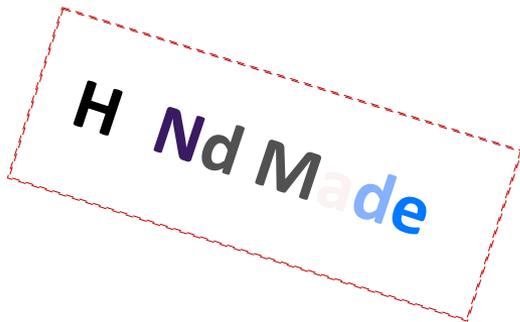


Классификация архитектур вычислительных систем



Автор - Никитин А.
А.

Введение

Общие принципы функционирования параллельных машин

- векторно-конвейерные
- массивно-параллельные
- компьютеры с широким командным словом
- систолические массивы
- гиперкубы
- спецпроцессоры
- мультипроцессоры
- иерархические
- кластерные компьютеры
- dataflow
- матричные ЭВМ

Наиболее важные параметры

- организация памяти
- топология связи между процессорами
- синхронность работы отдельных устройств
- способ исполнения арифметических операций

Основной вопрос - что заложить в основу классификации

- стоимость компьютера
- потоки данных и команд
- пословный и поразрядный параллелизм

Правильная классификация помогает:

- разобраться с тем, что представляет собой каждая архитектура
- понять как архитектуры взаимосвязаны между собой
- выяснить на какой класс архитектур ему следует ориентироваться для решения требуемого класса задач
- показать возможные пути совершенствования компьютеров

Классификация по мнению Скилликорна

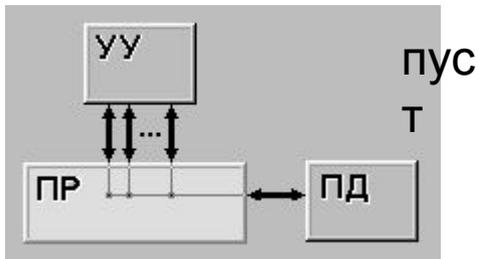
- **должна** способствовать пониманию того, что достигнуто на сегодняшний день в области архитектур вычислительных систем, и какие архитектуры имеют лучшие перспективы в будущем;
- подсказывать новые пути организации архитектур - речь идет о тех классах, которые в настоящее время по разным причинам пусты;
- показывать, за счет каких структурных особенностей достигается увеличение производительности различных вычислительных систем; с этой точки зрения, классификация может служить моделью для анализа производительности.

Классификация Флинна

dataflow и векторно-конвейерные машины пролетают

самая ранняя и наиболее известная классификация

Конвейерные машины



MISD - множественный поток команд и одиночный поток данных.

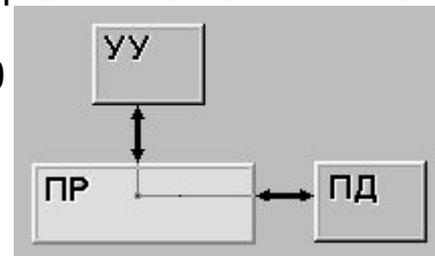
Классификация базируется на понятии потока - последовательности элементов обрабатываемых процессором

Однопроцессорные

последовательные компьютеры

VAX 11/780

CRAY-1

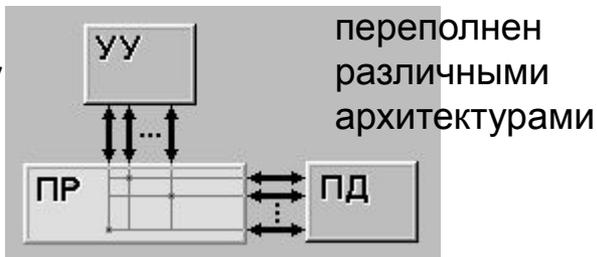


SISD - одиночный поток команд и одиночный поток данных

Мультипроцессорные системы

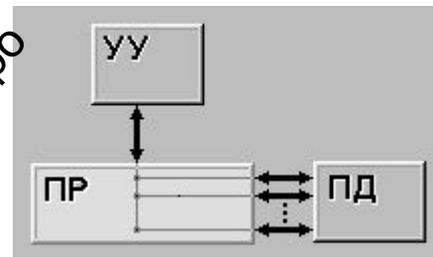
BBN Butterfly

Intel Paragon



MIMD - множественный поток команд и множественный поток данных.

