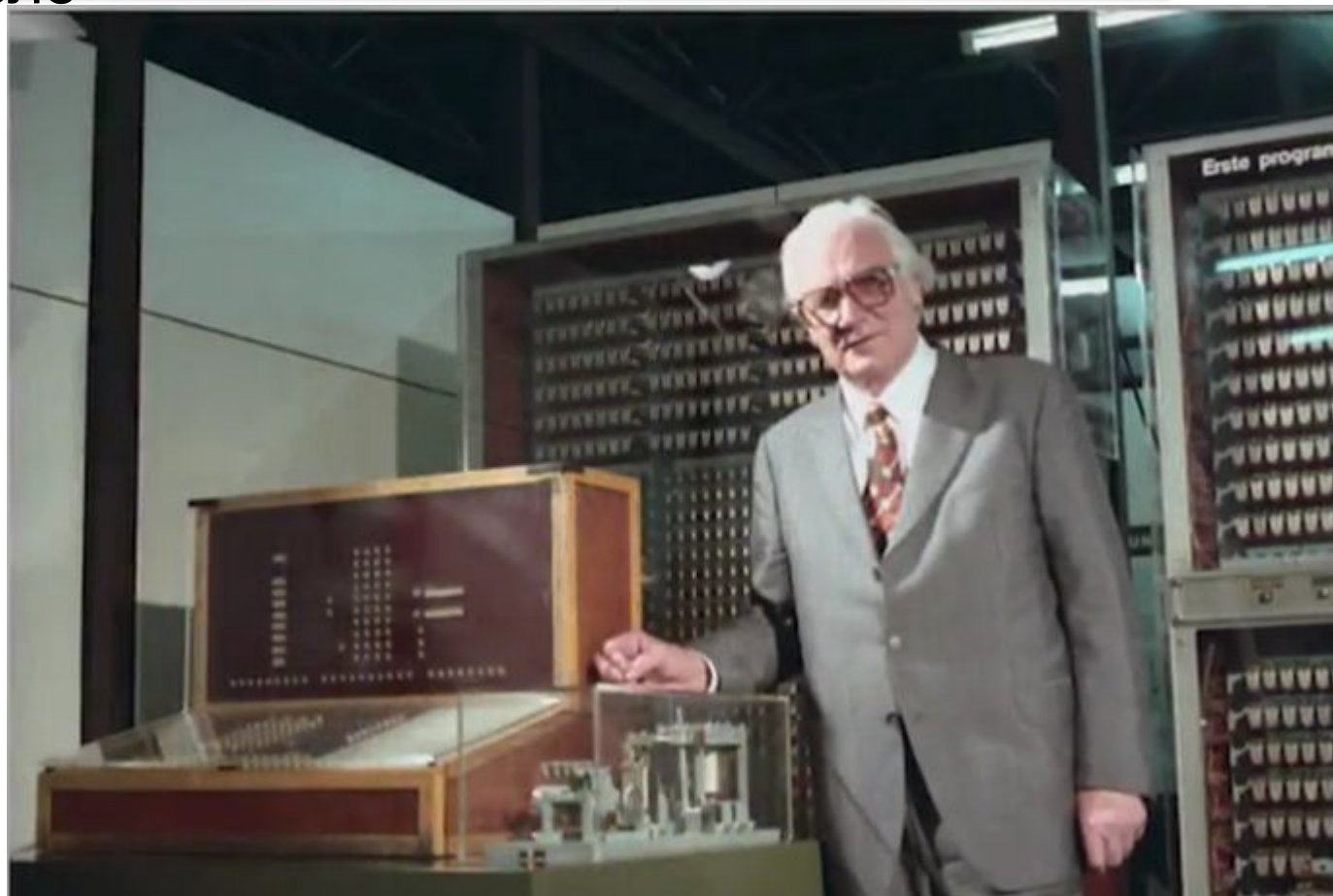


Лекция 2

Принципы построения
параллельных вычислительных
систем

История. Конрад Цузе и Z3 на механических реле

1941г.



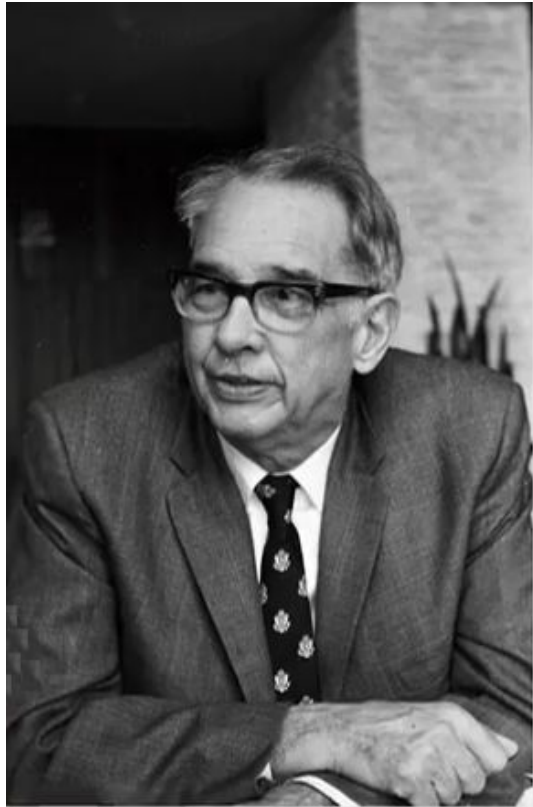
Средняя вычислительная производительность: 1 FLOPS

Потребление энергии: 4 кВт.

Масса: 1 т.

История. Первый в США электронный цифровой компьютер

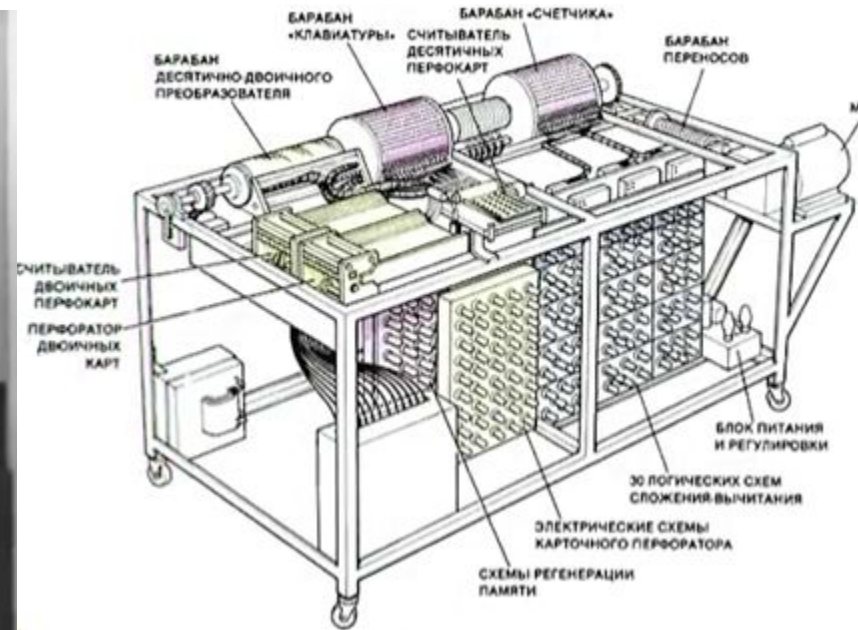
1942г.



Джон Атанасов,
Университет штата
Айова



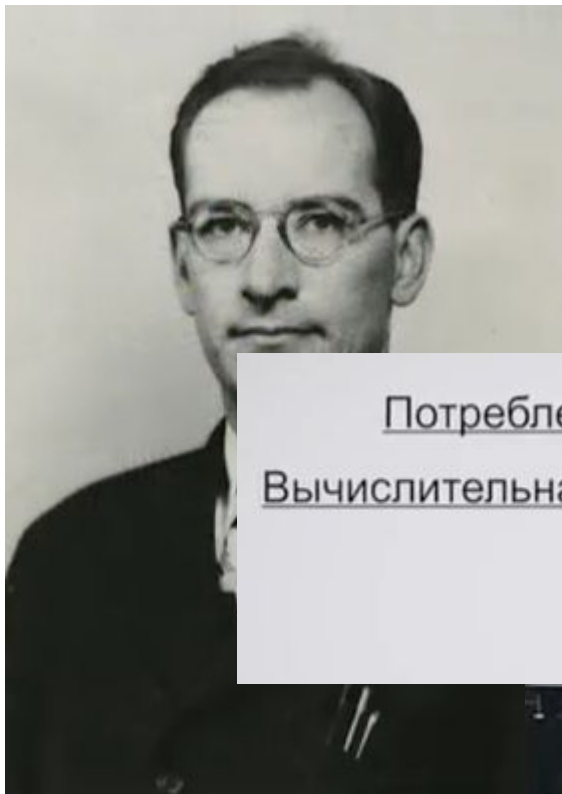
аспирант Клиффорд
Берри



Первый в США
электронный цифровой
компьютер

История. ЭВМ ЭНИАК

1945г.



