**有哪些分布式数据库书籍或论文比较好？**

首先要搞清楚，目前实践中的数据库系统，OLTP (偏事务） 和 OLAP(偏分析）的数据库架构其实差别很大。现代分布式数据库当然都会支持，但是性能优化取向的不同会导致存储，优化和执行的巨大区别。

我举几个相关领域最先进的而且已经进入production的分布式数据库系统的例子，论文自己搜吧：

OLTP: Spanner, VoltDB, Calvin, Aurora 这四个个系统是毫无疑问的State-of-Art. 需要认识到的是分布式事务有着天然的局限性，如果Single-Node能解决的事情完全没必要用分布式。还有一点是事务是ad-hoc 查询还是stored procedure也能造成系统设计的不同。

OLAP: 在大家都迈向Cloud Native的时代，State-of-Art其实只有三个系统， Dremel/BigQuery (Google), Snowflake (Snowflake) 和 Redshift (Amazon). 仔细比较下他们的系统架构，你会发现还是有很多共通的地方的，仔细想想这些系统为什么这样设计，跟传统的数据库的设计有什么不一样，你会对数据库系统架构的理解更深一层。

至于 Sharding, PAXOS这些具体的技术，当然也是需要深入学习的。

分布式数据库系统的书市面上只有一本有名的（来自Waterloo的Tamer大神），不幸的是这本书写的很烂而且也比较脱离时代。

PS: 别问我为什么没列HBase这种因为Hadoop兴起一时走红，迟早会被扔进历史的垃圾堆的东西。

<https://www.zhihu.com/question/37647788>