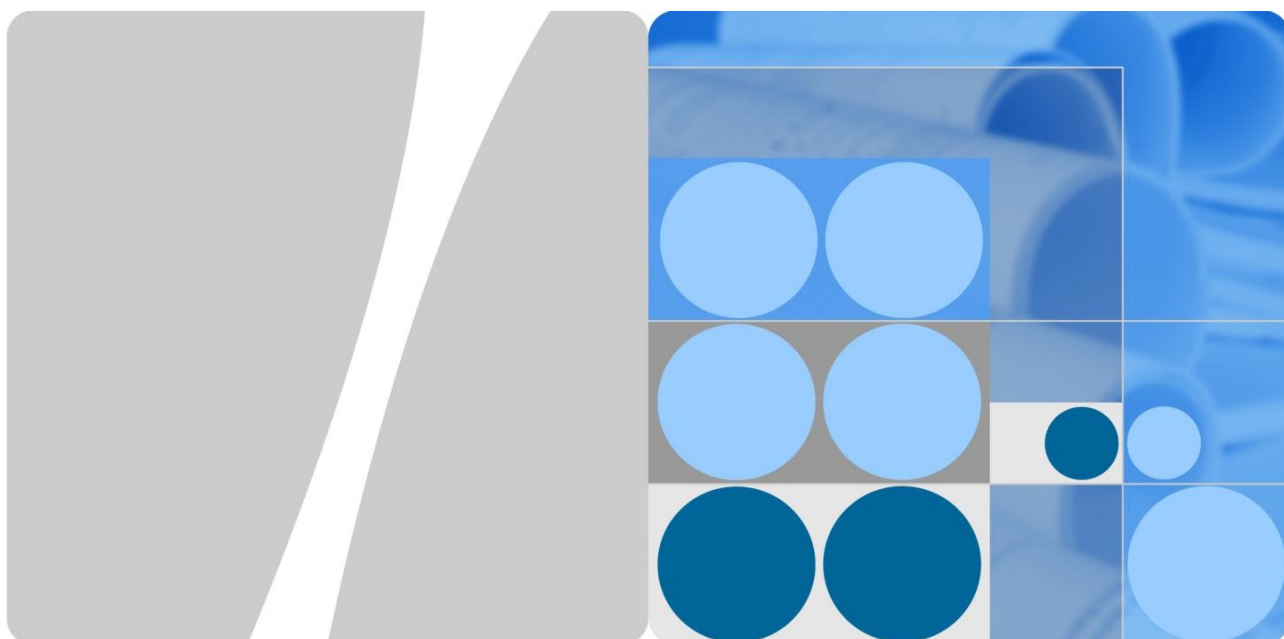




Ethernet Network Processor: Архитектурные инновации для реализации SDN в Agile-коммутаторах Huawei S12700

Материалы документа предназначены для технических специалистов, а также будут интересны продавцам, маркетологам и руководителям компаний-партнеров Huawei Enterprise для ознакомления с инновационной разработкой Huawei – программируемым сетевым процессором Ethernet ENP (Ethernet Network Processor), на котором реализована функциональность Agile-коммутаторов Huawei S12700.

Документ может передаваться клиентам.

Множество показателей, таких как пропускная способность, задержки, управляемость и безопасность используются для оценки производительности сетевых устройств. Среди других показателей, наиболее важный – технология форвардинга. Микросхемы ASIC (Application Specific Integrated Circuit) позволили увеличить емкость коммутаторов с 10/100 Мбит/с до N Гбит/с и N x 10 Гбит/с, тогда как сегодня новые сервисы электронной коммерции, онлайн медиа, BYOD, продолжают увеличивать объемы трафика, требования к функциональности и гарантиям качества.

В этом документе мы расскажем о существующих архитектурных решениях и технологиях, увеличивающих производительность, в частности о сетевых процессорах NP (Network Processor) и их эволюции к ENP (Ethernet NP) – решению Huawei Enterprise для развития новой линейки оборудования – Alige-коммутаторов S12700.

