

Лекция 1

Базовые понятия операционных систем

“Операционные системы”-2019, ИВК, УлГТУ.

Назначение и функции операционных систем

Программное обеспечение:

- системное;
- прикладное.

Системное ПО – такие программы и комплексы программ, без которых невозможно выполнение и создание других программ.

Создание системного ПО началось с осознания, что основа ЭВМ - операции ввода-вывода.

Назначение и функции операционных систем

Ввод-вывод данных может быть различным.

Пример ввода: чтение из файла, с клавиатуры, со звукового устройства и т.д.

Пример вывода: запись в файл, вывод на экран, на принтер, передача по сети и т.д.

Написать программу для работы с каждым видом ввода-вывода
трудозатратно, и просить программистов каждый раз это делать - не
эффективно.

Поэтому были разработаны специальные подпрограммы, в которых эта
функциональность была заложено. Так и появилось системное ПО.

Назначение и функции операционных систем

В 50е гг. взаимодействие пользователя с вычислительными комплексами было совсем иным.

Разделяли программистов-кодиров (отлично знали архитектуру ЭВМ и языки программирования) и программистов-алгоритмистов (отлично знали алгоритмы).

Программа при помощи операторов вводилась на специальные устройства, чаще всего на перфокарты и перфоленты.

Далее оператор ВК (ВЦ) выполнял следующие действия:

Назначение и функции операционных систем

- 1) Оператор ВК вводил (в регистры центрального процессора и оперативную память) программу для трансляции исходных кодов в машинные (двоичные) коды.
- 2) Далее транслятор считывал исходную программу и выводил промежуточные результаты трансляции опять же на перфокарты (перфоленты), т.к. трансляция - очень затратный процесс, и оперативной памяти не хватало. Процесс мог проводиться несколько раз.
- 3) Оператор загружал в ОП полученные двоичные коды оттранслируемой программы и подгружал коды для управления вводом-выводом. Только после этого исходная программа могла сама считывать данные и выполнять вычисления.

Назначение и функции операционных систем

Для автоматизации программиста-кодера стали разрабатывать специальные алгоритмические языки высокого уровня.

Для автоматизации труда оператора ВК была разработана специальная управляющая программа, загрузив которую один раз в оперативную память, оператор мог использовать её многократно и не обращаться к программированию ЭВМ с пульта оператора.

Именно эту программу и называли **операционной системой**.

Назначение и функции операционных систем

Затем на ОС стали назначать всё больше функций (не только загрузку трансляторов).

В 60е гг. ОС позволяли организовывать параллельное выполнение нескольких программ.

В 70е гг. у ОС начали появляться интерфейсные возможности.

Интерфейс - набор спецификаций, определяющих конкретный способ взаимодействия человека и компьютера.

В 70е гг. начал преобладать диалоговый режим работы с компьютерами.

Назначение и функции операционных систем

На сегодняшний день операционная система – комплекс системных управляющих и обрабатывающих программ, которые:

- выступают как интерфейс между прикладным ПО и аппаратным обеспечением компьютера;
- предназначены для наиболее оптимального расходования ресурсов вычислительной системы;
- предназначены для организации надёжных вычислений.

Назначение и функции операционных систем

Основные функции ОС [1]:

- Приём от пользователя заданий (команд), сформированных на специальном языке, и их обработка. Пример: запуск программы или работа с файлами.
- Загрузка в ОП подлежащих исполнению программ;
- Распределение памяти и, чаще всего, организация виртуальной памяти;
- Запуск программы (передача управления программе);
- Приём и исполнение различных запросов от выполняющихся приложений при помощи интерфейса прикладного программирования (API);
- Обслуживание всех операций ввода-вывода;

Назначение и функции операционных систем

Основные функции ОС [2]:

- Обеспечение работы систем управления файлами и/или систем управления базами данных (СУБД);
- Обеспечение режима мультипрограммирования (организация параллельного выполнения нескольких программ);
- Планирование и диспетчеризация задач;
- Организация обмена сообщениями между выполняющимися программами;
- Обеспечение работы систем программирования;
- Удовлетворение ограничениям по времени ответа в режиме реального времени (для ОСРВ, не применимо для настольных ОС).

Основные понятия операционных систем

Прикладные программы, созданные пользователями, выполняются в **операционной среде**. Это понятие отлично от понятия “операционная система”, т.к. одна операционная система можно реализовывать несколько операционных сред.

