Methodology for measurement of QoS KPIs for digital financial services

Joachim Pomy, Consultant, Germany

Dr. Wolfgang Balzer, Focus Infocom GmbH

ITU Workshop on Network Performance, QoS and QoE
Kigali, Rwanda, 4-5 March 2019

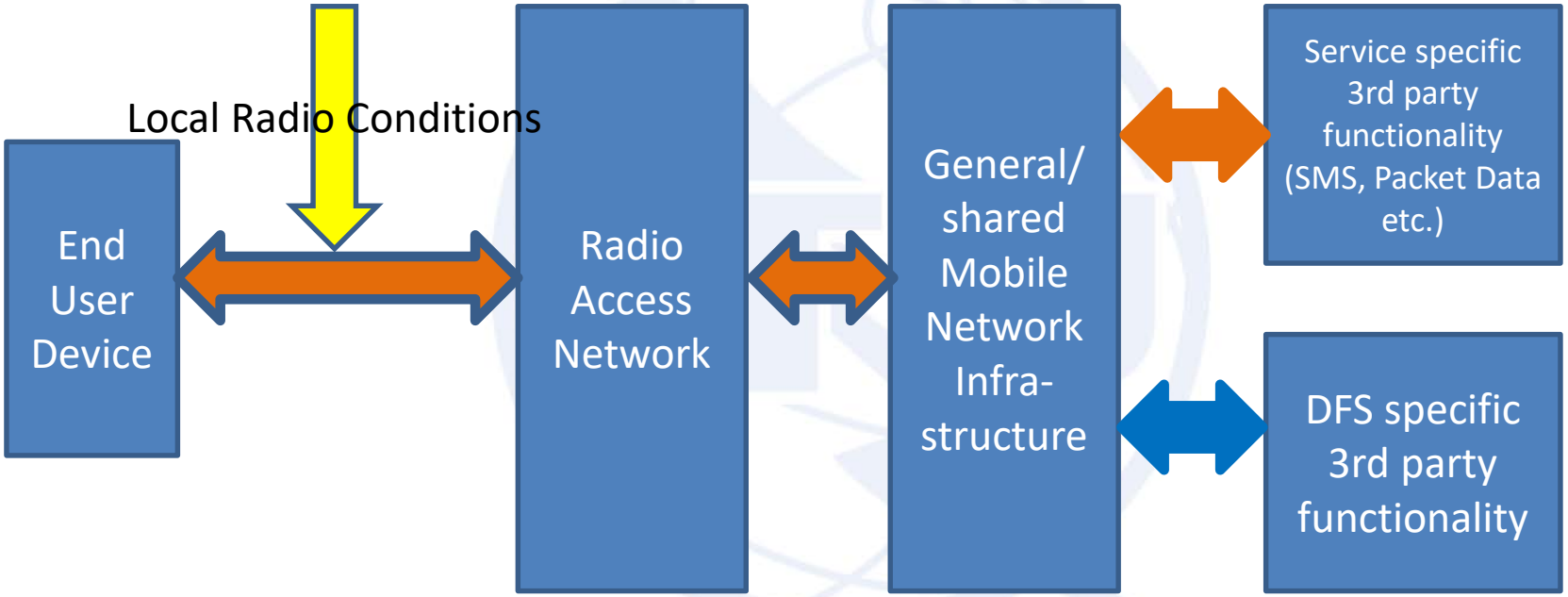
Motivation

- To identify parameters that relate to quality of mobile money transactions
- To establish means of measurement of end to end quality of mobile money transactions
- To benchmark results and set targets
- To understand state of the mobile money market

Overview

- Basic concept of DFS vs Mobile Network QoS
- Key results of the Ghana Test Campaign
- Methodology for end to end QoS Testing of Digital Financial Services

Elements of Service Delivery



QoS vs elements affected by service degradation

	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

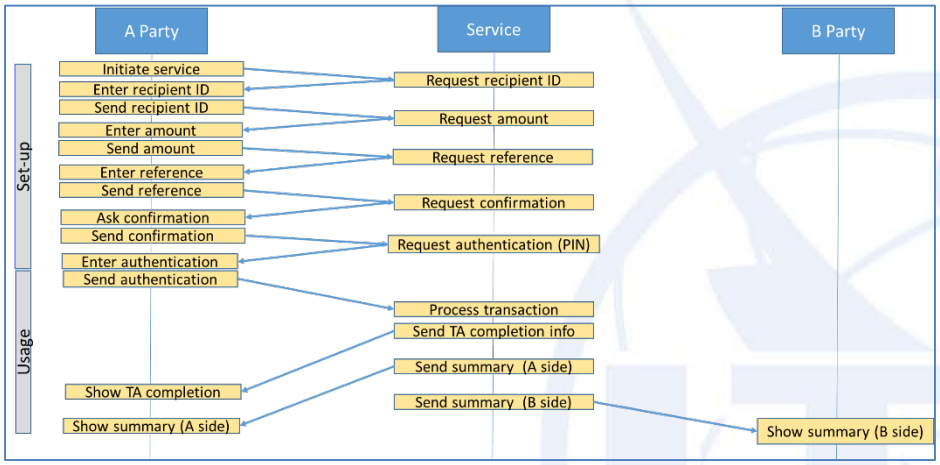
- Unless DFS specific infrastructure is dominantly affected by poor performance, QoS will be determined by mobile network performance
- End to end DFS QoS can only be achieved with good mobile network performance (coverage AND functionality)

