Тестирование Rspamd — AntiSpam решения с открытым кодом

Цели проекта

- Тестирование эффективности продукта Rspamd в части блокирования известных спам-экземпляров.
- Автоматизация процесса тестирования, в том числе:
  - создание тестового среды (генератор писем, почтовый сервер назначения, почтовый шлюз с функциями AntiSpam & AntiVirus);
  - отправки спам-экземпляров и формирования отчета о заблокированном проценте писем.

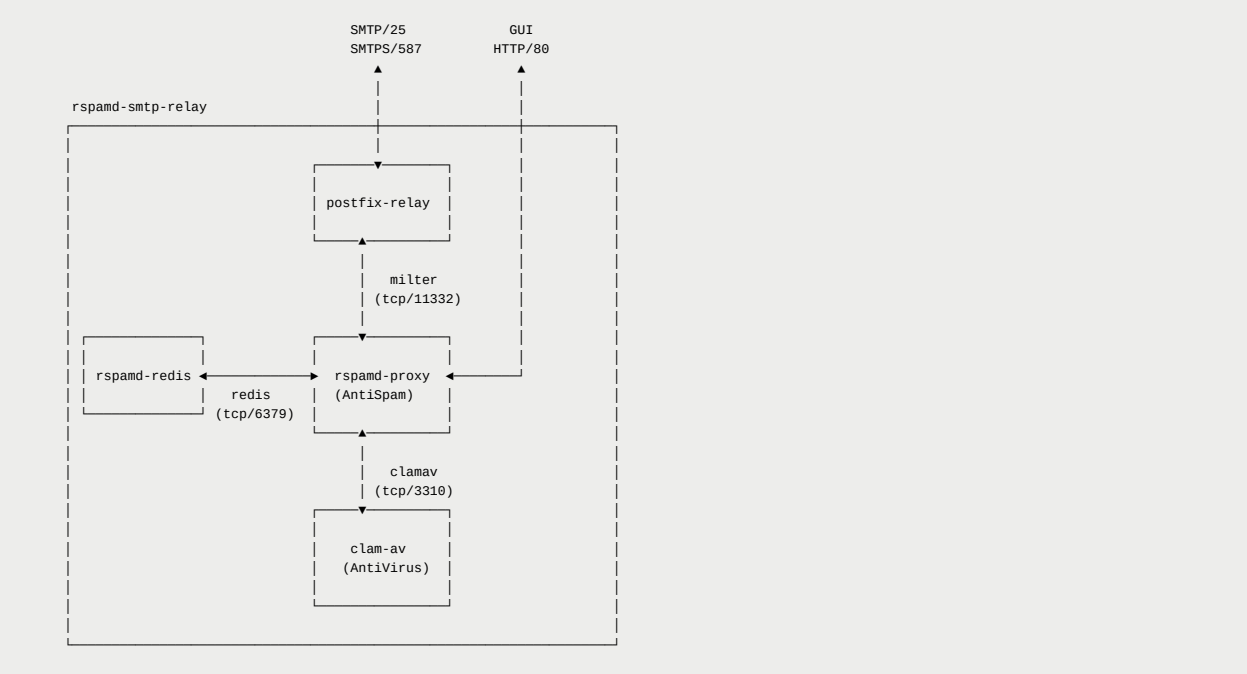
Результаты

1. Структура AntiSpam решения

AntiSpam решение (далее — **rspamd-smtp-relay**) состоит из следующих структурных компонентов:

- **postfix-relay** — почтовый сервер с функцией Relay (MTA).
- **rspamd-proxy** — программный модуль для выявления признаков спама в письмах.
- **rspamd-redis** — база данных rspamd-proxy, предназначенная для кеширования, хранения байесовской модели и модели нейронной сети.
- **clam-av** — программный модуль для выявления вредоносного кода во вложениях письма (AntiVirus).

