

第 15 届 ChinaSys 会议手册

会议时间：2018. 12. 07-2018. 12. 08

会议地点：长沙世纪金源大饭店合肥厅

Keynotes (Chair: 董德尊)

December 7th, 2018 (Morning)

08:30-08:40	Opening (陈文光)
08:40-09:20	Keynote 1 Speaker: Kai Lu (National University of Defense Technology) Title : <i>E 级高性能计算机设想</i>
09:20-10:00	Keynote 2 Speaker: Shouyi Yin (Tsinghua University) Title: <i>Embedding AI in Everything : μW-level Neural Network Processor</i>
10:00-10:20	Tea Break

卢凯

简介：卢凯，国防科技大学计算机学院教授、副院长，博士生导师，科技部中青年领军人才、教育部新世纪优秀人才支持计划人选。国防科大系统软件方向学术带头人，长期从事大规模并行系统软件研制，是银河系列高性能计算机、天河一号和天河二号的副总设计师。首次提出和实现多实例操作系统架构、高性能虚拟域等多项技术创新，成果均达到同期世界前列，保证了银河、天



河高性能计算机的高效率、高安全和易管理，为天河高性能计算机系统排名世界第一起到关键的技术支撑。获国家科技进步特等奖 1 项（4）、国家科技进步一等奖 2 项（10、14）、军队科技进步一等奖 3 项（1、1、7），湖南省科技进步一等奖 1 项（5），获得国家发明专利 1 项、国防发明专利 4 项，软件著作权 2 项，发表论文 100 余篇。获“求是”奖、中创软件人才奖、中国青年科技奖等。兼任 ACM 长沙分部主席、CCF 长沙主席等职。

题目：E 级高性能计算机设想

尹首一

简介 : Dr. Shouyi Yin is associate professor (Tenured) and vice director of Institute of Microelectronics in Tsinghua University. His research interests include reconfigurable computing, domain-specific reconfigurable architecture design and high level synthesis. He has published more than 100 journal papers and more than 50 conference papers. He has received Second Prize of China's State Technological Innovation Award (2015), China's Patent Golden Award (2015), First Prize of Technological Innovation Award of Ministry of Education, China (2014), Second Prize of Science and Technology Advancement Award of Jiangxi Province, China (2014) and Best Paper Award in China Communications IC Technology and Application Conference (2011).



Dr. Shouyi Yin is the Secretary-General of EDA Chapter in Chinese Institute of Electronics. He is also the technical committee member of Asia Pacific Signal and Information Processing Association. Dr. Shouyi Yin has been served as program committee member and organizer in the top VLSI and EDA conferences such as A-SSCC, DAC, ICCAD and ASPDAC.

题目 : Embedding AI in Everything : μ W-level Neural Network Processor

摘要 : Deep neural networks (DNNs) have achieved great success in many applications of artificial intelligence (AI). To enable always-on and pervasive AI applications in mobile and IoT devices, ultra-low power neural network processors are required. With the progresses of both neural network algorithms and computing architectures, it is possible to design μ W-level neural network processors. In this talk, we introduce Thinker processors which have the potential to embed AI in everything.

Session 1: Storage (Session Chair: 包云岗) December 7th, 2018 (Morning)	
10:20-10:40	Yibo Zhu (ByteDance) <i>Accelerating Multi-Tenant Storage Systems with Group-Based RDMA (SIGCOMM 2018)</i>
10:40-11:00	Lingfang Zeng (Huazhong University of Science and Technology), Xi Li (DDN/Whamcloud), Wen Cheng (Huazhong University of Science and Technology), Andre Brinkmann (Johannes Gutenberg University Mainz) and Dan Feng (Huazhong University of Science and Technology) <i>I/O Quality-of-Service for Lustre-based Exascale High Performance Computing</i>
11:00-11:20	Hao Fan (Huazhong University of Science and Technology), Song Wu (Huazhong University of Science and Technology), Shadi Ibraim (Inria, IMT), Ximing Chen (Huazhong University of Science and Technology) and Hai Jin (Huazhong University of Science and Technology) <i>NCQ-Aware I/O Scheduling for Conventional Solid State Drives</i>
11:20-11:40	Yunshan Tu, Jiajia Chu, Yao Zhang and Chuliang Weng (East China Normal University) <i>A Customized Distributed Storage Architecture for High-performance Scenarios</i>
11:40-12:00	Jian Zhang, Yuzhuo Jing, Guanzhou Hu and Shu Yin (ShanghaiTech University) <i>RosFS: A Parallel Data-indexing File System</i>
12:00-13:30	Lunch (合肥厅)

朱亦博

简介 : Yibo is a senior research scientist at AI Lab, ByteDance. Prior to joining ByteDance, he was a researcher at Microsoft Research, Redmond. Yibo received his PhD degree from UCSB in 2016 and his bachelor degree from Tsinghua University in 2011. Yibo's research is focused on datacenter networks and systems. In the research community, Yibo is known for exploring new networking hardware capabilities to improve the system performance and reliability. Examples include some of the earliest-in-the-community work on large-scale RDMA, 60GHz wireless, network switch telemetry, etc. Yibo is a recipient of Microsoft Research Fellowship (2015) and MSR Redmond Labs Exemplary Collaboration Award (2017), authored 10+ SOSP/SIGCOMM/NSDI/MobiCom papers, and served on SIGCOMM'18/CoNEXT'18 TPC.

Personal page: <http://yibozhu.com>.

