

# ReCA: Integrated Acceleration for Real-Time and Efficient Cooperative Embodied Autonomous Agents

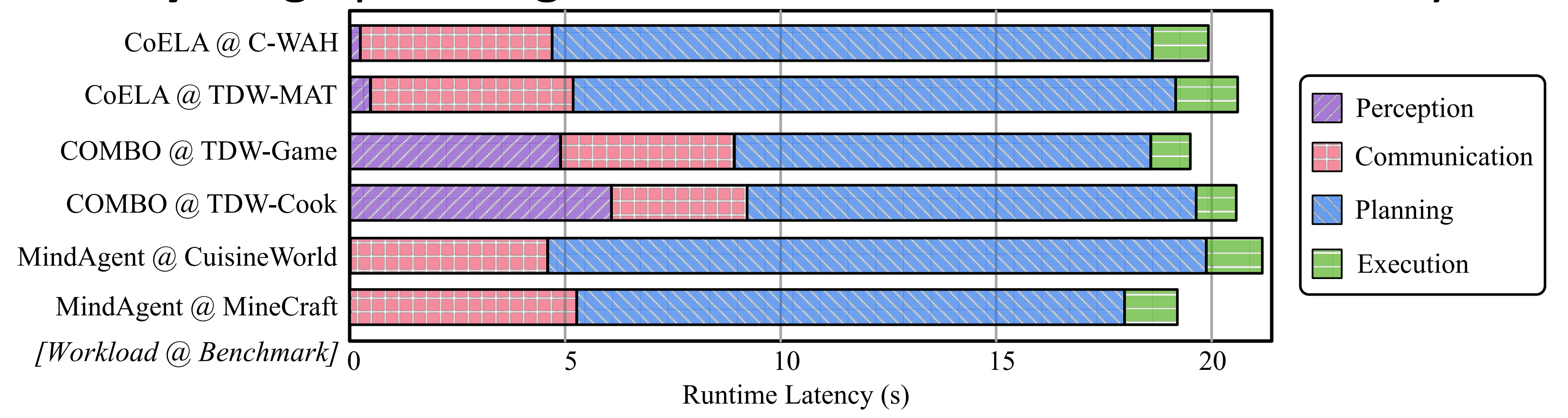
Zishen Wan<sup>1</sup>, Yuhang Du<sup>2</sup>, Mohamed Ibrahim<sup>1</sup>, Jiayi Qian<sup>1</sup>, Jason Jabbour<sup>3</sup>, Yang (Katie) Zhao<sup>2</sup>, Tushar Krishna<sup>1</sup>, Arijit Raychowdhury<sup>1</sup>, Vijay Janapa Reddi<sup>3</sup>  
<sup>1</sup>Georgia Tech, GA <sup>2</sup>University of Minnesota, Twin Cities, MN <sup>3</sup>Harvard University, MA

