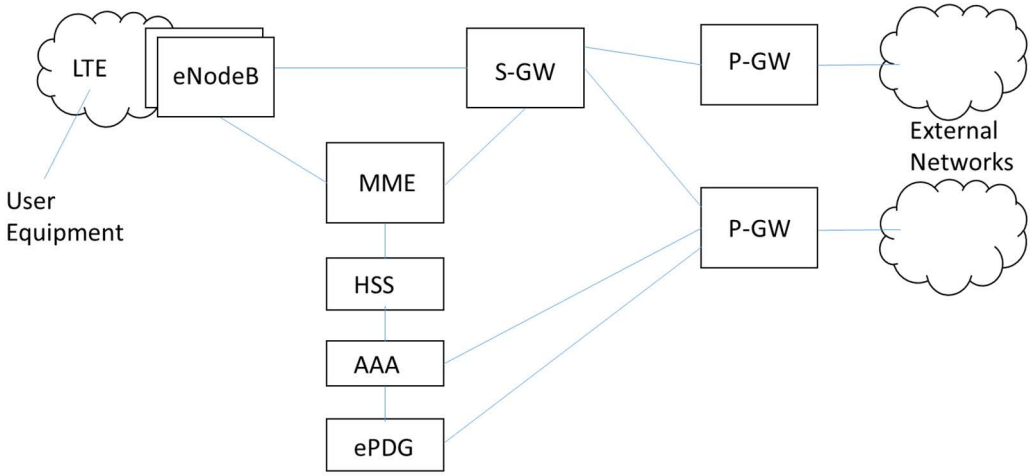

Figure 21: Typical elements of Evolved Packet Core within mobile network

The hardware configuration used for this proof of concept is shown in figure 22.

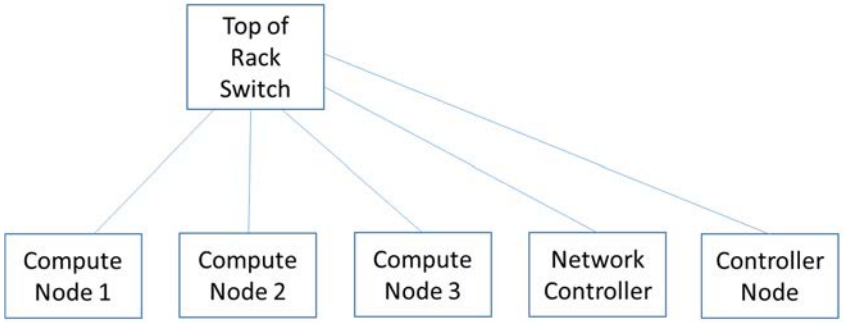

Figure 22: Hardware elements of virtual mobile network NFVI Node

For consistency with the proof of concept documentation, the terms "Network Controller" and "Controller Node" have been included in the previous figure. These, however, are not new types of NFVI Nodes, but rather, logical functions that serve to control the virtual environment. These functions were implemented on commercial processing hardware. In field deployments, NFVI control/management functions would likely serve more than one NFVI Node.

In the demo, three identical 8-core compute nodes provide resource functionality to the VNFs. Both the eNodeB and SGW/PGW applications are allowed to scale with capacity by adding additional VNF instances. The other VNFs are statically allocated. A top of rack switch (Network Node) provides network functionality and is controlled by a network controller function running OpenStack Neutron on another server. The controller node function provides the remainder of the OpenStack control services for the proof-of-concept and runs on another server.

Additional information about this application and proof-of-concept can be found in [i.14].

7.3 Access node

This clause discusses how a telecommunications access architecture can be refactored in fit within the NFVI Node architecture. It is based on information originally presented at the Open Compute Engineering workshop in March 2015 [i.15].

A typical legacy access architecture is shown in figure 23. A gigabit passive optical network (GPON) optical line terminator (OLT) provides termination to multiple customer broadband and telecommunications endpoints and serves as a first-level aggregation point. Multiple OLTs are aggregated together by an Ethernet aggregation switch, which in turn forwards the traffic to a broadband network gateway. This is shown in figure 23.

