# **Final Program**

## April 20, 2022 (Japan Standard Time)

#### 9:00-10:30 <u>Special Invited Lecture 1</u> Chair: Takatsugu Ono (Kyushu Univ.)

## 9:00-10:30 Closing the Gap between Quantum Algorithms and Machines with Hardware-Software Co-Design

Fred Chong (Univ. of Chicago and Super.tech, USA)

Abstract: Quantum computing is at an inflection point, where 127-qubit (quantum bit) machines are deployed, and 1000-qubit machines are perhaps only a few years away. These machines have the potential to fundamentally change our concept of what is computable and demonstrate practical applications in areas such as quantum chemistry, optimization, and quantum simulation. Yet a significant resource gap remains between practical quantum algorithms and real machines. A promising approach to closing this gap is to design architectural interfaces that selectively expose to programming languages and compilers some of the key physical properties of emerging quantum technologies. I will describe some of our recent work that focuses on techniques that break traditional abstractions, including compiling programs directly to analog control pulses, adapting programs for machine variations, computing with ternary quantum bits, efficient error correction using 2.5D structures, and designing just enough tunability in qubits to avoid crosstalk.

#### 10:30-11:00 Break

# **11:00-12:30Special Invited Lecture 2**<br/>Chair: Takatsugu Ono (Kyushu Univ.)

11:00-12:30 **Computer Architecture Challenges for the Security of Persistent Memory** *Yan Solihin (Univ. of Central Florida, USA)* 

> Abstract: Non-volatile or persistent memory (PM) is increasingly integrated into the main memory of computer systems. PM allows programmers to keep persistent data in data structures in memory instead of files in storage. Such structures can be wrapped into containers that we refer to as Persistent Memory Objects (PMO). While PMO allows fine-grain access at low latency to persistent data, it also presents several challenges. Among these challenges are: (1) new abstraction is needed to define its use, access, and sharing, (2) security vulnerabilities that arise from keeping them in memory instead of files, and (3) how it affects memory encryption and integrity verification. In this talk, I will give a broad landscape of computer architecture challenges with PM and discuss our approaches in solving these problems.

- 12:30-13:20 Lunch Time Break
- 13:20-14:10 Special Session I
- 13:20-14:10 <u>Keynote Presentation 1</u> Co-chairs: Tohru Ishihara (Nagoya Univ.), Takumi Uezono (Hitachi)

# Universal Chiplet Interconnect Express (UCIe): Poised to Change the Compute Landscape

Debendra Das Sharma (Intel, USA)

Abstract: High-performance workloads demand on-package integration of heterogeneous processing units, on-package memory, and communication infrastructure to meet the demands of the emerging compute landscape. Applications such as artificial intelligence, machine learning, data analytics, 5G, automotive, and high-performance computing are driving these demands to meet the needs of cloud computing, intelligent edge, and client computing infrastructure. On-package interconnects are a critical component to deliver the power-efficient performance with the right feature set in this evolving landscape. Universal Chiplet Interconnect Express (UCIe), is an open industry standard with a fully specified stack that comprehends plug-and-play interoperability of chiplets on a package; similar to the seamless interoperability on board with well-established and successful off-package interconnect standards such PCI Express®, Universal Serial Bus (USB)®, and Compute Express Link (CXL)®. In this talk, we will discuss the usages and key metrics associated with different technology choices in UCIe. We will also delve into the different layers as well as the software model associated with UCIe along with the compliance and interoperability mechanisms. We will also discuss how this open standard could potentially evolve to incorporate additional usage models in the future.

14:10-14:30 Break

#### 14:30-15:20 Special Session II

#### 14:30-15:20 <u>Keynote Presentation 2</u> Co-chairs: Akihiro Hashiguchi (Sony), Teruaki Sakata (Hitachi)

# RISC-V-based Parallel Processor IP with Vector Extension for Embedded Systems

Shotaro Shintani (NSITEXE)

