HCLTech | Supercharging Progress™

5G Network APIs: An ERS whitepaper

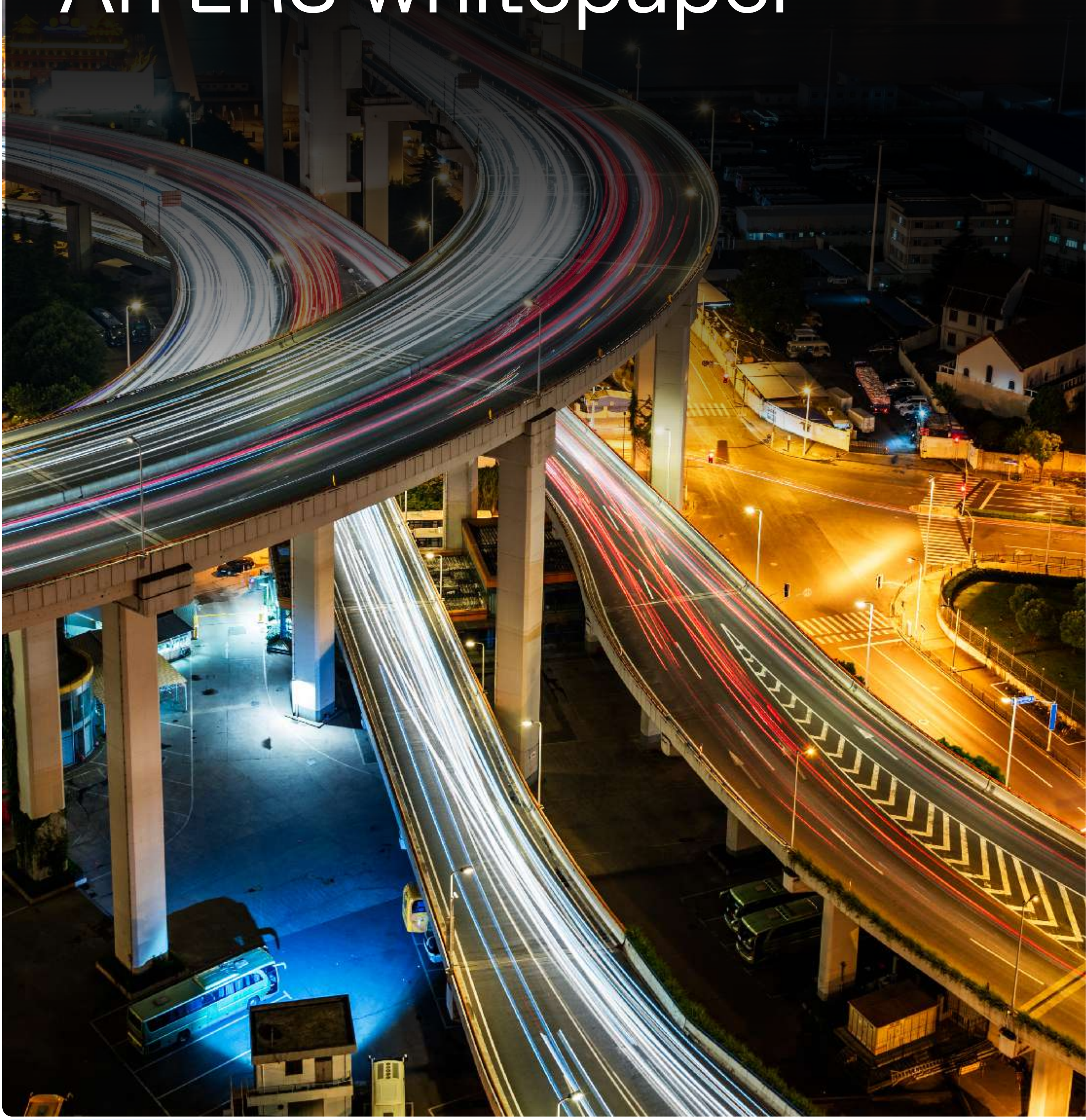


Table of contents

1. Abbreviations

AF	Application Function
AMF	Access and Mobility Management Function
AOA	Angle of Arrival
API	Application Program Interfaces
AS	Application Server
CSP	Communication Service Provider
DN	Data Network
ECID	Enhanced Cell ID
EDGE	Enhanced Data for Global Evolution
EES	Edge Enabler Server
GNSS	Global Navigation Satellite System
HTTP	Hypertext Transfer Protocol
IMEI	International Mobile Equipment Identity
IOT	Internet of Things
LCS	Location Services
LMF	Location Management Function
MEC	Multi access Edge Computing
MICO	Mobile Initiated Connection Only
MSISDN	Mobile Station Integrated Services Digital Network
NEF	Network Exposure Function
NF	Network Function
NSSAI	Network Slice Selection Assistance Information
O&M	Operation & Maintenance

