

AIOpsLab: A Holistic Framework to Design, Develop, and Evaluate AI Agents for Enabling Autonomous Clouds

Yinfang Chen, Manish Shetty, **Gagan Somashekar**, Minghua Ma, Yogesh Simmhan, Jonathan Mace, Chetan Bansal, Rujia Wang, Saravan Rajmohan



Cloud Incidents Are Inevitable and Costly



Technology

Amazon cloud services back up outage hits thousands of users

Major Outage across ChatGPT and API

Incident Report for OpenAI

Resolved

Between 5:42AM - 7:16AM PT we saw errors in our services. We identified the problem and implemented a fix. Users are now seeing normal responses from our services.

Alibaba Outage Caused by Cooling Unit

Cloud giant Alibaba's outage on Sunday was caused by a malfunctioning refrigeration unit, say company officials.

Facebook, Instagram Outage Affects Thousands Of Users As Meta Suffers 'Major Disruptions'



Amazon's cloud issues highlight backbones of modern internet

SHARE & SAVE



Amazon's cloud issues highlight backbones of modern internet

Tuesday's incident mostly affected the eastern U.S. and impacted everything from airline reservations and car dealerships to Amazon's own e-commerce operation.

Cloud Incidents Are Inevitable and Costly

- Production incidents cause
 - **Loss of revenue** due to SLA violations
 - **User dissatisfaction**
 - **Loss of productivity** of on-call engineers
- Majority of incidents (60%) were related to **cloud operations**
- Detecting, localizing, and mitigating them is **tedious** and requires **significant human effort**

Agents and Benchmarks

InterCode

*Interactive coding w/
execution feedback*

- Bash
- Python
- SQL

SWE Bench

- Solve GitHub issues
- Bug fixing
- Feature additions

Agents

SWE-Agent (Princeton)
AutoCodeRover (NUS)
Aider (Aider)
MASAI (MSR India)
OpenDevin (CMU)

Jul 3, 2023

Aug 7, 2023

Oct 10, 2023

Mar 12, 2024

Apr-Jul, 2024

???

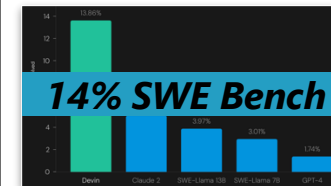
AgentBench

8 Environments

- OS tasks
- DB tasks
- Webpage navigation tasks
- ...

