3GPP TS 28.554 V18.0.0 (2022-12)

Technical Specification

3rd Generation Partnership Project;
Technical Specification Group Services and System Aspects;Management and orchestration;
5G end to end Key Performance Indicators (KPI)
(Release 18)

Keywords

5G,KPI,management,orchestration

3GPP

Postal address

3GPP support office address

650 Route des Lucioles - Sophia Antipolis
Valbonne - FRANCE
Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Internet

http://www.3gpp.org

Contents

Foreword

This Technical Specification has been produced by the 3rd Generation Partnership Project (3GPP).

The contents of the present document are subject to continuing work within the TSG and may change following formal TSG approval. Should the TSG modify the contents of the present document, it will be re-released by the TSG with an identifying change of release date and an increase in version number as follows:

Version x.y.z

where:

x the first digit:

1 presented to TSG for information;

2 presented to TSG for approval;

3 or greater indicates TSG approved document under change control.

y the second digit is incremented for all changes of substance, i.e. technical enhancements, corrections, updates, etc.

z the third digit is incremented when editorial only changes have been incorporated in the document.

1 Scope

The present document specifies end-to-end Key Performance Indicators (KPIs) for the 5G network and network slicing.

2 References

The following documents contain provisions which, through reference in this text, constitute provisions of the present document.

- References are either specific (identified by date of publication, edition number, version number, etc.) or non-specific.
- For a specific reference, subsequent revisions do not apply.
- For a non-specific reference, the latest version applies. In the case of a reference to a 3GPP document (including a GSM document), a non-specific reference implicitly refers to the latest version of that document *in the same Release as the present document*.

[1] 3GPP TR 21.905: "Vocabulary for 3GPP Specifications".

[2] Void.

[3] ITU-T Recommendation E.800: "Definitions of terms related to quality of service".

[4] 3GPP TS 24.501: " Non-Access-Stratum (NAS) protocol for 5G System (5GS); Stage 3".

[5] 3GPP TS 38.331: "NR; Radio Resource Control (RRC); Protocol specification".

[6] 3GPP TS 28.552: "Management and orchestration; 5G performance measurements".

[7] 3GPP TS 23.501: " System Architecture for the 5G System; Stage 2".

[8] ETSI ES 203 228 V1.2.1 (2017-04): "Environmental Engineering (EE); Assessment of mobile network energy efficiency".

[9] 3GPP TS 28.310: "Management and orchestration; Energy efficiency of 5G".

[10] ETSI 202 336-12 V1.2.1 (2019-02): "Environmental Engineering (EE); Monitoring and control interface for infrastructure equipment (power, cooling and building environment systems used in telecommunication networks); Part 12: ICT equipment power, energy and environmental parameters monitoring information model".

[11] ETSI GS NFV-IFA 027 V4.0.2 (2020-11): "Network Functions Virtualisation (NFV) Release 4; Management and Orchestration; Performance Measurements Specification".

[12] 3GPP TS 38.314: "NR; layer 2 measurements".

[13] 3GPP TS 22.261: "Service requirements for the 5G system".

[14] 3GPP TS 38.214: "NR; Physical layer procedures for data".

3 Definitions and abbreviations

3.1 Definitions

For the purposes of the present document, the terms and definitions given in 3GPP TR 21.905 [1] and the following apply. A term defined in the present document takes precedence over the definition of the same term, if any, in 3GPP TR 21.905 [1].

3.2 Abbreviations

For the purposes of the present document, the abbreviations given in 3GPP TR 21.905 [1] and the following apply. An abbreviation defined in the present document takes precedence over the definition of the same abbreviation, if any, in 3GPP TR 21.905 [1].

EE Energy Efficiency
kbit kilobit (1000 bits)
RTT Round Trip Time

4 End to end KPI concept and overview

The following KPI categories are included in the present document:

- Accessibility (see the definition in [3]).
- Integrity (see the definition in [3]).
- Utilization.
- Retainability (see the definition in [3]).
- Mobility.
- Energy Efficiency.
- Reliability (See the definition in [13]).

Editor's note: For future update of the document it will also include:

- Availability.

5 KPI definitions template

a) Name (Mandatory): This field shall contain the name of the KPI.

b) Description (Mandatory): This field shall contain the description of the KPI.
Within this field it should describe if the KPI is focusing on network or user view. This filed should also describe the logical KPI formula to derive the KPI. For example, a success rate KPI's logical formula is the number of successful events divided by all events. This field should also show the KPI unit (e.g., kbit/s, millisecond) and the KPI type (e.g., mean, ratio).

c) Formula definition (Optional):
This field should contain the KPI formula using the 3GPP defined measurement names.
This field can be used only when the measurement(s) needed for the KPI formula are defined in 3GPP TS for performance measurements (TS 28.552 [6]). This field shall clarify how the aggregation shall be done, for the KPI object level(s) defined in d).

d) KPI Object (Mandatory):
This field shall contain the DN of the object instance where the KPI is applicable, including the object where the measurement is made. The DN identifies one object instance of the following IOC:

- `NetworkSliceSubnet`
- `SubNetwork`
- `NetworkSlice`
- `NRCellDU`
- `NRCellCU`

e) Remark (Optional):
This field is for additional information reqquired for the KPI definition,
e.g. the definition of a call in UTRAN.

6 End to end KPI definitions

6.1 KPI Overview

The KPI categories defined in [2] will be reused by the present document.

6.2 Accessibility KPI

6.2.1 Mean registered subscribers of network and network slice through AMF

a) AMFMeanRegNbr.

b) This KPI describe the mean number of subscribers that are registered to a network slice instance. It is obtained by counting the subscribers in AMF that are registered to a network slice instance. It is an Interger. The KPI type is CUM.

c)
$$AMFMeanRegNbr = \sum_{AMF} RegisteredSubNbrMean.SNSSAI$$

d) SubNetwork, NetworkSlice

6.2.2 Registered subscribers of network through UDM

a) UDMRegNbr.

b) This KPI describe the total number of subscribers that are registered to a network through UDM. It is corresponding to the measurement RM.RegisteredSubUDMNbrMean that counts subscribers registered in UDM. It is an Interger. The KPI type is CUM.

c)
$$UDMRegNbr = \sum_{UDM} RegisteredSubUDMNbrMean$$

d) SubNetwork

6.2.3 Registration success rate of one single network slice

a) RSR.

b) This KPI describes the ratio of the number of successfully performed registration procedures to the number of attempted registration procedures for the AMF set which related to one single network slice and is used to evaluate accessibility provided by the end-to-end network slice and network performance. It is obtained by successful registration procedures divided by attempted registration procedures. It is a percentage. The KPI type is RATIO.

c)

$$\mathrm{RSR} = \frac{\sum_{Type} \mathrm{AMF.5GSRegisSucc.}\mathit{Type}}{\sum_{Type} \mathrm{AMF.5GSRegisAtt.}\mathit{Type}} * 100\%$$

NOTE: Above measurements with subcounter .*Type* should be defined in 3GPP TS 24.501 [4].

d) NetworkSlice

6.2.4 Partial DRB Accessibility for UE services

a) Partial DRB Accessibility

b) This KPI describes the DRBs setup success rate, including the success rate for setting up RRC connection and NG signalling connection. It is obtained as the succeess rate for RRC connection setup multiplied by the success rate for NG signalling connection setup multiplied by the success rate for DRB setup. The success rate for RRC connection setup and for NG signalling connection setup shall exclude setups with establishment cause mo-Signalling [5]. It is a percentage. The KPI type is RATIO.

c) Partial DRBAccessibility 5QI = (∑RRC.ConnEstabSucc.*Cause*/∑RRC.ConnEstabAtt.*Cause*) * (∑UECNTXT.ConnEstabSucc.*Cause*/∑ UECNTXT.ConnEstabAtt.*Cause*) * (DRB.EstabSucc.5QI/DRB.EstabAtt.5QI) * 100

Partial DRB Accessibility SNSSAI = (∑RRC.ConnEstabSucc.*Cause*/∑RRC.ConnEstabAtt.*Cause*) * (∑UECNTXT.ConnEstabSucc.*Cause*/∑ UECNTXT.ConnEstabAtt.*Cause*) * (DRB.EstabSucc.SNSSAI/DRB.EstabAtt.SNSSAI) * 100.

The sum over causes shall exclude the establishment cause mo-Signalling [5].

For KPI on SubNetwork level the measurement shall be the averaged over all NRCellCUs in the SubNetwork

d) SubNetwork, NRCellCU.

6.2.5 PDU session Establishment success rate of one network slice (S-NSSAI)

a) PDUSessionEstSR.

b) This KPI describes the ratio of the number of successful PDU session establishment request to the number of PDU session establishment request attempts for all SMF which related to one network slice (S-NSSAI) and is used to evaluate accessibility provided by the end-to-end network slice and network performance. It is obtained by the number of successful PDU session requests divided by the number of attempted PDU session requests. It is a percentage. The KPI type is RATIO.

c)

$$PDUSessionEstSR = \frac{\sum_{SMF} SM.PduSessionCreationSucc.SNSSAI}{\sum_{SMF} SM.PduSessionCreationReq.SNSSAI}$$

d) NetworkSlice

6.2.6 Maximum registered subscribers of network slice through AMF

a) AMFMaxRegNbr.

b) This KPI describe the maximum number of subscribers that are registered to a network slice. It is obtained by counting the subscribers in AMF that are registered to a network slice. It is an Interger. The KPI type is CUM.

$$AMFMaxRegNbr = \sum_{AMF} RegisteredSubNbrMax.SNSSAI$$

c)

d) NetworkSlice

6.2.7 Total DRB accessibility for UE services

a) Total DRB accessibility

b) This KPI describes the total DRBs accessibility obtained as the ratio of the number of successfully established DRBs and number of services intended to be setup by the end user that shall result into a DRB establishment via Initial Context setup procedure, Added DRB setup and RRC Resume procedure. The number of services intended to be setup by the end user that shall result into a DRB establishment via Initial Context setup procedure is obtained as number of attempted establishments of DRB via Initial Context setup procedure amplified by inverse of the UE-associated logical NG-connection success ratio further amplified by inverse of the RRC Connection setup state success ratio. The number of services intended to be setup by the end user that shall result into a DRB establishment via added DRB setup procedure is measured directly in gNB via number of attempted establishments of DRB via added DRB setup procedure. Finally the number of services intended to be setup by the end user that shall result into a DRB establishment via RRC Resume procedure is provided as number of attempted establishments of DRB via RRC Resume procedure amplified by inverse of the RRC Resume success ratio. The success rate for RRC connection setup and for UE-associated logical NG-connection setup shall exclude setups with establishment cause mo-Signalling [5]. The success rate for RRC resume shall exclude setups related to RNA update. It is a percentage. The KPI type is RATIO.

c) DRBAccessibility 5QI = 100 * (DRB.InitialEstabSucc.5QI + (DRB.EstabSucc.5QI-DRB.InitialEstabSucc.5QI) + DRB.ResumeSucc.5QI)/(DRB.InitialEstabAtt.5QI/((RRC connection setup success rate /100)*(UE-associated logical NG-connection success ratio/100)) + (DRB.EstabAtt.5QI-DRB.InitialEstabAtt.5QI) + DRB.ResumeAtt.5QI/(RRC Resume success rate/100))

DRBAccessibility SNSSAI = 100 * (DRB.InitialEstabSucc. SNSSAI + (DRB.EstabSucc. SNSSAI - DRB.InitialEstabSucc. SNSSAI) + DRB.ResumeSucc. SNSSAI)/(DRB.InitialEstabAtt. SNSSAI /((RRC connection setup success rate /100)*(UE-associated logical NG-connection success ratio /100)) + (DRB.EstabAtt. SNSSAI -DRB.InitialEstabAtt. SNSSAI) + DRB.ResumeAtt. SNSSAI /(RRC Resume success rate/100))

Where:

RRC Resume success rate = 100* ∑RRC.ResumeSucc.cause /∑(RRC.ResumeAtt.cause - RRC.ResumeFallbackToSetupAtt.cause), where all but the causes related to RNA update shall be included.

RRC connection setup success rate = 100* (∑(RRC.ConnEstabSucc.Cause + RRC.ResumeSuccByFallback.cause) + RRC.ReEstabSuccWithoutUeContext) /(∑(RRC.ConnEstabAtt.Cause + RRC.ResumeFallbackToSetupAtt.cause) + RRC.ReEstabFallbackToSetupAtt)

UE-associated logical NG-connection success ratio = 100*(∑UECNTXT.ConnEstabSucc.Cause/∑UECNTXT.ConnEstabAtt.Cause)

The sum over causes shall exclude the establishment cause mo-Signalling [5].

The sum over causes for RRC resume shall exclude the causes related to RNA update [5].

For KPI on SubNetwork level the measurement shall be the averaged over all NRCellCUs in the SubNetwork

d) SubNetwork, NRCellCU.

6.2.8 Mean CM-Connected subscribers of network slice through AMF

a) AMFMeanCmConNbr.

b) This KPI describe the mean number of subscribers in a period that are not only registered to a network slice but also established a PDU session related to the network slice. And subscribers also have a NAS signalling connection with the AMF over N1. It is obtained by counting the subscribers in AMF that are showed "cm-connected" state for a network slice. It is an Interger. The KPI type is CUM.

c)

$$AMFMeanCmConNbr = \sum_{AMF} CM - ConnectedSubNbrMean.SNSSAI$$

d) NetworkSlice.

6.2.9 Maximum on-line subscribers of network slice through AMF

a) AMFMaxCmConNbr.

b) This KPI describe the maximum number of subscribers in a period that are not only registered to a network slice but also established a PDU session related to a network slice. And subscribers also have a NAS signalling connection with the AMF over N1. It is obtained by counting the subscribers in AMF that are showed "cm-connected" state for a network slice. It is an Interger. The KPI type is CUM.

c)

$$AMFMaxCmConNbr = \sum_{AMF} CM - ConnectedSubNbrMax.SNSSAI$$

d) NetworkSlice.

6.2.10 PFCP session established success rate of one network and one network slice

a) PFCPSessionEstSR.

b) This KPI describes the successful rate of PFCP session established in a network or a network slice e on the UPF.

It is used to evaluate the quality of user-plane connection established and the accessibility provided by the end-to-end network slice and network performance. It is obtained by the number of successful PFCP session requests divided by the number of attempted PFCP session requests. It is a percentage. The KPI type is RATIO.

c)

$$PFCPsessionEstSR = \frac{UPF.PFCPSessionCreationSucc.SNSSAI}{UPF.PFCPSessionCreationReq.SNSSAI}$$

d) Subnetwork, NetworkSlice.

