

Лекция 4

Эволюция архитектуры вычислительных систем

Ефимов Александр Владимирович
E-mail: alexander.v.efimov@sibguti.ru

Курс «Архитектура вычислительных систем»
СибГУТИ, 2018

Понятие архитектуры

Архитектура ЭВМ – совокупность свойств и характеристик ЭВМ, призванных удовлетворить потребности пользователей



Архитектура вычислительного средства – это совокупность его свойств и характеристик



Архитектура вычислительного средства – это концепция взаимосвязи и функционирования его аппаратурных (Hardware) и программных (Software) компонентов

Модель вычислителя

$$c = \langle h, a \rangle$$

h – конструкция вычислителя

a – алгоритм его работы

$$a(p(D))$$

D – исходные данные

p – программа

$$h = \langle U, g \rangle$$

$U = \{u_i\}$ – множество устройств u_i , $i = \overline{1, k}$

g – структура сети связи между устройствами

Конструкция вычислителя

1. Последовательная обработка информации
2. Фиксированность структуры
3. Неоднородность составляющих устройств и связей между ними

Каноническая ЭВМ Дж. фон Неймана соответствует модели вычислителя

Принципы организации ЭВМ

Дж. фон Неймана

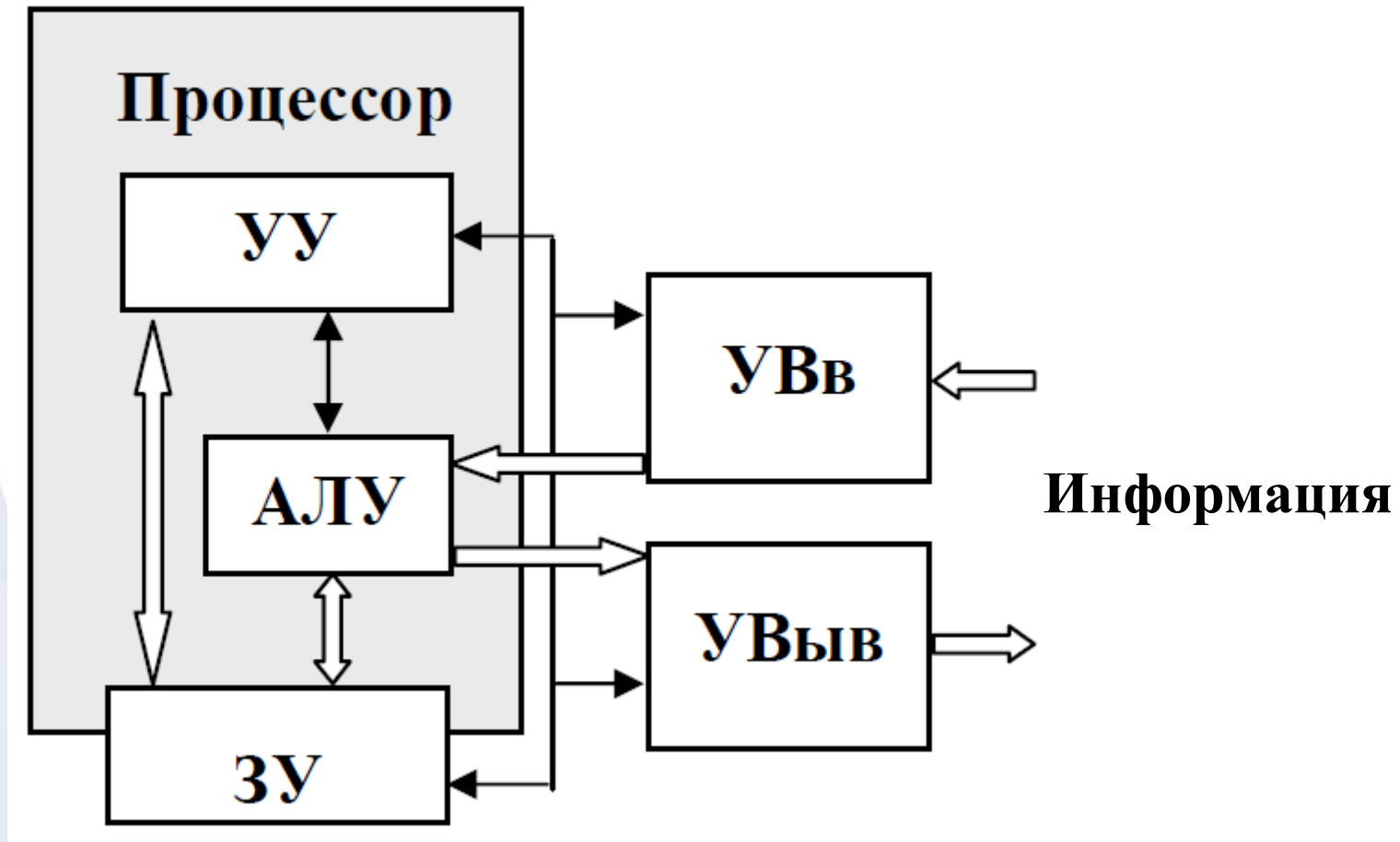


1903 – 1957

- Программное управление работой ЭВМ
- Условный переход
- Принцип хранимой программы
- Использование двоичной системы счисления
- Иерархичность запоминающих устройств

Функциональная структура ЭВМ

Дж. фон Неймана



Поколения ЭВМ

➤ Эффективность

$$E = \{\omega, v, \vartheta, \sigma\},$$

- ω – показатель производительности (в опер./с) или среднее число операций, выполняемых в секунду ЭВМ (процессором при работе с оперативной памятью);
- v – емкость оперативной памяти (в битах);
