

阿里开源经验分享

阿里云
朱照远（叔度）
2015-01-10

自我介绍

- 真名朱照远，花名叔度（sudo）
- 2009年加入淘宝，现负责阿里云-核心系统-Web平台
- 技术方向为Web架构优化，CDN，云计算
- 开源爱好者，Tengine项目发起人
- 阿里巴巴开源委员会成员，见证了阿里巴巴开源过程

个人开源的经验

开源对于工程师的意义

- 东西做好了，回报自然会来
- 通过开源，可以得到很多开发人员和高手的帮助，不断提高自己的能力
- 通过开源，可以得到更多的用户
- 通过开源，可以让自己的代码有更长的生命周期
- 用户的用法，会超出工程师的想象

如何参与别人开源项目

- 从小事做起，写文档做翻译都是做贡献
- 遇到问题提bug report，有能力就提patch
- 主动贡献，从小到大，逐步获得信任
- 不被接受也不要气馁

自己做开源项目

- 需要很好的解决某一个问题的
 - 如Memcached和Redis各自诞生并流行的原因
- 项目要容易上手
 - Quickstart文档
 - 简单的步骤可编译运行，减少外部依赖
- 代码、架构良好，容易扩展
- 文档齐全

开源社区的运营

- 制定规则（社区的“宪法”）
 - 如何提patch
 - code review
- 打造金字塔型的社区
 - 用户/开发者/核心
- 宣传
 - 用户的口碑是最核心和最重要的传播手段
 - 文章、会议介绍

贵在坚持

- 开放了源代码只是开源的第一步
- 持续更新，包括修正问题，增加功能等
- 保持较快的发布节奏（更新次数）

案例：Tengine开源

Tengine简介

- 阿里巴巴基于Nginx开发的开源Web服务器
- 针对大访问量网站的需求，添加了很多高级功能和特性
- 从2011年12月开始开源
- 据W3Techs统计，全球用户量排名第9，使用量0.2%

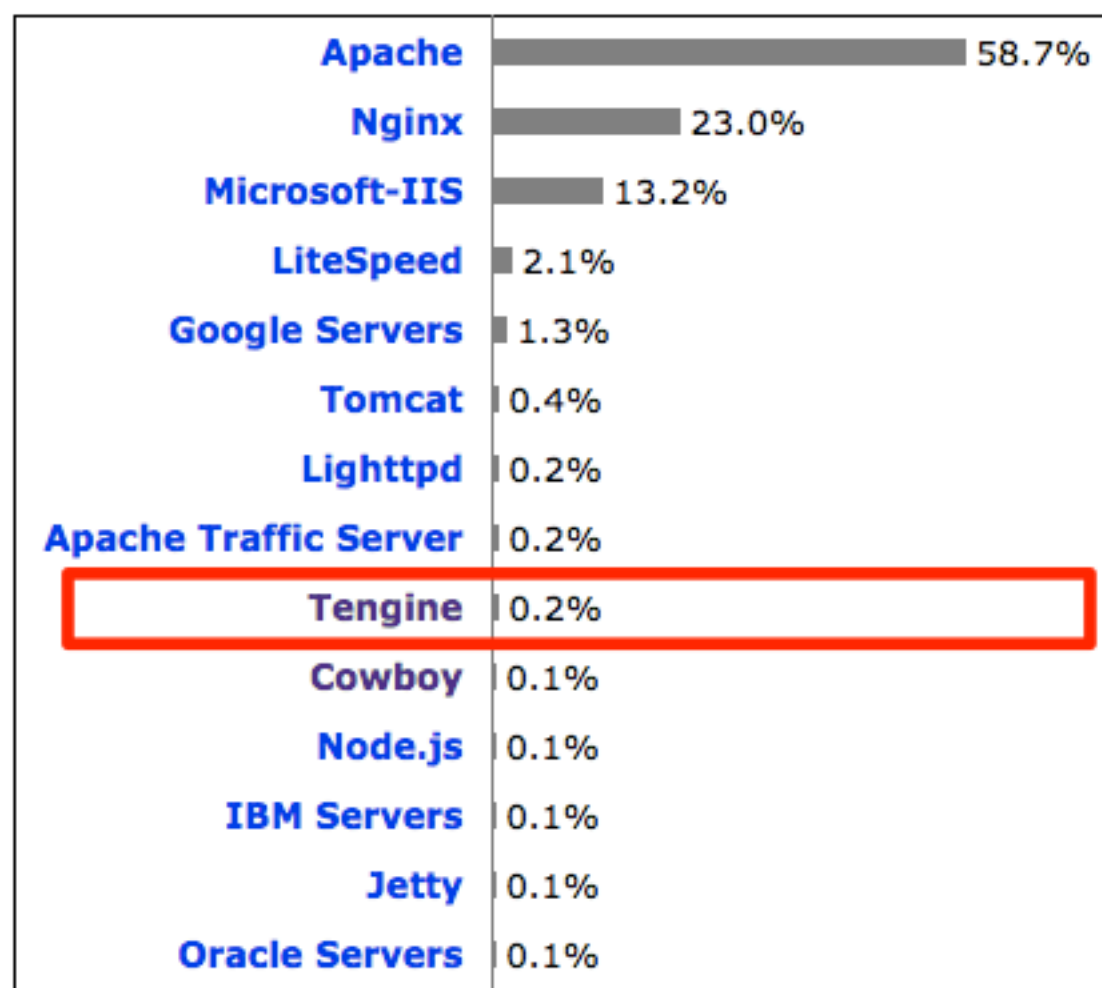
The word "Tengine" is displayed in a stylized, italicized font with a blue-to-white gradient and a 3D shadow effect.

