

Перспективы развития информационно-вычислительных и радиифотонных систем на базе 3D М ФЭ ПМ

Часть 1

Валерий Сведе-Швец (sooovs@mail.ru), Владислав Сведе-Швец, Станислав Сведе-Швец (Москва)

В статье представлены предложения по развитию трёхмерных фотон-электронных информационно-вычислительных и радиифотонных систем на базе 3D матричных фотон-электронных процессорных модулей (3D М ФЭ ПМ).

В первой части представлен обзор технологий построения архитектур вычислительных систем для выбора и создания 3D М ФЭ ПМ с сетевой архитектурой.

Микроэлектронная технология последнего поколения увеличила интеграцию систем-на-кристалле, частично перенесла электрические связи, ранее располагавшиеся на печатной плате, и, как следствие, повысила производительность различных устройств, однако не решила проблему соединений – связь по-прежнему остаётся электрической. Сегодня, чтобы обеспечить внешние соединения с интегральной системой-на-кристалле, необходимо реализовать электрические связи на уровне корпуса, которые насчитывают уже более 1000 электрических выводов. Рост их числа обусловлен увеличением количества линий сигнальных интерфейсов на кристалле, а также необходимостью многократных связей по питанию и земле. При этом тактовая частота внутри кристалла превышает значение тактовой частоты вне кристалла в 4–10 раз. Проблема электрических связей снова перенесена на многослойную печатную плату, она остаётся центральной и в современных программируемых логических интегральных схемах (ПЛИС), и в микропроцессорах с многоядерной архитектурой.

Микропроцессор содержит миллионы транзисторов, соединённых между собой тончайшими проводниками из алюминия или меди и используемых для обработки данных. В результате микропроцессор выполняет множество функций – от математических и логических операций до управления работой других микросхем и всего компьютера. Главные параметры функционирования микропроцессора – частота

работы кристалла, определяющая количество операций за единицу времени, частота работы системной шины, объём внутренней кэш-памяти SRAM. Частота работы кристалла определяется частотой переключений транзисторов. Возможность транзистора переключаться быстрее определяется технологией производства кремниевых пластин, из которых изготавливаются чипы. При использовании техпроцесса формируются слои кремния. Каждый слой имеет свой собственный рисунок, в совокупности все эти слои образуют трёхмерную электронную схему процессора – например, при 90 нм техпроцессе было сформировано 7 слоёв кремния.

В настоящее время разработанная IBM совместно с GF и Samsung технология 5 нм практически получила теоретически достижимый минимум для кремния. При этом возможно дальнейшее наращивание ядер, хотя особого смысла для обычных вычислений в этом нет. Предельное количество ядер на 5,6707 нм – 36 у Intel, 40 у AMD. Дальнейшее увеличение просто невозможно, т.к. требует мощных схем питания и системы водяного охлаждения. На этом микропроцессорное развитие вычислительной техники останавливается и начинается поиск новых архитектурных решений.

Технология кремний-на-сапфире (КНС) рассматривается как одна из перспективных для изготовления высокочастотных интегральных схем (ИС) с повышенной плотностью элементов. Структуры, изготовленные по этой

технологии, более долговечны, имеют высокую стойкость к радиации и потребляют меньше энергии по сравнению со структурами, изготовленными на массивном кремнии. По технологии КНС сейчас изготавливают только МОП- и КМОП-микросхемы.

С точки зрения использования в ИС преимущества КМОП-технологии на основе КНС таковы:

- уменьшение паразитных ёмкостей по причине использования изолирующей подложки, что обеспечивает увеличение быстродействия и уменьшение искажений;
- улучшение характеристик линейности, возможность использования меньших напряжений;
- высокий уровень развязки и уменьшение уровня помех благодаря изолирующим свойствам подложки;
- возможность ввода фотонных сигналов через подложку.

В случае ВЧ цифроаналоговых применений приборы на основе КНС обеспечивают большую величину максимальной частоты F_{\max} . Оценки показывают, что технология КМОП КНС с разрешением 0,5 мкм способна обеспечить достижение $F_{\max} = 50$ ГГц, что превышает аналогичный параметр для приборов, изготавливаемых по КМОП-технологии на объёмном кремнии с разрешением 0,13 мкм. Технология КНС остаётся весьма перспективной для массового выпуска надёжной и недорогой продукции благодаря простоте использования изолирующего материала подложки и ввода/вывода фотонных сигналов через подложку, но она не имеет такого высокого разрешения, как на массивном кремнии.

