ETSI TS 103 597-3 V1.1.1 (2021-01)

Methods for Testing and Specification (MTS); Test Specification for MQTT; Part 3: Performance Tests

Reference
DTS/MTS-TSTMQTT-3

Keywords
performance, testing

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

The present document can be downloaded from:
http://www.etsi.org/standards-search

Copyright Notification

Contents

Intellectual Property Rights

Essential patents

IPRs essential or potentially essential to normative deliverables may have been declared to ETSI. The information pertaining to these essential IPRs, if any, is publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available from the ETSI Secretariat. Latest updates are available on the ETSI Web server (https://ipr.etsi.org/).

Pursuant to the ETSI IPR Policy, no investigation, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Trademarks

The present document may include trademarks and/or tradenames which are asserted and/or registered by their owners. ETSI claims no ownership of these except for any which are indicated as being the property of ETSI, and conveys no right to use or reproduce any trademark and/or tradename. Mention of those trademarks in the present document does not constitute an endorsement by ETSI of products, services or organizations associated with those trademarks.

Foreword

This Technical Specification (TS) has been produced by ETSI Technical Committee Methods for Testing and Specification (MTS).

The present document is part 3 of a multi-part deliverable. Full details of the entire series can be found in part 1 [2].

Modal verbs terminology

In the present document "**shall**", "**shall not**", "**should**", "**should not**", "**may**", "**need not**", "**will**", "**will not**", "**can**" and "**cannot**" are to be interpreted as described in clause 3.2 of the ETSI Drafting Rules (Verbal forms for the expression of provisions).

"**must**" and "**must not**" are **NOT** allowed in ETSI deliverables except when used in direct citation.

Introduction

Technology advancements are bringing ever-increasing computing power and network speed in the communication domain. The number of communicating devices is expected to increase by 2 orders of magnitude in the following decade and with that several challenges emerge. A main challenge pertains to efficiency regarding resource consumption and overall performance.

As existing communication protocols evolve and new ones are created to fit the current technological capabilities and societal needs and the standards that serve the basis for interoperability and compliance. This is most relevant in the foreseen context of the Internet of Things (IoT) which envisions a very high density of connected devices in the near future. The Message Queuing Telemetry Transport (MQTT) protocol is one such example of evolution.

While many IoT components communicate over standardized protocols, communication protocols for IoT like MQTT or CoAP evolved over time without a holistic approach for quality assurance. Although there are many published evaluations of various MQTT implementations, a lack of common language, methods and presentation of results is slowing down the adoption rate and overall evolution of the protocol.

In the present document the performance testing is presented. It provides a basis for benchmark testing and performance evaluation for the MQTT protocol.

1 Scope

The present document provides a test specification, i.e. an overall test suite structure and catalogue of test purposes for the Message Queuing Telemetry Transport (MQTT) protocol. It will be a reference base for both client side test campaigns and server side test campaigns addressing the performance issues.

2 References

2.1 Normative references

References are either specific (identified by date of publication and/or edition number or version number) or non-specific. For specific references, only the cited version applies. For non-specific references, the latest version of the referenced document (including any amendments) applies.

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

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

The following referenced documents are necessary for the application of the present document.

[1] OASIS Standard: "MQTT Version 3.1.1".

NOTE: Available at http://docs.oasis-open.org/mqtt/mqtt/v3.1.1/os/mqtt-v3.1.1-os.html.

[2] ETSI TS 103 597-1: "Methods for Testing and Specification (MTS); Test Specification for MQTT; Part 1: Conformance Tests".

[3] ETSI ES 203 119-4: "Methods for Testing and Specification (MTS); The Test Description Language (TDL); Part 4: Structured Test Objective Specification (Extension)".

2.2 Informative references

References are either specific (identified by date of publication and/or edition number or version number) or non-specific. For specific references, only the cited version applies. For non-specific references, the latest version of the referenced document (including any amendments) applies.

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

The following referenced documents are not necessary for the application of the present document but they assist the user with regard to a particular subject area.