- ϑ – среднее время безотказной работы ЭВМ (или средняя наработка до отказа, в часах);
- σ – «цена операций», определяемая как отношение цены ЭВМ к ее показателю производительности (измеряется в долларах, отнесенных к опер./с).

➤ Архитектурные свойства ЭВМ

- способы и режимы обработки информации;
- конструктивные особенности (составы устройств и структуры);
- алгоритмы управления вычислительными процессами или алгоритмы функционирования машин;
- возможности программного обеспечения (языки, ОС и т.п.);
- свойства элементной базы, характер проектирования и производства ЭВМ.

1-е поколение (до 1950-х гг.)

- ✓ Операционные системы отсутствовали
- ✓ Механические реле → электронные лампы
- ✓ Коммутационные панели → перфокарты



Разработчики:

Конрад Цузе (Konrad Zuse)

Говард Айкен (Howard Aiken)

Джон фон Нейман
(John von Neumann)

Дж. Преспер Эккерт
(J. Presper Eckert)

Вильям Мочли (William Mauchley)

Первое поколение ЭВМ (1949 – 1951)

➤ Показатели эффективности

ω = 10 опер./с

ν = 10 бит

ϑ = 1–10 ч

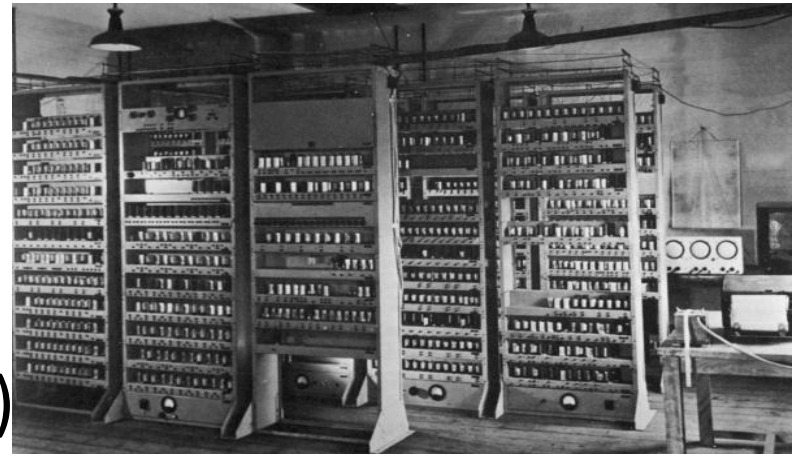
σ = 10 дол./опер. · с⁻¹

➤ Архитектурные свойства

- Последовательная обработка информации в монопрограммном режиме
- Состав вычислительных устройств и структура – канонические
- Алгоритм управления вычислительными процессами – универсальный и последовательный, адаптирован под фиксированную структуру ЭВМ
- Возможности программного обеспечения ЭВМ первого поколения: машинные языки (двоичные коды) для записи алгоритмов обработки информации и стандартные подпрограммы (например, для вычисления элементарных функций)
- Элементную базу составляли электронные лампы.
- Проектирование машин было «ручным», а их производство – индивидуальным.

2-ое поколение (1950 – сер. 1960-х гг.)

