Оглавление

[Основные понятия ОС, архитектура и процессы в ОС 1](#_Toc165649585)

[Понятие ОС 1](#_Toc165649586)

[Эволюция операционных систем. 2](#_Toc165649587)

[Функции операционной системы. 6](#_Toc165649588)

[Операционная система как расширенная машина и менеджер ресурсов. 6](#_Toc165649589)

[ОС как расширенная машина 7](#_Toc165649590)

[ОС как система управления ресурсами 7](#_Toc165649591)

[Классификация операционных систем 8](#_Toc165649592)

# Основные понятия ОС, архитектура и процессы в ОС

## Понятие ОС

Под ОС обычно понимают комплект управляющих программ, которые вступают как интерфейс между аппаратурой компьютера, и предназначены для наиболее эффективного использования ресурсов вычислительной системы, организации надежных вычислений. Любой из компонентов программного обеспечения работает под управлением ОС, и ни один из компонентов ПО, за исключением самой ОС не имеет доступа непосредственно к аппаратуре.

ОС – это программа, контролирующая работу пользовательской программы и систем приложений и исполняемая роль интерфейса между приложениями и аппаратным обеспечением компьютера. Её предназначения можно разделить на три основные составляющие:

1. удобство: операционная система делает исполнение компьютера простым и удобным;

2. эффективность: операционная система позволяет эффективно использовать ресурсы компьютерной системы;

3. возможность развития: операционная система должна допускать разработку тестирования новых приложений и системных функций без нарушения нормального функционирования вычислительной системы.

Поколения ОС также как и аппаратные средства отражают достижения в области электронных компонентов:

1. поколение – лампы,
2. поколение – транзисторы,
3. поколение – интегральные микросхемы (ИС),
4. поколение большие и сверхбольшие интегральные схемы – БИС и СБИС.

Поколения компьютеров отличаются:

* стоимостью, габаритами, мощностью;
* быстродействием и объемом памяти.

## Эволюция операционных систем.

**0 поколение (40 - 50 годы)**- ОС отсутствует. Полный доступ к ресурсам ЭВМ на машинном языке, все программы разрабатываются в двоичном коде. Этот период характеризуется высокой стоимостью приобретения и эксплуатации компьютеров и низкой стоимостью труда программистов. Компьютеры использовались в монопольном интерактивном режиме. Основная цель – максимизировать использование аппаратного обеспечения. Основной режим работы компьютера – простой и ожидание каких-либо действий программиста. При этом наблюдается недостаточное использование дорогостоящего вычислительного оборудования.

**1 поколение (60-е годы).**Возникла*система пакетной обработки***.**В каждый момент времени ЭВМ использовалась для решения одной прикладной программы. Пользователи уже не работают с ЭВМ на физическом уровне. Программы и данные принимает человек-оператор. Задание на обработку формируется в виде пакета, представляющего собой совокупность отдельных программ и данных, разделенных специальными символами. Пакет может быть составлен на перфокартах или МЛ. В машине постоянно находится управляющая программа, которая считывает последовательно пользовательские программы и данные из пакета и загружает их на выполнение. Результаты выводятся на внешнее устройство, и составляют пакет результатов. Управляющая программа должна фиксировать время исполнения каждой программы, реагировать на определенные ситуации по управлению программами по мере их исполнения (прекращать выполнение). Другими словами, она должна выполнять внутрисистемные операции управления, которые пользователь осуществлял ранее на физическом уровне. Дополнительно эта программа должна автоматически переключать машину на исполнение программ из пакета по последовательной схеме (*FIFO*). Такую системную программу можно назвать простейшей ОС, обеспечивающей обработку программ в *однопрограммном пакетном режиме*.

***Недостатки:***

1. Использование части машинного времени (времени процессора) на выполнение системной управляющей программы.
2. Программа, получившая доступ к процессору, обслуживается до ее завершения. При этом если возникает потребность в передаче данных между внешними устройствами и ОП, то процессор простаивает, ожидая завершения операции обмена. С другой стороны при работе процессора простаивают внешние устройства. Для персонального компьютера проявление фактора простоя процессора не столь существенно, так как стоимость его не велика, чего не скажешь о больших и дорогих ЭВМ.

