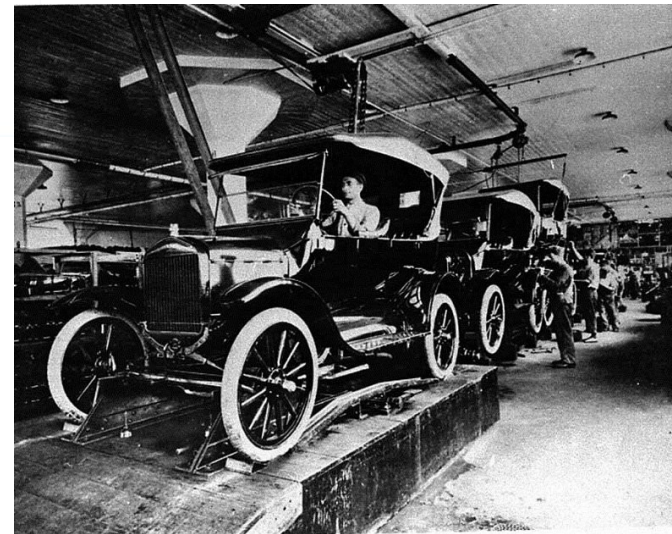
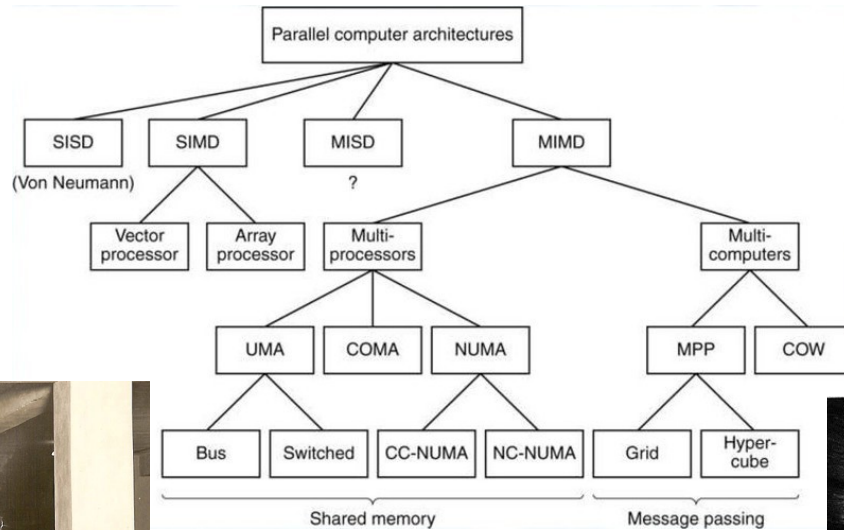


Архитектуры параллельных ВС (структурные аспекты)



Определения

Параллельная вычислительная система - вычислительная система, у которой имеется по меньшей мере более одного устройства управления или более одного центрального обрабатывающего устройства, которые работают одновременно.

Головкин А.С.

Определения

Параллелизм - воспроизведение в нескольких копиях некоторой аппаратной структуры, что позволяет достигнуть повышения производительности за счет одновременной работы всех элементов структуры, осуществляющих решение различных частей этой задачи.

Коуги П.М.

Параллелизм - способность к частичному совмещению или одновременному выполнению операций.

Хокни, Джессхоуп

Многопроцессорность и многопоточность

Многопроцессорность (Мультипроцессорность, Многопроцессорная обработка, Multiprocessing) — использование двух или большего количества физических процессоров в одной компьютерной системе.

Устройство называется многопроцессорным, если в его составе используется два или более физических процессора.

Многопотóчность (Multithreading) — свойство платформы (процессора, операционной системы, виртуальной машины и т. д.) или приложения, состоящее в том, что процесс, порождённый в операционной системе, может состоять из нескольких потоков, выполняющихся «параллельно», то есть без предписанного порядка во времени. При выполнении некоторых задач такое разделение может достичь более эффективного использования ресурсов вычислительной машины.

Такие *потоки* называют также *потоками выполнения (thread of execution)*; иногда называют «*нитями*» (*thread*).

Области применения ПВС

- предсказания погоды, климата и глобальных изменений в атмосфере;
- науки о материалах;
- построение полупроводниковых приборов;
- сверхпроводимость;
- структурная биология;
- разработка фармацевтических препаратов;
- генетика;
- квантовая хромодинамика;
- астрономия;
- транспортные задачи;
- гидро- и газодинамика;
- управляемый термоядерный синтез;

Области применения ПВС

- эффективность систем сгорания топлива;
- геоинформационные системы;
- разведка недр;
- наука о мировом океане;
- распознавание и синтез речи;
- распознавание изображений;
- военные цели.

Ряд областей применения находится на стыках соответствующих наук

Производительность параллельных вычислительных систем

Пиковая производительность - величина, равная произведению пиковой производительности одного процессора на число таких процессоров в данной машине.

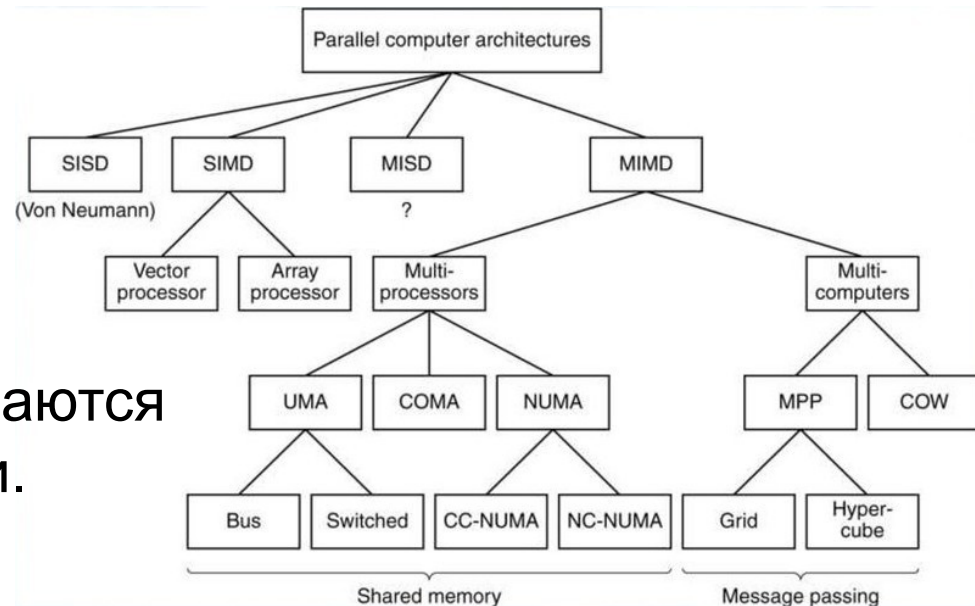
Закон Амдала (Amdahl's law) — В случае, когда задача разделяется на несколько частей, суммарное время её выполнения на параллельной системе не может быть меньше времени выполнения самого медленного фрагмента.

Согласно этому закону, ускорение выполнения программы за счёт распараллеливания её инструкций на множестве вычислителей ограничено временем, необходимым для выполнения её последовательных инструкций.

Классификация Флинна

Предложена в 1966 г.

Основана на том, как в машине увязываются команды с обрабатываемыми данными.



Поток - последовательность элементов (команд или данных), выполняемая или обрабатываемая процессором.

