Methodology for Benchmarking SIP Networking Devices

ABSTRACT
This document describes the methodology for benchmarking Session Initiation Protocol (SIP) performance as described in Terminology document [Po06]. The methodology and terminology are to be used for benchmarking SIP control plane performance with varying control and media load. Both scale and establishment rate are measured by control plane performance. The SIP Devices to be benchmarked may be a single device under test (DUT) or a system under test (SUT). Benchmarks can be obtained and compared for different types of devices such as SIP Proxy Server, SBC, P-CSCF, and Server paired with a Firewall/NAT device.
# Table of Contents

1. Introduction ............................................. 2
2. Existing definitions .................................... 2
3. Test Setup................................................ 3
   3.1 Test Topologies........................................ 3
   3.2 Test Considerations.................................... 5
   3.3 Reporting Format....................................... 7
4. Test Cases................................................ 7
   4.1 Registration Rate.................................... 7
   4.2 Session Setup Rate.................................. 8
   4.3 Session Setup Rate with Loop Detection Enabled.... 9
   4.4 Session Setup Rate with Forking..................... 9
   4.5 Session Setup Rate with Forking and Loop Detection.9
   4.6 Session Setup Rate with Media....................... 9
   4.7 Session Capacity........................................ 10
   4.8 Session Capacity with Media........................ 11
   4.9 Session Setup Rate with TLS Encrypted SIP........... 12
   4.10 Session Setup Rate with IPsec Encrypted SIP....... 12
   4.11 Session Setup Rate with SIP Flooding............... 12
   4.12 IM Rate................................................ 12
   4.13 Presence Rate......................................... 12
5. IANA Considerations........................................ 13
6. Security Considerations................................. 13
7. Acknowledgements........................................ 13
8. References................................................ 14
9. Author’s Address......................................... 15
10. Full Copyright Statement................................. 16
1. Introduction

This document describes the methodology for benchmarking Session Initiation Protocol (SIP) performance as described in Terminology document [Po06]. The methodology and terminology are to be used for benchmarking SIP control plane performance with varying control and media load. Both scale and establishment rate are measured by control plane performance.

The SIP Devices to be benchmarked may be a single device under test (DUT) or a system under test (SUT). The DUT is a SIP Server, which may be any RFC 3261 [Ro02] conforming device. The SUT can be any device or group of devices containing RFC 3261 conforming functionality along with Firewall and/or NAT functionality. This enables benchmarks to be obtained and compared for different types of devices such as SIP Proxy Server, SBC, P-CSCF, Proxy Server paired with a Firewall/NAT device, and P-CSCF paired with a Firewall/NAT device. SIP Associated Media benchmarks can also be made when testing SUTs.

The test cases covered in this methodology document provide benchmarks metrics of Registration Rate, SIP Session Setup Rate, Session Capacity, IM Rate, and Presence Rate. These can be benchmarked with or without associated Media. Some cases are also included to cover Forking, Loop detection, Encrypted SIP, and SIP Flooding. The test topologies that can be used are described in the Test Setup section. Topologies are provided for benchmarking of a DUT or SUT. Benchmarking with Associated Media can be performed when using a SUT.

SIP permits a wide range of configuration options that are also explained in the Test Setup section. Benchmark metrics could possibly be impacted by Associated Media. The selected values for Session Duration and Media Streams per Session enable benchmark metrics to be benchmarked without Associated Media. Session Setup Rate could possibly be impacted by the selected value for Maximum Sessions Attempted. The benchmark for Session Setup Rate is measured with a fixed value for Maximum Sessions Attempted.

2. Existing definitions

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14, RFC 2119. RFC 2119 defines the use of these key words to help make the intent of standards track documents as clear as possible. While this document uses these keywords, this document is not a standards track document. The term Throughput is defined in RFC 2544.

Terms specific to SIP Performance benchmarking are defined in [Po06].
3. Test Setup

3.1 Test Topologies

Figures 1 through 5 below provide various topologies to perform the SIP Performance Benchmarking. These figures show the Device Under Test (DUT) to be a single server or a System Under Test (SUT). Test Topology options to include benchmarking with Associated Media require use of a SUT and are shown in Figures 4 and 5.

![Figure 1. Basic SIP Test Topology](image1.png)

DUT
---------
|       |       |
|       |       |
| Server|<-------|Tester|
|       |       |
|-------|       |

SIP

---------

Figure 1. Basic SIP Test Topology

![Figure 2. SIP Test Topology with Firewall](image2.png)

SUT
-------------------------------
---------
|       |   |       |
|       |   |       |
| Server|Fire-|SIP   |
|       |    |Wall  |
|       |    |      |
|-------|   |      |

Tester

---------

Figure 2. SIP Test Topology with Firewall

![Figure 3. SIP Test Topology with NAT Device](image3.png)

SUT
-------------------------------
---------
|       |   |       |
|       |   |       |
| Server|NAT |SIP   |
|       |    |      |
|-------|   |      |

Tester

---------

Figure 3. SIP Test Topology with NAT Device
3.2 Test Considerations

3.2.1 Selection of SIP Transport Protocol
Test cases may be performed with any transport protocol supported by SIP. This includes, but is not limited to, SIP TCP, SIP UDP, and TLS. The protocol used for the SIP transport protocol must be reported with benchmarking results.

