INTERNATIONAL TELECOMMUNICATION UNION

TELECOMMUNICATION STANDARDIZATION SECTOR

STUDY PERIOD 2017-2020

SG13-C918

STUDY GROUP 13

Original: English

Question(s):	19/13	Virtual, 20 - 31 July 2020

CONTRIBUTION

Source:	Telecommunication Engineering Centre, Ministry of Communications (India)
Title:	Draft ITU-T Recommendation Y.e2efapm: " Cloud Computing - End-to-end fault and performance management framework of inter-cloud network services": Proposal to add some terms and abbreviations.
Purpose:	Proposal

Contact:	Abhay Shanker Verma Telecommunication Engineering Centre India	Tel:+91 9999554900 Email:as.verma@gov.in

Keywords: end-to-end; fault; inter-cloud; management; NF; NS; N-PoPs; performance; Y.e2efapm; 20-31 July 2020

Abstract: This contribution proposes to add some terms and abbreviations in the draft ITU-T Recommendation Y.e2efapm: “Cloud Computing - End-to-end fault and performance management framework of inter-cloud network services”.

1. Introduction

While going through the draft ITU-T Recommendation Y.e2efapm (**TD540-WP2**), it was felt that some terms need to be defined for improved readability and more clarity.

Accordingly, in this contribution, it is proposed to add some terms defined elsewhere in clause 3.1. Further, the reference to ETSI GS NFV 003, wherever appearing may be suitably replaced with a latest document ETSI GR NFV 003. It is proposed to add them into Bibliography of Y.e2efapm. At the same time, some abbreviations and acronyms should be added in clause 4. In addition, general formatting of the document has been done.

2. Proposal

It is proposed to modify clause 3, clause 4 and Bibliography of Y.e2efapm. The proposed modifications are in track change mode in **Annexure-I**.

3. Reference

[1] T17-SG13-200313-TD-WP2-0540!!MSW-E : Base Document for this contribution

[2] [ETSI GR NFV 003] ETSI GR NFV 003 V1.5.1 (2020-01), *Network functions virtualisation (NFV); Terminology for main concepts in NFV.*

*****

Annexure-I

Introduction

Cloud computing is an essential ingredient of all modern telecommunications services, including 5G. A use case that service providers are globally interested in, is deployment of their service offerings as Network Services (NS), using Network Function Virtualization (NFV), over multiple clouds. This gives them a number of advantages including freedom from proprietary solutions, reduced time to market, agility of service, proximity to customers and lower cost of deployment and operation. However, today the virtual deployments do not match the five nines (99.999%) availability, or the performance of the traditional physical networks based on dedicated and custom-built integrated hardware and software. A standards based Fault, Configuration, Accounting, Performance and Security (FCAPS) framework for NS over multiple clouds would help attain the level of availability and performance that service providers and subscribers expect from the traditional networks. This recommendation focuses mainly on the Fault and Performance (FP) aspects of NS deployments. For these aspects alone, ensuring proper operation of the NS is more complex, as compared to traditional services, because of two main reasons: a) more layers of abstraction i.e. physical, virtual resources, virtual network functions, service function chains and NSs, and b) complex interaction of the involved management platforms, i.e., Inter-Cloud management platform (MCMP) of the cloud service provider, Operation Support Systems (OSS) of the service provider and Management and Orchestration platform (MANO) of NFV. These platforms together have the responsibility of managing the Inter-Cloud resources and the life cycles of NSs and their components. For this, the FP management functionality must collect and process all the alarms, notifications and performance metrics from different layers, e.g., NS, SFC, VNF and EMS (**Note:** somewhere in the description we may refer to criticality of the alarms as defined ITU X.733 Recommendations). Four Critical aspects of end-to-end fault & performance management system are:

i) Fault and Performance issues detection sub-system: carries out detection of fault & performance issues, both impending and manifest faults. This is done in two steps: Step 1 involves classification of a situation as ‘fault’ or ‘no-fault’ and Step 2 involves further classification of fault problems as ‘manifest’ or ‘impending.

ii) Fault and performance localization sub-system: carries out localization of manifest faults in two steps: coarse-grain and fine-grain localization. For impending faults, it predicts the intensity and likely location of the problem.

iii) Performance Management: Fix the fault that degrades network performance i.e. troubleshoot fault to restore network performance to original or improved condition;

iv) Maintaining QoS (Quality of Service): Adhere to SLA (Service Level Agreement) for achieving 99.999% availability of network & business critical applications. (Availability requirement of service provider is five nines. Subscribers may have their own SLAs)

Draft new Recommendation ITU-T Y.e2efapm

Cloud Computing – End-to-end fault and performance management framework of network services in inter-cloud

Summary

This recommendation provides end-to-end fault and performance management framework of network services (NSs) in inter-cloud computing and relevant use cases. In particular, the aspects of faults detection and localization of affected area in inter-cloud environments is presented.

Keywords: inter-cloud, end-to-end, fault, performance, management

Table of Content

Draft new Recommendation ITU-T Y.e2efapm

Cloud Computing – End-to-end fault and performance management framework of network services in inter-cloud

1. Scope

This Recommendation specifies an end-to-end fault and performance management framework and relevant use cases of network services (NSs) in inter-cloud computing. The scope of this Recommendation includes:

- overview of end-to-end fault and performance management of NSs;
- functional requirements of end-to-end fault and performance management of NSs;
- use cases relevant to end-to-end fault and performance management of NSs.

2. References

The following ITU-T Recommendations and other references contain provisions, which, through reference in this text, constitute provisions of this Recommendation. At the time of publication, the editions indicated were valid. All Recommendations and other references are subject to revision; users of this Recommendation are therefore encouraged to investigate the possibility of applying the most recent edition of the Recommendations and other references listed below. A list of the currently valid ITU-T Recommendations is regularly published. The reference to a document within this Recommendation does not give it, as a stand-alone document, the status of a Recommendation.

