

Барабанов А.Б.

## Борьба со спамом, как фактор, снижающий надежность почтовой доставки.

*Часто бывает так, что системный администратор, ограничивая прохождение нежелательной корреспонденции, становится причиной нарушения информационного обмена с деловыми партнерами обслуживаемого предприятия. Какой же путь решения данной проблемы приведет к успеху с минимальными потерями?*

Электронная почта (далее просто почта) является незаменимым и очень важным информационным каналом, используемым в бизнес-процедурах современного офиса. Будучи однажды настроенной и включенной в практическое использование, она постепенно вовлекает в свою сферу все возможные деловые контакты, в силу следующих присущих ей свойств:

- изначальная интеграция в компьютерный офисный документооборот;
- низкая стоимость и отсутствие отходов от применения;
- высокая скорость и возможность информационного обмена вне зависимости от расстояния, государственных границ и без предварительных договоренностей.

Но в силу этих же причин таким путем удобно пользоваться для несанкционированного внедрения в системы документооборота компаний информационных материалов самого разнообразного содержания и с очень низким соотношением затрат на единицу передаваемой информации. Как известно,...

### **1. Если какая-нибудь неприятность может произойти, она случается. Закон Мерфи.**

И находится очень много лиц и даже специальных организаций, которые, используя указанную возможность, внедряют в офисные системы, подключенные к почтовым каналам Интернета, всякого рода рекламные материалы, вирусные коды и троянские программы. Объем нежелательной корреспонденции, называемой спамом, растет лавинообразно, отвлекая внимание операторов компьютеров, расходуя ресурсы серверов и сужая полосу соединительных каналов. Все это делает задачу противодействия спаму одной из важнейших не только в технологическом плане, но и с точки зрения экономической эффективности функционирования офисных систем. Можно даже сказать, что борьба со спамом является чуть ли не единственной деятельностью системного администратора, которая может быть оценена как составляющая положительного баланса, то есть является в чистом виде возвратом инвестиций в информационные ресурсы.

Казалось бы, что трудного - изъять спам из входящей почты компании? Почему бы не поступить так же, как с бумажной почтой и факсимильным спамом - перебрать все и удалить лишнее. Очень простая процедура – взять, оценить, поступить согласно вердикту. Эта формула универсальна! Все системы противодействия спаму работают по такому принципу. Но есть ряд существенных отличий от ручной сортировки сообщений ортодоксальных каналов, основанных на ранее перечисленных свойствах именно электронной почты:

- из-за низкой стоимости электронной почты соотношение спама к полезной почте может быть сколь угодно высоким, а, значит, затраты на сортировку могут значительно превосходить объем полезного «выхлопа»;
- из-за естественного встраивания электронной почты в офисный документооборот фильтрация спама приводит к созданию подсистем, игнорирующих как приватный так и служебный статус электронных посланий, или даже производящих обратное перераспределение почтовых сообщений от пользователей к серверным службам

(общий почтовый архив, карантины и прочее).

- информационный обмен без первоначальных договоренностей не позволяет построить эффективный и простой автоматический фильтр, поскольку не дает возможности заранее оценить число нужных в процессе работы контактов и, соответственно, обеспечить бесперебойное прохождение только полезной почты.

Иначе говоря...

## **2. Все не так легко, как кажется. Следствие 1 из Закона Мерфи.**

Рассмотрим подробнее, какие почтовые обмены предусматриваются в бизнес-процедурах и составляют очевидный информационный поток:

- - отправление регулярной почты из офиса удаленным деловым партнерам;
- - получение регулярной почты от удаленных деловых партнеров;
- - отправление почты новым респондентам;
- - получение деловой почты от новых респондентов.

Из перечисленного оставим только то, что относится к входящим потокам. Далее, поток получения почты от деловых партнеров можно легко формализовать очевидным образом – по адресам отправителей. Получается, что единственная неизвестная составляющая – почта от новых респондентов!

А теперь, точно также попробуем охарактеризовать спам:

- получение нежелательной почты от неизвестного респондента, или внешний спам;
- неавторизованное отправление почты некоторому респонденту, или внутренний спам (тройная рассылка).

Опять же, внутренний спам, устраняемый путем внедрения антивирусных систем и запретом прямой рассылки из офисной сети, далее рассматривать не будем, и обратим внимание на спам внешний. Характер данного потока полностью совпадает с ранее описанным как «получение деловой почты от неизвестного респондента». И именно в этом заключается конфликт. Запретить получение почты от новых отправителей невозможно и, таким образом, создается возможность удаленному отправителю внедрять свой неавторизованный информационный поток, подмешивая его к очевидному информационному потоку офиса. Обратная сторона данной проблемы заключается в том, что даже малейшее вмешательство в процесс получения писем такого рода создает вероятность ошибки, которая в деловой практике может привести к потере выгодного заказчика или к недоставке важной информации иного рода. Рассмотрим два этих схожих потока подробнее.

