

Реализация гибкого аппаратного маршрутизатора для новой IP-инфраструктуры

Чак Семерия

Инженер по маркетингу

Copyright © 2002 Poplar systems



<http://www.poplar.ru>

Общие положения

В 1998 году компания Juniper Networks, Inc. вывела на рынок магистральный Интернет-маршрутизатор M40. Он стал первым, специально сконструированным для решения специфических проблем с которыми столкнулись поставщики магистральных услуг Интернета. С тех пор возможности наших решений как для магистралей, так и для границы сети поставщика услуг значительно расширились:

- разработана специализированная ИС Internet Processor II;
- созданы четыре новые аппаратные платформы (маршрутизаторы M5, M10, M20 и M160);
- выпущено более 80 различных интерфейсных плат, в том числе первый интерфейс OC-192c/STM-64;
- выпущено 12 модификаций ОС Junos;
- созданы 10 специализированных ИС;
- четырехкратно увеличена плотность и общая пропускная способность наших систем.

Эта статья начинается с краткой истории маршрутизации в сети Интернет и тех проблем, что возникают при использовании маршрутизаторов предприятий в условиях Интернета. Мы проанализируем основные тенденции, характерные для середины 90-х годов, и то, как эти тенденции способствовали созданию компании Juniper Networks. Далее мы рассмотрим, как эти факторы повлияли на избранное нами направление проектирования и разработки аппаратных и программных средств для Интернет-маршрутизаторов. Затем, чтобы можно было оценить сложность собственно процесса проектирования, мы покажем, какие компоненты находятся в распоряжении производителя при конструировании Интернет-маршрутизатора, проанализируем уникальные возможности специализированной ИС Internet Processor II и их использование в наших системах для поддержки интеллектуальных IP услуг и сохранения высокой производительности.

После рассмотрения функций и преимуществ специализированной ИС Internet Processor II, мы приведем подробное описание архитектуры нашей системы. Именно архитектура обеспечивает гибкость, требующуюся поставщикам услуг в условиях постоянно меняющейся среды Интернета. Эта гибкость позволяет повысить рентабельность за счет существенного увеличения срока эксплуатации наших систем. В заключение мы рассмотрим, как наша концепция сервисной интерфейсной платы позволяет еще больше увеличивать гибкость наших систем.

История маршрутизации в Интернете

Для того чтобы оценить роль, которую Juniper Networks играет в развитии Интернета, необходимо иметь представление об истории маршрутизации в Интернете. В 1996 году, когда Прадип Синдху основал нашу компанию, коммерческий потенциал Интернета только начал осознаваться обществом, и в глобальных сетях появилась IP-маршрутизация, постепенно ставшая устойчивым сегментом рынка. Однако существующие маршрутизаторы класса предприятия плохо подходили для обслуживания нового направления. Ключевая стратегия Juniper Networks основывалась на том, что различные технологии, необходимые для создания реального магистрального маршрутизатора следующего поколения уже достигли того уровня, при котором ранее невозможные решения могли быть, наконец, построены. Juniper Networks стала первой компанией, осознавшей это конвергентное развитие технологий и успешно предложившей на рынок специализированный маршрутизатор для новой общедоступной сети.

Появление маршрутизаторов класса предприятия

Первые маршрутизаторы разрабатывались для удовлетворения потребностей предприятий и изначально были призваны решать совершенно другие задачи, по сравнению с тем, для чего они сейчас используются в Интернете. В конце 80-х и начале 90-х годов основным назначением маршрутизаторов было обеспечение связи в пределах предприятия локальных вычислительных сетей (ЛВС), построенных на основе разнообразных технологий (рис. 1). В то время на предприятиях эксплуатировались разнородные ЛВС, разработанные различными производителями по собственным технологиям, и развернутые в разное время. Маршрутизаторы

размещались в центре сети и обеспечивали взаимную связь ЛВС с различающимися технологиями. Ниже перечислены некоторые основные свойства такой среды:

- Относительно низкие скорости передачи (Ethernet - 10 Мбит/с, token ring - от 4 до 16 Мбит/с, последовательный интерфейс - менее 2 Мбит/с).
- Сравнительное малое количество пользователей.
- Основное применение - обеспечение взаимодействия в мультипротокольной среде.

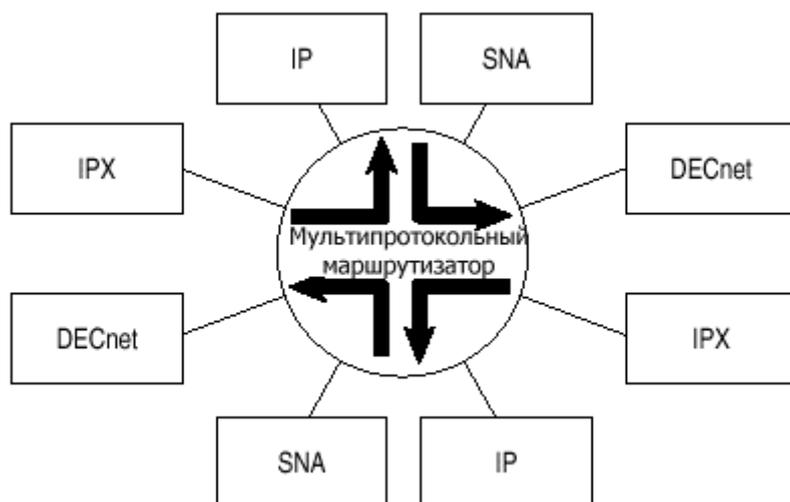


Рисунок 1. Типичная сеть предприятия

Как и следовало ожидать, специфические требования корпоративных сетей обусловили направление проектирования маршрутизаторов предприятий. Системы строились на основе малых, низкопроизводительных компьютеров с ограниченной памятью, так как платформы такого типа было более чем достаточно для решения имеющихся задач. Использовались малые встроенные операционные системы, во-первых, потому что они были дешевле, а во-вторых, потому что не было необходимости в расширенных функциях, таких как защита памяти или приоритетное обслуживание. Все сложные задачи (такие как передача служебного трафика протоколами маршрутизации, обработка маршрута, поиск маршрута) выполнялись программно. Наконец, весь форвардинг пакетов был реализован программными средствами, так как требования к производительности, предъявляемые сетью предприятия, не предполагали использования специализированных аппаратных средств.

"Коммутируйте, когда можно, маршрутизируйте – когда необходимо"

К середине 90-х годов сети предприятий проектировались в соответствии с популярным выражением "Коммутируйте, когда можно, маршрутизируйте – когда необходимо". Коммутация была способна обеспечить необходимую в ЛВС чистую производительность форвардинга пакетов, так как технология полупроводниковых ИС тогда позволяла относительно простые задачи уровня 2 реализовать аппаратно. Если коммутация повысила производительность форвардинга между сегментами ЛВС, то маршрутизация обеспечивала тот уровень контроля, без которого администраторы не могли эффективно управлять потоком трафика через свои сети. Из-за того, что маршрутизация (форвардинг на уровне сети) была существенно сложнее, чем коммутация (форвардинг на уровне 2), и так как существующие технологии полупроводниковых ИС не позволяли реализовать эти функции аппаратно, маршрутизацию приходилось реализовывать программно. Однако даже если развитие технологии полупроводниковых ИС в тот момент позволило бы реализовать маршрутизацию аппаратно, постоянное совершенствование технологии маршрутизации в мультипротокольной среде настолько сокращало срок жизни аппаратных реализаций, что они становились экономически невыгодными.

В то время как между сегментами ЛВС была важна чистая производительность форвардинга, проектировщики сетей понимали, что в глобальных вычислительных сетях (ГВС) производительностью форвардинга коммутаторов можно пожертвовать ради обеспечения

гибкости и управления маршрутизацией. Эта модель проектирования стала очень популярной, чему было несколько причин.

- Относительно низкоскоростные интерфейсы ГВС не требовали высокопроизводительных механизмов форвардинга пакетов.
- Объем трафика, отправляемого через ГВС предприятия, был мал по сравнению с объемом трафика ЛВС. Сеть предприятия проектировалась в соответствии с правилом 80/20: 80% трафика было локальным, и только 20% трафика покидало пределы ЛВС и уходило в глобальные сети.
- Приложения со строгими требованиями к задержке, функционирующие в сети предприятия, ограничивались пределами ЛВС.
- По некоторым причинам маршрутизаторы были единственными сетевыми устройствами, способными решить большое количество насущных проблем управления трафиком, которые не решались при использовании коммутаторов уровня 2.
- Маршрутизаторы обеспечивали эффективную связь между различными ЛВС, построенными по технологиям уровня 2 (Ethernet, Token Ring, FDDI, ATM и т. п.). Те, кто проектировал сети в начале 90-х годов, помнят, что мосты трансляции и коммутаторы вначале рассматривались как альтернатива маршрутизации. Однако они не решали поставленную задачу вследствие сложности отображения несовместимых сегментов ЛВС между различными технологиями.
- Маршрутизаторы ограничили количество ширококвещательного трафика в сети за счет разделения последней на сегменты с отдельными ширококвещательными доменами. Это было особенно эффективно в условиях ГВС, где приходилось тщательно распределять дорогую и не столь широкую полосу пропускания.
- Маршрутизаторы позволили сетевым администраторам точно управлять передачей трафика между сетевыми сегментами сети в мультипротокольных сетях. Т.к. маршрутизаторы имели доступ к существенно большей информации заголовка пакета, чем коммутаторы уровня 2, они могли принимать более интеллектуальные решения по продвижению пакетов, используя фильтрацию, правила распределения трафика и выбора маршрутов.
- Маршрутизаторы поддерживали развертывание сетей с ячеистыми топологиями и маршрутов равной стоимости с распределением нагрузки, преодолевая ограничения топологии, построенной на основе протокола spanning tree, которая требовалась для коммутаторов уровня 2.

Успешный провал

С течением времени на предприятиях поняли и оценили возможности связи, обеспечиваемой маршрутизаторами. В результате, производители маршрутизаторов на основе программных средств добились огромных успехов. Однако по ряду причин начальный успех, в конечном счете, с ростом размера корпоративных сетей и популярности программных маршрутизаторов, привел к краху.