Tengine的一些特性

- 动态模块加载（DSO）支持
- 支持SO_REUSEPORT选项，建连性能提升为官方Nginx的三倍
- 支持SPDY v3协议，自动检测同一端口的SPDY请求和HTTP请求
- 流式上传到HTTP后端服务器或FastCGI服务器，大量减少机器的I/O压力
- 更加强大的负载均衡能力，包括一致性hash模块、会话保持模块，主动健康检查
- 动态脚本语言Lua支持，增加灵活性
- 输入过滤器（input body filter）机制支持
- 支持管道（pipe）和syslog（本地和远端）形式的日志以及日志抽样
- 组合多个CSS、JavaScript文件的访问请求变成一个请求
- 自动去除空白字符和注释从而减小页面的体积
- 监控系统的负载和资源占用从而对系统进行保护
- 更强大的防攻击（访问速度限制）模块

Tengine的外部用户

- Internet Archive
- 腾讯
- 土豆
- 京东
- PPTV
- 国美
- 凤凰网
- 太平洋电脑网
- 开源中国
- ...



Tengine社区

- 网站
 - <http://tengine.taobao.org>（中英文）
- 邮件列表
 - <http://code.taobao.org/mailman/listinfo/tengine-cn>
- GitHub上开发
 - <https://github.com/alibaba/tengine>
 - star数目2364，fork数目719
 - 社区通过pull request方式驱动，做代码review

Tengine的开发和发布

- 开源后发布21个版本
- 几乎每周都有代码提交
- 核心开发者8人
- 国内外代码贡献者共37人
 - 中国、美国、荷兰、匈牙利、葡萄牙



企业开源的经验

开源前的思考

- 商业公司的目的是盈利
- 知道什么不能开源，才能做好开源
 - 不要暴露核心竞争力和核心机密
- 不伤害自己，却有利于他人，何乐而不为
- 做开源并不困难，任何公司都可以

开源与否都要遵守规则

- 使用开源软件和修改开源软件都要遵守游戏规则
- 开源软件最基本的游戏规则就是版权和协议
 - 宽松型：BSD/Apache/MIT
 - 严格型：GPL（国内有不少公司没有遵守）

开源对于企业很有好处

- 社会公益
- 技术影响力
- 工程师文化
- 吸引人才
- 提高开源出去的软件的质量

开源是一种商业模式

- 小企业对抗大玩家的杠杆
 - 通过开源免费做大用户基数
- 开源免费，增值收费
 - MySQL/Nginx
- 软件免费，服务收费
 - Red Hat
- 组成生态
 - Docker/OpenStack

如何内部弘扬开源文化

- 不要制止工程师的开源冲动，但可以制定流程规范
- 开源项目的成就感是对工程师的最大鼓励
- 内部的开源项目经验交流
- 对优秀的开源项目进行奖励

我看中国开源的未来

- 开源大大减少了中国IT界同国外的技术差距
 - 开源软件本身以及技术被直接使用
 - 工程师通过对开源软件的学习可迅速提高自身水平
- 道路崎岖，但前景光明
 - 人口红利
 - 创新与技术爆炸
 - 一定会涌现出很多优秀的开源企业和开源个人、项目
 - 中国的开源影响力最终会非常大