题目 : Accelerating Multi-Tenant Storage Systems with Group-Based RDMA

摘要 : Storage systems in data centers typically perform replicated transactional operations for high data availability and integrity. Today, however, such operations suffer from high tail latency even with recent kernel bypass networking optimizations. We observe that the root cause is the involvement of the CPU in the critical path of replicated transactions. To this end, we propose HyperLoop, a new framework that removes CPU from the critical path of replicated transactions by offloading them to commodity RDMA NICs. To achieve this, we develop new and general network primitives that can perform memory operations on all nodes in a replication group while guaranteeing ACID properties without CPU involvement. Microbenchmark and application benchmark show that HyperLoop can reduce 99th percentile latency by ~800× with nearly 0% CPU consumption on replicas.



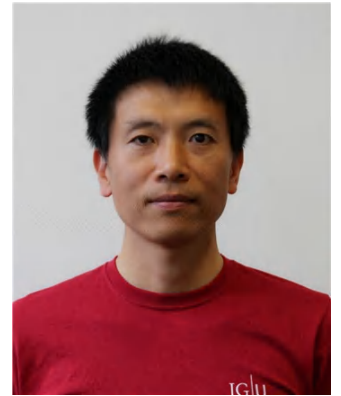
曾令仿

简介：曾令仿博士，华中科技大学武汉光电国家研究中心副教授，德国美因茨大学数据中心 Visiting Professor。主要从事高性能存储系统、大数据与计算智能等研究工作。曾在新加坡国立大学从事 4 年研究工作。在 FAST、SC、TPDS、ToS 等 CCF 推荐的 A 类国际会议和期刊上发表多篇学术论文。2006 年获 IEEE/ACM 超级计算机大会 (SC) 举办的 SC06 高性能存储挑战赛 Finalist Award (排名第一)。2011 年“基于主动对象的海量存储系统与技术”获湖北省技术发明一等奖 (排名第四)。CCF 高级会员，IEEE 及 ACM 会员。

个人主页：<https://lingfangzeng.github.io/>

题目：基于 Lustre 的艾级高性能计算 I/O 服务质量

摘要：报告将系统地介绍 Lustre 系统 I/O 服务质量方面的成果，研究发表于 IEEE/ACM 超级计算机大会 (SC'17)，Lustre 用户峰会 (LUG'17)，和 Lustre 管理和开发者大会 (LAD'18)，部分代码已被纳入 Lustre 正式发行版本中，剩余代码也被计划纳入 Lustre 2.13/2.14 版本。



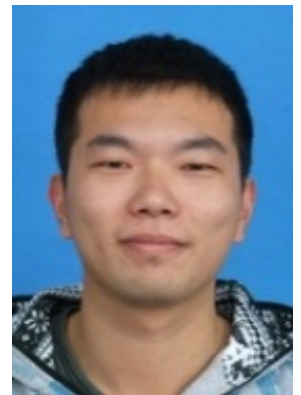
樊浩

简介：樊浩，华中科技大学在读博士，主要研究方向为 I/O 虚拟化。

联系方式：15527367311，邮箱：fanh@hust.edu.cn

题目：NCQ-Aware I/O Scheduling for Conventional Solid State Drives

摘要：现有的公平性 I/O 调度算法忽视了 NCQ 的影响。讲者通过实验分析发现 NCQ 竞争和低 NCQ 使用率会影响并发负载的公平性和设备使用率。讲者设计了一种调度框架。此框架可以感知并限制引起 NCQ 竞争和低 NCQ 使用率的负载。在保证高设备利用率的同时提升了系统公平性。



涂云山

简介：涂云山，华东师范大学研究生在读，研究方向为并行与分布式系统。

联系方式：13636370649，邮箱：yunshantu@stu.ecnu.edu.cn

题目：A Customized Distributed Storage Architecture for High-performance Scenarios

摘要：针对新型存储和网络硬件的特点，我们提出一种轻量化的定制分布式数据存储架构，消除传统分布式存储软件栈在新硬件下的过高开销，满足大数据处理应用对分布式数据访问的低时延和高吞吐性能需求，基于 NVMe SSD 和 RDMA/DPDK 等硬件的原型系统验证了该架构的有效性。



胡冠洲

简介：胡冠洲，上海科技大学 2016 级（大三）本科生，主要研究方向为并行与分布式存储系统、操作系统和高性能计算。并行与分布式存储系统实验室（LION Group）研究员，联系方式：15052208971，邮箱：huguanzhou123@gmail.com

题目：RosFS: A Parallel Data-indexing File System

摘要：应用正在产生越来越多的标签化数据。不同标签的数据被杂糅在一起存入传统的文件系统中，而频繁读取使用时却又往往需要提取特定标签的数据，这样一来数据的本地性就不能发挥作用了。我们初步实现了 RosFS 文件系统中间层，对应用而言提供了一个分析器用于分割标签数据，对后端而言将不同标签的数据聚类地、结构化地存入并行文件系统。通过写入时的一次拆分，RosFS 使后续的文件打开和读取无需遍历、索引的过程，提高了读性能（未投稿和发表）。



Session 2: Processor and Memory (Session Chair: 曾令仿)

December 7th, 2018 (Afternoon)

