ETSI TS 102 250-3 V1.4.1 (2008-12)

Technical Specification

Speech and multimedia Transmission Quality (STQ);
QoS aspects for popular services in GSM and 3G networks;
Part 3: Typical procedures for Quality of Service
measurement equipment

ETSI

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

Important notice

Individual copies of the present document can be downloaded from:
http://www.etsi.org

The present document may be made available in more than one electronic version or in print. In any case of existing or
perceived difference in contents between such versions, the reference version is the Portable Document Format (PDF).
In case of dispute, the reference shall be the printing on ETSI printers of the PDF version kept on a specific network drive
within ETSI Secretariat.

Users of the present document should be aware that the document may be subject to revision or change of status.
Information on the current status of this and other ETSI documents is available at
http://portal.etsi.org/tb/status/status.asp

If you find errors in the present document, please send your comment to one of the following services:
http://portal.etsi.org/chaircor/ETSI_support.asp

Copyright Notification

Contents

Intellectual Property Rights

IPRs essential or potentially essential to the present document 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 (http://webapp.etsi.org/IPR/home.asp).

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.

Foreword

This Technical Specification (TS) has been produced by ETSI Technical Committee Speech and multimedia Transmission Quality (STQ).

NOTE: Version 2.1.1 of the present document was published with a wrong version number. Please note that the present version 1.4.1 has been re-published to correct this and therefore superceeds the previous versions V1.1.1, V1.2.1 and 2.1.1 which have been withdrawn to avoid confusion.

The present document is part 3 of a multi-part deliverable covering the QoS aspects for popular services in GSM and 3G networks, as identified below:

Part 1: "Identification of Quality of Service aspects";

Part 2: "Definition of Quality of Service parameters and their computation";

Part 3: "Typical procedures for Quality of Service measurement equipment";

Part 4: "Requirements for Quality of Service measurement equipment";

Part 5: "Definition of typical measurement profiles";

Part 6: "Post processing and statistical methods";

Part 7: "Network Based Quality of Service Measurements".

Part 1 identifies QoS aspects for popular services in GSM and 3G networks. For each service chosen QoS indicators are listed. They are considered to be suitable for the quantitatively characterization of the dominant technical QoS aspects as experienced from the end-customer perspective.

Part 2 defines QoS parameters and their computation for popular services in GSM and 3G networks. The technical QoS indicators, listed in part 1, are the basis for the parameter set chosen. The parameter definition is split into two parts: the abstract definition and the generic description of the measurement method with the respective trigger points. Only measurement methods not dependent on any infrastructure provided are described in the present document. The harmonized definitions given in the present document are considered as the prerequisites for comparison of QoS measurements and measurement results.

Part 3 describes typical procedures used for QoS measurements over GSM, along with settings and parameters for such measurements.

Part 4 defines the minimum requirements of QoS measurement equipment for GSM and 3G networks in the way that the values and trigger-points needed to compute the QoS parameter as defined in part 2 can be measured following the procedures defined in part 3. Test-equipment fulfilling the specified minimum requirements, will allow to perform the proposed measurements in a reliable and reproducible way.

Part 5 specifies test profiles which are required to enable benchmarking of different GSM or 3G networks both within and outside national boundaries. It is necessary to have these profiles so that when a specific set of tests are carried out then customers are comparing "like for like" performance.

Part 6 describes procedures to be used for statistical calculations in the field of QoS measurement of GSM and 3G networks using probing systems.

Part 7 describes the methodology to be used to get statistical samples that could be representative of the population of all existent calls so the statistical inferences made from the data collected by test calls are statistically correct and the error margin could be calculated.

Introduction

All the defined quality of service parameters and their computations are based on field measurements. That indicates that the measurements were made from customers point of view (full End-to-end perspective, taking into account the needs of testing).

It is assumed that the end customer can handle his mobile and the services he wants to use (operability is not evaluated at this time). For the purpose of measurement it is assumed that:

- the service is available and not barred for any reason;

- routing is defined correctly without errors;

- the target subscriber equipment is ready to answer the call;

- voice quality values measured should only be employed by calls ended successfully for statistical analysis;

- need to define similar for video calls;

- however, measured values from calls ended unsuccessfully (e.g. dropped) should be available for additional evaluations and therefore, must be stored;

- further preconditions may apply when reasonable.

1 Scope

The present document describes typical procedures used for QoS measurements on mobile communication networks, e.g. GSM or UMTS, along with settings and parameters for such measurements.

Where possible existing ITU-T or ETSI definition are referenced. In some cases ITU-T or ETSI definitions do not exist or are considered as too generic, then a more service and mobile network specific definition is chosen.

2 References

References are either specific (identified by date of publication and/or edition number or version number) or non-specific.

- For a specific reference, subsequent revisions do not apply.

- Non-specific reference may be made only to a complete document or a part thereof and only in the following cases:

  - if it is accepted that it will be possible to use all future changes of the referenced document for the purposes of the referring document;

  - for informative references.

Referenced documents which are not found to be publicly available in the expected location might be found at http://docbox.etsi.org/Reference.

For online referenced documents, information sufficient to identify and locate the source shall be provided. Preferably, the primary source of the referenced document should be cited, in order to ensure traceability. Furthermore, the reference should, as far as possible, remain valid for the expected life of the document. The reference shall include the method of access to the referenced document and the full network address, with the same punctuation and use of upper case and lower case letters.

NOTE: While any hyperlinks included in this clause were valid at the time of publication ETSI cannot guarantee their long term validity.

2.1 Normative references

The following referenced documents are indispensable for the application of the present document. For dated references, only the edition cited applies. For non-specific references, the latest edition of the referenced document (including any amendments) applies.

[1] ITU-T Recommendation P.56: "Objective measurement of active speech level".

[2] ETSI TS 102 250-1: "Speech Processing, Transmission and Quality Aspects (STQ); QoS aspects for popular services in GSM and 3G networks; Part 1: Identification of Quality of Service criteria".

[3] ETSI TS 102 250-2: "Speech Processing, Transmission and Quality Aspects (STQ); QoS aspects for popular services in GSM and 3G networks; Part 2: Definition of Quality of Service parameters and their computation".

2.2 Informative references

The following referenced documents are not essential to the use of the present document but they assist the user with regard to a particular subject area. For non-specific references, the latest version of the referenced document (including any amendments) applies.

[i.1] IETF RFC 3501: "Internet message access protocol - version 4rev1".

[i.2] IETF RFC 2177: "IMAP4 IDLE command".

[i.3]   IETF RFC 2821: "Simple Mail Transfer Protocol".

3   Definitions and abbreviations

3.1   Definitions

For the purposes of the present document, the following terms and definitions apply:

A party: in direct transactions, the party initiating the transaction (calling party)

> NOTE: In store-and-forward transactions, the party sending content.

B party: in direct transactions, the termination or counterpart of a transaction

> NOTE: In store-and-forward transactions, the party receiving content.

content: entirety of information transferred within a transaction, seen from the user's perspective

> NOTE: In case of services requiring entrance procedures (e.g. server login with FTP), information flow to achieve the state of being able to transfer actual user data is not counted as content.

> EXAMPLE: Single SMS in SMS service; single multimedia message consisting of video, audio, and text components in MMS service.

direct transaction: real-time transaction between two entities

maximum expected delivery time: for store-and-forward services, this defines the time span within which a message shall be received by the B party to rate the transaction successful from the user's perspective

service family: group of services having main characteristics in common

> EXAMPLE: Speech and Video Telephony, as well as SMS and MMS, are assumed to form a service family.

store-and forward transaction: transaction where information is sent from one party A to another party B using an entity C to store information sent from A and attempting to deliver it to B

transaction: single, complete, typical usage of a particular service

> NOTE 1:   At the beginning of each clause describing a particular service or family of services, the typical transaction for this particular service is described.

> NOTE 2:   Each type of transaction has parameters. The sum of all parameters describes the transaction completely. A parameter set is assumed to be complete if, under constant outer conditions, all transactions using this parameter set provide the same result.

transaction result: set (list) of possible outcomes for a particular transaction

> NOTE: Services belonging to the same service family share the same set of transaction results.

3.2   Abbreviations

For the purposes of the present document, the following abbreviations apply:

AOV	Angles Of View
APN	Access Point Name
CS	Circuit Switched
CSD	Circuit Switched Data
DNS	Domain Name Server
FTP	File Transfer Protocol
GGSN	Gateway GPRS Support Node
GPRS	General Packet Radio Service
HTTP	Hyper Text Transfer Protocol

IP	Internet Protocol
KPI	Key Performance Indicator
MEDT	Maximum Expected Delivery Time
MMS	Multimedia Messaging Service
MO	Mobile Originated
MOF	Mobile Originated to Fixed
MOM	Mobile Originated to Mobile
MOS	Mean Opinion Score
MS	Mobile station
MT	Mobile Terminated
MTM	Mobile Terminated, originator is also a Mobile unit
MTU	Maximum Transmission Unit
PC	Personal Computer
PDU	Packet Data Unit
PS	Packet Switched
PSD	Packet Switched Data
QoS	Quality of Service
SMS	Short Message Service
SMSC	Short Message Service Centre
TCP/IP	Transmission Control Protocol/Internet Protocol
UDP	User Datagram Protocol
UE	User Equipment (terminal, mobile station)
URL	Uniform Resource Locator
WAP	Wireless Application Protocol

4 Goal of measurement

The goal of measurements described in the present document is to assess the network under test for its quality parameters as defined in TS 102 250-1 [2]. This is, to determine the network quality for the respective transactions from subscribers view.

5 Classification of services

5.1 Classification guidelines

For the purpose of the present document, services are classified using what is considered to be their dominating property. The first distinction is made between direct and store-and-forward services:

- Direct-transaction services are services where there is - in the user's perception - a direct end-to-end connection.

- Store-and-forward services are services where content is stored in the network and delivered to the recipient to a later point in time.

As a technically usable distinguishing question, a service is considered to be direct if it is possible to decide on end-to-end content transfer success from the initiating party (A party) of the connection within the scope of the transaction itself.

For a second classification within direct services, the typical nature of content flow is used for classification:

- In telephony services, content flow is typically symmetrical between the parties.

- In data-access services, content flow is typically - within a single transaction - unsymmetrical.

Of course there are borderline cases, such as voice mailbox services, voice over IP or messaging. Such services are, for the time being, not treated within the scope of the present document.

