



ЗАКОНУ МУРА 50 ЛЕТ: МАСШТАБИРОВАНИЕ ЭЛЕМЕНТОВ ИС

FIFTIETH ANNIVERSARY OF MOORE'S LAW: SCALING OF IC ELEMENTS

В.Вернер, Е.Кузнецов, А.Сауров / vdv@tcen.ru
V.Verner, E.Kuznetsov, A.Saurov

В первой части серии публикаций, посвященных 50-летию закона Мура, рассмотрена его связь с масштабированием элементов ИС.

The first part of the series of publications devoted to the 50th anniversary of Moore's Law discusses its correlation with the scaling of IC elements.

На заре истории интегральной электроники на базе кремния Г.Муру, в то время руководившему компанией Fairchild Semiconductor, удалось сформулировать тенденции ее дальнейшего развития. Нужно было доказать, что интегральные схемы более экономически привлекательны, чем решения на дискретных элементах, поэтому была исследована зависимость цены ИС от условий ее производства, то есть от технологии. Наблюдения показывали, что минимальная цена

зависит от уровня технологии, который меняется со временем (рис.1).

Уровень технологии был оценен числом транзисторов в микросхеме, получившим в дальнейшем название "степень интеграции" (рис.2).

СУЩНОСТЬ И ЗНАЧЕНИЕ ЗАКОНА МУРА

Мур первоначально обнаружил, что число транзисторов при выполнении условия минимизации стоимости ежегодно удваивается. Последующий анализ развития ИС, выполненный уже в 1975 году, подсказал Муру, что удвоение степени интеграции происходит через два

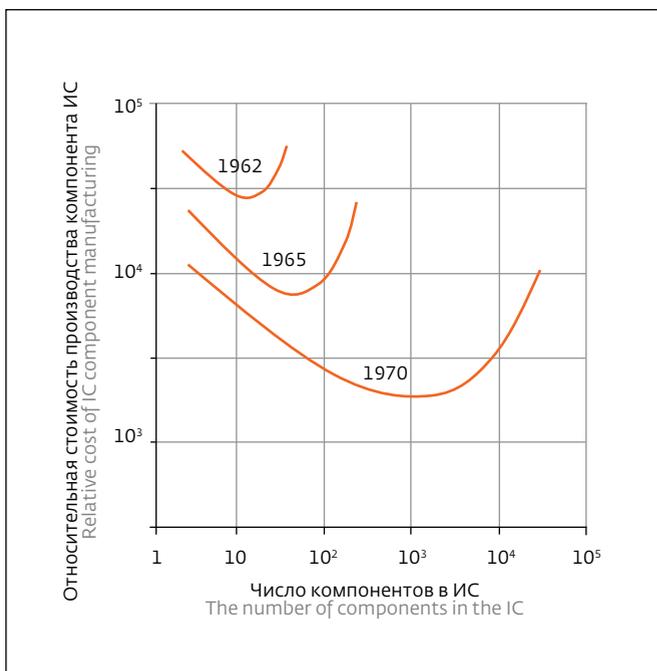


Рис.1. Минимизация стоимости компонента ИС в зависимости от уровня производства [1]

Fig.1. Minimizing the cost of the IC component depending on the level of production [1]

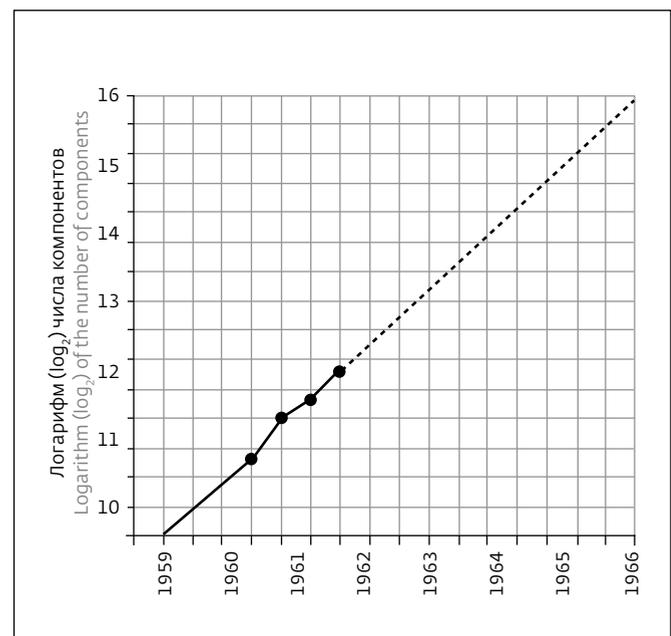


Рис.2. Рост степени интеграции ИС [1]

Fig.2. The growth of IC integration [1]

года. Заслуга Г.Мура заключается в том, что он сформулировал требования к темпу развития разработки и производства ИС. Именно задание темпа развития ИС было отмечено в [2] как главное достижение закона Мура, которое позволило сравнить его появление с таким эпохальным явлением как Интернет.

В [2] закон Мура был назван "боевым барабаном", который определял темп развития. Или ты следуешь этому темпу, или проигрываешь. Менее агрессивным, но имеющим тот же смысл, было название закона Мура "метрономом Кремниевой долины" [3].

Закон Мура определял развитие микроэлектроники по экспоненте, которое, в частности, выразилось в двухкратном изменении цены ИС, как и степени интеграции, каждые два года (рис.3) [4, 5].

Но, как оказалось, двухлетний темп изменений имеет ограничения - последующий анализ показал, что он замедляется (рис.4). Несмотря на это можно констатировать, что пока имеется возможность сформулировать закономерность изменения темпов развития микроэлектроники, закон Мура продолжает действовать.

Различия в оценке закона Мура связаны с его двойственностью. Изначально он был сформулирован как экономический закон экспоненциального развития нового рынка ИС, но опорой этого стал аналогичный характер развития технологии ИС. Требования изменений в технологии в определенном темпе стали главными воспринимае-

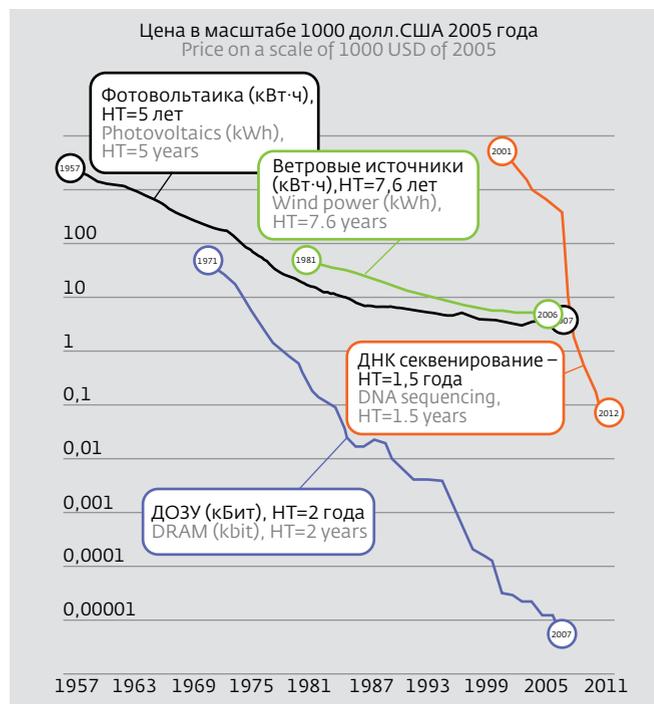


Рис.3. Время, необходимое для снижения в два раза цены характеристики изделия (Halving Time - HT) [4, 5]

Fig.3. The time required to reduce twice the price of the product specifications (Halving Time - HT) [4, 5]

мыми признаками закона Мура. Экономическая составляющая, оставаясь важной, отошла в оценке закона на второй план.

Действие закона Мура должно вызывать уменьшение цены, но чего именно: функционального

In the early days of integrated electronics based on silicon, G.Moore, who headed Fairchild Semiconductor at the time, formulated the further development trend of electronics. It was necessary to prove that integrated circuits were more profitable than solutions on discrete elements. Therefore, research was conducted around the correlation between the price of ICs and conditions of their manufacture, i.e. technology. According to observations, the least price depended on the technological level, which was changing over time (fig.1).

The technological level was determined by the number of transistors in a circuit, which later was known as "the degree of integration" (fig.2).

