

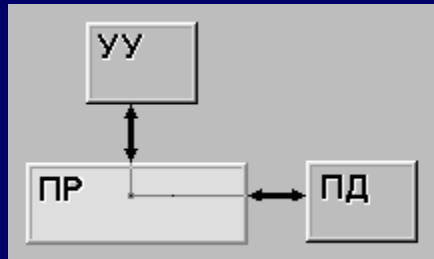
# **Архитектура параллельных вычислительных систем**

## **Лекция 2**

# 1. Классификация Флинна (1966)

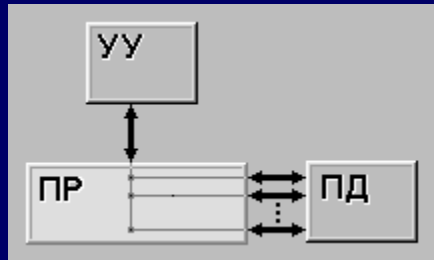
- Поток инструкций (команд)
- Поток данных
- **ОКОД (SISD** — Single Instruction, Single Data stream) — компьютер с единственным ЦП
- **ОКМД (SIMD** — Single Instruction, Multiple Data stream) — матричная обработка данных
- **МКОД (MISD** — Multiple Instruction, Single Data stream) — ?
- **МКМД (MIMD** — Multiple Instruction, Multiple Data stream) — множество процессоров одновременно выполняют различные последовательности команд над своими данными

# 1.1 ОКОД (SISD)



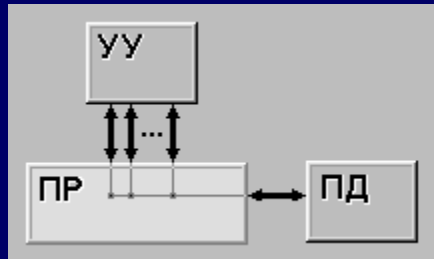
К этому классу относятся, прежде всего, классические последовательные машины, или иначе, машины фон-неймановского типа. В таких машинах есть только один поток команд, все команды обрабатываются последовательно друг за другом и каждая команда инициирует одну операцию с одним потоком данных. Не имеет значения тот факт, что для увеличения скорости обработки команд и скорости выполнения арифметических операций может применяться конвейерная обработка.

# 1.2 ОКМД (SIMD)



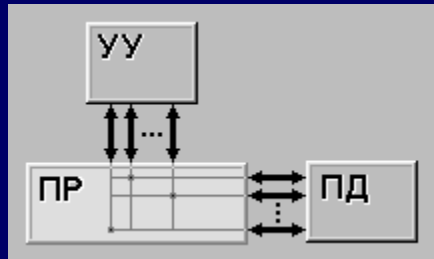
В архитектурах подобного рода сохраняется один поток команд, включающий, в отличие от предыдущего класса, векторные команды. Это позволяет выполнять одну арифметическую операцию сразу над многими данными - элементами вектора.

# 1.3 МКОД (MISD)



Определение подразумевает наличие в архитектуре многих процессоров, обрабатывающих один и тот же поток данных. Однако ни Флинн, ни другие специалисты в области архитектуры компьютеров до сих пор не смогли представить убедительный пример реально существующей вычислительной системы, построенной на данном принципе. Будем считать, что пока данный класс пуст.

# 1.4 МКМД (MIMD)



Этот класс предполагает, что в вычислительной системе есть несколько устройств обработки команд, объединенных в единый комплекс и работающих каждое со своим потоком команд и данных.

## 2.1 Классификация MIMD-систем

**MIMD**

```
graph TD; MIMD[MIMD] --> Shared[Системы с общей оперативной памятью]; MIMD --> Distributed[Системы с распределённой оперативной памятью]; Shared --> UMA[UMA — система с однородным доступом в память (Uniform Memory Access)]; Shared --> NUMA[NUMA — системы с неоднородным доступом в память (Non-Uniform Memory Access)]; Distributed --> MPP[MPP — (Massively Parallel Processors) — процессоры с массовым параллелизмом]; Distributed --> COW[COW — (Cluster Of Workstations — кластер рабочих станций)];
```

**Системы с общей оперативной памятью**

**UMA** — система с однородным доступом в память (Uniform Memory Access)

**NUMA** — системы с неоднородным доступом в память (Non-Uniform Memory Access)

**SMP** — симметричная мультипроцессорная система (symmetric multiprocessor)

**Системы с распределённой оперативной памятью**

**MPP** — (Massively Parallel Processors) — процессоры с массовым параллелизмом

**COW** — (Cluster Of Workstations — кластер рабочих станций)

# Иерархия MIMD-систем

Системы делятся по принципу организации работы с ОЗУ

**В системе с общей оперативной памятью** имеется ОЗУ, и любой процессорный элемент имеет доступ к любой точке общего ОЗУ, то есть любой адрес может быть исполнительным для любого процессора

