

# 《开源软件设计与开发》课程总结

姓名：夏小雅

学号： 51195100018

## 1、开源软件

这学期的课上下下来，也接触到了不同的导师，还有表伟老师、建胜老师的各种内容分享，我感觉开源最大的精神核心是自由。这种哲学意味也很符合自由软件运动的精神领袖 Richard Stallman 的风格。



大胡子哲学家 Richard Stallman

开源起于 GNU 运动，这个运动的起初是为用户的计算自由而战斗，它是一场道德运动【1】；而开源发展至今，它的哲学并不涉及对与错，而是一种实际的价值观，这种价值观认为，如果我们使用“开源”这种协作开发的模式开发，将会更容易得到一个高质量的软件。从哲学层面上，“开源”是“自由软件”的更务实的一种应用，开源软件是与 GNU 的精神一脉相承的。

顾名思义，开源的本质是开放源代码。这种软件发布模式使得代码可以被免费下载和更改。同时，不同的产品、项目往往会衍生出相应的开源社区，社区的形式使得开发者们被相同的兴趣连接在一起，和产品本身形成正反馈：既促进产品的功能和性能的优化，也能吸引新用户和挽留老用户。可以说，开源软件和开源社区是密不可分的。开源社区可以是针对某个产品，也可以是对整个软件行业的社区。

GitHub 是最有名的开源代码托管平台和分布式管理系统，也是全世界开发者们的社交平台。GitHub 的风格也是自由与友好的，这一点从网站的交互界面风格

就能提现，由于它是如此用户友好，有些人甚至使用 GitHub 来管理其他类型的项目 – 比如写书。除此之外，还有很多其他有意思的项目，用户可以自由访问他们的代码，感受到世界上其他开发者们的奇思妙想或者开发方式，甚至 fork 别人的项目，或提交 PR 来成为这个项目的贡献者。可以说，GitHub 的自由协作是“开源”精神的完美体现。

## 2、开源贡献

我参与的开源项目是 ServiceComb，它是业界第一个 Apache 微服务顶级项目，是一个开源微服务解决方案。致力于帮助企业、用户和开发者将企业应用轻松微服务化上云，并实现对微服务应用的高效运维管理。其提供一站式开源微服务解决方案，融合 SDK 框架级、0 侵入 ServiceMesh 场景并支持多语言。

ServiceComb 的项目负责导师是姜宁老师，我们在开源课的 slack 中开了一个频道来交流和围观开源社区提出 issue 和提交 PR 的过程。围观的 PR 链接为：  
<https://github.com/apache/servicecomb-pack/issues/571>

提出 issue 是在 <https://github.com/apache/servicecomb-pack/issues/571>  
具体修复：<https://github.com/apache/servicecomb-pack/pull/574>

围观了这个 PR 后，我们讨论了围观这个 PR 的收获，具体内容见图

The screenshot shows a GitHub interface with a pull request (PR) titled "#571 omega 开启 hystrix, 使用Feign调用bug". The PR is by user "pengyu0929". The description mentions a bug related to Hystrix and Feign. The code snippet shows Feign configuration for Hystrix, including thread pool settings and command timeout. The discussion includes comments from "Willem Jiang" and "方孝君" (Fang Xiaojun), who discuss the technical details of ThreadLocal and Hystrix. The PR is linked to issue #571.

**Thread**  
# servicecomb

**夏小雅** 2 months ago  
这个PR中提问者在提出issue时给出了bug的详细描述，包括代码片段和测试日志。提问者说在调用次数大于coreSize时，之后的每次调用，omega设置到feign调用的header中的globalTxId均为第三次的值。但是贴出来的测试日志中第三次和第四次的globalTxId是不一样的，导致我也不太理解这个问题。在PR中，我看到问题的原因是线程复用，由于Hystrix有隔离策略，线程无法更新，导致globalTxId不能被传递，于是提问者通过参考Spring Cloud使用Hystrix传播ThreadLocal对象的策略，添加了一个处理线程变量的hystrix callable接口，自定义并发策略，解决了这个问题。

**Willem Jiang** 2 months ago  
这里面有一个技术上下文是Thread Local传递是有局限性的，在线程发生切换的时候，我们需要将ThreadLocal的变量拷贝到新的线程中，才能保证之前涉及的业务逻辑能够继续执行。

**Willem Jiang** 2 months ago  
大家在后续提PR的时候也要想想如何描述清楚问题，让更多人能够通过阅读PR就能理解其中的上下文。

**方孝君** 2 months ago  
这两天我看了一下serviceComb-saga的文档（其中有Omega、Alpha的概念），了解了一下Hystrix、Feign的相关概念和基本原理。我了解到Hystrix是Netflix开源的一款容错框架，它提供了线程池隔离和信号量隔离两种资源隔离策略（资

Servicecomb 在 github 上的开源代码中给出了许多 demo，最后我们从 saga-spring-demo 入手，跑了其中的预定系统“booking”的 demo，进一步加深对微服务框架的理解。

## 从 GitHub 的开源数据认识开源

除了参与到 ServiceComb 项目中，我和开源进一步的连接是通过对 GitHub 的全域数据的理解与分析。前面已经介绍了 GitHub 与开源的关系，作为一个功能强大和用户集群庞大的社区和管理系统，它所拥有的项目数据和用户行为数据都有很大的研究价值。我首先阅读了一些论文，看看别人都在做什么，包括但不限于 GitHub 的数据集相关研究：

git2net - Mining Time-Stamped Co-Editing Networks from Large git Repositories

GH Archive - A project to record the public GitHub timeline, archive it, and make it easily accessible for further analysis.