13:30-13:50	Zihao Yu, Zhigang Liu, Yiwei Li, Bowen Huang, Sa Wang, Ninghui Sun and Yungang Bao (Institute of Computing Technology, Chinese Academy of Sciences) <i>基于 Labeled RISC-V 的芯片敏捷开发实践</i>
13:50-14:10	Yu Ji (Tsinghua University), Youyang Zhang (Tsinghua University), Xinfeng Xie (University of California, Santa Barbara), Shuangchen Li (University of California, Santa Barbara), Peiqi Wang (Tsinghua University), Xing Hu (University of California, Santa Barbara), Youhui Zhang (Tsinghua University) and Yuan Xie (University of California, Santa Barbara) <i>FPSA: A Full System Stack Solution for Reconfigurable ReRAM-based NN Accelerator Architecture (ASPLOS 2019)</i>
14:10-14:30	Quan Chen and Minyi Guo (Shanghai Jiao Tong University) <i>Bandwidth and Locality Aware Task-stealing for Many-core Architectures with Bandwidth-Asymmetric Memory (TACO 2018)</i>
14:30-14:50	Fang Zhou (Huazhong University of Science and Technology), Song Wu (Huazhong University of Science and Technology), Xiang Gao (Huazhong University of Science and Technology), Hai Jin (Huazhong University of Science and Technology) and Jinglei Ren (Microsoft Research) <i>Dual-Page Checkpointing: An Architectural Approach to Efficient Data Persistence for In-Memory Applications (TACO 2018)</i>
14:50-15:10	Yaocheng Xiang (Peking University), Chencheng Ye (Huazhong University of Science and Technology), Xiaolin Wang (Peking University), Zeyu Wang (Peking University), Yingwei Luo (Peking University) and Zhenlin Wang (Michigan Technological University) <i>On Modelling Memory Bandwidth Contention to Improve Throughput and Fairness on Commodity Multicore Processor</i>
15:10-15:30	Tea Break

余子濠

简介：余子濠，中科院计算所在读博士生，主要研究方向为计算机系统结构。参与国家重点研发计划“软件定义云计算基础理论与方法”项目，是 Labeled RISC-V 项目的核心负责人。

联系方式：18925083352，邮箱：yuzihao@ict.ac.cn

题目：基于 Labeled RISC-V 的芯片敏捷开发实践

摘要：报告将分享 Labeled RISC-V 项目中的如下经验：1) 对指令集的选择；2) Chisel 语言对项目敏捷开发的帮助；3) 与传统开发的效率和质量进行对比，通过这些经验展示 RISC-V、Rocket Chip 和 Chisel 是如何推动芯片敏捷开发的。



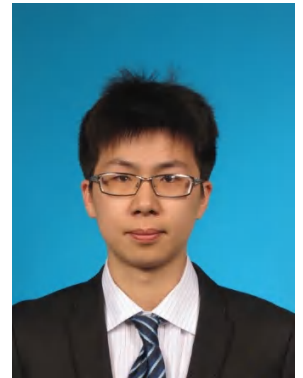
季宇

简介：季宇，清华大学在读博士生，主要研究方向为体系结构和神经网络。在国际期刊会议发表论文 6 篇，其中第一作者发表 CCF A 类 4 篇。

联系方式：18810301432，邮箱：jiy15@mails.tsinghua.edu.cn

题目：FPSA: A Full System Stack Solution for Reconfigurable ReRAM-based NN Accelerator Architecture

摘要：该论文发表在 Asplos19 上，通过软硬件结合，按照 RISC 的设计思想设计的基于 ReRAM 的 NN 芯片架构，相比现有的其他基于 ReRAM 的架构有 1000 倍的性能提升。



陈全

简介：陈全博士现为上海交通大学计算机科学与工程系 Tenure-Track 特别研究员。主要研究方向包括计算机系统、计算机体系结构、数据中心资源管理等。陈全已发表英文专著 1 部、论文超过 30 篇，多数发表在 Science (特刊)、IEEE Trans、ACM Trans 等著名期刊以及 ASPLOS、ISCA、IPDPS、ICS 等计算机系统领域顶级会议上。其研究在相关领域得到广泛关注，Google Scholar 引用超过 1500 次。陈全曾获 2015 年 CCF 优博、2016 年上海市优博、2017 年教育部自然科学一等奖 (排名: 3)、2018 年上海市技术发明奖一等奖 (排名: 6)、2017 年 IEEE TCSC Award for Excellent (Early Career Researcher) 等多项荣誉，并于 2017 年入选微软铸星计划。

题目：Bandwidth and Locality Aware Task-stealing for Manycore Architectures with Bandwidth-Asymmetric Memory

摘要：Parallel computers now start to adopt Bandwidth-Asymmetric Memory architecture that consists of traditional DRAM memory and new High Bandwidth Memory (HBM) for high memory bandwidth. However, existing task schedulers suffer from low bandwidth usage and poor data locality problems in bandwidth-asymmetric memory architectures. To solve the two problems, we propose BATS, a task scheduling system that consists of an HBM-aware data allocator, a bandwidth-aware traffic balancer, and a hierarchical task-stealing scheduler. Leveraging compile-time code transformation and run-time data distribution, the data allocator enables HBM usage automatically without user interference. According to data access hotness, the traffic balancer migrates data to balance memory traffic across memory nodes proportional to their bandwidth. The hierarchical scheduler improves data locality at runtime without priori program knowledge. Experiments on an Intel Knights Landing server that adopts bandwidth-asymmetric memory show that BATS reduces the execution time of memory-bound programs up to 83.5% compared with traditional task-stealing schedulers.



