

鹏云网络
陈靓博士

高性能
软件定义存储系统
架构设计

InfoQ官网 全新改版上线

促进软件开发领域知识与创新的传播



关注InfoQ网站
第一时间浏览原创IT新闻资讯



免费下载迷你书
阅读一线开发者的技术干货

极客邦科技 会议推荐2019

5月

QCon 北京

全球软件开发大会

大会：5月6-8日

培训：5月9-10日

QCon 广州

全球软件开发大会

培训：5月25-26日

大会：5月27-28日

6月

GTLC
GLOBAL
TECH LEADERSHIP
CONFERENCE

上海

技术领导力峰会

时间：6月14-15日

GMTC 北京

全球大前端技术大会

大会：6月20-21日

培训：6月22-23日

7月

ArchSummit

深圳

全球架构师峰会

大会：7月12-13日

培训：7月14-15日

10月

QCon

上海

全球软件开发大会

大会：10月17-19日

培训：10月20-21日

11月

GMTC

深圳

全球大前端技术大会

大会：11月8-9日

培训：11月10-11日

AiCon

北京

全球人工智能与机器学习大会

大会：11月21-22日

培训：11月23-24日

12月

ArchSummit

北京

全球架构师峰会

大会：12月6-7日

培训：12月8-9日



目录

COMPANY

01

SDS系统介绍

02

SDS系统数据流

03

经典SDS和鹏云系统架构比较

04

总结



大数据和云计算时代，数据呈指数形式增长。海量数据如何存储？

传统的磁盘阵列



成本高

削弱业务竞争力



扩展性差

难以支撑业务迅速增长



管理复杂

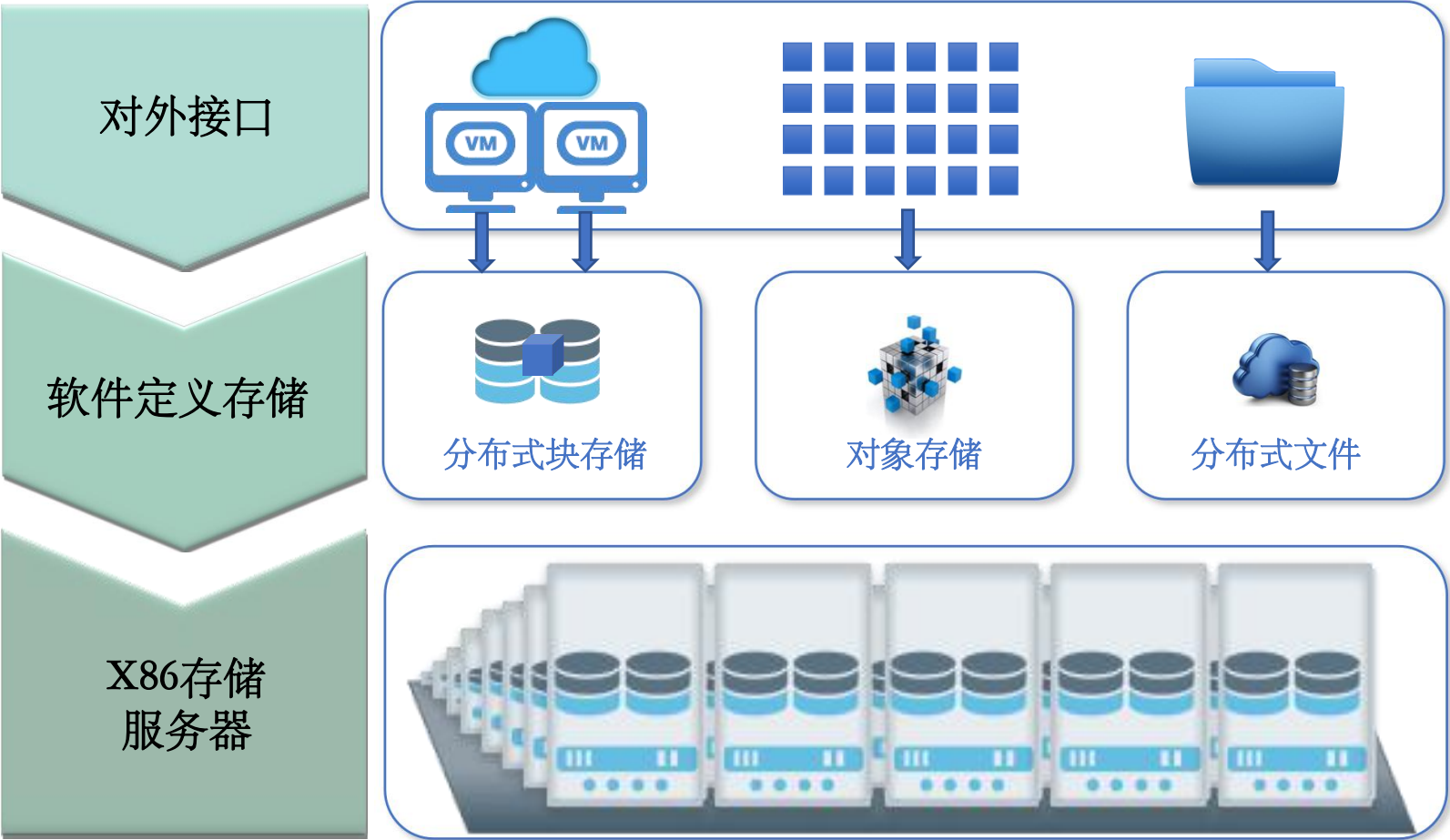
学习周期长，
无法快速响应业务系统需求



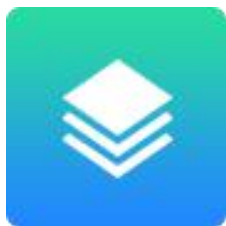
难运维

影响业务稳定

软件定义存储是什么



用软件定义存储 (SDS) 存储海量数据



SCALABILITY 灵活扩展

横向扩展，按需随时扩容
海量存储，支撑数据增长



AGILITY 高效敏捷

弹性伸缩，适应业务变化
按需供给，快速开通业务