[ITU-T X.733] Recommendation ITU-T X.733 (1992), *Information technology – Open Systems Interconnection – Systems Management: Alarm reporting function*

[ITU-T Y.3500] Recommendation ITU-T Y.3500 (2014) | ISO/IEC 17788:2014, *Information technology – Cloud computing – Overview and vocabulary*

[ITU-T Y.3501] Recommendation ITU-T Y.3501 (2013), *Cloud computing framework and high-level requirements*

[ITU-T Y.3502] Recommendation ITU-T Y.3502 (2014) | ISO/IEC 17789:2014, *Information technology – Cloud computing – Reference architecture*

[ITU-T Y.3503] Recommendation ITU-T Y.3503 (2014), *Requirements for Desktop As A Service*

[ITU-T Y.3510] Recommendation ITU-T Y.3510 (2016), *Cloud computing infrastructure requirements*

[ITU-T Y.3512] Recommendation ITU-T Y.3512 (2014), *Cloud computing - functional requirements of Network As A Service*

[ITU-T Y.3513] Recommendation ITU-T Y.3513 (2014), *Cloud computing - functional requirements of Infrastructure-As-A-Service*

[ITU-T Y.3515] Recommendation ITU-T Y.3515 (2017), *Cloud computing – Functional architecture of Network as a Service*

3. Definitions

3.1. Terms defined elsewhere

This Recommendation uses the following terms defined elsewhere:

3.1.1 cloud service [ITU-T Y.3500]: One or more capabilities offered via cloud computing invoked using a defined interface. It may comprise the hardware & hypervisor layers delivering individual servers, border routers, firewalls, load balancers & switches.

3.1.2 cloud service customer [ITU-T Y.3500]: Party which is in a business relationship for the purpose of using cloud services.

3.1.3 cloud service provider [ITU-T Y.3502]: party which makes cloud services available.

3.1.4 hypervisor [ITU-T Y.3510]: A type of system software that allows multiple operating systems to share a single hardware host.

3.1.5 Infrastructure as a Service (IaaS) [ITU-T Y.3500]: Cloud service category in which the cloud capabilities type provided to the cloud service customer is an infrastructure capabilities type.

NOTE – The cloud service customer does not manage or control the underlying physical and virtual resources, but does have control over operating systems, storage, and deployed applications that use the physical and virtual resources. The cloud service customer may also have limited ability to control certain networking components (e.g., host firewalls).

3.1.6 Network as a Service (NaaS) [ITU-T Y.3500]: Cloud service category in which the capability provided to the cloud service customer is transport connectivity and related network capabilities.

NOTE – NaaS can provide any of the three cloud capabilities types.

3.1.7 network function (NF) [b-ETSI GR NFV 003]**:** Functional block within a network infrastructure that has well-defined external interfaces and well-defined functional behaviour.

NOTE – In practical terms, a Network Function is today often a network node or physical appliance.

3.1.8 network functions virtualization (NFV) [b-ETSI GS NFV 003]: Principle of separating network functions from the hardware they run on by using virtual hardware abstraction.

3.1.~~8~~9 network functions virtualisation infrastructure (NFVI) [b-ETSI GS NFV 003]: Totality of all hardware and software components that build up the environment in which VNFs are deployed.

NOTE –~~:~~The NFV-Infrastructure can span across several locations, e.g. places where data centres are operated. The network providing connectivity between these locations is regarded to be part of the NFV-Infrastructure. NFV-Infrastructure and VNF are the top-level conceptual entities in the scope of Network Function Virtualisation. All other components are sub-entities of these two main entities.

3.1.10 network point of presence (N-PoP) [b-ETSI GR NFV 003]: Location, where a Network Function is implemented as either a Physical Network Function (PNF) or a Virtual Network Function (VNF).

3.1.~~9~~11 network service [ITU-T Y.3515]: A collection of network functions with a well specified behaviour.

NOTE – Examples of network services include content delivery networks (CDNs) and IP multimedia subsystem (IMS).

See also [b-ETSI GR NFV 003], which defines network service (NS) as composition of network function(s) and/or network service(s), defined by its functional and behavioural specification.

NOTE – The Network Service contributes to the behaviour of the higher layer service, which is characterized by at least performance, dependability, and security specifications. The end-to-end network service behaviour is the result of the combination of the individual network function behaviours as well as the behaviours of the network infrastructure composition mechanism.

3.1.~~10~~12party [ITU-T Y.3500]: Natural person or legal person, whether or not incorporated, or a group of either.

3.1.~~11~~13 virtualized network function [ITU-T Y.3515]: A network function that can be deployed as a software on a NaaS cloud service provider infrastructure.

NOTE – Examples of virtualized network functions include virtual switches and virtual routers.

3.2 Terms defined in this Recommendation

This Recommendation defines the following terms:

4. Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

API	Application Programming Interface
CDN	Content Delivery Network
CMIP	Common Management Information Protocol
CPE	Customer Premises Equipment
CSC	Cloud Service Customer
CSP	Cloud Service Provider
DaaS	Desktop as a Service
DAS	Direct-Attached storage
DoT	Department of Telecommunications
EMS	Element Management System
FC	Fibre Channel
FCAPS	Fault, Configuration, Accounting, Performance and Security
IaaS	Infrastructure as a Service
IOT	Internet of |Things
IP	Internet Protocol
N-PoP	Network Point of Presence
NaaS	Network as a Service
NF	Network Function
NFV	Network Function Virtualisation
NFVI	Network Functions Virtualisation Infrastructure

NFVIaaS	Network Function Virtualisation Infrastructure as a Service
NFV-MANO	NFV Management Orchestrator
NS	Network Service
OSS/BSS	Operations Support Systems/Business Support Systems
PaaS	Platform as a Service
P/PE router	Provider/Provider Edge router
PNF	Physical Network Function
PoP	Point of Presence
QoS	Quality of Service
SD-WAN	Software Defined Wide Area Network
SFC	Service Function Chain
SLA	Service Level Agreement
SSL	Secure Socket Layer
SVM	Support Vector Machine
TEC	Telecommunication Engineering Centre
TSP	Telecom Service |Provider
vCPE	Virtual Customer Premises Equipment
VNF	Virtual Network Function
VNFaaS	VNF as a Service
VNPaaS	Virtual Network Platform as a Service
VLAN	Virtual LAN
VM	Virtual Machine
VNFaaS	Virtual Network Function (VNF) as a Service
VPN	Virtual Private Network
WAN	Wide Area Network

5. Conventions

In this Recommendation:

The keywords “**is required to**” indicate a requirement which must be strictly followed and from which no deviation is permitted if conformance to this document is to be claimed.

The keywords “**is prohibited from**” indicate a requirement which must be strictly followed and from which no deviation is permitted if conformance to this document is to be claimed.

The keywords “**is recommended**” indicate a requirement which is recommended but which is not absolutely required. Thus this requirement need not be present to claim conformance.

The keywords "**is not recommended**" indicate a requirement which is not recommended but which is not specifically prohibited. Thus, conformance with this specification can still be claimed even if this requirement is present.

The keywords "**can optionally**" indicate an optional requirement which is permissible, without implying any sense of being recommended. This term is not intended to imply that the vendor's implementation must provide the option and the feature can be optionally enabled by the network operator/service provider. Rather, it means the vendor may optionally provide the feature and still claim conformance with the specification.

6. Overview of end-to end fault and performance management of network services

[Contributor's note] This clause is providing overview of fault and performance management of NSs.

