Oracle® Communications

Cloud Native Core, Network Repository Function Benchmarking Guide

Release 24.2.6
G11905-05
October 2025

Oracle Communications Cloud Native Core, Network Repository Function Benchmarking Guide, Release 24.2.6

G11905-05

Contents

My Oracle Support

My Oracle Support (https://support.oracle.com) is your initial point of contact for all product support and training needs. A representative at Customer Access Support can assist you with My Oracle Support registration.

Call the Customer Access Support main number at 1-800-223-1711 (toll-free in the US), or call the Oracle Support hotline for your local country from the list at http://www.oracle.com/us/support/contact/index.html. When calling, make the selections in the sequence shown below on the Support telephone menu:

- For Technical issues such as creating a new Service Request (SR), select **1**.
- For Non-technical issues such as registration or assistance with My Oracle Support, select **2**.
- For Hardware, Networking and Solaris Operating System Support, select **3**.

You are connected to a live agent who can assist you with My Oracle Support registration and opening a support ticket.

My Oracle Support is available 24 hours a day, 7 days a week, 365 days a year.

Acronyms

The following table provides information about the acronyms and the terminology used in the document.

Table Acronyms

Term	Definition
ASM	Aspen Service Mesh
B	Bytes
CA	Certificate Authority
CNE	Cloud Native Environment
CPU	Central Processing Unit
KB	Kilo Bytes
LB	Load Balancer
NF	Network Function
NRF	Oracle Communications Cloud Native Core, Network Repository Function
OSO	Operations Services Overlay
RAM	Random Access Memory
RTT	Round Trip Time
SBA	Service Based Architecture
Set	A set comprises of multiple Georedundant NRF sites
TLS	Transport Layer Security
XFCC	x-forwarded-client-cert

What's New in This Guide

This section lists the documentation updates for release 24.2.x.

Release 24.2.6- G11905-04, October 2025

There are no changes made to this document in this release.

Release 24.2.4- G11905-03, June 2025

- Updated the pod replica count and the min and max CPU or Pod count for the following services in the NRF Services Resource Requirements section:
  - <helm-release-name>-nfdiscovery
  - <helm-release-name>-ingressgateway
  - <helm-release-name>-egressgateway
- Updated the average CPU percentage in the NRF Resource Utilization section for the following services:
  - nfdiscovery
  - ingressgateway
  - egressgateway
- Added a note in the NF Count Per NF Types section to mention the impact of NF Profile size increase in performance.
- Updated the Targeted TPS value of NRF Forwarded Discovery (Incoming) request from 4150 to 5160 in the Call-Mix section.

Release 24.2.4- G11905-02, April 2025

- Updated the NRF version test constraint in the Software Test Constraints section.
- Modified the NRF Features section with the following details:
  - Updated the notes for the following features:
    - Subscriber Location Function
    - Access Token Caching during SLF Query
    - Preferred Locality Feature Set
    - NRF Forwarding per NF Type/Service Type
    - Maximum Number of Profiles in Discovery Response
  - Added the following features:
    - NF Screening Feature (Rules)
    - NRF Roaming support (hNRF + vNRF)
    - Access-token Expiry Time
    - Discovery Validity Feature
    - Load Change Threshold for Notifications from NRF
    - Monitoring the Availability of SCP Using SCP Health APIs
    - NRF to Retry Alternate Producer to Support Maximum Diversity

    - Error Response Enhancement
    - Error Log Enhancements
    - NRF support for A_PFD DataSetId in registration and discovery (DataSetId Enhancements)
    - NFService Priority Update
    - Subscriber Identity in messages (3GPP-SBI-Correlation-Info)
    - Egress Gateway Pod Throttling
- Updated the Observations for 51.6K TPS section to include the following sections for 51.6K TPS.
  - Modified the table in the STATE Data Quantum at NRF before Performance Run Begins section.
  - Added the following sections:
    - NF Count Per NF Types
    - Discovery TPS per NfType and Number of Profiles in Response

Release 24.2.0- G11905-01, August 2024

- Updated the versions of OpenSearch and OpenSearch Dashboard in the Observability Services section.
- Updated the cnDBTier version in the System Software section.
- Updated the NRF version test constraint in the Software Test Constraints section.
- Added the SLF Routing using SCP feature in the NRF Features section.
- Updated the Observations for 38.17K TPS section to include the following sections for 38.17K TPS.
  - Updated the NRF microservices resource utilization for benchmarking in the NRF Resource Utilization section for the following microservices:
    - nfsubscription
    - nrfauditor
    - egressgateway
  - Updated the cnDBTier microservices resource utilization for benchmarking in the cnDBTier Services Resource Utilization section for the following microservices:
    - SQL (ndbappmysqld)
    - DB (ndbmtd)
    - SQL (ndbmysqld)
    - MGMT (ndbmgmd)
    - Monitor Service (db-monitor-svc)
  - Corrected the value of SQL (ndbappmysqld) service pod count from 4 to 2 in the cnDBTier Services Resource Utilization section.
  - Latency Observations

1 Introduction

Oracle Communications Cloud Native Core Network Repository Function (NRF) is one of the main components of the 5G Service Based Architecture (SBA). NRF maintains an updated repository of all the Network Functions (NFs) available in the operator's network along with the services provided by each NFs in the 5G core. NRF instantiates, scales, and terminates the services with minimal or no manual intervention.

For more information about NRF architecture, see *Oracle Communications Cloud Native Core, Network Repository Function User Guide.*

1.1 Purpose and Scope

This document describes the measurements that are used to perform the performance evaluation of NRF, microservices, and deployment environment setup software such as Aspen Service Mesh (ASM), Operations Services Overlay (OSO), Cloud Native Environment (CNE), and so on.

This document provides the following information:

- Benchmark NRF performance and capacity
- Benchmark data from our labs
- Key metrics used to manage NRF performance and capacity

> **Note**
>
> The performance and capacity of the NRF system may vary based on the call model, Feature or Interface configuration, and underlying CNE and hardware environment.

1.2 References

Following are the reference documents:

- *Oracle Communications Cloud Native Core, Cloud Native Environment Installation, Upgrade, and Fault Recovery Guide*
- *Oracle Communications Cloud Native Core, Operations Services Overlay Installation and Upgrade Guide*
- *Oracle Communications Cloud Native Core, cnDBTier User Guide*
- *Oracle Communications Cloud Native Core, Network Repository Function User Guide*
- *Oracle Communications Cloud Native Core, Network Repository Function REST Specification Guide*
- *Oracle Communications Cloud Native Core, Network Repository Function Troubleshooting Guide*
- *Oracle Communications Cloud Native Core, Network Repository Function Network Impact Report*

- *Oracle Communications Cloud Native Configuration Console User Guide*
- *Oracle Communications 5G Automated Test Suite Guide*
- *Oracle Communication Cloud Native Core, Data Collector User Guide*

2
Deployment Environment

This section provides information about the cloud native infrastructure used for Oracle Communications Cloud Native Core, Network Repository Function (NRF) benchmarking.

