Proceedings of the 2019 IEEE International Conference on Information Technologies (InfoTech-2019)
19-20 September 2019, St. St. Constantine and Elena, Bulgaria

Toward Service Based Radio Access Network

Evelina Pencheva, Denitsa Velkova, Ivaylo Atanasov, Ventsislav Trifonov
Department of Communication Networks, Faculty of Telecommunications
Technical University of Sofia
8 Kliment Ohridski blvd., 1000 Sofia, Bulgaria
enp@tu-sofia.bg, denitsa.velkova@balkantel.net, {iia,vgt}@tu-sofia.bg

Abstract – **Service Oriented Architecture (SOA) capture the idea of network composed of services. SOA enables service virtualization and it is applied to the fifth generation (5G) core network design. In this paper, we propose an approach to define 5G Radio Access Network functionality as a set of services. The paper focus is on functions related to transferring of Non Access Stratum (NAS) signaling between the user equipment and the core network.**

Keywords – **Radio resource management; Representational State Transfer; Service Oriented Architecture; State models.**

I. INTRODUCTION

Future fifth generation (5G) networks are expected to be flexible and programmable. The 5G core network design is service-based and the communication between services uses HTTP based Application Programming Interfaces (APIs) replacing telecommunication protocols [1]. Network programmability facilitates introduction of new functions and empowers flexible and dynamic deployment of new services. Services may be run as virtual machines in different network segments. These virtual machines which decompose monolithic network functionality will represent interworking units in next generation systems [2].

The APIs enable service exposure to third parties which in turn may provide innovative applications. These applications may improve network utilization by customization of the behavior of network resources.

One of the main challenges of network programmability is to identify elementary functions which compose network functionality. In this paper, we propose an approach to design function of Next Generation Application Protocol (NGAP) as services. The NGAP is defined between 5G Radio Access Network (RAN) and AMF (Access and mobility Management Function) in the core network [3].

In [4], the authors study network slicing as a key enabler for transition to RAN-as-a-service. Each network slice includes an independent set of logical network functions that support specific use case. Design of RAN as a set of services provides multiservice environment which can be customized and programmed for different slices. In [5], the authors propose an end-to-end network architecture which follows the principles of software defined networking for control plane separation and centralization. The control plane functions are decomposed and defined as services which guarantees end-to-end quality of service and improves user's quality of experience. Service-based design facilitates the transition to virtualized access network where network infrastructure is separated from control functionality. Each component, which plays an essential role, may be regarded as a service and supplied by different service providers and operators. The resulting system architecture considers Everything-as-a-Service [6].

The focus of this paper is on RAN functionality for transferring of Non Access Stratum (NAS) signaling between a User Equipment (UE) and Core Network (CN) which is presented as a service. First we describe the proposed service by typical use cases. Then the resource structure and the supported HTTP method are presented. Service implementation aspects are considered in terms of modeling the resource state as supported by RAN and CN.

II. SERVICE DESCRIPTION BY TYPICAL USE CASES

The proposed RAN service enables NAS message transfer. The interaction between the service and the AMF follows the Representational State Transfer (REST) architectural style. In REST, each entity is represented as a resource which is uniquely identified. The resource has a state, which can be manipulated by four operations, namely CREATE, READ, UPDATE and DELETE. These operations may be implemented by HTTP requests. HTTP is used over TCP. The interaction between the proposed RAN service and AMF supports "request-response" and "subscribe-notify" mechanisms.

In order to receive notifications from the RAN node, the AMF need have a subscription. The AMF may subscribe for receiving Initial UE messages, uplink NAS messages and for receiving indications about non NAS delivery. In order to make a subscription, the AMF sends a POST request with message body containing {**notificationSubscription**) data structure to the resource representing the respecttive subscription. The {**subscriptionType**} is replaced with the one of the following: **initialUEmessages** (for receiving notifications about initial UE messages), **uplinkMessages** (for receiving notifications about uplink NAS messages), or **nonNASdeliveries** (for receiving notifications about non NAS delivery). The {**notificationSubscription**} data structure is replaced with the respective data type specified for the different subscriptions. The respective {**notificationSubscription**} data structure contains the filter criteria for sending notifications and the address at which the AMF wants to receive notifications. The RAN node answers with "201 Created" response, which contains the address of the resource created. Fig.1 shows the flow for subscribing to NAS transfer related events.