上面两个是对通过监控 GitHub 的时间轴来对数据进行存档。当然，GitHub 最有说服力的数据还是来自于官网的 API，毕竟数据是会改变的，存档下来的数据无法避免更新的问题，但是 API 最大的缺点 - 数据访问会有 Rate Limiting。

我和实验室的学长一起，抽取了 GitHub 在 2019 年活跃过的用户，并尝试进一步得到用户的详细信息，部分数据如图所示。

user_name	node_id	avatar_url	gravatar_id	type	site_admin	name	company	blog	location	email	hireable
mojombo	MDQ6VXNlcjE=	https://avatar		User	0	Tom Pres	(Null)	http://to	San Francisco	tom@moj	(Null)
wycats	MDQ6VXNlcjQ=	https://avatar		User	0	Yehuda K Tilde, Inc.		http://ye	San Francisco	wycats@	(Null)
ivey	MDQ6VXNlcjY=	https://avatar		User	0	Michael E	RiotGames	http://gw	Tuscumbia,	ivey@gw	(Null)
evanphx	MDQ6VXNlcjE=	https://avatar		User	0	Evan Phoenix	hashicorp	http://bl	Los Angeles	evan@ph	(Null)
vanpelt	MDQ6VXNlcjE3	https://avatar		User	0	Chris Van	crowdfunder	wandb.c	San Francisco	vanpelt@	(Null)
wayneesequin	MDQ6VXNlcjE4	https://avatar		User	0	Wayne E.	http://starkar		Buffalo, NY	wayneese	(Null)
brynary	MDQ6VXNlcjE5	https://avatar		User	0	Bryan Hel	Code Climate	http://cc	New York Ci	(Null)	(Null)
kevinclark	MDQ6VXNlcjIw	https://avatar		User	0	Kevin Clark	Cue	http://gl	(Null)	kevin.clar	(Null)
technoweenie	MDQ6VXNlcjIx	https://avatar		User	0	risk dang	GitHub	http://te	Louisville,	Ctechnowe	(Null)
takeo	MDQ6VXNlcjIy	https://avatar		User	0	Toby Ster	@zendesk	https://z	Portland, O	tobysterr	(Null)
caged	MDQ6VXNlcjIz	https://avatar		User	0	Justin Pal	(Null)	http://lal	Portland, O	encyteme	(Null)

hireable	bio	public_repos	public_gists	followers	following	created_at	updated_at
(Null)	None	61	62	21737	11	2007-10-20 05:24:19	2019-12-10 01:01:00
(Null)	None	224	751	9449	4	2008-01-12 05:38:33	2019-12-21 00:00:00
(Null)	None	86	79	97	2	2008-01-12 15:15:00	2019-11-22 15:00:00
(Null)	None	148	441	1258	6	2008-01-12 16:46:24	2019-12-20 18:00:00
(Null)	None	54	49	101	15	2008-01-13 05:57:18	2019-12-16 21:00:00
(Null)	None	97	95	700	16	2008-01-13 06:02:21	2019-12-11 11:00:00
(Null)	Co-founder and CEO, Code Climate	164	68	611	27	2008-01-13 10:19:47	2019-12-09 17:00:00
(Null)	None	23	27	88	4	2008-01-13 18:33:26	2019-10-22 06:00:00
(Null)	:metal:	173	106	2512	17	2008-01-14 04:33:35	2019-12-27 17:00:00
(Null)	None	19	10	78	11	2008-01-14 11:25:49	2019-11-08 01:00:00
(Null)	Engineering, Data Visualization, Design. P 127	98	2041	42	2008-01-15 04:47:24	2019-12-29 07:00:00	
1	Entrepreneur, developer, designer, teach 61	43	1205	217	2008-01-15 05:40:05	2019-12-18 02:00:00	
(Null)	None	91	50	151	31	2008-01-15 07:49:30	2019-12-27 05:00:00
(Null)	Founder of wandb.com	29	7	283	14	2008-01-15 12:50:02	2019-12-27 22:00:00
1	None	78	241	31	12	2008-01-15 15:44:31	2019-12-02 15:00:00
(Null)	None	2	2	37	0	2008-01-18 14:09:11	2019-12-19 01:00:00
(Null)	I enjoy learning, solving problems, and be 39	14	38	21	2008-01-19 05:50:12	2019-10-17 14:00:00	
(Null)	No.	74	96	376	10	2008-01-19 08:01:02	2019-12-02 06:00:00
(Null)	None	168	178	1157	165	2008-01-22 09:14:11	2019-11-25 06:00:00
(Null)	None	309	80	479	114	2008-01-24 04:40:22	2019-12-07 23:00:00
(Null)	None	156	165	1244	19	2008-01-24 04:44:30	2019-12-06 21:00:00
1	👋	110	40	670	88	2008-01-25 02:28:12	2019-12-29 15:00:00

如果能将这些数据建一个用户关系网络图,并检测高活跃用户所属的组织或者操作所针对的项目,以此得到有影响力的用户和项目,相信对开源社区和开源协作能有更深一步的理解。

### 3、课程反馈

我认为该课程的初衷十分新颖,即让每个同学都参与到实际的开源项目中去。这门课合作的开源项目来自于国内顶尖企业中的明星项目,如果每位同学都能实际深度的参与到这些项目中,将对无论是开源还是技术的理解都有进一步地提高;可惜后续地推动力不足,提前没有预设好具体的参与协作方式,对应的开源项目 owner 由于自身也都很忙,没有太多精力真正持续敦促或推动这件事情,导致很多同学其实没有参与到项目中去。

但是这种方式十分有意思,每位导师来宣讲地过程都是同学接触到大牛的机会,导师们风格各异,并且实力都很强。相信如果规划好具体的项目参与方式,强制性地敦促产出,可以让我们学到更多东西。

### 参考文献

【1】Richard Stallman. 为什么开源错失了自由软件的重点