Этот недостаток однопрограммного режима был устранен *в мультипрограммном пакетном режиме*(65-75 г.г.) в ОС второго поколения.

**ОС второго поколения**. Главной целью и критерием эффективности систем пакетной обработки является максимальная пропускная способность, то есть решение максимального числа задач в единицу времени. Для этой цели в начале работы формируется пакет заданий, содержащих требования к системным ресурсам. Из этого пакета формируется мультипрограммная смесь из заданий, предъявляющих отличающиеся требования к ресурсам для того, чтобы обеспечить сбалансированную нагрузку всех устройств ЭВМ. Так, например, в мультипрограммную смесь желательно вводить вычислительные задачи и задачи, интенсивно использующие ввод/вывод. В системах пакетной обработки переключение процессов с выполнения одной задачи на другую происходит только в том случае, если активная задача сама отказывается от процессора, например, для выполнения операции ввода/вывода, что делает невозможным выполнение интерактивных задач.

Особенностью этого поколения было и то, что ОС создавались как ***системы коллективного пользования с мультипрограммным режимом работы*** и как первые *системы мультипроцессорного типа*. В таких системах несколько пользовательских программ одновременно присутствуют в оперативной памяти компьютера и центральный процессор (ЦП) динамически переключается с решения одной задачи на решение другой задачи. При этом появляется возможность одновременной загрузки ЦП и устройств ввода-вывода. В случае многопроцессорных систем единый вычислительный комплекс (ВК) содержит несколько процессоров, что повышает вычислительную мощность вычислительной системы (ВС).

В то же время появились методы, обеспечивающие ***независимость программирования от внешних устройств***. В программе необходимо указывать только тип устройства, а не программировать его физический адрес. Конкретный адрес устройства определялся в процессе выполнения программы. Это было серьезное достижение в развитии ОС.

В этот период были разработаны ***системы с разделением времени (РВ),*** предоставляющие пользователям возможность непосредственно взаимодействовать с компьютером при помощи терминалов телетайпного типа (электронная пишущая машинка, имеющая интерфейс с ЭВМ), а в последующем и с помощью дисплея. При работе с такими ОС используется диалоговый или интерактивный режим. Каждому пользователю системы разделения времени предоставляется терминал, с которого он может вести диалог со своей программой. Пользователь вводит запрос, который обрабатывается и ответ выводится на терминал, что позволяет увеличить эффективность и удобство работы пользователя. Пакетный режимобеспечивал увеличение пропускной способности и максимальную загрузку процессора*,*отлучая пользователя от ЭВМ, в то время как в режиме разделения времени (РРВ) каждый пользователь имеет непосредственный доступ к ЭВМ через свой терминал. Суть разделения времени максимально проста. Каждой программе, готовой к исполнению, для работы выделяется фиксированный, заранее определенный интервал времени, называемый квантом. Программа в течение одного кванта может быть не выполнена до конца, тогда она прерывается в момент окончания кванта и помещается в конец очереди. Из начала очереди извлекается другая программа, которой планируется фиксированный интервал. При этом ни один из пользователей, работающих за дисплеем, параллельно друг с другом, никак не ощущает, что процессор *мультиплексируется* несколькими программами.

К этому же периоду относится появление первых ***систем реального времени* (*СРВ),***в которых ЭВМ применяется для управления техническими объектами, такими, например, как станок, спутник, научная экспериментальная установка или технологическими процессами, такими, как гальваническая линия, доменный процесс и т.п. Во всех этих случаях существует предельно допустимое время, в течение которого должна быть выполнена та или иная программа, управляющая объектом, в противном случае может произойти авария: спутник сойдет с орбиты, экспериментальные данные могут быть потеряны, толщина гальванического покрытия не будет соответствовать норме. Характерным для СРВ является обеспечение заранее заданных интервалов времени реакции на предусмотренныесобытия для получения управляющего воздействия. Поскольку в технологических процессах промедление может привести к не желаемым и даже опасным последствиям, СРВ работают со значительной недогрузкой, так как важнейшей характеристикой является постоянная готовность системы – ее реактивность.

