MOBILE NETWORK EXPOSURE FUNCTION AS A SERVERLESS FUNCTION

INVENTOR:

RAHIL GANDOTRA

Date: Aug 30, 2023

Submitter: Rahil Gandotra

Title: A method for architecting mobile network exposure function as a serverless function

Background:

Mobile network exposure functions

Mobile network cores consist of a network exposure function that enables exposing the network capabilities to authorized third-party applications. The corresponding function for 4G mobile networks is the Service Capability Exposure Function (SCEF), while that for 5G mobile networks is the Network Exposure Function (NEF).

From 3GPP TS 23.682 [1] –

The Service Capability Exposure Function (SCEF) provides a means to securely expose the services and capabilities provided by 3GPP network interfaces. The SCEF provides a means for the discovery of the exposed services and capabilities. The SCEF provides access to network capabilities through homogenous network application programming interfaces (e.g. Network APIs) defined over T8 interface. The SCEF abstracts the services from the underlying 3GPP network interfaces and protocols.

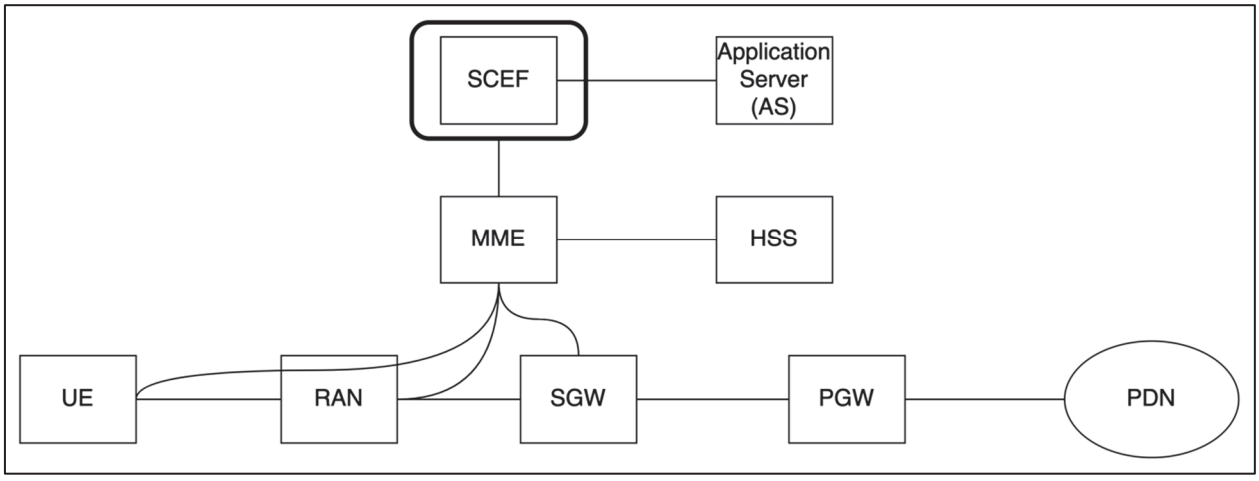

Fig. 1 LTE (4G) Network Architecture depicting SCEF

From 3GPP TS 23.501 [2] –

The Network Exposure Function (NEF) supports external exposure of capabilities of network functions. External exposure can be categorized as Monitoring capability, Provisioning capability, Policy/Charging capability, Analytics reporting capability and UE Member selection capability. The Monitoring capability is for monitoring of specific event for UE in 5G System and making such monitoring events information available for external exposure via the NEF. The Provisioning capability is for allowing external party to provision of information which can be used for the UE in 5G System. The Policy/Charging capability is for handling access and mobility management, QoS and charging policies for the UE based on the request from external party. The Analytics reporting capability is for allowing an external party to fetch or subscribe/unsubscribe to analytics information generated by 5G System. The UE Member selection capability is for allowing an external party to acquire one or more list(s) of candidate UE(s) (among the list of target member UE(s) provided by the AF) and additional information that is based on the assistance information generated by 5G System based on some defined filtering criteria.

