TENTAMEN TDDD07 Realtidssystem

DATUM: 19 August 2017
TID: 14-18
PLATS: TER2
ANSVARIG JOURLÄRARE: Simin Nadjm-Tehrani (0702 282412)

Material: English-Swedish-English dictionary
Calculator

No of assignments: 6

Total no. of points: 40

Preliminary grade limits for grades: 3, 4 and 5

3:	20 - 26 p
4:	27 - 33 p
5:	34 - 40 p

INSTRUCTIONS:
Please write your anonymous ID on each sheet of paper that you hand in. Pages should only contain answer to **one question per page** (answers to sub-questions can be on the same page). You are asked to only write on one side of each paper. Please **sort all the sheets** that you hand in, in the order of question numbers.

Make sure that **all** answers are **motivated** and supported by **clear** explanations. Figures or charts can be used to provide a clearer explanation but should be accompanied by a **textual description**. Points will not be given to answers for which the reasoning cannot be followed or that cannot be read due to bad handwriting. Wrong answers/reasoning which is embedded in partially correct ones will lead to deduction of points.

Hints: Read the question carefully to find the focus of the question. Make sure your answer is to the point and relevant for the question asked. Take the opportunity of asking questions about unclear issues during the exam session. Otherwise, whenever in doubt about the question, write down your interpretation and assumptions, and answer the question based on that interpretation. Try to dispose of your time on each question in proportion of the assignment points.

Results are reported no later than 5th September 2017.

Good luck!

Simin Nadjm-Tehrani

Q1: Scheduling

a) ABB's current industrial robots are a major improvement over the previous ones, since the same control system is now able to manage several robots at the same time, making the production process more efficient. For example, one robot can hold a metal surface that is supposed to be welded in position while another robot performs the welding. Then the holding robot can move the surface during operation so that the welding point is always facing up. A computer can steer several activities:

- Positioning of the robots, for which regulator processes with a sampling interval of 20 ms are used.
- Trajectory definition, by a process P_T which creates the sequence of desired positions for the positioning regulator.
- Fault monitoring, done by a process P_F that looks for deviations from a robot expected behaviour, and adjusts or signals the operator if needed.

The computer used for running these processes can be used to steer several robots. In the first version of the system an analysis of the utilisation of the processor for *two* robots is performed, whereby each robot has three positioning processes, one fault monitoring process, and one trajectory process (altogether 10 processes in the system).

Assume that every positioning process (i.e. in X, Y, Z axis for each robot) has a maximal computational time of 1 ms. Assume further that the trajectory generation and the fault monitoring processes have max. computation times of 5ms and 10 ms respectively. The sampling interval for trajectory generation is 30 ms.

1) Assume that faults do not take place more often than once every 100 ms and the fault monitoring process can take care of one fault at a time. Is the task set schedulable with rate monotonic scheduling? Provide all application-based assumptions that you make for the analysis.

(5 points)

2) Now assume that the minimum inter-arrival time for faults is 1 second. Assume further that positioning and trajectory generation processes can be optimised so that their maximal computation times are halved. Using the same processor and with EDF as a scheduling mechanism, how many robots in total can be served by the same processor?

(2 points)

3) Assume that a new task P_D is to be added to the task set in order to update the control panel data displayed for the human operator every second. The display panel can be considered as a shared resource on which also the trajectory generation and the fault monitoring processes can write to display their current output. Using the rate-monotonic policy, consider that the six positioning processes are run as one process P_P, the two trajectory processes are run as one process P_T, and the two fault monitoring processes as one process P_F, all with the same periods as in part 1) above. The fourth process P_D has a max computation time 6 ms of which 2 ms is for exclusively accessing the display. Assume that the locking of the display for the write operations by P_T and P_F is 1 ms. The four processes use the immediate ceiling protocol to avoid priority inversion and deadlock related problems. Compute the maximum blocking time for each of the tasks P_D, P_T, and P_F in the system.

(3 points)

b) Describe two advantages and one disadvantage of scheduling task sets with the cyclic scheduling method.

(3 points)

c) Based on the course literature and the invited industry lectures in the course describe how industry/business applications benefit from the scheduling knowledge that you obtained during this course by exemplifying 3 applications, where *at most* one example considers CPU (and its utilization) as a resource. Relate each application to a specific part of the knowledge acquired.

(3 points)

Q2: Dependability and predictability

Use the terminology of IFIP Working Group 10.4 (from the course literature) to answer the following question.

The Guardian reported on 3rd June 2017: "The two graduate students [Barrs and Chen] at the University of California, Berkeley, have devised a system that would have tightly-packed clusters of autonomous vehicles zipping past local traffic at speeds of more than 100mph, all on existing roadways. They call it Hyperlane, and it works a lot like high-speed toll lanes already do, only with a central computer controlling everything.

