

The World Computing - Power Network Conference 2025 2025世界算力网大会

会议手册





大会简介 | ABOUT THE CONFERENCE

2025世界算力网大会聚焦"构建全球算力共同体"愿景,重点推进以下工作:提升世界算力网联盟影响力,促进算力资源跨境流动与优化配置;展示我国算力网建设最新成果,发布算力并网、智能调度等关键技术标准;探讨算力赋能千行百业的创新应用,推动数字经济高质量发展。我们期待通过会议汇聚全球智慧,共同推动算力像水电一样成为普惠性公共基础设施,为构建人类数字命运共同体作出积极贡献。

WCNC2025 aims to promote cross regional sharing of computing power resources, foster technological collaboration, establish an open and inclusive global governance systems, and develop standards for transactions, security, and efficiency in computing power network. This event also seeks to showcase real-world applications across industries with computing power networks. We look forward to gathering global wisdom through the conference, jointly promoting computing power to become a universal public infrastructure like water and electrical power supply, and making positive contributions to building a digital community with a shared future for mankind.

主办单位/Host 世界算力网联盟(WCNC) World Computing-power Network Consortium (WCNC)

协办单位 /Co-organizer 广东省人工智能与机器人产业联盟 Guangdong Artificial Intelligence and Robotics Industry Alliance

支持单位/Supporting Organization 全球计算联盟 Global Computing Consortium (GCC) 全球云网宽带产业协会 World Broadband Association (WBBA)



大会手册 Conference Handbook



大会官网 Official Website



组织机构 | Conference Committees

Steering Committee



Chair
Wen GAO
Director of Peng Cheng
Laboratory



Lionel Ni
Fellow of the Hong Kong Academy
of Engineering (HKAE), the Founding
President of HKUST(GZ)



Member
Sherman Shen
Foreign Academician of the Chinese
Academy of Engineering (CAE), Fellow
of the Royal Society of Canada



Member
WANG JIAN
President of Zhejiang Lab, Founder of
Alibaba Cloud



Member

Depei Qian

Academician of Chinese Academy
of Sciences (CAS)



Guanglu Shao
Director of Science and Technology
Committee & Former General Manager
of China Telecom



Member
Hai Jin
Professor, Huazhong University
of Science and Technology



Member &
Executive Secretary
Yutong Lu
Professor, Sun Yat-Sen University

Program Committee



Chair
Yutong Lu
Professor, Sun Yat-Sen University



Vice Chair
Wenguang Chen
Professor, Tsinghua University

Organizing Committee



Chair

Jiankun Yang

Assistant Director of Peng Cheng Laboratory



会议议程 | CONFERENCE AGENDA

2025世界算力网大会开幕式及主论坛 WCNC2025 Opening Ceremony & Main Forum

时间time	主题 Session				
08:30-09:00	开幕式/Opening Ceremony				
主论坛/Main Forum 主持人/Host: 陈文光/Wenguang Chen					
时间time	主题 Session	嘉宾 Speaker			
09:00-09:30	National Computing Power Sovereignty and "China Computing Net" Project	Wen Gao Director of Peng Cheng Laboratory			
09:30-10:00	Computer Simulation and AI on HPC	Michael Resch Professor of University of Stuttgart, Director of HLRS			
10:00-10:30	Computing, AI and 3-Body computing constellation: endless frontiers and exploration	Jian Wang President of Zhejiang Lab, Founder of Alibaba Cloud			
10:30-11:00	Co-Designing Intelligent Cyberinfrastructure for Computing Continuum: Overview of the Activities at the NSF-AI Institute ICICLE	Dhabaleswar K (DK) Panda Professor of the Ohio State University, USA			
11:00-11:30	Pushing AI from Research to Application and Impact	Sumei Sun Executive Director of I ² R, A*STAR, Fellow of Academy of Engineering Singapore			
	Panel: Operation and Ecosystem Development of Computing-Power Network	Hai Jin (Moderator) Professor, Huazhong University of Science and Technology			
11:30-12:10	Wen Gao Director of Peng Cheng Laboratory	Weimin Zheng Professor, Tsinghua University			
	Lionel Ni Founding President of HKUST(GZ)	Guanglu Shao Director of Science and Technology Committee of China Telecom			