Abstract: There are a variety of controls in automobiles, and the amount of processing is continuously on the increase due to the more complex algorithms. Moreover, the progress of autonomous driving has been remarkable in recent years, and the performance requirements have increased dramatically in addition to the diversification of processing. There are also traditional requirements in automotive systems, such as hard real-time performance, functional safety, power efficiency, software portability, and so on. On the other hand, there is the problem that software development will become more complex and less portable when several processors or accelerators are adopted to meet the automotive requirements. In order to solve it, we develop new processors with enough performance and flexibility in a single architecture that can handle the various types of processing for autonomous driving while also satisfying the traditional automotive requirements such as functional safety. In this presentation, our first-generation Data Flow Processor (DFP), DR1000C, is introduced, and it is targeted at safety-critical systems and delivers the real-time performance and high-throughput data processing required for microcontrollers for embedded systems. DR1000C is a RISC-V-based multiple instruction stream, multiple data stream (MIMD) processor with a vector processing unit. It also supports a hardware multi-threading mechanism that allows up to 16 threads which simultaneously share the vector processing unit, resulting in highly efficient resource use. Furthermore, DR1000C has the necessary functional safety modules and meets ISO 26262 ASIL B to D safety requirements without additional external special safety mechanisms. Our evaluation results show that DR1000C can efficiently process control algorithms such as Model Predictive Control (MPC). DR1000C is also the world's first RISC-V processor with vector extensions which has achieved ISO 26262 ASIL D Ready certification. Lastly, our product roadmap and portfolio are referred to for addressing the area of autonomous driving, which is the target of the next-generation DFP.

- 15:20-15:30 Break
- 15:30-16:20 Special Session III
- 15:30-16:20 <u>Keynote Presentation 3</u> Co-chairs: Yukinori Sato (Toyohashi Univ. of Tech.), Makoto Ikeda (Univ. of Tokyo)

## NanoBridge-based FPGA for Space Applications

*Makoto Miyamura (NanoBridge Semiconductor)* 

**Abstract:** We demonstrate FPGA based on NanoBridge, a novel resistive-change switch. NanoBridge, which is integrated in the back end of line (BEOL), features a high ON/OFF conductance ratio, weak temperature dependence of its resistance, non-volatility, and endurance against soft errors. In place of SRAM and a pass transistor, NanoBridge is utilized as a configuration switch in the FPGA. In this presentation, we will report the NanoBridge-based FPGA (NBFPGA) for applications in low-power and harsh environments, such as space application. Under JAXA's innovative satellite technology demonstration program, NBFPGA was mounted on a camera module and installed in RAPIS-1 (RAPid Innovative payload demonstration Satelite-1). Successful one-year operation in space, soft error reliability of Nanobridge device, and low-power operation will also be reported.

16:20-17:20	Poster Short Speeches
	Chair: Koji Hashimoto (Fukuoka Univ.)

- Poster 1 Simulation Platform of Computation-in-Memory with Memory Device Non idealities for DNN Kazuhide Higuchi, Chihiro Matsui, Naoko Misawa, Ken Takeuchi (Univ. of Tokyo)
- Poster 2 Study of General-Purpose CGRA Architecture Shigeyuki Takano, Hideharu Amano (Keio Univ.)
- Poster 3 Compact and Tunable ReRAM Computation-in-Memory for Log-encoding Simulated Annealing Naoko Misawa, Chihiro Matsui, Ken Takeuchi (Univ. of Tokyo)
- Poster 4 Fusion of Multiple Core and Just-in-Time Compilable CGRA Tomoya Akabe, Hidenari Inamasu, Renyuan Zhang, Yasuhiko Nakashima (Nara Institute of Science and Tech.)
- Poster 5 LLVM Instruction Latency Measurement for Software-Hardware Interface for Multi/many-core Hiro Mikami, Kei Torigoe, Makoto Inokawa, Masato Edahiro (Nagoya Univ.)
- Poster 6 Towards Network-Wide Malicious Traffic Detection with Power-Effective Hardware NIDS Design Zhenguo Hu, Hirokazu Hasegawa, Yukiko Yamaguchi, Hajime Shimada (Nagoya Univ.)
- Poster 7 Just-In-Time Code Fusion of Unary Mathematical Calculation for Supercomputer Fugaku

Kentaro Kawakami<sup>1</sup>, Shigeo Mitsunari<sup>2</sup>, Kouji Kurihara<sup>1</sup>, Kazuhito Matsuda<sup>1</sup>, Masafumi Yamazaki<sup>1</sup>, Fuyuka Yamada<sup>1</sup>, Tsuguchika Tabaru<sup>1</sup>, Ken Yokoyama<sup>1</sup> (<sup>1</sup>Fujitsu, <sup>2</sup>Cybozu Labs)

