

Recommendation

ITU-T P.1503 (03/2023)

SERIES P: Telephone transmission quality, telephone installations, local line networks

Methods for objective and subjective assessment of quality of services other than speech and video

Extended methodology for cross-country and inter-operator digital financial services testing

ITU-T P-SERIES RECOMMENDATIONS

TELEPHONE TRANSMISSION QUALITY, TELEPHONE INSTALLATIONS, LOCAL LINE NETWORKS

Vocabulary and effects of transmission parameters on customer opinion of transmission quality	P.10–P.19
Voice terminal characteristics	P.30–P.39
Reference systems	P.40–P.49
Objective measuring apparatus	P.50–P.59
Objective electro-acoustical measurements	P.60–P.69
Measurements related to speech loudness	P.70–P.79
Methods for objective and subjective assessment of speech quality	P.80–P.89
Voice terminal characteristics	P.300–P.399
Objective measuring apparatus	P.500–P.599
Measurements related to speech loudness	P.700–P.709
Methods for objective and subjective assessment of speech and video quality	P.800–P.899
Audiovisual quality in multimedia services	P.900–P.999
Transmission performance and QoS aspects of IP end-points	P.1000–P.1099
Communications involving vehicles	P.1100–P.1199
Models and tools for quality assessment of streamed media	P.1200–P.1299
Telemeeting assessment	P.1300–P.1399
Statistical analysis, evaluation and reporting guidelines of quality measurements	P.1400–P.1499
Methods for objective and subjective assessment of quality of services other than speech and video	P.1500–P.1599

For further details, please refer to the list of ITU-T Recommendations.

Recommendation ITU-T P.1503

Extended methodology for cross-country and inter-operator digital financial services testing

Summary

Recommendation ITU-T P.1503 is based on Recommendation ITU-T G.1033, which standardizes a conceptual framework for quality of service and quality of experience (QoE) aspects of digital financial services (DFSs), and on Recommendation ITU-T P.1502, which specifies a methodology for QoE testing of DFSs for the basic person-to-person (P2P) money transfers between two devices using the same network and DFS operator.

Recommendation ITU-T P.1503 has three main parts.

First, the methodological framework and specifications for a generalized P2P money transfer use case are given. In this framework, the DFS operator used to send money, and the operator receiving the same funds (i.e., the A and B side of a money transfer) are parameters in the use case, which integrates all variations (same operator or inter-operators; intra- or international) into the same methodological context.

Second, Recommendation ITU-T P.1503 provides a comprehensive framework for data elements and related processing, including tools and procedures providing operational robustness and a high level of data quality. The data objects specified here support test planning and management as well as provide the input data foundation for efficient processing of data. Guidance is also given on how data processing can be done in a consistent and efficient way.

Third, this Recommendation describes a new tool and the corresponding methodology, designed to assist field test teams in data collection. This "multi-stopwatch" tool is conceptually suggested in Recommendation ITU-T P.1502. It is an electronic time-taking tool similar to a stopwatch but supporting testers to record the events within a DFS test case and upload data entered by testers directly to a central location. This tool eliminates the needs and weaknesses of manual entering of time readings and the process of transferring them to post-processing through multiple transformation stages, e.g., from handwritten notes to entries in a spreadsheet.

History

Edition	Recommendation	Approval	Study Group	Unique ID*
1.0	ITU-T P.1503	2023-03-01	12	11.1002/1000/15471

Keywords

Digital financial services, QoE, QoS, quality of experience, quality of service.

* To access the Recommendation, type the URL <http://handle.itu.int/> in the address field of your web browser, followed by the Recommendation's unique ID. For example, <http://handle.itu.int/11.1002/1000/11830-en>.

FOREWORD

The International Telecommunication Union (ITU) is the United Nations specialized agency in the field of telecommunications, information and communication technologies (ICTs). The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of ITU. ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The World Telecommunication Standardization Assembly (WTSA), which meets every four years, establishes the topics for study by the ITU-T study groups which, in turn, produce Recommendations on these topics.

The approval of ITU-T Recommendations is covered by the procedure laid down in WTSA Resolution 1.