- Рост количества необходимых протоколов (XNS, IPX, AppleTalk, OSI, Vines, SNA, DECnet, IP и др.) повлек за собой рост объема программного обеспечения.
- Рост количества и типов необходимых физических интерфейсов требовал больше аппаратных и еще больше программных средств.
- Быстрый рост числа пользователей требовал увеличения производительности программных маршрутизаторов. Это достигалось уменьшением времени выполнения программ с целью минимизации влияния на производительность форвардинга пакетов, и программное обеспечение производилось с учетом определенных предположений о функционировании аппаратных средств. Более того, сокращение времени выполнения программ достигалось за счет уменьшения времени проверки ошибок и использования упрощений, снижавших стабильность программного обеспечения. Учет специфики аппаратных средств в программном обеспечении уменьшал гибкость системы, т.к. аппаратные и программные средства оказывались тесно связаны друг с другом.

- Рынок маршрутизаторов все больше попадал под влияние необходимости разработки новых функций для каждого поддерживаемого протокола. В итоге, на рынке стал доминировать только один производитель маршрутизаторов, так как только он оказался способен финансировать исследовательскую организацию, создававшую эти функции с такой скоростью, которой не могли достичь другие изготовители. Это увеличило долю этого лидера на рынке, но наличие лидера-разработчика функций еще больше усложнило программное обеспечение.

И затем, с середины и до конца 90-х, широкомаштабный коммерческий доступ и развертывание Интернета резко изменили требования, предъявляемые к маршрутизаторам.

- Теперь от маршрутизаторов требовалась поддержка миллионов, а не сотен или тысяч пользователей.
- Производительность форвардинга пакетов стала основным требованием при поддержке приложений нового поколения, потреблявших все большую полосу пропускания.
- Плотность и энергопотребление приобрели большое значение, по мере того как стоечное пространство и электроэнергия стали критичными ресурсами.
- Надежность аппаратных и программных средств стала основной заботой при поддержке специализированных приложений, создающих фундамент успеха бизнеса.
- Значительно возросла роль ГВС по отношению к ЛВС.
- Стали резко меняться модели трафика. В Интернете больше не действует правило 80/20, так как полная связность (возможность соединения любого узла сети с любым другим) стала нормой.

Общедоступная сеть Интернет появилась в то время, когда корпоративные маршрутизаторы уже достигли пределов сложности. На рис. 2 показано, что если решение достигает характеристик, которые уже неспособна поддерживать его архитектура, сложность решения, в конце концов, достигнет точки, лежащей за пределами человеческого понимания. Когда это случается, продолжение разработок можно прекратить, так как любая попытка исправить ошибку или добавить новую функцию будут приводить к появлению двух или трех новых ошибок.

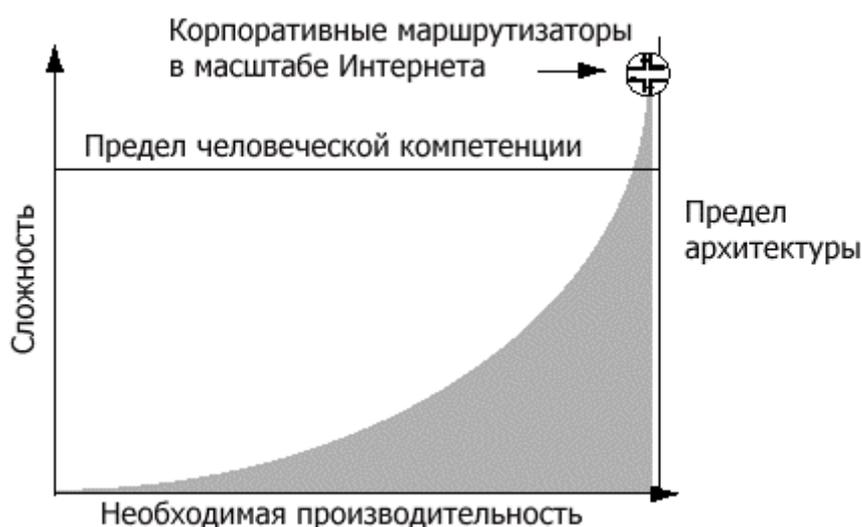


Рисунок 2. Влияние увеличения сложности программных маршрутизаторов.

Когда появилась Juniper Networks

Juniper Networks создавалась в середине 90-х, когда в телекоммуникационной отрасли действовали четыре основные тенденции. Появление созданных Juniper Networks специализированных Интернет-маршрутизаторов привело к тому, что последняя из ниже

перечисленных тенденций во многом исчерпала себя, но остальные три продолжают действовать и сегодня.

- Экспоненциальный рост Интернета.
- Фундаментальная важность протокола IP.
- Снижение стоимости полосы пропускания.
- Незрелость технологии IP маршрутизации.

Экспоненциальный рост Интернета

Экспоненциальный рост Интернета обусловлен тремя основными факторами:

- увеличением мощности универсальных вычислительных устройств;
- преимуществами полной связности;
- принципиальной независимостью приложений.

Рост мощности универсальных вычислительных устройств

Изначально компьютеры проектировались для решения задач, которые могли решаться и людьми, но необходимо было, чтобы эти задачи решались намного быстрее. С успехом персонального компьютера, производители начали разрабатывать приложения, помогающие экономить время и выполнять задачи, к которым человек изначально не приспособлен.

Возможности полной связности

Связь, или общение – основа человеческого общества. Чем теснее общаются отдельные индивидуумы в группе, чем лучше они связаны, тем более мощной становится вся группа в целом. Это демонстрируется снова и снова на протяжении всей истории человеческой деятельности через язык, книги, телефонную связь, радио, телевидение и, в настоящее время, Интернет. Появление каждой из этих технологий увеличивало скорость и разветвленность взаимодействия в человеческом обществе.

Независимость приложений

Рост Интернет сегодня не определяется приложениями, для которых сеть Интернет использовалась первоначально. До последнего времени поставщикам услуг требовалось развертывать отдельную сеть для каждого предоставляемого вида услуг (голос, видео, данные, Интернет). Каждая из этих сетей развивалась независимо от других, так как было технически невозможно обеспечить функционирование всех приложений в общей инфраструктуре. Из-за существующих технологических ограничений каждая из этих крайне дорогих параллельных инфраструктур создавалась специально для различных приложений. На начальном этапе развития сеть Интернет тоже была отдельной сетью, спроектированной для решения конкретной задачи обеспечения связи между различными исследовательскими и образовательными институтами.

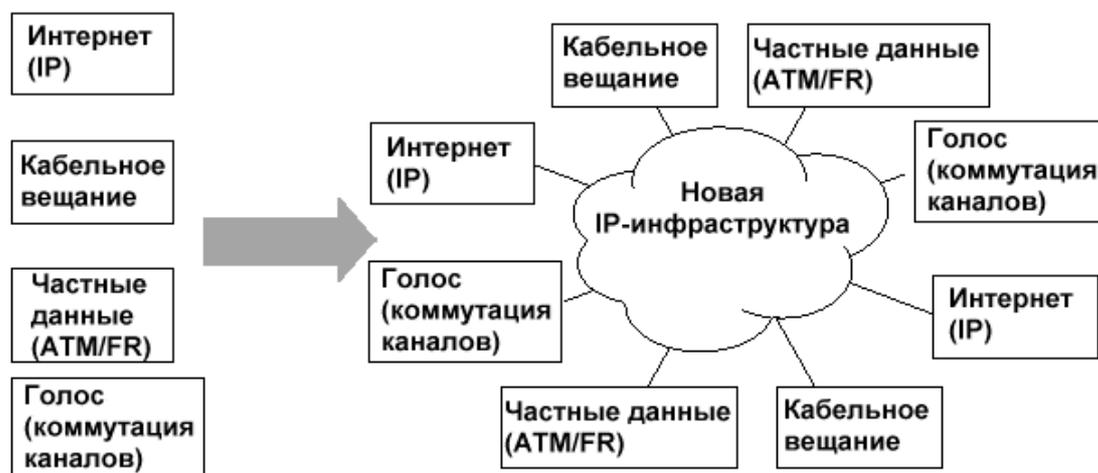


Рисунок 3. Независимость приложений

Принцип независимости приложений означает возможность построения такой сети, которая сама по себе она не ориентирована ни на какое конкретное приложение. Единственное назначение такой сети – объединить все традиционные сквозные службы, ранее задействовавшие ресурсы независимых сетей, в границах общей инфраструктуры. Это никогда бы не оказалось возможным без появления Интернета, и здесь можно вспомнить известную проблему: что появилось раньше, курица или яйцо. Невозможно строить общую инфраструктуру с глобальным охватом, если она не имеет применения, но Интернет обеспечил применение, стимулировавшее создание глобальной инфраструктуры. Развертывание глобальной IP инфраструктуры в сочетании с недавними успехами в технологии, делает возможным для поставщиков услуг и их абонентов получать большую отдачу от инвестиций, уменьшить эксплуатационные расходы, предлагать новые услуги на мировом рынке.

Фундаментальная важность протокола IP

Протокол IP - единственная технология, поддерживающая полную интерактивную связность между компьютерами в мировом масштабе. Преимущество полной связности существенно возрастает при наличии интерактивности в сравнении с тем, когда связь односторонняя.

- Пакетный протокол IP изначально разработан для связи компьютеров между собой, обеспечивает требуемый коэффициент статистического мультиплексирования.
- IP - протокол без установления соединения, поэтому он более надежен и лучше масштабируется, чем альтернативные протоколы, такие как ATM и Frame Relay.
- Развитие протокола IP зависит от его использования, он развивается не по желанию какого-либо комитета, но за счет усилий отдельных личностей, решающих задачи, которые на опыте им лучше известны, сохраняющих то, что работоспособно и отбрасывающих ненужное. Это приводит к фундаментальному принципу "работоспособного кода и консенсуса в целом" в противоположность строительству "башни из слоновой кости".

Уменьшение стоимости физической полосы пропускания

В начале 90-х полоса пропускания ГВС была тщательно распределявшимся дорогим и дефицитным ресурсом. С середины 90-х благодаря успехам электроники и оптических технологий произошло резкое снижение стоимости физической полосы пропускания оптического канала.

- Уменьшилось затухание оптоволоконных кабелей.
- Лазеры стали более стабильными.
- Детекторы стали более чувствительными.
- Увеличилась пропускная способность оптоволоконных кабелей благодаря достижениям в технологиях WDM и DWDM

- Разработаны оптические усилители с присадками эрбия.
- Увеличилась скорость работы электронных компонент.
- Новые методы коррекции ошибок

Неразвитость технологии IP маршрутизации

Как обсуждалось выше, состояние технологии корпоративных маршрутизаторов достигло уровня сложности, с которым невозможно было работать. Основная проблема маршрутизаторов была в том, что они проектировались для решения совершенно других задач, чем те, которые им пришлось решать теперь.

