

AmzTrends x TiDB Serverless：通过云原生改造实现全局成本降低 80%

 TiDB社区小助手 发表于 2024-04-11

转载

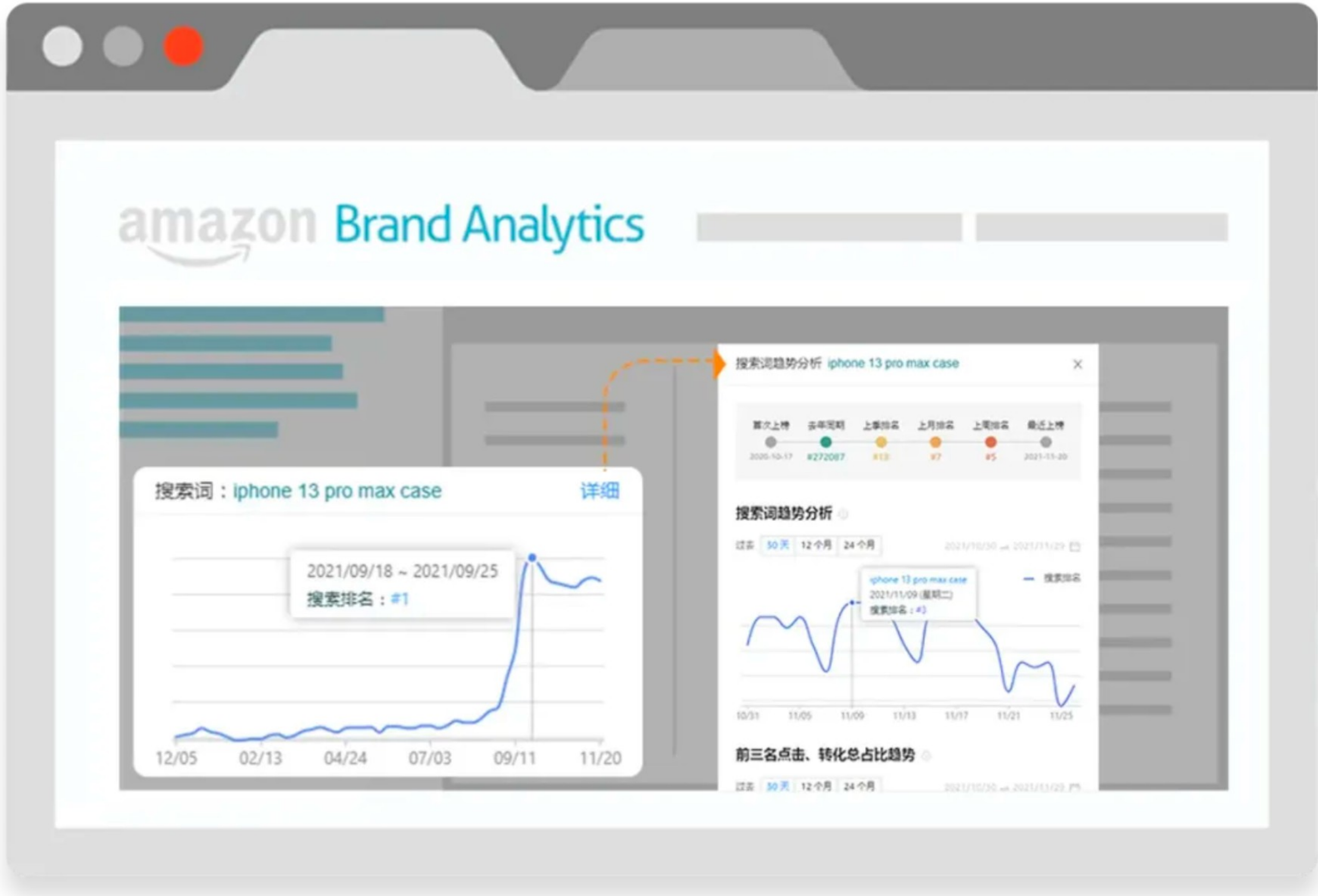
TiDB Cloud # 实践案例

本文作者：Steve | 笛卡尔数据创始人，5 年亚马逊大数据分析 with 产品开发经验，深度研究品牌分析、商机探测、类目分析，深度研究 COSMO 算法底层思维与数据构建逻辑