Сегодня архитектура микропроцессоров стала единообразной и система становится критичнее.

Процессор – это транзистор современности, который имеет 3 ограничения:

1. Энергетическая стена. Старая истина: энергия не стоит ничего, тран-

Сравнение машины фон Неймана с антимашинной

	Машина фон Неймана	Антимашина
	Языки потока команд	Языки потока данных
Базисные последовательные элементы	Следующая команда	Следующий элемент данных
	Безусловный переход (goto) по адресу команды	Безусловный переход (goto) по адресу данных
	Переход на произвольный шаг (jump) по адресу команды	Переход на произвольный шаг (jump) по адресу данных
	Цикл команды	Цикл данных
	Переход в исходное состояние (escape)	Переход в исходное состояние (escape)
	Команды условных переходов	Команды условных переходов
	Параллелизм внутренних циклов невозможен	Параллелизм внутренних циклов возможен

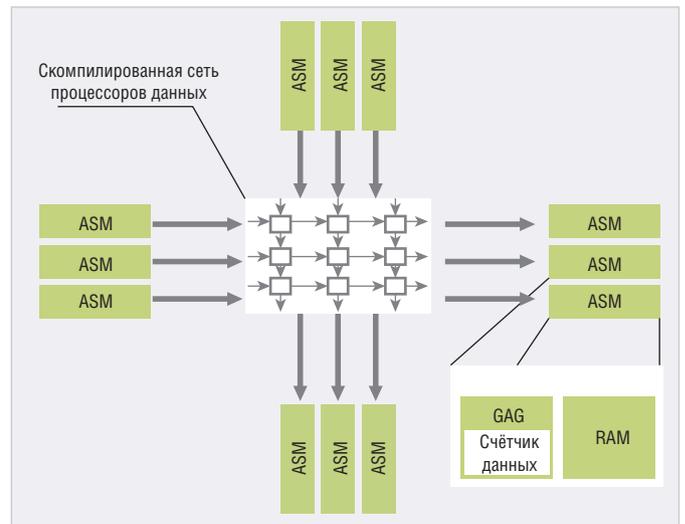


Рис. 1. Общая схема антимашинной

зисторы дороги. Новая истина: дорога́ энергия, транзисторы не стоят ничего.

2. Стена памяти. Старая истина: память работает быстро, а операции с плавающей запятой медленны. Новая истина: системную производительность сдерживает память, операции выполняются быстро.
3. Стена параллелизма на уровне команд. Старая истина: производительность можно повысить за счёт улучшения качества компиляторов и таких архитектурных усовершенствований, как конвейеры, внеочередное выполнение команд, сверхдлинное командное слово, явный параллелизм команд и др. Новая истина: естественный параллелизм; команды и длинные, и короткие, но выполняются они параллельно на разных ядрах.

Одноядерные процессоры бесперспективны, но и многоядерные процессоры, если всё сводится к размещению большего числа классических простых ядер на одной подложке, не решают всех проблем, т.к. их сложно программировать.

Существуют различные схемы построения вычислительных систем. Гарвардская схема основывается на том, что процессом вычислений управляют входные потоки данных, которые на входе системы попадают в подготовленную вычислительную инфраструктуру, обладающую естественным параллелизмом. Схема фон Неймана предполагает, что вычислительным процессом управляет поток команд, а данные, в основном статичные, выбираются из каких-то систем хранения или из памяти. Современные процессорные

