ETSI TS 128 552 V15.7.0 (2024-07)

5G;
Management and orchestration;
5G performance measurements
(3GPP TS 28.552 version 15.7.0 Release 15)

Reference
RTS/TSGS-0528552vf70

Keywords
5G

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - APE 7112B
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° w061004871

Intellectual Property Rights

Essential patents

IPRs essential or potentially essential to normative deliverables may have been declared to ETSI. The declarations pertaining to these essential IPRs, if any, are publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available from the ETSI Secretariat. Latest updates are available on the ETSI Web server (https://ipr.etsi.org/).

Pursuant to the ETSI Directives including the ETSI IPR Policy, no investigation regarding the essentiality of IPRs, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Trademarks

The present document may include trademarks and/or tradenames which are asserted and/or registered by their owners. ETSI claims no ownership of these except for any which are indicated as being the property of ETSI, and conveys no right to use or reproduce any trademark and/or tradename. Mention of those trademarks in the present document does not constitute an endorsement by ETSI of products, services or organizations associated with those trademarks.

DECT™, **PLUGTESTS™**, **UMTS™** and the ETSI logo are trademarks of ETSI registered for the benefit of its Members. **3GPP™** and **LTE™** are trademarks of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners. **oneM2M™** logo is a trademark of ETSI registered for the benefit of its Members and of the oneM2M Partners. **GSM**® and the GSM logo are trademarks registered and owned by the GSM Association.

Legal Notice

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

The present document may refer to technical specifications or reports using their 3GPP identities. These shall be interpreted as being references to the corresponding ETSI deliverables.

The cross reference between 3GPP and ETSI identities can be found under https://webapp.etsi.org/key/queryform.asp.

Modal verbs terminology

In the present document "**shall**", "**shall not**", "**should**", "**should not**", "**may**", "**need not**", "**will**", "**will not**", "**can**" and "**cannot**" are to be interpreted as described in clause 3.2 of the ETSI Drafting Rules (Verbal forms for the expression of provisions).

"**must**" and "**must not**" are **NOT** allowed in ETSI deliverables except when used in direct citation.

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

This document specifies the performance measurements for 5G networks including network slicing. Performance measurements for NG-RAN (clause 5.1) as well as for 5GC (clause 5.2 to 5.6) are defined in this document. Related KPIs are defined to those measurements in TS 28.554 [8].

The performance measurements for NG-RAN applies also to NR option 3 in many cases, but not to the RRC connection related measurements which are handled by E-UTRAN for NR option 3 (those are measured according to TS 32.425 [9] and related KPIs in TS 32.451 [10]).

The performance measurements are defined based on the measurement template as described in TS 32.404 [3].

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] 3GPP TS 32.401: "Telecommunication management; Performance Management (PM); Concept and requirements".

[3] 3GPP TS 32.404: "Performance Management (PM); Performance measurements - Definitions and template".

[4] 3GPP TS 23.501: "System Architecture for the 5G System".

[5] IETF RFC 5136: "Defining Network Capacity".

[6] 3GPP TS 38.473: "NG-RAN; F1 Application Protocol (F1AP)".

[7] 3GPP TS 23.502: "Procedures for the 5G System".

[8] 3GPP TS 28.554: “Management and orchestration; 5G end to end Key Performance Indicators (KPI)”.

[9] 3GPP TS 32.425: “Performance Management (PM); Performance measurements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN)”

[10] 3GPP TS 32.451: "Key Performance Indicators (KPI) for Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Requirements”

3 Definitions, abbreviations and measurement family

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].

IP Latency: the time it takes to transfer a first/initial packet in a data burst from one point to another.

Mapped 5QI: In case when a single 5QI is assigned to the DRB, the mapped 5QI refers to the 5QI that is used for a DRB within the gNB.

NOTE 1: In this case the mapped 5QI is used for separating certain measurements per QoS class.

NOTE 2: Individual QoS flows into a common 5QI is specified in TS 38.473 [6].

Packet Delay: the time it takes to transfer any packet from one point to another.

Packet Drop Rate: share of packets that were not sent to the target due to congestion or traffic management and should be seen as a part of the packet loss rate.

Packet Loss Rate: share of packets that could not be received by the target including packets droped, packets lost in transmission and packets received in wrong format.

3.2 Abbreviations

For the purposes of the present document, the abbreviations given in 3GPP TR 21.905 [1], 3GPP TS 23.501 [4] 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] and 3GPP TS 23.501 [4].

kbit	kilobit (1000 bits)
NG-RAN	Next Generation Radio Access Network
NSI	Network Slice Instance

3.4 Measurement family

The measurement names defined in the present document are all beginning with a prefix containing the measurement family name. This family name identifies all measurements which relate to a given functionality and it may be used for measurement administration.

The list of families currently used in the present document is as follows:

- DRB (measurements related to Data Radio Bearer)
- RRC (measurements related to Radio Resource Control)
- UECNTX (measurements related to UE Context)
- RRU (measurements related to Radio Resource Utilization)
- RM (measurements related to Registration Management)
- SM (measurements related to Session Management)
- GTP (measurements related to GTP Management)
- IP (measurements related to IP Management)
- PA (measurements related to Policy Association)

4 Concepts and overview

5 Performance measurements for 5G Network Functions

5.1 Performance measurements for gNB

5.1.1 Performance measurements valid for all gNB deployment scenarios

5.1.1.1 Packet Delay

5.1.1.1.1 Average delay DL air-interface

a) This measurement provides the average (arithmetic mean) time it takes to get a reponse back on a HARQ transmission in the downlink direction. The measurement is optionally split into subcounters per QoS level (mapped 5QI or QCI in NR option 3).

b) DER (n=1)

c) This measurement is obtained as: sum of (time when the last part of an RLC SDU packet was received by the UE according to received HARQ feedback information for UM mode or time when the last part of an RLC SDU packet was received by the UE according to received RLC ACK for AM mode, minus time when corresponding RLC SDUs arriving at MAC lower SAP) divided by total number of RLC SDUs transmitted to UE successfully. Separate counters are optionally maintained for each mapped 5QI (or QCI for option 3).

d) Each measurement is an integer representing the mean delay in microseconds. The number of measurements is equal to one. If the optional QoS level measurement is perfomed, the number of measurements is equal to the number of mapped 5QIs.

e) The measurement name has the form DRB.AirIfDelayDl or optionally DRB. AirIfDelayDl.*QOS*, where *QOS* identifies the target quality of service class.

f) NRCellDU

g) Valid for packet switched traffic

h) 5GS

i) One usage of this measurement is for performance assurance within integrity area (user plane connection quality).

5.1.1.2 Radio resource utilization

5.1.1.2.1 DL Total PRB Usage

a) a) This measurement provides the total usage (in percentage) of physical resource blocks (PRBs) on the downlink for any purpose.

b) b) SI

c) c) This measurement is obtained as: $M(T) = \left\lfloor \frac{M1(T)}{P(T)} * 100 \right\rfloor$, where $M(T)$ is the DL total PRB usage, which is percentage of PRBs used, averaged during time period T with value range: 0-100%; $M1(T)$ is a

count of full physical resource blocks and all PRBs used for DL traffic transmission shall be included; $P(T)$ is total number of PRBs available for DL traffic transmission during time period T ; and T is the time period during which the measurement is performed.

d) d) A single integer value from 0 to 100.

e) e) RRU.PrbTotDl, *which indicates the DL PRB Usage for all traffic*

f) f) NRCellDU

g) g) Valid for packet switched traffic

h) h) 5GS

i) i) One usage of this measurement is for monitoring the load of the radio physical layer.

5.1.1.2.2 UL Total PRB Usage

a) a) This measurement provides the total usage (in percentage) of physical resource blocks (PRBs) on the uplink for any purpose.

b) b) SI

c) c) This measurement is obtained as: $M(T) = \left\lfloor \frac{M1(T)}{P(T)} * 100 \right\rfloor$, where $M(T)$ is the UL total PRB usage, which is percentage of PRBs used, averaged during time period T with value range: 0-100%; $M1(T)$ is a count of full physical resource blocks and all PRBs used for UL traffic transmission shall be included; $P(T)$ is total number of PRBs available for UL traffic transmission during time period T ; and T is the time period during which the measurement is performed

d) d) A single integer value from 0 to 100.

e) e) RRU.PrbTotUl, *which indicates the UL PRB Usage for all traffic*

f) f) NRCellDU

g) g) Valid for packet switched traffic

h) h) 5GS

i) i) One usage of this measurement is for monitoring the load of the radio physical layer.