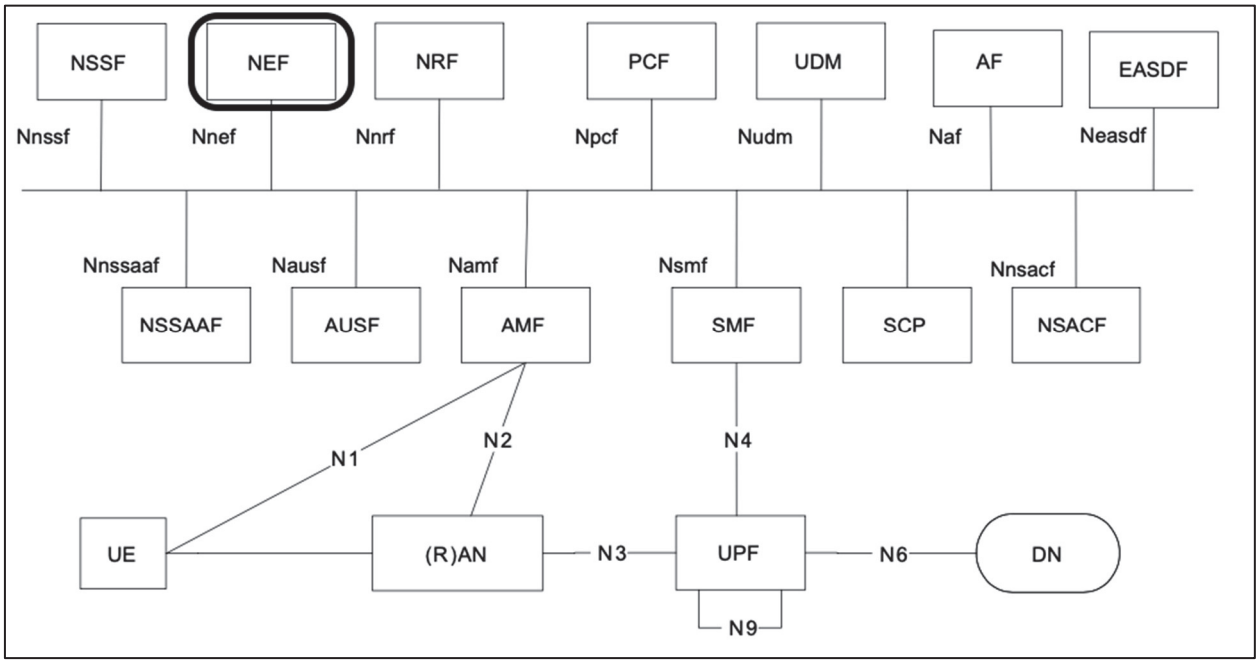

Fig. 2 5G System Architecture depicting NEF [2]

In the scenario of edge computing, a local-NEF (L-NEF) can be deployed in addition to L-PSA-UPF in order to provide local exposure of network information.

Serverless computing

Serverless computing is a new model for cloud-native computing that allows developers to define applications with a set of functions with access to a common data store [3]. This enables different customers to share common pools of servers managed by the cloud provider, thereby eliminating the effort required for server management. Serverless runtimes offer event-driven compute runtimes where the application logic is deployed in containers that is executed on demand. This allows the developed applications to be stateless, ephemeral, and event triggered. Serverless applications find their use for web applications, stream processing, file processing, and IoT and mobile backends.

