

IPWorks 2 Characteristics

USER GUIDE

Copyright

© Ericsson AB 2018. All rights reserved. No part of this document may be reproduced in any form without the written permission of the copyright owner.

Disclaimer

The contents of this document are subject to revision without notice due to continued progress in methodology, design and manufacturing. Ericsson shall have no liability for any error or damage of any kind resulting from the use of this document.

Trademark List

All trademarks mentioned herein are the property of their respective owners. These are shown in the document Trademark Information.



Contents

1	Introduction	1
1.1	Related Information	1
2	System Overview	3
2.1	System Architecture	3
2.2	System Configuration	3
2.2.1	IPWorks Deployment Configuration on CEE	3
2.2.2	IPWorks Deployment Configuration on Native	4
2.3	Key Configuration Parameters	5
2.3.1	Storage Server	5
2.3.2	AAA	6
2.3.3	DNS/ENUM	6
2.3.4	DHCP Server	7
3	System Characteristics	9
3.1	System Capacity	9
3.1.1	Standard Configuration	9
3.1.2	Compact Configuration	11
3.1.3	Single Server Configuration	12
3.2	AAA System Characteristics	13
3.2.1	Traffic Scenarios	13
3.2.2	Processing Capacity	16
3.2.3	Provisioning Capacity	27
3.2.4	Disk Capacity	27
3.2.5	Scalability	28
3.2.6	Quality of Service	28
3.2.7	Dependability	29
3.2.8	Severability	30
3.2.9	Management Impacts	30
3.3	DNS/ENUM System Characteristics	30
3.3.1	Traffic Scenarios	30
3.3.2	Processing Capacity	33
3.3.3	Provisioning Capacity	41
3.3.4	Disk Capacity	41
3.3.5	Scalability	42
3.3.6	Quality of Service	42
3.3.7	Dependability	43
3.3.8	Serviceability	43
3.3.9	Management Impacts	44
3.4	DHCP System Characteristics	44
3.4.1	Traffic Scenarios	44
3.4.2	Processing Capacity	44



3.4.3	Disk Capacity	46
3.4.4	Scalability	46
3.4.5	Quality of Service	46
3.4.6	Dependability	47
3.4.7	Severability	47
3.4.8	Management Impacts	48
4	VNF Life Cycle Management	49
	Reference List	51



1 Introduction

This document describes the performance characteristics for IPWorks system. Dimensioning rules are described in [IPWorks Dimensioning Guideline, Reference \[6\]](#).

Scope

This document presents the characteristics for IPWorks AAA and DNS/ENUM services included in the IPWorks 2 release. The document provides input for IPWorks part in [CANDI tool](#) for dimensioning purpose.

1.1 Related Information

Trademark information, typographic conventions, and definition and explanation of abbreviations and terminology can be found in the following documents:

- Trademark Information, Reference [1]
- Typographic Conventions, Reference [2]
- Glossary of Terms and Acronyms, Reference [3]

Some important definitions:

Table 1 Important Term Definition

Name	Definition
Dependability	Availability performance of the system.
Dimensioned Capacity	Maximum load a system is planned to be exposed. The dimensioned capacity for a system is: 'loadability - safety margin'
Engineered Capacity ⁽¹⁾	Maximum number of TPS at which the system fulfills the performance requirements.
Loadability	The maximum CPU load of the system at which the system fulfills the performance requirements.
Safety margin	Margin must be prepared against bursts or peaks of unexpected traffic. The safety margin is recommended to be 20% of the capacity of the system.
Serviceability	The ability of a service to be obtained.

(1) Throughout the specification, the engineered capacity (processing capacity) is measured with IPv4 transport. It is assumed that with IPv6 transport, the processing capacity may decrease for certain services or traffics but no more than 10% in the worst case.





2 System Overview

This section describes the system overview, including:

- System Architecture, see Section 2.1 on page 3.
- System Configuration, see Section 2.2 on page 3.
- Key Configuration Parameters, see Section 2.3 on page 5.

2.1 System Architecture

IPWorks 2 release includes AAA, DNS, ENUM/DNS, and DHCP services which can be deployed in EPC network or IMS core network. IPWorks 2 is a software only release which is verified on Cloud Execution Environment (CEE) on BSP8100 and HDS8000 (NFVI R4) hardware platform. IPWorks 2 can also be deployed on 3rd party cloud software and Commercial Off-the-Shelf (COTS) hardware. And IPWorks virtual deployment for KVM configuration verified on HP DL380 Gen9 and HP DL380 Gen10.

2.2 System Configuration

This section describes the deployment configurations supported by IPWorks 2.

2.2.1 IPWorks Deployment Configuration on CEE

The characteristics of IPWorks 2 are measured on Ericsson CEE on BSP8100 and HDS8000 hardware platform. SC VM and PL VM are defined for IPWorks 2. In the verification, one VM is dedicated one corresponding computer node using GEP5 or GEP7L board. The specifications of the two VM types for different deployment configurations are listed in Table 2 and Table 3.

Table 2 VNF Configurations for DNS, ENUM, AAA

Deployment Type	VM Node	vCPUs	Memory (GB)	Hard Disk (GB)	vNICs
Standard	SC	14	40	280	3
	PL	14	8	N/A	3
Compact	SC	2	4	75	3
	PL	2	6	N/A	3



Table 3 VNF Configuration for DHCP

Deployment Type	VM Node	vCPUs	Memory (GB)	Hard Disk (GB)	vNICs
Standard Lease Count is 5 M.	SC	8	8	280	3
	PL VM	8	16	N/A	3
Compact Lease Count is 300 K.	SC	2	4	75	3
	PL	2	6	N/A	3

For the Flexible configuration, the HW resources (such as vCPUs, memory, etc.) can be configured in a given range, see Table 4 for details.

Table 4 Flexible Configuration VNF Size Range

VM Node	vCPUs	Memory (GB)	Hard Disk (GB)	vNICs
SC	2 ~ 14	4 ~ 40	75 ~ 280	3
PL	2 ~ 14	6 ~ 16	none	3

IPWorks supports the following deployment configuration types:

- Standard configuration: 2 SC VMs + 2 PL VMs for AAA or DNS/ENUM services. Standard configuration can be scaled out to max 10 PLs depending on different IPWorks service resource requirement. Since SC does not support scalability, the number of SC is always 2. IPWorks DNS and DHCP services do not support scalability.
- Compact configuration: 2 SC VMs + 2 PL VMs for AAA or DNS/ENUM services. Compact configuration does not support scalability.
- Flexible configuration: 2 SC VMs + 2 PL VMs for AAA, DNS/ENUM, or DHCP services. vCPU number is from 2 to 14, PL and SC memory and hard disk usage will be changed according to the required capacity.

eVIP is deployed on the 2 PL VMs to provide single IP interface for AAA, DNS/ENUM, or DHCP service. Traffic load is evenly distributed among the 2 PL VMs through eVIP. Measured capacity is the capacity of the deployed cluster.

2.2.2

IPWorks Deployment Configuration on Native

The characteristics of IPWorks for native configuration are measured on IPWorks VNF on HP DL380 Gen9 and HP DL380 Gen10. IPWorks configuration 4 VMs (2 SCs + 2 PLs) is used in the verification.

The following tables show the IPWorks VNF configuration information of Virtual Machines for DNS, ENUM, AAA, and DHCP in different deployments.



Table 5 Configuration for DNS, ENUM, AAA on Native (G9)

Deployment Type	VM Node	vCPUs	Memory (GB)	Hard Disk (GB)	vNICs
Basic	SC	14	40	280	3
	PL	14	8	N/A	3
Single Server	SC	6	18	75	3
	PL	8	6	N/A	3

Table 6 Configuration for DHCP on Native (G9)

Deployment Type	VM Node	vCPUs	Memory (GB)	Hard Disk (GB)	vNICs
Basic	SC	8	8	280	3
	PL	8	16	N/A	3
Single Server	SC	6	8	75	3
	PL	8	16	N/A	3

Table 7 Configuration for DNS, ENUM, AAA, and DHCP on Native (G10)

Deployment Type	VM Node	vCPUs	Memory (GB)	Hard Disk (GB)	vNICs
Basic	SC	22	80	280	3
	PL	22	32	N/A	3
Single Server	SC	10	40	140	3
	PL	12	16	N/A	3

IPWorks supports the following deployment configuration types:

- Basic configuration: 2 SC VMs + 2 PL VMs for AAA, DNS/ENUM, or DHCP services. Basic configuration can be scaled out to max 10 PLs depending on different IPWorks service resource requirement. Since SC does not support scalability, the number of SC is always 2. IPWorks DNS and DHCP services do not support scalability.
- Single Server configuration: 2 SC VMs + 2 PL VMs for AAA, DNS/ENUM, or DHCP services. Single Server configuration does not support scalability.

2.3 Key Configuration Parameters

2.3.1 Storage Server

Storage Server provides the CLI provisioning interface for DNS, ENUM, AAA, and DHCP. The maximum supported concurrent number of CLI sessions for IPWorks Standard configuration and Basic configuration is 40. In the rest of this specification, the provisioning capacity is measured as sum of 5 concurrent CLI sessions.



2.3.2 AAA

Radius AAA

The performance is dependent on:

- The number of provisioned user records.

All provisioned user records are loaded into memory by the Network Database (NDB) cluster. The more the user records are, the worse the query performance is.

- Whether the session function is enabled or disabled.

If the session function is enabled, the protocol server updates the NDB cluster. The update operation affects the performance significantly.

When the AAA server is used for the **Basic Layered (Front-End) Solution**, the performance is dependent on:

- LDAP connectivity towards the external DB (CUDB). CUDB is serving as the central storage.
- Network latency between AAA-FE and CUDB.

When the AAA server is used for the **Wi-Fi SIM/USIM-based (EAP-AKA, EAP-SIM) Solution**, the performance is dependent on:

- SS7 stack

The communication between the AAA for Wi-Fi and HLR is through SIGTRAN or Signaling Link. IPWorks uses the SS7 stack, and the SS7 stack introduces long latency when it is handling traffic.

EPC AAA

When the AAA server is used for the EPC network, the performance depends on the following condition:

- The number of active sessions (STa/SWm/S6b).

All the active sessions are stored into memory by the Network Database (NDB) cluster. The more the sessions are, the worse the query performance is.

2.3.3 DNS/ENUM

The DNS server performance is highly dependent on the configuration of zones. For the same total number of records, different number of zones results in different performance. Configure less than 1000 zones to achieve the DNS engineering capacity described in this document.