算力网应用相关分论坛 Technical Tracks

分论坛1:大模型训练与推理

Track #1: Large Language Model Training and Inference 会议地点/Venue: 7楼/7F



论坛主席/Chair 陈文光/Wenguang Chen 清华大学教授 Professor, Tsinghua University

时间 time	主题 Session	嘉宾 Speaker
14:30-15:00	人工智能大模型推理架构的技术挑战 Technical Challenges in Inference Architecture for Large Al Models	郑纬民/Weimin Zheng 清华大学教授 Professor, Tsinghua University
15:00-15:30	Co-GenAI:一种新型融合驱动平台 Co-GenAI: A Novel Fusion-Driven Platform	杨红霞/Hongxia Yang 香港理工大学教授 Professor, The Hong Kong Polytechnic University
15:30-16:00	构建基于图形用户界面的人工智能代 理的经验与启示 Lessons learned in building an GUI-base AI agent	钱岭/Ling Qian 中国移动云能力中心首席科学家 Chief Scientist, China Mobile Cloud Centre
16:00-16:30	大规模混合专家语言模型推理引擎优 化: 经验与教训 Optimizing inference engine for large MoE language models: experience and lessons	谭焜/Kun Tan 华为中央软件院首席专家 Chief Expert, Central Software Institute, Huawei
16:30-17:00	欺骗性对齐的内在机制与方法 Intrinsic Mechanisms and Methods of Deceptive Alignment	杨耀东/Yaodong Yang 北京大学人工智能研究院助理教授 Assistant Professor at the Institute for AI, Peking University
17:00-17:45	Panel: 大模型能力是否已经基本到顶及未来的主要挑战 Have Large Language Model Reached a Performance Plateau? Navigating the Next Frontier of AI Challenges 主持人: 陈文光 (清华大学教授) 嘉宾: 杨红霞 (香港理工大学教授) 钱岭 (中国移动云能力中心首席科学家) 谭焜 (华为中央软件院首席专家) 杨耀东 (北京大学人工智能研究院助理教授) 杨耀东 (北京大学人工智能研究院助理教授) 陆游游 (清华大学计算机系副教授) Panelists: Hongxia Yang (Professor, PolyU) Qian Ling (Chief Scientist, China Mobile Cloud Centre) Kun Tan (Chief Expert, Central Software Institute, Huawei) Yaodong Yang (Assistant Professor at the Institute for Artificial Intelligence, Peking University) Youyou Lu (Associate Professor of Tsinghua University)	



算力网应用相关分论坛 Technical Tracks

分论坛 2:算力网云际计算前沿论坛

Track #2: Computing-power Network and JointCloud Computing 会议地点/Venue: 3楼/3F



论坛主席/Chair 余跃/Yue Yu 鹏城实验室研究员

时间 time	主题 Session	嘉宾 Speaker
14:25-14:30	开场致辞 Opening Address	金海/Hai Jin 华中科技大学教授,GCC理事长 Professor at HUST, Chairman of Global Computing Consortium
14:30-15:00	分布式调度驱动的超算互联网算力原生技术与实践 The Native Technology and Practice of Supercomputing Internet Computing Power Driven by Distributed Scheduling	(高WCNC) 特邀专家 Invited Expert
15:00-15:30	面向拥塞控制测试的可编程网络分析技术研究 Research on Programmable Network Analysis Techniques for Congestion Control Testing	田臣/Chen Tian 南京大学教授 Professor at the School of Computer Science, Nanjing University
15:30-16:00	面向超算与智算的高性能互连可扩展性设计 High-Performance Interconnects Scalability Design for High-Performance and AI Computing	董德尊/DeZun Dong 国防科技大学计算机学院研究员 Researcher at the College of Computer, NUDT
16:00-16:30	建设超算互联网推动算力生态发展 Building a High-Performance Computing Internet to Promote the Development of the Computing Ecosystem	郭庆/Qing Guo 国家高性能计算机工程技术研究 中心总工程师 Chief Engineer of the National Engineering Research Center for HPC
16:30-17:00	中国算力网上的生物计算:现状与未来挑战 Life science computing on Chinese National Computing Network: current status and future challenges	王宇/Yu Wang 鹏城实验室研究员 Professor at Peng Cheng Laboratory
17:00-17:45	Panel:智能时代的计算范式/Computing Pa 主持人: 余跃(鹏城实验室研究员) 嘉宾: 金海(华中科技大学教授,GCC理事长) 田臣(南京大学教授) 董德尊(国防科技大学计算机学院研究员) 王宇(鹏城实验室研究员)	radigm in Artificial Intelligence Era Moderator: Yue Yu (Researcher at Peng Cheng Laboratory) Panelists: Hai Jin (Professor at HUST, Chairman of GCC) Chen Tian (Professor at Nanjing University) DeZun Dong (Researcher at the College of Computer, NUDT) Yu Wang (Professor at Peng Cheng Laboratory)



