March 2021

# TECHNIQUES TO FACILITATE SELECTION OF NEXT HOP NRF IN A HIERARCHICAL NETWORK REPOSITORY FUNCTION (NRF) DEPLOYMENT IN 5G NETWORKS

Ravi Shekhar

Satyabrata Seth

Shantanu Bhate

Ameo Ghosh

# TECHNIQUES TO FACILITATE SELECTION OF NEXT HOP NRF IN A HIERARCHICAL NETWORK REPOSITORY FUNCTION (NRF) DEPLOYMENT IN 5G NETWORKS

AUTHORS:
Ravi Shekhar
Satyabrata Seth
Shantanu Bhate
Ameo Ghosh

## ABSTRACT

When a Fifth Generation (5G) Network Repository Function (NRF) deployed in a Third Generation Partnership Project (3GPP) hierarchical model architecture receives a Network Function (NF) Discovery or Subscription or Access Token request, the NRF may not have the required information to serve the request on account of the information being divided among multiple NRF instances in the network. In such cases, the NRF instance that receives the request from a consumer NF may choose to forward or redirect the request to another NRF instance in the hierarchy. The conditions and policies to decide when to redirect or forward the request to another NRF are not defined by 3GPP. Further, the procedures and policies for selecting the next hop NRF are left to implementation. This is a major gap in 3GPP standards as an NRF at each hierarchical level needs rules to determine whether it can serve a request locally or needs to forward/redirect the request to another NRF instance in the hierarchy. This proposal involves defining a granular, flexible and policy-driven system for NRF selection and message routing between hierarchical NRFs to minimize latency and signaling overhead.

## DETAILED DESCRIPTION

A hierarchical Network Repository Function (NRF) architecture, as explained in Third Generation Partnership Project (3GPP) Technical Specification (TS) 29.510, can be deployed in Fifth Generation (5G) core networks. Other NFs in the network will be served by NRFs at various levels, depending on deployment choices, e.g. some (e.g., such as Network Slice Selection Function (NSSF)) may be served by a Public Land Mobile Network (PLMN) level NRF, whereas some others (e.g., such as Session Management Function (SMF)) may be served by a regional or slice-specific NRF.

When a 5G NRF deployed in 3GPP hierarchical model architecture receives a NF Discovery, Subscription or Access Token request, that NRF may not have the required information to serve the request. In such cases, the NRF that receives request from a consumer NF may choose to forward or redirect the request to another NRF instance in the hierarchy. The conditions and policies to decide when to redirect or forward the request to another NRF are not defined by 3GPP. Further, the procedures and policies for selecting the next hop NRF are left to implementation. This is a major gap in 3GPP standards as an NRF at each hierarchical level needs rules to determine whether it can serve a request locally or needs to forward/redirect the request to another NRF instance in the hierarchy.

Additionally, selection of the next hop is also costly and could be complex in larger deployments in which many slices or deployment regions may be involved. Given the lack of a proper routing mechanism, the selection of a next hop NRF as well as the forwarding/redirection decision of packets to the final NRF could significantly affect the overall signaling latency in hierarchical architectures.

This proposal involves defining a granular, policy-driven system for NRF selection and message routing between hierarchical NRFs to minimize latency and signaling overhead. In particular, techniques herein provide for defining a policy-based system that can facilitate routing 5G Nnrf requests between multiple NRF instances deployed in the same PLMN. The system can intelligently identify an optimal route to reduce the number of hops to reach an appropriate NRF and reduce end-to-end latency. Additionally, the system also provides a flexible policy mechanism that can be adapted to different deployment models (e.g. geographically distributed with multiple deployment regions, slicing-based deployments with PLMN level, slice-shared and slice-dedicated NFs, etc.) such that network operators will have flexibility to define, update, and/or remove policies at any time on any chosen NRF in which all the NRFs in a hierarchy will be updated accordingly.

Consider Figure 1, below, which illustrates an example hierarchical NRF deployment model in which multiple NRFs are deployed in a network.

