

SpamCheetah - revolution in spam control

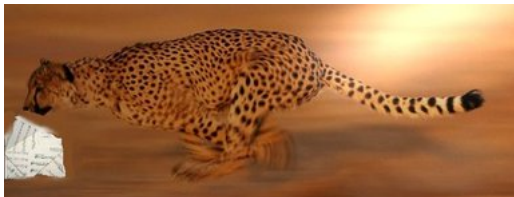
OpenBSD greylisting based spam control

Girish Venkatachalam

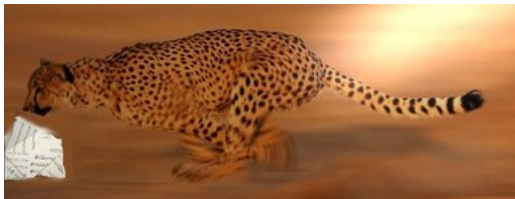
Chennai

girish1729@gmail.com

November 3, 2009



What is SpamCheetah?



- Not greylisting alone, very different from traditional meaning of greylisting

What is SpamCheetah?



- Not greylisting alone, very different from traditional meaning of greylisting
- **The tarpit does not even allow spam to enter your network, protect your network resources from dirty spammers!**

What is SpamCheetah?



- Not greylisting alone, very different from traditional meaning of greylisting
- The tarpit does not even allow spam to enter your network, protect your network resources from dirty spammers!
- **Revolution in spam control**

What is SpamAssassin?

- **Several products built with or around Spamassassin**

What is SpamAssassin?

- Several products built with or around Spamassassin
- **There is another rubbish called dspam that is better since it does some statistical analysis**

What is SpamAssassin?

- Several products built with or around Spamassassin
- There is another rubbish called dspam that is better since it does some statistical analysis
- **Several other alternatives exist for spam control**

What is SpamAssassin?

- Several products built with or around Spamassassin
- There is another rubbish called dspam that is better since it does some statistical analysis
- Several other alternatives exist for spam control
- **But content filtering does not help you with spam control**

What is SpamAssassin?

- Several products built with or around Spamassassin
- There is another rubbish called dspam that is better since it does some statistical analysis
- Several other alternatives exist for spam control
- But content filtering does not help you with spam control
- **But content filtering does not help you with spam**

What is SpamAssassin?

- Several products built with or around Spamassassin
- There is another rubbish called dspam that is better since it does some statistical analysis
- Several other alternatives exist for spam control
- But content filtering does not help you with spam control
- But content filtering does not help you with spam
- **Most expensive spam control appliances(big brand names) are utter rubbish since they result in false positives**

What is SpamAssassin?

- Several products built with or around Spamassassin
- There is another rubbish called dspam that is better since it does some statistical analysis
- Several other alternatives exist for spam control
- But content filtering does not help you with spam control
- But content filtering does not help you with spam
- Most expensive spam control appliances(big brand names) are utter rubbish since they result in false positives
- **In other words content scanning is the wrong way to attack the spam problem**

Why is spam control so important?

- E-mail the most important application on the Internet

Why is spam control so important?

- E-mail the most important application on the Internet
- **Every mail server needs a spam filtering engine**

Why is spam control so important?

- E-mail the most important application on the Internet
- Every mail server needs a spam filtering engine
- **But I don't like the term "spam filter". I call SpamCheetah as a spam blocker or spam control application**

Why is spam control so important?

- E-mail the most important application on the Internet
- Every mail server needs a spam filtering engine
- But I don't like the term "spam filter". I call SpamCheetah as a spam blocker or spam control application
- **Massive worldwide market with huge potential**

Why is spam control so important?

- E-mail the most important application on the Internet
- Every mail server needs a spam filtering engine
- But I don't like the term "spam filter". I call SpamCheetah as a spam blocker or spam control application
- Massive worldwide market with huge potential
- **This talk speaks about the technical aspects of SpamCheetah and how it does its job so well**

Status of SpamCheetah

- Close to 5 downloads a day

Status of SpamCheetah

- Close to 5 downloads a day
- **Selling as a hardware appliance in India**

Status of SpamCheetah

- Close to 5 downloads a day
- Selling as a hardware appliance in India
- **Every mail server needs a spam filtering engine**

Status of SpamCheetah

- Close to 5 downloads a day
- Selling as a hardware appliance in India
- Every mail server needs a spam filtering engine

Status of SpamCheetah

- Close to 5 downloads a day
- Selling as a hardware appliance in India
- Every mail server needs a spam filtering engine

Everyone with a mail server needs spam protection!