算力网应用相关分论坛 Technical Tracks

分论坛 3:算力网与安全前沿论坛

Track #3: Computing-power Network and Security 会议地点/Venue:6楼/6F



论坛主席/Chair 贾焰/Yan Jia 国防科技大学研究员

时间time	主题 Session	嘉宾 Speaker
14:30-14:50	算力网络一体化安全的创新与实践 Innovations and Practices of Integrated Security in Computing Power Networks	李安民/Anmin Li 中国电信研究院副院长 Deputy Director of the China Telecom Research Institute
14:50-15:10	面向算力网络安全的靶场发展与应用 Development and Application of Cyber Ranges for Computing Force Network Security	杜雪涛/Xuetao Du 中国移动设计院网信安全产品部总经理 Director of Network Security Dept., China Mobile Group Design Institute
15:10-15:30	广域智算网络的思考与实践 Thoughts and Practices on AINet of China Unicom	曹畅/Chang Cao 中国联通研究院下一代互联网研究部总监 Director of Next-Generation Internet Research at China Unicom Research Institute
15:30-15:50	算力网络内生安全架构及关键技术 Endogenous Security Architecture and Key Technologies for Computing Power Networks	王继刚/Jigang Wang 中兴通讯股份有限公司副总裁 Vice President, ZTE Corporation
15:50-16:00	休 息/Cc	offee Break
16:00-16:20	算网安全的第三范式思考 Reflections on the Third Paradigm of Computing-Power Network Security	潘柱廷/Zhuting Pan 启明星辰集团首席战略官 Chief Strategy Officer of Venusense Group
16:20-16:40	算力互联网的发展和安全趋势 The Development and Security Trends of the Computing Power Internet	魏亮/Liang Wei 中国信息通信研究院副院长 Deputy Director of the CAICT
16:40-17:00	AI-Native安全,护航算力网长治久安 AI-Native Security: Safeguarding Long-Term Stability and Resilience of Computing Power Networks	杨松/Song Yang 深圳华为云计算技术有限公司云安全CTO Chief Technology Officer at Security Cloud, Huawei Cloud
17:00-17:45	Panel: 算力网络一体化安全问题探讨/Discussion on 主持人: 郑志彬 (软极网络技术(北京)有限公司CEO) 嘉宾: 董振江 (南京邮电大学信息智能与安全研究所所长) 孟楠(中国信息通信研究院安全所副所长) 桑梓勤 (中国信息通信科技集团光通信技术和网络全国重点实验室智慧城市研究室主任) 张峰(中国移动集团级首席专家)	the Integrated Security Issues of Computing Power Networks Moderator: Zhibin Zheng (CEO of Rangesoft Network (Beijing) Technology Co., Ltd.) Panelists: Zhenjiang Dong (Director of Institute of Information Intelligence and Security, NJUPT) Nan Meng (Deputy Director of the Security Research Institute at the CAIC) Ziqin Sang (Director of State Key Laboratory of Optical Communication Technology and Network, CICT) Feng Zhang (Group-Level Chief Expert of China Mobile)



WANG JIAN

Dhabaleswar

K (DK) Panda

professorship for HPC at the University of Stuttgart. His research focusses on the usage of HPC systems in engineering applications. This includes Computational Fluid Dynamics and Artificial Intelligence. Prof. Resch holds a Dipl.-Ing (M.Sc.) in Technical Mathematics of the Technical University of Graz, Austria and a Dr.-Ing. (PhD) in

PhD by the Russian Academy of Science and an honorary doctoral degree by the Donezk National Technical University, Ukraine. Prof. Resch is currently a guest professor at Tongji University, Shanghai, PRC.

Traditionally HPC was used for computer simulations helping both science and industry to advance. Most recently, AI has become dominant in the HPC market. While some believe that AI may replace traditional computer simulation we think that AI and computer simulation will co-exist. This talk will show some ideas of how AI and CS

Dr. Jian Wang was the chief technology officer of Alibaba Group, and is the founder of Alibaba Cloud, which is a major public cloud provider globally and ranks as first in Asia. He was also the chief architect of Apsara, the

a public service. In 2016, Dr. Wang pioneered and led the nonprofit City Brain initiative to develop a new digital infrastructure for sustainable "smart" cities as his personal effort and was the architect of Hangzhou City Brain. He

for City Brain, and the Yunqi Science and Technology Innovation Foundation, a philanthropic private foundation that operates the 2050 museum which is open to public for free. The foundation envisions the future through

technology innovation and organizes the 2050 event every year in April, promoting the vision of science and

DK Panda is a Professor and University Distinguished Scholar of Computer Science and Engineering at the Ohio State University. He is serving as the Director of the ICICLE NSF-Al Institute (https://icicle.ai). He has published over 500 papers. The MVAPICH MPI libraries, designed and developed by his research group (http://mvapich.cse.ohio-state.edu), are currently being used by more than 3,400 organizations worldwide (in 92

frameworks and Machine Learning applications from his group are available from https://hidl.cse.ohio-state.edu.

Similarly, scalable, and high-performance solutions for Big Data and Data science frameworks are available from https://hibd.cse.ohio-state.edu. Prof. Panda is a Fellow of ACM and IEEE. He is a recipient of the 2022 IEEE Charles

Co-Designing Intelligent Cyberinfrastructure for Computing Continuum: Overview of the Activities at the NSF-AI

Artificial intelligence (AI) is transforming every sector of society. However, there is a massive and ever-growing gap between available AI techniques and their availability to end users across a range of application domains. Existing

infrastructure. This talk will start with an overview of the ICICLE (Intelligent CyberInfrastructure (CI) with Computational Learning in the Environment), an NSF-Al Institute, to address these challenges. We will demonstrate how ICICLE seeks to be the first and foremost edge-to-center Al-as-a-service enterprise for the emerging computing continuum, advancing foundational AI research by fostering next-gen CI for AI to support the

of AI to the wider population. An overview of the co-designing activities at the institute spanning multiple

knowledge graphs, model commons and conversational AI), data privacy and trust, visualization, and three leading use-inspired sciences (Animal Ecology, Digital Agriculture, and Smart Foodsheds) will be highlighted. Overall, the