导读

本文介绍了厦门笛卡尔数据（AmzTrends）在面临数据存储挑战时，选择将其数据分析服务迁移到 TiDB Serverless 的思路和实践。通过全托管的数据库服务，AmzTrends 实现了全局成本降低 80% 的效果，同时也充分展示了 TiDB Serverless 在简化架构、提升性能和降低成本方面的优势。未来，AmzTrends 计划继续利用 TiDB Serverless 的优势，扩展业务并提升竞争力。

厦门笛卡尔数据是一家专注于跨境电商数据分析的 SaaS 公司，AmzTrends 为亚马逊卖家提供品牌分析(ABA)、商机探测以及广告数据的可视化分析。目前，AmzTrends 主要以 SAAS 和 Chrome、紫鸟浏览器的插件形式为客户提供数据服务，以订阅模式为美国、日本、中国数万计的跨境电商卖家提高数据服务，帮助卖家在选品、运营、广告等运营环节提供专业的数据分析决策价值。



1. 业务挑战

AmzTrends 的数据主要以大单表的形式进行存储，最大的表数据量超过 22 亿，字段较多且某些字段很长的大宽表，单表中存在结构化与非结构化的数据结构，因此需要建立大量的索引，占用大量存储空间，而且过期数据还需要定期清理，经常使用 BATCH 进行批量操作，一旦遇到异常无法无法事务的一致性，因此数据维护压力巨大。

2. 技术痛点

在业务初期，AmzTrends 选择了在百度云上自建 TiDB 集群，资源按月付费。集群规模包含 1 个 TiDB 节点、1 个 PD 节点（此种部署方式会损害 PD 的高可用性，为了节约成本的部署方式，是官方不推荐的高风险方案）、3 个 TiKV 节点，技术人员通过将亚马逊下载的原始 CSV 格式的数据批量写入到 TiDB 中进行数据分析。此外还单独配置了 3 台服务器部署 Spark，进行全量复杂的数据计算分析。但 Spark SQL 与关系型数据库不同，需要专人运维，复杂的业务架构造成了资源冗余，使得运维成本较高。