OTT	Over The Top
PCC	Policy and Charging Control
PCF	Policy and Control Function
PFCP	Packet Forwarding Control Function
PFD	Packet Flaw Descriptor
QoS	Quality of Service
SBA	Service Based Architecture
TDOA	Time Difference of Arrival
UDM	Unified Data Management
UDR	Unified Data Repository
UE	User Equipment
UPF	User Plane Function
URLLC	Radio Access Network

2. Introduction

5G is creating opportunities for enterprise businesses to benefit from wireless communication over fixed local area networks, enabling communication service providers (CSPs) and enterprises to progress together. By exposing the 5G network APIs in a secured and flexible manner, the enterprise application logic can access, request, update and leverage the network analytics information to solve business challenges and facilitate digital transformation. Once enterprise business applications can query 5G network information like user status, traffic, quality of service (QoS), bandwidth, losses and other network statistics, numerous use cases will emerge.

To enable the ecosystem to define APIs and leverage 5G network-as-a-platform rather than just a connectivity pipe, a disaggregated and distributed architecture of 5G is required. The 3rd Generation Partnership Project (3GPP) has introduced the Network Exposure Function (NEF) to enable operators and enterprises to offer APIs that can be used to develop applications. Furthermore, Open-Radio Access Network(O-RAN) alliance has introduced the concept of developing applications in the RAN domain to deploy and leverage 5G network-as-a-platform.

This whitepaper presents use cases that can be built using 5G network APIs.

3. Need for network API

Network API is a secure method that enables communication between two network functions and facilitates the exchange of information. When one network function sends a request, the other function responds or acts accordingly. Many enterprise business applications require access to 5G networks to query and leverage them in their business logic.

However, in the absence of the network APIs being exposed by the CSPs, enterprises face the following challenges:

- Needing Operation and Maintenance (O&M) access to CSP network, but providing dedicated O&M terminals for hundreds of enterprise customers requesting 5G network data from a specific CSP is impractical.
- Raising service requests to CSP office for each of their needs, leading to delays in raising multiple requests towards CSP O&M staff and manually tracking them for closure.

To avoid manual intervention for each request and have real-time processing of enterprise demands, the industry required a solution that enhanced enterprise users' service experience. Enterprise applications should be able to query required real-time data processing for business and trigger 'allowed changes' on the enterprise's dedicated network. This is where 5G network APIs play a crucial role, facilitating the exposure of 5G network data in real-time.

The diagram below gives a view of some of the possible services. A list of 5G network data services is highlighted inside the 5G bubble, where services like location, subscription, billing, status are some of the examples that require real-time processing. On the other hand, hroughput, quality and latency require good optimization of network resources. CSPs monitor these data in real-time for ensuring service delivery as per the service level agreement and billing.

By exposing these as a Network API in a secure and specialized way for enterprises, it will enable multiple business use cases for enterprises.

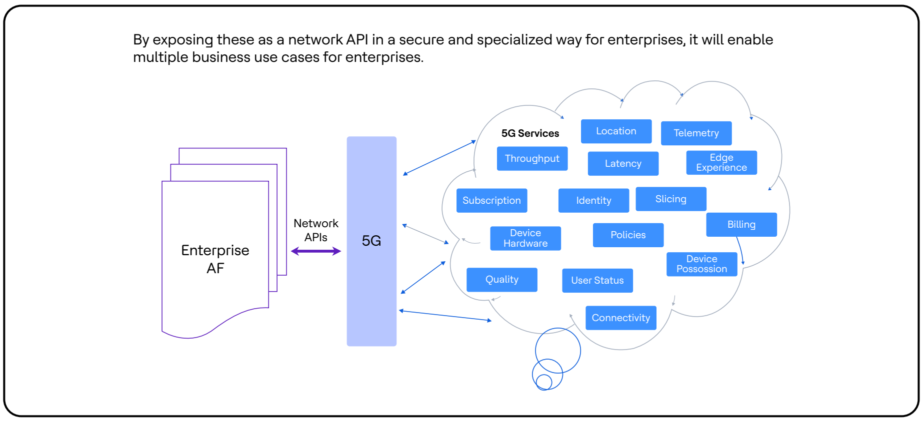


Figure 1: 5G services access by enterprise Application Function (AF)

4. Introduction to 5G network APIs

Network APIs are used for interactions between CSPs' network and the enterprise business' applications using APIs that are secured and exposed by the 5G network. Network APIs enable enterprises and operators to network as a platform, enabling them to build applications that can make changes in the network based on the business logic.

