

# Разработка 3D фотон-электронной матричной нейросетевой реконфигурируемой платформы для высокопроизводительной обработки информации

Валерий Сведе-Швец (ooooes@mail.ru)

В статье рассматривается проект по созданию нейросетевого инструмента для высокопроизводительной обработки информации общегосударственного применения на основе фотонной технологии.

## Задачи проекта:

- изготовление по кремниевой фотонной технологии ООО «ОЭС» трёхмерной фотон-электронной матрицы СБИС с нейроморфной/нейросетевой архитектурой – 3D ФЭ МНП СБИС с многоканальными аналогово-цифровыми фотонными и электрическими коммутационными связями;
- изготовление по кремниевой фотонной технологии ООО «ОЭС» трёхмерной фотон-электронной матрицы СБИС лазеров вертикального излучения – 3D ФЭ МЛВ СБИС;
- изготовление по кремниевой фотонной технологии ООО «ОЭС» трёхмерной фотон-электронной матрицы СБИС стандарта SW – 3D ФЭ MSW СБИС;
- изготовление 3D фотон-электронных матричных процессорных модулей – 3D ФЭ МПМ;
- изготовление механических разъёмных корпусов с матричными оптическими линзовыми растрами, многоканальных волоконных и призматических мультиплексных элементов;
- адаптация программного обеспечения с открытым кодом;

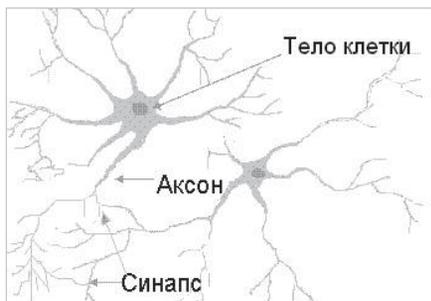


Рис. 1. Модель биологического нейрона

- изготовление 3D фотон-электронной матричной нейроморфной/нейросетевой платформы – 3D ФЭ МНП – и её модификаций для высокопроизводительной обработки информации общегосударственного применения.

В октябре 2019 года президентом РФ была принята Национальная стратегия развития искусственного интеллекта на период до 2030 года. Борьба за лидерство в развитии искусственного интеллекта предполагает рассмотрение методов и средств машинного обучения и связанных с ними технологий для всестороннего и успешного развития задач национальной безопасности.

Несмотря на совершенствование интегральных технологий, традиционные процессоры по-прежнему не способны эффективно решать нелинейные и неформализованные задачи. В то же время мозг животных и человека с такими задачами справляется отлично. Объяснение здесь кроется в качественных архитектурных различиях

между традиционным процессором и мозгом живых существ. Механизмы передачи информации в мозгу до сих пор не до конца изучены, и нет готового решения по созданию идеальной искусственной нейронной сети.

## Оценка биологической активности человеческого мозга

У человека порядка 100 миллиардов клеток мозга. Число связей нейрона с соседними оценивается в 1000, и при каждой активизации нейрона его импульс достигает тысячи других нейронов и далее по цепочке (рис. 1). По примерным оценкам нейрон связывается со своим соседом каждые 5 миллисекунд, что приблизительно равно 200 раз в секунду.

Оценим активность человеческого мозга: 100 миллиардов (количество нейронов) умножаем на 200 (секундная работоспособность), умножаем на 1000 (количество соединений), получаем скорость работы мозга, равную 20 квадриллионам операций в секунду.

Мозг вырабатывает энергию, равную лампочке 10 Ватт.

Предположительная ёмкость мозга в электронных терминах составляет около 1000 терабайт.

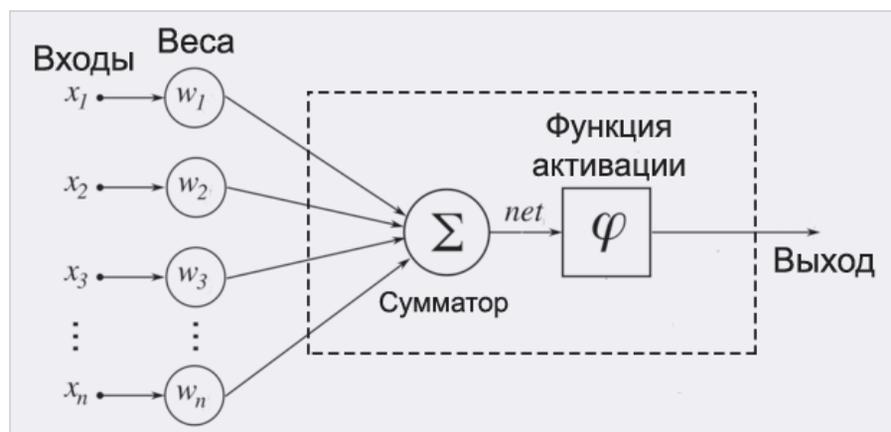


Рис. 2. Искусственный нейрон

## Структура искусственного нейрона

Искусственный нейрон является упрощённой моделью естественного нейрона.

Математически искусственный нейрон обычно представляют как некоторую нелинейную функцию от единственного аргумента: линейной комбинации всех входных сигналов.

Данную функцию называют функцией активации. Полученный результат посылается на единственный выход. Такие искусственные нейроны объединяют в сети: соединяют выходы одних нейронов со входами других (рис. 2).

Поступившие на входы искусственного нейрона сигналы умножаются на свои веса (веса изображены кружками). Сигнал первого входа  $x_1$  умножается на соответствующий этому входу вес  $w_1$ . В итоге получаем  $x_1w_1$ . И так до  $n$ -го входа. В итоге на последнем входе получаем  $x_nw_n$ . Теперь все произведения передаются в сумматор, а он просто суммирует все входные сигналы, умноженные на соответствующие веса:  $x_1w_1 + x_2w_2 + \dots + x_nw_n = \sum_{i=1}^n x_iw_i$ .

Искусственные нейроны объединяются между собой определённым образом, образуя искусственную нейронную сеть (ИНС) с различным уровнем слоёв (рис. 3).

Многослойная ИНС позволяет решать задачи любой формы и сложности. При этом преимущество нейронных сетей – универсальность: они превращают разные задачи в однотипные.

## Проблемы в аппаратной реализации ИНС

Синапс – сетевое соединение: количество синапсов растёт квадратично с ростом числа нейронов.

Вес синапсов: массы должны быть определены с высокой точностью, для того чтобы обеспечить правильную сходимость алгоритмов.

## Преимущества нейрокомпьютера

Искусственные нейроны и сети являются основными элементами нейрокомпьютера.

Все алгоритмы нейроинформатики высокопараллельны, а это залог высокого быстродействия.

Нейросистемы можно сделать устойчивыми к помехам и разрушениям.

Устойчивые и надёжные нейросистемы могут создаваться и из ненадёж-

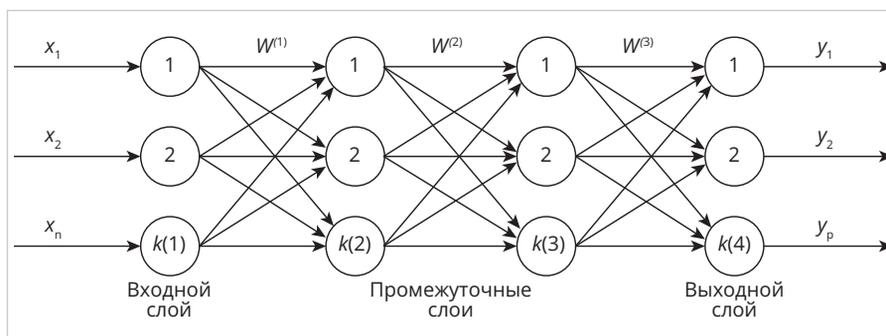


Рис. 3. Искусственная нейронная сеть

ных элементов, имеющих значительный разброс параметров.

Сегодня нет готовых решений по созданию идеальной искусственной нейронной сети.

## Системный анализ архитектуры нейропроцессора и классификация нейрочипов

Элементной базой нейровычислителей служат нейрочипы. Большинство из них ориентировано на конкретные специализированные управляющие системы.

Нейрочипы делятся:

- по типу логики – на цифровые, аналоговые и гибридные;
- по типу реализации нейроалгоритмов – с полностью аппаратной и с программно-аппаратной реализацией (когда нейроалгоритмы хранятся в ПЗУ);
- по характеру реализации нелинейных преобразований – на нейрочипы с жёсткой структурой нейронов (аппаратно-реализованные) и нейрочипы с настраиваемой структурой нейронов (перепрограммируемые);
- по возможностям построения нейросетей – нейрочипы с жёсткой и переменной нейросетевой структурой (т.е. нейрочипы, в которых топология нейросетей реализована жёстко или гибко).

Процессорные матрицы (систолические процессоры) обычно близки к обычным RISC-процессорам; они объединяют в своём составе некоторое число процессорных элементов, вся же остальная логика, как правило, должна быть реализована на базе периферийных схем.

В отдельный класс следует выделить так называемые нейросигнальные процессоры, ядро которых представляет собой типовой DSP-процессор.

Нейронным процессорам, построенным на основе сигнальных, тензорных,

систолических чипов для формирования и управления нейронной сетью, требуется внешняя память и управляющий процессор.

Существует огромное количество способов соединения нейронов, растущее с увеличением числа нейронов в сети. Наиболее употребительной является слоистая архитектура, в которой нейроны располагаются «слоями». В наиболее общем случае аксоны каждого нейрона одного слоя направлены к нейронам следующего слоя. Таким образом, нейроны первого слоя являются входными (принимающими информацию из внешнего мира), нейроны последнего слоя – выходными (выдающими информацию во внешний мир). Другой вид архитектуры – полносвязная, когда каждый нейрон соединён с каждым, в том числе сам с собой.

