FIGI FINANCIAL INCLUSION GLOBAL INITIATIVE

FIGI Symposium
18 May - 24 June, 2021

Methodology for Assessing QoS/QoE KPIs for Digital Financial Services

figi.itu.int
#financialinclusion

Organized by

Committee on Payments and Market Infrastructures | BIS | THE WORLD BANK IBRD • IDA | ITU

What is meant by quality of service key performance indicators for digital finance and how relevant is this for emerging economies?

- KPIs are indicators that provide a simple way to describe how well a service works.
- Their purpose is to help in making such services better.
- The process of standardization means to share knowledge and experience from situations where such know-how has been successfully used. This also helps to avoid re-inventing the wheel.
- Secondly, standardization makes performance comparable, between providers or countries. Also, this gives a realistic picture of what can be achieved.

How will the work on measurement of quality of service key performance indicators for digital finance help telecom regulators and DFS providers?

- KPIs act like a compass, guiding regulators and service providers to improvement of those services.

- Also, the methodologies that come with them, contain best practice on how to measure those KPIs. This helps to make the process of measuring those KPIs efficient, i.e. cost and time saving.

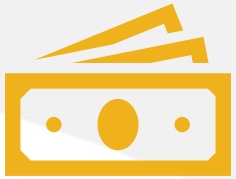
Overall DFS performance is the result of two subsystems: Mobile network and DFS infrastructure

	Poor DFS specific infrastructure performance	Good DFS specific infrastructure performance
Poor mobile network performance	Poor DFS QoS	Poor DFS QoS
Good mobile network performance	Poor DFS QoS	Good DFS QoS

- Both subsystems need to have good performance for overall good results
- Poor performance in one system may be hard to detect if the other system does also not perform well (risk of “evasive finger pointing”)
- Overall good quality can be achieved easier when respective regulators co-operate

Why and how should telecom regulators assess the quality of service key performance indicators for digital finance in their country?

- Financial issues are essential for every economy. Digital financial services are – or will become soon - a significant part of finance.
- When these services are provided through mobile networks, their performance depends also on the coverage and performance of these networks. A government will therefore have a strong interest in ensuring high quality of service of these networks.
- Telecom regulators are the institutions which provide the means to that end because they ensure that consumers are receiving good quality services from the networks.
- **The "how" will be subject of this session.**



Overview

- First, we will show some results from the two pilot campaigns we have been performing so far
  - Only the most necessary basics: Definition of KPI
  - Some results, pointing out just the major take-aways
- Then, we will explain how it was done, i.e. the metrics and methodologies.

Field-validate methodology and practical guidelines in a limited DFS campaign

Pilot Campaign Overview

Measured in a total of 78 locations in the larger area of Accra, Ghana

Measurements were made Mid-June to Mid-July 2018

Use case: P2P, most common use case to date

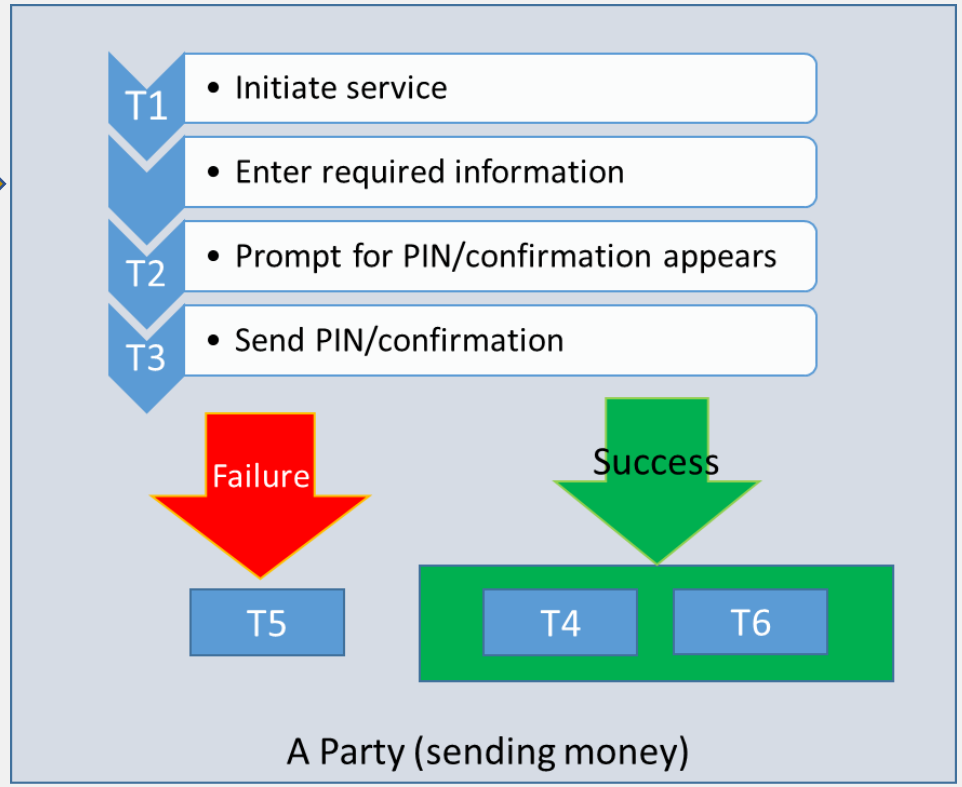
2 teams, each with 4 devices for DFS tests plus 1 “observer device” doing background measurements

Using a “practical set” of KPI due to the characteristics of the campaign (manual testing and time-taking)

Typically, 48 transactions per location

Use case and core KPI for DFS, basic

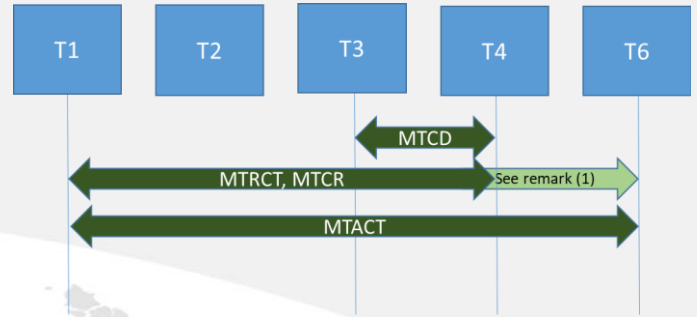
Single transaction



Indicator	Abbreviation	Computation with events used in this project
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
Money Transfer A-side Completion Time [s]	MTACT	T6-T1

Formal KPI definitions (for computation)

„timerflag“ definitions

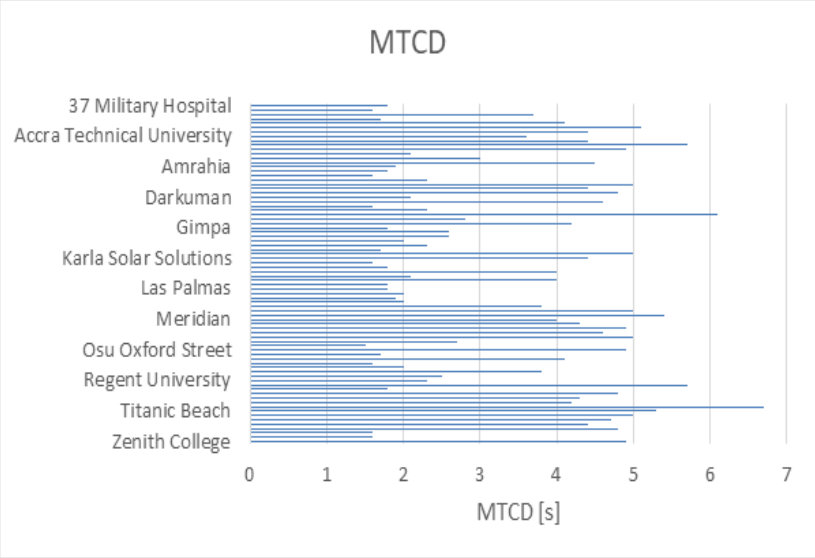


Mobile network KPI

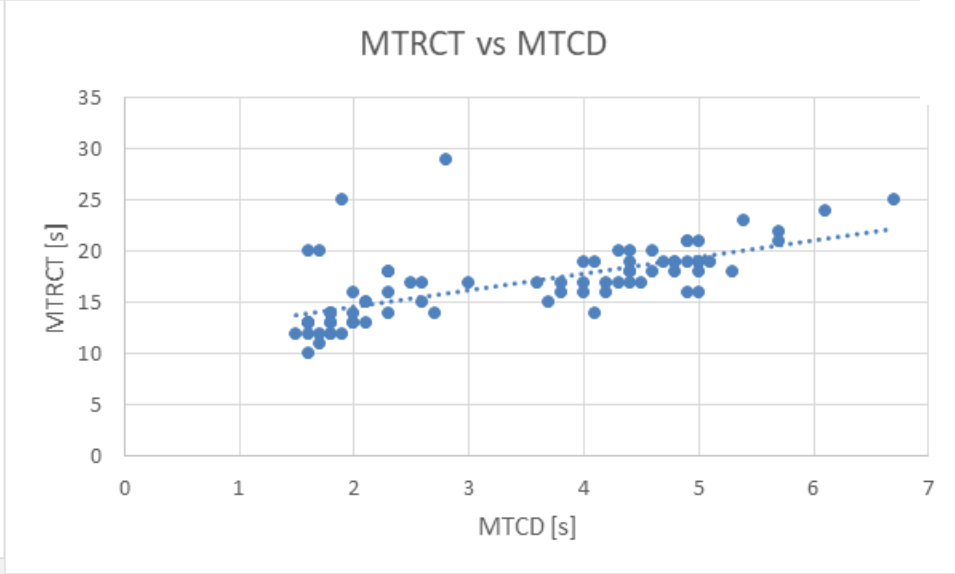
- **Data rate** (from download or upload test cases)
  - Transfer a fixed amount of data (typically a specially prepared file)
  - Measure the time for transfer
  - Data rate = $\frac{File\ size}{Required\ time}$ Unit: kbit/s or Mbit/s
- **Session time** (for web browsing test cases)
  - Access a web site (live web site, or specially prepared reference web site)
  - Measure the time it takes until the content is fully downloaded

Examples from Pilot No. 1: wider Accra area, multi-location

Each point is the average from 24 or 48 individual tests



Money Transfer Core Duration

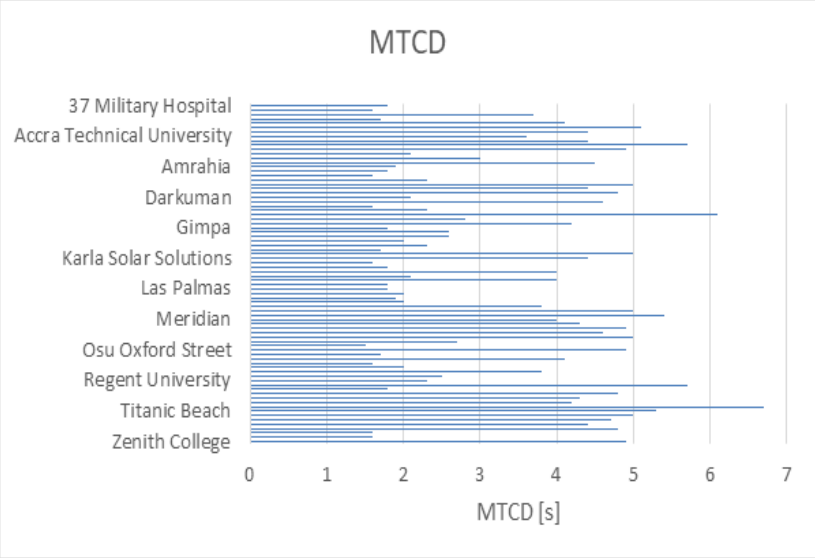


Money Transfer Raw Completion Time

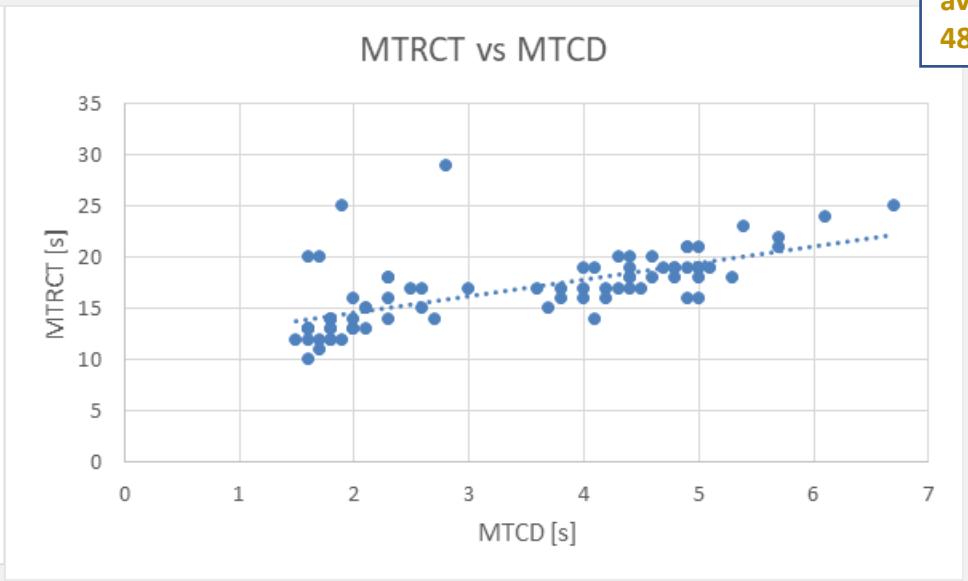
- Left: Variation of MTCD (transaction core time) over locations
- Right: MTRCT (raw completion time including manual operation) vs MTCD. This image shows a rather narrow distribution of manual execution times (only a few outliers).
- Each data point represents the average of all values at a given location.

