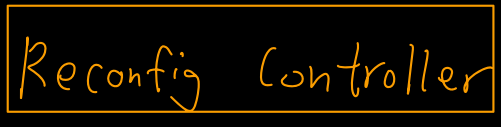
1. Tuplelization: read from a file line by line and then split each into words
2. Calculate the reconfiguration time (from graph 3-6 blow), the time between stream stop and stream resume.

1. Implement a time based tumbling window to calculate the through put, and dump it into a log file and using another program to plot the through put change, or port them into a realtime visualization app with throughput stats plotted lively.

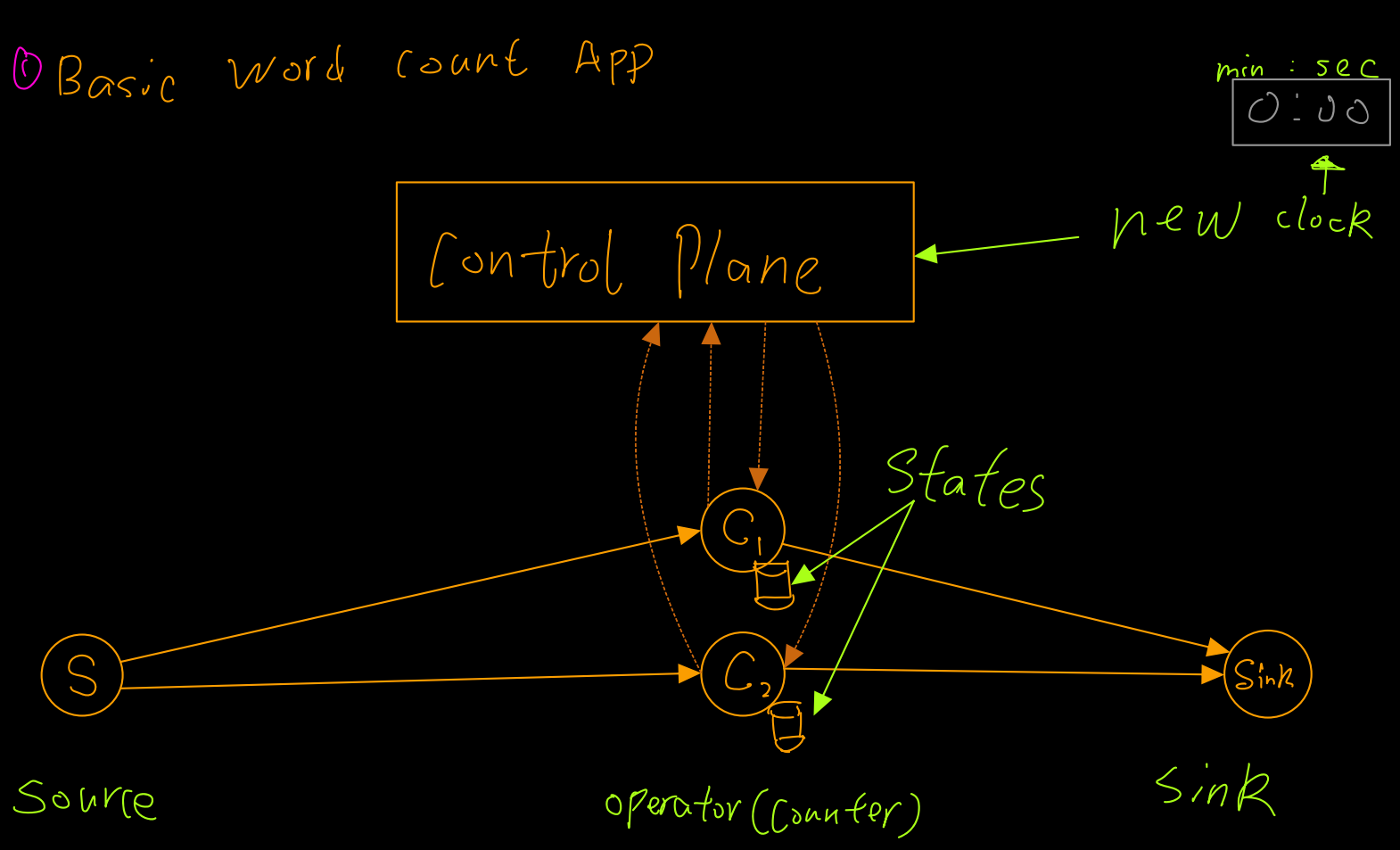
1. Update the word counter (value of the key-value pair) and pass it to the downflow(sink)
2. Can transfer part of the state to another worker
3. Can received and combine state from another worker



