

基本信息

姓 名	祁超	生 日	1988-12-29
邮 箱	xxqi1229@126.com	电 话	18610961116
学 历	湖北工业大学(本科)	专 业	网络工程
意 向	JAVA 研发	工作地点	武汉

工作经历

时 间：2017-09 ~ 至今

公 司：武汉镜像科技有限公司

部 门：数据中心

工作内容：数据中心 web 系统开发；hadoop 平台维护；

主要技能：Java Web: nodejs、express、SpringMVC、SpringBoot、MyBatis、mysql、mongodb、Bootstrap、css；
大数据：kafka、hadoop、spark、hbase、elasticsearch；

时 间：2014-12 ~ 2017-09

公 司：百纳（武汉）信息技术有限公司

部 门：CommonService

工作内容：运营与客服 web 系统开发；

主要技能：SpringMVC、MyBatis、mysql、hbase、Elasticsearch、angularjs、nodejs、html、css；

时 间：2013-12 ~ 2014-12

公 司：搜狐视频（北京）

部 门：移动开发事业部

工作内容：移动视频 Html5 的功能开发和维护（m.tv.sohu.com）

主要技能：jsp、servlet、redis、mysql、mongodb、hibernate、html、css

时 间：2012-02 ~ 2013-12

公 司：人人网（北京）

部 门：安全中心

工作内容：运营平台(admin.renren.com)和安全中心(safe.renren.com)的开发和维护工作

主要技能：jsp、springMVC、mysql、memcached、thrift、zookeeper、Kafka、elasticsearch

个人技能

熟练掌握：熟练运用 Java 语言，对集合、IO/NIO、并发与锁、线程池、内存模型、GC 原理

与优化等有深入的了解。熟练使用 mysql 和 mongodb 数据库，了解事物和索引优化，有使用 InnoDB 和 Mongo 分片集群处理大数据的经验。熟练使用 SpringMVC、SpringBoot、MyBatis、Hibernate，了解 SpringMVC 的源码。熟练使用 redis 和 elasticsearch。熟练使用 kafka、hbase 等开源工具，阅读过部分源码。熟悉常用的数据结构和算法，熟悉 TCP/IP。

基本掌握：python、JavaScript 编程语言。熟悉 jquery、html、css 等 WEB 开发技术。有使用 nodejs 和 express 搭建 web 前端的经验。有使用 hive、azkaban、spark 和 sparksql 的经验。有 netty 开发经验。

基本了解：scala 编程语言；了解 Spring Cloud。

项目经验

项目名称：数据中心的建立（镜像）

项目描述：数据中心主要包括：数据收集、数据分析、数据展示三部分。数据收集：flume+kafka+hadoop；数据分析：spark+mongodb+elasticsearch；数据展示：bootstrap+echats+nodejs+express+SpringMVC+SpringBoot+mysql+mongodb。

主要职责：

- a) 独自完成了大数据平台的搭建和 web 系统架构搭建；
- b) 基于 kafka 和 mongodb 分片集群的大数据查询系统；mongodb 优化；
- c) 基于 webSocket 的客服聊天系统开发；
- d) spark 任务编写，web 需求开发；

技能列表：nodejs+redis+SpringMVC+SpringBoot+mybatis+mongodb；kafka+hadoop+spark；

项目名称：运营&客服平台的建立（百纳）

项目描述：为了游戏精细化运营，每天会生成多维度报表，给游戏运营和制作人提供数据支持；为客服提供客服后台对用户进行管理。报表用 spark 任务生成数据，结果保存在 mysql 和 hbase；实时需求采用 kafka+SparkStreaming 写 mysql 或者 elasticsearch 供查询使用；前端展示采用 nodejs + angularjs+echats，后端接口调用采用 ajax + http restful api 的形式。基于 netty 构建自己的 http 服务器提供 web 服务(后来升级为 SpringMVC)。

