JULY 2014

virus BULLETIN

Covering the global threat landscape

VBSPAM COMPARATIVE REVIEW JULY 2014

INTRODUCTION

On 1 July 2014, exactly 25 years after the first issue of *Virus Bulletin* was published, all of *VB*'s content – both past and future – was made freely available. Not only does this mean that the full archive and all new feature articles are available free of charge (and without registration), it also means that our test reports can be read by everyone as soon as they are published.

Because the reports are read by people with a wide range of interests – from product developers interested in the tiny details, to those looking for updates on 'the state of spam' – we have decided to split this month's report into two: an easy-to-digest summary of results, which also focuses on the state of spam, and this full report which contains all the technical details.

Alongside this change, there has been another major development behind the scenes: a new Security Test Engineer, Tony Oliveira, has joined *Virus Bulletin*, and will focus specifically on the VBSpam tests. Tony's long experience in systems administration has already proven very useful, and together we have fixed a number of hidden bugs that had been plaguing the test set-up for some time.

It goes without saying that we always make sure that such bugs don't affect the products in the test. In fact, this month each of the 15 participating full solutions achieved a VBSpam award – making it the second 'full house' in a row. What's more, there were eight solutions that achieved a VBSpam+ award for blocking more than 99.5% of spam, while generating no false positives at all in the ham corpus, and very few in the newsletter corpus.

THE TEST SET-UP

The VBSpam test methodology can be found at http://www.virusbtn.com/vbspam/methodology/. As usual, emails were sent to the products in parallel and in real time, and products were given the option to block email pre-DATA (that is, based on the SMTP envelope and before the actual email was sent). Three products chose to make use of this option.

For those products running on our equipment, we use *Dell PowerEdge* machines. As different products have different hardware requirements – not to mention those running on their own hardware, or those running in the cloud – there is little point comparing the memory, processing power or hardware the products were provided with; we followed the developers' requirements and note that the amount of email we receive is representative of that received by a smaller organization.

To compare the products, we calculate a 'final score', which is defined as the spam catch (SC) rate minus five times the weighted false positive (WFP) rate. The WFP rate is defined as the false positive rate of the ham and newsletter corpora taken together, with emails from the latter corpus having a weight of 0.2:

WFP rate = (#false positives + 0.2 * min(#newsletter false positives , 0.2 * #newsletters)) / (#ham + 0.2 * #newsletters)

Products earn VBSpam certification if the value of the final score is at least 98:

SC - (5 x WFP) ≥ 98

Meanwhile, products that combine a spam catch rate of 99.5% or higher with a lack of false positives and no more than 2.5% false positives among the newsletters earn a VBSpam+ award.

THE EMAIL CORPUS

The stability of the test network has been a little problematic in recent months, and this month was no exception: there

ISSN 1749-7027

was a one-hour period on 28 June when we decided the network wasn't stable enough for the required messages to be filtered directly and thus temporarily suspended the test, and a serious crash on 5 July meant we had to suspend the test for another 20 hours. The main reason for suspending the tests during these periods of instability was to follow a 'better safe than sorry' approach: we simply don't want to send test emails when circumstances are sub-optimal.

Thankfully, we were able to run the test for a day longer than initially planned, and as mentioned in the introduction, we have since found a bug that is the likely root cause of these problems. Indeed, as is typical for systems administration, one has a certain amount of control over how lucky one gets.

The test period thus started at 12am on Saturday 21 June and, with the aforementioned gaps, ended at 12am on Tuesday 8 July 2014.

The test corpus consisted of 132,881 emails. 122,485 of these were spam, 56,999 of which were provided by *Project Honey Pot*, with the remaining 65,486 emails provided by *spamfeed.me*, a product from *Abusix*. They were all relayed in real time, as were the 10,057 legitimate emails ('ham') and 339 newsletters.

Figure 1 shows the catch rate of all full solutions throughout the test. To avoid the average being skewed by poorly performing products, the highest and lowest catch rates have been excluded for each hour.

Comparing this graph with that published in the last test (carried out in May), we immediately noticed that the products' catch rate was worse than it was two months ago. Indeed, the average catch rate dropped from 99.78% to 99.70%, which means that a spam email was more than a third more likely to be missed this time around than in May.