Examples from Pilot No. 1: wider Accra area, multi-location

Each point is the average from 24 or 48 individual tests



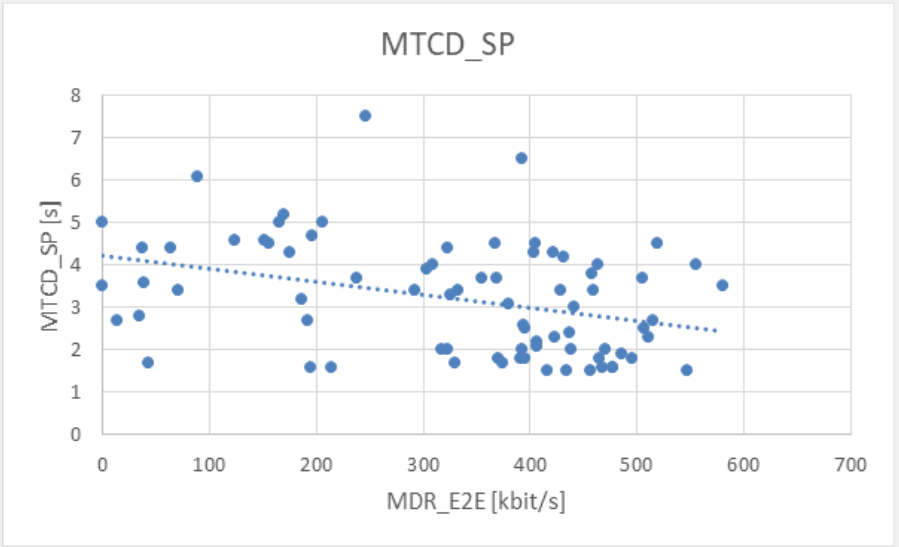
Money Transfer Core Duration

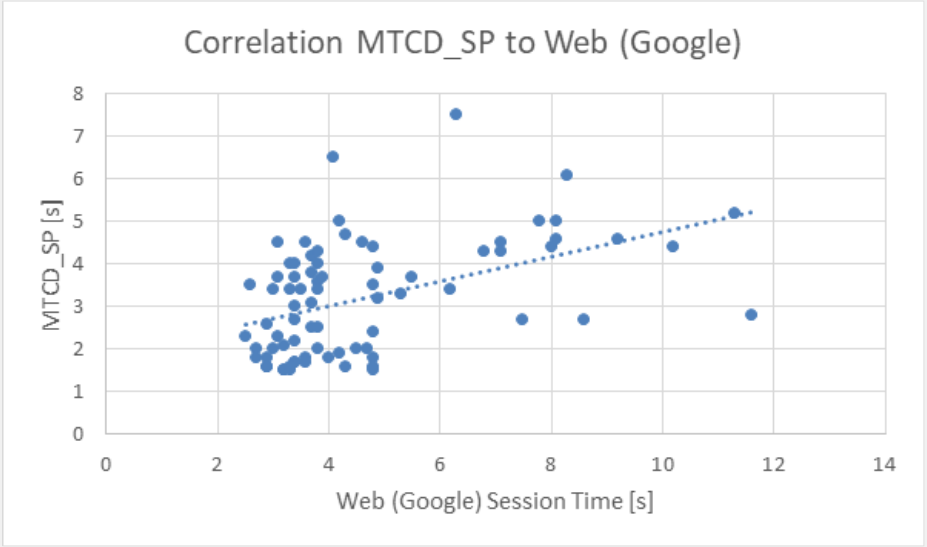


Money Transfer Raw Completion Time

- MTCD varies widely between locations, showing the influence of network coverage/performance
- MTCD only contains the system's response. MTRCT includes human action (entering required information for transfer, including also system response times). Values show good correlation nevertheless, demonstrating good performance of testers and high data quality in testing.

Examples from Pilot No. 1 (wider Accra area, multi-location)





- MTCD correlates reasonably well with network KPI for packet-data transfer performance, but there is considerable variance
- Possible reasons are fluctuations due to shared-medium use (other users in the same cell), or non-mobile network components

FIGI FINANCIAL INCLUSION GLOBAL INITIATIVE

figi.itu.int
#financialinclusion

The multi-country pilot for the extended methodology

Extensions to the P2P use case: cross-country, inter-operator

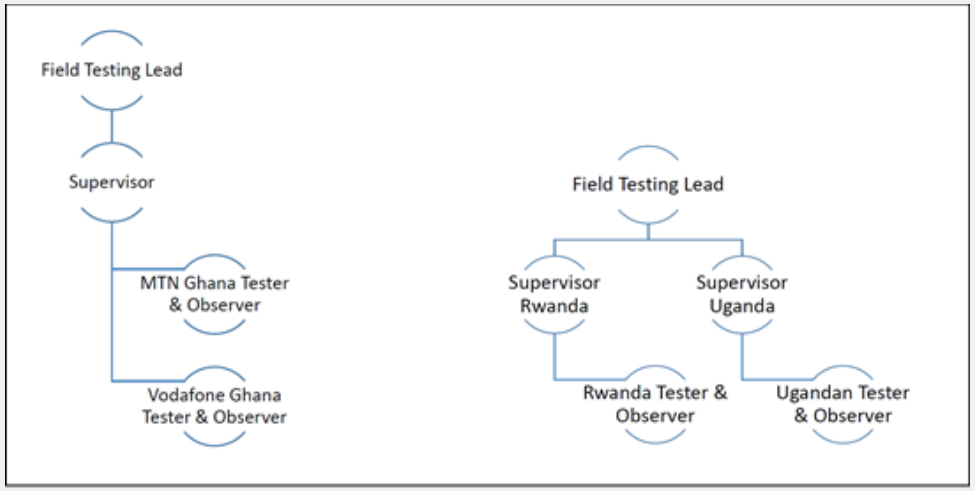
Extended Measurement methodology

- Asynchronous mode (as introduced in slide General flow of testing (asynchronous mode))

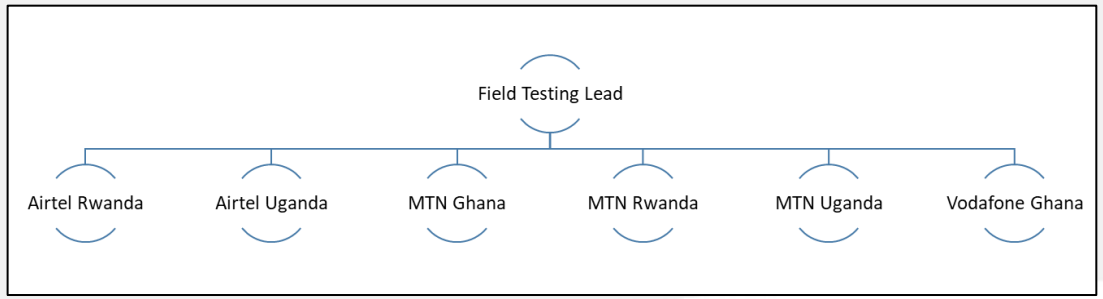
Additional restrictions due to the Corona pandemic

- “If life hands you lemons, make lemonade”: Chance to define and validate additional elements of robustness

Originally planned vs. actual structure



Originally planned structure:
4 teams, Rwanda/Uganda teams mobile
(same person in the Field Test Lead role)

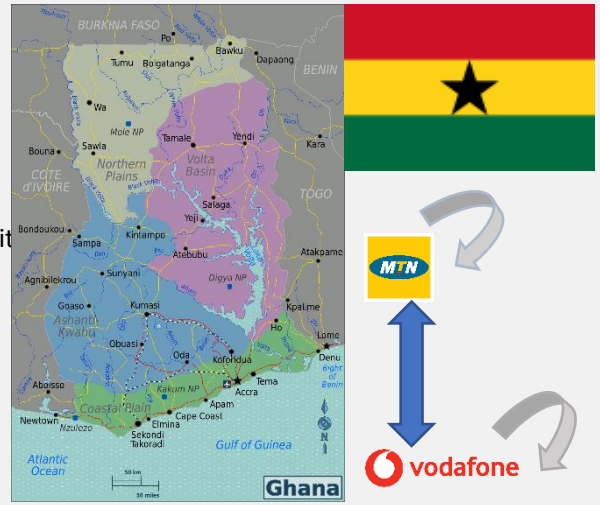


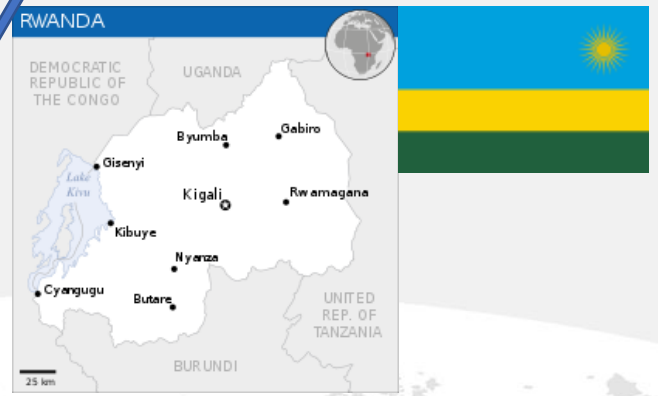
Actual structure:
6 teams, each in a fixed location

- Limited visibility of teams to each other
  - Communication bandwidth limitations
  - Testing at different times of the day

New DFS use cases at a glance - actual

- Image credit





Extended Methodology

Some cases not possible: services not supported

Extended Methodology Overview

Measured in a total of six locations;

two locations per country (Uganda, Rwanda and Ghana)

