

Slide 1

Mac OS X Snow Leopard

MacSysAdmin 2009  
- European Macintosh System Administrators Meeting 2009 -

SpamAssassin

SpamAssassin

Way more than the Mac OS X Server GUI shows

Presented by: Kevin A. McGrail  
Project Management Committee Member  
of the Apache Software Foundation  
SpamAssassin Project & President, PCCC

September 16, 2009

© TemplatesWow.com

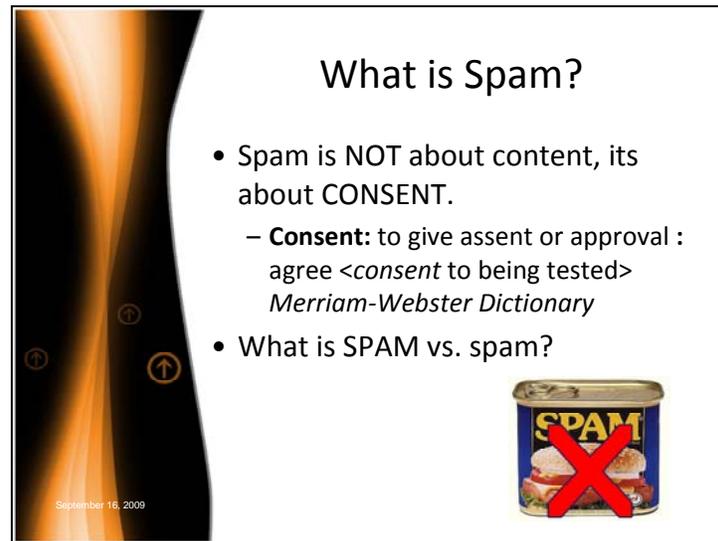
Good Afternoon,

My name is Kevin A. McGrail. If you read my biography<sup>1</sup> for this conference, you'll know already that I hate Spam and enjoy greatly fighting spammers.

You'll also know that I love all types of computers and use a wide variety of machines & operating systems. But I'm definitely old-school in my love for the command line interface. This doesn't mean I don't think that Apple's OS X is the most beautiful pairing of a rock-solid CLI with a beautifully polished GUI.

But it does mean that while we are here to talk about Mac system administration, the configuration of SpamAssassin is largely not server specific and most of the heavy-handed configuration changes will be done behind the scenes using the CLI.

So let's get started by talking about the definition of Spam.



**What is Spam?**

- Spam is NOT about content, its about CONSENT.
  - **Consent:** to give assent or approval : agree <consent to being tested> *Merriam-Webster Dictionary*
- What is SPAM vs. spam?



September 16, 2009

Chris Santerre gave the best definition of Spam I've ever seen. He based the definition of Spam on CONSENT not content because consent is when you give approval to someone to send you e-mails.

Many people try and use various legal definitions such as CAN-SPAM in the US. This definition clearly has problems for the global community of the internet.

However, I find the definition of CONSENT to be best. It leaves the content of any e-mail in the eye of the beholder. In short, if you consent to receive e-mails about XYZ, then those e-mails can't by definition be Spam.

OK, so let's touch on what isn't Spam.

SPAM, all caps is a canned meat product and a trademark of the Hormel Corporation. So SPAM is not the same as Spam.

So SPAM, all capitals, is not junk e-mail but instead a tasty, pork product that is yummy when fried in a sandwich. It's also quite tasty with Eggs, SPAM, Bacon, SPAM, SPAM & SPAM.

Which leads me to why junk e-mail is called Spam. The name refers to a skit by Monty Python's Flying Circus which involves what I can only describe as an overload of unwanted items offered at a café all including SPAM. A patron who does like SPAM certainly won't like "SPAM SPAM SPAM SPAM SPAM SPAM SPAM baked beans SPAM SPAM SPAM and SPAM". Hence, an overload of unwanted items became known as spam (but not SPAM).

At the end of my presentation, I'll have a list of resources including a link to the skit on YouTube. This presentation and my notes will also be made available later on the web.

## Slide 3

### Disagreements on Spam

- Hormel – SPAM Brand and the Internet
- CAN-SPAM!
- URIBLSucks



September 16, 2009

The image is a slide titled "Disagreements on Spam". It features a list of three items: "Hormel – SPAM Brand and the Internet", "CAN-SPAM!", and "URIBLSucks". To the right of the list is a cartoon illustration. The cartoon is titled "Uribl Cabalist..." and depicts a character with a white, pointed hat and a white body, standing behind a lemonade stand. The stand has a sign that says "URIBL LEMONADE". On the stand, there is a pitcher of yellow liquid, a stack of lemons, and a stack of papers. The background is a simple landscape with a blue sky and green grass. In the bottom left corner of the slide, there is a date: "September 16, 2009".

Disagreements on Spam and definitions regarding Spam are quite common.

Hormel, naturally, defines SPAM as their spiced ham product. And the use of their trademark to refer to junk e-mail has necessitated their lawyers to come up with an entire position on the matter. The link for this is in the resources.

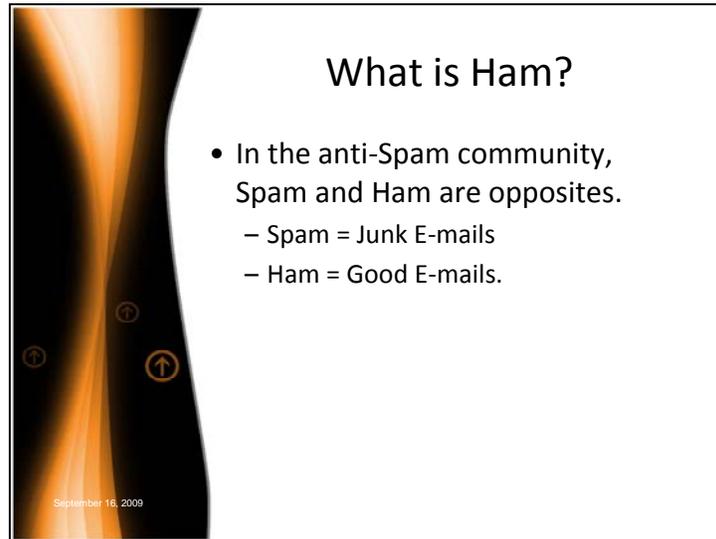
And I can tell you that Spammers, naturally, often disagree with anti-Spammers. The argument that “my e-mail can’t be Spam because it’s CAN-SPAM compliant!” is something I read several times a day. Spammers routinely fail to understand that the Internet is much larger than one country’s laws.

I can even tell you that I’ve been branded an “unAmerican, anti-capitalistic, cabal thug” for my help with URIBL stopping spammers. As you can see in the picture, they even put up a website and likened us to the Klu Klux Klan. Words can’t convey how offended I was by this but I did feel better because, as another anti-Spam peer wrote, he wished they had called him a cabal thug because he always liked the word cabal. And one of my office mates gave the guy props for standing up for his side of the argument and

wondered why I was so upset about being a “friendly ghost” apparently selling lemonade for Halloween.

Anyway, if you would like to debate these definitions, you won't be the first or the last. So let's move on to a few more definitions that won't be as controversial.

## Slide 4



**What is Ham?**

- In the anti-Spam community, Spam and Ham are opposites.
  - Spam = Junk E-mails
  - Ham = Good E-mails.

September 16, 2009

In the anti-Spam community, Spam and Ham are opposites. Spam refers to Junk E-mails and Ham refers to Good E-mails. And, while I promised to move on to non-controversial definitions, I have to admit there is even controversy about these definitions.

Since Ham is “Treif”, or non-kosher for those who follow Jewish dietary laws, some people have proposed using the word Yam, instead of Ham.

However, I’m firmly anti-vegetable and won’t stand for such nonsense. So let’s move on to truly non-controversial definitions.



## What is an FP?

- False-Positives (FPs)
  - E-mails incorrectly tagged as Junk
- False-Negatives (FNs)
  - E-mails incorrectly NOT tagged as Junk

September 16, 2009

Slide 6



The slide features a decorative background on the left side with a vertical orange-to-black gradient and three circular icons: a downward arrow, a downward arrow, and an upward arrow. The main content is on the right side.

## Blacklist / Whitelist

- Blacklist = Bad Items
- Whitelist = Good Items
- Greylist = Undecided Items

September 16, 2009

If you can't remember which is which, think of the old Westerns where the bad guys wore the black hats.

We'll talk about Greylists and Greylisting further along in the presentation.

## What is SpamAssassin?

- SpamAssassin is a mail filter & API used to identify junk e-mail.
- SpamAssassin powers the Junk Mail Filtering tool in OS X.
- SpamAssassin is also the basis of several other commercial products.



September 16, 2009

The slide features a title 'What is SpamAssassin?' at the top. Below it is a bulleted list of three points. At the bottom center is the SpamAssassin logo, which consists of several overlapping, colorful arrows pointing to the right, with the text 'SpamAssassin' underneath. In the bottom left corner, the date 'September 16, 2009' is printed.

SpamAssassin is a mail filter and programming interface that identifies junk e-mail. And last week, eWeek chose SpamAssassin as one of the Apache Technologies that have changed computing in the past 10 years.

One of the reasons SA deserves this honor is because of the flexibility built-in to the initial design. Not everyone will want to use SpamAssassin to do the actual e-mail processing. So the design was that you don't have to use SpamAssassin to filter your mail.

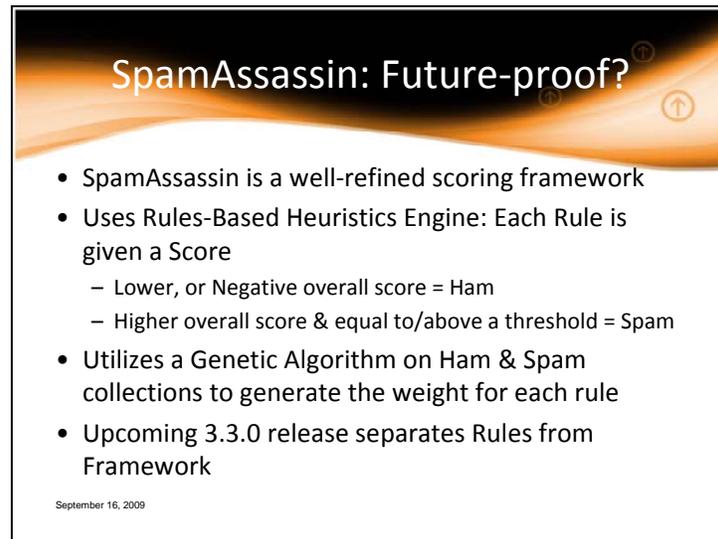
Instead, you can also use SpamAssassin from other programs. In this way, SpamAssassin can be used to return a "this is spam/this is ham", a score, a list of rules that lead to the score, and a detailed report. It's then the job of the program that calls SpamAssassin to decide what to do with this information.

So now you know a little bit about SpamAssassin and some definitions. So if you've never heard of SpamAssassin before, you might be thinking why do I care? And why am I listening to this speech.

Well, SpamAssassin is the all powerful wizard behind the curtains powering the Junk Mail filter in OS X.

And more specifically, the Mac OS X implementation of SpamAssassin is using AMaViS (A Mail Virus Scanner) to call SpamAssassin.

