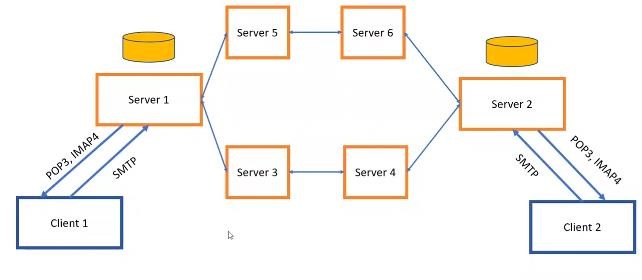
Электронная почта сама по себе не маршрутизироваться, потому что она маршрутизируется средствами TCP/IP, то есть почтовый сервер получив запрос от клиента на посылку электронной почты сам соединяется с целевым сервером.

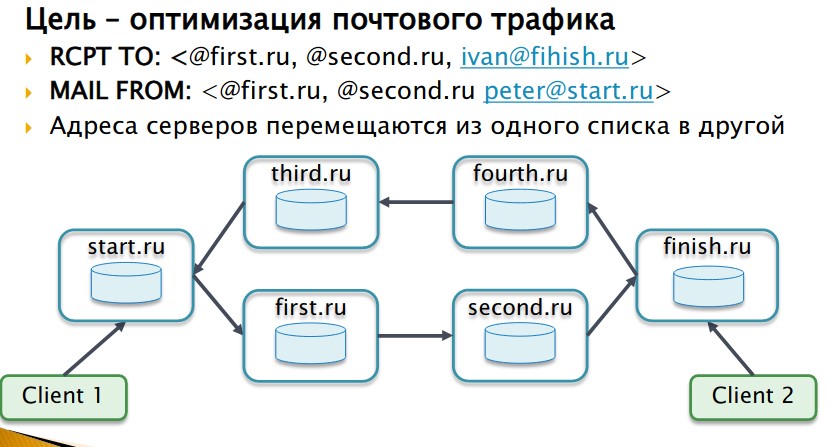
Почтовый **трафик** может быть довольно таки **большой** и может существенно **забивать** канал. Была одна очень неприятная история в начале 2000 х годов с компанией P&G компания одна из крупнейших в мире компаний. у них по моему сейчас работает чуть ли не 900 тысяч человек и гигантская корпоративная сеть даже в начале 2000 х это была огромная компания сотни тысяч сотрудников. история была следующая что какая то секретарша в какой то момент решила по корпоративной почте послать письмо о том что у нее будет вечеринка но и по ошибке вместо того чтобы послать письмо десяти своим подругам она воспользовалась адресом все клиенты P&G. Ну и несколько сотен тысяч человек получили письмо личного содержания о том что будет вечеринка. Многие возмутились ответили секретарше что вообще то недопустимо использовать адрес рассылки общей для посылки писем личного характера, чтобы больше она так не делала. причем примерно %10 из тех кто и ответил они точно так же как она ошиблась и ответили на общую рассылку. И после этого была еще. В ответ на это возмущенные пользовались стали писать что не нужно использовать письма для личных целей корпоративную почту и тоже примерно 10 процентов из них ответила кнопка ответить всем в том числе в эту самую рассылку. В итоге в корпоративной почте P&G одновременно существовала. Миллиарды писем одинакового содержания которые заняли весь трафик P&G и все серверы только и занимались тем что передавали электронную почту между собой. это на несколько дней вывела из строя все коммуникации компании P&G. у нее был полнейший электронный коллапс. пока администраторы не смогли собственную рассылку этих миллиардов писем остановить и очистить почтовый сервер то есть даже такое простое не преднамеренное действие действие если оно сделано не слишком умно может привести к тому что электронный трафик электронной почты может забить основные каналы.

