

Preface



## Preface: Special Issue on Advances in System Software

Yanjun Wu (武延军)

(Institute of Software, Chinese Academy of Sciences, Beijing 100190, China)

Corresponding author: Yanjun Wu, yanjun@iscas.ac.cn

**Citation** Wu YJ. Preface: Special issue on advances in system software, *International Journal of Software and Informatics*, 2021, 11(2): 117–119. <http://www.ijsi.org/1673-7288/00245.htm>

System software, the software that manages the hardware and supports the applications in computing systems, is a core component of IT ecosystem that covers operating system, programming language, compiler, runtime environment, integrated development environment, etc. In industrial competition, system software is the “moat” and “accelerator” of CPU and the “soil” of APPs.

The prosperity of academic research on system software in one country often represents its leadership in core software and hardware technologies in the world. Compared with the prosperity of Internet software or smartphone applications, there is still a large gap in the development of system software in China. The research topics on system software usually come from large processor vendors and large application platforms. In the past, there were few research groups in Chinese universities and institutes on system software due to the lack of demand from domestic industrial giants. People have no choice but use the ISAs, micro-architectures, and platforms from international IT giants, e.g. IBM, Intel, and Qualcomm. Therefore, a virtuous cycle of “industrial demand–academic research–production application–new industrial demand” did not exist in China in past decades.

The situation has changed in recent years, and system software has become a hot area in both academia and industry for the first time in the history of China. One reason is the rise of domestic chip vendors. The vendors represented by Huawei and Cambricon have taken the lead in the design of smartphone chips and AI chips, and developed their own system software products, and even make them open source, such as openEuler, OpenHarmony, MindSpore, and BANG. Another reason is the expansion of business stack of Internet companies. Internet giants such as Baidu, Tencent, Alibaba, and ByteDance have begun to develop their own chips or invest heavily in the chip industry, and actively joined the game of open source on system software, e.g. Apollo autopilot platform and PaddlePaddle deep learning framework. The cutting-edge demands from chip vendors and Internet giants have ignited the interest of academic research on system software. Many research results have gained international influence, such as the virtualization, IPC/RPC, RDMA, and TEE conducted by Prof. Haibo Chen’s group (IPADS Lab) at Shanghai Jiao Tong University, the graph computing and distributed parallel computing conducted by Prof. Wenguang Chen’s group at Tsinghua University, and the NVM/SSD storage system conducted by Prof. Jiwu Shu’s group at Tsinghua University/Xiamen University.

In addition to being driven by the demands of domestic industries, some emerging fields that attracted continuous attention worldwide in recent years have also promoted system software in China.

The first is the AI accelerators and application frameworks for deep learning. AI accelerators such as Cambricon MLU series, Huawei Ascend, and Nvidia GPU need to be enabled by system software. AI frameworks need to be simplified in application programming and model optimization. To deal with these problems, the Institute of Software, Chinese Academy of Sciences (ISCAS) launched the FactOS and the Institute of Computing Technology, Chinese Academy of Sciences (ICT-CAS) introduced the BANG language. Meanwhile, many studies focus on model auto-selection and multi-level optimization, which lead to the AI compilers and toolchains, such as TVM/NNVM by Prof. Tianqi Chen from University of Washington, as well as the AutoML and MLIR from Google. These researches have achieved more efficient development of AI applications.

The second is RISC-V open Instruction Set Architecture (ISA). Hopefully RISC-V will drive system software over the next decades. It has been 11 years since the birth of RISC-V in 2010 by Prof. David Patterson and his colleagues at University of California, Berkeley, though only in the past three years RISC-V had been recognized as an important player among existing ISAs. RISC-V is an advanced, open-source, modular, and diverse instruction set design and implementation, which, just like Linux in 2000 (nearly 10 years after its birth by Linus Torvalds), is favored and embraced by many system researchers. RISC-V gives us the opportunity to “reinvent the wheel” and to break the barrier between software and hardware. With the progress in RISC-V ecosystem, one day the so-called “full-stack engineer” will be redefined, i.e. an engineer who is able to handle from software development to chip design. ISCAS is working on the porting and optimization of the core RISC-V system software, such as the Linux kernel, LLVM and GCC toolchains, JDK, and V8 engine. A large group headed by Prof. Yungang Bao at ICT-CAS is researching on the agile development of CPU and has launched the “One Student One SoC” program at University of Chinese Academy of Sciences, requiring undergraduates to complete a CPU design for graduation. Prof. Xuehai Zhou’s group at University of Science and Technology of China is trying to transfer the instruction sets of neural network processors to RISC-V extensions.