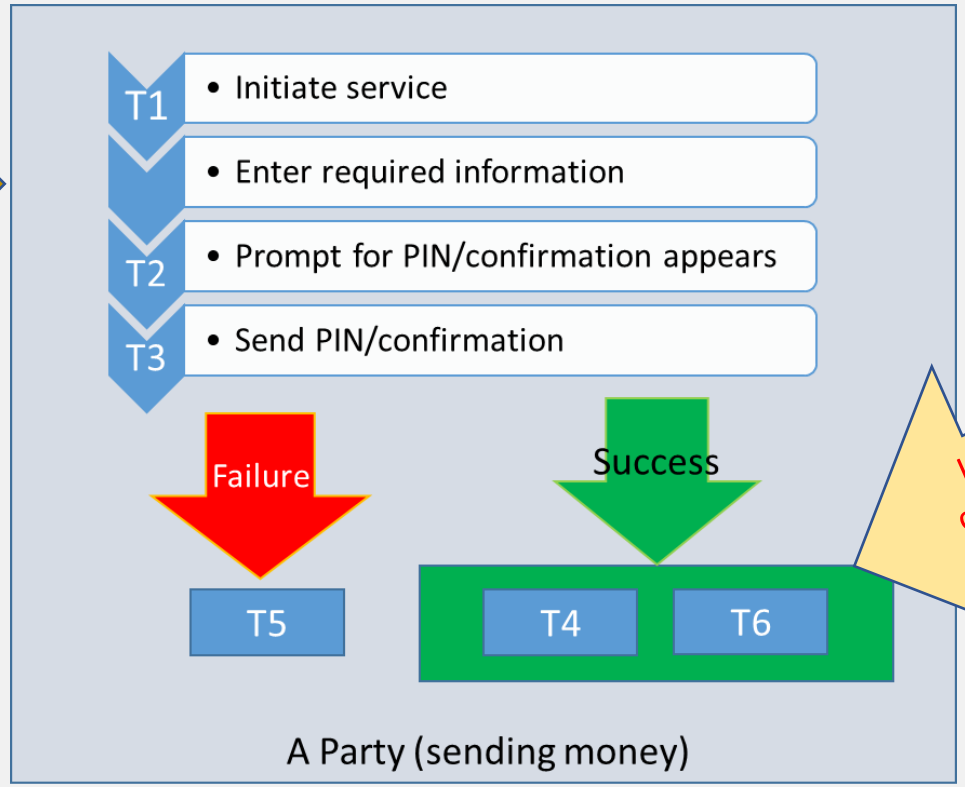
Measurements were taken May 2020

2 teams per country; each with 2 devices

Typically, 30 transactions per location

Use case and core KPI for DFS, multi-implementation

Single transaction

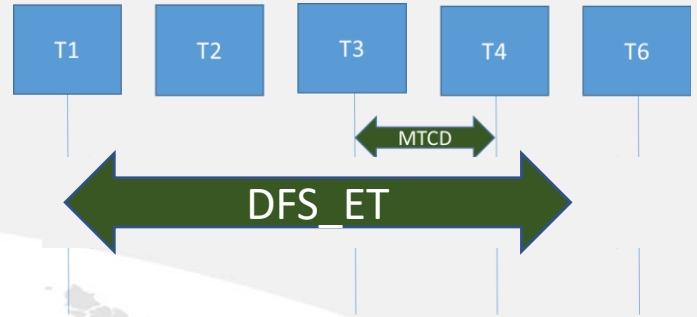


Indicator	Abbreviation	Computation with events used in this project
Money Transfer Core Duration [s]	MTCD	T4-T3
DFS Execution Time [s]	DFS_ET	{T4/T6}-T1
Money Transfer completion rate [%]	MTCR	T1 present, T4 or T6 present: success

Formal KPI definitions (for computation)

In some cases (different by country/operator): only T4 or only T6

„timerflag“ definitions



- DFS Execution time (ET) is a pragmatic indicator which has not been standardized yet

Results: Relation between DFS and network KPI

Scenario	DFS_avg_ET	DFS_Median_ET	DFS_SR	Web_SR	Web_ST_E2E	DL_SR	DL_ST_E2E	DL_MDR_E2E	UL_SR	UL_ST_E2E	UL_MDR_E2E
Cross-country Different network MTN Rwanda to Airtel Uganda	2,7	2	71,2	53,8	3,4	100	5,5	4,7	100	12,2	1,9
Cross-country Group network Airtel Rwanda to Airtel Uganda	8,7	7,4	100	50	5,2	94,2	13,1	2,1	98,1	8,3	1,4
Cross-country Group network Airtel Uganda to Airtel Rwanda	12,7	5,9	64,9	96,6	2,9	[illegible]	[illegible]	[illegible]	[illegible]	9,8	1,1
Cross-country Group network MTN Rwanda to MTN Uganda	1,9	1,7	61,3	19,8	8,1	80,3	7,4	3,5	93,7	19,3	0,4
Cross-country Group network MTN Uganda to MTN Rwanda	3,4	2,4	95,1	64,6	2,7	[illegible]	[illegible]	3,2	100	[illegible]	[illegible]
Inter-network Airtel Uganda to MTN Uganda	12,1	9	98,6	94,8	1,8	83,2	6,9	4,6	98,2	6,1	1,8
Inter-network MTN Ghana to Vodafone Ghana	2,8	2,3	91,8	42,9	4,6	79,3	14,1	[illegible]	100	[illegible]	[illegible]
Inter-network MTN Uganda to Airtel Uganda	2,5	1,8	91,4	47,6	3,6	93,3	8,4	3,5	100	9	1,1
Inter-network Vodafone Ghana to MTN Ghana	2,9	2,7	98,1	48,2	4,9	[illegible]	14,1	1,9	100	5,9	[illegible]
Same Network Airtel Rwanda to Airtel Rwanda	9,9	7,2	99,7	35,4	5,2	83,6	16,4	1,6	99,3	9,2	1,2
Same Network Airtel Uganda to Airtel Uganda	11,4	8,8	98,6	74,7	2,1	79,3	11,5	2,7	94,7	8,7	1,4
Same Network MTN Ghana to MTN Ghana	1,9	1,2	92,7	51,9	5,7	51,9	14,8	1,4	100	13,6	0,7
Same Network MTN Rwanda to MTN Rwanda	5,6	1,9	95,3	58,6	3	83,3	9,4	3,6	88,6	7,2	1,7
Same Network MTN Uganda to MTN Uganda	1,8	1,6	92,1	51,2	2,9	93,2	8,5	3,3	100	10,8	1
Same Network Vodafone Ghana to Vodafone Ghana	4,9	4,7	97	69,3	6,5	61,1	17,1	1,3	100	8,2	1,3

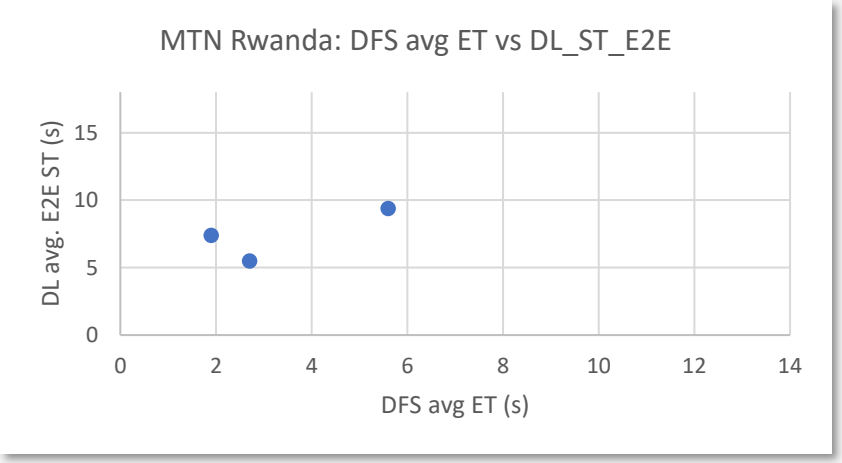
Multi-country/Cross-country and inter-operator test cases let us see the differences in performance between operators, and the margins for improvement.

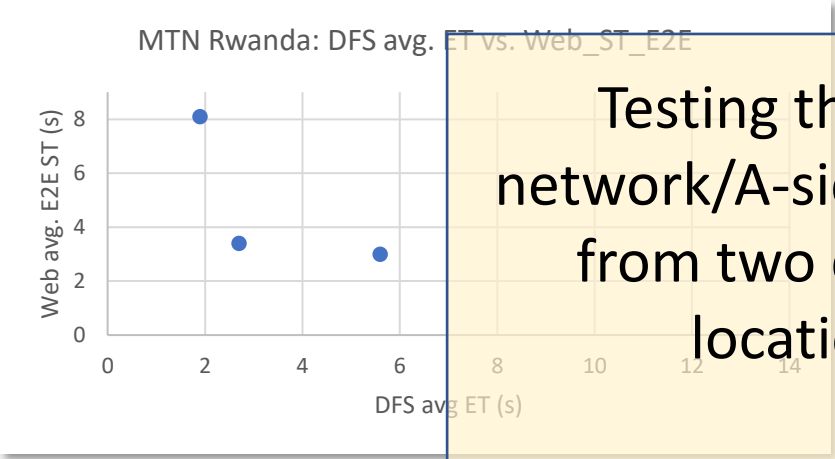
For instance, in some cases better performance was achieved in cross-country or inter-operator scenarios than in intra-operator/country cases.

Example of local mobile network performance having a strong effect

Configuration	Owner Team	Scenario	nSuccess	nTA	avg_ET
MTN Rwanda	Airtel Rwanda	Same Network MTN Rwanda to MTN Rwanda	37	41	14,4
MTN Rwanda	MTN Rwanda	Same Network MTN Rwanda to MTN Rwanda	85	87	1,7

Pilot 2 (cross-country/inter-operator)





Testing the same network/A-side operator from two different locations:

Clear visibility of network coverage/performance related effects in DFS QoS

- The configuration „MTN Rwanda“ has been tested by **two** teams, i.e. from two different locations)
- Unusual large spread of values also visible for other use cases
- Results show that location/network performance **does** have an effect

FIGI FINANCIAL INCLUSION GLOBAL INITIATIVE

figi.itu.int
#financialinclusion

Methodology: Essential points

1

QoS should reflect the user‘s point of view: Design of KPI

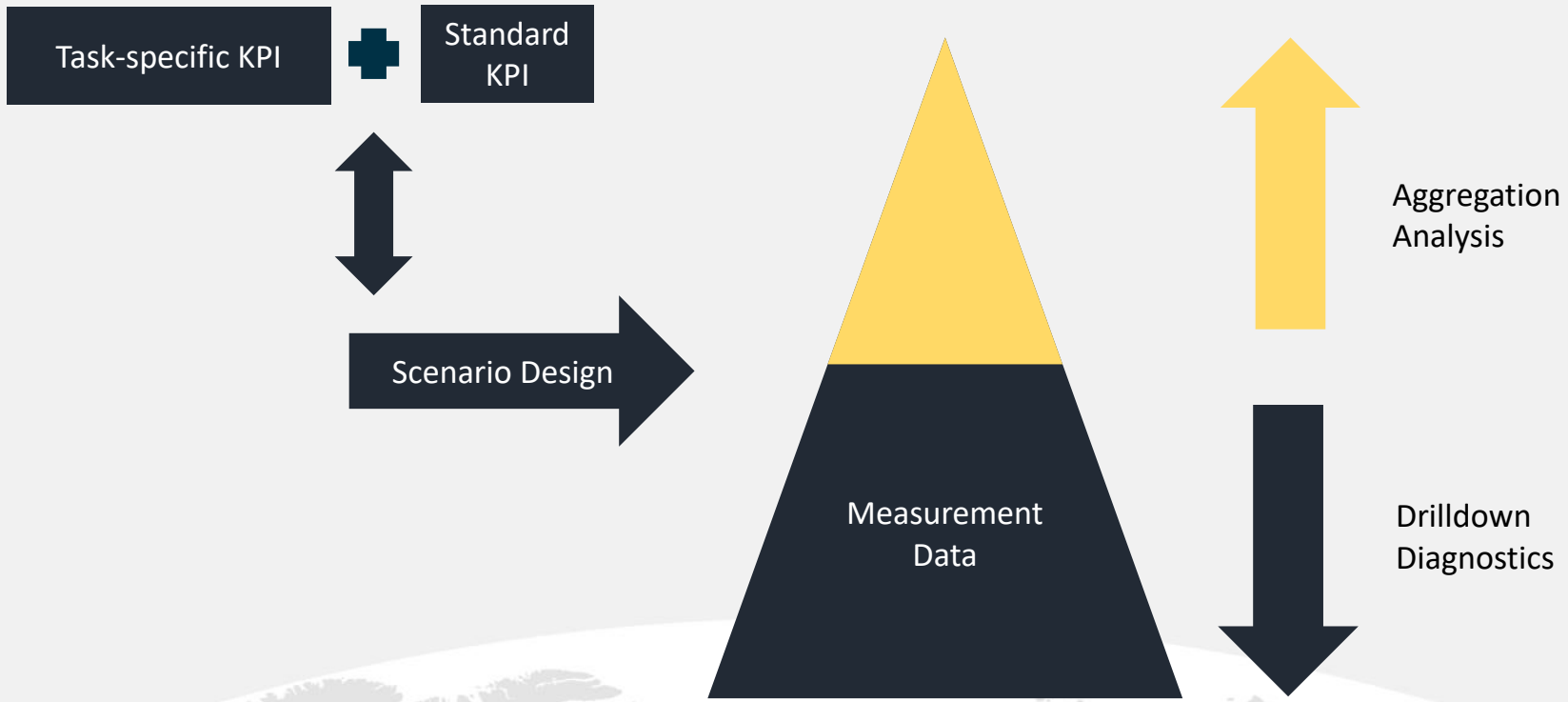
2

Data quality and comparability of results: Standardize how the measurements are done

