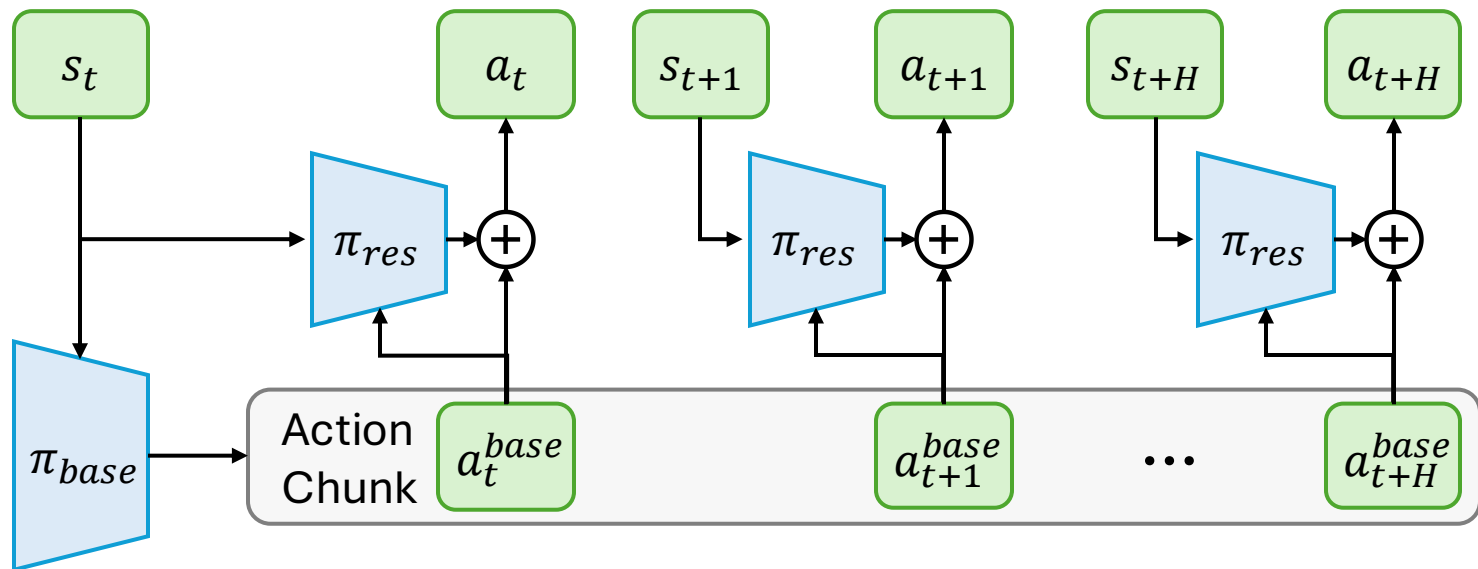


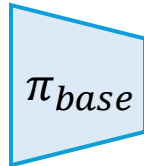
Inference



Offline Buffer



Train

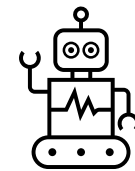


Phase 2: RL

Online Buffer



Train



Rollout