KUBERNETES COMMUNITY DAYS BEIJING 2025

多集群环境中AI驱动的故障诊断

AI-Driven Troubleshooting in Multi-Cluster Environments

闫猛（Meng yan）

Software Engineer @Red Hat

Content

目录

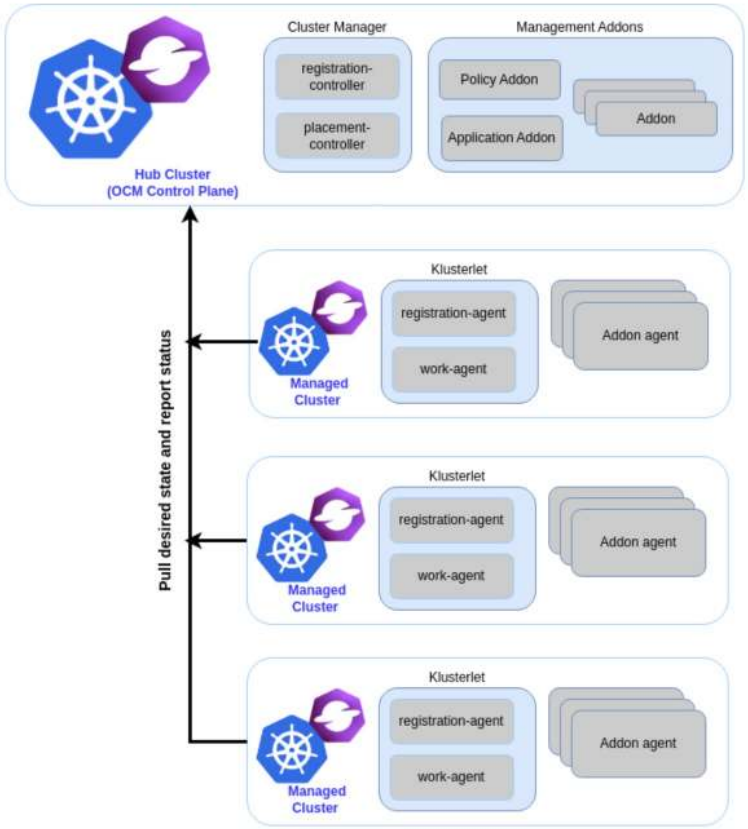
KCD
GROWING CLOUD NATIVE TOGETHER
BEIJING

Part 01

多集群管理-OCM概述

Open Cluster Management

多集群管理平台 - Open Cluster Management



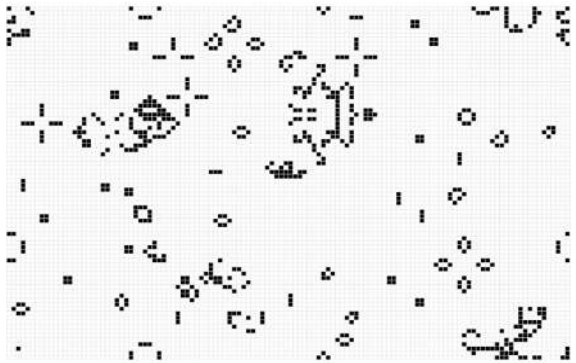
- Kubernetes Multi-Cluster Orchestration: CNCF Sandbox Project
- Architecture: Hub-Spoke, derived from the Hub-Kubelet pattern in Kubernetes, aligning with its native design
- Scalability: Offloads workload to Spoke clusters via agent pulling
- Robustness: Klusterlet and Hub operate independently and autonomously
- Modularity and Extensibility: Pluggable design for customization and further development
  - Example: Placement enables dynamic cluster selection and supports extension or replacement for advanced orchestration.
- More Detail: [Open Cluster Management Document]

