

# JuiceFS 在 Kubernetes 环境中数干节点数据集的应用实践

苏锐 - Juicedata 合伙人



## 分享大纲



- JuiceFS 是什么
- JuiceFS 在 Kubernetes 上的几种使用姿势
- 数据在AI + Kubernetes中遇到的挑战
- 提升 JuiceFS 在大型 Kubernetes 集群中的体验



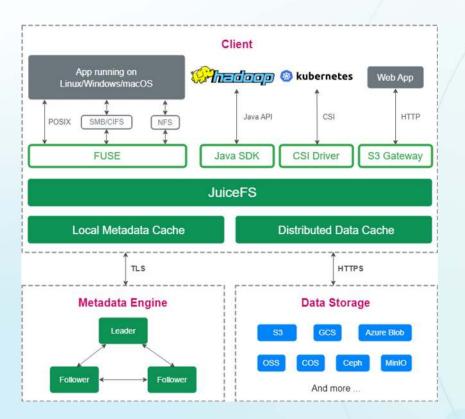
## 苏锐

- 2017 年作为联创开始 JuiceFS 的创业之旅
- 18年IT工作,做过 Tech Lead、PM、CEO
- 西电三系校友



## JuiceFS 是什么?





- 2017 年发布云服务;
- 支持(几乎)所有全球公有云;
- 生产最大规模单卷近千亿文件, 百 PB 容量 , 聚合吞吐数百 GBps;
- 100% POSIX 兼容。

















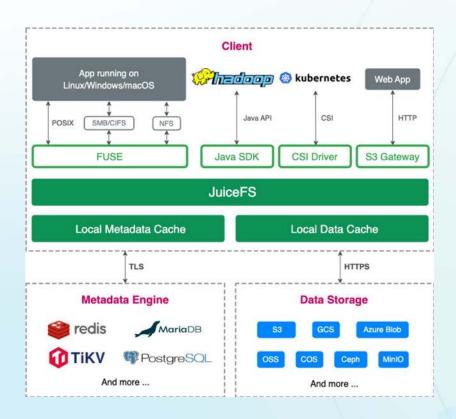


**DP**Technology 深势科技



## JuiceFS 是什么?





- 2021年发布;
- GitHub 11.3K 🔆;
- 胖客户端模式,简单上手,简单运维;
- Golang 开发 + CSI 完善,得到云原生开发 者支持;
- 使用最多:
  - AI 平台
  - Kubernetes PV
  - 大数据存算分离



## JuiceFS 在 Kubernetes 上的几种使用姿势



**hostPath** 

CSI - MountPod

CSI - Sidecar



#### hostPath



- 将 JuiceFS 挂载到所有宿主机的同一个目录,比如/mnt/data;
- 挂载参数在宿主机管理;
- Pod 声明 hostPath,指定挂载路径/mnt/data;
- 灵活性不足;
- 挂载点故障后无法无感恢复。

apiVersion: v1

kind: Pod metadata:

name: hostpath-pod

spec:

#### containers:

 name: test-container image: busybox

volumeMounts:

- mountPath: /data

name: hostpath-volume

#### volumes:

- name: hostpath-volume

hostPath:

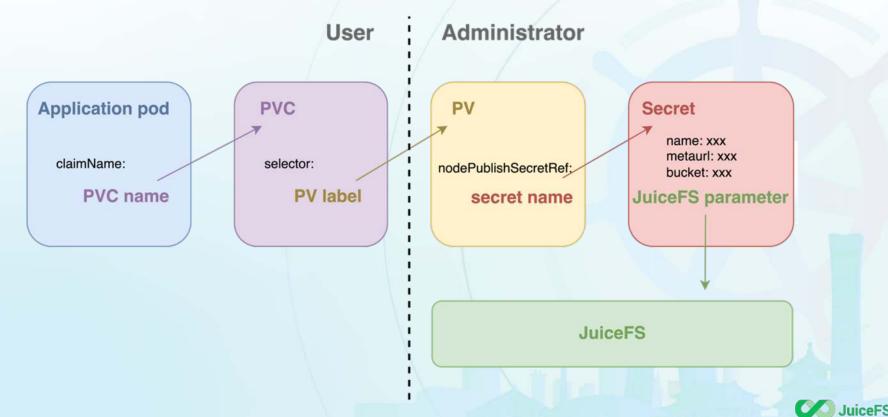
path:/mnt/data

type: Directory



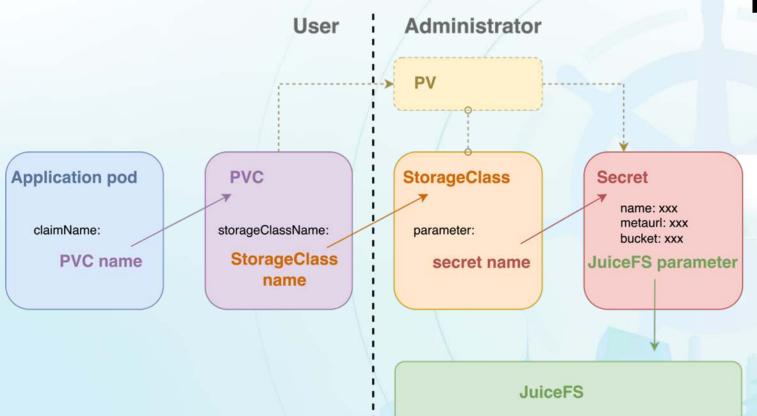
### **CSI - MountPod - Static Provision**





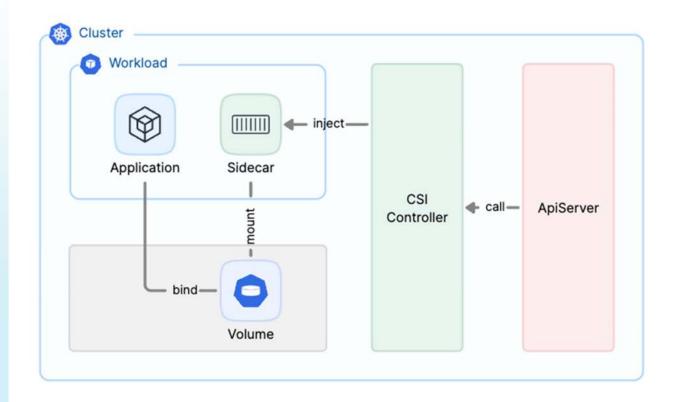
## **CSI - MountPod - Dynamic Provision**





## **CSI - Sidecar**





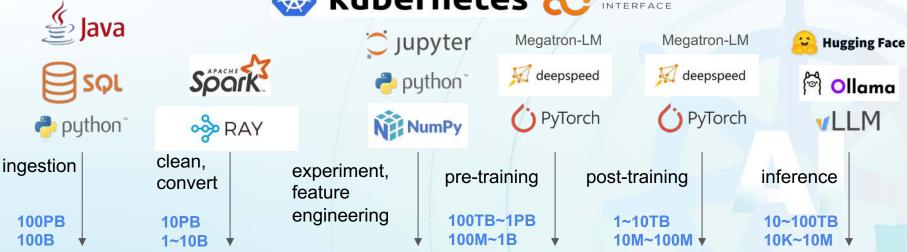


## 数据在 AI + Kubernetes 中的挑战









#### **JuiceFS Unified Storage**

Parquet, Avro, ORC, Lance

Structured data

Log, CSV, JSON ...

