软件学院导师团队与招生意向信息表

团队名称	SPAIL（System Performance Analytics and Intelligence Lab）			**团队负责人**	周经森
联系人	赵海亮	**邮箱**	hliangzhao@zju.edu .cn	**电话**	15172392385
意向学生需求数					

主要团队成员			
姓名	**职称**	**研究方向**	**个人主页**
周经森	研究员	System Performance Analytics and Intelligence	教授/博导/硕导 https://kingsum.github.io/
吴克强	研究员	Software Hardware Codesign & Cooptimization	https://person.zju.edu.cn/keqiang
赵海亮	“百人计划”研究员	learning-augmented systems, system performance optimization	博导/硕导 http://hliangzhao.me
常志豪	特聘研究员	软硬件性能优化、序列表示学习	https://proudc.github.io/

团队带头人介绍	Kingsum Chow is the head of this SPAIL lab. He graduated with a PhD from the University of Washington and had almost 30 years of experience working in the industry, spanning more than 20 years at Intel (USA) and more than 5 years at Alibaba (China), also rising to the positions of senior principal engineer and chief scientist. He has published more than 100 papers. 28 patents were issued to him. He is well known for speaking in the industry, including 4 appearances in JavaOne（The highest rated Java conference in the world） keynotes. While working at Intel and Alibaba, he collaborated with Amazon, Arm, Ampere, BEA（acquired by Oracle), Google, IBM, Microsoft, Oracle, Siebel (acquired by Oracle), Sun (acquired by Oracle) and Tencent, resulting in many software-hardware co-optimized solutions. In addition, he represented Alibaba in the world-wide Java standard committee, JCP EC（Java Community Process Executive Commission）election and took one of the 18 seats in the committee. As of today, Alibaba is still the only Chinese company with a seat there. Since joining the school in March 2023, Kingsum has established a course in System Performance Analytics at Scale and led 7 industry projects with Alibaba, Ampere, Bytedance, Huawei, and Kuaishou.
吴克强介绍	吴克强,美国明尼苏达大学双城分校的电气与计算机工程博士,美国伊利诺伊大学厄巴纳-香槟分校的工程学硕士学位。在计算机芯片,人工智能,大数据, 及数据中心计算方面有着深厚的专业知识和丰富的工程经验。作为第一作者及关键发明者拥有 10 余项芯片设计美国专利。作为英特尔软硬件协同设计技术领导者,吴博士发现现有芯片缓存设计的不足,主导下一代芯片缓存的软硬件协同设计与性能优化,设计并推动下一代 CPU 缓存替换方案,作为第一作者申请多项专利。在人工智能及大语言模型方面,他解决了深度学习模型中的众多瓶颈，实现最优的性能功耗并已落地英特尔 oneAPI compiler 2023.2 release 及之后的版本。值得强调的是,其缓存工作于 2023 初入选英特尔人工智能与分析（AIA）2022 年十大成就之一。在英特尔硅谷任职期间（2005-2018 年）,吴博士长期在甲骨文总部工作,主导英特尔与甲骨文诸多团队（Java, Database, Business Intelligence, Analytics 等）合作,成功解决甲骨文 Java 应用、数据库、Business Intelligence 等关键性能问题,持续推动芯片设计架构变革,在多款甲骨文产品实现显著性能提升,其成就多次获英特尔与甲骨文主席等高管的公开认可（OOW2016、OOW2014、JavaOne 2010）。在霍尼韦尔美国任职期间（2004 - 2005 年）,因首创太空通讯模拟器获得霍尼韦尔技术成就奖。