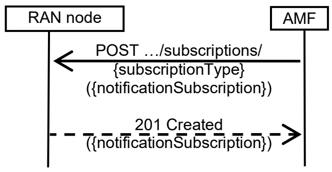

Fig.1 Flow of subscribing to NAS transfer related events

978-1-7281-3274-7/19/$31.00©2019 European Union

The AMF may modify an existing subscription by sending a PUT request to the resource representing the subscription with message body containing the requested subscription modifications (e.g. extension of the subscription duration). The AMF may terminate an existing subscription by sending a DELETE request to the resource representing the subscription.

In order to respond to paging message, or to initiate outgoing communications, the UE uses the Initial UE Message procedure which transfers the first uplink NAS message dedicated to the AMF. The UE needs to establish a RRC connection with the network. The first uplink NAS message is included in the request for RRC connection setup.

If the AMF has a subscription for notifications about initial UE messages, the RAN node sends a POST request to notify about the event to the callback address provided by the AMF. The notification contains **InitialUEMessage** data structure including the NAS message, UE user location information and RRC establishment cause. The AMF sends “204 No content” response to the RAN node.

Over established RRC connection the RAN node and UE may transfer NAS messages. When the AMF wants to send a NAS message to the UE it sends a POST request to the resources representing the downlink NAS messages. The request contains **dlNASMessage** data structure including the NAS message. The RAN node responds with “201 Created” containing the address of the resource created and the NAS message type. The RAN node uses **DLInformationTransfer** message to deliver the message to the UE. When the UE wants to send a NAS message to the AMF it uses **ULInformationTransfer** message. If the AMF has subscription for receiving notifications about Uplink NAS messages, the RAN node sends a POST request to the AMF to notify about the event uplink NAS message. The notification contains **ulNASMessage** data structure containing the NAS message and UE location. The AMF sends “204 No content response to the RAN node.

Fig.2 illustrates the flow for initial UE message transfer, downlink and uplink NAS message transfer.

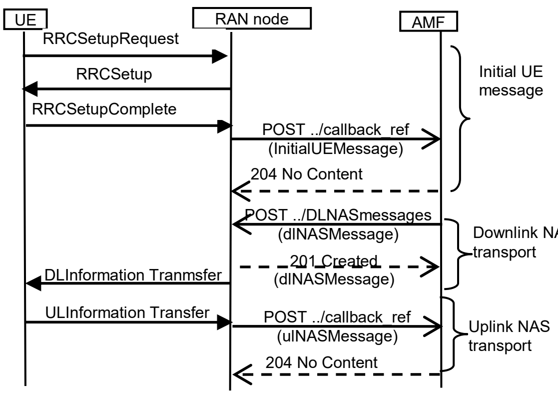

Fig.2 Initial UE transfer message and subsequent downlink and uplink transport of NAS messages

The RRC connection is suspended by the network when the UE is inactive for some time. When the UE wants to send a NAS message to the network, it first performs RRC resume procedure as shown in Fig.3.

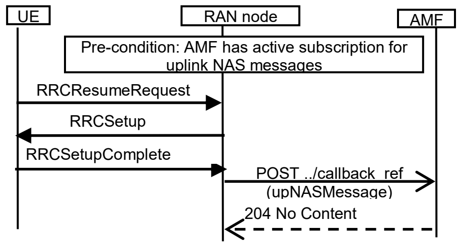

Fig.3 Flow of sending uplink NAS message after RRC connection resume

The RAN node may not be able to deliver a downlink NAS message or it may be not sure that the message is received by the UE, e.g. due to intra systems handover or intersystem handover. In this case, the RAN node has to notify the AMF by sending NAS non delivery indication. The precondition for notification is that the AMF has active subscription for receiving NAS non delivery indications. The RAN node sends a POST request to the callback address provided by the AMF with message body containing **failedNASMessag**e data structure. The **failedNASMessage** data structure contains the NAS message and the cause. Fig.4 shows the message flow for notification about NAS non delivery.

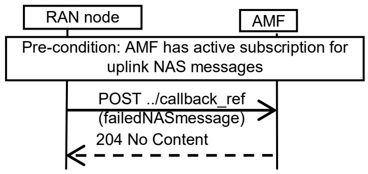

Fig.4 Flow of notification about NAS non delivery

The AMF may request from RAN node to reroute the Initial UE message to another AMF. For this purpose, the AMF sends a POST request to the resource representing NAS reroute requests with message body containing **rerouteNAS** data structure. The **rerouteNAS** data structure contains the Initial UE message itself and the ID of AMF set within the AMF region. The RAN node answers with “201 Created” containing the request ID. Fig.5 shows the respective message flow.