5.2   General structure of service descriptions

In the following, each service family description will contain the following structural elements:

- A "family general" part defining:

    - the basic transaction definition and if applicable, transaction types;

    - a description of the transaction phase combined with a table of parameters governing transaction behaviour in this phase;

    - a description of all possible outcomes of a single transaction;

    - a description of content quality measurement definitions (if applicable).

- For each service within the service family, if its respective description differs from the "family general" one, a set of descriptions for this service, having the same structure as above.

6   General aspects for all types of services

6.1   Set-up and control

Measurements should be conducted in a way that user behaviour is emulated, with a number of parameters under control of the measurement equipment.

The test case design (configuration and user profile) - to the degree necessary to fully reproduce the test - shall be part of the report.

It is assumed that for all types of services under test, a testcase consists of a number of single identical transactions. The measurement equipment and control must ensure that the starting conditions are the same for each transaction. This includes, among other things, that pause times are sufficiently long that the equipment is in a stable (idle) state again. The parameter "guard time" sets a minimum value for the pause between transactions.

It is assumed that the measurement is performed by a mobile measurement unit; depending on type of measurement, other equipment connected to the fixed network is added.

It is assumed that all QoS-relevant transaction parameters are recorded for proper post-processing and are kept constant during a test. If a test contains more than one parameter set, evaluation shall be made for each parameter set separately.

6.2   Phase and result classification

In order to ensure common wording, the following clause defines terms and definitions for services testing.

It is assumed that each transaction can be described at least by one seamless sequence of phases. There may exist several Angles Of View (AOV), each leading to a different phase description.

    EXAMPLE: Internet services (as described by KPI defined in TS 102 250-2 [3] model A and B). AOV differ here by different assumptions on start of service usage. Each AOV, however, is a consistent description by seamlessly connected phases.

Phases maybe further be described having sub-phases.

Pauses between transactions are not explicitly mentioned in this picture, but are relevant with respect to parameter reporting. Typically, there is a minimum pause (guard time) ensuring that the system under test is in a stable starting condition for the next test. Values are technology-dependent.

6.2.1   Phase and result classification for direct services

A direct transaction consists of two top-level phases: Service access and service usage.

Table 1

Phase	Sub-phase	Definition
Service access		All steps leading to the technical ability to do actual user-perspective content transport between A and B party. Service access may consist of different sub-phases, e.g. Network access, IP service access and Internet access. Which sub-phases actually exist depends on the particular service.
	Network access	Basic access to the network under test. Successful network access is assumed when the UE is able to do as much basic communication with the network as is necessary to initiate the next phase in the service access procedure.
	IP service access	Basic access to the generic packet-data transfer capabilities the particular service is based upon.
	Internet access	Basic access to those internet services the service is meant to provide.
Service usage		Content transfer between A and B party.

A direct transaction may have one of the following overall results.

Table 2

Result	Definition
Failed	Phase of service usage not reached. Successful or failed service access may be broken down into diagnostic sub-categories. The general name-forming rule is: <name of sub-phase>result. Example: Network access failure; IP service access success.
Completed	Data-transfer transactions: All content intended to be transferred has been successfully transferred. Conversational transactions: The intended transaction duration has been reached.
Dropped	Service usage was ended before completion.
NOTE:	If a transaction being in the service usage phase is stopped due to some timeout or due to other criteria by the measurement system, e.g. to enhance test rate, this shall be treated as a dropped transaction. This behaviour has to be recorded by the measurement system.

6.2.2 Phase and result classification for store-and-forward services

A S&F transaction consists of two top-level phases: Content sending and content delivery.

Table 3

Phase	Sub-phase	Definition
Content sending		All steps required to move the content to the network, up to the point where the network is able to start delivery. This phase is completed when there is nothing more the A party can/needs to do to move content to the B party. It is assumed that the A party gets information sufficient to judge if sending has been successful or not.
Content delivery		All steps to transfer the content to the B party. Delivery may consist of two main sub-phases: **notification** and **retrieval**.
	Notification	Information to the B party that content is ready for transfer.
	Retrieval	Transport of content from network to B party, initiated by the B party.
NOTE: Technically, existing S&F transaction types both SMS and MMS have these phases; however, notification in SMS is integral part of L3 protocol while in MMS two different media (SMS and WAP) are used.		

A S&F transaction may have one of the following overall results.

Table 4

Result	Sub-category	Definition
Completed		Content was successfully transferred from A to B party.
Failed		Content was not successfully transferred from A to B part. Depending on particular services and available in formation, there may be a number of possible sub-categories for this result.
	Undelivered	Content was successfully sent to the network, but was never (or: not within a given period of time, see note) delivered.
	Send failed	Content was not successfully delivered to the network.
	Lost	Content was successfully sent to the network, but notification was never received by A party. Diagnostic sub-category in case notification can/shall be technically identified within the delivery process.
NOTE: If a Transaction was completed, but content delivery time was above a given threshold, it shall be treated as Failed, sub-category Undelivered.		

7 Telephony measurements

This clause deals with telephony services. In general, the term "content" will be used throughout this clause for the information flow exchanged between subscribers during a call. Depending on the type of service, content can be audio or audio and video.

7.1 General aspects

7.1.1 Transaction definition and transaction types

The basic transaction for telephony testing is equivalent to a single call to a counterpart extension. It is assumed that the call partner is typically a fixed-network type extension to avoid uncertainties related to a second mobile connection.

Type is either Mobile Originated (MO) or Mobile Terminated (MT).

It is assumed that once a connection has been established, for further measurements it does not matter which side has triggered it. Therefore, the audio data flow parameter will not be logically linked to the call type.

7.1.2   Parameter overview

Table 5

Phase	Parameters
Service Access	Call counterpart. This includes the type of equipment (dedicated unit, automatic reply with taped message, etc.).
	Call type
	Time-out for failed-call condition
Service Usage	Call duration
	Content flow direction: This is an inner parameter for a transaction. Basically, all combinations of uplink/downlink dynamics are possible: - Uplink only; - Downlink only; - Conversational (alternating uplink and downlink). This is the recommended standard testing mode. Other testing modes are considered to be used only for special purposes; - "Duplex" (uplink and downlink flow simultaneously).
	Codec settings
Call clear-down	Guard time
Pause	Pause duration. For current GSM equipment, a pause time of at least 15 seconds (guard time) is recommended. However, this duration may be adjusted to local conditions or special testing goals, but this must be reported. If the pause duration is too short, side effects may occur, resulting in all kinds of transient effects and distortions in measurement data. It should be made certain that all the QoS parameters to be measured are not affected by the pause time.

The last transaction within a test sequence does not require a pause.

7.1.2.1   Additional considerations

Regarding call timing including behaviour in case of dropped calls. For reasons of comparability in density, it is required that all call attempts which form part of the statistics must follow a constant time pattern. Additional call attempts must not affect the pattern of the statistically relevant call attempts (see figure 1).
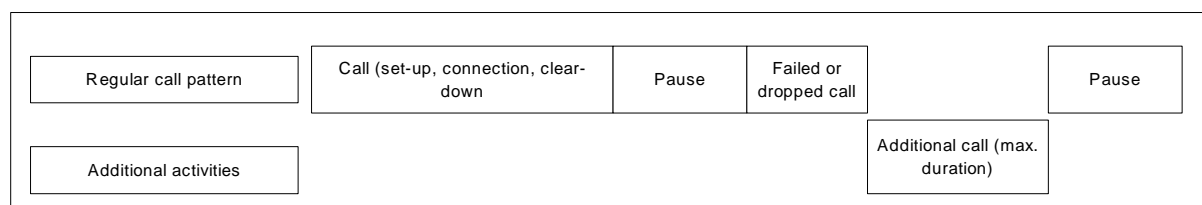
Figure 1

7.1.3   Additional transaction result definitions

For call set-up assessment beyond QoS data acquisition, typically a state model driven by Layer 3 information combined with information from the call control engine is being used. This state model may also be used to determine timing information for each phase.

Service usability, i.e. presence of a usable two-way connection, shall be verified by a procedure based on content test transmissions within a given time window. If within this time window no connection can be verified, the setup attempt shall be considered to be failed and the call attempt be terminated.

A call is active only as long as both sides consider it to be active. A call is therefore considered to be dropped if either side detects a dropped call.

Above definitions lead to the following decision tree for the outcome of a call (figure 2 includes the end-of-call cases).
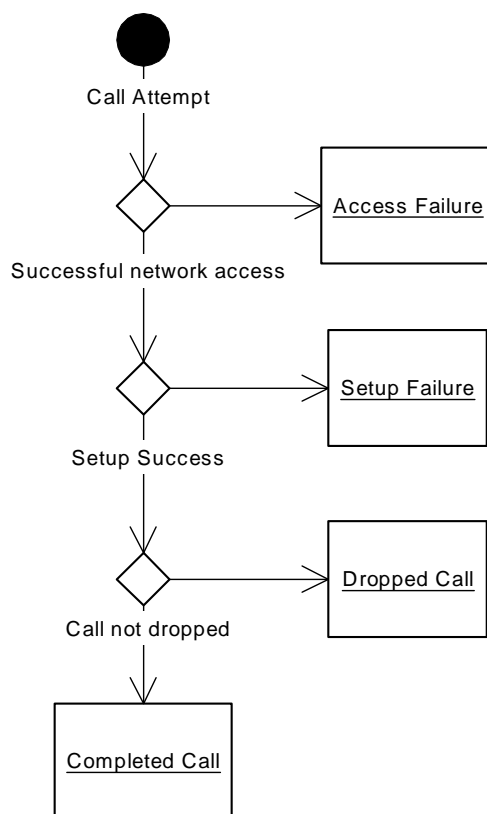
Figure 2

7.1.4   Content quality

For content quality assessment, data is generated at the receiving end. For downlink content, data storage is therefore straightforward; quality-assessment data is simply included with other data items making up the result file. For uplink content, at some point in time results have to be integrated with the rest of the data.

For assessing content quality of complete transmitted speech samples, at least the following methods are possible:

- Real-time assessment (streaming mode), where the speech quality assessment algorithm continuously outputs MOS_LQO data items.

- "Offline" assessment, where content is first recorded in some way and later being processed.

Data processing must make sure that only such content quality data is used which lies inside the "connection active" time window and is in line with one of the two different content quality parameters.

