This document describes the methodology for benchmarking Session Initiation Protocol (SIP) performance as described in Terminology document [5]. The methodology and terminology are to be used for benchmarking signaling plane performance with varying signaling and media load. Both scale and establishment rate are measured by
signaling 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. Terminology .................................................. 4
2. Introduction .................................................. 4
3. Test Setup .................................................... 5
   3.1. Test Topologies .......................................... 5
   3.2. Test Considerations ....................................... 8
       3.2.1. Selection of SIP Transport Protocol ............... 8
       3.2.2. Server ............................................... 8
       3.2.3. Associated Media .................................... 8
       3.2.4. Selection of Associated Media Protocol ............ 8
       3.2.5. Number of Associated Media Streams per SIP Session .. 8
       3.2.6. Session Duration ...................................... 9
       3.2.7. Attempted Sessions per Second ....................... 9
       3.2.8. Stress Testing ........................................ 9
   3.3. Reporting Format .......................................... 9
       3.3.1. Test setup Report .................................. 9
       3.3.2. Device Benchmarks .................................. 9
4. Test Cases .................................................... 10
   4.1. Maximum Session Attempt Rate ........................... 10
   4.2. Maximum Session Attempt Rate with Media ............... 10
   4.3. Maximum Session Attempt Rate with Loop Detection
        Enabled ................................................... 11
   4.4. Maximum Session Attempt Rate with Forking ................ 11
   4.5. Maximum Session Attempt Rate with Forking and Loop
        Detection ................................................. 12
   4.6. Maximum Session Attempt Rate with TLS Encrypted SIP .... 12
   4.7. Maximum Session Attempt Rate with IPsec Encrypted SIP ... 13
   4.8. Maximum Session Attempt Rate with SIP Flooding .......... 13
   4.9. Maximum Registration Rate ................................ 14
   4.10. Maximum IM Rate ......................................... 15
   4.11. Maximum Presence Rate ................................... 15
   4.12. Maximum Session Establishment Rate ...................... 15
   4.13. Maximum Session Establishment Rate with media .......... 16
5. IANA Considerations ............................................ 17
6. Security Considerations ....................................... 17
7. Acknowledgments ............................................... 17
8. References .................................................... 17
   8.1. Normative References .................................... 17
   8.2. Informational References ................................. 18
Authors’ Addresses ................................................. 18
1. Terminology

In this document, the key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" are to be interpreted as described in BCP 14, conforming to RFC 2119 [1] and indicate requirement levels for compliant implementations.

Terms specific to SIP Performance benchmarking are defined in [5].

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.

2. Introduction

This document describes the methodology for benchmarking Session Initiation Protocol (SIP) performance as described in Terminology document [5]. The methodology and terminology are to be used for benchmarking signaling plane performance with varying signaling and media load. Both scale and establishment rate are measured by signaling 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 [6] 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.

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](image-url)
Figure 2. SIP Test Topology with Firewall

Figure 3. SIP Test Topology with NAT Device
Figure 4. SIP Test Topology with Media through Firewall

Figure 5. SIP Test Topology with Media through NAT Device
3.2. Test Considerations

3.2.1. Selection of SIP Transport Protocol
Discussion:
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
Discussion:
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 is assumed for all tests. The Server may be the DUT 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
Discussion:
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
Discussion:
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
Discussion:
Benchmarking results may 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
Discussion:
SUT performance benchmarks may 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
Discussion:
DUT and SUT performance benchmarks may vary with the the rate of attempted sessions offered by the Tester. Attempted Sessions per Second must be reported with benchmarking results.

3.2.8. Stress Testing
Discussion:
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

3.3.1. Test setup Report
SIP Transport Protocol = _____________________________
IS Duration = _____________________________
Maximum Sessions Attempted = __________________________
Media Streams per Session = __________________________
Media Protocol = ___________________________

3.3.2. Device Benchmarks
Failed Session Attempts = __________________________
Session Capacity = __________________________
Maximum Session Establishment Rate = __________________________
Maximum Retransmits = _______________________________

Mean Session Setup Delay = __________________________

Mean Session Disconnect Delay = ______________________

4. Test Cases

4.1. Maximum Session Attempt Rate

Objective:
To benchmark the maximum session attempt rate 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 Maximum Session Establishment Rate is obtained.

Expected Results:

4.2. Maximum Session Attempt Rate with Media

Objective:
To benchmark the maximum session establishment rate 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 (eq 1) Offered Load per Media Stream = Throughput / Maximum Sessions Attempted, where Throughput is defined in [3].
3. Start Tester to initiate SIP Session establishment with the SUT and transmit media through the SUT to a destination other than the server.
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:
Maximum Session Establishment 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.

4.3. Maximum Session Attempt Rate with Loop Detection Enabled

Objective:
To benchmark the maximum session attempt rate of the DUT/SUT with zero failures when the Loop Detection option is enabled.

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. Turn on the Loop Detection option in the DUT or SUT.
4. Start Tester to initiate SIP Session establishment with the DUT.
5. Measure Failed Session Attempts and Total Sessions Established at the Tester.
6. If a Failed Session Attempt is recorded then reduce the Attempted Session Rate configured on the Tester by 50%.
7. If no Failed Session Attempt is recorded then increase the Attempted Session Rate configured on the Tester by 50%.
8. Repeat steps 4 through 7 until the Maximum Session Establishment Rate is obtained.

