SpamAssassin

Наемный убийца

Спам становится интернет-чумой XXI века, и это уже не преувеличение. Спамеры не делают различий между корпоративными клиентами и частными пользователями, бессмысленная реклама валится во все ящики.

Не волнует спамеров и то, какой операционной системой вы пользуетесь. И тут каждый сам выбирает свой способ борьбы. Linux в данном случае не исключение, для этой ОС есть свои утилиты и пакеты, предназначенные для борьбы со злом.

В данной статье мы рассмотрим локальную установку и настройку достаточно мощного комплекса для борьбы со спамом — пакета SpamAssassin, а также его интеграцию с утилитами фильтрации входящей почты.

Нанимаем киллера

Итак, средства, которые нам понадобятся:

- Собственно сам дистрибутив Linux. В данном случае мы будем пользоваться для примера отечественным дистрибутивом ASPLinux 9.
- Последняя версия пакета SpamAssassin в виде файла .src.rpm с домашней страницы проекта (http://spamassassin.org/released/RPMs), который и будет непосредственно заниматься расфасовкой почты на спам и не спам.
- Также необходимо убедиться, что у нас установлены требуемые пакеты для фильтрации почты — fetchmail и procmail. Проверить их наличие в системе можно, задав в консоли команду:

```
rpm -q fetchmail procmail
```

Если все, что нам потребуется, в системе есть, мы увидим примерно следующее сообщение:

```
fetchmail-6.2.2-1asp
procmail-3.22-9
```

Если же данные пакеты отсутствуют, нам необходимо найти их на дисках дистрибутива и установить командами:

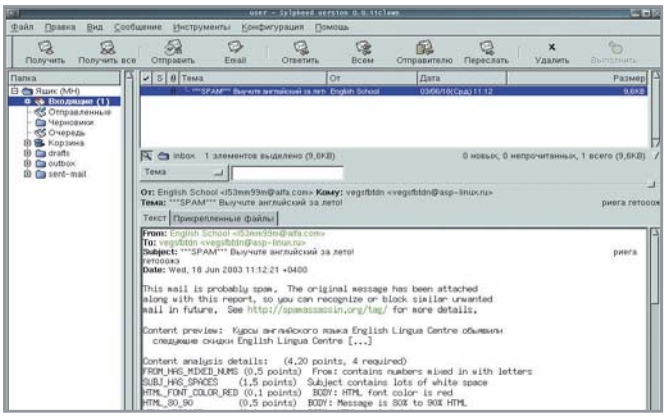
```
rpm -ivh fetchmail-<версия>
rpm -ivh procmail-<версия>
```

- Установленные пакеты для разработки, такие как gcc, perl и прочие. Для того чтобы не задумываться о наличии данных пакетов, рекомендуется выбирать набор «Разработка» при первоначальной установке системы.

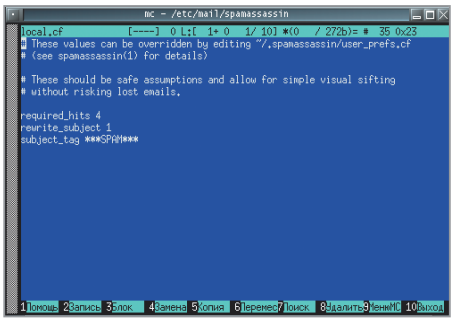
Подгоняем снаряжение

Теперь, когда сделаны все необходимые приготовления, можно приступать к сборке и установке пакета SpamAssassin.

Примечание: все указанные ниже действия выполняются в консольном режиме. »



От таких писем не застрахован никто. Ничего кроме раздражения они не вызывают, но приходят регулярно



Настраиваем параметры Spam Assassin, от которых будет зависеть определение письма как спама

Для начала скачаем последнюю версию пакета с помощью wget:

```
wget www.spamassassin.org/released/
RPMs/spamassassin-2.55-1.7.3.src.rpm
```

Теперь приступим к сборке. Находясь в каталоге с данным пакетом, вводим команду:

```
rpmbuild --rebuild spamassassin-2.54-
1.src.rpm
```

Начнется проверка необходимых зависимостей:

```
./configure
checking for gcc... gcc
checking for C compiler default output...
a.out
checking whether the C compiler works...
yes
checking whether we are cross compi-
ling... no
checking for suffix of executables...
checking for suffix of object files... o
checking whether we are using the GNU C
compiler... yes
checking whether gcc accepts -g... yes
```

Если на данном этапе не случилось никаких ошибок и все требуемые зависимости разрешены, начинается непосредственно построение пакетов.