Part 02
Agent介绍

ABM - ML - LLM

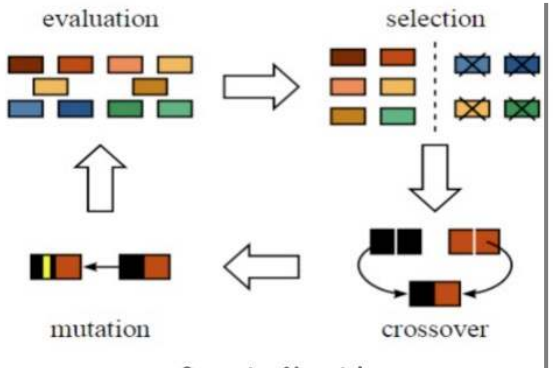
Agent 介绍

智能模拟

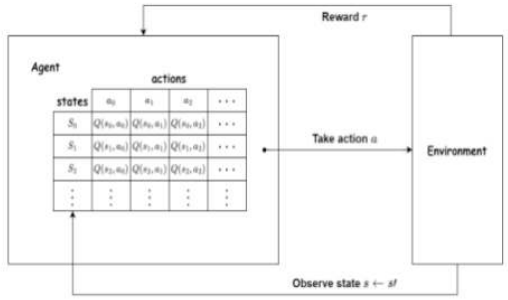

Game of Life

Rule-Based Agent

策略学习



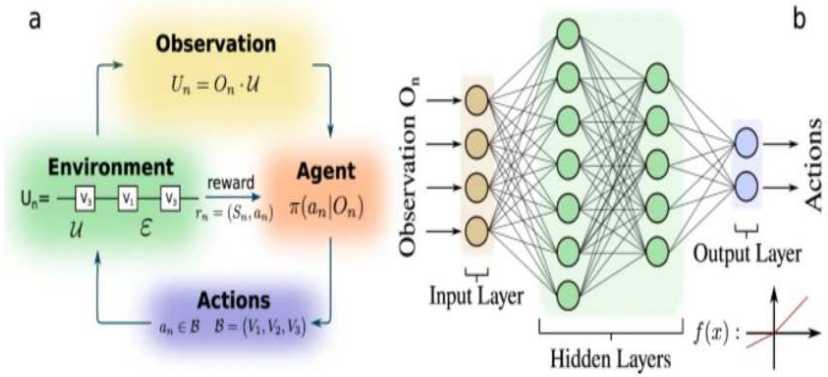
Genetic Algorithm



Q-Learning

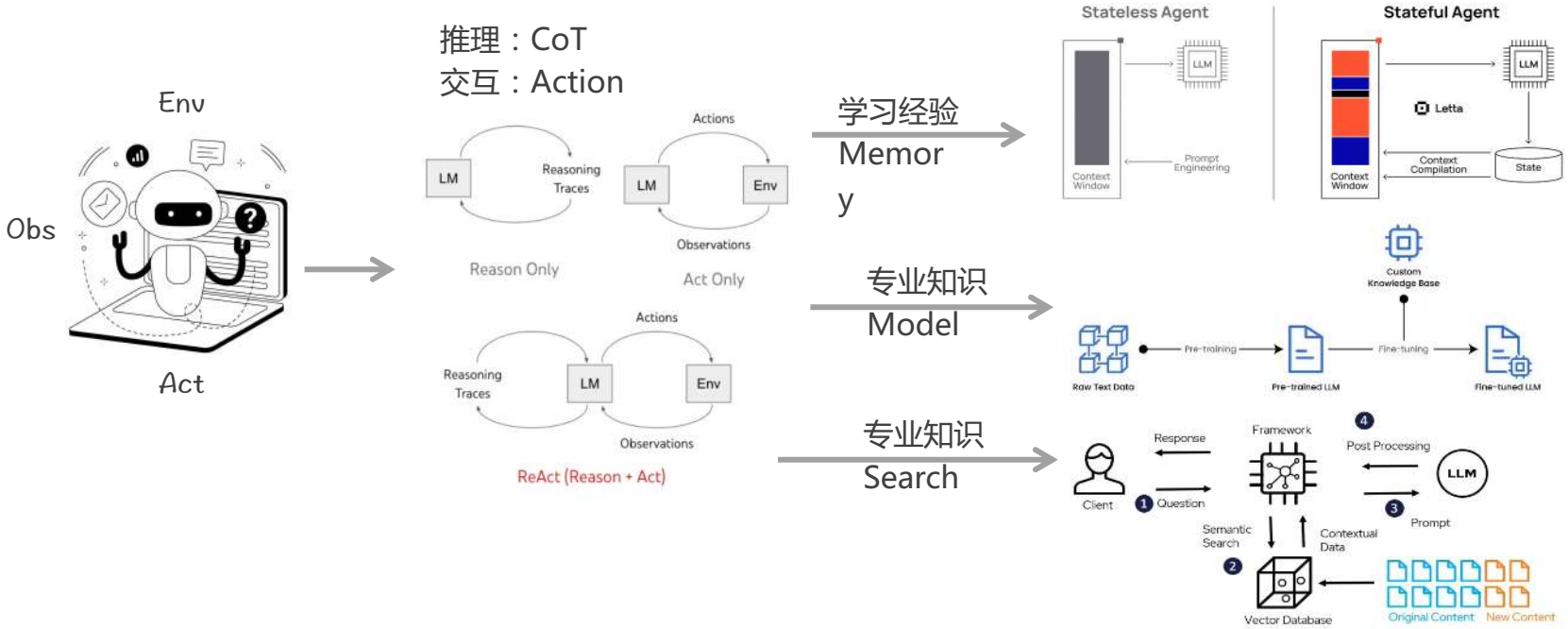
Heuristic Agent

深度学习：高维度决策



Deep Reinforcement Learning Agent

Agent 介绍 - GenAI: LLM



LangChain

AutoGen

Swarm

Letta

LlamaIndex

- ReAct: Synergizing Reasoning and Acting in Language Models (2022)
- MemGPT: Towards LLMs as Operating Systems(2023)
