

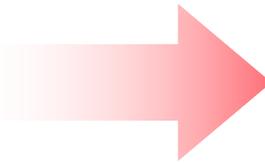
Экосистема российского ПО для серверов ARM



Михаил Никитин,  HUAWEI

Распределенные вычисления. Текущие вызовы

Новые вызовы ИТ индустрии Интеллектуальные вычисления



приложения



- Мобильные приложения
- Интернет вещей
- Беспилотный транспорт

данные



- текст
- изображение
- голос
- видео

вычисления



- Обработка текста
- Анализ больших данных
- Научные вычисления
- Обработка видео

Сочетание нескольких вычислительных архитектур - оптимальный путь решения

CPU

сопроцессор

x86

ARM



GPU

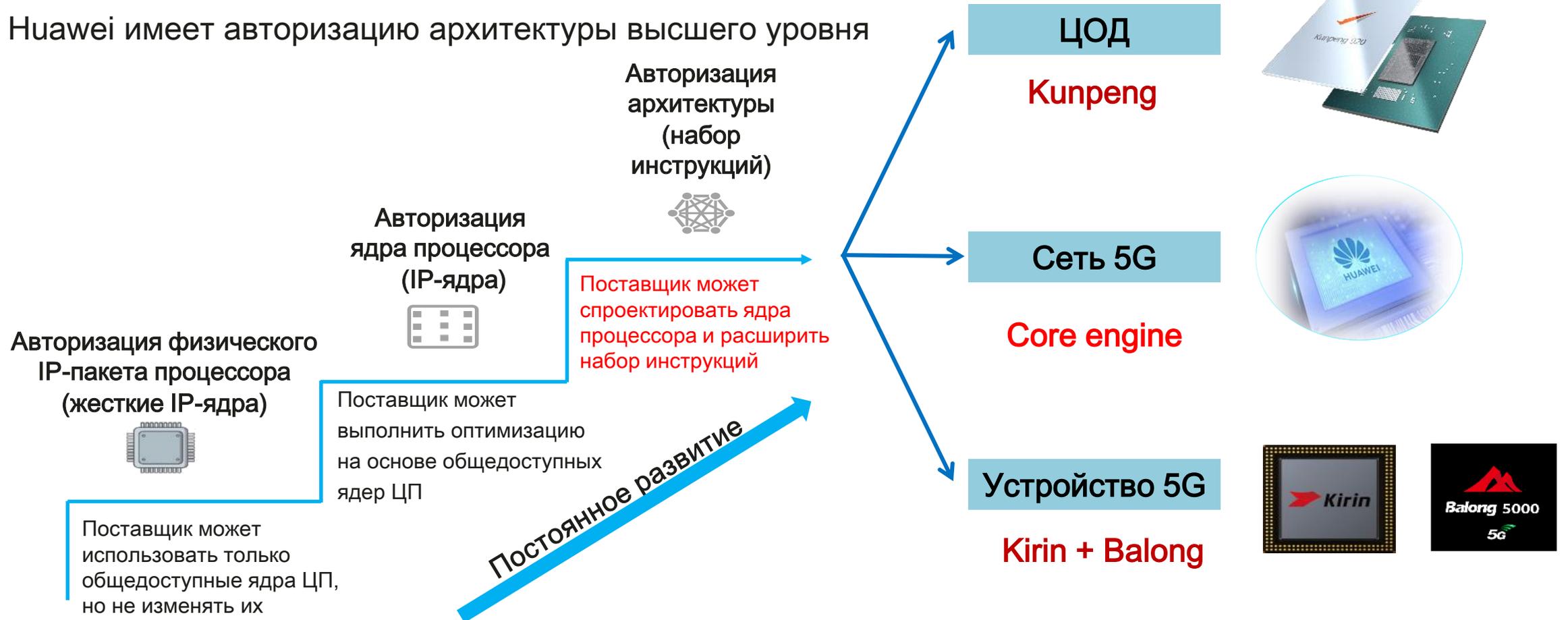
FPGA

NPU

- | | | | | |
|---|---|---|--|---|
| <ul style="list-style-type: none">• Тяжелые ядра• Производительность• Универсальность• Развитая экосистема | <ul style="list-style-type: none">• Многоядерность• Параллелизм• Энергоэффективность• Открытая платформа | <ul style="list-style-type: none">• Многоядерная архитектура• Вычисления с плавающей запятой• Обработка изображений | <ul style="list-style-type: none">• Гибкая логика• Гибкий и программируемый | <ul style="list-style-type: none">• ASIC чип• Низкое энергопотребление и низкие задержки• Глубокое обучение |
|---|---|---|--|---|

У Huawei есть постоянная лицензия на Armv8

Huawei имеет авторизацию архитектуры высшего уровня



Huawei полностью освоила ключевые технологии ядер процессоров Arm64 и SoC-дизайн, а также имеет постоянную авторизацию на архитектуру Armv8 для поддержки общей стратегии Huawei.