Матрицы процессоров



ILLIAC IV

ICL DAP

SIMD - одиночный поток команд и множественный поток данных.

Дополнения Ванга и Бриггса

Схема классификации наиболее применима при оценке базовых принципов работы компьютера.

Детализирует классы архитектур :
SIMD, SISD, MIMD

Класс SISD разбивается на два подкласса:

- архитектуры с единственным функциональным устройством, например, PDP-11;
- архитектуры, имеющие в своем составе несколько функциональных устройств - CDC 6600, CRAY-1, FPS AP-120B, CDC Cyber 205, FACOM 200.

В классе MIMD авторы различают

- вычислительные системы со слабой связью между процессорами, к которым они относят все системы с распределенной памятью, например, Cosmic Cube,
- и вычислительные системы с сильной связью (системы с общей памятью), куда попадают такие компьютеры, как C.mmp, BBN Butterfly, CRAY Y-MP, Denelcor HEP.

В класс SIMD вводится два подкласса:

- архитектуры с пословно-последовательной обработкой информации - ILLIAC IV, PEPE, BSP;
- архитектуры с разрядно-последовательной обработкой - STARAN, ICL DAP

Пустой класс MISD полезен для разработки принципиально новых концепций в теории и практике построения вычислительных систем.

Классификация Фенга

Любая ВС - точка на плоскости координатами (n, m)

Интегральная характеристика потенциала параллельности P - максимальная степень параллелизма $BC P(C)=mn$, это некоторый аналог пиковой производительности.

Все вычислительные системы можно разделить на четыре класса :

Разрядно-последовательные
пословно-последовательные
($n=m=1$). В каждый момент времени такие компьютеры обрабатывают только один двоичный разряд. MINIMA (1, 1)

Разрядно-параллельные
пословно-последовательные ($n > 1, m = 1$).
Последовательные компьютеры :

IBM 701 (36,1) ; VAX 11/780 (32,1).
Разрядно-последовательные пословно-параллельные ($n = 1, m > 1$). Много одноразрядных процессоров, каждый независимо обрабатывает данные. STARAN (1, 256) и MPP (1,16384)

Классификация вычислительных систем на основе двух простых характеристик

- Число бит n в машинном слове, обрабатываемых параллельно при выполнении машинных инструкций
 - числу слов m , обрабатываемых одновременно данной вычислительной системой
- Функционирование любого компьютера можно представить как параллельную обработку n битовых слоев, на каждом из которых независимо преобразуются m бит.

Разрядно-параллельные пословно-параллельные ($n > 1, m > 1$). Большая часть существующих параллельных вычислительных систем BBN Butterfly (32, 256) , ILLIAC IV (64, 64).

Недостатки

- процессорные матрицы = векторно-конвейерные = многопроцессорные системами
- не исследуется за счет чего можно одновременно обрабатывать более одного слова.

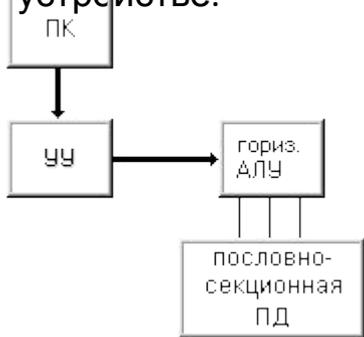
Достоинство - *введение единой числовой метрики* для всех типов компьютеров

Классификация Шора

Выделение типичных способов компоновки вычислительных систем на основе фиксированного числа базисных блоков: устройства управления, арифметико-логического устройства, памяти команд и памяти данных.

Выборка из памяти данных может осуществляться словами

Машина I - это вычислительная система, которая содержит устройство управления, арифметико-логическое устройство, память команд и память данных с пословной выборкой. Считывание данных осуществляется выборкой всех разрядов некоторого слова для их параллельной обработки в арифметико-логическом устройстве.



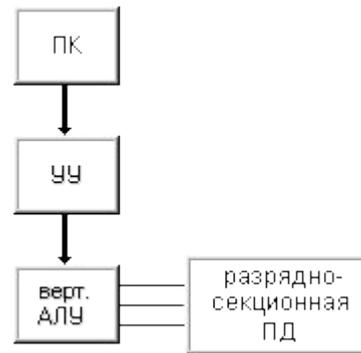
IBM 701

PDP-11

CRAY-1

классические последовательные машины
конвейерные скалярные машины
векторно-конвейерные машины

Машина II осуществляет выборку содержимого одного разряда из всех слов. Слова в памяти данных располагаются горизонтально. Доступ к словам идет через последовательную обработку битовых слоев при параллельной обработке множества слов.