MEANING AND FOUNDATION OF MOORE'S LAW

In the beginning, Moore concluded that the number of transistors in the condition of cost minimization doubled each year. The subsequent analysis of IC development, conducted in 1975, convinced Moore that the degree of integration doubled every two years. G.Moore's merit was that he formulated the target pace

in development of IC design and manufacture. Setting the IC development pace was described in [2] as the main achievement of Moore's Law, and it was compared to such epochal phenomenon as Internet.

In [2], Moore's Law was called a "battle drum" that set up the pace of development. Either you keep up with this pace or you lose. A less aggressive title of Moore's Law, though having the same meaning, was "the Metronome of the Silicon Valley" [3].

Moore's Law determined the development pace of

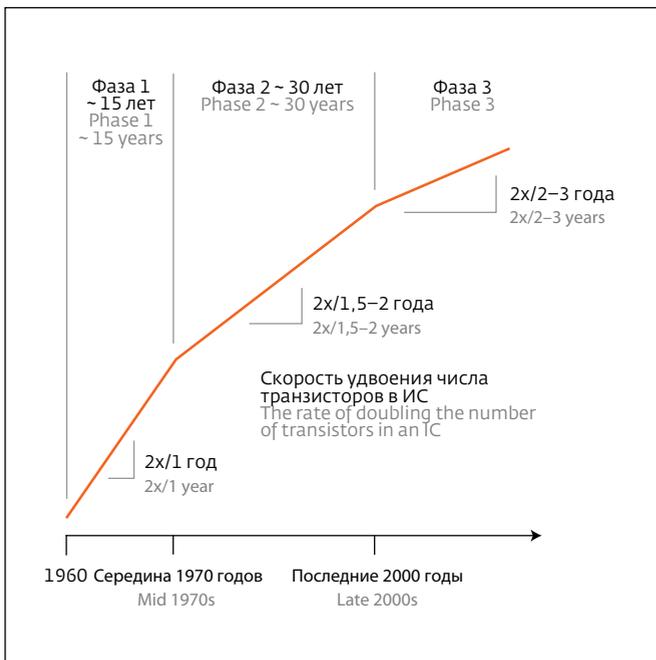


Рис.4. Три фазы закона Мура (IC Insights)

Fig.4. Three phases of Moore's law (IC Insights)

элемента – транзистора – или более объемной функции? Начнем с первого.

Увеличение степени интеграции, помимо снижения стоимости элементов, сопровождается улучшением рабочих характеристик ИС. Так, производительность микропроцессоров при таком темпе интеграции возрастает в два раза каждые 18 месяцев, что является результатом роста числа

транзисторов и снижения задержки сигнала на элементах ИС. Как следствия этого закона можно сформулировать аналогичные улучшения и для других характеристик ИС, таких как потребляемая мощность, стоимость, быстродействие и т.д. Поэтому очень часто под законом Мура подразумевают и все следствия его основной формулировки. В этой форме закон Мура перешел из области экономических в сферу технических прогнозов. И именно в этой форме он получил известность и стал стимулом развития микроэлектроники.

В практической деятельности закон Мура предполагает планирование и организацию научно-технической деятельности, определение ее направлений и постановку проблем так, чтобы продолжать определенные тенденции в развитии микроэлектроники и, в более широком смысле, всего научно-технического прогресса. Таким образом, закон Мура задал стремительное развитие микроэлектроники, что привело к формированию нового технологического уклада – информационного общества, и стало по существу основой экономического прогресса за последние 50 лет. Это развитие повлияло и продолжает оказывать неоспоримое влияние на все отрасли производства, изменило социальную и повседневную жизнь. Закон Мура в микроэлектронике привел к ускорению научно-технической революции и зарождению новых междисциплинарных направлений, таких как нанотехнологии, искусственный интеллект и когнитивные технологии.

Повышение степени интеграции ИС естественным образом связано с необходимостью уменьше-

microelectronics exponentially, which was expressed in the twofold change in the price of ICs and in the degree of integration, occurring every two years (fig.3) [4, 5].

However, it turned out that the two-year pace of change had limitations. Further analysis showed that it slowed down (fig.4). Despite this, it can be certainly stated that there is still a possibility to formulate the change pattern in the microelectronics development rate, so Moore's Law is working.

The differences in the evaluation of Moore's Law are related

to its ambivalence. In the beginning, it was formulated as an economic law of exponential development of the new IC market, but it was based on the exponential growth of technology. The requirements for introduction of changes in the technology at a certain pace became the main perceived signs of Moore's Law. The economic component, while remaining important, was pushed in the evaluation of the Law to the background.

The effect of Moore's Law is supposed to result in a reduction of price. But the price of

what? The price of the main functional element, i.e. the transistor, or of a more dimensional function? Let us start with the first one.

An increase in the degree of integration, besides the reduction of the cost, is accompanied by improved performance of ICs. Thus, the performance of microprocessors at this pace of integration doubles every 18 months, which is due to the increase in the number of transistors and the reduction of delays in IC signals. Because of this Law, similar improvements can be formulated too for other

ния площади элементов, например, транзисторов. Это уменьшение (масштабирование) происходит пропорционально фактору масштабирования S . Необходимо отметить, что если в первое десятилетие в микроэлектронике сосуществовали различные схемотехнические базисы и, соответственно, различные технологии (биполярные и на основе МОП-транзисторов), то, начиная с 1970-х годов, первенство прочно захватил КМОП-базис на основе комбинированных (р- и n-канальных) МОП-транзисторов. Поэтому под законом Мура предполагают именно масштабирование КМОП-технологий, и, прежде всего, ее элемента – МОП-транзистора.

Исследование масштабирования полевых МОП-транзисторов Деннардом [6, 7] показало, что если не меняется напряженность поля, то напряжение и время задержки уменьшаются в S раз. Следовательно, быстродействие растет, а энергопотребление падает. Таким образом, масштабирование влияет на развитие микроэлектроники через изменение степени интеграции и цены функции ИС, а также повышение рабочих характеристик ИС. Исторически масштабирование ИС прошло несколько этапов, которые были отражены в их определениях в изданиях ITRS. В несколько адаптированном виде они приведены ниже.

ГЕОМЕТРИЧЕСКОЕ МАСШТАБИРОВАНИЕ

Геометрическое масштабирование (Geometrical Scaling) направлено на постоянное сокращение горизонтальных и вертикальных физических размеров областей функций логики и памяти для повышения

плотности элементов и функциональных характеристик (скорость, мощность и надежность), необходимых для приложений у конечного заказчика.

Масштабирование параметров производится после выбора одного из них постоянным. В качестве постоянного фактора при масштабировании обычно принимались постоянная напряженность поля или постоянное напряжение [8] (табл.1). В последнем случае были рассмотрены три варианта учета частоты, а также возможность учета многоядерной структуры.

Казалось, что масштабирование полностью решает проблемы миниатюризации при последовательном переходе к все более малым размерам структурных элементов МОП-транзистора. И действительно, вплоть до проектных норм 90 нм (2004 г.), происходила масштабная геометрическая миниатюризация при постоянном электрическом поле – все размеры транзистора и напряжения в нем изменялись с таким же коэффициентом, как и минимальный литографический размер. Этот период называют "золотой эрой миниатюризации", когда уменьшение размеров транзистора было относительно простой задачей, с малыми изменениями в его структуре при переходе от одного технологического уровня к другому. Однако ниже 90 нм дальнейшее простое геометрическое масштабирование приводило к ухудшению транзисторных характеристик.

ЭКВИВАЛЕНТНОЕ МАСШТАБИРОВАНИЕ

С уменьшением толщины затворного оксида ниже 1,6 нм резко возрастает ток прямого

characteristics of ICs, such as power consumption, cost, performance etc. Thus, Moore's Law frequently implies all the impacts of its formulation. In this form, Moore's Law made a transition from the sphere of economics to the sphere of technical forecasts. In this form, it gained fame and became a stimulus for microelectronics development.

In practice, Moore's Law involves planning and organizing scientific and technological activity, determining its areas and setting problems in a way that helps to continue the

certain trends in the development of microelectronics and, in a broader sense, of all scientific and technical progress. Thus, Moore's Law set up the rapid development of microelectronics, which resulted in the formation of a new technological structure, i.e. an informational society, and was essentially at the foundation of economic progress over the past 50 years. This development has had an undeniable impact on all sectors of production and changed the social and daily life. In microelectronics, Moore's Law have

resulted in the acceleration of scientific and technological revolution and the emergence of new interdisciplinary areas such as nanotechnology, artificial intelligence and cognitive technologies.

