

存储分层 企业数据存储类型选择与优化

主讲人：周皓

阿里云对象存储OSS

面向互联网的非结构化数据存储服务，适合视频、图像、网页、日志、文本等类型的文件，广泛应用在移动应用、企业级备份、音视频、基因测序等领域



可靠



安全

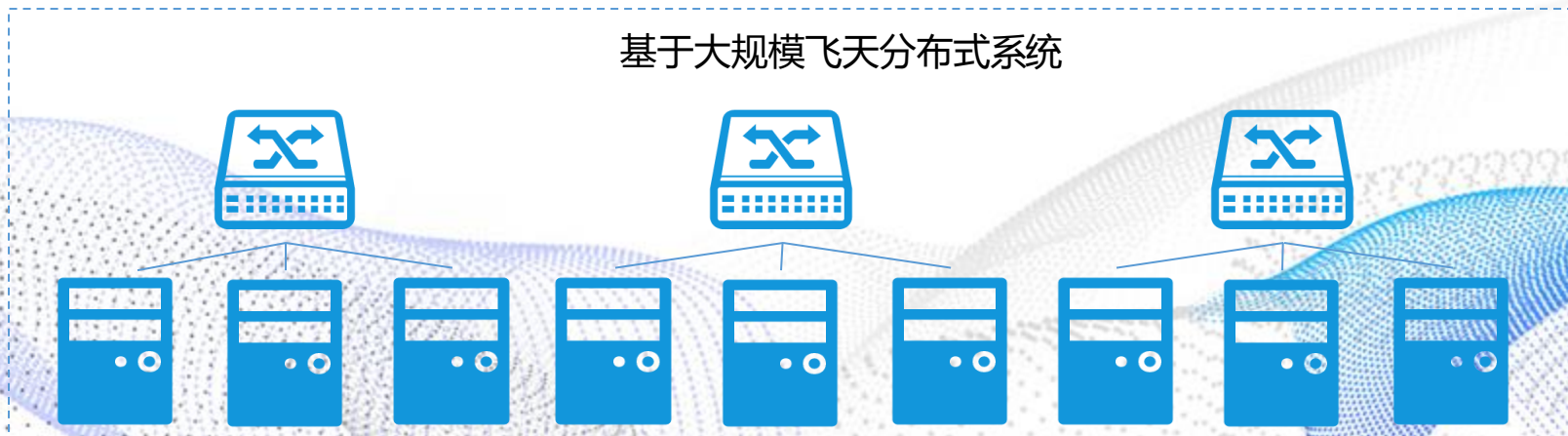


易用

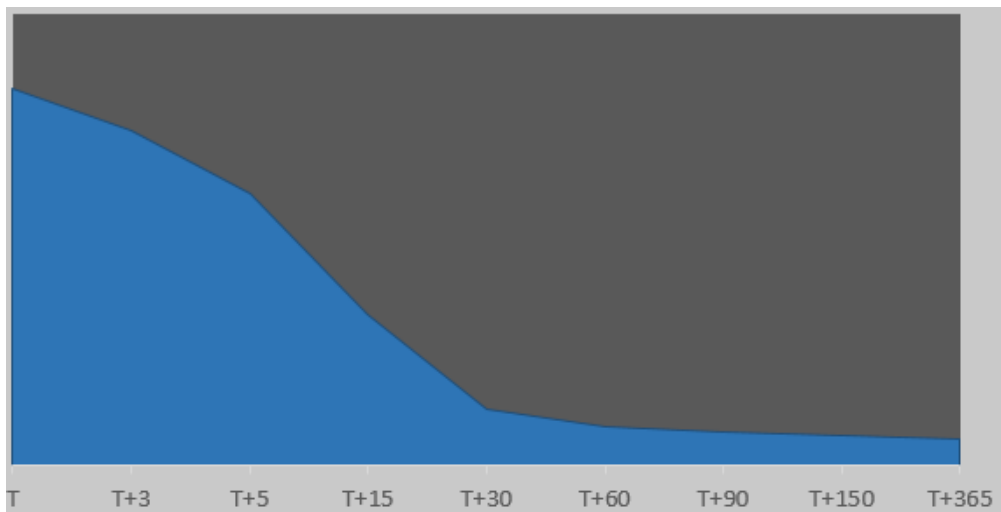


弹性

基于大规模飞天分布式系统



互联网数据访问特点



- 数据访问的热度集中在数据创建后一周内
- 一个月之内访问热度大幅度下降
- 超过3个月以上访问频率趋向稳定

归档数据特点

Long Term
Low Cost
Latency tolerance
High Security

- 存储周期长，几年、几十年甚至更长
