Edge Cloud Deployment with 3GPP 5G Stand Alone

Introduction

The 3GPP Release 15 specification (3GPP TS 23.501) introduced Edge Computing. Edge Compute is highlighted as a key deployment mechanism for delivering services to end users by placing applications closer to the user. Network and Enterprise operators are taking advantage of this advancement to provide edge services that are low-latency, user-centric, and secure.

This white paper focuses on some key challenges for 5G Standalone (SA) mode edge deployments and outlines how OpenNESS helps to address them. The next version of this white paper will address the 5G Non-standalone (NSA) mode.

Two key challenges in edge deployments are:

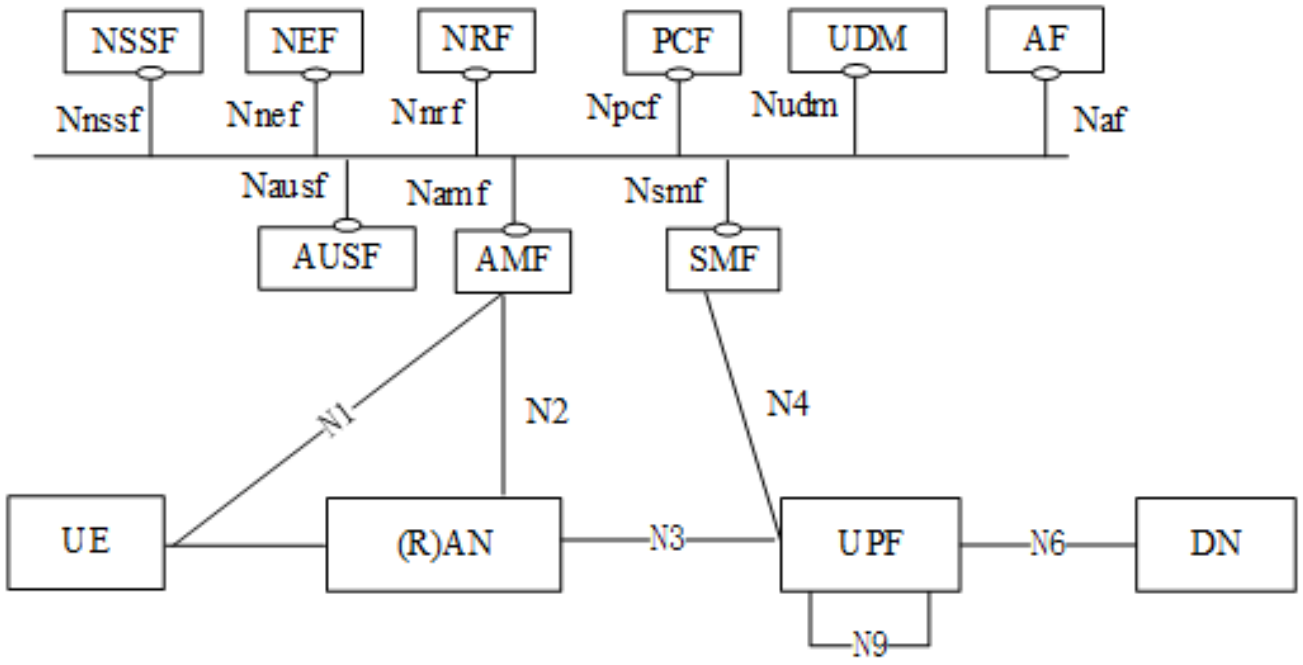
1. **UPF selection for UEs**
    - 3GPP standards have multiple references for the UPF selection procedure, which provides flexibility for implementation. Therefore, choosing the right implementation is a key factor in edge deployments.
2. **UE Traffic identification and steering within the UPF towards edge node interfaces**
    - The standard clearly outlines REST-based APIs for this purpose in 5G.

OpenNESS provides REST-based reference APIs along with 3GPP standard traffic influencing APIs (using the Application Function) to address some of these major challenges in 5G edge deployments.

5G Systems Architecture

The 5G system architecture specified by the 3GPP standard 3GPP_23501 addresses various use cases (ranging from serving simple IoT devices to critical services) where there is a need for a high bitrate and reliability. Although the 5G systems architecture includes enhancements to every component in end-to-end connectivity, this document will focus on reviewing some of the main features in Core networking components. For additional details, refer to the latest 3GPP standards.

The picture below depicts the 3GPP 5G Core networking components connected in a Service-Based Interface (SBI) architecture.



3GPP TS 23.501 Rel15.3 Sec 4.2.3 : 5G System Architecture

Edge 5G Architecture view

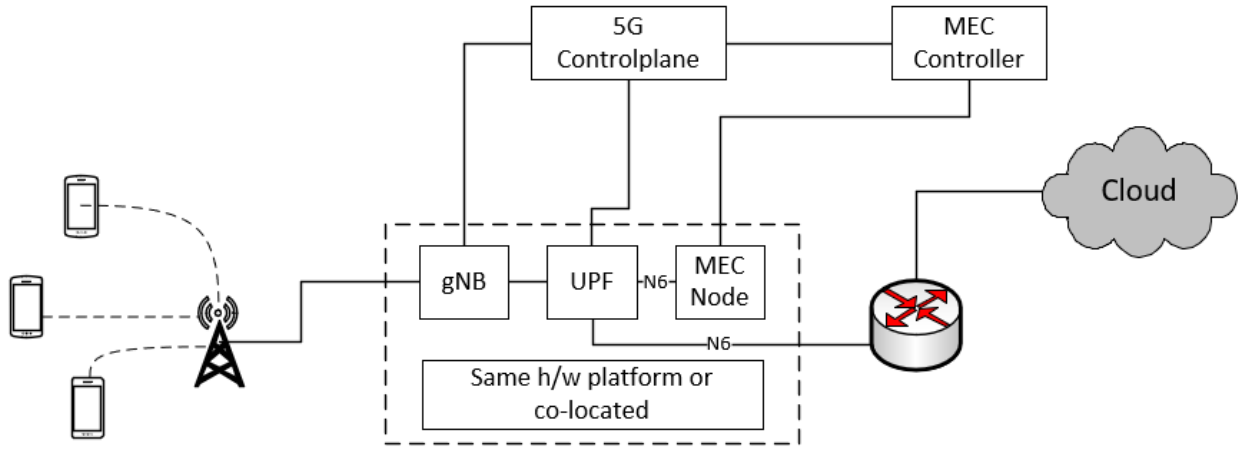
As discussed in the ETSI White paper “MEC in 5G Networks” ETSI_2018a, a new set of functional enablers introduced in 5G are essential for Edge deployments. The following is a list of highlights within the context of this document :

- Multiple Local Data Networks connected to the UPF and traffic steering of selected data traffic for a PDU session towards a local data network interface in UPF are key enablers of 5g edge deployments.
- Influencing the traffic steering rules in the UPF through external components like OpenNESS/MEC Controllers using an Application Function (AF) provides another level of flexibility for on-demand application deployments on edge nodes as described in 3GPP_23501 Release 15 Sec. 5.6.7.
- Session and Service Connectivity (SSC) and Local Area Data Networks (LADN) play an important role in edge deployments.

Edge deployment scenarios in 5G

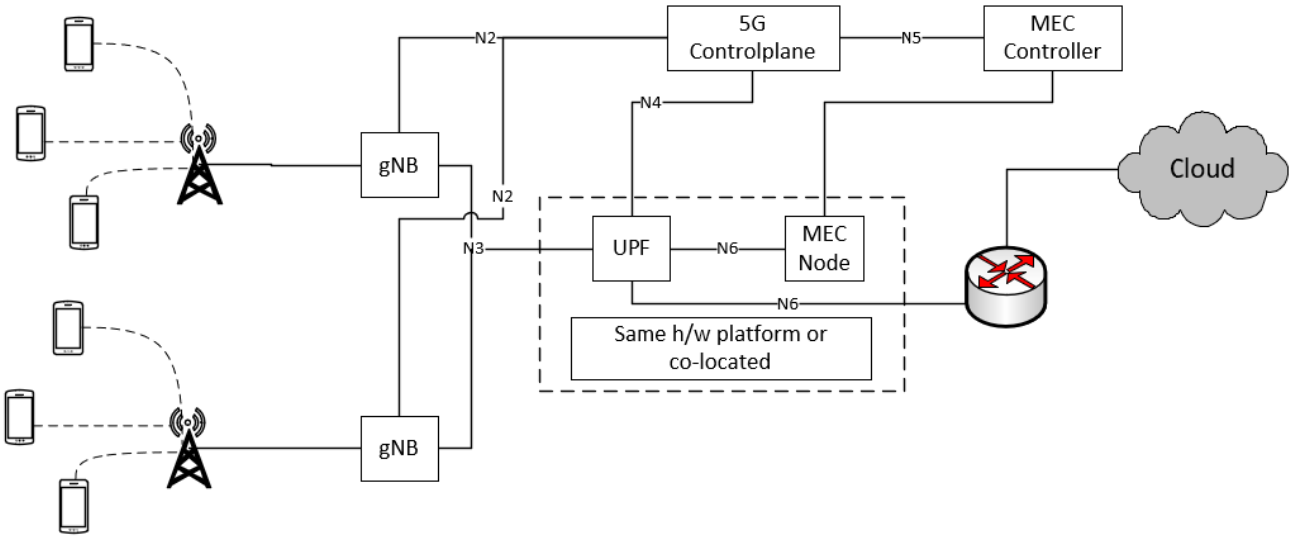
The following section outlines the various deployment scenarios in 5G:

1. The edge node hosts the edge applications and is co-located with the Base Station and the UPF:

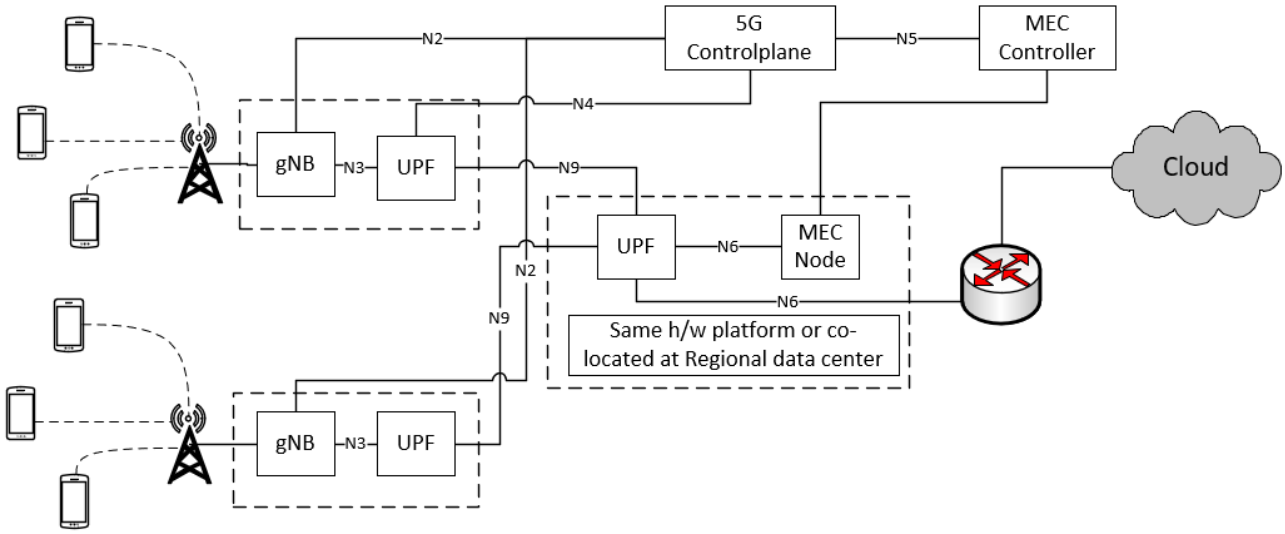


2.

