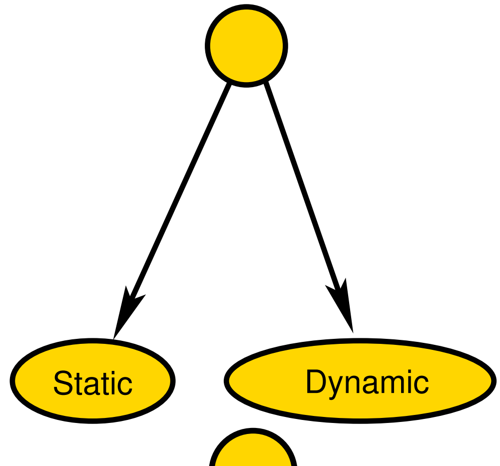
Embedded Systems: An Application-Centered Approach

Robert Dick

http://robertdick.org/esaca/
Office: 2417-E EECS
Department of Electrical Engineering and Computer Science
University of Michigan




Taxonomy of real-time systems



Static

- Task arrival times can be predicted.
- Static (compile-time) analysis possible.
- Allows good resource usage (low processor idle time proportions).
- Sometimes designers shoehorn dynamic problems into static formulations allowing a good solution to the wrong problem.

Dynamic

- Task arrival times unpredictable.
- Static (compile-time) analysis possible only for simple cases.
- Even then, the portion of required processor utilization efficiency goes to 0.693.
- In many real systems, this is very difficult to apply in reality (more on this later).
- Use the right tools but don't over-simplify, e.g.,

  > We assume, without loss of generality, that all tasks are independent.

  If you do this people will make jokes about you.

Soft real-time

- More slack in implementation
- Timing may be suboptimal without being incorrect
- Problem formulation can be much more complicated than hard real-time
- Two common (and one uncommon) methods of dealing with non-trivial soft real-time system requirements
  - Set somewhat loose hard timing constraints
  - Informal design and testing
  - Formulate as optimization problem

Hard real-time

- Difficult problem. Some timing constraints inflexible.
- Simplifies problem formulation.

Periodic

- Each task (or group of tasks) executes repeatedly with a particular period.
- Allows some nice static analysis techniques to be used.
- Matches characteristics of many real problems...
- ... and has little or no relationship with many others that designers try to pretend are periodic.

Periodic → Single-rate

- One period in the system.
- Simple.
- Inflexible.
- This is how a *lot* of wireless sensor networks are implemented.

Periodic → Multirate

- Multiple periods.
- Can use notion of circular time to simplify static (compile-time) schedule analysis E. L. Lawler and D. E. Wood.

  Branch-and-bound methods: A survey.
  Operations Research, pages 699–719, July 1966.
- Co-prime periods leads to analysis problems.

Periodic → Other

- It is possible to have tasks with deadlines less than, equal to, or greater than their periods.
- Results in multi-phase, circular-time schedules with multiple concurrent task instances.
  - If you ever need to deal with one of these, see me (take my code). This class of scheduler is nasty to code.

Aperiodic

- Also called sporadic, asynchronous, or reactive
- Implies dynamic
- Bounded arrival time interval permits resource reservation
- Unbounded arrival time interval impossible to deal with for any resource-constrained system

Definitions

- Task
- Processor
- Graph representations
- Deadline violation
- Cost functions

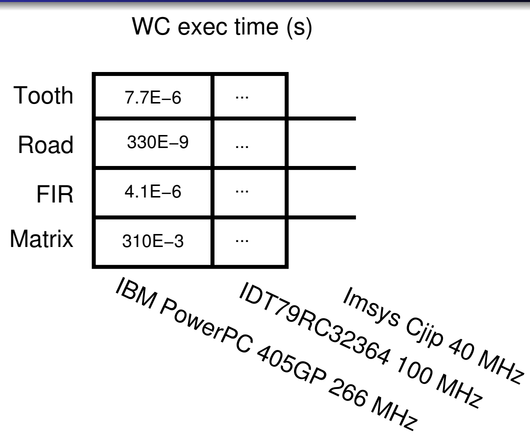
Task

- Some operation that needs to be carried out
- Atomic completion: A task is all done or it isn't
- Non-atomic execution: A task may be interrupted and resumed

