

NAT, который мы потеряли

2002 год, \$700:

Equipment designed for network edge and core applications requires high levels of processing performance to support value-added network services at OC-192/10 Gbps line rates.

*Supports 60 million enqueue/dequeue packet operations per second, enabling deep packet processing of minimum 46-byte Packet-over-Sonet (PoS)
Delivers 10 Gbps packet forwarding and traffic management on a single chip*

2010 год, \$40 000:

"Сможет ли Cisco ASR1002-F сNATить 2 Гбит/сек?"

Данная статья является, в каком-то роде, продолжением изысканий о трансляции сетевых адресов на дешевых (и не очень) платформах, от PC-маршрутизаторов, до аппаратных платформ типа Juniper MS-DPC или SRX, Cisco ACE, и других.

Как показала моя личная практика, при использовании PC-маршрутизатора для NATа весьма критичной является частота шины процессора и скорость обмена с памятью. Так, при использовании HP DL160G5 на Xeon QC 5430 2.66/1333MHz FSB/12M L2 cache загрузка CPU на NAT 100 Кппс FD (100 вход + 100 выход, 200 с точки зрения Cisco Systems) составляет 40%. При использовании HP DL140 Dual Xeon 2.8/800/4 - те же 100 Кpps FD дают загрузку под 100% CPU. Берем Xeon 5504 (2.0/800/4), сделанный по новой технологии относительно 5430 - загрузка будет 80% при 100 Кппс. А если взять 5570 (2.93/1.33/8) - то на этом же потоке трафика загрузка до 20%.

В разговорах в курилке с "Прохожим", он постоянно упоминал, что "сетевой чип на 2 или 4 гигабитных дырки со stateful firewall внутри стоит 250 долларов", и надо только написать под него софт и прибить к сетевому чипу CPU/Flash/PHY". В разговорах с представителями компании D-Link я часто жаловался на отсутствие дешевых производительных NAT-решений, после чего прозвучала некая магическая фраза о "сетевых чипах Cavium".

Тогда я начал изыскания в поисках того самого "священного Грааля" - сетевого чипа, который в себе несет stateful firewall на 1М потоков, и на котором можно сделать аппаратный NAT.

Простой поиск в Google с разными ключевыми словами привел к интересным и неожиданным результатам:



V-Series SSN Network Appliances какая-то сказочная вещь с wirespeed NAT, L7 traffic management, AAA и DPI на скоростях до 32 гбит/сек, при 250K новых сессий в секунду. Отсюда появилось название первого чипа - [AceNet](#) SSPP, и достаточно интересное описание возможных систем, построенных на нем. На базе этого чипа предлагаются различные продукты. Кстати, обратите внимание, что продукт активно продвигается [на китайском рынке](#). Сходу удалось найти ценник в 800 тысяч - 1 млн юаней за верхнюю модель - AG5000. 10 юаней - 44 рубля, 1 миллион юаней - 4.4 млн руб(?). Ясно, что за эти деньги можно купить SCE 8000 и Cisco ACE-8G заодно...

Потом нашелся еще один термин - Router-on-a-chip. Тут тоже есть свои представители, например [Vitesse G-RocX](#). Тайваньская компания [CyberTAN](#) даже выпустила некий маршрутизатор на базе этого чипа, но, надо понимать, что речь идет о SOHO-маршрутизаторе, а не о решении для провайдера.

Словосочетание "wirespeed NAT" приводит нас к разнообразным продуктам, например - NAT-акселератор на 100 Мбит/сек, но всего на 4K сессий. Узнаем совершенно случайно, что [Huawei NE80](#) может всего 1 Гбит/сек на NATе, а так же, что Marvel [представил первый wirespeed NAT](#) на 1 Гбит/сек в далеком 2001 году. Вот описание этого чипа - 128K сессий. На нем, к удивлению, Alcatel Lucent строил DWDM мультимплексоры.

В итоге даже удается найти более-менее известных производителей не просто чипов, а полуготовых платформ - [GE Intelligent Platforms](#) - с производительностью до 10GE. Ключевым моментом для меня стало то, что такие платформы часто оформлены [как сетевые карты](#). Для простоты построения решений на их базе логично предложить воткнуть их в PC с Linux, дать API - и вперед. Цен на эти платформы найти, к сожалению, не удалось.

Удивительно, что в 2009 году, на 11м собрании AfriNIC Huawei [называет себя](#) единственным вендором с wiresspeed NATом на гигабитных скоростях.