2.1 Deployed Components

This section provides details about the deployed components.

2.1.1 Hardware Details

This section describes the hardware details.

Table 2-1 Hardware Details

Nodes	Type
Master Nodes	ORACLE SERVER X8-2
Top of Rack Switch	Cisco 93108TC-EX
Worker Nodes	ORACLE SERVER X8-2

2.1.2 Observability Services

This section describes the observability services used for NRF benchmarking.

Table 2-2 Observability Services

Service Name	Version
OpenSearch	2.11.0
OpenSearch Dashboard	2.11.0
Fluentd OpenSearch	1.16.2
Kyverno	1.9.0
Grafana	9.5.3
Prometheus	2.51.1
Jaeger	1.52.0
MetalLB	0.14.4

2.1.3 System Software

This section describes the system software details.

Table 2-3 System Software

System Software	Details
Operating System (+Kernel Version)	5.15.0-105.125.6.2.2.el8uek.x86_64 Oracle Linux Server 8.8
Hypervisor	BareMetal Server
CNE	23.3.1
OSO	NA
Kubernetes	1.26.x
ASM	1.14.6
cnDBTier	24.2.4

2.2 Resource Requirements

The CPU and RAM resources that each service consumes are constrained, so that they do not consume excess resources that could be used by applications. Each service is initially allocated a CPU and RAM at the time of deployment and is allowed to grow to a specified upper limit of each resource while it continues to run. For services where little growth is expected, or where increasing the CPU or RAM underneath a running application might cause an unacceptable service disruption, the initial allocation and upper limit are set to the same value. The resource requests and limits are given below:

> **Note**
>
> Max replica count of a service must be adjusted as per the target TPS.

2.2.1 CNE Common Applications

This section describes the CNE common applications.

Table 2-4 CNE Common Applications

Service	Pod replica #	CPU/Pod		Memory (Mi)	
		Requests	Limit	Requests	Limit
Prometheus	2	2000	2000	4096	4096
Prometheus Node Exporter	1 per node	800	800	512	512
Prometheus Operator	1	100	200	100	200
Prometheus AlertManager	2	20	20	64	64
Prometheus Kube State Metrics	1	20	20	32	100
Promxy	1	100	100	512	512
OpenSearch Master	3	1000	1000	2048	2048
OpenSearch Data	3	1000	1000	16384	16384

Table 2-4 (Cont.) CNE Common Applications

Service	Pod replica #	CPU/Pod		Memory (Mi)	
		Requests	Limit	Requests	Limit
OpenSearch Client	3	1000	1000	2048	2048
OpenSearch Dashboard	1	100	100	512	512
occne-metrics-server	1	100	100	200	200
occne-alertmanager-snmp-notifier	1	100	100	128	128
Fluentd OpenSearch	1 per worker node	500	500	1024	1024
Jaeger Agent	1 per worker node	256	500	128	512
Jaeger Collector	1	500	1250	512	1024
Jaeger query	1	256	500	128	512
MetalLB Controller	1	100	100	100	100
MetalLB Speaker	1 per worker node	100	100	100	100
LB Controller (vCNE only)	1	10	500	128	1024
Egress Controller (vCNE only)	1 per worker node	10	1000	200	500
Bastion Controller	1	10	200	128	256
Kyverno	3	100	200	256	512

2.2.2 NRF Services

Table 2-5 NRF Services Resource Requirements

Service Name	Min Pod Replica Count per NRF
Helm test	1
`<helm-release-name>-nfregistration`	2
`<helm-release-name>-nfdiscovery`	2
`<helm-release-name>-nfsubscription`	2
`<helm-release-name>-nrfauditor`	2
`<helm-release-name>-nrfconfiguration`	1
`<helm-release-name>-nfaccesstoken`	2
`<helm-release-name>-nrfartisan`	1
`<helm-release-name>-nrfcachedata`	2
`<helm-release-name>-ingressgateway`	2
`<helm-release-name>-egressgateway`	2
`<helm-release-name>-alternate-route`	2
`<helm-release-name>-appinfo`	2
`<helm-release-name>-perfinfo`	2

Where, <helm-release-name> is prefixed in each microservice name. For example, if helm-release-name is "ocnrf", then nfregistration microservice name is "ocnrf-nfregistration".

For more information about other resource requirements, see *Oracle Communications Cloud Native Core, Network Repository Function Installation, Upgrade, and Fault Recovery Guide*.

2.2.3 cnDBTier Services

Table 2-6 cnDBTier Services Resource Requirements

Service Name	Min Pod Replica Count per cnDBTier
MGMT (ndbmgmd)	2
DB (ndbmtd)	4
SQL (ndbmysqld)	2
SQL (ndbappmysqld)	2
Monitor Service (db-monitor-svc)	1
Backup Manager Service (db-backup-manager-svc)	1
Replication Service - Leader	1
Replication Service - Other	0

> **Note**
>
> There is a change in the Memory/Pod resource requirement for Monitor Service (db-monitor-svc) and SQL (ndbappmysqld) services.

For more information about other resource requirements and recommended parameter values, see *Oracle Communications Cloud Native Core, Network Repository Function Installation, Upgrade, and Fault Recovery Guide*.

3
Test Topology

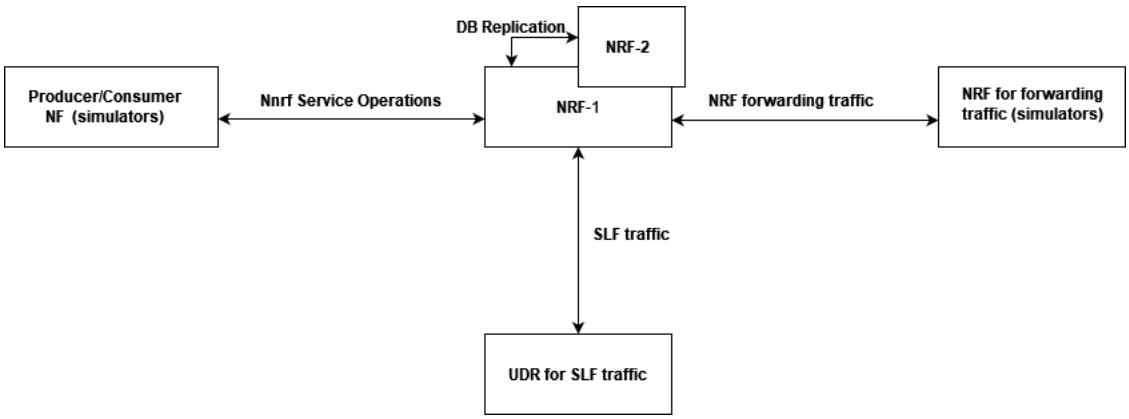
This section describes the topology used for testing Oracle Communications Cloud Native Core, Network Repository Function (NRF).

