

Keynote Speakers

Keynote I: Manish Parashar

» University of Utah, USA

Chair: Leonel Sousa, University of Lisbon, Portugal

Big Data and Extreme-Scales: Computational Science in the 21st Century

- September 1, 10:30 AM New York | 3:30 PM Lisbon | 4:30 PM Brussels | 10:30 PM Beijing

Abstract: Extreme scales and big data are essential to computational and data-enabled science and engineering in the 21st century, promising dramatic new insights into natural and engineered systems. However, data-related challenges are limiting the potential impact of application workflows enabled by current and emerging extreme scale, high-performance, distributed computing environments. These data-intensive application workflows involve dynamic coordination, interactions and data coupling between multiple application processes that run at scale on different resources, and with services for monitoring, analysis and visualization and archiving, and present challenges due to increasing data volumes and complex data-coupling patterns, system energy constraints, increasing failure rates, etc. In this talk I will explore some of these challenges and investigate how solutions based on data sharing abstractions, managed data pipelines, data-staging service, and in-situ / in-transit data placement and processing can be used to help address them. This research is part of the DataSpaces project at the Scientific Computing and Imaging (SCI) Institute, University of Utah.

Short bio: Manish Parashar is the Director of the Scientific Computing and Imaging (SCI) Institute, Chair in Computational Science and Engineering, and Professor at the School of Computing at the University of Utah. He is currently on an IPA appointment at the National Science Foundation where he is serving as Office Director of the NSF Office of Advanced Cyberinfrastructure. His research interests are in the broad areas of Parallel and Distributed Computing and Computational and Data-Enabled Science and Engineering and has published extensively in these areas. He has also deployed software systems that are widely used. Manish is the founding chair of the IEEE Technical Consortium on High Performance Computing (TCHPC), Editor-in-Chief of the IEEE Transactions on Parallel and Distributed Systems and serves on the editorial boards and organizing committees of several journals and international conferences and workshops. He has received numerous awards for his research and leadership, and is Fellow of AAAS, ACM, and IEEE/IEEE Computer Society.

For more information, please visit <http://manishparashar.org>.



Keynote II: Alba Cristina Melo

» University of Brasilia, Brazil

Chair: Domenico Talia, University of Calabria, Italy

HPC for Bioinformatics: The Genetic Sequence Comparison Quest for Performance

- September 2, 10:30 AM New York / 3:30 PM Lisbon / 4:30 PM Brussels / 10:30 PM Beijing

Abstract: Genetic Sequence Comparison is an important operation in Bioinformatics, executed routinely worldwide. Two relevant algorithms that compare genetic sequences are the Smith-Waterman (SW) algorithm and Sankoff's algorithm. The Smith-Waterman algorithm is widely used for pairwise comparisons and it obtains the optimal result in quadratic time - $O(n^2)$, where n is the length of the sequences. The Sankoff algorithm is used to structurally align two sequences and it computes the optimal result in $O(n^4)$ time. In order to accelerate these algorithms, many parallel strategies were proposed in the literature. However, the alignment of whole chromosomes with hundreds of millions of characters with the SW algorithm is still a very challenging task, which requires extraordinary computing power. Likewise, obtaining the structural alignment of two sequences with the Sankoff algorithm requires parallel approaches. In this talk, we first present our MASA-CUDAlign tool, which was used to pairwise align real DNA sequences with up to 249 millions of characters in a cluster with 512 GPUs, achieving the best performance in the literature in 2021. We will present and discuss the innovative features of the most recent version of MASA-CUDAlign: parallelogram execution, incremental speculation, block pruning and score-share balancing strategies. We will also show performance and energy results in homogeneous and heterogeneous GPU clusters. Then, we will discuss the design of our CUDA-Sankoff tool and its innovative strategy to exploit multi-level wavefront parallelism. At the end, we will show a covid-19 case study, where we use the tools discussed in this talk to compare the SARS-CoV-2 genetic sequences, considering the reference sequence and its variants.

Short bio: Alba Cristina Magalhaes Alves de Melo received her PhD in Computer Science from the Institut National Polytechnique de Grenoble (INPG), France, in 1996. Since 1997, she is with the Department of Computer Science at the University of Brasilia (UnB), Brazil, where she is now Full Professor. Prof. Melo is a CNPq/Brazil Research Fellow level 1C. She is Senior Member of the IEEE Computer Society, Member of the Brazilian Computer Society Council and Brazilian Delegate for the BRICS (Brazil, Russia, India, China and South Africa) Working Group on High Performance Computing. She advised 6 PhD Theses and 23 MsC Dissertations. In 2016, she received the Brazilian Capes Award on "Advisor of the Best PhD Thesis in Computer Science", which is the most prestigious award for PhD Theses in Brazil. The awarded Thesis describes the development of the MASA-CUDAlign tool. Prof. Melo was member of the Editorial Board of IEEE Transactions of Parallel and Distributed Systems (2015-2019) and she is currently member of the Editorial Board of IEEE Transactions on Computers and Journal of Parallel and Distributed Computing. Her current research interests are High Performance Computing, Bioinformatics and Cloud Computing.



Keynote III: Keshav Pingali

» Katana Graph & The University of Texas at Austin, USA

Chair: Fernando Silva, University of Porto, Portugal

Knowledge Graphs, Graph AI, and the Need for High-performance Graph Computing

- September 3, 10:30 AM New York / 3:30 PM Lisbon / 4:30 PM Brussels / 10:30 PM Beijing

Abstract: Knowledge Graphs now power many applications across diverse industries such as FinTech, Pharma and Manufacturing. Data volumes are growing at a staggering rate, and graphs with hundreds of billions edges are not uncommon. Computations on such data sets include querying, analytics, and pattern mining, and there is growing interest in using machine learning to perform inference on large graphs. In many applications, it is necessary to combine these operations seamlessly to extract actionable intelligence as quickly as possible. Katana Graph is a start-up based in Austin and the Bay Area that is building a scale-out platform for seamless, high-performance computing on such graph data sets. I will describe the key features of the Katana Graph Engine that enable high performance, some important use cases for this technology from Katana's customers, and the main lessons I have learned from doing a startup after a career in academia.

Bio: Keshav Pingali is the CEO of Katana Graph, a start-up in the area of graph computing backed by Intel Capital, Dell Technologies Capital, Redline Capital and Walden International, and a professor in the Department of Computer Science at the University of Texas at Austin where he holds the W.A."Tex" Moncrief Chair of Computing. He is a Foreign Member of the Academia Europaea, a Distinguished Alumnus of IIT Kanpur, India, and a Fellow of the ACM, IEEE and AAAS. He has served on the NSF CISE Advisory Committee (2009-2012), and he was co-Editor-in-Chief of the ACM Transactions on Programming Languages and Systems (2007-2010). He is the author of more 200 papers in the area of graph computing, parallel and distributed systems, and programming languages.



SPONSORS

