

分布式高性能图数据库 在知识图谱领域的应用

CCKS2022
CCIVIS2022

引言

知识图谱已成为当前人工智能和大数据分析处理的关键技术。为了应对图数据规模的快速增长和知识图谱领域分析算法的复杂性，向用户提供高效、实时的分析结果，解决已有系统无法很好支撑大规模图数据库的存储管理和高效分析计算的问题，海致自研新一代分布式高性能图数据库AtlasGraph。

方法

为了支撑大规模的图数据，AtlasGraph通过构建分布式图原生存储格式实现高效紧凑的数据存储和点边查询；通过多版本乐观并发控制技术实现分布式事务，保证数据库的一致性；通过链式复制协议实现数据的高可用；通过shared nothing 实现图查询的动态扩容；通过MPP架构实现图计算的高性能。

结果

AtlasGraph支持针对cypher图查询语言的高效图查询，可通过直观的查询语句来匹配图中点、边或路径，并实现毫秒级响应。

AtlasGraph还内置了高性能的图计算和图机器学习模块，支持社区划分、相似度计算、环路识别、图神经网络、图嵌入等知识图谱领域常用算法的高效计算。

应用

依托于分布式高性能图数据库提供的大规模图数据存储和分析能力，海致星图在金融方面适用场景包括**供应链关系识别、产业链图谱、量化潜客挖掘、集团派系挖掘、经济依存群体识别**等；零售金融方面适用场景包括**申请团伙反欺诈、客户亲密度分析、交易反欺诈、智能催收**等；风险管理方面适用场景包括**风险事件影响分析、风险传导、黑名单关联分析**等；审计合规方面适用场景包括**反洗钱团伙挖掘、疑似借户融资识别、异常担保识别、异常资金流向分析**等。