周放

简介：周放，华中科技大学在读博士，主要研究方向为：新型非易失性存储器。
联系方式：18627077360，邮箱：cnzhoufang@qq.com

题目：Dual-Page Checkpointing: An Architectural Approach to Efficient Data Persistence for In-memory Applications (TACO 2018)

摘要：本文探索了一种新的内存数据持久化方案，利用高频率、全系统、硬件级 checkpoint 技术，保证应用执行的 Crash Consistency，提供内存数据持久化，以降低系统复杂性与软件开销，大幅提升应用性能。提出了新的 checkpoint 技术 Dual-Page Checkpointing，打破了传统检查点技术中的 Tradeoff，同时实现了最小化 NVM 写次数与最少的元数据开销。



向耀程

简介：向耀程，北京大学在读博士生，主要研究方向为计算机系统结构。曾以第一作者发表过 Eurosys 一篇。
联系方式：18810705833，邮箱：jsbyxyc@me.com

题目：On Modelling Memory Bandwidth Contention to Improve Throughput and Fairness on Commodity Multicore Processor

摘要：该工作目前还处在实验阶段尚未投稿，希望得到与会专家们的批评指正。



Session 3: GPU (Session Chair: 沈立)**December 7th, 2018 (Afternoon)**

15:30-15:50	Chen Li (National University of Defense Technology, University of Pittsburgh), Rachata Ausavarungnirun (Carnegie Mellon University, King Mongkut's University of Technology North Bangkok), Christopher J. Rossbach (University of Texas Austin, VMware Research), Youtao Zhang (University of Pittsburgh), Onur Mutlu (ETH Zurich), Yang Guo (National University of Defense Technology), Jun Yang (University of Pittsburgh) <i>A Framework for Memory Oversubscription Management in Graphics Processing Units (ASPLOS 2019)</i>
15:50-16:10	Xiuhong Li (Peking University), Yun Liang (Peking University), Shengen Yan (SenseTime), Liancheng Jia (Peking University) and Yinghan Li (SenseTime) <i>A Coordinated Tiling and Batching Framework for Efficient GEMM on GPUs (PPoPP 2019)</i>
16:10-16:30	Chen Zhang, Shijie Cao and Lidong Zhou (Microsoft) <i>Sparse Computing for Efficient AI Inference on Heterogeneous Platform</i>
16:30-16:50	Zhaonian Tan, Weixing Ji, Jianhua Gao, Yizhuo Wang and Feng Shi (Beijing Institute of Technology) <i>2D Partitioning and Hybrid Encoding of Sparse Matrix</i>
16:50-17:10	Zhen Zheng, Jidong Zhai and Wenguang Chen (Tsinghua University) <i>CPU-GPU 异构平台上流水线程编程框架 (ASPLOS 2019)</i>
17:10-18:00	Panel (青年学者的博士之路, Chair: 包云岗)
18:00-19:30	Dinner (福州 1 厅)
20:00-21:00	ChinaSys Council & Committee Meeting

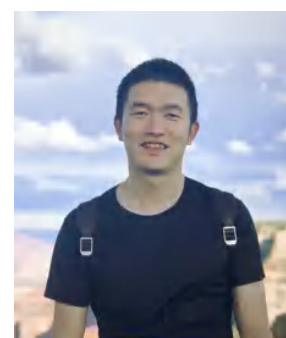
李晨

简介: 李晨, 国防科大在读博士生, 导师郭阳研究员, 研究方向 gpu 体系结构和片上网络。2016 年 10 月至 2018 年 10 月在美国匹兹堡大学联合培养, 导师 Jun Yang, Youtao Zhang。博士期间在 CCF 推荐的期刊会议上发表论文 4 篇, 包括一篇 A 类会议, 一篇 A 类期刊和两篇 B 类会议。

邮箱: lichen@nudt.edu.cn。

题目: A Framework for Memory Oversubscription Management in Graphics Processing Units

摘要: 主要介绍 GPU 统一虚拟内存和内存超额配置优化的相关工作。



李秀洪

简介：李秀洪，北京大学高效计算与应用中心五年级博士生，导师为梁云老师，主要研究方向为 GPU 代码优化和高性能计算。在国际期刊会议发表论文 9 篇，SCI 9 篇，其中第一作者发表 CCF A 类、B 类论文期刊 5 篇，先后参与过自然科学基金项目（高效图像与视频处理技术：理论实现和应用研究）、北京大学数字媒体所 863 项目（4K 视频编码标准 AVS2 的高性能实现）和商汤科技公司卷积运算核心代码库设计工作。

联系方式：15536090642，邮箱：lixihong@pku.edu.cn。

题目：面向 GPU 的基于分块和批处理的矩阵乘法优化技术

摘要：该技术通过优化批执行方案，可以显著提高小矩阵乘法的计算效率，从而克服现实应用中诸多小矩阵乘法无法充分发挥 GPU 计算能力的问题。



张宸

简介：张宸，微软亚洲研究院研究员，主要研究方向为“面向人工智能的高性能计算”。在国际期刊会议发表论文 8 篇，包括 AAAI, FPGA, DAC 等。曾获 FPGA 2015 Best Paper Candidate, ACM FPGA Library Most Cited Articles (第二名)。

