Workgroup:
Benchmarking Methodology Working Group
Internet-Draft:
draft-ietf-bmwg-ngfw-performance-10
Obsoletes: 3511 (if approved)
Published: September 2021
Intended Status: Informational
Expires: 30 March 2022
Authors: B. Balarajah   C. Rossenhoevel   B. Monkman
                        EANTC AG   NetSecOPEN

Benchmarking Methodology for Network Security Device Performance

Abstract

This document provides benchmarking terminology and methodology for next-generation network security devices including next-generation firewalls (NGFW), next-generation intrusion prevention systems (NGIPS), and unified threat management (UTM) implementations. This document aims to improve the applicability, reproducibility, and transparency of benchmarks and to align the test methodology with today's increasingly complex layer 7 security centric network application use cases. The main areas covered in this document are test terminology, test configuration parameters, and benchmarking methodology for NGFW and NGIPS.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on 5 March 2022.

Copyright Notice

Table of Contents

1. Introduction

18 years have passed since IETF recommended test methodology and
terminology for firewalls initially ([RFC3511]). The requirements
for network security element performance and effectiveness have
increased tremendously since then. Security function implementations
have evolved to more advanced areas and have diversified into
intrusion detection and prevention, threat management, analysis of
encrypted traffic, etc. In an industry of growing importance, well-
defined, and reproducible key performance indicators (KPIs) are
increasingly needed as they enable fair and reasonable comparison of
network security functions. All these reasons have led to the
creation of a new next-generation network security device
benchmarking document and this document obsoletes [RFC3511].

2. Requirements

The keywords "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT",
"SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and
"OPTIONAL" in this document are to be interpreted as described in
BCP 14 [RFC2119], [RFC8174] when, and only when, they appear in all
capitals, as shown here.

3. Scope

This document provides testing terminology and testing methodology
for modern and next-generation network security devices that are
configured in Active ("Inline") mode. It covers the validation of
security effectiveness configurations of network security devices,
followed by performance benchmark testing. This document focuses on
advanced, realistic, and reproducible testing methods. Additionally,
it describes testbed environments, test tool requirements, and test
result formats.

4. Test Setup

Test setup defined in this document applies to all benchmarking
tests described in Section 7. The test setup MUST be contained
within an Isolated Test Environment (see Section 3 of [RFC6815]).

4.1. Testbed Configuration

Testbed configuration MUST ensure that any performance implications
that are discovered during the benchmark testing aren't due to the
inherent physical network limitations such as the number of physical
links and forwarding performance capabilities (throughput and
latency) of the network devices in the testbed. For this reason,
this document recommends avoiding external devices such as switches
and routers in the testbed wherever possible.

In some deployment scenarios, the network security devices (Device Under Test/System Under Test) are connected to routers and switches, which will reduce the number of entries in MAC or ARP tables of the Device Under Test/System Under Test (DUT/SUT). If MAC or ARP tables have many entries, this may impact the actual DUT/SUT performance due to MAC and ARP/ND (Neighbor Discovery) table lookup processes. This document also recommends using test equipment with the capability of emulating layer 3 routing functionality instead of adding external routers in the testbed.

The testbed setup Option 1 (Figure 1) is the RECOMMENDED testbed setup for the benchmarking test.

```
+-----------------------+                   +-----------------------+
| +-------------------+ |    +-----------+   | +-------------------+ |
| | Emulated Router(s)| |    |           |   | | Emulated Router(s)| |
| |    (Optional)     | +----- DUT/SUT  +-----+    (Optional)     | |
| +-------------------+ |    |           |   | +-------------------+ |
| +-------------------+ |    +-----------+   | +-------------------+ |
| |     Clients       | |                    | |     Servers       | |
| +-------------------+ |                    | +-------------------+ |
|                       |                    |                       |
|    Test Equipment     |                    |    Test Equipment     |
+-----------------------+                    +-----------------------+
```
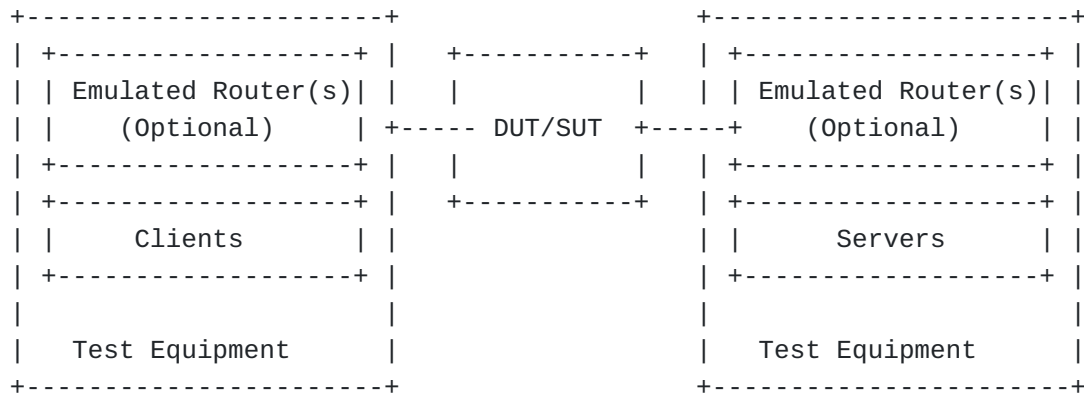
Figure 1: Testbed Setup - Option 1

If the test equipment used is not capable of emulating layer 3 routing functionality or if the number of used ports is mismatched between test equipment and the DUT/SUT (need for test equipment port aggregation), the test setup can be configured as shown in Figure 2.

```
+--------------------+        +-----------+       +--------------------+
|Aggregation Switch/|       |           |       | | Aggregation Switch/|
| Router            +------+  DUT/SUT  +------+ Router             |
|                   |       |           |       |                    |
+----------+--------+       +-----------+       +--------+-----------+
           |                                             |
           |                                             |
+----------+----------+                     +----------+-----------+
|                     |                     |                      |
| +------------------+ |                    | +------------------+ |
| | Emulated Router(s)| |                   | | Emulated Router(s)| |
| |     (Optional)    | |                   | |     (Optional)    | |
| +------------------+ |                    | +------------------+ |
| +------------------+ |                    | +------------------+ |
| |     Clients      | |                    | |     Servers      | |
| +------------------+ |                    | +------------------+ |
|                     |                     |                      |
|     Test Equipment  |                     |     Test Equipment   |
+---------------------+                     +----------------------+
```

Figure 2: Testbed Setup - Option 2

4.2. DUT/SUT Configuration
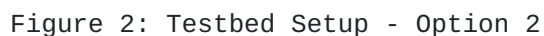
A unique DUT/SUT configuration MUST be used for all benchmarking
tests described in Section 7. Since each DUT/SUT will have its own
unique configuration, users SHOULD configure their device with the
same parameters and security features that would be used in the
actual deployment of the device or a typical deployment in order to
achieve maximum network security coverage. The DUT/SUT MUST be
configured in "Inline" mode so that the traffic is actively
inspected by the DUT/SUT. Also "Fail-Open" behavior MUST be disabled
on the DUT/SUT.

Table 1 and Table 2 below describe the RECOMMENDED and OPTIONAL sets
of network security feature list for NGFW and NGIPS respectively.
The selected security features SHOULD be consistently enabled on the
DUT/SUT for all benchmarking tests described in Section 7.

To improve repeatability, a summary of the DUT/SUT configuration
including a description of all enabled DUT/SUT features MUST be
published with the benchmarking results.

```
                  +------------------------+
                  |          NGFW          |
   +-------------- +-------------+----------+
   |               |             |          |
   |DUT/SUT Features| RECOMMENDED | OPTIONAL |
   |               |             |          |
   +---------------+-------------+----------+
   |SSL Inspection |      x      |          |
   +---------------+-------------+----------+
   |IDS/IPS        |      x      |          |
   +---------------+-------------+----------+
   |Anti-Spyware   |      x      |          |
   +---------------+-------------+----------+
   |Anti-Virus     |      x      |          |
   +---------------+-------------+----------+
   |Anti-Botnet    |      x      |          |
   +---------------+-------------+----------+
   |Web Filtering  |             |    x     |
   +---------------+-------------+----------+
   |Data Loss      |             |          |
   |Protection (DLP)|            |    x     |
   +---------------+-------------+----------+
   |DDoS           |             |    x     |
   +---------------+-------------+----------+
   |Certificate    |             |    x     |
   |Validation     |             |          |
   +---------------+-------------+----------+
   |Logging and    |      x      |          |
   |Reporting      |             |          |
   +---------------+-------------+----------+
   |Application    |      x      |          |
   |Identification |             |          |
   +---------------+-------------+----------+

        Figure 3: Table 1: NGFW Security Features
```

```
                +------------------------+
                |         NGIPS          |
      +---------------+------------+----------+
      |               |            |          |
      |DUT/SUT Features| RECOMMENDED | OPTIONAL |
      |               |            |          |
      +---------------+------------+----------+
      |SSL Inspection |     x      |          |
      +---------------+------------+----------+
      |Anti-Malware   |     x      |          |
      +---------------+------------+----------+
      |Anti-Spyware   |     x      |          |
      +---------------+------------+----------+
      |Anti-Botnet    |     x      |          |
      +---------------+------------+----------+
      |Logging and    |     x      |          |
      |Reporting      |            |          |
      +---------------+------------+----------+
      |Application    |     x      |          |
      |Identification |            |          |
      +---------------+------------+----------+
      |Deep Packet    |     x      |          |
      |Inspection     |            |          |
      +---------------+------------+----------+
      |Anti-Evasion   |     x      |          |
      +---------------+------------+----------+
```

                Figure 4: Table 2: NGIPS Security Features

The following table provides a brief description of the security
features.

DUT/SUT Features	Description
SSL Inspection	DUT/SUT intercepts and decrypts inbound HTTPS traffic between servers and clients. Once the content inspection has been completed, DUT/SUT encrypts the HTTPS traffic with ciphers and keys used by the clients and servers.
IDS/IPS	DUT/SUT detects and blocks exploits targeting known and unknown vulnerabilities across the monitored network.
Anti-Malware	DUT/SUT detects and prevents the transmission of malicious executable code and any associated communications across the monitored network. This includes data exfiltration as well as command and control channels.
Anti-Spyware	Anti-Spyware is a subcategory of Anti Malware. Spyware transmits information without the user's knowledge or permission. DUT/SUT detects and block initial infection or transmission of data.
Anti-Botnet	DUT/SUT detects traffic to or from botnets.
Anti-Evasion	DUT/SUT detects and mitigates attacks that have been obfuscated in some manner.
Web Filtering	DUT/SUT detects and blocks malicious website including defined classifications of website across the monitored network.
DLP	DUT/SUT detects and prevents data breaches and data exfiltration, or it detects and blocks the transmission of sensitive data across the monitored network.
Certificate Validation	DUT/SUT validates certificates used in encrypted communications across the monitored network.
Logging and Reporting	DUT/SUT logs and reports all traffic at the flow level across the monitored network.
Application Identification	DUT/SUT detects known applications as defined within the traffic mix selected across

```
|                    | the monitored network.                        |
+--------------------+-----------------------------------------------+
```

Figure 5: Table 3: Security Feature Description

Below is a summary of the DUT/SUT configuration:

  *DUT/SUT MUST be configured in "inline" mode.