Processor

- Processors execute tasks
- Distributed systems
  - Contain multiple processors
  - Inter-processor communication has impact on system performance
  - Communication is challenging to analyze
- One processor type: Homogeneous system
- Multiple processor types: Heterogeneous system

Task/processor relationship



Relationship between tasks, processors, and costs

Cost functions

- Mapping of real-time system design problem solution instance to cost value
- I.e., allows price, or hard deadline violation, of a particular multi-processor implementation to be determined

Back to real-time problem taxonomy: Jagged edges

- Some things dramatically complicate real-time scheduling
- These are horrific, especially when combined
  - Data dependencies
  - Unpredictability
  - Distributed systems
- These are irksome
  - Heterogeneous processors
  - Preemption

Central areas of real-time study

- Allocation, assignment and **scheduling**
- Operating systems and **scheduling**
- Distributed systems and **scheduling**
- **Scheduling is at the core or real-time systems study**

Allocation, assignment, and scheduling

How does one best

- Analyze problem instance specifications
  - E.g., worst-case task execution time
- Select (and build) hardware components
- Select and produce software
- Decide which processor will be used for each task
- Determine the time(s) at which all tasks will execute

Allocation, assignment, and scheduling

- In order to efficiently and (when possible) optimally minimize
  - Price, power consumption, soft deadline violations
- Under hard timing constraints
- Providing guarantees whenever possible
- For all the different classes of real-time problem classes

This is what I did for a Ph.D.

Operating systems and scheduling

How does one best design operating systems to

- Support sufficient detail in workload specification to allow good control, e.g., over scheduling, without increasing design error rate
- Design operating system schedulers to support real-time constraints?
- Support predictable costs for task and OS service execution

Distributed systems and scheduling

How does one best dynamically control

- The assignment of tasks to processing nodes...
- ... and their schedules

for systems in which computation nodes may be separated by vast distances such that

- Task deadline violations are bounded (when possible)...
- ... and minimized when no bounds are possible

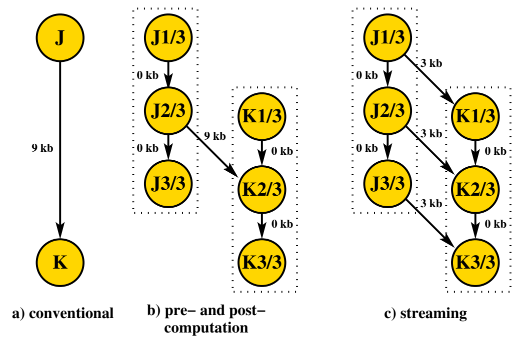
The value of formality: Optimization and costs

- The design of a real-time system is fundamentally a cost optimization problem
- Minimize costs under constraints while meeting functionality requirements
  - Slight abuse of notation here, functionality requirements are actually just constraints
- Why view problem in this manner?
- Without having a concrete definition of the problem
  - How is one to know if an answer is correct?
  - More subtly, how is one to know if an answer is optimal?

Optimization

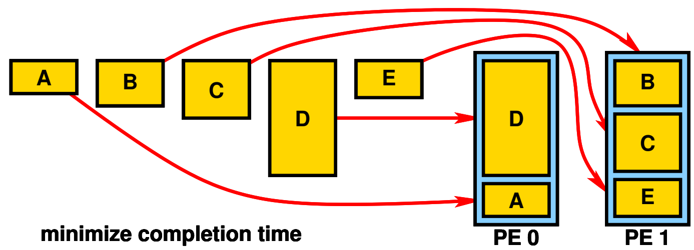
Thinking of a design problem in terms of optimization gives design team members objective criterion by which to evaluate the impact of a design change on quality.
Know whether your design changes are taking you in a good direction

Graph extensions



Allows pipelining and pre/post-computation
In contrast with book, not difficult to use if conversion automated

