

Fault Localization in NFV Framework

Byung Yun Lee*, Bhum Cheol Lee*

* ETRI(Electronics and Telecommunications Research Institute), Korea

bylee@etri.re.kr, bcleee@etri.re.kr

Abstract— Network function virtualization is quickly gaining acceptance as the new approach to delivering communication services. The promises of greater flexibility and dramatically reduced time to introduce new services coupled with the cost advantages, are driving communications service providers(CSP) around the world to begin deploying NFV-based services. But NFV service is actually operated based on virtualization environment, it has greater impact on disability virtual resources, when the physical devices had a fault. In this paper, when physical faults occurred in the NFV Framework, or logical failure occurs on a specific logical device, we will present a structure for Fault Localization method to ensure the continuity of service in NFVI

Keywords— NFV(Network Function Virtualization), Fault Localization

I. INTRODUCTION

Cloud-based NFV platform is a dynamic pool of virtualized computing resources and offers an elastic scheme matching the user demand, so that allocated resources can be scaled up or down on a per-use basis. To make the most of its advantages, distributed cloud-based NFV platform offers its high quality & proximity services to the end user from the closest cloud platform so called NFV HW platform. But NFV service is actually operated based on virtualization environment, it has greater impact on disability virtual resources, when the physical devices had a fault.

II. NFV MANAGEMENT FRAMEWORK

ETSI ISG has created the industry standards required in a short time to target specific technology areas that industry requires, for NFV management framework [1].

NFV management framework consists of three functional entities including NFV Orchestrator, VNF Manager and Virtualized Infrastructure Manager.

The NFVO(NFV Orchestrator) is responsible for the lifecycle management of Network Services across the entire Operator's domain (e.g. multiple VIMs: Virtualized Infrastructure Manager) and the orchestration of NFVI(Network Function Virtualization Infrastructure) resources across multiple VIMs fulfilling the Resource Orchestration functions[2].

The VNF Manager is responsible for the lifecycle management of VNF instances. Each VNF instance is assumed to have an associated VNF Manager. A VNF manager may be assigned the management of a single VNF

instance, or the management of multiple VNF instances of the same type or of different types.

The Virtualized Infrastructure Manager is responsible for controlling and managing the NFVI compute, storage and network resources, usually within one operator's infrastructure domain. A VIM may be specialized in handling a certain type of infrastructure resource (e.g., compute-only, storage-only, networking-only), or may be capable of managing multiple types of infrastructure resources.

NFV framework provides a virtual container interface in order to provide independent execution environment, and required resources (network, compute, storage, accelerator) rely on NFVI[3]. Each VNF is given all the necessary resources through a virtual container interface.

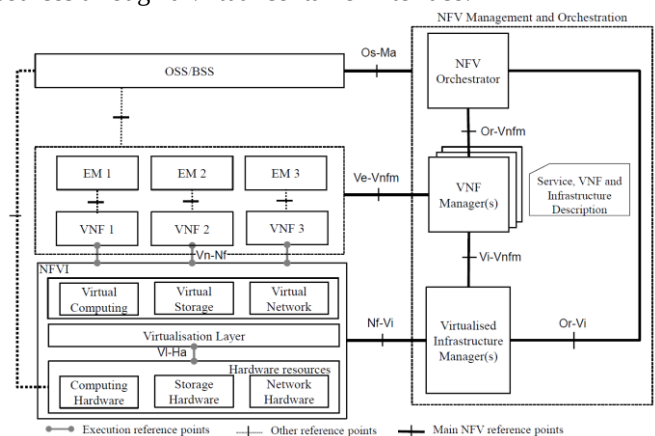


Figure.1 NFV MANAGEMENT FRAMEWORK (ETSI ISG)

III. FAULT LOCALIZATION IN NFV FRAMEWORK

In a NFVI environment, it is necessary to provide fault localization framework for NFVI and VIM layers focusing on network, and fault localization target is to deduce the exact source of a failure from a set of observed failure indications while reducing the time spent on analysing problems.