[i.1] IETF RFC 2544: "Benchmarking Methodology for Network Interconnect Devices".

[i.2] ETSI TR 101 577: "Methods for Testing and Specifications (MTS); Performance Testing of Distributed Systems; Concepts and Terminology".

[i.3] ISO/IEC 9646-1: " Information technology -- Open Systems Interconnection -- Conformance testing methodology and framework -- Part 1: General concepts".

[i.4] ETSI ES 202 951: "Methods for Testing and Specification (MTS); Model-Based Testing (MBT); Requirements for Modelling Notations".

3 Definition of terms, symbols and abbreviations

3.1 Terms

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

active server: connected server with at least 1 registered topic

active subscriber: connected client with at least 1 topic subscription

application messages: data carried by the MQTT protocol across the network for the application as defined in the OASIS Standard: "MQTT Version 3.1.1" [1]

NOTE: When an Application Message is transported by MQTT it contains payload data, a Quality of Service (QoS), a collection of Properties, and a Topic Name.

benchmark test: procedure by which a test system interacts with a System Under Test to measure its behaviour and produce a benchmark report

benchmark test report: document generated at the conclusion of a test procedure containing the metrics measured during

client: program or device that uses MQTT

NOTE: A Client always establishes the Network Connection to the Server [1], it can:

- Publish Application Messages that other Clients might be interested in.
- Subscribe to request Application Messages that it is interested in receiving.
- Unsubscribe to remove a request for Application Messages.
- Disconnect from the Server.

conformance: extent to which an implementation of a standard satisfies the requirements expressed in that standard

conformance testing: process to verify to what extent the IUT conforms to the standard

Design Objective Capacity (DOC): largest load an SUT can sustain while not exceeding design objectives defined for a use-case

disallowed unicode code point: set of Unicode Control Codes and Unicode Noncharacters which should not be included in a UTF-8 Encoded String

Implementation Under Test (IUT): implementation of one or more Open Systems Interconnection (OSI) protocols in an adjacent user/provider relationship, being the part of a real open system, which is to be studied by testing (ISO/IEC 9646-1 [i.3])

malformed packet: control packet that cannot be parsed according to the protocol specification

MQTT Protocol Packet: packet of information that is sent across the Network Connection as defined in the OASIS Standard: "MQTT Version 3.1.1" [1]

NOTE: The MQTT specification defines fourteen different types of Control Packet, one of which (the PUBLISH packet) is used to convey Application Messages.

parameter: attribute of a SUT, test system, system load, or traffic set whose value is set externally and prior to a benchmark test, and whose value affects the behaviour of the benchmark test

protocol error: error that is detected after the packet has been parsed and found to contain data that is not allowed by the protocol or is inconsistent with the state of the Client or Server as defined in the OASIS Standard

server: program or device that acts as an intermediary between Clients which publish Application Messages [1] and Clients which have made Subscriptions

NOTE: A Server can:

- Accepts Network Connections from Clients.
- Accepts Application Messages published by Clients.
- Process Subscribe and Unsubscribe requests from Clients.
- Forwards Application Messages that match Client Subscriptions.

session: stateful interaction between a Client and a Server

NOTE: Some Sessions last only as long as the Network Connection, others can span multiple consecutive Network Connections between a Client and a Server [1].

subscription: interaction between client and server within a session, where the client receives messages based on a topic filter

NOTE: A Subscription is associated with a single Session. A Session can contain more than one Subscription. Each Subscription within a session has a different Topic Filter [1]. Subscription comprises a Topic Filter and a maximum QoS as defined in the OASIS Standard: "MQTT Version 3.1.1" [1].

system under test: real open system in which the implementation under test resides (ETSI ES 202 951 [i.4])

test scenario: specific path through a use-case, whose implementation by a test system creates a system load

Topic name: label attached to an Application Message which is matched against the Subscriptions known to the Server as defined in the OASIS Standard: "MQTT Version 3.1.1" [1]

NOTE: The Server sends a copy of the Application Message to each Client that has a matching Subscription [1].

topic filter: expression contained in a Subscription, to indicate an interest in one or more topics as defined in the OASIS Standard: "MQTT Version 3.1.1" [1].

