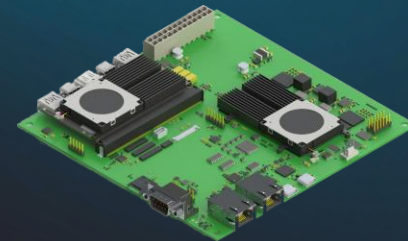


**Отечественные процессоры и решения на их базе  
для аппаратного ускорения работы нейронных сетей**

## Отечественные вычислители для запуска нейронных сетей

Бренд LinQ объединяет дизайн-центр электроники и российскую разработку решений для искусственного интеллекта на базе отечественного процессора собственной архитектуры и полного набора необходимого программного обеспечения.

Вычислители LinQ предназначены для запуска ИИ-приложений и применяются в области периферийных вычислений, туманных вычислений, а также в стандартной инфраструктуре центров обработки данных.



# 10 лет истории, 6 лет разработки...

- 2013 год**  
основание компании «ХайТэк»
- 2016 год**  
идея разработки тензорного процессора
- 2017-2018 годы**  
формирование команды и начало разработки
- 2020 год**  
тестирование архитектуры на независимой площадке MLperf
- 2021 год**  
получение статуса дизайн-центра
- 2022 год**  
первые чипы и модули
- 2023 год**  
три новых типа модулей
- 2024 год**  
обновленная версия чипа и модулей на его базе



## Собственные ресурсы:

- ✓ команда разработки ядра
- ✓ команда разработки ПО и компиляции
- ✓ команда разработки модулей

# Технические характеристики процессора H

**28 нм**

Техпроцесс

**16МБ**

Встроенная память

**BGA1296**

Корпус

**500-812 МГц**

Частота блока TPU

**≤ 25-30 Вт**

Энергопотребление

**MIPS64**

Управляющий процессор

**Gen3, EndPoint,  
8 линий**

Контроллер PCIe

**2 канала 72bit  
до 32Gb**

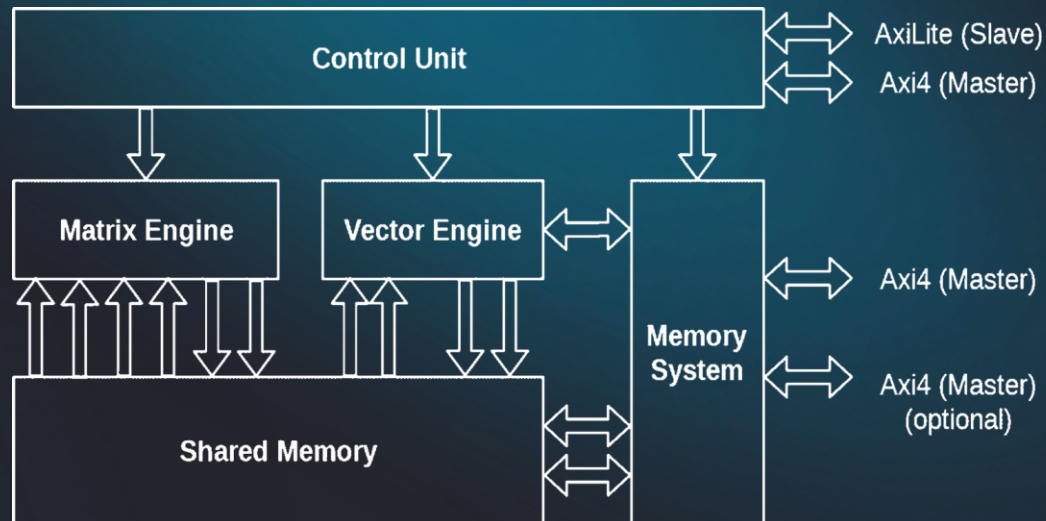
Контроллер DDR4+ECC

**Однопроцессорные  
модули HP****Многопроцессорные  
модули HPD/HPQ (2/4 чипа)****Модули HX для  
встраиваемых решений**