Что такое NP? Программируемая микросхема, выполняющая различные задачи телекоммуникационной отрасли: обработка пакетов, анализ протоколов, поиск маршрутов, агрегирование видео-данных, файрвол, QoS. NP – это высокоскоростной программируемый процессор для обработки и передачи данных. Отличие от специализированных микросхем

ASIC состоит в широких возможностях программирования, сходство – в высокой скорости обработки данных. Программная настройка позволяет производителю оборудования изменить алгоритмы анализа трафика, а также реализовать новые сетевые функции и протоколы потоковой обработки данных.

NP сочетаются со всеми типами интерфейсов, протоколов L2-L7 и участвуют в различных технических решениях. Поддерживаются форматы инкапсуляции в ячейки, фреймы, пакеты, а также идентификация и обработка потоков данных. Для настройки NP используются проприетарные языки программирования, либо универсальный сокращенный набор команд (RISC – Reduced Instruction Set Computer), что обеспечивает возможность совместного использования различных NP. Для достижения высокой производительности и предотвращения потерь, структура NP включает несколько RISC ядер и автоматов пакетной обработки, оптимизированных для работы друг с другом. Таким образом в NP создается ускоритель для параллельного выполнения задач обработки данных.

Основываясь на более чем 20-летнем опыте R&D в микроэлектронике, Huawei разработал следующее поколение NP – Ethernet NP (ENP), объединяющие показатели производительности и энергопотребления ASIC и гибкость коммерческих NP.

Эволюция к ENP

До конца 1990 годов большинство коммутаторов строилось на архитектуре, аналогичной персональным компьютерам.

1995 – первая микросхема для коммутации данных на основе центрального процессора x86;

1997 – встроенный процессор заменил центральный процессор;

2001 – L2 и L3 ASIC микропроцессоры Broadcom;

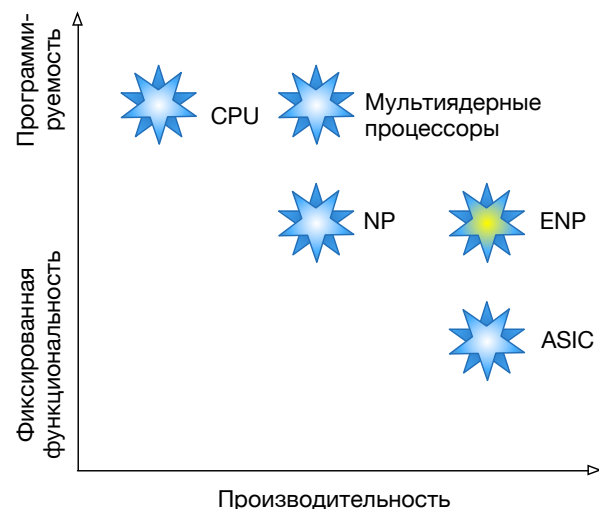
2004 – микропроцессор Broadcom BCM1250 с мультijядерной технологией;

2008 – Juniper представил программируемые ASIC, затем появляется аналогичная разработка Cisco;

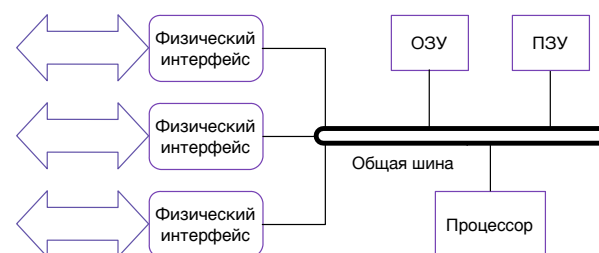
2009 – разработаны NP – высокая гибкость при высокой производительности;

2013 – Huawei выпустил ENP – еще большая производительность и емкость коммутации.

Дизайн коммутаторов с **Центральным процессором** был логичным продолжением в цепочке развития компьютерной техники. Центральный процессор обслуживал интерфейсы, выполнял коммутацию, маршрутизацию и пакетную фильтрацию. Аналогично компьютерам, CPU коммутатора получал инструкции от операционной системы и выполнял действия как предписывал программный код. Преимуществом такого подхода было то, что все инструкции хранились в программном обеспечении, новые функции легко добавлялись, поскольку требовался только программный апгрейд. Однако, универсальные микросхемы CPU разрабатывались для широкого применения и



