



DAMS

中国数据智能管理峰会

DATA & AI MANAGEMENT SUMMIT

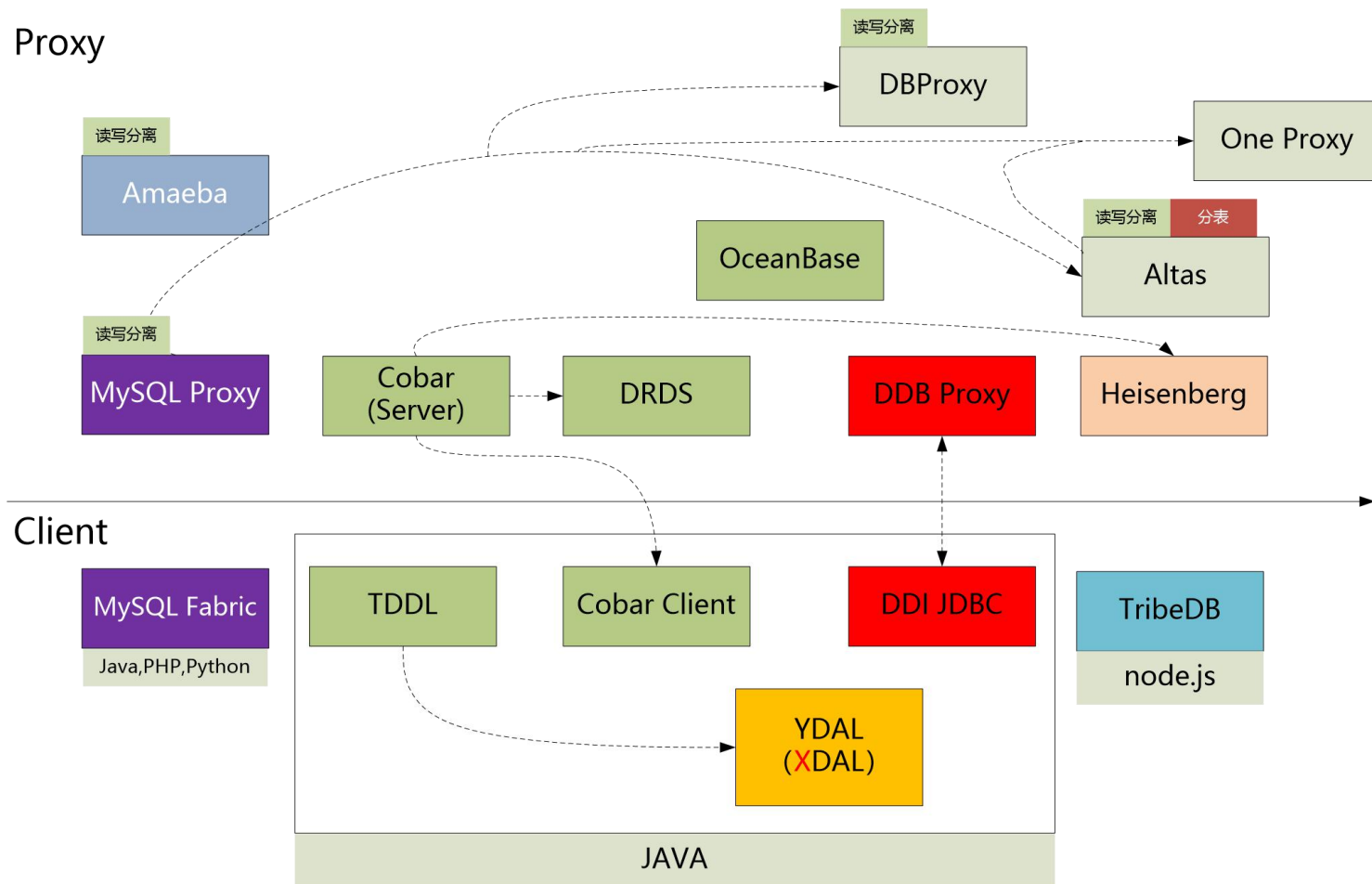
从银联自研看分布式数据库

演讲人：周家晶

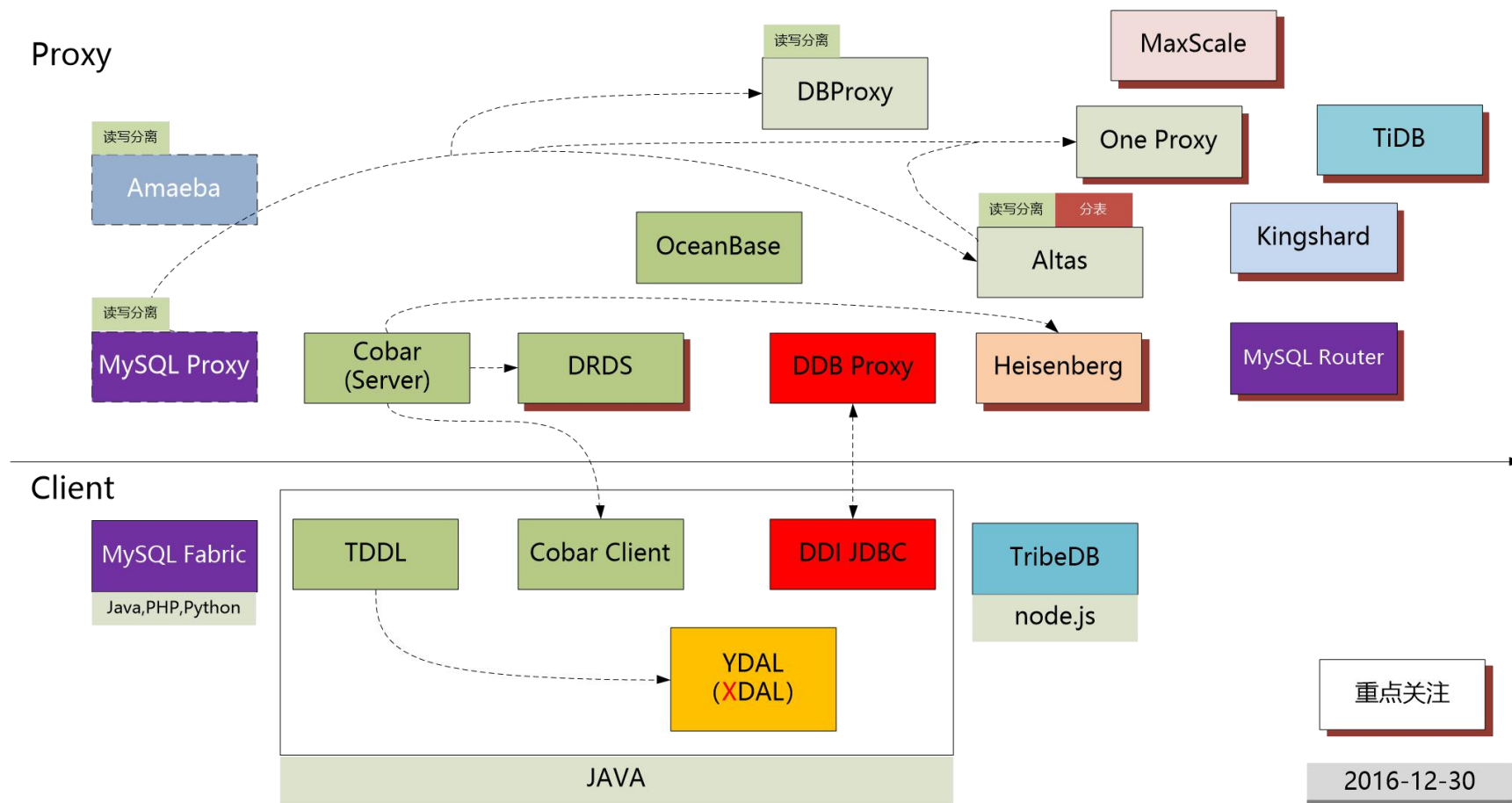


1. 业界分布式数据库产品变化

2015年分布式数据库相关产品



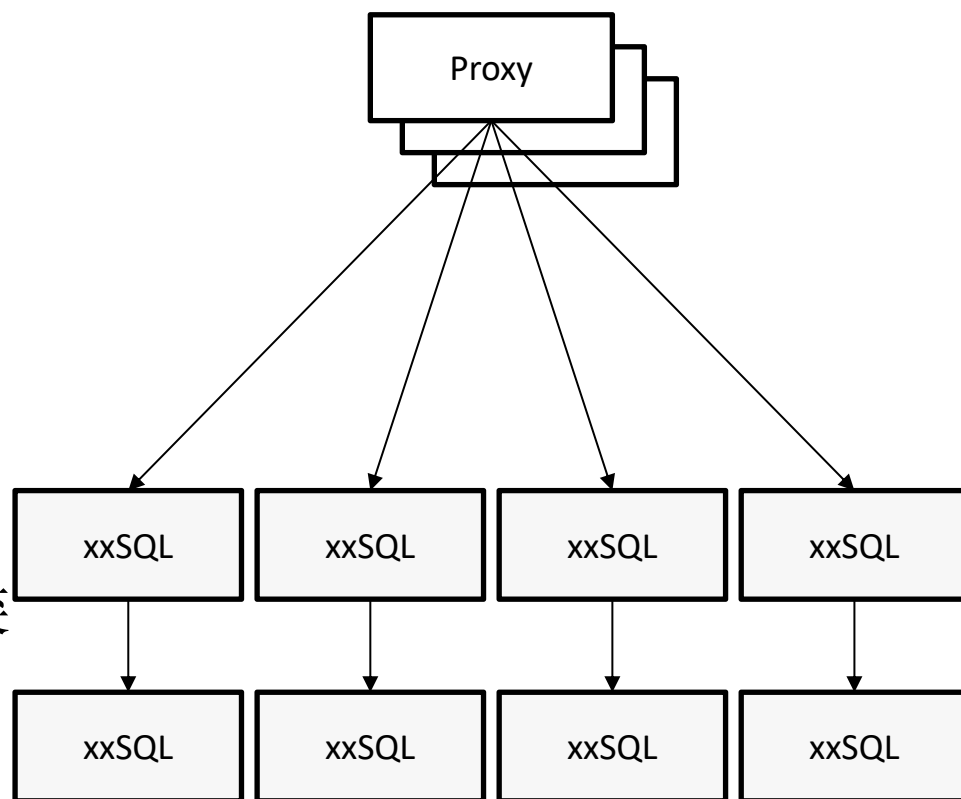
2016年分布式数据库相关产品



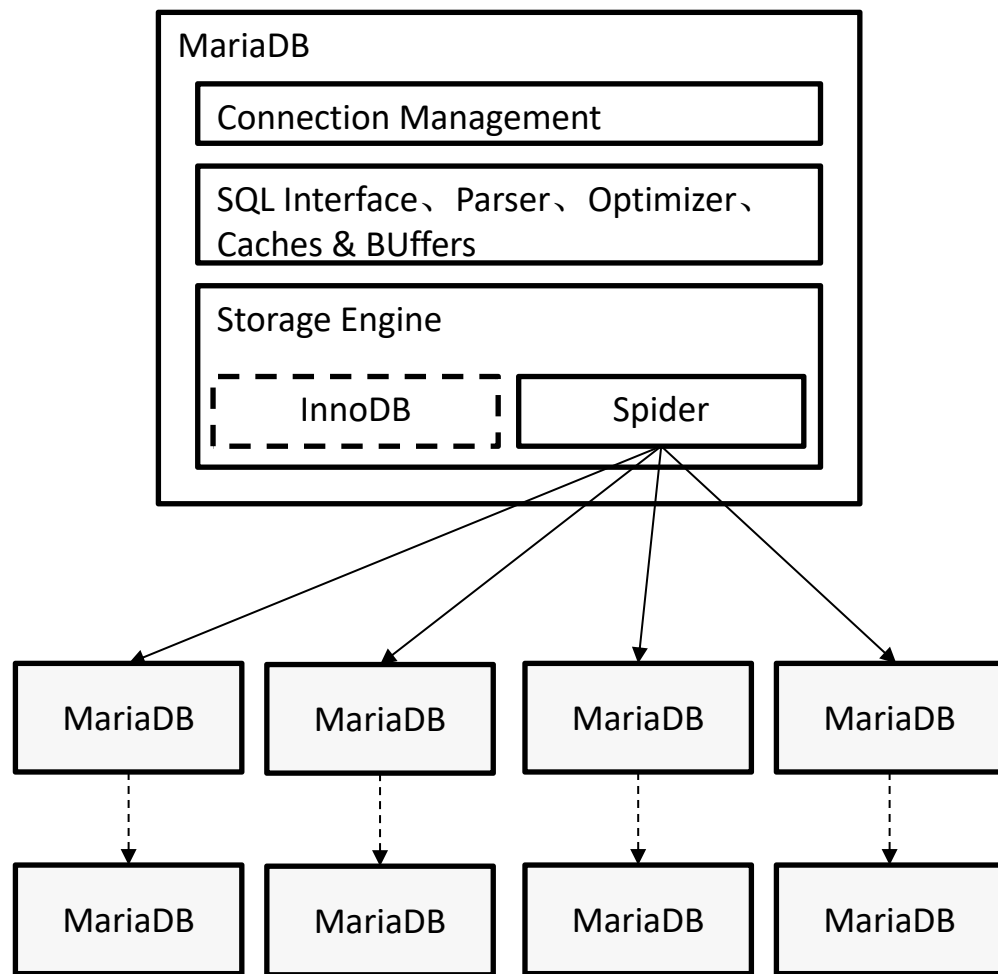
现在？

- 腾讯：
 - TDSQL (DCDB)
 - 阿里：
 - DRDS
 - PetaData (HybridDB for MySQL)
 - MariaDB:
 - Spider
 - PingCAP:
 - TiDB
 - 巨杉
 - SequoiaDB
 - 亚马逊、阿里
 - Aurora、PolarDB
- 1.数据存储：MySQL
 - 2.高可用：MySQL复制
- 1.数据存储：KV
 - 2.高可用：Multi-Raft
- 1.数据存储：KV
 - 2.高可用：类似MySQL复制
 - 3.分布式存储引擎
- 1.单点写、多点读
 - 2.存储与计算分离

- 技术路线
 - 丰富完善Proxy层的SQL解析能力
- 主要技术状态
 - 分布式事务支持
 - TDSQL支持
 - DRDS支持(多种分布式实现模式)
 - 复杂语句
 - 聚合类有限支持
 - 不支持部分数据类型



- 技术路线
 - 实现分布式的存储引擎
- 主要技术状态
 - 支持分布式事务
 - 复杂语句
 - 支持复杂语句
 - 并增加特性
 - DDL支持
 - 不支持单点DDL
 - 性能：
 - 与数据节点之间为同步调用
 - 集群性能低于单机MariaDB

