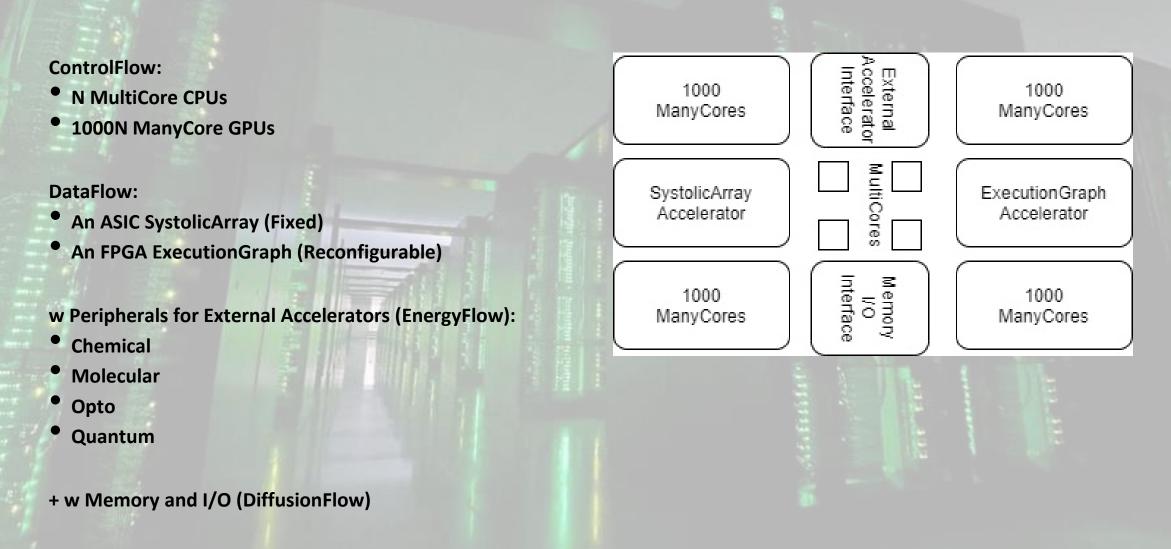
## The Ultimate SuperComputer-on-a-Chip for Massive Big Data and Highly Iterative Algorithms

Veljko Milutinović

## The Optimal Architecture



## **Optimal Distribution of Transistor Budget**

- Ceiling-Dependent:
  - VLSI: 100BTr
  - WSI: 1TTr
- Application-Dependent:
  - SW DataFlow
  - ML ControlFlow
- Strategy-Dependent:
  - CF-Oriented
  - DF-Oriented
- Memory-Dependent!

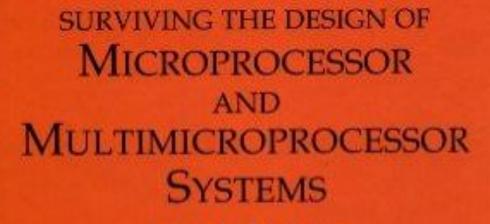
## **Pioneering Efforts (Before BIREN and CEREBRAS)**

Company	Product name	Country	City	Number of CPU cores	Number of GPU cores	CPU clock rate	Launched
Alibaba	ХТ910	China	Hangzhou	1/2/4 per cluster	N/A	2.0 – 2.5 GHz	2020
RISC-V	Micro Magic RISC-V Core [39]	US	San Francisco	1	N/A	4.25 – 5.19 GHz	2020
SiFive	FU740 RISC-V SoC [25]	US	San Francisco	4	1	1.4 – 1.5 GHz	2020
	AMD Ryzen™ 5 3400G with Radeon™ RX Vega 11 Graphics [24]	US	Santa Clara	4	11	3.7 – 4.2 GHz	2019
Nvidia	Tegra Xavier [28]	US	Santa Clara	8	512 CUDA	N/A	2019
Esperanto Technologies	N/A		Mountain View	16 ET-Maxion cores	4096 ET-Minion cores + ET- Graphics [26,27]	2+GHz	2018
Intel	Intel Sandy Bridge [23]	US	Santa Clara	1-4 (4-6 Extreme, 2-8 Xeon)	6	1.60 – 3.60 GHz	2011
Moscow Center of SPARC Technologies (MCST)	Elbrus-2S+ (1891ВМ7Я) [29]	Russia	Moscow	2 Elbrus 2000 cores	4 DSP Elcore-09 cores	300 – 800 MHz	2011

## Past Experiences: MultiCore

Split Cache

 Milutinovic, V. (1996). The Split Temporal/Spatial Cache: Initial Performance Analysis. SCIzzL-5, March 1996, 63-69.