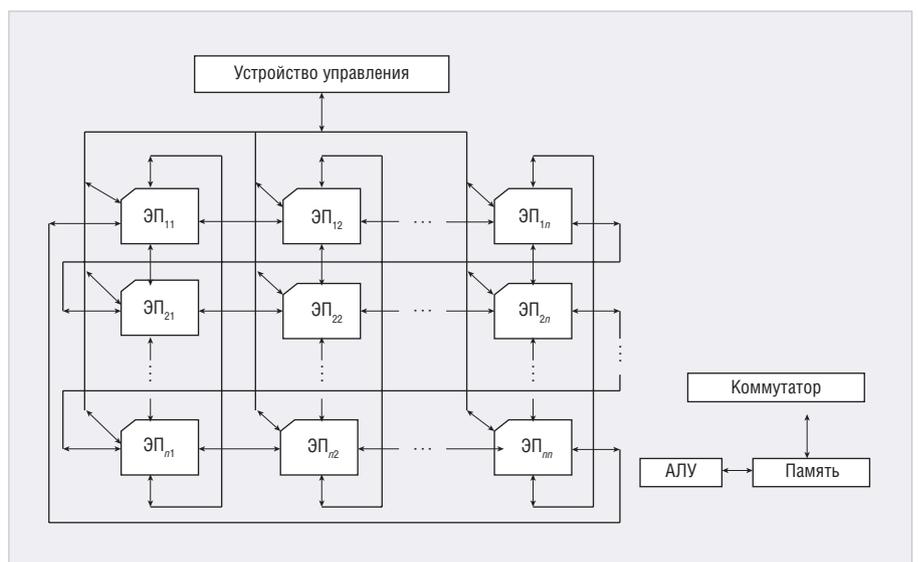


Рис. 2. Структура матричной системы

и вычислительные системы в основном используют принцип управления вычислениями фон Неймана.

Схема с архитектурой «антимашина» от машины фон Неймана отличается наличием одного или нескольких счётчиков, управляющих потоками данных. Она программируется с использованием потокового обеспечения, а роль центрального процессора в ней играют один или несколько процессоров данных. Центральной частью антимашинной может стать память с автоматической последовательностью.

Антисимметрия между машиной фон Неймана и антимашинной наблюдается во всём, за исключением того, что антимашина допускает параллелизм внутренних циклов (см. табл.), а это значит, что в ней решается проблема параллельной обработки данных.

Общая схема антимашинной приведена на рисунке 1.

В антимашине доступ к памяти обеспечивается не по адресу команды или фрагмента данных, записанному в соответствующий регистр, а посредством универсального генератора адресов. Его преимущество в том, что он позволяет передавать блоки и потоки данных. В то же время компиляция, посредством которой создаётся специализированная под определённую задачу система, заключается в объединении нужного количества настроенных процессоров данных в общий массив для выполнения определённых алгоритмов, как в реконфигурируемых матричных процессорах.

Матричные вычислительные системы (см. рис. 2) обладают более широкими архитектурными возможностями, чем конвейерные фон-неймановские системы, их каноническая архитектура относится к классу SIMD – одиночный

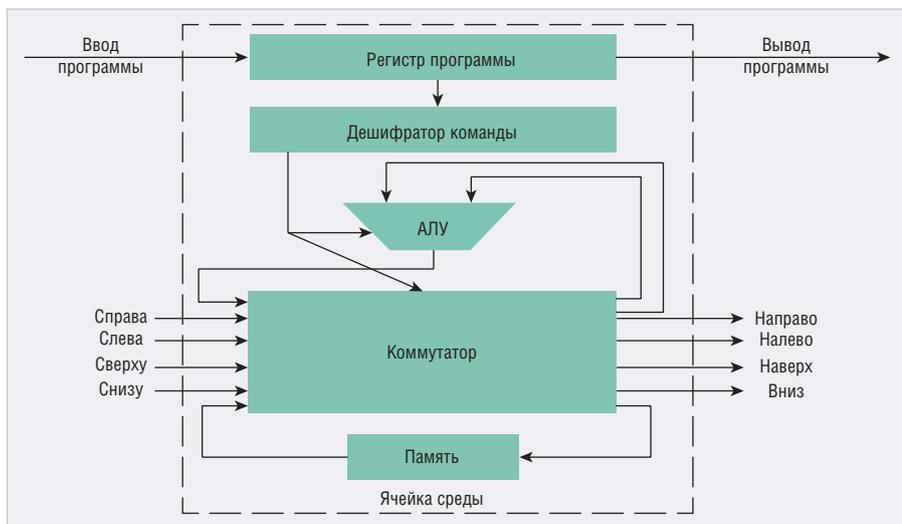


Рис. 3. Структура ячейки ПЭ матричной системы

поток команд множественных потоков данных.

В основе матричных вычислительных систем лежит матричный процессор, состоящий из массива процессорных элементов (ПЭ), которые осуществляют параллельную обработку множественных элементов данных. Единый поток команд, управляющий обработкой данных в ПЭ, которые передаются по шине широкополосной рассылки, генерируется управляющим процессором (УП). УП соединяет матричную SIMD-систему с внешним миром, используя для этого сетевой интерфейс (СИ).

