



KUBERNETES
COMMUNITY DAYS CHINA 2022

云原生存储CubeFS在 机器学习平台的技术实践

OPPO 贺迟

01

CubeFS项目背景

02

CubeFS架构设计

03

CubeFS在OPPO机器学习平台的实践

04

CubeFS的生态建设

CubeFS 是面向大规模容器平台的分布式文件系统，在 Apache 2.0 开源协议下发布。

Github地址: <https://github.com/cubeFS/cubefs>

1. 大量用户同时访问。
2. 数据量呈现翻倍增长的趋势。
3. 文件大小可以从几KB到几TB。
4. 多变的文件访问方式。

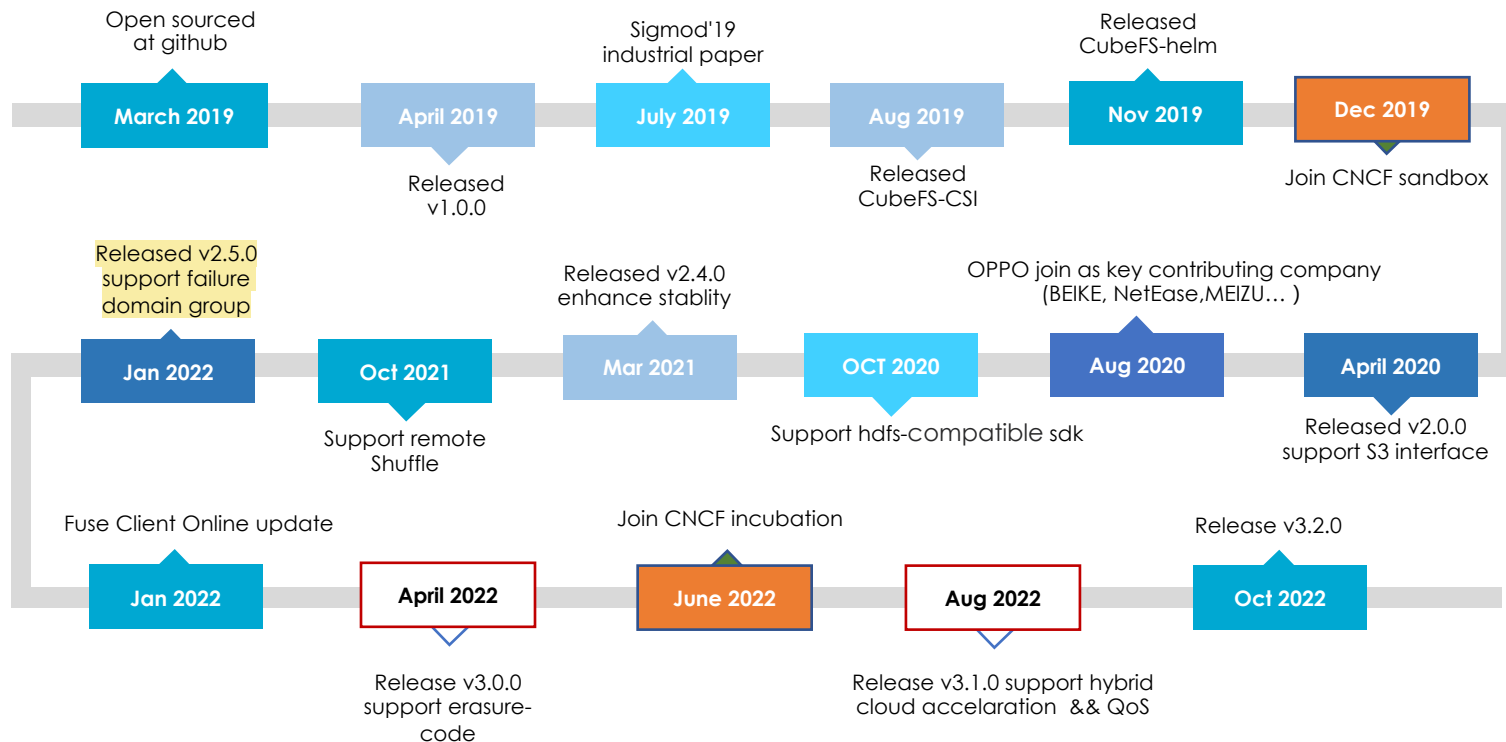


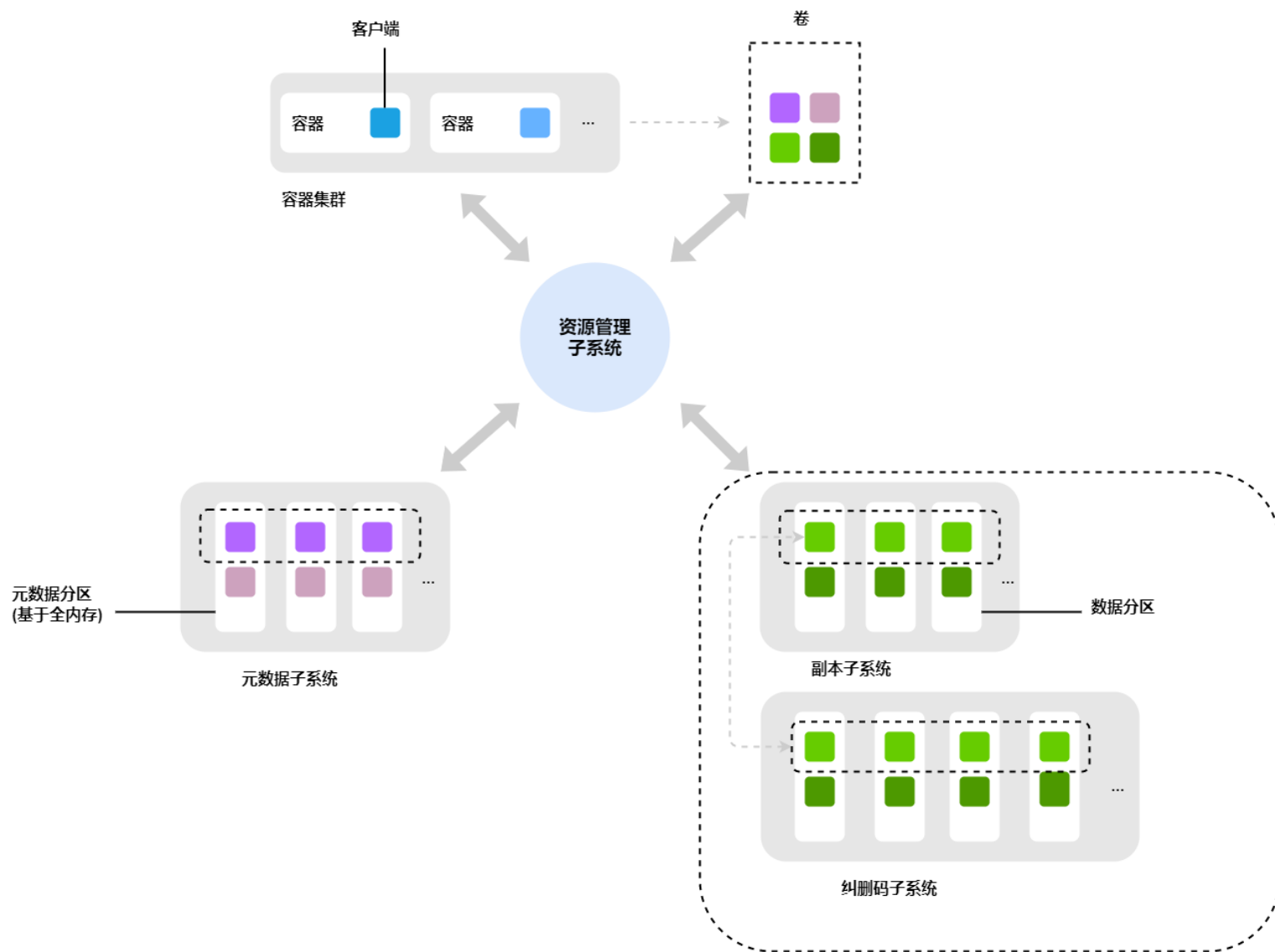
面临的挑战

1. 多租户和灵活的权限管理。
2. 数据和元数据都支持扩展。
3. 针对大/小文件进行优化存储。
4. 支持POSIX、HDFS、S3等多种文件系统接口。
5. 支持纠删码。
6. 多级缓存。