The edge node hosts the edge applications and is co-located with the UPF:



3. The edge node hosts the edge applications and is co-located at Regional-Office:



There may be other edge deployment scenarios in addition to those highlighted above but they are not the focus of this document.

In all of the above scenarios, the UPF has a dedicated N6 interface associated with each edge node hosting multiple applications. In some cases, the UPF may have multiple logical N6 interfaces (one for each application) associated with the edge node.

Application Traffic steering in 5G towards Edge

Traffic steering in the context of edge refers to routing the UE (or end user) application's traffic towards applications deployed at the Edge location. The UPF provides the data plane functionality in 5G networks: 1) ***assignment of proper UPF to UE*** and 2) within the UPF ***identification of particular application traffic and routed towards a proper N6 interface to reach to an application deployed locally***. These are the two critical steps for successful edge deployments.

The 3GPP standard provides multiple references for implementing the selection of UPF for UE, providing flexibility in the implementation. Edge deployments will be more meaningful when the UPF and Edge node platforms (physical compute resources) are co-located and are deployed per location.

The standard also defines a set of procedures in the Application Function (AF) to influence the traffic routing in the UPF as well as the selection of UPF to reach UE traffic on the local data network [3GPP_23501 Release 15 Sec. 5.6.7].

OpenNESS integration with 5G systems

As a reference software solution kit, OpenNESS addresses some of the key challenges in the 5G edge deployment scenarios referred to above through the Application Function (AF) microservice and REST-based APIs. The rest of the sections will focus on explaining about the components in OpenNESS, how they interact with each other, supported REST-based APIs (to interact with the 5G NGC solution), and end-user APIs (to integrate with orchestration and/or solutions).

OpenNESS scope

In the context of 5G edge deployments, OpenNESS interacts with 5G NGC through the AF Network Function microservice as defined in the 3GPP standard SBI interface. Additionally, OpenNESS proposes a reference REST-based API endpoint (OAM interface) to configure the 5G Control Plane elements with the information about UPF deployed at the edge nodes. In most cases, 5G NGC solutions may have this configuration path implemented. In the scope of OpenNESS integration with the 5G Core, the OAM interface is a point of discussion based on the existing 5G core interface.

OpenNESS implementation

The key challenges for Edge deployments in 5G networks have been outlined in the section Introduction. OpenNESS tries to address them in compliance with the standards by:

1. UPF selection

- For deployment scenarios #1 and #2, where the serving UPF and edge node are co-located with the RAN, proper UPF selection for UE is critical. If the 5G Core considers UE location and requesting DNN (i.e. TAC, DNN, DNAI, SNSSAI, SSC) in UPF selection, it would make the Edge deployment more efficient. To enable this capability, OpenNESS proposes an OAM REST-based API interface to inform the 5G core about the UPF info (upf-ip, tac, dnn, dnai, snssai, dns-ip) co-located with edge node.

- In the case of an edge node deployed at regional centers (#3), the selection of the serving UPF is done by 5G Core (SMF). However, UE application traffic needs to be steered from the serving UPF to the UPF co-located at the edge node through an N9 interface. To achieve this, traffic influencing rules need to be pushed in both UPFs to identify proper N9 and N6 interfaces for the data traffic to reach applications deployed on the edge node.

1. Traffic steering

- The 5G standard exposes REST-based APIs defined through Network Exposure Function [3GPP TS 23.502-f30 Sec. 5.2.6] for the AF to configure
    - The traffic flow rules to identify the application traffic, ie. Packet Flow Descriptor (PFD)
    - Operations create/modify/delete and traffic influencing subscription APIs for steering application traffic towards edge node N6 interfaces and more. OpenNESS AF will support these APIs in multiple phases, starting with traffic influencing subscription APIs in OpenNESS Rel 19.12.

1. DNS service

- For UE traffic to reach applications deployed at the edge, the DNS plays a major role. Resolving the DNS entry for applications running on Edge is always a topic for discussion with multiple options available and the choice is always influenced by the required deployment scenario.

  Two immediate options are (1) Use the DNS server maintained by the Network operator (2) Use the DNS services provided by the OpenNESS edge node.

    1. DNS server provided by the network operator
        - Pros: Central DNS server for multiple edge nodes hosting applications, one-stop-shop.
        - Cons: Challenging to keep the DNS records database up to date with the dynamic nature of application deployment at edge nodes.
    2. DNS server provided by the OpenNESS edge node
        - Pros: Keeps DNS records up to date for applications deployed dynamically, supports DNS forwarding functionality for unresolved DNS queries. An ideal solution for single-edge node deployment scenarios.
        - Cons: If a single UPF connects to multiple edge nodes, then some sort of daisy-chaining activity of the DNS server configuration within edge nodes is needed. Assigning an Edge Node DNS server IP to UEs is also a challenge in some implementations. Through the OAM reference APIs, OpenNESS proposes a path to configure the edge DNS server (associated with UPF) to the 5G core but may not be able to address all scenarios.

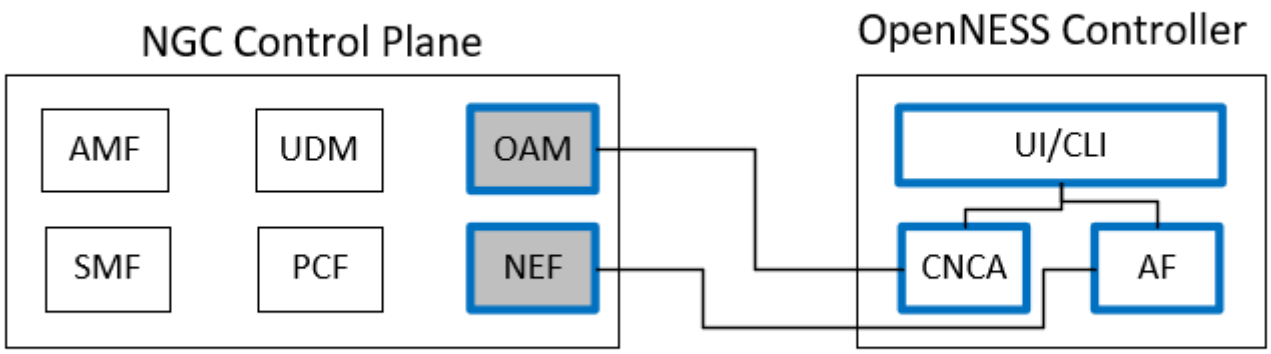
1. UE mobility

- How to support UE mobility in Edge scenarios is an important question for operators since 4G/LTE timelines. Thus, the 3GPP 5G standard has captured it during the functional requirements stage while defing the spec with enhanced features like Notification procedures, Session and Service Continuity (SSC) modes, etc.. to leave enough opopertunity for operators and Edge solution developers to achieve Edge Key Performance Indicators (KPIs) for end users during mobility. However, mobility in edge applications requires support in the end to end path, i.e.:
    - The 5G core has to notify the UE mobility events towards MEC platforms.
    - The MEC platforms should have the capability to register for notifications and act accordingly to re-configure the traffic influence subscription rules towards serving UPF, if applicable.
    - Application on the Edge node are capable of application context transfer from one edge node to another running similar application instances.
    - UE applications should also be aware of and honor the application context switch for an unintereputed service.

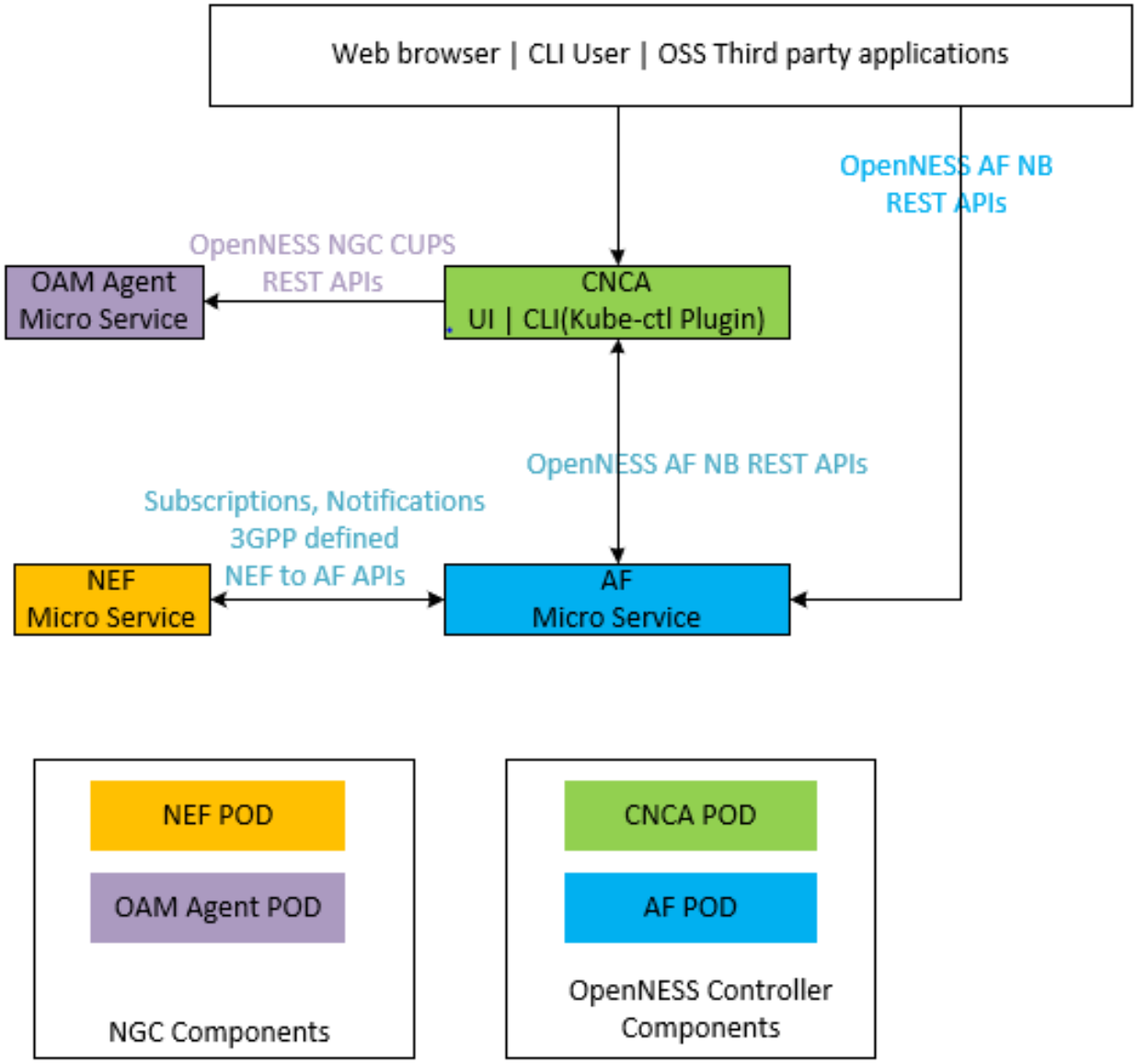
  Technical challenges aside, as Edge services are mostly location-based, the visualization of mobility might not be applicable in all edge deployment scenarios.