## COOPERATIVE EMBODIED AI AGENT SYSTEMS

## SYSTEM CHARACTERISTICS AND CHALLENGES

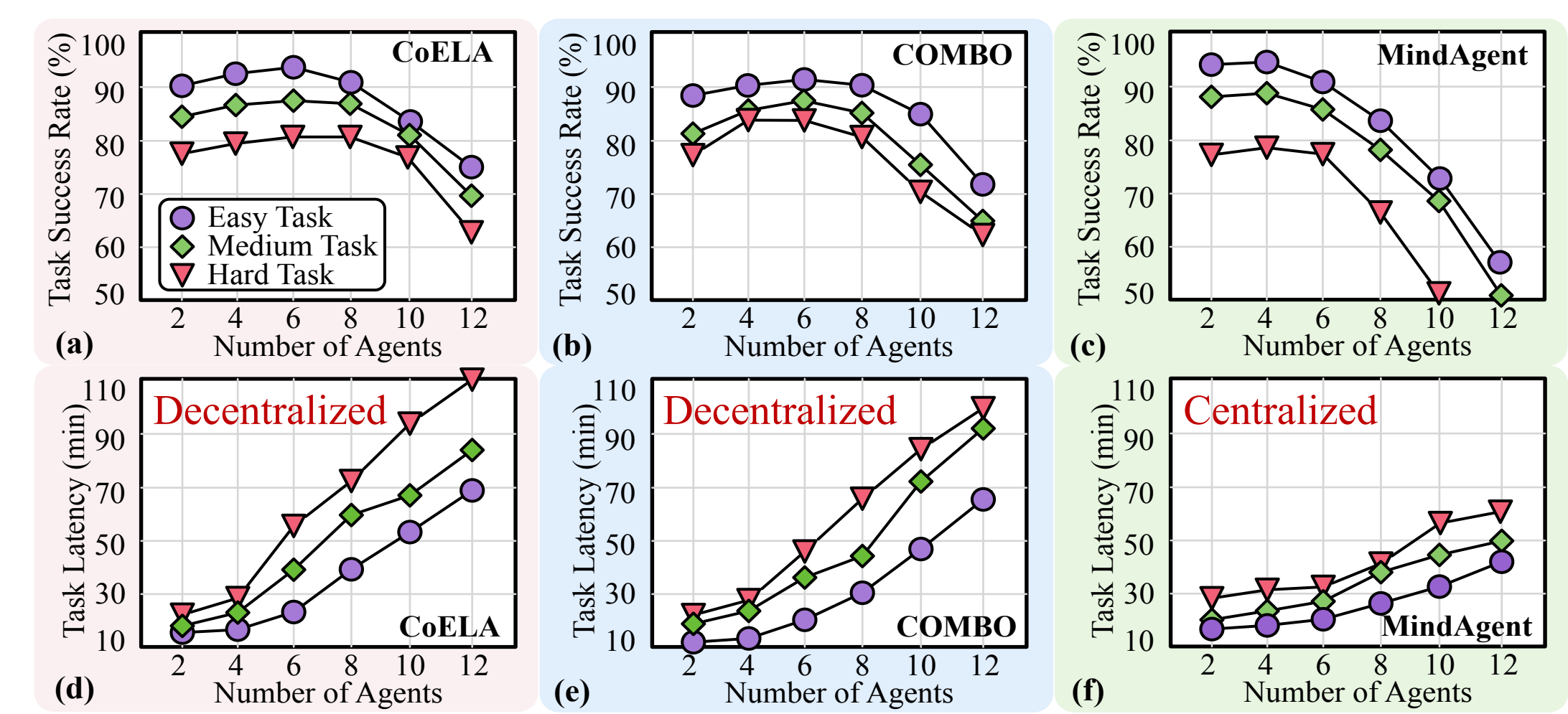
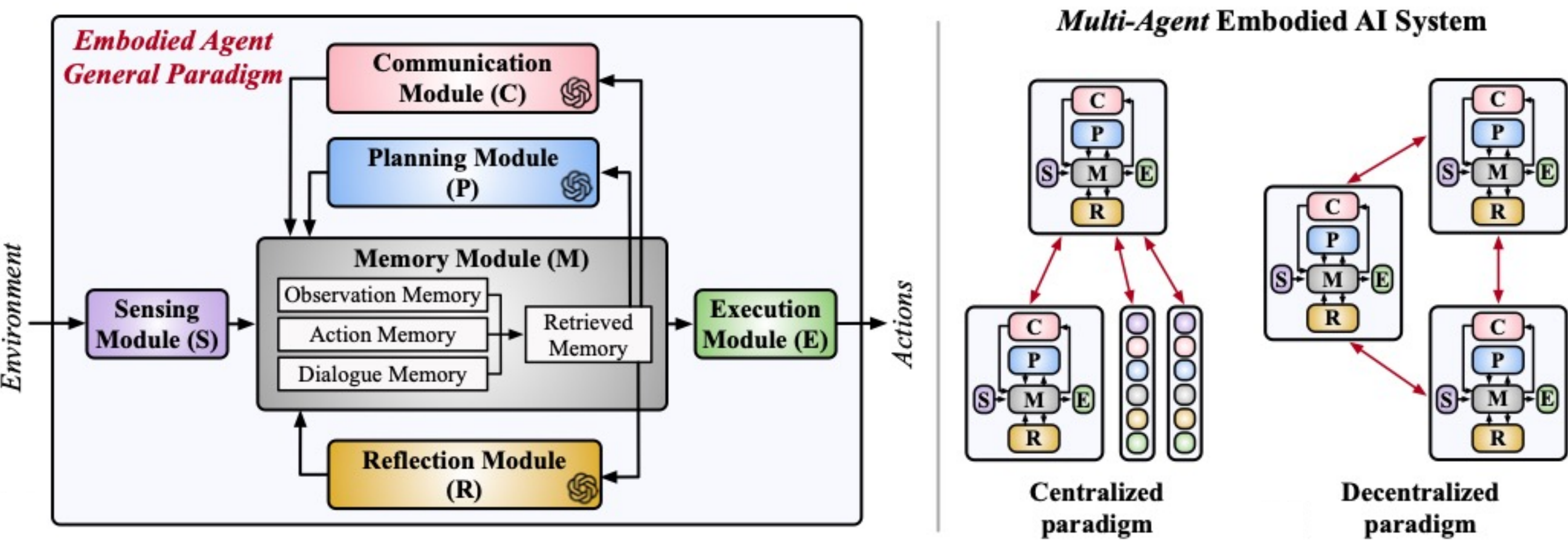
- ❖ **Task:** Long-horizon multi-objective task & planning
- ✓ **Applications:** complex household task, object transport, etc.

