

Shape-NAT-Netflow в больших сетях. Часть 2

Автор: технический директор ООО ["Ниеншанц-Хоум"](#)

[Первая часть](#)

Netflow

В настоящее время исторически сложилось так, что большинство биллингов принимают информацию о трафике пользователей по протоколу Netflow v5. Практически любой биллинг, кроме жестко ориентированных на диалог-пользователей, имеет в своем составе netflow-коллектор. Органы государственной безопасности требуют наличия у провайдера полной базы данных с Netflow за три года.

Требование к биллингу по поддержке Netflow так же зачастую выставляют сами провайдеры, которые вводят "классы трафика" (трафик на яндекс, трафик на локальные ресурсы, трафик на серверы Google, трафик в интернет). При использовании PC-серверов как терминаторов для VPN-соединений (популярный в России и странах СНГ метод авторизации пользователей), радиус-аккаунтинг с таких терминаторов не несет в себе информации о классе трафика, что при повальной безграмотности пользователей несет еще и дополнительную нагрузку на отделы по работе с клиентами ("почему у вас написано, что я скачал так много трафика и вы меня отключили от интернета, я вчера только новости на сайте посмотрел и почту на мыл.ру, а потом скачал с Васи из локальной сети 25 порнофильмов").

На самом деле, любому биллингу с поддержкой классов трафика, в упрощенном виде достаточно информации "кто - какой класс трафика - сколько скачал - сколько закачал". В случае применения тарификации трафика в зависимости от времени суток ("ночью в два раза дешевле"), необходимо добавить механизм настройки частоты отправки таких данных с сетевых устройств. Понятно, что такая информация не обязана быть в формате netflow, но стандартов на протокол такого плана просто нет. В случае использования больших железных терминаторов пользователей (BRAS от Juniper, Huawei, ISG от Cisco) для радиус-аккаунтинга можно описать несколько классов трафика, которые будут передаваться отдельно. Каждый класс задается определенным ACL на BRAS'e, а подсчет по "закачано - скачано" ведет как раз BRAS.

Но если у вас есть тарифы по трафику или требование от органов государственной безопасности о наличии трехлетней детальной базы обо всем прошедшем трафике (кто-когда-куда-откуда-сколько), то от полного детального Netflow (или его аналогов) не уйти. Таким образом, комплексная задача об учете трафика разбивается на две подзадачи: передать информацию в биллинг и сохранить информацию для хранения.

PC-серверы

На PC-серверах задача о получении Netflow решается достаточно просто. Есть как готовые решения, типа Vyatta или Mikrotik, есть решения, базирующиеся на используемых файрволах (pflog для FreeBSD, например). Также на серверах можно сразу производить агрегирование данных, необязательно использовать Netflow, зачастую для решения как задачи передачи информации в биллинг вида "кто-куда-откуда-сколько-когда", так и детализации для органов, достаточно снимать 1 раз в 5-10 минут ip accounting. При использовании Mikrotik, например, для снятия Netflow, на скоростях от 500 мбит/сек, может не хватать максимального количества одновременно отслеживаемых Netflow-сессий (там это число ограничено 512K).

Софт-маршрутизаторы

При использовании софт-маршрутизаторов снятие Netflow, как правило, не представляет собой проблемы, и все упирается только в производительность маршрутизатора. Например, Cisco 7204-NPE-G1 практически не увеличивает загрузку CPU при включении Netflow - все действия для обеспечения данного функционала там сделаны на уровне CEF, то-есть практически "в железе". Использование общей памяти на софт-маршрутизаторах как раз позволяет обрабатывать очень большое количество сессий. Для гигабитных потоков данных практически единственным софт-маршрутизатором на сегодняшний день является Cisco ASR1000. Возможно, что его аналогом можно назвать Huawei NE40E, так как там заявлено как раз очень много чего и на хороших скоростях, плюс анонсирован маршрутизатор на его замену - Huawei CX600 (?), но тестирование данных устройств провести не довелось, к сожалению. Еще очень интересно будет посмотреть на Juniper MX80, так как его реализация на новом чипе, как мне кажется, весьма близка к Cisco ASR1000, исходя из заявленного функционала.

Аппаратные платформы

Рассмотрим варианты организации статистики по трафику пользователей на аппаратных платформах. Напомню, что к аппаратным платформам мы отнесли файрволы по типу Cisco ASA,

Juniper ISG/NS, маршрутизаторы по типу Cisco 7600/6500, Juniper M/MX, другие специальные устройства ограниченного функционала, построенные, как правило, не на универсальных процессорах.

Первое, что приходит на ум при установке такого устройства: “Это же АППАРАТНЫЙ маршрутизатор, большой-дорогой-тяжелый и красивый, он ДОЛЖЕН обработать весь этот трафик и выдать мне то, что я с него хочу”. На самом деле, аппаратные маршрутизаторы подвержены очень и очень большим ограничениям как по функционалу, так и по возможностям расширения. В первую очередь, они предназначены для решения задач обработки трафика, и уж практически в последнюю - для задач ведения его статистики.

