

Лекция 2.

Принципы построения параллельных вычислительных систем



Содержание

- ❑ 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 однотипных 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 или при помощи конвейерных и векторных обрабатывающих устройств),
 - *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*



□ 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. МВС-1000...

(Межведомственный Суперкомпьютерный Центр РАН)

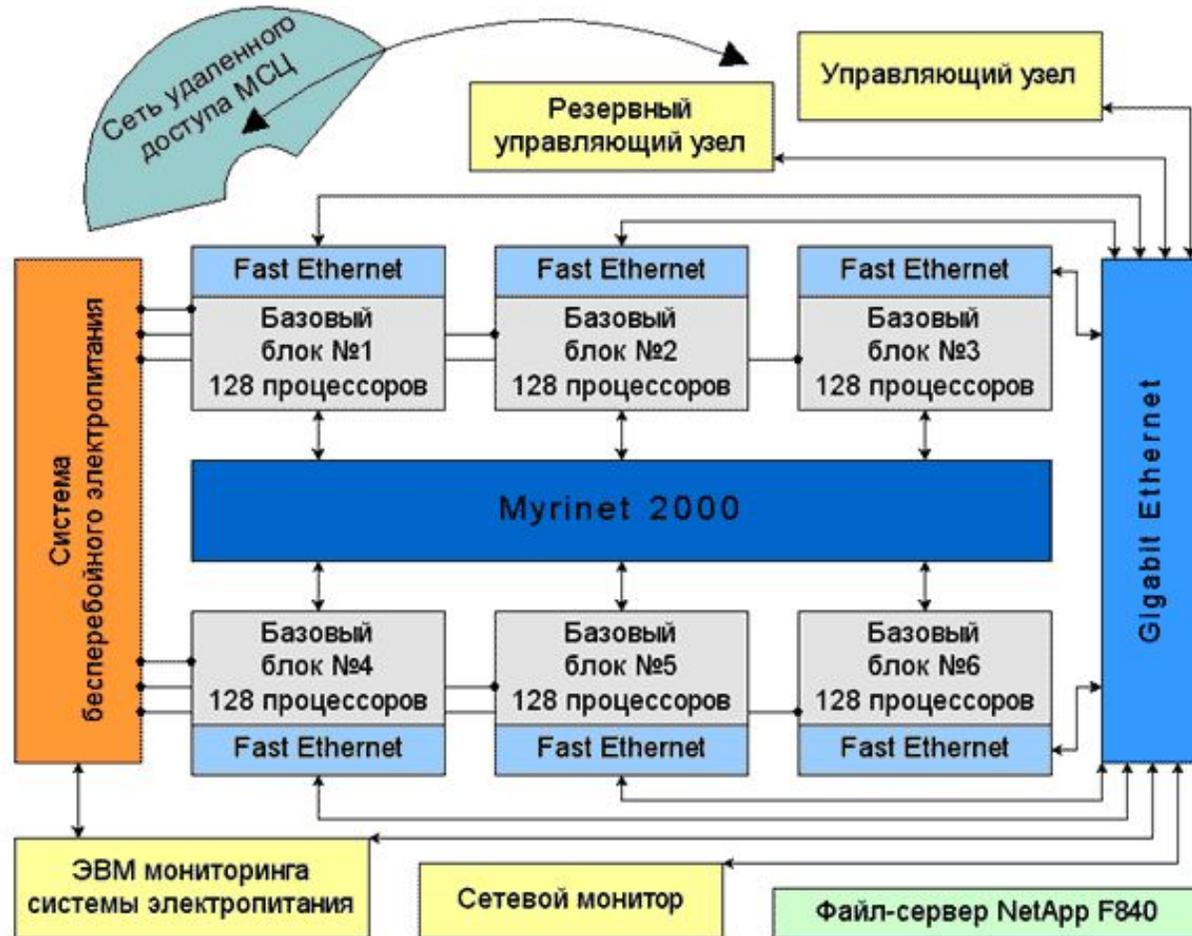
2001

- 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



□ Суперкомпьютеры. МВС-15000...

