Урок 3. Потоки

Поток является последовательностью команд, обрабатываемых процессором. В рамках одного процесса могут находиться один или несколько потоков.
Традиционный подход, при котором каждый процесс представляет собой единый поток выполнения, называется ***однопоточным***. Например, MS-DOS поддерживает один однопоточный пользовательский процесс. Некоторые ОС семейства UNIX поддерживают процессы множества пользователей, но при этом каждый из процессов содержит один поток.

Многопоточностью (multithreading) называется способность ОС поддерживать в рамках одного процесса выполнение нескольких потоков. Примерами многопоточных систем являются среда выполнения Java, Windows, Linux, Solaris и другие. На рис. 2.4 представлены варианты однопоточных и многопоточных процессов.

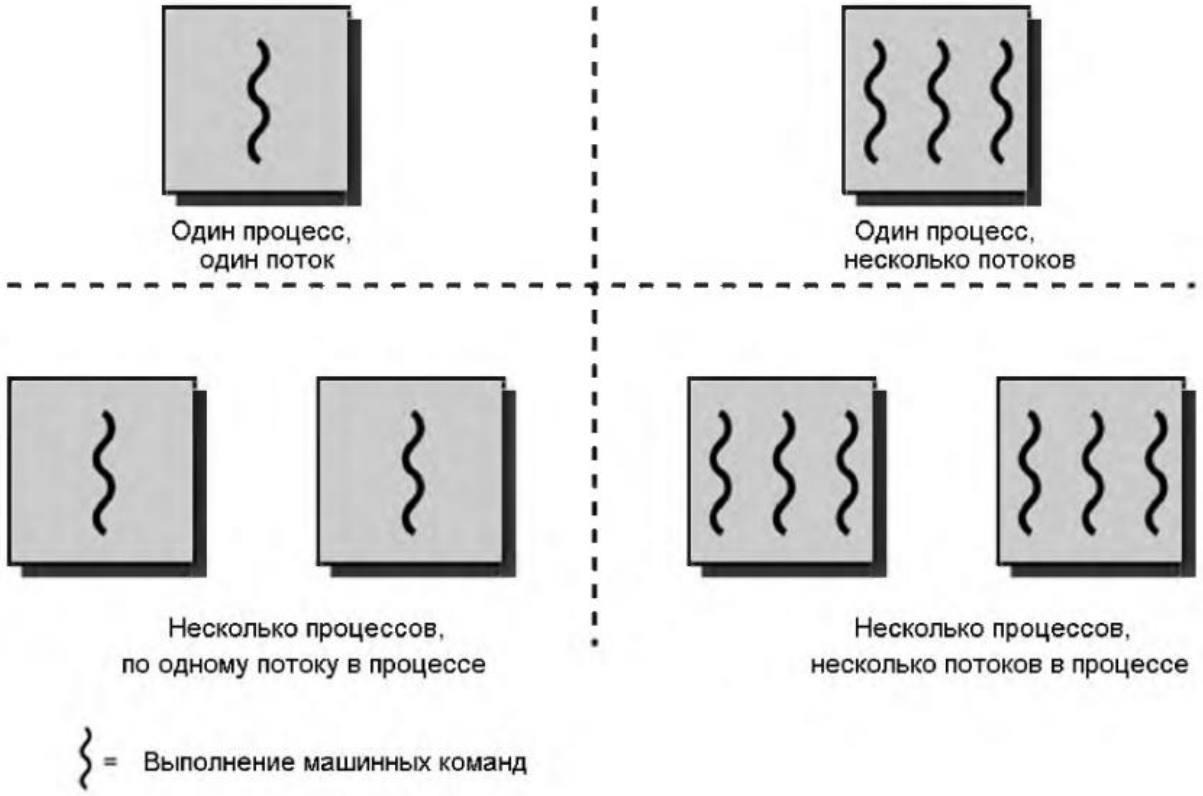

Рис 2.4. Однопоточные и многопоточные процессы

В многопоточной среде процесс можно рассматривать как структурную единицу объединения ресурсов и структурную единицу защиты. Ресурсами являются адресное пространство, открытые файлы, дочерние процессы, обработчики сигналов и многое другое. Под защитой подразумевается защищенный режим доступа к процессорам, другим процессам, файлам и ресурсам ввода-вывода. С другой стороны, процесс можно рассматривать как общий поток исполняемых команд, состоящий из нескольких отдельных потоков. У каждого потока есть свой счетчик команд, регистры и стек. Таким образом, в многопоточной среде процессы используются для группирования ресурсов, а потоки являются объектами, поочередно выполняющимися на процессоре (в случае однопроцессорной вычислительной системы), создавая впечатление параллельной работы потоков. Все потоки одного процесса разделяют между собой ресурсы процесса.

