

## 优势亮点

- 1.具有扎实的专业知识，技术全面。
  - 2.具有多年的软件设计和研发经验，熟练掌握Node JS、c/c++, python等前后端编程语言。
  3. 熟练使用TensorFlow, Caffe, PyTorch等深度学习开发框架，并评估硬件平台和算法的整体性能。
  - 4.对常见的CNN,YOLO,FCN等检测和分割网络有实际经验;
  - 5.良好的系统分析能力和架构经验,具有多年大型分布式系统及AI场景调度和优化经验，熟练 k8s, docker等虚拟化技术，并在k8s, kubeflow等社区有所贡献。
  - 6.为人乐观，开朗，能够承受较大的工作压力，文字功底强。
  - 7.善于沟通，注重团队合作。
- 
- 

## 工作经历

### 阿里巴巴及蚂蚁集团

2022/08-至今

**技术专家** 技术栈：Node JS、c/c++, python

职责业绩：1. 负责蚂蚁集团 AI infra 基础平台架构的 GPU 集群调度、容量管理及 GPU 虚拟化产品建设等

2. 负责蚂蚁集团 GPU 国产化的产品适配，可行性、兼容性评估，以及相应 dw 工具链对接与集成。

3. 在 GPU 集群的调度策略上，设计并实现 GBank 的基于银行模式的 GPU 容量管理模式，使得资源的分配率提升20%，整体利用率提升8%

4. 主导并推动 GPU 资源池化，并加入 vGPU 虚拟化资源池的能力，在资源混部，在线调试等任务场景下都能极大的提升资源使用。

**平安科技** 2018/08-2022/09

### 技术专家

下属人数：10

职责业绩：平安科技深度学习平台负责人，负责平台建设，技术实施和项目落地等工作。

1. 参与研发面向AutoML的深度学习框架，支持AI领域建模的定制化和深度调优，加快AI模型建立的速度和质量。针对模型压缩，量化剪枝进行优化。
2. 主导基于Docker, k8s, hadoop yarn的大规模深度学习模型训练平台的搭建及优化，支撑数千异构GPU规模的调度和训练任务。针对不同硬件平台实现任务调度算法，使单卡加速比提升7倍；并支撑团队在cv, nlp领域等多个国际性顶会比赛中取得多项冠军。
3. 主导搭建基于TensorRT, CUDA c/c++分布式高性能推断引擎
4. 应用部署ich(脑出血)，青光眼等检测模型
5. 主导建设AI推理平台，从作坊式到平台式演进，实现上百个模型推理服务上线，对外输出产品级服务。
6. 与销售、BD部门紧密合作，共同推进产品及解决方案的售卖及私有化部署和超卖方案的落地实施，及时反馈市场需求，推动产品及解决方案的不断迭代
7. 参与了多个集群管理，人工智能领域开源社区的贡献，并成为contributer，通过google developer 认证
8. 从0到1进行平台建设和人员团队组建，异地团队管理等

**HST, Healthcare Systems & Technologies Co., Ltd.** 2015/05-2018/05

### 高级软件工程师

职责业绩：1. 与美国团队合作，研发基于js-konckout, typescript, Bootstrap, .net signalR, wcf等框架结构的实时web医疗系统。

2. 带领团队研发基于React Native的Mobile App业务系统
3. 作为高级工程师负责软件的整个开发生命周期，把控产品进度与产品质量。
4. 作为北京开发团队负责人，负责团队的技术培训、敏捷开发实践和综合管理，研发过程跟进、任务验收、代码review、质量控制以及解决项目研发进程中的关键技术及事务。
5. 负责公司产品/项目的整体技术架构设计工作，包括系统的总体规划、架构设计、系统分析与设计、核心代码编写；

- 6. 协调各个部门，把控与跟进产品从需求、设计、开发、测试、验收、发布和运营的整个过程；
- 7. 其他团队管理工作，招聘，团建及绩效考核等等

C# SQL Javascript Python Shell C++ 技术管理 软件开发 系统设计  
架构设计

**中国农业新闻网** 2013/05-2015/05

**高级软件工程师**

- 职责业绩：
- 1. 负责项目整体需求分析，系统功能架构设计，参与核心代码及模块编写；
  - 2. 负责协调各有关部门的前期业务需求及后期人员培训；
  - 3. 完成对下级程序员的任务分配、协调及制定详细的任务计划表；
  - 4. 解决研发过程中的技术难题，指导团队成员进行相关技术方案研发；
  - 5. 配合完成日常行政事务。
  - 6. 研发基于微信平台的站点系统及业务系统

**北大方正** 2011/08-2013/05

**高级软件工程师**

- 职责业绩：
- 1.使用UML进行系统设计和需求分析，viso，EA等工具对项目文档进行设计和编写；
  - 2.协助项目的日常管理工作，参与项目前期调研、制定项目计划，控制项目进度、产品质量，有效保障项目的有效实施；
  - 3.参与系统平台的架构搭建,负责系统核心代码的编写，系统集成，项目部署和测试；
  - 3.参与设计和开发基于ActiveMQ的中间件的通信系统。
  - 4.参与开发基于C++的核心任务采集系统。
  - 4.设计并架构基于.net SOAP的Monitor Client & Server。
  - 5.优化基于.net的信号通信系统的传输性能和系统稳定性。
  - 6.基于WPF的通信客户端的性能调优（memory&cpu）。
  - 7.参与部署HP服务器设备的虚机部署。
  - 8.协调用户的开发需求。