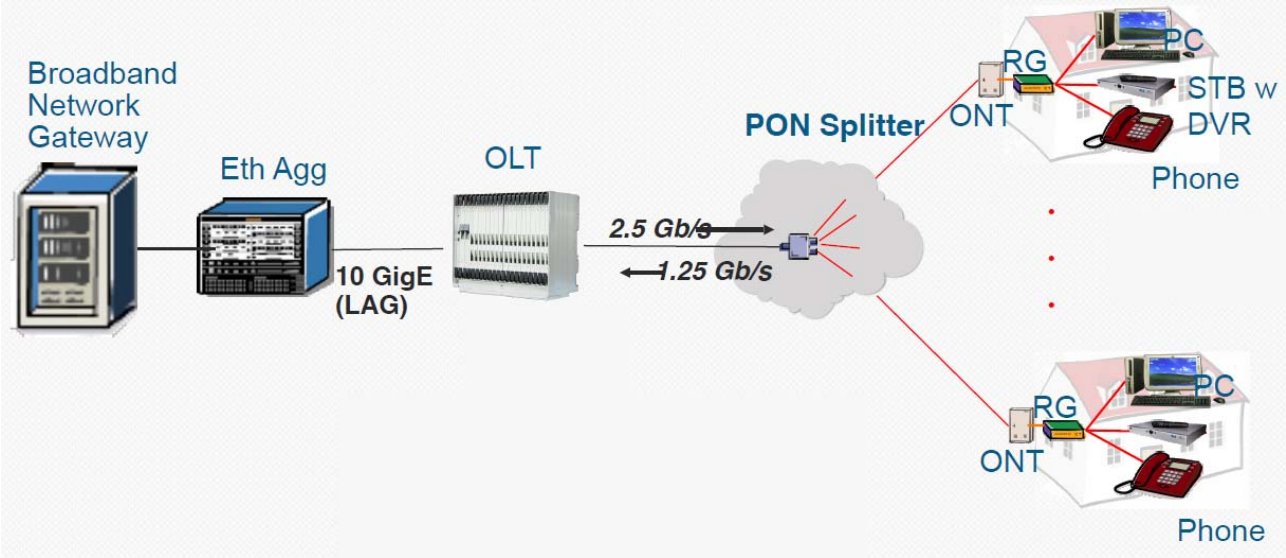

Figure 23: Legacy GPON access architecture

This architecture can be refactored into one that uses commercial network and compute nodes with the resulting architecture looking very similar to figure 20. Management and routing functions are placed in VNFs on compute nodes, switching functions map to network nodes. Gateway nodes provide I/O links to upstream and downstream elements within the network. This is shown in figure 24.

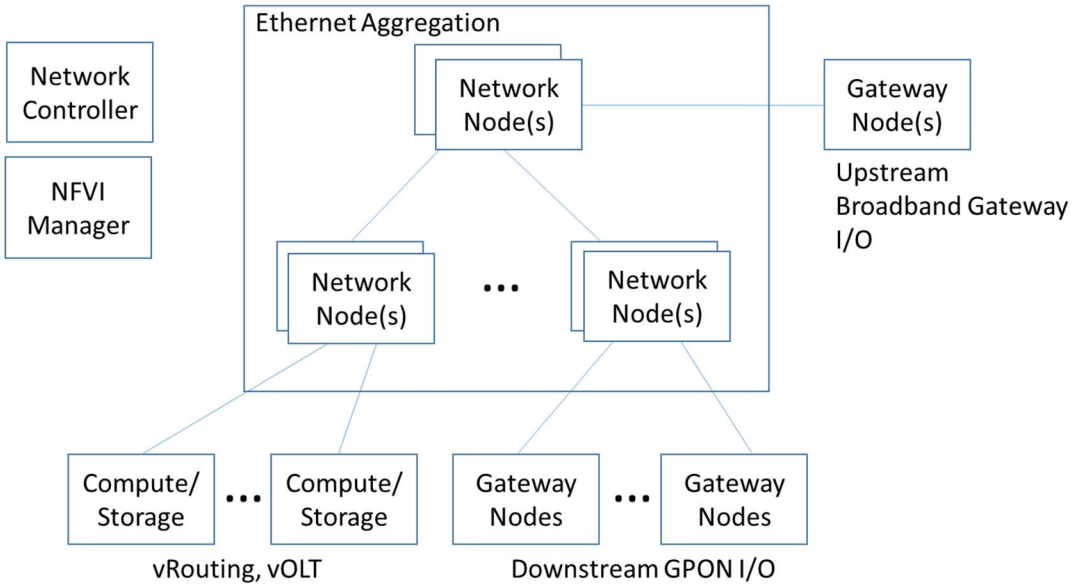

Figure 24: Architecture of access NFVI Node

All elements within this architecture can be implemented with commonly available commercial equipment with the exception of the GPON gateway nodes. These, however, can be constructed from merchant silicon. Common interconnects, drivers, and mechanical form-factor of this device would facilitate a multi-vendor commercial ecosystem.