Системное программное обеспечение (СПО) ОС этого периода решало множество проблем, связанных с защитой данных и результатов работы различных программ, защитой данных в оперативной памяти и распределением устройств. Кроме того, ОС должна управлять новыми устройствами, входящими в состав аппаратного обеспечения. Для решения этих задач системное программное обеспечение сформировалось в сложную систему, требующую для реализации своих возможностей значительных вычислительных ресурсов.

**ОС третьего поколения**(70-80 г.г.) были многорежимными системами, обеспечивающими пакетную обработку, разделение времени, режим реального времени и мультипроцессорный режим. Они были громоздкими, дорогостоящими (монстры операционных систем). Например, фирме IBM разработка ОС/360 стоила 6 млрд. долларов, что соизмеримо с затратами американской программы NASA высадки человека на Луне. Такие ОС, будучи прослойкой, между пользователем и аппаратурой ЭВМ, привели к значительному усложнению вычислительной обстановки. Для выполнения простейшей программы необходимо было изучать сложные языки управления заданием (*JCL – Job Control Language*). К этому периоду относится появление ***вытесняющей многозадачности (Preemptive scheduling)***, и использование концепции баз данных для хранения больших объемов информации для организации распределенной обработки. Программисты перестали использовать перфокарты и магнитные ленты для хранения своих данных. Вводится ***приоритетное планирование (Prioritized scheduling)*** и выделение квот на использование ограниченных ресурсов компьютеров (процессорного времени, дисковой памяти, физической (оперативной) памяти). При использовании компьютеров широкое распространение получила концепция распределения времени (*time sharing*) , но ограниченность ресурсов приводила к перегрузке компьютеров и к неприемлемому времени ожидания ответа или результатов работы. Программистам приходилось компенсировать это неудобство работой в ночное время.

**Четвертое поколение**с середины 70-х. Период характеризуется уменьшением стоимости компьютеров и увеличением стоимости труда программиста. Появление персональных компьютеров позволило установить компьютер практически каждому пользователю на рабочем столе. Благодаря широкому распространению вычислительных сетей и средств оперативной обработки (режим *on-line),* пользователи получают доступ к территориально распределенным компьютерам. Появились микропроцессоры, на основе которых создаются все новые и новые PC, которые могут быть использованы, как автономно, так и в качестве терминалов более мощных вычислительных систем. При передаче информации по линиям связи усложняются проблемы защиты информации, шифрования данных. Возникло ***понятие сетевого компьютера (Network computer)***, способного получать все ресурсы через компьютерную сеть. Понятие файловой системы распространяется на данные, доступные по различным сетевым протоколам.

Число людей, пользующихся компьютером, значительно возросло, что выдвигает требование дружественного интерфейса пользователя, ориентации на неподготовленного пользователя. Появились системы с управлением с помощью меню и элементов графического интерфейса. Начала широко распространяться ***концепция виртуальных машин***. Пользователь более не заботится о физическихдеталях построения ЭВМ или сетей. Он имеет дело с функциональным эквивалентом компьютера, создаваемым для него ОС, представляющимвиртуальную машину. Таким образом, возникла концепция виртуализации ресурсов ЭВМ. *Виртуальным ресурсом* называется функциональный, программно моделируемый эквивалент реального монопольного ресурса, допускающий его совместное использование многими процессами. ***Мультипрограммирование – виртуализация центрального процессора (ЦП – CPU)***. Буферный ввод/вывод – виртуализация устройств ввода и вывода.

В настоящее время концепция виртуальных машин находит все большее распространение. Виртуальная машина – это функциональный эквивалент реальной ЭВМ, обеспечивающий пользователей на основе одной ЭВМ множество функционально эквивалентных ей ЭВМ.

Широкое внедрение получила концепция *распределенной обработки данных*. Развитием распределенной обработки данных стала ***технология Клиент - Сервер***, в которой серверный процесс предоставляет возможность использовать свои ресурсы клиентскому процессу по соответствующему протоколу взаимодействия. Название сервера отображает вид ресурса, который предоставляется клиентским системам (сервер печати, сервер вычислений, сервер баз данных, сервер новостей, сервер FTP, сервер WWW и т.д.)