## Процессор NeuroMatrixR NM6404

NeuroMatrixR NM6404 (рис. 4) представляет собой высокопроизводительный DSP-ориентированный RISC-микروпроцессор. В его состав входят два основных блока: 32-разрядное RISC-ядро и 64-разрядное VECTOR-сопроцессор для поддержки операций над векторами с элементами переменной разрядности. NM6404 по системе команд совместим с предыдущей версией NM6403. Имеются два идентичных программируемых интерфейса для работы с внешней памятью различного типа и два коммуникационных порта, аппаратно-совместимых с портами ЦПС TMS320C4x, для возможности построения многопроцессорных систем.

Особенности:

- тактовая частота – 133 МГц (8 нс – время выполнения любой инструкции);
- технология КМОП 0,25 мкм;
- корпус PQFP256;

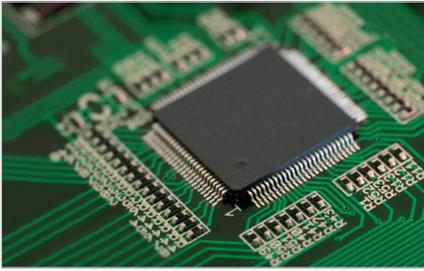


Рис. 4. Конструктивная реализация NeuroMatrixR NM6404

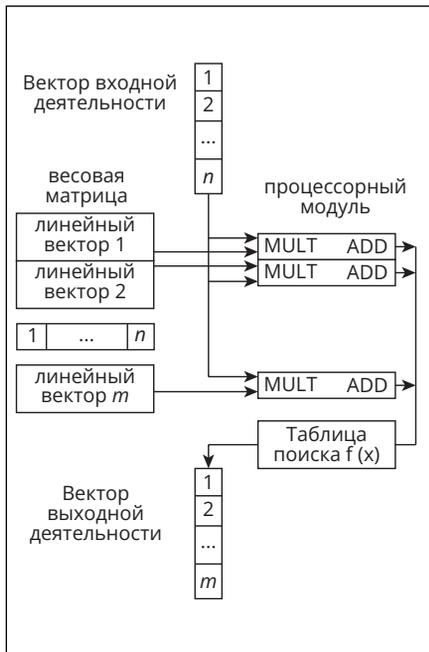


Рис. 5. Структура систолического процессора SAND

- напряжение питания от 2,5 В, 3,3 В, 5 В;
- потребляемая мощность – около 1,0 Вт;
- условия эксплуатации:  $-40...+80^{\circ}\text{C}$ . RISC-ядро:
- пятиступенчатый 32-разрядный конвейер;
- 32- и 64-бит команды (обычно выполняется две операции в одной команде);
- 2 Мбит внутреннее ОЗУ;
- доступ к внутренней памяти соседей;
- два адресных генератора, адресное пространство – 16 ГБ;
- два 64-разрядных программируемых интерфейса с SDRAM/SRAM/DRAM/Flash ROM разделяемой памятью;
- 4 одновременных доступа к внутренней памяти; широкополосный режим доступа к внешней памяти;
- 64k Boot ROM; формат данных – 32-разрядные целые; 4 канала DMA;
- два коммуникационных порта ввода/вывода, аппаратно-совместимых с портами TMS320C4x;

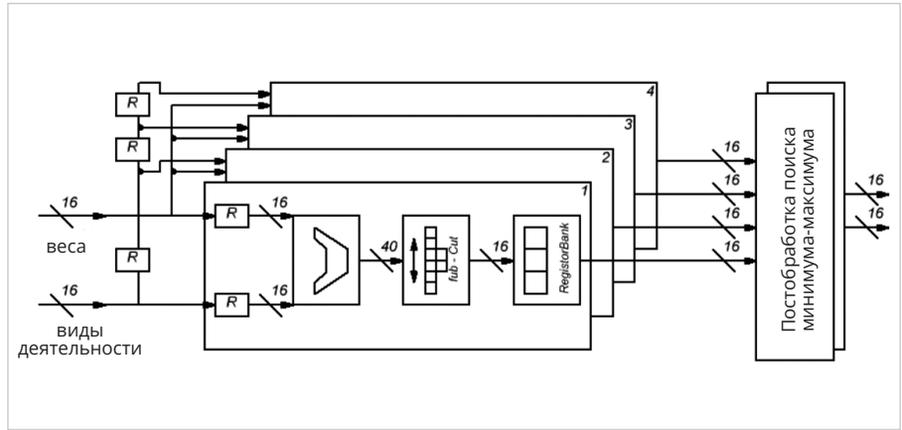


Рис. 6. Структура чипа SAND

- JTAG-совместимый отладочный интерфейс;
  - система управления потребляемой мощностью.
- VECTOR-процессор:*
- от 1 до 64-разрядная длина векторных операндов и результатов;
  - формат данных – целые числа, упакованные в 64-разрядные блоки, в форме слов переменной длины от 1 до 64 разрядов каждое;
  - поддержка векторно-матричных и матрично-матричных операций; 16 тактов на перезагрузку матрицы коэффициентов;
  - свопирование рабочей и теневой матриц; два типа функций насыщения на кристалле.

*Производительность*  
 Скалярные операции:

- 133 MIPS;
  - 399 MOPS для 32-разрядных данных;
- Векторные операции:
- от 133 до более чем 38 000 ММАС (миллионов умножений с накоплением в секунду);
  - I/O и интерфейсы с памятью:
    - пропускная способность двух 64-разрядных интерфейсов с памятью 2128 Мбайт/с;
    - I/O коммуникационные порты – до 20 Мбайт/с каждый.

### Систолические нейропроцессоры

Этот класс вычислительных систем создавался с ориентацией на применение в области нейросетей. Основная идея построения систолических процессоров состоит в использовании специальных обрабатывающих элементов, простых по своим функциям и структуре. Эти элементы образуют процессорную матрицу, через которую идёт поток данных, изменяемых каждым элементом. При этом может быть достигну-

та высокая степень параллельности обработки данных, если отработавший элемент сразу же считывает следующую порцию данных для обработки. Сеть работает циклически, и в каждом последующем цикле каждый обрабатывающий элемент считывает и обрабатывает новую порцию данных, независимо от работы остальных обрабатывающих элементов. За счёт этого достигается высокая степень распараллеливания процессов обработки информации и в результате высокая скорость работы всей сети в целом. При этом в нейропроцессорах используется ограниченный набор вычислений – преимущественно свёртка и перемножение матриц, что открывает большой простор для оптимизаций.

### Базовые принципы построения систолических архитектур:

- систола представляет собой сеть связанных вычислительных ячеек, обычно простых;
- каждая ячейка содержит в себе буферный входной регистр, защёлкивающий данные, и вычислитель, оперирующий с содержимым этого регистра. Выход вычислителя может подаваться на входы других ячеек;
- операции в систоле производятся по типу конвейерной обработки;
- вычисления в систоле регулируются с помощью общего тактового сигнала;
- результатом правильного построения систолы должна быть простая регулярная разводка с простой топологией связей.

### Систолический процессор SAND

Чип SAND (Simple Applicable Neural Device) разработан для применения в промышленных и исследователь-

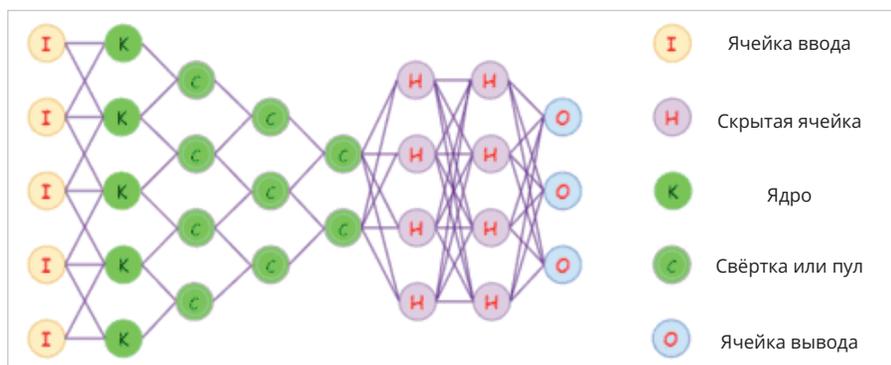


Рис. 7. Архитектура свёрточных нейронных сетей

ских системах реального времени. Чип способен реализовывать нейросети с максимальным числом входов 512 (рис. 5, 6).

Чип SAND содержит четыре параллельных обрабатывающих элемента PE (Processing Elements), каждый из которых снабжён АЛУ и блоками отсека (auto-cut). АЛУ используется для умножения векторов. Так как АЛУ накапливает входные значения, выходная шина имеет разрядность 40 бит (это ограничивает число входных нейронов до 512). Блок отсека снижает разрядность до 16 с контролем переполнения и потери точности. Окно, выбирающее 16 бит из 40, может быть смещено пользователем по своему усмотрению.

Нейрокомпьютеры с нейроморфными процессорами представляют собой одну из перспективных разработок в области вычислительной техники, так как они намного эффективнее решают нелинейные и неформализованные задачи по сравнению с традиционными процессорами.

Нейроморфные вычисления отличаются от классических подходов к ИИ, которые основаны на свёрточных нейронных сетях (CNN), тем, что они гораздо точнее имитируют мозг с помощью импульсных нейронных сетей (SNN).

### Свёрточные нейронные сети (CNN)

Свёрточные нейронные сети (Convolutional Neural Network, CNN) – это нейронные сети, которые показали высокую точность в классификации и кластеризации изображений, а также в распознавании объектов, хотя применяются практически везде. CNN состоят из двух видов слоёв: слоёв свёртки и пулинга. Слой пулинга необходим для уменьшения размерности. Преимущество свёрточных сетей заключается в

их свойстве инвариантности, т.е. объект на изображении может находиться в любом месте, но сеть его всё равно найдёт (рис. 7).

