

СКАТ DPI переходит на порты 25,40 и 100 Гбит/с

Мы знаем, что среди наших читателей множество клиентов, кто использует решение СКАТ DPI. Большинство крупных операторов давно задаётся вопросом о поддержке интерфейсов выше 10Гбит/с. Постоянное объединение портов в группы, организация кабельного хозяйства, выделение портов на коммутаторах - это те проблемы от которых также хотелось бы избавиться.

"Наш Road Map является очень гибким и в первую очередь реагирует на потребности клиентов, поэтому поддержка новых карт откладывалась в пользу развития функциональности. Наконец, после значительных работ по совершенствованию BRAS и поддержки IPv6, мы занялись развитием вертикального масштабирования включая работу со скоростными интерфейсами", - комментирует технический директор VAS Experts Дмитрий Молдаванов.

Ниже, мы рассмотрим основные факторы, которые помогли разработчикам добиться конечного результата, а также технологии, применяемые в работе.

Новости процессорного рынка

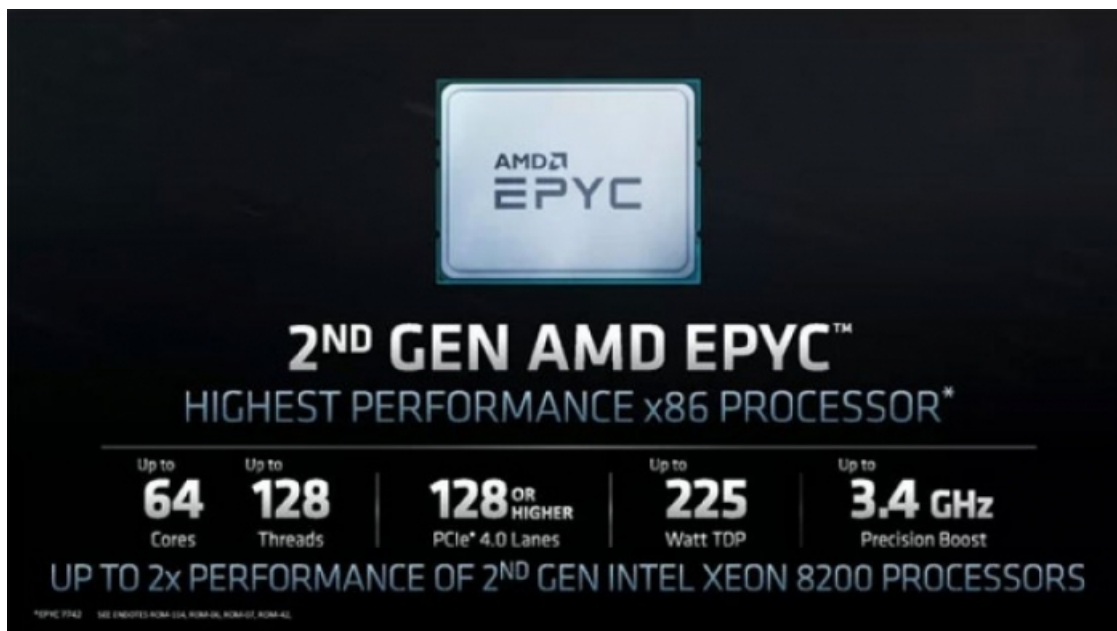
В основе работы платформы СКАТ DPI лежат сложные вычисления, что требует бесспорно больших вычислительных мощностей. На сегодняшний день используются процессоры Intel архитектуры x86. Также, готовы к работе процессоры AMD второго поколения, для массовой обработки данных.

Изначально VAS Experts сделали ставку на то, что производительность CPU будет расти и становится доступнее, поэтому при разработке в архитектуру был заложен принцип вертикального масштабирования - увеличения производительности на одно устройство путем наращивания CPU cores, RAM, NIC.

В первом квартале 2020 года произошло то, чего мы так все ждали - компания Intel анонсировала значительное расширение линейки процессоров Xeon Scalable второго поколения. В первую очередь, речь идет о Xeon Gold, Silver и Bronze: их производительность заметно увеличилась, а цена - сильно снизилась. Общие данные можно увидеть в таблице.