Increasing the degree of integration in IC is naturally associated with the reduction of the area of the elements, e.g., transistors. This reduction (scaling) is proportionate to the scaling factor S . It is noteworthy that if the first decade of microelectronics saw various circuit designs and, hence, all kinds of technologies (bipolar,



Таблица 1. Сценарии масштабирования [8]

Table 1. Scenarios of scaling [8]

	Постоянная напряженность Const field	Постоянное напряжение const V			
		Максимальная частота Max f	Постоянная частота Const f	Постоянная частота и число транзисторов в ядре Const f, N _{tran}	Многоядерная структура Multi core
Длина затвора, L _{gate} Ширина затвора, W; длина межсоединений, L _{wire} Напряжение, V	S S S	S S 1	S S 1	S 1 1	S S 1
Емкость, C Запасенная энергия, U _{stor} = ½ CV ²	S S ³	S S	S S	1 1	S S
Частота, f	1/S	1/S	1	1	1
Число транзисторов в ядре, N _{tran} /core Число ядер/площадь кристалла, N _{core} /A	1/S ² 1	1/S ² 1	1/S ² 1	1 1	1 1/S
Мощность цепи прибора, P _{ckr} Мощность на единицу поверхности, P/A	S ² 1	1 1/S ²	S 1/S	1 1	S 1
f N _{tran} N _{core}	1/S ³	1/S ³	1/S ²	1	1/S

квантово-механического туннелирования, и затворный диэлектрик из традиционного диоксида кремния теряет свои изолирующие свойства. Выход был найден в использовании

диэлектриков с высокой диэлектрической постоянной (high-k -диэлектриков).

С точки зрения управляемости транзистора, то есть возможности контролировать ток в канале

MOSFET transistor based), but since 1970s, the champion has been CMOS basis of complementary (p- and n- channel) MOSFET transistors. Therefore, Moore's Law implies the scaling of CMOS technologies, and, above all, its MOSFET transistor element.

Dennard's study of MOSFET field-effect transistor scaling [6, 7] showed that if the field strength does not change, the voltage and delay time reduce by S times. Consequently, the speed increases and the power consumption decreases. Thus, the scaling affects the

development of microelectronics via the change in the degree of integration, the IC price and the improvement of the IC performance. Historically, IC scaling had several stages described by their definitions in the ITRS publications. These definitions slightly adapted are given below.

GEOMETRICAL SCALING

Geometrical scaling means continuous reduction of physical horizontal and vertical dimensions of the functional logic and memory areas due to a higher density of elements and

improvement of functional characteristics (speed, capacity and reliability) required for applications of the end customer.

Parameters are scaled after selecting one of them as a constant. The constant field strength or constant voltage was usually taken as a constant factor for scaling [8] (Table 1). Three variants of frequency consideration and a possibility of multi-core structure consideration were studied in the latter case.

It seemed that the scaling resolved miniaturization problems by gradual transition

затворным напряжением, необходимо чтобы при масштабной миниатюризации удельная емкость затворного диэлектрика C_{ox} увеличивалась. При геометрическом масштабировании это увеличение обеспечивалось уменьшением толщины затворного диэлектрика. Однако можно не уменьшать толщину, а увеличить диэлектрическую проницаемость затворного диэлектрика, используя другой материал, что обеспечит необходимое значение C_{ox} . Таким образом исключается увеличение вертикального электрического поля и предотвращается затворная утечка. Используемый диэлектрик характеризуется значением "эквивалентной толщины оксида кремния" (ЭОТ-ЕОТ). Этот подход начали реализовывать еще с проектных норм 180 нм. В качестве альтернативного диэлектрика до проектных норм 65 нм использовался оксинитрид кремния с увеличивающейся долей нитрида по мере масштабирования (и, соответственно, увеличением диэлектрической проницаемости до 8-9). Начиная с проектных норм 45 нм в качестве затворных диэлектриков стали применять такие материалы, как HfO_2 и ZrO_2 . Замена SiO_2 на другие диэлектрики была первым примером "эквивалентного масштабирования".

Используемый в "золотую эру" масштабирования в качестве затвора легированный поликремний – вырожденный полупроводник с максимальной концентрацией порядка 10^{20} см^{-3} , поэтому на границе с затворным диэлектриком при рабочих напряжениях в нем образуется слой обеднения носителями и соответствующая ему

емкость обеднения. Так как эта емкость не масштабируется, необходим переход на новые материалы затвора – металлические. При этом новый материал должен обладать определенной работой выхода, необходимой для задания порогового напряжения транзистора. Такой переход был осуществлен на рубеже проектных норм 45 нм (2007 г.) одновременно с заменой оксинитрида кремния на новые диэлектрики.

Созданный таким образом "эквивалентный" затворный узел в англоязычной литературе получил обозначение НКМГ (High-K/Metal Gate – металлический затвор на диэлектрике с высокой диэлектрической постоянной К). Например, в исследовании ИМЕС были использованы затворы из TiN (для n-канального транзистора) и TaN (для p-канального), а в качестве затворного диэлектрика – HfO_2 , что обеспечивало ЭОТ менее 1 нм.

"Эквивалентный" затворный узел позволил при миниатюризации сохранять вертикальное поле в транзисторе постоянным, несмотря на замедление планомерного уменьшения напряжения питания. Поэтому, начиная с проектных норм 90 нм, дальнейшее масштабирование транзистора осуществляется при постоянной вертикальной и увеличивающейся горизонтальной составляющих электрического поля. Такой сценарий масштабирования резко обозначил паразитные "эффекты короткого канала", суть которых сводится к увеличению влияния напряжений на электродах стока и истока транзистора на протекание тока в канале, что приводит к деградации рабочих

to smaller dimensions of the structural elements of a MOSFET transistor. Indeed, before the design standards of 90 nanometers (2004), large-scale geometrical miniaturization had occurred with a constant electric field, when all the dimensions of a transistor and the voltage changed with the same ratio as the minimum lithographic size. This period was called the Golden Era of miniaturization, because reducing the size of a transistor was a relatively simple task, with small changes in its structure made during the

transition from one technological level to another. However, below 90 nanometers, further simple geometrical scaling resulted in deterioration of the transistor characteristics.

EQUIVALENT SCALING

When the thickness of gate oxide decreases below 1.6 nm, the current of direct quantum-mechanical tunneling increases dramatically and the gate dielectric insulator conventionally made of silicon dioxide loses its insulating properties. The solution was found in the use of dielectrics with a high-k of dielectrics.

From the perspective of transistor controllability, i.e. the possibility to control the current in the channel with gate power, it is necessary to have the specific capacitance of gate dielectric C_{ox} increased when scaling miniaturization. In geometrical scaling, this increase was ensured by reducing the thickness of the gate dielectric. However, rather than reducing the thickness, it is possible to increase the dielectric permeability of the gate using another material, which would provide the required value of the C_{ox} ,



характеристик транзистора. Для подавления этих паразитных эффектов были разработаны специальные методы "канальной инженерии" – создание определенного профиля легирования в области канала и стоков/истоков путем высокоэнергичной ионной имплантации под разными углами примесных атомов.

Другой эффект, который сильно ограничивал рабочие характеристики транзистора – сопротивление стоков/истоков, которое также не поддается геометрическому масштабированию. Проблема была решена "эквивалентным масштабированием" – последовательным переходом на новые материалы в шунтировании этого сопротивления, начиная от $TiSi_2$ (проектные нормы 350–250 нм) к $CoSi_2$ (90 нм) и далее к $NiSi$ (65–45 нм), а также на использование "приподнятых" эпитаксиальных стоков/истоков и $SiGe$ как промежуточного слоя между силицидом и кремнием.