In some areas of information technology which fall within ITU-T's purview, the necessary standards are prepared on a collaborative basis with ISO and IEC.

NOTE

In this Recommendation, the expression "Administration" is used for conciseness to indicate both a telecommunication administration and a recognized operating agency.

Compliance with this Recommendation is voluntary. However, the Recommendation may contain certain mandatory provisions (to ensure, e.g., interoperability or applicability) and compliance with the Recommendation is achieved when all of these mandatory provisions are met. The words "shall" or some other obligatory language such as "must" and the negative equivalents are used to express requirements. The use of such words does not suggest that compliance with the Recommendation is required of any party.

INTELLECTUAL PROPERTY RIGHTS

ITU draws attention to the possibility that the practice or implementation of this Recommendation may involve the use of a claimed Intellectual Property Right. ITU takes no position concerning the evidence, validity or applicability of claimed Intellectual Property Rights, whether asserted by ITU members or others outside of the Recommendation development process.

As of the date of approval of this Recommendation, ITU had not received notice of intellectual property, protected by patents/software copyrights, which may be required to implement this Recommendation. However, implementers are cautioned that this may not represent the latest information and are therefore strongly urged to consult the appropriate ITU-T databases available via the ITU-T website at <http://www.itu.int/ITU-T/ipr/>.

© ITU 2023

All rights reserved. No part of this publication may be reproduced, by any means whatsoever, without the prior written permission of ITU.

Table of Contents

		Page
1	Scope	1
2	References.....	1
3	Definitions	2
	3.1 Terms defined elsewhere	2
	3.2 Terms defined in this Recommendation	2
4	Abbreviations and acronyms	2
5	Conventions	3
6	Test scenario under consideration	3
	6.1 Roles, entities and action or /event flow	3
	6.2 Test parameterization and neutral starting state	4
	6.3 Re-initialization after unsuccessful transactions	5
	6.4 Disappearing money.....	5
	6.5 Automation of tests	5
7	Transaction model and digital financial service key performance indicators	6
8	Creating the use case model from actual use case examples.....	8
9	Data sources.....	8
10	Special procedures in the field.....	9
	10.1 Operational protocol if party A and B do not communicate directly.....	9
11	Testing modes.....	11
	11.1 Basics	11
	11.2 Synchronized and asynchronous testing mode.....	11
	11.3 Dual-function set-up and operation.....	13
12	Recording events of DFS test cases	14
	12.1 Basic functionality.....	14
	12.2 Practical application	14
13	Measurements in the background.....	15
14	General considerations about errors in measurements.....	16
15	Data validation and processing.....	17
	15.1 Overview	17
	15.2 Plausibility and validity checks	18
	15.3 Data processing.....	19
	15.4 Time profile	20
16	Background testing of mobile networks	20
	Appendix I – Working example for creation of a use case model from actual implementations	22
	Appendix II – Best-practice example of data structures and procedures used in DFS testing campaigns	23

	Page
II.1 Basic considerations for data collection and processing.....	23
II.2 Additional considerations for cross-border and inter-network testing.....	23
II.3 Data structure overview	25
II.4 Naming and formatting conventions.....	26
II.5 Naming and formatting conventions for data objects	27
II.6 Local log sheets	27
II.7 Team assignment list.....	28
II.8 Device assignment list.....	30
Appendix III – Best-practice examples for background testing of mobile networks	33
Bibliography	35

Electronic attachment: Spreadsheet log file templates

Recommendation ITU-T P.1503

Extended methodology for cross-country and inter-operator digital financial services testing

1 Scope

This Recommendation¹ describes the quality of experience (QoE) assessment methodology for the person-to-person (P2P) money transfer use case, i.e., from party A to party B. It covers the full range of possible combinations of operator of digital financial services (DFSs) and country where the parties are located, i.e., for the case where parties A and B use the same or different operators, as well as intra- and international transfers. Therefore, this Recommendation is an extension of [ITU-T P.1502].

This Recommendation also specifies extensions of testing methodology and introduces a new type of tool to assist in testing.

The methodology is designed to be easily extended to other use cases in future editions of this Recommendation.

