

# 百万亿级云存储系统挑战 与实践

TI-CloudStorage@ByteDance

# QCon+ 案例研习社



扫码学习大厂案例

## 学习前沿案例，向行业领先迈进

40个  
热门专题

—  
行业专家把关内容筹备，  
助你快速掌握最新技术发展趋势

200个  
实战案例

—  
了解大厂前沿实战案例，  
为 200 个真问题找到最优解

40场  
直播答疑

—  
40 位技术大咖，每周分享最新  
技术认知，互动答疑

365天  
持续学习

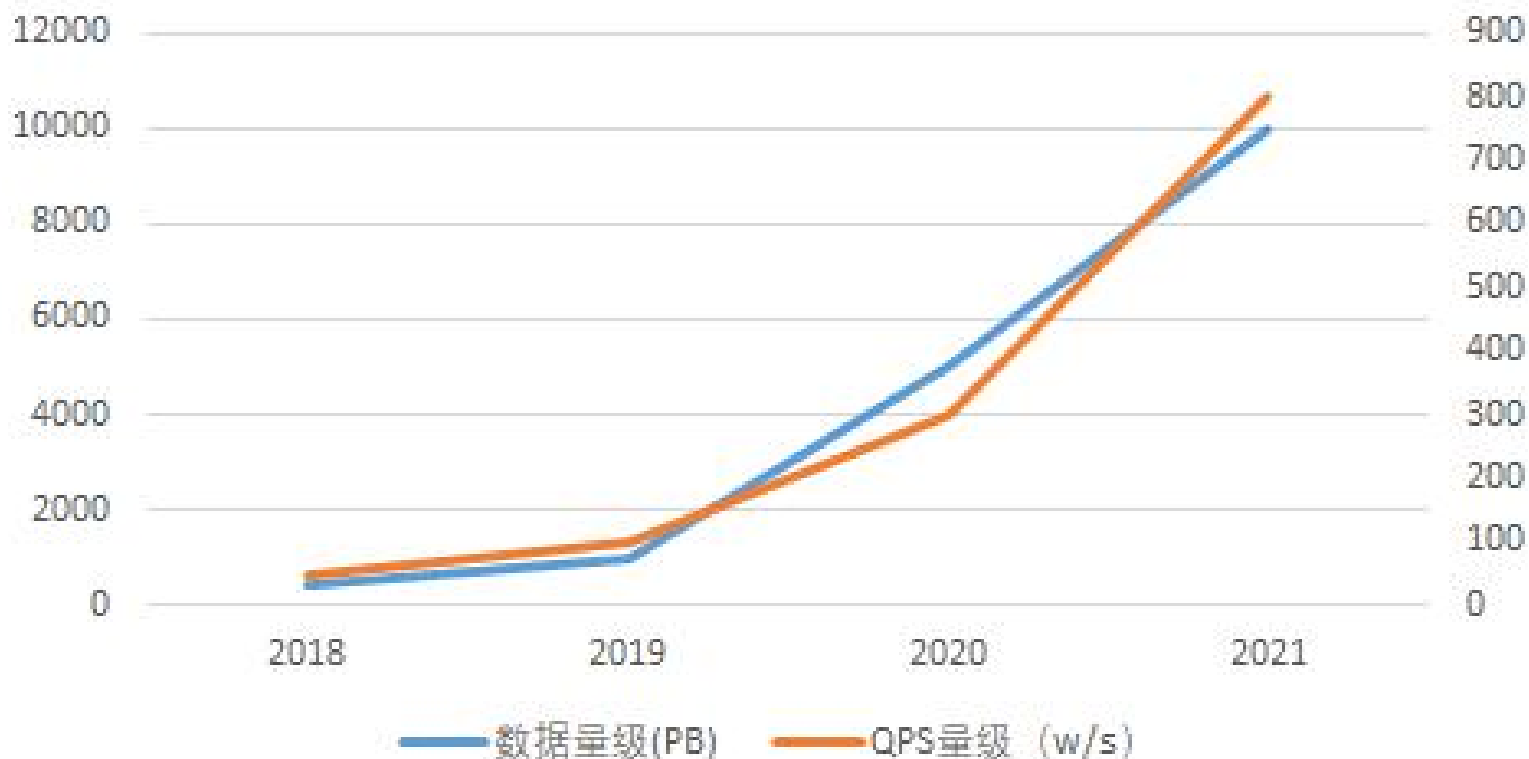
—  
视频结合配套 PPT  
畅学 365 天

# Agenda

- 问题和挑战
- 云存储架构
- 优秀实践
- 展望与探索
- Q&A

# 数据量爆发式增长

数据增长量级变化



- 数据量年化平均增长  
**200%**
- 带宽消耗平均读写  
**25+TB/s**
- QPS增长年化平均增长  
**180%**

# 海量数据的背后

20,000+



工程师

900,000+



服务器

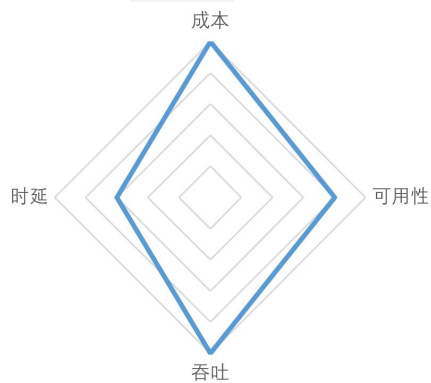
60PB+



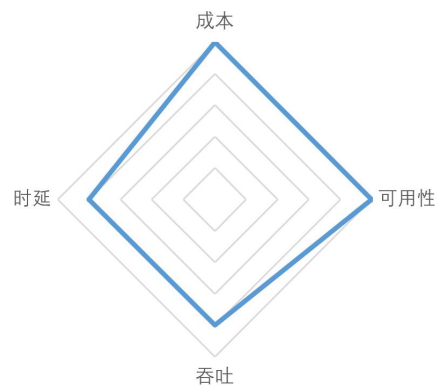
日增数据量

# 业务的挑战

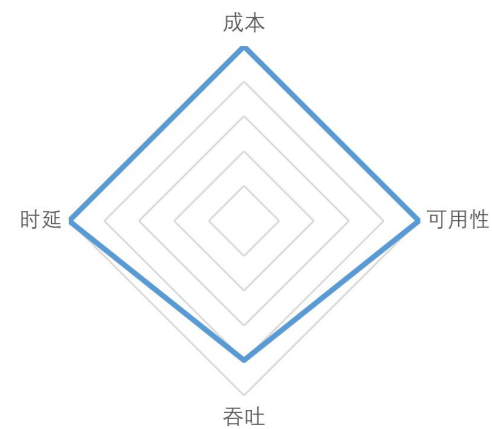
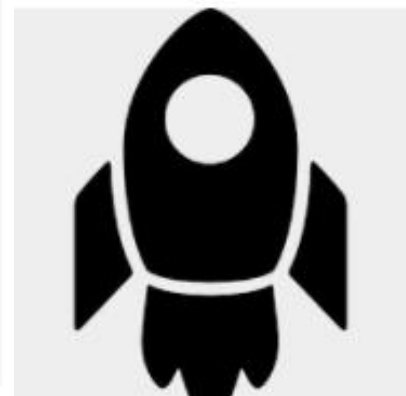
离线