The ENUM Server performance is highly dependent on the number of provisioned ENUM records. The performance decreases when the number of ENUM records increases. Since all the ENUM records are stored in the NDB Cluster, the number of zones configured for the ENUM records does not affect the performance much.

ENUM supports data layered architecture with cache mechanism. The performance of ENUM FE depends on the cache hit rate and network latency to back-end database.

IPWorks DNS can initiate recursive query toward other DNS servers if no answer can be found for the incoming query locally. In such case, the performance of IPWorks DNS highly depends on the following:

- **Network latency** to authoritative DNS servers or other caching DNS server which is configured as a forwarder. It is recommended to set the value of BIND option recursive-clients to 10,000 through CLI for better recursive performance.
- **Cache hit rate.** If the cache size or average TTL of the cached records is small, the cache hit rate decreases, which results in more recursive queries and degraded DNS performance.

If DNS clients are merely end-user devices, it is recommended to set the value of BIND option minimal-response to yes through CLI. The throughput and DNS performance improve when the packet size decreases.

For more information about DNS/ENUM configuration, refer to [Configure DNS and ENUM](#), Reference [4].

2.3.4 DHCP Server

The DHCP server performance is highly dependent on the following configurations:

- **Failover Configuration:** It causes synchronization between the primary and the secondary DHCP server. If the traffic load is high, the quantity of the synchronization information is huge. This has a significant impact on the DHCP server.
- **Lease Time:** The lease time set for the DHCP affects the frequency of the lease renewal. The shorter the lease time is, the higher load on the DHCP server is, which is because clients try renewals more frequently.
- **Pool Number:** The bigger the pool number is, the lower the performance is.
- **Client Class Number:** The bigger the client class number is, the lower the performance is
- **Lease Overlapping:** The more IP pools with IP address are overlapped in, the lower the performance is.





3 System Characteristics

3.1 System Capacity

This section describes the maximally supported AAA sessions and DNS/ENUM records per system configuration. It is possible to enlarge the system capacity by adding more memory to IPWorks VM. However, that also results in prolonged backup/restore time and is out of the scope of IPWorks CVE measurement.

3.1.1 Standard Configuration

This section describes IPWorks system capacity in standard configuration. AAA and DNS/ENUM are described separately because they don't reside in the same standard configuration.

3.1.1.1 AAA

Radius AAA

A maximum of 8 million user records and 4 million concurrent IP sessions for AAA for GPRS, AAA for Wi-Fi, AAA for Fixed Access.

- For CEE and G9, a maximum of 8 million user records and 4 million concurrent IP sessions for AAA for GPRS, AAA for Wi-Fi, AAA for Fixed Access.
- For G10, a maximum of 16 million user records and 8 million concurrent IP sessions for AAA for GPRS, AAA for Wi-Fi, AAA for Fixed Access.

Platform	User Record	IP Session
CEE, Standard	8 million	4 million
G9, Basic	8 million	4 million
G10, Basic	16 million	8 million

EPC AAA

The system capacity for AAA for EPC varies for the SIM and non-SIM services:

— SIM-based service

Subscriber profiles are provisioned to HSS for IPWorks Classic, or provisioned to CUDB for IPWorks Layered.

- For CEE and G9, IPWorks AAA supports maximal **2 million** attached SIM devices and **4 million** AAA sessions (including S6b session).
- For G10, IPWorks AAA supports maximal 4 million attached SIM devices and 8 million AAA sessions (including S6b session)



Platform	SIM Devices	IP Session
CEE, Standard	2 million	4 million
G9, Basic	2 million	4 million
G10, Basic	4 million	8 million

— Non-SIM based service

Subscriber profiles are provisioned to IPWorks Storage Server for IPWorks Classic, or provisioned to CUDB for IPWorks Layered.

- For CEE and G9, IPWorks AAA supports maximal **12 million** provisioned non-SIM device profiles and **6 million** AAA sessions.
- For G10, IPWorks AAA supports maximal 20 million provisioned non-SIM device profiles and 10 million AAA sessions.

Platform	SIM Devices	IP Session
CEE, Standard	12 million	6 million
G9, Basic	12 million	6 million
G10, Basic	20 million	10 million

IPWorks AAA supports maximal **12 million** provisioned non-SIM device profiles and **6 million** AAA sessions.

In case of mixed deployment of both SIM and non-SIM services, the maximal number of supported AAA sessions is between 4 million and 6 million. For G10 native, the maximal number of supported AAA sessions is between 8 million and 10 million.

3.1.1.2

DNS/ENUM

There is no hard limit on the number of DNS records per system configuration. However, it is recommended to configure no more than 100,000 DNS resource records for optimal performance in packet core or IMS networks.

ENUM records are provisioned to IPWorks Storage Server for IPWorks Classic. IPWorks supports maximal **24 million** ENUM records for classic architecture. For IPWorks G10 Basic and Basic specialized configuration, it can support 48 million ENUM records.

ENUM records are provisioned to CUDB for IPWorks Layered. The limit of supported number of ENUM records lies on CUDB. However, IPWorks ENUM cannot handle the generated queries per second by ENUM clients beyond its processing capacity.



3.1.2 Compact Configuration

This section describes IPWorks system capacity in compact configuration. AAA and DNS/ENUM are described separately because they don't reside in the same compact configuration.

3.1.2.1 AAA

Radius AAA

Compact configuration: A maximum of 0.4 million user records and 0.2 million concurrent IP sessions for AAA for GPRS, AAA for Wi-Fi, AAA for Fixed Access.

EPC AAA

The system capacity for AAA varies based on the SIM and non-SIM services:

— SIM-based service

Subscriber profiles are provisioned to HSS for IPWorks Classic, or provisioned to CUDB for IPWorks Layered.

IPWorks AAA supports maximal **0.2 million** attached SIM devices and **0.4 million** AAA sessions (including S6b session).

— Non-SIM based service

Subscriber profiles are provisioned to IPWorks Storage Server for IPWorks Classic. IPWorks Layered does not support non-SIM based service.

IPWorks AAA supports maximal **0.4 million** provisioned non-SIM device profiles and **0.2 million** AAA sessions.

In case of mixed deployment of both SIM and non-SIM services, the maximal number of supported AAA sessions is between **0.2 million** and **0.4 million** active attached subscribers.

3.1.2.2 DNS/ENUM

There is no hard limit on the number of DNS records per system configuration. However, it is recommended to configure no more than 40,000 DNS resource records for optimal performance in packet core or IMS networks.

ENUM records are provisioned to IPWorks Storage Server for IPWorks Classic. IPWorks supports maximal **0.6 million** ENUM records for classic architecture.

ENUM records are provisioned to CUDB for IPWorks Layered. The limit of supported number of ENUM records lies on CUDB. However, IPWorks ENUM cannot handle the generated queries per second by ENUM clients beyond its processing capacity.



3.1.3 Single Server Configuration

This section describes IPWorks system capacity in Single Server configuration. AAA and DNS/ENUM are described separately because they don't reside in the same configuration.

3.1.3.1 AAA

Radius AAA

A maximum of 4 million user records and 2 million concurrent IP sessions for AAA for GPRS, AAA for Wi-Fi, AAA for Fixed Access.

EPC AAA

The system capacity for AAA for EPC varies based on the SIM and non-SIM services:

— SIM-based service

Subscriber profiles are provisioned to HSS for IPWorks Classic, or provisioned to CUDB for IPWorks Layered.

IPWorks AAA supports maximal **1 million** attached SIM devices and **2 million** AAA sessions (including S6b session).

— Non-SIM based service

Subscriber profiles are provisioned to IPWorks Storage Server for IPWorks Classic, or provisioned to CUDB for IPWorks Layered.

IPWorks AAA supports maximal **6 million** provisioned non-SIM device profiles and **3 million** AAA sessions.

In case of mixed deployment of both SIM and non-SIM services, the maximal number of supported AAA sessions is between **2 million** and **3 million**.

3.1.3.2 DNS/ENUM

There is no hard limit on the number of DNS records per system configuration. However, it is recommended to configure no more than 50,000 DNS resource records for optimal performance in packet core or IMS networks.

ENUM records are provisioned to IPWorks Storage Server for IPWorks Classic. IPWorks supports maximal **12 million** ENUM records for classic architecture.

ENUM records are provisioned to CUDB for IPWorks Layered. The limit of supported number of ENUM records lies on CUDB. However, IPWorks ENUM cannot handle the generated queries per second by ENUM clients beyond its processing capacity.



3.2 AAA System Characteristics

3.2.1 Traffic Scenarios

3.2.1.1 Use-cases of AAA for GPRS Classic (Monolithic) Server

An AAA Basic server connects a Gateway GPRS Support Node (GGSN-MPG) and an external Public Data Network, enabling the Mobile Stations to exchange IP packets with the external network. IPWorks AAA services are used to authenticate a user, provide subscriber information, and provide accounting to activate the backend services. The Classic (monolithic) scenario is the legacy deployment of AAA server where the AAA subscriber data are provisioned in the IPWorks storage server and replicated to the AAA server data node.

An AAA Basic server can also be used for IP address allocation. The IP address pool selection can be based on user or GGSN.

An AAA Basic server supports the following scenarios, which are defined in 3GPP TS 29.061:

- Authentication and Accounting for IP PDP type
- Authentication and Accounting for PPP PDP type
- Accounting update
- AAA-Initiated PDP context termination

3.2.1.2 Use-cases of AAA for GPRS Layered (Front-End) Server

AAA for GPRS Front-End is the Data Layered Architecture which IPWorks AAA is interworking with external DB (CUDB). CUDB works as the back-end database. Use-cases of AAA for GPRS Front-End server is the same as that of the Classic (monolithic) (see Section 3.1.1.1 on page 9), the only difference is the place where the AAA subscriber data are stored:

- AAA subscriber-related data about authentication and authorization is stored in the central storage CUDB.
- AAA configuration parameter, Session management related data, and Accounting management-related data are still in the IPWorks AAA server.

3.2.1.3 Use-cases of AAA for Wi-Fi SIM-based (EAP-AKA and EAP-SIM) Classic/Layered Server

The Wi-Fi technology is built on IEEE 802.11 standards. A Wi-Fi enabled device such as a personal computer or smart phone can connect to any IP-based network. IPWorks AAA for Wi-Fi is used to offer mobile users equivalent security levels to GSM, GPRS, and WCDMA network when the user accesses a WLAN network. Access authentication to the WLAN is performed using the authentication vectors

stored in the HLR for GSM/GPRS/WCDMA authentication. The authentication vectors are retrieved by the Mobile Application Part (MAP).