7.1.5   Verification of usable two-way connection

Only calls with a valid two-way end-to-end information connection shall be considered for content quality assessment (valid calls).

A non-valid call shall be treated like a dropped call, with a modifier indicating this particular cause.

7.2   Speech telephony

7.2.1   Transaction definition and transaction types for speech telephony

See family definitions.

7.2.2   Parameter overview for speech telephony

None. See family definitions.

7.2.3   Additional transaction results for speech telephony

None. See family definitions.

7.2.4   Content quality for speech telephony

None. See family definitions.

7.2.5   Verification of usable two-way connection for speech telephony

This shall be verified by a procedure based on audio test transmissions within a given time window. If within this time window no audio connection can be verified, the setup attempt shall be considered to be failed and the call attempt be terminated.

> NOTE: To make sure an audio connection is valid, it is assumed that an appropriate kind of data analysis on audio flow is performed (see ITU-T Recommendation P.56 [1]) Video Telephony.

7.2.6   Void

7.2.7   Transaction definition and transaction types for video telephony

The basic transaction for video telephony testing is equivalent to a single call to a counterpart extension.

Due to existing usage and hardware, typical video calls will be between UE, so no further call types are distinguished.

Unlike other services, it is currently assumed that video-call testing will require a high degree of abstraction in the sense that testing-system architectures may differ quite thoroughly from those found in "real connections". At the time of writing, there are no known terminal-based protocol stacks delivering data necessary for testing or QoS assessment. Therefore, it is assumed that PC-based implementations of video protocol stacks will have to be used, where the UE serves only as the modem part of the connection.

Currently it is assumed that QoS-testing architecture will use a mobile-to-fixed CSD connection, with video telephony protocol being implemented on PC on both sides.

It is further assumed that typical tests will use data which serve as load and carry diagnostic data at the same time.

7.2.8   Parameter overview for video telephony

> NOTE: Content flow will typically be governed by the video protocol and is assumed to involve simultaneous data flow in both directions.

Table 6

Phase	Parameters
Service access	Connection type (CSD bearer type).
	Video-telephony protocol used.
	Call counterpart technology and architecture. This includes the type of test being made (e.g. "transparent" test of modem connection.).
	Time-out for failed-call condition.
Service usage	Call duration.
Call clear-down	Guard time.
Pause	Pause duration. For current GSM equipment, a pause time of at least 15 seconds (guard time) is recommended. However, this duration may be adjusted to local conditions or special testing goals, but this must be reported. If the pause duration is too short, side effects may occur, resulting in all kinds of transient effects and distortions in measurement data. It should be made certain that all the QoS parameters to be measured are not affected by the pause time.

The last transaction within a test sequence does not require a pause.

7.2.9   Additional transaction result definitions for video telephony

It is assumed that the A and B party video-telephony protocol stacks are compatible, and given a usable end-to-end connection of proper type, a video-telephony connection can be established. It is assumed that if the network can not provide a connection with the required bearer properties, this will either result in failure to provide a usable end-to-end connection, or the video-telephony protocol stack will detect this by indicating connection failure.

It is conceivable that the network provides the wrong bearer type. Due to limited "decision power" determining the usability of an end-to-end connection, this could go undetected at first, but lead to quality deficiencies due to inappropriate bearer properties.

7.2.10   Content quality for video telephony

Currently no stable specification for video MOS exists. It is assumed that the audio part of connection will be tested essentially the same way as for speech telephony.

7.2.11   Verification of usable two-way connection for video telephony

It is assumed that the video telephony protocol being used implicitly checks for useful connection.

8   Store-and-forward services measurement

8.1   General aspects

The basic transaction for store-and-forward services testing is equivalent to transmission of a single unit of content (which may, however, contain several parts such as audio, video and text) between two parties.

Typically, content is, in the end-to-end perspective, being sent either from some fixed-network source to a UE (mobile terminated, MT) or from one UE to another (mobile originated to mobile, MOM). However, MOM testing with a two-UE setup may cause additional uncertainty due to coverage and signalling effects in the destination UE.

Basically, there are two recommended methods of testing:

a)   using a destination UE in a fixed location with virtually 100 % transfer probability. Part of the measurement procedure should be cyclical reference checks to assure reception quality respectively to create baseline data; or

b)   using a fixed-network destination such as a Large Account server which is accessed by suitable means from a fixed-network location.

In all cases, testing methodology requires that measurement data are being collected in at least two locations and need to be integrated before the final QoS evaluation can be made.

For practical reasons, measurements are typically limited to a certain time window. Technically it is possible that content being sent in previous test session arrive at the system after the session has been finalized. Technical provisions are needed to ensure proper handling of such content, either by ignoring it or by correctly assigning it in post-processing.

Content used by the measurement system shall carry appropriate signature to identify it as belonging to the measurement, to identify the test session within it was sent, and to enable unambiguous correlation between sent and received content.

8.1.1   Transaction phase and parameter overview

Table 7

Phase	Parameters
Sending	Destination equipment type.
Notification	Time-out for notification.
Retrieval	
Pause	Pause duration. For current GSM equipment, a pause time of at least 15 seconds (guard time) is recommended.

8.1.2   Additional transaction result definitions

The nature of store and forward services implies that loss rates need precise definition. From the user's point of view, there is expectation that content is delivered within the MEDT. Even if technically an outstanding content item may be delivered later, it should be considered lost anyway. This leads, in turn, to the problem of over-aged content items which may disturb subsequent measurements. Suitable means of eliminating such effects should be taken, e.g. if the service allows for it, setting a reasonable lifetime for the content item, and using unique sequence numbering in a way to prevent aforementioned effects.

Post-processing must ensure that, if the measurement system is shut down before a given maximum waiting time has expired, all pending units of content, i.e. those with an age less than MEDT are completely removed from QoS evaluation since there is no way to predict if these elements will arrive within MEDT or not.

The influence of over-aged content should be reduced to a minimum by using a selective retrieval. Only the content delivered within the MEDT should be downloaded when possible.

8.1.3   Content quality for Store-and-Forward Services

8.1.3.1   SMS

It is assumed that SMS services protect their payload sufficiently so it can be assumed that no distortions in content will occur.

8.1.3.2   MMS

Due to the fact that transcoding is integral part of MMS which is an explicit change of content, no direct comparison of content sent an received is possible. Rather, validation shall be done based on presence of objects as a whole. See further definitions in later clauses.

8.1.3.3   e-Mail

Reference content enabling validation of transmitted vs. received data should be used.

The performance of all parameters except the access parameters depends on the e-mail content concerning the results. Therefore it is only allowed to aggregate test results which used the same e-mail content setup, except the access parameters.

A proposal on reference content is currently in the works.

8.2   SMS measurements

8.2.1   General aspects of SMS measurements

8.2.2   Transaction definitions and transaction types for SMS

The basic transaction for SMS testing is equivalent to transmission of a single SMS.

Typically, user SMS are being sent either from some fixed-network source to a mobile (SMS mobile terminated, SMS-MT), or from one mobile to another (SMS mobile originated, SMS MO, with regard to mobile-based testing). However, uplink SMS testing with a two-mobile setup may cause additional uncertainty due to coverage and signalling effects in the destination mobile. Basically, there are two recommended methods of testing:

- using a destination mobile in a fixed location with virtually 100 % transfer probability. Part of the measurement procedure should be cyclical reference checks to assure reception quality respectively to create baseline data; or

- using a fixed-network destination such as a Large Account SMS server which is accessed by suitable means from a fixed-network location.

In all cases, testing methodology requires that measurement data is being collected in at least two locations and needs to be integrated before the final QoS evaluation can be made.

There are two modes for SMS transmission:

- text mode;

- PDU mode.

Basically, these modes should be equivalent including options available. In practice, mode support are mobile-dependent.

For each SMS, a "lifetime" can be set after which a SMS is deleted in the SMSC. While in PDU mode, this parameter is part of the parameter structure, in text mode it is set in the mobile, which may not be supported by all mobile types. This leads to the recommendation that PDU mode should be used.

To ensure comparability and statistical validity of transactions, the following outer conditions need to be constant throughout a test case:

- SMS mode.

- Call counterpart.

- SMSC being used.

- Call timing including behaviour in case of delivery failure to the network.

For reasons of comparability in density, it is required that all SMS sending attempts which form part of the statistics must follow a constant time pattern. Additional SMS sending attempts e.g. in the case of failure of delivery to the network must not affect the pattern of the statistically relevant call attempts.

SMS being sent from either side should contain information enabling validity and integrity checking during evaluation. The recommended minimum set includes:

- signature assuring that only test system SMS are being considered;

- sequence number for SMS delivery timing.

If required, a test SMS may contain additional data (padding) to obtain SMS with a length assumed to be typical for a particular network.

If required, more powerful means of ensuring that only SMS generated by the test system can be added. For example, the SMS may contain a code which is created from checksum content and a seed value delivered by the receiving side. This ensures that even if a static signature is duplicated by operational errors or malfunction of other test systems, the system will be robust against it.

Further for general design of SMS testing, communication between the mobile testing system and its stationary counterpart is required to ensure certain starting conditions:

- Prior to actual testing, the SMS storage of the MS should be cleared to exclude SMS delivery failure due to memory shortage.

- The stationary side should start sending SMS only after a "go" from the mobile side, ensuring that no transient effects occur.

8.2.3   Testing mode for SMS-MT

For this type of testing, SMS are being generated at a constant rate by the stationary side.

All SMS received which have not been generated by the stationary side shall be ignored for quality assessment.

For reception behaviour, two modes exist; it is mobile-dependent which modes are supported.

a)   On SMS reception, the mobile generates an indication message (CMTI message) on its data connection to the PC. The actual SMS can then be requested by a command.

b)   The mobile outputs the SMS directly on its data connection to the PC.

From control and trigger point accuracy considerations, mode a) is preferred.

8.2.4   Testing mode for SMS-MO

For this type of testing, SMS are being generated at a constant rate by the mobile side, plus eventual extra SMS in case of failure of delivery to the network.

8.2.5   Transaction phase and parameter overview for SMS

Table 8

Phase	Parameters
Sending	Type and Destination equipment type (depends on type, e.g. To Mobile, to Large Account, etc.)
	SMSC used
	Content size
Notification	Time-out for notification
	Notification mode for incoming SMS
Retrieval	
Pause	Pause duration. For current GSM equipment, a pause time of at least 15 seconds (guard time) is recommended

8.2.6   Possible transaction results for SMS

None. See family definitions.