3.1 Single Set Test Topology

The following image represents the test topology consisting of the following components:

- Two-site NRF
- UDR for SLF traffic
- Producer/Consumer NF for Nnrf service operations (simulators)
- NRF for forwarding to a different segment (simulators)

Figure 3-1 Single Set Test Topology



1. Producer NF sends NfRegister, NfHeartbeat, NfUpdate, and NfDeregister requests to NRF.
2. Consumer NF sends NfStatusSubscribe, NfStatusSubscribe (PATCH), NfStatusUnsubscribe, NfDiscover, NfAccessToken, NfProfileRetrieval, and NfListRetrieval requests to NRF.
3. NRF sends request to UDR for processing SLF based discovery query.
4. NRF sends response to all service requests it receives.
5. NRF triggers NfStatusNotify to all the Consumer NFs who have created subscription.
6. NRF forwards NfDiscovery requests to the NRF simulator.

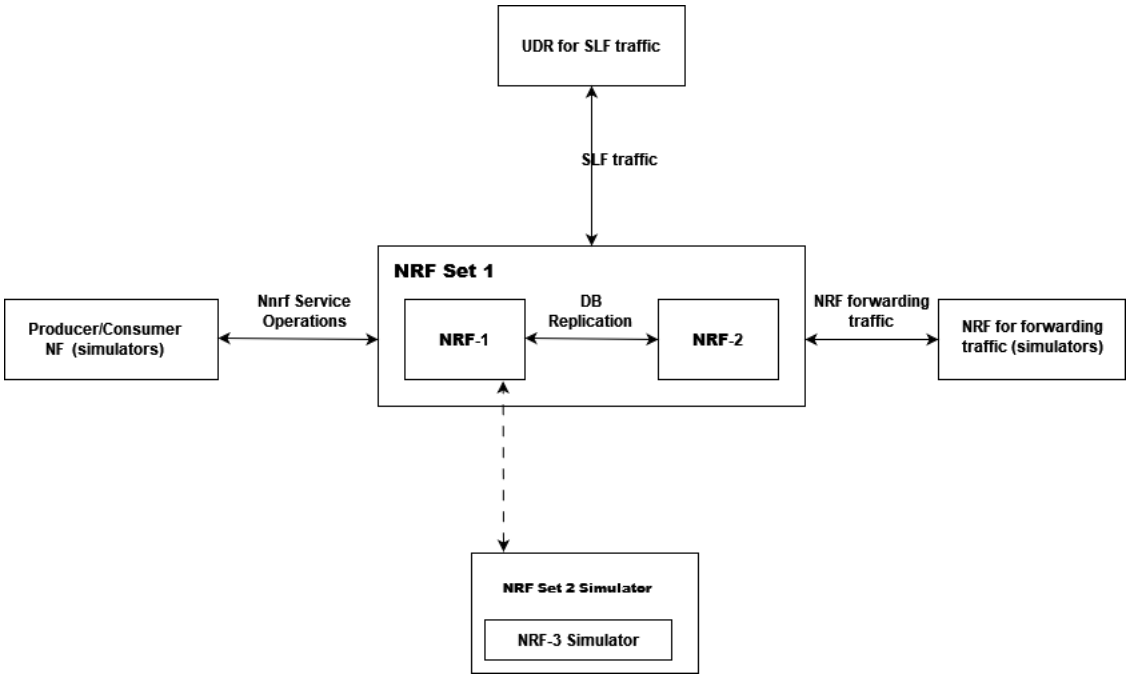
3.2 Multiset Test Topology (NRF Growth)

The following image represents the test topology consisting of the following components:

- Two NRF Sets with one NRF Set having two NRF sites and the other set with one NRF site.

- UDR for SLF traffic
- Producer/Consumer NF for Nnrf service operations (simulators)
- NRF for forwarding to a different segment (simulators)

Figure 3-2 Multiset Test Topology (NRF Growth)



1. Producer NF sends NfRegister, NfHeartbeat, NfUpdate, and NfDeregister requests to NRF.
2. Consumer NF sends NfStatusSubscribe, NfStatusSubscribe (PATCH), NfStatusUnsubscribe, NfDiscover, NfAccessToken, NfProfileRetrieval, and NfListRetrieval requests to NRF.
3. NRF sends request to UDR for processing SLF based discovery query.
4. NRF sends response to all service requests it receives.
5. NRF triggers NfStatusNotify to all the Consumer NFs who have created subscription.
6. NRF forwards NfDiscovery requests to the NRF simulator.
7. NRF Set 1 retrieves the remote set state data from NRF Set 2 and vice versa.

4 Benchmark Testing

This section describes the environment used for benchmarking Oracle Communications Cloud Native Core, Network Repository Function (NRF). Benchmarking is performed with the settings described in this section.

The default values or recommendations for any required software or resource are available from third-party vendors. The operators may choose different values.

4.1 Overview

To qualify the test run, the following elements are considered:

- Pod restart
- CPU and Memory utilization
- Error rate
- Ingress and egress traffic rate
- Success rate
- Message request and response processing time
- Infrastructure resource requirements and utilization

4.2 NRF Features

The following table lists the features that are enabled during benchmarking.

Table 4-1 NRF Features

NRF Features	Feature Status	Notes
Subscriber Location Function	ENABLED	SLF is enabled for 3 NfTypes (AUSF, UDR, and UDM) in slfLookupConfig.
SLF Selection from Registered NF Profiles	ENABLED	Dynamic SLF feature is enabled.
Access Token Caching during SLF Query	ENABLED	Access Token Caching during SLF Query is enabled with 24 hour validity period based caching. For more information about the `accessTokenCacheEnabled` under `slfOptions`, see *Oracle Communications Cloud Native Core, Network Repository Function Rest Specification Guide*
Overload Control Based on Percentage Discards	ENABLED	For more information about the configuration, see *Oracle Communications Cloud Native Core, Network Repository Function Rest Specification Guide*.
Preferred Locality Feature Set	ENABLED	Extended Preferred locality is configured with 2 locationTypes, 8 locationMappings, 56 preferredLocationDetails.
NRF Forwarding per NF Type/ Service Type	ENABLED for UDM	Forwarding is enabled for discovery services. 1 host is present in nrfHostConfig and 1 rule is present in forwardingRulesConfig.