Some of the commercial products that have used SpamAssassin are McAfee SpamKiller, Symantec MailScanner, & Kerio.



The slide features a title 'SpamAssassin: Future-proof?' in white text on a dark background with a wavy orange and black gradient. Below the title is a bulleted list of four points. The second point has two sub-points. At the bottom left of the slide, the date 'September 16, 2009' is written in small text.

- SpamAssassin is a well-refined scoring framework
- Uses Rules-Based Heuristics Engine: Each Rule is given a Score
  - Lower, or Negative overall score = Ham
  - Higher overall score & equal to/above a threshold = Spam
- Utilizes a Genetic Algorithm on Ham & Spam collections to generate the weight for each rule
- Upcoming 3.3.0 release separates Rules from Framework

September 16, 2009

SpamAssassin at its heart is a Scoring Framework. This framework allows virtually any anti-Spam technologies to be added and used.

The engine for SpamAssassin is sometimes called a “Rules-Based Heuristics Engine” because the filter typically looks for patterns such as common phrases or known senders. Heuristics is the application of experience-based techniques for problem solving. Virus scanners use the same technique and are also heuristics engines.

As an e-mail goes through the engine, each of the rules for SpamAssassin are run and generate a score. These individual scores are then totaled to provide an overall score for the e-mail. Some rules are positive and add to the score. Some rules are negative take away from the score.

The lower the overall score is, the more likely the e-mail is Ham. The higher the score, the more likely the e-mail is Spam.

But not each rule has the same weight. The weight each rule is given is determined first by initial “guesses” and later refined through optimization. We use a Genetic Algorithm

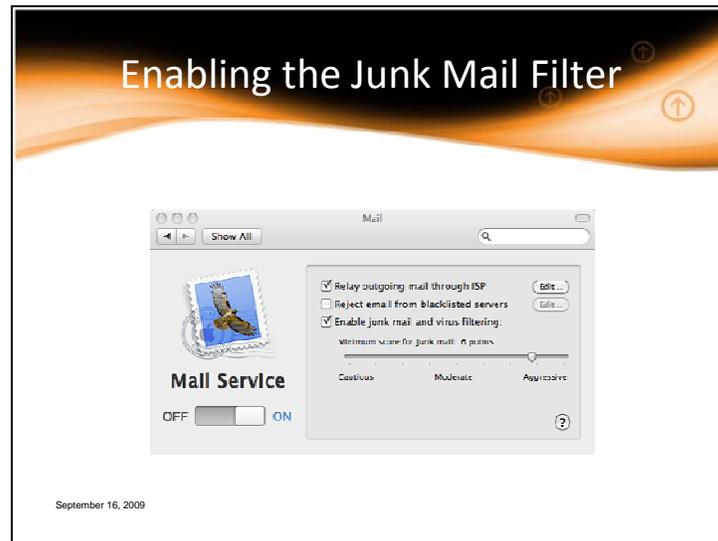
to optimize the scores. Discussing Genetic Algorithms is not going to be in the scope of this conference especially since Wikipedia defines it as “a search technique used in computing to find exact or approximate solutions to optimization and search problems. Genetic algorithms are categorized as global search heuristics. Genetic algorithms are a particular class of evolutionary algorithms (EA) that use techniques inspired by evolutionary biology such as inheritance, mutation, selection, and crossover.” The key point is that we use a scientific process to obtain the scores and you can read more about Genetic Algorithms on Wikipedia.

So one of the reasons I have devoted so much time to working with SpamAssassin is it is future-proof. If anyone identifies an algorithm/program/magic wand to detect spam (or ham), it can be quickly integrated into the framework as another test with an appropriately weighted score.

And because of just this design, the upcoming 3.3.0 release of SpamAssassin will not include any rules with the source distribution. Instead, the administrator will need to run the command ‘sa-update’ to install the latest rules and complete the installation. We anticipate that a future revision of the Mac OS X installation will seamlessly handle this for users, of course.

Removing the rules from the framework will allow the SA team to develop and adapt the rules more quickly to respond to the ever-changing techniques that spammers use to avoid detection without causing the rules to be delayed waiting for a development cycle.

Slide 9

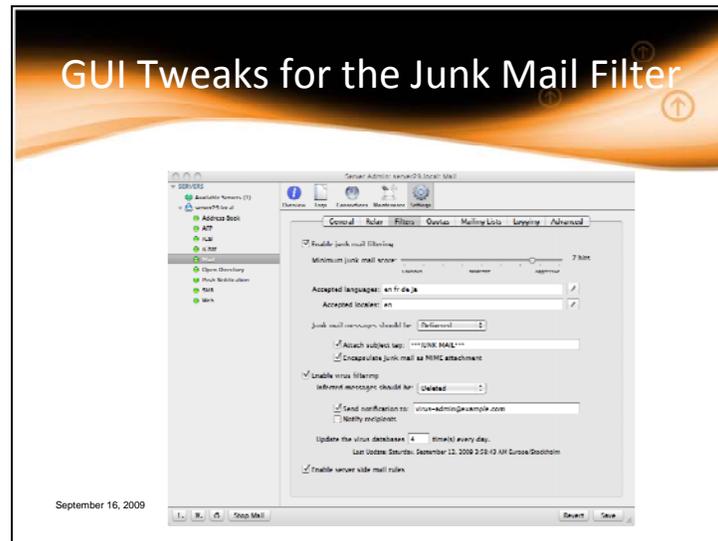


One of the reasons we are here today is because the Apple OS X GUI provide very little interface to the real inner workings of SpamAssassin. As you can see above, the Server Preferences in Snow Leopard only allow for two options related to Junk Mail.

Enabling the junk mail and virus filter

&

Setting the threshold that is used to mark an e-mail as Spam.



Under Server Admin, choosing the Mail service Settings page still shows only a few tweaks.

Here you can enable or disable the junk mail and virus filters individually, change or disable the subject tag that is used, or “Encapsulate junk mail as MIME attachments.”

This last cryptic option will take the spam and move it to an attachment on the original mail. This has two big benefits: First, users can view the spam unmodified by opening the attachment. Second, the users don't have to see whatever is contained in the e-mail without explicitly opting to do so.

While OS X uses **\*\*\*JUNK MAIL\*\*\***, the default SpamAssassin subject tag for Spam is **\*\*\*\*\*SPAM\*\*\*\*\***.

As for the Accepted languages and locales option, besides the fact that it's missing se, the configuration won't even work! Snow Leopard is using configuration options that are invalid for the version of SA that is included! We'll talk about fixing that later.



The slide features a vertical orange and black gradient on the left side with three small circular icons containing arrows. The main content is a white box with a black border containing the title and a bulleted list. The date 'September 16, 2009' is printed in small text at the bottom left of the slide area.

### SpamAssassin with Snow Leopard

- Snow Leopard (10.6.1) uses SA version 3.2.1
- 3.2.1 was released June 11<sup>th</sup>, 2007!
- 3.2.5 was released June 12<sup>th</sup>, 2008!
- 3.3.0 is beyond 2<sup>nd</sup> alpha with optimization for the score weighting in progress for imminent full release.

September 16, 2009

First, thanks to Tycho Sjögren for his assistance in getting me access to a Snow Leopard test box. As we all know, Snow Leopard was released very recently. However, there are very few changes in this release regarding the Junk Mail filter.

And I am fairly disappointed how old the version of SpamAssassin in use even with Snow Leopard remains. It's a testament to how long-standing our releases can be but I'd prefer to see some movement forward on this. I've reached out to our PMC to open a dialog with Apple about this issue or to find out more reasoning behind their decision.

But let's move on to improving the Junk Mail Filter as installed in OS X.



The slide features a decorative background on the left side with a vertical orange and black gradient and three circular icons containing arrows. The main content is on the right, starting with the title 'Improve Your Mail Setup' and a bulleted list of recommendations. A date 'September 16, 2009' is visible in the bottom left corner of the slide frame.

## Improve Your Mail Setup

- Improving your overall mail setup will assist other mail servers to identify legitimate senders:
  - rPTRs
  - Smart Host?
    - MSA/MUA/MTA
  - SPF
  - Reduction of DSNs

September 16, 2009

The first step in identifying junk mail is to understand that your overall mail setup is “judged” by other mail servers to help identify Spammers.

On the internet, computers speak in numbers called in IP addresses. Humans use names via DNS to translate into the numbers. However, you can also translate a number into a name. This is called a reverse lookup. Of course, real SysAdmins don't use DNS. We memorize all of our IP addresses.

But despite our good memories, having a valid answer for a reverse lookup, called a reverse PTR or rPTR makes your mail server more legitimate. Some ISPs, like AOL, will not even accept e-mail from a server that doesn't have one. So if you don't have a reverse pointer for your IP address, talk to your ISP and get one that is non-generic and doesn't reference your IP address. For example, mail.macsysadmin.se instead of static-71-163-15-129.washdc.fios.verizon.net.

If you don't have a static IP, you really **MUST** be using a Smart Host. This is the feature in the mail preferences to Relay E-mail through your ISP. If you send your mail directly, you will definitely look like a spammer!

Mark Martinec, another member of the SA Project Management Committee, also points out that you should really, really use the Smart Host or MSA of a domain used in your From address.

The acronym MSA stands for the Mail Submission Agent. It's the middle-man between the Mail User Agent or MUA -- That's SysAdmin speak for an e-mail client such as Mail or Outlook -- and the Mail Transport Agent or MTA. The MTA is the agent that sends mail from one server to the next.

So, via your user's Mail User Agent, they should be submitting the e-mail to your domain's Mail Submission Agent, preferably on a standard submission port 587 and authenticated through AUTH or POP-Before-SMTP or similar. In the same vein, when using @gmail.com in the From address, the e-mail should be submitted through that domains MSA, smtp.gmail.com:587. The Gmail web interface, technically just another MUA, would automatically be setup to use a proper submission method.

Submitting mail through anything but the MSA for the From address will likely contribute points towards the threshold to tag an e-mail as Spam.

The reason away from using an ISP's Smart Host to an MSA for the domain is due to the widespread e-mail address forgery in Junk Mail. This fraud has necessitated increasingly stricter methods to distinguish the valid addresses from the forged address.

One of these stricter methods is SPF. We'll touch on other methods like DKIM and Whitelist\_From\_Rcvd in a few slides.

Using these methods, like SPF or Sender Policy Framework by adding an SPF Record for your domain will help control forged e-mail. SPF allows other servers that receive your e-mails to check your domain's DNS. There you can set policies that tell the MTA

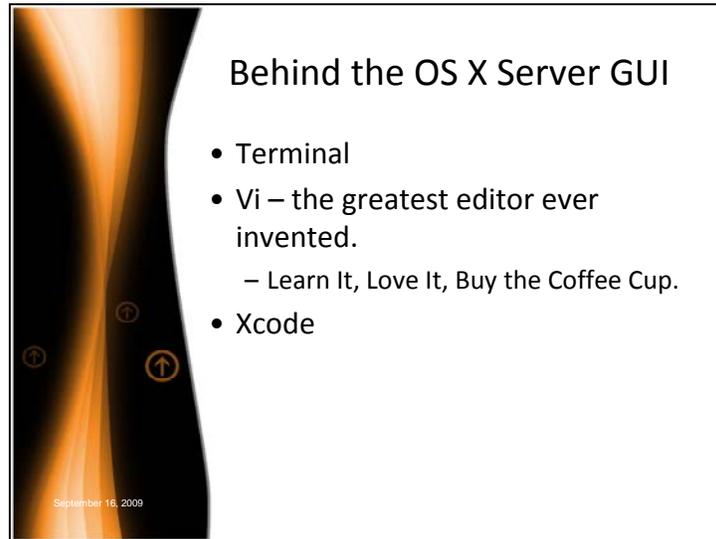
on the receiving end what MTAs are allowed to transmit e-mail for your domain. The website [www.openspf.org](http://www.openspf.org) can assist you in configuring this record.