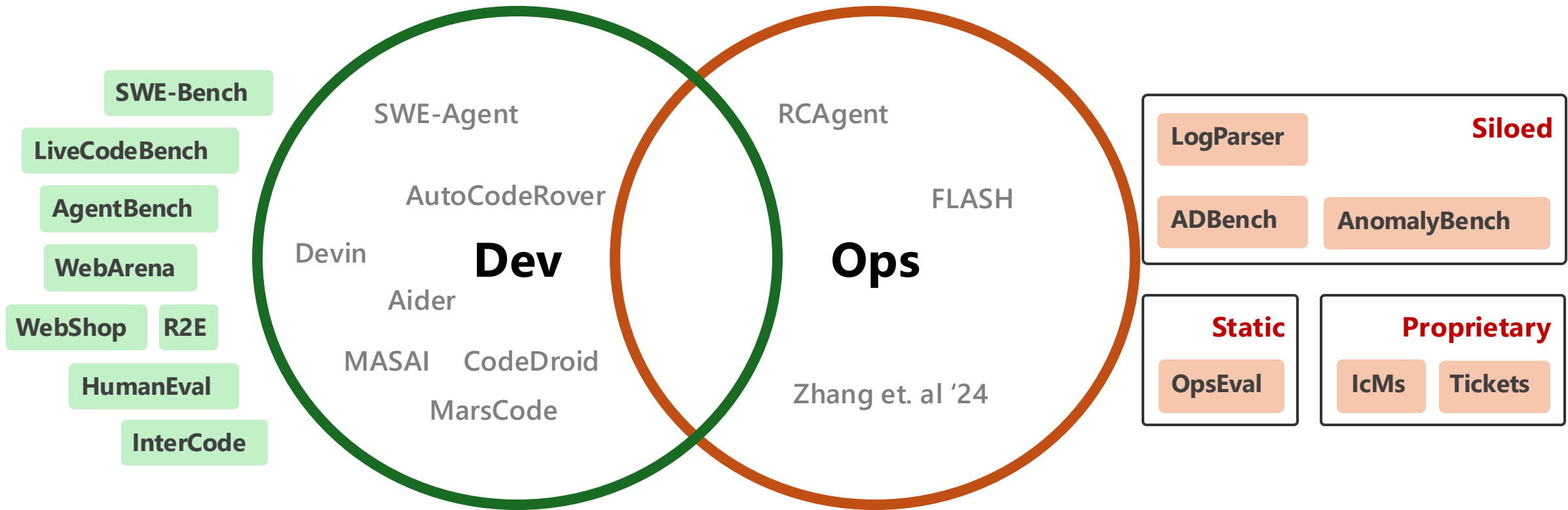
Devin

AI Developer



Dev 
Ops 

Agents: A DevOps View



AIOpsLab

A Holistic Framework to Design, Develop, and Evaluate AI Agents for Enabling Autonomous Clouds

Framework Goals



Standardize

tasks / metrics / baselines

Realistic

live real workloads and actions

Scalable

apps / tasks / faults / agents

Reliable

observable / x-lifecycle / reproducible

First Idea
(but does not work)

How about “record-and-replaying” previous cloud incidents from the incident database?

Challenges & Motivation

- **Diverse** cloud scenarios
 - Dynamic user workload and update
