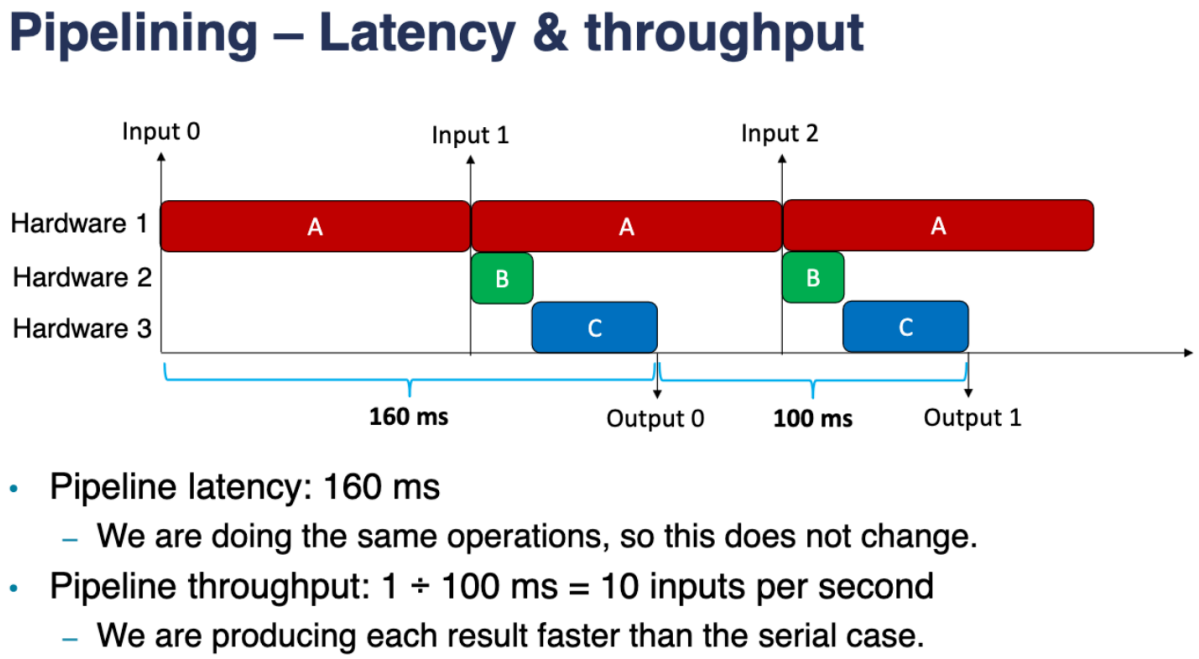
Serial Operations

- the operations all took very different amounts of time
- **latency** Time between the input arriving and the output being produced
- **throughput** Total number of inputs that can be processed per unit time.

Pipelining

- pipelining latency doesn't change
- throughput increases



Limitations

- requires OP to use different hardware units
- always has overhead (setup individual hardware units, may override benefit from pipelining)
- requires extra memory to support double buffering
  - avoid reading and writing into the same array at the same time
  - can also use circular buffer

Example