Table 4-1 (Cont.) NRF Features

NRF Features	Feature Status	Notes
EmptyList in Discovery Response	ENABLED	EmptyList feature is enabled and discoveryValidityPeriodCfg under nfDiscoveryOptions is configured with 5 rules.
Access Token Request Authorization	ENABLED	AuthFeature is enabled with 8 rules in authRulesConfig under nfAccessTokenOptions.
Pod Protection Support for NRF Subscription Microservice	ENABLED	For more information about the configuration, see *Oracle Communications Cloud Native Core, Network Repository Function Rest Specification Guide*.
Subscription Limit	ENABLED	Set the subscription limit to 1000. For more information about the configuration, see *Oracle Communications Cloud Native Core, Network Repository Function Rest Specification Guide*.
User-Agent Header for Outgoing Requests	ENABLED	For more information about the configuration, see *Oracle Communications Cloud Native Core, Network Repository Function Rest Specification Guide*.
NRF to Pick Port from ipEndpoints and FQDN from the NF Service or NF level when Selecting SLF	ENABLED	For more information about the configuration, see*Oracle Communications Cloud Native Core, Network Repository Function Rest Specification Guide*.
Ingress Gateway Pod Protection	ENABLED	For more information about the configuration, see *Oracle Communications Cloud Native Core, Network Repository Function Rest Specification Guide*.
ASM Sidecar	ENABLED	For more information about the configuration, see *Oracle Communications Cloud Native Core, Network Repository Function Rest Specification Guide*.
Maximum Number of Profiles in Discovery Response	ENABLED	Set the value as 12 for this parameter. For more information about the configuration, see *Oracle Communications Cloud Native Core, Network Repository Function Rest Specification Guide*. **Note**: If the value is greater than 12, there will be a performance impact.
NRF Georedundancy	ENABLED	For more information about the configuration, see *Oracle Communications Cloud Native Core, Network Repository Function Rest Specification Guide*.
NRF Growth	ENABLED	For more information about the configuration, see *Oracle Communications Cloud Native Core, Network Repository Function Rest Specification Guide*. **Note**: This feature is enabled only for 38.17K TPS Call-Mix.
SLF Routing using SCP	ENABLED	For more information about the configuration, see*Oracle Communications Cloud Native Core, Network Repository Function Rest Specification Guide*.
Discovery Validity	ENABLED	Discovery validity feature is enabled with a value as 1 minute (for all NF Types). For more information about the configuration, see *Oracle Communications Cloud Native Core, Network Repository Function Rest Specification Guide*.
Access Token Expire Time	NA	The expiry duration for a NRF signed certificate. The limit is 24 hours. For more information about NRF Access Token Service Usage Details, see *Oracle Communications Cloud Native Core, Network Repository Function User Guide*.

Table 4-1 (Cont.) NRF Features

NRF Features	Feature Status	Notes
Load Change Threshold for Notifications from NRF	NA	This is nfNotifyLoadThreshold. This is set as 10%. For more information about the configuration, see *Oracle Communications Cloud Native Core, Network Repository Function Rest Specification Guide*.
Monitoring the Availability of SCP Using SCP Health APIs	ENABLED	For more information about the configuration, see *Oracle Communications Cloud Native Core, Network Repository Function Rest Specification Guide*.
NRF to Retry Alternate Producer to Support Maximum Diversity	NA	The value of ***maxSlfAttempts*** parameter is set to 1 and the value of ***useAlternateScp*** parameter is set to true. For more information about the configuration, see *Oracle Communications Cloud Native Core, Network Repository Function Rest Specification Guide*.
Error Response Enhancement	NA	For more information about the configuration, see *Oracle Communications Cloud Native Core, Network Repository Function Rest Specification Guide*.
Error Log Enhancements	ENABLED	For more information about the configuration, see *Oracle Communications Cloud Native Core, Network Repository Function Rest Specification Guide*.
NRF support for A_PFD DataSetId in registration and discovery (DataSetId Enhancements)	ENABLED	For more information about the configuration, see *Oracle Communications Cloud Native Core, Network Repository Function Rest Specification Guide*.
NFService Priority Update	ENABLED	For more information about the configuration, see *Oracle Communications Cloud Native Core, Network Repository Function Rest Specification Guide*.
Subscriber Identity in messages (3GPP-SBI-Correlation-Info)	ENABLED	For more information about the configuration, see *Oracle Communications Cloud Native Core, Network Repository Function Rest Specification Guide*.
Egress Gateway Pod Throttling	ENABLED	For more information about the configuration, see *Oracle Communications Cloud Native Core, Network Repository Function Installation, Upgrade and Fault Recovery Guide*.

Note

Apart from these features being enabled, rest of the feature configurations have been set to their default values. For more information about the feature configurations, see *Oracle Communications Cloud Native Core, Network Repository Function REST Specification Guide*. For more information about the features, see *Oracle Communications Cloud Native Core, Network Repository Function User Guide*.

4.3 Software Test Constraints

Table 4-2 Software Test Constraints

Test Constraint	Details
NRF Version	24.2.4
Ingress incoming connections	For more information on this, see • Test Observations for 38.17K TPS
Ingress outgoing connections	

Table 4-2 (Cont.) Software Test Constraints

Test Constraint	Details
Egress incoming connections	• Test Observations for 51.6K TPS
Egress outgoing connections	
Average HTTP Request Packet Size	~630 to 11300 bytes (Based on NfRegister Request)
Average HTTP Response Packet Size	~2520-44000 bytes (Discovery Response with nfProfiles)
Average NF Heartbeats Timer per NF	30s
Average size of NF Profiles	~630 to 11300 bytes
Deployment with service mesh enabled	1.14.6

Note

Following simulators are used for performance measurement and their induced latency:

- NF simulator to NRF: 50ms
- NRF to forwarding simulator and forwarding simulator to NRF: 150ms
- NRF to another remote set NRF: 150 ms
- NRF to SLF: 150 ms

4.4 Service Operations

This section list the various service operations with the required query parameters.

NfDiscover Service Operation

The NfDiscover service operations are sent with the following query parameters:

- Target-nf-type, guami, and requester-nssais
- Target-nf-type, requester-nf-type, routing-indicator, and group-id-list
- Target-nf-type, requester-nf-type, supi, and service-names
- Target-nf-type, requester-nf-type, and service-names
- Target-nf-type, requester-nf-type, and group-id-list
- Target-nf-type and pgw
- Target-nf-type, requester-nf-type, supi, and service-names
- Target-nf-type, requester-nf-type, and service-names
- Target-nf-type, dnn, snssai, TAI, and pgw-ind
- Target-nf-type, SUPI, group-id-list, and service-names
- Target-nf-type, dnn, snssai, preferred-locality, and service-names
- Target-nf-type and guami
- Target-nf-type, requester-nf-type, and preferred-locality

- Target-nf-type, requester-nf-type, preferred-locality, and SUPI
- Target-nf-type, requester-nf-type, SUPI, group-id-list, requester-nf-type, and data-set
- Target-nf-type, requester-nf-type, routing-indicator, group-id-list, and service-names

4.5 CNE Common Services Resource Utilization

The following table describes CNE microservices and their utilization.