潜在客户



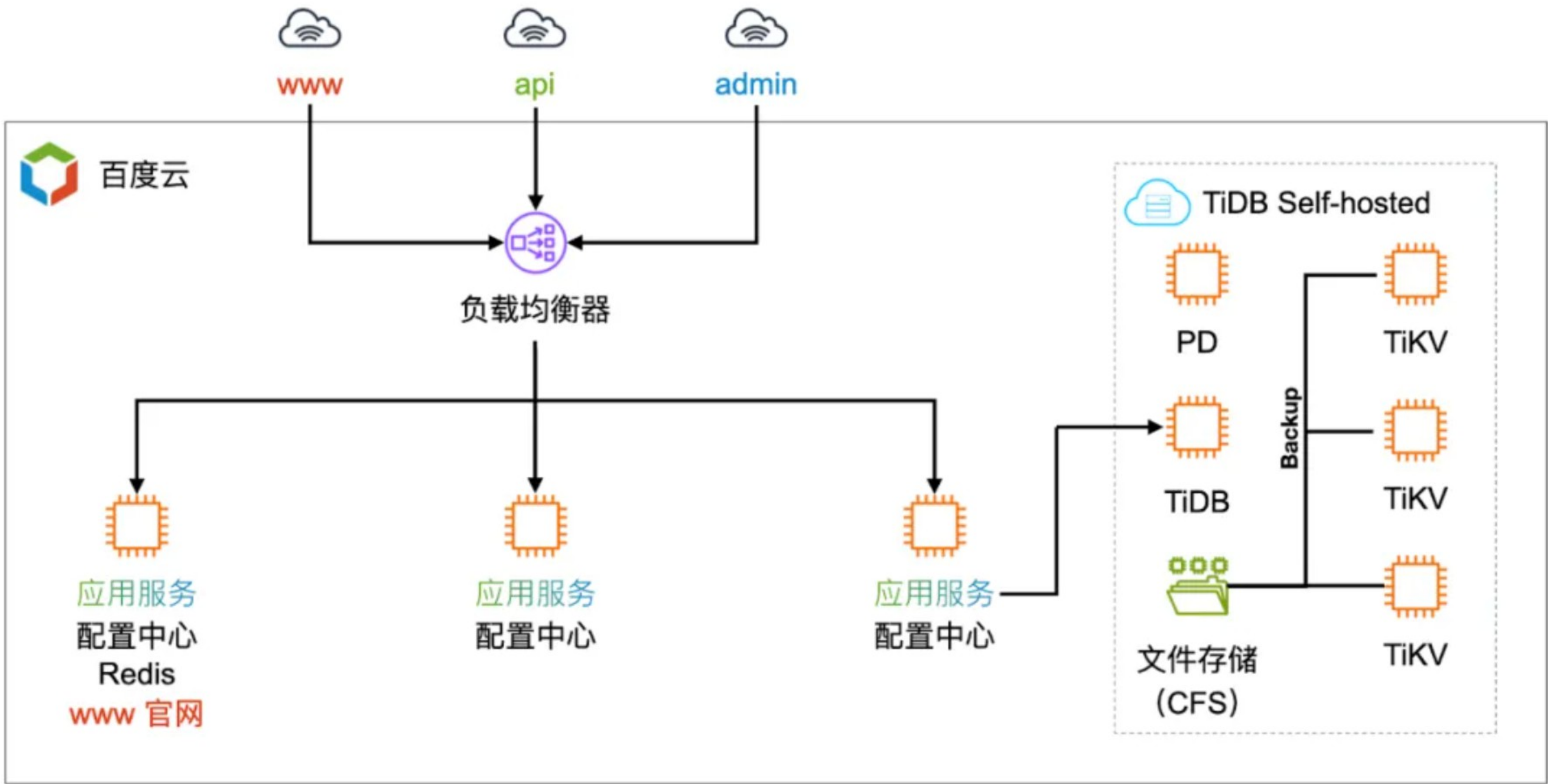
终端客户



管理员

导读

1. 业务挑战
2. 技术痛点
3. 为什么选择 TiDB Serverless?
4. 解决方案
5. 项目收益
6. 总结
7. 未来



由于集群配置不够加上专业运维团队的缺失，弃用了 Spark，大量复杂且数据计算量巨大的任务由 Spark 转到 TiDB 直接运行，集群越来越不稳定，数据安全面临严重挑战。在这种情况下，AmzTrends 不得不寻找对技术要求更低且更安全的运维解决方案。在接触到 TiDB Serverless 后，AmzTrends 认为由于都是 TiDB 产品体系，全托管的一栈式数据库服务 TiDB Serverless 不仅可以充分发挥 TiDB 数据库原有的特性和优势，还可以帮助公司简化架构，提升系统的整体性能和健壮性。所以，AmzTrends 决定将整体应用从百度云部署方案迁入 TiDB Serverless，不仅风险更低且兼容性和性能都能得到有效保证，另外通过数据容量与请求量的成本预估，AmzTrends 发现迁移后成本能比现有的云服务器部署更低，因此有了这样一次充满挑战的数据迁移过程。

