

孙海洲

中国科学院大学硕士研究生 · ServiceMesher 成员

中国科学院网络数据科学与技术重点实验室

☎ (+86) 1851-0343-050 | ✉ sunhaizhou.ai@gmail.com

👤 haiker2011 (24 followers)

个人简介

目前是中国科学院大学人工智能学院的研究生，ServiceMesher 成员，参与社区多个开源项目。中国科学院计算技术研究所学习和研究自然语言处理和云原生相关。

工作经历

才云科技

中国上海

合作研究

2017 年 11 月至今

- 研究 Kubernetes 上对于机器学习基准测试的系统 kubeflow/kubebench，研究成果发表在 IEEE AI4I'18 会议，在 KubeCon China 2018 发表时长 30 分钟的演讲。
- 实现和维护 Kubeflow 在 Jupyter 上的内核项目 caicloud/ciao，支持从 Jupyter 中发起分布式机器学习训练任务
- 为 Kubeflow 社区维护 TensorFlow 分布式训练支持 kubeflow/tf-operator 和超参数训练系统 kubeflow/katib
- 研究分布式机器学习任务在大规模机器集群上的调度，研究成果发表在 ICA3PP'18 会议

Processing 基金会

中国上海

Google Summer of Code 学生

2017 年 3 月至 2017 年 9 月

- 参与 Google Summer of Code 活动，本次 GSoC 申请的接收率为 6%(1318/20651)
- 基于 R 语言在 Java 虚拟机上的解释器，设计并实现了 Processing 在 R 语言模式 Processing.R，获得 94 stars，成为本次编程之夏 star 最多的项目

摩根士丹利

中国上海

项目实习生

2017 年 2 月至 2017 年 8 月

- 为开源容器调度管理框架 treadmill 实现与 Kubernetes 类似的调度模型
- 调度延迟在 100 节点规模下与原本的调度器相比下降 12%，但增强了其可配置性，支持对节点上硬件资源的动态监控

才云科技

中国上海

合作研究

2015 年 11 月至 2017 年 1 月

- 实现基于 Docker 的持续集成与持续部署系统 Cyclone

上海触宝信息技术有限公司

中国上海

实习大数据工程师

2015 年 9 月至 2015 年 9 月

- 移植爬虫代码到新的平台，优化重写部分过期的爬虫

蚂蚁金服（杭州）网络技术有限公司

中国浙江杭州

实习 Java 研发工程师

2015 年 7 月至 2015 年 9 月

- 在支付宝国际事业部创新业务组任职，从事海外直购业务
- 实现部分包裹清关的逻辑和后台管理的逻辑

演讲与分享

部分文章

论文阅读笔记

GitHub

2016 年 11 月至今

- 记录了 100 余篇学术论文的阅读笔记，论文的方向多为分布式系统，虚拟化，安全，机器学习，超参数训练，网络模型结构搜索等领域

机器学习平台漫谈

gaocegege.com

2018 年 7 月 24 日

- 随着深度学习的兴起，机器学习在最近几年以星火燎原之势席卷了整个科技行业。而在整个机器学习的工作流中，模型训练的代码只是其中的一小部分。除此之外，训练任务的监控，日志的回收，超参数的选择与优化，模型的发布与集成，数据清洗，特征提取等等，都是流程中不可或缺的部分。因此，有一些工具和公司的产品，致力于为机器学习从业者提供一个统一的平台，帮助用户更好地完成其机器学习业务的落地。这篇文章是关于机器学习平台产品的分析对比，由于利益相关性只放出国外的产品

- Kubernetes 已经成为了集群调度领域最炙手可热的开源项目之一。而多工作负载支持，是讨论到集群调度时不得不谈的一个话题。CRD 是 Kubernetes 的一个特性，通过它，集群可以支持自定义的资源类型，这是在 Kubernetes 集群上支持多工作负载的方式之一。本文希望讨论在实现一个 Kubernetes CRD Operator 时可能遇到的问题以及解决方案，抛砖引玉，探索实现的最佳实践

Katib: Kubernetes native 的超参数训练系统

- 这篇文章主要介绍了 Katib，一个由 NTT 贡献到 Kubeflow 社区的超参数训练系统。面向人群为对在 Kubernetes 上运行机器学习负载感兴趣的同学。

教育经历

中国科学院大学

计算机技术专业硕士学位

- 研究方向为自然语言处理，云原生

中国北京

2017 年至 2020 年

电子科技大学

软件工程专业学士学位

- 主修专业（软件工程方向）

中国成都

2011 年至 2015 年

开源项目

awesome-nlp-sentiment-analysis

GitHub

2019 年 4 月

- 收集NLP领域相关的数据集、论文、开源实现，尤其是情感分析、情绪原因识别、评价对象和评价词抽取方面，23 stars

getting-started-with-knative

GitHub

2019 年 4 月

- 《Knative 入门中文版》第三章翻译，其他章节 review，56 stars

trans

GitHub

2017 年 5 月

- ServiceMesher 社区中文资料库，249 stars