

СИСТЕМЫ УПРАВЛЕНИЯ ТЕХНОЛОГИЧЕСКИМИ ОБЪЕКТАМИ: ПРОЦЕССНО-ОРИЕНТИРОВАННЫЙ ПОДХОД К АНАЛИЗУ И ПРОЕКТИРОВАНИЮ

CONTROL SYSTEMS FOR TECHNOLOGICAL OBJECTS: PROCESS-ORIENTED APPROACH FOR ANALYSIS AND DESIGN

Санкт-Петербург, Санкт-Петербургский государственный технический университет
S-Petersburg, S-Petersburg state technical university

The paper deals with control systems for technological objects. We consider them as real-time systems and give their classification. We discuss also some problems of process-oriented analysis of such systems: periodic and sporadic processes, temporal scope, schedulability tests, priority assignment, predictability of their behavior, process synchronization, the class of priority inheritance protocols (PCP, SRP). When developing multiprocessing real-time systems, schedulability tests are used to formally prove that a given process set will meet its deadlines.

Системы управления технологическими объектами (ТО) представляют специфическую область применения вычислительной техники. Они предназначены для контроля и управления процессами, происходящими на этих объектах, и могут быть реализованы с использованием компьютеров или микропроцессоров, встроенных в ТО. В первом случае говорят о компьютерной системе управления, ориентированной на взаимодействие с человеком-оператором. Во втором случае - о встроенной (embedded) системе управления, в которой оператор либо отсутствует, либо его участие сведено к минимуму.

С точки зрения степени использования системных программных средств системы управления ТО подразделяются на несколько классов. В простейшем варианте программная часть системы реализуется прямым программированием микропроцессора на языке низкого уровня. Программы небольшие и предназначены для управления встроенными микропроцессорами. В более сложных системах разработка программного обеспечения ведется в инструментальной среде с высоким сервисом, а исполнение осуществляется в целевых компьютерных системах. Программное обеспечение, используемое в целевой системе, включает в себя системное ядро, реализующее основные системные функции. В наиболее развитых системах разработка и исполнение программного обеспечения осуществляется в рамках одной и той же операционной среды с полным сервисом.

Отличительными особенностями функционирования систем управления ТО являются:

- непрерывный характер функционирования,
- отказоустойчивость (как аппаратная, так и программная),
- способность адаптироваться к текущей ситуации,
- режим реального времени,
- предсказуемость поведения.

Рассмотрим подробнее последние два свойства. Функционирование систем управления ТО эффективно в том случае, если они успевают следить и выявлять особые события в развитии процессов и вовремя реагировать на эти события, оказывая управляющие воздействия на ТО. Другими словами объект управления и управляющий вычислительный комплекс должны функционировать в одном темпе. Режим функционирования системы, при котором события, возникающие на ТО, должны фиксироваться и обрабатываться системой по мере того, как они возникают и в установленные предельные сроки, называется режимом реального времени (РВ).

Программное обеспечение в системах РВ структурировано гораздо лучше, чем в системах общего назначения. Обычно какое-либо внешнее устройство, часы или датчик генерируют запрос для компьютера, который должен быть обработан до некоторого момента времени, называемого критическим сроком (deadline). Запросы могут поступать периодически или спорадически (случайно). В связи с этим различают два типа процессов в компьютерных системах управления ТО: периодические и спорадические, возникающие в ответ на некоторые асинхронные события (срабатывание датчиков) на ТО.

Важную роль для оценки предсказуемости поведения системы РВ [1] играют такие понятия как:

- совместное использования разделяемых ресурсов,
- характер обслуживания запросов (допускающий вытеснение выполняемого процесса или нет),
- жесткость критических сроков,
- детерминированность объекта управления и среды,
- используемая схема назначения приоритетов.

Системы РВ делятся на системы жесткого и мягкого РВ. Для первых нарушение режима грозит катастрофой, как например для систем управления полетом или автоматизированных химических производств. Для вторых - только снижением качества обслуживания, например, при передаче "живого" голоса по сети Интернет.

Важной частью систем управления РВ является планировщик заданий, который используя некоторый алгоритм диспетчеризации решает, какой из процессов должен выполнять в системе в каждый момент времени. Приемлемыми для режима РВ являются лишь алгоритмы, основанные на приоритетной многозадачности с вытеснением, основная идея которых состоит в том, что более высокоприоритетный процесс вытесняет низкоприоритетный процесс как только появляется запрос на его выполнение. Диапазон схем назначения приоритетов процессам достаточно широк: от полностью статических (все процессы заранее определены и имеют неизменные приоритеты) до динамических схем, в которых набор процессов, их приоритеты и даже сами схемы могут меняться в процессе функционирования системы в зависимости от сложившейся ситуации [2]. Наибольшее распространение в системах РВ получили схемы RMA, EDF, LLF. В первой схеме приоритеты периодическим процессам назначаются пропорционально частоте появления запросов на их выполнение, во второй - наиболее высокий приоритет у процесса, обрабатывающего запрос с ближайшим критическим сроком, в третьей - наиболее высокий приоритет у процесса, обрабатывающего запрос с наименьшим временным резервом до критического срока.