Table 4-3 CNE Common Services Resource Utilization

CNE Observability Services	Pod Count	Average CPU	Average Memory
occne-prometheus-server-1	2	8.07%	64%
occne-elastic-elasticsearch-master-0	1	0.53%	65%
occne-elastic-elasticsearch-master-1	1	0.56%	68%
occne-elastic-elasticsearch-master-2	1	0.53%	65%
occne-elastic-elasticsearch-data-1	1	0.31%	53%
occne-elastic-elasticsearch-data-2	1	0.72%	93%
occne-prometheus-node-exporter	16	0.80%	8%

4.6 Observations for 38.17K TPS

This section provides the details of observations noticed for 38.17K TPS.

4.6.1 NRF Services Resource Requirements

This section provides the resource requirement for NRF deployment.

Table 4-4 NRF Services Resource Requirements

Service Name	Pod Replica	Min CPU/Pod	Max CPU/Pod	Min Memory/Pod (in G)	Max Memory/Pod (in G)	Min (Mi) Ephemeral Storage	Max (Gi) Ephemeral Storage
Helm test	1	1	2	1	2	78.1	1
<helm-release-name>-nfregistration	2	2	2	3	3	78.1	1
<helm-release-name>-nfdiscovery	68	8	8	3	3	78.1	2
<helm-release-name>-nfsubscription	2	2	2	3	3	78.1	1
<helm-release-name>-nrfauditor	2	2	2	3	3	78.1	1
<helm-release-name>-nrfconfiguration	1	2	2	2	2	78.1	1
<helm-release-name>-nfaccesstoken	2	2	2	2	2	78.1	1
<helm-release-name>-nrfartisan	1	2	2	2	2	78.1	1
<helm-release-name>-nrfcachedata	2	4	4	3	3	78.1	1

Table 4-4 (Cont.) NRF Services Resource Requirements

Service Name	Pod Replica	Min CPU/Pod	Max CPU/Pod	Min Memory/Pod (in G)	Max Memory/Pod (in G)	Min (Mi) Ephemeral Storage	Max (Gi) Ephemeral Storage
<helm-release-name>-ingressgateway	34	4	4	4	4	78.1	1
<helm-release-name>-egressgateway	30	4	4	4	4	78.1	1
<helm-release-name>-alternate-route	2	2	2	4	4	78.1	1
<helm-release-name>-appinfo	2	1	1	1	1	78.1	1
<helm-release-name>-perfinfo	2	1	1	1	1	78.1	1

4.6.2 STATE Data Quantum at NRF before Performance Run Begins

The following table describes the parameters considered for the performance run.

Table 4-5 STATE Data Quantum at NRF before Performance Run Begins

Parameters	Values	Details
Number of NF Profiles in DB	125*	Following are the NfTypes: • AMF: 40 • AUSF: 4 • BSF: 2 • CBCF: 5 • CHF: 2 • DRA: 3 • GMLC: 3 • LMF: 3 • NEF: 3 • PCF: 5 • SCP: 4 • SMF: 40 • SMSF: 3 • UDM: 4 • UDR: 10
Number of Subscriptions in DB	500	Subscriptions are created with the below conditions: • Subscription to a set of NF Instances identified by their NF Type. (NfTypeCond). • NotificationCondition with monitored attributes set to “/priority” or “/load”. • reqNotifEvents set to NF_REGISTERED,NF_DEREGISTERED, NF_PROFILE_CHANGE.

> **Note**
>
> * indicates that Set1 hosts 125 NFs and Set 2 (simulator) hosts 125 NFs. When these two sets are combined together with NRF Growth feature enabled, we will have 250 NFs.

4.6.3 Call-Mix

The following table describes the various type of service requests with the required traffic distribution and targeted TPS.

Table 4-6 Call-Mix

S.No	Type of Request	Traffic Distribution (in %)	Targeted TPS
A	NF Register	0.00002	1
B	NF Update	0.00282	1
C	NF Heartbeat	Not considered	26
D	NF Subscribe + Update Subscription	0.01416	5
E	NF Discover	98.45	33,900
F	NRF Forwarded Discovery (Incoming)	10% of Incoming discovery traffic	3,760
G	% of traffic forwarded to remote PLMN (case Visited NRF)	0%	Not included in this test.
H	% of traffic received from remote PLMN (case Home NRF)	0%	Not included in this test.
I	SLF Query	49% of Discovery	18,605
J	NF Access Token for SLF	Extra	0.1
K	NF Profile Retrieval	0.00141	1
L	NF List Retrieval	0.00141	1
M	NF Status Notify	1.52328	480
N	NF Status UnSubscribe	0.00006	1
O	NF DeRegister	0.00002	1
P	NF Access Token for Consumers	0.00141	1
Q	NRF Forwarding (Outgoing)	10% of discovery traffic	3,760

Table 4-7 Total Traffic

Traffic Type	Message Considered	Targeted TPS
Total Incoming Traffic	A + B + C + D + E + F + G + H + J + K + L + N + O + P	37,698
Total Outgoing Traffic	I + M + Q	22,845

4.6.3.1 NF Count Per NF Types

The following table lists the NF count per NF type along with the NF profile size.

Table 4-8 NF Count Per NF Types

Target-NF Types	NF Count	Profile Size (with pretty printing) (in bytes)
AMF	58	8700-11300
AUSF	6	4640
BSF	3	1700
CHF	3	1800
GMLC	6	0630
LMF	6	2400
PCF	30	1500-2500
SMF	94	5700-7400
UDM	6	3200
UDR	36	1700-2000
SLF	3	2100
SMSF	3	11300
NEF	3	11300
CBCF	3	11300

Total **NF Count** is **260**.

4.6.3.2 Discovery TPS per NfType and Number of Profiles in Response

The following table indicates the discovery TPS per NfType and the number of profiles sent in the response.

Table 4-9 Discovery TPS per NfType and Number of Profiles in Response

Target-Nf Types	TPS (in %)	Absolute Discovery TPS	SLF Queries Posted	Number of Profiles in Response
AUSF	0.39	209	209	6
BSF	0.00	0	0	3
CBCF	0.00	0	0	3
CHF	0.077	42.3	0	3
GMLC	0.00	0	0	6
LMF	0.025	14.5	0	6
PCF	0.39	211	0	10
SMF	8.09	4368	0	10
UDM	85.5	46209	46209	6
UDR	1.21	652	652	12
SLF	0	0	0	3
NEF	0.026	14.1	0	3
SMSF	0.025	13.5	0	3

4.6.4 Test Observations

The following table provides the details of test duration and TPS used for the benchmarking test.

Table 4-10 Test Observations

Parameter	Values
Test Duration	60 hours
TPS Achieved	38.17K (considered all Ingress and NfStatusNotify traffic)
Ingress incoming connections	144
Ingress outgoing connections	1.440K
Egress incoming connections	144
Egress outgoing connections	720

4.6.5 NRF Resource Utilization

The following table describes the list of NRF microservices and their utilization.