Network APIs are restful and use HTTP or HTTP2 as a method to obtain or alter user data. These are broadly categorized as:

- Control plane APIs
- User plane APIs
- Policy changes APIs
- Telemetry APIs
- Event exposure APIs

In a 5G network, enterprise application functions can have exposure of 5G network via either Network Exposure Function (NEF) or Policy and Control Function (PCF). Most of the CSPs prefer to provide access through NEF only. The figure below shows a network placement of these elements.

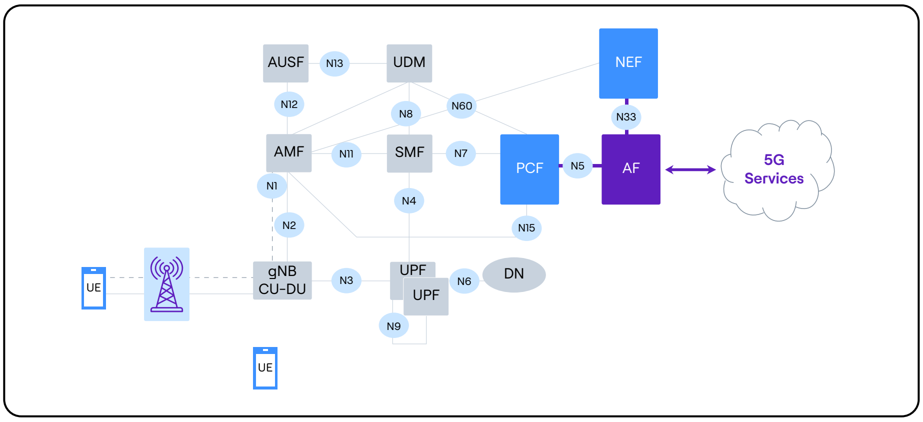


Figure 2: Network API access in 5G network

4.1 The role of NEF

The Network Exposure Function (NEF) plays a crucial role in enabling enterprise customers to access network capabilities and events managed by network functions. NEF's primary functions include:

- Facilitating secure information transmission from AFs to the network, which may involve authentication, authorization and regulation of traffic.
- Safeguarding the exposure of network capabilities and events from network functions to AFs.
- Converting information from AFs into a format that network functions can interpret and make information from network functions available to AFs.
- Implementing a Packet Flow Description (PFD) function that enables Afs to manage PFDs and interact with the Unified Data Repository (UDR) for storing and retrieving PFDs. NEF also manages the provisioning of PFDs to the Session Management Function (SMF).

The NEF instance may offer one or several of these capabilities and may support a selection of the APIs outlined for capability exposure. Enterprise AFs might utilize APIs for event exposure, traffic influence or location services. NEF is expected to host a variety of services to accommodate numerous potential northbound interactions, with some of the key services highlighted here.

- Nnef_EventExposure
- Nnef_TrafficInfluence
- Nnef_Location

Nnef_EventExposure service

Service operation functionalities applicable to application events include subscribe, unsubscribe and notify. AF acts as the service consumer for these operations. Whenever an event is detected, this service notifies AF accordingly. The service can observe events including service experience, user's mobility and communication events of a user, such as service start, termination and many others.

NEF can interact with the UDM, AMF and SMF, as well, by using event exposure functions like:

- Nudm_EventExposure
- Namf_EventExposure
- Nsmf_EventExposure

Nnef_TrafficInfluence service

This facilitates the UP-path management events, enabling AF to submit this service requests to NEF for influencing traffic, either for a single user or multiple users. This event can be leveraged to modify QoS or charging-related attributes. Supported service operations include create, update, delete and notify.

Nnef_Location service

It supports sharing the user's location with the application service function. Enterprise customers can utilize it for their own enterprise users served by a service provider. It supports notify operation.

4.2 Role of PCF

PCF is crucial in creating and deploying policies within the service provider's 5G network. It supports QoS and charging functions and is instrumental in implementing network slicing — a desirable solution in 5G — through proper utilization of policies. Enterprise customers can now also benefit from this service upon request.

