An Integrated SLA Management Framework in a 5G Environment

Marios Touloupou, Evgenia Kapassa, Chrysostomos Symvoulidis, Panagiotis Stavrianos, Dimosthenis Kyriazis
Department of Digital Systems
University of Piraeus
Piraeus, Greece
{mtouloup, ekapassa, simvoul, pstav, dimos}@unipi.gr

***Abstract*—A key feature of fifth generation (5G) and Software Defined Networking (SDN) is the assurance of high levels of the quality of service (QoS). To this end, Service Level Agreements (SLAs) are introduced in order to fulfill the gap between network operators and their customers. An SLA is a contract between the operator and the internal or external customer, which determines what Network Services (NSs) are offered and the guaranteed level of performance. Taking into consideration the above-mentioned needs, in this paper, we are introducing a fully integrated SLA management framework in a real 5G environment. In this demonstration we aim to bind business requirements as Service Level Objectives (SLOs) between network operators and the customers, with measurable recourse attributes. To achieve this, we allow network operators to choose between different SLOs during the SLA Template generation, and then automatically formulate an Agreement, based on each network slice instantiation and the corresponding NS. Finally, we provide a monitoring system in order to detect and alert for any violations.**

Keywords*—, 5G networks, SLA management, quality of service, slicing, monitoring

I. INTRODUCTION

5G networks are anticipated to provide exponentially more capacity, lower energy consumption, lower latency, ubiquitous connectivity, as well as increased reliability and availability. Therefore, 5G comes as a solution to network operators for the optimum management of their infrastructures, allowing higher flexibility on the way the resources are handled [1]. Challenges such as hosting several services under the same infrastructure with conflicting requirements and at the same time provide the optimal QoS, may now be solved through network slicing; a method that enables the partition of traditional network resources into virtual elements for the creation of multiple virtual networks on top of the physical infrastructure.

Here lies the significance of SLAs in the evolving 5G environments. The customers can negotiate specific QoS levels with the telecom operators, hence leading to the creation of SLAs that state the desired quality levels given by them. Consequently, an estimation of the required resources, in order to fulfill the above-mentioned demands per network slice, is necessary.

The advantages of enabling SLAs exist for both the operators, and the customers, as they come as an assurance that NSs and their corresponding Virtual Network Functions (VNFs) can be controlled efficiently. The abilities of such technologies can be maximized through the proper instantiation of Network Slices among with the corresponding SLA.

The remaining of the paper is structured as follows. Section II describes the end-to-end workflow of the SLA Management Framework, where Section III presents the demonstration of the proposed framework. In this demonstration, as soon as the generated SLAs are on-boarded to the 5GTANGO service platform, we illustrate the deployment of an example NS on a network slice and monitor its performance [2]. Section IV concludes the paper.

II. SLA MANAGEMENT IN 5G ENVIRONMENTS

It is adoptable that the SDN explosion could not leave unaffected the evolution of SLA models, and their flexibility in adapting more demanded parameters. The SLA Management Framework presented in this demo paper is a web, multi-platform application that allows to manage the whole lifecycle of SLAs, from the template creation to the agreement violation detection. It is a plugin-based decoupled component that can be adapted and extended to work on different 5G-oriented service platforms. Additionally, it is an open source project, released under the Apache 2 license.

Figure 1 presents the overall workflow which consists of four discrete phases. All the interactions and management requests are based on Rest APIs and include the following: a) Definition and advertisement of the capabilities of network operators in SLA Template forms, b) Creation and management of the agreements per network slice, c) Monitoring the agreement compliance during runtime and d) Management of the SLA violations records. It is worth mentioning that an agreement between a network operator and a customer specifies one or more SLOs, that are expressions of the requirements of the customers and of the assurances of the network operators.

