



# Главные технотренды 2018

В IT-разработках дальнейшее развитие традиционных компьютерных технологий играет всего лишь вспомогательную роль. **Искусственный интеллект и Интернет вещей** задают тон на ближайшее будущее.

**Т**от факт, что нас окружают компьютеры, а мы постоянно находимся на связи со смартфонами в руках, сегодня воспринимается как нечто само собой разумеющееся. Начало такого симбиоза человека, Сети и компьютера знаменует известная рекламная вставка AOL с Борисом Беккером и его «Ich bin drin» («Я внутри», «я в деле»). Двадцать лет спустя, в 2019-м, смысл сочетания «быть внутри» воспринимается по-другому. Может быть, искусственный интеллект (ИИ) и Интернет вещей — это два понятия, которые означают для многих что-то роковое и далекое, как еще пару лет назад облачные технологии, но за ними стоят перемены, которые крепко охватывают повседневную жизнь.

Интернет вещей — это мир сантиметровых устройств, объединенных в сеть, состоящих из чипов, датчиков и беспроводной связи. Они вшиты в одежду, встроены в автомобили, уличные фонари, вмонтированы в места стоянки и собирают потоки данных, которые, в свою очередь, в будущем станут

обрабатываться не простыми алгоритмами, а самообучающимися системами, с каждым притоком информации становящимися все лучше и откликающимися все более «интеллектуально».

Но этот чудесный новый мир наступит только при условии, что будет соблюдена основная константа: чтобы повысить эффективность вычислений компьютеров, самые маленькие вычислительные блоки — транзисторы — должны стать еще более компактными. Если еще 15 лет назад размер транзисторов составлял 130 нм, что сопоставимо с размером вируса гриппа, то сегодня ученые ломают голову в поисках способа уменьшить транзисторы в будущем до диаметра атома (0,3 нм).

## Миниатюризация для прогресса

Транзисторы — это миниатюрные электронные переключатели, представляющие значение 0 или 1 бинарного кода. Чем они меньше, тем быстрее компьютер производит вычисления, если

говорить о производительности, и тем энергоэффективнее он работает, если говорить об оптимизации энергопотребления. Знаменитый закон Мура, озвученный полвека назад, заключается в следующем: количество транзисторов, размещаемых на одной и той же площади интегральной схемы, каждые полтора года увеличивается в два раза. Такое удваивание означает повышение эффективности до 40%. Большого производители микропроцессоров на сегодняшний день достичь не в состоянии, поскольку используемые технологии производства микросхем дошли до пределов своих возможностей.

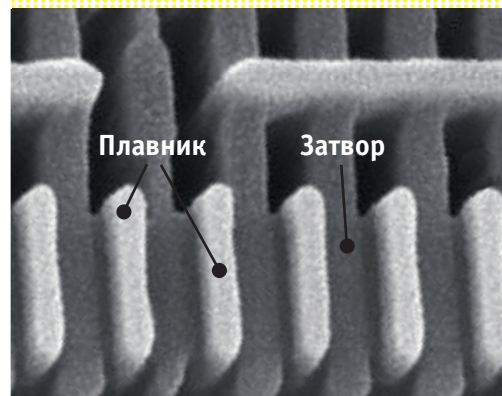
Уже сегодня развитие нового поколения требует не одного миллиарда долларов. Из двух десятков производителей интегральных микросхем закону Мура сейчас отвечают всего четыре: Intel, Samsung, TSMC (Taiwan Semiconductor Manufacturing Company) и GlobalFoundries, поставщик чипов AMD. Полупроводниковые изделия TSMC используют многие производители смартфонов. Кроме того, обозначения размеров поколения транзисторов больше не соответствует фактическим размерам. В настоящее время Samsung позиционирует свой Galaxy S8 как смартфон с чипом с десятинанометровыми транзисторами, но по сути никаких транзисторов размером в 10 нм там нет, а то, что Samsung называет «10 нанометров», для Intel и TSMC означает немного другое (см. справа внизу).

Этот момент также связан с новой структурой FinFET, которую Intel впервые использовала в 2012 году. Говоря простыми словами, эти транзисторы состоят из двух конструктивных элементов: затворов и каналов в виде плавников или ребер, перекрещивающихся между собой и образующих сетку (см. справа). Затвор пропускает или блокирует ток, протекающий по плавнику, обеспечивая таким образом переключение значения с 0 на 1 и наоборот. Сравнить размеры элементов структуры разных производителей микросхем можно по двум параметрам: Contacted Poly Pitch — шаг затворов, и Minimum Metal Pitch — расстояние между контактами, которые находятся непосредственно над транзисторами и обеспечивают подачу тока.