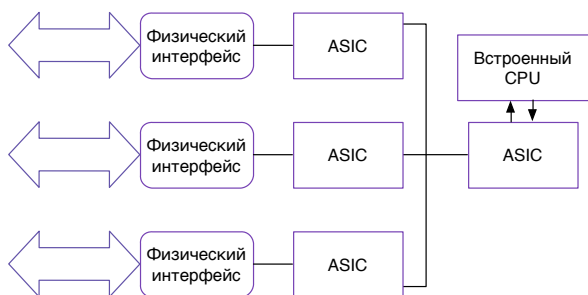
использовали расширенный набор команд (CISC, complex instruction set computer), избыточный для сетевых задач. Взаимодействие с памятью выполня-



лось стандартными средствами, что вносило задержки в производительности. А совместное выполнение функций управления и коммутации было решено за счет механизма прерываний для привлечения CPU к обработке входящего трафика. Результат – деградация производительности на 70% и рост утилизации при обработке большого числа коротких пакетов.

Встроенные CPU решали часть проблем, поскольку использовали отдельные звенья для интерфейсов ввода/вывода и операционной системы для передачи данных. Основанные на RISC архитектуре, встроенные CPU не обладали избыточностью центрального процессора, а выполняли лишь функции пакетной обработки трафика – прием, обработка и передача. За счет этого, встроенные CPU имели меньшее энергопотребление, меньшие размеры и меньшую стоимость. Однако, встроенные CPU не решали проблемы взаимодействия с памятью и производительностью в обработке данных.

Решения ASIC позволили аппаратно реализовать задачи обработки и передачи данных. Эти микросхемы характерны меньшими размерами, меньшим энергопотреблением, высокой надежностью, производительностью и низкой стоимостью.



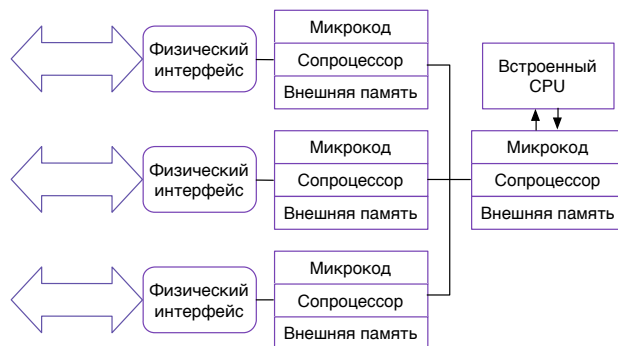
В микросхемах ASIC сервисы разбора пакетов, анализа, установки атрибутов и передачи, встроены в микросхему. Аппаратный форвардинг более эффективен, чем программный, а обработка коротких пакетов не влияет на производительность. Модули памяти интегрированы в ASIC, оптимизированы по функциональности и реализуют множественный доступ, ускоряя процесс чтения. Однако процессоры ASIC не обладают прежней гибкостью, поскольку внедрение новых аппаратных конфигураций требует замены оборудования. Даже программируемые микросхемы ASIC позволяют внести незначительные корректировки в функциональность, однако не могут кардинально менять методы форвардинга, например, реализовать IPv6 исключительно программными средствами. Важно отметить, что функциональность ASIC распространяется только на передачу данных, но не реализацию протоколов маршрутизации или сетевого взаимодействия – эти функции выносятся на уровень управления и реализуются на процессорный модуль коммутатора или встроенный CPU, а сервисная обработка трафика на уровнях L4-L7 выполняются на специализированных сервисных модулях.

Мультиядерные процессоры интегрируют вместе два или больше вычислительных механизма (ядра). Будучи расширением технологии CPU, решения обладают большой гибкостью и превосходят традиционную архитектуру по производительности. Мультиядерные процессоры разделяют уровень управления

от передачи данных – механизм прерываний более не требуется, поскольку функциональность делится между ядрами. Несколько пакетов могут обрабатываться параллельно, увеличивая производительность в разы, однако доступ к памяти выполняется последовательно. В результате, равные в стоимости многоядерные процессоры обладают меньшей производительностью, чем процессоры ASIC и коммерческие NP.

Коммерческие NP – это высокоскоростные программируемые процессоры, компромисс между ASIC и CPU. Таблица пересылки и процесс передачи данных реализованы программно на выделенных аппаратных ресурсах. Пользователь может настроить таблицы пересылки, и определить собственные правила поиска маршрута.