По этому поводу необходимы механизмы как это самый трафиком управлять чтобы можно было **трафик электронной почты пускать отдельно по тем каналам которым хотелось бы его послать.** ну и подход здесь примерно следующий:



В одну сторону почта ходит. по такому маршруту в другую сторону почта ходит по такому маршруту и специальный маршрут выбран таким образом чтобы например не забивать обычной магистрали а использовать для них например какие то дешевые но может быть не самые быстрые каналы. То есть подразумевается что мы имеем в данном случае два маршрута один маршрут из сервера 1 в 3 в 4 во 2 и один маршрут из 2го сервера в 6й в 5 в 1ый, в этом случае мы можем разделить трафики и управлять им в зависимости от корпоративной политики, в зависимости от типологии сетей и так далее и так далее.

Как это технически сделать.



Есть расширение SMTP которая называется SMTP маршрутизация, сделано следующим образом, изменяется формат адреса назначения, если раньше в **RCPT TO** у нас был только адрес назначения то теперь перед ним через запятую добавляются список почтовых сервисов через который мы хотим послать почту. например здесь написано что для того чтобы посылать в домен gmail, мы сначала должны послать в сервер first, а он должен послать сервер second. и обратная система **MAIL FROM** будет постепенно помещаться те адреса через которые уже почта прошла. то есть поместив изначально нужный список адресов команда **RCPT TO** мы управляем как будет идти почта в одну сторону и если специально сконфигурировано не будет, то обратно она пойдет ровно в обратном направлении потому что в **MAIL FROM** указаны все те маршруты через которые данное почтовое сообщение прошло.

При работе с электронной почтой не всегда это является проблемой. иногда трафик почты по умолчанию так же как делается и в обычном маршрутизации. но принципиальная возможность есть. Протокол поддерживать есть необходимость маршрутизации почты есть, то это можно реализовать. вот наверное основные расширения протокола SMTP.

**Вопрос** Откуда взялся спам. почему он вдруг появился?

**Ответ** хотят делать дешевую рекламу. То есть основная причина спама в том что технологии электронной почте позволяют очень дешево рассылать рекламу очень большому количеству людей.

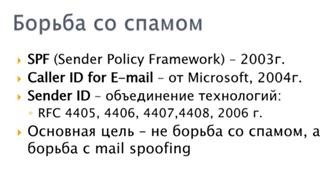
Например бумажный спам в почтовый ящик кинули. его нужно напечатать он стоит денег пусть небольших но стоит денег. например какая нибудь компания захочет сделать рассылку. например компания по доставке пиццы захочет например по всему петербургу разослать такие сообщения бумажные. ей придется напечатать пять миллионов этих самых баннеров и это будет на самом деле очень дорого. При этом же если они делают спам рассылку - то это очень дешево.

**Вопрос** как можно от спама защититься?

Идея следующая давайте мы заставим спамера выполнять ресурс съемки вычисления. например мы будем каждый кто хочет послать мне письмо он должен будет вычислить какую то сложную функцию или выполнить какую то другую сложную операцию которая кушает определенное время его почтового сервиса. Например **каждый кто будет посылать письмо мне должен например потратить например одну секунду машинного времени** например современного компьютера. Стандартная рассылка 10 миллионов адресов а где я возьму ресурсов на то чтобы сделать 10 миллионов секундах расчетов. это дикая вычислительная мощность требуется для этого всего. поэтому для меня обычного пользователя это абсолютно бесплатная вещь. мой почтовый сервер или мой клиент например вычисляют какую то функцию. это заметили только тогда когда рассылки массовые рассылки на сотни тысяч миллионы адресов и в этот момент выяснится что эта операция перестанет работать. Сейчас разрабатывается определенное количество протоколов. Единственное что начинает работать только в тот момент когда **весь мир** переходит на такие почтовые сервисы которые требуют выполнения этих самых тяжелых операций. сейчас же используются другие способы борьбы со спамом и мы на следующей лекции с вами об этом поговорим.

Серьезно озаботились борьбой со спамом примерно в конце прошлого - начале этого века, когда это стало серьезной проблемой и было одновременно разработано несколько технологий которые серьезно помогают в этом.

