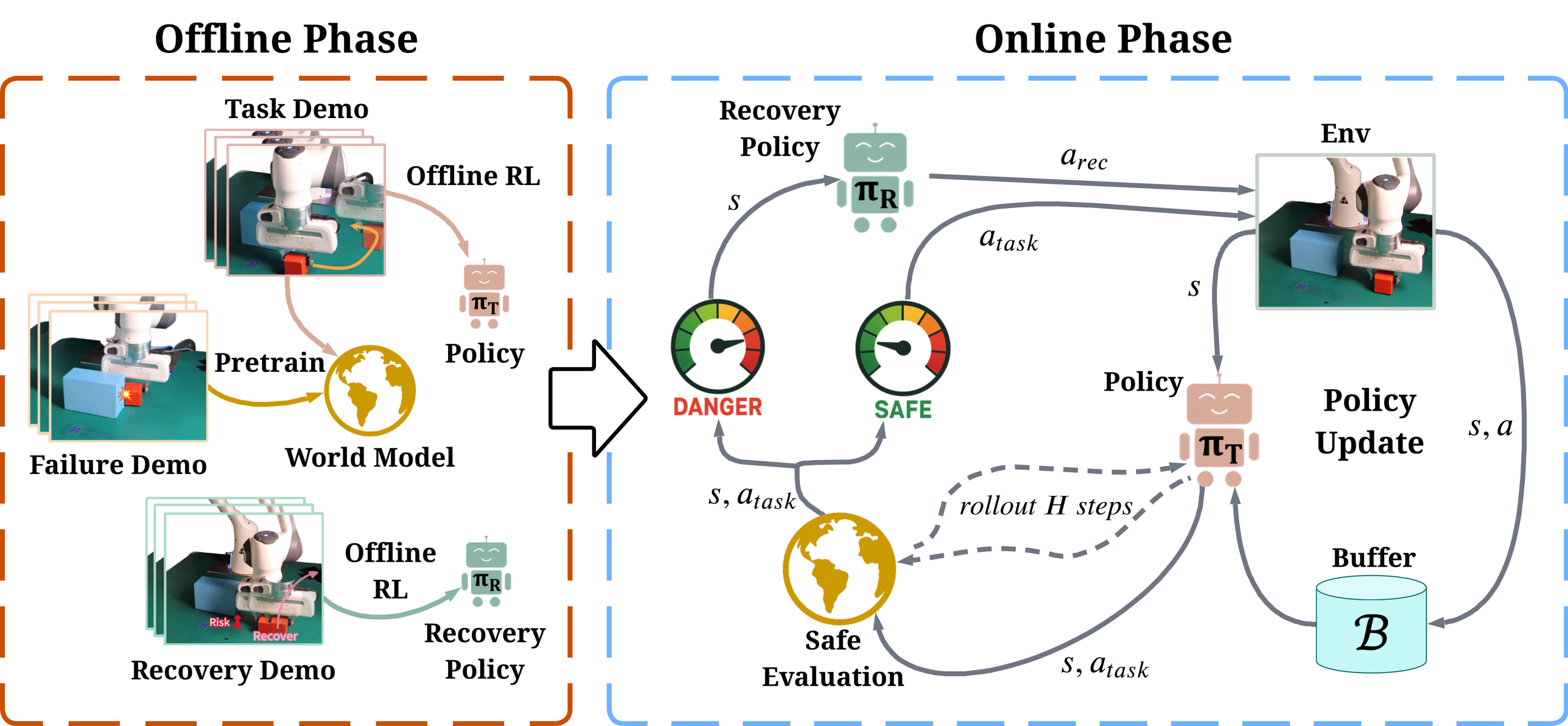
Offline Phase
Online Phase
Task Demo
Offline RL
Policy
Pretrain
Failure Demo
World Model
Risk
Recover
Offline RL
Recovery Demo
Recovery Policy
Recovery Policy
s
a_{rec}
Env
a_{task}
s
DANGER
SAFE
Policy
Policy Update
s, a
s, a_{task}
rollout H steps
Buffer
Safe Evaluation
s, a_{task}