- ✓ Системы пакетной обработки
- ✓ Электронные лампы → транзисторы
- ✓ ОС: FMS (Fortran Monitor System)
IBSYS (для IBM 7094)



**Electronic Delay Storage
Automatic Calculator (EDSAC)**



**Малая Электронная Счётная Машина
(МЭСМ)**



**Сергей Алексеевич
Лебедев**



Морис Уилкс

Второе поколение ЭВМ (1955 – 1960)

➤ Показатели эффективности

$$\omega = 10^6 \text{ опер./с}$$

$$\nu = 10^7 \text{ бит}$$

$$\vartheta = 10^2 \text{ ч}$$

$$\sigma = 10 \text{ дол./опер.} \cdot \text{с}^{-1}$$

➤ Архитектурные свойства

- Последовательная обработка информации
- *Мультипрограммирование* - режим обработки данных, при котором ресурсы ЭВМ одновременно используются более чем одной программой
- В процессор введены структурные решения, ускорившие процесс реализации арифметических операций, а также схемы прерывания, обеспечившие работу ЭВМ в реальном масштабе времени
- Возможность подключения каналов связи для обеспечения теледоступа
- Универсальный алгоритм управления вычислительными процессами стал последовательно-параллельным
- Программное обеспечение: диспетчеры, языки для записи алгоритмов обработки информации (АЛГОЛ 60, ФОРТРАН) и соответствующие трансляторы
- Основы элементной и логико-конструктивной баз ЭВМ второго поколения составляли соответственно полупроводниковые приборы и вентили
- Мелкосерийное производство

Система пакетной обработки



Ранняя система пакетной обработки: а — программист приносит карты для IBM 1401; б — IBM 1401 записывает пакет заданий на магнитную ленту; в — оператор приносит входные данные на ленте к IBM 7094; г — IBM 7094 выполняет вычисления; д — оператор переносит ленту с выходными данными на IBM 1401; е — IBM 1401 печатает выходные данные

3-е поколение (1965 – 1980)

- ✓ Транзисторы → интегральные микросхемы
- ✓ Системы разделения времени
- ✓ ОС: OS /360, CTSS, MULTICS



PDP-1



IBM System
/360



БЭСМ-6



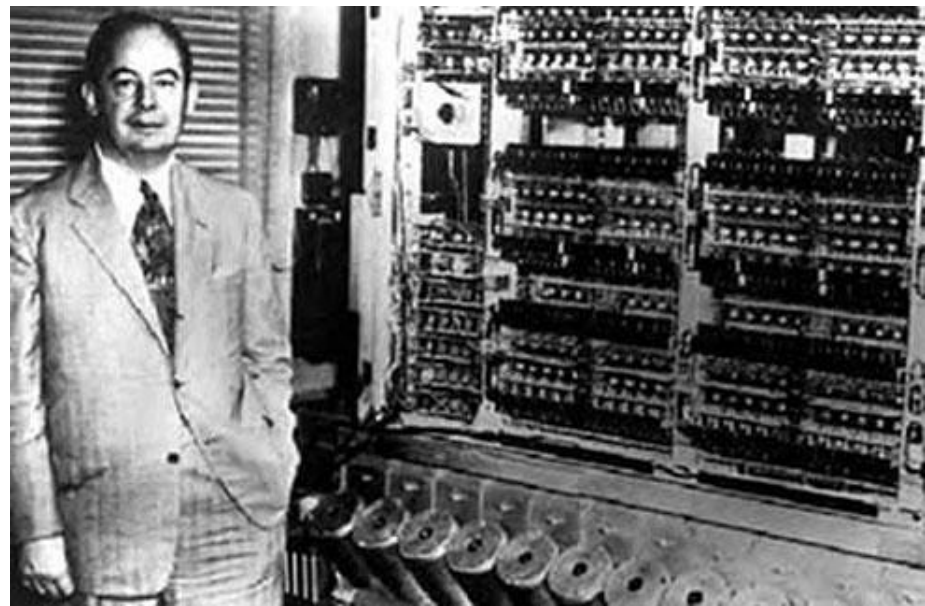
БЭСМ-1

БЭСМ-1 и БЭСМ-2: 1953 – 1958
 М-20: 1957
 М-40 и М-50: 1957 – 1959
 БЭСМ-4: 1962

EDVAC: 1950 – 1952
 MADAM: 1951
 JOHNIAC: 1953



БЭСМ-4



EDVAC

Третье поколение ЭВМ (1963 – 1965)

