ETSI TS 128 554 V15.1.0 (2019-04)

5G;
Management and orchestration;
5G end to end Key Performance Indicators (KPI)
(3GPP TS 28.554 version 15.1.0 Release 15)

Reference
RTS/TSGS-0528554vf10

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 - NAF 742 C
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° 7803/88

Intellectual Property Rights

Essential patents

IPRs essential or potentially essential to normative deliverables may have been declared to ETSI. The information pertaining to these essential IPRs, if any, is 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 IPR Policy, no investigation, 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.

Foreword

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, UMTS identities or GSM identities. These should be interpreted as being references to the corresponding ETSI deliverables.

The cross reference between GSM, UMTS, 3GPP and ETSI identities can be found under http://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

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".

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].

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

For future update of the document it will also include:

- Retainability (see the definition in [3]).
- Availability.
- Mobility.

5 KPI definitions template

a) Long name (Mandatory): This field shall contain the long and descriptive name of the KPI.

b) Description (Mandatory): This field shall contain the description of the KPI.
Within this field it should be given if the KPI is focusing on network or user view.

c) Logical formula definition (Mandatory):
The logical formula should describe what the KPI formula is in logical way. The description of the formula is given in a written textual format without any measurement or counter names. E.g. a success rate KPI's logical formula is the successful event divided by all event.

d) Physical formula definition (Optional):
This field should contain the KPI formula description using the 3GPP defined counter names.
This field can be used only if the counters needed for the KPI formula is defined in any of the 3GPP TS for performance measurements (TS 28.552 [1], TS 28.553 [2]).

e) Measurement names used for the KPI (Optional):
This clause should list the measurement names used for the KPI.
This clause can be filled out only when the underlying measurements for the KPI formula can be defined, i.e. physical formula definition is available.

f) KPI Object (mandatory)
This clause shall describe the object of the KPI. The object of the KPI is one or some of the following:

- NR and NG-RAN;
- 5GC
- 5GS

g) KPI category (mandatory)
This clause contains the classification of the KPI into one of the KPI categories listed in clause 4.

h) Unit of the KPI (mandatory)
This clause describes the unit of the KPI. The unit can be one of the following:

- percentage;
- time interval (second or millisecond or microsecond);
- Erlang;
- kbit/s.

i) Type of the KPI (Mandatory)
This clause describes the type of the KPI. The KPI type can be one of the following:

- MEAN: This KPI is produced to reflect a mean measurement value based on a number of sample results.
- RATIO: KPI is produced to reflect the percentage of a specific case occurrence to all the cases.
- CUM: This KPI is produced to reflect a cumulative measurement which is always increasing.

j) Remark: (Optional)
This field is for any further information that is needed for the KPI definition.
Here it is proposed to define any additional information that would be needed 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 Registered Subscribers of Network and Network Slice Instance through AMF

a) Registered Subscribers of Single Network Slice Instance through AMF.

b) This KPI describe the total number of subscribers that are registered to a network slice instance.

c) This KPI is obtained by counting the subscribers in AMF that are registered to a network slice instance.

d) $RSSNSI = \sum_{AMF} RegisteredSubNbrMean$

e) RegisteredAMFSubNbrMean

f) 5GS

g) Accessibility

h) Interger

i) CUM

6.2.2 Registered Subscribers of Network and Network Slice Instance through UDM

a) Registered Subscribers of Single Network Slice Instance through UDM

b) This KPI describe the total number of subscribers that are registered to a network slice instance.

c) This KPI is obtained by counting the subscribers in UDM that are registered to a network slice instance.

d) $RSSNSI = \mathrm{Re}\, gisteredSubUDMNbrMean$

e) RegisteredSubUDMNbrMean

f) 5GS

g) Accessibility

h) Integer

i) CUM

6.2.3 Registration success rate of one single network slice instance

a) Registration success rate of one single network slice instance.