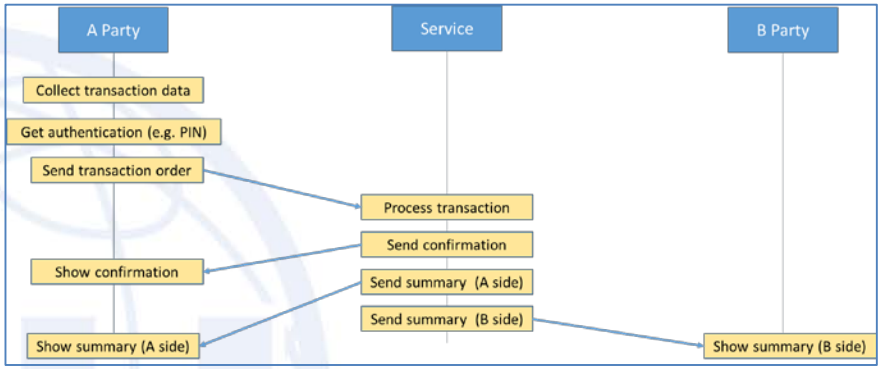
Basic Concept: DFS QoS vs Carrier Service QoS

- If mobile network performance is dominant for overall QoS, proxies can be used to predict DFS end to end QoS which are easier to measure than actual end to end performance.
  - Performance of DFS specific infrastructure may vary over time but does not depend on (is not aware of) the user's location
  - Optimizing/bug fixing DFS subsystem performance can be assumed to be relatively easy (as compared to achieving good area coverage and service performance of a mobile network)
- Results for carrier service measurements provide guidance to regulators w/r to minimum required mobile network performance

Remark: Definitions of and boundaries between QoS and QoE are a nontrivial matter. Some more details – in particular from the viewpoint of ITU-T SG12 – are given here.

End to end QoS KPI





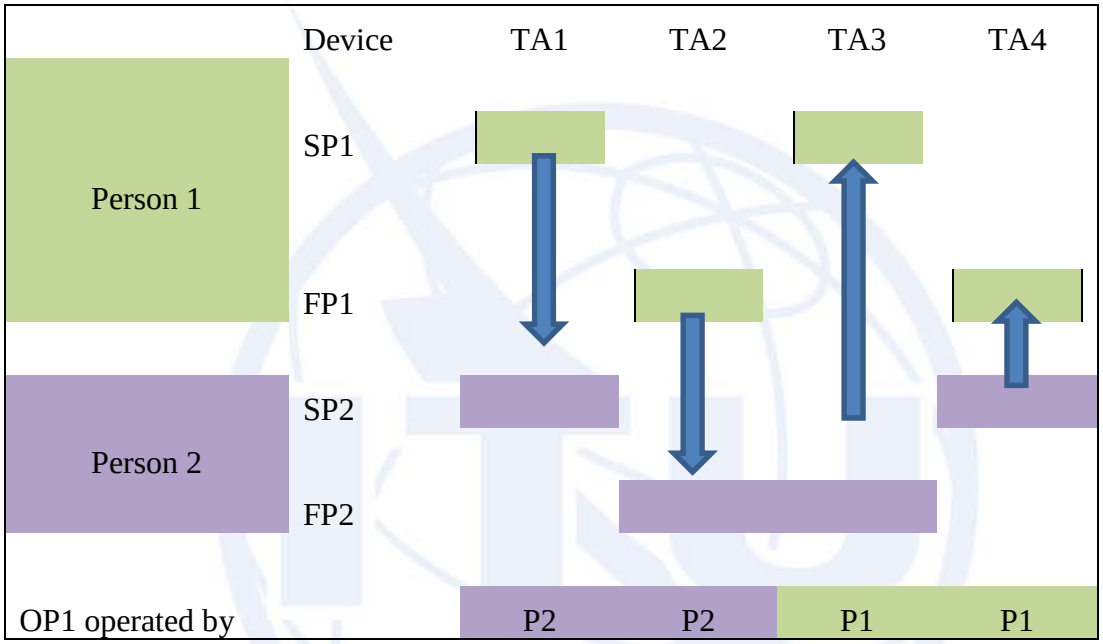
- Left: Detailed sequence of events for a DFS transaction
- Component events may not visible on the user interface, and/or required level of time resolution may not be achievable for human observation
- Right: Simplified sequence, only key events
- KPI used in Ghana as a subset of full KPI; human-measurement perspective

