

Масштабируемый вычислительный тракт высокопроизводительных сетевых коммутаторов для перспективных подходов обработки трафика

Авторы: Елизаров С.Г., Гончаров С.В.,
Лукьянченко Г.А., Монахов А.М., Смелянский Р.Л.

- 2000-е и ранее: автоматическая конфигурация сетевого оборудования, в лучшем случае, на основе шаблонов
- все производители имеют свой неповторимый синтаксис конфигурации (хотя и тяготеют к Cisco-style)
- изменения конфигурации централизованным образом практически невозможны или ограничиваются очень малым набором операций
- решение - единый протокол управления для всех сетевых устройств
- OpenFlow – прекрасное решение (для институтских лабораторий)

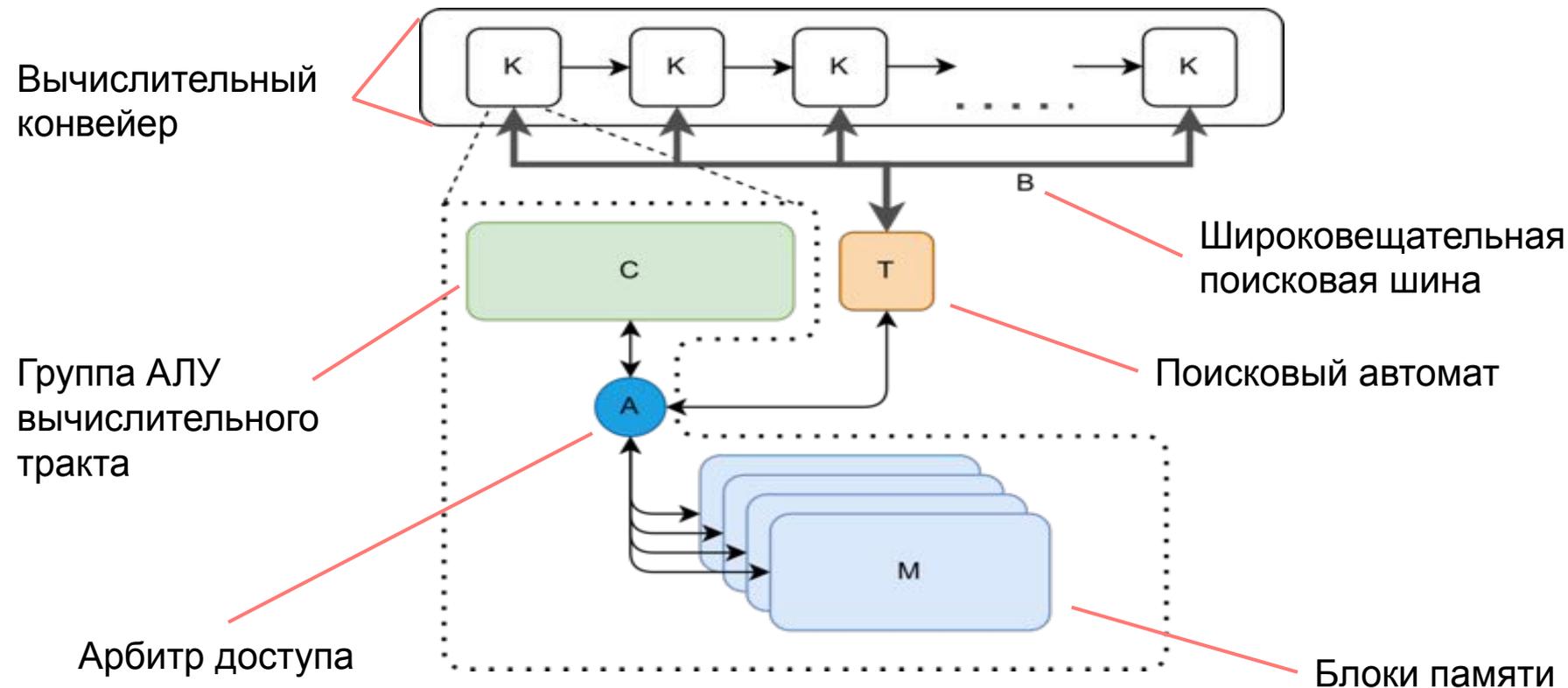
- универсальный протокол управления требует универсальной расширяемости
- производители оборудования за последнее десятилетие сделали множество несовместимых друг с другом оркестровок для централизованного управления сетью, хотя появились универсальные разметки (YANG/REST-модели...)
- программная сеть не имеет более преимуществ...?
- P4 - король умер, да здравствует король!
- что такое эффективность?