Коммерческие NP являются универсальными средствами для разработки новых приложений и устройств, поскольку обеспечивают высокую производительность, а их возможности программирования позволяют использовать их там, где раньше приходилось разрабатывать ASIC. Время внедрения уменьшается на несколько лет, а операционный процесс более не требует организации производства специализированных микросхем, поскольку вместо них используется универсальный NP. Простота внедрения новых L4-L7 сервисов на базе коммерческих NP открыла им области применения в коммутаторах и маршрутизаторах, файрволах, пограничных контроллерах доступа (SBC), устройствах предотвращения вторжений (IPD), решениях криптования и агентах мониторинга трафика.



Ресурсы NP параллельно включаются в обработку трафика, хотя каждый выполняет одинаковые задачи со своими потоками данных. В каждом модуле NP, сопроцессор и интегрированный процессор выполняют задачи разбора (parsing) и редактирования пакетов, ускоряя обработку. Микрокод обрабатывает данные протоколов маршрутизации, формирует маршрутные таблицы и определяет маршруты передачи трафика. Как и CPU, коммерческие NP сохраняют маршрутные записи и создают буфер обмена в памяти с произвольным доступом и удвоенной скоростью DDR, таким образом, размер памяти определяется программно. NP превосходит CPU в скорости чтения памяти, используя множественный доступ, аналогично ASIC. Программирование NP сводится к задаче разработки микрокода – менее дружественного чем язык C, но эффективного в своем классе.

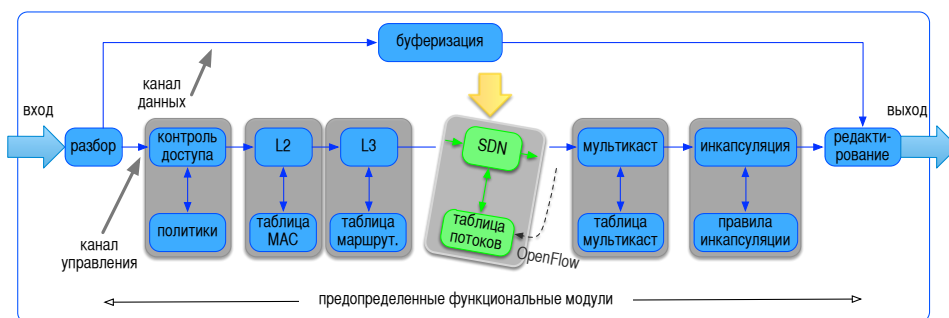
ENP – следующий шаг в развитии программируемых ASIC микросхем, призванный решать объемные тех-

нологические задачи корпоративных сетей – безопасность, мобильность, видео, QoS, облачные приложения и программно-определяемые сети SDN (Software Defined Networks). Huawei разработал решение ENP так, чтобы снизить стоимость и энергопотребление на уровень решений ASIC, но при этом поддерживать гибкое программирование, большие массивы данных и высокую скорость обработки дан-

ных. Кроме того, решение обладает 1.5 Gb буфером данных для снижения потерь трафика при всплесках. По сравнению с коммерческими NP, ENP обладает расширенной областью инструкций, возможностью равномерного использования всех вычислительных ресурсов, оптимизированную схему разрешения задачи маршрутизации, улучшенную производительность.

Предпосылки ENP

При высокой производительности, ASIC микросхемы поддерживают только predetermined алгоритмы обслуживания и фиксированный набор протоколов. Поэтому, наиболее существенная проблема в ASIC технологии – невозможность гибкого расширения функциональности. В примере ниже показана фиксированная схема аппаратного форвардинга в L3 коммутаторах. Добавление нового алгоритма, например MPLS или SDN требует корректировки цепочки, однако, изменение аппаратного дизайна микросхемы может занять годы.



Коммерческие NP разработаны, как альтернатива ASIC для реализации сложных схем программной обработки трафика, безопасности и качества обслуживания – например L4-L7 анализа, криптографии, управление очередями. Вместо аппаратной логики ASIC, NP предлагают программно конфигурируемые вычислительные ресурсы для быстрого внедрения новых приложений. Смогут ли NP стать долгосрочным решением задачи развития?