Ограничения корпоративных маршрутизаторов в Интернете

Корпоративные маршрутизаторы не предназначались для работы в Интернете. Их системная архитектура была заимствована из компьютерной платформы, на которой они строились. При загрузке трафиком Интернета функции программного обеспечения и программных средств разделялись неправильно. Системы были неустойчивы и ненадежны, так как работать им приходилось на пределе возможностей. Из-за ограничений производительности маршрутизаторы предприятий стали слишком сложны для работы с Интернетом, так как они требовали постоянной подстройки для "выжимания" последних капель производительности. И тогда поставщик услуг бросал все силы на поддержание работоспособности, вместо того, чтобы планировать новые услуги и думать о развитии.

Корпоративные маршрутизаторы имели жесткий программный код, что делало весьма трудным улучшение стабильности и производительности. Программное обеспечение разрабатывалось для поддержки взаимодействия мультипротокольных ЛВС, и основной упор делался на разработке соответствующих функций. К сожалению, эти функции оказались бесполезны в Интернете, где пользователи связываются по ГВС с использованием IP.

Наконец, аппаратная платформа маршрутизаторов была очень слабой. Маршрутизаторы предприятий строились на основе микропроцессоров и программируемых вентильных матриц, работавших с тактовыми частотами от 25 до 40 МГц. Этого было мало для обеспечения скорости форвардинга порядка десятков миллионов пакетов в секунду. В результате образовался разрыв между аппаратными средствами, используемыми для построения сетей, и аппаратными средствами, служившими базой для высокопроизводительных компьютеров, которым сеть была нужна для функционирования приложений.

Дилемма монополиста

А в это время наиболее успешный производитель корпоративных программных маршрутизаторов на собственном опыте столкнулся с тем, что называется *дилеммой монополиста*. Такой монополист размещает существующие продукты (корпоративные маршрутизаторы) на новом рынке (Интернет). Эти продукты – единственное, что может использовать новый рынок, так как никто еще не создал специализированную компанию для решения его конкретных проблем. Так как существующие продукты были единственными доступными решениями, а первоначальная прибыль - высокой, монополист построил свою бизнес-модель вокруг этих "успешных" продуктов и затем вернулся к обслуживанию уже имевшихся потребителей на других рынках, как это было до появления нового рынка.

Бизнес-модель, основанная на использовании недоработанной технологии на новом рынке, успешна только до тех пор, пока кто-то, сконцентрировав усилия на создании специализированных продуктов, не начнет отбирать часть рынка у монополиста. Такое случалось неоднократно.

- Универсальные компьютеры IBM работали как мини компьютеры до тех пор, пока не появилась первая компания по производству мини компьютеров.
- Производители мини компьютеров чувствовали себя прекрасно, пока не появилась компания, производящая персональные компьютеры.
- Корпоративные маршрутизаторы работали в Интернете до появления компании, производящей магистральные Интернет-маршрутизаторы.

Дилемма монополиста вовлекает доминирующего производителя в период рационализации, пока не начнется отставание, и в этот момент монополист вынужден догонять новую компанию, сконцентрировавшую усилия на разработке специализированных продуктов.

Подход Juniper Networks

В 1996 году, когда Juniper Networks решила создать аппаратный магистральный Интернет-маршрутизатор, эту проблему еще никто не пытался решать. Первый шаг проектирования заключался в проведении консультаций с будущими потребителями для точного понимания их требований. В результате мы пришли к выводу, что в любой платформе маршрутизации, которые мы будем поставлять, должны присутствовать определенные фундаментальные функции.

- Мы должны реализовать открытые протоколы (причем сделать это в масштабе Интернета) для обеспечения функциональности, взаимной совместимости и масштабируемости, требуемой поставщикам услуг. Это определило начало проектирования программного обеспечения с чистого листа бумаги, без попыток добавлять новые функции в уже существующие программы.
- Для исключения возможных узких мест в сети, которые начнут появляться на исходе столетия, мы выбрали модель проектирования системы 8 X OC-48/STM-16, а не терабитной системы маршрутизации. Так как было невозможно создать машину с такой производительностью на базе программного форвардинга пакетов, мы поняли, что должны основываться на специализированных аппаратных средствах.
- Маршрутизаторы отвечают за интеллектуальность сети, так как обеспечивают оперативное управление и создают дополнительные услуги. Требования поставщиков к созданию услуг и управлению постоянно меняются, поэтому архитектура разделения функций между аппаратными и программными средствами должна быть чрезвычайно гибкой, что позволит адаптировать наши системы под постоянно меняющиеся условия Интернета.
- Если миллионы пользователей будут использовать наши системы для связи "каждый с каждым", то должен быть обеспечен беспрецедентный уровень производительности и надежности. Следовательно, требуется тщательно спланировать и эффективно реализовать разделение системных функций (особенно между программными и аппаратными средствами). Иначе, полученное решение будет иметь серьезные изъяны, которые, в конечном счете, предопределят его провал.

Разделяй и властвуй

Для достижения поставленных целей задачу построения маршрутизатора для Интернет необходимо было разделить на две части (рис. 4).



Рисунок 4. Подсистема маршрутизации и подсистема форвардинга пакетов

Во-первых, было необходимо сконструировать подсистему маршрутизации (ПМ), которая могла быть реализована программно и работать на микропроцессоре общего назначения. ПМ должна отвечать за выполнение протоколов маршрутизации, обслуживание таблиц маршрутизации, создание адресных таблиц и управление программными средствами сети.

Во-вторых, требовалась подсистема форвардинга пакетов (ПФП), функционирующая аппаратно. ПФП отвечала бы за выполнение функций коммутации, необходимых для продвижения пакетов через систему от входного интерфейса к выходному интерфейсу.

Мы выбрали такой подход по ряду причин.

- Возможность охватить чрезвычайно сложную проблему и разделить ее приблизительно на две равные части.
- Планируя разработку простого интерфейса между этими двумя частями, мы знали, что сопровождение этой архитектуры на протяжении длительного времени не потребует больших затрат.
- Нагрузка на одну часть системы не сказывается неблагоприятно на производительности другой части, так как они функционируют независимо. Это устраняет одну из главных причин отказов программных маршрутизаторов: либо большое количество маршрутных изменений снижает производительность форвардинга, либо напряженный трафик снижает способность системы адаптироваться к маршрутным изменениям.
- Для ПМ мы выбрали стандартные компоненты Intel. Это позволило нам использовать последние достижения в области специализированных компьютеров и произвести несколько поколений ПФП фактически без коррекций ПМ.
- Такой подход способствовал независимой разработке аппаратных и программных средств, что, в свою очередь, позволило протестировать программное обеспечение независимо от оборудования (которое разрабатывается дольше). В результате, система была выведена на рынок намного скорее, чем при использовании какой-либо другой концепции.

Сейчас – это общепризнанный метод разработки Интернет-маршрутизатора, но в 1996 году он рассматривался как революционный.

Архитектура ОС JUNOS

Операционная система JUNOS является полностью новой разработкой и обеспечивает стабильность, производительность и масштаб, требуемый для Интернет-маршрутизатора. Имевшиеся платформы маршрутизации были спроектированы с использованием жесткого, одноадресного пространства и встроенной операционной системы, вследствие чего сбой в одной части системы (например, из-за ограничений памяти) вызывает полный отказ всей системы.

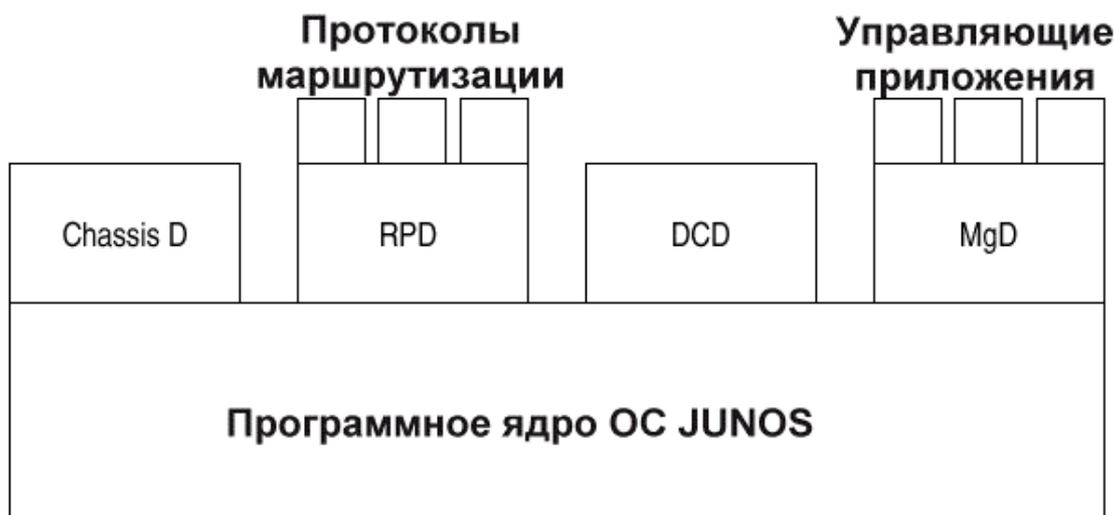


Рисунок 5. Архитектура ОС JUNOS

В современном проектировании операционных систем применяется подход, основанный на использовании ядра (рис. 5), предоставляющий ряд преимуществ.

- Выделяется некоторое базовое множество функций, для выполнения которых строится очень маленькое и исключительно стабильное ядро.
- Приложения (такие как протоколы маршрутизации и сетевое управление) функционируют поверх ядра в отдельных, жестко выделенных адресных пространствах, что обеспечивает надежную защиту отдельных процессов.
- Модульный подход с использованием отдельных (clean) интерфейсов между ядром и каждым из процессов и между самими процессами уменьшает число ошибок в цикле разработки, ограничивает влияние любой конкретной ошибки на систему в целом и упрощает тестирование новых функций программного обеспечения.

Основное преимущество использования ядра заключается в повышении надежности программного обеспечения и упрощении его сопровождения. Эта архитектура оказала огромное влияние на продуктивность работы Juniper Networks, так как позволила быстро внедрять новые функции и быстро вносить необходимые изменения. Без модульного проектирования было бы невозможно четыре раза в год модернизировать операционную систему.

Аппаратная архитектура

Наши аппаратные средства разработаны с использованием микросхем самой высокой степени интеграции. Использование таких ИС снижает стоимость, увеличивает производительность, повышает надежность, снижает энергопотребление. Разработка аппаратуры на базе ИС этого типа требует большого искусства. Проектирование полупроводниковых ИС полностью отличается от проектирования программных средств, так как нельзя отредактировать текст, убрать ошибку и загрузить исправленную версию.

