

# Метод снижения статической мощности КМОП-схем на основе отключающих транзисторов с контролем быстродействия

П.С. Волобуев, С.В. Гаврилов, Д.И. Рыжова

Институт проблем проектирования в микроэлектронике РАН (ИППМ РАН), [synopsis@bk.ru](mailto:synopsis@bk.ru)

**Аннотация** — При проектировании интегральных микросхем на основе технологий с проектными нормами 90нм и ниже приоритетной задачей становится уменьшение статической потребляемой мощности. Существующие средства САПР обеспечивают уменьшение потребляемой мощности на основе различных схемотехнических решений, в частности, применяющих отключающие транзисторы. Существующий маршрут на основе библиотечных элементов не обеспечивает требуемый контроль быстродействия при применении таких решений. Данная статья посвящена решению проблем контроля быстродействия в маршруте автоматизированного синтеза схемотехнических решений по снижению статической мощности сложно-функциональных блоков (СФ-блоков). Предложен подход на основе метода отключения питания с анализом пикового тока в активном режиме работы схемы. Подход позволяет автоматизировать выбор требуемой ширины отключающих транзисторов на основе точного анализа быстродействия КМОП схемы в активном режиме.

**Ключевые слова** — сложно-функциональный блок (СФ-блок), статическая потребляемая мощность.

## I. ВВЕДЕНИЕ

Порядок значений токов утечек в общем балансе энергопотребления с использованием современных полупроводниковых технологий сравним с порядком значений динамического тока, возникающего при зарядке емкостей. Проблема токов утечки в режиме бездействия усугубляется с течением времени, и при использовании нанометровых технологий потребляемая мощность в режиме ожидания становится больше мощности, потребляемой в активном режиме [1]. Кроме технологических методов, таких как использование диэлектриков с высокой диэлектрической проницаемостью, уменьшающих токи утечки через подзатворный диэлектрик, или FinFET транзисторов, обладающих трехмерным каналом для уменьшения подпороговых токов утечек, достаточно эффективны динамические методы уменьшения статических токов. Среди таких методов наибольшее распространение в реальных маршрутах проектирования получили методы отключения напряжения питания. Класс динамических методов значительно уменьшает токи утечки в режиме

сна, но сложен в практической реализации. При проектировании требуется учитывать несколько временных параметров: время пробуждения, или время перехода из режима сна в активный режим, а также время переключения схемы в активном режиме. Все они являются важными, но особое внимание в статье уделяется разработке схемотехнических решений по оптимизации статической мощности на основе отключающих транзисторов [2] с учетом контроля задержки в активной фазе работы схемы.

Очевидно, что размеры отключающих транзисторов оказывают существенное влияние как на время пробуждения схемы, так и на время срабатывания схемы в активном режиме. Поскольку в существующих маршрутах проектирования не предусмотрен анализ быстродействия библиотечных элементов с учетом отключающих транзисторов, то на практике используются завышенные размеры транзисторов, гарантирующие соблюдение требований по быстродействию схемы. В данной статье предлагается подход, позволяющий уменьшить размер отключающего транзистора на основе более точного анализа быстродействия.

## II. СТАТИЧЕСКАЯ ПОТРЕБЛЯЕМАЯ МОЩНОСТЬ

Статическая мощность для КМОП схем не зависит от рабочей частоты и потребляется в схеме всегда:

$$P_{stat} = I_{leak} V_{DD}.$$

Рассматривая статический ток утечки  $I_{leak}$  можно выделить три основные компоненты токов утечек: подпороговые токи утечек, токи утечек через затвор и токи утечки через p-n переход. Отключенные КМОП приборы имеют нулевое напряжение затвор-исток  $V_{gs}$ , а  $V_{ds} = V_{DD}$ , таким образом токи утечки текут от истока к стоку, через затвор и подложку. Подпороговые токи являются важной компонентой общего статического тока потребления МОП транзистора в отключенном состоянии. Если к транзистору в этом режиме приложено напряжение  $V_{ds}$ , то из-за разности концентраций носителей заряда в инверсионном слое у выводов стока и истока возникает диффузионный ток, называемый подпороговым. Подпороговые токи утечки

экспоненциально зависят от напряжений  $V_{gs}$ ,  $V_{ds}$  и  $V_{TH}$  через параметры концентрации носителей:

$$I_{sub} = \mu_0 C_{ox} \frac{1}{L_{eff}} (n-1) \phi_T^2 \times e^{\frac{V_{gs}-V_{TH}}{n\phi_T}} \times (1 - e^{-\frac{V_{ds}}{\phi_T}}),$$

где  $\mu_0$  – подвижность при нулевом смещении,  $C_{ox}$  – емкость подзатворного окисла,  $L_{eff}$  – эффективная длина канала,  $n$  – коэффициент влияния подложки.