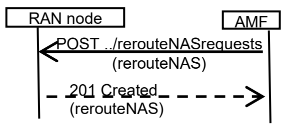

Fig.5 Flow of request for rerouting Initial UE Message

III. RESOURCE STRUCTURE AND SUPPORTED HTTP METHODS

Fig.6 shows the proposed resource structure, where all resources follow the same root. The subscriptions resource contains all subscriptions. The **initialUEmessages** resource contains subscriptions about receiving initial UE messages, while the **initialUEMessageID** represents an existing subscription for initial UE messages. The **uplinkNASmessages** resource contains subscriptions for receiving notifications about uplink NAS messages, while the **uplinkMessageID** represents an existing subscription of this type. The **nonNASdeliveries** resource contains subscriptions for receiving notifications about non NAS deliveries, while the **nonNASdeliveryID** represents an existing subscription for indications about failed NAS deliveries. The **downlink-Messages** resource is a container for downlink messages. The **rerouteNASrequests** resource is a container for requests for NAS rerouting.

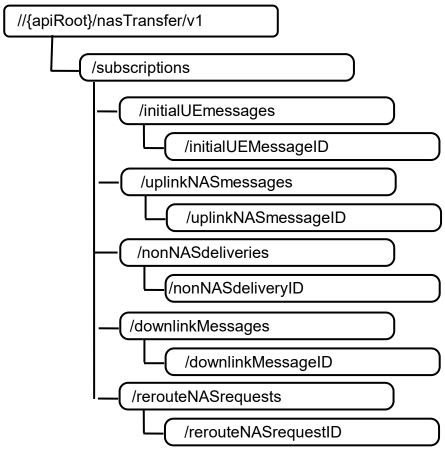

Fig.6 Resource structure supported by the proposed service

The container resources support HTTP GET method which retrieves a list of their children resources, and HTTP POST method, which creates a new child resource. The leaf resources support HTTP GET method which retrieves information about current resource, HTTP PUT method, which modifies the existing resource and HTTP DELETE method, which deletes the respective resource.

III. State Models

The implementation of the proposed service requires modeling the state of NAS message transfer as seen by the RAN node and by the AMF. Both models need to be synchronized.

Fig.7 shows the simplified NAS message transfer state as seen by the RAN node. The model reflects the RRC procedures [7] and the respective NAS message procedures. The timer is set with the maximum UE inactivity period for maintaining the established RRC connection. The RRC connection is released when the AMF orders the release of UE context in the RAN node.

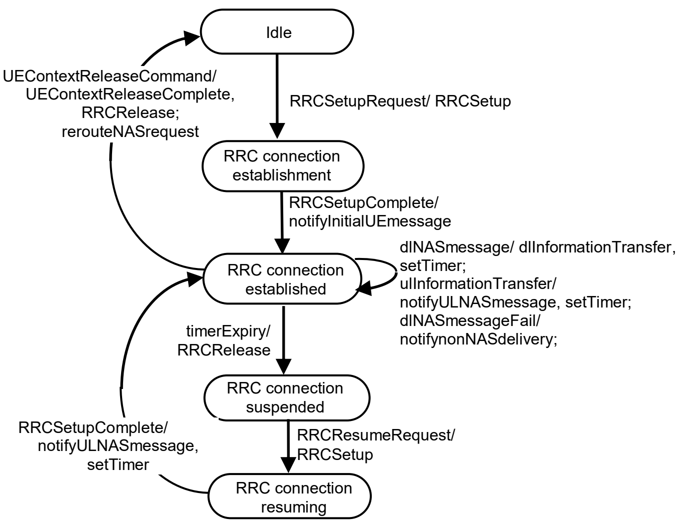

Fig.7 Simplified NAS transfer state model supported by the RAN node

Fig.8 shows the simplified NAS transfer state model supported by the AMF. The AMF requests re-routing of initial UE messages, if it is overloaded.

In order to provide that the both models expose equivalent behavior we formalize the models' descriptions. The model descriptions are formalized using the notion of Labeled Transition System (LTS).

A Labeled Transition System is represented as quadruple of a set of states, a set of actions, a set of transitions, and a set of initial states.