3

Efficiency and value-for-money: Optimize yield

QoS = Economical Value or Value for Society



- A well-designed set of KPI expresses the value of a service

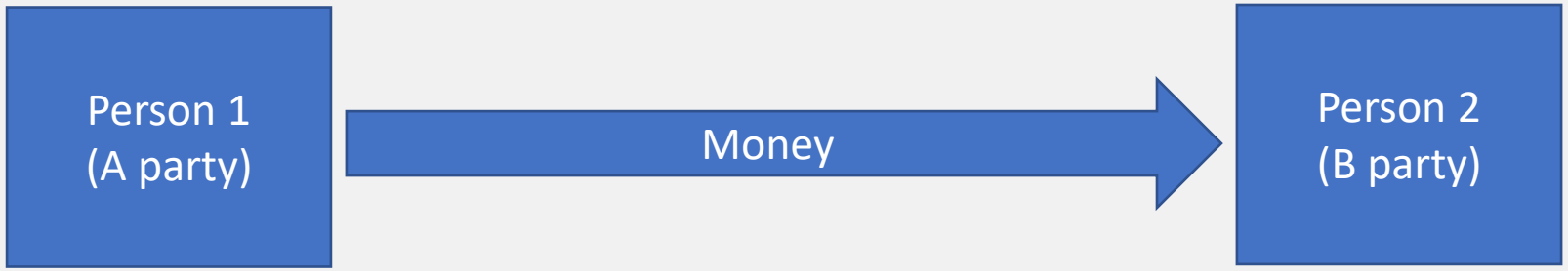
QoS/QoE KPI should be simple

- Typically, end to end customer perspective can be expressed in simple terms
  - The „Philadelphia test“: It should be possible to explain the concept to a five-year old
  - Use a small number of powerful KPI rather than too many of them

Defining the use case

- What is the essence, i.e. the core function and sequence of events?
  - Actors/parties („entities“)
  - Stages (phases)
  - Activities/events („relations“)
    - Events visible to the „common user“
      - (there may be other events, e.g. signalling-layer events, which are not perceivable directly. In QoS, try to avoid such events as far as possible!)

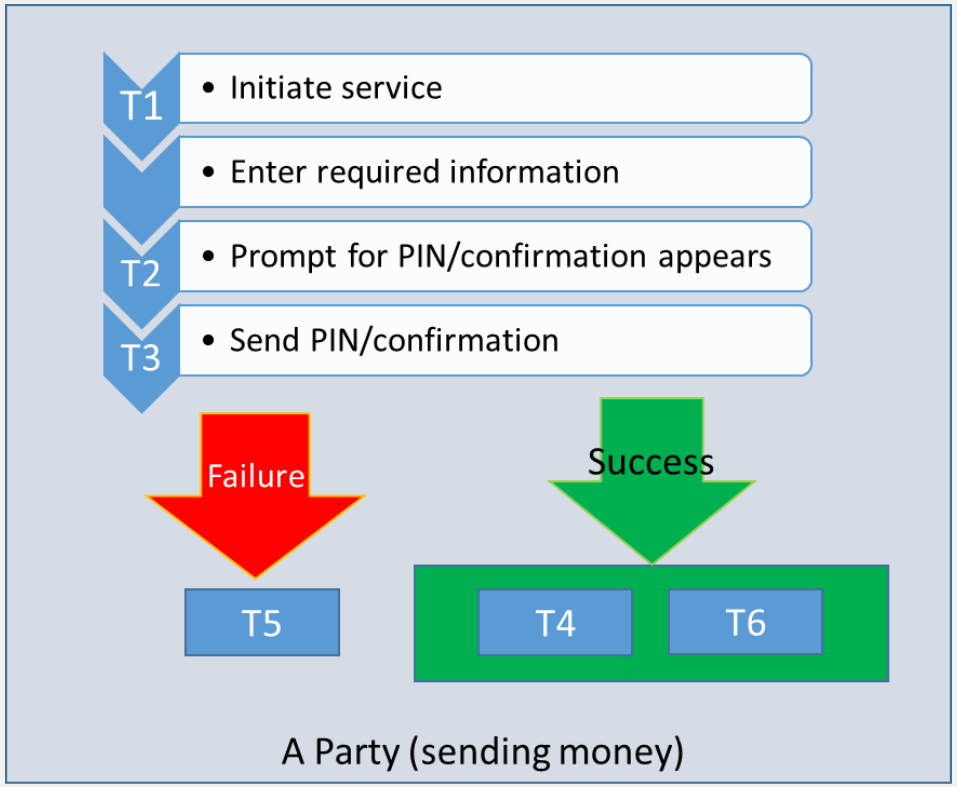
Person-to-person money transfer

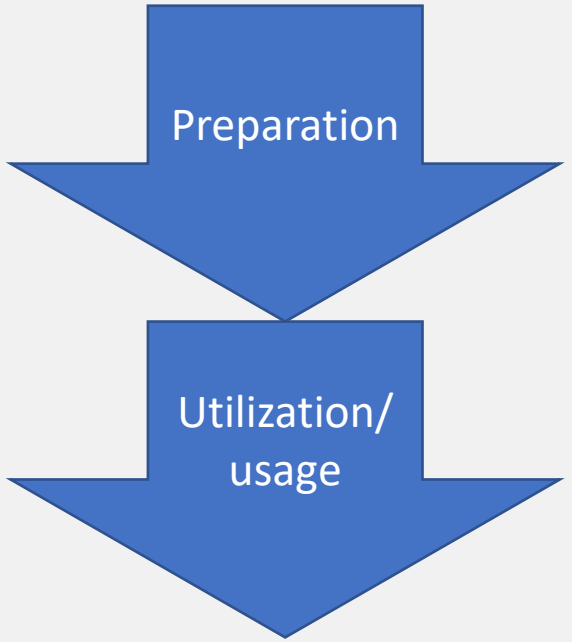


„Purpose“ plane: A sends money to B

„Functional“ plane: A service or system is used

Phases

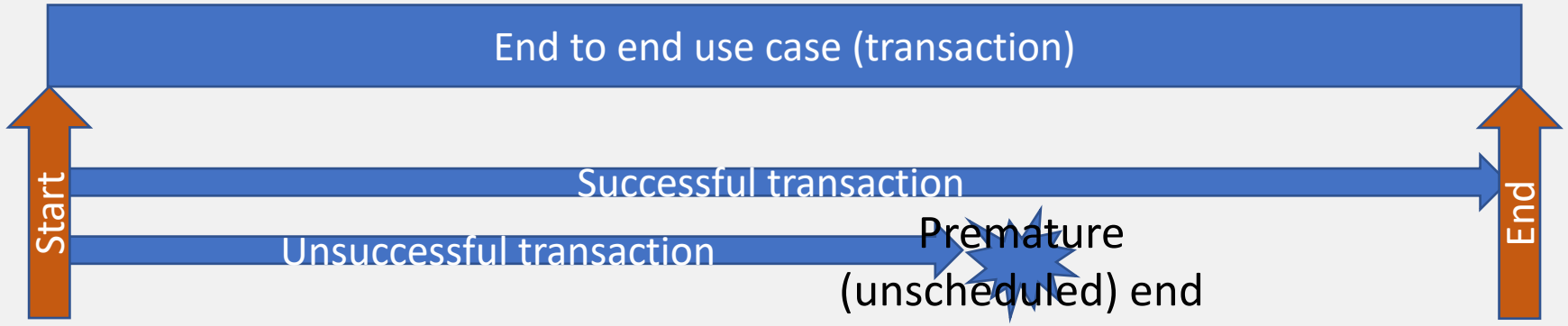




Telephone call:
- Dial/set up the call
- Talk: Use the connection

Money transfer:
- Enter required information
- Do the transfer

Formalized (essential) Transaction



- Top-level „end to end“ view
- Events are meaningful, observable entities (points in time). Dual use:
  - Progress markers. Success = „End“ event observed
  - Timers. Execution time = time of „End“ minus time of „Start“

Use case and KPI modelling – what to observe

- Assume the service is a „black box“: Design generic models
  - Using knowledge about the inner working of the service is tempting (more information or easier testing) but can be dangerous (creating wrong results) if implementation details change

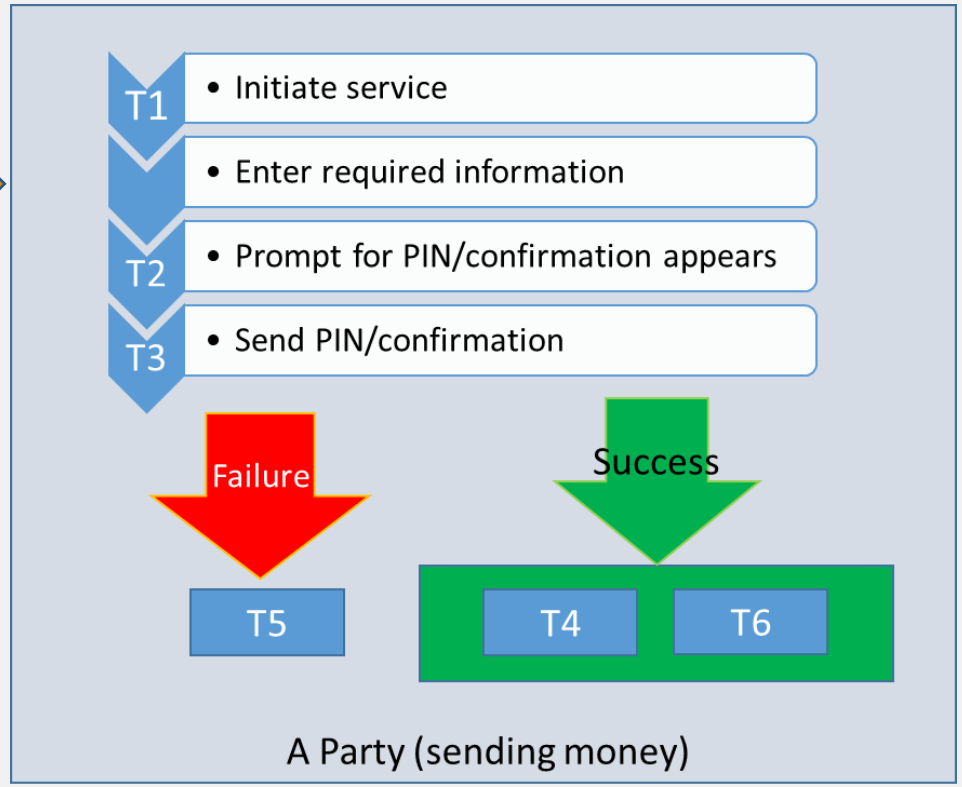
- Use „positive“ success criteria
  - Find a clear definition of a successful transaction. When using negative definitions, there is the risk of blind spots and artefacts (paths leading to non-success which have not been thought of before)

FIGI FINANCIAL INCLUSION GLOBAL INITIATIVE

figi.itu.int
#financialinclusion

Use case and formal definitions (core KPI)

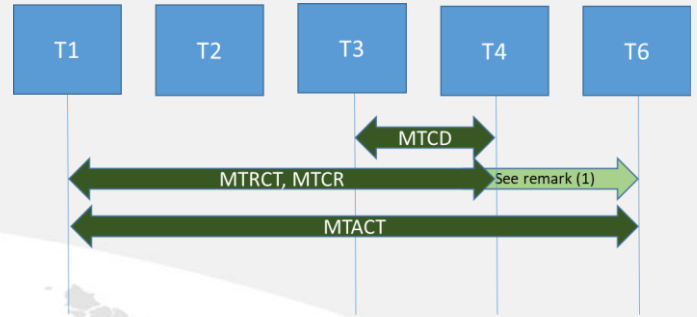
Single transaction



Indicator	Abbreviation	Computation with events used in this project
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
Money Transfer A-side Completion Time [s]	MTACT	T6-T1

