

■ 基本信息

姓 名: 俞景治

毕业学校: 西南交通大学

专 业: 信息管理与信息系统

英语水平: CET6

工作经历:

2019.9~至今 数澜科技有限公司

2019.5~2019.9 华为技术有限公司

2017.7~2019.5 数梦工场科技有限公司

2015.07~2017.07 新华三集团

求职岗位: Java软件工程师

联系方式: 842387737@qq.com Tel: 137-7758-1812



■ 个人技能

主攻语言: Java

技术框架: Spring(boot+MVC) + MyBatis + ActiveMQ + MySQL

数 据 库: 熟悉 MySQL 数据库及调优, 精通图数据库 JanusGraph 源码及调优, 熟悉搜索引擎 Elasticsearch 使用

其 他: 熟悉 Redis 开发及源码, 熟悉 nginx 开发; 可使用shell及python做脚本开发

■ 项目经历

1. 新华三集团-高端路由器

技能概括: 熟悉Linux下的网络协议栈的驱动开发; 熟悉Linux内核网络原理

2. 数梦工场-大数据产品线-数据中台产品

工作内容:

Java 软件开发。基于云计算平台, 进行数仓建设的元数据管理; 元数据管理包括元数据的建模、实体化、同步, 并基于底层图数据库相关技术, 对上层提供个性化数据服务管理

图数据库同步及性能优化:

概述: 数仓平台同步云计算平台表数据 (10w 表) 用时 10 小时, 且同步后查询效率低

项目职责: 定位 JVM 内存泄露及分析性能瓶颈; 使用 nginx 反向代理技术实现读写分离, 使用 zookeeper 实现租户级别的数据同步; 研究图数据库源码, 优化查询效率

项目成果: 优化后同步用时 2 小时, 优化后查询速度 ms 级

技术点: Atlas + JanusGraph + nginx+ zookeeper

数据服务开发: 数据血缘、大盘及目录服务

概述: 提供多叉目录树服务, 提供数据资产高级查询服务; 提供数据大盘指标服务, 包括总量、排行、任务等指标; 提供数据血缘服务。

项目实现: 通过 ActiveMQ 解耦上层模块调用底层基础服务, 通过 Redis 加速数据血缘等热点数据查询

技术点: Redis+ActiveMQ+线程池

调研表导入导出优化：

将大项目空间表字段模型按过滤条件进行批量导出，将调研数据按照调研模板导入并支持指定数据目录导入

项目实现：使用分布式事务同步模块数据、元数据、目录数据，确保导入导出操作的原子性

技术点：zookeeper + ActiveMQ + 线程池

3. 数澜科技有限公司-数据中台

Spark SQL字段级血缘解析

工作成果

实现Spark SQL字段级血缘解析及产品服务化，可实现insert /create等语法的解析

概述

在没有现成解决方案的大背景下，调研Spark SQL字段级血缘解析功能，在Spark SQL逻辑计划基础之上，根据逻辑语法树抽取字段血缘关键信息，实现Spark SQL高级复杂语句的解析功能；其中字段级表级血缘在理论上准确率可达到100%

数据中台用户及权限管理

工作成果

实现用户管理（ldap）、权限管理（ranger,sentry）功能在不同集群上的方案设计、代码实现

概述

通过插件及spi方式适配底层不同厂商hadoop集群的权限功能，在数据中台侧达成统一的用户及权限管理。

技术架构：

Spring boot + Dubbo + MyBatis + MySQL

Ambari 产品化

工作成果

通过阅读调研ambari产品部署上的可能性，并解决了ambari通过tarball部署的可能性，并完成了hadoop，zookeeper，hive部署的编码。

概述

阅读ambari开源代码，总结ambari技术架构，分析ambari部署框架及原理；摒弃rpm安装方式，采用tarball部署hadoop组件方案可行性分析；完成hadoop部分组件的安装部署工作。