- Poster 8 Implementation of HARK Sound source localization by M-KUBOS Hou Zhongyang<sup>1</sup>, Wei Kaijie<sup>1</sup>, Hideharu Amano<sup>1</sup>, Kazuhiro Nakadai<sup>2</sup> (<sup>1</sup>Keio Univ., <sup>2</sup>Tokyo Institute of Tech.)
- Poster 9 **Multi-FPGA Implementation of Distributed Computing System for Solving the Transportation Optimization Problem** *Huang Pengyu<sup>1</sup>*, Wei Kaijie<sup>1</sup>, Hideharu Amano<sup>1</sup>, Kaori Ohkoda<sup>2</sup>, Masashi Aono<sup>1,2</sup> (<sup>1</sup>Keio Univ., <sup>2</sup>Amoeba Energy)
- Poster 10 A Job Scheduling Method Considering the Effects of Last-Level Cache Conflict Satoshi Iwata, Naoto Fukumoto (Fujitsu)
- Poster 11 An implementation of Geometric High-order Dicorrelation-based Source Separation on an FPGA board Qin Ziquan<sup>1</sup>, Wei Kaijie<sup>1</sup>, Hideharu Amano<sup>1</sup>, Kazuhiro Nakadai<sup>2</sup> (<sup>1</sup>Keio Univ., <sup>2</sup>Tokyo Institute of Tech.)
- Poster 12 A Shared Cache Architecture for VVC Coding Yoshiaki Kondo<sup>1</sup>, Masayuki Sato<sup>1</sup>, Ken Nakamura<sup>2</sup>, Yuya Omori<sup>2</sup>, Daisuke Kobayashi<sup>2</sup>, Hiroe Iwasaki<sup>1</sup>, Ryusuke Egawa<sup>3</sup>, Kazuhiko Komatsu<sup>1</sup>, Hiroaki Kobayashi<sup>1</sup> (<sup>1</sup>Tohoku Univ., <sup>2</sup>NTT, <sup>3</sup>Tokyo Denki Univ.)
- Poster 13 **Optimal Ternarization Method for Federated Model Compression** *Qian Zhao<sup>1</sup>*, *Tianyu Zhang<sup>2</sup>*, *Zhen Liu<sup>1</sup>*, *Xiaoyan Zhu<sup>3</sup>*, *Haibin Lu<sup>1</sup>*, *Yukikazu Nakamoto<sup>4</sup>* (<sup>1</sup>Bright Dream Robotics, <sup>2</sup>Huizhong Guanhua Technology, <sup>3</sup>Tsinghua *Univ.*, <sup>4</sup>Univ. of Hyogo)
- Poster 14 Distant-aware Compression for interconnection network of many-core systems Zhou Yuqing, Naoya Niwa, Hideharu Amano (Keio Univ.)
- Poster 15 An implementation of image filters on an FPGA board Chen Yuchen, Wei Kaijie, Hideharu Amano (Keio Univ.)
- Poster 16 A CNN implementation on a multi-FPGA system with system-C description Shao Ningyu<sup>1</sup>, Hiroaki Suzuki<sup>1</sup>, Hideharu Amano<sup>1</sup>, Wataru Takahashi<sup>2</sup>, Kazutoshi Wakabayashi<sup>3</sup> (<sup>1</sup>Keio Univ., <sup>2</sup>NEC, <sup>3</sup>Univ. of Tokyo)
- Poster 17 Accelerating Graph-Based SLAM on an M-KUBOS Board towards Multiaccess Edge Computing Hajime Takishita, Yuan He, Masaaki Kondo, Hideharu Amano (Keio Univ.)
- Poster 18 **Posit-Based Vision Transformer (ViT) Exploration at Edge Sites** Mery Diana, Hiroki Murota, Motoki Amagasaki, Masato Kiyama (Kumamoto Univ.)
- Poster 19 **Task Mapping with Considering both Memory and Heterogeneity** *Yifan Jin<sup>1</sup>, Mulya Agung<sup>2</sup>, Keichi Takahashi<sup>1</sup>, Yoichi Shimomura<sup>1</sup>, Hiroyuki Takizawa<sup>1</sup>* (<sup>1</sup>Tohoku Univ., <sup>2</sup>Univ. of Edinburgh)
- Poster 20 System-on-Chip based Edge Device for Speech Commands Recognition Haris Gulzar<sup>1</sup>, Muhammad Shakeel<sup>1</sup>, Kenji Nishida<sup>1</sup>, Katsutoshi Itoyama<sup>1</sup>, Kazuhiro Nakadai<sup>1</sup>, Hideharu Amano<sup>2</sup> (<sup>1</sup>Tokyo Institute of Tech., <sup>2</sup>Keio Univ.)
- Poster 21 Evaluation of IMAX2 with Sparse Matrix-matrix Multiplication Units