- Retrieval-Augmented Generation (2020)

Part 03

多集群中Agent设计

Open Cluster Management + Multi-Agent Modeling

多集群中Agent的设计

动机 Motivation

- 多集群线上发生故障时，因为时区等问题，专业工程师无法及时响应
- 具备一些背景知识的工程师可以借助 Agentic Workflow 进行实时诊断与故障恢复，提高运维效率和系统稳定性

LLM应用面临的挑战 Challenges

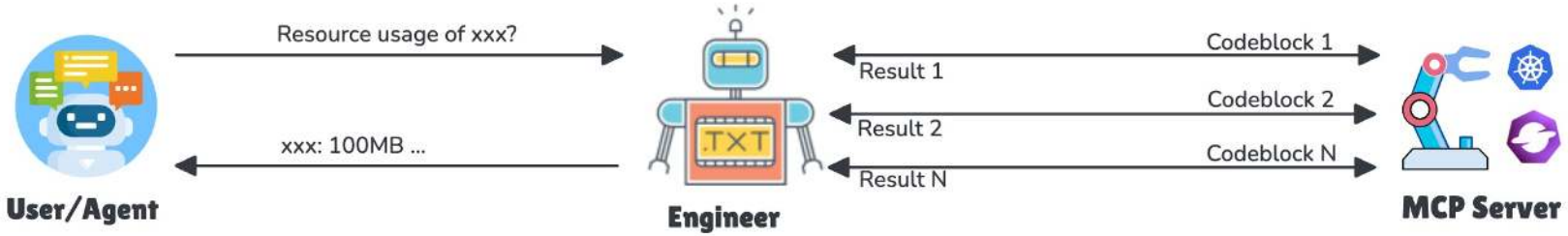
- 准确性 - 幻视（Hullucination）可能导致错误决策
- 领域知识 - 需要实时信息和专业知识的支撑
- 安全性 - 需要严格控制操作权限， 防止误用

应对策略 Solutions

- 提高准确性 - ReAct(CoT)，Multi-Agent System，Model Temperature，Model Type
- 增强领域知识 - Runbook, Search，RAG
- 保障安全性 - 权限控制(Action Permission Control)，从线上日志快照中获取集群上的资源信息