➤ Показатели эффективности

$$\omega = 10^7 \text{ опер./с}$$

$$\nu = 10^8 \text{ бит}$$

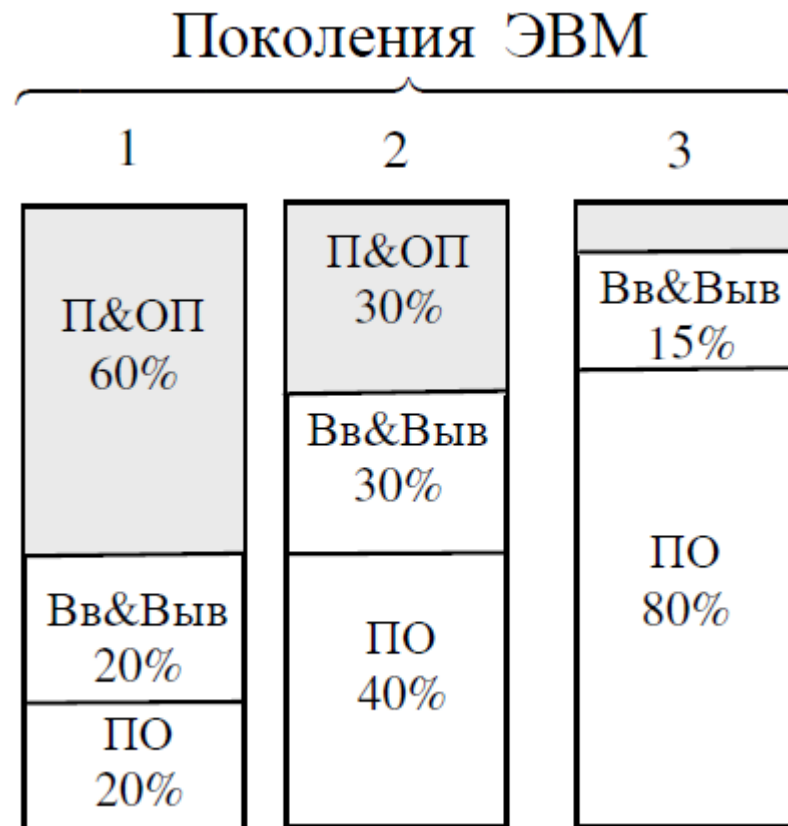
$$\vartheta = 10^3 \text{ ч}$$

$$\sigma = 10^{-1} \text{ дол./опер.} \cdot \text{с}^{-1}$$

➤ Архитектурные свойства

- Последовательная обработка информации
- Мультипрограммные режимы: *пакетная обработка и разделение времени.*
- Состав вычислительных устройств дополнен спецпроцессорами, оптическими устройствами ввода-вывода информации, накопителями (на магнитных лентах и дисках) большой емкости и др.
- Единый ресурс, через который осуществлялись взаимодействия между процессором и остальными устройствами
- Процедурный и структурный способы вычислений
- Последовательно-параллельный алгоритм управления вычислительными процессами
- Операционные системы и системы автоматизации программирования
- Элементная база опиралась на интегральную технологию (комплекты ИС)
- Серийное и автоматизированное производство

Распределение стоимости между компонентами ЭВМ



Роль СО РАН в развитии ВТ

- ✓ 1962 – 1965 гг. – разработка концептуальных основ построения вычислительных средств, базирующихся на новых принципах обработки информации (с нефоннеймановской архитектурой). Эти средства стали называть *вычислительными системами* (ВС) или параллельными ВС.
- ✓ Первая параллельная ВС (с программируемой структурой) – это система «Минск-222». Она была разработана и построена Институтом математики Сибирского отделения АН СССР (г. Новосибирск) совместно с Конструкторским бюро завода им. Г.К. Орджоникидзе (г. Минск) в 1965 – 1966 гг.
- ✓ Американская система ILLIAC-IV была построена в 1972 г. Реализованная единственная конфигурация ILLIAC-IV многие годы оставалась самой мощной ВС (2×10^8 опер./с), однако она по архитектурным свойствам и функциональной гибкости уступала конфигурациям системы “Минск-222”.

