

简历

熊伟

应聘岗位：大数据开发相关岗位

工作经验：5 年

联系电话：18339955330

邮箱：xiongv.ever@outlook.com

毕业院校：中原工学院

个人信息

姓 名：熊伟

工作经验：5 年

目前状况：在职

专 业：机械电子工程

年龄：27

学历：本科

到岗时间：一个月

个人技能

- 掌握 Java 和 Scala 编程，熟悉使用 Python 开发；
- 掌握 Hadoop 架构，熟悉 HDFS 的基本操作与原理，熟悉 Yarn 的任务分配调度流程；
- 掌握 HiveSQL 的编写，及数据分层、建模、优化等；
- 掌握数据仓库模型建设；
- 熟悉 Flink 工作流程以及了解运行原理；能够使用 Flink 进行流式计算；
- 熟悉 JVM 内存区域划分及垃圾回收过程；
- 掌握 Kafka 相关原理，使用 Kafka 构建实时数据流管道，与 Flink 实时数据流整合；
- 掌握 Zookeeper 的运行机制，能够搭建 zk 集群，了解基于 Zookeeper 的特性的典型应用；
- 掌握 Flume 日志数据收集框架，根据业务需求对配置文件进行相应的修改，满足业务需求；
- 掌握使用 Sqoop/DataX 对数据和数据库之间进行出导入导出；
- 了解 Flink 的水位线和背压机制，使用过 Flink CEP 进行代码开发；
- 了解使用 Flink SQL
- 了解 Druid, Kylin 等 OLAP 分析引擎；

工作经历

| 在职时间 | 所属行业 | 公司名称 | 担任职位 |
|------------------|------|-----------------|--------------|
| 2017.06--2018.04 | 电子商务 | 郑州万国优品保税进出口有限公司 | 大数据开发工程师[实习] |
| 2018.07--2020.02 | 智能硬件 | 宁波舜宇光学科技有限公司 | 大数据开发工程师 |
| 2020.05--至今 | 手机通讯 | 华勤技术股份有限公司 | IT 开发工程师 |

项目经历

2023.2-至今

项目名称：SSO 实时数据分发

项目描述：

SSO 是华勤单点登陆系统，基于上游 HR 系统的性能考虑，由 SSO 额外承担用户、组织、岗位数据下发功能。目前使用的是离线方案：每日增量分发数据给下游系统，两小时下发一次，凌晨下发全量数据。

离线方案目前已无法满足下游系统越来越高的实时要求，且数据拉取过程中，上游 HR 系统性能开销较大，在上游 HR 系统要求和下游场景驱动下，改为获取 Oracle 数据库的变更日志，Flink 实时补全变更数据的维度信息，并下发给下游系统。

软件架构：DataWorks + E-MapReduce + OSS + Replicate + Flink + Hologres + Kafka + MySQL

责任描述：

数据探查：探查 HR 系统数据质量、业务逻辑，与上游探讨一些改造方案的可行性；

规则探讨：由于不可能所有字段都做到实时，因此和下游探讨并确认离线字段清单和实时字段清单；

平台运维：负责阿里云平台的运维保障，配置调整，成功解决过一次集群内存泄漏问题，并给出预警办法；

实时数据采集：联系 Oracle DBA 开通需要的权限，分别测试了阿里云实时采集和 Replicate 实时采集，最终确认使用 Replicate；

实时方案开发：由于 HR 表结构设计，需要所有 state 不可以存在 TTL，且为了防止占用内存一直增长，设计了两套方案用以解决 state 问题；

接入下发规则：不同系统需要的数据密级不同，因此为每个系统制定一套规则，用于定义数据是否下发、下发字段清单、特定字段加密，特定行特定字段置空等规则；

项目经历

2022.2-2022-6

项目名称：华勤 HR BI 离线项目

项目描述

为了聚焦于业务部门较为关注的指标，由系统取代人工进行可视化报表的输出，供管理者查阅，提升数据输出的及时性、准确性，提升业务人员的工作效率。启动 HRBI 项目，主要解决以下痛点：

- 1、解决执行层：减少日常重复性人力加工报表数据的成本，提高数据时效性和准确性。
- 2、解决管理层：直观、即时、高效的获取数据和分析，更好的辅助管理者决策。

3、通过推进数据应用场景的不断深化，反推业务和流程中存在的问题，促进业务流程和系统的不断迭代优化。

个人主要负责员工团队以及异动数据模块的仓库模型构建，并输出 ADS 指标供帆软端调用。同时根据业务需求，将历史未上系统数据（202008 以前）从线下 Excel 报表抽取并加载至大数据平台，以相同逻辑输出 ADS 指标。

