New Thoughts On Distributed File System In The Cloud Native Era

Speaker: Shuoran Liu

Architect@JD.com & Maintainer@ChubaoFS

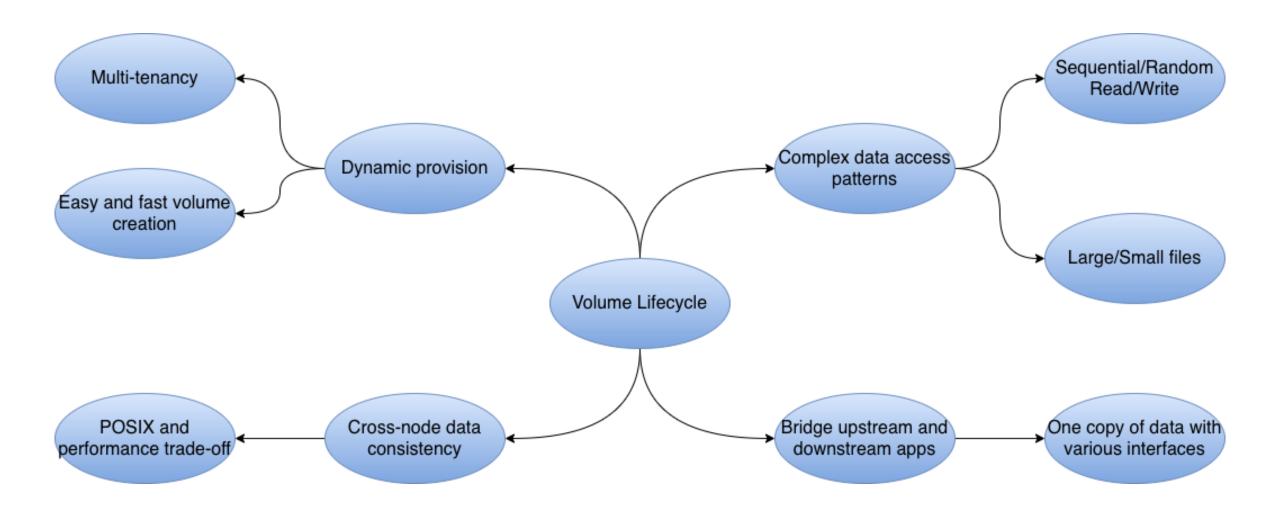
Start with CSI – volume lifecycle

https://github.com/container-storage-interface/spec/blob/master/spec.md

```
CreateVolume +----+ DeleteVolume
 ----->| CREATED +-----+
    Controller | | Controller
+++ Publish | Unpublish +++
  Volume | | Volume
|X|
   +---V----+
           | NODE_READY |
          Node | | Node
         Stage | Unstage
         Volume | | Volume
            VOL_READY |
          Node | Node
        Publish | Unpublish
         Volume | | Volume
            PUBLISHED
```