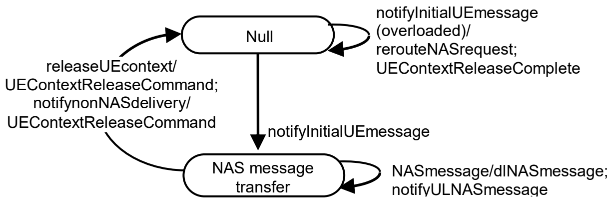

Fig.8 Simplified NAS transfer state model supported by the AMF

By $T_A = (S_A, Act_A, \rightarrow_A, s_0^A)$ it is denoted an LTS, representing the NAS message transfer state model supported by the AMF, where:

- S_A = {Null [s_1^A], NASmessageTransfer [s_2^A]};
- Act_A = {notifyInitialUEmessage [t_1^A], NASmessage [t_2^A], notifyULNASnessage [t_3^A], releaseUEcontext[t_4^A], UEContextReleaseComplete [t_5^A], notifyInitialUEmessage(overloaded)[t_6^A], notifynonNASdelivery [t_7^A]};
- $\rightarrow_A = \{(s_1^A\ t_1^A\ s_2^A), (s_2^A\ t_2^A\ s_2^A), (s_2^A\ t_3^A\ s_2^A), (s_2^A\ t_4^A\ s_1^A), (s_1^A\ t_5^A\ s_1^A), (s_1^A\ t_6^A\ s_1^A), (s_2^A\ t_7^A\ s_1^A)\}$;
- $s_0^A = \{ s_1^A \}$.

Short notations are given in brackets.

By $T_{RAN} = (S_{RAN}, Act_{RAN} \rightarrow_{RAN}, s_0^{RAN})$ it is denoted an LTS, representing the ERAB state model supported by RAN, where:

- S_{RAN} ={ Idle [s_1^N], RRCConnectionEstablishment [s_2^N], RRCConnectionEstablished [s_3^N], RRCConnection-Suspended [s_4^N], RRCConnectionResuming [s_5^N]};
- Act_{RAN} = {RRCSetupRequest [t_1^N], RRCSetupComplete [t_2^N], dlNASmessage [t_3^N], ulInformationTransfer [t_4^N], dlNASmessageFail [t_5^N], timerExpiry [t_6^N], RRCResumeRequest [t_7^N], UEContextRelease-Command[t_8^N], rerouteNASrequest[t_9^N]};
- $\rightarrow_{RAN} = \{ (s_1^N\ t_1^N\ s_2^N), (s_2^N\ t_2^N\ s_3^N), (s_3^N\ t_3^N\ s_3^N), (s_3^N\ t_4^N\ s_3^N), (s_3^N\ t_5^N\ s_3^N), (s_3^N\ t_6^N\ s_4^N), (s_4^N\ t_7^N\ s_5^N), (s_5^N\ t_2^N\ s_3^N), (s_3^N\ t_8^N\ s_1^N), (s_3^N\ t_9^N\ s_1^N)\}$;

- $s_0^{RAN} = \{ s_1^N \}$.

The formal proof that both models expose equivalent behavior uses the concept of weak bisimulation. The LTSs expose equivalent behavior if they display the same result. In strong bisimulation there has to exist a strong mapping between the states and transitions of both LTSs. In weak bisimulation, internal transitions may be ignored. A precise definition of bisimulation may be found in [8].

Proposition: T_A and T_{RAN} are weakly bisimilar.

Proof: In order to prove the LTSs bisimilarity, we identify the relation U_{ARAN} between their states. We prove that for any event triggering transitions from one state in U_{ARAN} to other state in U_{ARAN} for the one LTS, there are respective transitions from the corresponding state in U_{ARAN} to the corresponding state in U_{ARAN} for the other LTS.

$U_{ARAN} = \{(s_1^A, s_1^N), (s_2^A, s_3^N)\}$. Then, for each of the following events the following transitions for the U_{ARAN} states are identified:

1. The UE sends an Initial UE message which is successfully delivered to the AMF: for $(s_1^N\ t_1^N\ s_2^N)$, $(s_2^N\ t_2^N\ s_3^N) \exists (s_1^A\ t_1^A\ s_2^A)$.
2. The AMF sends a NAS message successfully: for $(s_2^A\ t_2^A\ s_2^A) \exists (s_3^N\ t_3^N\ s_3^N)$.
3. The AMF sends a NAS message and it is notified for non NAS message delivery: for $(s_2^A\ t_2^A\ s_2^A)$, $(s_2^A\ t_7^A\ s_1^A) \exists (s_3^N\ t_3^N\ s_3^N), (s_3^N\ t_5^N\ s_3^N), (s_3^N\ t_8^N\ s_1^N)$.
4. The UE sends a NAS message: for $(s_3^N\ t_4^N\ s_3^N) \exists$ $(s_2^A\ t_3^A\ s_2^A)$.
5. The RRC connection is suspended due to UE inactivity and then resumed for sending an uplink NAS message: for $(s_3^N\ t_6^N\ s_4^N)$, $(s_4^N\ t_7^N\ s_5^N)$, $(s_5^N\ t_2^N\ s_3^N) \exists$ $(s_2^A\ t_3^A\ s_2^A)$.
6. The UE sends an Initial UE message, but the AMF is overloaded and request rerouting the NAS message: for $(s_1^N\ t_1^N\ s_2^N)$, $(s_2^N\ t_2^N\ s_3^N)$, $(s_3^N\ t_9^N\ s_1^N) \exists$ $(s_1^A\ t_6^A\ s_1^A)$.
7. The AMF decides to release the UE context: for $(s_2^A\ t_4^A\ s_1^A), (s_1^A\ t_5^A\ s_1^A) \exists (s_3^N\ t_8^N\ s_1^N)$.

Therefore T_A and T_{RAN} are weakly bisimilar, i.e. they expose equivalent behavior. ■

IV. Conclusion

In this paper, we present an approach to define transport of NAS messages procedure supported by NGAP protocol as a service. The proposed service is described by typical use cases illustrating its functionality. The message exchange between the proposed service and AMF is based on REST and supports both request-response, and subscribe-notify interaction mechanisms. Some service implementation issues are considered in terms of modeling the NAS message transfer state. We provide two NAS transfer models that may be supported by the RAN node and by the AMF and formally prove that both models expose equivalent behavior.

With the proposed approach the interaction between NGAP services and AMF services, which are based on the same protocol, is seamless, faster and consistent as any protocol translation may lead to inconsistence and at least delays. Further, the approach improves modularity and facilitates the transition to RAN function virtualization.

The level of delay predictability is entirely dependant on the predictability of delays introduced by the IP based transport between the core network and RAN.

The adoption of programmable infrastructure although it facilitates the orchestration and management procedures requires additional security considerations. Decoupling of control plane from user plane makes the control plane elements susceptible to attacks such as a denial of service. In order to properly secure the 5G network, security and trust paradigm is more completed than in previous generations. Apart from having trust and security domains associated to the participants in communication links, there are also domains associated with virtualized and non-virtualized domains. Among the others, network slicing and overload control mechanisms might be applied to protect against signaling denial of service attacks and compromised RAN elements.

Acknowledgment

The research is conducted under the grant of project DH07/10-2016, funded by Bulgarian National Science Fund, Ministry of Education and Science.

The authors would like to thank to the Research and Development Sector at Technical University of Sofia for the financial support.

References

[1] H. Rahimi, A. Zibaeenejad, A.A. Safavi, "A Novel IoT Architecture based on 5G-IoT and Next Generation Technologies", arXiv e-prints, arXiv:1807.03065, 2018

[2] J. Ni, X. Lin and X. S. Shen, "Efficient and Secure Service-Oriented Authentication Supporting Network Slicing for 5G-Enabled IoT," IEEE Journal on Selected Areas in Communications, vol. 36, no. 3, pp. 644-657, March 2018.

[3] 3GPP TS 38.413 Technical Specification Group Radio Access Network; NG-RAN; NG Application Protocol (NGAP), v15.2.0, 2019, Release 15.

[4] N. Nikaein, C.-Y. Chang, "Slicing and Orchestration in Service-Oriented RAN Architecture," IEEE Softwarization, December, 2017.

[5] M. Yang, Y. Li, B. Li, D. Jin, S. Chen, "Service‐oriented 5G network architecture: an end‐to‐end software defining approach", International Journal of Communication Systems, vol. 29, issue 10, pp. 1645-1657, 2016

[6] Z. Chang, Z. Zhou, S. Zhou, T. Chen and T. Ristaniemi, "Towards Service-Oriented 5G: Virtualizing the Networks for Everything-as-a-Service," IEEE Access, vol. 6, pp. 1480-1489, 2018.

[7] 3GPP TS 38.331 Technical Specification Group Radio Access Network; NR; Radio Resource Control (RRC) protocol specification, Release 15, v15.4.0, 2019.

[8] Chen, X. J., De Nicola, R. Algebraic characteristics of trace and decorated trace equivalences over tree-like structures, Theoretical Computer Science, vol.254, issue 1-2, 2001, pp.337-361.