МЕТОДЫ СБОРА ДАННЫХ О ТРАФИКЕ

Сабина Р.А., группа ТКС-01н
Руководитель ст. пр. каф. АТ Бойко В.В.

Главной проблемой при любом моделировании сети является проблема сбора данных о существующей сети. [1]. Сбор данных о трафике при проектировании сети, осуществляемом при помощи имитационного моделирования, т.е. включающим в себя создание достоверной модели, является необходимым условием для создания такой модели. Так как достоверная модель требует обязательной верификации данных.

Существуют различные методы и продукты для решения задач сбора данных о трафике, а задачи мониторинга и анализа сетевого трафика могут решаться на разных уровнях — начиная от мониторинга загрузки сетевых интерфейсов и заканчивая анализом пакетов, собранных с критических участков исследуемой сети.

В настоящее время широкое распространение получили методы, основанные на сборе и обработке детализированной сетевой статистики. В этом случае множество наблюдаемых пакетных заголовков агрегируется по схожим признакам в так называемые информационные записи, которые в дальнейшем и являются предметом анализа. Такой подход сочетает хорошую масштабируемость с подробной информацией о структуре потоков данных в исследуемой сети и позволяет решать целый ряд задач как исследовательского, так и административного характера [2].

Продукты используемые для решения задач мониторинга и анализа трафика — это программный комплекс NeTAMS [3]; мультиплатформенный комплекс для сбора статистики NeTraMet [4]; IOS Cisco NetFlow [5], с помощью которого можно обеспечить практически любой уровень детализации трафика в сети; MRTG (Multi Router Traffic Grapher) [6], сервис, позволяющий посредством протокола SNMP получать из нескольких устройств информацию, и отображать графики загруженности канала (входящий трафик, исходящий,

максимальный, средний) с шагом в минуты, часы, дни и за год; для сбора трафика и построения отчетов в ОС FreeBSD, применяется пакет trafd. Как инструмент системы имитационного моделирования вычислительных сетей COMNET III, таким продуктом является COMNETBaseline, который позволяет создавать разнообразные фильтры, с помощью которых можно извлечь нужную для моделирования информацию и импортировать ее в единую модель трафика. Также в настоящее время разрабатывается единый стандарт IPFIX (Internet Protocol Flow Information eXport) [7], описывающий потоки трафика.

Для задач имитационного моделирования, в результате анализа статистических данных о трафике, можно сформировать основные критерии эффективности работы сети, такие как — пропускная способность, задержка и джиттер задержки.

Были проведены натурные эксперименты, сбор данных о трафике, а именно подсчет количества пакетов и их размер. Рис. 1 схематично отображает исследуемую сеть. В действительности в ядре системы находятся несколько маршрутизаторов, но чтобы не перегружать схему, были приведены только два — маршутизаторы «внешний» («A») и «внутренний» («B»).