8.2.7 Content quality for SMS

Not defined for SMS.

8.3 MMS

8.3.1 General aspects of MMS measurements

With respect to general aspects of store-and-forward systems, in MMS testing special care has to be taken for proper handling of incoming notification SMS not belonging to the active session. They shall be treated as follows:

- Incoming SMS not being a valid push SMS shall be ignored.

- Incoming valid push SMS identified as been sent by the testing system but not belonging to the current test session may be ignored (see general remarks).

- For all other push SMS, it is optional to download the MMS and to check if it belongs to the measurement.

NOTE: If the network under test may modify push SMS content in a way that it cannot be detected if the push SMS belongs to the measurement, ignoring the push SMS contains the risk that a MMS successful from the user's perspective is reported as unsuccessful by the measurement system.

8.3.2 Transaction definitions and transaction types for MMS

The basic transaction for MMS testing is equivalent to transmission of a single MMS, which may however contain several parts such as picture or video, audio and text.

Typically, user MMS are being sent either from a mobile to another mobile (MOM or MTM depending on measurement setup), or from a mobile to a fixed-network location (MOF). In the case of MMS the destination is typically one or more e-mail addresses.

In case a destination mobile is classified by the network as not capable of MMS reception, a "legacy SMS" is transmitted instead. This SMS will contain an information that and how MMS content can be downloaded via internet. The format of such legacy SMS is not standardized.

In mobile-to-mobile scenarios there is always the risk of ambiguity with respect to end to end quality assessment. In the case of store and forward measurements, this is less critical due to the definition of QoS parameters which allow to distinguish uplink and downlink components. Nevertheless, QoS parameters expression end-to-end performance will be affected. Therefore, two methods of testing are recommended:

a) using a B party mobile in a fixed location with very good network coverage. Part of the measurement procedure should be appropriate cyclical reference checks to obtain baseline data; or

b) using an e-mail destination.

To ensure comparability and statistical validity of transactions, the following outer conditions need to be constant throughout a test case:

- MMS content;

- call counterpart;

- MMS Proxy being used;

- send timing including behaviour in case of delivery failure to the network.

For reasons of comparability in density, it is required that all MMS sending attempts which form part of the statistics must follow a constant time pattern. Additional MMS sending attempts e.g. in the case of failure of delivery to the network must not affect the pattern of the statistically relevant call attempts.

MMS being sent from either side should contain information enabling validity and integrity checking during evaluation. The recommended minimum set includes:

- signature assuring that only test system MMS are being considered;

- sequence number for MMS delivery timing.

If required, a test MMS may contain additional data (padding) to obtain content with structure and size assumed to be typical for QoS measurement purposes.

Further for general design of MMS testing, communication between the mobile testing system and its stationary counterpart is required to ensure that prior to actual testing, the MMS and SMS storage of the MS should be cleared to exclude failures due to memory shortage or other limitation.

8.3.3 Transaction phase and parameter overview for MMS

Unlike SMS, posting and retrieving MMS is handled via a network service (WAP) rather than embedded in low-level functionality of the network. Also, uploading or downloading MMS is a process taking considerable time. Therefore, the simple phase scheme - and, consequently, the possible-result definitions, have been extended by inserting subphases.

Table 9

Phase	Subphase	Parameters
Sending	Network access	Equipment Types and Capabilities
	MMS gateway access	MMS gateway address (Proxy used)
	Upload	Content composition Content size User Agent String (identification of receiving UE)
Notification		Time-out for notification
Retrieval	Network access	Equipment Type and Capabilities
	MMS gateway access	MMS gateway address (Proxy used)
	Download	
Pause		Time between MMS postings

8.3.4 Additional transaction result definitions for MMS

The following definitions are based on the assumption that, from technical feedback during his service usage, the user notices (and cares about) different ways a MMS transaction can fail. The following definitions are to be seen in conjunction with family result definitions:

- Service access failed: The MMS transaction could not be started due to inability to access the network or the service.

- Dropped while posting: Content could not be uploaded successfully.

- Lost: The content package was successfully uploaded, but the recipient did not get notification within the MEDT.

- Dropped while retrieving: The recipient received a notification, but was unable to retrieve the MMS.

  NOTE: Unlike for upload, this includes network or service access failure.

- Completed: Content has been successfully delivered to the B-party (which includes completeness of the package).

8.3.5 Content quality for MMS

In MMS transmission, transcoding may be used by the MMS infrastructure to adapt content from the format of the sending UE to a format supported by the receiving UE.

For the time being, it is assumed that correct transcoding is validated by other means than QoS measurements. Therefore, it is not subject of the present document.

8.4   E-Mail

8.4.1   Transaction definitions and transaction types for E-Mail

E-Mails can be send from an e-mail client to another e-mail client. Both clients can independent use a wireless or wired connection to the server. E-Mails can be small and contain only text (plain text, rich text or html coded) without attachments or they can have big sizes with several pages of text body. The attachment content can be documents (pdf), pictures (jpg), software, music, ZIP files, etc.

8.4.1.1   Standardized E-Mail Content

The following three standardized e-mail setups shall be used to allow an objective benchmark. The performance of all parameters except the access parameters depends on the e-mail content concerning the results. Therefore it is only allowed to aggregate test results which used the same e-mail content setup, except the access parameters.

   NOTE: A standardized content is under study.

8.4.2   Transaction scenarios

Two different transaction scenarios are possible to test the service from end customer's perspective.

The first one examines the complete service usage from the sender side (e-mail upload) until the receiver side (e-mail download). The second scenario only downloads e-mails. The test shall support at least one scenario. Both scenarios have advantages and disadvantages – see clauses 8.4.2.1 End-to-End scenario and 8.4.2.2 Download scenario.

8.4.2.1   End-to-End scenario

This scenario should be used to get performance results for the complete service usage. This includes a test procedure for e-mail upload, e-mail notification and e-mail download.

It is necessary to notice the dependency of the notification and download parameters to the successful upload before. It is only possible to notify and download e-mails if the upload was successful before. This means that we only get performance data for the notification and download parameters in phases of successful e-mail uploads. If the e-mail server or the mobile network has a malfunction we would not get results for e-mail notification and download that causes in end results which are better than they should be. These failures are only seen within the results of the upload parameters.

The recommendation is to calculate only upload and end-to-end parameters in this case and to use the "Download scenario" additionally to get results for the Download-Parameters.

The test e-mail server where the accounts are hosted should be the same for up- and downloads to prevent side effects by the internet.

This End-to-End scenario can be shortened by only using the first upload procedure. The tests will only then generate results for the upload parameters. The "Download scenario" can expand also this upload procedure.

A test cycle proposal contains the clause 8.4.9.
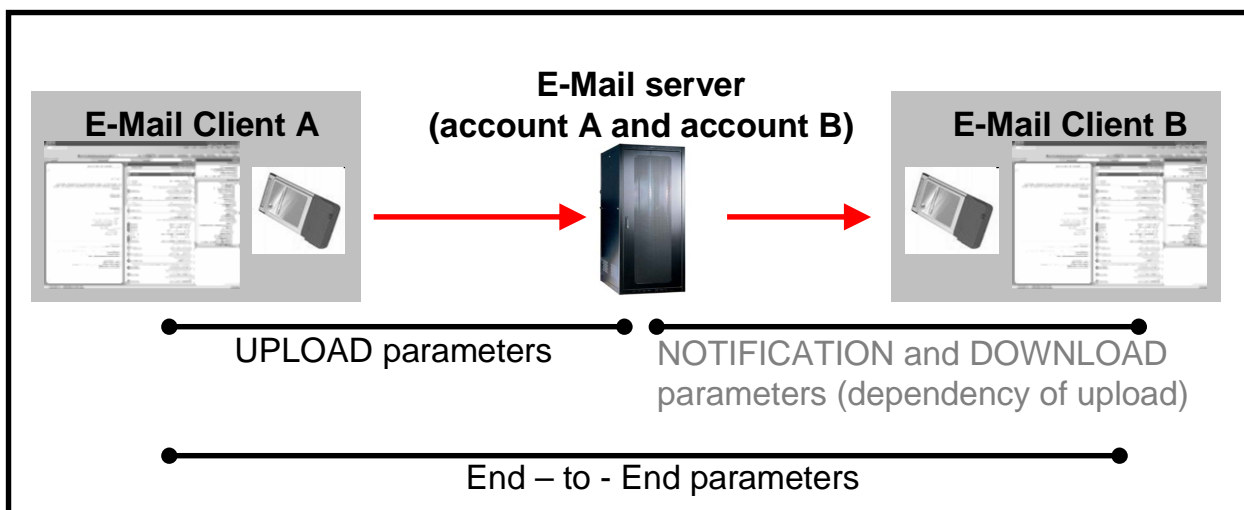
Figure 3

8.4.2.2 Download scenario

This scenario should be used to get performance results for the notification and download parameters. It can be seen as an extension to the "End-to-End scenario" because it is possible here to get valid notification and download parameter results instead of the "End-to-End scenario" because the upload of e-mails is here independent from the used communication technology by the e-mail download.

The tested e-mail server needs therefore a second communication interface to control the parameter precondition which means a new e-mail needs to be placed in the tested inbox. This second communication channel shall be independent from the tested technology and shall be very reliable.