OpenNESS functional elements

Below is a list of functional elements provided through the OpenNESS solution to address the key 5G integration challenges.



The following pictures show the microservice architectural view of the OpenNESS solution with 5G integration components:



Application Function

An Application Function (AF) is a microservice in the OpenNESS edge controller solution, and it is developed in golang. AF supports the Traffic influencing subscription and Packet Flow Description Management functionality to help steer the Edge-specific traffic in UPF towards the applications deployed on the OpenNESS edge node.

Other AF functionalities as discussed in 3GPP 5G standard [3GPP_29122], Changing chargeable party Section 4.4.4, configuration QoS for AF sessions Section 4.4.13, Monitoring Section 4.4.2, Device triggering Section 4.4.6, and resource management of Background Data Transfer (BDT) Section 4.4.3 are under consideration for future OpenNESS releases.

The OpenNESS AF microservice provides a northbound (NB) REST-based API interface for other microservices that provide a user interface (i.e., CNCA/UI or CLI). Also, these NB APIs can be invoked from external services that provide infrastructure for automation and/or orchestration.

Traffic steering NB APIs

- API Endpoint: */af/v1/subscriptions*
- Supported methods: POST, PUT, PATCH, GET, DELETE
- Request/Response body: *5G AF North Bound APIs schema at openness.org*

AF supported Traffic steering API (South bound)

- API Endpoint: */3gpp-traffic-influence/v1/{afId}/subscriptions*
- Supported methods: POST, PUT, PATCH, GET, DELETE
- Request/Response body: *5G NEF North Bound APIs schema at openness.org*

PFD Management NB APIs

- API Endpoint: */af/v1/pfd/transactions*
- Supported methods: POST, PUT, PATCH, GET, DELETE
- Request/Response body: *5G AF North Bound APIs schema at openness.org*

AF supported PFD management API (South bound)

- API Endpoint: */3gpp-pfd-management/v1/{scsAsId}/transactions*
- Supported methods: POST, PUT, PATCH, GET, DELETE
- Request/Response body: *5G NEF North Bound APIs schema at openness.org*

NGC notifications

As part of the traffic subscription API exchange, SMF generated notifications related to DNAI change can be forwarded to AF through NEF. NEF Reference implementation has place holders to integrate with 5G Core control plane.

Network Exposure Function

According to 3GPP 5G System Architecture [3GPP_23501 Release 15], NEF is a functional component in the 5G Core network. However, the reason for including NEF as a microservice in the OpenNESS solution is two-fold.

- For validation of AF functionality in OpenNESS before integrating with the 5G Core. This could enable OpenNESS partners to validate their interfaces before integrating with their 5G Core partner. Hence, the NEF microservice scope in OpenNESS is limited and in line with the AF functional scope.
- It may be helpful for 5G Core partners who are looking for NEF service to be added in their solution for OpenNESS integration.

The OpenNESS provided NEF reference implementation for Traffic influence and PFD management is as

per 3GPP TS 23.502 Section 5.2.6. Supported API endpoints are Nnef_TrafficInfluence {CREATE,UPDATE,DELETE} and Nnef_PfdManagement {CREATE, UPDATE, DELETE}.

OAM Interface

OAM agent functionality is another component that should be part of the 5G Core solution to add/update certain configuration information outside the scope of standards. For example, the Configuration of UPF parameters such as UPF IP address, DNS configuration, and DNNs are supported. In the case of edge deployments, when the UPF is deployed as an NFV service on the edge node platform, MEC controllers may need to update the 5G Core control-plane components about the edge associated user-planes. 5G solutions may have some sort of interface to address this requirement. However, to provide a unified interface for integrated solutions, OpenNESS proposes REST-based OAM interface APIs to configure certain UPF related parameters. The use of the OAM agent is optional and can be replaced with an OAM interface of the 5G Core solution if any exists.

Edge service registration

- OpenNESS suggested OAM API endpoint: */ngcoam/v1/af/services*
- Supported methods: POST, GET, PUT, DELETE

NOTE1: Because the OAM agent is a component in 5G Core and may need to interact with multiple AF instances (i.e., the OpenNESS edge controllers), the above API endpoint may need enhancements to incorporate `afId` to distinguish between AFs. THE updated API endpoint could be */ngcoam/v2/af/{afId}/services*, which is an interest for OpenNESS in future enhancements.

NOTE2: Registration of AF instance (OpenNESS controller) with 5G Core network could also be a topic for discussion during implementation. Customers can choose the suggested 3GPP method using NRF functionality or the existing OAM functionality can be extended with an additional API (for example, */ngcoam/v2/af/register*) to register and obtain `afId` from 5G Core.

Core Network Configuration Agent

Core Network Configuration Agent (CNCA) is a microservice that provides an interface for end users of the OpenNESS controller to interact with the 5G Core network solution. CNCA provides a CLI (kube-ctl plugin) interface to interact with the AF and OAM services.

Security between OpenNess 5GC micro-services

The security among OpenNESS 5GC microservices is supported through HTTPS and OAuth2.

HTTPS support

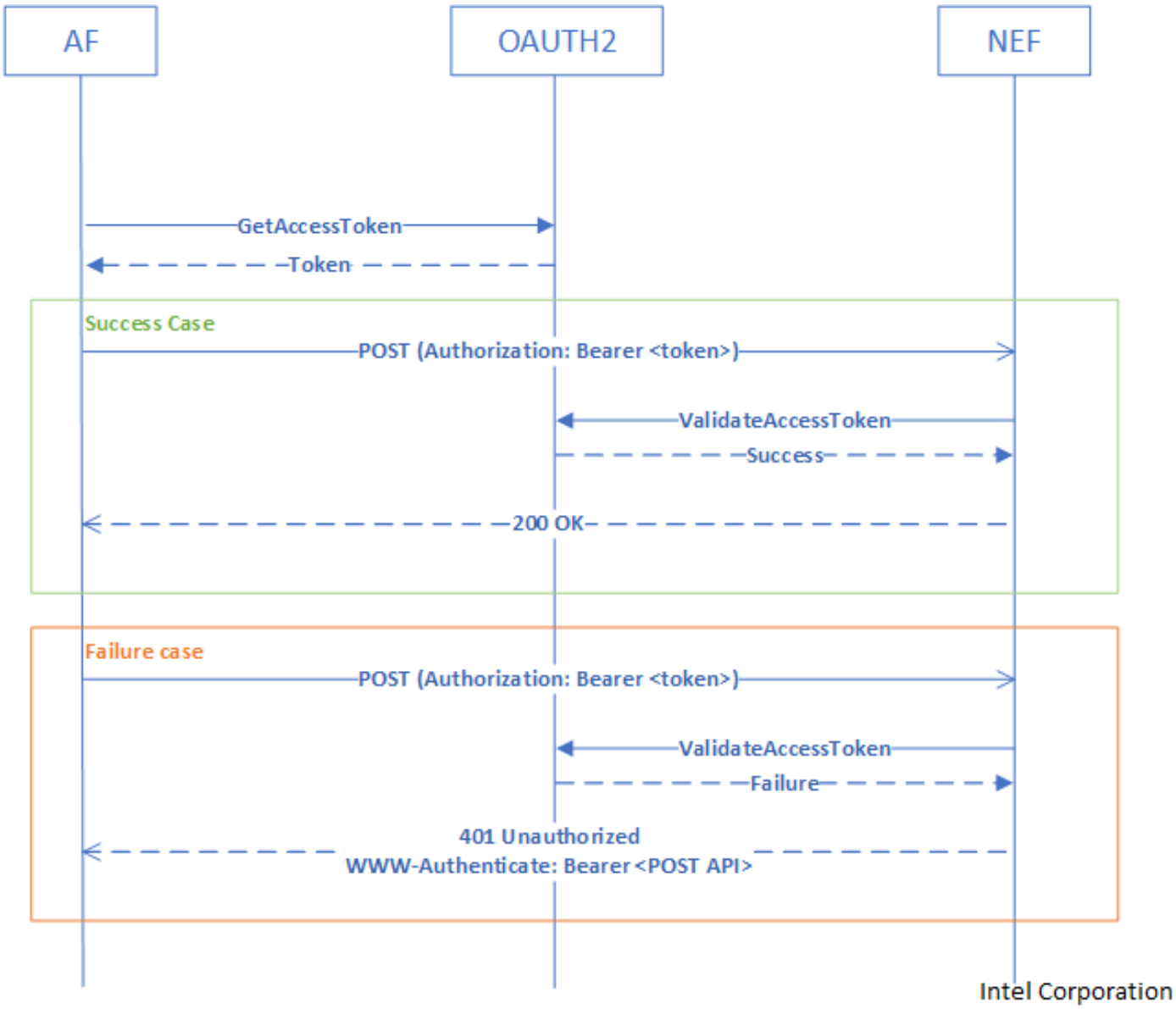
The OpenNESS 5GC microservices such as OAM, CNCA-UI, CLI kube-ctl, AF, and NEF communicate using REST API’s over the HTTPS interface among them.

OAuth2 Support

The AF and NEF microservices support the OAuth2 with grant type as “client_credentials” over an HTTPS interface. This is per the subclause 13.4.1 of 3GPP TS 33.501 (also refer 3GPP 29.122, 3GPP 29.500, and 3GPP 29.510). A reference OAuth2 library that generates the OAuth2 token and validates it is provided.

NOTE: When using 5GC core from any vendor, the OAuth2 library needs to be implemented as described by the vendor.

The OAuth2 flow between AF and NEF is as shown in below diagram.