我们的优势





卷: 客户端可以从容器访问数据的文件系统。

数据分区: 文件数据分片的最小管理单位。

元数据分区: 文件元信息的最小管理单位。

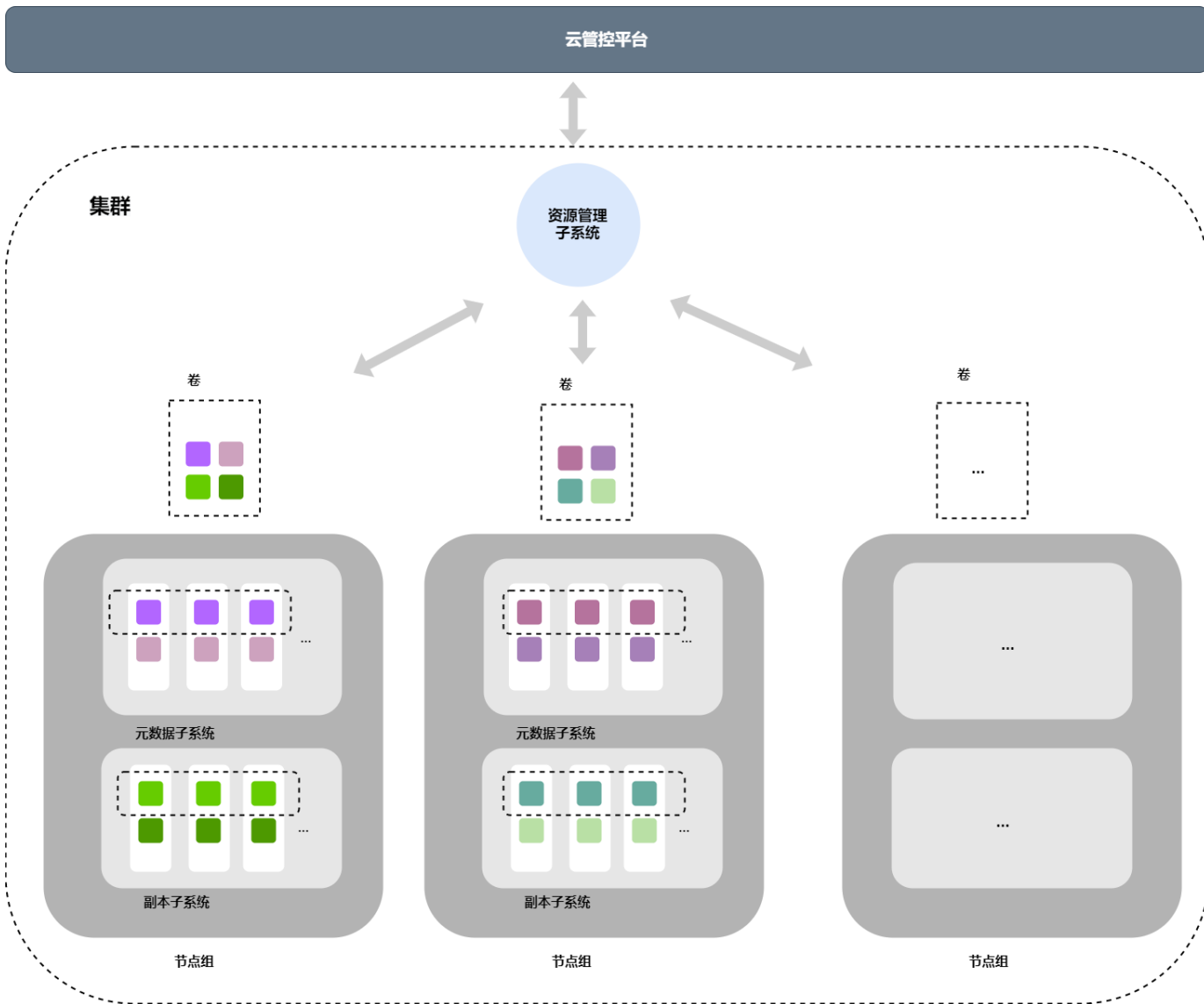
资源管理子系统: 用于管理集群中的节点资源。

元数据子系统: 管理集群中的元数据分区。

数据子系统: 管理集群中的温数据分区。

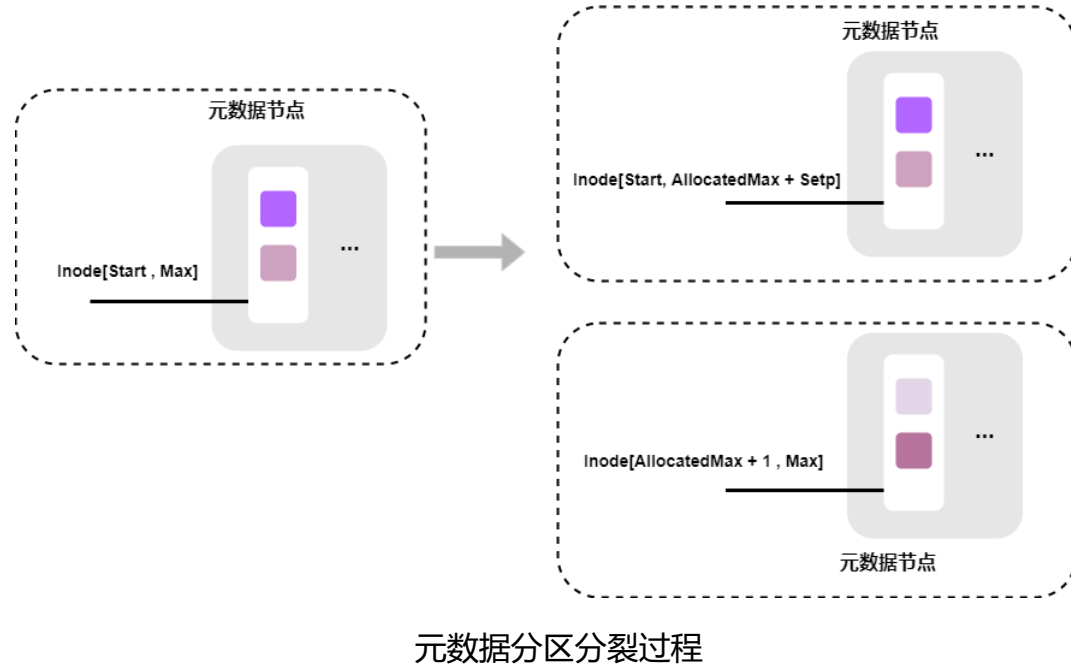