2ND GEN INTEL XEON SCALABLE PROCESSOR PORTFOLIO ENHANCEMENT							
Processor Number	Cores	Base Clock Speed (GHz)	Turbo Boost (GHz)	Cache (MB)	TDP (W)	Intel Optane Persistent Memory Support	Pricing (RCP)
Xeon Gold 6258R	28	2.7	4.0	38.5	205	Yes	\$3,950
Xeon Gold 6256	12	3.6	4.5	33	205	Yes	\$3,900
Xeon Gold 6250	8	3.9	4.5	35.75	185	Yes	\$3,400
Xeon Gold 6248R	24	3.0	4.0	35.75	205	Yes	\$2,700
Xeon Gold 6246R	16	3.4	4.1	35.75	205	Yes	\$3,286
Xeon Gold 6242R	20	3.1	4.1	35.75	205	Yes	\$2,529
Xeon Gold 6240R	24	2.4	4.0	35.75	165	Yes	\$2,445
Xeon Gold 6238R	28	2.2	4.0	38.5	165	Yes	\$2,612
Xeon Gold 6230R	26	2.1	4.0	35.75	150	Yes	\$1,894
Xeon Gold 6226R	16	2.9	3.9	22	150	Yes	\$1,300
Xeon Gold 6208U	16	2.9	3.9	22	150	Yes	\$989
Xeon Gold 5220R	24	2.2	4.0	35.75	150	Yes	\$1,555
Xeon Gold 5218R	20	2.1	4.0	27.5	125	Yes	\$1,273
Xeon Silver 4215R	8	3.2	4.0	11	130	Yes	\$749
Xeon Silver 4214R	12	2.4	3.5	16.5	100	Yes	\$694
Xeon Silver 4210R	10	2.4	3.2	13.75	100	Yes	\$501
Xeon Silver 4210T	10	2.3	3.2	13.75	95	No	\$554
Xeon Bronze 3206R	8	1.9	1.9	11	85	No	\$306

В свою очередь, извечный противник Intel, компания AMD выпустила свою линейку процессоров второго поколения по привлекательным ценам.



Благодаря подобным противостояниям, компании-разработчики софта могут использовать в своих решениях возможности вертикального масштабирования при росте производительности оборудования.

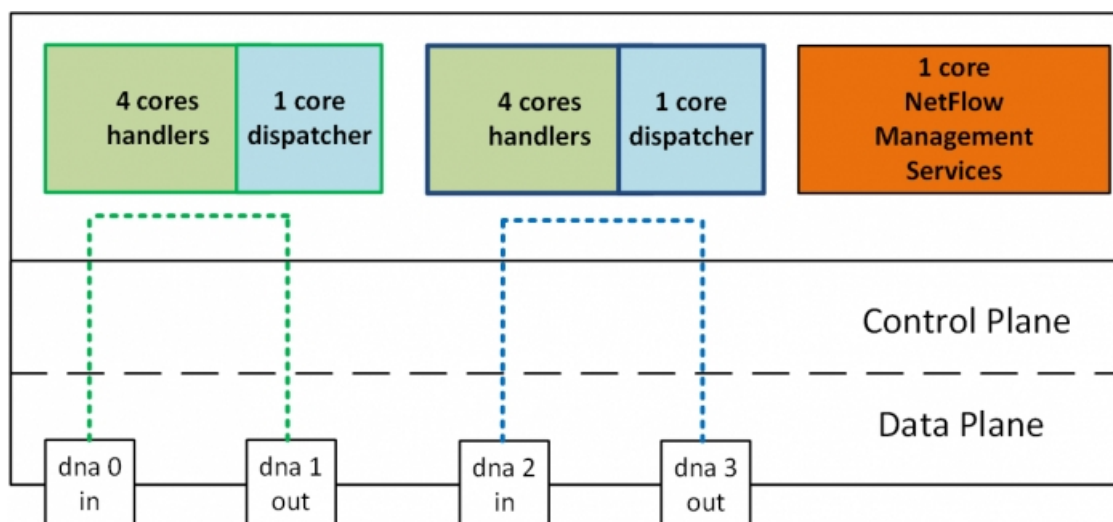
В конечном итоге ценовая война позволила снизить стоимость топовых платформ на 40%, и сегодня мы можем использовать не только Intel, но и AMD - более оптимальный по соотношению цена/количество ядер.

