

Лекция 7

Многомашинные вычислительные системы

Многомашинная ВС

Многомашинная ВС (ММС) содержит несколько ЭВМ, каждая из которых имеет свою ОП и работает под управлением своей операционной системы, а также средства обмена информацией между машинами. Компьютеры могут находиться рядом друг с другом или быть удаленными друг от друга на некоторое, иногда значительное расстояние (например, вычислительные сети). Реализация обмена информацией происходит, в конечном счете, путем взаимодействия операционных систем машин между собой. Это ухудшает динамические характеристики процессов межмашинного обмена данными. Применение многомашинных систем позволяет повысить надежность вычислительных комплексов. При отказе в одной машине обработку данных может продолжать другая машина комплекса. Однако можно заметить, что при этом оборудование комплекса недостаточно эффективно используется для этой цели.

Различия

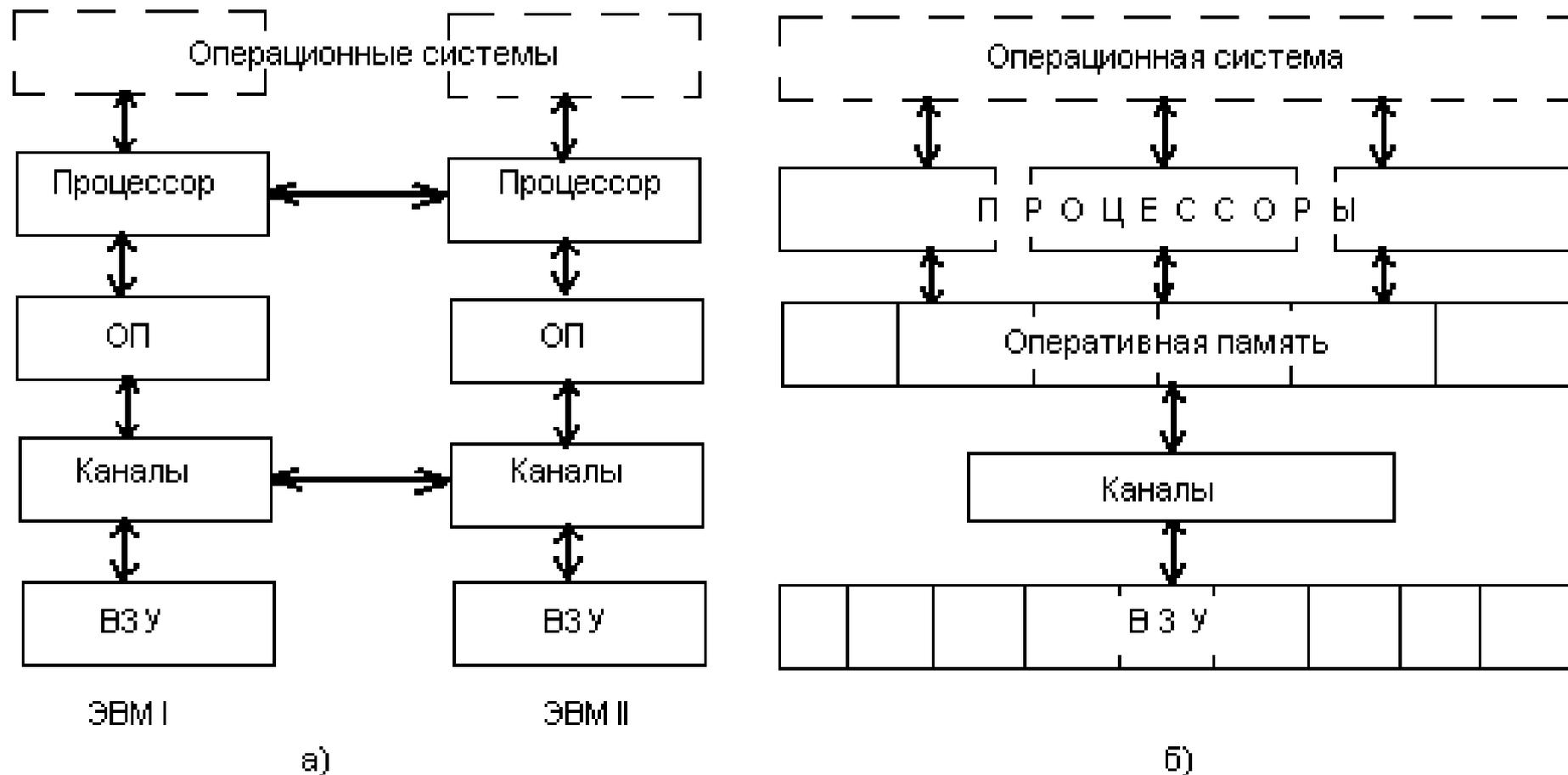


Рис. 6.1 Многомашинные(а) и многопроцессорные(б) системы.

Виды взаимодействия

Информационное взаимодействие компьютеров в многомашиной ВС организовывается на следующих уровнях:

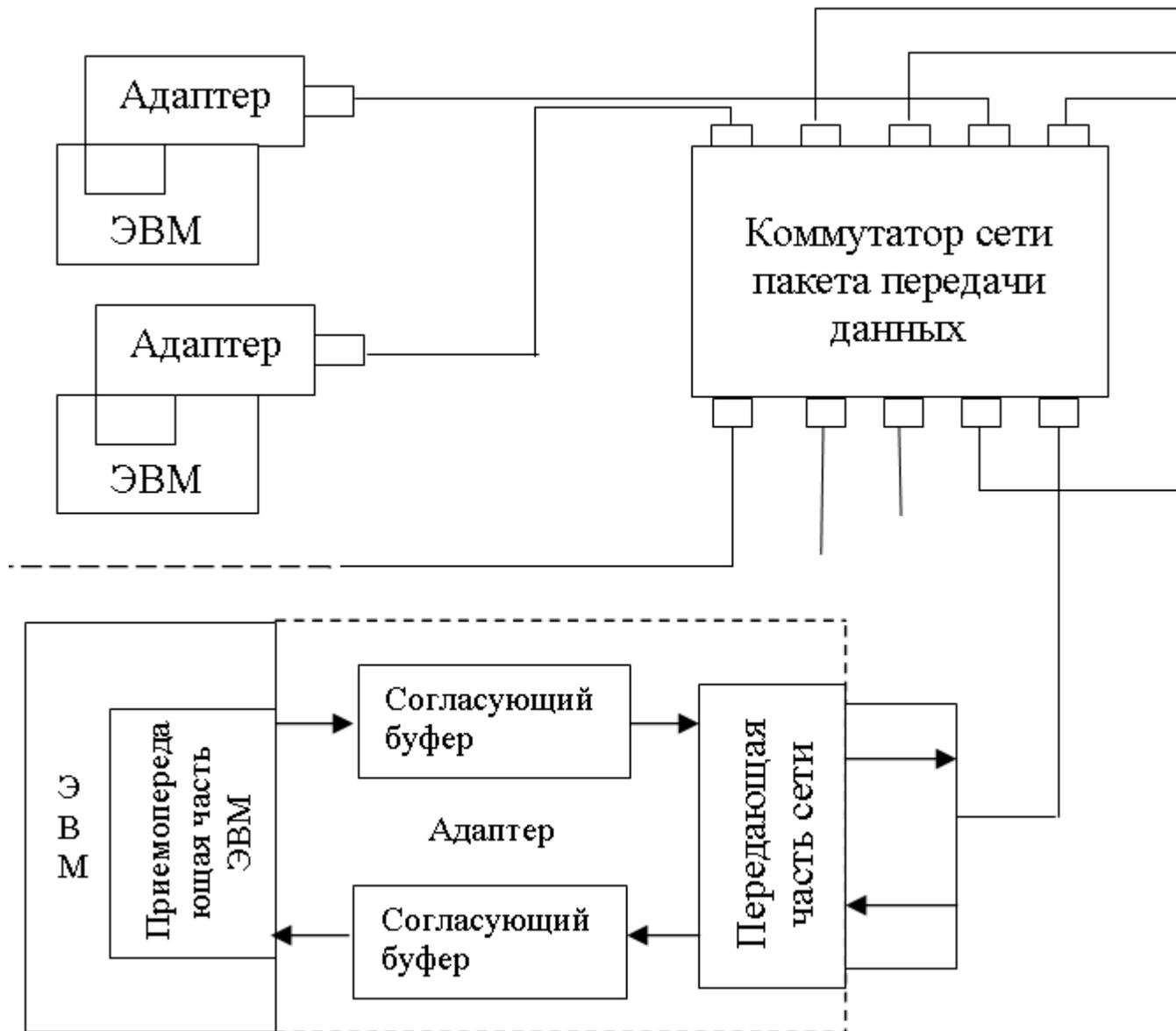
- процессоров;
- оперативной памяти;
- каналов связи.

