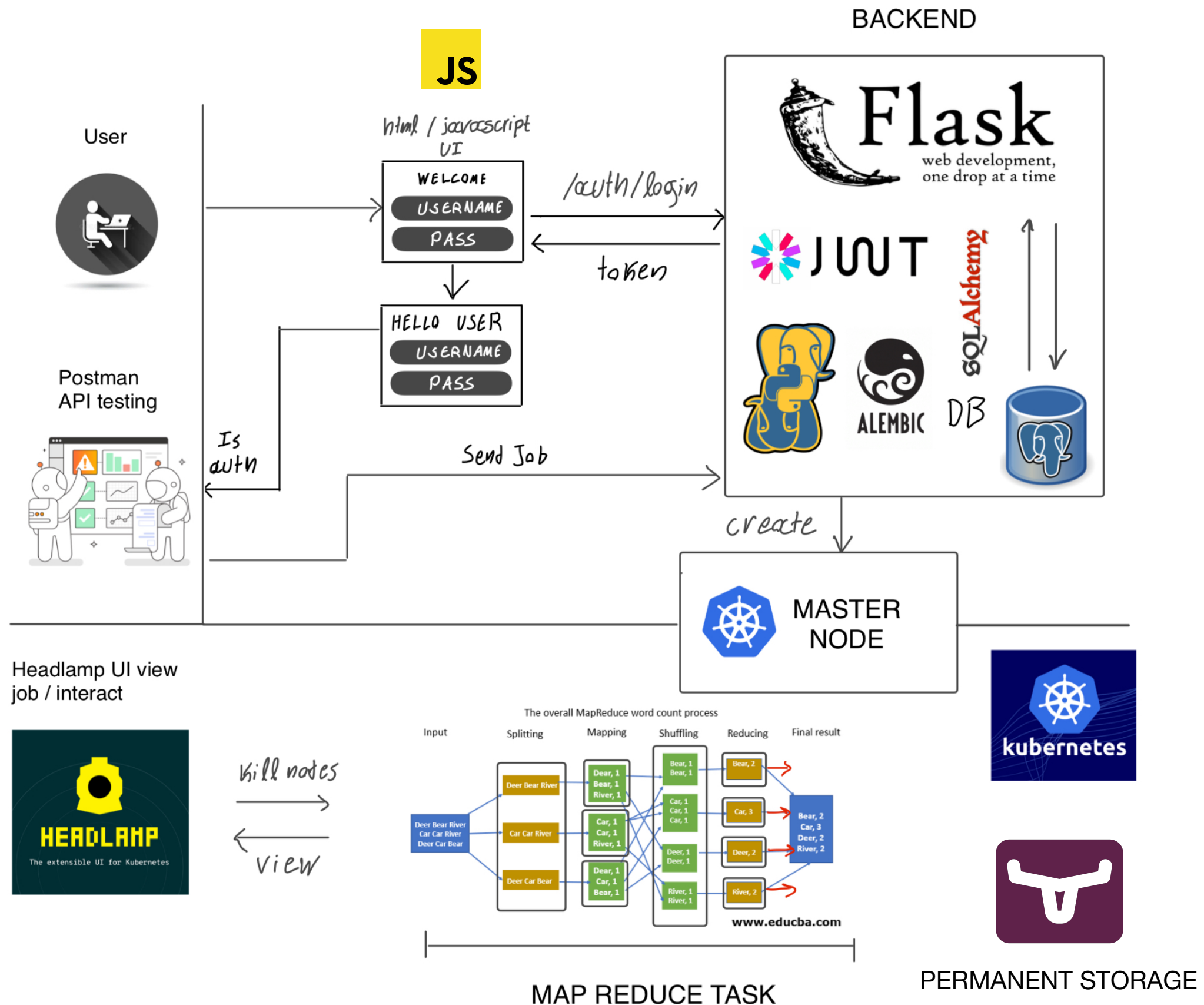


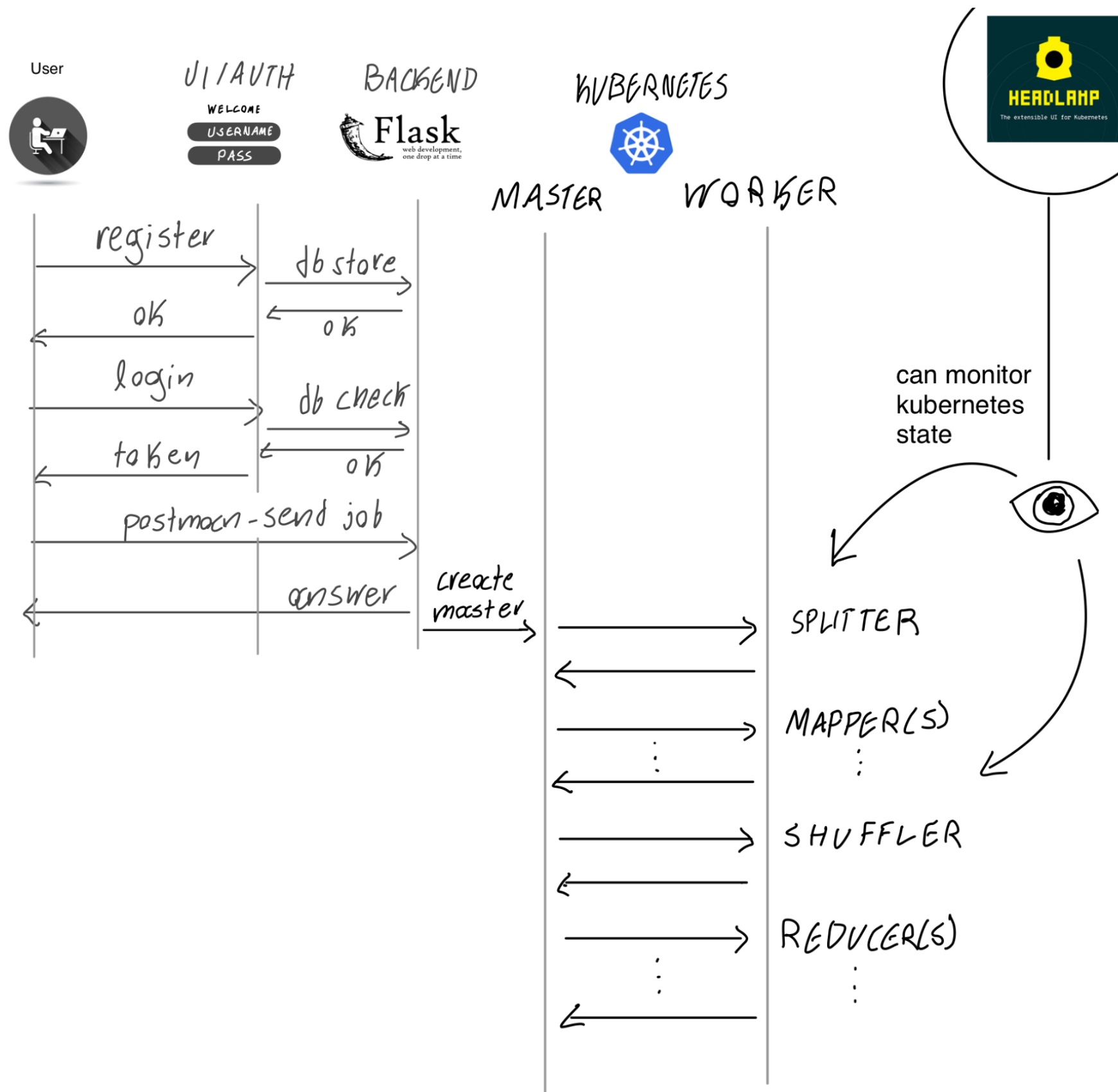
Distributed Systems project

Μπιμπίρης Πέτρος
Σταυγιανουδάκης Γεώργιος
