

Operators in flink

source-stream = kafka ①

mapped = map(source-stream) ②

filter = filter(mapped) ③

sink = kafka(filter) ④



Running 1 operator = 1 task

Parallelism in flink

↳ We can set parallelism on job level ①

↳ We can set parallelism on operator level ②

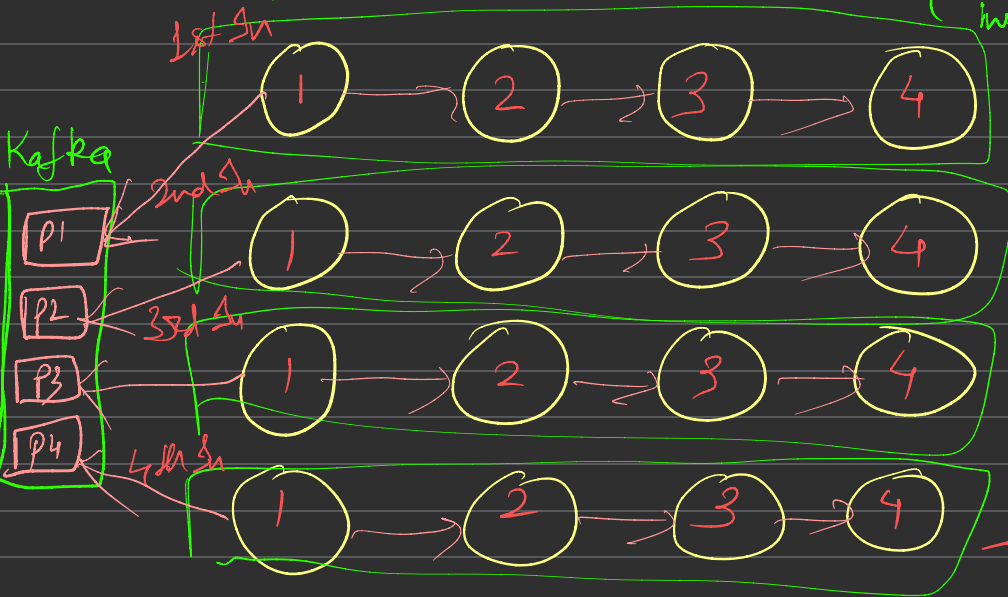
① Example \Rightarrow env.set-parallelism(4)

② Example \Rightarrow read-stream = from-source(kafka)
set-parallelism(2)

Dataflow Graph if parallelism is 4

Dataflow Graph if parallelism is 4

4 parallel
instance of task
flow



if Source operator has 2 parallelism and
Sink has 3 and transformation operators
have 4

Kafka