При достижении проектных норм 28–22 нм традиционная структура МОП-транзистора себя исчерпала. Сильно возросший туннельный ток сток/исток-подложка приводит к существенному ухудшению рабочих характеристик. Был осуществлен переход на так называемые "полностью обедненные" структуры – КНИ (FD SOI) транзисторы и FinFET-транзисторы (fin – плавник). Здесь "эквивалентное" масштабирование заключается в использовании новых структур с наноразмерной толщиной слоя рабочего кремния (области под затво-

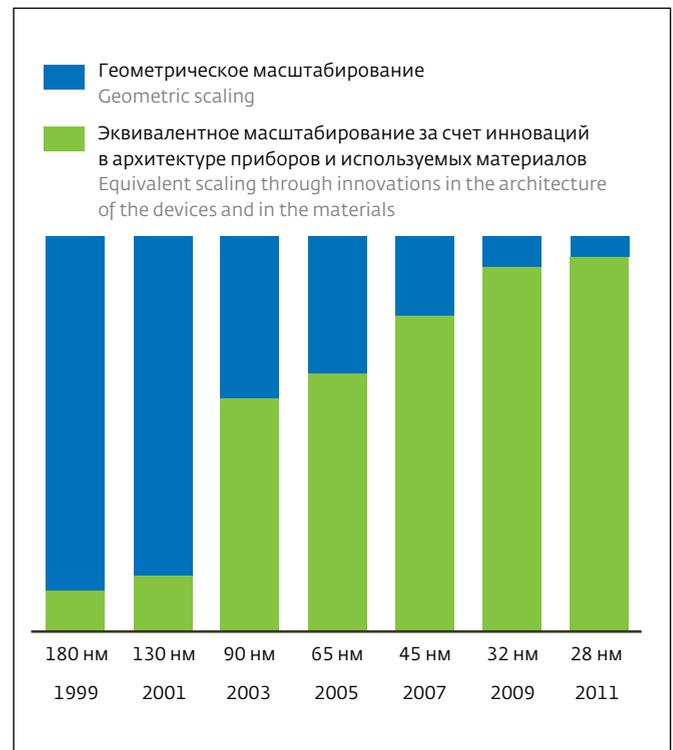


Рис.5. Относительные изменения функциональных характеристик [9]

Fig.5. Relative changes of the functional characteristics [9]

ром). Полностью обедненные транзисторы были названы "многозатворными" (MugFET). Условно FD SOI можно считать однозатворным транзистором, тогда как FinFET – двухзатворный. Предполагается, что по мере дальней-

thereby avoiding the increase in the vertical electric field and preventing gate leakage. The used dielectric is characterized by a value "Equivalent to silicon oxide thickness" (EOT). This approach was used even with the design standards of 180 nm. As an alternative dielectric before the standard of 65 nm there was silicon oxynitride with the share of nitride increased in scaling (and thus increasing the dielectric permeability up to 8–9). Starting with the design standards 45 nm, the gate dielectrics were made of such materials as HfO_2

and ZrO_2 . The replacement of SiO_2 with other dielectrics was the first example of "equivalent scaling".

Gate doped polysilicon used during the "Golden Era" of scaling is degenerate semiconductor with a maximum concentration of about 10^{20} cm^{-3} , therefore, it creates a carrier depletion layer at the boundary with the gate dielectric at working voltages and the corresponding depletion capacitance. Because this capacitance was not scalable, it was necessary to make transition to new metallic materials for gate. The new

material had to have a specific output needed for setting the threshold voltage of the transistor. This transition was made at the time of design standards of 45 nm (2007) together with the simultaneous replacement of silicon oxynitride with new dielectrics.

The newly created "equivalent" gate node in the English literature was designated as HKMG (High-K/Metal Gate). For example, in the IMEC study, valves from a TiN (for an n-channel transistor) and TaN (a p-channel) were used, and the gate was made from HfO_2 ,

шего масштабирования "эффективная многозатворность" транзисторных структур и "эквивалентная" составляющая в масштабировании элементов будут расти.

Таким образом, эквивалентное масштабирование (Equivalent Scaling) – улучшение рабочих характеристик 2D и 3D-структур не за счет изменения характерных геометрических параметров, а путем использования новых материалов, новых процессов и новых интегральных структур. Эквивалентное масштабирование может использоваться как без геометрического масштабирования, так и совместно с ним, дополняя друг друга. На рис.5 приведено соотношение вкладов геометрического и эквивалентного масштабирования в повышение рабочих характеристик ИС в период 1999–2011 годов.

ФУНКЦИОНАЛЬНОЕ МАСШТАБИРОВАНИЕ

Система, реализованная для выполнения специфических функций с помощью определенной технологии, называется функционально масштабированной (Functional Scaling) если может быть реализована с использованием альтернативной технологии так, что ее функции будут идентичны функциям оригинальной системы, и по крайней мере одна из ее рабочих характеристик улучшается, а остальные не ухудшаются. На практике функциональное масштабирование предполагает такое проектирование ИС, при котором ее последующее изготовление возможно по различным проектным нормам.

ЭКВИВАЛЕНТНОЕ МАСШТАБИРОВАНИЕ НА ОСНОВЕ ПРОЕКТИРОВАНИЯ

Масштабирование элементов ИС выдвигает новые проблемы, непосредственно касающиеся проектирования ИС как системы. Поэтому необходима разработка новых методов проектирования, которые позволяют при масштабировании достичь рабочих характеристик ИС, включая потребляемую мощность и стоимость в соответствии с законом Мура. Такой подход получил название "эквивалентное масштабирование на основе проектирования" (Design-based Equivalent Scaling). Оно включает геометрическое и эквивалентное масштабирование элементов ИС и позволяет достичь улучшения рабочих характеристик, плотности упаковки и других показателей негеометрическими методами. Взаимосвязь типов масштабирования приведена на рис.6.

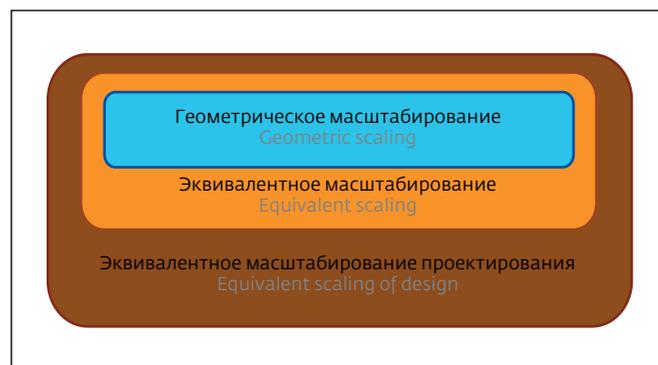


Рис.6. Соотношение типов масштабирования [10]
Fig.6. Correlation of scaling types [10]

which provided EOT less than 1 nm.

In miniaturization, the "equivalent" gate node was used to keep the vertical field constant in the transistor, despite a systematic slowdown in the reduction of the power supply voltage. Therefore, starting with design standards of 90 nanometers, further scaling of the transistor was carried out with constant vertical and increasing horizontal components of the electric field. This scenario of scaling was sharply marked by spurious "short-channel effects",

which is essentially an increase in the influence of voltages at the drain and source electrodes of the transistor on the current flowing in the channel that leads to degradation of the functional characteristics of the transistor. Special methods of "channel engineering" were developed to suppress these spurious effects, which was the creation of a specific doping profile in the channel and drains/sources by high-energy ion implantation at different angles of the dopant atoms.

Another effect, which severely limited the

performance of the transistor, was the resistance of drains/sources, which did not allow geometrical scaling either. The problem was resolved by "equivalent scaling", i.e. gradual transition to new materials in the shunting of this resistance, from $TiSi_2$ (design standards of 350–250 nm) to $CoSi_2$ (90 nm) and then to $NiSi$ (65–45 nm), and the use of "elevated" epitaxial drains/sources and SiGe as an intermediate layer between silicide and silicon.

Upon reaching the design standards of 28–22 nm, the conventional structure of a MOS



Различные методы эквивалентного масштабирования (как интегральных элементов, так и на основе проектирования) могут быть применены для различных типов ИС. Это привело к тому, что улучшения рабочих характеристик в соответствии с законом Мура стали характерными для определенного типа ИС. Поэтому по мере масштабирования происходит дальнейшая дифференциация и специализация ИС на уровне проектирования, технологических процессов, а также проектных норм.

Таким образом, в общем смысле масштабирование следует понимать как методы повышения рабочих и функциональных характеристик прибора, начиная от транзистора и заканчивая всей системой.

Среди рассмотренных методов масштабирования важное значение имеет уменьшение размеров, то есть геометрическое масштабирование. В этом случае фактор масштабирования S дает величину относительных изменений геометрических размеров.

Абсолютные значения могут быть получены, если определена мера минимального характеристического размера – f . Определение характеристического размера приводится в изданиях ITRS для горизонтальных или вертикальных сечений элементов транзистора [11]. Наиболее часто используется величина, равная половине зазора между линиями проводников (half-pitch – $h-p$) первого слоя металлизации для ИС ДОЗУ или половина шага полукремниевых шин для флэш-ЗУ.