Ryotaro Funai, Hidenari Inamasu, Renyuan Zhang, Yasuhiko Nakashima (Nara Institute of Science and Tech.)

Poster 22 Strong Lottery Ticket Exploration for Hidden Neural Network Accelerator Yasuyuki Okoshi, Angel Lopez Garcia-Arias, Kazutoshi Hirose, Kota Ando, Kazushi Kawamura, Thiem Van Chu, Masato Motomura, Jaehoon Yu (Tokyo Institute of Tech.)

## April 21, 2022 (Japan Standard Time)

- 9:00-9:10 Session I
- 9:00-9:10 <u>Welcome and Opening Remarks</u> Co-chairs: Yuki Kobayashi (NEC), Hiroki Matsutani (Keio Univ.) Makoto Ikeda Chair of the Organizing Committee Tadao Nakamura Chair of the Steering Committee

# Tadao NakamuraChair of the Steering CommitteeJose RenauChair of IEEE/CS TCMMMinoru FujishimaVice President and President-Elect of IEICE/ES

#### 9:10-10:00 Session II

#### 9:10-10:00 <u>Keynote Presentation 4</u> Co-chairs: Takuya Nakaike (IBM), Hiroki Matsutani (Keio Univ.)

#### The IBM Telum Enterprise-class Processor

Christian Jacobi (IBM, USA)

**Abstract:** The IBM Telum Processor powers scalable performance for enterprise workloads, embedded AI acceleration, and industry leading security and RAS. In this presentation, learn how Telum balances high performance and high frequency design, power efficiency, and reliability using a multi-faceted approach including micro-architecture, advanced power and thermal management and optimization of processor and system design. Telum-based systems will grow total workload capacity, provide new functionality like real-time embedded AI inferencing, and improve power efficiency compared to prior-generation IBM Z systems.

- 10:00-10:10 Break
- 10:10-11:40 Session III: Panel Discussions

#### **Topics: "The Future of Mission-critical Mixed-criticality High-performance Embedded Systems"** *Organizer and Moderator: Masaki Gondo (eSOL) Panelists: Christian Jacobi (IBM, USA) Timothy Mattson (Intel. USA)*

Timothy Mattson (Intel, USA) Makoto Miyamura (NanoBridge Semiconductor) Juanjo Noguera (Xilinx, USA)

- 11:40-11:50 Break
- 11:50-12:40 Session IV

#### 11:50-12:40 <u>Keynote Presentation 5</u> Co-chairs: Yuki Kobayashi (NEC), Yasuo Unekawa (Toshiba Electronic Devices & Storage)

Heterogeneity: A Play in Two Acts Joseph Curley, Timothy Mattson (Intel, USA)

Abstract: Heterogeneous computing promises high performance with optimized

energy consumption. In Act 1 programmers write code with one API which our system maps onto any single processor from a range of processor-types. We'll show, for example, a single codebase moving between CPUs and GPUs. In Act 2 we look to the future where many heterogeneous processors work together to solve a single problem. We still focus on one API mapping onto many processor-types, but now we run parallel programs over distributed heterogeneous systems. Our name for this amazing system... oneAPI of course.