NOTE: A Topic Filter can include wildcard characters [1].

traffic-time profile: evolution of the average scenario over a time interval

use case: description of a goal that a user has in interacting with a system, the various actors and the SUT

3.2 Symbols

Void.

3.3 Abbreviations

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

CoAP	Constrained Application Protocol
CPU	Central Processing Unit
DOC	Design Objective Capacity
GTW	Gateway
IoT	Internet of Things
IUT	Implementation Under Test
KPI	Key Performance Indicator
LAN	Local Area Network
MQTT	Message Queuing Telemetry Transport
NoS	Number of Subscribers
OS	Operating System
PICS	Protocol Implementation Conformance Statement
QoS	Quality of Service

RAM	Random Access Memory
SoC	System on a Chip
SSD	Solid State Drive
SUT	System Under Test
TCP	Transmission Control Protocol
TDL	Test Description Language
TDL-TO	Test Description Language - Test Objectives
TS	Test System
WLF	Workload Factor

4 Performance metrics

4.0 Introduction

The performance metrics specified herein pertain to the specifics of a MQTT IUT. As such, the objective is to use these metrics in order to determine how well the MQTT component (be it client or server) is performing its functions. As MQTT is a transport protocol, the metrics will be focused on how fast, reliable and efficient the transport is handled. The metrics are designed to fit this purpose while covering multiple use-case scenarios. Following below are the specific messages of the MQTT protocol for which the performance metrics are defined.

4.1 Concepts

4.1.1 Test Equipment Considerations

For measuring performance of a given Test System (TS), a comprehensive description of the test environment is required. This includes but it is not limited to:

- TS hardware infrastructure: resource specification, type, capacity and distribution.
- test environment type and resources (virtualization technology, allocated resources).
- Measurement equipment hardware/software infrastructure, measurement probe distribution/placement, clock synchronization precision, allocated resources.
- Communication infrastructure: transport network specification, number of switches/hops between TS components, bandwidth capacity.

Additional to the specific characteristics of the SUT, the MQTT protocol specifies sessions as stateful interactions between clients and servers. Because of this, additional performance session-based metrics are considered.

4.1.2 Measurement Preliminary Considerations

In order for the collected measurement data to be useful, special consideration needs to be given to the TS setup. Given that the performance evaluation is targeting one or several IUTs same TS setup characteristics are required in order for the evaluation results to allow valid comparisons between them. Some of the characteristics may refer to infrastructure, hardware, physical or virtual resources as well as network connectivity resources.

4.2 Measurement Methodology

4.2.0 Introduction

This clause presents the test methodology for MQTT performance evaluation. From the performance perspective, all measurable metrics related to the protocol should be considered. Although not exhaustive, these metrics can be categorized as follows.

Powerfulness metrics, as defined in ETSI TR 101 577 [i.2] include 3 sub categories: Responsiveness, Capacity and Scalability. From the Responsiveness category the response time, roundtrip time and latency time metrics are used. From the Capacity category the arrival capacity, peak capacity, in progress capacity, streaming capacity and Throughput capacity metrics are used. From the Scalability category the scaling capacity metric is used.

Reliability metrics, as defined in ETSI TR 101 577 [i.2] include 6 sub categories: Quality of Service (QoS), Stability, Availability, Robustness, Recovery, and Correctness. The Quality of Service sub category refers to well defined requirements which may include acceptable values or ranges for metrics from other categories. Stability refers to the capacity of the System to deliver acceptable performance over time. From the Availability sub category, the logical availability metric is used. From the Robustness sub category, the service capacity reduction and service responsiveness deterioration metrics are used. From the Recovery sub category, the service restart characteristics metric is used. Correctness metrics cover the ability of delivering correctly processed requests under high or odd load conditions.