Принципы взаимодействия DPI с процессорными ресурсами

Напомним следующие моменты:

- система DPI чувствительна к частоте ядер (минимальное пороговое значение 2,5 ГГц),
- система может задействовать только физические ядра (Hyper Threading нужно отключать),
- для оптимальной работы системы используется мульти-кластер - дробление системы на кластера с целью распределения нагрузки по ядрам и роста производительности на процессор.

Наглядно это выглядит так:



Пример из практики по загрузке процессора Intel Xeon Gold 6126 (2,60ГГц, 12 ядер) из первого поколения Scalable.

Активированные опции:

- Полисинг общего канала и ограничение скорости на абонента по классам.
- Снятие статистики на СОРМ-3.



Для общей информации: производитель рекомендует использовать данный процессор на производительности до 20 Гбит/с в одну сторону.

В нашем случае, это решения SKAT-40. Несмотря на эти ограничения, оператор подал нагрузку на уровне 30 Гбит/с, чтобы проверить возможности системы. При этом, система мониторинга показала загрузку CPU всего лишь на уровне 60%.

Давайте перейдем к конкретным конфигурациям платформ и посмотрим какой конфиг рекомендуют разработчики.

Пиковый исходящий, Гбит/с	Пиковый входящий, Гбит/с	Версия SKAT DPI	CPU cores	RAM, ГБ
до 3	до 3	6	4	16
до 5	до 5	10	6	32
до 10	до 10	20	6	48
до 20	до 20	40	12	64
до 30	до 30	60	18	96
до 40	до 40	80	22	128
до 50	до 50	100	28	160
до 60	до 60	120	32	192
до 100	до 100	200	64	256

Из линеек берем более дорогие процессоры, исходя из требований по базовой частоте (смотреть на Turbo Boost не стоит, т.к. не все ядра можно разогнать до данных значений). Выбираем Intel Xeon Gold 6258R вместо Intel Xeon Gold 6238R, хотя цена очень соблазнительная.

CPU	Пропускная способность, Гбит/с	Конфигурация платформы	Типы интерфейсов	Цена платформы (зависит от типа интерфейсов и наличия bypass)
Intel	100	Supermicro SYS-5019P-WTR (Gold 6258R, 192GB RAM, 2x480GB SSD)	2.5G/5G/10G/25G/50G/40G/100G на чипсетах Intel XXV710, XL710, E810, FM10000 Полный перечень здесь.	от 8000\$
AMD	100	Supermicro AS-1014S-WTRT (AMD EPYC 7542 , 192GB RAM, 2x480GB SSD)		от 8000\$
AMD	200	Supermicro AS-1014S-WTRT (AMD EPYC 7742 , 256GB RAM, 2x480GB SSD)		от 15000\$

Реализация SKAT-100 на Intel второго поколения и AMD, сравнима по цене, но реализация SKAT-200 возможна только на AMD, притом, что цена меньше, чем простое удвоение.

Вы спросите, а почему не рассматривается линейка Intel Platinum?

1. Intel Platinum дороже чем AMD.
2. [Процессор Intel® Xeon® Platinum 9282](#) по факту состоит из двух CPU по 28 ядер с использованием NUMA. Данная конфигурация не поддерживается для работы DPI.

Поддержка новых чипов сетевых карт сопровождается переходом к технологии DPDK и ОС Centos 8 в роли Control Plane.

Уже доступна бета версия 9.0 SKAT DPI, релиз запланирован на 2 квартал 2020 года.

Следите за новостям VAS Experts, в рассылках НАГ. Как только компания VAS Experts будет готова, мы позовем вас на более детальный вебинар, посвященный данному вопросу.

Data Plane и сетевые карты

Ранее разработчики использовали в качестве Data Plane библиотеки прямого доступа к сетевому адаптеру от компании [NTOP](#), что было обусловлено стабильной работой на момент запуска продукта (2013 год) и независимостью от платформы. Но время не стоит на месте и Open Source продукты доказывают свое качество. Поэтому был сделан выбор в пользу DPDK.