### **2.1. Процедура получения деловой почты от нового респондента.**

Отправитель такой почты - сотрудник офиса другой компании, который, воспользовавшись известным ему почтовым адресом, произвел отправку с клиентской рабочей станции (далее компьютер, участвующий в сетевом обмене, будет именоваться хостом) через почтовый сервер офиса или почтовый сервер провайдера Интернета.

Опишем характеристики такой почты:

- в списке отправителей как минимум 2 хоста (станция и почтовый сервер);
- непосредственно передача производится с почтового сервера, настройка которого обязана удовлетворять сетевым стандартам;
- результаты процедуры отправки почты оцениваются оператором-человеком, который при неудаче имеет возможность изменить параметры почтового запроса;
- получатель такой почты полностью определен в соответствии с деловым назначением, то есть совмещение сеансов передачи подобных сообщений разным получателям маловероятно.

С точки зрения бизнес-процедур такая почта делится на два типа:

**Тип 1** – контакт с новым партнером, когда результатов ожидают и в случае неудачи используют альтернативные каналы связи;

**Тип 2** – деловая рассылка по диапазону адресов, когда отправитель, скорее всего, не будет заинтересован в повторении сеанса передачи сообщения в случае первоначальной неудачи, так как оценивает успех статистически (например, извещение о планируемой выставке или другом публичном событии).

Последний случай чрезвычайно схож с процедурами рассылки спама, что делает этот тип деловой почты очень уязвимым при применении мер противодействия спаму.

## **2.2. Процедура получения нежелательной почты от неизвестного респондента.**

Эта почта очень редко отправляется вручную оператором. Рассылка спама является экономически целесообразным действием, то есть оценивается в зависимости от объема, скорости доставки и её успешности. Поэтому используются всевозможные средства автоматизации, что, как следствие, приводит к неизбежности возникновения общих схожих черт в информационном потоке такой почты:

- используются как транзитные почтовые сервера, так и хосты с прямым подключением к Сети, то есть не всегда присутствуют 2 отправителя в списке пересылок;
- используется автомат отправки почты, который в силу большого объема адресов не имеет возможности рассматривать ответ от каждого сервера и часто допускает несоответствия сетевым стандартам;
- с одного хоста может отправляться почта, принадлежащая разным адресам отправителей;
- с одного хоста может отправляться почта на разные и, очень часто, не объединенные общими бизнес-процедурами внутренние адреса;
- объем почты от хоста-спамера зачастую превышает типичный информационный поток от хоста, управляемого оператором-человеком.

Итак, оба рассматриваемых потока поступают на вход почтовой системы. Изначально не известно, какому из них принадлежит входящее сообщение. В этом вся интрига и заключена. Потому, настало время обсудить, как конкретно производится прием и обработка почты.

## **2.3. Процедура приема почты на стороне офиса.**

Напоминаю, что обсуждается только почта от ранее неизвестных отправителей. То есть, некоторое сообщение, характер которого еще не определен, поступает на вход почтовой системы (далее сервера). На этапе соединения сервер обязан выработать решение - принимать данное сообщение или отвергнуть его. Здесь сервер имеет сведения лишь о сетевом адресе хоста-отправителя, почтовом адресе предполагаемого отправителя и почтовом адресе внутреннего получателя. Если сервер принимает решение о получении сообщения, то запускается процедура приема и сообщение поступает внутрь почтовой очереди сервера. Затем производится сортировка, и некоторый внутренний абонент получает данное сообщение в свой почтовый ящик на стороне сервера.

На этом этапе почта уже принята, но еще не прочтена оператором, которому предназначена. Будем считать, что возможны только два способа получения почты на стороне клиента – POP3 и IMAP, то есть всякие проприетарные почтовые протоколы рассматриваться не будут.

Первое, POP3 – протокол по которому все сообщения забираются почтовым клиентом и далее обрабатывается в автономном режиме.

Второе, IMAP – протокол, с помощью которого клиентская программа работает с сообщениями на стороне сервера, что позволяет отслеживать характер и последовательность манипуляций человека-оператора с почтовым контентом.

Таким образом, завершение пути почтового сообщения производится или по протоколу POP3 или по протоколу IMAP, после чего можно утверждать, что сообщение доставлено.