多集群中Agent的设计

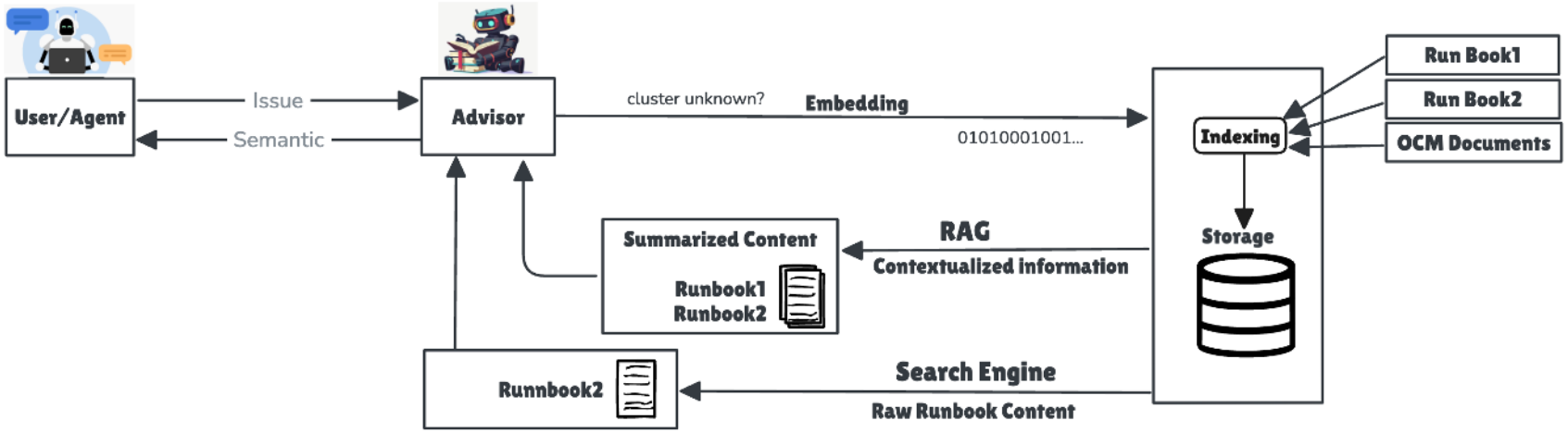
问题1：怎样与多集群交互?
How to Interact with Multiple Kubernetes Environments?



- 工程师：分析用户意图，与多集群进行交互
- Multicluster MCP Server - 构建Open Cluster Management 与 GenAI 的桥梁
  - kubectl解释器：实现对资源的增删查改等各种操作
  - OCM-ManagedServcieAccount: 使用kubernetes原生的RBAC机制，在多个集群间实现权限管理
  - 离线分析：借助其他工具，例如must-gather等和日志进行交互
  - 示例1: MCP - Multiple Kubernetes Operations
  - 示例2: MCP - Cluster Status, Resource Usage