Semi-structured data

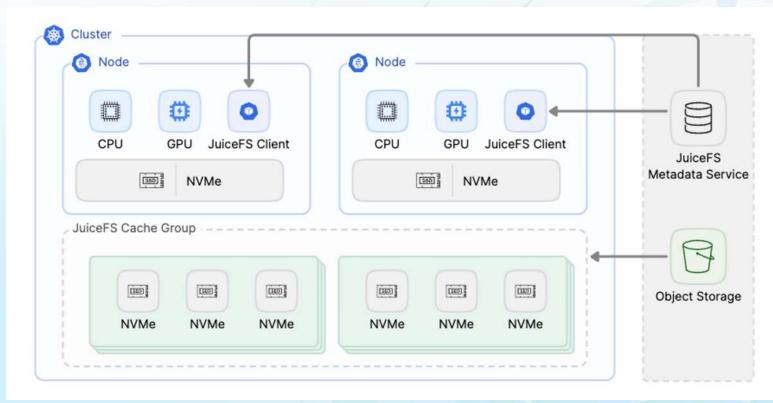
Image, Video, Audio, PCD, PDF, Fastq, XYZ, Bag, FFRecord ...

Unstructured data



## 大家都关心的:性能

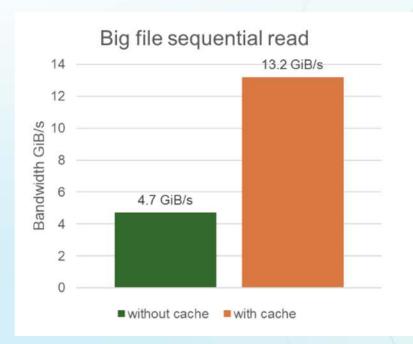


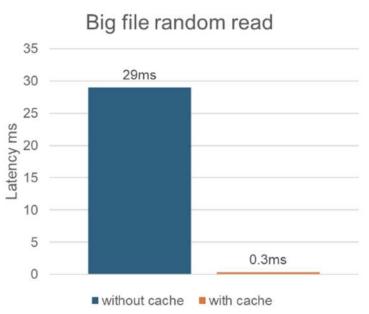




## 大家都关心的:性能









## 大家容易忽视的: 图形化可观测 - CSI Dashboard



JuiceFS CSI	http://localhost:8088						① · ② · ¾	(
Application Pod     System Pod	Application Pod							
PV PVC Storage Class	Name: Pinase enter Status: Pinase select		Namespace Fleate enter PV: Fleate en		Mount	t Pods Phrase enter		
	Application Pod List ⊙					CIS	A	
	Name	Namespace	PV	Mount Pods	Status	CSI Node	CreateTime	
	ce-dynamic-6c6b54478d- pwd6b	kube- system	<ul> <li>pvc-9a09305d-5004-45ec-83c9- 0e626ae3b59d</li> </ul>	<ul> <li>juicefs-cn-hangzhou,10.0.1.84-dynamic-ce- drkagd</li> </ul>	• Running	• juicets-esi-node- p6ccl	2024/8/2 14:25:50	
	ce-dynamic-6c6b54478d- mrh87	default	<ul> <li>pvc-a77b0ac6-7a24-4545-8484- 76f9f2fac5db</li> </ul>	<ul> <li>juicets-cn-hangzhou,10.0.1.84-dynamic-ce-drkagd</li> </ul>	• Running	• juicets-csi-node- p6cct	2024/8/2 11:11:44	
	ce-static-5445fc7dbd-r7s98	default	• ce-static	<ul> <li>juicefs-cn-hangzhou.10.0.1.84-ce-static-handle- bspgdx</li> </ul>	• Running	• juicets-csi-node- p6ccl	2024/7/31 17:46:45	
	normal-664f8b8846-mt4fb	default	• ce-static	• juicefs-cn-hangzhou,10.0.1.84-ce-static-handle-kuudke	<ul> <li>Running</li> </ul>	• juicefs-csi-node- p6ccl	2024/7/17 10:21:56	
	cn-wrong-767577678d-17pcz	default	• ce-static	juicefs-cn-hangzhou.10.0.1.84-ce-static-handle-kuudke	• CrashLoopBackOff	• juicets-csi-node- p6ccl	2024/7/17 10:21:56	
	PVC which it uses was not successfully bound, please click "PVC" to view details.	sidecar	• ce-sidecar	<ul> <li>juicefs-cn-hangzhou.10.0.1.84-ce-sidecar- handle-citobf</li> </ul>	* Running	• juicets-csi-node- p6cct	2024/7/17 10:21:56	
	pending o	default	<b>5</b> )	5.	• Pending	*	2024/6/24 11:38:15	
	res-err	default	• ce-static	*	<ul> <li>Pending</li> </ul>	0	2024/6/24 11:38:00	
						1-8	of 8 items < 1 >	