It is important to understand that this Recommendation only covers the methodology for tests done from an individual user's (end-to-end) perspective, acting within a given DFS ecosystem under current load conditions.

It may be desirable to extend the scope of testing to capacity tests, which would involve creation of specified load scenarios to a DFS ecosystem to determine the robustness of its functionality under these conditions. Such extensions can be easily created from the methodology described in this Recommendation. Their execution is mainly a matter of scale of required resources.

2 References

The following ITU-T Recommendations and other references contain provisions which, through reference in this text, constitute provisions of this Recommendation. At the time of publication, the editions indicated were valid. All Recommendations and other references are subject to revision; users of this Recommendation are therefore encouraged to investigate the possibility of applying the most recent edition of the Recommendations and other references listed below. A list of the currently valid ITU-T Recommendations is regularly published. The reference to a document within this Recommendation does not give it, as a stand-alone document, the status of a Recommendation.

[ITU-T E.804] Recommendation ITU-T E.804 (2014), *Quality of service aspects for popular services in mobile networks*.

[ITU-T E.840] Recommendation ITU-T E.840 (2018), *Statistical framework for end-to-end network-performance benchmark scoring and ranking*

[ITU-T G.1033] Recommendation ITU-T G.1033 (2019), *Quality of service and quality of experience aspects of digital financial services*.

[ITU-T P.1502] Recommendation ITU-T P.1502 (2020), *Methodology for QoE testing of digital financial services*.

¹ This Recommendation has an electronic attachment containing spreadsheet log file templates for project usage.

3 Definitions

3.1 Terms defined elsewhere

This Recommendation uses the following terms defined elsewhere:

3.1.1 quality of experience (QoE) [b-ITU-T P.10]: The degree of delight or annoyance of the user of an application or service.

3.1.2 quality of service (QoS) [b-ITU-T E.800]: Totality of characteristics of a telecommunications service that bear on its ability to satisfy stated and implied needs of the user of the service.

3.2 Terms defined in this Recommendation

This Recommendation defines the following term:

3.2.1 digital financial service (DFS): Methods to electronically store and transfer funds; to make and receive payments; to borrow, save, insure and invest; and to manage personal or business finances; and where these services are accessed via mobile communication devices.

4 Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

AVDTA	Accumulated Validated Data on Transaction
DAL	Device Assignment List
DFS	Digital Financial Service
FTL	Field Test Lead
GPS	Global Positioning System
ID	Identifier
IMEI	International Mobile Equipment Identity
KPI	Key Performance Indicator
MDR	Mean Data Rate
MSW	Multi-Stopwatch
MTACT	Money Transfer A-side Completion Time
MTCD	Money Transfer Core Duration
MTCR	Money Transfer Completion Rate
MTFCT	Money Transfer Full Completion Time
MTRCT	Money Transfer Raw Completion Time
P2P	Person-to-Person
PIN	Personal Identification Number
QoE	Quality of Experience
QoS	Quality of Service
SMS	Short Message Service
SQL	Structured Query Language
TA	Transaction

TAL	Team Assignment List
USSD	Unstructured Supplementary Service Data
UUID	Universally Unique Identifier
Wi-Fi	Wireless Fidelity

5 Conventions

None.

6 Test scenario under consideration

The basic scenario under consideration is the P2P money transfer between two parties in the following cases:

- in the same country, but using different DFS providers (inter-operator scenario);
- in different countries, using the same DFS providers (e.g., national branches of a multi-national network operator);
- in different countries, using different DFS providers.

It is important to state that the methodology for these variants is the same since in all cases money is transferred between two entities. There may be differences in the details of an operating sequence; these are, however, not greater than differences between same-operator, same-country operating sequences.

NOTE – In some countries, the DFS might be registered under the central bank, and money transfers considered to be bank transfers. This can have an effect on the appearance of such transactions where, in actual implementations, it has been shown that in some cases entities other than the sending party (e.g., agents) are used for international transactions. From an end-to-end perspective, this is still a P2P transfer; technically, money received in this case appears to come from that particular entity rather than from the actual sending entity. If reception notifications are used for data evaluation (which is not the case in the current methodology), this would have to be considered in data processing.

