

数据驱动 存储创新

# 2017企业存储技术峰会·北京站

2017 ENTERPRISE STORAGE TECHNOLOGY SUMMIT BEIJING

2017.01.18 · 北京东方丽晶酒店

主办方

日知录技术社区

# 数据方舟-构建云平台的TimeMachine服务

- 彭晶鑫 (UCloud块存储研发经理)

## Agenda

- UCloud 数据方舟介绍
- UCloud 数据方舟基础架构
- 数据方舟的应用实践
- 数据方舟的未来规划

## Agenda

- UCloud 数据方舟介绍
- UCloud 数据方舟基础架构
- 数据方舟的应用实践
- 数据方舟的未来规划

## UCloud数据方舟介绍

- 提供块设备连续数据保护 (CDP) 的产品。
- 为UCloud 云硬盘/ UCloud 云主机提供 CDP 能力。

## UCloud数据方舟特性

- 可以将块设备回滚至12小时内任一秒，24小时内任一个小时，3天内任一个零点。
- 可以将块设备从12小时内任一秒，24小时内任一个小时，3天内任一个零点复制到新的块设备。
- 可以制作快照。
- 不关心上层是什么应用。

## 块设备传统备份及恢复流程

- 手动或者定期对块设备制作快照，例如每天 03:00 制作一个快照。
- 故障或者操作不当写入脏数据（14:56分）回滚至 03:00。
- 从 03:00 到 14:56 分所有的有用数据全都丢失。

## UCloud数据方舟恢复流程

- 故障或者操作不当写入脏数据 (14:56:56) 回滚至 14:56:46
- 故障或者操作不当写入脏数据 (14:56:56) 将 14:56:46 秒的状态 Clone 至新的块设备/主机

## Agenda

- UCloud 数据方舟介绍
- **UCloud 数据方舟基础架构**
- 数据方舟的应用实践
- 数据方舟的未来规划

## UCloud 数据方舟基础架构

NetDisk IO接入

LocalDisk IO接入

实时IO 存储层

IO Minor Merge 存储层

IO Major Merge 存储层

## UCloud 数据方舟基础架构

**IO接入层:** 将网络盘/本地盘的写IO 实时转发至服务后端实时IO存储层。

**实时IO存储层:** 负责云平台下块设备的实时IO写的存储。

**IO Minor Merge 存储层:** 从实时IO存储层拉取数据，并进行合并计算生成分钟或者小时级别粒度的增量数据。

**IO Major Merge 存储层:** 从IO Minor Merge存储层拉取数据，并进行合并计算生成天级别粒度的base数据或者增量数据。

## UCloud 数据方舟的挑战

- 云平台下大量的随机IO写
- IO 数据任一点不能丢失
- 流量突增 过载问题
- 后端存储成本

## Agenda

- UCloud 数据方舟介绍
- UCloud 数据方舟基础架构
- 数据方舟的应用实践
- 数据方舟的未来规划

## 硬件选型

- 为了应对不同的 IO 需求，数据方舟采取了分层的架构设计。
- 实时 IO 存储层选型 SSD。
- IO Minor Merge 存储层 和 IO Major Merge 存储层 选型 SATA，并在程序设计上充分使用顺序读写发挥 SATA 的顺序 IO 能力

## 容量问题

- 单个接入实例的 IOPS 容量 。
- 单个接入实例的 带宽容量
- 单个接入实例的 磁盘容量

单个实例的负载超过其自身容量时 – 在线动态迁移至其它实例

## 云平台下大量随机写

- 实时 IO 存储层对于大量块设备来说是随机写操作，采用随机写能力更强的 SSD。
- 实时 IO 存储层对于同一个块设备进行 IO Journal 的 append 操作
- IO 接入层 可将 1s 内的 IO 合并为一个请求转发给后端。

## IO数据不能丢失

- 任何一份数据（实时 IO 数据， Minor 数据， Major 数据）采用主副本双份存储。
- 网络波动 / 后端服务重启 IO 流队列
- 实时 IO 接入层容量问题（IOPS 撑不住， 实时 Journal 容量不足）

## 后端存储成本

- 采用 Base 数据 + 增量数据的存储方式。
- 实际存储文件进行压缩

## 后端存储文件格式

0-1G 寻址范围

Block header	Block Data
Block header	Block Data

...

N-1 - N G 寻址范围

Block header	Block Data
Block header	Block Data

...

...

Block header	Block Data
--------------	------------

Block header	Block Data
--------------	------------

## 按容量划分set, 按set部署和运营



## Agenda

- UCloud 数据方舟介绍
- UCloud 数据方舟基础架构
- 数据方舟的应用实践
- 数据方舟的未来规划

## 未来半年规划

- RTO 加速（1T盘半小时可以恢复完成，测试中，3月份全面上线）
- 更智能的 IO 流接入层调度（更好的保障 IO 流不会因容量问题中断）

## Q&A

