

《开源软件设计与开发》课程总结

姓名：黄陈超

学号：51194507005

1、开源理解

我理解的开源就是一个项目的代码可以免费的提供给所有人，允许所有人获取、使用、修改和发行。开源的意义主要在于合作，通过合作，才能形成围绕某个开源项目的软件生态。相信大家对 Github 肯定都不陌生，GitHub 是一个面向软件项目的托管平台，可以用于托管各种类型的软件项目，包括开源项目和私有项目。很多人会把自己写的一些项目传到 Github 上分享给大家。我也经常从 Github 上下载一些牛人关于某个算法实现的代码，可以从中学学习到了大佬对于算法的理解，怎么样更加简洁方便实现某个功能等。开源可以说程序猿的福利，很多程序猿可以在 Github 上下载到代码并做出相应的修改从而实现自己的预期功能。对于我们学生来说，我们可以在这里找到一些开源项目并参与开源项目，可以很大程度上锻炼我们自身的能力。对于以后找工作也是很有帮助的。

开源社区又称开放源代码社区，一般由拥有共同兴趣爱好的人所组成，根据相应的开源软件许可证协议公布软件源代码的网络平台，同时也为网络成员提供一个自由学习交流的空间。由于开放源码软件主要被散布在全世界的编程者所开发，开源社区就成了他们沟通交流的必要途径，因此开源社区在推动开源软件发展的过程中起着巨大的作用。比如说我在上一次作业里写到的 ubports 社区，在开发初期，开发人员可以在这里分享想法并相互指导，希望将 Ubuntu Touch 平台引入越来越多的设备。最初的 Ubuntu Touch 是由 Canonical 设计和开发的。快速和平稳的发展让这个项目有了起色。在 2017 年四月 Canonical 宣布终止对操作系统的支持之后，它被社区接手了。反响强烈，社区正在迅速发展。出于对 Ubuntu Touch 的热爱让他们聚集在了社区，共同完成和实现 Ubuntu Touch 应用软件生态。

2、开源贡献

我参与的是 TiDB 的开源项目，TiDB 是 PingCAP 公司设计的开源分布式 HTAP (Hybrid Transactional and Analytical Processing) 数据库，结合了传统的 RDBMS 和 NoSQL 的最佳特性。TiDB 兼容 MySQL，支持无限的水平扩展，具备强一致性和高可用性。TiDB 的目标是为 OLTP (Online Transactional Processing) 和 OLAP (Online Analytical Processing) 场景提供一站式的解决方案。

我做的任务是将一个函数向量化。对于一个 builtin 函数，比较重要的是进行语法解析以及如何求值。其中语法解析部分需要了解如何写 yacc 以及如何修改 TiDB 的词法解析器，较为繁琐，社区已经将这部分工作提前做好，大多数 builtin 函数的语法解析工作已经做完。对 builtin 函数的求值需要在 TiDB 的表达式求值框架下完成，每个 builtin 函数被认为是一个表达式，用一个 ScalarFunction 来表示，每个 builtin 函数通过其函数名以及参数，获取对应的函数类型以及函数签名，然后通过函数签名进行求值。我选择了 builtinBitCountSig 函数并将其进行向量化。该函数的功能是实现一个 64 位的整型数里面的一的个数的统计。

我学习了一下关于 go 语言的语法和 git 相关的知识，试着去实现这个功能。在导师的帮助下成功的提交了一个 pr，但是不幸的是，我并没有使函数变得优化，反而使功能退化了。收到了社区 reviewer 如下的评价：



我重新查看了代码逻辑，并且在导师的指导下修改了代码，中间遇到了很多问题，比如没有进行 go fmt 导致上传 push 的时候测试点没用通过，好在有导师和同学的帮助，最终还是解决了问题。但是社区的 reviewer 觉得我统计位数的方法没有原先的好，要求我用社区原先的方法。我通过 batchmark 做发现确实是我更新后的统计算法好并且给予回复。下面是社区 reviewer 关于修改统计方法的回复。



社区 reviewer 要求我将统计位数的代码放在一个新的函数里面，并给它取个名字，作简单的功能描述，我按照他们的需求再次修改了代码并重新 commit。最后成功的 merge 到了项目里。



这个过程中我要感谢这个项目的导师，他给了我很多帮助，因为我对开源这方法包括 go 语言、git 的相关操作都不是很熟悉他给了我很大的帮助。下面是我 commit 之后的代码地址：
<https://github.com/pingcap/tidb/commit/d744bc41d84034d24e658ddbddd9944d685e7297>

3、课程反馈

希望能在课程开始就能向学生讲述一些关于如何参与一个开源社区的项目,包括怎么提 pr、还有关于 git 的一些操作。最好是还能有一些实例或者课堂小组间讨论交流,这样对新手更加友好。

4、参考文献

《十分钟成为 TiDB Contributor | 添加内建函数》