近线



在线



# 基础设施的复杂性



## 网络设施复杂

- 地域间带宽不稳定
- 机房间带宽不均衡
- 机房内异构环境复杂

## 混部环境复杂

- 资源争抢
- 节点异构

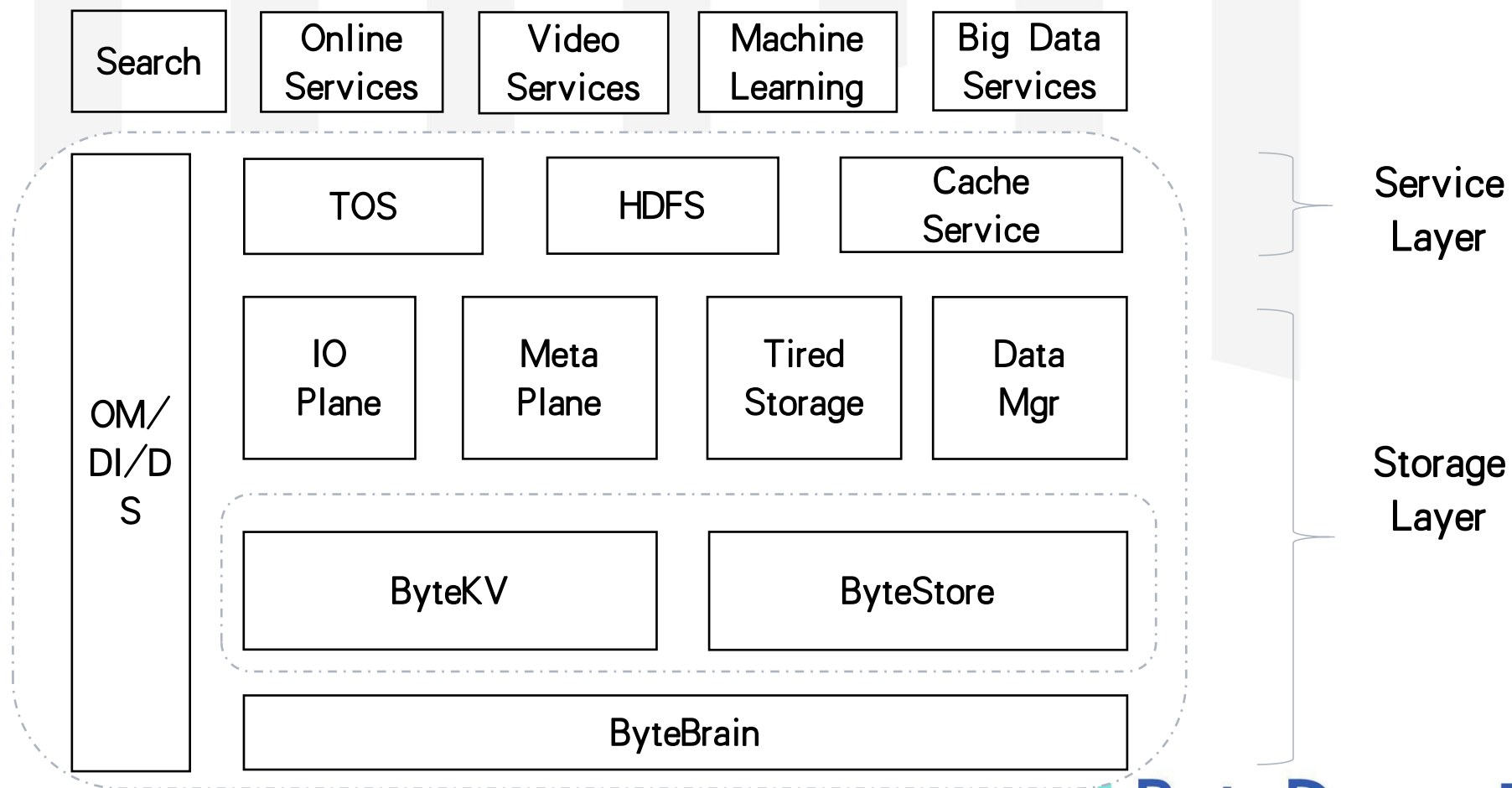


## 资源分布不均衡

- 存量不均衡