Бороться со статическим током утечки требует его рост при увеличении рабочей температуры кристалла (таблица 1). Учет токов утечек через затвор и подпороговых токов в равной степени важен и при комнатной температуре. На более высоких температурах составляющая подпороговых токов утечек становится преобладающей в общем балансе статического энергопотребления, так как рост температуры снижает пороговое напряжение транзисторов, приводя к деградации его крутизны [3].

Таблица 1

Компоненты токов утечек

Компоненты токов утечек	Температурная зависимость
Подпороговые токи утечки	2x/100°C
Токи утечки через затвор	8-12x/100°C
Токи утечки через p-n переход	50-100x/100°C

### III. СНИЖЕНИЕ СТАТИЧЕСКОЙ МОЩНОСТИ НА ОСНОВЕ МЕТОДА ОТКЛЮЧЕНИЯ ПИТАНИЯ

В последнее время активно развиваются схемотехнические решения, обеспечивающие возможность более раннего влияния на конечное энергопотребление всего проекта [4].

Метод отключения схемы от шин земли-питания является широко используемым в промышленности методом уменьшения токов утечек. Он включен в проекты ползуказных СБИС на стандартных ячейках, разрабатываемых с помощью коммерческих программных систем проектирования в кооперации с полупроводниковыми фабриками (TSMC), и принят в качестве промышленного стандарта.

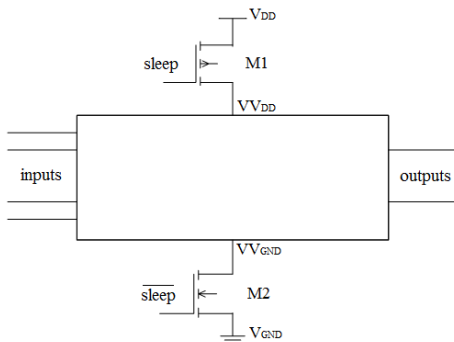


Рис. 1. Схема метода отключения питания с множественным пороговым напряжением

На рисунке 2 представлена базовая схема метода отключения питания с множественным пороговым напряжением (МТСМОС). Ключи на транзисторах M1 и M2 в этом подходе принадлежат всем вентилям схемы. Вентилю, в свою очередь, подключаются к виртуальным шинам  $VV_{DD}$  и  $VV_{GND}$  вместо настоящих шин  $V_{DD}$  и  $V_{GND}$ . Низкое пороговое напряжение используется внутри схемы для достижения максимальной производительности. Транзисторы с высоким пороговым напряжением используются в ключах с целью минимизации подпорогового тока утечки. В активном режиме сигнал sleep держится на низком логическом уровне, транзисторы M1 и M2 открыты и значения напряжений  $VV_{DD}$  и  $VV_{GND}$  близки к  $V_{DD}$  и  $V_{GND}$ , соответственно. В режиме сна сигнал sleep держится на высоком логическом уровне, транзисторы M1 и M2 закрыты и значения напряжений  $VV_{DD}$  и  $VV_{GND}$  имеют «плавающие» значения, а ток утечки схемы, содержащей транзисторы с низким пороговым напряжением, ограничивается ключом с высоким пороговым напряжением [4].

