COMPARATIVE REVIEW

VBSPAM COMPARATIVE REVIEW SEPTEMBER 2010

Martijn Grooten

My spam isn't the same as your spam, which then isn't the same as the spam of the man playing with his *iPhone* next to you on the bus. That isn't too surprising: our respective email addresses may have ended up on different spammers' lists, and different spammers send different spam. But spam sent to addresses on one domain also differs from that sent to addresses on a different domain, and even groups of domains – where one might expect such differences to average out – receive spam that differs significantly.

We have always kept this in mind when running our anti-spam tests. Since we wanted the tests to provide a measure of performance that would be relevant to any organization, we didn't want to use the spam sent to a single domain, or even group of domains. This is the reason why we have been using *Project Honey Pot*'s spam feed for our tests. *Project Honey Pot* receives spam sent to a large number of spam traps on a large number of domains, distributed all over the world. By using this feed, we can be sure that products are being tested against spam that isn't any more likely to be received by someone in the UK than by someone in, say, New Zealand.

However, we always like to see things from a different perspective, and this is why we are very pleased to have developed a relationship with *Abusix*, a German company that also manages a large number of spam traps. From this test onwards, *Abusix* will provide us with a second spam corpus; in this test, and in all future tests, products will see spam from both streams (as well as a number of legitimate emails) and will be required to filter all of these emails correctly.

This month's test included 19 full solutions and one partial solution. For various reasons, a number of products that have participated in previous tests decided to sit this one out, but most of them expect to be back on the test bench next time. All of the full solutions tested this month achieved a VBSpam award. However, for several products there is still significant room for improvement and no doubt their developers will be working hard to see their products move towards the top right-hand corner of the VBSpam quadrant.

THE TEST SET-UP

The test methodology can be found at http://www.virusbtn.com/vbspam/methodology/. Email was sent to the products in parallel and in real time, and products were given the option to block email pre-DATA. Five products chose to make use of this option.

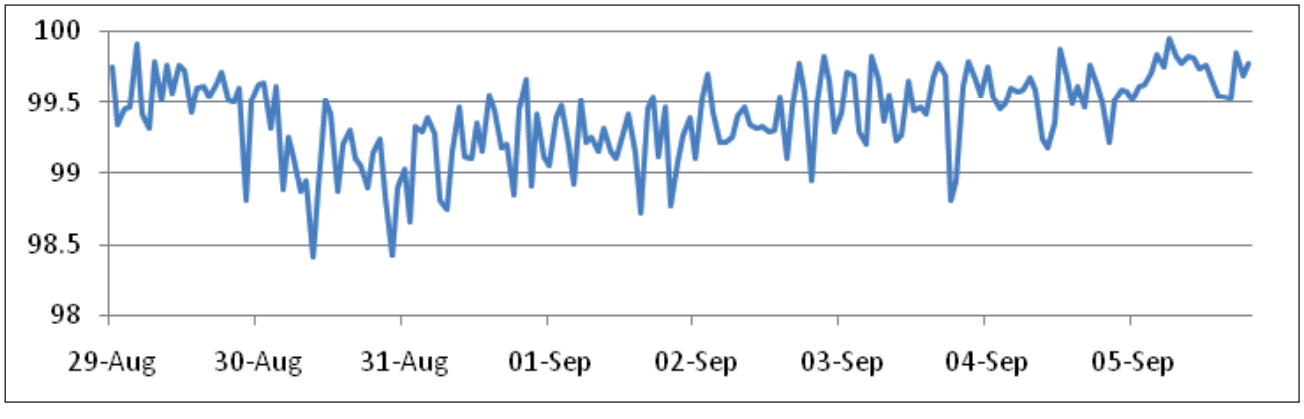
As in previous tests, the products that needed to be installed on a server were installed on a *Dell PowerEdge R200*, with a 3.0GHz dual core processor and 4GB of RAM. The *Linux* products ran on *SuSE Linux Enterprise Server 11*; the *Windows Server* products ran on either the 2003 or the 2008 version, depending on which was recommended by the vendor.