1. A process that start timing once instantiated.
2. Notify and trigger reconfiguration
3. Instantiate new worker

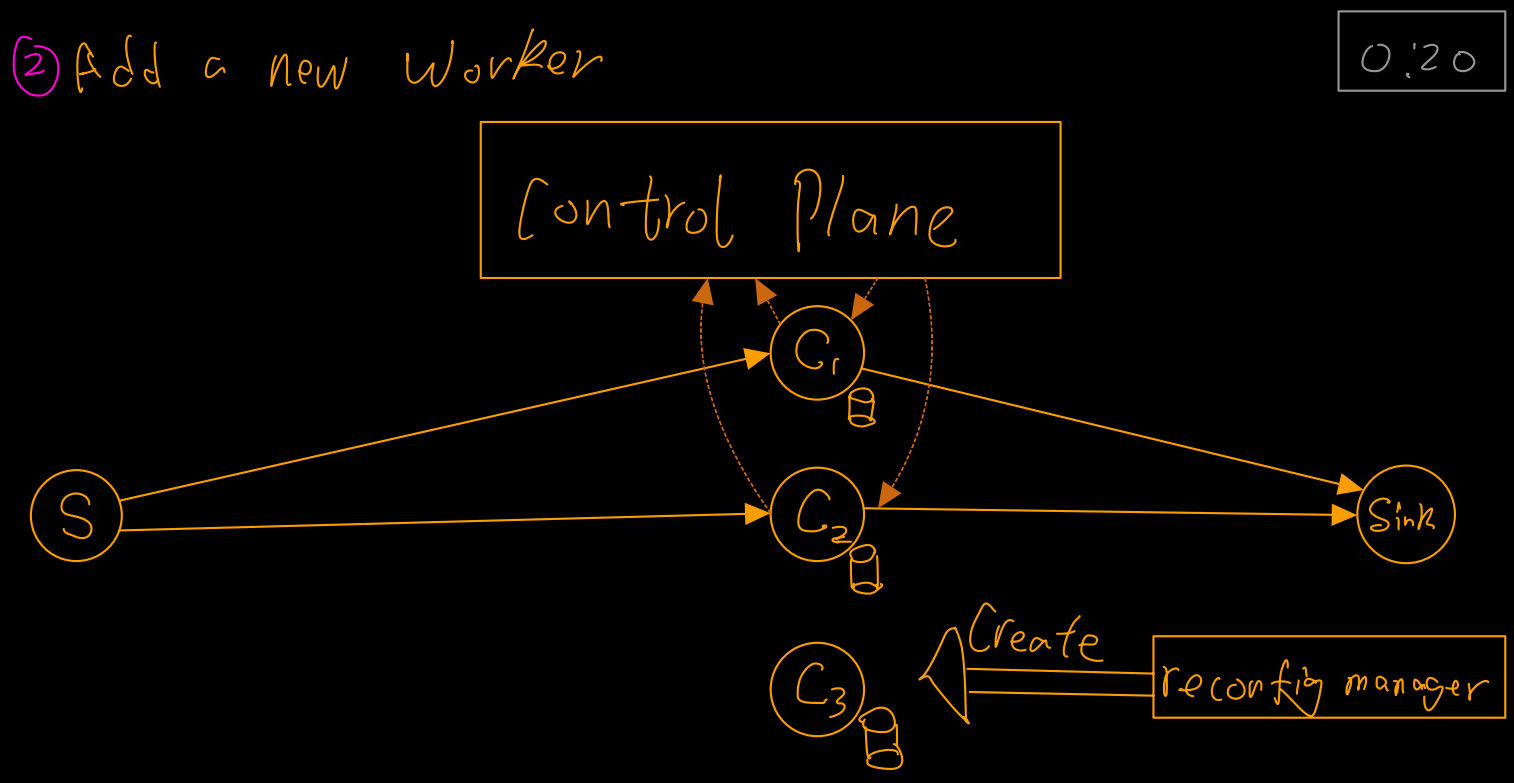
1. Maintain a state table for workers
2. Worker can ask control plane where's my states.
3. Can check which tuple belongs to which worker.