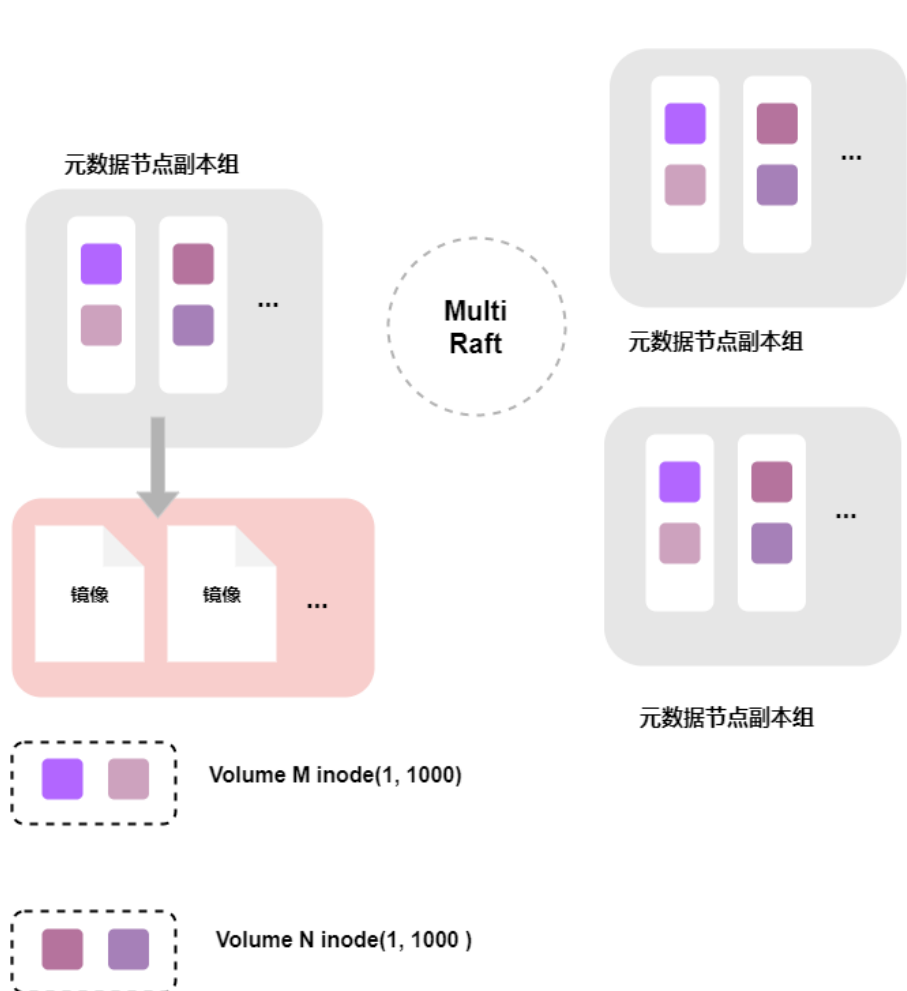
纠删码子系统: 管理集群中的冷数据分区。

客户端: 提供挂载文件系统的访问接口。



丰富的资源管理接口: 极大提高了集群的资源管理效率, 降低运维难度。

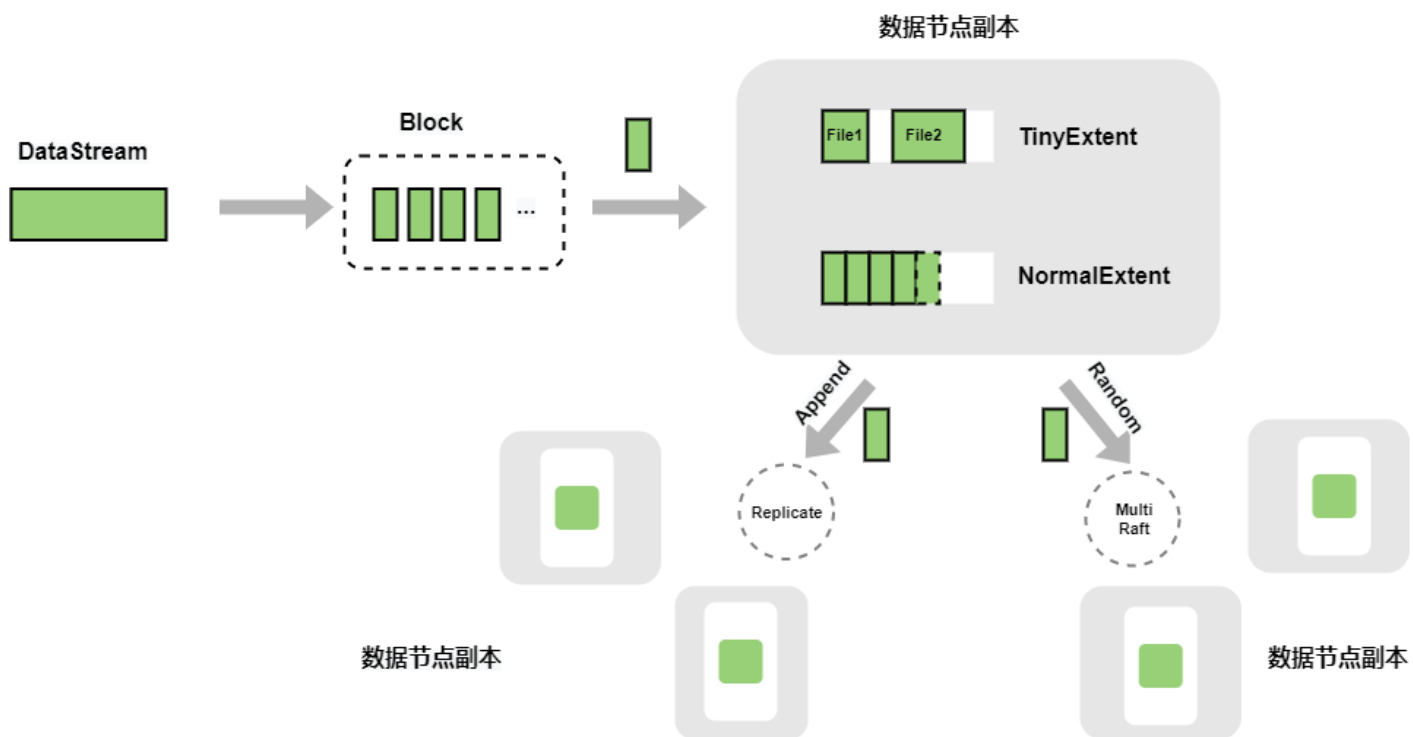
基于利用率的分布策略: 根据节点的内存/磁盘使用情况放置元数据和数据内容, 避免新节点加入时引入的数据迁移。