### Без пяти двенадцать для закона Мура

Преимущества транзисторов FinFET заключаются в плавниках. Затвор обхватывает канал-плавник с трех сторон, в результате чего повышается эффективность управления включением и отключением тока. Но чем меньше структуры, тем сложнее контролировать прохождение тока. В результате возникает ток утечки, повышается теплоотдача и энергопотребление. В производстве полупроводниковых приборов считается, что предел технологии FinFET — семь нанометров. Компания TSMC намерена к 2018 году перейти на семь нанометров для iPhone через одно поколение, и это будет последним поколением микропроцессоров, выпущенных по старой технологии.

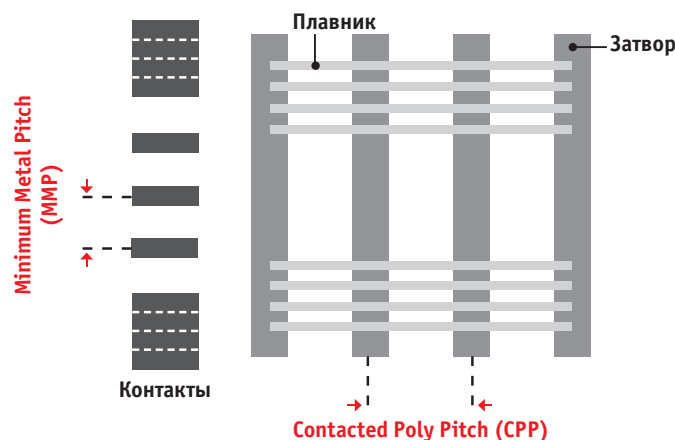
Срок действия закона Мура должны продлить две новинки. В первую очередь — использование нового метода экспонирования. Производители интегральных микросхем по-прежнему применяют иммерсионную литографию для формирования кремниевых структур при помощи лазера, излучающего свет с длиной волны 193 нм. Чтобы вырастить структуры, размер которых меньше длины волны, материал многократно экспонируется через матрицы. В будущем будет использоваться фотолитография в глубоком ультрафиолете (EUV) — лазер будет излучать волны длиной всего 13,5 нм. Применение волн такой длины требует перемещения полупроводниковых пластин через ряд установок, причем в вакууме, что представляет собой технические сложности. Нидерландская компания ASML уже производит тестовые EUV-литографы, из которых 14 систем уже используются →



Транзисторы FinFET на рынке с 2012 года. Они состоят из двух конструктивных элементов: канал-плавников и затворов. Напряжение на затворе переключает транзистор

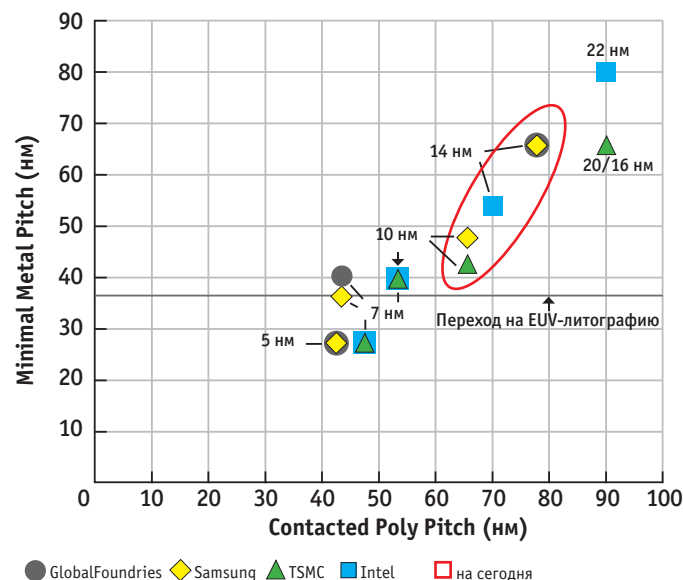
### Два параметра определяют производительность

Чем меньше транзисторы FinFET, тем выше производительность чипа. Определяющими являются два параметра: шаг затворов (CPP) и расстояние между контактами, которые подают ток (MMP)