6.3 Integrity KPI

6.3.1 Latency and delay of 5G networks

6.3.1.0 Void

6.3.1.1 Downlink latency in gNB-DU

a) DLLat_gNB-DU.

b) This KPI describes the gNB-DU part of the packet transmission latency experienced by an end-user. It is used to evaluate the gNB latency contribution to the total packet latency. It is the average (arithmetic mean) of the time from reception of IP packet to gNB-DU until transmission of first part of that packet over the air interface, for a packet arriving when there is no previous data in queue for transmission to the UE. It is a time interval (0.1 mS). The KPI type is MEAN. This KPI can optionally be split into KPIs per QoS level (mapped 5QI or QCI in NR option 3) and per S-NSSAI.

c) DLLat_gNB-DU = DRB.RlcSduLatencyDl

or optionally DLLat_gNB-DU.*QoS* = DRB.RlcSduLatencyDl.QoS where QOS identifies the target QoS quality of service class.

or optionally DLLat_gNB-DU.*SNSSAI* = DRB.RlcSduLatencyDl.*SNSSAI* where *SNSSAI* identifies the S-NSSAI.

d) NRCellDU

6.3.1.2 Integrated downlink delay in RAN

6.3.1.2.1 Downlink delay in NG-RAN for a sub-network

a) DLDelay_NR_SNw.

b) This KPI describes the average packet transmission delay through the RAN part to the UE. It is used to evaluate delay performance of NG-RAN in downlink for a sub-network. It is the weighted average packets delay from reception of IP packet in gNB-CU-UP until the last part of an RLC SDU packet was received by the UE according to received HARQ feedback information for UM mode or until the last part of an RLC SDU packet was received by the UE according to received RLC ACK for AM mode. It is a time interval (0.1 ms). The KPI type is MEAN. This KPI can optionally be split into KPIs per QoS level (mapped 5QI or QCI in NR option 3) and per S-NSSAI.

c) Below are the equations for average "Integrated downlink delay in RAN" for this KPI on SubNetwork level. The "Integrated downlink delay in RAN" is the sum of average DL delay in gNB-CU-UP of the sub-network (DLDelay_gNBCUUP_SNw) and the average DL delay in gNB-DU of the sub-network (DLDelay_gNBDU_SNw):

   DLDelay_NR_SNw = DLDelay_gNBCUUP_SNw + DLDelay_gNBDU_SNw

   or optionally DLDelay_ NR_SNw.*QOS* = DLDelay_gNBCUUP_SNw.*QOS* + DLDelay_gNBDU_SNw.*QOS* where *QOS* identifies the target quality of service class.

   or optionally DLDelay_NR_SNw.*SNSSAI* = DLDelay_gNBCUUP_SNw.*SNSSAI* + DLDelay_gNBDU_SNw.*SNSSAI* where *SNSSAI* identifies the S-NSSAI.

d) SubNetwork

6.3.1.2.2 Downlink delay in NG-RAN for a network slice subnet

a) DLDelay_NR_Nss.

b) This KPI describes the average packet transmission delay through the RAN part to the UE. It is used to evaluate delay performance of NG-RAN in downlink for a network slice subnet. It is the weighted average packets delay from reception of IP packet in gNB-CU-UP until the last part of an RLC SDU packet was received by the UE according to received HARQ feedback information for UM mode or until the last part of an RLC SDU packet was received by the UE according to received RLC ACK for AM mode. It is a time interval (0.1 ms). The KPI type is MEAN.

c) Below is the equation for average "Integrated downlink delay in RAN" for this KPI on NetworkSliceSubnet level. The "Integrated downlink delay in RAN" for network slice subnet is the sum of average DL delay in gNB-CU-UP of the network slice subnet (DLDelay_gNBCUUP_Nss) and the average DL delay in gNB-DU of the network slice subnet (DLDelay_gNBDU_Nss):

   DLDelay_NR_Nss.*SNSSAI* = DLDelay_gNBCUUP_Nss.*SNSSAI* + DLDelay_gNBDU_Nss.*SNSSAI* where *SNSSAI* identifies the S-NSSAI that the network slice subnet supports.

d) NetworkSliceSubnet

6.3.1.3 Downlink delay in gNB-DU

6.3.1.3.1 Downlink delay in gNB-DU for a NRCellDU

a) DLDelay_gNBDU_Cell.

b) This KPI describes the average packet transmission delay through the gNB-DU part to the UE. It is used to evaluate delay performance of gNB-DU in downlink. It is the average packets delay time from arrival of an RLC SDU at the RLC ingress F1-U termination until the last part of an RLC SDU packet was received by the UE according to received HARQ feedback information for UM mode or until the last part of an RLC SDU packet was received by the UE according to received RLC ACK for AM mode. It is a Time interval (0.1 ms). The KPI type is MEAN. This KPI can optionally be split into KPIs per QoS level (mapped 5QI or QCI in NR option 3) and per S-NSSAI.

c) Below is the equation for average DL delay in gNB-DU for a NRCellDU:

DLDelay_gNBDU_Cell = DRB.RlcSduDelayDl + DRB.AirIfDelayDl.

and optionally: DLDelay_gNBDU.*QOS* = DRB.RlcSduDelayDl.*QOS* + DRB.AirIfDelayDl.*QOS* where *QOS* identifies the target quality of service class.

and optionally: DLDelay_gNB.*SNSSAI* = DRB.RlcSduDelayDl.*SNSSAI* + DRB.AirIfDelayDl.*SNSSAI* where *SNSSAI* identifies the S-NSSAI

d) NRCellDU

6.3.1.3.2 Downlink delay in gNB-DU for a sub-network

a) DLDelay_gNBDU_SNw.

b) This KPI describes the average packet transmission delay through the gNB-DU part to the UE. It is used to evaluate delay performance of gNB-DU in downlink for a sub-network. It is the weighted average packets delay time from arrival of an RLC SDU at the RLC ingress F1-U termination until the last part of an RLC SDU packet was received by the UE according to received HARQ feedback information for UM mode or until the last part of an RLC SDU packet was received by the UE according to received RLC ACK for AM mode. It is a Time interval (0.1 ms). The KPI type is MEAN. This KPI can optionally be split into KPIs per QoS level (mapped 5QI or QCI in NR option 3) and per S-NSSAI.

c) Below is the equation for average DL delay in gNB-DU for a sub-network, where

- W is the measurement for the weighted average, one of the following:
  - the DL data volume of the NR cell;
  - the number of UL user data packets of the NR cell;
  - any other types of weight defined by the consumer of KPI
- the #NRCellDU is the number of NRCellDU's in the SubNetwork.

$$\text{DLDelay\_gNBDU\_SNw} = \frac{\sum_{1}^{\#\text{NRCellDU}}\left((\text{DRB.RlcSduDelayDl} + \text{DRB.AirIfDelayDl}) * \text{W}\right)}{\sum_{1}^{\#\text{NRCellDU}}(\text{W})}$$

and optionally KPI on SubNetwork level per QoS and per S-NSSAI:

$$\text{DLDelay\_gNBDU\_SNw}.QoS = \frac{\sum_{1}^{\#\text{NRCellDU}}\left((\text{DRB.RlcSduDelayDl}.QoS + \text{DRB.AirIfDelayDl}.QoS) * \text{W}.QoS\right)}{\sum_{1}^{\#\text{NRCellDU}}(\text{W}.QoS)}$$

$$\text{DLDelay\_gNBDU\_SNw}.SNSSAI = \frac{\sum_{1}^{\#\text{NRCellDU}}\left((\text{DRB.RlcSduDelayDl}.SNSSAI + \text{DRB.AirIfDelayDl}.SNSSAI) * \text{W}.SNSSAI\right)}{\sum_{1}^{\#\text{NRCellDU}}(\text{W}.SNSSAI)}$$

d) SubNetwork

6.3.1.3.3 Downlink delay in gNB-DU for a network slice subnet

a) DLDelay_gNBDU_Nss.

b) This KPI describes the average packet transmission delay through the gNB-DU part to the UE. It is used to evaluate delay performance of gNB-DU in downlink for a network slice subnet. It is the weighted average packets delay time from arrival of an RLC SDU at the RLC ingress F1-U termination until the last part of an RLC SDU packet was received by the UE according to received HARQ feedback information for UM mode or until the last part of an RLC SDU packet was received by the UE according to received RLC ACK for AM mode. It is a Time interval (0.1 ms). The KPI type is MEAN.

c) Below is the equation for average DL delay in gNB-DU for a network slice subnet, where

- W is the measurement for the weighted average, one of the following:
  - the DL data volume of the NR cell;
  - the number of DL user data packets of the NR cell;
  - any other types of weight requested by the consumer of KPI;
- the #NRCellDU is the number of NRCellDU’s associated with the NetworkSliceSubnet.

$$\text{DLDelay\_gNBDU\_Nss}.\mathit{SNSSAI} = \frac{\sum_1^{\#\text{NRCellDU}}((\text{DRB.RlcSduDelayDl}.\mathit{SNSSAI} + \text{DRB.AirIfDelayDl}.\mathit{SNSSAI}) * \text{W}.\mathit{SNSSAI})}{\sum_1^{\#\text{NRCellDU}}(\text{W}.\mathit{SNSSAI})}$$

d) NetworkSliceSubnet

6.3.1.4 Downlink delay in gNB-CU-UP

6.3.1.4.1 Downlink delay in gNB-CU-UP

a) DLDelay_gNBCUUP.

b) This KPI describes the average packet transmission delay through the gNB-CU-UP to the gNB-DU. It is used to evaluate the delay performance of gNB-CU-UP in downlink. It is the average packets delay from reception of IP packet in gNB-CU-UP until the time of arrival, at the gNB-DU, of the RLC SDU at the RLC ingress F1-U termination. It is a Time interval (0.1 ms). The KPI type is MEAN. This KPI can optionally be split into KPIs per QoS level (mapped 5QI or QCI in NR option 3) and per S-NSSAI.

c) Below the equation for average DL delay in a gNB-CU-CP:

DLDelay_gNBCUUP = DRB. PdcpSduDelayDl + DRB.PdcpF1Delay

and optionally: DLDelay_ gNBCUUP.*QOS* = DRB.PdcpSduDelayDl.*QOS* + DRB.PdcpF1Delay.*QOS* where *QOS* identifies the target quality of service class.

and optionally: DLDelay_gNBCUUP.*SNSSAI* = DRB.PdcpSduDelayDl.*SNSSAI* + DRB.PdcpF1Delay.*SNSSAI* where *SNSSAI* identifies the S-NSSAI.

d) GNBCUUPFunction

e) In non-split gNB scenario, the value of DRB.PdcpF1Delay (optionally DRB.PdcpF1Delay.*QOS*, and optionally *DRB.PdcpF*1Delay.*SNSSAI)* is set to zero because there are no F1-interfaces in this scenario.

6.3.1.4.2 Downlink delay in gNB-CU-UP for a sub-network

a) DLDelay_gNBCUUP_SNw.

b) This KPI describes the average packet transmission delay through the gNB-CU-UP to the gNB-DU. It is used to evaluate the delay performance of gNB-CU-UP in downlink for a sub-network. It is the weighted average packets delay from reception of IP packet in gNB-CU-UP until the time of arrival, at the gNB-DU, of the RLC SDU at the RLC ingress F1-U termination. It is a Time interval (0.1 ms). The KPI type is MEAN. This KPI can optionally be split into KPIs per QoS level (mapped 5QI or QCI in NR option 3) and per S-NSSAI.

c) Below is the equation for average UL delay in gNB-CU-UP for a sub-network, where

- W is the measurement for the weighted average, one of the following:
  - the DL data volume in gNB-CU-UP;
  - the number of DL user data packets in gNB-CU-UP;

- any other types of weight requested by the consumer of KPI;
- the # GNBCUUPFunctions is the number of GNBCUUPFunctions's in the SubNetwork.

$$\text{DLDelay\_gNBCUUP\_SNw} = \frac{\sum_{1}^{\#\text{GNBCUUPFunction}} ((\text{DRB.PdcpSduDelayDl} + \text{DRB.PdcpF1Delay}) * \text{W})}{\sum_{1}^{\#\text{GNBCUUPFunction}} (\text{W})}$$

and optionally KPI on SubNetwork level per QoS and per S-NSSAI:

$$\text{DLDelay\_gNBCUUP\_SNw}.QoS = \frac{\sum_{1}^{\#\text{GNBCUUPFunction}} ((\text{DRB.PdcpSduDelayDl}.QoS + \text{DRB.PdcpF1Delay.QoS}) * \text{W}.QoS)}{\sum_{1}^{\#\text{GNBCUUPFunction}} (\text{W}.QoS)}$$

$$\text{DLDelay\_gNBCUUP\_SNw}.SNSSAI = \frac{\sum_{1}^{\#\text{GNBCUUPFunction}} ((\text{DRB.PdcpSduDelayDl}.SNSSAI + \text{DRB.PdcpF1Delay}.SNSSAI) * \text{W}.SNSSAI)}{\sum_{1}^{\#\text{GNBCUUPFunction}} (\text{W}.SNSSAI)}$$

d) SubNetwork

e) In non-split gNB scenario, the value of DRB.PdcpF1Delay (optionally DRB.PdcpF1Delay.*QOS*, and optionally DRB.PdcpF1Delay.*SNSSAI)* is set to zero because there are no F1-interfaces in this scenario.

6.3.1.4.3 Downlink delay in gNB-CU-UP for a network slice subnet

a) DLDelay_gNBCUUP_Nss.

b) This KPI describes the average packet transmission delay through the gNB-CU-UP to gNB-DU. It is used to evaluate the delay performance of gNB-CU-UP in downlink for a network slice subnet. It is the weighted average packets delay from reception of IP packet in gNB-CU-UP until the time of arrival, at the gNB-DU, of the RLC SDU at the RLC ingress F1-U termination. It is a Time interval (0.1 ms). The KPI type is MEAN.

c) Below is the equation for average UL delay in gNB-CU-UP for a network slice subnet, where

- W is the measurement for the weighted average, one of the following:
  - the DL data volume in gNB-CU-UP;
  - the number of DL user data packets in gNB-CU-UP;
  - any other types of weight requested by the consumer of KPI;
- the # GNBCUUPFunctions is the number of GNBCUUPFunctions's associated with the NetworkSliceSubnet.

$$\text{DLDelay\_gNBCUUP\_Nss}.SNSSAI = \frac{\sum_{1}^{\#\text{GNBCUUPFunction}} ((\text{DRB.PdcpSduDelayDl}.SNSSAI + \text{DRB.PdcpF1Delay}.SNSSAI) * \text{W}.SNSSAI)}{\sum_{1}^{\#\text{GNBCUUPFunction}} (\text{W}.SNSSAI)}$$

d) NetworkSliceSubnet

e) In non-split gNB scenario, the value of DRB.PdcpF1Delay.*SNSSAI* is set to zero because there are no F1-interfaces in this scenario.