(Межведомственный Суперкомпьютерный Центр РАН)

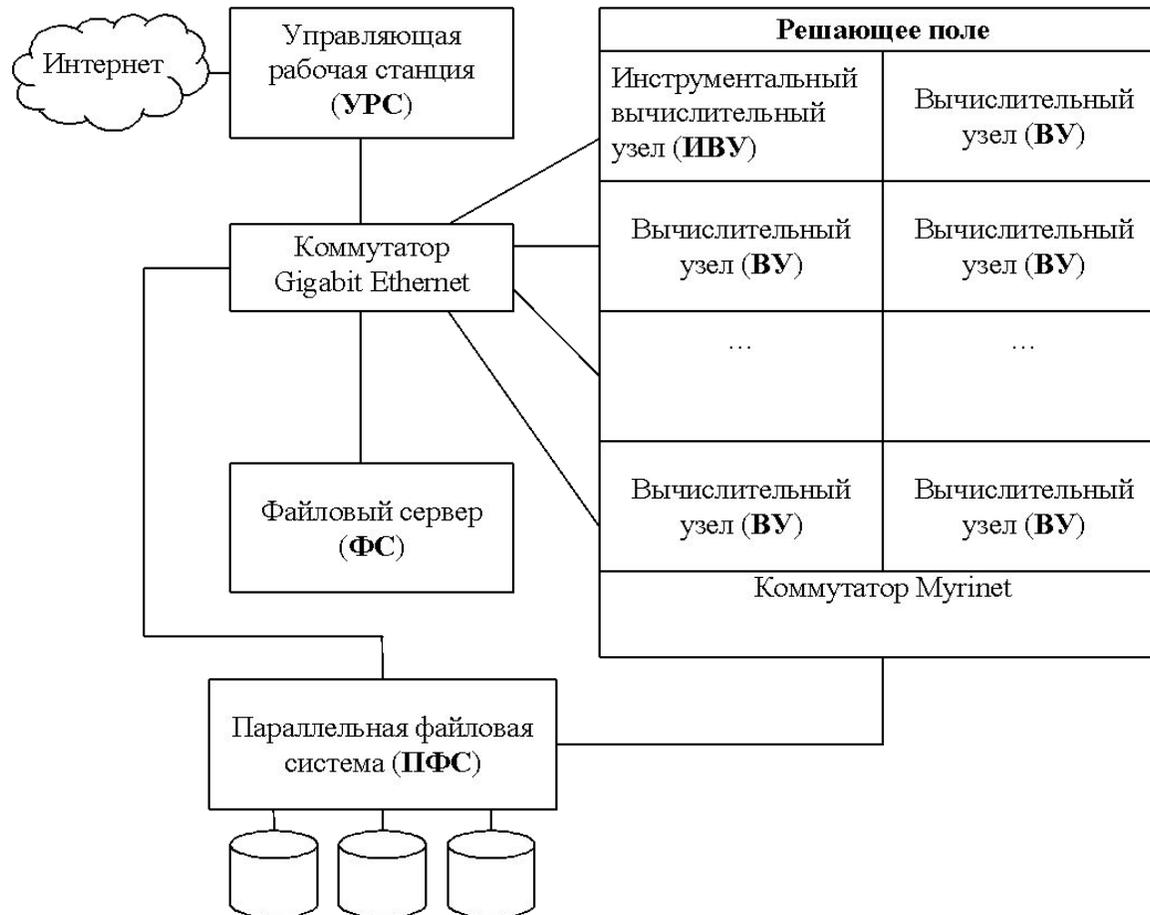
2005

- 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



□ 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).



□ 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.