- 在存储周期内的存储成本控制
- 实时访问要求不高，可以接受等待时间
- 安全性高，很多数据需要符合合规

OSS的存储类型



标准类型Standard



低频访问类型IA



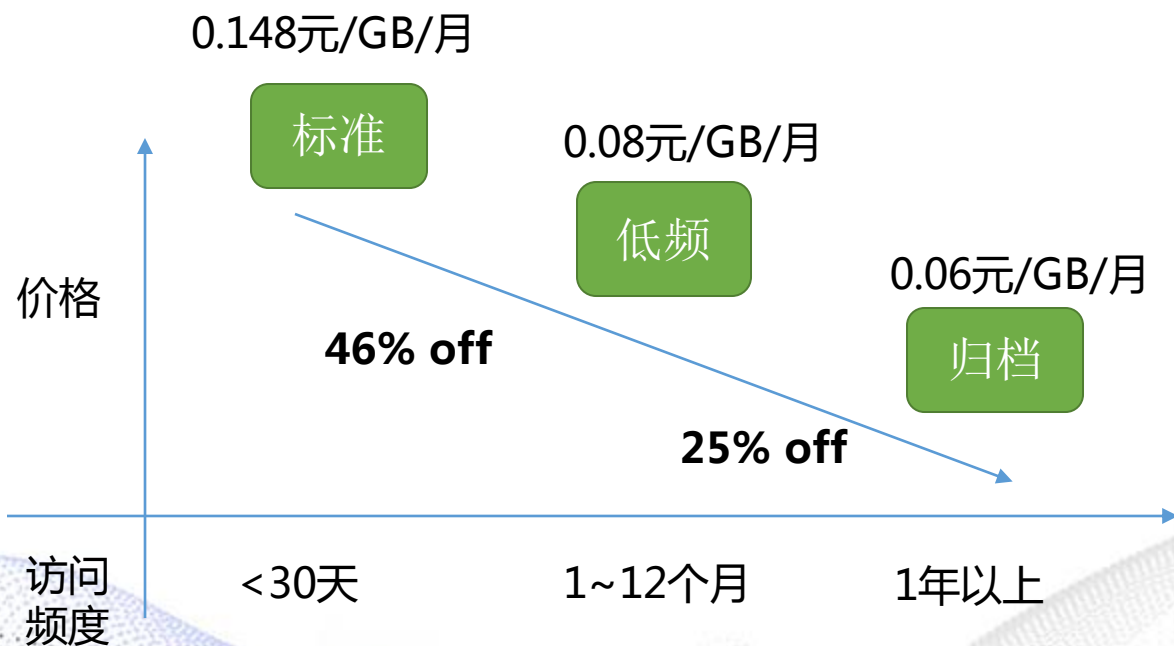
归档类型Archive

热数据




很少访问的数据

归档冷数据

存储类型存储单价对比



存储类型特点

 标准类型	热数据、 临时数据	实时响应 ms延迟	0.148元/GB/月	≥ 0 天	> 0 KB
 低频类型	温数据、 长期存储	实时响应 ms延迟 数据取回 0.0325元/GB	0.08元/GB/月	≥ 30 天	≥ 128 KB
 归档类型	冷数据、 长期存储	数据读取需要解冻 1h~4h 数据取回 0.06元/GB	0.06元/GB/月	≥ 30 天	≥ 128 KB