To compare the products, we calculate a 'final score', defined as the spam catch (SC) rate minus three times the false positive (FP) rate. Products earn VBSpam certification if this value is at least 96:

$$SC - (3 \times FP) \geq 96$$

THE EMAIL CORPUS

The test ran from midnight on 29 August 2010 to midnight on 6 September 2010, a period of eight full days. This was



Average spam catch rate throughout the test.

a shorter testing period than usual. A number of system crashes had caused the test network to be unreliable for several days after the test was initially started, and rather than using results from periods between the crashes when the network appeared to be working well, we decided to err on the side of caution and restart the whole test. The addition of a second spam stream and an increase in the size of the ham corpus gave us quantities of email comparable to those of previous tests despite the shorter test period.

The corpus contained 211,968 emails, 209,766 of which were spam. Of these spam emails 148,875 were provided by *Project Honey Pot* and 60,891 were provided by *Abusix*; in both cases they were relayed in real time, as were all legitimate messages, of which there were 2,202. The introduction of some new mailing lists (see *VB*, May 2010, p.24 for details), some of which are in foreign languages not previously included, means that seven out of the ten most commonly spoken languages in the world are now represented in the ham corpus.

The graph on the previous page shows the average spam catch rate for all products during every hour that the test ran (with the best and worst performing products removed from the computation of the averages). The graph shows that spam was harder to filter in certain periods than in others; for instance new spam campaigns tend to be harder to filter than ones that have been running for a while.

RESULTS

Anubis Mail Protection Service

SC rate: 99.93%
SC rate (image spam): 99.77%
SC rate (large spam): 99.63%
SC rate pre-DATA: N/A
FP rate: 0.05%
Final score: 99.80

Lisbon-based *AnubisNetworks*, the largest email security provider in Portugal, made a good debut in the previous test. The product's developers, however, were only mildly satisfied with the test results as they believed the product was capable of better. They were right – this month the product's spam catch rate increased, and the false positive rate was reduced to just a single missed email. With the second highest final score of this test, the developers should be very pleased with these results and the accompanying VBSpam award.

BitDefender Security for Mail Servers 3.0.2

SC rate: 99.91%
SC rate (image spam): 99.81%
SC rate (large spam): 99.20%
SC rate pre-DATA: N/A
FP rate: 0.00%
Final score: 99.91

I like it when developers have confidence in their product, and *BitDefender*'s developers demonstrated plenty of confidence when they were among the first to submit their product to the VBSpam tests in the early days. Despite this, they have never stopped trying to find ways to improve the product and have always been eager to hear feedback on its performance. *BitDefender* is the only product to have won a VBSpam award in every single VBSpam test – and with one of the highest catch rates in this test, and no false positives, it outperforms all other products and achieves the highest final score this month.

Fortinet FortiMail

SC rate: 98.44%
SC rate (image spam): 97.34%
SC rate (large spam): 95.98%
SC rate pre-DATA: N/A
FP rate: 0.05%
Final score: 98.30

One of the products that has been filtering mail quietly ever since its introduction to the tests, *FortiMail* wins its eighth VBSpam award in as many attempts. It does so with a nicely improved performance, demonstrating that the product's developers are keeping up with the latest spam campaigns.

Kaspersky Anti-Spam 3.0

SC rate: 98.30%
SC rate (image spam): 98.25%
SC rate (large spam): 97.37%
SC rate pre-DATA: N/A
FP rate: 0.05%
Final score: 98.16

Neither the new ham nor the new spam stream proved to be a problem for *Kaspersky*.

The company's *Linux* product saw its false positive rate improve, while barely compromising on the spam catch rate. *Kaspersky* easily wins another VBSpam award.