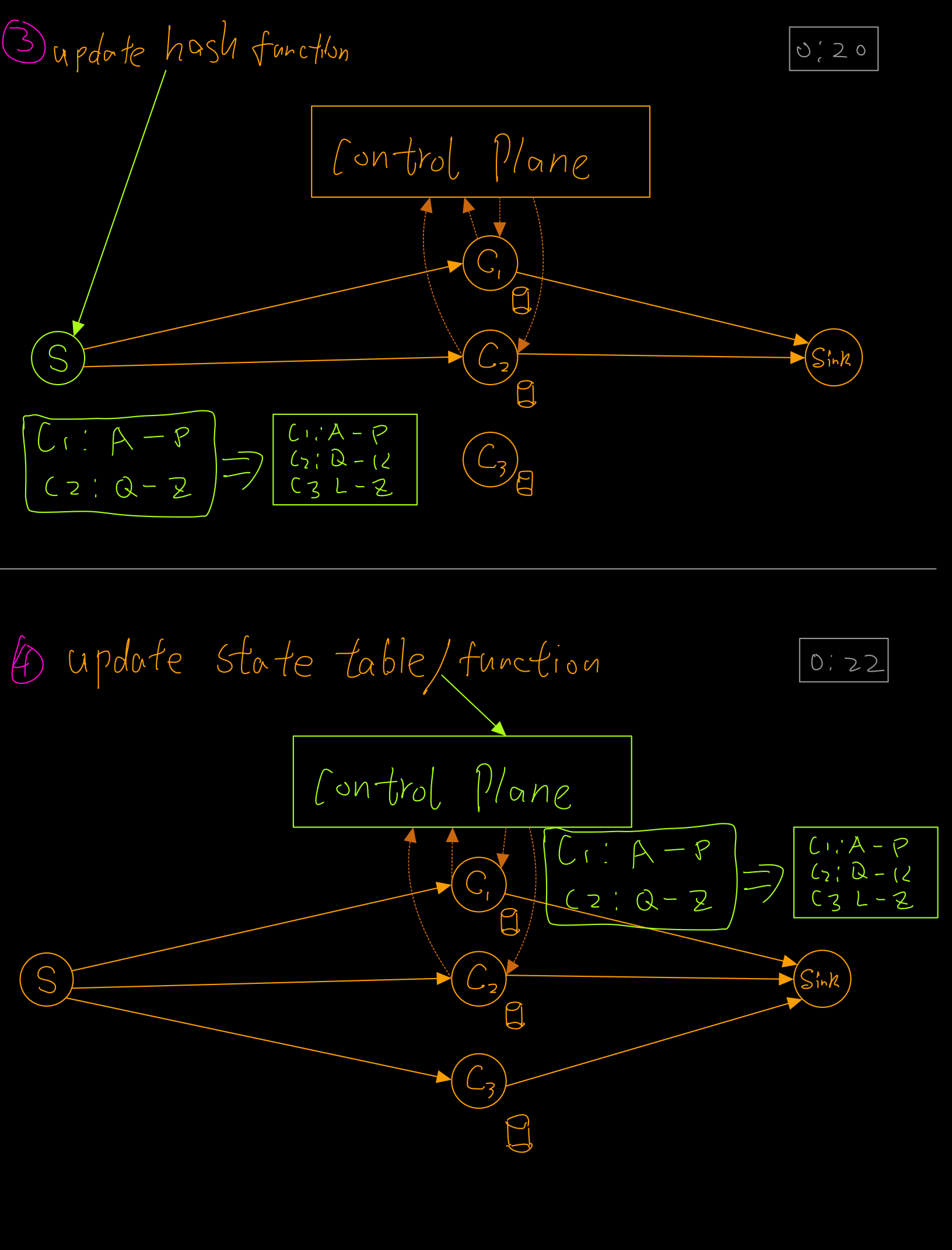
① Basic Word Count App



---

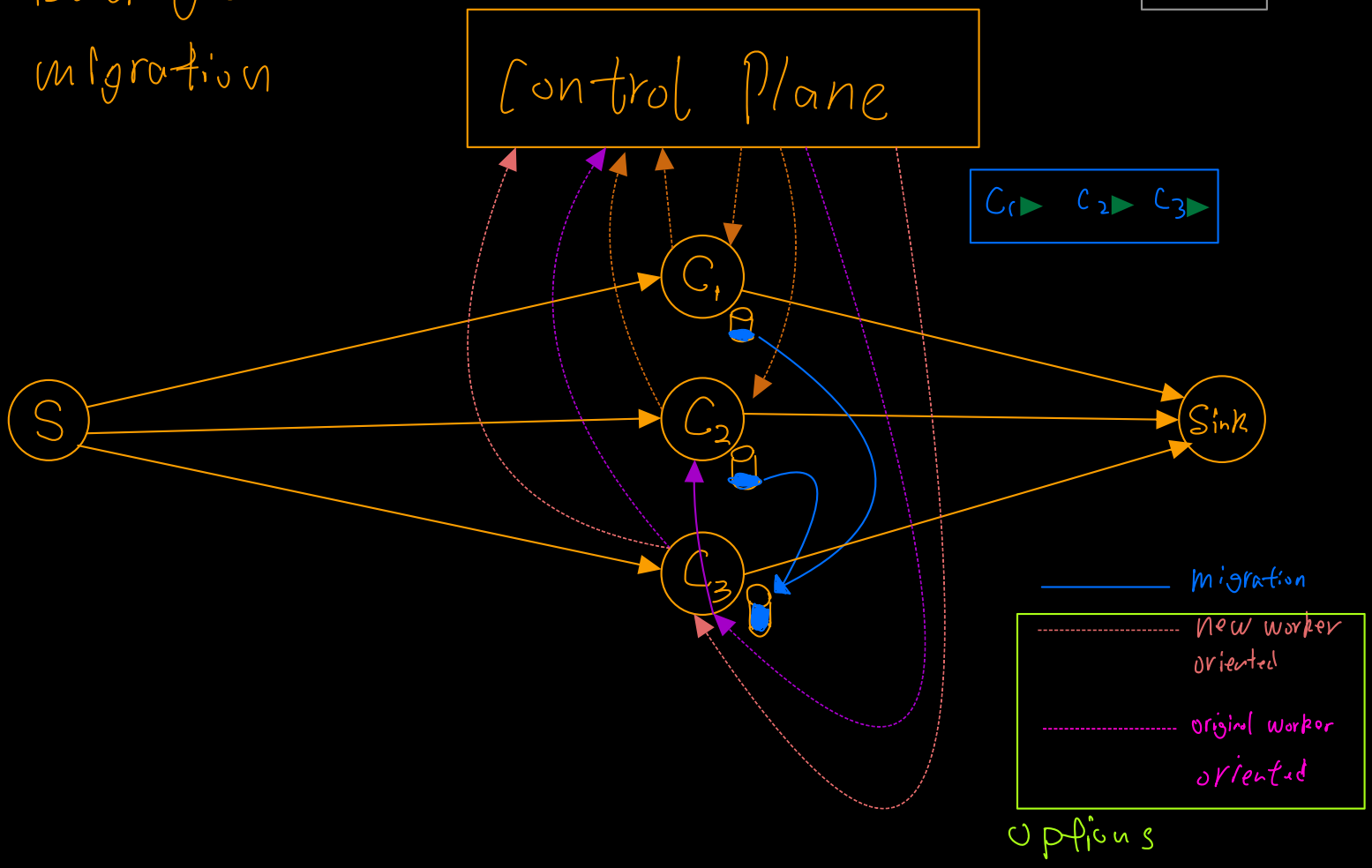
② Add a new Worker



