



Fluid CSI Plugin重构

徐之浩 (Fluid项目Maintainer)

南京大学 PASA Lab

1 Fluid CSI Plugin工作原理

2 现存问题

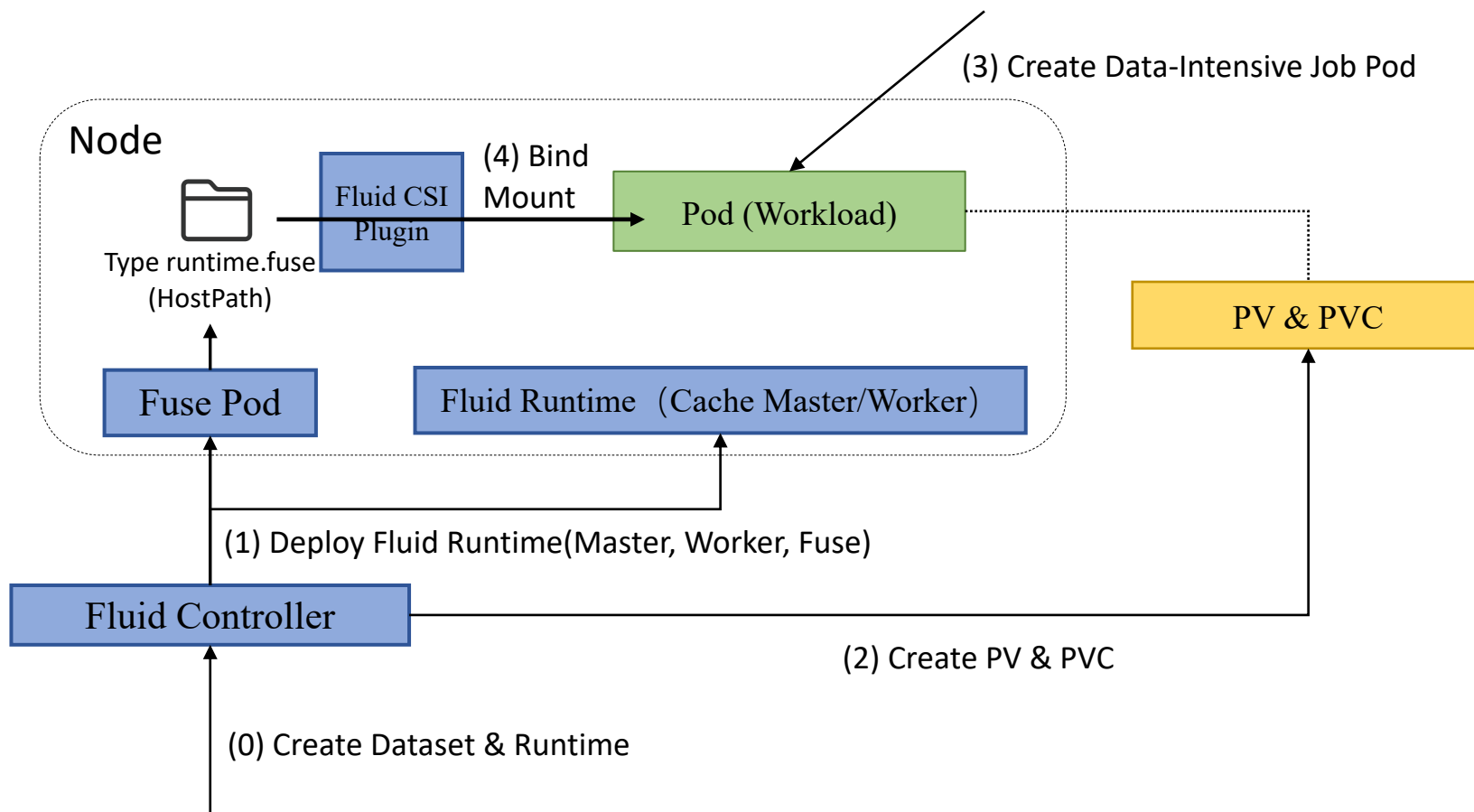
3 重构设计方案

4 系统演示

Fluid CSI Plugin工作原理



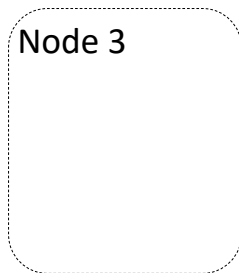
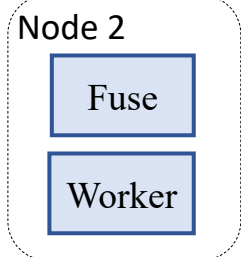
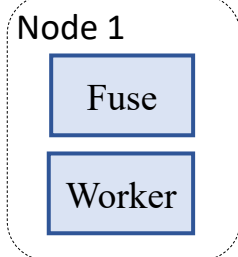
现有架构



Fuse部署模式

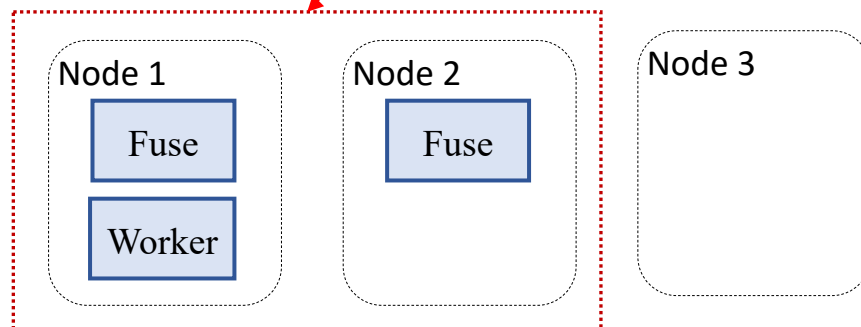
Default (Non-global)

```
apiVersion: data.fluid.io/v1alpha1
kind: AlluxioRuntime
metadata:
  name: mydata
spec:
  replicas: 1
  ...
  fuse:
    global: false # default
```



