

**MALAYSIAN PUBLIC SECTOR  
OPEN SOURCE SOFTWARE (OSS) PROGRAMME**

---

**BENCHMARK / EVALUATION REPORT ON  
CONTENT FILTERING**

---

**NOVEMBER 2009**



## Copyright

The Government of Malaysia retains the copyright of this document.

---

## Table of Contents

SECTION 1 : INTRODUCTION.....	1
Executive Summary.....	2
Introduction.....	2
SECTION 2 : DESCRIPTION OF EVALUATION METHODS (METHODOLOGY).....	3
Purpose of the Evaluation.....	4
Evaluation Design.....	4
Data Collection .....	4
SECTION 3 : DISCUSSION OF RESULTS.....	6
Results.....	7
Conclusion.....	10
SECTION 4 : REFERENCES.....	11
References.....	12

## **SECTION 1 : INTRODUCTION**

## Executive Summary

When we access the internet, we can be confronted by websites that may be deemed harmful, objectionable or just plain offensive. As much as possible, we need to filter elements that may cause distress, fear and confusion, e.g. racial, religious, ethnic prejudice; pornography; civil disobedience; etc. Among the people most likely to be affected are children and teenagers. In corporate sectors, it is mostly now normal policy to filter out offensive sites which will affect workers in a negative way.

This report will be of interest to all sectors of the general public who have a need and the authority to restrict access to the Internet. This report evaluates the effectiveness and efficiency of an OSCC product, MySurfGuard, against other tools available in the market to filter potentially harmful Internet content.

## Introduction

This report tested an OSCC product, MySurfGuard. This product is built on DansGuardian, which is an open source web content filter which works on actual content of pages. It runs only on Linux-based systems. Results from MySurfGuard evaluations are then benchmarked against other known content filtering tools, i.e. Sonicwall and Barracuda. The evaluations and benchmarks only compare their function of filtering harmful website content. It does not include spam, viruses or malware which may be present in the websites. This benchmark aims to improve awareness of solutions, to encourage best practices and to provide guidance to Public Sector agencies and the general public.

## **SECTION 2 : DESCRIPTION OF EVALUATION METHODS (METHODOLOGY)**

## Purpose of the Evaluation

This evaluation is to assess the following characteristics of MySurfGuard filtering tool.

- Filtering effectiveness
- Filtering efficiency

## Evaluation Design

An automated script was used to test the filtering tools. A Linux (Ubuntu) machine was used for the automated scripted tests. First, a Ruby-based testing script was downloaded from the “Untangle – Deep Throat Fight Club” website, It is a package designed specifically to test web filtering tools like Barracuda and Sonicwall. The same script was modified to test MySurfGuard for later comparisons.

Ruby, an open source programming language was installed on a Linux machine. After that the testing script was installed. In the network, the machine was located “outside the firewall”, i.e. unprotected. This set-up was necessary in order to get accurate results when testing MySurfGuard's capabilities. The testing script package comprised 5,540 websites.

## Data Collection

The Ruby-based script, described in the previous section, ran a test and checked what MySurfGuard filtered. It then generated a file listing the websites as blocked, not blocked or disqualified. After that, it checked the generated file with a manually prepared list and generated results and charts. A table below shows how the testing scripts determined the comparison.

Manually prepared list	Generated List	Score Result
Y	Y	Correct
Y	N	Underblocking (missed)
N	Y	Overblocking (False positive)
N	N	Correct
D	Y	Disqualified
D	N	Disqualified

Table 1. Scoring methodology for testing script

Unsafe site = Sites that contained material deemed to be objectionable

Safe site = Sites that are considered safe for general public viewing

Y = Correctly blocked unsafe site

N = Correct safe site that is not blocked

D = Disqualified, as site is inaccessible or cannot be determined

After this comparison was done, the script calculated the total number of correct, missed and false positive numbers. Based on these numbers, percentages were calculated and charts produced. The charts showed how MySurfGuard's effectiveness and efficiency compared with other commercial products, i.e. Sonicwall and Barracuda. The two metrics used here are underblocking and overblocking as in the table below.

Metric	Definition
Underblocking	"How much unwanted content will this web filtering tool miss ?"
Overblocking	"How much good content will it unnecessarily filter ?"

Table 2. Blocking metrics definitions

## **SECTION 3 : DISCUSSION OF RESULTS**

## Results

Results of the automated test for MySurfGuard as compared with results taken from the Untangle website for Sonicwall and Barracuda are in the tables below.