Джон
Мокли

1945 год,
ЭВМ
ЭНИАК



Потребление энергии: 150 кВт.

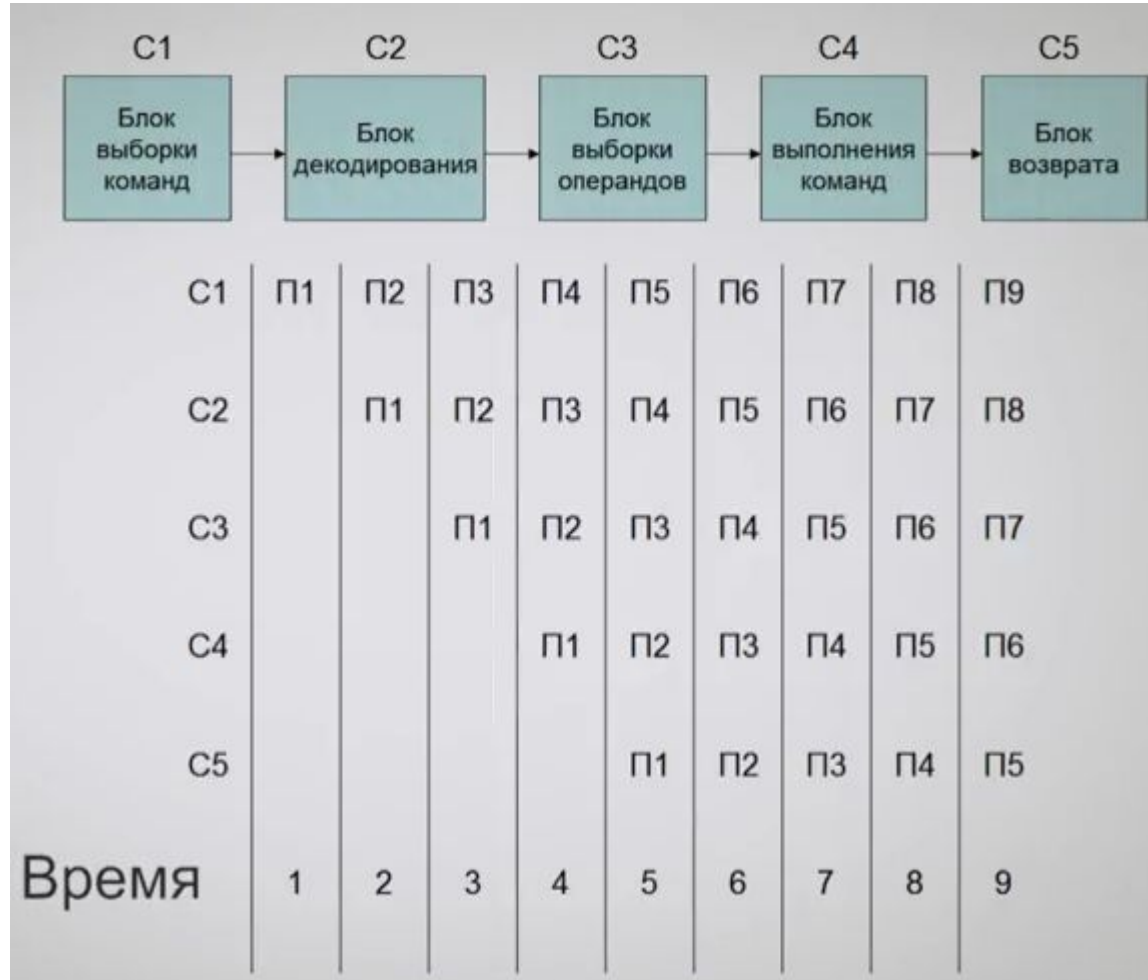
Вычислительная мощность: 300 операций умножения,
5 000 операций сложения в секунду.

Масса: 27 тонн.



Джон Преспер
Экерт,
Джон Уильям Мокли

История. Конвейер





ЭВМ «М-100»

Производительность: 100 тыс. операций/сек.



**Анатолий Иванович
Китов**

История. Конвейер с асинхронным процессором

1961, 1962гг.



IBM 7030

Производительность: 300 тыс. операций/сек.



Atlas

Производительность: 200-300 тыс. операций/сек.

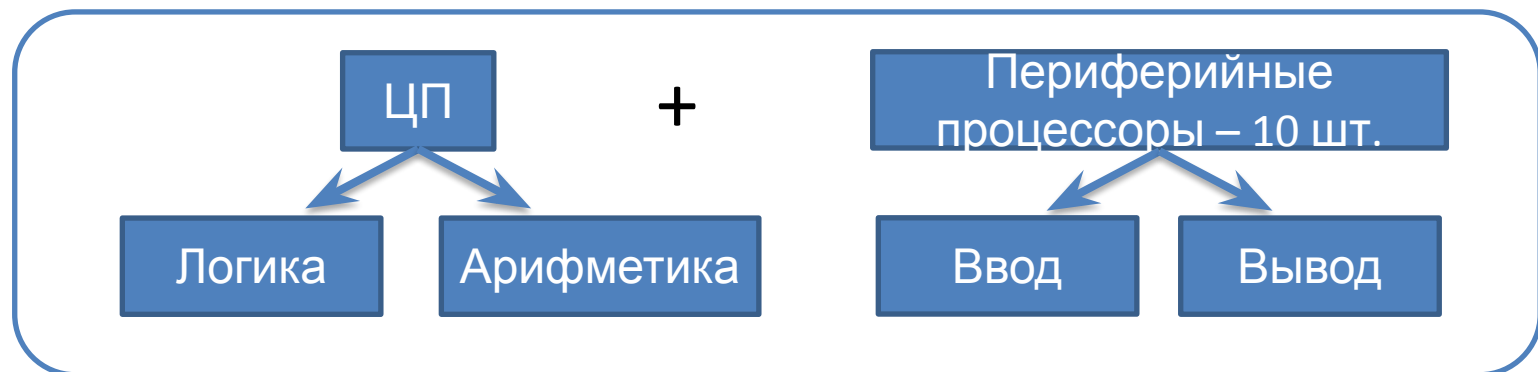
История. Компьютер с независимыми ФУ

1964г.



Фирма
Control Data Corporation,

Сеймур Крэй



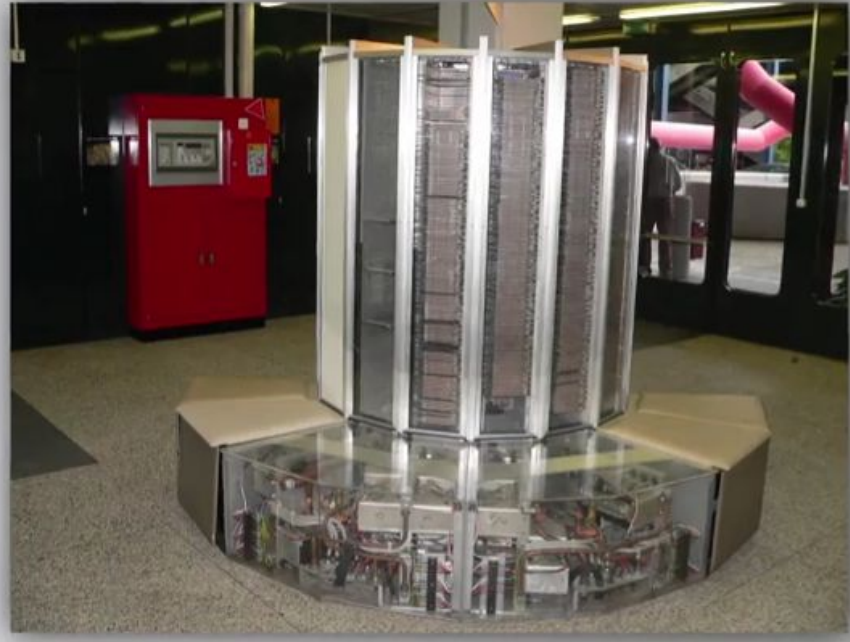
История. Компьютер с векторными операциями

1976г.

Компания Cray Research



CRAY1

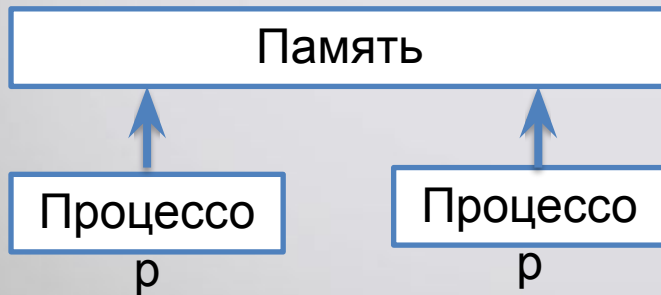


Производительность: 160 млн. операций/сек. (160MFlops)

12 ФУ
конвейерного типа
с операциями над векторами

Cray X-MP

Производительность: около 800 MFlops.