Against this, however, stood a decrease in the false positive rates for both the ham corpus and the newsletter corpus. This isn't visible in the graph, but is at least as important, and it would be wrong to conclude that the drop in average catch rate meant that the overall performance of the spam filters was a little bit 'worse' than in the last test.

The first of the two low points in the graph was caused by the two emails with which products had most problems: two very brief 419 scams sent from an IP address belonging to a university in Venezuela. A number of seemingly unrelated emails were the cause of the second low point in the graph.

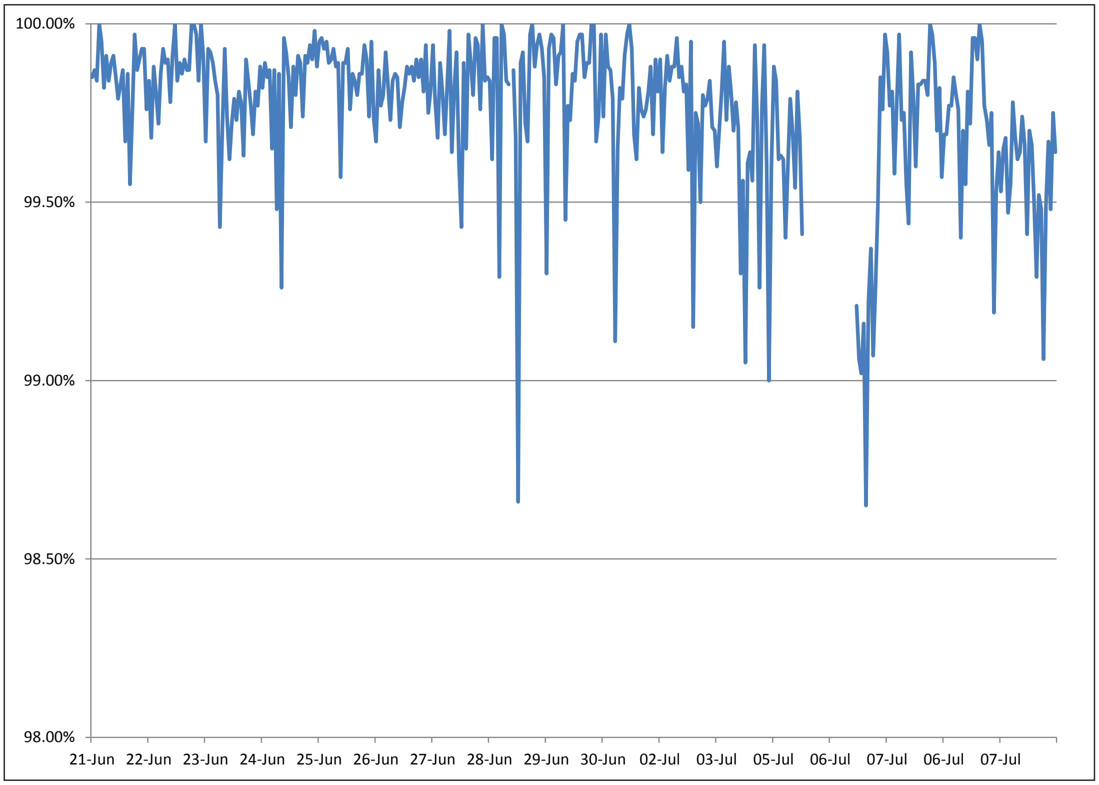


Figure 1: Spam catch rate of all full solutions throughout the test period.

RESULTS

In the text that follows, unless otherwise specified, 'ham' or 'legitimate email' refers to email in the ham corpus – which excludes the newsletters – and a 'false positive' is a message in that corpus that has been erroneously marked by a product as spam.

Because the size of the newsletter corpus is significantly smaller than that of the ham corpus, a missed newsletter has a much greater effect on the newsletter false positive rate than a missed legitimate email has on the false positive rate for the ham corpus (e.g. one missed email in the ham corpus results in an FP rate of less than 0.01%, while one missed email in the newsletter corpus results in an FP rate of almost 0.3%).

It should also be noted that, because of changes to the formula used to calculate the final score, these scores are not comparable with those documented prior to the March 2014 report.