Если же ошибки все-таки возникли, прочитайте последние сообщения процесса сборки, это поможет вам установить, что именно стало причиной сбоя и какие действия необходимо предпринять для ее устранения.

После того как процесс сборки пакетов завершится, вы сможете найти их в каталоге /usr/src/asplinux/RPMS/i386.

Перейдем в этот каталог:

```
cd /usr/src/asplinux/RPMS/i386/
```

В нем должны находиться три пакета, получившиеся после предыдущего шага: perl-Mail-SpamAssassin, spamassassin-tools, spamassassin-2.55.

Вводим команду, которая установит эти пакеты:

```
rpm -ivh perl-Mail-SpamAssassin-2.55-
1.i386.rpm spamassassin-tools-2.55-
1.i386.rpm spamassassin-2.55-1.i386.rpm
```

Теперь внесем необходимые изменения, которые улучшат фильтрацию, в конфигурационные файлы пакета. Откроем в режиме редактирования файл /etc/mail/spamassassin/local.cf, в котором изменим и при необходимости добавим следующие строки:

- **required_hits 4** — письмо будет считаться спамом, если по сумме баллов в тестах наберет более четырех очков. Подробнее о системе тестов мы расскажем далее.
- **rewrite_subject 1** — переписывать поле «Тема» писем, получивших статус спама, для более легкой и корректной их фильтрации в почтовом клиенте. Я не рекомендую сразу удалять письма, которые были помечены как спам, потому что при определенных обстоятельствах программа может допустить ошибку и посчитать ненужным письмо, которое на самом деле является нормальной корреспонденцией.
- **subject_tag ***SPAM*** — собственно сам текст переписанного поля «Тема» сообщения. Например, если тема была: «**American Language Center**», то она будет изменена на «*****SPAM*** American Language Center**».

После выполнения этих действий минимальную настройку SpamAssassin можно считать завершенной. При стандартных установках будет отфильтрована большая часть спама.

Командная игра

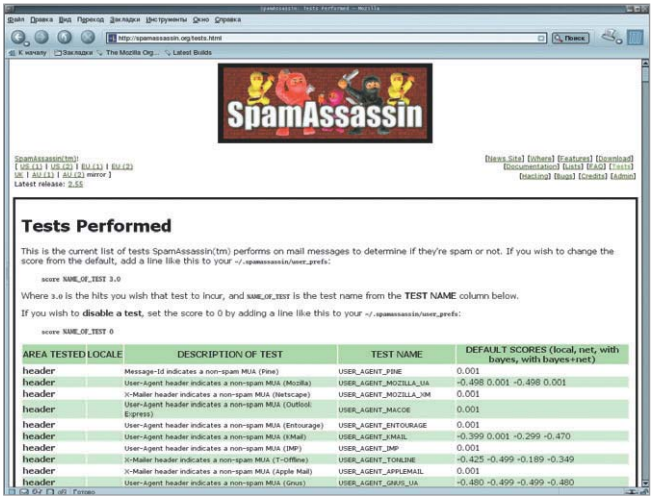
Сам по себе пакет SpamAssassin будет не слишком полезен на машине пользователя. Его следует использовать вместе с почтовым клиентом и другими почтовыми утилитами, например с программами фильтрации почты.

Для того чтобы установить между нашими «защитниками» взаимопонимание, сделаем необходимые настройки fetchmailrc и procmail.

В домашнем каталоге создадим файл .fetchmailrc командой **touch .fetchmailrc**, где сделаем необходимые записи о серверах, с которых мы хотим получать почту. Например, для сервера server.ru можно сделать следующую конструкцию:

- **server pop.server.ru** — собственно адрес почтового сервера, с которого мы получаем почту;
- **proto pop3** — протокол, используемый для получения почты;
- **auth password** — тип аутентификации с использованием пароля;
- **user login** — учетное имя на сервере;
- **pass password** — пароль на сервере;
- **flush** — удалять почту с сервера;
- **mda "/usr/bin/procmail -d %T"** — передавать полученные сообщения procmail для фильтрации.

