

Введение в методы параллельного программирования

Содержание

- Понятие параллельных вычислений
 - Необходимость параллельных вычислений
 - Сдерживающие факторы
 - Характеристика необходимых знаний и умений
 - Структура учебного плана
 - Содержание курса
 - Литература
-

Понятие параллельных вычислений...

Под *параллельными вычислениями (parallel or concurrent computations)* можно понимать процессы решения задач, в которых в один и тот же момент времени могут выполняться одновременно несколько вычислительных операций

Понятие параллельных вычислений

- Параллельные вычисления обычно сводятся к использованию многопроцессорных вычислительных систем
 - Одновременные выполняемые операции должны быть направлены на решение общей задачи
 - Параллельные вычисления следует отличать от многозадачных (многопрограммных) режимов работы последовательных ЭВМ
-

Необходимость параллельных вычислений...

- Опережение потребности вычислений быстрогодействия существующих компьютерных систем
(ex., Problems of Grand Challenge – проблемы большого вызова)
 - моделирование климата,
 - генная инженерия,
 - проектирование интегральных схем,
 - анализ загрязнения окружающей среды,
 - создание лекарственных препаратов и др.

Оценка min необходимой производительности –
 10^{12} операций (1 Tflops)

Необходимость параллельных вычислений

- Теоретическая ограниченность роста производительности последовательных компьютеров
 - Резкое снижение стоимости многопроцессорных (*параллельных*) вычислительных систем
 - 1 Cray T90 processor – 1.8 GFlops (\$2 500 000),
 - 8 Node IBM SP2 using R6000 - 2.1 GFlops (\$500 000)
 - Смена парадигмы построения высокопроизводительных процессоров - *многоядерность*
-

Сдерживающие факторы...

- **высокая стоимость параллельных систем –**
в соответствии с *законом Гроша (Grosch)*,
производительность компьютера
возрастает пропорционально квадрату его
стоимости
?!
-

Сдерживающие факторы...

- **потери производительности для организации параллелизма** – согласно гипотезе Минского (*Minsky*), ускорение, достигаемое при использовании параллельной системы, пропорционально двоичному логарифму от числа процессоров
?!
-

Сдерживающие факторы...

- **постоянное совершенствование последовательных компьютеров** – в соответствии с *законом Мура (Moore)* мощность последовательных процессоров возрастает практически в два раза каждые 18 месяцев
?!!!
-

Сдерживающие факторы...

- **существование последовательных вычислений** –
в соответствии с *законом Амдаля (Amdahl)*
ускорение процесса вычислений при использовании p
процессоров ограничивается величиной

$$S \leq 1/(f+(1-f)/p) \leq 1/f,$$

где f есть доля последовательных вычислений в
применяемом алгоритме обработки данных

?!

Сдерживающие факторы...

- зависимость эффективности параллелизма от учета характерных свойств параллельных систем (**отсутствие мобильности для параллельных программ**)
?!
-

Сдерживающие факторы

- существующее программное обеспечение ориентировано в основном на последовательные ЭВМ
?!
-

Структура учебного плана

- Лекции – 18 часа
 - Лабораторный практикум –
 - ☞ 6 лабораторных работ (7)
 - Экзамен !
-

Литература...

- **Гергель В.П.** Теория и практика параллельных вычислений. - М.: Интернет-Университет, БИНОМ. Лаборатория знаний, 2007.
 - **Богачев К.Ю.** Основы параллельного программирования. - М.: БИНОМ. Лаборатория знаний, 2003.
 - **Воеводин В.В., Воеводин Вл.В.** Параллельные вычисления. - СПб.: БХВ-Петербург, 2002.
 - **Немнюгин С., Стесик О.** Параллельное программирование для многопроцессорных вычислительных систем — СПб.: БХВ-Петербург, 2002.
 - **А.С.Антонов.** Параллельное программирование с использованием технологии MPI.
 - **К.Ю. Богачев.** Основы параллельного программирования.
-

Принципы построения параллельных вычислительных систем

- Пути достижения параллелизма
 - Примеры параллельных вычислительных систем
 - Суперкомпьютеры
 - Кластеры
 - Классификация многопроцессорных вычислительных систем
 - Мультипроцессоры
 - Мультикомпьютеры
 - Типовые схемы коммуникации процессоров
 - Системные платформы для построения кластеров
 - Заключение
-

Пути достижения параллелизма...

Под *параллельными вычислениями* понимаются процессы обработки данных, в которых одновременно могут выполняться несколько операций компьютерной системы

Пути достижения параллелизма...

- Достижение параллелизма возможно только при выполнении следующих требований:
 - **независимость функционирования отдельных устройств ЭВМ** (устройства ввода-вывода, обрабатывающие процессоры и устройства памяти),
 - **избыточность элементов вычислительной системы**
 - *использование специализированных устройств* (например, отдельные процессоры для целочисленной и вещественной арифметики, устройства многоуровневой памяти),
 - *дублирование устройств ЭВМ* (например, использование нескольких одноплатных обрабатывающих процессоров или нескольких устройств оперативной памяти),
 - Дополнительная форма обеспечения параллелизма - **конвейерная** реализация обрабатывающих устройств
-

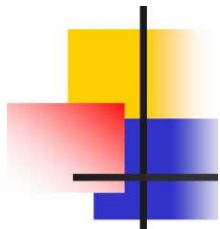
Пути достижения параллелизма...

- Возможные режимы выполнения независимых частей программы:
 - *многозадачный режим (режим разделения времени)*, при котором для выполнения нескольких процессов используется единственный процессор (данный режим является псевдопараллельным, в каждый момент времени исполняемым может быть единственный процесс),
 - *параллельное выполнение*, когда в один и тот же момент времени может выполняться несколько команд обработки данных (обеспечивается при наличии нескольких процессоров),
 - *распределенные вычисления*, при которых для параллельной обработки данных используется несколько обрабатывающих устройств, достаточно удаленных друг от друга, а передача данных по линиям связи приводит к существенным временным задержкам.
-

Пути достижения параллелизма

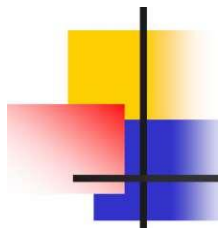
Основное внимание будем уделять второму типу организации параллелизма, реализуемому на многопроцессорных вычислительных системах

Классификации вычислительных систем



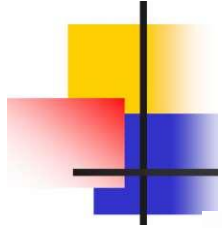
- По назначению
 - По модели программирования
 - Потоки данных/команд
(классификация Флинна)
-

Классификации по назначению



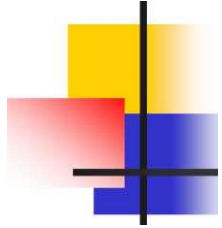
- Персональный компьютер (рабочая станция)
 - Тонкий клиент (X-терминал)
 - Сервер
 - Мейнфрейм
 - Суперкомпьютер
 - Блэйд-системы
-

Персональный компьютер (ПЭВМ)



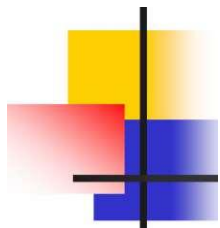
- Малый размер
- Серийное производство
- Доступность (~\$1000)
- Универсальность
- Низкая производительность
- Единственный пользователь

Тонкий клиент



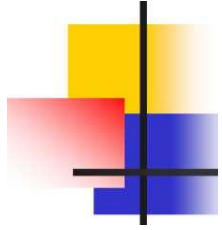
- Минимальный размер
- Дешевизна (~\$150)
- Отсутствие ПЗУ
- Необходимость сервера/мейнфрейма

Сервер



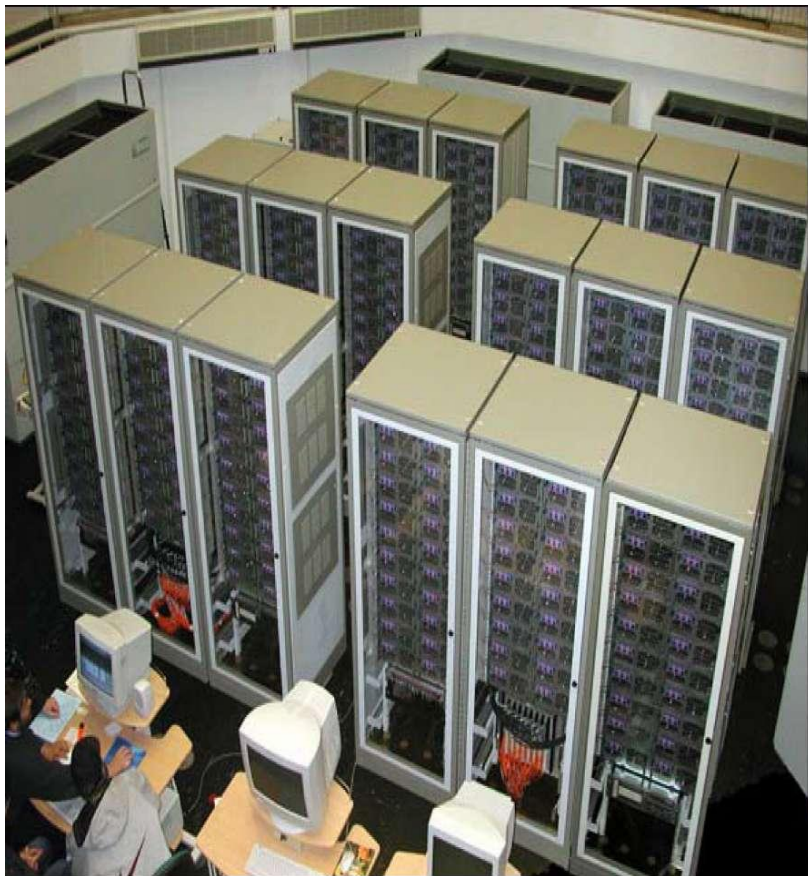
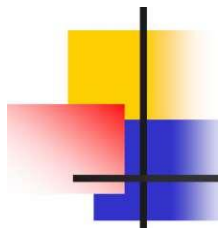
- Схожая с ПК платформа
 - (аналогичные
 - комплектующие)
 - Специализированные задачи (БД, файлы, приложения и т.д.)
 - Увеличенная производительность
 - Отсутствие пользователя
 - Серийное производство
-