Axway MailGate 5.3.1

SC rate: 99.41%
FP rate: 0.05%
Final score: 99.04
Project Honey Pot SC rate: 99.35%
Abusix SC rate: 99.47%
Newsletters FP rate: 3.8%

Going somewhat against the general trend, *Axway MailGate* blocked more spam emails in this test than it did in the last one. At the same time, the false positive rate increased slightly – although the product did block fewer newsletters than last time.

Four out of the five false positives were Russian emails from the same sender, and several of the blocked newsletters used a non-Latin character set. At the same time, a surprising amount of the spam the product missed used the Latin character set. This suggests that there may be a slight bias towards certain character sets in the filter – which is something for the developers to look at. For now, they should be happy with their well deserved third VBSpam award.

Bitdefender Security for Mail Servers 3.1.2

SC rate: 99.81%
FP rate: 0.00%
Final score: 99.80
Project Honey Pot SC rate: 99.65%
Abusix SC rate: 99.95%
Newsletters FP rate: 0.3%

Bitdefender missed more spam on this occasion than it has done since March 2013. That may sound bad, but in fact it still missed fewer than one in 500 spam emails (a lot of which used East-Asian character sets). Moreover, with no false positives and only one FP in the newsletter corpus, the product earned a VBSpam+ award.

There is thus plenty of reason for the developers to celebrate, especially since this is not only the product's 32nd VBSpam award in as many tests (the product has yet to miss a test), but also its tenth VBSpam+ award in a row.

Egedian Mail Security

SC rate: 99.45%
FP rate: 0.04%
Final score: 99.11
Project Honey Pot SC rate: 99.38%
Abusix SC rate: 99.51%
Newsletters FP rate: 4.4%

In this test, *Egedian Mail Security* proved that its very respectable performance on its debut in May wasn't a one-off: once again, the French product from *Profil Technology* achieved a VBSpam award. In comparison with the last test, the product's catch rate did drop a little – as part of the overall trend this month – but there were only four false positives, three of which were from the same Malawi-based sender.

Of course, we hope that *Egedian*'s developers will work towards getting the product to block the remaining 0.55% of spam – a lot of which had 'adult' content – but with a final score of 99.11, its second VBSpam award is certainly well deserved.

ESET Mail Security for Microsoft Exchange Server

SC rate: 99.87%
FP rate: 0.00%
Final score: 99.84
Project Honey Pot SC rate: 99.79%
Abusix SC rate: 99.93%
Newsletters FP rate: 0.9%

The results of this test should put smiles on the faces of *ESET*'s developers: the product slashed almost two-thirds off its false negative rate, missing just 163 spam emails (quite a few of which were written in German). To make this even more impressive, the product didn't block any of the more than 10,000 legitimate emails (and it blocked just three newsletters).

	True negatives	False positives	FP rate	False negatives	True positives	SC rate	Final score
Axway	10052	5	0.05%	718	121767	99.41%	99.04
Bitdefender	10057	0	0.00%	234	122251	99.81%	99.80
Egedian	10053	4	0.04%	672	121813	99.45%	99.11
ESET	10057	0	0.00%	163	122322	99.87%	99.84
FortiMail	10057	0	0.00%	122	122363	99.90%	99.86
GFI	10056	1	0.01%	227	122258	99.81%	99.72
IBM	10056	1	0.01%	318	122167	99.74%	99.69
Kaspersky LMS	10057	0	0.00%	180	122305	99.85%	99.83
Libra Esva	10057	0	0.00%	47	122438	99.96%	99.91
Netmail Secure	10057	0	0.00%	223	122262	99.82%	99.77
OnlyMyEmail	10057	0	0.00%	15	122470	99.99%	99.96
Scrollout	10049	8	0.08%	850	121635	99.31%	98.64
Sophos	10056	1	0.01%	654	121831	99.47%	99.40
SpamTitan	10056	1	0.01%	601	121884	99.51%	99.42
ZEROSPAM	10057	0	0.00%	421	122064	99.66%	99.59
Spamhaus DBL*	10057	0	0.00%	87620	34865	28.46%	28.46
Spamhaus ZEN*	10057	0	0.00%	12960	109525	89.42%	89.42

*Spamhaus is a partial solution and its performance is not to be compared with that of other products.

(Please refer to the text for full product names.)

Not only does this give the product the fourth-highest final score, it also means it maintains a clean sheet after a dozen tests – and now adds its fifth VBSpam+ award to its tally.

