

2022 IEEE/ACM International Symposium on Code Generation and Optimization (CGO 2022)

**Seoul, South Korea
2 – 6 April 2022**



**IEEE Catalog Number: CFP22CGO-POD
ISBN: 978-1-6654-0585-0**

**Copyright © 2022 by the Institute of Electrical and Electronics Engineers, Inc.
All Rights Reserved**

Copyright and Reprint Permissions: Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. copyright law for private use of patrons those articles in this volume that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923.

For other copying, reprint or republication permission, write to IEEE Copyrights Manager, IEEE Service Center, 445 Hoes Lane, Piscataway, NJ 08854. All rights reserved.

****** This is a print representation of what appears in the IEEE Digital Library. Some format issues inherent in the e-media version may also appear in this print version.***

IEEE Catalog Number:	CFP22CGO-POD
ISBN (Print-On-Demand):	978-1-6654-0585-0
ISBN (Online):	978-1-6654-0584-3

Additional Copies of This Publication Are Available From:

Curran Associates, Inc
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: (845) 758-0400
Fax: (845) 758-2633
E-mail: curran@proceedings.com
Web: www.proceedings.com

CURRAN ASSOCIATES INC.
proceedings
.com



**Association for
Computing Machinery**

Advancing Computing as a Science & Profession

April 02–06, 2022
Seoul, South Korea



CGO '22

Proceedings of the 2022 IEEE/ACM International Symposium on

Code Generation and Optimization

Edited by:

Jae W. Lee, Sebastian Hack, and Tatiana Shpeisman

Sponsored by:

ACM, ACM SIGPLAN, ACM SIGMICRO, IEEE Computer Society

Supported by:

ARM, Meta, Huawei, Microsoft, Samsung, Google, Seoul National University

Message from the General Chair

Welcome to the 20th edition of the IEEE/ACM International Symposium on Code Generation and Optimization (CGO) organized in Seoul, South Korea, but attended from all around the globe!

CGO is a premier venue to bring together researchers and practitioners working at the interface of hardware and software on a wide range of optimization and code generation techniques and related issues. The conference spans the spectrum from purely static to fully dynamic approaches, and from pure software-based methods to specific architectural features and support for code generation and optimization.

I would like to dedicate my deepest gratitude to the CGO Program Committee, co-chaired by Tatiana Shpeisman and Sebastian Hack, for putting together a fantastic technical program including 27 high-quality research papers. I also thank Taewook Oh, the Workshops and Tutorials Chair, for gathering a highly stimulating set of workshops and tutorials.

Continuing the tradition, this year's CGO is held in conjunction with three "sister" conferences: the 27th International Symposium on Principles and Practice of Parallel Programming (PPoPP), the 28th International Symposium on High-Performance Computer Architecture (HPCA), and the 31st International Conference on Compiler Construction (CC). This co-location is intended to foster collaboration between the communities by bringing together software and hardware architects, programming language designers, and compiler writers. The three keynotes are jointly hosted by all the conferences. The side-by-side technical program talks, workshops and tutorials are all open to attendees from any of the other three conferences.

CGO would not have been possible without the efforts of our dedicated and hard-working organizing committee, most of whom have served for two terms (2021 and 2022). My tremendous thanks go to Yongjun Park and Mehrzad Samadi (Treasurer & Finance Co-Chairs), Hanjun Kim and Nick Johnson (Local/Global Arrangement Co-Chairs), Jubi Taneja and Roland Leißa (Artifact Evaluation Co-Chairs), Hyojin Sung (SRC Chair), Dongyoon Lee (Publicity Chair), Bernd Burgstaller (Registration Chair), Jeehoon Kang (Proceedings Chair), Tae Jun Ham (Web Chair), and Shivam Bharuka (Social Media Chair). I would also like to thank Paula Anderson, our Conference Manager from the IEEE Computer Society, the IEEE/ACM Conference Support staff, and Conference Publishing. On behalf of the organizers of CGO 2022, I would like to extend my gratitude to Fabrice Rastello, the CGO Steering Committee Chair, and the rest of the Steering Committee members for their continuous advice and support, as well as all the student volunteers.