Npcf_EventExposure service

This service is used by the application function to get exposure of events occurring inside the service provider's PCF about PDU sessions of enterprise users under access of specific PCF, PCC rules and policy control request triggers.

5. Network architecture

Figure 3 depicts a typical 5G network architecture, illustrating how enterprise customers maintain their own applications, developers and administration services. They prioritize deploying edge computing on-premises to fully capitalize on the 5G high-speed, high-throughput and low-latency services. Accessing 5G services requires implementing network APIs both at their premises and on the service providers' network.

Network APIs can interact with enterprise applications with the help of network functions like NEF or PCF as explained in Section 4. Utilizing NEF, communication service providers can expose 5G network services in a secure manner to enterprise customers.

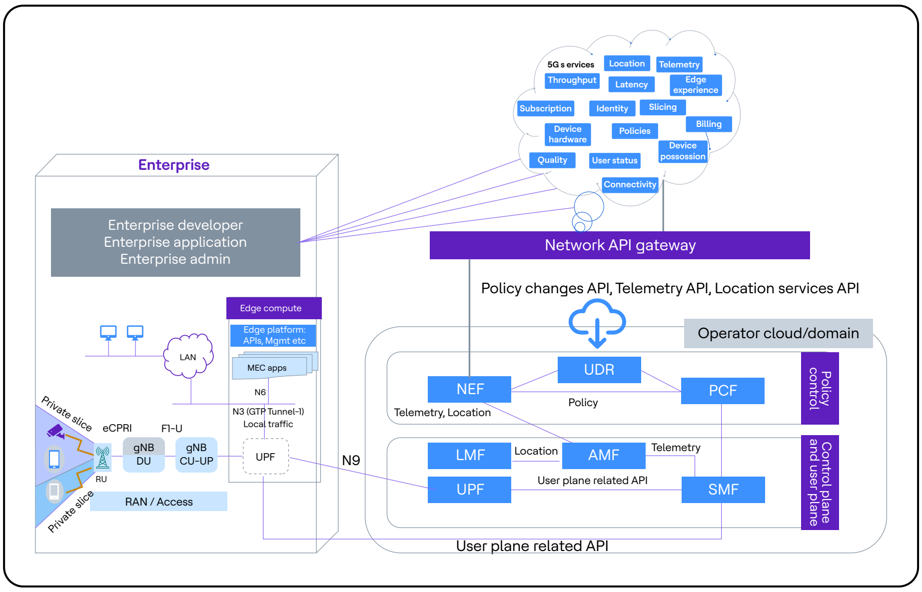


Figure 3: Network architecture for 5G services through network API

6. Use case illustration of network API

Enterprise businesses can leverage network APIs to transform their business functions, allowing elastic scaling as and when required and ensuring applicable quality to their users on demand. This section covers many use cases that operators and enterprises can develop today.

6.1 Traffic steering

In this use case, a central application server is supported by edge application servers situated across different regions. 5G users who utilize such applications require clients supported by application server, enabling access to the central application server to invoke services which encompasses any OTT or gaming application server.

According to the standard 5G practice, users are allocated an anchor at the edge of network upon session establishment. So, in order to enjoy a seamless OTT or gaming application, the user should get connected to nearest edge for high-rated experience supported by low latency and high throughput.

While service providers intend to deploy edge servers for application users at different edge locations, practical constraints may limit the availability of supported server at all edge locations. In the figure below we can see that various supported edge servers are placed at limited locations only. Hence, whenever service is needed, access to nearest server is required.

