

Разработка быстродействующего блока памяти с ассоциативной выборкой

Ю.М. Потовин¹, С.А. Соин²

¹Национальный исследовательский университет «МИЭТ»

²ОАО «ИНЭУМ им. И.С. Брука», yurapotovin@mail.ru

Аннотация — Рассматривается блок памяти с ассоциативной выборкой. Описываются модернизации схем, позволяющие улучшить характеристики. Блок реализован в соответствии с нормами и правилами для технологии 28нм фирмы TSMC и может быть использован в последующих проектах.

Ключевые слова — ассоциативная память, статическая память, снижение мощности, ячейка, высокоскоростной, помехоустойчивость.

I. ВВЕДЕНИЕ

Блок содержит 64 регистра, каждый из которых состоит из 3 частей: ассоциативной (адресной), числовой и бита значимости. В режиме опроса происходит параллельное сравнение содержимого адресных частей всех регистров с адресом, пришедшим на соответствующий вход блока. При совпадении адресов в каком-либо регистре и установленном в нем бите значимости, происходит считывание числовой части этого регистра и выдача ее на выход блока. Так же на выход подается результат сравнения – есть совпадение или нет.

Запись в блок производится по номеру регистра и затрагивает все 3 его части.

Предусмотрено 2 режима обнуления битов значимости: одновременное их обнуление во всех регистрах, или же обнуление бита значимости в регистре, выбранном по результату сравнения адресных частей с входящим адресом.

Широкое применение устройств, включающих ассоциативную выборку, ограничено тем, что дополнительная схема сравнения увеличивает физический размер и рассеивает относительно большую мощность. В работе предложены решения, позволяющие улучшить характеристики, и существенно уменьшить мощность ассоциативного опроса.

II. КЛАССИЧЕСКИЙ ВАРИАНТ ИСПОЛНЕНИЯ АССОЦИАТИВНОЙ ЧАСТИ

На рис. 1 показана упрощенная электрическая схема классической ассоциативной памяти (рис. 1а) [1]. Основным элементом этой памяти (ЭАП) является 10-транзисторная ячейка (рис. 1б), состоящая из 6-

транзисторной ячейки оперативной памяти (6Т) и схемы сравнения на четырех транзисторах. Схема сравнения представляет собой две пары последовательно соединенных транзисторов.

Выходы схем сравнения одного слова объединены при помощи шины совпадения (match line). Шины, объединяющие входы данных одноименных разрядов слов массива, носят название «шины поиска» (search lines). Заметим, что на шины поиска подается инвертированное слово, в силу чего при совпадении исходного слова с записанным ранее проводящий путь между шиной совпадения и землей отсутствует. Информация о результате сравнения поступает на выходы ассоциативной памяти Q0...QN.

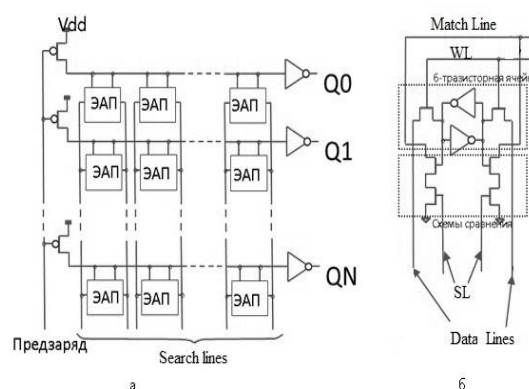


Рис. 1. Обобщенная структура ассоциативной памяти (а – классическая организация ассоциативной памяти, б – основной ЭАП)

Временная диаграмма работы этой схемы состоит из двух фаз. На фазе предварительного заряда (предзаряда) осуществляется заряд шин совпадения до напряжения питания через р-канальные транзисторы (рис. 1а). На выходах устанавливается низкий уровень.

На фазе опроса предзаряд отключается, а на шины поиска подаются данные для сравнения. В случае совпадения этих данных с ранее записанным словом потенциал соответствующей шины совпадения не изменяется. Для всех слов, в которых имеется несовпадение хотя бы в одном разряде, появляется ток, разряжающий шины совпадения до потенциала «земли».