[Editor's note in March 2020]: The draft Rec. Y.e2efafm should be check about references to legacy technologies and technics, which are out of scope of work this draft Recs. Especially, aspects of ITU-T standards about events and alarms should be verified. Contributions are invited.

[Editor's note in March 2020]: Figure 6-1 and Figure 6-2 should be redraw. Contributions are invited.

6.1. Background

The traditional telecommunication network deployment largely involves use of physical network appliances like routers, switches, broadband remote access servers, and middle-boxes like firewalls, deep packet inspectors or load balancers. These integrated hardware and software solutions are normally closed and proprietary leading to vendor lock-in, thereby making expansions and deployment of new services difficult and time consuming. Such equipment are also not amenable to easy scaling or redeployment of resources. The power and space requirements as well as the total cost of operation are higher in physical element based networks.

In traditional networks, time-tested standards relating to fault, configuration, accounting, performance and security (FCAPS) are embodied in ISO Common Management Information Protocol (CMIP) and ITU TMN M.3010 and M.3400 recommendations. Network management based on relevant standards provides five nines (99.999%) availability and carrier grade reliability.

Inter-cloud computing, coupled with network function virtualization (NFV), provides numerous advantages to cloud service providers (CSPs) including ease of deployment, ease of scaling, ease of introducing and switching off services and reduced cost of operation. This may increase viability of telecommunications business and lead to thriving telecommunication sectors. However, there are a number of reasons as to why the combination of inter-cloud & NFV i.e. inter-cloud NS needs a strong fault & performance management system to be a viable replacement for traditional networks. For carrier grade availability & reliability of up-to five nines (99.999%) for inter-cloud NS, there is a need for standardization of techniques for fault and performance detection and localization to deal with complexity in such networks as the anomalous behaviour could be in the hardware, virtual machines, virtual network functions (VNFs), service chains (SCs) or at the service levels.

6.2. Network Services in the virtualized environment

[Editor's note in March 2020] This clause introduces the concept of NS. Figure 6-1 and Figure 6-2 should be redraw. Contributions are invited.

According to [ITU-T Y.3515], NS is a collection of network functions with a well specified behaviour, examples of network services include content delivery networks (CDNs) and IP multimedia subsystem (IMS). When supporting for NaaS connectivity services, NS can be described as an abstracted transport connectivity between two end points in a virtualised environment where the end points may be located in one or more clouds. A NS utilises a SC or Virtual Network Function Forwarding Graph (VNFFG) for interconnecting virtual network functions end to end.

A NS can be described as an end to end implementation using SFC or VNFFG, interconnecting the virtual network resources. SFC or VNFFG is an ordered set of VNFs in the virtual environment that represent functions like routers and broadband network gateways or middle-boxes like load balancers and firewalls, which act on the traffic in the sequence they appear in the chain. Such VNFs are hosted on VMs instantiated over physical data centre and network resources. An example of end-to-end Inter-Cloud NS is shown in figure below.

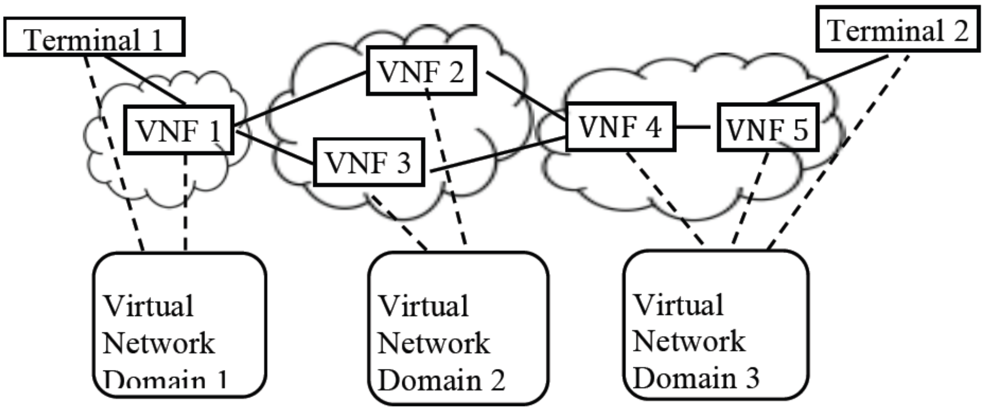


Figure 6-1 – Inter domain end to end network service

The detailed network and VNF connectivity diagram in Inter-Cloud scenario is depicted below in figure-8. In this diagram, two VNFs in first cloud, two VNFs in second cloud and three VNFs in third cloud are connected as per the VNF Graph to provide the end to end network service.