Expected Results:

4.4. Maximum Session Attempt Rate with Forking

Objective:
To benchmark the maximum session attempt rate of the DUT/SUT with zero failures when the Forking option is enabled.
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. Turn on the Forking option in the DUT or SUT.
4. Start Tester to initiate SIP Session establishment with the DUT.
5. Measure Failed Session Attempts and Total Sessions Established at the Tester.
6. If a Failed Session Attempt is recorded then reduce the Attempted Session Rate configured on the Tester by 50%.
7. If no Failed Session Attempt is recorded then increase the Attempted Session Rate configured on the Tester by 50%.
8. Repeat steps 4 through 7 until the Maximum Session Establishment Rate is obtained.

Expected Results:

4.5. Maximum Session Attempt Rate with Forking and Loop Detection

Objective:
To benchmark the maximum session attempt rate of the DUT/SUT with zero failures when both forking and loop detection are enabled.

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. Turn on both the forking and the loop detection options.
5. Measure Failed Session Attempts and Total Sessions Established at the Tester.
6. If a Failed Session Attempt is recorded then reduce the Attempted Session Rate configured on the Tester by 50%.
7. If no Failed Session Attempt is recorded then increase the Attempted Session Rate configured on the Tester by 50%.
8. Repeat steps 4 through 7 until the Maximum Session Establishment Rate is obtained.

Expected Results:

4.6. Maximum Session Attempt Rate with TLS Encrypted SIP

Objective:
To benchmark the maximum session attempt rate 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 TCP, enable TLS, 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 Maximum Session Establishment Rate is obtained.

Expected Results:

4.7 Maximum Session Attempt Rate with IPsec Encrypted SIP

Objective: To benchmark the maximum session attempt rate 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 TCP, enable IPsec, 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 Maximum Session Establishment Rate is obtained.

Expected Results:

4.8 Maximum Session Attempt Rate with SIP Flooding

Objective:
To benchmark the maximum session attempt rate 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.

4.9. Maximum Registration Rate

Objective:
To benchmark the maximum registration rate of the 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 Registration Rate = 100 SPS, Maximum Registrations Attempted = 100,000.
3. At the Tester measure Failed Registration Attempts, Total Registrations and Packet Loss.
4. If a Failed Registration Attempt or Packet Loss is recorded then reduce the Attempted Registration Rate configured on the Tester by 50%.
5. If no Failed Registration or Packet Loss is recorded then increase the Attempted Registration Rate configured on the Tester by 50%.
6. Repeat steps 3 through 6 until the Session Setup Rate is obtained.
4.10. Maximum IM Rate

Objective:
To benchmark the maximum IM rate of the 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 IM Rate = 100 SPS, Maximum IM Attempted = 100,000.
3. At the Tester measure Failed IM Attempts, Total IM and Packet Loss.
4. If a Failed IM Attempt or Packet Loss is recorded then reduce the Attempted IM Rate configured on the Tester by 50%.
5. If no Failed IM or Packet Loss is recorded then increase the Attempted IM Rate configured on the Tester by 50%.
6. Repeat steps 3 through 6 until the Session Setup Rate is obtained.

Expected Results:

4.11. Maximum Presence Rate

Objective:
To benchmark the Maximum Presence Rate of the 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 Presence Rate = 100 SPS, Maximum Registrations Attempted = 100,000.
3. At the Tester measure Failed Presence Attempts, Total Presence Attempts and Packet Loss.
4. If a Failed Presence Attempt or Packet Loss is recorded then reduce the Attempted Presence Rate configured on the Tester by 50%.
5. If no Failed Presence Attempt or Packet Loss is recorded then increase the Attempted Registration Rate configured on the Tester by 50%.
6. Repeat steps 3 through 6 until the Session Setup Rate is obtained.

Expected Results:

4.12. Maximum Session Establishment Rate

Objective:
To benchmark the 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:

4.13. Maximum Session Establishment Rate with media

Objective:
To benchmark the Maximum Session Establishment Rate of the DUT/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 a Session Attempt 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 (eq 1) Offered Load per Media Stream = Throughput / Maximum Sessions Attempted, where Throughput is defined in [3].
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 and Total Sessions Established 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: Session establishment rate 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.

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. Acknowledgments

The authors would like to thank Keith Drage and Daryl Malas for their contributions to this document.

8. References

8.1. Normative References


8.2. Informational References

Authors’ Addresses

Scott Porersky
Reef Point Networks
8 New England and Executive Park
Burlington, MA  08103
USA

Phone: +1 508 439 9008
Email: sporetsky@reefpoint.com

Vijay K. Gurbani
Bell Laboratories, Alcatel-Lucent
2701 Lucent Lane
Rm 9F-546
Lisle, IL  60532
USA

Phone: +1 630 224 0216
Email: vkg@alcatel-lucent.com
Carol Davids  
Illinois Institute of Technology  
201 East Loop Road  
Wheaton, IL 60187  
USA  

Email: davids@iit.edu
Full Copyright Statement

Copyright (C) The IETF Trust (2007).

This document is subject to the rights, licenses and restrictions contained in BCP 78, and except as set forth therein, the authors retain all their rights.

This document and the information contained herein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY, THE IETF TRUST AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Intellectual Property

The IETF takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights. Information on the procedures with respect to rights in RFC documents can be found in BCP 78 and BCP 79.

Copies of IPR disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at http://www.ietf.org/ipr.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at ietf-ipr@ietf.org.

Acknowledgment

Funding for the RFC Editor function is provided by the IETF Administrative Support Activity (IASA).