5.1.1.2.3 Distribution of DL Total PRB Usage

a) a) This measurement provides the distribution of samples with total usage (in percentage) of physical resource blocks (PRBs) on the downlink in different ranges. This measurement is a useful measure of whether a cell is under high loads or not in the scenario which a cell in the downlink may experience high load in certain short times (e.g. in a second) and recover to normal very quickly.

b) b) CC

c) c) Each measurement sample is obtained as: $M[n] = \left\lfloor \frac{M1[n]}{P[n]} * 100 \right\rfloor$, where $M[n]$ is total PRB usage at sample n for DL, which is a percentage of PRBs used, averaged during time period t_n with value range: 0-100%; $M1[n]$ is a count of full physical resource blocks and all PRBs used for DL traffic transmission shall be included; $P(n)$ is the total number of PRBs available for DL traffic transmission during time period t_n and n is the sample with time period t_n during which the measurement is performed.

d) Distribution of total PRB usage is calculated in the time-frequency domain only. The reference point is the Service Access Point between MAC and L1. The distribution of PRB usage provides the histogram result of the samples collected during time period T.

e) Depending on the value of the sample, the proper bin of the counter is increased. The number of samples during one measurement period is defined by the vendor.

f) d) A set of integers. Each representing the (integer) number of samples with a DL total PRB percentage usage in the range represented by that bin.

g) e) RRU.PrbTotDlDist.BinX, which indicates the distribution of DL PRB Usage for all traffic.

h) f) NRCellDU

i) g) Valid for packet switched traffic

j) h) 5GS

k) i) One usage of this measurement is for monitoring the load of the radio physical layer.

5.1.1.2.4 Distribution of UL Total PRB Usage

a) a) This measurement provides the distribution of samples with total usage (in percentage) of physical resource blocks (PRBs) on the uplink in different usage ranges. This measurement is a useful measure of whether a cell is under high loads or not in the scenario which a cell in the uplink may experience high load in certain short times (e.g. in a second) and recover to normal very quickly.

b) b) CC

c) c) Each measurement sample is obtained as: $M[n] = \left\lfloor \frac{M1[n]}{P[n]} * 100 \right\rfloor$, where $M[n]$ is total PRB usage at sample n for UL, which is a percentage of PRBs used, averaged during time period t_n with value range: 0-100%; $M1[n]$ is a count of full physical resource blocks and all PRBs used for UL traffic transmission shall be included; $P(n)$ is the total number of PRBs available for UL traffic transmission during time period t_n and n is the sample with time period t_n during which the measurement is performed.

d) Distribution of total PRB usage is calculated in the time-frequency domain only. The reference point is the Service Access Point between MAC and L1. The distribution of PRB usage provides the histogram result of the samples collected during time period T.

e) Depending on the value of the sample, the proper bin of the counter is increased. The number of samples during one measurement period is defined by the vendor.

f) d) A set of integers, each representing the (integer) number of samples with a UL PRB percentage usage in the range represented by that bin.

g) e) RRU.PrbTotUlDist.BinX, which indicates the distribution of UL PRB Usage for all traffic.

h) f) NRCellDU

i) g) Valid for packet switched traffic

j) h) 5GS

k) i) One usage of this measurement is for monitoring the load of the radio physical layer.

5.1.1.3 UE throughput

5.1.1.3.1 Average DL UE throughput in gNB

a) This measurement provides the average UE throughput in downlink. This measurement is intended for data bursts that are large enough to require transmissions to be split across multiple slots. The UE data volume refers to the total volume scheduled for each UE regardless if using only primary- or also supplemental aggregated carriers. The measurement is optionally split into subcounters per QoS level (mapped 5QI or QCI in NR option 3).

b) DER(N=1)

c) This measurement is obtained according to the following formula based on the "ThpVolDl" and "ThpTimeDl" defined below. It is optionally split into subcounters for each QoS level.

If $\sum_{UEs}\sum ThpTimeDl > 0$, $\frac{\sum_{UEs}\sum ThpVolDl}{\sum_{UEs}\sum ThpTimeDl}$×1000 [kbit/s]

If $\sum_{UEs}\sum ThpTimeDl = 0$, 0 [kbit/s]

For small data bursts, where all buffered data is included in one initial HARQ transmission, $ThpTimeDl = 0$, otherwise $ThpTimeDl = T1 - T2\ [ms]$

ThpTimeDl	The time to transmit a data burst excluding the data transmitted in the slot when the buffer is emptied. A sample of "ThpTimeDl" for each time the DL buffer for one DataRadioBearer (DRB) is emptied.
$T1$	The point in time after T2 when data up until the second last piece of data in the transmitted data burst which emptied the RLC SDU available for transmission for the particular DRB was successfully transmitted, as acknowledged by the UE.
$T2$	The point in time when the first transmission begins after a RLC SDU becomes available for transmission, where previously no RLC SDUs were available for transmission for the particular DRB.
ThpVolDl	The RLC level volume of a data burst, excluding the data transmitted in the slot when the buffer is emptied. A sample for ThpVolDl is the data volume, counted on RLC SDU level, in kbits successfully transmitted (acknowledged by UE) in DL for one DRB during a sample of ThpTimeDl. (It shall exclude the volume of the last piece of data emptying the buffer).

d) Each measurement is a real value representing the throughput in kbits per second. The number of measurements is equal to one. If the optional QoS level measurement is perfomed, the number of measurements is equal to the number of mapped 5QIs.

e) The measurement name has the form DRB.UEThpDl, or optionally DRB.UEThpDl.*QOS*, where *QOS* identifies the target quality of service class.

f) NRCellDU

g) Valid for packet switched traffic

h) 5GS

i) One usage of this measurement is for performance assurance within integrity area (user plane connection quality).

5.1.1.3.2 Distribution of DL UE throughput in gNB

a) This measurement provides the distribution of the UE throughput in downlink. This measurement is intended for data bursts that are large enough to require transmissions to be split across multiple slots. The UE data volume refers to the total volume scheduled for each UE regardless if using only primary- or also supplemental aggregated carriers.

b) CC

c) Considering there are n samples during measurement time T and each sample has the same time period tn, the measurement of one sample is obtained by the following formula for a measurement period tn:

If $\sum_{UEs}\sum ThpTimeDl > 0$, $\frac{\sum_{UEs}\sum ThpVolDl}{\sum_{UEs}\sum ThpTimeDl}\times 1000$ [kbit/s]

If $\sum_{UEs}\sum ThpTimeDl = 0$, 0 [kbit/s]

For small data bursts, where all buffered data is included in one initial HARQ transmission, $ThpTimeDl = 0$, otherwise $ThpTimeDl = T1 - T2\ [ms]$

ThpTimeDl	The time to transmit a data burst excluding the data transmitted in the slot when the buffer is emptied. A sample of "ThpTimeDl" for each time the DL buffer for one DataRadioBearer (DRB) is emptied.
$T1$	The point in time after T2 when data up until the second last piece of data in the transmitted data burst which emptied the RLC SDU available for transmission for the particular DRB was successfully transmitted, as acknowledged by the UE.
$T2$	The point in time when the first transmission begins after a RLC SDU becomes available for transmission, where previously no RLC SDUs were available for transmission for the particular DRB.
$ThpVolDl$	The RLC level volume of a data burst, excluding the data transmitted in the slot when the buffer is emptied. A sample for ThpVolDl is the data volume, counted on RLC SDU level, in kbits successfully transmitted (acknowledged by UE) in DL for one DRB during a sample of ThpTimeDl. (It shall exclude the volume of the last piece of data emptying the buffer).

Alternatively, for small data bursts, that are successfully transmitted in any given slot (i.e. the requirement that data bursts need to span across several slots excluding transmission of the last piece of the data in a data burst does not apply). where all buffered data is included in one initial HARQ transmission, fraction of the slot time ($ThpTimeDL$) may be counted and obtained by the formula:

$$ThpTimeDl = slot \times \frac{(TBVol - PaddingVol)}{TBVol}\ [ms]$$

slot	Duration of the slot
TBVol	Volume of the TB related to one slot burst
PaddingVol	Volume of padding bits added into Transport Block related to one slot burst.

For each measurement sample, the bin corresponding to the DL throughput experienced by the UE is incremented by one.

d) A set of integers, each representing the (integer) number of samples with a DL UE throughput in the range represented by that bin.

e) The measurement name has the form
DRB.UEThpDlDist.BinX where BinX represents the bin.

NOTE: Number of bins and the range for each bin is left to implementation

f) NRCellDU

g) Valid for packet switched traffic

h) 5GS

i) One usage of this measurement is for performance assurance within integrity area (user plane connection quality).