L
I
F
E
C
Y
C
L
E

互联网UGC/PGC应用存储特点

文件上传后的一个阶段内访问集中，需要支撑的大请求量

超过30天后请求量下降明显，90天后几乎没有访问

对于访问实时性有要求，即使是很久前的数据，也只能接受几秒钟左右的等待时间

有些应用对访问资源的URL要求整个生命周期内始终保持不变

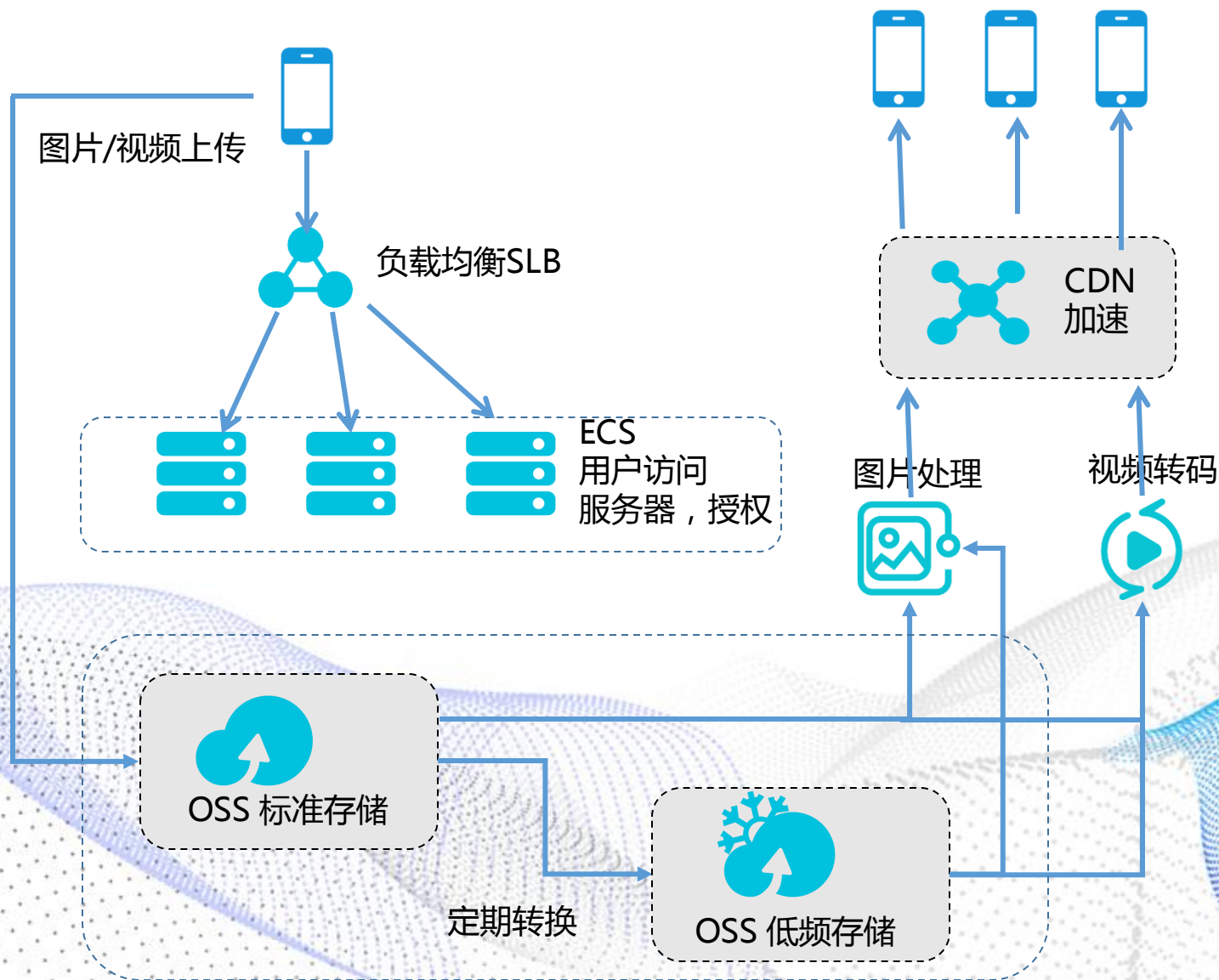
UGC应用通常会配合使用图片处理、视频转码等功能

何时转储成IA最省钱

50TB数据	平均每月访问频率1次	平均每月访问频率10次	标准存储类型
存储费用	4096.0	4096.0	7577.6
请求费用	26.2	262	2.6
数据取回费用	1664	16640	0
总和	5786	20998	7580