A test cycle proposal contains the clause "8.4.10 Test sequence for Download scenario".
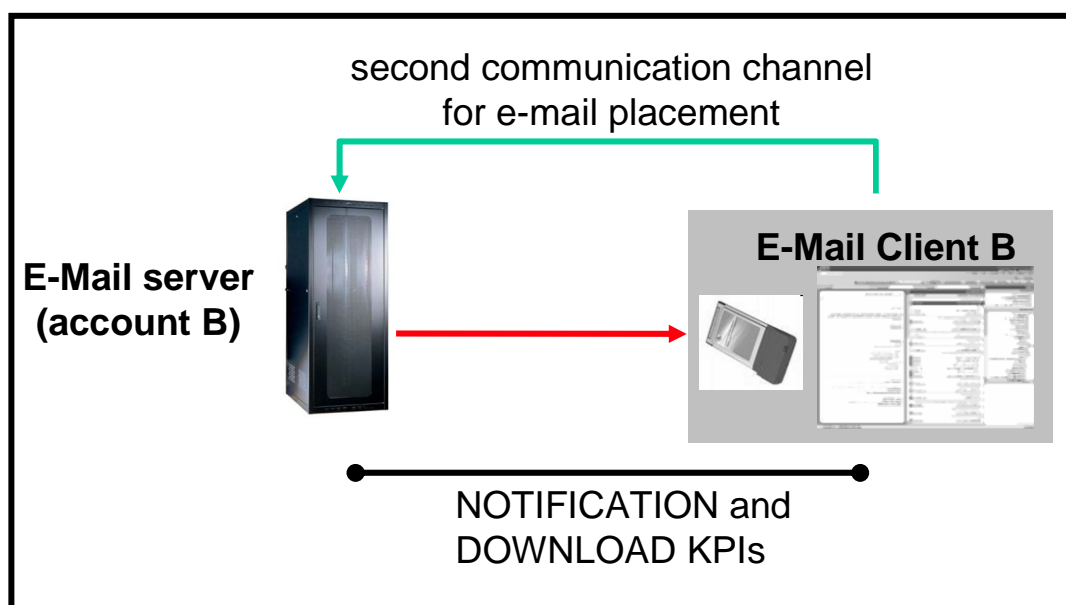
Figure 4

8.4.3 Protocols

The test shall support the following protocols:

- For Download: IMAP4 (Internet Message Access Protocol, version 4) with "Idle" (push functionality), RFC 3501 [i.1] and RFC 2177 [i.2];

- For Upload: SMTP (Simple Mail Transfer Protocol) RFC 2821 [i.3].

NOTE: The popular POP3 protocol does not follow the "Store-and-Forward" approach because it is not capable to push which means "Forward" e-mails to the receiver side. By using the POP3 protocol the receiver has to poll the e-mail server to transmit new e-mails which cause further delays – see also clause 8.4.4.

8.4.4   Push function

To get meaningful results for the End-to-End parameters it is mandatory to use the push function of the e-mail protocols (e.g. IMAP idle function). Otherwise the test results would not reflect the real performance in respect to the end-to-end time parameters.

Take into account that e-mails can be transmitted compressed or uncompressed which can influence the test results, especially the transmission times. The choice of small text length and compressed attachment types can minimize these effects.

8.4.5   Support for header download

In respect to the functional range of popular e-mail clients the measurement should support to download the header information of new e-mails with the following different settings:

1) Complete download of all e-mails.

2) At first download of header and then download of all e-mail data (body + attachment).

3) Only download of header.

If the e-mail inbox has more than one new e-mail waiting for download after the login, the header download should be done for all new e-mails as one step to reflect the end customer's perspective. After that the e-mail can be downloaded one after another.

To differentiate between new content within MEDT and over-aged content or junk e-mails the header information downloaded should contain at least the following content:

- To.

- From.

- Subject.

Test results which are based on over-aged and junk e-mails shall be excluded in post processing.

8.4.6   Timeout

Every used parameters shall have an own configurable timeout setting. This applies also to parameters which overlaps with others concerning the trigger points - accessing the IP-Service (e.g. 7.12.4 4 E-Mail Client Upload IP-Service Access Failure Ratio {Service} Message Upload Access Failure Ratio), transferring the data (e.g. 7.12.6 9 {Service} E-Mail Client Upload Data Transfer Cut-off RatioMessage Upload Data Transfer Cut-off Ratio) and the combination of the access and transmission (e.g. 7.12.2 6 E-Mail Client Upload Session Failure Ratio{Service} Message Upload Failure Ratio). The timeout settings depend on the e-mail sizes and the used technology.

8.4.7   Further mandatory requirements

Each e-mail uploaded shall have a unique identifier inside the subject and e-mail body text to identify each e-mail from the other e-mails. This unique identifier shall be used to find test results with the same setup and e-mail size to aggregate them together.

Each e-mail uploaded shall have an atomic time stamp (time of upload start) inside the subject and the e-mail body text. Sender and receiver need a time synchronization if they use different systems.

If the upload and download part use different locations they need a synchronization link for the parameter calculation. This is only important if a complete End-to-End scenario is tested.

8.4.7.1   Guard times

For a valid End-to-End time calculation it shall be used one of the following procedures to prevent that other e-mails will influence the download time performance:

1)   Only one e-mail is sent at once per test cycle (one cycle means one connection to a server). The guard time between each cycle shall be equal or higher than the sum of the timeout settings of each single parameter. This means the guard time is the same as the timeout for the "7.12.24 17 { E-Mail Client to Client End-to-End Failure RatioService} End-to-End Failure Ratio" parameter.

2)   If several e-mails are uploaded at once within one connection, they need a guard time to prevent download influences between the different e-mails. The guard time shall be equal or higher than the sum of the timeout settings of each single parameter. This means the guard time is the same as the timeout for the "7.2.17 { E-Mail Client to Client End-to-End Failure Ratio7.1.24 {Service} End-to-End Failure Ratio" parameter.

3)   If several e-mails are uploaded at once, they shall be sent to different e-mail accounts – one account per e-mail. Each receiver e-mail account needs to be connected to a separate receiver station with separate wireless connections. A guard time between the e-mail uploads can be short. Only the a configurable guard time (default 15 s) between two e-mail upload attempts is necessary to be sure that the radio resource is released; otherwise there will be no "channel request" but immediately a "packet uplink assignment".

8.4.8   Further optional requirements

It should be ensured that the influence on the measurements of additional (unwanted) e-mails on the Account B is reduced to a minimum. Appropriate provisions should be made (e.g. configuration of account, junk mail filter, etc.).

Each e-mail uploaded should have a unique machine code inside e-mail body text and subject to identify the uploading system. This will be needed when different systems are sending to a common used receiving account.

The implemented E-Mail Client should be comparable to popular e-mail clients used by customers in respect of behaviour and performance.

A flexible configuration concerning the number of up- and download within one connection should be possible.

8.4.9   Test sequence for End-to-End scenario

The following test sequence uses a wireless connection between the upload client and the e-mail server and download client and the e-mail server. This test sequence is a recommendation. See also clause 8.4.2.1 End-to-End scenario for further information.

Target: Client A uploads a set of different e-mails while connected wireless from Account A to Account B. Simultaneous the Client B is already connected to Account B and downloads new e-mails as soon as they arrive in the Account B inbox.

This test scenario uses the preconfiguration "2. At first download of header and then download of all e-mail data (body + attachment)" (clause 8.4.5 "Support for header download") for the header download. For the optional choice of the header setup "1. Complete download of all e-mails" the Notification Phase is not used and for the header setup "3. Only download of header" the Download Phase is not used.
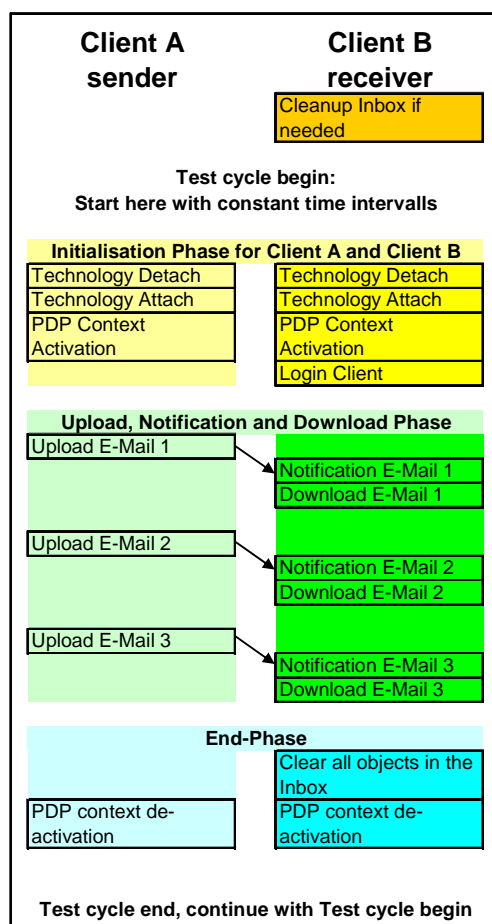
Figure 5

Precondition: Technology (e.g. UMTS PS) is available.

Cleanup the Client B Inbox somehow if the test starts after a longer pause to delete over-aged content and junk e-mails.

The test sequence shall be started with a constant time interval also if the test case fails. This avoids an over evaluation during phases of bad quality.

Initialization Phase for Client A and Client B side:

1) Perform the Technology Detach (optional if Technology Attach should be tested).

2) Configurable pause (15 s default).

3) Perform the Technology Attach and (optional if Technology Attach should be tested):

   - measure success/failure (TS 102 250-2 [3], **5.3 Attach Failure Ratio**) and measure time (TS 102 250-2 [3], **5.4 Attach Setup Time**).

4) Configurable pause (15 s default).

5) Activate PDP context for e-mail transfer:

   - measure success/failure (TS 102 250-2 [3], **5.5 PDP Context Activation Failure Ratio**) and measure time (TS 102 250-2 [3], **5.6 PDP Context Activation Time**).

6) Configurable pause (15 s default).

7) *Just in case of Intranet access:*

   - Establish the VPN tunnel for establishment.

   - measure success/failure (tbd.) and measure time (tbd.).

8) Just in case of Intranet access:

- Configurable pause (15 s default).

9) Login Client B to Account B (e-mails will be send from Account A to another Account B).

Measure failure ratio and time for service access.

→ success/failure (TS 102 250-2 [3], **7.12.26 11 E-Mail Client Login Non-AccessibilityLogin Non-Accessibility**) of each service access and measure the time (TS 102 250-2 [3], **7.12.27 12 E-Mail Client Login Access TimeLogin Access Time**).

Upload-Phase

10) Attempt to upload defined e-mails of configurable list with a configurable pause after each upload and a configurable time-out value. Each e-mail in the list has a configurable index or indicator. Parameter values for e-mails with the same indicator-index will be evaluated together. (With this technique the number and type of different e-mails in between one PDP-Context can vary in different campaigns and the number of reporting values is configurable too.)

Perform an e-mail service access and data upload for each e-mail message.

- Success/failure (TS 102 250-2 [3], 7.12.4 E-Mail Client Upload IP-Service Access Failure Ratio{Service} Message Upload Access Failure Ratio) of each e-mail service access and measure the time (TS 102 250-2 [3], 7.12.5 E-Mail Client Upload IP-Service Access Failure Ratio{Service} Message Upload Access Time).