#### 12:40-13:30 Lunch Time Break

- 13:30-14:20Session V: Power Management and Analysis<br/>Co-chairs: Yuetsu Kodama (Riken), Satoshi Imamura (Fujitsu)
- 13:30-13:55 Body Bias Control on a CGRA based on Convex Optimization Takuya Kojima<sup>1</sup>, Hayate Okuhara<sup>2</sup>, Masaaki Kondo<sup>3</sup>, Hideharu Amano<sup>3</sup> (<sup>1</sup>Univ. of Tokyo, <sup>2</sup>National Univ. of Singapore, <sup>3</sup>Keio Univ.)
- 13:55-14:20 **Power Analysis of Directly-connected FPGA Clusters** *Kensuke lizuka, Haruna Takagi, Aika Kamei, Kazuei Hironaka, Hideharu Amano (Keio Univ.)*
- 14:20-15:10 Poster Break
- 15:10-16:00 Session VI
- 15:10-16:00 <u>Keynote Presentation 6</u> Co-chairs: Fumio Arakawa (Univ. of Tokyo), Masato Suzuki (Socionext)

#### **Software-defined Architecture and Platforms - Automotive and beyond** *Masaki Gondo (eSoL)*

Abstract: The embedded system industries, notably the automotive but not limited to, are all talking about "Software-Defined." They typically mean "Software-Defined Architecture (SDA)" but the meaning is quite vague. In this talk, we start by reviewing the different use of the term in the industry. Then, using the nextgeneration automotive software platform based on AUTOSAR Adaptive as an example, we look into the background of SDA trends, and the design approach of such platforms that needs to tackle the challenge of developing highly-energy efficient, high-performance, safe, and yet affordable systems. Also, the business and market implications of such trends and approaches will be discussed briefly if the time allows.

16:00-16:10 Break

#### 16:10-17:00 Session VII: Machine Learning I - Circuit based

*Co-chairs: Shunsuke Sasaki (Toshiba), Renyuan Zhang (Nara Institute of Science and Tech.)* 

# 16:10-16:35 A Memcapacitive Spiking Neural Network with Circuit Nonlinearity-aware Training

Reon Oshio<sup>1</sup>, Takuya Sugahara<sup>1</sup>, Atsushi Sawada<sup>1</sup>, Mutsumi Kimura<sup>1,2</sup>, Renyuan Zhang<sup>1</sup>, Yasuhiko Nakashima<sup>1</sup> (<sup>1</sup>Nara Institute of Science and Tech., <sup>2</sup>Ryukoku Univ.,)

 16:35-17:00 A 1036 TOp/s/W, 12.2 mW, 2.72 μJ/Inference All Digital TNN Accelerator in 22 nm FDX Technology for TinyML Applications Moritz Scherer<sup>1</sup>, Alfio Di Mauro<sup>1</sup>, Georg Rutishauser<sup>1</sup>, Tim Fischer<sup>1</sup>, Luca Benini<sup>1,2</sup> (<sup>1</sup>ETH Zurich, Switzerland, <sup>2</sup>Univ. of Bologna, Italy)

## April 22, 2022 (Japan Standard Time)

#### 9:00-9:40 Session VIII

#### 9:00-9:40 <u>Invited Presentation</u> Co-chairs: Ryusuke Egawa (Tokyo Denki Univ.), Daisuke Kobayashi (NTT)

## AMD Ryzen 6000 $^{\rm TM}$ Series For Mobile

Jim Gibney (AMD, USA)

Abstract: AMD will present the AMD Ryzen<sup>™</sup> 6000 Series processor for laptops, bringing the new "Zen 3+" core architecture together with AMD RDNA<sup>TM</sup> 2 architecture-based on-chip graphics. Fabricated using TSMC's 6nm process technology, the processor SoC delivers higher single and multi-threaded CPU performance than the last generation, new levels of built-in graphics performance, new platform features and long battery life. The updated AMD "Zen 3+" core is optimized to deliver high frequency and performance-per-watt. The SoC features up to eight high-performance cores, delivering 16 threads of processing, with clock speeds up to 5 GHz. These are the first notebook processors to feature RDNA<sup>TM</sup> 2 architecture-based built-in graphics, with performance up to 2X of the last generation. The SoC includes an all-new integrated display engine, allowing for ultra-high resolutions and refresh rates. New power management features include deep sleep states that save power and greatly extend battery life. The platform is entirely new with the SoC including support for DDR5 memory, PCIe® 4.0, USB4, and AI noise cancellation. This is the first x86 processor to fully support advanced Windows® 11 security features with the integrated Microsoft® Pluton security processor.

