# 概述

TiDB是PingCAP公司自主设计、研发的开源分布式关系型数据库，是一款同时支持在线事务处理与在线分析处理（Hybrid Transactional and Analytical Processing, HTAP）的融合型分布式数据库产品，具备水平扩容或者缩容、金融级高可用、实时HTAP、云原生的分布式数据库、兼容MySQL 5.7协议和MySQL生态等重要特性。目标是为用户提供一站式OLTP（Online Transactional Processing）、OLAP（Online Analytical Processing）、HTAP解决方案。TiDB适合高可用、强一致要求较高、数据规模较大等各种应用场景。

## 核心特性

**一键水平扩容或者缩容**

得益于TiDB存储计算分离的架构的设计，可按需对计算、存储分别进行在线扩容或者缩容，扩容或者缩容过程中对应用运维人员透明。

**金融级高可用**

数据采用多副本存储，数据副本通过Multi-Raft协议同步事务日志，多数派写入成功事务才能提交，确保数据强一致性且少数副本发生故障时不影响数据的可用性。可按需配置副本地理位置、副本数量等策略满足不同容灾级别的要求。

**实时HTAP**

提供行存储引擎TiKV、列存储引擎TiFlash两款存储引擎，TiFlash通过Multi-Raft Learner协议实时从TiKV复制数据，确保行存储引擎TiKV和列存储引擎TiFlash之间的数据强一致。TiKV、TiFlash可按需部署在不同的机器，解决HTAP资源隔离的问题。

**云原生的分布式数据库**

专为云而设计的分布式数据库，通过TiDB Operator可在公有云、私有云、混合云中实现部署工具化、自动化。

**兼容 MySQL 5.7协议和MySQL生态**

兼容MySQL 5.7协议、MySQL常用的功能、MySQL生态，应用无需或者修改少量代码即可从MySQL迁移到TiDB。提供丰富的数据迁移工具帮助应用便捷完成数据迁移。

## 基本功能

## TiDB 6.1 Release Notes

## 与MySQL的兼容性

## 使用限制

# 架构

## 计算/SQL引擎

## 调度/路由

## 存储引擎

TiDB是一款分布式HTAP数据库，它目前有两种存储节点，分别是TiKV和TiFlash。TiKV采用了行式存储，更适合TP类型的业务；而TiFlash采用列式存储，擅长AP类型的业务。TiFlash通过raft协议从TiKV节点实时同步数据，拥有毫秒级别的延迟，以及非常优秀的数据分析性能。它支持实时同步TiKV的数据更新，以及支持在线DDL。我们把TiFlash作为Raft Learner融合进TiDB的raft体系，将两种节点整合在一个数据库集群中，上层统一通过TiDB节点查询，使得TiDB成为一款真正的HTAP数据库。