主要职责：

- a) spark 任务编写，报表生成和展示；
- b) 基于 Hive 的大数据处理；
- c) 基于 Innodb 分区表的大数据处理以及 TokuDB 引擎优化；
- d) 基于 elasticsearch 的用户聊天与后台搜索相关需求开发；
- e) 基于 Canal 和 Hbase 的用户画像系统；

技能列表：nodejs、echats、angularjs、netty、SpringMVC、spark、kafka、mysql、ES、Hbase

项目名称：大数据平台 Groot 维护优化（百纳）

项目描述：用户行为日志收集、分析、存储的平台支持。用户行为日志通过 flume 进入数据平台，在 Kafka 聚合后根据不同的需求存储到不同的存储设备，例如：hdfs、es、mysql 等。在 hdfs 上提供 spark 计算框架、hive 数据仓库和 hbase 分布式存储服务。在 mysql 和 es 上提供实时查询服务。

主要职责：a) Groot 系统的维护；

b)基于 SparkStreaming 的分布式分拣程序的升级

c)日志数据校验和自动补漏系统开发;

技能列表: flume、kafka、hadoop、hive、hbase、spark、canal、kylin

项目名称: m.tv.sohu.com 开发维护 (搜狐视频)

项目描述: m.tv.sohu.com 是搜狐视频四大流量入口之一。站点视频日播放量 4500 万, 日 PV 达到 3 亿。是一个多用户高并发的系统。前端采用 CDN 加速, 后端使用 haproxy 和 nginx 做负载均衡, redis 和 ehcache 做系统缓存。后端主要使用 jsp+servlet+redis+ehcache+hibernate+mysql 等。

主要职责: a)新版本需求开发、活动开发、版本发布;

b) nginx 页面静态化与异步刷新, varnish 的 http 加速;

c) redis 集群的升级, 包括集群扩展, 哨兵监控等; 数据拆分。

d)后端 API 优化, 线程池优化, 高并发接口的实现; 压力测试;

技能列表: html、css、jsp、servlet、redis、mysql、varnish、tcpcopy、nginx。

项目名称: 一键删除 & UGC 检索 (人人网)

项目描述: 为了监控站点舆情、广告等非法内容, 需要对用户产生的所有内容进行索引, 根据关键字一键搜索所有的历史记录, 并快速删除。前期采用 lucene 自己构建索引, 开发维护成本大, 索引不稳定。后升级到采用 ES 索引, 系统可持续稳定运行。由于需要同时删除大量数据, 采用分布式调度、异步化删除的方式删除历史记录。

个人职责: a) 负责 Elasticsearch 集群搭建、维护;

b) 索引建立, 采用 DelayQueue 异步增量扫描业务 mysql 表创建索引;

c) 任务处理系统: 基于 Zookeeper 的分布式调度、异步化删除;

技能列表: ES、DelayQueue、MySQL、Zookeeper

项目名称: URL 爬虫 (人人网)

项目描述: 为了增强用户体验, 在人人网上发布的外部链接地址都会转化为 <http://rrurl/xxx> 的短链形式。大量淘宝客到人人网发布推广链接的时候, 严重影响用户体验。爬虫系统会对用户发布的 url 进行爬虫, 找到用户链接的最终目的地, 并判断链接是否非法

个人职责: 主要包括: HTMLParser 页面解析, url 跳转判断

技能列表: HTMLParser、HTTP 协议

项目名称: UGC 过滤系统 & 人人小秘书自动问答系统 (人人网)

项目描述: 人人网各业务 (相册、状态、日志...) 将用户行为日志推送到安全中心 kafka, 消费者实时消费数据, 对用户发布内容和样本库中的样本进行比对, 判断相似度, 如果相似度达到阈值, 执行对应的操作。

我的职责: 全权负责, 主要包括: Lucene 分词、文本相似度 Thrift 服务、用户处理等。

技能列表: kafka、lucene、thrift、mysql