В настоящее время МТСМОС часто называется «power gating», а для простоты используется только один ключевой транзистор: либо верхний M1, либо нижний M2. На рисунке 2 схематично показана последовательность выключения и пробуждения схемного блока с р-МОП отключающим транзистором. Схема контролируется управляющим сигналом sleep. На рисунке показаны временные диаграммы токов в режиме сна и активном режиме. Последовательность операций разделена на следующие фазы:

1. Первая активная фаза
2. Фаза отключения
3. Фаза сна
4. Фаза пробуждения
5. Вторая активная фаза

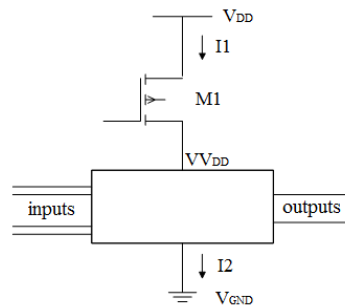


Рис. 2. Схема функционирования метода отключения питания с р-МОП отключающим транзистором

#### А. Активный режим работы

В активном режиме отключающий транзистор находится в проводящем состоянии, при этом между его стоком и истоком возникает небольшое падение напряжения  $V_{DROP}$ . Оно возникает из-за ненулевого сопротивления  $R_S$  открытого р-МОП транзистора. Последовательное соединение отключающего

транзистора и схемы образует делитель напряжения, у которого одно сопротивление мало, а другое – достаточно велико. Оба сопротивления изменяются с течением времени из-за смены значений сигналов на входах логического блока. Как следствие, эффективное напряжение питания схемы уменьшается. Часто в литературе оно называется виртуальным напряжением питания  $VV_{DD}$ . При использовании n-канального МОП транзистора вводится понятие виртуального напряжения земли  $VV_{GND}$ . Падение напряжения  $V_{DROP}$  на отключающем транзисторе можно определить как  $V_{DROP} = V_{DD} - VV_{DD}$  для p-МОП транзистора и  $V_{DROP} = VV_{GND}$  для n-МОП транзистора.

В обоих случаях эффективное напряжение питания для блока определяется как:

$$V_{DD}^{eff}(t) = V_{DD} - V_{DROP}(t).$$

Важным следствием является увеличение задержки схемы, так как задержка является функцией напряжения питания, которое, в свою очередь, уменьшается. На рисунке 2 представлены два различных сценария активного режима работы. При первом активном сценарии данные на входах логического блока фиксированы. Напряжение виртуальной шины  $VV_{DD}$ , как и ток утечки в активном состоянии, не изменяется. Во втором активном сценарии сигналы блока изменяют свои значения до тех пор, пока не достигается стабильное состояние первого активного сценария. Генерируемый в этом режиме ток влияет на границы значений  $VV_{DD}$  и  $V_{DROP}$ .

#### В. Временная фаза отключения блока

От начала момента активации ключа напряжение  $VV_{DD}$  падает до тех пор, пока не достигнет напряжения, близкого к  $V_{GND}$ . В то же самое время токи  $I_1$  и  $I_2$  уменьшаются до тока  $I_{off}$ , который ограничен компонентой подпорогового тока утечки через отключающий транзистор. Токи  $I_1$  и  $I_2$  уменьшаются с разной скоростью. В то время как  $I_1$  резко падает в момент достижения сигналом sleep значения  $V_{TH}$ , ток  $I_2$  медленно уменьшается и разряжает емкости блока. В течение этой фазы значения выходных сигналов теряются, если не используются какие-либо методы его сохранения. Время, требуемое на отключение схемы, обозначено как  $t_{powerdown}$ .

#### С. Фаза сна

В режиме сна ток  $I_1 = I_{sleep}$  определяется только токами утечки и зависит от значения напряжения насыщения при напряжении  $VV_{DD}$ , которое, в свою

очередь, зависит от отношения сопротивлений отключающего транзистора и логического блока. Как следствие,  $I_{sleep}$  является функцией нескольких параметров, таких как: технология изготовления транзисторов, типа самой отключаемой схемы, значений входных сигналов схемы, параметров отключаемого транзистора ( $W, L, V_{TH}$ ), рабочей температуры. Ток  $I_2$  начинает убывать при немного более высоком значении, нежели ток  $I_1$ . Это происходит при протекании тока утечки через затвор тех транзисторов, которые управляются высоким логическим уровнем сигнала.