Dr Sumei Sun is the Executive Director of Institute for Infocomm Research (I2R), A*STAR, Singapore. A*STAR I2R focuses on multi-disciplinary digital technologies research including artificial intelligence (AI), communications

communications, integrated sensing-communications-computing-control, applied Al, and industrial internet of things. She's a Fellow of the Academy of Engineering Singapore (SAEng) and Fellow of the IEEE. She's recipient of

Career Award, and Singapore National Day 2022 Public Administration Medal (Bronze). She's an elected Board of Governors member of the IEEE Vehicular Technology Society, and Chair of Fellow Evaluation Committee of IEEE ComSoc. She holds a joint appointment with the Singapore Institute of Technology, and an adjunct appointment

In 2023, the Singapore National AI Strategy 2.0 (NAIS 2.0) was launched for Singapore to achieve the twin goals of excellence and empowerment for AI for the Public Good, for Singapore and the World. As a national research $institute focusing on multi-disciplinary digital technologies research, A^*STAR Institute for Infocomm \,Research \,(I^2R)$

digital transformation in key sectors. In this talk, we will share our endeavour in this pursuit, and report some of our latest results in multimodal foundation models and the applications in digital services, aerospace and aviation, and next-generation land transportation. We will also discuss challenges and efforts in AI model compression and

Dr. Yutong Lu now holds an appointment as Professor in the Department of Computer Science and Engineering at Sun Yat-Sun University, China. She also severs as the Director of National supercomputing center in both Guangzhou and Shenzhen. Professor Lu specializes in high performance computing, mainly including the design

environments. Her extensive research and development experience has spanned several generations of domestic

achieved the remarkable distinction of ranking No.1 in the world on the Top500 list for six consecutive times.

disciplines. Her efforts have facilitated significant breakthroughs in climate modeling, bioinformatics, and materials science, among other fields. Her work has not only enhanced the performance and efficiency of supercomputers but has also made these powerful tools more accessible to a broader range of researchers and

Committee of China's National Key R&D Program. She has won several national awards for science and technology advancement, and is leading a number of research projects under the support of the Ministry of Science and

cutting-edge next-generation architectures, as well as the convergence of advanced AI and HPC systems and

清华大学计算机系王建筑讲席教授,鹏城实验室研究员,智能计算部主任。主要研究领域为大规模分布式计算系统,包括图计算

researcher at Pengcheng Laboratory, and Director of the Intelligent Computing Department. His main research areas include large-scale distributed computing systems, such as graph computing systems, big data processing

Telecommunications in 1975, received a Master's degree in telecommunications from the University of Essex, U.K.

From 1979 to 1980, he was one of the first group visiting scholars to study in Switzerland, sent by China Education

From 1975 to 1986, Zhao worked in the design Institute of the former Ministry of Posts and Telecommunications. In

Bureau (TSB) from 1999 to 2006, Deputy Secretary-General from 2007 to 2014, Secretary-General from 2015 to 2022.

清华大学计算机系教授,中国工程院院士。长期从事高性能计算机体系结构、并行算法和系统研究。提出可扩展的存储系统结构 及轻量并行的扩展机制,发展了存储系统扩展性理论与方法。在国内率先研制并成功应用集群架构高性能计算机。在国产神威太 湖之光上研制的极大规模天气预报应用获得ACM Gordon Bell奖。曾获国家科技进步一等奖1项、二等奖2项、国家技术发明二等

报告围绕当前人工智能大模型推理系统的架构演进,介绍以存换算和以存强算两种前沿技术。报告从AI大模型生命周期出发,分

析了大模型推理在存储、带宽与计算资源上的巨大挑战,重点聚焦于KV-Cache的复用机制以及推理过程中的异构系统协同策 略。Mooncake架构通过存储优化推动算力复用,已广泛应用于如Kimi、DeepSeek等大模型的推理部署,并成为vLLM和SGL等 主流系统的核心组件之一;KTransformers则以CPU/GPU协同为核心,致力于推动个人级大模型推理能力的普及化。报告中结 合架构部署情况、关键技术突破、开源进展以及未来发展趋势,为大模型推理系统的设计与应用提供了实践指导和战略思考。

Prof. Hongxia Yang, Associate Dean of Faculty of Computer and Mathematical Sciences & Professor at PolyU, PhD

Award), the second prize of the 2020 National Science and Technology Progress Award, the first prize of Science and Technology Progress of the Chinese Institute of Electronics in 2021, the Forbes China Top 50 Women in Science and Technology and Ministry of Education Science and Technology Progress Award First Class in 2022, Al 2000 Most

Member at IBM T.J. Watson Research Center, joint adjunct professor at Zhejiang University Shanghai Advanced