### TiKV

参考：<https://docs.pingcap.com/zh/tidb/stable/tikv-overview>

### TiFlash

参考：<https://docs.pingcap.com/zh/tidb/stable/tiflash-overview>

# 分布式事务

# 数据分布

# 复制/一致性

# 备份恢复

## 备份TiDB集群快照

## 使用BR恢复集群

# 容错/故障切换

# 兼容性

# 扩展性/扩容

## 使用TiUP扩容缩容TiDB集群

## 使用TiDB Operator

# 高并发

# 高可用

# 数据安全

## 安全加固

### 为TiDB客户端服务端间通信开启加密传输

### 为TiDB组件间通信开启加密传输

### 生成自签名证书

### 静态加密

### 为TiDB落盘文件开启加密

### 日志脱敏

## 权限

### 与MySQL安全特性差异

### 权限管理

### TiDB用户账户管理

### 基于角色的访问控制

### TiDB证书鉴权使用指南

# 数据压缩

# 数据迁移

## 概述

数据迁移方案如下：

全量数据迁移。

数据导入：使用TiDB Lightning将Aurora Snapshot，CSV文件或Mydumper SQL文件的数据全量导入到TiDB集群。

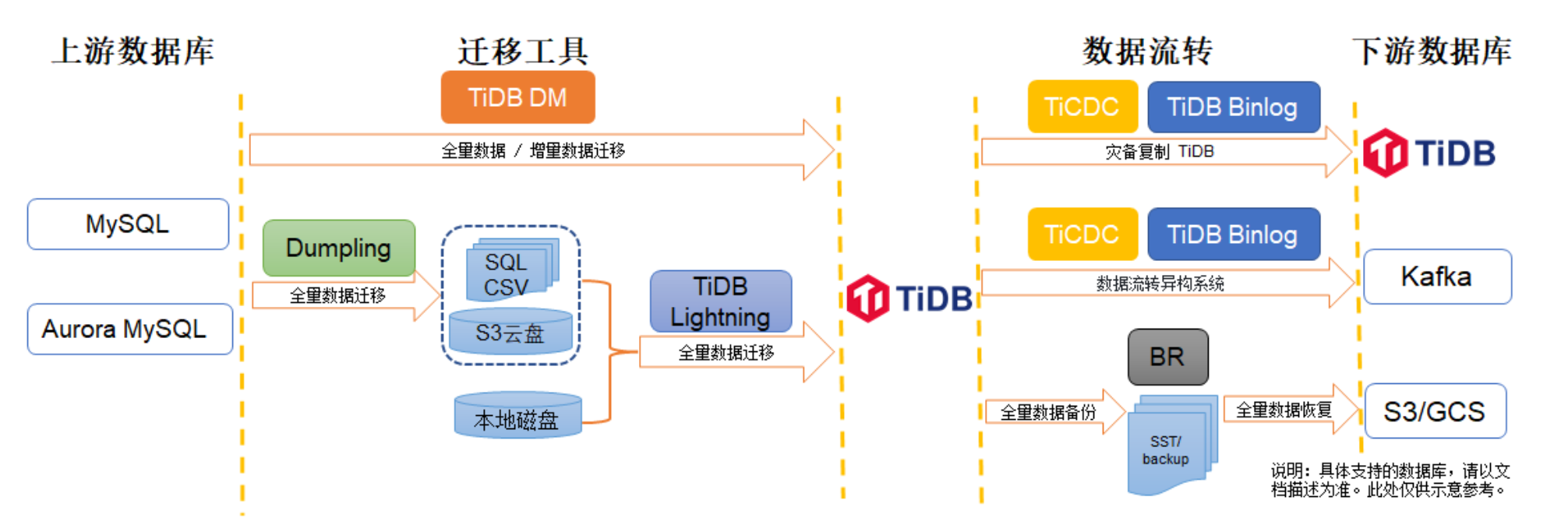
数据导出：使用Dumpling将TiDB集群的数据全量导出为CSV文件或Mydumper SQL文件，从而更好地配合从MySQL数据库或MariaDB数据库进行数据迁移。

TiDB DM（Data migration）也提供了适合小规模数据量数据库（例如小于1 TiB）的全量数据迁移功能。

## 迁移工具

TiDB提供了丰富的数据迁移相关的工具，用于全量迁移、增量迁移、备份恢复、数据同步等多种场景。

下图显示了各迁移工具的使用场景：



### TiDB DM

### Dumpling

### TiDB Lightning

### Backup & Restore (BR)

### TiCDC

### TiDB Binlog

### sync-diff-inspector

# 数据集成

数据集成一般是指数据在各个独立的数据源之间流动、转换和汇集。随着数据量的爆炸式增长和数据价值被深度挖掘，对数据集成的需求越来越普遍和迫切。为了避免TiDB成为数据孤岛，顺利与各个数据系统进行集成，TiCDC提供将TiDB增量数据变更日志实时同步到其他数据系统的能力。

## 与Confluent Cloud进行数据集成

你可以使用TiCDC将TiDB的增量数据同步到Confluent Cloud，并借助Confluent Cloud的能力最终将数据分别同步到ksqlDB、Snowflake、SQL Server。

## 与Apache Kafka和Apache Flink进行数据集成

可以使用TiCDC将TiDB的增量数据同步到Apache Kafka，并使用Apache Flink消费Kafka中的数据。

# 性能调优

## 系统调优

## 软件调优

### 配置

#### TiDB内存调优

#### TiKV线程调优

#### TiKV内存调优

#### TiKV Follower Reader

#### TiFlash调优

### 下推计算结果缓存

## SQL性能调优

# 运维/监控告警

## 监控指标

### Overview面板

### TiDB面板

### PD面板

### TiKV面板

### TiFlash监控指标

### TiCDC监控指标

# 故障诊断