Efficiency metrics, as defined in ETSI TR 101 577 [i.2], cover resource utilization. The metrics cover the characteristics of resource usage, linearity, scalability and bottleneck. The efficiency metrics used in the present document are referring to the service level and not covering the platform level.

4.2.1 Metric Post-processing

The collection of metric values from a SUT is performed by multiple agents and/or directly by the IUT. Often a better insight into the IUT performance is gained by post-processing these values in order to get more meaningful results. To this scope, the data samples can be aggregates over time intervals in the experiment. From such common practices, the following are used for the metrics listed in this clause:

- Mean Average: $\frac{1}{n}\sum_1^n x_i$, where n is the number of samples and x is a sample value.
- Standard deviation: $\sqrt{\frac{1}{n}\sum_1^n (x_i - \bar{x}_i)^2}$, where n is the number of samples, x is a sample value and $\bar{x}$ is the mean average.
- Minimum: min(x_i), the smallest sample value, relative to the rest of the samples.
- Maximum: max(x_i), the greatest sample value, relative to the rest of the samples.

4.2.2 Message Types

Table 1 contains the set control packet messages specified by the MQTT [1] standard.

Table 1: Message Types

Control Packet Name	Description	Client-> Server	Server-> Client	Payload
CONNECT	client requests a connection to the server	✓		Required
CONNACK	acknowledge connection request		✓	None
PUBLISH	Publish message	✓	✓	Optional
PUBACK	Publish acknowledgement	✓	✓	None
PUBREC	Publish received (QoS 2 publish received)	✓	✓	None
PUBREL	Publish release	✓	✓	None
PUBCOMP	Publish complete	✓	✓	None
SUBSCRIBE	Subscribe to topics	✓		Required
SUBACK	Subscribe acknowledgement		✓	Required
UNSUBSCRIBE	Unsubscribe from Topics	✓		Required
UNSUBACK	Unsubscribe acknowledgement		✓	Required
PINGREQ	Ping request	✓		None
PINGRESP	Ping response		✓	None
DISCONNECT	Disconnect notification	✓	✓	None
AUTH	Authentication Exchange	✓	✓	None

4.2.3 Test parameters

The benchmark test parameters are used to control the behaviour of the test script. The data elements required to configure the test system are listed in table 2.

Table 2 is a non-exhaustive list of test parameters defined for the benchmark standard. The list is expected to grow over time, as additional subsystems and system configurations are developed.

Table 2: Test parameters

Parameter	Description
Duration	Amount of time that a system load is presented to a SUT
Type of call	Type of messages contained within a workload
NoC	number of clients generating or subscribing to data/control traffic
NoS	Number of servers handling data/control traffic
Transport interface	Underlying transport interface
WLF for GTW	Work load factor for gateway expressed in number messages received per second, by type of message
Payload	Size of the data in Bytes carried within a message
Monitoring Window(s)	The time interval window for which the monitored metrics are recorded. This reflects the measuring accuracy (e.g. per second, minute, hour, etc.)
Validation threshold(s)	The specific metric thresholds used for validating whether a system performs at specifications

Table 3: Test output

Metric	Description
Minimum call duration	The minimum duration of a successful message request/response interaction within a Monitoring Window
Maximum call duration	The maximum duration of a successful message request/response interaction within a Monitoring Window
Average call duration	The average duration of a successful message request/response interaction within a Monitoring Window
Total number of calls	The total number of workload specified request/response type operations executed during the test
Success rate	Percentage number of successful workload operations relative to the total workload operations
Error rate	Percentage number of failed workload operations relative to the total workload operations
Requests processed per time unit	This metric reflects the average number of successfully processed requests per preferred time unit (second/minute/etc.)