In the Wi-Fi offload scenario, IPWorks AAA for Wi-Fi enables 3GPP devices to connect to Internet via the Wi-Fi hotspot and residential Wi-Fi gateways, and access the 3GPP services via Wi-Fi access. IPWorks AAA in this scenario is used for SIM/USIM based authentication by using this authentication mechanism in the Wi-Fi Access scenario, which minimizes the interaction from an end-user perspective and offload the traffic data to the Wi-Fi network.

IPWorks AAA implements W_a , W_m reference point, and D'/Gr' reference point according to 3GPP TS 23.234. IPWorks supports the following AAA for Wi-Fi functions:

- EAP-SIM and EAP-AKA authentication mechanism in the WLAN AAA Solution.
- EAP-AKA/SIM full authentication and fast re-authentication procedures.
- EAP-SIM authentication for a 3G user. In this case, if AAA receives a quintuplet vector from HLR within EAP-SIM authentication, it translates it to triplet vectors.
- During the EAP-AKA/SIM authentication, as the authentication vectors are stored in the HLR Server, IPWorks AAA supports to fetch the authentication vectors from HLR via SS7 MAP protocol.
- WLAN Access authorization for the UE based on the subscriber data received from HLR. If the user has a contract for accessing the WLAN Network, HLR returns a flag in Operator Determined Barring HPLMP data.
- Receiving notifications from the HLR about the user status change and deciding whether to send Disconnect-Request according to the notifications contents to terminate the accounting session automatically.
- Accounting functions in WLAN scenario.

3.2.1.4 Use-cases of AAA for Wi-Fi Server for Trusted Wi-Fi Access (Classic/Layered)

When UE attaches from trusted Wi-Fi access, the subscriber profile is retrieved from HLR/HSS. IPWorks AAA authenticates UE based on the subscriber profile. Once authenticated, UE can be authorized with access to EPC via S2a GTP tunnel, or offload the traffic to Internet via Wi-Fi GW (Non-Seamless WLAN Offload, or NSWOW).

For SIM/USIM based UE, EAP-SIM/AKA is used for authentication:

Subscriber profile and authentication vectors are retrieved from HLR via Gr' interface. If authentication succeeds, AAA sends offload-Indication back to Wi-Fi GW in Access-Accept message indicating authorization type. If access to EPC is granted, parameters to set up S2a GTP tunnel are also returned.



3.2.1.5 Use-cases of AAA for Wi-Fi (Classic/Layered) Server Supporting HSS Integration

AAA for Wi-Fi server supports HSS integration by retrieving authentication vectors and subscriber profile from HSS in case of LTE/4G subscriber without changing the existing capability of Wi-Fi Access Network (for example, using RADIUS for EAP authentication).

3.2.1.6 Use-cases of AAA for EPC Server for Trusted/Untrusted Non-3GPP Access

The AAA is a Diameter-based server used for trusted/untrusted non-3GPP access scenarios, which are specified in 3GPP TS 23.402 and 3GPP TS 29.273. IPWorks 3GPP AAA supports the access authentication and authorization from the non-3GPP access networks, such as CDMA2000, WiMAX and Wi-Fi, as well as mobility between these networks and LTE access.

- For trusted non-3GPP access, STa with EAP-AKA' is used to authenticate and authorize SIM-based UE.
- For untrusted non-3GPP access, SWm with EAP-AKA is used to authenticate and authorize SIM-based UE.
- For both cases, authentication vectors and user profiles are fetched from HSS through SWx interface.

IPWorks AAA also supports authentication and authorization for Multi-Device solution (Classic/Layered). Proprietary SWm interface with EAP-TLS is used in such scenario. This function enables non-SIM-based Wi-Fi calling.

AAA supports STa/SWm/S6b, and SWx interfaces as per 3GPP TS 23.402 and 29.273, and the proprietary SWm interface. Following functions/procedures are supported:

- EAP-AKA' as the authentication mechanism for SIM-based UE which is attached to trusted non-3GPP access (STa Procedure).
- EAP-AKA as the authentication mechanism for SIM-based UE which is attached to untrusted non-3GPP access (SWm Procedure).
- EAP-TLS as the authentication mechanism for non-SIM devices in Multi-Device solution which are attached to untrusted non-3GPP access (SWm Procedure).
- Fetching the authentication vectors and user profiles from HSS through SWx interface during EAP-AKA'/EAP-AKA authentication.
- Updating PDN-GW information into the HSS and downloading the QoS profiles and possibly the mobility-related parameters for non-3GPP accesses.

3.2.1.7 Use-cases of AAA for Fixed Access

IPWorks AAA supports user Authentication, Authorization, and Accounting for fixed access network based on RFC and BBF standards. eSAPC can be integrated for policy control in the fixed access network scenario.

The following functionalities are supported by IPWorks AAA:

- Authentication, Authorization and Accounting for PPPoX users
- Accounting update
- Accounting mediation, IPWorks AAA supports forwarding accounting message to multiple target server groups based on configuration, and it can be used by regulatory service.
- AAA-Initiated user termination

AAA for Fixed Access uses the PAP/CHAP authentication. For the information of processing capacity, see Section 3.2.2.1 AAA for GPRS Classic (Monolithic) on page 16.

3.2.2 Processing Capacity

This section describes the processing capacity of IPWorks with deployment of standard system configuration given in Section 2.2.

AAA supports TCP or SCTP as transport protocol. The performance of AAA is the same for TCP and SCTP.

With eVIP deployed, the processing capacity of the system is the sum of the 2 PL VMs. If 1 PL VM fails, the processing capacity of the system decreases by 50%. The dimensioned capacity is recommended to be set as 80% of the system engineered capacity. The safety margin to use is to be decided by the customer. Depending on the traffic scenario, the system capacity varies.

3.2.2.1 AAA for GPRS Classic (Monolithic)

An AAA for GPRS Classic server handles the authentication, authorization and accounting functionality for the remote clients or systems. When the system receives a request from a client, it checks the received information, does the PAP or CHAP authentication for the specified user and assigns a list of reply attributes according to a pre-configured authorization reply list. If an accounting message comes, AAA server will record the attributes as CSV files based on the configured format.

AAA for GPRS Classic is the basic and traditional deployment of AAA Server where both application logic and data reside in the AAA node. The system must be configured with a number of users. And it is recommended to attach a sort of users to the specified user groups to ease the complexity for management.



The authorization is achieved through a checklist for request and a reply list for what to be sent back. The checklist and reply list can be applied to both a single user and a user group.

The engineered capacity is obtained based on 99% reply rate in 10ms, except the mixed traffic with session enabled is obtained based on 99% reply rate in 15ms. It is based on the maximum capacity that the system could achieve. The safety margin is recommended to be 80% of the system engineered capacity. The safety margin to use is to be decided by the customer. Depending on the traffic scenario, the system capacity varies.

The measurements taken in the following subsections are based on IPWorks standard configuration, and with the following provisioned user data:

- 8 Million Records, with 40 groups and 200,000 user records per group

For Compact, with the following provisioned user data:

- 0.4 Million Records, with 20 groups and 20,000 user records per group

3.2.2.1.1 Accounting Only

Table 8 shows the performance of accounting traffic only:

Table 8 Accounting Traffic only Performance

Platform	Engineered Capacity (QPS)	Dimensioned Capacity (QPS)	Max. Latency (ms)	Min. Latency (ms)	Avg. Latency (ms)	Reply Rate
CEE Standard	20,000	16,000	439	2.5	3	99.99%
G10 Basic	30,000	24,000	474	1	5	99.99%

The engineered capacity with IPv6 transport interface is the same as the one with IPv4 transport interface.

3.2.2.1.2 Authentication + Authorization + Accounting with Session

The traffic mixes are:

- 25% Access-Request
- 25% Accounting-Start
- 25% Accounting-Update
- 25% Accounting-Stop

Table 9 shows the performance of AAA for GPRS with PAP/CHAP authentication, authorization, accounting and session enabled:



Table 9 AAA for GPRS PAP/CHAP Mixed Traffic with Session

Platform	Engineered Capacity (QPS)	Dimensioned Capacity (QPS)	Max. Latency (ms)	Min. Latency (ms)	Avg. Latency (ms)	Reply Rate
CEE Standard	14,000	11,200	434	2.3	10.8	99.99%
G9 Basic	16,000	12,800	-	-	-	99.99%
Compact	670	536	-	-	-	99.99%
G10 Basic	28,000	22,400	227	1.5	7.5	99.99%
G10 Single	22,000	17,600	264	1.5	6.2	99.99%

3.2.2.1.3

Authentication and Authorization Proxy

Table 10 shows the performance of authentication and authorization proxy messages:

Table 10 Authentication and Authorization Proxy Performance

Platform	Engineered Capacity (QPS)	Dimensioned Capacity (QPS)	Max. Latency (ms)	Min. Latency (ms)	Avg. Latency (ms)	Reply Rate
CEE Standard	20,000	16,000	5,460	0	119	99.99%
G9 Basic	20,000	16,000	-	-	-	99.99%
G10 Basic	25,000	20,000	3,036	0.8	18	99.99%

3.2.2.1.4

Accounting Proxy

IPWorks AAA supports accounting proxy over Radius protocol. When configured, AAA can proxy/forward Radius accounting messages (including start, stop and interim update) to one or multiple destination nodes/groups based on pre-configured criteria, such as AVPs included in the Accounting messages. During the measurement, 9 AVPs are configured. Note that with increased number of configured AVPs, the performance of AAA proxy might drop.

The engineered capacity is obtained based on the maximum capacity that the system can achieve. The safety margin is recommended to be 80% of the system engineered capacity. The safety margin to use is to be decided by the customer. Depending on the traffic scenarios and subscriber profile configuration, the system capacity varies.

The performance of AAA accounting proxy over Radius is measured with different number of destination nodes, the measurement results are outlined in Table 11.



Table 11 AAA Accounting Proxy (CEE, Standard)

Scenario	Engineered Capacity (QPS)	Dimensioned Capacity (QPS)	Max. Latency (ms)	Min. Latency (ms)	Avg. Latency (ms)	Reply Rate
1 target nodes	7,000	5,600	2,482	1.0	1.8	99.99%
15 target nodes (5 groups * 3 servers)	1,000	800	78	2.0	5.8	99.99%

Table 12 AAA Accounting Proxy (G10 Basic)

Scenario	Engineered Capacity (QPS)	Dimensioned Capacity (QPS)	Max. Latency (ms)	Min. Latency (ms)	Avg. Latency (ms)	Reply Rate
1 target nodes	20,000	16,000	3,056	0.8	18	99.99%

3.2.2.1.5

AAA IP Allocation with Sessions with Accounting to Database Traffic Performance

The traffic mixes are:

- 25% Access-Request
- 25% Accounting-Start
- 25% Accounting-Update
- 25% Accounting-Stop

The performance is measured with protocol servers working at active-standby mode, and with scenario that assigns IP from a specific IP address pool.

