

АССОЦИАТИВНАЯ ПАМЯТЬ: НОВОЕ - ХОРОШО ЗАБЫТОЕ СТАРОЕ

Кафтанников Игорь Леопольдович, к.т.н., доцент, Южно-Уральский государственный университет, кафедра ЭВМ

kil@is74.ru

В настоящее время активно продолжаются поиски путей увеличения производительности вычислительных средств различного назначения для решения современных задач, требующих гигантских вычислительных ресурсов с точки зрения традиционных архитектур. Возможные решения рассматриваются по многим направлениям:

- уменьшение размеров и параметров транзисторов,
- применение других электронных компонентов (мемристоры и т.п.),
- модификации традиционных архитектур и структур - конвейеризация на всех уровнях, расширение регистровой (SIMD) структуры универсальных процессоров с соответствующим расширением систем команд, разработка многоядерных архитектур микропроцессоров,
- разработка распределенных и суперкомпьютерных архитектур,
- использование новых физических принципов, процессов и свойств для обработки и хранения информации,
- разработка нетрадиционных архитектур - нейросетевые, перестраиваемые и т.п.

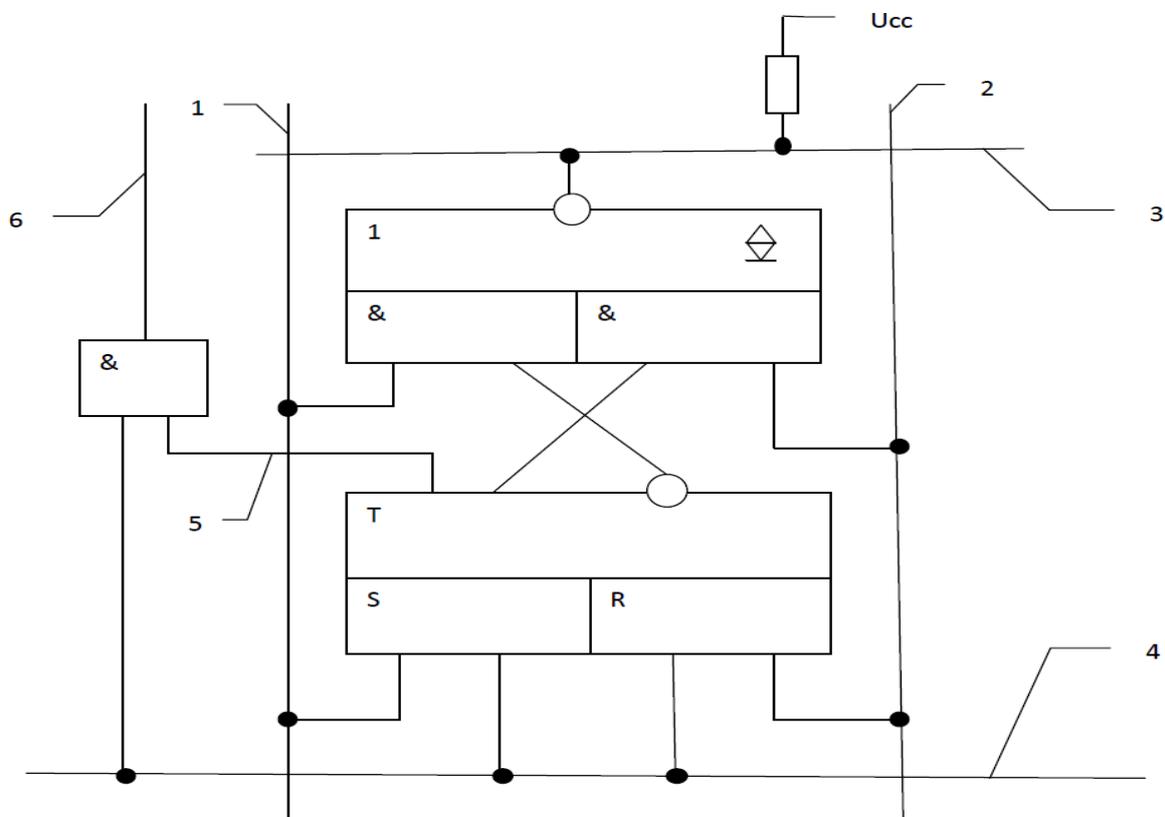
Следует также отметить, что существенно расширились типы и форматы данных, и, соответственно, функциональность решений задач. Увеличилась доля нечисловой обработки, поисков, сравнений и т.п. Вследствие этого есть необходимость рассмотрения структур и архитектур, более производительных в данном плане.