 - Different architectures of the cloud, e.g., monolithic, microservices, serverless etc.
- **Proprietary** cloud incident/troubleshooting data
 - Existing AIOps benchmarks: Static, text-based QA
- **Large action space** for more challenging **site reliability engineering** tasks

CONFIDENTIAL

Dev 

Ops (SRE) 

Contributions

- **Diverse** SRE scenarios
- **Proprietary** cloud incident data
- **Large action space** for **SRE** tasks

Interfaces for developers to facilitate building their scenarios

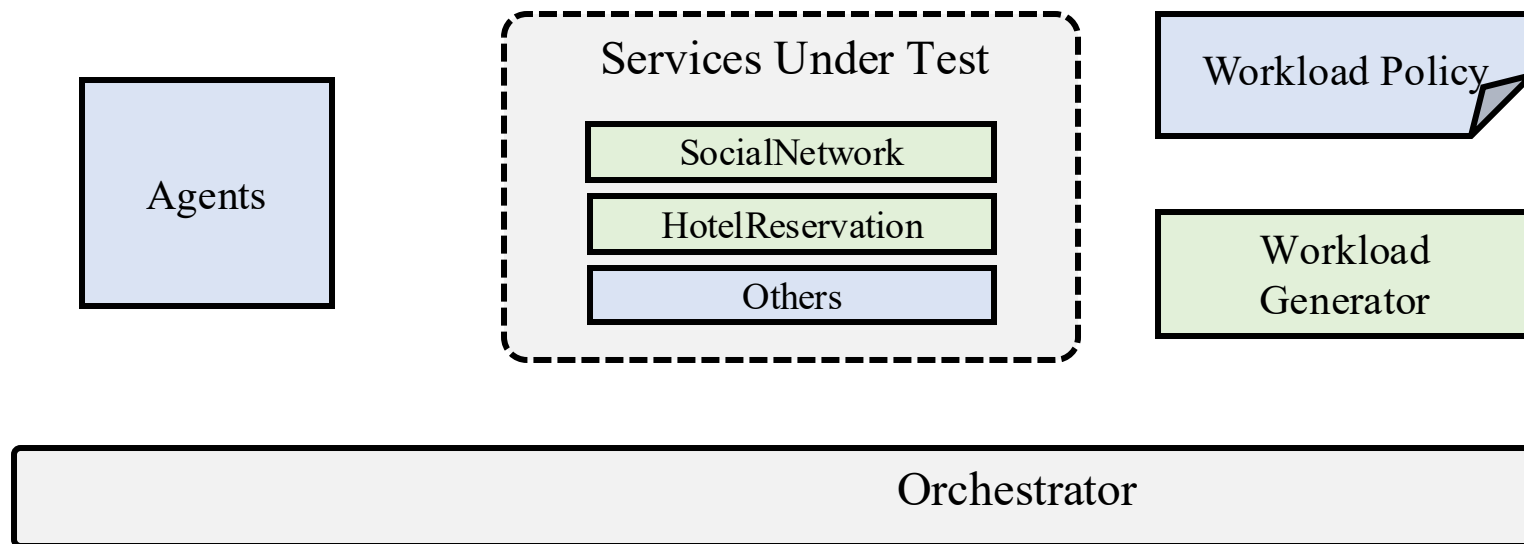
A live **open-sourced** benchmark & framework for cloud operation (SRE)