3.2.2 Server
The Server is a SIP-speaking device that complies with RFC 3261. The purpose of this document is to benchmark SIP performance, not conformance. Conformance to RFC 3261 [Ro02] is assumed for all tests. The Server may be the DU or a component of a SUT that includes Firewall and/or NAT functionality. The components of the SUT may be a single physical device or separate devices.
3.2.3 Associated Media
Some tests may require associated media to be present for each SIP session. The Server is not involved in the forwarding of media. Associated Media can be benchmarked only with a SUT in which the media traverses a Firewall, NAT, or Firewall NAT device. The test topologies to be used when benchmarking SUT performance for Associated Media are shown in Figures 4 and 5.

3.2.4 Selection of Associated Media Protocol
The test cases specified in this document provide SIP performance independent of the protocol used for the media stream. Any media protocol supported by SIP may be used. This includes, but is not limited to, RTP, RTSP, and SRTP. The protocol used for Associated Media must be reported with benchmarking results.

3.2.5 Number of Associated Media Streams per SIP Session
Benchmarking results can possibly vary with the number of media streams per SIP session. When benchmarking a SUT for voice, a single media stream is used. When benchmarking a SUT for voice and video, two media streams are used. The number of Associated Media Streams must be reported with benchmarking results.

3.2.6 Session Duration
SUT performance benchmarks can possibly vary with the duration of SIP sessions. Session Duration must be reported with benchmarking results. A Session Duration of zero seconds indicates transmission of a BYE immediately following successful SIP establishment indicate by receipt of a 200 OK. An infinite Session Duration indicates that a BYE is never transmitted.

3.2.7 Attempted Sessions per Second
DUT and SUT performance benchmarks can possibly vary with the rate of attempted sessions offered by the Tester. Attempted Sessions per Second must be reported with benchmarking results.

3.2.8 Stress Testing
The purpose of this document is to benchmark SIP performance, not system stability under stressful conditions such as a high rate of Attempted Sessions per Second.
3.3 Reporting Format

Test Setup
SIP Transport Protocol = ___________
Session Duration = ___________
Attempted Session Rate = ___________
Maximum Sessions Attempted = ___________
Media Streams per Session = ___________
Media Protocol = ___________

Device Benchmarks
Failed Session Attempts = ___________
Session Capacity = ___________
Zero-Failure Session Setup Rate = ___________
Maximum Retransmits = ___________
Mean Session Setup Delay = ___________
Mean Session TearDown Delay = ___________

4. Test Cases

4.1 Registration Rate
4.2 Session Setup Rate

Objective:
To benchmark the maximum session setup rate performance
of the DUT/SUT with zero failures.

Procedure:
1. Configure the DUT in the test topology shown in Figure 1 or
   SUT as shown in Figures 2 or 3.

2. Configure Tester for SIP UDP with an Attempted Session Rate =
   100 SPS, Session Duration = 0 sec, Maximum Sessions Attempted
   = 100,000 and media streams per session=0.

3. Start Tester to initiate SIP Session establishment with the
   DUT.

4. Measure Failed Session Attempts and Total Sessions
   Established at the Tester.

5. If a Failed Session Attempt is recorded then reduce the
   Attempted Session Rate configured on the Tester by 50%.

6. If no Failed Session Attempt is recorded then increase the
   Attempted Session Rate configured on the Tester by 50%.

7. Repeat steps 3 through 6 until the Session Setup
   Rate is obtained.

Expected Results:
4.3 Session Setup Rate with Loop Detection Enabled

4.4 Session Setup Rate with Forking

4.5 Session Setup Rate with Forking and Loop Detection

4.6 Session Setup Rate with Media

Objective:
To benchmark the maximum session setup rate performance of the SUT with zero failures when Associated Media is included in the benchmark test.

Procedure:
1. Configure the SUT in the test topology shown in Figure 4 or 5.

2. Configure Tester for SIP UDP with an Attempted Session Rate = 100 SPS, Session Duration = 30 sec, Maximum Sessions Attempted = 100,000 and media streams per session = 1. The rate of offered load for each media stream SHOULD be

\[
\text{Offered Load per Media Stream} = \frac{\text{Throughput}}{\text{Maximum Sessions Attempted}},
\]

where Throughput is defined in [Ba99].

3. Start Tester to initiate SIP Session establishment with the SUT and transmit media through the SUT to a destination other than the server.

4. At the Tester measure Failed Session Attempts, Total Sessions Established, and Packet Loss [Ba99] of the media.

5. If a Failed Session Attempt or Packet Loss is recorded then reduce the Attempted Session Rate configured on the Tester by 50%.

6. If no Failed Session Attempt or Packet Loss is recorded then increase the Attempted Session Rate configured on the Tester by 50%.