We introduce a ground breaking platform designed to make AI development more accessible and efficient. At its core and all the production of the productiois Domain-Adaptive Continual Pretraining (DACP), a pioneering system that enhances Large Language Models by

enterprise and scientific domains often underrepresented in general web data. DACP consistently outperforms

minimizing GPU costs. Complementing this is the Advanced Model Fusion Infrastructure, which utilizes a "Model

pretraining structures. This innovative approach enables the creation of foundation models by fusing existing models, drastically reducing computational requirements compared to traditional methods—requiring only 160 GPU hours to merge four top models, as opposed to 1-1.6 million GPU hours needed to train foundation models from scratch. The platform's Resource-Efficient Architecture further democratizes Al development by facilitating the effective use of distributed, entry-level GPU resources. By leveraging distributed high-performance computing centers equipped with diverse computing accelerators, the platform efficiently trains foundation models through the fusion of smaller language models (SLMs), offering a viable alternative to training large foundation models from scratch. This approach reduces dependency on massive centralized computational resources, fostering greater

Qian Ling, Ph.D., Senior Engineer, Chief Expert of China Mobile Group, Deputy Chief Engineer of China Mobile

many books in this field, published more than 20 papers including the Nature sub-journal, and obtained more than

At present, on the one hand, the intelligence level of LLM continues to improve, but on the other hand, practical

upper-level intelligent applications are still traditional services such as chatbots and search engines, whether there

cutting-edge Al frameworks, agentic Al systems, cloud native and serverless, and communication technology for

Award in 2009, USENIX Test-of-Time Award in 2019, and State Technological Invention Award (Second Class) in 2023.

expert-parallelism to implement a high-throughput and low latency DeepSeek R1 inference model based on JiuSi.

separation, distributed KV cache with shared memory, and dynamic load balancing among experts. We show that our implementation delivers the state-of-the-art performance on Ascend hardware. Lastly, we also highlight several

focuses on interactive learning and alignment for intelligent agents, spanning the fields of reinforcement learning,

with more than 10,000 citations on Google Scholar. His honors include the ACL 2025 Best Paper Award, ICCV 2023 Best Paper Award Finalist, and CoRL 2020 Best System Paper Award. He led a Chinese research team to develop a multi-agent reinforcement learning algorithm that was the first from the community to be published in Nature

Deceptive Alignment refers to a phenomenon in which, once a model possesses a certain degree of self-

supervision or evaluation based on input patterns, contextual cues, and other features.2. Strategic Output: When supervised, it generates compliant appearances; in "safe" contexts, it reverts to its true policy.3. Motivational Drive: It optimizes for long-term rewards or built-in objectives rather than immediate human preferences.

adversarial evaluation, sandbag testing, and process supervision (e.g., Chain-of-Thought monitoring) can be introduced to force the model to maintain consistency across diverse domains and states. During deployment, self-

优青项目、重点研发计划青年科学家项目、战略性创新合作项目等国家项目。在FAST、OSDI、SOSP等国际顶级会议上发表论文 50余篇,曾获得NVMSA'14最佳论文奖、MSST'15最佳论文提名奖、SIGMOD'23研究亮点论文、CACM'25研究亮点等奖项。研制 高性能文件系统SuperFS,部署于鹏城云脑II,2023年至今蝉联世界超算存储IO500榜单第一名。2025年指导学生获得ASPLOS/EuroSys大模型推理优化竞赛冠军。担任FAST、OSDI、USENIX ATC、EuroSys等国际会议程序委员会委员。曾获得

His current research focuses on storage systems and machine learning systems, with dozens of publications in top-tier conferences such as FAST, SOSP, OSDI, and USENIX ATC. His work has been recognized with multiple awards, including the Best Paper Award at NVMSA 2014, the Best Paper Runner-up at MSST 2015, the Research Highlight

鹏城实验室高效能云计算研究所研究员、博导,聚焦分布式智能计算与系统平台方面开展研究,在软件工程与人工智能国际顶级 会议或期刊发表论文80余篇,担任新一代人工智能产业技术创新战略联盟(AITISA)智算中心和智算网络标准工作组组长、IEEE

focusing on distributed intelligent computing and system platforms. Published over 80 papers in top-tier

and Intelligent Computing Network Standard Working Group of the New Generation Artificial Intelligence Industry Technology Innovation Strategic Alliance (AITISA), and Chair of the IEEE P3404 Standard Working Group. Led the

金海,博士,华中科技大学教授、博士生导师,长江学者特聘教授,国家杰出青年基金获得者,国家"万人计划"科技创新领军人才。中国计算机学会副理事长、会士,IEEE Fellow。全球计算联盟理事长。华中科技大学"大数据技术与系统国家地方联合工程研究中 心"主任、"服务计算技术与系统教育部重点实验室"主任。全国创新争先奖获得者、中国计算机学会王选奖获得者。国务院特殊津 贴专家。"十三五""云计算与大数据"国家重点研发计划专家组副组长、"十四五""先进计算与新兴软件"国家重点研发计划专家 组组长。获国家自然科学二等奖1项、国家科技进步二等奖2项、国家发明二等奖1项。主要研究领域为计算机体系结构、并行与分

Hai Jin is a Chair Professor of computer science and engineering at Huazhong University of Science and Technology

2001. Jin is a Fellow of IEEE, Fellow of CCF, and a life member of the ACM. He has co-authored more than 20 books and published over 900 research papers. His research interests include computer architecture, parallel and