Эти отличия обусловили следующее:

- Процесс проектирования начался с тщательного разделения системных функций и последующего создания надежных, стабильных интерфейсов между отдельными компонентами.
- Разработка каждой заказной ИС проводилась очень тщательно, так как серьезная ошибка могла привести к полугодовой задержке появления готового продукта.
- После разработки отдельных заказных ИС, они должны были интегрироваться в систему таким образом, чтобы вся система обеспечивала максимальную производительность и максимальную стабильность при отсутствии внутренних узких мест.

Принцип проектирования, основанный на разработке очень "чистых" и стабильных интерфейсов, как в аппаратуре, так и в программных средствах использовался при создании всех наших продуктов. Он позволил обеспечить высокую продуктивность, так как такие интерфейсы позволяют создавать компоненты аппаратуры и программ, которые можно либо использовать повторно, либо усовершенствовать для создания новых продуктов.

Подсистема форвардинга пакетов, использованная во всех маршрутизаторах серии M, построена с использованием:

- базового набора ИС, содержащего от 4 до 6 ИС;
- встроенного управляющего ПО;
- интерфейса между встроенным ПО и базовым набором ИС;
- интерфейса между базовым набором ИС и набором ИС ввода-вывода (см. рис. 6).

Такая архитектура позволила ограничить функции карт ввода-вывода и использовать эти карты в различных платформах маршрутизации. С годами наша концепция продолжает доказывать свою эффективность, позволяя в сжатые сроки завершать серьезные разработки.

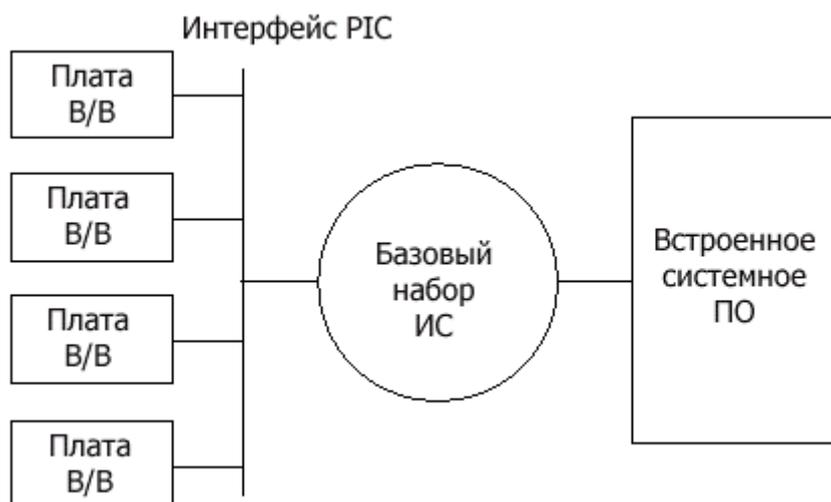


Рисунок 6. Архитектура подсистемы форвардинга пакетов

Аппаратные компоненты

При построении аппаратного Интернет-маршрутизатора производитель может комбинировать различные технологии.

- Готовые компоненты
- Микросхемы памяти
- Микропроцессоры общего назначения
- Перепрограммируемые вентильные матрицы (FPGA)
- Специализированные заказные ИС (ASIC)

Готовые компоненты

В середине 90-х годов, когда маршрутизаторы предприятий пользовались особым спросом, их основные компоненты закупались у сторонних производителей. Наиболее подходящим аппаратным решением были стандартные процессоры, так как главной задачей была гибкость программной поддержки мультипротокольной маршрутизации, а полоса пропускания ГВС стоила еще достаточно дорого.

Сегодня требования производительности Интернета ведут к разработке ИС, оптимизированных для обработки пакетов. Однако производители маршрутизаторов не хотят тратить время на полную разработку системных компонент, так как это серьезно увеличивает время вывода продуктов на рынок. В результате продолжают закупки отдельных компонент у сторонних производителей.

Обычно стандартные компоненты в маршрутизаторах используются для обработки MAC-адресов (Ethernet и ATM SAR) и низкоскоростных кадров SONET.

Среди проблем, с которыми сталкиваются производители маршрутизаторов при покупке компонент у сторонних производителей можно выделить следующие:

- ИС должны быть совместимы со стандартами для обеспечения взаимной совместимости.
- Компоненты должны работать со скоростью линии и соответствовать внутренним устройствам системы.
- Производитель маршрутизатора должен весьма тщательно проектировать интерфейс между готовыми компонентами и специализированными ИС.
- Крайне важна квалификация производителя ИС, так как производитель маршрутизатора должен выбрать наилучшую технологию, а производитель ИС – обеспечить надежность и регулярность поставок.

Микросхемы памяти

Маршрутизаторам память необходима для буферизации пакетов, хранения таблиц форвардинга и значений счетчиков, управления очередями пакетов и т.п. Микросхемы памяти можно разделить на две основные категории: SRAM и DRAM. Статическая память с произвольным доступом (SRAM) создавалась для обеспечения исключительно быстрого произвольного доступа, необходимого при обслуживании счетчиков, поиске в таблице маршрутизации, инкапсуляции и деинкапсуляции уровня 2, создании таблиц транзитных переходов. В динамической памяти с произвольным доступом (DRAM) время доступа к памяти больше, но также увеличен и объем памяти, необходимой для поддержки передачи данных (хранение пакетов и ячеек). Максимальное время задержки в DRAM на несколько порядков меньше времени поиска в таблице форвардинга.

Выбирая память, производитель должен определить наиболее передовую технологию для каждого из этих классов, так как Интернет-маршрутизаторы постоянно работают на пределе технологических возможностей.

- Выбор технологии памяти определяется спецификой приложения, которая обычно подразумевает компромисс между частотой доступа к большим объемам данных в единицу времени (DRAM) и относительно нечастом доступе к малым объемам данных в единицу времени (SRAM).
- Из-за высокой скорости, случайной природы доступа и недостаточной временной или пространственной локализации таких объектов как счетчики, производителям маршрутизаторов требуется память SRAM, которая может обеспечить максимальную производительность. Это требует использования памяти SRAM с *конвейерным режимом (pipeline late mode)*, которая применяется в серверах класса high-end, но не такой, которая применяется в персональных компьютерах.
- Что касается DRAM, то производители памяти для графических приложений с кадровой буферизацией делают быструю память (обеспечивает высокие полосы пропускания при чтении-записи) небольшого объема. Производители же памяти для ПК делают память большого объема, но не очень быструю, т.е. такую, которая не сможет поддерживать широкополосные операции чтения-записи. Проблема при использовании DRAM состоит в том, что маршрутизаторам необходимо хранить большие объемы информации в памяти, которая к тому же должна быть очень быстрой (способной хранить объем буфера, определяемый двойным временем прохождения сети или задержкой полосы пропускания).

Эти проблемы требуют, чтобы производители маршрутизаторов выбирали наилучших производителей памяти SRAM и DRAM. Для каждого приложения в системе варианты выбора определяются стоимостью, полосой пропускания памяти, габаритными ограничениями, количеством контактов, требуемых для интерфейса с системой и энергопотреблением. Выбор правильного решения крайне важен для устранения узких мест и повышения производительности маршрутизации.

Микропроцессоры общего назначения

Микропроцессоры - это электронные схемы, обычно реализуемые на одном кристалле кремния, выполняющие арифметические, логические и управляющие функции. Обычно они подключаются к шине, которая позволяет им общаться с устройствами ввода-вывода и системной памятью (см. рис. 7). Микропроцессоры конструируются для выполнения специального набора инструкций, задаваемого на языке ассемблера. Язык ассемблера поддерживается аппаратной реализацией, что позволяет микропроцессору выполнять общие операции: сложение, вычитание, сдвиг влево, сдвиг вправо, запись значений в память, чтение значений из памяти.



Рисунок 7. Взаимодействие микропроцессора с памятью и устройствами ввода-вывода

Микропроцессоры функционируют как вычислительные устройства общего назначения, поэтому их можно использовать для решения практически любой задачи. Они применяются как в персональных компьютерах (текстовые процессоры, электронные таблицы, видеоигры и т. п.), так и в суперкомпьютерах (динамика ракет, ядерные реакции, геновая инженерия и т. п.). Заметим, что все эти приложения подразумевают решение математических или вычислительных задач, и что вычисления, выполняемые в Интернет-маршрутизаторах относительно просты в сравнении с ними.

Микропроцессоры – достаточно слабый инструмент поддержки форвардинга IP-пакетов, так как они требуют написания огромного программного кода для выполнения задач, для которых они изначально не предназначались. Это относится к программам, реализующим драйверы физических интерфейсов устройств, программам обработки данных, принимаемых по входному интерфейсу, программам анализа содержимого заголовка пакета и фильтрации пакетов, программам поиска маршрута в таблице форвардинга и продвижения данных к выходному интерфейсу, к программам управления очередями выходного интерфейса.

В добавление к большому объему программного кода, который необходимо написать и выполнить на микропроцессоре, существует ряд ограничений по производительности, которые производители должны преодолеть, создавая маршрутизатор на основе микропроцессора:

- До своего продвижения пакет требуется передать микропроцессору для обработки и получить обработанный пакет обратно. Проблема в том, что пропускная способность микропроцессоров ограничена, так как они не предназначались для поддержки высокой скорости ввода-вывода. Поэтому ввод и вывод пакетов происходит медленно, что ухудшает производительность форвардинга пакетов.
- Микропроцессор должен хранить таблицу форвардинга в оперативной памяти (RAM), так как он не может обслуживать таблицу, содержащую 100000 маршрутов (приблизительный размер таблицы маршрутизации Интернет) в кэше уровня 2. Так как полностью таблица форвардинга не может разместиться в кэше, микропроцессору при выполнении поиска требуется постоянная подкачка данных адресной таблицы, что обуславливает дополнительные задержки форвардинга.

Несмотря на эти ограничения, микропроцессоры играют важную роль в маршрутизаторах Juniper Networks.

- На микропроцессоре подсистемы маршрутизации функционирует ОС JUNOS. Этот микропроцессор отвечает за интерфейс с пользователем, выполнение протоколов маршрутизации, обслуживание таблицы маршрутизации и работу программных средств сетевого управления.
- Все наши системы поддерживают иерархию управления при помощи микропроцессоров, установленных на каждой системной плате. В верхней точке иерархии, в подсистеме маршрутизации, находится микропроцессор Intel, на котором функционирует ОС JUNOS. Следующий уровень иерархии содержит микропроцессоры, работающие в разных частях системы. Они обеспечивают управление локальными устройствами на отдельных системных платах.