#### Д. Временная фаза пробуждения

В самом начале этой фазы сигнал управления отключающим транзистором меняет свой логический уровень на противоположный (появляется управляющий фронт), что приводит к переходу этого транзистора к проводящему состоянию. Через него начинают заряжаться емкости схемы. Время пробуждения, необходимое для достижения стабильного состояния со значением напряжения виртуальной земли, близким к  $VV_{DD}$ , определено как  $t_{wakeUp}$ . При пробуждении, максимальный рабочий ток  $I_{ON}$  ограничен максимальным током через отключающий транзистор. Общую энергию, необходимую для зарядки емкостей, можно определить, как:

$$E_{sw} = \int_{t_{wakeUp}} V_{DD} \cdot I_1(t).$$

Для n-МОП отключающего транзистора данная формула имеет некоторые отличия, но предыдущие заключения являются схожими. Сигнал отключения блока также должен быть инверсным.

В активном режиме напряжение питания схемы близко к напряжению  $V_{DD}$  за вычетом небольшого падения напряжения на отключающем транзисторе. Схемный узел между транзистором и схемой (виртуальная земля) имеет потенциал  $V_{GND}$ . При отключении напряжения в схеме продолжают протекать остаточные токи. Если обозначить  $I_1$  как ток, текущий через отключающий транзистор, то выражение  $E_{sw}$  можно применить как для n-МОП, так и для p-МОП отключающего транзистора.

#### Е. Результативность применения метода отключения питания

Если фаза сна у схемы обладает достаточно большой временной длительностью, то метод отключения питания является эффективным. Определив энергию, затрачиваемую блоком на процесс смены состояния по формуле для  $E_{sw}$ , а также зная значения токов утечек в активном режиме  $I_{ACTIVE}$  и в

режиме сна  $I_{SLEEP}$ , становится возможным расчет необходимого времени бездействия  $t_{be}$  блока. Но такой расчет будет неточным. К примеру, р-МОП или п-МОП отключающие транзисторы, необходимые интерфейсные схемы, буферы обладают внутренним током утечки и забирают от источника некоторое количество энергии в процессе смены состояния схемы или блока. В любом случае, общее выражение для времени  $t_{be}$  записывается через параметры  $E_{OVERHEAD}$  и  $I_{OVERHEAD}$ , которые отвечают за все дополнительные энергии и токи, соответственно:

$$t_{be} = \frac{E_{sw} + E_{OVERHEAD}}{(I_{ACTIVE} - (I_{OFF} + I_{OVERHEAD})) \cdot V_{DD}}$$

На рисунке 3 представлены данные, полученные на основе исследования сумматора с последовательным переносом. С помощью программы схемотехнического моделирования Synopsys HSPICE [5]-[6] и доступной учебной технологии разработки схем с топологическими нормами 90нм были рассчитаны значения параметра энергии переключения – задержка схемотехнических реализаций схемы с наличием отключающих транзисторов.

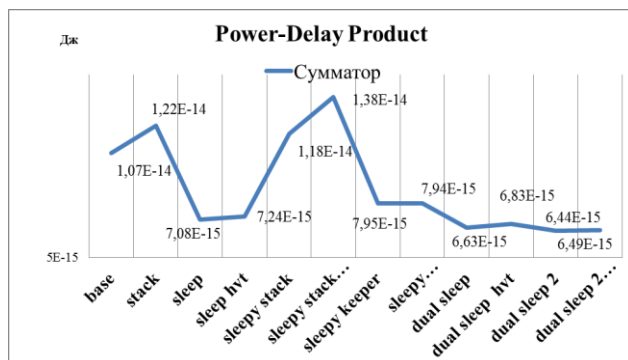


Рис. 3. Данные исследования энергетической характеристики PDP для схем 4-битного сумматора на основе метода отключения питания

Важный вывод состоит в том, что реализация на основе одного отключающего транзистора (sleep и sleep hvt) позволяет обеспечить хорошую энергоэффективность, что и обуславливает популярность метода.

#### IV. КОНТРОЛЬ БЫСТРОДЕЙСТВИЯ В СХЕМАХ С ОТКЛЮЧЕНИЕМ ПИТАНИЯ

Контроль быстродействия в схемах с отключением питания осуществляется путем задания размеров ключевой структуры. При этом затрагиваются такие параметры, как занимаемая площадь, скорость работы отключаемого блока и энергия переключения. Наиболее простым подходом служит суммирование ширин всех транзисторов в схеме и выбор требуемой ширины с помощью таблицы соответствия, но такой подход не учитывает падение напряжения на ключе, вследствие этого контроль быстродействия

логического блока будет некорректным из-за сильно завышенной оценки. Более сложные методы определения ширины отключающих транзисторов можно разделить на два класса:

- методы, основанные на определении пикового тока;
- методы, основанные на определении среднего тока.