Мейнфрейм



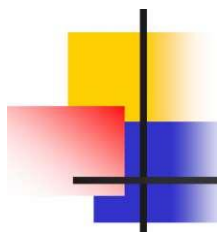
- Значительная производительность
 - Значительный объем ресурсов
 - Мелкосерийное производство
 - Повышенная устойчивость
 - Высокая рабочая нагрузка (80-95% пиковой)
-

Суперкомпьютер

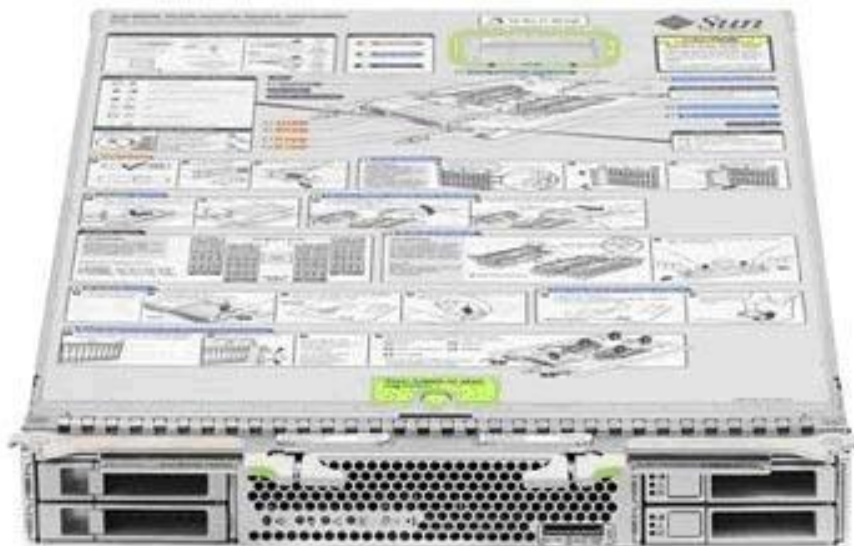


- Уникальная платформа
- Штучное производство (под заказ)
- Большой размер
- Огромная производительность
- Большая стоимость

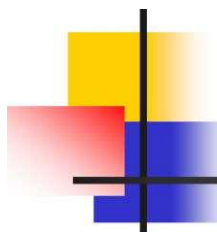
Блэйд-системы



Новый тип серверов XXI века — модульные, чаще называемые Blade-серверами, или серверами-лезвиями (blade — лезвие). Преимущества Blade-серверов, первые модели которых были разработаны в 2001 г. изготовители описывают с помощью правила "1234". "По сравнению с обычными серверами при сравнимой производительности Blade-серверы занимают в два раза меньше места, потребляют в три раза меньше энергии и обходятся в четыре раза дешевле".

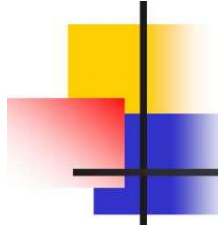


Блэйд-системы



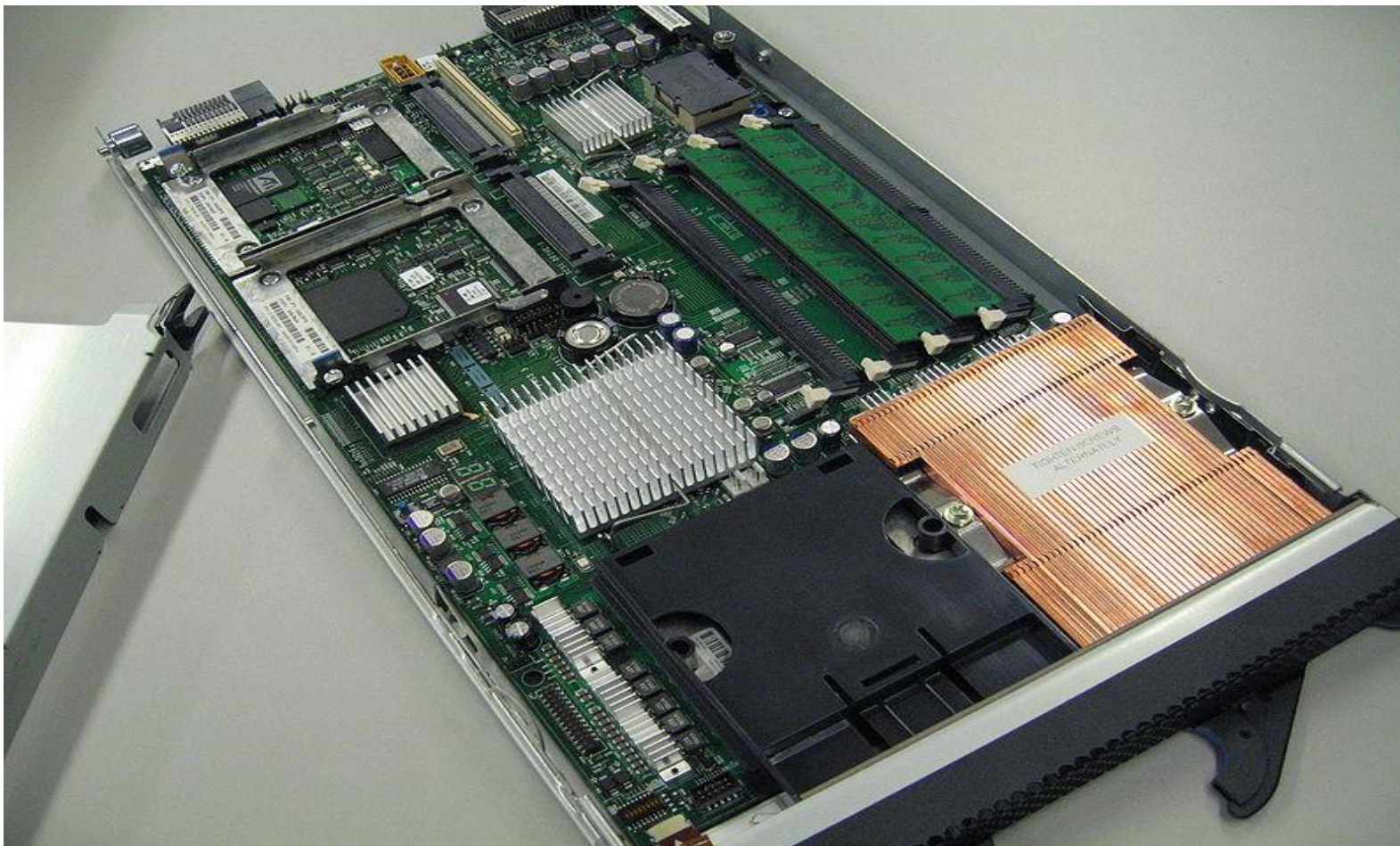
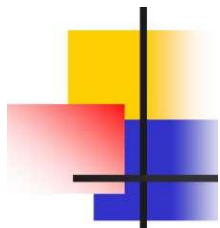
Blade-сервер или лезвие - это модульная одноплатная компьютерная система, включающая процессор и память. Лезвия вставляются в специальное шасси с объединительной панелью (backplane), обеспечивающей им подключение к сети и подачу электропитания. **Это шасси с лезвиями, является Blade-системой.** Оно выполнено в конструктиве для установки в стандартную 19-дюймовую стойку и в зависимости от модели и производителя, занимает в ней 3U, 6U или 10U (один U - unit, или монтажная единица, равен 1,75 дюйма). За счет общего использования таких компонентов, как источники питания, сетевые карты и жесткие диски, Blade-серверы обеспечивают более высокую плотность размещения вычислительной мощности в стойке по сравнению с обычными тонкими серверами высотой 1U и 2U.

Блэйд-системы



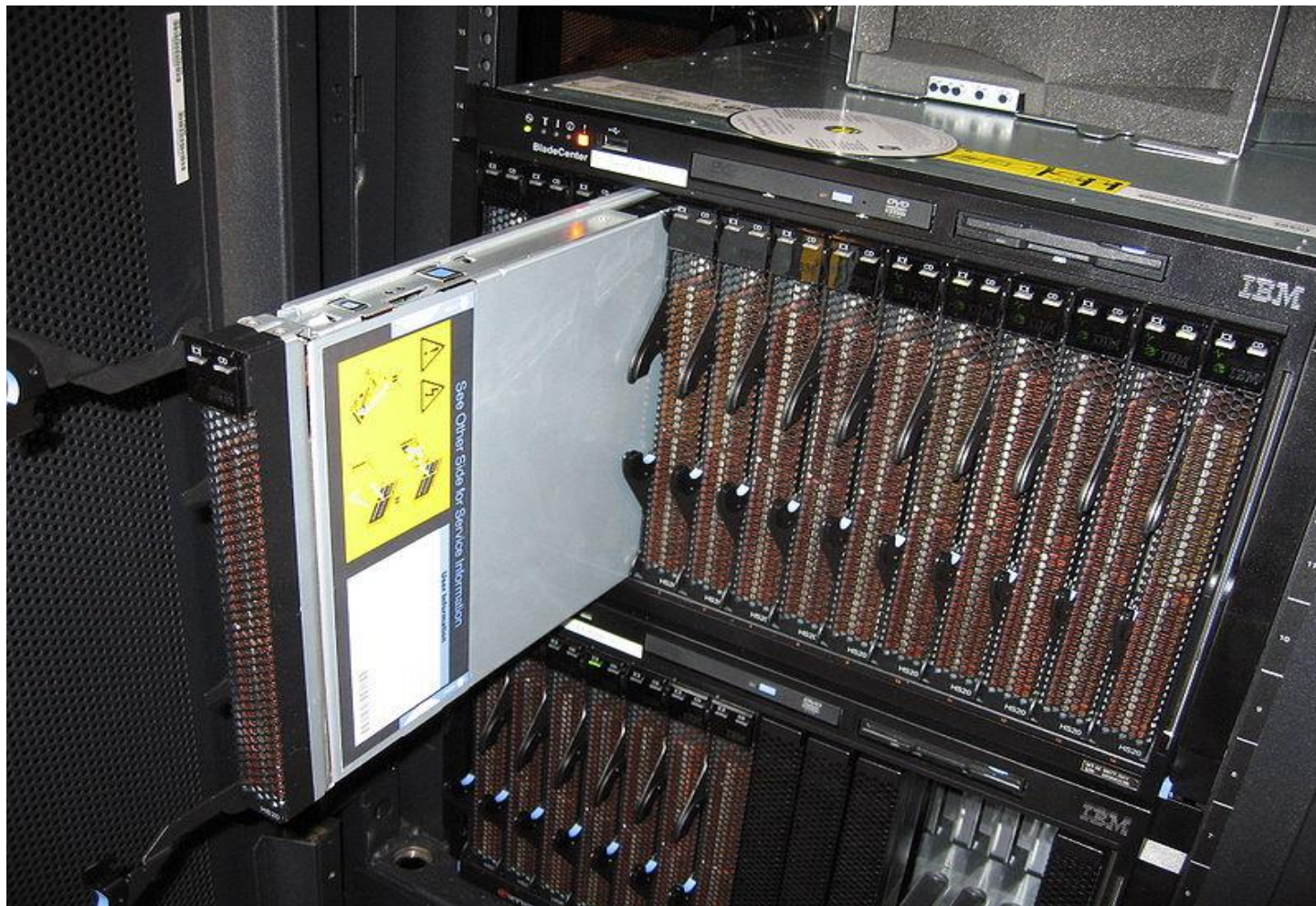
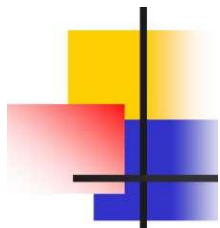
Типичное 10U шасси для 10 Blade-серверов