5.1.1.3.3 Average UL UE throughput in gNB

a) This measurement provides the average UE throughput in uplink. This measurement is intended for data bursts that are large enough to require transmissions to be split across multiple slots. The UE data volume refers to the total volume scheduled for each UE regardless if using only primary- or also supplemental aggregated carriers. The measurement is optionally split into subcounters per QoS level (mapped 5QI or QCI in NR option 3).

b) DER(N=1)

c) This measurement is obtained according to the following formula based on the "ThpVolUl" and "ThpTimeUl" defined below. It is optionally split into subcounters for each QoS level.

If $\sum_{UEs}\sum ThpTimeUl > 0$, $\frac{\sum_{UEs}\sum ThpVolUl}{\sum_{UEs}\sum ThpTimeUl}$×1000 [kbit/s]

If $\sum_{UEs}\sum ThpTimeUl = 0$, 0 [kbit/s]

For small data bursts, where all buffered data is included in one initial HARQ transmission $ThpTimeUl = 0$
otherwise:

$$ThpTimeUl = T1 - T2\ [ms]$$

ThpTimeUl	The time to transmit a data burst excluding the data transmitted in the slot when the buffer is emptied. A sample of "ThpTimeUl" for each time the UL buffer for one DataRadioBearer (DRB) is emptied.
$T1$	The point in time when the data up until the second last piece of data in data burst has been successfully received for a particular DRB
$T2$	The point in time when transmission is started for the first data in data burst for a particular DRB.
$ThpVolUl$	The RLC level volume of a data burst, excluding the data transmitted in the slot when the buffer is emptied. A sample for ThpVolUl is the data volume counted on RLC SDU level in kbits received in UL for one DRB during a sample of ThpTimeUl, (It shall exclude the volume of the last piece of data emptying the buffer).

a)

d) Each measurement is a real value representing the throughput in kbits per second. The number of measurements is equal to one. If the optional QoS level measurement is perfomed, the number of measurements is equal to the number of mapped 5QIs.

e) The measurement name has the form
DRB.UEThpUl, or optionally DRB.UEThpUl.*QOS*, where *QOS* identifies the target quality of service class.

f) NRCellDU

g) Valid for packet switched traffic

h) 5GS

i) One usage of this measurement is for performance assurance within integrity area (user plane connection quality).

5.1.1.3.4 Distribution of UL UE throughput in gNB

a) This measurement provides the distribution of the UE throughput in uplink. This measurement is intended for data bursts that are large enough to require transmissions to be split across multiple slots. The UE data volume refers to the total volume scheduled for each UE regardless if using only primary- or also supplemental aggregated carriers.

b) CC

c) Considering there are n samples during measurement time T and each sample has the same time period tn, the measurement of one sample is obtained by the following formula for a measurement period tn:

If $\sum_{UEs}\sum ThpTimeUl > 0$, $\frac{\sum_{UEs}\sum ThpVolUl}{\sum_{UEs}\sum ThpTimeUl}$×1000 [kbit/s]

If $\sum_{UEs}\sum ThpTimeUl = 0$, 0 [kbit/s]

For small data bursts, where all buffered data is included in one initial HARQ transmission $ThpTimeUl = 0$ otherwise:

$$ThpTimeUl = T1 - T2\ [ms]$$

ThpTimeUl	The time to transmit a data burst excluding the data transmitted in the slot when the buffer is emptied. A sample of "ThpTimeUl" for each time the UL buffer for one DataRadioBearer (DRB) is emptied.
T1	The point in time when the data up until the second last piece of data in data burst has been successfully received for a particular DRB
T2	The point in time when transmission is started for the first data in data burst for a particular DRB.
ThpVolUL	The RLC level volume of a data burst, excluding the data transmitted in the slot when the buffer is emptied. A sample for ThpVolUl is the data volume counted on RLC SDU level in kbits received in UL for one DRB during a sample of ThpTimeUl, (It shall exclude the volume of the last piece of data emptying the buffer).

Alternatively, for small data bursts, that are successfully transmitted in any given slot (i.e. the requirement that data bursts need to span across several slots excluding transmission of the last piece of the data in a data burst does not apply). where all buffered data is included in one initial HARQ transmission, fraction of the slot time ($ThpTimeUL$) may be counted and obtained by the formula:

$$ThpTimeUl = slot \times \frac{(TBVol - PaddingVol)}{TBVol}\ [ms]$$

slot	Duration of the slot
TBVol	Volume of the TB related to one slot burst
PaddingVol	Volume of padding bits added into Transport Block related to one slot burst.

For each measurement sample, the bin corresponding to the UL throughput experienced by the UE is incremented by one.

d) A set of integers, each representing the (integer) number of samples with a UL UE throughput in the range represented by that bin.

e) The measurement name has the form DRB.UEThpUlDist.BinX where BinX represents the bin.

NOTE: Number of bins and the range for each bin is left to implementation

f) NRCellDU

g) Valid for packet switched traffic

h) 5GS

i) One usage of this measurement is for performance assurance within integrity area (user plane connection quality).

5.1.1.3.5 Percentage of unrestricted DL UE data volume in gNB

a) This measurement provides the percentage of DL data volume for UEs in the cell that is classified as unrestricted, i.e., when the volume is so low that all data can be transferred in one slot and no UE throughput sample could be calculated. The UE data volume refers to the total volume scheduled for each UE regardless if using only primary- or also supplemental aggregated carriers. The measurement is optionally split into subcounters per QoS level (mapped 5QI or QCI in NR option 3).

b) SI

c) For periods when no data is transferred at all *Percentage Unrestricted Volume DL = 0*, otherwise:

$$Percentage\ Unrestricted\ Volume\ DL\ = \frac{\sum_{UE} ThpUnresVolDl}{\sum_{UE}(ThpUnresVolDl + ThpVolDl)} * 100$$

ThpUnresVolDl	The volume of a data burst that is transmitted in the slot when the buffer is emptied (which could be the only slot needed to transmit the data burst) and not included in the UE throughput measurement. A sample for ThpUnresVolDl is the data volume counted on RLC SDU level in kbit sent in DL for one DRB.
ThpVolDl	The volume of a data burst, excluding the data transmitted in the slot when the buffer is emptied. A sample for ThpVolDl is the data volume

	counted on RLC SDU level in kbit sent in DL for one DRB.

d) Each measurement is a single integer value from 0 to 100. The number of measurements is equal to one. If the optional QoS level measurement is perfomed, the number of measurements is equal to the number of mapped 5QIs.

e) The measurement name has the form DRB.UEUnresVolDl or optionally DRB.UEUnresVolDl.*QOS*, where *QOS* identifies the target quality of service class.

f) NRCellDU

g) Valid for packet switched traffic

h) 5GS

i) One usage of this measurement is for performance assurance within integrity area (user plane connection quality).

5.1.1.3.6 Percentage of unrestricted UL UE data volume in gNB

a) This measurement provides the percentage of UL data volume for UEs in the cell that is classified as unrestricted, i.e., when the volume is so low that all data can be transferred in one slot and no UE throughput sample could be calculated. The UE data volume refers to the total volume scheduled for each UE regardless if using only primary- or also supplemental aggregated carriers. The measurement is optionally split into subcounters per QoS level (mapped 5QI or QCI in NR option 3).

b) SI

c) For periods when no data is transferred at all *Percentage Unrestricted Volume UL = 0*, otherwise:

$$Percentage\ Unrestricted\ Volume\ UL\ = \frac{\sum_{UE} ThpUnresVolUl}{\sum_{UE}(ThpUnresVolUl + ThpVolUl)} * 100$$

ThpUnresVolUl	The volume of a data burst that is transmitted in the slot when the buffer is emptied (which could be the only slot needed to transmit the data burst) and not included in the UE throughput measurement. A sample for ThpUnresVolUl is the data volume counted on RLC SDU level in kbit received in UL for one DRB.
ThpVolUl	The volume of a data burst, excluding the data transmitted in the slot when the buffer is emptied. A sample for ThpVolUl is the data volume counted on RLC SDU level in kbit received in UL for one DRB.

d) Each measurement is a single integer value from 0 to 100. The number of measurements is equal to one. If the optional QoS level measurement is perfomed, the number of measurements is equal to the number of mapped 5QIs.

e) The measurement name has the form DRB.UEUnresVolUl or optionally DRB.UEUnresVolUl.*QOS,* where *QOS* identifies the target quality of service class.

f) NRCellDU

g) Valid for packet switched traffic

h) 5GS

i) One usage of this measurement is for performance assurance within integrity area (user plane connection quality).

