INTERNATIONAL TELECOMMUNICATION UNION

TELECOMMUNICATION STANDARDIZATION SECTOR

STUDY PERIOD 2017-2020

16286-C86 (180409)

Study Group 13

Original: English

Question(s):	Q19/13	Switzerland [Geneva], 2018-04-9/18
C		
Source:	India	
Title:	Proposal of new work item on "Predictive Fault and performance management to ensure carrier grade reliability and availability of virtual network services on multiple clouds"	
Purpose:	Proposal	

Contact:	Lav Gupta India	Tel: +91 11 23320317 Fax: +911123323157 E-mail: ddgm.tec@gov.in
	A.S. Verma TEC, DoT India	

Keywords: Multiple cloud, predictive fault, performance management

Abstract: This contribution proposed new work item on "Predictive Fault and performance management to ensure carrier grade reliability and availability of virtual network services on multiple clouds"

Background:

Cloud Computing is being studied in the Telecommunication Engineering Center (TEC) of the Department of Telecommunications (DoT) under Ministry of Communications, Govt. of India, collaboratively in the divisions of Future Networks, Information Technology, Mobile Technologies and Telecommunications Security. Among themselves these divisions are engaged in the study of cloud computing as an essential paradigm for 5G, placement of virtual network functions on clouds and security in virtual network services deployment on multiple clouds. The Future Networks division is presently coordinating the activities. National working group (NWG) - 13 was formed in TEC to coordinate ITU-T SG-13 related activities in India. It is having members from Industries and the Government.

Proposal

This contribution relates to deployment of telecommunication network services over multiple clouds. It proposes to discuss the paradigm of end-to-end fault and performance detection and localization in cloud based virtual network services as a means to achieve carrier grade reliability and availability. The proposal is generic in nature and is applicable to all member states.

The details of new work item are as per the Annexure A attached.

Annexure A

Use case:

1. Title of the use case

a) Name of the use case: Predictive Fault and performance management to ensure carrier grade reliability and availability of virtual network services on multiple clouds.
b) ID of the use case:
c) Version/revision history: 22/October/2017

2. Source: India/ MoC/ TEC Objective of the use case

This use case describes the problem of fault and performance management in telecommunication networks deployed over clouds, specifically multiple clouds. Such a deployment introduces complexities beyond IT services deployment over clouds where problems can be traced to virtual machines or the

physical infrastructure underlying virtualization. In virtual network services the virtual network functions may themselves randomly malfunction in addition to the already existing complexities in multi-cloud deployments. Identification of the fault and performance problem and elements of the solution would help in reaping myriad benefits that such deployments could potentially bring.

3. Background

a) Countries' specific telecommunication network deployment scenario

The India specific information is available here http://www.dot.gov.in/telecom-glance
The use-case described here is generic in nature and applicable to all member states interested in ushering state-of-the-art cloud based telecommunications network deployment in their countries. Such deployment are expected to give a number of benefits over traditional deployments using physical applicances. Some of the benefits are: flexibility of obtaining resources, ease of scaling and descaling, freedom from proprietary hardware and software, ease of redeploying resources, risk mitigation, ease of deploying new services and reduced total cost of operation.

b) Current Practice

The current practice largely involves use of physical network applicances like routers, switches, broadband remote access servers or middleboxes like firewalls, deep packet inspectors or load balances. These appliances and associated software are normally proprietary and leads to vendor lock-in making expansions and deployment of new services difficult and time consuming. They are also not amenable to easy scaling or redeployment of resources. The power and space requirements as well as the total cost of operation is higher in physical networks.
In traditional networks described above, 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 availability and carrier grade reliability.

c) Need for Use Case

Network function virtualization (NFV) is a promising framework. NFV coupled with multi-cloud computing provides numerous advantages to the service providers including ease of deployment, ease of scaling, ease of introducing and switching off services and reduced cost of operation. This would increase viability of telecommunications business and lead to a thriving telecommunication sectors in the member states. NFV over multiple clouds has not yet attained the level of performance to be a viable replacement for traditional networks. One of the main reasons is the absence of a standard based Fault, Configuration, Accounting, Performance and Security (FCAPS) framework for the virtual network services. Since deployments of such networks are in nascent stages in various countries, there is a need for standardization of techniques for fault and performance detection and localization in such networks.