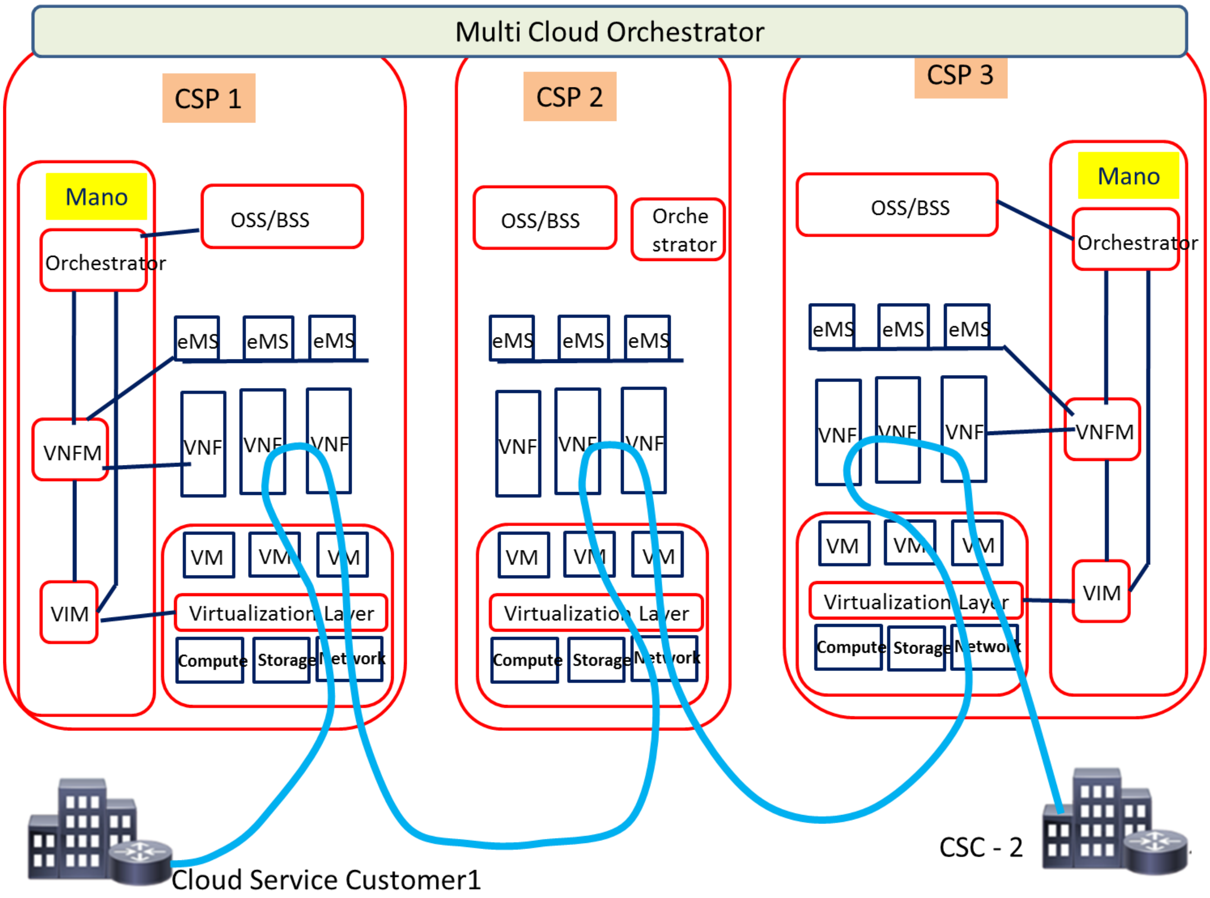


Figure 6-2 – VNF Forwarding Graph for Inter-cloud end to end service

6.3. Challenges of network service management in inter-cloud

The telecommunication's networks have traditionally been designed to provide high availability and standards-based quality of service. In inter-cloud, network services deployment over multiple clouds identifies new challenges to equip inter-cloud management systems to deal with management issues. Especially, those NSs relay over underlying both physical and "software" infrastructure (NFV-based infrastructure). Therefore, the end to end management is related to physical, virtual layer or the VNFs of inter-cloud environments where virtual machines are instantiated, on which particular VNFs are hosted.

In fact, the traditional deterministic methods fail to deliver in virtual environments in which virtual resources can be dynamically scaled, migrated or destroyed. It is important to use predictive techniques to identify and resolve management issues before or after they have occurred.

In hybrid telecommunication networks with physical and "software" infrastructures, the deterministic methods ensure carrier grade availability and reliability. On the other side, the NSs using virtual resources over multiple clouds provide a number of complex factors and make it imperative to use predictive methods for assuring carrier grade availability.

Some of the key challenges for NS management in inter-cloud are as follows:

- Absence of an FCAPS framework.
- Non-applicability of traditional rule based techniques when used in today's networks.
- Multiple layers of implementation: physical infrastructure, NFVI, VNF and NSs.
- Massive distribution of network functions over disparate clouds.

- Multiple control centres: cloud management systems, operators' OSS/BSS and NFV-MANO and inter-cloud management platforms.

6.4. End-to-end fault and performance management of network services in inter-cloud

[Editor's note in June 2019:] This sub-clause should illustrate the overview of end-to-end fault and performance management of NSs in inter-cloud. Contributions are invited.

One of the main challenges identified in NFV-based systems is related to fault management (including single or cascade faults) and performance issues, which have strong impact on whole environment. The precise detection of source of fault and area affected by faults are key aspects in telecommunication's software infrastructure, whose performance starts to be comparable to performance achieved over traditional networks.

6.4.1. Fault management of network services in inter-cloud

The fault management of NSs in inter-cloud would be a collaborative process among the elements constituting the service and the management systems involved. Modern communication systems produce large volumes of high-dimensional operational data. In such a case, analysing the data to get an actionable understanding of the situation becomes difficult. The fault management should be able to identify a fault that would require resources to restore the service parameters. In particular, the key challenges of fault management in virtualized environments are related to proper classification of reasons of fault in proper service operation:

- Fault detection to notify impending or actual fault and performance issues caused by resource failure.
- Determination of the root cause of the problem by identifying the inter-cloud resources that are malfunctioning or the severity with which they may malfunction in the future.
- Performance detection to notify impending or actual performance issues caused by service overload.
- Determination of the root cause of the problem by identifying the inter-cloud resources that are overloaded or the severity with which they may be overloaded in the future.
- Configuration detection to notify impending or actual performance issues caused by service configuration.
- Determination of the root cause of the problem by identifying the service configuration parameters that cause the severity or which they may cause it in the future.

6.4.2. Performance management of network services in inter-cloud

Performance management of NSs in inter-cloud is based on monitoring of certain Key Performance Indicators (KPIs), that are to be maintained at certain level of values e.g. above the threshold, below the threshold, between or outside the boundaries. KPIs are described by metrics, which are defined as measurable items. Dedicated capabilities which are implemented in functionalities allow to monitor the KPI's values in real time or in defined points in time. Changes of the values depend on the two groups of reasons:

- Performance constraints: service provider defines particular values for provided services to assure technical parameters e.g. throughput, delay, CPU load, capacity. Reaching the limit of the value results in degradation of performance parameters and triggers appropriate lifecycle operations like scaling.
- Failure constraints: service provider defines particular values for provided services to assure technical parameters e.g. throughput, delay, CPU load, capacity. In case of technical failure of elements of cloud environment, the performance parameters may be degraded and should be properly identified to trigger the appropriate lifecycle operation like healing.

The problem of detection and diagnostic for given conditions that degrade network performance deals with the detection of any condition that has already led to or could lead to degraded performance or failure as well as identification and localization of manifest and impending faults or elements to be scaled. The performance management is based on Quality of Service (QoS) metrics, which measure if the network behaves according to expectations, or Quality of Experience (QoE) metrics, which ensure the user perception of the network and service quality.

7. Functional requirements for end-to-end fault and performance management of network services

[Contributor's note] This clause will provides functional requirements related to end-to-end fault and performance management of network services based on ones derived from use cases.

[Editor's note in March 2020]: The text in this clause comes from the original clause 6.4.3. The style of the text should be modified as requirements and the related use case should be provided. Contributions are invited.

Faults happen due to physical or algorithmic causes. Faults may occur for a number of reasons, prominent amongst which are malfunctioning or failed devices because of hardware or software failures in VMs or VNFs, failure of links and configuration errors. There could be other reasons like cyber-attacks, disasters or environment factors. Faults appear as errors. Errors in turn are deviations of a system from normal operations. Errors are reported through system alarms. Alarms are notifications about specific events that may or may not be errors. The degradation of a service can be detected through notifications, counters or meters. The Fault detection and Performance Management system should be able to identify which issues are potential performance hazards or may result in a fault that would require resources to rectify.