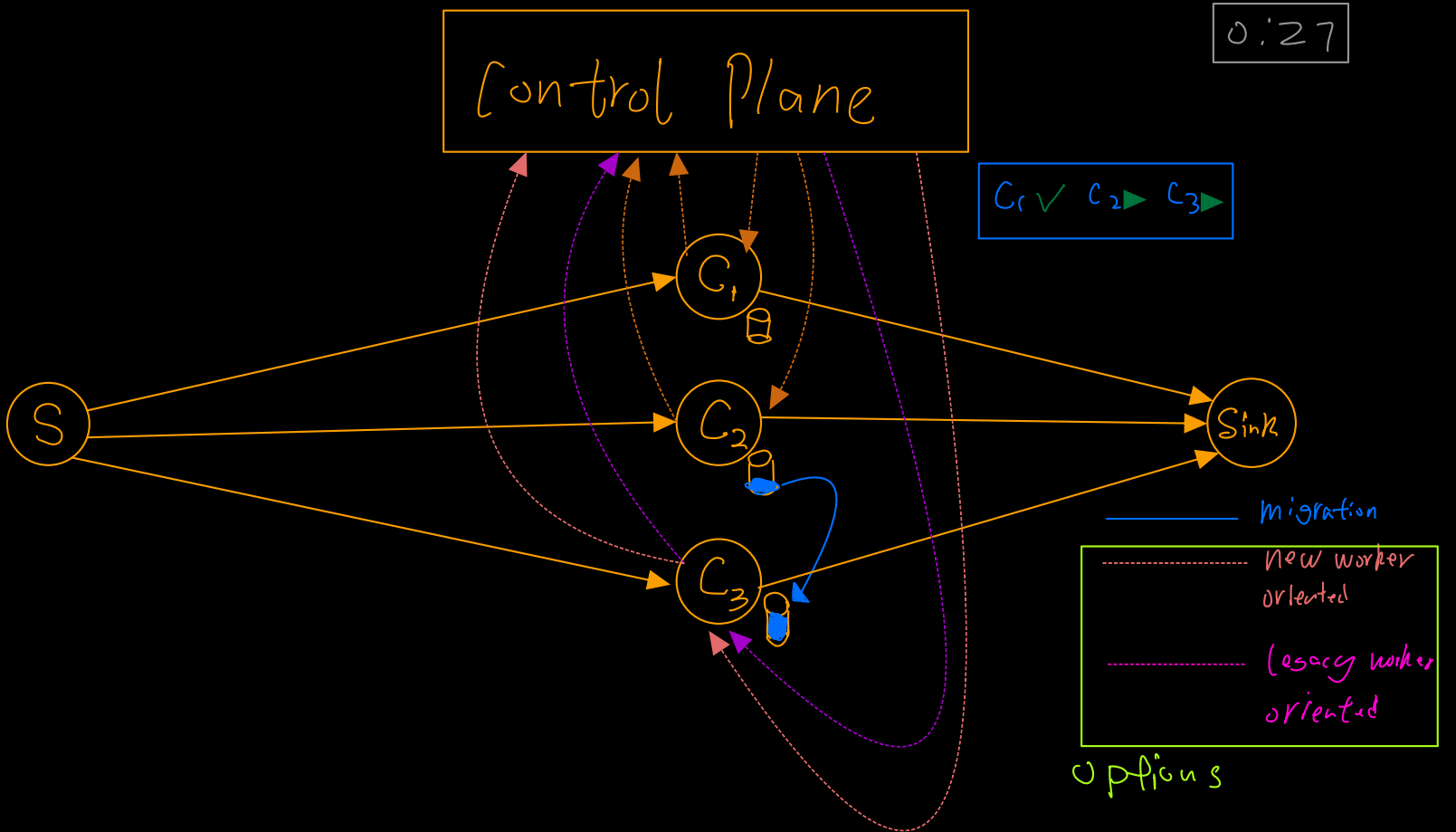

3 update hash function
0:20
Control Plane
C1
C2
S
Sink
C1 : A - P
C2 : Q - Z
C1 : A - P
C2 : Q - K
C3 L - Z
C3
4 update state table / function
0:22
Control Plane
C1
C2
C3
S
Sink
C1 : A - P
C2 : Q - Z
C1 : A - P
C2 : Q - K
C3 L - Z