Problem definition



- Given a set of tasks,
- a cost function,
- and a set of resources,
- decide the exact time each task will execute on each resource

Types of scheduling problems

- Discrete time – Continuous time
- Hard deadline – Soft deadline
- Unconstrained resources – Constrained resources
- Uni-processor – Multi-processor
- Homogeneous processors – Heterogeneous processors
- Free communication – Expensive communication
- Independent tasks – Precedence constraints
- Homogeneous tasks – Heterogeneous tasks
- One-shot – Periodic
- Single rate – Multirate
- Non-preemptive – Preemptive
- Off-line – On-line

Discrete vs. continuous timing

System-level: Continuous

- Operations are not small integer multiples of the clock cycle

High-level: Discrete

- Operations are small integer multiples of the clock cycle

Implications:

- System-level scheduling is more complicated. . .
- . . . however, high-level also very difficult.
- Can we solve this by quantizing time? Why or why not?

Hard deadline – Soft deadline

Tasks may have hard or soft deadlines

- Hard deadline
  - Task must finish by given time or schedule invalid
- Soft deadline
  - If task finishes after given time, schedule cost increased

Real-time – Best effort

- Why make decisions about system implementation statically?
  - Allows easy timing analysis, hard real-time guarantees
- If a system doesn't have hard real-time deadlines, resources can be more efficiently used by making late, dynamic decisions
- Can combine real-time and best-effort portions within the same specification
  - Reserve time slots
  - Take advantage of slack when tasks complete sooner than their worst-case finish times

Unconstrained – Constrained resources

- Unconstrained resources
  - Additional resources may be used at will
- Constrained resources
  - Limited number of devices may be used to execute tasks

Uni-processor – Multi-processor

- Uni-processor
  - All tasks execute on the same resource
  - This can still be somewhat challenging
  - However, sometimes in $\mathcal{P}$
- Multi-processor
  - There are multiple resources to which tasks may be scheduled
- Usually $\mathcal{NP}$-complete

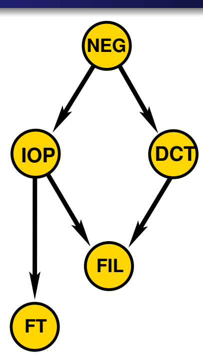
Homogeneous – Heterogeneous processors

- Homogeneous processors
  - All processors are the same type
- Heterogeneous processors
  - There are different types of processors
  - Usually $\mathcal{NP}$-complete

Free – Expensive communication

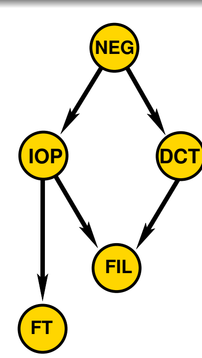
- Free communication
  - Data transmission between resources has no time cost
- Expensive communication
  - Data transmission takes time
  - Increases problem complexity
  - Generation of schedules for communication resources necessary
  - Usually $\mathcal{NP}$-complete

Independent tasks – Precedence constraints



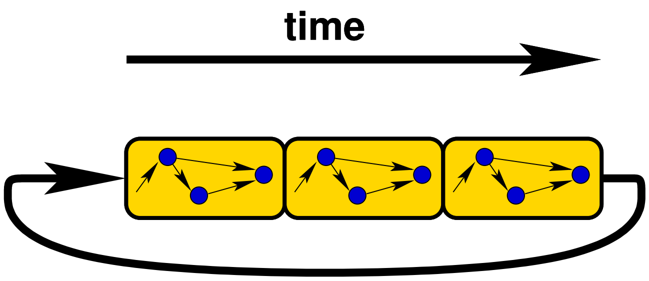
- Independent tasks: No previous execution sequence imposed
- Precedence constraints: Weak order on task execution order