- Success/failure (TS 102 250-2 [3], 7.12.6 9 E-Mail Client Upload Data Transfer Cut-off Ratio{Service} Message Upload Data Transfer Cut-off Ratio) of each e-mail data and measure the time (TS 102 250-2 [3], 7.12.7 10 E-Mail Client Upload Data Transfer Cut-off Ratio{Service} Message Upload Data Transfer Time).

- Success/failure (TS 102 250-2 [3], 7.12.2 6 E-Mail Client Upload Session Failure Ratio{Service} Message Upload Failure Ratio) of each transfer and measure the time (TS 102 250-2 [3], 7.12.3 7 E-Mail Client Upload Session {Service} Message Upload Time).

NOTE 1: A configurable guard time (see clause 8.4.7.1 Guard times) between two e-mail upload attempts is necessary to be sure that the radio resource is released; otherwise there will be no "channel request" but immediately a "packet uplink / downlink assignment".

11) Perform PDP context de-activation for the client A.

12) Configurable pause (see clause 8.4.7.1   Guard times) before starting next e-mail cycle.

Notification-Phase

This phase runs in parallel to the Upload-Phase.

Header information of all incoming e-mails are downloaded here. This includes also junk e-mail and old incoming e-mail of previous tests. The impact to the performance parameters should be minor, if all precautions are used to reduce this effect.

13) Download Header Information.
Perform an e-mail service access for header download and download incoming header data.

→ measure success/failure (TS 102 250-2 [3], **7.12.8 13 E-Mail Client Notification Start Failure Ratio{Service} Notification Start Failure Ratio**) of each notification start and measure the time (TS 102 250-2 [3], **7.12.9 14 E-Mail Client Notification Start {Service} Notification Start Time**).

→ measure success/failure (TS 102 250-2 [3], **7.2.4 E-Mail Client Header Download IP-Service Access Failure Ratio7.1.12 {Service} Notification Download Access Failure Ratio**) of each service access and measure the time (TS 102 250-2 [3], **7.2.5 E-Mail Client Header Download IP-Service Setup Time7.1.13 {Service} Notification Download Access Time**).

→ measure success/failure (TS 102 250-2 [3], **7.12.14 9 E-Mail Client Header Download Data Transfer Cut-off Ratio{Service} Notification Data Transfer Cut-off Ratio**) of each e-mail header transfer and measure the time (TS 102 250-2 [3], **7.12.15 10 E-Mail Client Header Download Data Transfer {Service} Notification Data Transfer Time**).

→ measure success/failure (TS 102 250-2 [3], **7.12.10 6 E-Mail Client Header Download Session Failure Ratio{Service} Notification Download Failure Ratio**) of each header download and measure the time (TS 102 250-2 [3], **7.12.11 7 E-Mail Client Header Download Session Failure Ratio{Service} Notification Download Time**) Configurable pause (15 s default).

NOTE 2: Take care about the quality of the notification parameters because they depend on the upload performance. See clause 8.4.9 Test sequence for End-to-End scenario for details.

Download Header Information has to be successful to continue with the Download Phase, otherwise count as notification failure and proceed with next (possible) transactions or cycle respectively.

Download-Phase

This phase runs in parallel to the Notification-Phase.

14) Attempt to download all new and relevant e-mails related to the actual test case notified e-mails (see note3) with a configurable pause after each download or within pause triggered by the notification (see clause 8.4.5 Support for header download) and a configurable time-out value. Each e-mail in the list has a configurable index of indicator. Parameter values for e-mails with the same indicator-index will be evaluated together. E-Mails of the same type and length shall be evaluated together, auto assign of the indicator-index.

It should be only downloaded new e-mails here. Old e-mails of previous tests which arrived too late should be ignored and deleted at the end of the test cycle.

NOTE 3: Relevant e-mails means only to download new e-mails which are sent before related to the actual running test. E-Mails which are notified before during the notification phase need to be differentiated between junk e-mails, old e-mails which belongs to the previous tests before and new e-mails which are belongs to the actual running test by interpret unique identifier inside the subject. All e-mails will be deleted at the end of the test.

Perform an e-mail service access and download e-mail data.

→ measure success/failure (TS 102 250-2 [3], **7.12.18 4 E-Mail Client Download IP-Service Access Failure Ratio{Service} Message Download Access Failure Ratio**) of each service access and measure the time (TS 102 250-2 [3], **7.12.159 E-Mail Client Download IP-Service Setup Time{Service} Message Download Access Time**).

→ measure success/failure (TS 102 250-2 [3], **7.12.20 9 E-Mail Client Download Data Transfer Cut-off Ratio{Service} Message Download Data Transfer Cut-off Ratio**) of each e-mail download data transfer and measure the time (TS 102 250-2 [3], **7.12.21 10 E-Mail Client Download Data Transfer{Service} Message Download Data Transfer Time**).

→ measure success/failure (TS 102 250-2 [3], **7.12.16 6 E-Mail Client Download Session Failure Ratio{Service} Message Download Failure Ratio**) of each e-mail download and measure the time (TS 102 250-2 [3], **7.12.717 E-Mail Client Download Session {Service} Message Download Time**).

→ measure success/failure (TS 102 250-2 [3], **7.12.1522 E-Mail Client Header and E-Mail Download Failure Ratio{Service} Notification and Message Download Failure Ratio**) of "header and e-mail" download and measure the time (TS 102 250-2 [3], **7.12.23 16 E-Mail Client Header and E-Mail Download {Service} Notification and Message Download Time**).

NOTE 4: Take care about the quality of the download parameters because they depend on the upload performance. See clause 8.4.9 Test sequence for End-to-End scenario for details.

→ measure success/failure (TS 102 250-2 [3], **7.12.24 17 E-Mail Client to Client End-to-End Failure Ratio{Service} End-to-End Failure Ratio**) of complete test (upload, notification and download) and measure the time (TS 102 250-2 [3], **7.12.25 18 E-Mail Client to Client End-to-End Time{Service} End-to-End Time**).

NOTE 5: After each download it should be checked, if the downloaded e-mail is received completely and is identical to the data on the server (including attachment, if any). Such a failure should be included in the analysis.

NOTE 6: Only if the download is not triggered by the notification: A configurable guard time (default 15 s) between two e-mail download attempts is necessary to be sure that the radio resource is released; otherwise there will be no "channel request" but immediately a "packet uplink / downlink assignment".

End-Phase

15) Clear all objects in the Inbox of the Client B.

16) Perform PDP context de-activation.

17) Configurable pause (15 s default) before starting next e-mail cycle.

8.4.10   Test sequence for Download scenario

The following test sequence uses a wireless connection between the download client and the e-mail server. This test sequence is a recommendation. See also clause 8.4.2.2 for further information.
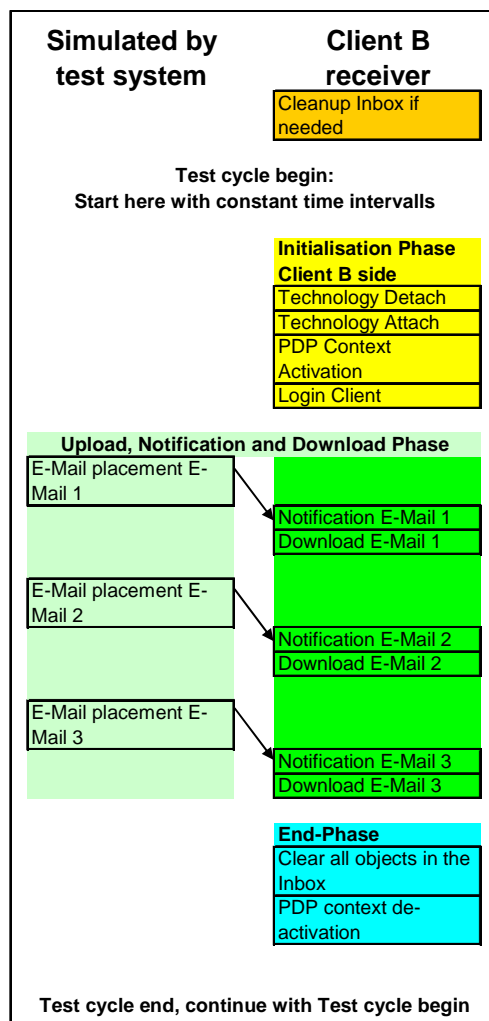
Figure 6

Target: The new e-mails are placed directly in the Account B inbox by the test system which is connected to the e-mail server with a second communication way. Client B connects wireless to Account B and downloads new e-mails as soon as they arrive in the Account B inbox.

This test scenario uses the preconfiguration "At first download of header and then download of all data" (see clause 8.4.5 Support for header download) for the header download. For the optional choice of the header setup "Complete download of all e-mails" the Notification Phase is not used and for the header setup "Only download of header" the Download Phase is not used.

Precondition: Technology (e.g. UMTS PS) is available.

Cleanup the Client B Inbox somehow if the test starts after a longer pause to delete over-aged content and junk e-mails.

The test sequence shall be started with a constant time interval also if the test case fails. This avoids an over evaluation during phases of bad quality.

Initialization Phase for Client B side

1) Perform the Technology Detach (optional if Technology Attach should be tested).

2) Configurable pause (15 s default).

3) Perform the Technology Attach and (optional if Technology Attach should be tested):

    - measure success/failure (TS 102 250-2 [3], **5.3 Attach Failure Ratio**) and measure time
      (TS 102 250-2 [3], **5.4 Attach Setup Time**).

4) Configurable pause (15 s default).

5) Activate PDP context for e-mail transfer:

    - measure success/failure (TS 102 250-2 [3], **5.5 PDP Context Activation Failure Ratio**) and measure
      time (TS 102 250-2 [3], **5.6 PDP Context Activation Time**).

6) Configurable pause (15 s default).

7) *Just in case of Intranet access:*

    - Establish the VPN tunnel for establishment.

    - measure success/failure (tbd.) and measure time (tbd.).

8) Just in case of Intranet access:

    - Configurable pause (15 s default).

9) Login Client B to Account B.