Four levels of severity of events & alarms have been defined in ITU standard X.733: Critical, Major, Minor, and Warning [ITU92]. The critical alarm comes when the service can no longer be provided to the user. Major alarm indicates the service affected condition while minor means no current degradation is there, but if not corrected may develop into a major fault. A warning is an impending service affecting fault or performance issue. It is for the predictive capabilities of the Fault detection and Performance Management system to predict what faults will develop and with what severity levels.

Communication networks are widely distributed and are complex. The variety of FCAPS issues that can afflict them is large. The system to detect, diagnose and localize any condition that degrades network performance requires:

- Detection of any condition that has already led to or could lead to degraded performance or failure. The reasons could be manifest faults, hidden faults or inconspicuous deviations. The goal of such detection would be to sense and notify impending or actual fault and performance issues.
- Identification and localization of manifest and impending faults as well as performance problems. The goal of such localization would be to determine the root cause of the problem by identifying the resources that are malfunctioning or the severity with which they may malfunction in the future.

Any end-to-end fault and performance management system should take into account all the markers including alarms, notifications, warnings, observed behaviour, counter readings and measured values of performance indicators to carry out the above functions.

8. Framework of end-to-end fault and performance management of network services in inter-cloud

[Contributor's note] This clause will provide framework of end-to-end fault and performance management of network services in inter-cloud. At the moment, existing material is illustration only and allows better positioning aspects of network services in general network architecture. This material will be updated accordingly. Contributions are invited.

[Editor's note in March 2020]: The aspects of AI/ML should be considered as optionally here, as scope of work this draft Rec. is broad enough. Please consider to start new work item on ML for network management (including e.g. fault prediction) as alternative option. Contributions are invited.

[Editor's note in March 2020]: The number of subclauses (especially for clause 8) should be reduced to make the structure of this draft more simple and clear. Contributions are invited.

In NFV, faults and performance issues can have complex geneses within virtual resources, compute, storage and networking, as well as virtual network functions and cannot be effectively handled by traditional rule-based systems. To be able to make use of the Inter-Cloud paradigm effectively, it is more important to fix Fault and Performance issues. Without a robust mechanism for handling Fault and Performance, service providers would find meeting service level agreements (SLAs) difficult and growth of the promising technology of NFV might get hampered. The framework should contain mechanisms for handling both manifest and latent fault and performance issues.

The framework is intended to facilitate effective end-to-end fault and performance management in Inter-Cloud NSs. In telecommunication networks with physical appliances, deterministic methods ensure carrier grade availability and reliability. However, when telecom service providers' SCs are using virtual resources over multiple clouds, a number of complex factors make it imperative to use predictive methods for assuring carrier grade availability and reliability. This recommendation provide a model based on a judicious combination of shallow as well as deep structures / architectures in machine learning to ensure this objective.

8.1. Model for Fault Detection and Localisation

In general, predictive approach is recommended that takes a learning route to solve the problem of the complex interaction of features of fault detection and localization. More specifically, however, a model based on a judicious combination of shallow as well as deep structures / architectures in machine learning, can be used for prediction of fault & performance issues along-with the severity levels of impending faults with a high level of accuracy.

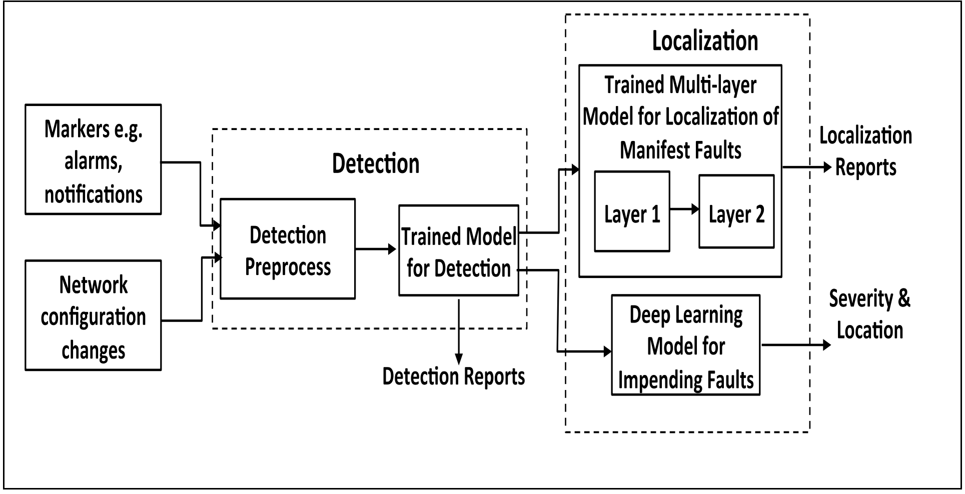

Markers e.g. alarms, notifications
Network configuration changes
Detection
Detection Preprocess
Trained Model for Detection
Detection Reports
Localization
Trained Multi-layer Model for Localization of Manifest Faults
Layer 1
Layer 2
Deep Learning Model for Impending Faults
Localization Reports
Severity & Location

Figure 8-1 – Fault and Performance Issues Detection and Localization Model

The proposed model approach has predictive and deductive properties to meet the fault and performance management requirements. Run time monitoring and measurements, alarms, notifications and warnings, configuration changes, measurements and environmental factors are all used along with the models trained with historical data to draw inferences about the manifest performance and fault issues. Additionally, decision about impending faults is taken using these inputs and the predictive properties of machine learning models. The detection system first decides whether there is a manifest or an impending fault or a performance issue. Based on this, the system will launch into identification and localization. Detection is essentially a two-stage binary classification problem that first classifies the outcome into 'normal performance' and 'abnormal performance' or 'faulty' and 'not faulty' classes. Then for the 'faulty' or 'abnormal' cases, it decides whether the problem is manifest or impending. Failure prediction needs to be accompanied with a high probability of correctness as actions following such a prediction involve cost. For localization, the model uses a multi-layered strategy. First, the broad category of the fault is determined (Layer 1). The system then does fine grain localisation (Layer 2) within the broad category and identifies the actual device(s) having a fault or suffering from performance degradation as well as their severity levels. Location & Severity of impending faults need deeper predictive structures.

8.2. Markers and Metrics for Fault Detection and Localisation

There are events relating to communication, quality of service, processing, equipment and environment that produce alarms, notifications, warning or error messages, measurements, counter values and conditions. Of course, many of the markers will appear in more than one type of fault or performance issue. Once, trained, detection and localisation algorithms would be able to pick out relevant markers and use them to predict the type of condition that may have arisen.

Some of the markers related to mobile, fixed and broadband networks are given in Table 8-1 below.