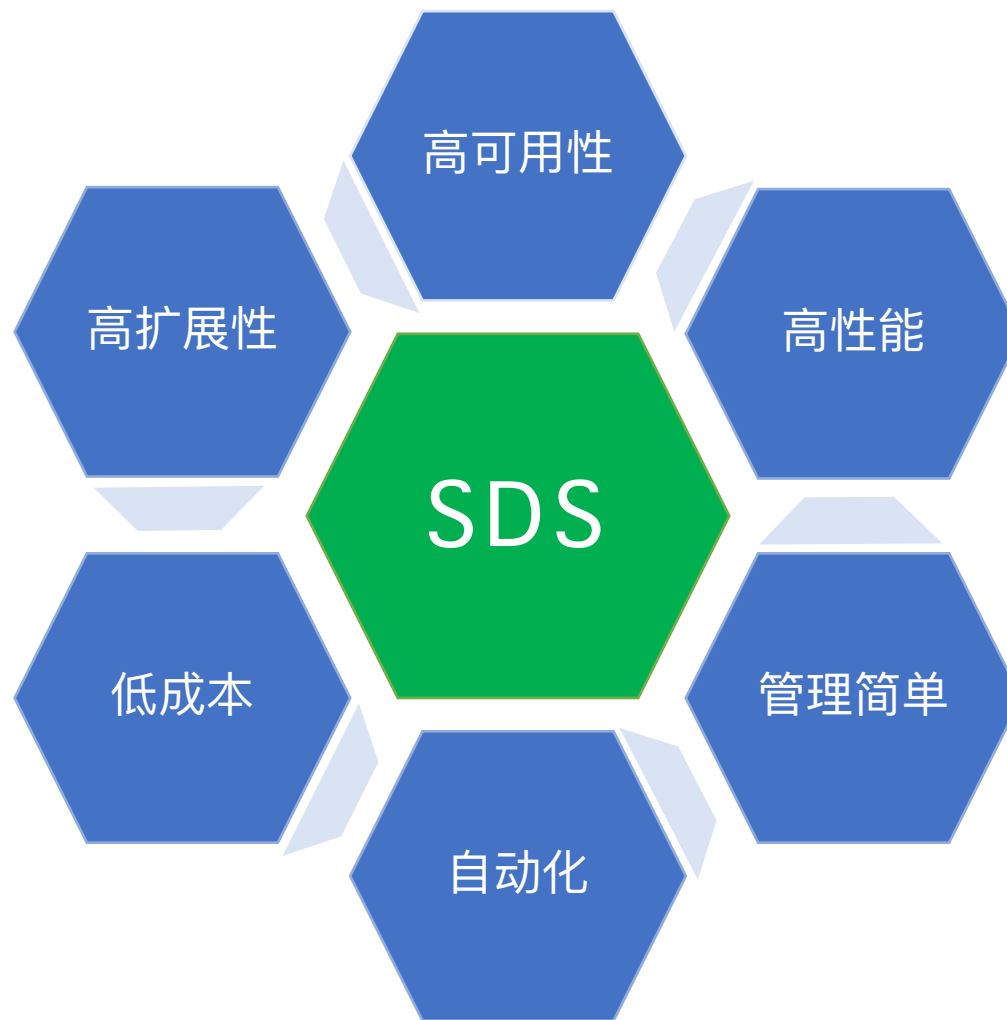
SIMPLICITY 管理简单

智能自愈，无需人工干预
架构简单，无高技能要求



LOW-COST 更低成本

通用硬件，更低建设成本
软件定义，避免厂商绑定





目录

COMPANY

01

SDS系统介绍

02

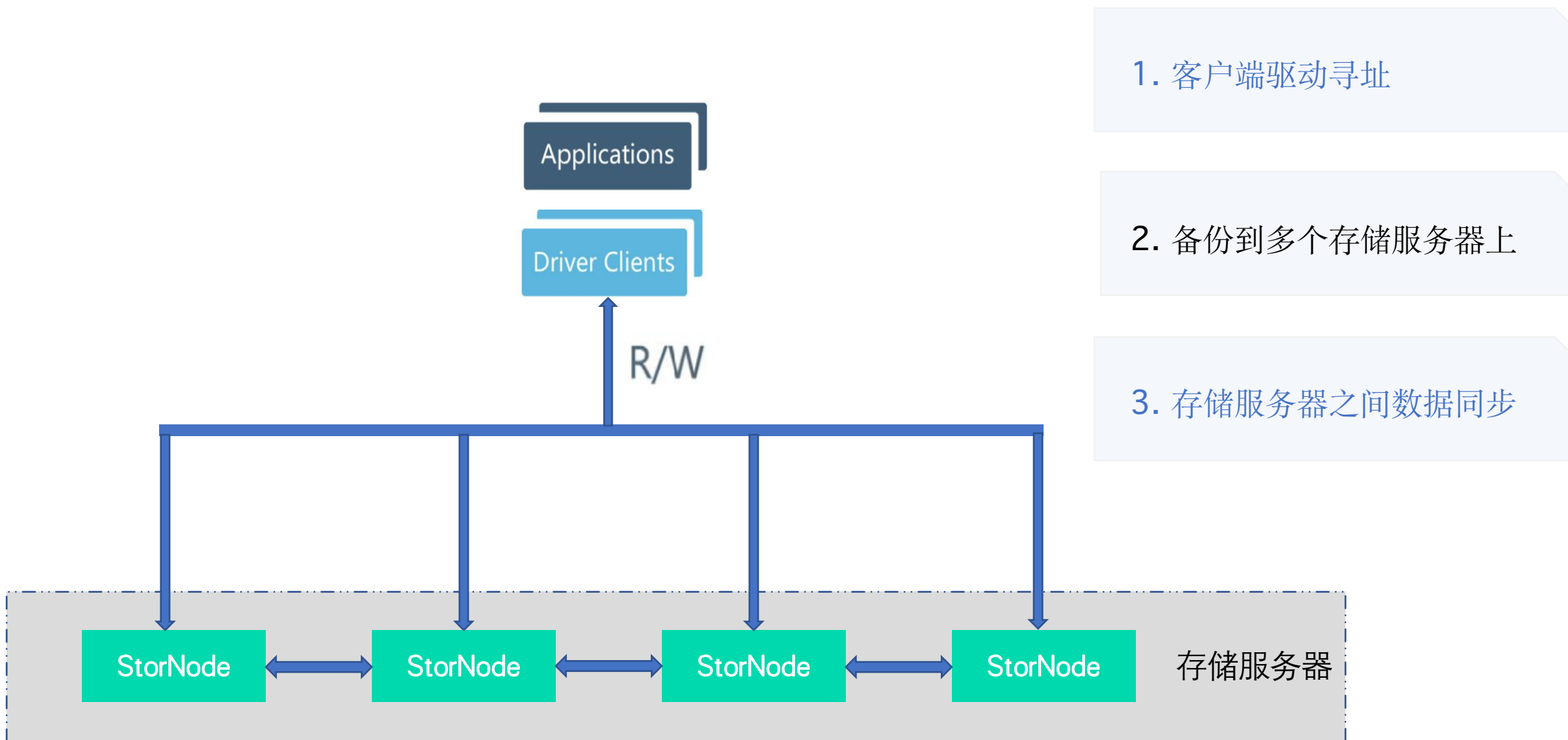
SDS系统数据流

03

经典SDS和鹏云系统架构比较

04

总结





目录

COMPANY

01

SDS系统介绍

02

SDS系统数据流

03

经典SDS和鹏云系统架构比较

04

总结

- 地址存放在元数据中

闭源产品使用

缺点：维护元数据很复杂

优点：系统的容错性、稳定性和可扩展性高

- 通过算法直接计算数据地址，无需元数据

DHT（Distributed Hash Table），GlusterFS等在使用

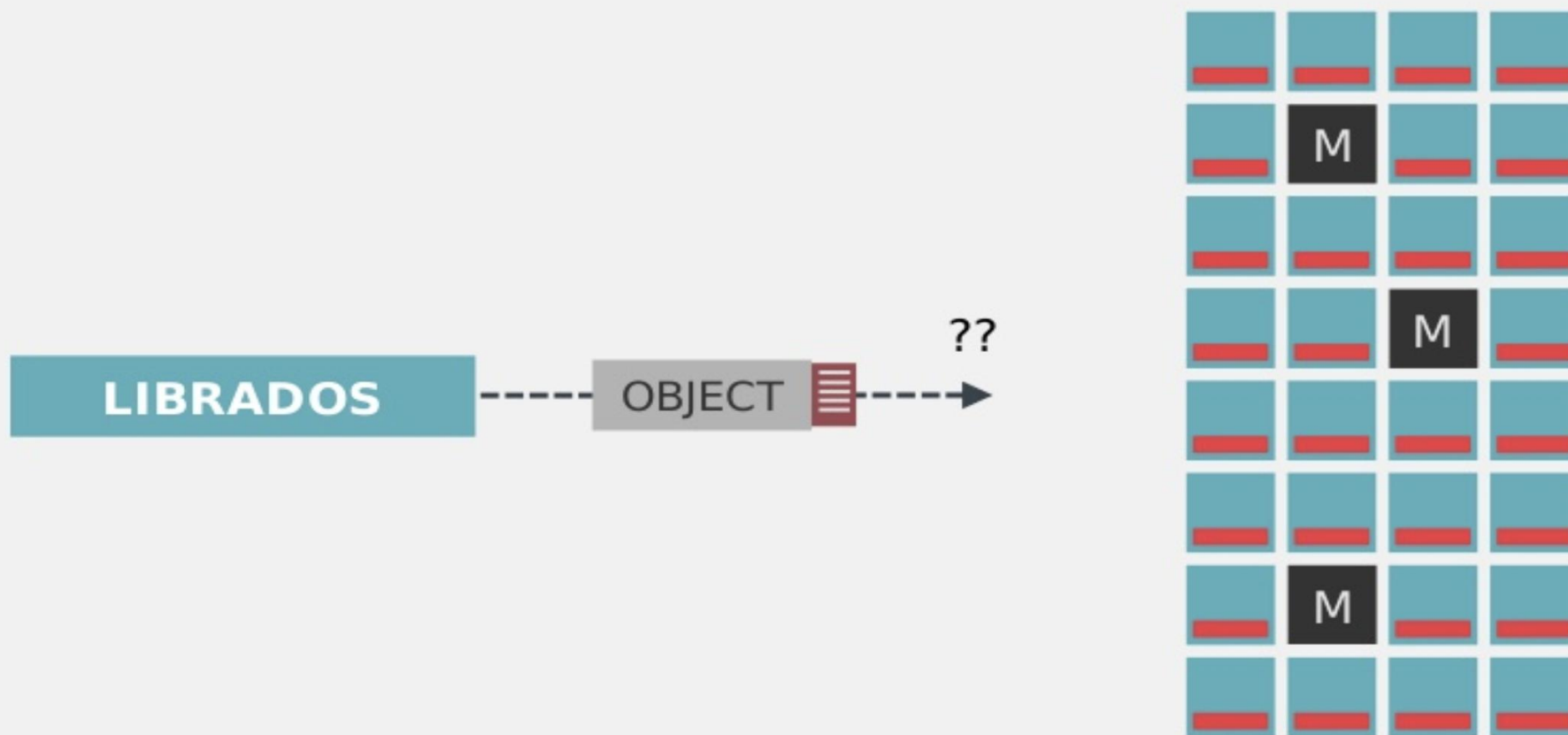
CRUSH（Controlled Replication Under Scalable Hashing）Ceph
在使用