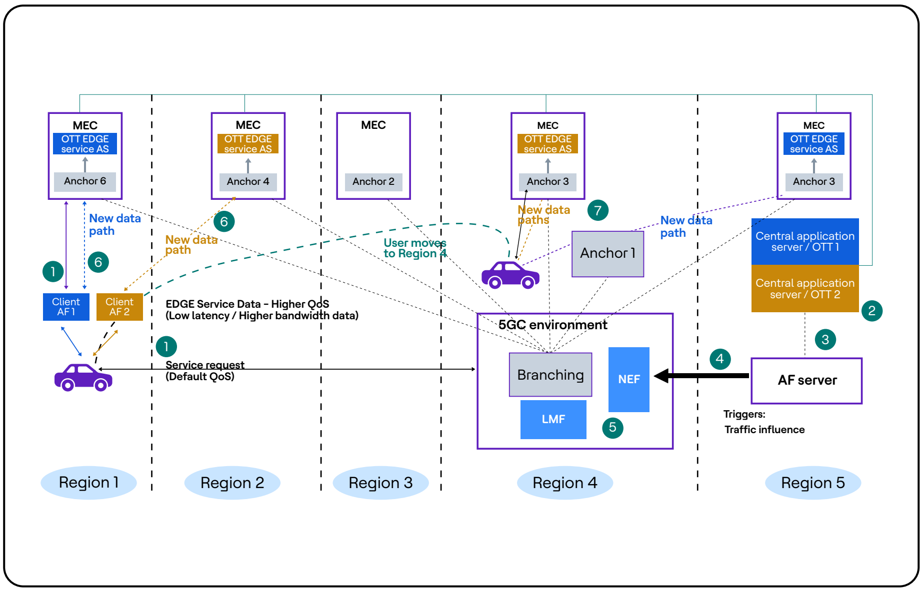


Figure 4: Traffic steering in 5G network through network API

The diagram in figure 4 shows numbers that can help in tracking the process flow.

- When a user submits a service request, it is sent to 5GC, which then allocates a default QoS. The request for allocation is then sent to the nearest MEC anchor 6 from region 1. This allows the user to transfer Data-to-Data Network (DN) through this anchor.
- When the user needs to access an OTT application, they log into the application installed by the client on their mobile handset. This notifies the central application server (2) that the user wants to access application data.
- The OTT application server checks user's subscription. If the user has subscribed to a higher quality service, then it shares desired QoS detail with AF server. AF server uses Network API to trigger traffic influence in 5GC.
- The AF server identifies the user through user identification parameters like IP address or MSISDN from OTT central application server and utilizes service operation traffic influence via N33 interface to notify about this user. AF is acquainted with user identification parameters like IP address or MSISDN from OTT central application server.
- NEF leverages 5G SBA architecture to notify the PCF about the required policy and use policy update procedure with SMF. The user plane reconfiguration operation is initiated using PFCP protocol and QoS required is set.
- In the example in figure 4, for OTT 1 edge server is in region 1 with anchor 6 so appropriate QoS can get set for this. But in case of OTT2, nearest edge server is available in region 2 with anchor 5 initiating a modification in user data path for the user to get connected with anchor 5. PDU session management is as per standard defined by 3GPP in TS 23.502.
- In case user moves into another region like in our example region 6, then a new user path needs to be established for low latency and high throughput. Based on location of user, decisions can be taken to have revised policies and traffic migration on anchor 3 and 5 respectively.

Figure 5 and 6 below shows message flow to achieve traffic steering in 5GC using network APIs.

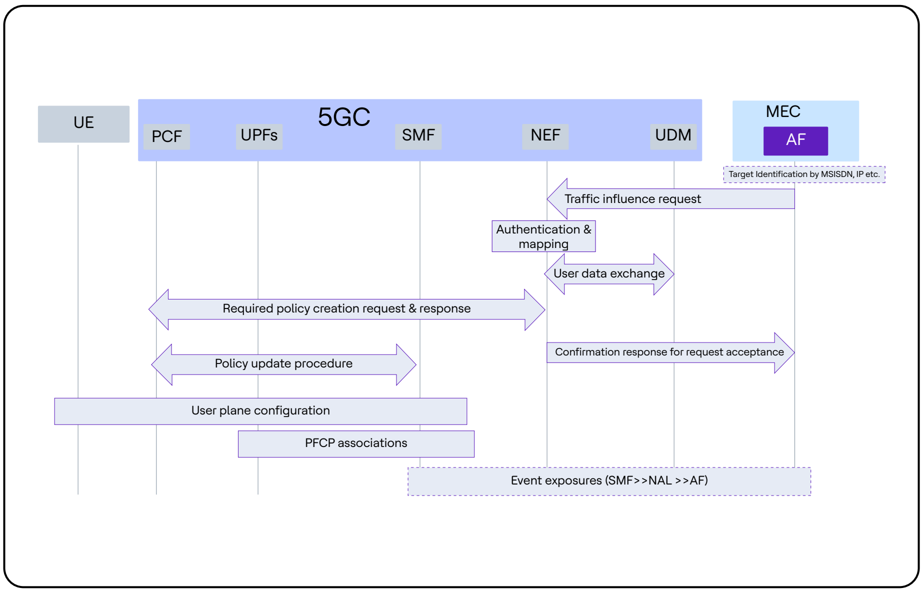


Figure 5: Conceptual message flow in 5GC for traffic steering through network API