Libra Esva 2.0

SC rate: 99.96%
SC rate (image spam): 99.92%
SC rate (large spam): 99.71%
SC rate pre-DATA: 97.93%
FP rate: 0.32%
Final score: 99.01

Once again, *Libra Esva* had one of the highest spam catch rates of all products. Compared to previous tests, the product scored a slightly higher false positive rate – whilst this is something for the developers to pay attention to, the FP rate was still only average. With another very respectable final score, the Italian product wins its third consecutive VBSpam award.

M86 MailMarshal SMTP

SC rate: 99.97%
SC rate (image spam): 99.96%
SC rate (large spam): 99.93%
SC rate pre-DATA: N/A
FP rate: 0.5%
Final score: 98.47

M86's *MailMarshal* blocked the second largest amount of spam of all the products in this test, which is quite an achievement. Unfortunately, the product also missed almost a dozen legitimate emails, which lowered its final score quite significantly. It was still decent though, earning the product its sixth VBSpam award, but the developers will need to concentrate on reducing the FP rate, while not compromising too much on the amount of spam caught.

McAfee Email Gateway (formerly IronMail)

SC rate: 97.81%
SC rate (image spam): 93.08%
SC rate (large spam): 96.85%
SC rate pre-DATA: N/A
FP rate: 0.45%
Final score: 96.45

McAfee's *Email Gateway Appliance* suffered what appeared to be a temporary glitch during this test – with a disappointing spam catch rate towards the start of the test improving to see scores of over 99% during the final days of the test. Despite a small number of false positives, the product still earns a VBSpam award, but the product's developers will no doubt be working hard to determine the cause of the earlier problems and to ensure its spam catch rate remains consistently high in future.

McAfee Email and Web Security Appliance

SC rate: 99.05%
SC rate (image spam): 92.98%
SC rate (large spam): 90.20%
SC rate pre-DATA: N/A
FP rate: 0.27%
Final score: 98.23

McAfee's *Email and Web Security Appliance* achieved a VBSpam award in the previous test, but with a rather low final score. It was good to see that this appears to have been a one-off dip, rather than a serious problem with the installation; a high spam catch rate combined with a small handful of false positives easily earns the product its seventh VBSpam award.

MessageStream

SC rate: 99.08%
SC rate (image spam): 99.68%
SC rate (large spam): 99.56%
SC rate pre-DATA: N/A
FP rate: 0.09%
Final score: 98.81

As a product whose customers are based mostly in the Anglo-Saxon world, correctly filtering email in foreign languages may not be a high priority for *MessageStream*. However, in an industry where the devil is in the details, the developers have taken good care of even these details: a spam catch rate of over 99%, combined with just two false positives, means that the hosted solution more than deserves its eighth VBSpam award.

	True negative	False positive	FP rate	False negative	True positive	SC rate	Final score
AnubisNetworks	2201	1	0.05%	144	233321	99.93%	99.80
BitDefender	2202	0	0.00%	192	233232	99.91%	99.91
FortiMail	2201	1	0.05%	3281	229721	98.44%	98.30
Kaspersky	2201	1	0.05%	3567	229709	98.30%	98.16
Libra Esva	2195	7	0.32%	76	233387	99.96%	99.01
M86 MailMarshal	2191	11	0.50%	60	233408	99.97%	98.47
McAfee Email Gateway	2192	10	0.45%	4587	207284	97.81%	96.45
McAfee EWS	2196	6	0.27%	1999	231362	99.05%	98.23
MessageStream	2200	2	0.09%	1931	231179	99.08%	98.81
Messaging Architects M+Guardian	2182	20	0.91%	111	233291	99.95%	97.22
Pro-Mail	2201	1	0.05%	3607	229309	98.28%	98.15
Sophos	2199	3	0.14%	173	233273	99.92%	99.51
SPAMfighter	2199	3	0.14%	2795	230473	98.67%	98.26
SpamTitan	2188	14	0.64%	1942	231321	99.07%	97.17
Symantec Brightmail	2202	0	0.00%	745	232511	99.64%	99.64
The Email Laundry	2197	4	0.18%	392	233014	99.81%	99.27
Vade Retro	2194	8	0.36%	1175	232230	99.44%	98.35
Vamsoft ORF	2194	8	0.36%	1418	231616	99.32%	98.24
Webroot	2188	14	0.64%	18	233294	99.99%	98.08
Spamhaus ZEN	2202	0	0.00%	18119	211304	91.36%	91.36