Они находятся в общем адресном пространстве и имеют доступ к одним и тем же данным. Например, если один поток изменяет данные в памяти, то другие потоки имеют возможность отследить эти изменения. Если один поток открывает файл с правом чтения, другие потоки данного процесса могут читать из этого файла.

Использование потоков имеет следующие преимущества с точки зрения производительности:

1. Создание нового потока в уже существующем процессе занимает существенно меньше времени, чем создание нового процесса.
2. Поток можно завершить быстрее, чем процесс.
3. Переключение потоков в рамках одного процесса происходит намного быстрее.
4. Использование потоков повышает эффективность обмена информацией между процессами. В большинстве операционных систем обмен информацией между процессами происходит с участием ядра, обеспечивающего механизм осуществления обмена и защиту. Обмен информацией между потоками может производиться без участия ядра.

Многопоточная модель процесса

Однопоточная модель процесса включает управляющий блок процесса, пользовательское адресное пространство (программы и данные), а также стеки ядра и пользователя, с помощью которых осуществляются вызовы процедур и возвраты из них при выполнении процесса.

Многопоточная модель процесса, помимо управляющего блока процесса и адресного пространства пользователя, включает дополнительно для каждого потока счетчик команд, регистры, стек, локальную память, состояние выполнения.

1. Счетчик команд - отслеживает порядок выполнения.
2. Регистры - используются для хранения текущих переменных.
3. Стек - содержит протокол выполнения процесса, где для каждой вызванной, но еще не вернувшейся процедуры, отведен отдельный фрейм. Во фрейме находятся локальные переменные процедуры и адрес возврата. Каждый поток может вызывать различные процедуры и соответственно иметь различный протокол выполнения процесса.
4. Локальная память - статическая память, выделяемая потоку для локальных переменных.
5. Состояние выполнения потока - одно из четырех основных состояний (готовый к выполнению, выполняющийся, блокированный, завершающийся).

В многопоточном режиме процессы обычно запускаются с одним потоком. Этот поток может создавать новые потоки, определив их указатели команд и аргументы. Новые потоки создаются со своим собственным контекстом регистров и стековым пространством, после чего помещаются в очередь готовых к выполнению потоков. В случае необходимости ожидания некоторого события поток блокируется (при этом сохраняется содержимое его пользовательских регистров, счетчика команд и указателя стека). После этого процессор может перейти к выполнению другого потока. После завершения потока его контекст регистров и стеки удаляются.

Многопоточная модель процесса охватывает две категории потоков: потоки на уровне пользователя (user-level threads - ULT) и потоки на уровне ядра (kernel - level

threads - KLT). Потоки на уровне пользователя управляются самим приложением. Обычно приложение в начале своей работы состоит из одного потока, с которого начинается выполнение данного приложения, и который размещается в процессе, управляемом ядром. Приложение может создать новый поток при помощи вызова библиотечной процедуры работы с потоками, например thread_create(). В результате создается структура данных для нового потока и управление передается к одному из готовых к выполнению потоков данного процесса в соответствии с заданным алгоритмом планирования. При этом контекст текущего потока сохраняется. При возврате управления к данному потоку его контекст восстанавливается. Все управление ULT осуществляется в пользовательском пространстве в рамках одного процесса. Связь с ядром отсутствует. Ядро продолжает осуществлять планирование процесса как единого целого. Однако существует взаимосвязь между планированием потоков и планированием процессов. В случае возникновения прерывания процесса, например по вводу-выводу или таймеру, выполняющийся поток процесса продолжает оставаться в состоянии выполнения, хотя перестает выполняться на процессоре. При возврате управления процессу возобновляется выполнение потока на процессоре.

Использование потоков на уровне пользователя обладает определенными преимуществами по сравнению с использованием потоков на уровне ядра. К числу таких преимуществ относятся:

• Для управления потоками процессу не нужно переключаться в режим ядра, что позволяет избежать дополнительных расходов.

• Планирование потоков может производиться в зависимости от специфики приложения. Для одних приложений лучше подходит простой алгоритм круговой диспетчеризации, а для других - алгоритм планирования с использованием приоритетов.