Неотъемлемыми компонентами ПЭ в большинстве вычислительных систем являются:

- арифметико-логическое устройство (АЛУ);
- регистры данных;
- СИ, который может включать в свой состав регистры пересылки данных;

- номер процессора;
- регистр флага разрешения маскирования (F);
- локальная память;
- коммутатор.

Структура ячейки ПЭ матричной системы представлена на рисунке 3.

В реконфигурируемой матричной системе (см. рис. 4) вершины графа вычислительного процесса задаются между ПЭ матрицы, а указанные графом связи по управлению и данным реализуются с помощью соединительной сети.

Эффективность функционирования такой системы требует однородности этапов по времени выполнения и требуемым ресурсам.

ПЭ выполняет команды над операндами, поступающими из информационных каналов или из внутренней памяти.

Настройка ячеек может быть статической или динамической.

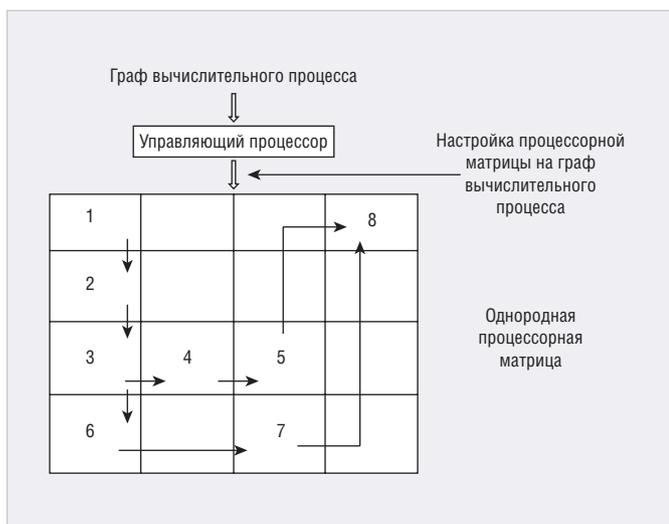


Рис. 4. Реконфигурируемая матричная система

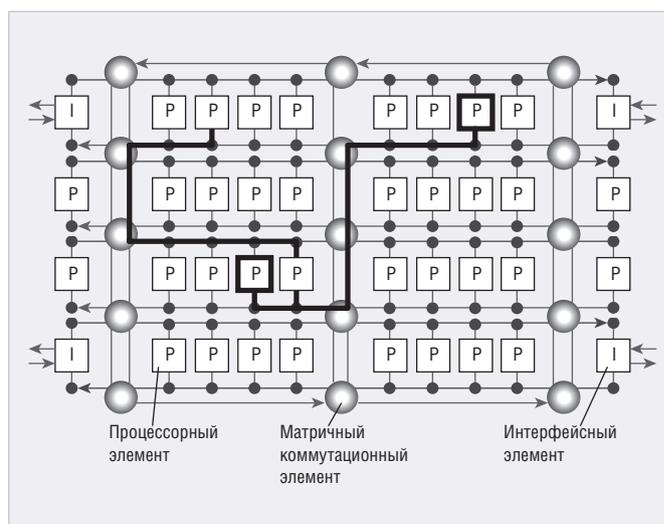


Рис. 5. Архитектура PicoArray

Конфигурирование является чисто локальной операцией. В ней участвуют два ПЭ: тот, который имеет новое содержимое, и тот, в который это содержимое записывается. Благодаря локальности операции конфигурирования могут производиться одновременно во множестве различных областей матрицы ПЭ.

Реконфигурируемая матричная система с архитектурой PicoArray (см. рис. 5) содержит 430 ПЭ, причём каждый работает на частоте 160 МГц и выполняет три команды за один цикл. Все ПЭ взаимодействуют посредством шин с суммарной пропускной способностью 5 Гбит/с. В состав архитектуры PicoArray входят ПЭ трёх типов: для быстрой реализации процедур в алгоритмах типа Витерби, для быстрого умножения с накоплением и для быстрой реализации контроллерных процедур. Кроме того, в состав архитектуры входят коммутационный и интерфейсный элементы. Программирование задач реализуется посредством индивидуальной настройки каждого ПЭ на выполняемую функцию. Архитектура PicoArray может увеличивать свою производительность путём каскадного наращивания. PicoArray изготовлен по 130 нм технологии на предприятии Taiwan Semiconductor Manufacturing Co.