Здесь самое время вспомнить выше заявленную формулу – «взять, оценить, поступить согласно вердикту». В пути доставки почтового сообщения можно указать только три места, в которых возможно употребление этого славного и простого алгоритма признания почты нежелательной:

- на этапе обращения удаленного отправителя к офисному серверу – отказ в приеме;
- удаление на пути до почтового ящика пользователя – контентная фильтрация;
- путем ручной сортировки на клиентской машине.

Последнее означает, что спам был принят пользователем к прочтению. Поэтому рассматривать, как возможные альтернативы, следует лишь первые два. Даже с этой поправкой ...

### **3. Всякая работа требует больше времени, чем вы думаете. Следствие 2 из Закона Мерфи.**

Итак, выше были определены две точки по пути следования почты от отправителя к получателю, в которых целесообразна фильтрация спама. Рассмотрим их в порядке применения.

#### **3.1. Отказ в приеме нежелательной почты.**

На этом этапе сервер может получить очень немного сведений о предполагаемой почтовом сообщении:

- сетевой адрес и доменные имена хоста-отправителя;
- почтовый адрес отправителя;
- почтовый адрес получателя.

Решение о принадлежности почты к категории нежелательной должно быть принято лишь по перечисленным критериям. Достаточно ли этого для получения абсолютно правильной резолюции? Скорее всего, нет. Но этого полностью достаточно для определения той почты, которую НЕ нужно фильтровать на спам! Как правило, именно по адресу отправителя или по хосту-отправителю производится построение «белых» списков в почтовых системах.

Но перехват спама на данном этапе вместе с тем имеет очень много положительных сторон:

- отказ в приеме почтового сообщения оформляется в соответствии с сетевыми стандартами, и отправитель получает стандартное уведомление с указанием причины;
- почтовое сообщение, причисленное к спаму, не принимается, то есть не использует ресурсы сервера и не расходует трафик, что не позволяет хосту-спамеру блокировать сервер частыми или крупноразмерными посылками.

Учитывая вышесказанное, именно в этом способе противодействия спаму предпринималось наибольшее число инициатив разных компаний и общественных организаций. Это всем известный SPF (sender policy framework) [1], затем Sender ID от Microsoft [2], и, наконец, недавнее предложение от IETF почтовой аутентификации на основе DKIM [3]. Все перечисленные сетевые стандарты создавались с целью более полного представления полномочий хоста-отправителя, чтобы сервер мог квалифицированно решать вопрос о приеме или отказе в отношении конкретного почтового сообщения.

Если такой фильтр вынес неверную резолюцию, то для сообщений типа 1, которые отправляются оператором, и ответ на которые также ожидается конкретной персоной, не происходит ничего фатального, так как сервер отправителя сформирует служебное письмо с указанием причины отказа. А вот для сообщений типа 2, рассылаемых автоматически или через списки рассылки, все складывается гораздо хуже, если результаты рассылки оцениваются лишь по совокупности, а не каждая доставка в отдельности.

Иначе говоря, даже если предположить, что фильтрация на входе изначально была включена в состояние запрета всех доставок, то в самые короткие сроки связь с постоянными контрагентами можно восстановить путем настройки исключений ( «белый» список ). И все новые респонденты также будут в самые короткие сроки получать доступ к каналу доставки электронной почты.

Единственной проблемой остаются неавторизованные рассылки. Но эту проблему можно частично преодолеть путем изучения типичного входного трафика данного сервера. Тогда будут отсекается лишь рассылки, предпринимаемые вновь.

Особо надо заметить, что, как правило, подобный фильтр не включается на тотальное подавление почтового трафика. Априорно отказывается в приеме лишь тем хостам, чьи регистрационные данные в доменных серверах не соответствуют стандартам. То есть запретительная часть такого фильтра строится на основе статистики полученного спама, или, говоря иначе, по требованию пользователей.

### **3.2. Контентная фильтрация.**

На этом этапе сервер имеет полный набор сведений об отправителе, получателе и характере самого почтового сообщения, что позволяет с максимально возможной точностью определить является ли данное почтовое сообщение нежелательной рассылкой. Но вместе с тем для отправителя почтовое сообщение уже выглядит как полученное, поскольку сервер принял его к доставке. И если на этапе доставки спам-фильтрами будет принято решение о том, что сообщение воспринимается как спам, то при изъятии его из почтового трафика складывается конфликт ожиданий: отправитель считает, что сообщение получено, а получатель не видит его в своем почтовом ящике. Поэтому все системы контентной фильтрации не изымают почтовые сообщения из потока, а лишь помечают их соответствующим образом. Иначе говоря, контентная фильтрация на уровне сервера работает лишь при соответствующим образом настроенной фильтрации на уровне клиентских компьютеров.