全内存缓存策略：提高元数据的访问速度。

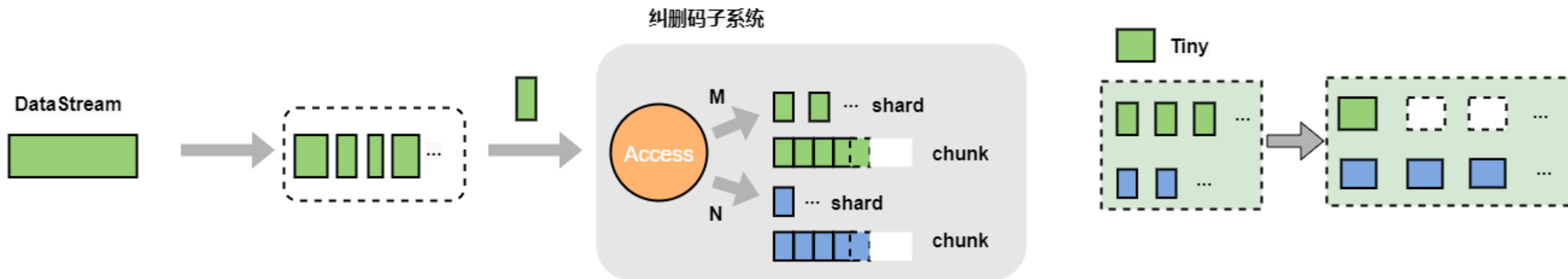
定期快照：元数据以分区为单位定期持久化到磁盘，用于备份和回复。

元数据分区分裂：通过将拆分元数据分区的管理范围，实现动态扩容。



大/小文件存储优化:通过分片/聚合的方式存储。

场景感知复制: 副本之间根据不同的写入方式采取不同的复制策略, 提高复制效率。

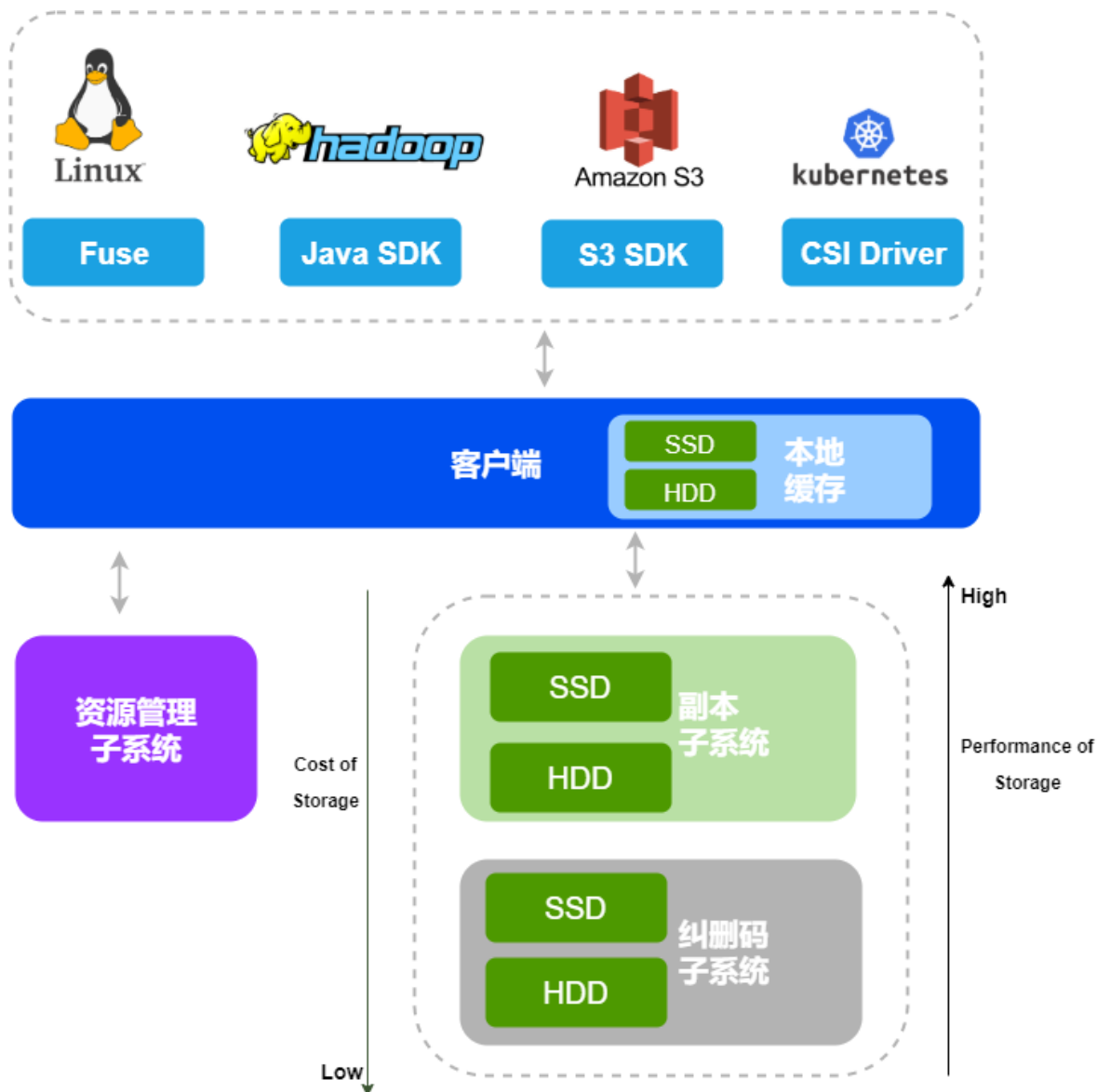


Quorum机制:允许一定写入失败，有效解决拖尾时延问题。

小文件EC优化:以空间换时间，提高读取性能。

高效的垃圾回收:通过稀疏语义回收空间，减少IO开销。