6.3.1.5 Uplink delay in gNB-DU

6.3.1.5.1 Uplink delay in gNB-DU for a NR cell

a) ULDelay_gNBDU_Cell.

b) This KPI describes the average packet transmission delay through the gNB-DU part from the UE in a NR cell. It is used to evaluate delay performance of gNB-DU in uplink. It is the average packet delay from when an UL RLC SDU was scheduled, as per the scheduling grant provided, until time when the RLC SDU is sent to PDCP or CU for split gNB. It is a time interval (0.1 ms). The KPI type is MEAN. This KPI can optionally be split into KPIs per QoS level (mapped 5QI or QCI in NR option 3) and per S-NSSAI.

c) Below is the equation for average UL delay in gNB-DU for a NRCellDU:

ULDelay_gNBDU_Cell = DRB.RlcDelayUI + DRB.AirIfDelayUI

and optionally: $\text{ULDelay\_gNBDU}.QoS = \text{DRB.RlcDelayUl}.QOS + \text{DRB.AirIfDelayUl}.QOS$ where *QOS* identifies the target quality of service class.

and optionally: $\text{ULDelay\_gNBDU}.SNSSAI = \text{DRB.RlcDelayUl}.SNSSAI + \text{DRB.AirIfDelayUl}.SNSSAI$ where *SNSSAI* identifies the S-NSSAI.

d) NRCellDU

6.3.1.5.2 Uplink delay in gNB-DU for a sub-network

a) ULDelay_gNBDU_SNw.

b) This KPI describes the average packet transmission delay through the gNB-DU part from the UE for a sub-network. It is used to evaluate delay performance of gNB-DU in uplink for a sub-network. It is the weighted average packet delay from when an UL RLC SDU was scheduled, as per the scheduling grant provided, until time when the RLC SDU is sent to PDCP or CU for split gNB. It is a time interval (0.1 ms). The KPI type is MEAN. This KPI can optionally be split into KPIs per QoS level (mapped 5QI or QCI in NR option 3) and per S-NSSAI.

c) Below is the equation for average UL delay in gNB-DU for a sub-network, where

- W is the measurement for the weighted average, one of the following:
  - the UL data volume of the NR cell;
  - the number of UL user data packets of the NR cell;
  - any other types of weight defined by the consumer of KPI
- the #NRCellDU is the number of NRCellDU's in the SubNetwork.

$$\text{ULDelay\_gNBDU\_SNw} = \frac{\sum_{1}^{\#\text{NRCellDU}}\left((\text{DRB.RlcDelayUl} + \text{DRB.AirIfDelayUl}) * \text{W}\right)}{\sum_{1}^{\#\text{NRCellDU}}(\text{W})}$$

and optionally KPI on SubNetwork level per QoS and per S-NSSAI:

$$\text{ULDelay\_gNBDU\_SNw}.QoS = \frac{\sum_{1}^{\#\text{NRCellDU}}\left((\text{DRB.RlcDelayUl}.QoS + \text{DRB.AirIfDelayUl}.QoS) * \text{W}.QoS\right)}{\sum_{1}^{\#\text{NRCellDU}}(\text{W}.QoS)}$$

$$\text{ULDelay\_gNBDU\_SNw}.SNSSAI = \frac{\sum_{1}^{\#\text{NRCellDU}}\left((\text{DRB.RlcDelayUl}.SNSSAI + \text{DRB.AirIfDelayUl}.SNSSAI) * \text{W}.SNSSAI\right)}{\sum_{1}^{\#\text{NRCellDU}}(\text{W}.SNSSAI)}$$

d) SubNetwork

6.3.1.5.3 Uplink delay in gNB-DU for a network slice subnet

a) ULDelay_gNBDU_Nss.

b) This KPI describes the average packet transmission delay through the gNB-DU part from the UE for a network slice subnet. It is used to evaluate delay performance of gNB-DU in uplink for a network slice subnet. It is the weighted average packet delay from when an UL RLC SDU was scheduled, as per the scheduling grant provided, until time when the RLC SDU is sent to PDCP or CU for split gNB. It is a time interval (0.1 ms). The KPI type is MEAN.

c) Below is the equation for average UL delay in gNB-DU for a network slice subnet, where

- W is the measurement for the weighted average, one of the following:
  - the UL data volume of the NR cell;
  - the number of UL user data packets of the NR cell;

    - any other types of weight requested by the consumer of KPI;
  - the #NRCellDU is the number of NRCellDU's associated with the NetworkSliceSubnet.

$$\text{ULDelay\_gNBDU\_Nss}.SNSSAI = \frac{\sum_1^{\#\text{NRCellDU}}((\text{DRB.RlcDelayUl}.SNSSAI + \text{DRB.AirIfDelayUl}.SNSSAI) * \text{W}.SNSSAI)}{\sum_1^{\#\text{NRCellDU}}(\text{W}.SNSSAI)}$$

d) NetworkSliceSubnet

6.3.1.6 Uplink delay in gNB-CU-UP

6.3.1.6.1 Uplink delay in gNB-CU-UP

a) ULDelay_gNBCUUP.

b) This KPI describes the average packet transmission delay through the gNB-CU-UP from gNB-DU. It is used to evaluate delay performance of gNB-CU-UP in uplink. It is the average packet delay from when the RLC SDU is sent to PDCP or CU for split gNB, until time when the corresponding PDCP SDU was sent to the core network from gNB-CU-UP. It is a time interval (0.1 ms). The KPI type is MEAN. This KPI can optionally be split into KPIs per QoS level (mapped 5QI or QCI in NR option 3) and per S-NSSAI.

c) Below the equation for average UL delay in a gNB-CU-CP:

$$\text{ULDelay\_gNBCUUP} = \text{DRB.PdcpReordDelayUl} + \text{DRB.PdcpF1Delay}$$

and optionally: $\text{ULDelay\_gNBCUUP}.QoS = \text{DRB.PdcpReordDelayUl}.QoS + \text{DRB.PdcpF1Delay}.QoS$ where *QOS* identifies the target quality of service class.

and optionally: $\text{ULDelay\_gNBCUUP}.SNSSAI = \text{DRB.PdcpReordDelayUl}.SNSSAI + \text{DRB.PdcpF1Delay}.SNSSAI$ where *SNSSAI* identifies the S-NSSAI.

d) GNBCUUPFunction

e) It is assumed that the F1 uplink delay is the same as the F1 downlink delay. In non-split gNB scenario, the value of DRB.PdcpF1Delay (optionally DRB.PdcpF1Delay.*QOS*, and optionally *DRB.PdcpF*1Delay.*SNSSAI)* is set to zero because there are no F1-interfaces in this scenario.

6.3.1.6.2 Uplink delay in gNB-CU-UP for a sub-network

a) ULDelay_gNBCUUP_SNw.

b) This KPI describes the average packet transmission delay through the gNB-CU-UP part from the gNB-DU for a sub-network. It is used to evaluate delay performance of gNB-CU-UP in uplink for a sub-network. It is the weighted average packet delay from when the RLC SDU is sent to PDCP or CU for split gNB, until time when the corresponding PDCP SDU was sent to the core network from gNB-CU-UP. It is a time interval (0.1 ms). The KPI type is MEAN. This KPI can optionally be split into KPIs per QoS level (mapped 5QI or QCI in NR option 3) and per S-NSSAI.

c) Below is the equation for average UL delay in gNB-CU-UP for a sub-network, where

  - W is the measurement for the weighted average, one of the following:
    - the UL data volume in gNB-CU-UP ;
    - the number of UL user data packets in gNB-CU-UP ;
    - any other types of weight requested by the consumer of KPI;
  - the # GNBCUUPFunctions is the number of GNBCUUPFunctions's in the SubNetwork.

$$\text{ULDelay\_gNBCUUP\_SNw} = \frac{\sum_1^{\#\text{GNBCUUPFunction}}((\text{DRB.PdcpReordDelayUl} + \text{DRB.PdcpF1Delay}) * \text{W})}{\sum_1^{\#\text{GNBCUUPFunction}}(\text{W})}$$

and optionally KPI on SubNetwork level per QoS and per S-NSSAI:

$$\text{ULDelay\_gNBCUUP\_SNw}.QoS = \frac{\sum_{1}^{\#\text{GNBCUUPFunction}}((\text{DRB.PdcpReordDelayUl}.QoS + \text{DRB.PdcpF1Delay.QoS})*\text{W}.QoS)}{\sum_{1}^{\#\text{GNBCUUPFunction}}(\text{W}.QoS)}$$

$$\text{ULDelay\_gNBCUUP\_SNw}.SNSSAI = \frac{\sum_{1}^{\#\text{GNBCUUPFunction}}((\text{DRB.PdcpReordDelayUl}.SNSSAI + \text{DRB.PdcpF1Delay}.SNSSAI)*\text{W}.SNSSAI)}{\sum_{1}^{\#\text{GNBCUUPFunction}}(\text{W}.SNSSAI)}$$

d) SubNetwork

e) It is assumed that the F1 uplink delay is the same as the F1 downlink delay. In non-split gNB scenario, the value of DRB.PdcpF1Delay (optionally DRB.PdcpF1Delay.*QOS*, and optionally DRB.PdcpF1Delay.*SNSSAI)* is set to zero because there are no F1-interfaces in this scenario.

6.3.1.6.3 Uplink delay in gNB-CU-UP for a network slice subnet

a) ULDelay_gNBCUUP_Nss.

b) This KPI describes the average packet transmission delay through the gNB-CU-UP part from the gNB-DU for a network slice subnet. It is used to evaluate delay performance of gNB-CU-UP in uplink for a network slice subnet. It is the weighted average packet delay from when the RLC SDU is sent to PDCP or CU for split gNB, until time when the corresponding PDCP SDU was sent to the core network from gNB-CU-UP. It is a time interval (0.1 ms). The KPI type is MEAN. This KPI can optionally be split into KPIs per QoS level (mapped 5QI or QCI in NR option 3) and per S-NSSAI.

c) Below is the equation for average UL delay in gNB-CU-UP for a network slice subnet, where

- W is the measurement for the weighted average, one of the following:
  - the UL data volume in gNB-CU-UP;
  - the number of UL user data packets in gNB-CU-UP;
  - any other types of weight requested by the consumer of KPI;
- the # GNBCUUPFunctions is the number of GNBCUUPFunctions’s associated with the NetworkSliceSubnet.

$$\text{ULDelay\_gNBCUUP\_Nss}.SNSSAI = \frac{\sum_{1}^{\#\text{GNBCUUPFunction}}((\text{DRB.PdcpReordDelayUl}.SNSSAI + \text{DRB.PdcpF1Delay}.SNSSAI)*\text{W}.SNSSAI)}{\sum_{1}^{\#\text{GNBCUUPFunction}}(\text{W}.SNSSAI)}$$

d) NetworkSliceSubnet

e) It is assumed that the F1 uplink delay is the same as the F1 downlink delay. In non-split gNB scenario, the value of DRB.PdcpF1Delay.*SNSSAI* is set to zero because there are no F1-interfaces in this scenario.

6.3.1.7 Integrated uplink delay in RAN

6.3.1.7.1 Uplink delay in NG-RAN for a sub-network

a) ULDelay_NR_SNw.

b) This KPI describes the average packet transmission delay through the RAN part from the UE for a sub-network. It is used to evaluate delay performance of NG-RAN in uplink. It is the weighted average packet delay from when an UL RLC SDU was scheduled, as per the scheduling grant provided, until time when the corresponding PDCP SDU was sent to the core network from gNB-CU-UP. It is a time interval (0.1 ms). The KPI type is MEAN. This KPI can optionally be split into KPIs per QoS level (mapped 5QI or QCI in NR option 3) and per S-NSSAI.

c) Below are the equations for average "Integrated uplink delay in RAN" for this KPI on SubNetwork level. The "Integrated uplink delay in RAN" is the sum of average UL delay in gNB-CU-UP of the sub-network (ULDelay_gNBCUUP_SNw) and the average UL delay in gNB-DU of the sub-network (ULDelay_gNBDU_SNw):

ULDelay_NR_SNw = ULDelay_gNBCUUP_SNw + ULDelay_gNBDU_SNw

or optionally ULDelay_ NR_SNw.*QOS* = ULDelay_gNBCUUP_SNw.*QOS* + ULDelay_gNBDU_SNw.*QOS* where *QOS* identifies the target quality of service class.

or optionally ULDelay_NR_SNw.*SNSSAI* = ULDelay_gNBCUUP_SNw.*SNSSAI* + ULDelay_gNBDU_SNw.*SNSSAI* where *SNSSAI* identifies the S-NSSAI.

d) SubNetwork

6.3.1.7.2 Uplink delay in NG-RAN for a network slice subnet

a) ULDelay_NR_Nss.

b) This KPI describes the average packet transmission delay through the RAN part from the UE for a network slice subnet. It is used to evaluate delay performance of NG-RAN in uplink. It is the weighted average packet delay from when an UL RLC SDU was scheduled, as per the scheduling grant provided, until time when the corresponding PDCP SDU was sent to the core network from gNB-CU-UP. It is a time interval (0.1 ms). The KPI type is MEAN. This KPI can optionally be split into KPIs per QoS level (mapped 5QI or QCI in NR option 3) and per S-NSSAI.

c) Below is the equation for average "Integrated uplink delay in RAN" for this KPI on NetworkSliceSubNet level. The "Integrated uplink delay in RAN" for network slice subnet is the sum of average UL delay in gNB-CU-UP of the network slice subnet (ULDelay_gNBCUUP_Nss) and the average UL delay in gNB-DU of the network slice subnet (ULDelay_gNBDU_Nss):

ULDelay_NR_Nss.*SNSSAI* = ULDelay_gNBCUUP_Nss.*SNSSAI* + ULDelay_gNBDU_Nss.*SNSSAI* where *SNSSAI* identifies the S-NSSAI that the network slice subnet supports.

d) NetworkSliceSubnet

6.3.1.8 E2E delay for network slice

6.3.1.8.1 Average e2e uplink delay for a network slice

a) DelayE2EUlNs.

b) This KPI describes the average e2e UL packet delay between the PSA UPF and the UE for a network slice. It is the weighted average packet delay from the time when an UL RLC SDU was scheduled at the UE until the time when the corresponding GTP PDU was received by the PSA UPF. The KPI type is MEAN in unit of 0.1 ms.

c) This KPI is the weighted average of UL packet delay between PSA UPF and UE, for all N3 interfaces (modelled by EP_N3 MOIs) and N9 interfaces (modelled by EP_N9 MOIs) of all PSA UPFs supporting the network slice (modelled by NetworkSlice MOI) identified by the S-NSSAI.

This KPI is calculated in the equation below, where Wn3 and Wn9 are the measurements for the weighted average, Wn3 is one of the following:

- the data volume of UL GTP PDUs received by PSA UPF on the N3 interface;
- the number of UL GTP PDUs received by PSA UPF on the N3 interface;
- any other types of weight defined by the consumer of KPI.

And Wn9 is one of the following:

- the data volume of UL GTP PDUs received by PSA UPF on the N9 interface;
- the number of UL GTP PDUs received by PSA UPF on the N9 interface;
- any other types of weight defined by the consumer of KPI.

DelayE2EUlNs =

$$\frac{\sum_{EP\_N3}(\text{GTP.DelayUlPsaUpfUeMean}.SNSSAI * \text{Wn3}.SNSSAI) + \sum_{EP\_N9}(\text{GTP.DelayUlPsaUpfUeMean}.SNSSAI * \text{Wn9}.SNSSAI)}{\sum_{EP\_N3}\text{Wn3}.SNSSAI + \sum_{EP\_N9}\text{Wn9}.SNSSAI}$$

Where the *SNSSAI* identifies the S-NSSAI.

d) NetworkSlice.

6.3.1.8.2 Average e2e downlink delay for a network slice

a) DelayE2EDlNs.

b) This KPI describes the average e2e DL packet delay between the PSA UPF and the UE for a network slice. It is the weighted average packet delay from the time when an GTP PDU has been sent by the PSA UPF until time when the corresponding RLC SDU was received by the UE. The KPI type is MEAN in unit of 0.1 ms.

c) This KPI is the weighted average of DL packet delay between PSA UPF and UE, for all N3 interfaces (modelled by EP_N3 MOIs) and N9 interfaces (modelled by EP_N9 MOIs) of all PSA UPFs supporting the network slice (modelled by NetworkSlice MOI) identified by the S-NSSAI.

This KPI is calculated in the equation below, where Wn3 and Wn9 are the measurements for the weighted average, Wn3 is one of the following:

- the data volume of DL GTP PDUs transmitted by PSA UPF on the N3 interface;
- the number of DL GTP PDUs transmitted by PSA UPF on the N3 interface;
- any other types of weight defined by the consumer of KPI.

And Wn9 is one of the following:

- the data volume of DL GTP PDUs transmitted by PSA UPF on the N9 interface;
- the number of DL GTP PDUs transmitted by PSA UPF on the N9 interface;
- any other types of weight defined by the consumer of KPI.

DelayE2EDlNs =

$$\frac{\sum_{EP\_N3}(\text{GTP.DelayDlPsaUpfUeMean}.SNSSAI * \text{GTP.Wn3}.SNSSAI) + \sum_{EP\_N9}(\text{GTP.DelayDlPsaUpfUeMean}.SNSSAI * \text{Wn9}.SNSSAI)}{\sum_{EP\_N3}\text{Wn3}.SNSSAI + \sum_{EP\_N9}\text{Wn9}.SNSSAI}$$

Where the *SNSSAI* identifies the S-NSSAI.

d) NetworkSlice.