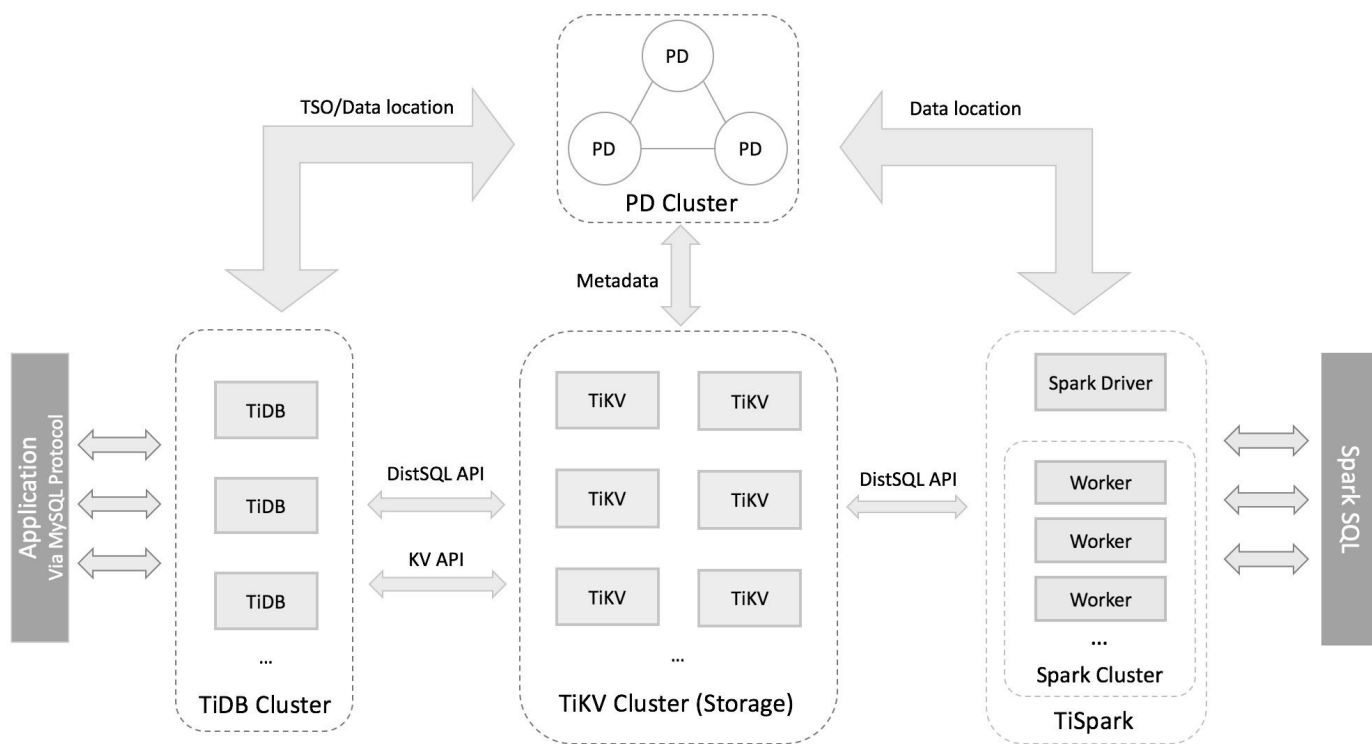


• 技术路线

- 借鉴Google Spanner和F1
- 使用KV进行数据存储
- 使用Multi-Raft实现高可用与数据迁移

• 主要技术状态

- 支持分布式事务
- 乐观锁机制：不适合热点数据
- 较好的MySQL兼容性
- 时延较高
- 成本较高

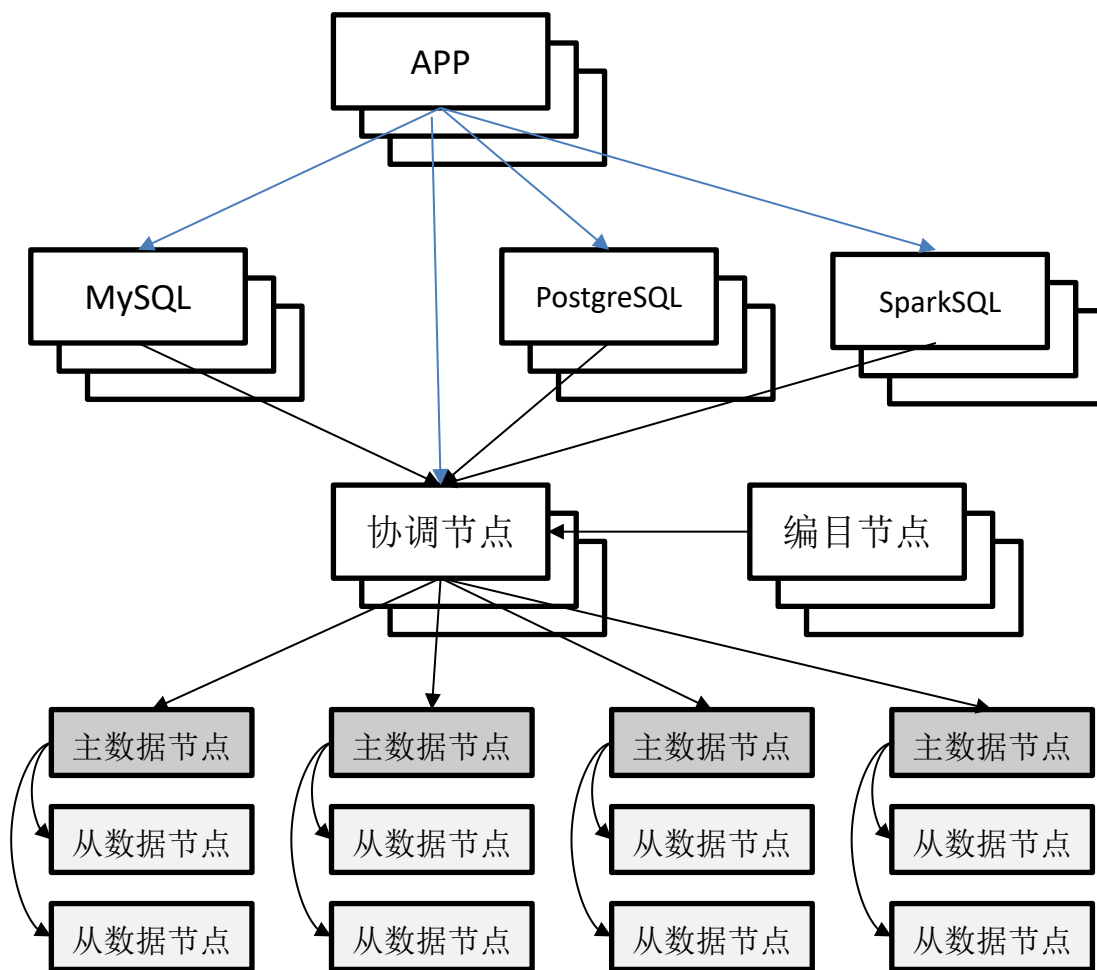


- 技术路线

- 使用KV进行数据存储
- 实现分布式存储引擎

- 主要技术状态

- 支持分布式事务
- 支持复杂语句
- 悲观锁，不适合高并发场景
- 较好的MySQL语法兼容
- 支持多种xSQL



- 技术路线