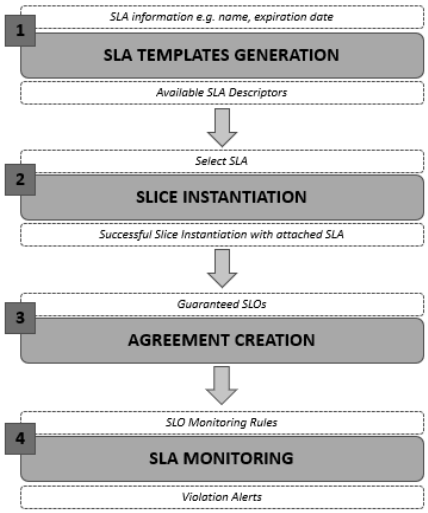

Fig. 1: SLA Management Overall Workflow

A. SLA Template Generation

To start with, the definition of an SLA template by a commercial officer, on behalf of the network operator, is taking place. During the SLA Template generation four (4) parameters are mandatory for the successful creation. The most critical one is the selection of the NS that is going to be correlated with the newly created SLA template. At the same time an SLA name among with a valid future expiration date and at least one SLO are needed. The result is the generation of the SLA template in form of a descriptor, and the storage of it in a NoSQL database [3,4]. The aforementioned SLA Descriptor (SLAD) is based on WS-Agreement specification [5,6], as depicted in Figure 2. The WS-Agreement specification describes an XML schema for specifying service level agreements (both applicable to SLA Templates, and Agreements). In addition, the SLA management framework keeps a record in an internal database, PostgreSQL [7], with the corresponding correlation.

B. Network Slice instantiation with SLA

At the next phase, let's assume that a slice customer visits the commercial officer asking for a NS. The commercial officer presents the available NSs with the corresponding SLA templates as the operator's initial offers. The customer is giving some personal details and the flow continues in order to instantiate the NS in a specific network slice of the network operator's 5G infrastructure. The request is published in a Message Queue System (MQ) [8]. Except from the MANO framework [9], that is responsible to consume this request-message and instantiate the service, the SLA management framework consumes the message as well. A correlation between the SLA and the customer is created and stored in an internal database (DB) with status *"pending"*.

```
{
  "name": "silver-template-example",
  "vendor": "UPRC",
  "version": "2.0",
  "author": "Evgenia Kapassa, Marios Touloupou",
  "description": "This is a silver SLA Template for Haproxy
  "sla_template": {
    "template_name": "Silver",
    "offer_date": "2018-11-14T11:35:10Z",
    "expiration_date": "2019-11-14T11:35:10Z",
    "provider_name": "Telefonica",
    "template_initiator": "Evgenia Kapassa",
    "service": {
      "ns_uuid": "0e69ccfd-d9ba-4439-99b8-cd4f2a059457",
      "ns_name": "ns-squid-haproxy",
      "ns_vendor": "eu.5gtango",
      "ns_version": "0.2",
      "guaranteeTerms": [
        {
          "guaranteeID": "g1",
          "guarantee_name": "Availability",
          "guarantee_threshold": "98",
          "guarantee_operator": "greater",
          "guarantee_unit": "%",
          "guarantee_period": "Daily",
          "guarantee_definition": "",
          "guarantee_service_level": "1728sec/24h",
          "target_slo": [
            {
              "target_kpi": "Downtime",
              "target_value": "1728s",
              "target_operator": "less",
              "target_duration": "10s",
              "target_period": "24h",
              "target_service_level":"Downtime less 1728s"
            }
          ]
        }
      ]
    }
  }
}
```

Fig. 2: SLA Descriptor Example

C. Agreement Creation

As long as the MANO framework instantiate the service successfully, another message is produced giving the status of the instantiation. When a message with status *"ready"* is consumed by the SLA Manager, a method is triggered where the status of the correlation between the SLA and the customer is set to *"ready"*. The agreement is now officially attached to the slice instance, and the attached NS starts being monitored.