We thank all our conference sponsors for their generous support: IEEE, ACM, ARM, Meta, Huawei, Microsoft, Samsung, Google, Seoul National University, SIGMICRO, and SIGPLAN.

Finally, my biggest thanks go to all of our speakers and attendees, who represent the best of CGO. Please enjoy the conference!

Jae W. Lee, Seoul National University
CGO 2022 General Chair

Welcome from the Program Chairs

We are pleased to welcome you to CGO 2022. On behalf of the Program Committee, we are pleased to present an exciting and stimulating program for the 2022 International Symposium on Code Generation and Optimization.

This year the conference attracted 99 submissions, 12 of which were Tools and Practical Experience papers. Based on the critical reviews of the Program Committee and subsequent online discussions, 27 high-quality research papers were accepted, with an acceptance rate of 27 %. Four of the accepted papers were Tools and Practical Experience papers. The accepted papers cover a diverse range of topics with several papers having both industry and university co-authors.

The Program Committee consisted of 47 members, which included 17 % women, 25 % industry/government members and members from 11 different countries. Each paper received at least four reviews during the double-blind review process. In a few cases where the PC was less confident in its expertise, the papers received additional expert reviews. Using the HotCRP software, each member received 8–9 papers to review. The next step was a rebuttal period with the authors of the papers. Discussions of the papers, reviews, and rebuttals then took place online in three meetings to accommodate the time zone differences of the PC members. The first two meetings discussed papers which were reviewed by reviewers of neighboring geographies (America/Europe, America/Asia). A final third meeting with all PC members discussed the remaining papers and came to a conclusion on papers that could not be decided by the previous meetings. The three online PC meetings were prepared by an online discussion period of one week before the meetings. We thank all program committee members for their comprehensive reviews and hard work leading to the selected CGO '22 program.

We also thank Jubi Taneja and Roland Leißa who co-chaired the Artifact Evaluation Committee, for their diligent work and collaboration, addressing 22 submissions. All twelve submitted Tool and Practical Experience papers were first evaluated by the Program Committee, based on practical criteria. The acceptance of Tool papers was conditional on successful evaluation of their artifact thereafter.

The final program includes papers that focus on traditional topics such as compiler infrastructures and intermediate representations, compiler correctness, program analysis and optimization, parallelism, DSLs, as well as new topics such as ML approaches to code optimization problems, optimizations for code originating from ML tasks, performance modeling, natural language based synthesis, and more!

Have a great conference experience!

Sebastian Hack, Saarland University
Tatiana Shpeisman, Google
CGO 2022 Program Chairs

Committees of CGO 2022

Organizing Committee

General Chair

[Jae W. Lee](#) (Seoul National University, South Korea)

Treasurer & Finance Co-Chairs

[Yongjun Park](#) (Hanyang University, South Korea)

[Mehrzaad Samadi](#) (NVIDIA, USA)

Local/Global Arrangement Co-Chairs

[Nick Johnson](#) (D. E. Shaw Research, USA)

[Hanjun Kim](#) (Yonsei University, South Korea)

Workshop & Tutorial Chair

[Taewook Oh](#) (Meta)

Artifact Evaluation Co-Chairs

[Roland Leißa](#) (University of Mannheim, Germany)

[Jubi Taneja](#) (Microsoft Research, USA)

Student Research Competition Chair

[Hyojin Sung](#) (POSTECH, South Korea)

Publicity Chair

[Dongyoon Lee](#) (Stony Brook University, USA)

Proceedings Chair

[Jeehoon Kang](#) (KAIST, South Korea)

Registration Chair

[Bernd Burgstaller](#) (Yonsei University, South Korea)

Social Media Chair

[Shivam Bharuka](#) (Meta, USA)

Web Chair

[Tae Jun Ham](#) (Seoul National University, South Korea)

Program Committee

Sebastian Hack (Saarland University, Germany), Co-chair

Tatiana Shpeisman (Google, USA), Co-chair

Erik Altman (IBM)

