



GOTC 2023

全球开源技术峰会

THE GLOBAL OPENSOURCE TECHNOLOGY CONFERENCE

OPEN SOURCE, INTO THE FUTURE

「基础设施与软件架构」专场

本期议题：文件存储在 AI 时代的实践与演进

苏锐 2023年5月28日

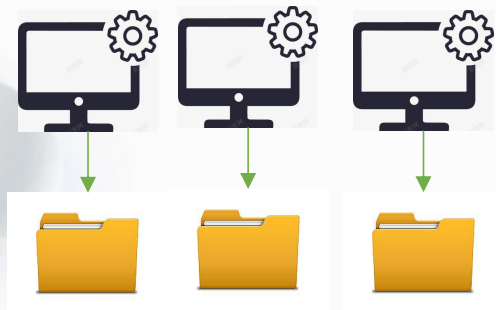
自我介绍

- 苏锐
- 西安电子科技大学 北京航空航天大学（退学）
- 16 年 IT 领域 做过 Tech Lead, 产品经理, CEO
- 2017 参与 Juicedata, 联创



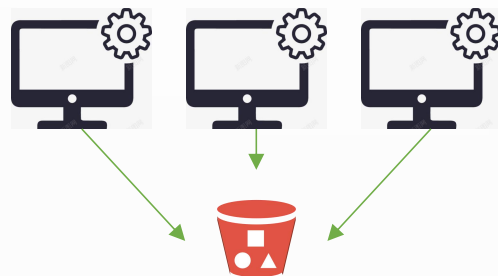
块存储 (云盘 / EBS)

- + POSIX
- + 和本地盘使用体验一样
- 单机访问
- 不弹性
- 有容量上限
- 贵



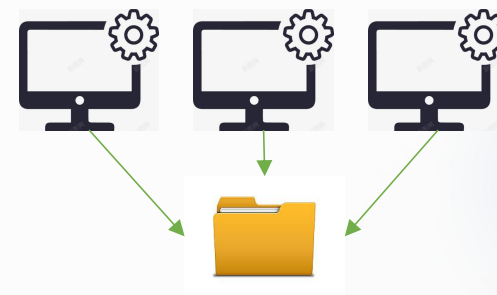
对象存储 (S3 / OSS)

- + 多机访问
- + 弹性
- + 高可用
- + 便宜
- 只有 RESTful API
- 没有原生目录结构
- 元数据访问性能弱
- 不能原子 Rename



文件存储 (EFS / NAS)