- динамическое распределение ресурсов памяти и фиксированного вычислительного времени между задачами классификации, поиска и преобразования элементов сетевых пакетов
- сегодня нужно разобрать десять вложенных заголовков MPLS, завтра хочется использовать изобретенную для собственных задач транспортную обертку
- для большого числа меток/классификаторов и их комбинаций требуется большое количество счетчиков

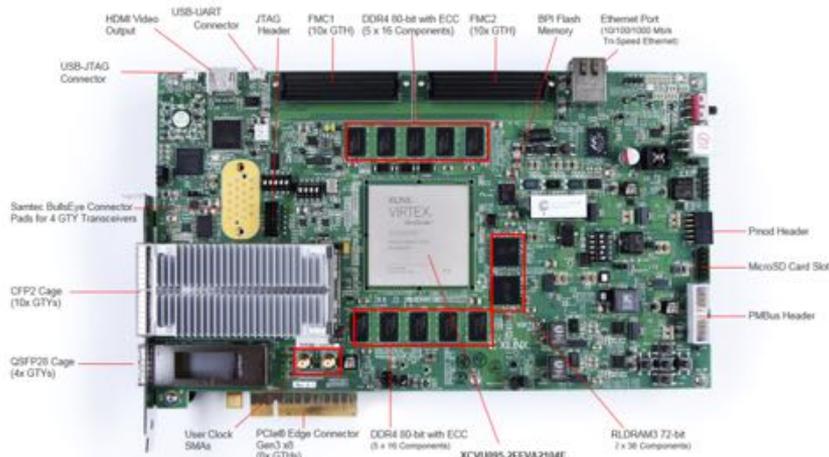


Архитектура, которая (быстро) не устареет



- поддержка классических протоколов – через обратную связь с контроллером с стейтмашинами ядра Linux и юзерспейсным ПО
- ASIC “разучивается” делать специализированные вещи, например, аппаратно обновлять таблицу форвардинга
-зато “обучается” любой разметке трафика и любому L2/L3 протоколу без потерь в характеристиках, а программируемость дает возможность использовать микрокод, обновляющий себя или свои поисковые таблицы (опять научились обновлять fdb аппаратно)





ПЛИС для прототипирования: Virtex Ultrascale с платами расширения FMC <-> 4x1G

Архитектурные упрощения:

- снижен объем общей буферной памяти
- уменьшено количество стадий конвейера
- использованы универсальные подходы к формированию микрокода на базе open-source решений, не обеспечивающие предельную плотность кода или времени вычислений, но позволившие за короткое время прототипировать связку control plane с data plane.

Протокол управления	OpenFlow 1.3, (P4 1.1.0 - WIP)
Скорость передачи	1 Гбит/с, line rate
Количество портов	4 (8 - WIP) 1000Base-T

- деагрегация не на словах, а на деле – распределенные поисковые таблицы и правила обработки пакетов и их когерентная модификация в рамках сотен наносекунд (+ скорость света)
- content-based network – сеть с миллиардом правил
- вычисления-в-проводе, требующие низких задержек (*), (**)
- вспомогательные функции для NFV

* P4xos: Consensus as a Network Service, doi: 10.1109/TNET.2020.2992106

** In-Network Computation is a Dumb Idea Whose Time Has Come, doi: 10.1145/3152434.3152461



Сети 5G



Датацентровые
сети



Инфраструктура
провайдеров
СВЯЗИ



Спасибо за внимание!



МЕЖДУНАРОДНЫЙ ФОРУМ
МИКРОЭЛЕКТРОНИКА-2020