联系方式：18811464662，邮箱：zhac@microsoft.com。

题目：Sparse Computing for Efficient AI Inference on Heterogeneous Platform

摘要：介绍我们提出的一种新颖的神经网络稀疏化方法，相关论文发表在 FPGA 2019 和 AAAI 2019。



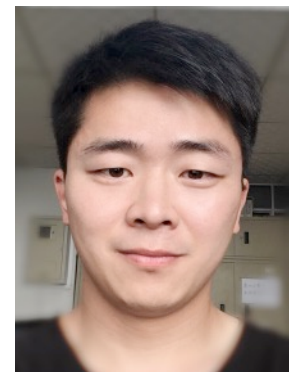
谈兆年

简介：谈兆年，北京理工大学在读硕士生，主要研究方向为高性能计算。

联系方式：15501161082，邮箱：tan_zn@163.com

题目：2D Partitioning and Hybrid Encoding of Sparse Matrix

摘要：稀疏矩阵广泛应用于科学和工程领域，本次报告主要介绍稀疏矩阵的划分和混合编码问题。我们充分考虑了矩阵数据布局，使用图像处理的方法对矩阵进行了 2D 划分，提升了 SpMV 在 GPU 上的运行效率。



郑祯

简介：郑祯，清华大学在读博士生，主要研究方向为 GPU 程序性能优化。在国际期刊会议发表论文 3 篇，其中第一作者发表 CCF A 类会议 2 篇，先后参与过国家 863 项目、高性能计算机评测项目等。

联系方式：18813080596，邮箱：zzchman@gmail.com

题目：CPU-GPU 异构平台上流水线程序编程框架

摘要：该工作主要介绍 GPU 上流水线编程优化方法及 CPU 与 GPU 之间流水线通信的优化。



Session 4: Deep Learning (Session Chair: 王蕾)

December 8th, 2018 (Morning)

08:20-08:40	Jian Ouyang (Baidu) <i>Baidu Kunlun - Make Compute More Intelligent</i>
08:40-09:00	Wencong Xiao (Beihang University & Microsoft Research Asia), Fan Yang (Microsoft) and Lidong Zhou (Microsoft) <i>Gandiva: Introspective Cluster Scheduling for Deep Learning (OSDI 2018)</i>
09:00-09:20	Xinhui Tian, Biwei Xie and Jianfeng Zhan (Institute of Computing Technology, Chinese Academy of Sciences) <i>Cymbalo: An Efficient Graph Processing Framework for Machine Learning</i>
09:20-09:40	Yemao Xu, Dezun Dong, Weixia Xu and Xiangke Liao (National University of Defense Technology) <i>SketchDLC: A Sketch on Distributed Deep Learning Communication via Trace Capturing (TACO 2019)</i>
09:40-10:00	Wanling Gao, Jianfeng Zhan and Lei Wang (Institute of Computing Technology, Chinese Academy of Sciences) <i>Data Motifs: A Lens Towards Fully Understanding Big Data and AI Workloads (PACT 2018)</i>
10:00-10:20	Tea Break

欧阳剑

简介: 欧阳剑，百度主任架构师，智能芯片总经理，对计算机体系结构和系统有一定的研究，曾在 ASPLOS 2014, Hotchips 2014/2016/2017 上发表过论文。

题目: Baidu Kunlun - Make Compute More Intelligent

摘要: 百度是国内最早大规模投入 AI 研究和产业化的企业，同时有非常多样化的 AI 应用，如智能云，自动驾驶，DuerOS 等。百度 XPU 是在多样化应用驱动下，经过多年迭代的 AI 处理器架构，本次报告将给大家分享基于 XPU 架构的百度昆仑芯片-面向云端的全功能 AI 芯片。



肖文聪

简介: 肖文聪，微软亚洲研究院和北京航空航天大学博士联合培养项目五年级博士生，主要研究方向为分布式机器学习系统平台。在国际期刊会议发表论文 11 篇，其中第一作者发表 OSDI 和 NSDI 各 1 篇。

联系方式: 18611201750，邮箱：xiaowencong@gmail.com。

题目: Gandiva: Introspective Cluster Scheduling for Deep Learning

摘要: 本文发表在 OSDI2018，本文介绍了一个全新的针对深度学习训练任务设计的 GPU 计算集群调度器，它提供一系列全新的调度机制设计来加速深度学习训练，并且显著提升了 GPU 集群利用率。常多样化的 AI 应用，如智能云，自动驾驶，DuerOS 等。百度 XPU 是在多样化应用驱动下，经过多年迭代的 AI 处理器架构，本次报告将给大家分享基于 XPU 架构的百度昆仑芯片-面向云端的全功能 AI 芯片。



田昕晖

简介：田昕晖，中科院计算所在读博士生，主要研究方向为分布式图计算。在国际期刊会议发表论文 9 篇，SCI 1 篇。

联系方式：13810295430，邮箱：tianxinhui@ict.ac.cn

题目：Cymbalo: An Efficient Graph Processing Framework for Machine Learning

摘要：介绍一种新型的面向机器学习算法的分布式图计算框架。



徐叶茂

简介：徐叶茂，国防科技大学在读博士生，主要研究方向为分布式深度学习的通信优化。

联系方式：15575961342，邮箱：xuyemaovip@nudt.edu.cn

题目：SketchDLC: A Sketch on Distributed Deep Learning Communication via Trace Capturing

摘要：本文提出通过分布式通信 trace 截取的方式对网络的通信特征进行分析，并利用 trace 计算在分布式训练过程中涉及的各种开销，分析开销与训练参数设置、集群规模等之间的关系，同时也为模拟实验的开展生成相应的通信负载。