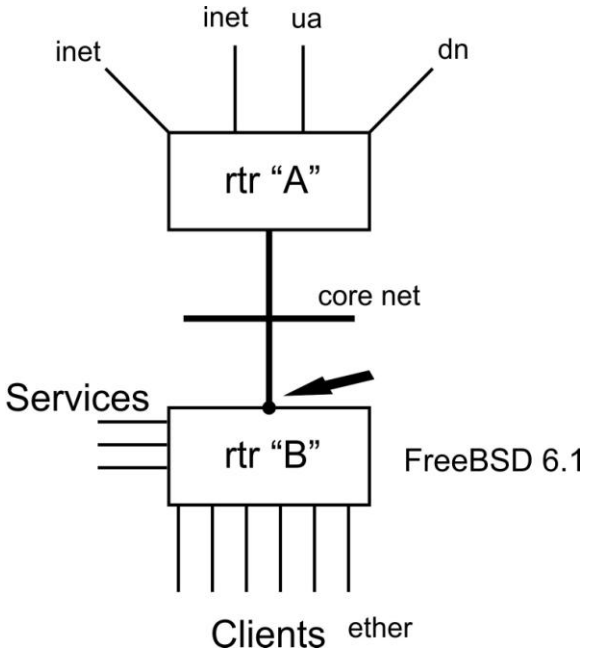

Рисунок 1 — Схематическое изображение сети

Corenet представляет собой несколько объединенных маршрутизаторов, в том числе маршутизаторы «A» и «B», «внешний» маршрутизатор «A» подключен к точкам обмена (ua, dn), а к маршрутизатору «B» подключены клиенты и ряд сервисов, такие как почта, новости, ftp-архивы, игровые серверы и некоторые другие. Интересующий нас, маршутизатор «B», работает под управлением ОС FreeBSD 6.1, клиенты подключены к нему ether-интерфейсами. Клиенты — это крупные и малые предприятия, с собственными корпоративными сетями, удаленные офисы, а также частные лица. Клиенты в большинстве своем применяют internet messengers, ip-телефонию (в т.ч. Skype), пользуются услугами интернет-радио и YouTube (потоковое видео), используют сетевые средства для организации видео-конференции, а также ряд других услуг. Стрелкой отмечена «точка сбора» статистических данных, а именно количество прошедших UDP-пакетов и их размер.

Был выбран учет именно UDP-пакетов, так как приложения и сервисы реального времени в большинстве своем используют именно этот протокол, некритичный к задержкам.

На языке Perl был написан простейший скрипт считающий проходящие через ядро firewall’а пакеты, выбирая их из счетчиков. Учитывался как входящий, так и исходящий трафик. IPFW (IP-firewall), — это системное устройство ОС FreeBSD, которое позволяет выполнять фильтрацию, переадресацию и другие операции с IP-пакетами, проходящими через системный интерфейс. Выбор пакетов производится путем применения упорядоченного списка правил образцов к каждому пакету до тех пор, пока не будет найден подходящий, после чего выполняются соответствующие действия по передаче этого пакета. Команда COUNT не производит никаких действий. Но, как и для любого правила в firewall, в счетчики заносится количество и суммарный объем пакетов, удовлетворяющих этому правилу, поэтому единственное логичное применение этому правилу — подсчет трафика [8]. По умолчанию UDP-пакеты считает правило под номером 5. Скрипт был помещен в CRON системы FreeBSD, данные «снимались» каждые 10 минут, натурный эксперимент продолжался неделю. Пример части firewall’а приведен ниже:

ipfw add 5 count udp

from any to any in

recv <if> …

Проанализировав данные, был сделан вывод о низкой степени детализации трафика. Вне всяких сомнений, использованный метод является сильно упрощенным (без внимания остался и внутренний служебный трафик), однако раскрывает суть метода сбора данных о трафике и дает возможность создать модель трафика для верификации модели сети.

Конечно, существуют средства для сбора данных с высокой степенью детализации трафика. Однако, в этом случае объем собираемой статистики может оказаться довольно большим и, более того, плохо предсказуемым. Типовое решение, связанное с уменьшением детализации собираемых данных, неизбежно ухудшает точность измерительной системы и не всегда позволяет достичь желаемого результата.

Перечень ссылок

1. Н.А. Олифер, В.Г. Олифер. Средства анализа и оптимизации локальных сетей. —Центр Информационных Технологий, 1998.

2. В.В. Коноплев, Д.Ю. Захаров, М.Н. Боярский, Р.Р. Назиров. «Схема адаптивного агрегирования для кластеризации данных сетевого трафика». Институт Космических Исследований РАН. Электронный журнал «Исследовано в России»/ Электронный ресурс. Способ доступа: URL: http://zhurnal.ape.relarn.ru/articles/2003/220.pdf .

3. Электронный ресурс. Способ доступа: URL: http://www.netams.com.

4. Электронный ресурс. Способ доступа: URL: http://www.auckland.ac.nz/net/NeTraMet.

5. Электронный ресурс. Способ доступа: URL: http://www.cisco.com/en/US/products/ps6601/products_ios_protocol_group_home.html.

6. Электронный ресурс. Способ доступа: URL: http://oss.oetiker.ch/mrtg.

7. Электронный ресурс. Способ доступа: URL: http://www.ietf.org/html.charters/ipfix-charter.html.

8. Майкл Эбен, Брайан Таймэн. FreeBSD. Администрирование: искусство достижения равновесия. Энциклопедия пользователя. — ДиаСофтЮП, 2003.