(click on images to show a larger version)

Outer conditions

- Direct observability
- Circular money flow
- Subsequent design of DFS test
- Background tests
- Requirements to equipment in the field
- Robustness and repeatability: Design of data handling and data quality assurance
- Special focus on identification and compensation of artifacts caused by “human errors”

Circular money transfer scheme



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

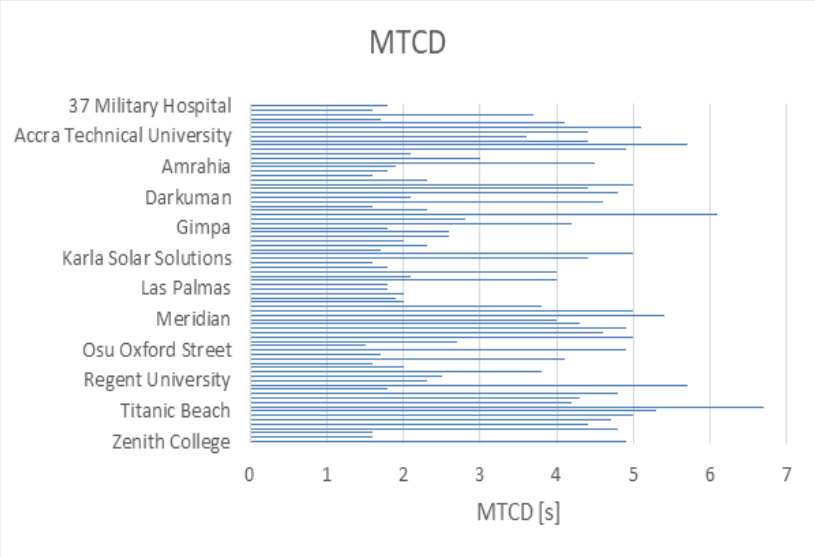
Pilot Campaign Overview

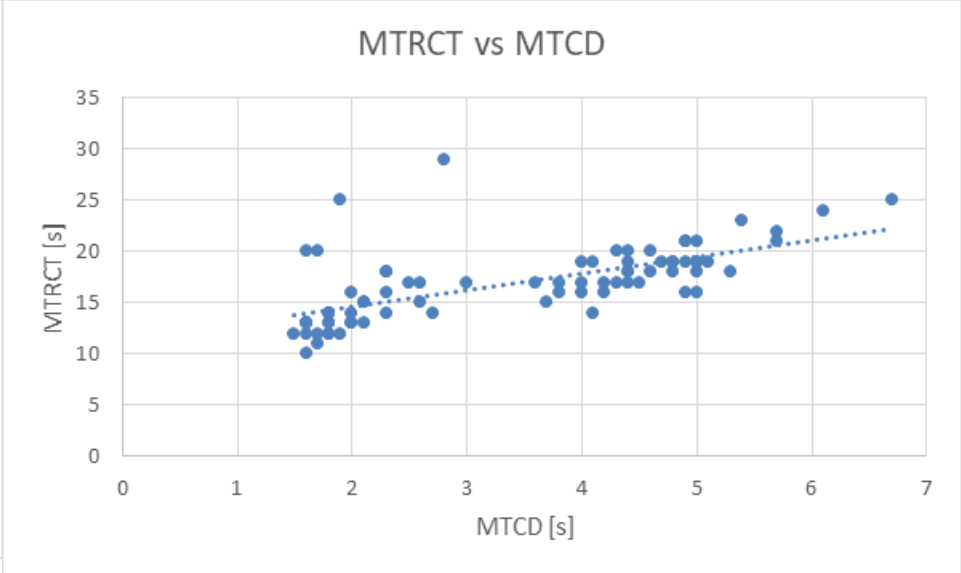
- Total of 78 locations in the larger area of Accra, Ghana
- Measurements made June / July 2018
- Use case: P2P
- 2 teams, each with 4 devices
  - plus 1 “observer device” doing background measurements
- Using a “practical set” of KPI due to the characteristics of the campaign (manual testing and time-taking)
- Typically 48 transactions per location

KPI used

- The set of KPI used for the campaign is explained in detail further down in this presentation. The following slides shows key results using those KPI.