Применение CNN:

- распознавание образов,
- «компьютерное зрение» (computer vision),
- видеоанализ.

### Структура нейроморфного процессора

Нейроморфные процессоры – это устройства, аппаратно моделирующие работу импульсных нейронных сетей (SNN). В таких сетях моделируемые нейроны, как и реальные биологические нейроны, общаются друг с другом, используя электрические импульсы – спайки. При этом SNN обладают большей «вычислительной мощностью», чем сети предыдущих поколений (рис. 8).

В нейроморфных процессорах искусственные нейроны объединяются по модели импульсных нейронных (спайковых) сетей SNN, особенностью которых является передача с помощью разнесённых по времени коротких импульсов равной амплитуды. По аналогии с биологическим образцом искусственный нейрон – один выход (аксон), сигнал с которого может поступать на большое количество входов других нейронов и тем самым изменять их состояние.

### Обработка информации нейроморфным процессором в импульсной нейронной сети (spiking neural network – SNN)

Каждый «нейрон» в сети SNN может срабатывать независимо от других: он отправляет импульсные сигналы другим нейронам в сети, которые напрямую изменяют электрические состояния этих нейронов. Кодирова

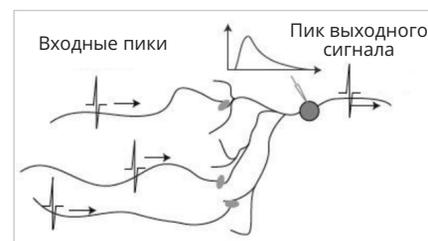


Рис. 8. Структура нейроморфного процессора

информацию в самих сигналах, SNN моделируют естественные процессы обучения, динамически переключая синапсы между искусственными нейронами в ответ на стимулы.

Хотя нейроморфные чипы, как правило, всё ещё являются цифровыми, они в большинстве случаев работают на основе асинхронных цепей, что означает отсутствие глобальной синхронизации. В зависимости от конкретного приложения нейроморфные вычисления могут быть на несколько порядков быстрее и требуют меньше энергии.

В настоящее время лидером по разработке полупроводникового нейроморфного процессора является корпорация INTEL с процессором Loihi 2. Процессор организован по матричной архитектуре ядер. Ядро содержит вычислительные элементы и память в качестве нейронов и синапсов. Системы на NPU Intel Loihi 2 учатся быстрее и эффективнее, что открывает перед ними массу перспектив.

При решении задач оптимизации и поиска выяснилось, что NPU Loihi 2 может решать задачи более чем в 1000 раз эффективнее и в 100 раз быстрее по сравнению с традиционными процессорами (рис. 9).

### Основные характеристики чипа Intel Loihi 2:

- техпроцесс – Intel 4;
- площадь кристалла – 31 мм<sup>2</sup>;
- площадь ядра – 0,21 мм<sup>2</sup>;
- количество транзисторов – 2,3 млрд;
- количество нейронных ядер на чип – 128;
- количество процессоров на чип – 6;
- количество нейронов на чип – 1 млн;
- количество синапсов на чип – 120 млн;
- память на нейронное ядро – 192 кБ, гибкое размещение;
- модели нейронов – полностью программируемые;
- область состояния нейрона – 0...4096 байт на нейрон;

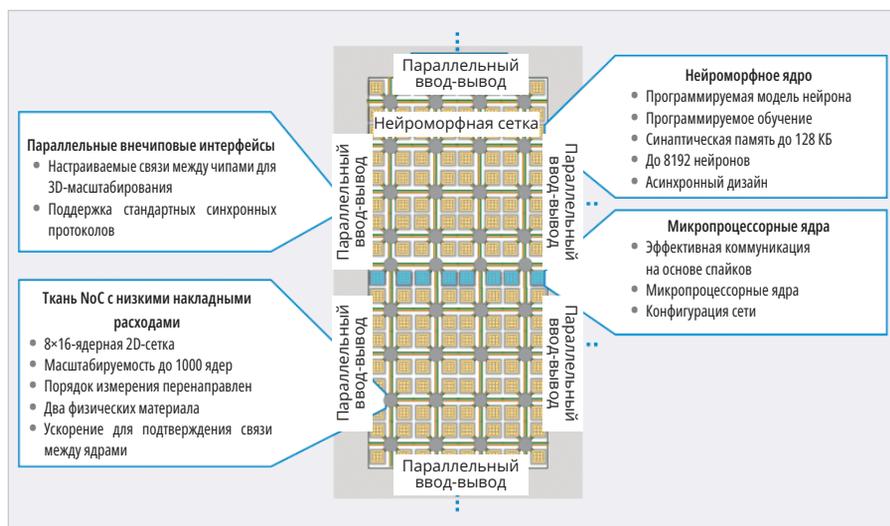


Рис. 9. Архитектура чипа Intel Loihi 2

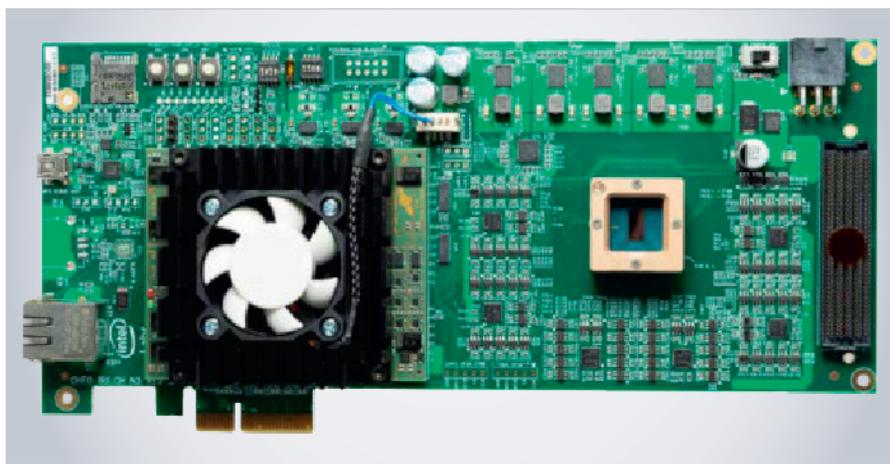


Рис. 10. Базовое устройство для демонстрации возможностей Loihi 2 – Oheo Gulch

- кодирование информации – ступенчатое состояние спайка, до 32 бит на импульс;
- внешние интерфейсы – стандартные протоколы SPI и AER, GPIO, Ethernet 1000BASE-KX, 2500BASE-KX.

Готово базовое устройство для демонстрации возможностей Loihi 2. Оно называется Oheo Gulch и представляет собой плату с одним чипом Loihi 2, а также FPGA Intel Arria 10, предоставляющим интерфейс и удалённый доступ через Ethernet к нейроморфному чипу (рис. 10).

Российская компания «Мотив НТ» разрабатывает собственный нейроморфный процессор «Алтай».

### Нейроморфный чип «Алтай»

Энергоэффективный нейропроцессор для интеллектуальных устройств «Алтай» – «вычислительное» устройство, функционирующее на принципах, схожих с биологическими нейронными системами.

Ключевые преимущества:

- высокая эффективность по энергопотреблению, производительности и размерам;
  - возможность решения неформализуемых и плохо формализуемых задач;
  - высокая масштабируемость, ограниченная только требованиями по энергопотреблению и массогабаритным параметрам;
  - отказоустойчивая архитектура;
  - отечественный.
- Основные модели применения:
- обработка большого потока разнородных сигналов с целью обнаружения аномалий в системах киберфизической безопасности и системах мониторинга физических объектов;
  - обработка видео- и аудиоизображений;
  - интегрирование и обработка сенсорной информации от различных датчиков и сенсоров в робототехнике;
  - обработка физиологических сигналов.
- Низкое энергопотребление, малый размер и высокая производительность

позволяют применять нейрочип «Алтай» в очень широком спектре интеллектуальных устройств (рис. 11–13).

Ключевые архитектурные решения чипа «Алтай»:

- полная цифровая реализация на современной монокристаллической кремниевой технологии;
- универсальная, но простая в реализации на кристалле цифровая модель нейрона;
- параметры функционирования нейронов и структура сети формируются вне кристалла;
- многоядерная архитектура; нейрочип представляет собой масштабируемую сеть нейроядер;
- нейрочип проектируется по модели GALS. Ядра являются синхронными схемами, каждое из которых функционирует в своём домене синхронизации. Все коммуникативные блоки нейрочипа являются асинхронными. Возможности, предоставляемые НП «Алтай»:

- исполнение произвольных импульсных нейронных сетей;
- неограниченная масштабируемость сети;
- обучение в процессе работы устройства;
- низкое энергопотребление по сравнению с классическими вычислительными устройствами;
- высокая производительность – обработка до 2000 кадров в секунду в задачах технического зрения;
- компактный и недорогой чип.

Изготовлен модуль нейроморфного акселератора с 8 прототипами нейроморфного процессора «Алтай» (рис. 14).

Недостатки современных полупроводниковых нейрочипов для нейропроцессоров:

- кристаллы изготовлены с применением кремниевой 2D-технологии с планарными электрическими связями;
- кристаллы устанавливаются на плату и ограничены по масштабированию электрическими связями;
- кристаллы изготовлены на наноразмерной, не существующей в России технологической базе.

### Фотоника для ИИ и нейроморфные фотонные системы

Созданные на сегодняшний день нейронные сети и нейроморфные процессоры весьма приближённо «копируют» работу биологических мозгов

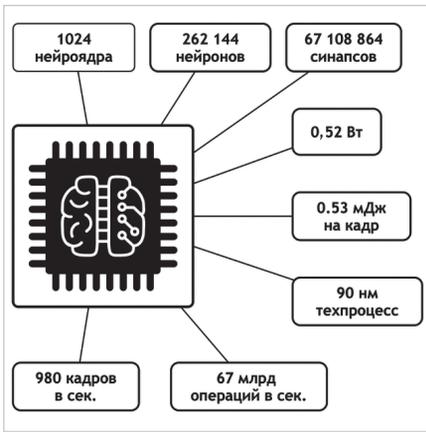


Рис. 11. Функциональные характеристики нейрочипа «Алтай»

даже животных, не говоря уже о функциональности головного мозга человека.

