

个人信息	联系方式
<ul style="list-style-type: none">易斌鑫/男/1993学历: 西安电子科技大学/数学系/本科期望职位: 开发工程师期望薪资: 9k-15k	<ul style="list-style-type: none">Tel: 18710849538Email: ybxnice@gmail.comGitHub: https://github.com/two8g现居住地: 西安

工作经历

- 西安博创软件(2014,7月-2017,今)

工作项目

1. K12题库搜索系统(QSS)

负责solr分布式搜索集群建设,进行分词算法优化,包括词典建设,分词算法及LaTeX公式分词优化.将搜索响应从3s提升至200ms,稳定服务C端每日100万PV,文本搜索查全率提升到90%,图像识别搜索差准率提升至70%.与团队分享倒排索引的原理与向量空间模型等技术知识.系统数据量从百万级增长至千万级,从1G索引上升到现在10G.同时,给其它有搜索需求的项目快速提供实现及接口.

2. 自动化部署与分布式架构

首先自建Maven仓库,方便代码共享,加速项目开发,然后部署Jenkins自动部署服务,建服务部署机器自动化,实现应用快速部署.另外,自建GitLab代码仓库,实现代码内部透明与CodeReview;学习git-flow工作流,规划团队协作;并引入issue协作方式,解决责任明确及问题跟踪.并尝试采用GitLab-CI完成自动打包与自动测试.
鉴于团队项目业务复杂度提升,由垂直应用架构向分布式架构升级.我推荐并使用Dubbox分布式RPC框架.和团队学习新的技术并抽取核心业务,完成系统核心服务的分布式升级.

3. 日志收集与监控系统(LOGs)

项目采用storm流式计算实时处理用户日志并加工,后通过flume收集系统,最终存储至hadoop集群.我在此项目搭建了CDH5.7分布式集群,并通过hive进行数据统计分析,为用户推荐项目提供数据服务.该项目解决就日志系统采用关系型数据库存储所面临的容量及利用问题.期间,同时维护6台云Linux主机,目前总数据量已达300G.
使用流行的ELK应用组合,为主要服务提供日志实时监控功能,填补服务访问日志监控空白,随时了解服务运行状态和服务压力,使得尽早发现和解决服务问题.

4. 文档转码自动化(AutoConverter)

项目主要解决复杂HTML内容(包含LaTeX,Markdown等)到文档(DOC,DOCX,PDF,PNG)的自动化格式转换难题.由于我最熟悉LaTeX语法和MathJax渲染工具的后台开发,项目由我全圈负责.最大问题为渲染问题和排版问题.对于问题一,我实验过Webkit和V8渲染引擎,效果不尽理想,而且部分引擎存在浏览器应用先决条件,部分引擎内存和CPU消耗太大.最终采用了Linux下Xvfb虚拟显示方案,解决了网页至图片的转码需求.排版问题包括Office文档及PDF排版美化问题,通过一月的PANDOC转码工具,Office XML和xelatex排版系统的研究,最终达到90+%的满意效果.

技能

英语: CET-6 掌握:Java, Spring, Hibernate, Linux, Redis, Solr, Git, IDEA. 涉猎: Docker, Hadoop, Spark, Hive, DevOps等

自我描述

热爱编程,喜欢独立思考,对新世界充满好奇心的小码农