## Функции операционной системы.

### Операционная система как расширенная машина и менеджер ресурсов.

Функции современной ОС поддерживают: мультипрограммный режим, работа с виртуальной памятью, многооконный интерфейс.

2 основных задачи ОС – расширение возможностей машины и управление её ресурсами

### ОС как расширенная машина

Использование большинства компьютеров на уровне машинного языка затруднительно, особенно это касается ввода-вывода. Например, для организации чтения блока данных с гибкого диска программист может использовать 16 различных команд, каждая из которых требует 13 параметров, таких как номер блока на диске, номер сектора на дорожке и т. п. Когда выполнение операции с диском завершается, контроллер возвращает 23 значения, отражающих наличие и типы ошибок, которые, очевидно, надо анализировать. Даже если не входить в курс реальных проблем программирования ввода-вывода, ясно, что среди программистов нашлось бы не много желающих непосредственно заниматься программированием этих операций. При работе с диском программисту-пользователю достаточно представлять его в виде некоторого набора файлов, каждый из которых имеет имя. Работа с файлом заключается в его открытии, выполнении чтения или записи, а затем в закрытии файла. Вопросы подобные таким, как следует ли при записи использовать усовершенствованную частотную модуляцию или в каком состоянии сейчас находится двигатель механизма перемещения считывающих головок, не должны волновать пользователя. Программа, которая скрывает от программиста все реалии аппаратуры и предоставляет возможность простого, удобного просмотра указанных файлов, чтения или записи - это, конечно, операционная система. Точно также, как ОС ограждает программистов от аппаратуры дискового накопителя и предоставляет ему простой файловый интерфейс, операционная система берет на себя все малоприятные дела, связанные с обработкой прерываний, управлением таймерами и оперативной памятью, а также другие низкоуровневые проблемы. В каждом случае та абстрактная, воображаемая машина, с которой, благодаря операционной системе, теперь может иметь дело пользователь, гораздо проще и удобнее в обращении, чем реальная аппаратура, лежащая в основе этой абстрактной машины.

С этой точки зрения функцией ОС является предоставление пользователю некоторой расширенной или виртуальной машины, которую легче программировать и с которой легче работать, чем непосредственно с аппаратурой, составляющей реальную машину.

### ОС как система управления ресурсами

Идея о том, что ***ОС*** прежде всего ***система, обеспечивающая удобный интерфейс*** пользователям, соответствует рассмотрению сверху вниз. **Другой взгляд**, снизу вверх, дает представление об ***ОС*** как о некотором ***механизме, управляющем всеми частями сложной системы.*** Современные вычислительные системы состоят из процессоров, памяти, таймеров, дисков, накопителей на магнитных лентах, сетевых коммуникационной аппаратуры, принтеров и других устройств. В соответствии со вторым подходом функцией ОС является распределение процессоров, памяти, устройств и данных между процессами, конкурирующими за эти ресурсы. ОС должна управлять всеми ресурсами вычислительной машины таким образом, чтобы обеспечить максимальную эффективность ее функционирования. Критерием эффективности может быть, например, пропускная способность или реактивность системы. ***Управление ресурсами включает решение двух общих, не зависящих от типа ресурса задач:***

* планирование ресурса - то есть определение, кому, когда, а для делимых ресурсов и в каком количестве, необходимо выделить данный ресурс;
* отслеживание состояния ресурса - то есть поддержание оперативной информации о том, занят или не занят ресурс, а для делимых ресурсов - какое количество ресурса уже распределено, а какое свободно.

Для решения этих общих задач управления ресурсами разные ОС используют различные алгоритмы, что в конечном счете и определяет их облик в целом, включая характеристики производительности, область применения и даже пользовательский интерфейс. Так, например, алгоритм управления процессором в значительной степени определяет, является ли ОС системой разделения времени, системой пакетной обработки или системой реального времени.

### Классификация операционных систем

Операционные системы различаются

* особенностями реализации внутренних алгоритмов управления основными ресурсами компьютера (процессорами, памятью, устройствами),
* особенностями использованных методов проектирования,
* типами аппаратных платформ,
* критериями эффективности,
* особенностями реализации сетевых решений
* и многими другими свойствами.

