周鹏

。 男, 汉族, 1989年10月生, 现居深圳

。 电话: 138-2653-4430

邮箱: zhoupengblack@icloud.com

教育背景

。 2008.09 - 2012.07 南昌大学, 软件工程专业, 本科

工作经历

。 2013.07 - 2016-07 博雅互动-数据开发技术经理

1.有ETL完整链路相关经验

2.熟练使用,开发过Hadoop,Hive,Hbase,Spark应用

3.有完善的数据展示系统搭建经验

4.具备初步的数据分析,用户画像相关经验

2012.07 - 2013.07 华为-软件开发工程师

1.工作中使用 C/C++开发智能路由业务软件

2.CI(持续集成)项目负责人,保证CI过程自动完整的执行

3.问题定位排查

项目经验

○ 2016.02 — 2016.07 数据展示系统前端开发

JavaScript,AngularJS,CSS

技术: HTML + JavaScript + AngularJS

背景:之前前端开发都是通过后端模板渲染,开发效率低,并且和后端耦合很大。

导致前端开发效率严重影响了整体开发效率,因此开始了前后端分离方案

过程:

1.使用AngularJS作为前端开发框架,极大的提高了开发效率

2.使用Grunt构建,降低了前端部署的复杂度

3.页面展示选用Highcharts,极大的丰富了页面展现形式

。 2015.10 - 2016.03 数据ETL方案的讨论与设计

Heka,Flume,Kafka,Hadoop

数据ETL方案采用 Heka + Flume + Kafka + Hadoop 方案

1.Heka负责数据的收集和处理

2.Flume负责一个区的日志汇总,提供缓存防止日志丢失

3.Kafka作为消息系统提供日志的接收和消费

4.Hadoop 负责日志的存储和计算

2015.11 - 2016.01 游戏渠道注册、活跃、收入实时计算

Scala,Spark

技术: Spark + Scala

背景:博雅游戏合作的渠道有几百个,这些渠道商包括我们自己都希望实时看到游戏推广的

效果。这样方便对渠道做正确的评估

过程:

1.主要使用 Spark Streaming从Kafka中读取数据并进行各种计算,最后将结果存入

PostgreSQL数据库中

2.如果是需要去重的数据则会用到 Bloom Filter进行去重

2015.07 - 2015.10 用户宽表和用户画像开发

Hbase.Impala

技术: Hbase + Impala

背景: 我们只有离线统计的数据, 缺乏单个用户的属性信息

过程:

1.将用户唯一标识ID作为Hbase的key, value存放用户的各种属性和一些状态

2.这就建立了一个基本的用户画像,后续可以进行更多元的业务开发

○ 2014.11 - 2015.06 Oracle数据库到Hadoop的迁移

Hadoop, Hive, Sqoop, PostgreSQL

技术: Hadoop + Hive + Sqoop + PostgreSQL

背景: 随着数据量的增大, 我们决定从Oracle数据库迁移到Hadoop集群



过程:

- 1.将Oracle的存储过程全部替换为基于Python的Hive统计脚本
- 2.自研发调度系统运行每天离线统计的Hive统计脚本
- 3.将Hive中的统计结果通过Sqoop同步到PostgreSQL数据库
- 4.使用工具对比Oracle和PostgreSQL中数据的差异

。 2013.07 - 2014.10 数据展示系统后端开发

Python, Flask, Tornado, Redis

技术: Python + Flask + Tornado + Redis

背景:对统计数据提供RESTful APIs供前端页面调用,展示成各种图表

过程:

- 1.主要系统采用Flask作为Web framework 提供API接口
- 2.部分系统(高并发)使用Tornado作为Web framework提供API接口
- 3.采用Redis作为缓存层,供页面访问
- 4.系统部署采用: supervisor + gunicorn + gevent + Flask

作品展示

- 。 2015年读的书
- 。 2016年读的书

自我描述

- 。 高效的持续学习者
- 。 爱好阅读,相信阅读可以改变人的思想
- 。 管理小心得: 以结果为导向, 发自内心关心他人的成长
- 。 有良好的逻辑思维和表达能力