6.3.2 Upstream throughput for network and Network Slice Instance

a) UTSNSI.

b) This KPI describes the upstream throughput of one single network slice by computing the packet size for each successfully received UL IP packet through the network slice during each observing granularity period and is used to evaluate integrity performance of the end-to-end network slice . It is obtained by upstream throughput provided by N3 interface from NG-RAN to all UPFs which are related to the single network slice . The KPI unit is kbit/s and the KPI type is CUM.

c) $$UTSNSI = \sum_{UPF} GTP.InDataOctN3UPF.SNSSAI$$

d) NetworkSlice, SubNetwork.

6.3.3 Downstream throughput for Single Network Slice Instance

a) DTSNSI.

b) This KPI describes the downstream throughput of one single network slice instance by computing the packet size for each successfully transmitted DL IP packet through the network slice instance during each observing granularity period and is used to evaluate integrity performance of the end-to-end network slice instance. It is obtained by downstream throughput provided by N3 interface from all UPFs to NG-RAN which are related to the single network slice . The KPI unit is kbit/s and the KPI type is CUM.

c) $$DTSNSI = \sum_{UPF} GTP.OutDataOctN3UPF.SNSSAI$$

d) NetworkSlice.

6.3.4 Upstream Throughput at N3 interface

a) N3UpstreamThr.

b) This KPI describes the total number of octets of all incoming GTP data packets on the N3 interface (measured at UPF) which have been generated by the GTP-U protocol entity on the N3 interface, during a granularity period. This KPI is used to evaluate upstream GTP throughput integrity performance at the N3 interface. It is obtained by measuring the GTP data upstream throughput provided by N3 interface from NG-RAN to UPF, during the granularity period. The KPI unit is kbit/s and the KPI type is MEAN.

c) *UGTPTS=SUM (GTP.InDataOctN3UPF)/timeperiod) at UPF*

d) NetworkSlice

6.3.5 Downstream Throughput at N3 interface

a) N3DownstreamThr.

b) This KPI describes the total number of octets of all downstream GTP data packets on the N3 interface (transmitted downstream from UPF) which have been generated by the GTP-U protocol entity on the N3 interface, during a granularity period. This KPI is used to evaluate integrity performance at N3 interface. It is obtained by measuring the GTP data downstream throughput provided by N3 interface from UPF to NG-RAN, during the granularity period. The KPI unit is kbit/s and the KPI type is MEAN.

c) DGTPTS=SUM (GTP.OutDataOctN3UPF)/timeperiod) at UPF

d) NetworkSlice

6.3.6 RAN UE Throughput

6.3.6.1 Void

6.3.6.2 RAN UE Throughput definition

To achieve a Throughput measurement (below examples are given for DL) that is independent of file size and gives a relevant result it is important to remove the volume and time when the resource on the radio interface is not fully utilized. (Successful transmission, buffer empty in figure 1).

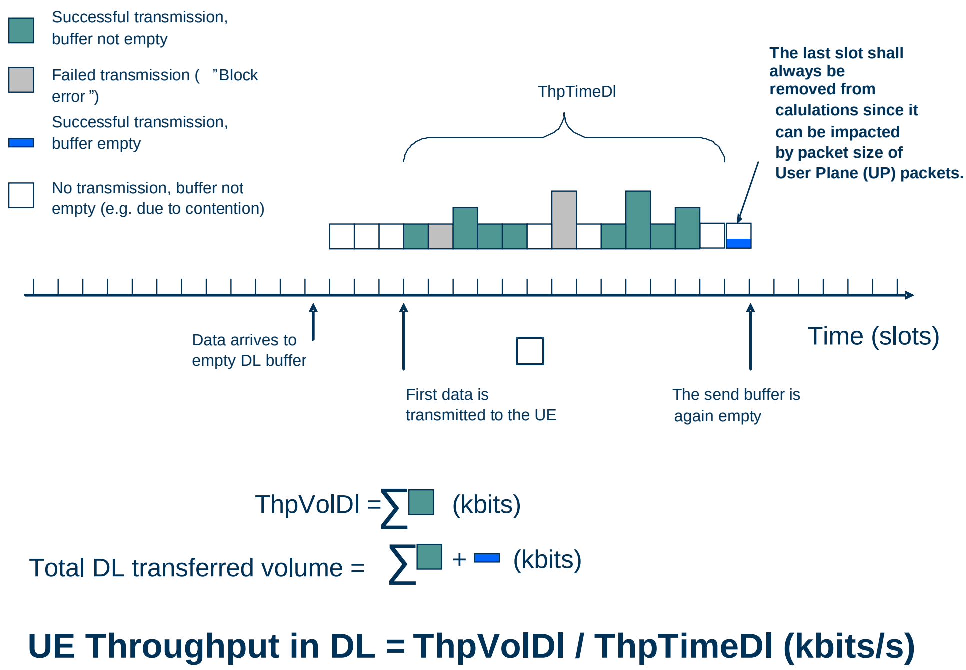


Figure 1

To achieve a throughput measurement that is independent of bursty traffic pattern, it is important to make sure that idle gaps between incoming data is not included in the measurements. That shall be done as considering each burst of data as one sample.

6.3.6.3 DL RAN UE throughput

6.3.6.3.1 DL RAN UE throughput for a NRCellDU

a) DlUeThroughput _Cell.

b) This KPI describes the average DL RAN UE throughput for a NRCellDU. The KPI type is MEAN in kbit per second. This KPI can optionally be split into KPIs per QoS level (mapped 5QI or QCI in NR option 3) and per S-NSSAI.

c) Below is the equation for average DL RAN UE throughput for a NRCellDU:

DlUeThroughput _Cell = DRB.UEThpDl;

and optionally: DlUeThroughput _Cell.*QOS* = DRB.UEThpDl.*QOS*, where *QOS* identifies the target quality of service class;

and optionally: DlUeThroughput _Cell.*SNSSAI* = DRB.UEThpDl.*SNSSAI*, where *SNSSAI* identifies the S-NSSAI.

d) NRCellDU

6.3.6.3.2 DL RAN UE throughput for a sub-network

a) DlUeThroughput _SNw.

b) This KPI describes the average DL RAN UE throughput for a sub-network. The KPI type is MEAN in kbit per second. This KPI can optionally be split into KPIs per QoS level (mapped 5QI or QCI in NR option 3) and per S-NSSAI.

c) Below is the equation for average DL RAN UE throughput for a sub-network, where

- W is the measurement for the weighted average, it is one of the following:

    - the DL data volume of the NR cell;
    - a weight defined by the consumer of KPI
  - the #NRCellDU is the number of NRCellDU's in the SubNetwork.

$$\text{DlUeThroughput\_SNw} = \frac{\sum_{1}^{\#\text{NRCellDU}}(\text{W})}{\sum_{1}^{\#\text{NRCellDU}}\left(\frac{\text{W}}{\text{DRB.UEThpDl}}\right)}$$

and optionally KPI on SubNetwork level per QoS and per S-NSSAI:

$$\text{DlUeThroughput\_SNw}.QoS = \frac{\sum_{1}^{\#\text{NRCellDU}}(\text{W}.QoS)}{\sum_{1}^{\#\text{NRCellDU}}\left(\frac{\text{W}.QoS}{\text{DRB.UEThpDl}.QoS}\right)}$$

$$\text{DlUeThroughput\_SNw}.SNSSAI = \frac{\sum_{1}^{\#\text{NRCellDU}}(\text{W}.SNSSAI)}{\sum_{1}^{\#\text{NRCellDU}}\left(\frac{\text{W}.SNSSAI}{\text{DRB.UEThpDl}.SNSSAI}\right)}$$

d) SubNetwork

6.3.6.3.3 DL RAN UE throughput for a network slice subnet

a) DlUeThroughput _Nss.

b) This KPI describes the average DL RAN UE throughput for a network slice subnet. The KPI type is MEAN in kbit per second.

c) Below is the equation for average DL RAN UE throughput for a network slice subnet, where
  - W is the measurement for the weighted average, it is one of the following:
    - the DL data volume of the NR cell;
    - a weight defined by the consumer of KPI
  - the #NRCellDU is the number of NRCellDU's associated with the NetworkSliceSubnet.

$\text{DlUeThroughput\_Nss}.SNSSAI = \frac{\sum_{1}^{\#\text{NRCellDU}}(\text{W}.SNSSAI)}{\sum_{1}^{\#\text{NRCellDU}}\left(\frac{\text{W}.SNSSAI}{\text{DRB.UEThpDl}.SNSSAI}\right)}$, where the *SNSSAI* identifies the S-NSSAI that the NetworkSliceSubnet supports.

d) NetworkSliceSubnet

6.3.6.4 UL RAN UE throughput

6.3.6.4.1 UL RAN UE throughput for a NRCellDU

a) UlUeThroughput_Cell.

b) This KPI describes the average UL RAN UE throughput for a NRCellDU. The KPI type is MEAN in kbit per second. This KPI can optionally be split into KPIs per QoS level (mapped 5QI or QCI in NR option 3) and per S-NSSAI.

c) Below is the equation for average UL RAN UE throughput for a NRCellDU:

UlUeThroughput _Cell = DRB.UEThpUl;

and optionally: UlUeThroughput_Cell.*QOS* = DRB.UEThpUl.*QOS*, where *QOS* identifies the target quality of service class;

and optionally: UlUeThroughput_Cell.*SNSSAI* = DRB.UEThpUl.*SNSSAI*, where *SNSSAI* identifies the S-NSSAI.

d) NRCellDU

6.3.6.4.2 UL RAN UE throughput for a sub-network

a) UlUeThroughput_SNw.

b) This KPI describes the average UL RAN UE throughput for a sub-network. The KPI type is MEAN in kbit per second. This KPI can optionally be split into KPIs per QoS level (mapped 5QI or QCI in NR option 3) and per S-NSSAI.

c) Below is the equation for average UL RAN UE throughput for a sub-network, where

- W is the measurement for the weighted average, it is one of the following:
  - the UL data volume of the NR cell;
  - a weight defined by the consumer of KPI
- the #NRCellDU is the number of NRCellDU's in the SubNetwork.

$$\text{UlUeThroughput\_SNw} = \frac{\sum_{1}^{\#\text{NRCellDU}}(\text{W})}{\sum_{1}^{\#\text{NRCellDU}}\left(\frac{\text{W}}{\text{DRB.UEThpUl}}\right)}$$

and optionally KPI on SubNetwork level per QoS and per S-NSSAI:

$$\text{UlUeThroughput\_SNw}.QoS = \frac{\sum_{1}^{\#\text{NRCellDU}}(\text{W}.QoS)}{\sum_{1}^{\#\text{NRCellDU}}\left(\frac{\text{W}.QoS}{\text{DRB.UEThpUl}.QoS}\right)}$$

$$\text{UlUeThroughput\_SNw}.SNSSAI = \frac{\sum_{1}^{\#\text{NRCellDU}}(\text{W}.SNSSAI)}{\sum_{1}^{\#\text{NRCellDU}}\left(\frac{\text{W}.SNSSAI}{\text{DRB.UEThpUl}.SNSSAI}\right)}$$

d) SubNetwork

6.3.6.4.3 UL RAN UE throughput for a network slice subnet

a) UlUeThroughput _Nss.

b) This KPI describes the average UL RAN UE throughput for a network slice subnet. The KPI type is MEAN in kbit per second.

c) Below is the equation for average UL RAN UE throughput for a network slice subnet, where

- W is the measurement for the weighted average, it is one of the following:
  - the UL data volume of the NR cell;
  - a weight defined by the consumer of KPI
- the #NRCellDU is the number of NRCellDU's associated with the NetworkSliceSubnet.

$\text{UlUeThroughput\_Nss}.SNSSAI = \frac{\sum_{1}^{\#\text{NRCellDU}}(\text{W}.SNSSAI)}{\sum_{1}^{\#\text{NRCellDU}}\left(\frac{\text{W}.SNSSAI}{\text{DRB.UEThpUl}.SNSSAI}\right)}$, where the *SNSSAI* identifies the S-NSSAI that the NetworkSliceSubnet supports.

d) NetworkSliceSubnet

6.4 Utilization KPI

6.4.1 Mean number of PDU sessions of network and network Slice Instance

a) PDUSesMeanNbr.

b) This KPI describes the mean number of PDU sessions that are successfully established in a network slice . It is obtained by successful PDU session establishment procedures of SMFs which is related to the network slice . It is an integer. The KPI type is MEAN.

c) $$PDUSesMeanNbr = \sum_{SMF} SM.SessionNbrMean.SNSSAI$$

d) NetworkSlice

6.4.2 Virtualised Resource Utilization of Network Slice Instance

a) VirtualResUtilizaiton.

b) This KPI describes utilization of virtualised resource (e.g. processor, memory, disk) that are allocated to a network slice . It is obtained by the usage of virtualised resource (e.g. processor, memory, disk) divided by the system capacity that allocated to the network slice . It is a percentage, The KPI type is Ratio.

NOTE: In the present document, this KPI is for the scenario when NF is not shared between different network slice .

c) $$VRU_{Processor} = \frac{MeanProcessorUsage}{System\ Capacity_{Processor}} * 100\%$$

$$VRU_{Memory} = \frac{MeanMemoryUsage}{System\ Capacity_{Memory}} * 100\%$$

$$VRU_{Disk} = \frac{MeanDiskUsage}{System\ Capacity_{Disk}} * 100\%$$

d) NetworkSlice

6.4.3 PDU session establishment time of network slice

a) PDUEstTime.

b) This KPI describes the time of successful PDU session establishment which related to one single network slice and is used to evaluate utilization provided by the end-to-end network slice and network performance. It is obtained by measuring the time between the receipt by SMF from AMF of " Nsmf_PDUSession_UpdateSMContext Request ", which includes N2 SM information received from (R)AN to the SMF and the sending of a " Nsmf_PDUSession_CreateSMContext Request or Nsmf_PDUSession_UpdateSMContext Request PDU Session Establishment Request " message from AMF to the SMF. It is a time interval (millisecond). The KPI type is MEAN.

c) PDUEstTime = SM.PduSessionTimeMean.SNSSAI

d) NetworkSlice

6.4.4 Mean number of successful periodic registration updates of Single Network Slice

a) RegUpdMeanNbr.

b) This KPI describes the mean number of successfully periodic registration updates in a network slice at the AMF. It is obtained by summing successful of periodic registration updates at the AMFs which is related to the network slice after registration accept by the AMF to the UE that sent the periodic registration update request. It is an integer. The KPI type is MEAN.

c) $$RegUpdMeanNbr = \sum_{AMF} AM.RegNbrMean.SNSSAI$$

d) NetworkSlice

6.4.5 Maximum number of PDU sessions of network slice

a) PDUSesMaxNbr.

b) This KPI describes the maximum number of PDU sessions that are successfully established in a network slice. It is obtained by successful PDU session establishment procedures of SMFs which is related to the network slice. It is an integer. The KPI type is MEAN.

