COMPARATIVE REVIEW

ANTI-SPAM COMPARATIVE REVIEW SEPTEMBER 2009

Martijn Grooten

This month's VBSpam comparative review sees an increase in the field of competitors for the third time in a row. Starting out with a modest six products on the test bench in the first test (see *VB*, May 2009, p.S5), this month sees that figure doubled, with a total of 12 products lined up on the bench: eight of the products that took part in the last VBSpam test (see *VB*, July 2009, p.25) are joined by four new ones. To date, there have been few other anti-spam tests with as many participating products.

We hope that our tests will make a valid contribution to the anti-spam community, helping the community answer questions about which anti-spam methods work better than others, and helping developers find ways to improve their products. For me, the best part of conducting these tests is hearing developers say that they have made improvements to their product upon receiving our feedback on its performance.

A total of nine VBSpam awards were given out this month, but only one of these was at the Platinum level, leaving most developers with something to improve upon. But even those achieving a Platinum award have good reason to look carefully at their product's performance: with spam changing constantly, a filter that isn't kept up to date – even a very good one – will soon start to fall behind.

THE TEST SET-UP

A few changes were made to the set-up after the last test. However, these were mostly of a technical nature and designed to help the test run more smoothly, and thus should not have affected the test itself. The full methodology can be found at http://www.virusbtn.com/vbspam/methodology/. Readers who wonder about the relatively high false positive rates measured in our tests (compared with those seen in other tests and those claimed by the developers themselves) are advised to consult the last review (see *VB*, July 2009, p.25) for explanation. Finally, as has been mentioned previously, the nature of this test is comparative, and as such it is important to note that it is not so much the absolute performance of a product that matters, but the relative performance compared to that of its competitors.

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. Those running on *Linux* ran on *SuSE Linux Enterprise Server 11*; the *Windows Server* products ran either the 2003 or the 2008 version, depending on which was recommended by the vendor.

THE EMAIL CORPUS

The test ran from 16:45 h on 7 August 2009 to 08:00 h on 27 August 2009. The corpus consisted of all emails – ham and spam – sent to '@virusbtn.com' addresses mixed with a spam stream provided by Project Honey Pot. Emails from both sources were sent through the products in real time.

While the test was running, we noticed a downside to its popularity: with so many products running on the same network, many of which perform regular Internet look-ups, the Internet connection was put under considerable strain and during some periods wasn't as reliable as we would have liked it to have been. It is interesting to see how different products react to this situation – which could easily occur in a real-world environment – and how much their performance suffers.

But of course, we do not want uncontrolled and unannounced circumstances to influence our test. Therefore we looked carefully at the network's performance while the test was running; emails that were received during periods for which we cannot be absolutely sure the network performance was reliable have been eliminated from the test. It should be noted that this has been done without looking at whether or how this affected individual products, and the final corpus used was still large enough for the results to give a good reflection of the products' performances.

This corpus contained 1,275 ham messages and 19,401 spam messages sent to *VB* addresses. It also contained 294,338 spam messages from Project Honey Pot; these emails reflect the global nature of spam and the fact that different addresses and domains do sometimes receive different kinds of spam. The total corpus thus contained 315,014 messages.

BitDefender Security for Mail Servers 3.0.2

SC rate (total): 98.74%
SC rate (Project Honey Pot corpus): 99.29%
SC rate (VB spam corpus): 90.41%
FP rate: 0.87%
FP rate of total VB mail corpus: 0.053%

Romanian company *BitDefender* submitted its *Linux* product for the third time this month, and its developers were eager to improve upon the Silver VBSpam award they won in July. On that occasion the product missed out on the higher-level awards because it was eager to block some

legitimate emails from countries that use different character sets – a tempting idea perhaps, as a large volume of such mail is spam, but the practice could, in fact, lead to end-users missing important messages. Fixing this issue saw the product's false positive rate drop significantly, while still retaining a high spam catch rate, and as such it is the deserved winner of a VBSpam Gold award, only narrowly missing out on a top-level Platinum award.