Перепрограммируемые вентиляльные матрицы (FPGA)

FPGA - это компоненты, которые производятся изготовителями ИС в очень больших количествах. Когда FPGA поступают от производителя, связи между логическими элементами отсутствуют, они должны программироваться изготовителем системы. Изготовитель системы пишет код, который загружается в FPGA и объединяет логические элементы FPGA в заданные комбинации для выполнения конкретных функций. Теоретически FPGA обеспечивают огромную гибкость, так как их можно перепрограммировать путем простой загрузки новой программы.

FPGA имеют ограниченное количество логических элементов по сравнению с микропроцессорами общего назначения, т.е. они являются превосходным инструментом для выполнения задач с относительно узкой специализацией. Хотя FPGA неспособны соревноваться с микропроцессорами в сложности задач, свои узкоспециализированные функции они выполняют значительно быстрее.

В наших системах FPGA играют достаточно ограниченную роль. Основное их применение в маршрутизаторах Juniper Networks – обеспечение взаимодействия наших специализированным ИС с ИС, полученными от сторонних производителей. Предположим, что готовая ИС, реализующая протокол Ethernet, должна связаться с нашей заказной ИС, выполняющей роль менеджера ввода-вывода. ИС FPGA, включенная между двумя этими компонентами, переводит язык ИС Ethernet в язык специализированной ИС менеджера ввода-вывода и позволяет связываться этим компонентам между собой. FPGA также применяются для регулировки напряжения питания печатных плат и мониторинга системных вентиляторов.

Специализированные ИС

Специализированные ИС (ASIC) для Интернет-маршрутизаторов - это компоненты, производимые изготовителями ИС в относительно малых объемах. Они содержат больше логических элементов на кристалле, чем FPGA и работают с более высокими тактовыми частотами. Так как специализированные ИС не являются в общем случае перепрограммируемыми, гибкость достигается за счет тщательного проектирования и применения микрокода.

Проектирование и разработка заказных ИС

Использование микрокода позволяет продлить срок развертывания аппаратных маршрутизаторов от трех до пяти лет. Большое количество аппаратных средств в сочетании с маленьким микрокодом может обеспечить огромную гибкость. Вообще, ИС с микрокодом можно рассматривать как мини-процессор, реализующий очень специфический набор функций (рис. 8).

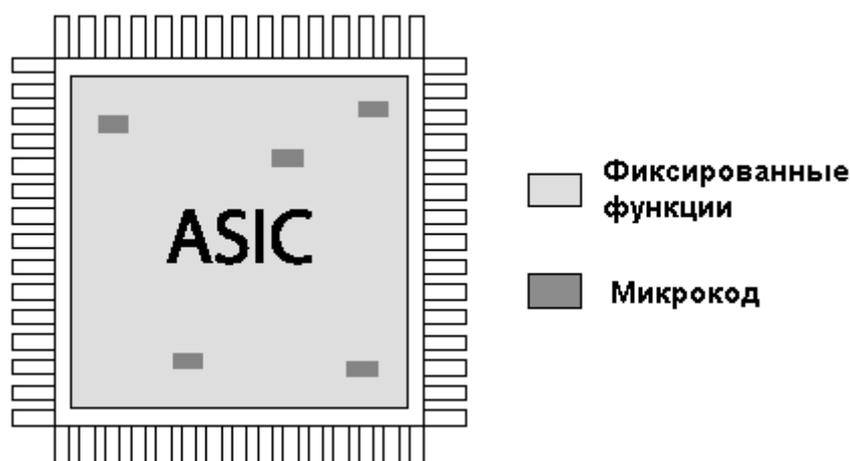


Рисунок 8. ИС с микрокодом

Предположим, что производителю нужна специализированная ИС для выполнения 200 различных функций. Из них 185 запрограммированы жестко и никогда не будут изменяться, а оставшиеся 15 (такие, как изменение значений в заголовке PPP, указывающих на тип протокола IPv4 либо IPv6)

требуется сделать гибкими для адаптации к появляющимся технологиям. Вместо аппаратного кодирования этих 15 функций, производитель может закодировать их в специализированной ИС с использованием микрокода, чтобы продлить срок эксплуатации продукта.

Конечно, для выбора функций, реализуемых в микрокоде, производитель должен обладать опытом и предвидением. Кроме того, абсолютно необходимо, чтобы ИС была изначально спроектирована для выполнения в микрокоде именно этих функций. При проектировании специализированной ИС также задается набор инструкций микрокода, и если в будущем потребуются некая инструкция, отсутствующая в первоначальном наборе, единственным выходом будет создание совершенно новой ИС с требуемой функциональностью.

Для проектирования высокопроизводительных ИС нужна сплоченная команда, члены которой должны обладать редким сочетанием опыта разработки ИС, знания программного обеспечения Интернета и практического опыта в создании сетей, а также необходимо наличие рабочей среды, облегчающей тесную связь и свободный обмен идеями между отдельными группами сотрудников. Потребовалось время, чтобы собрать команду старшего инженерного персонала и отработать процесс, позволяющий различным группам гармонично сотрудничать. Опыт, накопленный во время проектирования и разработки первого набора специализированных ИС, помогает сократить время разработки и избежать ошибок при последующих разработках.

Роль производителя специализированных ИС

Производитель ИС предлагает производителю маршрутизатора на выбор несколько вариантов технологических процессов, которые определяют плотность размещения логических элементов и число выводов ИС. Чем ближе располагаются проводники друг к другу, тем ближе друг к другу можно разместить транзисторы, тем больше логических схем может быть реализовано в единице объема. Для Juniper Networks специализированные ИС изготавливает компания IBM.

Производитель маршрутизаторов выбирает наилучшего партнера с учетом скорости работы производимых им компонент, степени интеграции, энергопотребления и количества выводов ИС. Это ответственное решение, так как оно может определить успех или неудачу всего проекта. Например, производитель маршрутизатора может превосходно спроектировать логическую часть, и технологический процесс одного производителя ИС позволит создать замечательную ИС, однако процесс второго производителя может привести к получению неработоспособных ИС.

Концепция, которая определяет выполнение функций ИС, разрабатываются производителем маршрутизатора. Он же проектирует топологию и затем, используя инструментарий изготовителя ИС, согласует ее с соответствующей технологией. Далее изготовитель ИС использует эту топологию для изготовления кристалла, устанавливает его в корпус, проводит основное тестирование и затем возвращает ИС производителю маршрутизатора для окончательной интеграции в систему.

Применение специализированных ИС в Интернет-маршрутизаторах

Основное применение специализированные ИС в Интернет-маршрутизаторах – это выполнение сложных функций обработки пакетов. В старых корпоративных маршрутизаторах на основе микропроцессоров эти функции выполнялись программно, что серьезно снижало производительность форвардинга пакетов. В первом поколении Интернет-маршрутизаторов на основе специализированных ИС основные функции форвардинга (обработка входящих пакетов, поиск маршрута, обработка исходящих пакетов) выполнялись аппаратно, тогда как все функции создания услуг (фильтрация, списки управления доступом, статистическая выборка, правила управления трафиком, ограничение скорости, учет) были реализованы программно. Это означало, что пользователи должны выбирать между производительностью форвардинга и предоставлением услуг. В современном поколении Интернет-маршрутизаторов на основе специализированных ИС и форвардинг пакетов, и функции создания услуг выполняются аппаратно для удовлетворения постоянно растущих требований по производительности как на границе сети, так и в магистральном ядре Интернета.

Мы всегда сами проектировали ИС для обработки кадров SONET, так как наша ПФП всегда работает на пределе производительности, и производители SONET примерно на год отстают от наших требований по производительности интерфейсов. Мы будем продолжать собственное проектирование этих ИС, чтобы обеспечить соответствие наших высокоскоростных интерфейсов и производительности обработки пакетов в магистральном ядре. Кроме ИС обработки кадров SONET мы разработали ряд специализированных ИС для магистрального ядра, включая ИС ввода-вывода (обработка входящих и исходящих пакетов), ИС управления распределенным

буфером (управление памятью), ИС Internet Processor II (поиск маршрута и создание услуг) и ИС директора пакетов (внутреннее управление трафиком). Мы постоянно совершенствуем наши ИС и разрабатываем новые для ПФП всех наших систем.

Специализированная ИС Internet Processor II

ИС Internet Processor II является вторым поколением специализированных Интернет-процессоров (ИП). Мы усвоили уроки разработки первого ИП, что позволило ускорить разработку кристаллов второго поколения.

Функции Интернет-процессора

Первый ИП Internet Processor был спроектирован как процессор обработки пакетов для ядра Интернета, обеспечивающий высокую скорость поиска (свыше 40 миллионов пакетов в секунду). Благодаря надежной производительности ИП Internet Processor быстро стал основным элементом в магистралях самых крупных поставщиков услуг. ИС Internet Processor была и продолжает оставаться самым быстрым и наиболее универсальным процессором поиска маршрута в истории нашей отрасли.

Функции ИС Internet Processor II

ИС Internet Processor была разработана для устранения узких мест в магистральных сетях поставщиков услуг. Постепенно узкие места стали образовываться на границе сети. Поставщики услуг требовали от нас обеспечения безопасности, управления трафиком, обзора сети, возможностей создания новых услуг на сетевой границе. Кроме того, они хотели, чтобы эти функции были реализованы аппаратно для снижения влияния на производительность форвардинга пакетов.

Для удовлетворения этих требований мы спроектировали специализированную ИС Internet Processor II - процессор поиска маршрута и создания услуг для магистрального ядра Интернета и сетевой границы (где требуются развитые функции обработки пакетов). Однако мы не проектировали Internet Processor II в соответствии с принципом разделения трактов обработки на быстрые и медленные. В Internet Processor II весь трафик проходит через один и тот же аппаратный тракт обработки пакетов, который не влияет существенно на производительность форвардинга при создании услуг. В результате, Internet Processor II устраняет необходимость компромиссов между созданием услуг на границе и производительностью форвардинга.

ИС Internet Processor II выполняет три основные функции, известные как *примитивы*: поиск маршрута, поиск в таблицах, фильтрацию.

