

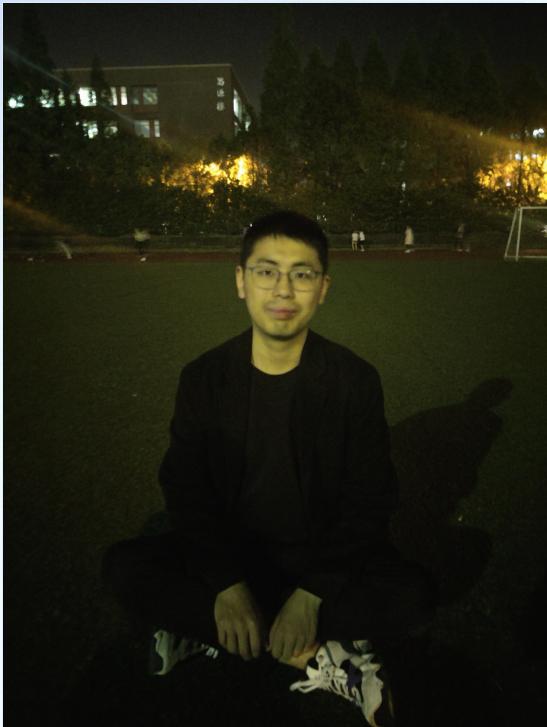


Ray History Server的历史任务可观测性实践

吴昆

阿里巴巴高级研发工程师

2025/12/20



吴昆

阿里巴巴 高级研发工程师

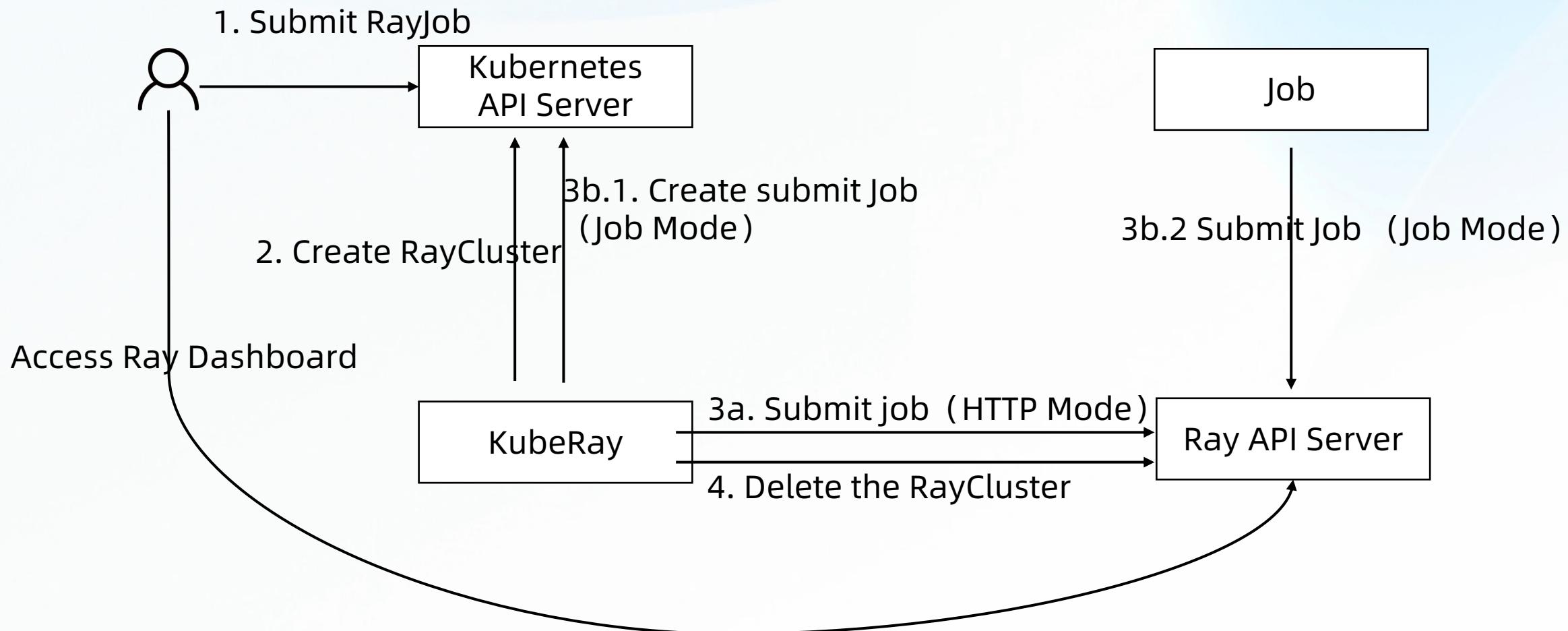
个人简介：吴昆就职于阿里巴巴容器服务团队，是容器服务调度领域负责人以及Ray on ACK研发，同时也是Koordinator, KubeRay, Kueue, Kubernetes等社区贡献者。

