

简介

- 12年多工作经验:
 - 平安银行: Elasticsearch相关开发和集群运维性能调优以及大数据相关开发
 - 万达网络科技 1年4月: 私有云PAAS平台开发和信息流推荐系统开发
 - Autodesk 10年: 早期App开发, 后期后台及搜索业务相关开发
- 擅长Elasticsearch及其生态, 丰富的开发运维和性能调优
- 一定的垂直领域搜索, 推荐系统构架
- 一定的大数据相关开发经验
- 一定的云计算和DevOps相关经验, 基于Docker/Kubernetes的PASS开发, 熟悉AWS各产品

技能

- **语言及框架:** Java(Spring Boot, Play, Swagger), Scala, Python, Bash/Linux, Go, C++
- **数据库及中间件:** Elasticsearch, MySQL, Redis, Couchbase
- **大数据:** Spark, Presto, Hive, Hbase, Zookeeper, Kafka
- **DevOps:** AWS, Docker, Kubernetes, Jenkins, Apache/Tomcat

经历

- **平安银行 - 零售大数据部门大数据基础平台 助理架构师 2018/2 to 现在**
 - 专注Elasticsearch及其生态开发和运维, 主要职责及成果:
 1. Elasticsearch SQL插件开发, 以及插件服务化(ES 6.3版本以前不支持SQL语法查询)
 - 支持标准SQL语法查询ES(性能原因, 不支持多表联合查询)
 - 支持扩展的函数, 方便聚合统计运算
 - 查询语句校验, 缓存及大查询并发控制
 2. ES集群运维和性能调优
 - 单集群规模: 日增2亿文档, 数据1TB, 历史数据20T
 - 解决频繁宕机集群不稳定问题等
 - 大查询性能问题: 20亿文档+字段维度100万上做聚合统计保证集群稳定, 查询时间小于10秒
 3. ES技术评审和支持
 - 接入规范, 数据规模, 资源评审
 - 开发/运维规范及咨询答疑
 - 搭建基于ELK和Kafka, Hive, Presto等大数据技术的日志采集和归档平台
 1. 负责系统架构设计, 平台搭建以及日志采集SDK及相关ETL等脚本开发
 2. 定制日志规范, 和业务接入等
- **万达网络科技 - 信息流和广告技术中心 资深工程师 2017/5 to 2018/1 www.ffan.com**

一款类似于今日头条的APP, 聚合网络中和品质生活相关的文章, 根据用户的兴趣和历史行为, 精准投放内容, 达到千人千面的效果。

 - 主要职责:
 1. 负责内容团队Sprint快速迭代开发, 任务版本分解, 新技术预研及研发流程规范优化。
 2. 信息流个性化推荐系统架构设计和核心模块编码、存储系统设计, 确保业务的高并发高可用。
 3. 规划实施基于Thrift实现跨语言服务, 基于spring cloud的微服务架构设计。
 - 主要成果:
 1. 作为内容团队技术负责人, 带领3人团队开发日均优质文章增量5万+的分布式爬虫系统。
 - 基于Kafka消息分发/HBase存储/webmagic爬虫框架的分布式系统

- 支持模拟人工刷新页面抓取, 支持动态页面爬取等高级反爬虫技术, 支持微信公众号抓取
- 基于Elasticsearch/Kibana的报表统计分析
- 2. 负责信息流系统的整体架构设计, 包括内容入库Pipeline(文档处理, NLP特征标注, simhash去重, 入库), 倒排索引构建, 召回策略(NLP特征, 协同过滤)和排序算法(L2 Rank, L3 Rank)工程化等。
- 3. 基础架构搭建: 线上文章库Elasticsearch集群化, 倒排索引库Redis集群化, 基于ELK的分析报表, 以及基于Elastic Job的分布式任务调度。