В принципе подобные определения могут быть использованы и для логических ИС, в том числе для микропроцессоров (МП). Но чаще в качестве f для МП используют длину затвора: физическую (GLph – physical gate length) или литографическую (GL – Printed gate length). Такое различие в определении f вызвано различием в критических параметрах ЗУ и МП. Для ИС ЗУ критична степень интеграции, которая может быть охарактеризована через величину ($h-p$). Для МП критично быстроедействие, которое определяется через GLph.

Изменение ($h-p$) происходит на 0,71 за один временной цикл или на 0,5 за два временных цикла. В 2013 году ($h-p$) менялся по трехгодичному циклу – за 6 лет на 0,5. Ожидается, что с 2017 года вновь начнется двухгодичный цикл.

Если при геометрическом масштабировании минимальный характеристический размер f полностью определял уровень технологии, то по мере возрастания роли эквивалентных типов масштабирования он теряет это свое значение. Поэтому была введена характеристическая мера масштабирования, называемая "узлом" (Node Naming – N), которая отражает условный размер минимального элемента, как он бы изменялся при полностью геометрическом масштабировании. Характеристики f для разных типов ИС через значения $h-p$ и N приведены в табл.2.

Разные типы масштабирования стали использоваться на разных этапах развития

transistor was exhausted. A strong increase in the tunneling current of the drain/source-substrate resulted in a significant deterioration of performance. The industry made transition to the so-called "fully depleted" structures of silicon-on-insulator (FD SOI) transistors and FinFET transistors. Here, the "equivalent" scaling is related to the use of new structures with nanoscale thickness of the working silicon layer (the area under the gate). Fully depleted transistors were named "multi-gate" transistors (MugFET). Conventionally,

FD SOI may be viewed as a one-gate transistor, whereas FinFET is a two-gate transistor. It is expected that with further scaling, the effective multi-gate feature of transistor structures and the "equivalent" component in scaling elements will only grow.

Thus, the equivalent scaling means improved performance of 2D and 3D structures, not due to changes in the typical geometrical parameters, but due to the use of new materials, new processes and new integrated structures. Equivalent scaling can be used either

without geometrical scaling or with it, complementing each other. fig.5 shows the ratio of the contributions of geometric and equivalent scaling to the improvement of the IC performance in 1999-2011

FUNCTIONAL SCALING

A system implemented to perform specific functions using a particular technology is called functional scaling if it can be implemented using alternative technologies so that its functions are identical to the original functions of the system, while at least one of its performance



Таблица 2. Целевые показатели ИС [10]

Table 2. Target specifications of IC [10]

Год производства Year of production	2013	2015	2017	2019	2021	2023	2025	2028
N	"16/14"	"10"	"7"	"5"	"3.5"	"2.5"	"1.8"	
Логика ½(p-h), нм Logic ½(p-h), nm	40	32	25	20	16	13	10	7
Флэш (2D) ½(p-h), нм Flash (2D) ½(p-h), nm	18	15	13	11	9	8	8	8
ДОЗУ ½(p-h), нм DRAM is ½(p-h), nm	28	24	20	17	14	12	10	7.7
FinFET ½(p-h), нм	30	24	19	15	12	9.5	7.5	5.3
FinFET, ширина Fin, нм FinFET, Fin width, nm	7.6	7.2	6.8	6.4	6.1	5.7	5.4	5.0
6Т СОЗУ размер ячейки (60f²), мкм² 6T SRAM the cell size (60f²), μm²	0.096	0.061	0.038	0.024	0.015	0.010	0.0060	0.0030
Генерация флэш ЗУ (бит на кристалл) Generation of flash memory (bits per chip)	64G/128G	128G/256G	256G/512G	512G/1T	512G/1T	1T/2T	2T/4T	4T/8T
Генерация ДОЗУ (бит на кристалл) Generation of DRAM (bits per chip)	4G	8G	8G	16G	32G	32G	32G	32G
V _{dd}	0.86	0.83	0.80	0.77	0.74	0.71	0.68	0.64
МП высокого уровня, GLpr, нм High-level MP, GLpr, nm	28	22	18	14	11	9	7	5
МП высокого уровня, GLph, нм High-level MP, GLph, nm	20	17	14	12	10	8	7	5
ASIC низкой мощности, GLph, нм Low power ASIC, GLph, nm	23	19	16	13	11	9	8	6

parameters improves and others deteriorate. In practice, functional scaling involves designing an IC for which subsequent manufacturing is possible using different design standards.

EQUIVALENT SCALING BASED ON THE DESIGN

Scaling of IC elements poses challenges that are directly related to the design of ICs as a system. Therefore, it is necessary to develop new design methods that allow scaling with better IC performance parameters, including power consumption and cost, in line with Moore's

Law. This approach is called "design-based equivalent scaling". It includes geometrical and equivalent scaling of IC elements and allows achieving improved performance, packing density and other indicators, using non-geometrical methods. The correlation between types of scaling is shown in fig.6.

Various methods of equivalent scaling (both integral elements and design-based scaling) can be applied for different types of ICs. This resulted in improved performance according to Moore's Law becoming a characteristic for a specific

type of ICs. Thus, with further scaling, there is further differentiation and specialization of ICs at the level of design, technological processes and design standards.

Therefore, in a general sense, scaling should be understood as methods of improving the working and functional characteristics of a device, from a transistor to the whole system.

Among the reviewed methods of scaling, the important one is related to reducing the size, i.e. the geometrical scaling. In this case, the scaling factor S provides the magnitude



Таблица 3. Сопоставление этапов развития полупроводниковой промышленности и последовательности этапов масштабирования
Table 3. A comparison of the stages of semiconductor industry development and stages of scaling

	Этапы развития полупроводниковой промышленности Development of the semiconductor industry	Этапы развития масштабирования Development of scaling
1960-е годы 1960s		1965 год – формулировка закона Мура 1965 – formulation of Moore's law
1970-е годы 1970s	Производство компонентов: • эффективных ЗУ для ВТ; • ASIC для реализации специальных функций Production of components: • efficient memory for computing; • ASIC for special functions	1975 год – масштабирование по Деннарду 1975 – Dennard scaling
1980-е годы 1980s	Спецификация ИС в руках системных интеграторов. Новые технологии вводятся каждые 3 года, вначале для ЗУ, а затем для логических ИС (МП) Specification of IC in the hands of system integrators. New technologies are introduced every 3 years, first for memory, and then for logic ICs	Первая эра – эра классического геометрического масштабирования The first era – the era of the classical geometric scaling
1990-е годы 1990s	Логические ИС развиваются ускоренно в связи с введением новых технологий каждые 2 года. Корреляция между улучшением технологии при масштабировании и повышением показателей продукта вызывает частичный переход к контролю системной функциональности и доходов к производителям ИС. Прибыль растет на 17% в год Logic ICs are developing rapidly due to the introduction of new technology every 2 years. The correlation between the improvements of technology and products causes a partial shift to control of system functionality and income to producers of IC. Profits is growing at 17% per year	Вторая эра – эра эквивалентного масштабирования (конец 1990-х годов). Введение новых материалов для новых структур элементов транзисторов The second era – the era of equivalent scaling (late 1990s). The introduction of new materials for new structures of transistors

of relative changes in geometrical dimensions.

Absolute values can be obtained if the measurement of the minimum characteristic size f is determined. The determination of the characteristic size is described in the publications of the ITRS on horizontal or vertical cross sections of transistor elements [11]. The most often used value is equal to half of the pitch between the lines of conductors (half-pitch – $h-p$) of the first layer of metalization for DOSE IC or half step of polysilicon buses for flash memory.

In principle, this determination may be used for logic ICs, including microprocessors (MPs). However, the length of the gate is more often used as f for MPs: physical gate length (GLph) or printed gate length (GL). This difference in the determination of f is caused by the difference in the critical parameters of the memory and MPs. For IC memory, the degree of integration is critical, which can be characterized by the value ($h-p$). For MP, speed is critical, determined via GLph.

The change of ($h-p$) occurs at 0.71 per one temporal cycle

or 0.5 per two temporal cycles. In 2013 ($h-p$) was changed in a three year cycle – by 0.5 in 6 years. It is expected that 2017 the two-year cycle will resume.