Saman Amarasinghe (Massachusetts Institute of Technology, USA)

Derek Bruening (Google)

Simone Campanoni (Northwestern University, USA)

Jerónimo Castrillon (TU Dresden, Germany)

Albert Cohen (Google, France)

Huimin Cui (Institute of Computing Technology at Chinese Academy of Sciences, China)

Delphine Demange (University of Rennes / Inria / CNRS / IRISA, France)

Yufei Ding (University of California at Santa Barbara, USA)

Johannes Doerfert (Argonne National Laboratory, USA)

Christophe Dubach (McGill University, Canada)

Nalia Farooqui (NVIDIA)

Xiaobing Feng (Institute of Computing Technology at Chinese Academy of Sciences, China)

Justin Gottschlich (Intel Labs / University of Pennsylvania, USA)

Tobias Grosser (University of Edinburgh, UK)

Ben Hardekopf (University of California at Santa Barbara, USA)

Bettina Heim (Microsoft)

Wei-Chung Hsu (National Taiwan University, Taiwan)

Vijay Janapa Reddi (Harvard University, USA)

Alexandra Jimborean (University of Murcia, Spain)

Timothy M. Jones (University of Cambridge, UK)

Changhee Jung (Purdue University, USA)

Hanjun Kim (Yonsei University, South Korea)

Fredrik Kjolstad (Stanford University, USA)

Hugh Leather (University of Edinburgh, UK)

Jaejin Lee (Seoul National University, South Korea)

Xu Liu (North Carolina State University / Oak Ridge National Laboratory, USA)

Charith Mendis (University of Illinois at Urbana-Champaign, USA)

Madan Musuvathi (Microsoft Research)

Santosh Nagarakatte (Rutgers University, USA)

Guilherme Ottoni (Meta, USA)

Yongjun Park (Hanyang University, South Korea)

Jacques Pienaar (Google, USA)

Markus Püschel (ETH Zurich, Switzerland)

Fabrice Rastello (Inria, France)

John Regehr (University of Utah, USA)

Saday Sadayappan (University of Utah, USA)

Michel Steuwer (University of Edinburgh, UK)

Ben L. Titzer (Independent Consultant, USA)

Wenwen Wang (University of Georgia, USA)

Zheng Wang (University of Leeds, UK)
Christian Wimmer (Oracle Labs, USA)
Weng-Fai Wong (National University of Singapore, Singapore)
Peng Wu (Meta, USA)
Antonia Zhai (University of Minnesota, USA)

Steering Committee

Jack Davidson (University of Virginia, USA)
Carol Eidt (Microsoft)
Vijay Janapa Reddi (Harvard University, USA)
Teresa Johnson (Google)
Fabrice Rastello (Inria, France)
Aaron Smith (Microsoft, USA)
Jingling Xue (UNSW Sydney, Australia)

Artifact Evaluation Committee

Jubi Taneja (Microsoft Research, USA), Artifact Evaluation Co-Chair
Roland Leißa (University of Mannheim, Germany), Artifact Evaluation Co-Chair
Ajay Brahmakshatriya (Massachusetts Institute of Technology, USA)
Anastasiia Izycheva (Technical University of Munich, Germany)
Andrew Habib (University of Luxembourg, Luxembourg)
Angelica Aparecida Moreira (Federal University of Minas Gerais, Brazil)
Antonio Sciarappa (Leonardo Labs, Italy)
Dongjie He (UNSW Sydney, Australia)
Guangli Li (Institute of Computing Technology at Chinese Academy of Sciences, China)
Jie Zhou (University of Rochester, USA)
Joao Rivera (ETH Zurich, Switzerland)
Johannes Lenfers (University of Münster, Germany)
Julian Rosemann (Saarland University / Saarland Informatics Campus, Germany)
Mahesh Lakshminarasimhan (University of Utah, USA)
Mathieu Fehr (University of Edinburgh, UK)
Meghan Cowan (Microsoft Research, USA)
Nader Boushehrinejad Moradi (University of Utah, USA)
Pankaj Kumar Kalita (IIT Kanpur, India)
Robert Khasanov (TU Dresden, Germany)
Sangeeta Chowdhary (Rutgers University, USA)
Seonyeong Heo (ETH Zurich, Switzerland)
Shilei Tian (Stony Brook University, USA)
Sujit Kumar Muduli (IIT Kanpur, India)
Utpal Bora (IIT Hyderabad, India)
Xiaoyang Sun (University of Leeds, UK)
Yangyu Zhang (Institute of Computing Technology at Chinese Academy of Sciences, China)
Yann Herklotz (Imperial College London, UK)

