

软工实验个人总结

一、概述

本次实验我们选择的实验目标是 Hadoop 中的 MapReduce 模块，实验小组由四名成员组成，实验内容主要分为前期调研、需求分析、测试分析、测试等四个阶段，以及贯穿整个实验的项目管理，下面我将对实验的具体内容以及过程中遇到的主要问题进行个人分析。

二、实验内容

（一）前期调研

这部分的工作内容主要是确定实验目标和实验范围，这期间小组成员调研了多种开源软件，最后将目标定为目前应用非常广泛的 Hadoop，也将实验范围初步确定为 Hadoop，包括分布式计算 MapReduce 和分布式存储系统 HDFS 两部分。

这部分出现的主要问题是调研不全面，对任务量估计不充分。最初我们确定的是 MapReduce 和 HDFS 两部分，在后续的分析当中，我们发现任务量过大，于是进行了调整，将目标缩减为 MapReduce 模块。

（二）需求分析

这部分工作内容是确定用例图和用例规约，我们采取的方式是首先每个人通过学习调研对于 MapReduce 有个整体的认识，提炼出整个工作流程，之后根据整个的流程确定用例图，将每个小的模块抽象成一个用例，再之后四人分工，分别调研每个功能模块，确定用例规约。

这部分出现的主要问题是 Actor 选择不合理，功能主次不分。Actor 的选择我们调整过多次，主要原因是我们对用例图的了解不深入，对于应该选择将什么对象设定为 Actor 不明确。功能主次不分主要是因为我们最初的设想是将 MapReduce 的功能进行一个全面的分析，但是分析的过于全面就会导致主体功能不突出，我们最初一版的用例图过于繁杂，主体功能的计算只占其中极少一部分，之后我们根据老师的建议调整了我们的方案，采用剥离主体功能，次要功能服务于主体功能的方式进行描述，这样一来对于 MapReduce 的分析就更加清晰。

（三）测试分析

这部分工作的内容是确定测试用例和测试用例规约。测试用例我们采用的是匹配需求用例的方式，每个测试用例对应一个需求用例。测试用例规约针对需求用例的前提条件、基本流、分支流进行设计。

这部分的主要问题是对于非功能需求考虑不充分，由于设备等限制，我们没有特别好的办法对非功能需求进行验证。

（四）测试

测试是占用时间最多的一个阶段，由于需求配置环境、编写代码。我们采用了 KNN 等案例对功能需求进行了测试，采用了单机模式与 Hadoop 模式对比的方式进行了非功能需求的测试。

这部分的主要问题是我们发现单机模式反而比分布式模式的运行速度要快，我们考虑主要的原因可能是网络带宽、结点太少等原因，但由于环境问题等限制，我们最后也没有进行较好的测试。

（五）项目管理

此次实验的项目管理我们选择了工作日志、工作量分析、变更记录等方式，对我们整个的工作进展以及工作中出现的问题进行管理与分析。

此部分的主要问题是我们的计划的落实并不是十分的好。

三、 总结分析

本次实验的收获主要有以下几点。

首先是对 Hadoop 的工作机制有了较好的认识，熟悉了 Hadoop 的工作流程以及部分内部实现。

第二是对 GitHub、MS 等管理工具有了了解，这对于日后的学习工作都有较好的帮助。

第三是整个实验过程小组内成员间的协作，以及小组间的交流过程都是非常好的一次体验。通过观察小组成员的表现，认识到了自身的不足，通过观察其他小组的表现，发现自己所在小组的问题。

本次实验觉得不足的地方主要是小组返工次数太多，内容量也很大，究其原因一部分是由于我们对于整个实验没有一个特别好的宏观认识，所以真正动手去做的时候不是很明确，另一部分原因是觉得整个的实验进度多少存在着些问题，前期调研留用的时间过少，在对项目还没有很好的认识的时候已经着手去做，所以出现了很多问题。

最后感谢各位老师的指导、小组成员的帮助以及其他组同学的质疑指正，我们最后顺利的完成了本次实验！