Модель коллектива вычислителей

Архитектура ВС основывается на структурной и функциональной имитации коллектива людей-вычислителей

$$S = \langle H, A \rangle$$

H – конструкция

A – алгоритм работы коллектива вычислителей.

$$H = \langle C, G \rangle$$

$C = \{c_i\}$ – множество вычислителей c_i , $i = \overline{0, N-1}$

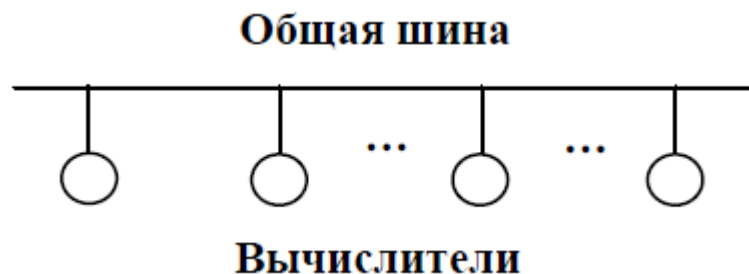
N – мощность множества C

G – описание макроструктуры коллектива вычислителей, т.е. структуры сети связей между вычислителями $c_i \in C$ (или структура коллектива)

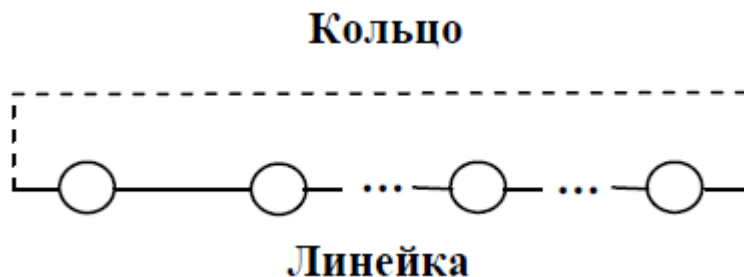
Структура ВС

(Макро)структура коллектива вычислителей представляется *графом G* , вершинам (узлам) которого сопоставлены вычислители $c_i \in C$, а ребрам – линии связи между ними

➤ Нульмерные структуры

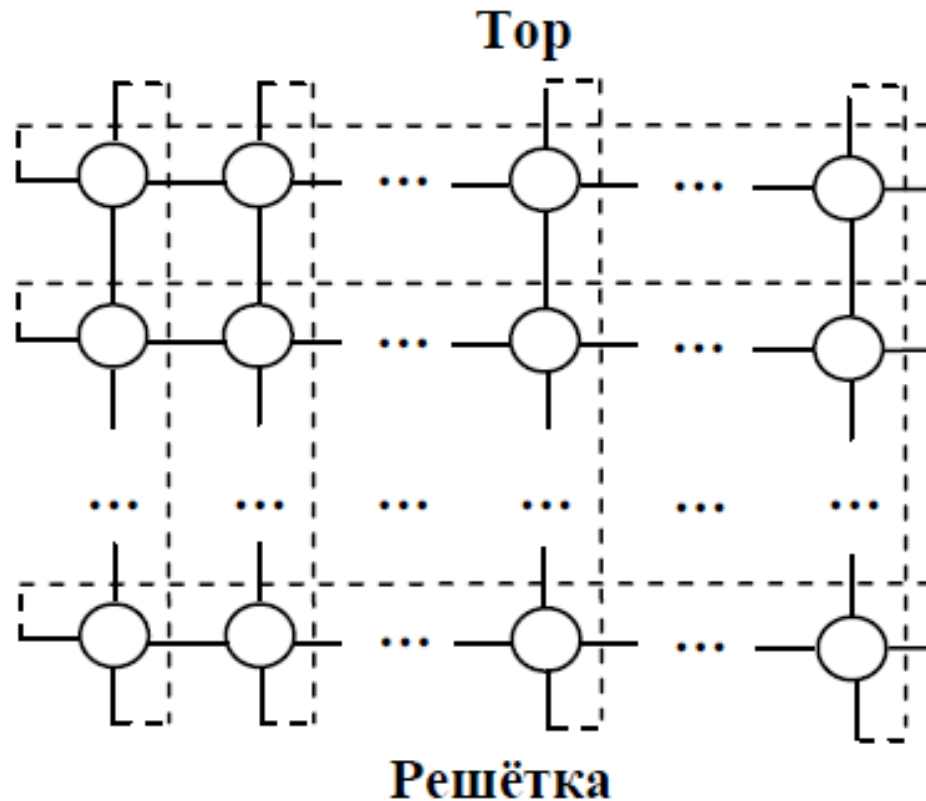


➤ Одномерные структуры



Структура ВС

➤ Двумерные структуры



**Увеличение размерности структуры
повышает структурную надёжность ВС**

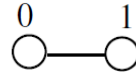
Структура ВС

➤ Гиперкубы (структуры в виде булевых n -мерных кубов)

$$n = \log_2 N$$

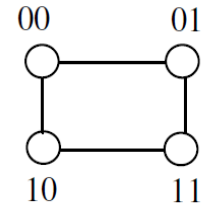
N – количество вершин
 n – число ребер,
выходящих из вершины
(размерность)