## Собственное IP-ядро

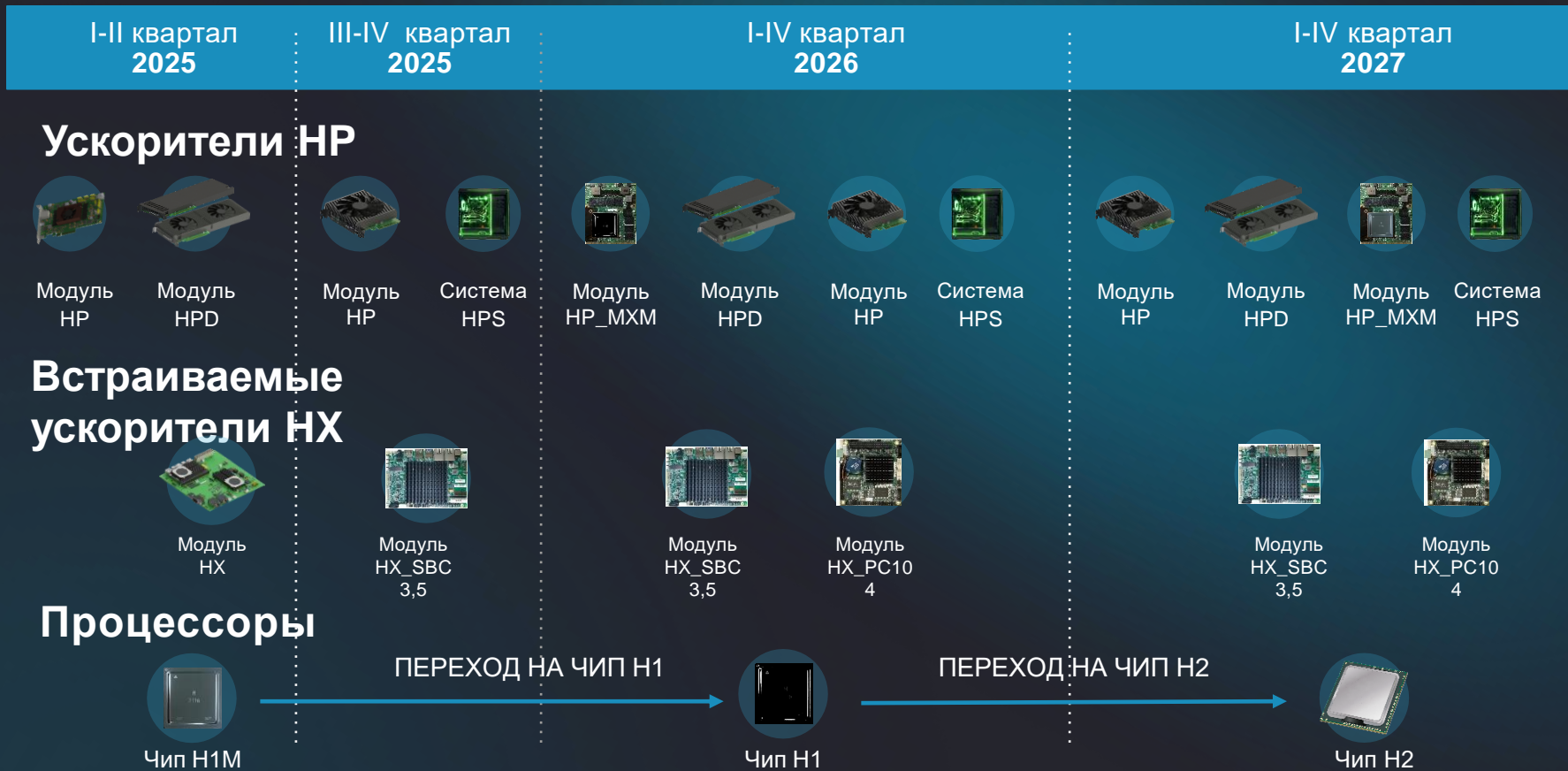
- **Полностью российское AXI\_4** совместимое IP-ядро
- Расчет (inference) сверточных (**CNN**) и рекуррентных (**RNN**) нейронных сетей
- Вычислительных элементов **16К**
- Максимальная рабочая частота **1 ГГц**
- Тип данных **int8**
- Встроенная память **до 16МБ**
- Производительность **24 TOPs (int8)** на систолическом массиве 128x128