Удивительно так же то, что совершенно незамеченным на российском рынке б/у телеком-оборудования остался модуль WS-X6066-SLB, делающий NAT на 2 Гбит/сек FullDuplex, 1М соединений, 1.25 Мппс. Стоит этот модуль на Ebay - \$700+ доставка, снят с производства в далеком 2007м году. Т.е. 3 года назад сняли с производства железку, за которую многие российские ШПД операторы могли бы «душу продать». =) На самом деле интересно, будет ли этот модуль действительно осуществлять NAT.

Но что мы получили в итоге, на самом деле? Совершенно неожиданно нашлась ссылка на публикацию корейских исследователей про NAT на программируемых сетевых процессорах. [Опубликовано в 2005 году](#), в Санкт-Петербурге. Товарищи получили гигабитный NAT на более чем 500K одновременных сессий (Also, we demonstrate that our NAT subsystem can support more than five hundreds of thousands of concurrent TCP/UDP sessions and sustain the full line rate on two Gigabit Ethernet links.). Из введения к этой статье мы получаем название чипа - Intel IXP2400. Бельгийцы в том же 2005 году [тоже пишут](#) про этот чипсет применительно к NATу.



Тут же находится описание чипа, даташит на него, название продукта на этом чипе - Radisys ENP-2611, божеская цена в районе \$600-1000, и все остальные первичные признаки “священного Грааля”.

Оказывается, Intel в 2002 году предложил линейку сетевых процессоров, IXP2350-2400-2600-2800-2850. Стоимость этих процессоров составила от \$400 до \$700.

Топовый из них, 2850, по datasheet может обрабатывать поток до 10 Гбит/сек и содержит аппаратный ускоритель IPSec. Появилось громадное количество тестовых и разработческих платформ, попытка написания компилятора с некоего языка для сетевых приложений, сравнения чипов, тестовые результаты на 1-2 гбит/сек NATа или файрвола, есть сетевые детекторы/генераторы трафика на этих чипах. Но НЕТ МАРШРУТИЗАТОРОВ! Нет коробочных готовых продуктов, которые решали бы простую сетевую задачу.

Китайцы [пишут про этот чип](#) применительно к NATу в 2004 году, анонсируются какие-то маршрутизаторные платформы на базе этого чипа за \$4 тыс. list price. Но чем все это в итоге заканчивается, почему же нет готовых продуктов на этой базе?

В 2004 году Intel, выступая на North American IPv6 summit, предложил эти чипы для организации маршрутизаторов на IPv4 - IPv6 NAT, для плавного перехода между ними.

Так как с 2002 года прошло очень много времени, то логично было бы предположить, что должны быть маршрутизаторы на базе чипа IXP2400/2800, с разведенными портами, с неким набором программного обеспечения, позволяющее решить задачи маршрутизации пакетов.

Вводим в Google "IXP2400 router". Что видим? “Дизайн файрвола в маршрутизаторе на базе 2400”, “предоставление QoS в реал-тайм сервисах в IPv6 сетях на базе процессоров 2400”, “моделирование производительности NAT”, “DDoS-защищенная архитектура интернета”, “Исследование производительности”, “Высокоскоростная гибкая система защиты от сетевых атак”... Нашлись даже индусы, которые в 2003 году собирались сделать файрвол на базе этих сетевых карт.



На сегодняшний день Intel в линейке своих продуктов предлагает [блейд-модуль на базе IXP2850](#). Это 4 гигабитных порта (8 Гбит/сек производительность), аппаратная поддержка IPSec, stateful firewall... Но продается это почему-то только как коробка для дальнейших исследований, а не как законченный продукт.

Некая компания CloudShield в 2004 году выпустила устройство DPI-анализа и защиты от сетевых атак на базе IXP2800 – CS-2000.



Продукт получил восторженные отзывы о производительности и функционале, но реальных упоминаний о массовых продажах этого устройства найти практически не удалось. Кроме того, устройство было предназначено для анализа трафика, а не функций NAT. Цена устройства составила \$50 000.

Увеличение глубины поиска привело, наконец, к положительным результатам, нашлись «умельцы», которые выпустили продукты на базе этих процессоров.



"Натекс", по косвенным данным, в 2007 году представил маршрутизатор 7908 как раз на базе IXP2400, ровно на 1 миллион NAT-сессий, со скоростью изменения в 20K сессий в секунду, можно установить 3 Gigabit Ethernet порта. Но вот беглый просмотр цен показал, что с ценой все очень плохо – \$9000 за коробку, \$9000 за управляющий модуль и 2 гигабитных порта по \$3300 каждый – это перебор.