- ❖ **System Challenges:** Latency, scalability, module sensitivity
- ✓ **Latency:** large planning and communication runtime latency.



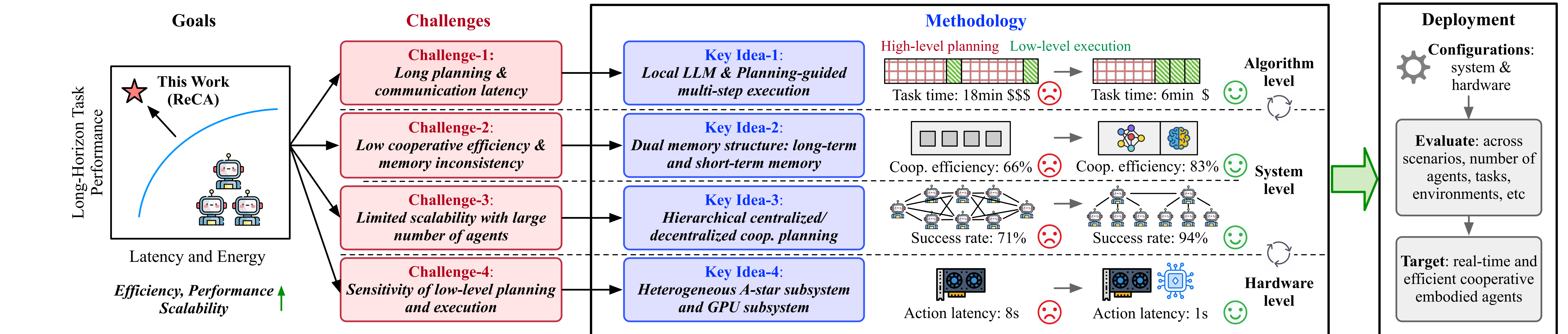
- ✓ **Modular framework** for embodied AI agents.

- ✓ **Scalability:** Decentralized EAI systems suffer from exploded latency; Centralized EAI systems suffer from reduced task success rate.



- ✓ **Module Sensitivity:** planning and execution modules are critical.