Классификация Флинна

ОКОД - один поток команд, один поток данных

SISD - *single instruction stream / single data stream*

ОКМД - один поток команд, много потоков данных

SIMD - *single instruction stream / multiple data stream*

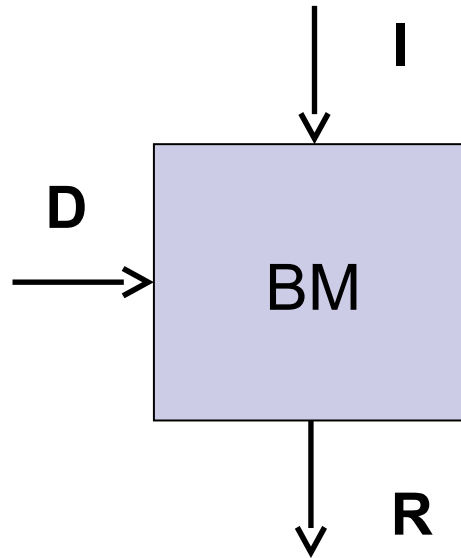
МКОД - много потоков команд, один поток данных

MISD - *multiple instruction stream / single data stream*

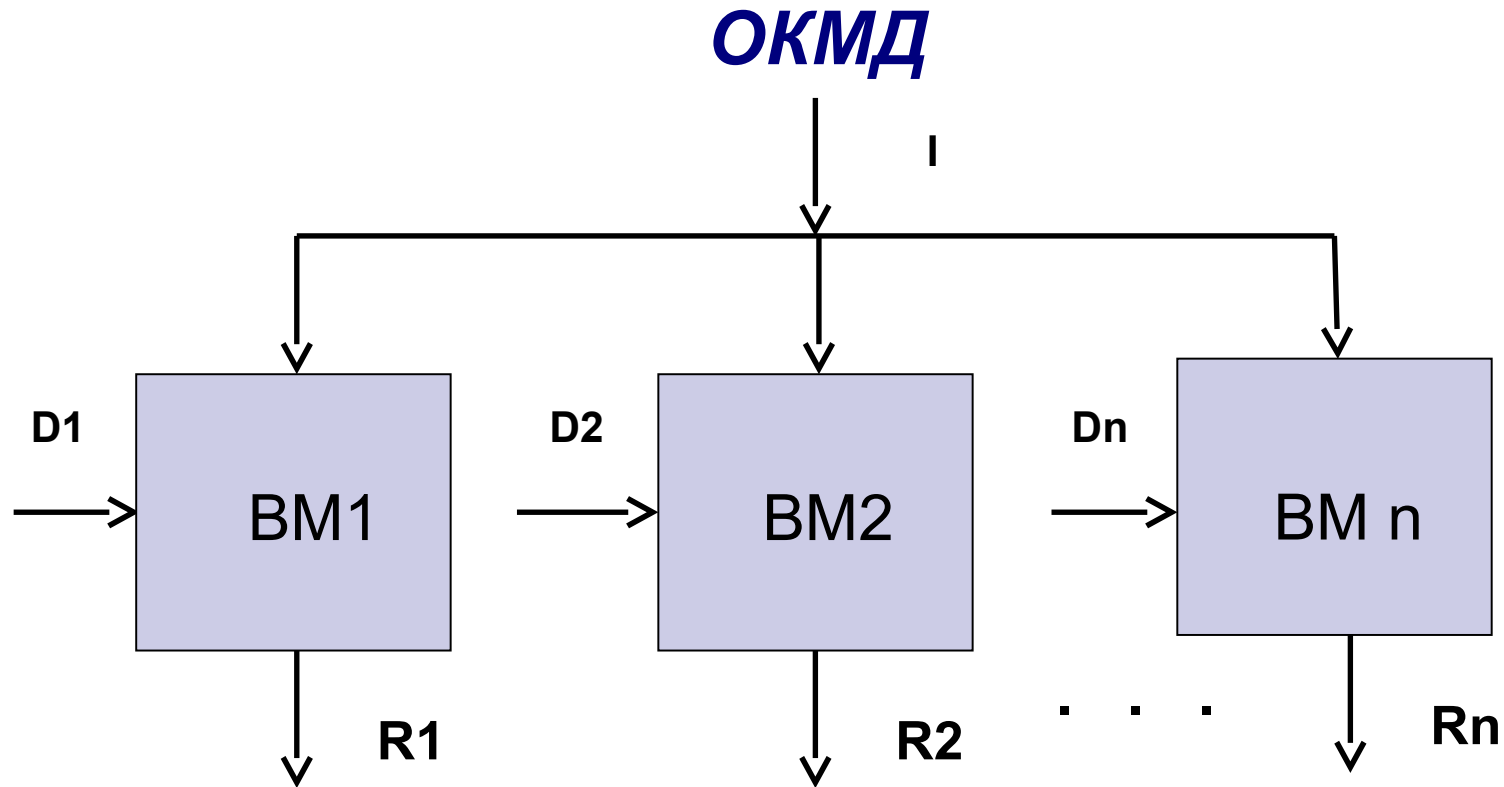
МКМД - много потоков команд, много потоков данных

MIMD - *multiple instruction stream / multiple data stream*

ОКОД

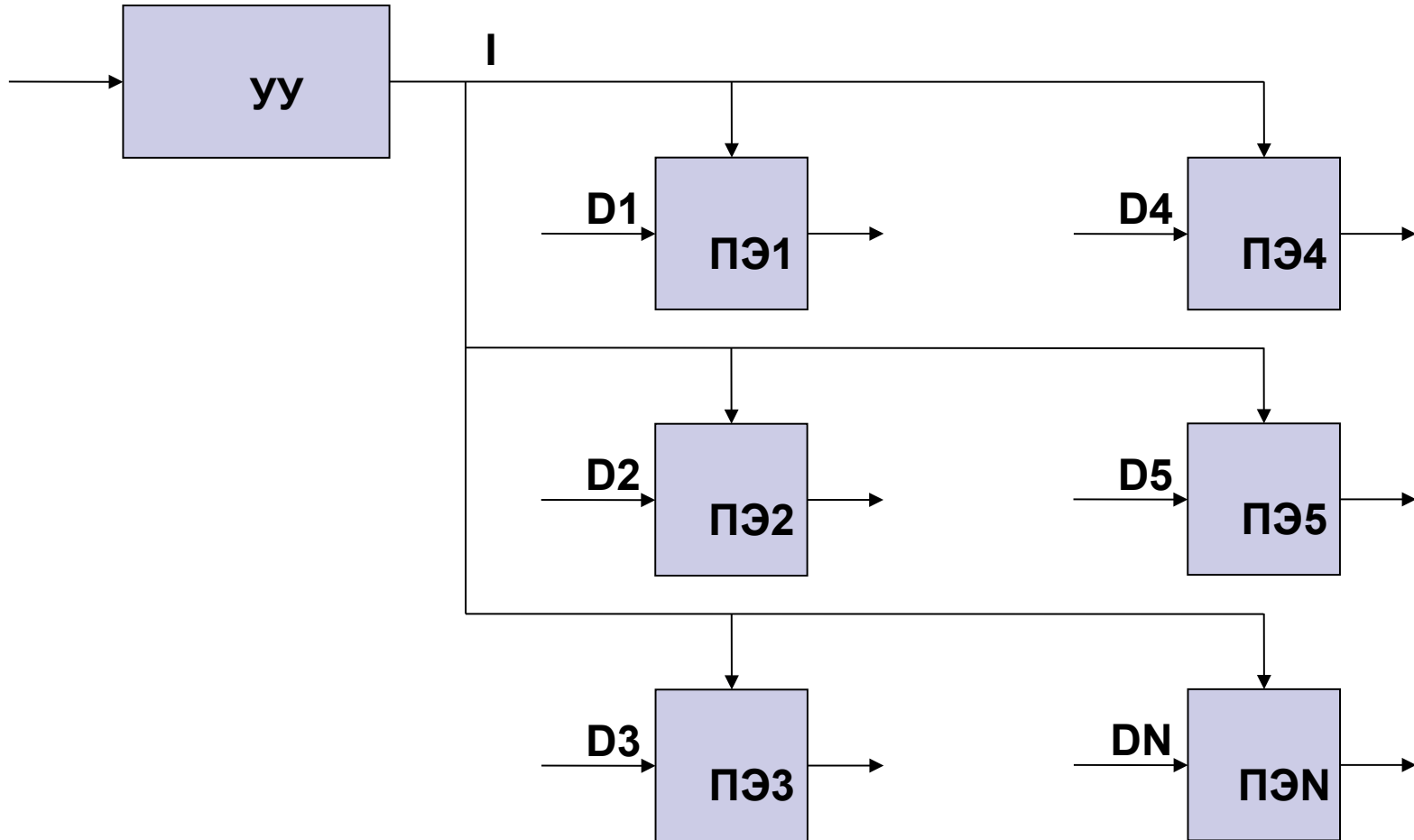


Классические последовательные машины, или иначе, машины фон-неймановского типа. В таких машинах есть только один поток команд, все команды обрабатываются последовательно друг за другом и каждая команда инициирует одну операцию с одним потоком данных.



Один поток команд, включает векторные (массивные) команды. Это позволяет выполнять одну арифметическую операцию сразу над многими данными - элементами вектора (массива).