**Особенности алгоритмов управления ресурсами**

От эффективности алгоритмов управления ресурсами компьютера во многом зависит эффективность всей ОС в целом. Поэтому, характеризуя ОС, часто приводят важнейшие особенности реализации функций ОС по управлению процессорами, памятью, внешними устройствами

По режиму обработки задач различают ОС, обеспечивающие однопрограммный или мультипрограммный режим работы.

*Однопрограммные ОС* в основном выполняют функцию предоставления пользователю виртуальной машины, делая более простым и удобным процесс взаимодействия пользователя с компьютером.

Они включают средства управления периферийными устройствами, средства управления файлами, средства общения с пользователем.

*Мультипрограммные ОС,* кроме функций однопрограммных ОС, управляют разделением совместно используемых ресурсов, таких как процессор, оперативная память, файлы и внешние устройства.

По числу одновременно работающих пользователей ОС делятся на:

* однопользовательские (MS-DOS, Windows 3.x, ранние версии OS/2);
* многопользовательские (UNIX, Windows NT).

Главным отличием многопользовательских систем от однопользовательских является наличие средств защиты информации каждого пользователя от несанкционированного доступа других пользователей.

*!!! Следует заметить, что не всякая многозадачная система является многопользовательской, и не всякая однопользовательская ОС является однозадачной.*

Важнейшим разделяемым ресурсом является процессорное время. Способ распределения процессорного времени между несколькими одновременно существующими в системе процессами (или нитями) во многом определяет специфику ОС. Среди множества существующих вариантов реализации многозадачности можно выделить две группы алгоритмов:

* невытесняющая многозадачность (NetWare, Windows 3.x);
* вытесняющая многозадачность (Windows NT, OS/2, UNIX).

При***невытесняющей многозадачности*** активный процесс выполняется до тех пор, пока он сам, по собственной инициативе, не отдаст управление операционной системе для того, чтобы та выбрала из очереди другой готовый к выполнению процесс.

При ***вытесняющей многозадачности*** решение о переключении процессора с одного процесса на другой принимается операционной системой, а не самим активным процессом.

Важным свойством операционных систем является возможность распараллеливания вычислений в рамках одной задачи.

ОС подразделяются на поддерживающие и неподдерживающие многонитевость.

***Многонитевая ОС***разделяет процессорное время не между задачами, а между их отдельными ветвями (нитями).

Еще однимважным свойством ОС является отсутствие или наличие в ней средств поддержки многопроцессорной обработки - *мультипроцессирование*.

***Мультипроцессорная обработка*** – это способ организации вычислительного процесса в системах с несколькими процессорами, при котором несколько задач (процессов, потоков) могут одновременно выполняться на разных процессорах системы.

Мультипроцессирование приводит к усложнению всех алгоритмов управления ресурсами.

В современных ОС введены функций поддержки многопроцессорной обработки данных. Такие функции имеются в операционных системах Solaris 2.x фирмы Sun, Open Server 3.x компании Santa Crus Operations, OS/2 фирмы IBM, Windows NT фирмы Microsoft и NetWare 4.1 фирмы Novell.

Многопроцессорные ОС могут классифицироваться по способу организации вычислительного процесса в системе с многопроцессорной архитектурой: ***асимметричные ОС и симметричные ОС***.

Асимметричная ОС целиком выполняется только на одном из процессоров системы, распределяя прикладные задачи по остальным процессорам.

Симметричная ОС полностью децентрализована и использует все процессоры, разделяя их между системными и прикладными задачами.

**Особенности аппаратных платформ**

На свойства операционной системы непосредственное влияние оказывают аппаратные средства, на которые она ориентирована.

По типу аппаратуры различают операционные системы

* + персональных компьютеров,
  + мини-компьютеров,
  + мейнфреймов,
  + кластеров
  + и сетей ЭВМ.

Наряду с ОС, ориентированными на совершенно определенный тип аппаратной платформы, существуют операционные системы, специально разработанные таким образом, чтобы они могли быть легко перенесены с компьютера одного типа на компьютер другого типа, так называемые *мобильные* ОС (аппаратно-независимые). Наиболее ярким примером такой ОС является популярная система UNIX.