Finally, while some administrators might think they are being neighborly and helping out a fellow netizen, most Spam and Viri is delivered using forged from addresses. These forged addresses sometimes receive a lot of Delivery Status Notifications or DSNs. This is called backscatter.

So, please, don't use the OS X feature for the junk mail and virus filter that bounces the e-mail instead of discarding it as this causes backscatter! This can also get you blacklisted.

Another similar technique that causes DSNs is called Challenge and Response. This is a system where a new sender is sent an email by an automated system asking them to verify that they sent the original email. Unfortunately, because of forgeries, C+R systems are often sometimes considered spam. Use it at your own risk!

Finally, during the conference, Matt Wynne reported that Snow Leopard introduces Vacation Messages. If these vacation messages are also sent to the e-mails tagged as Junk Mail, you are simply going to respond to spammers. At this time, I do not know if the feature ignores tagged Junk Mail or responds to it.



Diving behind the GUI is a necessary evil for any administrator. Here are 3 musts for doing so to enhance your Junk Mail filter on OS X.

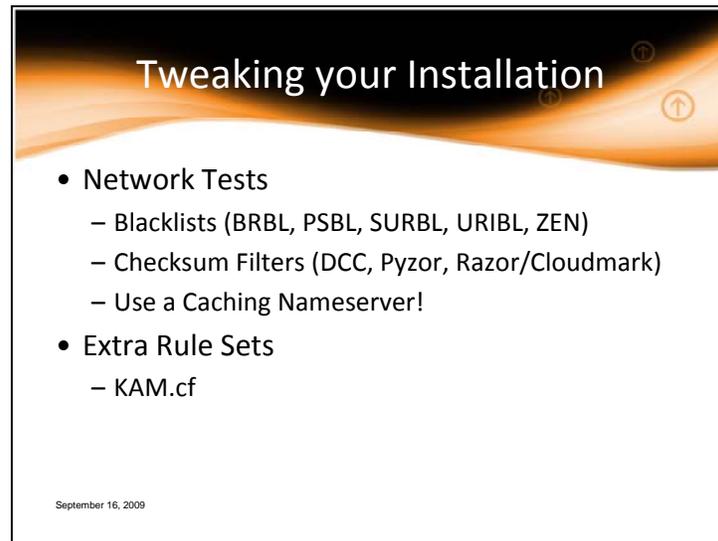
First, I won't say much about terminal but I hope you are familiar with a CLI as a system administrator.

So I'll move on to Vi. The single greatest editor ever invented. I won't say it's easy but I can tell you that every administrator (or programmer) I have ever trained to use Vi, still uses it as their primary editor. Vi, or more correctly, Vi IMproved (VIM), is an archaic but very powerful and very fast text editor. It will exist on virtually ever Unix-based box you will ever work on.

To become familiar with it, run the command 'vimtutor' and follow the step by step tutorial.

Finally, as administrators, you should have Apple's Development Environment, Xcode, on your boxes. It's free and available from [developer.apple.com](http://developer.apple.com). Xcode adds things such as the GNU compiler collection (gcc) and other tools which are invaluable in

working with open-source software such as SpamAssassin which you may want to compile yourself from time to time.



**Tweaking your Installation**

- Network Tests
  - Blacklists (BRBL, PSBL, SURBL, URIBL, ZEN)
  - Checksum Filters (DCC, Pyzor, Razor/Cloudmark)
  - Use a Caching Nameserver!
- Extra Rule Sets
  - KAM.cf

September 16, 2009

The network tests for SpamAssassin can be one of the most significant ways to improve your Junk Mail filter. Unfortunately, each one of these Blacklists has different rules for usage: Some are available for commercial usage without charge. Some are available only for home users. Some are limited to a specific number of queries. Unfortunately, the rules and limits vary so please research the Blacklists before enabling them but this is a list I recommend you consider (in alphabetical order):

Barracuda Reputation Block List (BRBL) - <http://www.barracudacentral.org/rbl>

Passive Spam Block List (PSBL) - <http://psbl.surriel.com/> (Public Mirror)

SURBL – <http://www.surbl.org/> (Public Mirror)

URIBL- <http://www.uribl.com/> (Public Mirror)

ZEN – <http://www.spamhaus.org/zen/>

To enable the blacklists, they are typically enabled by adding a few configuration lines or a \*.cf file in the directory '/etc/mail/spamassassin'. Some are also available in the default SA rules but that depends on your version of SpamAssassin. Viewing the headers from a few e-mails filtered by your installation can be helpful in determining if you have a duplicate blacklist configured.

To use network tests with SpamAssassin on OS X, you have to make sure that network tests are enabled in AMaViS. Edit `/etc/amavisd.conf` and check that the line for local tests reads `'$sa_local_tests_only = 0;'`

In OS X back to 10.4, it was the default to run local tests only, however Snow Leopard appears to have changed this default.

SpamAssassin also supports several Checksum Filters. This is another network test that calculates a checksum on the message and compares it to a database of reported Spam. Checksum Filters have varying difficulty of installation and availability for older versions of SA. So I would consider these an advanced tweak and something to attack for more experienced SA administrators. More information can be found out by using the commands `'perldoc Mail::SpamAssassin::Plugin::DCC'`, `'perldoc Mail::SpamAssassin::Plugin::Pyzor'` and `'perldoc Mail::SpamAssassin::Plugin::Razor2'`. Cloudmark is a commercial derivative of the Razor checksum filter.

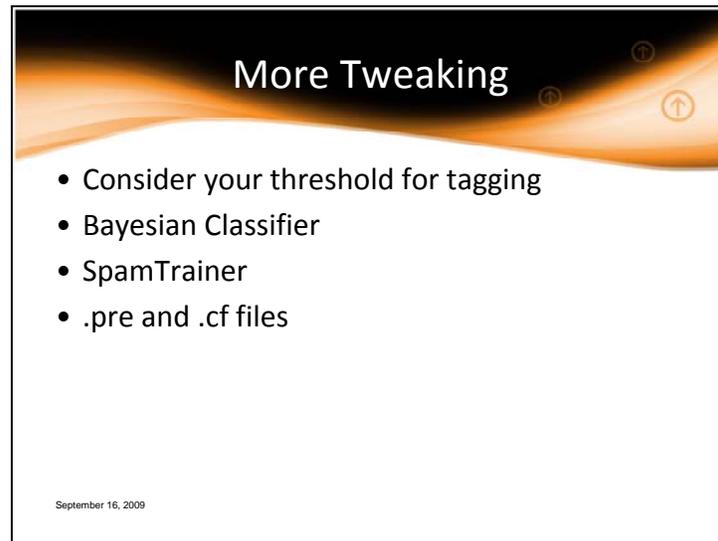
The network tests will run significantly faster with a local caching nameserver. This is easy to do in OS X. Simply turn on DNS in Server Admin and change your DNS resolution to use the DNS server on your localhost (127.0.0.1) as the first DNS entry in System Preferences -> Network.

You can also consider Extra Rule Sets. Now that sa-update provides for much faster rule updates, the need for extra rulesets has diminished greatly. For example, this has lead to the lack of updates for the Rules Emporium (SARE) rules.

And I currently run very few extra rule sets on my servers while just a few years ago I ran quite a few. I rely on sa-update to provide the latest rules coupled with my

development ruleset, KAM.cf. KAM.cf is available at  
<http://www.pccc.com/downloads/SpamAssassin/contrib/KAM.cf>

But you may find with a little research that someone else is maintaining rules that can help your Junk Mail filter.



SpamAssassin is a scoring framework that uses your threshold limit to tag an e-mail as Ham or Spam. Carefully consider the threshold you use. For example, a threshold of 5 is used as the threshold for the Genetic Algorithm for weighting the rules. And I generally recommend using a 6.0 to 7.0 score as the threshold for my users. But I still know users who have considered their options carefully and are using thresholds from 5.0 to 20! And yes, despite the indications of Junk Mail preferences in the OS X GUI, you can use threshold that are not integers such as 6.5.

The Bayesian Classifier in SpamAssassin learns tokens - words or short sequences - that are commonly found in spam or ham.

The SA command `sa-learn` is a way of teaching the Bayesian Classifier. Mail includes a “Junk Button”. This button is an interface to `sa-learn`.

Which leads me to the script in `/etc/mail/learn_junk_mail` that runs `sa-learn` on ham folder (called `notjunkmail`) and a spam folder (called `junkmail`).

This script runs every 24 hours (Note: the option for how often this was run seems to be removed in Snow Leopard. I've been unable to confirm when and how often it runs on Snow Leopard.)

To use this script, though, you first need to create the accounts junkmail and notjunkmail with e-mail addresses. They don't need home directories. Then redirect Spam that gets through untagged to the junkmail address and Ham that is improperly tagged to the notjunkmail address. The system will learn from those messages using the Bayesian Classifier. If you use IMAP, you should also use a shared IMAP folder to permit dragging ham and spam to the notjunkmail/junkmail mailboxes as well!

A replacement/addition for the learn\_junk\_mail script is SpamTrainer from <http://osx.topicdesk.com/spamtrainer>

Because the Bayes rules will not be used until there are at least 200 tokens. SpamTrainer can also help you manually jumpstart or intervene on a server by "learning" a few folders of Spam and Ham that you have classified by hand. Once you have 200 tokens, you should start seeing BAYES\_\* rules in your email headers and reports.

Finally, as I talked about before, SpamAssassin is a framework that allows for constantly integrating new technologies. Look at the \*.pre files and the \*.cf files in /etc/mail/spamassassin/

For example, the local.cf is your local configuration file for SA. It can have far more options than the GUI shows such as the configuration for Auto-Whitelisting where SA tracks the long-term average score for each sender and then shifts the score of new messages towards that long-term average. However, you may want to create a new file such as my.cf and rename local.cf. SpamAssassin will load all of the \*.cf files in a directory and you may not want the GUI interface to accidentally change something.

