**华东师范大学本科生毕业论文（设计）开题报告**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **论文题目** | **基于数据分段的高性能批量数据加载** | | | | | | 学院 软件工程学院 | | | | 专业 软件工程 | | | | | | | |
| 学生姓名 | | 李俊贤 | | 学号 | | | 10195101505 | | | | |
| **一、选题的背景与意义** | | | | | | | | | | | | | | | | | | |
| 选题背景  键值型数据库：  键值型数据库或键值型存储，是一种专门为了存储、查询以及管理键值映射表而设计的存储结构。由于其简单、高性能的特性，近年来键值型数据库成为了许多软件系统的基础设施的重要组成部分。就像文件系统对操作系统而言，图数据库，分布式数据库，流处理引擎等分布式系统也相当依赖于键值型数据库提供的存储能力。因此，键值型数据库的任何性能提升都可能给现代的软件系统带来极大的收益。  批量数据加载：  批量数据加载是指从CSV等半结构化文本文件中读取数据并将其插入数据库中的过程。这个操作常见于数据库的升级换代过程中，迁移过程大致是首先将旧有数据导出为CSV文件，然后将数据加载到新的数据库中。很多时候导入会在一个有限的运维窗口中才能进行，如果批量价在数据的时间过长，可能会冲击正常的业务压力，因此，批量加载的速度会至关重要。  LSM树：  LSM树是由谷歌设计的一个存储结构，其特点在于利用顺序写操作来提高写性能，由于其优秀的写入性能，许多键值型数据库，例如LevelDB,RocksDB等，选择LSM树作为核心的存储结构。  LevelDB：  LevelDB是来自谷歌的一个基于LSM树实现的键值型数据库，常被用作高性能的基础设施。且LevelDB并未提供批量数据加载的接口。一般而言要实现在LevelDB上的批量数据加载，在不修改LevelDB的情况下，只能在LevelDB的外部进行批量的插入操作，而插入过程中存在许多对于批量数据加载而言无意义的操作，这对性能有极大的影响。 | | | | | | | | | | | | | | | | | | |
| **二、研究的主要内容和预期目标** | | | | | | | | | | | | | | | | | | |
| **研究主要内容：在LevelDB上利用数据分段、SSTable层级细分等手段** | | | | | | | | | | | | | | | | | | |
| **三、拟采用的研究方法、步骤** | | | | | | | | | | | | | | | | | | |
|  | | | | | | | | | | | | | | | | | | |
| **四、研究的总体安排与进度** | | | | | | | | | | | | | | | | | | |
| **第1周：熟悉LevelDB的系统架构以及代码**  **第2周：做好开发的基础设施；首先准备好测试数据集以及单元测试，然后为LevelDB的LSM树实现可视化方便后续算法调试**  **第3周：细化算法设计；并实现作为对比基线的数据加载**  **第4-7周：开始实现前期设计的方案**  **第8周到结束：进行实验和理论分析，撰写并修改论文** | | | | | | | | | | | | | | | | | | |
| **五、参考文献** | | | | | | | | | | | | | | | | | | |
|  | | | | | | | | | | | | | | | | | | | |
| **论文题目** | |  | | | | | | 学院 | | | | | 专业 | | | | | |
| 学生姓名 | |  | | | 学号 | |  | | | |
| **六、指导教师意见** | | | | | | | | | | | | | | | | | | |
|  | | | | | | | | | | | | | | | | | | |
|  | | | | | 签字： |  | | | | 年 | | 月 | | | | 日 |  | |
| **七、开题答辩小组意见** | | | | | | | | | | | | | | | | | | |
|  | | | | | | | | | | | | | | | | | | |
|  | | | 小组成员签字： |  | | | | | | 年 | | 月 | | | | 日 | |  |

教务处编制