Report from the Artifact Evaluation Committee

CGO 2022 included two separate categories of papers—main conference papers and tool & practical experience papers—as it did in the previous year. The acceptance criterion of the tool papers continued to adopt the minimum requirement to test the functionality of the tool from the Artifact Evaluation Committee.

All authors of accepted CGO 2022 standard (main conference) papers and conditionally accepted Tools & Practical Experience papers were given an opportunity to participate in the artifact evaluation process by submitting a research artifact.

ACM defines an artifact as “*a digital object that is either created by the authors to be used as part of the study or generated by the experiment itself*”. The artifact evaluation process validates if the submitted artifact supports the claims made in the paper. The main goal of the process is intended to encourage researchers to take special care in conducting reproducible experiments and to package experimental workflows including related materials for making them accessible for others.

Our philosophy of artifact evaluation is that it should act as a mechanism to help authors prepare their materials and replicating their experimental results. Thus, the artifact evaluation process is intentionally designed to be interactive for both authors and reviewers. This process enables frequent communication to overcome technical issues leading to a successful evaluation of the artifact by the reviewer.

We received 20 artifact submissions from a total of 27 accepted papers, making the submission rate to 74.1% this year. All 20 artifacts, including 4 artifacts for tools papers were successfully evaluated. Each artifact received at least two reviews from the artifact evaluation committee which consisted of 25 researchers and engineers from 10 different countries across the globe.

The artifact evaluation of CGO 2022 has followed the revised ACM’s Review and Badging guidelines¹ to award papers up to three badges to indicate:

- if an artifact is publicly *available*
- if an artifact has been *functionally* evaluated or if base functionality has been exceeded and the artifact is considered easily *reusable*
- if the experimental results have been independently *reproduced*

These badges are added to the first page of the papers.

We are delighted to share that we awarded the *Reusable* and *Reproducible* badges to 14 artifacts. Five artifacts were awarded the *Functional* and *Reproducible* badges and we awarded the *Functional* badge to one artifact. The conference publishing team verified the publicly available artifacts and we

¹<https://www.acm.org/publications/policies/artifact-review-and-badging-current>

awarded the *Available* badge to 19 artifacts this year. For the tools papers, it was only required to validate the functionality criterion. However, we are thrilled to share that all tool papers that contained experimental results also qualified for the reproducible badge, indicating that the main experimental results of their papers were reproduced by our reviewers.

We would like to thank all authors and reviewers for their efforts in making the artifact evaluation process a huge success this year.

Jubi Taneja, Microsoft Research, USA
Roland Leissa, University of Mannheim, Germany
CGO 2022 Artifact Evaluation Chairs

CGO 2022 Sponsors



Association for
Computing Machinery



IEEE
COMPUTER
SOCIETY



SIGPLAN



sigmicro

arm

∞ Meta



HUAWEI



Microsoft



Google



SEOUL
NATIONAL
UNIVERSITY

Contents

Frontmatter

Welcome from the General Chair	iii
Welcome from the Program Chairs	iv
CGO 2022 Organization	v
Report from the Artifact Evaluation Committee	viii
CGO 2022 Sponsors	x

GPU

A Compiler Framework for Optimizing Dynamic Parallelism on GPUs

Mhd Ghaith Olabi, Juan Gómez Luna, Onur Mutlu, Wen-mei Hwu, and Izzat El Hajj — <i>American University of Beirut, Lebanon; ETH Zurich, Switzerland; University of Illinois at Urbana-Champaign, USA; NVIDIA, USA</i>	1
--	---