  *"Fail-Open" behavior MUST be disabled.

  *All RECOMMENDED security features are enabled.

  *Logging SHOULD be enabled. DUT/SUT SHOULD log all traffic at the
   flow level - Logging to an external device is permissible.

  *Geographical location filtering, and Application Identification
   and Control SHOULD be configured to trigger based on a site or
   application from the defined traffic mix.

In addition, a realistic number of access control rules (ACL) SHOULD
be configured on the DUT/SUT where ACLs are configurable and
reasonable based on the deployment scenario. This document
determines the number of access policy rules for four different
classes of DUT/SUT: Extra Small (XS), Small (S), Medium (M), and
Large (L). A sample DUT/SUT classification is described in Appendix
B.

The Access Control Rules (ACL) defined in Table 4 MUST be configured
from top to bottom in the correct order as shown in the table. This
is due to ACL types listed in specificity decreasing order, with
"block" first, followed by "allow", representing a typical ACL based
security policy. The ACL entries SHOULD be configured with routable
IP subnets by the DUT/SUT. (Note: There will be differences between
how security vendors implement ACL decision making.) The configured
ACL MUST NOT block the security and measurement traffic used for the
benchmarking tests.

| Rules Type | Match Criteria | Description | Action | DUT/SUT Classification # Rules | | | |
				XS	S	M	L
Application layer	Application	Any application not included in the measurement traffic	block	5	10	20	50
Transport layer	SRC IP and TCP/UDP DST ports	Any SRC IP subnet used and any DST ports not used in the measurement traffic	block	25	50	100	250
IP layer	SRC/DST IP	Any SRC/DST IP subnet not used in the measurement traffic	block	25	50	100	250
Application layer	Application	Half of the applications included in the measurement traffic (see the note below)	allow	10	10	10	10
Transport layer	SRC IP and TCP/UDP DST ports	Half of the SRC IPs used and any DST ports used in the measurement traffic (one rule per subnet)	allow	>1	>1	>1	>1
IP layer	SRC IP	The rest of the SRC IP subnet range used in the measurement traffic (one rule per subnet)	allow	>1	>1	>1	>1

Figure 6: Table 4: DUT/SUT Access List

Note: If half of the applications included in the measurement
traffic is less than 10, the missing number of ACL entries (dummy
rules) can be configured for any application traffic not included in
the measurement traffic.

4.2.1. Security Effectiveness Configuration

The Security features (defined in table 1 and 2) of the DUT/SUT MUST
be configured effectively to detect, prevent, and report the defined
security vulnerability sets. This section defines the selection of
the security vulnerability sets from Common vulnerabilities and
Exposures (CVE) list for the testing. The vulnerability set SHOULD
reflect a minimum of 500 CVEs from no older than 10 calendar years
to the current year. These CVEs SHOULD be selected with a focus on
in-use software commonly found in business applications, with a
Common vulnerability Scoring System (CVSS) Severity of High (7-10).

This document is primarily focused on performance benchmarking.
However, it is RECOMMENDED to validate the security features
configuration of the DUT/SUT by evaluating the security
effectiveness as a prerequisite for performance benchmarking tests
defined in the section 7. In case the benchmarking tests are
performed without evaluating security effectiveness, the test report
MUST explain the implications of this. The methodology for
evaluating security effectiveness is defined in Appendix A.

4.3. Test Equipment Configuration

In general, test equipment allows configuring parameters in
different protocol layers. These parameters thereby influence the
traffic flows which will be offered and impact performance
measurements.

This section specifies common test equipment configuration
parameters applicable for all benchmarking tests defined in Section
7. Any benchmarking test specific parameters are described under the
test setup section of each benchmarking test individually.

4.3.1. Client Configuration

This section specifies which parameters SHOULD be considered while
configuring clients using test equipment. Also, this section
specifies the RECOMMENDED values for certain parameters. The values
are the defaults used in most of the client operating systems
currently.

4.3.1.1. TCP Stack Attributes

The TCP stack SHOULD use a congestion control algorithm at client
and server endpoints. The IPv4 and IPv6 Maximum Segment Size (MSS)

SHOULD be set to 1460 bytes and 1440 bytes respectively and a TX and
RX initial receive windows of 64 KByte. Client initial congestion
window SHOULD NOT exceed 10 times the MSS. Delayed ACKs are
permitted and the maximum client delayed ACK SHOULD NOT exceed 10
times the MSS before a forced ACK. Up to three retries SHOULD be
allowed before a timeout event is declared. All traffic MUST set the
TCP PSH flag to high. The source port range SHOULD be in the range
of 1024 - 65535. Internal timeout SHOULD be dynamically scalable per
RFC 793. The client SHOULD initiate and close TCP connections. The
TCP connection MUST be initiated via a TCP three-way handshake (SYN,
SYN/ACK, ACK), and it MUST be closed via either a TCP three-way
close (FIN, FIN/ACK, ACK), or a TCP four-way close (FIN, ACK, FIN,
ACK).

4.3.1.2. **Client IP Address Space**

The sum of the client IP space SHOULD contain the following
attributes.

  *The IP blocks SHOULD consist of multiple unique, discontinuous
   static address blocks.

  *A default gateway is permitted.

  *The DSCP (differentiated services code point) marking is set to
   DF (Default Forwarding) '000000' on IPv4 Type of Service (ToS)
   field and IPv6 traffic class field.

The following equation can be used to define the total number of
client IP addresses that will be configured on the test equipment.

Desired total number of client IP = Target throughput [Mbit/s] /
Average throughput per IP address [Mbit/s]

As shown in the example list below, the value for "Average
throughput per IP address" can be varied depending on the deployment
and use case scenario.

(Option 1) DUT/SUT deployment scenario 1 : 6-7 Mbit/s per IP (e.g.
     1,400-1,700 IPs per 10Gbit/s throughput)

(Option 2) DUT/SUT deployment scenario 2 : 0.1-0.2 Mbit/s per IP
     (e.g. 50,000-100,000 IPs per 10Gbit/s throughput)

Based on deployment and use case scenario, client IP addresses
SHOULD be distributed between IPv4 and IPv6. The following options
MAY be considered for a selection of traffic mix ratio.

(Option 1) 100 % IPv4, no IPv6

(Option 2)   80 % IPv4, 20% IPv6

(Option 3) 50 % IPv4, 50% IPv6

(Option 4) 20 % IPv4, 80% IPv6

(Option 5) no IPv4, 100% IPv6

Note: The IANA has assigned IP address range for the testing purpose as described in Section 8. If the test scenario requires more IP addresses or subnets than the IANA assigned, this document recommends using non routable Private IPv4 address ranges or Unique Local Address (ULA) IPv6 address ranges for the testing.

4.3.1.3. Emulated Web Browser Attributes

The client emulated web browser (emulated browser) contains attributes that will materially affect how traffic is loaded. The objective is to emulate modern, typical browser attributes to improve realism of the result set.

For HTTP traffic emulation, the emulated browser MUST negotiate HTTP version 1.1 or higher. Depending on test scenarios and chosen HTTP version, the emulated browser MAY open multiple TCP connections per Server endpoint IP at any time depending on how many sequential transactions need to be processed. For HTTP/2 or HTTP/3, the emulated browser MAY open multiple concurrent streams per connection (multiplexing). If HTTP/3 is used the emulated browser MUST open Quick UDP Internet Connections (QUIC). HTTP settings such as number of connection per server IP, number of requests per connection, and number of streams per connection MUST be documented. This document refers to [RFC8446] for HTTP/2 and [RFC9000]for QUIC. The emulated browser SHOULD advertise a User-Agent header. The emulated browser SHOULD enforce content length validation. Depending on test scenarios and selected HTTP version, HTTP header compression MAY be set to enable or disable. This setting (compression enabled or disabled) MUST be documented in the report.

For encrypted traffic, the following attributes SHALL define the negotiated encryption parameters. The test clients MUST use TLS version 1.2 or higher. TLS record size MAY be optimized for the HTTPS response object size up to a record size of 16 KByte. If Server Name Indication (SNI) is required in the traffic mix profile, the client endpoint MUST send TLS extension Server Name Indication (SNI) information when opening a security tunnel. Each client connection MUST perform a full handshake with server certificate and MUST NOT use session reuse or resumption.

The following TLS 1.2 supported ciphers and keys are RECOMMENDED to use for HTTPS based benchmarking tests defined in [Section 7](#).

1. ECDHE-ECDSA-AES128-GCM-SHA256 with Prime256v1 (Signature Hash Algorithm: ecdsa_secp256r1_sha256 and Supported group: sepc256r1)

2. ECDHE-RSA-AES128-GCM-SHA256 with RSA 2048 (Signature Hash Algorithm: rsa_pkcs1_sha256 and Supported group: sepc256r1)

3. ECDHE-ECDSA-AES256-GCM-SHA384 with Secp521 (Signature Hash Algorithm: ecdsa_secp384r1_sha384 and Supported group: sepc521r1)

4. ECDHE-RSA-AES256-GCM-SHA384 with RSA 4096 (Signature Hash Algorithm: rsa_pkcs1_sha384 and Supported group: secp256r1)

Note: The above ciphers and keys were those commonly used enterprise grade encryption cipher suites for TLS 1.2. It is recognized that these will evolve over time. Individual certification bodies SHOULD use ciphers and keys that reflect evolving use cases. These choices MUST be documented in the resulting test reports with detailed information on the ciphers and keys used along with reasons for the choices.

[RFC8446] defines the following cipher suites for use with TLS 1.3.

1. TLS_AES_128_GCM_SHA256

2. TLS_AES_256_GCM_SHA384

3. TLS_CHACHA20_POLY1305_SHA256

4. TLS_AES_128_CCM_SHA256

5. TLS_AES_128_CCM_8_SHA256

4.3.2. Backend Server Configuration

This section specifies which parameters should be considered while configuring emulated backend servers using test equipment.

4.3.2.1. TCP Stack Attributes

The TCP stack on the server side SHOULD be configured similar to the client side configuration described in [Section 4.3.1.1](#). In addition, server initial congestion window MUST NOT exceed 10 times the MSS. Delayed ACKs are permitted and the maximum server delayed ACK MUST NOT exceed 10 times the MSS before a forced ACK.

4.3.2.2. Server Endpoint IP Addressing

The sum of the server IP space SHOULD contain the following
attributes.

  *The server IP blocks SHOULD consist of unique, discontinuous
   static address blocks with one IP per server Fully Qualified
   Domain Name (FQDN) endpoint per test port.

  *A default gateway is permitted. The DSCP (differentiated services
   code point) marking is set to DF (Default Forwarding) '000000' on
   IPv4 Type of Service (ToS) field and IPv6 traffic class field.

  *The server IP addresses SHOULD be distributed between IPv4 and
   IPv6 with a ratio identical to the clients distribution ratio.

Note: The IANA has assigned IP address range for the testing purpose
as described in Section 8. If the test scenario requires more IP
addresses or subnets than the IANA assigned, this document
recommends using non routable Private IPv4 address ranges or Unique
Local Address (ULA) IPv6 address ranges for the testing.