Fortinet FortiMail

SC rate: 99.90%
FP rate: 0.00%
Final score: 99.86
Project Honey Pot SC rate: 99.80%
Abusix SC rate: 99.98%
Newsletters FP rate: 1.2%

vb
VERIFIED
SPAM +
virusbtn.com

This test was an ‘average’ one for *Fortinet*’s *FortiMail* appliance: it almost repeated the scores it achieved in the last test, and the spam messages it missed were the kind of emails that many other products missed too.

Of course, in the case of *FortiMail*, ‘average’ actually means very good: the appliance has sat in our lab filtering emails for more than five years, never having missed a VBSpam award and typically ending up among the better performing products in each test. This month, *Fortinet*’s developers not only earn their 31st VBSpam award, but also, with the third-highest final score, a well-deserved VBSpam+ award (their fourth).

	Newsletters		Project Honey Pot		Abusix		pre-DATA[‡]		STDev[†]
	False positives	FP rate	False negatives	SC rate	False negatives	SC rate	False negatives	SC rate	
Axway	13	3.8%	369	99.35%	349	99.47%			1.14
Bitdefender	1	0.3%	199	99.65%	35	99.95%			0.99
Egedian	15	4.4%	351	99.38%	321	99.51%			1.03
ESET	3	0.9%	118	99.79%	45	99.93%			0.26
FortiMail	4	1.2%	112	99.80%	10	99.98%			0.24
GFI	5	1.5%	178	99.69%	49	99.93%			0.3
IBM	0	0.0%	251	99.56%	67	99.90%			0.42
Kaspersky LMS	2	0.6%	147	99.74%	33	99.95%			0.26
Libra Esva	5	1.5%	35	99.94%	12	99.98%	110384	90.12%	0.16
Netmail Secure	5	1.5%	186	99.67%	37	99.94%	110050	89.85%	0.28
OnlyMyEmail	3	0.9%	1	100.00%	14	99.98%			0.14
Scrollout	27	8.0%	158	99.72%	692	98.94%			0.83
Sophos	2	0.6%	431	99.24%	223	99.66%			0.65
SpamTitan	4	1.2%	297	99.48%	304	99.54%			1.02
ZEROSPAM	7	2.1%	331	99.42%	90	99.86%	108760	88.79%	0.45
Spamhaus DBL[*]	0	0.0%	31353	44.99%	56267	14.08%			10.92
Spamhaus ZEN[*]	0	0.0%	11776	79.34%	1184	98.19%			3.77

[*] Spamhaus is a partial solution and its performance is not to be compared with that of other products.

[‡] pre-DATA filtering was optional and was applied on the full corpus. All false positives occurred post-DATA.

[†] The standard deviation of a product is calculated using the set of its hourly spam catch rates.

(Please refer to the text for full product names.)

GFI MailEssentials

SC rate: 99.81%
FP rate: 0.01%
Final score: 99.72
Project Honey Pot SC rate: 99.69%
Abusix SC rate: 99.93%
Newsletters FP rate: 1.5%

This was a good test for *GFI MailEssentials*. The *Windows*-based solution blocked a lot more spam than it did in the last test, while at the same time reducing the number of false positives to a single blocked email.

Among the spam messages that were missed, there were some that probably linked to malware – although with malware links changing rapidly, it is impossible to know for certain. This isn't a particular problem for *GFI* though, more a mere reminder that even if your spam filter blocks an impressive 99.81% of spam (as *GFI* did in this test), clicking links in emails is never a good idea. With a final score of 99.72, *GFI* earns yet another VBSpam award to add to its collection.

Hosted solutions	Anti-malware	IPv6	DKIM	SPF	DMARC	Multiple MX-records	Multiple locations
OnlyMyEmail	Proprietary (optional)		√	√	*	√	√
ZEROSPAM	ClamAV			√		√	√

* OnlyMyEmail verifies DMARC status but doesn't provide feedback at the moment.

(Please refer to the text for full product names.)

