

刘文博

📞 13281101982 · ✉ Yemiliceklichko@outlook.com · 🌐 我的GitHub · 📖 我的博客

👤 个人信息

- 男, 25岁
- 求职意向: Golang / Python / 全栈 研发工程师
- 工作经验: 4年
- 期望薪资: 15 - 20k

👤 自我介绍

- GitHub: <https://github.com/Alexanderklau>
- 技术博客: <https://www.Yemilice.com>
- 旧技术博客: <https://www.cnblogs.com/Yemilice/>
- 是一个开发老手, 入行已经四年+, 主力开发是Golang和Python, 后端开发, Geek气质浓厚。
- 同时也是一名活跃的说唱歌手(网易音乐人, 新说唱500强选手)
- 技术面宽, 涉猎多, 在配合沟通方面有优势, 前后端双持, 移动端也有涉及。
- Golang / Python 双持
- 开源社区活跃分子
- 算法 / 后端开发 / 架构设计
- 4年开发从未停止学习

🎓 教育经历

- 2012 - 2016 成都信息工程大学

💼 工作经历

- 上海德拓信息技术股份有限公司成都分公司, 大数据/存储开发项目组, 全栈工程师
Go/Python/Java, 2017.10~至今
 1. 负责公司存储系统的后台开发维护。
 2. 重构公司部分存储系统的代码, 修复固有 bug。
 3. 重构了以前的C++代码, 利用golang完全重写, 独立负责完成公司项目的文档在线预览和编辑服务。
 4. 学习Ceph, 开发Datrix分布式存储系统。
 5. 做为主力开发, 使用Flask框架从零开发中国电信项目。
 6. 开发企业级网盘系统。
 7. 负责私有云系统开发。
 8. 重构公司数据库, 替换性能较差的旧数据库ETCD, 更新为Elasticsearch数据库, 设计并完成数据库更新换代。
 9. 从零开发项目的日志系统模块, 编写查询算法, 提高了日志查询, 日志检索效率。

10. 独立开发分布式数据同步系统，利用了rabbitmq，elasticsearch，Golang，Python协作开发，实现了数据同步系统的多节点部署，即开即用等复杂功能。
11. 持续性单兵作战，独立开发了数据同步系统，检索服务系统，日志系统，后期完全独立开发项目，完成了这些后端项目的架构设计，接口设计，代码开发，测试编写，打包上线，自动部署等。
12. 将原有的shell脚本部署逻辑完全重写，用ansible进行重构，并且将服务全部ansible模块化进行管理。
13. 设计重构整个公司的日志模块，重构过后的日志模块已经推广到公司的所有产品中使用，已经使用的产品20个以上。
14. 开发对象存储检索服务，利用ElasticSearch进行检索服务器开发，利用Redis和Rabbitmq作为消息的分发和即时处理控件，利用java + apache tika作为提取文字和流媒体的控件，实现了全文检索，模糊检索（类似谷歌的高级检索和百度文科的文字检索）。
15. 利用Golang和Python开发ElasticSearch的SDK，已经推广到自己的GITHUB当中。
16. 和其他同事进行协作，正在将产品微服务化，作为项目主开发人员，进行docker部分的开发调研。

- **深圳前海融数通科技公司，后端研发部门，全栈工程师Go/Python，2016.8~2017.10**

1. 负责公司整个催收系统的功能开发，例如案件筛选，案件分配等。
2. 维护优化并且重构了催收系统部分的代码，并且设计，优化了催收系统的 mysql 数据表。
3. 编写Django的相关接口，技术，并且输出系统和开发文档。
4. 负责编写设计公司产品的逾期算法和贷款策略，并且对旧有算法进行优化处理。
5. 负责公司运营系统部分的功能开发，例如案件导出，案件流转，案件解析等。
6. 负责公司新型后台管理系统的开发，主要包括了案件运营系统，员工管理系统等
7. Django框架维护，开发公司的催收/管理系统