Коммерческие NP разделяют задачи между отделенными вычислительными ресурсами NPU (Network Processing Unit) – набором процессоров, собранном вместе для увеличения производительности. Каждой группе процессоров соответствует выделенная область памяти инструкций, описывающих алгоритмы выполнения задач. Сходные задачи размещаются на одной группе NPU, при необходимости NP выполнит операции используя ряд NPU последовательно или в несколько циклов.

Увеличение производительности NP до уровня ASIC обеспечивается за счет аппаратного наращивания тактовой частоты и конфигурации процессоров. Тогда как функциональность NP определяется программным кодом, что дает возможность внедрить новые сетевые протоколы и сценарии обработки данных путем программирования, без необходимо-

сти изменения аппаратного дизайна микросхем. К примеру, IPv6 и соответствующие таблицы маршрутов, новые типы туннелирования или шифрование трафика могут быть реализованы в NP за счет изменения микрокода, тогда как использование технологии ASIC потребует замену оборудования.

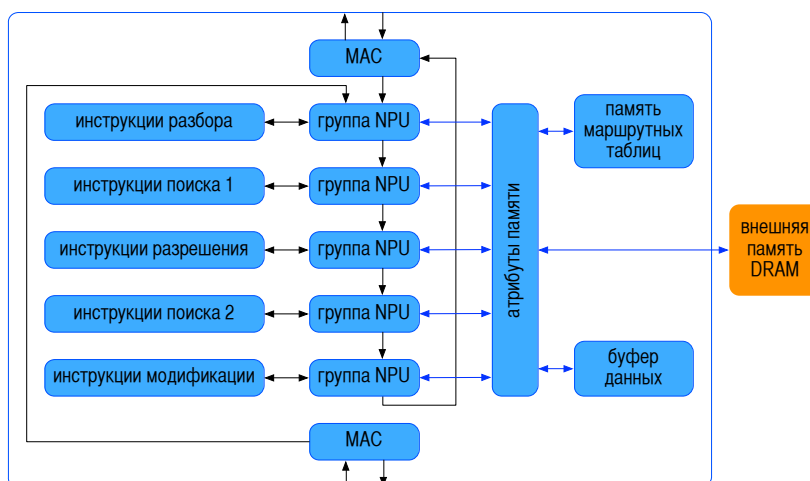
Однако, неправильное планирование ресурсов может привести к простоям одних и перегрузке других NPU, приводит к появлению «узких мест» и деградации производительности. Еще одним недостатком

NP является ограниченность размера инструкций и производительности группы NPU, особенно при необходимости внедрения более ресурсоемких задач, чем планировалось при разработке NP.

Ответом Huawei Enterprise на проблему производительности NP стал разработанный в 2013 году процес-

сор **ENP** – развитие NP с учетом:

- функциональных требований, в частности, технологии SDN (Software Defined Networking), которая предполагает коммутацию большого количества



потоков данных и использование отдельных маршрутных таблиц SDN;

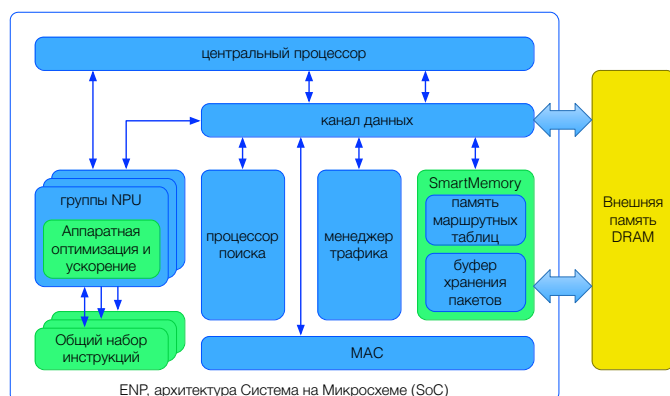
- снижения энергосбережения, в противовес существенному росту мощности устройств, основанных на коммерческих NP;
- снятия ограничений на объем инструкций и улучшения быстродействия памяти.

Высокая производительность и гибкость в ENP

Уникальными решениями в дизайне Huawei ENP стали: единая область инструкций, оптимизированная память Smart Memory, таблица пересылки на миллионы записей и энергосбережение.

