

УВАЖАЕМЫЕ СТУДЕНТЫ! Изучите приведенную лекцию, законспектируйте основные понятия, дайте ответы на контрольные вопросы.

Ответы на вопросы, фотоотчет, предоставить преподавателю на e-mail: r.bigangel@gmail.com **до 17.04.2023.**

При возникновении вопросов по приведенному материалу обращаться по следующему номеру телефона: (072)111-37-59, (Viber, WhatsApp), vk.com: <https://vk.com/daykini>

ВНИМАНИЕ!!! При отправке работы, не забывайте указывать ФИО студента, наименование дисциплины, дата проведения занятия (по расписанию).

Лекция

Тема Алгоритмы планирования, основанные на приоритетах, смешанные алгоритмы планирования.

1. Алгоритмы планирования, основанные на приоритетах

Другой важной концепцией, лежащей в основе многих вытесняющих алгоритмов планирования, является **приоритетное обслуживание.**

Приоритетное обслуживание предполагает наличие у потоков некоторой изначально известной характеристики — приоритета, на основании которой определяется порядок их выполнения. **Приоритет** — это число, характеризующее степень привилегированности потока при использовании ресурсов вычислительной машины, в частности процессорного времени: чем выше приоритет, тем выше привилегии, тем меньше времени будет проводить поток в очередях.

Приоритет может выражаться целым или дробным, положительным или отрицательным значением.

В большинстве операционных систем, поддерживающих потоки, приоритет потока непосредственно связан с приоритетом процесса, в рамках которого выполняется данный поток. Приоритет процесса назначается операционной системой при его создании. Значение приоритета включается в описание процесса и используется при назначении приоритета потокам этого процесса.

Во многих ОС предусматривается возможность изменения приоритетов в течение жизни потока. Изменение приоритета могут происходить по инициативе самого потока, когда он обращается с соответствующим вызовом к операционной системе, или по инициативе пользователя, когда он выполняет соответствующую команду. Кроме того, ОС сама может изменять приоритеты потоков в зависимости от ситуации, складывающейся в системе. В последнем случае приоритеты называются *динамическими* в отличие от неизменяемых, *фиксированных*, приоритетов,

От того, какие приоритеты назначены потокам, существенно зависит эффективность работы всей вычислительной системы.

Существуют две разновидности приоритетного планирования: обслуживание с относительными приоритетами и обслуживание с абсолютными приоритетами.

В обоих случаях выбор потока на выполнение из очереди готовых осуществляется одинаково: выбирается поток, имеющий наивысший приоритет. Однако проблема определения момента смены активного потока решается по-разному. В системах с относительными приоритетами активный поток выполняется до тех пор, пока он сам не покинет процессор, перейдя в состояние ожидания

В системах с абсолютными приоритетами выполнение активного потока прерывается кроме указанных выше причин, еще при одном условии: если в очереди готовых потоков появился поток, приоритет которого выше приоритета активного потока.

В системах с абсолютными приоритетами время ожидания потока в очередях может быть сведено к минимуму, если ему назначить самый высокий приоритет. Такой поток будет вытеснять из процессора все остальные потоки (кроме потоков, имеющих такой же наивысший приоритет). Это делает планирование на основе абсолютных приоритетов подходящим для систем управления объектами, в которых важна быстрая реакция на событие.

2. Смешанные алгоритмы планирования

Во многих операционных системах алгоритмы планирования построены с использованием как концепции квантования, так и приоритетов. Например, в основе планирования лежит квантование, но величина кванта и/или порядок выбора потока из очереди готовых определяется приоритетами потоков. Именно так реализовано планирование в системе Windows NT. На выполнение выбирается готовый поток с наивысшим приоритетом. Ему выделяется квант времени. Если во время выполнения в очереди готовых появляется поток с более высоким приоритетом, то он вытесняет выполняемый поток. Вытесненный поток возвращается в очередь готовых, причем он становится впереди всех остальных потоков имеющих такой же приоритет.

Каждый процесс в зависимости от задачи, которую он решает, относится к одному из трех определенных в системе приоритетных классов: классу реального времени, классу системных процессов или классу процессов разделения времени. Назначение и обработка приоритетов выполняются для разных классов по-разному. Процессы системного класса, зарезервированные для ядра, используют стратегию фиксированных приоритетов. Уровень приоритета процессу назначается ядром и никогда не изменяется.

Процессы реального времени также используют стратегию фиксированных приоритетов, но пользователь может их изменять. Характеристики планирования процессов реального времени включают две величины: уровень глобального приоритета и квант времени. Для каждого уровня приоритета по умолчанию имеется своя величина кванта времени. Процессу разрешается захватывать процессор на указанный квант времени, а по его истечении планировщик снимает процесс с выполнения.

Состав класса процессов разделения времени наиболее неопределенный и часто меняющийся в отличие от системных процессов и процессов реального времени. Для справедливого распределения времени процессора между процессами в этом классе используется стратегия динамических приоритетов. Величина приоритета, назначаемого процессам разделения времени, вычисляется пропорционально значениям двух составляющих:

пользовательской части и системной части. Пользовательская часть приоритета может быть изменена администратором и владельцем процесса, но в последнем случае только в сторону его снижения.

В операционной системе OS/2. Планирование здесь основано на использовании квантования и абсолютных динамических приоритетов. На множестве потоков определены приоритетные классы — критический (time critical), серверный (server), стандартный (regular) и остаточный (idle), в каждом из которых имеется 32 приоритетных уровня. Потоки критического класса имеют наивысший приоритет. В этот класс могут быть отнесены, например, системные потоки, выполняющие задачи управления сетью. Следующий по приоритетности класс предназначен, как это следует из его названия, для потоков, обслуживающих серверные приложения. К стандартному классу могут быть отнесены потоки обычных приложений. Потоки, входящие в остаточный класс, имеют самый низкий приоритет. К этому классу относится, например, поток, выводящий на экран заставку, когда в системе не выполняется никакой работы.

Поток из менее приоритетного класса не может быть выбран для выполнения, пока в очереди более приоритетного класса имеется хотя бы один поток. Внутри каждого класса потоки выбираются также по приоритетам. Потоки, имеющие одинаковое значение приоритета, обслуживаются в циклическом порядке.

Приоритеты могут изменяться планировщиком в следующих случаях:

1) Если поток находится в ожидании процессорного времени дольше, чем это задано системной переменной MAXWAIT, то его уровень приоритета будет автоматически увеличен операционной системой. При этом результирующее значение приоритета не должно превышать нижней границы диапазона приоритетов критического класса.

2) Если поток ушел на выполнение операции ввода-вывода, то после ее завершения он получит наивысшее значение приоритета своего класса.

3) Приоритет потока автоматически повышается, когда он поступает на выполнение.

ОС динамически устанавливает величину кванта, отводимого потоку для выполнения. Величина кванта зависит от загрузки системы и интенсивности подкачки. Параметры настройки системы позволяют явно задать границы изменения кванта. В любом случае он не может быть меньше 32 мс и больше 65 536 мс. Если поток был прерван до истечения кванта, то следующий выделенный ему интервал выполнения будет увеличен на время, равное одному периоду таймера (около 32 мс), и так до тех пор, пока квант не достигнет заранее заданного при настройке ОС предела.

Благодаря такому алгоритму планирования в OS/2 ни один поток не будет «забыт» системой и получит достаточно процессорного времени.