Measure failure ratio and time for service access.
→ success/failure (TS 102 250-2 [3], **7.2.11 E-Mail Client Login Non-Accessibility7.1.26 Login Non-Accessibility**)
of each service access and measure the time (TS 102 250-2 [3], **7.2.12 E-Mail Client Login Access Time7.1.26 Login
Non-Accessibility).**

E-Mail placement

10) To simulate the upload phase the test system shall place defined e-mails of configurable list in the tested Inbox
    with a configurable pause after each upload. Each e-mail in the list has a configurable index or indicator.

Notification-Phase

This phase runs in parallel to the "E-Mail placement".

Header information of all incoming e-mails are downloaded here. This includes also junk e-mail and old incoming
e-mail of previous tests. The impact to the performance parameters should be minor, if all precautions are used to
reduce this effect.

11) Download Header Information.
    Perform an e-mail service access for header download and download incoming header data.

→ measure success/failure (TS 102 250-2 [3], **7.2.13 E-Mail Client Notification Start Failure Ratio7.1.8 {Service}
Notification Start Failure Ratio**) of each notification start and measure the time (TS 102 250-2 [3], 7.**2.14 E-Mail
Client Notification Start Time7.1.9 {Service} Notification Start Time).**

→ measure success/failure (TS 102 250-2 [3], **7.2.4 E-Mail Client Header Download IP-Service Access Failure
Ratio7.1.12 {Service} Notification Download Access Failure Ratio**) of each service access and measure the time
(TS 102 250-2 [3], **7.2.5 E-Mail Client Header Download IP-Service Setup Time7.1.13 {Service} Notification
Download Access Time).**

→ measure success/failure (TS 102 250-2 [3], **7.2.9 E-Mail Client Header Download Data Transfer Cut-off Ratio7.1.14 {Service} Notification Data Transfer Cut-off Ratio**) of each e-mail header transfer and measure the time (TS 102 250-2 [3], **7.2.10 E-Mail Client Header Download Data Transfer Time7.1.15 {Service} Notification Data Transfer Time**).

→ measure success/failure (TS 102 250-2 [3], **7.2.6 E-Mail Client Header Download Session Failure Ratio7.1.10 {Service} Notification Download Failure Ratio**) of each header download and measure the time (TS 102 250-2 [3], **7.2.7 E-Mail Client Header Download Session Failure Ratio7.1.11 {Service} Notification Download Time**) Configurable pause (15 s default).

12) Download of header information has to be successful to continue with the Download Phase, otherwise count as notification failure and proceed with next (possible) transactions or cycle respectively.

Download-Phase

This phase runs in parallel to the Notification-Phase:

13) Attempt to download all new and relevant e-mails related to the actual test case notified e-mails (see note 7) with a configurable pause after each download or within pause triggered by the notification (see clause 8.4.5 Support for header download) and a configurable time-out value. Each e-mail in the list has a configurable index of indicator. Parameter values for e-mails with the same indicator-index will be evaluated together. E-Mails of the same type and length shall be evaluated together, auto assign of the indicator-index.

It should be only downloaded new e-mails here. Old e-mails of previous tests which arrived too late should be ignored and deleted at the end of the test cycle.

NOTE 7: Relevant e-mails means only to download new e-mails which are sent before related to the actual running test. E-Mails which are notified before during the notification phase need to be differentiated between junk e-mails, old e-mails which belongs to the previous tests before and new e-mails which are belongs to the actual running test by interpret unique identifier inside the subject. All e-mails will be deleted at the end of the test.

Perform an e-mail service access and download e-mail data.
→ measure success/failure (TS 102 250-2 [3], **7.2.4 E-Mail Client Download IP-Service Access Failure Ratio7.1.18 {Service} Message Download Access Failure Ratio**) of each service access and measure the time (TS 102 250-2 [3], **7.2.5 E-Mail Client Download IP-Service Setup Time7.1.19 {Service} Message Download Access Time**).

→ measure success/failure (TS 102 250-2 [3], **7.2.9 E-Mail Client Download Data Transfer Cut-off Ratio7.1.20 {Service} Message Download Data Transfer Cut-off Ratio**) of each e-mail download data transfer and measure the time (TS 102 250-2 [3], **7.2.10 E-Mail Client Download Data Transfer Time7.1.21 {Service} Message Download Data Transfer Time**).

→ measure success/failure (TS 102 250-2 [3], **7.2.6 E-Mail Client Download Session Failure Ratio7.1.16 {Service} Message Download Failure Ratio**) of each e-mail download and measure the time (TS 102 250-2 [3], **7.2.7 E-Mail Client Download Session Time7.1.17 {Service} Message Download Time**).

→ measure success/failure (TS 102 250-2 [3], **77.2.15 E-Mail Client Header and E-Mail Download Failure Ratio.1.22 {Service} Notification and Message Download Failure Ratio**) of "header and e-mail" download and measure the time (TS 102 250-2 [3], **7.2.16 E-Mail Client Header and E-Mail Download Time7.1.23 {Service} Notification and Message Download Time**).

→ measure success/failure (TS 102 250-2 [3], **7.2.17 E-Mail Client to Client End-to-End Failure Ratio7.1.24 {Service} End-to-End Failure Ratio**) of complete test (upload, notification and download) and measure the time (TS 102 250-2 [3], **7.2.18 E-Mail Client to Client End-to-End Time7.1.25 {Service} End-to-End Time**).

NOTE 8: After each download it should be checked, if the downloaded e-mail is received completely and is identical to the data on the server (including attachment, if any). Such a failure should be included in the analysis.

NOTE 9: Only if the download is not triggered by the notification: A configurable guard time (default 15 s) between two e-mail download attempts is necessary to be sure that the radio resource is released; otherwise there will be no "channel request" but immediately a "packet uplink / downlink assignment".

End-Phase

14) Perform PDP context de-activation.

15) Clear all objects in the Inbox of the Client B.

16) Configurable pause (15 s default) before starting next e-mail cycle.

9   Data measurements

In the following clauses, data measurements for internet-related services are described. While process of obtaining a connection is different between Circuit-switched and Packet-switched access, there are many similarities for the actual data transfer phase.

9.1   Common aspects

9.1.1   Transaction definition and transaction types for data measurements

A transaction consists of access to a server to obtain content which is a closed unit from the user's perspective, e.g. a single e-mail, a single downloaded file, or a web-site viewing access which may consist of several single objects which form the desired web page to be viewed.

9.1.2   Server types

The following categories of servers are distinguished, it is assumed that data service access is made to a server or entity within the general internet domain:

- Third-party content in the public internet. We will term this type of counterpart A-Servers.

- Accounts on internet servers which are under control of the testing system or under control of testing personnel (e.g. web domains assigned for testing purposes). This type of counterpart shall be termed B-Servers.

- Special servers not reachable via public internet (e.g. in the GGSN domain), or equipped with additional instrumentation e.g. for IP-level tracing or non-internet capabilities (e.g. UDP testing). This type of counterpart shall be termed C-Servers.

Access to a particular type of server will be termed using the same letter, e.g. A-access for access to an A-server.

It is assumed that A-servers are generally outside the sphere of control of testing systems, in particular, no baseline information for load or other conditions is obtainable. In case of B-servers and C-servers, such control may exist, but will typically be limited in some way due to IP-security reasons, in particular if IP access occurs intra-network.

Typically, when testing internet services, availability may depend on influences other than those under test, and performance will be affected by third-party traffic. Meaningful tests therefore shall contain appropriate measures to exclude such effects from QoS assessment, or make sure that all networks under test are affected the same way. It is assumed that such tests are performed by fixed network units. Suggested methods are:

- Cyclical availability checks on target servers or domains.

- Cyclical access-time tests.

It is assumed that for services where login is required, accounts used by the test system are valid, give positive login, and are good for full access to all activities forming the test.

NOTE: This covers read/write privileges and directory-access rights e.g. for FTP).

User's point of view typically includes assumptions of a time the user is ready to wait before an action is considered to be failed. At the same time, IP service access typically goes with timeout windows at several levels, e.g. for in activity over a certain period of time. Test design must combine these two aspects to a reasonable, technically feasible set of parameters.

9.1.3   Test data content

When using web content (web sites or single pages) for testing, it must be taken into account that such content will typically change frequently (e.g. content of popular web portals) and therefore performance tests may give varying results over time. Test design must ensure that such effects are excluded from QoS assessment. Preferably, standardized and constant web content shall be used.

The degree of control a testing system has further depends on the type of service.

With E-mail and FTP (assuming appropriate access privileges), data content can be determined exactly (e.g. by uploading files to be downloaded as part of subsequent tests.

9.1.4   Transaction phase and parameter overview

9.1.4.1   General

To ensure comparability and statistical validity of transactions, the following outer conditions need to be constant throughout a test case:

- Access timing including behaviour in case of failure to obtain IP access. For reasons of comparability in density, it is required that all access attempts which form part of the statistics must follow a constant time pattern. Additional attempts e.g. in the case of network or service unavailability must not affect the pattern of the statistically relevant access attempts.

For services using buffering, such as video streaming, there may be the situation that while actual data transfer is already running, the visual appearance is that of still waiting. This period of time shall be considered part of service usage.

Typically, in IP services periods of inactivity can occur with the session still intact. On the other hand, it must be taken into account that a useful PDP context is indicated by the system, which is in fact not useful. Therefore, testing shall include cyclical "lifecheck" measures. A possible method are ICMP pings to the DNS, not if such pings are blocked by the network, DNS accesses with dummy URL. Due to possible URL/IP address storage, it must be assured that actual DNS access takes place.

Independently of the type of access, the following general information elements need to be logged in order to ensure reproducibility and comparability of tests.

- Operating system (type and version).

- MTU size.

- Logical location of server (public internet, GGSN, etc.) with respect to effects possibly created by other traffic.

- Maximum server throughput inbound and outbound, per session response per connection.

9.1.4.2 Packet-switched access

Table 10

Phase	Subphase	Parameters
Internet access	Network access	Equipment Types and Capabilities Type of access used (CSD/PSD. UE initialization) APN and other initial settings
	Session access	Access and authentication parameters for basic internet access
	DNS access	DNS (in case of non-automatic DNS assignment during session access procedure)
	Domain access	Target URL, or server IP address Account being used (where appropriate)
Data transfer		Content composition and size
Cleardown		Pause between access attempts

9.1.4.3 Circuit-switched access

Table 11

Phase	Subphase	Parameters
Internet access	Connection set-up	Equipment Types and Capabilities Parameters for dial-up connection
	Session access	Access and authentication parameters for basic internet access
	DNS access	DNS (in case of non-automatic DNS assignment during session access procedure)
	Domain access	Target URL, or server IP address Account being used (where appropriate)
Data transfer		Content composition and size
Cleardown		Pause between access attempts

9.1.5 Possible transaction results

Commonly, a data transaction is termed "session" in contrast to "call" regardless of the type of connection (CS or PS). However, the general result term "dropped" shall be used, leading to "dropped session" for non-completed service usage.