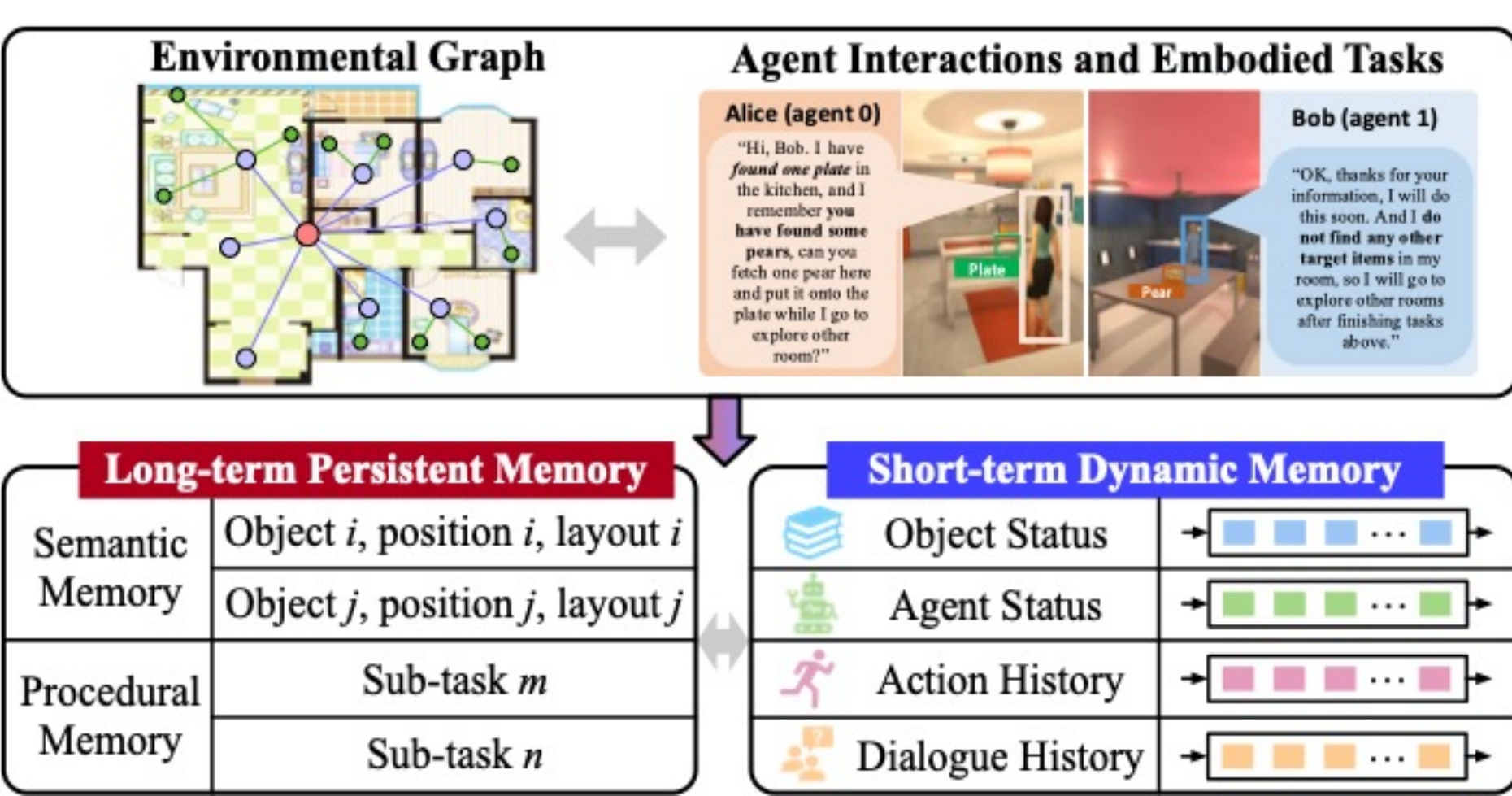
## RECA: ACCELERATION FRAMEWORK FOR COOPERATIVE EMBODIED AI AGENTIC SYSTEMS



## SYSTEM-LEVEL OPTIMIZATION

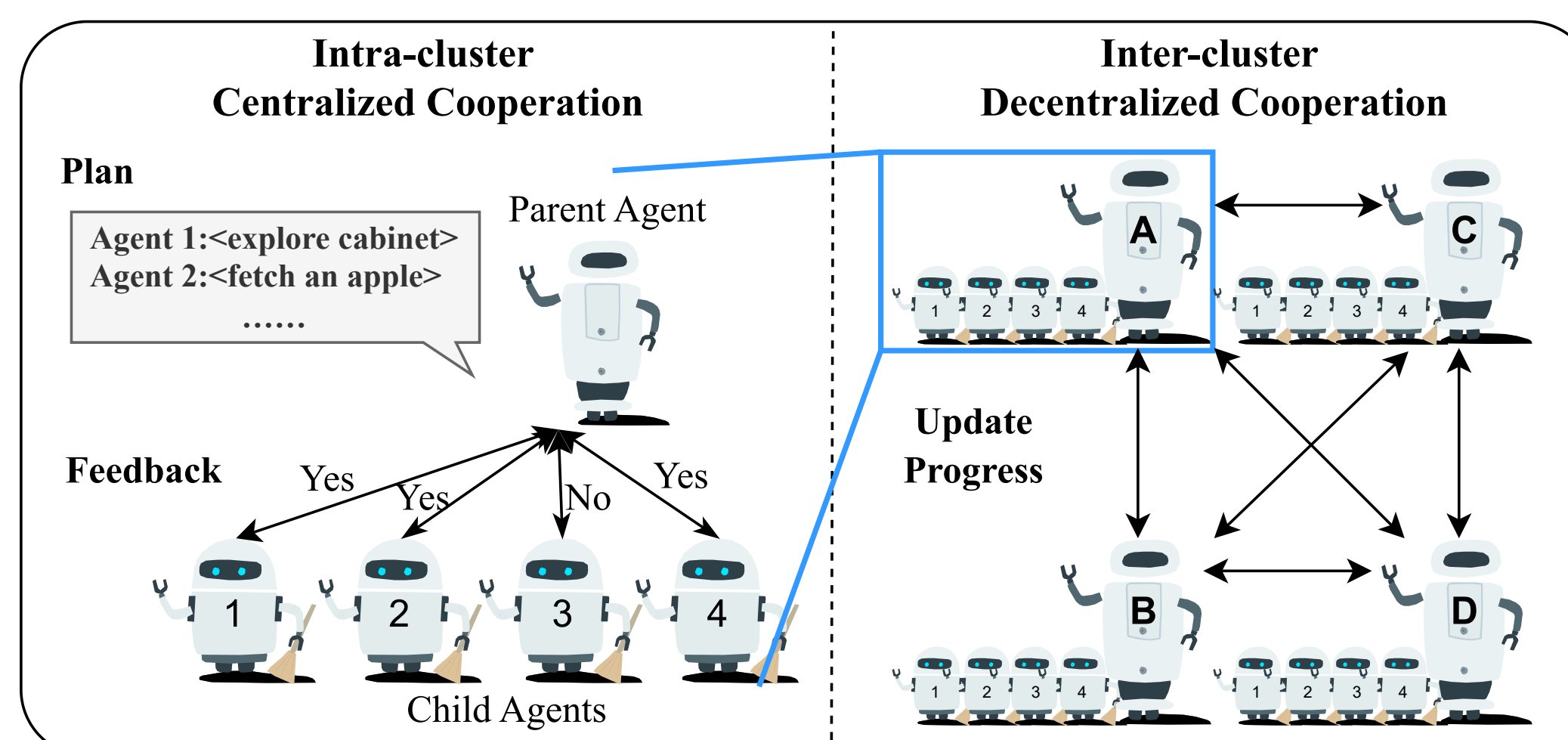