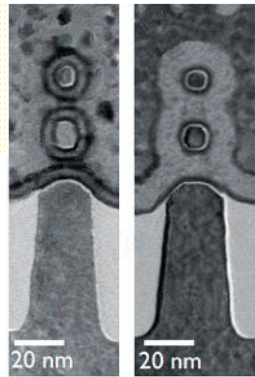
### Размеры транзисторов: маркетинг и реальность

Обозначения размеров производителей микросхем не соответствуют действительности. Также неясно, удастся ли им с 2018 года наладить серийное производство с помощью EUV-литографии



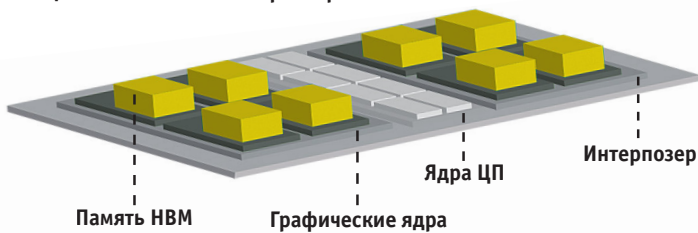


Для поколения чипов размером 7 нм будут использоваться нанопроводные транзисторы, в которых затвор полностью обхватывает плавник. Справа — прототип, разработанный в исследовательском центре IMEC



## Высокопроизводительный ПК размером с ладонь

Компания AMD представила модель процессора EHP. Все компоненты размещаются на кремниевой подложке и соединяются при помощи тысяч линий в интерпозере



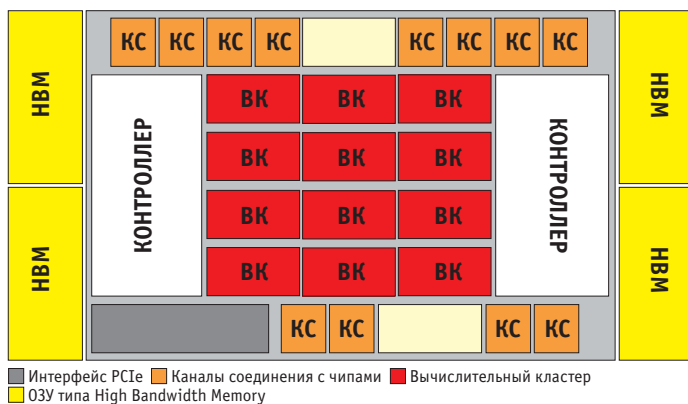
## Узлы нейронной сети

Узлы (нейроны) нейронной сети расположены слоями, каждый узел выполняет простые вычисления. В итоге результаты разных рядов нейронов сверяются и сеть выдает вероятностный результат. На рисунке показана нейронная сеть, распознающая изображения.



## Lake Crest: чип Intel для вычислений ИИ

Микросхема Lake Crest содержит 12 кластеров для нейронов, которые будут хранить данные в стеках памяти HBM (пропускная способность 1 Тбит/с). Кластеры соединяются с другими чипами через каналы с пропускной способностью до 100 Гбит/с



по всему миру. По всей видимости, к 2019 году ASML созреет для серийного производства EUV-установок.

К 2020-му появится поколение пятинанометровых транзисторов с новой структурой. В нанопроводных транзисторах затвор полностью обхватывает плавник, что позволяет контролировать ток лучше, чем FinFET. Летом 2016 года Межуниверситетский центр микроэлектроники, один из крупнейших научно-исследовательских центров в области микро- и наноэлектроники, расположенный в бельгийском городе Левене и насчитывающий около двух тысяч сотрудников, представил прототип такого устройства в восемь нанометров в поперечнике. Специалисты считают, что нанопровода можно использовать вплоть до трехнанометрового поколения, а потом придется отказаться от кремния в качестве материала для транзисторов.

## Высокоскоростные вычислительные ядра для компьютеров будущего

Уменьшение размеров транзисторов позволит разработать новые форм-факторы домашних ПК в будущем. Как они могут выглядеть, недавно показала AMD, пролив свет на дизайн своего нового процессора с гетерогенной архитектурой эксафлопсного класса (EHP) с высокой вычислительной мощностью. Микросхема EHP включает все основные компоненты ПК: центральный и графический процессоры, память. Многие центральные процессоры уже сегодня наряду с вычислительными ядрами ЦПУ содержат графическое ядро — его хватает на декодирование видео, но для игр оно слабовато. В рамках EHP же подразумевается выделение графическому ядру достаточного количества места на кристалле для осуществления высокопроизводительных вычислений (см. схему справа).