Cisco ASA 5505 - 5550 и Juniper ISG/NS не имеют средств по сбору статистики о прошедшем через них трафике, так что даже если вы собираетесь на них выполнять функции NAT, то о шейпинге пользователей и снятии статистики придется позаботиться заранее. Отправка статистики о трафике сделана в Cisco ASA5580, причем по результатам теста сотовых операторов - эта функция нормально работает на синтетическом потоке в 600 тысяч сессий, что дает шанс предположить, что 4-гигабитный поток данных ASA сможет не только транслировать в интернет, но и обеспечить статистику по нему. Правда, следует иметь в виду, что Cisco ASA5580 дает статистику по трафику по протоколу Netflow V9, а не V5, но можно настроить их на использование сислог-сервера и отдачи этой информации на удаленный сислог.

Очень хочется получить Netflow с коммутаторов или маршрутизаторов Cisco 6500/7600. К сожалению, их функционал в плане Netflow ограничен очень малым количеством записей в кэше сессий - буквально 128K или 256K для BXL версий. При этом еще следует учесть, что для рейт-лимита пользователей в случае UBRL используются те же самые ресурсы, что для Netflow, то есть для реальных применений Cisco Netflow с 6500/7600 неприменим по умолчанию.

Есть варианты использовать sampled netflow

(<http://www.cisco.com/en/US/docs/routers/7600/ios/12.2SXF/configuration/guide/nde.html#wp1148270>) - когда учитывается не каждый пакет, а с какой-то периодичностью. Тогда, правда, есть определенный шанс упустить часть данных (особенно для коротких сессий, в несколько пакетов), посчитать объем с приблизительной точностью, и данные о трафике необходимо умножать перед заливкой в биллинг. В случае с постоянно открытым большим количеством сессий, чем “грешат” P2P клиенты, данный функционал может не принести ожидаемого результата, и все равно будет постоянное переполнение таблицы сессий Netflow.

В случае использования маршрутизаторов Juniper, M или MX серий, функционал Netflow можно обеспечить с использованием сервисных модулей (MS-100, MS-200, ..., MS-DPC). Juniper не рекомендует делать одновременно NAT и Netflow на одном сервисном модуле, хотя я лично не понимаю таких ограничений, наоборот, с использованием одновременно NAT и Netflow можно было бы сэкономить ресурсы - и там, и там есть понятие сессии. При использовании MS-модулей на M-серии следует учитывать, что шина на один слот в маршрутизаторе - 1 гбит/сек Full Duplex, то-есть для Netflow на 2-4 гигабитах необходимо устанавливать 2-4 сервисных модуля. MS-DPC с 40 гбит/сек на слот позволяет до определенного момента не заботиться о производительности =), но стоимость в 120K по GPL превосходит границы разумного.

В случае использования BRAS мы можем настроить радиус-аккаунтинг на отдачу отдельной статистики по нескольким видам/классам трафика, и вопрос будет только в организации приема этой информации биллингом. Но это не решает задачи сохранения детальной статистики по трафику для органов государственной безопасности.

Очень неплохим решением для Netflow, особенно при терминции PPPoE/L2TP туннелей является Cisco 10K. На форуме forum.nag.ru был отчет о том, что с использованием PRE2 был снят поток Netflow в 10 mbit/sec, то-есть до 10 тысяч сессий в секунду. Для справки - при домосетевом трафике порядка 6-6.5 гбит/сек скорость создания новых сессий - около 30 тысяч сессий в секунду.

Интересным решением для статистики по трафику может стать syslog по сессиям на Cisco ACE, который я рекомендовал для использования в качестве NAT на больших потоках трафика. Детально вопрос о логировании сессий с количеством трафика, прошедшим по сессии, не рассматривался, но для случаев детализации информации для органов - это вполне приемлемый вариант, так как сразу получается информация “когда-кто-куда-откуда”, а вопрос “сколько” органы госбезопасности не очень волнует, как правило. Технические возможности логирования на модуле ACE - порядка 100 тысяч сессий в секунду, что для усредненного домосетевого трафика - гораздо выше его пропускной способности.

Достаточно естественным решением в случае использования на сети устройств DPI выглядит получение статистики о прошедшем через них трафике прямо с этих устройств. Анализ потоков трафика эти устройства все равно выполняют, подсчет количества трафика - тоже, разделение на классы трафика есть. Поэтому в случае использования Cisco SCE можно получить с нее Transaction RDR, данные в ее внутреннем формате, включающие в себя “логин подписчика - класс трафика - IP сервера - протокол - скачано - закачано” по каждой сессии. Более того, для HTTP трафика данное устройство сохраняет там же поля HTTP Host и HTTP Request URI, что для органов госбезопасности вообще “манна небесная”. Transaction RDR можно сохранять сразу на винчестер, и по запросам уже их обрабатывать. Для статистики в биллинг можно использовать Subscriber RDR, которые представляют собой агрегированные потоки Transaction RDR, и включают в себя “логин подписчика - класс трафика - скачано - закачано”. Эта информация идет по каждому подписчику и по каждому отдельно сконфигурированному классу трафика (счетчики сервисов) каждые 10 (20, 30, зависит от

конфигурации) минут. Этих данных хватает для конвертации в Netflow 5 и “скармливания” данных в биллинг, имеющий Netflow-коллектор. Наверняка, другие модели DPI устройств имеют в себе подобный функционал.