Точно оценить время выполнения процесса сложно, поэтому говорят о наихудшем времени его выполнения (worst case execution time - WCET) p . Критический срок завершения для периодических процессов часто совпадает с концом периода их повторения T . Время ответа на запрос представляет собой интервал от момента появления запроса до момента завершения его обработки. Переполнение наступает тогда, когда какой-либо из процессов, обрабатывающий некоторый запрос не успевает завершиться до своего критического срока.

Для периодического процесса i можно определить коэффициент использования процессора как отношение WCET к периоду процесса p_i/T_i . Понятно, что для эргодичности функционирования однопроцессорной системы необходимо и достаточно, чтобы сумма коэффициентов использования процессорного времени для всех процессов не превосходила единицы. Это соотношение однако не гарантирует завершения всех процессов до их критических сроков т.е. отсутствие переполнений и полной **диспетчируемости** системы. Не существует соотношения для оценки диспетчируемости системы в общем случае. Получены тесты диспетчируемости для ряда частных случаев [3]: если рассматривается множество n периодических

$$\sum_{i=1}^n p_i / T_i \leq 1$$

независимых процессов и используется схема EDF, то система диспетчируема при загрузке процессора до 100%,

Если же используется схема RMA, то диспетчируемость гарантируется только при уровне загрузки процессора до 69 %:

$$\sum_{i=1}^n p_i / T_i \leq n(2^{1/n} - 1)$$

Из этого примера видно, какой выигрыш можно получить при использовании той или иной схемы назначения приоритетов.

Другая важная проблема, как влияет синхронизация процессов на показатели функционирования систем РВ. Синхронизировать процессы необходимо в ряде случаев:

- когда один процесс готовит данные для другого и второй не должен быть запущен раньше завершения первого (существует отношение предшествования между процессами),
- когда процесс должен быть запущен в строго определенный момент астрономического времени,
- когда процесс запускается для обработки запроса от некоторого датчика,
- когда необходимо упорядочить доступ нескольких процессов к разделяемому общему ресурсу.

Последний случай требует дополнительных пояснений. Большинство процессов в системе оказываются фактически связанными использованием общих ресурсов. Если ресурсы, из-за которых конкурируют процессы, можно использовать только в монопольном режиме, то необходимо упорядочить доступ к разделяемому процессами ресурсу с помощью введения критических секций, семафоров или блокировок *MutEx*. Критической секцией процесса называется весь интервал его работы с разделяемым ресурсом. При входе в критическую секцию процесс должен проверять значение связанного с используемым ресурсом семафора (переменной, характеризующей состояние ресурса). Если семафор открыт, то процесс может работать с ресурсом в рамках данной секции, но должен переключить семафор в состояние "закрыто". При завершении критической секции процесс переключает семафор в состояние "открыто", разрешая другим процессам использовать данный ресурс.

Время выполнения высокоприоритетного процесса, ожидающего освобождения разделяемого ресурса, в общем случае, трудно предсказать и оценить поскольку существует неприятная возможность захвата разделяемого ресурса низкоприоритетным процессом. В этом случае высокоприоритетный процесс будет блокирован и вынужден ожидать завершения работы с ресурсом ряда низкоприоритетных процессов. Эта ситуация известна под названием "инверсия приоритетов" [4]. Для упорядочения отношений между процессами используется семейство специальных протоколов наследования приоритетов, включающее протоколы PCP и SRP. Оба протокола успешно работают со схемами назначения приоритетов RMA и EDF. Протоколы позволяют ограничить и оценить время, в течение которого возможна инверсия приоритетов и соответственно блокировка высокоприоритетного процесса, а также предотвращают возможность возникновения тупиковых ситуаций.

Тест диспетчируемости для множества n периодических процессов, использующих общий ресурс, схему назначения приоритетов RMA и протокол синхронизации PCP, может быть представлен в виде:

$$\sum_{i=1}^n p_i/T_i + \max \{b_1/T_1, \dots, b_{n-1}/T_{n-1}\} \leq n(2^{1/n} - 1)$$

где b_i - длительность самой длинной критической секции процессов, приоритет которых ниже, чем приоритет процесса i .

Выход в печать статьи [3] положил начало развитию новой теории оценки поведения систем РВ. В настоящее время на эту тему опубликовано большое количество работ и разработана прототипная среда для моделирования и анализа систем РВ, что позволяет применять результаты академических исследований в области процессно-ориентированного подхода в инженерной практике проектирования систем управления ТО.

Литература

1. М.В.Хлудова. Предсказуемость функционирования систем реального времени. В материалах научно-технической конференции "Фундаментальные исследования в технических университетах", СПб, СПбГТУ, 1998
2. Khloudova M. "Classification of scheduling algorithms for real-time systems". International Workshop on Nondestructive Testing and Computer Simulations in Science and Engineering, Proceedings of SPIE Vol. 3687 pp. 228-231 (1999)
3. C.L.Liu and J.W.Layland, "Scheduling algorithms for multiprogramming in a hard real-time environment," JACM, vol.20, no.1, pp.46-61 (1973)
4. М.В.Хлудова. Диспетчеризация вычислительных процессов в многопроцессорных распределенных системах управления интеллектуальными мобильными роботами. В материалах X научно-технической конференции "Экстремальная робототехника", СПб, СПбГТУ, 1999