Table 13 shows the performance of AAA messages with session:

Table 13 AAA IP Allocation with Accounting with Session

Platform	Engineered Capacity (QPS)	Dimensioned Capacity (QPS)	Max. Latency (ms)	Min. Latency (ms)	Avg. Latency (ms)	Reply Rate
CEE Standard	7,000	5,600	197	2.2	8.0	99.99%
G10 Basic	20,000	16,000	86	1.8	10	99.99%

3.2.2.2

AAA for GPRS Layered (Front-End)

The AAA for GPRS Front-End server provides the same functionality as the Classic Basic AAA does.

In the AAA for GPRS Front-End deployment, AAA acts as a data-less front end with only the application logic, the AAA user data are provisioned in an external backend database, for example, CUDb, which can be accessed by the AAA FE

through the LDAP interface. However, the AAA session data and accounting data are still stored in the local MySQL NDB Cluster.

The AAA users stored in the external backend database are provisioned by EDA. And it is still recommended to attach a sort of users to the specified user groups to ease the complexity for management.

The authorization is still achieved through a checklist for request and a reply list for what to be sent back. The checklist and reply list can be applied to both a single user and a user group, which retrieves by AAA server from backend database through the LDAP interface.

In the Classic Basic AAA server, the data are stored and loaded into the memory by the MySQL NDB Cluster. Comparing to the Classic Basic AAA server, Front-End the basic server queries the user data from the external DB (CUDB) by the LDAP interface. The traffic performance of the LDAP and External DB (CUDB) has significant impact on the AAA server performance. Generally, the local data access has less delay and latency than the external DB, so the AAA for GPRS Front-End performance is expected a decrement.

The engineered capacity is obtained based on 99% reply rate in 10ms, except the mixed traffic with session enabled is obtained base on 99% reply rate in 10 ms. The engineered capacity is obtained based on the maximum capacity that the system could achieve. The safety margin is recommended to be 80% of the system engineered capacity. The safety margin to use is to be decided by the customer. Depending on the traffic scenario, the system capacity varies.

The performance of proxy and account is similar to that of the AAA Basic Class.

3.2.2.2.1 Authentication + Authorization + Accounting with Session

The traffic mixes are:

- 25% Access-Request
- 25% Accounting-Start
- 25% Accounting-Update
- 25% Accounting-Stop

There is little performance difference between generating accounting information directly to CSV file and storing the accounting information in the MySQL NDB cluster.

Table 14 shows the performance of AAA for GPRS Front-End with PAP/CHAP authentication, authorization, accounting and session enabled.



Table 14 AAA for GPRS Front-End PAP/CHAP Mixed Traffic with Session (OWD is 0 ms)

Platform	Engineered Capacity (QPS)	Dimensioned Capacity (QPS)	Max. Latency (ms)	Min. Latency (ms)	Avg. Latency (ms)	Reply Rate
CEE Standard	11,000	8,800	116	3.1	23	99.99%
G9 Single Server	4,000	3,200	809	0.7	28.3	99.99%
G10 Basic	18,000	14,400	210	1.5	5	99.99%
G10 Single	16,000	12,800	46	1.6	5	99.99%

3.2.2.3

AAA for Wi-Fi SIM-based

The performance data in this session are measured with the condition that IPWorks has cached part of user profiles in the NDB database. Therefore, IPWorks does not need to retrieve user profiles from HLR all the time. The cache mechanism is enabled in IPWorks by default and is not configurable. When the cache is clear, IPWorks needs to request user profiles from HLR, and the engineered capacity is lower, especially in TDM scenarios.

3.2.2.3.1

EAP-AKA/EAP-SIM Full Authentication + Accounting, SIGTRAN

Upon reception of an EAP-Request, AAA for Wi-Fi is triggered to perform EAP-AKA/EAP-SIM authentication procedure. The authentication vectors are fetched from HLR via D'/Gr' interface. Once the end-user passes the authentication, the access right of 3GPP network is granted. There is no user data stored in MySQL NDB cluster locally. AAA session data and accounting data are still stored in MySQL NDB Cluster locally.

The engineered capacity is obtained based 99% reply rate in 400 ms. The safety margin is recommended to be less than 80% of the system engineered capacity. The safety margin to use is to be decided by customers. Depending on the traffic scenarios, the system capacity varies.

This scenario, which is based on SIGTRAN, includes 1 EAP full authentication and 1 accounting procedure. The performance of EAP-AKA and EAP-SIM is the same.

Table 15 shows the performance of EAP full authentication session on SIGTRAN based on the detailed test result:

Table 15 EAP-AKA/EAP-SIM Full Authentication Session Performance on SIGTRAN

Platform	Engineered Capacity (TPS)	Dimensioned Capacity (TPS)	Max. Latency (ms)	Min. Latency (ms)	Avg. Latency (ms)	Reply Rate
CEE Standard	2,000	1,600	412	0.6	6.2	99.99%
G10 Basic	3,000	2,400	229	0.4	5	99.99%

3.2.2.3.2 1 (Full Authentication + Accounting) + 1 (Fast Re-Authentication + Accounting), SIGTRAN

Upon the reception of an EAP-Request, AAA for Wi-Fi is triggered to perform the EAP-AKA/EAP-SIM authentication procedure. The authentication vectors are fetched from HLR through the D'/Gr' interface. Once the end user passes the authentication, the access right of 3GPP network is granted.

This scenario, which is based on SIGTRAN, includes one EAP full authentication, one Fast re-authentication, and one accounting procedure. The performance of EAP-AKA and EAP-SIM is the same.

Table 16 shows the performance of EAP full authentication session on SIGTRAN based on the detailed test result:

Table 16 EAP-AKA/EAP-SIM Fast Re-authentication Session Performance on SIGTRAN (CEE)

Platform	Engineered Capacity (TPS)	Dimensioned Capacity (TPS)	Max. Latency (ms)	Min. Latency (ms)	Avg. Latency (ms)	Reply Rate
CEE Standard	1,200	960	314	0.6	6.1	99.99%
G10 Basic	1,700	1,260	285	0.5	4.9	99.99%

3.2.2.4 AAA for Wi-Fi for Trusted Wi-Fi Access

When UE attaches from trusted Wi-Fi access, EAP-Request is transported over EAPoL between UE and Wi-Fi AP, and over RADIUS from Wi-Fi GW to IPWorks AAA. Based on the subscriber profile, IPWorks AAA grants UE access to EPC via S2a GTP tunnel, or ask Wi-Fi GW to offload the traffic to Internet (Non-Seamless WLAN Offload, or NSWO).

For SIM/USIM based UE: The subscriber profile and authentication vectors are retrieved from HLR via Gr' interface. From the prefix of APN in subscriber profile, IPWorks AAA makes authorization decision. Offload-Indication is sent back to Wi-Fi GW in Access-Accept message indicating authorization type. If access to EPC is granted, parameters to set up S2a GTP tunnel are also returned.

3.2.2.4.1 EAP-SIM + Accounting with S2a to GTP Tunnel Traffic Performance

In this case, SIM-based UE attaches to trusted Wi-Fi access and is authenticated by IPWorks AAA with EAP-SIM. Accounting is enabled. Based on the subscriber profile retrieved from HLR via Gr' interface, IPWorks AAA grants the UE access to EPC via S2a and sends related AVPs to Wi-Fi GW for S2a GTP tunnel setup.

Table 17 shows the performance of EAP-SIM authentication with accounting, and authorization of S2a GTP tunnel to P-GW. For the same scenario, EAP-AKA has the same performance as that of EAP-SIM.



Table 17 EAP-SIM with Accounting, Authorization of S2a

Platform	Engineered Capacity (TPS)	Dimensioned Capacity (TPS)	Max. Latency (ms)	Min. Latency (ms)	Avg. Latency (ms)	Reply Rate
CEE Standard	1,400	1,120	300	0.61	6.8	99.99%
Single Server	1,300	1,040	282	0.03	15.3	99.99%
G10 Basic	1,700	1,360	253	0.19	3.6	99.99%
G10 Single	1,500	1,200	298	0.3	4.5	99.99%

3.2.2.5

AAA for Wi-Fi with HSS Integration

The solution of HSS integration supports the retrievable of authentication vectors and the subscriber profile from HSS in case of LTE/4G subscriber without changing the existing capability of Wi-Fi Access Network (for example, using RADIUS for EAP authentication).

- For EAP-AKA authentication (3G/4G user), IPWorks AAA communicates with HSS through SWx Diameter interface. If no subscription in HSS, IPWorks performs a fallback to HLR for retrieval of user authentication vector and 3G user profile data.
- For EAP-AKA' authentication (4G user), only HSS is selected for retrieving user authentication vector and profile.
- For EAP-SIM authentication (2G/3G user), only HLR is selected for retrieving user authentication vector and profile.

3.2.2.5.1

EAP-AKA Full Authentication with HSS Integration

When UE attaches to Wi-Fi access network, EAP-AKA procedure is executed over RADIUS from UE to IPWorks AAA via Wi-Fi GW for authentication. IPWorks AAA firstly tries to retrieve authentication vectors and the subscriber profile from HSS via SWx interface, and continues authentication when they are found in HSS.

Table 18 shows the performance test results of EAP-AKA Full Authentication with HSS Integration.

Table 18 EAP-AKA Full Authentication with HSS Integration

Platform	Engineered Capacity (TPS)	Dimensioned Capacity (TPS)	Max. Latency (ms)	Min. Latency (ms)	Avg. Latency (ms)	Reply Rate
CEE Standard	1,000	800	379	2.4	30.9	99.99%
G10 Basic	2,000	1,600	181	1.3	11.1	99.99%



3.2.2.5.2 EAP-AKA' Full Authentication with HSS Integration

When UE attaches to Wi-Fi access network, EAP-AKA' procedure is executed over RADIUS from UE to IPWorks AAA via Wi-Fi GW for authentication. With EAP-AKA', IPWorks AAA only tries to retrieve authentication vectors and subscriber profile from HSS via SWx interface.

Table 19 shows the performance test results of EAP-AKA' Full Authentication with HSS Integration.