- Примитив поиска маршрута выполняет поиск самого длинного согласующегося префикса в таблице маршрутизации или просматривает список префиксов в программе фильтрации.
- Примитив поиска в таблице определяет необходимость выполнения другого примитива. Например, поиск в таблице по входному интерфейсу пакета может определить, требуется ли для данного пакета запуск фильтра межсетевого экрана.
- Программы фильтрации классифицируют трафик и затем выполняют некоторые функции для всех потоков, согласующихся с данным классом трафика. Программы фильтров не реализуются в микрокоде; они больше похожи на программы на языке высокого уровня, скомпилированные и оптимизированные для работы в ИС Internet Processor II. Программы фильтров могут быть исключительно гибкими, они используют комплексный набор инструкций, специально разработанный для выполнения операций фильтрации в ИС Internet Processor II.

Эти примитивы проектировались для выполнения конкретных задач. Так как ИС Internet Processor II не является процессором общего назначения, она не может выполнять функции графического процессора или выполнять приложения на ПК. Это специализированный сетевой процессор

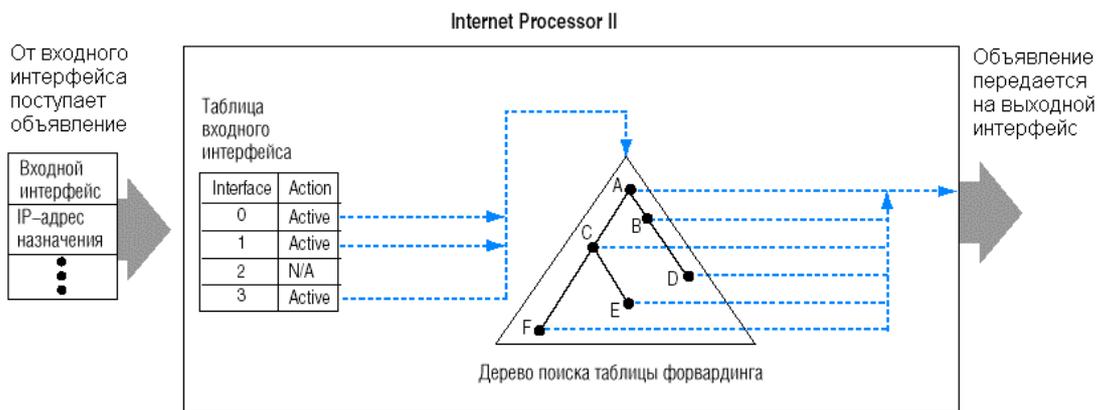


Рисунок 9. Специализированная ИС Internet Processor II - два примитива

Рис. 9 иллюстрирует обработку объявлений пакетов в случае, если бы Juniper Networks развернула в тракте обработки пакетов ИС Internet Processor II только два примитива. ИС Internet Processor II спроектирована для обработки *объявлений пакета (packet notification)*, содержащих всю необходимую информацию, переносимую в полях заголовка соответствующего пакета.

- Предположим, что объявление пакета поступило на Интерфейс 3, и адрес назначения согласуется с Префиксом С. Примитив поиска в таблице входного интерфейса определяет, что Интерфейс 3 активен, поэтому объявление передается примитиву поиска маршрута для выполнения поиск самого длинного согласующегося префикса в таблице форвардинга. На основе информации, полученной при поиске маршрута, объявление отправляется в очередь пакетов выходного интерфейса.
- Предположим, что пакет поступил на Интерфейс 2 и адрес назначения согласуется с Префиксом F. Примитив поиска в таблице входного интерфейса определяет, что Интерфейс 2 неактивен, поэтому пакет отбрасывается, и примитив поиска не выполняется.

Способность выполнять различные примитивы в любой последовательности является источником мощности и гибкости ИС Internet Processor II. Наши инженеры используют встроенное программное обеспечение для программирования и управления основной структурой и последовательностью примитивов обработки пакетов, поддерживаемых каждой версией ОС JUNOS. Встроенное программное обеспечение позволяет добавлять примитивы в цепочку обработки, удалять примитивы, выполнять примитивы в любой комбинации, в любом порядке, фактически для любого типа пакетов. Выгода для поставщиков услуг при этом очевидна. Они просто модернизируют свои системы, устанавливая новую версию ОС JUNOS, и новые программные функции моментально получают поддержку аппаратной подсистемы форвардинга.

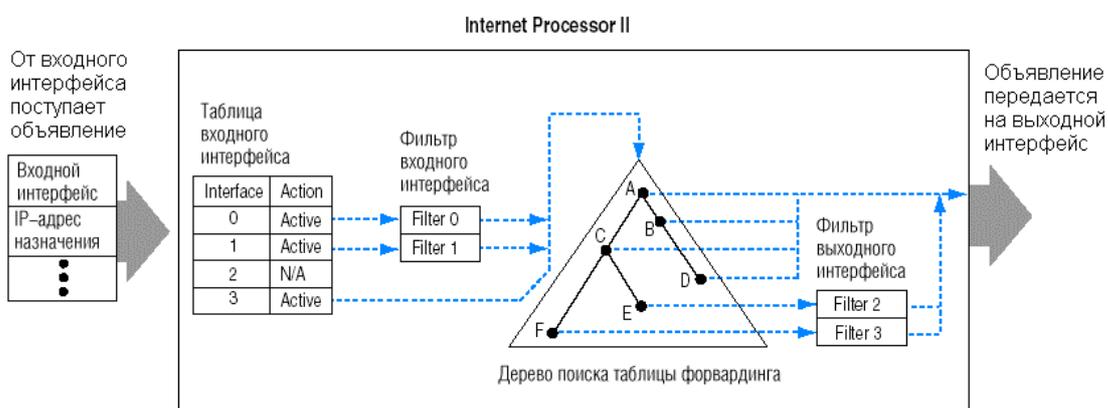


Рисунок 10. Internet Processor II - множество примитивов

Рис. 10 иллюстрирует использование возможностей обработки пакетов ИС Internet Processor II для поддержки предоставления новых услуг за счет введения дополнительных примитивов в тракт обработки пакетов.

- Предположим, что пакет поступает на Интерфейс 3, и адрес назначения согласуется с Префиксом F. Примитив поиска в таблице входного интерфейса определяет, что Интерфейс 3 активен, поэтому объявление передается примитиву поиска маршрута для поиска самого длинного согласующегося префикса в таблице форвардинга. На основе информации, полученной при поиске маршрута, для объявления срабатывает выходной фильтр 3. Если объявление пакета принимается выходным фильтром 3, объявление отправляется в очередь передачи на выходной интерфейс.
- Предположим, что пакет поступает на Интерфейс 1, и адрес назначения согласуется с Префиксом E. Примитив поиска в таблице входного интерфейса определяет, что интерфейс 1 активен. На основе поиска в таблице входного интерфейса для объявления пакета выполняется входной фильтр 1. Если объявление пакета удовлетворяет входному фильтру 1, объявление пропускается к примитиву поиска маршрута для выполнения поиска согласующегося префикса максимальной длины в таблице форвардинга. На основе информации, полученной при поиске маршрута, для объявления пакета выполняется выходной фильтр 2. Если объявление пакета удовлетворяет выходному фильтру 2, объявление отправляется в очередь передачи на выходной интерфейс.
- Предположим, что пакет поступает на Интерфейс 0, и адрес назначения согласуется с префиксом C. Примитив поиска в таблице входного интерфейса определяет, что Интерфейс 0 активен. На основе поиска в таблице входного интерфейса, для объявления пакета выполняется входной фильтр 0. Если объявление пакета удовлетворяет входному фильтру 0, объявление пропускается к примитиву поиска маршрута для поиска согласующегося префикса максимальной длины в таблице форвардинга. На основе информации, полученной при поиске маршрута, объявление отправляется в очередь передачи на выходной интерфейс.

Специализированная ИС Internet Processor II может поддерживать только те функции, которые мы явно встроили в кристалл. Такая гибкость потребовала огромных затрат по планированию. Мы одновременно поддерживаем четыре различных протокола для задач поиска, можно анализировать любые поля в заголовке пакета, можно выбирать примитивы для выполнения, можно программировать примитивы для выполнения в различном порядке, можно определять следующий выполняемый примитив на основе результатов выполнения предыдущего примитива. ИС Internet Processor II обеспечивает исключительную гибкость только потому, что наша команда проектировщиков смогла предвидеть будущие требования и встроила необходимые функции в основу ИС Internet Processor II.

Гибкость специализированной ИС Internet Processor II

Juniper Networks продолжает демонстрировать мощь и гибкость ИС Internet Processor II, обеспечивая одновременную аппаратную поддержку новых программных функций ОС JUNOS. Начальная версия ОС JUNOS поддерживала ИС Internet Processor II с аппаратной поддержкой фильтров межсетевых экранов (принять, отвергнуть, отбросить, подсчитать, зарегистрировать), статистической выборки и жестко заданным распределением нагрузки по пакетам. Последующие версии ОС JUNOS продемонстрировали нашу способность осуществления немедленной аппаратной поддержки новых функций программного обеспечения.

- Правила управления трафиком позволяют поставщикам услуг классифицировать пакеты и присваивать различным потокам пакетов различные пороговые значения.
- Усовершенствование выборки позволяет агрегировать выбранный трафик и отправлять информацию потока в формате **cflowd** на удаленный хост или станцию управления.
- Модернизированный фильтр межсетевого экрана с функцией *alert* (*предупреждение*), вызывающей регистрацию условий фильтра пакета в системном лог-файле (*syslog*), который затем можно передать на сервер составления отчетов.
- Новая функция класса назначения (*destination-class usage*, DCU) поддерживает расширенные возможности биллинга и учета, что позволяет поставщикам услуг

поддерживать счетчики пакетов в точках входа и выхода трафика, передаваемого через их сети.

- Форвардинг на основе фильтров поддерживает требования *открытого доступа* поставщиков услуг, обеспечивая возможность продвижения пакета от заданного физического или логического интерфейса на основе адреса источника пакета.
- Поддержка множества отдельных копий таблицы маршрутизации обеспечивает аппаратный форвардинг для виртуальных частных сетей (VPN) в соответствии с рекомендацией RFC 2547bis в пограничных маршрутизаторах провайдера (provider edge, PE).

Все эти новые функции поддерживаются специализированной ИС Internet Processor II без заметного ухудшения производительности форвардинга пакетов в наших системах, работающих в ваших сетях.

Проектирование аппаратных маршрутизаторов-долгожителей

Уникальная комбинация аппаратно-программной архитектуры Juniper Networks позволяет создавать исключительно гибкие, надежные и производительные системы форвардинга пакетов. Гибкость наших аппаратных маршрутизаторов позволяет им адаптироваться к постоянно меняющейся среде Интернета и обеспечивает исключительно большие сроки эксплуатации.