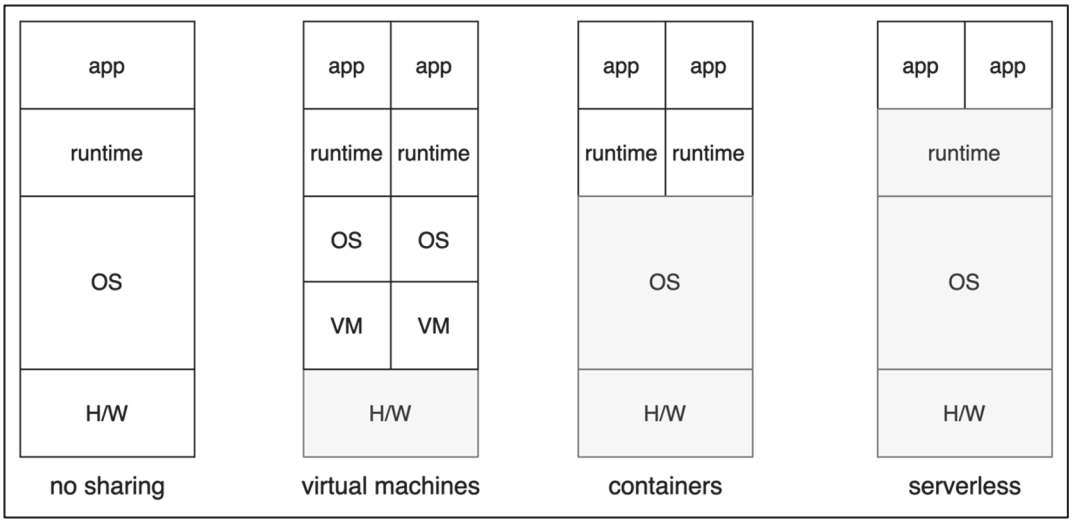

Fig. 3 Evolution of compute sharing models

Each serverless function is defined in terms of event, trigger, and action. Fig. 4 depicts a high-level serverless architecture [4]. Triggers are requests sent to events which are input data sources either using a request/response protocol like HTTP or streaming protocols. Based on the input, the service invokes a user-defined function to take action and provide a response. Serverless computing provides an abstraction of the short-running stateless function by being expressive

enough to build intent-driven applications and simple enough to enable application-agnostic autoscaling.

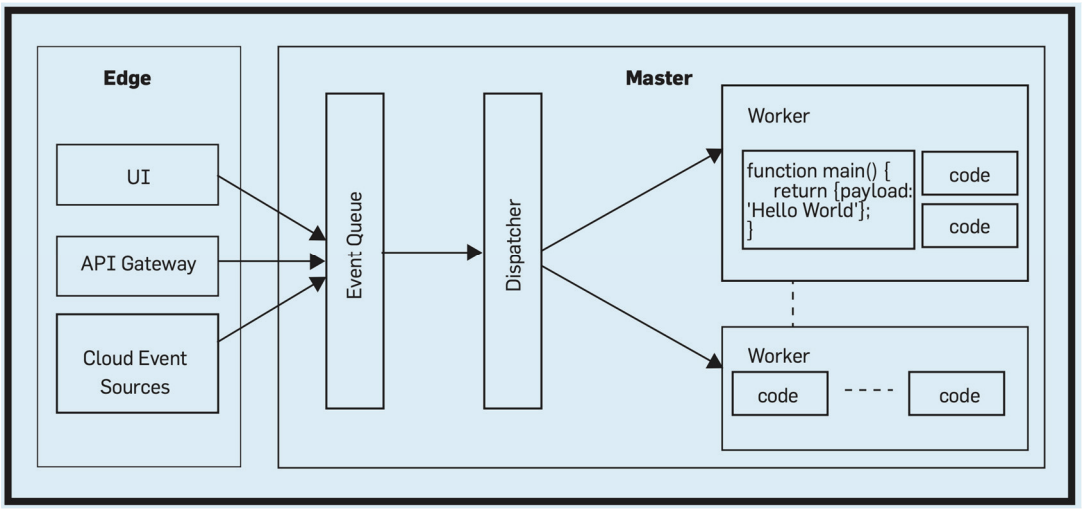

Fig. 4 High-level serverless architecture [4]

Problem statement:

Currently no method exists to develop and operate the mobile network exposure function (like NEF) as a serverless function, allowing for separation of the API functionality from the exposure processing functionality.