Например, программу, написанную для ОС MS DOS можно запустить в ОС Microsoft Windows XP, т.к. в Windows XP существует реализация 16-битной операционной среды от MS DOS: NTVDM. В Windows 10 она тоже есть, но по умолчанию выключена.

Основные понятия операционных систем

Прикладные программисты могут и не знать деталей управления конкретными ресурсами, т.к. обращаются к некоторой программной подсистеме с конкретными **системными вызовами**.

Эта программная подсистема - и есть ОС, а набор функций и сервисов, к которым можно обратиться - это программная среда.

Набор системных вызовов и правил их использования и образуют интерфейс прикладного программирования (API).

Программная среда, образуемая непосредственно кодом ОС - **native** (основная), остальные программные (или системные) среды могут эмулироваться. Пример с MS DOS на Windows XP подходит.

Основные понятия операционных систем

ОС предоставляет вида интерфейсов: пользовательский и программный.

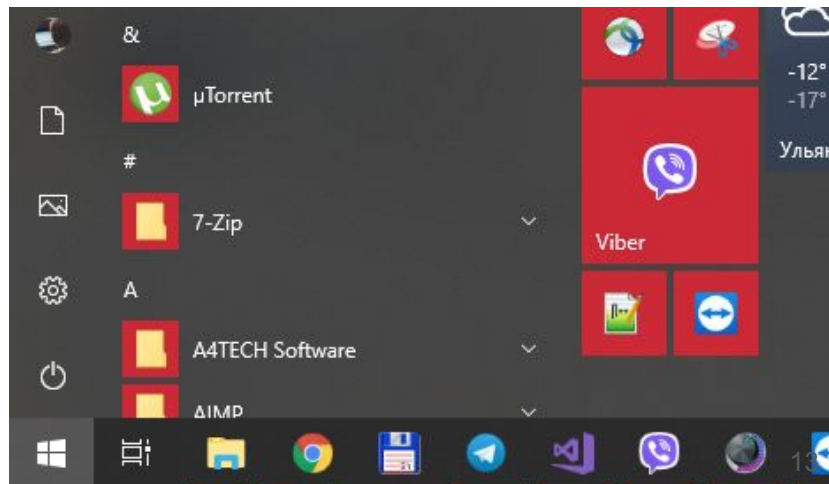
Пример пользовательских интерфейсов: командная строка или графический интерфейс. В среде Linux распространены Bash, X Window.

```
Windows PowerShell
(C) Корпорация Майкрософт (Microsoft Corporation). Все права защищены.

PS C:\Users\Владислав> pwd

Path
----
C:\Users\Владислав

PS C:\Users\Владислав> 
```



Основные понятия операционных систем

Много устройств в ЭВМ работают параллельно. Как реагировать на события от этих устройств? Например, при нажатии на клавишу клавиатуры или движении мышью.

Или, например, необходимо отреагировать на особое состояние процессора.

Для этого был изобретён механизм “прерываний”.

Прерывание — это принудительная передача управления от выполняемой программы к системе (а через неё — к соответствующей программе обработки прерывания), происходящая при возникновении определённого события.

Основные понятия операционных систем

Механизм прерываний реализуется аппаратно-программными средствами. Структуры систем прерывания могут быть самыми разными, но все они имеют одну общую особенность — прерывание непременно влечёт за собой изменение порядка выполнения команд процессором.

Механизм обработки прерываний состоит из следующих шагов [1]:

- 1) Установление факта прерывания и идентификация прерывания.
- 2) Запоминание состояния прерванного процесса вычислений (контекста процесса).
- 3) Управление аппаратно передаётся на подпрограмму обработки прерывания.

Основные понятия операционных систем

Механизм обработки прерываний состоит из следующих шагов [2]:

- 4) Программное сохранение информации о прерванной программе, которую не удалось “спасти” на этапе 2.
- 5) Собственно выполнение программы прерывания.
- 6) Восстановление информации о прерванной программе.
- 7) Возврат на прерванную программу.

Шаги 1-3 реализуются аппаратно, 4-7 - программно.

На самом деле алгоритм обработки прерываний почти всегда намного сложнее, но указанные выше шаги в нём всегда присутствуют.

Основные понятия операционных систем

Прерывания, возникающие при работе вычислительной системы, можно разделить на два основных класса: внешние (их иногда называют асинхронными) и внутренние (синхронные).

Пример внешних прерываний:

- от таймера;
- от внешних устройств (по вводу-выводу);
- по нарушению питания;
- с пульта оператора вычислительной системы.