\$ helm install juicefs-csi-driver juicefs/juicefs-csi-driver



## 大家容易忽视的:缓存组配置简单灵活了



#### 之前,我们要通过 StatefulSet 或 DaemonSet 的方式创建缓存组,但存在以下问题:

- 无法在同一集群内针对不同节点类型或资源(如挂载参数、缓存组权重等)进行单独配置;
- 需要依赖人工监控并手动添加或移除节点,操作繁琐,容易出错;
- 缓存清理需手动执行,不能自动化。

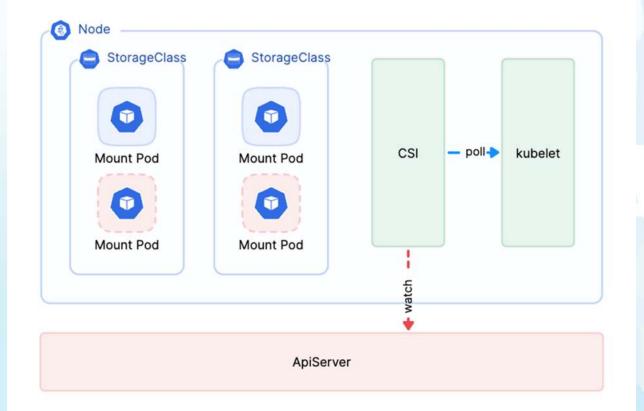
#### JuiceFS 用 Cache Group Operator 改善了上面的问题:

- 在同一集群中配置不同的节点类型和资源;
- 支持平滑添加或移除节点,尽可能减小加减节点期间缓存命中率波动;
- 缓存自动清理;
- Dashboard 中可以管理缓存组。



# 大家容易忽视的:给 ApiServer 减负

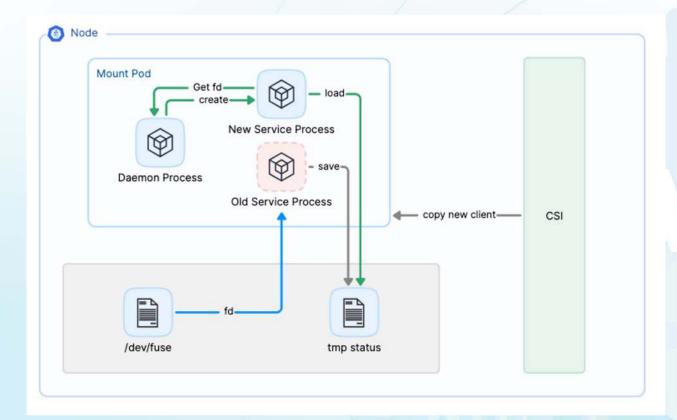






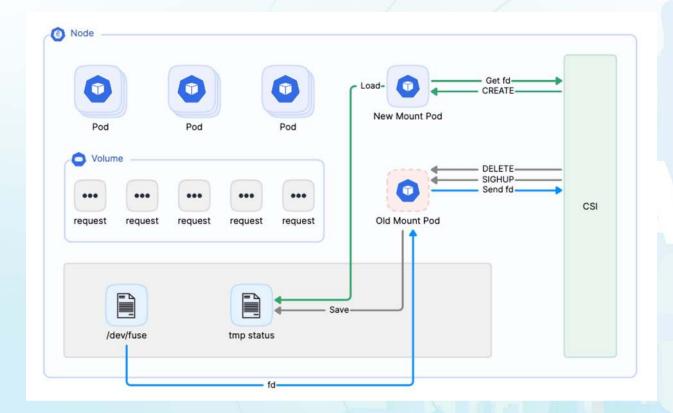
## 大家容易忽视的: JuiceFS 平滑升级,业务不中断





## 大家容易忽视的: JuiceFS 平滑升级,业务不中断







# 谢谢大家