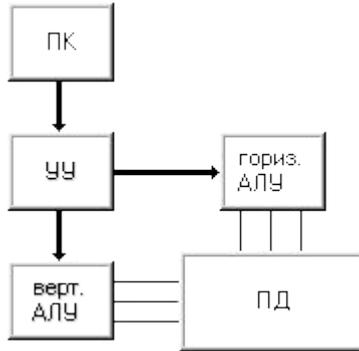
STARAN

ICL DAP

ассоциативные компьютеры
Матричные системы

Классификация Шора

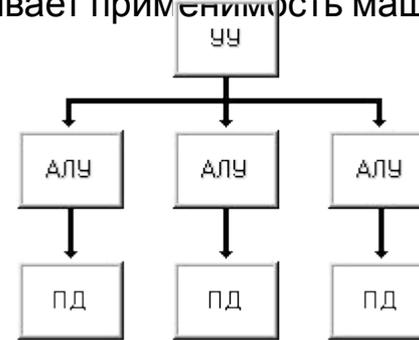
Если объединить принципы построения машин I и II, то получим **машину III**. Эта машина имеет два арифметико-логических устройства - горизонтальное и вертикальное, и модифицированную память данных, которая обеспечивает доступ как к словам, так и к битовым слоям.



Вычислительные системы семейства OMEN-60 фирмы Sanders Associates, построенные в прямом соответствии с концепцией ортогональной машины.

Если память представлять как матрицу слов, то доступ к данным осуществляется в направлении, "ортогональном" традиционному - не по словам (строкам), а по битовым слоям (столбцам)

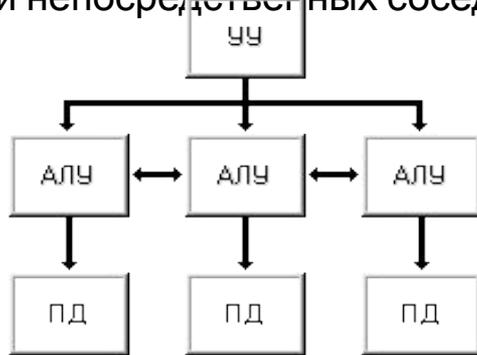
Если в машине I увеличить число пар арифметико-логическое устройство \Leftrightarrow память данных (иногда эту пару называют *процессорным элементом*) то получим **машину IV**. Единственное устройство управления выдает команду за командой сразу всем процессорным элементам. С одной стороны, отсутствие соединений между процессорными элементами делает дальнейшее наращивание их числа относительно простым, но с другой, сильно ограничивает применимость машин этого класса.



Такую структуру имеет вычислительная система PERE, объединяющая 288 процессорных элементов.

Классификация Шора

Если ввести непосредственные линейные связи между соседними процессорными элементами машины IV, например в виде матричной конфигурации, то получим схему **машины V**. Любой процессорный элемент теперь может обращаться к данным как в своей памяти, так и в памяти непосредственных соседей.



Классический матричный компьютер
ILLIAC IV

Заметим, что все машины с I-ой по V-ю придерживаются концепции разделения памяти данных и арифметико-логических устройств, предполагая наличие шины данных или какого-либо коммутирующего элемента между ними. **Машина VI**, названная *матрицей с функциональной памятью* (или памятью с встроенной логикой), представляет собой другой подход, предусматривающий распределение логики процессора по всему запоминающему устройству.