Максимальное расстояние
(число ребер) между двумя
вершинами совпадает с
размерностью гиперкуба



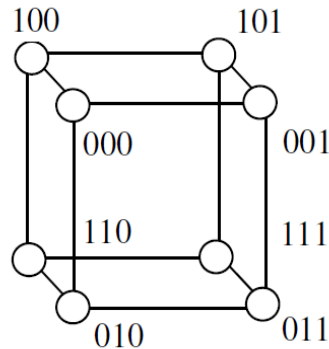
$N = 2, n = 1$

1D – куб



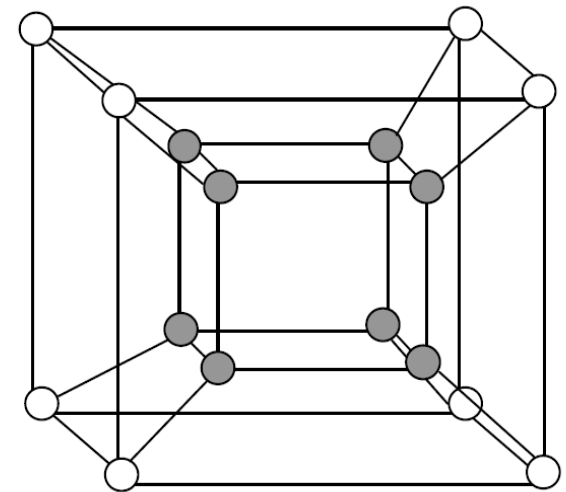
$N = 4, n = 2$

2D – куб



$N = 8, n = 3$

3D – куб



$N = 16, n = 4$, 4D – куб

Архитектурные принципы коллектива вычислителей

- *Параллелизм при обработке информации* - параллельное выполнение операций на множестве S вычислителей, взаимодействующих через связи структуры G ;
- *Программируемость структуры* - настраиваемость структуры G сети связей между вычислителями, достигаемая программными средствами;
- *Однородность конструкции H* - однородность вычислителей $s_i \in S$ и структуры G .

Понятие о ВС

Вычислительная система – средство обработки информации, базирующееся на модели коллектива вычислителей

Вычислительная система – совокупность взаимосвязанных и одновременно функционирующих аппаратурно-программных вычислителей, которая способна не только реализовать (параллельный) процесс решения сложной задачи, но и априори и в процессе работы автоматически настраиваться и перестраиваться с целью достижения адекватности между своей структурно-функциональной организацией и структурой и характеристиками решаемой задачи.

Классы архитектур ВС

Классификация архитектур средств обработки информации была предложена профессором Стенфордского университета США М. Дж. Флинном в 1966 г.

SISD (Single Instruction stream / Single Data stream) или
ОКОД (Одиночный поток Команд и Одиночный поток Данных)

ЭВМ

SIMD (Single Instruction stream / Multiple Data stream) или
ОКМД (Одиночный поток Команд и Множественный поток Данных)

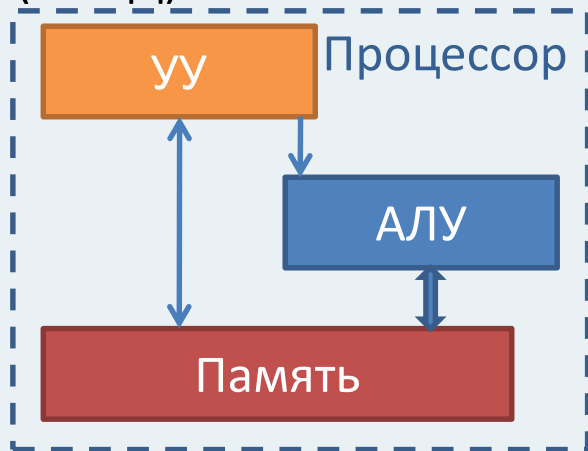
MISD (Multiple Instruction stream / Single Data stream) или
МКОД (Множественный поток Команд и Одиночный поток Данных)

MIMD (Multiple Instruction stream / Multiple Data stream) или
МКМД (Множественный поток Команд и Множественный поток Данных)