For example, as in fig.2, VNF #2 indicates that it is not working (no sessions, no network connectivity etc.), there are several causes may result this: iptables may not configured or MTU size misconfigured in vSwitch, or NIC or ToR switch have a failure. In case of physical switch or NIC are down, there are several symptoms as follow (management port down, neighbor switches port down, neighbor hosts port down, VMs connectivity lost, Apps connectivity lost). To localize the fault, it is diagnosed by the API, where and what was the problem to make happen.

- Get physical topology – Find out all existing switches in the domain, connectivity, connection to racks, connection to hosts
- Get virtual topology
- Get mappings: VMs to hosts, Hosts to racks, racks to switch ports, apps to VMs
- Get switch status/ event
- Get switch port status /event
- Get NIC status /event
- Get VMs status /Event
- Activate link OAM tool

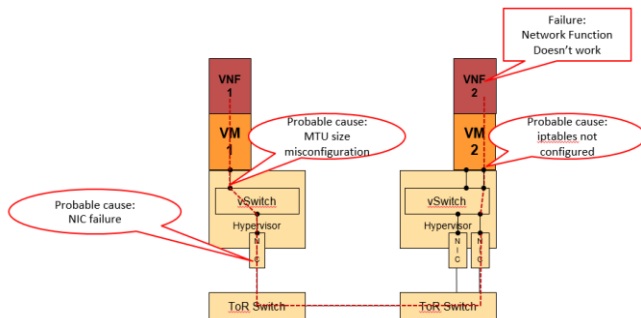


Figure.2 Fault in NFVI

As shown in Figure. 3, to localize the fault, it is necessary to monitor various events, alarms, statistics from the exact sources, and analysing the related entities based on the configuration of the NFVI.

After analysing the fault, Fault Localization Framework finds root causes, and correlated failures caused by the fault.

There are several processes to get the information of the exact resources.

1) System OAM tools

- Ping: This command shows how long it takes for packets to reach host, it is possible to check the resources are working and reachable.
- SNMP: This protocol is widely used in network management systems to monitor network-attached devices for conditions that warrant administrative attention.
- Trace Route : This command traces the route of packets to destination host from specific server, it is possible to check the route from specific server to the resources are working and reachable

2) Fault information Source

NFV platform consists of interrelated components that control hardware pools of processing, storage, and networking resources throughout a data center. The fault information source are modules like Neutron, Nova, Cinder.

- Events : Event notification occurred from the source when it has some information to notify
- Alarms : Alarm occurred from the source when it has some information to alarm, as critical, major, minor.
- Statistics : It collects the information about the statistics of the specific source.

- Logfiles : It records either events that occur in an operating system or other software runs, or messages between different users of a communication software. In the simplest case, messages are written to a single logfile.
 - Local fault correlators : It represents correlation information about a fault. (Ceilometer, Monasca, Vitrage)
- ① Ceilometer provides single point of contact for billing systems, providing all the counters they need to establish customer billing, across all current and future OpenStack components. The delivery of counters is traceable and auditable, the counters must be easily extensible to support new projects, and agents doing data collections should be independent of the overall system.
 - ② Monasca is a open-source multi-tenant, highly scalable, performant, fault-tolerant monitoring-as-a-service solution that integrates with OpenStack. It uses a REST API for high-speed metrics processing and querying and has a streaming alarm engine and notification engine.
 - ③ Vitrage is the Openstack RCA (Root Cause Analysis) Engine for organizing, analyzing and expanding OpenStack alarms & events, yielding insights regarding the root cause of problems and deducing the existence of problems before they are directly detected.
- IETF LIME – Layer Independent OAM Management in the Multi-Layer Environment.