Блэйд-системы



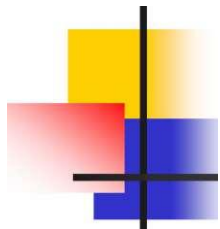
Блейд-сервер IBM HS20

Блэйд-системы



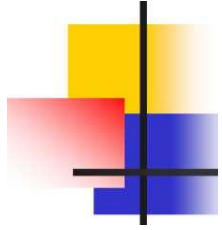
Корзина с блейд-серверами

Классификации по «модели программирования»



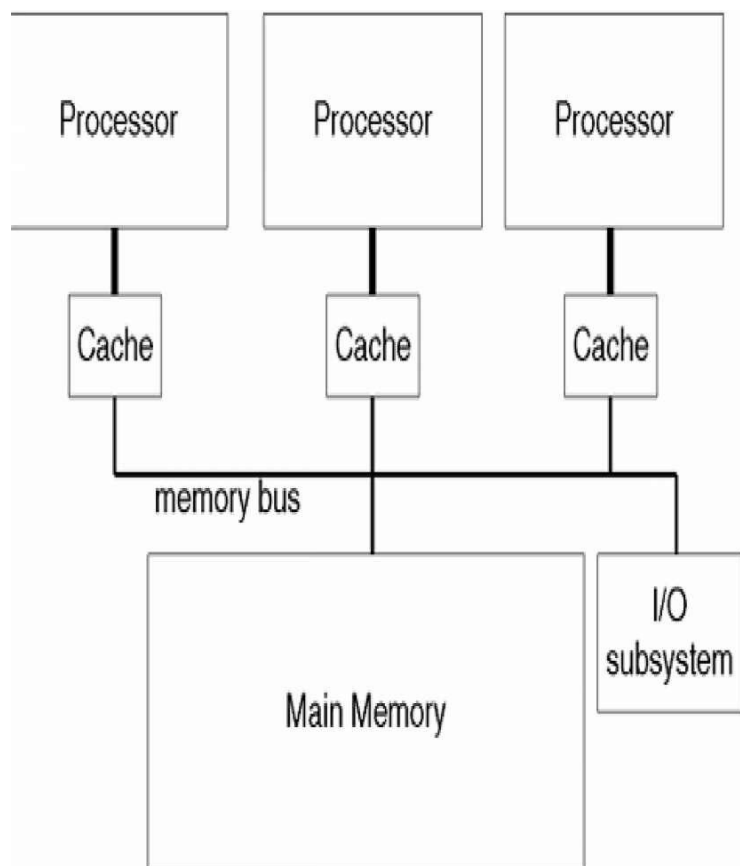
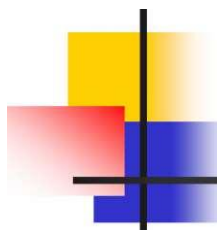
- Последовательная система (тривиальный случай)
 - Симметричные параллельные системы (SMP) - общей памятью
 - Массивно-параллельные системы - с распределенной памятью
 - Кластерные системы
-

Ядро, поток, процесс



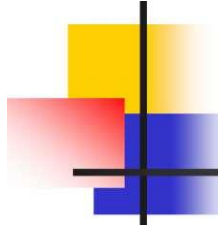
- Вычислительное ядро - совокупность программно-аппаратных средств, способная выполнять поток команд
 - Процесс - совокупность ресурсов системы (адресное пространство, дисковое пространство, вычислительные ядра), выделенная для решения некоторой задачи
 - Поток - последовательность инструкций, выполняемая вычислительным ядром в рамках процесса
-

Симметричное мультипроцессирование



- Несколько вычислительных ядер
- Единая физическая память (единая шина данных)
- Одна копия ОС
- Параллелизм на основе распределения ресурсов средствами ОС
- Ограниченная масштабируемость

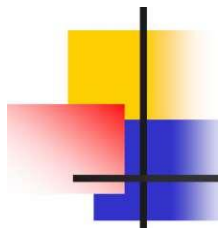
Модели программирования SMP



- Программирование в терминах ОС (API)
- Использование библиотечных классов и функций (VCL, MFC, Boost)
- Средства языков программирования (Java, Ada)
- Расширения языков программирования (OpenMP)

Любая модель в конечном итоге основывается на средствах ОС

Кластерные системы



- Множество вычислительных ядер
- Каждое вычислительное ядро управляется собственной копией ОС
- Вычислительные ядра соединены аутентичным коммуникационным интерфейсом сложной топологии
- Затраты на взаимодействие между ядрами

Программирование в терминах передачи сообщений средствами встроенной ОС

Примеры параллельных вычислительных систем...

- **Суперкомпьютеры**

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

Примеры параллельных вычислительных систем...

- **Суперкомпьютеры. Программа ASCI**

(Accelerated Strategic Computing Initiative)

- **1996**, система **ASCI Red**, построенная Intel, производительность 1 TFlops,
 - **1999**, **ASCI Blue Pacific** от IBM и **ASCI Blue Mountain** от SGI, производительность 3 TFlops,
 - **2000**, **ASCI White** с пиковой производительностью свыше 12 TFlops (реально показанная производительность на тесте LINPACK составила на тот момент 4938 GFlops)
-

Примеры параллельных вычислительных систем...

- **Суперкомпьютеры. ASCI White...**
 - система с 512-ю симметричными мультипроцессорными (SMP) узлами, каждый узел имеет 16 процессоров,
 - процессоры IBM RS/6000 POWER3 с 64-х разрядной архитектурой и конвейерной организацией с 2 устройствами по обработке команд с плавающей запятой и 3 устройствами по обработке целочисленных команд, они способны выполнять до 8 команд за тактовый цикл и до 4 операций с плавающей запятой за такт, тактовая частота 375 MHz,
 - оперативная память системы – 4 TB,
 - емкость дискового пространства 180 TB
-

Примеры параллельных вычислительных систем...

- **Суперкомпьютеры. ASCI White**
 - Операционная система представляет собой версию UNIX – IBM AIX,
 - Программное обеспечение ASCI White поддерживает смешанную модель программирования – передача сообщений между узлами и многопоточность внутри SMP-узла,
 - Поддерживаются библиотеки MPI, OpenMP, потоки POSIX и транслятор директив IBM, имеется параллельный отладчик IBM.
-

- **Суперкомпьютеры. Система BlueGene**
 - Первый вариант системы представлен в 2004 г. и сразу занял 1 позицию в списке Top500
 - Расширенный вариант суперкомпьютера (до ноября 2007 г.) на 1 месте в перечне наиболее быстродействующих вычислительных систем:
 - 212992 двухядерных 32-битных процессоров PowerPC 440 0.7 GHz,
 - пиковая производительность около 600 TFlops, производительность на тесте LINPACK – 478 TFlops
-

Примеры параллельных вычислительных систем...

- **Суперкомпьютеры. МВС-1000...**

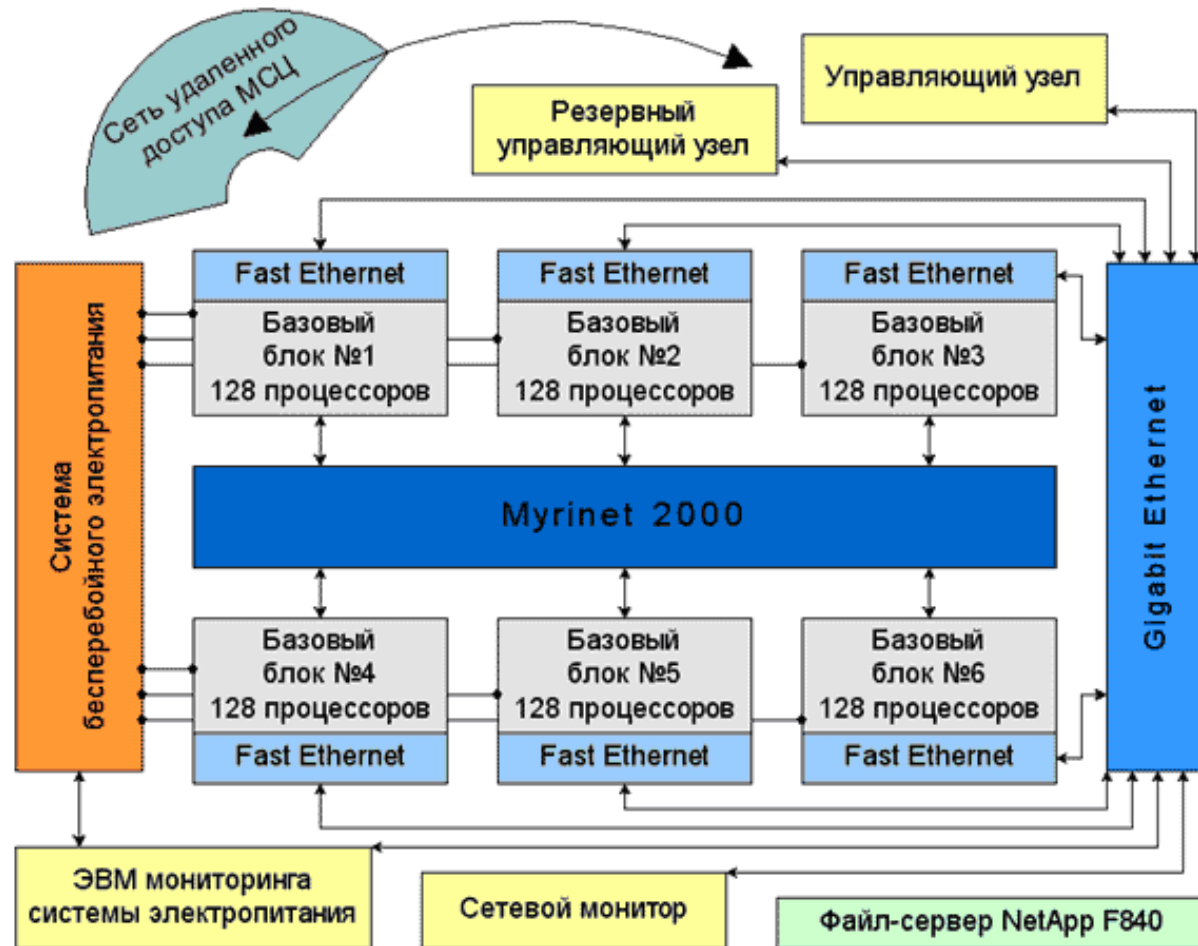
(Межведомственный Суперкомпьютерный Центр РАН)