- + 多机访问
- + POSIX / NFS
- + 体验类似本地盘
- + (部分) 弹性
- 性能与容量线性相关
- Inode 天花板
- 贵



随着数据量增长和硬件的发展，文件存储的变迁



1980 - 2003
局域网多机共享,
Web 1.0 数据共享

NAS

EMC / NetApp

- 单点故障
- 控制器瓶颈
- 共享受限
- 横向扩展困难

2004 - 2010
满足 Web 2.0 数据增长,
面向物理机的架构设计

分布式文件系统

Ceph

- 难掌握
- 横向扩展难
- 小文件性能受限

HDFS

- 单点故障
- Inodes 受限
- GC 影响可用性

2006 - 2016
移动互联网数据增长,
公有云时代到来

对象存储

S3

- 缺少完整文件系统语义
- List、Rename 代价大
- 最终一致性
- 扁平空间管理难
- 时延高

2017 - 现在
复杂的数据使用场景
云原生 & 容器化

云上分布式文件存储

JuiceFS

- 多访问协议兼容
- 强一致性
- 弹性伸缩
- 高性能

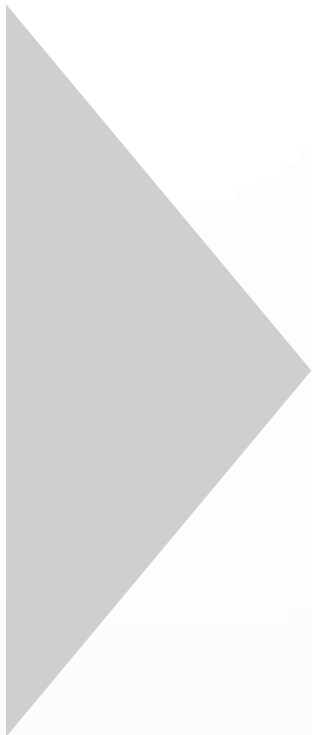
数据量大

性能需求难以预测

GPU资源昂贵，吞吐性能要求高