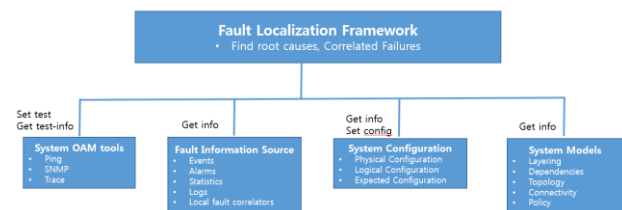


Figure.3 Architecture of Fault Localization Framework

3) System Configuration

- Physical Configuration : Information about Physical Configuration
- Logical Configuration : Information about Logical Configuration

4) System Models

- Network topology

① Legacy world :

- MTOSI(Multi-Technology Operations Systems Interface) is the NBI of the NMS that will provide the topology up to the TOR switch. MTOSI is an XML-based Operations System (OS)-to-OS interface suite. The Network Management System-to-Element Management System communications is a special

case and is defined by the Multi-Technology Network Management (MTNM) standards.

- ToR switch provides Learned Mac addresses via SNMP
- IPMI(Intelligent Platform Management Interface) manager can provide all the required information about the hosts. IPMI is a set of computer interface specifications for an autonomous computer subsystem that provides management and monitoring capabilities independently of the host system's CPU, firmware (BIOS or UEFI) and operating system.

② SDN World

- SDN controller has a global view of the virtual network and NBI should provide topology for all SDN switches.
- SDN controller had learnt MAC addresses of all the VM in the management network domain
- IPMI manager can provide all the required information about the hosts

IV. FAILURE ON THE NFVI

Some VNFCs may have a requirement for being deployed on a compute node with two physical interface cards which mutually protect each other to prevent a VNFC failure in case of a physical NIC, physical link, or physical switch failure. Therefore, this clause illustrates remediation and recovery alternative for failures of a physical NIC, a physical link between physical NIC and adjacent switch, or the failure of an adjacent switch. The failure of an adjacent switch or the link to this switch is assumed to be detected by the NIC leading to an alarm generated by the NIC driver. Failures of the physical NIC will also generate an alarm by the NIC driver. For this illustration, it is further assumed that the compute node has dedicated NICs connecting the node with the VIM and at least two NICs which can be assigned to the tenant's virtual environment[4]. The failure is assumed to happen on one of the NICs used by the tenant.

A. Physical NIC bonding

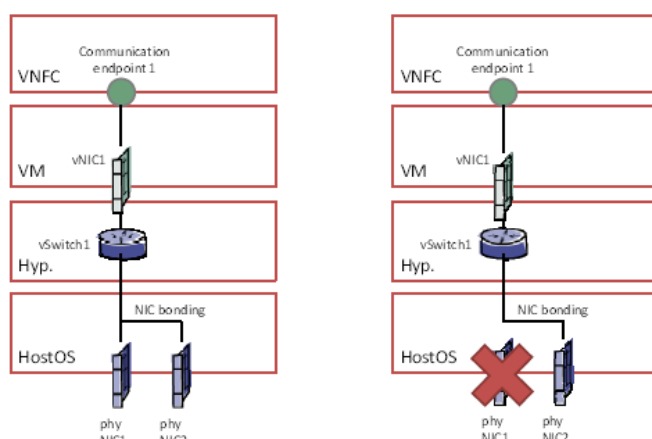


Figure.4 System configuration for physical NIC bonding

When physical NIC1 or its upstream resources fail, the physical NIC is removed from the bonding. Traffic is automatically transferred to the backup NIC to continue service delivery. A corresponding alarm has to be generated and forwarded via VIM to VNF Manager to decide which further steps to take - the alarm content changes depending on the abstraction level. Details on the alarm content are provided below.

In an optional step the VNF manager can initiate to trigger the VNFC instance to fail-over to a standby VNFC instance with full NIC redundancy in case of VNFC level protection. Afterwards, recovery to normal operation is performed; for this the VNFC is re-located to a compute node with two or more functioning physical NICs. The flow diagram below shows VM migration as one option

Within the node, a host_alarm message informs the HostOS about failed hardware, e.g. link-down event. This message is OS specific.