Серверы TaiShan: вычислительная техника на основе процессоров Kunpeng

17 лет инженерной экспертизы



Уникальный термодизайн

Тройная вибрационная защита и пассивный бэкаплайн

Высокая теплоотдача

Низкие задержки

Надежные компоненты

Контроль Качества

Скоростная шина обмена данных: 56G на уровне системной платы

На 15% меньше отказов, чем в среднем по отрасли

Ориентированные на хранение

5290
Сервер СХД
4U 72 диска



5280
Сервер СХД
4U 40 дисков



2280
Универсальный
Сервер
2U 2S



ЦОД

Ориентированные на вычисления

1280
Высокоплотный
Сервер
1U 2S



2480
Высокопроизводительный
Сервер
2U 4S



X6000
Супер плотный сервер
2U 4-node



Граничные
вычисления

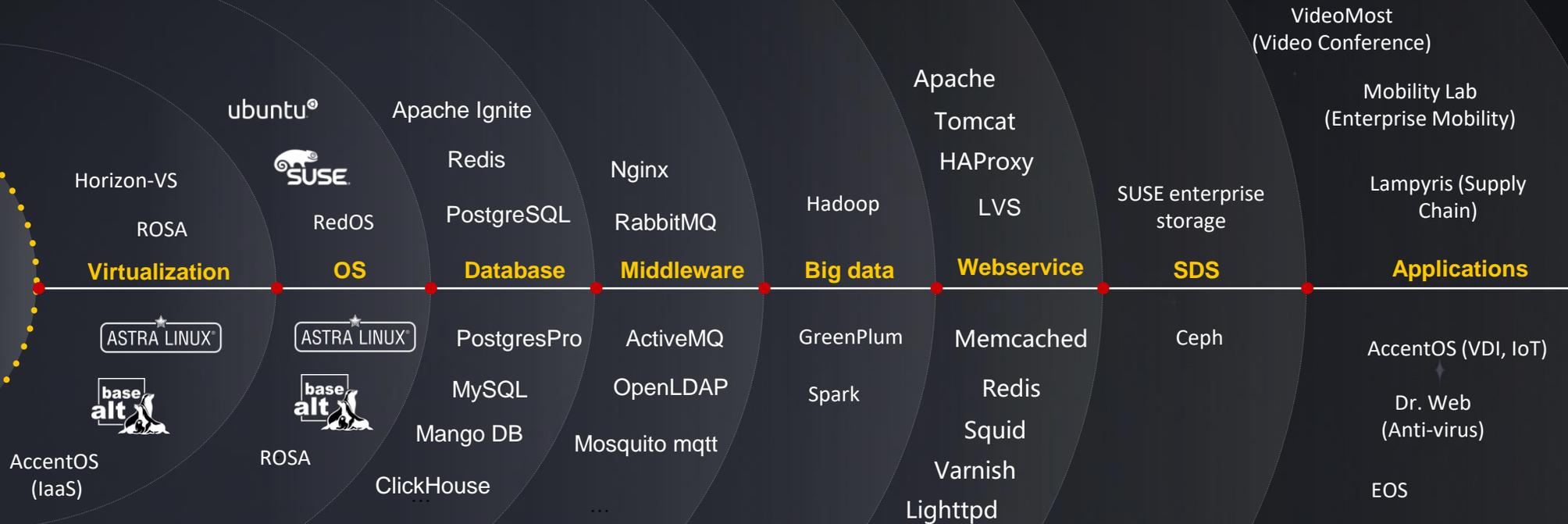
2280E
Edge сервер



2480 и 5290 доступны после первого квартала 2020

Экосистема ПО под ARM

Kunpeng



Купреng обладает гибкостью в подборе к разным сценариям



Big data

- В рамках проекта China Telecom Jiangsu, посвященного большим данным, первый проект гибридных развертываний больших данных на Kупреng 916 (E5-2620) операторского класса мигрирует и оптимизируется. После гибридного развертывания производительность улучшается на 50% по сравнению с исходным кластером X86.