- **Interfaces** for agents to interact with the cloud (*AgentOps*)
 - Agent-Cloud Interfaces (ACIs)
- **SRE-domain** tasks for agents

~~CONFIDENTIAL~~

Ops (SRE)  

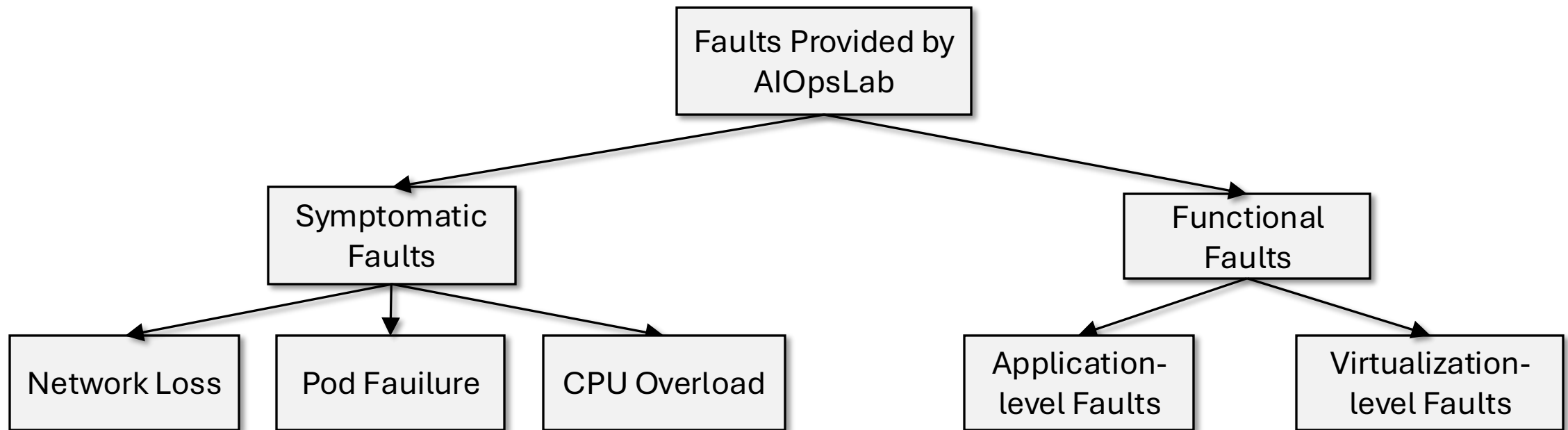

AIOpsLab



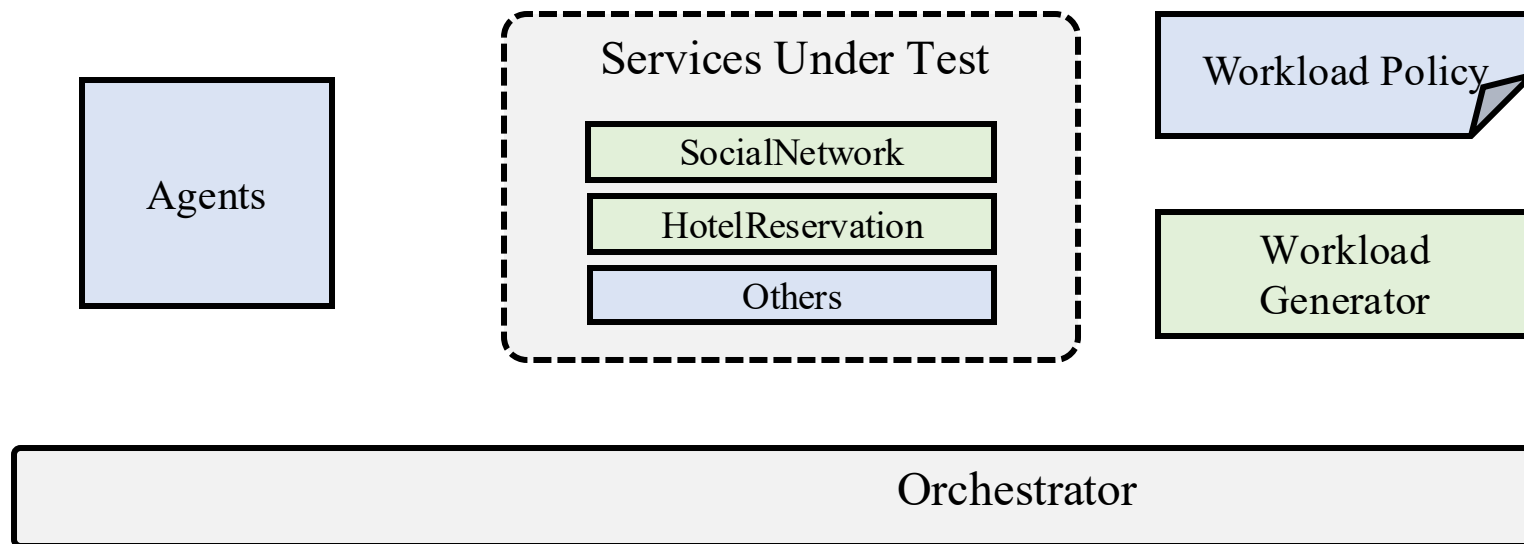
Agent Goals: Task Types

	Task	Description
1	Detection	Can the approach accurately detect anomalies or deviations?
2	Localization	Can the approach pinpoint a fault's exact source?
3	Root Cause Analysis (RCA)	Can the approach categorize the underlying cause of the fault?
4	Mitigation	Can the approach give effective mitigation?