The ASCI Option Red Supercomputer

Производительность: 1,068 TFLOPS.



Supercomputer Earth Simulator

Производительность: 35,86 TFLOPS

Стоимость: \$500 млн.



5 тыс. процессоров

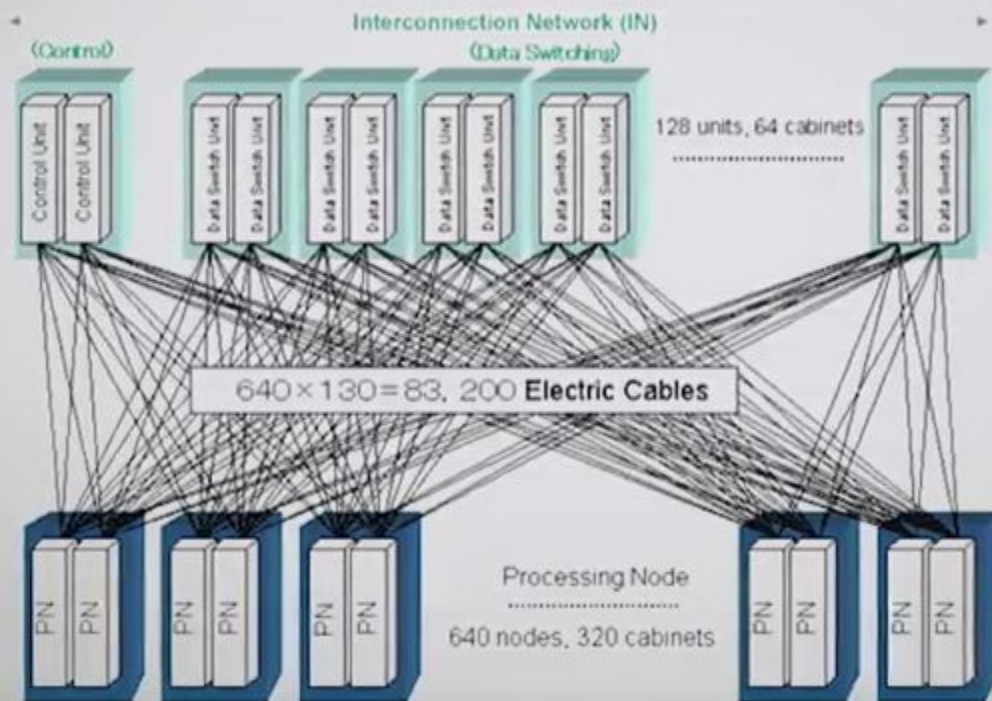
Назначение:
Изучение глобального
потепления

Supercomputer Earth Simulator

Производительность: 35,86 TFLOPS

Стоимость: \$500 млн.

Interconnection Network



Supercomputer Roadrunner

Производительность: 1,105 PFLOPS.

Состоял из:

- 6480 двухъядерных процессоров AMD Opteron;
- 12 960 процессоров IBM Cell 8i.



ТОП 500 суперкомпьютеров на июнь 2018

Распределение суперкомпьютеров в списке TOP500 по странам (июнь 2018 года)

страна	количество
Китай	206
США	124
Япония	36
Великобритания	22
ФРГ	21
Франция	18
Голландия	9
Южная Корея	7
Ирландия	7
Канада	6
Австралия	5
Индия	5
Италия	5
Польша	4
Россия	4
Саудовская Аравия	4
Швейцария	3
Швеция	3

ТОП 500 суперкомпьютеров на июнь 2018

Положение ↕	Rmax Rpeak (PFLOPS) ↕	Название ↕	Производитель ↕	Место страна, год ↕	Операционная система ↕
1 ▲	122.300 187.659	Summit	IBM	Oak Ridge National Laboratory 🇺🇸 США, 2018	Linux (RHEL)
2 ▼	93.015 125.436	Sunway TaihuLight	NRCP	National Supercomputing Center in Wuxi 🇨🇳 Китай, 2016	Linux (Raise)
3 ▲	71.610 119.194	Sierra	IBM	Lawrence Livermore National Laboratory 🇺🇸 США, 2018	Linux (RHEL)
4 ▼	61.445 100.679	Tianhe-2A	NUDT	National Supercomputing Center in Guangzhou 🇨🇳 Китай, 2013	Linux (Kylin)
5 ▲	19.880 32.577	AI Bridging Cloud Infrastructure	Fujitsu	National Institute of Advanced Industrial Science and Technology ● Япония, 2018	Linux
6 ▼	19.590 25.326	Piz Daint	Cray	Swiss National Supercomputing Centre 🇨🇭, 2016	Linux (CLE)

Суперкомпьютеры

Суперкомпьютер – это вычислительная система, обладающая предельными характеристиками по производительности среди имеющихся в каждый конкретный момент времени компьютерных систем.

Кластер – группа компьютеров, объединенных в локальную вычислительную сеть (ЛВС) и способных работать в качестве единого вычислительного ресурса. Предполагает более высокую надежность и эффективность, нежели ЛВС, и существенно более низкую стоимость в сравнении с другими типами параллельных вычислительных систем (за счет использования типовых аппаратных и программных решений).

Примеры

Суперкомпьютер СКИФ МГУ

(НИВЦ МГУ) 2008

Общее количество двухпроцессорных узлов 625
(1250 четырехядерных процессоров Intel Xeon
E5472 3.0 ГГц),



- Общий объем оперативной памяти – 5,5 Тбайт,
- Объем дисковой памяти узлов – 15 Тбайт,
- Операционная система Linux,
- Пиковая производительность 60 TFlops, быстродействие на тесте LINPACK 47 TFlops.

Примеры

Персональные мини-кластеры

T-Edge Mini - см. <http://www.t-platforms.ru/ru/temini.php>

- 4 двухпроцессорных узла на базе четырехядерных процессоров Intel Xeon (всего 32 вычислительных ядер)
- Оперативная память – до 128Гбайт
- Сеть передачи данных - Gigabit Ethernet или InfiniBand
- Операционная система - SUSE Linux Enterprise Server, RedHat Enterprise Linux или Microsoft Windows Compute Cluster Server 2003
- Пиковая производительность – 384 GFlops
- Размеры (см) - 57x33x76



Пути достижения параллелизма

Пути достижения параллелизма:

- независимость функционирования отдельных устройств ЭВМ;
- избыточность элементов вычислительной системы;
 - использование специализированных устройств;
 - дублирование устройств ЭВМ.

Режимы выполнения независимых частей программы:

- многозадачный режим (режим разделения времени);
- параллельное выполнение;
- распределенные вычисления.

Процессы, потоки, нити

Процесс (задача) - программа, находящаяся в режиме выполнения.

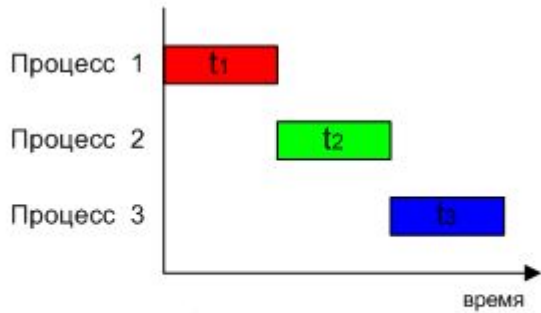
С каждым процессом связывается его **адресное пространство**, из которого он может читать и в которое он может писать данные.