LESSONS LEARNED



Veljko Milutinović Foreword by Michael J. Flynn



## Past Experiences: ManyCore

- GaAs Microprocessor
  - Milutinovic, V., Fura, D., & Helbig, W. (1986).
     An Introduction To GaAs Microprocessor Architecture for VLSI.
     Computer, (3), 30-42.

#### SURVIVING THE DESIGN OF A 200 MHz RISC MICROPROCESSOR LESSONS LEARNED



Veljko Milutinović Foreword by Michael Flynn

Convertien Society

P 2 a partition of according to an

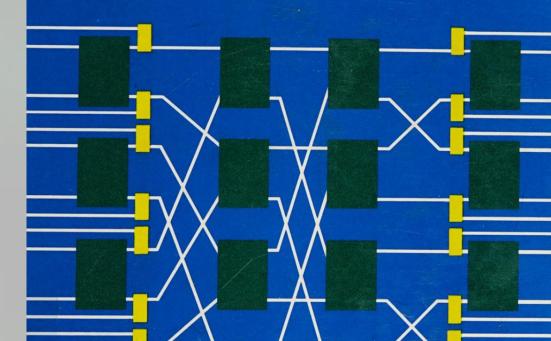
## Past Experiences: SystolicArrays

- GaAs Systolics
  - Fortes, J. A., Milutinovic, V., (1986, January).
     A High-Level Systolic Architecture for GaAs.
     In Proc. 19th Ann. Hawaii Int'l Conf.

System Sciences (pp. 253-258).

## **COMPUTER ARCHITECTURE**

Concepts and Systems Edited by Veljko M. Milutinović



## Past Experiences: ExecutionFlow

- DataFlow
  - Milutinović, V., Salom, J., Trifunović, N., & Giorgi, R. (2015). Guide to DataFlow Supercomputing. Cham: Springer Nature.

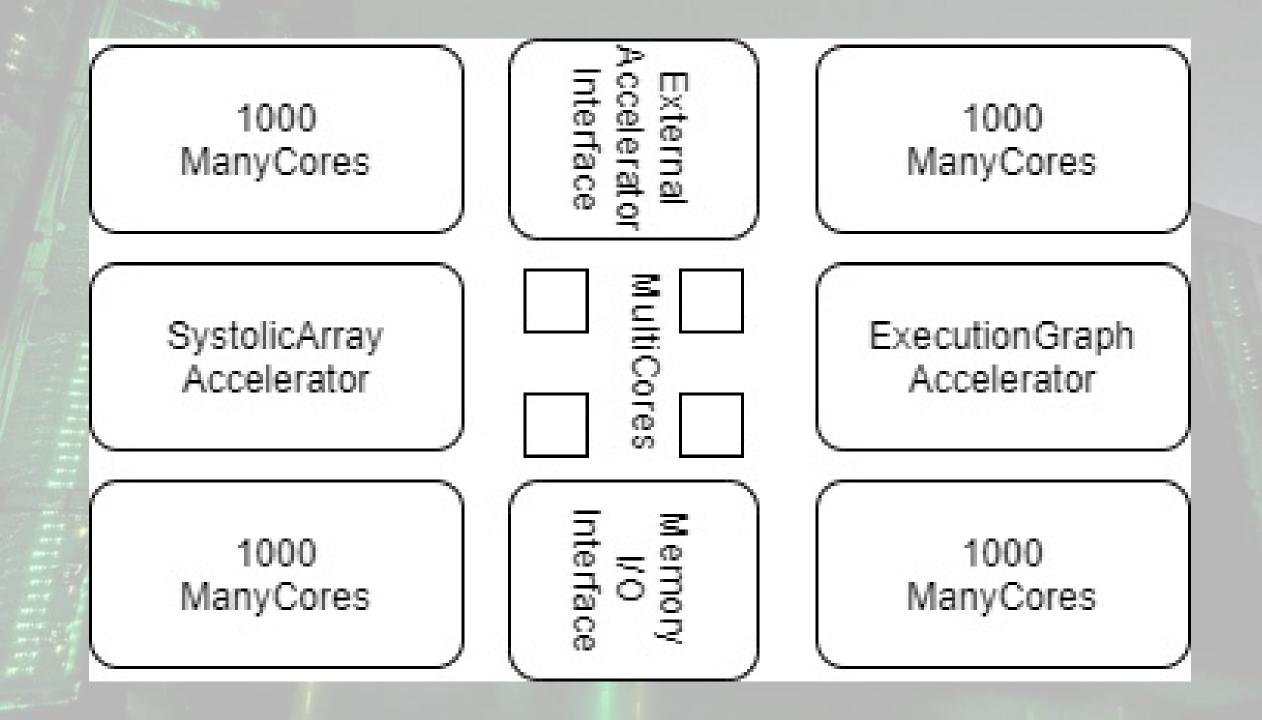
computer communications and recording

Veljko Milutinović Jakob Salom Nemanja Trifunovic Roberto Giorgi

Guide to DataFlow Supercomputing

Basic Concepts, Case Studies, and a Detailed Example





## **Current Research:**

January 31, 2020

#### The Ultimate DataFlow for Ultimate SuperComputers-on-a-Chips

Veljko Milutinovic, Erfan Sadeqi Azer, Kristy Yoshimoto, Indiana University, Bloomington, Indiana, USA

