JuiceFS 整合minio导出POSIX文件系统

官网手册: https://juicefs.com/docs/zh/community/introduction/

JuiceFS介绍

JuiceFS 是一款面向云原生设计的高性能共享文件系统,在 Apache 2.0 开源协议下发布。提供完备的 POSIX 兼容性,可将几乎所有对象存储接入本地作为海量本地磁盘使用,亦可同时在跨平台、跨地区的不同主机上挂载读写。

JuiceFS 采用「数据」与「元数据」分离存储的架构,从而实现文件系统的分布式设计。使用 JuiceFS 存储数据,数据本身会被持久化在<u>对象存储</u>(例如,Amazon S3),相对应的元数据可以按需持久化在Redis、MySQL、TiKV、SQLite 等多种数据库中。

JuiceFS 提供了丰富的 API,适用于各种形式数据的管理、分析、归档、备份,可以在不修改代码的前提下无缝对接大数据、机器学习、人工智能等应用平台,为其提供海量、弹性、低价的高性能存储。运维人员不用再为可用性、灾难恢复、监控、扩容等工作烦恼,专注于业务开发,提升研发效率。同时运维细节的简化,也让运维团队更容易向 DevOps 团队转型。

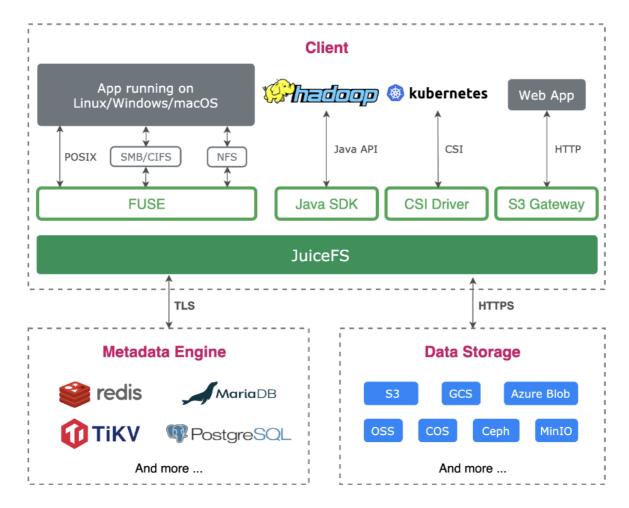
核心特性

- 1. POSIX 兼容: 像本地文件系统一样使用,无缝对接已有应用,无业务侵入性;
- 2. HDFS 兼容: 完整兼容 HDFS API, 提供更强的元数据性能;
- 3. **S3 兼容**: 提供 S3 网关 实现 S3 协议兼容的访问接口;
- 4. 云原生: 通过 CSI Driver 轻松地在 Kubernetes 中使用 JuiceFS;
- 5. 分布式设计: 同一文件系统可在上干台服务器同时挂载, 高性能并发读写, 共享数据;
- 6. 强一致性: 确认的文件修改会在所有服务器上立即可见, 保证强一致性;
- 7. 强悍性能: 毫秒级延迟, 近乎无限的吞吐量 (取决于对象存储规模), 查看性能测试结果;
- 8. **数据安全**: 支持传输中加密 (encryption in transit) 和静态加密 (encryption at rest) , <u>查看详</u> <u>情</u>;
- 9. 文件锁: 支持 BSD 锁 (flock) 和 POSIX 锁 (fcntl);
- 10. 数据压缩:支持 LZ4 和 Zstandard 压缩算法,节省存储空间。

技术架构

JuiceFS 文件系统由三个部分组成:

- 1. **JuiceFS 客户端**:协调对象存储和元数据存储引擎,以及 POSIX、Hadoop、Kubernetes CSI Driver、S3 Gateway 等文件系统接口的实现;
- 2. 数据存储: 存储数据本身, 支持本地磁盘、公有云或私有云对象存储、HDFS 等介质;
- 3. **元数据引擎**:存储数据对应的元数据 (metadata) 包含文件名、文件大小、权限组、创建修改时间和目录结构等,支持 Redis、MySQL、TiKV 等多种引擎



安装部署

这里测试部署选用的是etcd, etcd提供强一致的读写访问,所有操作都涉及到多机事务更新和数据持久化,性能肯定不能和redis作为元数据引擎相比,但redis作为元数据引擎只支持单机redis(或sentinel集群),有其自身局限,有数据丢失风险(即使开启rbd和AOF,fsync的默认间隔为1s,可能会有1s的元数据丢失风险)