While, in geometrical scaling, the minimum characteristic size f fully determined the technological level, it loses its value with the increasing role of equivalent scaling. Therefore characteristic scaling measure was introduced, called a "node" (Node Naming – N), which reflects how the conditional minimum element size would change in full geometrical scaling. The f characteristics of



	Этапы развития полупроводниковой промышленности Development of the semiconductor industry	Этапы развития масштабирования Development of scaling
2000-е годы 2000s	<p>Последнее десятилетие – возникновение новой экосистемы полупроводниковой промышленности:</p> <ul style="list-style-type: none"> • агрессивный ввод новых технологий каждые два года, эффективные комплексные системы в одном корпусе (СНК, СВК); • развитие системы foundry позволяет изготавливать ИС по приемлемой цене, что обеспечило возможность развития fabless; • развитие полупроводниковых технологий обеспечило активный рост производства и рынка новых типов приборов (МЭМС, плоские панели, РЧ и т.д.) 	<p>2005 год – прекращение действия закона Деннарда Ограничения 2D-масштабирования Переход к 3D-структурам 2005 – the termination of the Dennard's law Limitations of 2D scaling The transition to 3D structures</p>
2010-е годы 2010s	<p>The past decade – the emergence of a new ecosystem in the semiconductor industry:</p> <ul style="list-style-type: none"> • aggressive introduction of new technology every two years, effective integrated systems in a single case; • development of foundry allows the manufacturing of IC at a reasonable price, which enabled the fabless development; • the development of semiconductor technologies has provided rapid growth of production and market for new types of devices (MEMS, flat panel, RF, etc.) 	<p>Третья эра масштабирования – комбинация 3D архитектуры с низкой мощностью приборов (3D масштабирование мощности – 3D Power Scaling). Увеличение числа транзисторов на 2D слое дополняется увеличением числа слоев. Плотность возрастает без необходимости уменьшения размера элемента The third era of scaling: combination of 3D architecture with low power devices (3D Power Scaling). The increase in the number of transistors on a 2D layer is complemented by an increase in the number of layers. The density increases without the need to reduce the size of the element</p> <p>Главным становится не увеличение функциональности, а снижение мощности. Минимизация мощности – главный стимул проектирования The main thing is not increasing the functionality but reducing power. The power minimization is the main stimulus of the design</p>

different types of ICs via the values of h - p and N are given in Table 2.

Different types of scaling were used at different stages of development of the semiconductor industry, shown in a comparative table (Table 3).

The power consumption of MOS transistor P , which transforms into heat when switching with the frequency f , is derived from dynamic P_{dyn} and static P_{stat} powers:

$$P = P_{dyn} + P_{stat}, \quad (1)$$

$$P_{dyn} = CV_{dd}^2 f, \quad (2)$$

$$P_{stat} = V_d I_{off}, \quad (3)$$

where C is the load capacitance, I_{off} is the current in off state (leakage current), V_{dd} is the supply voltage IC.

The dynamic power consumption can be diminished by reducing the system's frequency f . The static power of idle blocks can be reduced to zero by disconnecting them from the power supply. These two techniques, which have been widely used in the development of ICs with the design standards below 180 nm, are the first examples

of design-based equivalent scaling.

Power consumption is a restricting factor primarily in systems working on accumulators (laptops, tablets, cell phones, etc.). ICs for them are put into a separate low-power class. Circuits, for which the primary factor is the performance and the maximum power consumption is determined by the heat sink, are put into a high performance class.

An example of design-based equivalent scaling is the use of multiprocessor systems. Parallelization of computation



полупроводниковой промышленности, поэтому можно составить сравнительную таблицу (табл.3).

Потребляемая мощность МОП-транзистора P , которая выделяется в виде тепла, при переключении с частотой f складывается из динамической P_{dyn} и статической P_{stat} мощностей:

$$P = P_{\text{dyn}} + P_{\text{stat}}, \quad (1)$$

$$P_{\text{dyn}} = CV_{\text{dd}}^2 f, \quad (2)$$

$$P_{\text{stat}} = V_{\text{d}} I_{\text{off}}, \quad (3)$$

где C – емкость нагрузки, I_{off} – ток в выключенном состоянии (ток утечки), V_{dd} – напряжение питания ИС.

Динамическую потребляемую мощность можно снизить, уменьшив частоту f системы. Статическую мощность неработающих блоков можно свести к нулю, отключив от них напряжение питания. Эти два приема, которые стали широко использоваться при разработке ИС проектных норм ниже 180 нм, – первые примеры эквивалентного масштабирования на основе проектирования.

Потребление мощности является ограничительным фактором, в первую очередь, в системах, работающих от аккумуляторов (ноутбуки, планшеты, сотовые телефоны и т.д.). ИС для них выделились в отдельный класс низко-потребляющих (low-power – LP). Схемы, для которых первостепенным фактором является производительность, а максимальная потребляемая мощность определяется теплоотводом, выде-

лились в класс высокопроизводительных (high performance – HP).

Примером эквивалентного масштабирования на основе проектирования является использование многопроцессорных систем. Распараллеливание вычислительного процесса с сохранением частоты привело к увеличению производительности системы без значительного увеличения потребляемой мощности.

Анализ данных о росте количества транзисторов на кристалле за последние годы показывает, что наметилось его замедление (рис.4). Хотя прогресс в области литографии обеспечивал по крайней мере до 2013 года возможности геометрического масштабирования в соответствии с законом Мура, реальный рост плотности транзисторов, начиная с 2007 года, уменьшился до кратности 1,6 с каждым новым технологическим "узлом". Такое отставание объясняют как экономическими причинами, так и техническими ограничениями, накладываемыми надежностью, условиями эксплуатации и архитектурой проекта, а также увеличивающейся вариативностью технологических процессов. Этот разрыв между обусловленной литографическим шагом "возможной" и "реализуемой" плотностью элементов поставил вопрос о действенности закона Мура в его классической формулировке.

Переход на новые технологии ведет к существенному усложнению проектирования ИС. Это связано с существенными ограничениями, накладываемыми на взаиморасположение слоев, а также с мульти-масочными литографиями, приводящими

with preservation of frequency resulted in increased performance without great increase in power consumption.

Analysis of data on the increasing number of crystal-based transistors in recent years has shown that there was a slowdown (fig.4). Although the progress in the field of lithography ensured, at least until 2013, geometrical scaling in accordance with Moore's Law, the real growth in transistor density decreased to a ratio of 1.6 since 2007 with each new technological "node". This lag is explained by both

economic factors and technical constraints related to reliability, operating conditions and design architecture, as well as the increasing variability of technological processes. This gap between the lithographic phases of "possible" and "feasible" densities of elements raised the question of the validity of Moore's Law in its classical formulation.

Adoption of new technologies leads to a significant complication of IC design. This is connected with significant constraints laid on relative positions of layers, as well as

multi-mask lithographs, resulting in a large variability of spurious elements. Consequently, there is an increase in the risks for project profitability.

The situation is redeemed by the methods of design-based equivalent scaling (DES). According to an optimistic forecast, for server and desktop systems processors, DES may replace one node of scaling in the period up to 2020. It is suggested that DES may potentially reduce the size of logic circuits up to 63% of the current size over the next six years and compensate for the 1.6-fold increase

к большой вариабельности паразитных элементов. Вследствие этого увеличиваются риски окупаемости проектов.

Ситуацию спасают методы эквивалентного масштабирования на основе проектирования (DES). Согласно оптимистическому прогнозу, для серверных процессоров и процессоров для настольных систем, DES может возместить один узел масштабирования в период до 2020 года. Предполагается, что потенциально DES может снизить площадь логических схем до 63% от существующего уровня в течение последующих шести лет и компенсирует 1,6-кратный рост транзисторной плотности, то есть "спасает" закон Мура на ближайший период времени.

ПЕРСПЕКТИВЫ

В настоящее время размерное и функциональное масштабирование КМОП-технологии является основой развития и реализации в различных приложениях технологий процессинга информации – приема, преобразования, хранения, обработки и передачи данных. Многие из этих приложений стали реализуемыми благодаря возросшей производительности и сложности систем, что обеспечивалось и обеспечивается масштабированием ИС. Поскольку геометрическое и эквивалентное масштабирование КМОП приближается к фундаментальным пределам, необходим поиск как альтернативных приборов, так и альтернативных микроархитектур для процессинга информации, которые бы расширяли функциональность системы. Это позволило бы поддержать

масштабирование при снижении стоимости и увеличении функциональности систем процессинга информации. Здесь возможны два подхода:

1. расширение функциональности КМОП-платформы за счет интеграции КМОП-технологии с новыми технологиями разных типов;
2. разработка абсолютно новой парадигмы процессинга информации.