D. SLA Monitoring

It is worth mentioning that the agreement is automatically created after the successful slice instance creation. Thus, the monitoring manager should be informed in order to provide monitoring data to the SLA management framework. To this end, when the NS is successfully instantiated, monitoring rules depending on the signed SLO are formulated and sent to the monitoring manager. The monitoring manager receives the rules and start capturing data for the specific NS instance [10]. When a violation of the SLA is realized by the monitoring manager, an alert is produced and published to the MQ system. The SLA manager which is a consumer of that queue, consumes the alert and updates the status of the SLA to *"violated"*. The violation data is stored in the internal database in order to be available for visualization in a unified portal.

III. DEMONSTRATION

In order to demonstrate the functionality of the proposed workflow in terms of efficiency and ease of use, the SLA Management Framework was included in the 5G infrastructure environment of the 5GTANGO Service Platform [2]. 5GTANGO project, a 5G Development and Validation Platform for global Industry specific Network Services and Apps, is an EU funded Innovation Action, which enables the flexible programmability of 5G networks with, among others, a modular Service Platform with an innovative orchestrator, so it can bridge the gap between business needs and network operational management systems [11]. As it is presented in Figure 3, Sally, the commercial officer formulates and links an SLA template with an available NS. For the demonstration purposes, we use an elastic proxy network service, which is consisted of two chained VNFs: a) a HAProxy VNF, configured as a load balancer and b) a Squid VNF configured as a proxy server. The end-users use the ingress interface of the HAProxy as

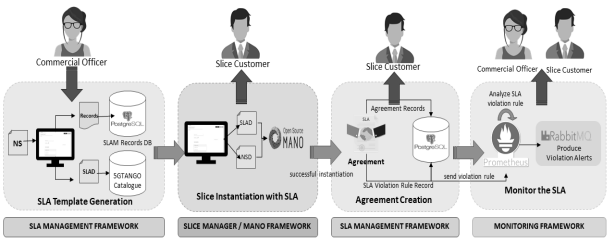

Fig. 3: SLA Management Framework Demonstration

Proxy IP, and sequentially the HAProxy forwards the incoming requests to one of the Squids in its backend pool [12,13].

The commercial officer is now able to define guarantees like availability of 99,95% calculated over 24 hours. In our case, we assume that the availability of the service is the total time where Squids are available to response to incoming requests from the HAProxy. The latter is depicted in Figure 4.

As a next step, the instantiation process of the NS to the 5GTANGO Service Platform is being triggered. Thus, the commercial officer instantiates a network slice where the selected NS is attached with the corresponding SLA template. When the network slice – service is deployed and ready for use, the final agreement is automatically generated by the SLA manager while it consumes the messages from the MQ system coming from MANO framework [14]. As part of the demonstration an SLA violation should be also occurred and captured by the SLA manager. For this reason, we access the Point of Presence (PoP) that the service is deployed into, and we force delete the last chained Squid VNF. At this point, the monitoring manager realizes the backend downtime of the Squid and triggers the rule where an alert is produced and sent towards the SLA manager. Consuming this alert, means that although the monitoring data informs about breaches, violations only will be raised when the mean of these measurements over the specified interval exceeds the defined threshold by the SLA management framework. Once the SLA becomes violated, a violation record is created and both stakeholders (i.e. network operator and customer) are informed.

The entire workflow described in the above paragraphs can be executed on a regular laptop and the code of all involved software components as well as their install instructions and documentation are available on the project's official web site [15], as well as on GitHub [16]. A video showing the demonstration is also available on YouTube [17].

IV. CONCLUSIONS

In 5G environments, network operators with clear definition of business terms defined in the provided SLA can increase the reliability and trust level of operator-customer relationship. Thus, the demonstrated SLA Management Framework is an important part of the provision of a NS, especially in 5G networks. After the successful NS instantiation into a network slice, the SLA framework is populated with infrastructure's recourse information, in order to monitor the agreement with real-time usage data, to efficiently avoid or/and manage unexpected violations. Hence, network operators can focus more on how to please their customers and provide them high quality by any mean. Nevertheless, provisioning of resources in a virtualized 5G infrastructure is a challenging task, that still needs a lot of investigation and testing. Therefore, we plan to extend the framework in order to enable Quality of Experience (QoE) enforcement. Additionally, the proposed framework is able to monitor and manage business guarantees in a single NS per slice. Therefore, we envision to manage SLA with multiple NSs per slice, enabling higher level of abstraction.