7. Repeat steps 3 through 6 until the Session Setup Rate is obtained.

8. Repeat steps 1 through 7 for multimedia in which media streams per session = 2.

Expected Results:
Session Setup Rate results obtained with Associated Media with any number of media streams per SIP session will be identical to the Session Setup Rate results obtained without media.
4.7 Session Capacity

Objective:
To benchmark the SIP Control Session Capacity of the DUT/SUT.

Procedure:
1. Configure the DUT in the test topology shown in Figure 1 or SUT as shown in Figures 2 or 3.

2. Configure Tester for SIP UDP with an Attempted Session Rate = 100 SPS, Session Duration = 30 sec, Maximum Sessions Attempted = 100,000 and media streams per session = 1. The rate of offered load for each media stream SHOULD be

\[
\text{(eq 1) Offered Load per Media Stream} = \frac{\text{Throughput}}{\text{Maximum Sessions Attempted}},
\]

where Throughput is defined in [Ba99].

3. Start Tester to initiate SIP Session establishment with the SUT and transmit media through the SUT to a destination other than the server.

4. Measure Failed Session Attempts, Total Sessions Established, and at the Tester.

5. If a Failed Session Attempt is recorded then reduce the Maximum Sessions Attempted configured on the Tester by 5,000.

6. If no Failed Session Attempt is recorded then increase the Maximum Sessions Attempted configured on the Tester by 10,000.

7. Repeat steps 3 through 6 until the Session Capacity is obtained.

Expected Results:
4.8 Session Capacity with Media

Objective:
To benchmark the SIP Control Session Capacity of the SUT with Associated Media.

Procedure:
1. Configure the DUT in the test topology shown in Figure 1 or SUT as shown in Figures 2 or 3.
2. Configure Tester for SIP UDP with an Attempted Session Rate = Zero-Failure Session Setup Rate, Session Duration = 0 sec, Maximum Sessions Attempted = 10,000 and media streams per session = 0.
3. Start Tester to initiate SIP Session establishment with the DUT.
5. If a Failed Session Attempt or Packet Loss is recorded then reduce the Maximum Sessions Attempted configured on the Tester by 5,000.
6. If no Failed Session Attempt or Packet Loss is recorded then increase the Maximum Sessions Attempted configured on the Tester by 10,000.
7. Repeat steps 3 through 6 until the Session Capacity is obtained.
8. Repeat steps 1 through 7 for multimedia in which media streams per session = 2.

Expected Results:
Session Capacity results obtained with Associated Media with any number of media streams per SIP session will be identical to the Session Capacity results obtained without media.
4.9 Session Setup Rate with TLS Encrypted SIP

4.10 Session Setup Rate with IPsec Encrypted SIP

4.11 Session Setup Rate with SIP Flooding

Objective:
To benchmark the maximum session setup rate performance of the SUT with zero failures when SIP Flooding is occurring.

Procedure:
1. Configure the DUT in the test topology shown in Figure 1 or the SUT as shown in Figure 2.

2. Configure Tester for SIP UDP with an Attempted Session Rate = 100 SPS, Session Duration = 0 sec, Maximum Sessions Attempted = 100,000, Associated Media Streams per session = 0, and SIP INVITE Message Flood = 500 per second.

3. Start Tester to initiate SIP Session establishment with the SUT and SIP Flood targeted at the Server.

4. At the Tester measure Failed Session Attempts, Total Sessions Established, and Packet Loss of the media.

5. If a Failed Session Attempt or Packet Loss is recorded then reduce the Attempted Session Rate configured on the Tester by 50%.

6. If no Failed Session Attempt or Packet Loss is recorded then increase the Attempted Session Rate configured on the Tester by 50%.

7. Repeat steps 3 through 6 until the Session Setup Rate is obtained.

8. Repeat steps 1 through 7 with SIP INVITE Message Flood = 1000 per second.

Expected Results:
Session Setup Rate results obtained with SIP Flooding may be degraded.

4.12 IM Rate

4.13 Presence Rate
5. IANA Considerations

This document requires no IANA considerations.

6. Security Considerations

Documents of this type do not directly affect the security of Internet or corporate networks as long as benchmarking is not performed on devices or systems connected to production networks. Security threats and how to counter these in SIP and the media layer is discussed in RFC3261, RFC3550, and RFC3711 and various other drafts. This document attempts to formalize a set of common methodology for benchmarking performance of SIP devices in a lab environment.

7. Acknowledgements
8. References

8.1 Normative References


8.2 Informative References

None

Poretsky, Gurbani, Davids
9. Author’s Address

Scott Poretsky
Reef Point Systems
8 New England Executive Park
Burlington, MA 01803
USA
Phone: + 1 508 439 9008
EMail: sporetsky@reefpoint.com

Vijay Gurbani
2000 Lucent Lane
Lucent Technologies
Room 6G-440
Naperville, IL 60566
USA
Phone: + 1 630 224 0216
Email: vkg@lucent.com

Carol Davids
Illinois Institute of Technology
Rice Campus
201 East Loop Road
Wheaton, IL 60187
USA
Phone: + 1 630 682 6000
Email: davids@iit.edu