The third is open-source mode. Open source has become the mainstream of system software development. The popular system software products emerging in the last decade, such as Hadoop/Spark for big data and OpenStack/K8S for cloud computing, have mostly been developed in an open-source way. Open source is profoundly changing the paradigm of software development and reshaping the software industry. Behind a typical open-source operating system are usually tens of thousands of upstream open-source software, linked by their dependencies and function definitions, which is also a supply chain but made of software. The supply chain of open-source software is more complex than that of hardware devices such as cell phones and laptops. How to ensure the reliability of open-source software supply for large complex system software, is an interesting interdisciplinary topic of system software and software engineering. Prof. Minghui Zhou’s group from Peking University conducted extensive research on the sociological relationship of the open-source supply chain. The group led by Prof. Wei Wang at East China Normal University performed big data analysis and mining on GitHub projects. Prof. Ge Li’s group at Peking University and Prof. Xin Peng’s group at Fudan University deeply investigated the automatic completion and generation of source code. To explore the open-source software supply chain, ISCAS is now constructing an open-source software knowledge graph database and a platform for source code synthesis, validation, integration, and release, as well as risk assessment of supply chain software in terms of quality, license, and maintainability.

Hopefully this research will lead to an infrastructure for the open-source software supply chain and make open-source software into public goods, thereby ensuring the continuity and reliability of business and supporting innovative research.

Many research groups are engaged in system software in China, with many perspectives on the trend in system software, which are not detailed here for lack of space.

In this special issue, six representative papers focusing on the above topics were selected.

In the paper on “*Static Checking of Array Index Out-of-Bounds Defects in C Programs Based on Taint Analysis*”, a static checking method of array index out-of-bounds was proposed based on taint analysis and data flow analysis, to address the problem of low accuracy of static checking of array index out-of-bounds defects in the C language, and the automatic static analysis tool Cararraybound was implemented.

In the paper on “*Autonomous Learning System Towards Mobile Intelligence*”, a machine learning training paradigm (autonomous learning) for data privacy protection in mobile devices was proposed, and the autonomous learning system AutLearn was designed and implemented to improve the accuracy and speed of training on the devices.

In the paper on “*Accelerator Virtualization Framework Based on Inter-VM Exitless Communication*”, an accelerator virtualization framework that supports rapid delegated execution across virtual machines was proposed based on C/S architecture to improve the performance of AI accelerator virtualization, and a prototype system for NVIDIA GPU was implemented on the mainstream QEMU-KVM virtualization platform.

In the paper on “*Hybrid Access Cache Indexing Framework Adapted to GPU*”, a hybrid access cache indexing framework adapted to GPU was proposed to reduce memory access and bus transfer and improve the performance of traditional hash tables, which was implemented and experimentally verified on GPU hardware.

In the paper on “*Open Source Software Supply Chain for Reliability Assurance of Operating Systems*”, a knowledge graph-based method for open-source software supply chain management, as well as a reliability risk assessment and management method were proposed, which were applied to and verified with the construction of a mainstream Linux operating system distribution as an example.

In the paper on “*Neural Network Instruction Set Extension and Code Mapping Mechanism*”, a specific instruction set, RV-CNN, was designed based on RISC-V architecture for convolutional neural network, and a code-to-instruction mapping mechanism was proposed and verified on the FPGA platform.

The papers included in this special issue cover programming language, programming framework, virtualization, operating system, open-source software, artificial intelligence, and RISC-V instruction set. They reflect the achievements of Chinese scholars in these fields and can serve as a reference for researchers in system software.



**Yanjun Wu**, Ph.D., distinguished professor, deputy chief engineer, and director of Intelligent Software Research Center at Institute of Software, Chinese Academy of Sciences, doctoral supervisor, and senior member of CCF, is mainly engaged in the research on operating systems and system security. He has presided over key/major projects such as the Knowledge Innovation Project of the Chinese Academy of Sciences, the Strategic Priority Research Program of the Chinese Academy of Sciences, and the National Science and Technology Major Project. He had also been funded by Tencent University Cooperation Program for four consecutive years. He was awarded the Outstanding Graduate of Chinese Academy of Sciences, Beijing Science and Technology Rising Star, and Outstanding Member of Youth Innovation Promotion Association, the Chinese Academy of Sciences. He has published more than 50 papers, including those in internationally renowned conferences such as ASPLOS, EuroSys, ICSE, IJCAI, and AAAI, and applied for more than 30 patents and software copyrights.