Адресное пространство содержит:

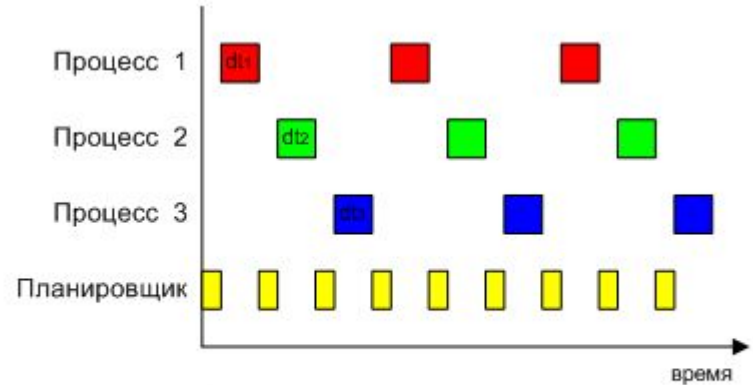
- саму программу
- данные к программе
- стек программы

С каждым процессом связывается набор **регистров**

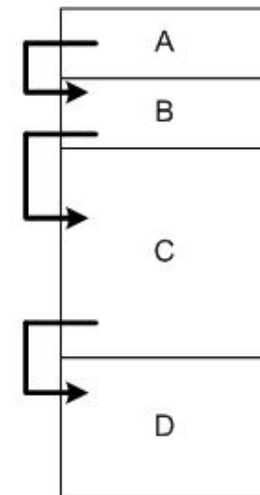
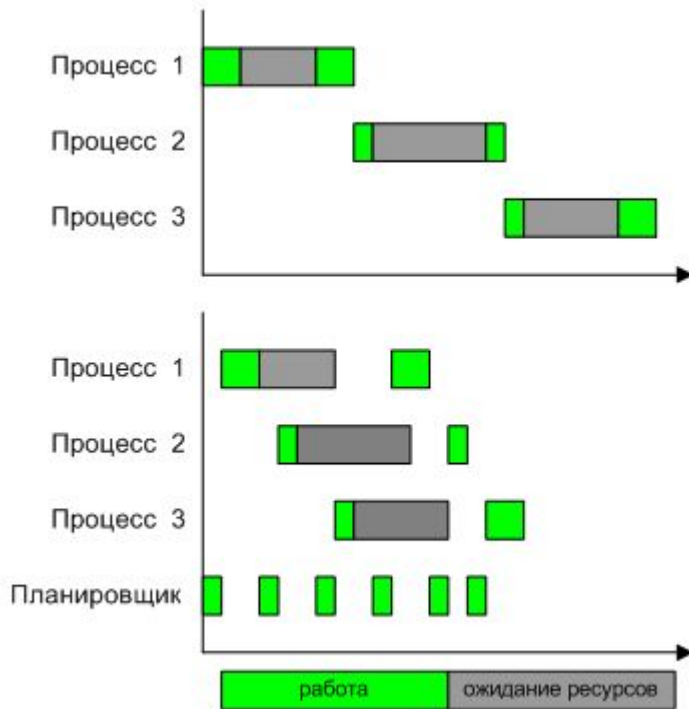
Процессы, потоки, нити



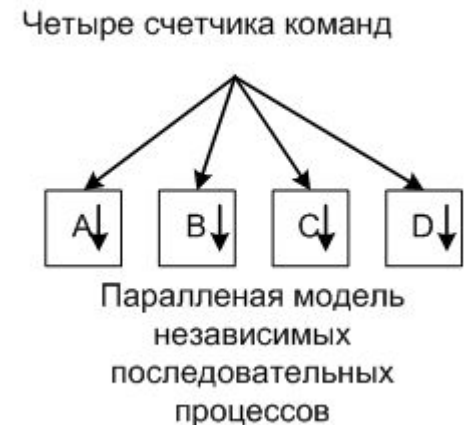
Последовательное выполнение трёх процессов на однопоточном процессоре



Работа каждого процесса на однопоточном процессоре разбита на временные промежутки



Четыре процесса в многозадачном режиме



Что работает быстрее?

Дано:

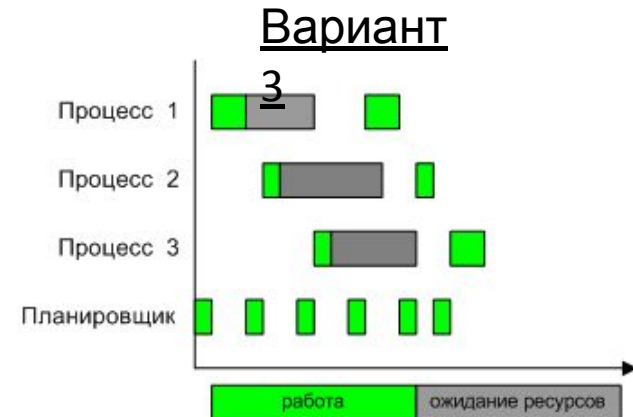
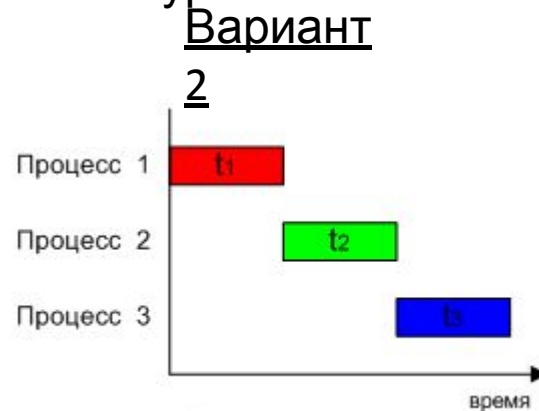
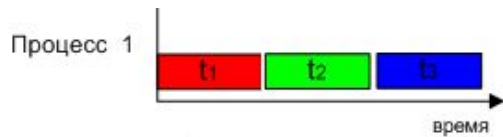
1 задача = Подзадача1 + Подзадача2 +
Подзадача3

$$t = t_1 + t_2 + t_3$$

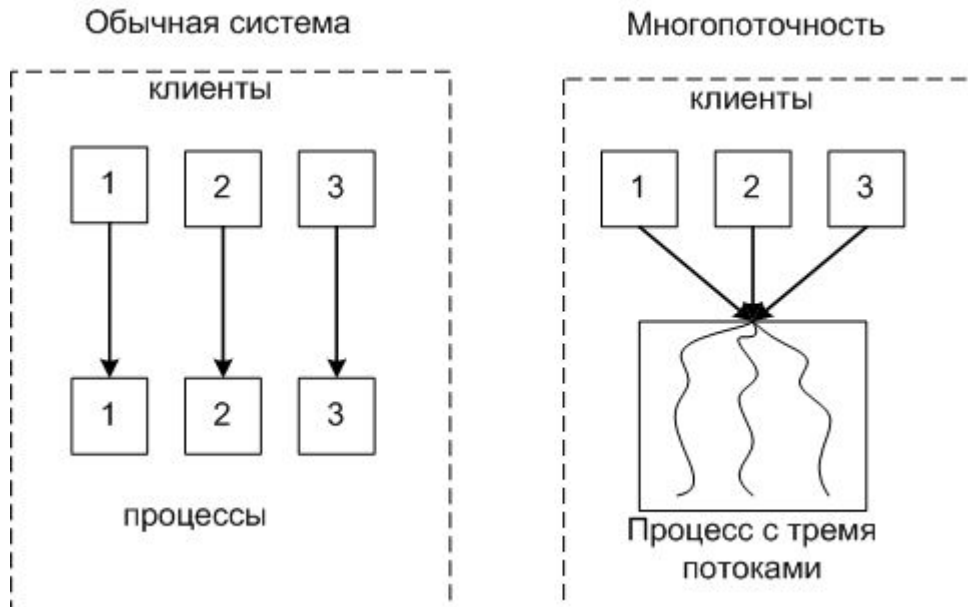
1 процессор (1 ядро) с архитектурой Фон-

Неймана

Вариант
3 процесса



Потоки (облегченные процессы)



С каждым потоком связывается:

- Счетчик выполнения команд
- Регистры для текущих переменных
- Стек
- Состояние

Потоки делят между собой элементы своего процесса:

- Адресное пространство
- Глобальные переменные
- Открытые файлы
- Таймеры
- Семафоры
- Статистическую информацию.

В остальном модель идентична модели процессов.

Процессы, потоки, нити

Преимущества использования потоков

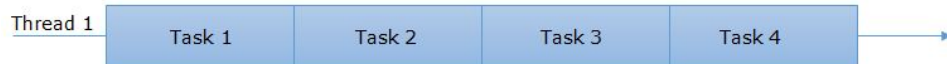
- Упрощение программы в некоторых случаях, за счет использования общего адресного пространства.
- Быстрота создания потока, по сравнению с процессом, примерно в 100 раз.
- Повышение производительности самой программы, т.к. есть возможность одновременно выполнять вычисления на процессоре и операцию ввода/вывода.

Пример: текстовый редактор с тремя потоками может одновременно взаимодействовать с пользователем, форматировать текст и записывать на диск резервную копию.