2001

- Пиковая производительность 1024 GFlops, максимально показанная на тесте LINPACK производительность 734 GFlops,
 - 384 двухпроцессорных модуля на базе Alpha 21264 667 MHz (кэш L2 4 Mb), собранные в виде 6 базовых блоков, по 64 модуля в каждом,
 - Каждый вычислительный модуль имеет по 2 Gb оперативной памяти, HDD 20 Gb, сетевые карты Myrinet (2000 Mbit) и Fast Ethernet (100 Mbit),
 - Операционные системы управляющего сервера и вычислительных модулей – ОС Linux RedHat 6.2 с поддержкой SMP.
-

Примеры параллельных вычислительных систем...

- Суперкомпьютеры. МВС-1000



Примеры параллельных вычислительных систем...

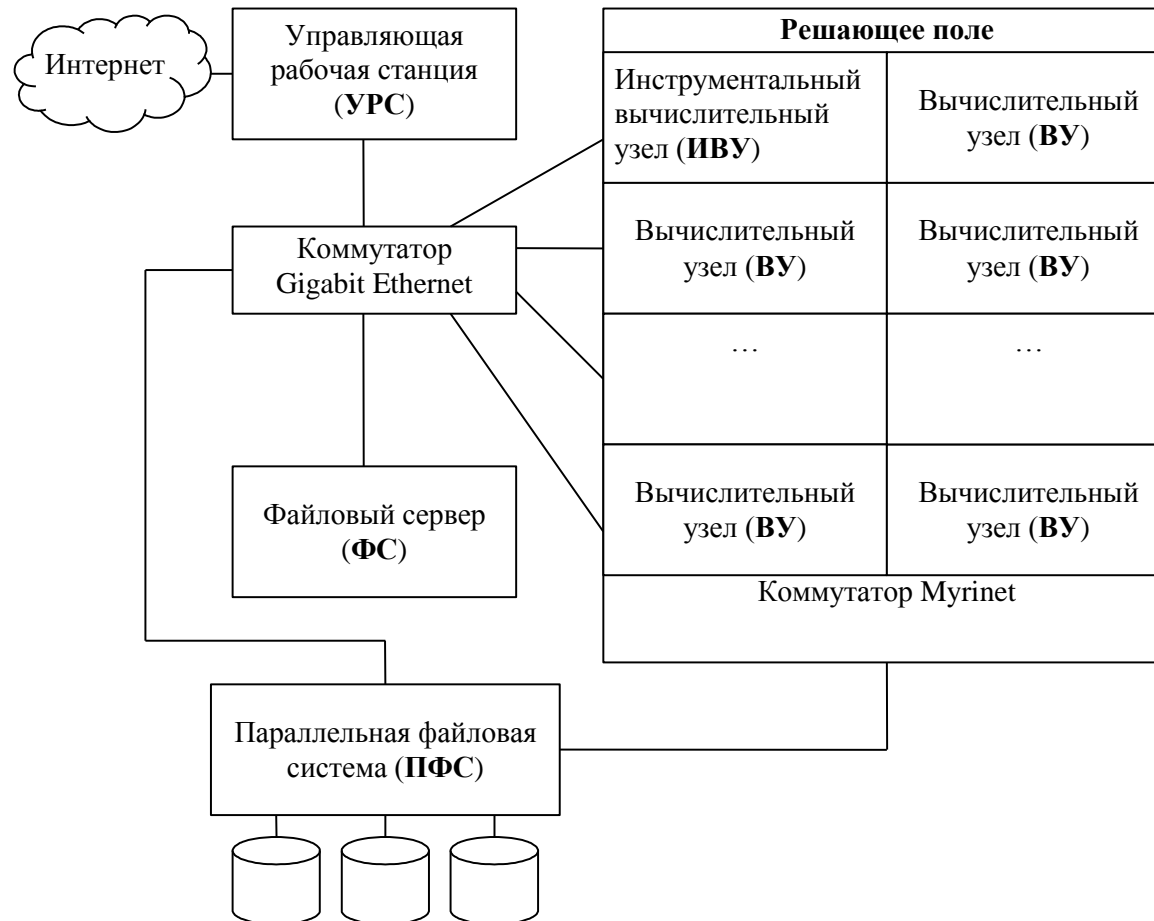
- **Суперкомпьютеры. МВС-15000...**

(Межведомственный Суперкомпьютерный Центр РАН)
2005

- Общее количество узлов 276 (552 процессора).
Каждый узел представляет собой:
 - 2 процессора IBM PowerPC 970 с тактовой частотой 2.2 GHz, кэш L1 96 Kb и кэш L2 512 Kb,
 - 4 Gb оперативной памяти на узел,
 - 40 Gb жесткий диск IDE,
 - Операционная система SuSe Linux Enterprise Server версии 8 для платформ x86 и PowerPC,
 - Пиковая производительность 4857.6 GFlops и максимально показанная на тесте LINPACK 3052 GFlops.
-

Примеры параллельных вычислительных систем...

- **Суперкомпьютеры. МВС-15000**



Примеры параллельных вычислительных систем...

- **Кластеры**

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 параллельных вычислительных систем...

- **Кластеры. Beowulf...**

- В настоящее время под кластером типа “*Beowulf*” понимается вычислительная система, состоящая из одного серверного узла и одного или более клиентских узлов, соединенных при помощи сети Ethernet или некоторой другой сети передачи данных. Это система, построенная из готовых серийно выпускающихся промышленных компонент, на которых может работать ОС Linux/Windows, стандартных адаптеров Ethernet и коммутаторов.

Примеры параллельных вычислительных систем...

- **Кластеры. Beowulf...**

- **1994**, научно-космический центр NASA Goddard Space Flight Center, руководители проекта - Томас Стерлинг и Дон Бекер:

- 16 компьютеров на базе процессоров 486DX4, тактовая частота 100 MHz,
 - 16 Mb оперативной памяти на каждом узле,
 - Три параллельно работающих 10Mbit/s сетевых адаптера,
 - Операционная система Linux, компилятор GNU, поддержка параллельных программ на основе MPI.
-

Примеры параллельных вычислительных систем...

- **Кластеры. Beowulf**

- **1998**, Система **Avalon**, Лос-Аламосская национальная лаборатория (США) ,
руководители проекта - астрофизик Майкл Уоррен:

- 68 процессоров (позднее расширен до 140) Alpha 21164A с тактовой частотой 533 MHz,
 - 256 Mb RAM, 3 Gb HDD, Fast Ethernet card на каждом узле,
 - Операционная система Linux,
 - Пиковая производительность в 149 GFlops, производительность на тесте LINPACK 48.6 GFlops.
-

Примеры параллельных вычислительных систем...

- **Кластеры. AC3 Velocity Cluster**

- **2000**, [Корнельский университет](#) (США), результат совместной работы университета и Advanced Cluster Computing Consortium, образованного компаниями Dell, Intel, Microsoft, Giganet:
 - 64 четырехпроцессорных сервера Dell PowerEdge 6350 на базе Intel Pentium III Xeon 500 MHz, 4 GB RAM, 54 GB HDD, 100 Mbit Ethernet card,
 - 1 восьмипроцессорный сервер Dell PowerEdge 6350 на базе Intel Pentium III Xeon 550 MHz, 8 GB RAM, 36 GB HDD, 100 Mbit Ethernet card,
 - Операционная система Microsoft Windows NT 4.0 Server Enterprise Edition,
 - Пиковая производительность AC3 Velocity 122 GFlops, производительность на тесте LINPACK 47 GFlops.
-

Примеры параллельных вычислительных систем...

- Кластеры. NCSA [NT Supercluster](#)
 - 2000, Национальный центр суперкомпьютерных технологий (National Center for Supercomputing Applications):
 - 38 двухпроцессорных систем [Hewlett-Packard Kayak XU PC workstation](#) на базе Intel Pentium III Xeon 550 MHz, 1 Gb RAM, 7.5 Gb HDD, 100 Mbit Ethernet card,
 - Операционная система ОС Microsoft Windows,
 - Пиковая производительностью в 140 GFlops и производительность на тесте LINPACK 62 GFlops.
-

Примеры параллельных вычислительных систем...

- Кластеры. Thunder

- 2004, Ливерморская Национальная Лаборатория (США):

- 1024 сервера, в каждом по 4 процессора Intel Itanium 1.4 GHz,
 - 8 Gb оперативной памяти на сервер,
 - общая емкость дисковой системы 150 Tb,
 - операционная система CHAOS 2.0,
 - пиковая производительность 22938 GFlops и максимально показанная на тесте LINPACK 19940 GFlops (5-ая позиция списка Top500).
-

Классификация вычислительных систем...