- 通过Proxy层进行复杂语句路由

- OLTP与OLAP混合

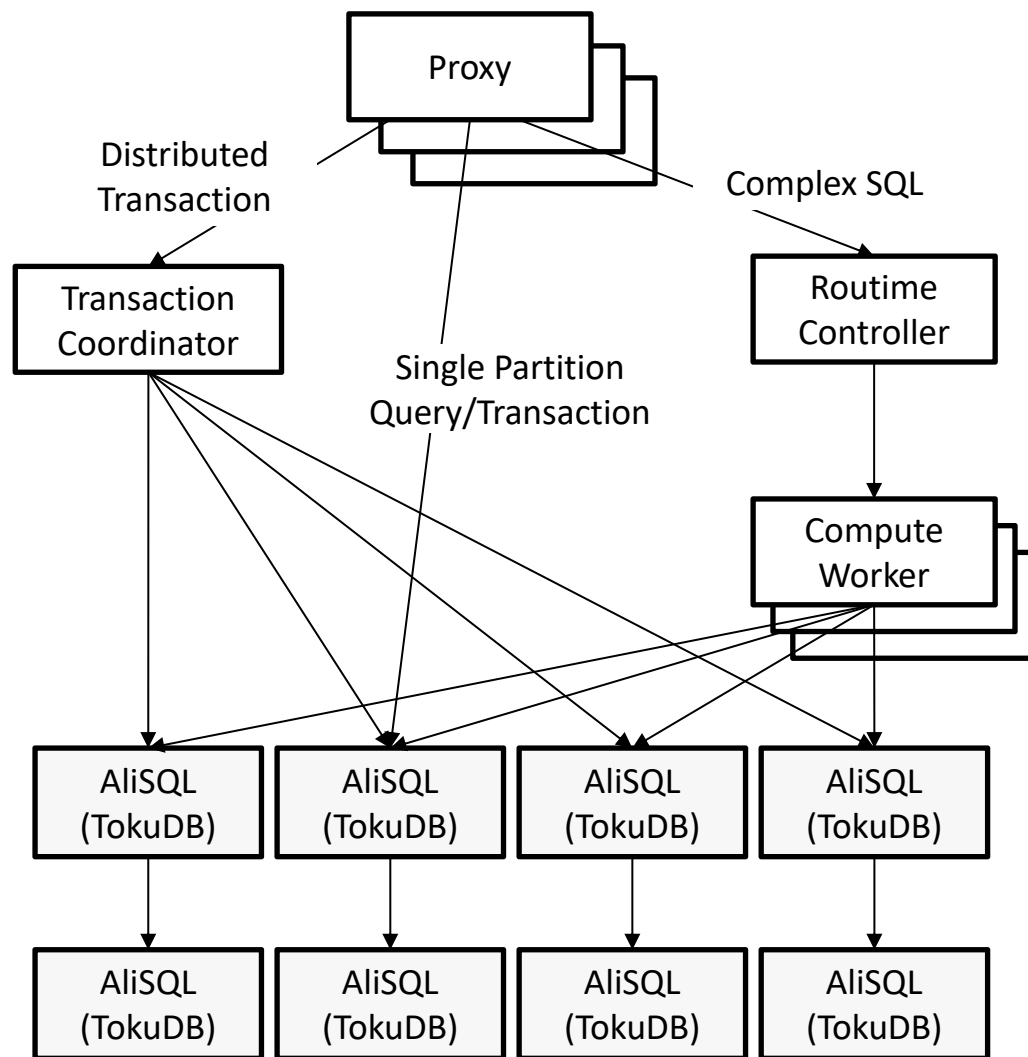
- 主要技术状态

- 分布式事务

- 简单语句支持分布式事务
 - 复杂语句目前不支持事务

- 复杂语句

- 支持MySQL单机语句
 - 增加了OLAP特性语法支持
 - TPC-H
 - TPC-DS

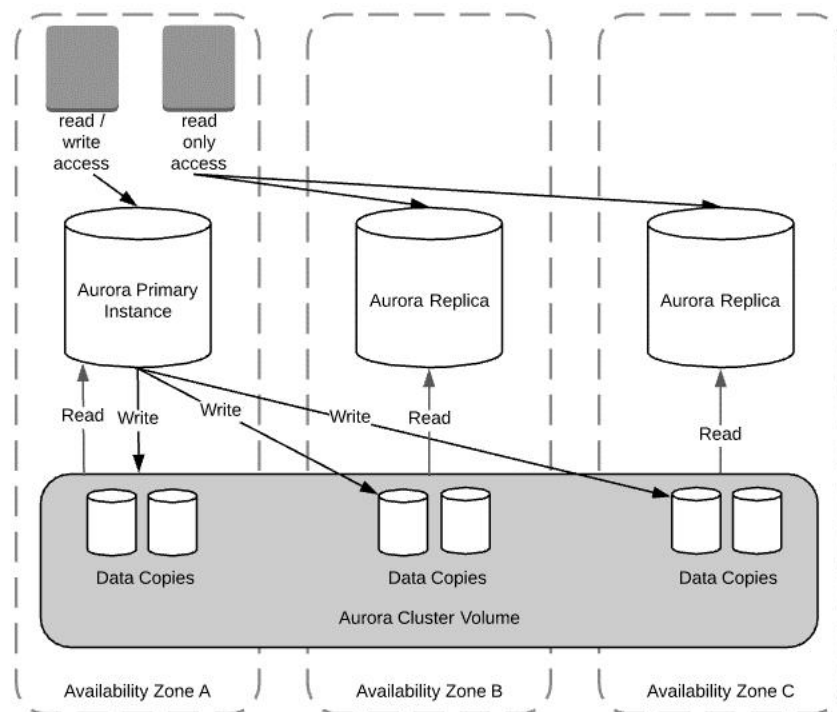


- 技术路线

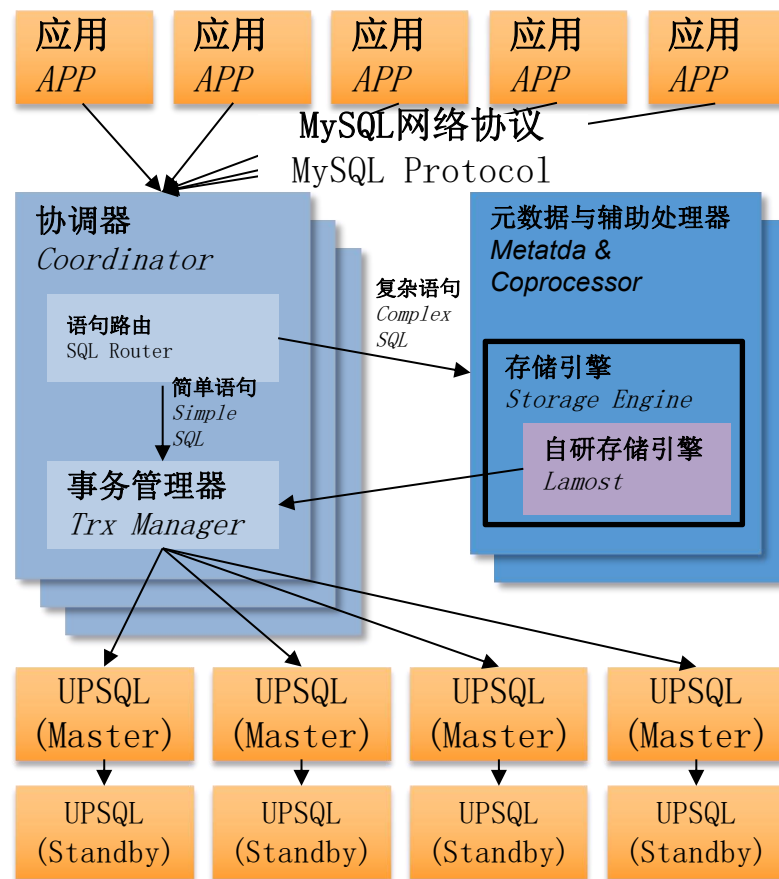
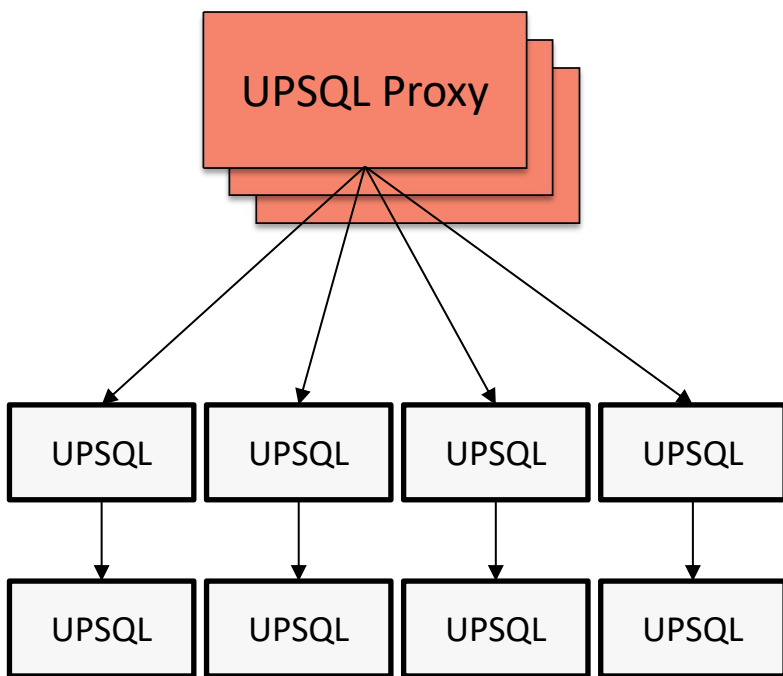
- 单点写、多点读