And making matters more confusing, not all of the options present will affect the way SpamAssassin works in OS X because AMaViS is calling SA. It's the job of the AMaViS daemon to determine what to do if an e-mail is identified by SA as Spam. And even more confusing, some of the comments in the amavisd.conf are quite old. Mark Martinec pointed out that the \$sa\_auto\_whitelist option has no effect on SA since version 3.0.0 - SA now uses the configuration file option 'use\_auto\_whitelist';

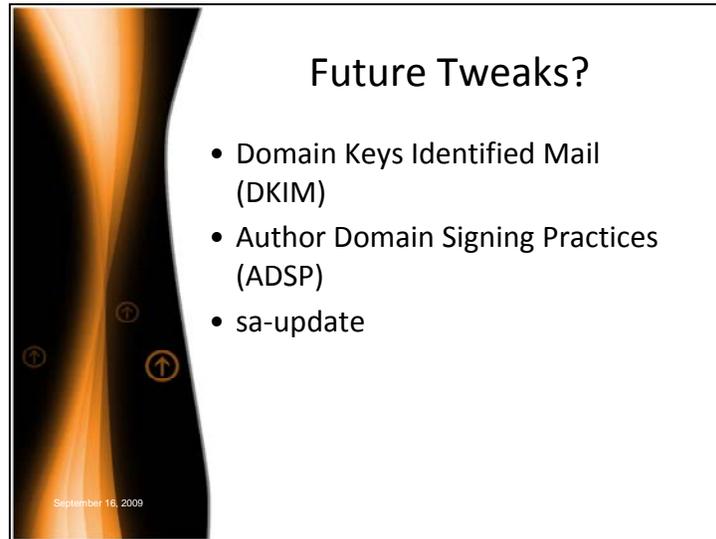
You can also statically blacklist and whitelist senders with configuration entries such as whitelist\_from kmcgrail@apache.org. Or "disable" SA for a single user on your system with the whitelist\_to bob@<mydomain>.com. AMaViS also has an interface in amavisd.conf for achieving this as well.

Though I should mention whitelist\_from is becoming more of a no-no these days because of the forged Junk Mail. New whitelist options that are based on some form of verification might be even more helpful such as whitelist\_auth, whitelist\_from\_spf, whitelist\_from\_dkim, and whitelist\_from\_rcvd.

You can find out more about the configuration file options for SA by using the command 'man Mail::SpamAssassin::Conf'

And the \*.pre files are where plugins are implemented. You'll note that there are plenty that aren't enabled in a default installation. For example, remember when I told you the accepted locale option in the GUI wasn't even going to be used? That feature is implemented by the TextCat plugin now. If you want to use TextCat, edit the file '/etc/mail/spamassassin/v310.pre' and remove the # that is commenting the line 'loadplugin Mail::SpamAssassin::Plugin::TextCat'!

One final note. To me, SpamAssassin is less about sender whitelist and blacklists (though it supports them). I always look at what can be tweaked BEFORE using a whitelist/blacklist for the sender. I have less than two dozen whitelist\_from & blacklist\_form entries on our firm's primary mail server!



DKIM allows a sender's server to sign an e-mail to prove it came from them.

I predict we'll see a massive shift in Spam in the future as the DKIM is embraced and the default DKIM Plug-in for SA 3.3.0 starts scoring e-mails from domains which sign all their mail with DKIM.

In the past, while places such as Google, eBay and PayPal were often the targets of Phishing, these organization would not say "we sign every e-mail with a DKIM signature". They have finally begun saying "if our e-mail is not DKIM signed, it's fake!".

As 3.3.0 is not released yet, I do not have stats to back up this prediction.

As a side note, AMaViS can be used to sign your e-mails with DKIM but that would be a fairly advanced hack. Let's push to implement this as a simple feature of the GUI interface in a future release!

ADSP is an optional extension to DKIM where a domain administrator can specify the DKIM signing practice.

And the current DKIM plugin using Mail::DKIM v0.34 or greater implements ADSP – RFC 5617 adopted as a standards track in August of this year.

Domains that publish their ADSP record in DNS will be honored and the SA 3.3.0 rules DKIM\_ADSP\_\* will hit on a valid or missing signature which will contribute to the proper classification of e-mails.

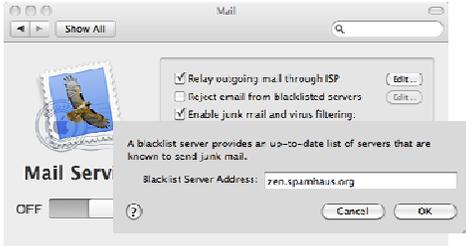
I believe we'll see a big push to implement ADSP and DKIM as these technologies start to reduce the forgeries littering our inboxes.

Finally, you may want to use sa-update to keep your rules up to date on the server. Unfortunately, sa-update can't be run because the OS doesn't ship with GNU privacy guard or GPG. GPG can be found at: <http://macgpg.sourceforge.net/> and there is a simple, step by step install instructions available at <http://macgpg.sourceforge.net/docs/howto-build-gpg-osx.txt.asc> assuming you have installed Xcode.

You have installed Xcode haven't you? You might also want to edit /etc/profile and put /usr/local/bin and /usr/local/sbin at the front of your path. Then logout and log back in. This will make new self-compiled programs run instead of the default installations because many open-source projects default to installing in /usr/local.

## More Spam Reduction Techniques?

- Using Blacklists to Outright Blocking E-mail
  - RBL100 - 0.0.0.0 to 255.255.255.255
- Greylisting



The screenshot shows the Mail Server preferences window. It includes a 'Mail Serv' section with a 'Blacklist Server Address' field containing 'zen.spamhaus.org'. There are checkboxes for 'Relay outgoing mail through ISP', 'Reject email from blacklisted servers', and 'Enable junk mail and virus filtering'. A 'Mail Serv' status indicator is set to 'OFF'. A date stamp 'September 16, 2009' is visible in the bottom left corner of the window.

I labeled this slide with a question mark because these are techniques you have to question before you implement.

One of my specialties in the anti-Spam community is promoting techniques that “do no harm”. In other words, don’t cause False Positives.

For example, the easiest way to block 100% of your Spam (guaranteed) is to blacklist the following IPs from sending you e-mail:

0.0.0.0 to 255.255.255.255.

This block has a 100% guarantee to block 100% of the Spam. Of course, this block has a 100% false positive rate, too. (Assuming you use IPv4)

For me, as a business-owner, 10,000 Spam e-mails are better than 1 critical e-mail being accidentally blocked! But for a system designed to keep Spam away from a child, the ratio might be 10,000 blocked e-mails is better than 1 Spam e-mail getting through.

Blacklists, such as the zen.spamhaus.org, list can be used to block servers from even connecting to send e-mails. In Mac OS X, the option under Server Preferences for Mail

shows this as “Reject e-mail from blacklisted servers”. This then prompts you for a single blacklist server.

I don't recommend this option is used unless your server is simply too overloaded. A better option is to let the e-mail through and use the Network Tests that can check blacklists, including zen.spamhaus.org, and give the e-mail a score weighted appropriately for each blacklist. Otherwise, using Blacklists will have False Positives!

Which leads to a question I received about why someone would NOT want to use the Junk Filter. The most common answer is that the Junk Filter can be fairly resource intensive leading to server resource allocation issues or even denial of service problems.

Greylisting, which we touched on earlier in the definitions, is essentially where you don't make a decision to put an IP address on the Blacklist or the Whitelist but rather temporarily Greylist the sender. The theory behind a greylist is this:

- Much of the Spam in the world is sent by software that isn't written very well.
- Much of the Spam in the world is sent by software that will not properly handle a “temporary error” SMTP Response Code of 4XX. This software is often called “Ratware”
- Therefore, give a temporary error to each and every new IP Address connecting to your server requiring them to connect twice to send.
- Weed out all of the Spam sent by Ratware because they fail to resend the e-mail

However, the technique of greylisting has some issues.

First, it purposefully introduces a delay in the sending of e-mail.

Second, this technique of greylisting can have severe implications for servers that don't handle temporary errors well.

This includes even prominent vendors such as Microsoft Exchange in Exchange 2003 without certain patches.

See KB934709 - **On a Windows Server 2003-based SMTP gateway server, some messages may remain in the queue folder until the SMTP service is restarted** and KB950757 - **E-mail senders do not receive an indication that some messages have been held by Exchange Server 2003 until the SMTP service, the Microsoft Exchange Information Store service, or the Exchange server is restarted).**

I've seen e-mails that are MONTHS old get delivered just because someone rebooted an Exchange 2003 based server because of greylisting.

Third, mail servers that have a lot of e-mail can take a long time before re-attempting the delivery. I've seen retries exceeding 24-hours for this issue as well.

Finally, some ISPs such as AOL, have used the number of errors in sending you e-mail to judge your server. Too many errors and you can be blacklisted or purposefully rate-limited & delayed when dealing with that ISP.

So, in conclusion, I have personally seen greylisting cause unacceptable delays in the e-mail delivery and do not recommend it.



## Common Problems

- Don't whitelist your own domain!
- FPs with Mailing Lists / Newsletters / Digests

September 16, 2009

The vast majority of Spam uses forged sender addresses as we talked about with the Reduction of DSNs. If you whitelist your own domain, you are just going to whitelist a whole bunch of Spam! Deploying SPF (as well as DKIM & ADSP) will help combat these forgeries and there are rules for these technologies that can decrease AND increase the overall score for an e-mail.

Because of the length of content having a higher probability of matching the heuristic tests for Spam, the most common FPs are newsletters and mailing list digests. For example, real estate mortgage scams abound in Spam trying to sucker people into refinancing their debt. These rules, unfortunately, may have a propensity to fire on legitimate real estate mailing lists discussing mortgages.

Unfortunately, these lists are the most likely candidates for whitelist entries!



Why is an E-Mail Tagged /  
Not Tagged as Spam?

- Review the Headers
  - Content vs. Pathway Analysis
- Checking Blacklists

September 16, 2009

Any time I am asked to review why an e-mail was or was not tagged, I start by viewing the headers of the e-mail. Too often, I find simple problems such as whitelisting enabled for a Spammer or misconfigured network rules.

More complex issues may involve how an e-mail was relayed. Since anti-Spam is as much about seeing HOW the e-mail got from point A to B as it is reviewing the content of the e-mail, the Received headers can be crucial to answer this question as well.

And if you suspect a relay is being caught by a real-time Blacklist, there are some excellent tools to check an IP address. For example, I've used <http://toolbox.webhotel.net/cgi-bin/rbl.cgi>

This tool can be invaluable if your site or someone sending to you has been infected by viruses/malware. When this occurs, a site can end up in blacklists very quickly and you'll need to work carefully through each and every blacklist's requirements for de-listing. Though don't pay any of the blacklists. No blacklists should charge for de-listing. That's a scam!