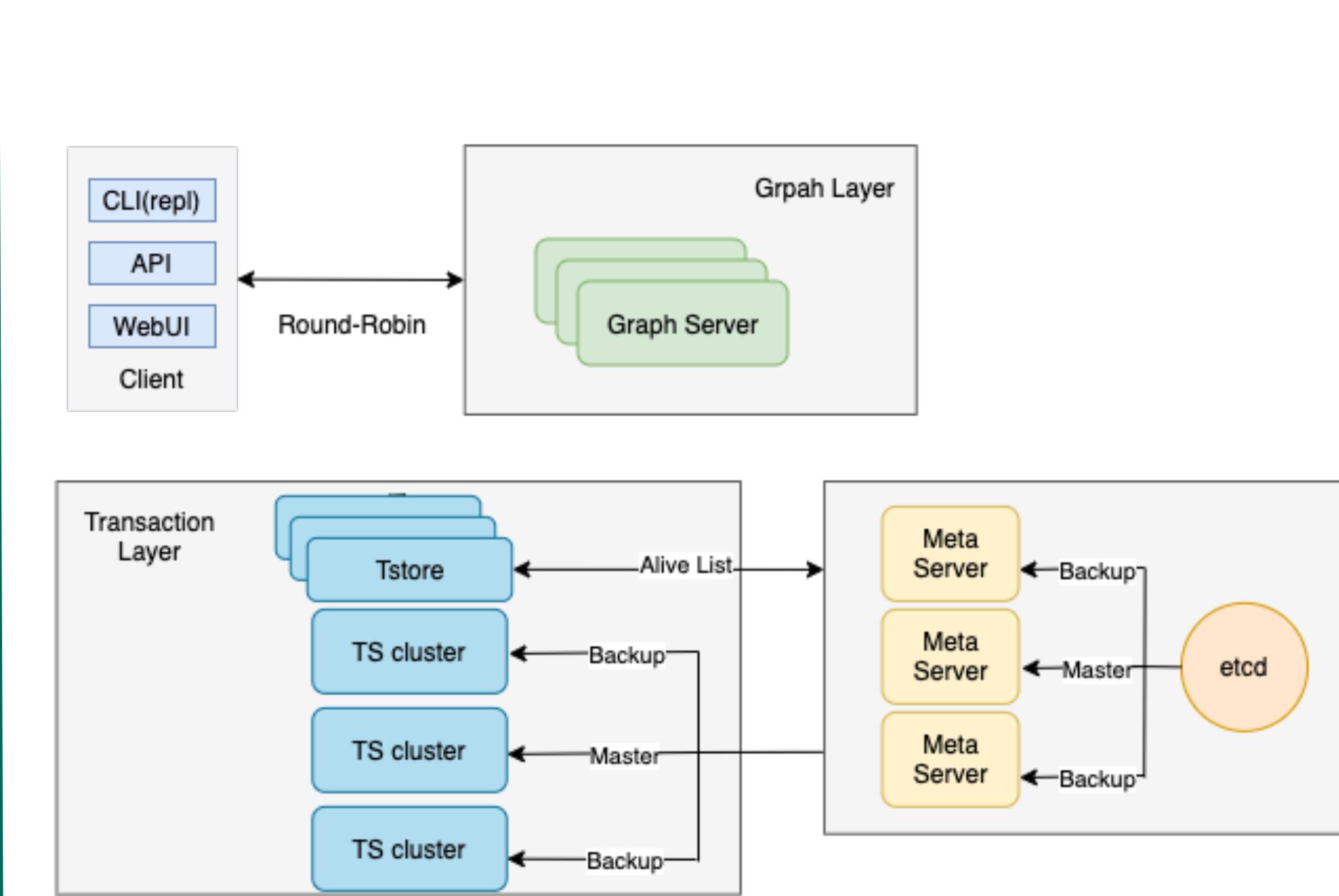


图1: AtlasGraph服务高可用

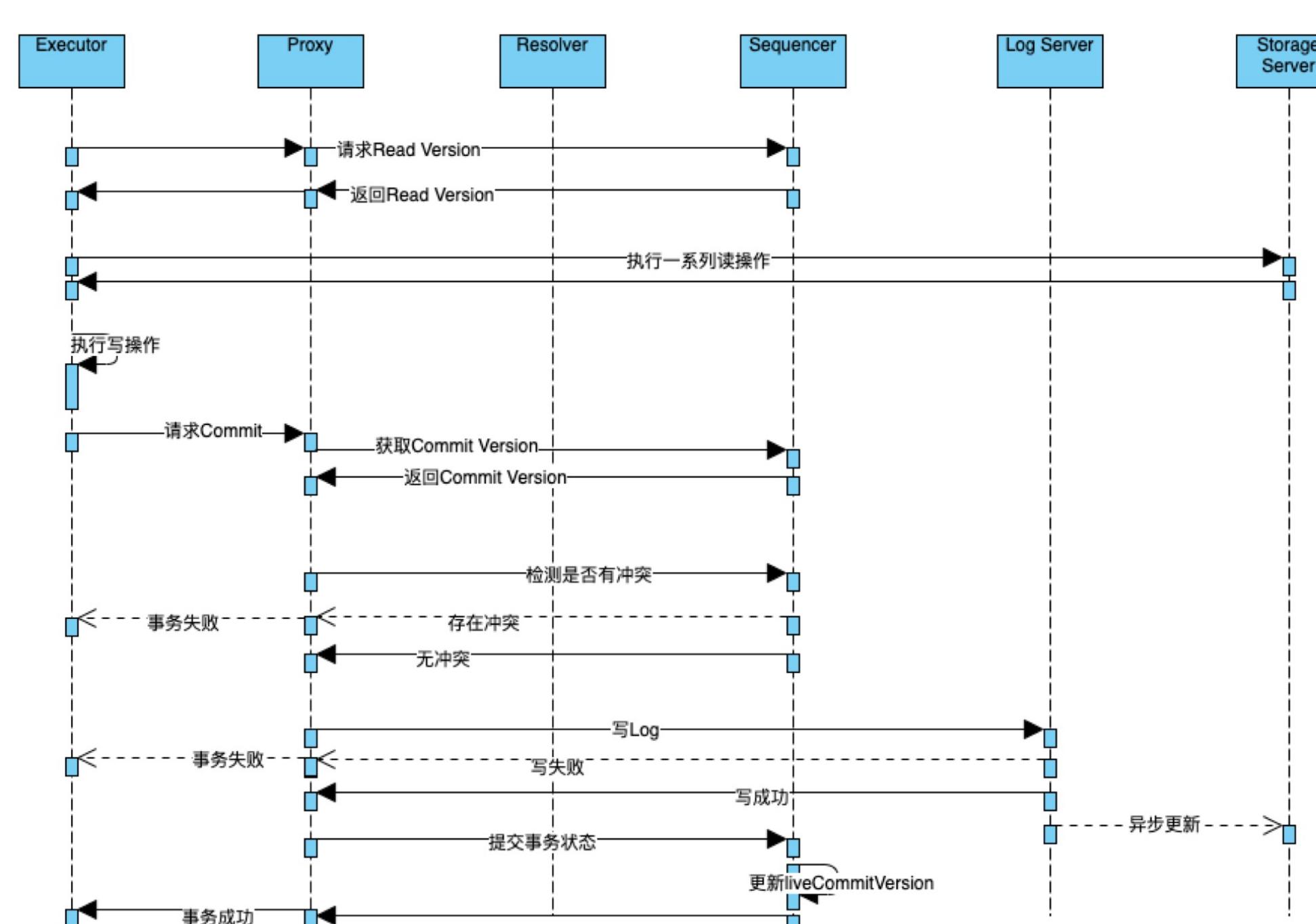


图2: AtlasGraph分布式事务

AtlasGraph架构概览