Table 19 EAP-AKA' Full Authentication with HSS Integration

	Engineered Capacity (TPS)	Dimensioned Capacity (TPS)	Max. Latency (ms)	Min. Latency (ms)	Avg. Latency (ms)	Reply Rate
CEE Standard	1,000	800	197	2.55	13.8	99.99%
G10 Basic	2,000	1,600	335	0.8	14.7	99.99%

3.2.2.6 AAA for EPC Server for Trusted/Untrusted Non-3GPP Access

3.2.2.6.1 STa Full Authentication with EAP-AKA' and S6b Authorization

When an EAP-Response/Identity message is received, AAA is triggered to perform AKA (EAP-AKA', defined RFC 5448) authentication and authorization procedure. The authentication and authorization is based on Diameter protocol. The authentication vectors and user profile are downloaded from HSS through SWx interface.

The update location procedure is triggered by PDN GW when the UE attaches to the EPC using the S2a reference point in the PMIPv6 mode. The AAA server updates the PDN GW address information to HSS and the AAA server also retrieves and updates the mobility-related parameters to authorize the PDN GW. The procedure is described in S6b interface.

The engineered capacity is obtained based on the maximum capacity that the system can achieve. The safety margin is recommended to be 80% of the system engineered capacity. The safety margin to use is to be decided by the customer. Depending on the traffic scenarios and subscriber profile configuration, the system capacity varies.

Table 20 shows the performance of the AKA' full authentication with S6b authorization:

Table 20 STa Full Authentication with S6b Performance

Platform	Engineered Capacity (TPS)	Dimensioned Capacity (TPS)	Reply Rate
CEE Standard	1200	960	99.99%
G9 Basic	1500	1200	99.99%



Platform	Engineered Capacity (TPS)	Dimensioned Capacity (TPS)	Reply Rate
Single Server	780	624	99.99%
G10 Basic	2500	2000	99.99%
G10 Single	2000	1600	99.99%

3.2.2.6.2

SWm Full Authentication with EAP-AKA and S6b Authorization

When UE attaches to untrusted Wi-Fi access, an IPSec tunnel is established between UE and ePDG. ePDG sends EAP-AKA authentication messages to AAA through SWm interface. The authentication vectors and user profile are downloaded from HSS through the SWx interface.

If the authentication is successful, ePDG sets up S2b GTP tunnel towards PDN GW for UE data transportation. The AAA server updates the PDN GW address information to HSS and the AAA server also retrieves and updates the mobility-related parameters to authorizing the PDN GW. The procedure is described in S6b interface.

The engineered capacity is obtained based on the maximum capacity that the system can achieve. The safety margin is recommended to be 80% of the system engineered capacity. The safety margin to use is to be decided by the customer. Depending on the traffic scenarios and subscriber profile configuration, the system capacity varies.

Table 21 shows the performance results of SWm full authentication with EAP-AKA and S6b authorization.

Table 21 SWm Full Authentication with S6b Performance

Platform	Engineered Capacity (TPS)	Dimensioned Capacity (TPS)	Reply Rate
CEE Standard	1200	960	99.99%
G9 Basic	1500	1200	99.99%
IMSI Masking Standard	950	760	99.99%
IMSI Masking, G9 Single Server	780	624	99.99%
G10 Basic	2500	2000	99.99%

3.2.2.6.3

SWm Full Authentication with EAP-TLS

For UE without SIM/USIM, AAA server authenticates the UE via EAP-TLS. Non-SIM UE must obtain a valid certificate from trusted Certificate Authority (CA) before connecting to EPC. Subscriber profiles are pre-provisioned to IPWorks storage server.

When UE attaches to untrusted Wi-Fi access and establishes IPSec tunnel with ePDG, UE initiates access request to AAA through ePDG. IPWorks AAA authenticates the UE via EAP-TLS over SWm interface between ePDG and AAA.



If authentication succeeds, ePDG sets up S2b tunnel towards PDN GW and UE has access to the packet core network.

The engineered capacity is obtained based on the maximum capacity that the system can achieve. The safety margin is recommended to be 80% of the system engineered capacity. The safety margin to use is to be decided by the customer. Depending on the traffic scenarios and subscriber profile configuration, the system capacity varies.

Table 22 shows the performance results.

Table 22 SWm Full Authentication with EAP-TLS Performance

Platform	Engineered Capacity (TPS)	Dimensioned Capacity (TPS)	Reply Rate
CEE Standard	1500	1200	99.99%
G9 Basic	1800	1440	99.99%
CEE, IMSI masking	805	565	99.99%
G9 Single Server	900	720	99.99%
G10 Basic	3000	2400	99.99%
G10 Single	2500	2000	99.99%

To support 20M provisioned non-SIM devices and 10.5M active SWm sessions capacity, 2 PLs need to be scaled out to 4 PLs.

Table 23 SWm Full Authentication with EAP-TLS

Scenario	Engineered Capacity (TPS)	Dimensioned Capacity (TPS)	Reply Rate
20M users	3600	2880	99.97%
Compact	240	192	99.99%

3.2.2.6.4

SWm Full Authentication with EAP-TLS (Front-End)

The EPC SWm Full Authentication with EAP-TLS (Front-End) provides the same functionality as the Classic AAA EPC does. In the Front-End deployment, AAA acts as a dataless front end with only application logic, the AAA user profile data are provisioned in an external backend database, for example, CUDB, which can be accessed by the AAA-FE through LDAP interface.

Table 24 shows the performance results.

Table 24 SWm Full Authentication with EAP-TLS

Platform	Engineered Capacity (TPS)	Dimensioned Capacity (TPS)	Reply Rate
CEE Standard	1450	1150	99.99%
G9 Basic	1500	1200	99.99%
G9 Single Server	900	720	99.99%



Platform	Engineered Capacity (TPS)	Dimensioned Capacity (TPS)	Reply Rate
G10 Basic	1600	1280	99.99%
G10 Single	1400	1120	99.99%

3.2.3 Provisioning Capacity

AAA for EPC supports both SIM and non-SIM subscribers. For SIM subscribers, their profiles are not provisioned to IPWorks. Only the profiles of non-SIM subscribers are provisioned to IPWorks.

The provisioning performance of update, modify, and delete operations on 3M non-SIM users is collected while 3M active sessions are kept in memory. The performance results are outlined in Table 25 .

Table 25 Provisioning Capacity (CEE)

Case	Total (Rec/S)
Create	180
Modify	90
Delete	90

AAA for Radius. The provisioning performance of update, modify, and delete operations on 200K users with IMSI is collected while 8M user in DB. The performance results are outlined in Table 26.

Table 26 Provisioning Capacity (CEE)

Case	Total (Rec/S)
Create	140
Modify	85
Delete	85

3.2.4 Disk Capacity

3.2.4.1 Storage Server

IPWorks storage server runs in SC VM. Disk capacity usage in SC VM for AAA is outlined in Table 27.

Table 27 Disk Space Used in SC VM for AAA

Scenario	Total (GiB)	Used (GiB)
12 million non-SIM devices and 6 million sessions	280	81



3.2.4.2 Protocol Server

IPWorks AAA service runs on PL VM, and has access to NDB in SC VM for provisioned records or IP sessions. There is no disk use on PL board.

3.2.5 Scalability

IPWorks VNF standard configuration can be scaled out to max 10 PLs depending on different IPWorks service resource requirement. For AAA, the max supported PL number after scaling-out operation and corresponding critical resource are outlined in Table 28.

Table 28 Max Supported PL Number and Critical Resource for Scalability (AAA)

	AAA for GPRS	AAA for GPRS Front-End	AAA for EPC	AAA PKI	AAA PKI Front End
Critical Resource	SC Disk I/O ⁽¹⁾	SC Disk I/O ⁽¹⁾	SC Memory ⁽²⁾	SC Memory ⁽²⁾	N/A
Number of SC	2	2	2	2	2
Max Number of PL after scale-out	3	3	4	4	10

(1) For AAA GPRS, accounting on PL generates disk I/O operation on SC, which limits the number of PL that can be scaled-out.

(2) Memory on SC limits the max number of subscribers IPWorks VNF can support, which in turn limits the max number of PL needed to process request generated correspondingly.

IPWorks VNF processing capacity in terms of TPS is increased linearly by adding PL until PL number reaches the max supported number.

For AAA-FE (Radius and PKI) solution, the max supported subscribers depends on CUDB capacity.

The processing capacity of single PL is half of the processing capacity of IPWorks VNF outlined in Section 3.2.2 Processing Capacity on page 16.

3.2.6 Quality of Service

3.2.6.1 Delay

When AAA SIM based is working, the server needs to contact HSS to get the user profile. The latency of AAA depends on the reply latency of HSS. If the HSS is deployed as HSS FE, extra latency between HSS FE and CUDB is introduced.

3.2.6.2 Quality of Service Impacts on Overload Situations

IPWorks AAA server has overload protection mechanism over TCP. AAA running over SCTP does not have overload protection mechanism.



At a load above the engineered capacity, IPWorks AAA server can handle the traffic with some reduced capacity, and it handles (that is, maintains and finishes) all established sessions as under normal load.

— System Robustness

IPWorks AAA server can sustain constant overload situations and return to normal operations once traffic has reached normal levels.

— Performance degradation during overload situation

- At an offered load of 150%, there is at least a throughput of 90%.
- At an offered load of 300%, there is at least a throughput of 60%.
- At an offered load of 500%, there is at least a throughput of 30%.
- At an offered load of 1000%, the throughput is higher than 0%.

3.2.7 Dependability

3.2.7.1 Reliability

Reliability is the probability that the IPWorks can perform a function over a given period. Within this period, the system must be available.

Stability test is performed in the IPWorks system with mixed AAA traffic for 48 hours with no issue detected.

3.2.7.2 Availability

Availability indicates the probability that the system is in a state ready to perform a function.

The system is not available at certain conditions, such as:

- Restart
- Cold backup and restore

The total restart of a protocol server takes about 4 minutes. If a protocol server is expected to have one failure with restart once a year, which means availability per protocol server is:

$$\left(1 - \frac{240}{31536000}\right) * 100 = 99,99924$$

In a system with 2 protocol servers, for the functionality of protocol servers to be unavailable, both servers must be down at the same time. The availability of the protocol-severs functionality in a system with 2 protocol servers is:



$$\left(1 - \left(\frac{240}{31536000}\right)^2\right) * 100 = 99,99999$$

If one of the storage server hosts is unavailable, then the other host takes over. The time needed for the stand-by storage server to take over is about 30 seconds. With one expected failover per year, the availability of the storage server is:

$$\left(1 - \frac{30}{31536000}\right) * 100 = 99,99999$$

The protocol servers are available for traffic during the backup dump and restore.

The storage server is available for provisioning during the backup dump and restore.