4.2.4 Operation Message Flows

The IUT will be evaluated based on the metric values obtain as a result of the service operations using the messages described in clause 4.1.1. The set of messages exchanged triggered by the initial client message are further referred as operations. For the tests, the metrics use operations rather than specific messages because it is easier to handle the measurements. E.g. for a ping message, the roundtrip time is calculated as the duration between the time the PINGREQ is sent and the time when PINGRESP is received. If specific test system network measurements are available, by subtracting the measured network delayed from the duration, the operation processing time can be deducted:

1) CONNECT: This clause describes the MQTT operation types and message sequences required for test execution using a Client Connection example:

   Preconditions:

   - Unregistered Client, Server/Broker.
   - TCP connection between Client and Server/Broker established.

   Operation sequence:

   - Client sends CONNECT message.
   - Client receives CONNACK message.

   Measurement:

   - Time period expressed in milliseconds between the moment client forwards the CONNECT Message and the moment Client receives CONNACK message from server.

Figure 1: MQTT Server-Client CONNECT message flow

2) PINGREQ: This section describes the MQTT operation types and message sequences required for test execution using a Client ping example:

   Preconditions:

   - Connected Client, Server/Broker.
   - TCP connection between Client and Server/Broker established.

   Operation sequence:

   - Client sends PINGREQ message.
   - Client receives PINGRESP message.

   Measurement:

   - Time period expressed in milliseconds between the moment client forwards the PINGREQ Message and the moment Client receives PINGRESP message from server.

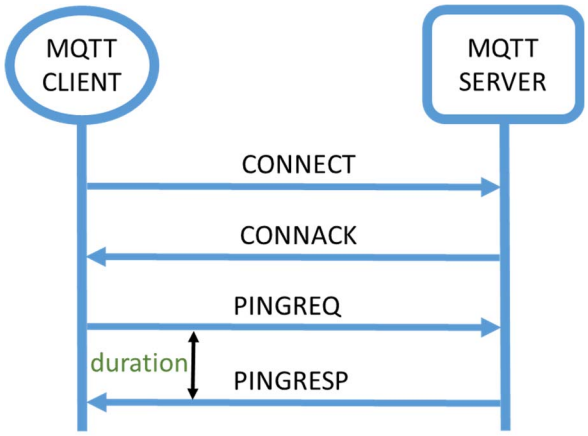

Figure 2: MQTT Server-Client PING message flow

3) SUBSCRIBE: This section describes the MQTT operation types and message sequences required for test execution using a Client SUBSCRIBE example:

Preconditions:

- Connected Client, Server/Broker.
- TCP connection between Client and Server/Broker established.

Operation sequence:

- Client sends SUBSCRIBE message.
- Client receives SUBACK message.

Measurement:

- Time period expressed in milliseconds between the moment client forwards the SUBSCRIBE Message and the moment Client receives SUBACK message from server.

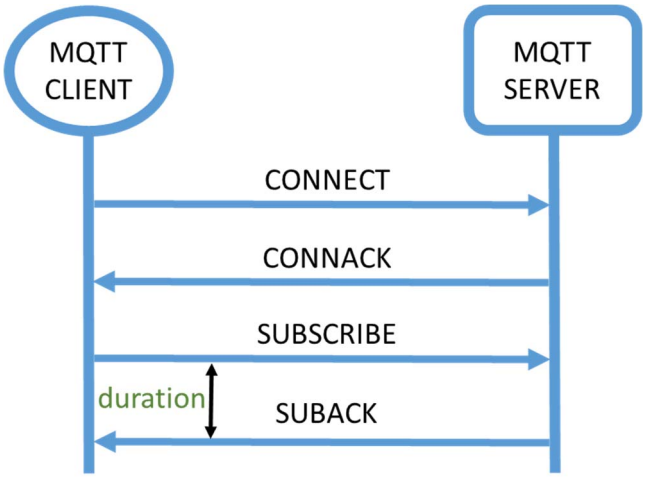

Figure 3: MQTT Server-Client SUBSCRIBE message flow

4) PUBLISH: This section describes the MQTT operation types and message sequences required for test execution using a Client PUBLISH example:

Preconditions:

- Connected Client, Server/Broker.
- TCP connection between Client and Server/Broker established.