Table 8-1 – Illustrative list of markers

Mobile Network	Fixed Network	Broadband
Carrier/Interference Ratio	No Dial Tone	Intermittent Connection
Radio Link Time Out	Channel Noisy	Low Data Rate
Time Slot Shortage	MDF Jumper Disconnection	Phone Works Broadband Down
Occupied Bandwidth	Line Card Port Faulty	Repeated Training
RX Noise Floor	Primary Cable Fault	LAN Lamp Off
Radio Power	Distribution Cable Fault	Line Noisy
Frequency Error	DP Fault	DSLAM Port Mismatch
Antenna Tilt	House Wiring	No Ping
Signal Strength	MDF Fuse Blown	ADSL Lamp Flashes/Off
BTS Down	Customer Instrument Faulty	No Line Sync
Handover Failure	Dis in One Limb	Browsing Issues
Roaming Failure	Earth Contact	Micro-Filter Faulty
Packet Loss	Drop Wire Fault	No Comms
Hypervisor Alarm	Ring Tone Fault	Dropouts
Registration Failure	Message Fault	No Authentication

Low CSSR	Delayed Dial Tone	

Many of the markers could appear in more than one type of fault or performance issue. Some examples are shown in table 8-2.

Table 8-2 – Example showing many-to-many relationship between faults and markers

	Phase Error	Power	EVM	Rx Noise Floor	Origin Offset	Occupied BW	Frequency Error	C/I Ratio
Call Drop*	Y	Y	Y	Y	Y		Y	Y
Call Blocked**		Y	Y	Y	Y	Y		
*Radio link timeout; **Time Slot Short; EVM: Error Vector Magnitude; Rx: Receiver; C/I: Carrier to Interference Ratio; Y: Marker Present								

8.3. Training Datasets

The quality & quantity of the datasets affect the learning and prediction performance of machine learning algorithms. Information about faults, observations and restoration details in the telecommunication networks is contained in the fault dockets, test reports, central office system logs, outdoor maintenance staff logs, cable maintenance staff diaries and docket closure reports. Fault severity has three categories with 0 indicating no faults, 1 indicating a few faults and 2 indicating many faults. There are datasets for event type, the features logged, the resource affected and the severity type. The severity type is different from fault severity and classifies the warning given by the system.

8.4. Shallow and Deep Learning Methods

Shallow structures are simpler with one stage of non-linear operation, e.g., one hidden layer in neural networks. Here, the Support Vector Machine (SVM) Learning Method is a supervised learning method that analyses data and recognizes patterns. However, Deep learning architectures through stacked auto encoders would have more than one level of the composition of non-linear operations in the function learned. One of the key advantages of deep learning is the automatic extraction of high-level features from the given dataset. This is a distinct advantage over the difficult feature engineering in shallow structures that require human intervention. In deep learning, higher-level features are learned as a composite of lower level features. In this way, features are learned at many levels of abstraction, making it easier to grasp complex functions that map the input to the output directly from data.

In the above model for detection & localization of manifest & impending fault & performance issues of NSs in inter-cloud, some of the aspects of detection and localization of faults could be implemented using shallow and deep structures respectively. Simpler detection can effectively be handled by shallow machine learning structures like SVM. However, deeper structure i.e. the stacked auto encoder can be used for a more complex localization function where a large amount of information needs to be worked through to get to the root cause of the problem.

8.5. Detection of Fault and Performance

Fault and Performance issues may range from simple single point failures to multiple correlated or uncorrelated events. A fault presents itself in the form of system malfunction and notifications from faulty and other connected devices. The failure detection mechanism should be able to filter out dependent and routine operational events so that resources are not wasted in localizing these problems. In NFV, the faults in VM, VNF & Virtual Network cause NS to behave abnormally. For example, failure of a Gigabit Ethernet interface on the core router may cause some or all of the virtual private network (VPN) links of many customers to be non-functional. In this context, the goal of the Fault & Performance detection mechanism is to correlate alarms, notifications, measurements and other markers generated by events to infer manifest or predict impending performance and fault conditions. Some errors may be cleared by the system, others may produce warnings that may signal impeding problems while still another may produce faults that bring down functionalities and make themselves evident. The trained shallow machine learning models learn from the past events relating to faults and their resolutions. The models work in two stages: the first stage just makes a decision between ‘fault’ and ‘no-fault’ conditions, while the second stage does a more detailed examination of the markers to choose between ‘manifest’ and ‘impending’ faults. Minor faults & warnings would be the main contributors to the impending faults and need to be analysed to make this decision. With correct segregation, the localization stage would be able to carry out its functions properly.

8.6. Localization of Fault and Performance

The severity level of the faults indicates whether they are warnings, minor, major or critical. In the case of major & critical faults, devices degrade performance or stop working and need immediate action. Minor faults do not affect service and can be scheduled for localization accordingly. Warnings, along with the state information, provide insight into the degrading health of devices and could signal a major impending fault. In multi-layer fault identification and localization system, at Layer I, it detects the brand category of fault and then at Layer 2, does a fine grain classification. In the case of impending faults, the system predicts the locations and severity levels of the developing faults.

9. Security consideration

Security aspects for consideration within the cloud computing environment, including inter-cloud computing, are described in [ITU-T X.1601], which analyses security threats and challenges, and describes security capabilities that could mitigate these threats and meet the security challenges.

Appendix I

Use case of end-to-end fault and performance management of network services in inter-cloud

(This appendix does not form an integral part of this Recommendation.)

I.1 Use case template

The use cases developed in Appendix I should adopt the following unified format for better readability and convenient material organization.

Title	Note: The title of the use case
Description	Note: Scenario description of the use case
Roles	Note: Roles involved in the use case
Figure (optional)	Note: Figure to explain the use case, but not mandatory
Pre-conditions (optional)	Note: The necessary pre-conditions that should be achieved before starting the use case.
Post-conditions (optional)	Note: The post-condition that will be carried out after the termination of current use case.
Derived requirements	Note: Requirements derived from the use cases, whose detailed description is presented in the dedicated chapter

[Editor’s note in March 2020]: The use cases should be reconsidered to provide the derived requirements for clause 7. Contributions are invited.