The website <http://spamlinks.net/filter-dnsbl-lookup.htm#general-sites> is a great resource for finding these tools! Here is the current list:

MultiRBL - [multirbl.valli.org/](http://multirbl.valli.org/)

Dr. Mønsted - [www.moensted.dk/spam/](http://www.moensted.dk/spam/)

SenderBase -

[senderbase.org/senderbase\\_queries/main?searchBy=ipaddress&searchString=](http://senderbase.org/senderbase_queries/main?searchBy=ipaddress&searchString=)

MXTToolBox Blacklist Lookup - [www.mxtoolbox.com/blacklists.aspx](http://www.mxtoolbox.com/blacklists.aspx)

Anti-SPAM site lookup - [tools.web-max.ca/dsbl.php](http://tools.web-max.ca/dsbl.php)

RBL Lister - [www.loosenut.com/russ-bin/rbl.pl](http://www.loosenut.com/russ-bin/rbl.pl)

Realtime Blackhole List Lookup - [www.mob.net/~ted/tools/rbl.php3](http://www.mob.net/~ted/tools/rbl.php3)

Multi-RBL check - [www.robtext.com/rbl/](http://www.robtext.com/rbl/)

Black list DB check - [www.rbl.jp/ckdb/](http://www.rbl.jp/ckdb/) - Japanese

DNSBL Info - [www.dnsbl.info/](http://www.dnsbl.info/)

Sprawdzanie RBL-i - [nospam-pl.net/rbl.php](http://nospam-pl.net/rbl.php)

rbcheck clone - [andrew.triumf.ca/cgi-bin/nph-rbcheck.cgi?addr=](http://andrew.triumf.ca/cgi-bin/nph-rbcheck.cgi?addr=)

rbcheck clone - [www.chem.utoronto.ca/cgi-bin/rbcheck.cgi](http://www.chem.utoronto.ca/cgi-bin/rbcheck.cgi)

Query RBLs - [antispam.rothen.com/rbl.php](http://antispam.rothen.com/rbl.php)

Webhostingtalk.nl DNSBL check - [www.webhostingtalk.nl/whtdnsblcheck.cgi](http://www.webhostingtalk.nl/whtdnsblcheck.cgi)

Karmasphere - [www.karmasphere.com/](http://www.karmasphere.com/)

Multi RBL Checker - [checker.msrbbl.com/v/1/](http://checker.msrbbl.com/v/1/)

MyIPTest DNSBL check - [www.myiptest.com/staticpages/index.php/check-Blacklisted-IP-DNSBL](http://www.myiptest.com/staticpages/index.php/check-Blacklisted-IP-DNSBL)

Multi-RBL Lookup Tool - <https://toolbox.webhotel.net/cgi-bin/rbl.cgi>

DNSBL Testing Services (via e-mail) - [www.crynwr.com/spam/](http://www.crynwr.com/spam/)

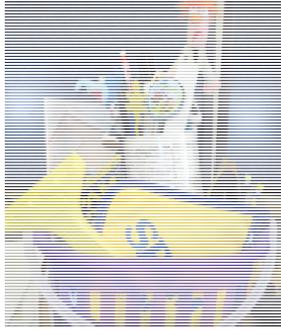
VeryNiceTools DNSBL Lookup - [verynicetools.com/blacklist](http://verynicetools.com/blacklist)

Deep VI DNSBL Lookup - [d6tech.com/support-tools/dnsbl-query.php](http://d6tech.com/support-tools/dnsbl-query.php)

EmailStuff Blacklist Lookup - [emailstuff.org/bl/](http://emailstuff.org/bl/)

## Who is Sending the Spam?

- Who is sending the Spam?
- How did they get my e-mail address?
- Why do they send it?
- Creative Spam (to a Spam Junkie)



September 16, 2009

I became interested in Spam research during a Thanksgiving holiday about a decade ago. Why this is important is that Thanksgiving is a purely American holiday. It isn't celebrated anywhere else in the world.

But strangely, I noticed that the Spam mail STOPPED. I mean zero Spam e-mails for the entire day. This indicated a very interesting fact. The Spam ALL had to be originating with Americans.

A short while later, I discovered the SpamAssassin project. I started off originally submitting patches, then assisted them with stabilizing their DNS problems. Eventually, I hosted the US mirror and then website prior to writing rules and becoming a Project Management Committee member.

These days, I see more and more Spam being sent by compromised computers. In particular, unpatched and unlicensed desktops that can't install the patches for the operating system. This means a lot of Spam is sent from pockets of the world that have more of these computers.

But this doesn't really tell you WHO is sending the Spam because more and more of the Junk Mail seems to be sent by organized crime. "Bot networks" of compromised computers that can launch Spam, malware and denial of service attacks have become quite prevalent. They can be rented and Blackhat conferences have included ex-Spammers that talk about being paid by the "200". A 2XX is the SMTP Response Code for message received OK.

And How do they get your e-mail address? They use a number of methods. For example, dictionary attacks of email to lists of names from adam@... To zeke@.... Dictionary attacks can send hundreds of thousands of emails with the possibility that only a tiny fraction, if any, will be delivered.

And, they use unsubscribe links to verify your e-mail address works! So think twice before you employ an unsubscribe link.

Also, posting your email address on a website or using a publicly archived mailing list can result in the information being "scraped" into Spammer's lists.

Most sinisterly, Spammers employ malware with another goal: getting your entire address book off your machine. Unfortunately, I've seen malware harvest these accounts from other machines on the network. And because I have many administrative accounts that are rarely if ever used on servers, I know when they have been compromised.

In fact, just in the past 3 years, I've seen Spam rapidly being sent by what appears to be more organized entities that are using much higher caliber programming as well as actively monitoring anti-Spam communities to attempt to thwart our techniques. They have also launched outright attacks to disrupt anti-Spam resources.

So why do they do send the e-mail? The main 3 reasons that come to mind for sending Junk Mail are: to make money by selling you something, to make money by stealing it

from you, and to trick you into installing malware which let's then send more Junk Mail repeating the cycle.

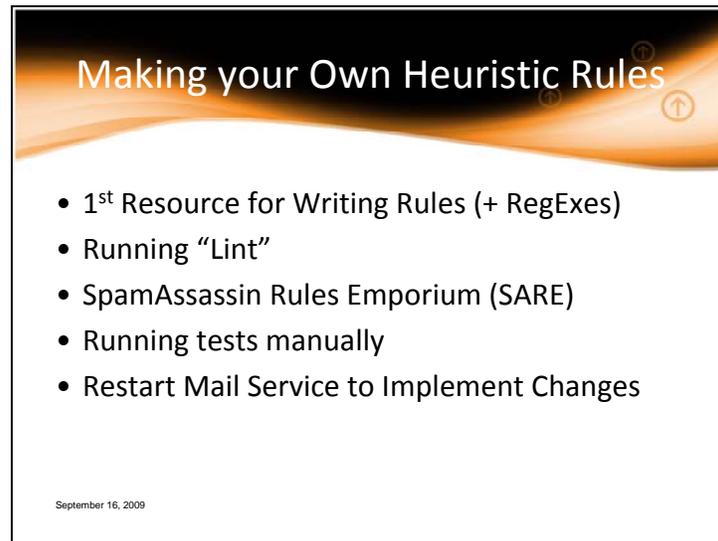
So please, let me stress a point about Why they send these e-mails. As a capitalist myself, the strongest voice I have is where I choose to spend my money. Do not spend your money on ANYTHING or ANY COMPANY that advertises via Junk Mail. And educate your users/family/friends about Junk Mail and the scams. These are really the best ways to modify the behavior and protect those you care about!

OK, so let's finish up this slide and talk about some Creative Spam. I'm a Spam junkie so I can't help but be impressed when I see a particularly creative way of sending Spam. Sometimes it's a new technology or a twist on an old scam.

Scams that target specific individuals and use the web to guess at specifics will be tough to catch. They are individually crafted messages (note that I don't require the term Bulk to be consider Spam), use online record databases, Spam using your address, nicknames and information about your neighbors.

How many people would fall for something like? "Hi Barb, this is your neighbor Carol. It's Bob's birthday on Thursday and I'm trying to buy something online with PayPal. Can you send \$20 to ItrickedYou@<free e-mail>.com, please, and I'll write you a check tonight?"

And here's another creative advertisement that was forwarded to me because of my interest in unusual Junk Mail. It contains no return address but a hand-written recipient on the outside. Inside, you have a newspaper clipping with an Advertisement for a local car dealer with a hand-written Post-it note. However, the clipping is not from any real newspaper.



**Making your Own Heuristic Rules**

- 1<sup>st</sup> Resource for Writing Rules (+ RegExes)
- Running “Lint”
- SpamAssassin Rules Emporium (SARE)
- Running tests manually
- Restart Mail Service to Implement Changes

September 16, 2009

A great place to start if you are writing rules is  
<http://wiki.apache.org/spamassassin/WritingRules>

But the first thing you have to know about writing rules and changing SpamAssassin configuration files is Lint.

Lint was a program that looked for bugs in C language source. It now refers generically to programs that look for bugs in source code. If you write your own rules or even change your configuration files, does your system pass lint? You can check with the command ‘spamassassin -D --lint’. This step is IMPORTANT.

SARE’s hints on writing new rules are invaluable. I’ve used them as the basis for this list:

Some of the rules you write may already be (or will be) incorporated into standard versions of SA. It isn't good to duplicate rules. Just raise the default score. Scoring is an art! But a general idea is to write many small rules and score them low. Let 10 rules add up to 2 points rather than 1 rule with all 10 things in it scoring 2 points.

This will help minimize false positives. I personally use a lot of sub-rules with a master “meta” rule which requires multiple conditions to trigger.

Many spam can be caught using Bayes, and we strongly recommend the use of Bayes in any/all SpamAssassin systems.

**Always** run spamassassin -D --lint before making your rule changes live.

It is **very** important to write rules for **your** circumstances.

Write some negative scoring rules!!! Example: If you work in a furniture store, then the words furniture, bed, chair, drawer, desk and lamp in an e-mail should be worth negative points!

Phrases are **much** better than single words.

Don't get too aggressive. Think about how the rule might hit on legitimate e-mail!

Once you've written a rule, test it by running spamassassin in test mode against a text file containing the entire e-mail with the command:

```
spamassassin -t -D < [message]
```

Or, for a slightly faster test which doesn't run the network tests:

```
spamassassin -t -L -D < [message]
```

NOTE: -D in the commands above prints out lots of debug information. You might find it useful or an overload of information.

Finally, any time you change a rule or configuration file, you likely need to stop and start the Mail service in Server Admin. This will restart the AMaViS daemon.



## Excuses, Excuses!

- Best Spam RBL Delisting Request
  - "Please delist us. We promote cosmic peace as in the poem below. This world badly needs it. Why do some hate peace and have us blacklisted?"
- Best Unsubscribe Tag
  - If you do not have Internet access, please send an e-mail to delete@<removed to protect the guilty>.com

September 16, 2009

I want to love you all

My name is Love.

I want to love you all.

I really do.

But many of you do not bother  
to open their heart to me.

I can only love you  
when you allow me to do so.

I can only love you  
when your heart is a place  
full of warmth and peace.  
When I am in your heart  
you can feel love always.  
You can feel my energy.  
You can feel my power.

You can feel life as never before.

I cannot love you though  
when your heart is a place  
of hate and bad feelings.

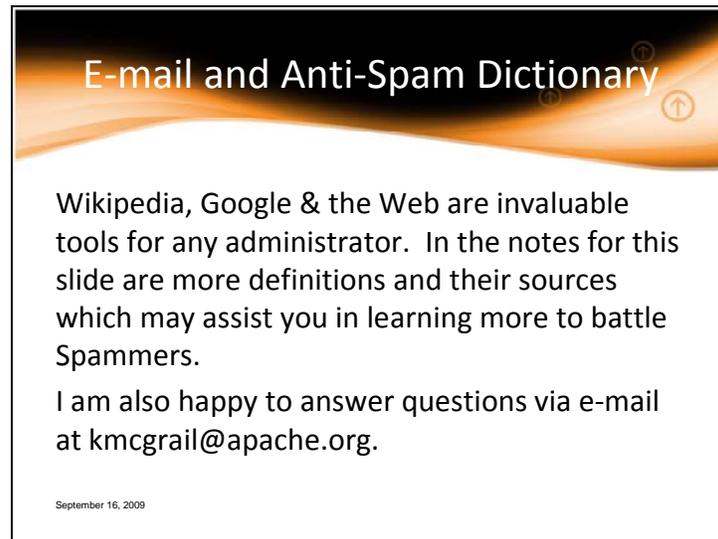
Please let me love you  
and make you a king or a queen  
in my kingdom of bountiful love.  
Please do not let hate and bad feelings  
abuse you as their pitiful slave.

