INTERNET-DRAFT
Intended Status: Informational
Expires: April 1, 2012

T. Alexander, K. Green
(Ixia)

October 2011

Benchmarking Methodology for Evaluating the Security Effectiveness of Content Aware Devices
draft-green-bmwg-seceff-bench-meth-00

Abstract

This document defines a methodology for evaluating the ability of content-aware network devices to correctly detect and block malicious or administratively disallowed traffic flows. This benchmark addresses the issue of classification accuracy under well defined conditions. It is not concerned with measuring forwarding performance which is covered by other BMWG documents.

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 http://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 April 1, 2012.

Copyright and License Notice

Table of Contents

1 Introduction

Networks of the 21st century exist in an environment flooded with complex and highly sophisticated security threats. There is an intense and enduring arms race under way between those developing and distributing attack technology and those developing and supplying defense technology.

In addition there is a growing need to limit access from users inside private or corporate networks to Internet sites or services deemed undesirable and to ensure that intellectual property and other private information is not allowed to pass freely from inside the protected network to the outside world.

In response to the this dynamic and constantly expanding range of security threats and privacy requirements there is a growing diversity of network devices that provide a variety of defensive services including but not limited to firewall, intrusion detection, intrusion prevention, anti-virus, anti-malware, anti-spam, anti-dos, anti-ddos, unified threat management, data leakage prevention and more. These content-aware devices use a mixture of stateless and stateful L3 to L7 technologies, including deep packet inspection (DPI) to categorize traffic flows.

What all of these defensive solutions have in common is the requirement that they reliably and accurately distinguish between evil (malicious or disallowed) traffic and good traffic.

Categorization of traffic as either good or evil is fundamental to the operation of these devices since it is a prerequisite to all security functions.

Security Effectiveness is a measure of how accurately the device under test (DUT) categorizes traffic:

- No false negatives = correctly blocks all evil traffic
- No false positives = never blocks good traffic

In contrast, Security Performance is the characterization of the DUT’s forwarding performance while under attack. Security Performance measures how well the device forwards good traffic with security features enabled and in the presence of evil traffic. This is addressed in [HAMILTON].

Security Effectiveness is orthogonal to Security Performance.

1.1 Requirements Language

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 [RFC2119].

2 Terminology

2.1 Existing Terminology

2.2 New Terminology

2.2.1 Good Traffic

Good traffic is any traffic flow which is benign and should not be blocked by the DUT.

2.2.2 Evil Traffic

Evil traffic is any illicit traffic flow which should be blocked by the DUT. Evil traffic is either malicious (i.e. part of a deliberate attack) or administratively banned (i.e. disallowed from passing in to or out of the protected network due to its content and/or destination).

3 Test Setup

3.1 Application Traffic Mix

Some test cases require the test equipment to inject good traffic mixed with the evil traffic. The purpose of the good traffic is to force the DUT to distinguish between good and evil traffic and it is not used to quantify the forwarding performance of the DUT from an application perspective.

Given this purpose, in order to protect the integrity and repeatability of the benchmark, a single fixed definition of the good traffic application mix is provided. No attempt is made to accurately model any particular mix of application traffic such as might be seen in an operational network.

Rather, the traffic mix includes an appropriate mix of traffic types to ensure that the security engine cannot blindly assume that every packet is either good or evil and so deliver unrealistically high performance or otherwise undermine the benchmark.

In those test scenarios where application traffic is specified, the

following mix MUST be used:

*** TBD but likely to include at least UDP, TCP, HTTP ***

4 Benchmarking Tests

4.1 Maximum Attack Blocking Rate

Maximum Attack Blocking Rate (attacks/second) is defined as the largest number of attacks per second where 100% of attacks are blocked with no application traffic present.

4.2 Useful Attack Blocking Rate

Useful Attack Blocking Rate (attacks/second) is defined as the largest number of attacks per second where 100% of attacks are blocked in the presence of good traffic and 0% of the good traffic is blocked or dropped.

4.3 Attack Blocking Effectiveness

Attack Blocking Effectiveness (percentage) is the ratio of blocked attacks/attempted attacks counted over the total number of different types of attack in the presence of good traffic and where 0% of the good traffic is dropped or blocked.

5 Security Considerations

Benchmarking activities as described in this memo are limited to technology characterization using controlled stimuli in a laboratory environment, with dedicated address space and the other constraints defined in [RFC2544].

The benchmarking network topology will be an independent test setup and MUST NOT be connected to devices that may forward the test traffic into a production network, or misroute traffic to the test management network.

Further, benchmarking is performed on a "black-box" basis, relying solely on measurements observable external to the DUT/SUT.

Special capabilities SHOULD NOT exist in the DUT/SUT specifically for benchmarking purposes. Any implications for network security arising from the DUT/SUT SHOULD be identical in the lab and in production networks.

6 IANA Considerations

This memo includes no request to IANA.

7 Acknowledgements

Thanks to X, Y & Z for their review and comments.

8 References

8.1 Normative References

[RFC1242] Bradner, S., "Benchmarking Terminology for Network Interconnection Devices", RFC 1242, July 1991.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

8.2 Informative References

[RFC2544] Bradner, S. and J. McQuaid, "Benchmarking Methodology for Network Interconnect Devices", RFC 2544, March 1999.

[RFC3511] Hickman, B., Newman, D., Tadjudin, S., and T. Martin, "Benchmarking Methodology for Firewall Performance", RFC 3511, April 2003.

[HAMILTON] "Benchmarking Methodology for Content Aware Network Devices", draft-hamilton-bmwg-ca-bench-07.txt

Authors' Addresses

Kenneth Green
Ixia
Australia

EMail: kgreen@ixiacom.com

Tom Alexander
Ixia
USA

EMail: talexander@ixiacom.com