Hyperlane works a lot like existing dedicated commuter lanes, only the separate lanes are only for autonomous vehicles. After entering an acceleration lane, Hyperlane's central computer takes over the car's functions and finds a slot for it in the already fast-moving traffic in the dedicated lanes. Barrs and Chen said vehicles would travel at speeds up to 120mph, and that the centralized computer control – which would be in constant communication with each vehicle using emerging 5G technology – would allow for a more tightly-packed traffic pattern. Sensors in the road would evaluate traffic density, weather hazards, accidents and other changes, prompting the system to adjust vehicle speed as necessary."

Identify two possible faults that can be treated using techniques for *fault tolerance* and that are specific to this scenario (relate it to the text above) and describe what would be the consequent error that would appear in a system of this kind. Motivate your choice of fault-error by relating to the appropriate fault tolerance technique.

(6 points)

Q3: Real-time Communication

a) Describe what is the notion of "node" in a time-triggered architecture (TTA). That is, what are hardware elements and implemented functions that comprise a node in a TTA system?

(2 points)

b) Describe the principle for allocating message priorities in a CAN bus scheduling context. How does future changes in an application and addition of new messages affect the work to be done for (re)analyzing the schedulability of communication in a real-time application with hard constraints?

(2 points)

Q4: Application design & RTOS

a) Take a stand (true/false) on each of the following statements and motivate your answer!

(1) Non-functional requirements need specification languages that have specific stereotypes for those types of requirements.

(2) AUTOSAR is an industrial standard for interfaces in the critical networks domain.

(3) OSEK is a time-triggered operating system widely deployed in one industrial sector.

(4) The PIM part of a model driven architecture approach proposed by OMG is useful since it relates to the specification of aspects that do not change from one platform to another.

(4 points)

b) How is the application developer expected to estimate the scheduler's overhead in a real-time OS, so that utilization-based approaches to analyzing schedulability can be applied with a margin for the overhead?

(2 points)

Q5: Distributed systems, Quality of Service (QoS)

a) List three metrics for measuring Quality of Service (QoS) in Internet-based applications and provide three techniques for QoS enforcement such that each method uses one of the listed metrics for evaluation.

(3 points)

b) In order to manage the virtual machine (VM) loads to adapt to the dynamic changes in arriving tasks the Xiao et al. 2013 paper proposes live migration of tasks. What is the cost of live migration mostly dependent on? How is this used when deciding which task to migrate when adapting the physical machine (PM) loads?

(3 points)

c) What is the difference between elastic applications and tolerant applications in the terminology of Internet-based QoS enforcement methods?

(2 points)

Q6: Bonus points

In this question you state if you have any bonus points allocated to your attempts at bonus exercises 1, 2, and 3 during the course. Please sum up all three (if any) of your attempted exercises and write the total attained points here.

Notation for Processes

C = Worst-case execution time
B = Worst-case blocking time
D = Relative deadline
n = Number of processes
T = Period
R = Worst-case response time
J = Release jitter

Schedulability test for Rate Monotonic:

$$\sum_{i=1}^{n}\left(\frac{C_i}{T_i}\right) \leq n(2^{1/n} - 1)$$

Schedulability test Earliest Deadline First:

$$\sum_{i=1}^{n}\left(\frac{C_i}{T_i}\right) \leq 1$$

RMS Response time analysis

$$w_i = C_i + B_i + \sum_{\forall P_j \in hp(P_i)} \left\lceil \frac{w_i + J_j}{T_j} \right\rceil C_j$$

$$R_i = w_i + J_i$$

$hp(P_i)$ is the set of processes with a higher priority than process P_i.

Timing Analysis of CSMA/CR

B = blocking time
C = transmission time of entire frame
T = period
τ_{bit} = transmission time of one bit
w = response time for the first bit of a frame to be sent
R = total response time
J = Jitter
t = Longest busy interval
$lp(m)$ = set of frames with lower priority than m.
$hp(m)$ = set of frames with higher priority than m.
$hep(m)$ = set of frames with higher or equal priority than m.
n = number of bytes in message (data field)

$$R_m = \max_{q=0..Q_m-1}(R_m(q))$$

$$R_m(q) = J_m + w_m(q) - q \cdot T_m + C_m$$

$$w_m(q) = B_m + q \cdot C_m + \sum_{\forall j \in hp(m)} \left\lceil \frac{w_m(q) + J_j + \tau_{bit}}{T_j} \right\rceil \cdot C_j$$

(with $w_m{}^0(q) = B_m + C_m q$)

$$Q_m = \left\lceil \frac{t_m + J_m}{T_m} \right\rceil$$

$$t_m = B_m + \sum_{j \in hep(m)} \left\lceil \frac{t_m + J_j}{T_j} \right\rceil \cdot C_j \quad \text{(with } t_m{}^0 = C_m \text{)}$$

$$C_m = \left(8n + 47 + \left\lfloor \frac{34 + 8n - 1}{4} \right\rfloor \right) \tau_{bit}$$

$$B_m = \max_{j \in lp(m)}(C_j)$$