### ❖ Dual-Memory Structure

- ✓ Long-term & short-term memory



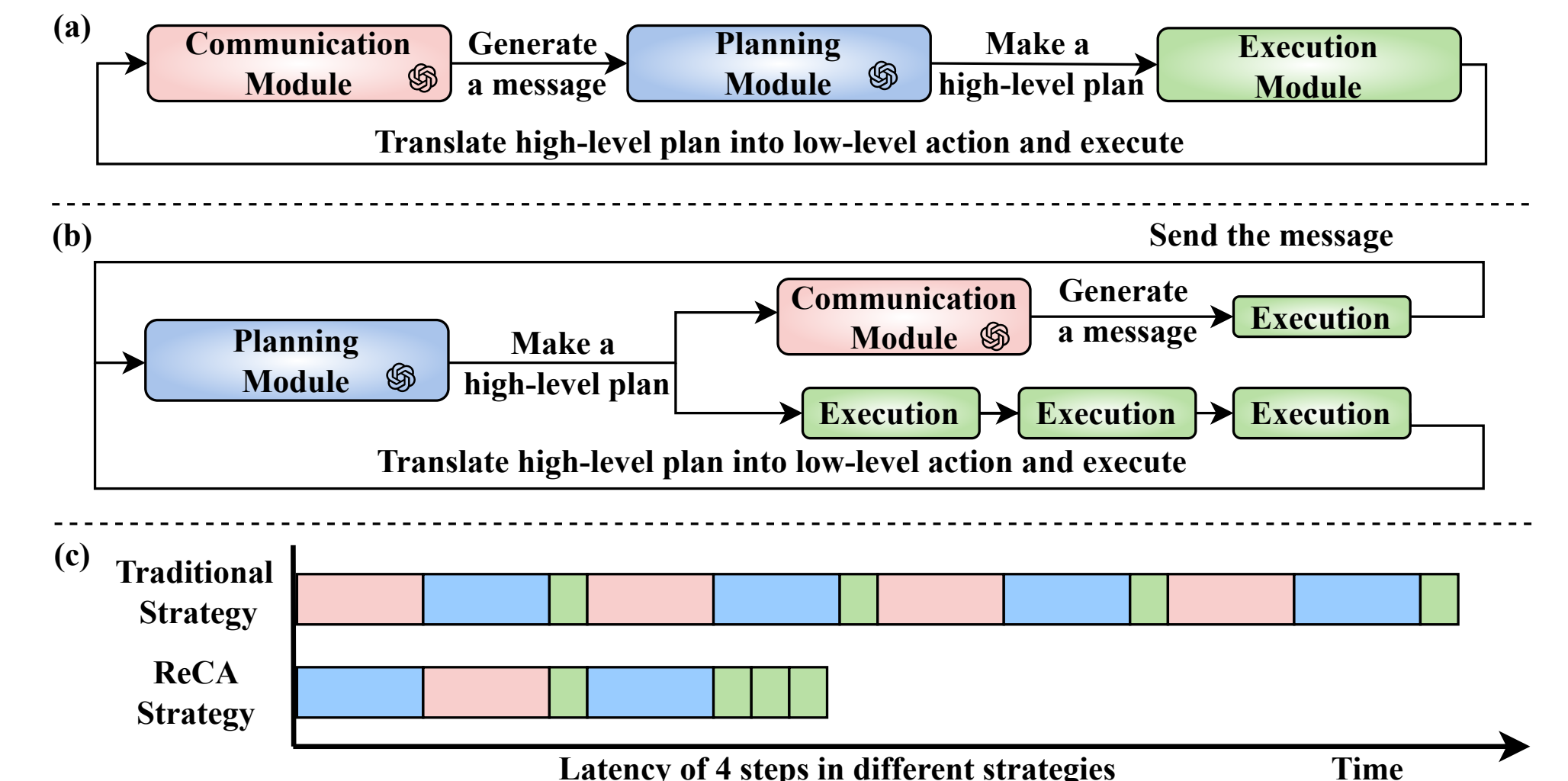
### ❖ Hierarchical Cooperative Planning

- ✓ Inter-cluster