Assignment 2: Dataflow analysis using Intel Pin

Perform a data dependency characterization of the different SPEC CPU2017 workloads using Intel Pin as described below.

Consider the following metrics:

1. RAW distance: Suppose an instruction `j` consumes the register value produced by instruction `i`. The RAW distance of this dependence `i→j` is defined as the number of instructions between `i` and `j` (`i` not included).
2. WAW distance: Suppose an instruction `j` writes to the same architectural register as an earlier instruction `i`, and there are no intervening writes to the same register. The WAW distance of this hazard `i→j` is defined as the number of instructions between `i` and `j` (`i` not included).
3. WAR distance: Suppose an instruction `j` writes to the same architectural register as that read by an earlier instruction `i`, and there are no intervening writes or reads to the same register. The WAR distance of this hazard `i→j` is defined as the number of instructions between `i` and `j` (`i` not included).
4. Store-load distance: Suppose a load instruction `j` consumes the data produced by a store instruction `i`. The Store-Load distance of this dependence `i→j` is defined as the number of instructions between `i` and `j` (`i` not included).

Plot distributions of the above metrics for the different SPEC CPU2017 benchmarks. Your report must contain the plots, your observations of different patterns in the plots, and your inferences based on these observations.

Deadline: 23:55, 30/08/2023 (2 weeks)
Marks: 5