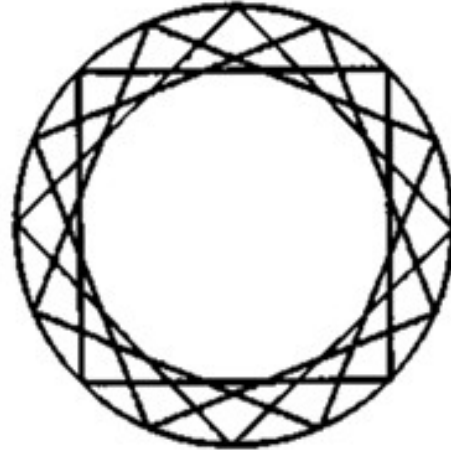
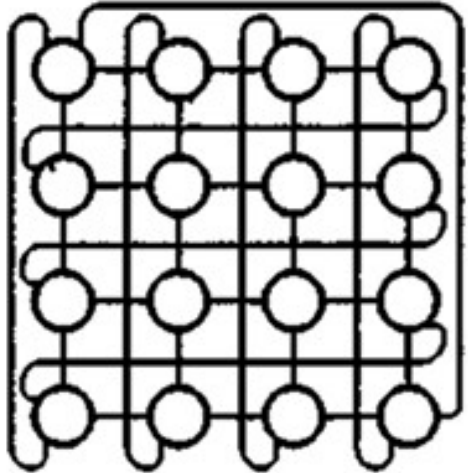
ОКМД



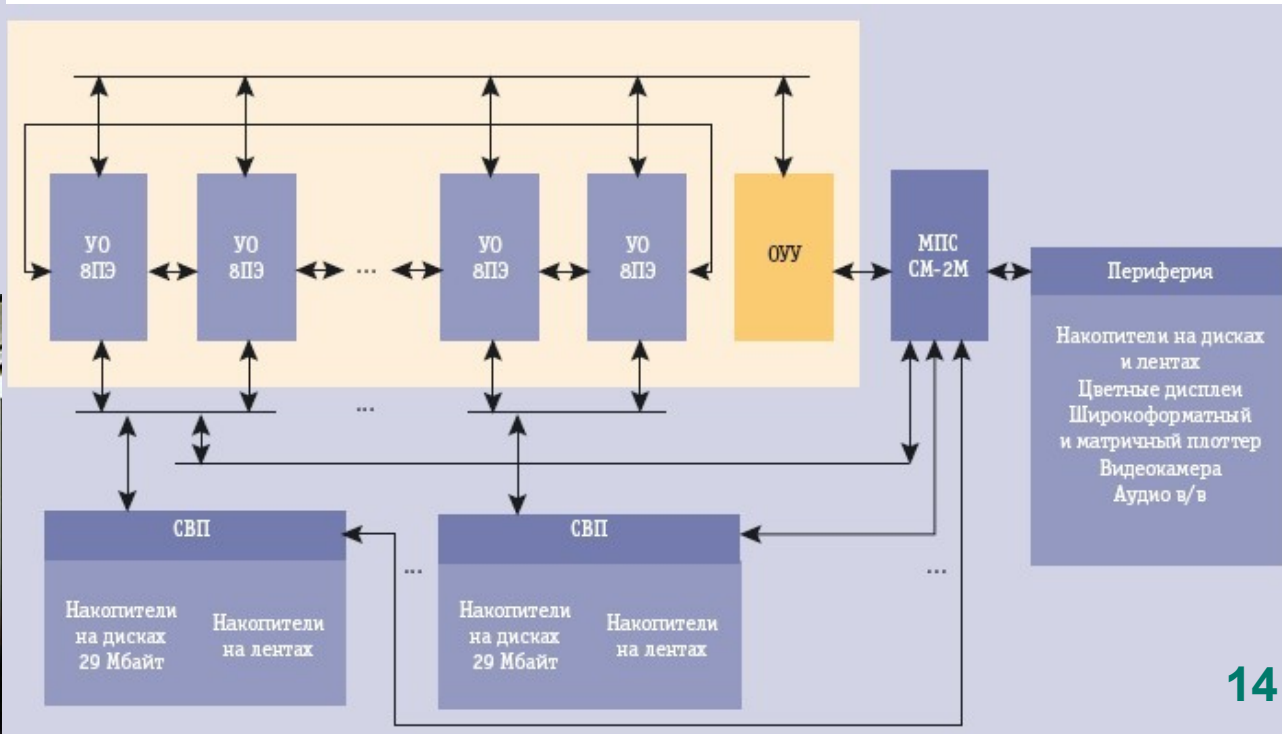
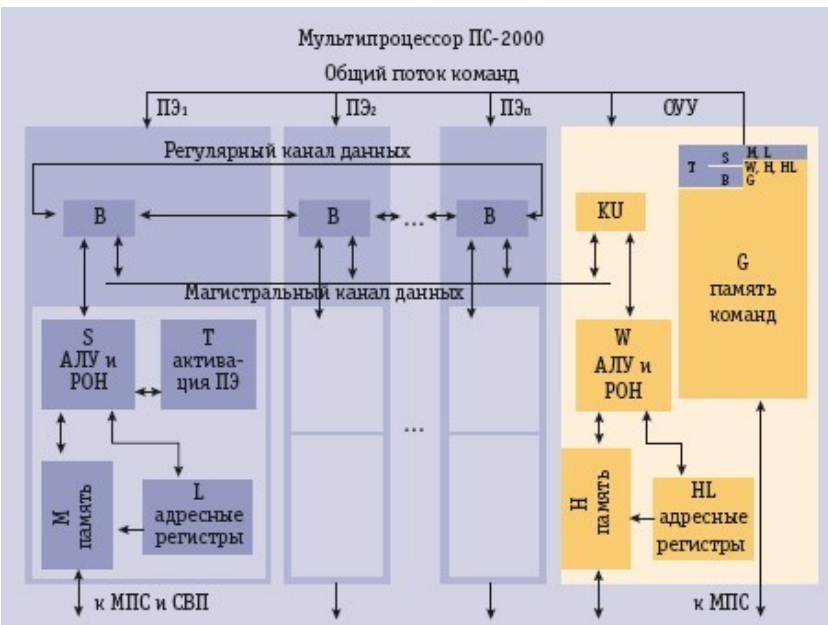
ILLIAC - IV

Систолическая матрица - массив связанных с соседями специфических вычислительных элементов.

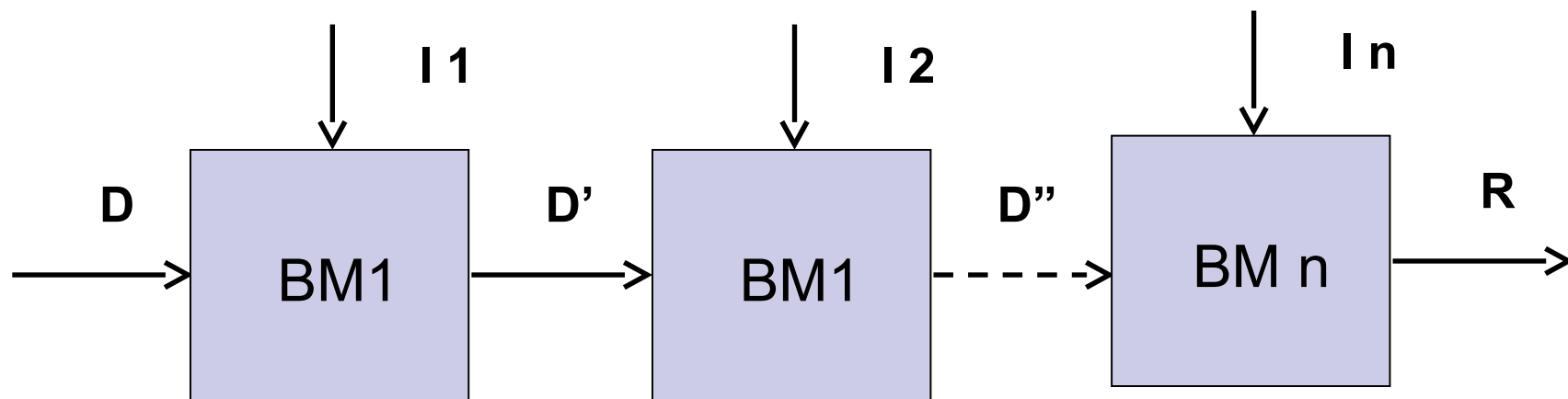
Часто использовалась и в многомашинных системах.



ПС-2000 (перенастраиваемая структура)



МКОД



Подразумевает наличие в архитектуре многих процессоров, обрабатывающих один и тот же поток данных.

Трудно представить убедительный пример реально существующей вычислительной системы, построенной на данном принципе. Иногда относят конвейерные машины к данному классу, однако окончательного консенсуса нет (parallel.ru).