 - Aurora已支持多点写
- 存储计算分离
- 分布式文件系统
- 用户态读写网络、文件

- 主要技术状态

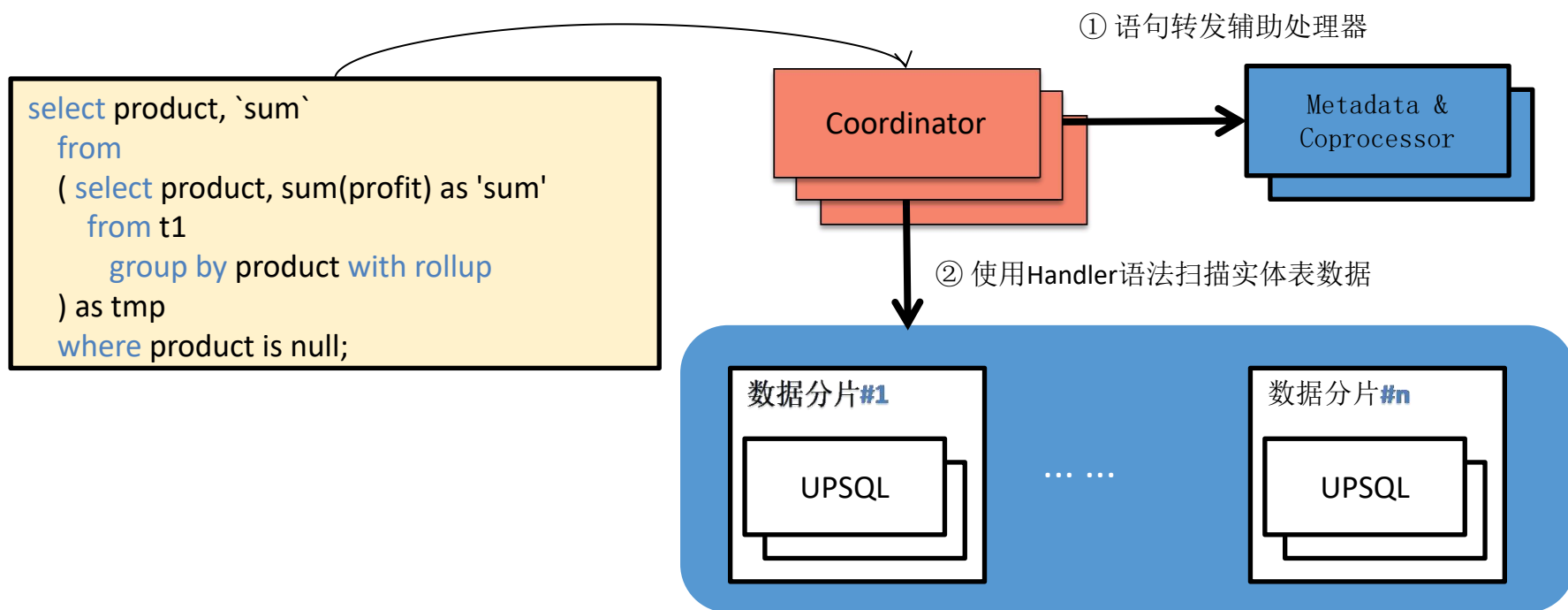
- 读性能线性扩展
- 写性能可以达到MySQL单机性能的6倍以上



2. 银联架构演进



- UPDRDB能支持跨分片join、子查询、视图、UDF等复杂语句，其一般处理流程为：
 - 协调器(Coordinator)将复杂语句转发给辅助处理器(Coprocessor)
 - 如果操作的表为Lamost引擎，Lamost引擎会将相应操作转换为Handler语句，从实体表获取数据
 - 辅助处理器根据获取到的数据进行语句处理，并将最终结果返回协调器