Messaging Architects M+Guardian

SC rate: 99.95%
SC rate (image spam): 99.94%
SC rate (large spam): 99.85%
SC rate pre-DATA: 94.89%
FP rate: 0.91%
Final score: 97.22

Quite understandably, *M+Guardian*'s developers were not happy with their product's performance in the last test – in which it failed to achieve a VBSpam award. They looked into the settings of the appliance and among the changes they made was to turn on XCLIENT; this way they could use pre-DATA filtering, which they believe is one of the core benefits of the product.

Indeed, almost 94.9% of the spam was blocked this way, while the subsequent content filtering left less than 0.1% of spam unfiltered. An excellent spam catch rate, and *M+Guardian* easily reclaims its VBSpam award. However, there will be some disappointment for the developers over an incorrectly blocked domain which accounted for 15 of the 20 false positives.

Pro-Mail (Prolocation)

SC rate: 98.28%
SC rate (image spam): 99.66%
SC rate (large spam): 93.93%
SC rate pre-DATA: N/A
FP rate: 0.05%
Final score: 98.15

Like several anti-spam solutions, *Pro-Mail*, the hosted solution that debuted in the last test, classifies email into not

	Project Honey Pot		Abusix		Image spam*		Large spam*		pre-DATA†		St. dev‡
	FN	SC rate	FN	SC rate	FN	SC rate	FN	SC rate	FN	SC rate	
AnubisNetworks	138	99.91%	6	99.99%	11	99.77%	5	99.63%			0.14
BitDefender	112	99.93%	80	99.87%	9	99.81%	11	99.20%			0.15
FortiMail	2449	98.36%	832	98.63%	126	97.34%	55	95.98%			0.89
Kaspersky	2680	98.21%	887	98.54%	83	98.25%	36	97.37%			2.38
Libra Esva	58	99.96%	18	99.97%	4	99.92%	4	99.71%	4452	97.93%	0.09
M86 MailMarshal	39	99.97%	21	99.97%	2	99.96%	1	99.93%			0.11
McAfee Email Gateway	4576	96.94%	11	99.98%	328	93.08%	43	96.85%			2.10
McAfee EWS	1891	98.73%	108	99.82%	333	92.98%	134	90.20%			1.38
MessageStream	1066	99.29%	865	98.58%	15	99.68%	6	99.56%			0.69
Messaging Architects M+Guardian	100	99.93%	11	99.98%	3	99.94%	2	99.85%	10970	94.89%	0.13
Pro-Mail	2886	98.07%	721	98.82%	16	99.66%	83	93.93%			1.60
Sophos	144	99.90%	29	99.95%	17	99.64%	3	99.78%			0.22
SPAMfighter	1567	98.95%	1228	97.98%	139	97.07%	94	93.12%			2.55
SpamTitan	1423	99.05%	519	99.15%	2	99.96%	24	98.24%			1.10
Symantec Brightmail	429	99.71%	316	99.48%	4	99.92%	4	99.71%			0.36
The Email Laundry	336	99.78%	56	99.91%	2	99.96%	4	99.71%	11136	94.82%	0.24
Vade Retro	645	99.57%	530	99.13%	10	99.79%	27	98.02%			0.81
Vamsoft ORF	1232	99.18%	186	99.69%	46	99.03%	40	97.07%			0.51
Webroot	15	99.99%	3	100.00%	2	99.96%	9	99.34%	77506	63.92%	0.05
Spamhaus ZEN	10749	92.69%	7370	86.55%	378	92.03%	126	90.78%	18119	91.36%	3.35