Formal KPI definitions (for computation)

„timerflag“ definitions



Methodology

Define how a measurement is done

- Efficiency: Optimal usage of devices/minimization of hardware effort
- Repeatability/reproducibility
- Documentation
- Data quality and yield of testing
- Tool support

Prepare

- Training
- Define how to respond to a wide variety of situations

Special requirements of money transfers

Plan for a cyclical transfer (minus fees) but prepare for lost transfers (top-up capability)

Make sure that the service access is undisturbed

- Sufficient credit for data transfers or other network usage (e.g., SMS if required)

Preparation and Training/Pre-testing phase

Make testers familiar with their tools

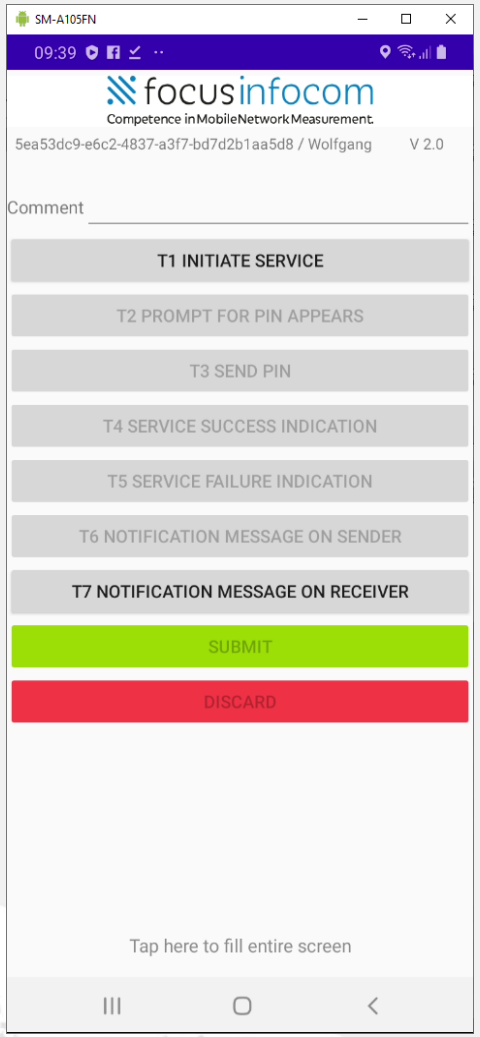
Are there limitations to handle?

- Is there a limit to the daily or hourly number of transactions?
- Is there a limit to the total fund transfer sum, e.g., a daily or weekly maximum?
- What can be done if a transaction fails (retrieval of funds?)

Are there fixed or temporal limitations of interoperability between networks or providers?

Tools

- Reduce margin of error and relieve testers from dull, repeating work
  - Fully automated mobile network performance testing
  - Automatic “multi-stopwatch: Take time of events by push of a button
- Reduce the risk that data is lost or compromised
  - Automatic upload of data, back-up strategies
- Support testers in staying alert
  - Use check lists to ensure operating conditions (e.g. make sure mobile device has sufficient battery charge, or DFS account has sufficient credit



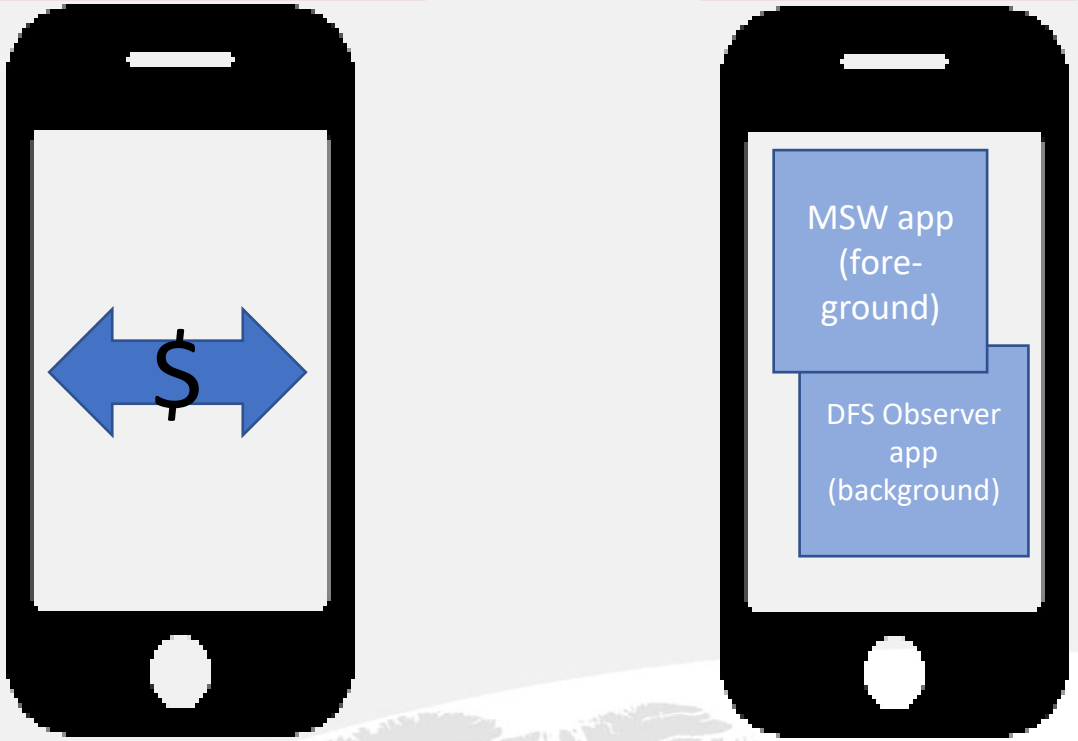
Test Setup and preparation

- Use log sheets/check lists
- Basic checks before Testing:
  - Is the Phone charged and can it be charged (wall plugs/power banks)?
  - Is there Data for uploading results?
  - Is there enough money on the wallet?
  - What is the signal strength?

Devices and roles: Basic scheme

Device for running the test case

Device for data acquisition

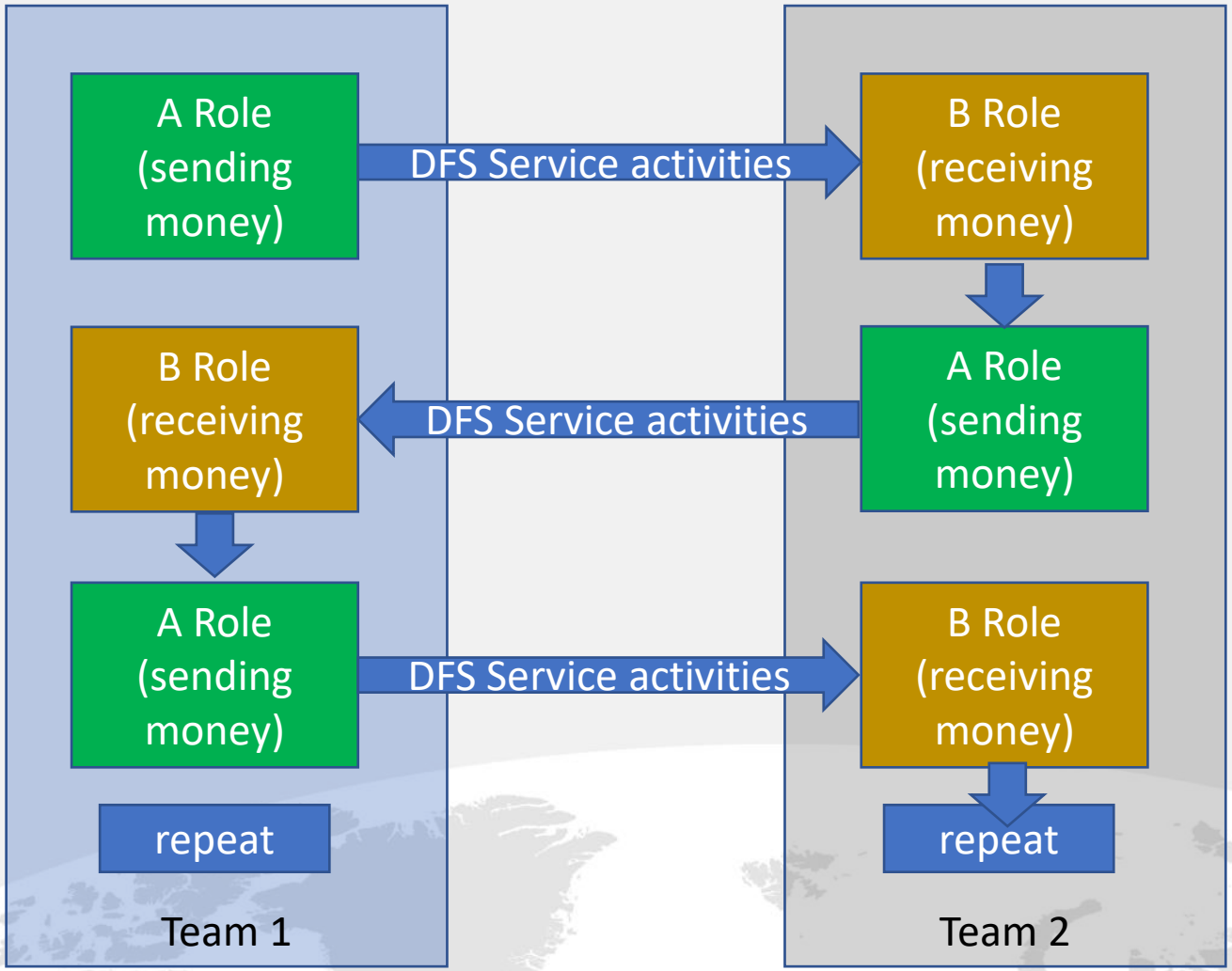


Multi-stopwatch app (time-taking)

Mobile network performance/coverage measurement

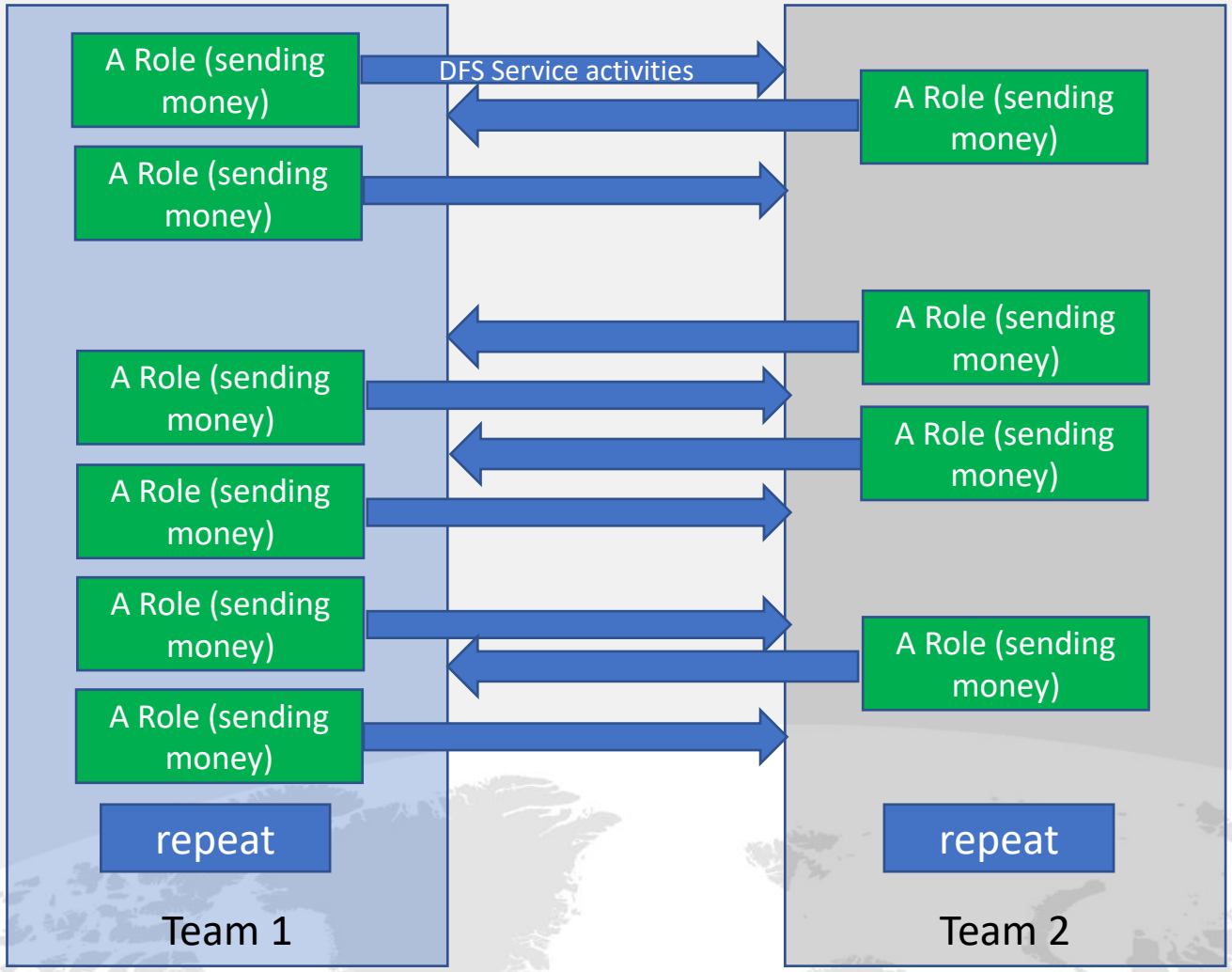
- Try to avoid frequent re-configuration of devices: Fixed allocation of functions/roles for devices

