



万昆

ID:72311412



目前正在找工作



15618758658



wankunde@163.com



男 29 岁 (1988/06/08) 现居住上海-松江区 8 年工作经验

最近工作

职位: 高级大数据工程师
公司: 平安证券
行业: 金融/投资/证券

最高学历/学位

专业: 计算机科学与技术
学校: 南京工程学院
学历/学位: 本科

求职意向

个人标签: HADOOP HIVE HBASE SPARK 源码

期望薪资: 45-55 万元/年

地点: 上海

职能/职位: 高级软件工程师 系统架构设计师 大数据平台工程师

行业: 计算机软件 网络游戏 互联网/电子商务

到岗时间: 1 个月内

工作类型: 全职

自我评价:

1. 对 Hadoop 生态圈内各个软件都比较熟悉;
2. 有快速学习的能力, 快速跟踪开源软件的最新性能, 功能等;
3. 熟悉 Hive, Hadoop 源码, 遇到问题可以到源码中跟踪, 分析问题;
4. 有过 Yarn,Hive Patch 经验.
5. github : <https://github.com/wankunde>
6. Blog: <http://blog.csdn.net/wankunde>

职位期望:

拥有 200 台以上大数据集群的平台。

工作经验

2015/4-至今 高级大数据工程师|经济事业部
平安证券 [2 年 9 个月]
金融/投资/证券|5000-10000 人|民营企业

工作描述: 负责公司大数据平台维护, 保证大数据平台的平稳运行;
开发平台的基础功能, 包括 HIVE UDF 等;

跟踪测试最新的开源大数据功能组件，并尝试结合部分业务进行实施应用；
开发 spark, hive, 程序处理等相关程序。

2014/4-2015/4 **大数据开发工程师**|数据中心
巨人网络科技有限公司 [1 年]
互联网/电子商务|5000-10000 人|民营企业

- 工作描述:
1. 研究最新云计算相关技术;
 2. 参与数据部项目架构设计, 根据业务特点, 进行技术选型, 性能测试等;
 3. 参与相关项目开发, 完成技术难点和风险点攻关;
 4. 使用 SPARK, Hadoop Yarn, Impala, Hive 等框架, 完成项目开发, 技术与业务整合落地。

2013/2-2014/4 **软件工程师**|集成服务部
亚信联创有限公司 [1 年 2 个月]
计算机软件|10000 人以上|民营企业

- 工作描述:
- 负责云计算研发, 研究 Storm, hadoop 等相关新技术, 对大数据相关技术进行研究和方案实施。
 - 熟悉 HDFS, MapReduce 内部实现原理, 读过 Hive 源码及其中的编译原理。
 - 研发了 Hive, Hadoop 任务信息统计, Hadoop 资源使用计费系统。
 - 熟悉 JVM 模型及性能调优知识。

2013 年获得优秀个人奖

2010/7-2013/2 **软件工程师**|产品中心
四方精创咨询有限公司 [2 年 7 个月]
计算机软件|500-1000 人|民营企业

- 工作描述:
- 负责完成公司产品功能开发, 对于 J2EE 开发有过大型项目经验。
 - 熟悉并能熟练使用各种 Java 开源框架。
 - 熟悉 Oracle 相关开发, 有过大批量数据处理的实际开发经验。

项目经验

2017/1-至今 **大数据平台维护**

所属公司: 平安证券

项目描述: 大数据平台监控, 维护和调优, 部分 HIVE SQL 优化后, 性能提升 90%, 其余为平台的 BUG 修复等。

HIVE-18362 Introduce a parameter to control the max row number for map join conversion

HIVE-12710 add better logging for Tez session creation thread failures

HIVE-14690 Query fail when hive.exec.parallel=true, with conflicting session dir

HIVE-11969 start Tez session in background when starting CLI

HIVE-18055 cache like pattern object using map object in like function

HIVE-11013 : MiniTez tez_join_hash test on the branch fails with NPE (initializeOp not called?) (Sergey Shelukhin, reviewed by Vikram Dixit K)

HIVE-12563 NullPointerException with 3-way Tez merge join

HIVE-11583 When PTF is used over a large partitions result could be corrupted

HIVE-11016 MiniTez mergejoin test fails with Tez input error (issue in merge join under certain conditions)

HIVE-11837 comments do not support unicode characters well.

HIVE-10542 Full outer joins in tez produce incorrect results in certain cases

HIVE-12384 Union Operator may produce incorrect result on TEZ

HIVE-12207 Query fails when non-ascii characters are used in string literals

YARN-3933 FairScheduler: Multiple calls to completedContainer are not safe

YARN-3619 ContainerMetrics unregisters during getMetrics and leads to ConcurrentModificationException

HDFS-7765 FSOutputSummer throwing ArrayIndexOutOfBoundsException

2016/6-至今

指标系统后台 Hive 计算处理

项目描述: 根据 Hive 中 Fact Table 数据和统计维度计算指标的最大值, 最小值, 平均值, 累加值, 以及相应记录在发生时的相关属性。

责任描述: 负责 UDF 函数开发, 后台 HIVE 数据处理, 任务性能调优。

2014/9-2015/2

分布式日志实时分析系统

项目描述: Kafka + SparkStreaming + SparkSQL + Oracle

使用 kafka 实时收集日志文件内容;

使用 Spark Streaming 从 Kafka 中读出日志文件内容, 抽象 DStream 后在使用 Spark SQL 抽象为 DafaFrame 进行 SQL 计算, 最后将计算结果写入 HBase 系统, 共前端系统查询。

GIT: <https://github.com/wankunde/logcount>

责任描述: 1. 负责系统架构;
2. 负责 Kafka, Spark, Hadoop 等集群软件的安装和部署;
3. 负责系统的编码和开发工作;

2013/2-2013/7

流量经营系统

项目描述: ue(开发工具)

Linux 集群(硬件环境)

hadoop+hive+tcl(软件环境)为中国移动完成网络流量的分析, 并生成多样报表数据。

责任描述： 主要研究 Hive 的新功能，对生产环境中对 Hive 进行性能调优。
完成 hadoop CDH 软件版本的安装部署和 HA 方案的实施。
完成 hadoop map/reduce 相关的编程开发。
完成 hive 的 UDF 函数的编写以满足项目开发需求。
完成系统性能调优，通过多种方式加快 hadoop 集群的数据处理速度。

2011/6-2013/2 **业务监控分析系统**

项目描述： UE +XShell(开发工具)
 AIX(硬件环境)
 ProC+Oracle +gcc(软件环境)完成对每日银行交易流水等数据的清洗，根据业务需求，对清洗后的数据进行数据分析和采集并产生风险预警数据，使用 J2EE 平台完成风险数据的管理和后期数据分析。

责任描述： 根据项目需求完成后台数据库的构架，对数据处理进行调优。
 根据项目需求，完成对各类银行交易风险模型的风险数据采集。
 完成后台批量程序的调度，性能优化。
 完成银行数据基础表数据的导入和基础数据处理。
 此项目要求熟练使用 oracle 开发技术，熟练使用 oracle 报表统计分析函数。

教育经历

2007/9-2011/6 **南京工程学院**

本科|计算机科学与技术

专业描述： 从事计算机基础知识学习，包括 C，Java，数据结构，操作系统等

技能特长 （包括 IT 技能、语言能力、证书、成绩、培训经历）

技能/语言

英语良好

证书

2015/5 **高级工程师** （系统分析师）

2008/1 **大学英语四级**