Максимальная задержка будет в случае, если не совпадает только один разряд. Традиционным способом уменьшения задержки является увеличение транзисторов в схемах сравнения. Однако это приводит к увеличению мощности, затрачиваемой на перезаряд шин совпадения. Для исключения сквозных токов до начала предзаряда на всех шинах поиска должен быть установлен нулевой потенциал. Именно эти две составляющие определяют мощность, рассеиваемую устройством, и ограничивают допустимое количество слов и количество разрядов в каждом слове.

III. ИСПОЛЪЗУЕМЫЙ ВАРИАНТ АССОЦИАТИВНОЙ ЧАСТИ

В данном блоке используется схемотехническое и топологическое решение ассоциативной памяти. На рис. 2 приведена электрическая схема одного слова такой памяти.

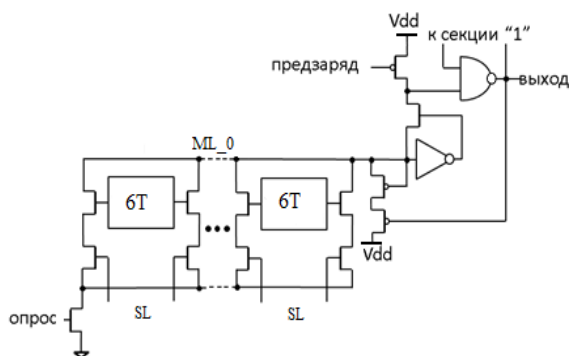


Рис. 2. Электрическая схема одного слова ассоциативной памяти

Для классической схемы в фазе предзаряда все шины совпадения заряжаются до напряжения питания V_{dd} . В фазе опроса – все шины, кроме той, где имело место совпадение, разряжаются до уровня «земли». Таким образом, энергия, затрачиваемая на перезаряд шин совпадения, равна $(N-1) * C_{ml} * V_{dd}^2$, где N – количество шин совпадения, C_{ml} – емкость шины совпадения.

В предложенном варианте предзаряд осуществляется до напряжения переключения выходного инвертора, которое не превышает $V_{dd}/2$. Следовательно, энергия, затрачиваемая на перезаряд шин совпадения, будет в четыре раза меньше.

Существенно, что в случае несовпадения только в одном разряде, ток, разряжающий емкость шины совпадения, должен понизить потенциал этой шины на величину порядка зоны переключения выходного инвертора, что уменьшает задержку выходного сигнала. С целью дальнейшего уменьшения задержки шина совпадения разделена на две секции, которые объединены через схему «И-НЕ».

Рассмотрим процесс перезаряда шин поиска. Для исключения сквозных токов во время предзаряда шин совпадения в классической схеме все шины поиска должны иметь потенциал земли. Во время опроса

половина этих шин заряжается до напряжения V_{dd} . В предложенном варианте схемы сравнения подключаются к «земле» через ключ, управляемый сигналом опроса, тем самым исключается необходимость предварительной зарядки шин поиска, следовательно, за такт шины поиска заряжаются только один раз. В среднем будет перезаряжаться половина шин поиска. В результате средняя энергия, затрачиваемая на перезаряд шин поиска в предлагаемой схеме, будет в 4 раза меньше.

Таким образом, данная реализация ассоциативной памяти в режиме ассоциативного опроса рассеивает в четыре раза меньше мощности по сравнению с классической схемой.

IV. ЧИСЛОВАЯ ЧАСТЬ

Классическая схема основана на 6-ти транзисторной ячейке статической памяти (6Т) и представлена на рис. 3 [2].

Линия WL (Word Line) управляет двумя транзисторами доступа. Линии BL, BLB (Bit Line) – битовые линии, используются и для записи данных, и для чтения данных. Помимо того, что эти операции должны быть разнесены во времени, возникает проблема соотношения ширины транзисторов. С одной стороны, транзисторы M1 и M2 должны быть «больше» транзисторов M5 и M6, в противном случае, во время считывания триггер может переключиться. Для записи же требуется обратное.