Одним из таких направлений является универсальная вычислительная среда, называемая ассоциативной памятью (АП). В России (СССР) в 80-х годах осуществлялся ряд работ в этом направлении, в частности, в ИПУ АН СССР под руководством академика И.В.Прангишвили. Определенное участие в этих работах принимали и сотрудники Челябинского политехнического института (ныне

ЮУрГУ), К сожалению, отсутствие соответствующей технологической базы не позволило реализовать ряд интересных проектов.

Сейчас развитие новой элементной и технологической базы позволяет заново рассмотреть уже с современных представлений возможности АП.

Напомним, что АП относится к SIMD структурам, причем уникальной возможностью АП является возможность анализа и преобразования данных непосредственно в памяти и с различной степенью локализации: бит, строка, массив, группа массивов. Схематехническая структура ячейки АП представлена на рис.1, а набор микроопераций в табл.1.



- 1 - шина опроса/записи «1» (BQ1);
- 2 - шина опроса/записи «0» (BQ2);
- 3 - шина совпадения (BE);
- 4 - шина записи (BW);
- 5 - выдача данных(BD);
- 6 - шина считывания (BR);

Рисунок 1. Ячейка ассоциативной памяти.

Таблица1.

шина опроса/записи «1»	шина опроса/записи «0»	шина записи	операция
0	0	0	маскирование опроса
0	1	0	опрос по «0»
1	0	0	опрос по «1»

1	1	0	запрещено (блокировка сравнения)
0	0	1	маскирование записи
0	1	1	Запись «0»
1	0	1	Запись «1»
1	1	1	запрещено

Набор микроопераций ячейки АП

Очевидно, что элементный размер ячейки больше, чем, например, ячейка SRAM, но применение мемристоров или МРАМ, по нашим представлениям может сократить объем и энергопотребление ячейки.