d) Ecosystem Specifics

Some of the key challenges in NFV over multi-clouds are as follows:

• Absence of an FCAPS framework
• Non-applicability of traditional rule based techniques used in today's networks
• Multiple layers of implementation: physical infrastructure, NFVI (Virtual Machines), Virtual Network Functions (VNF) and Virtual Network Services.
• Massive distribution of network functions over disparate clouds
• Multiple control centers: cloud management systems, operators' OSS/BSS and NFV-MANO (Management and Orchestration)

4. Description

a) Summary

Telecommunications networks have traditionally been designed to provide five nines availability and standards based quality of service. In virtual network services deployment over multiple clouds it is challenging to equip multi-cloud management systems to deal with fault and performance issues. Virtual network services have underlying physical and network function virtualization infrastructure. The telecommunications network functions in the virtualized form go into the virtual machines provided by the NFVI. When performance deviates from normal or a fault occurs, there is no access to the physical hardware for telecom operators to test. The root cause of the problem could be in the physical, virtual layer or the virtual network functions. The traditional deterministic methods fail to deliver in virtual environments in which virtual resources can be constantly scaled, migrated or destroyed. It is proposed to make use of predictive techniques to be able to identify fault or performance issues before or after they have occurred.

b) Scope

This Recommendation is intended to facilitate effective end-to-end fault and performance management in multi-cloud virtual network services. In telecommunication networks with physical appliances, deterministic methods ensure carrier grade availability and reliability. However, when telecom service providers service function chains using virtual resources over multiple clouds a number of complex factors make it imperative to use predictive methods for assuring carrier grade availability. This recommendation discusses open source non-discriminatory procedure to ensure this objective.

c) Ecosystem Description

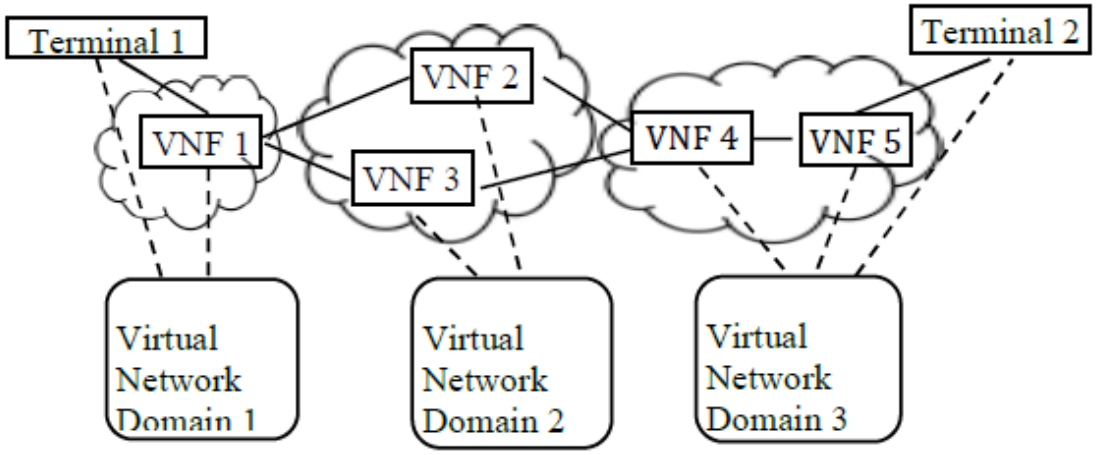
Network Function Virtualization (NFV) provides the advantages of breaking free from proprietary network appliances and brings in ease of scaling. When NFV is deployed over multiple clouds then there are added advantages like greater flexibility in obtaining resources, avoiding total outages, proximity to customers and lower total cost of operation. Cloud technology can multiply the benefits of NFV [2], [3], [59]. It could provide greater flexibility in obtaining resources, bring Network Service Provider's (NSP's) points of presence close to customers, provide an opportunity to optimize performance and control cost. However, NFV over multiple clouds has not yet attained the level of performance to be a viable replacement for traditional networks. One of the main reasons is the absence of a standard based Fault, Configuration, Accounting, Performance and Security (FCAPS) framework for the virtual network services. Traditional networks have time-tested standards relating to fault, configuration, accounting, performance and security (FCAPS) as embodied in ISO Common Management Information Protocol (CMIP) and ITU TMN M.3010 and M.3400 recommendations. In NFV Concerns regarding five nines availability and quality of service parameters, like latency and packet loss, still remain.
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 multi-cloud paradigm effectively, it is important that we fix the FCP issues for this environment. Without a robust mechanism for handling FCP, 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 manifest and latent fault and performance issues.