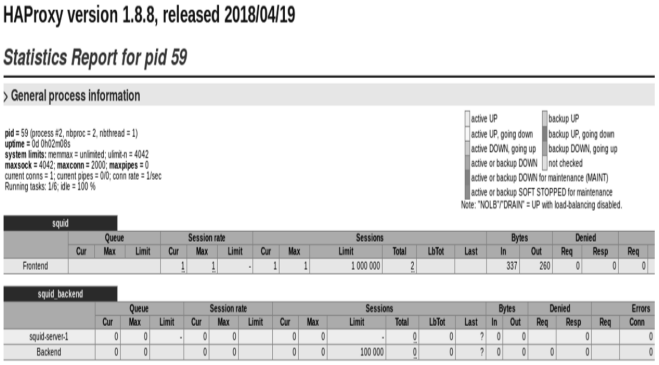

Fig. 4: HAProxy GUI indicating general process information

ACKNOWLEDGMENT

This work has been partially supported by the 5GTANGO project, funded by the European Commission under Grant number H2020ICT-2016-2 761493 through the Horizon 2020 and 5G-PPP programs (http://5gtango.eu).

REFERENCES

[1] F. Hu, Opportunities in 5G Networks: A Research and Development Perspective, CRC Press, 2016.

[2] C. Parada et al., "5Gtango: A beyond-mano service platform," European Conference on Networks and Communications (EuCNC), pp. 26-30, 2018.

[3] F. Gessert, W. Wingerath, S. Friedrich, and N. Ritter, "NoSQL database systems: a survey and decision guidance," Computer Science-Research and Development, vol. 32, pp. 353-365, 2017.

[4] P. Twamley et al., "5Gtango: An approach for testing NFV deployments," European Conference on Networks and Communications (EuCNC), pp. 1-218, 2018.

[5] A. Andrieux et al., "Web Services Agreement Specification (WS Agreement)", Grid Resource Allocation Agreement Protocol (GRAAP) WG, 2011.

[6] Kabert, R., Katsaros, G., and Wang T. "A RESTful implementation of the WS-agreement specification". In Proceedings of the Second International Workshop on RESTful Design, 67-72, 2011.

[7] PostgreSQL:The world's wost advanced open source relational database, https://www.postgresql.org/

[8] RabbitMQ, https://www.rabbitmq.com/

[9] M. Ersue, "ETSI NFV Management and Orchestration - An Overview", Online: https://www.ietf.org/proceedings/88/slides/slides-88-opsawg-6.pdf

[10] P. Trakadas et al., "Scalable monitoring for multiple virtualized infrastructures for 5G services," The International Symposium on Advances in Software Defined Networking and Network Functions Virtualization (SoftNetworking), pp. 1-4, 2018.

[11] 5GTANGO Project Consortium, "5GTANGO: 5G Development and Validation Platform for global Industry-specific Network Services and Apps", Online: https://5gtango.eu/

[12] HAProxy: The reliable, high performance TCP/HTTP load balancer, http://www.haproxy.org/

[13] Squid: Optimising Web Delivery, www.squid-cache.org/

[14] T. Soenen, M. Peuster, and F. Vicens, "SONATA's Service Platform MANO Framework", https://github.com/sonata-nfv/son-mano-framework

[15] 5GTANGO Consortium, Documentation, https://5gtango.eu/software/documentation.html

[16] E. Kapassa and M. Touloupou, "The 5GTANGO SP SLA Management Repository," https://github.com/sonata-nfv/tng-sla-mgmt/tree/v4.0, 2018.

[17] M. Touloupou, E. Kapassa, "5GTANGO SLA Manager" Demo, https://www.youtube.com/watch?v=_dtSAi6KjN4, 201