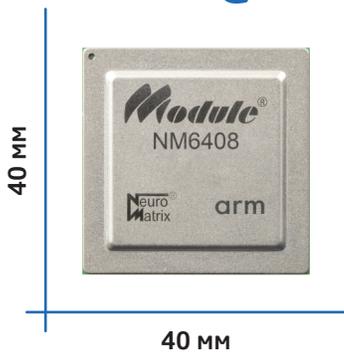


Module

NM Quad



на базе
4 СБИС K1879ВМ8Я
(NM 6408)



Высокопроизводительный вычислительный модуль на основе четырех многоядерных DSP-процессоров на базе оригинальной архитектуры NeuroMatrixCore 4 с управляющим RISC-процессором ARM Cortex A5 в форм-факторе PCIe add-in card x4 (2 слота).

Решает задачи цифровой обработки больших массивов данных в режиме реального времени. Работает с 32- и 64-разрядными данными с плавающей точкой одинарной и двойной точности.

Технические характеристики

Процессор и топология:	Интерфейсы и память:	ПО и NN:
<p>DSP: 64 ядра NMC4 – 1 ГГц</p> <p>RISC: 20 ядер Arm Cortex A5 – до 800 МГц L2 Cache – 512 Кб</p> <p>Производительность: FP32 – 2 TFLOP/s FP64 – 0.5 TFLOP/s</p> <p>Техпроцесс: 28 нм КМОП</p> <p>Мощность (типичная/максимальная): 50 Вт / 100 Вт</p> <p>Температурный диапазон: -20С +50С</p> <p>Корпус СБИС: 1444 HFCBGA</p> <p>Размеры: 277 x 143 x 39 мм</p>	<p>20 Гб памяти DDR3L (до 32 Гб/с) PCIe 2.0 (4 lanes)</p> <p>Форм-фактор PCIe x16, 2 слота расширения</p> <p>Суммарная пропускная способность интерфейсов межпроцессорного обмена одного процессора: 160 Гбит/с</p> <p>Внутренняя память одного процессора: 76 Мбит</p>	<p>NMC SDK: IDE (open VS-code), компиляторы, отладчики ARM и NMC Vare Metal SDK БЗИО, БУПВ (MPI, GAS), OpenCL Набор специализированных библиотек математических функций (BLAS, ЦОС)</p> <p>Драйверы: Windows (7,10), Linux (включая "AstraLinux" и "Эльбрус ОС")</p> <p>Neuro Matrix Deep Learning: Фирменный компилятор ИНС DarkNet и ONNX. Библиотека для применения ИНС. Библиотека трансляции моделей. Библиотека для подготовки изображений. Набор утилит для применения.</p>

Решает задачи



Benchmark NN inference

Ниже приведены значения производительности модуля при реализации (Inference) некоторых современных глубоких нейронных сетей.

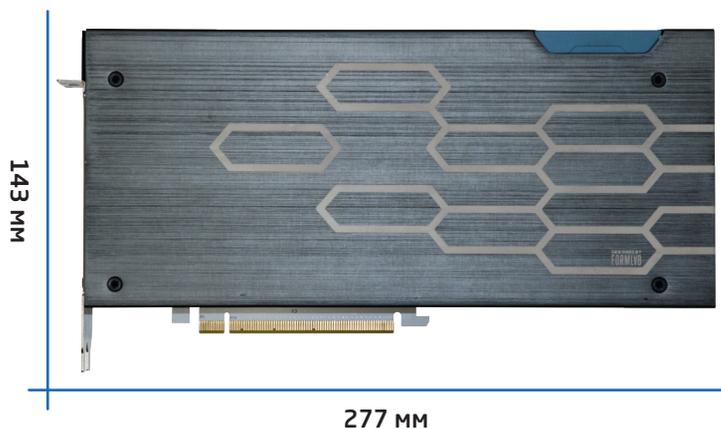
FPS - число обработанных кадров изображения в секунду.

Latency - задержка до получения обратного кадра.

Batch-mode - режим одновременной обработки нескольких кадров.

Нейронная сеть (размер изображения в пикселях)	FPS NM Quad (multi unit mode)	Latency NM Quad (multi unit mode)	FPS NM Quad (batch mode)	Latency NM Quad (batch mode)
alexnet (227x227)	50,4	79	52	308
inception v3 (299x299)	51,2	78	81,2	197
inception v3 (512x512)	15,72	254	21,76	735
resnet 18 (224x224)	100	40	188	85
resnet 50 (224x224)	48,8	82	82,4	194
squeezenet (224x224)	297,6	13	400	40
u-net (512x512)*	8	500	8	2000
yolo v2 tiny (416x416)	84	47	121,6	132
yolo v3 (416x416)	14,8	270	18	889
yolo v3 tiny (416x416)	109,2	36	141,2	113
yolo v5s (640x640)	13,2	303	13,6	1176

Габариты модуля



Масса модуля < 1,5 кг.