В чем преимущества данной технологии:

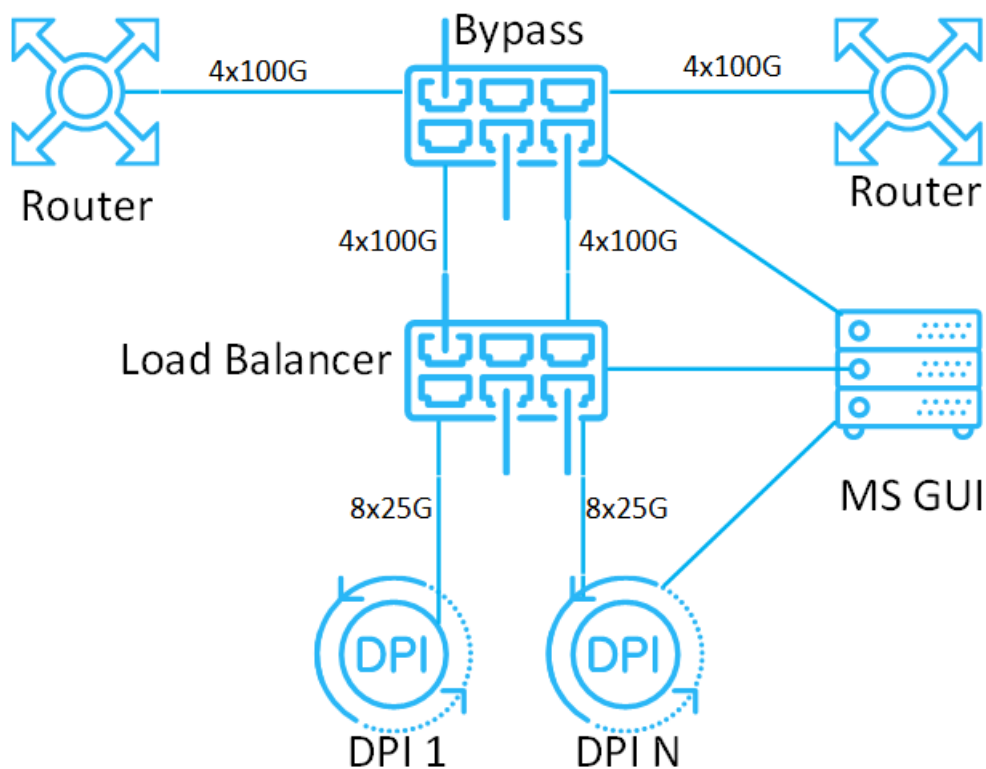
1. Нет лицензирования, соответственно, снижается стоимость платформы.
2. Поддержка множества сетевых [чипсетов](#).
3. Открытость кода - можно править ошибки, что VAS Experts и практикует :)
4. Открытое Community - легче разобраться с проблемами.
5. Гибкая конфигурация - на практике дает большую производительность.
6. Поддержка виртуальных интерфейсов - работа на гипервизоре в облачной инфраструктуре.

DPDK сделал SKAT DPI еще более доступным и снизил требования к оборудованию. Напомним, что большая часть клиентов использует собственное оборудование для установки ПО. Масштабирование железа проходит также силами оператора с участием технической поддержки разработчика.

Решения для систем больше 1 Тбит/с

Как было отмечено выше, разработаны устройства для скоростей до 200Гбит/с. Что же делать, если этого недостаточно? Ждать более производительных решений от AMD или Intel, использовать современные методы ускорения серверных вычислений, такие как FPGA или GPU? Это все требует длительной разработки программного кода, оптимизации решения, а также непонятно, когда наши процессорные мощности достигнут таких возможностей, чтобы обрабатывать подобные вычисления.

В нашем случае все гораздо проще: задача решается путем установки специального устройства - балансировщика, который при необходимости подключается через внешний bypass-модуль. Балансировщик выделяет пользовательские сессии и направляет симметричный трафик на DPI-модуль.



Конфигурация Load Balancer может включать 10G/25G/100G-порты, что в пределе дает возможность обработать с помощью двух 64-портовых свичей, ориентировочной стоимостью 10 тыс. у.е. каждый, 32x100G каналов. Общая пропускная способность комплекса составит 6,4 Тбит/с общего трафика или 3,2 Тбит/с full duplex. Этого достаточно для большинства клиентов.

Итог

Таким образом, уже сегодня можно готовиться к внедрению технологии DPI на высоких скоростях и с доступным бюджетом.