



Fluid EFCRuntime 介绍

杨俊

2023-05-25

1、EFC介绍

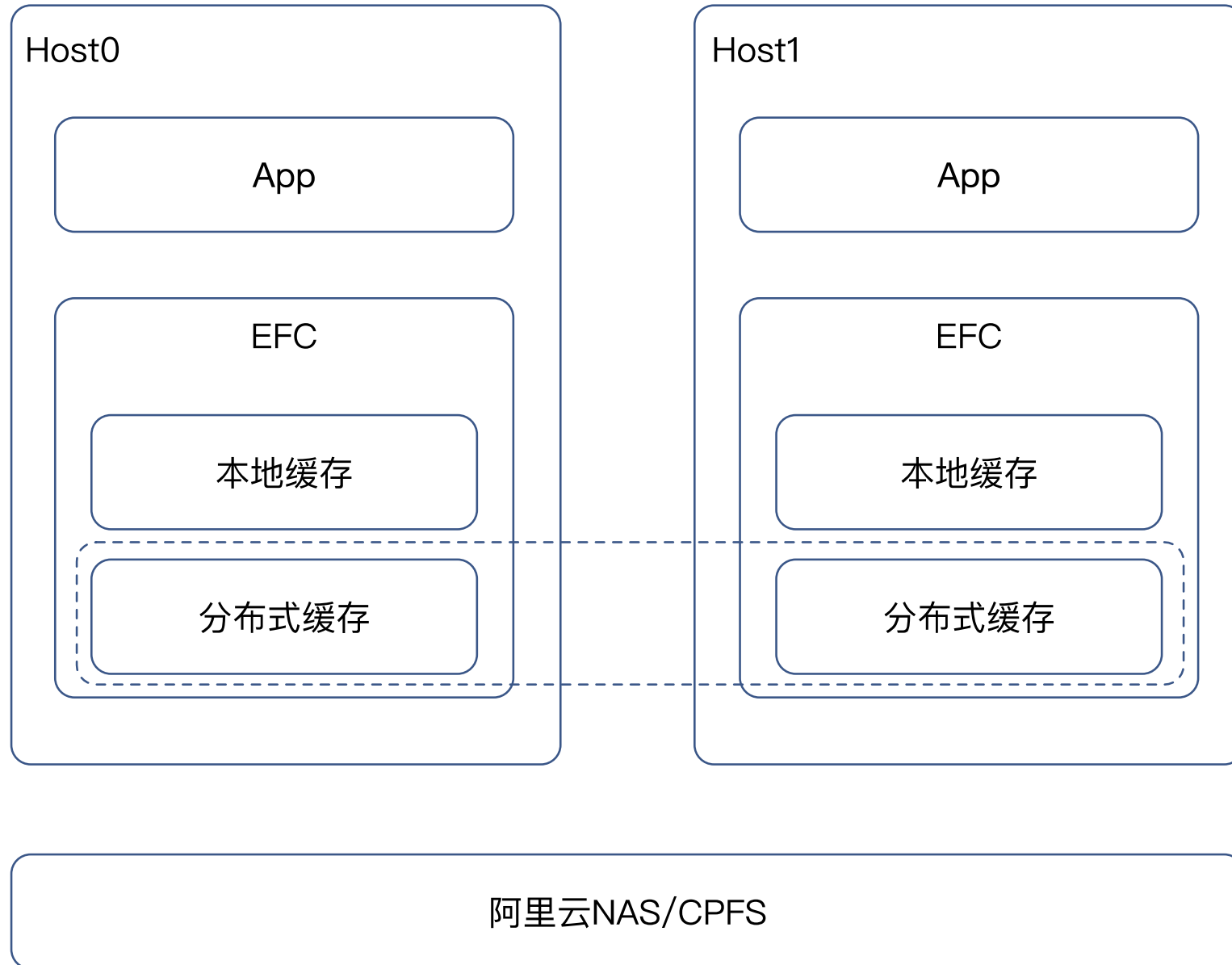
2、EFCRuntime介绍

3、EFCRuntime使用



EFC介绍

EFC架构



EFC弹性文件客户端（Elastic File Client）是阿里云文件存储团队开发的基于FUSE的用户态POSIX客户端。它可以替代传统的内核态NFS客户端，提供多链接访问、元数据缓存、分布式数据缓存等加速能力，并结合阿里云Prometheus监控提供端侧性能监控能力。