Homogeneous – Heterogeneous tasks



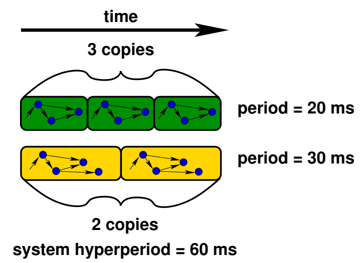
- Homogeneous tasks: All tasks are identical
- Heterogeneous tasks: Tasks differ

One-shot – Periodic



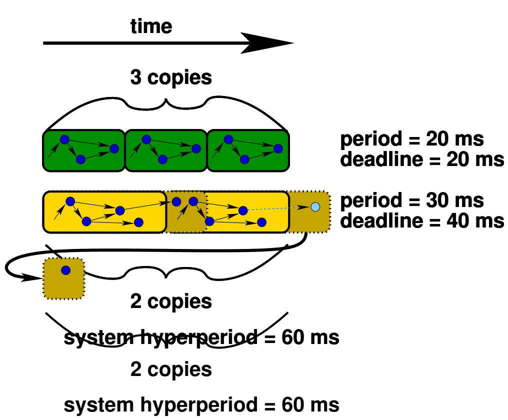
- One-shot: Assume that the task set executes once
- Periodic: Ensure that the task set can repeatedly execute at some period

Single rate – Multirate



- Single rate: All tasks have the same period
- Multirate: Different tasks have different periods
  - Complicates scheduling
  - Can copy out to the least common multiple of the periods (hyperperiod)

Periodic graphs



Aperiodic/sporadic graphs

- No precise periods imposed on task execution
- Useful for representing reactive systems
- Difficult to guarantee hard deadlines in such systems
  - Possible if minimum inter-arrival time known

Periodic vs. aperiodic

Periodic applications

- Power electronics
- Transportation applications
  - Engine controllers
  - Brake controllers
- Many multimedia applications
  - Video frame rate
  - Audio sample rate
- Many digital signal processing (DSP) applications

However, devices which react to unpredictable external stimuli have aperiodic behavior

Many applications contain periodic and aperiodic components

Aperiodic to periodic

Can design periodic specifications that meet requirements posed by aperiodic/sporadic specifications

- Some resources will be wasted

Example:

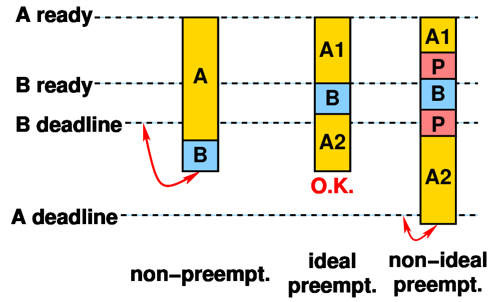
- At most one aperiodic task can arrive every 50 ms
- It must complete execution within 100 ms of its arrival time

Aperiodic to periodic

- Can easily build a periodic representation with a deadline and period of 50 ms
  - Problem, requires a 50 ms execution time when 100 ms should be sufficient
- Can use overlapping graphs to allow an increase in execution time
  - Parallelism required

The main problem with representing aperiodic problems with periodic representations is that the tradeoff between deadline and period must be made at design/synthesis time

Non-preemptive – Preemptive



- Non-preemptive: Tasks must run to completion
- Ideal preemptive: Tasks can be interrupted without cost
- Non-ideal preemptive: Tasks can be interrupted with cost

Off-line – On-line

Off-line

- Schedule generated before system execution
- Stored, e.g., in dispatch table. for later use
- Allows strong design/synthesis/compile-time guarantees to be made
- Not well-suited to strongly reactive systems

On-line

- Scheduling decisions made during the execution of the system
- More difficult to analyze than off-line
  - Making hard deadline guarantees requires high idle time
  - No known guarantee for some problem types
- Well-suited to reactive systems

Hardware-software co-synthesis scheduling

Automatic allocation, assignment, and scheduling of system-level specification to hardware and software

Scheduling problem is hard

- Hard and soft deadlines
- Constrained resources, but resources unknown (cost functions)
- Multi-processor
- Strongly heterogeneous processors and tasks
  - No linear relationship between the execution times of a tasks on processors

Hardware-software co-synthesis scheduling

- Expensive communication
  - Complicated set of communication resources
- Precedence constraints
- Periodic
- Multirate
- Strong interaction between $\mathcal{NP}$-complete allocation-assignment and $\mathcal{NP}$-complete scheduling problems
- Will revisit problem later in course if time permits

Behavioral synthesis scheduling

- Difficult real-world scheduling problem
  - Not multirate
  - Discrete notion of time
  - Generally less heterogeneity among resources and tasks
- What scheduling algorithms should be used for these problems?