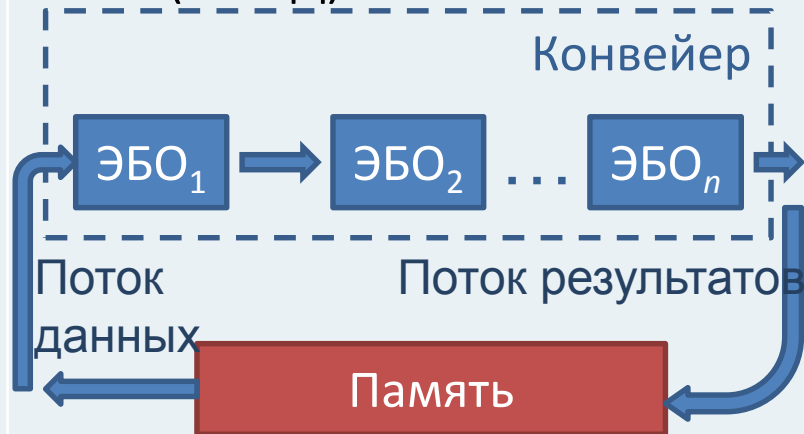
ВС

Классы архитектур ВС

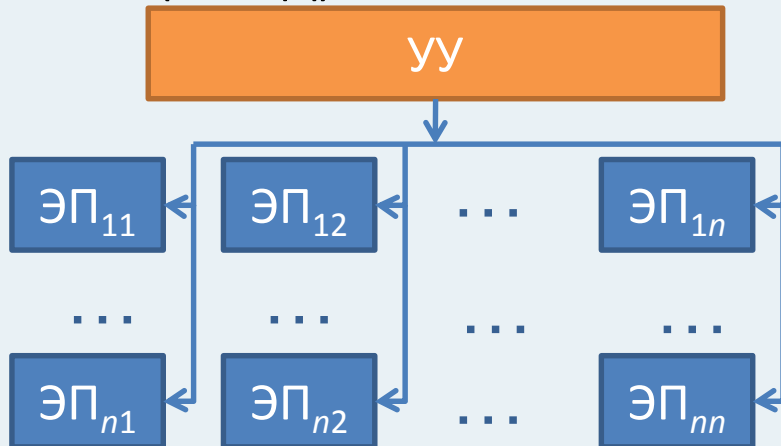
SISD (ОКОД)



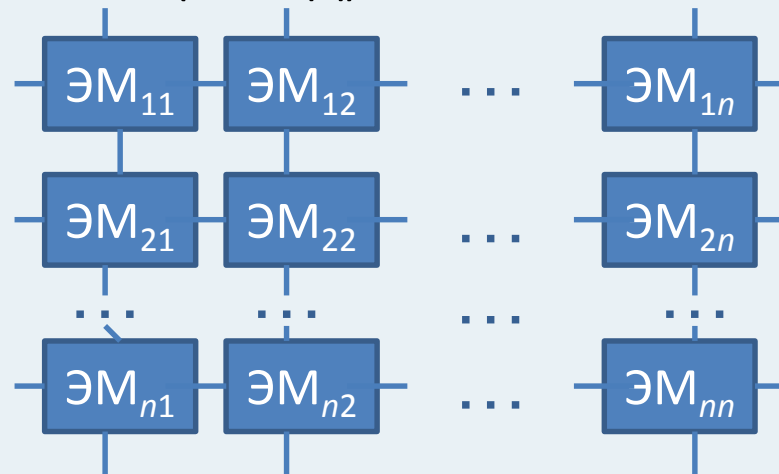
MISD (МКОД)



SIMD (ОКМД)



MIMD (МКМД)



Конвейерные ВС (MISD)

Конвейерные ВС – системы, архитектура которых является предельным вариантом эволюционного развития последовательной ЭВМ и простейшей версией модели коллектива вычислителей

- Конвейерный способ обработки информации
- Функциональная структура представляется в виде «последовательности» связанных элементарных блоков обработки (ЭБО) информации
- Все блоки работают параллельно, но каждый из них реализует лишь свою операцию над данными одного и того же потока

Матричные ВС (SIMD)

Матричные ВС основываются на принципе массового параллелизма, в них обеспечивается возможность одновременной реализации большого числа операций на элементарных процессорах (ЭП), «объединенных» в матрицу