缺点：系统的容错性、稳定性和可扩展性低

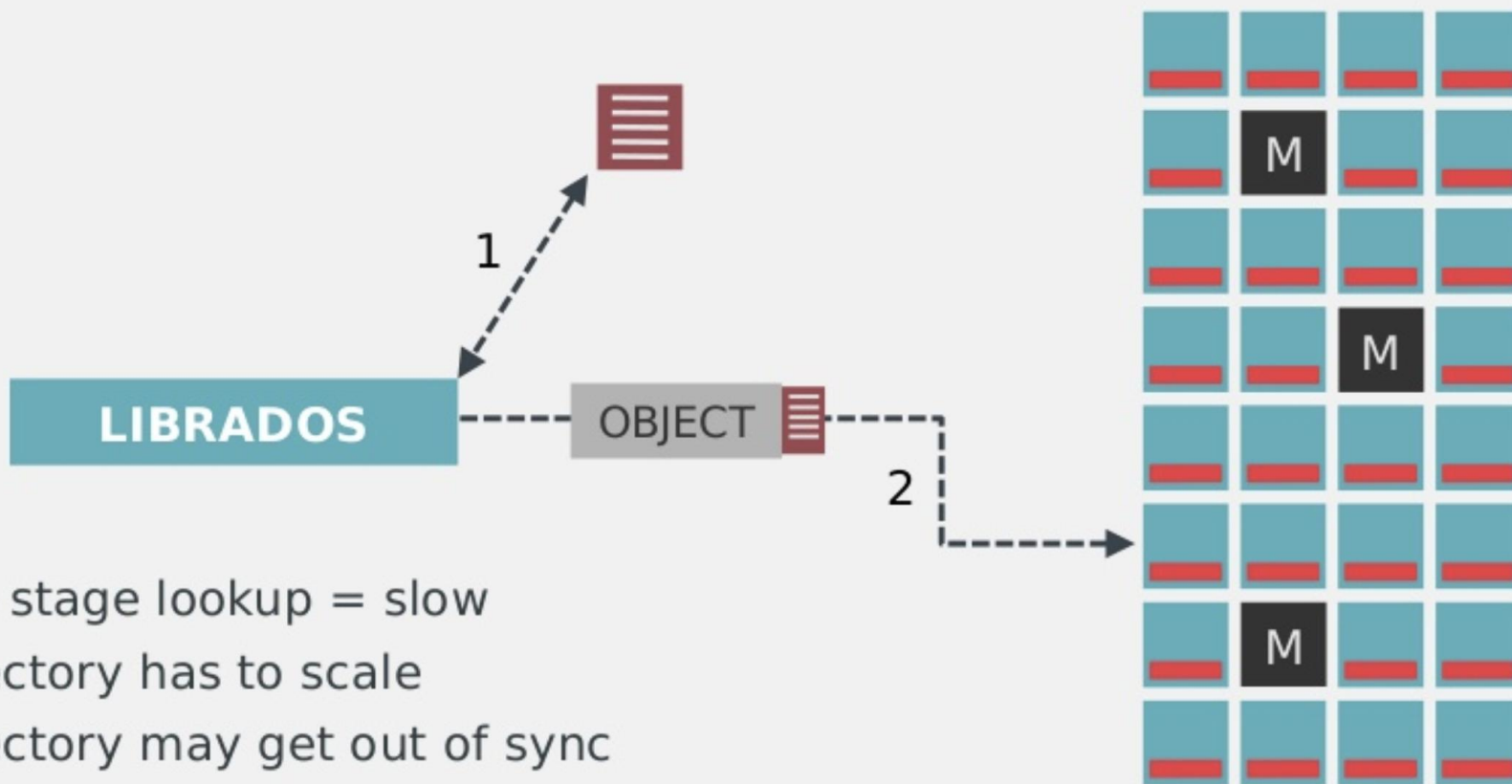
优点：实现简单