同时负责输出员工团队信息模块的数据质量，以及 HR 项目在 Hive 端的报错问题。

软件架构：DataWorks + E-MapReduce + OSS + Ranger + ADB PostgreSQL + FineReport

责任描述：

指标开发：根据业务需求建模，构建从 ODS 到 ADS 一整套开发流程并输出指标，主要涉及员工团队领域和组织发展领域；

历史数据补充：由于 HR 系统 2018 年以前线上数据丢失，按照业务提供的线下 excel 数据生成员工信息数据，并根据员工信息数据创建拉链表生成员工异动数据。最后旧拉链表和新拉链表合并；

需求沟通：同业务人员沟通指标逻辑，确认需求可行性；

数据质量：检测 HR 系统原始数据质量，检测结果输出看板晾晒；

数据平台运维：解决 HR 项目在 Hive 端遇到的问题；

项目经历

2020.8-至今

项目名称：华勤质量云中枢

项目描述

为对公司生产、售后过程中产品质量问题进行监控、可视化展示供管理层决策并对制造部门考核而发起。从业务系统中抽取生产、检验、售后信息上云，以 ERP 主数据作为维度信息实现分析，分析结果传入关系型数据库供帆软调用；

主要负责公司质量域的开发工作，由于质量部门没有自己的业务系统，所以需要与业务沟通数据来源，并进行数据抽取、探查等工作。完成模型 ODS 到 ADS 数据开发，其中字段翻译需要遵守公司词素库，模型在开发并评审完成后需要上传公司 DDM 模型库中，并更新指标字典；

同时担任公司大数据平台运维工作，针对慢 SQL 任务会进行专项优化调整，同时对开发同事的一些 SQL 报错（非语法报错）找出解决办法；

维护部分公共域主数据表，并对业务系统提供数据接口服务；

软件架构：DataWorks+EMR+OSS+Ranger+ADB PG+Serverless+FineReport

责任描述：

模型开发：根据业务方需求，抽取业务系统数据上云，确定业务逻辑，实现从 ODS、CDM 到 ADS 以及 ADB PG 一整套模型开发；

数据分析：创建生产、检验、售后信息事实表，根据提供指标进行相关分析，满足业务需求；

平台运维：负责阿里云平台的运维保障，配置调整，成功解决过一次集群内存泄漏问题，并给出预警办法；

任务调优：对项目重点任务调优，帮助解决部分 Hive 任务频繁 OOM、以及运行时间过长问题；

平台开发：阿里云平台二次开发，对 Hive 任务产生的 Yarn 任务进行分析，如分析任务时长、或执行频率等；

临时任务：根据要求开发 Python 脚本；

项目经历

2019.5-2020.2

项目名称：舜宇工业大数据项目[实时]

项目描述：

舜宇工业大数据项目是舜宇光电为提升产品良率，采集工业生产各个环节实时信息，实时分析处理后发送反馈信息，通过收集原料、机台、操作日志信息等，分析处理后及时得到实时生产信息，帮助生产部门领导及现场技术人员得到各环节良率信息，便于在良率降低时及时响应，快速定位问题，降低企业生产成本；

软件架构：MySQL+Flink+Kafka+Canal+Redis+HBase+Phoenix+Druid+Flume

责任描述：

数据同步：对接公司 MES 系统，产品状态信息存储于 MySQL，Canal 获取 binlog 信息下沉到 Kafka 的 ods 层；存于 MySQL 中的维度信息如果发生更改，由 Canal 发送到 Flink，Flink 处理后同步修改 Redis 中的维度数据

数据处理：主要利用 Flink 程序消费 Kafka 中 ods 层字段，对数据进行过滤拉宽，处理完成的数据向 Kafka 的 dw 层和 HBase 各发送一份；

数据分析：Druid 预处理 Kafka 中 dw 层数据，保存计算后数据；

项目经历

2018.7-2019.4

项目名称：舜宇工业大数据项目[离线]

项目描述：

为分析深入分析良率影响因素，搭建离线数仓分析数据，离线分析有助于部门领导及现场技术人员通过查看某一周期的生产数据，以直观的统计数据查找到对于生产影响较大的信息，如机台 Mac、供应商来料、不良集中工序等；

软件架构：MySQL+Hadoop+Hive+SparkSQL +Kylin+HBase+DataX

责任描述：

- 数据同步：使用 DataX 同步 MySQL 数据到 Hive 的 ods 层，供后续数据处理；
Flume 获取生产日志信息，同步到 HDFS，再使用 SparkSQL 将 HDFS 数据同步到 hive 的 ods 层；
- 数据处理：使用 SparkSQL 进行数据 ETL 操作；
Spark 范围分区解决数据倾斜问题；
- 数据分析：Kylin 预计算作为数据仓库的 ads 层；

自我评价

- 1、有很强的团队合作意识，能较好的处理人际关系；
- 2、有较强的适应能力和接触新事物的能力，学习能力较强；
- 3、能很快适应新的工作环境,融入团队中去，善于学习新的知识；
- 4、乐于与用户以及同事和领导沟通，以便快速解决项目遇到的问题；
- 5、有一定的抗压能力；