环境说明

IP	服务	配置
192.168.251.102	minio/etcd	16C16G 2HDD* 80G
192.168.251.88	minio/etcd	16C16G 2HDD* 80G
192.168.251.98	minio/etcd	16C16G 2HDD* 80G
192.168.251.126	ECM 6.0、JuiceFS客户端	8C12G 200G

下载安装

```
JFS_LATEST_TAG=$(curl -s
https://api.github.com/repos/juicedata/juicefs/releases/latest | grep 'tag_name'
| cut -d '"' -f 4 | tr -d 'v')
wget
"https://github.com/juicedata/juicefs/releases/download/v${JFS_LATEST_TAG}/juice
fs-${JFS_LATEST_TAG}-linux-amd64.tar.gz"
tar -zxf "juicefs-${JFS_LATEST_TAG}-linux-amd64.tar.gz"
sudo install juicefs /usr/local/bin
```

etcd连接信息

```
[root@node1 minio]# ETCDCTL_API=3 etcdctl member list
6284eeb773af23cc, started, etcd3, http://192.168.251.98:2380,
http://192.168.251.98:2379
793ba0aaa9c9a810, started, etcd1, http://192.168.251.133:2380,
http://192.168.251.102:2379
80c8b20abbb616fc, started, etcd2, http://192.168.251.88:2380,
http://192.168.251.88:2379
```

minio连接信息

```
http://192.168.251.102:9000,http://192.168.251.88:9000,http://192.168.251.98:900
```

格式化文件系统

```
juicefs format \
    --storage minio \
    --bucket http://192.168.251.102:9000/edoc3 \
    --access-key 1KuNsQIqF91ZK1]E \
    --secret-key i86K9M4z1kQj1Ng33b2zjGzqxYhywA4y \
    etcd://192.168.251.102:2379,192.168.251.88:2379,192.168.251.98:2379/jfs \
    myjfs
```

```
[root@andj opt]# juicefs format --storage minio --bucket http://192.168.251.102:9000/edoc3 --access-key 1KuNsQlqF91ZK1IE --secret-key i86K9M4zlkQj1Ng33bZZjGzqxYhywA4y et cd://192.168.251.102:2379, 192.168.251.102:2379, 192.168.251.98:2379/jfs [interface.go:400]
2822/07/13 15:23:38.27747 juicefs[18295] <INFO: Data use minio://192.168.251.102:2900/edoc3/myjfs/ [format.go:437]
2822/07/13 15:23:38.277747 juicefs[18295] <INFO: Volume is formatted as {
    "Name": "myjfs",
    "ULID": "A79C5c38-ad59-4573-a251-7b16adace00a",
    "Storage": "minio",
    "Bucket": "http://192.168.251.102:9000/edoc3",
    "AccessAey": "AkusQufe91ZK1IE",
    "SecretKey": "removed",
    "Blocksize": 4096,
    "Compression": "none",
    "KeyfEncythed": true,
    ""TrashDays": 1,
    "MetaVersion": 1
    [format.go:474]
```

格式化一个名为myifs的文件系统,元数据引擎使用的是etcd,数据引擎使用的是minio

文件系统挂载到本地目录

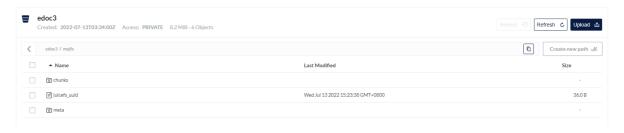
```
juicefs mount \
    --background \
    --cache-dir /mycache \
    --cache-size 512 \
    etcd://192.168.251.102:2379,192.168.251.88:2379,192.168.251.98:2379/jfs \
    /mnt/jfs
```

指定缓存目录 /mycache,缓存的大小512mb,指定元数据引擎地址,挂载位置为本机 /mnt/jfs

```
[root@and] mycache]# ps -ef | grep juicefs
root 4992 14132 0 15:32 pts/1 00:00:00 grep --color=auto juicefs
root 39243 1 0 15:30 ? 00:00:00 juicefs mount --background --cache-dir /mycache --cache-size 512 etcd://192.168.251.102:2379,192.168.251.88:2379,192.168.251.98:2379/jfs /mnt
//fs
[root@and] mycache]# pwd
/mycache
[root@and] mycache]# 15
470:56:38-ad50-4573-a251-7516adace00a
[root@and] mycache]# 15 /mnt/jfs/
[root@and] mycache]# 15 /mnt/jfs/
```