3.2.8 Severability

3.2.8.1 Accessibility

For IPWorks AAA server, the server is accessible under overload situation.

3.2.8.2 Retainability

An AAA server keeps all the session data in the MySQL NDB Cluster, 2 PL VMs share the session data.

3.2.9 Management Impacts

3.2.9.1 PM/FM Impacts

PM/FM measurements do not have any impact on the traffic.

3.3 DNS/ENUM System Characteristics

3.3.1 Traffic Scenarios

3.3.1.1 Server Node Lookup

End-user devices or network nodes as DNS clients look up the IP address of the server that provides network service.



IMS Registration	<p>A User Equipment(UE) supporting IMS sends a SIP REGISTER request to the P-CSCF. The client needs the IP address of the serving P-CSCF, DNS lookup of P-CSCF from the client to the DNS server is required.</p> <p>The P-CSCF server needs to route SIP messages to the next hop I-CSCF located in the Home Network. DNS lookup is required in the P-CSCF to obtain the IP address of a SIP-based I-CSCF server located in the Home Network. This type of operation continues through the SIP network (For example, lookup of S-CSCF). This applies as well to HSS lookup and RADIUS server lookup.</p> <p>The intermediate nodes must locally cache DNS results to reduce DNS server load and effective DNS query latency.</p>
AP and XCAP Server	A UE or server can look up the AP server. The AP server needs to look up the XCAP server of client.
SIP service Location	A UE or server can look up the SIP services supported for a node. Generally the client looks up the NAPTR records based on the URI (for example tel URI). The resulting URI points to a set of SRV records. The DNS SRV records return a set of domain names and ports where SIP services are located. The domain names point to A or AAAA records, which the client looks up to get a list of IP addresses. Use the same method to look up the other services (such as HTTP, WAP, and SCAP).
GPRS	The SGSN performs a DNS lookup to obtains the IP address of the GPRS GGSN node located in the Home Network associated with the MCC and MNC.
MMS Server	A client or server obtains IP address of an MMS server located in the network associated with the MCC and MNC.

3.3.1.2

Lookup of Client Related Information

Tel URI lookups (ENUM records)⁽¹⁾	Servers or UEs can look up the SIP URI for a user telephone number (E.164 address). This is required when more than one service provider exists in the network, and the telephone number is used to identify the user. The domain for the user must be determined, so the SIP request can be made toward the I-CSCF of that domain. The same principle is applicable to XCAP queries and presence for the latter, see for example RFC 3861).
General client services	SRV queries can also be defined for each subscriber. Besides SIP services, the SRV record can be used to include an email contact address, MMS home server URI, XCAP service URI location, or the home web page of the user.
Security	Reverse lookups records can be done to reduce the chances of spoofed domain names.

(1) This capability defined for IMS can also be reused for other services, for example MSS, which requires similar DNS lookups to obtain IP addresses of servers located in the network associated with MCC and MNC.

3.3.1.3

Caching

DNS information cached in the DNS clients (especially in the IMS core servers) reduces the fraction of the DNS queries. Caching can be done in end-user devices, core network nodes, and the DNS resolving servers. Active select DNS requires small value of TTL (Time To Live) to perform effectively.

3.3.1.4

General Use-Cases of DNS Server

IPWorks DNS server plays at least four distinct roles in a DNS query:



Use Case	Role	Description
1	Authoritative Server	An authoritative server with response based solely on its locally stored database. For example, requests are entirely processed by the DNS server and requests that to a different DNS server are not used. The 'named.conf' file includes entries like: zone "foo.com">{type master; file "db.foo.com";};
2	Delegation	Each query only receives a response that includes only the NS records and 'A' records of the DNS server responsible for the subdomains. For example, server only points to the authoritative servers serving the subdomains. When DNS client sends a DNS query, it must send a new DNS query to the name server in the response.
3	Forwarding	Each query results in the DNS query being forwarded to a specific IP address and the response is relayed back to the requester. This is similar to delegation but the DNS server does the work of the lookup instead of the client. The 'named.conf' file includes entries like: Zone "sales.foo.com" {forwards only; forwarders {10.2.3.4; 192.167.2.5; };}; This is normally used between a set of DNS servers for either of the following: <ul style="list-style-type: none">• Providing local control. For example, DNS servers for one company with different organizations. Each DNS has its own zones.• Providing scalability by breaking large set of domains into smaller zones with different zones handled by different DNS servers. Forwarding requires recursion which turns on caching.
4	Recursive resolver	The recursive resolver can be enabled by setting the option recursion to yes in the IPWorks BIND configuration file. The server does the work of the client by looking up the information outside itself and doing one or more DNS queries as needed to get the requested information. For example, <ul style="list-style-type: none">• A recursive-resolver DNS server sends a DNS request to root TLD DNS servers to find a DNS server that is authoritative for a domain;• Then the recursive-resolver DNS forms another DNS query toward that server to find another DNS server, and so on. The difference between a forwarding DNS server and a recursive-resolver DNS server is that only one DNS query is sent by the DNS server doing the forwarding, while recursive-resolver can do several queries to resolve the request.

(1) BIND does not make a major distinction between these two functions (For example, forwarding requires recursion though the depth of the recursion can be controlled).

If the DNS server (or intermediate server) is not an authoritative or a recursive-resolver, the client must act as an iterative resolver. Beyond the basic uses described earlier, an IPWorks DNS server has more properties. Authoritative servers are either master or secondary. Secondary DNS servers synchronize their zone data to the master DNS server through AXFR or IXFR transfer requests. This keeps the data of the secondary DNS server updated with that of the Master DNS server. Furthermore, a server can employ a cache (keeping frequent and recent DNS responses in a cache for reuse). The cache function is most valuable for the recursive-resolver use case.

Note: Each function is different in complexity, and the capacity of the DNS server is different in each role.

From an IPWorks product portfolio point of view, the use Case 1 and 3 described earlier are most likely to be used in the ICS. Use Case 4 is more typically used close



to the clients in the network, the actual end-user clients would probably be served by DNS servers as part of an operators existing IP access network. However, certain IMS server DNS user cases use more advanced functions, such as ENUM lookups and/or NAPTR records lookups. In those cases, the system solution benefits from having the IPWorks DNS server as a resolver with cache.

3.3.1.5 General Use-Cases of ENUM Server

IPWorks ENUM Server acts as an authoritative server. When the ENUM Server cannot resolve the query, it forwards the request to DNS server. For more information about ENUM specification, see Reference [7].

3.3.2 Processing Capacity

3.3.2.1 DNS Function

The DNS function is not expected to be heavily used in IMS Core or packet core network, as the ENUM records are handled by the ENUM function. The number of records to be handled by the DNS server is expected to be low. Characteristics are obtained for the DNS function when handling other type of records than ENUM, including:

- A
- PTR
- SRV
- NAPTR (Not related to ENUM)

The performance for the DNS function highly depends on the configuration of zones. For the same total number of records, different number of zones results in different performance. IPWorks DNS server can support maximal 18000 zones per system.

It is recommended to have a medium number of zones, for example, to configure 10,000 records, you have lots of options:

- 1 zone with 10,000 records
- 10 zones with 1000 records each
- 100 zones with 100 records each
- 10 000 zones with 1 record each

The configurations with one zone or 10,000 zones are not recommended. It is recommended to have a configuration with for example, 100 zones, as the number of zones affects the DNS server performance.



The replies to the 99% of the queries must be received in less than 5 ms for non-recursive queries.

The following tables show the DNS characteristics with 2000 zones and 1000 records per zone for IPv4 traffic.

Table 29 System Capacity for DNS IPv4 Traffic (Standard and Basic)

Platform	Number of Zones	Records per Zone	Engineered Capacity (TPS)	Dimensioned Capacity (TPS)	Average Latency (ms)	Reply Rate (5 ms)
CEE Standard & G9, Basic	2000	1000	70000	56000	<1	100%
G10, Basic	200	1000	75000	60000	1.2	99.99%
G10, Single	200	1000	75000	60000	1.2	100%

Table 30 shows the DNS characteristics with 1000 zones and 400 records per zone for IPv4 traffic.

Table 30 System Capacity for DNS IPv4 Traffic (CEE, Compact)

Number of Zones	Records per Zone	Engineered Capacity (TPS)	Dimensioned Capacity (TPS)	Average Latency (ms)	Reply Rate (5 ms)
1000	400	8900	7200	<1	99.99%

3.3.2.2

Internet DNS

IPWorks VNF for KVM (Native) DNS server and Single Server can be deployed as Internet DNS server. In such scenario, IPWorks DNS works as cache DNS server. When DNS client sends a query of a domain name to IPWorks DNS server for resolution, IPWorks DNS server checks if it is found in its local cache. If found, the result is returned. If not, IPWorks DNS server starts recursive query from the root DNS server until an authoritative name server for the queried domain is found, then the result from the authoritative name server is returned to DNS client and inserted into IPWorks DNS local cache. When TTL associated with one DNS resource record expires, the resource record is removed from the local cache.

Cache hit rate and network latency for recursive query have significant impact on the processing capacity of IPWorks DNS server. For individual DNS resource record, large TTL means high probability of hit in the cache. Small TTL means high probability of recursive query to external DNS server. The capacity of IPWorks DNS server is measured in two corresponding scenarios.

— Internet DNS with 100% cache hit

2 million DNS resource records and default cache size (unlimited) are configured at IPWorks DNS server. The TTL for the 2 million resource records is set to 86400 (24 hours), and they are completely cached. Queries for the configured 2 million resource records are generated towards IPWorks DNS server, and cache hit rate is 100%. The performance capacity of IPWorks DNS server is measured in 15-min time frame.



— Internet DNS with 100% recursive query

2 million DNS resource records and default cache size (unlimited) are configured at IPWorks DNS server. The TTL for the 2 million resource records is set to 0. Queries for the configured 2 million resource records are generated towards IPWorks DNS server. Cache hit rate is zero and all incoming queries are forwarded to external DNS server for recursive query. Network latency to external DNS server is set as OWD=10 ms, 20 ms and 50 ms respectively. The processing capacity of IPWorks DNS server is measured in 15-min time frame.

The following tables outline the measurement results of Internet DNS with 100% cache hit.

Table 31 Performance of Internet DNS with 100% Cache Hit (G9, Basic)

Engineered Capacity (TPS)	Dimensioned Capacity (TPS)	Reply Rate
95000	76000	99.99%

The following tables outline the measurement results of Internet DNS with 100% recursive query. The measurements are taken with default value of BIND option recursive-clients. Increasing the number of recursive-clients can improve the performance of DNS recursive query.