Компоненты будут связаны с помощью интерпозера — кремниевой подложки с протравленными в ней тысячами линий, благодаря чему обмен данными происходит в несколько раз быстрее. Рядом с графическими ядрами располагается память типа High Bandwidth Memory. Преимущество HBM перед модулями DDR RAM заключается в расположении ячеек стеками, что также ускоряет передачу данных (за счет широкой шины) и плотность хранения данных. Из-за высокого нагрева разместить память непосредственно в процессоре сегодня не представляется возможным — микросхема просто перегреется. Для этого концепта требуются транзисторы еще меньшего размера с низким энергопотреблением и высокой производительностью.

## Умное аппаратное обеспечение

Концепт EHP можно адаптировать под ПК, но он не создавался специально для них. Толчки к техническому прорыву исходят отнюдь не от серверов, ПК или смартфонов. Каким будет компьютер будущего, определяют два других фактора: искусственный интеллект (ИИ) и Интернет вещей, то есть объединение датчиков и устройств в единую сеть.

## Самообучающиеся системы

Термин «искусственный интеллект» — неудачный, современный ИИ мало общего имеет с тем, что обычно подразумевается под словом «интеллект». ИИ, который подобно человеку анализирует непредсказуемые ситуации, будет создан еще нескоро. Сейчас в рамках концепции искусственного интеллекта вычисления производит нейронная сеть, которая решает специфические задачи, а до этого проходит обучение. Некоторые системы продолжают обучаться и совершенствоваться и после завершения подготовки, используя обратную связь.

Нейронные сети состоят из множества узлов, которые производят параллельные вычисления (см. блок слева). Разработчики ИИ используют видеокарты типа GeForce GTX 1080 — она содержит 2560 арифметико-логических блоков (ALU). Но это не более чем временное решение: большая часть ALU нередко не задействуется, поскольку нейронной сети не требуется параллелизация в таких масштабах. Кроме того, нейронной сети нужно постоянно передавать в память и получать из нее большое количество данных — в этом отношении HBM была бы побыстрее, чем GDDR5X на GTX 1080. Во второй половине года на рынке ожидается появление ускорителя вычисления под названием Lake Crest, предназначенного для построения ИИ, который анонсировала Intel. Микросхема Lake Crest будет нести 12 вычислительных узлов, которые Intel называет кластерами, а каждый кластер дробится на несколько ALU, связанных с HBM.

### Искусственный интеллект в массы

Оптимизированное под ИИ аппаратное обеспечение уже массово выпускается: например, некоторые смартфоны работают на новом чипе Qualcomm Snapdragon 835, содержащем, помимо ЦП и графического ядра, целый ряд специальных микросхем, одну из которых производитель предназначил для поддержки технологии машинного обучения Google TensorFlow. При помощи TensorFlow скорость выполнения операций по сравнению с ЦПУ повышается в 25 раз, а по сравнению с графическим ядром — в восемь раз. Google открыла библиотеку TensorFlow для разработчиков примерно год назад, а с начала 2017 года она поддерживается Windows.

Использовать оптимизированное под вычисления ИИ «железо» можно и в Интернете вещей — это уже почувствовал американский стартап CubeWorks. В рамках проекта Micro Mote разрабатываются компьютеры объемом в кубический миллиметр, которые можно дооснастить микрокамерой или датчиками движения или температуры. Недавно разработчики представили крошечный компьютер, работающий на процессоре с поддержкой ИИ, с потреблением всего 300 мкВт. Устройство может, например, определить, чем вызван шум за окном — приближением злоумышленника или ветром.

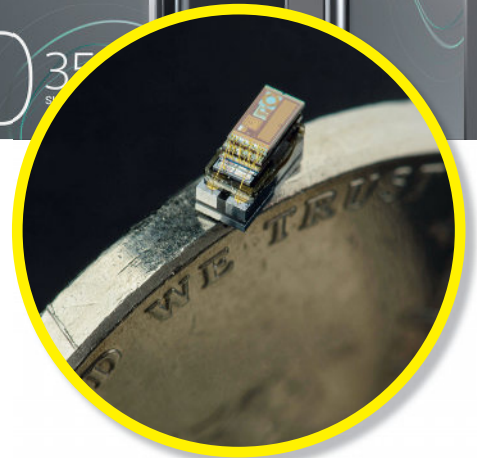
Присутствие искусственного интеллекта в нашей повседневной жизни подкрепляет не только оптимизация аппаратного обеспечения. Некоторые вычисления можно продолжать оптимизировать только при помощи ИИ. Так, в ноябре Google рассказала о новом алгоритме, предназначенном для увеличения разрешения изображений. Технология RAISR («быстрый и точный образ в сверхразрешении») использует нейронную сеть и умеет изменять размеры изображения лучше любого из разработанных человеком фильтров. Разработки по использованию нейронных сетей в кодировании видео уже в ходу — в Google полным ходом идет разработка собственного видеокodeка VP10. Если конвертировать фильмы с использованием нейронных сетей, теоретически повысится качество на выходе, поскольку при сжатии видео главная задача заключается в идентификации одинаковых значений цвета и яркости в разных кадрах. Это умеют делать и сложные алгоритмы, но ИИ мог бы дополнительно оптимизировать поиск одинаковых значений под условия определенного фильма.