Таким образом, контентная фильтрация является платформенно-зависимой, во-первых. Более трудоемкой – настройка сервера и программы-клиента, во-вторых. И в-третьих, возрастает вероятность потери почтового сообщения в результате ошибочной фильтрации, так как отправитель уже не извещается о резолюции фильтрующей системы. И если фильтр срабатывает на сообщениях из категории потока от новых отправителей, то практически со 100% вероятностью в такой системе недоставка будет обнаружена гораздо позже, чем в фильтрации на этапе получения. Более того, получатель будет вынужден вручную пересмотреть все сообщения, помеченные как спам, в поисках потерянного.

С другой стороны, можно утверждать, что при контентной фильтрации почтовые сообщения не теряются вовсе, так как все они оказываются у получателя тем или иным образом.

Итак, оба подхода к фильтрации не исключают ошибок. Оба подхода не идеальны, но каждый имеет свои специфические преимущества. Попробуем сделать выбор между ними, исходя из риска потери важной деловой почты, поскольку ...

#### **4. Из всех неприятностей произойдет именно та, ущерб от которой больше. Следствие 3 из Закона Мерфи.**

Предположим, сотрудник компании обратился к системному администратору с жалобой на то, что некоторое очень важное для него письмо было «съедено» почтовым сервером. Причем, он даже, как правило, не знает обстоятельств произошедшего, максимум – ему известен адрес отправителя.

##### **4.1. Отказ в приеме.**

Нам известно, что отправитель должен был получить уведомление и самостоятельно предпринять очень простые меры к исправлению ошибочной фильтрации – сообщить получателю по альтернативным каналам связи или направить письмо системному администратору. Поэтому, данная ситуация вызвана скорее не техническими проблемами, а сложностью отношений между отправителем и получателем. Например, получатель не может вторично обратиться к отправителю и попросить еще раз послать нужное письмо, а отправитель не считает нужным разбираться в тонкостях «письмохождения». Часто такое происходит из-за чрезмерной заинтересованности получателя, что характерно для коммерческих компаний, в которых маркетинговые отделы играют определяющую роль в отношениях с системными администраторами. В тех компаниях, где основным «заказчиком» информационных служб являются финансовые отделы, более склонны мириться с такими неудобствами ради сокращения потока нежелательной почты.

##### **4.2. Контентная фильтрация.**

Так как в данном методе фильтрации письма не изымаются из потока, а лишь только метятся, то теперь нам уже точно известно, что получатель получил данное письмо. Значит, он его не смог найти по каким-то причинам. Это дает основания сразу «завернуть» подобную жалобу. Если цель системного администратора в сокращении своих забот любой ценой, то здесь ему представляется прекрасный случай переложить труды по поиску пропавшего сообщения на озабоченного сотрудника.

В том же случае, когда системный администратор в самом деле попытается решить данную проблему, не заставляя несчастного менеджера перечитывать все спамовые письма в поисках случайно помеченного, то действовать ему придется точно теми же методами, что и в первом варианте. То есть, сначала узнать адрес отправителя, затем найти в протоколах путь обработки данного сообщения и так выйти на то место, где оно хранится.

В чем же разница? Кажется, что вторая ситуация гораздо легче. Во-первых, поиск происходит «в две руки». Во-вторых, получатель, потерявший письмо, «сохраняет лицо» перед отправителем, так как ему не приходится просить о повторной отправке. Но на самом деле, если учесть фактор времени, то получается, что потеря письма в случае контентной фильтрации будет обнаружена существенно позднее. Если почтовому сообщению отказано в приеме, то отправитель получает служебную резолюцию в течении нескольких минут. А вот если письмо затерялось в недрах спамного карантина, то узнать о пропаже можно лишь после того как истечет разумное время ожидания, обусловленное темпами тех бизнес-процедур, которые будут нарушены из-за недоставки данного сообщения. Иначе говоря, письмо потерянное в результате ошибки контентной фильтрации в силу особенностей данного метода может быть восстановлено только после нанесения максимального вреда от его недоставки.

Следующим критерием сравнения выберем эффективность применения каждого из методов, так как...

#### **5. Если четыре причины возможных неприятностей заранее устранимы, то**

## **всегда найдется пятая. Следствие 4 из Закона Мерфи.**

Попробуем предположить модель рассылки спама. Пусть разнородный спам, представленный некоторым множеством сообщений, рассылается с определенного множества хостов. В такой ситуации, «отказ в приеме» будет фильтровать хосты спамерской сети, а вот «контентная фильтрация» будет наполнять фильтровую базу дайжестами всех сообщений. Задумаемся, насколько эффективно противодействие спаму в каждой из точек фильтрации. Для этого зададимся вопросом, может ли один спамерский хост рассылать спам разного типа. Безусловно, да! Далее, задумаемся, какой процесс протекает более динамично – изменение числа спамерских хостов или модификация набора нежелательной корреспонденции. Опять, ответом будет, что, скорее всего, спамерская сеть является более устойчивым множеством, чем совокупность нежелательной электронной корреспонденции. Все это дает основания считать применение методов «отказа в приеме» более эффективными по сравнению с методами «контентной фильтрации».

