



中立安全·赋能产业

云计算基础-RSSD块存储技术详解和应用场景

UCloud优刻得 技术总监 彭晶鑫

01

块存储基础介绍

02

RSSD云盘技
术详解

03

块存储备份
方案

04

RSSD云盘应
用案例介绍

云计算-块存储介绍

块存储是为云服务器ECS提供的随机的可随机读写的块设备存储，具备持久化存储，高可靠，低时延。块存储

采用多副本冗余，防止网络，硬件故障而导致的数据不可用。并且使用方式和硬盘一致

利用分布式块存储，可以实现计算和存储分离



在线迁移

由于计算和存储分离，将虚机从一台宿主机在线迁移至另一台宿主机时，只需迁移cpu和内存，可在几十秒内完成



宕机恢复

如果宿主机宕机，由于计算和存储分离，可以快速在另外一台宿主机上拉起。宕机恢复速度可提速10倍



高可用，高可靠

通过分布式，可以发现节点故障并剔除，
从而避免节点硬件故障，网络故障等带来
的存储数据问题和可用性问题



存储池化

存储组成较大池化容量，按需分配存储空
间，避免本地存储碎片过多问题，可提供
大容量存储，例如40TB, 100TB

UCloud-云盘

UCloud目前提供2种类型的云盘，即普通云盘和

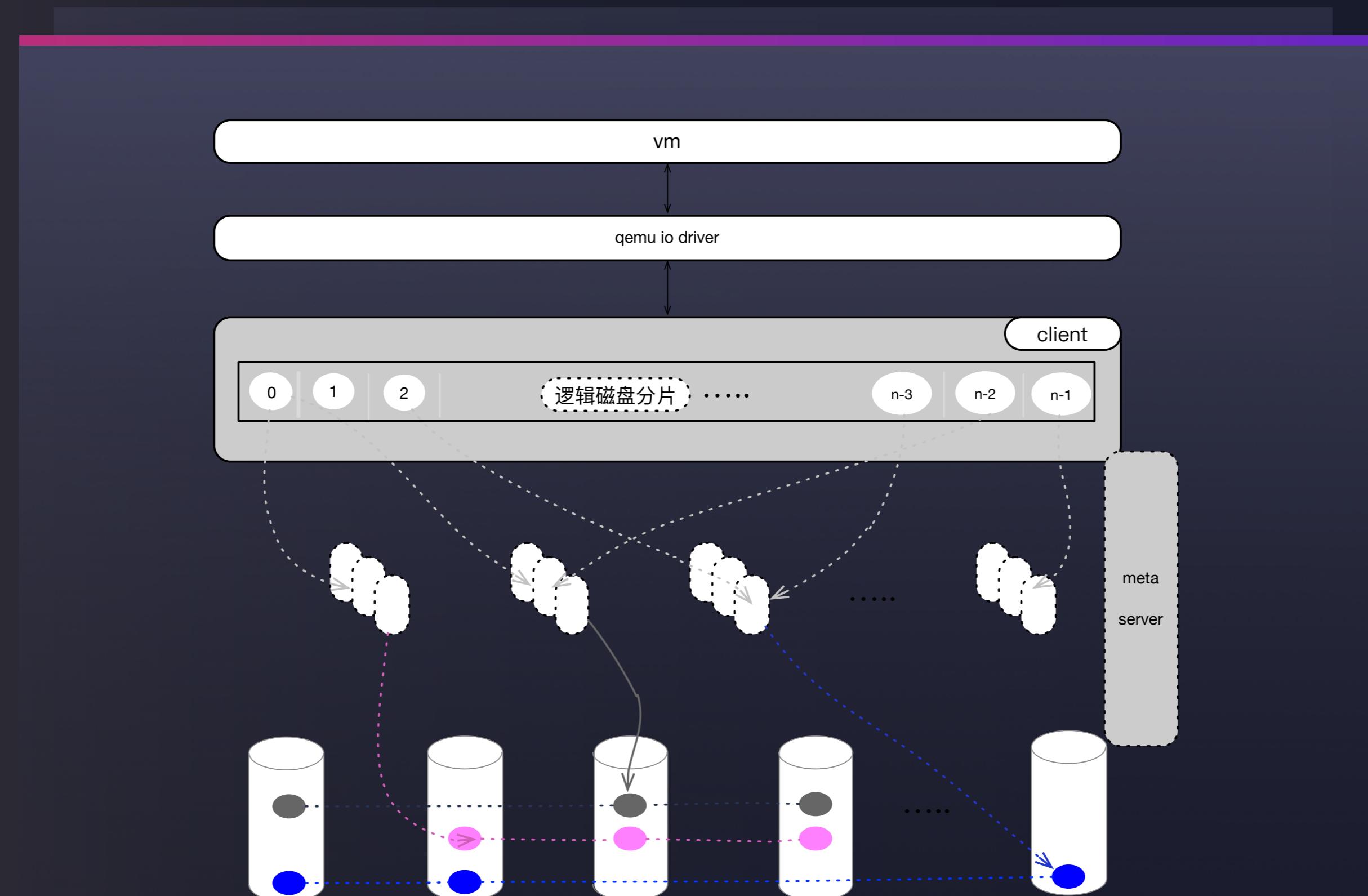
SSD云盘