高婉铃

简介：高婉铃，中科院计算所博士生，主要研究方向为数据中心计算、基准测试和体系结构。

联系方式：13126729336，邮箱：gaowanling@ict.ac.cn

题目：Data Motifs : A Lens Towards Fully Understanding Big Data and AI Workloads

摘要：本报告介绍了大数据分析和 AI 负载的统一计算抽象 data motif，基于统一计算抽象构建的适用于早期处理器设计验证的仿真基准测试集和支持大规模系统评测的可扩展基准测试集 BigDataBench 4.0。



Session 5: Data Center (Session Chair: 董德尊)
December 8th, 2018 (Morning)

10:20-10:40	Xingda Wei, Rong Chen and Haibo Chen (Shanghai Jiao Tong University) <i>Unifying Timestamp with Transaction Ordering for MVCC with Decentralized Scalar Timestamp</i>
10:40-11:00	Cheng Li (University of Science and Technology of China) <i>Language and System Support for Fine-grained Distributed Consistency (ATC 2018)</i>
11:00-11:20	Ran Shu (Microsoft), Peng Cheng (Microsoft), Guo Chen (Hunan University), Zhiyuan Guo (Microsoft Research & Beihang University), Lei Qu (Microsoft), Yongqiang Xiong (Microsoft), Derek Chiou (Microsoft Azure) and Thomas Moscibroda (Microsoft Azure) <i>Direct Universal Access: Making Data Center Resources Available to FPGA (NSDI 2019)</i>
11:20-11:40	Dongjie Tang, Yun Wang, Linsheng Li, Zhengwei Qi and Haibing Guan (Shanghai Jiao Tong University) <i>gRemote: Remoting Multi-tenant Cloud Workloads on GPU clusters</i>
11:40-12:00	Chen Zheng (Institute of Computing Technology, Chinese Academy of Sciences), Lei Wang (Institute of Computing Technology, Chinese Academy of Sciences), Jianfeng Zhan (Institute of Computing Technology, Chinese Academy of Sciences), Lixin Zhang (Institute of Computing Technology, Chinese Academy of Sciences) and Sally A. McKee (Chalmers University of Technology) <i>XOS: An Application-Defined Operating System for Datacenter Computing</i>
12:00-13:30	Lunch (合肥厅)

魏星达

简介：魏星达，上海交通大学在读博士生，主要研究方向为分布式数据系统，数据存储系统。在国际期刊会议发表论文 4 篇，其中第一作者发表 CCF A 类 3 篇，曾获奖 2018 MSRA Fellowship。

题目：Unifying Timestamp with Transaction Ordering for MVCC with Decentralized Scalar Timestamp

摘要：介绍如何为数据库提供高效且可扩展的多版本支持。



李诚

简介：李诚，德国马普学会软件系统所(MPI-SWS)博士，现为中国科学技术大学计算机科学与技术学院特任研究员，CCF《技术动态》编委。回国前曾任美国甲骨文公司瑞士苏黎世实验室 (Oracle Labs Swiss) 高级技术研究员，葡萄牙计算机系统工程研究所 (INESC-ID, Portugal) 访问研究员。李诚研究员一直从事分布式计算、一致性模型、系统性能优化、容灾容错等方面的基础理论和系统构建的研究工作。其成果发表于 OSDI, USENIX ATC, EuroSys, Hotstorage, DSN 等系统领域的知名国际会议上。曾担任 ACM TURC 2018 SIGOPS 会议程序委员会共同主席，ACM SOSP 2017 Poster Session 程序委员会共同主席。

题目：参数服务器架构的可扩展性研究



摘要：参数服务器架构 (Parameter Server Architecture) 是最常用的分布式深度学习计算架构，为 MxNet、Tensorflow 等主流深度学习系统所采用。然而，该架构的可扩展性受到参数模型体量大、参数同步开销成本高、计算与通信协同欠缺等问题的严重制约，从而使得深度学习模型的训练过程占用大量网络和计算资源，且收敛时延较高，无法满足业务的时效性要求。本报告将系统地阐述参数服务器架构存在的性能瓶颈以及初步的性能优化方案。

舒然

简介：舒然，微软亚洲研究院副研究员，主要研究方向为数据中心网络。舒然于 2018 年获得清华大学博士学位。他在国际期刊及会议发表论文十余篇。

题目：Direct Universal Access: Making Data Center Resources Available to FPGA

摘要：该论文已被录用，将发表于 2019 年 NSDI。本次报告的主要内容为：随着 FPGA 被广泛部署在数据中心的中心，FPGA 应用程序对访问数据中心中异构资源（如 CPU、GPU 和存储）的需求越来越高。基于现有的技术，FPGA 无法高效地访问这些资源。本研究为 FPGA 访问数据中心资源提供了一套统一而高效的通信体系结构，包含资源编址、路由和应用编程接口。



汤冬劫

简介：汤冬劫，上海交通大学计算机系在读博士生，主要研究方向为 GPU 虚拟化；曾在美国工作两年，后回国在上海兆芯集成电路有限公司任职架构师。曾参与 Elite2000（现应用在东方有限机顶盒中）以及 Elite2500（中国第一款手机 GPU 芯片）的研发。

