

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	大数据 ETL		

工作经历

时 间：2017-09 ~ 至今

公 司：武汉镜像科技有限公司

部 门：数据平台

工作内容：游戏数据中心的搭建与维护，数据需求的分析与开发

主要技能：flume、kafka、hdfs、spark、yarn、mongodb、hbase、logstash、elasticsearch、nodejs、springMVC

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：游戏运营平台的搭建，游戏大数据平台的维护

主要技能：flume、kafka、Hadoop、spark、hive、hbase、zookeeper、azkaban、canal、sqoop、kylin、netty、angularjs、nodejs

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 m.tv.sohu.com 的需求开发与系统维护

主要技能：java、jsp、servlet、redis、mysql、nginx、resin、mongodb

时 间：2012-02 ~ 2013-12

公 司：人人网（北京）

部 门：安全中心

工作内容：维护人人网运营平台（admin.renren.com）和安全中心（safe.renren.com）的需求开发和维护工作

主要技能：java、memcached、thrift、zookeeper、elasticsearch、spring、mysql、resin、nginx

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+hdfs；数据分析：spark+yarn；数据展示：nodejs+springMVC+mongodb。新接手的海外游戏数据中心主要使用 flume+kafka+logstash+elasticsearch 等技术栈。

主要职责：大数据 ETL 平台搭建与维护、数据分析任务编写、数据展示需求开发。

项目人数：2 人

技能列表：HDFS 技术栈（flume+kafka+hdfs+spark+yarn）；ELK 技术栈（logstash、elasticsearch、kibana）；web 技术栈（nodejs+redis+springMVC+mongodb）

项目名称：运营平台的建立（百纳）

项目描述：为了游戏精细化运营，每天会例行运行 spark 任务，分析不同维度数据，生成报表，提供运营和游戏制作人。Spark 任务采用 azkaban 调度，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，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。

主要职责：运营平台建立，spark 脚本编写，azkaban 调度，mysql 调优，elasticsearch 搭建

项目人数：2 人

技能列表：flume、kafka、hadoop、hive、hbase、spark、canal、kylin、nodejs、angularjs、netty、spark、azkaban、kafka、mysql、elasticsearch

项目名称：m.tv.sohu.com 开发维护（搜狐视频）

项目描述：m.tv.sohu.com 和搜狐视频客户端一样是搜狐视频移动设备的主要入口之一。站点需要提供基本的视频网站的所有功能；最主要的是需要改变用户的使用习惯，拉拢用户，提高搜狐视频的用户量和客户端的下载量。（站点每天的视频播放量（VV）达到 4500W；每天拉起客户端 100W；每天通过站点下载客户端达到 5W；）

主要职责：主要是维护和新版发布的工作；

- 1、nginx 缓存升级，性能调优，varnish 的 http 加速等；
- 2、redis 集群的改造，包括集群扩展，哨兵监控等。
- 3、后端编码，高并发接口的实现；压力测试

技能列表：nginx、resin、jsp、servlet、java、mysql、redis、ehcache、hibernate、varnish、tcpcopy、Executors。

项目名称：一键删除 & UGC 检索（人人网）

项目描述：为了快速获取用户的 UGC，采用 elasticsearch 对用户的 UGC 进行索引，供客服和运营查找非法 UGC，一旦确定，可以快速删除。

个人职责：负责 Elasticsearch 集群搭建、索引建立

技能列表：elasticsearch、DelayQueue、Executors

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：全权负责，主要包括：HTMLParser 页面解析，url 跳转判断

技能列表：HTMLParser、HTTP

项目名称：任务处理系统（人人网）

项目描述：安全中心每天会执行大量的任务（例如：删除用户状态、照片，封禁用户等），为了防止任务突增带来的接口安全隐患，将所有任务执行异步化。将任务封装存储数据库，启动多个 resin 服务，通过 zookeeper 选举 leader，leader 负责分发任务，follower 负责执行任务，每个 resin 服务中的任务执行线程可以通过 zookeeper 动态调控。

个人职责：全权负责，主要包括：多 resin 集群的 leader 管理，动态线程调控等

技能列表：zookeeper、resin、DelayQueue & LinkedBlockingDeque

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

个人技能

熟练掌握：flume，kafka，hdfs 阅读过源代码，对项目中出现的问题可以快速定位；7 年 java 开发经验，熟练掌握 java 集合、IO/NIO、并发、线程池，了解 GC 原理与优化，喜欢 TroubleShoulting；熟悉 mysql，了解 Innodb 索引、事务原理，有使用 Innodb/TokuDB 处理大数据（50 亿）的经验；熟悉 mongodb，了解 wireTiger 引擎，有使用 mongodb 分片集群处理大数据的经验；熟悉 hbase，有使用 hbase+canal 做用户画像的经验；

基本掌握：熟练使用 logstash、elasticsearch、kibana、zookeeper、hive、spark、python、shell；有运维 hadoop 集群及相关组件的经验；了解 servlet、spring、netty

基本了解：nodejs、angularjs、jquery、html、css；有使用 nodejs+express 搭建 web 前端的经验；有使用 nginx+resin 处理高并发请求的经验；