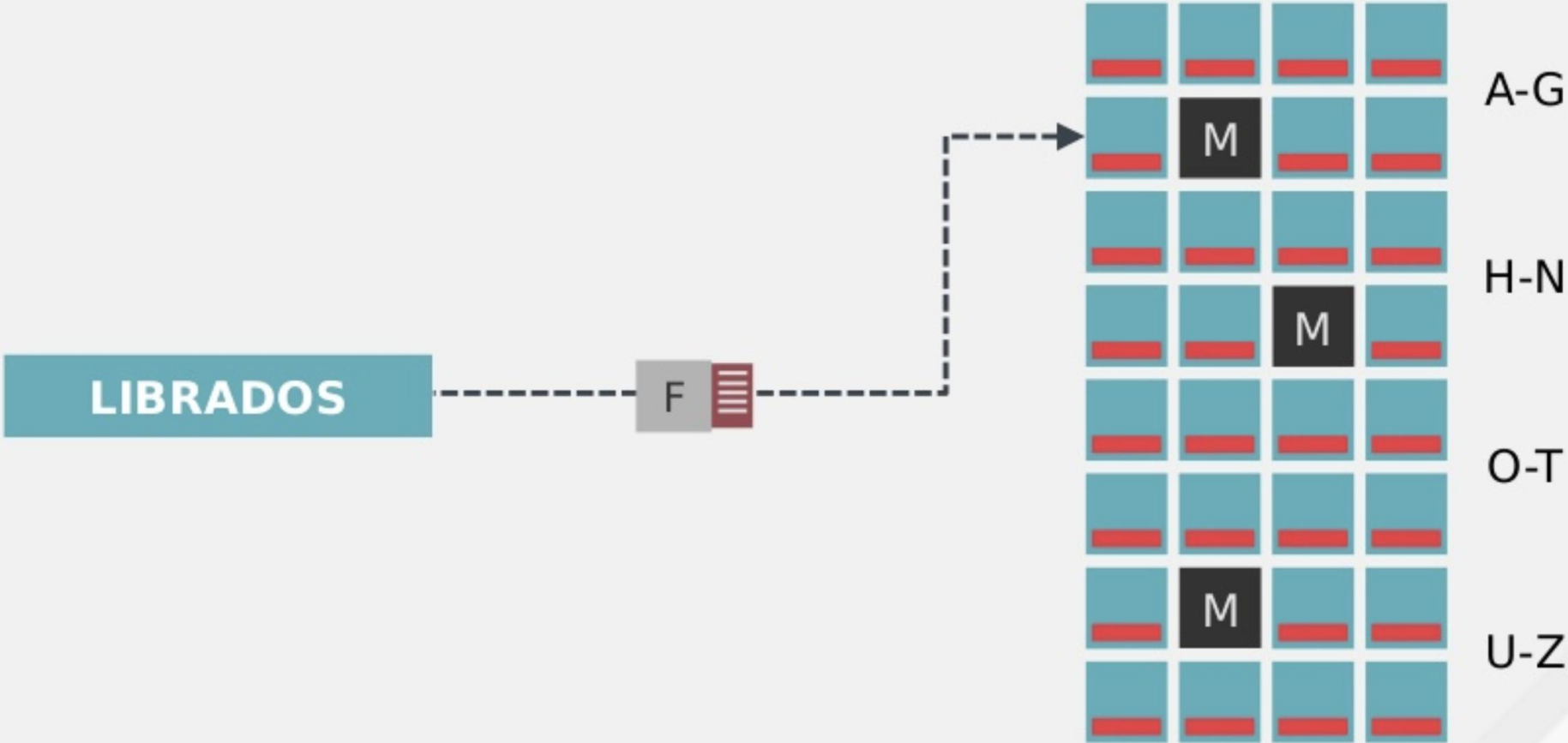
WHERE DO OBJECTS LIVE?