赵海亮介绍	赵海亮，博士，浙江大学软件学院“百人计划”研究员，ACM/IEEE/CCF 专业委员，启真优秀青年学者。2024 年博士毕业于浙江大学计算机学院，曾于新加坡南洋理工大学访问交流（2022-2023）。研究方向为服务计算、调度优化、工业智能等，成果包括资源调度优化（DPE@TPDS '21、ESDP@TPDS '22）、学习增强算法（PFSUM@NeurIPS '24、Guard@NeurIPS '25）、模型量化及推理加速技术（SegQuant@CVPR '26）等。近年来，在 Proceedings of the IEEE、TPDS、TMC、TSC、NeurIPS、CVPR、ICWS、MM 等国际顶级期刊/会议上发表/录用论文 40 余篇，含 ESI 热点论文/高被引论文 2 篇，总计被引 2200 余次。长期担任 TMC、TSC、TKDD、计算机学报、FGCS、NeurIPS、CVPR 等期刊/会议审稿人及会议程序委员会成员，担任计算机工程与科学青年编委。获浙江省优博提名（2026）、CCF 服务计算专委优博（2025）、浙江大学优博（2024）、领域顶级会议 IEEE ICWS 2019 最佳学生论文等奖项及荣誉。主持国家重点研发计划课题、国家自然科学基金青年项目，参与浙江省自然科学基金重大项目、浙江省“领雁”科技计划项目，主持/参与华为、快手、字节跳动、航天五院、西飞等多项企事业单位技术合作项目。
常志豪介绍	常志豪，2024 年于浙江大学获得计算机博士学位，现担任浙江大学软件学院特聘研究员，浙江大学启真优秀青年人才项目获得者。当前主要研究方向为软硬件协同性能优化、序列表示学习，主持及参与国家重点研发计划子课题，阿里巴巴、华为技术合作项目等多项课题。在 VLDB、VLDBJ、TKDE、ICDE、AAAI 等 CCF-A 类会议及期刊上已第一作者发表高水平学术论文 5 篇，作为共同作者在 USENIX SRECon21 上发表软硬件协同性能优化演讲:“Need for SPEED: Site Performance Efficiency, Evaluation and Decision”，并针对 Intel 超线程技术在 OS 统计 CPU 利用率时造成的干扰，以第一作者身份在阿里巴巴集团内部提交“一种单容器 CPU 利用率的修正算法”创新提案，并成功应用于阿里巴巴集群性能优化。
团队介绍	The team is comprised of researchers specializing in software engineering, distributed computing, system performance optimization, and artificial intelligence. We possess deep expertise in algorithm design, software and hardware co-optimization, and system tuning, backed by research experience at top institutes and extensive industry roles with international high-tech companies. Our core mission drives us to be all in on agentic AI, developing autonomous agents that leverage our system-level strengths for complex reasoning and execution. Robotics stands as a central pillar of our work. We actively engineer intelligent robotic systems using ROS (Robot Operating System) and RTOS (Real-Time Operating Systems), seamlessly integrating AI for Robotics to achieve advanced perception, decision-making, and control. By combining our foundational strengths in system optimization with cutting-edge AI, we bridge the gap between theoretical algorithms and real-world embodied intelligence. The team is fluent in both English and Chinese communications. Our team maintains close collaborations with multiple enterprises, research institutions, and overseas universities. With abundant research funding and a vibrant scientific atmosphere, we provide numerous opportunities for students to engage directly in projects spanning system performance tuning, hardware-software optimization, Agentic AI, and advanced robotics. This environment enables them to showcase their individual capabilities, pursue academic aspirations, and enhance their engineering expertise, fostering both academic growth and engineering proficiency across these critical domains.

项目情况	SPAIL has solid projects spanning from the academia to the industry. SPAIL has healthy funding to continue the current projects in computer system performance optimization and explore new ones. Currently, SPAIL is engaged in collaborative projects with several high-tech companies, focusing on performance data collection, analysis, modeling, and prediction across different instruction sets and CPU microarchitectures. The projects aim to enhance the understanding of performance metrics and optimize system efficiency through advanced algorithm design, data analytics and modeling techniques.
团队与企业合作情况	As Kingsum was a senior leader in the industry for almost 30 years, SPAIL has established research collaborations with major tech giants in the industry, spanning from cloud computing to software hardware co-optimization. Students can work on research projects solving real-world problems. There are opportunities to work in school and intern at both well-established industry leaders and flourishing startups in system performance optimization
对学生的要求	1、 诚实、有激情、有好奇心，追求真实的可复现的工作 2、 追求体系结构和性能分析方向的研究和工程实践 3、 Read and write in English. 4、 较好的数学基础、计算机基础知识和一定的动手能力 5、 良好的工程能力，熟悉常用编程语言如 Python, C/C++, Rust 6、 良好的数据分析能力, Python packages, Jupyter, Excel, (assisted by ChatGPT, Qwen, etc) 7、 对分布式计算、并行计算有一定的理解和认识
团队可以在宁波开设专业课程情况	Kingsum teaches “System Performance at Scale”: In-depth understanding of software and hardware co-optimization at the speed of software and hardware evolution in September 2024. This class will present the theory and practice of doing system performance analysis he did for almost 30 years. 吴克强博士计划开设《RISC-V 计算机体系架构和性能》课程，旨在系统讲授 RISC-V 架构的核心原理、处理器实现及产业应用知识。 赵海亮博士已开设《机器学习优化理论》课程。这门课介绍机器学习及机器学习系统领域中的典型优化理论与方法，并结合系统性能优化中的典型研究问题介绍对应的建模方法、问题分析、算法设计与实现。 常志豪博士计划开设《软硬件性能优化》课程，该课程将介绍软硬件性能优化领域的相关技术，包括相关基础理论以及相关会议上的最新研究进展。