测试文件上传

已用 可用 已用% 挂载点 1516 466 77% / 0 5.76 0% /dev 9 5.76 0% /dev/shm 9.7M 5.76 1% /run 0 5.76 0% /sys/fs/cgroup 145M 876M 15% /boot 5.7G 0 5.7G 5.7G 0 5.7G 5.7G 9.7M 5.7G 5.7G 0 5.7G 1014M 145M 870M levtmpfs tmpfs dev/vda1 mpfs overlay 1.2G 197G 0 151G 1.2G 46G 0% /run/user/0 77% /var/lib/docker/overlay2/58eb37e57bda6f0fa9375b28e17699437494c0e90ef76e verlay 197G 197G 151G 151G 46G 77% /var/lib/docker/overlay2/7154ba784fd96b94f60a380e18e9772141a4b8a58ee1507ca4310301282bf2db/merged 77% /var/lib/docker/overlay2/de3dc8cd554ff1c3b2c1053c78274ccba36e09c4cd530bae16df66676d024a18/merged //% /Var/lib/docker/overlay2/de3dc8dc354ff1c3b2c1b3c2R2/4CcDa3beb9c4cCd33bbae1a0ffbbb/6db2/4418/mmggd /7% /Var/lib/docker/overlay2/372a8e649f9016dadee8a0e40008fb65e69fd38f47229ff9c96be23b3caa73062/merged /7% /Var/lib/docker/overlay2/372a8e649f9016dadee8a0e40008fb65e69fd38f47229ff9c9be23b3caa73062/merged /7% /Var/lib/docker/overlay2/ad7c0df6b3ae7lbd6fbb4b2f24edf084b414843a65d3a795e553b0348ae9g37/merged /7% /Var/lib/docker/overlay2/30x59952bc671aa81337256673c0a1733d24474fld05ebe13eeC716f5cbe5615/merged /7% /Var/lib/docker/overlay2/b8201f2b9e55f11ddcca1e2364afae8aca7edc20b91a8387e07d5ad35e02da14/merged 151G 151G 151G verlay 197G 197G 197G verlav 197G 197G 151G 151G verlay 77% /var/lib/docker/overlay2/b820ff2b9e55f11ddcca1e2364afae8aca7edc20b91a8387e07d5ad35e02da14/merged
77% /var/lib/docker/overlay2/77de47466b425e8d2819397bcd5e8e9727a0ff6e043-6291a631235d82805d/merged
77% /var/lib/docker/overlay2/d69783733299ba78e395fe1c19855c5c18c85097e113f830ee96606516c106c8/merged
77% /var/lib/docker/overlay2/dc4aca2281b7a2fa03435cdf74f1b6dff9a98d5c98f5bf034730efac25a42d65/merged
77% /var/lib/docker/overlay2/fol101dsf799f473f958e7a97f4e531c78ec40b4470933bf19fb57e0e71c3d3758ea8f/merged
77% /var/lib/docker/overlay2/753eaf9bf9f7deda790abdf21771065a28bc582087da9a7ee9dfb5407e7fdf2f/merged
77% /var/lib/docker/overlay2/2a573d35881d55ead7ce6a13892c28b5c25bb50c096e53732cbb4ca7cb5d88ea/merged
77% /var/lib/docker/overlay2/74d6f86c21283e7d95ac8ba1c5bf683e9c33ef2162096511b737a03cf41a2133/merged
77% /var/lib/docker/overlay2/74d6f86c21283e7d95ac8ba1c5bf683e9c33ef2162096511b737a03cf41a2133/merged
77% /var/lib/docker/overlay2/7e02e065dd2b9b902148d32b03ebbe8da5eaaac0c2d39e3e7bda6f946f8ad5f/merged 151G 151G 151G 151G 151G 151G 151G verlay verlay 197G 197G 46G 46G verlay verlay 197G 197G 46G verlay 197G 197G 197G 46G 46G verlay verlay verlay 197G 197G 151G 151G 46G 46G verlay uiceFS:myjfs 197G 1.0P 151G 23M 46G 1.0P 77% /var/lib/docker/overlay2/5e1c7188c8e18ce65cab57b7741104f4e71b5701cc2d3088cb2b52ad4056e182/merged 1% /mnt/jfs



dedoc3 / myjfs / chunks / 0 / 0		Create new path ://o
☐ ▲ Name	Last Modified	Size
☐ ☑ 1_0_1769472	Wed Jul 13 2022 15:34:38 GMT+0800	1.7 MiB
☐ ☑ 10_0_1835008	Wed Jul 13 2022 15:35:23 GMT+0800	1.8 MiB
☐ ☑ 100_0_16384	Wed Jul 13 2022 15:39:26 GMT+0800	16.0 KiB
☐ ☑ 101_0_4096	Wed Jul 13 2022 15:39:26 GMT+0800	4.0 KIB
☐ ☑ 102_0_4194304	Wed Jul 13 2022 15:39:26 GMT+0800	4.0 MiB
☐ 102_1_892928	Wed Jul 13 2022 15:39:26 GMT+0800	872.0 KiB
☐ ☑ 103_0_290816	Wed Jul 13 2022 15:39:26 GMT+0800	284.0 KIB

ecm添加本地存储



126节点创建DefaultStorage2目录, juicefs挂载到此路径





测试上传没有问题

20并发压测 (etcd)