In case of dropped sessions, the testing system shall indicate which of the possible principal causes occurred for the purpose of deciding of the network under test is to be blamed or not:

- Loss of radio connection.

- Loss of basic IP connection. If DNS access is still possible after a session loss, it is assumed that basic IP connection is still intact.

- Loss of internet access: If access to another domain or server is still possible, it shall be assumed that basic internet.

- Loss of connection to the server.

For the single phases of internet access (as part of general service access), the following additional definitions are given:

- **Session access:** Access and authentication procedure for basic internet access. Successful access is assumed when a temporary IP address good for data transactions has been assigned to the testing system.

- **DNS access:** Obtaining an IP address from a URL (web site name, server name). Successful DNS access is assumed when an IP address has been obtained from a given, valid URL.

- **Domain access:** This is the actual internet access as seen from the user's perspective. For testing purposes, a successful domain access can be assumed when communication with the target server has been proved. The respective action will depend on the service under test.

9.1.6   Content quality

Content quality relates to the quality as perceived by the user.

Content integrity is correctness and completeness.

Technical precautions should be taken to make sure that the content received is the one expected.

9.2   FTP

9.2.1   Transaction definition and transaction types for FTP

The basic transaction for FTP testing consists of internet access to a FTP server followed by either downloading or uploading a single file of given size.

It is understood that basic access procedures such as server login shall not be considered as part of a transaction.

For download tests, it must be assured that the file to be downloaded is actually available on the server. For upload tests, it must be assured that no storage-size limitations prevent successful upload.

9.2.2   Transaction phase and parameter overview for FTP

Same as family definitions.

In addition, the following parameters shall be logged:

- type of FTP client used;

- protocol used (TCP/IP or UDP);

- type used (active or passive).

9.2.3   Possible transaction results for FTP

Same as family definitions.

9.2.4   Content quality for FTP

Same as family definitions.

9.2.5   Content integrity for FTP

Same as family definitions It is recommended to use file size and checksum comparison.

9.3   HTTP

9.3.1   Transaction definition and transaction types for HTTP

The basic transaction for HTTP testing consists of internet access followed by downloading a web site for a given URL response of given structure.

9.3.2   Transaction phase and parameter overview for HTTP

Same as family definitions.

In addition, the following parameters shall be logged:

- Type of browser used (including version number/build). For test-system browsers, maximum number of parallel socket connections.

- Reference web site used (structure, size of documents, etc.).

9.3.3   Possible transaction results for HTTP

Same as family definitions.

9.3.4   Content quality for HTTP

Same as family definitions.

9.3.5   Content integrity for HTTP

It is assumed that a typical HTTP downloaded page consists of a main page and a number of sub-elements contained in this page, which may have sub-elements themselves. Based on the family definition, recommended method for content integrity checking is to verify that the expected main page has been loaded, and to verify the expected number and type of sub-elements contained in each element.

9.4   E-mail

9.4.1   Transaction definition and transaction types for E-mail

The basic transaction for e-mail testing consists of internet access to a e-mail server followed by either downloading or uploading a given number of e-mails of given size.

It is understood that basic access procedures such as server login shall not be considered as part of a transaction.

Proper authentication must be assured. For download tests, it must be assured that the mails to form download traffic are actually present on the server. For upload tests, it must be assured that no storage-size limitations prevent successful upload.

9.4.2   Transaction phase and parameter overview for E-mail

Same as family definitions.

In addition, the following parameters shall be logged:

- Type of e-mail client used (including version number/build). For test-system clients, maximum number of parallel socket connections.

- Authentication mode used.

9.4.3   Possible transaction results for E-mail

Same as family definitions.

9.4.4   Content quality for E-mail

Same as family definitions.

9.4.5   Content integrity for E-mail

Same as family definitions.

9.5   WAP

9.5.1   Transaction definition and transaction types for WAP

The basic transaction for WAP testing consists of WAP access and download of a WAP page.

To guarantee comparability and statistical validity of transactions, the following parameters have to be logged:

- Protocol (WAP1.x or WAP2.0).

- Access-Point (Gateway).

- Type of mobile used / emulated (user-agent string, accept-header).

- Mobile browser (version number, build).

- Depending on the test goal one of the following clauses should be used:

    - A given reference WAP page (static WAP content)
      Picture count and text size have to be stable.

    - A static url e.g. main page of the WAP-portal (variable WAP content)
      Picture count and text size have to fulfil minimum requirements depending on the url.

  NOTE 1: It is required that the WAP page has a size greater than one packet.

- It is required that all WAP session attempts follow a constant time pattern.

  NOTE 2: The size of the mobile cache can be neglected, because all tests should be performed with an empty mobile cache.

The mobile specific effects (rendering) are not represented in case of non application level testing.

9.5.2   Transaction phase and parameter overview for WAP

Same as family definitions.

In addition, the following parameters shall be logged:

- Protocol (WAP1.x or WAP2.0).

- Access-Point (Gateway).

- mobile used (user-agent string, accept-header, browser type).

- Structure of the WAP page (size of content, number of pictures, etc.).

- URL of the WAP page.

When measuring the parameter using drive tests or field tests, it is recommended that the following general method should be used.
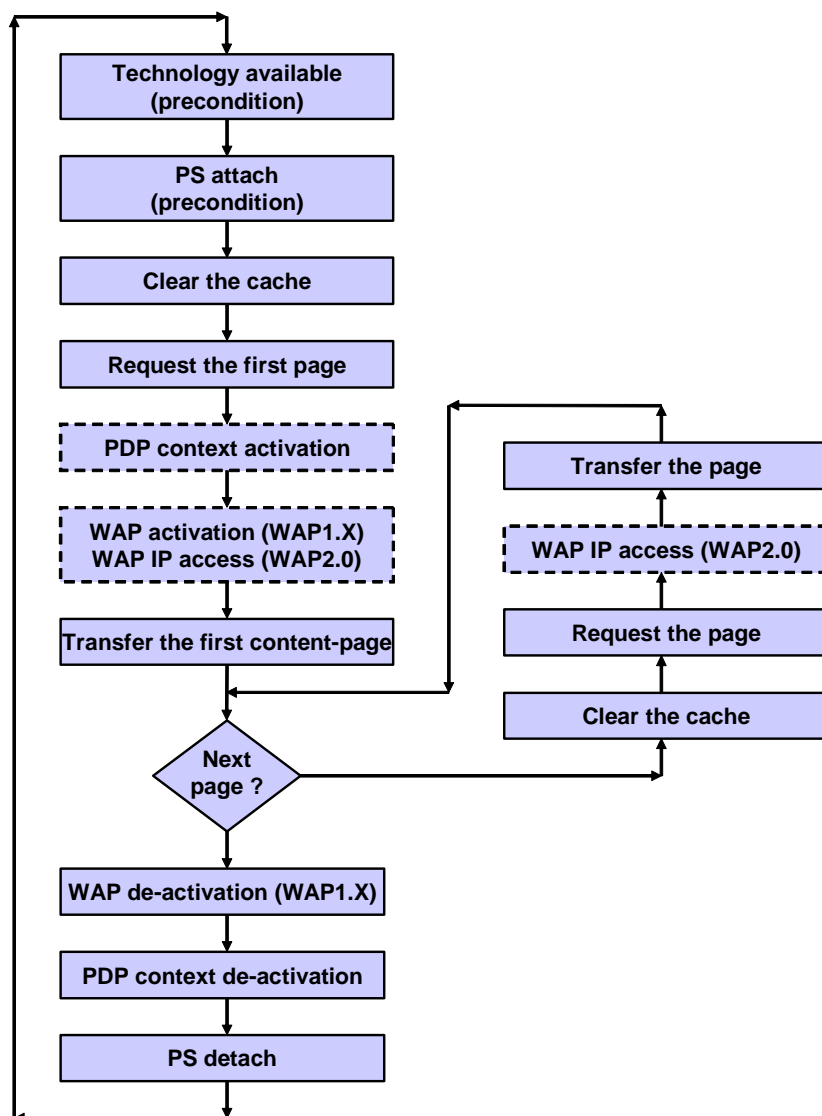

Figure 7

9.5.3   Possible transaction results for WAP

Same as family definitions.

9.5.4   Content quality for WAP

Same as family definitions.

9.5.5   Content integrity for WAP

The following is to verify that the received page is not an "error message" page and that the requested page is downloaded.

It is assumed that a typical WAP downloaded page consists of a main page and a number of sub-elements contained in this page, which may have sub-elements themselves. Based on the family definition, recommended method for content integrity checking is to verify that all elements of the expected main page have been completely loaded.

- In case of testing a given reference WAP page (static WAP content) the whole (known) content should be verified.

In case of testing public WAP pages with variable WAP content it is necessary to define and check the page regarding minimum criteria (e.g. five pictures and text size greater than 150 Byte). For identifying the content of the requested page it is necessary to check static page content (logo, keyword).

9.6 Streaming Video

9.6.1 Transaction definition and transaction types for streaming video

The basic transaction for Streaming Video testing consists of internet access to a Streaming server followed by replay access to a given content from this server.

9.6.2 Transaction phase and parameter overview for streaming video

Same as family definitions.

In addition, the following parameters shall be logged:

- streaming Server and Client versions used;

- OS and relevant configuration of streaming server;

- streaming protocol used.

9.6.3 Possible transaction results for Streaming Video

Same as family definitions.

9.6.4 Content quality for Streaming Video

To be decided.

9.6.5 Content integrity for Streaming video

To be decided.

History

Document history		
V1.1.1	October 2003	Publication (Withdrawn)
V1.2.1	March 2004	Publication (Withdrawn)
V2.1.1	February 2005	Publication (Withdrawn)
V1.3.1	April 2005	Publication
V1.3.2	September 2005	Publication
V1.4.1	December 2008	Publication