□ Кластеры. Thunder

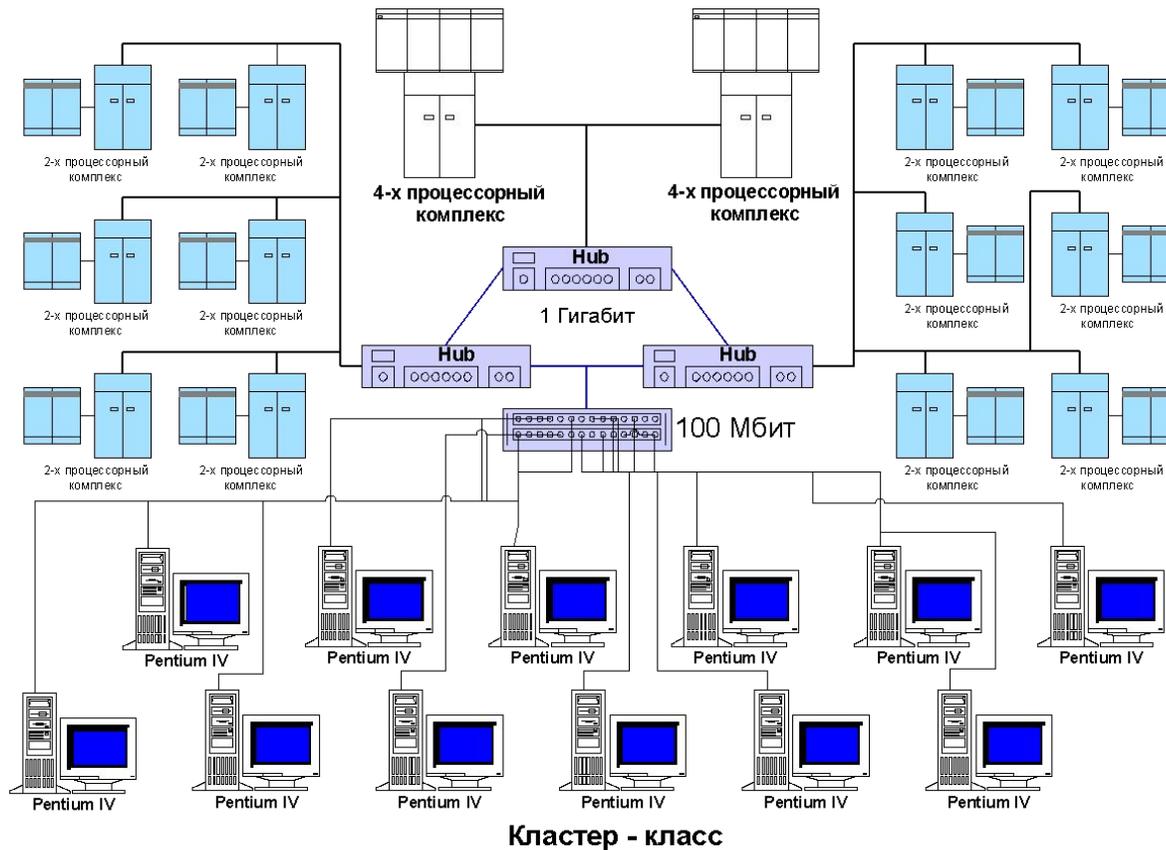
– 2004, Ливерморская Национальная Лаборатория (США):

- 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).



Примеры параллельных вычислительных систем...

Кластеры. Вычислительный кластер ННГУ...



□ Кластеры. Вычислительный кластер ННГУ

– 2007, Модернизация в рамках Инновационной образовательной программы ННГУ:

- 64 вычислительных сервера, каждый из которых имеет 2 двухядерных процессора Intel Core Duo 2,66 GHz, 4 GB RAM, 100 GB HDD, 1 Gbit Ethernet card,
- Пиковая производительность ~3 Tflops
- Операционная система Microsoft Windows.



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***



Классификация вычислительных систем...