CONTENT

目录

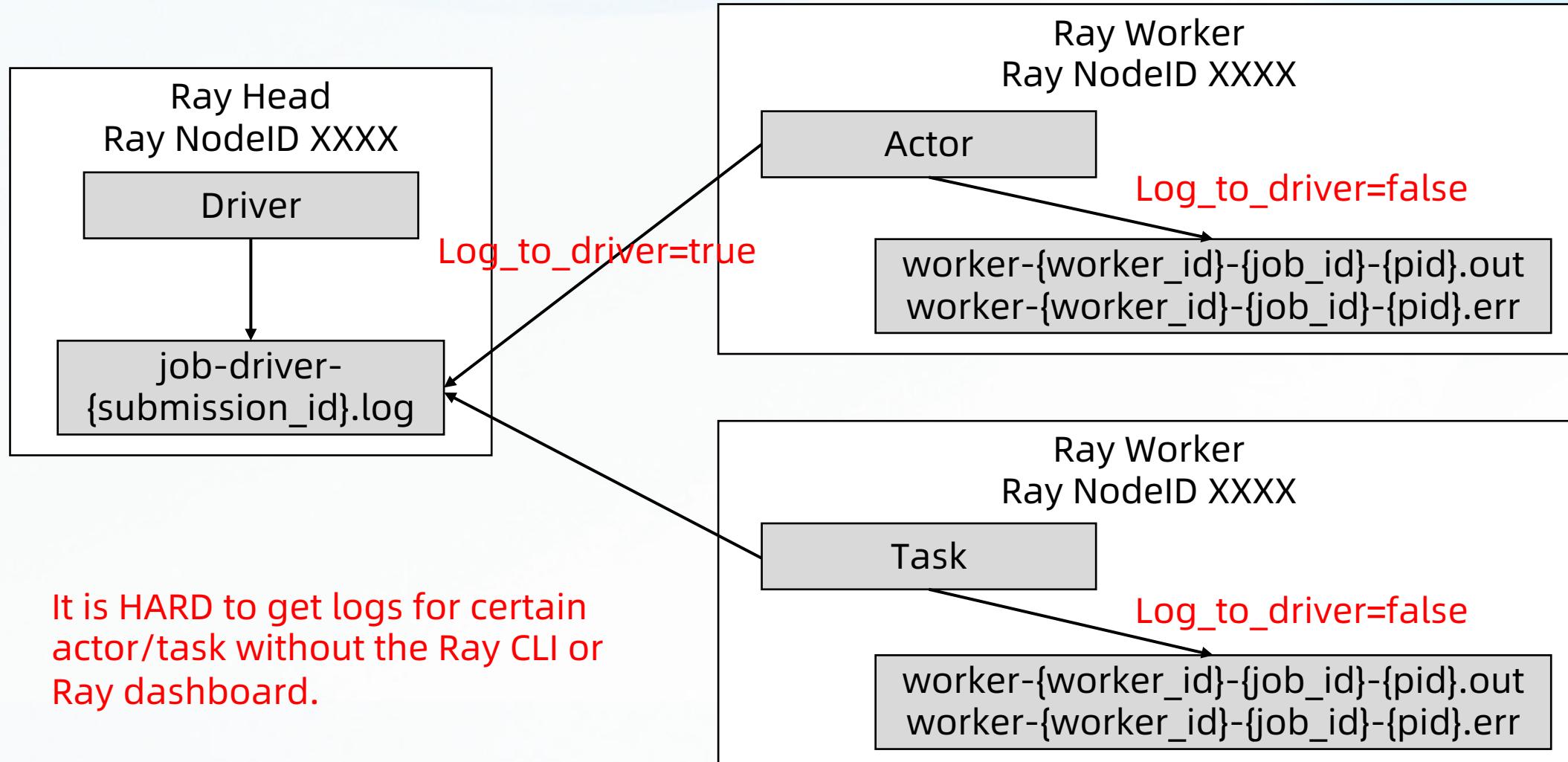
- 01 为什么需要 History Server
- 02 History Server 的设计与实现
- 03 History Server 开源 RoadMap

