



伍俊波

15828256946 wujunbo1111@yeah.net
mobile email

个人介绍

熟悉etl流程，数据仓库知识，pentaho kettle ETL工具,数据库开发，可以独立完成数据etl处理；
熟悉python,mysql,hive ,MapReduce以及优化； 熟练使用hadoop hive系统，能独立搭建hadoop环境；
熟悉linux shell命令，了解kettle日常作业的部署和调度，可以搭建IBM的整套BI环境，包括PCI和PMQ以及单独的cognos环境；
熟悉sql编程,熟悉正则表达式,excel, notepad++。
乐观自信，有很好人沟通能力，较强的责任感和义务感，注重团队协作，有较强抗压能力，学习能力强。

职业技能

ETL	熟练
Python	掌握
Mysql	熟练
Hive	熟练
sql	精通

教育背景

2008.09 - 2011.09 成都工业学院（专科）



工作经历

2011.10 至今 广州天懋信息系统股份有限公司（全职）
数据工程师

项目信息

- 2017.08 至今 大数据平台LUDP EDW层建设
- 1.根据源系统在数据仓库建立对应的模型
 - 2.ETL Job开发，测试，上线，运维，优化。
 - 3.通过Python脚本提高ETL开发过程中的工作效率
 - 4.迁移优化历史ETL Job到新的开发平台
 - 5.目前开发了EDW层400+个 ETL JOB以及相关的表
 - 开发工具：Hadoop hive，python,Leap_process,Talend,sql,notepad++ ,excel

- 2016.12 - 2017.03 **抓取二手闲鱼网站的便宜货**
- 通过python 开发的爬虫，抓取淘宝闲鱼网站上我想购买的产品，如果价格在我指定的范围之内，就会发一封邮件给我，提示我购买，我只需要在我的函数里面加上商品型号，价格区间，就可以让整个系统运行起来，并实现翻页和过滤
 - 开发工具:python,linux,crontab
 - 用到的库：os,sys,BeautifulSoup,smtplib,re,urllib2
- 2015.04 - 2017.07 **各类业务线报表**
- ETL项目的上线，部署，调度，测试，后期的维护
 - 利用ETL工具kettle对hive,impala,mysql,mongoDB,kudu的数据源的抽取，整合，加载到报表数据库mysql中或者通过邮件发送给各业务方的相关的同学,以及每个项目的上线，部署，测试，后期的维护，业务变更
 - 开发工具:kettle,sql,hive,spark,mysql,mongoDB,kudu,linux,正则表达式
- 2015.04 - 2017.07 **负责全线的业务数据统计任务**
- 负责公司产品/市场运营/财务/物流/管理层/战略部/法务部/客服部等部门的数据统计分析任务
 - 开发工具：sql,excel,notepad++ ,正则表达式
- 2014.08 - 2014.12 **报表数据补跑脚本**
- 按照需要开发报表后，往往需要补历史数据，这个脚本主要使用python来实现，执行时传入时间参数，自动完成按天维度的或小时维度的报表历史数据补跑。
 - 开发工具:python,linux
 - 用到的库：datetime,os,sys
- 2013.01 - 2014.06 **深圳物流中心办公仓储网络系统**
- 1.配合研发部门进行mysql优化，负责测试服务器和开发服务器的日常维护和故障处理。
 - 2.维护和优化talend EBS搭建的ETL数据仓库，保证日志系统，商品系统，仓储系统的正常流转。配合业务部门优化和搭建Bi系统。
 - 3.物流中心仓储中心和办公区的有线和无线网络环境搭建，以及网络设备，服务器安装调试
 - 安装调试设备：10台无线AP，一台无线控制器，一台文件共享服务器，3台思科交换机，一台路由器
- 2012.01 - 2013.03 **连锁便利店数据分析**
- cognos报表开发。echarts图表开发。数据跟新和维护。
 - 为该便利连锁开发可视化报表，支持运营决策。
- 2011.10 - 2011.12 **磁盘重复文件查找器**
- 用python开发了一个exe文件小工具，实现对当前目录下扫描文件夹以及子文件夹下的重复的文件并生成一份日志文件，报告结果
 - 开发工具：python
 - 用到的库：os，hashlib，time