Попытки повторить на кремниевых электронных чипах структуру связей, существующую в биологическом мозге, до сих пор не увенчались успехом вследствие высочайшей сложности её построения из сотен миллионов искусственных нейронов, соединённых друг с другом так, что эти соединения в процессе обучения или получения нового опыта «переконфигурируются», создавая новые устойчивые сетевые структуры. Основная сложность повторения вроде бы понятного «узора» биологической нейронной сети заключается в физических ограничениях, присутствующих в полупроводниковой микро- и нанoeлектронике. Выход из этого «технологического тупика» учёные видят в развитии фотоники – оптическом аналоге электроники, в котором носителями сигналов являются фотоны света.

Технология оптоэлектроники уже достигла успеха в ряде областей, продемонстрировав черты самостоятельного приборостроения. Уже широко используются волоконно-оптические линии связи, оптическая и голографическая память большой ёмкости, сенсоры изображения и другие устройства.

Фотонные пространственные соединения в архитектурах вычислительных устройств и систем имеют ряд преимуществ по сравнению с электронными соединениями.

Основными из них являются:

- отсутствие электрического проводника, позволяющее осуществлять эффективные пространственные соединения «кристалл-кристалл», «плата-плата»;

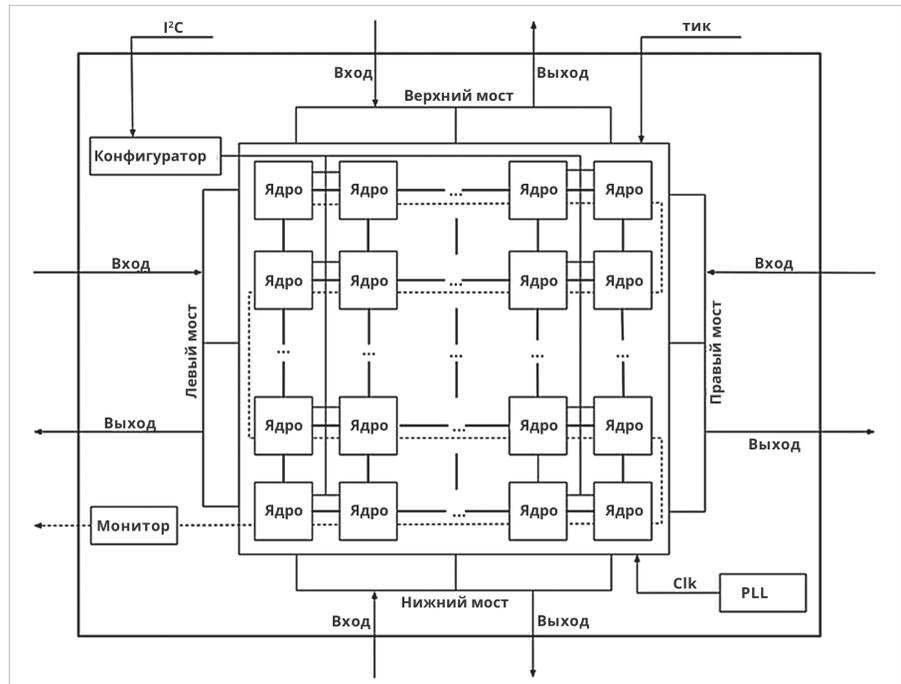


Рис. 12. «Алтай»: структурная схема нейропроцессора

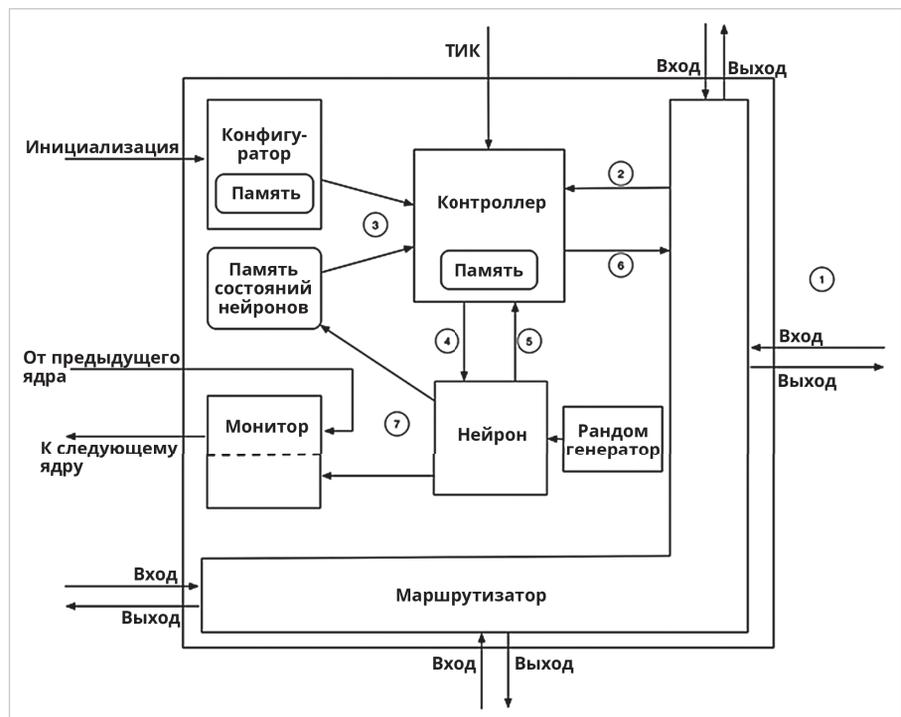


Рис. 13. «Алтай»: структурная схема нейроядра



Рис. 14. Модуль нейроморфного акселератора с 8 прототипами

- отсутствие взаимного влияния при пересечении лучей света в пространстве;
- несколько каналов с высокой пропускной способностью могут сосуществовать в одном пространстве;
- параллельная природа светового потока и способность к объединению обеспечивают создания гибких параллельных архитектур;
- отсутствие возможности перехватывать информацию, поскольку опти-

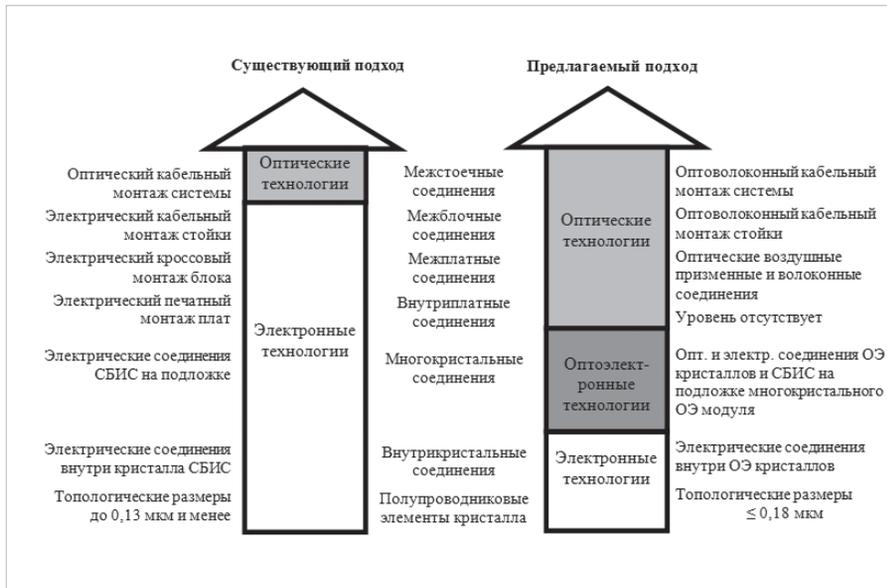


Рис. 15. Технология межсоединений

ческая система ничего не излучает в окружающую среду;

- устойчивость к электромагнитным помехам.

Кремниво-фотонная технология стала многообещающей КМОП-совместимой альтернативой для реализации нового поколения устройств и систем, которые могут использовать свет как для связи, так и для вычислений.

Кремниевая фотонная технология становится реальностью, а полупроводниковые микросхемы с интегральными микролазерами позволяют создавать 3D-процессоры с бесконтактными и волоконными многоканальными фотонными связями, осуществляя обмен данными как внутри, так и снаружи ПК с большой пропускной

способностью. Устройства на основе кремниевой фотоники могут обеспечить беспрецедентный уровень энергоэффективности и параллелизма. В настоящее время ведутся разработки в области оптоэлектронных устройств, плазмоники и нанофотоники, а также фотонно-аналоговой обработки информации и нейроморфных вычислений.

### В США создан первый фотонный нейроморфный чип

Одно из последних прорывных достижений в сфере создания нейроморфных чипов недавно продемонстрировали учёные Национального института стандартов и технологий США, представив двухслойную трёхмерную систему, состоящую из матриц

оптических волноводов, взаимодействующих при передаче оптических сигналов в разных направлениях. Это решение позволило многократно увеличить количество связей между активными компонентами (оптическими логическими элементами) по сравнению с электронными нейросетями, благодаря чему удалось повысить сложность схем маршрутизации сигналов и скорость обмена данными между узлами сети. В результате достигнута рекордная «схожесть» архитектуры нейронной фотонной сети со строением биологического мозга.

Суть изобретения фотонного нейроморфного чипа американскими учёными заключается в формировании на подложке из кремния двух слоёв матриц, состоящих из тончайших нитрид-кремниевых световодов (их поперечные сечения находятся в пределах 800×400 нм).