案例：阿里的开源情况

阿里开源大体情况

- 共开源自主研发软件100余个
- 涵盖前端、后端、数据库、文件系统、内核、硬件等多方面
- 对阿里使用的若干项目贡献了代码



典型项目-开源软件增强

- Linux Kernel (定制内核)
- Hadoop (大数据生态)
- MySQL (数据库)
- JVM (Java虚拟机)
- LVS (负载均衡)
- Tengine (Web服务器)
- TrafficServer (HTTP缓存软件)

| | | |
|-----------|------------------------------|------------|
| ⊕ No. 112 | MathEmbedded Consulting | 160(0.06%) |
| ⊕ No. 113 | General Electric | 158(0.06%) |
| ⊕ No. 114 | Real-Time Remedies | 156(0.06%) |
| ⊕ No. 115 | Tao Bao | 140(0.05%) |
| ⊕ No. 116 | Open Nandra | 137(0.05%) |
| ⊕ No. 117 | Tilera | 135(0.05%) |
| ⊕ No. 118 | US National Security Agency | 134(0.05%) |
| ⊕ No. 119 | Barco | 129(0.05%) |
| ⊕ No. 120 | CSR | 128(0.05%) |
| ⊕ No. 121 | secunet Security Networks AG | 124(0.05%) |
| ⊕ No. 122 | MSC Vertriebs GmbH | 123(0.05%) |
| ⊕ No. 123 | RisingTide Systems | 118(0.04%) |
| ⊕ No. 124 | VIA Technologies, Inc. | 115(0.04%) |
| ⊕ No. 125 | OKI SEMICONDUCTOR | 114(0.04%) |
| ⊕ No. 126 | Realtek | 113(0.04%) |
| ⊕ No. 127 | Ksplice | 109(0.04%) |
| ⊕ No. 128 | Eukrea Electromatique | 108(0.04%) |
| ⊕ No. 129 | Solid Boot | 106(0.04%) |
| ⊕ No. 130 | Nuvoton Technology | 104(0.04%) |
| ⊕ No. 131 | Fusion-io | 103(0.04%) |
| ⊕ No. 132 | Adaptec | 101(0.04%) |
| ⊕ No. 133 | AudioScience | 98(0.04%) |
| ⊕ No. 134 | Selenic Consulting | 94(0.03%) |
| ⊕ No. 135 | CISCO | 89(0.03%) |
| ⊕ No. 135 | MEV Limited | 89(0.03%) |
| ⊕ No. 135 | Gaisler Research | 89(0.03%) |
| ⊕ No. 135 | Holoscopio Tech. | 89(0.03%) |
| ⊕ No. 139 | Candela Tech. | 88(0.03%) |
| ⊕ No. 140 | GNU | 80(0.03%) |
| ⊕ No. 141 | Telargo | 79(0.03%) |
| ⊕ No. 141 | Tower Technologies | 79(0.03%) |
| ⊕ No. 141 | IDT | 79(0.03%) |
| ⊕ No. 144 | OMICRON electronics | 77(0.03%) |
| ⊕ No. 145 | Apple | 73(0.03%) |
| ⊕ No. 145 | Digi International | 73(0.03%) |
| ⊕ No. 147 | Tensilica | 70(0.03%) |
| ⊕ No. 147 | SMSC | 70(0.03%) |
| ⊕ No. 149 | Wacom | 67(0.02%) |
| ⊕ No. 149 | Infor | 67(0.02%) |
| ⊕ No. 149 | University of Queensland | 67(0.02%) |
| ⊕ No. 152 | KFKI Research Institute | 66(0.02%) |
| ⊕ No. 152 | emlix GmbH | 66(0.02%) |
| ⊕ No. 154 | Coraid | 65(0.02%) |
| ⊕ No. 154 | Microgate | 65(0.02%) |

典型项目-自主研发

- TFS（分布式文件系统）
- Tair（K/V存储）
- OceanBase（分布式数据库）
- Kissy（前端UI库）
- WebX（Java框架）
- Druid（中间件）
- RocketMQ（消息中间件）
- TDDL（分布式数据层）
- Automan（测试框架）

阿里开源的目的

- 回馈社会，阿里用了很多开源软件，也为开源作一些贡献。来而不往非礼也
- 跟业界一起进行技术创新，一起改善软件
- 建立阿里的技术影响力，汇聚更多的人才，迎接更大的技术挑战

阿里开源里程碑事件

- 2009年，章文嵩博士（花名正明）加入阿里，带来开源的种子
 - 章文嵩博士是著名开源软件LVS的作者，开源经验丰富
 - 开源的种子 + 阿里开放的文化土壤
- 2010年，淘蝌蚪平台建立
- 2010年6月份，Tair第一个开源
 - TFS等更多软件在其后陆续开源
- 2011年，成立阿里开源委员会
 - 完善了开源的政策和流程