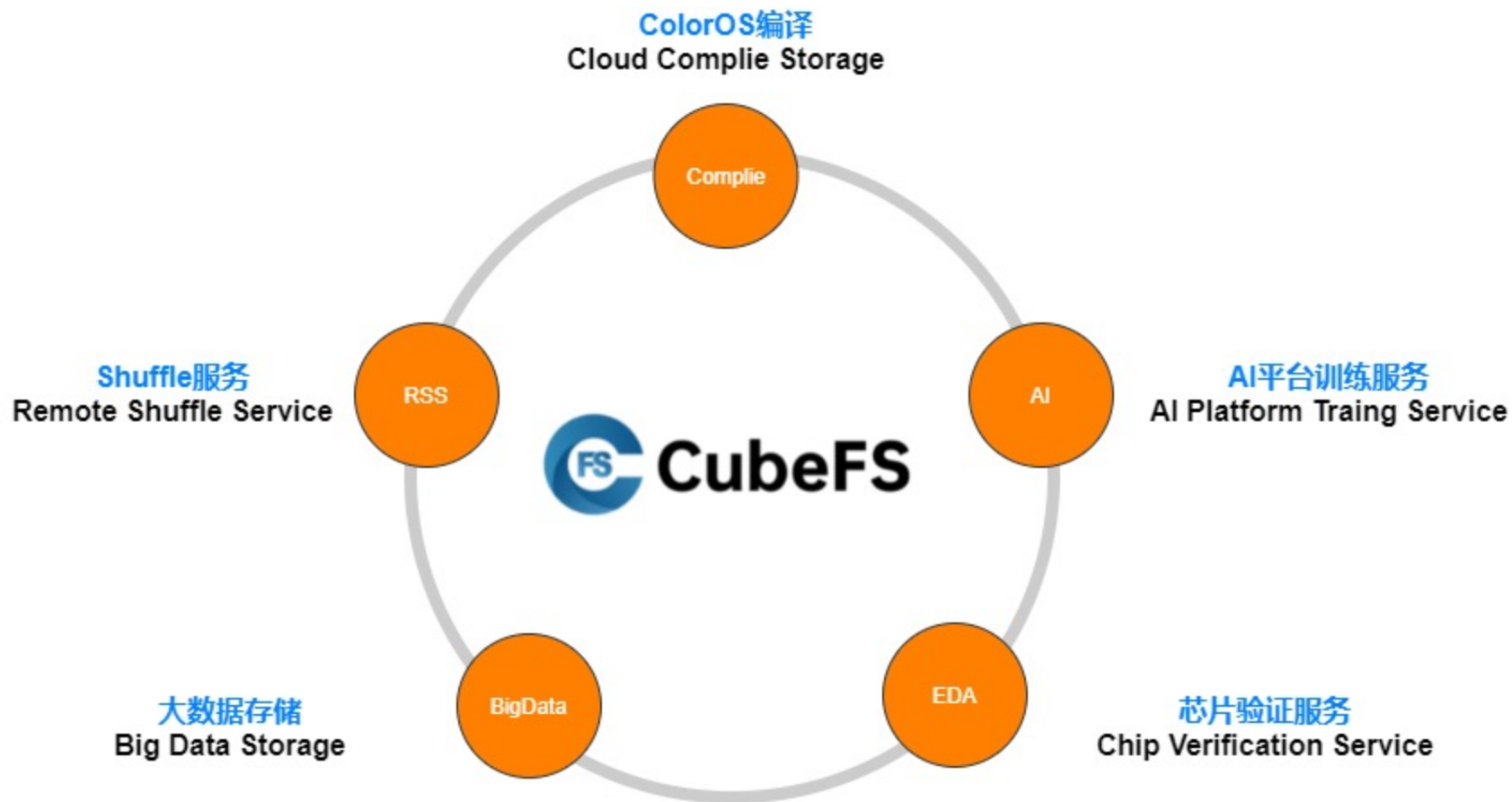
客户端架构



丰富的文件系统接口:支持Fuse, HDFS等应对不同的业务场景。

通信地址缓存:减少与其他子系统之间的通信。

多级缓存:多种缓存策略以提高数据读取效率。





海量小文件

训练任务的训练集通常都是从几KB到几MB的小文件，包含图片、视频、音频等，数据量规模亿级别。并且业务每年数据量呈现翻倍整张的趋势。



超大目录

训练任务的训练集或者测试集目录大小不可控，可以是百万级别或者千万级别。



热点目录

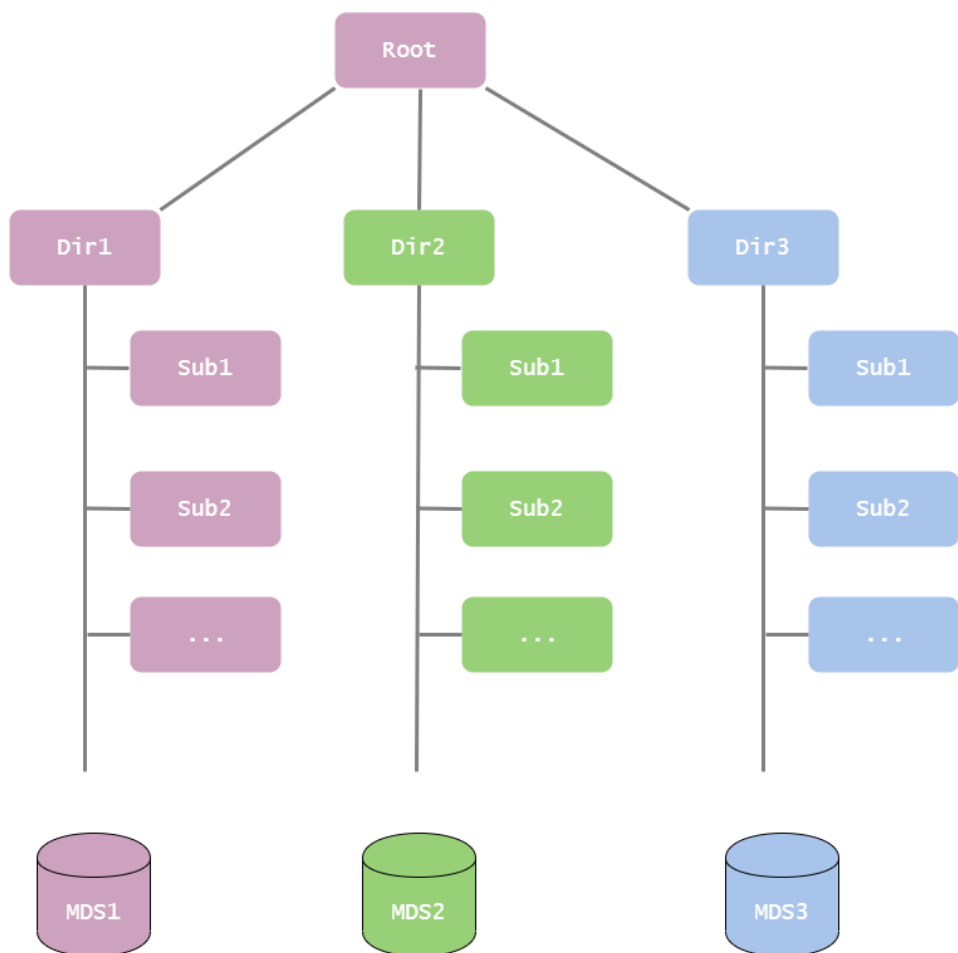
多用户并行训练任务对公共数据的访问，容易使得数据节点成为性能瓶颈，无法充分利用集群性能。