More information about this application can be found in [i.15].

7.4 Transport node

A transport network element serves the function of routing high volumes of traffic within the operator's network. Optical transport nodes switch traffic based on wavelengths of light. Packet transport nodes, switch traffic based on packet header and flow information. Historically optical and packet network elements have been distinct pieces of equipment, however, more recently there has been a trend to converge both types of transport into the same physical device.

A typical transport element consists of line cards, which receive and transmit traffic; fabric switches, which direct traffic from one line card to another; and control elements that control the operation of the transport element. This is shown in figure 25.

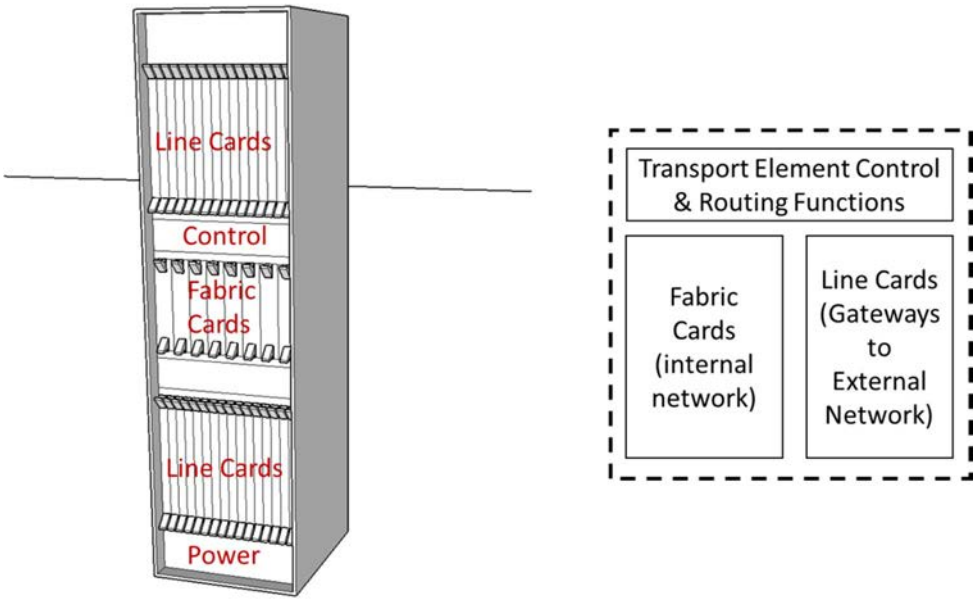

Figure 25: Representative transport network element

This network element architecture can be refactored nicely into the NFVI Node architecture. The switch fabric maps onto network nodes and the line cards map onto gateway nodes. The vTransport control and routing functions map onto one or more compute nodes. This is shown in figure 26.

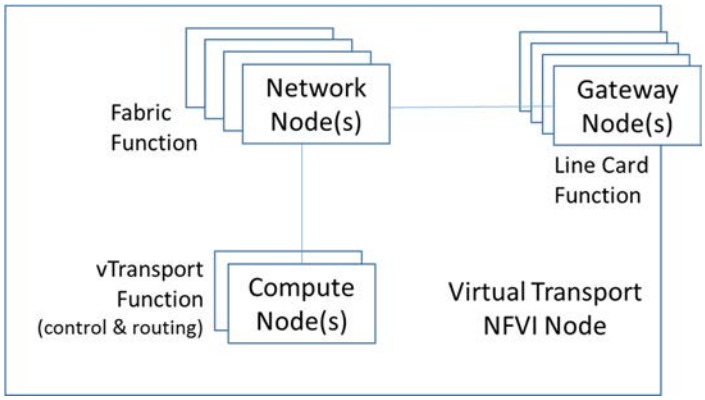

Figure 26: Virtual transport NFVI Node hardware decomposition showing network, gateway and compute nodes

The switching requirements for a high-capacity transport network element are quite stringent. It is envisioned that today's commercial enterprise-grade switches are not capable of meeting all the requirements; however, lower capacity vTransport devices could be built using commonly available commercial switches. Higher performance switches can be developed using available merchant silicon. Standardization of interface requirements could facilitate the emergence of an ecosystem of commercial components for higher performance transport deployments.

