Lecture Notes in Computer Science 13159

Founding Editors

Gerhard Goos Karlsruhe Institute of Technology, Karlsruhe, Germany

Juris Hartmanis Cornell University, Ithaca, NY, USA

Editorial Board Members

Elisa Bertino Purdue University, West Lafayette, IN, USA

Wen Gao Peking University, Beijing, China

Bernhard Steffen D TU Dortmund University, Dortmund, Germany

Moti Yung D Columbia University, New York, NY, USA More information about this series at https://link.springer.com/bookseries/558

Stephen Poole · Oscar Hernandez · Matthew Baker · Tony Curtis (Eds.)

OpenSHMEM and Related Technologies

OpenSHMEM in the Era of Exascale and Smart Networks

8th Workshop on OpenSHMEM and Related Technologies, OpenSHMEM 2021 Virtual Event, September 14–16, 2021 Revised Selected Papers



Editors Stephen Poole Los Alamos National Laboratory Los Almos, NM, USA

Matthew Baker Oak Ridge National Laboratory Oak Ridge, TN, USA Oscar Hernandez NVIDIA Corporation Santa Clara, CA, USA

Tony Curtis D Stony Brook University Stony Brook, NY, USA

ISSN 0302-9743 ISSN 1611-3349 (electronic) Lecture Notes in Computer Science ISBN 978-3-031-04887-6 ISBN 978-3-031-04888-3 (eBook) https://doi.org/10.1007/978-3-031-04888-3

© Springer Nature Switzerland AG 2022

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

OpenSHMEM is a portable specification API that implements a partitioned global address space (PGAS) programming model that focuses on low-latency one-sided communication across nodes within a system. OpenSHMEM is a modern derivative of the SGI SHMEM API, originally developed by Cray Research, Inc., for efficient programming of large-scale systems. Because of its strong following, Open Source Software Solutions, Inc. (OSSS) was licensed to drive a specification called OpenSHMEM that has portable implementations among vendors, research laboratories, and academia. The OpenSHMEM and Related Technologies Workshop (OpenSHMEM Workshop) was established in 2014 when OpenSHMEM 1.1 was released to the community and has enjoyed success in being the main workshop where users, vendors, and researchers share their experiences and publish their latest results. The community has developed the specification to version 1.5, adding new features such as teams to group together subsets of processing elements (PEs), non-blocking atomic memory operations (AMOs), and a profiling interface.

This year's workshop (OpenSHMEM 2021) included topics ranging from new applications, benchmarks, and libraries experiences to new OpenSHMEM implementations on novel hardware, programming models, and low-level communication framework extensions. The workshop agenda can be found at http://www.openshmem.org/worksh ops/openshmem2021/program.html. This year's keynotes included talks on Bale 3.0, a collection of applications with many-to-many communication patterns; NVSHMEM, an implementation of OpenSHMEM for NVIDIA accelerators; the latest advances of Chapel, an asynchronous PGAS programming language developed by Cray, which now supports multi-resolution aggregated communication to improve the message rates of applications; and Arkouda, a Chapel application that provides Python interfaces to key NumPy and Pandas operations for data science applications.

This book constitutes the proceedings of the 6th OpenSHMEM and Related Technologies Workshop. The conference was held virtually and organized by Los Alamos National Laboratory, having 102 attendees from around the world. In total, 12 papers were selected from the 18 submissions (66% acceptance rate) and presented at the workshop. The Technical Program Committee members and the chairs reviewed all the papers submitted to the workshop. The papers were organized as follows: Applications and Implementations; Tools and Benchmarks; and Programming Models and OpenSHMEM Extensions.