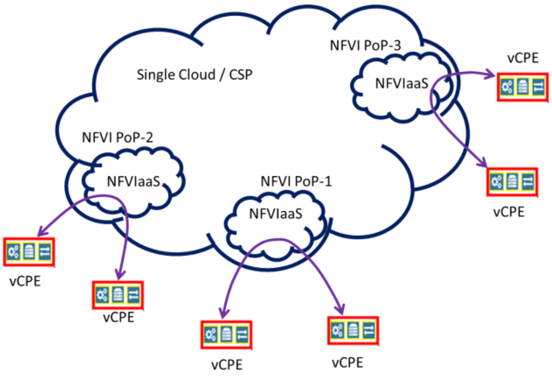
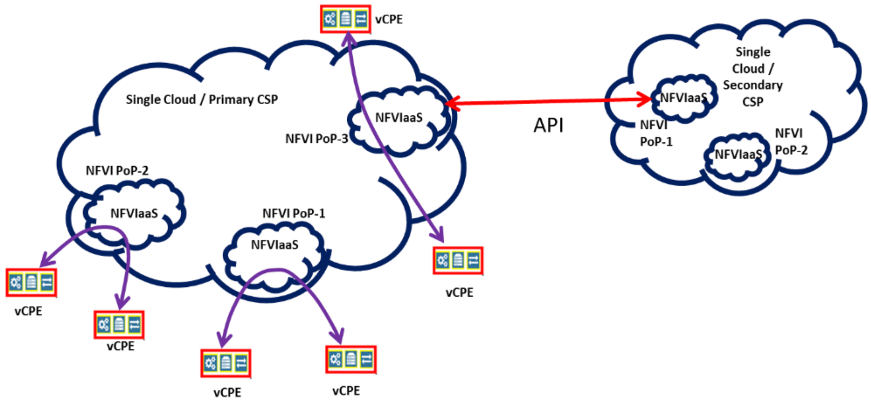
I.2. NS in inter-cloud scenario

I.2.1 NS within a NFVI-PoP in a single CSP cloud

This use case illustrates the NS built between two virtual Customer Premises equipment (vCPEs), which are placed at different CSC locations, and are connected to the same NFVI-PoP of a CSP. A single CSP cloud is the cloud infrastructure owned by one CSP.

Table I.2.1 NS within the same NFVI-PoP in a single CSP cloud

Title	NS within a NFVI-PoP in a single CSP cloud
Description	A CSP is fully responsible for the creation, scaling, termination (life cycle management) of a NS. A primary CSP may avail NFVIaaS from secondary CSP. A CSP shall be responsible for the Fault and performance management of the NS which includes the CSC links.
Relevant roles	CSC and CSP

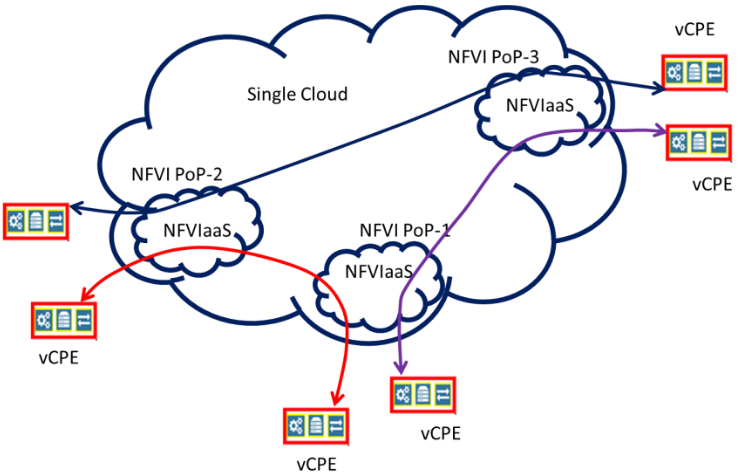
High-level figure describing the use case	 Single Cloud / CSP NFVI PoP-3 NFVIaaS vCPE vCPE NFVI PoP-2 NFVIaaS NFVI PoP-1 NFVIaaS vCPE vCPE vCPE vCPE Fig – A (above) : Single Cloud  vCPE Single Cloud / Primary CSP NFVIaaS NFVI PoP-3 API Single Cloud / Secondary CSP NFVIaaS NFVI PoP-1 NFVIaaS NFVI PoP-2 NFVI PoP-2 NFVIaaS NFVI PoP-1 NFVIaaS vCPE vCPE vCPE vCPE vCPE vCPE Fig – B (above): Single cloud availing NFVIaaS from secondary CSP
Pre-conditions	
Post-conditions	
Derived requirements for the cloud capability	

I.2.2. NS among different NFVI-PoPs in a single CSP cloud

This use case illustrates NS built between two vCPEs connected to different NFVI-PoPs within a single CSP cloud.

Table I.2.2 NS among different NFVI-PoPs in a single CSP cloud

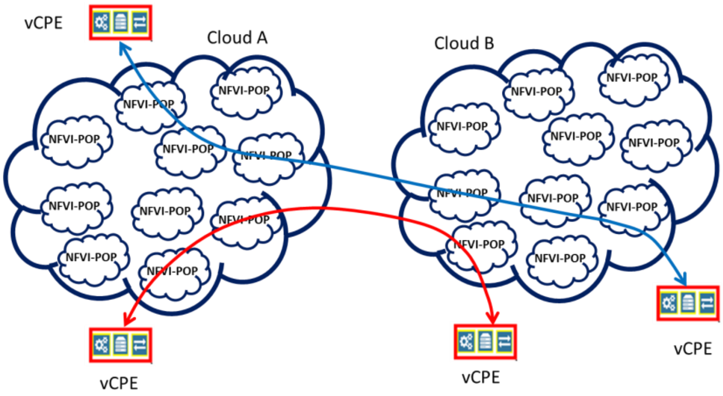
Title	NS among different NFVI-PoPs in a single CSP cloud
Description	A CSP is fully responsible for the creation, scaling, termination (life cycle management) of a NS. A primary CSP may avail NFVIaaS from secondary CSP. A CSP (primary) shall be responsible for the Fault and performance management of the NS which includes the CSC links.
Relevant roles	CSC and CSP

High-level figure describing the use case	 Single Cloud NFVI PoP-3 NFVIaaS NFVI PoP-2 NFVIaaS NFVI PoP-1 NFVIaaS vCPE vCPE vCPE vCPE vCPE
Pre-Conditions	
Post-Conditions	
Derived requirements for the cloud capability	

I.2.3. NS among NFVI-PoPs in separate clouds

This use case illustrates the NS built between two vCPEs connected to NFVI-PoPs located in two separate clouds administered by two different CSPs.