普通云盘后端利用机械硬盘介质

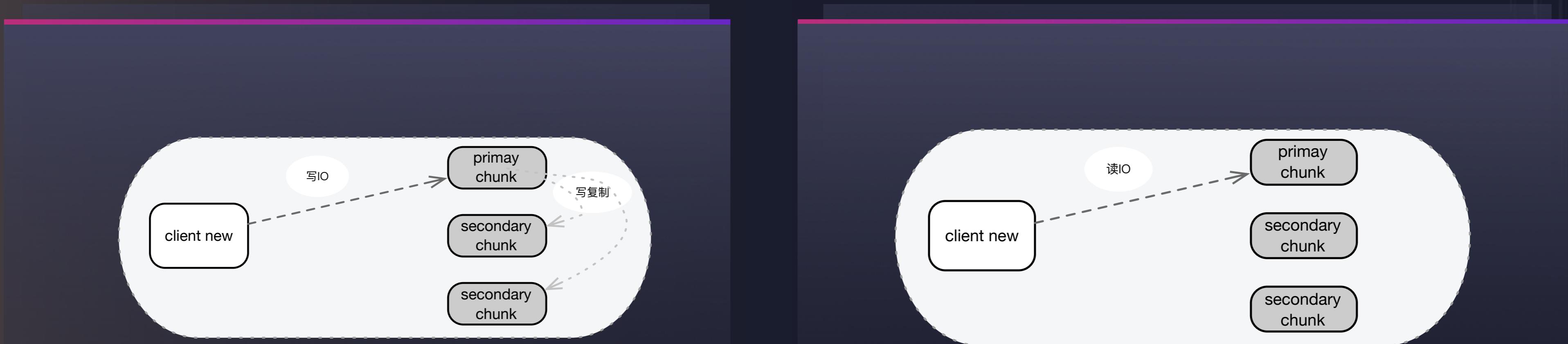
SSD云盘后端利用NVME固态硬盘介质

	普通云盘	SSD云盘
IOPS	1000	2.4W
带宽	100MB/s	260MB/s
时延	5-10ms	0.5-3ms

SSD云盘架构



SSD云盘读写



行业发展

存储

hdd -> sata ssd ->
nvme ssd

性能百倍提升

网络

1G -> 10G -> 25G ->
100G

网络接口也迎来了跨越式
发展

CPU

CPU主频继续停留在2-
3GHZ

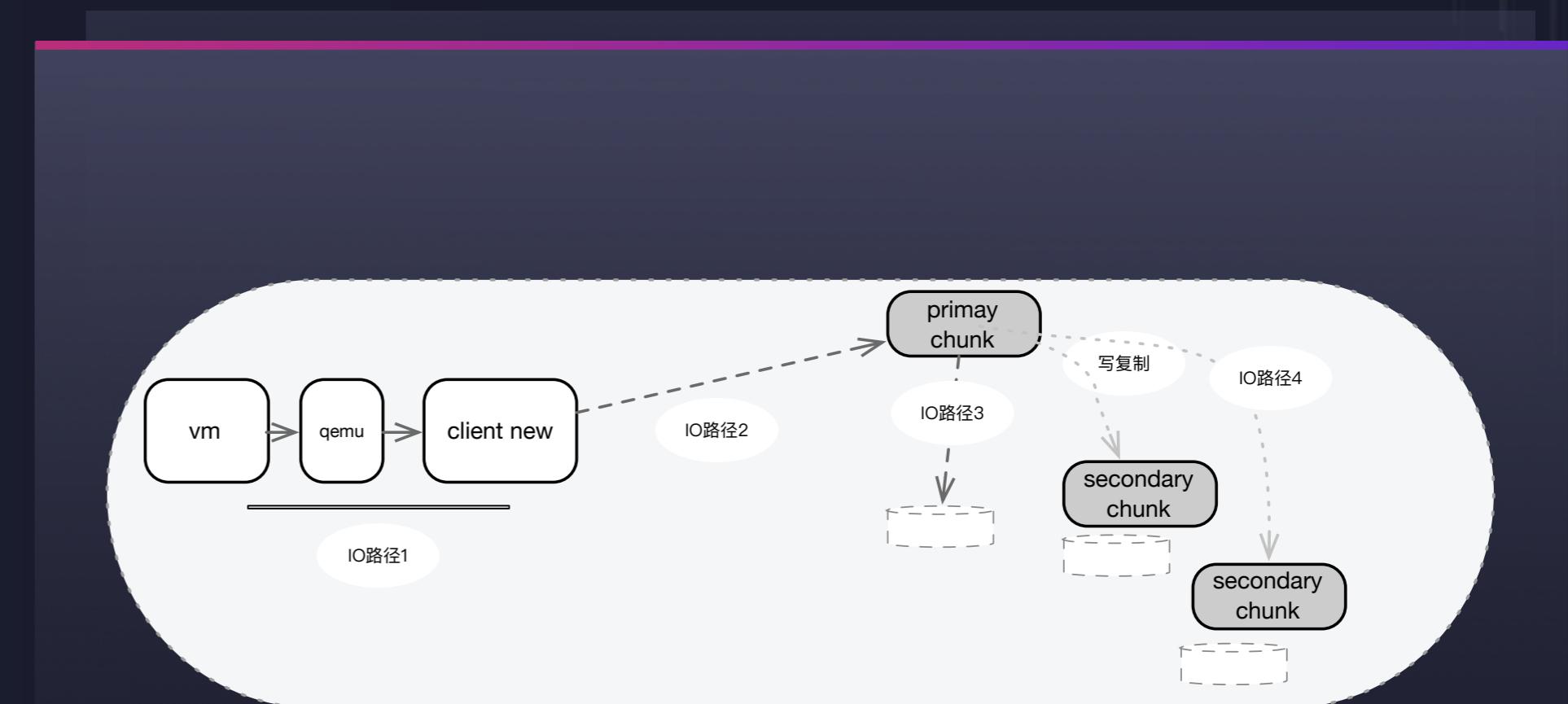
未出现跨越发展

超高性能一块存储

- 一台物理机可以6–8块NVME盘，意味着一台物理机NVME固态盘可以提供4K 300–500万的IOPS
- 传统应用服务器软件模式下，基于TCP的epoll loop, 网卡的收发包, IO的读写要经过用户态，内核态多层拷贝和切换，并且需要靠kernel的中断来唤醒
- 增加线程软件很难压榨出硬件的全部能力，例如在IOPS, 和时延上

RSSD 云盘

- 能否将网络的时延控制的更低?
- 能否压榨出nvme固态盘的时延和IOPS能力?
- 能否整体IO上更好的软件消耗?