RayJob on Kubernetes



We can not access the dashboard after the job complete

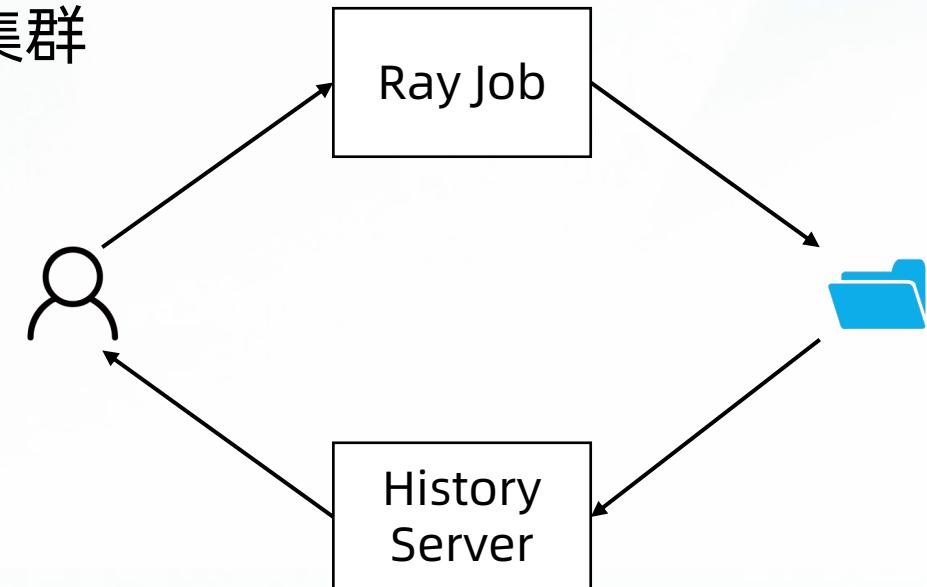
Ray的日志管理方式



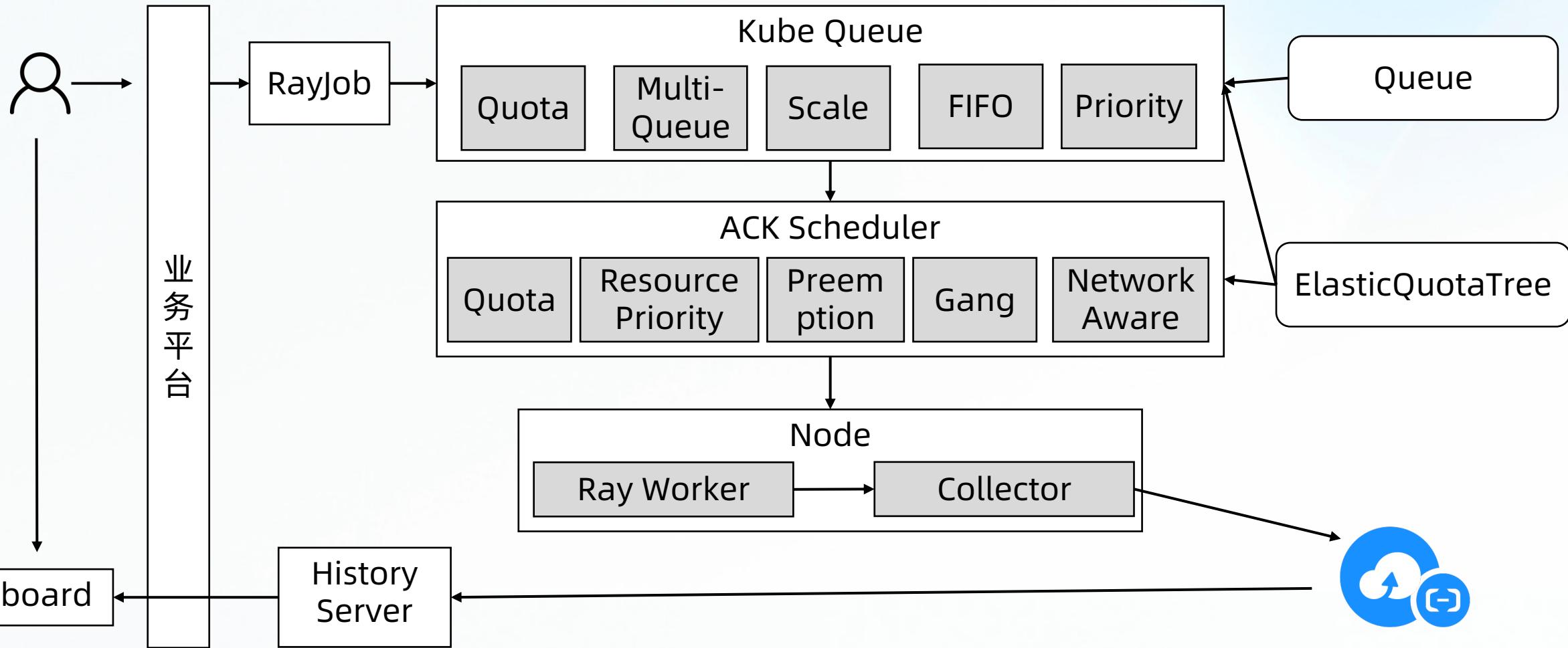
Why History Server ?



