

2025 IEEE Symposium on High-Performance Interconnects (HOTI 2025)

**Virtual Symposium
20-22 August 2025**



IEEE Catalog Number: CFP25HIS-POD
ISBN: 979-8-3315-4889-6

**Copyright © 2025 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:	CFP25HIS-POD
ISBN (Print-On-Demand):	979-8-3315-4889-6
ISBN (Online):	979-8-3315-4888-9
ISSN:	1550-4794

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

2025 IEEE Symposium on High-Performance Interconnects (HotI) **HotI 2025**

Table of Contents

Message from the General Chairs	vii
Message from the Program Committee Chairs	ix
Organizing Committee	x
Technical Program Committee	xii
Steering Committee	xiii
Keynotes	xiv
Technical Sessions	xvi
Invited Talks	xvii
Sponsor Talks	xix
Tutorials	xxx
Patrons	xxxvii
Sponsors	xxxix
Media Partner	xl

Technical Paper Session A: Network Synchronization and Routing

bittide: Control Time, Not Flows	1
<i>Martijn Bastiaan (QBayLogic, Netherlands), Christiaan Baaij (QBayLogic, Netherlands), Martin Izzard (Google), Felix Klein (QBayLogic, Netherlands), Sanjay Lall (Stanford University, USA; Google DeepMind), and Tammo Spalink (Google DeepMind)</i>	
Deadlock-Free Routing for Full-Mesh Networks without using Virtual Channels	12
<i>Alejandro Cano (Universidad de Cantabria, Spain), Cristóbal Camarero (Universidad de Cantabria, Spain), Carmen Martínez (Universidad de Cantabria, Spain), and Ramón Beivide (Universidad de Cantabria, Spain; Barcelona Supercomputing Center, Spain)</i>	

Technical Paper Session B: Co-packaged Optics

Accelerating Frontier MoE Training with 3D Integrated Optics	25
<i>Mikhail Bernadskiy (Lightmatter, USA), Peter Carson (Lightmatter, USA), Thomas Graham (Lightmatter, USA), Taylor Groves (Lightmatter, USA), Ho John Lee (Lightmatter, USA), and Eric Yeh (Lightmatter, USA)</i>	
Silicon Photonic Accelerated Memory Pooling for Efficient Compute Resource Allocation	36
<i>Zhenguo Wu (Columbia University, USA) and Keren Bergman (Columbia University, USA)</i>	

Technical Paper Session C: Interconnect Software -- Runtimes & Workloads

Demystifying NCCL: An In-Depth Analysis of GPU Communication Protocols and Algorithms	48
<i>Zhiyi Hu (ETH Zurich, Switzerland), Siyuan Shen (ETH Zurich, Switzerland), Tommaso Bonato (ETH Zurich, Switzerland), Sylvain Jeaugey (NVIDIA Corporation), Cedell Alexander (Broadcom Inc.), Eric Spada (Broadcom Inc.), James Dinan (NVIDIA Corporation), Jeff Hammond (NVIDIA Corporation), and Torsten Hoefer (ETH Zurich, Switzerland)</i>	
Characterizing Communication Patterns in Distributed Large Language Model Inference	60
<i>Lang Xu (The Ohio State University, US), Kaushik Kandadi Suresh (The Ohio State University, US), Quentin Anthony (The Ohio State University, US), Nawras Alnaasan (The Ohio State University, US), and Dhabaleswar K. Panda (The Ohio State University, US)</i>	
FPsPIN: An FPGA-Based Open-Hardware Research Platform for Processing in the Network	71
<i>Timo Schneider (ETH Zurich), Pengcheng Xu (ETH Zurich), and Torsten Hoefer (ETH Zurich)</i>	

Technical Paper Session D: Scale-up Interconnects

Micro-LED Data Interconnect for Scale-Up Networks with Record Energy Efficiency	82
<i>Bardia Pezeshki (AvicenaTech Corp, USA), Alex Tselikov (AvicenaTech Corp, USA), Rob Kalman (AvicenaTech Corp, USA), Suresh Rangarajan (AvicenaTech Corp, USA), Ivan Huang (AvicenaTech Corp, USA), Jeff Pepper (AvicenaTech Corp, USA), Sarah Zou (AvicenaTech Corp, USA), Howard Rourke (AvicenaTech Corp, USA), and Alasdair Fikouras (AvicenaTech Corp, USA)</i>	
On-Package Memory with Universal Chiplet Interconnect Express (UCIe): A Low Power, High Bandwidth, Low Latency and Low Cost Approach	87
<i>Debendra Das Sharma (Intel Corporation), Swadesh Choudhary (Intel Corporation), Peter Onufryk (Intel Corporation), and Rob Pelt (AMD Corporation)</i>	
Author Index	97