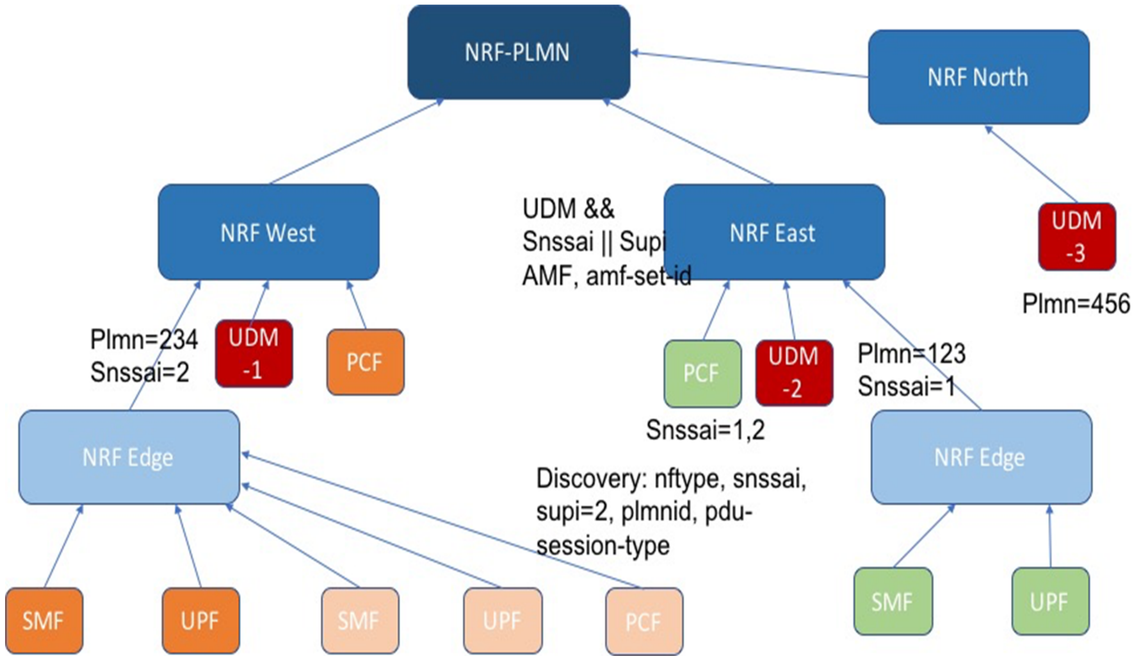



*Figure 1: Hierarchical NRF Deployment Model*

For Figure 1, a granular policy-driven system can be provided for routing of Nnrf messages between multiple NRF instances deployed in the same PLMN. The system defines a tree-structured deployment topology for NRF instances. It also defines the relationship between a parent NRF node and its children. The system further defines policies that can be configured and used by operators to determine routes for NRF to NRF communication within the hierarchy.

In particular, for the techniques herein, the system defines a framework for a granular policy to define conditions that should be satisfied to forward/redirect a NF Discovery, Subscription or Access Token request to other NRF(s) within the network. The parameters used in the policy for each type of message (Discovery, Subscription, Access Token) are based on the Application Programming Interface (API) Data Model defined by 3GPP for various Nnrf operations.

For example, the policy is used by an NRF node to forward/redirect an Nnrf request to its parent. The policy can be enabled for any NF-type based on the customer deployment model. The policy further allows an NRF instance to either immediately forward/redirect the message to its parent NRF node or to check if the request can be served using locally available information before forwarding/redirecting the request to its parent NRF node.

The policy further defines criteria that should be satisfied to consider a request for forwarding/redirection within the NRF hierarchy. The criteria may be based on the presence or absence of certain mandatory, conditional or optional request parameters defined by 3GPP for Nnrf service operations (i.e., NF Discovery, Status Subscription, and Access Token).