Также разработано ПО, обеспечивающее указание автоматического направления сигналов по нужным путям в схеме и регулировку уровней взаимосвязей между отдельными нейронами.

### 3D кремниво-фотонная технология ООО «ОЭС»

В объёмной (3D) фотон-электронной технологии, разработанной ООО «ОЭС», многоканальные оптические и электрические каналы реализуются непосредственно на обеих поверхностях полупроводниковых кристаллов.

На рис. 15 показан существующий и разработанный перспективный уровень технологий соединений.

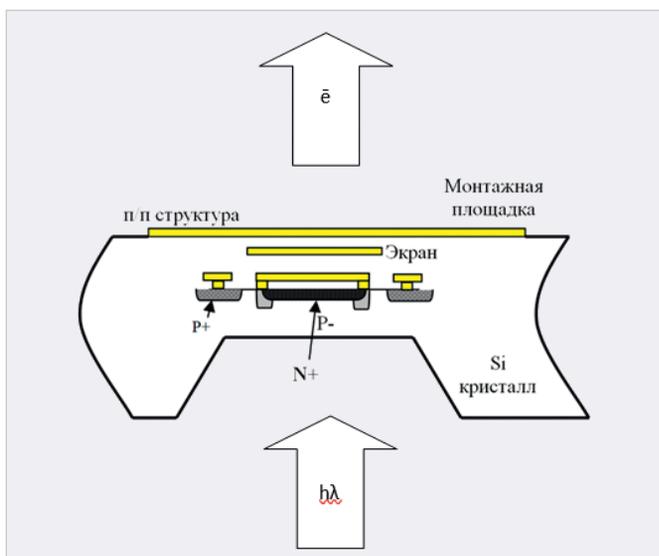


Рис. 16. Фотонный и электронный ввод-вывод информации в Si-кристалл 3D ФЭ матрицы функциональной СБИС

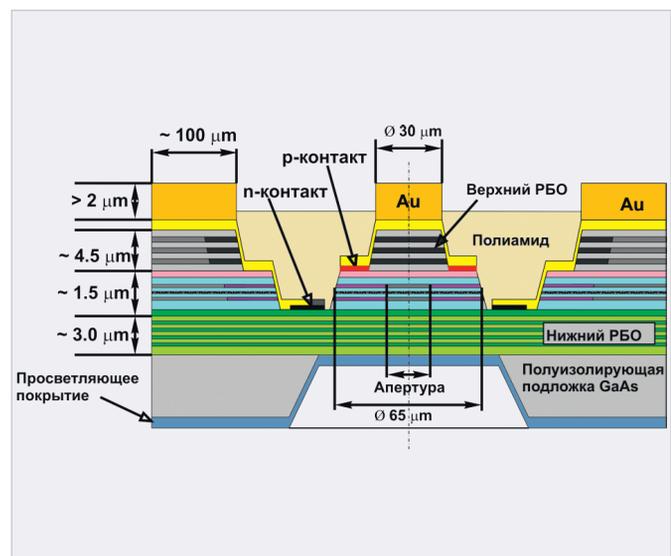


Рис. 17. Структурная схема 3D ФЭ МВЛ СБИС Si-кристалл 3D ФЭ матрицы функциональной СБИС

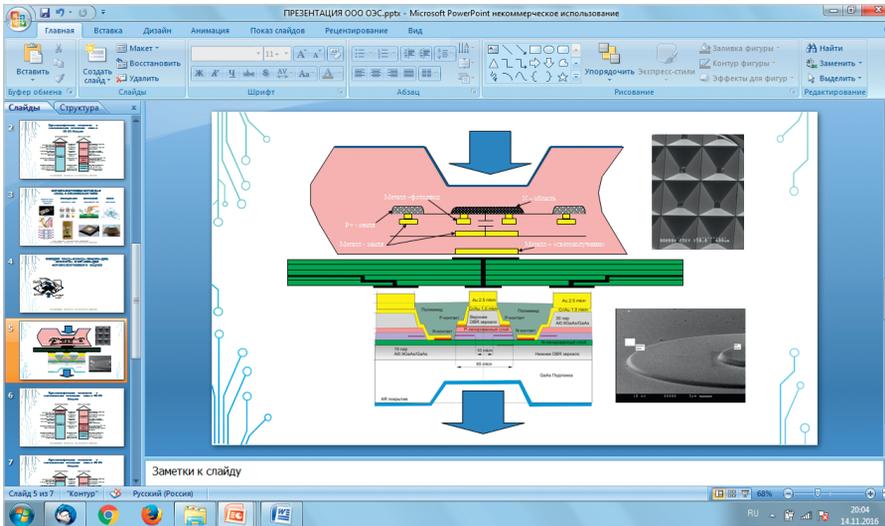


Рис. 18. Структура 3D ФЭ матрицы функциональной СБИС с оптическими и электрическими связями

Реализация матриц с многоканальными 3D ФЭ процессорными элементами требует разработки фотон-электронного интерфейса, обеспечивающего многоканальный оптический и электрический ввод-вывод информации на обеих поверхностях всей интегральной схемы. Реализация такого фотон-электронного интерфейса, обеспечивающего многоканальный ввод-вывод информации в интегральную схему, возможен, если для фотон-электронного многоканального ввода/вывода информации используются обе поверхности кристалла наряду с планарным электрическим интерфейсом ввода/вывода.

Такой подход позволяет реализовать принцип трёхмерной многоканальной связи и создавать функциональные 3D фотон-электронные полупроводниковые элементы с многоканальными (матричными) оптическими и электрическими каналами с потоковой обработкой информации для различных устройств и строить высокопроизводительные информационно-вычислительные системы с множеством датчиков.

Организация фотонного и электронного ввода-вывода информации с использованием двух сторон кристалла позволяет перейти к разработке кремниево-фотонной объёмной технологии и созданию многоканальных фотон-электронных соединений на уровнях «чип-чип», «плата-плата», «распределённая объектовая связь» для различных устройств и строить высокопроизводительные информационно-вычислительные системы со множеством датчиков (рис. 16, 17).

### 3D ФЭ матрицы функциональных СБИС с оптическими и электрическими связями для потоковой обработки информации

Матрицы функциональных СБИС с оптическими и электрическими связями имеют внутренние и внешние многоканальные связи. Многоканальная оптическая связь направлена ортогонально к плоскости кристалла по оси Z, а электрические связи по осям X, Y и Z относительно плоскости полупроводниковых СБИС.

Такая реализация многоканальных оптических и электрических связей в СБИС переводит их в категорию полупроводниковых приборов с объёмной, трёхмерной фотон-электронной архитектурой.

Трёхмерные фотон-электронные (3D ФЭ) матричные функциональные СБИС с оптическими и электрическими связями представляют собой прибор с интегральными полупроводниковыми схемами, фокальные плоскости которых обеспечивают ввод или вывод информации по многоканальным оптическим и электрическим каналам связи и потоковой обработкой информации.

Полупроводниковые кристаллы изготавливаются с применением кремниевой и арсенид-галлиевой технологий (рис. 18).

Кристаллы 3D ФЭ матриц функциональных СБИС с оптическими и электрическими связями по технологии «кристалл на плате» монтируются с двух сторон на LTCC-плату и закрываются герметичным корпусом с механическим разъёмом и многоканальными оптическими линзовыми растрами (рис. 19, 20).

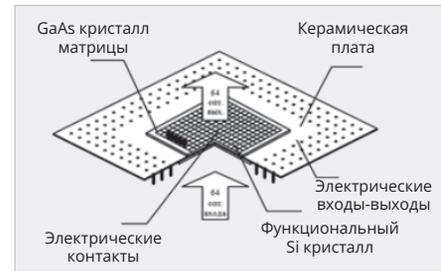


Рис. 19. Связь в 3D МОЭМ

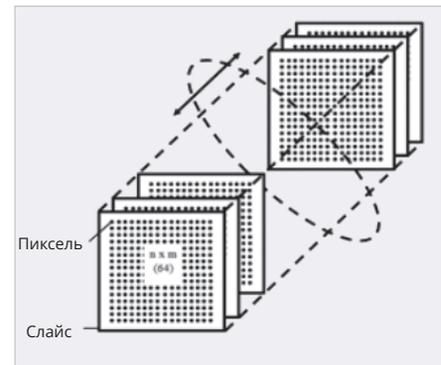


Рис. 20. Организация потоковой «вертикальной обработки»

3D фотон-электронная технология является комплексной технологией для реализации 3D ФЭ матриц функциональных СБИС с оптическими и электрическими связями на базе кремниевых и арсенид-галлиевых интегральных схем (рис. 21–25), которые реализуют следующие виды обмена информацией:

- оптический приём – логическая обработка, коммутация – оптический выход;
- оптический приём – логическая обработка, коммутация – электрический выход;
- электрический приём – логическая обработка, коммутация – оптический выход;
- электрический приём – логическая обработка, коммутация – электрический выход.

Основные конструктивные единицы 3D ФЭ МПМ:

- мезонинная LTCC-плата для монтажа компонентов 3D ФЭ МПМ;
- 3D ФЭ СБИС МНП в составе четырёх кристаллов;
- 3D ФЭ СБИС МВЛ в составе четырёх кристаллов;
- 3D ФЭ СБИС SW МНК в составе одного кристалла;
- корпус 3D ФЭ МКП с многоканальными линзовыми растрами;
- 3D ФЭ МКЦ – корпус цилиндрический с многоканальными линзовыми растрами;



### Концепция архитектуры потокового нейросетевого процессора на основе 3D фотон-электронной матричной элементной базы и пространственной многоканальной фотонной сети

Все современные нейронные вычислители построены по планарной технологии с электрическими интерфейсными связями, а это фундаментальный их недостаток. Биологическое строение мозга имеет объёмную конструкцию и многомерные всесторонние нейронные связи.

