

Тема 4. РЕСУРСЫ ВЫЧИСЛИТЕЛЬНОЙ СИСТЕМЫ

Под **ресурсом** понимается любой аппаратный или программный объект, который может понадобиться для работы процессов и доступ к которому может при этом вызвать конкуренцию процессов.

К важнейшим ресурсам любой системы относятся, основная память, периферийные устройства, файлы.

Общую классификацию ресурсов можно представить в виде:

По реальности существования: физический и виртуальный;

Под физическим понимают ресурс, который реально существует и при распределении его между пользователями обладает присущими ему физическими характеристиками. Виртуальный ресурс - это некоторая модель физического ресурса.

- По возможности расширения свойств: эластичный и жесткий

Признак "возможность расширения свойств" характеризует ресурс с точки зрения возможности построения на его основе некоторого виртуального ресурса. Физический ресурс, который допускает "виртуализацию", т.е. воспроизведение и (или) расширение своих свойств, называют эластичным. Жестким называется физический ресурс, который по своим внутренним свойствам не допускает виртуализацию.

- По степени активности: активный и пассивный

При использовании активного ресурса он способен выполнять действия по отношению к другим ресурсам или процессам, которые в общем случае приводят к изменению последних. Пассивный ресурс не обладает таким свойством. ЦП - активный ресурс, область памяти, выделяемая по требованию - пассивный ресурс.

- По времени существования: постоянный, временный

Если ресурс существует в системе до момента порождения процесса и доступен для использования на всем проведении интервала существования процесса, то такой ресурс является постоянным для данного процесса. Временный ресурс может появляться или уничтожаться в системе динамически в течение времени существования рассматриваемого процесса. Ресурсы, которые являются постоянными для одних процессов, могут быть временными для других, и наоборот.

- По степени важности: главный и второстепенный

Ресурс является главным по отношению к конкретному процессу, если без его выделения процесс принципиально не может развиваться. К таким ресурсам относятся прежде всего ЦП и ОП. Ресурсы, которые допускают некоторую альтернативу развития процесса, если они не будут выделены, называются второстепенными.

- По характеру использования: параллельно используемый, последовательно используемый;

Последовательная схема предполагает, что в отношении некоторого ресурса допустимо строго последовательное во времени выполнение цепочек действий "запрос-исполнение-освобождение" каждым процессом-потребителем этого ресурса. Параллельная схема предполагает одновременное использование одного ресурса более чем одним процессом.

При необходимости использовать какой-то ресурс (оперативную память, устройство ввода/вывода, массив данных и т.п.) задача обращается к супервизору ОС и сообщает о своем требовании. При этом указывается вид ресурса, его объем (например, количество адресуемых ячеек оперативной памяти, количество дорожек или секторов на системном диске, устройство печати и объем выводимых данных и т.п.), требование монопольного или совместного использования. Например, с файлом можно работать монопольно, а можно и совместно с другими задачами. Директива обращения к ОС передает ей управление, переводя процессор в привилегированный режим работы.

Ресурс может быть выделен задаче, если:

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

Получив запрос, операционная система либо удовлетворяет его и возвращает управление задаче, выдавшей данный запрос, либо, если ресурс занят, ставит задачу в очередь к ресурсу, переводя её в состояние ожидания.

После окончания работы с ресурсом задача сообщает операционной системе об отказе от ресурса, или ОС забирает ресурс сама.

Если в системе имеется некоторая совокупность ресурсов, то управляют их использованием на основе определенной стратегии. Стратегия подразумевает четкую формулировку целей, следуя которым можно добиться эффективного распределения ресурсов:

- быстрое обслуживание отдельных наиболее важных запросов;
- предоставление всем процессам равных возможностей;
- обслуживание максимально возможного количества процессов и наиболее полное использование всех ресурсов.

Основные виды ресурсов:

1) Процессор (процессорное время). Процессорное время делится попеременно (параллельно).

2) Оперативная память (ОП), которая может быть разделена и одновременным способом (в памяти одновременно может располагаться

несколько потоков), и попеременно (в разные моменты времени память может предоставляться для разных вычислительных процессов).

3) Внешняя память (например, память на магнитных дисках) и доступ к ней считаются разными видами ресурса, каждый из которых может предоставляться независимо от другого. Но для полной работы с внешней памятью необходимо иметь оба этих ресурса.

4) Внешние устройства могут разделяться параллельно в случае, если используются механизмы прямого доступа. Если же устройство работает с последовательным доступом, то оно не может считаться разделяемым ресурсом. Например, принтером невозможно воспользоваться попеременно двум параллельно выполняющимся потокам.

5) Программные модули, и прежде всего – системные программные модули. Однократно используемые модули правильно выполняются только один раз (выполняются на этапе загрузки ОС) и являются неделимым ресурсом. Они обычно вообще не распределяются как ресурс системы. Повторно используемые программные модули могут быть непривилегированными, привилегированными, реентерабельными и повторно входимыми.

Привилегированные программные модули работают при отключенной системе прерываний, т.е. никакие внешние события не могут нарушить естественный порядок вычислений. После окончания своей работы данный модуль может быть вызван снова из другой задачи. Такой модуль выступает как попеременно разделяемый ресурс. В первой его секции выключается система прерываний, а в последней – снова включается.

Непривилегированные программные модули – обычные модули, которые могут быть прерваны во время своей работы. Их нельзя считать разделяемыми, поскольку после прерывания такого модуля его нельзя запустить ещё раз из другого вычислительного процесса, т.к. промежуточные результаты для прерванных вычислений могут быть потеряны.

Реентерабельные (reenterable) программные модули допускают многократное прерывание своего исполнения и повторный их запуск из других задач. Они создаются таким образом, чтобы было обеспечено сохранение промежуточных вычислений в случае прерывания и возврат к сохраненным результатам при возобновлении процесса с прерванной ранее точки. Для этого могут использоваться статический или динамический способы выделения памяти под сохраняемые значения.

Повторно входимые (re-entrance) модули также допускают многократное параллельное использование, но, в отличие от реентерабельных, их нельзя прерывать. Они состоят из привилегированных секций (в которых запрещены прерывания) и имеют четко определенные точки входа. Они встречаются чаще реентерабельных.

6) Информационные ресурсы, т.е. в качестве ресурсов могут выступать данные. Они могут существовать как в виде переменных, так и в виде файлов.

Если потоки используют данные только для чтения, то такие информационные ресурсы можно разделять.