Parallelism

There are several levels of the parallelism:

- Instruction-level parallelism: The processor with the instructions can re-order, pipeline, split …
  In order to increase the parallelism here we could have deeper pipelines or use a **multiple issue** technique.
- Thread-level parallelism (TLP): Each thread can do a specific task.

There are 2 main types of multiple issue:

- Static multiple issue: The compiler chooses to group instruction to be issued together, issue packet, the compiler packages them into issue slots and detects and avoids hazards. Issue packets are considered "very long instructions".
- Dynamic multiple issue: The CPU examines the instruction stream and chooses instructions to issue each cycle, the compiler can help reordering instructions and the CPU resolves hazards using advanced techniques at runtime. "Superscalar" processors are the ones that use this.

Some other types of parallelism are:

- Multithreading: Performing multiple threads of execution in parallel. There is **fine-grain** multithreading which switch threads after each cycle, interleave instruction execution and if one thread stalls, others are executed. There is also **coarse-grain** multithreading only switch on long stalls, simplifies hardware, but doesn't hide short stalls.
- Multicore.