# Линейка решений для инференса

Характеристики	HP2	HPD2	HPQ2	HX с сопроцессором общего назначения
Внешний вид				
Производительность пиковая	24 (int8) ТОПС	48 (int8) ТОПС	96 (int8) ТОПС	24 (int8) ТОПС
ResNet-50 (batch 1)	460 FPS	920 FPS	1840 FPS	460 FPS
ResNet-50 (batch 8)	1046 FPS	2092 FPS	4184 FPS	1046 FPS
Частота ускорителя тензорных вычислений	до 812 МГц	до 812 МГц	до 812 МГц	до 812 МГц
Интерфейсы	PCIe 2.0 x8 в Endpoint режиме	PCIe 3.0 x16 в Endpoint режиме	PCIe 3.0 x16 в Endpoint режиме	802.11a/b/g/n/ac; MIPI-SCI (2 канала); MIPI_DSI + TP; eDP; SDMMC; Audio codec
Форм-фактор	PCIe x16, 1 слот	PCIe x16, 1 слот	PCIe x16, 1 слот	Mini-ITX
Энергопотребление	20-23 Вт	<55 Вт	<105 Вт	<35 Вт

# Планы по развитию продукта



## Поддерживаемые сети

- ✓ densenet121
- ✓ efficientnet\_b0
- ✓ efficientnet\_b1
- ✓ efficientnet\_b2
- ✓ efficientnet\_b3
- ✓ inception\_v1
- ✓ inception\_v2
- ✓ inception\_v3
- ✓ inception\_v4
- ✓ lenet5
- ✓ mobilenet\_v1
- ✓ mobilenet\_v2
- ✓ nasnet\_mobile
- ✓ pnasnet\_large
- ✓ pnasnet\_mobile
- ✓ resnet152
- ✓ resnet34
- ✓ resnet50
- ✓ resnet50\_mlperf
- ✓ resnet50\_v2
- ✓ squeezenet
- ✓ ssd\_mobilenet\_v1
- ✓ ssd\_mobilenet\_v2
- ✓ scrfd 2.5g
- ✓ tiny\_yolo2
- ✓ tiny\_yolo3
- ✓ torch\_densenet169
- ✓ torch\_mobilenet\_v2
- ✓ torch\_resnet50
- ✓ vgg19
- ✓ xception
- ✓ yolo2
- ✓ yolo3
- ✓ yolo4
- ✓ Yolo5s
- ✓ ... и другие



# Производительность

Сети (выборочно)		Batch	Модуль HP2		Модуль HPQ	
			Performance, FPS	Latency, ms	Performance, FPS	Latency, ms
1	Resnet-50 v1.5 MLperf	8	818,126	-	3272,504	-
		1	424,902	3.923	1699,608	3,923
2	Resnet-50	8	910,461	-	3641,844	-
		1	431,559	3.877	1726,236	3,877
3	Yolo2	8	-	-	-	-
		1	88,095	17.322	352,38	17,322
4	Yolo3	8	-	-	-	-
		1	31,315	42.712	125,26	42,712
5	Yolo4	8	-	-	-	-
		1	1,824	560.907	7,296	560,907
6	Yolo5s	8	-	-	-	-
		1	10.768	108,132	43,072	108,132
7	SSD Mobilenet v2	8	457,762	-	1831,048	-
		1	141,980	9,366	567,92	9,366
8	SCRFD 2,5g	8	110,823	-	443,292	-
		1	75,145	23,033	300,58	23,033

Измерения на частоте 750 МГц

## Сравнение удельной производительности

Модель	Производительность	Потребление	Удельная пр-ть	Техпроцесс
LinQ HP2 (1 чип)	24 TOPS	20 Вт	1,2 TOPS/Вт	28 nm
LinQ HPD (2 чипа)	48 TOPS	40 Вт	1,2 TOPS/Вт	28 nm
Nvidia T4	130 TOPS/65 TFLOPS	70-100 Вт	1,85-1,3 TOPS/Вт	7 nm
Nvidia Jetson Xavier NX	1 TFLOPS	7,5 Вт	0,1 TOPS/Вт	16 nm
Nvidia Jetson Xavier NX	22 TOPS	30 Вт	0,7 TOPS/Вт	12 nm
SOPHON BM1684	17,6 TOPS / 2,2TFLOPS	19 Вт	0,92 TOPS/Вт	12 nm
SOPHON BM1684X	32 TOPS/16TFLOPS	55 Вт	0,58 TOPS/Вт	12 nm