Теперь проанализируем возможность ложного срабатывания вопреки действующим настройкам. Применительно к «отказу в приеме», который построен на списке заблокированных и/или разрешенных хостов, это звучит как – а может ли хост из участвующего в деловой переписке превратиться в спамерский, и наоборот. В первом случае – да, так как это равносильно тому, что в офисе удаленного респондента заведется затрояненный компьютер, включенный в сеть спамерской рассылки. Будет ли это основанием для изменения статуса данного отправителя? Вряд ли! Скорее всего, надо будет своевременно информировать администрацию удаленного офиса о проблемах. Во втором случае сложнее, так как это значит, что кто-то из заблокированной сети пытается вступить в деловую переписку и получает в ответ обидное послание о блокировке как спамера. Вероятность такого события не исключена. И практика показывает, что именно с такими проблемами приходится чаще всего сталкиваться при обслуживании информационной системы коммерческого предприятия.

Если же подобное случается в процессе контентной фильтрации, то ошибка подобного рода может вызываться двумя причинами:

- получен спам от делового респондента;
- получено письмо, ошибочно принятое за спам по совокупности его параметров.

Здесь подразумевается, что источник обсуждаемого сообщения не входит в «белый список», то есть информационный поток от него подвергается обычной процедуре фильтрации. Тогда, ситуация с сообщениями первого рода полностью аналогична выше описанному случаю с затрояненным компьютером в сети отправителя. А вот письма, совпавшие по сигнатуре со спамом, составят статистику риска «by design» для данного метода фильтрации. Причем, этот риск не устраним способами, положенными в основу контентных методов фильтрации. Для преодоления данной проблемы и защиты деловой почты важных контрагентов используются не специальным образом вычисленные сигнатуры «правильных» писем, а все те же «белые» списки, состоящие из адресов отправителей – почтовых и сетевых.

Еще хуже будет результат сравнения двух этих способов противодействия спаму по ресурсоемкости реализации. Что не удивительно, поскольку ...

## **6. Предоставленные самим себе события имеют тенденцию развиваться от плохого к худшему. Следствие 5 из Закона Мерфи.**

В реальном почтовом сервере всегда существуют правила ограничения приема, то есть, стандартно настроенный сервер уже имеет некоторый фильтр на входе, принимающий от определенной категории хостов всю почту и отказывающий остальным в определенной части своего сервиса. Это неизбежно. Почтовый сервер, не отказывающий никому, называется открытым релеем. Кроме того, на практике всегда встречаются контрагенты, ошибающиеся

адресом. Иначе говоря, штатные уведомления о трудностях доставки обязательно должны обрабатываться на стороне отправителя и, аналогично, принимаются по альтернативным каналам на стороне получателя для устранения причины недоставки сообщений электронной почты. Подобное поведение почтовых систем закреплено стандартами. Значит, внедрение любой системы фильтрации, работающей по принципу «отказ в приеме» не потребует никаких переделок и доработок как на стороне отправителя так и на стороне получателя сообщений, но заставит более внимательно и оперативно обрабатывать информацию, поступающую по альтернативным каналам для устранения коллизий ошибочной фильтрации.

Внедрение контентной фильтрации потребует увеличения мощностей почтового сервера в расчете на работу фильтрующих программ и создания дополнительной емкости средств хранения данных для спам-карантина. Если стоимость памяти падает, то вот первый нагрузочный параметр явно не поспевает за темпами роста пересылаемых объемов данных, вызванных увеличением пропускной полосы каналов связи. Почтовое сообщение, включающее многомегабайтный архив, может обрабатываться в системе контентной фильтрации десятки минут.