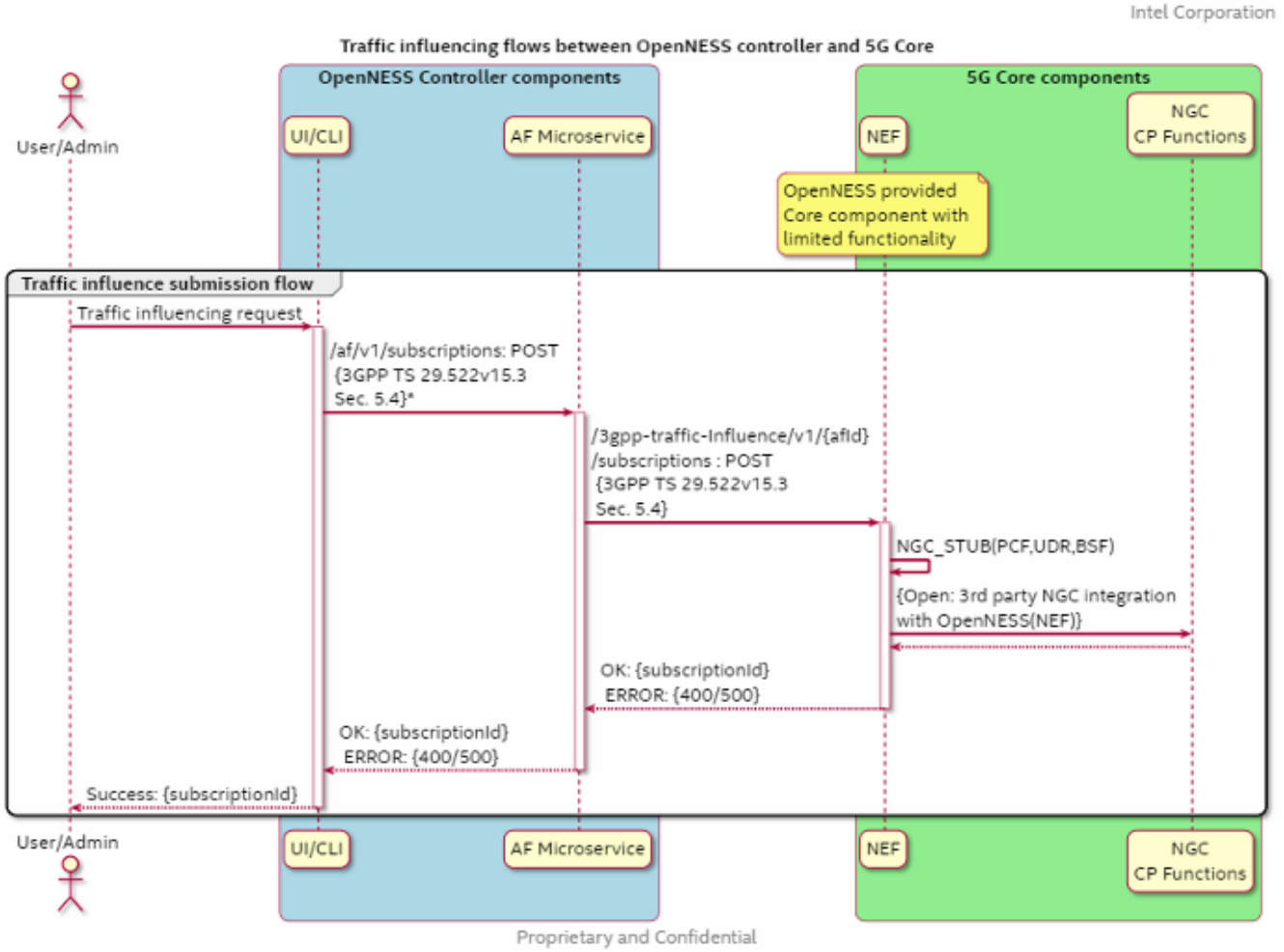
REST based API flows

The flow diagrams below depict the scenarios for the traffic influence subscription operations from an end user of OpenNESS controller towards 5G core.