September 2021

Stephen Poole Matthew Baker Oscar Hernandez Tony Curtis

Organization

General Co-chairs

Matthew Baker	Oak Ridge National Laboratory, USA
Oscar Hernandez	NVIDIA Corporation, USA
Stephen Poole	Los Alamos National Laboratory, USA
Tony Curtis	Stony Brook University, USA

Technical Program Co-chairs

United States Department of Defense, USA
United States Department of Defense, USA
Hewlett Packard Enterprise, USA
Arm Ltd., USA
NVIDIA Corporation, USA

Technical Program Committee

Matthew Baker	Oak Ridge National Laboratory, USA
Pavan Balaji	Facebook, USA
Swen Boehm	Oak Ridge National Laboratory, USA
Camille Coti	University of Oregon, USA
Tony Curtis	Stony Brook University, USA
James Dinan	NVIDIA Corporation, USA
Marcel Fallet	United States Department of Defense, USA
Megan Grodowitz	Arm Ltd., USA
Max Grossman	Georgia Tech, USA
Khaled Hamidouche	Advanced Micro Devices, Inc., USA
Oscar Hernandez	NVIDIA Corporation, USA
Curtis Hughey	United States Department of Defense, USA
Bryant Lam	United States Department of Defense, USA
Alex Margolin	Huawei, Israel
Naveen Namashivayam	Hewlett Packard Enterprise, USA
Dhabaleswar Panda	Ohio State University, USA
Nick Park	United States Department of Defense, USA
Stephen Poole	Los Alamos National Laboratory, USA
Wendy Poole	Los Alamos National Laboratory, USA
Swaroop Pophale	Oak Ridge National Laboratory, USA

Howard Pritchard Michael Raymond Thomas Rolinger Gilad Shainer Pavel Shamis Sameer Shende Min Si Jessica Steffy Mitsuhisa Sato Manjunath Gorentla Venkata **Brody Williams**

Logistics

Los Alamos National Laboratory, USA Hewlett Packard Enterprise, USA United States Department of Defense, USA NVIDIA Corporation, USA Arm Ltd., USA University of Oregon, USA Argonne National Laboratory, USA United States Department of Defense, USA RIKEN, Japan NVIDIA Corporation, USA Texas Tech University, USA

Stony Brook University, USA
Los Alamos National Laboratory, USA
Los Alamos National Laboratory, USA

Web Chair

Tony Curtis

Stony Brook University, USA

Contents

Applications and Implementations

A Study in SHMEM: Parallel Graph Algorithm Acceleration with Distributed Symmetric Memory	3
Michael Ing and Alan D. George	
OpenFAM: A Library for Programming Disaggregated Memory Sharad Singhal, Clarete R. Crasta, Mashood Abdulla, Faizan Barmawer, Dave Emberson, Ramya Ahobala, Gautham Bhat, Rishi kesh K. Rajak, and P. N. Soumya	21
OpenSHMEM over MPI as a Performance Contender: Thorough Analysis and Optimizations	39
Tools and Benchmarks	
SKaMPI-OpenSHMEM: Measuring OpenSHMEM Communication Routines Camille Coti and Allen D. Malony	63
A Tools Information Interface for OpenSHMEM Md. Wasi-ur-Rahman, David Ozog, and Kieran Holland	81
CircusTent: A Tool for Measuring the Performance of Atomic Memory Operations on Emerging Architectures	92
SHMEM-ML: Leveraging OpenSHMEM and Apache Arrow for Scalable, Composable Machine Learning	111
Programming Models and Extensions	
OpenSHMEM Active Message Extension for Task-Based Programming Wenbin Lu, Tony Curtis, and Barbara Chapman	129
UCX Programming Interface for Remote Function Injection and Invocation Luis E. Peña, Wenbin Lu, Pavel Shamis, and Steve Poole	144

x Contents

Can Deferring Small Messages Enhance the Performance of OpenSHMEM Applications?	160
David Ozog, Md. Wasi-ur-Rahman, and Kieran Holland	
Remote Programmability Model for SmartNICs in HPC Workloads Morad Horany and Alex Margolin	178
Dynamic Symmetric Heap Allocation in NVSHMEM Akhil Langer, Seth Howell, Sreeram Potluri, Jim Dinan, and Jiri Kraus	187
Author Index	199