

1. 主要工作

我们组主要是研究了 Apache Spark 这一款流式实时分布式计算框架，主要是运用软件工程的方法对 Spark 进行分析，并在此基础上，将 Spark 扩展应用。在本项目中，我主要担任项目主管的角色，对项目实施过程中的项目进度计划文档进行管理与维护，同时参与整个项目过程的需求、设计、测试、评审等工作。

2. 遇到的困难及解决方法

在项目中遇到的主要困难有：1. 项目初期对于 Spark 的了解不够，只能通过小组同学之间的研讨以及查看一些资料，才逐渐的了解了一些皮毛。2. 在项目过程中，管理 mpp 文件时，没有对项目的一个很好的把握，导致的一些问题，比如资源分配不合理，项目时间不合理，基线设置不正确等情况。通过老师的提醒，以及自己在网上查找了一些方法只会，才将 mpp 文档调整过来。

3. 需要改进的努力的地方

对于项目的了解仍不够，很多的编程和对比测试工作都是其他组员完成的，自己的参与度不高；在一些问题上，和组员的共同协作还不够。

4. 对本课题的感想及建议

本课程时一门很好的将实际与理论结合的课程，是很好的一次锻炼自己参与完整的软件项目的机会。从软件项目的课题选取、调研、需求分析到软件实现，到测试。虽然我们使用的是开源项目，进行的改进的地方也不是很多，但是在参与项目的整个过程中，我们了解了软件工程的方法，也学到了除了编程以外的一些软件项目的管理的知识，对以后的学习和工作都有益。