Homework 2

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. Write a fully parametric LC program implementing the producer-consumer paradigm with $N \geq 1$ and $M \geq 1$ producer and consumer processes respectively.

Exercise 2. Write the LC program of a process $P$ receiving non-deterministically messages from $N > 1$ input channels with the same data type (integers – one word each). The process has an output channel of type integer. The semantics of $P$’s computation is the following:
- let $x_{i,j}$ the $j$-th stream element received from the $i$-th input channel. The $j$-th output result produced by $P$ is computed as follows, $y_j = \sum_{i=0}^{N-1} x_{i,j}$.

Hint: the student must use properly the features and expressive power of the alternative command.

Exercise 3. Write the LC program implementing at the process level a generic switch unit for a uni-directional ring network (see slide 28, slide block 5). The LC version must be able to serve the same number of messages per service time of the equivalent firmware implementation.

Exercise 4. A process $P$ receives a stream of $10^5$ non-zero integers (let $x$ be a generic input element) with inter-arrival time $T_a = 1M\tau$. The process encapsulates an array of $M = 1$ Mega integers statically initialized. The sequential algorithm executed on each input element is the following:

```plaintext
receive(ch_in, x);

x' = F(x, ...);

y = 0;

for i=0 to M-1 do
    if(A[i]%x' == 0) { A[i]++; y++; }

y' = G(y);

send(ch_out, y');
```

The average calculation time of the two functions are $T_F = 1.5M\tau$ and $T_G \approx 0$. The parameters of the communication cost model are $T_{setup} = 10^2\tau$ and $T_{transm} = 10\tau$.

- evaluate the performance analysis of the system (ideal/effective service time, bandwidth, utilization factor, optimal parallelism degree, latency, efficiency and completion time);
- if needed, write a parallelization of $P$ and evaluate the performance analysis (same parameters of the previous point) for the whole system and for the parallelized module. Evaluate the scalability of the parallel implementation.

Hint: the student must be careful in taking into account the effect of data partitioning in the exploitation of the reuse property.

Assume the following architectural specifications:

1. multiprocessor with $N = 32$ PEs and 32 shared memory macro-modules;
2. D-RISC CPU with mean service time per instruction equal to $4\tau$;
3. each PE has on-demand 32K primary cache (16K instruction cache + 16K data cache) and secondary cache of 512K. Block size equal to 8 words;
4. the base memory access latency per cache block is equal to $L_{read-C1}(\sigma) = 72\tau$; the latency for transferring a block from the secondary cache to the primary one is $L_{C2-C1}(\sigma) = (\sigma + 2)\tau$;
5. the architecture allows overlapping of communication with calculation.