Local solutions	Anti-malware	IPv6	DKIM	SPF	DMARC	Interface			
						CLI	GUI	Web GUI	API
Axway MailGate	Kaspersky; McAfee	√	√	√				√	
Bitdefender	Bitdefender	√				√		√	√
Egedian	Bitdefender, ClamAV	√				√		√	√
ESET	ESET Threatsense					√	√		
FortiMail	Fortinet	√	√	√		√		√	
GFI	Five anti-virus engines	√		√				√	
IBM	Sophos; IBM Remote Malware Detection			√		√		√	
Kaspersky LMS	Kaspersky	√		√		√		√	
Libra Esva	ClamAV; others optional		√	√		√		√	
Netmail Secure	Proprietary	√	√	√		√		√	
Scrollout	ClamAV			√		√		√	
Sophos	Sophos							√	
SpamTitan	Kaspersky; ClamAV	√	√	√		√		√	√

(Please refer to the text for full product names.)

IBM Lotus Protector for Mail Security

SC rate: 99.74%

FP rate: 0.01%

Final score: 99.69

Project Honey Pot SC rate: 99.56%

Abusix SC rate: 99.90%

Newsletters FP rate: 0.0%

VERIFIED SPAM virusbtn.com

In this test, *IBM* was the only full solution not to miss a single newsletter, which is no easy achievement. Unfortunately, *Lotus Protector for Mail Security* did miss a single legitimate email from the ham corpus, which was the only thing standing in the way of it earning its first VBSpam+ award.

An 'ordinary' VBSpam award was easily earned though – *IBM*'s 16th – and its developers should be pleased with the all-around improved performance.

Kaspersky Security 8 for Linux Mail Server

SC rate: 99.85%

FP rate: 0.00%

Final score: 99.83

Project Honey Pot SC rate: 99.74%

Abusix SC rate: 99.95%

Newsletters FP rate: 0.6%

After two VBSpam+ awards in a row, there wasn't much for *Kaspersky*'s *Linux*-based solution to improve upon, but in a test in which many products blocked less spam than on previous occasions, it managed to increase its catch rate to 99.85%, which means that only one in 680 spam emails was missed.

At the same time, it managed to avoid false positives and only missed two newsletters (written in Hebrew). Another VBSpam+ award is thus earned by the security giant – its sixth, and with the fifth highest final score in this test, one to be very proud of.

Libra Esva 3.3.2.0

SC rate: 99.96%

FP rate: 0.00%

Final score: 99.91

Project Honey Pot SC rate: 99.94%

Abusix SC rate: 99.98%

Pre-DATA SC rate: 90.12%

Newsletters FP rate: 1.5%

Only 47 spam emails were missed by *Libra Esva* in this test (fewer than all but one other product), each of which was missed by many other products as well. Alongside an impressive catch rate, there was a zero false positive rate for the Italian product in the ham corpus.

With only five missed newsletters, the product not only earns yet another VBSpam award (it now has one for every letter in the alphabet), it also achieves its ninth VBSpam+ award – one for each letter in its name.

Messaging Architects Netmail Secure

SC rate: 99.82%

FP rate: 0.00%

Final score: 99.77

Project Honey Pot SC rate: 99.67%

Abusix SC rate: 99.94%

Pre-DATA SC rate: 89.85%

Newsletters FP rate: 1.5%

In the last test, three false positives prevented *Netmail Secure* from achieving a VBSpam+ award. This time, however, it didn't block any emails from the ham corpus. What's more, the product actually blocked more spam emails than it did in the last test.

The Canadian company earns its fifth VBSpam+ award.

OnlyMyEmail's Corporate MX-Defender

SC rate: 99.99%

FP rate: 0.00%

Final score: 99.96

Project Honey Pot SC rate: 100.00%

Abusix SC rate: 99.98%

Newsletters FP rate: 0.9%

OnlyMyEmail's *Corporate MX-Defender* missed just 15 spam emails out of the more than 120,000 in this test. While this might be something that any other participant would be jealous of, for the Michigan-based hosted solution, it actually meant that it missed more spam than it has done in any test since March 2011. Of the false negatives, 14 were part of the same campaign and sent in very short succession – and subsequent emails from the same campaign were blocked.

There were yet again no false positives, and the product erroneously blocked just three newsletters – two in Hebrew and one in English – giving it the highest final score in the test, and thus well deserving of its sixth VBSpam+ award.