业务流程长，一份数据多次使用

涉及技术栈多样



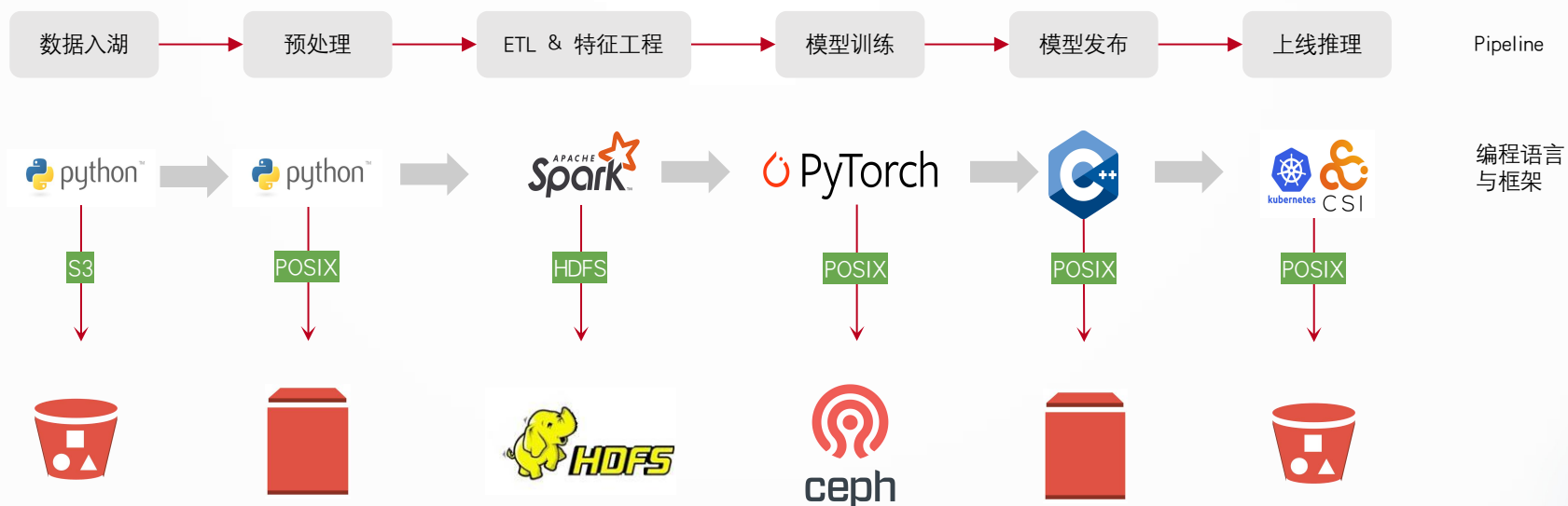
如何打造一个
弹性、高效、高性能的数据存储平台？

多套存储系统
难管理

吞吐扩展难

跨云数据访问
效率低

POSIX 接口，解决AI业务流程多技术栈痛点



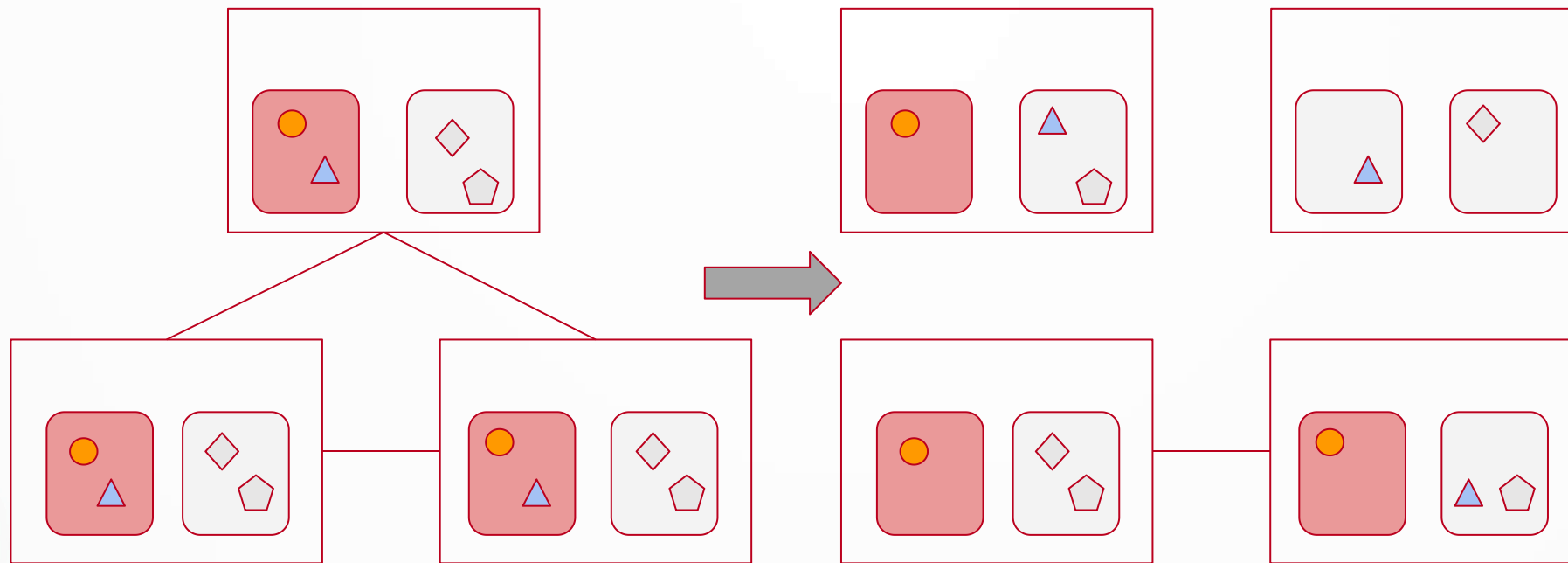
多套存储系统
难管理

吞吐扩展难

跨云数据访问
效率低

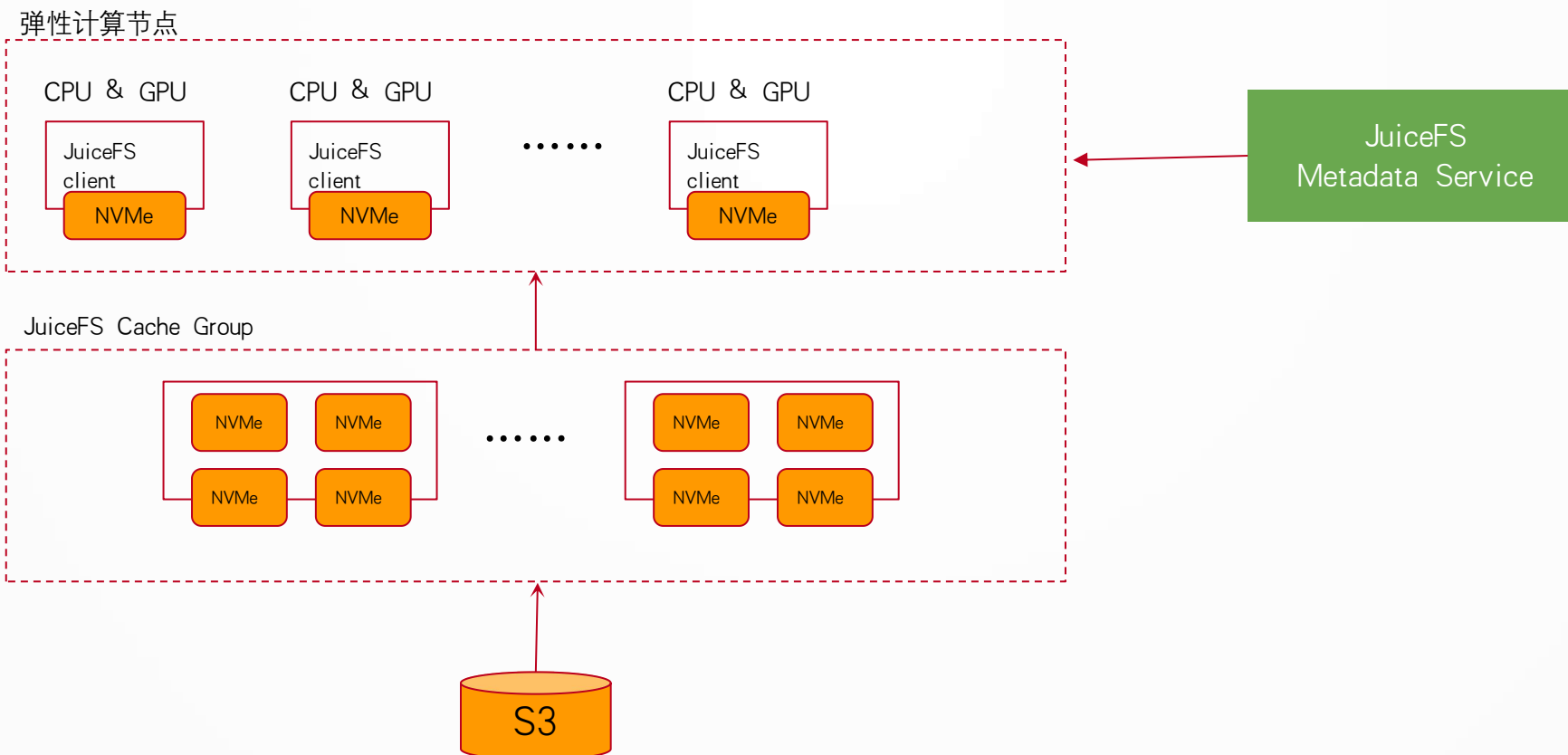
AI业务吞吐扩展难的原因