7.5 Customer Premises Equipment

In historical network deployments, equipment placed on the customer premises provided the primary gateway to the operator network as well as services related to voice, video, and data. Multiple customer premises devices might have been present, each providing functions and interfaces for the specific services supported. This is shown in figure 27.

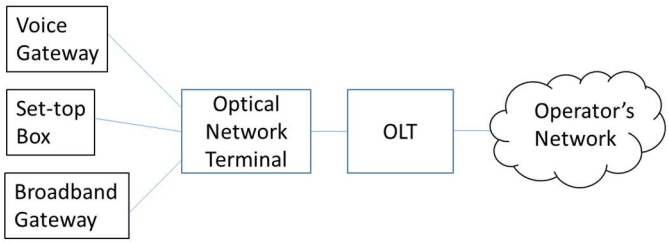

Figure 27: Traditional Customer Premises Equipment

By virtualising, the customer premises equipment, the services are moved into the operator's cloud leaving only a very simple gateway device at the customer site. This facilitates network agility by allowing new customers and services to be added virtually. This configuration is shown in figure 28.

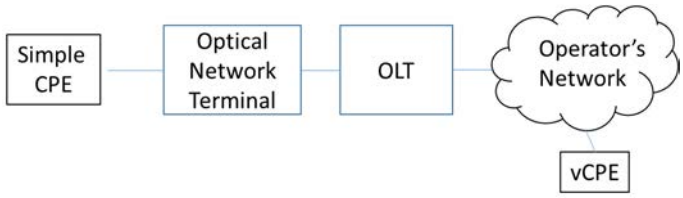

Figure 28: Virtual Customer Premises Equipment

Implementation of virtual Customer Premises Equipment with NFVI Node hardware is straightforward. Network and gateway functions provide connectivity between the vCPE functions (running on compute nodes) and the Simple CPE gateway nodes located at the customer site. This is shown in figure 29.

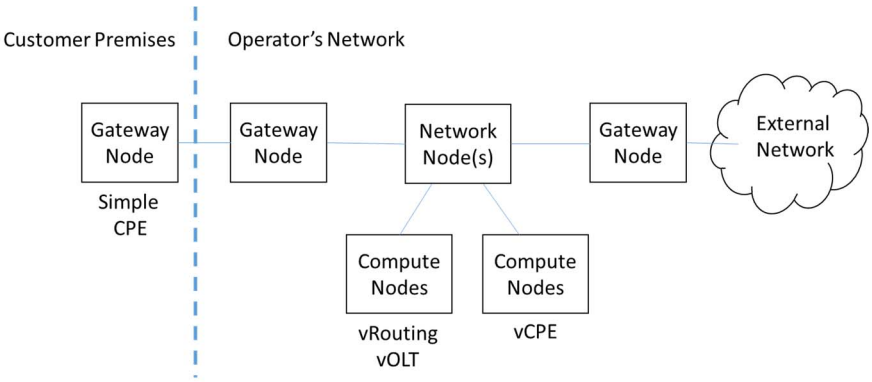

Figure 29: Virtual customer Premises Equipment using NFVI Node hardware

Annex A (informative): Bibliography

Open Compute Project™.

NOTE: Available at http://www.opencompute.org/.

The Xen Project™.

NOTE: Available at http://www.xenproject.org/.

Annex B (informative): Authors & Contributors

The following people have contributed to this specification:

Rapporteur:

Percy S. Tarapore, AT&T

Contributors:

Cheng Chen, Huawei

Don Clarke, CableLabs

Jim Darroch, Artesyn

Hiroshi Dempo, NEC

Doug Sandy, Artesyn

Xu Yang, Huawei

Valerie Young, Intel

History

Document history		
V1.1.1	January 2016	Publication