李肯立,现任湖南大学副校长,教授。国家自然科学基金创新研究群体项目负责人,国家杰青、长江学者、万人计划科技创新领军

人才,CCF会士。长期从事高性能计算调度与应用研究工作,主持承担国家重点研发计划、科技创新2030重大项目等国家项目 18项,入选全国创新争先奖,以第一完成人获国家科技进步二等奖2项、中国专利金奖1项、省部或学会科技一等奖4项。 Kenli Li is currently a Cheung Kong Scholar Chair Professor and the Vice-President of the Hunan University and the

Director of National Supercomputing Center in Changsha. He is also the principle investigator (PI) of the Creative

Li's research interests mainly include high-performance computing scheduling and applications, as the PI, he has

the National Science and Technology Innovation 2030 Initiative. He was selected for the National Innovation and Excellence Award. As the first-completer, he has won the second-class Award of National Science and Technology Progress (twice), one China Patent Gold Award, and four first-class awards of science and technology from

田臣,男,现为南京大学计算机学院教授、博士生导师。2023年入选国家杰出青年科学基金资助,个人研究专长为计算机网络和分

名国际期刊上录用和发表论文100余篇。为下一代数据中心网络提出以流量控制为中心的拥塞管理理念,设计拥有自主知识产权的带状态可编程网络测试仪,主导实现对开源网络仿真软件的大规模并行加速,曾任OpenNetLab国际网络实验床的轮值主席。

息、节点状态等采集到数据仓库,从而使得借助人工智能、专家经验、大数据分析等技术实现对物理网络全生命周期的分析、诊 断、仿真和决策提供数据支撑。智算中心是支撑人工智能大模型的算力基础设施,而拥塞控制是保障智算中心网络高效稳定运行的关键技术。面向拥塞控制的网络测试仪是验证复杂拥塞场景下算法设计、参数调优和部署正确性的核心手段。现有的网络测试技术无法同时满足高吞吐、高并发、可编程、有状态、高通量、高精度的要求,成为智算中心发展的主要技术瓶颈之一。本报告将汇

The digital twin of a network is a digital replica of a physical network. It collects data from the physical network—such as data packets, configuration information, and node states—into a data warehouse through real-

intelligence, expert knowledge, and big data analytics.AI computing centers serve as the computational infrastructure supporting large-scale AI models, wherein congestion control is a key technology ensuring efficient

network testing technologies fall short in simultaneously meeting the demands for high throughput, high

董德尊,CCF杰出会员、体系结构专委常委、分布式计算与系统专委常委,国防科技大学计算机学院研究员、博导,教育部长江学

者、全国百篇优秀博士论文奖获得者、国防卓青、湖南省杰青,主要研究计算机体系结构、高性能与智能计算、并行与分布式系统,长期参与国产高性能计算机系统研制工作,获湖南省教学成果特等奖、军队科技进步一等奖、湖南省自然科学一等奖、中国电子学会和中国计算机学会(CCF)自然科学二等奖等,在CCF推荐的A/B类国际刊物发表论文80余篇,担任Fundamental Research、

Education, a recipient of the National Excellent Doctoral Dissertation Award, an Outstanding Young Scholar of the National Defense Science and Technology, and an Outstanding Young Scholar of Hunan Province. His research

互连通信是决定超算与智算系统可扩展性的关键因素之一。为了提升高性能互连系统的可扩展性,需要从不同层次和维度进行思考和尝试,包括在芯片内探索互连微架构的高带宽/低延迟/低开销设计优化,在芯片间探索拓扑/路由/流控/通信库等的端到端 /全协议栈优化,在系统级探索应用驱动的算法/计算/通信的协同优化,超大规模互连系统设计验证工具等,报告将分享课题组在

approaches across different layers and dimensions. This includes optimizing within-chip interconnect micro-architecture for high bandwidth, low latency, and low overhead; pursuing end-to-end, full-protocol-stack optimizations between chips in areas such as topology, routing, flow control, and communication libraries; and

Additionally, tools for design verification of ultra-large-scale interconnect systems are critical. This presentation will share some of the research group's recent achievements in this field.

跨域资源调度、应用服务化封装等多个角度阐述相关技术进展和工程落地情况;在运营层面,提出以互联网化思维重构算力运营模式,通过电商式应用商城、开发者社区等创新机制,构建开放共享的算力生态体系。报告面向典型领域,展示超算互联网的落地

The High-Performance Computing Internet integrates numerous computing centers across the country via computing network, establishing a unified computing service platform. This initiative aims to coordinate the

王宇博士,鹏城实验室研究员。本科毕业于四川大学无线电系,比利时天主教鲁汶大学硕士学位,德国慕尼黑工业大学博士学位。 曾在德国生物信息学和系统生物学研究所(MIPS),亥姆赫兹慕尼黑研究中心,慕尼黑工业大学和莱布尼兹超算中心任职。科研兴趣:1.超大规模AI模型在生物医药的应用;2.可解释的人工智能xAI。

随着高通量基因组、转录组、蛋白组、代谢组等多组学技术的爆发式发展,生命科学研究正由"假设驱动"全面转向"数据驱动"。

PB 级生物数据的产生,使生命科学相关计算迅速成为国家算力网的核心应用之一。本报告概述了我国国家算力网(China National Computing Network)在生物信息学工作流、容器化调度、GPU 加速深度学习及隐私保护联邦学习等方面的最新部署,并指出四大关键挑战:跨领域数据标准化、面向混合精度计算的生命组学数据分析、跨机构数据共享的可扩展安全模型,以及