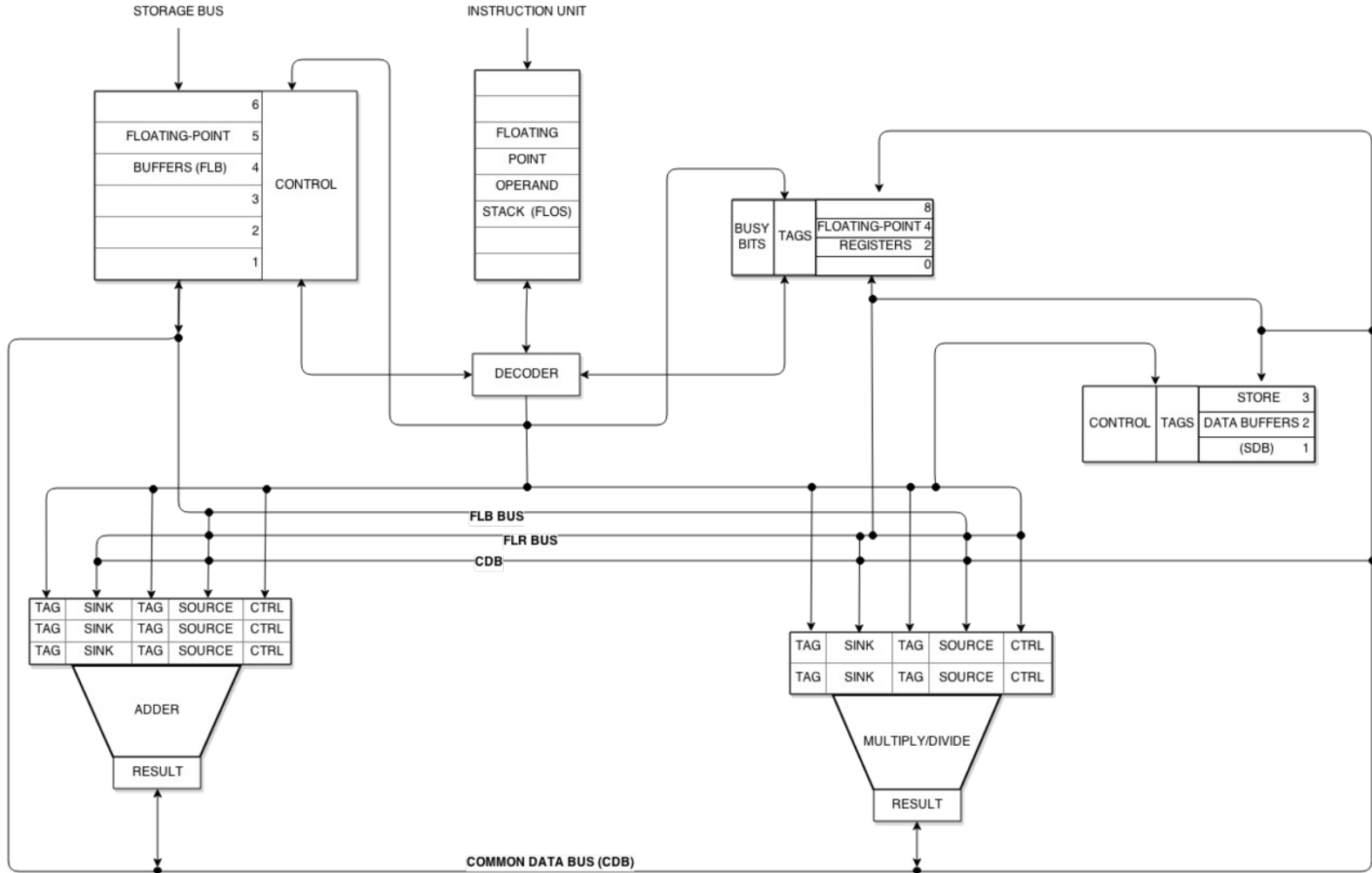
МКОД – конвейерные системы



МКОД – коллективные вычисления

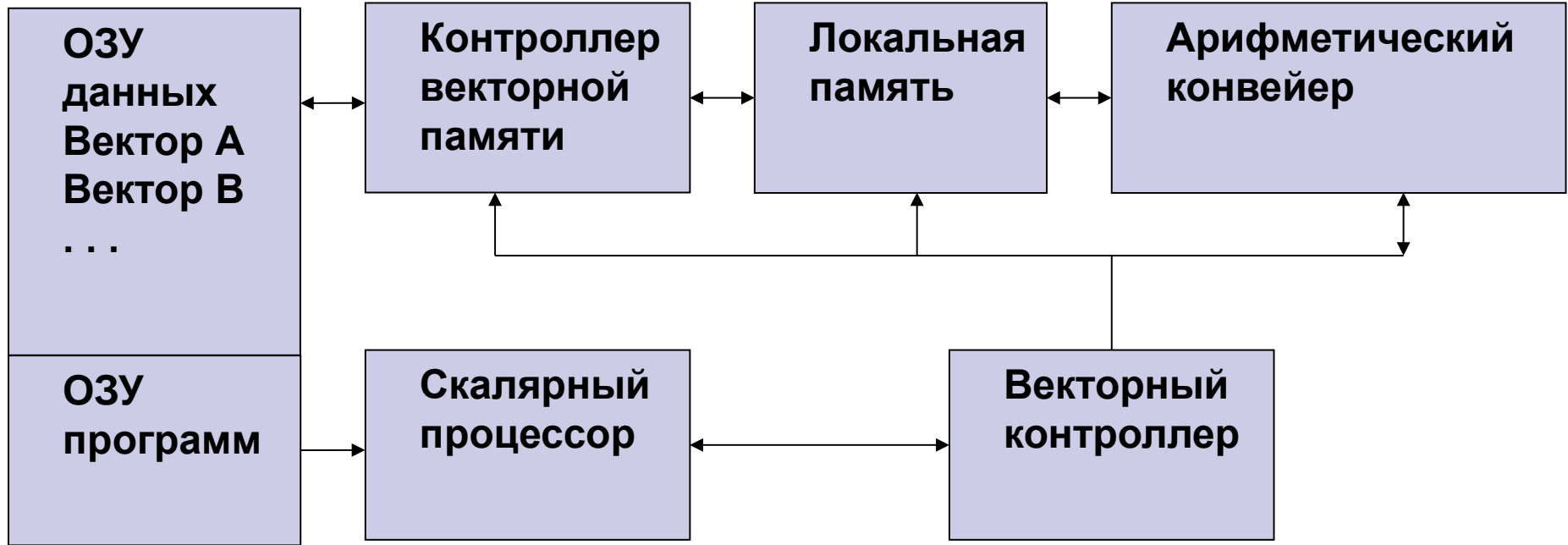


МКОД - конвейеризация команд



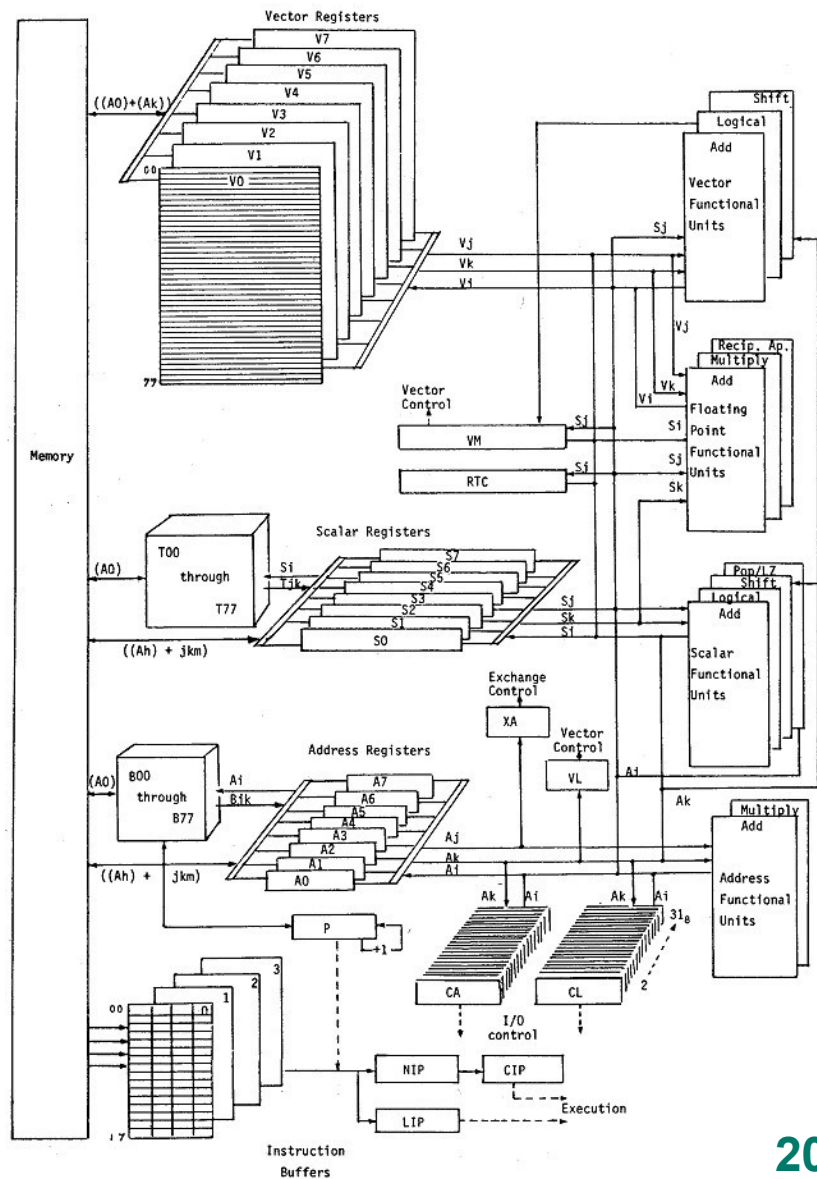
Конвейер команд в IBM 360/91

МКОД – векторно-конвейерные вычисления



$$\bar{A} * \bar{B} + \bar{C}$$

МКОД – Cray 1



МКОД – макроконвейерные ВС

Каждому отдельному процессору на очередном шаге вычислений дается такое задание, которое позволяет ему длительное время работать автономно без взаимодействия с другими процессорами.

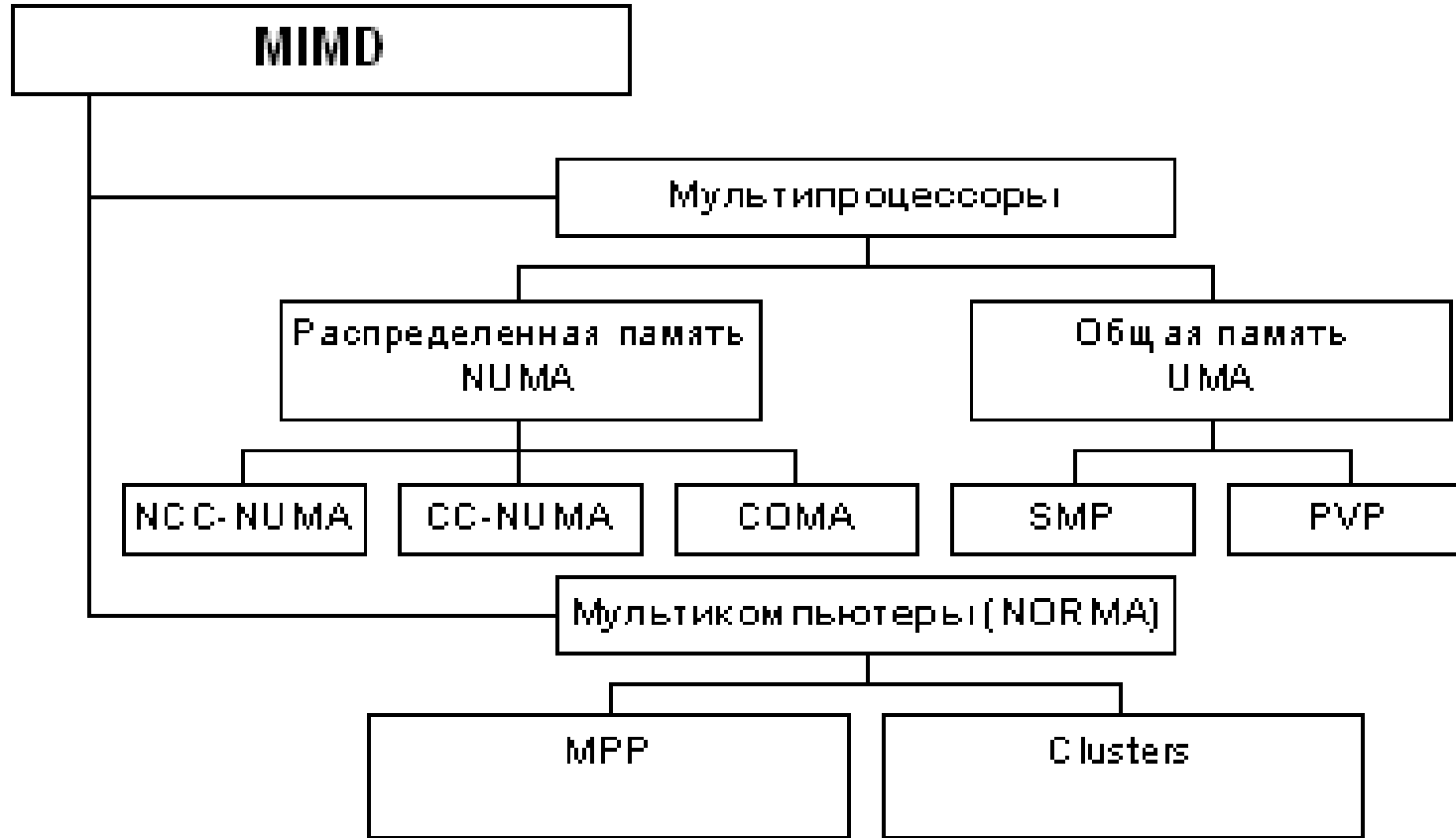
Глушков В.М.

Необходимо специальное представление распределенной программы, позволявшее соблюсти баланс между числом операций обмена и вычислительных операций для задействованных процессоров.