The policy rule may include at least one of the following parameters (all of these parameters are standard Information Elements (IEs) defined by 3GPP for Nnrf operations):

- nfInstanceId: NF instance ID of the NF for which a message is requested;
- xyzInfo (e.g. udmInfo, pcfInfo, etc.): The NF info parameters for the particular nf-type as specified in its NF Profile registered to NRF. Applicable nf-type and respective nfInfo parameters are provided in Figure 2, below;
- sNssai: Single-Network Slice Selection Assistance Information (sNSSAI) of the NF for which the message is requested;
- nsi: Network Slice Instance (NSI) of the NF for which the message is requested; and
- plmn: PLMN ID of the NF for which the message is requested.

| target-nf-type | Query Parameters |
|---|---|
| UDR | supi, gpsi, external-group-identity, data-set, group-id-list |
| UDM | supi, gpsi, external-group-identity, routing-indicator, group-id-list |
| AUSF | supi, routing-indicator, group-id-list |
| AMF | tai, amf-region-id, amf-set-id, guami |
| SMF | dnn, tai, pgw-ind, pgw, access-type |
| UPF | dnn, smf-serving-area, dnai-list, pdu-session-types |
| PCF | supi, dnn |
| BSF | dnn, ue-ipv4-address, ip-domain, ue-ipv6-prefix |
| CHF | supi, gpsi, chf-supported-plmn |

*Figure 2: Query Parameters*

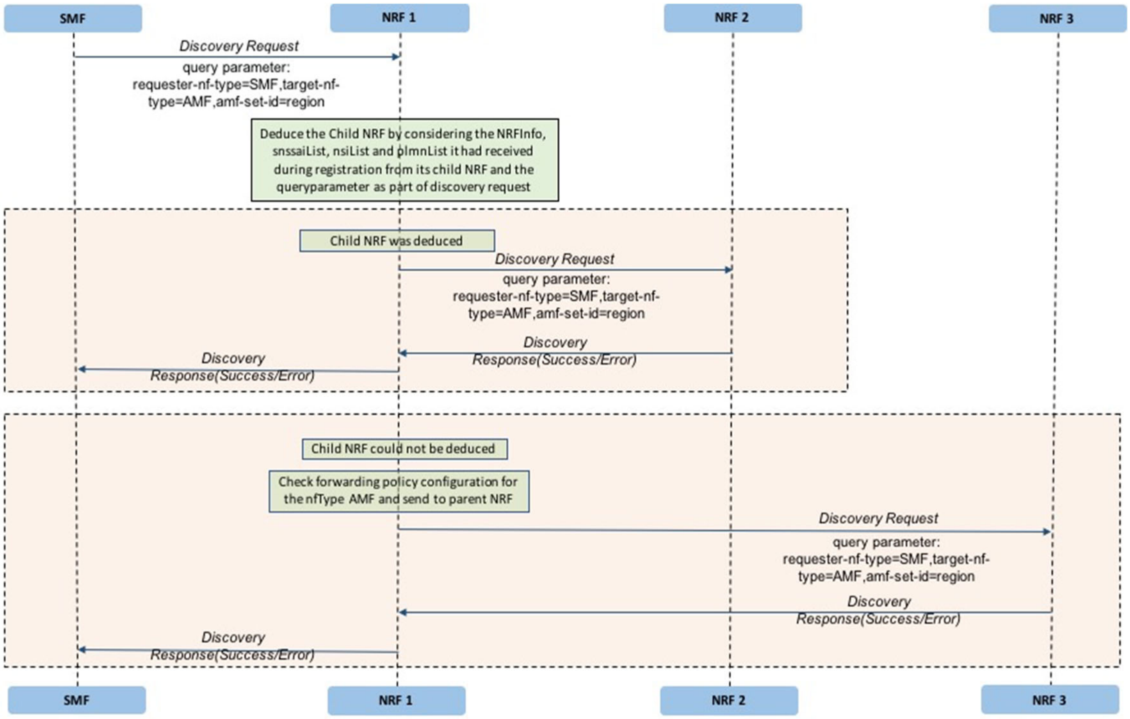
The system of Figure 1 provides for defining a decision-making process that can used to determine a route, i.e., a next-hop NRF for forwarding/redirecting a message. The decision-making process can be different for each type of message (Discovery, Subscription, and Access Token).

During operation, if a leaf NRF node receives a request from a consumer NF, the leaf NRF is to route the request to its parent (if applicable). If an intermediate level NRF receives a request from a consumer NF or from another NRF, the request can be routed to either a child NRF node or parent NRF node of the intermediate NRF, depending on the request parameters and information available locally. If the root NRF receives a request from a consumer NF or from another NRF, the request is to be routed to one of the children of the root NRF, if applicable. A next-hop NRF node can be determined using the policy as discussed above along Nnrf information of the child nodes (applicable for intermediate and root NRF).