# Оригинальный framework



# Web-интерфейс для удаленной компиляции и квантования

## Подготовка к инференсу:

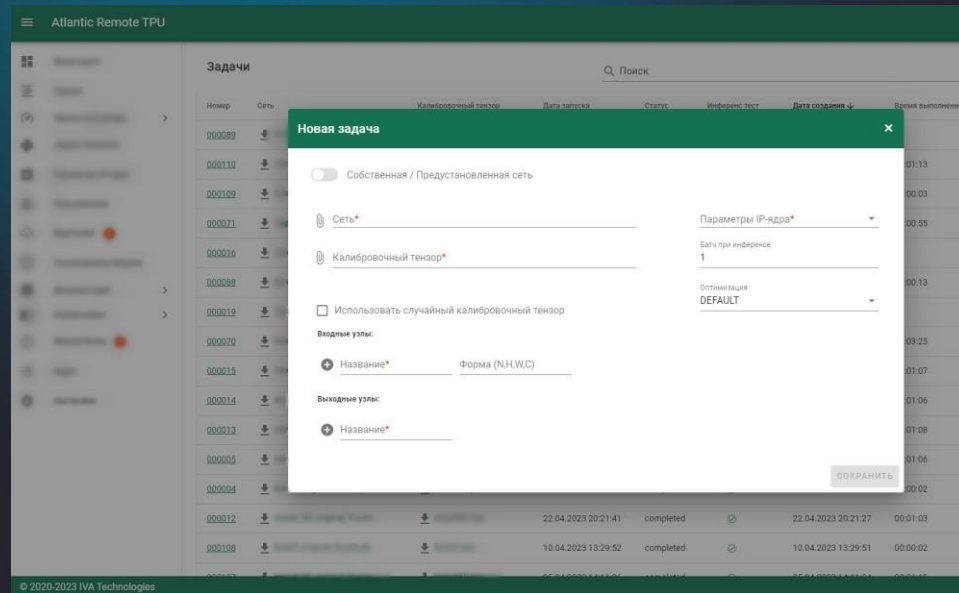
- ✓ Квантование модели ИНС
- ✓ Проверка точности квантованной модели
- ✓ Компиляция модели ИНС для TPU

## Инференс:

- ✓ Эмуляция пре/пост-процессинга данных (Jupyter Notebook)
- ✓ Инференс ИНС на TPU

## Аналитика:

- ✓ Снятие метрик производительности и точности инференса на TPU
- ✓ Проверка совместимости программных стеков





## Области применения



Беспилотный транспорт,  
автономная навигация



Интеллектуальная  
дорожная  
инфраструктура



Помощь в принятии  
решений, ЦОДы общего и  
специального назначения



Интеллектуальная  
видеоаналитика



Биометрия



Информационная  
безопасность



Управление БПЛА,  
группами объектов и т.п.



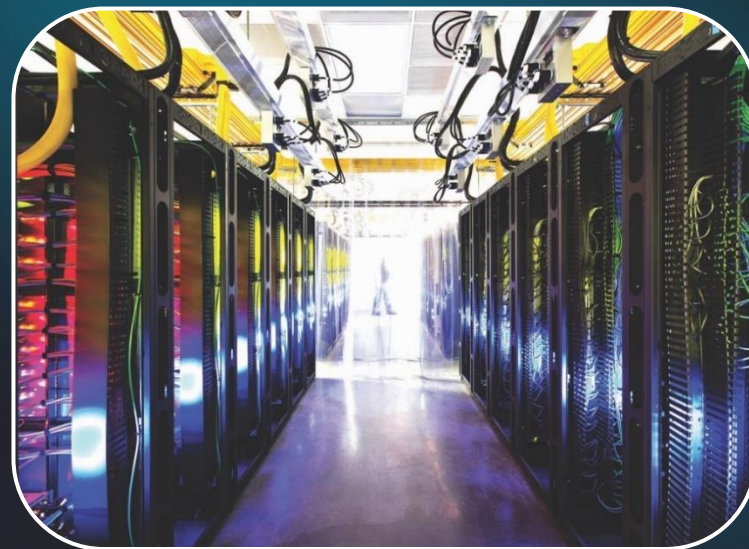
Умный город

# Области применения

Обеспечение принятия решений  
в центрах управления  
общего и специального назначения