В нейронных вычислителях основная вычислительная нагрузка ложится на сетевые архитектуры. Кто создаст вычислительную платформу с идеальной нейронной сетевой архитектурой с фотонной многоканальной полносвязной реконфигурируемой связью, тот получит доступ к рыночному разнообразию внедрения изделия, как для гражданского, так и для специального применения (рис. 27).

### Матричный 3D ФЭ нейровычислитель на основе аналоговых оптических вычислений, элементов оптических систем и многоканальной, пространственной фотон-электронной и волоконно-оптической связи (рис. 28)

Состав нейровычислителя:

- три 3D ФЭ М ПМ с 256 ПЭ и многоканальной пространственной фотонной связью;
- функция обработки – трёхуровневая, аналогово-цифровая;
- базовый 3D ФЭ М ПМ – 3 шт.;
- оптический волоконный многоканальный пространственный преобразователь – 6 шт.;
- фотонный сумматор – оптическая линза – 3 шт.

### 3D ФЭ матричная нейроморфная среда ООО «ОЭС» – 3D ФЭ МНС с фотонной многоканальной импульсной нейроморфной сетью (SNN) на основе 3D ФЭ МПП и 3D ФЭ МНП СБИС с реформируемым нейроморфным чипом «Алтай»

Нейроморфный чип российской разработки «Алтай» и его матрица нейроядер конструкторско-технологически реформируется под объёмный

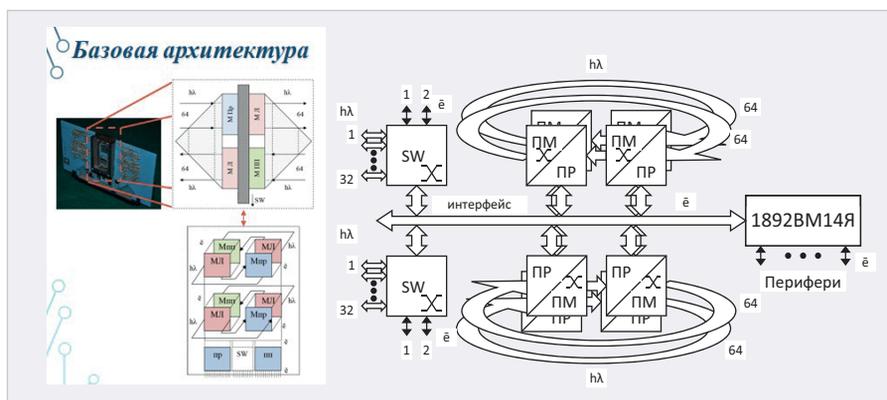


Рис. 24. Архитектура 3D ФЭ МПП с 3D ФЭ матрицами функциональных СБИС с оптическими и электрическими связями

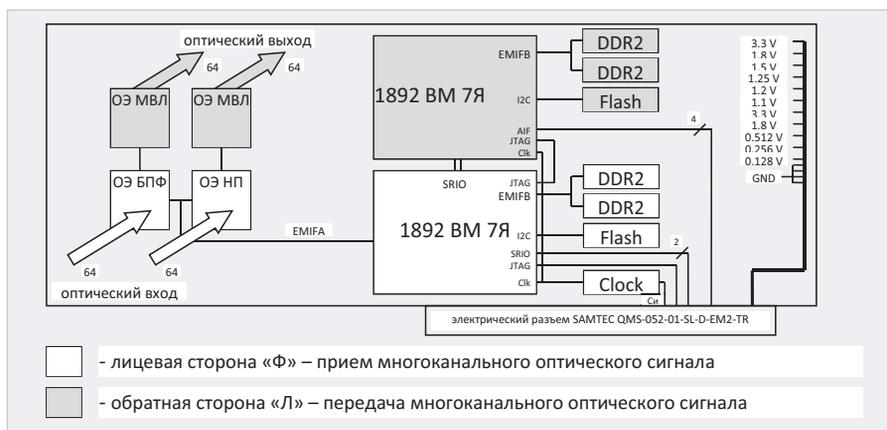


Рис. 25. Структурная схема 3D ФЭ МПП с 3D ФЭ матрицами функциональных СБИС с оптическими и электрическими связями

фотон-электронный матричный нейроморфный процессорный СБИС – 3D ФЭ МНП СБИС с матрицей АЦП/ЦАП и неблокирующим коммутатором матрицы выходных фотонных сигналов по кремниево-фотонной технологии ООО «ОЭС».

Чип «Алтай» получает в каждое нейроядро матрицы входной плоскости чипа фотонный импульсный сигнал с каналной аналого-цифровой величиной информационного сигнала до 256 бит.

После алгоритмической обработки входных сигналов из каждого нейроядра матрицы выходной плоскости чипа передаются цифро-аналоговые электрические импульсные сигналы для матрицы лазеров вертикального излучения с каналной величиной информационного сигнала до 256 бит.

3D ФЭ МНП СБИС устанавливается на объёмные фотон-электронные матричные процессорные платы – 3D ФЭ МПП.

Архитектура матричной пространственно-временной нейроморфной среды – 3D ФЭ МНС с шестнадцатью 3D ФЭ МПП каждая, с четырьмя/восе-



Рис. 26. Фотография 3D ФЭ МПП с 3D ФЭ матрицами функциональных СБИС с оптическими и электрическими связями

мью 3D ФЭ МНП СБИС, 3D ФЭ МВЛ СБИС и оптическими призмными мультиплексными элементами представлена на рис. 29–31.

### Фотонные информационные связи в матричной пространственно-временной нейроморфной среде – 3D ФЭ МНС

- В матричной пространственно-временной потоковой нейроморфной среде 3D ФЭ МНС располагается шестнадцать 3D ФЭ МПП. Каждая

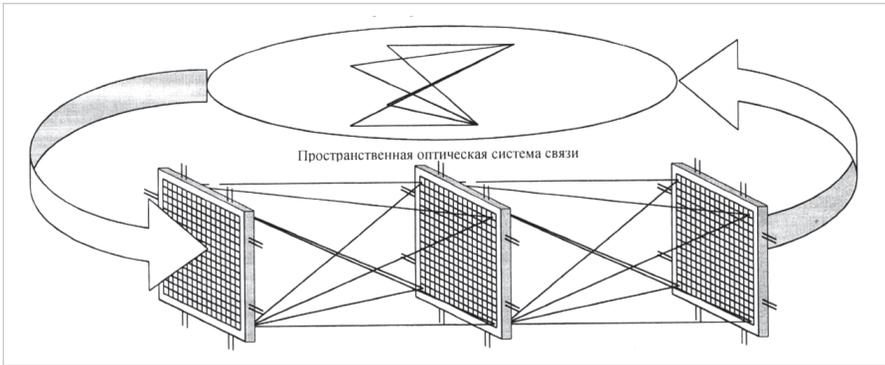


Рис. 27. Концепция архитектуры матричного нейронного процессора

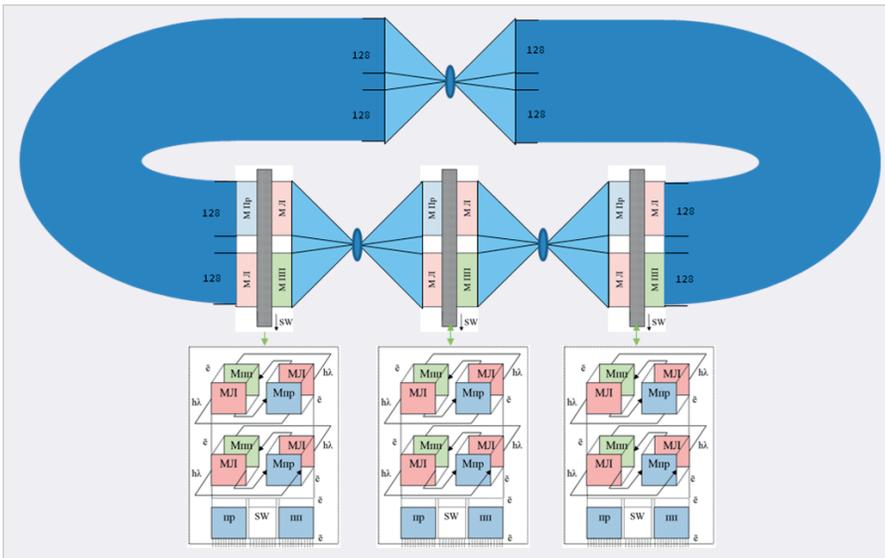


Рис. 28. Матричный 3D ФЭ 256-ядерный нейровычислитель

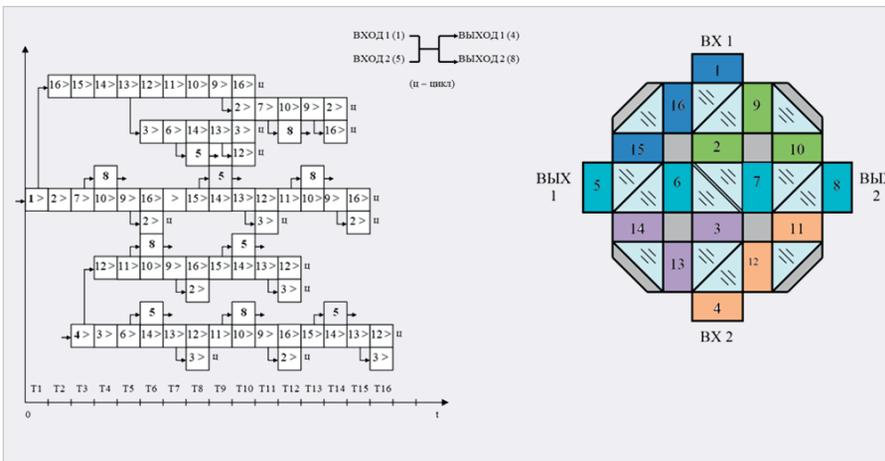


Рис. 29. Структурная схема и информационные потоки 3D ФЭ МНС

3D ФЭ МПП содержит четыре/восемь 3D ФЭ МНП СБИС с матрицами 32x32 нейронных ядер и 3D ФЭ МВЛ СБИС с матрицами 32x32 лазеров вертикального излучения. Каждый выход матриц нейронных ядер 3D ФЭ МНП СБИС в 3D ФЭ МПП связан через оптическую призмную систему с входом с одной или двумя матрицами нейронных ядер 3D ФЭ МНП СБИС в 3D ФЭ МПП и их 1024

или 2048 нейронными ядрами. Так же организована оптическая связь и у других 3D ФЭ МПП с их матрицами нейронных ядер.