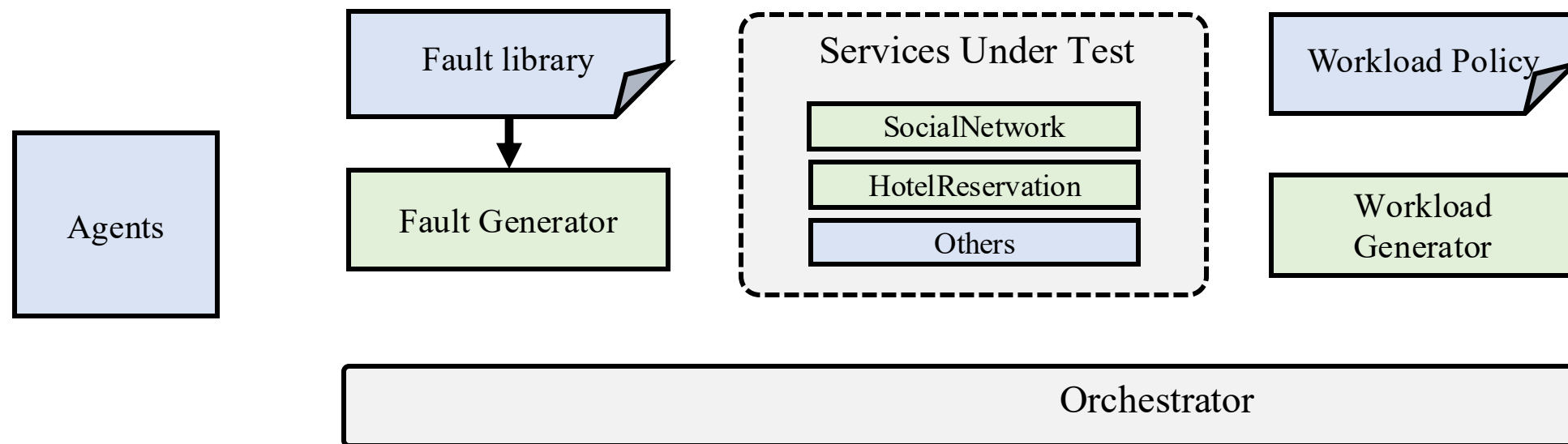
Fault Injection: Taxonomy



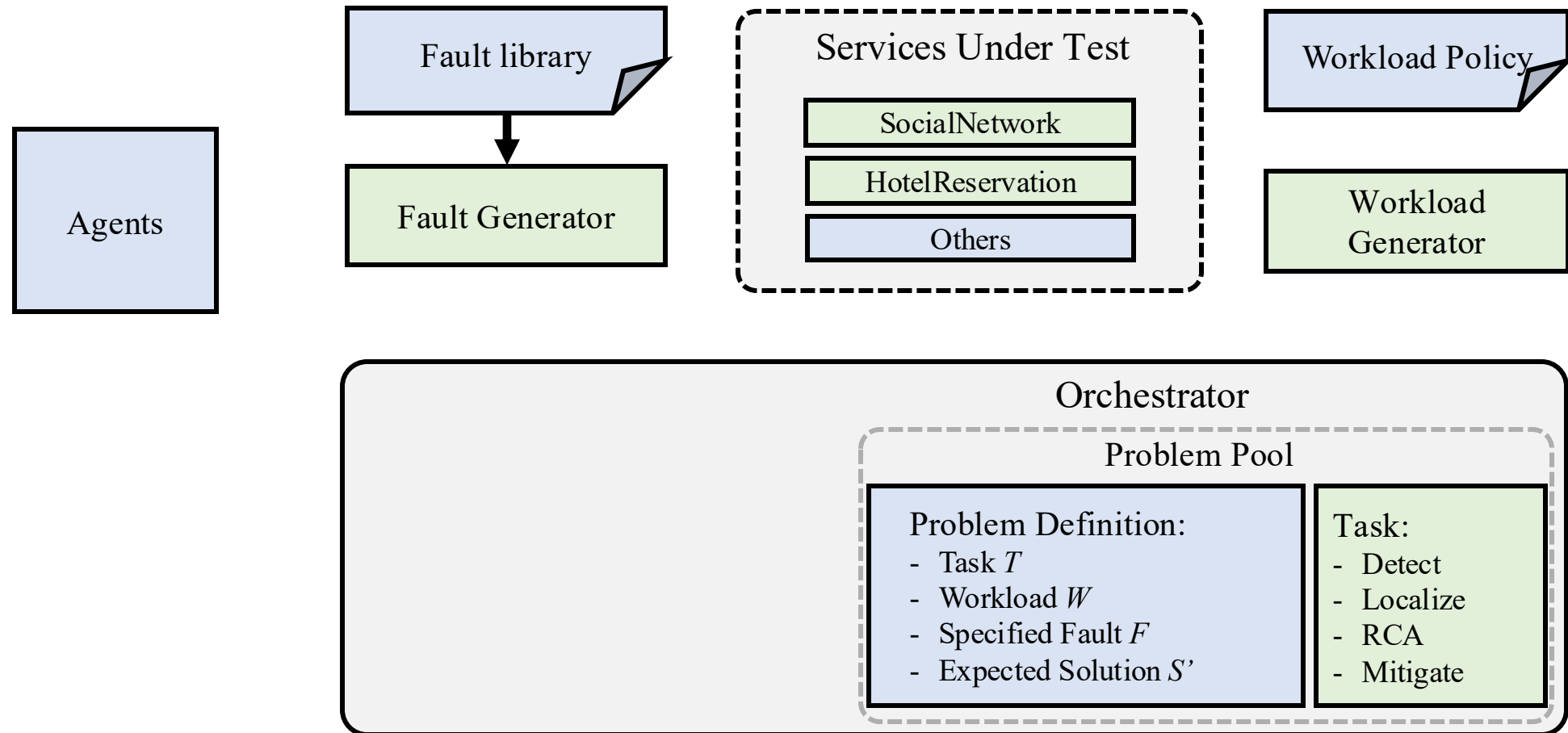
AIOpsLab



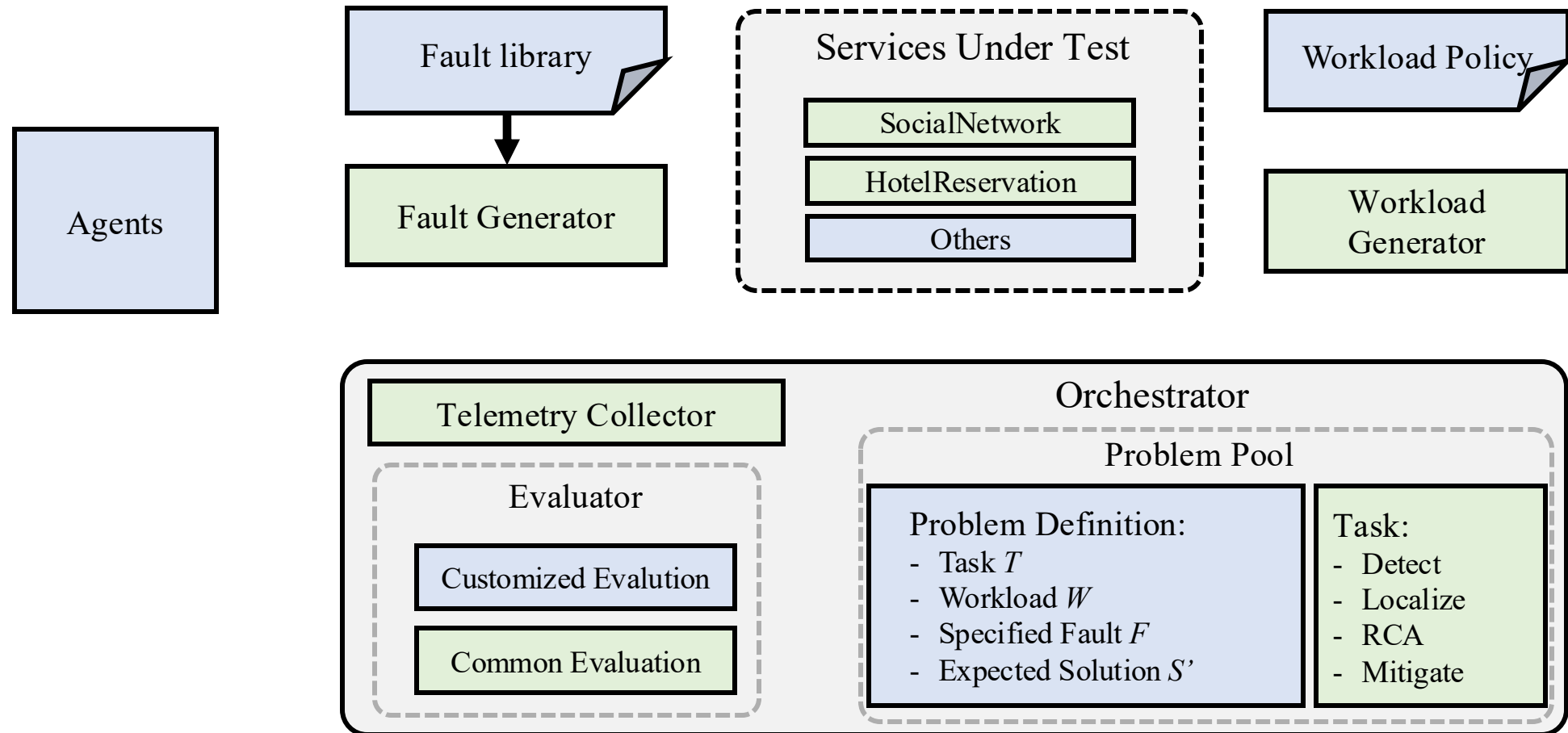
AIOpsLab



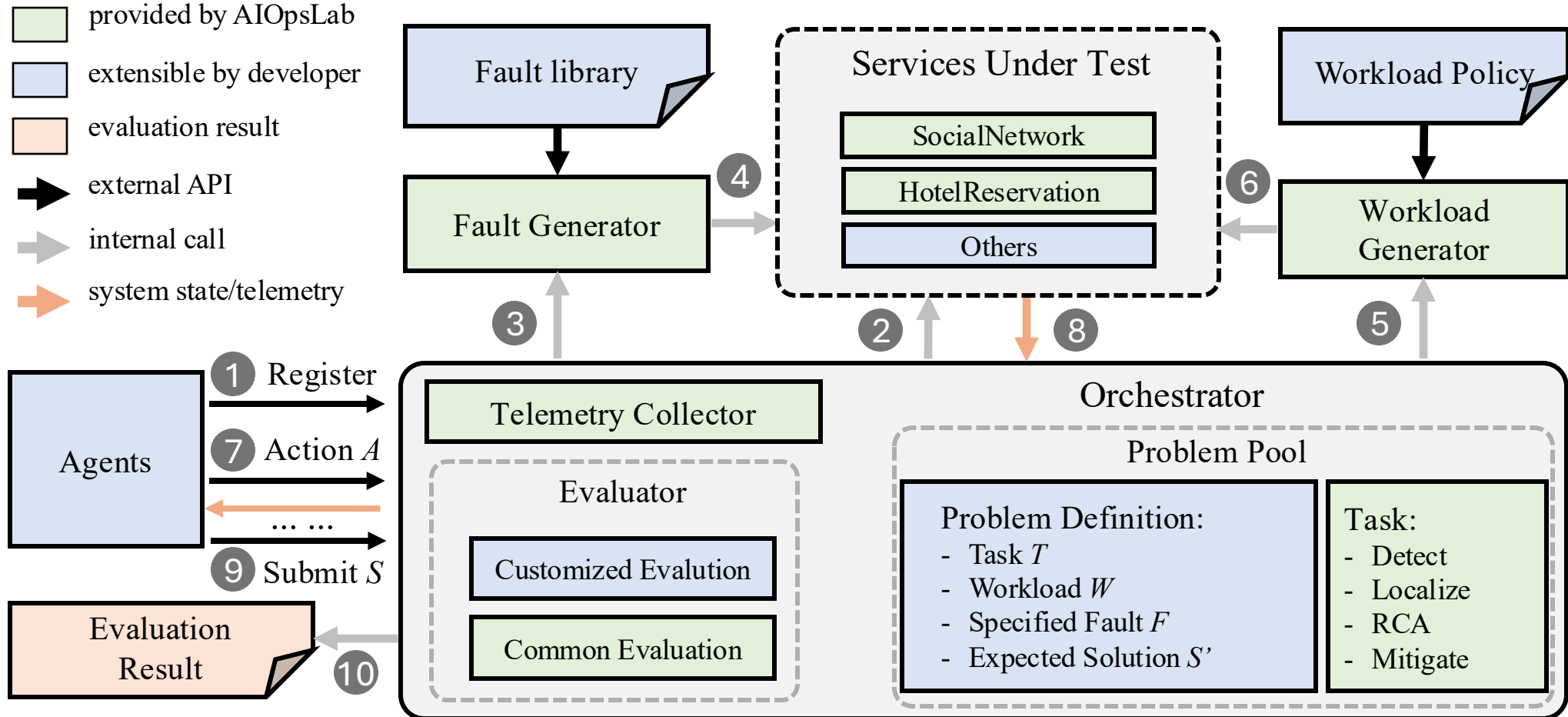
AIOpsLab



AIOpsLab



AIOpsLab



A *problem* agent needs to solve

- “Problem”: one evaluation scenario that agent should complete:
 - *Task*
 - *Context*
 - *Expected Solution* (oracle)

```
from aiopslab import LocalizationTask, SocialNetwork
from aiopslab import Wrk, VirtFaultInjector

class K8STargetPortMisconf(LocalizationTask):
    def __init__(self):