General flow of testing (synchronized mode)



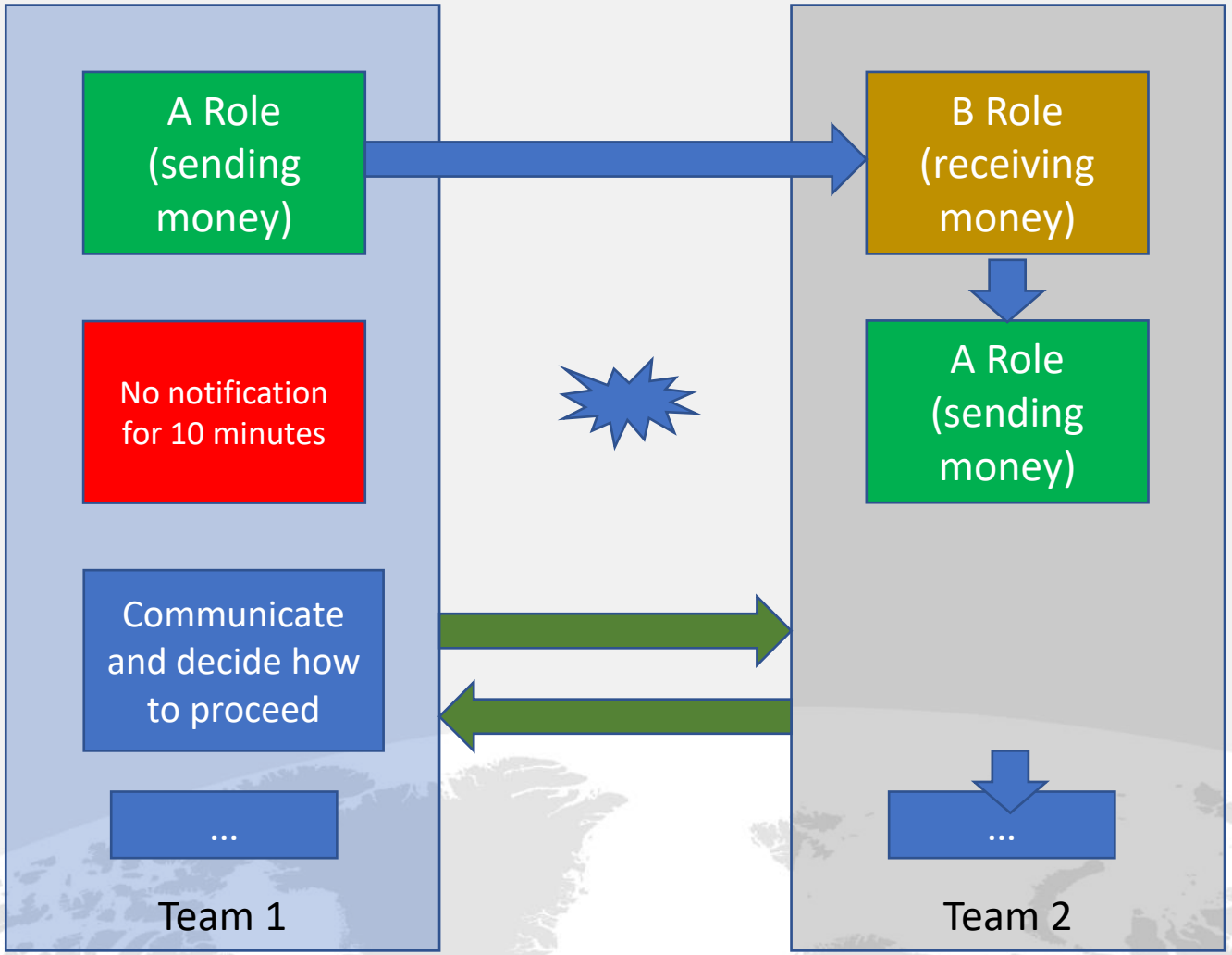
Working also if teams don't see each other directly

General flow of testing (asynchronous mode)

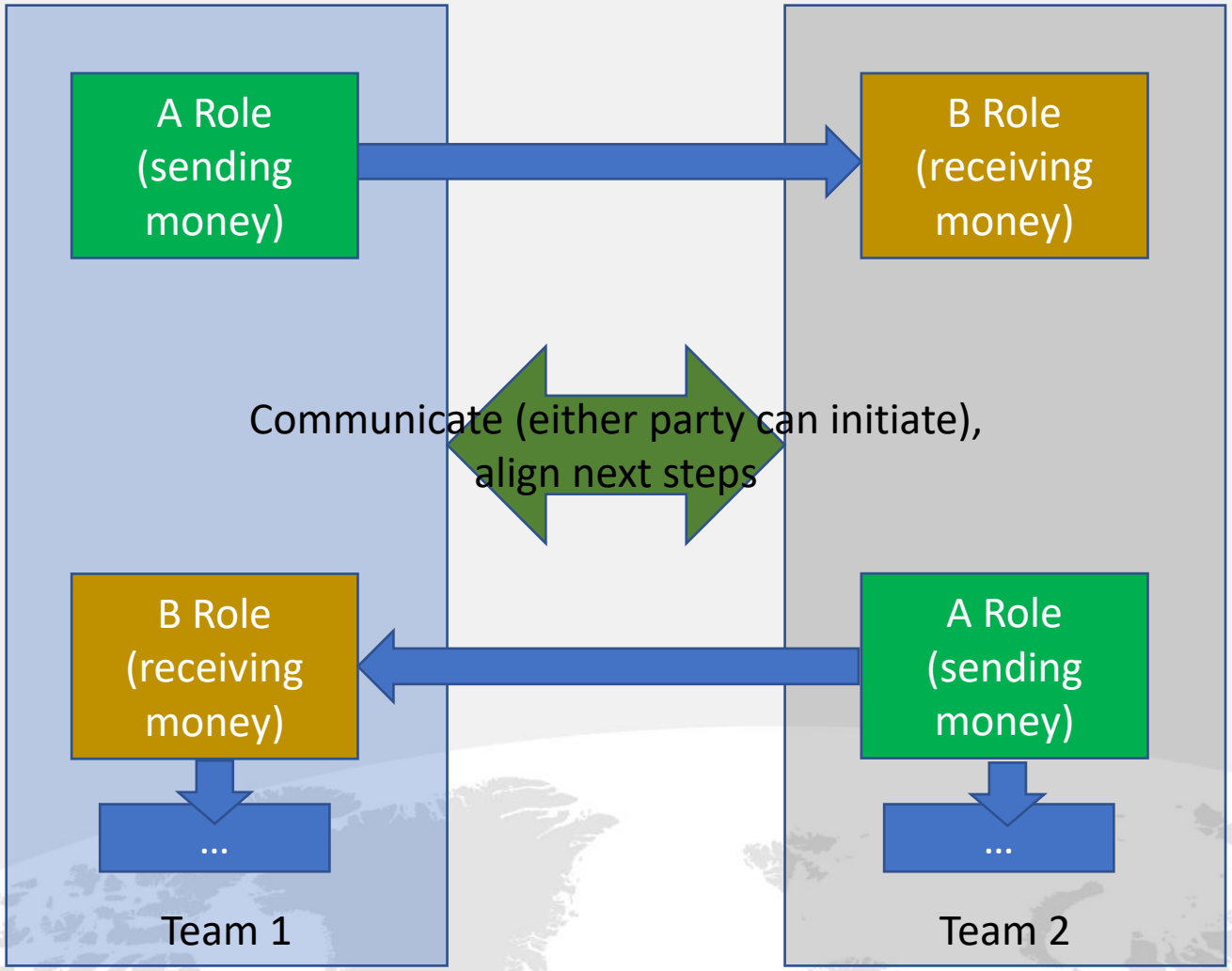


For situations where T7-related KPI are not needed, and/or teams work different hours or in very different speeds.
T7 is not recorded on B side.
Caution: may require larger buffer of credit on A/B side!

Robustness: Planned handling of unplanned interruptions



Planned interruptions/pauses



Wrap-up: KPI and Methodology

Standardize KPI and Methodology

- Ensure that performance indicators are comparable
  - Measurements of the same service: Monitor progress
  - Benchmark with other providers: Ensure to get best-in-class
- Utilize cumulated experience for best data quality and robustness

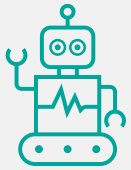
Lessons learned from each campaign contribute to further improvement of standards

FIGI FINANCIAL INCLUSION GLOBAL INITIATIVE

figi.itu.int
#financialinclusion

Ongoing evolution

Continuous improvement and enrichment of existing tools and methodologies from experience gained



Further automation of test and measurement

Create a repository of reference campaigns

Establish a meaningful set of target values

SG12: The place where QoS standards are created

- ITU-T Study Group 12 is the expert group responsible for the development of international standards (ITU-T Recommendations) on performance, quality of service (QoS) and quality of experience (QoE).
- This work spans the full spectrum of terminals, networks, and services, ranging from speech over fixed circuit-switched networks to multimedia applications over mobile and packet-based networks.

DFS related standardization in SG12

- ITU-T Recommendations G.1033 (“Quality of service and quality of experience aspects of digital financial services”) and P.1502 (“Methodology for testing of digital financial services”) have already been approved and are in force.
- The extended methodology that has been described in this session is on its way into an upgraded ITU-T recommendation.
- SG12 has adopted a new Question (Q20) which will study and develop other standards related to DFS QoS.

Questions to the DFS Regulators

- How do the implementations shown match the one in your country? Does the modelling apply?
- How do the KPIs match your requirements?
- Are the insights matching the ones you already have?
- What would be modifications/extra requirements to optimally cover the situation in your country?

End of the tutorial – next: Q&A and panel disscussion

FIGI FINANCIAL INCLUSION GLOBAL INITIATIVE

FIGI Symposium
18 May - 24 June, 2021

Thank you for your attention. Questions?

