TIDB 和 OB 区别

TIDB：存储和计算分离

OB：单机和分布式一体化

分布式系统

是一种其组件位于不同的联网计算机上的系统，然后通过互相传递消息来进行通信和协调。

为了达到共同的目标，这些组件会相互作用

分布式数据库系统

分布式数据库是用计算机网络将物理上分散的多个数据库单元连接起来组成的一个逻辑上统一的数据库。每个被连接起来的数据库单元称为站点或节点。分布式数据库有一个统一的数据库管理系统来进行管理，称为分布式数据库管理系统

google三驾马车

GFS(google file system):解决分布式文件系统的问题

Google Bigtalbe：解决数据存储问题

Google MapReduce: 大量数据读取的一个算法

分布式系统里著名的CAP理论 CA AP CP

一致性 Consistency：线性一致性，假设在T时刻写入了一个值，那么在T时候之后读取一定能读到这个最新的值

ACID的一致性指一致性的状态，指数据变化之后的前后一致性，比如：AB转账

A原来是100，B也是100，A向B转账50，成功后，此时A为50，B为150,他们两个账户的和与转账之前的和是一样的

可用性 Availability:指的服务高可用，比如主从，当主节点故障之后，从节点可以立刻替换为主节点继续工作而不影响线上业务

分区容忍性 Partition Tolerance: 不同数据存在不同的节点，当数据节点出现故障，还可以提供正常的服务，比如脑裂、网络隔离等导致的节点故障

TIDB 是CP+HA系统，多数时候也是一个CA系统，如果出现了P，也就是刚好对外服务的leader被隔离了，新leader大概需要10S+以上的 时间才能被 选举出来对外提供服务。

分布式数据库价值：

高可用

易扩展

高性能

newSql：原生分布式关系型数据库

支持分布式、支持sql、支持事务

是一类关系数据库，它寻求为在线交易处理（OLTP）工作提供NoSQL的系统的可扩展性，同时维护传统数据库系统的ACID保证

未来：

云原生部署，数据库云，一栈式，实时HTAP

什么是事务

它的本质就昌指并发控制的单元，是用户定义的一个操作序列。这些操作要么都做，要么都不做，是一个不可分割的工作单位

你用哪些数据库的中间件？

有什么分库分表方案？