4.3.2.3. HTTP / HTTPS Server Pool Endpoint Attributes

The server pool for HTTP SHOULD listen on TCP port 80 and emulate
the same HTTP version and settings chosen by the client (emulated
web browser). The Server MUST advertise server type in the Server
response header [RFC2616]. For HTTPS server, TLS 1.2 or higher MUST
be used with a maximum record size of 16 KByte and MUST NOT use
ticket resumption or session ID reuse. The server SHOULD listen on
TCP port 443. The server SHALL serve a certificate to the client.
The HTTPS server MUST check host SNI information with the FQDN if
SNI is in use. Cipher suite and key size on the server side MUST be
configured similar to the client side configuration described in
Section 4.3.1.3.

4.3.3. Traffic Flow Definition

This section describes the traffic pattern between client and server
endpoints. At the beginning of the test, the server endpoint
initializes and will be ready to accept connection states including
initialization of the TCP stack as well as bound HTTP and HTTPS
servers. When a client endpoint is needed, it will initialize and be
given attributes such as a MAC and IP address. The behavior of the
client is to sweep through the given server IP space, generating a
recognizable service by the DUT. Sequential and pseudorandom sweep
methods are acceptable. The method used MUST be stated in the final
report. Thus, a balanced\ mesh between client endpoints and server
endpoints will be generated in a client IP and port to server IP and
port combination. Each client endpoint performs the same actions as

other endpoints, with the difference being the source IP of the
client endpoint and the target server IP pool. The client MUST use
the server IP address or FQDN in the host header [RFC2616].

4.3.3.1. Description of Intra-Client Behavior

Client endpoints are independent of other clients that are
concurrently executing. When a client endpoint initiates traffic,
this section describes how the client steps through different
services. Once the test is initialized, the client endpoints
randomly hold (perform no operation) for a few milliseconds for
better randomization of the start of client traffic. Each client
will either open a new TCP connection or connect to a TCP
persistence stack still open to that specific server. At any point
that the traffic profile may require encryption, a TLS encryption
tunnel will form presenting the URL or IP address request to the
server. If using SNI, the server MUST then perform an SNI name check
with the proposed FQDN compared to the domain embedded in the
certificate. Only when correct, will the server process the HTTPS
response object. The initial response object to the server is based
on benchmarking tests described in Section 7. Multiple additional
sub-URLs (response objects on the service page) MAY be requested
simultaneously. This MAY be to the same server IP as the initial
URL. Each sub-object will also use a canonical FQDN and URL path, as
observed in the traffic mix used.

4.3.4. Traffic Load Profile

The loading of traffic is described in this section. The loading of
a traffic load profile has five phases: Init, ramp up, sustain, ramp
down, and collection.

1. Init phase: Testbed devices including the client and server
   endpoints should negotiate layer 2-3 connectivity such as MAC
   learning and ARP. Only after successful MAC learning or ARP/ND
   resolution SHALL the test iteration move to the next phase. No
   measurements are made in this phase. The minimum RECOMMENDED
   time for Init phase is 5 seconds. During this phase, the
   emulated clients SHOULD NOT initiate any sessions with the DUT/
   SUT, in contrast, the emulated servers should be ready to
   accept requests from DUT/SUT or from emulated clients.

2. Ramp up phase: The test equipment SHOULD start to generate the
   test traffic. It SHOULD use a set of the approximate number of
   unique client IP addresses to generate traffic. The traffic
   SHOULD ramp up from zero to desired target objective. The
   target objective is defined for each benchmarking test. The
   duration for the ramp up phase MUST be configured long enough
   that the test equipment does not overwhelm the DUT/SUTs stated

performance metrics defined in [Section 6.3](#) namely, TCP Connections Per Second, Inspected Throughput, Concurrent TCP Connections, and Application Transactions Per Second. No measurements are made in this phase.

3. Sustain phase: Starts when all required clients are active and operating at their desired load condition. In the sustain phase, the test equipment SHOULD continue generating traffic to constant target value for a constant number of active clients. The minimum RECOMMENDED time duration for sustain phase is 300 seconds. This is the phase where measurements occur. The test equipment SHOULD measure and record statistics continuously. The sampling interval for collecting the raw results and calculating the statistics SHOULD be less than 2 seconds.

4. Ramp down phase: No new connections are established, and no measurements are made. The time duration for ramp up and ramp down phase SHOULD be the same.

5. Collection phase: The last phase is administrative and will occur when the test equipment merges and collates the report data.

5. Testbed Considerations

This section describes steps for a reference test (pre-test) that control the test environment including test equipment, focusing on physical and virtualized environments and as well as test equipments. Below are the RECOMMENDED steps for the reference test.

1. Perform the reference test either by configuring the DUT/SUT in the most trivial setup (fast forwarding) or without presence of the DUT/SUT.

2. Generate traffic from traffic generator. Choose a traffic profile used for HTTP or HTTPS throughput performance test with smallest object size.

3. Ensure that any ancillary switching or routing functions added in the test equipment does not limit the performance by introducing network metrics such as packet loss and latency. This is specifically important for virtualized components (e.g., vSwitches, vRouters).

4. Verify that the generated traffic (performance) of the test equipment matches and reasonably exceeds the expected maximum performance of the DUT/SUT.

5. Record the network performance metrics packet loss latency introduced by the test environment (without DUT/SUT).

6. Assert that the testbed characteristics are stable during the
      entire test session. Several factors might influence stability
      specifically, for virtualized testbeds. For example, additional
      workloads in a virtualized system, load balancing, and movement
      of virtual machines during the test, or simple issues such as
      additional heat created by high workloads leading to an
      emergency CPU performance reduction.

   The reference test SHOULD be performed before the benchmarking tests
   (described in section 7) start.

6. Reporting

   This section describes how the benchmarking test report should be
   formatted and presented. It is RECOMMENDED to include two main
   sections in the report, namely the introduction and the detailed
   test results sections.

6.1. Introduction

   The following attributes SHOULD be present in the introduction
   section of the test report.

   1. The time and date of the execution of the tests

   2. Summary of testbed software and hardware details

        a. DUT/SUT hardware/virtual configuration

             *This section SHOULD clearly identify the make and model
              of the DUT/SUT

             *The port interfaces, including speed and link
              information

             *If the DUT/SUT is a Virtual Network Function (VNF),
              host (server) hardware and software details, interface
              acceleration type such as DPDK and SR-IOV, used CPU
              cores, used RAM, resource sharing (e.g. Pinning details
              and NUMA Node) configuration details, hypervisor
              version, virtual switch version

             *details of any additional hardware relevant to the DUT/
              SUT such as controllers

        b. DUT/SUT software

             *Operating system name

             *Version

*Specific configuration details (if any)

     c. DUT/SUT enabled features

          *Configured DUT/SUT features (see Table 1 and Table 2)

          *Attributes of the above-mentioned features

          *Any additional relevant information about the features

     d. Test equipment hardware and software

          *Test equipment vendor name

          *Hardware details including model number, interface type

          *Test equipment firmware and test application software
           version

     e. Key test parameters

          *Used cipher suites and keys

          *IPv4 and IPv6 traffic distribution

          *Number of configured ACL

     f. Details of application traffic mix used in the
        benchmarking test ["Throughput Performance with Application
        Traffic Mix"](Section 7.1) ([Section 7.1](#))

          *Name of applications and layer 7 protocols

          *Percentage of emulated traffic for each application and
           layer 7 protocols

          *Percentage of encrypted traffic and used cipher suites
           and keys (The RECOMMENDED ciphers and keys are defined
           in [Section 4.3.1.3](#))

          *Used object sizes for each application and layer 7
           protocols

3. Results Summary / Executive Summary

     a. Results SHOULD resemble a pyramid in how it is reported,
        with the introduction section documenting the summary of
        results in a prominent, easy to read block.

6.2. Detailed Test Results

In the result section of the test report, the following attributes
SHOULD be present for each benchmarking test.

   a. KPIs MUST be documented separately for each benchmarking test.
      The format of the KPI metrics SHOULD be presented as described
      in Section 6.3.

   b. The next level of details SHOULD be graphs showing each of
      these metrics over the duration (sustain phase) of the test.
      This allows the user to see the measured performance stability
      changes over time.

6.3. Benchmarks and Key Performance Indicators

This section lists key performance indicators (KPIs) for overall
benchmarking tests. All KPIs MUST be measured during the sustain
phase of the traffic load profile described in Section 4.3.4. All
KPIs MUST be measured from the result output of test equipment.

   *Concurrent TCP Connections

     The aggregate number of simultaneous connections between hosts
     across the DUT/SUT, or between hosts and the DUT/SUT (defined in
     [RFC2647]).

   *TCP Connections Per Second

     The average number of successfully established TCP connections
     per second between hosts across the DUT/SUT, or between hosts and
     the DUT/SUT. The TCP connection MUST be initiated via a TCP
     three-way handshake (SYN, SYN/ACK, ACK). Then the TCP session
     data is sent. The TCP session MUST be closed via either a TCP
     three-way close (FIN, FIN/ACK, ACK), or a TCP four-way close
     (FIN, ACK, FIN, ACK), and MUST NOT by RST.

   *Application Transactions Per Second

     The average number of successfully completed transactions per
     second. For a particular transaction to be considered successful,
     all data MUST have been transferred in its entirety. In case of
     HTTP(S) transactions, it MUST have a valid status code (200 OK),
     and the appropriate FIN, FIN/ACK sequence MUST have been
     completed.

   *TLS Handshake Rate

The average number of successfully established TLS connections
per second between hosts across the DUT/SUT, or between hosts and
the DUT/SUT.

  *Inspected Throughput

   The number of bits per second of examined and allowed traffic a
   network security device is able to transmit to the correct
   destination interface(s) in response to a specified offered load.
   The throughput benchmarking tests defined in Section 7 SHOULD
   measure the average Layer 2 throughput value when the DUT/SUT is
   "inspecting" traffic. This document recommends presenting the
   inspected throughput value in Gbit/s rounded to two places of
   precision with a more specific Kbit/s in parenthesis.

  *Time to First Byte (TTFB)

   TTFB is the elapsed time between the start of sending the TCP SYN
   packet from the client and the client receiving the first packet
   of application data from the server or DUT/SUT. The benchmarking
   tests HTTP Transaction Latency (Section 7.4) and HTTPS
   Transaction Latency (Section 7.8) measure the minimum, average
   and maximum TTFB. The value SHOULD be expressed in milliseconds.

  *URL Response time / Time to Last Byte (TTLB)

   URL Response time / TTLB is the elapsed time between the start of
   sending the TCP SYN packet from the client and the client
   receiving the last packet of application data from the server or
   DUT/SUT. The benchmarking tests HTTP Transaction Latency (Section
   7.4) and HTTPS Transaction Latency (Section 7.8) measure the
   minimum, average and maximum TTLB. The value SHOULD be expressed
   in millisecond.

7. Benchmarking Tests

7.1. Throughput Performance with Application Traffic Mix

7.1.1. Objective

   Using a relevant application traffic mix, determine the sustainable
   inspected throughput supported by the DUT/SUT.

   Based on customer use case, users can choose the relevant
   application traffic mix for this test. The details about the traffic
   mix MUST be documented in the report. At least the following traffic
   mix details MUST be documented and reported together with the test
   results:

      Name of applications and layer 7 protocols

Percentage of emulated traffic for each application and layer 7
protocol

Percentage of encrypted traffic and used cipher suites and keys
(The RECOMMENDED ciphers and keys are defined in Section
4.3.1.3.)