Distributed storage

- Исследовательский центр Сучжоу: кластер распределенного хранилища, Kупреng 920 (5230, 4214), улучшена пропускная способность ввода / вывода чтения и записи на 36,4% и IOPS на 6,5%



Database OLAP

- После оптимизации, производительность виртуализации и ЦП больших данных в проекте миграции системы запросов кредитных карт China Construction Bank превышает производительность серверов x86 (C1.6148) более чем на 20%, а производительность базы данных превышает производительность серверов x86 более чем на 10%.



Web applications

- В проекте оптимизации производительности веб-системы JD Mobile APP решение по ускорению SSL Kупреng построено на основе встроенного механизма ускорения RSA Kупреng 920. Производительность HTTPS одного узла в 2,4 раза выше, чем у решения аппаратного ускорения x86 (4110).



Cloud platform

- В проекте развертывания контейнерного решения China Merchants Bank производительность контейнера Kупреng 920 составляет от 120% до 130% от производительности x86 (6138) благодаря оптимизации компонентов Nginx и Kafka.

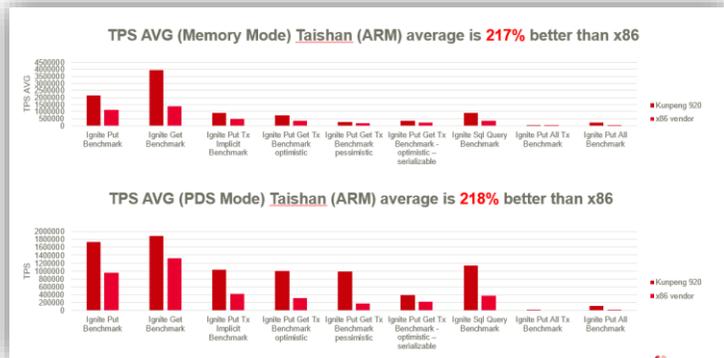
Достижения в производительности Taishan и планы на 2021



2021

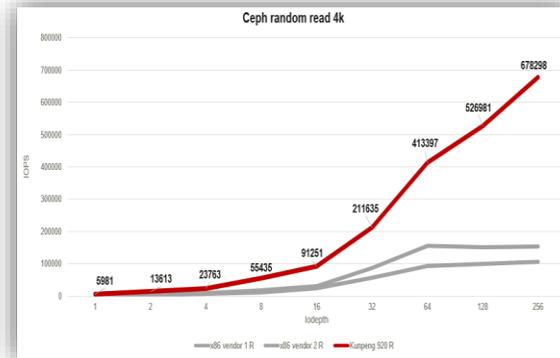
Database

Apache ignite: **+ 217%**



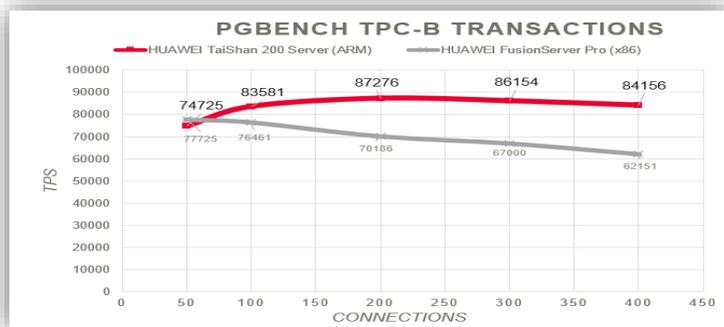
SDS

Ceph: **+231%**



Machine Learning

Apache Spark: **>100%**



Algorithm	Kunpeng VS X86	New alg.+Kunpeng VS Open alg.+X86
Rand Forest	-2.9%	112.9%
GBDT (Gradient Boosting Decision Tree)	-7.01%	281.04%
MCE (Maximal Clique Enumeration)	-44%	598.07%
TC (Triangular count)	-4.3%	150.9%
Modularity	-8.3%	738.4%



Thank you.

Все вопросы можно задать по электронной почте:

Mikhail.nikitin@huawei.com