AF-NEF interface for traffic influence

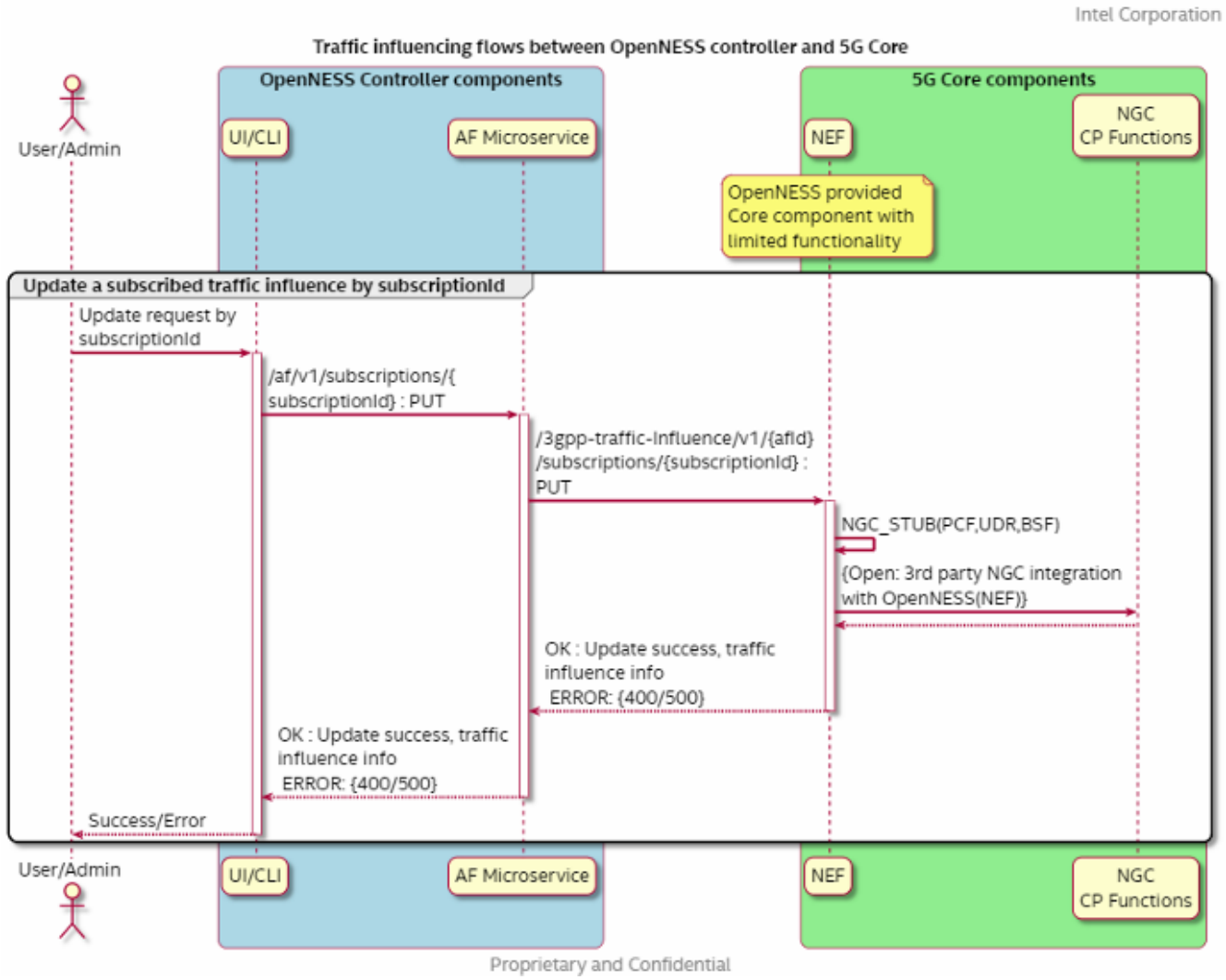
-

Addition of traffic influencing rules subscription through AF



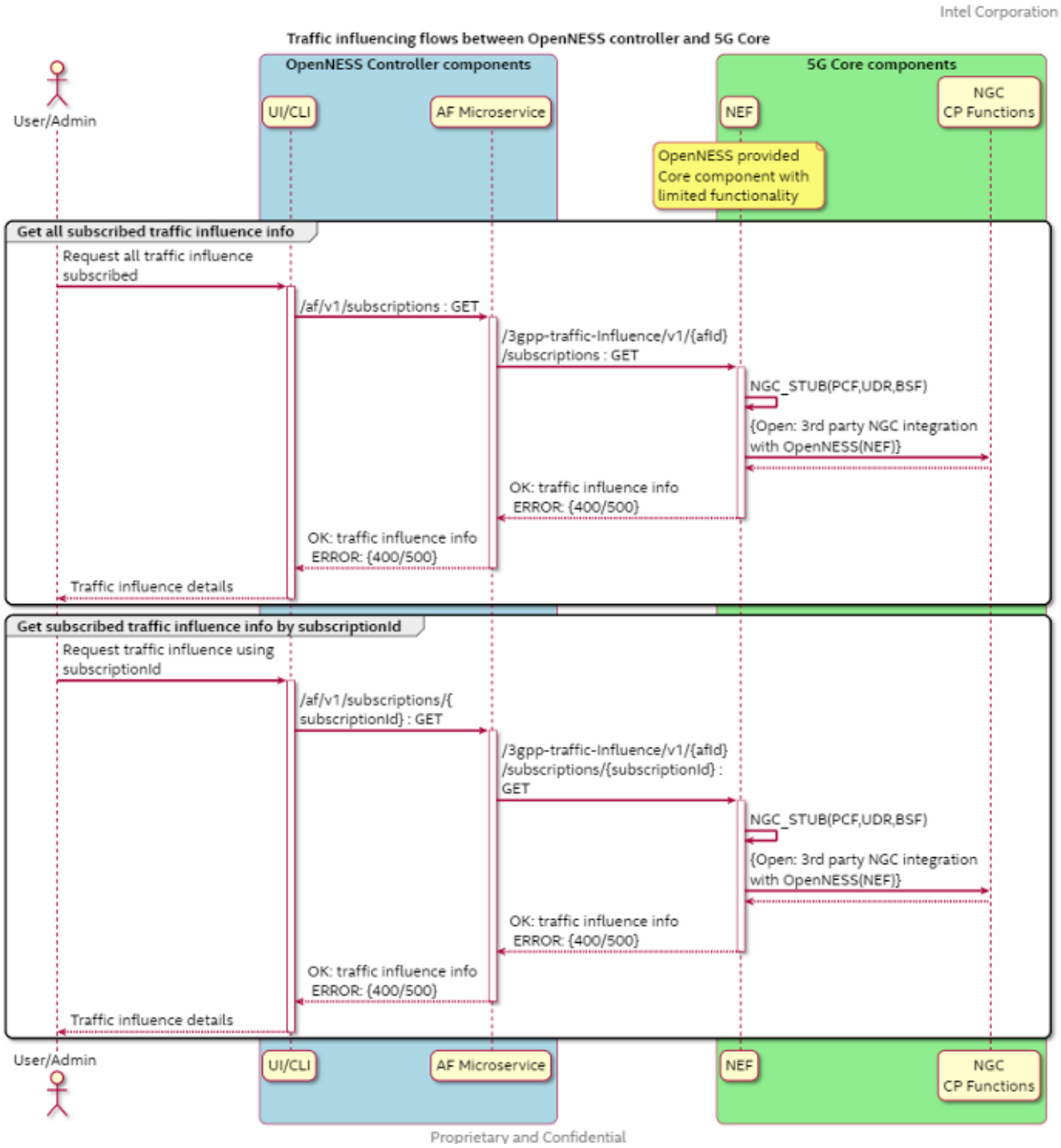
-

Update of traffic influencing rules subscription through AF



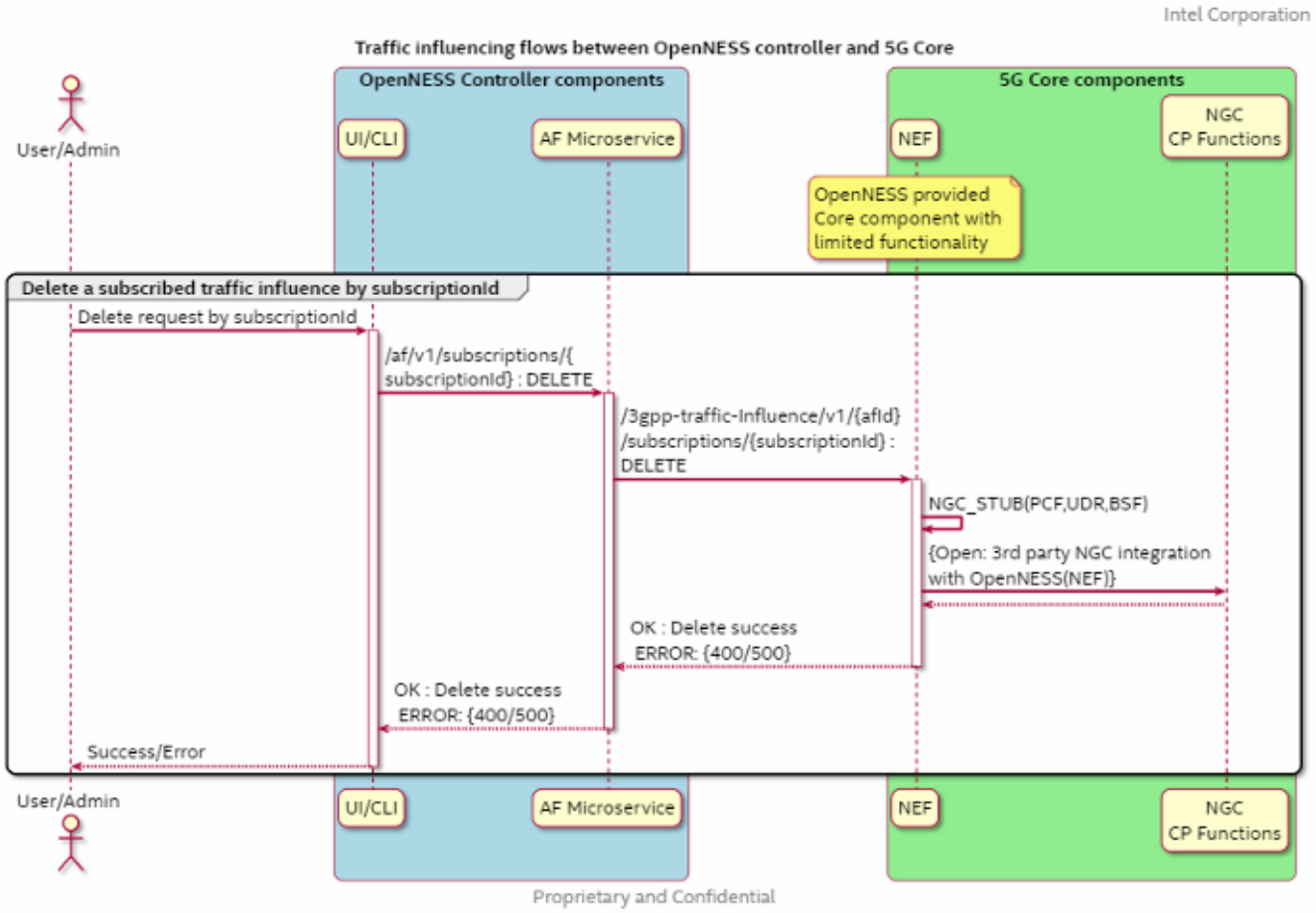
-

Get traffic influencing rules subscription through AF



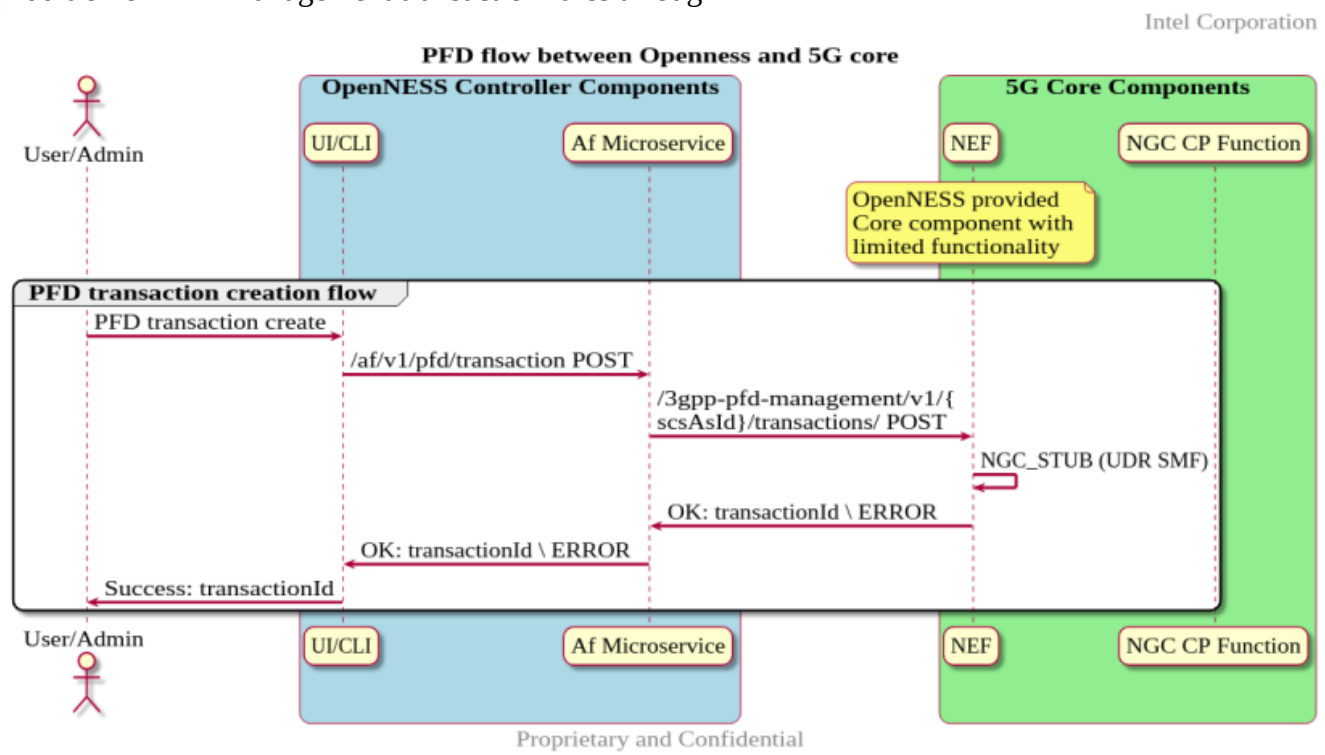
-

Deletion of traffic influencing rules subscription through AF