- 增量不均衡
- 供货不确定性增大

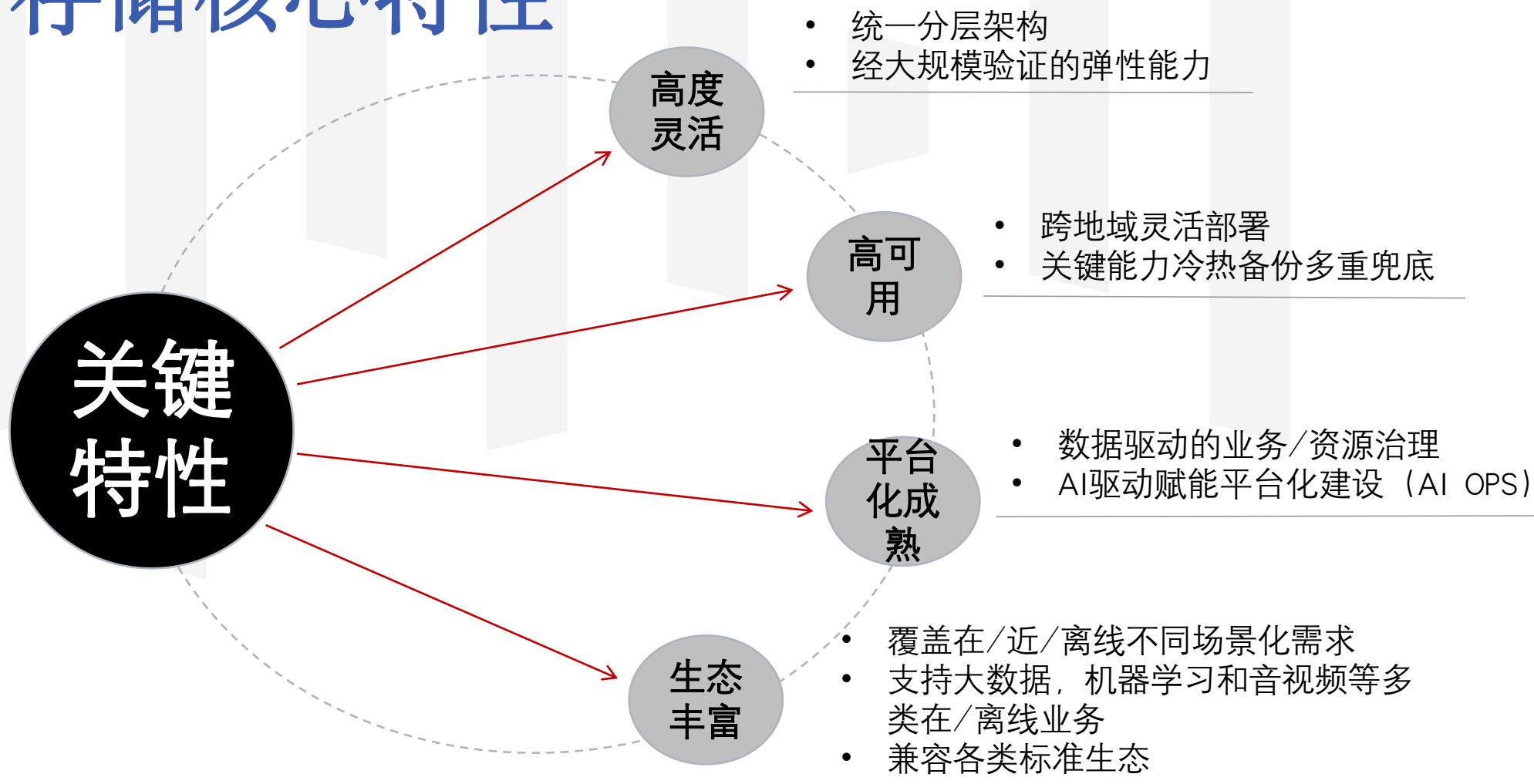


# 云存储架构简介

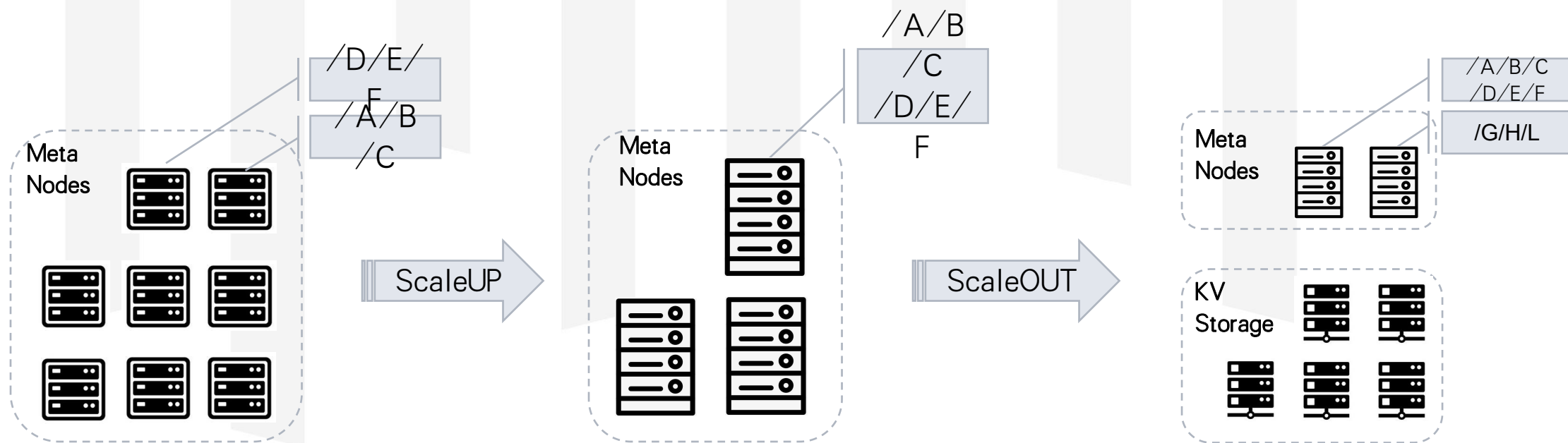




# 云存储核心特性



# 优秀实践 - 离线元数据体系优化



痛点:

- 不稳定
- 性能差

解法:

- C++化
- 关键算法重新设计

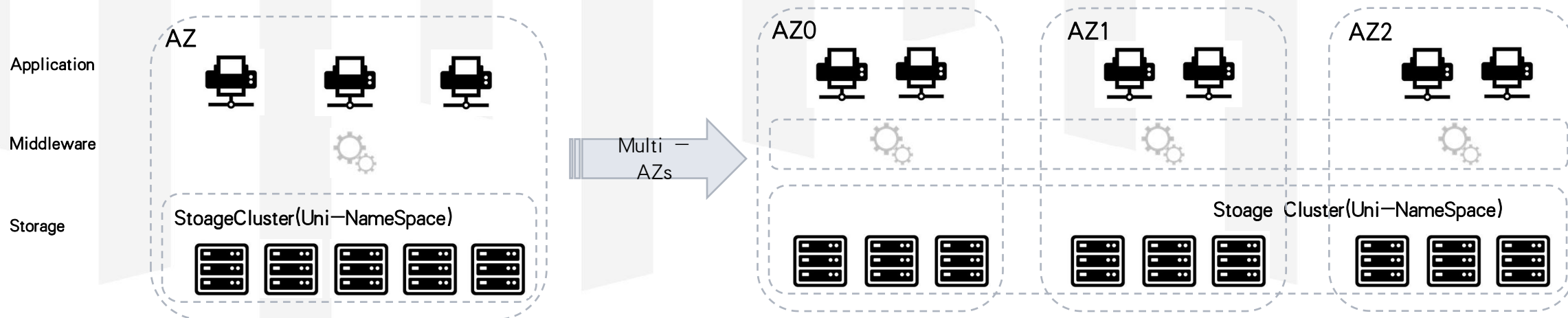
痛点:

- 可扩展性差
- 运维成本高
- 利用率低
- 成本高

解法:

- 存算分离
- 分层架构

# 优秀实践 - 离线容灾体系建设



## 痛点:

- 多AZ容灾需求
- 多AZ资源分布不均衡
- 机房内外带宽管理难度大

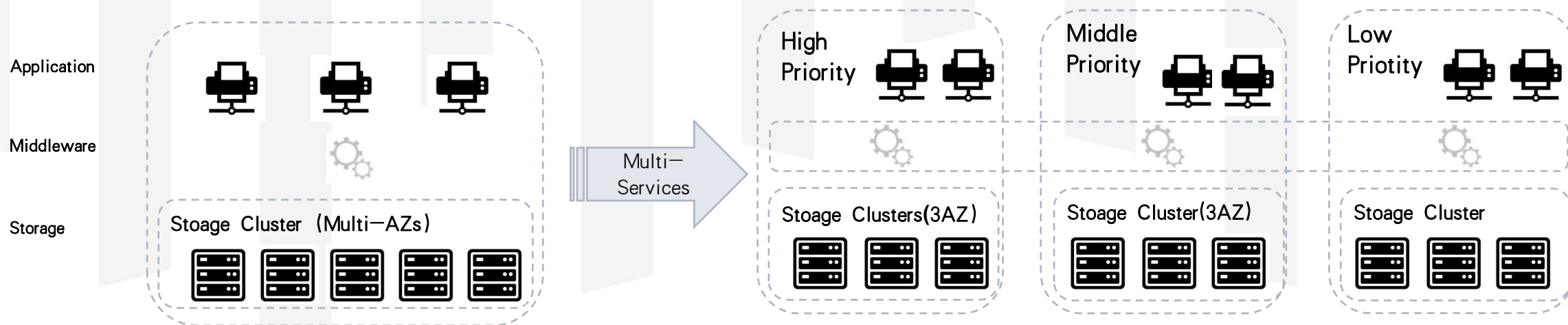
## 解法:

- 跨AZ高可用架构 (元数据)
- 离线统一目录视图屏蔽基础设施复杂性
- 带宽分级管理

## 持续优化:

- Data Management帮助业务无感伸缩
- 增强元数据处理能力实现多AZ统一元数据管理

# 优秀实践 - 在线高可用架构



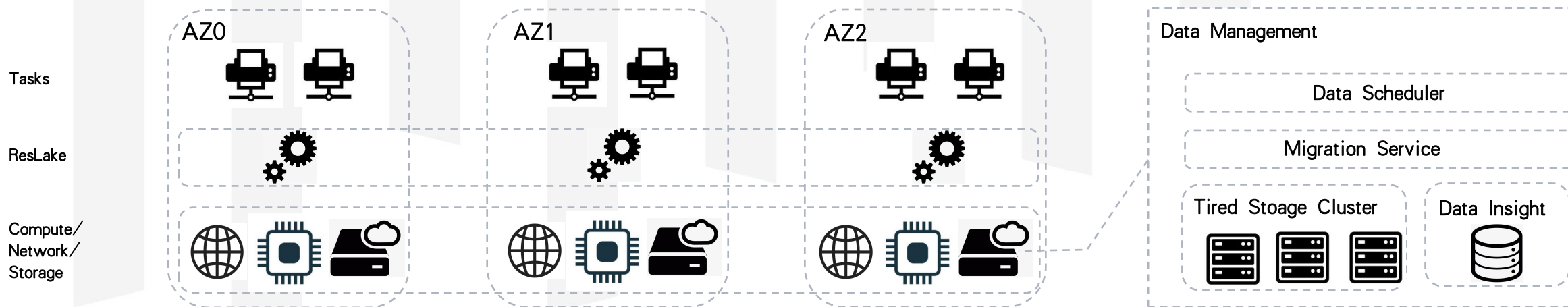
## 痛点:

- 不同业务不同可用性/成本需求
- 基础设施频繁变更
- 运营运维代价大，业务心智负担重

## 解法:

- 小集群垂直化建设
- 分级保障（差异化服务）
- 加强灾备体系建设
- 运营运维平台化建设

# 优秀实践 - 离线数据治理



## 痛点:

- 存储成本高企
- 多AZ基础设施频繁变更
- 计算任务和存储强相关, 业务扩容和迁移心智负担太重

## 解法:

- 构建分级存储
- 构建DataInsight能力
- 构建资源迁移平台化能力
- 构建计算存储网络三位一体调度系统

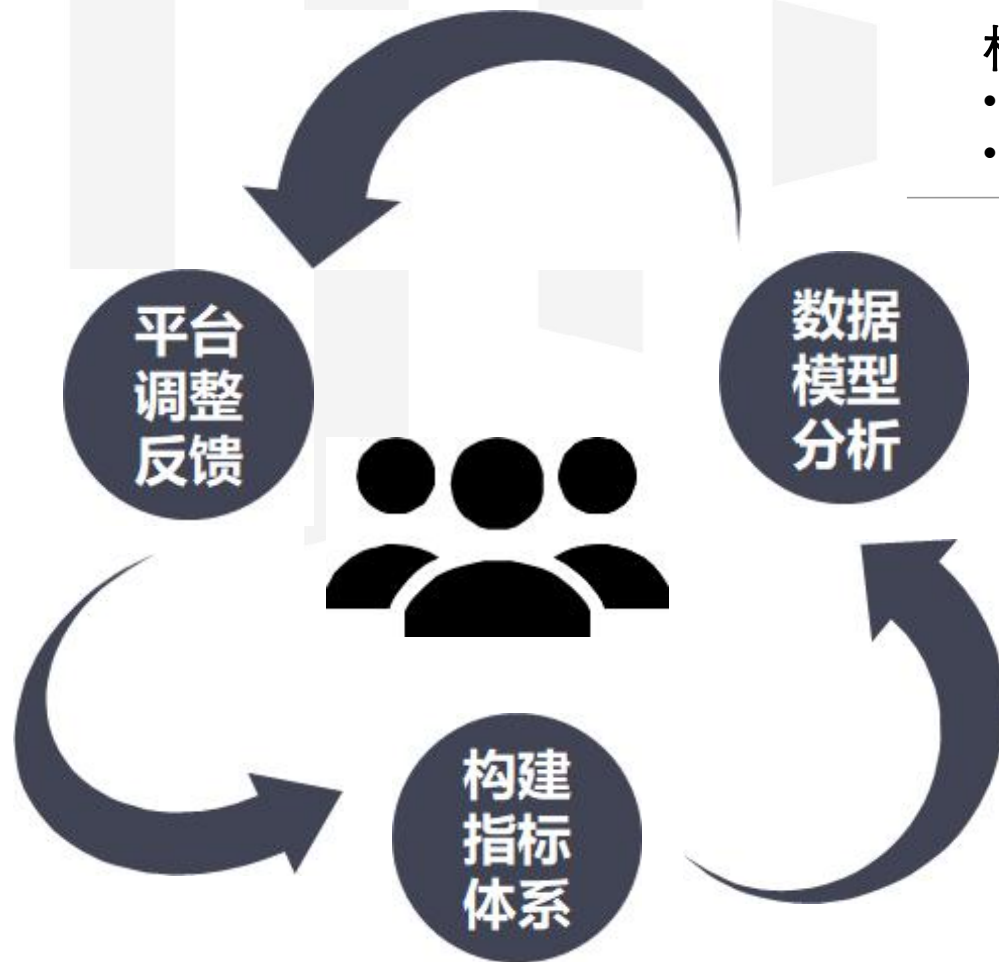
# 优秀实践 - 在线业务服务治理

## 痛点:

- 业务访问模型多样化
- 业务SLA需求多样化
- 基础设施资源管理难度大

## 解法:

- 构建数据驱动的服务治理体系



## 构建指标体系

- 建模不同业务访问指标
- 建模关键资源使用指标

## 数据模型分析

- 实时/离线分析业务指标变化
- 分析关键资源使用指标变化