Please let me love you  
and make you a king or a queen.

## Resources from this Speech

- SpamAssassin – <http://spamassassin.apache.org/>
- SpamAssassin – Writing Rules - <http://wiki.apache.org/spamassassin/WritingRules>
- Monty Python SPAM Skit – <http://www.youtube.com/watch?v=cFrtpT1mKy8>
- Hormel SPAM and the Internet – <http://www.spam.com/about/internet.aspx>
- Xcode – <http://developer.apple.com/technology/Xcode.html>
- Sender Policy Framework – <http://www.openspf.org/>
- The SpamHaus Project – <http://www.spamhaus.org/>
- URIBL – <http://www.uribl.com/>
- SURBL – <http://www.surbl.org/>
- PCCC – Peregrine Computer Consultants Corporation – <http://www.pccc.com/>
- KAM.cf – <http://www.pccc.com/downloads/SpamAssassin/contrib/KAM.cf>
- SpamTrainer – <http://osx.topicdesk.com/spamtrainer>
- Genetic Algorithm – [http://en.wikipedia.org/wiki/Genetic\\_algorithm](http://en.wikipedia.org/wiki/Genetic_algorithm)
- Template thanks to <http://www.templateswise.com/detail/link-105.html>

September 16, 2009



**E-mail and Anti-Spam Dictionary**

Wikipedia, Google & the Web are invaluable tools for any administrator. In the notes for this slide are more definitions and their sources which may assist you in learning more to battle Spammers.

I am also happy to answer questions via e-mail at [kmcgrail@apache.org](mailto:kmcgrail@apache.org).

September 16, 2009

**Phishing** - A social engineering technique used by bodies posing as a trustworthy source to steal information (i.e. - usernames/passwords, bank/PayPal account information, any information that can be used to assist in data and/or identity theft.). These attacks are typically carried out via e-mail or instant messaging and in some cases through bogus accounts on social networking services such as MySpace, Facebook, etc.

E-mail deception usually takes the form of a link in a bogus e-mail or IM (though the e-mail can be VERY convincing down to being an almost Exact copy of a legitimate e-mail from the company being "spoofed" using company logos and appearing very 'official looking'.) which leads to a website where the user being scammed will input sensitive information thus completing the theft. Requests to "verify your account" are often used in these fraudulent messages with links that appear to lead to the legitimate website to 'verify your account' but actually lead to the websites mentioned previously.

Phishers employ various other tactics in attempts to trick you into following their links and submitting confidential information. These techniques include Link Manipulation such as, Misspelled URLs([www.paypa1.com](http://www.paypa1.com) rather than [www.paypal.com](http://www.paypal.com)), and Masked URL, making the anchor text of a link appear valid but on mouseover in the tooltip that appears you will see where the link is Actually pointing to the scam site. Also, Filter

Evasion is a common technique, using images with text on them instead of actual text in an e-mail to avoid anti-phishing filters searching for text commonly used in phishing e-mails.

Not all phishing attacks require a fake website. Scammers can employ VOIP numbers to call users claiming to need Account and PIN numbers for different services. The caller ID can be spoofed to show a legitimate company or organization name. This technique of Voice Phishing is called "Vishing".

Damage caused by phishing ranges from denial of access to e-mail to substantial financial loss.

Information/Websites used in creation of this text -

<http://www.microsoft.com/protect/yourself/phishing/identify.msp>

<http://en.wikipedia.org/wiki/Phishing>

**Backscatter** (outscatter and misdirected bounces, blowback, or collateral spam, et al.)- A side effect of e-mail spam, viruses, and worms where e-mail servers receiving spam and other mail send bounce messages to innocent users. These usually come in the form of "Your mail could not be delivered.." or "Your mail contained a virus.." messages. These messages are classified as spam because they aren't solicited by the recipient and are delivered in bulk quantity. The vast majority of SPAM comes from forged e-mail addresses.

Information/Websites used in creation of this text -

<http://www.spamresource.com/2007/02/backscatter-what-is-it-how-do-i-stop-it.html>

<http://en.wikipedia.org/wiki/Outscatter>

**rPTR (see rDNS)** - Reverse DNS is a way of associating an IP address with its domain name. The reverse DNS identifier is contained in the PTR portion of the IP Zone File.

The IP Zone File contains all the different ways that your IP and domain name can be associated; each association serves a different need.

Information/Websites used in creation of this text -

<http://postmaster.aol.com/info/rdns.html>

**DNS** - The Domain Name System (DNS) is an internet directory service. DNS's most basic service is to translate hostnames into IP addresses, and DNS also controls e-mail delivery. If your computer cannot access DNS, your web browser will not be able to find web sites, and you will not be able to receive or send e-mail. The DNS system consists of three components: DNS data (called resource records), servers (called name servers), and Internet protocols for fetching data from the servers. The billions of resource records in the DNS are split into millions of files called zones. Zones are kept on authoritative servers distributed all over the Internet, which answer queries based on the resource records stored in the zones they have copies of. Caching servers ask other servers for information and cache any replies. Most name servers are authoritative for some zones and perform a caching function for all other DNS information. Large name servers are often authoritative for tens of thousands of zones, but most name servers are authoritative for just a few zones.

Information/Websites used in creation of this text -

<http://www.dns.net/dnsrd/docs/whatis.html>

[http://en.wikipedia.org/wiki/Domain\\_Name\\_System](http://en.wikipedia.org/wiki/Domain_Name_System)

**SMTP** - Simple Mail Transfer Protocol (SMTP) is the standard for e-mail transmission across the internet. A relatively simple text-based protocol, SMTP is limited in its ability to queue messages at the receiving end, it is usually used with one of two other protocols, POP3 or IMAP, that let the user save messages in a server mailbox and

download them periodically from the server. Extended SMTP(ESMTP) is the protocol used today and allows for multimedia files to be delivered as e-mail.

Information/Websites used in creation of this text -

<http://en.wikipedia.org/wiki/SMTP>

[http://searchexchange.techtarget.com/sDefinition/0,,sid43\\_gci214219,00.html](http://searchexchange.techtarget.com/sDefinition/0,,sid43_gci214219,00.html)

**Real-time Blacklist (RBL)** - The first DNSBL (DNS Blacklist) was the Real-time Blackhole List (RBL). Initially, the RBL was not a DNSBL, but rather a list of commands that could be used to program routers so that network operators could blackhole, a routing term to send all the packets into nothingness, all TCP/IP traffic for machines used to send spam or host spam supporting services, such as a website. The purpose of the RBL was not simply to block spam—it was to educate Internet service providers and other Internet sites about spam and related problems, such as open SMTP relays, spamvertising, etc. The RBL was also released in a DNSBL form and authors of Sendmail and other mail software were urged to implement RBL clients. This allowed the mail software to query the RBL and reject mail from listed sites on a per mail server basis instead of “blackholing” all traffic.

Information/Websites used in creation of this text -

[http://en.wikipedia.org/wiki/Real-time\\_Blackhole\\_List#Terminology](http://en.wikipedia.org/wiki/Real-time_Blackhole_List#Terminology)

**DNS Blacklist (DNSBL)** - A published list of IP addresses that can be queried through the Internet. DNSBLs are used to publish IP addresses associated with e-mail spam and spamming. Most mail servers can be configured to reject messages from addresses on a DNSBL. An address found in a DNSBL may be directly associated with spam, or may have made the list due to Web server vulnerabilities that can be used by spammers. There are many DNSBLs available, each published and maintained by different individuals and organizations.

Information/Websites used in creation of this text -

<http://www.webopedia.com/TERM/D/DNSBL.html>

<http://en.wikipedia.org/wiki/DNSBL>

**Reverse DNS (rDNS)** - rDNS is a process to determine the hostname or host associated with a given IP address or host address. Reverse DNS is setup by configuring PTR records (Pointer Records) in your DNS server. The Domain Name System is used to determine what IP address is associated with a given domain name. So, to reverse DNS lookup an IP address is to look up what host and domain name belongs to that IP address. There are many reverse DNS lookup tools available for free on the internet at various sites.

Information/Websites used in creation of this text -

<http://www.tech-faq.com/reverse-dns.shtml>

[http://en.wikipedia.org/wiki/Reverse\\_DNS](http://en.wikipedia.org/wiki/Reverse_DNS)

**HELO greeting (pertaining to AntiSpam)** - Spam can be greatly reduced by a number of checks confirming compliance with standard addressing and MTA operation. In many situations, simply requiring a valid FQDN (Fully Qualified Domain Name) in the SMTP EHLO statement is enough to block 25% of incoming spam.

- \* Refusing connections from hosts that begin transmission prior to presentation of the receiving host's HELO banner.

- \* Refusing connections from hosts that give an invalid HELO - for example, a HELO that is not an FQDN or is an IP address not surrounded by square brackets

Example:

Invalid: HELO localhost

Invalid: HELO 127.0.0.1

Valid: HELO domain.tld

Valid: HELO [127.0.0.1]

\* Refusing connections from hosts that give an obviously fraudulent HELO - for example, issuing a HELO using the FQDN or an IP address that doesn't match the IP address of the connecting host

Fraudulent: HELO friend

Fraudulent: HELO -232975332

\* Refusing to accept e-mail claiming to be from a hosted domain when the sending host has not authenticated

\* Refusing to accept e-mail whose HELO/EHLO argument does not resolve in DNS. Unfortunately, some e-mail system administrators ignore section 3.6 of RFC2821 and administer the MTA to use a nonresolvable argument to the HELO/EHLO command.

All of the examples above are fairly simple checks, all conform to existing standards and RFCs, and all are missing from most commercial MTA implementations available today.

Information/Websites used in creation of this text -

[http://en.wikipedia.org/wiki/Anti-spam\\_techniques\\_%28e-mail%29#HELO.2FEHLO\\_checking](http://en.wikipedia.org/wiki/Anti-spam_techniques_%28e-mail%29#HELO.2FEHLO_checking)

postmaster.aol.com - presents a set of standards, guidelines, and best practices regarding e-mail policy.

Information/Websites used in creation of this text -

<http://postmaster.aol.com/guidelines/>

**Sender Policy Framework (SPF)** - SPF is an anti-spam approach in which the Internet domain of an e-mail sender can be authenticated for that sender, thereby discouraging spam mailers, who routinely disguise the origin of their e-mail, a practice known as e-mail spoofing. SPF allows the owner of an Internet domain to use a special format of DNS TXT records to specify which machines are authorized to transmit e-mail for that domain. For example, the owner of the example.org domain can designate which machines are authorized to send e-mail whose sender e-mail address ends with "@example.org". Receivers checking SPF can reject messages from unauthorized machines before receiving the body of the message. Principles of operations are quite similar to those of DNSBL, except that SPF exploits the authority delegation scheme of the real Domain Name System. SPF and other authentication-based measures are designed to redress a vulnerability in Simple Mail Transfer Protocol (SMTP), the main protocol used in sending e-mail, which does not include an authentication mechanism.