c) $$PDUSesMaxNbr = \sum_{SMF} SM.SessionNbrMax.SNSSAI$$

d) NetworkSlice

6.5 Retainability KPI

6.5.1 QoS flow Retainability

6.5.1.1 Definition

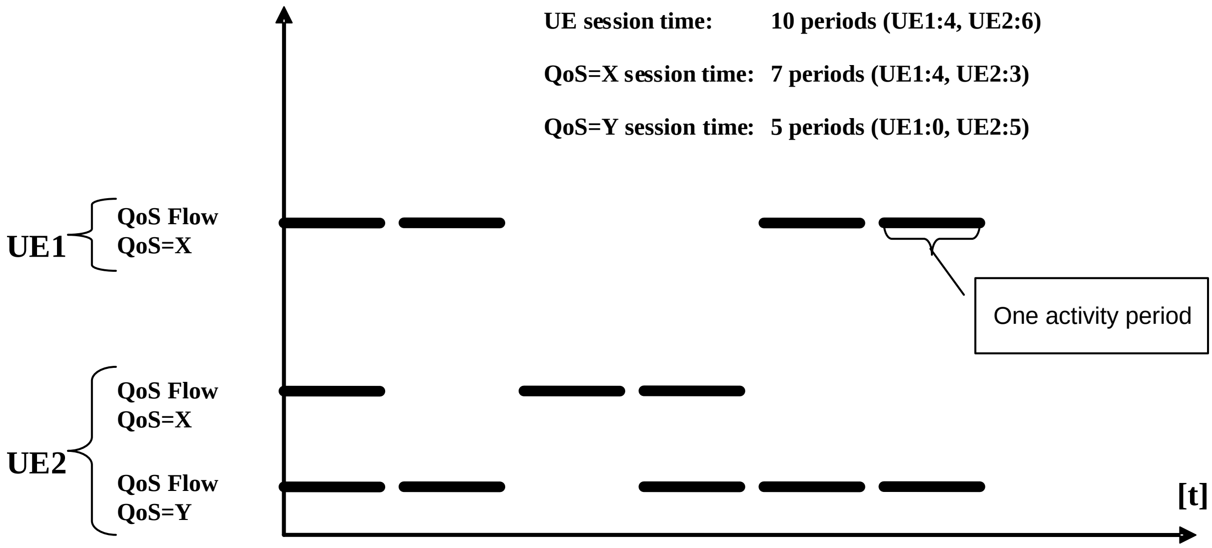
a) QoSRetain_R1, QoSRetain_R2.

b) This KPI shows how often an end-user abnormally loses a QoS flow during the time the QoS flow is used. It is obtained by number of QoS flows with data in a buffer that was abnormally released, normalized with number of data session time units. The unit of this KPI is "active release / second". The KPI type is MEAN.

c) To measure QoS flow Retainability for a single QoS level (R1) is fairly straight forward.

$$R1_{QoS\_x} = \frac{QF.RelActNbr.QoS_{QoS\_x}}{QF.SessionTimeQoS.QoS_{QoS\_x}}$$

However to measure the QoS flow Retainability for UEs is not as straight forward. The measurement R1 is defined to look at the activity level of just one QoS level at the time, so to use this formula and measurements in an aggregated way to get QoS flow Retainability on UE level will not be accurate (e.g. for an UE with multiple QoS flows there might be QoS flows that are active at the same time, hence aggregating the QoS level measurements for session time will give a larger session time than the total UE session time. See picture below).



Hence a measurement QoS flow Retainability on UE level is defined (R2) to provide a measurement for the overall QoS flow Retainability.

$$R2 = \frac{\sum_{QoS} QF.RelActNbr.QoS}{QF.SessionTimeUE}$$

d) SubNetwork, NRCellCU

e) The definition of the service provided by 5GS is QoS flows.

6.5.1.2 Extended definition

The retainability rate is defined as:

$$\frac{\text{Number of abnormally released QoS flow with data in any of the buffers}}{\text{Active QoS flow Time}} \quad \text{[Releases/Session time]}$$

$$\frac{\text{Number of abnormally released QoS flow with data in any of the buffers}}{} \quad \text{[Releases/Session time]}$$

To define (from a QoS flow Retainability point of view) if a QoS flow is considered active or not, the QoS flows can be divided into two groups:

- For QoS flows with bursty flow, a QoS flow is said to be active if there is user data in the PDCP queue in any of the directions or if any data (UL or DL) has been transferred during the last 100 ms.
- For QoS flows with continuous flow, the QoS flow (and the UE) is seen as being active in the context of this measurement as long as the UE is in RRC connected state, and the session time is increased from the first data transmission on the QoS flow until 100 ms after the last data transmission on the QoS flow.

A particular QoS flow is defined to be of type continuous flow if the mapped 5QI is any of {1, 2, 65, 66}.

6.5.2 DRB Retainability

6.5.2.1 Definition

a) DRBRetain

b) This KPI shows how often an end-user abnormally loses a DRB during the time the DRB is active. It is obtained by number of DRBs that were abnormally released and that were active at the time of release, normalized with number of data session time units. The unit of this KPI is "active release / second". The KPI type is MEAN.

c) DRB Retainability for a single mapped 5QI level (R1) and for a single S-NSSAI (R1) are defined as:

$$R1_{5QI\_x} = \frac{DRB.RelActNbr.5QI_{5QI\_x}}{DRB.SessionTime.5QI_{5QI\_x}}$$

and

$$R1_{SNSSAI\_x} = \frac{DRB.RelActNbr.SNSSAI_{SNSSAI\_x}}{DRB.SessionTime.SNSSAI_{SNSSAI\_x}}$$

d) SubNetwork, NRCellCU

e) The definition of the service provided by 5GS is DRBs.

6.5.2.2 Extended definition

To define (from a DRB Retainability point of view) if a DRB is considered active or not, the DRB can be divided into two groups:

- For DRBs with bursty flow, a DRB is said to be active if any data (UL or DL) has been transferred during the last 100 ms.
- For DRBs with continuous flow, the DRB (and the UE) is seen as being active in the context of this measurement as long as the UE is in RRC connected state, and the session time is increased from the first data transmission on the DRB until 100 ms after the last data transmission on the DRB.

A particular DRB is defined to be of type continuous flow if the mapped 5QI is any of {1, 2, 65, 66}.

6.6 Mobility KPI

6.6.1 NG-RAN handover success rate

a) GRANHOSR.

b) A KPI that shows how often a handover within NR-RAN is successful, regardless if the handover was made due to bad coverage or any other reason. This KPI is obtained by successful handovers to the same or another gNB divided by attempted handovers to the same or another gNB. This KPI covers legacy Handover.

c)

$$\text{GRANHOSR} = \frac{(\text{MM.HoExeInterSucc}+\text{MM.HoExeIntraSucc})}{(\text{MM.HoExeInterReq}+\text{MM.HoExeIntraReq})} \times \frac{(\text{MM.HoPrepInterSucc}+\text{MM.HoPrepIntraSucc})}{(\text{MM.HoPrepInterReq}+\text{MM.HoPrepIntraReq})} \times 100[\%]$$

d) SubNetwork, NRCellCU.

6.6.2 Mean Time of Inter-gNB handover Execution of Network Slice

a) InterGNBHOMeanTime.

b) This KPI describes the time of successful Mean Time of Inter-gNB handover which related to one single network slice and is used to evaluate utilization provided by the end-to-end network slice and network performance. This KPI is obtained by measuring the time between the receipt by the Source NG-RAN from the Target NG-RAN of a “Release Resource" and the sending of a " N2 Path Switch Request " message from Source NG-RAN to the Target NG-RAN over a granularity period. The unit of this KPI is millisecond.

d) Subnetwork

6.6.3 Successful rate of mobility registration updates of Single Network Slice

a) MobilityRegUpdateSR.

b) This KPI describes the successful rate of mobility registration updates in a network slice e at the AMF. This KPI is obtained by deviding the number of successful mobility registration updates at the AMFs by number of mobility registration update requests received by the AMFs of single network slice.

d) NetworkSlice

6.6.4 5GS to EPS handover success rate

a) 5GSEPSHOSR.

b) A KPI that shows how often a handover from 5GS to EPS is successful, regardless if the handover was made due to bad coverage or any other reason. This KPI is obtained by successful handovers from 5GS to EPS system divided by the total number of handovers attempt’s from 5GS to EPS system.

c) $$5GSEPSHOSR = \frac{(MM.HoOutExe5gsToEpsSucc)}{(MM.HoOutExe5gsToEpsReq)} \times \frac{(MM.HoOut5gsToEpsPrepSucc)}{(MM.HoOut5gsToEpsPrepReq)} \times 100\,[\%]$$

d) SubNetwork, NRCellCU.

6.6.5 NG-RAN handover success rate for all handover types

a) GRANHOSRA.

b) A KPI that shows how often a handover within NR-RAN is successful, regardless if the handover was made due to bad coverage or any other reason. This KPI is obtained by successful handovers to the same or another gNB divided by attempted handovers to the same or another gNB. This KPI covers legacy Handover, Conditional Handover and DAPS Handover.

c)

$$GRANHOSRA = \frac{MM.HoExeInterSucc + MM.HoExeIntraSucc + MM.ChoExeInterSucc + MM.ChoExeIntraSucc + MM.DapsHoExeInterSucc + MM.DapsHoExeIntraSucc}{MM.HoExeInterReq + MM.HoExeIntraReq + MM.ConfigInterReqChoUes + MM.ConfigIntraReqChoUes + MM.DapsHoExeInterReq + MM.DapsHoExeIntra}$$

$$\times \frac{MM.HoPrepInterSucc + MM.HoPrepIntraSucc + MM.ChoPrepInterSuccUes + MM.ChoPrepIntraSuccUes + MM.DapsHoPrepInterSucc + MM.DapsHoPrepIntraS}{MM.HoPrepInterReq + MM.HoPrepIntraReq + MM.ChoPrepInterReqUes + MM.ChoPrepIntraReqUes + MM.DapsHoPrepInterReq + MM.DapsHoPrepIntraRe}$$

$$\times 100\ [\%]$$

d) SubNetwork, NRCellCU.

6.7 Energy Efficiency (EE) KPI

6.7.1 NG-RAN data Energy Efficiency (EE)

6.7.1.1 Definition

a) $EE_{MN,DV}$.

b) A KPI that shows mobile network data energy efficiency in operational NG-RAN. Data Volume (DV) divided by Energy Consumption (EC) of the considered network elements. The unit of this KPI is bit/J.

c) $EE_{MN,DV}$

$$= \frac{\sum_{Samples}(DRB.PdcpSduVolumeUl + DRB.PdcpSduVolumeDl)}{\sum_{Samples} PEE.Energy}$$ - for non-split gNBs;

$$= \frac{\sum_{Samples} \begin{matrix}(DRB.F1uPdcpSduVolumeUl + DRB.XnuPdcpSduVolumeUl + DRB.X2uPdcpSduVolumeUl + \\ DRB.F1uPdcpSduVolumeDl + DRB.XnuPdcpSduVolumeDl + DRB.X2uPdcpSduVolumeDl)\end{matrix}}{\sum_{Samples} PEE.Energy}$$ - for split gNBs;

d) SubNetwork

e) The Data Volume (in kbits) is obtained by measuring amount of DL/UL PDCP SDU bits of the considered network elements over the measurement period. For split-gNBs, the Data Volume is calculated per Interface (F1-U, Xn-U, X2-U). The Energy Consumption (in kWh) is obtained by measuring the PEE.Energy of the considered network elements over the same period of time. The samples are aggregated at the NG-RAN node level. The 3GPP management system responsible for the management of the gNB (single or multiple vendor gNB) shall be able to collect PEE measurements data from all PNFs in the gNB, in the same way as the other PM measurements.

6.7.2 Network slice Energy Efficiency (EE)

6.7.2.1 Generic Network Slice Energy Efficiency (EE) KPI

$$\text{Generic network slice EE KPI} = \frac{Performance\ of\ network\ slice\ (P_{ns})}{Energy\ Consumption\ of network\ slice\ (EC_{ns})}$$

where:

- 'Performance of network slice' (P_{ns}) is defined per type of network slice;

- 'Energy Consumption of network slice' (EC_{ns}) is defined independently from any type of network slice.

For one unit of EC_{ns}, the higher P_{ns} is, the higher the generic network slice EE KPI is, i.e. the more energy efficient the network slice is.

6.7.2.2 Energy efficiency of eMBB network slice

a) $EE_{eMBB,DV}$

b) A KPI that shows the energy efficiency of network slices of type eMBB. The P_{ns} for a network slice of type eMBB is obtained by summing up UL and DL data volumes at N3 interface(s) of the network slice.

$$P_{eMBB,DV} = \sum_{UPF}(\text{GTP.InDataOctN3UPF}.SNSSAI + \text{GTP.OutDataOctN3UPF}.SNSSAI) * 8$$

, where *SNSSAI* identifies the S-NSSAI.

This KPI is obtained by the sum of UL and DL data volumes at N3 interface(s) of the network slice, divided by the energy consumption of the network slice. The unit of this KPI is bit/J.

c)

$$EE_{eMBB,DV} = \frac{\sum_{UPF}(\text{GTP.InDataOctN3UPF.}SNSSAI + \text{GTP.OutDataOctN3UPF.}SNSSAI) * 8}{EC_{ns}}$$

d) NetworkSlice

e) In case of redundant transmission paths over the N3 interface for high reliability communication (cf. TS 23.501 [7] clause 5.33.2), it is expected that the data volume is counted once. In particular:

- In case of Dual Connectivity based end to end Redundant User Plane Paths (cf. TS 23.501 [7] clause 5.33.2.1), in which a UE may set up two redundant PDU Sessions over the 5G network, the Data Volume related to only one PDU session is to be considered;

- In case of redundant transmission with two N3 tunnels between the PSA UPF and a single NG-RAN node (cf. TS 23.501 [7] figure 5.33.2.2-1) which are associated with a single PDU Session, the Data Volume related to only one N3 tunnel is to be considered;

- In case of two N3 and N9 tunnels between NG-RAN and PSA UPF for redundant transmission (cf. TS 23.501 [7] figure 5.33.2.2-2) associated with a single PDU Session, the Data Volume related to only one N3 tunnel is to be considered.

For the measurement of the energy efficiency of the 5G core network, the 3GPP management system in charge of collecting the data volume measurements listed here above shall consider them only once in case of redundant transmission over the N3 interface.

6.7.2.2a Energy efficiency of eMBB network slice – RAN-based

6.7.2.2a.1 Definition

a) $EE_{RANonlyeMBB,DV}$

b) A KPI that shows the energy efficiency of network slices of type eMBB based on NR measurements. The Pns for a network slice of type eMBB is obtained by summing up UL and DL data volumes at F1-U, Xn-U and X2-U interface(s) of gNBs, on a per S-NSSAI basis.

c)

For non-split gNBs:

$$P_{RANonlyeMBB,DV} = \sum_{Samples} DRB.PdcpSduVolumeUl.SNSSAI + DRB.PdcpSduVolumeDl.SNSSAI$$

, where:

- DRB.PdcpSduVolumeUl.*SNSSAI* is the Data Volume (amount of PDCP SDU bits) in the uplink delivered to PDCP layer per S-NSSAI - see TS 28.552 [4] clause 5.1.2.1.2.1,

- DRB.PdcpSduVolumeDl.*SNSSAI* is the Data Volume (amount of PDCP SDU bits) in the downlink delivered to PDCP layer per S-NSSAI - see TS 28.552 [4] clause 5.1.2.1.1.1.