Основные понятия операционных систем

Пример внутренних прерываний:

- при нарушении адресации;
- при неизвестном коде операции;
- при делении на 0;
- при переполнении;
- от средств контроля (нарушение контроля чётности, возникновение ошибок при работе устройств).

Также особым видом прерываний является программное прерывание, которое вызывается по соответствующей команде прерывания.

Основные понятия операционных систем

У прерываний есть приоритет: (от наивысшего к низшему):

- средства контроля процессора;
- системный таймер;
- магнитные диски;
- сетевое оборудование / внешние устройства;
- терминалы;
- программные прерывания.

Основные понятия операционных систем

Не всегда прерывание должно быть выполнено немедленно.

Процессоры обладают функциями отключения системы прерываний и маскирования (запрета) отдельных сигналов прерывания.

Дисциплины обслуживания прерываний:

- с абсолютными приоритетами;
- с относительными приоритетами;
- по дисциплине LCFS (Last Come First Served — последним пришел, первым обслужен) или по принципу стека.

Основные понятия операционных систем

Управление ходом выполнения задач со стороны операционной системы заключается в организации реакций на прерывания, в организации обмена информацией (данными и программами), в предоставлении необходимых ресурсов, в динамике выполнения задачи и в организации сервиса.

Причины прерываний определяет операционная система (модуль, который называют супервизором прерываний), она же и выполняет действия, необходимые при данном прерывании и в данной ситуации.

Т.к. часто при обработке прерываний необходимо отключать прерывания или их маскировать, а затем обратно включать, то эти секции подпрограмм обработки прерываний и были выделены в супервизор прерываний.

Основные понятия операционных систем

Последовательный вычислительный процесс (или просто "процесс, также иногда называемый "задачей") — это отдельная программа с её данными, выполняющаяся на последовательном процессоре.

Последовательный процессор – процессор, где следующая команда выполняется после предыдущей.

Концепция процесса предполагает два аспекта:

- 1) он является носителем данных;
- 2) он собственно и выполняет операции, связанные с обработкой этих данных.

Основные понятия операционных систем

Концепция процесса преследует цель выработать механизмы распределения и управления ресурсами.

Термин ресурс обычно применяется по отношению к многократно используемым, относительно стабильным и часто недостающим объектам, которые запрашиваются, задействуются и освобождаются в период их активности.

Другими словами, ресурсом называется всякий объект, который может распределяться внутри системы.

Основные понятия операционных систем



Основные понятия операционных систем

Операционная система может поддерживать мультипрограммирование (многопроцессность).

В этом случае она должна стараться эффективно использовать имеющиеся ресурсы путём организации к ним очередей запросов, составляемых тем или иным способом.

Это требование достигается поддержанием в памяти более одного вычислительного процесса, ожидающего процессор, и более одного процесса, готового использовать другие ресурсы, как только последние станут доступными.

Основные понятия операционных систем

Общая схема выделения ресурсов такова. При необходимости использовать какой-либо ресурс (оперативную память, устройство ввода-вывода, массив данных и т. п.) вычислительный процесс путём обращения к супервизору (supervisor) операционной системы посредством специальных вызовов сообщает о своём требовании. При этом указывается вид ресурса и, если надо, его объем.

Например, при запросе оперативной памяти указывается количество адресуемых ячеек, необходимое для дальнейшей работы.

Основные понятия операционных систем

Команда обращения к операционной системе передает ей управление, переводя процессор в привилегированный режим работы, если такой существует.

Большинство компьютеров имеют два (и более) режима работы: привилегированный (режим супервизора) и пользовательский.

В привилегированном режиме могут выполняться вызовы, не допустимые в пользовательском режиме. Например, в пользовательском режиме нельзя выключить систему прерываний.

Основные понятия операционных систем

Ресурс может быть выделен вычислительному процессу (задаче), обратившемуся к операционной системе с соответствующим запросом, если:

- ресурс свободен и в системе нет запросов от задач более высокого приоритета к этому же ресурсу;
- текущий запрос и ранее выданные запросы допускают совместное использование ресурсов;
- ресурс используется задачей низшего приоритета и может быть временно отобран (разделяемый ресурс).

Основные понятия операционных систем

После окончания работы с ресурсом задача опять с помощью специального вызова супервизора сообщает операционной системе об отказе от ресурса, либо ОС забирает ресурс сама.