central & inter-cluster decentral



### ❖ Planning-Guided Efficient Execution

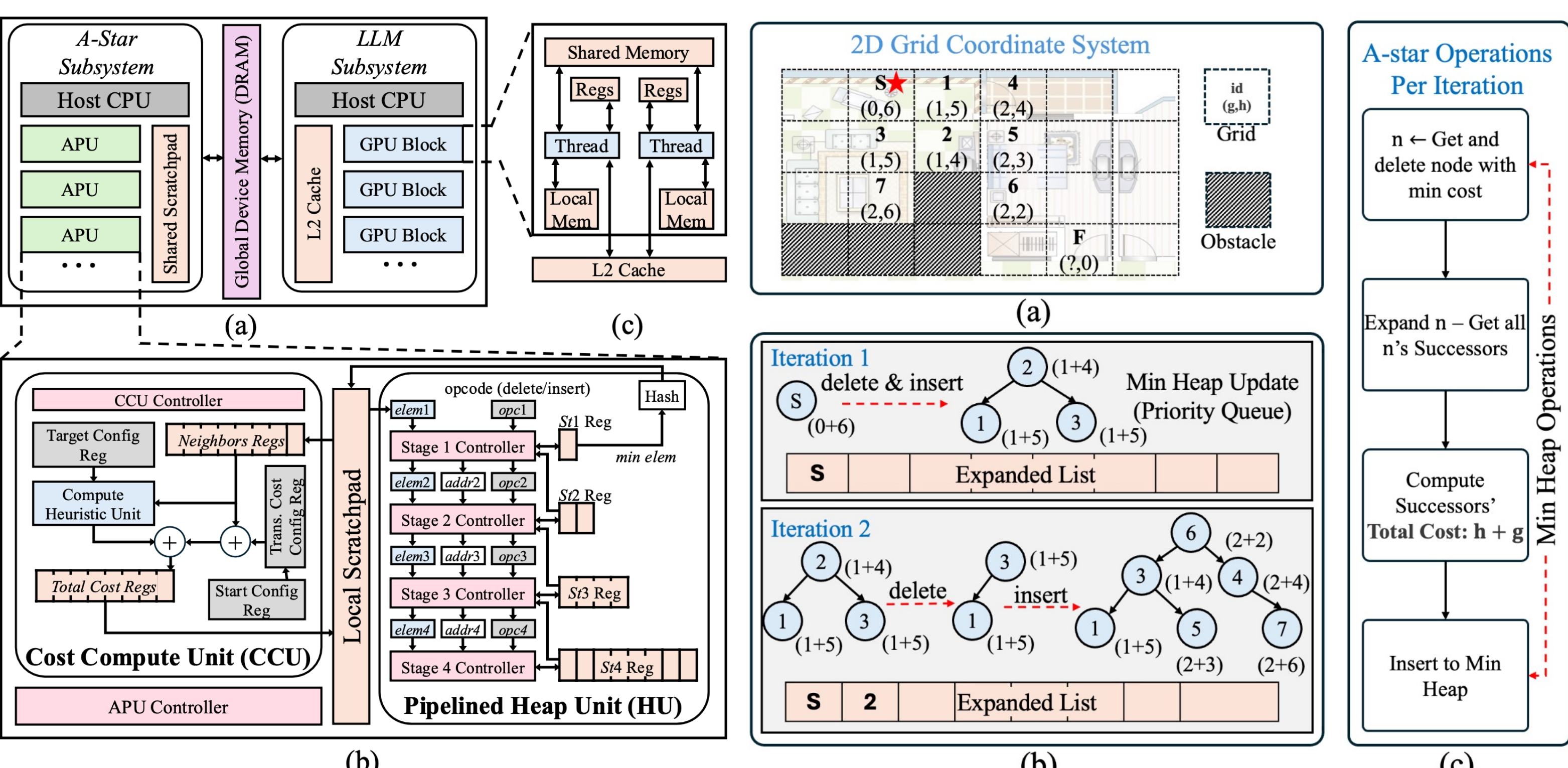
- ✓ Plan-then-comm; plan-guided multi-step exe