Информационная безопасность,  
доверенные ПАК КИИ



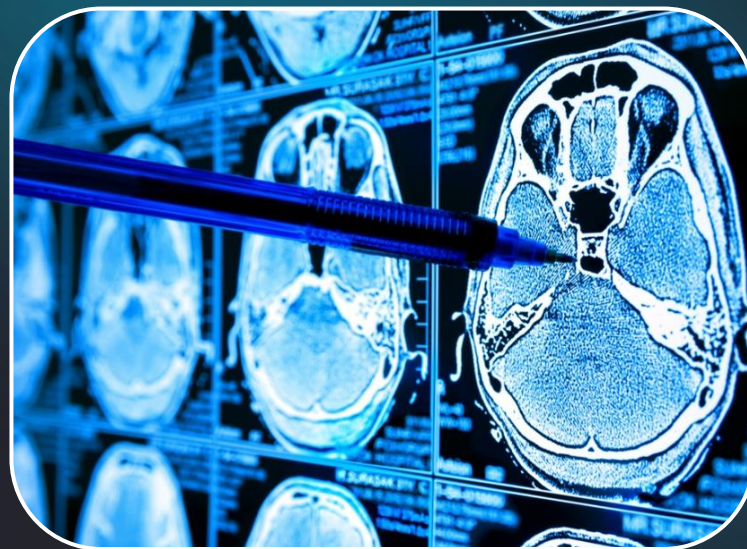


# Области применения

Контроль доступа, предупреждение о запрещенных действиях, нахождение в запрещенных пространствах



Обработка медицинских данных, помощь врачу в выявлении и прогнозировании болезней



# Области применения

Городские биометрические системы



Повышение эффективности досмотровых систем безопасности





# Области применения

Контроль соблюдения ПДД



Система Безопасный город

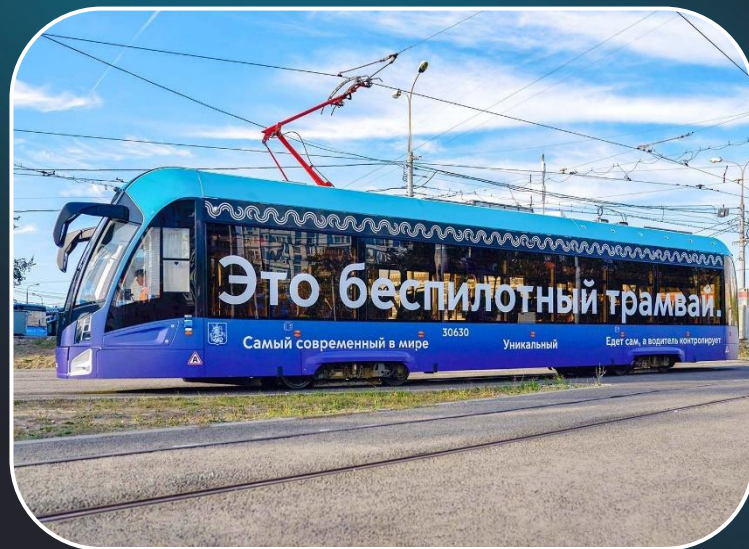


# Области применения

Метро без машиниста



Беспилотный наземный транспорт





ДИЗАЙН-ЦЕНТР  
МИКРОЭЛЕКТРОНИКИ



+7 (495) 137 59 95

[info@lin-q.ru](mailto:info@lin-q.ru)

[lin-q.ru](http://lin-q.ru)

Москва, Волгоградский пр-т, д. 42,  
Технополис «Москва»