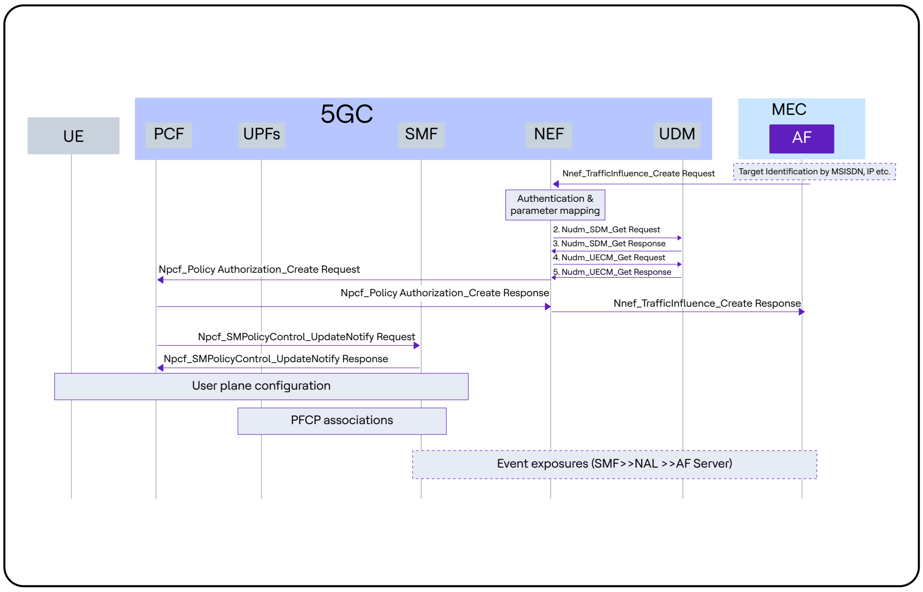


Figure 6: Signaling message flow in 5GC for traffic steering through network API

This reflects automatic service provision in 5GC from an external enterprise application. When user does not require OTT applications then default QoS can be set, freeing up some bandwidth and the same can be used to serve other desired user. This is effective real-time bandwidth utilization by enterprises without any intervention of communications services provider's personnel.

6.2 Location services on demand

In today's world of services, getting precise locations based on demand is critical. This capability can be utilized to provide emergency services, geofencing, location-based advertisements, location-based messaging and QoS. Using network APIs, it is possible to take service decisions based on locations outside of 5GC. Business logic gets implemented at the enterprise business level who have access to AF.

Figure 8 below depicts signaling message flow of such use cases where AF and NEF use Network API to trigger location identification request to User Equipment (UE). AF can further implement logic based on UE's location.

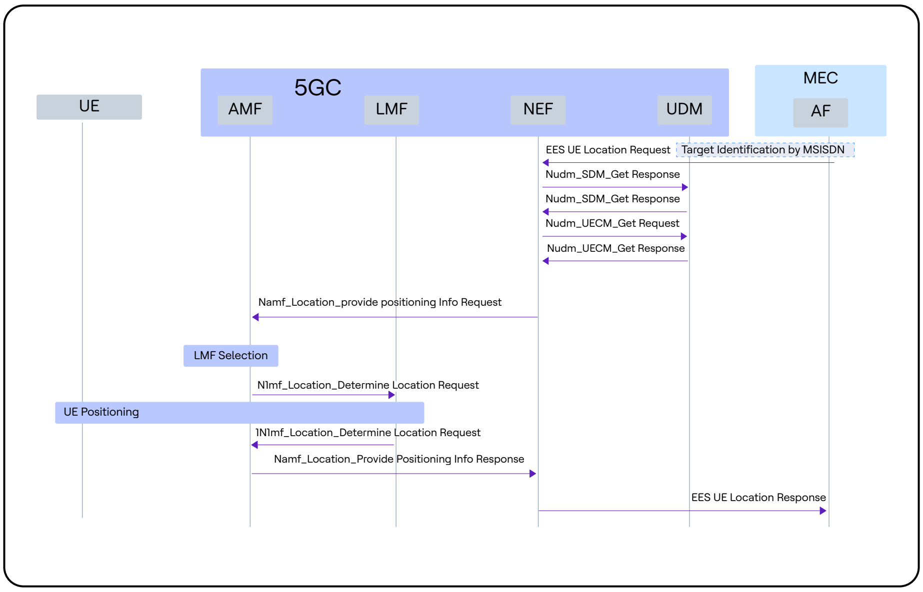


Figure 8: Location services on demand

Figure 8 illustrates the mobile termination location request procedure. When there is a need to determine the location of a particular user, EES UE Location Request message will get triggered from AF. The request is received by NEF whichuses Nudm_SDM and Nudm_UECM operations for user's subscription. Namf_Location_ProvidePositioningInfo service operation is used to request the current location of the UE towards AMF. The service operation includes the SUPI, and client type and may include the required QoS and Supported GAD shapes. AMF performs LMF selection to determine location estimation of the target UE.