ClamAV using Sanesecurity signatures

SC rate (total): 85.40%
SC rate (Project Honey Pot corpus): 86.43%
SC rate (VB spam corpus): 69.83%
FP rate: 0.39%
FP rate of total VB mail corpus: 0.024%

From the outset, the developer of the *Sanesecurity* signatures that work with *ClamAV* did not expect to win a VBSpam award for his product, which is generally used together with other solutions, but he was eager to hear feedback on its performance. After receiving feedback from the last test, the developer made some adjustments to the product, resulting in a significantly improved performance this time around. A spam catch rate of over 85%, together with fewer false positives than all but one product, is a good score indeed; not quite enough to win an award, but nevertheless impressive for a product based on many hours of voluntary work.

Fortinet FortiMail

SC rate (total): 99.04%
SC rate (Project Honey Pot corpus): 99.20%
SC rate (VB spam corpus): 96.64%
FP rate: 2.25%
FP rate of total VB mail corpus: 0.135%

Fortinet's *FortiMail* achieved a VBSpam Silver award for its first performance in our tests in July. An extensive logging system enabled us to provide feedback to its developers on the types of emails that were being blocked incorrectly and which anti-spam tests these emails had failed. As a result of subsequent tweaks made to the product, *FortiMail*'s false positive rate dropped, while its spam catch rate remained high; in fact, its performance against the spam received directly by *VB* (the VB spam corpus) was better than that of any other product. The product's scores were not quite sufficient to earn a Gold award, but another VBSpam Silver award should spur the developers on to do even better next time.

Kaspersky Anti-Spam 3.0

SC rate (total): 98.39%
SC rate (Project Honey Pot corpus): 99.01%
SC rate (VB spam corpus): 88.92%
FP rate: 0.63%
FP rate of total VB mail corpus: 0.039%

Kaspersky's anti-spam product, usually referred to as *KAS* and running on *Linux*, is an excellent demonstration of the fact that using a locally installed product does not necessarily mean a lot of extra work for the system administrator: from the point at which it was first set up in our test network in June, we have had hardly any reason to look at the product. As in the previous test, the product's false positive rate was very low. The spam catch rate was slightly higher this time than in the previous test, but with the benchmarks having become stricter this month, it was not quite enough to earn a Platinum award. A VBSpam Gold award is thus earned by *Kaspersky* in recognition of a better-than-average performance on both accounts.

McAfee Email Gateway (formerly IronMail)

SC rate (total): 99.60%
SC rate (Project Honey Pot corpus): 99.87%
SC rate (VB spam corpus): 95.62%
FP rate: 1.27%
FP rate of total VB mail corpus: 0.077%

McAfee Email Gateway appliance, previously known as both *Secure Mail* and *IronMail*, was originally developed by *Secure Computing*, which was bought by *McAfee* in 2008. The product was set up easily using a web interface. I was particularly interested in seeing how well the product performed in the blocking of spam as this was the only product in the test that was configured to scan the contents of incoming email during the SMTP transaction and block those emails it was certain were spam. When done well, this can save a significant number of resources.

Helped by this two-layered blocking, the product's spam catch rate of 99.60% was higher than that of any other. The cost of this was a number of false positives though, including a few non-bulk emails, although it should be added that none of these were blocked at the SMTP level and would thus have ended up in the quarantine rather than being discarded outright. Moreover, the number of false positives was still below average, which means that *McAfee Email Gateway* debuts with a well-deserved VBSpam Gold award.

McAfee Email and Web Security Appliance

SC rate (total): 99.39%
SC rate (Project Honey Pot corpus): 99.63%
SC rate (VB spam corpus): 95.72%
FP rate: 0.24%
FP rate of total VB mail corpus: 0.015%

McAfee's *Email and Web Security Appliance* can do a lot more than just filter spam – which perhaps isn't surprising for an appliance made by a well-known anti-virus vendor. We did not look beyond the product's email filtering capacity, but even here the extensive web interface gives the interested system administrator many settings to tinker with.