Requests		Executions					Throughput	Network (F	(B/sec)				
Label *	#Samples +	FAIL \$	Error % \$	Average \$	Min ¢	Max +	Median ¢	90th pct +	95th pct +	99th pct +	Transactions/s \$	Received ¢	Sent ¢
Total	2538	0	0.00%	786.90	26	6433	683.00	1417.30	1691.25	3016.48	25.23	15.72	7.07
企业库上传文件 信息_20并发	1259	0	0.00%	340.30	129	1560	290.00	542.00	654.00	1122.60	12.65	10.29	7.10
企业库上传文件 流_20并发	1259	0	0.00%	1245.15	455	6433	1135.00	1696.00	2181.00	3653.60	12.56	5.52	0.00
登录获取token	20	0	0.00%	53.80	26	126	33.50	126.00	126.00	126.00	25.25	5.92	6.13

Requests		Executions				1	Throughput	Network (H	(B/sec)				
Label 🔺	#Samples \$	FAIL \$	Error % \$	Average \$	Min ◆	Max \$	Median ¢	90th pct +	95th pct +	99th pct \$	Transactions/s \$	Received \$	Sent ¢
Total	8796	0	0.00%	225.56	10	1156	140.00	513.00	578.00	799.09	87.81	4407.89	9.25
下载检查_20并 发	4388	0	0.00%	24.61	10	888	21.00	30.00	37.00	109.33	44.10	52.07	9.26
文件下载_20并 发	4388	0	0.00%	427.14	90	1156	414.00	578.00	671.55	861.22	44.10	4385.68	0.00
登录获取token	20	0	0.00%	86.10	20	217	48.00	217.00	217.00	217.00	24.88	5.83	6.04

元数据引擎更换redis

```
juicefs format \
    --storage minio \
    --bucket http://192.168.251.102:9000/edoc3 \
    --access-key 1KuNsQIqF91ZK1lE \
    --secret-key i86K9M4zlkQj1Ng33b2zjGzqxYhywA4y \
    redis://:zgDsv0oY5Q@192.168.251.126:30005/2 \
    myjfsredis
```

```
juicefs mount \
   --background \
   redis://:ZgDSvOoY5Q@192.168.251.126:30005/2 \
   /home/edoc2/macrowing/edoc2v5/data/edoc2Docs/DefaultStorage2
```

20并发压测 (redis)

Requests		Executions					Throughput	Network (KB/sec)					
Label -	#Samples •	FAIL ¢	Error % •	Average ¢	Min ◆	Max ◆	Median ◆	90th pct ◆	95th pct ◆	99th pct •	Transactions/s •	Received Φ	Sent •
Total	5720	0	0.00%	347.64	24	1480	321.00	555.90	692.00	1033.43	57.13	35.73	16.05
企业库上传文件 信息_20并发	2853	0	0.00%	232.75	67	762	219.00	342.00	384.30	516.14	28.54	23.22	16.03
企业库上传文件 流_20并发	2847	0	0.00%	464.85	122	1480	418.00	692.00	853.20	1127.16	28.54	12.54	0.00
登录获取token	20	0	0.00%	49.95	24	99	40.00	99.00	99.00	99.00	24.57	5.76	5.96

Requests		Executions					Throughput	Network (K	(B/sec)				
Label 🔺	#Samples \$	FAIL \$	Error % \$	Average \$	Min ¢	Max ¢	Median	90th pct \$	95th pct \$	99th pct \$	Transactions/s \$	Received \$	Sent ¢
Total	5599	0	0.00%	356.53	8	1572	45.00	762.00	802.00	998.00	55.48	2771.92	5.87
下载检查_20并 发	2799	0	0.00%	699.33	18	1572	695.00	802.00	855.00	1232.00	27.79	32.81	5.84
文件下载_20并 发	2780	0	0.00%	13.38	8	146	12.00	16.00	18.00	27.19	28.04	2788.79	0.00
登录获取token	20	0	0.00%	78.95	21	195	40.50	195.00	195.00	195.00	24.72	5.79	6.00

可以看到上传下载的流的操作,比etcd有提升,这个下载检查和存储没多大关系。