- 建议对存储时间超过3个月以上的转储成IA
- 每月平均访问频率超过3次的不建议存储成IA
- 可以通过开启OSS访问日志、分析Object的访问频率变化，确定转换周期

UGC应用存储优化-示例



数据备份

备份类数据通常保存时间长、在保存周期内访问频率少

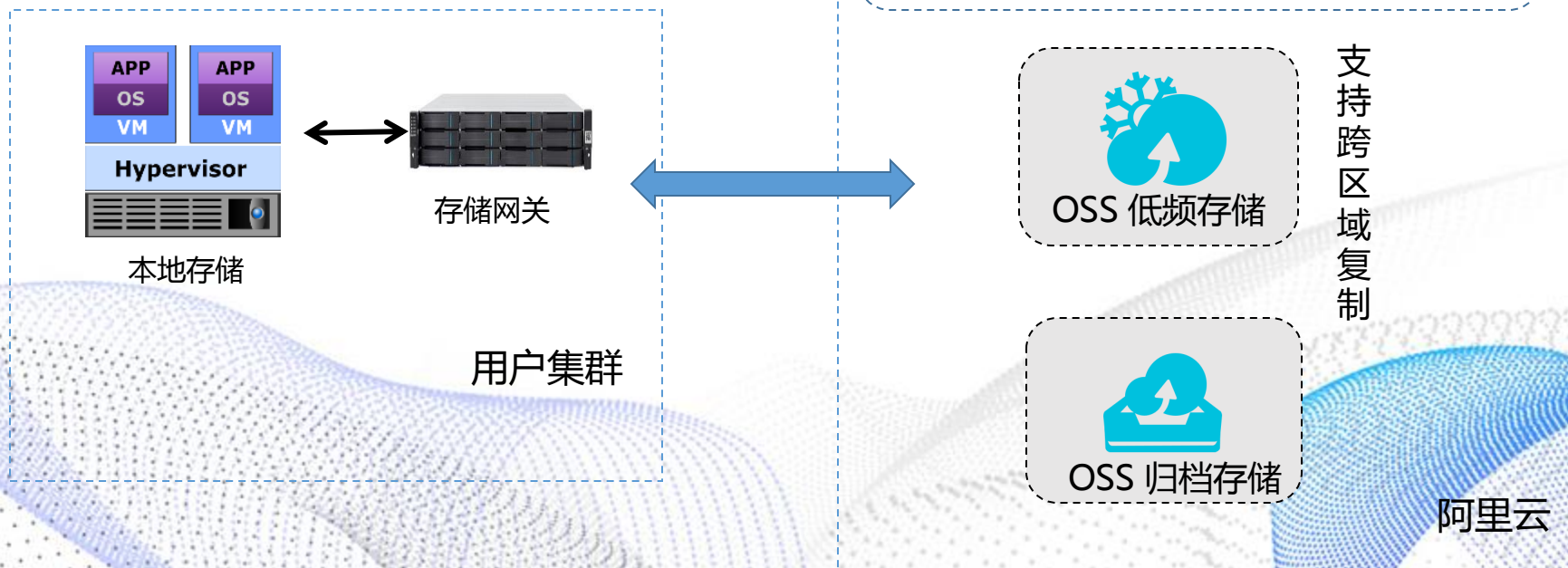
在线类数据的备份需要恢复数据时、备份数据需要被及时访问到

当出现本地集群故障时，能配合其他云产品，在云上继续提供业务

离线类备份数据可以容忍较长的恢复时间

混合云数据备份-示例

