

张志强

✉ zhangzhiqiangcs@outlook.com · ☎ (+86) 131-2120-0924 · 🌐 zhangzhiqiangcs

教育经历

南开大学	软件学院/软件工程，本科	2012 – 2016
中科院计算所	计算机体系结构/区块链，硕士	2017 – 2020

工作经历

旷视, 北京	软件工程师	2020 年 6 月 – 至今
--------	-------	-----------------

AI PaaS 平台

- 参与 AI PaaS 平台的研发，构建集成**算力、算法、智能体**管理的统一平台，实现推理系统的标准化和扩展性；平台设计符合背景安全防范行业协会（北安协）**中心解析平台**的要求，确保在智慧城市、公共安全等领域的广泛应用。
- 扩展算法管理模块，支持多种**大模型**的托管和推理，基于 **vLLM** 优化推理流程，通过 AI PaaS Gateway 提供 OpenAI API 代理，实现高效的模型推理服务。
- 支持多种智能体接入，提供统一 API 服务，实现会话管理、历史消息管理功能，提升对话质量和并发处理能力。
- 适配 Kubernetes 运行环境，实现模型推理任务的弹性调度和负载均衡，提高系统的稳定性和资源利用率。

私有云平台 MDP

- 基于 Kubekey 设计并实现标准化 Kubernetes 交付方案，简化异构环境下的集群部署流程，实现 Kubernetes 在多种环境中的快速落地，提升系统可维护性与交付效率。
- 构建 Kubernetes 交付平台，作为公司内部**软件发布中心**和现场交付的核心能力平台，集成镜像中心、控制面高可用、监控告警、应用回滚等关键能力，支持 Helm Chart 模块化组合与交付；同时建设现场**集群管理与业务 Dashboard**，实现集群状态可视化、业务监控与运维诊断，提升交付效率与客户体验。
- 提供覆盖**基础设施**可观测性、备份恢复、弹性扩缩容等能力的基础设施最佳实践，支撑业务快速稳定地落地于私有云架构中。

私有云平台 DevOps

- 基于 Docker 和 Supervisor 自研容器编排引擎、集群自动化运维核心功能。在 ToB 领域、私有化、重交付场景中实现标准化交付、自动化运维和最佳实践。
- 包管理器**：自研内部包管理器，包含**命令行工具**，**统一制品仓库与控制器**组件。命令行工具提供依赖解析、配置传递、模板渲染、提交资源；制品仓库具备版本管理、检索与存储能力；控制器监听 apiserver 的资源，控制器监听 apiserver 资源变更并驱动部署至 DevOps 平台。
- 加密授权体系**：基于 CodeMeter 与 Sentinel 设计并实现产品授权机制，提供授权生成、验证与管理能力，有效防止未授权使用，增强核心系统的安全性与商业防护能力。

腾讯, 容器云开发组, 实习生, 北京	2019 年 6 月 – 2019 年 8 月
---------------------	-------------------------

- 参与腾讯云容器存储（Ceph 集群）的开发和维护，优化存储服务的可靠性和性能。
- 排查线上 Ceph 集群异常问题，通过日志分析和 ceph status 监控诊断问题，修复存储节点异常问题，提高系统稳定性。

技能

- 精通 Golang，熟悉 Python，具备大型分布式系统的架构设计与开发经验。
- 熟练使用 Linux 作为主要开发与部署平台，擅长性能优化（Profiling、GC 调优）与线上故障排查。
- 深入研究 **AI Infra**（vLLM、TensorRT），**云原生技术**（K8s、Operator、Service Mesh），**内核技术**（eBPF、cgroup）。
- 持有 Kubernetes CKA 认证，具备 Kubernetes Operator 开发经验，能独立编写 CRD、Controller。