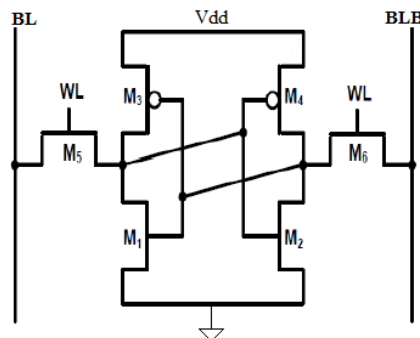


Рис. 3. Электрическая схема 6-ти транзисторной ячейки статической памяти

Рассмотрим статический запас помехоустойчивости при считывании (SRAM Read Static Noise Margin(SNM)) таких ячеек[3]-[5].

На рис. 4 представлен способ определения SNM.

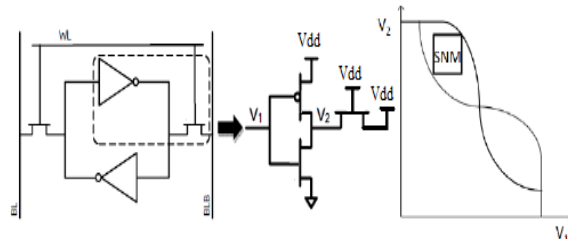


Рис. 4. Схема моделирования процесса считывания

Для номинального напряжения питания 0.9V при уменьшении соотношения транзисторов M6 к M2 видно, что SNM увеличивается (рис. 5), однако увеличивается и площадь ячейки.

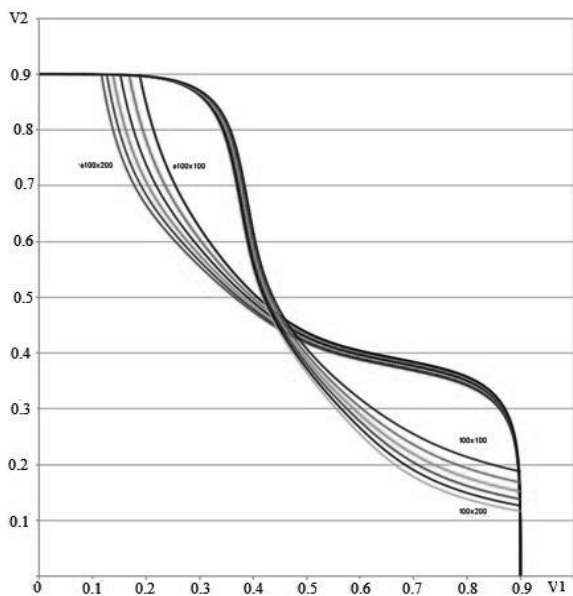


Рис. 5. Результат моделирования 6Т при номинальном напряжении питания 0.9V

КМОП схемы должны нормально функционировать при снижении питания до 2/3 от номинала, однако уменьшив питание до 0.6V, видно, что SNM становится очень малым (рис. 6).

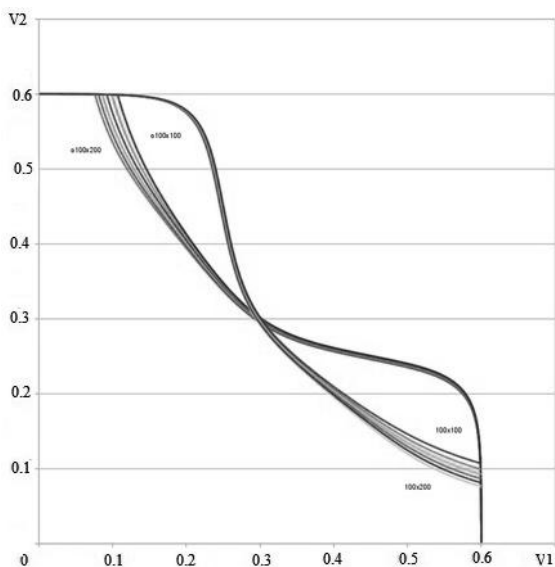


Рис. 6. Результат моделирования 6Т при уменьшении напряжении питания до 0.6V

Решением этих проблем является использование 8-ми транзисторной ячейки (8Т) (рис.7), в которой операции записи и чтения разделены. Временная диаграмма состоит из двух фаз. На фазе предзаряда осуществляется заряд битовых линий (RBL). На фазе считывания предзаряд отключается и подается сигнал на линию RWL (Read Word Line). В зависимости от

того, что хранилось в ячейке, на RBL сохраняется потенциал, либо появляется ток, разряжающий до потенциала земли. Существенно, что одновременно по разным адресам может производиться и запись и считывание.