延时敏感

千万级别的文件规模对延时误差放大，最终导致训练时长过长，影响业务效率。

Ceph在机器学习平台的痛点



MDS性能:主备MDS无法线性扩容，高负载导致延时上升，降低集群GPU利用率；需要从Rados读取元数据再进行缓存，无法满足低延时的需求。

MDS稳定性:超大目录遍历容易OOM，单点故障定位困难大；拆分目录难度系数高，无法解决热点问题。

CubeFS的元数据痛点解决方案

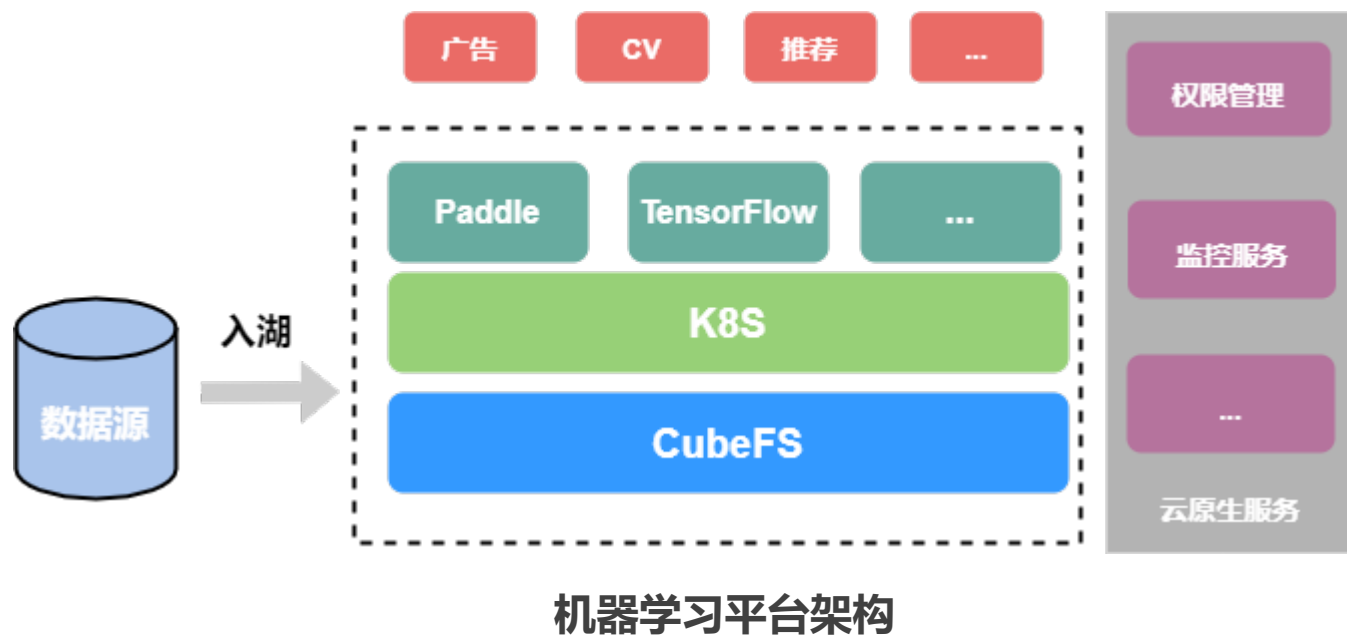


降低热点目录概率:元数据分区的选择和创建分布策略打散元数据。

元数据节点线性扩展:添加更多的元数据节点来创建元数据分区。

低延时:全内存缓存的元数据策略, 操作延时更低更平稳。

OPPO机器学习平台整体解决方案



项目收益



可用性

99.9%提升至**99.99%**



运维成本

