

# 火山引擎云搜索服务 大规模技术应用实践

桂峰 | 火山引擎云搜索资深架构师

自主 · 创新 · 引领

# 目录 | Contents

- 01 云搜索服务
- 02 高可用架构
- 03 向量搜索引擎
- 04 混合搜索

## Chapter 01

# 云搜索服务

## 火山引擎云搜索服务

- 全托管一站式分布式搜索服务，完全兼容搜索引擎 OpenSearch、ElasticSearch。
- 多年来持续为包括头条、抖音、飞书、懂车帝等字节跳动内部核心业务提供搜索服务支持。历经多年业务发展，云搜索服务累积丰富的业务实践并提供给内外部客户。



全文检索



向量检索



Geo



结构化搜索



其他



100W+ Core



100PB+ 存储

## Chapter 02

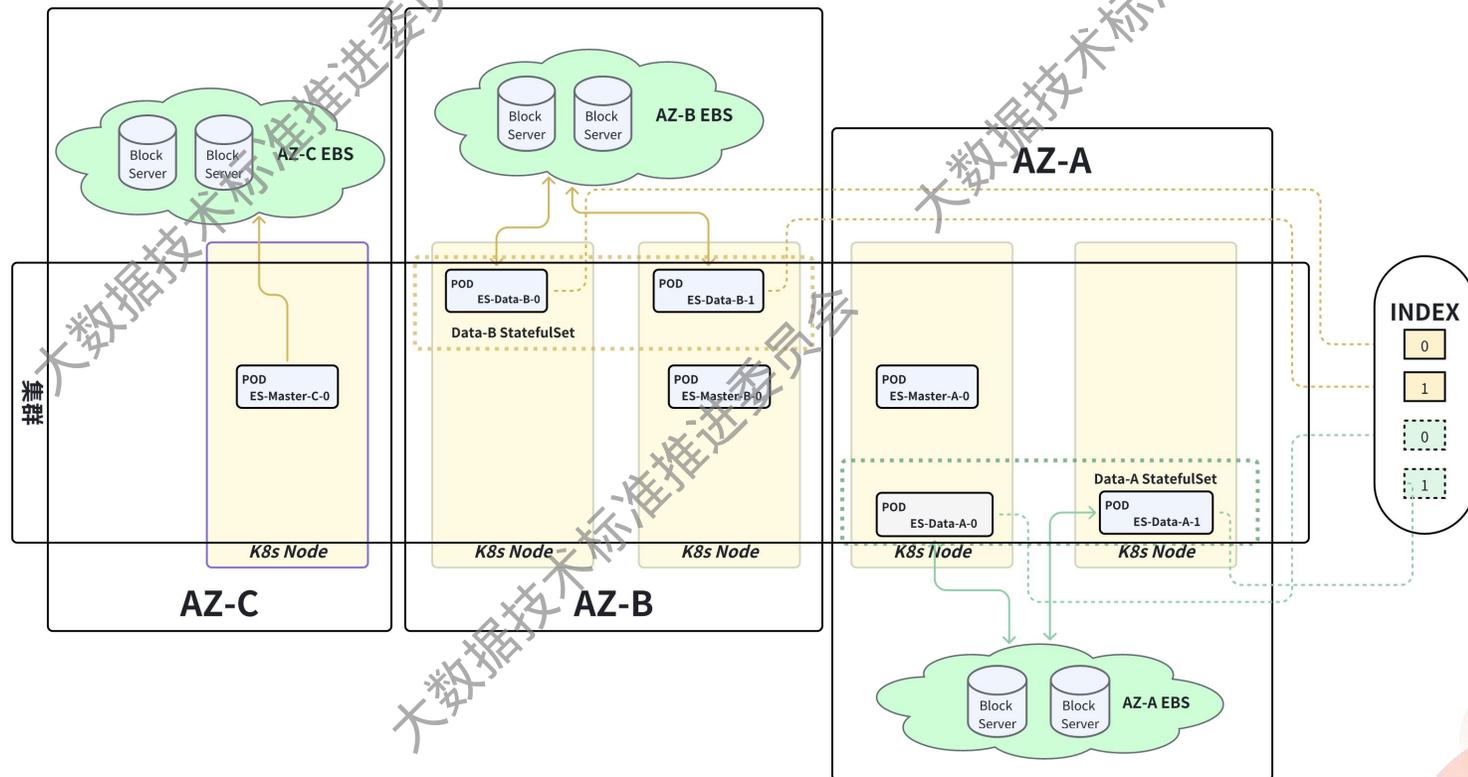
# 高可用架构

## AZ 容灾方案

- 数据节点多 AZ 部署：每个AZ一组 k8s StatefulSet，存储使用同AZ的EBS云盘
- 结合 ES Allocation Awareness，保证同一 Shard 的主、副分片分配到不同AZ