Χριστοφιλογιάννης Ιωάννης

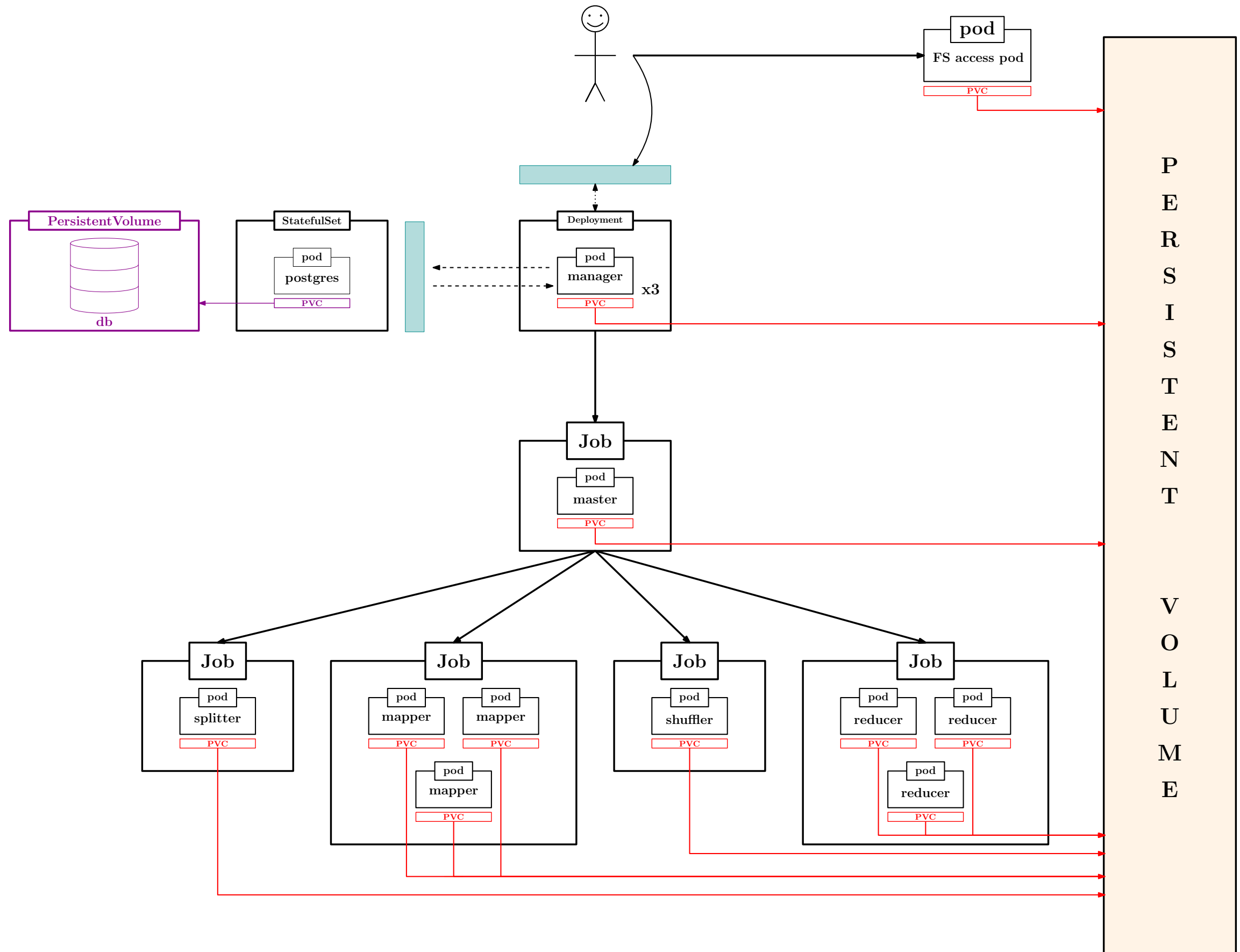
Implementation



Sequence diagram



Cluster Architecture



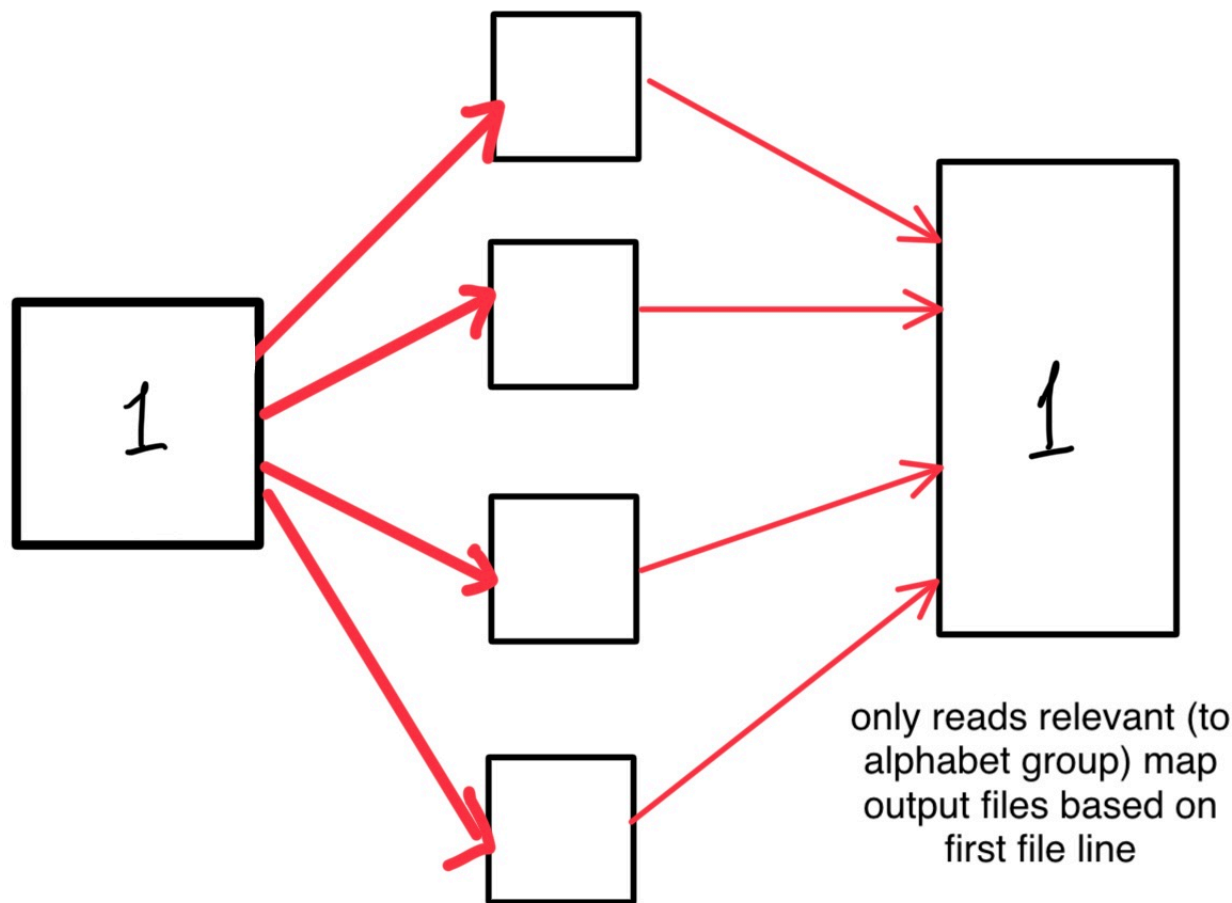