The system further provides for defining information that is configured in each NRF and conveyed by a child NRF to its parent do help in determining the next hop NRF. For example, a child NRF node may construct Nnrf information using the information of NF profiles registered with the node (as defined in TS 29.510) and the Nnrf information of its children (if applicable). The NF profile of a child NRF registered with its parent may include the sNSSAI, NSI, and PLMN ID (as defined in TS 29.510). The value of each of these IEs is the superset of all values of the same IE in NF profiles registered with the child NRF. For example, the sNSSAI of a child NRF would be the superset of sNSSAI values of all the NF profiles registered with the child NRF.

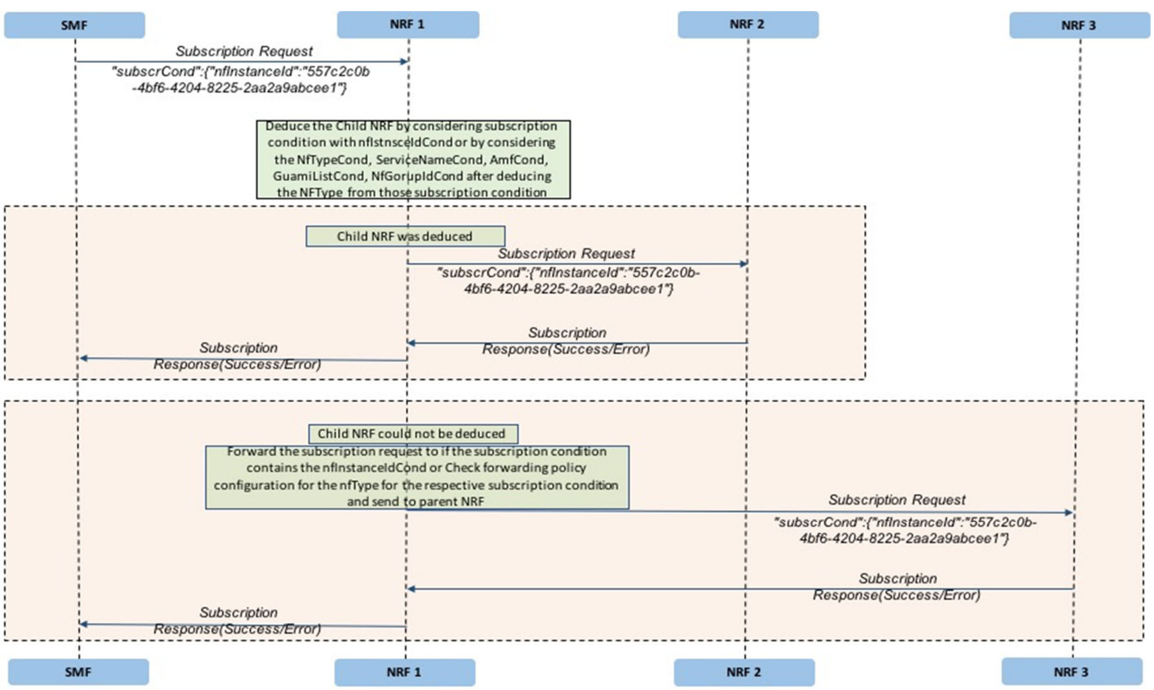
Finally, the system also provides for defining two types of behavior for an intermediate NRF when multiple candidates for a next-hop NRF are identified. In one instance, if an intermediate NRF determines that more than one of its child NRF nodes can potentially serve the request, the least loaded child NRF node is chosen as the next-hop NRF. In another instance (only for NF Discovery requests), if an intermediate NRF determines that more than one of its child NRF nodes can potentially serve the request, the request can be sent to all candidate child NRF nodes and responses can be merged together and sent back.

Various flow diagrams illustrating example details of the techniques herein are illustrated below in Figures 3 (Discovery Forwarding), 4 (Subscription Forwarding), and 5 (Access Token Forwarding).