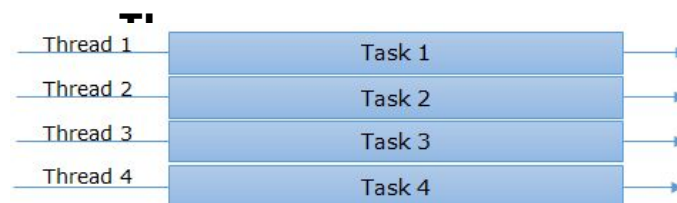
Синхронность и асинхронность потоков

Синхронная программная модель

Однопоточность



Многопоточность



Асинхронная программная модель

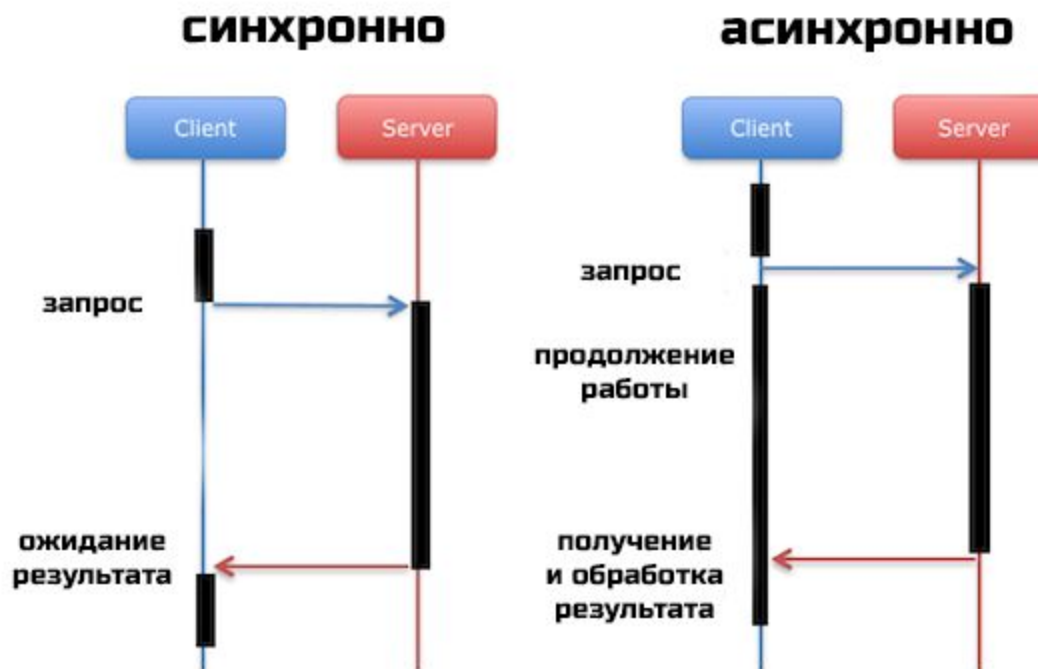
Однопоточность



Многопоточность



Синхронность и асинхронность потоков



Объекты синхронизации и проблемы потоков

1. Критическая секция (CriticalSection)
2. Взаимоисключение (мьютекс, mutex - от MUTual EXclusion)
3. Событие (Event)
4. Семафор

Проблемы потоков

Условия гонки [Race condition];

Конкуренция за ресурс [Resource contention];

Вечная блокировка [Deadlock];

Голодание [Starvation] ;

Инверсия приоритетов [Priority Inversion];

Неопределенность и справедливость [Non-deterministic and Fairness].

Последовательная обработка

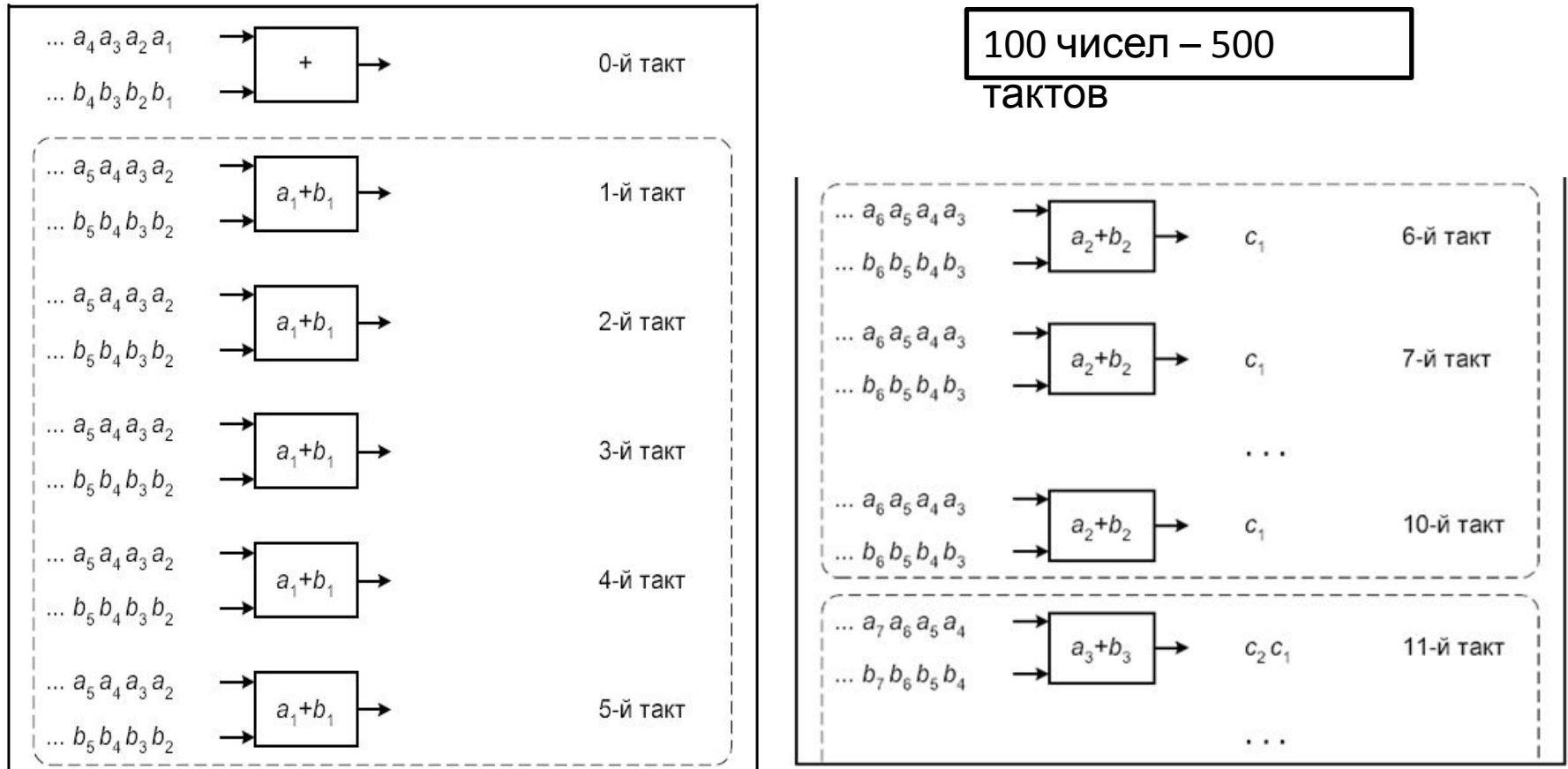


Рис. Суммирование векторов $C = A + B$ с помощью последовательного устройства, выполняющего одну операцию за пять тактов

Параллельная обработка

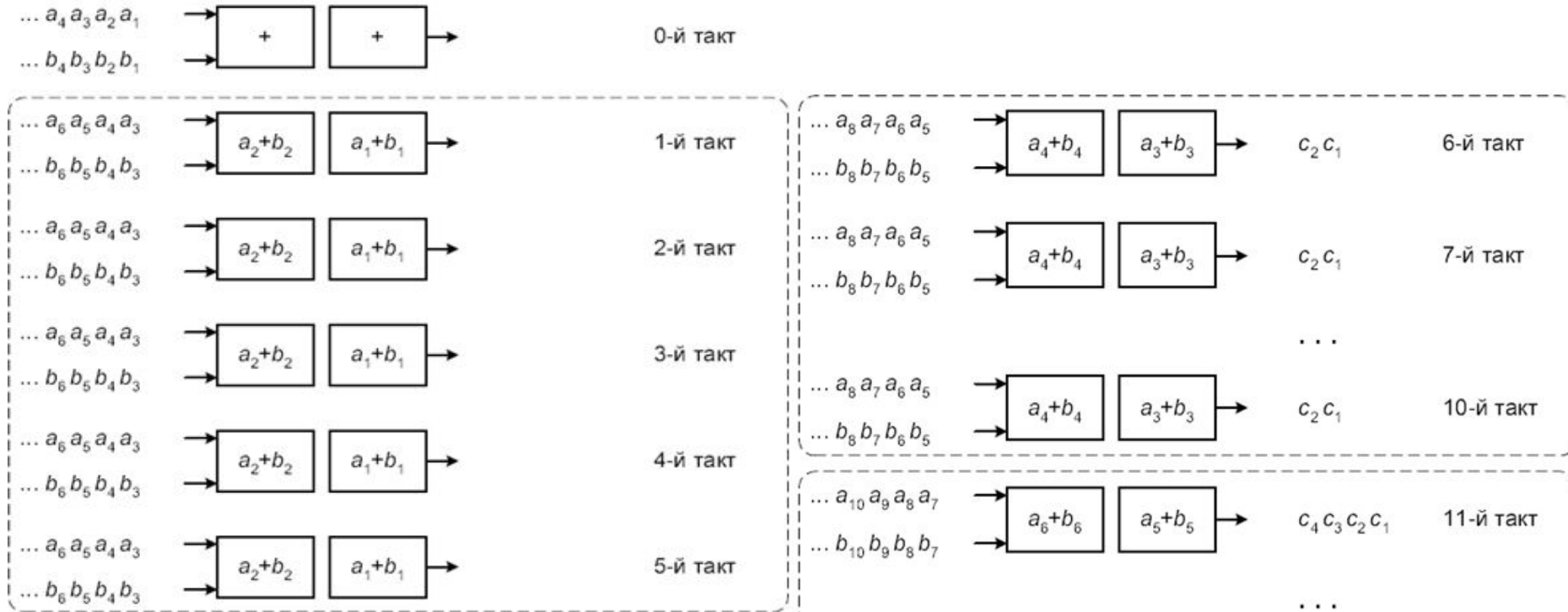


Рис. Суммирование векторов $C = A + B$ с помощью двух одинаковых последовательных устройств, выполняющих операцию за пять тактов каждое

100 чисел – 250

ТАКТОВ

Конвейерная обработка

Определения.