Automatic Horizontal Fusion for GPU Kernels

Ao Li, Bojian Zheng, Gennady Pekhimenko, and Fan Long — <i>Carnegie Mellon University, USA; University of Toronto, Canada</i>	14
---	----

DARM: Control-Flow Melding for SIMT Thread Divergence Reduction

Charitha Saumya, Kirshanthan Sundararajah, and Milind Kulkarni — <i>Purdue University, USA</i>	28
--	----

Efficient Execution of OpenMP on GPUs

Joseph Huber, Melanie Cornelius, Giorgis Georgakoudis, Shilei Tian, Jose M Monslave Diaz, Kuter Dinel, Barbara Chapman, and Johannes Doerfert — <i>Oak Ridge National Laboratory, USA; Illinois Institute of Technology, USA; Lawrence Livermore National Laboratory, USA; Stony Brook University, USA; Argonne National Laboratory, USA; Düzce University, Turkey</i>	41
--	----

Domain-Specific Compilation

GraphIt to CUDA Compiler in 2021 LOC: A Case for High-Performance DSL Implementation via Staging with BuildSL

Ajay Brahmakshatriya and Saman Amarasinghe — <i>Massachusetts Institute of Technology, USA</i>	53
--	----

A Compiler for Sound Floating-Point Computations using Affine Arithmetic

Joao Rivera, Franz Franchetti, and Markus Püschel — <i>ETH Zurich, Switzerland; Carnegie Mellon University, USA</i>	66
---	----

Aggregate Update Problem for Multi-clocked Dataflow Languages

Hannes Kallwies, Martin Leucker, Torben Scheffel, Malte Schmitz, and Daniel Thoma — <i>University of Lübeck, Germany</i>	79
--	----

Performance

CompilerGym: Robust, Performant Compiler Optimization Environments for AI Research

Chris Cummins, Bram Wasti, Jiadong Guo, Brandon Cui, Jason Ansel, Sahir Gomez, Somya Jain, Jia Liu, Olivier Teytaud, Benoit Steiner, Yuandong Tian, and Hugh Leather — <i>Meta, USA</i>	92
---	----

PALMED: Throughput Characterization for Superscalar Architectures

Nicolas Derumigny, Théophile Bastian, Fabian Gruber, Guillaume Iooss, Christophe Guillon, Louis-Noël Pouchet, and Fabrice Rastello — <i>Grenoble Alps University, France; Inria, France; CNRS, France; Grenoble INP, France; Colorado State University, USA; STMicroelectronics, France</i>	106
---	-----

SRTuner: Effective Compiler Optimization Customization by Exposing Synergistic Relations

Sunghyun Park, Salar Latifi, Yongjun Park, Armand Behroozi, Byungsoo Jeon, and Scott Mahlke — <i>University of Michigan at Ann Arbor, USA; Hanyang University, South Korea; Carnegie Mellon University, USA</i>	118
---	-----

Binary Techniques

Recovering Container Class Types in C++ Binaries

Xudong Wang, Xuezheng Xu, Qingan Li, Mengting Yuan, and Jingling Xue — <i>UNSW, Australia; Wuhan University, China</i>	131
--	-----

Automatic Generation of Debug Headers through BlackBox Equivalence Checking

Vaibhav Kiran Kurhe, Pratik Karia, Shubhani Gupta, Abhishek Rose, and Sorav Bansal — <i>IIT Delhi, India</i>	144
--	-----

Gadgets Splicing: Dynamic Binary Transformation for Precise Rewriting

Linan Tian, Yangyang Shi, Liwei Chen, Yanqi Yang, and Gang Shi — <i>Institute of Information Engineering at Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China</i>	155
--	-----