- 在线类数据备份到低频存储
- 出现机房级故障，使用云端ECS恢复在线业务
- 长期不访问的冷数据离线备份可以使用归档存储



归档存储优化

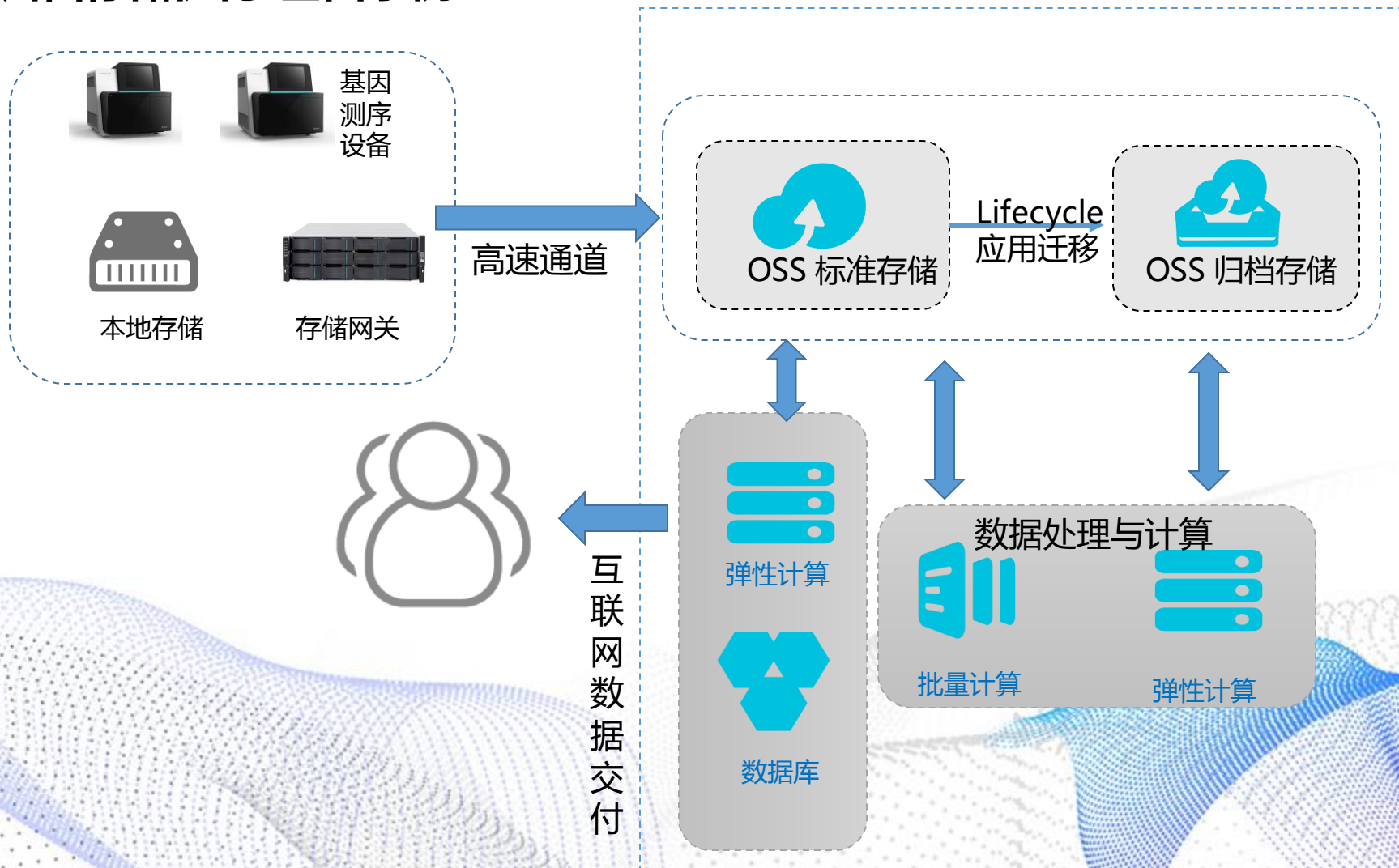
业务形态对归档类数据能容忍较长的恢复时间

绝大部分数据在存储周期内不会被读取到

文件大小如果过小，可以合并成大文件存储

具备文件生命周期内跟踪访问情况的能力

归档存储优化-基因示例





飞天·智能

APSARA INTELLIGENCE

2017云栖大会·成都峰会

5月23日 成都世纪城天堂洲际大酒店