простые ассоциативные запоминающие
устройства
сложные ассоциативные
процессоры

Классификация

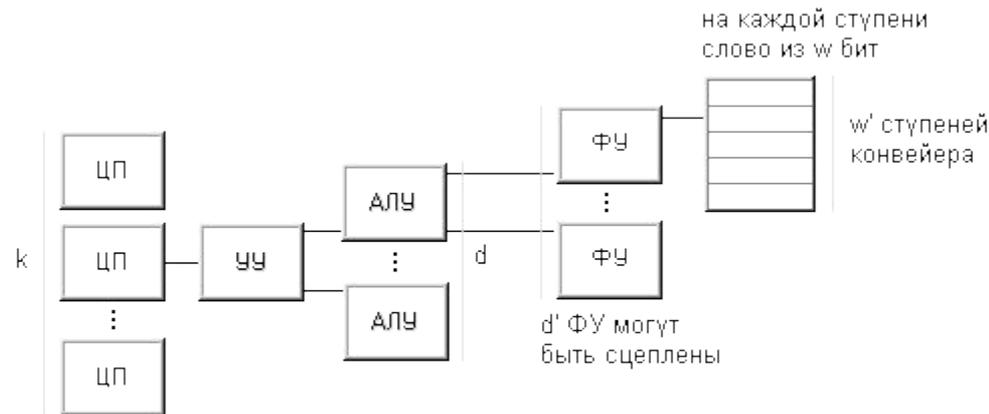
Детализация
Фенга

Хендлера

Три уровня обработки данных в процессе выполнения программ:

- уровень выполнения программы - опираясь на счетчик команд и некоторые другие регистры, устройство управления (УУ) производит выборку и дешифрацию команд программы;
- уровень выполнения команд - арифметико-логическое устройство компьютера (АЛУ) исполняет команду, выданную ему устройством управления;
- уровень битовой обработки - все элементарные логические схемы процессора (ЭЛС) разбиваются на группы, необходимые для выполнения операций над одним двоичным разрядом.

основа классификации - явное описание возможностей параллельной и конвейерной обработки информации вычислительной системой
способы связи между процессорами и блоками памяти не рассматриваются



Тройка для описания вычислительной системы С:

$$t(C) = (k \times k', d \times d', w \times w'), \text{ где}$$

k - число процессоров (каждый со своим УУ), работающих параллельно

k' - глубина макроконвейера из отдельных процессоров

d - число АЛУ в каждом процессоре, работающих параллельно

d' - число функциональных устройств АЛУ в цепочке

w - число разрядов в слове, обрабатываемых в АЛУ параллельно

w' - число ступеней в конвейере функциональных устройств АЛУ

k' ЦП из k могут работать в макроконвейере

Исчезает проблема с определением векторно-конвейерных машин

- $t(\text{ILLIAC IV}) = (1, 64, 64)$
- $t(\text{STARAN}) = (1, 8192, 1)$
- $t(\text{C.mmp}) = (16, 1, 16)$
- $t(\text{PRIME}) = (5, 1, 16)$
- $t(\text{BBN Butterfly GP1000}) = (256, \sim 1, \sim 32)$

Классификация Хокни

Детализация классификации и Флинна

систематизации компьютеров, попадающих в класс MIMD по систематике Флинна.

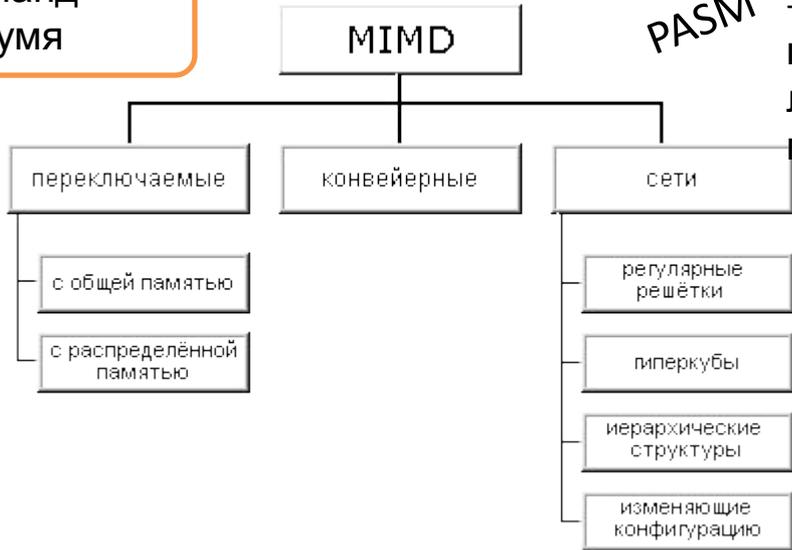
Множественный поток команд может быть обработан двумя способами:

-одним конвейерным устройством обработки, работающем в режиме разделения времени для отдельных потоков

Конвейерные MIMD

-каждый поток обрабатывается своим собственным устройством

- ВС, где возможна прямая связь каждого процессора с каждым, реализуемая с помощью переключателя;
- ВС, где прямая связь каждого процессора возможна только с ближайшими соседями по сети, взаимодействие удаленных процессоров через процессоры-посредники.