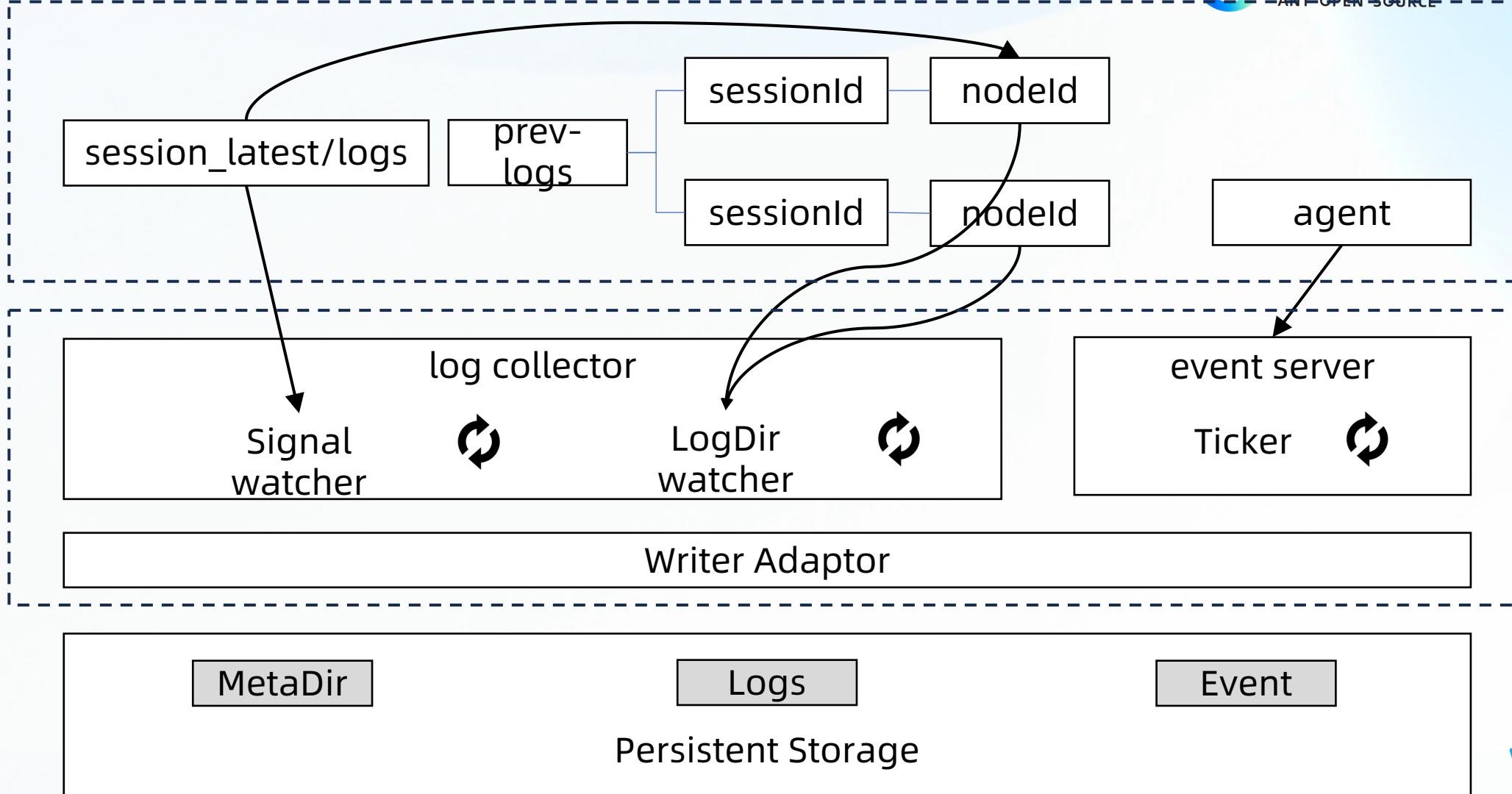
- 支持集群销毁后通过Ray Dashboard访问日志以及集群状态，保持一致的用户体验
- 日志以及集群数据存储在后端，避免部署多个Ray API Server，节省集群资源
- 通过Event收集，HistoryServer可以支持集群状态回放等功能
- 通过Event收集以及重放，减少对Ray API Server的访问，提升Ray API Server性能