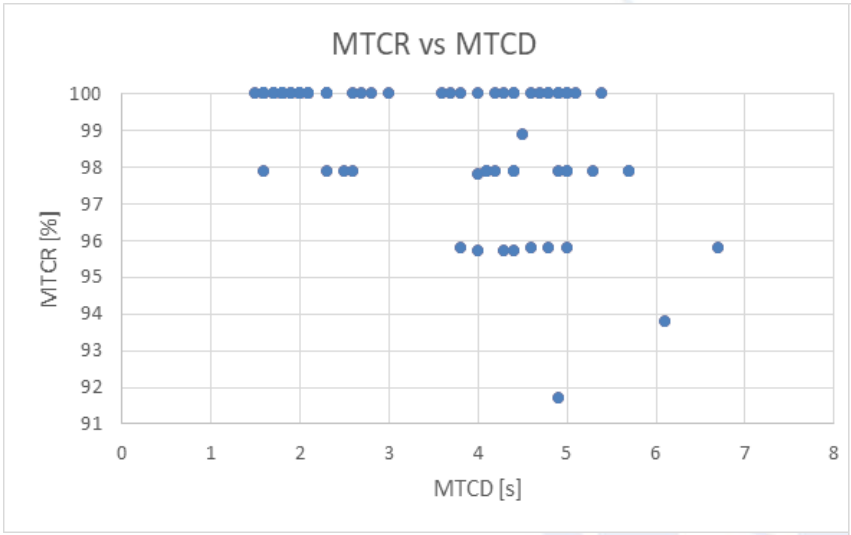
Examples of Results #1

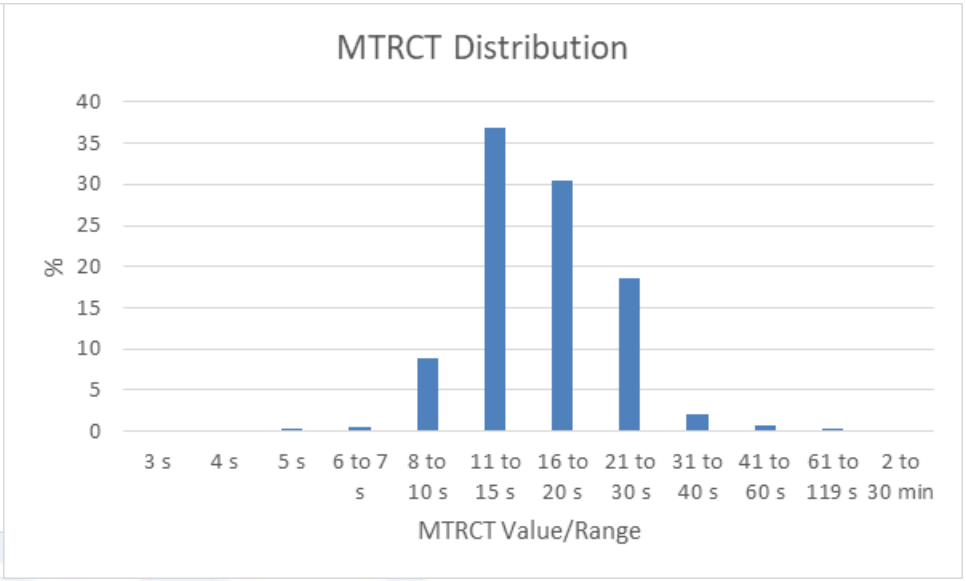




- Left: Variation of MTCD (transaction core time) over locations
- Right: MTRCT (raw completion time including manual operation) vs MTCD
  - Rather narrow distribution of manual execution times (only a few outliers).
- Each data point represents the average of all values at a given location.

