

# Массивно-параллельные системы (МРР).



# MPP архитектура (massive parallel processing) - массивно-параллельная архитектура.

- **Особенность** - память физически разделена. Система строится из отдельных модулей (процессор, операционная память(ОП), два коммуникационных процессора для передачи команд и данных, или сетевой адаптер, иногда - жесткие диски и/или другие устройства ввода/вывода). По сути это - полнофункциональные компьютеры. Модули соединяются специальными коммуникационными каналами.

# MPP архитектура (massive parallel processing) - массивно-параллельная архитектура.

- **Главное преимущество:** хорошая масштабируемость (отличие от SMP-систем каждый процессор имеет доступ только к своей локальной памяти, в связи с чем не возникает необходимости в потактовой синхронизации процессоров). Практически все рекорды по производительности на сегодняшний день устанавливаются на машинах именно такой архитектуры, состоящих из нескольких тысяч процессоров (ASCI Red, ASCI Blue Pacific).

# MPP архитектура (massive parallel processing) - массивно-параллельная архитектура.

- **Недостатки:**
- •Отсутствие общей памяти заметно снижает скорость межпроцессорного обмена.
- •Каждый процессор может использовать только ограниченный объем памяти.
- •Архитектурные недостатки => трудо усилия, чтобы максимально использовать системные ресурсы => высокая цена.

# Операционная система и способ программирования

- Существуют два основных варианта:
- Полноценная ОС работает только на управляющей машине (front-end), на каждом узле работает сильно урезанный вариант ОС, обеспечивающие только работу расположенной в нем ветви параллельного приложения. Пример: Cray T3E.
- На каждом узле работает полноценная UNIX-подобная ОС (вариант, близкий к кластерному подходу). Пример: IBM RS/6000 SP + ОС AIX, устанавливаемая отдельно на каждом узле.
- Модель программирования Программирование в рамках модели передачи сообщений ( MPI, PVM, BSPlib)

# Немного истории=)

- С самого начала компьютерной эры существовала необходимость во все более и более производительных системах. В основном это достигалось в результате эволюции технологий производства компьютеров. Наряду с этим имели место попытки использовать несколько процессоров в одной вычислительной системе в расчете на то, что будет достигнуто соответствующее увеличение производительности. Первой такой попыткой, осуществленной в начале 70-х годов, является ILLIAC IV. Сейчас имеется масса параллельных компьютеров и проектов их реализации.

# ILLIAC IV



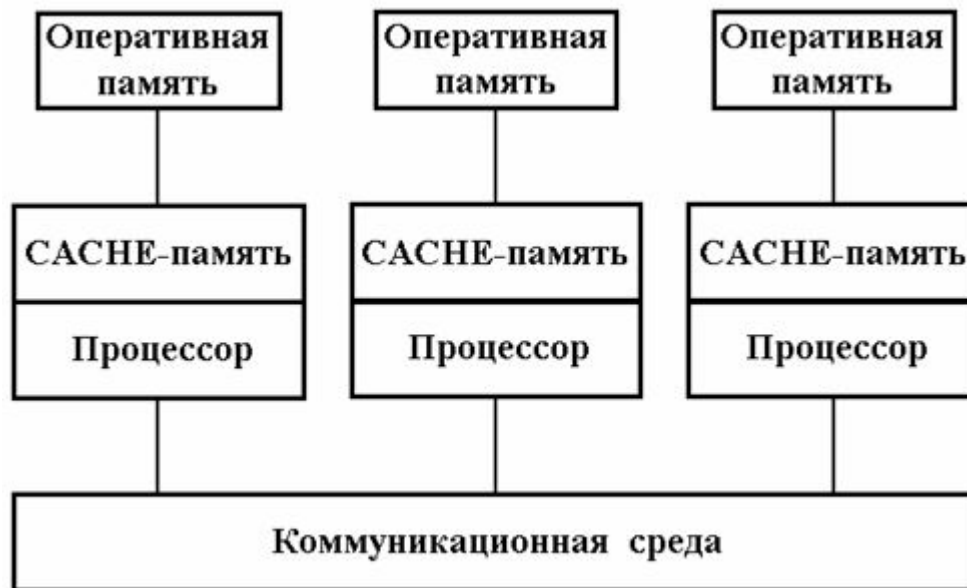
Copyright (c) 2005 Robert A. Walker

## Способ доступа к модулям памяти

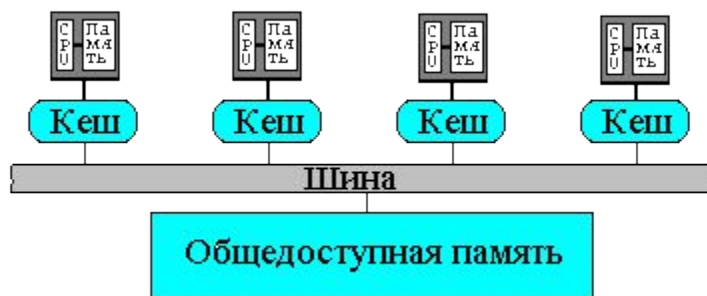
Имеет ли каждый процессор локальную память и обращается к другим блокам памяти, используя коммутирующую сеть\*, или коммутирующая сеть соединяет все процессоры с общей памятью\*. Исходя из способа доступа к памяти, различают следующие (довольно условные) типы параллельных (MIMD) архитектур:



# Система с распределенной памятью



# Коммутативная сеть с общей памятью.



# Способ доступа к модулям памяти

- 1) Компьютеры с распределенной памятью (Distributed memory)
  - Каждый процессор имеет доступ только к локальной собственной памяти. Процессоры объединены в сеть. Доступ к удаленной памяти возможен только с помощью системы обмена сообщениями.