Information/Websites used in creation of this text -

[http://searchsecurity.techtarget.com/sDefinition/0,,sid14\\_gci953520,00.html](http://searchsecurity.techtarget.com/sDefinition/0,,sid14_gci953520,00.html)

[http://en.wikipedia.org/wiki/Sender\\_Policy\\_Framework](http://en.wikipedia.org/wiki/Sender_Policy_Framework)

**DomainKeys Identified Mail (DKIM)** - DomainKeys Identified Mail is a method for E-mail authentication. It offers almost end-to-end integrity from a signing to a verifying Mail transfer agent (MTA). In most cases the signing MTA acts on behalf of the sender by inserting a DKIM-Signature header, and the verifying MTA on behalf of the receiver, validating the signature by retrieving a sender's public key through the DNS.

Information/Websites used in creation of this text -

<http://en.wikipedia.org/wiki/DKIM>

**Post Office Protocol (POP3)** - an application-layer Internet standard protocol, to retrieve e-mail from a remote server over a TCP/IP connection. The design of POP3 and its procedures supports end-users with intermittent connections (such as dial-up

connections), allowing these users to retrieve e-mail when connected and then to view and manipulate the retrieved messages without needing to stay connected. Although most clients have an option to leave mail on server, e-mail clients using POP3 generally connect, retrieve all messages, store them on the user's PC as new messages, delete them from the server, and then disconnect. This standard protocol is built into most popular e-mail products, such as Eudora and Outlook Express. It's also built into the Netscape and Microsoft Internet Explorer browsers. POP3 is designed to delete mail on the server as soon as the user has downloaded it. However, some implementations allow users or an administrator to specify that mail be saved for some period of time. POP can be thought of as a "store-and-forward" service. An alternative protocol is Internet Message Access Protocol (IMAP). IMAP provides the user more capabilities for retaining e-mail on the server and for organizing it in folders on the server. IMAP can be thought of as a remote file server.

Information/Websites used in creation of this text -

[http://searchexchange.techtarget.com/sDefinition/0,,sid43\\_gci212805,00.html](http://searchexchange.techtarget.com/sDefinition/0,,sid43_gci212805,00.html)

<http://en.wikipedia.org/wiki/POP3>

**Internet Message Access Protocol (IMAP)** - a method of accessing electronic mail or bulletin board messages that are kept on a (possibly shared) mail server. In other words, it permits a "client" e-mail program to access remote message stores as if they were local. For example, e-mail stored on an IMAP server can be manipulated from a desktop computer at home, a workstation at the office, and a notebook computer while traveling, without the need to transfer messages or files back and forth between these computers. IMAP's ability to access messages (both new and saved) from more than one computer has become important as reliance on electronic messaging and use of multiple computers increase, but this functionality cannot be taken for granted: the widely used Post Office Protocol (POP) works best when one has only a single computer, since it was designed to support "offline" message access, wherein messages are downloaded and then deleted from the mail server.

Information/Websites used in creation of this text -

<http://en.wikipedia.org/wiki/IMAP>

<http://www.imap.org/about/whatisIMAP.html>

**<Mail Protocol>(i.e. - POP,IMAP,etc.) before SMTP** - <Mail Protocol>(POP will be used in this text) before SMTP is a method of authorization used by mail server software which helps allow users the option to send e-mail from any location, as long as they can demonstrably also fetch their mail from the same place. Users are allowed to use SMTP from an IP address as long as they have previously made a successful login into the POP service at the same mail hosting provider, from the same IP address, within a predefined timeout period. The main advantage of this process is that it's generally transparent to the average user who will be connecting with an e-mail client, which will almost always make a connection to fetch new mail before sending new mail. The disadvantages include a potentially complex setup for the mail hosting provider (requiring some sort of communication channel between the POP service and the SMTP service) and uncertainty as to how much time users will take to connect via SMTP (to send mail) after connecting to POP. Those users not handled by this method need to resort to other authorization methods. Also, in cases where users come from externally controlled dial-up addresses (more specifically, all dynamically assigned IP addresses), the SMTP server must be careful about not giving too much leeway when allowing unauthorized connections, because of a possibility of race conditions leaving an open mail relay unintentionally exposed.

Information/Websites used in creation of this text -

[http://en.wikipedia.org/wiki/POP\\_before\\_SMTP](http://en.wikipedia.org/wiki/POP_before_SMTP)

<http://popbsmtp.sourceforge.net/>

**SMTP AUTH** - SMTP-AUTH is an extension of the Simple Mail Transfer Protocol (SMTP) to include an authentication step through which the client effectively logs in to the mail server during the process of sending mail. Servers which support SMTP-AUTH can usually be configured to require clients to use this extension, ensuring the true identity of the sender is known. SMTP-AUTH is defined in RFC 4954.

Information/Websites used in creation of this text -

[http://en.wikipedia.org/wiki/SMTP\\_AUTH](http://en.wikipedia.org/wiki/SMTP_AUTH)

**Request for Comments (RFC)** - Documents published as a series of memos encompassing new research, innovations, and methodologies applicable to internet technologies for review by peers or to convey new technologies/protocols. The Internet Engineering Task Force (IETF) adopts RFCs as Internet Standards.

Information/Websites used in creation of this text -

[http://en.wikipedia.org/wiki/Request\\_for\\_Comments](http://en.wikipedia.org/wiki/Request_for_Comments)

**Lightweight Directory Access Protocol (LDAP)** - LDAP is an Internet protocol that e-mail and other programs use to look up information from a server. LDAP is used to look up encryption certificates, pointers to printers and other services on a network, and provide "single sign-on" where one password for a user is shared between many services. LDAP is appropriate for any kind of directory-like information, where fast lookups and less-frequent updates are the norm. As a protocol, LDAP does not define how programs work on either the client or server side. It defines the "language" used for client programs to talk to servers (and servers to servers, too). On the client side, a client may be an e-mail program, a printer browser, or an address book. The server may speak only LDAP, or have other methods of sending and receiving data—LDAP may just be an add-on method.

LDAP also defines:

Permissions: set by the administrator to allow only certain people to access the LDAP database, and optionally keep certain data private.

Schema: a way to describe the format and attributes of data in the server.

Information/Websites used in creation of this text -

<http://www.gracion.com/server/whatldap.html>

<http://en.wikipedia.org/wiki/LDAP>

**Blacklist (See DNSBLs)** - a basic access control mechanism that allows every access, except for the members of the black list (i.e. list of denied accesses). The opposite is a whitelist, which means allow nobody, except members of the white list. As a sort of middle ground, a greylist, contains accesses that are temporarily blocked.

Information/Websites used in creation of this text -

[http://en.wikipedia.org/wiki/Blacklist\\_%28computing%29](http://en.wikipedia.org/wiki/Blacklist_%28computing%29)

**POP vs. IMAP** - POP and IMAP both have various Pros and Cons associated with each protocol.

Advantages of POP3 include:

Message storage is limited only by the capacity of your computer.

Minimum use of connect time.

Minimum use of server resources.

It is less likely to exhaust disk space on the server.

Disadvantages to POP3 include:

Reading your e-mail from multiple computers or e-mail programs results in messages scattered about.

Messages are stored on your computer. If your computer fails you may lose all your e-mail.

You are not able to preview new messages before downloading, nor do you have control over which messages can be downloaded.

Once delivered, e-mail messages are stored on your local computer and deleted from the mail server.

Advantages of IMAP include:

Messages are stored on the server and are not affected if your computer fails.

Easily use multiple computers or e-mail programs to read mail.

Faster start-up time, as only message headings are transferred initially

Optimization for low-speed connections.

Disadvantages to IMAP include:

Mail is not usually available if you are offline.

Sensitive to size and requires periodic archival of e-mail messages

Subject to storage quotas

Very few ISPs and e-mail providers offer IMAP as it is considered a high end option and it's complex for them to support

Not all e-mail programs support it properly

Information/Websites used in creation of this text -

<http://saturn.med.nyu.edu/it/help/email/imap/index.html>

[http://www.washington.edu/computing/windows/issue13/imap\\_pop.html](http://www.washington.edu/computing/windows/issue13/imap_pop.html)

<http://www.uoregon.edu/~mcshtml/email/popvsimap.html>

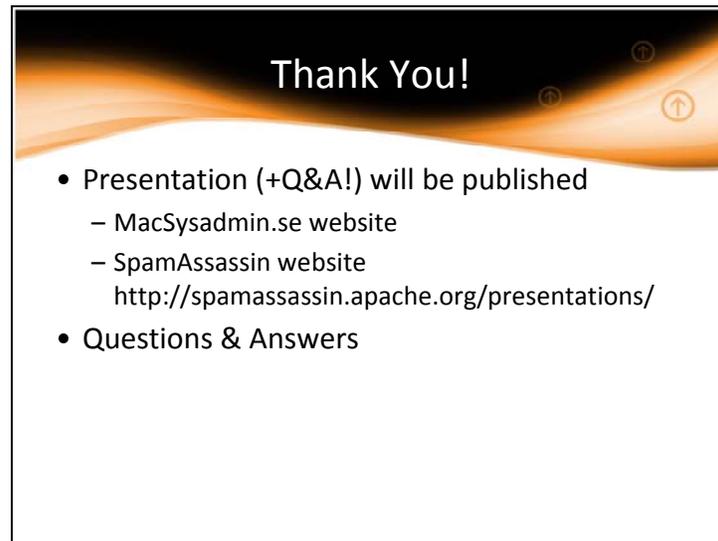
<http://www.it.northwestern.edu/accounts/email/imap/pop-imap-comparison.html>

**TTL - (Time to Live)** - Occur in the Domain Name System (DNS), where they are set by an authoritative nameserver for a particular resource record. When a caching nameserver queries the authoritative nameserver for a resource record, it will cache that

record for the time (in seconds) specified by the TTL. Shorter TTLs can cause heavier loads on an authoritative nameserver, but can be useful when changing the address of critical services like web servers or MX records, and therefore are often lowered by the DNS administrator prior to a service being moved, in order to minimize disruptions.

Information/Websites used in creation of this text -

[http://en.wikipedia.org/wiki/Time\\_to\\_live#Time\\_to\\_live\\_of\\_DNS\\_records](http://en.wikipedia.org/wiki/Time_to_live#Time_to_live_of_DNS_records)



**Q&A:**

Since the GUI options are small and I am covering far too much information for 60 minutes, I'll also do my best to answer questions via email. I will also work to add Q&A's that I believe will be helpful to others to the notes for this presentation when published on the web. Besides that macsysadmin.se website, the presentation will also be found at <http://spamassassin.apache.org/presentations/>

One of the questions I received from an early review of my presentation was:

**Q:** You talk about Apple using an old version of SpamAssassin. Do you have any info if upgrading breaks the GUI?

**A:** The interaction of the GUI is fairly small. I'm running the latest SA on an OS X Tiger (10.4.11) machine. However, after using the GUI once to turn on the filtering, I then just bypass the GUI using only the CLI. I've also looked at a new install of Snow Leopard (10.6.1). In the end, I believe the OS X GUI interface for the Junk Filter will not break as it only does the following:

Edits a few options in /etc/mail/spamassassin/local.cf

Edits a few options in /etc/mail/amavisd.conf

Changes the startup status for the various services

The next question I received was if people really are that gullible and fall for these scams via e-mail. The answer is yes so I remind administrators to PLEASE warn users and help educate them! Some good places to start are:

About.com: Internet For Beginners: **The Top 10 Internet/Email Scams**

<http://netforbeginners.about.com/od/scamsandidentitytheft/ss/top10inetscams.htm>

FTC Facts for Consumers: **You've Got Spam: How to "Can" Unwanted Email**

<http://www.ftc.gov/bcp/edu/pubs/consumer/tech/tec02.shtm>

The other questions I received dealt with a feature Kerio recommended called a "Greeting Pause" and the use of invalid HELO strings to block email.