Used object sizes for each application and layer 7 protocols

7.1.2. Test Setup

Testbed setup MUST be configured as defined in Section 4. Any
benchmarking test specific testbed configuration changes MUST be
documented.

7.1.3. Test Parameters

In this section, the benchmarking test specific parameters SHOULD be
defined.

7.1.3.1. DUT/SUT Configuration Parameters

DUT/SUT parameters MUST conform to the requirements defined in
Section 4.2. Any configuration changes for this specific
benchmarking test MUST be documented. In case the DUT/SUT is
configured without SSL inspection, the test report MUST explain the
implications of this to the relevant application traffic mix
encrypted traffic.

7.1.3.2. Test Equipment Configuration Parameters

Test equipment configuration parameters MUST conform to the
requirements defined in Section 4.3. The following parameters MUST
be documented for this benchmarking test:

Client IP address range defined in Section 4.3.1.2

Server IP address range defined in Section 4.3.2.2

Traffic distribution ratio between IPv4 and IPv6 defined in
Section 4.3.1.2

Target inspected throughput: Aggregated line rate of interface(s)
used in the DUT/SUT or the value defined based on requirement for
a specific deployment scenario

Initial throughput: 10% of the "Target inspected throughput"
Note: Initial throughput is not a KPI to report. This value is
configured on the traffic generator and used to perform Step 1:

"Test Initialization and Qualification" described under the
[Section 7.1.4](#).

One of the ciphers and keys defined in [Section 4.3.1.3](#) are
RECOMMENDED to use for this benchmarking test.

7.1.3.3. Traffic Profile

Traffic profile: This test MUST be run with a relevant application
traffic mix profile.

7.1.3.4. Test Results Validation Criteria

The following criteria are the test results validation criteria. The
test results validation criteria MUST be monitored during the whole
sustain phase of the traffic load profile.

   a. Number of failed application transactions (receiving any HTTP
      response code other than 200 OK) MUST be less than 0.001% (1
      out of 100,000 transactions) of total attempted transactions.

   b. Number of Terminated TCP connections due to unexpected TCP RST
      sent by DUT/SUT MUST be less than 0.001% (1 out of 100,000
      connections) of total initiated TCP connections.

7.1.3.5. Measurement

Following KPI metrics MUST be reported for this benchmarking test:

Mandatory KPIs (benchmarks): Inspected Throughput, TTFB (minimum,
average, and maximum), TTLB (minimum, average, and maximum) and
Application Transactions Per Second

Note: TTLB MUST be reported along with the object size used in the
traffic profile.

Optional KPIs: TCP Connections Per Second and TLS Handshake Rate

7.1.4. Test Procedures and Expected Results

The test procedures are designed to measure the inspected throughput
performance of the DUT/SUT at the sustaining period of traffic load
profile. The test procedure consists of three major steps: Step 1
ensures the DUT/SUT is able to reach the performance value (initial
throughput) and meets the test results validation criteria when it
was very minimally utilized. Step 2 determines the DUT/SUT is able
to reach the target performance value within the test results
validation criteria. Step 3 determines the maximum achievable
performance value within the test results validation criteria.

This test procedure MAY be repeated multiple times with different IP types: IPv4 only, IPv6 only, and IPv4 and IPv6 mixed traffic distribution.

7.1.4.1. Step 1: Test Initialization and Qualification

Verify the link status of all connected physical interfaces. All interfaces are expected to be in "UP" status.

Configure traffic load profile of the test equipment to generate test traffic at the "Initial throughput" rate as described in Section 7.1.3.2. The test equipment SHOULD follow the traffic load profile definition as described in Section 4.3.4. The DUT/SUT SHOULD reach the "Initial throughput" during the sustain phase. Measure all KPI as defined in Section 7.1.3.5. The measured KPIs during the sustain phase MUST meet all the test results validation criteria defined in Section 7.1.3.4.

If the KPI metrics do not meet the test results validation criteria, the test procedure MUST NOT be continued to step 2.

7.1.4.2. Step 2: Test Run with Target Objective

Configure test equipment to generate traffic at the "Target inspected throughput" rate defined in Section 7.1.3.2. The test equipment SHOULD follow the traffic load profile definition as described in Section 4.3.4. The test equipment SHOULD start to measure and record all specified KPIs. Continue the test until all traffic profile phases are completed.

Within the test results validation criteria, the DUT/SUT is expected to reach the desired value of the target objective ("Target inspected throughput") in the sustain phase. Follow step 3, if the measured value does not meet the target value or does not fulfill the test results validation criteria.

7.1.4.3. Step 3: Test Iteration

Determine the achievable average inspected throughput within the test results validation criteria. Final test iteration MUST be performed for the test duration defined in Section 4.3.4.

7.2. TCP/HTTP Connections Per Second

7.2.1. Objective

Using HTTP traffic, determine the sustainable TCP connection establishment rate supported by the DUT/SUT under different throughput load conditions.

To measure connections per second, test iterations MUST use different fixed HTTP response object sizes (the different load conditions) defined in Section 7.2.3.2.

7.2.2. Test Setup

Testbed setup SHOULD be configured as defined in Section 4. Any specific testbed configuration changes (number of interfaces and interface type, etc.) MUST be documented.

7.2.3. Test Parameters

In this section, benchmarking test specific parameters SHOULD be defined.

7.2.3.1. DUT/SUT Configuration Parameters

DUT/SUT parameters MUST conform to the requirements defined in Section 4.2. Any configuration changes for this specific benchmarking test MUST be documented.

7.2.3.2. Test Equipment Configuration Parameters

Test equipment configuration parameters MUST conform to the requirements defined in Section 4.3. The following parameters MUST be documented for this benchmarking test:

Client IP address range defined in Section 4.3.1.2

Server IP address range defined in Section 4.3.2.2

Traffic distribution ratio between IPv4 and IPv6 defined in Section 4.3.1.2

Target connections per second: Initial value from product datasheet or the value defined based on requirement for a specific deployment scenario

Initial connections per second: 10% of "Target connections per second" (Note: Initial connections per second is not a KPI to report. This value is configured on the traffic generator and used to perform the Step1: "Test Initialization and Qualification" described under the Section 7.2.4.

The client SHOULD negotiate HTTP and close the connection with FIN immediately after completion of one transaction. In each test iteration, client MUST send GET request requesting a fixed HTTP response object size.

The RECOMMENDED response object sizes are 1, 2, 4, 16, and 64 KByte.

7.2.3.3. Test Results Validation Criteria

The following criteria are the test results validation criteria. The Test results validation criteria MUST be monitored during the whole sustain phase of the traffic load profile.

   a. Number of failed application transactions (receiving any HTTP response code other than 200 OK) MUST be less than 0.001% (1 out of 100,000 transactions) of total attempted transactions.

   b. Number of terminated TCP connections due to unexpected TCP RST sent by DUT/SUT MUST be less than 0.001% (1 out of 100,000 connections) of total initiated TCP connections.

   c. During the sustain phase, traffic SHOULD be forwarded at a constant rate (considered as a constant rate if any deviation of traffic forwarding rate is less than 5%).

   d. Concurrent TCP connections MUST be constant during steady state and any deviation of concurrent TCP connections SHOULD be less than 10%. This confirms the DUT opens and closes TCP connections at approximately the same rate.

7.2.3.4. Measurement

TCP Connections Per Second MUST be reported for each test iteration (for each object size).

7.2.4. Test Procedures and Expected Results

The test procedure is designed to measure the TCP connections per second rate of the DUT/SUT at the sustaining period of the traffic load profile. The test procedure consists of three major steps: Step 1 ensures the DUT/SUT is able to reach the performance value (Initial connections per second) and meets the test results validation criteria when it was very minimally utilized. Step 2 determines the DUT/SUT is able to reach the target performance value within the test results validation criteria. Step 3 determines the maximum achievable performance value within the test results validation criteria.

This test procedure MAY be repeated multiple times with different IP types: IPv4 only, IPv6 only, and IPv4 and IPv6 mixed traffic distribution.

7.2.4.1. Step 1: Test Initialization and Qualification

Verify the link status of all connected physical interfaces. All interfaces are expected to be in "UP" status.

Configure the traffic load profile of the test equipment to
establish "Initial connections per second" as defined in Section
7.2.3.2. The traffic load profile SHOULD be defined as described in
Section 4.3.4.

The DUT/SUT SHOULD reach the "Initial connections per second" before
the sustain phase. The measured KPIs during the sustain phase MUST
meet all the test results validation criteria defined in Section
7.2.3.3.

If the KPI metrics do not meet the test results validation criteria,
the test procedure MUST NOT continue to "Step 2".

7.2.4.2. Step 2: Test Run with Target Objective

Configure test equipment to establish the target objective ("Target
connections per second") defined in Section 7.2.3.2. The test
equipment SHOULD follow the traffic load profile definition as
described in Section 4.3.4.

During the ramp up and sustain phase of each test iteration, other
KPIs such as inspected throughput, concurrent TCP connections and
application transactions per second MUST NOT reach the maximum value
the DUT/SUT can support. The test results for specific test
iterations SHOULD NOT be reported, if the above-mentioned KPI
(especially inspected throughput) reaches the maximum value.
(Example: If the test iteration with 64 KByte of HTTP response
object size reached the maximum inspected throughput limitation of
the DUT/SUT, the test iteration MAY be interrupted and the result
for 64 KByte SHOULD NOT be reported.)

The test equipment SHOULD start to measure and record all specified
KPIs. Continue the test until all traffic profile phases are
completed.

Within the test results validation criteria, the DUT/SUT is expected
to reach the desired value of the target objective ("Target
connections per second") in the sustain phase. Follow step 3, if the
measured value does not meet the target value or does not fulfill
the test results validation criteria.

7.2.4.3. Step 3: Test Iteration

Determine the achievable TCP connections per second within the test
results validation criteria.

7.3. HTTP Throughput

7.3.1. Objective

Determine the sustainable inspected throughput of the DUT/SUT for HTTP transactions varying the HTTP response object size.

7.3.2. Test Setup

Testbed setup SHOULD be configured as defined in Section 4. Any specific testbed configuration changes (number of interfaces and interface type, etc.) MUST be documented.

7.3.3. Test Parameters

In this section, benchmarking test specific parameters SHOULD be defined.

7.3.3.1. DUT/SUT Configuration Parameters

DUT/SUT parameters MUST conform to the requirements defined in Section 4.2. Any configuration changes for this specific benchmarking test MUST be documented.

7.3.3.2. Test Equipment Configuration Parameters

Test equipment configuration parameters MUST conform to the requirements defined in Section 4.3. The following parameters MUST be documented for this benchmarking test:

Client IP address range defined in Section 4.3.1.2

Server IP address range defined in Section 4.3.2.2

Traffic distribution ratio between IPv4 and IPv6 defined in Section 4.3.1.2

Target inspected throughput: Aggregated line rate of interface(s) used in the DUT/SUT or the value defined based on requirement for a specific deployment scenario

Initial throughput: 10% of "Target inspected throughput" Note: Initial throughput is not a KPI to report. This value is configured on the traffic generator and used to perform Step 1: "Test Initialization and Qualification" described under Section 7.3.4.

