

SFWR ENG 4AA4

Kemal Ahmed
Fall 2015
Dr. Down

Note: information from the pre-requisite, [SFWR ENG 3DX4](#) will not be included in this summary (although corrections will be).

Contents

Real-Time Systems	1
Classifications.....	1
Task optimization.....	2
Types of Scheduling	2
Static	2
FIFO	2
Dynamic	3
Multiprocessor.....	3
Task Interactions.....	3
Sporadic Server	4
Clocks	5
PID Control.....	6

Real-Time Systems

Classifications

What happens upon failure to meet deadlines:

- **Soft:** performance is degraded but not destroyed
- **Firm:** a few times will simply degrade performance, but after may lead to system failure
- **Hard:** complete and catastrophic system failure
 - **Safety Critical:** may cause injury / death (a type of hard)

Forward difference method: derivatives using $f'(x) = \frac{f(x+h) - f(x)}{h}$

Backwards Difference method: derivatives using $f'(x) = \frac{f(x) - f(x-h)}{h}$

Controller [C(s)]:

Input [E(s)]:

Output [U(s)]:

$$U(s) = C(s)E(s)$$

Task optimization

Task [T]: $T_i = (p_i, r_i, e_i, d_i)$

Period [p]: time between tasks are repeatedly released

Release time [r]: time it takes to release task

Execution time [e]: slowest time task could take to be completed (but assume the tasks will take this long no matter what)

Deadline [d]: when task needs to be completed

Number of tasks [n]:

Processor Utilization [U]: used as a priority level $U = \sum_{i=1}^n \frac{e_i}{p_i}$

If $U > 1$, nothing is feasible

If $r_i = 0$ and $p_i = d_i$, then write $T_i = (p_i, e_i)$

Types of Scheduling

Static

Static Scheduling:

- task's priority is assigned before execution and does not change
- If a task misses its deadline, you mess up all the deadlines after it like an airport at Christmas

FIFO

First In First Out (FIFO):

- Could cause problems for tasks whose execution time is significantly shorter than the rest when there are deadlines
 - E.g. $T_1 = (100, 3)$; $T_2 = (2, 1)$
- A.K.A. **First Come, First Served (FCFS)**

Schedule: the order in which tasks will be executed

Hyperperiod [H]: the entire length of a cycle, least common multiple

Harmonic: every task period evenly divides every longer period

Frame Size [f]:

- The best way for computers to segment the schedule in a way that it verify that the appropriate tasks have been executed
- Constraints:
 1. $f \geq \max_{1 \leq i \leq n} (e_i)$
 2. $H \% f = 0$
 3. $2f - \gcd(p_i, f) \leq d_i$

Least Compute Time (LCT): tasks with smallest execution times executed first

- Think *greedy*
- Works poorly; worse than RR

Rate Monotonic (RM): shorter period, higher priority

- Think: tasks requiring frequent attention should have higher priority
- If harmonic, feasible as long as $U \leq 1$
- If non-harmonic, guaranteed feasible if $U \leq n \left(2^{\frac{1}{n}} - 1 \right)$
 - If the equation fails, it still might be, so draw the whole thing to be safe.

Dynamic

Pre-empting: splitting a task up into multiple mini tasks. Also, if a task misses its deadline, halt the task at the deadline

The only two optimal dynamic priorities are:

- **Earliest Deadline First (EDF):**
 - more flexible, better U
 - If deadlines < periods, still optimal, but determining feasibility is NP-hard
 - Always feasible if $U \leq 1$
- **Least Slack Theorem (LST):** not as popular as EDF

Multiprocessor

Once you have multiple processors, neither EDF nor RM are guaranteed to work.

Look into first-fit algorithms

Task Interactions

Suspended: active choice, of access prevention until algorithm allows it to

Blocked: as a result of waiting for a resource to be free

How to do the timing diagrams with locks:

- $S_1 = \text{lock}(S_1)$
- $S_1^{\wedge} = \text{unlock}(S_1)$

One-shot Tasks: non-periodic tasks

Critical Section: when a task tries to acquire an already locked by another task resource

Priority Inversion: a method of avoiding deadlock by telling high priority tasks to share their resources with the lower priority tasks even when it's not their turn

- Allocate time, where T_1 has access to shared resource, so the time not allocated can be pre-empted
- Connect the pre-empted by T_1 when T_1 wants to access the resource
- Protect the resource with a semaphore

- You can make it so that tasks can use the resource even after they release the semaphore, but you risk overwriting in that time

Priority Inheritance Protocol (PIP):

- Temporarily raise the priority of a task only if and when it actually blocks a higher priority task; on leaving the critical section, the task priority reverts to its original value
- Issues:
 - If only one shared resource, there's only one possible schedule
 - If more than one resource blocking:
 - Blocking time may be excessively long
 - Deadlock may occur
 - If accessing multiple resources, you can only use them in the same order

Priority Ceiling Protocol (PCP):

- Which tasks require which resources?
- Doesn't give a shit about when they were released.
- **Priority Ceiling (PC):** maximum priority that tasks will be given
 - For a current task, the PC doesn't matter
- "The state of the art when resolving resource-contention issues"
- "Deadlock free for an arbitrary number of tasks with an arbitrary number of resources acted upon in an arbitrary way."
- Main points:
 - No locked resources, so free access
 - If resource is locked by other tasks, S_2 needs to have priority of T_2 higher than the PC (S_2). S_1 is (suspended)
 - Priority higher than $PC(S_2)$
 - If any task needs priority higher than the priority ceiling, it's suspended
- When entering critical sections, check if any other tasks have resources

Sporadic Server

Execution Budget [e_s]: periodic tasks aren't flexible...

Execution time [e_i]: ...sporadic tasks are

Deadline [d_i]: absolute deadline

Release Time [r_i]:

Set of Sporadic Tasks [θ]:

Sporadic Task [S_i]:

- Non-periodic task
- (r_i, e_i, d_i)
- Typically interrupt-driven

Rules [ρ]: set of rules regulating a sporadic server

Sporadic Server [Φ_s]: (p_s, e_s, θ, ρ)

Periodic Task: (p_s, e_s)

Φ_s scheduled with T_i according to RM

We don't use K_d because it looks at the derivative regardless of the size of the error function. If your error is a sine function with a small amplitude, K_d will only take the derivative into account and it will overcompensate.

Open loop response: plant with no control

Ziegler-Nichols Tuning Rule: a PID tuning rule

Look at the *open loop response*. It could have a longer rise time / overshoot than preferred.

1. Tangent to curve on upslope

High sample rate \rightarrow lots of high frequency noise

Clocks

Computer Clock [C]:

Attributes:

- Correctness
- Bounded Drift
- Monotonicity
- Chronoscopia

Drift [ρ]: rate of change of the clock value away from a perfect clock (each second)

There's usually a reason why a clock drifts

$$\left| \frac{dC(t)}{dt} - 1 \right| \leq \rho$$

(EPS):

Monotonicity: Clock will always have a consistent spacing and will only move in one order (forward / backwards)

SSL certs will fail signature if your clock is wrong as to ensure this

Chronoscopia [γ]: changing drift

second derivative of stuff $\left| \frac{d^2C(t)}{dt^2} \right| \leq \gamma$

Error bound [ϵ]:

Acceptance Test: $\omega(S_k, t) = \left\lfloor \frac{d_k - t}{p_s} \right\rfloor e_s - e_k - \sum_{S_i \in \theta, d_i < d_k} e_i - \xi_i$

PID Control