Коммуникационная среда

На структурном уровне коммуникационная среда сети состоит из трех компонентов:

- адаптеров, осуществляющих интерфейс между вычислительной машиной и сетью передачи пакетов;
- коммутаторов сети передачи пакетов;
- кабелей, служащих для подсоединения входных и выходных каналов (линков) адаптеров к портам коммутаторов и соединения коммутаторов друг с другом для обеспечения необходимого числа портов и заданной конфигурации сети.

Коммуникационная среда

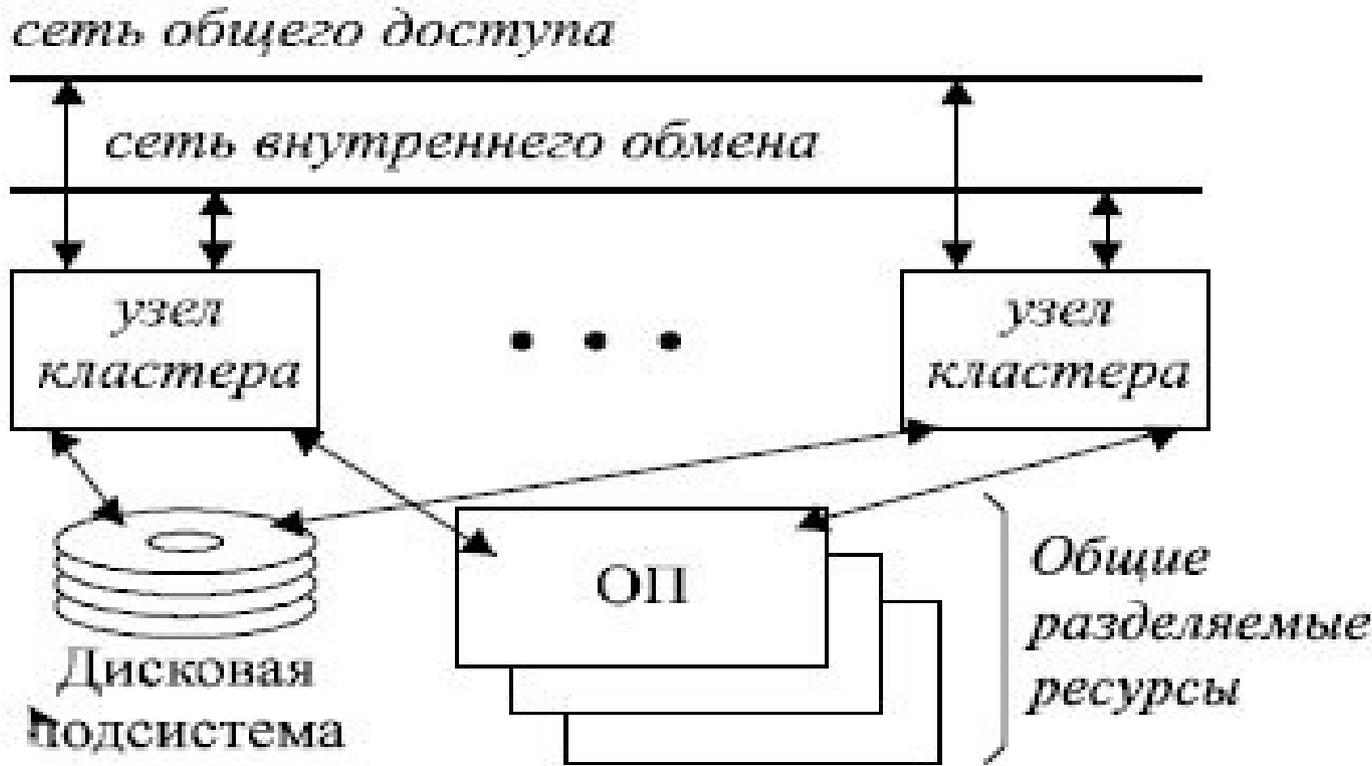


Кластеры

Кластер представляет собой систему из нескольких компьютеров (в большинстве случаев серийно выпускаемых), имеющих общий разделяемый ресурс для хранения совместно обрабатываемых данных (обычно набор дисков или дисковых массивов) и объединенных высокоскоростной магистралью.

Обычно в кластерных системах не обеспечивается единая ОС для работы общего набора приложений на всех узлах кластера. То есть каждый компьютер - это автономная система с отдельным экземпляром ОС и своими, принадлежащими только ей системными ресурсами. Приложение, запущенное на нем, может видеть только общие диски или отдельные участки памяти. На узлах кластера работают специально написанные приложения, параллельно обрабатывающие общий набор данных. Таким образом, кластерное ПО - это лишь средство для взаимодействия узлов и синхронизации доступа к общим данным. Кластер как параллельная система формируется на прикладном уровне, а не на уровне ОС.

Кластеры



Кластеры

В настоящее время такие системы имеют две основные области применения: *параллельные серверы баз данных* и *высоконадежные вычислительные комплексы*. Рынок параллельных СУБД и есть основной рынок кластеров приложений. Высоконадежные комплексы представляют собой группу узлов, на которых независимо друг от друга выполняются некоторые важные приложения, требующие постоянной, непрерывной работы. То есть в такой системе на аппаратном уровне фактически поддерживается основной механизм повышения надежности - резервирование. Узлы находятся в так называемом "горячем" резерве, и каждый из них в любой момент готов продолжить вычисления при выходе из строя какого-либо