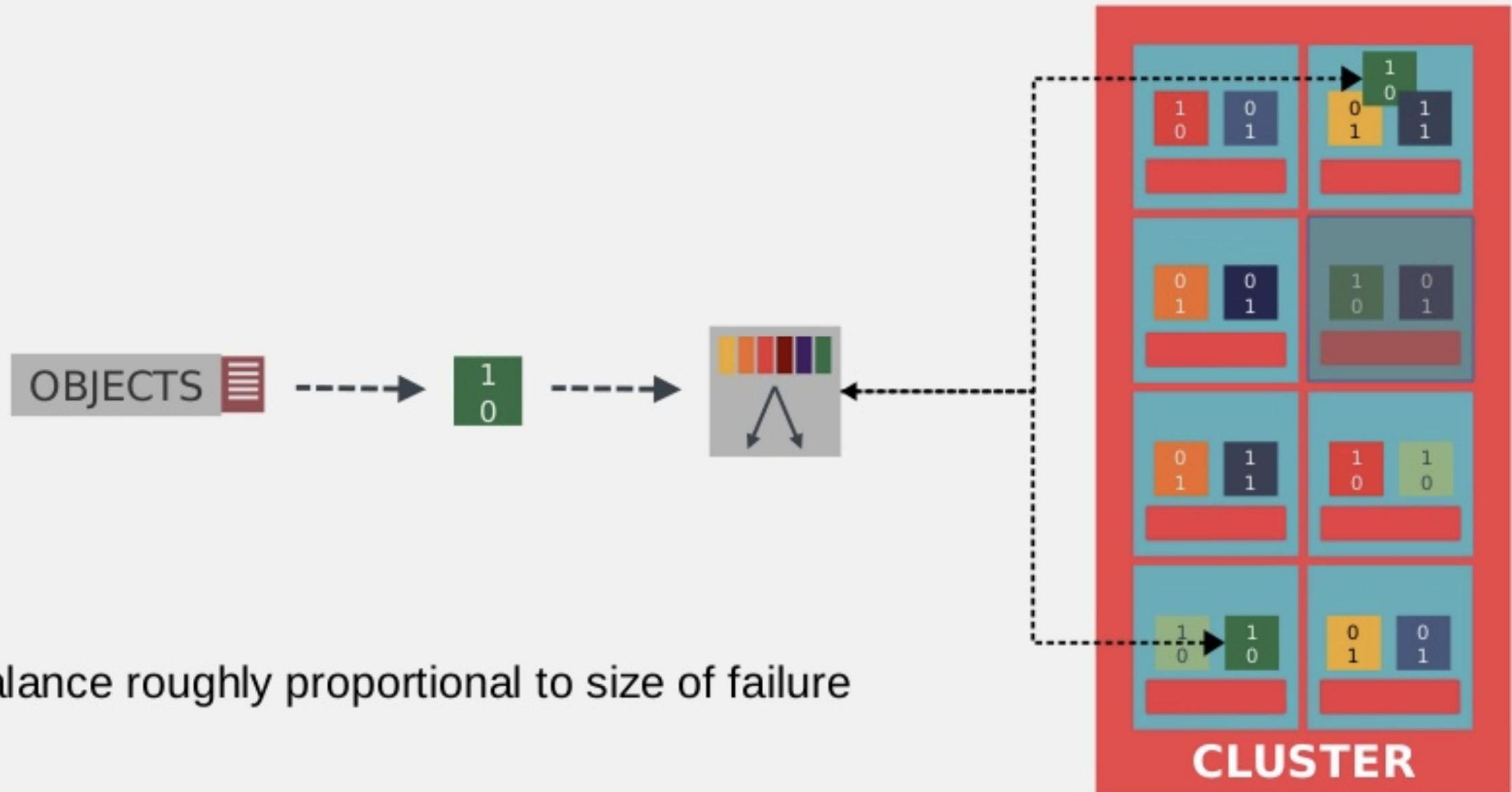
MAINTAIN A DIRECTORY?