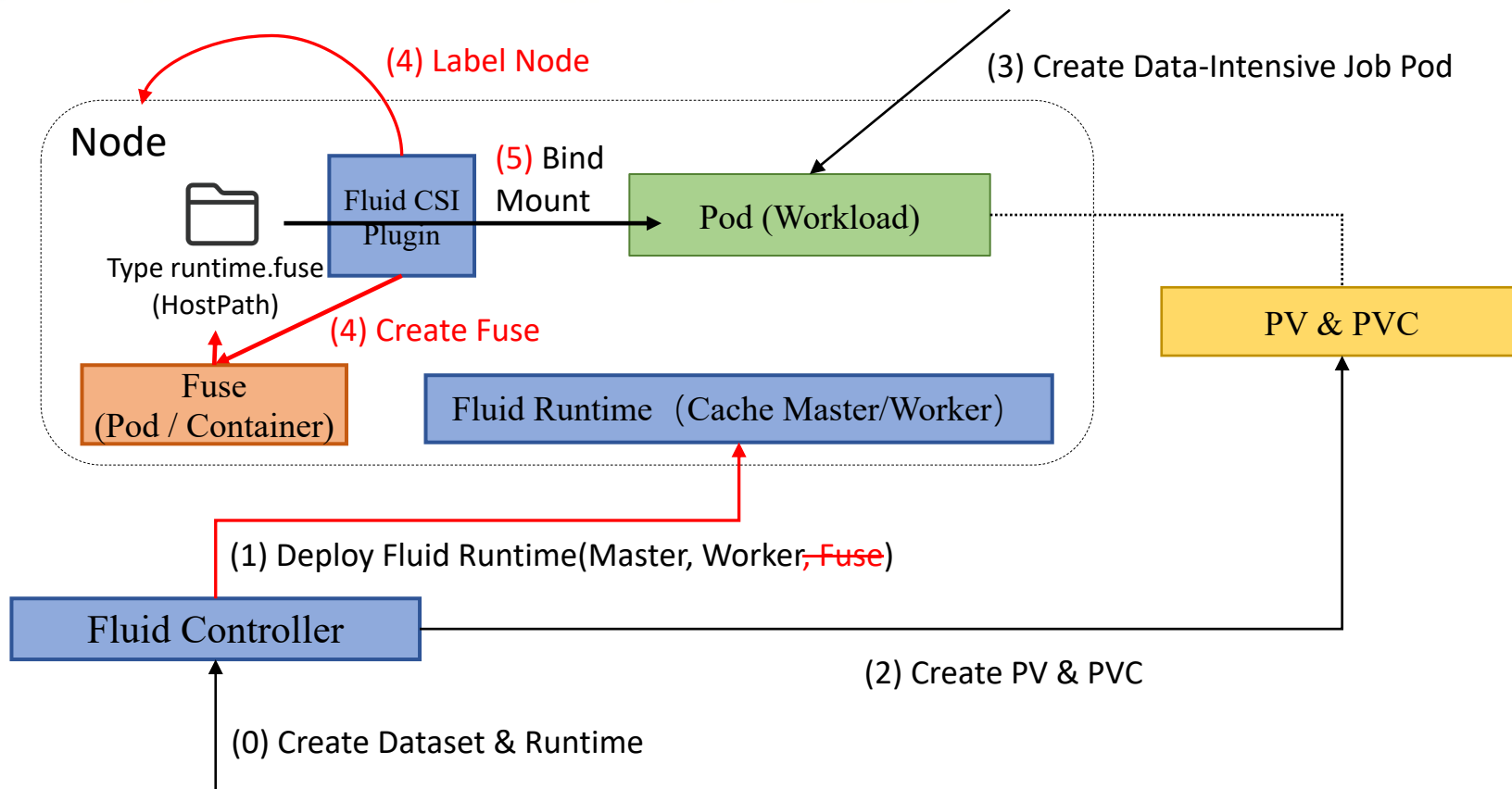
Global

```
apiVersion: data.fluid.io/v1alpha1
kind: AlluxioRuntime
metadata:
  name: mydata
spec:
  replicas: 1
  ...
  fuse:
    global: true
    nodeSelector:
      foo: bar
```



- Fuse Pod被K8s驱逐，使得应用数据访问崩溃
- 闲置Fuse Pod空占集群资源，资源利用率低
- Fuse Pod数量固定，无法灵活适应数据应用实际使用情况

重构方案



- Lazy启动Fuse容器或Pod(k8s version ≥ 1.17 , system-node-critical)
- 支持不同Container Runtime(e.g. Containerd, Docker) 启动Fuse容器
- 通过Node Label暴露Fuse正在某节点上运行的状态

Fuse部署模式更改

- 取消Fuse部署模式Non-Global、Global配置
- Fuse NodeSelector保留，控制可启动Fuse的节点

调度信息注入

- 向Workload Pod中注入节点亲和信息
- 优先调度到正在运行Fuse的节点 (Preferred Affinity)
- 优先调度到Cache Worker所在节点 (Preferred Affinity)

Thank You!