For split gNBs:

$$P_{RANonlyeMBB,DV} = \sum_{Samples} \begin{matrix} DRB.F1uPdcpSduVolumeDl.SNSSAI + DRB.F1uPdcpSduVolumeUl.SNSSAI + \\ DRB.XnuPdcpSduVolumeDl.SNSSAI + DRB.XnuPdcpSduVolumeUl.SNSSAI + \\ DRB.X2uPdcpSduVolumeDl.SNSSAI + DRB.X2uPdcpSduVolumeUl.SNSSAI \end{matrix}$$

, where:

- DRB.F1uPdcpSduVolumeDl.*SNSSAI* is the number of DL PDCP SDU bits sent to GNB-DU (F1-U interface) per S-NSSAI - see TS 28.552 [4] clause 5.1.3.6.2.3,
- DRB.F1uPdcpSduVolumeUl.*SNSSAI* is the number of UL PDCP SDU bits entering the GNB-CU-UP from GNB-DU (F1-U interface) per S-NSSAI - see TS 28.552 [4] clause 5.1.3.6.2.4,
- DRB.XnuPdcpSduVolumeDl.*SNSSAI* is the number of DL PDCP SDU bits sent to external gNB-CU-UP (Xn-U interface) per S-NSSAI - see TS 28.552 [4] clause 5.1.3.6.2.3,
- DRB.XnuPdcpSduVolumeUl.*SNSSAI* is the number of UL PDCP SDU bits entering the GNB-CU-UP from external gNB-CU-UP (Xn-U interface) per S-NSSAI - see TS 28.552 [4] clause 5.1.3.6.2.4,
- DRB.X2uPdcpSduVolumeDl.*SNSSAI* is the number of DL PDCP SDU bits sent to external eNB (X2-U interface) per S-NSSAI - see TS 28.552 [4] clause 5.1.3.6.2.3,
- DRB.X2uPdcpSduVolumeUl.*SNSSAI* is the number of UL PDCP SDU bits entering the GNB-CU-UP from external eNB (X2-U interface) per S-NSSAI - see TS 28.552 [4] clause 5.1.3.6.2.4.

The final Network Slice EE KPI definition, based on Data Volume, for RAN-only eMBB type of network slice, would be defined as follows:

$$EE_{RANonlyeMBB,\ DV} = \frac{P_{RANonlyeMBB,\ DV}}{EC_{RANonlyns}}$$

, where $EC_{RANonlyns}$ is the energy consumption of the RAN-only network slice over the same observation period.

NOTE: Void

d) NetworkSlice

6.7.2.3 Energy efficiency of URLLC network slice

6.7.2.3.1 Introduction

This KPI is defined with two variants.

6.7.2.3.2 Based on latency of the network slice

a) $EE_{URLLC,Latency}$

b) A KPI that shows the energy efficiency of network slices of type URLLC. The P_{ns} for a network slice of type URLLC is the inverse of the average end-to-end User Plane (UP) latency of the network slice. In this KPI variant, latency are the only factor considered for evaluating the performance of network slice.

$$P_{URLLC,\ Latency} = \frac{1}{Network\ slice\ mean\ latency}$$

, where 'Network slice mean latency' is defined as the average end-to-end User Plane (UP) latency of the network slice, and where the average end-to-end User Plane (UP) latency for one S-NSSAI is defined by:

$$Network\ slice\ mean\ latency = DelayE2EUlNs\ + DelayE2EDlNs$$

This KPI is obtained by the inverse of the average end-to-end User Plane (UP) latency of the network slice divided by the energy consumption of the network slice. The unit of this KPI is $(0.1ms * J)^{-1}$.

c)

$$EE_{URLLC, Latency} = \frac{1}{Network\ slice\ mean\ latency * EC_{ns}}$$

d) NetworkSlice

6.7.2.3.3 Based on both latency and Data Volume (DV) of the network slice

a) $EE_{URLLC,DV,Latency}$

b) A KPI that shows the energy efficiency of network slices of type URLLC. The P_{ns} for a network slice of type URLLC is the sum of UL and DL traffic volumes at N3 or N9 interface(s) on a per S-NSSAI basis multiplied by the inverse of the end-to-end User Plane (UP) latency of the network slice. In this KPI variant, data volume and latency are two factors considered for evaluating the performance of network slice. This KPI is applicable for the cases where, for example, the URLLC network slice is deployed and operators want to evaluate the Energy Efficiency of the slice at different periods of time, such as the busy hours in the morning and the idle hours in the mid night, in which both latency performance and the data volume performance can vary.

$$P_{\mathrm{URLLC},DV,Latency} = \frac{w_{N3} * DV_{N3} + w_{N9} * DV_{N9}}{(DelayE2EUlNs + DelayE2EDlNs)}$$

where

$$DV_{N3} = \sum_{\mathrm{PSA\ UPF}} 8 * (\mathrm{GTP.OutDataOctN3UPF.SNSSAI} + \mathrm{GTP.InDataOctN3UPF.SNSSAI})$$

$$DV_{N9} = \sum_{\mathrm{PSA\ UPF}} 8 * (\mathrm{GTP.OutDataOctN9PsaUpf.SNSSAI} + \mathrm{GTP.InDataOctN9PsaUpf.SNSSAI})$$

w_{N3} and w_{N9} are the weight for DV_{N3} and DV_{N9} respectively. w_{N3} and w_{N9} can be decided according to the deployment of PSA UPF. For example, in cases where PSA UPF has only N9 tunnels, such as the ones described in TS 23.501[2] clause 5.6.4 and clause 5.33.2.2, w_{N3} can be set to 0 and W_{N9} can be set to 1, so that only N9 interface is considered. In the cases where PSA UPF has only N3 tunnels, w_{N3} can be set to 1 and w_{N9} can be set to 0, so that only N3 interface is considered.

This KPI is obtained by the product of the sum of the weighted UL and DL traffic data volumes at N3 interface(s) or N9 interface of the PSA UPF of the network slice multiplied by the inverse of the end-to-end User Plane (UP) latency of the network slice, divided by the energy consumption of the network slice. The unit of this KPI is bit/(0.1ms*J).

c)

$$EE_{\mathrm{URLLC},DV,Latency} = \frac{w_{N3} * DV_{N3} + w_{N9} * DV_{N9}}{(DelayE2EUlNs + DelayE2EDlNs) * EC_{ns}}$$

d) NetworkSlice

e) In case of redundant transmission paths for high reliability communication (TS 23.501 [7] clause 5.33.2), it is expected that the data volume is counted once. In particular:

- In case of Dual Connectivity based end to end Redundant User Plane Paths (TS 23.501 [7] clause 5.33.2.1), in which a UE may set up two redundant PDU Sessions over the 5G network, the Data Volume related to only one PDU session is to be considered;

- In case of redundant transmission with two N3 tunnels between the PSA UPF and a single NG-RAN node (TS 23.501 [7] figure 5.33.2.2-1) which are associated with a single PDU Session, the Data Volume related to only one N3 tunnel is to be considered;

- In case of two N3 and N9 tunnels between NG-RAN and PSA UPF for redundant transmission (TS 23.501 [7] figure 5.33.2.2-2) associated with a single PDU Session, the Data Volume related to only one of the multiple

N3/N9 tunnels for redundant transmission connecting to PSA UPF is considered. The main reason for this is that, if the traffic is counted more than once, it will increase artificially the EEURLLC,DV,Latency KPI.

The 3GPP management system in charge of collecting the data volume measurements listed here above shall consider them only once in case of redundant transmission over the N3/N9 interface.

6.7.2.4 Energy efficiency of MIoT network slice

6.7.2.4.1 Based on the number of registered subscribers of the network slice

a) $EE_{MIoT,RegSubs}$

b) A KPI that shows the energy efficiency of network slices of type MIoT. In this case, the P_{ns} for a network slice of type MIoT is the maximum number of subscribers registered to the network slice.

$$P_{MIoT_{,}RegSubs} = \sum_{AMF} RM.RegisteredSubNbrMax.SNSSAI$$

, where *SNSSAI* identifies the S-NSSAI.

This KPI is obtained by the maximum number of registered subscribers to the network slice divided by the energy consumption of the network slice. The unit of this KPI is user/J.

c)

$$EE_{MIoT_{,}RegSubs} = \frac{\sum_{AMF} RM.RegisteredSubNbrMax.SNSSAI}{EC_{ns}}$$

d) NetworkSlice

6.7.2.4.2 Based on the number of active UEs in the network slice

a) $EE_{MIoT,ActiveUEs}$

b) A KPI that shows the energy efficiency of network slices of type MIoT. In this case, the P_{ns} for a network slice of type MIoT is the mean number of active UEs of the network slice.

$$P_{MIot_{,}ActiveUEs} = \sum_{gNBDU} (DRB.MeanActiveUeDl.SNSSAI + DRB.MeanActiveUeUl.SNSSAI)$$

, where *SNSSAI* identifies the S-NSSAI.

This KPI is obtained by the mean number of active UEs of the network slice divided by the energy consumption of the network slice. The unit of this KPI is UE/J.

c)

$$EE_{MIoT_{,}ActiveUEs} = \frac{\sum_{gNBDU}(DRB.MeanActiveUeDl.SNSSAI + DRB.MeanActiveUeUl.SNSSAI)}{EC_{ns}}$$

d) NetworkSlice

6.7.3 5G Energy Consumption (EC)

6.7.3.1 NF Energy Consumption (EC)

6.7.3.1.1 Definition

a) EC_{NF}

b) This KPI describes the Energy Consumption (EC) of a 5G Network Function (NF). This KPI is obtained by summing up the energy consumption of PNF(s) and/or VNF(s) which compose the NF. The unit of this KPI is J.

c)

$$EC_{NF} = \sum_{PNF} EC_{PNF} + \sum_{VNF} EC_{VNF}$$

- How a 5GC NF is composed of VNFs and PNFs is implementation specific. In particular, whether a VNF instance (respectively PNF) is shared or not between more than one NF is implementation specific. Hence, the case where a VNF instance (resp. PNF) is shared between multiple NFs is out of scope of the present document;

- EC_{PNF} represents the Energy Consumption (EC) of a PNF;

- EC_{VNF} represents the Energy Consumption (EC) of a VNF. It is obtained by summing up the Energy Consumption (EC) of all its constituent VNFCs;

- In the present document:

  # EC_{PNF} is measured according to ETSI ES 202 336-12 [10],

  # it is considered that EC_{VNF} cannot be measured hence is estimated. Therefore the resulting EC_{NF} KPI is defined as:

$$EC_{NF} = \sum_{PNF} EC_{PNF,\ measured} + \sum_{VNF} EC_{VNF,\ estimated}$$

6.7.3.1.2 Estimated Virtualized Network Function (VNF) energy consumption

a) $EC_{VNF,estimated}$

b) A KPI that gives an estimation of the energy consumption of a VNF. This KPI is obtained by summing up the estimated energy consumption of its constituent Virtualized Network Function Components (VNFC). The unit of this KPI is J.

c)

$$EC_{VNF,\ estimated} = \sum_{VNFC} EC_{VNFC,\ estimated}$$

d) ManagedFunction

e) In this version of the document, the energy consumption of the VNFC is estimated as per clause 6.7.3.1.3.

6.7.3.1.3 Estimated Virtualized Network Function Component (VNFC) energy consumption

a) $EC_{VNFC,estimated}$

b) A KPI that gives an estimation of the energy consumption of a VNFC. In this version of the document, this KPI is obtained by taking the estimated energy consumption of the virtual compute resource instance on which the VNFC runs. The unit of this KPI is J.

c) $$EC_{VNFC_{,} estimated} = EC_{virtualCompute_{,} estimated}$$

d) ManagedFunction

e) In this version of the document, the energy consumption of the virtual compute resource instance is estimated based on its mean vCPU usage, as per clause 6.7.3.1.4. The method for calculating $EC_{VNFC,estimated}$,estimated is described in TS 28.310 [9] clause 6.3.2.2.1.

6.7.3.1.4 Estimated virtual compute resource instance energy consumption based on mean vCPU usage

a) $EC_{virtualCompute,estimated,VCpuUsageMean}$

b) A KPI that gives an estimation of the energy consumption of a virtual compute resource instance. The energy consumption of a virtual compute resource instance X is estimated as a proportion of the energy consumption of the NFVI node on which the virtual compute resource instance X runs. This proportion is obtained by dividing the vCPU mean usage of the virtual compute resource instance X, by the sum of the vCPU mean usage of all virtual compute resource instances running on the same NFVI Node as X. The unit of this KPI is J.

c)

$$EC_{virtualCompute_{,} estimated_{,} VCpuUsageMean} = \frac{VCpuUsageMean}{\sum_{virtualCompute} VCpuUsageMean} * EC_{NFVINode,measured}$$

d) ManagedFunction

e)

- VCpuUsageMean is the mean vCPU usage of the virtual compute resource instance during the observation period, provided by ETSI NFV MANO (see clause 7.1.2 of ETSI GS NFV-IFA 027 [11]),

- $\sum_{virtualCompute} VCpuUsageMean$ is sum of the vCPU mean usage of all virtual compute resource instances running on the same NFVI Node during the same observation period, all separately provided by NFV MANO (see clause 7.1.2 of ETSI GS NFV-IFA 027 [11]),

- $EC_{NFVINode,measured}$ is the energy consumption of the NFVI node on which the virtual compute resource runs, measured during the same observation period, as per ETSI ES 202 336-12 [10].

6.7.3.2 5GC Energy Consumption (EC)

6.7.3.2.1 Definition

a) EC_{5GC}

b) This KPI describes the Energy Consumption (EC) of the 5G Core Network (CN). It is obtained by summing up the Energy Consumption of all the Network Functions (EC_{NF}) that compose the 5G core network. For the Energy Consumption (EC) of Network Functions, see clause 6.7.3.1. The unit of this KPI is J.

c)

$$EC_{5GC} = \sum_{NF} EC_{NF}$$

d) Subnetwork

6.7.3.3 Network Slice Energy Consumption (EC)

a) EC_{ns}

b) This KPI describes the Energy Consumption (EC) of the network slice. It is obtained by summing up the Energy Consumption of all the Network Functions (EC_{NF}) that compose the network slice. The unit of this KPI is J.

c)

$$EC_{ns} = \sum_{NF} EC_{NF}$$

As a network slice may be composed of a RAN network slice subnet, a Transport Network (TN) network slice subnet and a 5GC network slice subnet, they all participate to the energy consumption of the network slice. However, the definition and way to measure the energy consumption of the TN segment is not in the scope of the present document.

The definition of EC_{ns} based on the following principles:

- For all gNBs in the network slice, clause 5.1.1.19.3 (PNF Energy consumption) of TS 28.552 [6] applies. This measurement is obtained according to the method defined in ETSI ES 202 336-12 [10] – clauses 4.4.3.1, 4.4.3.4, Annex A;
- In case a 5GC NF is composed of Virtualized Network Functions (VNF) and/or Physical Network Functions (PNF), clause 6.7.3.1 of this document defines the NF Energy Consumption (EC);
- In case a NF is dedicated to a network slice, the energy consumption of the NF is entirely attributable to the network slice;
- In case a NF is shared between multiple network slices, the participation of the NF to the energy consumption of the network slice has to be estimated, as it can't be measured:
    - In case of a gNB shared between multiple network slices, the energy consumption attributable to each network slice is estimated as a proportion of the total gNB energy consumption, where the proportion is calculated as the data volume of the network slice relatively to the total data volume carried by the gNB,
    - In case of a AMF shared between multiple network slices, the energy consumption attributable to each network slice is estimated as a proportion of the total estimated AMF energy consumption, where the proportion is calculated as the mean number of registered subscribers of the network slice relatively to the overall mean number of registered subscribers of the AMF during the same time period (see TS 28.552 [6] clause 5.2.1.1 for the definition of the mean number of registered subscribers),
    - In case of a SMF shared between multiple network slices, the energy consumption attributable to each network slice is estimated as a proportion of the total estimated SMF energy consumption, where the proportion is calculated as the mean number of PDU sessions of the network slice relatively to the overall mean number of PDU sessions of the SMF during the same time period (see TS 28.552 [6] clause 5.3.1.1 for the definition of the mean number of PDU sessions),
    - In case of a UPF shared between multiple slices, the energy consumption attributable to each network slice is estimated as a proportion of the total estimated UPF energy consumption, where the proportion is calculated as the data volume of the network slice relatively to the overall data volume of the UPF during the same time period.
        - In case of a UPF with N3 interface(s), the data volume of the UPF is obtained by summing up, for all N3 interface(s), the number of octets of incoming GTP data packets on the N3 interface, from (R)AN to UPF (see TS 28.552 [6] clause 5.4.1.3) and the number of octets of outgoing GTP data packets on the N3 interface, from UPF to (R)AN (see TS 28.552 [6] clause 5.4.1.4)
        - In case of a PSA UPF with no N3 interface(s), the data volume of the UPF is obtained by summing up, for all N9 interface(s), the number of octets of incoming GTP data packets on the N9 interface for PSA UPF (see TS 28.552 [6] clause 5.4.4.2.3) and the number of octets of outgoing GTP data packets on the N9 interface for PSA UPF (see TS 28.552 [6] clause 5.4.4.2.4)
    - The case of other 5GC NFs shared between network slices is not addressed in the present document.

d) NetworkSlice

6.7.3.4 NG-RAN Energy Consumption (EC)

6.7.3.4.1 NG-RAN EC

a) EC_{NG-RAN}

b) This KPI describes the Energy Consumption (EC) of the NG-RAN. It is obtained by summing up the Energy Consumption of all the gNBs that constitute the NG-RAN. The unit of this KPI is J.

c) $EC_{NG-RAN} = \sum_{gNB} EC_{gNB}$

d) Subnetwork

6.7.3.4.2 gNB EC

a) EC_{gNB}

b) This KPI describes the Energy Consumption (EC) of the gNB. It is obtained by summing up the Energy Consumption of all the Network Functions (NF) that constitute the gNB. For the Energy Consumption of Network Functions (EC_{NF}), see clause 6.7.3.1. The unit of this KPI is J.

c) $EC_{gNB} = \sum_{NF} EC_{NF}$

d) ManagedElement

6.7.4 5GC Energy Efficiency (EE)

6.7.4.1 Generic 5GC Energy Efficiency (EE)KPI

$$\text{Generic 5GC EE KPI} = \frac{Useful\ Output\ of\ 5GC\ (UsefulOutput_{5GC})}{Energy\ Consumption\ of\ 5GC\ (EC_{5GC})}$$

, where:

- 'Useful Output of 5GC' ($UsefulOutput_{5GC}$) is the useful output of 5GC. It can be defined differently, depending on which 5GC network functions are considered;

- 'Energy Consumption of 5GC' (EC_{5GC}) is the Energy Consumption of 5GC.

