Frontispiece to ACM HPDC Talk Abstract

This abstract has no directly linked paper. It has two presentations and several related papers listed below

Presentations

1. Geoffrey Fox, "HPC, Big Data, and Machine Learning Convergence" Presentation in Washington DC July 16 2019
   http://dsc.soic.indiana.edu/presentations/HPC-BigDataConvergenceJuly16-2019.pdf
2. Geoffrey Fox, "Perspectives on High-Performance Computing in a Big Data World", http://dsc.soic.indiana.edu/presentations/HPDC%20Presentation.pptx ACM HPDC 2019 The 28th International Symposium on High-Performance Parallel and Distributed Computing, Phoenix, Arizona, USA - June 27, 2019. This uses a video nugget http://dsc.soic.indiana.edu/presentations/nancoconfinement_movie.avi. There is a Google Slides version https://docs.google.com/presentation/d/1HQ1QPKO6QPEHEg9-h7y3wUZgKEspzvTqMasRObL6NF4/edit?usp=sharing and a YouTube video (short) https://www.youtube.com/playlist?list=PLy0VLh_GFyz8VEJa1syIwzsY4zI61kKYq or longer version https://www.youtube.com/playlist?list=PLy0VLh_GFyz8QWhdAO0IKfc5Mjc62dK5i, each in 5 parts.

Papers

3. Geoffrey Fox, "A review of the Learning Everywhere area or the Intersection of Machine Learning, Big Data and HPC", July 28, 2019.
   http://dsc.soic.indiana.edu/publications/Learning%20EverywhereResource.pdf
   https://www.researchgate.net/publication/333815238_Learning_Everywhere_Resource_for_BDEC_General_Links
4. Geoffrey Fox, James A. Glazier, JCS Kadupitiya, Vikram Jadhao, Minje Kim, Judy Qiu, James P. Sluka, Endre Somogyi, Madhav Marathe, Abhijin Adiga, Jiangzhuo Chen, Oliver Beckstein, and Shantenu Jha, "Learning Everywhere: Pervasive Machine Learning for Effective High-Performance Computation," presented at the HPDC Workshop at IPDPS, Rio de Janeiro [Online]. Available: https://arxiv.org/abs/1902.10810, http://dsc.soic.indiana.edu/publications/Learning_Everywhere_Summary.pdf https://www.researchgate.net/publication/331152466_Learning_Everywhere_Pervasive_Machine_Learning_for_Effective_High-Performance_Computation
5. Geoffrey Fox, James A. Glazier, JCS Kadupitiya, Vikram Jadhao, Minje Kim, Judy Qiu, James P. Sluka, Endre Somogyi, Madhav Marathe, Abhijin Adiga, Jiangzhuo Chen, Oliver Beckstein, and Shantenu Jha, "Learning Everywhere: Pervasive Machine Learning for Effective High-Performance Computation: Application Background," Feb. 2019 [Online]. Available: http://dsc.soic.indiana.edu/publications/Learning_Everywhere.pdf https://www.researchgate.net/publication/331075093_Learning_Everywhere_Pervasive_Machine_Learning_for_Effective_High-Performance_Computation_Application_Background

Perspectives on High-Performance Computing in a Big Data World

Geoffrey Fox
Indiana University
Bloomington, IN, USA
gcf@indiana.edu

ABSTRACT

High-Performance Computing (HPC) and Cyberinfrastructure have played a leadership role in computational science even since the start of the NSF computing centers program. Thirty years ago parallel computing was a centerpiece of computer science research. Naively Big Data surely requires HPC to be processed, and transformational Big Data technology such as Hadoop and Spark exploit parallelism to success. Nevertheless, the HPC community does not appear to be thriving as a leader in Data Science while parallel computing is no longer a centerpiece. Some reasons for this are the dominant presence of Industry in technology futures and the universal fascination with Artificial Intelligence and Machine Learning. Maybe the pendulum will swing back a bit, but I expect the "AI first" philosophy to dominate in the foreseeable future. Thus I describe a future where HPC thrives in collaboration with Industry and AI. In particular, I discuss the promise of MLforHPC (AI for systems) and HPCforML (systems for AI).

CCS Concepts/ACM Classifiers

• **Computing methodologies~Artificial intelligence**
• **Computing methodologies~Massively parallel and high-performance simulations** • **Computer systems organization~Parallel architectures**

Author Keywords

HPC, Big Data, Computational Science, Data Science

BIOGRAPHY

Fox received a Ph.D. in Theoretical Physics from Cambridge University where he was Senior Wrangler. He is now a distinguished professor of Engineering, Computing, and Physics at Indiana University where he is the director of the Digital Science Center. He previously held positions at Caltech, Syracuse University, and Florida State University after being a postdoc at the Institute for Advanced Study at Princeton, Lawrence Berkeley Laboratory, and Peterhouse College Cambridge. He has supervised the Ph.D. of 72 students and published around 1300 papers (over 500 with at least ten citations) in physics and computing with an hindex of 77 and over 35000 citations. He is a Fellow of APS (Physics) and ACM (Computing) and works on the interdisciplinary interface between computing and applications. Current work is in Biology, Pathology, Sensor Clouds and Ice-sheet Science, Image processing, Deep Learning, and Particle Physics. His architecture work is built around High-performance computing enhanced Software Defined Big Data Systems on Clouds and Clusters. The analytics focuses on scalable parallel machine learning. He is an expert on streaming data and robot-cloud interactions. He is involved in several projects to enhance the capabilities of Minority Serving Institutions. He has experience in online education and its use in MOOCs for areas like Data and Computational Science.

REFERENCES

1. Geoffrey Fox, James A. Glazier, JCS Kadupitiya, Vikram Jadhao, Minje Kim, Judy Qiu, James P. Sluka, Endre Somogyi, Madhav Marathe, Abhijin Adiga, Jiangzhuo Chen, Oliver Beckstein, and Shantenu Jha, "Learning Everywhere: Pervasive Machine Learning for Effective High-Performance Computation", HPBDC workshop at IPDPS, Rio de Janeiro, Brazil, Monday, May 20th, 2019