Процессоры с многоядерной архитектурой (см. рис. 6) позволяют повышать производительность, они стали одним из основных направлений развития современной компьютерной техники. Многоядерный процессор – это центральный микропроцессор, содержащий 2 и более вычислительных ядер

на одном процессорном кристалле или в одном корпусе.

Преимущества многоядерных процессоров: возможность распределять работу программ, например основных задач приложений и фоновых задач операционной системы, по нескольким ядрам, увеличивать скорость работы программ и процессов, требующих интенсивных вычислений. Недостатки многоядерных процессоров: высокая себестоимость их производства по сравнению с одноядерными и возросшее энергопотребление, которое требует применения мощных схем питания и системы охлаждения. Количество программ, оптимизированных под многоядерность, ничтожно мало, поэтому они просто не могут задействовать вычислительную мощь дополнительных ядер. В настоящее время многоядерные процессоры используются крайне неэффективно. Проблема обеспечения высокопроизводительных вычислений перемещается теперь из области компьютерного оборудования в сферу параллельного программирования – и здесь нужны новые идеи и перспективные технологии для организации массового производства параллельных программ.

В многоядерных, многопроцессорных вычислительных системах для задач, требующих интенсивных, но нерегулярных межпроцессорных обменов и обращений к системе распределённой памяти, реальная производительность системы снижается. Одна из основных причин такого явления – несоответствие между «жёсткой» архитектурой многопроцессорной вычислительной системы и информационной структурой решаемых задач. Устранить этот недостаток можно путём построения вычислительной системы с реконфигурируемой структурой на основе ПЛИС (см. рис. 7).

Основные принципы создания реконфигурируемой вычислительной системы на основе полей ПЛИС заключаются в следующем:

1. Некоторое множество ПЛИС объединяется в единое вычислительное поле.
2. С помощью средств автоматического проектирования и структурного программирования в поле ПЛИС создаётся последовательная вычислительная структура, адекватная графу решаемой задачи.
3. Если аппаратные ограничения поля ПЛИС не позволяют отобразить весь