Первая технология это **SPF**(Sender Policy Framework), была выпущена в 2003-ем году, и почти одновременно Microsoft предложила свою похожую технологию, которая называется **Caller ID for E-mail** (2004 г). В итоге, между собой смогли договориться и объединенная технология, которая называется **Sender ID**,вошла в стандарты Internet. В 2006 году вышла серия RFC 4400, четыре RFC, и это сейчас один из новых методов борьбы со спамом. На самом деле, не совсем со спамом, а с **mail spoofing**-ом.



Что такое mail spoofing. **Mail spoofing** – это когда кто-то использует какие то разные слабости в протоколе, в настройке серверов или еще в чем то и посылает вам письма от чьего-то имени, то есть это подделка исходного адреса письма. Для чего это делается? Если бы спамер присылал почту все время с одного адреса, то мы бы просто отфильтровали этот адрес и спам бы не получали. Это было бы слишком просто. Поэтому спамерами было придумано, что спам приходит каждый раз с разных адресов и сами эти адреса могут быть от разных почтовых серверов, разных ящиков и так далее. И просто так простым фильтром это невозможно отфильтровать. И все это работает, потому что из-за особенностей организации протокола SMTP, а также из-за особенностей организации самих серверов, можно относительно несложно послать письма от чужого имени. Когда в протоколе SMTP не было аутентификации, это было совсем просто. Сейчас, когда есть аутентификация, это сложнее, но тоже возможно сделать.

В итоге, что из себя представляет эта технология (**Sender ID**), в чем же заложена идея. Когда мы получаем спамовские письма, мы их получаем, как письма с каких-то несуществующих или непонятных адресов, а может быть даже с существующих, но с известных доменов.

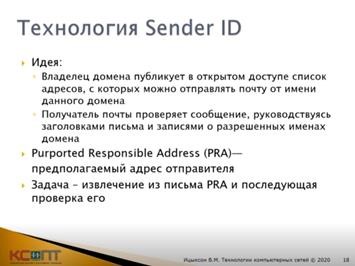
Идея в следующем - а давайте-ка мы сделаем некий сервис, где владелец домена публикует в открытом доступе те адреса, с помощью которых можно отправлять почту от имени данного домена.

Например, есть сервис mail.ru. Если кто-то хочет отправлять почту с сервиса mail.ru, он должен отправлять её исключительно с нескольких IP адресов. Например, у Mail.ru есть официальный сервис, с помощью которого можно посылать почту, он называется, по-моему, SMTP mail.ru или ещё MXS mail.ru, ну несколько всего адресов, которые обслуживают домен mail.ru. И вот только с них можно будет посылать почту от имени пользователя mail.ru.

Чем это упростит жизнь нам? Тем, что у нас не получится так, что мы можем воспользоваться особенностями протокола SMTP, тем, что отсылающий сервер никак не связан с доменом почтовым, от которого идёт отсылка. Если вы вспомните протокол SMTP, мы пишем **MAIL FROM**, **RCPT TO** и всё. И этот MAIL FROM может быть, собственно, с любого домена. И вот как раз так с этой особенностью хочется бороться.

При этом, получатель, получив сообщение, по заголовкам письма смотрит что, например, письмо с обратным адресом с mail.ru пришло с таких-то IP адресов и смотрит, разрешенные эти адреса были или нет. Если адреса разрешенные, то письма мы получаем, если нет, то не получаем.

Соответственно, появляется такая сущность как PRA (Purported Responsible Address) - предполагаемый адрес отправителя, ну и задача почтового сервера - это извлечение из текста письма этого самого PRA, и последующая проверка его на валидность.

● 

Вопрос следующий, а где хранить вот эти вот разрешенные домены? У нас, например, почтовый сервер, тот же mail.ru, где нам хранить IP адреса тех серверов, с которых мы считаем, что валидно принимать почту от нас?