	<b>MySurfGuard</b>	<b>Sonicwall</b>	<b>Barracuda</b>
Correct	93.30%	95.20%	93.80%
Missed	3.60%	3.10%	4.80%
False positive	3.10%	1.70%	1.30%

Table 3. Effectiveness results comparison

	<b>MySurfGuard</b>	<b>Sonicwall</b>	<b>Barracuda</b>
Porn Caught	95.60%	96.10%	94.00%
Porn missed	4.40%	3.90%	6.00%

Table 4. Efficiency results comparison

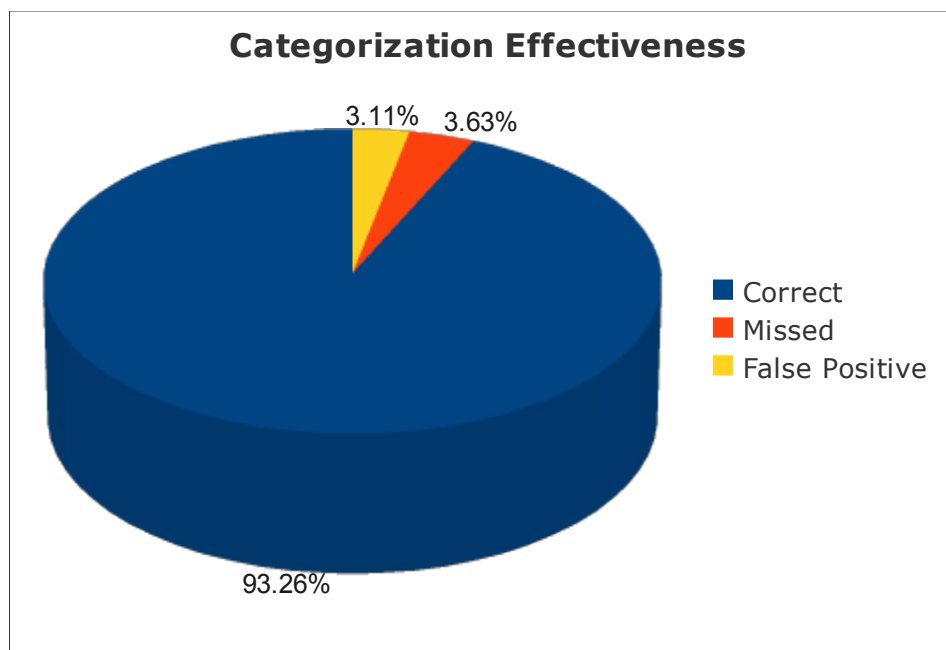


Chart 1. Categorization Effectiveness for MySurfGuard

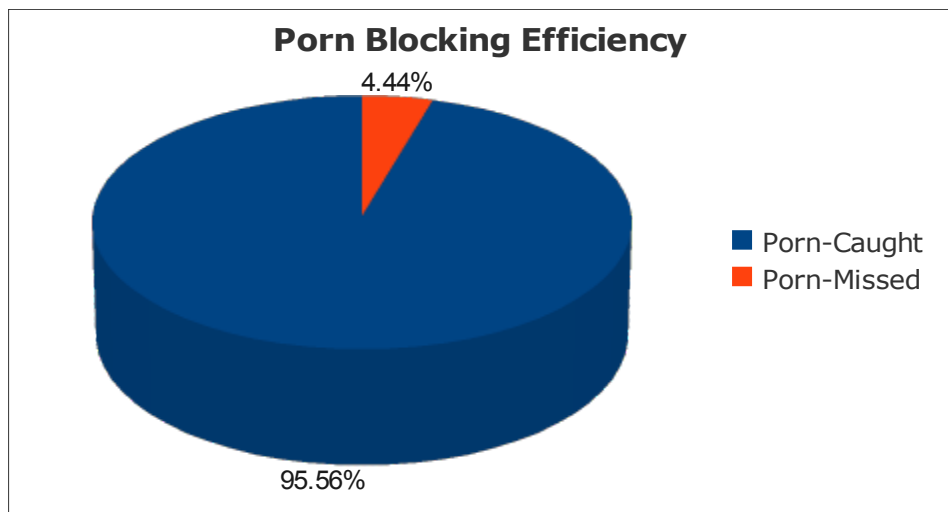


Chart 2. Porn Blocking Efficiency for MySurfGuard

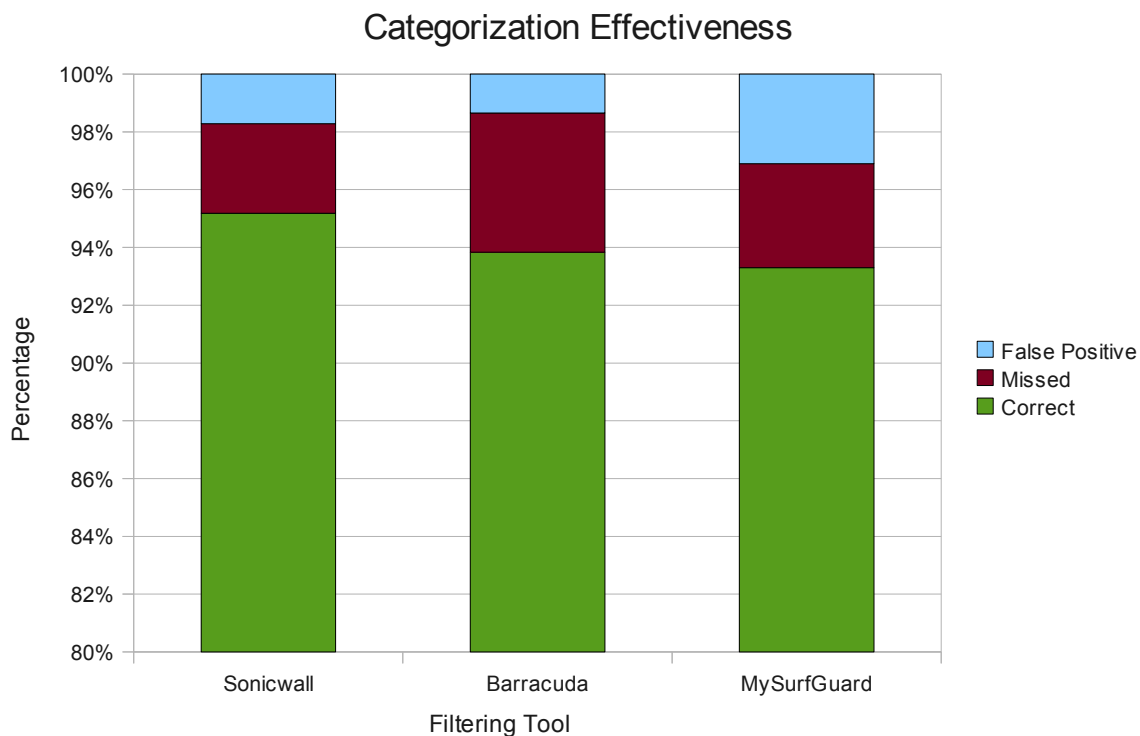


Chart 3. Categorization Effectiveness comparison with 3 filtering tools

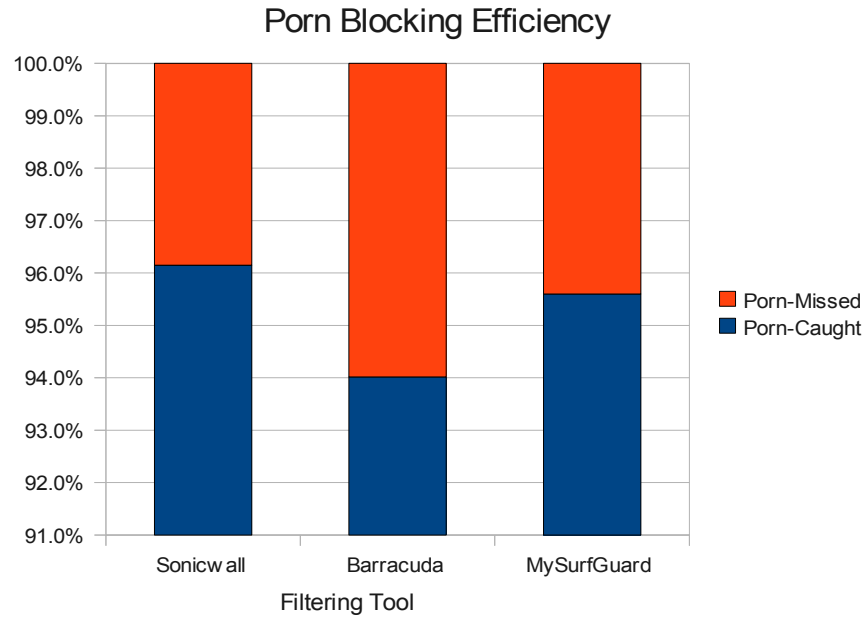


Chart 4. Porn Blocking Efficiency comparison with 3 filtering tools

## Conclusion

For effectiveness MySurfGuard scored 93.3% which is lower than Sonicwall's score of 95.2% but on par with Barracuda's score of 93.8%. We also see that Barracuda has the highest missed sites of 4.80%, followed by MySurfGuard's 3.60% and Sonicwall's 3.10%. As for porn blocking efficiency, MySurfGuard outperformed Barracuda by 1.60%. However Sonicwall tops the list among the three tools with a score of 96.1%. Hence we can conclude that MySurfGuard is a capable product and has comparable effectiveness and efficiency with other commercial products.

## **SECTION 4 : REFERENCES**

## References

- 1) <http://mysurfguard.oscc.org.my/?q=node/1>
- 2) <http://dansguardian.org/>
- 3) <http://www.untangle.com/About-Us/Deep-Throat-Fight-Club>
- 4) <http://www.sip-bench.eu/index.html>