CALCULATED PLACEMENT

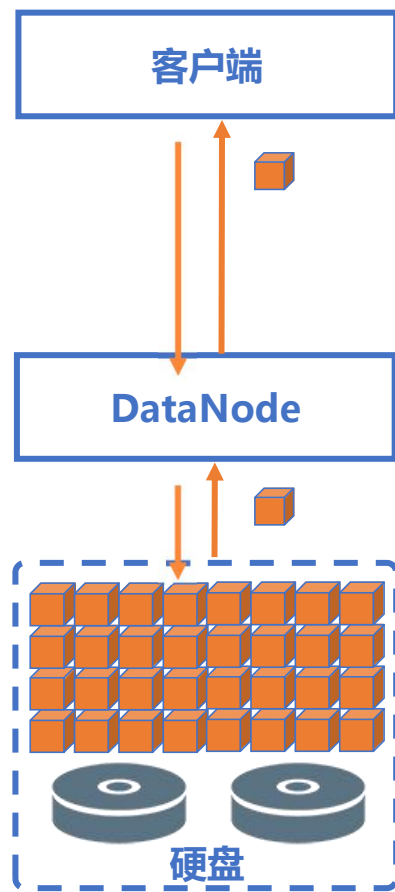


DEVICE FAILURES



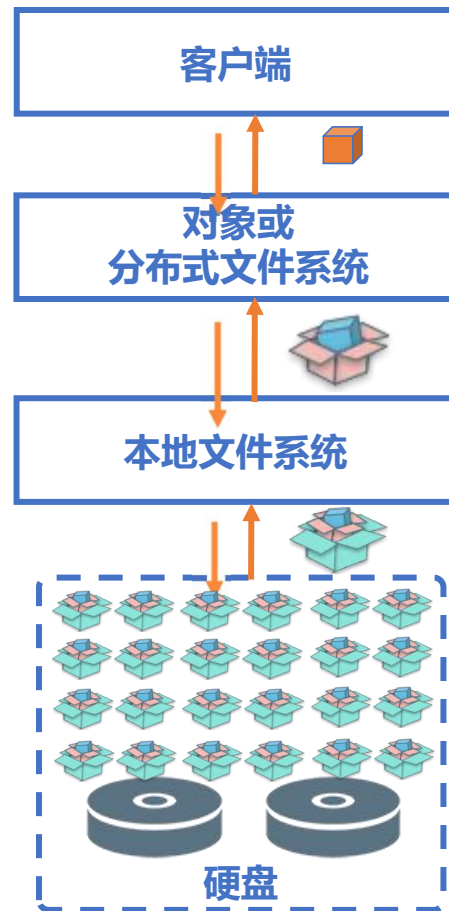
原生块存储

直接管理磁盘裸设备
不依赖对象或文件系统
无需复杂数据格式转换
低访问延迟
更低系统开销

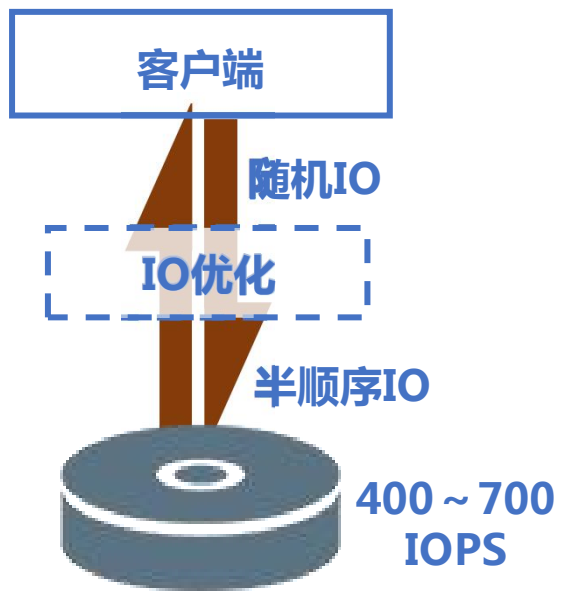


底层基于对象存储的产品

不能操作裸设备
底层依赖对象或文件系统
复杂数据格式转换
高访问延迟
高系统开销

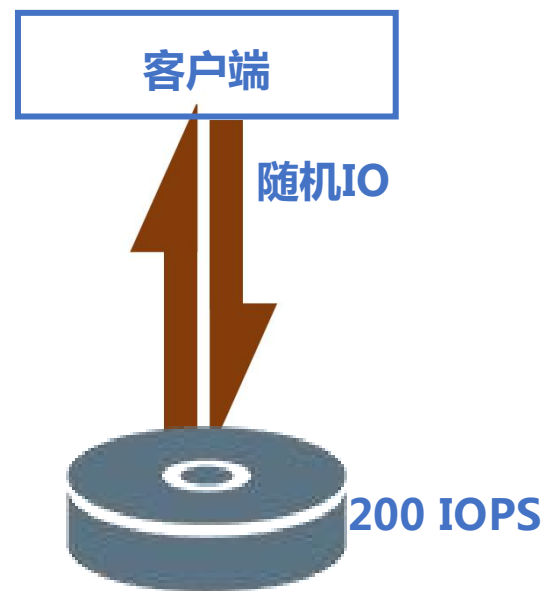


ZettaStor DBS



变随机IO为半顺序IO
IO性能提升2 ~ 3.5倍
更高性能，降低成本

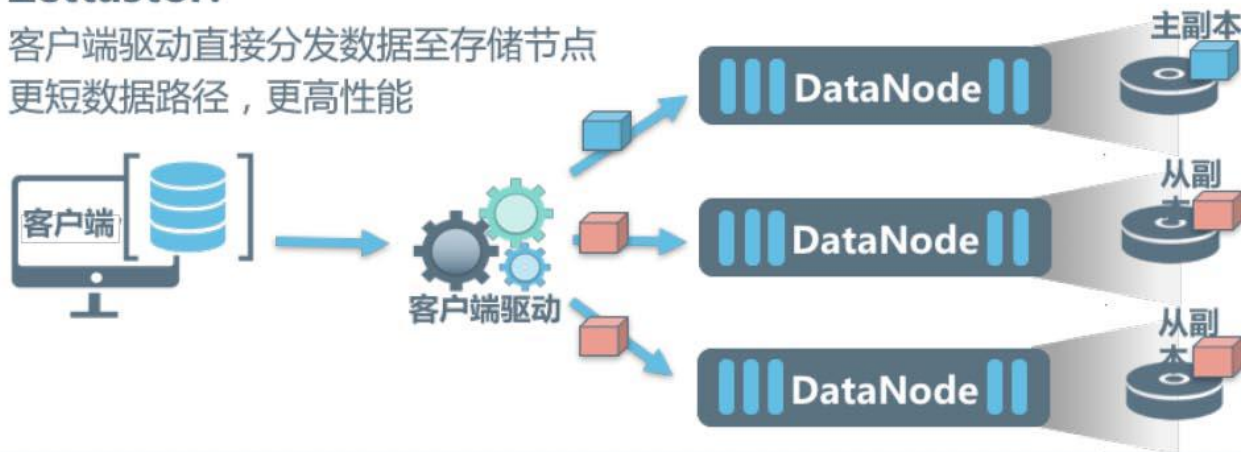
非优化的同类产品



磁盘随机IO性能不佳
需配置更多磁盘来提升性能
更低性能、更高成本

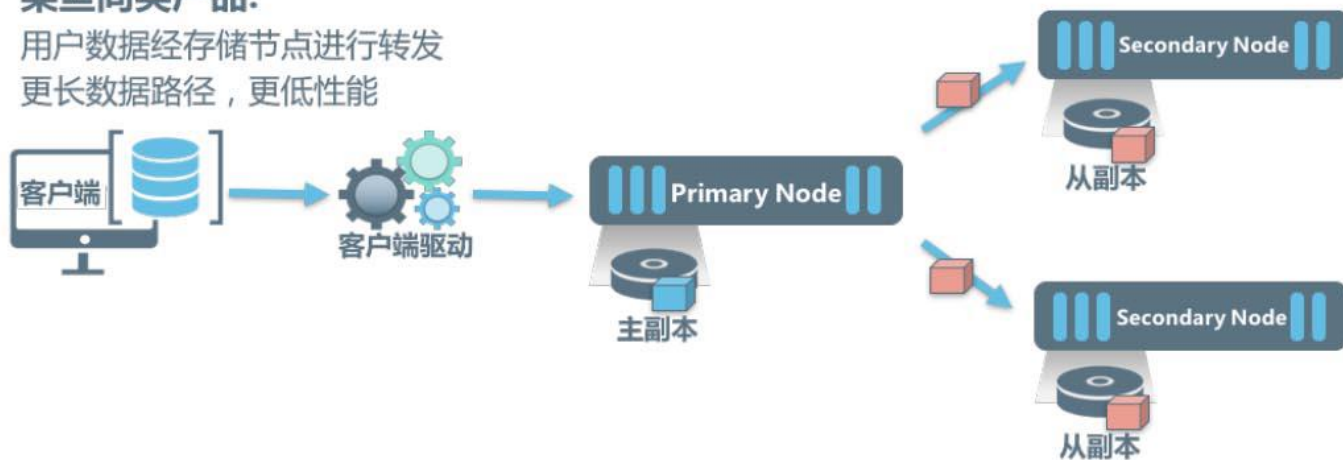
Zettastor:

客户端驱动直接分发数据至存储节点
更短数据路径，更高性能



某些同类产品:

用户数据经存储节点进行转发
更长数据路径，更低性能