Proposed solution:

Rearchitecting the network exposure function (NEF) from a monolithic function to the serverless paradigm separating the API functions acting as events and exposure processing functions acting as actions.

This would involve separating the NEF API functionality from the core processing functionality. Based on the number of incoming requests, the core processing functionality could be scaled up and down to efficiently use the server resources (such as on the edge).

Fig. 5 shows the serverless NEF operation in case of moderate number of incoming requests.

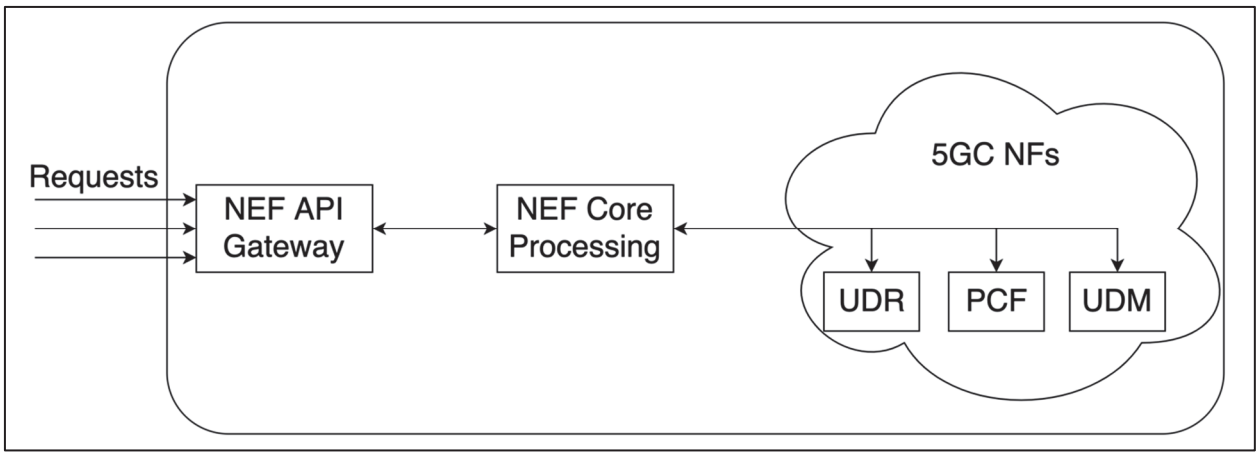

Fig. 5 Serverless NEF operating with moderate number of incoming requests

Fig. 6 shows the serverless NEF operation in case of high number of incoming requests. The NEF processing functionality instances are automatically scaled up to handle the high compute requirements.