Структура АП (решающего поля) представлена на рис.2

СТРУКТУРА МОДУЛЯ АССОЦИАТИВНОЙ ПАМЯТИ

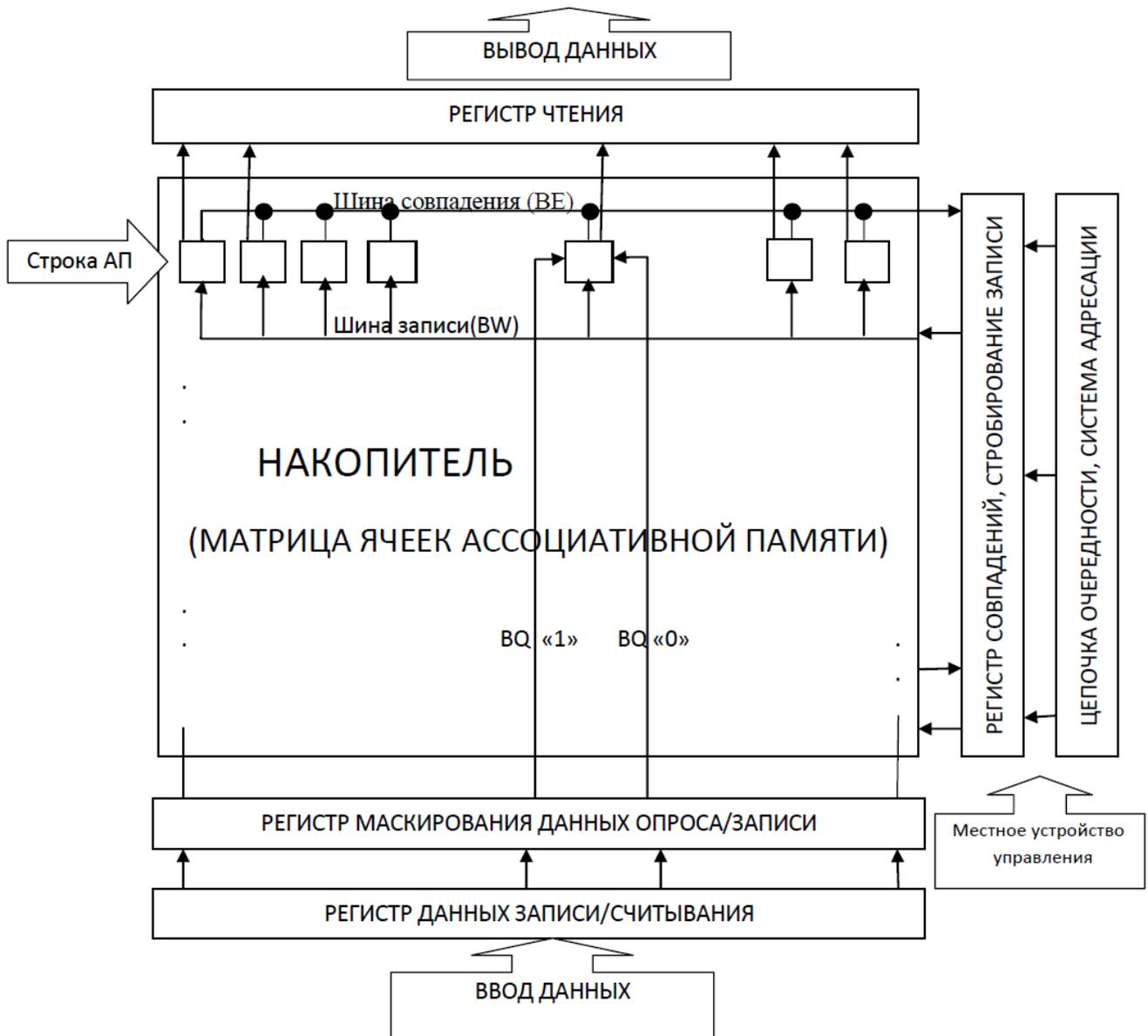


Рисунок 2.

результат $s(0)i := a(0)i + b(0)i$

	Операнд А						Операнд В						Результат S						служ раз					
	slices (разряды)																							
	15	14	13	...	2	1	0	15	14	13	...	2	1	0	15	14	13	...	2	1	0	2	1	0
Строка M-1	0	1	1	0	1	0	1	0	0	0	1	1	0	0	0	0	0	0	0	0	1	0	0	0
Строка M-2	0	1	0	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
Строка M-3	0	0	1	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Строка M-4	1	0	1	1	0	1	0	1	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	1
Строка M-5	1	1	1	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0
Строка M-6	0	0	0	1	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0
.																								
Строка 7	1	0	1	1	1	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Строка 6	1	1	0	0	0	0	0	1	0	1	1	1	0	0	0	0	0	0	0	0	1	0	0	0
Строка 5	0	0	1	1	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Строка 4	0	1	1	0	1	1	1	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	1
Строка 3	0	0	0	0	1	0	1	1	1	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0
Строка 2	0	1	0	1	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	1	0	0	0
Строка 1	1	0	1	1	1	1	0	0	1	1	1	0	0	0	0	0	0	0	0	0	1	0	0	0
Строка 0	1	0	0	0	0	1	0	1	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	1

сложение $s(1)i := a(1)i + b(1)i + p(1)i$

Опрос 1.1	регистр маски	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1				
	регистр данных	x	x	x	x	x	1	x	x	x	x	x	x	0	x	x	x	x	x	x	x	x	x	x	0			
Опрос 1.2	регистр маски	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1				
	регистр данных	x	x	x	x	x	0	x	x	x	x	x	x	1	x	x	x	x	x	x	x	x	x	x	0			
Опрос 1.3	регистр маски	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1				
	регистр данных	x	x	x	x	x	0	x	x	x	x	x	x	0	x	x	x	x	x	x	x	x	x	x	1			
Запись 1.1	регистр маски	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0			
	регистр данных	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	1	x	x	0	x			
Опрос 1.4	регистр маски	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1			
	регистр данных	x	x	x	x	x	0	x	x	x	x	x	x	1	x	x	x	x	x	x	x	x	x	x	1			
Опрос 1.5	регистр маски	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1			
	регистр данных	x	x	x	x	x	1	x	x	x	x	x	x	0	x	x	x	x	x	x	x	x	x	x	1			
Опрос 1.6	регистр маски	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1			
	регистр данных	x	x	x	x	x	1	x	x	x	x	x	x	1	x	x	x	x	x	x	x	x	x	x	x	0		
Запись 1.2	регистр маски	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0
	регистр данных	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	0	x	x	1	x

