

上海交通大学试卷 (A 卷)

(2022 至 2023 学年 第 1 学期)

班级号 _____ 学号 _____ 姓名 _____

课程名称 _____ 应用系统体系架构 _____ 成绩 _____

在本学期，同学们结合上课内容，对大二时开发的 E-BookStore 系统进行了升级和扩展。在系统设计与开发的过程中，大家碰到了很多问题，请你根据在本课程中学习到的知识，帮助他们分析处理以下问题（每题 5 分，共 100 分）：

所有题目请用中文作答

一、 请你帮助同学们解答在设计和开发服务器端程序时遇到的下列问题：

1. 在应用服务器中都会配置数据库连接池，对数据库连接进行池化管理。由于 E-BookStore 存在大量的用户，因此，有的同学认为需要把连接池的大小配置得很大，这样才能确保大量并发操作时不会因为等待数据库连接而导致性能变差。但是，很遗憾，这种配置方式不仅没有提高性能，反而降低了性能。请问，这是为什么？到底将数据库连接池配置成多大才合适？
2. 在大二开发的 E-BookStore 中，同学们并没有使用 `@Transactional` 注解进行事务管理，但是程序似乎也没有出现事务方面的异常。请问，你觉得这是为什么？
3. 同学们在设计订单支付服务的 `payOrder()` 时，在其中调用了两个方法：`pay()` 从用户账户余额中扣减订单金额；`shipping()` 生成订单配送信息。同学们希望 `pay()` 和 `shipping()` 在两个独立的事务中执行。请问，要怎样设计这三个方法的事务传播属性，以及怎样设计代码才能实现这个目标？

我承诺，我将严格遵守考试纪律。

承诺人：_____

题号	一	二							
得分									
批阅人(流水阅卷教师签名处)									

4. 同学们按照课程作业要求，将下订单功能改写为通过 Kafka 消息中间件来实现，这样可以在用户访问数量激增时更好地应对。请问，使用这种方式来下订单与之前同步下订单的方式相比，存在哪些缺点？你打算如何克服？
5. 在多线程编程中，我们使用 Synchronization 关键字对若干个方法进行了同步，但是有的同学认为应该在当前对象上使用 Synchronization 关键字进行语句同步。请问，你认为这些同学的想法是否正确？为什么？
6. 同学们使用 Redis 缓存来提高系统性能。具体做法是，将所有订单数据都放入 Redis。在双十一时，当用户生成订单时，先将订单放入 Redis，然后再同步写入数据库。请问，这种使用 Redis 的方式是否合理？为什么？
7. 同学们使用了 JWT 来进行用户认证，通过在不同服务器之间传递 JWT，可以实现单点认证。请问，JWT 的 token 为什么可以实现单点认证？

8. 同学们使用搜索引擎 Solr 来建立对图书信息的全文搜索功能。Solr 会使用反向索引来支持全文搜索，在 MySQL 中也有 FULLTEXT 类型的索引。请问，这种索引的结构是什么样的？为什么它能够支持全文搜索？
9. 同学们将采用微服务架构重构在线书店，将整个在线书店的功能拆分为多个微服务进行单独部署。在部署方案中，还有一个 Gateway 和一个注册中心。请问，部署注册中心的目的是什么？它如何可以帮助我们实现此目的？
10. 同学们在大二时开发的 E-BookStore 系统中，后端所有的 Controller 都是 @RestController，前后端传递的是 JSON 对象。因此，有的同学认为，这个版本后端暴露出来的已经是 RESTful Web Services 了。请问，你觉得这种看法正确吗？为什么？

二、 请你帮助同学们解答在配置和使用数据库，并进行系统优化时遇到的下列问题：

1. 大家在设计数据库时，需要遵循无损连接分解和最小化重复信息原则。请问，遵循这些原则的目的是什么？
2. MySQL 的缓存设置成了多个实例，而不是使用单个实例。显然，这样做每个实例的尺寸大小就受到了限制，既然如此，MySQL 为什么还要这样做？

3. 同学们在设计 MySQL 表时，对于书名字段，应该使用 VARCHAR 还是 CHAR 类型有些分歧。请你告诉同学们，使用 CHAR 能带来什么好处和缺点？
4. 同学们在对 MySQL 数据库执行物理备份还是逻辑备份时产生了歧义。有的同学认为物理备份省空间且速度快，所以应该使用物理备份。请问，物理备份为什么相对于逻辑备份省空间且速度快？并且，请你举出物理备份的缺点。
5. 同学们认为 E-BookStore 并没有太多的数据，不需要做分布式存储，因此无需对数据表进行分区操作。请你告诉同学们，除了分布存储之外，分区还能带来哪些好处？
6. 同学们希望使用 MongoDB 来存储系统中的所有数据，即从 MySQL 彻底迁移到 MongoDB。请问，你觉得这种方案是否可行？为什么？
7. 同学们打算使用 influxDB 来存储监控到的服务器状态。在 influxDB 中，在 field 上是不构建索引的。请问，这是为什么？这样会影响数据访问的效率吗？

8. 同学们构建了 MySQL 集群来提供更好的服务。在集群中，同学们对不同的实例赋予了不同的权重。请问，这个权重是指负载均衡的权重吗？如果是，它是如何进行负载均衡的？如果不是，那么它有什么作用？

9. 同学们打算使用数据仓库对 E-BookStore 的数据进行分析。请你举例说明，在构建数据仓库时为什么需要执行清洗、转换和加载（ETL）操作？

10. 同学们打算使用 Hadoop 的 MapReduce 计算范型来执行在线书店的大数据处理功能，例如对订单的统计和查询等。在编写程序时，Combiner 选择的是和 Reducer 一样的程序。请问，Combiner 的作用是什么？它和 Reducer 使用相同的代码的话，与 Reducer 到底有什么区别呢？