- **Систематика Флинна (Flynn)**

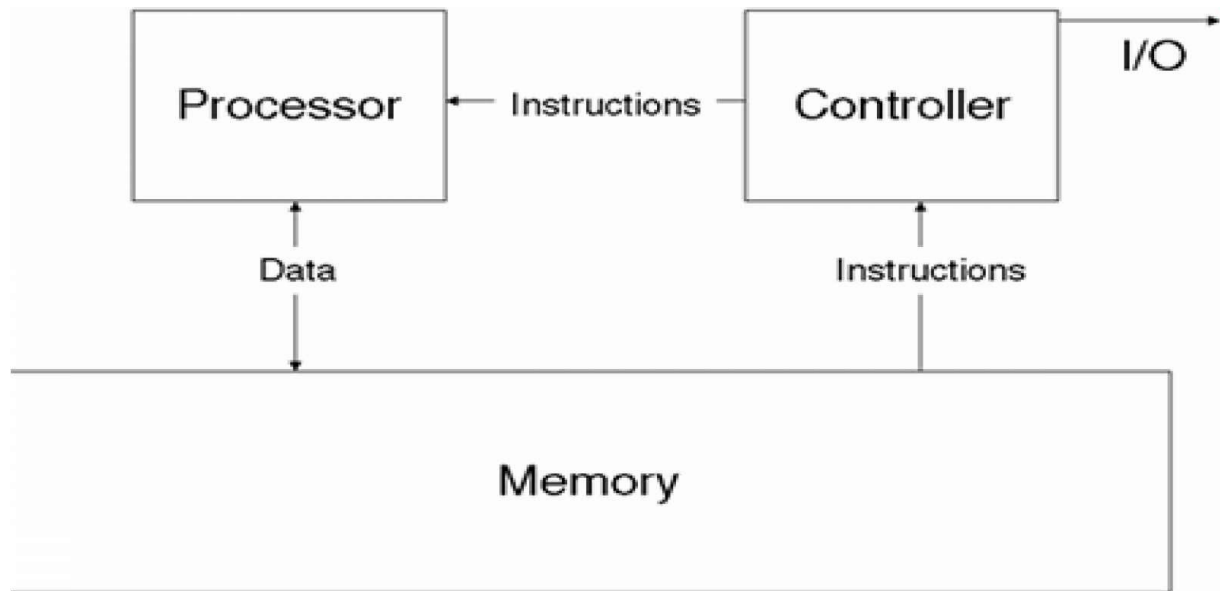
- классификация по способам взаимодействия последовательностей (*потоков*) выполняемых команд и обрабатываемых данных:

- **SISD** (Single Instruction, Single Data)
 - **SIMD** (Single Instruction, Multiple Data)
 - **MISD** (Multiple Instruction, Single Data)
 - **MIMD** (Multiple Instruction, Multiple Data)

Практически все виды параллельных систем, несмотря на их существенную разнородность, относятся к одной группе

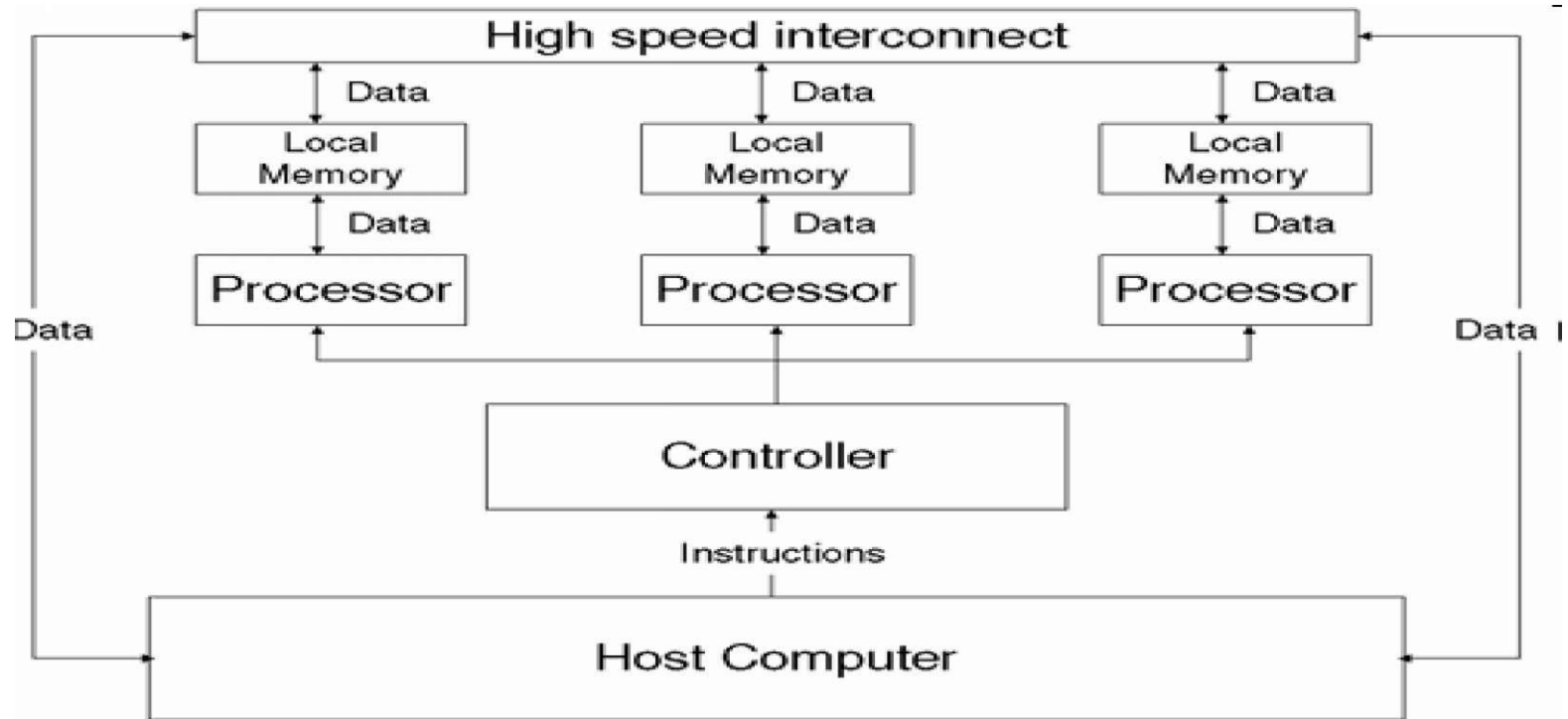
MIMD

Систематика Флинна SISD



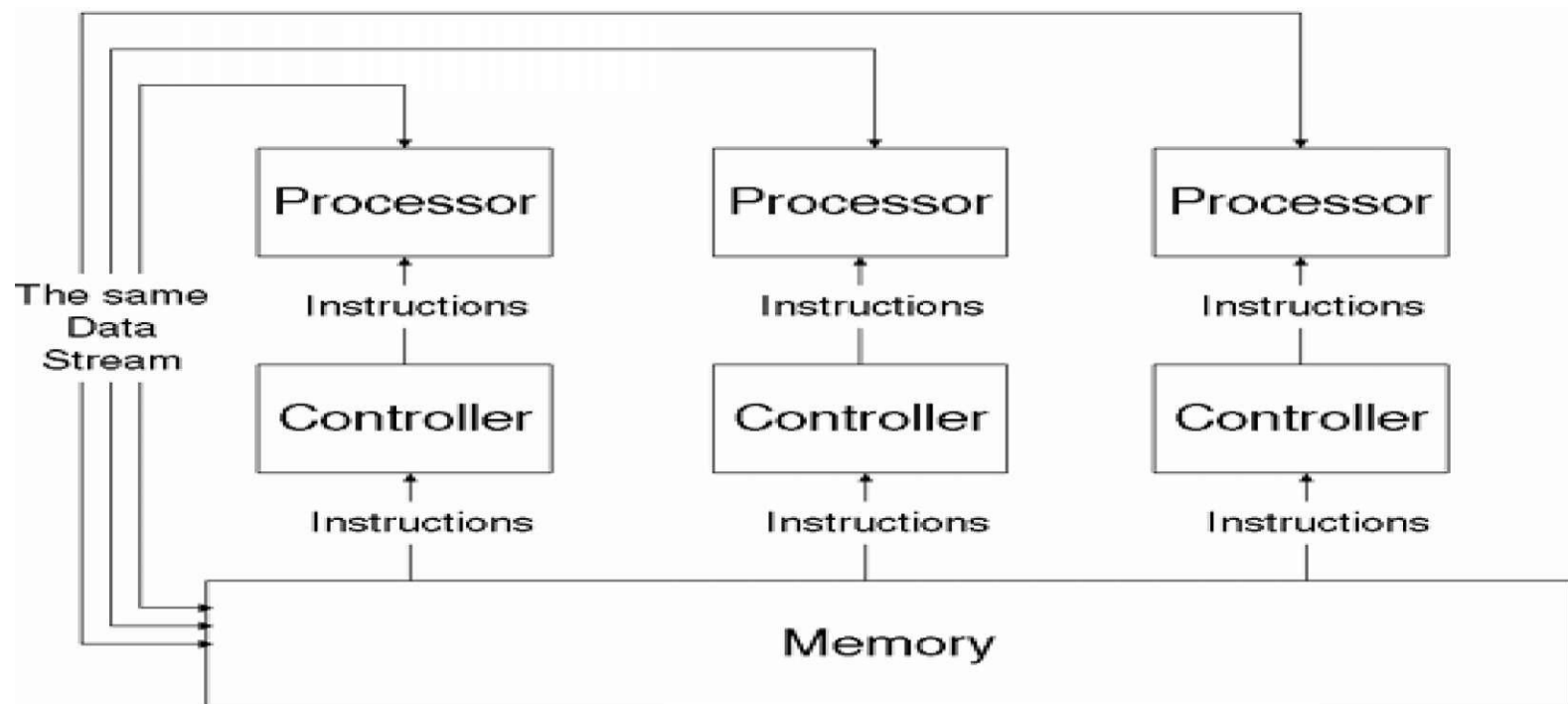
Single instruction stream / single data stream - одиночный поток команд и одиночный поток данных.