Супервизор операционной системы освобождает ресурс и проверяет, имеется ли очередь к освободившемуся ресурсу. Если очередь есть, то в зависимости от принятой дисциплины обслуживания и приоритетов заявок он выводит из состояния ожидания задачу, ждущую ресурс, и переводит её в состояние готовности к выполнению, после чего либо передаёт управление ей.

Диаграмма состояний процесса

Необходимо отличать системные управляющие вычислительные процессы, представляющие работу супервизора операционной системы и занимающиеся распределением и управлением ресурсами, от всех других процессов: задач пользователей и системных обрабатывающих процессов.

Последние, хоть и относятся к операционной системе, но не входят в ядро операционной системы и требуют общих ресурсов для своей работы, которые получают от супервизора.

В задаче распределения ресурсов между процессами исполнение системных управляющих программ обычно не рассматривается.

Диаграмма состояний процесса

Если некий вычислительный процесс (назовём его первым) в данный конкретный момент времени не исполняется, поскольку процессор занят исполнением какого-то другого процесса, то операционная система должна знать, что вычисления в первом процессе приостановлены.

Информация об этом заносится в специальную информационную структуру, сопровождающую каждый вычислительный процесс. Она называется дескриптором процесса.

Диаграмма состояний процесса

Процесс может находиться в двух основных состояниях: активном и пассивном (применим только для ОСРВ).

В активном состоянии процесс может конкурировать за ресурсы вычислительной системы.

В пассивном состоянии процесс не конкурирует за ресурсы, но операционной системе он известен.

Это необходимо для уменьшения накладных расходов при инициировании процесса и для обеспечения предсказуемости работы вычислительной системы.

Диаграмма состояний процесса

Активный процесс может быть в одном из следующих состояний:

- выполнения - все затребованные процессом ресурсы выделены (в этом состоянии в каждый момент времени может находиться только один процесс, если речь идёт об однопроцессорной вычислительной системе);
- готовности к выполнению - ресурсы могут быть предоставлены, и тогда процесс перейдёт в состояние выполнения;
- ожидания или блокирования - затребованные ресурсы не могут быть предоставлены, или не завершена операция ввода-вывода.

Диаграмма состояний процесса



Классификация операционных систем

Основным предназначением ОС является организация эффективных и надёжных вычислений, создание различных интерфейсов для взаимодействия с этими вычислениями и с самой вычислительной системой.

Выделяют классификации ОС:

- по назначению;
- по режиму обработки задач;
- по способу взаимодействия с системой;
- по способам построения (архитектурным особенностям системы).

Классификация операционных систем

Прежде всего, различают ОС общего и специального назначения.

По режиму обработки задач различают ОС, обеспечивающие однопрограммный и мультипрограммный (мультизадачный) режимы. К однопрограммным ОС относится, например, всем известная, хотя ныне уже практически и не используемая MS DOS.

По способу взаимодействия с компьютером, то можно говорить о диалоговых системах и системах пакетной обработки.

Классификация операционных систем

При организации работы с вычислительной системой в диалоговом режиме можно говорить об однопользовательских (однотерминальных) и мультитерминальных ОС. В мультитерминальных ОС с одной вычислительной системой одновременно могут работать несколько пользователей, каждый со своего терминала.

Основной особенностью операционных систем реального времени (ОСРВ) является обеспечение обработки поступающих заданий в течение заданных интервалов времени, которые нельзя превышать.

В ОС, не предназначенных для решения задач реального времени, имеются некоторые накладные расходы процессорного времени на этапе инициирования задач.

Архитектурные особенности операционных систем

По основному архитектурному принципу операционные системы разделяются на микроядерные и макроядерные (монолитные).

Монолитное ядро ОС

- Представляет богатый набор оборудования. Все компоненты монолитного ядра находятся в одном адресном пространстве.
- Эта схема ОС, когда все части ее ядра - это составные части одной программы.
- Монолитное ядро - самый старый способ организации ОС.
- Ошибка работы одного из компонентов ядра нарушает работу всей системы.

Архитектурные особенности операционных систем

Микроядро ОС

- Представляет только основные функции управления процессами и минимальный набор для работы с оборудованием. Классические микроядра дают очень небольшой набор системных вызовов.
- Устойчивость к сбоям и ошибкам оборудования и компонентов системы, высокая степень ядерной модульности, что упрощает добавление в ядро новых компонентов и процесс отладки ядра. Для отладки такого ядра можно использовать обычные средства.
- Архитектура микроядра увеличивает надежность системы.
- Передача информации требует больших расходов и большого количества времени.