Но и это не все! Кроме того, понадобится соответствующая настройка клиентских компьютеров. Если все письма после контентной фильтрации будут смешаны в одной папке «Входящие сообщения», тогда оператору придется самостоятельно принимать решение о чтении или игнорировании каждого сообщения на основании рекомендации фильтра. Значит, или почтовый клиент на компьютере получателя должен поддерживать сортировку по некоторому ключевому признаку на спам и рациональную почту, или такая сортировка должна производиться на стороне почтового сервера, а клиентская программа должна использовать протоколы для удаленной работы с сообщениями, хранящимися на стороне сервера, например IMAP. Хотя в тексте используется союз «или», но на практике сортировка на стороне клиента не имеет смысла, так как серверные компоненты системы контентной фильтрации лишаются простой возможности получения от пользователя подтверждения или опровержения правильности применяемых политик исходя из перемещений опротестованных сообщений по почтовым папкам. Например, если некоторое нежелательное сообщение спам-фильтром пропущено, то будучи помещенным в специально выделенную папку, оно сканируется повторно и его дайджест включается в поисковую базу. И наоборот, изъятое из папки со спамом деловое сообщение сканируется, и адрес его отправителя помещается в «белый список». Конечно, подобный функционал можно решить за счет внутренних пересылок (форвардов) на служебные адреса, но такое действие будет удачным лишь в случае корректной пересылки самим почтовым клиентом. И все равно, таким путем не решается главный вопрос – статистический анализ спама после отстоя его в клиентском карантине, после изъятия ошибочных сообщений и размещения в нем пропущенных и проч.

Тогда зачем заниматься контентной фильтрацией? Может выкинуть ее вовсе? В сущности сама идея, что можно априорно составить такую базу сигнатур почтовых сообщений или такой набор эвристических методов, которые исчерпывающе дадут ответ на вопрос принадлежности некоторой кодовой последовательности, называемой электронным письмом, к категории нежелательной корреспонденции чрезвычайно близка к теме создания вечных двигателей, философских камней и прочих почти магических предметов. Можно даже утверждать, что все (ВСЕ, я не шучу!) коммерческие продукты, основанные на контентной фильтрации спама, являются в чем-то мошенническими.

И окончательно контентная фильтрация проигрывает по показателю трудоемкости внедрения, потому что ...

**7. Как только вы принимаетесь делать какую-то работу, находится другая, которую надо сделать еще раньше. Следствие 6 из Закона Мерфи.**

Рассмотрим процедуру внедрения перечисленных методов фильтрации спама.

### **7.1. Отказ в приеме.**

Есть два пути настройки подобной блокировки. Первый заключается в том, чтобы воспользоваться публичными службами, коллекционирующими адреса хостов, замеченных в рассылке спама, так называемые RBL (realtime blackhole list). Это самый простой и неэффективный способ. Но именно с него, как правило, начинают все. Настройка занимает 15 минут, затем долго обрабатываются жалобы получателей на отказ в приеме деловой почты, что выливается в составление корпоративного «белого списка». Самое плохое, что «белый список» составляется не методично, а лишь в той степени как используемые провайдеры RBL перекрывают деловую корреспонденцию. А так как RBL пополняются динамически, то «борьба» за доступ деловой почты рискует затянуться до полного истощения сил одной из противодействующих сторон «сисадмин vs заказчики его работодателя». Не надо быть очень прозорливым, чтобы предсказать исход – среди используемых поставщиков RBL останутся только самые консервативные вроде тех, что описывают сети dial-up, что сведет эффективность данной фильтрации к нулю!

Второй способ заключается в составлении как «белого», так и «черного» списка фильтрации с учетом особенностей информационных связей конкретной компании. Здесь просто раздолье для всяких стратегий, основанных на изучении почтовых протоколов. В качестве примера можно привести решение, предложенное Максимом Чирковым [4] для почтовых серверов с большим числом пользователей. В основу этого метода положено предположение о том, что спам поступает сразу многим пользователям одного сервера, не связанным никакой общей бизнес схемой. Это позволяет блокировать распространителей спама сразу после первой рассылки. Возможен и другой подход к составлению списков фильтрации. Главное в этом методе то, что изначально блокировка отключена. Затем в течение некоторого периода происходит автоматическое накопление информации, на основе которой составляются списки блокировки. Собственно тем все и сказано – автоматически! Хотя, над самим скриптом сбора данных возможно и придется предварительно потрудиться. Отрицательным свойством можно считать постепенность внедрения – эффект фильтрации проявляется лишь со временем, в темпе пополнения «белых» и «черных» списков, и то, что такую систему надо постоянно поддерживать в работоспособном состоянии – предоставленные самим себе фильтрующие списки постепенно теряют актуальность.

### **7.2. Контентная фильтрация.**

Данный метод подобен «волшебной пилюле» - эффект проявляется практически моментально! Но вот внедрение требует значительных переделок, как почтовой системы, так и клиентских рабочих мест. Сам фильтр придется встроить в путь доставки почты. Причем это не пассивная таблица со списком фильтрации, а программная компонента, требующая собственных ресурсов, а значит, её внедрение нужно учесть в характеристиках серверного оборудования. Если масштабирование табличной фильтрации заключается в изменении размеров рабочих устройств памяти, то масштабирование активных программных компонентов (а в процессе внедрения может появиться такая необходимость) это уже вопрос иной ценовой категории. Затем, надо предусмотреть процедуру сортировки на основании резолюции фильтра. И если такую сортировку не получится выполнить в пределах IMAP директорий на сервере, то придется перенастраивать ВСЕ клиентские рабочие места!