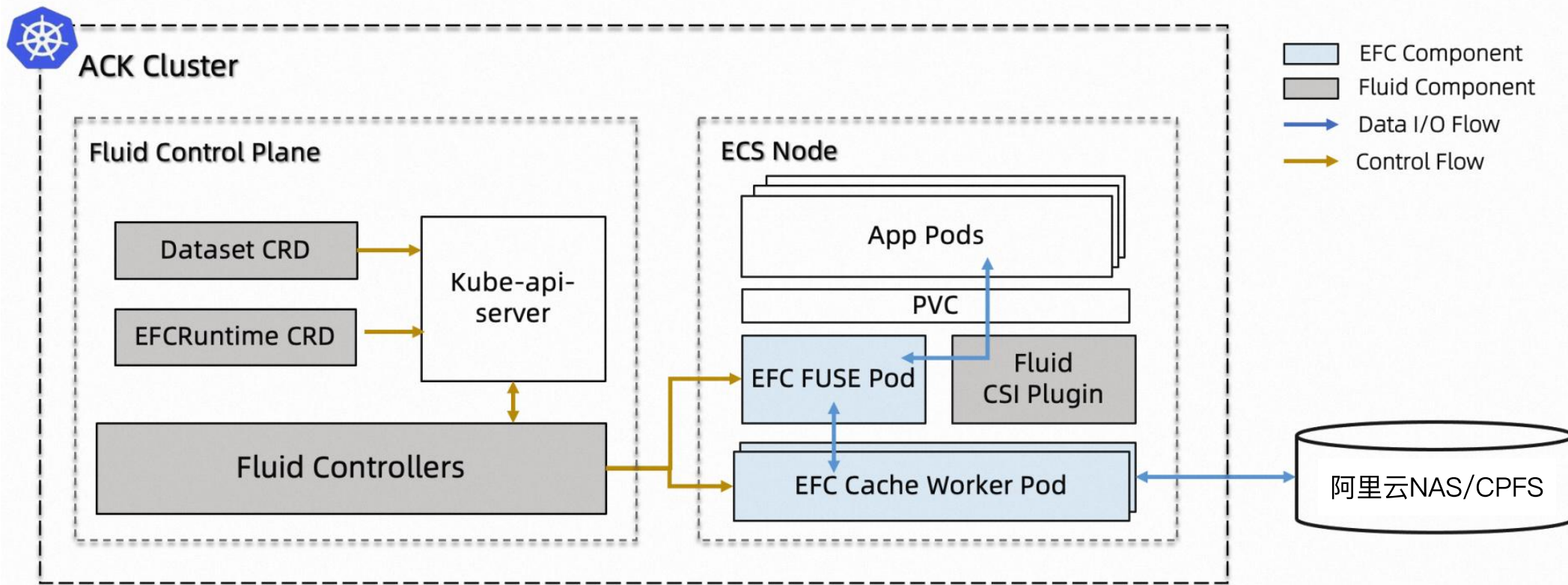
相比于内核态NFS v3、v4.x客户端和其他基于开源FUSE的客户端，EFC有以下优势：

- 强一致的语义：EFC通过强一致协议实现文件和目录的强一致；文件写入可以立刻被其他客户端读取；新文件创建出来后，就可以立刻让所有的其他客户端同步访问到，便于您在多节点间管理数据。
- 端上单机读写缓存：优化了FUSE的缓存逻辑，利用计算节点上的内存，提供了更好的小文件读写性能。相比于传统的NFS客户端，性能提升3-10倍。
- 分布式只读缓存：除了端上读写缓存，EFC提供分布式只读缓存的能力，利用多节点的内存建立缓存池，并随计算规模免运维自动扩展。
- 小文件预取：EFC会针对性的预取热目录下的数据，节省拉取数据的开销。
- 热升级和Failover：EFC支持秒级Failover特性，实现了对业务无影响的客户端版本的热升级能力。



EFCRuntime介绍

EFCRuntime架构





- Dataset: 由Fluid定义的一种Kubernetes CRD。它描述了逻辑上相关的一组数据的集合，会被上层运算引擎使用。
- EFCRuntime: 支撑Dataset访问加速能力的一种Runtime类型实现，其背后使用的缓存引擎为EFC。EFC缓存引擎包括EFC Fuse组件和EFC Cache Worker组件。
- EFC Fuse: 以POSIX协议标准暴露数据访问接口的EFC客户端组件。
- EFC Cache Worker: 通过一致性哈希提供缓存能力的服务端组件，根据不同的数据访问需求可选择关闭该组件。关闭该组件后，EFC将关闭分布式只读缓存功能。其他特性不受影响。
- EFC Master: 用户通过Fluid的API管理Fuse和Cache Worker缓存的组件。