узла. При этом все приложения с отказавшего узла автоматически переносятся на другие машины комплекса. Такая система также *формально* является кластером, хотя в ней отсутствует параллельная обработка общих данных. Эти данные обычно монопольно используются выполняемыми в рамках кластера приложениями и должны быть доступны для всех узлов.

Кластеры

Кластерные системы обладают следующими преимуществами:

- абсолютная масштабируемость.
- наращивание в процессе эксплуатации.
- высокая надежность.
- снижение соотношения цена / производительность.



Кластеры

Широко распространённым средством для организации межсерверного взаимодействия является библиотека MPI, поддерживающая языки C и Fortran. Она используется, например, в программе моделирования погоды MM5.

Solaris предоставляет программное обеспечение Solaris Cluster, которое служит для обеспечения высокой доступности и безотказности серверов. Для OpenSolaris существует реализация под названием OpenSolaris HA Cluster.

В GNU/Linux популярны несколько программ:

- distcc, MPICH и др. — специализированные средства для распараллеливания работы программ.
- Linux Virtual Server, Linux-HA — узловое ПО для распределения запросов между вычислительными серверами.
- MOSIX, openMosix, Kerrighed, OpenSSI — полнофункциональные кластерные среды, встроенные в ядро, автоматически распределяющие задачи между однородными узлами. OpenSSI, openMosix и Kerrighed создают среду единой операционной системы между узлами.

Кластеры

Компанией Microsoft выпускается HA-кластер для операционной системы Windows. Windows Compute Cluster Server 2003 (CCS), выпущенный в июне 2006 года разработан для высокотехнологичных приложений, которые требуют кластерных вычислений. Издание разработано для развертывания на множестве компьютеров, которые собираются в кластер для достижения мощностей суперкомпьютера. Каждый кластер на Windows Compute Cluster Server состоит из одного или нескольких управляющих машин, распределяющих задания и нескольких подчиненных машин, выполняющих основную работу. В ноябре 2008 был представлен Windows HPC Server 2008, призванный заменить Windows Compute Cluster Server 2003.

