

# 结束语 | 分布式数据库，未来可期

你好，我是高洪涛。课程到这里就要和你说再见了。

不知不觉三个月转瞬即逝，感谢你一直以来的陪伴。人总说“相识是缘分，相守是情投”，能坚持学完全部 24 讲内容和两节加餐，证明我们彼此之间情谊是相合的，而情谊的纽带我想就是对分布式数据库技术的欣赏与热爱吧。

我大概在 2015 年开始接触分布式数据库这个领域，刚开始就见识了数据库中间件技术的大繁荣，并有幸成为国内最早制作分布式中间件的专业人员。而后经过几年的积累，我目前正在设计和实现一款专业领域的 NoSQL 数据库。这一路走来，我愈发觉得当下热爱技术的同学对基础软件的认知水平远远不够，所以当拉勾教育找到我并希望开设这门课时，我是非常开心的，因为能有一次机会与你分享我对分布式数据库的理解。

最初设计课程的时候，我非常担心这类核心系统听众不多。因为中国虽然是目前世界上 IT 用户最多的国家，中文也是 IT 最大的受众群体。但我们的开发者中，绝大多数都是面向应用开发的，如果单纯从实用主义的角度看，核心系统，特别是数据库相关主题能否得到广大读者的接受，这一点我内心其实没有底。但是，课程上线以来，不论是订阅量还是你们的反馈，都让我越来越有动力把专栏好好写下去。

这门专栏是关于分布式数据库底层逻辑的，关注这个领域的专业人员逐步增多，这使我对国产数据库，乃至国产核心系统的发展抱有了极大的憧憬。事实也正逐步验证我的这种看法，特别是 TiDB 与 Oceanbase 开始在电信、银行等国家重点行业中商用，这预示着国产数据库，特别是 NewSQL 类国产数据库将承担越来越重要的任务。

虽然，NewSQL 数据库目前应用的案例还很有限，但它未来的想象空间是巨大的。而最关键的因素是，越来越多的组织和企业开始参与其中。我相信，只要人类社会持续将资源放置在一个领域内，最终该领域都会取得重大突破。从移动互联网到电动汽车，这些案例都侧面证明了这个说法。那么，现在都有哪些 NewSQL 数据库值得我们关注呢？

首先是具有创新架构的 NewSQL 类数据库。

这里以 Spanner 和 F1 为主，其中 F1 主要作为 SQL 引擎，而事务一致性、复制机制、可扩展存储等特性都是由 Spanner 完成的，所以我们有时会忽略 F1，而更多地提到 Spanner。

Spanner 和 F1 在 2017 年已经发生了新的变化，Spanner 的论文是“Spanner: Becoming a SQL System”。就像论文名字所说的，Spanner 完善了 SQL 功能，这样就算不借助 F1，它也能成为一个完整的数据库。这篇论文用大量篇幅介绍了 SQL 处理机制，同时在系统定位上相比 2012 版有一个大的变化，强调了兼容 OLTP 和 OLAP，也就是 HTAP。

对应的，Spanner 在存储层的设计也从 2012 版中的 CFS 切换到了 Ressi。Ressi 是类似 PAX 的行列混合数据布局（Data Layout）。F1 的新论文是“F1 Query: Declarative Querying at Scale”。这一版论文中 F1 不再强调和 Spanner 的绑定关系，而是支持更多的底层存储。非常有意思的是，F1 也声称自己可以兼顾 OLTP 和 OLAP。

由于 Spanner 的名声在外，一众借鉴 Spanner 概念的数据库需要被大家所了解。如 TiDB、YugabyteDB 等。但是每家数据库都有自己的特色，它们在时钟、分片和事务上都各不相同。

其次，除了 Spanner，OceanBase 目前在行业内部也得到了越来越多的认可，它具有优异的性能、完善的功能；同时最重要的是它由支付宝这个重量级的商业案例背书，所以得到了金融、银行客户的青睐。

此外，以 AWS 的 Aurora 和阿里云的 PolarDB 为代表的云原生分布式数据库采用了另一种策略。它采用了共享存储模式，放弃了一定的扩展能力，但是却换来了对传统式数据库优良的支撑能力。但是这种策略目前看仅仅能在云厂商落地，因为其技术方案比较特别，很难在一般的 IDC 中进行复制。

除了创新架构的 NewSQL 外，另一类分布式数据库可能你在之后的学习和工作中更容易接触到。那就是从数据库中间件发展而来的 NewSQL 数据库，如以 MySQL 为基础的 Vitess。但是对于咱们国内用户而言，更常见的 PGXC 类型的数据库，比如 TDSQL、

GoldenDB，等等。它们由于具有传统数据库的用户接口，目前越来越多的组织开始将它们落地。至于数据库中间件，目前单纯的中间件都已经到了发展的瓶颈，并逐步被淘汰，如果不引入更多偏向于 NewSQL 的架构，这些中间件终将消失在历史之中。

好了，我们回顾了整个分布式数据库目前的状态，那么未来在哪里？其实眼下我们就站在了历史的路口之上。

首先，分布式数据库的产品进入了高速发展时期，越来越多的组织正在或计划采用新一代架构的分布式数据库。特别是在 2021 年的 4 月份，Oracle 在 DB-Engine 的排名大幅下滑，而相对的 Postgre 和一种云原生数据库排名快速上升，这也预示着曾经不可撼动的商业帝国正在逐步褪色，新的时代在向我们招手。

其次，在中美竞争加剧的背景下，国产基础核心系统软件和国产数据库将会得到更多的机会，人们将会在电信、银行、能源等众多领域逐步完成国产替代。而新一代的国产数据库无非是分布式数据库。未来，从业者们也将有越来越多的机会直接使用它们。除了使用，国内的开发者也会有更多的机会参与数据库的研发过程，从而完成我们整个技术人才层次的升级。

基于对现状的认知和对未来的寄望，我把这门专栏定位成你在历史的十字路口中的向导，希望可以为你指明方向。

感谢你的支持和陪伴，山高路远，我们后会有期！

Hello 小伙伴，课程已经全部结束了，不知道这 24 讲和 2 节加餐对你的帮助如何？小编真诚邀请你根据自己的实际情况填写这份调查问卷，你的反馈就是我们进步的方向。同时也希望你未来能够乘风破浪、披荆斩棘，成为更好的自己！