What We Got – A mind map



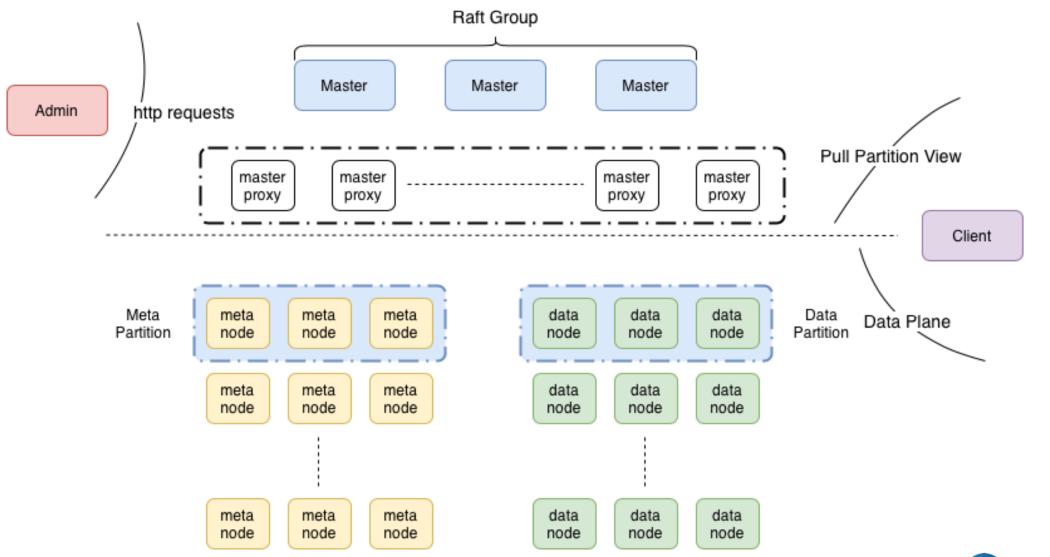


New Challenges - cloud native storage

- Multi-tenancy
- Elasticity, scalability and high availability
 - Avoid bottlenecks in a cluster
- Optimize for small files
 - High concurrent meta operations
 - Data aggregation and deletion
- POSIX compliance and performance trade-off
- One copy of data with various interfaces
 - Per upstream and downstream user needs