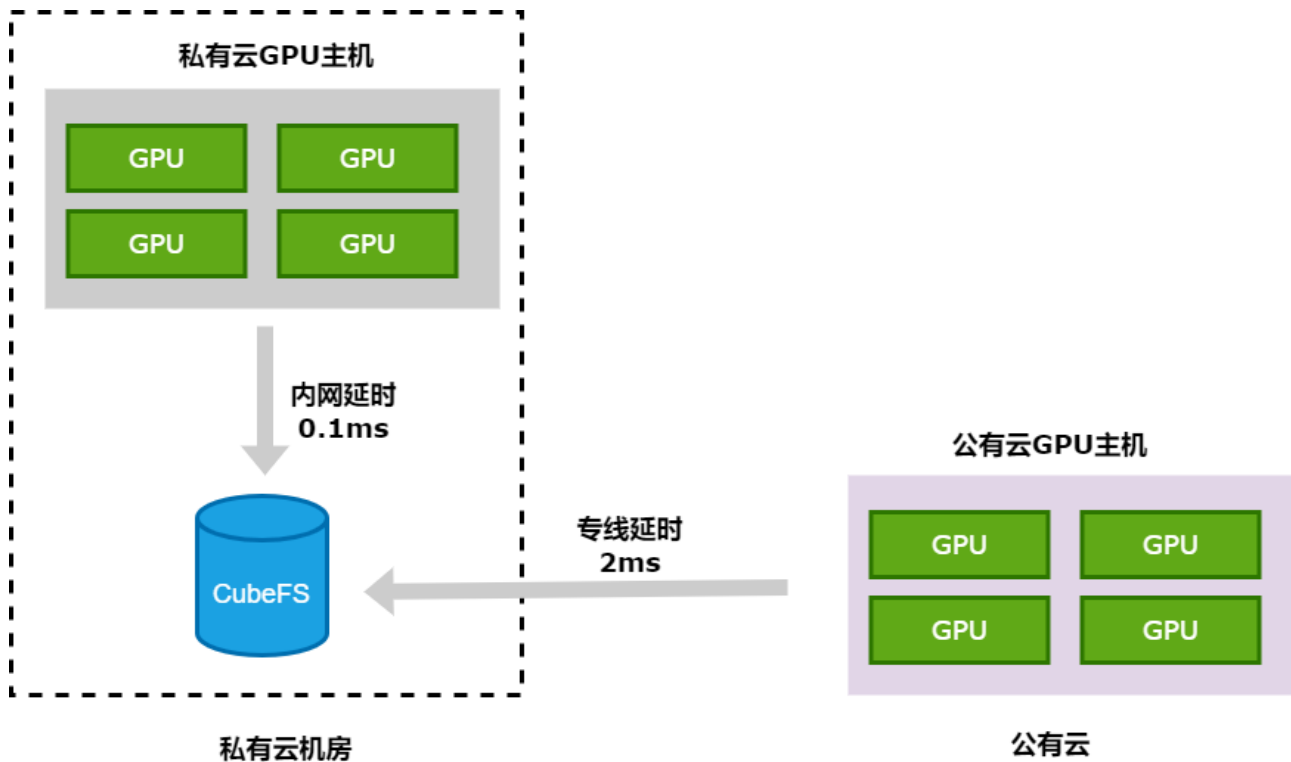
运维成本**大幅度降低**，扩容更
简便



元数据性能

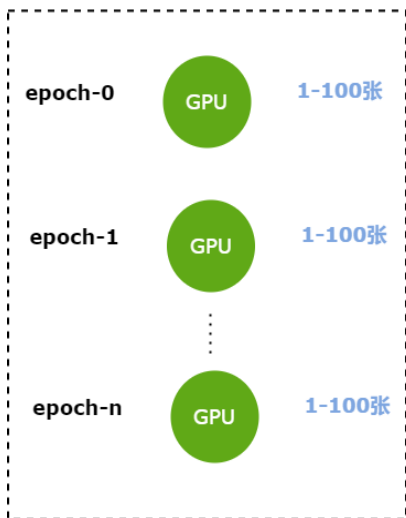
平均时延**缩短至1ms**

OPPO机器学习混合云的挑战

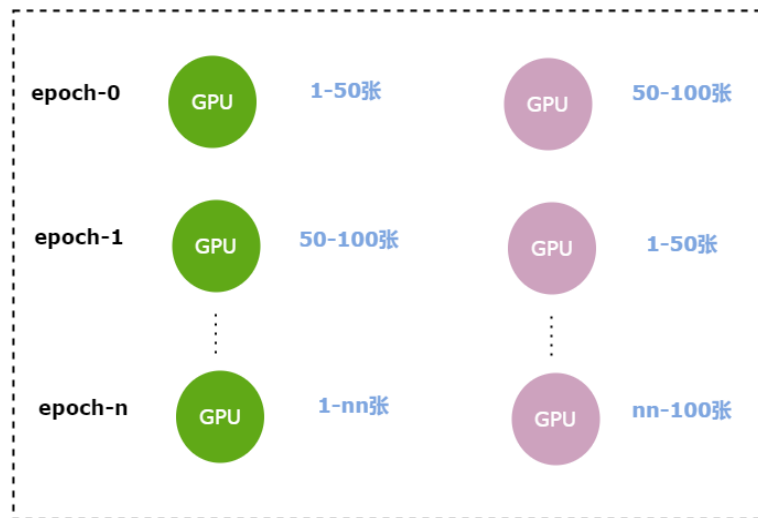


性能下降：高延时有AI海量文件训练特别是小文件不友好，相比私有云，公有云访问CubeFS性能**下降2-3倍**。

CubeFS缓存加速方案

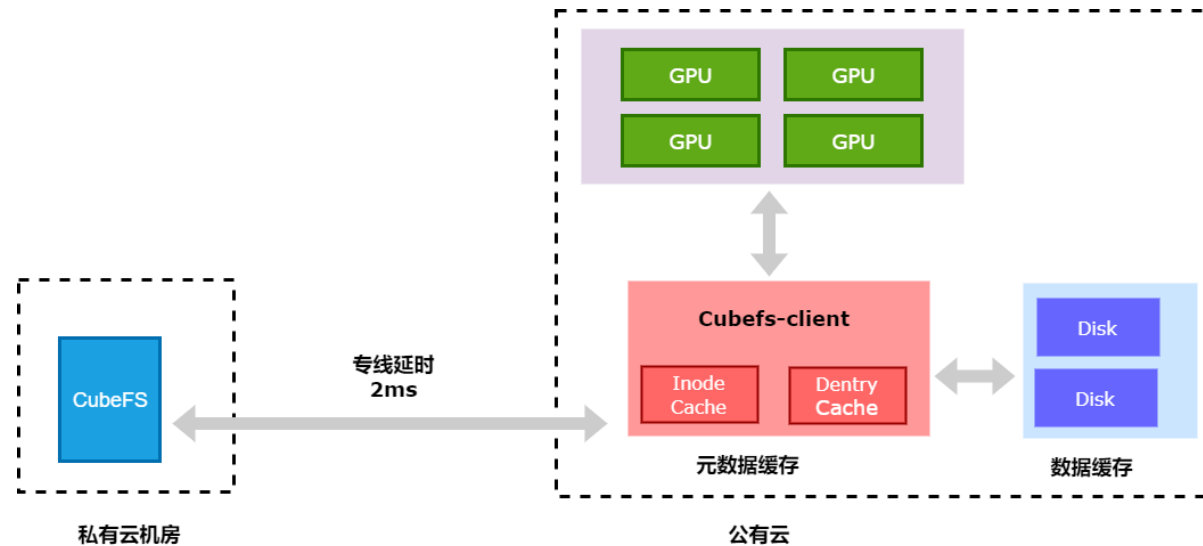


单机多卡训练任务



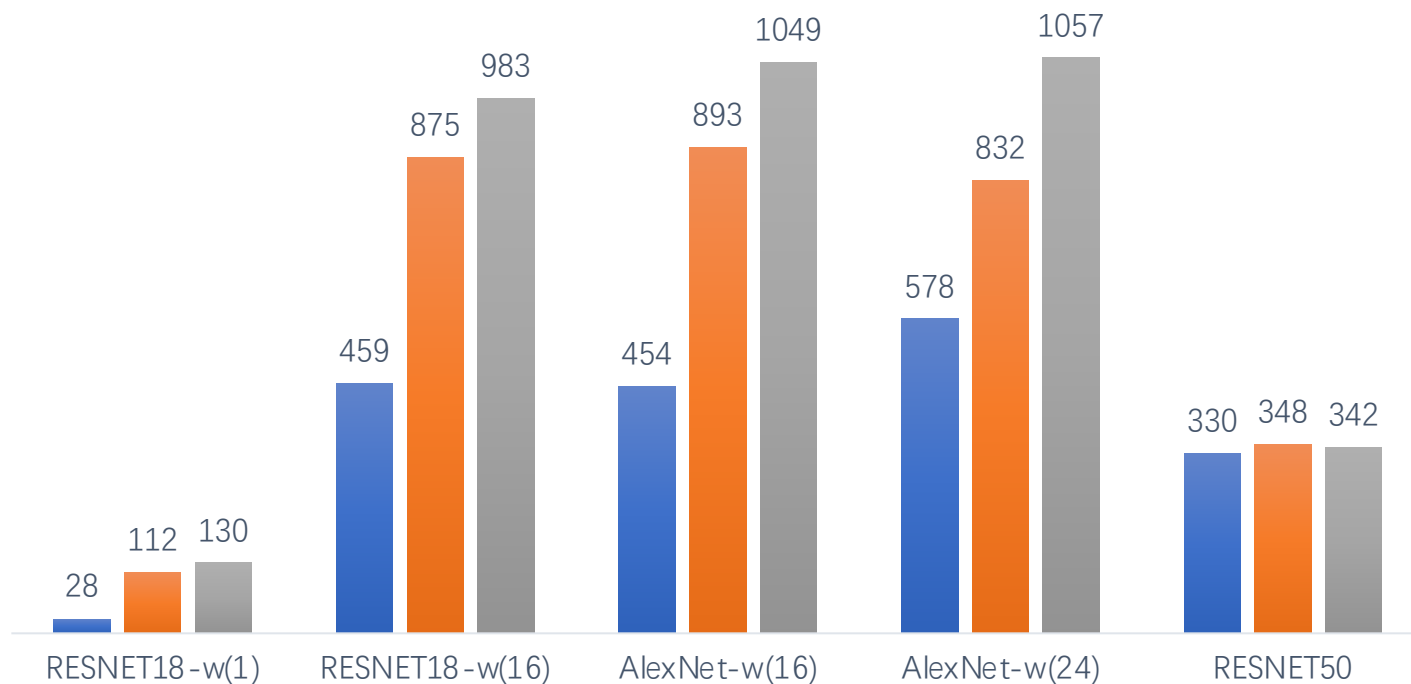
多机多卡训练任务

训练模型特点：针对某个数据集进行反复多次读取。



基准性测试