Единая область инструкций

В отличие от стандартной реализации на коммерческих NP, область инструкций ENP не закреплена за каждой из групп NPU, а единая на все группы. Каждая из групп NPU может исполнять любой набор инструкций – от анализа заголовков до модификации данных. Таким образом, функциональность не должна распределяться между различными группами NPU, а все ресурсы доступны для выполнения любой из задач. Решение Huawei исключает проблему распределения функций для планирования производительности, снижает время на разработку новой функциональности.



Кроме того, ENP имеет несколько параллельных цепочек обработки данных – так увеличивается производительность и оптимизируется работа интерфейсов ввода/вывода. В частности:

- ENP может выполнять параллельную обработку протоколов разных уровней L2, L3, MPLS, VPN;
- Ввиду того, что размер инструкций для NPU группы в коммерческих NP ограничен, сложные сервисы разносились на несколько NPU групп, которые выполняли свои задачи последовательно; единая область инструкций в ENP позволяет выполнять сервис на одной группе NPU, поскольку размер инструкций может быть существенно увеличен.

Smart Memory

Как микропроцессоры ASIC, так и большинство коммерческих NP разделяют функции обработки и хранения данных. Взаимодействие с элементами памяти создает задержку, а исполнение в отдельных модулях увеличивает энергопотребление. Поэтому в своем решении Huawei интегрировал элементы памяти Smart Memory в ENP и тем самым оптимизировал скорость обмена данными.

Smart Memory объединяет функциональность хранения маршрутных таблиц и поиска маршрутов, а также способен обрабатывать трафик, выполняя в частности ограничения скорости передачи данных или участвуя в сборе статистики.

Миллионы записей в таблице пересылки

Гибридное SDN решение Huawei Enterprise параллельно использует уровень OpenFlow и уровень традиционной маршрутизации для управления пересылкой данных. Такой подход позволяет осуществлять плавную миграцию к SDN архитектуре, не нарушая работу существующих сервисов. Аналогичное решение используется и другими участниками рынка, однако встроенные таблицы потоков в ASIC решениях обладают небольшим размером. Одним из промышленных решений является использование внешних модулей TCAM (Ternary Content Addressable Memory), которые комбинируют таблицу пересылки, QoS и пакетные фильтры. Однако такие решения также ограничены по быстродействию и размерам таблиц.

В решении ENP поддерживается до 16 миллионов потоков данных, что обеспечивает возможности гибкого управления сервисами для десятки тысяч сотрудников кампуса.

Энергосбережение

Энергопотребление микросхем ASIC фиксированное и обычно ниже, чем у коммерческих NP. Решение ENP уменьшает энергопотребление в зависимости от объема трафика и сложности обработки, поскольку ENP способен блокировать неиспользуемые группы NPU. Когда трафик увеличивается, ENP активирует ресурсы в течение нескольких наносекунд. Успешность этого механизма проверена тестовыми лабораториями, которые зафиксировали способность коммутаторов Huawei справляться с резким всплеском объема трафика.



Использование технологии энергосбережения приводит к снижению потребляемой мощности на 30%.

Энергопотребление микросхемы прямо пропорционально зависит от тактовой частоты. Поэтому в решении ENP, Huawei интегрировал спидометр, контроллеры напряжения и синхронизации. Спидометр находится на пути следования данных к группам NPU для мониторинга объема трафика. Изменения скорости данных приводят к изменениям тактовой частоты канала данных. В то же время, изменения числа пакетов приводят к регулировке количества активных пакетных процессоров. Если группа NPU заблокирована, то на процессоры не подается электропитание и

синхронизация. При необходимости, процессоры моментально включаются в работу.

Приложения ENP

Процессоры ENP стали основой инновационных Agile-коммутаторов S12700, представленных Huawei Enterprise в марте 2013 года. Процессоры ENP используют в модулях типа X1E с интерфейсами 1GE и 10GE, которые рекомендованы для подключения агрегирующих коммутаторов уровня доступа. Тогда как магистральные интерфейсы 10GE, 40GE и 100GE формируются модулями, использующими ASIC архитектуру для достижения максимальной производительности и поддержки MPLS. Процессоры ENP также планируют использовать в коммутаторах серии S9700/S7700 и младших моделях для обеспечения миграции сетей к технологии SDN.