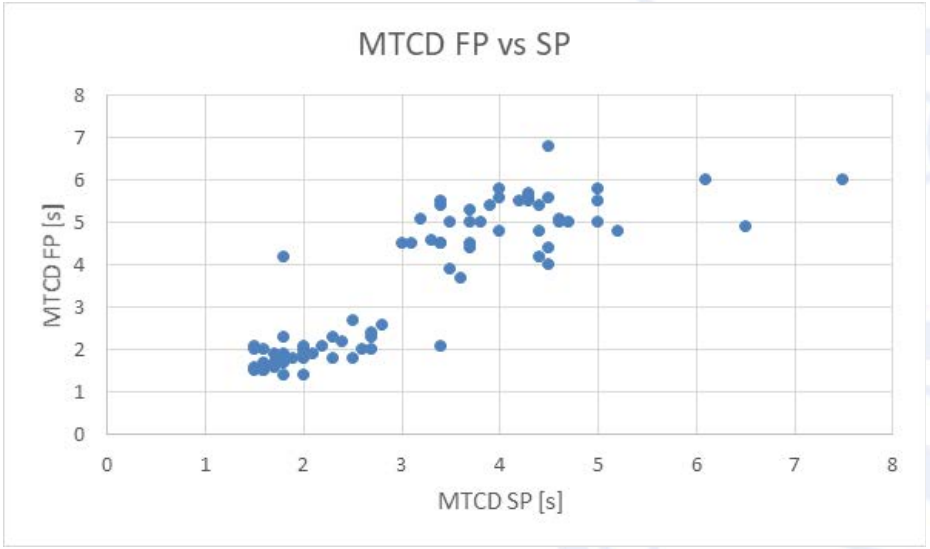
Examples of Results #2

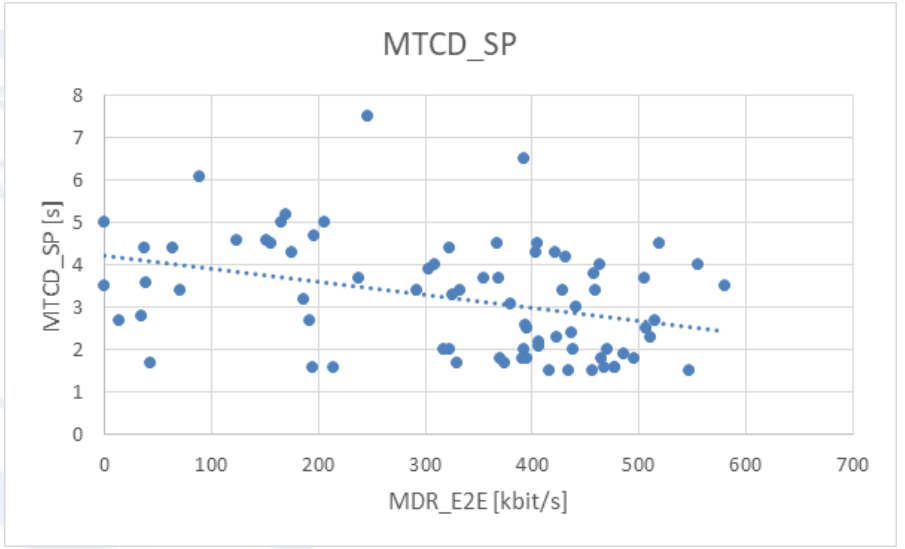




- Left: Completion Rate vs MTCD:
  - failed transactions were quite rare
- Right: MTRCT value distribution, all locations