## HARDWARE-LEVEL OPTIMIZATION

### ❖ Heterogeneous Hardware System

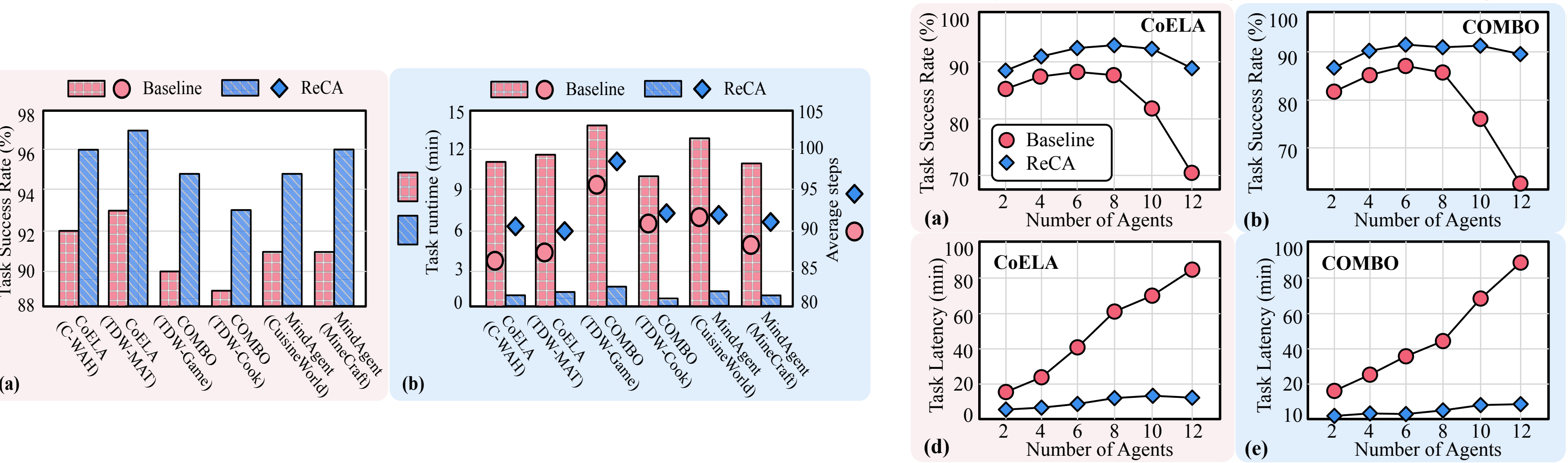
- ✓ LLM subsystem (high-level plan); A\* subsystem (low-level plan)



## EVALUATION RESULTS

### ❖ Efficiency-Performance-Scalability Improvement

- ✓ Task success rate and runtime
- ✓ Scalability under large num of agents



## ACKNOWLEDGMENTS

This work was supported in part by CoCoSys, one of seven centers in JUMP 2.0, a SRC program sponsored by DARPA, and NSF GRFP DGE-214074.