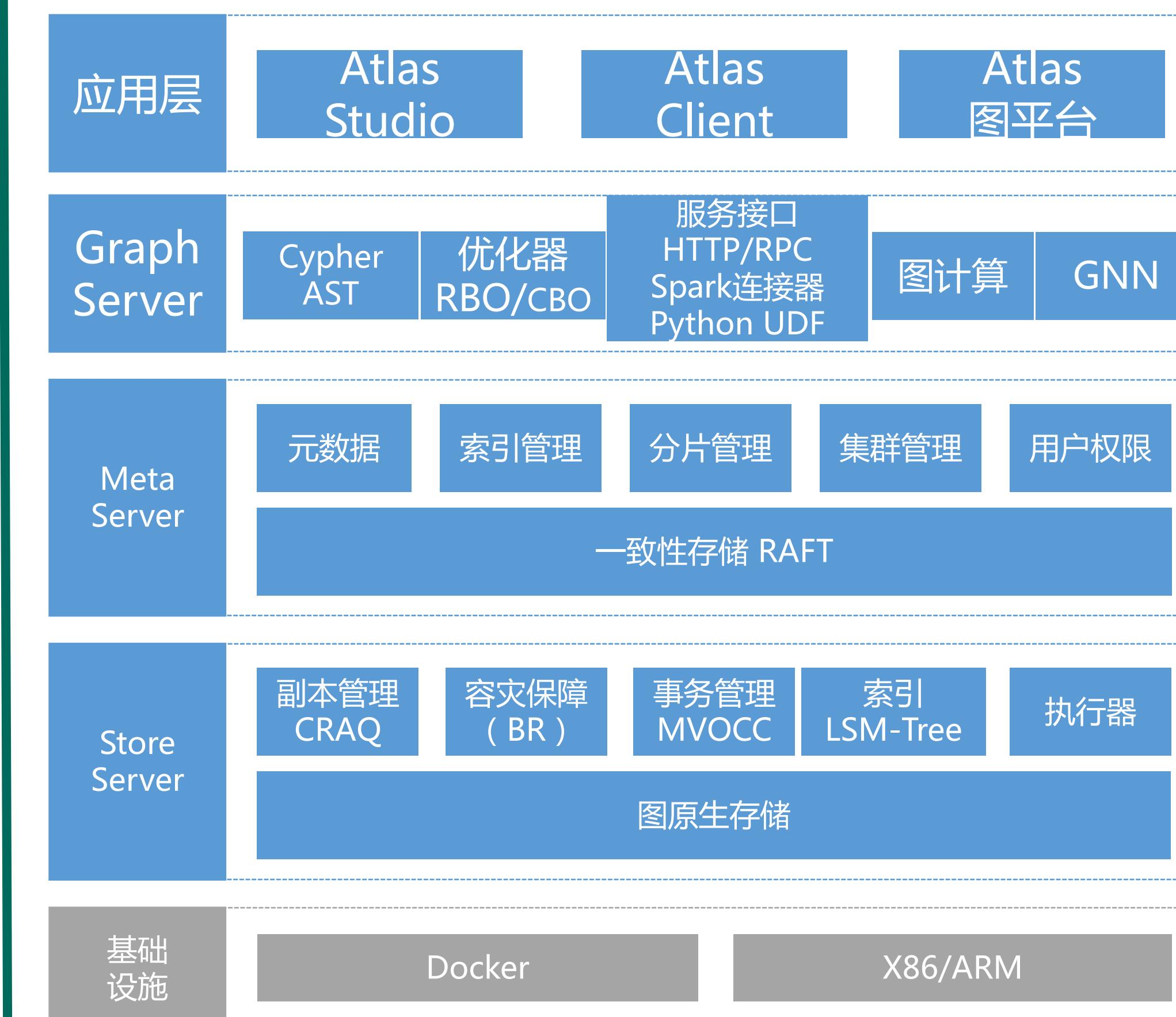


图3: AtlasGraph分布式查询性能

王芳
15811026226

2022全国知识图谱与语义计算大会
China Conference on Knowledge Graph and Semantic Computing
秦皇岛 8.24-8.27