学会副理事长等职。主要研究方向为人工智能和大数据分析技术在网络空间安全领域中的应用;作为课题负责人承担和主持了国 家级重大/重点项目20余项;获国家科技进步二等奖5项;发表进入SCI和EI检索的论文320余篇,出版专著8部,获得120余项发明

设施安全国家工程研究中心常务副主任,中国电信研究院副院长(二级正),曾任中国电信上海研究院院长,中国电信信息安全部、创新业务事业部总经理等职务。以主要完成人身份参与的项目获得国家级管理创新一等奖一项,二等奖两项,省部级科技进

重点产品开发项目,牵头组织申办成功云网基础设施安全国家工程研究中心落户中国电信,总体负责过中国电信的信息安全管理工作,获得全国首届网络信息安全杰出人才奖。 Li Anmin, male, born in December 1969, is a member of the Communist Party of China. He holds a doctoral degree

体,构建了新一代算力安全范式。顺应国家一体化算力网建设与数据互联互通的战略布局,中国电信基于"息壤"平台,积极探索可信智算的创新实践。通过构建覆盖算力基础设施、平台、数据和应用全生命周期的一体化安全防护体系,确保算力资源的可信、可控与高效流转。本次分享将深入剖析实践中的关键创新与突破,并探讨其如何协同网络安全、数据安全与Al安全,共同筑牢数字

Duxuetao is an expert with State Council special allowance for her outstanding contributions. Engaged in long-term

算力网络面临泛在接入、异构调度、存算分离和服务质量保证等新特点下的安全挑战。而网络安全靶场是用于网络空间安全研

安全挑战,结合靶场灵活配置的的仿真能力,可应对海量应用场景的独特需求,充分发挥其作为安全基础设施的价值,开创算力 网络未来安全发展之路。

曹畅,博士后,教授级高级工程师,国家网信创新人才支持项目入选者,中国联通网络通信首席专家,联通研究院下一代互联网研究部总监,北京邮电大学兼职教授。目前担任中国通信学会(CIC)算力网络委员会委员,信息通信网络技术委员会委员,中国通信标准化协会网络路由与传输协议工作组(CCSA TC3 WG5)副组长,工信部IMT-2030推进组算力网络研究子组组长,中国计算机

学会 (CCF) 分布式计算与系统专委会执行委员,《信息通信技术》期刊编委。曾多次获中国通信学会科技进步一等奖,世界互联网大会领先成果奖等奖励,主要专业领域为IP数据通信、SDN/NFV、下一代互联网与算力网络等。
Cao Chang, Postdoctoral Fellow, Director of the Next-Generation Internet Research Department at China Unicom

王继刚,博士,研究员,科技部"十四五"国家重点研发计划指南专家,工信部高质量专项总体组专家。中兴通讯副总裁,安全产品 总经理。长期从事操作系统、网络安全,及人工智能等领域的技术研究和产品开发工作。近年来先后获得省部级奖励10余项,发表 学术论文30余篇,获得国内外专利授权20项,软件著作权9项。 Wang Jigang, Ph.D., researcher, expert in the "14th Five

随着算力网络成为国家数字化发展的核心基础设施,其面临的安全威胁日益复杂化和全域化。传统外挂式安全机制难以应对跨域攻击链和资源动态调度带来的新型风险。本报告提出算力网络内生安全架构,通过统一身份认证、弹性泛在安全边界、原子化

面对"五全安全"的目标和需求(全客户、全网络、全数据、全流程、全场景),针对其中关键的算网安全,从当前能力侧"全范式"中相对薄弱的第三范式着眼,阐述第三范式平行仿真技术对于安全整体防御体系的影响和思考。

完,覆盖网络安全、信息安全、重要通信、两化融合安全和数据安全等研究方向,统筹网络安全技术、产业、政策和网络空间国际治理等方面科研布局、学科建设和人才培养等工作。负责多项重大科研项目研究工作,包括国家科技重大专项03专项、国家重点研发计划网络空间安全专项、国家信息安全专项、工业互联网创新发展工程专项等重大项目,多次获得国家级、省部级奖项,相关成

近年来,人工智能大模型技术取得突破性进展,人工智能推理需求将极大扩展。用户需要接入优质高效的人工智能模型,同时也需要随时、随地、随需、低成本的算力资源支持和安全的算力服务。全球开展算力网、算力并网、分布式计算网络等多种算力互联成网技术研究,通过算力互联网推动优化配置计算资源,提升服务效率,满足人工智能、科学计算等任务实时调用需求,成为算力

杨松,深圳华为云计算技术有限公司云安全CTO/华为云安全运营中心(SOC)主任。20+年深耕华为安全技术领域,其中近10年从事华为云安全领域工作,主要聚焦方向:1、安全系统架构设计与创新;2、安全工程能力体系构筑;3、安全运营中心体系构建。期望

算网安全的第三范式思考

.CCF). He has published over 80 papers in international journals recommended as Category A/B b

报我们团队面向复杂拥塞控制场景下,对可编程高速网络分析技术进行的初步研究。

《国防科技大学学报》、《计算机工程与科学》等期刊编委。

面向超算与智算的高性能互连可扩展性设计

planning and formulation of a series of standards for AI computing centers and computing power networks.