 client

利用vhost user技术

实现虚机到存储client的数据零拷贝

 网络

ROCE 代替 TCP

4K网络包收发10-20us

 存储读写

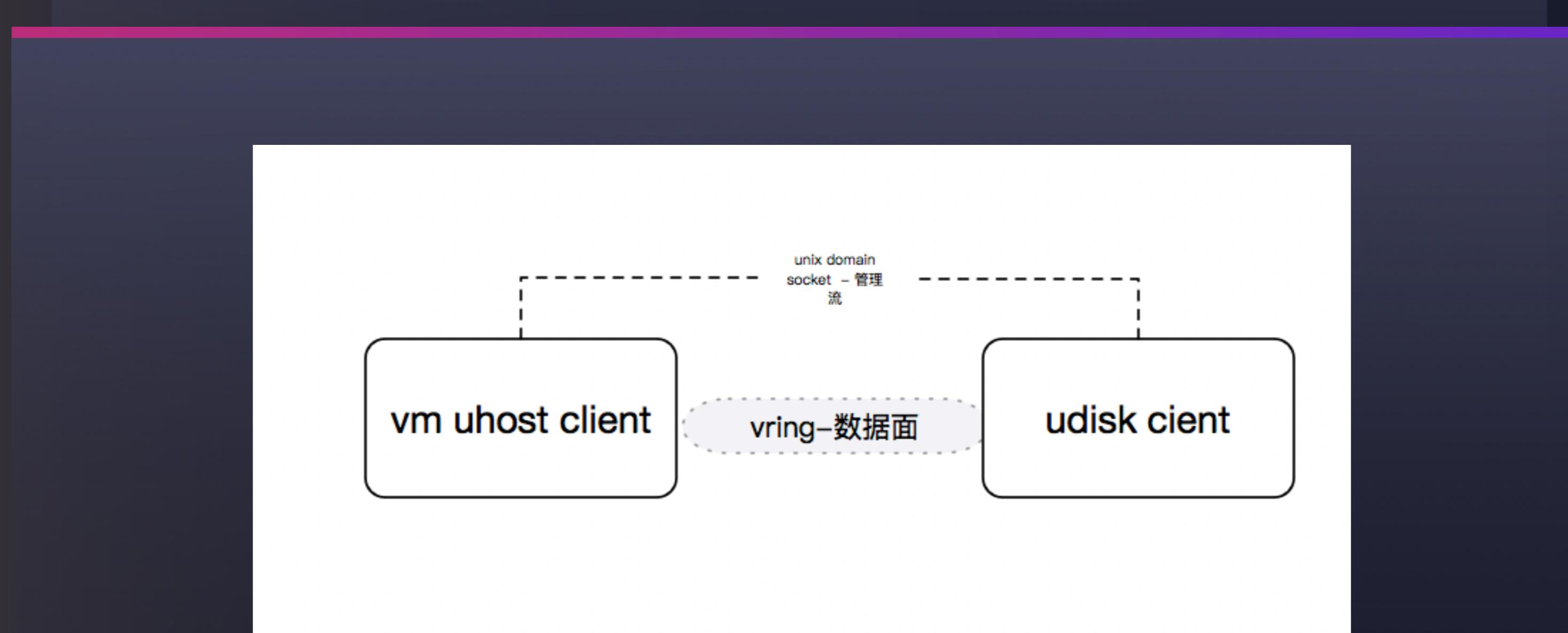
