## 案例背景

某个MySQL服务器的user表有100万条数据,请运用多线程并发编程等相关基础知识,将这100万条数据秒级(8秒内)导出到若干Excel文件.

实际需求如果产品经理说要弄到一个Excel文件的话,你可以跟产品经理要提建议的,因为你100万条数据往一个Excel里面保存的话那么你的文件会很大,而且看数据也不方便,也会很卡.

如果你把一百万条数据导出到多个Excel里面,一个Excel里面有两万条数据,这样看起来也稍微方便点.

## 百万数据串行从数据库导出到Excel中会产生什么问题?

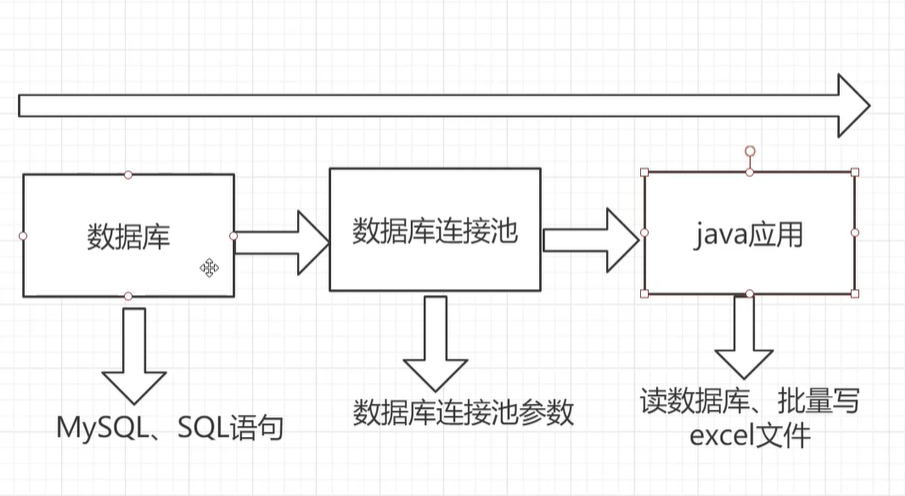
数据量比较大,串行的话耗时比较长,百万级别数据放到一个List集合里面可能会出现内存溢出的问题.就是list数据结构里面还没装到100W条数据之后就内存溢出了.

所以就需要并行分批次处理数据, 比如说一次拉取2W条数据处理,处理完了快速的释放掉内存,这样防止内存溢出,

## 如何秒级将百万数据并发写到Excel的文件里面

如果是串行的去做的话,肯定是很慢的,就得批量并行去做,就需要多线程了,每个线程处理若干笔数据,比如说每个线程处理2万条数据.这样处理起来就快很多了.

## 如何从架构角度优化性能,如何解决整体性能瓶颈



答案是:功能分层

上面是三个流程,只有每一次性能都很好,整体性能才能好,如果有一层性能不好的话,那么整体性能也不会太好.

比如说即使数据库和Java应用的性能好,但是你配置的数据库连接池配置的不到位,那么整体的性能就不好.

如果你数据库性能好,数据库连接池配置的到位,但是你Java应用性能不好,那么整体的性能也不会太好.

所以数据库层 数据库连接池 Java应用,这几层的性能都要好.

MySQL的SQL语句要有索引,数据库连接池再调优一下,Java应用这里代码再用多线程处理一下,就很快了.

如果哪一层慢的话,就每一层代码上下打个日志,看日志的时间,来确定哪里比较慢.

## MySQL数据库核心参数优化,Druid相关参数调优

MySQL 的 my.cfg参数调优

innodb\_thread\_concurrency 参数调整

druid相关参数调优

|  |
| --- |
| # 配置连接池的参数  initialSize=50  # 连接池的最大数据库连接数。设为0表示无限制。  maxActive=200  # 最大建立连接等待时间。如果超过此时间将接到异常。设为-1表示无限制。  maxWait=600000  # 连接池中的最小空闲连接数，Druid会定时扫描连接池的连接，如果空闲的连接数大于该值，则关闭多余的连接，反之则创建更多的连接以满足最小连接数要求。  minIdle=5 |

活跃数越高,性能就越高

## 代码使用方式

执行ExcelExporter类即可