Scheduling methods

- Clock
- Weighted round-robbin
- List scheduling
- Priority
  - EDF, LST
  - Slack
  - Multiple costs

Scheduling methods

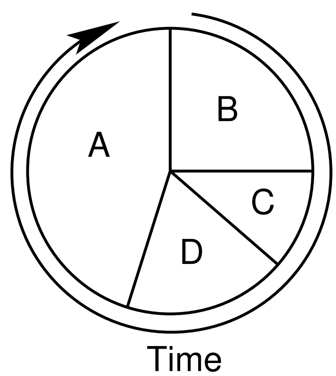
- MILP
- Force-directed
- Frame-based
- PSGA
- RMS

Clock-driven scheduling

Clock-driven: Pre-schedule, repeat schedule
Music box:

- Periodic
- Multi-rate
- Heterogeneous
- Off-line
- Clock-driven

Weighted round robbin



Weighted round-robbin: Time-sliced with variable time slots

List scheduling

- Pseudo-code:
  - Keep a list of ready jobs
  - Order by priority metric
  - Schedule
  - Repeat
- Simple to implement
- Can be made very fast
- Difficult to beat quality

Priority-driven

- Impose linear order based on priority metric
- Possible metrics
  - Earliest start time (EST)
  - Latest start time
    - Danger! LST also stands for least slack time.
  - Shortest execution time first (SETF)
  - Longest execution time first (LETF)
  - Slack (LFT - EFT)

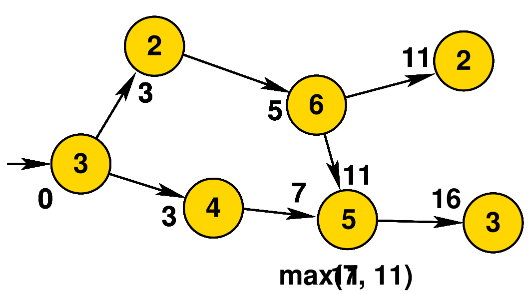
List scheduling

- Assigns priorities to nodes
- Sequentially schedules them in order of priority
- Usually very fast
- Can be high-quality
- Prioritization metric is important

Prioritization

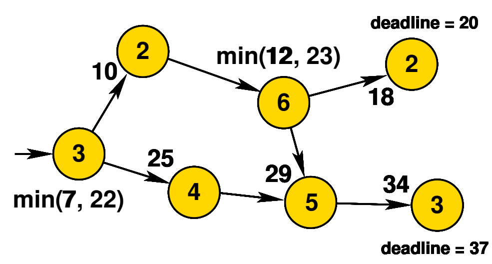
- As soon as possible (ASAP)
- As late as possible (ALAP)
- Slack-based
- Dynamic slack-based
- Multiple considerations

As soon as possible (ASAP)



- From root, topological sort on the precedence graph
- Propagate execution times, taking the max at reconverging paths
- Schedule in order of increasing earliest start time (EST)

As late as possible (ALAP)



- From deadlines, topological sort on the precedence graph
- Propagate execution times, taking the min at reconverging paths
- Consider precedence-constraint satisfied tasks
  - Schedule in order of increasing latest start time (LST)

Slack-based

- Compute EFT, LFT
- For all tasks, find the difference, LFT − EFT
- This is the *slack*
- Schedule precedence-constraint satisfied tasks in order of increasing slack
- Can recompute slack each step, expensive but higher-quality result
  - Dynamic critical path scheduling

Multiple considerations

- Nothing prevents multiple prioritization methods from being used
- Try one method, if it fails to produce an acceptable schedule, reschedule with another method

Effective release times

- Ignore the book on this
  - Considers simplified, uniprocessor, case
- Use EFT, LFT computation
- Example?