Итак, на одной чаше весов большие хлопоты сисадмина, а на другой обещания одним махом избавиться от спама! А уж если такую систему предлагают как коммерческий продукт, то у руководства компании создается иллюзорная уверенность в том, что проблема, которая их так раздражает, решается и быстро, и не дорого! Причем, если позитивные результаты работы фильтрации на входе никак не видны, а негативные сразу заметны, то в контентной фильтрации

все с точностью до наоборот. Здесь позитивные качества сразу заметны по наполнению директорий со спамом, а вот негативные и обнаруживаются не так быстро, да и количественно тонут, «тонут» буквально, внутри тех же спамовых карантинных.

Но, как уже было заявлено, ни одна из систем противодействия спаму не гарантирует безошибочной и правильной фильтрации. Иначе говоря...

## **8. Всякое решение плодит новые проблемы. Следствие 7 из Закона Мерфи.**

Предположим, все хлопоты позади. Фильтрация спама настроена и работает. И тут выясняется, что необходимо настроить дифференциальное применение. То есть единообразное применение фильтра ко всему входному потоку неэффективно! Например, почтовыми стандартами требуется без всяких ограничений принимать почту на служебные адреса, такие как `postmaster@` [5]. Далее, обычно так же «открываются» адреса `abuse@`, `webmaster@` и подобные. Кроме того, характер деятельности многих предприятий требует установления ряда совершенно открытых и публично доступных адресов. Очень часто это `info@`, или даже `mail@`. Почта по таким адресам должна проходить во чтобы то ни стало!

А может быть и наоборот. Многие руководители очень не любят, когда в приходящих к ним спамовых письмах они именуется полной уважительной формой (ФИО) в рекламной листовке предметов интимного досуга. И поскольку никакое объяснение, что информация о руководителях предприятий публично доступна, а значит, с легкостью может попасть в спамерскую базу, не помогает, то единственный выход – максимально избавить такие почтовые учетные записи от нежелательных посланий.

Таким образом, реальное применение фильтрующей системы невозможно без использования набора политик, оговаривающих требуемый уровень защиты от спама. На самом примитивном уровне это все те же «белые» списки, но уже не отправителей, а получателей, например, для Postfix [6]. Но для построения разветвленной и согласованной системы политик, управляющих фильтрацией и доступом к электронной почте, конечно же требуются специальные программные компоненты. Для того же Postfix-а их можно найти на странице [7].

Использование таких систем делает модель фильтрации не столь прямолинейной. Например, если применять `policy`-демоны на входе `smtpd` сервера Postfix, то можно проверять не только минимальный набор признаков `smtp`-соединения (хост, отправитель, получатель), но и SPF-таги отправителя. Кроме того, такая компонента может отслеживать статистику соединений, что используется для построения политик `greylist`, т.е. разрешения доступа к почте только с определенного числа попыток отправителя.

Вот, теперь задумаемся, как можно совместить политики с выбранным методом размещения спам-фильтра.

### **8.1. Отказ в приеме.**

Политика строится на основе все лишь трех упомянутых признаков сообщения – хост, отправитель, получатель. Например, для `greylist`, триада, характеризующая отдельную отправку, записывается в базу, что и позволяет строить историю взаимодействия с хостом, отправляющем сообщения. В проекте `apolicy` [8], на основании полей триады строятся традиционные для межсетевых экранов правила управления доступом. Обращаю внимание, в политиках используется ВСЯ информация, которой располагает почтовый сервер в данной точке.

### **8.2. Контентная фильтрация.**

А вот тут получаем парадокс! Механизм самой контентной фильтрации в policy никак не задействуется. Политики применяются ДО исполнения процедуры фильтрации. Собственно они ей чужды, поскольку в самом простом случае политики реализуются как исключения, построенные на все тех же «белых» или «черных» списках. В таких списках используются те же характеристики сообщения, что и выше – хост, отправитель, получатель. Но ведь резолюция контентного фильтра строится на основе сигнатуры, и если нельзя использовать такую сигнатуру в системе построения политики, то и получается, что политика в контентной фильтрации является внешней искусственной надстройкой. Да и как можно заранее указать сигнатуру уникального сообщения, если его прием еще только предполагается.