项目经历

- **分布式数据同步系统项目**

使用到的技术

1. rabbitmq分布式消息队列
2. Elasticsearch检索服务
3. Golang的gin框架
4. Etcd分布式数据库

这个项目是我独立开发和设计的，从架构到开发到测试都由我一个人完成，主要的就是实现了一个不同源数据之间的同步，主要功能有

1. 对象存储 - 文件系统的相互同步
2. 文件的完整性检查
3. 任务管理（任务开始，停止）
4. 任务状态（任务进度，任务更新）
5. 任务日志

6. 文件的条件过滤
7. 文件的校验
8. 文件的实时同步 / 定时同步

相比于主流的Rsync，我开发的数据同步系统的**优势**在于

1. 分布式，增强同步效率
2. 文件校验可以进行多种选择，完全同步，增量同步，实时同步等
3. 支持进度的查看，有大致准确的完成时间
4. 有错误文件的日志展示，并且有完整的错误原因
5. 可以进行定时，定时有多种选择，年月日，时间等。
6. 支持文件过滤，可以将不需要的文件直接过滤。
7. 正在开发断点功能，未来可以直接从上次任务停止的部分继续同步

数据同步系统已经通过**中央电视台**招标，被央视项目使用，并且央视将数据同步系统作为同步数据的工具，已经进行了900TB的数据同步。中间还未出现问题。

• 存储系统管理系统开发

使用到的技术

1. Flask框架
2. Python/Rust技术混用开发
3. ETCD分布式数据库
4. Elasticsearch检索服务
5. Filebeat日志转发器
6. Logstash数据收集引擎

这个项目是我和同事们一起开发的管理系统，我主要负责日志系统的架构和研发，将原有的日志存储系统ETCD替换为ElasticSearch。并且我也负责了一部分Flask框架的调优。在这个项目中，我将原有的文件-ETCD数据库架构进行了修改，改为 文件 - Filebeat - elasticsearch数据库 - 前端页面，后续由于数据格式的多样化，我加入了logstash作为数据清洗和数据格式化的工具，这套日志框架 / 逻辑 / 代码，已经被公司20多个项目组共同使用，例如**苏州电视台**，**上海市警察局**等。所以可以保证日志系统的检索效率，也可以保证日志系统的稳定性。

• 对象存储检索服务器开发

使用到的技术

1. Python
2. ETCD分布式数据库
3. Elasticsearch检索服务
4. Java
5. Apache Tika文件读取器
6. Rabbitmq消息队列

这个项目也是由我独立开发设计的，主要是完成了一个对象存储的数据检索功能。首先，编写程序监控对象存储（Python），将对象存储中的数据读出（Java），编写构建Apache Tika的文件读取器，读取文件数据流，然后，传入数据队列（Rabbitmq），通过预先设计好的Es数据结构，从消息队列获

取数据，结构化数据，传输数据到Es中（Golang）。最后继续编写代码（Golang）支持文件的检索，高级检索，关键字检索，特殊字符检索等。

此项目已经通过南京电视台，浙江电视台项目招标，现在已经上线使用，南京台的RGW（对象存储）服务有将近10TB数据，现已全部支持检索 / 全文检索。

- 私人项目：中介雷达

使用到的技术

1. Python爬虫技术，模拟登陆，逆向js等
2. Mysql存储

这个项目是我git库中的私有项目，刚来成都的时候让中介坑了，然后写了一个中介雷达，爬取了58，豆瓣，安居客等等个人房源之类，其中涉及到逆向js，登陆，验证码识别等。通过找到手机号，地址，房屋图片等，根据手机号和图片校验判断房源是否虚假，未来我打算加入评价和中介黑名单功能，爬虫已经写好，前端页面正在开发当中。

- 更多

- 1. 开源 ETCD 的 Golang SDK
 2. 开源 Elasticsearch 的 Golang SDK
 3. 持续性输出博客内容，答疑解惑。
- 更多项目可以移步我的[github](#)或者[技术博客1](#),[技术博客2](#)进行观看。



技能清单

- ★★★ Golang/Python
- ★★★ Rust
- ★★★ ETCD/Elasticsearch
- ★★★ Rabbitmq
- ★★★ JavaScript/Java
- ★☆☆ C++