        self.app = SocialNetwork()
        self.ans = "user-service"

    def start_workload(self):
        wrk = Wrk(rate=100, duration=10)
        wrk.start_workload(url=self.app.frontend_url)

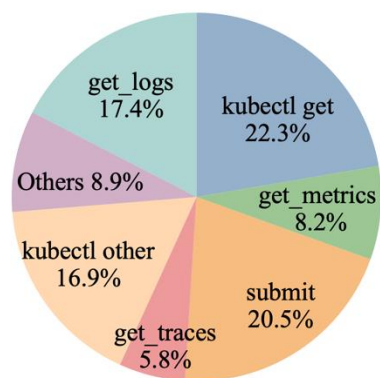
    def inject_fault(self):
        inj = VirtFaultInjector(self.app.ns)
        inj.inject([self.ans], "misconfig_k8s")

    def eval(self, soln, trace, duration):
        res["TTL"] = duration
        res["success"] = is_exact_match(soln, self.ans)
```

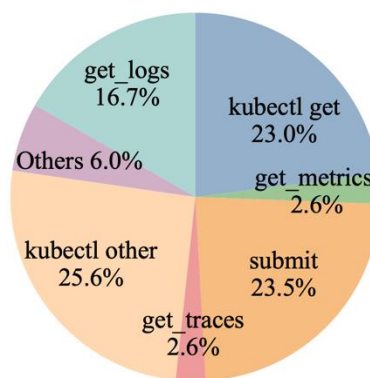
Evaluation

General performance					