5.1.1.4 RRC connection number

5.1.1.4.1 Mean number of RRC Connections

a) This measurement provides the mean number of users in RRC connected mode during each granularity period.

b) SI.

c) This measurement is obtained by sampling at a pre-defined interval, the number of users in RRC connected mode for each NR cell and then taking the arithmetic mean.

d) A single integer value.

e) RRC.ConnMean

f) NRCellCU

g) Valid for packet switched traffic

h) 5GS

i) One usage of this measurement is for monitoring the number of RRC connections in connected mode during the granularity period.

5.1.1.4.2 Max number of RRC Connections

a) This measurement provides the maximum number of users in RRC connected mode during each granularity period.

b) SI.

c) This measurement is obtained by sampling at a pre-defined interval, the number of users in RRC connected mode for each NR cell and then taking the maximum.

d) A single integer value.

e) RRC.ConnMax

f) NRCellCU

g) Valid for packet switched traffic

h) 5GS

i) One usage of this measurement is for monitoring the number of RRC connections in connected mode during the granularity period.

5.1.2 Performance measurements valid only for non-split gNB deployment scenario

5.1.3 Performance measurements valid for split gNB deployment scenario

5.1.3.1 Packet Loss Rate

5.1.3.1.1 UL PDCP SDU Loss Rate

a) This measurement provides the fraction of PDCP SDU packets which are not successfully received at gNB-CU-UP. It is a measure of the UL packet loss including any packet losses in the air interface, in the gNB-CU and on the F1-U interface. Only user-plane traffic (DTCH) and only PDCP SDUs that have entered PDCP (and given a PDCP sequence number) are considered. The measurement is optionally split into subcounters per QoS level (mapped 5QI or QCI in NR option 3).

b) SI.

c) This measurement is obtained as: 1000000* Number of missing UL PDCP sequence numbers, representing packets that are not delivered to higher layers, of a data radio bearer, divided by Total number of UL PDCP sequence numbers (also including missing sequence numbers) of a bearer, starting from the sequence number of the first packet delivered by UE PDCP to gNB-CU-UP until the sequence number of the last packet. If transmission of a packet might continue in another cell, it shall not be included in this count. Separate counters are optionally maintained for mapped 5QI (or QCI for NR option 3).

d) Each measurement is an integer value representing the loss rate multiplied by 1E6. The number of measurements is equal to one. If the optional QoS level measurements is perfomed, the measurements are equal to the number of mapped 5QIs.

e) The measurement name has the form DRB.PacketLossRateUl or optionally DRB.PacketLossRateUl.*QOS* where *QOS* identifies the target quality of service class.

f) GNBCUUPFunction.

NRCellCU.

g) Valid for packet switched traffic.

h) 5GS.

i) One usage of this measurement is for performance assurance within integrity area (user plane connection quality). NRCellCU measurement applies only for 2-split deployment.

5.1.3.1.2 UL F1-U Packet Loss Rate

a) This measurement provides the fraction of PDCP SDU packets which are not successfully received at gNB-CU-UP. It is a measure of the UL packet loss on the F1-U interface. The measurement is optionally split into subcounters per QoS level (mapped 5QI or QCI in NR option 3).

b) SI

c) This measurement is obtained as: 1000000* Number of missing UL GTP sequence numbers (TS 29.281), representing packets that are not delivered to higher layers, of a data radio bearer, divided by Total number of UL GTP sequence numbers (also including missing sequence numbers) of a bearer, starting from the GTP sequence number of the first packet delivered by gNB-DU to gNB-CU-UP until the GTP sequence number of the last packet. Separate counters are optionally maintained for mapped 5QI (or QCI for option 3).

d) Each measurement is an integer value representing the loss rate multiplied by 1E6. The number of measurements is equal to one. If the optional QoS level measurement is perfomed, the measurements are equal to the number of mapped 5QIs.

e) The measurement name has the form DRB.F1UPacketLossRateUl or optionally DRB.F1UPacketLossRateUl.*QOS* where *QOS* identifies the target quality of service class.

f) GNBCUUPFunction

g) Valid for packet switched traffic

h) 5GS

i) One usage of this measurement is for performance assurance within integrity area (user plane connection quality).

5.1.3.1.3 DL F1-U Packet Loss Rate

a) This measurement provides the fraction of PDCP SDU packets which are not successfully received at the gNB-DU). It is a measure of the DL packet loss on the F1-U interface. The measurement is optionally split into subcounters per QoS level (mapped 5QI or QCI in NR option 3).

b) SI

c) This measurement is obtained as: 1000000* Number of missing DL GTP sequence numbers (TS 29.281), representing packets that are not delivered to lower layers, of a data radio bearer, divided by Total number of UL GTP sequence numbers (also including missing sequence numbers) of a bearer, starting from the sequence number of the first packet delivered by gNB-CU-UP to gNB-DU until the GTP sequence number of the last packet. Separate counters are optionally maintained for mapped 5QI (or QCI for NR option 3).

d) Each measurement is an integer value representing the loss rate multiplied by 1E6. The number of measurements is equal to one. If the optional QoS level measurement is perfomed, the measurements are equal to the number of mapped 5QIs.

e) The measurement name has the form DRB.F1UPacketLossRateDl or optionally DRB.F1UPacketLossRateDl.*QOS* where *QOS* identifies the target quality of service class.

f) NRCellDU

g) Valid for packet switched traffic

h) 5GS

i) One usage of this measurement is for performance assurance within integrity area (user plane connection quality).

5.1.3.2 Packet Drop Rate

5.1.3.2.1 DL PDCP SDU Drop rate in gNB-CU-UP

a) This measurement provides the fraction of PDCP SDU packets which are dropped on the downlink, due to congestion, traffic management etc in the gNB-CU-UP. Only user-plane traffic (DTCH) is considered. A dropped packet is one whose context is removed from the gNB-CU-UP without any part of it having been transmitted on the F1-U or Xn-U or X2-U interface. The measurement is optionally split into subcounters per QoS level (mapped 5QI or QCI in NR option 3).

NOTE: this measurement may include packets that were supposed to be sent via the eUtran air interface if using NR split bearer option 3, 4 or 7.

b) SI.

c) This measurement is obtained as: 1000000*Number of DL packets, for which no part has been transmitted over the F1-U or Xn-U or X2-U interface, of a data radio bearer, that are discarded in the PDCP layer, divided by Number of DL packets for data radio bearers that has entered PDCP upper SAP. Separate counters are optionally maintained for mapped 5QI (or QCI for NR option 3).

d) Each measurement is an integer value representing the drop rate multiplied by 1E6. The number of measurements is equal to one. If the optional QoS level measurement is perfomed, the measurements are equal to the number of mapped 5QIs.

e) The measurement name has the form DRB.PdcpPacketDropRateDl or optionally DRB.PdcpPacketDropRateDl.*QOS*
where *QOS* identifies the target quality of service class.

f) GNBCUUPFunction.

NRCellCU.

g) Valid for packet switched traffic.

h) 5GS.

i) One usage of this measurement is for performance assurance within integrity area (user plane connection quality). NRCellCU measurement applies only for 2-split deployment.

5.1.3.2.2 DL Packet Drop Rate in gNB-DU

a) This measurement provides the fraction of RLC SDU packets which are dropped on the downlink, due to congestion, traffic management etc in the gNB-DU. Only user-plane traffic (DTCH) is considered. A dropped packet is one whose context is removed from the gNB-DU without any part of it having been transmitted on the air interface. The measurement is optionally split into subcounters per QoS level (mapped 5QI or QCI in NR option 3).

b) SI

c) This measurement is obtained as: 1000000*Number of DL packets, for which no part has been transmitted over the air, of a data radio bearer, that are discarded in the gNB-DU divided by Number of DL packets for data radio bearers that were received from gNB-CU-UP. Separate counters are optionally maintained for mapped 5QI (or QCI for NR option 3).

d) Each measurement is an integer value representing the drop rate multiplied by 1E6. The number of measurements is equal to one. If the optional QoS level measurement is perfomed, the measurements are equal to the number of mapped 5QIs.

e) The measurement name has the form DRB.RlcPacketDropRateDl or optionally DRB.RlcPacketDropRateDl.*QOS*
where *QOS* identifies the target quality of service class.

f) NRCellDU

g) Valid for packet switched traffic

h) 5GS

i) One usage of this measurement is for performance assurance within integrity area (user plane connection quality).