## 容灾能力

- 支持 EBS 云盘容灾，单 EBS 集群故障时集群可正常工作
- 支持 AZ 容灾，单 AZ 故障时集群可正常工作



## 常见多活方案

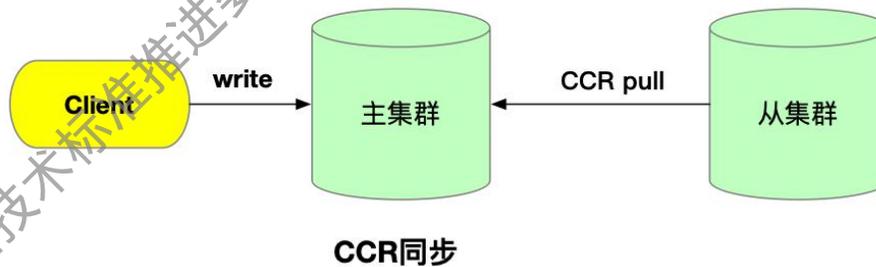
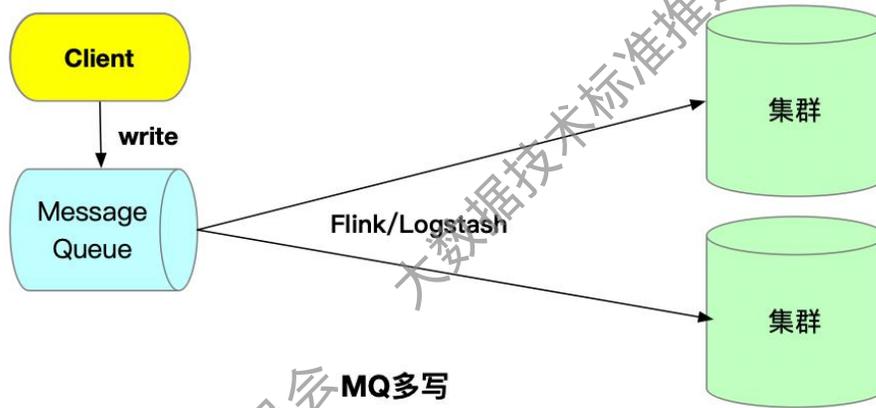
|        | 单向同步 | 双向同步 | 业务改造 | 系统复杂性 | 数据一致性 |
|--------|------|------|------|-------|-------|
| MQ 多写  | ✓    | ✓    | 中    | 高     | 弱     |
| 业务多写   | ✓    | ✓    | 难    | 高     | 弱     |
| CCR 同步 | ✓    | ✗    | -    | 低     | 高     |