Contact:

Kwame Baah-Acheamfuor
kwame.baah-acheamfuor@moc.gov.gh

Dr. Wolfgang Balzer
Wolfgang.balzer@focus-infocom.de

Fiona M Kamikazi Beyaraaza
fkamikazi@ucc.co.ug

Organized by

Committee on Payments and Market Infrastructures | BIS | THE WORLD BANK IBRD • IDA | ITU

figi.itu.int
#financialinclusion

FIGI FINANCIAL INCLUSION GLOBAL INITIATIVE

FIGI Symposium
18 May - 24 June, 2021

Back-up slides

figi.itu.int
#financialinclusion

Organized by

Committee on Payments and Market Infrastructures | BIS | THE WORLD BANK IBRD • IDA | ITU

Speakers

Kwame Baah-Acheamfuor

Ghana's Counsellor to ITU Council

Chairman, ITU-T Study Group (SG) 12: Performance, Quality of Service (QoS) and Quality of Experience (QoE)

Fiona M Kamikazi Beyaraaza

Senior Officer Telecom Compliance

Member of EACO Working Group 3 (QoS), and ITU-T S12 and SG12 RG AFR, focus on QoS and QoE testing

Dr. Wolfgang Balzer

Managing Director of Focus Infocom (Manufacturer of testing solutions for mobile networks)

Member of ITU-T SG 12, Specialist for QoS/QoE and testing strategies

FIGI FINANCIAL INCLUSION GLOBAL INITIATIVE

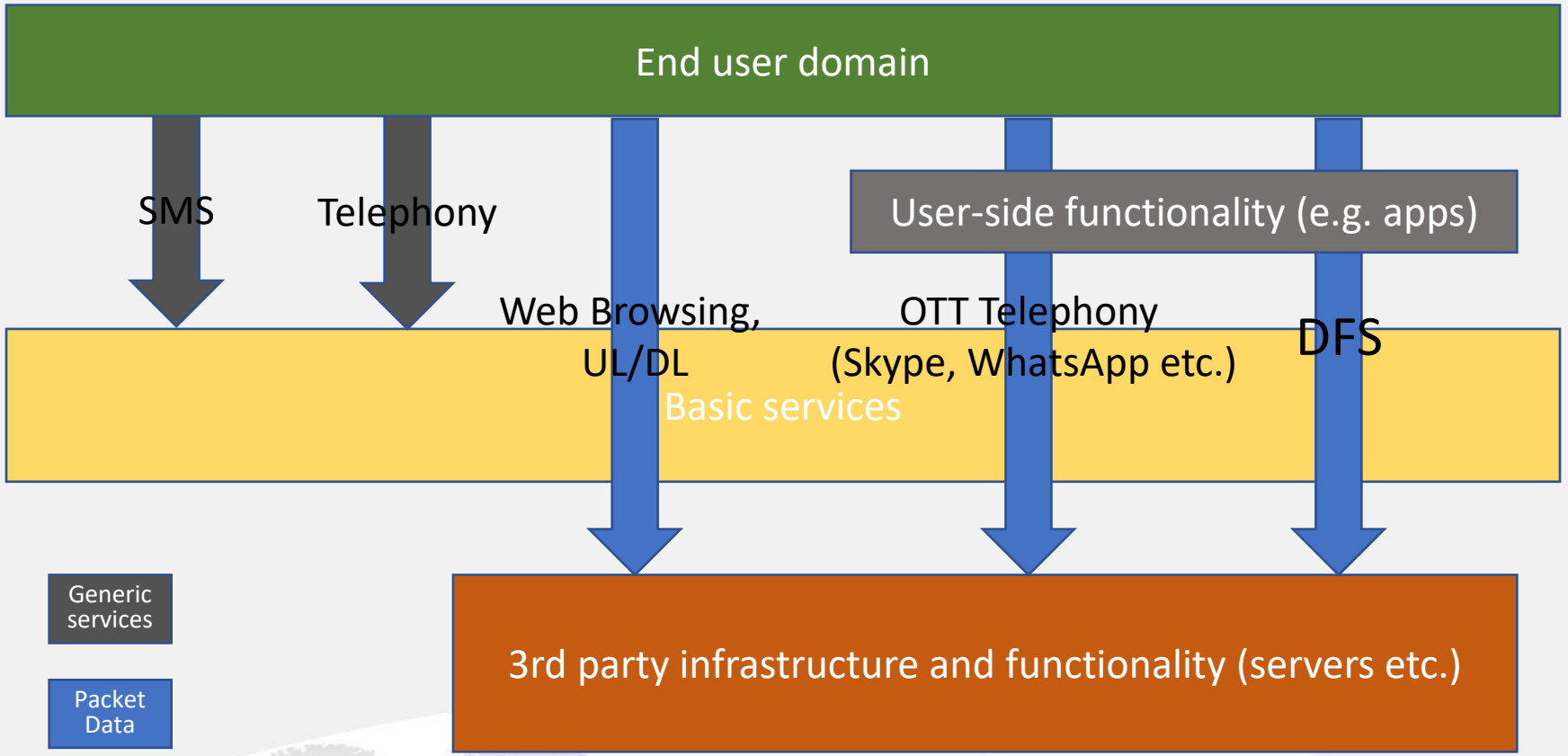
FIGI Symposium
18 May - 24 June, 2021

Image Credits

- Image credits:
- Ghana map: By Burmesedays, minor amendments by Joelf - Own work based on Perry-Castañeda Library Map Collection Ghana Maps, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=22745324
- Ghana flag: Public Domain, https://commons.wikimedia.org/w/index.php?curid=343073
- Uganda map: By OCHA, CC BY 3.0, https://commons.wikimedia.org/w/index.php?curid=34826076
- Uganda flag: By tobias - From the Open ClipArt Library website., Public Domain, https://commons.wikimedia.org/w/index.php?curid=433085
- Rwanda location map: By OCHA, CC BY 3.0, https://commons.wikimedia.org/w/index.php?curid=34826078
- Rwanda flag: By !Original:UnknownVector: Zscout370 - http://www.primature.gov.rw/component/option,com_docman/task,doc_download/gid,859/Itemid,95/, Public Domain, https://commons.wikimedia.org/w/index.php?curid=327857
- Logos:
- MTN Group: Von MTN - mtn.co.za, Gemeinfrei, https://commons.wikimedia.org/w/index.php?curid=37719378
- Vodafone Ghana: By Vodafone - Vodafone UK, Public Domain, https://en.wikipedia.org/w/index.php?curid=57428450
- Airtel Africa: By airtel - www.airtel.in, Public Domain, https://en.wikipedia.org/w/index.php?curid=30177516

Services and performance

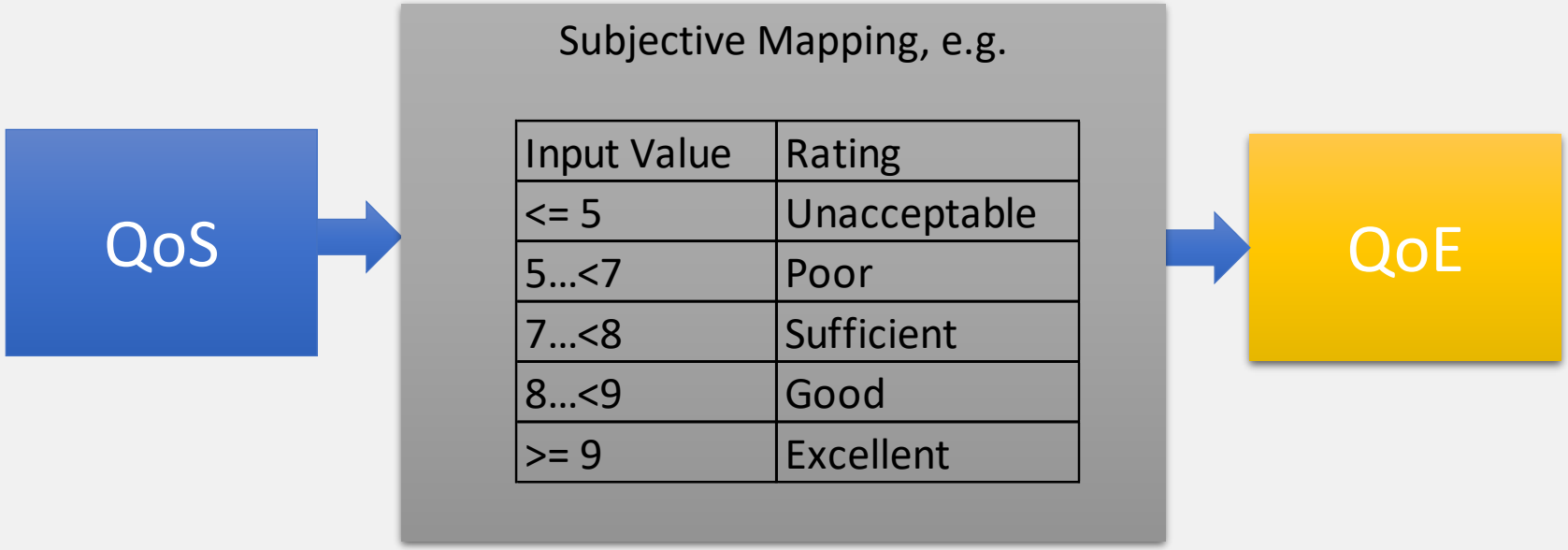
QoS and QoE: Basics



- A mobile network offers some generic services (e.g. telephony, SMS, basic packet data services). „Over the top“ offerings utilize these services as carrier services.

Practical definition of QoS vs QoE

QoS and QoE: Basics



- QoE can be determined from QoS by applying a mapping which contains elements from the opinion, experience, or expectation domain.

Full set of DFS KPI

Basic Methodology and KPI

Abbreviation	Type	Reference
MTCR	Rate/Probability	Money Transfer completion rate
MTCT	Time	Money Transfer completion time
MTFPR	Rate/Probability	Money Transfer False Positive Rate
MTFNR	Rate/Probability	Money Transfer False Negative Rate
MTFTRR	Rate/Probability	Money Transfer Failed Transaction Resolution Rate
MTASSR	Rate/Probability	Money Transfer Account Stabilization Success Rate
MTAST	Time	Money Transfer Account Stabilization Time
MTLR	Rate/Probability	Money Transfer Loss Rate
MTDR	Rate/Probability	Money Transfer Duplication Rate

- KPI defined in ITU-T Rec. P.1502

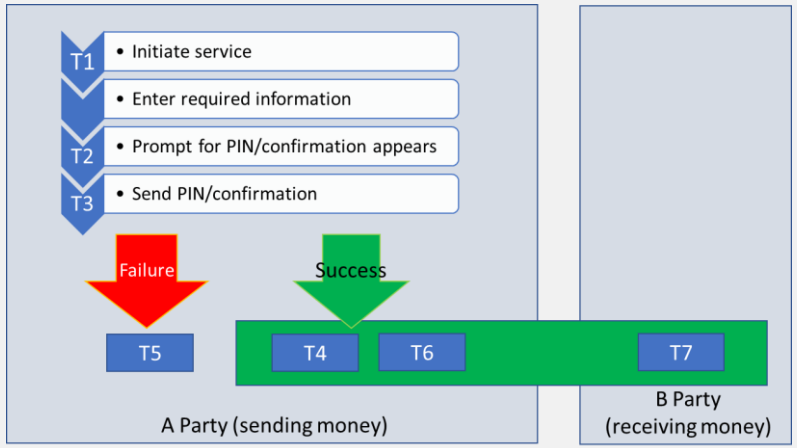
FIG
18 M

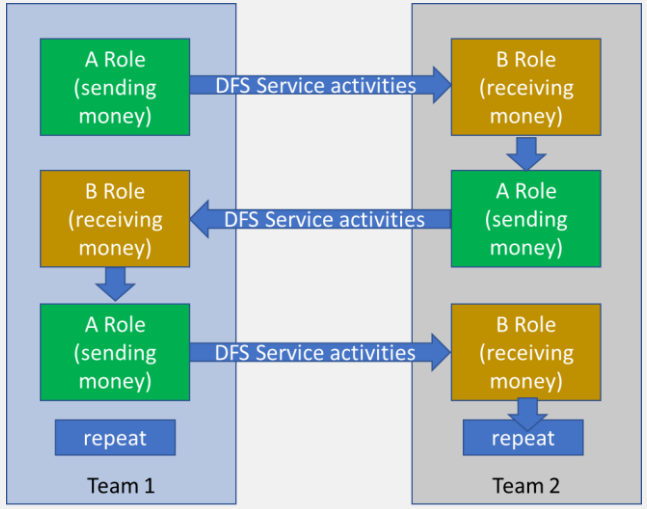
Basic Methodology and KPI

Indicator	Abbreviation	Computation	Reference to formal KPI
Money Transfer Core Duration	MTCD	T3-T2	
Money Transfer Raw Completion Time	MTRCT	T3-T1	MTCT
Money Transfer completion rate	MTCR	T1 present, T3 present: success Valid Try: T1 present	MTCR
Money Transfer Full Completion Time	MTFCT	T7-T1	
Money Transfer A-side Completion Time	MTACT	T6-T1	