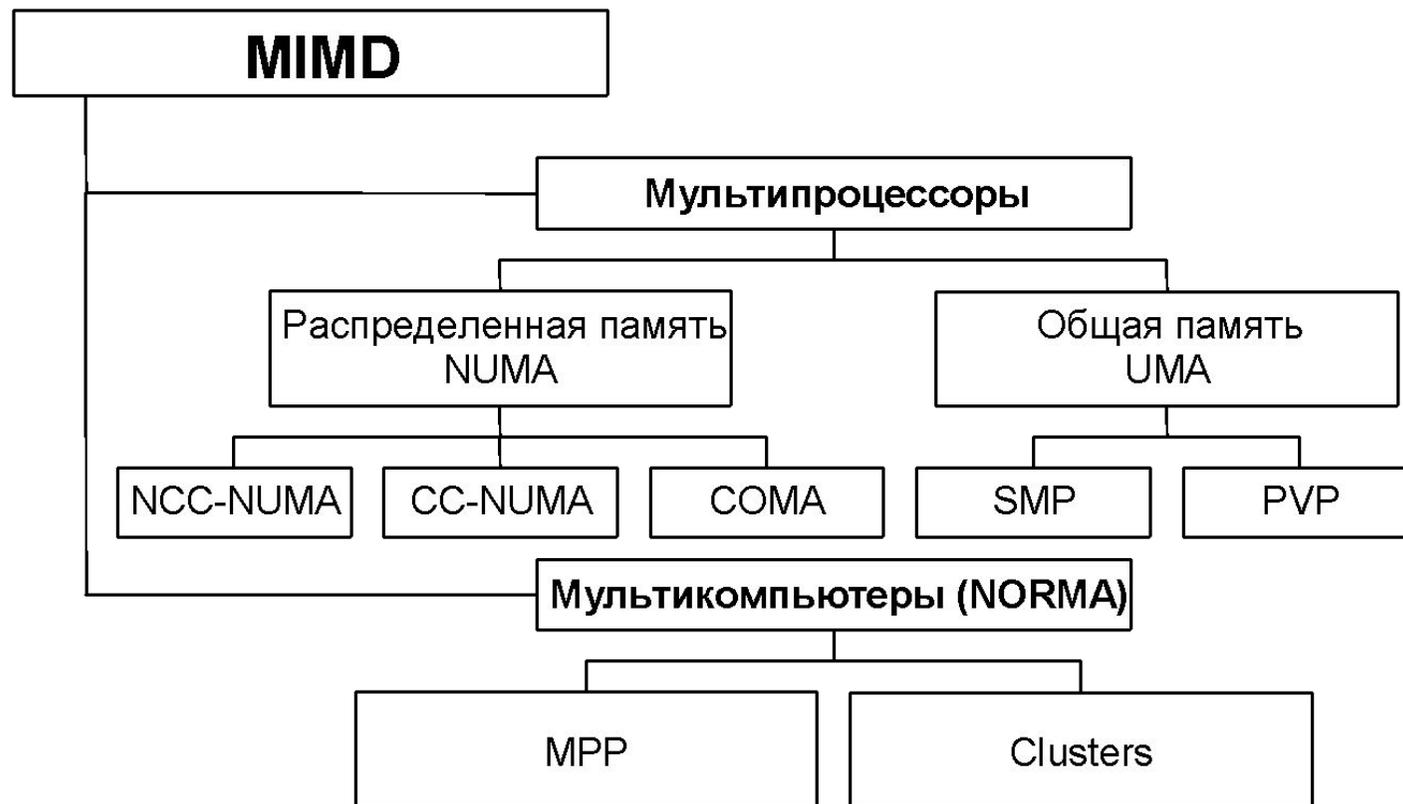
□ Детализация систематики Флинна...

- 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...