## 局限性：无法满足

业务单写场景

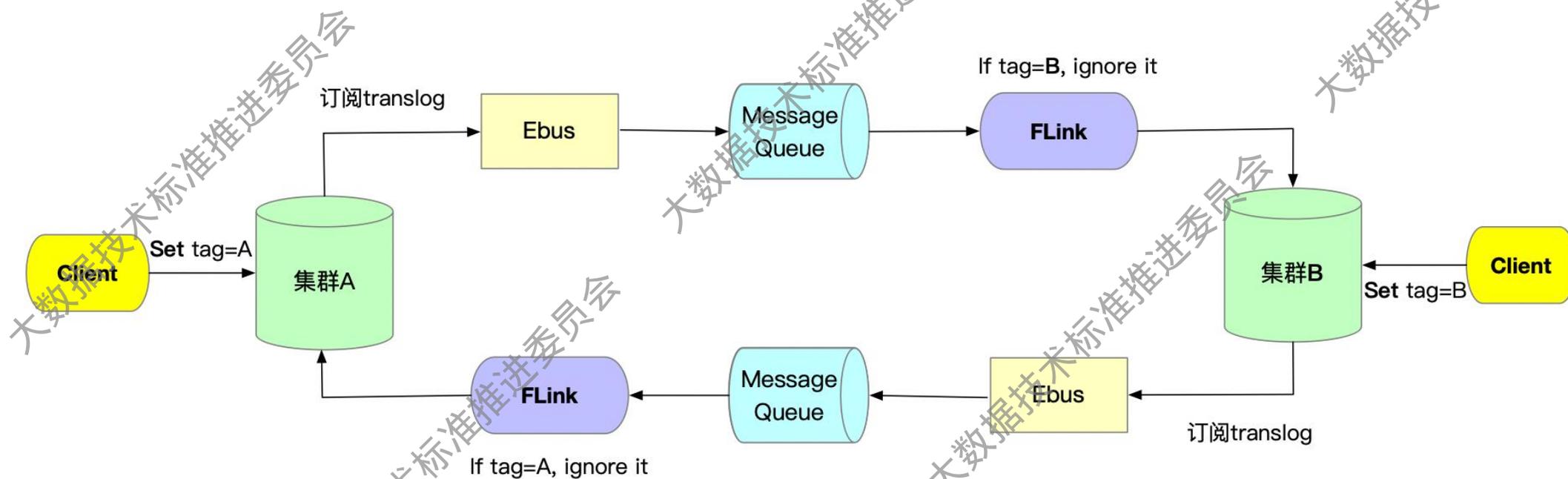
+

双向同步需求



## Ebus 多活双向同步方案

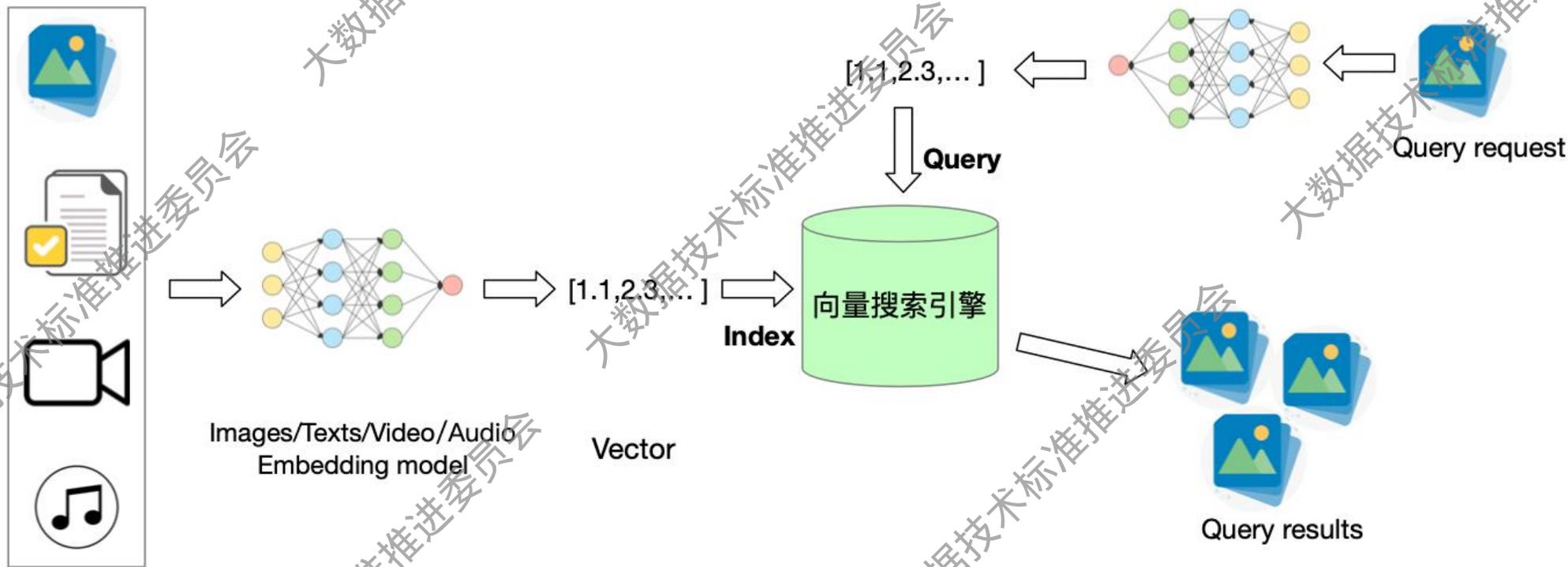
- Ebus组件: 订阅消费集群的Translog
- 防回环: 针对业务写入数据打Tag & 消费侧过滤
- 实时增量数据一致性检验



## Chapter 03

# 向量搜索引擎

## 向量搜索引擎：大模型时代重要的存储



以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/825341022123011312>