EDF, LST optimality

- EDF optimal if zero-cost preemption, uniprocessor assumed
  - Why?
  - What happens when preemption has cost?
- Same is true for slack-based list scheduling in absence of preemption cost

Breaking EDF, LST optimality

- Non-zero preemption cost
- Multiprocessor
- Why?

Multi-rate tricks

- Contract deadline
  - Usually safe
- Contract period
  - Sometimes safe
- Consequences?

Linear programming

- Minimize a linear equation subject to linear constraints
  - In $\mathcal{P}$
- Mixed integer linear programming: One or more variables discrete
  - $\mathcal{NP}$-complete
- Many good solvers exist
- Don't rebuild the wheel

MILP scheduling

P the set of tasks

t_{max} maximum time

$start(p, t)$ 1 if task p starts at time t, 0 otherwise

D the set of execution delays

E the set of precedence constraints

$$t_{start}(p) = \sum_{t=0}^{t_{max}} t \cdot start(p, t) \text{ the start time of } p$$

MILP scheduling

Each task has a unique start time

$$\forall_{p \in P}, \sum_{t=0}^{t_{max}} start(p, t) = 1$$

Each task must satisfy its precedence constraints and timing delays

$$\forall \{p_i, p_j\} \in E, \sum_{t=0}^{t_{max}} t_{start}(p_i) \geq t_{start}(p_j) + d_j$$

Other constraints may exist

- Resource constraints
- Communication delay constraints

MILP scheduling

- Too slow for large instances of $\mathcal{NP}$-complete scheduling problems
- Numerous optimization algorithms may be used for scheduling
- List scheduling is one popular solution
- Integrated solution to allocation/assignment/scheduling problem possible
- Performance problems exist for this technique

Force directed scheduling

- P. G. Paulin and J. P. Knight. Force-directed scheduling for the behavioral synthesis of ASICs. *IEEE Trans. Computer-Aided Design of Integrated Circuits and Systems*, 8(6):661–679, June 1989
- Calculate EST and LST of each node
- Determine the force on each vertex at each time-step
- Force: Increase in probabilistic concurrency
  - Self force
  - Predecessor force
  - Successor force

Self force

F_i all slots in time frame for i

F_i' all slots in new time frame for i

D_t probability density (sum) for slot t

δD_t change in density (sum) for slot t resulting from scheduling

self force

$$A = \sum_{t \in F_a} D_t \cdot \delta D_t$$

Predecessor and successor forces

pred all predecessors of node under consideration

succ all successors of node under consideration

predecessor force

$$B = \sum_{b \in \textbf{pred}} \sum_{t \in F_b} D_t \cdot \delta D_t$$

successor force

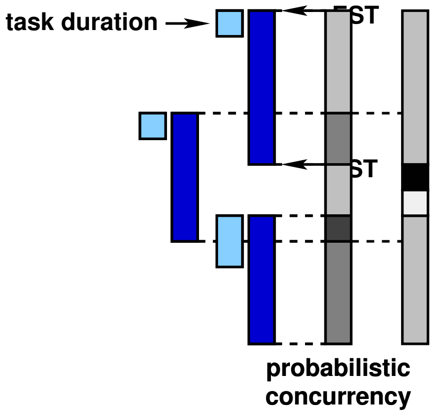
$$C = \sum_{c \in \textbf{succ}} \sum_{t \in F_c} D_t \cdot \delta D_t$$

Intuition

total force: $A + B + C$

- Schedule operation and time slot with minimal total force
  - Then recompute forces and schedule the next operation
- Attempt to balance concurrency during scheduling

Force directed scheduling



Force directed scheduling

- Limitations?
- What classes of problems may this be used on?