машины с

переключателем PASM -память распределена среди процессоров как их локальная

память это разделяемый ресурс, доступный всем процессорам через переключатель

-гибридные ВС

CRAY X-MP

Машины с сетевой структурой систематизируются в соответствии с топологией сети :

- звздообразная сеть (ICAP)
- гиперкубы (NCube, Intel iPCS)
- кластеры (Cm*, CEDAR)

Классификация Джонсона

Детализация
классификации
и Флинна

классификацию MIMD архитектур на основе структуры памяти и реализации механизма взаимодействия и синхронизации между процессорами

По структуре оперативной памяти существующие вычислительные системы делятся на две большие группы:

- системы с общей памятью, прямо адресуемой всеми процессорами
- системы с распределенной памятью, каждая часть которой доступна только одному процессору.

Для межпроцессорного взаимодействия существуют две альтернативы:

- через разделяемые переменные
- С помощью механизма передачи сообщений.

Четыре класса MIMD архитектур, уточняющих систематику Флинна:

- общая память - разделяемые переменные (GMSV);
- распределенная память - разделяемые переменные (DMSV);
- распределенная память - передача сообщений (DMMP);
- общая память - передача сообщений (GMMP).

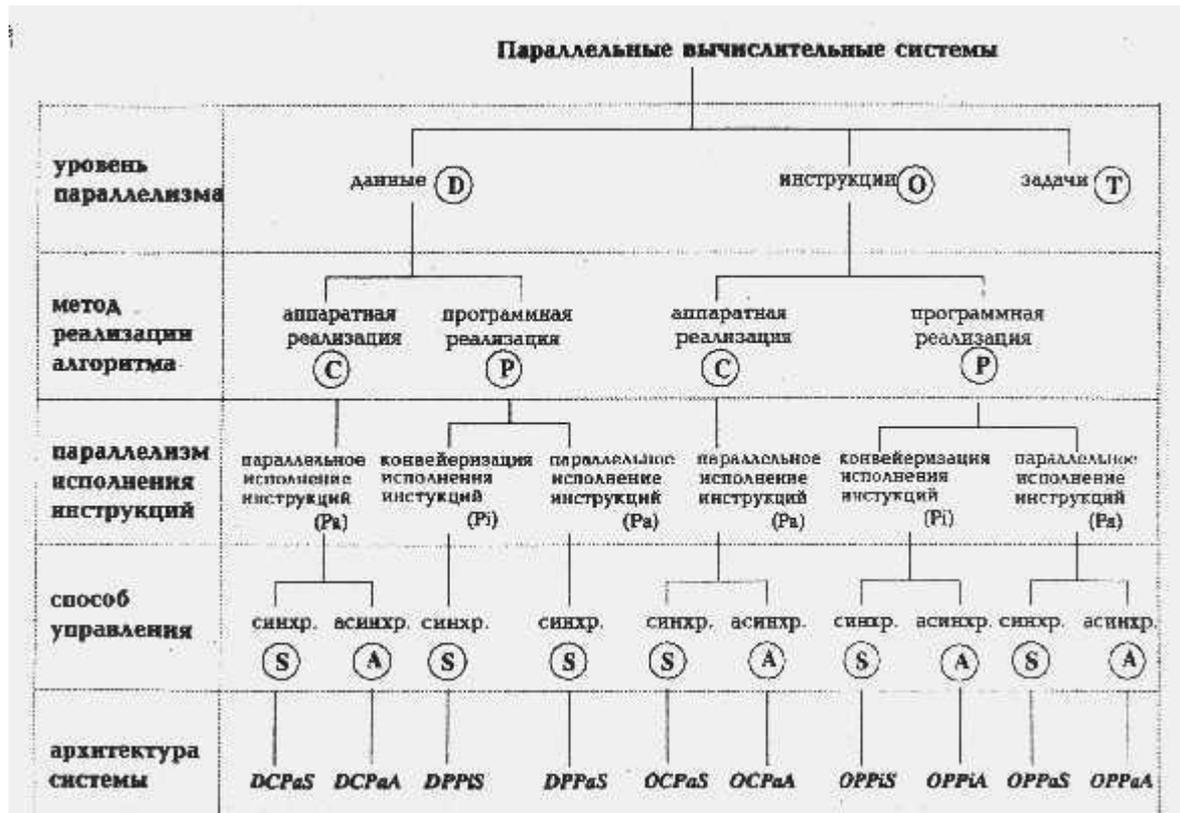
Вычислительные системы, использующие общую разделяемую память для межпроцессорного взаимодействия и синхронизации - системами с переменной, как в BBN Butterfly, называются разделяемой памятью, например, CRAY Y-MP (I) гибридными архитектурами (II)