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 instance and is used to evaluate accessibility provided by the end-to-end network slice instance and network performance.

c) This KPI is obtained by successful registration procedures divided by attempted registration procedures.

d)

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

e) AMF.5GSRegisAtt.*Type*
AMF.5GSRegisAttachSucc.*Type*

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

f) 5GS

g) Accessibility

h) Percentage

i) RATIO

6.3 Integrity KPI

6.3.1 End-to-end Latency of 5G Network

6.3.1.0 KPI categories

a) End-to-end latency of 5G network.

b) This KPI describes the end to end packet transmission latency through the RAN, CN, and TN part of 5G network and is used to evaluate utilization performance of the end-to-end network.

c) This KPI is the RTT end to end latency of UE IP packets transmitted from UE to the N6 interface in the 5G network. The N6 interface is the reference point between UPF and DN.

d) E2ELatency

e) End-to-end latency

f) 5GS

g) Integrity

h) Time interval (millisecond)

i) MEAN

6.3.1.1 Downlink latency in gNB-DU

a) Downlink latency for IP packets through gNB in split scenario.

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.

c) This KPI 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.

d) DownlinkLat = DRB.RlcSduLatencyDl or optionally DownlinkLat.QoSx = DRB.RlcSduLatencyDl.QoSx where QOS identifies the target quality of service class.

e) DRB.RlcSduLatencyDl, DRB.RlcSduLatencyDl.QoS,

f) NG-RAN

g) Integrity

h) Time interval (microsecond)

i) MEAN

6.3.2 Upstream Throughput for Network and Network Slice Instance

a) Upstream throughput for network and network slice instance.

b) This KPI describes the upstream throughput of one single network slice instance by computing the packet size for each successfully transmitted UL 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.

c) This KPI is obtained by upstream throughput provided by N3 interface from NG-RAN to UPF which is related to the single network slice instance.

d) $UTSNSI = \sum_{AMF} GTP.InDataOctN3UPF$

e) GTP.InDataOctN3UPF

f) 5GS

g) Integrity

h) Kbit/s

i) CUM

6.3.3 Downstream Throughput for Single Network Slice Instance

a) Downstream throughput for network and network slice instance.

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.

c) This KPI is obtained by downstream throughput provided by N3 interface from NG-RAN to UPF which is related to the single network slice instance.

d) $UTSNSI = \sum_{UPF} GTP.OutDataOctN3UPF$

e) GTP.OutDataOctN3UPF

f) 5GS

g) Integrity

h) Kbit/s

i) CUM

6.3.4 Upstream Throughput at N3 interface

a) Upstream GTP data throughput at N3 interface.

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.

c) This KPI is obtained by measuring the GTP data upstream throughput provided by N3 interface from NG-RAN to UPF, during the granularity period.

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

e) GTP.InDataOctN3UPF

f) 5GS

g) Integrity

h) Kbit/s

i) MEAN

6.3.5 Downstream Throughput at N3 interface

a) Downstream GTP data throughput at N3 interface.

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.

c) This KPI is obtained by measuring the GTP data downstream throughput provided by N3 interface from UPF to NG-RAN, during the granularity period.

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

e) GTP.OutDataOctN3UPF

f) 5GS

g) Integrity

h) Kbit/s

i) MEAN

6.3.6 RAN UE Throughput

6.3.6.1 Definition

a) RAN UE Throughput.

b) A KPI that shows how NG-RAN impacts the service quality provided to an end-user.

c) Payload data volume on RLC level per elapsed time unit on the air interface, for transfers restricted by the air interface.

d) RAN UE Throughput DL = DRB.UEThpDl and

RAN UE Throughput UL = DRB.UEThpUl

or optionally RAN UE Throughput DL for single mapped 5QI or QCI = DRB.UEThpDl.QoS and