Table I.2.3 NS among NFVI-PoPs in separate clouds

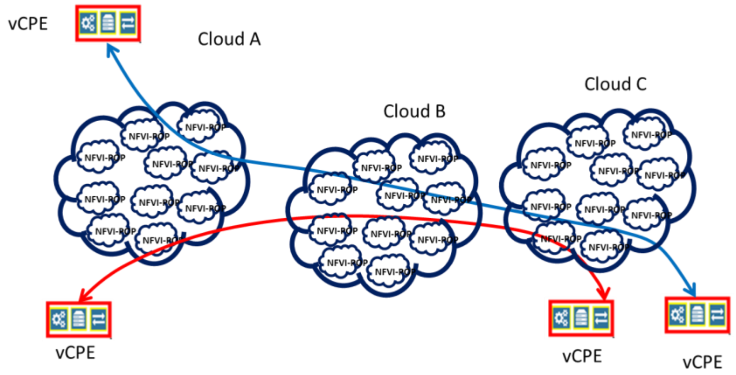
Use case title	NS among NFVI-PoPs in separate clouds
Use case description	Both the CSPs responsible for the creation, scaling, termination (life cycle management) of a NS. Both CSPs shall be responsible for the Fault and performance management of the NS in their administrative domain area including the CSC link connected to their cloud.
Relevant roles	CSC and CSP
High-level figure describing the use case	 vCPE Cloud A Cloud B NFVI-POP vCPE vCPE vCPE

Pre-Conditions	
Post-Conditions	
Derived requirements for the cloud capability	

I.2.4. NS among NFVI-PoPs in separate clouds with an intermediary cloud

This use case illustrates the NS built between two vCPEs connected to NFVI-PoPs located in two separate clouds with another cloud as an intermediary.

Table I.2.4 NS among NFVI-PoPs in separate clouds with an intermediary cloud

Title	NS among NFVI-PoPs in separate clouds with an intermediary cloud
Use case description	All the CSPs responsible for the creation, scaling, termination (life cycle management) of a NS. All the CSPs shall be responsible for the Fault and performance management of the NS in their administrative domain area. All the intermediary CSPs are not responsible for the CSC links.
Relevant roles	CSC and CSP
High-level figure describing the use case	 vCPE Cloud A Cloud B Cloud C NFVI-PoP vCPE vCPE vCPE
Pre-conditions	
Post-conditions	
Derived requirements for the cloud capability	

[Editor's note in June 2019:] The content of Appendix II comes from C34 in June 2019's SG13 meeting. Although we agree to remove NFVIaaS and VNFaaS related content from clause 6, some expert think the relationship of NFVIaaS and VNFaaS with NaaS capabilities is critical for this draft Rec. So we add this Appendix to remind this issue and decide whether it should be reserved in the future.

Appendix II

Clarification on NFVIaaS, VNFaaS and the relationships with NaaS capabilities

II.1 Concept of NFVIaaS and VNFaaS

Both NFVIaaS and VNFaaS are use cases describing service models (or cloud service categories) which are defined by ETSI. These use cases are intended to clarify the roles and interactions of the various types of commercial entities acting in a marketplace for services delivered by CSPs. But these two cloud service categories are not defined in ITU-T yet.

In NFVIaaS, the NFV Infrastructure(NFVI) can be considered as a service providing the capability or functionality to support an environment in which VNFs can execute, and a CSP could run VNF instances inside an NFVI which is operated as a service by a different CSP. From the cloud capabilities type of view, NFVIaaS provides the infrastructure capabilities type.

In VNFaaS, the VNF can be considered as a service provided by CSP to CSC. From the cloud capabilities type of view, VNFaaS provides the application capabilities type.

II.2 Concept of NaaS

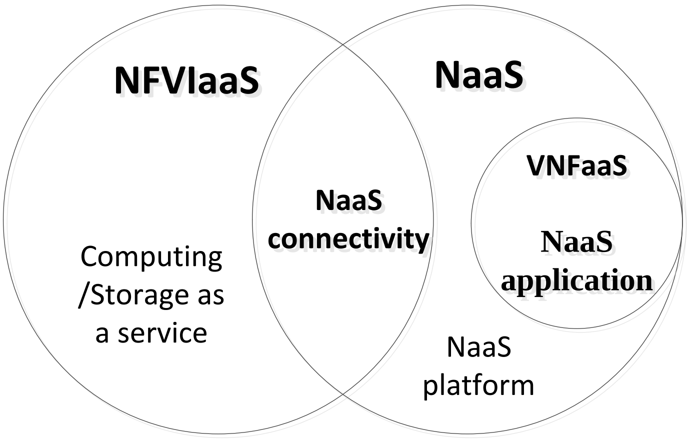
NaaS is a cloud service category in which the capability provided to the CSC is transport connectivity and related network capabilities. NaaS can provide any of the three cloud capabilities types(i.e. infrastructure capabilities type, platform capabilities type, application capabilities type). NaaS services are therefore divided into NaaS application service, NaaS platform service and NaaS connectivity service.

NaaS application: application capabilities type of service where NaaS CSC can use network applications provided by NaaS CSP. These network applications are considered and used as a VNF provided by NaaS CSP. Examples of NaaS applications include virtual router, virtual content delivery network (vCDN), virtualised evolved packet core (vEPC) and virtual firewall (vFW).

NaaS connectivity: infrastructure capabilities type of service where NaaS CSC can provision and use networking connectivity resources provided by NaaS CSP. This includes for example flexible and extended VPN, bandwidth on demand (BoD), etc.

II.3 The relationships of NFVIaaS and VNFaaS with NaaS capabilities

Base on the above's discussion, we can see there are some overlaps existing in NFVIaaS , VNFaaS and NaaS. The relationships of NFVIaaS and VNFaaS with NaaS are illustrated in below figure:



[Editor’s note in June 2019:]Some expert suggests that NaaS application is one of the VNFaaS similarly as NaaS Connectivity is one of the NFVIaaS. Discussion about this issue will based on contribution. Contributions are invited.

NFVIaaS, VNFaaS and NaaS are among various types of cloud service categories. They can provide different types of cloud services to CSC. The overlap between NFVIaaS service and NaaS service is NaaS connectivity which is an infrastructure capabilities type of service where CSC can use networking connectivity resources provided by CSP. And VNFaaS service is almost as same as NaaS application service which CSC can use network applications/VNFs provided by CSP.

II.4 Conclusion

Base on the above’s discussion, we can summarize the relationships of NFVIaaS and VNFaaS with NaaS as below:

- NFVIaaS, VNFaaS and NaaS are different kind of cloud service category;
- The capabilities of NFVIaaS and NaaS have an overlap as both of them can provide transpor t connectivity and related network capabilities;
- The capability of VNFaaS is almost as same as NaaS application capability which can provide network applications/VNFs.

Bibliography

[b-DMTF OVF] DMTF Standard DSP0243 Version 1.0.0 (2009), Open virtualization format specification.

[b-ETSI ~~GS~~ GR NFV 003] ETSI ~~GS~~ GR NFV 003 V1.~~4~~5.1 (~~2018~~2020-01), *Network functions virtualisation (NFV); Terminology for main concepts in NFV*.