## **Виды систем с общей оперативной памятью**

- UMA — Uniform Memory Access
  - SMP — Symmetrical MultiProcessing
- NUMA — Non-Uniform Memory Access

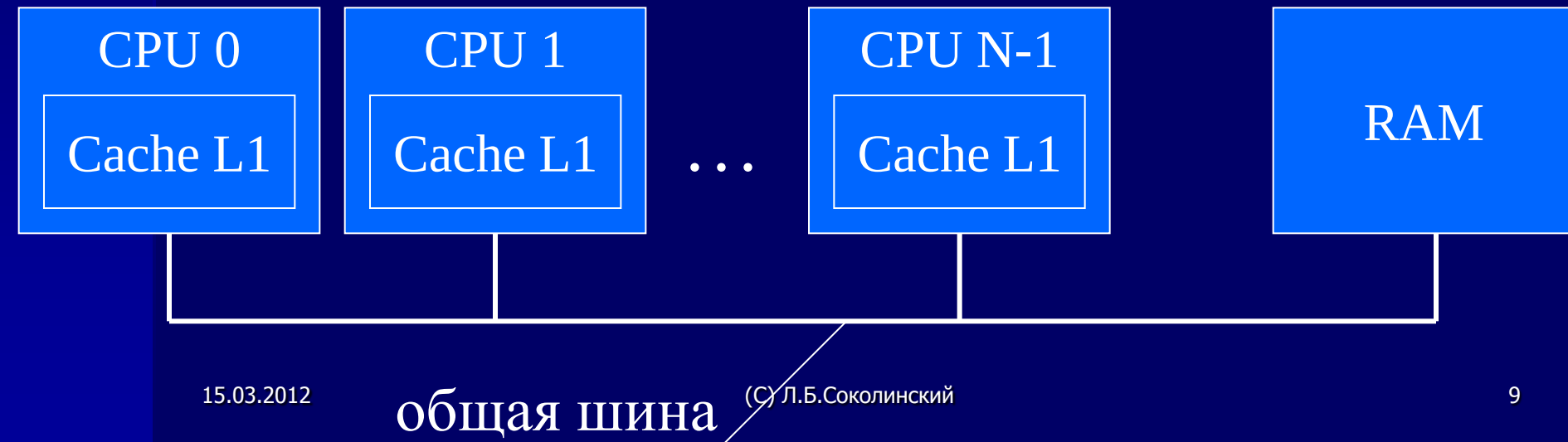


# UMA и SMP системы

**UMA:** характеристики доступа любого процессорного элемента в любую точку ОЗУ не зависят от конкретного элемента и адреса (Все процессоры равноценны относительно доступа к памяти).

**SMP-системы** являются подвидом UMA-систем.

## Схема SMP-системы



# SMP-системы

## Преимущества SMP

- Простота реализации

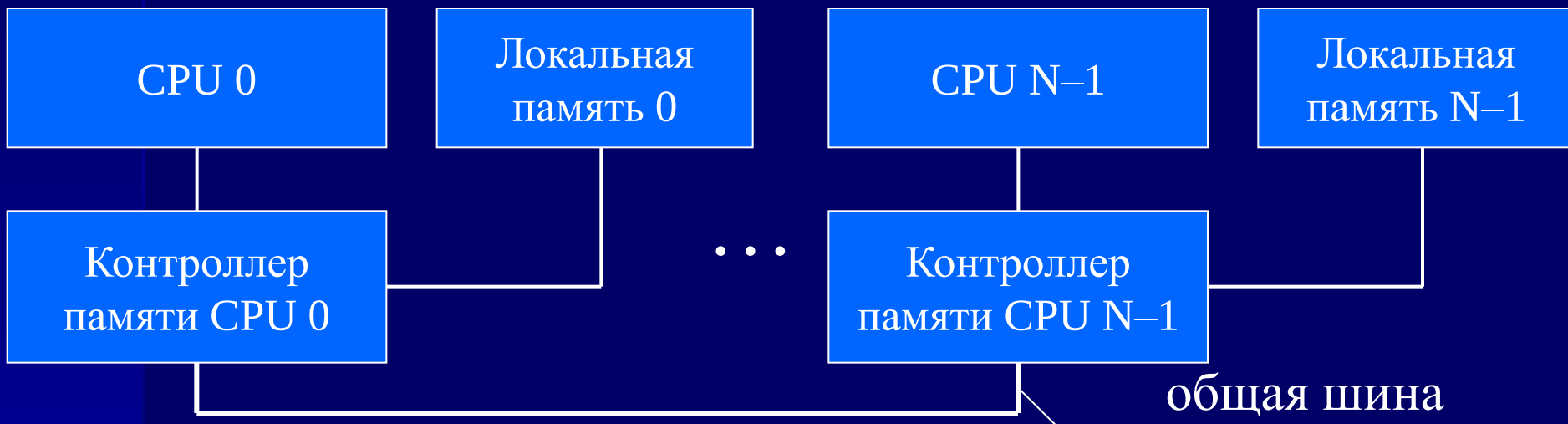
## Недостатки SMP

- Задержки при доступе к памяти
- Система с явной централизацией — общая шина является «узким местом»
- Проблема синхронизации кэша (решением является кэш-память с отслеживанием)
- Ограничение на количество процессорных элементов (как следствие централизации)

# NUMA-системы

**NUMA:** Процессорные элементы работают на общем адресном пространстве, но характеристики доступа процессора к ОЗУ зависят от того, куда он обращается.