The P2P basic scenario and its modelling is described in detail in [ITU-T G.1033] and [ITU-T P.1502]. For the sake of convenience of reading, the essential parts are explained here. For more details, see [ITU-T G.1033] and [ITU-T P.1502].

6.1 Roles, entities and action or /event flow

In the P2P money transfer scenario, money is transferred from party A (the active party which is sending money) to party B (the receiving party).

In a practical test implementation, each party is represented by one testing team. In practice, money is transferred in a cyclic fashion, so teams switch roles after each transfer, as shown in Figure 1.

NOTE – Figure 1 only shows the basic case. In order to take care of the whole spectrum of possible cases during testing, some additional, derived cases need to be considered.

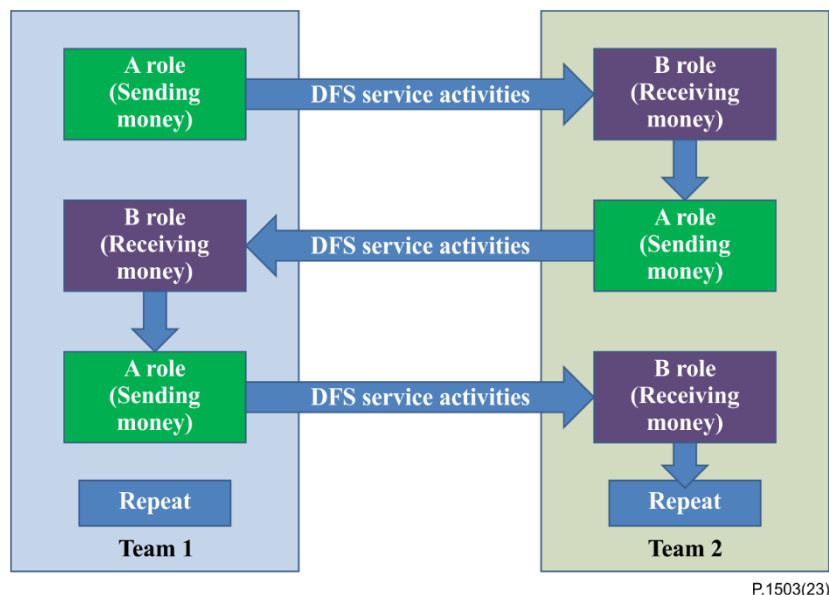


Figure 1 – P2P money transfer roles and team activities

6.2 Test parameterization and neutral starting state

A particular property of systematic service tests is frequency of usage that is significantly higher than that created by a typical end user.

While a high testing frequency leads to a high yield of samples for computation of quality of service (QoS) key performance indicators (KPIs), it is conceivable that the system has a certain dead time after each transaction, where the system does not accept a new transaction or creates unexpected results of a transaction attempted within this time period. It is advisable to be aware of this possibility and obtain relevant information before actual parameters of a test campaign are determined.

The testing frequency can be controlled by a pause between transactions, which also acts as a guard time to allow the service under test to reach its neutral state again. Respective considerations are in full analogy to testing e.g., of telephony.

A testing campaign, therefore, should contain a pre-testing phase with systematic tests to make sure that usage frequencies typical for testing do not affect testing results with respect to the end-user perspective.

NOTE – In actual implementation, DFS providers can apply limitations on the number of transactions per day, or the total amount of transferred money per day or another time period. It is also conceivable that mechanisms exist that limit the frequency of testing. When setting up tests, it is important to check for such conditions. This starts with making testers aware of the effects of such mechanisms and their preparation to adapt data processing accordingly after detection during a testing campaign. Otherwise, misclassification can result, such as ascribing effects of such limitation to functional failures of a DFS.

As the starting point for systematic testing, the guard time should typically be in the range 10 s to 30 s.

When testing is done manually, it is assumed that the system can handle all testing speeds that can be realized by human testers, as even an experienced tester will not work significantly faster than an experienced regular user of DFS. Therefore, no special requirements to slow down testing are applied.