SPDK代替libaio

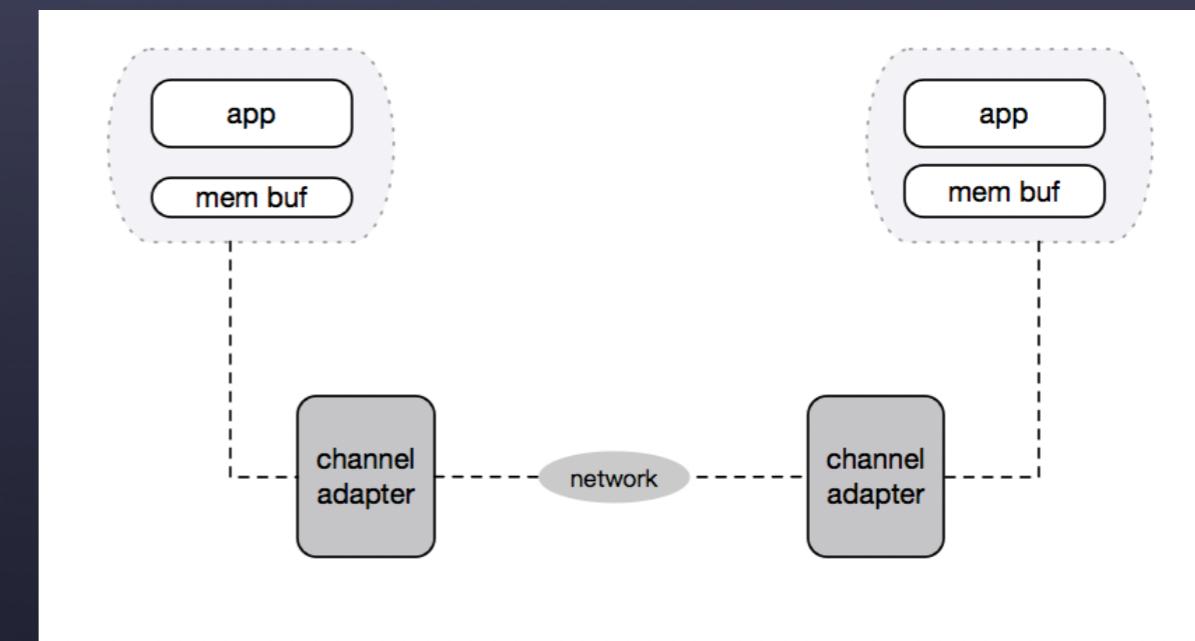
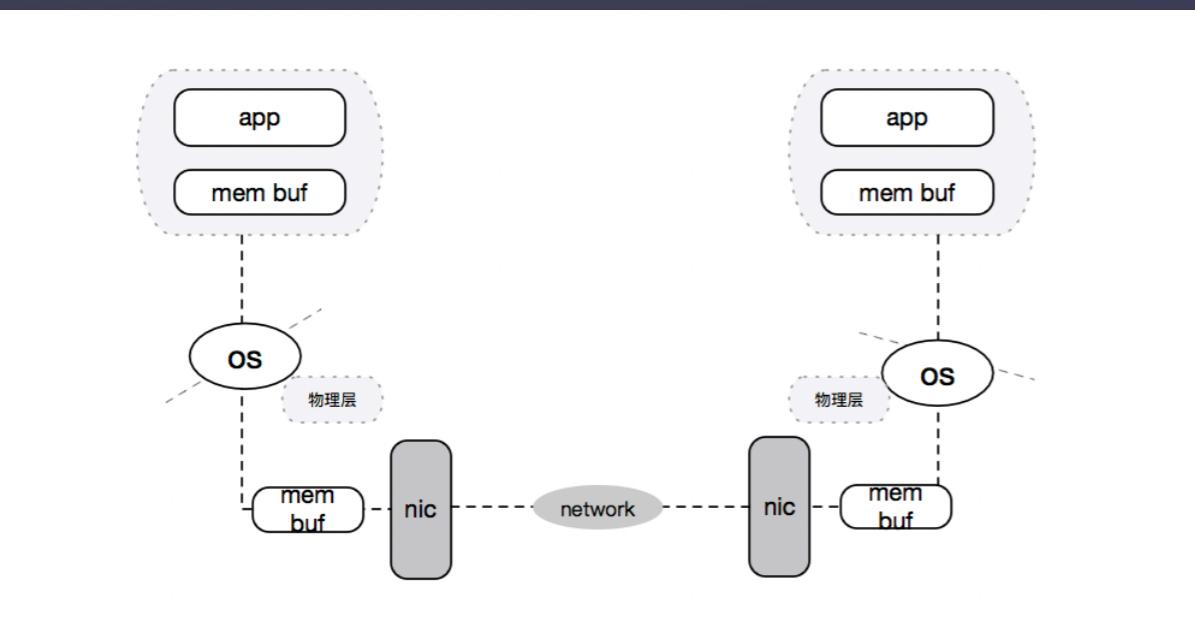
用户态IO读写，高并发下，依然保持较低的时延