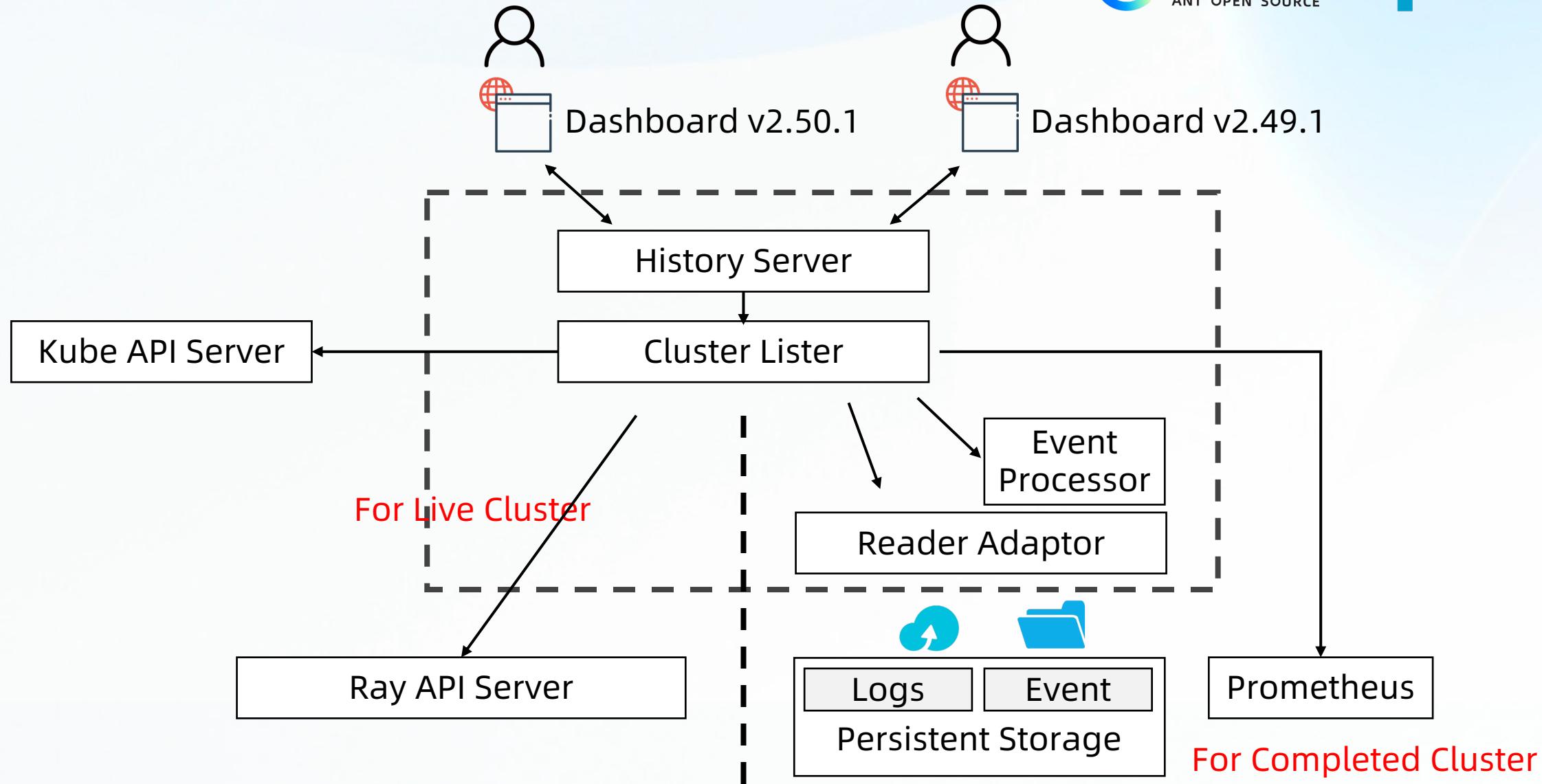
Ray on ACK使用场景



History Server 的设计与实现



History Server 的设计与实现



History Server 的设计与实现



KubeRay HistoryServer Dashboard

Instructions: Select a cluster from the list below, choose a dashboard version, and click "Enter Cluster Dashboard" to view the cluster details.

Available Clusters

Cluster Name	Namespace	Created	Session	Dashboard Version	Action
wukun	default	2025-11-13 14:10:27 +0800 CST	live	Ray Dashboard	Enter Cluster Dashboard
wukun	default	2025-11-12T22:10:30Z	session_2025-11-12_22-10-30_318513_1	Ray Dashboard	Enter Cluster Dashboard
wukun	default	2025-11-12T21:48:38Z	session_2025-11-12_21-48-38_112021_1	1. Select a version v2.51.0	Enter Cluster Dashboard
wukun	default	2025-11-12T21:46:37Z	session_2025-11-12_21-46-37_444456_1	Ray Dashboard	Enter Cluster Dashboard
wukun	default	2025-11-12T20:01:28Z	session_2025-11-12_20-01-28_947657_1	Ray Dashboard	Enter Cluster Dashboard
wukun	default	2025-11-12T19:59:23Z	session_2025-11-12_19-59-23_509375_1	Ray Dashboard	Enter Cluster Dashboard

1. Select a version
2. Enter the dashboard



localhost:8080/#/overview

Overview Jobs Serve Cluster Actors Metrics Logs

Cluster utilization

1. Set up Prometheus and Grafana for better Ray Dashboard experience

Time-series charts are hidden because either Prometheus or Grafana server is not detected. Follow [these instructions](#) to set them up and refresh this page.