In the way in which it was set up for this test, there was little need to think about modifying the settings. Not only did the product have one of the highest spam catch rates, it combined that with the lowest false positive rate of all products. As a result of this stellar performance, the *Email and Web Security Appliance* earns a VBSpam Platinum award, the only one of its kind in this test.

MessageStream

SC rate (total): 98.65%
SC rate (Project Honey Pot corpus): 99.01%
SC rate (VB spam corpus): 93.24%
FP rate: 0.78%
FP rate of total VB mail corpus: 0.048%

The developers of *MessageStream*, the hosted solution from British company *Giacom*, used some of our feedback from the last test to change the settings in their spam filter and reduce the number of false positives. As a result, *MessageStream*'s false positive rate halved compared to that of the previous test, while barely affecting the spam catch rate. Thus, despite even stricter benchmarks in this test, the product completes a hat-trick of three VBSpam Gold awards in a row.

Messaging Architects M+Guardian

SC rate (total): 98.78%
SC rate (Project Honey Pot corpus): 99.20%
SC rate (VB spam corpus): 92.41%
FP rate: 1.11%
FP rate of total VB mail corpus: 0.068%

M+Guardian, a hardware appliance developed by Canadian company *Messaging Architects*, achieved VBSpam Platinum awards in both of the preceding tests, giving a good indication of its capabilities. Of course, the spam landscape changes over time and success in the past (or even in the present) does not guarantee success in the future. Indeed, a slightly higher false positive rate and a slightly lower spam catch rate mean that on this occasion the product wins a VBSpam Gold award. A commendable achievement, but to regain a Platinum-level award the product's developers have some work to do.

Microsoft Forefront Security for Exchange Server v.11

SC rate (total): 99.51%
SC rate (Project Honey Pot corpus): 99.77%
SC rate (VB spam corpus): 95.53%
FP rate: 2.00%
FP rate of total VB mail corpus: 0.121%

The founder of *Microsoft* once famously predicted that spam would be a thing of the past within two years, but thankfully *Microsoft* is a realistic company and its *Forefront* product is one of many solutions available to protect our inboxes. The product runs on a *Windows* server, where it is an enhancement of *Microsoft Exchange*; the version we tested ran on *Windows Server 2008*. In a normal situation, email that is thought to be ham is sent to the user's inbox and email that is believed to be spam is discarded, with an in-between category stored in quarantine. As is the case with many products, the thresholds for emails ending up in each category can be adjusted by system administrators. In our test, both the discarded mail and the quarantined mail were considered to have been marked as spam.

With a spam catch rate of over 99%, *Forefront* is among the best spam catchers in this test. However, its false positive rate was slightly higher than average – the product misclassified several emails discussing spam as well as a number of newsletters – which means that it debuts in our tests with a VBSpam Silver award.

SPAMfighter Mail Gateway

SC rate (total): 98.03%
SC rate (Project Honey Pot corpus): 98.40%
SC rate (VB spam corpus): 92.46%
FP rate: 3.16%
FP rate of total VB mail corpus: 0.189%

SPAMfighter's free product protects the inboxes of many a home-user, but the Danish company also has a server product for businesses. It can run together with *Microsoft Exchange* or *Lotus Domino*, but the version we tested runs as a stand-alone MTA on *Windows Server 2003*. Installation is smooth and the product can be set up easily through a simple web-interface. I was charmed by a graph that showed how many spam emails had been caught and, at $0.04 per email, how much money was thus being saved: while these numbers are nothing but a rough estimate, they show how essential a spam filter is in a business setting.