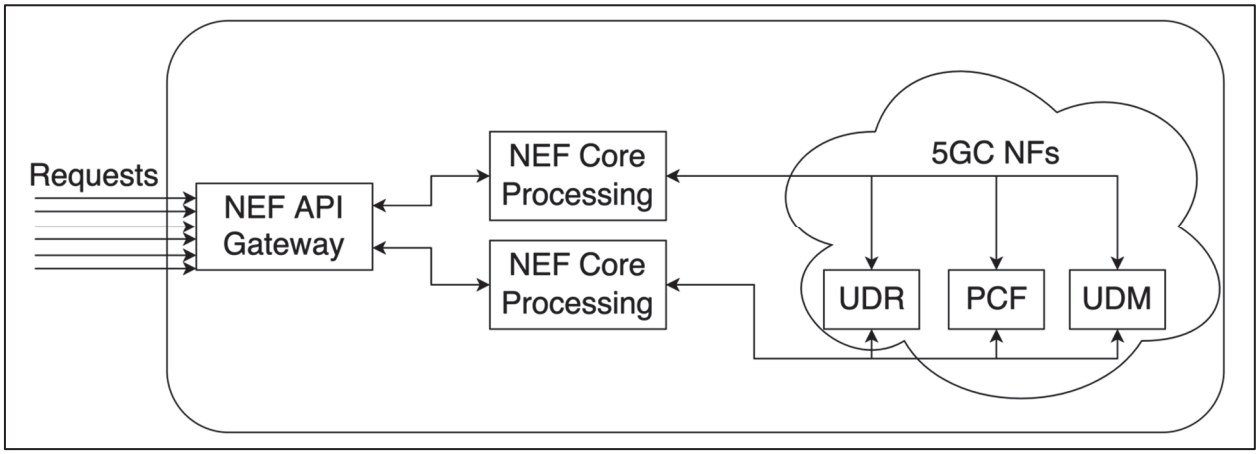

Fig. 6 Serverless NEF operating with high number of incoming requests

Fig. 7 shows serverless NEF operation in case of no incoming requests. The NEF processing functionality instances are automatically scaled down to zero to save on compute resources which could be used by other applications.

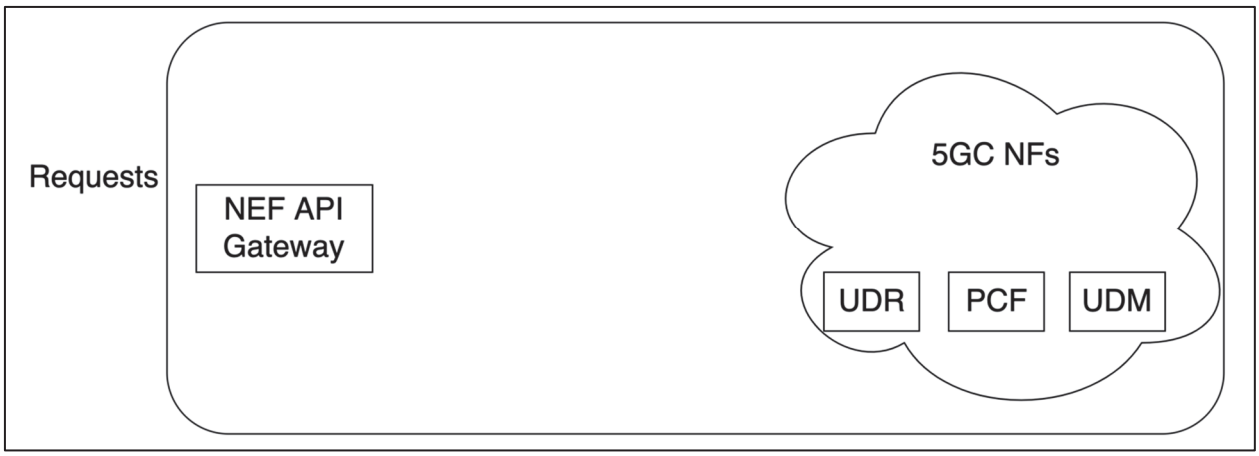

Fig. 7 Serverless NEF operating with no incoming requests

Additionally, the NEF core processing functionality could be subdivided into smaller independent services. This would enable them to be independently deployed and scaled.

Benefits:

- Efficient use of available compute in mobile networks
- Enable utilizing resource-constrained environments such as the edge
- Cost efficiency through systematic scaling of resources

Existing solutions:

None

Interested parties:

MNOs, mobile core vendors, edge infrastructure providers

Impact:

Efficient edge architecture. Evolution of mobile core network.

References:

[1] 3GPP TS 23.682, v18.0.0, Technical Specification Group Services and System Aspects; Architecture enhancements to facilitate communications with packet data networks and applications; Stage 2 (Release 18), (2023-03).

[2] 3GPP TS 23.501, v18.2.2, Technical Specification Group Services and System Aspects; System architecture for the 5G System (5GS); Stage 2 (Release 18), (2023-07).

[3] S. Hendrickson et al., “Serverless computation with OpenLamda,” in the *8th USENIX workshop on hot topics in cloud computing (HotCloud)*, 2016.

[4] P. Castro, V. Ishakian, V. Muthusamy, and A. Slominski, “The rise of serverless computing,” in the *Communications of the ACM*, vol. 62, no. 12, pp.44-54, Nov. 2019.