Implementation: Frame-based scheduling

- Break schedule into (usually fixed) frames
- Large enough to hold a long job
  - Avoid preemption
- Evenly divide hyperperiod
- Scheduler makes changes at frame start
- Network flow formulation for frame-based scheduling
- Could this be used for on-line scheduling?

Problem space genetic algorithm

- Let's finish off-line scheduling algorithm examples on a bizarre example
- Use conventional scheduling algorithm
- Transform problem instance
- Solve
- Validate
- Evolve transformations

Rate mononotic scheduling (RMS)

- Single processor
- Independent tasks
- Differing arrival periods
- Schedule in order of increasing periods
- No fixed-priority schedule will do better than RMS
- Guaranteed valid for loading $\leq \ln 2 = 0.69$
- For loading $> \ln 2$ and < 1, correctness unknown
- Usually works up to a loading of 0.88

Rate monotonic scheduling

Main idea

- 1973, Liu and Layland derived optimal scheduling algorithm(s) for this problem
- Schedule the job with the smallest period (period = deadline) first
- Analyzed worst-case behavior on any task set of size n
- Found utilization bound: $U(n) = n \cdot (2^{1/n} - 1)$
- 0.828 at $n = 2$
- As $n \to \infty$, $U(n) \to \log 2 = 0.693$
- Result: For any problem instance, if a valid schedule is possible, the processor need never spend more than 31% of its time idle

Optimality and utilization for limited case

- Simply periodic: All task periods are integer multiples of all lesser task periods
- In this case, RMS/DMS optimal with utilization 1
- However, this case rare in practice
- Remains feasible, with decreased utilization bound, for in-phase tasks with arbitrary periods

Rate monotonic scheduling

- Constrained problem definition
- Over-allocation often results
- However, in practice utilization of 85%–90% common
  - Lose guarantee
- If phases known, can prove by generating instance

Critical instants

Main idea:

> A job's critical instant a time at which all possible concurrent higher-priority jobs are also simultaneously released

Useful because it implies latest finish time

Definitions

- Period: T.
- Execution time: C.
- Process: i.
- Utilization: $U = \sum_{i=1}^{m} \frac{C_i}{T_i}$.
- Assume Task 1 is higher priority than Task 2, and thus $T_1 < T_2$.

Case 1 I

All instances of higher-priority tasks released before end of lower-priority task period complete before end of lower-priority task period.

1. $C_1 \leq T_2 - T_1 \left\lfloor \frac{T_2}{T_1} \right\rfloor$.
2. I.e., the execution time of Task 1 is less than or equal to the period of Task 2 minus the total time spent within the periods of instances of Task 1 finishing within Task 2's period.
3. Now, let's determine the maximum execution time of Task 2 as a function of all other variables.
4. $C_{2,max} = T_2 - C_1 \left\lceil \frac{T_2}{T_1} \right\rceil$.
5. I.e., the maximum execution time of Task 2 is the period of Task 2 minus the total execution time of instances of Task 1 released within Task 2's period.

Case 1 II

6. In this case,

$$\begin{aligned} U &= U_1 + U_2 \\ &= \frac{C_1}{T_1} + \frac{C_{2,max}}{T_2} \\ &= \frac{C_1}{T_1} + \frac{T_2 - C_1 \left\lceil \frac{T_2}{T_1} \right\rceil}{T_2} \\ &= \frac{C_1}{T_1} + 1 - \frac{C_1 \left\lceil \frac{T_2}{T_1} \right\rceil}{T_2} \\ &= 1 + C_1 \left(\frac{1}{T_1} - \frac{1}{T_2} \left\lceil \frac{T_2}{T_1} \right\rceil \right) \end{aligned}$$

7. Is $\frac{1}{T_1} - \frac{1}{T_2} \left\lceil \frac{T_2}{T_1} \right\rceil < 0$?
8. Thus, U is monotonically decreasing in C_1.

Case 2 I

Instances of higher-priority tasks released before end of lower-priority task period complete after end of lower-priority task period.