旧方案：增加更多的磁盘，才能提供更大的吞吐。

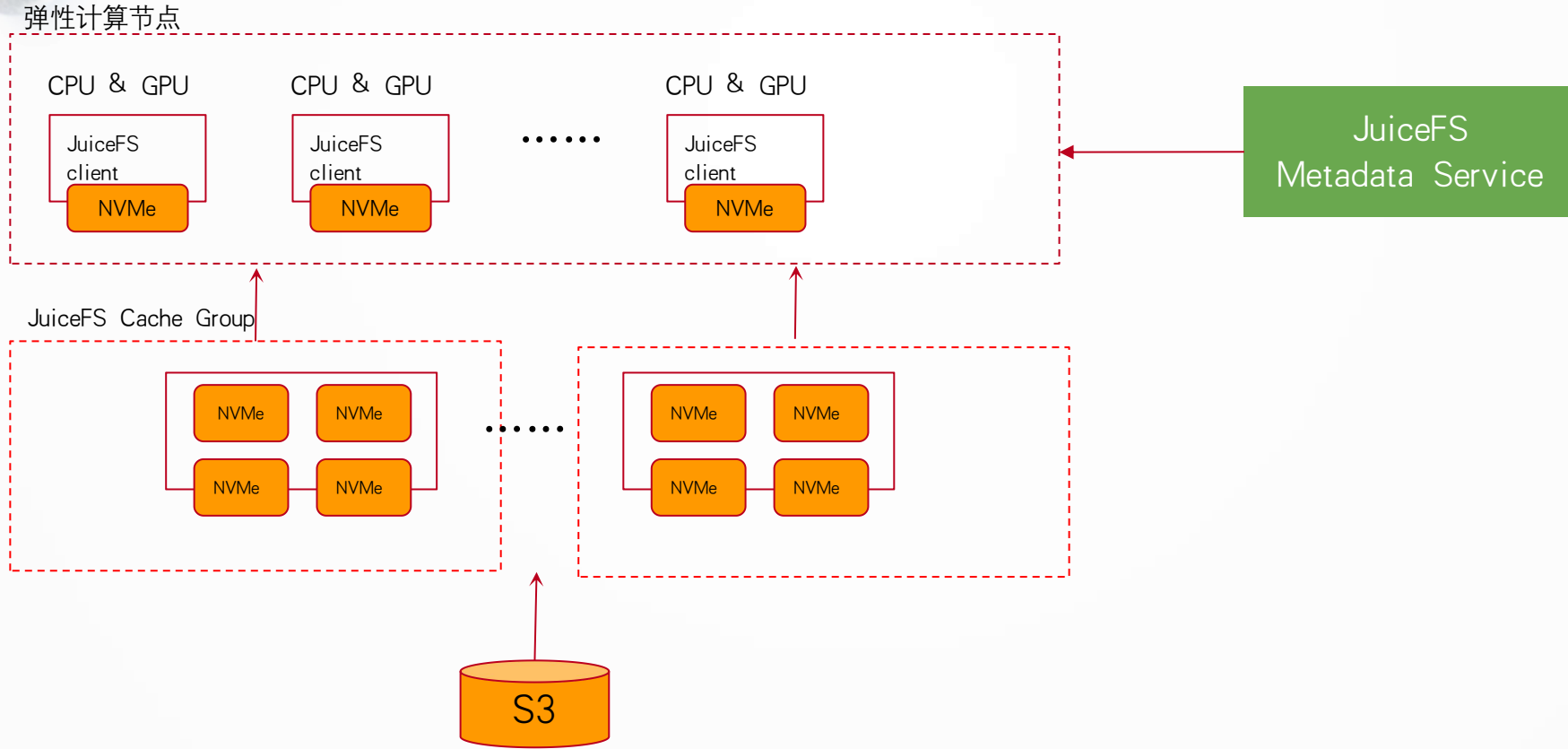


性能不足引发的存储扩容，造成存储空间闲置

如何实现弹性吞吐- 缓存



如何实现弹性吞吐- 缓存分组，放大吞吐



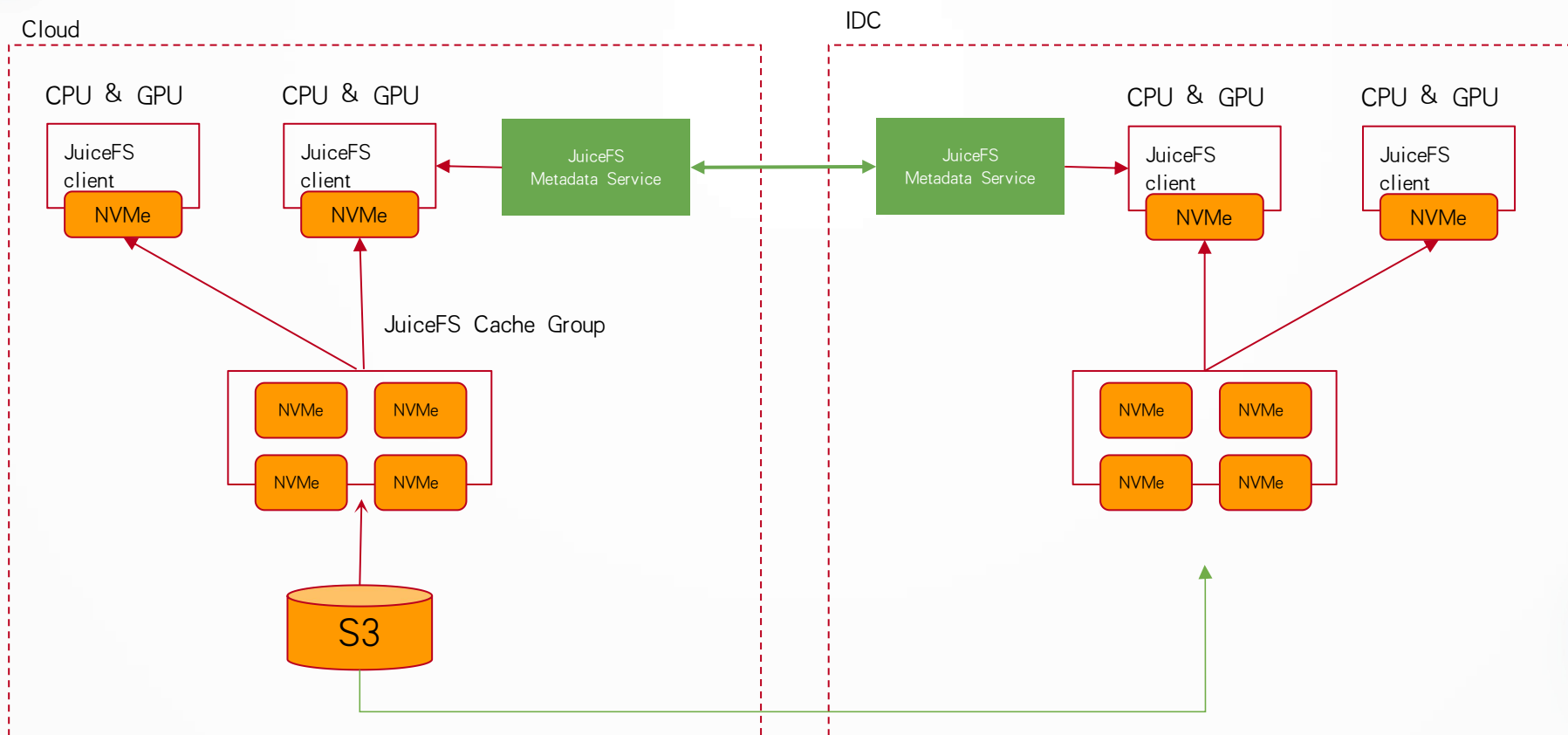
JuiceFS 缓存分组架构图

多套存储系统
难管理

吞吐扩展难

跨云数据访问
效率低

高效实现跨云数据访问：数据镜像



JuiceFS 数据镜像架构图

THANKS



公众号



企业微信