**В зависимости от выбранного критерия эффективности операционные системы делятся на:**

* системы пакетной обработки (например, OC EC),
* системы разделения времени (UNIX, VMS),
* системы реального времени (QNX, RT/11).

*Системы пакетной обработки* служат для решения задач в основном вычислительного характера, не требующих быстрого получения результатов.

Главной целью и критерием эффективности таких систем является максимальная пропускная способность, то есть решение максимального числа задач в единицу времени.

Схема функционирования систем пакетной обработки данных:

* в начале работы формируется пакет заданий, каждое задание содержит требование к системным ресурсам;
* из этого пакета заданий формируется мультипрограммная смесь, то есть множество одновременно выполняемых задач. Для одновременного выполнения выбираются задачи, предъявляющие отличающиеся требования к ресурсам, так, чтобы обеспечивалась сбалансированная загрузка всех устройств вычислительной машины; (в мультипрограммной смеси желательно одновременное присутствие вычислительных задач и задач с интенсивным вводом-выводом).

Таким образом, выбор нового задания из пакета заданий зависит от внутренней ситуации, складывающейся в системе, то есть выбирается "выгодное" задание.

* В таких ОС невозможно гарантировать выполнение того или иного задания в течение определенного периода времени.
* В системах пакетной обработки переключение процессора с выполнения одной задачи на выполнение другой происходит только в случае, если активная задача сама отказывается от процессора, например, из-за необходимости выполнить операцию ввода-вывода. Поэтому одна задача может надолго занять процессор, что делает невозможным выполнение интерактивных задач.

Взаимодействие пользователя с вычислительной машиной, на которой установлена система пакетной обработки, сводится к тому, что он приносит задание, отдает его диспетчеру-оператору, а в конце дня после выполнения всего пакета заданий получает результат. Очевидно, что такой порядок снижает эффективность работы пользователя.

Основной недостаток систем пакетной обработки - изоляция пользователя-программиста от процесса выполнения его задач.

*Системы разделения времени* призваны исправить недостаток систем пакетной обработки данных.

Каждому пользователю системы разделения времени предоставляется терминал, с которого он может вести диалог со своей программой.

Каждой задаче выделяется только квант процессорного времени, ни одна задача не занимает процессор надолго, и время ответа оказывается приемлемым.

Если квант выбран достаточно небольшим, то у всех пользователей, одновременно работающих на одной и той же машине, складывается впечатление, что каждый из них единолично использует машину.

Системы разделения времени обладают меньшей пропускной способностью, чем системы пакетной обработки, так как

1. на выполнение принимается каждая запущенная пользователем задача, а не та, которая "выгодна" системе,
2. увеличивается время работы, так как выполняется более частое переключение процессора с задачи на задачу.

*Таким образом, критерием эффективности систем разделения времени является не максимальная пропускная способность, а удобство и эффективность работы пользователя.*

*Системы реального времени* применяются для управления различными техническими объектами, такими, например, как станок, спутник, научная экспериментальная установка или технологическими процессами, такими, как гальваническая линия, доменный процесс и т.п.

Существует предельно допустимое время, в течение которого должна быть выполнена та или иная программа, управляющая объектом, в противном случае может произойти авария: спутник выйдет из зоны видимости, экспериментальные данные, поступающие с датчиков, будут потеряны, толщина гальванического покрытия не будет соответствовать норме.

Таким образом, *критерием эффективности для систем реального времени является их способность выдерживать заранее заданные интервалы времени между запуском программы и получением результата (управляющего воздействия).*

Это время называется временем реакции системы, а соответствующее свойство системы - реактивностью.

Для этих систем мультипрограммная смесь представляет собой фиксированный набор заранее разработанных программ, а выбор программы на выполнение осуществляется исходя из текущего состояния объекта или в соответствии с расписанием плановых работ.

Замечание. Некоторые операционные системы могут совмещать в себе свойства систем разных типов, например, часть задач может выполняться в режиме пакетной обработки, а часть - в режиме реального времени или в режиме разделения времени. В таких случаях режим пакетной обработки часто называют фоновым режимом.