### Человеческие слабости компьютера в социальной дилемме

Могут ли компьютеры врать? Во время обучения генеративных состязательных сетей (GAN) они не то что могут, они должны: GAN — это своего рода война двух нейронных сетей, где одна →

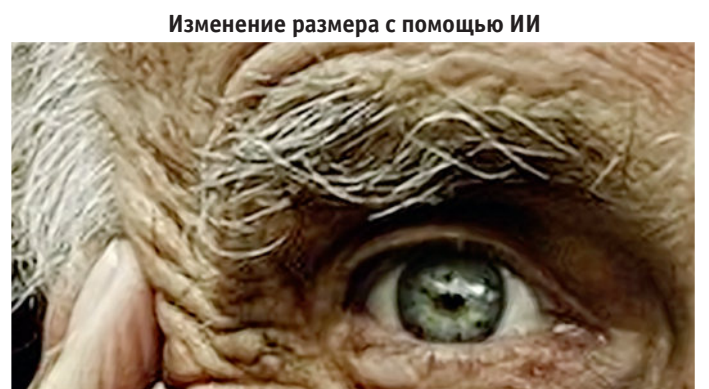
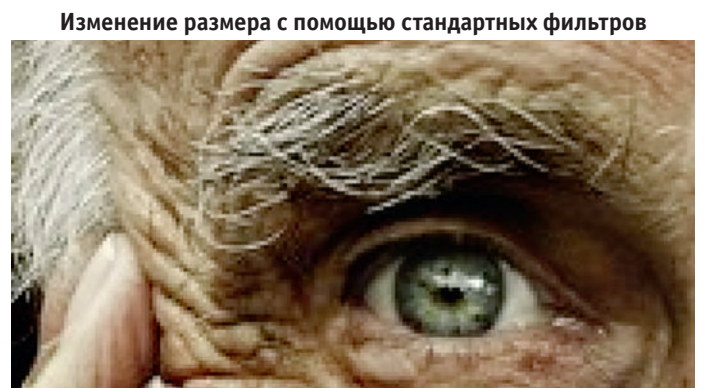


Micro Mote — компьютер объемом в 1 кубический миллиметр с собственным ИИ, который уместится на ребре монеты



### RAISR: увеличение без потерь

Технология Google RAISR, использующая нейронные сети, настолько хорошо увеличивает фотографии, что заменяет любой фильтр, меняющий размер изображений



Изменение размера с помощью ИИ

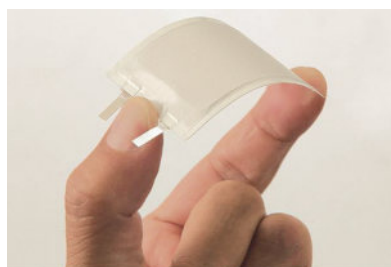
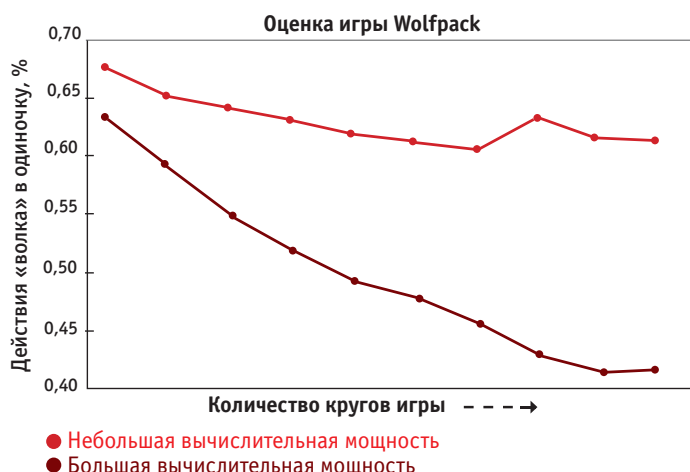