Выводы

Итак, в двух статьях мы рассмотрели варианты решения трех основных задач при предоставлении услуг широкополосного доступа в интернет. Резюмируя все вышесказанное, есть три комплексных подхода к их решению:

использование дешевых PC-серверов, что является оптимальным по соотношению цена/производительность;

использование универсального высокопроизводительного софт-маршрутизатора - на текущий день альтернативы Cisco ASR1000 практически нет;

использование нескольких решений, каждое для своей задачи - например, Cisco SCE для шейпинга и статистики, Cisco ACE для NAT.

На PC-серверах можно решить задачи сетевой трансляции, шейпинга (особенно в случае использования полисинга на линуксе), статистики по трафику. Опыт и консолидация данных с форума показывает, что оптимальным решением является использование ОС Linux, с HTB для шейпинга. Но так же стоит учитывать, что при использовании PC-серверов на 10 гбит/сек потоках достаточно часто придется добавлять новые серверы, производить переконфигурацию балансировщика трафика перед ними (в связи с ростом определенных сегментов сети), и тратить ощутимые деньги на поддержку данного решения. Стоимость инсталляции гигабита обработанных данных в случае использования PC-серверов - около 1500 долларов.

Использование софт-маршрутизатора Cisco ASR1000 становится насущным как при нежелании плодить стойку серверов и постоянно заниматься балансировкой нагрузки, резервированием серверов по схемам N+1 и т.п., так и при архитектуре сети вида “vlan на пользователя”, использовании сервиса ISG, использовании PPPoE. Особенно актуальна Cisco ASR1000 как раз для PPPoE-сетей. Примеров конфигурации Cisco на решение всех перечисленных задач в сети очень и очень много, техподдержку на коммюнити-форумах получить достаточно просто и быстро. Для 10GE решения состав Cisco будет таков: шасси, ESP20, два SIP, два модуля SPA-10GE. Примерный GPL на такое решение - около 80-90K, что дает 4500 долларов инсталляции на гигабит трафика. Но - надо заметить - что данное решение при всей своей дороговизне может быть оптимальным в ряде случаев, так как это же устройство можно держать как единственное головное в сети, решая на нем функции BPG Border (а уже только это в случае нескольких аплинков может обойтись минимум в 2000 долларов инсталляции за гигабит трафика). На тестовом стенде и в боевых инсталляциях Cisco ASR показывает очень и очень хорошие результаты как универсальный маршрутизатор для Ethernet-сетей, и может быть применим для решения практически всех задач.

Использование аппаратных платформ может быть неоптимальным с точки зрения соотношения цена/производительность, но может добавить ряд дополнительного функционала, который сэкономит нервы, время и деньги при решении проблем с трафиком на/от пользователей, например, вирусы, DDOS, уведомление пользователей о необходимости оплаты, а также послужить отправной точкой для реализации дополнительных услуг на сети. Cisco SCE8000/2x10GE + Cisco ACE решают проблемы трансляции, статистики, шейпинга до 7 гбит/сек. Стоимость решения составляет порядка 110K + 60K, то есть цена инсталляции в данном случае максимальна. Но - Cisco 7600, поставляемая в бандле с Cisco ACE может использоваться и как бордер-маршрутизатор, и как коммутатор уровня ядра сети (при инсталляции интерфейсных модулей). Гибкость данного решения будет максимальной - именно за счет применения DPI устройства.

Если еще вчера задачи домосетестроения на гигабитных скоростях выхода в интернет могли решаться только двумя способами - или на PC-серверах, или на аппаратных платформах, то сейчас производительность универсальных процессоров в софт-маршрутизаторах стала стремительно расти, догоняя рынок, и это возвращает софт-маршрутизаторы на свое место в иерархии устройств средней производительности. Cisco на сегодняшний день с маршрутизатором ASR1000 - наиболее оптимальный вариант для строительства “с нуля”, или при полной модернизации сети. Но - “завтра” выходит Juniper MX 3D, который заявляет такой же функционал на новых линейных картах, причем ценник данного решения для 60 гбит/сек фулл-дуплекса может составить около 300K GPL, возвращая цену инсталляции гигабита к 2.5 - 3 тысячам долларов. Для “малых” сетей, до 20 Гбит/сек, предлагается Juniper MX80.

Таким образом, подбор оптимального решения будет всегда зависеть не только от текущих нужд, но и нужд с учетом прогноза на 2-3 года вперед, хотя на 2-3 года с текущим ростом трафика загадывать не приходится. Увеличение абонентской базы на 30% за год может привести к росту внешнего трафика в 3 раза, например, и этому есть немало примеров.