---

**项目经验**

**深度学习平台/AI平台** 2018/11-至今

**技术专家/项目负责人**

项目描述: 1. 研发大规模GPU集群算力调度平台, 及针对AI任务场景的深度学习平台。  
2. 构建平安科技深度学习平台生态, 服务于各BU的AI场景和业务  
3. 针对训练和推理场景的平台建设和功能研发, 训练场景下的AI任务调度、调试及资源管理和超卖功能研发; 推理场景下, 支持GPU虚拟化, auto scaling、灰度发布等特性  
4. 针对AI业务的全生命周期的pipeline流程设计与研发, 研发基于airflow, argo, kubeflow等流程化调度平台

### **中国农业新闻网系列产品** 2014/04-至今

项目描述: 1. 针对中国农业新闻网已有的内容管理系统 (TRS) 的日常维护及新的业务功能拓展;  
2. 开发国农优品、涉农导航、展会、美丽乡村等站点系统;  
3. 开发农机频道、资讯财经频道专栏, 从业务需求到功能研发的整体设计分析;  
4. 设计针对手机终端的数据对接模块及相关接口。

项目职责: 配合并指导团队成员进行系统的整体设计, 跟踪研发进度, 保证产品质量。完成关键技术的攻关对部门总监直接负责。

### **基于微信平台的站点系统及业务系统** 2014/01-至今

项目描述: 1. 开发针对微信平台及接口的微网站部署系统, 实现微信服务号的自动生成及部署;  
2. 研发针对微信接口的业务系统, 根据具体的业务需求进行各模块的整合, 涉及酒店、珠宝等行业;  
3. 研发基于微信支付接口的在线支付功能, 实现业务系统对接。

项目职责: 独立进行整体项目的需求分析、设计及实现。完成系统核心模块的研发及系统整体架构的设计。

### **“核高基”科技重大专项** 2012/10-至今

项目描述: Arm, x86(硬件环境)1. “核高基”科技重大专项分课题中对Dalvik虚拟机, Android C运行库和其他函数库的分析及安全性评估。

2.对Android系统整体架构及运行原理进行深入分析; 对Dalvik GC, zygote机制, Dalvik中的JIT编译及C运行库等核心模块进行深入的分析 and 阐述;

3.对Android系统架构及Dalvik虚拟机的移植研究; Dalvik虚拟机与主流Java虚拟机的对比分析与测试等;

4.对Android系统Bionic C库, SQLite, Webcore等核心模块的分析研究及与其标准库的对比测试等;

4.采用专业的测评工具针对Dalvik内部模块性能与平台依赖关系进行相关测评;

5.在分析研究的基础上取得多项科研成果和发明专利。

- 项目职责: 1.对专项任务的调研分析与研究;  
2.对各子任务进行模块划分和任务分配;  
3.负责关键技术攻关和关键问题解决;  
4.负责对团队成员的相关系统知识培训和相关技能分享;  
5.负责相关研究成果的收集与发表。

### **CCTV总控系统** 2011/03-至今

项目描述: vs2008/2010(开发工具)

三菱, 大洋信号接收与控制设备; 思科与hp服务器设备(硬件环境)

unix/windows平台, c/c++/.net运行环境(软件环境),

- 1.基于wcf, wpf的MC Task Execution System。
- 2.基于Devops的MC Monitor System。
- 3.播出信号语音报警系统。

- 项目职责: 1.负责相关系统的研发与技术攻关。  
2.协调用户需求。  
3.软件性能与效率优化。

### **基于Deep-Learning的智能AI内容生成器** 2017/08-2018/08

项目描述: 1. 搭建基于tensorflow, opencv和keras的RNN模型环境

2. 搭建网络爬虫系统, 用于对目标内容的训练数据搜集与修正

3. 使用基于搭建的RNN, LSTM模型结构对内容产生模型进行数据训练

### **全国山洪灾害分布式水文模型并行分析模拟系统** 2015/02-2017/11

项目描述: vs,xmlspy,ea,mockplus,viso(开发工具)开发基于全国水文模型的分布式并行模拟分析系统, 并整合ArcGIS相关资源服务, 进行实时雨水情的并行分析与实时模拟, 给出山洪灾害的实时预警信息。

- 项目职责: 1.负责根据分布式水文模型整理应用解决方案;  
2.参与编写业务需求分析文档, 完成软件项目的设计、系统模块划分、功能设计、代码编写等工作;  
3.参与软件架构设计和新技术研究及论证;  
4.编写产品技术手册、功能及操作手册, 管理项目产品的技术文档;

### **酒店中央预定系统** 2013/10-2014/01

项目描述: vs,viso,xmlspider(开发工具)1. 设计并开发基于.net面向服务的酒店中央预定系统服务端程序, 完成对各分店的数据集中处理及分布式部署;

2. 研发基于IE浏览器的在线语音播报ActiveX插件, 配合服务端服务程序实现服务端无须声卡的语音播报功能;
3. 架构基于多点信息采集及显示的CMS系统, 适配多终端浏览器, 尤其是移动终端浏览器;

项目职责: 针对业务需求独立开发完成核心模块研发、插件开发及前台页面布局等, 配合美工进行页面UE的优化处理; 部署项目系统并进行系统优化。

---

## 教育经历



**北京航空航天大学**

硕士·软件工程

2011/09-2014/06



**河南大学**

本科·计算机科学与技术

2007/09-2011/06

---

## 语言能力

普通话