邮箱：018033210001@sjtu.edu.cn

题目：GPU 指令迁移到远端服务器集群上的调度与应用

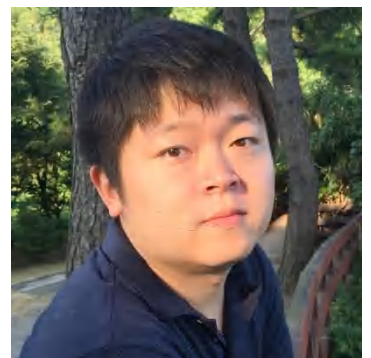


郑晨

简介：郑晨，中国科学院计算技术研究所，博士后。2017 年中科院计算所博士毕业，参与多项国家、中科院课题和华为 A 类项目，主要研究方向为操作系统，虚拟化，大数据，基准测试程序，性能优化，网络安全等。六项国内发明专利，多项国际专利。主持或参与了数据中心操作系统 (XOS, Rainforest) 和大数据基准测试程序 BigDataBench 的研发。

题目：XOS: An Application-Defined Operating System for Datacenter Computing

摘要：中心操作系统面临着性能、可扩展性和隔离性的等问题的挑战，尤其是在大数据和多核时代，数据中心应用高并发，多应用混部，深软件栈的情况下。本文介绍了 XOS，一种应用定义的操作系统。XOS 将资源管理从内核中分离，在用户态定制实现用户态资源管理服务，并弹性的划分节点硬件资源。从而在性能、隔离性和可扩展性上得到较大提升 (BigData 2018)。



Session 6: Security and Performance (Session Chair: 叶晨成)
December 8th, 2018 (Afternoon)

13:30-13:50	Peinan Li (Institute of Information Engineering, Chinese Academy of Sciences), Lutan Zhao (Institute of Information Engineering, Chinese Academy of Sciences), Rui Hou (Institute of Information Engineering, Chinese Academy of Sciences), Lixin Zhang (Institute of Computing Technology, Chinese Academy of Sciences), and Dan Meng (Institute of Information Engineering, Chinese Academy of Sciences) 安全依赖与条件推测：浅谈幽灵漏洞的硬件防御机制 (HPCA 2019)
13:50-14:10	Panlong Yang (University of Science and Technology of China) 人工智能驱动下的物联网—从有源到无源、从节点安全到体系安全 (MobiCom)
14:10-14:30	Jiaxin Li (National University of Defense Technology), Yuxi Chen (The University of Chicago), Haopeng Liu (The University of Chicago), Shan Lu (The University of Chicago), Yiming Zhang (National University of Defense Technology), Haryadi Gunawi (The University of Chicago), Xiaohui Gu (North Carolina State University), Xicheng Lu (National University of Defense Technology) and Dongsheng Li (National University of Defense Technology) PCatch: Automatically Detecting Performance Cascading Bugs in Cloud Systems (Eurosys 2018)
14:30-14:50	Zhenyu Guo (Ant Financial) A Kernel of Fuse Computing for Internet Services
14:50-15:10	Zichen Xu (Nanchang University), Christopher Stewart (The Ohio State University) and Jiacheng Huang (Nanchang University) An Elastic, Location-diverse Raft
15:10-15:20	Concluding Remarks

李沛南

简介：李沛南，中国科学院信息工程研究所信息安全国家重点实验室博士生二年级学生，主要研究方向为计算机体系结构安全，导师是侯锐研究员。在 CCF A 类会议 HPCA 2019 发表论文 1 篇。

联系方式：13161031941，邮箱：lpeinan@iie.ac.cn

题目：安全依赖与条件推测：浅谈幽灵漏洞的硬件防御机制

摘要：主要介绍我们关于幽灵漏洞防御的相关工作。在传统数据依赖与控制依赖基础上，我们首次提出“安全依赖”概念，刻画程序动态执行中潜在的安全隐患，进一步提出“条件推测”执行机制，有效防御推测执行漏洞。



杨盘龙

简介：杨盘龙博士，男，1977年8月出生。现为中国科学技术大学教授。江苏省杰出青年基金获得者。2005年毕业于解放军理工大学，获博士学位。南京大学博士后，香港科技大学访问学者。迄今为止已经在国际会议与期刊杂志累计发表文章80余篇。包括本领域最权威的国际期刊《IEEE Transactions on Networking》《IEEE Transactions on Mobile Computing》，《IEEE Transactions on Parallel and Distributed Systems》，《IEEE Transactions on Information Forensics and Security》，《IEEE Transactions on Wireless Communication》，《IEEE Transactions on Knowledge and Data Engineering》，《IEEE Transactions on Vehicular Technology》；以及顶级国际学术会议如 ACM Mobicom, MobiHoc, IEEE INFOCOM。申请发明专利8项。申请人作为负责人承担国家重点研发计划项目课题，国家自然科学基金面上项目，青年基金项目；作为子课题负责人参加重点项目。另外主持包括省自然科学基金和高技术研究项目，以及参加973，国家自然科学基金重大项目等。多年来担任 ACM MobiHoc, IEEE MASS 等国际会议的程序委员会委员，并担任 MSCC 2014 国际会议程序委员会区域主席, IEEE MASS, MSN 等国际会议分会主席。获得 ACM MobiCom 演示论文荣誉提名奖, ACM MobiHoc 最佳论提名奖, IEEE MSN 最佳论文奖, IEEE MASS 最佳演示论文奖。另外，获得军队科技进步三等奖3项。CCF 杰出演讲者（2015），发表 CCF A 类论文总计 29 篇。主要研究方向：无线网络；群智感知；无源感知与组网；无线射频成像。
邮箱: plyang@ustc.edu.cn