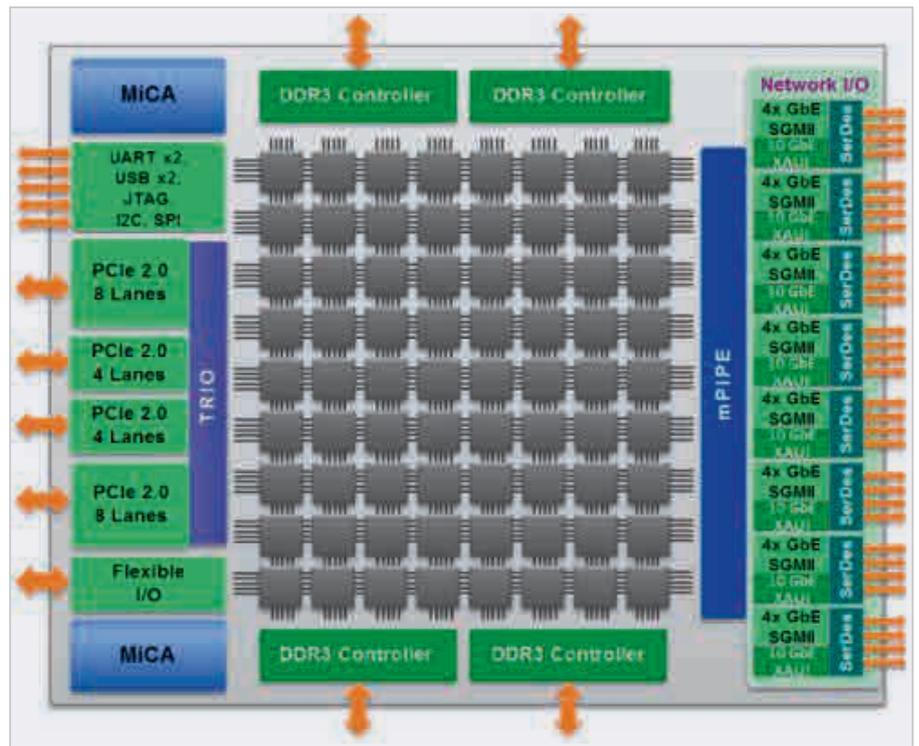


Рис. 6. Структурная схема многоядерной архитектуры процессоров Telera

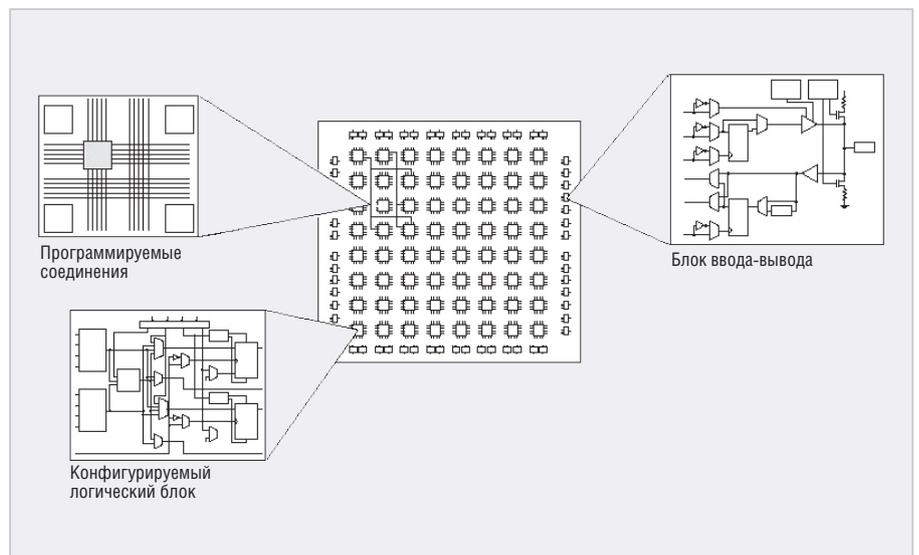


Рис. 7. Обобщённая структура ПЛИС – FPGA

граф решаемой задачи, то последний предварительно разрезается на непересекающиеся подграфы таким образом, чтобы каждый из них мог быть структурно реализован в имеющемся поле ПЛИС. С помощью ПЛИС строится вычислительное поле путём их объединения в некоторую структуру, например ортогональную решётку. В рамках этого поля каждый раз будет формироваться проблемно-ориентированная структура, которая наилучшим образом отвечает структуре решаемой задачи.

Квантовые компьютеры в корне отличаются от классических по прин-

ципу работы. Их архитектура построена вокруг кубитов, а не бит. Главное отличие квантового компьютера от обычного – это принцип суперпозиции, благодаря которому вместо выбора между 1 и 0 (бит) устройство способно вычислять задачи, одновременно учитывая и 1, и 0. Таким образом, во внимание одновременно принимается вся совокупность возможных решений.

Чтобы принцип суперпозиции действовал, квантовые компьютеры следует эксплуатировать при максимально низких температурах, например внутри системы D-Wave 2000Q температура в

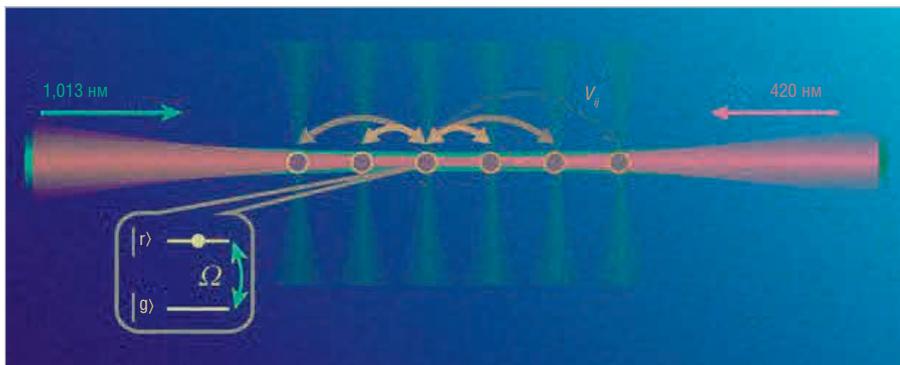


Рис. 8. Атомы-кубиты в оптической ловушке

180 раз ниже, чем в межзвёздном пространстве.