Table 32 System Capacity for IPWorks DNS Server with 100% Recursive Query (G9, Basic)

OWD (ms)	Engineered Capacity (TPS)	Dimensioned Capacity (TPS)
10	5000	4000
20	4600	3680
50	1500	1200

In real deployment, the system capacity can be approximated by setting a cache hit rate p and combining the measured capacities of 100% cache hit with 100% recursive query.

For example, if cache hit rate $p=80\%$, and OWD=20 ms, the system capacity can be derived by the following:

$$76000 * 0.8 + 3680 * 0.2 = 61536 \text{ TPS}$$

3.3.2.3

ENUM Monolithic

ENUM service resolves ENUM queries by following the standard S-NAPTR procedure. ENUM records are stored in database as NAPTR Resource Records which are provisioned to IPWorks storage server.

The performance for the ENUM Server is different from the performance for the DNS server. The number of zones for the ENUM Server does not have such a significant impact as for the DNS server.



Standard measurements are taken for:

- 24 million records, with 20 zones and 1,200,000 records per zone

Compact measurements are taken for:

- 0.6 million records, with 40 zones and 15,000 records per zone

The performance with 20,000 zones for the same total number of records is similar to the performance with 40 zones. There is a slight difference that is considered insignificant. The number of zones used to distribute the records does not affect the performance of the ENUM Server.

The number of records provisioned in the system has minor impact on the loadability of the system. The loadability is lower when the system is provisioned with a larger number of records.

Table 33 shows the ENUM monolithic server system capacity:

Table 33 Processing Capacity for ENUM Monolithic Server

Platform	Engineered Capacity (TPS)	Dimensioned Capacity (TPS)	Average Latency (ms)	Reply Rate (10 ms)
CEE, Standard	38000	30400	0.88	100%
CEE, Compact	10000	8000	0.88	100%
G10, Basic (48M)	55000	44000	1.7	100%

3.3.2.4

ENUM Front-End

ENUM supports the data layered architecture with cache mechanism. ENUM records are provisioned to the back-end database, such as CUDB, and are accessible through LDAP operations. Upon receiving ENUM query, ENUM FE checks its local cache for a hit first. If it is not found in local cache, ENUM FE initiates LDAP search toward the back-end database. Once found, the corresponding ENUM record is returned from the back-end database and inserted into the local cache of ENUM FE before it is returned to the ENUM client. TTL is associated with each record stored in the cache. Once TTL expires, the record is removed from cache. If the subscription changes, the back-end database must notify ENUM FE through SOAP notification. ENUM FE then refetches the changed record and refreshes its cache.

The performance of ENUM FE highly depends on the cache hit rate and network latency to the back-end database. Two test cases are performed to measure the corresponding performance of ENUM FE with 100% cache hit and 100% outgoing query respectively. Table 34 summarizes the measurement results for ENUM FE with 100% cache hit.



Table 34 Processing Capacity for ENUM FE with 100% Cache Hit

Platform	Engineered Capacity (TPS)	Dimensioned Capacity (TPS)	Average Latency (ms)	Reply Rate (10 ms)
CEE Standard	35000	28000	1.741	100%
G9 Basic	45000	36000	1.59	99.77%
G10 Basic	55000	44000	3.9	99.69%

The following tables summarize the measurement results for ENUM FE with 100% outgoing query. Table 36 summarizes the result with OWD = 10 ms.

Table 35 Processing Capacity for ENUM FE with 100% Outgoing Query (Single Server)

Platform	Engineered Capacity (TPS)	Dimensioned Capacity (TPS)	Average Latency (ms)	Reply Rate (10 ms)
G9	9000	7200	-	99.77%

Table 36 Processing Capacity for ENUM FE with 100% Outgoing Query

Platform	OWD (ms)	Engineered Capacity (TPS)	Dimensioned Capacity (TPS)	Average Latency (ms)	Reply Rate
CEE Standard	10	10000	8000	26.79	99.95%
G10 Basic	10	10000	8000	31.07	99.95%
G10 Single	10	10000	8000	29.6	99.95%

3.3.2.5

ENUM Traffic Using ENUM Views and ACLs

IPWorks allows the creation of ENUM views. A view is a set of ENUM zones that specific clients are allowed to retrieve information from. Access Control List (ACL) is a list of IP addresses representing the clients that have access to specific ENUM views.

Two types of ENUM views are listed as follows:

- Overlapping, for example, different ENUM views include the same ENUM zone
- Not overlapping, for example, different ENUM views include different ENUM zones

For more information about the handling of ENUM views and ACL, refer to [Configure DNS and ENUM, Reference \[4\]](#).

For more information about the EnumAcl and EnumView objects, refer to [IPWorks DNS, ASDNS, ENUM Parameter Description, Reference \[5\]](#).

The configuration of ENUM views and ACLs has a small impact on the ENUM traffic performance. The performance of the ENUM Server is similar to the performance without using ENUM views.



The use of overlapping ENUM views vs. non-overlapping ENUM views does not severely affect the performance of the ENUM Server.

Each ACL is configured with a set of client IP addresses. When the ACLs contain a greater number of client IP addresses (10 vs. 5), the performance is **slightly** deteriorated.

3.3.2.6 ENUM Traffic with Number Portability Queries over SS7

IPWorks ENUM Server handles the ENUM queries from ENUM clients. When IPWorks receives an ENUM query for an E.164 telephony subscriber number, it gets information from its own database and replies with the information to the subscriber.

When IPWorks does not have the information required for the subscriber, it forms an extra query to an external Number Portability Database (NPDB) or toll-free database to get the routing information for that subscriber. Then the routing information is used to compose the ENUM reply.

NP and toll-free queries to Service Control Function (SCF) nodes use the AIN protocol. NP query to SCF node use ETSI INAP protocol and to Mobile Number Portability Signaling Relay Function (MNP SRF) node use MAP protocol. For MAP protocol, two MAP services are supported, which are MAP ATI (Any Time Interrogation) and MAP SRI (Send Routing Information). Depending on the configuration, either MAP ATI or MAP SRI is used. The performance is the same for MAP ATI and MAP SRI.

Because the NP information is not stored in the local database, it takes more time to query the ENUM for the NP records. The NP traffic performance is greatly affected by the network latency.

Table 37 gives the characteristic measurement results for 100% NP traffic using MAP ATI/SRI:

Table 37 100% NP Traffic Through MAP ATI/SRI

Platform	NP Ratio	Engineered Capacity (TPS)	Dimensioned Capacity (TPS)
CEE	100%	10000	8000
G9 Basic	100%	10000	8000
G9 Single Server	100%	8000	6400
G10 Basic	100%	15000	12000
G10 Single	99.99%	10000	8000

Table 38 gives the characteristic measurement results for 100% NP traffic using INAP protocol:



Table 38 100% NP Traffic Through INAP

Platform	NP Ratio	Engineered Capacity (TPS)	Dimensioned Capacity (TPS)
CEE Standard	100%	10000	8000
G9, Basic	100%	10000	8000

Table 39 gives the characteristic measurement results for 100% NP traffic using AIN protocol:

Table 39 100% NP Traffic Through AIN

Platform	NP Ratio	Engineered Capacity (TPS)	Dimensioned Capacity (TPS)
CEE Standard	100%	10000	8000
G9, Basic	100%	10000	8000

3.3.2.7

ENUM Traffic with Number Portability Queries over LDAP

Lightweight Directory Access Protocol (LDAP) is an open, vendor-neutral, industry standard application protocol for accessing and maintaining distributed directory information services over an IP Network.

IPWorks DNS/ENUM server with ERH module supports NP queries toward NPDB either over SS7 or over LDAP.

3.3.2.7.1

100% NP Request over LDAP

In this case, 100% of requests sent to IPWorks DNS/ENUM server are NP queries. ERH is configured to query NPDB over LDAP protocol.

The following tables summarize the processing capacity of IPWorks DNS/ENUM server handling 100% NP requests over LDAP under different network latencies.

Table 40 100% NP Request over LDAP (CEE Standard)

OWD (ms)	Engineered Capacity (TPS)	Dimensioned Capacity (TPS)	Average Latency (ms)	Reply Rate
10	18000	14400	48.39	100%
20	15000	12000	62.90	100%
50	6800	5440	142.10	100%

Table 41 100% NP Request over LDAP (G9, Basic with Real CUDB)

OWD (ms)	Engineered Capacity (TPS)	Dimensioned Capacity (TPS)	Average Latency (ms)	Reply Rate
10	25000	20000	0.6	100%
20	20000	16000	0.6	100%
50	18000	14400	0.6	100%



Table 42 100% NP Request over LDAP (G9,Single Server)

OWD (ms)	Engineered Capacity (TPS)	Dimensioned Capacity (TPS)	Average Latency (ms)	Reply Rate
0	10000	8000	10.95	99.72%

Table 43 100% NP Request over LDAP (G10, Basic with Simulator)

OWD (ms)	Engineered Capacity (TPS)	Dimensioned Capacity (TPS)	Average Latency (ms)	Reply Rate
10	13000	10000	35.6	99.99%
20	9500	7600	104.6	100%
50	4800	3840	215.7	100%

3.3.2.8

IMS Interconnection

IMS Interconnection as an enhancement of standard DNS/ENUM infrastructure provides the E.164 number resolution in different operator domains. IPWorks supports IMS interconnection with MAP protocol.

There are different traffic scenarios for IMS interconnection function:

- Number is ported to another operator that is IMS interconnection participating operator.
- Number is ported to another operator that is not IMS interconnection participating operator or non-IMS network.

Table 44 IMS Interconnect with 100% NP traffic (CEE, Standard)

Transport	Engineered Capacity (TPS)	Dimensioned Capacity (TPS)
SIGTRAN	6000	4800

3.3.2.9

Critical Resource for ENUM Server

When running the ENUM traffic, the critical resource is the capacity of MySQL NDB. To improve the performance of NDB, use a higher frequency CPU to reach higher TPS.

IPWorks ENUM Server does not have an overload protection mechanism. This means that all the requests are to be processed. In an overload scenario, this leads to an increased latency.

During an overload scenario, some requests are discarded owing to the exhaustion of request buffers. The probability that a query is answered depends on the load of the system.

The provisioning of the system does not have impact on the ENUM performance, but high NP traffic affects the performance of the ENUM Server.



3.3.3 Provisioning Capacity

The following sections outline the provisioning capacity for DNS and ENUM under different scenarios.

3.3.3.1 DNS Provisioning Capacity

Provisioning performance of DNS is verified for Creation, Modification, and Deletion operations separately. During verification, all provisioned records are acknowledged within 5 seconds.