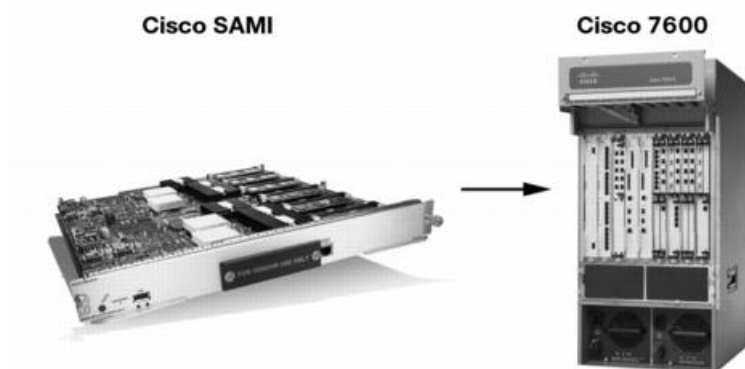
И, как обычно, китайцы не подвели.

Нашелся маршрутизатор BDCOM 7204, построенный на базе IXP2400. Естественно, заявлено все то, что умеет процессор – гигабитный NAT, 1М одновременных сессий. [Ценник гуманный - \\$2200](#).

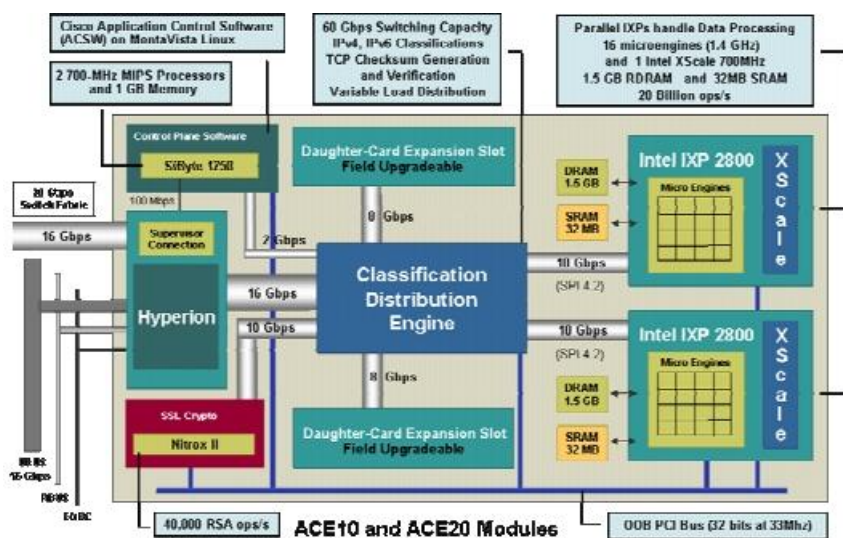
В 2007 году подразделение этих процессоров было продано в Marvel, а сейчас Intel предлагает линейку новых чипов, как продолжение тех сетевых процессоров. Из чисто сетевых они стали универсальными и резко "похудели", как в цене, так и в производительности. Большинство SOHO-маршрутизаторов делаются на базе Intel XScale ARM IXP425/435, младший из которых был анонсирован одновременно с IXP2400, в 2002 году.

Отсутствие полноценной линейки законченных продуктов по приемлемым ценам на данных сетевых чипах – это заговор производителей или какие-то скрытые проблемы с функциональностью под нагрузкой? Или полная неспособность к маркетингу, как в случае с компанией Натекс?

Когда статья уже готовилась к публикации, выяснились очень интересные подробности, которые отражают общую идею о заговоре производителей. Наиболее производительные модули NAT для маршрутизаторов Cisco построены на базе Intel IXP2800.



Это и Cisco Content Services Gateway, предназначенный для обработки пользователей мобильного интернета, до 500K подписчиков на один модуль, и Cisco ACE20-MOD-K9, сделанный на базе двух IXP2800.



Именно возможности IXP-чипов позволяют модулю ACE работать с потоками в 12 Гбит/сек трафика и 2 Мппс, держа до 8 миллионов одновременных сессий.

Соответственно, дешевые сетевые чипы плюс обвязка в китайском варианте стоят 2200 долларов, а при наличии на этом лейбла Cisco – 40 тысяч долларов.

Что же в итоге? “Пропал подъезд”. Для 2-гбит NATа рекомендуют покупать ASR1002-F, стоимостью 40K\$ GPL. А так бы... \$1000 сервер, \$1000 сетевая карта на IXP2400 (ну, пусть \$3000 на базе 2850), \$1000 за софт. \$5000 за 10 Гбит/сек производительности - по-моему, это был бы прорыв в сетестроении.

Реально хочется взять пару таких сетевушек, системного программиста, написать NAT-модуль для линукса на карте прямо. Хотя бы для себя.

Так что у кого найдется такая карточка - куплю недорого. NAT буду делать.

Примечание NAG. Карточка Radisys ENP-2611 уже куплена, и окажется у Кирилла примерно через месяц. ;-)