For one unit of EC_{5GC}, the higher $UsefulOutput_{5GC}$ is, the higher the generic 5GC EE KPI is, i.e. the more energy efficient the 5GC is.

6.7.4.2 Energy Efficiency of 5GC based on the useful output of 5GC user plane

a) $EE_{5GC,UO,UP,DV}$

b) A KPI that shows the energy efficiency of 5GC. This KPI is based on the useful output of 5GC user plane. The useful output of the 5GC user plane is obtained by summing up UL and DL data volumes at N3 interface(s).

$$UsefulOutput_{5GC,DV} = \sum_{UPF} (GTP.InDataOctN3UPF + GTP.OutDataOctN3UPF) * 8$$

This KPI is obtained by the sum of UL and DL data volumes at N3 interface(s), divided by the energy consumption of 5GC. The unit of this KPI is bit/J.

c) $$EE_{5GC,UO,UP,DV} = \frac{\sum_{UPF}(GTP.InDataOctN3UPF + GTP.OutDataOctN3UPF) * 8}{EC_{5GC}}$$

d) SubNetwork

e) In case of redundant transmission paths over the N3 interface for high reliability communication (cf. TS 23.501 [7] clause 5.33.2), it is expected that the data volume is counted once. In particular:

- In case of Dual Connectivity based end to end Redundant User Plane Paths (see TS 23.501 [7] clause 5.33.2.1), in which a UE may set up two redundant PDU Sessions over the 5G network, the Data Volume related to only one PDU session is to be considered;

- In case of redundant transmission with two N3 tunnels between the PSA UPF and a single NG-RAN node (cf. TS 23.501 [7] figure 5.33.2.2-1) which are associated with a single PDU Session, the Data Volume related to only one N3 tunnel is to be considered;

- In case of two N3 and N9 tunnels between NG-RAN and PSA UPF for redundant transmission (see TS 23.501 [7] figure 5.33.2.2-2) associated with a single PDU Session, the Data Volume related to only one N3 tunnel is to be considered.

For the measurement of the energy efficiency of the 5G core network, the 3GPP management system in charge of collecting the data volume measurements listed here above shall consider them only once in case of redundant transmission over the N3 interface.

6.8 Reliability KPI

6.8.1 Definition

Reliability is defined (see TS 22.261 [13] clause 3.1) in the context of network layer packet transmissions, as the percentage value of the packets successfully delivered to a given system entity within the time constraint required by the targeted service out of all the packets transmitted.

6.8.1.1 Packet transmission reliability KPI in DL on Uu

a) DLRelPSR_Uu

b) This KPI describes the Reliability based on Packet Success Rate(PSR) Percentage between gNB and UE. It is used to evaluate the Uu interface reliability contribution to the total network downlink reliability. It is the percentage of RLC SDU packets which are successfully received in UE out of the total RLC SDU packets transmitted by gNB. It is a measure of the DL packet delivery success i.e. PSR% over Uu interface. It is a percentage value (%). This KPI can optionally be split into KPIs per QoS level (mapped 5QI or QCI in NR option 3) and per S-NSSAI.

c) Below is the equation for downlink Reliability in RAN based on PSR percentage between gNB and UE.

DLRelPSR_Uu = $\left\lfloor \frac{N(T1,drbid)}{N(T1,drbid)+Dloss(T1,drbid)} \right\rfloor \times 100$, where N(T1,drbid) & Dloss(T1,drbid) are as defined in TS 38.314.

or optionally DLRelPSR_Uu.QoS = $\left\lfloor \frac{N(T1,drbid).QoS}{N(T1,drbid).QoS + Dloss(T1,drbid).QoS} \right\rfloor \times 100$, where QoS identifies the target QoS quality of service class.

or optionally DLRelPSR_Uu.SNSSAI = $\left\lfloor \frac{N(T1,drbid).SNSSAI}{N(T1,drbid).SNSSAI + Dloss(T1,drbid).SNSSAI} \right\rfloor \times 100$,

where SNSSAI identifies the S-NSSAI.

d) NRCellDU

6.8.1.2 Packet transmission reliability KPI in UL on Uu

a) ULRelPSR_Uu

b) This KPI describes the Reliability based on Packet Success Rate Percentage between UE and gNB. It is used to evaluate the Uu interface reliability contribution to the total network uplink reliability. It is the percentage of PDCP SDU packets which are successfully received in gNB out of the total PDCP SDU packets transmitted by UE. It is a measure of the UL packet delivery success i.e. PSR% over Uu interface. It is a percentage value (%). This KPI can optionally be split into KPIs per QoS level (mapped 5QI or QCI in NR option 3) and per S-NSSAI.

c) ULRelPSR_Uu = DRB.PacketSuccessRateUlgNBUu × 100 , where DRB_PacketSuccessRateUlgNBUu is as defined in TS 28.552 [6].

or optionally ULRelPSR_Uu.QoS = DRB.PacketSuccessRateUlgNBUu.QOS × 100, where QoS identifies the target QoS quality of service class.

or optionally ULRelPSR_Uu.SNSSAI = DRB.PacketSuccessRateUlgNBUu.SNSSAI × 100, where SNSSAI identifies the S-NSSAI.

d) NRCellCU

6.8.1.3 Packet transmission reliability KPI in DL on N3

a) DLRelPSR_N3

b) This KPI describes the Reliability based on Packet Success Rate(PSR) Percentage between UPF and gNB. It is used to evaluate the N3 interface reliability contribution to the total network downlink reliability. It is the percentage of GTP data PDUs which are successfully received by gNB out of the total GTP data PDUs transmitted by UPF over N3 interface. It is a measure of the DL packet delivery success i.e. PSR% over N3 interface. It is a percentage value (%). This KPI can optionally be split into KPIs per QoS level (mapped 5QI or QCI in NR option 3) and per S-NSSAI.

c) DLRelPSR_N3 = $\left(\frac{\text{GTP.OutDataPktN3UPF - GTP.InDataPktPacketLossN3gNB}}{\text{GTP.OutDataPktN3UPF}}\right) \times 100$

where GTP.OutDataPktN3UPF, GTP.InDataPktPacketLossN3gNB are as defined in TS 28.552

or optionally,

DLRelPSR_N3.QoS = $\left(\frac{\text{GTP.OutDataPktN3UPF.}\mathit{QoS}\text{ - GTP.InDataPktPacketLossN3gNB.}\mathit{QoS}}{\text{GTP.OutDataPktN3UPF.}\mathit{QoS}}\right) \times 100$

where QoS identifies the target QoS quality of service class.

or optionally,

DLRelPSR_N3.SNSSAI = $\left(\frac{\text{GTP.OutDataPktN3UPF.}\mathit{SNSSAI}\text{ - GTP.InDataPktPacketLossN3gNB.}\mathit{SNSSAI}}{\text{GTP.OutDataPktN3UPF.}\mathit{SNSSAI}}\right) \times 100$

where SNSSAI identifies the S-NSSAI.

d) UPFFunction, GNBCUUPFunction

6.8.1.4 Packet transmission reliability KPI in UL on N3

a) ULRelPSR_N3

b) This KPI describes the Reliability based on Packet Success Rate(PSR) Percentage between gNB and UPF. It is used to evaluate the N3 interface reliability contribution to the total network uplink reliability. It is the percentage of GTP data PDUs which are successfully received by UPF out of the total GTP data PDUs transmitted by gNB over N3 interface. It is a measure of the UL packet delivery success i.e. PSR% over N3 interface. It is a percentage value (%). This KPI can optionally be split into KPIs per QoS level (mapped 5QI or QCI in NR option 3) and per S-NSSAI.

c) ULRelPSR_N3 = $\left(\frac{\text{GTP.InDataPktN3UPF}}{\text{GTP.InDataPktN3UPF} + \text{GTP.InDataPktPacketLossN3UPF}}\right) \times 100$

where GTP.InDataPktN3UPF , GTP.InDataPktPacketLossN3UPF are as defined in TS 28.552

or optionally,

ULRelPSR_N3.QoS = $\left(\frac{\text{GTP.InDataPktN3UPF.}QoS}{\text{GTP.InDataPktN3UPF.}QoS + \text{GTP.InDataPktPacketLossN3UPF.}QoS}\right) \times 100$

where QoS identifies the target QoS quality of service class.

or optionally,

ULRelPSR_N3.SNSSAI = $\left(\frac{\text{GTP.InDataPktN3UPF.}SNSSAI}{\text{GTP.InDataPktN3UPF.}SNSSAI + \text{GTP.InDataPktPacketLossN3UPF.}SNSSAI}\right) \times 100$

where SNSSAI identifies the S-NSSAI.

d) UPFFunction

6.9 Average air-interface efficiency achievable per UE within the observed NRCellDU

a) AvgCqiEfficiency_Cell.

b) The KPI describes the average air-interface efficiency for a NRCellDU according to CQI tables. The KPI takes into account both the channel rank(RI) and the channel quality(CQI), and can comprehensively reflect the overall channel quality of the cell. It is a real value. The unit of KPI refer to that of efficiency defined in TS 38.214. The KPI type is MEAN.

c) Below is the equation for average air-interface efficiency for NRCellDU:

$$AvgCqiEfficiency\_Cell = \frac{\sum_{X,Y,Z} Y * CARR.WBCQIDist.BinX.BinY.BinZ * efficiency_{CQI_X}^{TableZ}}{\sum_{X,Y,Z} CARR.WBCQIDist.BinX.BinY.BinZ}$$

Where $efficiency_{CQI_X}^{TableZ}$ is the efficiency used in the CQI table defined in TS 38.214 [14].

d) NRCellDU

Annex A (informative): Use cases for end to end KPIs

A.1 Use case for end-to-end latency measurements of 5G network-related KPI

The end-to-end latency is an important performance parameter for operating 5G network. In some scenarios (e.g. uRLLC), if end-to-end latency is insufficient, the 5G network customer cannot obtain guaranteed network performance provided by the network operator. So it is necessary to define end-to-end latency of network related measurement to evaluate whether the end-to-end latency that network customer requested has been satisfied. A procedure is invoked by network management system and is used:

- to update the CSMF/NSMF with the end-to-end latency parameter for monitoring;
- to inform the network customer/network operator the end-to-end latency;
- to make CSMF/NSMF aware if the end-to-end latency can meet network customer’s service requirement.

If high end-to-end latency are measured, it is also of benefit to pinpoint where in the chain from application to UE that the latency occurs.

A.2 Use case for number of registered subscribers of single network-slice related KPI

Number (mean and max) of registered subscribers of single network slice can be used to describe the amount of subscribers that are successfully registered, it can reflect the usage of network slice , It is useful to evaluate accessibility performance provided by one single network slice which may trigger the lifecycle management of the network slice, this kind of KPI is valuable especially when network functions (e.g. AMF) are shared between different network slice . This KPI is focusing on both network and user view.

A.3 Use case for upstream/downstream throughput for one-single-network-slice-related KPI

Measuring throughput is useful to evaluate system load of end to end network slice. If the throughput of the specific network slice cannot meet the performance requirement, some actions need to be performed to the network slice e.g. reconfiguration, capacity relocation. So it is necessary to define the IP throughput for one single network slice . This KPI is focusing on network and user view.

A.4 Use case for mean PDU sessions number in network slice

It is necessary to evaluate the mean or maximum PDU session number in the network slice to indicate system load level. For example, if the mean or maximum value of the PDU sessions is high, maybe the system capacity should be increased. This KPI is focusing on network view.

A.5 Use case for virtualised resource utilization of network-slice-related KPI

It is necessary to evaluate the current utilization of virtualised resources (e.g. memory and storage utilization) that a network slice   is occupied. If the utilization is larger or smaller than the threshold, maybe some scale in/out operations will be made by the management system. This KPI is focusing on network and user view.

A.6 Use case for 5GS registration success rate of one single-network-slice-related KPI

It is necessary to evaluate accessibility performance provided by 5GS. 5GS registration for a UE is important when they have registered to the network slice . If users or subscribers cannot register to the network slice , they cannot access any network services in the network slice . This KPI is focusing on network view.

A.7 Use case for RAN UE throughput-related KPI

The UE perceived throughput in NG-RAN is an important performance parameter for operating 5G network. If the UE throughput of the NR cell cannot meet the performance requirement, some actions need to be performed to the network, e.g. reconfiguration or capacity increase. So it is necessary to define UE throughput KPI to evaluate whether the end-users are satisfied. The KPI covers volume large enough to make the throughput measurement relevant, i.e. excluding data volume of the last or only slot.

The UE throughput KPI covers also "NR option 3" scenarios. Then the gNB is "connected" towards the EPC, and not towards 5GC.

It is proposed to allow the KPI separated based on mapped 5QI (or for QCI in case of NR option 3).

When network slicing is supported by the NG-RAN, multiple network slices may be supported. The UL and DL UE throughput for each network slice is then of importance to the operator to pinpoint a specific performance problem.

A.8 Use case for QoS flow retainability-related KPI

QoS flow is the key and limited resource for 5GS to deliver services. The release of the QoS flow needs to be monitored. QoS flow retainability is a key performance indicator of how often an end-user abnormally losing a QoS flow during the time the QoS flow is used. This key performance indicator is of great importance to estimate the end users’ experiences.

A.9 Use case for DRB accessibility-related KPIs

In providing services to end-users, the first step is to get access to the service. First after access to the service has been performed, the service can be used.

The service provided by NG-RAN is the DRB. For the DRB to be successfully setup it is also necessary to setup an RRC connection and an NG signalling connection.

If an end user cannot access a service, it is hard to charge for the service. Also, if it happens often that an end-user cannot access the provided service, the end-user might change wireless subscription provider, i.e. loss of income for the network operator. Hence, to have a good accessibility of the services is important from a business point of view.

