SLA-controlled Proxy Service Through Customisable MANO Supporting Operator Policies

Thomas Soenen*, Felipe Vicens†, José Bonnet‡, Carlos Parada‡, Evgenia Kapassa§, Marios Touloupou§, Eleni Fotopoulou¶, Anastasios Zafeiropoulos¶, Ana Pol‖, Stavros Kolometsos**, George Xilouris**, Pol Alemany††, Ricard Vilalta††, Panos Trakadas‡‡, Panos Karkazis‡‡, Manuel Peuster[x] and Wouter Tavernier*

*Ghent University - imec, †Atos Research & Innovation, ‡Altice Labs, §University of Piraeus, ¶Ubitech, ‖Quobis, **NSCR Demokritos, ††Centre Tecnolgic de Telecomunicacions de Catalunya, ‡‡Synelixis, [x]Paderborn University

***Abstract*—Service Level Agreements are essential tools enabling clients and telco operators to specify required quality of service. The 5GTANGO NFV platform enables SLAs through policies and custom service lifecycle management components. This allows the operator to trigger certain lifecycle management events for a service, and the network service developer to define how to execute such events (e.g., how to scale). In this demo we will demonstrate this unique 5GTANGO concept using an elastic proxy service supported by a high availability SLA enforced through a range of traffic regimes.**

I. Introduction

In Network Function Virtualization, operators often provide network services from independent network service developers. Telecom operators commit to SLAs with customers to define adequate QoS levels for these services. However, only service developers are aware about the internal composition and behaviour of the service. Fulfilment and assurance of SLAs therefore requires mechanisms to execute service-specific management actions upon events induced by the SLA.

The 5GTANGO Service Platform [1] is a modular NFV platform that manages network services throughout their end-to-end lifecycle, from instantiation to termination. It provides extended support for SLA commitments between customers and the operator that owns the platform, and for their enforcement. The platform combines a policy mechanism with a customisable MANO Framework to allow the operator and the service developer to join forces in managing their lifecycles to fulfil negotiated SLAs. This shared service management is the novelty that we highlight and demonstrate in this paper.

II. 5GTANGO Service Platform

The three 5GTANGO components that convert monitoring data into an operational SLA mechanism (Fig 1) are: i) SLA Manager, ii) Policy Manager and iii) MANO Framework.

The SLA Manager documents and tracks the relation between the operator and the customer. It allows the operator to create a set of SLA templates associated to a network service. These templates are blueprints of an SLA and describe what type of QoS commitments the operator is willing to take, and which monitoring parameters capture whether they are satisfied. Once a customer requests to deploy a network service, he or she selects one of these templates which is then converted by the SLA Manager into an enforced SLA. The SLA Manager monitors the associated metrics and evaluates whether the SLA is being violated or not.

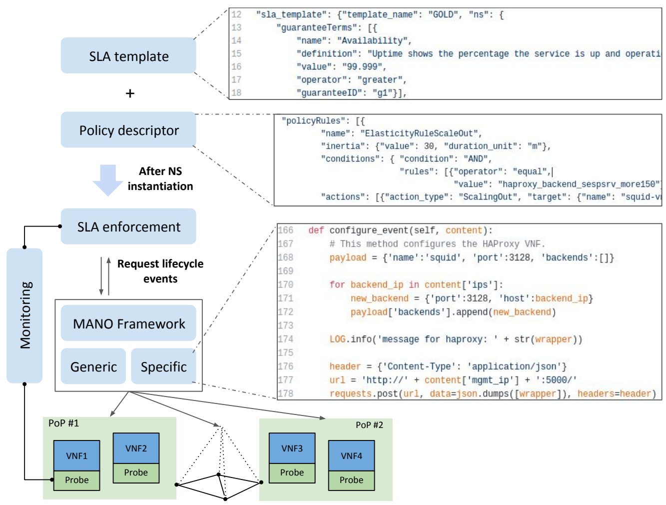
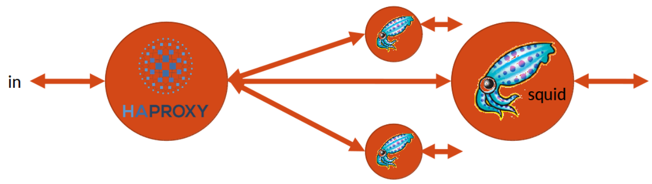


Fig. 1: SLA mechanism with SLA template, policy descriptor and FSM code snippet.

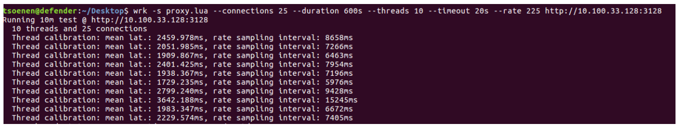
The Policy Manager allows the operator to describe which lifecycle events should be triggered when pre-defined regime changes are detected. Examples of such lifecycle events are scaling, migration and reconfiguration. This mapping is described in the policy descriptor and is defined by the operator. A policy descriptor is network service specific, and is activated when the service is instantiated. A policy descriptor can be associated to an SLA template. This gives operators the toolbox to create policies that work for an SLA, to prevent a violation. E.g., to satisfy a high availability SLA, the operator can create policies to scale the service when its load increases.