## Схема NUMA-системы



При обращении не к своей памяти контроллер выбрасывает запрос на общую шину, целевой контроллер его принимает и возвращает результат

# NUMA-системы

В NUMA-системах остаётся проблема синхронизации кэша. Существует несколько способов её решения:

- использовать процессоры без кэша (использовать только Cache L2)
- использовать **ccNUMA** (NUMA-системы с когерентными кэшами)

ccNUMA сложнее, но позволяет строить системы из сотен процессорных элементов.

# NUMA-системы

## Преимущества NUMA

- Степень параллелизма выше, чем в SMP

## Недостатки NUMA

- Централизация (ограничение ресурсом шины)
- Использование когерентных кэшей загружает шину служебной информацией

## Недостатки ccNUMA

- Загрузка общей шины служебной информацией

# Иерархия MIMD-систем

**Системы с распределённой оперативной памятью** представляются как объединение компьютерных узлов, каждый из которых состоит из процессора и ОЗУ, непосредственный доступ к которой имеет только «свой» процессорный элемент. Класс наиболее перспективных систем.

## **Виды систем с распределённой оперативной памятью**

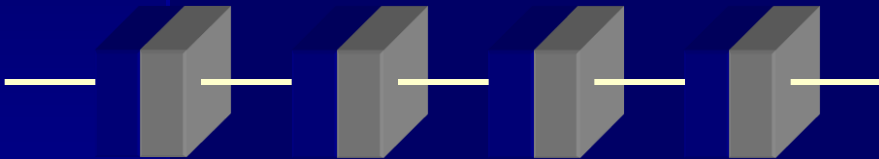
- MPP — Massively Parallel Processors
- COW — Cluster Of Workstations

# MPP-системы

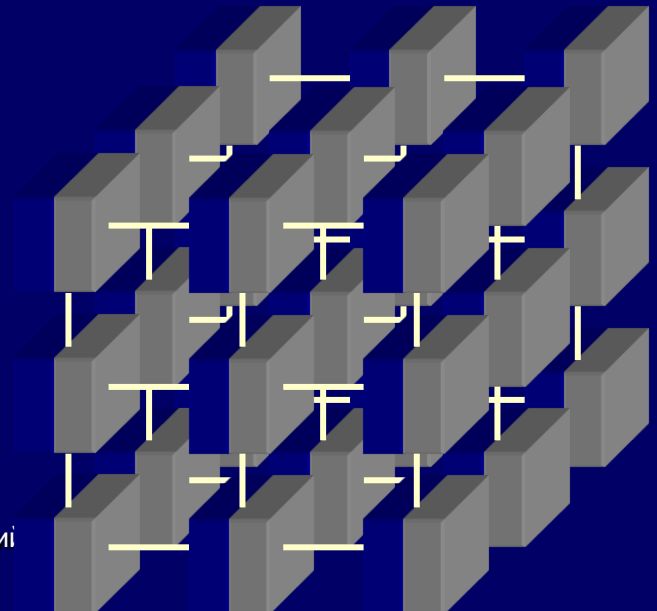
**MPP** — Специализированные дорогостоящие ВС. Эти компьютеры могут выстраиваться, процессорные элементы могут объединяться в различные топологии: макроконвейер, n-мерный гиперкуб и др.

## Примеры топологий MPP-систем

Макроконвейер:



3-мерный гиперкуб:



Процессорный элемент с  
локальной памятью

Межэлементные коммуникации,  
определяющие топологию

15.03.2012

(С) Л.Б.Соколинский

# МРР-системы

## Преимущества МРР

- Высокая эффективность при решении определённого класса задач

## Недостатки МРР

- Высокая стоимость
- Узкая специализация



# COW-системы

Кластеры имеют две ориентации на использование:

- Кластер как вычислительный узел  
(высокопроизводительная система)
- Кластеры, которые обеспечивают надёжность  
(сохранение работоспособности при возможном  
снижении производительности)

Для сопряжения вычислительных узлов в кластере используются специализированные сети.

Исторически кластеры представлялись в виде объединения системных блоков. Современные кластеры — специализированные системы, которые имеют в своём составе сотни и тысячи вычислительных узлов.

# COW-системы

## Преимущества COW

- Высокая эффективность при решении широкого круга задач

## Проблемы COW

- Тепловыделение
- Топология