- В матричной пространственно-временной потоковой нейроморфной среде – 3D ФЭ МНС, содержащей шестнадцать 3D ФЭ МПП с четырьмя или восемью 3D ФЭ МНП СБИС с матрицами 32x32 нейронных ядер. В 3D ФЭ МПП имеется 16 777 216 или

33 554 432 нейрона с 4 294 967 296 или 8 589 934 592 синапсами.

- В матричной пространственно-временной потоковой нейроморфной среде – 3D ФЭ МНС, содержащей четыре или восемь 3D ФЭ МНП СБИС в 3D ФЭ МПП, в случае реализации в 3D ФЭ МНП СБИС неблокирующего коммутатора фотонных выходных каналов 3D ФЭ МНС реализует архитектуру произвольного формирования структуры нейроморфных сетей с многократным увеличением фотонных информационных связей.
  - В матричной пространственно-временной потоковой нейроморфной среде – 3D ФЭ МНС, содержащей четыре или восемь 3D ФЭ МНП СБИС в 3D ФЭ МПП, возможна реализация интерфейсных электрических связей между 3D ФЭ МНП СБИС на 3D ФЭ МПП, что, соответственно, увеличивает гибкость информационных потоков в нейроморфной сети 3D ФЭ МНС.
  - 3D ФЭ матричная фотонная нейросетевая оптическая платформа ООО «ОЭС» – 3D ФЭ МНОП строится на 3D ФЭ МПП и 3D ФЭ МНП СБИС с тензорной, систолической, нейроморфной архитектурой и реализует свёрточные нейронные сети (CNN) или импульсные нейроморфные сети (SNN) (рис. 32).
- Конструктивная единица 3D ФЭ МНОП*
- В 3D ФЭ матричной нейросетевой оптической платформе – 3D ФЭ МНОП – располагается четыре 3D ФЭ МПП. Каждый 3D ФЭ МПП содержит четыре 3D ФЭ МНП СБИС с тензорной, систолической или нейроморфной архитектурой, связанных между собой по электрическому интерфейсу.
  - 3D ФЭ МНП СБИС содержит матрицу 32x32 нейронных ядер.
  - Каждый фотонный вход-выход четырёх матриц нейронных ядер 3D ФЭ МНП СБИС 3D ФЭ МПП соединён через фокон-линзовую оптическую систему и оптическую мультиплексную призму с фотонным входом-выходом двух 3D ФЭ МНП СБИС 3D ФЭ МПП и их 512 нейронными ядрами. Так же организована фоконно-линзовая оптическая система и у других матриц нейронных ядер 3D ФЭ МНП СБИС 3D ФЭ МПП.
  - В 3D ФЭ матричной нейросетевой оптической платформе – 3D ФЭ МНОП матрица выходных сигналов нейронных ядер 3D ФЭ МНП СБИС 3D ФЭ.

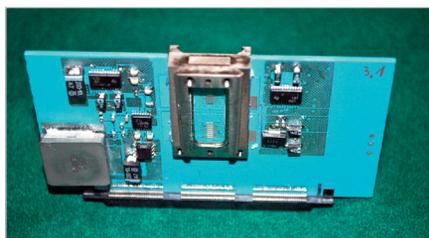


Рис. 30. Фотография конструкции 3D ФЭ МПП

● МПМ суммируется с помощью оптической линзовой системы и передаётся во входной канал принимающей нейронной матрицы. Программа решает, индивидуально или полностью принять результат выхода нейрона с матрицы выходных сигналов через фокусную оптическую систему: сразу с 512 соседними нейронами ядрами или индивидуально с помощью не блокирующего коммутатора. Так же организована многоканальная связь и с другими матрицами нейронов 3D ФЭ МНП СБИС.

### Фотонные информационные связи в 3D ФЭ матричной нейросетевой оптической платформе – 3D ФЭ МНБП(МНОП)

В 3D ФЭ МНБП(МНОП) каждый 3D ФЭ МПМ содержит четыре 3D ФЭ МНП СБИС с матрицей процессоров тензорной, систолической или нейроморфной архитектуры, связанных между собой по электрическому интерфейсу.

Матричные процессоры – 3D ФЭ МНП СБИС размерностью  $32 \times 32 = 1024$  ПЭ с нейронными ядрами.

$1024 \times 1024 \times 2 = 2\,097\,152$ . Прямые фотонные связи с нейронными ядрами.

$2\,097\,152 \times 4 = 8\,388\,608$ . Прямые фотонные связи между четырьмя кристаллами 3D ФЭ МНП СБИС с нейронными ядрами.

$8\,388\,608 \times 4 = 33\,554\,432$ . Прямые фотонные связи между четырьмя 3D ФЭ МПМ с четырьмя 3D ФЭ МНП СБИС размерностью  $32 \times 32 = 1024$  ПЭ с нейронными ядрами.

$8\,388\,608 \times 256 = 2\,147\,483\,648$ . Прямые фотонные связи с нейронами в 3D ФЭ МНБП(МНОП).

$2\,147\,483\,648 \times 256 = 549\,755\,813\,888$ . Прямые фотонные связи с синапсами в 3D ФЭ МНБП(МНОП).

В 3D ФЭ матричной нейросетевой оптической платформе – 3D ФЭ МНБП/МНОП реализуется:

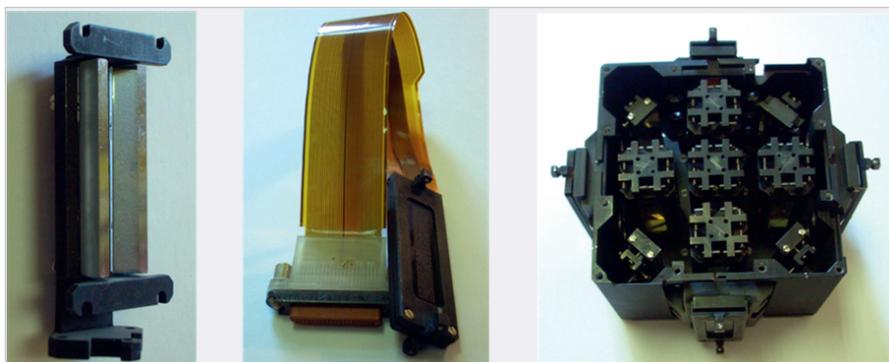


Рис. 31. Фотография конструкции матричной пространственно-временной потоковой нейроморфной среды – 3D ФЭ МНС с шестнадцатью функциональными 3D ФЭ МПП и оптическим призмным мультиплексным элементом

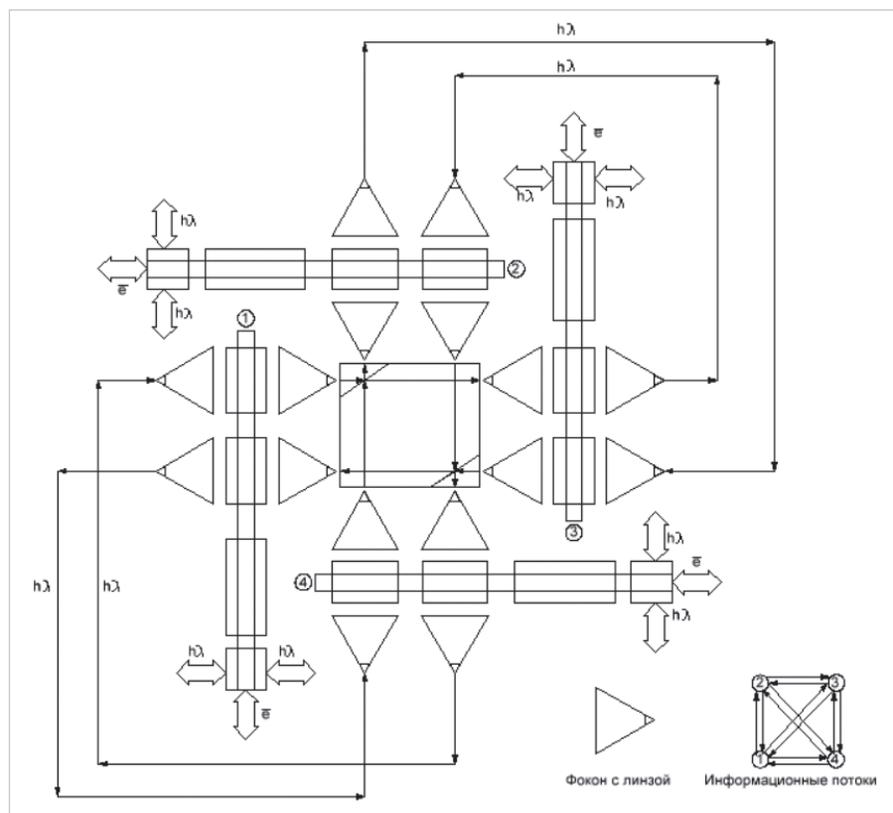


Рис. 32. Архитектура 3D ФЭ матричной нейросетевой оптической платформы ООО «ОЭС» – 3D ФЭ МНОП

● 2 147 483 648 (два миллиарда сто сорок семь миллионов четыреста восемьдесят три тысячи шестьсот сорок восемь) фотонных связей с нейронами;

● 549 755 813 888 (пятьсот сорок девять миллиардов семьсот пятьдесят пять миллионов восемьсот тринадцать тысяч восемьсот восемьдесят восемь) фотонных связей с синапсами.