 用户态和轮询

IO路径整体保持在用户态，并使用线程轮询

减少用户态内核态切换，减少中断带来的损耗



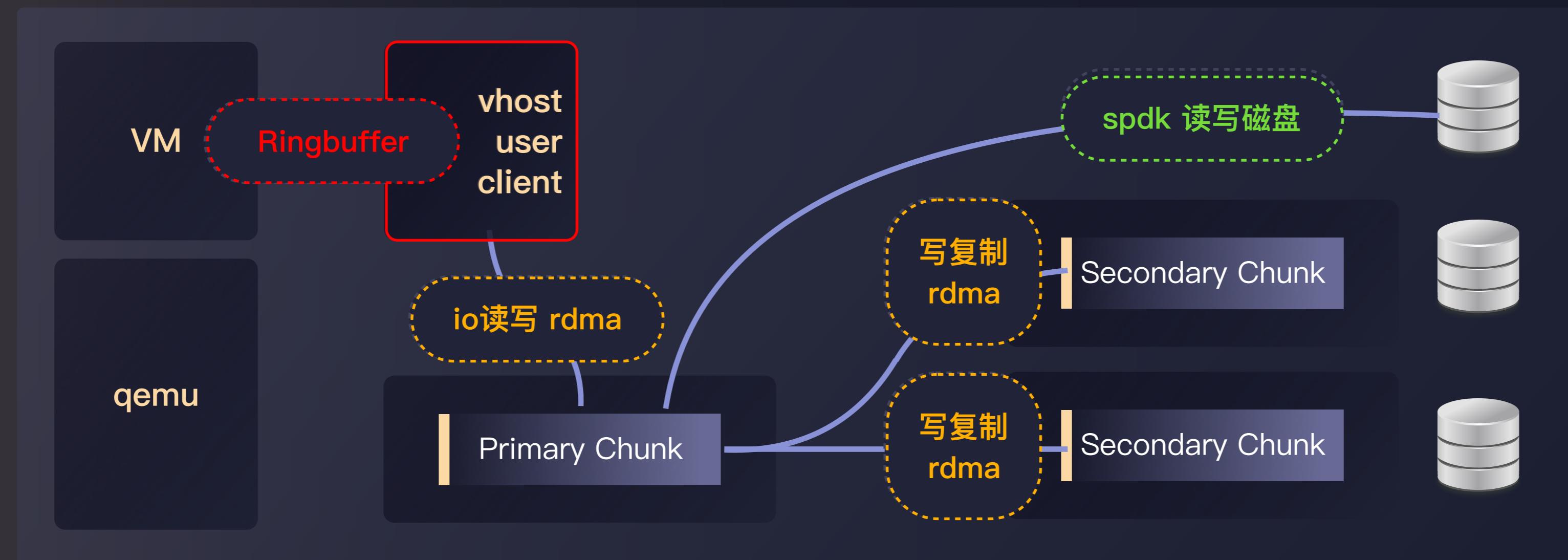


 SPDK

异步的, 无锁的, 轮询的, 用户态的 nvme driver

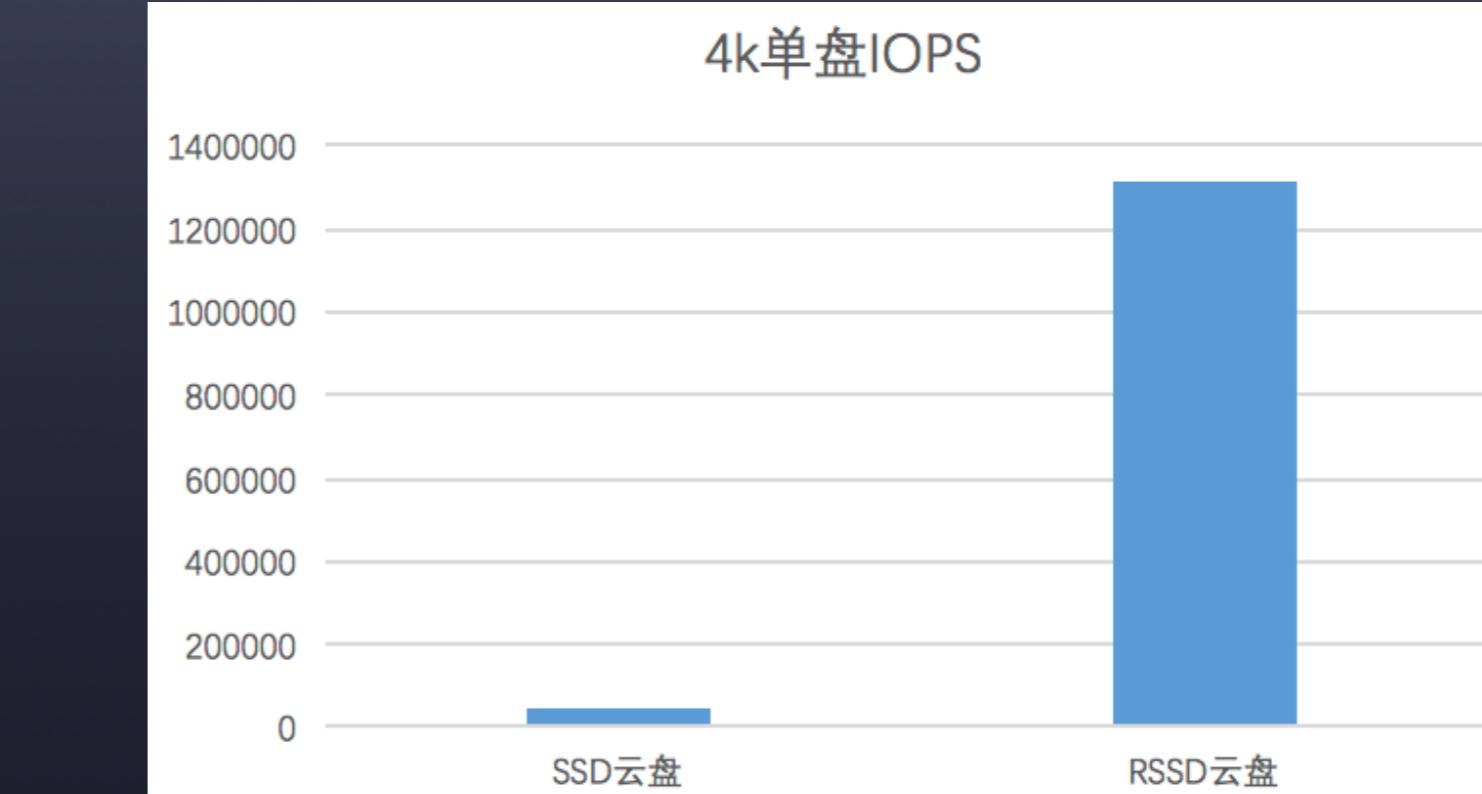
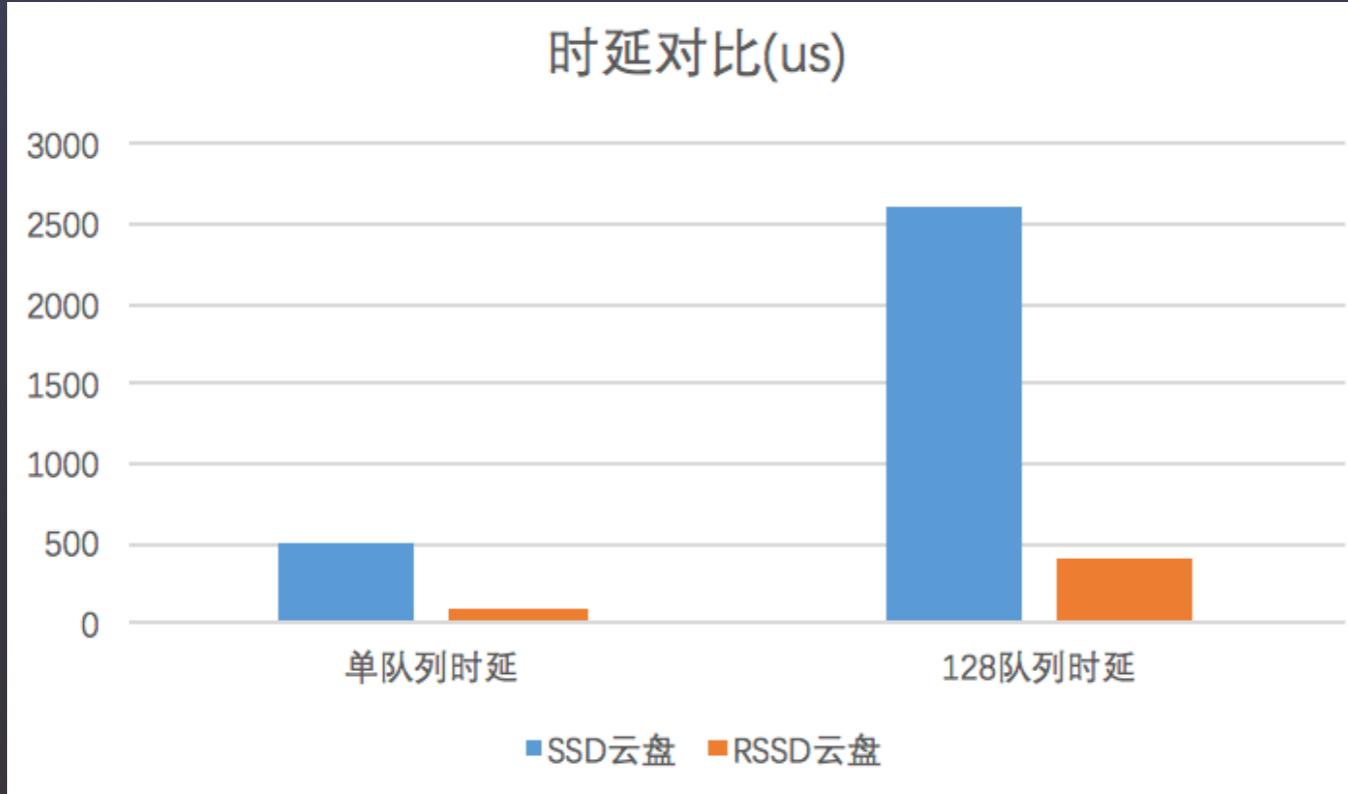
用户态应用程序直接用户态driver访问nvme设备

0拷贝, 可设置队列高并发访问nvme设备



全新软硬件栈的协同设计

RSSD云盘性能



2.4 万IOPS
0.5 ms延迟



120 万IOPS
0.1 ms延迟

云盘备份-数据方舟

- 快照备份点 (RPO)
- 恢复速度 (RTO)
- 多副本存储
- 异构解耦
- 可回滚任意一秒?
- 不影响原云主机性能?
- 操作简单, 无需理解复杂的虚拟化和快照实现?