Основная идея класса методов с определением пикового тока состоит в том, чтобы оценить максимальный мгновенный ток, протекающий через ключевой транзистор в процессе переключения входов схемы, т.е. ток активного режима работы. Такая оценка относится к наихудшему случаю поведения схемы. При этом требуется провести анализ по всем входным векторам, что является NP-полной задачей, поэтому данный метод неприменим для схем с большим количеством входов. Другой возможностью избежать полного перебора входных векторов при определении пикового тока служит введение различного рода эвристик или проведение статического временного анализа. Преимущество таких подходов в скорости оценки тока при незначительном ухудшении точности.

Второй класс методов работает с усредненными значениями токов. К примеру, алгоритмы задания размеров основаны на некотором ожидаемом значении тока, полученном с помощью параметра вероятности переключения входов схемы. Задание размеров ключевых транзисторов рассматривается как проблема распределения задержки. Такие подходы приводят к недооценке размеров ключевых транзисторов.

Промышленные программы (например, IC Compiler компании Synopsys) позволяют на этапе физического синтеза оценить размеры добавляемых транзисторов. Но доступные для синтеза транзисторные ключи обладают большими размерами и используются в основном для работы с большими логическими кластерами. Кроме того, отсутствует точный расчет временных характеристик таких схем, в частности, влияние добавляемых транзисторов на быстродействие.

В данной работе предлагается маршрут, позволяющий с достаточной степенью точности рассчитать задержки базовых логических элементов и СФ-блоков, использующих отключающие структуры.

Для этого предлагается точная оценка значения возможного тока, протекающего через СФ-блоки в активном режиме работы КМОП схемы. Наряду с оценкой тока происходит определение привносимой отключающими структурами задержки.

Предположим, что деградация задержки в схеме с отключающей структурой происходит при переключении выхода из 1 → 0 в случае п-МОП отключающего транзистора или из 0 → 1 в случае р-МОП отключающего транзистора. При этом увеличивается потенциал виртуальной земли или уменьшается потенциал виртуального питания. Предлагаемый маршрут состоит из нескольких шагов:

1. На первом шаге для каждого sleep транзистора из некоторого набора проводится характеристика максимального пропускного тока и соответствующего ему напряжения виртуальной шины.

Выберем для описываемого маршрута p-МОП отключающий транзистор (рассуждения при выборе n-МОП транзистора аналогичны). Для проведения характеристики используется схема эквивалентной нагрузки, которая подключается к терминалу виртуального питания. Она состоит из набора инверторов, переключающихся одновременно, и тем самым позволяет эмулировать пиковый нагрузочный ток (рис. 4).

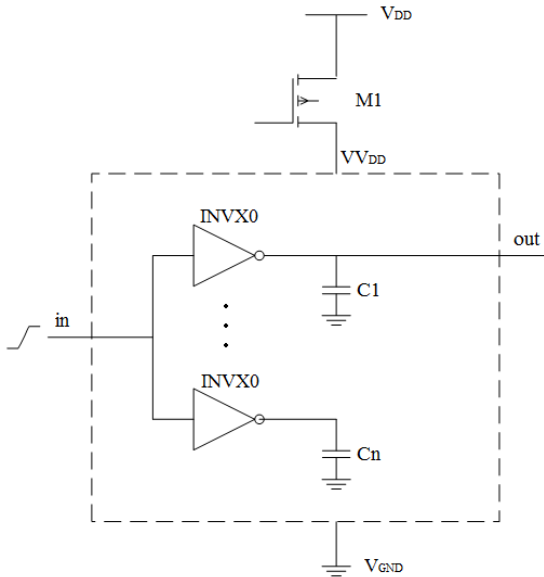


Рис. 4. Схема для характеристики максимального пропускного тока отключающего транзистора

2. На втором шаге анализируются полученные результаты для каждого размера отключающего транзистора. Данные по зависимости между током и напряжением на виртуальной шине ограничиваются максимально возможным допустимым значением падения напряжения. На рисунке 5 показана зависимость напряжения виртуальной шины питания от значения тока, проходящего через отключающий транзистор, для различных его размеров. К примеру, нижняя кривая определяет значение потенциала виртуальной шины для отключающего транзистора минимального размера.