AMF invokes the Nlmf_Location_DetermineLocation service operation towards the LMF to request the current location of the UE. LMF performs one or more of the positioning procedures like GNSS, TDOA, AOA or ECID. There can be LPP or NRPP protocols to support finding location of the user. LMF returns the Nlmf_Location_DetermineLocation Response towards the AMF to return the current location of the UE. AMF returns the Namf_Location_ProvidePositioningInfo Response and final response would reach AF with the help of NEF.

6.3 Device identifier

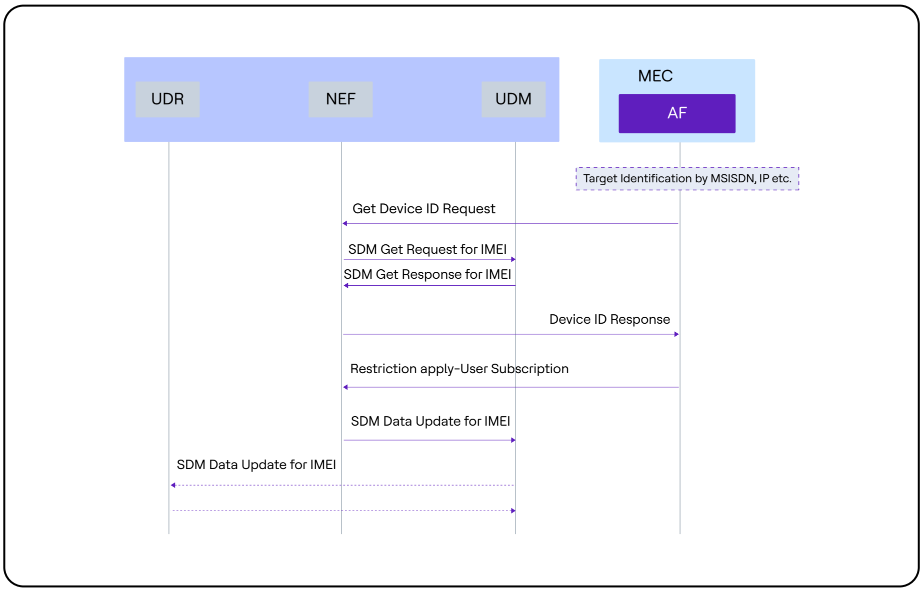


Figure 11: Knowing the device identification in 5G and applying restriction

This use case pertains to the enterprise business logic required to identify their mobile handset manufacturer. Every mobile device used in telephony possesses an International Mobile Equipment Identity (IMEI) number, which serves as a unique identifier for a user's device. While a user registers in the 5G network, they share this number with 5GC. The AF can then request this number and by using IMEI to filter the manufacturer detail database, retrieving information about the mobile handset manufacturer.

Business rules can help retrieve information about the manufacturer of the mobile handset. Within each network, some handsets may not be authorized to operate due to government policies and certain services may be blocked for these users based on external triggers.