■ 公有云GPU-未加速 ■ 私有云GPU ■ 公有云GPU-加速



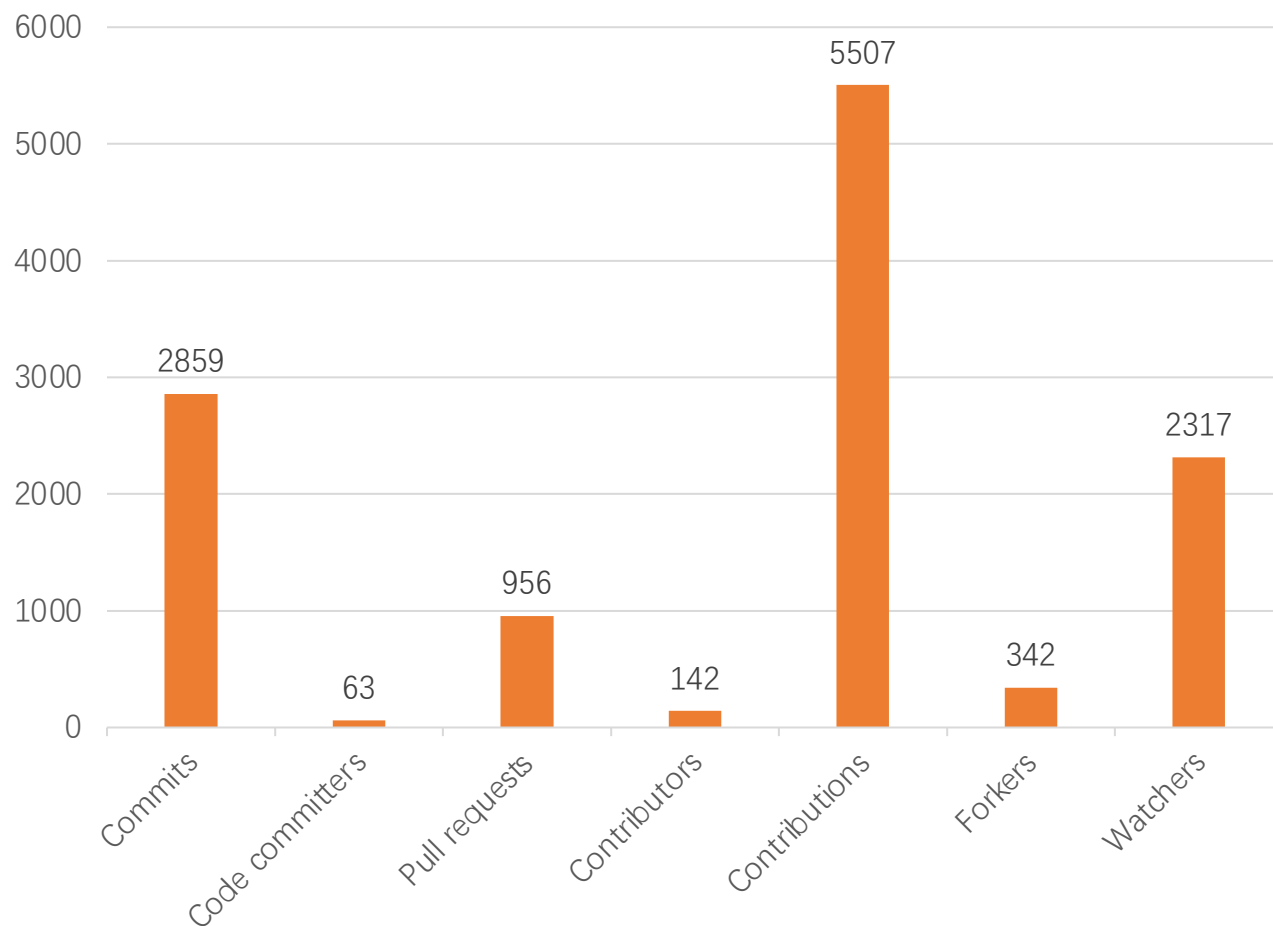
公有云开启CubeFS缓存

RESNET18在Dataloader worker=1、16时，性能分别提升**360%**、**114%**。

AlexNet在Dataloader worker=16、24时，性能分别提升**130%**、**80%**。

相比私有云部署，性能也有**12%-27%**的提升。

社区活跃度



合作伙伴



与众多项目生态集成



Prometheus



Kubernetes



Vitess



Helm



Harbor



Elastic Search



ClickHouse



KUBERNETES
COMMUNITY DAYS CHINA 2022

感谢观看

OPPO 贺迟