* There were 4,743 spam messages containing images and 1,367 considered large; the two are not mutually exclusive.
† Pre-DATA filtering was optional and was applied on the full spam corpus.
‡ The standard deviation of a product is calculated using the set of its hourly spam catch rates.

two but three categories: ham, spam and 'possibly spam'. Messages that fall into the 'possibly spam' category are not blocked by the product but, as a header is added, can be put into a separate folder. Emails in this category were considered to have been marked as ham in this test, which may explain the product's relatively low spam catch rate. It was still decent enough for the product to win a VBSpam award though, and with just one false positive, it would be interesting to see what effect a stricter filtering policy would have.

Sophos Email Appliance

SC rate: 99.92%
SC rate (image spam): 99.64%
SC rate (large spam): 99.78%
SC rate pre-DATA: N/A
FP rate: 0.14%
Final score: 99.51

There is a reason why we run an anti-spam test every two months: while one decent performance is certainly a promising sign, what really matters is that a product

manages to perform well repeatedly. With four good sets of results in as many VBSpam tests – each time achieving a final score among the top seven in the test – the *Sophos Email Appliance* certainly satisfies that criterion and adds another VBSpam award to its collection.

SPAMfighter Mail Gateway

SC rate: 98.67%
SC rate (image spam): 97.07%
SC rate (large spam): 93.12%
SC rate pre-DATA: N/A
FP rate: 0.14%
Final score: 98.26

It has been a while since I last needed to log into the admin interface of *SPAMfighter*. That is a good thing, but what is even better is that the product's developers have been working on their product in the meantime and upgrades have been downloaded automatically. This test saw improvements to both the spam catch rate and the false positive rate and, consequently, a significant improvement to the product's final score, winning *SPAMfighter* its sixth consecutive VBSpam award.

SpamTitan

SC rate: 99.07%
SC rate (image spam): 99.96%
SC rate (large spam): 98.24%
SC rate pre-DATA: N/A
FP rate: 0.64%
Final score: 97.17

SpamTitan is one of several products that suffered from more than a handful of false positives in this test. False positives are undesirable and customers are unlikely to accept them unless the spam catch rate of the product is exceptional. *SpamTitan*'s spam catch rate is very good – pushing the product's final score up to above the VBSpam threshold – but the developers will no doubt be spending some time scrutinizing the false positive samples in an attempt to improve the product's position on the VBSpam quadrant.

Symantec Brightmail Gateway 9.0

SC rate: 99.64%
SC rate (image spam): 99.92%
SC rate (large spam): 99.71%
SC rate pre-DATA: N/A
FP rate: 0.00%
Final score: 99.64

A product that manages to increase an already excellent spam catch rate, while eliminating the single false positive that pestered it in the previous test clearly deserves a VBSpam award. *Symantec*'s *Brightmail Gateway* virtual appliance did exactly that, completing this test with the third highest final score and the product's fifth VBSpam award.