Number of HTTP response object requests (transactions) per connection: 10

RECOMMENDED HTTP response object size: 1, 16, 64, 256 KByte, and
mixed objects defined in table 5.

```
+--------------------+--------------------+
| Object size (KByte) | Number of requests/ |
|                    | Weight             |
+--------------------+--------------------+
| 0.2               | 1                  |
+--------------------+--------------------+
| 6                 | 1                  |
+--------------------+--------------------+
| 8                 | 1                  |
+--------------------+--------------------+
| 9                 | 1                  |
+--------------------+--------------------+
| 10                | 1                  |
+--------------------+--------------------+
| 25                | 1                  |
+--------------------+--------------------+
| 26                | 1                  |
+--------------------+--------------------+
| 35                | 1                  |
+--------------------+--------------------+
| 59                | 1                  |
+--------------------+--------------------+
| 347               | 1                  |
+--------------------+--------------------+
```

Figure 7: Table 5: Mixed Objects

7.3.3.3. Test Results Validation Criteria

The following criteria are the test results validation criteria. The
test results validation criteria MUST be monitored during the whole
sustain phase of the traffic load profile.

   a. Number of failed application transactions (receiving any HTTP
      response code other than 200 OK) MUST be less than 0.001% (1
      out of 100,000 transactions) of attempt transactions.

   b. Traffic SHOULD be forwarded at a constant rate (considered as a
      constant rate if any deviation of traffic forwarding rate is
      less than 5%).

   c. Concurrent TCP connections MUST be constant during steady state
      and any deviation of concurrent TCP connections SHOULD be less
      than 10%. This confirms the DUT opens and closes TCP
      connections at approximately the same rate.

7.3.3.4. Measurement

Inspected Throughput and HTTP Transactions per Second MUST be reported for each object size.

7.3.4. Test Procedures and Expected Results

The test procedure is designed to measure HTTP throughput of the DUT/ SUT. The test procedure consists of three major steps: Step 1 ensures the DUT/SUT is able to reach the performance value (Initial throughput) and meets the test results validation criteria when it was very minimal utilized. Step 2 determines the DUT/SUT is able to reach the target performance value within the test results validation criteria. Step 3 determines the maximum achievable performance value within the test results validation criteria.

This test procedure MAY be repeated multiple times with different IPv4 and IPv6 traffic distribution and HTTP response object sizes.

7.3.4.1. Step 1: Test Initialization and Qualification

Verify the link status of all connected physical interfaces. All interfaces are expected to be in "UP" status.

Configure traffic load profile of the test equipment to establish "Initial inspected throughput" as defined in Section 7.3.3.2.

The traffic load profile SHOULD be defined as described in Section 4.3.4. The DUT/SUT SHOULD reach the "Initial inspected throughput" during the sustain phase. Measure all KPI as defined in Section 7.3.3.4.

The measured KPIs during the sustain phase MUST meet the test results validation criteria "a" defined in Section 7.3.3.3. The test results validation criteria "b" and "c" are OPTIONAL for step 1.

If the KPI metrics do not meet the test results validation criteria, the test procedure MUST NOT be continued to "Step 2".

7.3.4.2. Step 2: Test Run with Target Objective

Configure test equipment to establish the target objective ("Target inspected throughput") defined in Section 7.3.3.2. The test equipment SHOULD start to measure and record all specified KPIs. Continue the test until all traffic profile phases are completed.

Within the test results validation criteria, the DUT/SUT is expected to reach the desired value of the target objective in the sustain phase. Follow step 3, if the measured value does not meet the target value or does not fulfill the test results validation criteria.

7.3.4.3. Step 3: Test Iteration

Determine the achievable inspected throughput within the test results validation criteria and measure the KPI metric Transactions per Second. Final test iteration MUST be performed for the test duration defined in Section 4.3.4.

7.4. HTTP Transaction Latency

7.4.1. Objective

Using HTTP traffic, determine the HTTP transaction latency when DUT is running with sustainable HTTP transactions per second supported by the DUT/SUT under different HTTP response object sizes.

Test iterations MUST be performed with different HTTP response object sizes in two different scenarios. One with a single transaction and the other with multiple transactions within a single TCP connection. For consistency both the single and multiple transaction test MUST be configured with the same HTTP version

Scenario 1: The client MUST negotiate HTTP and close the connection with FIN immediately after completion of a single transaction (GET and RESPONSE).

Scenario 2: The client MUST negotiate HTTP and close the connection FIN immediately after completion of 10 transactions (GET and RESPONSE) within a single TCP connection.

7.4.2. Test Setup

Testbed setup SHOULD be configured as defined in Section 4. Any specific testbed configuration changes (number of interfaces and interface type, etc.) MUST be documented.

7.4.3. Test Parameters

In this section, benchmarking test specific parameters SHOULD be defined.

7.4.3.1. DUT/SUT Configuration Parameters

DUT/SUT parameters MUST conform to the requirements defined in Section 4.2. Any configuration changes for this specific benchmarking test MUST be documented.

7.4.3.2. Test Equipment Configuration Parameters

Test equipment configuration parameters MUST conform to the requirements defined in Section 4.3. The following parameters MUST be documented for this benchmarking test:

Client IP address range defined in Section 4.3.1.2

Server IP address range defined in Section 4.3.2.2

Traffic distribution ratio between IPv4 and IPv6 defined in Section 4.3.1.2

Target objective for scenario 1: 50% of the connections per second measured in benchmarking test TCP/HTTP Connections Per Second (Section 7.2)

Target objective for scenario 2: 50% of the inspected throughput measured in benchmarking test HTTP Throughput (Section 7.3)

Initial objective for scenario 1: 10% of "Target objective for scenario 1"

Initial objective for scenario 2: 10% of "Target objective for scenario 2"

Note: The Initial objectives are not a KPI to report. These values are configured on the traffic generator and used to perform the Step1: "Test Initialization and Qualification" described under the Section 7.4.4.

HTTP transaction per TCP connection: Test scenario 1 with single transaction and test scenario 2 with 10 transactions.

HTTP with GET request requesting a single object. The RECOMMENDED object sizes are 1, 16, and 64 KByte. For each test iteration, client MUST request a single HTTP response object size.

7.4.3.3. Test Results Validation Criteria

The following criteria are the test results validation criteria. The Test results validation criteria MUST be monitored during the whole sustain phase of the traffic load profile.

   a. Number of failed application transactions (receiving any HTTP response code other than 200 OK) MUST be less than 0.001% (1 out of 100,000 transactions) of attempt transactions.

b. Number of terminated TCP connections due to unexpected TCP RST sent by DUT/SUT MUST be less than 0.001% (1 out of 100,000 connections) of total initiated TCP connections.

c. During the sustain phase, traffic SHOULD be forwarded at a constant rate (considered as a constant rate if any deviation of traffic forwarding rate is less than 5%).

d. Concurrent TCP connections MUST be constant during steady state and any deviation of concurrent TCP connections SHOULD be less than 10%. This confirms the DUT opens and closes TCP connections at approximately the same rate.

e. After ramp up the DUT MUST achieve the "Target objective" defined in Section 7.4.3.2 and remain in that state for the entire test duration (sustain phase).

7.4.3.4. Measurement

TTFB (minimum, average, and maximum) and TTLB (minimum, average and maximum) MUST be reported for each object size.

7.4.4. Test Procedures and Expected Results

The test procedure is designed to measure TTFB or TTLB when the DUT/SUT is operating close to 50% of its maximum achievable connections per second or inspected throughput. The test procedure consists of two major steps: Step 1 ensures the DUT/SUT is able to reach the initial performance values and meets the test results validation criteria when it was very minimally utilized. Step 2 measures the latency values within the test results validation criteria.

This test procedure MAY be repeated multiple times with different IP types (IPv4 only, IPv6 only and IPv4 and IPv6 mixed traffic distribution), HTTP response object sizes and single and multiple transactions per connection scenarios.

7.4.4.1. Step 1: Test Initialization and Qualification

Verify the link status of all connected physical interfaces. All interfaces are expected to be in "UP" status.

Configure traffic load profile of the test equipment to establish "Initial objective" as defined in Section 7.4.3.2. The traffic load profile SHOULD be defined as described in Section 4.3.4.

The DUT/SUT SHOULD reach the "Initial objective" before the sustain phase. The measured KPIs during the sustain phase MUST meet all the test results validation criteria defined in Section 7.4.3.3.

If the KPI metrics do not meet the test results validation criteria,
the test procedure MUST NOT be continued to "Step 2".

7.4.4.2. Step 2: Test Run with Target Objective

Configure test equipment to establish "Target objective" defined in
Section 7.4.3.2. The test equipment SHOULD follow the traffic load
profile definition as described in Section 4.3.4.

The test equipment SHOULD start to measure and record all specified
KPIs. Continue the test until all traffic profile phases are
completed.

Within the test results validation criteria, the DUT/SUT MUST reach
the desired value of the target objective in the sustain phase.

Measure the minimum, average, and maximum values of TTFB and TTLB.

7.5. Concurrent TCP/HTTP Connection Capacity

7.5.1. Objective

Determine the number of concurrent TCP connections that the DUT/ SUT
sustains when using HTTP traffic.

7.5.2. Test Setup

Testbed setup SHOULD be configured as defined in Section 4. Any
specific testbed configuration changes (number of interfaces and
interface type, etc.) MUST be documented.

7.5.3. Test Parameters

In this section, benchmarking test specific parameters SHOULD be
defined.

7.5.3.1. DUT/SUT Configuration Parameters

DUT/SUT parameters MUST conform to the requirements defined in
Section 4.2. Any configuration changes for this specific
benchmarking test MUST be documented.

7.5.3.2. Test Equipment Configuration Parameters

Test equipment configuration parameters MUST conform to the
requirements defined in Section 4.3. The following parameters MUST
be noted for this benchmarking test:

Client IP address range defined in Section 4.3.1.2

Server IP address range defined in Section 4.3.2.2

Traffic distribution ratio between IPv4 and IPv6 defined in Section 4.3.1.2

Target concurrent connection: Initial value from product datasheet or the value defined based on requirement for a specific deployment scenario.

Initial concurrent connection: 10% of "Target concurrent connection" Note: Initial concurrent connection is not a KPI to report. This value is configured on the traffic generator and used to perform the Step1: "Test Initialization and Qualification" described under the Section 7.5.4.

Maximum connections per second during ramp up phase: 50% of maximum connections per second measured in benchmarking test TCP/HTTP Connections per second (Section 7.2)

Ramp up time (in traffic load profile for "Target concurrent connection"): "Target concurrent connection" / "Maximum connections per second during ramp up phase"

Ramp up time (in traffic load profile for "Initial concurrent connection"): "Initial concurrent connection" / "Maximum connections per second during ramp up phase"

The client MUST negotiate HTTP and each client MAY open multiple concurrent TCP connections per server endpoint IP.

Each client sends 10 GET requests requesting 1 KByte HTTP response object in the same TCP connection (10 transactions/TCP connection) and the delay (think time) between each transaction MUST be X seconds.

X = ("Ramp up time" + "steady state time") /10

The established connections SHOULD remain open until the ramp down phase of the test. During the ramp down phase, all connections SHOULD be successfully closed with FIN.