RAN UE Throughput UL for single mapped 5QI or QCI = DRB.UEThpUl.QoS

e) DRB.UEThpDl, DRB.UEThpUl, DRB.UEThpDl.QoS, DRB.UEThpUl.QoS

f) NG-RAN

g) Integrity

h) kbits/s

i) MEAN

6.3.6.2 Extended 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).

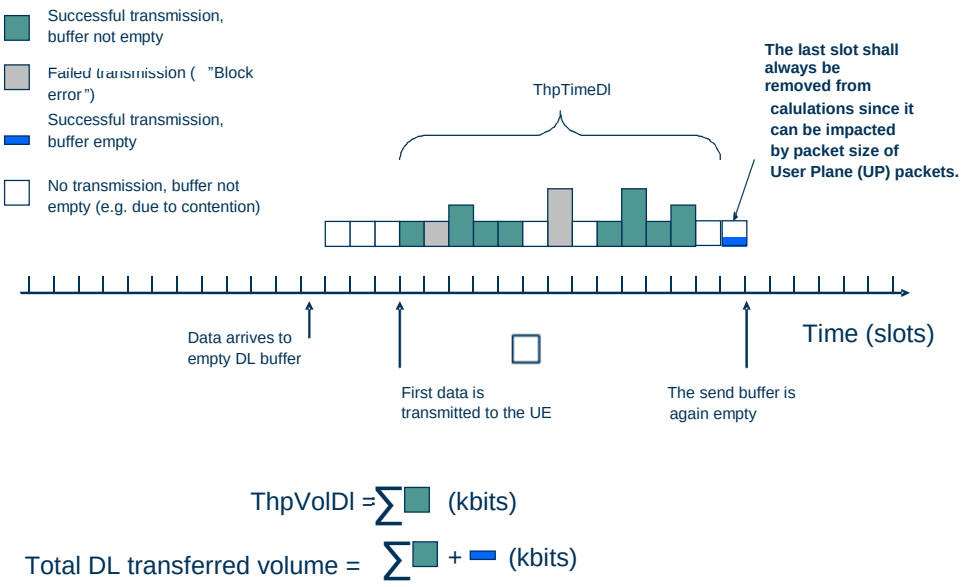

Figure 1

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.

6.4 Utilization KPI

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

a) Mean number of PDU sessions of Single Network Slice Instance.

b) This KPI describes the mean number of PDU sessions that are successfully established in a network slice instance.

c) This KPI is obtained by successful PDU session establishment procedures of SMFs which is related to the network slice instance.

d) PDUSesMeanNbr=Sum (SM.MeanNbrSession) over SMFs.

e) PDUSessionNum

f) 5GS

g) Utilization

h) Erlang

i) MEAN

6.4.2 Virtualised Resource Utilization of Network Slice Instance

a) Virtualised resource utilization of single network slice instance.

b) This KPI describes utilization of virtualised resource (e.g. processor, memory, disk) that are allocated to a network slice instance.

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

c) This KPI is obtained by the usage of virtualised resource (e.g. processor, memory, disk) divided by the system capacity that allocated to the network slice instance.

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

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

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

e) MeanProcessorUsage

MeanMemoryUsage

MeanDiskUsage

System capacity indicates amount of virtualised resource which allocated to the network slice instance.

f) 5GS

g) Utilization KPI

h) Percentage

i) Ratio

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 instance related KPI

Number of registered subscribers of single network slice instance can be used to describe the amount of subscribers that are successfully registered, it can reflect the usage of network slice instance, It is useful to evaluate accessibility performance provided by one single network slice instance 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 instances. This KPI is focusing on both network and user view.

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

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

A.4 Use case for mean PDU sessions number in Network Slice instance

It is necessary to evaluate the mean PDU session number in the network slice instance to indicate system load level. For example, if the mean 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 instance related KPI

It is necessary to evaluate the current utilization of virtualised resources (e.g. memory and storage utilization) that a network slice instance 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 instance 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 instance. If users or subscribers cannot register to the network slice instance, they cannot access any network services in the network slice instance. 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 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.

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

History

Document history		
V15.0.1	October 2018	Publication
V15.1.0	April 2019	Publication