Работа любой многомашинной системы определяется двумя главными компонентами: высокоскоростным механизмом связи процессоров и системным программным обеспечением, которое предоставляет пользователям и приложениям прозрачный доступ к ресурсам всех компьютеров, входящих в комплекс. В состав средств связи входят программные модули, которые занимаются распределением вычислительной нагрузки, синхронизацией вычислений и реконфигурацией системы. Если происходит отказ одного из компьютеров комплекса, его задачи могут быть автоматически переназначены и выполнены на другом компьютере. Если в состав многомашинной системы входят несколько контроллеров внешних устройств, то в случае отказа одного из них, другие контроллеры автоматически подхватывают его работу. Таким образом, достигается высокая отказоустойчивость комплекса в целом.

Помимо повышения отказоустойчивости, многомашинные системы позволяют достичь высокой производительности за счет организации параллельных вычислений. По сравнению с мультипроцессорными системами возможности параллельной обработки в многомашинных системах ограничены: эффективность распараллеливания резко снижается, если параллельно выполняемые задачи тесно связаны между собой по данным. Это объясняется тем, что связь между компьютерами многомашинной системы менее тесная, чем между процессорами в мультипроцессорной системе, так как основной обмен данными осуществляется через общие многоходовые периферийные устройства. Говорят, что в отличие от мультипроцессоров, где используются сильные программные и аппаратные связи, в многомашинных системах аппаратные и программные связи между обрабатывающими устройствами являются более слабыми.

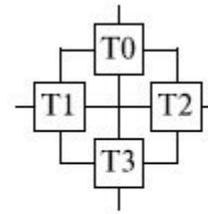
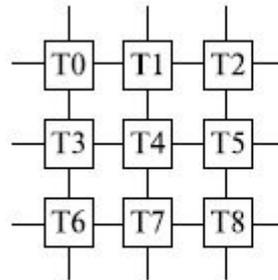
Транспьютеры

Для построения многопроцессорных систем могут быть использованы специально разработанные процессоры, называемые транспьютерами. Они были созданы в середине 1980-х годов фирмой INMOS Ltd.

Транспьютер - это микропроцессор со встроенными средствами межпроцессорной коммуникации, предназначенной для построения многопроцессорных систем. Его название происходит от слов TRansfer (передатчик) и computer (вычислитель).

Транспьютер включает в себя средства для выполнения вычислений (ЦП, АЛУ с плавающей точкой, внутрикристальную память) и 4 канала для связи (линка) с другими транспьютерами и/или другими устройствами. Каждый линк представляет собой 2 однонаправленных последовательных канала передачи информации. Встроенный интерфейс позволяет подключать внешнюю память емкостью до 4 Гбайт

Транспьютеры



Транспьютеры

Организация транспьютеров и многопроцессорных систем на их основе базируется на языке Оссам, специально разработанном в начале 1980-х годов группой ученых из Оксфорда под руководством Дэвида Мэя по заданию компании INMOS в рамках работ по созданию транспьютеров. Отметим основные характеристики этого языка с точки зрения его применения для программирования транспьютерных систем: выполнение задачи разбивается на процессы, которые могут выполняться параллельно; размещение процессов не привязано к конкретному оборудованию: необязательно каждому процессу должен ставиться в соответствие свой транспьютер - несколько процессов могут выполняться на одном транспьютере; язык позволяет описать размещение процессов по оборудованию; язык позволяет описать, как эти процессы обмениваются между собой данными (какой процесс принимает информацию, от какого процесса и в каком объеме). *(Развитие остановилось после появления PentiumPro)*

ИСТОЧНИКИ

1. Лекции Кашковского В.В.
2. http://www.cap-design.ru/ksptp/1_1_2_2.htm
3. <http://wiki.mvtom.ru/index.php>
/Мультипроцессорные_и_многомашинные_системы
4. <http://ru.wikipedia.org/wiki/>