技术栈: Java/Python, Spring boot, HBase, Redis, ELK, Kafka, Thrift, 机器学习/NLP处理

◦ **万达网络科技 - 云平台 资深工程师 2016/10 to 2017/5** www.ffan.com

基于Jenkins, Docker和Kubernetes技术, 开发万达私有云控制台&持续集成/部署系统。实现一键式交付, 方便万达快速实现服务和应用的容器化。

- 主要职责及成果:
 1. 作为核心人员, 负责后台框架搭建, 数据库设计, CI/CD工作流交互, API设计, 以及核心的服务创建部署等API开发。
 - 开发基于fabric8的K8s服务部署API, 自动创建、扩容服务
 - 制作docker base镜像给业务使用
 - 业务应用接入和运维
 2. 协助项目scrum master和PM保证后台开发的进度, 及时提供API支持。

技术栈: Java, Spring boot, Docker, Kubernetes, Jenkins, MySQL

◦ **Autodesk - Homestyler/设计家 高级开发 2015 to 2016/10**

基于HTML5和WebGL的家装3D云设计平台(原欧特克和居然之家合作项目, 后卖给居然之家www.shejijia.com)

- 设计和实现基于Elasticsearch的家装3D产品模型的后台搜索服务。支持的功能: 满足前端各种条件的查询和过滤, 产品名和类目名的全文搜索, 自定义权重排名(零售商权重, 品牌权重等), 自动补全, 拼音纠错, 实时索引等。

技术栈: Java/Scala, Play Framework, Elasticsearch, AWS SQS.

- 设计和搭建所有后台的网络基础架构, 以提高服务的可用性, 可靠性和安全性。采取的措施: 基于AWS ELB给每个服务搭集群, 解决单点故障问题, Couchbase数据库搭建集群, Elasticsearch索引数据库搭建集群, 给应用和服务设置正确的AWS VPC安全组。

技术栈: AWS ELB/EC2/SQS, HAProxy, Couchbase, Elasticsearch

- 设计并带领后端团队开发家装施工管理后台服务(欧特克和居然之家合作项目), 提供家装施工过程中的项目/计划/任务/人员/问题追踪相关的数据管理以及消息通知提醒。负责整体架构以及所有的技术方案, 从业务逻辑到底层的微服务框架, 文档工具、数据库选型、到日志以及CI & CD整合等, 方便开发部署和维护。现在该工程作为团队后台微服务的样板工程。

技术栈: Java, Spring boot, Jersey, Swagger, AWS, Couchbase, MySQL

- 设计并带领一个团队实现了一个基于前端AngularJS/后端Play框架的控制台,用以零售商后台人员管理产品模型数据。基于消息驱动的后台支持分布式的模型处理和产品实时索引。

技术栈: Java/Scala, Javascript, AngularJS, Bootstrap CSS, Html5, AWS S3

- 和运维人员紧密合作, 搭建后台的编译和部署系统, 实时解决线上问题, 以及统一整理和实现大量的后台脚本满足各种数据迁移要求。

◦ **Autodesk 其他产品 2007 to 2014**

- **2013 to 2014** Autodesk Tinkercad www.tinkercad.com - 基于浏览器的3D设计和建模的工具, 满足3D打印要求

- 设计和实现了核心的3D mesh处理算法, 包含这些原创算法: 图片内容打印到3D模型的表面, 3D模型表面雕刻字体, mesh切片等
- 实现后台API及产品的部署

- **2011 to 2013** Autodesk 123D Family app www.123dapp.com - Autodesk 123D产品系列的online版本

- 搭建123产品系列的网络基础架构, 设计和实现了CI&CD系统
- 设计和实现若干前端功能模块, 从UI到建模工具

- **2007 to 2010** Autodesk Sketchbook www.sketchbook.com - 欧特克最有名的绘图/图像处理软件

- 实现了若干笔刷和图像处理算法和功能

教育

- 2002 to 2006: 学士学位 华中农业大学 计算机科学与技术专业