1984 г.: ЕС 2701

1987 г.: ЕС-1766 - серия.

Реальная скорость вычислений на 48 процессорах составляла полмиллиарда операций в секунду. В построенном макроконвейере удалось добиться почти линейного роста производительности при наращивании вычислительных ресурсов.



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



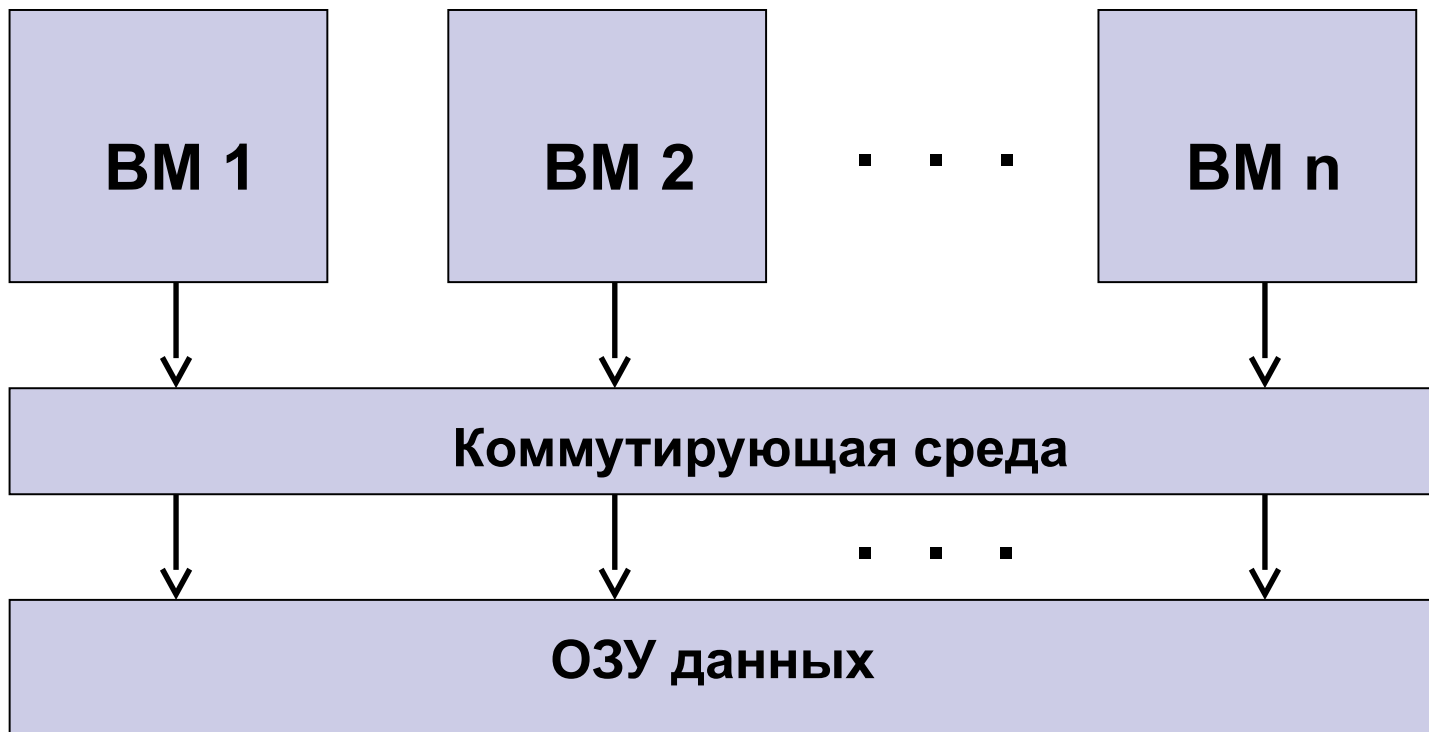
SMP – симметричные мультипроцессорные системы

Кластерные вычислительные системы

- Специализированные кластеры
- Кластеры общего назначения

MPP – массивно-параллельные системы

МКМД - SMP



Симметричные мультипроцессоры (SMP) - состоят из совокупности процессоров, обладающих одинаковыми возможностями доступа к памяти и внешним устройствам и функционирующих под управлением единой ОС.