I'll answer the HELO Strings question first. If you look in the Dictionary information under slide 24, you'll note I talk about invalid HELO string rejections. Postfix can do this as shown at [http://www.howtoforge.com/virtual\\_postfix\\_antispam](http://www.howtoforge.com/virtual_postfix_antispam). However, this technique has considerable FPs if this box also can except emails from your users and you have no way of bypassing this test for them. This is because end-user boxes often identify themselves as HELO <machinename> without a FQDN. Also, anti-virus software often is configured to proxy e-mail between the MUA and the MSA/MTA for the user. This software also has a high-rate of FPs for a HELO string compliance test.

The greeting pause is a technique where you introduce a delay of up to 30 seconds to the SMTP communication's initial greeting. Then, if the sending server ignores this delay and starts sending data, you close the connection. This is a great technique and I

use a greeting pause of 1.25s on my servers. Kerio's default setting of 15 seconds is a bit long but still effective. Anything higher than 15s and you will run in to issues. Whether there is a higher level of efficiency between 1.25s and 15s, I am unsure. However, this technique is excellent and has virtually no FPs. For sendmail, this option is added with a simple FEATURE(`greet\_pause', `1250') in your sendmail.mc. For Postfix, this feature is part of the reject\_unauth\_pipelining. From <http://www.postfix.org/postconf.5.html>:

### **sleep seconds**

Pause for the specified number of seconds and proceed with the next restriction in the list, if any. This may stop zombie mail when used as:

/etc/postfix/main.cf:

```
smtpd_client_restrictions =  
    sleep 1, reject_unauth_pipelining  
smtpd_delay_reject = no
```

This feature is available in Postfix 2.3.

---

i Speakers from the MacSysAdmin 2009 Conference:



Arek Dreyer

---

Consultant, Trainer and Author, [Dreyer Network Consultants](#)

Arek Dreyer is a trainer and consultant specializing in Mac OS X Server. In 2008 he helped update the Apple Global Training courses for Directory Services as well as Xsan 2 Administration. He is the author of the reference guide, Mac OS X Directory Services v10.5. He is currently updating the Mac OS X Server Essentials and Directory Services courses for Mac OS X 10.6.

Arek been an Apple Certified Trainer since 2002. Originally an expert in Sun systems, Dreyer shifted his professional focus to Mac OS X shortly after its release. President of Dreyer Network Consultants, Inc, Arek lives in Chicago but sometimes gets to visit other parts of the world as part of his work. He speaks, deliver courses, and implements solutions for customers.



Kevin A. McGrail aka KAM

Project Management Committee Member, [Apache Software Foundation SpamAssassin Project](#)  
President, [PCCC](#)

Kevin is based just outside of Washington, D.C. in the US. He hates spammers.

He takes borderline freakish delight in analyzing spam messages and combating spammers by writing rule sets for SpamAssassin.

Kevin enjoys using a heterogeneous environment of machines and operating systems in concert with one another but often prefers a command line interface to a GUI.

At PCCC, Kevin has had the pleasure of solving a myriad of computer problems for all types of customers since 1993. He has also served as CTO at a number of Internet startups including Romlight.com (acquired by MovieGallery), PickTheHits.com (lost money but made it up in volume) and ThoughtWorthy Media (website sold to Gemstar/TV Guide).

---

As the father of three beautiful children, one being a son with autism, Kevin volunteers much of his extra time & energy as the President and Webmaster for [POAC-NoVA](#).



Matthias Fricke

Training Manager / Apple Mentor Trainer, [Assense Software Solutions](#)

WebObjects, Cocoa and Mac OS X Instructor Matthias Fricke has more than 15 years experience in the IT sector. In the early 90s he worked for the German NeXT Distributor DART Software and co-founded later the WebObjects consulting company NetMatic Internet/Intranet Solutions.

Matthias spend more than 8 years in the US for his companies and worked the last years there at Apple as the Worldwide Training Delivery Manager. He moved back to Germany in 2007 and is now working for Assense Software Solutions in Hamburg. Since the end of 2007 he also teaches for Apple (EMEIA) as a T3 (Train the Trainer) Instructor and prepares Trainers to become Apple Certified Trainers for the Apple Certified IT classes. He cowrote the Mac OS X Advanced System Administration v10.5 course for Apple Global Training / PeachPit and also translated some of Apples IT-course material and examen to German.



Karl Kuehn

---

Lead Developer, [InstaDMG](#)

Over the last 10 years Karl Kuehn has worked for 3 large universities specializing in the Macintosh platform. During that time he has increasingly focused on the deployment systems for large numbers of computers. Going beyond the traditional system administrator roll Karl has become involved in a couple of open-source projects, and has created number of small programs for system administrators.

Karl has become the lead developer on the InstaDMG project, and was the sole developer on the InstaUp2Date add-on to that project. He is currently working on the second version of InstaDMG.



John Jones

Sales Engineer, [Kerio](#)

John has been a Sales Engineer at Kerio Technologies for 5 years. He provides High level technical support for the companies reseller network and helps solve the intricate problems resellers customers pose them. A intricate knowledge of programming languages coupled with Practical experience in Linux, Windows and Mac OS deployments means a sense a déjà vu is never far away and this serves to better understand challenges that developers and customers face. He has qualifications in Software Engineering Management and a Masters in Education from Cambridge University.

Prior to Kerio John worked at ARM and MIPS Technologies and has always been fascinated by simple, stable and secure digital communication and how it evolves.



P-M Lejon

Mac sysadmin,  
[Newspaper Expressen](#)

P-M has spent most of his career at evening newspaper Expressen, doing almost everything IT related spanning from helpdesk (anyone remember 300 baud modems?) to the current position as Mac sysadmin.

At Expressen, the guiding principle has always been that the computer should help streamline the editorial process, not slow it down. To facilitate this, and to create a consistent and unified Mac work environment in the newsroom, FileWave was chosen as a key system over ten years ago.

Since the migration to OS X, P-M has tweaked FileWave into a fine grained tool and integrated the Macintosh environment into the company AD.



Sascha Uhl

Head of Sales Engineering, EMEA, [Parallels](#)

Sascha is Head of Sales Engineering at Parallels, a worldwide leader in virtualization and automation software that optimizes computing for consumers, businesses, and cloud services providers. He also works as a technical consultant and project leader for international projects in Europe and APAC implementing Software as a Service and Virtual Desktop Infrastructures.

---

Sascha combines a comprehensive knowledge of the technical background and the business demands of the rapidly growing virtualization market. Before he joined Parallels in summer 2005, he worked several years for AVAYA and ITS.

## MacSysAdmin 2009

- European Macintosh System Administrators Meeting 2009 -



Matt Wynne

Technical Director,  
[Databubble Ltd](#)

Matt is an IT consultant, speaker and trainer, with specific emphasis in integrating Macs into 3rd Party networks. He has been working with Macs since 1992 and has been an Apple Authorised Trainer for 7 years, delivering all the IT courses that Apple provide. As the Technical Director of Databubble Ltd, he is often asked to provide consultancy on capital projects involving large scale deployments of Macs within the corporate and tertiary education sectors. He is also involved in Databubble's SME clients, usually getting his 'hands dirty' with server installations and MCX configurations.

In previous lives, Matt was in the Royal Air Force and served for 10 years as a Navigator and Air Traffic Controller, moving onto become IT Director at the change consultancy MCA in 1997.

He is married with 2 young children and currently lives in Cheshire. although his work takes him all over the UK and Europe.



---

Timo Sirainen

Lead Developer, [Dovecot](#)  
Software Developer, [Rackspace Hosting](#)

Timo has been developing the Dovecot IMAP server for the last seven years. Initially it started as a hobby, but as the program grew more mature several different companies have been funding his further development of Dovecot.

In January 2009 Timo moved from Finland to the US and has since then been working full time for the Email and Apps division of Rackspace, which is using Dovecot to serve over a million mailboxes.



Charles Edge

Consultant, Author & Engineer,  
[318](#) and [krypted.com](#)

Charles Edge is author of a number of titles on Mac OS X Server and systems administration topics, including the upcoming Enterprise Mac Administrator's Guide from Apress. Charles has spoken at a variety of conferences including DefCon, BlackHat, LinuxWorld, MacWorld and the WorldWide Developers Conference.

Charles is the Director of Technology at 318, based in Santa Monica, California, the largest Mac consultancy in the United States. At 318, Charles leads a team of over 40 engineers and has worked with network architecture, security and storage for various vertical and horizontal markets. Charles maintains the 318 corporate blog at [318.com/techjournal](#) as well as a personal site at [krypted.com](#).



Joakim Jardenberg

CEO, [Mindpark](#)

Old timer from the borderland between traditional and new media. CEO and president of the digital developing agency Mindpark owned by five strong media houses.

To the Mac community Joakim is known as the force behind the [Crack A Mac contest](#) in 1997, where he challenged the world to crack a Mac server.

Joakim participated in the construction of Sweden's first successful newspaper site on the web, [aftonbladet.se](#).

He is an inspiring speaker with the motto: Those who say that it cannot be done should get out of the way of those who are doing it!



Kjell Högström

Senior Systems Engineer,  
[Sun Microsystems Nordic](#)

Kjell has worked for Sun for more than eleven years. He has been working as a Solaris presales specialist for more than ten years. As a member of Sun's group of Solaris presales experts called the OS Ambassadors, which have approximately 60 members from the field, he has been in regular contact with Solaris engineering and product marketing. He also works with virtualization of Solaris environments using Solaris Containers and Logical Domains.

---

He is working with most of Sun's big customers in Sweden. Lately with some focus on bank and finance customers.



Alan Gordon

Chief Technology Officer, [Humac Group](#)

Alan is the CTO for the Humac Group, and has the responsibility for service, support, consultancy and also has the responsibility for Humac's internal IT services.

Alan has worked with solutions based around Apple technology in the past 14 years. Alan still get "hands on" with projects and his main emphasis is deploying Mac's in heterogeneous environments and software distribution & deployment.



Thomas Kaufmann

Marketing, [FileWave](#)

Thomas is responsible for marketing and communication of FileWave organizations worldwide. As a career changer he came from the publicity and advertising branch into the world of IT. Two years ago he started in sales at FileWave and learnt fast about the wishes and needs of the enterprise information technology from the administration as well as from the economical aspects.

Thomas made his apprenticeship in marketing and advertising with a qualification for universities. Today he is working extra-occupational on his business economist degree.



Serge Robe

Senior Product Marketing Manager, [VMware](#)

Serge is Senior Product Marketing Manager at VMware EMEA; he leads the VMware Fusion and the VMware vSphere offer for Small and Midsize Businesses for international.

Serge comes directly from Apple Europe where he spent the last 10 years in various consulting engineering, product marketing and business development roles. He was the Mac OS X client and Server solutions evangelist for Europe. He co-designed what became Podcast Producer on Mac OS X Server Leopard. Serge spent nine years with Digital Equipment Corporation too. Serge is a graduate from ESME Sudria Paris School of Engineering.