Recent jobs

No jobs yet...

[View all jobs →](#)

Cluster status and autoscaler

Node count

1. Set up Prometheus and Grafana for better Ray Dashboard experience

Time-series charts are hidden because either Prometheus or Grafana server is not detected. Follow [these instructions](#) to set them up and refresh this page.

No cluster status.

Serve Deployments

No Deployments yet...

[View all deployments →](#)

No cluster status.

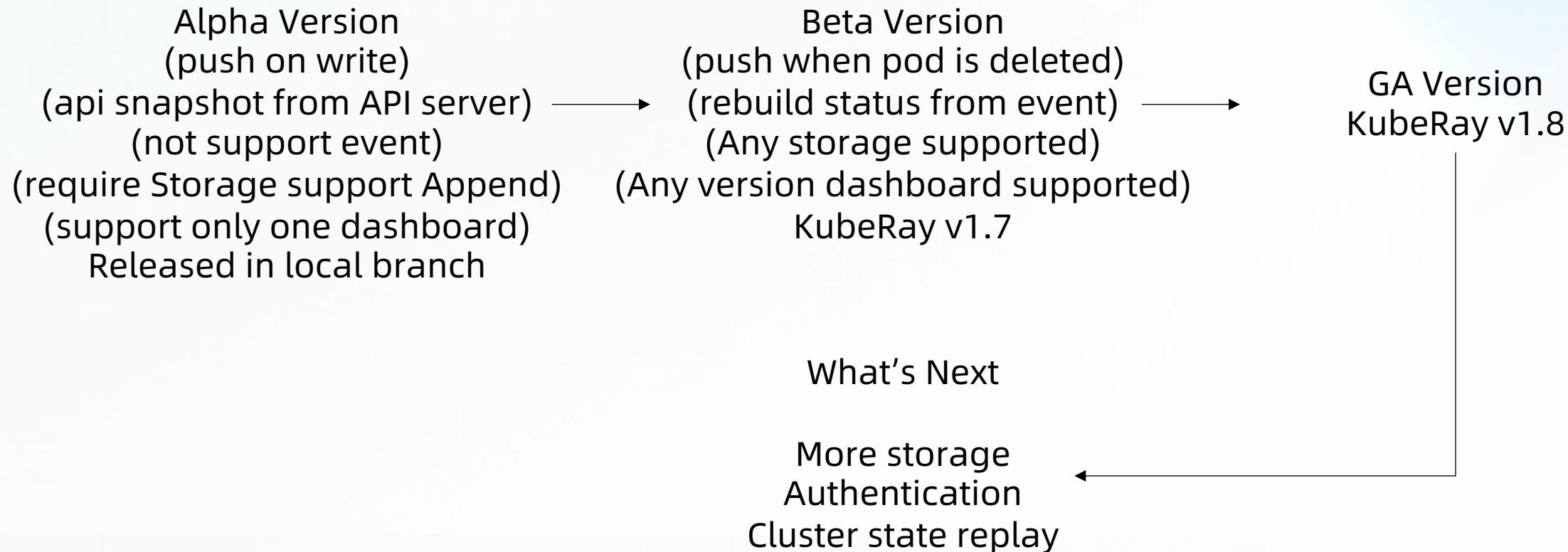
Demo视频



A screenshot of a Linux terminal window. The title bar says 'kuberay'. The command line shows the user is in a directory under 'ray-historyserver-demo'. A large white cursor arrow is positioned in the center of the screen. At the bottom, the text 're api is installed' is displayed in a large, bold, black font.

1. Ensure api is installed

History Server RoadMap





谢谢
欢迎大家进群交流