результат $s(1)i := a(1)i + b(1)i + p(1)i$																									
Операнд А						Операнд В						Результат S						служ раз							
slices (разряды)																									
	15	14	13	...	2	1	0	15	14	13	...	2	1	0	15	14	13	...	2	1	0	2	1	0	
Строка М-1	0	1	1		0	1	0	1	0	0		0	1	1	0	0	0		0	0	0	1	0	1	0
Строка М-2	0	1	0		1	1	1	1	1	0		0	0	0	0	0	0		0	0	0	1	1	0	0
Строка М-3	0	0	1		0	0	0	0	0	1		1	1	0	0	0	0		0	0	0	1	0	0	0
Строка М-4	1	0	1		1	0	1	0	1	1		0	1	1	0	0	0		0	0	0	0	0	0	1
Строка М-5	1	1	1		0	0	1	0	0	0		0	1	0	0	0	0		0	0	0	1	1	0	0
Строка М-6	0	0	0		1	0	1	0	1	0		1	0	0	0	0	0		0	0	0	0	1	0	0
.																								0	
.																								0	
.																								0	
Строка 7	1	0	1		1	1	0	0	1	1		0	1	0	0	0	0		0	0	0	0	0	0	1
Строка 6	1	1	0		0	0	0	0	1	0		1	1	1	0	0	0		0	0	0	1	1	0	0
Строка 5	0	0	1		1	1	0	0	0	1		0	0	0	0	0	0		0	0	0	1	0	0	0
Строка 4	0	1	1		0	1	1	1	0	1		1	0	1	0	0	0		0	0	0	0	0	0	1
Строка 3	0	0	0		0	1	0	1	1	1		0	0	1	0	0	0		0	0	0	1	1	0	0
Строка 2	0	1	0		1	0	0	0	0	0		1	0	1	0	0	0		0	0	0	0	1	0	0
Строка 1	1	0	1		1	1	1	0	0	1		1	1	0	0	0	0		0	0	0	0	1	0	1
Строка 0	1	0	0		0	0	1	0	1	1		0	1	1	0	0	0		0	0	0	0	0	0	1

Опросов 7, Записей 3

Из таблиц видно, что количество тактов, необходимых для сложения 16 битовых кодов – $15 \cdot 10 + 5 = 155$. Таким образом для получения преимущества в общем количестве тактов при покоординатном сложении 2-х векторов, без учета процедур ввода/вывода размерность векторов должна быть не менее 155, что существенно меньше возможностей современных, например, ПЛИС. Конечно, есть ряд решений, позволяющих сократить число тактов на разряд, но это требует более детальной проработки функциональности и структуры накопителя АП.

На каждый опрос нужна запись в служебный разряд и анализ: есть ли хотя бы одно совпадение. Реализация анализа может быть выполнена аппаратным образом, а суммарное число тактов операции зависит от n . При анализе всех разрядов потребуется $2n$ тактов для n разрядных операндов.

Для достижения высокой производительности необходимо учитывать особенности каждой операции, при этом часть особенностей реализовывать аппаратно в местном устройстве управления (МУУ) см. рис 1.

Более сложные и специализированные операции могут быть показаны при детальном рассмотрении возможностей АП.

Предварительный анализ показывает, что во многих существенных современных задачах (генный анализ, поиски новых структур и форм в фармацевтике, и других аналогичных направлениях, искусственный интеллект, и т.п.) применение АП может существенно сократить время их решения.

Следует также заметить, что создание модулей АП как компонентов или субмодулей высокопроизводительных систем может быть успешным только при комплексном решении задач по элементной базе, разработке соответствующей системы команд, языка и новых алгоритмов задач для реализации с применением АП. А, следовательно, необходим определенный уровень кооперации в этом направлении.