阿里开源委员会

- 隶属于阿里技术委员会
 - 章文嵩博士为开源委员会主席
- 人员构成（10人）
 - 工程师为主
 - 法务
 - 安全
 - 技术发展
- 协助推动开源的过程
 - 流程 & 辅导

阿里开源协议的选择

- 大部分项目使用GPL
- 一些开源的库用BSD或Apache等
- 非阿里初始项目遵守原项目的license
- 阿里集团是版权持有人, (C) 2007-2014 Alibaba Group Holding Limited

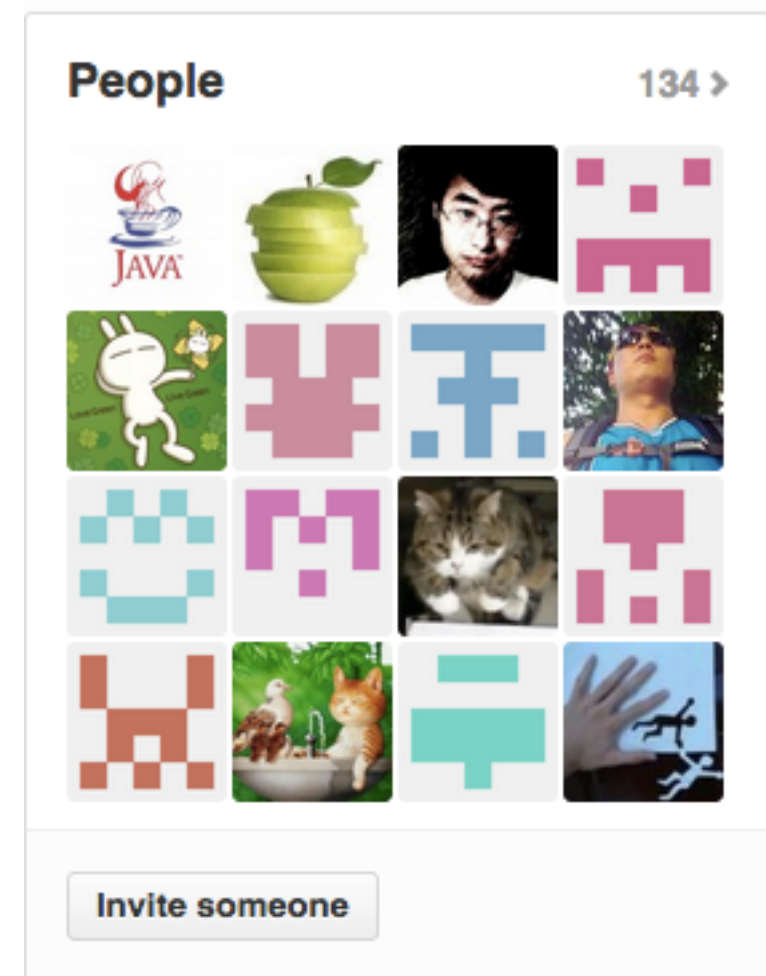
阿里开源平台

- 淘蝌蚪 (code.taobao.org)
 - 开源平台本身开源
 - 国内访问方便
- 现状
 - 现有开放项目10388个
 - 成熟项目以阿里的为主
 - 非阿里项目非常活跃，并出现了一些不错的项目



阿里的GitHub

- 地址: <https://github.com/alibaba>
- 有利于国际化合作
- 更突出工程师的贡献
- 现状
 - 74个项目
 - 134个工程师



阿里的收获

- 促进了开源软件本身的质量提高
 - 用户的意见反馈
 - 用户的bug报告
 - 用户的patch
- 阿里的技术能力和开放精神获得业界的肯定
 - 阿里技术工程师强烈的认同感
 - 吸引更多有能力的工程师加入

阿里开源总结

- 阿里是开源系统的受益者，并积极参与开源生态系统的建设，阿里公司希望以更开放的方式与业界一起进行技术创新
- 采用开源软件与自主开发相结合，有更好的可控性，更高的可扩展性
- 阿里在开源的过程中有付出，更有收获
- 阿里希望更多业内同行一起参与开源，打造一个开放、共赢的中国互联网

联系我

- 邮箱: shudu@taobao.com
- 新浪微博: @淘叔度
- 来往: 叔度



谢谢!

- Q & A