Scrollout F1

SC rate: 99.31%

FP rate: 0.08%

Final score: 98.64

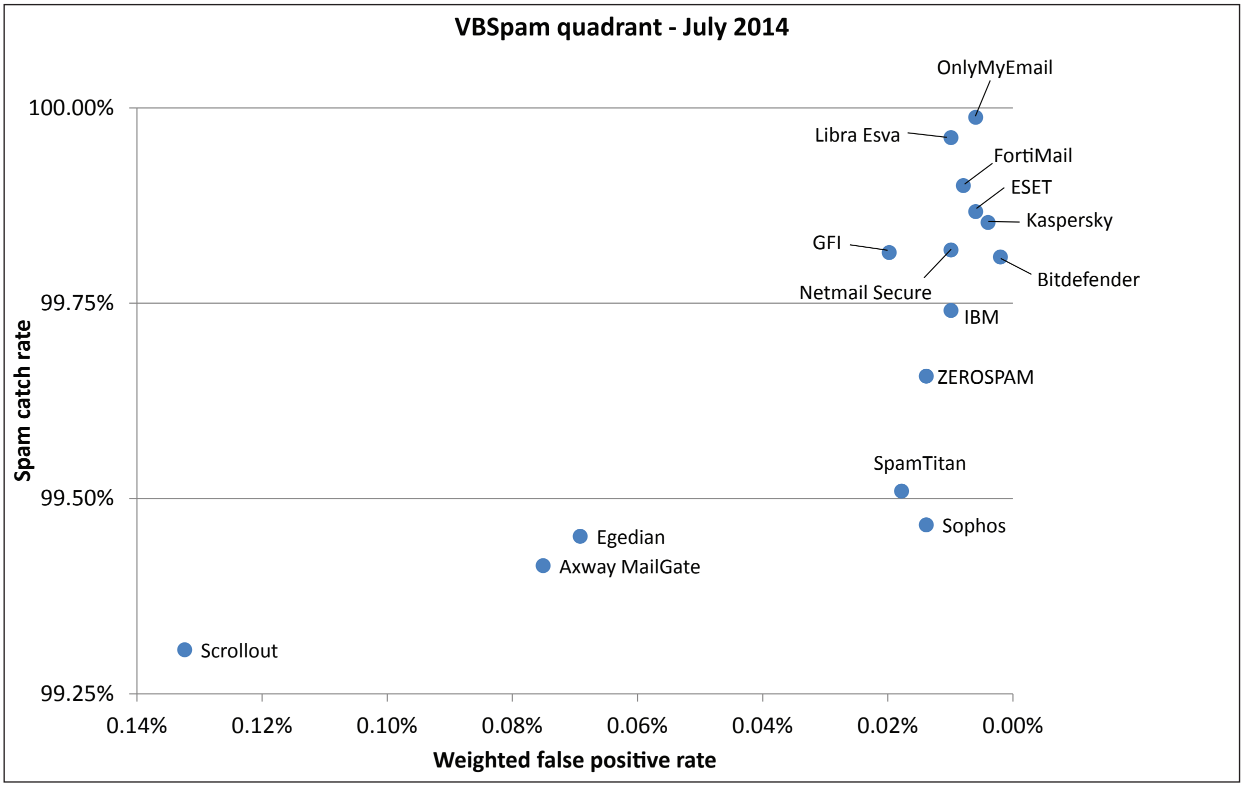
Project Honey Pot SC rate: 99.72%

Abusix SC rate: 98.94%

Newsletters FP rate: 8.0%

Before the start of this test, and at the request of the product's main developer, we installed a new version of *Scrollout F1* on a virtual machine in our lab. Installation is smooth and easy, showing that open source doesn't have to stand in the way of usability.

In previous tests, *Scrollout* has had some issues with a high false positive rate among the newsletters. While still blocking more of these than any other participant, its performance was a lot better on this occasion. At the same time, though, the product's spam catch rate took



(Please refer to full report for full product names and details.)

a bit of a nose-dive, and the product missed 850 spam emails – a surprising number of which were classic 'male enhancement' spam messages.

However, with the false positive rate stable at 0.08%, the product's final score remained well above the certification threshold, and thus *Scrollout* achieved another VBSpam award.

Sophos Email Appliance

SC rate: 99.47%
FP rate: 0.01%
Final score: 99.40
Project Honey Pot SC rate: 99.24%
Abusix SC rate: 99.66%
Newsletters FP rate: 0.6%

vb VERIFIED SPAM virusbtn.com

Sophos's *Email Appliance* was one of the products that saw its spam catch rate drop between the last test and this one, but here there is a simple explanation: the product had some difficulty parsing some unusual headers that had been added by the MTA we put between the product and the Internet. While it is ultimately the product's responsibility to make sure these headers, even if unusual, are parsed correctly, such a situation is unlikely to occur in real-world use of the product.