7.5.3.3. Test Results Validation Criteria

The following criteria are the test results validation criteria. The Test results validation criteria MUST be monitored during the whole sustain phase of the traffic load profile.

a. Number of failed application transactions (receiving any HTTP response code other than 200 OK) MUST be less than 0.001% (1 out of 100,000 transaction) of total attempted transactions.

b. Number of terminated TCP connections due to unexpected TCP RST sent by DUT/SUT MUST be less than 0.001% (1 out of 100,000 connections) of total initiated TCP connections.

c. During the sustain phase, traffic SHOULD be forwarded at a constant rate (considered as a constant rate if any deviation of traffic forwarding rate is less than 5%).

7.5.3.4. Measurement

Average Concurrent TCP Connections MUST be reported for this benchmarking test.

7.5.4. Test Procedures and Expected Results

The test procedure is designed to measure the concurrent TCP connection capacity of the DUT/SUT at the sustaining period of traffic load profile. The test procedure consists of three major steps: Step 1 ensures the DUT/SUT is able to reach the performance value (Initial concurrent connection) and meets the test results validation criteria when it was very minimally utilized. Step 2 determines the DUT/SUT is able to reach the target performance value within the test results validation criteria. Step 3 determines the maximum achievable performance value within the test results validation criteria.

This test procedure MAY be repeated multiple times with different IPv4 and IPv6 traffic distribution.

7.5.4.1. Step 1: Test Initialization and Qualification

Verify the link status of all connected physical interfaces. All interfaces are expected to be in "UP" status.

Configure test equipment to establish "Initial concurrent TCP connections" defined in Section 7.5.3.2. Except ramp up time, the traffic load profile SHOULD be defined as described in Section 4.3.4.

During the sustain phase, the DUT/SUT SHOULD reach the "Initial concurrent TCP connections". The measured KPIs during the sustain phase MUST meet all the test results validation criteria defined in Section 7.5.3.3.

If the KPI metrics do not meet the test results validation criteria, the test procedure MUST NOT be continued to "Step 2".

7.5.4.2. Step 2: Test Run with Target Objective

Configure test equipment to establish the target objective ("Target concurrent TCP connections"). The test equipment SHOULD follow the traffic load profile definition (except ramp up time) as described in [Section 4.3.4](#).

During the ramp up and sustain phase, the other KPIs such as inspected throughput, TCP connections per second, and application transactions per second MUST NOT reach the maximum value the DUT/SUT can support.

The test equipment SHOULD start to measure and record KPIs defined in [Section 7.5.3.4](#). Continue the test until all traffic profile phases are completed.

Within the test results validation criteria, the DUT/SUT is expected to reach the desired value of the target objective in the sustain phase. Follow step 3, if the measured value does not meet the target value or does not fulfill the test results validation criteria.

7.5.4.3. Step 3: Test Iteration

Determine the achievable concurrent TCP connections capacity within the test results validation criteria.

7.6. TCP/HTTPS Connections per Second

7.6.1. Objective

Using HTTPS traffic, determine the sustainable SSL/TLS session establishment rate supported by the DUT/SUT under different throughput load conditions.

Test iterations MUST include common cipher suites and key strengths as well as forward looking stronger keys. Specific test iterations MUST include ciphers and keys defined in [Section 7.6.3.2](#).

For each cipher suite and key strengths, test iterations MUST use a single HTTPS response object size defined in [Section 7.6.3.2](#) to measure connections per second performance under a variety of DUT/SUT security inspection load conditions.

7.6.2. Test Setup

Testbed setup SHOULD be configured as defined in [Section 4](#). Any specific testbed configuration changes (number of interfaces and interface type, etc.) MUST be documented.

7.6.3. Test Parameters

In this section, benchmarking test specific parameters SHOULD be defined.

7.6.3.1. DUT/SUT Configuration Parameters

DUT/SUT parameters MUST conform to the requirements defined in Section 4.2. Any configuration changes for this specific benchmarking test MUST be documented.

7.6.3.2. Test Equipment Configuration Parameters

Test equipment configuration parameters MUST conform to the requirements defined in Section 4.3. The following parameters MUST be documented for this benchmarking test:

Client IP address range defined in Section 4.3.1.2

Server IP address range defined in Section 4.3.2.2

Traffic distribution ratio between IPv4 and IPv6 defined in Section 4.3.1.2

Target connections per second: Initial value from product datasheet or the value defined based on requirement for a specific deployment scenario.

Initial connections per second: 10% of "Target connections per second" Note: Initial connections per second is not a KPI to report. This value is configured on the traffic generator and used to perform the Step1: "Test Initialization and Qualification" described under the Section 7.6.4.

RECOMMENDED ciphers and keys defined in Section 4.3.1.3

The client MUST negotiate HTTPS and close the connection with FIN immediately after completion of one transaction. In each test iteration, client MUST send GET request requesting a fixed HTTPS response object size. The RECOMMENDED object sizes are 1, 2, 4, 16, and 64 KByte.

7.6.3.3. Test Results Validation Criteria

The following criteria are the test results validation criteria. The test results validation criteria MUST be monitored during the whole test duration.

a. Number of failed application transactions (receiving any HTTP response code other than 200 OK) MUST be less than 0.001% (1 out of 100,000 transactions) of attempt transactions.

b. Number of terminated TCP connections due to unexpected TCP RST sent by DUT/SUT MUST be less than 0.001% (1 out of 100,000 connections) of total initiated TCP connections.

c. During the sustain phase, traffic SHOULD be forwarded at a constant rate (considered as a constant rate if any deviation of traffic forwarding rate is less than 5%).

d. Concurrent TCP connections MUST be constant during steady state and any deviation of concurrent TCP connections SHOULD be less than 10%. This confirms the DUT opens and closes TCP connections at approximately the same rate.

7.6.3.4. Measurement

TCP connections per second MUST be reported for each test iteration (for each object size).

The KPI metric TLS Handshake Rate can be measured in the test using 1 KByte object size.

7.6.4. Test Procedures and Expected Results

The test procedure is designed to measure the TCP connections per second rate of the DUT/SUT at the sustaining period of traffic load profile. The test procedure consists of three major steps: Step 1 ensures the DUT/SUT is able to reach the performance value (Initial connections per second) and meets the test results validation criteria when it was very minimally utilized. Step 2 determines the DUT/SUT is able to reach the target performance value within the test results validation criteria. Step 3 determines the maximum achievable performance value within the test results validation criteria.

This test procedure MAY be repeated multiple times with different IPv4 and IPv6 traffic distribution.

7.6.4.1. Step 1: Test Initialization and Qualification

Verify the link status of all connected physical interfaces. All
interfaces are expected to be in "UP" status.

Configure traffic load profile of the test equipment to establish
"Initial connections per second" as defined in Section 7.6.3.2. The
traffic load profile SHOULD be defined as described in Section
4.3.4.

The DUT/SUT SHOULD reach the "Initial connections per second" before
the sustain phase. The measured KPIs during the sustain phase MUST
meet all the test results validation criteria defined in Section
7.6.3.3.

If the KPI metrics do not meet the test results validation criteria,
the test procedure MUST NOT be continued to "Step 2".

7.6.4.2. Step 2: Test Run with Target Objective

Configure test equipment to establish "Target connections per
second" defined in Section 7.6.3.2. The test equipment SHOULD follow
the traffic load profile definition as described in Section 4.3.4.

During the ramp up and sustain phase, other KPIs such as inspected
throughput, concurrent TCP connections, and application transactions
per second MUST NOT reach the maximum value the DUT/SUT can support.
The test results for specific test iteration SHOULD NOT be reported,
if the above mentioned KPI (especially inspected throughput) reaches
the maximum value. (Example: If the test iteration with 64 KByte of
HTTPS response object size reached the maximum inspected throughput
limitation of the DUT, the test iteration MAY be interrupted and the
result for 64 KByte SHOULD NOT be reported).

The test equipment SHOULD start to measure and record all specified
KPIs. Continue the test until all traffic profile phases are
completed.

Within the test results validation criteria, the DUT/SUT is expected
to reach the desired value of the target objective ("Target
connections per second") in the sustain phase. Follow step 3, if the
measured value does not meet the target value or does not fulfill
the test results validation criteria.

7.6.4.3. Step 3: Test Iteration

Determine the achievable connections per second within the test
results validation criteria.

7.7. HTTPS Throughput

7.7.1. Objective

Determine the sustainable inspected throughput of the DUT/SUT for
HTTPS transactions varying the HTTPS response object size.

Test iterations MUST include common cipher suites and key strengths
as well as forward looking stronger keys. Specific test iterations
MUST include the ciphers and keys defined in Section 7.7.3.2.

7.7.2. Test Setup

Testbed setup SHOULD be configured as defined in Section 4. Any
specific testbed configuration changes (number of interfaces and
interface type, etc.) MUST be documented.

7.7.3. Test Parameters

In this section, benchmarking test specific parameters SHOULD be
defined.

7.7.3.1. DUT/SUT Configuration Parameters

DUT/SUT parameters MUST conform to the requirements defined in
Section 4.2. Any configuration changes for this specific
benchmarking test MUST be documented.

7.7.3.2. Test Equipment Configuration Parameters

Test equipment configuration parameters MUST conform to the
requirements defined in Section 4.3. The following parameters MUST
be documented for this benchmarking test:

Client IP address range defined in Section 4.3.1.2

Server IP address range defined in Section 4.3.2.2

Traffic distribution ratio between IPv4 and IPv6 defined in Section
4.3.1.2

Target inspected throughput: Aggregated line rate of interface(s)
used in the DUT/SUT or the value defined based on requirement for a
specific deployment scenario.

Initial throughput: 10% of "Target inspected throughput" Note:
Initial throughput is not a KPI to report. This value is configured
on the traffic generator and used to perform the Step1: "Test
Initialization and Qualification" described under the Section 7.7.4.

Number of HTTPS response object requests (transactions) per
connection: 10

RECOMMENDED ciphers and keys defined in Section 4.3.1.3

RECOMMENDED HTTPS response object size: 1, 16, 64, 256 KByte, and
mixed objects defined in Table 5 under Section 7.3.3.2.

7.7.3.3. Test Results Validation Criteria

The following criteria are the test results validation criteria. The
test results validation criteria MUST be monitored during the whole
sustain phase of the traffic load profile.

   a. Number of failed Application transactions (receiving any HTTP
      response code other than 200 OK) MUST be less than 0.001% (1
      out of 100,000 transactions) of attempt transactions.

   b. Traffic SHOULD be forwarded at a constant rate (considered as a
      constant rate if any deviation of traffic forwarding rate is
      less than 5%).

   c. Concurrent TCP connections MUST be constant during steady state
      and any deviation of concurrent TCP connections SHOULD be less
      than 10%. This confirms the DUT opens and closes TCP
      connections at approximately the same rate.

7.7.3.4. Measurement

Inspected Throughput and HTTP Transactions per Second MUST be
reported for each object size.

7.7.4. Test Procedures and Expected Results

The test procedure consists of three major steps: Step 1 ensures the
DUT/SUT is able to reach the performance value (Initial throughput)
and meets the test results validation criteria when it was very
minimally utilized. Step 2 determines the DUT/SUT is able to reach
the target performance value within the test results validation
criteria. Step 3 determines the maximum achievable performance value
within the test results validation criteria.

This test procedure MAY be repeated multiple times with different
IPv4 and IPv6 traffic distribution and HTTPS response object sizes.