5.1.3.3 Packet Delay

5.1.3.3.1 Average delay DL in CU-UP

a) This measurement provides the average (arithmetic mean) PDCP SDU delay on the downlink within the gNB-CU-UP, for all PDCP packets. The measurement is optionally split into subcounters per QoS level (mapped 5QI or QCI in NR option 3).

b) DER (n=1)

c) This measurement is obtained as: sum of (time when sending a PDCP SDU to the gNB-DU at the egress PDCP layer on F1-U/Xn-U, minus time of arrival of the same packet at NG-U ingress IP termination) divided by total number of PDCP SDUs arriving at NG-U ingress IP termination. Separate counters are optionally maintained for each mapped 5QI (or QCI for option 3).

d) Each measurement is an integer representing the mean delay in microseconds. The number of measurements is equal to one. If the optional QoS level measurement is perfomed, the number of measurements is equal to the number of mapped 5QIs.

e) The measurement name has the form DRB.PdcpSduDelayDl or optionally DRB.PdcpSduDelayDl.*QOS* where *QOS* identifies the target quality of service class.

f) GNBCUUPFunction

g) Valid for packet switched traffic

h) 5GS

i) One usage of this measurement is for performance assurance within integrity area (user plane connection quality).

5.1.3.3.2 Average delay on F1-U

a) This measurement provides the average (arithmetic mean) GTP packet delay on the F1-U interface. The measurement is optionally split into subcounters per QoS level (mapped 5QI or QCI in NR option 3).

b) DER (n=1)

c) This measurement is obtained as: the time when receiving a GTP packet delivery status message from the gNB-DU at the egress GTP termination, minus time when sending the same packet to gNB-DU at the GTP ingress termination, minus feedback delay time in gNB-DU, obtained result is divided by two. Separate counters are optionally maintained for each mapped 5QI (or QCI for option 3).

d) Each measurement is an integer representing the mean delay in microseconds. The number of measurements is equal to one. If the optional QoS level measurement is perfomed, the number of measurements is equal to the number of mapped 5QIs.

e) The measurement name has the form DRB.PdcpF1Delay or optionally DRB.PdcpF1Delay.*QOS* where *QOS* identifies the target quality of service.

f) GNBCUUPFunction

g) Valid for packet switched traffic

h) 5GS

i) One usage of this measurement is for performance assurance within integrity area (user plane connection quality).

5.1.3.3.3 Average delay DL in gNB-DU

a) This measurement provides the average (arithmetic mean) RLC SDU delay on the downlink within the gNB-DU, for initial transmission of all RLC packets. The measurement is optionally split into subcounters per QoS level (mapped 5QI or QCI in NR option 3).

b) DER (n=1)

c) This measurement is obtained as: sum of (time when the last part of an RLC SDU was scheduled and sent to the MAC layer for transmission over the air, minus time of arrival of the same packet at the RLC ingress F1-U termination) divided by total number of RLC SDUs arriving at the RLC ingress F1-U termination. If the RLC SDU needs retransmission (for Acknowledged Mode) the delay will still include only one contribution (the original one) to this measurement. Separate counters are optionally maintained for each mapped 5QI (or QCI for option 3). Each measurement is an integer representing the mean delay in microseconds.

d) The number of measurements is equal to one. If the optional QoS level measurement is perfomed, the number of measurements is equal to the number of mapped 5QIs.

e) The measurement name has the form DRB.RlcSduDelayDl or optionally DRB.RlcSduDelayDl.*QOS* where *QOS* identifies the target quality of service class.

f) NRCellDU

g) Valid for packet switched traffic

h) 5GS

i) One usage of this measurement is for performance assurance within integrity area (user plane connection quality).

5.1.3.4 IP Latency measurements

5.1.3.4.1 General information

The measurement named “IP Latency in DL (gNB-DU)” defines the DL latency in gNB-DU. DL latency measurements for CU-UP and F1-U are not defined.

5.1.3.4.2 IP Latency DL in gNB-DU

a) a) This measurement provides the average IP Latency in DL (arithmetic mean) within the gNB-DU, when there is no other prior data to be transmitted to the same UE in the gNB-DU. The measurement is optionally split into subcounters per QoS level.

b) b) DER (n=1)

c) c) This measurement is obtained as: sum of (time when the first piece of an RLC SDU transmitted on the air interface, minus time of arrival of the same packet at the RLC ingress F1-U termination, for IP packets arriving when there is no other prior data to be transmitted to the same UE in the gNB-DU) divided by total number of RLC SDUs arriving at the RLC ingress F1-U termination when there is no other prior data to be transmitted to the same UE in the gNB-DU. Separate counters are optionally maintained for each mapped 5QI (or QCI for option 3).

d) d) Each measurement is an integer representing the average latency in microseconds. The number of measurements is equal to one. If the optional QoS level measurement is performed, the number of measurements is equal to the number of supported mapped 5QIs.

e) e) The measurement name has the form DRB.RlcSduLatencyDl or optionally DRB.RlcSduLatencyDl.*QOS* where *QOS* identifies the target quality of service class

f) f) NRCellDU

g) g) Valid for packet switched traffic

h) h) 5GS

i) i) One usage of this measurement is for performance assurance within integrity area (user plane connection quality).

5.1.3.5 UE Context Release

5.1.3.5.1 UE Context Release Request (gNB-DU initiated)

a) This measurement provides the number of UE CONTEXT Release initiated by gNB-DU for each release cause.

b) SI

c) Transmission of an UE CONTEXT RELEASE REQUEST message initiated by gNB-DU. Each release request is to be added to the relevant cause measurement. The possible causes are defined in 38.473 [6]. The sum of all supported per causes measurements shall equal the total number of UE CONTEXT Release initiated by gNB-DU. In case only a subset of per cause measurements is supported, a sum subcounter will be provided first.

d) Each measurement is an integer value. The number of measurements is equal to the number of causes plus a possible sum value identified by the .sum suffix.

e) The measurement name has the form UECNTX.RelReq.*Cause*
where *Cause* identifies the release cause.

f) NRCellDU

g) Valid for packet switched traffic

h) 5GS

5.1.3.5.2 Number of UE Context Release Requests (gNB-CU initiated)

a) This measurement provides the number of UE CONTEXT RELEASE initiated by gNB-CU for each release cause.

b) SI

c) Transmission of an UE CONTEXT RELEASE COMMAND message initiated by gNB-CU. Each release request is to be added to the relevant cause measurement. The possible causes are defined in 38.473 [6]. The sum of all supported per causes measurements shall equal the total number of UE CONTEXT Release initiated by gNB-CU. In case only a subset of per cause measurements is supported, a sum subcounter will be provided first.

d) Each measurement is an integer value. The number of measurements is equal to the number of causes plus a possible sum value identified by the .sum suffix.

e) The measurement name has the form UECNTX.RelCmd.Cause where Cause identifies the release cause.

f) NRCellCU

g) Valid for packet switched traffic

h) 5GS

5.2 Performance measurements for AMF

5.2.1 Registered subscribers measurement

5.2.1.1 Mean number of registered subscribers

a) This measurement provides the mean number of registered state subscribers per AMF

b) SI

c) This measurement is obtained by sampling at a pre-defined interval the number of registered subscribers in an AMF and then taking the arithmetic mean. The measurement can be split into subcounters per NSI identifier (S-NSSAI).

d) A single integer value

a) e) RM.RegisteredSubNbrMean.*SNSSAI*

Where *SNSSAI* identifies the NSI

f) AMFFunction

g) Valid for packet switching

h) 5GS

5.2.1.2 Maximum number of registered subscribers

a) This measurement provides the maximum number of registered state subscribers per AMF

b) SI

c) This measurement is obtained by sampling at a pre-defined interval the number of registered subscribers in an AMF and then taking the maximum. The measurement can be split into subcounters per NSI identifier (S-NSSAI).

d) A single integer value

a) e) RM.RegisteredSubNbrMax.*SNSSAI*

Where *SNSSAI* identifies the NSI

f) AMFFunction

g) Valid for packet switching

h) 5GS

5.2.2 Registration procedure related measurements

5.2.2.1 Number of initial registration requests

a) a) This measurement provides the number of initial registration requests received by the AMF.

b) b) CC

c) c) On receipt by the AMF from the UE of Registration Request with the registration type indicating an initial registration (see clause 4.2.2.2.2 of 3GPP TS 23.502 [7]). Each initial registration request is added to the relevant subcounter per NSI identifier (S-NSSAI).

d) d) Each subcounter is an integer value

e) e) RM.RegInitReq.*SNSSAI*

f) Where *SNSSAI* identifies the NSI;

g) f) AMFFunction

h) g) Valid for packet switched traffic

i) h) 5GS

j) i) One usage of this performance measurements is for performance assurance.