The Email Laundry

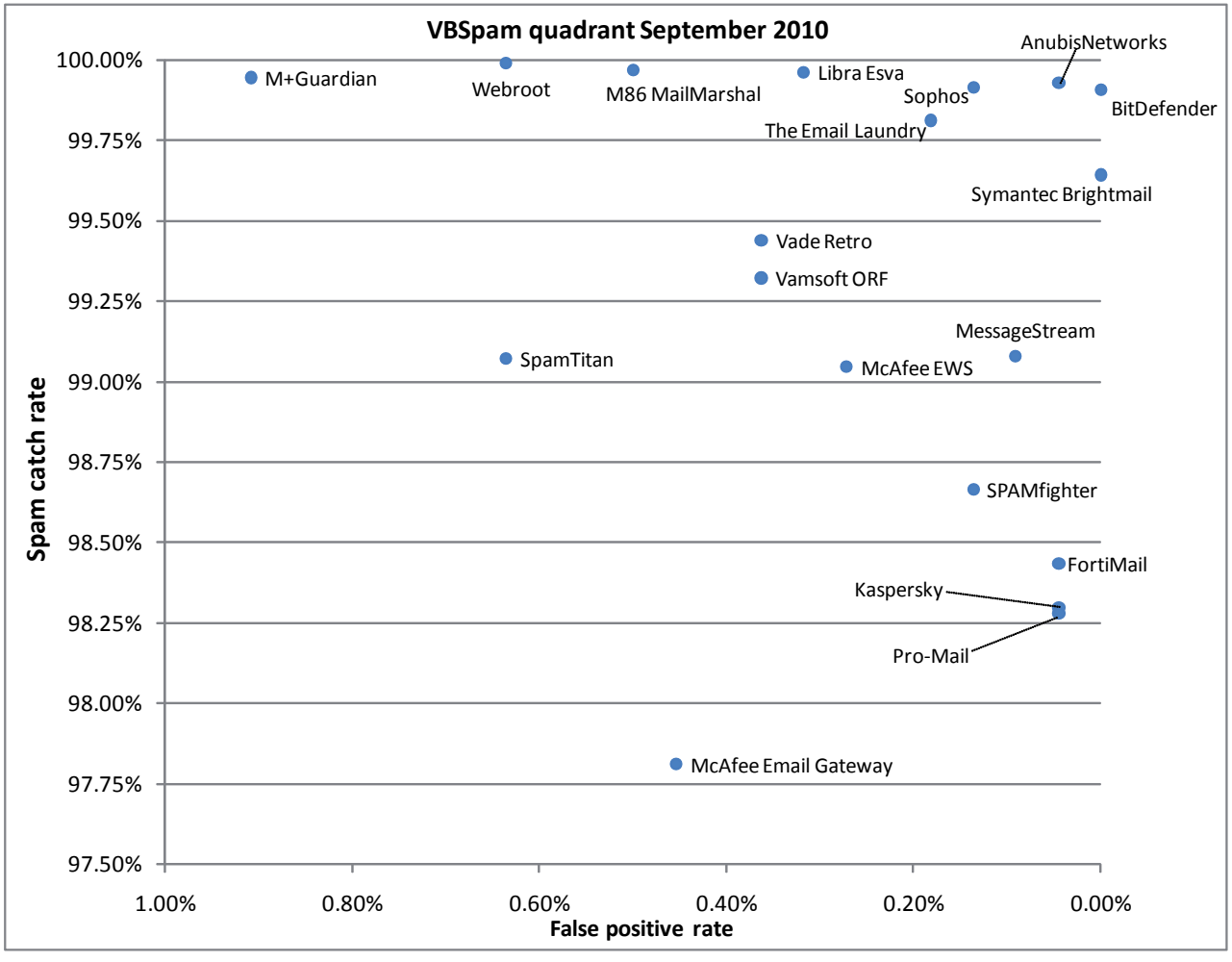
SC rate: 99.81%
SC rate (image spam): 99.96%
SC rate (large spam): 99.71%
SC rate pre-DATA: 94.82%
FP rate: 0.18%
Final score: 99.27

The significant drop in *The Email Laundry*'s pre-DATA catch rate since the last test deserves some explanation. The drop does not necessarily mean that the product's spam-filtering performance has worsened, but that spam has changed and, consequently, blocking on senders' domains and IP addresses wasn't as effective this month as it was in previous months.

What matters to the user is the percentage of spam that makes it to the inbox and this has decreased a fraction. There were a few false positives this time, but not enough to stop the hosted solution from achieving the fifth highest final score and earning a VBSpam award.