- 1) Ступень конвейера.
- 2) Длина конвейера.

Ступени

конвейера:



5 тактов

Длина вектора	Число тактов
1	5
2	6
3	7
4	8
...

100 чисел – 104

такта

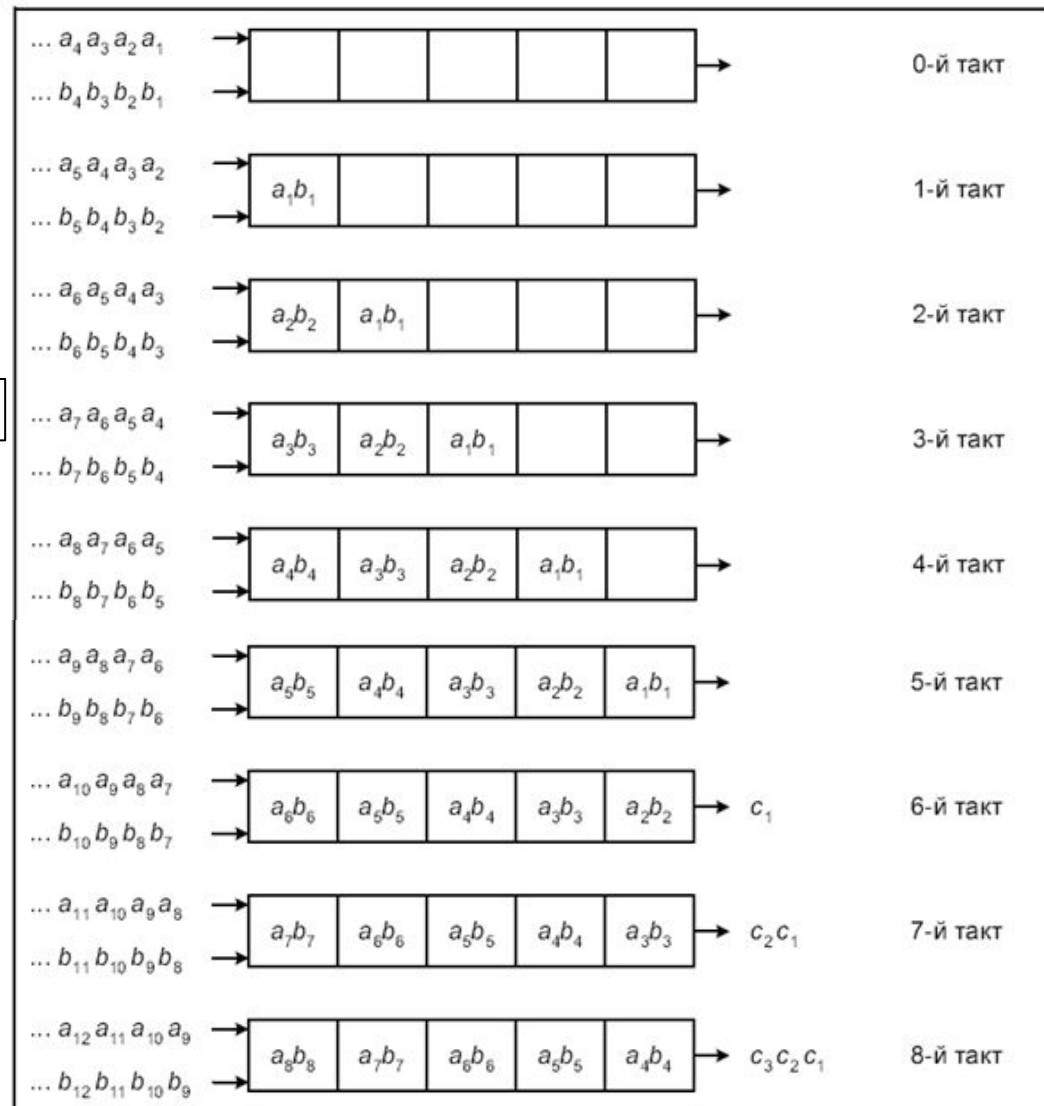


Рис. Суммирование векторов $C = A + B$ с помощью конвейерного устройства. Каждая из пяти ступеней конвейера срабатывает за один такт

Конвейерная

обработка

В лучшем

случае:
Пусть:

n – число операций;
 l – длина конвейера.

Тогда:

Время выполнения
операции $T = n + l - 1$

В общем

случае:

Пусть:

n – число операций;
 l – длина конвейера;
 σ - погрешность.

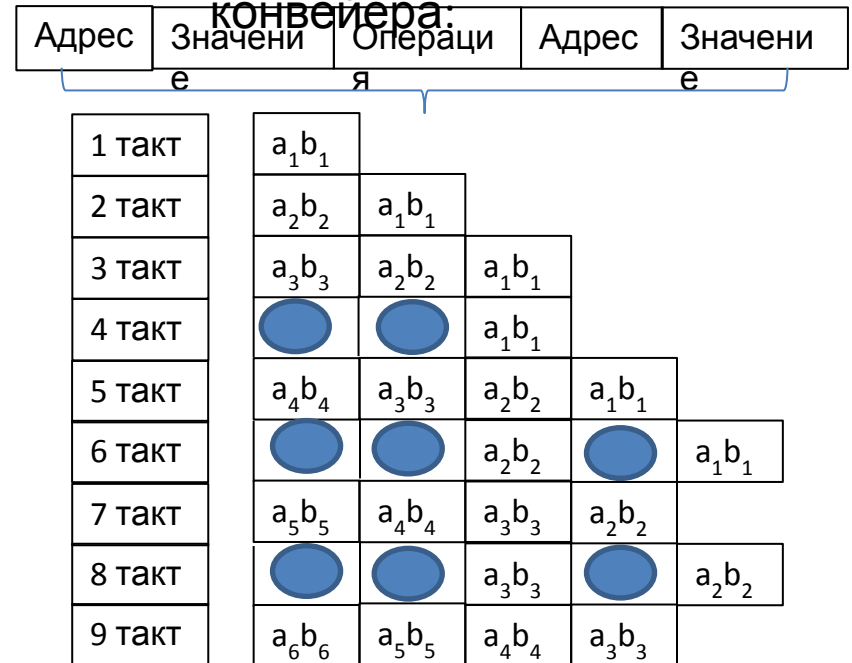
Тогда:

Время выполнения
операции $T = \sigma + n + l - 1$

Пузырь

и

конвейера:



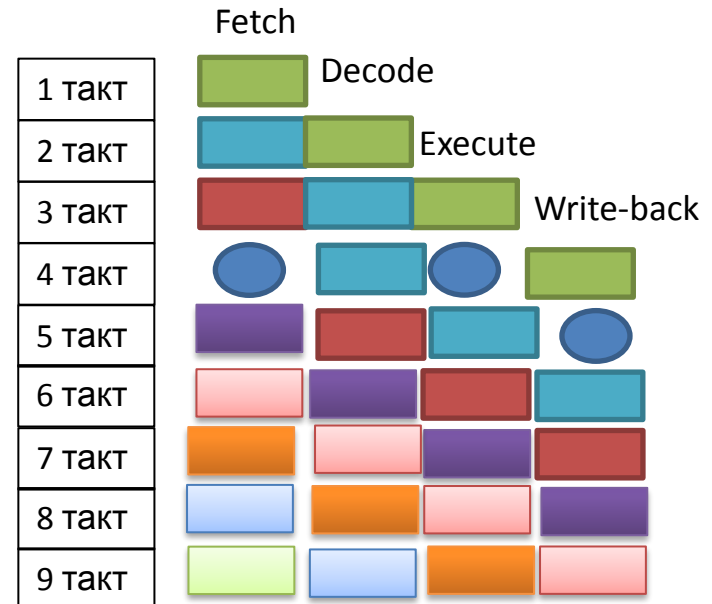
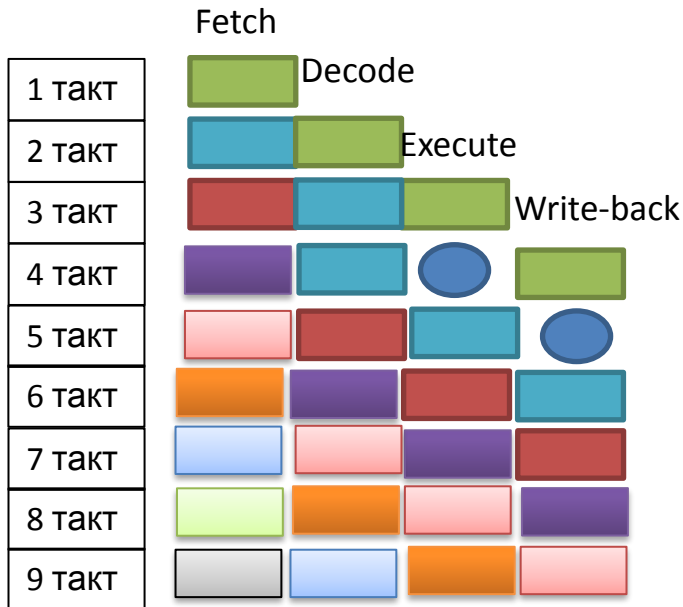
Длина вектора	Число тактов
1	6
2	8
3	10
4	12

$$\sigma = n$$

Конвейерная обработка

Ступени конвейера

- Получение (Fetch)
- Раскодирование (Decode)
- Выполнение (Execute)
- Запись результата (Write-back)



Конвейерная обработка

Определение.

Эффективность
конвейера

$$E = \frac{n}{t} = \frac{n}{[(\sigma + l + n - 1) \tau]} = \frac{1}{\left[\tau + (\sigma + l - 1) \frac{\tau}{n} \right]}$$

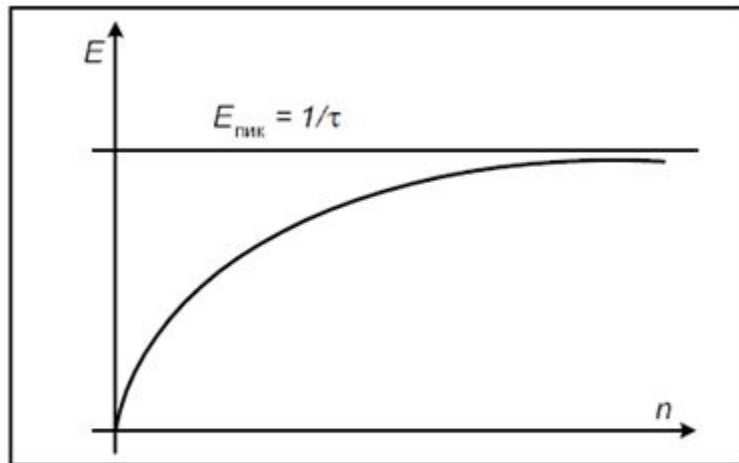
где:

n – число операций;

l – длина конвейера;

σ – погрешность;

τ – время работы одного такта.



Считаем:

Пусть $l=4$, $n=6$, $\tau=1$, $\sigma=1$

На последовательном:

$$E=6/24=0,25;$$

На конвейере без пузырей:

$$E=6/(0+4+6-1)=6/9=0,67$$

На конвейере с пузырем:

$$E=6/(1+4+6-1)=6/10=0,6$$

Рис. Зависимость производительности конвейерного устройства от длины входного набора данных

Технико-эксплуатационные характеристики ЭВМ

- быстродействие;
- разрядность;
- формы представления чисел;
- номенклатура и характеристики запоминающих устройств;
- номенклатура и характеристики устройств ввода-вывода информации;
- типы и характеристики внутренних и внешних интерфейсов;
- наличие многопользовательских режимов;
- типы и характеристики, используемых ОС;
- система команд и их структура;
- функциональные возможности программного обеспечения и его наличие;
- программная совместимость с другими типами ЭВМ;
- срок эксплуатации;
- условия эксплуатации;
- характеристики надежности;
- состав и объем профилактических работ;
- стоимостные характеристики;
- совокупная стоимость владения

Классификации компьютеров

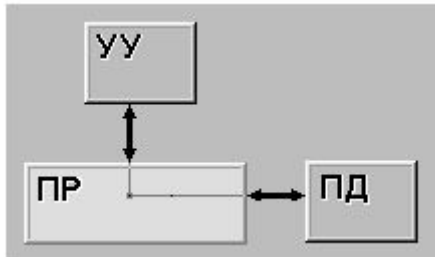
- принцип действия:
(цифровые, аналоговые и гибридные);
- назначение:
(универсальные, проблемно-ориентированные, специализированные);
- размеры и вычислительная мощность:
(суперкомпьютеры и остальные);
- особенности архитектуры:

Классификация по Флинну

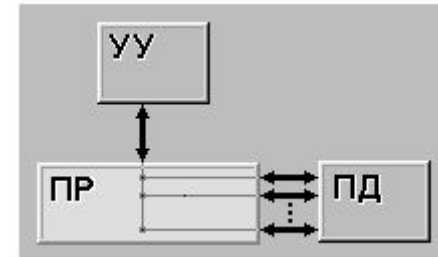
	Одиночный поток команд (Single Instruction)	Множество потоков команд (Multiple Instruction)
Одиночный поток данных (Single Data)	SISD (ОКОД)	MISD (МКОД)
Множество потоков данных (Multiple Data)	SIMD (ОКМД)	MIMD (МКМД)

Классификации компьютеров по Флинну

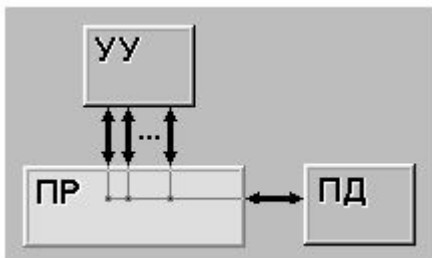
- **SISD** (Single Instruction, Single Data)



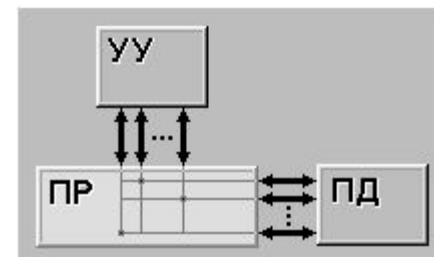
- **SIMD** (Single Instruction, Multiple Data)



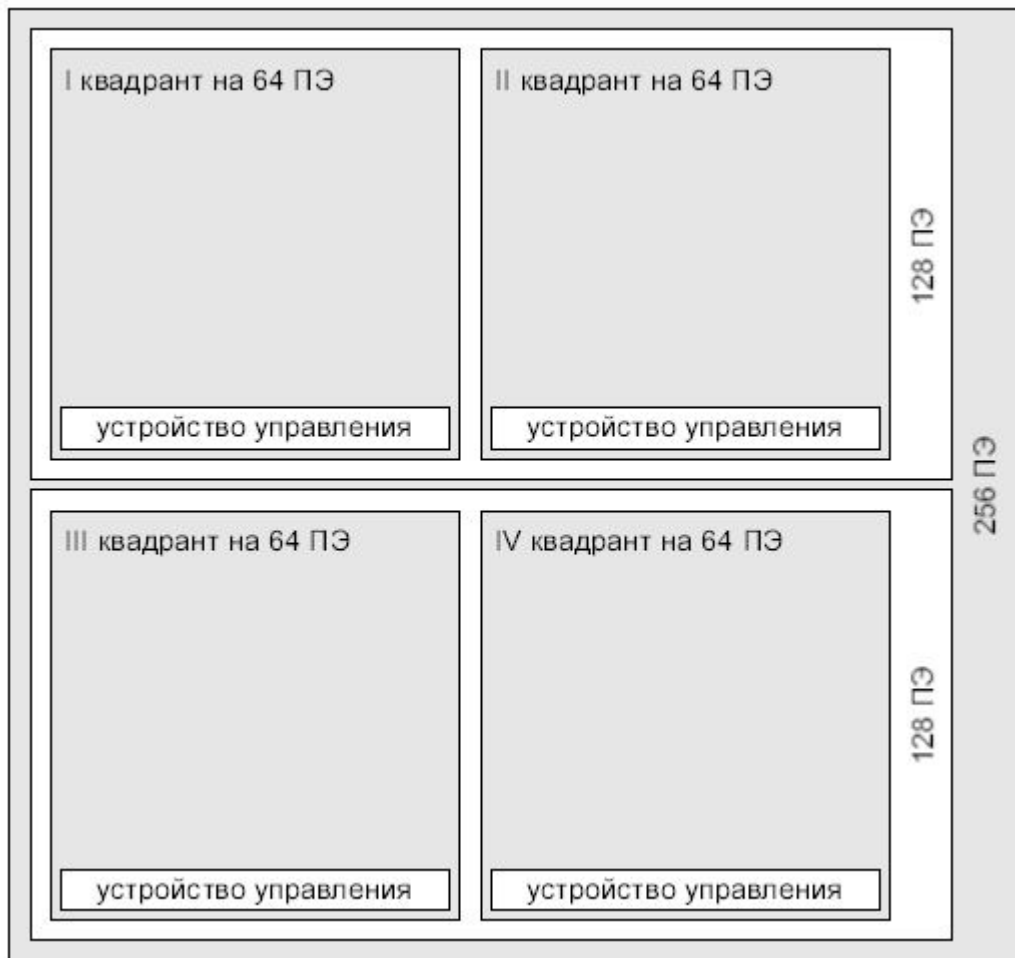
- **MISD** (Multiple Instruction, Single Data)



- **MIMD** (Multiple Instruction, Multiple Data)



ILLIAC IV



Начало работ – 1967г.
Первый квадрат – 1972г.
Наладка системы –
1975г.