МКМД - SMP

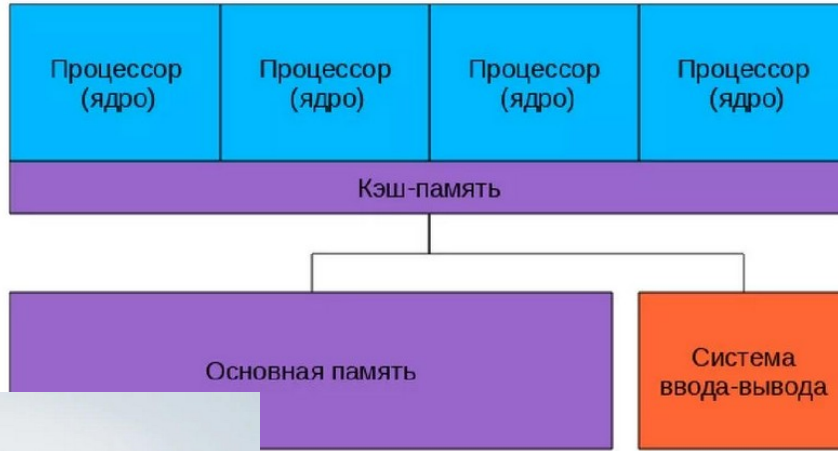
Операционная система. Система работает под управлением единой ОС (обычно UNIX-подобной, но для Intel-платформ поддерживается Windows). ОС автоматически распределяет процессы/нити по процессорам; но иногда возможна и явная привязка.

Модель программирования. С обменом данными через общую память (POSIX threads, OpenMP).

Масштабируемость. Наличие общей памяти упрощает взаимодействие процессоров между собой, однако накладывает сильные ограничения на их число - не более 32 в реальных системах.

МКМД - SMP

В настоящее время к этому классу относятся практически все многоядерные системы



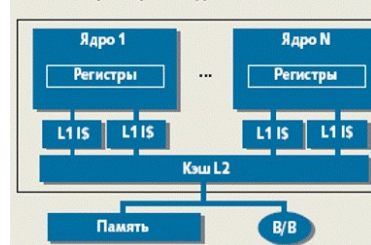
а) Традиционный процессор



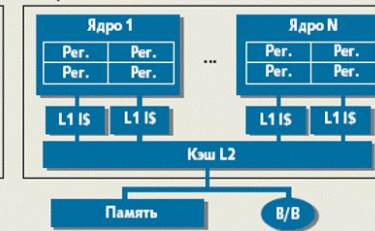
б) Простая многоядерная архитектура



в) Многоядерная архитектура с общей кэш-памятью



г) Многопоточковая многоядерная архитектура с общей кэш-памятью



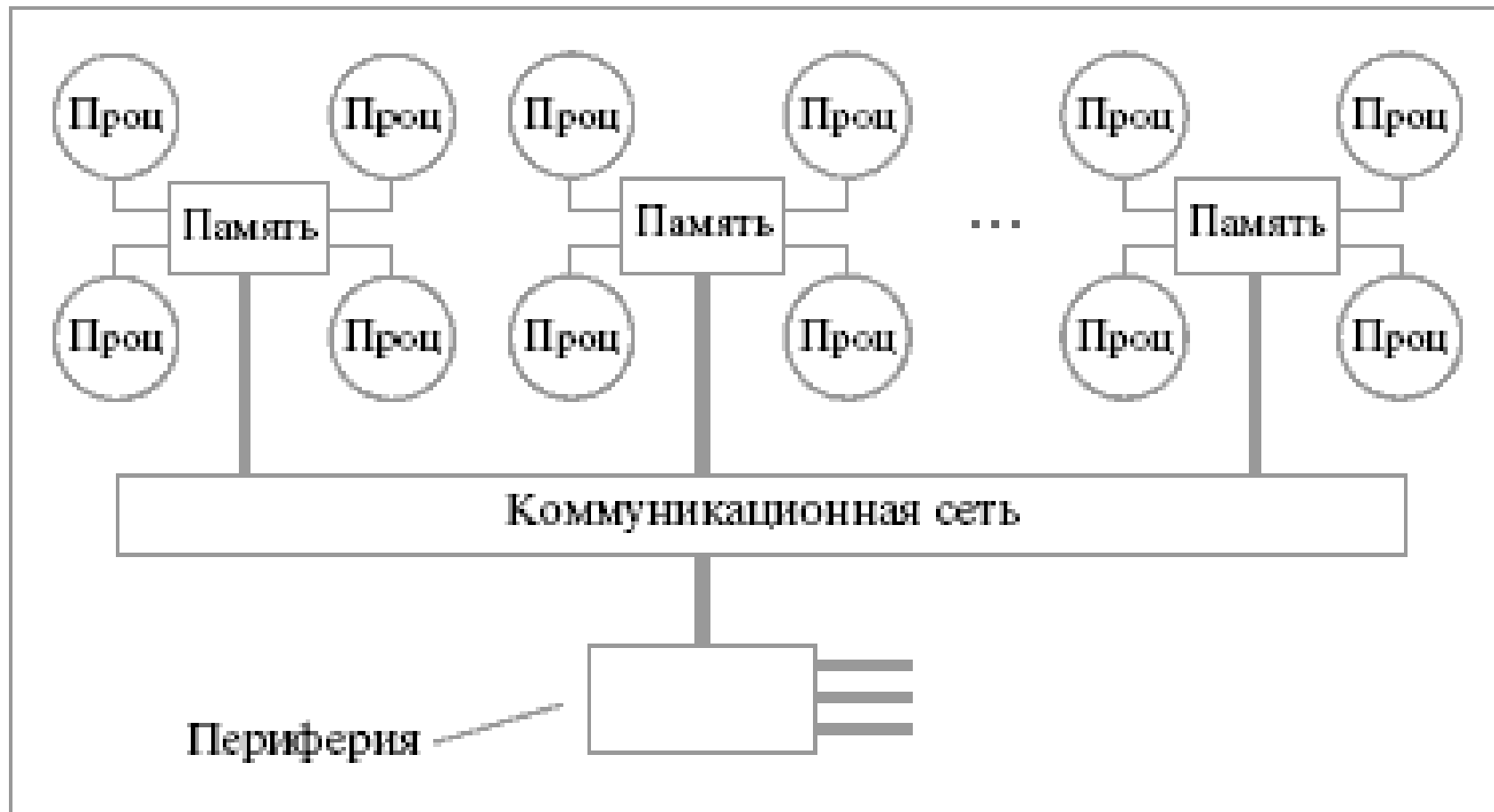
МКМД – Мультипроцессоры с распределенной памятью (NUMA – Non Uniform Memory Access)

Cache-Only Memory Architecture, СОМА - для представления данных используется только локальная кэш-память имеющихся процессоров.

