IEEE Hot Chips 20 Tutorial:

Scalable Parallel Programming with CUDA

Author/Presenter Biographies

John Nickolls is director of architecture at NVIDIA for GPU computing. He was previously with Broadcom, Silicon Spice, Sun Microsystems, and was a cofounder of MasPar Computer. His interests include parallel processing systems, languages, and architectures. He has a B.S. in electrical engineering and computer science from the University of Illinois, and M.S. and Ph.D. degrees in electrical engineering from Stanford University.

Massimiliano Fatica is a Senior Applied Engineer at NVIDIA where he works in the areas of GPU computing (high-performance computing and clusters). Prior to joining NVIDIA, he was a research staff member at the Center for Integrated Turbulence Simulations and Center for Turbulence Research at Stanford University where he worked on applications for the Stanford Streaming Supercomputer and a Senior Solution Architect at Clearspeed. He holds a Laurea in Aeronautic Engineering and PhD in Theoretical and Applied Mechanics from the University of Rome "La Sapienza".

Michael Garland is a research scientist with NVIDIA Research. Prior to joining NVIDIA, he was an assistant professor in the department of computer science at the University of Illinois at Urbana-Champaign. He received Ph.D. and B.S. degrees from Carnegie Mellon University. His research interests include computer graphics and visualization, geometric algorithms, and parallel algorithms and programming models.

Patrick LeGresley is a Senior Applied Engineer at NVIDIA where he works in the areas of Computational Fluid Dynamics (CFD) and structural mechanics. Prior to joining NVIDIA, he was a research staff member at the Center for Integrated Turbulence Simulations at Stanford University where he worked on developing CFD codes on multicore architectures such as the GPU. He holds a BS in Aerospace Engineering from the University of Kansas and a MS and PhD in Aeronautics and Astronautics from Stanford University.

Ian Buck is software director of GPU computing at NVIDIA. He completed his Ph.D. at the Stanford Graphics Lab in 2004. His thesis was titled "Stream Computing on Graphics Hardware," researching programming models and computing strategies for using graphics hardware as a general-purpose computing platform. His work included developing the Brook software tool chain for abstracting the GPU as a general-purpose streaming coprocessor.

Wen-mei W. Hwu is a Professor and holds the Walter J. ("Jerry") Sanders III-Advanced Micro Devices Endowed Chair in Electrical and Computer Engineering of the University of Illinois at Urbana-Champaign. His research interests are in the area of architecture, implementation, and compilation for parallel computer systems. He is the director of the IMPACT research group (www.crhc.uiuc.edu/Impact). For his contributions in research and teaching, he received the ACM SigArch Maurice Wilkes Award, the ACM Grace Murray Hopper Award, the Tau Beta Pi Daniel C. Drucker Eminent Faculty Award, and the ISCA Most Influential Paper Award. He is a fellow of IEEE and ACM. Hwu serves on the Executive Committee of the MARCO/DARPA C2S2 (www.c2s2.org) and GSRC (www.gigascale.org) Focus Research Centers. He leads the GSRC Concurrent Systems Theme. He is the hardware lead of the \$208M NSF Petascale Computer Project awarded to the University of Illinois and IBM in 2007. He directs the world's first CUDA Center of Excellence funded by NVIDIA and co-directs the Intel-Microsoft UIUC UPCRC with Marc Snir. Dr. Hwu received his Ph.D. degree in Computer Science from the University of California, Berkeley.