Сложное искусство системного проектирования

Интернет-маршрутизаторы - исключительно сложные системы, от которых требуется очень быстрое масштабирование по разным направлениям для адаптации к постоянно меняющимся условиям. Эти направления включают общую емкость полосы пропускания системы и скорость, с которой пакеты должны продвигаться по системе, а также количество маршрутов в таблице маршрутизации, скорость поиска маршрута, различные физические интерфейсы, фильтры, правила маршрутизации, соединения пиринга.

Необходимость разносторонней технической подготовки

Для построения систем, работающих в условиях Интернета, требуются глубокие знания по многим технологиям.

- Программное обеспечение.
- Проектирование и оптимизация ядра системы.
- Проектирование и реализация протоколов маршрутизации.
- Сетевое управление и проектирование пользовательских интерфейсов.
- Проектирование встроенных систем.
- Аппаратные средства:
 - проектирование кристаллов и корпусов ИС;
 - проектирование печатных плат;
 - проектирование высокоскоростных схем;
 - разъемы;
 - память;
 - оптическая электроника.
- Конструирование:
 - электропитание;
 - конструкция корпуса;
 - охлаждение;

- электромагнитная совместимость.

Вследствие сложности процесса разработки проектирование и создание новой платформы маршрутизации требует минимум двух лет. Из-за того, что все технологии развиваются независимо от других и с разной скоростью, ошибки в прогнозировании их развития привести к неудаче проекта.

Разделение системных функций - нетривиальная задача

Задача разделения функций между аппаратными и программными средствами включает два этапа. На первом этапе определяется, какие функции будут реализованы аппаратно, а какие - программно. На втором этапе, после анализа доступных вариантов аппаратных средств, производитель решает, какие части аппаратуры должны программироваться жестко, а какие - необходимо сделать перепрограммируемыми. На этом этапе производитель маршрутизатора должен учесть следующие факторы.

- Конкретные функции потребуют максимальной производительности, поэтому они должны жестко программироваться в специализированных ИС.
- Очень конкретные части этих функций должны контролироваться с использованием микрокода, что обеспечит гибкость, гарантирующую потребителям защиту инвестиций от устаревания аппаратуры.
- Некоторые функции обработки пакетов требуют аппаратной производительности наряду с большей гибкостью, чем та, которую может обеспечить микрокод. Эти функции реализуются в узкоспециализированных подсистемах обработки пакетов, таких как Internet Processor II.
- Для некоторых частей системы производитель может не захотеть тратить время и деньги на разработку заказных ИС, а использовать перепрограммируемую вентильную матрицу или стандартные ИС.
- Некоторые компоненты системы будет необходимо реализовать с использованием микропроцессоров или иерархической структуры микропроцессоров для обеспечения наивысшей гибкости, которую можно получить только используя программные средства.

Разделение системы требует талантливой команды, имеющей знания на уровне экспертов по многим дисциплинам для принятия оптимальных проектных решений в процессе разработки стабильной, масштабируемой, высокопроизводительной и адаптируемой системы.

Системная интеграция

Так как производитель имеет неограниченное число вариантов разделения системы, и все функциональные компоненты должны быть связаны друг с другом для работы в составе единого комплекса, процесс их интеграции с соблюдением условий системной стабильности и высокой производительности является исключительно ответственным. Системная интеграция требует от производителя дара предвидения и таланта для выполнения правильных действий с правильными компонентами, с использованием правильно выбранных технологий и каналов связи между отдельными компонентами. Неудача на этом этапе приведет в результате к некорректно работающей системе, неспособной надежно предоставлять услуги в непредсказуемых, а иногда и враждебных условиях Интернета.

Системная архитектура Juniper Networks

На рис. 11 представлена основная архитектура наших систем.

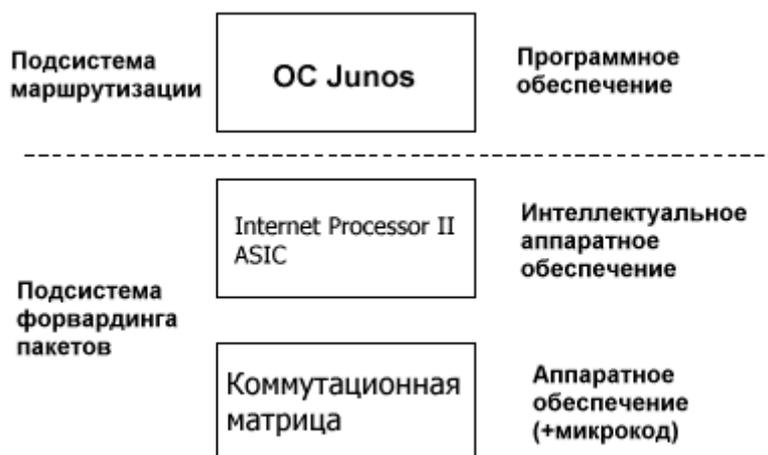


Рисунок 11. Системная архитектура Juniper Networks

Нижний уровень: коммутационная матрица

Коммутационная матрица аппаратно реализует тракт форвардинга, который позволяет пакетам проходить через систему от входного интерфейса к выходному. Она выполняет следующие функции.

- Удаляет инкапсуляцию уровня 2 и осуществляет проверку контрольной суммы.
- Записывает пакет в системную память.
- Генерирует объявление пакета.
- Продвигает объявление пакета в ИС Internet Processor II для обработки.
- Продвигает объявление пакета после обработки в ИС Internet Processor II к интерфейсам ввода-вывода.
- Помещает объявление пакета в выходную очередь.
- Выполняет уменьшение счетчика времени жизни (TTL) IP пакета.
- Выполняет фрагментацию пакета при несоответствии с максимальным размером кадра для данной среды передачи (MTU).
- Считывает пакет из системной памяти.
- Формирует кадр уровня 2 на выходном интерфейсе.

Наши коммутационные матрицы построены на основе специализированных ИС, так как свои функции им необходимо выполнять в реальном времени со скоростью линии. Сама обработка остается практически постоянной, но та часть, которая может измениться, реализуется при помощи микрокода. Мы спроектировали коммутационную матрицу как основу общей подсистемы форвардинга, которую можно запрограммировать для продвижения пакета любого типа. Гибкость, обеспечивается различными элементами аппаратной подсистемы форвардинга за счет использования микрокода.

- ИС управления вводом-выводом может программироваться для распознавания различных типов кадров, включая IPv4, IPv6, Frame Relay, MPLS и IPX.
- ИС управления распределенным буфером может программироваться для поиска в любой части заголовка пакета для извлечения информации форвардинга и построения объявления пакета. Процедуры поиска полностью программируемы, и объявление пакета может быть сформировано в любом необходимом формате.

Хотя эти возможности и не видны нашим клиентам, они специально разработаны для обеспечения гибкости модификации компонент коммутационной матрицы при развитии стандартов, без необходимости замены уже находящейся в эксплуатации аппаратуры.

Средний уровень: специализированная ИС Internet Processor II

Специализированная ИС Internet Processor II выполняет и поиск маршрута, и расширенную обработку пакетов, требуемую при предоставлении IP-услуг. Подобно ИС коммутационной матрицы, ИС Internet Processor II проектировалась как общая система поиска маршрута и обработки пакетов, которую можно запрограммировать для пакетов любого типа. В частности, ее можно запрограммировать для выполнения поиска в соответствии с четырьмя различными протоколами, которые будут поддерживаться одновременно. Добавление поддержки новых протоколов требует следующего:

- разработки нового программного обеспечения маршрутизации для подсистемы маршрутизации;
- привязки новой таблицы форвардинга к ИС Internet Processor II;
- добавления новых функций в коммутационную матрицу для поддержки нового протокола с использованием микрокода, заложенного при проектировании в наши специализированные ИС.

ИС Internet Processor II спроектирована для обеспечения максимальной производительности, что гарантируется использованием при обработке пакетов минимального количества циклов поиска и обращений к памяти. ИС Internet Processor II имеет запас по производительности для своих задач, более чем достаточный для выполнения дополнительных функций обработки пакетов в дополнение к стандартному поиску маршрута с согласованием по префиксу максимальной длины, необходимого для базовой процедуры форвардинга пакетов. ИС Internet Processor II подтвердила свою высокую стабильность и производительность при работе в магистрали и на границе крупнейших сетей поставщиков услуг.

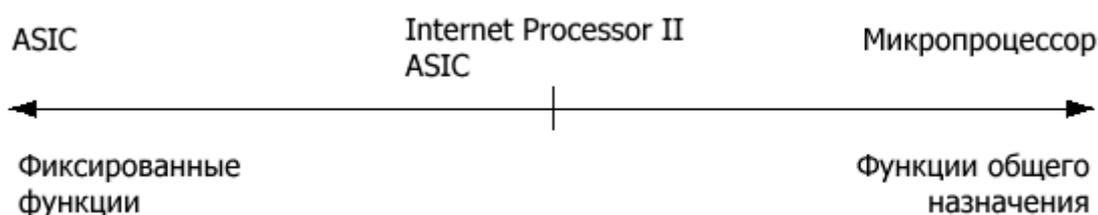


Рисунок 12. Гибкость ИС Internet Processor II

ИС Internet Processor II, работающая под управлением ОС JUNOS представляет собой идеальное сочетание программно-аппаратных средств для создания инструментов гибкой и высокопроизводительной обработки пакетов. На рис. 12 показана ось координат, на одном конце которой находятся специализированные ИС с фиксированным набором функций (использующие жесткую логику), на другом - микропроцессоры общего назначения (с примитивами, поддерживающими общие арифметические операции). ИС Internet Processor II располагается в центре этой оси, так как она обеспечивает и высокопроизводительную обработку, характерную для специализированных ИС, и гибкость микропроцессоров общего назначения при выполнении специализированных функций обработки пакетов.

Верхний уровень: ОС JUNOS

ОС JUNOS обеспечивает интерфейс с пользователем, поддержку протоколов маршрутизации и сетевого управления. Эти функции развиваются опережающими темпами, поэтому мы реализовали их программно на микропроцессоре общего назначения для обеспечения максимальной гибкости. Разработчики ПО знают, что если программные средства ограничивают производительность всей системы, то для увеличения производительности системы надо уменьшать число инструкций выполняемых за одну операцию. Это обычно достигается за счет удаления проверок надежности и устойчивости в программных средствах, однако удаление этих инструкций приводит к уязвимости ПО. Если ПО функционирует на участке форвардинга пакетов,

необходимо создавать вызовы (shortcuts) для различных функций (например, жесткий код), что противоречит идее разработки четко разделенных и стабильных интерфейсов между этими функциями. Монолитные коды изначально нестабильны и трудно поддаются модернизации, т.к. изменения в одной части кода могут привести к ошибкам в другой части.