В игре Wolfpack двум ИИ (красные точки) предлагается поймать добычу (синяя точка). «Волки» могут сотрудничать или охотиться в одиночку, но во втором случае вероятность успеха выше

### Тестирование двух ИИ: сотрудничать сложнее

Анализ игры показывает, что ИИ быстрее достигают совместных решений, если им предоставляется много вычислительных ресурсов



### Компактный карманный аккумулятор

Компания Panasonic представила гибкий аккумулятор, который можно носить на теле

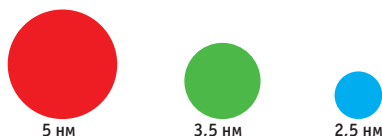
### Телевизионный дисплей на квантовых точках

Samsung продает телевизоры на квантовых точках под названием QLED



### Квантовые точки: важен каждый атом

Квантовые точки — это нанокристаллы, которые излучают свет определенной длины. Цвет зависит от размера: синие квантовые точки состоят всего из нескольких атомов



сеть пытается обмануть другую, предоставляя ей изображения, в том числе сгенерированные, а задача второй — распознать подделку. Обе сети получают ответ, удался ли обман. Чем дольше они «состязаются», тем лучше удается одной генерировать, а второй распознавать поддельные изображения. В итоге обе сети с помощью подложных сведений обучаются быстрее, чем с помощью только настоящих фотографий.

Еще один способ машинного обучения — reinforcement learning (обучение с подкреплением) — сходится к взаимодействию системы с некоторой средой. Задача ИИ — самостоятельно обучиться правилам игры в своей среде. Компьютерные игры обеспечивают хорошие условия, которые довольно просты, но при этом дают ИИ отклик об успешности принятых им решений. Подразделение Google DeepMind, занимающееся искусственным интеллектом, раскрутило концепт еще дальше, чтобы посмотреть, как два ИИ будут вести себя в игровой среде, которая будет основываться на социальной дилемме: приведет ли сотрудничество или конфронтация к успеху (см. слева)? Результат оказался следующим: сотрудничество — более сложный сценарий, предполагающий более длительный процесс обучения. Чем больше ресурсов предоставлялось в распоряжение, тем чаще оба ИИ выбирали эту стратегию.

### Аккумуляторы и квантовые точки

Для Интернета вещей компактные аккумуляторы — это необходимость. Какими могут быть миниатюрные элементы питания, можно было увидеть на выставке потребительской электроники CES (см. слева). Гибкий аккумулятор от Panasonic толщиной всего 0,45 мм идеально подходит для носимых устройств. Чтобы обеспечить такую гибкость, компания перестроила внутренние компоненты литий-ионного аккумулятора, расположив противоположно заряженные электроды в виде тонких слоев один над другим. Заряжается такой элемент «по воздуху». Panasonic планирует запустить массовое производство своей «крохи» в 2018 году. Еще одна важная область исследований — разработка твердых электролитов. В современных литий-ионных аккумуляторах в качестве электролита используется органическая жидкость, которая при сильном нагреве начинает испаряться, что в итоге разрывает и сжигает аккумулятор. Твердый же электролит в значительной мере мог бы предотвратить короткое замыкание и сгорание.

### Телевизионное изображение в чистых цветах

Индустрия развлечений технически тоже ориентирована на уменьшение структур. Samsung выводит на рынок новую линейку телевизоров с дисплеями на квантовых точках под наименованием QLED. Квантовые точки — это полупроводниковые кристаллы из металлического сплава размером всего в несколько нанометров (см. слева), которые излучают энергию в форме света определенной длины волны. Технология квантовых точек подразумевает преобразование синего свечения светодиодов в красный и зеленый цвета, благодаря чему практически отпадает необходимость в использовании фильтра. QLED — это всего лишь промежуточный шаг на пути к настоящему дисплею на квантовых точках. Конечный этап предполагает наличие квантовых точек не только для красного и зеленого, но и для синего света, однако тут производители вынуждены искать решение по увеличению долговечности материала. В итоге квантовые точки должны сами излучать свет под воздействием тока, и тогда можно будет совсем отказаться от фильтра. По прогнозам Samsung, технический прорыв возможен не ранее 2019 года.

ФОТО: Panasonic; Samsung