Квантовые вычисления позволяют решать специфические задачи в области оптимизации, создания новых лекарств и материалов, совершенствования энергосберегающих методов, а также осуществлять сверхбыстрый поиск в базе данных.

По проекту «Оптические квантовые симуляторы» первый российский квантовый компьютер должен быть создан к сентябрю 2021 года. Его разработкой занимается консорциум, в состав которого входят Фонд перспективных исследований, «Внешэкономбанк», ООО «ВЭБ Инновации», Московский государственный университет имени Ломоносова и АНО «Цифровая экономика».

Ожидается, что его мощность будет составлять не менее 50 кубит.

Одним из приоритетов отечественного квантового компьютера станет решение задач в сфере материаловедения. В частности, с его помощью можно будет более точно определять свойства новых материалов и фармакологических препаратов.

Российский компьютер будет создан на основе нейтральных атомов в оптических ловушках (см. рис. 8). Они представляют собой упорядоченные в пространстве образования, состоящие из отдельных атомов и удерживаемые в определённом положении с помощью лазерных пучков. Это наименее сложный подход, позволяющий выполнять квантовые вычисления при помощи линейных оптических устройств. 50 кубит – это как раз тот уровень, когда квантовый компьютер начинает превосходить по производительности современную классическую технику.

Стоимость создания такого компьютера составляет порядка 900 млн рублей.

Электронные многопроцессорные системы создаются в соответствии с

конфигурацией электрических связей между модулями – линейных или матричных структур, но не реализуются в электронных многопроцессорных реконфигурируемых трёхмерных структурах.

В процессорных системах с оптическими связями изменение межэлементных соединений осуществляется проще, чем в электронных системах, что позволяет разрабатывать трёхмерные многопроцессорные вычислительные системы, в том числе с архитектурой «антимашина».

Говоря о процессорных системах, нельзя не уделить внимание вопросу перспектив развития микроэлектроники в целом. Если проанализировать основные тенденции, сложившиеся в отрасли, можно попытаться выделить основные направления её эволюции:

- многопроцессорные системы; преобладание SnK над однопроцессорными системами;
- появление NoC – network-on-chip (сеть-на-кристалле) – высокопроизводительных устройств, представляющих собой группы процессоров, обмен между которыми осуществляется по пакетным протоколам; используются связи «точка-точка» и асинхронный способ передачи;
- блочное программирование (component-based software): число ядер в системе невелико, а блоки ПО разрабатываются индивидуально для каждой вычислительной ячейки специальной группой разработчиков компонентов и затем собираются вместе в одну многоядерную систему;
- единый код для многоядерной системы; разрабатываются новый принцип – SPMD (одна программа – много данных) и параллельные алгоритмы;
- отказ от FPGA в связи с тем, что небольшие многоядерные системы потребляют меньше энергии и имеют

более широкий арсенал возможностей для выполнения сложных алгоритмов, чем АЛУ и таблицы преобразования в FPGA;

- исчезновение ЦПУ: в многоядерных системах эффективность взаимодействия между ними поддерживается аппаратными ОС.

Кроме того, можно предположить, что системы станут многопроцессорными, а доступ ко всем ресурсам МП будет осуществляться за один такт. Будет задействовано третье измерение, т.е. схемы станут объёмными. Уже сейчас ведутся исследования в этом направлении, и появляются многослойные кристаллы (SiP – system in package).

Здесь же стоит сказать, что последний физический предел размеров транзисторов, которого теоретически можно достигнуть, – 1 нм. Дальше идёт уже атомарный уровень, и просто невозможно создать единицу машинной логики из нескольких атомов.

Новейшие современные техпроцессы используют технологию производства транзисторов FinFET, которую также называют 3D-структурой, поскольку расположение транзисторов на схеме производится не планарно, а в несколько слоёв, которые связаны друг с другом в вертикальном направлении. Исходя из этого структура должна стать действительно трёхмерной, представляя собой мультиполигональный объект, в котором каждый элемент граничит не с двумя (планарное размещение транзисторов) или с четырьмя-шестью, а с десятком других элементов. Поэтому процессоры также должны стать и гетероструктурными. Полная гетерогенность предполагает, что элементы всех сопроцессоров будут вписаны внутрь вычислительных ядер таким образом, что отделить их на плате будет просто невозможно.