题目：人工智能驱动下的物联网 - 从有源到无源、从节点安全到体系安全

摘要：从 1991 年马克·维塞尔提出普适计算的概念后，经过 20 多年的努力，物联网已被认为是新一代信息技术的重要组成部分。然而，由于物联网节点能量的制约，大规模的物联网应用尚未出现。因此，如何解决物联网能量问题，保证网络持久稳定地工作，成了物联网实现普适应用的重要挑战问题。近年来，“从环境中获取能量”受到了人们的广泛关注。这些能量可以是来自光能、温差供能、运动能以及泛在的电磁能等。受这一理念的启发，无源感知网络中的节点通常不配备电池，而是通过采集环境中的能量进行计算、感知、通信和组网，以解决物联网的能量问题。无源感知技术能够利用普适的反射无线通信信号，对用户和环境进行感知，具有跨域协同一体化的特点；同时不再依赖传统的针对感知对象的专用传感器来完成感知任务，具有易用和低成本的特点，适合普适化应用。

报告将介绍团队最新的研究成果，主要包括无源网络平台方面的建设，以及系统设计方面的研究成果。首先，基于无源无线信号的智能感知实现了人体动作的高精度识别，能够利用更普适的一般无线信号反射功能实现精准感知，可用于智能教育系统，实现教师和学生行为的普适化部署和精准感控。其次，将介绍基于无源标签的追踪和定位系统，以及无源物联网协议栈，充电路由优化方面的成果。最后，将介绍团队在物联网安全和智慧城市方面的研究构想和思路。

李佳鑫

简介：李佳鑫，国防科技大学在读博士生（2018年12月毕业），导师是卢锡城教授。博士期间曾赴美国芝加哥大学访学两年，师从 Shan Lu 副教授。研究方向为分布式系统可靠性。
联系电话：18684847787，微信：xincafe，邮箱：licyh@nudt.edu.cn。

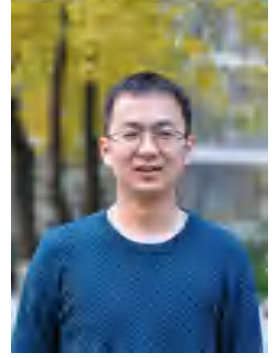
题目：PCatch:云系统中性能级联缺陷的自动检测方法

摘要：该报告的论文发表在 EuroSys2018，论文针对分布式性能级联缺陷（PCbug）违反系统的高可扩展性和性能隔离属性，并导致全局性能下降甚至威胁系统可用性的问题，提出了一个 PCbug 检测工——PCatch。PCatch 包含预测 PCbug 的三个关键组件：使用程序分析来识别执行时间可能会随着工作负载大小显著增加的代码区域；采用传统的 happens-before 模型来推理软件资源争用和性能依赖关系；使用动态跟踪来识别减速传播是否被包含在一个作业内部或之外。论文使用具有代表性的分布式系统 Cassandra、HadoopMapReduce、HBase 和 HDF 进行的实验评估表明，PCatch 可以在小规模的工作负载执行条件下高效、准确地预测 PCbug。



郭振宇

简介 : 郭振宇, 蚂蚁金服数据技术部资深技术专家, 发起并负责蚂蚁新计算引擎 arcOS [a:kəs] 的研发。研究兴趣包括分布式系统, 操作系统, 编译器, 以及机器学习系统等。在蚂蚁金服之前, 他在微软亚洲研究院系统组工作, 发表论文若干, 并发起了 rDSN 开源项目, 用于可靠分布式系统的快速搭建, 并在微软, 小米, 蚂蚁等公司实际部署。
联系方法 : imzhenyu (微信号)



题目 : ARCOS: A Kernel of Fuse Computing for Internet Services

摘要 : ARCOS 主要用于解决在复杂互联网服务内部如何融合不同的计算需求的问题, 该项目在过去的阿里双十一进行了部署并初步验证了可行性。

徐子晨

简介 : 徐子晨, 江西省双千, 计算机学会体系结构专委会委员, 俄亥俄州立大学博士, 南昌大学教授, 高层次人才, 学科带头人。主要研究内容包括数据密集计算, 智能计算, 高性能计算及分布式数据存储等方面的教研工作。以第一作者发表数据库、分布式系统及节能计算等方向旗舰期刊会议文章 20 余篇 (CCF-A/B), 是 VLDB/TC/TPDS 等期刊的文章审稿人及助理编辑, 曾服务于 ICAC, IWQoS, ICDCS 等国际会议的 TPC 及 TPC Chair 相关工作。学习工作期间, 完成多项基于 Google Cloud、Amazon AWS 等云平台的存储系统设计, 独立完成基于亚马逊云计算平台的高可靠性分布式弹性缓存系统设计及实现, 该系统目前运行于亚马逊底层存储服务中。获得 10 余项国家自然科学基金、教育部协同育人计划、江西省自然科学基金及微软、亚马逊等国内外纵、横向基金资助。

题目 : An Elastic, Location-diverse Raft

摘要 : Raft is a networked protocol used to synchronize data streams. It is widely used, especially by workloads that span geographically distributed sites. As these workloads grow, Raft's costs should grow proportionally. However, auto scaling approaches for Raft inflate costs by provisioning at all sites when one site exhausts its local resources. This paper presents Loa-Raft, an implementation that enables precise auto scaling. Loa-Raft extends Raft with the following abstractions: (1) secretaries which relieve log processing from leader and (2) observers which relieve read requests from followers. These abstractions are stateless, allowing for elastic auto scaling, even with unreliable spot instances. Loa-Raft preserves strong consistency guarantees provided by Raft. Loa-Raft improves goodput of 95th-percentile SLO by 9X.