7.7.4.1. Step 1: Test Initialization and Qualification

Verify the link status of all connected physical interfaces. All
interfaces are expected to be in "UP" status.

Configure traffic load profile of the test equipment to establish "Initial throughput" as defined in Section 7.7.3.2.

The traffic load profile SHOULD be defined as described in Section 4.3.4. The DUT/SUT SHOULD reach the "Initial throughput" during the sustain phase. Measure all KPI as defined in Section 7.7.3.4.

The measured KPIs during the sustain phase MUST meet the test results validation criteria "a" defined in Section 7.7.3.3. The test results validation criteria "b" and "c" are OPTIONAL for step 1.

If the KPI metrics do not meet the test results validation criteria, the test procedure MUST NOT be continued to "Step 2".

7.7.4.2. Step 2: Test Run with Target Objective

Configure test equipment to establish the target objective ("Target inspected throughput") defined in Section 7.7.3.2. The test equipment SHOULD start to measure and record all specified KPIs. Continue the test until all traffic profile phases are completed.

Within the test results validation criteria, the DUT/SUT is expected to reach the desired value of the target objective in the sustain phase. Follow step 3, if the measured value does not meet the target value or does not fulfill the test results validation criteria.

7.7.4.3. Step 3: Test Iteration

Determine the achievable average inspected throughput within the test results validation criteria. Final test iteration MUST be performed for the test duration defined in Section 4.3.4.

7.8. HTTPS Transaction Latency

7.8.1. Objective

Using HTTPS traffic, determine the HTTPS transaction latency when DUT/SUT is running with sustainable HTTPS transactions per second supported by the DUT/SUT under different HTTPS response object size.

Scenario 1: The client MUST negotiate HTTPS and close the connection with FIN immediately after completion of a single transaction (GET and RESPONSE).

Scenario 2: The client MUST negotiate HTTPS and close the connection with FIN immediately after completion of 10 transactions (GET and RESPONSE) within a single TCP connection.

7.8.2. Test Setup

Testbed setup SHOULD be configured as defined in Section 4. Any specific testbed configuration changes (number of interfaces and interface type, etc.) MUST be documented.

7.8.3. Test Parameters

In this section, benchmarking test specific parameters SHOULD be defined.

7.8.3.1. DUT/SUT Configuration Parameters

DUT/SUT parameters MUST conform to the requirements defined in Section 4.2. Any configuration changes for this specific benchmarking test MUST be documented.

7.8.3.2. Test Equipment Configuration Parameters

Test equipment configuration parameters MUST conform to the requirements defined in Section 4.3. The following parameters MUST be documented for this benchmarking test:

Client IP address range defined in Section 4.3.1.2

Server IP address range defined in Section 4.3.2.2

Traffic distribution ratio between IPv4 and IPv6 defined in Section 4.3.1.2

RECOMMENDED cipher suites and key sizes defined in Section 4.3.1.3

Target objective for scenario 1: 50% of the connections per second measured in benchmarking test TCP/HTTPS Connections per second (Section 7.6)

Target objective for scenario 2: 50% of the inspected throughput measured in benchmarking test HTTPS Throughput (Section 7.7)

Initial objective for scenario 1: 10% of "Target objective for scenario 1"

Initial objective for scenario 2: 10% of "Target objective for scenario 2"

Note: The Initial objectives are not a KPI to report. These values are configured on the traffic generator and used to perform the Step1: "Test Initialization and Qualification" described under the Section 7.8.4.

HTTPS transaction per TCP connection: Test scenario 1 with single
transaction and scenario 2 with 10 transactions

HTTPS with GET request requesting a single object. The RECOMMENDED
object sizes are 1, 16, and 64 KByte. For each test iteration,
client MUST request a single HTTPS response object size.

7.8.3.3. Test Results Validation Criteria

The following criteria are the test results validation criteria. The
Test results validation criteria MUST be monitored during the whole
sustain phase of the traffic load profile.

   a. Number of failed application transactions (receiving any HTTP
      response code other than 200 OK) MUST be less than 0.001% (1
      out of 100,000 transactions) of attempt transactions.

   b. Number of terminated TCP connections due to unexpected TCP RST
      sent by DUT/SUT MUST be less than 0.001% (1 out of 100,000
      connections) of total initiated TCP connections.

   c. During the sustain phase, traffic SHOULD be forwarded at a
      constant rate (considered as a constant rate if any deviation
      of traffic forwarding rate is less than 5%).

   d. Concurrent TCP connections MUST be constant during steady state
      and any deviation of concurrent TCP connections SHOULD be less
      than 10%. This confirms the DUT opens and closes TCP
      connections at approximately the same rate.

   e. After ramp up the DUT/SUT MUST achieve the "Target objective"
      defined in the parameter Section 7.8.3.2 and remain in that
      state for the entire test duration (sustain phase).

7.8.3.4. Measurement

TTFB (minimum, average, and maximum) and TTLB (minimum, average and
maximum) MUST be reported for each object size.

7.8.4. Test Procedures and Expected Results

The test procedure is designed to measure TTFB or TTLB when the DUT/
SUT is operating close to 50% of its maximum achievable connections
per second or inspected throughput. The test procedure consists of
two major steps: Step 1 ensures the DUT/SUT is able to reach the
initial performance values and meets the test results validation
criteria when it was very minimally utilized. Step 2 measures the
latency values within the test results validation criteria.

This test procedure MAY be repeated multiple times with different IP types (IPv4 only, IPv6 only and IPv4 and IPv6 mixed traffic distribution), HTTPS response object sizes and single, and multiple transactions per connection scenarios.

7.8.4.1. Step 1: Test Initialization and Qualification

Verify the link status of all connected physical interfaces. All interfaces are expected to be in "UP" status.

Configure traffic load profile of the test equipment to establish "Initial objective" as defined in the [Section 7.8.3.2](#). The traffic load profile SHOULD be defined as described in [Section 4.3.4](#).

The DUT/SUT SHOULD reach the "Initial objective" before the sustain phase. The measured KPIs during the sustain phase MUST meet all the test results validation criteria defined in [Section 7.8.3.3](#).

If the KPI metrics do not meet the test results validation criteria, the test procedure MUST NOT be continued to "Step 2".

7.8.4.2. Step 2: Test Run with Target Objective

Configure test equipment to establish "Target objective" defined in [Section 7.8.3.2](#). The test equipment SHOULD follow the traffic load profile definition as described in [Section 4.3.4](#).

The test equipment SHOULD start to measure and record all specified KPIs. Continue the test until all traffic profile phases are completed.

Within the test results validation criteria, the DUT/SUT MUST reach the desired value of the target objective in the sustain phase.

Measure the minimum, average, and maximum values of TTFB and TTLB.

7.9. Concurrent TCP/HTTPS Connection Capacity

7.9.1. Objective

Determine the number of concurrent TCP connections the DUT/SUT sustains when using HTTPS traffic.

7.9.2. Test Setup

Testbed setup SHOULD be configured as defined in [Section 4](#). Any specific testbed configuration changes (number of interfaces and interface type, etc.) MUST be documented.

7.9.3. Test Parameters

In this section, benchmarking test specific parameters SHOULD be defined.

7.9.3.1. DUT/SUT Configuration Parameters

DUT/SUT parameters MUST conform to the requirements defined in Section 4.2. Any configuration changes for this specific benchmarking test MUST be documented.

7.9.3.2. Test Equipment Configuration Parameters

Test equipment configuration parameters MUST conform to the requirements defined in Section 4.3. The following parameters MUST be documented for this benchmarking test:

Client IP address range defined in Section 4.3.1.2

Server IP address range defined in Section 4.3.2.2

Traffic distribution ratio between IPv4 and IPv6 defined in Section 4.3.1.2

RECOMMENDED cipher suites and key sizes defined in Section 4.3.1.3

Target concurrent connections: Initial value from product datasheet or the value defined based on requirement for a specific deployment scenario.

Initial concurrent connections: 10% of "Target concurrent connections" Note: Initial concurrent connection is not a KPI to report. This value is configured on the traffic generator and used to perform the Step1: "Test Initialization and Qualification" described under the Section 7.9.4.

Connections per second during ramp up phase: 50% of maximum connections per second measured in benchmarking test TCP/HTTPS Connections per second (Section 7.6)

Ramp up time (in traffic load profile for "Target concurrent connections"): "Target concurrent connections" / "Maximum connections per second during ramp up phase"

Ramp up time (in traffic load profile for "Initial concurrent connections"): "Initial concurrent connections" / "Maximum connections per second during ramp up phase"

The client MUST perform HTTPS transaction with persistence and each
client can open multiple concurrent TCP connections per server
endpoint IP.

Each client sends 10 GET requests requesting 1 KByte HTTPS response
objects in the same TCP connections (10 transactions/TCP connection)
and the delay (think time) between each transaction MUST be X
seconds.

X = ("Ramp up time" + "steady state time") /10

The established connections SHOULD remain open until the ramp down
phase of the test. During the ramp down phase, all connections
SHOULD be successfully closed with FIN.

7.9.3.3. Test Results Validation Criteria

The following criteria are the test results validation criteria. The
Test results validation criteria MUST be monitored during the whole
sustain phase of the traffic load profile.

   a. Number of failed application transactions (receiving any HTTP
      response code other than 200 OK) MUST be less than 0.001% (1
      out of 100,000 transactions) of total attempted transactions.

   b. Number of terminated TCP connections due to unexpected TCP RST
      sent by DUT/SUT MUST be less than 0.001% (1 out of 100,000
      connections) of total initiated TCP connections.

   c. During the sustain phase, traffic SHOULD be forwarded at a
      constant rate (considered as a constant rate if any deviation
      of traffic forwarding rate is less than 5%).

7.9.3.4. Measurement

Average Concurrent TCP Connections MUST be reported for this
benchmarking test.

7.9.4. Test Procedures and Expected Results

The test procedure is designed to measure the concurrent TCP
connection capacity of the DUT/SUT at the sustaining period of
traffic load profile. The test procedure consists of three major
steps: Step 1 ensures the DUT/SUT is able to reach the performance
value (Initial concurrent connection) and meets the test results
validation criteria when it was very minimally utilized. Step 2
determines the DUT/SUT is able to reach the target performance value
within the test results validation criteria. Step 3 determines the
maximum achievable performance value within the test results
validation criteria.

This test procedure MAY be repeated multiple times with different IPv4 and IPv6 traffic distribution.

7.9.4.1. Step 1: Test Initialization and Qualification

Verify the link status of all connected physical interfaces. All interfaces are expected to be in "UP" status.

Configure test equipment to establish "Initial concurrent TCP connections" defined in Section 7.9.3.2. Except ramp up time, the traffic load profile SHOULD be defined as described in Section 4.3.4.

During the sustain phase, the DUT/SUT SHOULD reach the "Initial concurrent TCP connections". The measured KPIs during the sustain phase MUST meet the test results validation criteria "a" and "b" defined in Section 7.9.3.3.

If the KPI metrics do not meet the test results validation criteria, the test procedure MUST NOT be continued to "Step 2".

7.9.4.2. Step 2: Test Run with Target Objective

Configure test equipment to establish the target objective ("Target concurrent TCP connections"). The test equipment SHOULD follow the traffic load profile definition (except ramp up time) as described in Section 4.3.4.

