Networked Embedded Systems Lab
Prof. Marco Zimmerling

Introduction to Embedded Systems – WS 2022/23

Exercise 8: Architecture Synthesis II

Task 1: Scheduling with Pipeline Resources

Pipeline-resources process data in time intervals that are smaller than the actual execution time w. As soon as after the start of a task v_1 the so-called *pipeline-interval* PI has elapsed, the next task v_2 can be started on the same resource (see Figure 1). Non-pipeline-resources are a special case of pipeline-resources with $PI = w$.

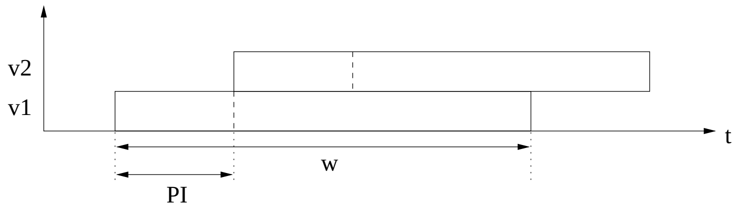

Figure 1: Tasks on pipeline-resource

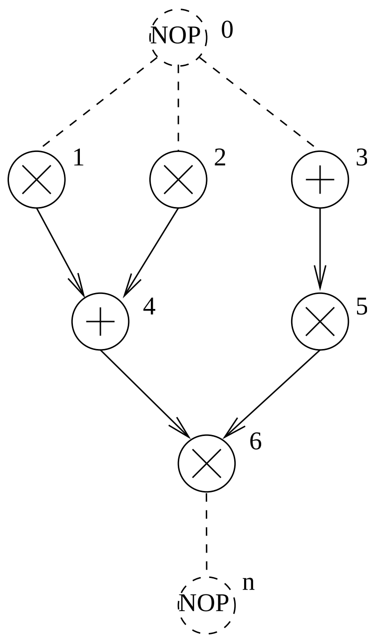

Figure 2: Sequence graph for Pipelining

a) Modify the LIST algorithm given in the lecture notes so that pipeline-resources are considered. Which step has to be reformulated and how? (Explain your answer!)

b) Perform the scheduling for the sequence graph given in Figure 2 using the modified algorithm. You can use Table 1. The multiplication (r_2) lasts 4 time units and the length of the pipeline-interval is 2 time units. The addition (r_1) lasts 2 time units and cannot be executed as pipeline-operation. 1 adder and 1 multiplier are available. Use the number of successor nodes as priority criterion. What is the resulting latency?

Task 2: Integer Linear Programming

Given the sequence graph $G_S = (V_S, E_S)$ in Fig. 3.

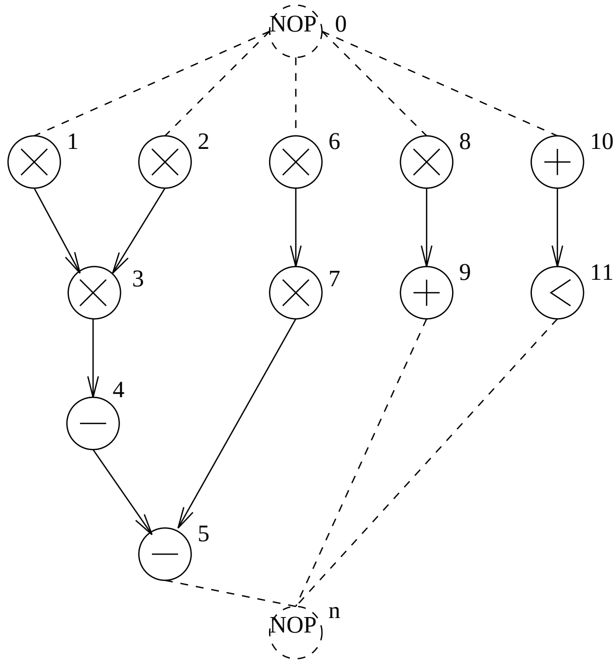

Figure 3: Sequence graph.

For the execution times of the operations assume: A multiplication operation (MULT) takes 2 time units and all other (ALU) operations take 1 time unit each. Two units of the resource type r_1 (multiplier) and two units of the resource type r_2 (ALU) are allocated.

(a) Apply the ASAP and ALAP algorithms to compute the earliest (l_i) and the latest (h_i) starting time of all operations $v_i \in V_s, i \in \{1, \ldots, 11\}$. For ALAP, assume the maximum latency $\overline{L} = 7$. Fill in the starting times in Table 2.

(b) Formulate the problem of latency minimization with restricted resources as an integer linear program (ILP). For this, you should introduce the binary variables $x_{i,t} \in \{0, 1\}$ $\forall v_i \in V_S$ and $\forall t \in \{t \in \mathbb{Z} \mid l_i \leq t \leq h_i\}$. $\tau(v_i)$ is used to denote the starting time of operation $v_i \in V_S$ and $\alpha(r_i)$ with $r_i \in V_R = \{\mathrm{MULT}, \mathrm{ALU}\}$ denotes the number of allocated resource instances. Given the above notations, write down the following equations/inequations without using the $\sum$ symbol.