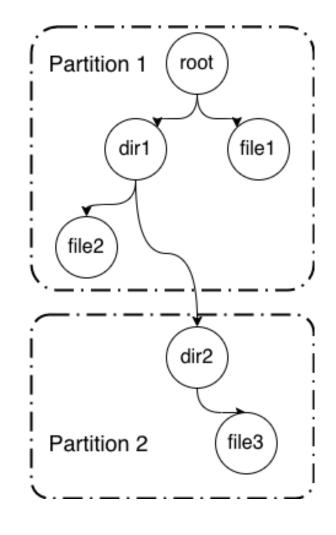
What does it look in a cluster

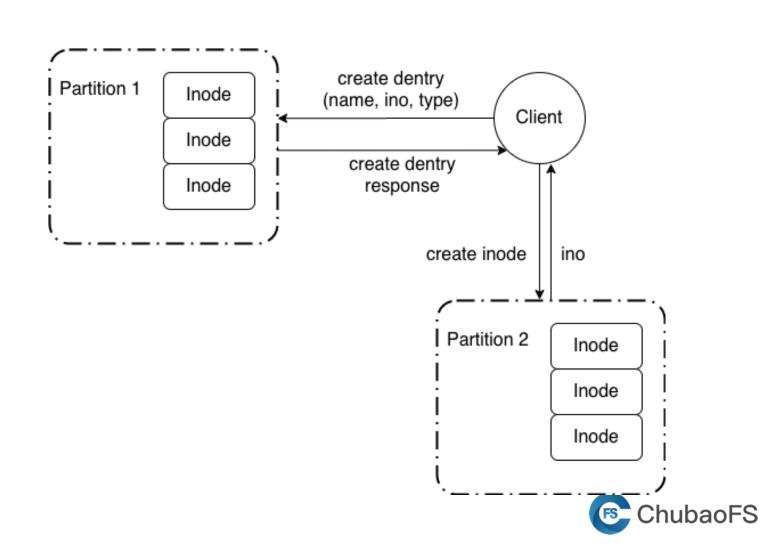




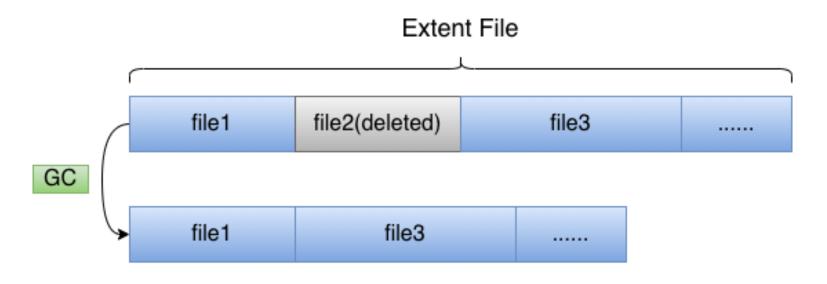
Small Files – meta concurrency

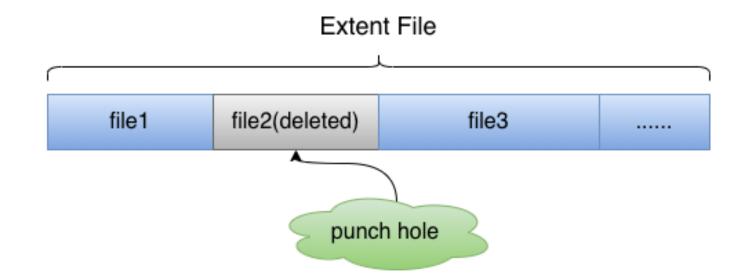
Locality vs Distribution





Small Files – data aggregation

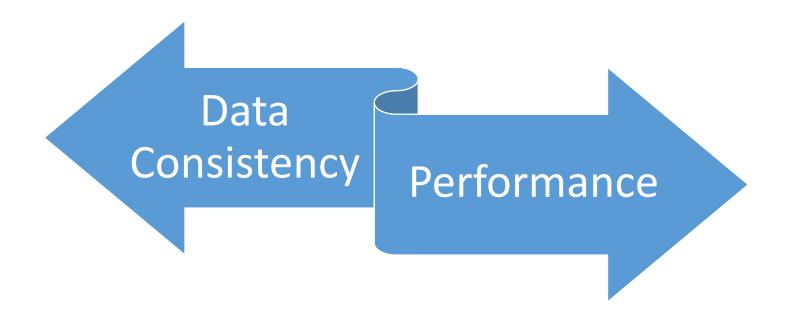






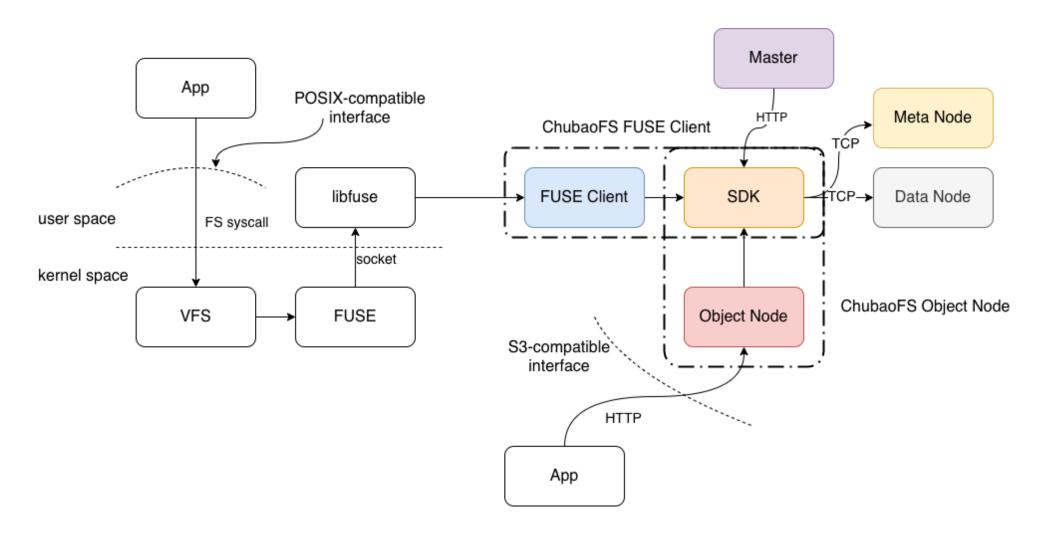
POSIX Compliance

To what extent shall we compromise?





Fusion Storage





Some other thoughts

- Benefit from low latency storage
 - Used for raft log storage to improve random write performance
- CSI plugin driver dilemma
 - Multiple volumes mounted in a single container



Thanks!

Github: github.com/chubaofs

Contact: @shuoranliu