- 数据写到多个备份节点时，用全量写

客户端驱动确保数据写到了所有的备份节点里，才返回给应用端

优点：实现简单

缺点：如果某台服务器某段时间变慢，会影响整个集群的写操作性能

- 数据写到多个备份节点时，用Quorum写

客户端驱动确保数据写到了主节点和 $(M - 1)$ 从节点便可以返回

优点：写操作性能受个别慢的服务器影响小

缺点：实现复杂



- Quorum写的实现

用日志记录下来所有的写操作

主节点拥有所有的日志

因为Quorum写，导致从节点缺失部分数据，通过增量地同步主节点上的日志，确保从节点有全部的数据



极客时间全部课程任学 喊老板来买单！

- ✔ 精选 13+ 热门职位的学习路径，包括架构、运维、前端工程师等
- ✔ 根据不同技术岗位能力模型匹配合适的课程
- ✔ 一键设置购买条件，成员按需选课，自主制定学习计划
- ✔ 享充值满赠优惠，帮老板省钱，团队免费学习



立即申请



TGO 鲲鹏会

汇聚全球科技领导者的高端社群

 全球12大城市

 850+高端科技领导者

使命
Mission

为社会输送更多优秀的
科技领导者

愿景
Vision

构建全球领先的有技术背景
优秀人才的学习成长平台



扫描二维码，了解更多内容



Thanks

南京鹏云网络科技有限公司