i) Network Service Structure

Based on the ETSI specifications and IETF RFC a network service can be described as a service function chain (SFC) or virtual network function (VNF) graph, interconnected by virtual network resources, as an ordered set of VNFs that represent functions like routers and broadband network gateways or middleboxes like load balancers and firewalls, which act on the traffic in the sequence they appear in the chain. VNFs are

hosted on virtual machines (VM) instantiated over physical data center and network resources. An example of end-to-end service is shown on the left side of Fig 1.

Fig. 1 Multi-domain End-to-End Service



Faults happen due to physical or algorithmic causes. They 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 FCP 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 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 affective 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 FCP 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. To detect, diagnose and localize any condition that degrades network performance becomes quite onerous. In this contribution we restrict our discussion to:

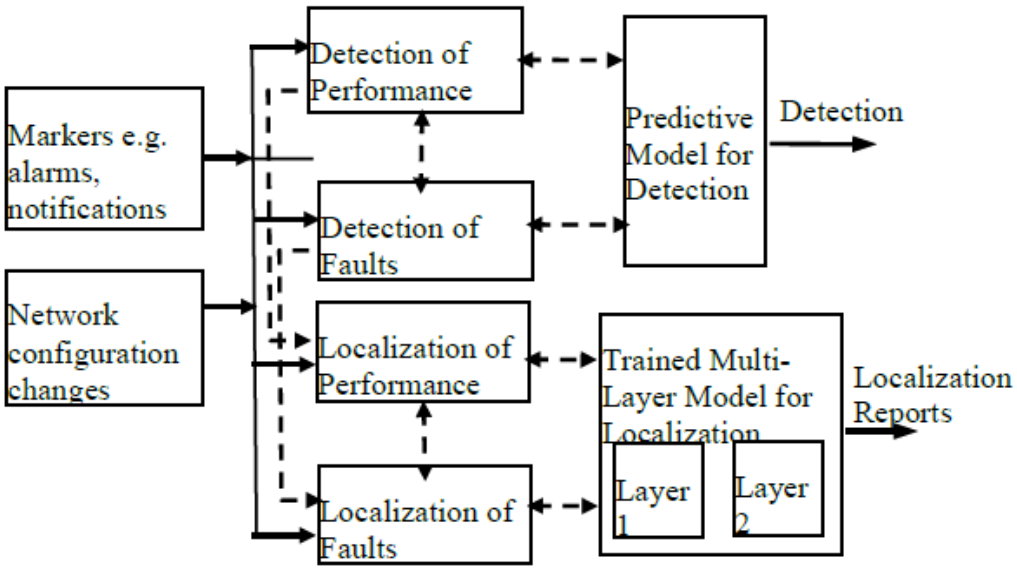
1. Detection of any condition that has already led to or could lead to degraded performance or failure. The reasons could be manifest faults, hidden faults or inconspicuous deviations. The goal of FCP detection would be to sense and notify impending or actual fault and performance issues.

2. Identification and localization of manifest and impending faults. The goal of FCP 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 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.

ii) FCP Solution Components
FCP management in cloud based network services 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, analyzing the data to get an actionable understanding of the situation becomes difficult. In general, the researchers agree on predictive approaches that take a learning route to solve the problem of the complex interaction of features of fault detection and localization.

Fig 2. Fault/Performance Detection and Localization Model



The proposed model has predictive and deductive properties to meet the FCP 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. The system then identifies the actual device(s) having a fault or suffering from performance degradation and their severity levels. Severity of impending faults need deeper predictive structures.

5. **Proposal:**
This contribution is being submitted for the information and benefit of member states. It may be taken in to account in the standardization work plan of future networks involving end-to-end management of cloud computing when applied to telecommunications networks.

_____