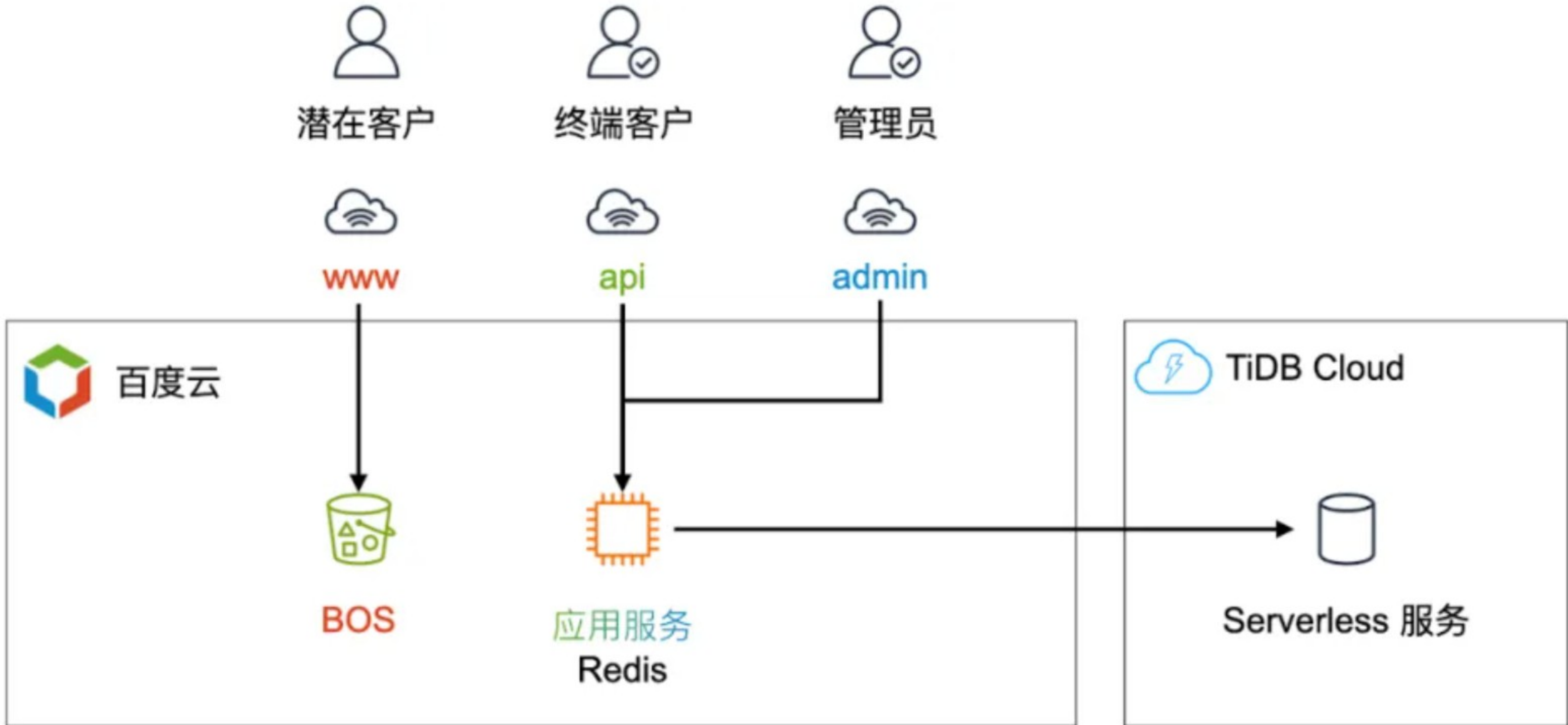
3. 为什么选择 TiDB Serverless？

TiDB Serverless 是一种全托管的、自动化的数据库服务（DBaaS），能够快速部署集群，并按实际使用资源量计费。结合自身业务架构特性，AmzTrends 在考虑 TiDB Serverless 时重点关注以下特性：