Table 45 shows the performance results:

Table 45 DNS Provisioning Capacity (CEE, Standard)

Case	Avg. Rate (RR/Sec)	Reply Rate (5 s)
DNS Create	65	100%
DNS Modify	63	100%
DNS Delete	69	100%

3.3.3.2 ENUM Provisioning Capacity

Provisioning performance of ENUM monolithic is verified for Creation, Modification, and Deletion operations separately. Each operation is verified for ENUM server with 24 million-based records respectively. During verification, all provisioned records are acknowledged within 5 seconds.

Table 46 shows the performance results:

Table 46 ENUM Provisioning Capacity (CEE, Standard)

Case	Avg. Rate (RR/Sec)	Reply Rate (5 s)
ENUM Create	148	100%
ENUM Modify	90	100%
ENUM Delete	92	100%

3.3.4 Disk Capacity

3.3.4.1 Storage Server

IPWorks storage server runs in SC VM. Disk capacity usage in SC VM for ENUM is outlined in Table 47.



Table 47 Disk Space Used in SC VM for ENUM

Scenario	Total (GiB)	Used (GiB)
24 million ENUM NAPTR records	280	38

3.3.4.2 Protocol Server

DNS/ENUM service runs on PL VM, and has access to NDB in SC VM for provisioned ENUM records and DNS configuration. There is no disk use on PL VM.

3.3.5 Scalability

IPWorks VNF standard configuration can be scaled out to max 10 PLs depending on different IPWorks service resource requirement. For DNS/ENUM, the max supported PL number after scaling-out operation and corresponding critical resource are outlined in Table 48.

Table 48 Max Supported PL Number and Critical Resource for Scalability (ENUM)

	ENUM	ENUM Front-End
Critical Resource	SC memory ⁽¹⁾	N/A
Number of SC	2	2
Max Number of PL after scale-out	4	10

(1) Memory on SC limits the max number of subscribers IPWorks VNF can support, which in turn limits the max number of PL needed to process request generated correspondingly.

IPWorks VNF processing capacity in terms of TPS is increased linearly by adding PL until PL number reaches the max supported number.

For ENUM-FE solution, the max supported subscribers depends on CUDB capacity.

The processing capacity of single PL is half of the processing capacity of IPWorks VNF outlined in Section 3.3.2 Processing Capacity on page 33.

IPWorks DNS server does not support scalability.

3.3.6 Quality of Service

3.3.6.1 Delay

The latency for a DNS server is low for small zone sets, but is a problem for large zone sets.

The ENUM Server is not affected by the number of zones. But with overload, the latency increases significantly.



3.3.6.2 Quality of Service Impacts on Overload Situations

No overload protection mechanism is available in IPWorks DNS/ENUM. Though the capacity of IPWorks DNS/ENUM is large enough in most scenarios, when any overload occurs, processing capacity dips and latency increases sharply.

3.3.7 Dependability

3.3.7.1 Reliability

Reliability is the probability that the IPWorks performs a function over a given period. Within this period, the system must be available.

Stability test is performed in the IPWorks system. One proposed stability test case is as follows:

- DNS/ENUM and NP enabled at the same time. 48 hours running with 30% of engineered capacity load for each solution.

The IPWorks system has passed this test case.

3.3.7.2 Availability

Availability is the probability that the system is in a state ready to perform a function. The system is not available under certain conditions, such as:

- Restart
- Cold backup and restore

IPWorks system is configured in a fully local redundant way. In case that service in 1 SC VM or PL VM fails, the service on redundant SC or PL VM in the system continues to handle the traffic. Failed IPWorks service can be restarted by underlying AMF component in seconds.

3.3.8 Serviceability

3.3.8.1 Accessibility

Accessibility Overload scenario is not tested. It will be investigated later.

3.3.8.2 Retainability

Retainability is not applicable to a DNS/ENUM Server.

DNS/ENUM provides query-and-answer type of service. Once provided, the service is not retained.



3.3.9 Management Impacts

3.3.9.1 PM/FM Impacts

PM/FM measurements do not have any impact on the traffic.

3.4 DHCP System Characteristics

3.4.1 Traffic Scenarios

3.4.1.1 General Use-cases of DHCP Server

IPWorks DHCP server supports most of the standard DHCP server and client options defined in IETF RFCs. The DHCP server supports DHCP relay and DHCP proxy clients typically found in routers, AAA servers, GGSNs, and PDSN nodes. The server is optimized for the carrier networks.

DHCP allows remote clients to connect to a network, obtain a temporary IP address, and retrieve client configuration parameters.

3.4.2 Processing Capacity

Dynamic Host Configuration Protocol (DHCP) handles the assignment of dynamic IP addresses for the remote end systems. When a request from a client is received, the system checks the received information, assigns an IP address to the client, and replies to the client indicating the IP address is assigned.

The system must be configured with an address space. And the DHCP server is configured indicating how that address space is to be managed. A subnet represents the configuration for a contiguous set of addresses in an IP subnet.

A pool is a contiguous set of IP addresses available for dynamic assignment. A pool must be contained in a subnet.

The configurations for the DHCP servers are listed as follows:

- Single mode

One DHCP server.

- Failover mode

Two DHCP servers, one acting as the primary and the other as the secondary. Each server updates the other with information about the DHCP leases. If one of them fails, the other takes over and continues serving the network.

The main types of the transaction request are listed as follows:



- Discover
- New lease
- Renew lease
- Release

When a client requests for a new IP address, it uses the transactions 'discover' and 'new lease'. The DHCP is configured with a lease renewal time. That is, the server indicates the client how often the client does a renewal of the IP address toward the server. Depending on the configuration of this parameter, the clients will determine the frequency of renewal.

Different types of transaction requests generate different loads in the system.

The traffic mixes are listed as follows:

- High: 40% discover, 40% new lease, 10% renew lease, 10% release
- Normal: 10% discover, 10% new lease, 70% renew lease, 10% release

The release transactions are not replied by the server, which is the reason why the number of the answered transactions is smaller than that of the offered transactions.

The DHCP test result is based on the below configurations:

- Pool Number: 1000
- Lease Number: 5 M
- Link Number: 1
- ClientClass Number: 0
- Lease Overlap: 0

3.4.2.1 Failover Configuration

In a failover configuration, the TPS is affected because each server must communicate with each other.

The following figures show the load for different transaction mixes, with 1000 pools, 5000000 addresses.

Table 49 shows the performance for the normal traffic model and high traffic model:



Table 49 System Capacity for DHCP

Platform	Model Name	Engineered Capacity	Dimensioned Capacity
CEE, Standard	Normal Traffic	8000	6400
	High Traffic	4000	3200
Native (G9)	Normal Traffic	9000	7200
	High Traffic	4000	3200
Native (G10, Basic and Single)	Normal Traffic	15000	12000
	High Traffic	5000	4000

Table 50 System Capacity for DHCP with Local Disk

Platform	Model Name	Engineered Capacity	Dimensioned Capacity
Native (G9, local disk)	Normal Traffic	11000	8800
	High Traffic	7000	5600
Native (G10, local disk)	Normal Traffic	15500	12400
	High Traffic	8500	6800

3.4.3

Disk Capacity

A DHCP server records all the allocated IP addresses both in the memory and in the file. When DHCP server is rebooted, it loads all the information from the lease file. DHCP server has two lease files named `dhcpd.leases` and `dhcpd.leases~.gz`. The `dhcpd.leases~.gz` file is used for rotating. The size of `dhcpd.leases~.gz` file depends on the quantity of active leases. The size of `dhcpd.leases` is also variable. The following is a snapshot of disk usage.

Table 51

File/Directory	Disk Space Used
<code>dhcpd.leases</code>	2000 M
<code>dhcpd.leases~.gz</code>	200M

3.4.4

Scalability

IPWorks DHCP server does not support scalability.

3.4.5

Quality of Service

3.4.5.1

Delay

In the overload scenario, the latency increases significantly. It is recommended that a DHCP server run with a capacity less than the engineered capacity.



3.4.5.2 Quality of Service Impacts on Overload Situations

At a load above the engineered capacity, IPWorks DHCP server can handle the traffic with reduced capacity.

- System Robustness

IPWorks DHCP server can sustain constant overload situations and return to normal operations once traffic has reached normal levels.

- Performance degradation during overload situation

- At an offered load of 200%, there is at least a throughput of 80%.
- At an offered load of 500%, there is at least a throughput of 20%.

3.4.6 Dependability

3.4.6.1 Reliability

Reliability is the probability that the IPWorks can perform a function over a given period. Within this period the system must be available.

One proposed stability test case is as follows:

- DHCP failover configuration. 48 hours running with 80% of engineered capacity load.

The IPWorks system has passed this test case.

3.4.6.2 Availability

Availability is the probability that the system is in a state ready to perform a function. The system is not available under certain conditions, such as:

- Restart
- Cold backup and restore

IPWorks system is configured in a fully local redundant way. In case that service in 1 SC VM or PL VM fails, the service on redundant SC or PL VM in the system continues to handle the traffic. Failed IPWorks service can be restarted by underlying AMF component in seconds.

3.4.7 Severability

3.4.7.1 Accessibility

For IPWorks DHCP server, the server is accessible under overload situation.



3.4.7.2 Retainability

A DHCP server keeps all the active lease information both in the memory and in the lease file. When the DHCP server is down and restarted, it loads the lease information from the lease file.

3.4.8 Management Impacts

3.4.8.1 PM/FM Impacts

PM/FM measurements do not have any impact on the traffic.



4 VNF Life Cycle Management

The VNF-LCM provides a workflow execution environment and a web-based application for managing VNF life cycle procedures. VNF life cycle procedures are realized by executing ordered sequences of steps, which is called workflows.

Table 52

Workflow	Duration	Comments
Instantiate	22 minutes	
Scale-in	5 minutes for 1 PL 15 minutes for 8 PLs	A less than 10 seconds partial traffic loss will happen.
Scale-out	5 minutes for one PL 15 minutes for 8 PLs	A less than 10 seconds partial traffic loss will happen.
Terminate	2 minutes	





Reference List

IPWorks Library Documents

- [1] Trademark Information
- [2] Typographic Conventions
- [3] Glossary of Terms and Acronyms
- [4] Configure DNS and ENUM
- [5] IPWorks DNS, ASDNS, ENUM Parameter Description

PCAT and Other Ericsson Documents

- [6] IPWorks Dimensioning Guideline, 1/192 02-FGC 101 3568

Standard

- [7] [The E.164 to URI Dynamic Delegation Discovery System Application \(ENUM\), RFC 3761, April 2004](#)