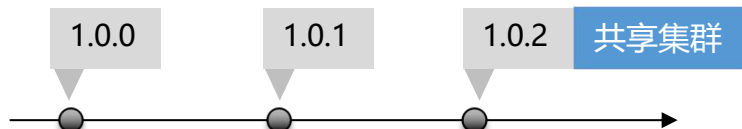


UPSQL Proxy 1

2015-08-31



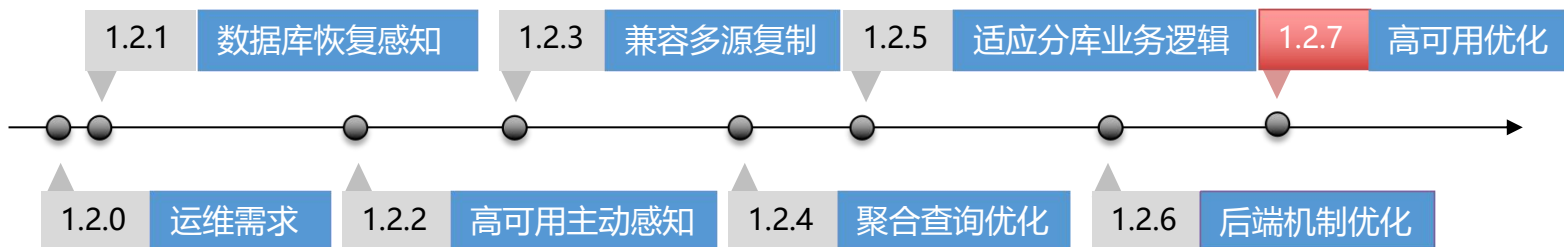
2015-11-18



2016-05-20



2016-06-27



UPSQL Proxy 2

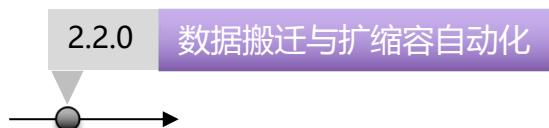
2017-07-07



2017-11-24



2018-08-28



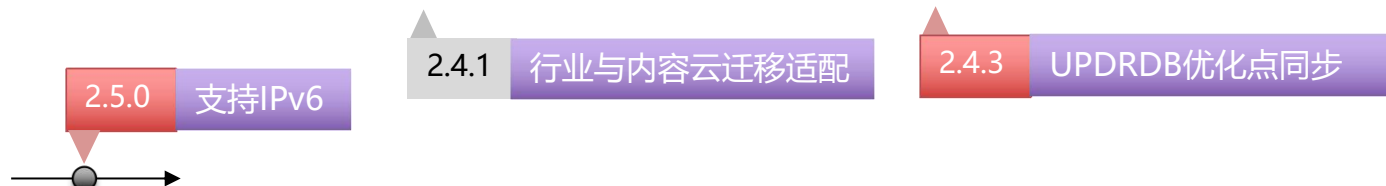
2018-09-28



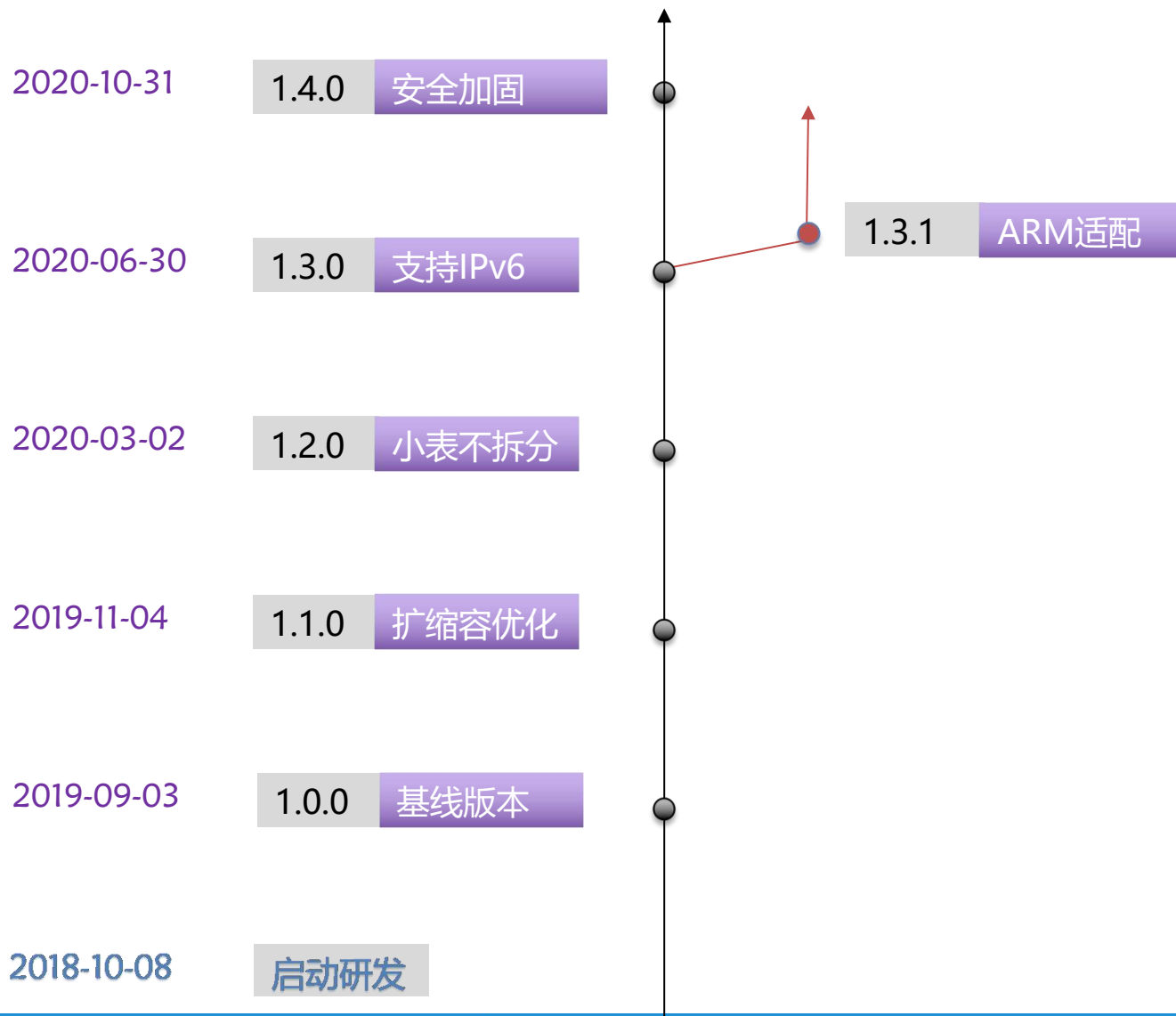
2018-12-19



2020-06-30



UPDRDB



3. 产品迭代的不同阶段

- 2017年之前，我们主要解决数据库高可用、简单的数据拆分
- 2017年开始，以分布式事务、异步连接池为重点主要提供分布式功能与性能
- 2018年底开始，开始论证技术路线，考虑更为统一的复杂语句、扩缩容等问题的解决方案

4. 分布式数据库取决于：架构 or 功能？

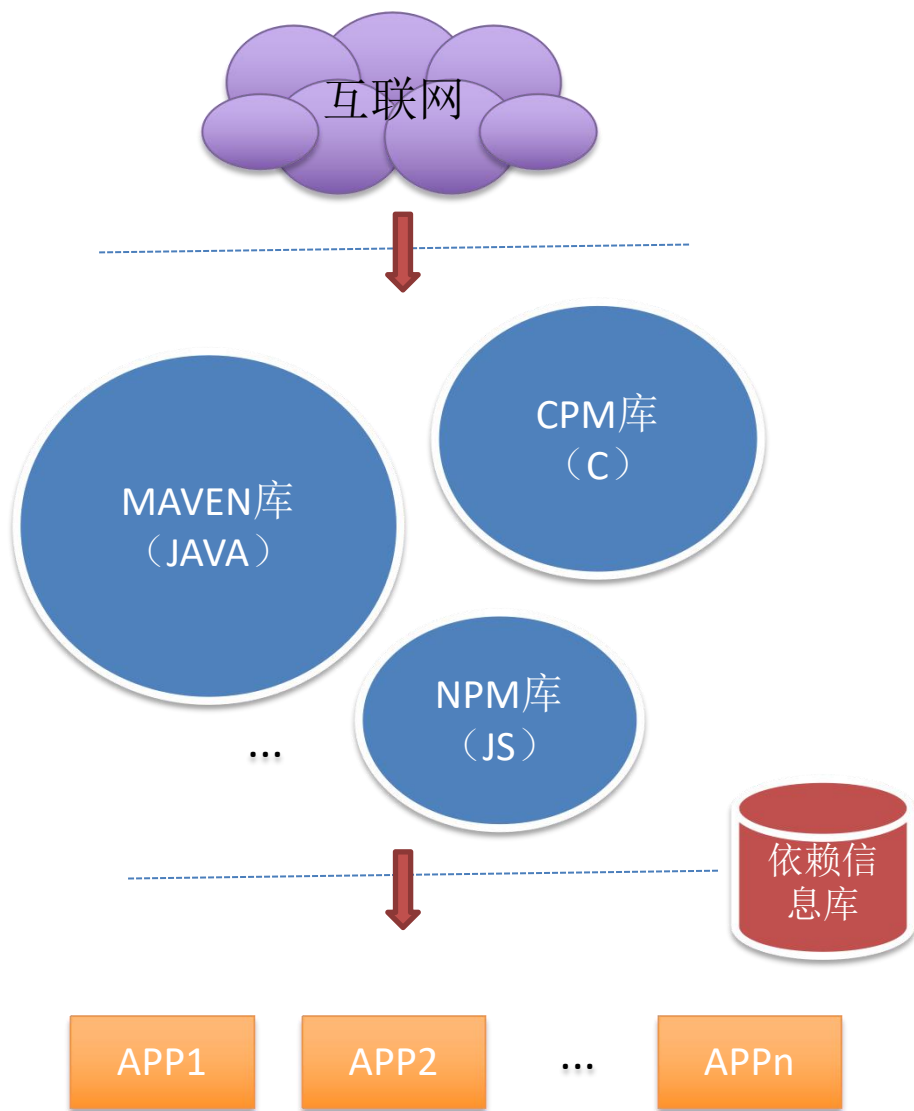
技术路线	含义	特征	产品	混合方案	产品
新架构	基于新架构创建的新型分布式数据库。	设计之初就考虑了分布式特征。具备较高的硬件要求，且存在稳定性担忧。	OceanBase TiDB		