9:40-9:50 Break

#### 9:50-10:40 Session IX

9:50-10:40 <u>Keynote Presentation 7</u> Co-chairs: Yasutaka Wada (Meisei Univ.), Koyo Nitta (Univ. of Aizu)

#### Xilinx 7nm Edge Processors

Juanjo Noguera (Xilinx, USA)

**Abstract:** In this presentation, Xilinx will provide an overview of the 7nm Versal architecture, focusing on Edge applications. This talk with introduce the key target use-cases focusing on AI/ML inference applications at the edge, give a device-level Versal architecture overview and provide details on the second-generation AIE architecture (AIE-ML). We'll provide an overview of the programing abstractions for this products, and finally demonstrate some system-level performance/W comparison to different architectures.

#### 10:40-11:30 Session X: Machine Learning II - Applications

Co-chairs: Hiroyuki Takizawa (Tohoku Univ.), Teppei Hirotsu (NSITEXE)

# 10:40-11:05 Reinforcement Learning based Efficient Mapping of DNN Models onto Accelerators

Shine Parekkadan Sunny, Satyajit Das (Indian Institute of Tech. Palakkad, India)

11:05-11:30	A Low-power and Real-time 3D Object Recognition Processor with Dense
	RGB-D Data Acquisition in Mobile Platforms
	Dongseok Im, Gwangtae Park, Junha Ryu, Zhiyong Li, Sanghoon Kang,
	Donghyeon Han, Jinsu Lee, Wonhoon Park, Hankyul Kwon, Hoi-Jun Yoo (KAIST,
	Korea)

- 11:30-12:00 Poster Break
- 12:00-12:50 Session XI: Circuit and Systems Co-chairs: Yasutaka Wada (Meisei Univ.), Ryuichi Sakamoto (Tokyo Institute of Tech.)
- 12:00-12:25 Zero-standby-power Nonvolatile Standard Cell Memory Using FiCC for IoT Processors with Intermittent Operations Yuki Abe<sup>1</sup>, Kazutoshi Kobayashi<sup>1</sup>, Jun Shiomi<sup>2</sup>, Hiroyuki Ochi<sup>3</sup> (<sup>1</sup>Kyoto Institute of Tech., <sup>2</sup>Osaka Univ., <sup>3</sup>Ritsumeikan Univ.)
- 12:25-12:50 DXT501: An SDR-based Baseband MP-SoC for Multi-protocol Industrial Wireless Communication Yang Chen, Lin Liu, Xuelin Feng, Jinlin Shi (Univ. of Chinese Academy of Sciences, China)
- 12:50-14:20 Lunch Time Break
- 14:20-15:10 Session XII: Broadcasting-tech Co-chairs: Kotaro Shimamura (Hitachi), Sugako Otani (Renesas Electronics)
- 14:20-14:45 An Efficient Reference Image Sharing Method for the Parallel Video Encoding Architecture Ken Nakamura<sup>1,2</sup>, Yuya Omori<sup>1</sup>, Daisuke Kobayashi<sup>1</sup>, Koyo Nitta<sup>1</sup>, Kimikazu Sano<sup>1</sup>, Masayuki Sato<sup>2</sup>, Hiroe Iwasaki<sup>2</sup>, Hiroaki Kobayashi<sup>2</sup> (<sup>1</sup>NTT, <sup>2</sup>Tohoku Univ.)
- 14:45-15:10 Ultra-low Latency 8K Video-transmission System Utilizing Disaggregation Configuration Yasuhiro Mochida, Daisuke Shirai, Koichi Takasugi (NTT)
- 15:10-15:20 Break
- 15:20-16:10 Session XIII: Domain Specific Accelerator Co-chairs: Yuichiro Shibata (Nagasaki Univ.), Ryohei Kobayashi (Univ. of Tsukuba)
- 15:20-15:45 Hardware Acceleration of Aggregate Signature Generation and Authentication by BLS Signature over BLS12-381 Curve Kaoru Masada, Ryohei Nakayama, Makoto Ikeda (Univ. of Tokyo)
- 15:45-16:10 Encoder-based Many-Pattern Matching on FPGAs Hoang-Gia Vu, Ngoc-Dai Bui (Le Quy Don Technical Univ., VietNam)
- 16:10-16:20 <u>Poster Award and Closing Remarks</u> Program Committee Co-chairs: Yasutaka Wada (Meisei Univ.), Ryusuke Egawa (Tokyo Denki Univ.),