A single false positive – on an email discussing *Alice in Wonderland* – means that *Sophos Email Appliance* would not have achieved a VBSpam+ award anyway. However, even with the extra missed spam (close to half the false negatives were caused by the issue with the unusual headers), a final score of 99.40 is sufficient to earn the product its 25th VBSpam award.

SpamTitan 6.00

SC rate: 99.51%
FP rate: 0.01%
Final score: 99.42
Project Honey Pot SC rate: 99.48%
Abusix SC rate: 99.54%
Newsletters FP rate: 1.2%

vb VERIFIED SPAM virusbtn.com

SpamTitan's catch rate dropped 0.3 percentage points to 99.51% in this test, which was a bit disappointing to see; we noticed a fairly large number of 'get-rich-quick' schemes among the missed emails. Of course, this catch rate still means that more than 199 out of every 200 emails were blocked.

In fact, it could still have earned the product a VBSpam+ award if it weren't for four newsletters in three European languages and a single legitimate email from Malawi that was incorrectly blocked. *SpamTitan* thus earns another VBSpam award – the product's 29th in as many tests.

ZEROSPAM

SC rate: 99.66%
FP rate: 0.00%
Final score: 99.59
Project Honey Pot SC rate: 99.42%
Abusix SC rate: 99.86%
Pre-DATA SC rate: 88.79%
Newsletters FP rate: 2.1%

vb
VERIFIED
SPAM +
virusbtn.com

In this month's separately published summary, we looked at the image spam that was circulating during the first few days of the test. In general, products blocked this type of spam more easily than they blocked the average spam email, but *ZEROSPAM* was the exception – it missed more of these emails than any other product. It may well be that the reason for the lower overall catch rate the product saw this month lies deeper though, as the rate was more or less stable throughout the testing period.

Against this lower catch rate stood a lack of false positives which, together with a small enough number of blocked newsletters, means the Canadian hosted solution not only earned its 15th VBSpam award in as many tests, but also its fourth VBSpam+ award.

Spamhaus DBL

SC rate: 28.46%
FP rate: 0.00%
Final score: 28.46
Project Honey Pot SC rate: 44.99%
Abusix SC rate: 14.08%
Newsletters FP rate: 0.0%

Spamhaus ZEN

SC rate: 89.42%
FP rate: 0.00%
Final score: 89.42
Project Honey Pot SC rate: 79.34%
Abusix SC rate: 98.19%
Newsletters FP rate: 0.0%

Spamhaus's two DNS-based blacklists – the IP-based *ZEN* and the domain-based *DBL* – performed about as well as they did in May: *ZEN* a fraction worse, *DBL* a fraction better. More importantly, there was a complete lack of false positives (including newsletters) for both lists, demonstrating that they are both useful additions to many an anti-spam solution.

CONCLUSION

It has been an exciting month for *Virus Bulletin* and so will it have been for all 15 participating full solutions, as they all achieved a VBSpam award in the first test since they became freely available to all.

Of course, when looking at the details, things are more mixed and many developers will already have started working on improving their products' performance based on the feedback we have provided them with.

In the meantime, we too have been working on improving the performance of our test lab. We are looking forward to the next test that will hopefully be a very smooth experience for all concerned.

The next VBSpam test will run in August 2014, with the results scheduled for publication in September. Developers interested in submitting products should email martijn.grooten@virusbtn.com.

Editor: Martijn Grooten
Chief of Operations: John Hawes
Security Test Engineers: Scott James, Tony Oliveira
Sales Executive: Allison Sketchley
Editorial Assistant: Helen Martin
Perl Developer: Tom Gracey
Consultant Technical Editors: Dr Morton Swimmer, Ian Whalley

© 2014 Virus Bulletin Ltd, The Pentagon, Abingdon Science Park, Abingdon, Oxfordshire OX14 3YP, England.
Tel: +44 (0)1235 555139. Fax: +44 (0)1865 543153
Email: editorial@virusbtn.com
Web: http://www.virusbtn.com/