Соответственно теперь нам необходимо настроить фильтрацию в procmail. Создадим еще один файл, теперь с именем .procmailrc и впишем в него следующие строки:



На сайте проекта SpamAssassin вы можете подробнее ознакомиться с тем, что и как оценивает каждый из тестов

```
:0 f: .spam_check.procmail.lock
| /usr/bin/spamassassin -L
```

Эти строки означают, что все сообщения, переданные procmail, будут обрабатываться с помощью SpamAssassin. После обработки они будут находиться в локальной почтовой директории на вашем компьютере. По умолчанию это каталог /var/spool/mail/<user>, где <user> — имя пользователя, для которого мы и определяем текущие настройки.

Если у вас есть постоянное подключение к Сети и вы желаете выполнять как можно меньше действий вручную, можно настроить автоматическую проверку почты через заданный промежуток времени, например, используя сервис cron.

Инструкции почтальону

Если все действия выполнены успешно, то мы имеем в своем распоряжении минимально необходимый комплекс средств, которые помогут по большей части забыть о такой проблеме, как спам. Осталась самая малость — настроить вашу почтовую программу.

Откроем диалог создания фильтров и создадим новое правило, в соответствии с которым все письма, в теме которых содержится фраза «*****SPAM*****», помещались бы в отдельную папку или, по желанию, просто удалялись.

Также в настройках учетной записи почтового клиента стоит указать получение почты из локальной директории /var/spool/mail/<user>.

Теперь для получения почты нам осталось лишь дать в консоли команду fetchmail, после чего письма будут доставлены на ваш компьютер.

Заберите их вашим почтовым клиентом, нажав на кнопку «Получить почту» или аналогичную, в зависимости от того, какой программой вы пользуетесь. Разумеется, использование этого комплекса не избавит вас от спама на все сто процентов.

К сожалению, спамеры придумывают все новые и новые средства для того, чтобы обойти фильтры почтовых программ, но это уже беда не программистов, а морали тех, кто спам рассылает. Тем не менее значительную часть «мусора» SpamAssassin уничтожит на корню.

Александр Быков

Фильтры и правила

Кодекс SpamAssassin

Фильтрация сообщений в SpamAssassin построена на системе правил, которые проверяют каждое письмо, просматривая такие его части, как тема, тело письма, адрес отправителя, общую структуру сообщения. Каждое правило ассоциируется с каким-либо количеством очков, которое получает письмо при соответствии этому правилу. При желании мы можем изменить количество очков для каждого из правил, указав требуемое значение в файле **~/.spam assassin/user_prefs.**

После прохождения набора правил, которые для каждого письма могут быть различны и определяются внутренней логикой программы, каждое письмо получает итоговое количество очков. Результат сверяется с максимально возможным, и при превышении этого порога письму присваивается статус спама, в результате чего сообщение принимает примерно следующий вид:

```
От:     American Language Center
        <ayhhjuw@galamail.com>
Кому:   Info <info@asplinux.ru>
Тема:   ***SPAM*** 105-5186
Дата:   Mon, 16 Jun 2003 18:06:16
        +0000
This mail is probably spam.
Content preview: AMERICAN LANGUAGE
CENTER
Content analysis details: (6.40 points,
4 required)
```

В данном примере мы можем видеть, что к письму были применены следующие правила определения статуса (в реальности правила могут быть и другими, это зависит от настроек комплекса):

```
BAYES_90 (4.0 points) BODY: Bayesian classi-
fier says spam probability is 90 to 99%
CALL_NOW (0.9 points) BODY: Urges you to
call now
```

Первое правило говорит о том, что анализ тела письма по методу Байеса показал, что вероятность того, что это письмо спам, от 90 до 99%. Соответственно, сообщению сразу было присвоено 4 очка по одному лишь тесту.

Второе правило сработало в ответ на призыв в теле письма позвонить по какому-либо номеру телефона. То есть, за данный тест письмо получило еще 0,9 балла. В итоге после прохождения остальных тестов письмо набрало 6,4 балла из 4 требуемых, что является даже перебором. Логический вывод: данное письмо однозначно является спамом.

Подробное описание всех тестов, а также количество максимально возможных очков (при настойках программы «по умолчанию») для каждого из них и другие материалы вы сможете прочитать на странице программы http://useast.spamassassin.org/tests.html.