Во второй части статьи пойдёт речь о построении устройств на базе 3D МФЭ ПМ с сетевой архитектурой и интеграцией оптических компонентов непосредственно на электронной плате.

ЛИТЕРАТУРА

1. www.osp.ru
2. www.present5.com
3. www.dammlab.com
4. www.cyberleninka.ru
5. www.studfiles.net
6. www.echo.msk.ru
7. www.stevsky.ru



ИЗГОТОВЛЕНИЕ КЕРАМИЧЕСКИХ ПЛАТ И ПОДЛОЖЕК

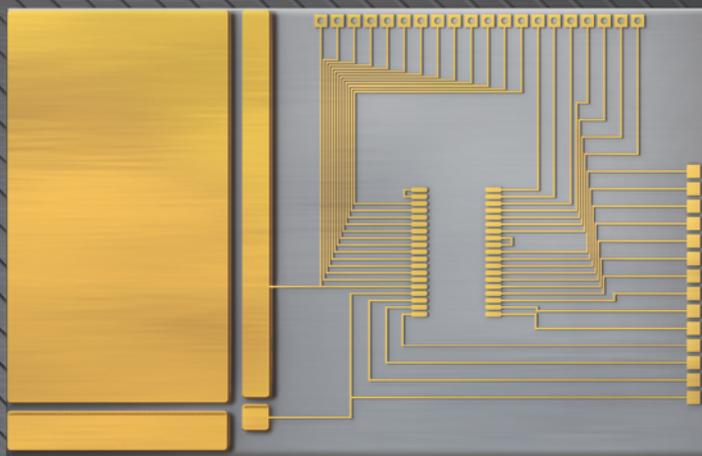
ТЕСТПРИБОР

ИСПОЛЬЗУЕМЫЕ МАТЕРИАЛЫ

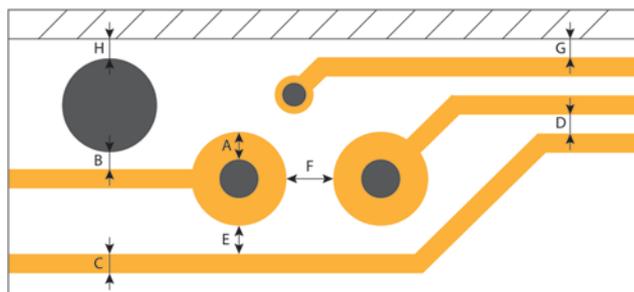
Al₂O₃ 96%

Al₂O₃ 99,6%

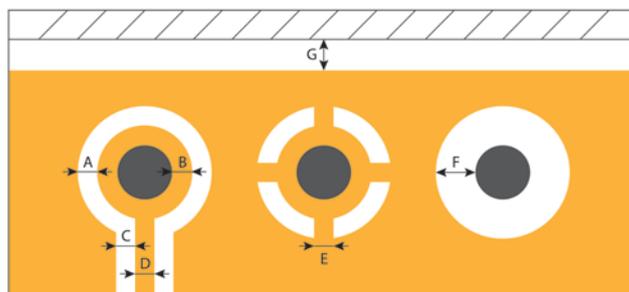
AlN



ТОПОЛОГИЧЕСКИЕ НОРМЫ СИГНАЛЬНЫХ СЛОЕВ



ТОПОЛОГИЧЕСКИЕ НОРМЫ СИЛОВЫХ СЛОЕВ



Наименование параметра	Керамические подложки с металлизацией на основе толстопленочной технологии	Керамические подложки с металлизацией на основе тонкопленочной технологии
Материал проводников и металлизации	W/Ni-Au или Mo/Ni-Au (Ni 5 мкм max, Au 0,5 мкм max)	TiW/Au; TaN/TiW/Au; TiW/Ni/Au; TaN/TiW/Ni/Au; TaN/NiW/Au/Cu/Ni/Au
Поверхностное сопротивление проводников	10,0 МОм/□	—
Сопротивление переходных отверстий размером (Ø0,2×0,25) мм	6,0 МОм	—
A	0,20	0,20/020
B	0,60	0,25/025
C	0,20	0,15/0,20
D	0,20/0,25	0,15/0,20
E	0,20/0,30	0,15/0,38
F	0,25/0,65	0,15/0,30
G	0,60/0,60	0,25/0,50
H	0,60/-	0,15/-