Систематика Флинна SIMD



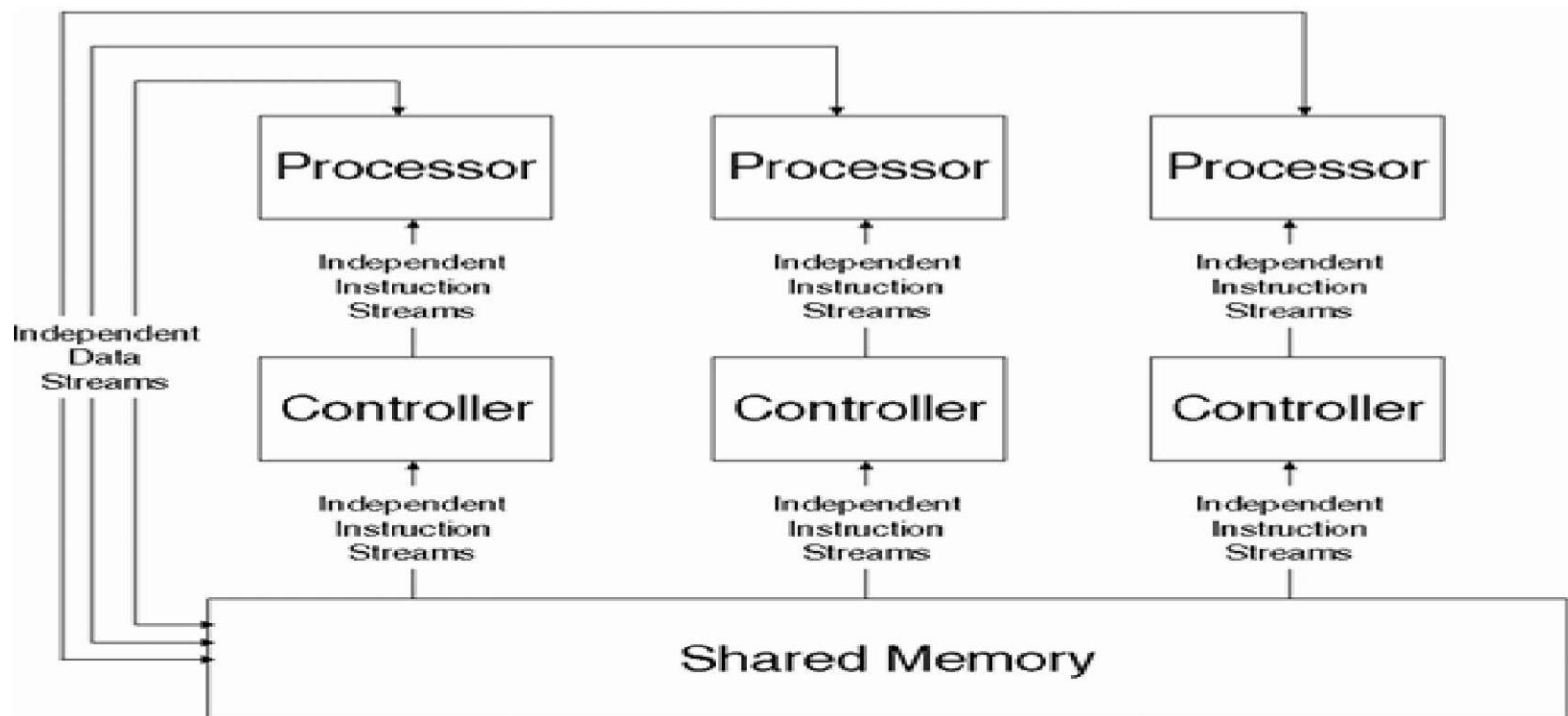
Single instruction stream / multiple data stream - одиночный поток команд и множественный поток данных

Систематика Флинна MISD



Multiple instruction stream / single data stream - множественный поток команд и одиночный поток данных
Машин этого класса реально не существует

Систематика Флинна MIMD



Multiple instruction stream / multiple data stream - множественный поток команд и множественный поток данных

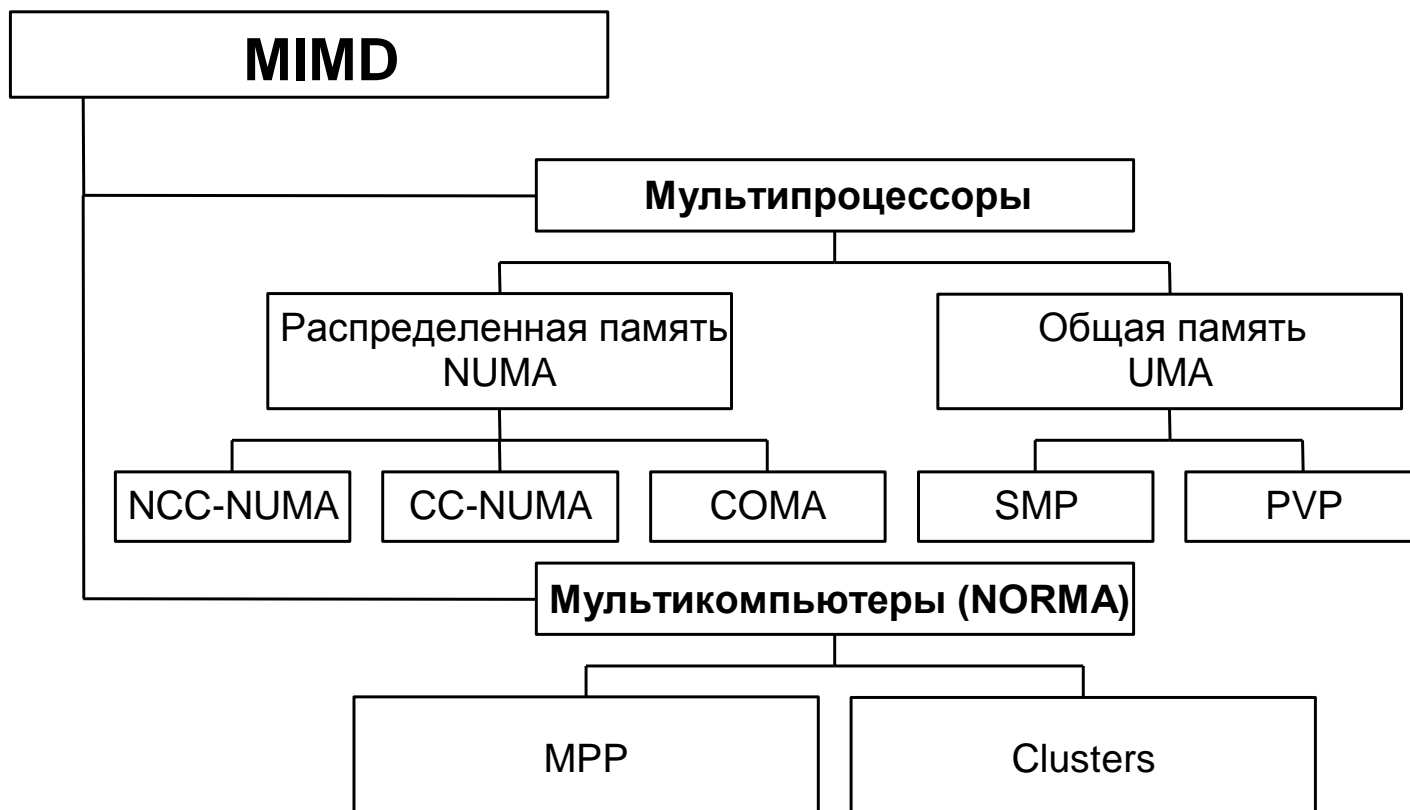
Классификация вычислительных систем...

- **Детализация систематики Флинна...**

- дальнейшее разделение типов многопроцессорных систем основывается на используемых способах организации оперативной памяти,
 - позволяет различать два важных типа многопроцессорных систем:
 - ***multiprocessors*** (**мультипроцессоры** или системы с общей разделяемой памятью),
 - ***multicomputers*** (**мультикомпьютеры** или системы с распределенной памятью).
-

Классификация вычислительных систем...

• Детализация систематики Флинна...



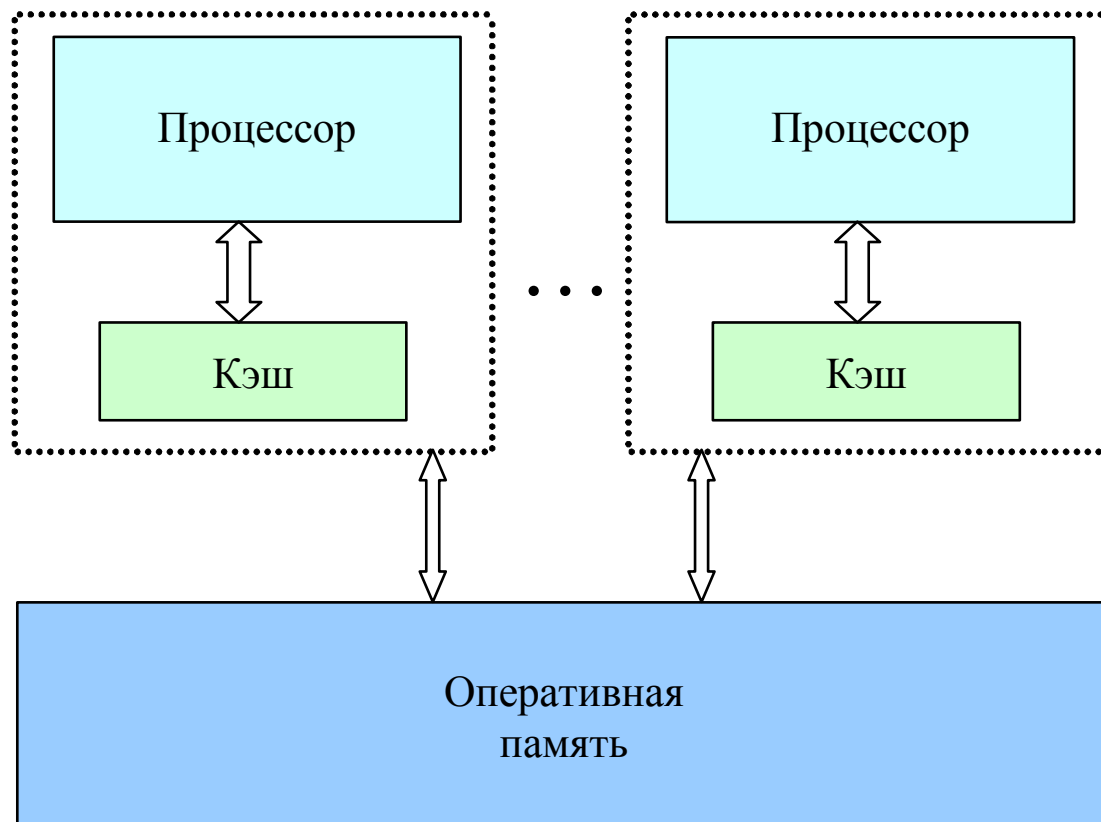
UMA – uniform memory access
NUMA – non-uniform memory access
SMP – symmetric multiprocessor
PVP – parallel vector processor
COMA – cache-only memory architecture
CC – cache coherent
NCC – non-cache coherent
NORMA – no-remote memory access
MMP – massively parallel processor

Классификация вычислительных систем...

- **Мультипроцессоры** с использованием единой *общей памяти (shared memory)*...
 - обеспечивается *однородный доступ к памяти (uniform memory access or UMA)*,
 - являются основой для построения:
 - *векторных параллельных процессоров (parallel vector processor or PVP)*. Примеры: Cray T90,
 - *симметричных мультипроцессоров (symmetric multiprocessor or SMP)*. Примеры: IBM eServer, Sun StarFire, HP Superdome, SGI Origin.
-

Классификация вычислительных систем...

