

Fluid JindoRuntime 新特性介绍

扬礼

2023-1-12

1、JindoRuntime 支持 PVC 挂载

2、JindoRuntime 性能优化

3、JindoFSx 5.0 规划



JindoRuntime 支持 PVC 挂载

JindoRuntime 支持 PVC 挂载



1、原理：

- 1、PVC 以 volume 的形式挂载到容器中
- 2、统一命名空间挂载容器内本地路径

3、优点：

支持更多存储，简化了访问云上和云下异构存储的过程，满足更广泛的缓存加速需求

2、使用方式：

```
apiVersion: data.fluid.io/v1alpha1
kind: Dataset
metadata:
  name: ceph
spec:
  mounts:
    - mountPoint: pvc://ceph-data
      name: ceph-data
```

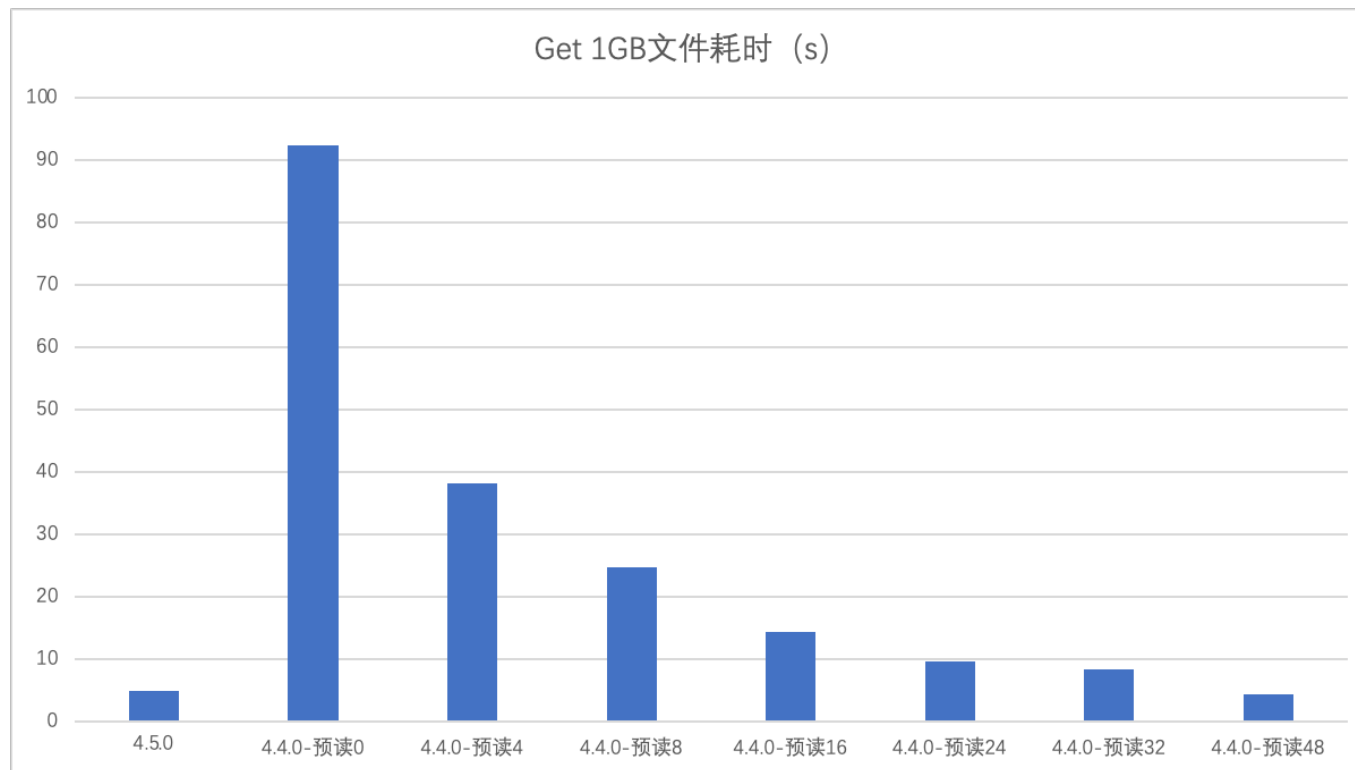
```
---
apiVersion: data.fluid.io/v1alpha1
kind: JindoRuntime
metadata:
  name: ceph
spec:
  replicas: 2
  tieredstore:
    levels:
      - mediumtype: MEM
        path: /dev/shm
        quota: 2Gi
        high: "0.95"
        low: "0.7"
```

JindoRuntime 性能优化

- 1、使用 low-level API 重新设计，大幅提高 readdir 等操作的性能。
- 2、使用 OSS-HDFS 进行 AI 训练最佳实践配置

```
fuse:
  properties:
    fs.oss.flush.interval.millisecond: "10000"
    fs.oss.randomwrite.sync.interval.millisecond: "10000"
  args:
    - -oattr_timeout=0
    - -oentry_timeout=0
    - -onegative_timeout=0
    - -ono_symlink
    - -ono_xattr
    - -ono_flock
    - -odirect_io
```

1、进一步改进自适应预读算法，提升预读效率



2、改进多副本读可能存在热点的问题，多副本访问随机均匀

JindoFSx 5.0 规划


- 1、支持更灵活的 dataset 定义，可配置路径级别的读写缓存策略（cacheset）
- 2、支持写缓存，支持多种写缓存策略
- 3、全新的分布式缓存加速策略，小文件训练场景性能杀手

预计 2023年2月发布敬请期待

JindoData 用户交流群

290人



 扫一扫群二维码，立刻加入该群。



Thank You!