Business model of SpamCheetah



- I did the online version on August 6th and rechristened Anjal as SpamCheetah

Business model of SpamCheetah



- I did the online version on August 6th and rechristened Anjal as SpamCheetah
- **Three companies based on online downloads model**
- Zohocorp Inc (~1000 employees), Vembu (~100 employees) and Jijitechnologies Kovilpatti (~10 employees)

Business model of SpamCheetah



- I did the online version on August 6th and rechristened Anjal as SpamCheetah
- Three companies based on online downloads model - Zohocorp Inc (~1000 employees), Vembu (~100 employees) and Jijitechnologies Kovilpatti (~10 employees)
- **Online business model requires competitive pricing and unique product expertise, Google Adwords and some patience**

SpamCheetah implementation details

SpamCheetah's internal implementation details

Uses the `dbopen(3)` using the binary tree abstraction. Uses `getpcap(3)` interface for accessing the fields in a line.

A typical line in `/var/db/spamd` database looks like this:

```
WHITE|208.82.16.192||1241069409|1241072910...
```

```
GREY|87.182.96.240|p57b660f0.dip.t-dialin.net| \  
<tennis5@pctcu.com>|<spamd@your_host.org> \  
|1200326584|1200337384|1200337384|1|0
```

Spamd supports synchronization messages between multiple `spamd(8)` daemons on the network. A simple message format is used for sending updates after applying SHA1 checksum.

`time_t` datatype and `time(2)` system call are used for initial stuttering. `setsockopt(2)` is used for setting low TCP window size.

SpamCheetah implementation details II

SpamCheetah's internal implementation details

This is only the printed format. Internally it is not stored like this. It is completely different. printf does this job when you run \$ spamdb command.

libpcap(3) is used for filtering and reading from the /dev/pflog0 virtual network interface in a loop. Read pcap(3) for details.

pf(4) <spamd-white> is the table using which IP addresses are whitelisted.

Without pf(4) redirection to spamd port TCP 8025 and to mail server port 25, all this magic will fall apart!

UNIX daemons in SpamCheetah

SpamCheetah - the daemon workhorses

[spamd(8)]

Update /var/db/spamd database

Track blacklisted IPs

Implement tarpit by
delaying response(stuttering)

[spamd-setup(8)]

Keep track of blacklisted IP netblocks
that send out spam right now.

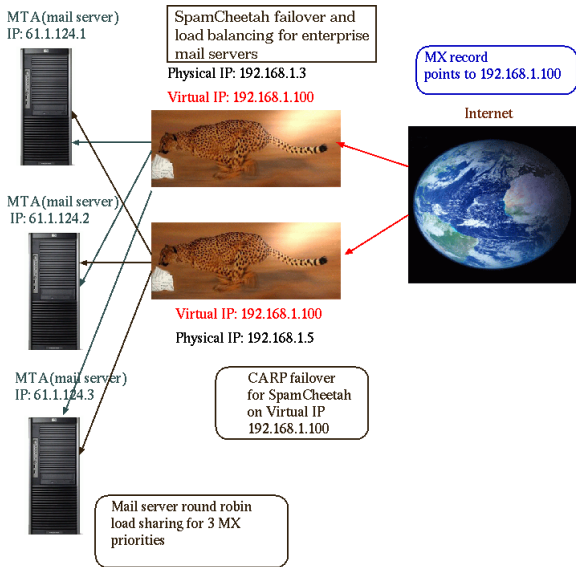
Updated every hour. Talks to spamd(8)
using the spam-cfg local UDP socket
interface. It is a simple line by line text
protocol

[spamlogd(8)]

Keep looping over pcap_loop()
on interface /dev/pflog0 log
interface for mail attempts
for whitelisting

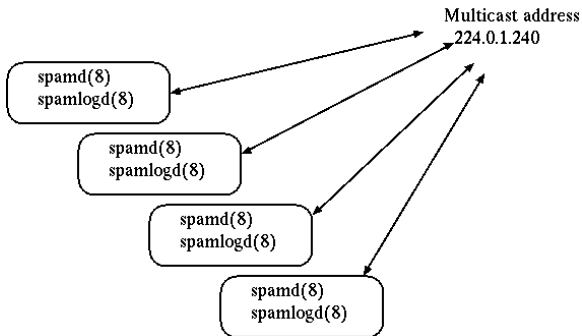


SpamCheetah load balancing and failover



SpamCheetah greylist sync

Multicast syncing in SpamCheetah
of the greylist database



Discussion and questions?

- 1 SpamCheetah's website
- 2 Gayatri Hitech - the company behind SpamCheetah
- 3 OpenBSD papers