Для уточнения классификации существует возможность учитывать вид связи между процессорами: общую шину, переключатели, разнообразные сети.

Классификация Базу

корень дерева - вычислительная система

Любую параллельную вычислительную систему можно однозначно описать последовательностью решений, принятых на этапе ее проектирования, а сам процесс проектирования представить в виде дерева



Систолические деревья, введенные Кунгом для вычисления арифметических выражений могут быть описаны как OSP_aS либо OSP_aA

IBM 360/91, Amdahl 470/6 и многие современные RISC процессоры, разбивающие исполнение всех инструкций на несколько этапов, в данной классификации имеют обозначение OPP_iS

Data-flow компьютеры, в зависимости от особенностей реализации, могут быть описаны либо как OPP_iA , либо OPP_aA .

аналог
Базу

Классификация Кришнамарфи

для классификации параллельных вычислительных систем
предлагает использовать четыре характеристики

- 1. степень гранулярности;
- 2. способ реализации параллелизма;
- 3. топология и природа связи процессоров;
- 4. способ управления процессорами.

Несмотря на то, что классификация Е. Кришнамарфи построена лишь на четырех признаках, она позволяет выделить и описать такие "нетрадиционные" параллельные системы, как систолические массивы, машины типа dataflow и wavefront.

Можно расширить классификацию если
Дополнить ее размещением задач по процессорам, способом маршрутизации сообщений, возможностью реконфигурации, аппаратной поддержки языков программирования

Векторно-конвейерные компьютеры:

- гранулярность - на уровне данных;
- реализация параллелизма - аппаратная;
- связь процессоров - простая топология со средней связностью;
- способ управления - синхронный.



Гиперкубы и ассоциативные процессоры
Определяются неоднозначно

Классификация Скилликорна

Классификация состоит из двух уровней. На первом уровне

она проводится на основе восьми характеристик:

1. тип переключателя между DP и DM;
2. количество процессоров команд (IP);
3. число запоминающих устройств (модулей памяти) команд (IM);
4. тип переключателя между IP и IM;
5. количество процессоров данных (DP);
6. число запоминающих устройств (модулей памяти) данных (DM);
7. тип переключателя между IP и DP;
8. тип переключателя между DP и DP.

На втором уровне классификации Д. Скилликорн просто уточняет описание, сделанное на первом уровне, добавляя возможность конвейерной обработки в процессорах команд и данных.

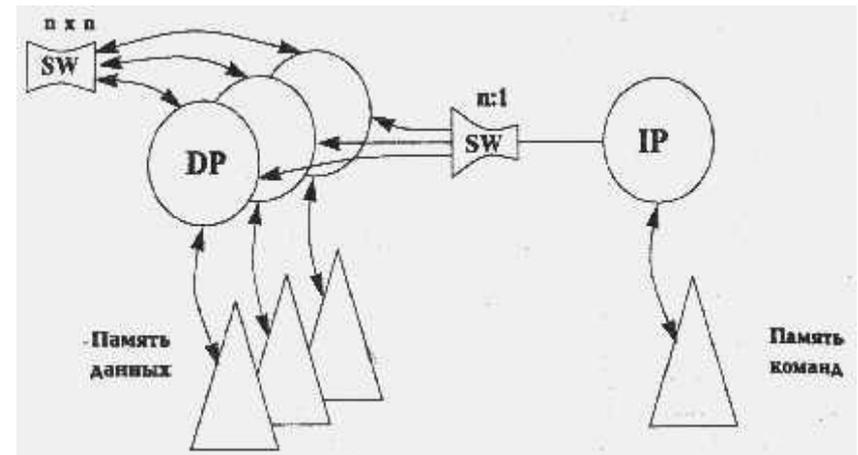
Основа - абстрактная структура, состоящая из четырех компонент:

- процессор команд (IP) - интерпретатор команд
- процессор данных (DP) - преобразователь данных, в соответствии с арифметическими операциями;
- иерархия памяти (IM, DM) - запоминающее устройство
- переключатель - связь между процессорами и

очередная попытка расширить классификацию Флинна

Рассмотрим компьютер Connection Machine 2

его можно описать как:
(1, 1, 1-1, n, n, n-n, 1-n, nxn)



Классификация

Дазгупты

Модификация метода
Скилликорна

Назовем *классом классификационной схемы* именованную группу объектов, которые по некоторым специально выделенным свойствам отличаются от объектов других классов. Множество классов образует *категорию*.

Предлагается следующая иерархия категорий:

1. Самый низкий уровень - категория КЭШ-процессора с памятью MCP (memory-cache processor).
Классами этой категории являются всевозможные различные архитектуры. Соответствующую архитектуре формулу можно рассматривать как имя класса.
2. Более высокий уровень - категория КЭШ-процессора (CP). Множество классов этой категории получается путем удаления из классов категории CP составляющих, описывающих память.
3. Самый высокий уровень - категория процессора (P). Классы получают удалением кэш-составляющих из классов категории CP.

Базовые элементы архитектуры:

- iM - память с расслоением - память, из которой можно выбрать несколько единиц информации за один цикл памяти;
- sM - простая память - память, из которой можно выбрать единицу информации за цикл памяти;
- C - программируемая или непрограммируемая кэш-память. Буферные регистры, подобные регистрам CRAY-1, также описываются, как кэш-память;
- sl - простой (неконвейерный) процессор для подготовки команды к исполнению;
- pl - конвейерный процессор для подготовки команды к исполнению;
- sX - простой процессор для исполнения команды;
- pX - конвейерный процессор для исполнения команды.

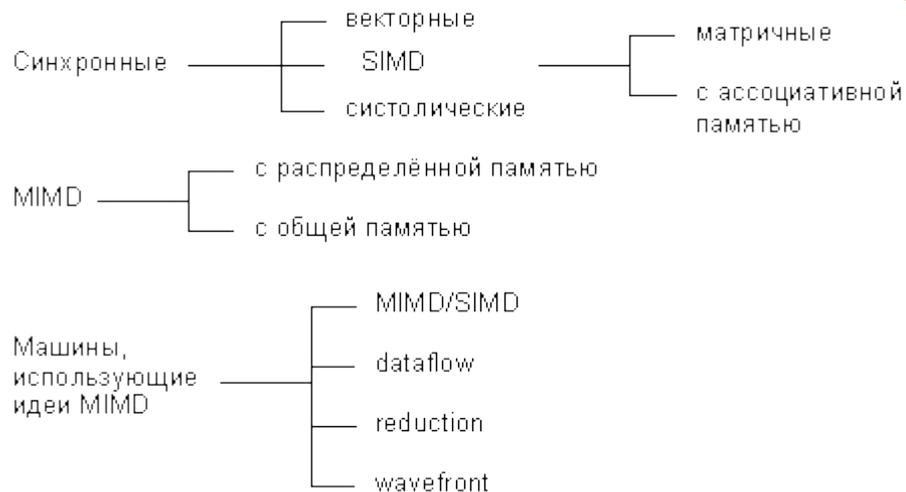
последние исследования
по классификации

Классификация Дункана

Включает
все

Параллельная архитектура - это такой способ организации вычислительной системы, при котором допускается, чтобы множество процессоров (простых или сложных) могло бы работать одновременно, взаимодействуя по мере надобности друг с другом.

Требование Дункана - не рассматривать компьютеры, использующие только параллелизм низкого уровня, наравне со всеми остальными, т.к., во-первых, практически все существующие системы будут классифицированы как "параллельные" (что заведомо не будет позитивным фактором для классификации), и, во-вторых, такие машины будут плохо вписываться в любую модель или концепцию, отражающую параллелизм высокого уровня.



Систематика очень простая: процессоры системы работают либо синхронно, либо независимо друг от друга, либо в архитектуру системы заложена та или иная модификация идеи MIMD. На следующем уровне происходит детализация в рамках каждого из этих **трех классов**.

K H e C