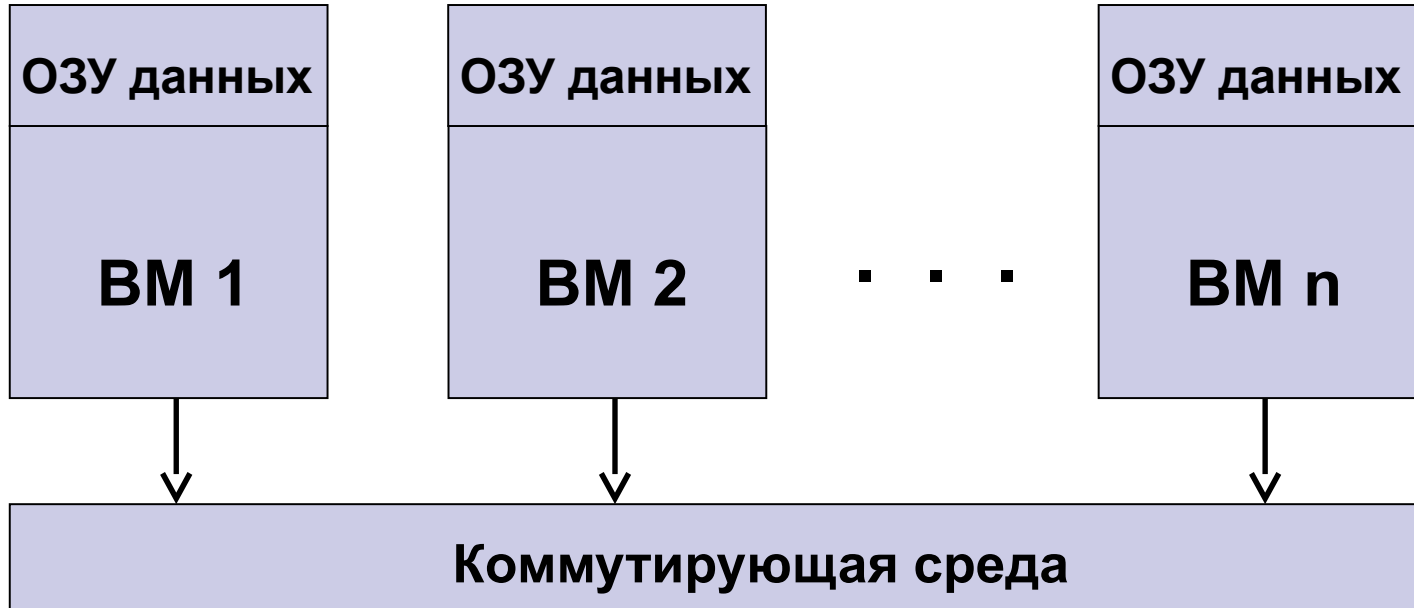
Cache-Coherent NUMA, СС-NUMA - обеспечивается однозначность локальных кэш-памятей разных процессоров.

Non-Cache Coherent NUMA, НСС-NUMA - обеспечивается общий доступ к локальной памяти разных процессоров без поддержки на аппаратном уровне когерентности кэша.

МКМД – Мультипроцессоры с распределенной памятью (NUMA – Non Uniform Memory Access)



МКМД – Кластеры и MPP



МКМД - Кластер

Кластерная система – параллельная вычислительная система, создаваемая из модулей высокой степени готовности, объединенных стандартной системой связи или разделяемыми устройствами внешней памяти.

МКМД - МРР

Массивно-параллельная система – высокопроизводительная параллельная вычислительная система, создаваемая с использованием специализированных вычислительных модулей и систем связи.

МКМД – кластеры

Архитектура. Набор элементов высокой степени готовности, рабочих станций или ПК общего назначения, объединяемых при помощи сетевых технологий и используемых в качестве массивно-параллельного компьютера.

Коммуникационная среда. Стандартные сетевые технологий (Fast/Gigabit Ethernet, Myrinet и др) на базе шинной архитектуры или коммутатора.

При объединении в кластер компьютеров разной мощности или разной архитектуры, говорят о **гетерогенных** (неоднородных) кластерах.

Узлы кластера могут одновременно использоваться в качестве пользовательских рабочих станций.

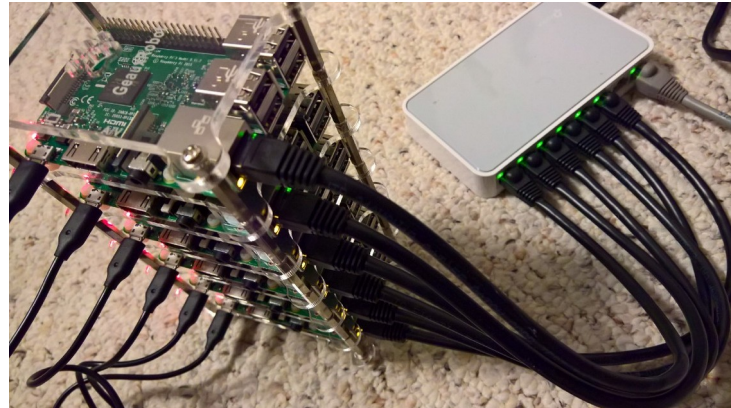
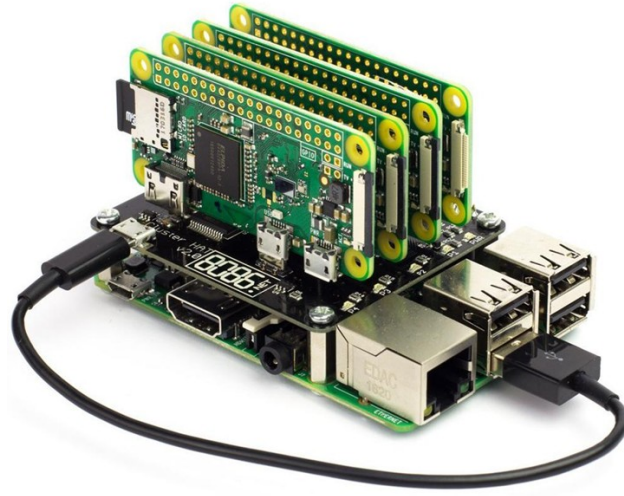
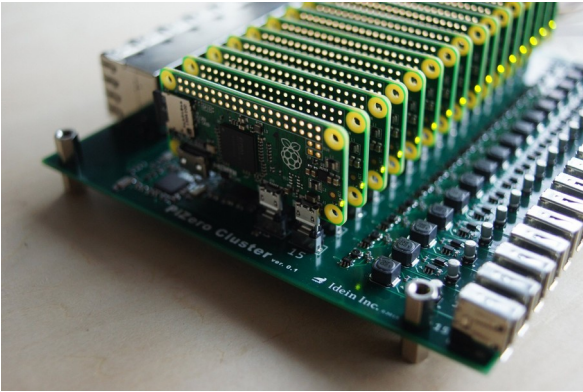
МКМД – кластеры

Операционная система - стандартные ОС - Linux/FreeBSD, вместе со средствами поддержки параллельного программирования и распределения нагрузки.

Модель программирования - с использованием передачи сообщений (PVM, MPI).

Основная проблема - большие накладные расходы на взаимодействие параллельных процессов между собой, что сильно сужает потенциальный класс решаемых задач.

МКМД – кластеры на Raspberry Pi



Архитектура. Система состоит из однородных вычислительных узлов, включающих:

- один или несколько центральных процессоров
- локальную память коммуникационный процессор или сетевой адаптер

Коммуникационная среда. Специально разработанная высокоскоростная.

Операционная система – два варианта:

1. Полноценная ОС работает только на управляющей машине (front-end), на каждом узле работает урезанный вариант ОС, обеспечивающие только работу расположенной в нем ветви параллельного приложения.
2. На каждом узле работает полноценная UNIX-подобная ОС аналогично кластерному подходу.

Модель программирования. Программирование в рамках модели передачи сообщений (MPI, PVM, BSPlib)

Масштабируемость. От нескольких тысяч до сотни тысяч и более узлов.

МКМД – Гибридные системы

Архитектура. Лежит в основе современных суперкомпьютеров. Состоит из неоднородных вычислительных узлов, включающих:

- множеств центральных процессоров, графические ускорители, спецпроцессоры на ПЛИС (программируемых логических интегральных схемах)
- разнообразные коммутационные средства и дополнительную память

Коммуникационная среда. Специально разработанная высокоскоростная.

Операционная система – два варианта:

1. Полноценная ОС работает только на управляющей машине (front-end), на каждом узле работает урезанный вариант ОС, обеспечивающие только работу расположенной в нем ветви параллельного приложения.
2. На каждом узле работает полноценная UNIX-подобная ОС аналогично кластерному подходу.

Модель программирования. Использование комплексных решений.

Масштабируемость. От нескольких тысяч до сотни тысяч и более узлов.

Суперкомпьютеры



Суперкомпьютеры

- Компьютер с производительностью свыше 100 000 млн. теоретических операций в сек.
- Компьютер стоимостью более 2 млн. долларов.
- Штучно или мелкосерийно выпускаемая вычислительная система, производительность которой многократно превосходит производительность массово выпускаемых компьютеров.
- Вычислительная система, сводящая проблему вычислений любого объема к проблеме ввода/вывода.