Table 4-11 NRF Resource Utilization

NRF Microservices	Number of Pods	Average CPU	Average Memory (in Gi)
nfregistration	2	0.28 (7%)	1.69 (28.17%)
nfdiscovery	44	113.79 (64.65%)	57.16 (43.3%)
nfsubscription	2	0.37 (9.17%)	2.18 (36.33%)
nrfauditor	2	0.01 (0.48%)	0.68 (11.33%)
nrfconfiguration	1	0.04 (1%)	0.65 (16.25%)
nfaccesstoken	2	0.03 (0.75%)	1.27 (31.75%)
nrfartisan	1	0 (0%)	0 (0%)
nrfcachedata	2	3.598 (37%)	2.55 (30.75%)
alternate-route	2	0 (0%)	0.36 (9%)
perf-info	2	0.09 (4.5%)	0.25 (12.5%)
appinfo	2	0.11 (5.5%)	0.25 (12.5%)
ingressgateway	18	32.59 (45.26%)	4.5 (8.33%)
egressgateway	18	21.3 (34.2%)	20.15 (27.99%)

> **Note**
>
> The features enabled for this benchmarking testing does not utilize alternate-route service.

4.6.6 cnDBTier Services Resource Utilization

The following table provides observed values of cnDBTier services.

Table 4-12 cnDBTier Services Resource Utilization

Service Name	Pod Count	Average CPU	Memory
SQL (ndbappmysqld)	2	13.8%	35.33%
DB (ndbmtd)	4	5.1%	88.84%

Table 4-12 (Cont.) cnDBTier Services Resource Utilization

Service Name	Pod Count	Average CPU	Memory
SQL (ndbmysqld)	2	0.2%	1.5%
MGMT (ndbmgmd)	2	0.2%	0.35%
Monitor Service (db-monitor-svc)	1	0.2%	10.1 %
Backup Manager Service (db-backup-manager-svc)	1	0.5%	8.9%

Table 4-13 cnDBTier Statistics

DBTier Statistics	Value
Disk Write operations per second	98.4
Disk Read operations per second	439
Transaction rates on data nodes	230

4.6.7 Latency Observations

The following table lists the details of latency observations.

Table 4-14 Latency Observations

Latency Parameter	Details		
	Min (s)	Max (s)	Avg (s)
Average Turnaround time at Ingress Gateway, Discovery Processing, and SLF Egress Transaction	0.145	0.147	0.145
Average Turnaround time for Discovery Processing and SLF Egress Transaction	0.130	0.156	0.131
Turnaround time for SLF Egress Transaction (see Note)	0.126	0.148	0.126

The latency value of Bucketed View, Max, and Average for discovery service operation is maximum among all the other service operations. For more information about the simulators used for performance measurement and their induced latency, see the #unique_19/unique_19_Connect_42_GUID-00ED7E51-66E0-48D8-ADC2-00B7571F60F0 section.

4.7 Observations for 51.6K TPS

This section provides the details of observations noticed for 51.6K TPS.

4.7.1 Observations for 51.6K TPS

This section provides the details of observations noticed for 51.6K TPS.

4.7.1.1 NRF Services Resource Requirements

This section provides the resource requirement for NRF deployment.

Table 4-15 NRF Services Resource Requirements

Service Name	Pod Replica	Min CPU/Pod	Max CPU/Pod	Min Memory/Pod (in G)	Max Memory/Pod (in G)	Min (Mi) Ephemeral Storage	Max (Gi) Ephemeral Storage
Helm test	1	1	2	1	2	78.1	1
<helm-release-name>-nfregistration	2	2	2	3	3	78.1	1
<helm-release-name>-nfdiscovery	68	8	8	3	3	78.1	2
<helm-release-name>-nfsubscription	2	2	2	3	3	78.1	1
<helm-release-name>-nrfauditor	2	2	2	3	3	78.1	1
<helm-release-name>-nrfconfiguration	1	2	2	2	2	78.1	1
<helm-release-name>-nfaccesstoken	2	2	2	2	2	78.1	1
<helm-release-name>-nrfartisan	1	2	2	2	2	78.1	1
<helm-release-name>-nrfcachedata	2	4	4	3	3	78.1	1
<helm-release-name>-ingressgateway	34	4	4	4	4	78.1	1
<helm-release-name>-egressgateway	30	4	4	4	4	78.1	1
<helm-release-name>-alternate-route	2	2	2	4	4	78.1	1
<helm-release-name>-appinfo	2	1	1	1	1	78.1	1
<helm-release-name>-perfinfo	2	1	1	1	1	78.1	1

4.7.1.2 STATE Data Quantum at NRF before Performance Run Begins

The following table describes the parameters considered for the performance run.

Table 4-16 STATE Data Quantum at NRF before Performance Run Begins

Parameters	Values	Details
Number of Subscriptions in DB	1000	Subscriptions are created with the below conditions: • Number of subscriptions to a set of NF Instances identified by their NF Type - 506 • Number of subscriptions to a set of NF Instances identified by their NF Instance Id - 260 • Number of subscriptions to a set of NF Instances identified by their NF Set Id - 156 • Miscellaneous subscriptions - 78 • NotificationCondition with monitored attributes set to “/priority” or “/load”. • reqNotifEvents set to NF_REGISTERED,NF_DEREGISTERED, NF_PROFILE_CHANGE.

4.7.1.3 Call-Mix

The following table describes the various type of service requests with the required traffic distribution and targeted TPS.

Table 4-17 Call-Mix

S.No	Type of Request	Traffic Distribution (in %)	Targeted TPS
A	**NF Register**	0.00002	1
B	**NF Update**	0.00282	1
C	**NF Heartbeat**	Not considered	52
D	**NF Subscribe + Update Subscription**	0.01416	7
E	**NF Discover**	98.8	51,600
F	**NRF Forwarded Discovery (Incoming)**	10% of Incoming discovery traffic	5,160
G	**% of traffic forwarded to remote PLMN (case Visited NRF)**	0%	Not included in this test.
H	**% of traffic received from remote PLMN (case Home NRF)**	0%	Not included in this test.
I	**SLF Query**	88.4% of Discovery	45600
J	**NF Access Token for SLF**	Extra	0.1
K	**NF Profile Retrieval**	0.00141	1
L	**NF List Retrieval**	0.00141	1
M	**NF Status Notify**	1.52328	750
N	**NF Status UnSubscribe**	0.00006	1
O	**NF DeRegister**	0.00002	1
P	**NF Access Token for Consumers**	0.00141	1

Table 4-17 (Cont.) Call-Mix

S.No	Type of Request	Traffic Distribution (in %)	Targeted TPS
Q	**NRF Forwarding (Outgoing)**	10% of discovery traffic	5,160

Table 4-18 Total Traffic

Traffic Type	Message Considered	Targeted TPS
Total Incoming Traffic	A + B + C + D + E + F + G + H + J + K + L + N + O + P	51666
Total Outgoing Traffic	I + M + Q	51511

4.7.1.3.1 NF Count Per NF Types

The following table lists the NF count per NF type along with the NF profile size.