Базовая платформа ООО «ОЭС» – 3D ФЭ МНБП(МНОП) с фотонной многоканальной связью на основе 3D ФЭ МПМ и 3D ФЭ МНП СБИС и оптических призмных мультиплексных элементов для систем искусственного интеллекта с приме-

нением как свёрточных нейронных сетей (CNN), так и импульсных нейроморфных сетей (SNN) является 3D ФЭ МНБП потоковой, пространственно-временной, перестраиваемой архитектурой и включает в себя (рис. 33):

- базовый 3D ФЭ МПМ – 4 шт.;
- оптический мультиплексор 3D ФЭ МОМ с призмами – 4 шт.

Число удалённых абонентов с двухсторонней связью и электромагнитной защитой – 256.

3D ФЭ матричный нейросетевой кластер ООО «ОЭС» – 3D ФЭ МНС с фотонной многоканальной связью на основе 3D ФЭ МПМ и 3D ФЭ МНП СБИС и оптических призмных мульт-

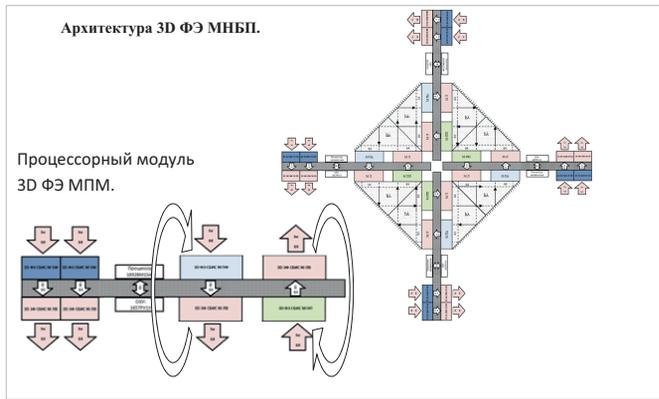


Рис. 33. 3D ФЭ матричная нейросетевая базовая платформа ООО «ОЭС» – 3D ФЭ МНБП

типлексных элементов для систем искусственного интеллекта с применением как свёрточных нейронных сетей (CNN), так и импульсных нейроморфных сетей (SNN) включает в себя 3D ФЭ МНБП – 4 шт (рис. 34).

### Фотонные информационные связи в 3D ФЭ матричном нейросетевом кластере ООО «ОЭС» – 3D ФЭ МНК

3D ФЭ МНК содержит четыре 3D ФЭ МНБП, каждый имеет 2 147 483 648 прямых фотонных связей с нейронами и 549 755 813 888 прямых фотонных связей с синапсами.

В 3D ФЭ матричном нейросетевом кластере ООО «ОЭС» – 3D ФЭ МНК реализуется:

- $2\ 147\ 483\ 648 \times 4 = 8\ 589\ 934\ 592$  (восемь миллиардов пятьсот восемьдесят девять миллионов девятьсот тридцать четыре тысячи пятьсот де-

вяносто две) фотонные связи с нейронами;

- $415\ 226\ 380\ 288 \times 4 = 1\ 660\ 905\ 521\ 152$  (один триллион шестьсот шестьдесят миллиардов девятьсот пять миллионов пятьсот двадцать одна тысяча сто пятьдесят две) фотонные связи с синапсами.

Сервер на базе 3D ФЭ с матричными нейросетевыми кластерами – 3D ФЭ МНК и потоковой, пространственно-временной, перестраиваемой архитектурой способен реализовать любой уровень ИИ.

3D кремниво-фотонная технология ООО «ОЭС» реализует 3D фотонно-электронные матричные, многоканальные связи в обеих плоскостях чипов, изготовленных по кремниевой или арсенид-галлиевой технологии. 3D ФЭ нейрочипы для 3D ФЭ нейропроцессоров более эффективны, чем нейрочипы для нейропроцессоров, изготовленных по

2D полупроводниковой технологии с планарными электрическими выводами. Изделия с 3D кремниво-фотонной технологией ООО «ОЭС» полностью реализуются на российских предприятиях.

3D кремниво-фотонная технология ООО «ОЭС» и 3D ФЭ конструкторские и технологические решения позволяют создавать 3D ФЭ изделия с ИИ и применять их в специальных системах, например, беспилотных летательных аппаратах и в гражданских системах для создания сервера с программно-сетевой адаптацией и большим объёмом памяти искусственного мозга.

Серверы с ИИ для медицинской нейроинформатики позволяют устранять непрофессиональные диагнозы врачей и избежать таких трагедий, которая произошла с моей женой, которую я потерял, прожив совместно 50 лет счастливой жизни.

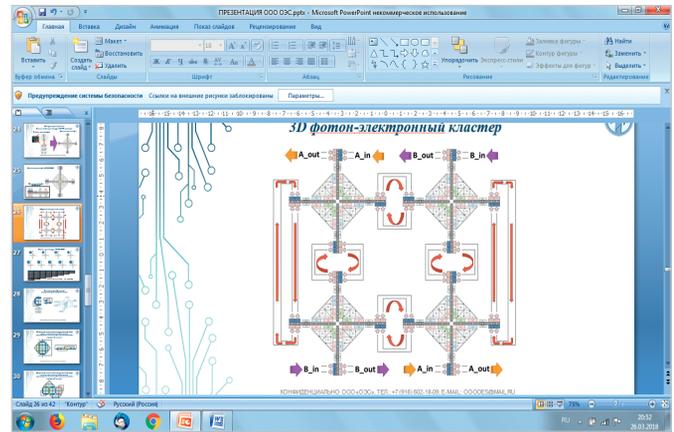


Рис. 34. 3D ФЭ матричный нейросетевой кластер – 3D ФЭ МНК

## НОВОСТИ МИРА

### США подготовили новые санкции в отношении 30+ компаний технологического сектора КНР

Американская администрация намерена включить в «чёрный список» китайского производителя чипов флэш-памяти Yangtze Memory Technologies Corp (YMTC) и свыше 30 других китайских компаний, что закроет им доступ к закупкам определённых типов продукции, сообщило в среду агентство Bloomberg.

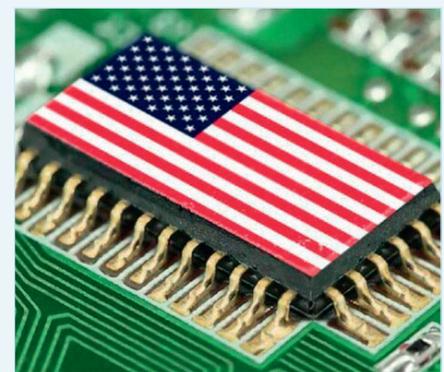
По информации источников Bloomberg, министерство торговли США внесёт YMTC и другие китайские компании в список организаций и лиц, действующих вопреки национальной безопасности и внешнеполитическим интересам

США (US Entity List), «уже на этой неделе».

Находящимся в списке компаниям блокируется доступ к американским технологиям. Поставлять их можно только при наличии специальной экспортной лицензии от министерства торговли США.

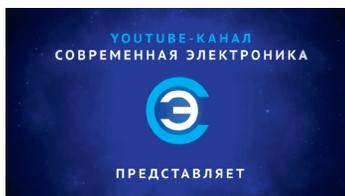
Напомним, в ноябре из-за политического давления Apple отказалась закупать чипы флэш-памяти у YMTC, альтернативным поставщиком была выбрана Samsung Electronics. В сенате США Apple пригрозили беспрецедентными проверками в случае покупки чипов у китайцев для iPhone 14.

Отметим, США последовательно подрыывают возможности КНР по созданию передовой полупроводниковой продукции. Вашингтон формирует антикитайские альянсы, переманивает мировых произво-



дителей чипов на свою территорию, объявляет компании КНР угрозой национальной безопасности, ограничивает доступ ведущим китайским производителям к американским технологиям.

industry-hunter.com



Смотрите на канале СОВРЕМЕННАЯ ЭЛЕКТРОНИКА



*Сергей Пилкин, ЭРЕМЕКС*



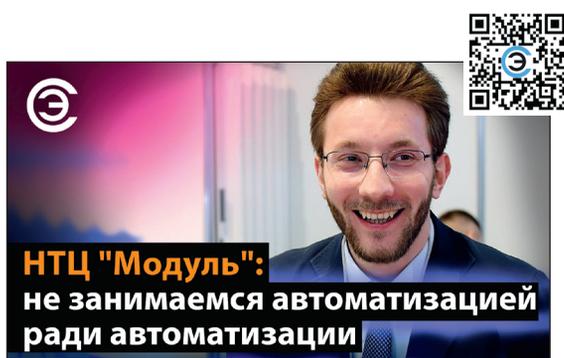
*Лев Теверовский, АСКОН*



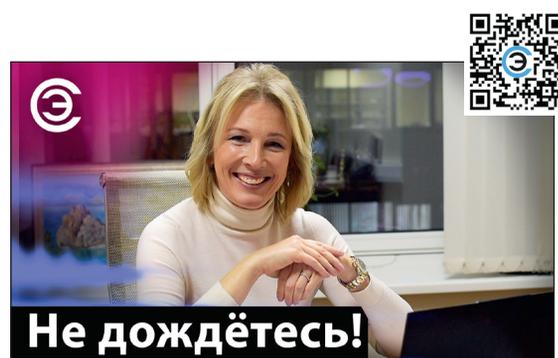
*Николай Александров, Dannie*



*Иван Ларионов, АРПЭ, компания «Третий тин»*



*Халиль Эль-Хажж, НТЦ «Модуль»*



*Светлана Легостаева, АНО «Консорциум «Вычислительная техника» / Часть 1*