5.2.2.2 Number of successful initial registrations

a) a) This measurement provides the number of successful initial registrations at the AMF.

b) b) CC

c) c) On transmission of Registration Accept by the AMF to the UE that sent the initial registration request (see 3GPP TS 23.502 [7]). Each accepted initial registration is added to the relevant subcounter per NSI identifier (S-NSSAI).

d) d) Each subcounter is an integer value

e) e) RM.RegInitSucc.*SNSSAI*

f) Where *SNSSAI* identifies the NSI;

g) f) AMFFunction

h) g) Valid for packet switched traffic

i) h) 5GS

j) i) One usage of this performance measurements is for performance assurance.

5.2.2.3 Number of mobility registration update requests

a) a) This measurement provides the number of mobility registration update requests received by the AMF.

b) b) CC

c) c) On receipt by the AMF from the UE of Registration Request with the registration type indicating a Mobility Registration Update (see clause 4.2.2.2.2 of 3GPP TS 23.502 [7]). Each mobility registration update request is added to the relevant subcounter per NSI identifier (S-NSSAI).

d) d) Each subcounter is an integer value

e) e) RM.RegMobReq.*SNSSAI*

f) Where *SNSSAI* identifies the NSI;

g) f) AMFFunction

h) g) Valid for packet switched traffic

i) h) 5GS

j) i) One usage of this performance measurements is for performance assurance.

5.2.2.4 Number of successful mobility registration updates

a) a) This measurement provides the number of successful mobility registration updates at the AMF.

b) b) CC

c) c) On transmission of Registration Accept by the AMF to the UE that sent the mobility registration update request (see 3GPP TS 23.502 [7]). Each accepted mobility registration update is added to the relevant subcounter per NSI identifier (S-NSSAI).

d) d) Each subcounter is an integer value

e) e) RM.RegMobSucc.*SNSSAI*

f) Where *SNSSAI* identifies the NSI;

g) f) AMFFunction

h) g) Valid for packet switched traffic

i) h) 5GS

j) i) One usage of this performance measurements is for performance assurance.

5.2.2.5 Number of periodic registration update requests

a) a) This measurement provides the number of periodic registration update requests received by the AMF.

b) b) CC

c) c) On receipt by the AMF from the UE of Registration Request with the registration type indicating a Periodic Registration Update (see clause 4.2.2.2.2 of 3GPP TS 23.502 [7]). Each periodic registration update request is added to the relevant subcounter per NSI identifier (S-NSSAI).

d) d) Each subcounter is an integer value

e) e) RM.RegPeriodReq.*SNSSAI*

f) Where *SNSSAI* identifies the NSI;

g) f) AMFFunction

h) g) Valid for packet switched traffic

i) h) 5GS

j) i) One usage of this performance measurements is for performance assurance.

5.2.2.6 Number of successful periodic registration updates

a) a) This measurement provides the number of successful mobility registration updates at the AMF.

b) b) CC

c) c) On transmission of Registration Accept by the AMF to the UE that sent the periodic registration update request (see 3GPP TS 23.502 [7]). Each accepted periodic registration update is added to the relevant subcounter per NSI identifier (S-NSSAI).

d) d) Each subcounter is an integer value

e) e) RM.RegPeriodSucc.*SNSSAI*

f) Where *SNSSAI* identifies the NSI;

g) f) AMFFunction

h) g) Valid for packet switched traffic

i) h) 5GS

j) i) One usage of this performance measurements is for performance assurance.

5.2.2.7 Number of emergency registration requests

a) a) This measurement provides the number of emergency registration requests received by the AMF.

b) b) CC

c) c) On receipt by the AMF from the UE of Registration Request with the registration type indicating an Emergency Registration (see clause 4.2.2.2.2 of 3GPP TS 23.502 [7]). Each emergency registration request is added to the relevant subcounter per NSI identifier (S-NSSAI).

d) d) Each subcounter is an integer value

e) e) RM.RegEmergReq.*SNSSAI*

f) Where *SNSSAI* identifies the NSI;

g) f) AMFFunction

h) g) Valid for packet switched traffic

i) h) 5GS

j) i) One usage of this performance measurements is for performance assurance.

5.2.2.8 Number of successful emergency registrations

a) a) This measurement provides the number of successful emergency registrations at the AMF.

b) b) CC

c) c) On transmission Registration Accept by the AMF to the UE that sent the emergency registration request (see 3GPP TS 23.502 [7]). Each accepted emergency registration is added to the relevant subcounter per NSI identifier (S-NSSAI).

d) d) Each subcounter is an integer value

e) e) RM.RegEmergSucc.*SNSSAI*

f) Where *SNSSAI* identifies the NSI;

g) f) AMFFunction

h) g) Valid for packet switched traffic

i) h) 5GS

j) i) One usage of this performance measurements is for performance assurance.

5.3 Performance measurements for SMF

5.3.1 Session Management

5.3.1.1 Number of PDU sessions (Mean)

a) a) This measurement provides the mean number of PDU sessions.

b) b) SI

c) c) The measurement is obtained by sampling at a pre-defined interval, the number of PDU sessions established by SMF, and then taking the arithmetic mean. The measurement is optionally split into subcounters per NSI identifier (S-NSSAI).

d) d) A single integer value

e) e) SM.SessionNbrMean.*SNSSAI*
Where *SNSSAI* identifies the NSI

f) f) SMFFunction

g) g) Valid for packet switched traffic

h) h) 5GS

5.3.1.2 Number of PDU sessions (Maximum)

a) a) This measurement provides the max number of PDU sessions.

b) b) SI

c) c) The measurement is obtained by sampling at a pre-defined interval, the number of PDU sessions established by SMF, and then selecting the maximum value. The measurement is optionally split into subcounters per NSI identifier (S-NSSAI).

d) d) A single integer value

e) e) SM.SessionNbrMax.*SNSSAI*
Where *SNSSAI* identifies the NSI

f) f) SMFFunction

g) g) Valid for packet switched traffic

h) h) 5GS

i)

5.3.1.3 Number of PDU session creation requests

a) This measurement provides the number of PDU sessions requested to be created by the SMF.

b) CC

c) On receipt by the SMF from AMF of Nsmf_PDUSession_CreateSMContext Request (see 3GPP TS 23.502 [7]). Each PDU session requested to be created is added to the relevant subcounter per NSI identifier (S-NSSAI) and the relevant subcounter per establishment cause.

d) Each subcounter is an integer value

e) SM.PduSessionCreationReqNSI.*SNSSAI*

Where *SNSSAI* identifies the NSI;

SM.PduSessionCreationReqCause.*cause*

Where *cause* indicates the establishment cause for the PDU session.

f) SMFFunction

g) Valid for packet switched traffic

h) 5GS

i) One usage of this performance measurements is for performance assurance.

5.3.1.4 Number of successful PDU session creations

a) This measurement provides the number of PDU sessions successfully created by the SMF.

b) CC

c) On transmission by the SMF to AMF of Nsmf_PDUSession_CreateSMContext Response that indicates a successful PDU session creation (see 3GPP TS 23.502 [7]). Each PDU session successfully created is added to the relevant subcounter per NSI identifier (S-NSSAI) and the relevant subcounter per establishment cause.

d) Each subcounter is an integer value

e) SM.PduSessionCreationSuccNSI.*SNSSAI*

Where *SNSSAI* identifies the NSI;

SM.PduSessionCreationSuccCause.*cause*

Where *cause* indicates the establishment cause for the PDU session.

f) SMFFunction

g) Valid for packet switched traffic

h) 5GS

i) One usage of this performance measurements is for performance assurance.

5.3.1.5 Number of failed PDU session creations

a) This measurement provides the number of PDU sessions failed to be created by the SMF.

b) CC

c) On transmission by the SMF to AMF of Nsmf_PDUSession_CreateSMContext Response that indicates a rejected PDU session creation (see 3GPP TS 23.502 [7]). Each PDU session rejected to be created is added to the relevant subcounter per rejection cause.

d) Each subcounter is an integer value

e) SM.PduSessionCreationFail.*cause*

Where *cause* indicates the rejection cause for the PDU session.

f) SMFFunction

g) Valid for packet switched traffic

h) 5GS

i) One usage of this performance measurements is for performance assurance.