数据方舟要点

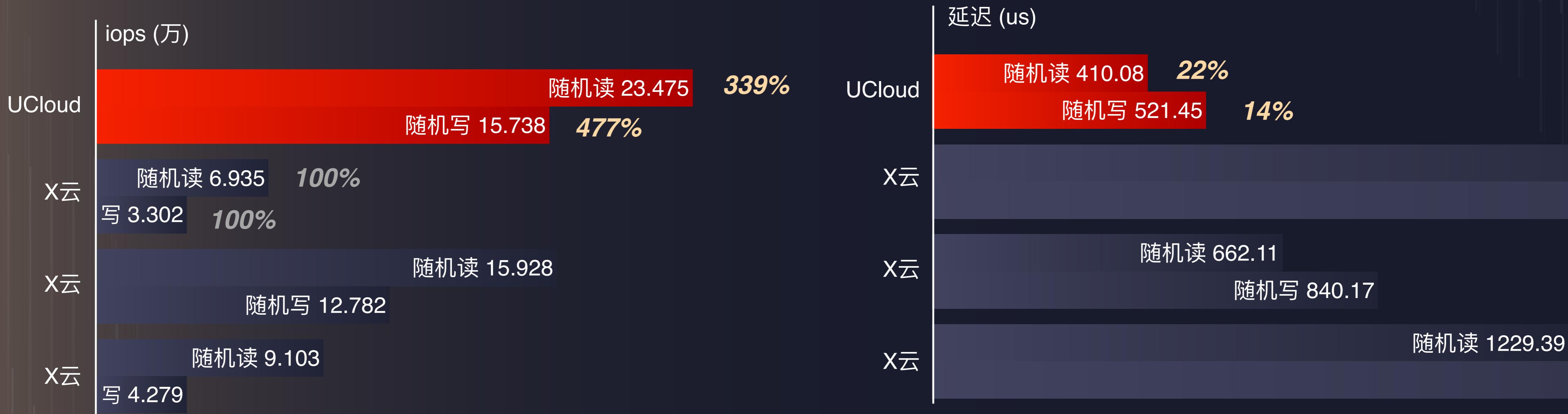
- 记录实时IO流，以恢复任意一秒（即RPO）
- 异构解耦的实现方式，不影响源主机，且运维过程依赖小持续海量随机IO带来的挑战
- 备份恢复恢复至全新的磁盘，而不是覆盖原盘，新意味着全盘数据重新写入，通过分布式计算和存储技术保证恢复时间（即RTO）