Связи между компонентами можно отобразить следующим образом:



2. Автоматизация создания тестовой среды

Для создания тестовой среды используются инструменты окружения Docker, в частности:

- генератор писем (**mail-generator**) — готовый к использованию образ Docker (ссылка);
- почтовый сервер назначения (**fakemail-server**) — готовый к использованию образ Docker (ссылка);
- почтовый шлюз с функциями AntiVirus & AntiSpam (**rspamd-smtp-relay**) — готовый к использованию файл docker-compose.yml, описывающий модули и связи между ними (ссылка).

Ниже представлена инструкция для запуска тестовой среды:

```
# create destination SMTP server -- its task is to accept all incoming e-mail
docker run --rm --name fakemail -d -it -p 172.18.0.1:25:25  jmtalavera/fakemail-server

# create SMTP Relay with AntiSpam & AntiVirus
git clone https://github.com/mdraevich/rspamd-smtp-relay.git && cd rspamd-smtp-relay
docker compose up -d

# create mail-generator instance
docker run -v /home/matvey/Downloads/mail_examples:/home/user/mail_examples --rm -it drmatthew/mail-generator:1.0 \
  /home/user/mail_examples test@test.test 172.22.0.1:25
```

3. Автоматизация отправки писем и формирования отчета

Автоматизация отправки писем и формирования отчета выполнена с использованием Bash-скрипта.

Bash-скрипт принимает на вход список директорий, каждая из которых должна содержать набор писем в формате `*.eml`. Далее для каждой директории выполняется процесс отправки писем с использованием утилиты swaks.

Если вызов swaks завершается с `exit code=0` , письмо считается успешно переданным, иначе — письмо считается заблокированным.

> ⚠️ Перед проверкой эффективности произвольного AntiSpam решения, необходимо настроить его следующим образом:
> 1) Отключите Greylisting и механизмы ограничения запросов (Rate Limiting).
> 2) Для спам-писем настройте действие Reject (чтобы почтовый сервер возвращал SMTP код, отличный от 250).

Документация по генератору писем доступна по ссылке.

4. Результаты тестирования эффективности

Поскольку Rspamd предполагает процесс обучения байесовской и нейронной моделей, было проведено 2 последовательных испытания на одном и том же наборе спам-писем. Для испытания **rspamd-smtp-relay** использовалась конфигурация по умолчанию.

Результаты следующие:

100% Спам-экземпляры (untroubled.org)	**Кол-во писем**	**Заблокировано** - испытание №1	**Заблокировано** - испытание №2
Январь 2022	1759	**66.8%**	**67.8%**
Февраль 2022	1591	**75.9%**	**76.8%**
Март 2022	2405	**81.2%**	**82.2%**
Апрель 2022	1826	**87.1%**	**88.0%**
Май 2022	1730	**85.3%**	**86.7%**
Июнь 2022	958	**89.9%**	**91.4%**
Итого	10269	**80.5%**	**81.5%**

Легитимные письма (бизнес-переписка)	**Кол-во писем**	**Заблокировано** - испытание №1	**Заблокировано** - испытание №2
Случайный набор	336	**2.9%**	**2.9%**
Итого	336	**2.9%**	**2.9%**

Выводы

Поставленные цели были достигнуты:

1. Были автоматизированы процессы создания и эксплуатации тестовой среды для оценки эффективности AntiSpam решений.
2. Получены результаты об эффективности Rspamd, а именно:

a. спам-письма — эффективность в среднем ниже, чем у аналогов FortiMail (заблокировано не менее 92%) и VadeSecure (заблокировано не менее 87%).
b. легитимные письма — точность аналогична продукту FortiMail (заблокировано в среднем 3%) и выше, чем у VadeSecure (заблокировано в среднем 15%).

💡 1) Есть возможность «увеличить чувствительность» применяемых методов по выявлению спам-писем, что позволит с одной стороны повысить процент заблокированного спама, но с другой — увеличить количество заблокированных легитимных писем.

2) Увеличение процента заблокированных писем для испытания №2 говорит о способности инструмента обучаться на ранее полученных спам-письмах, что увеличивает эффективность решения с каждым новым спам-письмом.