• Использование потоков на уровне пользователя применимо для любой операционной системы. Библиотека потоков представляет собой набор процедур, работающих на уровне приложения и совместно используемых всеми приложениями.

К числу недостатков использования ULT относятся:

• При выполнении операционной системой системного вызова блокируется не только данный поток, но и все другие потоки данного процесса.

• Поскольку ядро закрепляет за каждым процессом только один процессор, несколько потоков одного процесса не могут выполняться одновременно даже в случае многопроцессорной системы.

Потоки на уровне ядра управляются самим ядром через интерфейс прикладного программирования (API) средств ядра, управляющих потоками. Ядро поддерживает контекст процесса, а также контексты каждого отдельного потока процесса. Планирование выполняется ядром исходя из состояния потоков. Благодаря этому ядро может осуществлять планирование работы нескольких потоков одного процесса на нескольких процессорах. Кроме того, при блокировке одного из потоков процесса, ядро может выбрать для выполнения другой поток данного процесса. Основным недостатком использования потоков на уровне ядра являются дополнительные накладные расходы на переключения между потоками внутри одного процесса за счет переходов в режим ядра.

Если для большинства потоков приложения необходим доступ к ядру, то использование схемы KLT более целесообразно. Примерами использования потоков на уровне ядра являются ОС Windows, Linux. В не-которых операционных системах, например в ОС Solaris, используется комбинированный подход. Потоки на пользовательском уровне, входящие в приложение, отображаются в такое же или меньшее число потоков на уровне ядра (рис 2.5). При этом число потоков на уровне ядра является изменяемым параметром, что позволяет оптимизировать работу приложения.

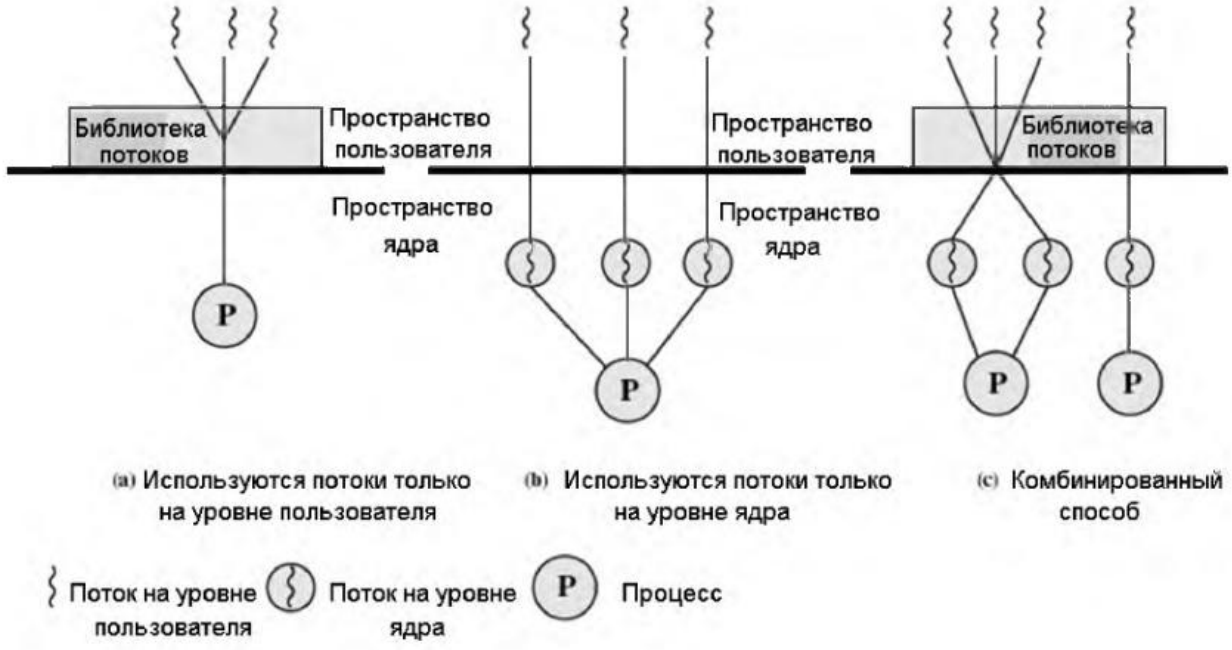

Рис 2.5. Потоки на уровне пользователя и ядра

Илюшкин Б.И. Операционные системы. Процессы и потоки: Учеб. пособие. – СПб.: СЗТУ, 2005, - 103 с.