存储引擎	在传统数据库之后，研发分布式存储引擎，以具备分布式数据库关键特性。	与传统数据库具备极好的兼容性，但性能损耗较大，部分产品存在明显提升上限。	SequoiaDB-MySQL(巨杉) MariaDB Spider Aurora (by Amazon) PolarDB (by Alibaba) CynosDB (by Tencent)	混合存储引擎和中间件方案，在兼容性、稳定性、扩展性上进一步发展。	PolarDB-X UPDRDB
透明中间件	在传统数据库之上构建中间件。	具备高性能、低成本和稳定性特征。但与传统数据库的兼容性最差。	Vitess(from Youtube) Sharding-Sphere UPSQL Proxy TDSQL GoldenDB		

1. Andrew Pavlo and Matthew Aslett. What's really new with newsql? 45(2):45-55, 2016.

2. Newsql—the new way to handle big data. <https://opensourceforu.com/2012/01/newsqldb-handle-big-data/>.

5. 开源 or 闭源





- 网络入口：严格受控
- 组件引入：专家评审
- 组件维护：专人负责
- 应用依赖：自动化提交



自研分布式数据库产品的出路

- 消亡
- 融入
- 引领

Q&A





DAMIS

中国数据智能管理峰会

DATA & AI MANAGEMENT SUMMIT

THANK YOU !