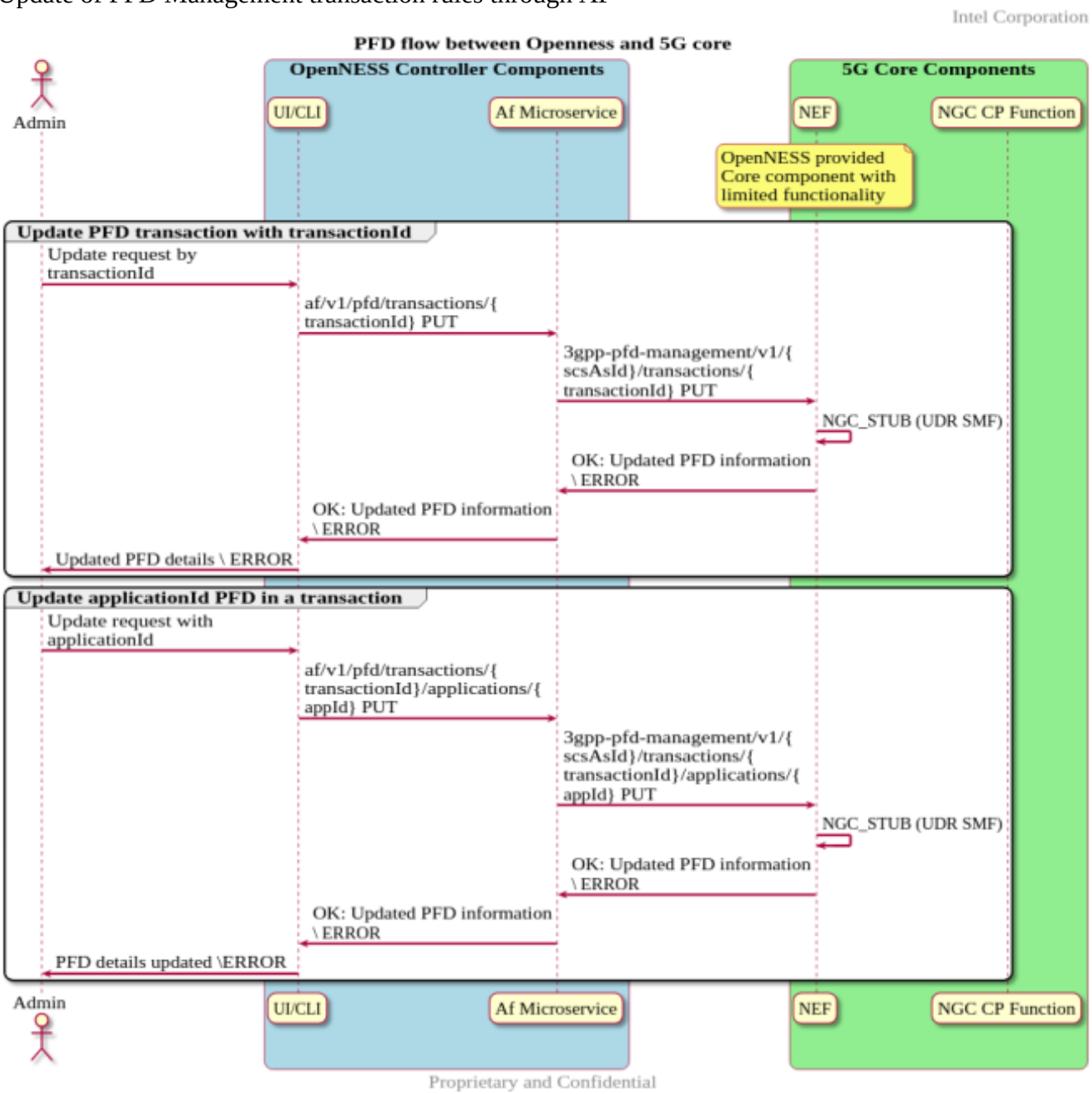


AF-NEF interface for PFD Management

- Addition of PFD Management transaction rules through AF

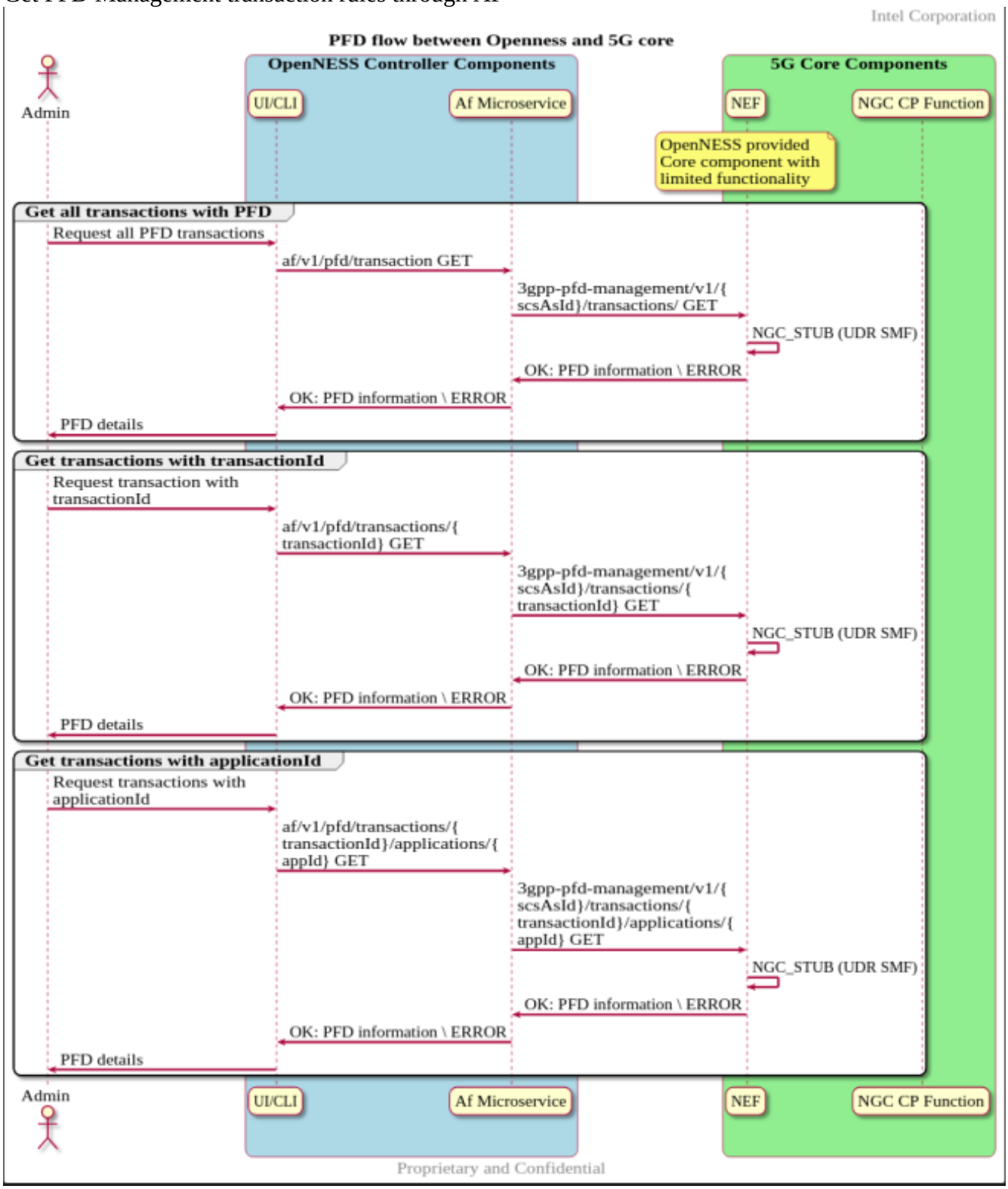


- Update of PFD Management transaction rules through AF



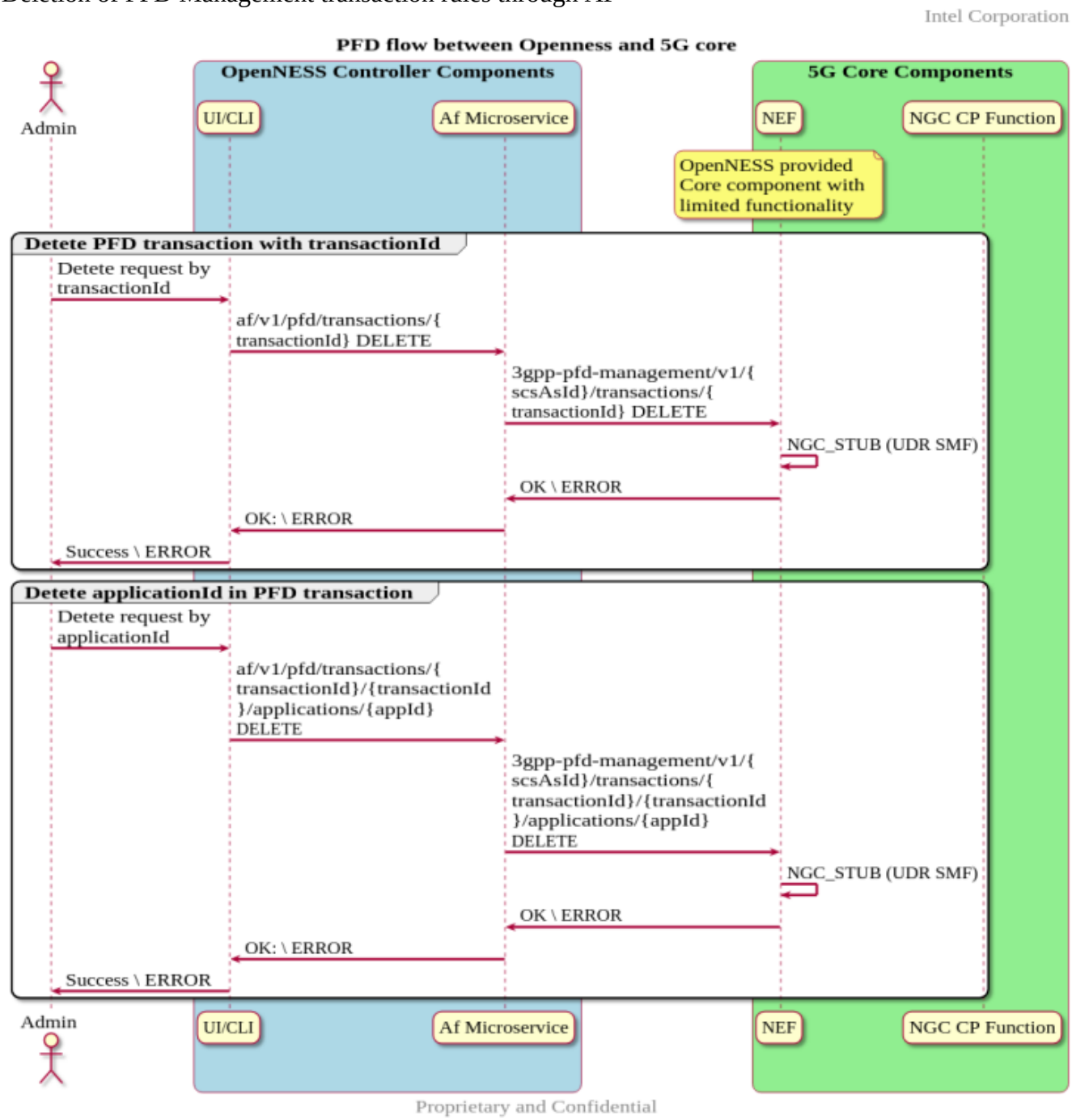
-

Get PFD Management transaction rules through AF



-

Deletion of PFD Management transaction rules through AF



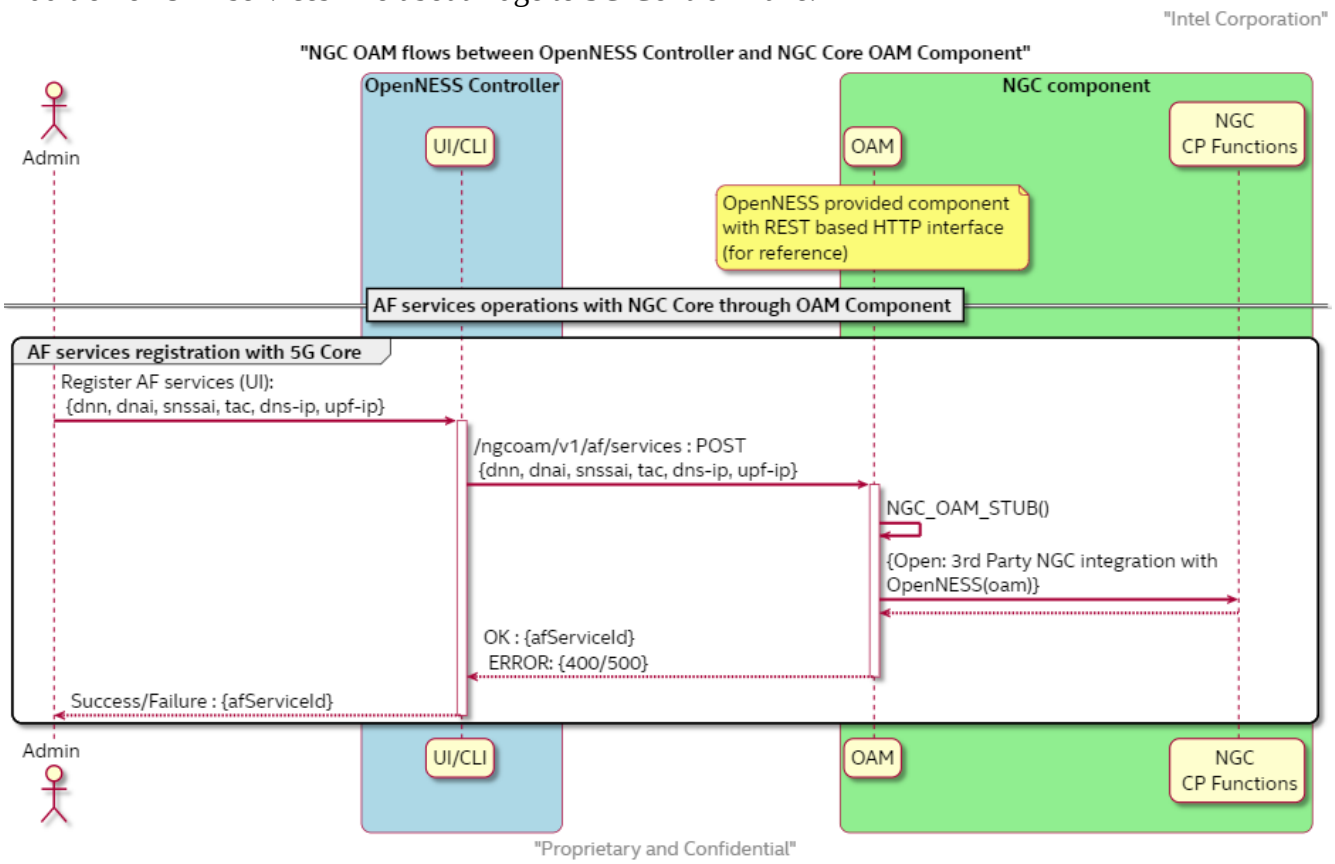
OAM interface for edge service registration