多集群中Agent的设计

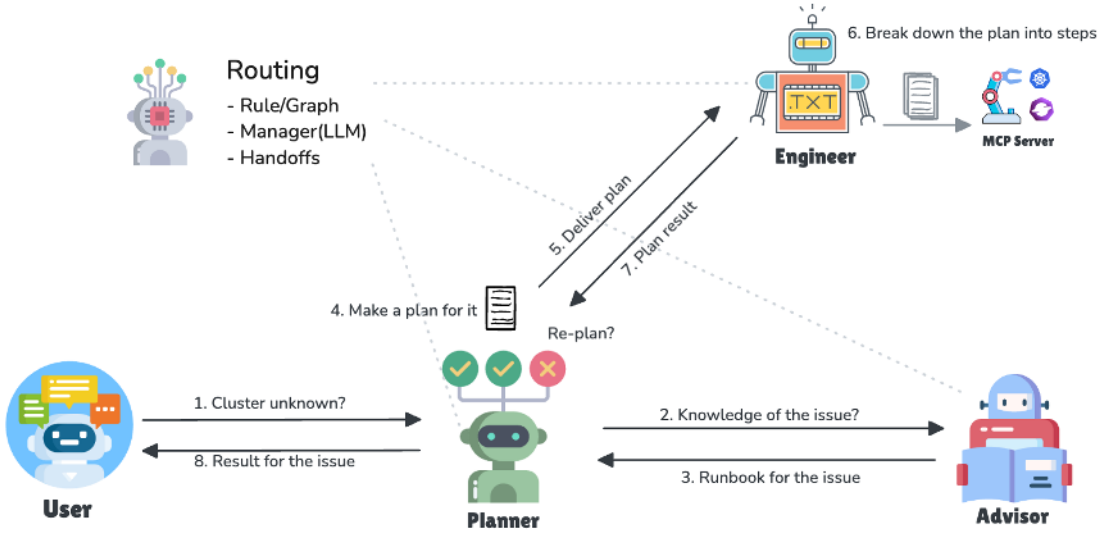
问题2：怎样整合OCM专业知识?
How to Integrate OCM Knowledge/Context into the System?



- 顾问师：提供与特定问题相关的背景知识、文档或操作手册，辅助问题理解
  - 搜索引擎：提供预定义操作手册，由工程师编写，包含可能原因和排查思路。可提高问题定位和修复的准确性，但依赖特定指导手册
  - RAG实现：结合 OCM 文档、Jira Issue、GitHub Issue 进行分析汇总，提供问题原因和解决思路。信息量大但可能有偏差，且依赖模型的文档分析能力，无需额外维护手册

多集群中Agent的设计

问题3：怎样组合调度这些Agents?
How to Orchestrate all Agents within the System?



- 规划师：针对特定的任务，咨询顾问智能体获取背景知识，起草执行计划来让工程师执行
- 执行顺序：使用Handoffs的方式控制信息流在不同Agents之间的流转

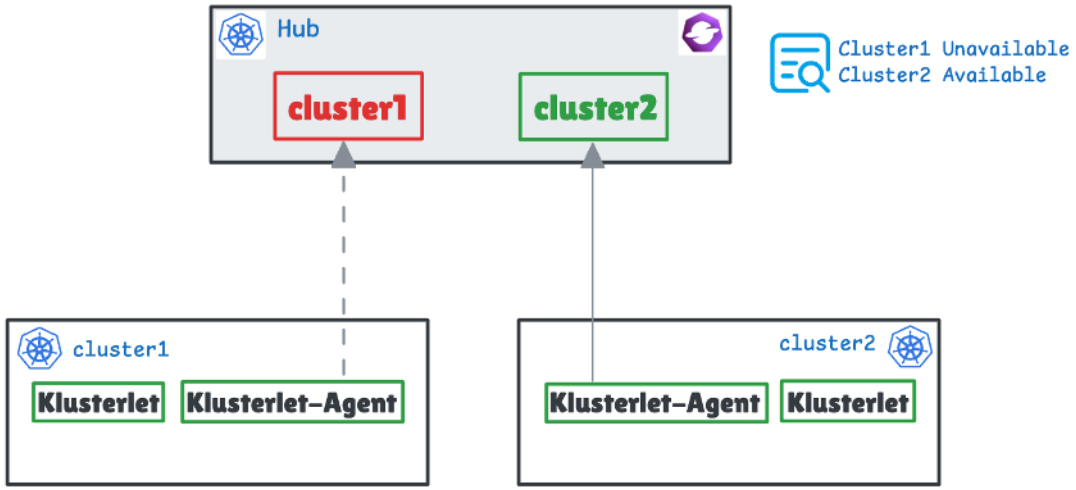
Part 04

样例展示

Demo

样例展示

- Issue 1: Cluster Status Unknown - Disable **klusterlet-agent** in Cluster1



- Issue 2: Cluster Status Unknown - **hub-kubeconfig** is invalid in Cluster2
- Issue 3: **Addons** Not Created

Thanks.