

互联网时代分布式关系数据库的有益探索

周傲英

(华东师范大学 上海 200241)

在实现分布式数据库的技术方案上,业界存在不同的选择。第一种方式需要对应用系统进行拆分,通过分库分表将原本单个数据库管理的数据分散到多个集中式数据库。分库分表方案要求应用系统重构,跨库访问效率较低,关系数据库的重要功能,如外键、全局唯一性约束、全局索引等无法使用。第二种方式是对传统集中式关系数据库进行分布式改造,增加分布式事务处理,小规模集群部署下的自动故障恢复等功能。这类分布式数据库由于存储系统、事务处理和SQL优化器等源自集中式架构,在分布式场景下面临功能和性能上的诸多限制。第三种方式是从头开始设计和实现一个原生分布式关系数据库,将分布式作为基本特性融入存储系统、事务处理和SQL优化器等关键组件。相比前两种方案,原生分布式数据库在高可用、数据一致性、事务性能、弹性伸缩、快速无损的故障恢复等方面有着更大的优势。

OceanBase 是一个从头开始设计与实现的分布式关系数据库系统。**OceanBase** 因淘宝而诞生,因支付宝而发展和壮大,如今已在金融、政务、通信和互联网等领域得到广泛应用。由 **OceanBase** 首席科学家阳振坤领衔的分布式数据库研发团队实现了多项技术创新和突破,该团队撰写的论文“**OceanBase** 分布式关系数据库架构与技术”介绍了 **OceanBase** 的分布式架构,分布式事务处理、存储引擎、SQL 优化、多租户机制等关键技术,具体总结如下:

- 1) 设计了强一致、高可用、可扩展的分布式事务处理机制,实现了单机/单机房故障的自动、无损、快速的故障恢复;
- 2) 提出了单机/分布式一体化关系数据库架构,实现了关系数据库容量和处理能力从单机数据库到分布式数据库的无缝切换和伸缩;
- 3) 实现了关系数据库的性能无损的高倍率数据压缩,论文实验展示了数据压缩倍率是主流关系数据库的 3 倍甚至更高;
- 4) 实现了单数据库系统同时支持高性能事务处理和实时分析处理,典型场景的事务处理性能和分析处理性能都高于 MySQL。

OceanBase 是迄今为止唯一同时获得了 TPC-C 和 TPC-H 性能榜首的数据库。尽管关系数据库的提出已经过去了半个世纪之久,真正意义上的分布式关系数据库时代才刚刚开始,论文不仅展示了 **OceanBase** 采用的分布式数据库关键技术,也对未来分布式数据库的发展方向提出了展望。我相信,这篇论文能引发很多关于数据库发展方向的思考,对于从事相关研究和开发的工程技术人员和数据库应用领域的专业人士都有重要的参考价值。

评述专家



周傲英,教授,博士生导师。主要研究方向为 Web 数据管理、数据密集型计算、内存集群计算、分布事务处理、大数据基准测试和性能优化。

亮点论文

阳振坤,杨传辉,韩富晟,王国平,杨志丰,成肖君.
OceanBase 分布式关系数据库架构与技术 [J]. 计算机研究与发展, 2024, 61(3): 540–554. DOI:10.7544/issn1000-1239.202330835