MapReduce

Splitter
out:
split-id.txt

Mappers
out:
mapper-id.txt

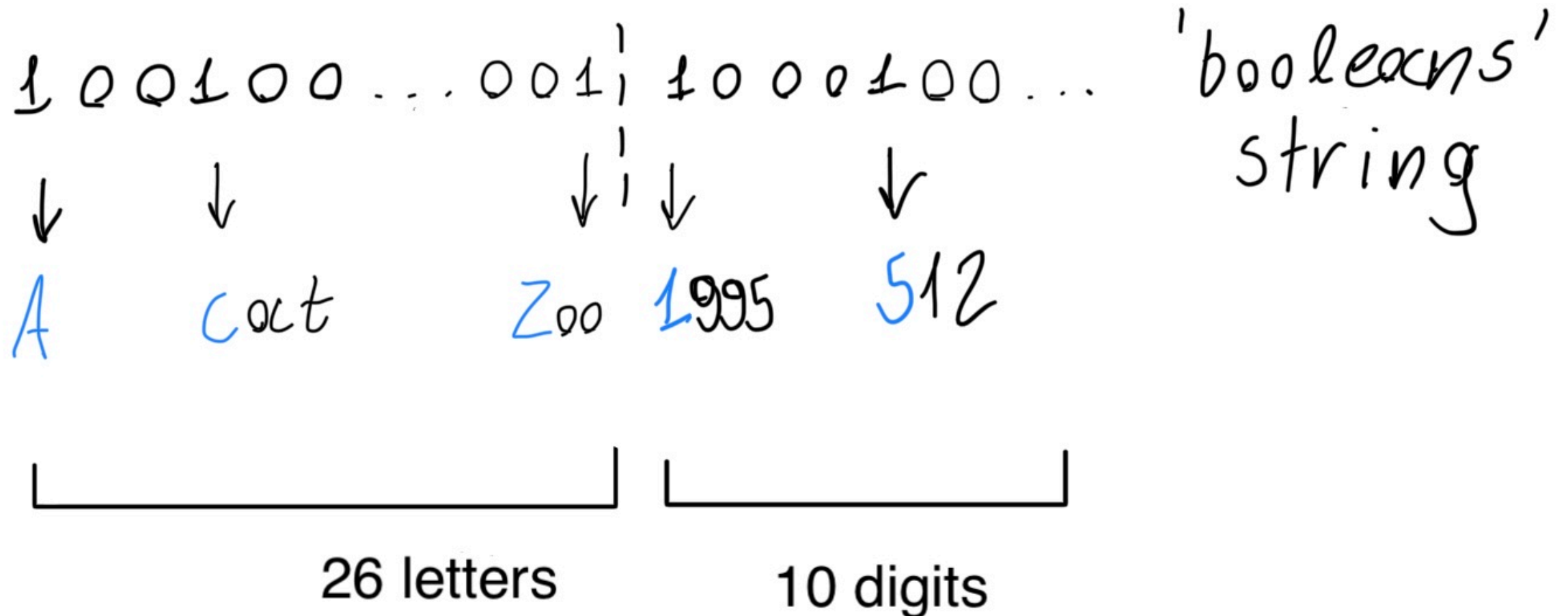
Shuffler
out:
shuffler-id.json

Reducers
out:
reduced-out-id.json



● Save Files to Longhorn

Mappers store metadata:



Now, shuffler knows which files contain data
relevant to the reducers