In fully automated testing, it would also be possible to use the high degree of repeatability of such control to determine the appropriate guard time by probing, i.e., by systematically varying the guard time and check for respective effects.

There is a second category of effects that needs to be considered, namely the possibility of a service-specific local memory (analogously to a browser's cache) that stores information related to previous

transactions. The effect would be that in subsequent transactions, such information would be read from local memory instead of obtaining it by an over the air request to the service. This could then impact related measurement values or KPIs.

As long as effects are quantitative rather than qualitative, it may not be practicable and is not necessarily required to exclude frequency-dependent effects entirely. However, such effects need to be recorded and documented carefully as part of the reporting in order to understand their impact on the testing conditions.

6.3 Re-initialization after unsuccessful transactions

If a transaction fails, in particular after a time-out condition has occurred, ensure that the service and the device or application are in the typical neutral starting state again, i.e., that no memory of previous error states remains in the system.

6.4 Disappearing money

It is possible that during a transaction, the amount of money deducted is not correct with respect to transferred amount and fees. This includes the case that the amount is correct but sent to a third party by an error in the system. From an end customer perspective, this is either a loss (if too much money is deducted), or an unjustified gain (if money is credited but not deducted on the other side of the transaction). For simplicity, the term "disappear" is used for both variants of this kind of effect.

There may be undelivered transactions where money is deducted from the sender's account along with transfer charges. In such cases, filing of a complaint with the DFS operator is typically required. If this complaint is resolved successfully, money is returned at a later point in time (depending on the process and the terms of service of the DFS operator, transfer charges may not be refunded).

Retrieval of lost money is understood as a second stream of activities outside the scope of this methodology. Functionally, even if money is returned later, it reduces the available credit for further tests. Therefore, in all cases of disappearing money, insertion of fresh funds may be necessary to keep up the necessary level of credit for further testing.

The matter of transaction failures needs special consideration. In that case, it is assumed that a typical user seeks confirmation, e.g., by calling or messaging the recipient (i.e., using an external means of communication). In many cases, the recipient also issues a receipt confirming incoming payments. The sending party might wait for that statement for a certain time period, and enquire about one from the recipient if none is received.

In any case, in particular in testing modes where party A has no direct visibility of events on party B (this issue is also discussed in clauses 11, 13, 15 and 16), reasonable and appropriate measures and conventions, adapted to the actual scope and goals, shall be considered and set-up as part of a testing campaign.

6.5 Automation of tests

The methodology in this Recommendation describes generic testing, i.e., service tests can be done manually as well as automatically. It is understood that automation of tests is desirable to achieve a greater degree of repeatability, and less variation in quantitative data values due to inaccuracy, e.g., of manual time measurements.

Automation can have different forms with respective degrees of automation up to fully automated testing. Using the multi-stopwatch (MSW) concept as described in this methodology is the next step of evolution, significantly improving the robustness of testing with respect to manual event logging.

The next step may be to still use manual operation of transactions but to record low-level activities on the DFS device itself, e.g., from recording of layer 3 messages or Internet protocol-level activities. The ultimate goal is a system that executes all DFS processes automatically. Such implementations

require, at least, an extended level of access to platform devices ("rooting", e.g., having system-level access) and substantial technical efforts, in particular because testing such systems requires actual use of DFSs. Making these systems fit for unsupervised operation or to enable operation on a larger number of mobile device types, would further increase the necessary effort.

Design or implementation of further steps of automation, or related evolution of the methodology lies outside the scope of this Recommendation.

7 Transaction model and digital financial service key performance indicators

The basic model is identical to that described in [ITU-T P.1502]. This clause describes the special variant for inter-operator and cross-border testing where teams do not have direct contact with each other.

In the following, the basic considerations and principal definitions from [ITU-T P.1502], as well as extensions, are described. For a more thorough background reading, as well as for the full set of DFS KPIs, see [ITU-T P.1502].

Figure 2 shows the basic structure and event flow of the DFS implementation; the collection of required details ("Collect transaction data") is shown summarily; the details are different between operators.