Разделение микроэлектроники на отдельные домены было предложено ITRS в 2005 году, и стало общепринятым (рис.7 и 8). Соотношения между обозначенными выше подходами схематично представлено на рис.9.

Развитие КМОП-платформы за счет различных типов масштабирования получило название "Больше Мура". Расширение функциональности интегральной системы получило название "Больше чем Мур". Новые приборы для процессинга информации, так же как и новые архитектурные подходы, стали называть "Вне КМОП"-технологии. Гетерогенная интеграция подходов "вне КМОП", как и подхода "Больше чем Мур" с развивающейся платформой "Больше Мура" образуют "расширение КМОП". При этом предполагается, что со временем роль технологической парадигмы "снизу-вверх" в изготовлении интегральных систем будет возрастать. Такая общая концепция гетерогенной интеграции предоставляет возможность дальнейшего масштабирования систем процессинга информации и обеспечивает переход от КМОП-платформы к новым областям нанoeлектроники.

in transistor density, i.e. it "saves" Moore's Law for the near future.

PROSPECTS

Currently, dimensional and functional scaling of CMOS technology is the basis for development and implementation in various applications of information processing technologies – data ingestion, transformation, storage, processing and transmission. Many of these applications have become feasible thanks to the increased performance and complexity of the systems that have been provided

by IC scalability. Since geometrical and equivalent CMOS scaling is approaching its fundamental limitation, we need to search for alternative devices and alternative microarchitectures for processing information that would expand the functionality of the system. This would support scaling while reducing the cost and increasing functionality of the information processing system. There are two possible approaches:

1) expansion of the CMOS platform functionality by integrating of CMOS technology

with new technologies of different types;

2) developing a brand new paradigm of information processing.

The division of microelectronics into the individual domains was proposed by ITRS in 2005, and it became widely accepted (fig.7 and 8). The ratio between the aforementioned approaches is schematically shown in fig.9.

The development of a CMOS platform due to different types of scaling is called "More Moore". The expansion of the integrated system is called

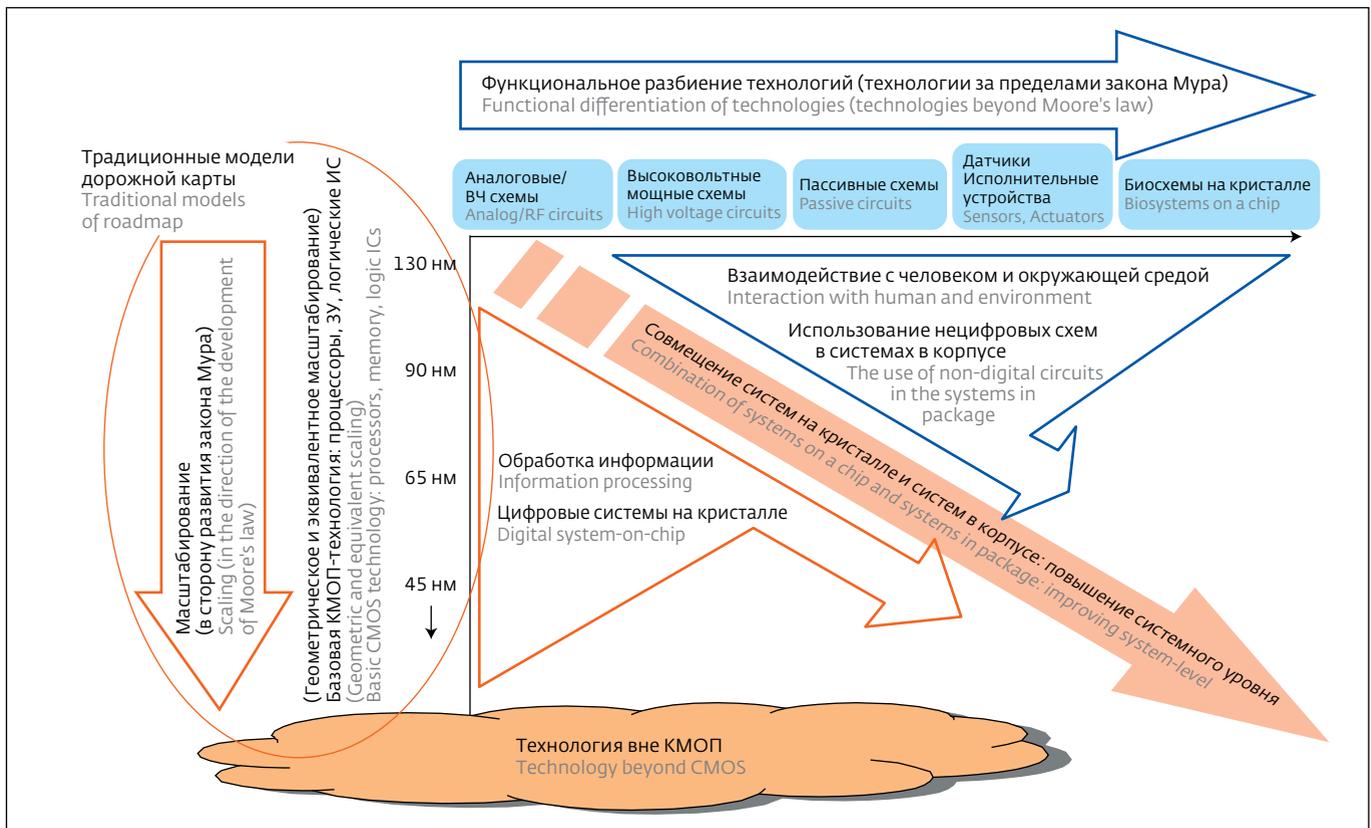


Рис.7. Развитие полупроводниковой технологии применительно к закону Мура [12]
 Fig.7. The development of semiconductor technology in relation to Moore's law [12]

В общем случае система процессинга информации включает несколько иерархических и взаимодействующих уровней. Так, например, при проектировании ИС эти уровни прослежи-

ваются сверху-вниз – из определения назначения системы и ее функции строится архитектура, и далее последовательно: микроархитектура, отдельные блоки (схемы), элементы (транзисторы

"More than Moore". New devices for processing information and new architectural approaches were called "Beyond CMOS" technologies. Heterogeneous integration of "beyond CMOS" approaches, as well as "More than Moore" approach, with the emerging "More Moore" platform shape the "CMOS extension". This means that over time the role of technological "bottom-up" paradigm in the production of integrated systems will increase. This general concept of heterogeneous integration provides the possibility of further scaling

of information processing systems and provides a transition from the CMOS platform to new nanoelectronics areas.

In general, the processing information system includes several hierarchical and interactive levels. For example, in the design of ICs, these levels may be traced top-down – the architecture is built on defining the purpose of the system and its functions, and then sequentially come: microarchitecture, individual blocks (circuits), elements (transistors and wires), a required technological process (which involves

the use of certain materials), which form individual elements and the system as a whole. Each level of the hierarchy has its own technology. However, one can also build bottom-up hierarchy by first determining the lowest physical level, i.e. the variable of the informational state, and ending in a system architecture (fig.10).

In this representation, the processing information system has a device level, and the fundamental unit of information is represented by a variable of state, for example, as it

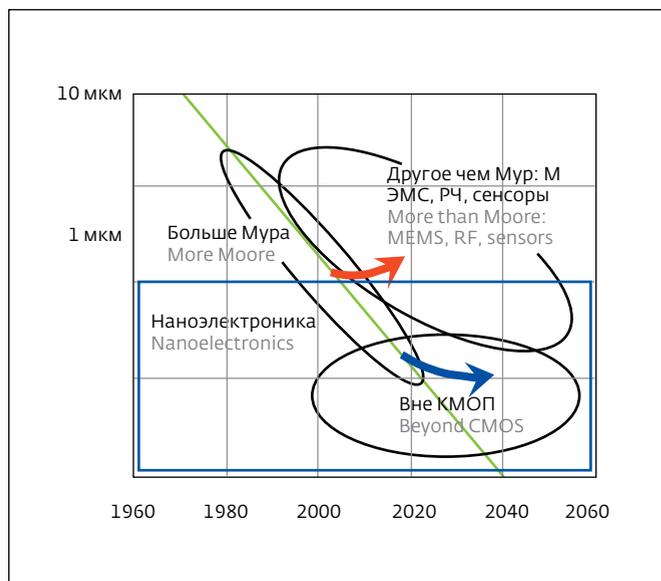


Рис.8. Домены развития микроэлектроники. Стрелки демонстрируют влияние домена "Больше Мура" на остальные [13]
 Fig.8. The domains of development of microelectronics. The arrows show the influence of More Moore domain on the other [13]