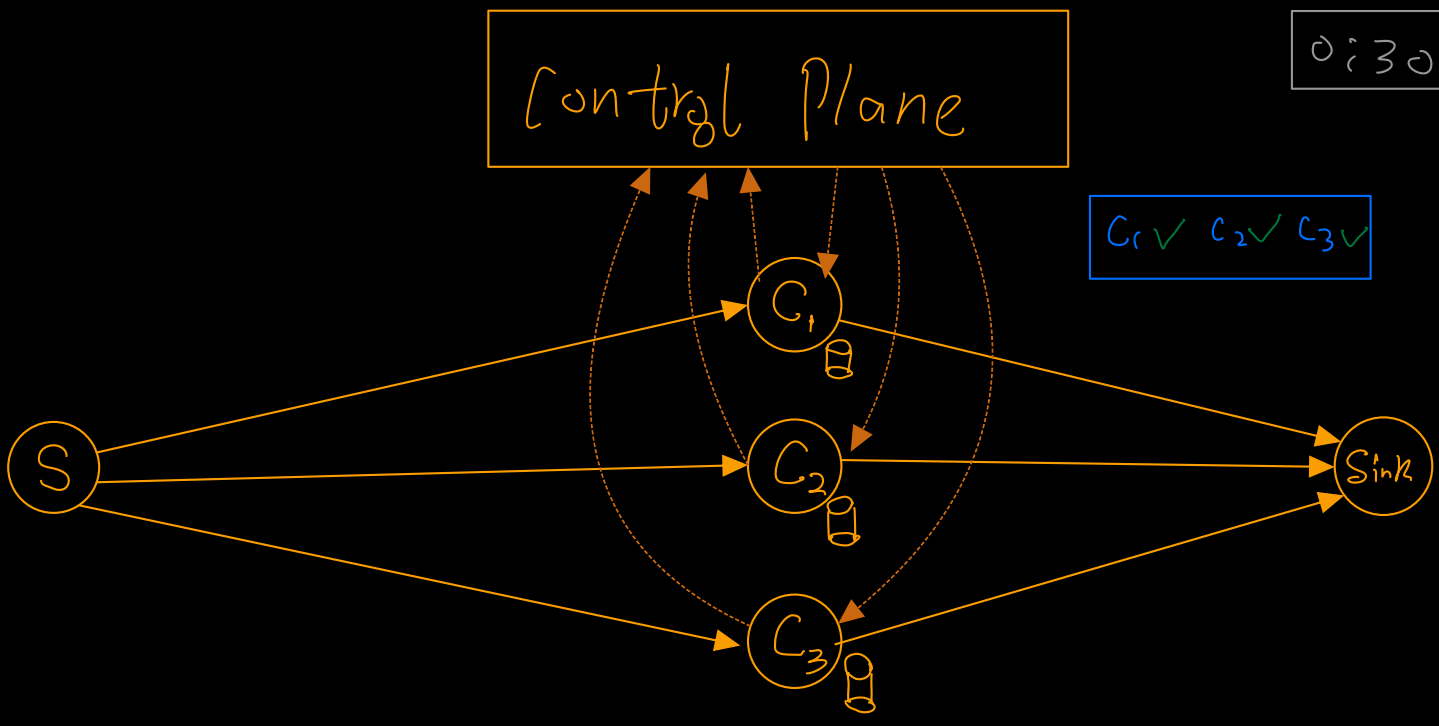
(5) Background migration

0:25



0:27




Control Plane
0:30
C1 ✓ C2 ✓ C3 ✓
S
C1
C2
C3
Sink