- ЭП – композиция из арифметико-логического устройства (АЛУ) и локальной памяти (ЛП)
- Поток команд на матрицу ЭП формируется устройством управления
- SIMD-архитектура в классическом виде

Мультипроцессорные ВС (MIMD с общей памятью)

Обширная группа систем, в которую, в частности, могут быть включены конвейерные и матричные ВС (а также многомашинные ВС)

Обычно к мультипроцессорным ВС относят системы с MIMD-архитектурой, которые состоят из множества (не связанных друг с другом) процессоров и *общей* памяти; взаимодействие между процессорами и памятью осуществляется через коммутатор (общую шину и т.п.), а между процессорами – через память

Распределённые ВС (MIMD)

Мультипроцессорные ВС с MIMD-архитектурой, в которых *нет* единого ресурса (общей памяти)

- Принципы модульности и близкодействия
- Пример промышленной реализации распределённых ВС – *транспьютерные ВС* – композиция из одинаковых взаимосвязанных микропроцессорных кристаллов, называемых *транспьютерами*

ВС с программируемой структурой (MIMD)

Полностью основываются на модели коллектива вычислителей и являются композицией взаимосвязанных элементарных машин

- Состав ЭМ: локальный коммутатор (ЛК), процессор, память
- Ориентированы на распределенную обработку информации; эффективны и при конвейерной, и при матричной обработке
- Рассчитаны на работу во всех основных режимах: решения сложной задачи, обработки наборов задач, обслуживания потоков задач, реализации функций вычислительной сети
- Концепция вычислительных систем с программируемой структурой была сформулирована в Сибирском отделении АН СССР, первая система («Минск-222») была построена в 1965 – 1966 гг.

Кластерные ВС

- Термин вычислительный кластер (cluster) был впервые введен компанией DEC (Digital Equipment Corporation):

Кластер – это группа компьютеров, которые связаны между собой и функционируют как единое средство обработки информации

- *Кластерная ВС* или кластер – это композиция множества вычислителей, сети связей между ними и программного обеспечения, предназначенная для параллельной обработки информации

Принципы построения ВС

- **Модульность** – принцип, предопределяющий формирование вычислительной системы из унифицированных элементов (называемых модулями), которые функционально и конструктивно закончены, имеют средства сопряжения с другими элементами и разнообразие которых составляет полный набор.

Обеспечивает:

- возможность использования любого модуля заданного типа для выполнения любого соответствующего ему задания пользователя;
- простоту замены одного модуля на другой однотипный;
- масштабируемость, т.е. возможность увеличения или уменьшения количества модулей без коренной реконфигурации связей между остальными модулями;
- открытость системы для модернизации, исключаящую ее моральное старение.

Принципы построения ВС

- **Близкодействие** – принцип построения ВС, обуславливающий такую организацию информационных взаимодействий между модулями-вычислителями, при которой каждый из них может непосредственно (без «посредников») обмениваться информацией с весьма ограниченной частью модулей-вычислителей.
- *Локальность связей и взаимодействий между вычислителями* – состояние $E_i(t+1)$ вычислителя c_i , $i \in \{0, 1, \dots, N-1\}$, на очередном временном шаге $t+1$ зависит от состояний (на предшествующем шаге t) непосредственно с ним связанных вычислителей $c_j \in C^*$, $C^* \subset C$:

$$E_i(t+1) = f(E_i(t), E_{i_1}(t), E_{i_2}(t), \dots, E_{i_m}(t));$$

- *Асинхронность функционирования ВС* обеспечивается, если порядок срабатывания ее модулей определяется не с помощью вырабатываемых тем или иным образом отметок времени, а достижением заданных значений определенных (как правило, логических) функций.

Принципы построения ВС

- *Децентрализованность управления ВС* достигается, если в системе нет выделенного модуля, который функционирует как единый для всей системы центр управления.

Позволяет:

- достичь живучести ВС, т. е. ее способности продолжать работу при отказах модулей (в том числе и тех, которые предназначены для принятия решений);
- избежать очередей при обслуживании «заявок» на управление.

- *Распределённость ресурсов ВС*

Распределённая ВС – система, в которой нет единого ресурса, используемого другими в режиме разделения времени.

Литература

Хорошевский В.Г. Архитектура вычислительных систем. Учебное пособие. – М.: МГТУ им. Н.Э. Баумана, 2005; 2-е издание, 2008.

Хорошевский В.Г. Инженерные анализ функционирования вычислительных машин и систем. – М.: “Радио и связь”, 1987.