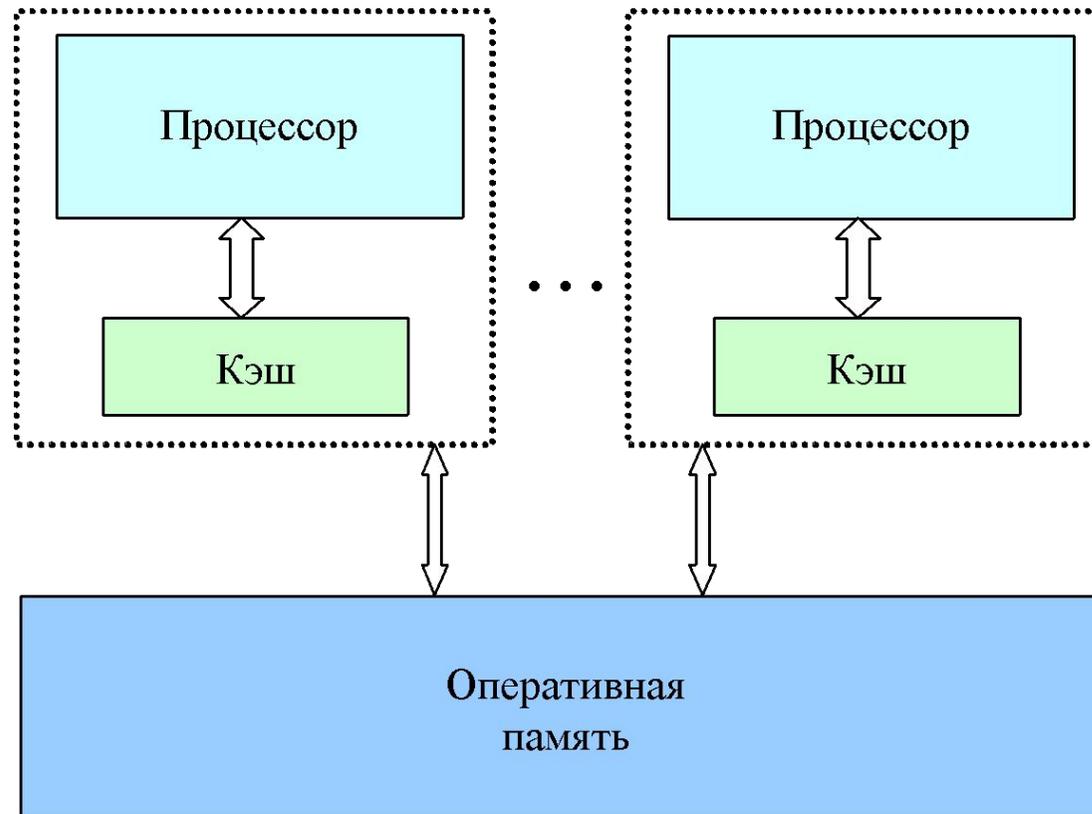
Классификация вычислительных систем...

- **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...

- ❑ **Мультимикропроцессоры с использованием единой общей памяти...**



Классификация вычислительных систем...

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

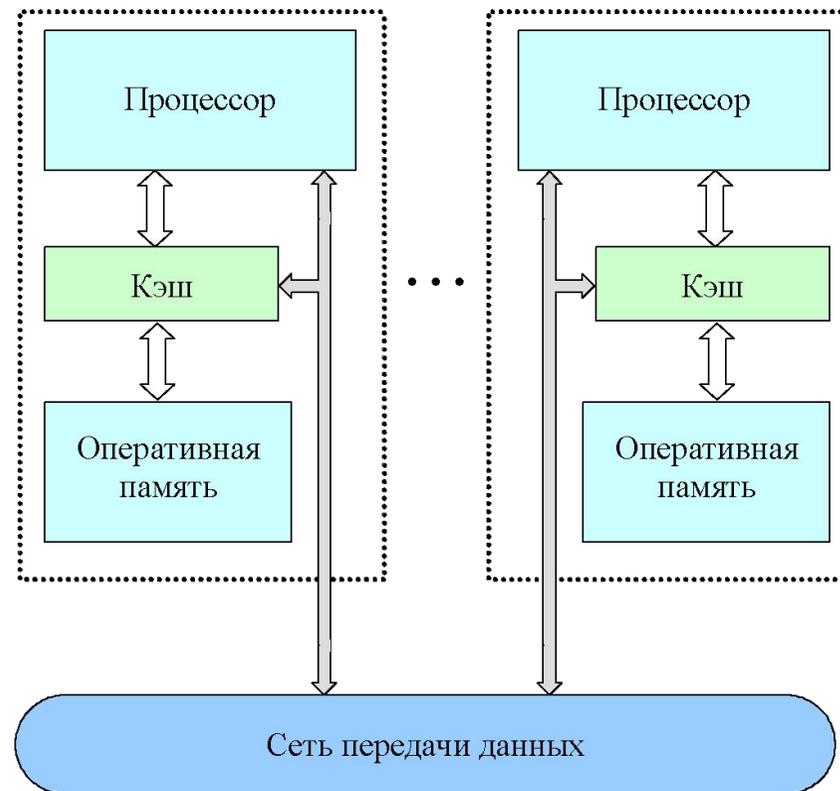
Проблемы:

- 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



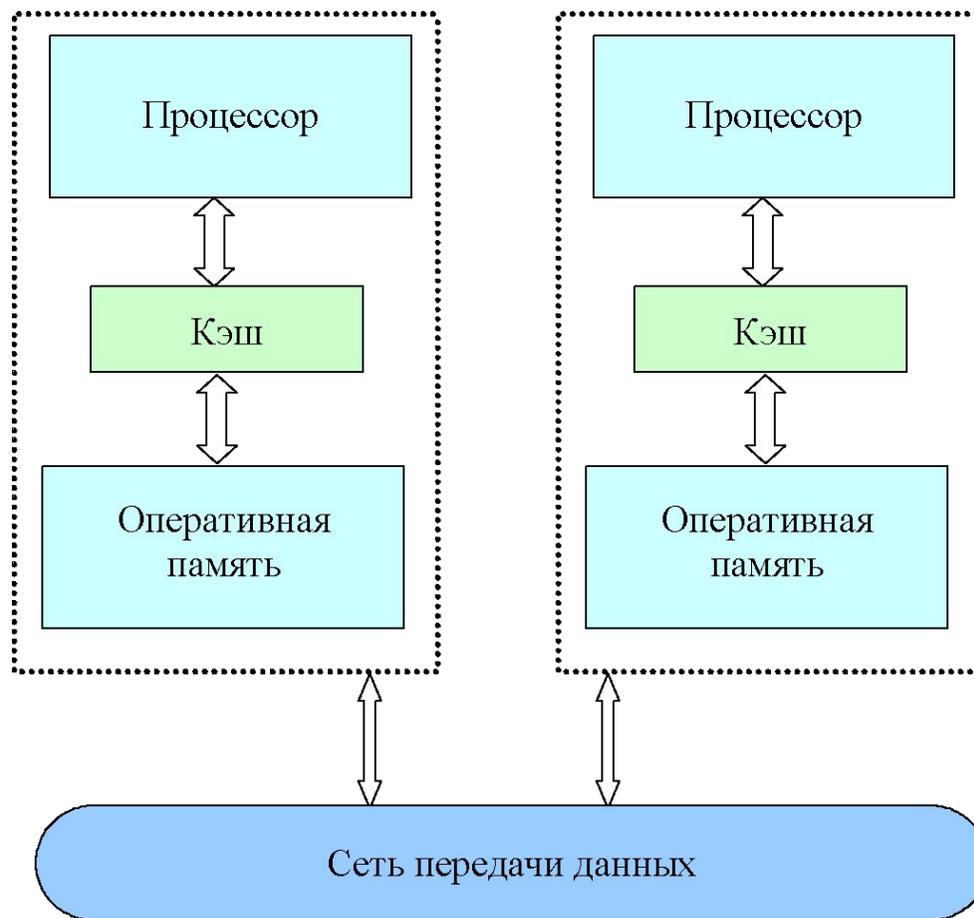
Классификация вычислительных систем...

- ❑ **Мультимикропроцессоры с использованием физически распределенной памяти...**



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.

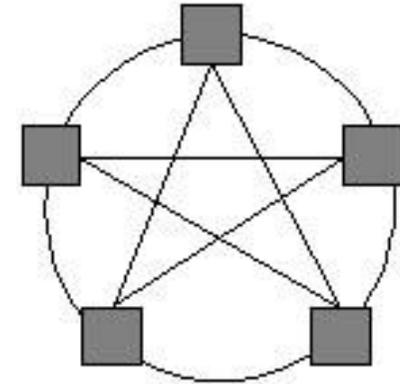
Топология сети передачи данных - структура линий коммутации между процессорами вычислительной системы



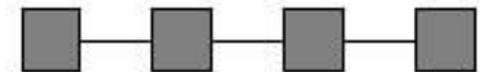
Характеристика типовых схем коммуникации...

□ Топология сети передачи данных...