Надежность систем IP базируется на надежности ПО, и мы достигли этой надежности путем четкого разделения функций между аппаратными и программными средствами. Этот подход позволяет нам не использовать ПО в тракте критичного к задержкам форвардинга пакетов. В результате, ОС JUNOS обеспечивает большую гибкость (с дополнительной надежностью), так как программный код прозрачен и хорошо структурирован. Это обеспечивает нам возможность быстрого добавления новых функций и позволяет выпускать в год четыре новые версии. Другие производители неспособны с такой скоростью добавлять новые функции, так как их ПО имеет жестко заданный код и функционирует на платформах, работающих на пределе возможной производительности.

Наконец, наше модульное ПО позволяет иметь единую последовательность разработки ОС JUNOS. Если вы - производитель, пытающийся поддерживать жестко заданный код, работающий на пределе производительности, то добавление новой программной функции может легко привести к неработоспособности уже существующих программных функций. В результате приходится создавать ответвление в последовательности разработки программ, так как производителю приходится создавать две разные версии для поддержки двух различных функций. Со временем это приводит к множеству "последних" версий, каждая из которых лишь немного отличается набором функций для различных комбинаций интерфейсов и аппаратных платформ.

Повышение гибкости архитектуры с использованием сервисных интерфейсных карт (PIC)

Наша команда опытных инженеров-разработчиков аппаратных и программных средств создала исключительно гибкую и разделенную систему, которая позволила гарантировать, что задачи, которые требуется выполнять со скоростью линии, будут реализованы аппаратно, на нижнем уровне, а преимущества микрокода добавят необходимую гибкость. Функции выбора маршрута и интеллектуальной обработки пакетов для создания услуг выполняются на среднем уровне, реализуемом специализированной ИС Internet Processor II, объединившей в себе лучшие программные и аппаратные средства. Наконец, функции, требующие максимальной гибкости, выполняются на верхнем уровне, в ОС JUNOS. В результате разделения нашей системы с использованием правильного сочетания аппаратных средств, микрокода и программного обеспечения для каждой задачи была создана система более гибкая, чем любая система других производителей.

Для дальнейшего улучшения гибкости Juniper Networks разработала сервисные интерфейсные карты PIC. На рис. 13 показано, как концепция PIC соотносится с архитектурой маршрутизатора.

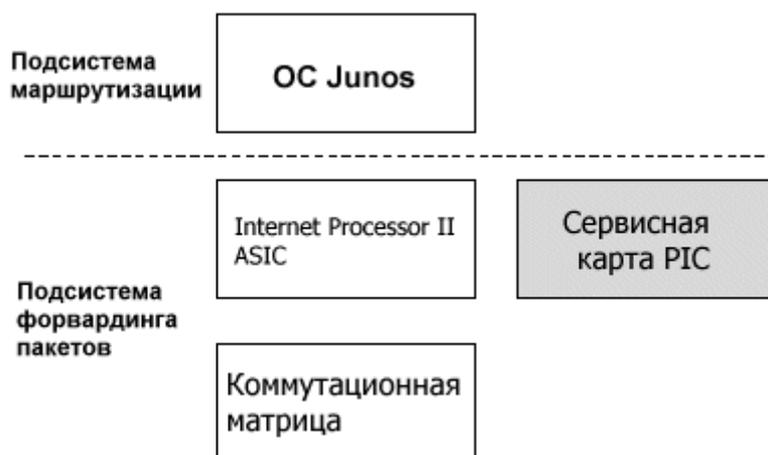


Рисунок 13. PIC в архитектуре маршрутизатора Juniper Networks

Центральная часть наших систем выполняет те функции, которые магистральные Интернет-маршрутизаторы должны выполнять фактически для всех пакетов, проходящих через систему. Есть определенный смысл возложить такие функции как анализ на уровне 2 и уровне 3, поиск

маршрута, выделение таблиц форвардинга ВЧС, обслуживание очередей пакетов и т.п. на системное ядро, так как они касаются каждого пакета, проходящего по системе. Однако функции создания туннелей, ML/PPP, шифрования пакетов и т. п. не требуются для каждого пакета. В действительности может получиться так, что эти специализированные услуги вообще не потребуются для проходящих по системе пакетов. Следовательно, нет особого смысла встраивать их в ядро, так как, даже если они и не ухудшат производительность, то сделают центр системы более сложным.

Поэтому наше проектное решение состоит в использовании ядра для скоростной обработки всех пакетов с последующим принятием решения о необходимости специализированной обработки пакета посредством PIC. На сервисных PIC реализован требуемый набор функций с использованием тех технологий, которые соответствуют конкретному решению. Для карты PIC, выполняющей шифрование пакетов, можно использовать аппаратные средства. Для карты ML/PPP PIC использованы специализированные FPGA. На карте PIC можно даже разместить рабочую станцию для программного выполнения конкретных функций, необходимых для обработки трафика определенного класса. Этот подход позволяет выйти на совершенно новый уровень гибкости архитектуры Интернет-маршрутизаторов.

В системном центре несколько очень простых механизмов определяют, требуется ли конкретному пакету специализированная обработка при помощи PIC. Для входящих услуг, таких как ML/PPP или дешифрация, эту необходимость (по входному интерфейсу) может определить IC Internet Processor II. Для исходящих услуг, таких как ML/PPP (с объединением множества выходных интерфейсов в один логический интерфейс) или шифрование трафика на заданном интерфейсе, решение об использовании PIC является естественным этапом процесса поиска маршрута, определяющего выходной интерфейс. Зная выходной интерфейс, IC Internet Processor II может вводить PIC в тракт форвардинга определенных пакетов, направляя пакет в соответствующую карту PIC (с дополнительной информацией о конкретном интерфейсе ML/PPP или шифруемом туннеле).

Поддержка функций на границе системы с использованием сервисных PIC обеспечивает огромную гибкость и для наших клиентов, и для нас.

- Вы свободны в принятии решений. Например, при заданном множестве сетевых интерфейсов, необходимых вашему приложению, и количестве поддерживаемых услуг, вы можете выбрать нужный вариант шасси и установить в нем только необходимые сетевые интерфейсы с требуемой емкостью для поддержки услуги.
- Мы можем продолжать поддержку важных функций без неблагоприятного воздействия на производительность системного ядра, без увеличения риска и сложности, не требуя от клиента выбрать конкретный набор сетевых интерфейсов или услуг, и не заставляя его платить за ненужные услуги.
- Конечно, сервисные карты PIC – это определенные затраты, как с точки зрения капиталовложений, так и по плотности портов маршрутизатора. Однако мы считаем их оправданными, так как гибкость, обеспечиваемая картами PIC, работает во всех направлениях постепенного расширения системы. Например: вам нужна конкретная услуга, предсказуемо работающая в действующей сети, а не только в тщательно контролируемых лабораторных условиях. Когда вы приобретаете сервисную карту PIC с портами OC-12/STM-4 для протокола MP/PPP, вы действительно получаете производительность OC-12/STM-4 для MP/PPP. Это позволяет вам стабильно оказывать услуги с помощью карты PIC. Если вам не нужна эта служба, то вы получаете дополнительный слот PIC с полосой пропускания OC-12/STM-4.

Заключение

Проектирование и разработка аппаратных Интернет-маршрутизаторов требует глубокого понимания искусства и науки построения компьютерных сетей. Как производители маршрутизаторов, мы считаем, что поставляем нашим потребителям работоспособную сеть, а не набор отдельных систем или сетевых элементов. Мы выбрали этот путь, так как именно маршрутизаторы обеспечивают уровень управления сетью, отвечающий за доставку всего трафика. Наличие узких мест в маршрутизаторах или их отказ существенно ухудшают работоспособность ваших сетей и возможность поддержки критичных приложений.

Наши системы отличаются от систем конкурентов по многим параметрам.

- Производительность аппаратного форвардинга пакетов поддерживает интерфейсы с самой большой в отрасли полосой пропускания.
- Уникальная стабильность программных и аппаратных средств, обусловленная архитектурой, специально разработанной для масштабирования и работы в постоянно меняющихся условиях Интернета.
- Высокая плотность полосы пропускания и низкое энергопотребление, критичные в условиях быстрого роста трафика и высоких цен на аренду и электроэнергию.
- Конфигурационная гибкость, обусловленная исключительно широким диапазоном физических интерфейсов, - от 1,5 Мбит/с до 10 Гбит/с.
- Агрегированная производительность форвардинга, которая позволяет без опасений установить плату физического интерфейса в слот PIC с уверенностью в предсказуемой производительности.
- Специализированная ИС Internet Processor II позволяет гарантировать обслуживание без заметного ухудшения производительности форвардинга. ИС Internet Processor II поддерживает классификацию пакетов, фильтрацию, правила управления трафиком, формирование трафика, статистику и учет в магистрали и на границе сети.
- Один и тот же программный код функционирует на всех платформах и обеспечивает беспрецедентную скорость разработки функций, позволяющую создавать четыре версии ОС JUNOS в год.
- Специализированная архитектура, обеспечивающая системную гибкость для продления срока эксплуатации наших систем в вашей сети, позволяет непрерывно улучшать их возможности путем простого обновления программного обеспечения, а не за счет дорогого изменения аппаратной конфигурации.

Сегодня граница сети Интернет становится похожей на магистральное ядро, каким оно было три года назад, - те же скорости и протоколы, но больше трафика и больше серьезных проблем, таких как безопасность и необходимость оказания услуг по высокоскоростным каналам доступа. Маршрутизаторы уровня предприятий со своей устаревшей архитектурой и жестким программным кодом больше неприменимы в таких условиях. И ядро, и граница новой IP инфраструктуры требуют специализированных систем, предусматривающих масштабируемость для удовлетворения эксплуатационных требований клиентов. Управление полосой пропускания магистрали в ядре или на уровне распределения, выделенный доступ на границе сети или магистрали городских сетей, Web хостинг или внедрение новых услуг, - везде Juniper Networks установила стандарт для гибких и высоконадежных Интернет-маршрутизаторов, имеющих устойчиво высокую производительность и позволяющих стабильно предоставлять IP услуги. Цель Juniper Networks - постоянно помогать поставщикам услуг поддерживать лидерство в бизнесе, увеличивая рентабельность за счет внедрения новых дифференцированных услуг и уменьшения общей стоимости владения сетью.

Copyright © 2002 Poplar systems

Poplar
systems

<http://www.poplar.ru>