RSSD 云盘案例评测

存储性能 云主机4C8G

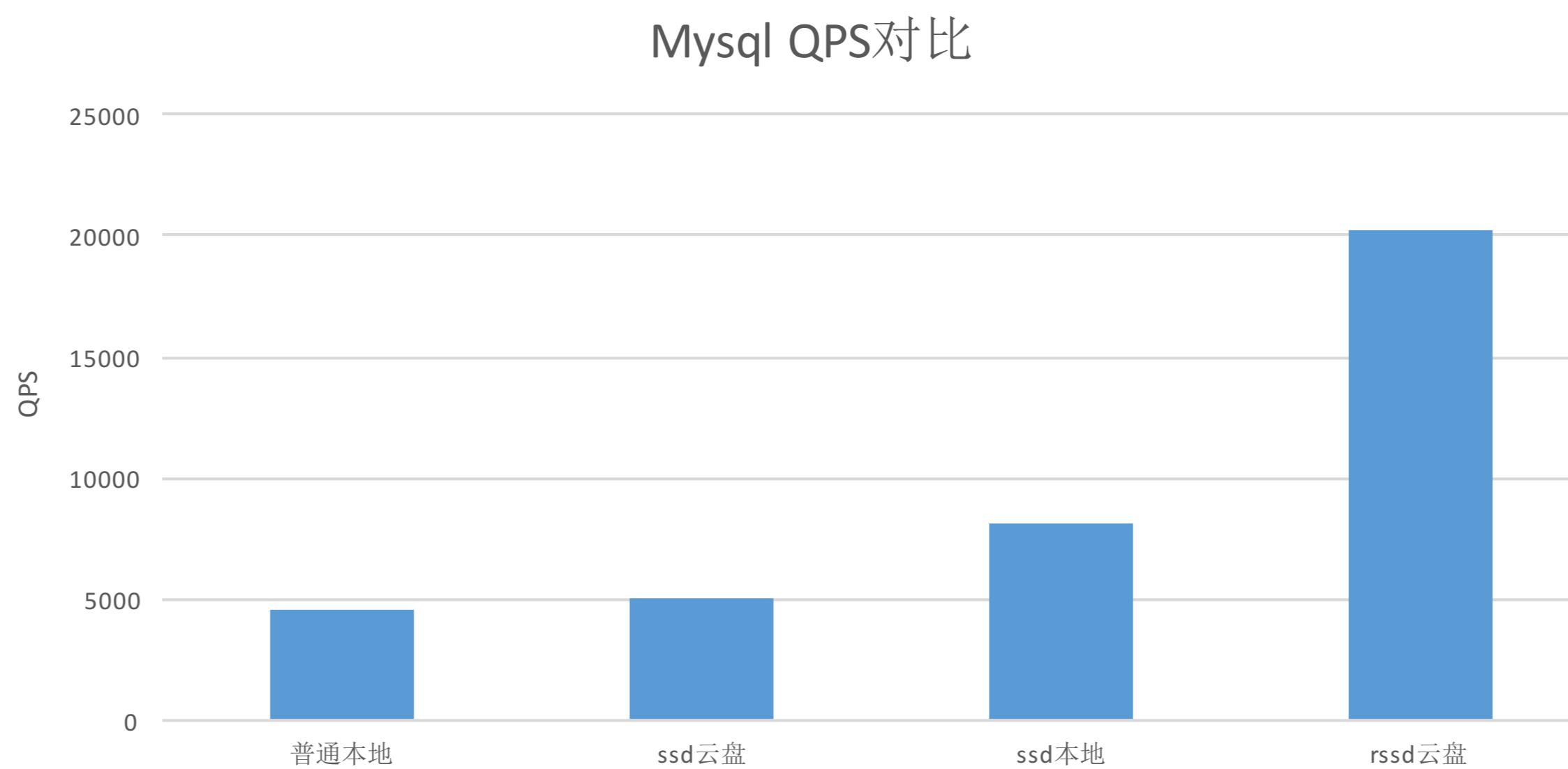
IODepth 4*32 / BS 4k



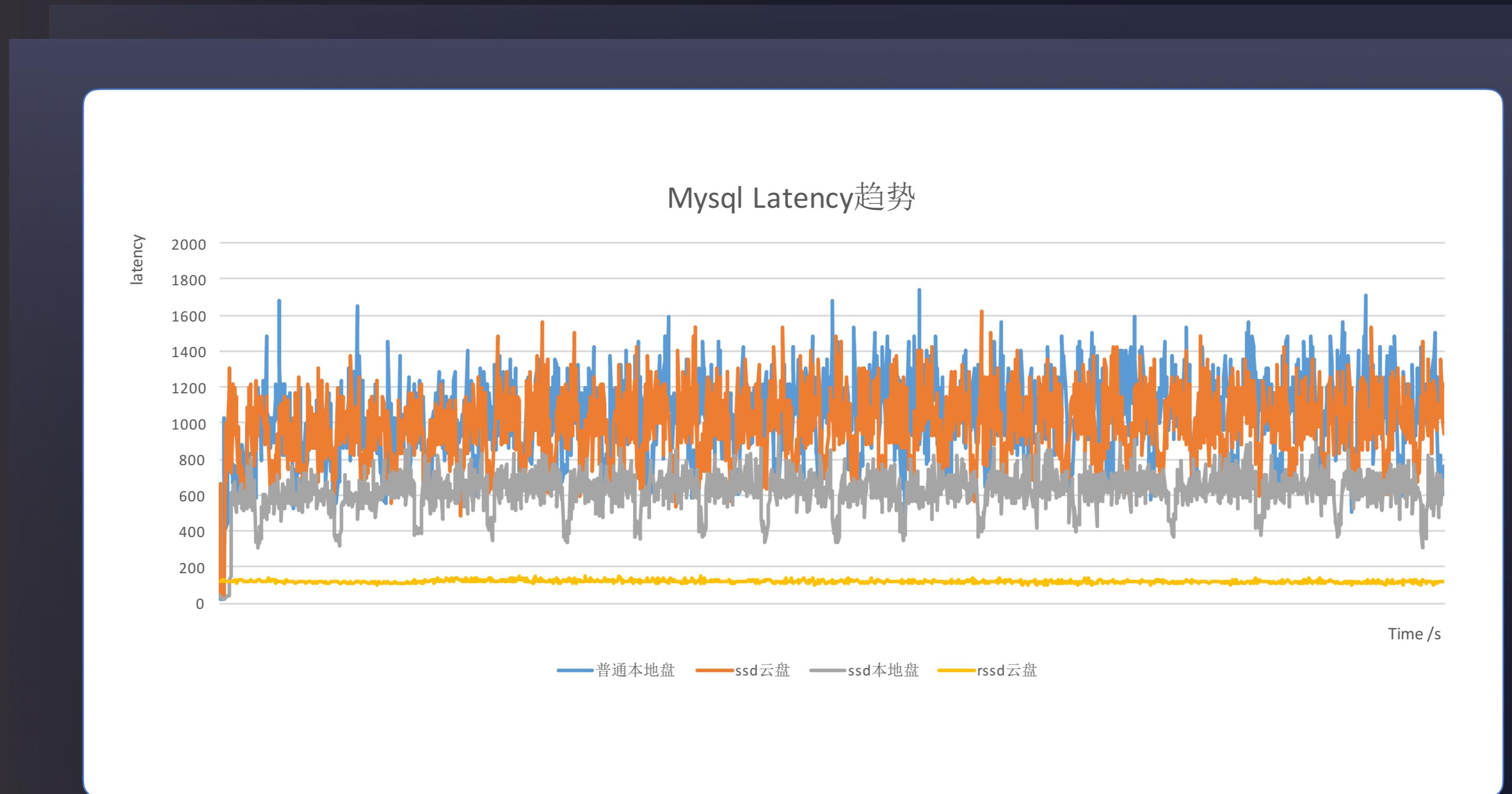
存储性能 云主机8C16G

IODepth 8*32 / BS 4k











THANKS