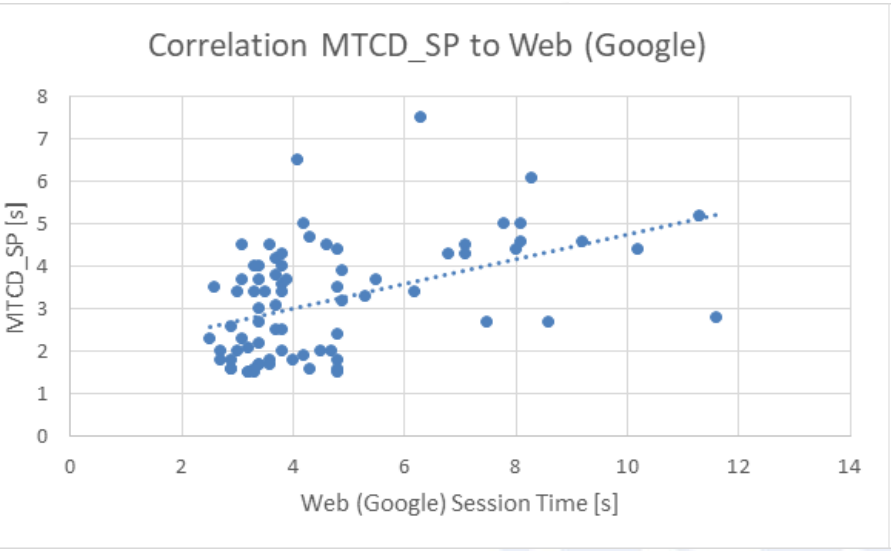
Examples of Results #3

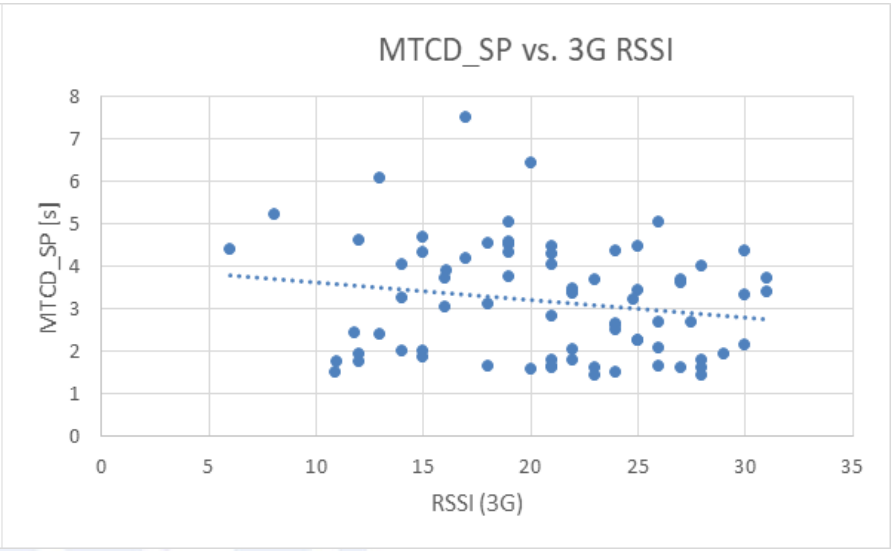




- Left: MTCD correlation by role (FP: restricted to 2G, SP: free RAT selection)
  - Good overall correlation - only some locations differ
    - Comparison with background testing results should be restricted to SP roles
- Right: Correlation between MTCD and Download Data Rate (E2E)
  - Visible correlation (E2E MDR is a useful proxy)
  - Slope of curve suggests that this DFS implementation does not use large volumes of transferred data

Examples of Results #4





- Left: Correlation between Session Time of a low-volume web site (local Google) and MTCD
  - Large variance - but significant correlation
- Right: Correlation between MTCD and 3G RSSI (level indicator)
  - Almost all locations had rather good network coverage
    - Not critically affecting required service performance

General Methodology

- KPI: Definition and computation
- Test design, repeatability
- Data handling (general rules to ensure robustness of tests)
- Assessment of statistical error margins
- So far, only P2P case is covered but the methodology is scalable and can easily be extended

Full set of DFS 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

Practical set of DFS 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	

KPI Definitions (Example)

- Practical Event Definitions

T1	Transaction started
T2	All data entered, trigger money transfer
T3	Reception of primary success indicator from system
T4	(aux): Failure of transaction received
T5	(aux): Time-out for transaction reached
T6	Reception of A-side confirmation SMS
T7	Reception of B-side confirmation SMS

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

Data Handling, Robustness

- Use intermediate back-up of collected data /take photos of completed forms)
- Run a series of plausibility checks on imported data
- Run cross checks on data from different sources