The MANO Framework is responsible to execute all lifecycle events for network services. It exposes an API where other components, like the Policy Manager, can request such lifecycle events. 5GTANGO introduces a unique service-specific management mechanism [2] enabling the developer to customise MANO behaviour with optimised workflows. This is a necessity to support non-trivial VNFs and network services. How to configure a VNF through its specific API or configuration interface can't be described in a generic way. Complex lifecycle events such as scaling require input from

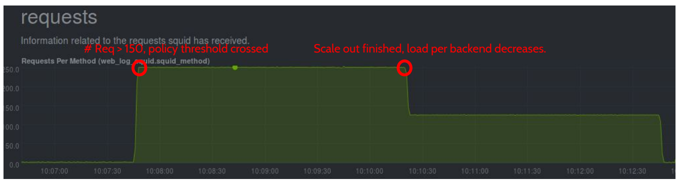
978-3-903176-15-7 © 2019 IFIP



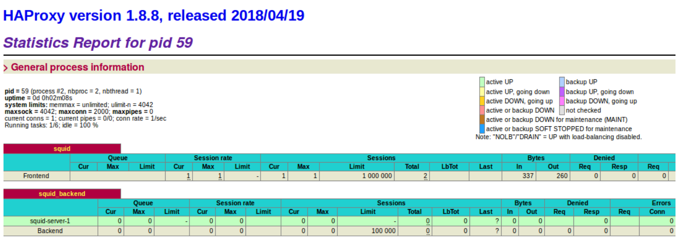
(a) Proxy service topology.



(b) Creating users with wrk.



(c) Requests per squid server before and after scale out event.



(d) HAProxy GUI indicating its current configuration and load.

Fig. 2: Prototype illustrations.

the developer that indicates which VNF to add, where to add additional VNF instances and how to reconfigure the networking between them. To this end, the MANO accepts Service (SSM) and Function Specific Managers (FSM) attached to network services and VNFs [2]. They overwrite the MANO logic for that specific component with logic of their own. This feature allows the 5GTANGO platform to support any network service, independent of the complexity of its lifecycle events, and therefore a wider variety of SLAs and policies, as they build on top of the extended set of lifecycle events it provides.

III. Demonstration Scenario

We demonstrate the 5GTANGO SLA feature by means of an elastic proxy service (Fig. 2a) that contains two VNFs: an HAProxy VNF configured as a load balancer and a Squid VNF configured as a proxy server. Clients use the ingress interface of the HAProxy as proxy IP, the HAProxy forwards incoming requests to one of the Squids in its backend pool. The service scales by adding or terminating a Squid instance and reconfiguring the HAProxy so its pool of backends reflects reality. This reconfiguration is HAProxy specific, as we need to obey its API, and thus requires an FSM that executes it when triggered by the MANO. The service also requires a scaling SSM to inform the MANO which scaling steps are needed. When consulted, this SSM instructs the MANO to i) add/remove a Squid VNF and ii) trigger the HAProxy FSM with the required config data. The proxy service with all its artefacts is available on GitHub[1] and in the catalogue of the 5GTANGO platform.

In the first part of the demo, we show how an operator configures the SLA templates and associated policies for a network service through the 5GTANGO Portal. The operator creates an SLA template for the proxy service that guarantees an availability of four nines, which converts in a maximum of 8 seconds downtime in each 24 hour period. Next, the operator creates the policies for this SLA template, describing how the service should be configured in various regimes to prevent SLA violations. The created policies describe that the MANO should scale the service to ensure the availability. Each policy describes either a scale out (adding a Squid) or a scale in (terminating a Squid) event when the number of users crosses a certain threshold and a relaxation period has passed.

In the second part of the demo, the proxy service is instantiated form the portal with this SLA template selected. The instantiation has two datacenters available, managed by an OpenStack setup. Once the deployment is completed, we show that i) the VNFs are spread among the datacenters in accordance with the selected placement policy, ii) the SLA template was promoted to an actual agreement that is being enforced and iii) the associated policies are active.

Next, we create load for the service to see the policies in action. We use the *wrk*[2] tool (Fig. 2b) to create enough users for the proxy service so that the scaling policy threshold (> 150) is crossed. At this moment, the Policy Manager instructs the MANO Framework to scale the service out. When completed, both the portal and the HAProxy GUI indicate that an additional Squid is available, that the scaling event was successful and that the average load per squid was reduced (Fig. 2c). By further increasing, or reducing, the number of users, additional scaling events are triggered. This shows how the Policy Manager is working to aid in guaranteeing its SLAs, and how it uses input from the network service developer through the scaling SSM and the HAProxy configuration FSM.

As the final step, we manually terminate the remaining Squid instances to emulate some kind of issue with the datacenters. At this moment, the proxy service is no longer available. When this unavailability takes longer than 8 seconds, the SLA becomes violated. Although the scale out policy will directly trigger the instantiation of a new Squid instance, this instantiation takes longer than 8 seconds. Therefore, the demo ends with the SLA being violated, as indicated on the portal.

Acknowledgement

Partially Funded by the EC through 5GTANGO (761493).

References

[1] 5GTANGO Project, "D5.1 First open-source release of the Service Platform," https://5gtango.eu/project-outcomes.html, Apr. 2018.
[2] T. Soenen *et al.*, "Empowering network service developers: enhanced nfv devops and programmable mano," *Accepted to IEEE Comm. Magazine*.

[1]https://github.com/sonata-nfv/tng-y1-demo
[2]https://github.com/wg/wrk