

# 郝有强 | 高级数据工程师

☎ 182-1070-5367 • ✉ windv587@gmail.com • 🌐 wind4869

## 工作经历

秒针系统  
后台研发工程师

数据平台组  
2016.07 – 2018.05

## 教育背景

哈尔滨工业大学  
硕士

计算机技术  
2014.09 – 2016.07

哈尔滨工业大学  
本科

计算机科学与技术  
2010.09 – 2014.06

## 专业技能

编程语言: 熟练使用 C/C++, 熟悉 Java、Python、Golang、Shell 等编程语言  
操作系统: 熟悉 Linux 开发环境, 使用 CentOS 和 OS X 进行开发, 熟悉服务端网络编程  
软件工具: 熟练使用 Vim、Git、GDB, 熟悉 Nginx、MySQL、Redis 等常用软件工具  
其他技能: 熟悉 Bitcoin、Ethereum 原理, 了解 Solidity 语言以及智能合约开发

## 外语能力

英语 CET6

流畅阅读英文文档

## 项目经验

- 在线广告投放服务的开发维护
  - 基于 Lighttpd 服务器的 C/C++ 项目, 通过开发插件实现业务需求。主要负责了对接今日头条、易车网等媒体的开发以及对接外部 DMP 的性能优化工作, 并维护了一个 Golang 实时报表。
- 高性能 API 服务的开发
  - 基于 Lighttpd 开发了 DMP 标签 API 服务以及反作弊 API 服务, 同时负责维护后端的 Redis Cluster 数据存储服务, 期间负责改进了 Redis 储存方案, 压缩了数据, 最终节省了近一半的内存。
- 基于 Druid 的实时查询服务的开发
  - 为了支持实时的数据报表服务, 团队引入了 Druid Cluster。主要负责了生成数据的 MapReduce Job 的开发、数据加载以及基于 Spring 的查询 API 的开发, 支撑了两个重要的实时数据报表服务。
- 数据 ETL 通用框架的开发
  - 为了同时支持多种数据源的 ETL, 参与开发了一个可配置的 ETL 框架, 并主导了在公司内部的两大监测数据源上的应用, 目前公司内部已经有近 10 种数据源的 ETL 程序是基于该框架开发的。
- 基于 Airflow 的工作流开发
  - 在团队内部推广了工作流管理工具 Airflow, 可以使用 Python 定义工作流, 将 crontab 中的任务都转移到该工具上统一管理, 也基于该工具开发了内部 ETL 的工作流及相应的 Airflow 定制化插件。