Homework 3

**Premise.** This homework is optional but warmly advised to check the understanding of the student lecture-by-lecture. It should be used to improve the student preparation during the course. All the answers must be properly and clearly explained.

**Exercise 1.** Given a farm parallelism form with emitter, workers and collector modules (at any level), provide a formal proof that the cost model of the ideal service time of the parallelization is the following: \( T_{id\_farm} = \max\{T_{id\_E}, T_{id\_C}, T_{id\_W}/n\} \) where \( n \geq 1 \) is the number of worker modules.

**Exercise 2.** Given a map parallelism form with a sequential distributor module (scatter or multicast), workers and gather modules (at any level), provide a formal proof that the cost model of the ideal service time of the parallelization is the following: \( T_{id\_map} = \max\{T_{id\_D}, T_{id\_G}, T_{id\_W}\} \).

**Exercise 3.** Given the acyclic computation graph with a single source in the figure:

![Computation Graph](image)

Provide the performance analysis of the graph by finding the values of:
- the effective service time, utilization factor and relative efficiency of each module;
- the ideal service time, effective service time and efficiency of the system;
- the processing latency per stream element and the completion time of the whole system by assuming a stream length of \( M = 10^4 \) input items.

**Exercise 4.** A process \( P_2 \) receives a stream of pairs of floating point numbers \((x_i, y_i)\) and for each pair executes the following computation, where \( s \) is an initialized floating point variable: \( \forall \text{input } (x, y) : z = \{r = F_1(x, y); s = F_2(y, s); t = F_3(y, s); z = F_4(r, t); \} \). The process receives the stream from a process \( P_1 \) with ideal service time \( 4.5 \times 10^2 \tau \). Assume that:
- the calculation times of the four functions are: \( T_{F_1} = 200\tau, T_{F_2} = 300\tau, T_{F_3} = 150\tau, T_{F_4} = 400\tau \);
- we can neglect the communication latency, i.e. \( L_{com} \approx 0 \).

Determine whether \( P_2 \) is a bottleneck and in case study a data-flow parallelization by discussing the achieved effective service time, relative efficiency and processing latency. Although the communication latency can be neglected in this problem, discuss how the data-flow graph is interfaced with the pre-existing process \( P_1 \), i.e. which kinds of collective communications \( P_1 \) executes to forward the necessary operands to the processes of the data-flow graph.