DEPARTMENT OF COMPUTER SCIENCE, INDIANA UNIVERSITY, BLOOMINGTON, USA

DataFlow SuperComputing for BigData DeepAnalytics

INSTRUCTOR
Veljko Milutinovic
vmilutin@iu.edu

CONTACT
Charles Pope
cepope@indiana.edu

COURSE DATES
Aug 30 - Sept 04 2020

LOCATION
Myles Brand 100A

COURSE OVERVIEW

This course presents the DataFlow SuperComputing paradigm, defines its advantages and sheds light on the related programming model with hands-on coding experience. DataFlow computers, compared to ControlFlow computers, offer speedups of 20 to 200 (even 2000 for some applications), and power and size reductions of up to 1/20. However, the programming paradigm is different, and has to be studied. The course explains the paradigm of programming in space, using Maxeler (a provider of multiscale dataflow computing systems) as an example, and gives an overview of ongoing research in the field. Examples include Data Engineering, Image Processing, Deep Learning, Financial Analytics, GeoPhysics, Genomics, etc. The course also covers advanced DataFlow issues like: Compilation, OS, and methods for speed-up maximization, using tools like WebIDE and MaxIDE. DataFlow potentials are discussed through the notions introduced by four Nobel Laureates: Richard Feynman, Ilya Prigogine, Daniel Kahneman, and Andre Geim. A recent study from Tsinghua University in China reveals that, for Shallow Water Weather Forecast, which is a BigData problem, on the 1U level, the Maxeler DataFlow machine was 14 times faster than the Tianhe machine. At that time, Tianhe was rated #1 on the Top 500 list of the fastest computers in the World (based on Linpack, which is a SmallData benchmark). This engine is used by leading banks and labs, and is now available from Amazon AWS.

The same course material was previously delivered at MIT, Harvard, Purdue, IU, Columbia, NYU, CMU, GATECH, ETH, EPFL, UNIWIE, TUWIEN, Ljubljana, Belgrade…

GRADING SYSTEM

This course is a 3-credit course for those who take a 60-hour homework (three HW assignments), a 2-credit course for a 40-hour homework (two HW assignments), and a 1-credit course for a 20-hour homework (one HW assignment). After the entire material is delivered in-class, according to the **SCHEDULE**, students get an oral open-book exam, and a tentative grade gets assigned. The remaining teaching hours till 45 course hours are delivered via Skype or Study, on the weekly basis, till semester's end. The tentative grade becomes final when all HW assignments compile and run successfully by the semester's end. If one of the HW assignments does not compile, the grade gets one step lower, if two do not compile – the grade becomes two steps lower. At least one HW assignment must compile for a passing grade. Those who like to get a grade one step better than assigned at the mid-term exam, have to do also the HW#4. HW#1 is related to math, HW#2 to image processing, HW#3 to Machine Learning, while HW#4 is on any topic selected by student.

Note! Senior level undergraduate students, master students, and PHD students are welcome: Computer Science, Information Systems, Data Science, Informatics, Computer Engineering, Electrical Engineering, Finance, MBA (onLine and inClass).

INSTRUCTOR

Prof. Veljko Milutinovic received his PhD from the University of Belgrade, spent about a decade on various faculty positions in the USA (mostly at Purdue University, and more recently at Indiana University in Bloomington), and was a co-designer of the DARPA's first GaAs RISC microprocessor and the DARPA's first 4096-processor GaAs systolic array. Later he taught and conducted research at the University of Belgrade, Serbia, in ECE, MATH, eBUSINESS, and SCIENCE. Now he serves as a Senior Advisor to Maxeler Technologies in London, UK. His research is mostly in datamining and dataflow computing, with the emphasis on mappings of algorithms onto architectures. His co-authored paper on matrix multiplication for dataflow received "The IET Premium Award for 2014" (meaning the single best paper in IET Computing for 2012 and 2013). He is a Fellow of the IEEE and a Member of Academia Europaea. He is a member of the Serbian National Academy of Engineering Sciences and a Foreign Member of the Montenegro National Academy of Sciences and Arts. He has over 100 SCI journal papers, well over 1000 Thomson-Reuters citations, well over 1000 Scopus citations, and about 4000 Google Scholar citations.

TEACHING ASSISTANT **Milos Kotlar** OriginTrail, Slovenia: kotlar.milos@gmail.com

SCHEDULE

***Aug 30 12pm-8pm*:**
- Introduction to DataFlow computing
- Concepts of DataFlow computing
- Applications of DataFlow computing

***Aug 31 4pm-8pm*:**
- Details of programming in space
- Selected examples with hands-on
- Open research problems

***Sep 01 4pm-8pm*:**
- Advanced issues in Math (Tensor Calculus)
- Advanced issues in Image Understanding
- Advanced issues in ML and AI for apps of interest

***Sep 2 4pm-8pm*:**
- Discussion of homework assignments
  Hands-on activities (maxeler.mi.sanu.ac.rs)
- Examples of interest for students

***Sep 3 OPEN*:**
- Sample midterm exam + Q&A

***Sep 4 OPEN*:**
- Midterm exam + Q&A

REFERENCES

Milutinovic, V., et al, **Guide to DataFlow SuperComputing**, Springer, 2015 (one textbook, part I) and 2017 (two textbooks, parts II and III).
Hurson, A., Milutinovic, V., editors, **Advances in Computers: DataFlow**, Elsevier, 2015 (one SCI textbook) and 2017 (two SCI textbooks).
Trifunovic, N., Milutinovic, V. et al, **"The AppGallery.Maxeler.com for BigData SuperComputing,"** Journal of Big Data, Springer, 2016.
Trifunovic, N., Milutinovic, V. et al, **"Paradigm Shift in SuperComputing: DataFlow vs ControlFlow,"** Journal of Big Data, 2015.
Milutinovic, V., **"The HoneyComb Architecture,"**
Proceedings of the IEEE, 1989.
Milutinovic, V. et all, **"Splitting Spatial and Temporal Localities for Entropy Minimiation"** Tutorial of the IEEE ISCA, 1995.
Jovanovic, Z., Milutinovic, V., **"FPGA Accelerator for Floating-Point Matrix Multiplication,"** The IET Computers and Digital Techniques Premium Award for 2014, IET (formerly IEE),
Volume 6, Issue 4, 2012 (pp. 249-256).
Milutinovic, V., **"A Comparison of Suboptimal Detection Algorithms (Suboptimal Algorithms for Data Analytics),"**
Proceedings of the IEE (now IET), 1988.
Flynn, M., Mencer, O., Milutinovic, V., at al, **Moving from PetaFlops to PetaData, Communications of the ACM**, May 2013.
Kotlar, M., Milutinovic, V., **"The Tensor Calculus Operations for the Data Flow Paradigm,"** The ExaComm Workshop of the International Supercomputing Conference, Frankfurt, Germany, June 28, 2018.
Milutinovic, V., **"The Ultimate DataFlow"**, Invited Talk at the ExaComm Workshop of the ISC, Frankfurt, Germany, June 28, 2018.

LINKS

Programming practices of this course are supported by: https://webide.maxeler.com and appgallery.maxeler.com
This course covers the know-how needed for the mastering of the programming model and the related technology that Intel describes in the Intel patent (US20180189063A1) about the next generation Intel products, portrayed by Intel as: "Intel has dreamed up a new architecture that could in one fell swoop kill off the general purpose processor as a concept and the X86 instruction set as the foundation of modern computing," as in QR code in the lower right-hand corner here:

At the U.S. News Report:
#1 for onLineMBA (USNewsReport2019)
#7 for Management Information Systems (USNewsReport2019)