Unfortunately, the product blocked what it thought was spam a little too eagerly and various legitimate emails were wrongly classified. In particular, newsletters and press releases were blocked and the number of false positives was about twice as high as that of the average product. Hence, despite a decent spam catch rate, the product failed to win an award on its first entry in the test.

Vircom modusGate

SC rate (total): 97.36%
SC rate (Project Honey Pot corpus): 97.68%
SC rate (VB spam corpus): 92.48%
FP rate: 4.42%
FP rate of total VB mail corpus: 0.261%

Vircom's *ModusGate*, a product that runs on *Windows Server 2003*, failed to win an award in the previous test and careful investigation by the developers determined that some scripts had not been working as well as they should have been. Fixing these scripts did indeed make a difference, and the product's spam catch rate on this occasion was comparable to that of most other products. At the same time, the product's false positive rate halved, but still this was not sufficiently low for it to win an award. For this, the product would have had to block fewer newsletters as well as emails that discuss spam and/or malware.

Webroot E-Mail Security SaaS

SC rate (total): 99.56%
SC rate (Project Honey Pot corpus): 99.81%
SC rate (VB spam corpus): 95.75%
FP rate: 1.84%
FP rate of total VB mail corpus: 0.111%

					Total spam			Project Honey Pot spam			VB corpus		
	True negative	FP	FP rate	FP/total VB corpus	False negative	True positive	SC rate	False negative	True positive	SC rate	False negative	True positive	SC rate
BitDefender Security	1264	11	0.87%	0.053%	3959	309780	98.74%	2099	292239	99.29%	1860	17541	90.41%
ClamAV	1270	5	0.39%	0.024%	45808	267931	85.40%	39954	254384	86.43%	5854	13547	69.83%
Fortinet FortiMail	1247	28	2.25%	0.135%	3014	310725	99.04%	2363	291975	99.20%	651	18750	96.64%
Kaspersky Anti-Spam	1267	8	0.63%	0.039%	5056	308683	98.39%	2907	291431	99.01%	2149	17252	88.92%
McAfee Email Gateway	1259	16	1.27%	0.077%	1244	312495	99.60%	394	293944	99.87%	850	18551	95.62%
McAfee Email & Web Security Appliance	1272	3	0.24%	0.015%	1927	311812	99.39%	1097	293241	99.63%	830	18571	95.72%
MessageStream	1265	10	0.78%	0.048%	4234	309505	98.65%	2922	291416	99.01%	1312	18089	93.24%
Messaging Architects M+Guardian	1261	14	1.11%	0.068%	3822	309917	98.78%	2350	291988	99.20%	1472	17929	92.41%
Microsoft Forefront	1250	25	2.00%	0.121%	1534	312205	99.51%	667	293671	99.77%	867	18534	95.53%
SPAMfighter Mail Gateway	1236	39	3.16%	0.189%	6184	307555	98.03%	4722	289616	98.40%	1462	17939	92.46%
Vircom modusGate	1221	54	4.42%	0.261%	8273	305466	97.36%	6814	287524	97.68%	1459	17942	92.48%
Webroot E-mail Security	1252	23	1.84%	0.111%	1377	312362	99.56%	553	293785	99.81%	824	18577	95.75%

Webroot's hosted solution won a VBSpam Silver award in the previous test, but the developers made it clear they wanted to do better this time.

The results indeed show a significant improvement in performance, both on blocking spam and on letting through ham, and its false positive rate would have been even higher had it not blocked several legitimate emails from the same sender discussing malware. Unfortunately for *Webroot*, other products did better this month too, resulting in stricter benchmarks for this test, and as a result *Webroot* wins another VBSpam Silver award.

AWARDS

As in the previous test, the levels of the awards earned by products are defined as follows:

- VBSpam Platinum for products with a total spam catch rate twice as high and a false positive rate twice as low as the average in the test.
- VBSpam Gold for products with a total spam catch rate at least as high and a false positive rate at least as low as the average in the test.
- VBSpam Silver for products whose total spam catch rate and false positive rates are no more than 50% worse than the average in the test.