The node sends a node_alarm message to VIM informing about the failed resource, i.e. link-down on NIC-ID or NIC-ID failed. The VIM identifies which VMs have been scheduled on the impacted node, have requested physical NIC redundancy, and have been allocated the failed NIC. The VIM informs the respective VNF Managers using a vnfm_alarm message; this message contains an indicator that the failure results in a service degradation provided by the hardware to the VM.

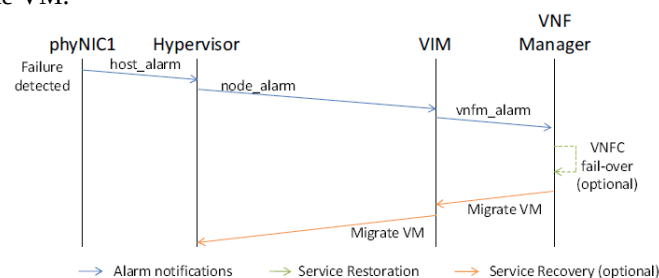


Figure.5 Recovery from NIC failure in case of physical NIC bonding

B. NIC bonding of virtual NICs

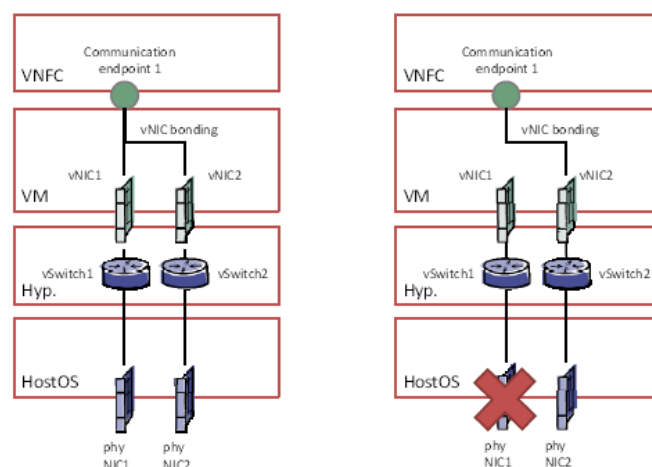


Figure.6 System configuration for virtual NIC bonding

When physical NIC1 or its upstream resources fail, the vSwitch attached to this physical NIC should be deactivated.

This would result in a NIC-down event for all attached vNICs. Though this, the vNIC is removed from the bonding. Traffic is automatically transferred to the backup vNIC. A corresponding alarm has to be generated and forwarded to VNF manager to decide which further steps to take.

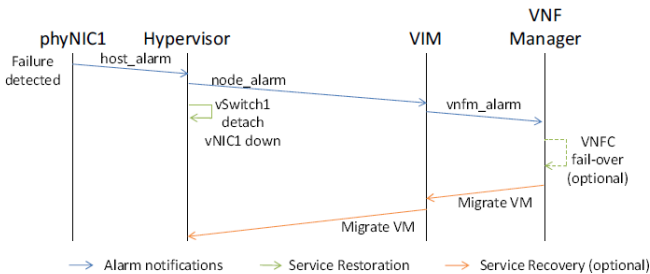


Figure.7 Recovery from NIC failure in case of virtual NIC bonding

The procedure also applies to setups with direct access from the VM to the physical NIC, e.g. SR-IOV or IOMMU. The only difference is the missing vSwitch and thus, the related deactivation of it[5].

C. VNF internal failover mechanism

When physical NIC1 or its upstream resources fail, the NFVI informs the VNF about the failure. This could be implemented as a link down command issued by the host OS to the virtual NIC. Once the virtual NIC is down the VNF will detect this event and will transfer communication channels from the failing NIC to the other NIC by performing a socket re-attach procedure, this can be implemented for example by opening a new socket and informing peer entities.