5.4 Performance measurements for UPF

5.4.1 N3 interface related measurements

5.4.1.1 Number of incoming GTP data packets on the N3 interface, from (R)AN to UPF

a) This measurement provides the number of GTP data PDUs on the N3 interface which have been accepted and processed by the GTP-U protocol entity on the N3 interface.

b) CC

c) Reception by the UPF of a GTP-U data PDU on the N3 interface from the (R)AN. See TS 23.501 [4].

d) A single integer value.

e) GTP.InDataPktN3UPF

f) EP_N3

g) Valid for packet switching.

h) 5GS

5.4.1.2 Number of outgoing GTP data packets of on the N3 interface, from UPF to (R)AN

a) This measurement provides the number of GTP data PDUs on the N3 interface which have been generated by the GTP-U protocol entity on the N3 interface.

b) CC

c) Transmission by the UPF of a GTP-U data PDU of on the N3 interface to the (R)AN. See TS 23.501 [4].

d) A single integer value.

e) GTP.OutDataPktN3UPF

f) EP_N3

g) Valid for packet switching.

h) 5GS

5.4.1.3 Number of octets of incoming GTP data packets on the N3 interface, from (R)AN to UPF

a) This measurement provides the number of octets of incoming GTP data packets on the N3 interface which have been generated by the GTP-U protocol entity on the N3 interface. The measurement can optionally be split into subcounters per S-NSSAI.

b) CC

c) Reception by the UPF of a GTP-U data PDU on the N3 interface from (R)AN. See TS 23.501 [4].

d) Each measurement is a single integer value, the number of measurements is equal to one. If the optional S-NSSAI subcounter measurements are perfomed, the number of measurements is equal to the number of supported S-NSSAIs.

e) GTP.InDataOctN3UPF and optionally GTP.OutDataOctN3UPF.*SNSSAI*, where *SNSSAI* identifies the S-NSSAI.

f) EP_N3

g) Valid for packet switching

h) 5GS

5.4.1.4 Number of octets of outgoing GTP data packets on the N3 interface, from UPF to (R)AN

a) This measurement provides the number of octets of outgoing GTP data packets on the N3 interface which have been generated by the GTP-U protocol entity on the N3 interface. The measurement can optionally be split into subcounters S-NSSAI.

b) CC

c) Transmission by the UPF of a GTP-U data PDU on the N3 interface to the(R)AN, .See TS 23.501 [4].

d) Each measurement is a single integer value, the number of measurements is equal to one. If the optional S-NSSAI subcounter measurements are perfomed, the number of measurements is equal to the number of supported S-NSSAIs.

e) GTP.OutDataOctN3UPF and optionally GTP.OutDataOctN3UPF.*SNSSAI*, where *SNSSAI* identifies the S-NSSAI.

f) EP_N3

g) Valid for packet switching

h) 5GS

5.4.2 N6 related measurements

5.4.2.1 N6 incoming link usage

a) This measurement provides the PDU-layer incoming link usage of N6 interface.

b) CC

c) See clause 2.3.4 for IP packet. Definition: IP-type-P (broad spectrum of packet types) Link Usage in IETF RFC 5136 [5].

NOTE: How to measure the unstructured data type is not specified in the present document.

d) Each measurement is an integer value.

e) IP.N6IncLinkUsage.*N6RP*
where *N6RP* identifies the N6 reference point of this UPF, the format of *N6RP* is vendor specific.

f) EP_N6

g) Valid for packet switched traffic.

h) 5GS

5.4.2.2 N6 outgoing link usage

a) This measurement provides the PDU-layer outcoming link usage of N6 interface.

b) CC

c) See clause 2.3.4 for IP packet. Definition: IP-type-P (broad spectrum of packet types) Link Usage in IETF RFC 5136 [5].

NOTE: How to measure the unstructured data type is not specified in the present document.

d) Each measurement is an integer value.

e) IP.N6OutLinkUsage.*N6RP*
where *N6RP* identifies the N6 reference point of this UPF, the format of *N6RP* is vendor specific.

f) EP_N6

g) Valid for packet switched traffic.

h) 5GS

5.5 Performance measurements for PCF

5.5.1 AM policy association related measurements

5.5.1.1 Number of AM policy association requests

a) This measurement provides the number of AM policy association requests received by the visiting PCF ((V-)PCF).

b) CC

c) On receipt by the PCF from the AMF of Npcf_AMPolicyControl_Create (see 3GPP TS 23.502 [7]). Each AM policy association request is added to the relevant subcounter per NSI identifier (S-NSSAI).

d) Each subcounter is an integer value

e) PA.PolicyAMAssoReq.*SNSSAI*

Where *SNSSAI* identifies the NSI;

f) PCFFunction

g) Valid for packet switched traffic

h) 5GS

i) One usage of this performance measurements is for performance assurance.

5.5.1.2 Number of successful AM policy associations

a) This measurement provides the number of successful AM policy associations at the visiting PCF ((V-)PCF).

b) CC

c) On transmission by the PCF to the AMF of Npcf_AMPolicyControl_Create response (see 3GPP TS 23.502 [7]). Each successful AM policy association is added to the relevant subcounter per NSI identifier (S-NSSAI).

d) Each subcounter is an integer value

e) PA.PolicyAMAssoSucc.*SNSSAI*

Where *SNSSAI* identifies the NSI;

f) PCFFunction

g) Valid for packet switched traffic

h) 5GS

i) One usage of this performance measurements is for performance assurance.

5.5.2 SM policy association related measurements

5.5.2.1 Number of SM policy association requests

a) This measurement provides the number of SM policy association requests received by the PCF.

b) CC

c) On receipt by the PCF from the SMF of Npcf_SMPolicyControl_Create (see 3GPP TS 23.502 [7]). Each SM policy association request is added to the relevant subcounter per NSI identifier (S-NSSAI).

d) Each subcounter is an integer value

e) PA.PolicySMAssoReq.*SNSSAI*

Where *SNSSAI* identifies the NSI;

f) PCFFunction

g) Valid for packet switched traffic

h) 5GS

i) One usage of this performance measurements is for performance assurance.

5.5.2.2 Number of successful SM policy associations

a) This measurement provides the number of successful SM policy associations at the PCF.

b) CC

c) On transmission by the PCF to the SMF of Npcf_SMPolicyControl_Create response (see 3GPP TS 23.502 [7]). Each successful SM policy association is added to the relevant subcounter per NSI identifier (S-NSSAI).

d) Each subcounter is an integer value

e) PA.PolicySMAssoSucc.*SNSSAI*

Where *SNSSAI* identifies the NSI;

f) PCFFunction

g) Valid for packet switched traffic

h) 5GS

i) One usage of this performance measurements is for performance assurance.

5.6 Performance measurements for UDM

5.6.1 Mean number of registered subscribers through UDM

a) This measurement provides the mean number of registered subscribers to UDM .

b) SI.

c) This measurement is obtained by sampling at a unified interval the number of registered subscribers in a UDM and then taking the arithmetic mean.

d) A single integer value.

e) RM.RegisteredSubUDMNbrMean.

f) UDMFunction.

g) Valid for packet switching.

h) 5GS.

5.6.2 Maximum number of registered subscribers through UDM

a) This measurement provides the maximum number of registered subscribers to UDM .

b) SI

c) This measurement is obtained by sampling at an unified interval the number of registered subscribers in the UDM and then taking the maximum.

d) A single integer value.

e) RM.RegisteredSubUDMNbrMax.

f) UDMFunction.

g) Valid for packet switching.

h) 5GS.

6 Measurements related to end-to-end 5G network and network slicing

6.1 Void

6.2 Virtualised resource usage measurement

a) This measurement provides the mean usage of virtualised resource (e.g. processor, memory, disk) in single network slice instance during the granularity period.

b) OM

c) This measurement is generated with .sum suffix for the usage of each virtualised NF (see 3GPP TS 32.426 [1]) related to single network slice instance by taking the weighted average. The algorithm of the weighted average is vendor specific.

d) Each measurement is an real value (Unit:%).

e) MeanProcessorUsage

MeanMemoryUsage

MeanDiskUsage

f) Performance measurement service.

g) Packet Switched.

h) 5GS

NOTE: The name of service in f) needs to align with the TS (e.g., 28.550) defining the management service.

Annex A (informative): Use cases for performance measurements

A.1 Monitoring of UL and DL user plane latency in NG-RAN

Satisfying low latency expectations for 5G services, such as URLLC, is one of the key tasks for the operator to meet service performance expectations. As the performance in UL and DL differs, it is important for operators to be able to monitor the UL and DL user plane latencies separately. With performance measurements allowing the operator to obtain or derive the UL and DL user plane latency information separately, the operators can pinpoint the services performance problems to specific problems in UL or DL.

The DL IP latency monitoring in NG-RAN refers to the transmission within gNB of IP packets arriving when there is no other prior data to be transmitted to the same UE in the gNB.

To further pinpoint performance problem detected, separate counters may be provided per mapped 5QI (which are particularly useful when the mapped 5QI is used by few services and users and the packet size does not vary much).