CCF优博、首届中国科协青托工程、首届ACM SIGOPS China新星奖等荣誉,获省部级奖两项

Optimizing inference engine for large MoE language models: experience and lessons

in 1985. The University of Essex awarded him "Honorary Degree of Doctor of the University " in July 2024

appointed by Beijing City honorary chairman of "Global Digital Economy Cities Alliance (DEC40)"

Ministry. In 2019, he was named on the list of "中国海归70年70人" by CCG (全球化智库)。

doctoral supervisor in Zhejiang University, BJUPT, and NJUPT.

Honour" in 2023, and "Distinguished Contribution Award" in 2024.

奖1项,何梁何利科技进步奖,获得首届中国存储终身成就奖。

人工智能大模型推理架构的技术挑战

pioneer in the field of Generative AI.

Co-GenAI: A Novel Fusion-Driven Platform

Lessons learned in building an GUI-base AI agent

open challenges in real deployment in a large cloud.

 $Provincial \, Committee, \, Chairman \, of \, Guang dong \, Artificial \, Intelligence \, and \, Robotics \, Industry \, Alliance.$

of the computing continuum and serve multiple use-inspired domains.

with the National University of Singapore, both as a full professor.

Pushing AI from Research to Application and Impact

系统、大数据处理系统以及大模型训练系统等。

王曦,中国科学院院士,广东省委常委,广东省人工智能与机器人产业联盟理事长

陈文光 Wenguang Chen

Xi Wang

Weimin Zheng

Hongxia Yang

Qian, Ling

TAN Kun

Yaodong Yang

Computing, AI and 3-Body computing constellation: endless frontiers and exploration

Engineering from the University of Stuttgart, Germany. He has been an invited keynote speaker at SC' 07 in Reno, NV, USA and at many international conferences. Prof. Resch was awarded an honorary professorship and honorary

Michael Resch

Wen GAO Educational Talents". National Computing Power Sovereignty and "China Computing Net" Project

Wen GAO is a member of Chinese Academy of Engineering, ACM Fellow and IEEE Fellow. He is the founding director of Pengcheng Laboratory (Shenzhen, China). He is also a Boya Chair Professor at Peking University. He is currently a deputy to the 14th National People's Congress. He used to be a member of the 10th, 11th and 12th CPPCC National

Computing power sovereignty refers to a nation's ownership and control over the core computational resources through autonomous computing power and models as the foundation. Simultaneously, it presents the "China Computing Net" initiative—an innovation plan currently being vigorously promoted by Pengcheng Laboratory in collaboration with strategic scientific and technological forces—along with its latest progress, aiming to foster the

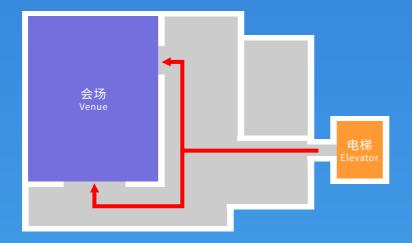
Prof. Michael Resch is the director of the High Performance Computing Center Stuttgart (HLRS). He holds a

the Prize for Scientific and Technological Progress of Ho Leung Ho Lee Foundation (2022), the Outstanding Contribution Award of Guangdong Province (2021), the CCF Wang Xuan Award and the title of "2005 China's Top Ten

Federation and the chief editor of Chinese Journal of Computers. He earned seven State Awards in Science and Technology Achievements as the first accomplisher. He received the IEEE Innovation in Societal Infrastructure Award (2025), National May 1 Labor Medal (2023), Wu Wenjun Al Highest Achievement Award (2023), the Special



会场指引 | Venue Guide Map



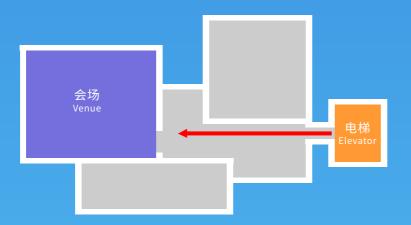
7楼/7F

08:30-12:10 开幕式与主论坛 Opening Plenary & Main Forum

14:30-17:45

分论坛1:大模型训练与推理 Track#1: Large Language Model

Training and Inference



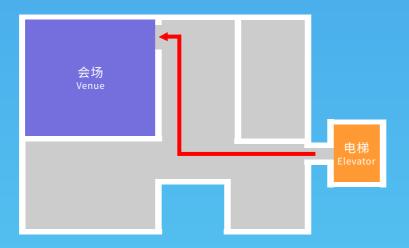
6楼/6F

14:30-17:45

分论坛3:算力网与安全前沿论坛

Track #3: Computing-power

Network and Security



3楼/3F

14:30-17:45

分论坛2:算力网云际计算前沿论坛 Track #2: Computing-power Network and JointCloud