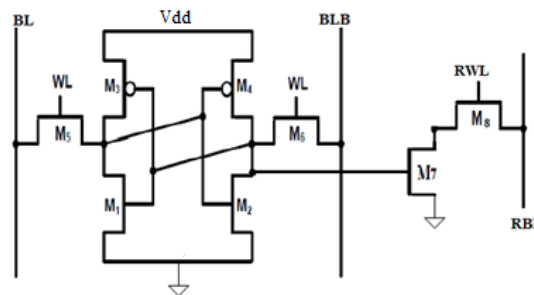


Рис. 7. Электрическая схема 8-ми транзисторной ячейки памяти

Аналогично промоделировав эту ячейку видно, что даже при уменьшенном питании и увеличении соотношения транзисторов M6 к M2, SNM остается высоким (рис.8).

Прибавка в площади компенсируется значительно лучшими временными характеристиками, а также возможностью использования такого принципа для реализации многопортовых памяти, например кэша первого уровня L1 [6].

Для числовой части данного блока были использованы именно 8Т ячейки.

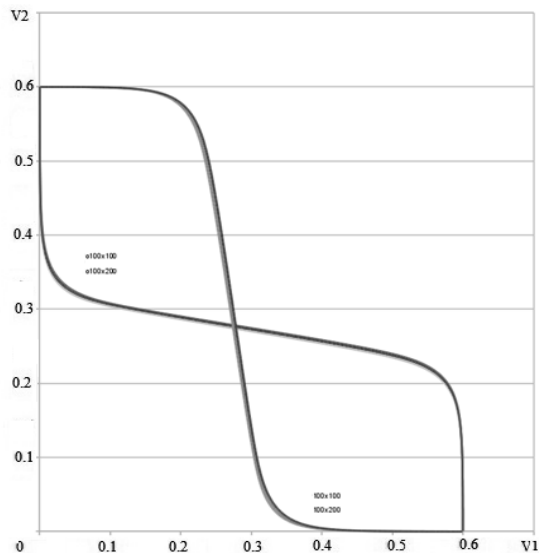


Рис. 8. Результат моделирования 8Т при напряжении питания 0.6V

V. ТОПОЛОГИЧЕСКАЯ РЕАЛИЗАЦИЯ

Топология блока (рис. 9) разработана в строгом соответствии с топологическими нормами и правилами для технологии 28nm фирмы TSMC, процесс НРМ.

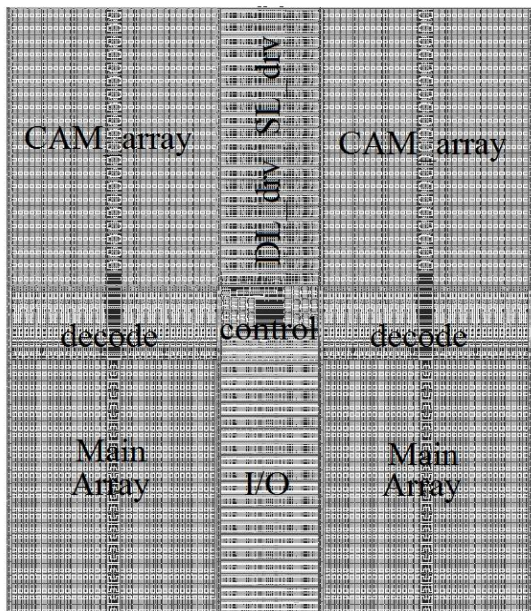


Рис. 9. Общая топология блока

Направление шин: WL, ML, RWL - вертикальное, параллельно затворам, SL, DL, BL, BLB — горизонтальное.