Vade Retro Center

SC rate: 99.44%
SC rate (image spam): 99.79%
SC rate (large spam): 98.02%
SC rate pre-DATA: N/A
FP rate: 0.36%
Final score: 98.35



Vade Retro is the market leader in France, but international spam is no problem for the product and it saw its catch rate improve significantly this month. With a small number of exceptions, legitimate email in foreign languages proved no problem either. The product thus wins its third VBSpam award in as many tests and with its best results to date.

Vamsoft ORF

SC rate: 99.32%
SC rate (image spam): 99.03%
SC rate (large spam): 97.07%
SC rate pre-DATA: N/A
FP rate: 0.36%
Final score: 98.24

No doubt *ORF*'s developers will be frustrated with two senders in this month's ham corpus, each of which caused four false positives, thus breaking their zero false positive record to date. However, it should be seen as a gentle reminder to all developers that no one can ignore the problem of false positives. Moreover, an improved spam catch rate means the product still achieved a decent final score and thus wins its third VBSpam award.

Webroot Email Security Service

SC rate: 99.99%
SC rate (image spam): 99.96%
SC rate (large spam): 99.34%
SC rate pre-DATA: 63.92%

FP rate: 0.64%

Final score: 98.08

Webroot was one of five products filtering email pre-DATA. It did not block as many emails during this stage as other products did, but this is not a sign that something is wrong with the product: it reflects a choice made by the developers as to where spam is filtered. And with more spam blocked than any other product, *Webroot*'s choice appears to be a good one. Unfortunately, there were a number of false positives this time, but the product easily earned another VBSpam award – its seventh to date.

Spamhaus ZEN

SC rate: 91.36%

SC rate (image spam): 92.03%

SC rate (large spam): 90.78%

SC rate pre-DATA: 91.36%

FP rate: 0.00%

Final score: 91.36

We owe an apology to *The Spamhaus Project*, as a bug on our side caused the *Spamhaus DBL* – the domain blacklist that in previous tests worked so well alongside *Spamhaus*'s *ZEN* blacklist – to fail during the running of this test. This is a shame, especially since *Spamhaus ZEN* – which combines three IP blacklists – performed significantly less well here than in previous tests.

It is important to realize that *Spamhaus* is a partial solution and is not supposed to be applied on its own. And while, together with the *DBL*, it is still recommended that the blacklists be supplemented with a content filter, the *DBL* is supposed to work especially well together with the organization's IP blacklists. What we can see is that during a period when pre-DATA filtering has produced worse results than during previous periods, *Spamhaus* is still a reliable first line of defence against spam – in particular because, once again, no legitimate emails were blocked.

CONCLUSION

For some products, the addition of a second spam stream and/or the new emails added to the ham corpus this month has given them something to work on; developers of other products will be trying to repeat this month's performance. As always, we will be working hard too – perhaps even harder than before. After nine successful tests, the VBSpam set-up is ready to go '2.0'.

Products ranked by final score	Final score
BitDefender	99.91
AnubisNetworks	99.80
Symantec Brightmail	99.64
Sophos	99.51
The Email Laundry	99.27
Libra Esva	99.01
MessageStream	98.81
M86 MailMarshal	98.47
Vade Retro	98.35
FortiMail	98.30
SPAMfighter	98.26
Vamsoft ORF	98.24
McAfee EWS	98.23
Kaspersky	98.16
Pro-Mail	98.15
Webroot	98.08
M+Guardian	97.22
SpamTitan	97.17
McAfee Email Gateway	96.45

For readers of the comparative reviews, little to nothing will change, but the new set-up will ensure greater system stability and allow room for the tests to grow bigger. Moreover, the provision of feedback on products' performance to the participants – most of which has been done manually until now – will be semi-automated, saving considerable time.

The next test is due to run throughout October, with results published in the November issue of *Virus Bulletin*. The deadline for submission of products will be Friday 24 September. Any developers interested in submitting a product should email martijn.grooten@virusbtn.com.