- **轻松扩展基础设施，按实际使用量付费：**商家只有在工作日时间才会使用 AmzTrends 的数据分析服务，而传统的云模式需要为预估的最大流量付费，容易造成闲时资源浪费。TiDB Serverless 可以秒级搭建起数据库环境，用户只需为实际的数据处理和存储付费，这让 AmzTrends 开发团队能够更专注于创新和发展。
- **轻松上手，无需手动分片：**以前，AmzTrends 的 20 亿数据需要定期进行清理或者分片，应用 TiDB Serverless 可以帮助用户摆脱数据分片等数据库繁琐的操作任务，相当于只维护一张大表即可，既可以满足更长时间维度的查询需求，又无需反复对数据进行整理操作。
- **自动伸缩，成本效益显著：**无论是存储、事务处理还是分析工作负载，AmzTrends 都能在几秒内根据应用需求自动伸缩，避免造成资源冗余浪费，既灵活又经济。
- **全面托管，无感运维：**为 AmzTrends 提供最优化的 TiDB 部署方案，免去了服务器配置维护的烦恼。以前一旦出现故障就需要人工进行紧急处理，使用全托管的 TiDB Serverless 后运维人员可以安心入眠。
- **熟悉的 SQL 接口，灵活应对业务增长：**TiDB Serverless 保持了开发者熟悉的 MySQL SQL 语言，AmzTrends 无需进行额外的学习即可高效地投入开发工作，同时 TiDB Serverless 具有一个高度灵活扩展的分布式架构，可以随需弹性扩缩容，应对业务的快速增长。

4. 解决方案

决定迁移后，AmzTrends 采用 TiDB Serverless 替代了云上自建的 TiDB 集群，将数据全量迁移到 TiDB Serverless 中，简化了架构，大大缓解了运维压力，降低总体 IT 成本。



迁移后，AmzTrends 在 PingCAP 官方服务的帮助下也进行了架构简化、优化建表、索引优化，提高了整体数据处理效率，并使 AmzTrends 业务架构演进至云原生模式：

- **索引优化：**通过发现和删除冗余数据索引，节省了 50% 的数据库成本；
- **SQL 查询性能提升：**优化索引后，不仅提高了查询速度，也显著减少了成本；

- 架构简化：自建配置中心和服务注册要求的最小三台云服务器转变为单台云服务器，降低了成本和复杂性；
- 业务缓存建立：在了解业务的基础上建立了业务缓存，进一步降低了实际查询量，节约成本。

5. 项目收益

- 全托管一栈式数据库服务库服务

TiDB Serverless 提供全托管的服务模式，充分发挥了 TiDB 数据库的原有特点和优势，简化了 Amztrends 的整体架构，提升了性能和系统健壮性，同时降低了总体成本。

- 数据架构成本节省 50%

通过云原生改造，Amztrends 在数据架构方面节省了 50% 的成本。（Serverless 采用全分离架构，Amztrends 只需为使用付费，无需浪费闲置资源）

- 整体服务资源节约 30%

简化架构后，Amztrends 前端服务资源实现了 30% 的资源节约。

- 全局成本降低 80%

综合上述优化，Amztrends 实现了整体成本降低 80%。

6. 总结

TiDB Serverless 作为全托管一站式的数据库服务，从根本上解决了 Startup 公司基础架构门槛高的难题，即使缺少专业运维人员，也可以拥有简化、健壮的数据基础架构，获得了更高的系统稳定性和可用性。同时，Serverless 的 “Pay as you go”模式提供了更灵活、更优化的资源利用率，AmzTrends 从百度云迁移至 AWS 之后，总体 IT 成本实现了大幅下降。在此基础上，这种灵活的资源使用模式也可以让 AmzTrends 探索更灵活、更具性价比的商业方案，为商业模式提供更多可能性。

AmzTrends PM Steve 表示：“迁移到 TiDB Serverless 后我们不会再去担心数据库的问题，完全放心交给 TiDB Serverless 就好了。用最简单的话来讲，没有问题就是最好的结果。”

7. 未来

未来，Amztrend 计划继续利用 TiDB Serverless 的优势，进一步扩展其业务并探索更多的数据分析和优化机会。公司还计划通过持续的技术改进和优化，维持并提升其在市场上的竞争力。

声明：本文转载于 https://mp.weixin.qq.com/s/Xnb5rDNxyP1dIKglzECYVw

评论

T

评论

暂无评论

互助与交流

活动

问答论坛

TiKV 社区

Chaos Mesh 社区

学习与应用

文档

专栏

视频课程

考试认证

典型案例

开发者指南

发现社区

TiDB User Group

问答之星

社区准则

联系我们

电子书