# Способ доступа к модулям памяти

## 2) Компьютеры с общей (разделяемой) памятью (True shared memory)

- Каждый процессор компьютера обладает возможностью прямого доступа к общей памяти, используя общую шину (возможно, реализованную на основе высокоскоростной сети). В таких компьютерах нельзя существенно увеличить число процессоров, поскольку при этом происходит резкое увеличение числа конфликтов доступа к шине.
- В некоторых архитектурах каждый процессор имеет как прямой доступ к общей памяти, так и собственную локальную память.

## Способ доступа к модулям памяти

- 3) Компьютеры с виртуальной общей (разделяемой) памятью (Virtual shared memory)
- В таких системах общая память как таковая отсутствует. Каждый процессор имеет собственную локальную память. Он может обращаться к локальной памяти других процессоров, используя "глобальный адрес". Если "глобальный адрес" указывает не на локальную память, то доступ к памяти реализуется с помощью сообщений с малой задержкой, пересылаемых по сети, соединяющей процессоры.

# И на последок

- Массивно-параллельные системы остаются достаточно дорогими. Однако заложенная в них идея - использование большого числа процессоров, каждый из которых имеет собственную локальную память и подключен к сети передачи данных, - может быть реализована иначе. Суперкомпьютером будет служить сеть рабочих станций. В ряде прикладных пакетов эта идея уже воплощена. Вычисления распределяются по сети и выполняются на станциях в фоновом режиме. Разумеется, далеко не все задачи удастся распараллелить таким образом, чтобы обычная локальная сеть могла обеспечить требуемый уровень передачи сообщений между процессами. Однако тот факт, что многие пакеты передачи сообщений, в том числе и MPI, существуют в варианте для сетей рабочих станций, говорит о том, что этот вариант имеет право на существование. Вполне вероятно, что с появлением на персональных компьютерах защищенных операционных систем, таких как Windows NT или OS/2, подобные пакеты будут реализованы и для них, тогда локальная сеть учреждения сможет функционировать как распределенный суперкомпьютер.

# Несколько примеров

1) - CRAY T3E - массивно-параллельный компьютер фирмы Тега



# Несколько примеров

## 1) - CRAY T3E - массивно-параллельный компьютер фирмы Тега

- Каждый узел в системе содержит один процессорный элемент (ПЭ), включающий процессор, память и средство коммутации, которое осуществляет связь между ПЭ. Система конфигурируется до 2048 процессоров. Пиковая производительность составляет 2,4 Тфлоп. Разделяемая, высокопроизводительная, глобально адресуемая подсистема памяти делает возможным обращение к локальной памяти каждого ПЭ в Cray T3E. Процессорные элементы в системе Cray T3E связаны в трехмерный тор двунаправленной высокоскоростной сетью с малым временем задержки, которая в шесть раз превосходит по скорости аналогичную сеть в Cray T3D. Также добавлена адаптивная маршрутизация, при которой возможен обход участков с высокой эффективностью передачи.
- Системы Cray T3E выполняют операции ввода/вывода через многочисленные порты на один или более каналов GigaRing. Каналы ввода/вывода интегрированы в трехмерную межузловую сеть и пропорциональны размеру системы. При этом при добавлении ПЭ пропускная способность каналов ввода/вывода увеличивается, и масштабируемые приложения могут выполняться на системах с большим числом процессоров так же эффективно, как на системах с меньшим числом процессоров.



# Несколько примеров

## 2) Cray T3D



# Несколько примеров

## Общая структура компьютера CRAY T3D

- Компьютер CRAY T3D - это массивно-параллельный компьютер с распределенной памятью, объединяющий от 32 до 2048 процессоров. Распределенность памяти означает то, что каждый процессор имеет непосредственный доступ только к своей локальной памяти, а доступ к данным, расположенным в памяти других процессоров, выполняется другими, более сложными способами.
- CRAY T3D подключается к хост-компьютеру (главному или ведущему), роль которого, в частности, может исполнять CRAY Y-MP C90. Вся предварительная обработка и подготовка программ, выполняемых на CRAY T3D, проходит на хосте (например, компиляция). Связь хост-машины и T3D идет через высокоскоростной канал передачи данных с производительностью 200 Мбайт/с.
- Массивно-параллельный компьютер CRAY T3D работает на тактовой частоте 150MHz и имеет в своем составе три основные компоненты: сеть межпроцессорного взаимодействия (или по-другому коммуникационную сеть), вычислительные узлы и узлы ввода/вывода.

# Несколько примеров

## 3) Сборка массивно-параллельной системы Red Storm