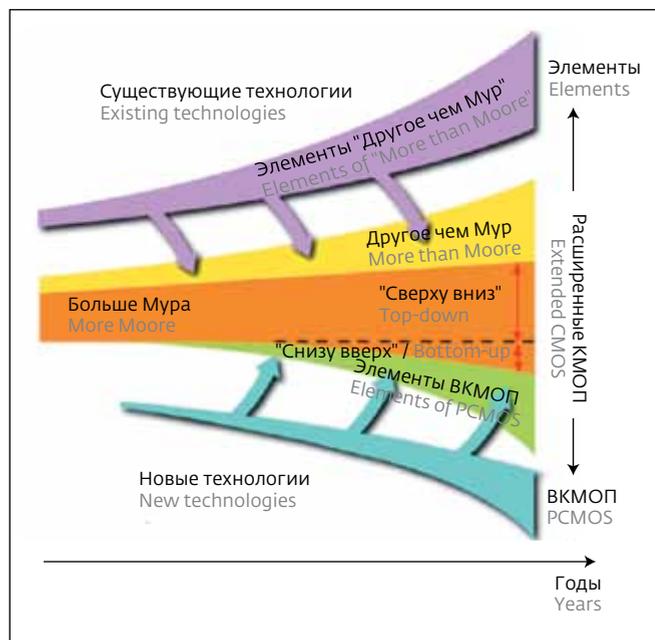


Рис.9. Эволюция расширенных КМОП [10]
 Fig.9. Evolution of extended CMOS [10]

и разводка), подбирается необходимый технологический процесс (подразумевает использование определенных материалов), по которому формируются отдельные элементы и система в целом. Каждому уровню иерархии соответствует своя технология. Но можно построить такую иерархию и снизу-вверх, определив сначала низший физический уровень – переменную информационного

состояния, и закончив архитектурой системы (рис.10).

В таком представлении в системе процессинга информации выделяют уровень прибора, а фундаментальная единица информации представлена переменной состояния, например, как это делается сейчас – зарядом (напряжением) в узле схемы КМОП, или, как это было в прошлом, распо-

is currently done, by the charge (voltage) at the node of the CMOS circuit, or, as it was in the past, by the location of the bone in the abacus. The device in this case is used to represent a physical value and ensures the control of transitions of the informational variable between two or more allowed conditions. It is a physical structure made of different materials with special properties through the sequence of technological operations. Thus, it is possible to identify the most important hierarchical level, which is – the level of materials science and technology. The

level of data representation in this hierarchy describes how the variable of state is encoded by a group of devices for processing data. The two best-known examples of data representation are digital and analog signals. The higher level of architecture can be divided into several subclasses. The subclass nano-architecture or functional primitives is used to build another subclass of a computational model of an information processing model, for example, devices for logical or arithmetic transformations, memory, cellular automata chains etc.

In fig.10, the yellow rectangles in a red box highlights the elements corresponding to the CMOS technological platform. It is characterized by the binary data representation based on a variable of state, i.e. electric charge, which serves as the basis for construction of Von Neumann's computer architecture. The elements drawn in white boxes in this fig. are grouped into five categories and they represent some of the most promising innovative technologies that can become the basis of a new scaling paradigm for processing information systems. ■

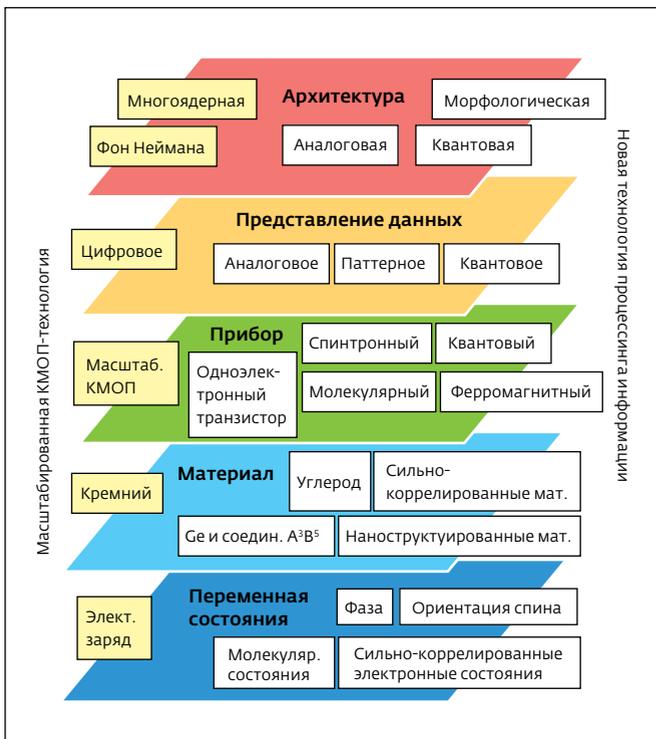


Рис.10. Систематика развития существующих и перспективных устройств процессинга информации
Fig.10. Systematics of the development of the existing and future devices of information processing

ложением косточки на счетах. Прибор в этом случае служит для представления физического значения и обеспечивает управление переходами информационной переменной между двумя или несколькими разрешенными состояниями. Он представляет собой физическую структуру, сформированную различными материалами со специальными свойствами посредством последовательности технологических операций. Таким образом, можно выделить важнейший иерархический уровень – уровень материаловедения и технологии. Уровень представления данных в такой иерархии характеризует то, как переменная состояния кодируется группой приборов для процессинга данных. Два наиболее известных примера представления данных – цифровой и аналоговый сигналы. Более высокий уровень архитектуры можно разбить на несколько подклассов. Подкласс нано-архитектуры или функциональных примитивов служит для построения другого подкласса – вычислительной модели или модели процессинга информации, например, устройства логических или арифметических преобразований, памяти, цепи клеточных автоматов и т.д.

На рис.10 желтыми прямоугольниками в красной рамке выделены элементы, соответствующие технологической платформе КМОП. Для нее характерно бинарное представление данных, основанное на переменной состояния – электрическом заряде, что служит основой для построения вычислительной архитектуры фон Неймана. Элементы, заключенные на рисунке в белые прямоугольники, сгруппированы в пять категорий и представляют некоторые наиболее перспективные инновационные технологии, которые могут стать основой новой масштабируемой парадигмы систем процессинга информации.

ЛИТЕРАТУРА

1. **Or-Bach Zvi.** Are we using Moore's name in vain? // <http://electroiq.com/blog/2013/11/are-we-using-moores-name-in-vain/>.
2. **Applewhite A.** Take away the semiconductor // IEEE Spectrum, November 2004? INT.
3. **Takahashi D.** Intel's Gordon Moore speculates on the future of tech and the end of Moore's Law // <http://venturebeat.com/2015/05/11/intels-gordon-moore-speculates-on-the-future-and-the-end-of-moores-law/>.
4. **Балякин А.А., Доминич А.С.** Конструирование будущего: нелинейная динамика в экономических моделях // Интеграл. №1. 2011. С. 33-35.
5. **Elert E.** Tech Trajectories: Four More Moore's Laws. Semiconductor technologies aren't the only ones that have gotten exponentially better // <http://spectrum.ieee.org/at-work/innovation/tech-trajectories-four-more-moores-laws;>
6. Закон масштабирования Деннарда // Открытые системы, №2, 2012;
7. **Dennard R.H.** Design Of Ion-implanted MOSFET's with Very Small Physical Dimensions // IEEE J. of SSC, VSC-9. 1974. №5. P. 256.
8. **Theis T.N., Solomon P.M.** In Quest of the "Next Switch": Prospects for Greatly Reduced Power Dissipation in a Successor to the Silicon Field-Effect Transistor // Proceeding of the IEEE. 2010. V. 98. №12. P. 2005-2014.
9. **Garrou P.** IFTLE 191 ITRS Echoes Dylan "The Times They are a-Changin'" // <http://electroiq.com/insights-from-leading-edge/2014/05/iftle-191-itrs-echoes-dylan-the-times-they-are-a-changin/>.
10. ITRS 2013 Edition // <http://www.itrs.net>.
11. ITRS // <http://www.itrs.net>.
12. ITRS 2005 Edition // <http://www.itrs.net>.
13. **Jonescu A.M.** Nanoelectronics roadmap: evaling Moor's law // NANOLAB.