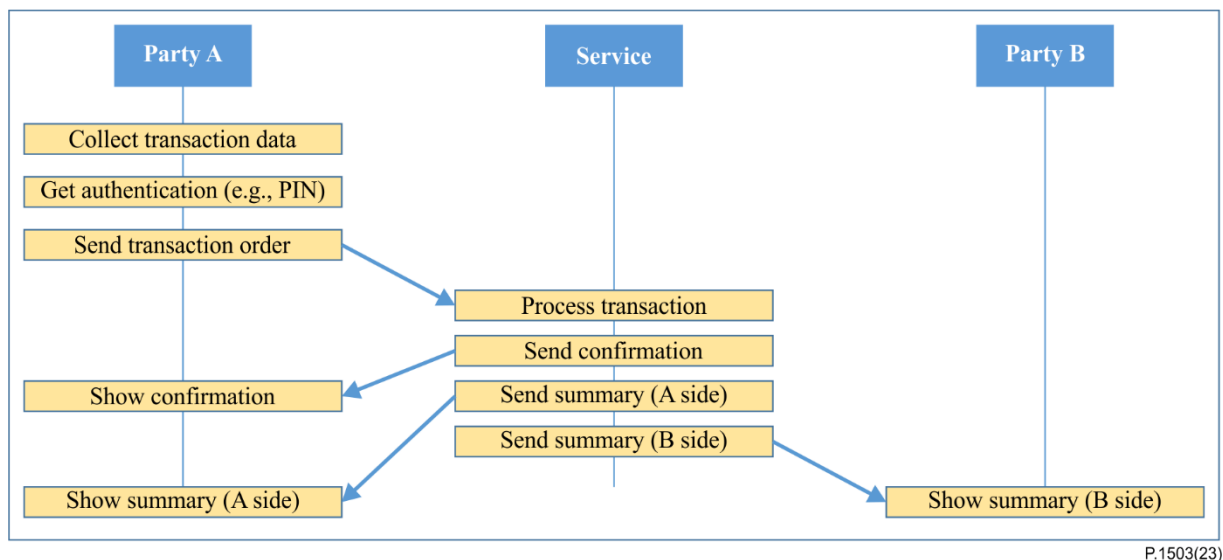


Figure 2 – Basic model of a P2P money transfer implementation

Figure 3 shows how the action and event flow is mapped to the "timer flag" elements that are then used to compute the KPIs. Again, the basic processing is the same as that for the basic P2P case.

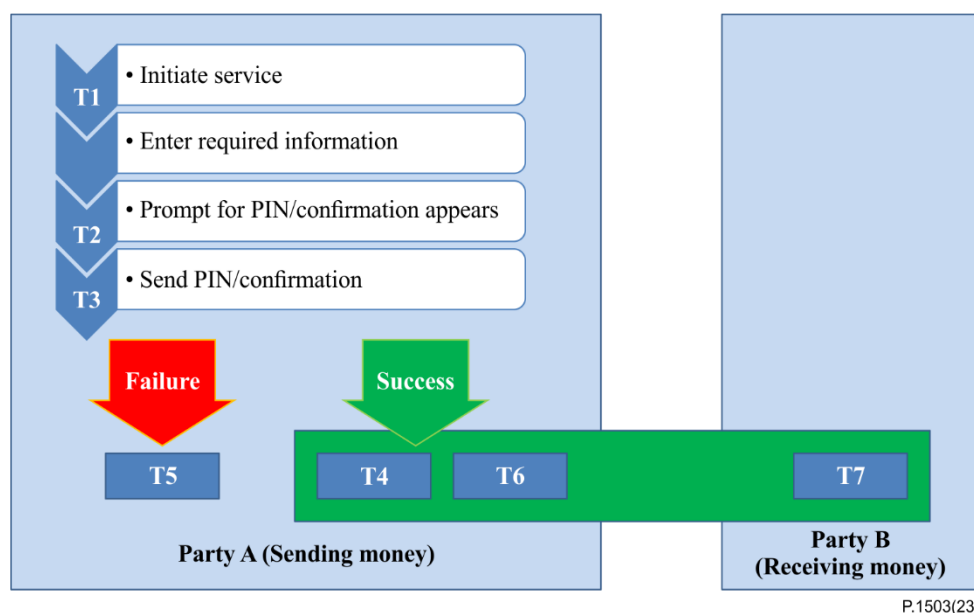


Figure 3 – Connection between events and timer flags (see text for details)

T4 is the primary (in-application) success criterion while T6 refers to the success criterion provided by receipt of a notification short message service (SMS). Please note that in this generalized case, T4 and T6 can appear in any combination and order.