Благодаря гибкому решению ENP, в реализации Huawei на одном интерфейсе могут параллельно работать :

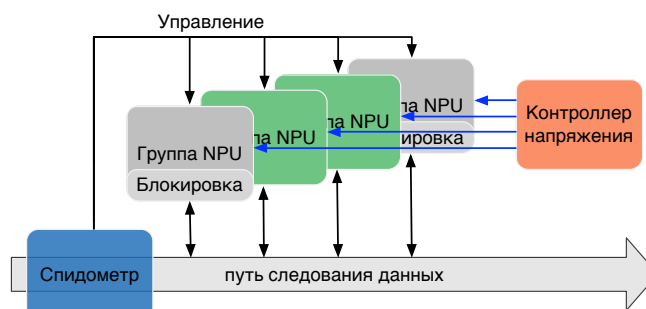
- функции коммутации и маршрутизации (3 млн. маршрутных записей IPv4, 1 млн. IPv6, 128 тыс. IP мультикаст, 1 млн. MAC и 256 тыс. пакетных фильтров),
- SDN для управления потоками данных (16 млн. потоков).

Реализация SDN в решении Huawei POF (Protocol Oblivious Forwarding) базируется на стандарте OpenFlow и обратно совместима с ним для возможности взаимодействия с другими производителями. Технология POF отличается возможностями работать с любыми протоколами и нестандартными пакетами, тогда как стандарт OpenFlow предполагает только лишь операции с IP. Процессоры ENP расширяют возможности Huawei по работе с новыми протоколами, которые могут быть внедрены программным путем, тогда как реализации SDN на базе ASIC (и опциональных TCAM) лишает производителей гибкости.

Практическим примером использования SDN решения Huawei Enterprise является инкапсуляция данных и добавление в заголовки пакетов пользовательской информации. В тестовой схеме собранной Miercom, было проверено формирование нестандартного пакета, содержащего информацию о месторасположении отправителя. Дополнительная информация в дальнейшем использовалась для идентификации потока данных и направления его по сети.

User-defined flow table rule		
building:	11 22	(Length: 2)
floor:	33 44	(Length: 2)
room:	55 66	(Length: 2)
seat:	02	(Length: 1)
Inbound interface:	GigabitEthernet 2/0/24	
Action: ☒ Encapsulation ☐ Forward ☐ Decapsulation		
outport interface:	GigabitEthernet 2/0/34	
Apply		

Кроме SDN, интерфейсные модули на основе ENP поддерживают расширенную функциональность агрегирования трафика беспроводного доступа, реализуют контроль качества и локализацию аварий. Решение по абонентскому агрегированию включает распределенный контроллер доступа AC (Access Controller) и сервер удаленного доступа BRAS (Broadband Remote Access



Server), что превращает коммутатор Huawei S12700 в высокопроизводительный концентратор трафика WiFi, способного объединить до 4 тысяч точек доступа AP (Access Point), обслуживающих до 65 тысяч одновременных подключений. Суммарная пропускная способность решения – 800 Гбит/с. Для удобства управления AP и удаленными портами агрегирующих коммутаторов в Agile-решении Huawei реализована технология объединяющей виртуальной фабрики SVF (Super Virtual Fabric). SVF представляет удаленные AP и порты коммутаторов доступа, как виртуализированные порты S12700, а технология управления и настройки беспроводных точек доступа CAPWAP (Control And Provisioning of Wireless Access Points), позволяет централизованно управлять программным обеспечением, конфигурациями и текущими настройками сети AP без необходимости работы с индивидуальными устройствами.

Интеграция решений абонентского агрегирования и SDN позволяет реализовать политики обслуживания пользователей и при необходимости оперировать с трафиком вертикально интегрированных приложений, выделяя его в виртуальные среды безопасного доступа. При этом, концепция потоков данных, реализованная в коммутаторах Huawei S12700, позволяет централизованно управлять качеством в масштабах всей сети. Технология изолированных сетевых потоков iPCA (Packet Conservation Algorithm for Internet) позволяет единожды определить признаки приоритетного трафика на контроллере кампуса, а затем передать требования на все сетевые устройства. Решения на базе ENP идентифицируют потоки приоритетного трафика и управляют его прохождением через сеть на каждом транзитном узле. А для контроля качества и локализации аварий, iPCA накапливает операционную статистику на базе каждого микропроцессора ENP.

Реализация iPCA – это еще один пример технологического лидерства Huawei и практического применения SDN.

