

简历承载内容有限，期待与您面谈

## 个人信息

---

操润贴(花名:青涯)/男/88年

大专 (如果您过于介意我的学历，勿扰)

工作年限:9年

微博:[卿本自由人](#)

技术博客(多年前技术博客已经长草，刚开辟的新博客，无过多内容): [飞鸟集](#)

公众号:nofootbird

GitHub: [runc BigDataMatrix](#)

期望职位: 架构师(应用和大数据皆可)

期望薪资: 主要看公司

期望地点: 杭州

个人标签和介绍:

- 1、多编程语言患者 (java/python/golang/scala/nodejs皆有涉猎)。
- 2、6年应用架构开发、3年大数据相关从业经验，熟悉并使用相关大数据技术栈，并有一定大数据架构相关经验。
- 3、熟悉分布式爬虫 (web/app) 和数据采集系统的设计和研发，并有大型项目实战经验。
- 4、技术栈全面，有热情，对分布式架构和代码范式，设计模式有一定研究，喜欢研究最新和流行的技术框架和类库。
- 5、有完整的技术产品从0到n的设计研发经验，以及两段 (<20人) 的技术团队建设和管理经验。
- 6、有比较超前的技术前瞻性和产品化意识和思维。

## 工作经历

---

### 杭州玳数科技有限公司

时间: 2016.01~至今

岗位: 技术专家/架构师/产品技术负责人

主要工作内容:

- 1、主导负责企业日志产品设计研发(EasyLog)[<https://log.dtstack.com/easylog.html>]
- 2、主导公司基础组件和框架产品, 服务设计研发
- 3、主导企业分布式数据采集产品(sdk/爬虫)设计开发
- 4、参与企业一栈式数据中台Paas产品设计研发(DtInsight)

## 杭州橙子信息技术有限公司

时间: 2014.06~2016.01

岗位: 技术合伙人/总监

主要工作内容:

- 1、从0到1搭建公司电商系统的设计, 研发(从app, web到后端全栈覆盖)。
- 2、技术选型, 技术规范和框架落地, 协助招聘建团队, rpc服务化订单, 物流, 商品, 支付等子系统。

## 阿里巴巴聚划算

时间: 2011.11~2014.6

岗位: 高级开发工程师

主要工作内容:

- 1、负责聚划算竞拍, 保证金, 订单等资金系统的设计研发工作
- 2、负责商家后台, 运营/营销等后台产品研发

## 最近重要产品和项目经历

### 1、企业分布式日志中心产品(EasyLog)

涵盖编程语言: java/python/golang/react

涵盖技术栈: Django/spring boot/spring cloud/ELK/Jlogstash/ElasticSearch/Kafka/ClickHouse/Flink等等

作为该技术产品技术leader, 带领团队同学多次对该产品从架构到具体实现上了多次迭代重构, 具体演变过程请参见早些时候写的技术沙龙分享ppt[日志中心建设那些事情], 里面主要讲到了几次重大的架构调整和优化, 不管是从技术覆盖面上还是部分技术模块的深度上都是对团队和自己的一个很好的学习和挑战。

更多内容了解详见以下[日志中心建设那些事情]

关于该技术产品更多介绍:

## 2、企业分布式数据采集产品(EasyCrawl/EasyTrace)

涵盖编程语言：python/java/kotlin

涵盖技术栈：scrapy/scrapyd/celery/docker等等

作为数据采集同步产品，公司涵盖的产品技术模块矩阵主要分为EasyCrawl/EasyTrace/EasyDatax三部分。其中EasyCrawl分为外部企业项目版本，内部孵化产品版本，但从产品形态上来讲，做的是同样一件事情，解决外部特别是web数据采集与内部数据整合打通的问题，

其中外部企业项目版本大致的解决方案见[[scrapy企业分布式爬虫采集架构](#)]，目前已经在数个大型电商公司稳定部署运行。

其中内部孵化产品，主要是基于java+大数据技术栈来构建，主要思路是采用web sql+BI将web数据转化为数据库表的思路。

其中用户接口层，为了应对多变的数据格式和用户使用成本上主要是采用sql的思路，

考虑到通用的antlr的解决方案的定制化成本问题，采用了h2的UDF的思路，使用成本更低，

在底层数据采集吞吐和横向可扩展性以及数据存储上采用并借鉴了nutch+hdfs的思路，

在最终价值呈现层面除了整合企业内部数据(ETL)，一个更直接的思路就是BI，除了打通公司内部EasyV可视化产品，也着手打通外部开源BI产品比如superset/metabase等。

关于该产品上更多个人总结：

[漫谈网络数据采集gitchat版本](#)

[漫谈网络数据采集github版本](#)

[scrapy企业分布式爬虫采集架构](#)

## 3、企业数据中台Paas(DtlInsight)

涵盖编程语言：java/golang/scala

涵盖技术栈：vertx/hadoop/flink/spark等等

该产品作为一站式企业数据中台paas产品，覆盖到的大数据技术栈相当广，同时也是一个正在成长的产品。

作为公司的核心产品，虽然不是该产品技术负责人，但作为初创团队有幸见证了该产品从0到1乃至到现在的商业化的全部过程，

熟悉该产品的产品架构和技术架构，并参与了其中部分模块的功能设计和开发。

除了集成目前大数据技术栈的开源产品，更主要是在之上提供了一套完整易用可靠的可视化大数据视窗，让大数据从业者能更好地解决大数据开发和企业大数据存通用的问题，

个人对此一些相关作品和研究：

[大数据相关业余项目](#)

早些时间对于数据中台理解成文 [浅谈数据中台](#)

## 4、分布式电商系统打包总结

覆盖技术栈：rpc/mq/spring(boot/cloud)/service mesh等等

鉴于前几年基本围绕电商系统的研发，设计，架构的工作经历，在此就不一把所有项目经历全部罗列在此了，该期间乃至到现在，完整经历了商品，支付，用户，物流，资金等所有环节的研发设计。

在阿里聚划算，经历了高并发下的分布式系统的架构设计开发(主要是围绕阿里内部技术栈hsf/tair/redis等等)，  
如何优化每一段代码，斟酌复杂的产品业务模型和多变的运营需求下如何做好合适的架构和设计。

在橙子科技，作为公司技术合伙人，从搭建团队，招聘，到后续的产品脑暴，迭代，沙盘演练，到最后技术框架选型，技术规范搭建等。

全程参与并完整经历了一个创业公司如何从0到1的过程，如何遇到问题并解决问题自我认知升级。

其中主导并参与了一个电商产品，从web,wap到app，到服务端的全栈解决方案落地推进，模块拆分，系统服务化改造等。

产品相关

[折疯了](#)

## 汇总作品和产品

[老github](#)

[最近维护github](#)

[商业产品EasyLog](#)