Transaction model and KPI

Single transaction

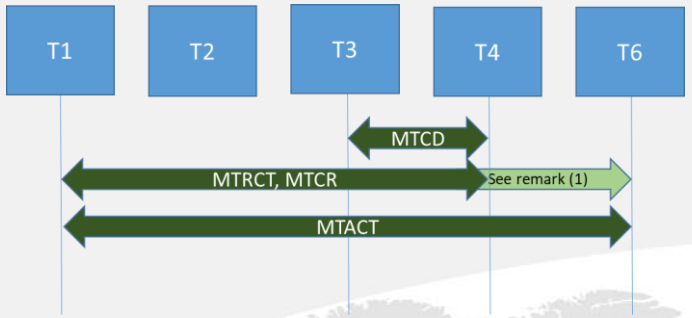




Basic Methodology and KPI

Testing sequence (multiple transactions)

„timerflag“ definitions



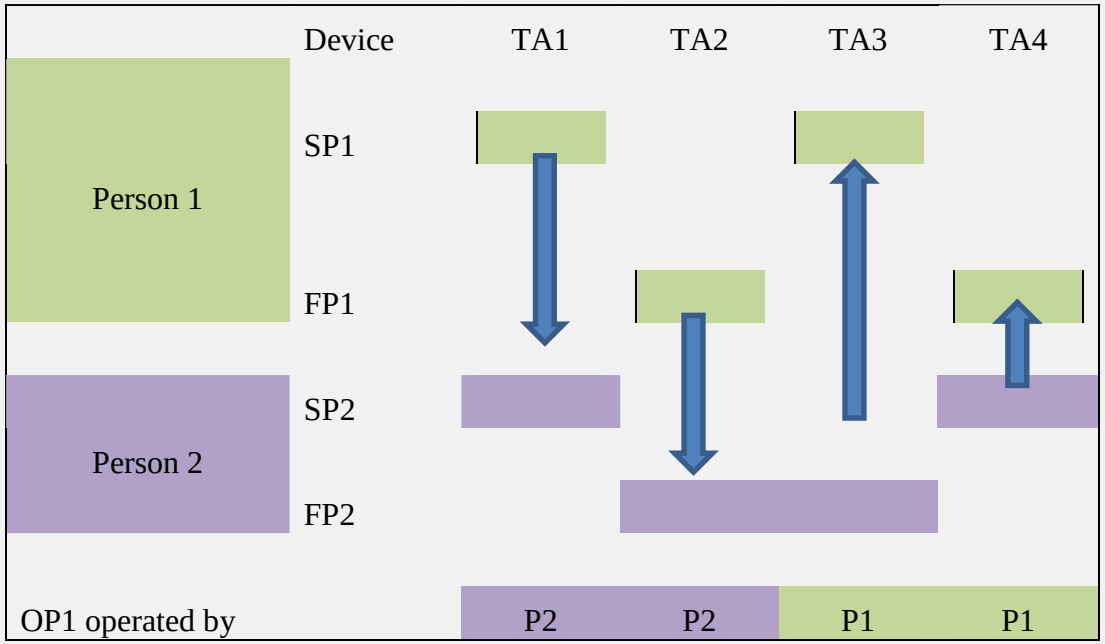
Indicator	Abbreviation	Computation with events used in this project
Money Transfer Core Duration [s]	MTCD	T4-T3
Money Transfer Raw Completion Time [s]	MTRCT	T4-T1
Money Transfer completion rate [%]	MTCR	T1 present, T4 present: success (see remark 1)
Money Transfer Full Completion Time [s]	MTFCT	T7-T1: Not reported due to testing mode (no B-side event tracking)
Money Transfer A-side Completion Time [s]	MTACT	T6-T1 (see remark 2)

Formal KPI definitions (for computation)

- In the first Ghana pilot, synchronous mode was used, so additional KPI were available (see Practical set of DFS KPI)

Methodology: Circular money transfer scheme

Basic Methodology and KPI



We need to make sure that there is sufficient credit on each device/account. Circular schemes are a way to do this with a moderate amount of credit on each device.

- 2 types of devices
- After 4 transfers, money (reduced by transaction fees) is back „by category“
- Full cycle completed after 8 transfers

Design of Test Campaigns

- Transparency
  - Clearly define what is measured
  - Limit uncontrollable environmental effects as much as possible
  - Validate test methods/use cases and data processing before
- Repeatability
  - Test question: Which information is needed to repeat the test?
    - Have an understanding of which outer effects affect results, e.g. platform (computer/mobile device)
  - If the system under test is the same, the test must produce the same result (within statistical accuracy) when repeated
  - Keep records, logs etc.
- Data Quality: Make sure data is meaningful
  - Minimize the effects of human error
  - Use automation and tool-assisted testing wherever it makes sense
  - Make sure relevant data is measured in a well-defined way
    - E.g. time-taking should always use the same tool and way of reading the values
- Robustness: Protect against loss of data
  - Make sure that measured data is preserved (e.g. back-ups, intermediate upload); use four-eye principle where reasonably feasible

Data Quality: make sure data is meaningful

- Ensure data quality at measurement time
  - Use check lists and logs
    - Support field testers to make sure that equipment works properly
    - Create a solid track of activities
    - Record information of events which may affect validity of data
  - Use pre-printed templates or electronic means
    - Reduce the amount of handwriting as much as possible
    - Consider direct entry to e.g. Excel® tables (make sure to include a good backup strategy). Consider however extra handling effort of computers
  - Use back-up copies
    - Analyze scenarios where data can be lost, and design strategies to minimize risk (e.g. take photos of filled-in sheets and email them to a backup location)

A major risk to data quality is, paradoxically, „smoothness“ of testing: repeated tasks can become monotonous and dull. Checking procedures can help to maintain alertness

Data Quality (2)

- Cross-checks during the measurement campaign
  - Schedule checks on uploaded/submitted data as early as possible
  - Have a chance to intervene if plausibility checks indicate potential trouble in processes
- Cross-checks during post processing (examples)
  - Run plausibility checks on times/dates
    - Unusually long or short durations of transactions
  - Check if the number of measurement data points (transactions) is consistent with frame conditions
    - E.g. typical duration of a single transaction vs. transactions recorded per hour of testing
    - E.g. plot transactions on a time axis and check against test planning

Field testing - Basics

- Field time is precious – avoid idle times or unclear conditions
- Use a careful „mental walk-through“ for the execution of tests; take nothing for granted
  - Is there enough battery level on the mobile devices to run the test?
  - Can batteries be recharged (wall plugs, chargers, power banks…)
  - Is there enough credit on the SIMs?
  - Special questions for MoMo service tests:
    - What happens if money „disappears“ from the loop? (sufficient reserve, an action plan for that case)
    - Shall tests be made in „mystery shopping“ mode? What happens if some security system becomes aware of atypical patterns (e.g. unusually frequent money transfers between the same accounts)

Test Design, Repeatability

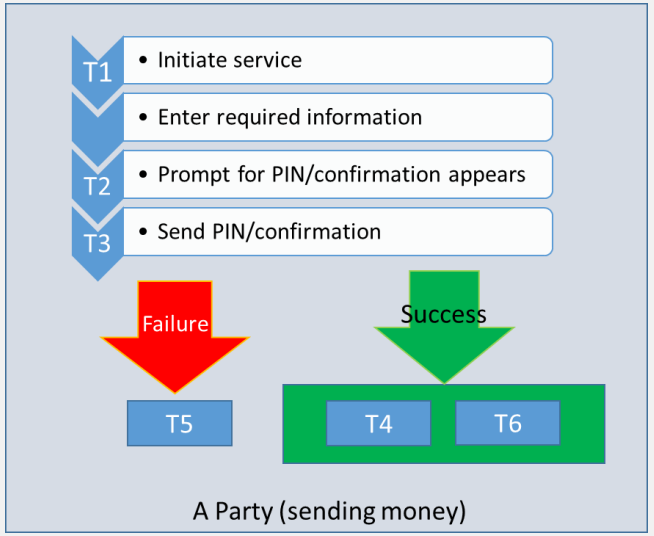
- Formalized guidance of testing operations
  - Use check lists to ensure valid initial conditions, and to regularly check integrity of testing conditions
  - Collect information on unusual outer conditions
- Formalized data collection and data transfer rules
  - Use forms to collect results
  - Transcribe to electronic tables

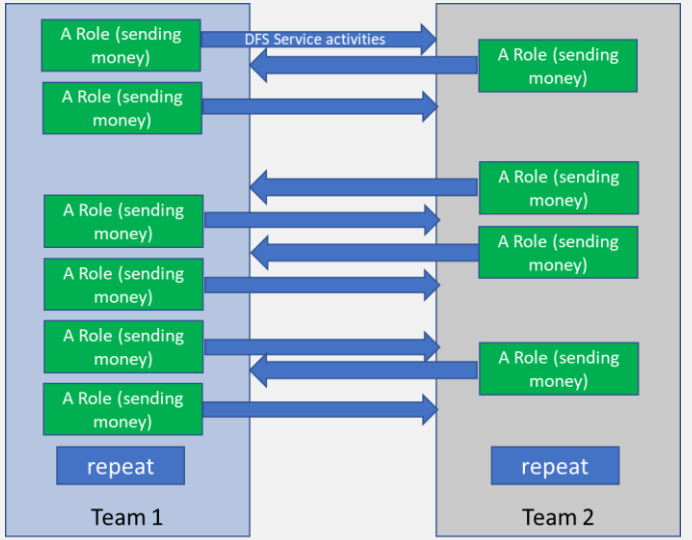
Error margins

- Tests produce a limited number of data points (samples)
- The accuracy of indicators is a function of sample count
  - Example: When a success rate is calculated from 100 measurements, the result has a statistical uncertainty of ~ +/- 3%
- Attempt to understand statistical error margins, use sample counts to assess confidence intervals
  - Example: 100 samples, 5% base unsuccessful rate; TS 102 250-6 Pearson-Clopper tables 4.4.4 and 4.4.5 give a range of 1.64 to 11.28% with a confidence level of 95%
  - Also refer to ITU-T Rec. E.840 (specifically: sections 9.3.2 and 9.3.3) for considerations on statistical error and statistical significance

Transaction model and KPI: same as P2P basis (but using async mode)

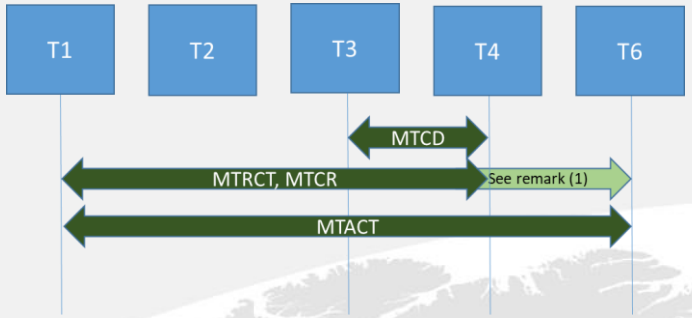
Single transaction





Testing sequence (multiple transactions)

„timerflag“ definitions



Indicator	Abbreviation	Computation with events used in this project
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 reported due to testing mode (no B-side event tracking)
Money Transfer A-side Completion Time [s]	MTACT	T6-T1 (see remark 2)

Formal KPI definitions (for computation)

- Extended use case: Using different source and destination accounts/identities

Potential of technical evolution: Increasing degree of automation

Future: Evolution

- Generic system for full automation of testing
  - “Robot finger”: external automation by a general machine programmed to do task-specific operations
  - Specialized DFS testing app (internal automation)
  - Challenge: Handling country-specific implementation details and service-specific properties (e.g. restrictions)
  - Challenge: Testing the test system (after all, it is real money being moved, and some features may only work in a specific country or region)
- Automated processing
  - Automated tool chains (data upload, data cleansing, reporting/dashboards)

Benchmarking

Future: Evolution

Best-practice examples to enhance the evolution of technical capabilities

Comparison of performance

- between operators
- between countries