Table 4-19 NF Count Per NF Types

Target-NF Types	NF Count	Profile Size (with pretty printing) (in bytes)
AMF	58	8700-11300
AUSF	6	4640
BSF	3	1700
CHF	3	1800
GMLC	6	0630
LMF	6	2400
PCF	30	1500-2500
SMF	94	5700-7400
UDM	6	3200
UDR	36	1700-2000
SLF	3	2100
SMSF	3	11300
NEF	3	11300
CBCF	3	11300

Total **NF Count** is **260**.

> **Note**
>
> An increase in profile size can result in more resource consumption and lower performance.

4.7.1.3.2 Discovery TPS per NfType and Number of Profiles in Response

The following table indicates the discovery TPS per NfType and the number of profiles sent in the response.

Table 4-20 Discovery TPS per NfType and Number of Profiles in Response

Target-Nf Types	TPS (in %)	Absolute Discovery TPS	SLF Queries Posted	Number of Profiles in Response
AMF	4.27	2163	0	6
AUSF	0.38	210	210	6
BSF	0.00	0	0	3
CBCF	0.00	0	0	3
CHF	0.077	10	0	3
GMLC	0.00	0	0	6
LMF	0.025	2	0	6
PCF	0.39	192	0	10
SMF	8.14	4144	0	10
UDM	85.4	45735	45735	6
UDR	1.22	651	651	12
SLF	0	0	0	3
NEF	0.026	14.8	0	3
SMSF	0.025	13.9	0	3

4.7.1.4 Test Observations

This section following table provides the details of test duration and TPS used for the benchmarking test.

Table 4-21 Test Observations

Parameter	Values
Test Duration	72 hours
TPS Achieved	51666
Ingress incoming connections	287
Ingress outgoing connections	2.169K
Egress incoming connections	240
Egress outgoing connections	1.200k

4.7.1.5 NRF Resource Utilization

The following table describes NRF microservices and their utilization.

Table 4-22 NRF Resource Utilization

NRF Microservices	Number of Pods	Average CPU	Average Memory (in Gi)
nfregistration	2	15.7%	37.9%
nfdiscovery	68	61%	52.7%
nfsubscription	2	8.98%	35.7%
nrfauditor	2	1.10%	32.2%
nrfconfiguration	1	1.79%	47%
nfaccesstoken	2	0.62%	35.8%

Table 4-22 (Cont.) NRF Resource Utilization

NRF Microservices	Number of Pods	Average CPU	Average Memory (in Gi)
nrfartisan	1	0.10%	23.1%
nrfcachedata	2	6.76%	22.7%
alternate-route	2	0 (0%)	0%
perf-info	2	12.3%	13.8%
appinfo	2	7.12%	26.4%
ingressgateway	34	59.8%	58%
egressgateway	30	59%	46.6%

> **Note**
>
> If you enable Message Feed feature at Ingress Gateway and Egress Gateway, approximately 33% pod capacity is impacted.
>
> The features enabled for this benchmarking testing does not utilize alternate-route service.

4.7.1.6 cnDBTier Services Resource Utilization

The following table provides observed values of cnDBTier services.

Table 4-23 cnDBTier Services Resource Utilization

Service Name	Pod Count	Average CPU	Memory
SQL (ndbappmysqld)	2	0.524 (13.1%)	4.134 (67.285%)
DB (ndbmtd)	4	0.988 (4.94%)	16.277 (79.477%)
SQL (ndbmysqld)	2	0.02 (0.2%)	0.51 (1.5%)
MGMT (ndbmgmd)	2	0.015 (0.187%)	0.061 (0.496%)
Monitor Service (db-monitor-svc)	1	0.003 (0.75%)	0.356 (71.2 %)
Backup Manager Service (db-backup-manager-svc)	1	0.001 (0.01%)	0.23 (179.69%)
Replication Service	1	0.0 (0.0%)	0.41 (20.5%)

Table 4-24 cnDBTier Statistics

DBTier Statistics	Value
Disk Write operations per second	83.6
Disk Read operations per second	521
Transaction rates on data nodes	257

4.7.1.7 Latency Observations

Table 4-25 Latency Observations

Latency Parameter	Details		
	Min (s)	Max (s)	Avg (s)
Average Turnaround time at Ingress Gateway, Discovery Processing, and SLF Egress Transaction	0.199	0.463	0.218
Average Turnaround time for Discovery Processing and SLF Egress Transaction	0.199	0.425	0.217
Turnaround time for SLF Egress Transaction (see Note)	0.199	0.404	0.210

The latency value of Bucketed View, Max, and Average for discovery service operation is maximum among all the other service operations. For more information about the simulators used for performance measurement and their induced latency, see the #unique_19/unique_19_Connect_42_GUID-00ED7E51-66E0-48D8-ADC2-00B7571F60F0 section.

4.7.2 Observations for 5K TPS

This section provides the details of observations noticed for 5K TPS.

4.7.2.1 NRF Services Resource Requirements

This section provides the resource requirement for NRF deployment.

Table 4-26 NRF Services Resource Requirements

Service Name	Pod Replica	Min CPU/Pod	Max CPU/Pod	Min Memory/Pod (in Gi)	Max Memory/Pod (in Gi)	Min (Mi) Ephemeral Storage	Max (Gi) Ephemeral Storage
Helm test	1	1	2	1	2	78.1	1
<helm-release-name>-nfregistration	2	2	2	3	3	78.1	1
<helm-release-name>-nfdiscovery	6	8	8	3	3	78.1	2
<helm-release-name>-nfsubscription	2	2	2	3	3	78.1	1
<helm-release-name>-nrfauditor	2	2	2	3	3	78.1	1
<helm-release-name>-nrfconfiguration	1	2	2	2	2	78.1	1
<helm-release-name>-nfaccesstoken	2	2	2	2	2	78.1	1
<helm-release-name>-nrfartisan	1	2	2	2	2	78.1	1
<helm-release-name>-nrfcachedata	2	4	4	3	3	78.1	1
<helm-release-name>-ingressgateway	3	4	4	4	4	78.1	1

Table 4-26 (Cont.) NRF Services Resource Requirements

Service Name	Pod Replica	Min CPU/Pod	Max CPU/Pod	Min Memory/Pod (in Gi)	Max Memory/Pod (in Gi)	Min (Mi) Ephemeral Storage	Max (Gi) Ephemeral Storage
<helm-release-name>-egressgateway	3	4	4	4	4	78.1	1
<helm-release-name>-alternate-route	2	2	2	4	4	78.1	1
<helm-release-name>-appinfo	2	1	1	1	1	78.1	1
<helm-release-name>-perfinfo	2	1	1	1	1	78.1	1

4.7.2.2 Call-Mix

The following table describes the various type of service requests with the required traffic distribution and targeted TPS.