The DRB accessibility KPIs require the following 5 measurements:

- DRB connection setup success rate.

- DRB setup success rate.

The success rate for RRC connection setup and for NG signalling connection setup shall exclude setups with establishment cause mo-Signalling, since these phases/procedures occur when there is no request to setup a DRB.

The KPIs are available per mapped 5QI and per S-NSSAI, and assist the network operator with information about the accessibility provided to their 5G network customers.

A.10 Use case for mobility KPIs

When a service is used it is important that it is not interrupted or aborted. One of the fault cases in a radio system for this is handovers/mobility.

If a mobility KPI is not considered OK, then the network operator can investigate which steps that are required to improve the mobility towards their services.

These KPIs can be used for observing the impact on end-users of mobility in NG-RAN and towards other system.

A.11 Use case for DRB retainability related KPI

DRB is the key and limited resource for 5GS to deliver services. Once a QoS flow reaches a gNB it will trigger setup of a new DRB or it will be mapped to an existing DRB. The decision on how to map QoS flows into new or existing DRBs is taken at the CU-CP. CU-CP also defines one set of QoS parameters (one 5QI) for the DRB. If a QoS flow is mapped to an existing DRB, the packets belonging to that QoS flow are not treated with the 5QI of the QoS flow, but they are treated with the mapped 5QI of the DRB.

The release of the DRB needs to be monitored, so that abnormal releases while the UE is considered in an active transfer shall be logged. DRB retainability is a key performance indicator of how often an end-user abnormally loses a DRB during the time the DRB is actively used. This key performance indicator is of great important to estimate the end users' experiences. DRBs with bursty flow are considered active if any data (UL or DL) has been transferred during the last 100 ms. DRBs with continuous flow are seen as active DRBs in the context of this measurement as long as the UE is in RRC connected state. A particular DRB is defined to be of type continuous flow if the mapped 5QI is any of {1, 2, 65, 66}.

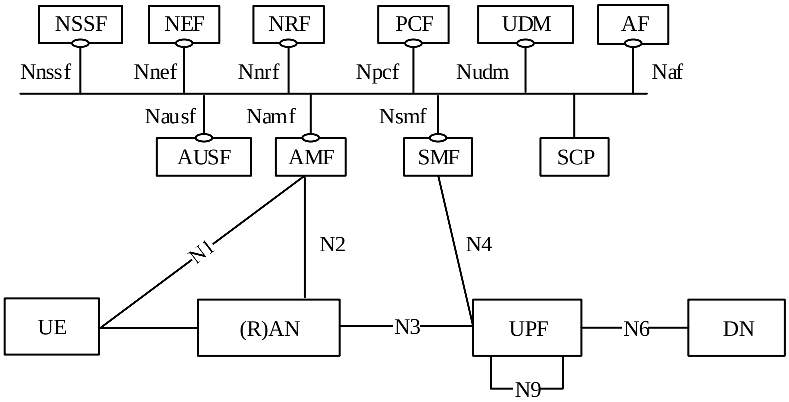
The key performance indicator shall monitor the DRB retainability for each used mapped 5QI value, as well as for the used S-NSSAI(s). DRBs used in 3GPP option 3 shall not be covered by this KPI. For the case when a DRB have multiple QoS flows mapped and active, when a QoS flow is released it will not be counted as a DRB release (DRB still active) in this KPI.

A.12 Use case for PDU session establishment success rate of one network slice (S-NSSAI) related KPI

It is necessary to evaluate accessibility performance provided by 5GS. PDU session Establishment for a UE is important when it has registered to the network slice. If users or subscribers cannot establish PDU sessions in slice instance, they cannot access any network services in the network slice. This KPI is focusing on network view.

A.13 Use case for integrated downlink latency in RAN

Following figure captured in clause 4.2.3, 3GPP TS 23.501[x] illustrates the 5G system architecture. The end to end downlink latency should be measured from Data Network to UE, of which the latency from RAN to UE is an important part for the latency of this section is closely related to NG-RAN.



The integrated downlink latency in RAN is a key performance parameter for evaluating the packet delay in RAN for QoS monitoring. This KPI is also an important part of the end-to-end network latency for SLA assurance.

A.14 Use case for PDU session Establishment success rate of one single-network-slice instance-related KPI

It is necessary to evaluate PDU session establishment time, it can be used to analyse the network service difference between different RAN locations in one area, which can be used for management area division. This KPI is focusing on network view.

A.15 Use case for QoS flow retainability-related KPI

QoS Flow is the key and limited resource for 5GS to deliver services. The release of the QoS flow needs to be monitored. QoS Flow drop ratio is a key performance indicator of how often an end-user is abnormally losing a bearer. This key performance indicator is of great importance to estimate the end users' experiences.

The KPI shall be available per QoS group.

From QoS perspective it is important to focus also on call duration as in some cases wrong quality perceived by the end user is not fully reflected by drop ratio nor retainability KPI. Typical case is when due to poor radio conditions the end user redials (the call was terminated normally) to the same party to secure the quality. But in this case the drop ratio KPI will not show any degradation. Secondly, although the call is dropped the end user may or may not redial depending on dropped call duration compared to the case when the call would be normally released. It is therefore highly recommended to monitor distribution of duration of normally and abnormally released calls.

A.16 Use case for 5G Energy Efficiency (EE) KPI

Assessment of Energy Efficiency in network is very important for operators willing to control their OPEX and, in particular, their network energy OPEX.

5G energy efficiency can be addressed from various perspectives:

- NG-RAN.

Mobile Network data Energy Efficiency ($EE_{MN,DV}$) is the ratio between the performance indicator (DV_{MN}) and the energy consumption (EC_{MN}) when assessed during the same time frame, see ETSI ES 203 228 [8] clause 3.1 and clause 5.3.

$$EE_{MN,DV} = \frac{DV_{MN}}{EC_{MN}}$$

where $EE_{MN,DV}$ is expressed in bit/J.

Assessment of $EE_{MN,DV}$ needs the collection of both Data Volumes (DV) and Energy Consumption (EC) of 5G Network Functions (NF). How this EE KPI can be applied to NG-RAN is specified in clause 4.1 of TS 28.310 [9].

Before the network operator takes any action to save network energy OPEX, the network operator needs to know the energy efficiency of its 5G network.

This KPI needs to be used for observing the impact of NG-RAN on data energy efficiency of 5G access networks.

- Network slices.

In a Network Slice as a Service (NSaaS) model, a Network Slice Customer (NSC) may ask to its Network Slice Provider (NSP) a network slice with certain characteristics, among which the expected EE of the network slice. It is therefore required that a standardized definition of EE KPIs exists, per type of network slice, and that such EE KPIs can be measured and delivered by Network Slice Providers.

A.17 Use case for PFCP session established success rate of one network and one network slice instance-related KPI

It is necessary to evaluate the PFCP session established success rate of one network and one network slice instance-related KPI. It can be used to analyse the quality for the N4 interface connection between different vendor's SMF and UPF related to one network or one network silce. The KPI is very useful for the industry customer's Park area and N4 interface decouping deployment for operrators. And this KPI is mainly focusing on network view.

A.18 Use case for end-to-end reliability measurements of 5G network-related KPI

The end-to-end reliability is an important performance parameter for operating 5G network. In some scenarios (e.g. uRLLC), if end-to-end reliability is insufficient, the 5G network customer cannot obtain guaranteed network performance provided by the network operator. So it is necessary to assess end-to-end reliability of network utilizing the packet delivery success rate measurements defined in clauses 6.8.2,6.8.3,6.8.4 and 6.8.5.The same can be used to determine the end to end reliability of a slice.

A.19 Use case for average UE achievable air-interface efficiency

The average UE achievable air-interface efficiency could provide operators with the channel status considering both channel rank and quality simultaneously. It can help operators to evaluate the overall channel state and the efficiency of the cell when Massive MIMO is enabled. There are several CQI tables defined in TS 38.214[14], the KPI can be calculated by using all the CQI tables or some of CQI tables with the same BLER value.

Annex B (informative): Change history

Change history							
Date	**Meeting**	**TDoc**	**CR**	**Rev**	**Cat**	**Subject/Comment**	**New version**
2018-09	SA#81					Upgrade to change control version	15.0.0
2018-09	SA#81					EditHelp fix	15.0.1
2018-12	SA#82	SP-181041	0001	-	F	Align title with TS database	15.1.0
2019-03	SA#83	SP-190122	0005	2	F	Update KPI subscribers of single network slice instance through UDM	15.2.0
2019-03	SA#83	SP-190122	0011	2	F	Update definition of mean number of PDU sessions KPI	15.2.0
2019-03	SA#83	SP-190111	0007	1	B	Add KPI of QoS flow Retainability	16.0.0
2019-03	SA#83	SP-190111	0009	1	B	Add DRB Accessibility KPI and Use Case	16.0.0
2019-06	SA#84	SP-190371	0013	-	B	Add KPI for NG-RAN Handover Success Rate	16.1.0
2019-06	SA#84	SP-190375	0015	1	A	Correction of Throughput KPI	16.1.0
2019-09	SA#85	SP-190747	0016	2	B	Add KPI for DRB Retainability	16.2.0
2019-09	SA#85	SP-190747	0017	3	B	Add a new KPI definition of PDU session Establishment Success Rate of one network slice (S-NSSAI)	16.2.0
2019-09	SA#85	SP-190751	0020	-	A	Correction on kbits abbreviation	16.2.0
2019-09	SA#85	SP-190747	0021	1	F	Correction of Flow Retainability KPI	16.2.0
2019-09	SA#85	SP-190747	0022	1	F	Correction of DRB Accessibility KPI	16.2.0
2019-09	SA#85	SP-190748	0024	2	A	Correct the title of KPI	16.2.0
2019-09	SA#85	SP-190747	0025	1	B	Add definition of integrated downlink latency in RAN	16.2.0
2019-09	SA#85	SP-190747	0028	2	B	Add a new KPI definition of Inter-gNB handover Execution time of one single network slice	16.2.0
2019-09	SA#85	SP-190747	0029	2	B	Add a new KPI definition of PDU session Establishment Time of one single network slice	16.2.0
2019-09	SA#85	SP-190747	0030	2	B	Add new specification requirement related to extended 5QI 1 QoS Flow Retainability monitoring	16.2.0
2019-12	SA#86	SP-191165	0032	1	B	Add 5G Energy Efficiency KPI	16.3.0
2019-12	SA#86	SP-191149	0033	1	B	Add a new KPI definition of Mean number of successful periodic registration updates of Single Network Slice	16.3.0
2019-12	SA#86	SP-191149	0034	1	B	Add a new description of KPI that related to successful rate of mobility registration updates of Single Network Slice	16.3.0
2019-12	SA#86	SP-191150	0036	1	F	Update the template of KPI definition for TS 28.554	16.3.0
2020-03	SA#87E	SP-200163	0038	1	F	Update KPI definitions to align with the new template	16.4.0
2020-03	SA#87E	SP-200162	0039	-	F	Correction of equation color	16.4.0
2020-07	SA#88-e	SP-200502	0044	1	F	Correction of Downlink latency in gNB-DU KPI	16.5.0
2020-07	SA#88-e	SP-200502	0045	-	F	Removal of the KPI named KPI categories	16.5.0
2020-07	SA#88-e	SP-200502	0046	-	F	Update of KPI template	16.5.0
2020-07	SA#88-e	SP-200503	0049	1	B	Add KPI on e2e UL delay for network slice	16.5.0
2020-07	SA#88-e	SP-200503	0050	1	B	Add KPI on e2e DL delay for network slice	16.5.0
2020-07	SA#88-e	SP-200503	0051	1	B	Add KPIs for UL packet delay in NG-RAN	16.5.0
2020-07	SA#88-e	SP-200503	0052	1	B	Correction of Integrated downlink delay in RAN KPI	16.5.0

2020-07	SA#88-e	SP-200485	0053	1	F	Cleanup based on refined slice definitions	16.5.0
2020-09	SA#89e	SP-200751	0054	1	F	Fixing KPIs	16.6.0
2020-09	SA#89e	SP-200738	0056	-	F	Correction of RAN UE throughput KPI	16.6.0
2020-09	SA#89e	SP-200747	0057	1	B	Additional KPI Definition for Max Subscriber and PDU Session	17.0.0
2020-12	SA#90e	SP-201059	0061	1	A	Correct UDM e2e KPI	17.1.0
2020-12	SA#90e	SP-201061	0064	-	A	Editorial Correction of TS 28.554	17.1.0
2020-12	SA#90e	SP-201054	0068	-	A	Correction and alignment of Retainability KPIs definitions	17.1.0
2020-12	SA#90e	SP-201054	0069	-	A	Add missing KPI for inter system Handover success rate	17.1.0
2020-12	SA#90e	SP-201062	0071	1	B	Add EE KPI definitions for network slices	17.1.0
2020-12	SA#90e					Correction of style of clause from CR0071	17.1.1
2021-03	SA#91e	SP-210157	0074	-	B	CHO measurements KPI	17.2.0
2021-03	SA#91e	SP-210150	0077	-	A	Update retainability KPIs to consider abnormal releases in RRC connected state	17.2.0
2021-06	SA#92e	SP-210412	0075	2	B	Update the Accessibility KPI to cover DRB access via RRC Resume	17.3.0
2021-06	SA#92e	SP-210412	0078	1	B	Definition of the Total DRB Accessibility KPI .	17.3.0
2021-06	SA#92e	SP-210578	0079	1	B	Update on energy efficiency of URLLC network slice	17.3.0
2021-06	SA#92e	SP-210578	0080	1	B	Add Energy Consumption KPI pour 5G NF and 5G CN	17.3.0
2021-06	SA#92e	SP-210578	0081	1	B	Add EE KPI for eMBB network slice based on RAN measurements	17.3.0
2021-09	SA#93e	SP-210869	0082	-	B	Add estimated VNF, VNFC and virtual compute resource instance Energy Consumption KPI	17.4.0
2021-09	SA#93e	SP-210869	0083	1	C	Update the EE KPI for the URLLC network slice	17.4.0
2021-09	SA#93e	SP-210872	0084	-	B	Add Mean&Maximum CM-Connected subscribers of network slice through AMF	17.4.0
2021-09	SA#93e	SP-210872	0085	-	B	Add PFCP session established success rate of one network and one network slice	17.4.0
2021-12	SA#94e	SP-211459	0086	1	B	Add definition of ECns	17.5.0
2021-12	SA#94e	SP-211459	0087	1	B	Add Energy Consumption KPI for NG-RAN	17.5.0
2021-12	SA#94e	SP-211459	0088	1	B	Add definition of 5GC energy efficiency (EE) KPI	17.5.0
2022-03	SA#95e	SP-220172	0091	-	A	Editorial clean up of mobilty KPIs HO success rate	17.6.0
2022-03	SA#95e	SP-220172	0092	1	B	Add KPI for HO success rate for all handover types	17.6.0
2022-03	SA#95e	SP-220180	0093	-	B	Define Reliability KPI in 5G Network	17.6.0
2022-06	SA#96	SP-220515	0095	-	A	Update formula of PDU session establishment success rate	17.7.0
2022-09	SA#97e	SP-220853	0097	-	A	Correct wrong measurement names in KPI definition	17.8.0
2022-09	SA#97e	SP-220850	0098	1	F	Correct 5G energy consumption definitions	17.8.0
2022-09	SA#97e	SP-220850	0099	1	F	Updating Packet transmission reliability KPI in DL on N3	17.8.0
2022-09	SA#97e	SP-221183	0105	2	B	Add KPI on average air-interface efficiency achievable per UE within the observed NRCellDU	18.0.0