7. HCL NAL (Network Abstraction Layer) solution and services

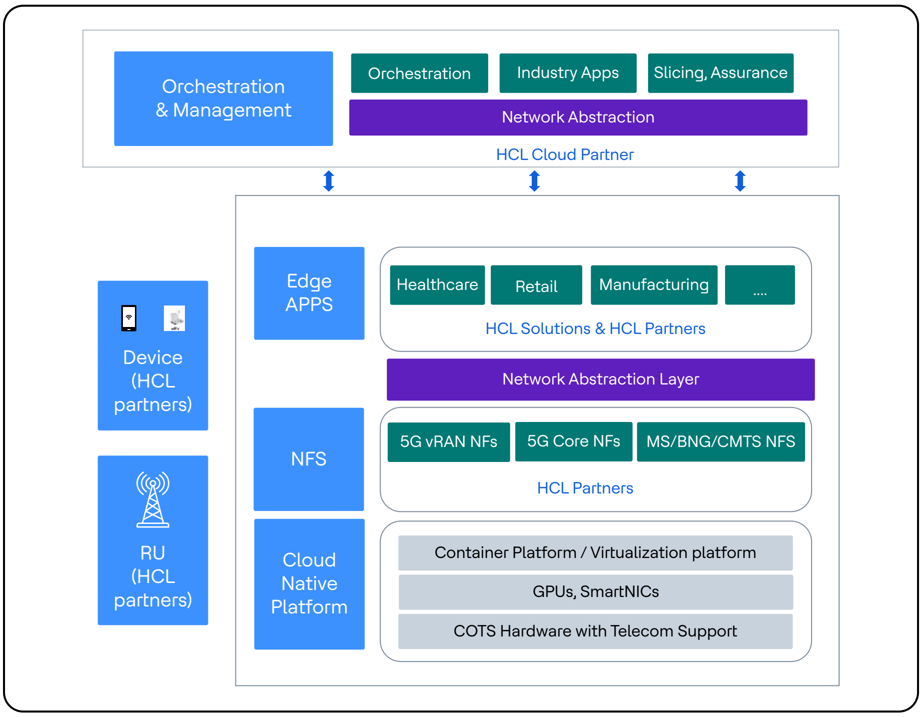
HCLTech's NAL solution enables operators and vendors to accelerate the development of APIs discussed in this document. HCLTech's NAL solution has all the key building blocks (APIs) required for the enterprise use case development.

HCLTech's NAL offers the following key benefits:

- Seamless integration with any vendor 5G Core functions
- Scalable solution with microservices-based architecture

- Seamless integration with enterprise applications

The following diagram shows a conceptual view of HCLTech NAL which can be positioned for both Edge Apps and Network Orchestration and Management.



8. Conclusion

Communications Service Providers can now provide services to an enterprise application by securely exposing the Network APIs with no delay and without additional expenses to set up O&M terminals in the enterprise premises. NAL offers users a great user experience with offerings such as on-demand quality and effective usage of subscriptions eliminating delays and errors per the user requirement. Enterprise applications can support their users with add-on services based on their status, device, location and other parameters.

Enterprise AF can interact with multiple 5G cores with the help of APIs and this will minimize the inter core complex configuration for multiple enterprise 5G use cases.

9. References

- 3GPP TS 23.501 – System architecture for the 5G System (5GS)
- 3GPP TS 23.502 – Procedures for the 5G System (5GS)
- 3GPP TS 23.273 – 5G System (5GS) Location Services
- 3GPP TS 29.503 – 5G System; Unified Data Management Services
- 3GPP TS 29.508 – 5G System; Session Management Event Exposure Service
- 3GPP TS 29.518 – 5G System; Access and Mobility Management Services
- 3GPP TS 29.522 – 5G System; Network Exposure Function Northbound APIs
- 3GPP TS 29.591 – 5G System; Network Exposure Function Southbound Services
- Camara Project – https://camaraproject.org/

10. Authors

Ranganatha Swamy

Principal Consultant

Ranganatha (Ranga) is part of HCLTech's ERS organization and is responsible for defining solutions to address HCLTech's customers in 5G ecosystem that include Semi, OEM, TSPs hyperscalers and enterprises. Ranga has architected many products in wireless and wireline domains in his 24+ years at HCLTech.

Sathish Kumar M

Global Program Director

Satish is part of HCLTech's ERS 5G team. He is responsible for development of HCLTech 5G solution accelerators and integrating with partner solutions to address the Telecom OEM, CSPs, semi and hyperscalers requirements globally. He brings 30 years of rich expertise in the Telecom domain, with 17 years at HCLTech focusing on the Telecom vertical.

Razat Trehan

Solution Architect

Razat is part of HCLTech's ERS organization and key architect of HCLTech's 5G solutions. Prior to joining HCLTech in 2014, Razat worked for Fujitsu, Alcatel-Lucent (Now Nokia) and contributed to many complex 3G and 4G programs. He has nearly 30 years of extensive experience across all generations of wireless and wireline networks. He has also worked with TEC (Telecommunication Engineering Centre, India) for Telecom equipment validation processes. Razat currently manages a program that HCLTech is executing for a leading OEM.

Kuldeep Sing Negi

Technical Architect

Kuldeep is part of HCLTech's ERS organization and is a key architect of HCLTech's 5G solutions. Prior to joining HCLTech in 2018, Kuldeep worked for various global clients including Ericsson, Vodafone, Bell Canada, Sunrise and contributed in many complex 3G and 4G programs. Kuldeep has 20+ years of extensive experience providing solutions in telecom and other domains. He has been involved with teams contributing to open-source telco projects like ONAP, ORAN. Kuldeep currently is associated with a program that HCLTech is executing for a leading OEM.