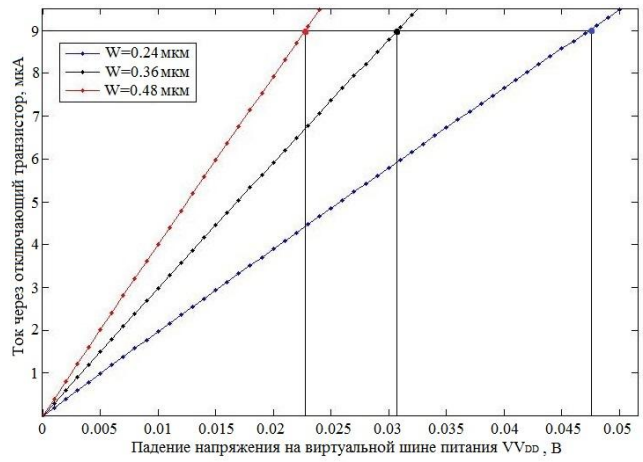


Рис. 5. График зависимости  $I_{\max}$  и  $VV_{DD}$  отключающего транзистора от параметра  $W$

3. На третьем шаге для данного виртуального напряжения требуется оценить задержку распространения сигнала для логических блоков, подключенных к виртуальным шинам, а также пиковый ток, протекающий в схеме. В случае p-МОП транзистора задержка распространения сигнала из 0- >1, влияющая на верхнюю границу рабочего напряжения питания, зависит от нескольких переменных, таких как время нарастания/спада фронта сигнала на входе логического блока, выходной емкостной нагрузки и значения напряжения питания. На рисунке 6 показана тестовая схема для проведения характеристики базовых ячеек с отключающими транзисторами при изменении значений входного фронта, выходной нагрузочной емкости и потенциала виртуальной шины питания. Для эмуляции падения напряжения на шине питания используется идеальный источник постоянного напряжения.

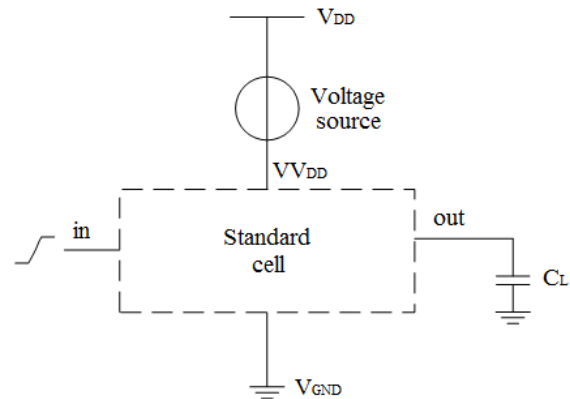


Рис. 6. Схема характеристики задержки прохождения сигнала для базовых элементов

В качестве примера представлена таблица нормализованной максимальной задержки распространения сигналов стандартных вентилях при изменении потенциала виртуальной шины. Полученные значения деградации задержек записываются в отдельную таблицу.

Таблица 2

Нормализованная задержка исследуемых базовых элементов и СФ-блоков

Падение напряжения на виртуальной шине $V_{DD}$ , мВ	Процент деградации задержки, %		
	NAND2X0	C17	C432
5	0,76	0,50	0,58
10	1,53	0,84	1,13
20	3,07	1,78	2,29
40	6,20	3,68	4,65
80	11,16	7,26	9,67
100	14,56	9,07	12,30

4. На четвертом шаге производится оценка пикового тока базовых логических вентилях или СФ блоков с учетом деградации напряжения на виртуальной шине. Применяемый метод изложен в статье [7]. Проведенные численные эксперименты по оценке пикового тока показали, что данный подход обеспечивает расчет пикового тока на уровне логического моделирования с точностью в пределах 7% по сравнению со схемотехническим моделированием.

Графики зависимости значения пикового тока от напряжения виртуальной шины представлены на рисунке 7.

