

马晓晖

电话：+86 13810000000  mason_joe
邮箱：joey.gq  joejoe.ma@gmail.com

技能

- Web开发：React、ReactNative(Expo)、Capacitor、Electron、TypeScript、Tailwind / CSS-in-JS、Redux、Zustand、Zod、Tanstack Query
后端: Node.js、Next.js、Nestjs封装CRUD RESTful API、GraphQL、WebSocket，Middleware、Serverless Vercel、Cloudflare, AWS Lambda/ECS冷启动、Edge分布式微服务，API网关、Redis、Kafka、gRPC
- Web框架：Medusajs、Better Auth、GSAP Motion animations、Aceternity /shadcn UI AntD，熟悉构建工具 Webpack、Vite、Turborepo、Monorepo、pnpm workspaces项目管理性能优化的实现例如代码优化、打包优化、资源优化，能结合实际业务场景进行优化
- DevOps：Prisma、Drizzle ORM、PostgreSQL/D1, R2/S3,并使用Docker镜像容器部署,Kubernetes集群，CDN缓存、负载均衡，Sentry.io事件抓捕、Elastic日志Kibana监控，CI/CD / MLOps、自动化Jest测试、灰度发布、A/B，熟悉 Azure 或其他云计算平台（如 AWS、Google Cloud、Vultr、阿里云等）

工作

Hong Kong Network (2023.8- NOW)

AI ITBS Senior software Engineer

- 从0到1的基建部署上线AI Portal赋权BS-CS、API-Dev-向量知识库体系的AI驱动工具链。

广东 物流股份有限公司 (2021.2-2023.7)

工程化 全栈架构组 信息主管

- 三端 (Web,APP,桌面端) ERP系统(框架重构)数据流并发高，跨部门协作集成，容器部署容灾挑战。

广州 跨境有限公司 (2020.2-2021.2)

DevOps 前端开发

- 基于业务跨境电商平台研发,埋点，关联分析，风控数据运营，优化落地上线。

项目

AI Portal业务数据统筹Prompt

Next.js RAG Pipeline Data Lake Kubeflow PyTorch Airflow GSAP

- 主导从0到1搭建基于Microservices + K8s的AI中台，API网关聚合+构建 Kafka + Redis 全链路异步与缓存 (Query/Vector/Prompt Cache) Data Lake + RAG Kubeflow 编排数据清洗、Embedding 及多模态索引构 pgvector + OpenSearch 分布式存储，HNSW + BM25 + RRF 的 Hybrid 检索结合Chunking上下文优化，落地 LLOps平台，集成Langfuse全链路埋点与观测，沉淀检索、Re-rank全栈指标。引入 A/B、灰度测试、检索策略、Prompt、top_k 参数进行持续迭代并整合1M Menu Agent模块需求和自定义规则完成了40+业务场景确保系统在真实流量、回答准确率与聚焦度提升至生产可用级别。
- 平安钟视频监控：自建本地视觉模型设计从数据标注关键帧抽取LoRA 微调 (Fine-tuning) 模型蒸馏自动化训练流水线，针对特定垂直识别包括人脸动作跌倒检测、异常告警、对象检测视频跟踪、构建轻量模型适配边缘计算实时端侧推理转化为结构化语义数据输出监控事件安全报告，误报率降低80%检出率98%的交付标准。
- Next.js + Server / Client状态管理 PPR/SSR/组件级Cache/Compiler 优化实现首屏加载秒级速度，配合lazy和Suspense按需加载，Transition以及图片API预加载优化。按Figma设计图还原及动画使用GSAP交付、解决渲染时出现的Hydration(水合)错误。

ERP/业务财务一体化/进销存物流管理SAAS

WebView PWA Service Worker monorepo Redux ant-design Apollo Cluster

- 基于 React + Nest.js + Capacitor + Electron 构建的企业级物流管理系统，Monorepo + 微服务三端部署响应式适配服务于省级/区域/分拨中心、手持终端、司机智控等解决千万级吨数运输业务动态调度订单性能痛点。
- 重构redux状态管理引入以请求缓存重试并发控制实现了实时同步+Optimistic更新+支持弱网离线缓存IndexedDB并封装分页/排序/虚拟滚动、表单验证等加入事件速率优化,解决初始渲染从17秒降低至3秒。
- 实现通信车辆,货车地图导航可视化拖拽、解析Excel、CSV、回单付水印章电子面单等组件结合业务逻辑实现从路由成本控制审批流转，四级地址结算体系，末端落地配派、拼单抢单模式。网约车智能合单、拆单配送。
- BFF层微服务分析聚合API并实施基于Redis Lua的分级限流与熔断机制，DataLoader解决了 GraphQL N+1查询难题包含JWT令牌动态判断权限、五层权限模型，支持1000+并发连接，处理600+辆车实时位置推送延迟<500ms。
- 构建Kafka数据管道+Kubernetes容器化部署，CI/CD自动化Jest/e2e集成测试与持续交付、Sentry性能监控、WebSocket解决轮询卡顿心跳重连机制，支持60+Topic集群扩展系统高可用架构。

基于node的独立站电商平台集POS线下一体购物系统

Express.js ORM react-i18next Stripe ChatWoot

- 全栈node跨平台集线上线下一体购物(POS)系统开发，RESTful API 和 MVC架构 ORM封装CRUD接口结合react-scanner扫码入库结账,Express.js定制Stripe支付网关，处理状态回调WebHook，支付路由转跳。驱动层实现扫码枪小票打印机的硬件级交互打通 O2O 交易闭环。
- 搭建覆盖率超 90% 的用户行为追踪系统，sendBeacon页面关闭时的数据发送，基于采集的各类指标及转化漏斗数据针对性优化LCP/CLS/PU/UV等指标给予运营决策。
- i18next实现多语言资源按需分块加载并定制开发基于ChatWoot的智能工单系统实现了7x24小时自动化客服响应转化率相比提升28%。

教育

华南理工大学-本科-计算机科学与技术-学士学位证书

全国英语等级考试 PETs-3证书 口语5/5