OAM API flows

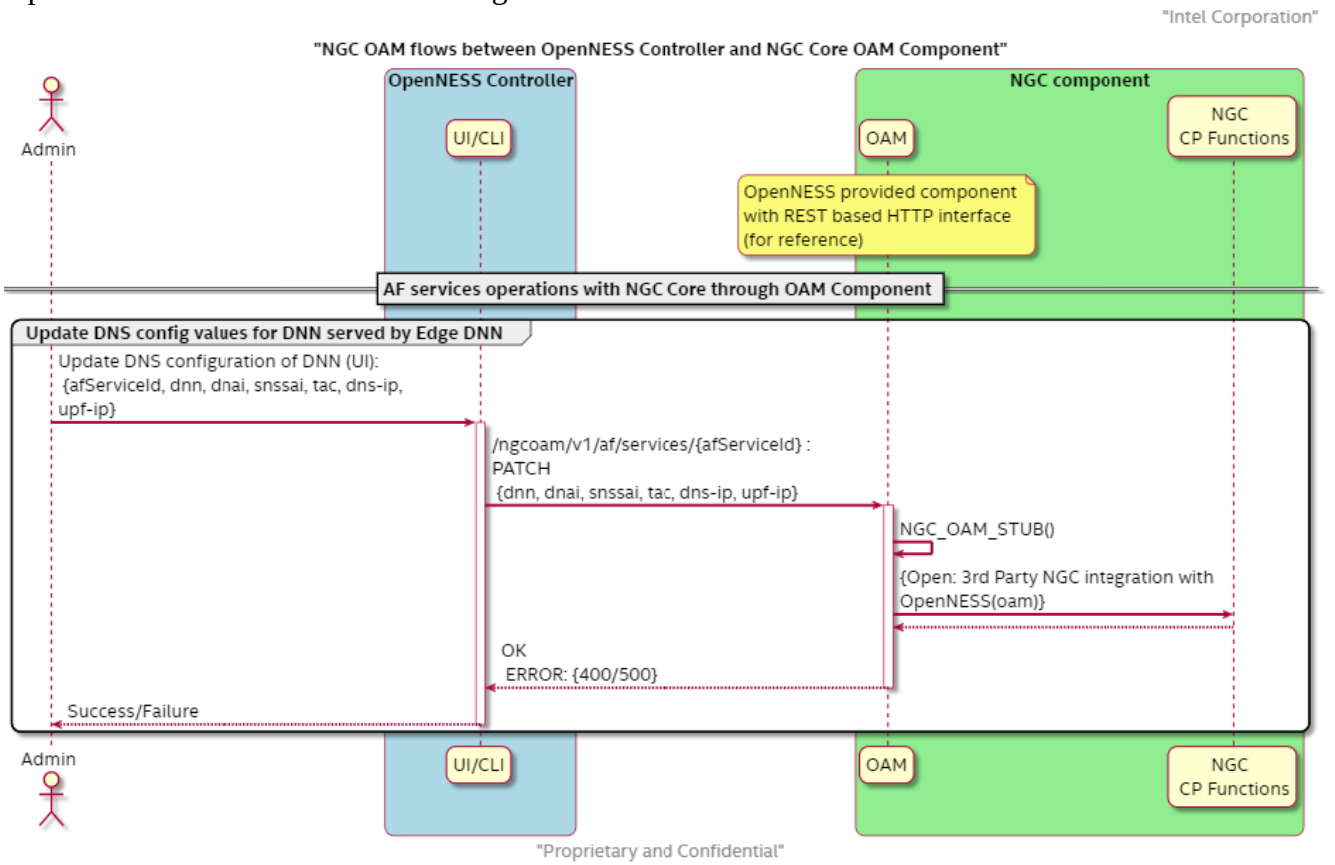
As discussed above, the need for configuring the 5G Control plane components with the information about UPF DNN information related to the edge. The flow diagrams below depict the API flow between various components to passdown the information towards 5G control plane. Detailed information about the OAM reference API endpoints can be found at 5G OAM API Schema in the documentation page at OpenNESS.org.

-

Addition of UPF services info about Edge to 5G Control Plane:

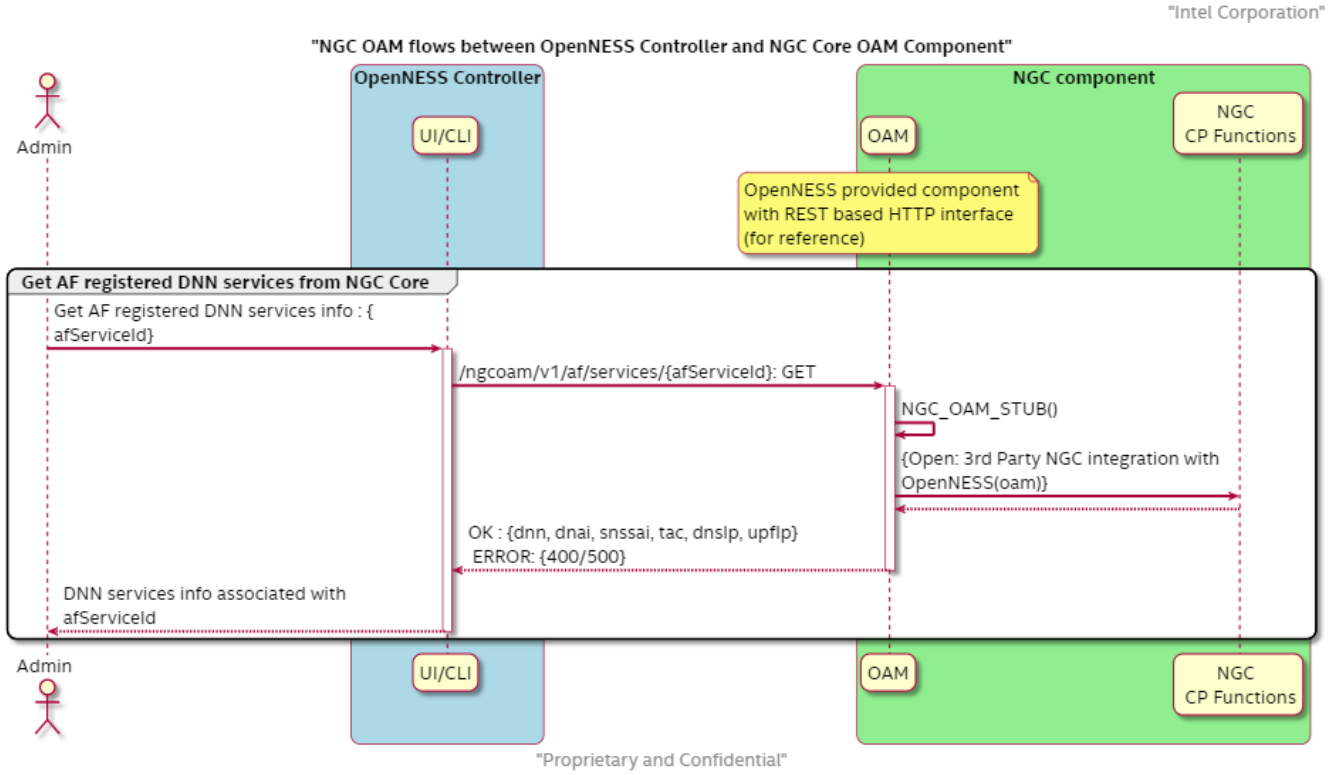


- Update of UPF services info about Edge to 5G Control Plane:

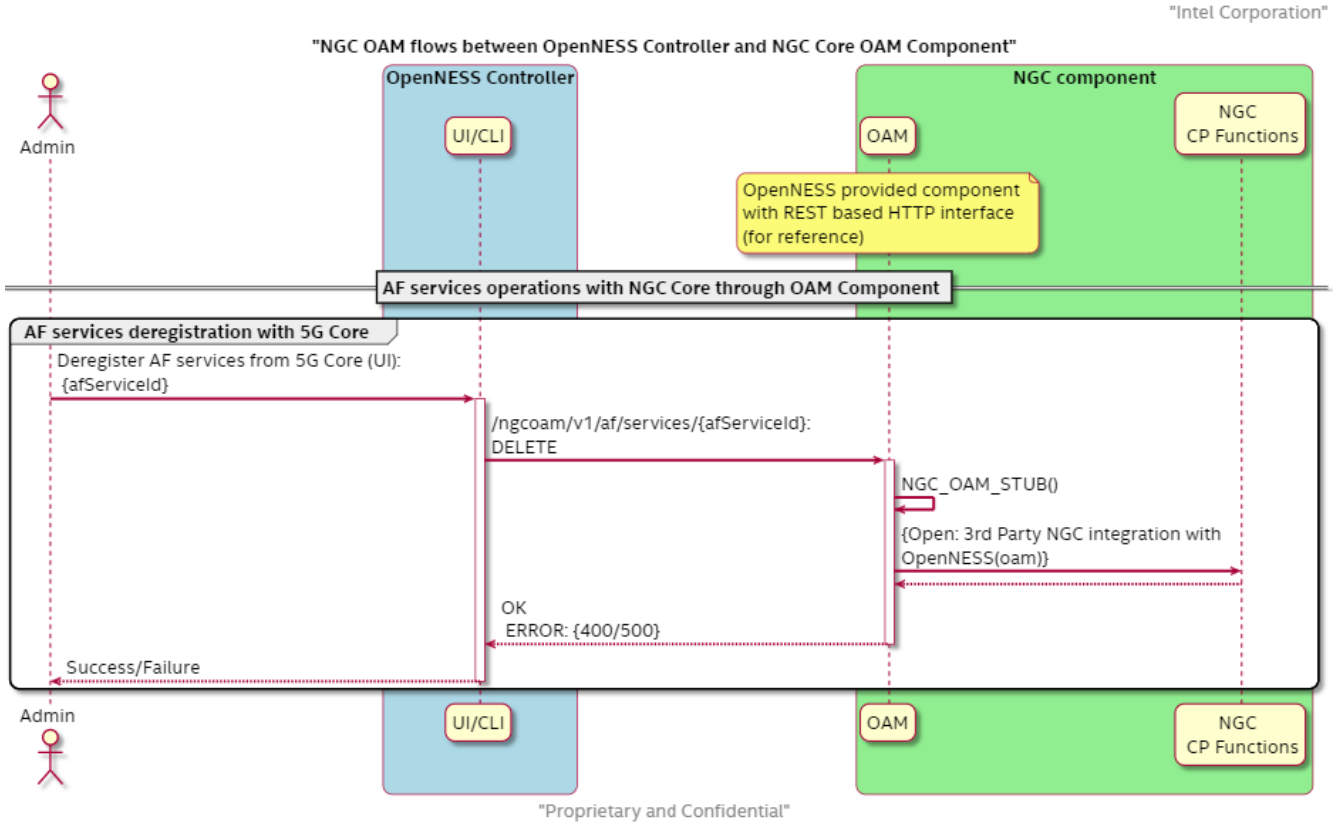


-

Get/Read UPF services info about Edge from 5G Control Plane:

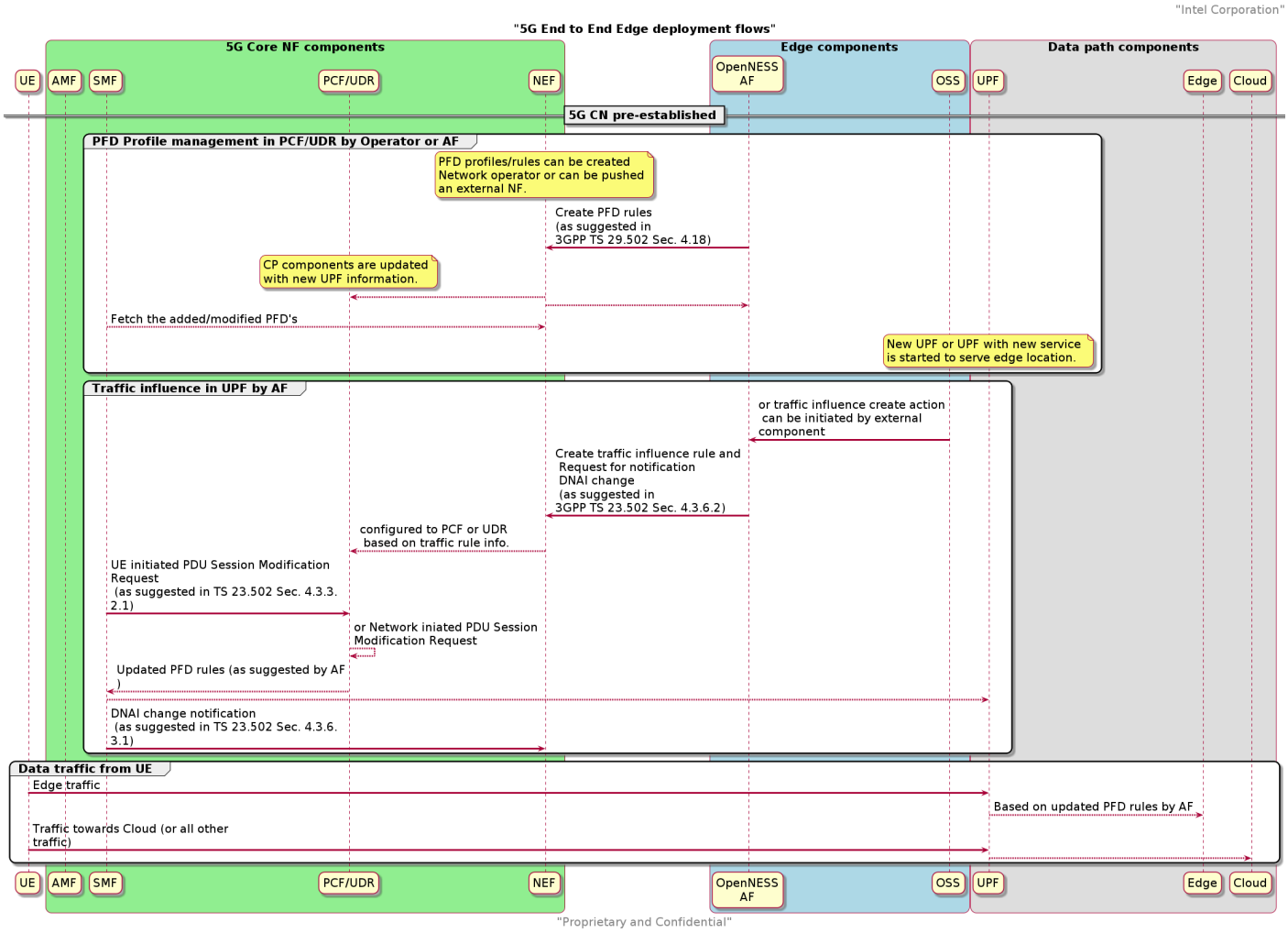


- Delete UPF services info about Edge from 5G Control Plane:



5G End to End flows for Edge by OpenNESS

The flow diagrams below depict a possible end-to-end edge deployment scenario including the PFD management, traffic influencing, and traffic routing in UPF towards Local DN.



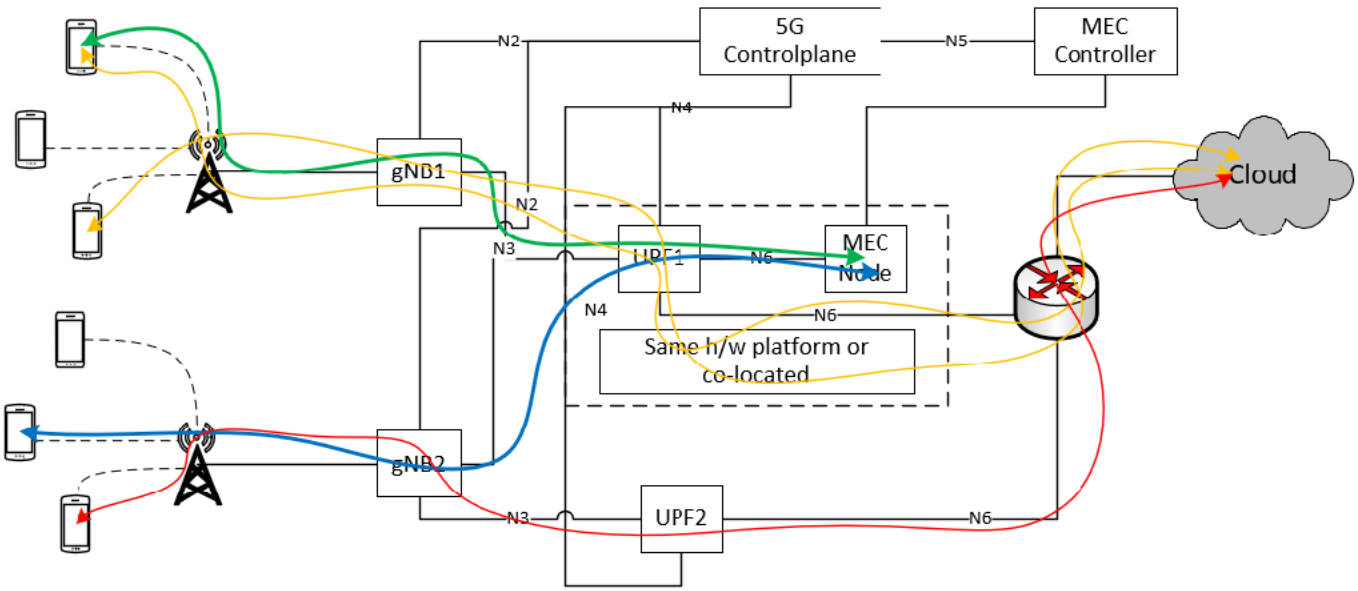
- AF authenticates and registers with the 5G Core

- PFD profile management
  - PFD profiles can be created based on the trigger in PCF/UDR by AF or the 5G Network operator can create PFD profiles.
  - SMF on getting notification from NEF on PFD's addition/modification, pulls the PFD's from NEF
- Traffic influence in UPF by AF
  - Traffic influence requests can be sent by AF towards the PCF (via NEF) for PFD profiles created in PCF. The action of a traffic influence request created in AF can be triggered by an external applications like OSS or a DNAI change notification events from the NEF or Device triggering events from NEF. AF registers for DNAI change notifications through the TrafficInfluence request API
  - Traffic influence requests will be consumed by PCF or UDR based on the requested information.
  - UE may initiate the PDU Session Modification procedure towards SMF, because of the location change event. Or the PCF may initiate a Network initiated PDU Session Modification request procedure towards SMF because of a traffic influence request generated by AF.
  - SMF may push this updated PFD profiles to UPF
  - When a new UPF is deployed in the 5G network or a new DN service is started on an existing UPF, SMF may generate a trigger to AF about the DNAI change notification.

- Data path from UE
    - Edge traffic sent by the UE reaches the UPF, the UPF routes the edge-traffic towards the local DN where the OpenNESS Edge Node is configured.
    - All other traffic sent by the UE that reaches the UPF will be sent to another UPF or to a remote gateway.

5G Edge Data paths supported by OpenNESS

The below picture shows multiple data paths that are supported in OpenNESS integrated edge deployment scenarios.



Each data path/scenario is represented by a colored line, which is described below.

All the UEs attached from Base Station gNB1 are assigned to UPF1 based on location (TAC) configuration. UEs that are attached from gNB2 are assigned to UPF1 or UPF2 based on TAC, DNN configuration while assigning UPF.

Green Colored data-path : UE application traffic that reaches UPF1 is routed to OpenNESS edge node through N6 interface. Traffic is served by applications deployed at the Edge node.

Orange colored data-path : UE application traffic that reaches UPF1, but non-edge traffic will be routed towards cloud through another N6 interface.

Blue colored data-path : UE attached to a Base Station (gNB2), but assigned to UPF1. UE application traffic that reaches UPF1 is routed towards edge node through N6 interface.

Red colored data-path : UE attached to a gNB2, but assigned to UPF2. UE application traffic reaches UPF2 is routed towards cloud through it's N6 interface.

NOTE All the above mentioned data paths also applicable to other two deployment scenarios described in the section Edge deployment scenarios in 5G.

5G Core Network functionality for OpenNESS integration

The following is the minimum functionality required to support the integration of the OpenNESS integrated MEC solution with a 5G Core network based on the required deployment scenarios:

- Control and Configuration in UPF selection functionality. Required in the case of location-based edge node deployments with associated 5G user-plane.
- Interface for dynamic configuration of a user-plane deployed or deleted from edge nodes. Referred to as the OAM interface in this document.
- Network Exposure Function (NEF) with minimal functionality to support Traffic influence subscription operations through the Application Function (AF).

Summary

This white paper highlights the Edge computing enhancements made in the 3GPP 5G standards along with the key implementation challenges. This document provides a description of the OpenNESS view of integration for Edge controllers with the 5G Core Network as well as the API endpoints and end-to-end flows required in edge deployments. The OpenNESS reference implementation was validated with a modified 5G Core in SA mode to support the APIs for edge deployments (per the 3GPP 5G Standard) and the 5G UPF on the edge node with multiple N6 interfaces connected towards local DNN interface.

Along with discussing the supported features in OpenNESS for 5G integration, the areas of interest for future enhancements are also outlined.

References

- [ETSI_MEC003]ETSI GS MEC 003 V1.1.1, “Mobile Edge Computing (MEC); Framework and Reference Architecture” (2016-03)
- [ETSI_2018]ETSI White Paper #24, “MEC Deployments in 4G and Evolution Towards 5G”, First Edition, February 2018, https://www.etsi.org/images/files/ETSIWhitePapers/etsi_wp24_MEC_deployment_in_4G_5G_FINAL.pdf .
- [ETSI_2018a] ETSI White Paper #28, “MEC in 5G Networks”, June 2018, https://www.etsi.org/images/files/ETSIWhitePapers/etsi_wp28_mec_in_5G_FINAL.pdf.
- [3GPP_29244]TS 29.244 Interface between the Control Plane and the User Plane of EPC Nodes.
- [3GPP_23501]3GPP TS 23.501 V15.3.0, “3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; System Architecture for the 5G System; Stage 2 (Release 15)”
- [3GPP_23502]3GPP TS 23.502 v15.3.0, “Procedures for 5G Systems; Stage2 (Release 15)”
- [3GPP_29122]3GPP TS 29.122 V15.3.0, “T8 reference point of Northbound APIs (Release 15)”
- [3GPP_29512]3GPP TS 29.512 V15.3.0, “5G System; Session Management Policy Control Service; Stage3; (Release 15)”

- [3GPP_29522]3GPP TS 29.522 V15.3.0, “5G System; Network Exposure Function Northbound APIs; Stage 3; (Release 15)”
- [3GPP_CUPS]”Control and User Plane Separation of EPC Nodes (CUPS)”, https://www.3gpp.org/cups
- [OpenNESS_2019]”OpenNESS Architecture and Solution”, white paper, 2019.

List of abbreviations

- 3GPP: Third Generation Partnership Project
- 5GC: 5G Core Network
- CUPS: Control and User Plane Separation of EPC Nodes
- NEF: Network Exposure Function
- AF: Application Function
- AMF: Access and Mobility Management Function
- CP: Control Plane
- DN: Data Network
- DNN: Data Network Name
- DNAI: DN Access Identifier
- LADN: Local Area Data Network
- APN: Access Point Name
- TAC: Tracking Area Code
- SUPI: Subscriber Permanent Identifier
- RAN: Radio Access Network
- S-NSSAI: Single Network Slice Selection Assistance Information
- SSC: Session and Service Continuity
- UL CL: Uplink Classifier
- UPF: User Plane Function
- UDM : Unified Data Management
- NRF: Network Repository Function
- PFD: Packet Flow Description
- UE: User Equipment (in the context of LTE)
- MCC: Mobile Country Code
- MME: Mobility Management Entity
- MNC: Mobile Network Code
- API: Application Programming Interface
- ETSI: European Telecommunications Standards Institute
- FQDN: Fully Qualified Domain Name
- HTTP: Hyper Text Transfer Protocol
- JSON: JavaScript Object Notation
- MEC: Multi-Access Edge Computing
- OpenNESS: Open Network Edge Services Software
- LTE: Long-Term Evolution
- OAM: Operations, Administration, and Maintenance
- PDN: Packet Data Network
- DNS: Domain Name Service
- REST: Representational State Transfer
- CNCA: Core Network Configuration Agent
- UI: User Interface
- CLI: Command Line Interface