VI. ИССЛЕДОВАНИЕ ХАРАКТЕРИСТИК РАЗРАБОТАННОГО БЛОКА

Для исследования характеристик блока использовалась программа симуляции электронных схем FineSim (Программа фирмы Synopsys типа HSPICE). Наличие топологии позволило экстрагировать модель блока с учетом всех паразитных емкостей и сопротивлений.

Пример временной диаграммы, использованной для моделирования, приведен на рис.10.

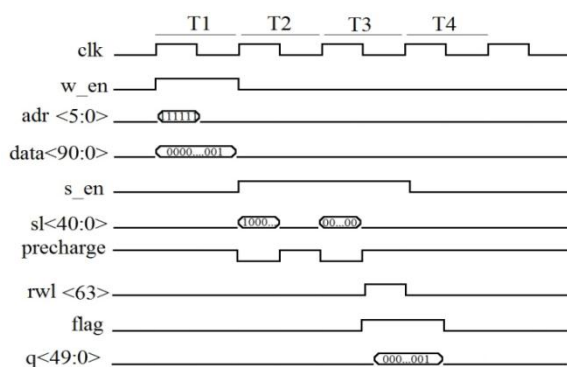


Рис. 10. Временная диаграмма

Моделирование проводилось с учётом возможных технологических разбросов параметров транзисторов. Для этого использовались следующие модели транзисторов: tt (типичные характеристики), ff («быстрые» n- и p-канальные), ss («медленные» n- и p-канальные). Блок устойчиво работает при всех

сочетаниях параметров транзисторов и отклонении напряжении питания на 10%.

Основные характеристики блока представлены в табл. 1.

Таблица 1

Характеристики блока

| Кол-во регистров | Разрядность (адрес /число) | Размеры, мкм | Частота (в углт), ГГц | Кол-во слоёв металла | Мощность, мВт |
|------------------|----------------------------|--------------|-----------------------|----------------------|---------------|
| 64 | 41/50 | 70x85 | 1,5 | 5 | 11 |

VII. ВЫВОДЫ

Уменьшение мощности ассоциативной части блока в 4 раза, использование 8 транзисторной ячейки для числовой части, которое, в свою очередь, позволяет увеличить помехоустойчивость за счёт разделения шин для записи и чтения данных, избавиться от усилителей считывания, используемых при классической реализации, что увеличивает общую надежность блока, повысить быстродействие - все это позволяет использовать данный блок в буфере ассоциативной трансляции (Translation lookaside buffer, TLB). TLB - специализированный кэш центрального процессора, используемый для ускорения трансляции адреса виртуальной памяти в адрес физической памяти. Каждая запись содержит соответствие адреса страницы виртуальной памяти адресу физической памяти. Если адрес отсутствует в TLB, процессор обходит таблицы страниц и сохраняет полученный адрес в TLB, что занимает в 10-60 раз больше времени, чем получение адреса из записи, уже закешированной TLB.

ЛИТЕРАТУРА

- [1] Кристовский Г.В., Погребной Ю.Л., Соин С.А. Разработка быстродействующих маломощных блоков ассоциативной памяти для микропроцессора «Эльбрус-4С+» // Вопросы радиоэлектроники. Сер. ЭВТ. 2013. Вып. 3. С. 169–177.
- [2] Ajay Gadhe, Ujwal Shirode Read stability and write ability analysis of different SRAM cell structures // International Journal of Engineering Research and Applications (IJERA). January-February, 2013. Vol. 3. Issue 1. P. 1073-1078.
- [3] Lecture 10: SRAM // Advanced Digital Integrated Circuits. EE241 — Spring 2011.
- [4] SRAM Circuit Design // Introduction to VLSI (361-1-3701).
- [5] Vishal Saxena. SRAM Static Characterization // Department of Electrical and Computer Engineering, Boise State University.
- [6] Gary S. Ditlow, Robert K. Montoye, Sherman M. Dance, Sebastian Ehrenreich, Bruce M. Fleischer, Thomas W. Fox, Kyle M. Holmes, 4R2W Register File for a 2.3GHz Wire-Speed POWER™ with Double-Pumped Write Operation // ISSCC 2011. Session 14. High-performance embedded memory. 14.2. P. 256-258.