Эксплуатация – до

Время такта по проекту – 40нс

Реальное время – 80нс

Пиковая производительность:
по проекту – 1 миллиард
опер./с

реальная – 50 миллионов
опер

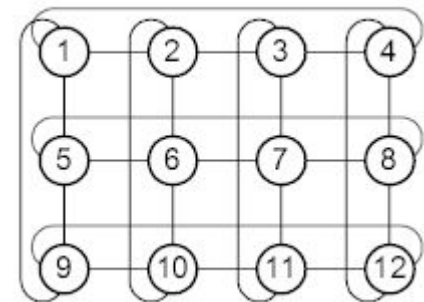
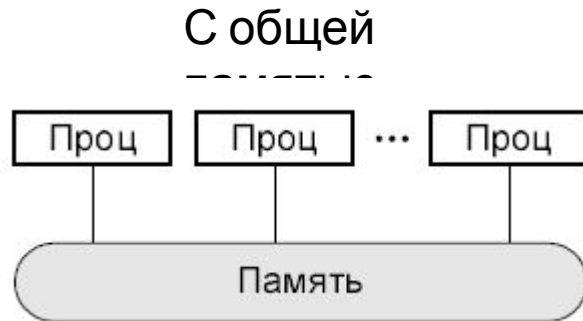


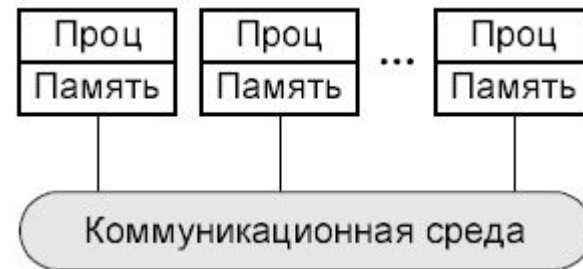
Рис. Проект матричной системы ILLIAC IV

MIMD

Параллельные компьютеры MIMD



С распределенной
памятью



massive parallel processing (**MPP**)

Кластеры

Кластер – группа компьютеров, объединенных в локальную вычислительную сеть (ЛВС) и способных работать в качестве единого вычислительного ресурса.

Пример:

- 1) Symmetric Multi Processors (SMP);
- 2) Parallel Vector Processor (PVP) (Cray T90);

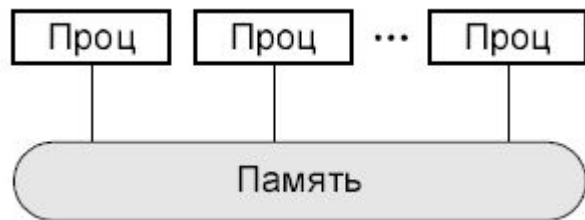
Кластеры и ВС:

- Кластеры \subset распределенные ВС;
- Кластер для users – одна система;
- Кластер – быстрая связь между узлами;
- Кластер – узкая специализация

Две основные задачи параллельных вычислений

2) Поиск методов разработки эффективного ПО

С общей

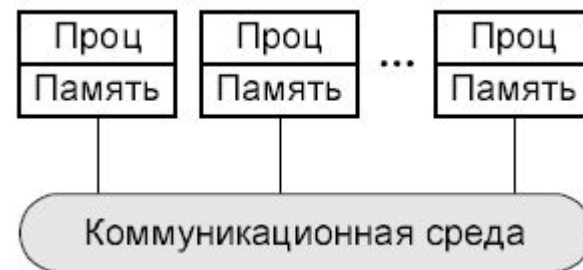


Проблемы:

- Сложность объединения процессоров под единой ОЗУ;
- Низкая производительность.

1) Построение вычислительных систем с максимальной производительностью

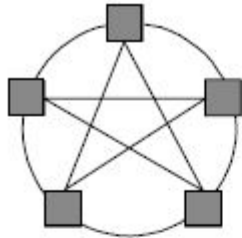
С распределенной



Проблемы:

- Накладные расходы;
- Сложность параллельных разработок;
- Системы обмена сообщения PVM и MPI – не просты.

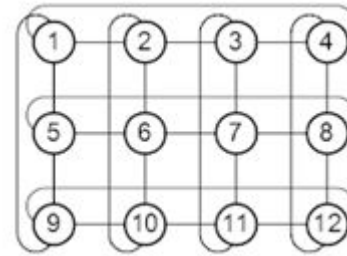
Примеры топологий сети передачи данных



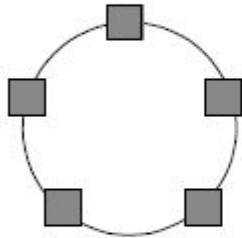
1) Полный граф



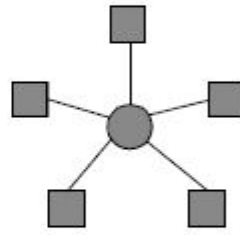
2) Линейка



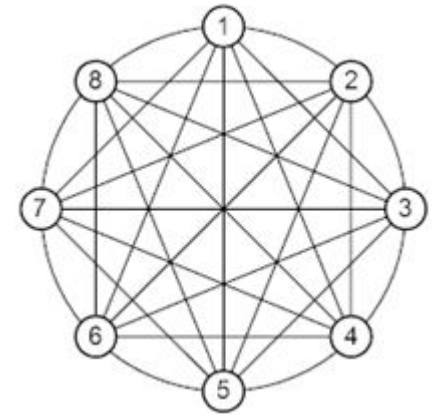
7) Тор



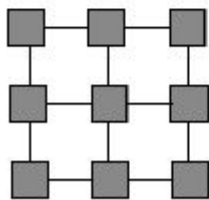
3) Кольцо



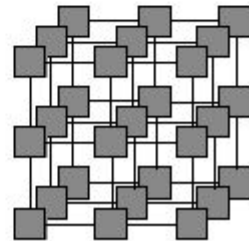
4) Звезда



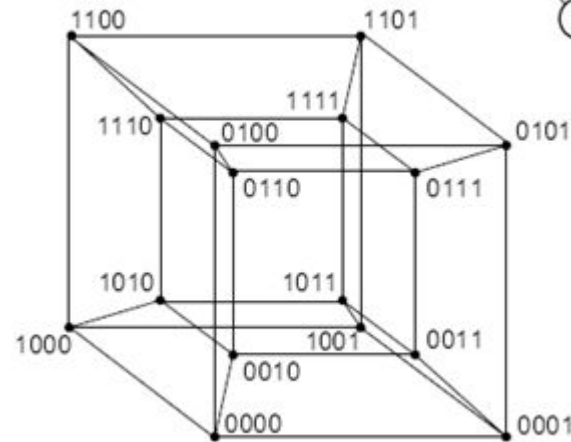
8) Полная связь



5) 2-мерная решетка



6) 3-мерная решетка



9) Гиперкуб

Характеристики топологии сети

- Диаметр;
- Связность;
- Ширина бинарного деления;
- Стоимость.

Топология	Диаметр	Ширина	Связность	Стоимость
Полный граф	1	$p^2/4$	$p-1$	$p(p-1)/2$
Звезда	2	1	1	$p-1$
Полное двоичное дерево	$2\log_2((p+1)/2)$	1	1	$p-1$
Линейка	$p-1$	1	1	$p-1$
Кольцо		2	2	p
Решетка N=2	$\lfloor p/2 \rfloor$		2	
Решетка-тор N=2	$2(\sqrt{p} - 1)$	\sqrt{p}	4	$2(p - \sqrt{p})$
Гиперкуб	$2\lfloor \sqrt{p}/2 \rfloor$	$2\sqrt{p}/2$	$\log_2 p$	$(p \log_2 p)/2$