A.2 Monitoring of UL and DL packet loss in NG-RAN

Keeping track of UL and DL packet loss in the NG-RAN is essential, since for certain services packets that are lost along the way through the system may have a noticeable impact on the end user. UL and DL packet loss measurements can be useful for evaluation, optimization and for performance assurance within the integrity area (user plane connection quality).

UL packet loss is a measure of packets dropped in the UE and the packets lost on the interfaces (air interface and F1-U interface). If parts of the gNB are deployed in a virtualized environment, it is important to measure also the F1-U UL interface packet loss in a separate measurement, to be able to pinpoint the reason for high packet loss.

A.3 Monitoring of DL packet drop in NG-RAN

Keeping track of DL packet drops in the NG-RAN is essential, since for certain services packets that are dropped along the way through the system may have a noticeable impact on the end user. DL packet drop measurements can be useful for evaluation, optimization and for performance assurance of the network.

For gNBs that are deployed in a split architecture, e.g. when parts of a gNB are deployed in a virtualized environment, the DL packet drops may occur in two parts; the gNB CU-UP and the gNB DU. Therefore, it is important to measure this separately.

A.4 Monitoring of UL and DL user plane delay in NG-RAN

Satisfying low packet delay is of prime concern for some services, particularly conversational services like speech and instant messaging. As the performance in UL and DL differs, it is important for operators to be able to monitor the UL and DL user plane delay separately. With performance measurements allowing the operator to obtain or derive the UL and DL user plane delay information separately, the operators can pinpoint the services performance problems to specific problems in UL or DL.

The DL delay monitoring in gNB refers to the average delay of any packet within NG-RAN, including air interface delay until the UE receives the packet. A gNB deployed in a split architecture, the user plane delay will occur in gNB-CU-UP, on the F1 interface, in gNB-DU and on the air interface. Therefore, four gNB related delay measurements needs to be monitored for the DL delay to pinpoint where end user impact from packet delay occurs.

To further pinpoint a detected delay performance problem, the packet delay measurement separation may be based on mapped 5QI (or for QCI in case of NR option 3).

A.5 Monitoring of UE Context Release Request (gNB-DU initiated)

In order to monitor the stability of the network and detect the service/connection interruption caused by NGRAN, monitoring the UE Context Release Request initiated by gNB-DU is an effective method. Collecting the measurement information of the message and analysing the releasing cause conveyed in the message, operators could detect the stability of NG-RAN, and could decide a specific means to improve the NG-RAN performance.

A.6 Monitoring of physical radio resource utilization

The physical radio resource utilization measurements could provide operators the load information of the radio network during the measurement time period. The physical radio resource utilization measurements should reflect the average usage and the usage distribution of the radio resource of the physical layer. The measurements can make the operator to be aware of whether a cell has ever experienced high load or not in the monitoring period, and is a key input to network capacity planning and load balancing.

A.7 Monitoring of RRC connection number

The number of the users in RRC connected and inactive mode need to be monitored as it reflects the load of the radio network, the operators can use this information for dynamic frequency resource allocation or load balance purpose. Moreover, it is an important factor to be evaluated in the radio network capacity enhancement decision-making.

A.8 Monitoring of UE Context Release

In order to monitor the stability of the network and detect the service/connection interruption caused by NG-RAN, monitoring the UE Context Release Request initiated by gNB-DU and UE Context Release Command initiated by gNB-CU is an effective method. Collecting the measurement information of the message and analysing the releasing cause conveyed in the message, operators could detect the stability of NG-RAN, and could decide a specific means to improve the NG-RAN performance.

A.9 Monitoring of UE Throughput in NG-RAN

Keeping track of UL and DL UE throughput in the NG-RAN is essential, to ensure end user satisfaction and well-functioning and well configured cells and scheduling features.

The restricted UE throughput per mapped 5QI will show the scheduling efficiency and QoS priority handling in the gNB and the ratio between unrestricted and restricted volume will show the gNB ability to handle small data transfers efficiently.

To be able to monitor the spread of throughput within the cell, and estimate the ratio of satisfied users, the throughput distribution measurement can be used.

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

A.10 Monitoring of Unrestricted volume in NG-RAN

Measuring the share of unrestricted user data volume in the NG-RAN is important, to show the gNB ability to handle small data transfers efficiently and to see how large share of the volume that is part of the UE throughput measurement. It is not meaningful to measure throughput for data transfers so small that they fit in one single slot but it is still important to know how much such transfers can be handled by the gNB.

When network slicing is supported by the NG-RAN, multiple NSIs may be supported. The share of unrestricted volume for each NSI is then of importance to the operator to pinpoint a specific performance problem.

A.11 N3 data volume related measurements

N3 related measurements are used to measure data volume on N3 interface including incoming and outgoing of GTP data packets without counting the mandatory part of the GTP-U header

It is useful to analyse transport bandwidth usage of N3 interface. If the transport bandwidth usage is too high, more bandwidth should be deployed, or load balance should be considered according to core network dimension if there are multiple UPFs connected to multiple gNodeBs.

So it is necessary to define N3 related measurements.

A.12 N6 related measurements

N6 related measurements are used to measure data volume on N6 interface including incoming and outgoing of IP data packets.

It is useful to analyse transport bandwidth usage of N6 interface. If the transport bandwidth usage is too high, more bandwidth should be deployed.

So it is necessary to define N6 related measurements.

A.13 Registration related measurements

A UE needs to register with the 5GS to get authorization to receive services, to enable mobility tracking and to enable reachability. The following registration types are defined:

- Initial Registration to the 5GS;
- Mobility Registration Update (upon changing to a new Tracking Area (TA) outside the UE's Registration Area in both CM-CONNECTED and CM-IDLE state, or when the UE needs to update its capabilities or protocol parameters that are negotiated in Registration procedure with or without changing to a new TA);
- Periodic Registration Update (due to a predefined time period of inactivity); and
- Emergency Registration (i.e. the UE is in limited service state).

The performance of registration for each registration type needs to be monitored by the operator since it is relevant to whether the end user can use the service of 5GS or a specific network slice.

A.14 PDU session establishment related measurements

The PDU session establishment is one of essential procedures for 5G network. The performance of PDU session establishment directly impacts the QoS of the network and the QoE of the end users. Therefore, the performance measurements are needed to reflect the performance of the PDU session establishment.

The number and success rate of PDU session creations, the number of PDU sessions running on the SMF are some of the basic performance measurements to monitor the performance of the PDU session establishment. And the performance measurements of failed PDU session creations are helpful to solve the network issues in case the performance is below the expectation.

A.15 Policy association related measurements

To ensure the UE properly use the services provided by 5GS, the UE needs to be associated with a policy. The policies are categorized into AM policy and SM policy, and are executed by AMF and SMF respectively. Both kinds of policies are provisioned by PCF.

The AM policy association needs to be established in case the UE initially registers to the network or the UE needs the AMF re-allocation.

The SM policy association needs to be established when the UE requests a PDU Session Establishment.

The policy association establishment is the essential steps allowing the UE to be served by the 5GS under the designed policies, therefore it needs to be monitored.

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
2018-12	SA#82	SP-181047	0002	1	F	Remove the redundant measurement of end-to-end latency KPI	15.1.0
2018-12	SA#82	SP-181047	0024	1	F	Correction of the Packet loss measurements	15.1.0
2019-03	SA#83	SP-190122	0038	2	F	Update performance measurements for UDM	15.2.0
2019-03	SA#83	SP-190122	0048	1	F	Clean-up	15.2.0
2019-03	SA#83	SP-190122	0073	-	F	Correction of percentage unrestricted volume measurements	15.2.0
2019-06	SA#84	SP-190375	0083	-	F	Correction of F1 measurements	15.3.0
2019-09	SA#85	SP-190748	0091	-	F	Correct the definition of Average delay DL air-interface measurement	15.4.0
2019-09	SA#85	SP-190751	0093	1	F	Correction on kbits abbreviation	15.4.0
2019-12	SA#86	SP-191174	0134	-	F	Correction of Registered subscribers measurement for AMF	15.5.0
2020-07	SA#88-e	SP-200492	0202	-	F	Adding Per Slice N3 Measurements	15.6.0
2024-06	SA#104	SP-240822	0540	-	F	Adding Per Slice N3 Measurements	15.7.0

History

Document history		
V15.0.0	October 2018	Publication
V15.1.0	April 2019	Publication
V15.2.0	May 2019	Publication
V15.3.0	June 2019	Publication
V15.4.0	October 2019	Publication
V15.5.0	January 2020	Publication
V15.6.0	August 2020	Publication
V15.7.0	July 2024	Publication