Суперкомпьютеры – **TOP 500**

29-я редакция Top500 от 27.06.2007

- 1** - прототип будущего суперкомпьютера IBM BlueGene/L с производительностью на Linpack 280.6 TFlop/s.
- 2** - Cray XT4/XT3, установленный в Oak Ridge National Laboratory, производительность на тесте Linpack составила 101.7 TFlop/s.
- 3** - Cray Red Storm с производительностью 101.4 TFlop/s на тесте Linpack.

Суперкомпьютеры – TOP 50. 58-я редакция Top50 от 16.11.2021

- **Rank** – порядковый номер в списке Top500
- **System** – название (тип) компьютера, указанное поставщиком, и организация, в которой установлен компьютер
- **Cores** – количество вычислительных ядер
- **Rmax (TFlop/s)** – максимальная полученная производительность по Linpack (TFlop/s)
- **Rpeak (TFlop/s)** – теоретическая пиковая производительность (TFlop/s)
- **Power (kW)** – электропотребление системы в кВт

Rank	System	Cores	Rmax (TFlop/s)	Rpeak (TFlop/s)	Power (kW)
1	Supercomputer Fugaku - Supercomputer Fugaku, A64FX 48C 2.2GHz, Tofu interconnect D, Fujitsu RIKEN Center for Computational Science Japan	7,630,848	442,010.0	537,212.0	29,899
2	Summit - IBM Power System AC922, IBM POWER9 22C 3.07GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband, IBM DOE/SC/Oak Ridge National Laboratory United States	2,414,592	148,600.0	200,794.9	10,096
3	Sierra - IBM Power System AC922, IBM POWER9 22C 3.1GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband, IBM / NVIDIA / Mellanox DOE/NNSA/LLNL United States	1,572,480	94,640.0	125,712.0	7,438
4	Sunway TaihuLight - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway, NRCPC National Supercomputing Center in Wuxi China	10,649,600	93,014.6	125,435.9	15,371
5	Perlmutter - HPE Cray EX235n, AMD EPYC 7763 64C 2.45GHz, NVIDIA A100 SXM4 40 GB, Slingshot-10, HPE DOE/SC/LBNL/NERSC United States	761,856	70,870.0	93,750.0	2,589
6	Selene - NVIDIA DGX A100, AMD EPYC 7742 64C 2.25GHz, NVIDIA A100, Mellanox HDR Infiniband, Nvidia NVIDIA Corporation United States	555,520	63,460.0	79,215.0	2,646
7	Tianhe-2A - TH-IVB-FEP Cluster, Intel Xeon E5-2692v2 12C 2.2GHz, TH Express-2, Matrix-2000, NUDT National Super Computer Center in Guangzhou China	4,981,760	61,444.5	100,678.7	18,482

Используемые источники

1. Таненбаум Э. Архитектура компьютера. 6-е изд. — СПб.: Изд. Питер, 2017. — 816 с.
2. Коуги П.М. Архитектура конвейерных ЭВМ. — М.: Радио и связь, 1985. — 360 с.
3. Все о мире суперкомпьютеров и параллельных вычислений:
<https://parallel.ru/>
4. Юрий Затуливетер, Елена Фищенко. Многопроцессорный компьютер ПС-2000. - Открытые системы. СУБД. № 09, 2007. -
<https://www.osp.ru/os/2007/09/4570286>
5. Как мы создавали кластер из Raspberry Pi: <https://habr.com/ru/post/170305/>

Суперкомпьютеры – TOP 500

На ноябрь 2021 года по числу систем в рейтинге страны распределяются:

Китай — 173 суперкомпьютера (при этом из Топ-10 по мощности 2 суперкомпьютера установлены в Китае).

США имеет в Топ-500 149 суперкомпьютеров (при этом из Топ-10 по мощности 5 суперкомпьютеров установлено в США).

Япония — 32 (при этом из Топ-10 только один суперкомпьютер, и одновременно самый мощный в мире суперкомпьютер Фугаку установлен в Японии),

Германия — 26, **Франция** — 19, **Великобритания** — 11,

Нидерланды — 11, **Канада** — 11, **Южная Корея** — 7, **Россия** — 7.

Суперкомпьютеры – TOP 50

Россия: 35-я редакция Top50 от 27.09.2021

№	Название Место установки	Узлов Проц. Ускор.	Архитектура: кол-во узлов: конфигурация узла сеть: вычислительная / сервисная / транспортная	Rmax Rpeak (Тфлоп/с)	Разработчик Область применения
1	«Кристофари» SberCloud (ООО «Облачные технологии»), СберБанк, Москва	75 150 1200	75: NVIDIA DGX-2 CPU: 2x Intel Xeon Platinum 8168 24C 2.7GHz, 1536 GB RAM Acc: 16x NVIDIA Tesla V100 EDR Infiniband / 100 Gigabit Ethernet / 10 Gigabit Ethernet	6669.0 8789.76	SberCloud (ООО «Облачные технологии») NVIDIA Облачный провайдер
2	«Ломоносов-2» Московский государственный университет имени М.В.Ломоносова, Москва	1696 1696 1856	1536: CPU: 1x Intel Xeon E5-2697v3, 64 GB RAM Acc: 1x NVIDIA Tesla K40M 160: CPU: 1x Intel Xeon Gold 6126, 96 GB RAM Acc: 2x NVIDIA Tesla P100 FDR Infiniband / Gigabit Ethernet / FDR Infiniband	2478.0 4946.79	Т-Платформы Наука и образование
3 new	«МТС GROM» ПАО "МТС", Лыткарино	20 40 160	20: CPU: 2x AMD EPYC 7742, 1024 GB RAM Acc: 8x NVIDIA NVIDIA A100 40GB InfiniBand / нд / нд	2258.0 3011.84	NVIDIA Mellanox NetApp Искусственный интеллект
4 ▾	ФГБУ 'ГВЦ Росгидромета', Москва	976 1952 н/д	976: CPU: 2x Intel Xeon E5-2697v4, 128 GB RAM Aries / Aries + Gigabit Ethernet / Aries + Infiniband	1200.35 1293.0	Т-Платформы Cray Исследования



- О Сервере
- Что нового?
- Компьютеры
- Технологии
- Исследования
- Метакомпьютинг
- Информация
- Конференции
- Центры
- История
- Россия
- Новости
- Общение
- Наш сервис
- ЦКП

Twitter [Читать @ParallelRu](#)

Главные Новости

В NASA Ames Research Center установлен суперкомпьютер Aitken HPE Apollo с пиковой производительностью 10.77 PFlop/s.

Atos и CEA выпустили суперкомпьютер EXA1 BullSequana XH2000 с пиковой производительностью 23.2 PFlop/s.

В CINES (National Computing Center for Higher Education) будет установлен суперкомпьютер Adastra HPE Cray EX с пиковой производительностью более 70 PFlop/s.

Некоторая статистика 58-ой редакции списка Top500 мощнейших суперкомпьютеров мира.

Размещена копия 58-ой редакции списка Top500 наиболее мощных суперкомпьютеров мира.



Открыт приём статей в специальный номер журнала "Supercomputing Frontiers and Innovations" по теме "Supercomputing in Weather, Climate and Environmental Prediction" (#4, 2021).

25 сентября – 2 октября 2021 г. в Москве состоится международный конгресс "Суперкомпьютерные дни в России". Конгресс объединяет научную конференцию, научные школы суперкомпьютерной академии, серию специализированных научных семинаров, экскурсии в ведущие суперкомпьютерные центры и множество других событий, проводимых на различных площадках Москвы и России.

Открыта вакансия специалиста для работы в НИВЦ МГУ имени М.В.Ломоносова в рамках научных проектов в области суперкомпьютерных технологий.

Опубликована версия 5.2 стандарта OpenMP.

В United Weather Centres – West будет установлен суперкомпьютер HPE Cray.

NVIDIA анонсирует новое поколение архитектуры InfiniBand под названием Quantum-2.

Появились сообщения, о том, что в Китае уже созданы два экзафлопсных суперкомпьютера, Oceanlite производства Sunway и Tianhe-3, пиковая производительность которых составляет примерно по 1.3 EFlop/s.

В Texas Advanced Computing Center (TACC) установлен суперкомпьютер Lonestar6 производства Dell Technologies на базе процессоров AMD EPYC с пиковой производительностью около 3 PFlop/s.

[Все новости](#)