Error margins

- 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

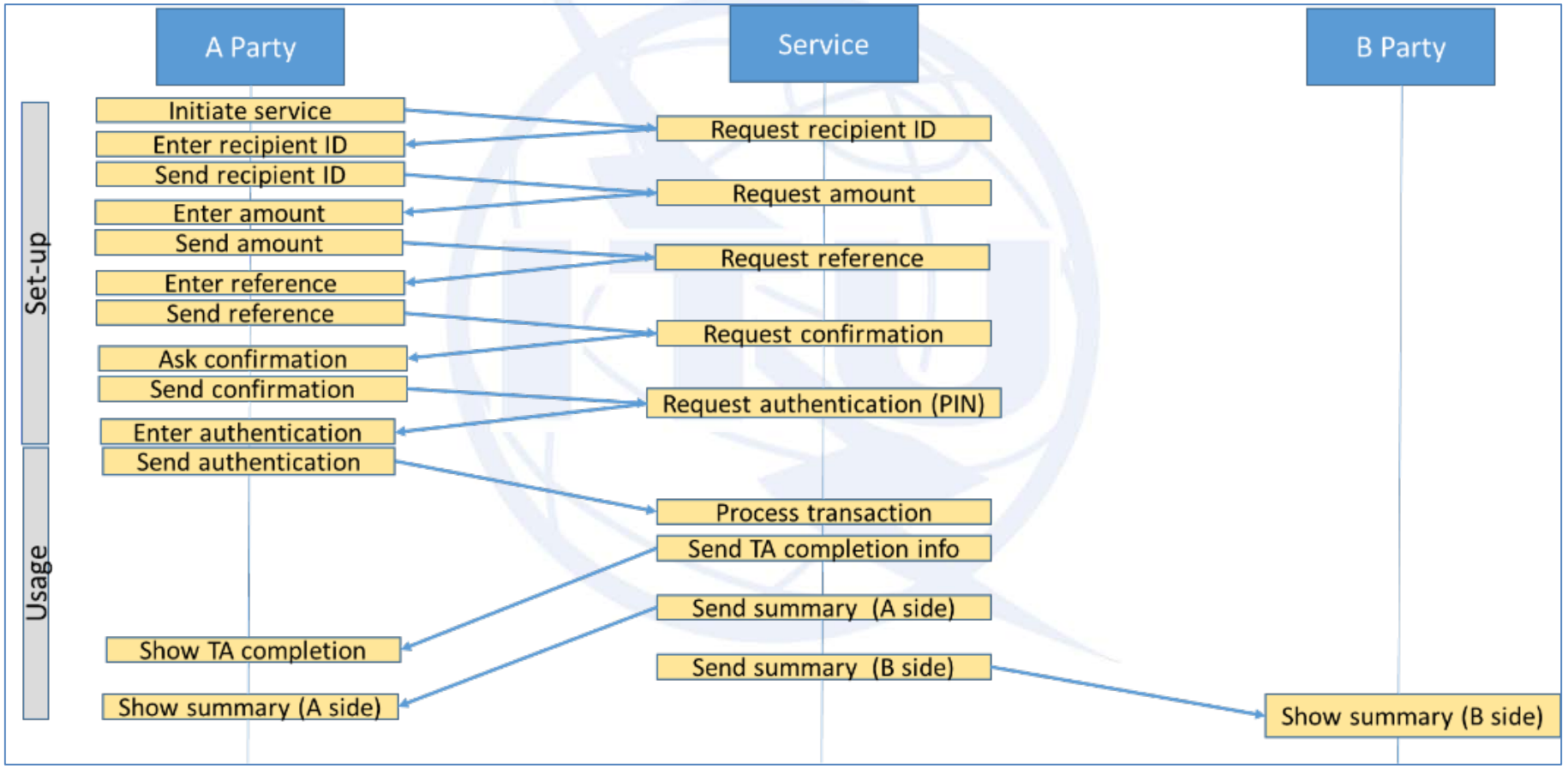
Conclusion and Way forward

- Standardize Methodology in ITU-T SG12
  - In progress
- Extended scenarios
  - Areas with poor or very poor coverage
  - Mobility (e.g. user in public transport vehicles)
- More use cases
  - Interoperability: e.g. national between different DFS operators, international
  - Transfer between mobile and fixed accounts
  - Bulk payments to multiple accounts (e.g. G2P)
- Toolset Expansion
  - App-supported time-taking (including automated upload of data)
  - Standardized data evaluation
  - Automated testing
- Extended range of platforms
  - Dedicated devices (e.g. IoT or Low Power Network based)
  - App-based implementations (e.g. secure transactions)

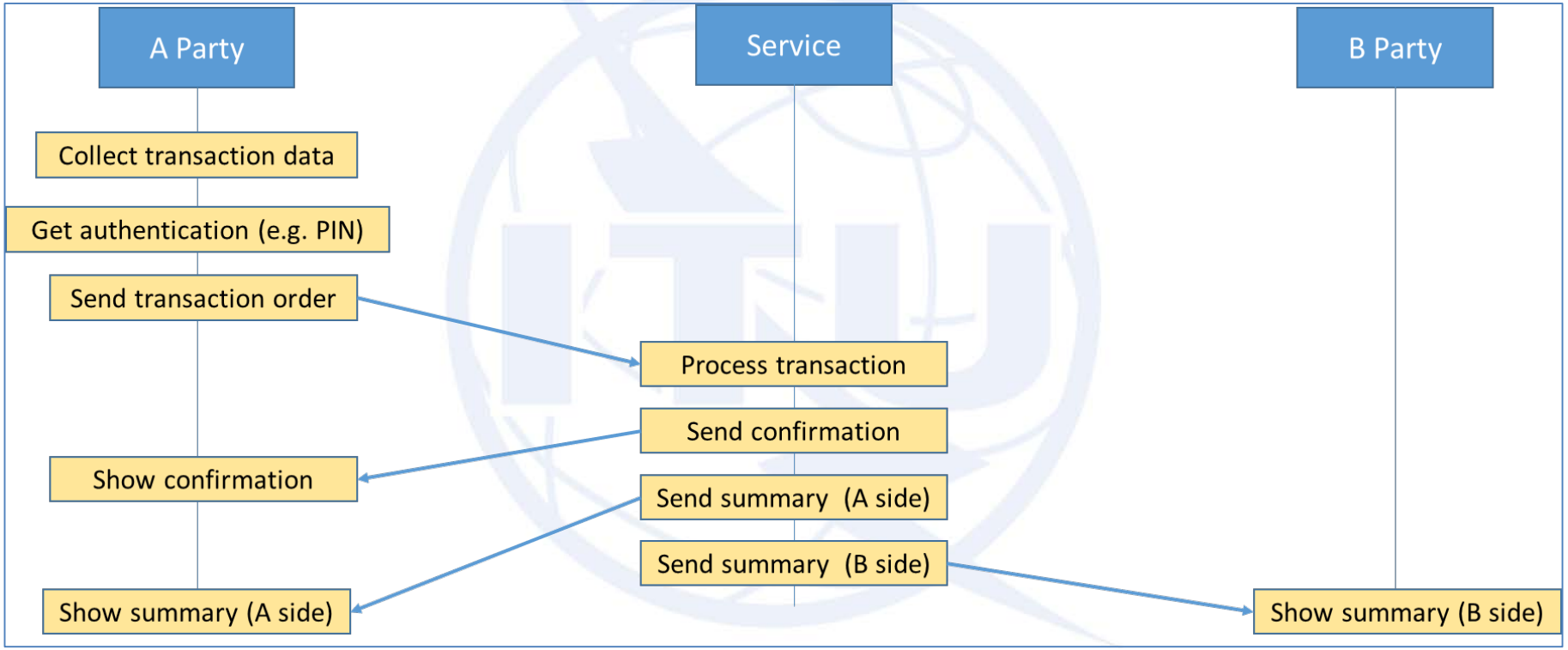
Thank you for your attention.
Questions?