Table 4-27 Call-Mix

S.No	Type of Request	Traffic Distribution (in %)	Targeted TPS
A	**NF Register**	0.00002	1
B	**NF Update**	0.00282	1
C	**NF Heartbeat**	Not considered	52
D	**NF Subscribe + Update Subscription**	0.01416	7
E	**NF Discover**	98.8	5000
F	**NRF Forwarded Discovery (Incoming)**	10% of Incoming discovery traffic	0
G	**% of traffic forwarded to remote PLMN (case Visited NRF)**	0%	Not included in this test.
H	**% of traffic received from remote PLMN (case Home NRF)**	0%	Not included in this test.
I	**SLF Query**	88.4% of Discovery	4420
J	**NF Access Token for SLF**	Extra	0.1
K	**NF Profile Retrieval**	0.00141	1
L	**NF List Retrieval**	0.00141	1
M	**NF Status Notify**	1.52328	500
N	**NF Status UnSubscribe**	0.00006	1
O	**NF DeRegister**	0.00002	1
P	**NF Access Token for Consumers**	0.00141	1
Q	**NRF Forwarding (Outgoing)**	10% of discovery traffic	500

Table 4-28 Total Traffic

Traffic Type	Message Considered	Targeted TPS
Total Incoming Traffic	A + B + C + D + E + F + G + H + J + K + L + N + O + P	5066
Total Outgoing Traffic	I + M + Q	5420

4.7.2.2.1 NF Count Per NF Types

The following table lists the NF count per NF type along with the NF profile size.

Table 4-29 NF Count Per NF Types

Target-NF Types	NF Count	Profile Size (with pretty printing) (in bytes)
AMF	58	8700-11300
AUSF	6	4640
BSF	3	1700
CHF	3	1800
GMLC	6	0630
LMF	6	2400
PCF	30	1500-2500
SMF	94	5700-7400
UDM	6	3200
UDR	36	1700-2000
SLF	3	2100
SMSF	3	11300
NEF	3	11300
CBCF	3	11300

Total **NF Count** is **260**.

> **Note**
>
> An increase in profile size can result in more resource consumption and lower performance.

4.7.2.2.2 Discovery TPS per NfType and Number of Profiles in Response

The following table indicates the discovery TPS per NfType and the number of profiles sent in the response.

Table 4-30 Discovery TPS per NfType and Number of Profiles in Response

Target-Nf Types	TPS (in %)	Absolute Discovery TPS	SLF Queries Posted	Number of Profiles in Response
AMF	4.27	2163	0	6
AUSF	0.38	210	210	6

Table 4-30 (Cont.) Discovery TPS per NfType and Number of Profiles in Response

Target-Nf Types	TPS (in %)	Absolute Discovery TPS	SLF Queries Posted	Number of Profiles in Response
BSF	0.00	0	0	3
CBCF	0.00	0	0	3
CHF	0.077	10	0	3
GMLC	0.00	0	0	6
LMF	0.025	2	0	6
PCF	0.39	192	0	10
SMF	8.14	4144	0	10
UDM	85.4	45735	45735	6
UDR	1.22	651	651	12
SLF	0	0	0	3
NEF	0.026	14.8	0	3
SMSF	0.025	13.9	0	3

4.7.2.3 Test Observations

This section following table provides the details of test duration and TPS used for the benchmarking test.

Table 4-31 Test Observations

Parameter	Values
Test Duration	72 hours
TPS Achieved	5195 (* considered all Ingress and NfStatusNotify traffic)
Ingress incoming connections	24
Ingress outgoing connections	240
Egress incoming connections	22
Egress outgoing connections	120

4.7.2.4 NRF Resource Utilization

The following table describes NRF microservices and their utilization.

Table 4-32 NRF Resource Utilization

NRF Microservices	Number of Pods	Average CPU	Average Memory
nfregistration	2	12.1%	37.1%
nfdiscovery	6	52.2%	50.8%
nfsubscription	2	4.47%	24.4%
nrfauditor	2	0.81%	24.4%
nrfconfiguration	1	0.96%	45.7%
nfaccesstoken	2	0.71%	33.8%
nrfartisan	1	0.10%	24.4%

Table 4-32 (Cont.) NRF Resource Utilization

NRF Microservices	Number of Pods	Average CPU	Average Memory
nrfcachedata	2	7.80%	31.2%
alternate-route	2	0.10%	8.29%
perf-info	2	11.8%	14.1%
appinfo	2	3.21%	25.6%
ingressgateway	3	49.4%	66.5%
egressgateway	3	48.6%	40.9%

> **Note**
>
> If you enable Message Feed feature at Ingress Gateway and Egress Gateway, approximately 33% pod capacity is impacted. The features enabled for this benchmarking testing does not utilize alternate-route service.

4.7.2.5 cnDBTier Services Resource Utilization

The following table provides the observed values of cnDBTier services.

Table 4-33 cnDBTier Services Resource Utilization

Service Name	Pod Count	Average CPU	Memory
SQL (ndbappmysqld)	2	0.48 (12.02%)	3.47 (57.83)
DB (ndbmtd)	4	0.56 (2.8%)	15.87 (79.35)
SQL (ndbmysqld)	2	0.047 (0.8%)	1.92 (19.2%)
MGMT (ndbmgmd)	2	0.01 (0.13%)	0.060 (0.48%)
Monitor Service (db-monitor-svc)	1	0.01 (2.5%)	0.354 (70.8%)
Backup Manager Service (db-backup-manager-svc)	1	0.001 (0.01%)	0.06 (46.875%)
Replication Service	1	0.0 (0.0%)	0.41 (20.5%)

Table 4-34 cnDBTier Statistics

DBTier Statistics	Value
Disk Write operations per second	45.4
Disk Read operations per second	166
Transaction rates on data nodes	102

4.7.2.6 Latency Observations

Table 4-35 Latency Observations

Latency Parameters	99th Percentile (ms)	95th Percentile (ms)	90th Percentile (ms)	50th Percentile (ms)
Average Turnaround time at Ingress Gateway, Discovery Processing, and SLF Egress Transaction	230	189	195	143
Average Turnaround time for Discovery Processing and SLF Egress Transaction	229	191	196	150
Average Turnaround time for SLF Egress Transaction (see Note)	222	196	191	150

Note

Following simulators are used for performance measurement and their induced latency:

- NF simulator to NRF: 50 ms
- NRF to forwarding simulator and forwarding simulator to NRF: 150 ms
- NRF to SLF: 150 ms