EFCRuntime使用

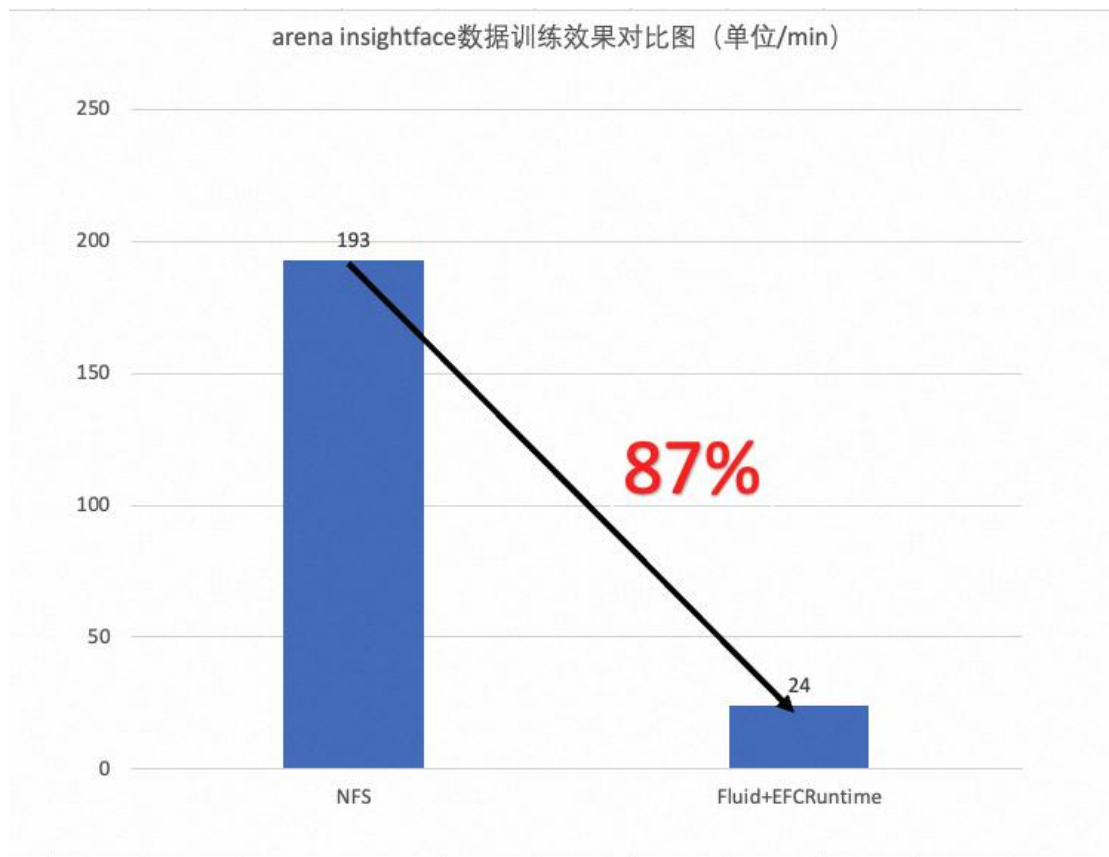
EFC支持“阿里云NAS/CPFS”



```
1  apiVersion: data.fluid.io/v1alpha1
2  kind: Dataset
3  metadata:
4    name: mydemo
5  spec:
6    mounts:
7      - mountPoint: "nfs://xxx.region.nas.aliyuncs.com/"
8    ---
9  apiVersion: data.fluid.io/v1alpha1
10 kind: EFCRuntime
11 metadata:
12   name: mydemo
13 spec:
14   replicas: 3
15   tieredstore:
16     levels:
17       - mediumtype: MEM
18         volumeType: emptyDir
19         path: /mnt/efc-worker-cache-path
20         quota: 10Gi
21   fuse:
22     properties:
23     master:
24       # disabled: true
25       networkMode: ContainerNetwork
26     # podMetadata:
27     # labels:
28     #   alibabacloud.com/eci: "true"
29     # annotations:
30     #   k8s.aliyun.com/eci-use-specs: ecs.g7.4xlarge
31   worker:
32     # disabled: true
33     networkMode: ContainerNetwork
34     # podMetadata:
35     # labels:
36     #   alibabacloud.com/eci: "true"
37     # annotations:
38     #   k8s.aliyun.com/eci-use-specs: ecs.g7.4xlarge
```

```
1  apiVersion: data.fluid.io/v1alpha1
2  kind: Dataset
3  metadata:
4    name: mydemo
5  spec:
6    mounts:
7      - mountPoint: "cpfs://cpfs-xxx-xxx.region.cpfs.aliyuncs.com:/share/"
8    ---
9  apiVersion: data.fluid.io/v1alpha1
10 kind: EFCRuntime
11 metadata:
12   name: mydemo
13 spec:
14   replicas: 3
15   tieredstore:
16     levels:
17       - mediumtype: MEM
18         volumeType: emptyDir
19         path: /mnt/efc-worker-cache-path
20         quota: 10Gi
21   fuse:
22     properties:
23     master:
24       # disabled: true
25       networkMode: ContainerNetwork
26     # podMetadata:
27     # labels:
28     #   alibabacloud.com/eci: "true"
29     # annotations:
30     #   k8s.aliyun.com/eci-use-specs: ecs.g7.4xlarge
31   worker:
32     # disabled: true
33     networkMode: ContainerNetwork
34     # podMetadata:
35     # labels:
36     #   alibabacloud.com/eci: "true"
37     # annotations:
38     #   k8s.aliyun.com/eci-use-specs: ecs.g7.4xlarge
```

EFCRuntime的性能表现



我们使用insightface数据集（包含约380万小文件）基于 Kubernetes 集群并使用 [Arena](#) 在此数据集上验证并发读取速度，基于 EFC 的 EFCRuntime 在开启本地缓存的情况下性能大幅度优于开源 NFS，训练耗时缩短了 87%，该测试场景会在后续文章中进行详细介绍。

Thank You!