NOTE: NRF2 is child NRF of NRF1
NRF3 is parent NRF of NRF1

*Figure 3: Discovery Forwarding*




NOTE: NRF2 is child NRF of NRF1
NRF3 is parent NRF of NRF1

*Figure 4: Subscription Forwarding*

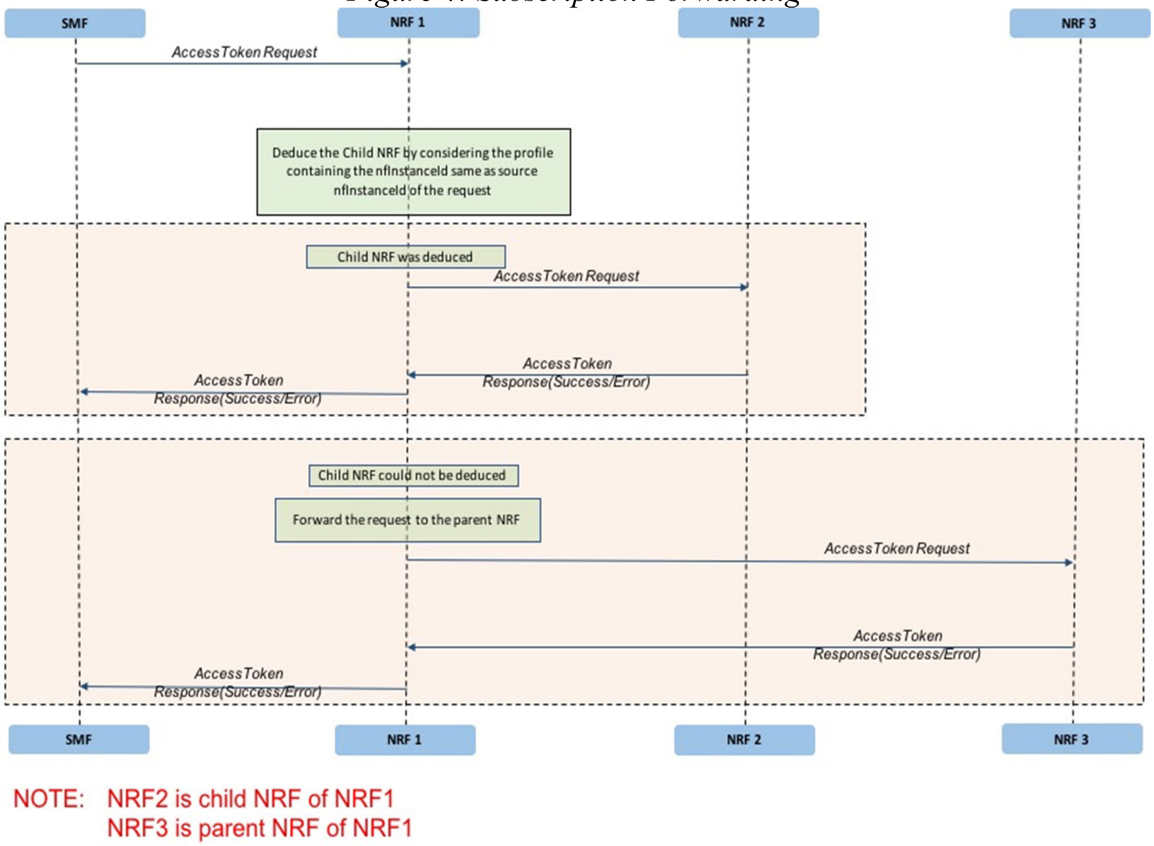



*Figure 5: Access Token Forwarding*

In summary, techniques described herein may provide a policy-based system that may facilitate routing 5G Nnrf requests between multiple NRF instances deployed in the same PLMN. The system intelligently identifies an optimal route to reduce the number of hops to reach the appropriate NRF and reduce end-to-end latency. The system also provides a flexible policy mechanism that can be adapted to different deployment models. Utilizing techniques described herein, network operators will have flexibility to define, update, and/or remove policies at any time and on any chosen NRF such that all the NRFs in the hierarchy will be updated accordingly.