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



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

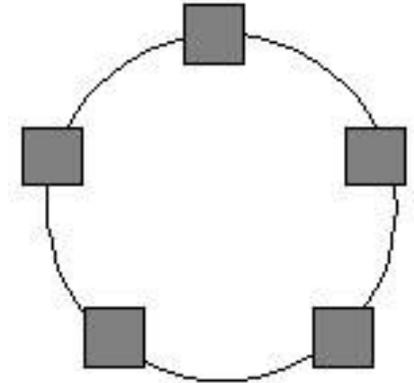


Линейка (*linear array or farm*)

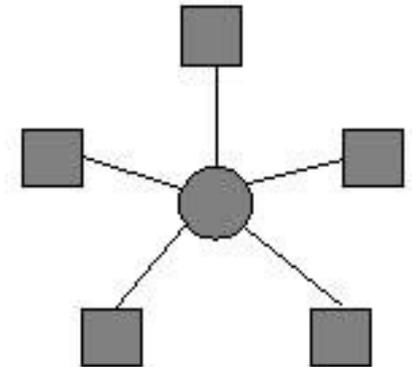
Характеристика типовых схем коммуникации...

□ 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*)

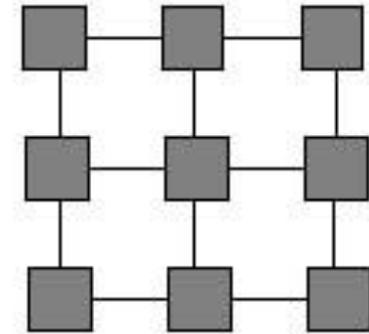


Звезда (*star*)

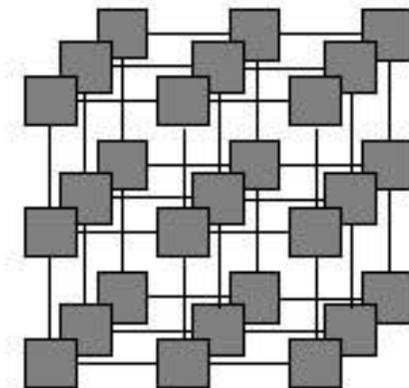
Характеристика типовых схем коммуникации...

□ Топология сети передачи данных...

- **решетка** (*mesh*) – система, в которой граф линий связи образует прямоугольную сетку,
- **гиперкуб** (*hypercube*) – данная топология представляет частный случай структуры решетки, когда по каждой размерности сетки имеется только два процессора.



Решетка (*mesh*)



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 | $2p$ |



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



Вопросы для обсуждения

- В чем заключаются основные способы достижения параллелизма?
- В чем могут состоять различия параллельных вычислительных систем?
- Что положено в основу классификация Флинна?
- В чем состоит принцип разделения многопроцессорных систем на мультипроцессоры и мультикомпьютеры?
- В чем состоят положительные и отрицательные стороны кластерных систем?
- Какие топологии сетей передачи данных наиболее широко используются при построении многопроцессорных систем?
- В чем состоят особенности сетей передачи данных для кластеров?
- Каковы основные характеристики сетей передачи данных?
- Какие системные платформы могут быть использованы для построения кластеров?



Темы заданий для самостоятельной работы

- ❑ Приведите дополнительные примеры параллельных вычислительных систем
- ❑ Выполните рассмотрение дополнительных способов классификации компьютерных систем
- ❑ Рассмотрите способы обеспечения когерентности кэшей в системах с общей разделяемой памятью
- ❑ Подготовьте обзор программных библиотек, обеспечивающих выполнение операций передачи данных для систем с распределенной памятью
- ❑ Рассмотрите топологию сети передачи данных в виде двоичного дерева
- ❑ Выделите эффективно реализуемые классы задач для каждого типа топологий сети передачи данных



Литература...

- ❑ **Гергель В.П.** Теория и практика параллельных вычислений. - М.: Интернет-Университет, БИНОМ. Лаборатория знаний, 2007.
- ❑ **Воеводин В.В., Воеводин Вл.В.** Параллельные вычисления. – СПб.: БХВ-Петербург, 2002.
- ❑ **Корнеев В.В.** Параллельные вычислительные системы. – М.: Нолидж, 1999.
- ❑ **Таненбаум Э.** (2002) . Архитектура компьютера. – СПб.: Питер.