Иначе говоря, построение политик фильтрации спама для систем, использующих «отказ в приеме», является гармоничным и логически обоснованным, что не скажешь о системах контентной фильтрации.

## 9. Выводы.

Вот и пришло время сделать выводы, как построить фильтрацию спама и не создать проблем для компании-нанимателя. Поскольку «фильтрация на входе», так или иначе, присутствует в любом почтовом сервере, то целесообразно использовать именно этот механизм. Даже более, его просто невозможно не использовать! Следует изучить структуру входящих почтовых потоков и выделить ряд «белых» коммерческих адресов, для которых гарантировать доставку до почтового ящика. Обычно такие адреса публикуются на вебсайте компании как контактные. Все остальные контакты должны подвергаться фильтрации. Очень хорошие результаты дает применение greylist-а. Фактически, без всяких специальных запретов нежелательная корреспонденция сокращается на 90% минимум! Использование контентной фильтрации, как дополнительной меры, тоже не помешает, так как в условиях минимального срабатывания она даст нужный эффект – даже если некоторый спам прорвется, он будет помещен в локальный карантин пользовательского ящика (например, папка .Spam в IMAP) и не создаст проблем в поиске нужного письма в случае ложного срабатывания. А вот применение такой фильтрации для группы открытых публичных адресов просто необходимо. Опять же, поскольку данных адресов не много, то и работы по настройке клиентских систем на срабатывание по признакам контентной фильтрации будет не много. Если использовать контентную фильтрацию только для «открытых» адресов, то её применение может носить не глобальный характер (не server-wide, а, например, через gmail), а, значит, использование программ фильтрации не создаст большой нагрузки на оборудование сервера.

Так как ни одно из технических решений, даже самое безупречное, не избавит от возможных проблем, то следует позаботиться о снижении негативных последствий от их возникновения. Во-первых, надо довести специально, или сделать доступной для интересующихся, политику фильтрации, используемую на сервере предприятия. Иначе, возможно бесконечное число жалоб, вызванных элементарным непониманием происходящего. Во-вторых, необходимо создать средство, позволяющее пользователям выяснить, как успешно была отправлена их почта (такой же фильтр может быть и на удаленной стороне!). В противном случае, они не смогут уверенно общаться с контрагентами, скрывающими факт получения корреспонденции. В-третьих, пользователи должны иметь возможность самостоятельно проверить, что удаленная сторона (например, важный контактер, отправивший очень важную почту) предпринимала попытки отправки, и узнать их результат (доставка, отказ и проч.). Если это не сделать, то на системного администратора будут сыпаться бесконечные претензии о вероятно и несуществующих проблемах. И в-четвертых, очень полезно предоставить пользователям возможность управления процедурой фильтрации – вносить собственные «белые» адреса или наоборот блокировать нежелательные. Последнее надо сделать максимально корректно в иерархии применения фильтров, чтобы не позволить менять глобальные параметры системы.

Все это сделает сотрудников предприятия максимально информированными о процессах

внутри почтовой системы и частично перенесет на них ответственность за возможную недоставку почты. Например, очень часто проблемы возникают из-за «на слух» надиктованных адресов или ошибочно «забитых» в почтовую программу, как адрес ответа. Система отправляет отправителю сообщение «такого пользователя нет», а он раздражается – я не могу вам отправить письмо из-за того, что ваш сервер его отклоняет. Надо сделать все возможное чтобы у сотрудников предприятия была информация для принятия мотивированных решений в отношении почтовых проблем. Только так можно превратить их из капризных и недалеких ламеров в тех клиентов, которые ценят работу своих системных администраторов. Ценят буквально ... ;)

### **Использованные ссылки:**

1. Сайт проекта SPF.

<http://www.openspf.org>

2. Страница проекта Sender ID.

<http://www.microsoft.com/mscorp/safety/technologies/senderid/default.msp>

3. Сайт проекта DKIM.

<http://www.dkim.org>

4. Система адаптивного блокирования спама после первого факта рассылки. Максим Чирков.

[http://www.opennet.ru/dev/spam\\_check/readme.txt](http://www.opennet.ru/dev/spam_check/readme.txt)

5. Обоснование необходимости безусловного приема почты по служебным адресам.

<http://rfc-ignorant.org/rfcs/rfc2821.php>

6. Владимир Камарзин. Реализация беспрепятственного доступа к адресу postmaster@.

<http://www.freesource.info/wiki/Dokumentacija/Postfix/antispam/postmaster&>

7. Дополнительное программное обеспечение к почтовому серверу Postfix.

<http://www.postfix.org/addon.html>

8. ACL Policy Daemon for Postfix (apolicy).

<http://home.gna.org/apolicy/>