IR, Encryption, and Compression

Lambda the Ultimate SSA: Optimizing Functional Programs in SSA Siddharth Bhat and Tobias Grosser — <i>IIT Hyderabad, India; University of Edinburgh, UK</i>	168
NOELLE Offers Empowering LLVM Extensions Angelo Matni, Enrico Armenio Deiana, Yian Su, Lukas Gross, Souradip Ghosh, Sotiris Apostolakis, Ziyang Xu, Zujun Tan, Ishita Chaturvedi, Brian Homerding, Tommy McMichen, David I. August, and Simone Campanoni — <i>Northwestern University, USA; Princeton University, USA</i>	179
HECATE: Performance-Aware Scale Optimization for Homomorphic Encryption Compiler Yongwoo Lee, Seonyeong Heo, Seonyoung Cheon, Shinnung Jeong, Changsu Kim, Eunkyung Kim, Dongyoon Lee, and Hanjun Kim — <i>Yonsei University, South Korea; ETH Zurich, Switzerland; Seoul National University, South Korea; Samsung SDS, South Korea; Stony Brook University, USA</i>	193
Unified Compilation for Lossless Compression and Sparse Computing Daniel Donenfeld, Stephen Chou, and Saman Amarasinghe — <i>Massachusetts Institute of Technology, USA</i>	205

Program Analysis and Optimization

Loop Rolling for Code Size Reduction Rodrigo C. O. Rocha, Pavlos Petoumenos, Björn Franke, Pramod Bhatotia, and Michael O’Boyle — <i>University of Edinburgh, UK; University of Manchester, UK; TU Munich, Germany</i>	217
Solving PBQP-Based Register Allocation using Deep Reinforcement Learning Minsu Kim, Jeong-Keun Park, and Soo-Mook Moon — <i>Seoul National University, South Korea</i>	230
F3M: Fast Focused Function Merging Sean Stirling, Rodrigo C. O. Rocha, Kim Hazelwood, Hugh Leather, Michael O’Boyle, and Pavlos Petoumenos — <i>Codeplay, UK; University of Edinburgh, UK; Facebook, USA; University of Manchester, UK</i>	242
Sound, Precise, and Fast Abstract Interpretation with Tristate Numbers Harishankar Vishwanathan, Matan Shachnai, Srinivas Narayana, and Santosh Nagarakatte — <i>Rutgers University, USA</i>	254

Natural-Language Techniques

M3V: Multi-modal Multi-view Context Embedding for Repair Operator Prediction Xuezhen Xu, Xudong Wang, and Jingling Xue — <i>UNSW, Australia</i>	266
Enabling Near Real-Time NLU-Driven Natural Language Programming through Dynamic Grammar Graph-Based Translation Zifan Nan, Xipeng Shen, and Hui Guan — <i>North Carolina State University, USA; University of Massachusetts at Amherst, USA</i>	278

AI Systems

SPNC: An Open-Source MLIR-Based Compiler for Fast Sum-Product Network Inference on CPUs and GPUs Lukas Sommer, Cristian Axenie, and Andreas Koch — <i>TU Darmstadt, Germany; Huawei Research, Germany</i>	290
Distill: Domain-Specific Compilation for Cognitive Models Jan Vesely, Raghavendra Pradyumna Pothukuchi, Ketaki Joshi, Samyak Gupta, Jonathan D. Cohen, and Abhishek Bhattacharjee — <i>Yale University, USA; Princeton University, USA</i>	301
Optimizing GPU Deep Learning Operators with Polyhedral Scheduling Constraint Injection Cédric Bastoul, Zhen Zhang, Harenome Razanajato, Nelson Lossing, Adilla Susungi, Javier de Juan, Etienne Filhol, Baptiste Jarry, Gianpietro Consolaro, and Renwei Zhang — <i>Huawei Technologies, France; Huawei Technologies, China</i>	313
Comprehensive Accelerator-Dataflow Co-design Optimization for Convolutional Neural Networks Miheer Vaidya, Aravind Sukumaran-Rajam, Atanas Rountev, and P. Sadayappan — <i>University of Utah, USA; Washington State University, USA; Ohio State University, USA</i>	325

Author Index