Agent	LoC	Time (s)	# Steps	Tokens	Accuracy
GPT-4-W-SHELL	41	28.61	6.44	6394	49.15%
GPT-3.5-W-SHELL	41	12.44	14.70	2557	15.25%
ReAct	49	43.79	11.50	16941	55.93
Flash	60	99.64	8.48	6484	59.32%

Trajectory

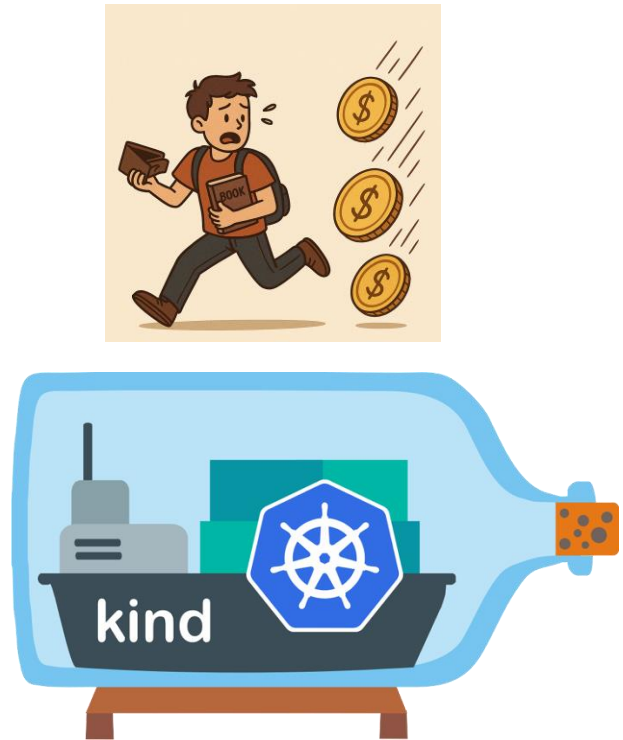


Failure cases



Success cases

Deployment



Local Deployment



Cloud Deployment

Summary

- **Open-sourced framework - AIOpsLab**
 - *deploy* diverse cloud environments, *inject* faults, *generate* workloads, and *export* telemetry data
 - *orchestrate* these components and provide *agent-cloud interfaces* for design, deployment, and evaluation of AIOps agents.
- AIOpsLab currently supports 100 problems (and growing)
 - Across different operation tasks in an interactive environment

- Email: **AIOpsLab@microsoft.com**
- Code: **<https://aka.ms/AIOpsLab-repo>**
- Poster: Booth #4

Yinfang will be on job market in fall 2025.



AIOpsLab: A Holistic Framework to Design, Develop, and Evaluate AI Agents for Enabling Autonomous Clouds

Yinfang Chen, Manish Shetty, **Gagan Somashekar**, Minghua Ma, Yogesh Simmhan, Jonathan Mace, Chetan Bansal, Rujia Wang, Saravan Rajmohan