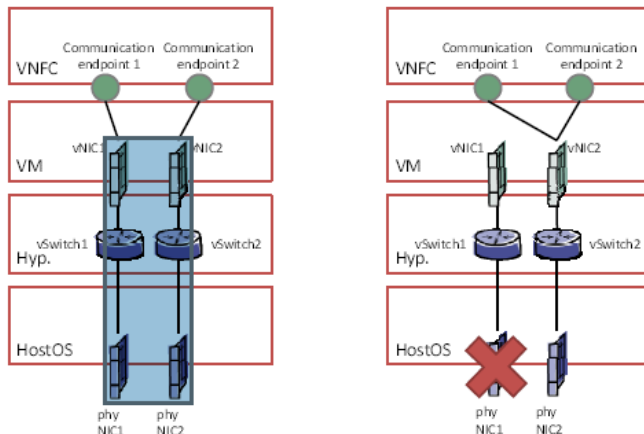


Figure.8 System configuration for virtual NIC bonding

A corresponding alarm has to be generated and forwarded to VNF manager to decide which further steps to take.

In order to realize this use case a new alarm message is required. The Hypervisor informs the VNF about the affected virtual resource mapped to the failed virtual resource by sending a vnf_alarm message.

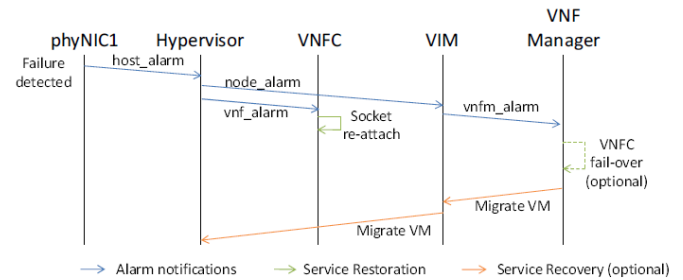


Figure.9 VNF internal Restoration and Recovery

V. CONCLUSION

Network function virtualization is quickly gaining acceptance as the new approach to delivering communication services. The promises of greater flexibility and dramatically reduced time to introduce new services coupled with the cost advantages, are driving communications service providers(CSP) around the world to begin deploying NFV-based services. But NFV service is actually operated based on virtualization environment, it has greater impact on disability virtual resources, when the physical devices had a fault. In this paper, when physical faults occurred in the NFV Framework, or logical failure occurs on a specific logical device, we will present a structure for Fault Localization method to ensure the continuity of service in NFVI.

ACKNOWLEDGMENT

This work was supported by Institute for Information & communications Technology Promotion(IITP) grant funded by the Korea government(MSIP) (B0101-15-233, Smart Networking Core Technology Development)

REFERENCES

- [1] ETSI GS NFV-INF 001, NFV Infrastructure Overview, 2015-01
- [2] ETSI GS NFV-MAN 001, NFV Management & Orchestration
- [3] ETSI GS NFV 002, NFV Architectural Framework, 2014-12
- [4] ETSI GS NFV-REL 001, NFV Resiliency Requirements
- [5] <https://wiki.opnfv.org/projects/pinpoint>



Byung Yun Lee is currently a Principal Member of Telecommunication Internet Research Division at Electronics and Telecommunication Research Institute (ETRI), Korea. He received the PhD degree in computer engineering from Chungnam National University, Korea, in 2003. Since joining ETRI in 1992, his work has focused on SDN/NFV technology, and network management.



Bhum Cheol Lee received M.S. and Ph.D. degree in Electric Engineering from Yonsei University, Korea in 1983 and 1997, respectively. He is currently Manager of Networking Computing Convergence Lab. in Electronics and Telecommunications Research Institute (ETRI), Korea. His research interests are Smart Network, Parallel Flow Processing and Network Virtualization