1. $C_1 \geq T_2 - T_1 \left\lfloor \frac{T_2}{T_1} \right\rfloor$.
2. $C_{2,max} = -C_1 \left\lfloor \frac{T_2}{T_1} \right\rfloor + T_1 \left\lfloor \frac{T_2}{T_1} \right\rfloor$.
3. $U = \frac{T_1}{T_2} \left\lfloor \frac{T_2}{T_1} \right\rfloor + C_1 \left(\frac{1}{T_1} - \frac{1}{T_2} \left\lfloor \frac{T_2}{T_1} \right\rfloor \right)$.

Minimal U

1. $C_1 = T_2 - T_1 \left\lfloor \frac{T_2}{T_1} \right\rfloor$.
2. $U = 1 - \frac{T_1}{T_2} \left(\left\lceil \frac{T_2}{T_1} \right\rceil - \frac{T_2}{T_1} \right) \left(\frac{T_2}{T_1} - \left\lfloor \frac{T_2}{T_1} \right\rfloor \right)$.
3. Let $I = \left\lfloor \frac{T_2}{T_1} \right\rfloor$ and
4. $f = \frac{T_2}{T_1}$.
5. Then, $U = 1 - \frac{f(1-f)}{I+f}$.
6. To maximize U, minimize I, which can be no smaller than 1.
7. $U = 1 - \frac{f(1-f)}{1+f}$.
8. Differentiate to find mimima, at $f = \sqrt{2} - 1$.
9. Thus, $U_{min} = 2\left(\sqrt{2} - 1\right) \approx 0.83$.
10. Is this the minimal U? Are we done?

Notes on RMS

- DMS better than or equal RMS when deadline $\neq$ period
- Why not use slack-based?
- What happens if resources are under-allocated and a deadline is missed?

Scheduling summary

- Scheduling is a huge area
- This lecture only introduced the problem and potential solutions
- Some scheduling problems are easy
- Most useful scheduling problems are hard
  - Committing to decisions makes problems hard: Lookahead required
  - Interdependence between tasks and processors makes problems hard

Mixing on-line and off-line

- Book mixes off-line and on-line with little warning
- Be careful, actually different problem domains
- However, can be used together
- Superloop (cyclic executive) with non-critical tasks
- Slack stealing
- Processor-based partitioning

Vehicle routing

- Low-price, slow, ARM-based system
- Long-term shortest path computation
- Greedy path calculation algorithm available, non-preemptable
- Don't make the user wait
  - Short-term next turn calculation
- 200 ms timer available

Mixing on-line and off-line

- Slack stealing
- Processor-based partitioning

Bizarre scheduling idea

- Scheduling and validity checking algorithms considered so far operate in time domain
- This is a somewhat strange idea
- Think about it and tell/email me if you have any thoughts on it
- Could one very quickly generate a high-quality real-time off-line multi-rate periodic schedule by operating in the frequency domain?
- If not, why not?
- What if the deadlines were soft?

Example problem: Static scheduling

- What is an FPGA?
- Why should real-time systems designers care about them?
- Multiprocessor static scheduling
- No preemption
- No overhead for subsequent execution of tasks of same type
- High cost to change task type
- Scheduling algorithm?

Problem: Uniprocessor independent task scheduling

- Problem
  - Independent tasks
  - Each has a period = hard deadline
  - Zero-cost preemption
- How to solve?

What to do by Friday I

1. Make a team folder under the interviews directory on Google Drive.
2. Enter your interviews so far.
3. Make a team folder in ESACA on Google Drive. Email me today if it hasn't been shared with you already. Pick a brief team name focused on your value props.
4. In the team folder, have a "value propositions" file, "hypotheses" file, and an "interviews" directory.
5. Name interviews as follows [date]-[title]. For example 19jan15-building-manager.
6. Have 22 interviews (total) in by Friday and adjust your value propositions and hypotheses based on the interview results.

What to do by Friday II

7. Also prepare a "system architecture" file that describes the embedded system design at a high level and makes clear which parts you are doing in this course, and which are building blocks you are taking from existing work. This can be left at the outline level for now if that is the only way to get the interviews done in time.