Contacts:
Joachim Pomy
consultant@joachimpomy.de
Dr. Wolfgang Balzer
Focus Infocom GmbH
www.focus-infocom.de
Wolfgang.balzer@focus-infocom.de

Detail 1 – Full Event Flow



Detail 2 – Simplified Event Flow



QoS vs. QoE

- QoE – Rec. ITU-T P.10/G.100
  - 6.209 quality of experience (QoE)
  - The degree of delight or annoyance of the user of an application or service. [b-Qualinet2013]
  - NOTE – Recognizing on-going research on this topic, this is a working definition which is expected to evolve for some time. (This note is not part of the definition.)
  - Source: https://www.itu.int/ITU-T/recommendations/rec.aspx?id=13408&lang=en

- QoS Rec. Rec. ITU-T G.1000
  - 3.2 quality of service (QoS): the collective effect of service performances, which determine the degree of satisfaction of a user of the service (ITU-T Rec. E.800).
  - Sourffce: https://www.itu.int/ITU-T/recommendations/rec.aspx?id=5597&lang=en

Back to main slides

QoS and QoE contd.

- QoE – Rec. ITU-T P.10/G.100
  - 6.210 QoE influencing factors
    Include the type and characteristics of the application or service, context of use, the user's expectations with respect to the application or service and their fulfilment, the user's cultural background, socio-economic issues, psychological profiles, emotional state of the user, and other factors whose number will likely expand with further research.
  - 6.211 QoE assessment
    The process of measuring or estimating the QoE for a set of users of an application or a service with a dedicated procedure, and considering the influencing factors (possibly controlled, measured, or simply collected and reported). The output of the process may be a scalar value, multi dimensional representation of the results, and/or verbal descriptors. All assessments of QoE should be accompanied by the description of the influencing factors that are included. The assessment of QoE can be described as comprehensive when it includes many of the specific factors, for example a majority of the known factors. Therefore, a limited QoE assessment would include only one or a small number of factors.

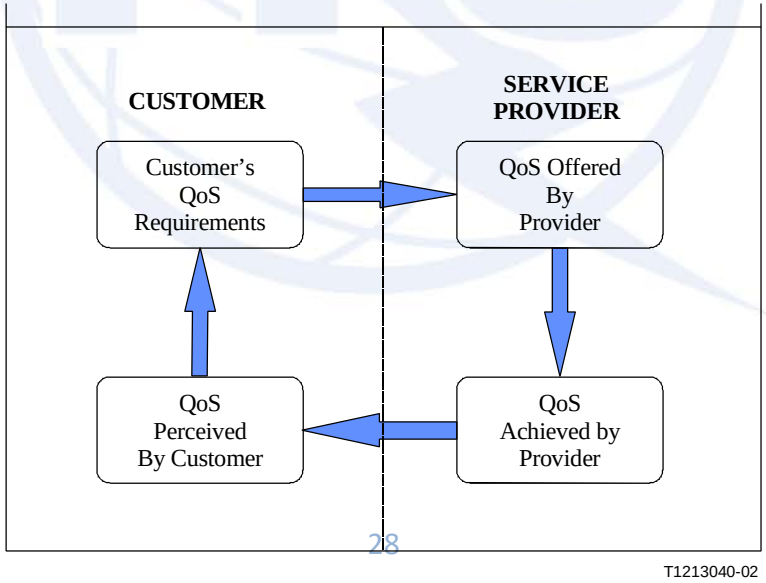


Figure 2/G.1000 – The four viewpoints of QoS

Back to main slides

Practical definition: QoS vs QoE

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