The principal difference to an intra-country, intra-network is that most likely teams are working in different locations.² Therefore, not all of timer flags are recorded in the same place. The practical consequence is that data has to be combined from different sources, which is described in more detail in clause 15.

It is important to keep in mind that due to the different paths for events T4, T6 and T7, they can appear in any order.

Due to the different types of implementation, it is possible that on the party A side, either T4 or T6 is missing. For data processing, this means that transaction success is indicated by either criterion.

The KPIs used are a subset of the simplified set defined in clause 10 and Table 5 of [ITU-T P.1502].

NOTE – The event names are those specified in this Recommendation; the set differs from that in [ITU-T P.1502].

² [ITU-T P.1502] assumes that both the party A and B devices are in the same place. Of course, even for national, intra-network testing it is possible that operation in different places is desired. Therefore, this Recommendation provides an extension to [ITU-T P.1502] even in this basic case.

Table 1 – Simplified set of key performance indicators used in this Recommendation, based on the set in [ITU-T P.1502]

Indicator	Abbreviation	Definition/Remark
Money transfer core duration [s]	MTCD	T4-T3
Money transfer raw completion time [s]	MTRCT	T4-T1
Money transfer completion rate [%]	MTCR	T1 present, T4 present: success (see remark 1)
Money transfer full completion time [s]	MTFCT	T7-T1: Not used when asynchronous mode is used
Money transfer A-side completion time [s]	MTACT	T6-T1 (see remark 2)

It is assumed that some data cleansing has to be applied here. With manual operation of the test case, input errors may occur. In data recording, a way to mark such transactions as "to be ignored" should be provided. When KPIs are calculated, such transactions need to be excluded from the set because they may cause distortions to other KPIs (e.g., longer response times in the case of non-existent account numbers or personal identification number (PIN) errors).

In addition, it is recommended to align KPI calculation with known general network problems. For instance, if there is a mobile network outage during a test, respective transactions should be ignored in KPI calculation also.

There is also the possibility that a transaction is reported to the party A as successful but actually is not. As part of data cleansing, the actual arrival of money should be checked (e.g., as a bulk action in post-processing) and appropriate corrections applied in case. This can, for example, be done by setting respective transaction attributes, or – in the simplest case – invalidating T4 and or T6. These procedures need to be aligned with the specific implementation of evaluation and KPI processing.

Remark 1: When a test campaign is designed, a decision should be made about whether the MTCR is reported at all, based on an assessment if the test is seen as representative.

Remark 2: If T6 is used as surrogate for T4 and T4 is missing, MTACT is typically not computed as there would be partial overlap with MTCD and the sample basis is smaller in any case. If computation of MTACT is nevertheless required, it needs to be made clear in the accompanying documentation (project report), that this is a secondary or auxiliary KPI.

Special consideration is required for the case of MTCR, as the reported completion rate depends on the number of unsuccessful transactions. There might be cases where either the network or the DFS infrastructure is temporarily down, i.e., a part of the service is systematically unavailable. MTCR is therefore not only a technical element, but also a matter of testing perspective and scope of a testing campaign. It needs to be clearly specified and documented how these cases are treated, i.e., if and which transactions shall be removed from the valid set.

In the case of a campaign having an explorative character or seeking a broad perspective, a solution might be to report different variants of MTCR to show the corridor of values, depending on respective decisions.

8 Creating the use case model from actual use case examples

This Recommendation does not prescribe explicit ways to create use case models. However, Appendix I provides a best-practice example of use case models from existing implementations.

9 Data sources

In order to compute DFS KPIs, relevant input data need to be collected.