- **Мультипроцессоры** с использованием единой *общей памяти...*



Классификация вычислительных систем...

- **Мультипроцессоры** с использованием единой *общей памяти*

Проблемы:

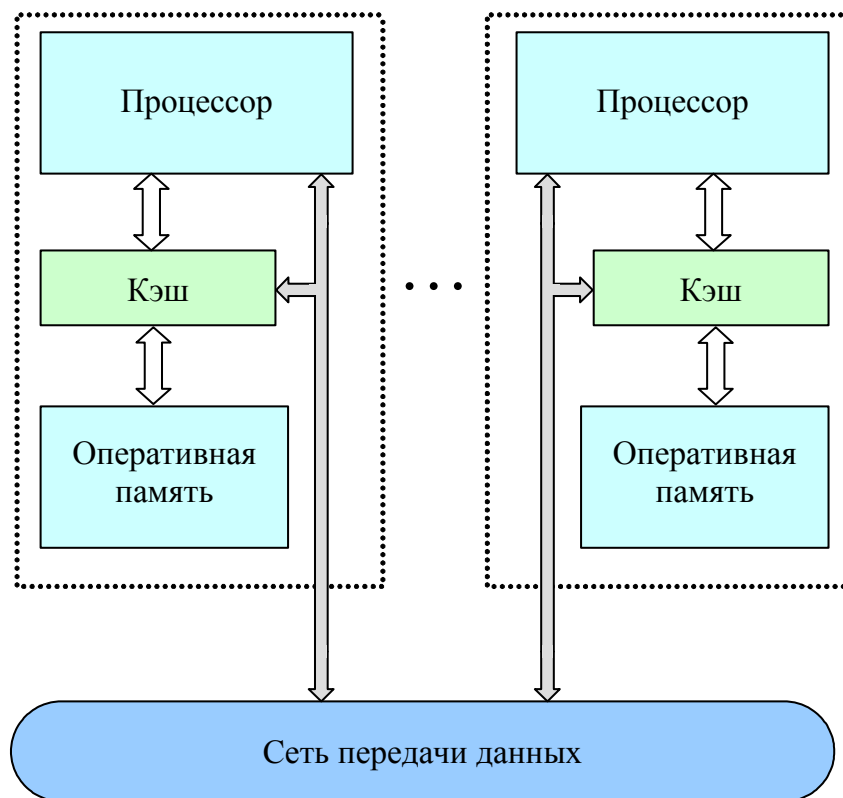
- Доступ с разных процессоров к общим данным и обеспечение, в этой связи, *однозначности (когерентности) содержимого разных кэшей (cache coherence problem)*,
 - Необходимость *синхронизации взаимодействия* одновременно выполняемых потоков команд
-

Классификация вычислительных систем...

- **Мультипроцессоры** с использованием физически распределенной памяти (*distributed shared memory or DSM*):
 - *неоднородный доступ к памяти (non-uniform memory access or NUMA)*,
 - Среди систем такого типа выделяют:
 - *Cache-only memory architecture or COMA* (системы KSR-1 и DDM),
 - *cache-coherent NUMA or CC-NUMA* (системы SGI Origin 2000, Sun HPC 10000, IBM/Sequent NUMA-Q 2000),
 - *non-cache coherent NUMA or NCC-NUMA* (система Cray T3E).
-

Классификация вычислительных систем...

- **Мультипроцессоры** с использованием физически распределенной памяти...



Классификация вычислительных систем...

- **Мультипроцессоры** с использованием физически распределенной памяти:
 - упрощаются проблемы создания мультипроцессоров (известны примеры систем с несколькими тысячами процессоров),
 - возникают проблемы эффективного использования распределенной памяти (время доступа к локальной и удаленной памяти может различаться на несколько порядков).
-

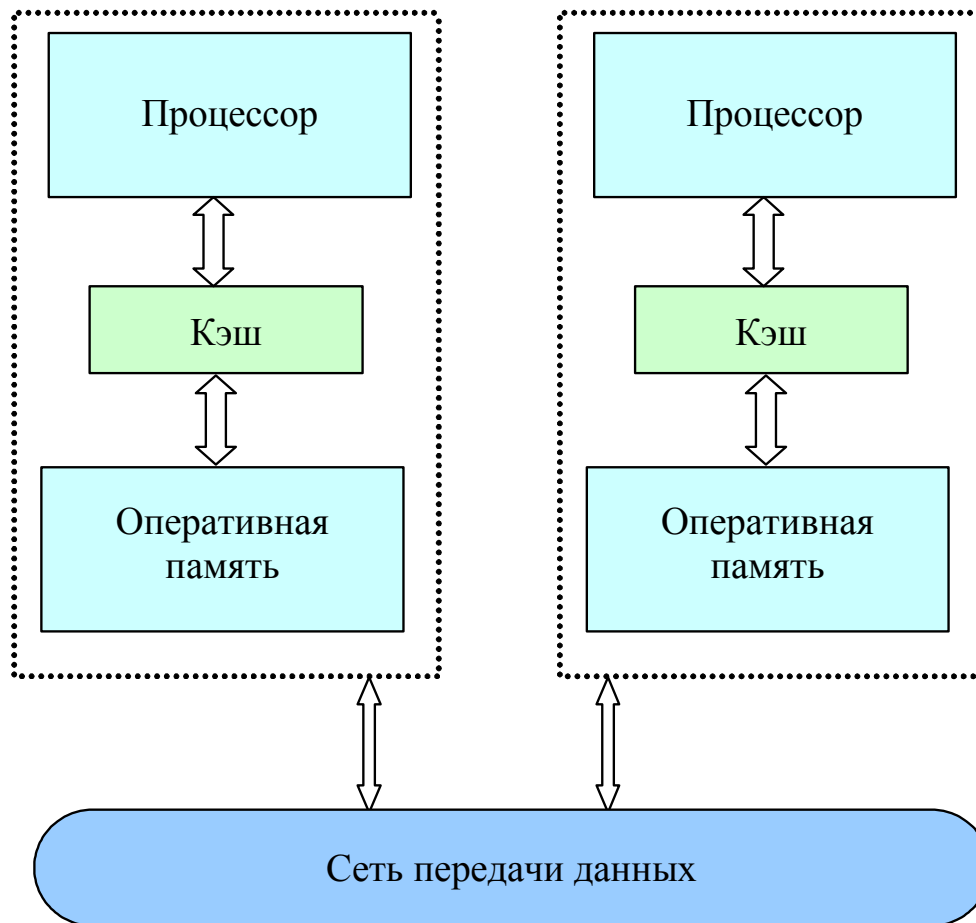
Классификация вычислительных систем...

- **Мультикомпьютеры...**

- не обеспечивают общий доступ ко всей имеющейся в системах памяти (*no-remote memory access or NORMA*),
 - каждый процессор системы может использовать только свою локальную память,
 - для доступа к данным, располагаемым на других процессорах, необходимо явно выполнить *операции передачи сообщений (message passing operations)*.
-

Классификация вычислительных систем...

- Мультикомпьютеры...



Классификация вычислительных систем...

- **Мультикомпьютеры**

Данный подход используется при построении двух важных типов многопроцессорных вычислительных систем:

- *массивно-параллельных систем (massively parallel processor or MPP)*, например: IBM RS/6000 SP2, Intel PARAGON, ASCI Red, транспьютерные системы Parsytec,
 - *кластеров (clusters)*, например: AC3 Velocity и NCSA NT Supercluster.
-

Классификация вычислительных систем...

- **Мультикомпьютеры. Кластеры...**

Кластер - множество отдельных компьютеров, объединенных в сеть, для которых при помощи специальных аппаратно-программных средств обеспечивается возможность унифицированного управления (single system image), надежного функционирования (availability) и эффективного использования (performance)

Классификация вычислительных систем...

- **Мультикомпьютеры. Кластеры...**

Преимущества:

- Могут быть образованы на базе уже существующих у потребителей отдельных компьютеров, либо же сконструированы из типовых компьютерных элементов;
 - Повышение вычислительной мощности отдельных процессоров позволяет строить кластеры из сравнительно небольшого количества отдельных компьютеров (*lowly parallel processing*),
 - Для параллельного выполнения в алгоритмах достаточно выделять только крупные независимые части расчетов (*coarse granularity*).
-

Классификация вычислительных систем

- **Мультикомпьютеры. Кластеры**

Недостатки:

- Организация взаимодействия вычислительных узлов кластера при помощи передачи сообщений обычно приводит к значительным временным задержкам,
 - Дополнительные ограничения на тип разрабатываемых параллельных алгоритмов и программ (*низкая интенсивность потоков передачи данных*)
-

Характеристика типовых схем коммуникации...

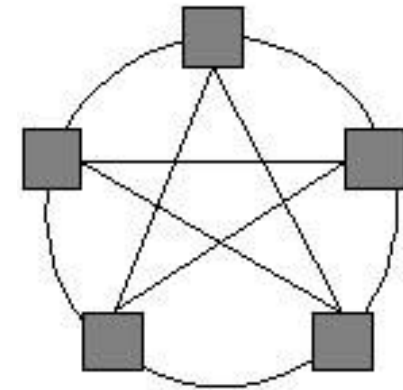
При организации параллельных вычислений в мультикомпьютерах для взаимодействия, синхронизации и взаимоисключения параллельно выполняемых процессов используется передача данных между процессорами вычислительной среды.

Топология сети передачи данных - структура линий коммутации между процессорами вычислительной системы

Характеристика типовых схем коммуникации...

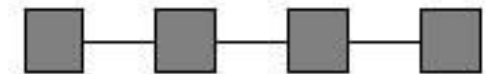
- Топология сети передачи данных...

- **полный граф** (*completely-connected graph* or *clique*) – система, в которой между любой парой процессоров существует прямая линия связи,



Полный граф (*completely-connected graph* or *clique*)

- **линейка** (*linear array* or *farm*) – система, в которой все процессоры перенумерованы по порядку и каждый процессор, кроме первого и последнего, имеет линии связи только с двумя соседними,

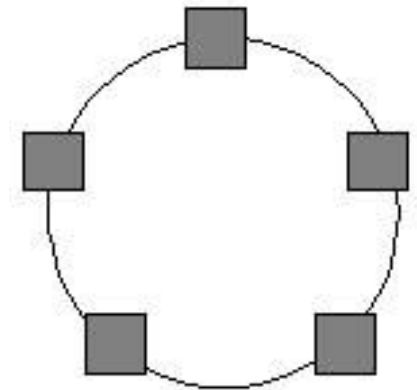


Линейка (*linear array* or *farm*)

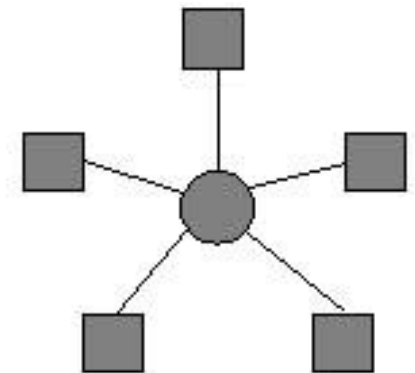
Характеристика типовых схем коммуникации...