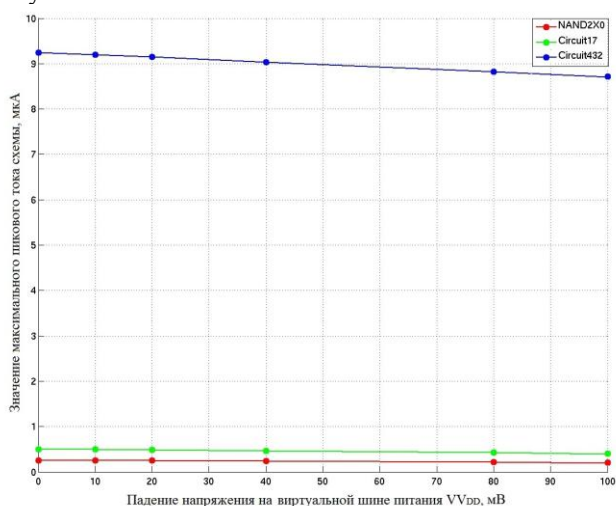


Рис. 7. Данные зависимости измеренного пикового тока исследуемых схем от падения напряжения на виртуальной шине питания  $V_{DD}$

5. На пятом шаге происходит выбор итогового значения ширины отключающего транзистора с использованием данных по его максимальному пропускному току, а также полученному значению пикового тока отключаемой КМОП схемы в активном режиме ее работы.

Для схемы C432 из набора ISCAS 85 результаты показывают (рис. 7), что при заданном падении напряжения на виртуальной шине питания в 5 мВ

пиковый ток схемы равен 9,15 мкА, а процент деградации задержки при этом будет равен 0,58%. По данным характеристики отключающих транзисторов (рис. 7) таким параметрам удовлетворяет транзисторный р-МОП ключ с шириной канала, равной 0,24 мкм.

## V. ПРИМЕНЕНИЕ РАЗРАБОТАННОГО АЛГОРИТМА КОНТРОЛЯ БЫСТРОДЕЙСТВИЯ

Для заданного пикового тока схемы существует множество вариантов выбора ширин отключающих транзисторов. Выбор той или иной структуры разработчик делает сам, исходя из требуемых параметров, например, быстродействия.

Предложенный маршрут может быть реализован в программном виде в составе программного комплекса временного анализа. Эффективность предложенного подхода оценивалась по результатам точного схемотехнического анализа схем.

## VI. ЗАКЛЮЧЕНИЕ

Предложенный маршрут анализа быстродействия базовых логических элементов и СФ-блоков на основе метода отключения питания обеспечивает снижение статической мощности КМОП схем на основе отключающих транзисторов с контролем быстродействия в активном режиме работы схемы. Контроль быстродействия обеспечивает компромисс в выборе минимальных размеров отключающих транзисторов с сохранением ограничений на период и частоту сигналов в КМОП схеме.

## ЛИТЕРАТУРА

- [1] Chandrakasan A., Sheng S., Brodersen R. Low-Power CMOS Digital Design // IEEE Journal of Solid-State Circuits. 1992. V. 27. № 4. P. 473–484.
- [2] Kaijian Shi and David Howard. Challenges in sleep transistor design and implementation in low-power designs // Proc. Of the 3rd annual conference on Design automation. 2006. P. 113–116.
- [3] Vaibhav Neema, Shailesh Singh Chouhan and Sanjiv Tokekar. Novel Circuit Technique for Reduction of Leakage Current in Series/Parallel PMOS/NMOS Transistor Stack // IETE Journal of Research. November-December 2010. Vol. 56.
- [4] Benton H. Calhoun, Frank A. Honore, and Anantha Chandrakasan. Design methodology for fine-grained leakage control in MTCMOS // In Proceedings of the 2003 international symposium on Low power electronics and design. ACM Press. 2003. P. 104-109.
- [5] HSPICE User Guide: Simulation and Analysis. // Release 2011.09, September 2011.
- [6] Synopsys Inc. Synopsys products. <http://www.synopsys.com/>. // March 2014.
- [7] Рыжова Д.И., Гаврилов С.В., Щелоков А.Н. Анализ пикового тока на основе результатов характеристики реальных библиотек логических вентилях // Тр. Международного конгресса по интеллектуальным системам и информационным технологиям – 2013. “IS&IT’13”. С. 251-252.