During the ramp up and sustain phase, the other KPIs such as inspected throughput, TCP connections per second, and application transactions per second MUST NOT reach to the maximum value that the DUT/SUT can support.

The test equipment SHOULD start to measure and record KPIs defined in Section 7.9.3.4. Continue the test until all traffic profile phases are completed.

Within the test results validation criteria, the DUT/SUT is expected to reach the desired value of the target objective in the sustain phase. Follow step 3, if the measured value does not meet the target value or does not fulfill the test results validation criteria.

7.9.4.3. Step 3: Test Iteration

Determine the achievable concurrent TCP connections within the test results validation criteria.

8. IANA Considerations

The IANA has assigned IPv4 and IPv6 address blocks in [RFC6890] that have been registered for special purposes. The IPv6 address block 2001:2::/48 has been allocated for the purpose of IPv6 Benchmarking [RFC5180] and the IPv4 address block 198.18.0.0/15 has been allocated for the purpose of IPv4 Benchmarking [RFC2544]. This assignment was made to minimize the chance of conflict in case a testing device were to be accidentally connected to part of the Internet.

9. Security Considerations

The primary goal of this document is to provide benchmarking terminology and methodology for next-generation network security devices. However, readers should be aware that there is some overlap between performance and security issues. Specifically, the optimal configuration for network security device performance may not be the most secure, and vice-versa. The cipher suites recommended in this document are for test purpose only. The cipher suite recommendation for a real deployment is outside the scope of this document.

10. Contributors

The following individuals contributed significantly to the creation of this document:

Alex Samonte, Amritam Putatunda, Aria Eslambolchizadeh, Chao Guo, Chris Brown, Cory Ford, David DeSanto, Jurrie Van Den Breekel, Michelle Rhines, Mike Jack, Ryan Liles, Samaresh Nair, Stephen Goudreault, Tim Carlin, and Tim Otto.

11. Acknowledgements

The authors wish to acknowledge the members of NetSecOPEN for their participation in the creation of this document. Additionally, the following members need to be acknowledged:

Anand Vijayan, Chris Marshall, Jay Lindenauer, Michael Shannon, Mike Deichman, Ryan Riese, and Toulnay Orkun.

12. References

12.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/ RFC2119, March 1997, <https://www.rfc-editor.org/info/ rfc2119>.

[RFC8174]   Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC
                 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174,
                 May 2017, <https://www.rfc-editor.org/info/rfc8174>.

12.2. Informative References

[RFC2544]   Bradner, S. and J. McQuaid, "Benchmarking Methodology for
                 Network Interconnect Devices", RFC 2544, DOI 10.17487/
                 RFC2544, March 1999, <https://www.rfc-editor.org/info/
                 rfc2544>.

[RFC2616]   Fielding, R., Gettys, J., Mogul, J., Frystyk, H.,
                 Masinter, L., Leach, P., and T. Berners-Lee, "Hypertext
                 Transfer Protocol -- HTTP/1.1", RFC 2616, DOI 10.17487/
                 RFC2616, June 1999, <https://www.rfc-editor.org/info/
                 rfc2616>.

[RFC2647]   Newman, D., "Benchmarking Terminology for Firewall
                 Performance", RFC 2647, DOI 10.17487/RFC2647, August
                 1999, <https://www.rfc-editor.org/info/rfc2647>.

[RFC3511]   Hickman, B., Newman, D., Tadjudin, S., and T. Martin,
                 "Benchmarking Methodology for Firewall Performance", RFC
                 3511, DOI 10.17487/RFC3511, April 2003, <https://www.rfc-
                 editor.org/info/rfc3511>.

[RFC5180]   Popoviciu, C., Hamza, A., Van de Velde, G., and D.
                 Dugatkin, "IPv6 Benchmarking Methodology for Network
                 Interconnect Devices", RFC 5180, DOI 10.17487/RFC5180,
                 May 2008, <https://www.rfc-editor.org/info/rfc5180>.

[RFC6815]   Bradner, S., Dubray, K., McQuaid, J., and A. Morton,
                 "Applicability Statement for RFC 2544: Use on Production
                 Networks Considered Harmful", RFC 6815, DOI 10.17487/
                 RFC6815, November 2012, <https://www.rfc-editor.org/info/
                 rfc6815>.

[RFC6890]   Cotton, M., Vegoda, L., Bonica, R., Ed., and B. Haberman,
                 "Special-Purpose IP Address Registries", BCP 153, RFC
                 6890, DOI 10.17487/RFC6890, April 2013, <https://www.rfc-
                 editor.org/info/rfc6890>.

[RFC8446]   Rescorla, E., "The Transport Layer Security (TLS)
                 Protocol Version 1.3", RFC 8446, DOI 10.17487/RFC8446,
                 August 2018, <https://www.rfc-editor.org/info/rfc8446>.

[RFC9000]   Iyengar, J., Ed. and M. Thomson, Ed., "QUIC: A UDP-Based
                 Multiplexed and Secure Transport", RFC 9000, DOI

10.17487/RFC9000, May 2021, <https://www.rfc-editor.org/info/rfc9000>.

Appendix A. Test Methodology - Security Effectiveness Evaluation

A.1. Test Objective

   This test methodology verifies the DUT/SUT is able to detect,
   prevent, and report the vulnerabilities.

   In this test, background test traffic will be generated to utilize
   the DUT/SUT. In parallel, the CVEs will be sent to the DUT/SUT as
   encrypted and as well as clear text payload formats using a traffic
   generator. The selection of the CVEs is described in Section 4.2.1.

      *Number of blocked CVEs

      *Number of bypassed (nonblocked) CVEs

      *Background traffic performance (verify if the background traffic
       is impacted while sending CVE toward DUT/SUT)

      *Accuracy of DUT/SUT statistics in term of vulnerabilities
       reporting

A.2. Testbed Setup

   The same testbed MUST be used for security effectiveness test and as
   well as for benchmarking test cases defined in Section 7.

A.3. Test Parameters

   In this section, the benchmarking test specific parameters SHOULD be
   defined.

A.3.1. DUT/SUT Configuration Parameters

   DUT/SUT configuration parameters MUST conform to the requirements
   defined in Section 4.2. The same DUT configuration MUST be used for
   Security effectiveness test and as well as for benchmarking test
   cases defined in Section 7. The DUT/SUT MUST be configured in inline
   mode and all detected attack traffic MUST be dropped and the session
   SHOULD be reset

A.3.2. Test Equipment Configuration Parameters

   Test equipment configuration parameters MUST conform to the
   requirements defined in Section 4.3. The same client and server IP
   ranges MUST be configured as used in the benchmarking test cases. In

addition, the following parameters MUST be documented for this
benchmarking test:

  *Background Traffic: 45% of maximum HTTP throughput and 45% of
   Maximum HTTPS throughput supported by the DUT/SUT (measured with
   object size 64 KByte in the benchmarking tests "HTTP(S)
   Throughput" defined in Section 7.3 and Section 7.7).

  *RECOMMENDED CVE traffic transmission Rate: 10 CVEs per second

  *It is RECOMMENDED to generate each CVE multiple times
   (sequentially) at 10 CVEs per second

  *Ciphers and keys for the encrypted CVE traffic MUST use the same
   cipher configured for HTTPS traffic related benchmarking tests
   (Section 7.6 - Section 7.9)

A.4. Test Results Validation Criteria

   The following criteria are the test results validation criteria. The
   test results validation criteria MUST be monitored during the whole
   test duration.

   a. Number of failed application transaction in the background
      traffic MUST be less than 0.01% of attempted transactions.

   b. Number of terminated TCP connections of the background traffic
      (due to unexpected TCP RST sent by DUT/SUT) MUST be less than
      0.01% of total initiated TCP connections in the background
      traffic.

   c. During the sustain phase, traffic SHOULD be forwarded at a
      constant rate (considered as a constant rate if any deviation
      of traffic forwarding rate is less than 5%).

   d. False positive MUST NOT occur in the background traffic.

A.5. Measurement

   Following KPI metrics MUST be reported for this test scenario:

   Mandatory KPIs:

  *Blocked CVEs: It SHOULD be represented in the following ways:

     -Number of blocked CVEs out of total CVEs

     -Percentage of blocked CVEs

*Unblocked CVEs: It SHOULD be represented in the following ways:

   -Number of unblocked CVEs out of total CVEs

   -Percentage of unblocked CVEs

*Background traffic behavior: It SHOULD be represented one of the
 followings ways:

   -No impact: Considered as "no impact'" if any deviation of
    traffic forwarding rate is less than or equal to 5 % (constant
    rate)

   -Minor impact: Considered as "minor impact" if any deviation of
    traffic forwarding rate is greater than 5% and less than or
    equal to10% (i.e. small spikes)

   -Heavily impacted: Considered as "Heavily impacted" if any
    deviation of traffic forwarding rate is greater than 10% (i.e.
    large spikes) or reduced the background HTTP(S) throughput
    greater than 10%

  *DUT/SUT reporting accuracy: DUT/SUT MUST report all detected
   vulnerabilities.

Optional KPIs:

  *List of unblocked CVEs

A.6. Test Procedures and Expected Results

The test procedure is designed to measure the security effectiveness
of the DUT/SUT at the sustaining period of the traffic load profile.
The test procedure consists of two major steps. This test procedure
MAY be repeated multiple times with different IPv4 and IPv6 traffic
distribution.

A.6.1. Step 1: Background Traffic

Generate background traffic at the transmission rate defined in
Appendix A.3.2.

The DUT/SUT MUST reach the target objective (HTTP(S) throughput) in
sustain phase. The measured KPIs during the sustain phase MUST meet
all the test results validation criteria defined in Appendix A.4.

If the KPI metrics do not meet the acceptance criteria, the test
procedure MUST NOT be continued to "Step 2".

A.6.2. Step 2: CVE Emulation

While generating background traffic (in sustain phase), send the CVE traffic as defined in the parameter section.

The test equipment SHOULD start to measure and record all specified KPIs. Continue the test until all CVEs are sent.

The measured KPIs MUST meet all the test results validation criteria defined in [Appendix A.4](#).

In addition, the DUT/SUT SHOULD report the vulnerabilities correctly.

Appendix B. DUT/SUT Classification

This document aims to classify the DUT/SUT in four different categories based on its maximum supported firewall throughput performance number defined in the vendor datasheet. This classification MAY help user to determine specific configuration scale (e.g., number of ACL entries), traffic profiles, and attack traffic profiles, scaling those proportionally to DUT/SUT sizing category.

The four different categories are Extra Small (XS), Small (S), Medium (M), and Large (L). The RECOMMENDED throughput values for the following categories are:

Extra Small (XS) - Supported throughput less than or equal to 1Gbit/s

Small (S) - Supported throughput greater than 1Gbit/s and less than or equal to 5Gbit/s

Medium (M) - Supported throughput greater than 5Gbit/s and less than or equal to 10Gbit/s

Large (L) - Supported throughput greater than 10Gbit/s

Authors' Addresses

Balamuhunthan Balarajah
Berlin
Germany

Email: bm.balarajah@gmail.com

Carsten Rossenhoevel
EANTC AG
Salzufer 14
10587 Berlin

Germany

Email: cross@eantc.de

Brian Monkman
NetSecOPEN
417 Independence Court
Mechanicsburg, PA 17050
United States of America

Email: bmonkman@netsecopen.org