To avoid the averages being skewed by one or more malperforming products, the scores for any product with a false positive rate of more than 10% and/or a spam catch rate of less than 70% are removed from the computation of the averages; this did not apply to any of the products this month.

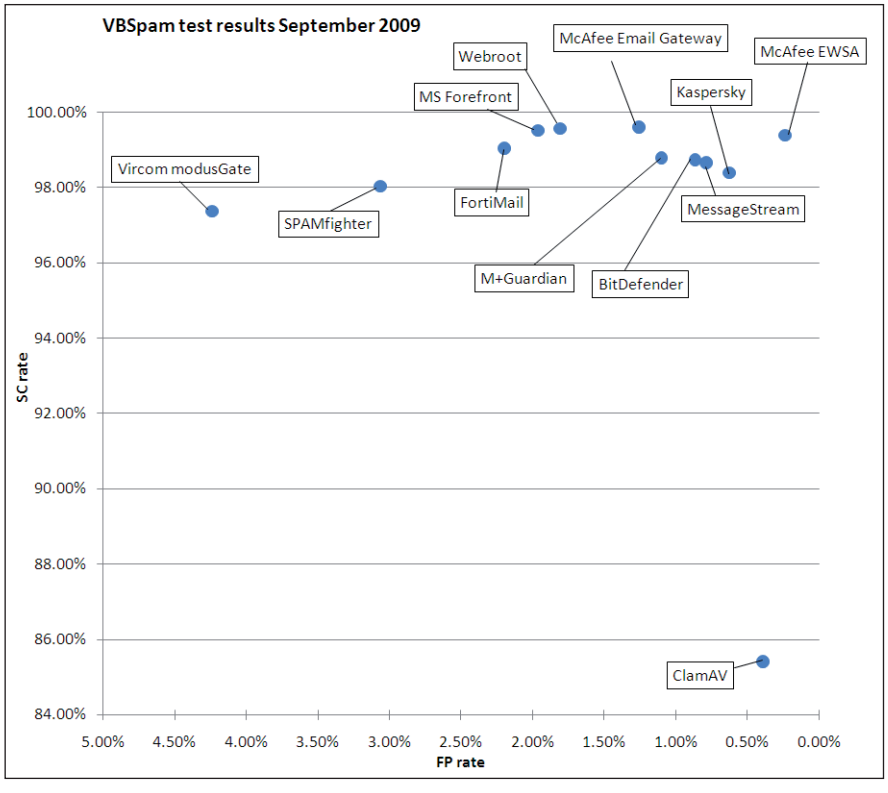
This month's benchmarks are then as follows:

- Platinum: SC 98.85%; FP 0.79%
- Gold: SC 97.70%; FP 1.58%
- Silver: SC 96.56%; FP 2.37%

The table on the previous page shows the scores for all of the products on test. The highlighted columns show the scores used for the benchmark calculations.

CONCLUSION

With three full anti-spam tests having been completed, a clearer picture is starting to emerge as to which products are the better performers. But, as seen in this test, those that do well cannot rest on their laurels and must work as hard as the others on keeping their products up to date.



We are already working on the next anti-spam test, which we hope will see even more products on the test bench. Just as spam filters need to be constantly updated to fight the latest threats, good anti-spam testers should always look for ways in which their tests can be improved. We welcome comments and suggestions and I hope to open the discussion on what a good anti-spam test entails in a presentation on the subject at VB2009 in Geneva, later this month (23–25 September 2009, see http://www.virusbtn.com/conference/vb2009/ for details).

The next anti-spam comparative review will run in October, with the results published in the November 2009 issue of *Virus Bulletin*. The deadline for product submission is 28 September 2009. Any developers interested in submitting a product are asked to email martijn.grooten@virusbtn.com.