- **Топология сети передачи данных...**

- **кольцо** (*ring*) – данная топология получается из линейки процессоров соединением первого и последнего процессоров линейки,
- **звезда** (*star*) – система, в которой все процессоры имеют линии связи с некоторым управляющим процессором,



Кольцо (*ring*)

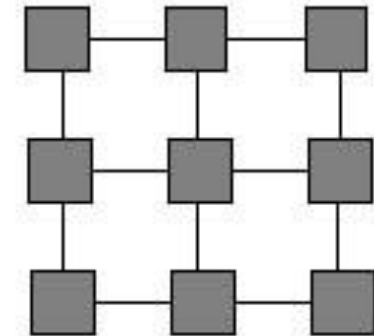


Звезда (*star*)

Характеристика типовых схем коммуникации...

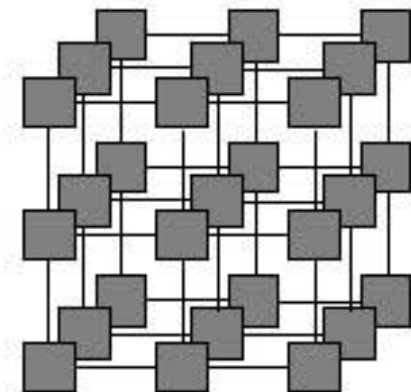
- Топология сети передачи данных...

- **решетка** (*mesh*) – система, в которой граф линий связи образует прямоугольную сетку,



Решетка (*mesh*)

- **гиперкуб** (*hypercube*) – данная топология представляет частный случай структуры решетки, когда по каждой размерности сетки имеется только два процессора.



Характеристика типовых схем коммуникации...

- **Топология сети вычислительных кластеров**

Для построения кластерной системы во многих случаях используют *коммутатор (switch)*, через который процессоры кластера соединяются между собой.

Одновременность выполнения нескольких коммуникационных операций является ограниченной.

*В любой момент времени каждый процессор может принимать участие только в одной операции приема -
Передачи данных*

Характеристика типовых схем коммуникации...

- **Характеристики топологии сети...**

- *диаметр* – максимальное расстояние между двумя процессорами сети; характеризует максимально-необходимое время для передачи данных между процессорами,
 - *связность (connectivity)* – минимальное количество дуг, которое надо удалить для разделения сети передачи данных на две несвязные области,
 - *ширина бинарного деления (bisection width)* – минимальное количество дуг, которое надо удалить для разделения сети передачи данных на две несвязные области одинакового размера,
 - *стоимость* – общее количество линий передачи данных в многопроцессорной вычислительной системе.
-

Характеристика типовых схем коммуникации

- Характеристики топологии сети

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

Характеристика системных платформ для построения кластеров...

- В качестве системной платформы для построения кластеров используют обе наиболее распространенные в настоящий момент операционные системы Unix/Linux и Microsoft Windows.
- Далее подробно будет рассмотрено решение на основе ОС семейства Microsoft Windows; характеристика подхода на базе ОС Unix может быть получена, например, в

Sterling, T. (Ed.) Beowulf Cluster Computing with Linux.
- Cambridge, MA: The MIT Press, 2002.

Характеристика системных платформ для построения кластеров...

Microsoft Compute Cluster Server 2003...

- Интегрированная платформа для поддержки высокопроизводительных вычислений на кластерных системах
 - CCS 2003 состоит из операционной системы Microsoft Windows Server 2003 и Microsoft Compute Cluster Pack (CCP) – набора интерфейсов, утилит и инфраструктуры управления
 - Вместе с CCP поставляется SDK, содержащий необходимые инструменты разработки программ для CCS, включая собственную реализацию MPI (Microsoft MPI)
-

Характеристика системных платформ для построения кластеров...

Microsoft Compute Cluster Server 2003...

- В качестве вычислительных узлов кластера могут быть использованы 64-битные процессоры семейства x86 с, как минимум, 512 Мб оперативной памяти и 4 Гб свободного дискового пространства
 - На вычислительных узлах кластера должна быть установлена операционная система Microsoft Windows Server 2003 (Standard, Enterprise или Compute Cluster Edition)
-

Характеристика системных платформ для построения кластеров

Microsoft Compute Cluster Server 2003

- В состав CCR входит удобная система планирования заданий, позволяющая просматривать состояния всех запущенных задач, собирать статистику, назначать запуски программ на определенное время, завершать "зависшие" задачи и пр.
 - В состав CCR входит Microsoft MPI – версия реализации стандарта MPI 2 от Argonne National Labs. MS MPI совместима с MPICH 2 и поддерживает полнофункциональный API с более чем 160 функциями
 - Microsoft Visual Studio 2005 включает параллельный отладчик, работающий с MS MPI
-

Заключение

- Приведена общая характеристика способов организации параллельных вычислений
 - Рассмотрено различие между многозадачным, параллельным и распределенным режимами выполнения программ
 - Приведен ряд примеров параллельных вычислительных систем
 - Дано описание одного из наиболее известных способов классификации вычислительных систем – *систематики Флинна*
 - Даны ключевые определения *мультипроцессора* и *мультикомпьютера*
 - Рассмотрены основные характеристики сетей передачи данных в многопроцессорных вычислительных системах
-

Анализ эффективности использования параллелизма:

- Принципиальный момент при разработке параллельных алгоритмов - **анализ эффективности использования параллелизма:**
 - Оценка эффективности распараллеливания конкретных выбранных методов выполнения вычислений,
 - Оценка максимально возможного ускорения процесса решения рассматриваемой задачи (анализ всех возможных способов выполнения вычислений)
-

Граф "операции-операнды"...

- Модель в виде графа "операции-операнды" используется для описания существующих информационных зависимостей в выбираемых алгоритмах
 - В наиболее простом виде модель основывается на предположениях:
 - время выполнения любых вычислительных операций является одинаковым и равняется 1,
 - передача данных между вычислительными устройствами выполняется мгновенно без каких-либо затрат времени.
-

Граф "операции-операнды"...

Множество операций, выполняемые в исследуемом алгоритме решения вычислительной задачи, и существующие между операциями информационные зависимости могут быть представлены в виде *ациклического ориентированного графа*

$$G = (V, R)$$

$V = \{1, \dots, |V|\}$ – множество вершин графа, представляющих выполняемые операции алгоритма,

R – множество дуг графа; дуга $r(i, j)$ принадлежит графу только если операция j использует результат выполнения операции i

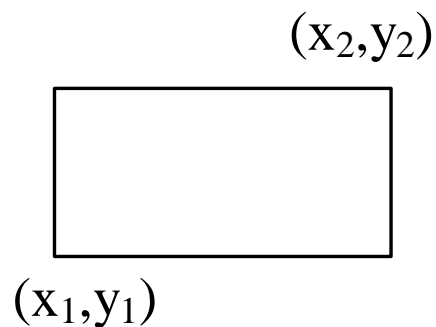
Вершины без входных дуг могут использоваться для задания операций ввода, а вершины без выходных дуг – для операций вывода.

\bar{V} – множество вершин графа без вершин ввода,

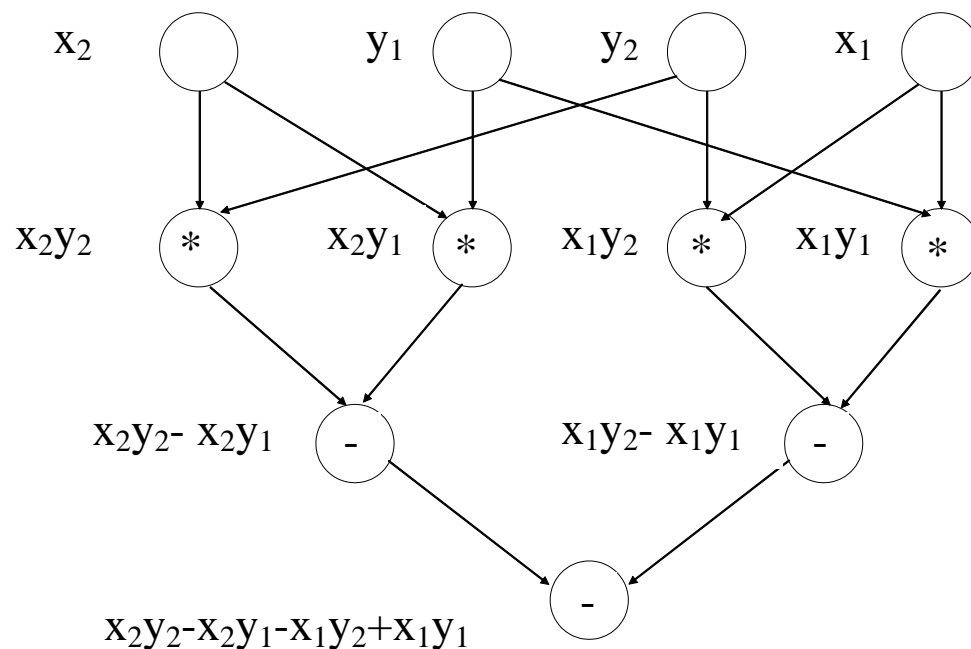
$d(G)$ – диаметр графа (длина максимального пути)

Граф "операции-операнды"...

Пример: граф алгоритма вычисления площади прямоугольника, заданного координатами двух противоположащих углов



$$\begin{aligned} S &= (x_2 - x_1)(y_2 - y_1) = \\ &= x_2 y_2 - x_2 y_1 - x_1 y_2 + x_1 y_1 \end{aligned}$$



Граф "операции-операнды"

- Схемы вычислений обладают различными возможностями для распараллеливания, при построении модели вычислений может быть поставлена **задача выбора наиболее подходящей** для параллельного исполнения вычислительной схемы алгоритма
 - Операции алгоритма, между которыми нет пути в рамках выбранной схемы вычислений, могут быть выполнены **параллельно**
-