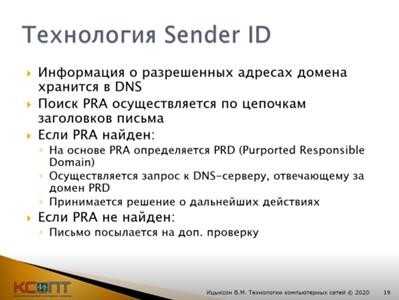
Где-то в облаке возможно, в пространстве пользователя?

Видите, в чём дело, речь идёт не о пользователе, а о сервере, то есть мы не для одного пользователя говорим, а для всего почтового домена, в данном случае mail.ru. И мы должны хранить это в таком месте, чтобы это место было доступно всем другим почтовым сервисам. Не знаете ли вы такой базы данных, которая доступна всем? Для этого используются записи **DNS**. Либо используют запись TXT, либо используются некоторые другие записи, там разные технологии существуют. Собственно, поиск PRA осуществляется по цепочкам заголовков письма. Если вы посмотрите какое-нибудь письмо, которое прошло через несколько почтовых серверов, там есть такое понятие “цепочки заголовков”, через которое полетел этот сервер (пролетело письмо?). По ним, собственно, вычисляется исходный сервер, с которого было переслано сообщение, и дальше через DNS это проверяется.

Как это делается? Во первых, не всегда этот PRA можно найти по цепочкам заголовков, но если нашли, то определяется так называемый PRD - домен предполагаемый, дальше осуществляется запрос к DNS-серверу, отвечающему за домен PRD, ну и в зависимости от ответа принимается решение о дальнейших действиях.

Если же не найден, тут есть две политики:

* жесткая политика - это мы просто это письмо игнорируем
* не жесткая – отсылаем письмо на дополнительную проверку на спам.



Ну и надо сказать что технологии эти действуют. Например, компания Microsoft уже много лет их использует, и вы не сможете послать письмо на почтовый сервер DNS, если ваш почтовый сервер не поддерживает технологию Sender ID. Вы просто не сможете им послать письмо.

Ну и это очень сильно уменьшает для пользователей поток спама. Технология стандартизирована, есть RFC, и возможность настройки технологии Sender ID есть сейчас во всех современных почтовых серверах, поэтому если вдруг с вашего почтового сервера кто-то перестал принимать почту, посмотрите в чём произошла проблема. Если действительно дело в том, что там сообщение, что не поддерживается Sender ID, надо его настроить у себя, настроить в DNS правильные реквизиты серверов и этим помогаете вашим партнёрам работать с вами не получая спам.

Как было сказано, технология Sender ID служит для борьбы с mail spoofing-ом, а для борьбы с самим спамом используются спам фильтры.

Обычно это довольно-таки интеллектуальные системы, сделанные с помощью технологии машинного обучения, которая использует одно большое преимущество почтовых серверов, особенно публичных почтовых серверов. В чем их преимущество, этих серверов, почему на них легко сделать фильтрацию от спама? Почему на вашем корпоративном сервере сложнее сделать фильтрацию от спама, чем на публичном сервере? Например, на яндексе сделать фильтрацию спама проще, чем на каком-нибудь почтовом сервере какой-нибудь небольшой фирмы.

Может быть потому что больше данных приходит?

Естественно, там ситуация очень простая. Спам, он чем хорош, одинаковое рекламное письмо отсылается огромному числу пользователей. Ну, предположим, у нас пять миллионов подписчиков в спам рассылке, предположим, из них полмиллиона попали на сервер яндекса и яндекс может проанализировав то, что одинаковое письмо или очень похоже письмо попало огромному количеству пользователей, он может с гораздо большей долей вероятности определить, что это спам. Ну а соответственно получив информацию, что одно и то же письмо или не сильно модифицированное письмо пришло большой группе, он может сделать по нему так называемый **шингл** - это слепок письма, по которому может находиться не только этот спам, но и его клоны. Спамеры они хитрые, они при рассылке не всем рассылают одинаковые письма, они рассылают похожие письма именно для того, чтобы фильтры не работали, но доступ к огромной базе пользователей, который есть в таких больших публичных почтовых серверах позволяет очень хорошо настраивать эти спам фильтры.