## 平台调整反馈

- 反馈SLA变化, 重保高优业务
- 预测关键资源变化趋势
- 指导调整资源热度变化 (Tiring)
- 资源变化帮助业务及时调整和基础设施资源管理

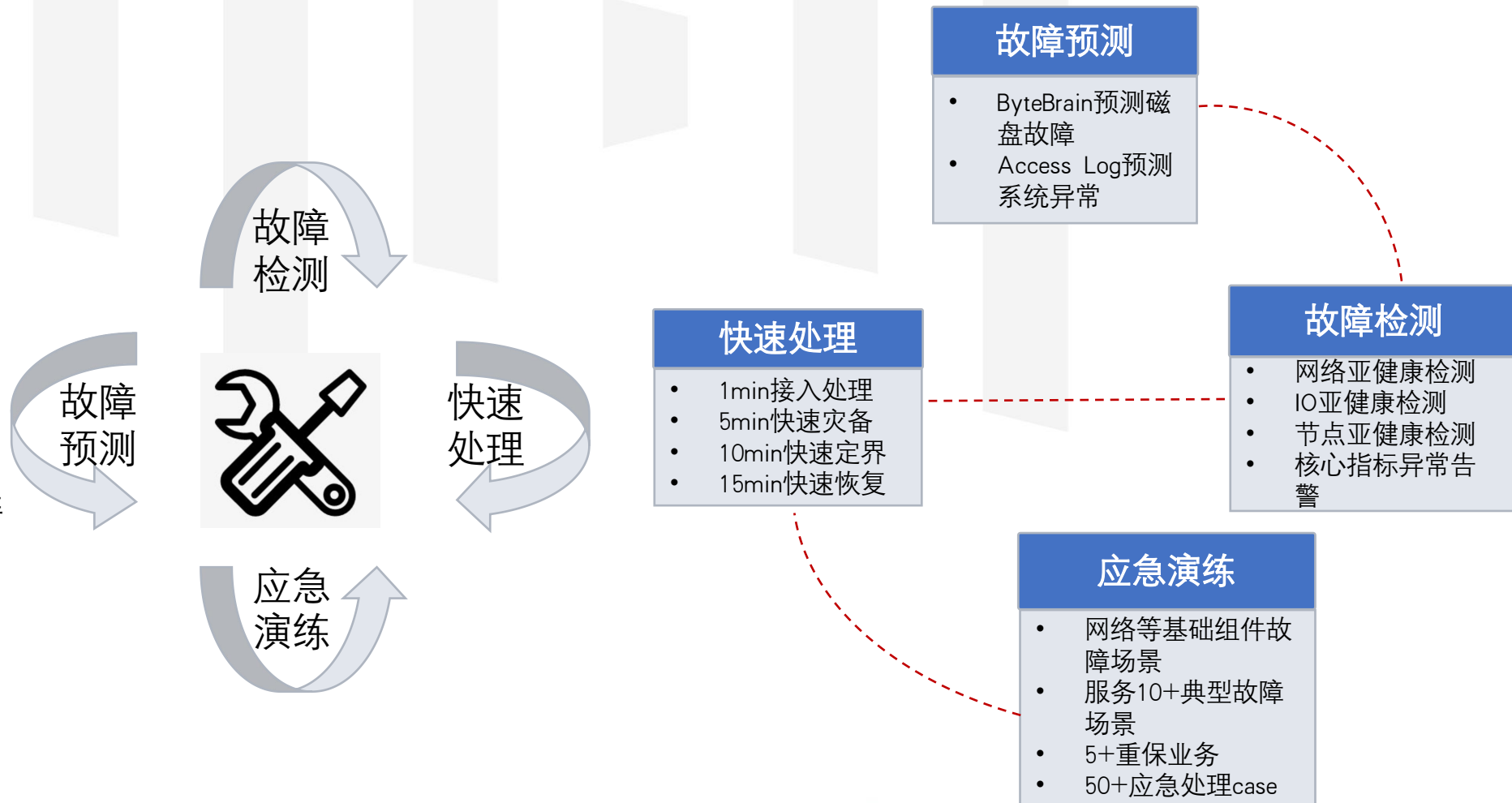
# 优秀实践 - 在线服务之故障应急

## 痛点:

- 在线高优业务对存储系统可用性要求高
- 系统规模大, 基础设施复杂, 稳定性不容易保障

## 解法:

- 加强平台化建设, 完善监控-告警-处理流程
- 增强故障预测能力
- 定期应急演练





# 优秀实践 - 离线生态服务

## MQ存算分离实践

- 1万+节点高可用集群设计和落地
- 极致计算存储独立横向扩展能力

## 平台化能力增强

- 监控运维自动化
- 运营服务平台化
- 安全服务能力平台化

## 数据湖深度实践

- 格式层能力增强
- IO优先级
- 海量元数据访问支持
- 缓存能力(类Alluxio)

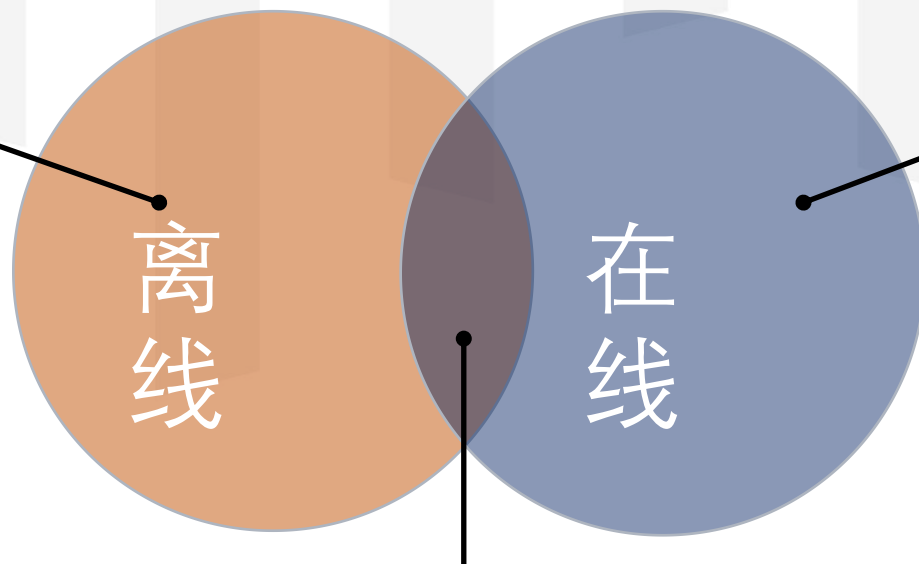
## 缓存能力增强

- 计算侧结构化数据访问加速
- 高效支持机器学习场景
- 计算引擎远程存储访问加速



# 优秀实践 - 成本治理

- 跨地域资源调度应对定期/即时任务
- 存储计算网络统一调度
- 数仓结合识别血缘实现智能流转/沉降



- 分布式GC能力
- 三机房EC
- ByteBrain智能预测资源使用变化
- Reblance能力增强

- 统一分级存储建设
- 统一分层架构演进
- 池化能力增强
- 统一资源管理/调度

# 展望与探索

## 存储计算结合更加紧密

- 计算引擎和存储系统趋于融合
- 格式层，缓存等存储能力进一步加强
- 计算能力进一步下推至存储侧



思考  
总结



## 非结构化数据处理在离线趋近统一

- 非/半结构化到结构化统一元数据
- 分层架构下统一提供底层存储能力

## 云原生化

- 存储能力的弹性
- 运营运维体系的统一
- 成本优化和使用效率进一步提升

# 为一线互联网公司核心技术 人员提供优质内容

☑ TGO专访

☑ 技术干货

☑ 每周精要

☑ 行业趋势



关注 InfoQ 公众号

THANKS