  (i) Express the objective function of the ILP
  (ii) Define $\tau(v_i)$ $\forall i \in \{1, \ldots, 11\}$ as a function of $x_{i,t}$, where $l_1 \leq t \leq h_1$
  (iii) Express all data dependencies
  (iv) Express all resource limitations

(c) In an analogous manner try to formulate an ILP that solves the problem of cost minimization with latency limitation. Hint: We assume that the cost of a realization is the sum of the costs c of the multipliers with $c(r_1) = 2$ per allocated unit, and of the ALUs with $c(r_2) = 1$ per allocated unit. For the latency bound, we choose $\bar{L} = 6$.

t	k			
0	r_1			
	r_2			
1	r_1			
	r_2			
2	r_1			
	r_2			
3	r_1			
	r_2			
4	r_1			
	r_2			
5	r_1			
	r_2			
6	r_1			
	r_2			
7	r_1			
	r_2			
8	r_1			
	r_2			
9	r_1			
	r_2			
10	r_1			
	r_2			
11	r_1			
	r_2			
12	r_1			
	r_2			
13	r_1			
	r_2			

Table 1: Table for Task 1

Task 3: Iterative Algorithms

Please answer the following questions considering the given video codec application specified as a marked graph in Figure 4.

	l_i (ASAP)	h_i(ALAP)
v_1		
v_2		
v_3		
v_4		
v_5		
v_6		
v_7		
v_8		
v_9		
v_{10}		
v_{11}		

Table 2: Earliest and latest starting times (Task 2a)

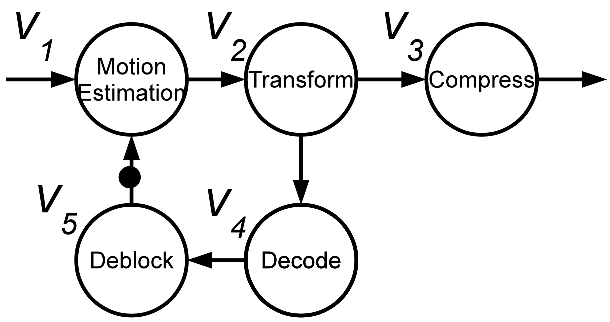

Figure 4: Video codec marked graph representation

	ν_1	ν_2	ν_3	ν_4	ν_5
$w(\nu_i)$	10	10	10	5	5

Table 3: Execution time of each function

(a) Formulate all existing dependencies in Figure 4 from ν_i to ν_j in the form of

$$\tau(\nu_j) - \tau(\nu_i) \geq w(\nu_i) - d_{ij} \cdot P,$$

where P is the minimum iteration interval. The execution time of each function is listed in Table 3.

(b) Assuming unlimited resources and only one token on the edge between ν_5 and ν_1, determine the minimum iteration interval P and the latency L. To justify your answer, draw the scheduling on the timeline given in Figure 5 with the dependency from ν_5 to ν_1 highlighted.

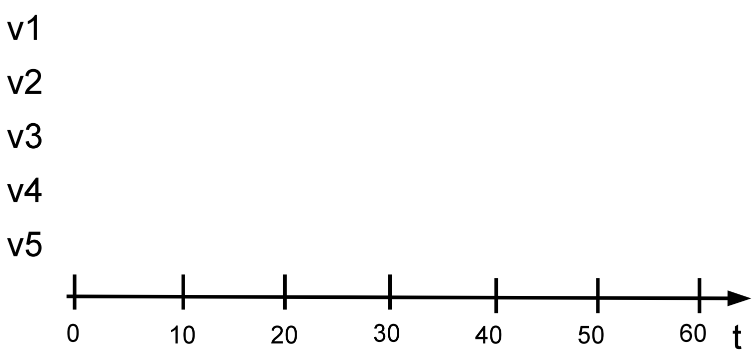

Figure 5: Scheduling result of the video codec

(c) The motion estimation function (ν_1) uses the result of the previous frame (See the dependency between ν_1 and ν_5). Let us now suppose that any arbitrary number of tokens can be inserted to reduce P using functional pipelining. Then, determine the minimum number of tokens that should be added on the edge $\nu_5 \to \nu_1$ to achieve $P = 10$? To justify your answer, draw the pipelined scheduling on the timeline given in Figure 6 with the dependency from ν_5 to ν_1 highlighted and calculate the latency L of the schedule.

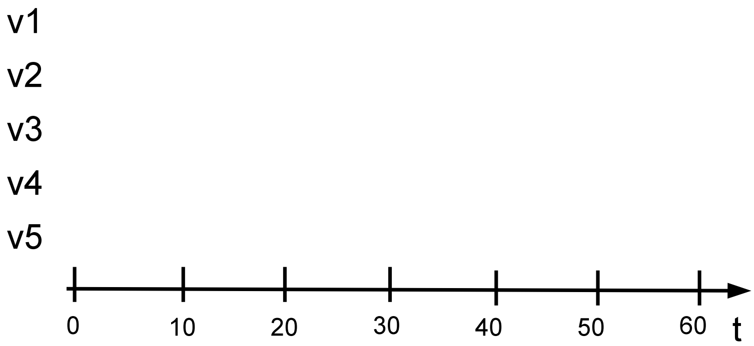

Figure 6: Pipelined scheduling result of the video codec