Gerhard Klimeck, Purdue University, IN, USA

Miljan Djordjevic, Milos Kotlar, Miroslav Bojovic, Bozidar Miladinovic, Nenad Korolija, and Stevan Stankovic, Universty of Belgrade, Serbia

Nenad Filipović, Zoran Babovic, Universty of Kragujevac, Serbia

Miroslav Kosanic, MIT, Cambridge, MA, USA

Akira Tsuda, Harvard University, Cambridge, Massachusetts, USA

Mateo Valero, BSC, Barcelona, Spain

Massimo De Santo, University of Salerno, Fisciano, Italy

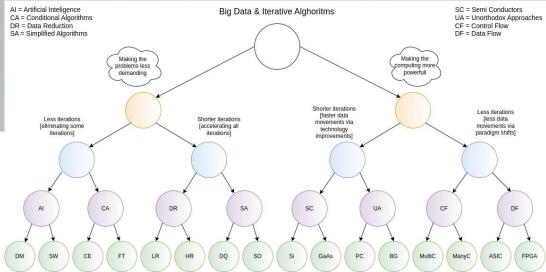
Erich Neuhold, UNIWIE and TUWIEN, Vienna, Austria

Jelena Skoručak, University of Zurich and ETH, Switzerland

Laura Dipietro, Highland Instruments, Cambridge, MA, USA

Ivan Ratkovic, Esperanto Technologies, Belgrade, Serbia and San Francisco, California, USA

This article starts from the assumption that near future 100BillionTransistor (100BT) SuperComputers-on-a-Chip will include N big multi-core processors, 1000N small many-core processors, an ASIC TPU-like fixed-structure systolic array accelerator for the most frequently used Machine Learning algorithms needed in bandwidth-bound applications and an FPGA flexible-structure re-programmable accelerator for less frequently used Machine Learning

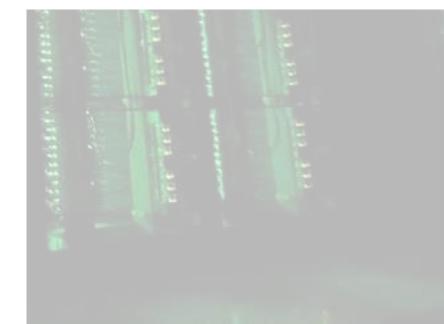


#### Making the problems less demanding:

DM = Data Mining SW = Semantic Web CE = Conditional Execution FT = Fast Transforms LR = Low Level Reductions HR = High Level Reductions DQ = Data Quantizations ( eg. binary arithmetic ) SO = Simplified Operations ( eg. no MLTP )

#### Making the computing more powerfull:

Si = Silicon GaAs = Gallium Arsenide PC = Physics/Chemistry BG = Biology/Genomics MultiC = Multi Cores ManyC = Many Cores ASIC = Aplication Specific Integrated Circuits (eg. Google TPU ) FPGA = Field Programmable Gate Arrays (eg. Maxeler DFE )

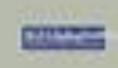


### Handbook of Research on Methodologies and Applications of Supercomputing

Handbook of Research on

Methodologies and Applications of Supercomputing

rite Wolfsen; and Wise, Keller



## Table of Contents #1

- An introduction to Controlflow and Dataflow Supercomputing
- Introduction to Control Flow
- Optimal Scheduling of Parallel Jobs With Unknown Service Requirements
- Intelligent Management of Mobile Systems Through Computational Self-Awareness
- Paradigms for Effective Parallelization of Inherently Sequential Graph Algorithms on Multi-Core Architectures
- Introduction to Dataflow Computing
- Data Flow Implementation of Erosion and Dilation
- Transforming the Method of Least Squares to the Dataflow Paradigm
- Forest Fire Simulation: Efficient Realization Based on Cellular Automata
- High Performance Computing for Understanding Natural Language
- Deposition of Submicron Particles by Chaotic Mixing in the Pulmonary Acinus: Aciner Chaotic Moxing

## Table of Contents #2

- Recommender Systems in Digital Libraries Using Artificial Intelligence and Machine Learning: A Proposal to Create Automated Links Between Different Articles Dealing With Similar Topics
- Unified Modeling for Emulating Electric Energy Systems: Toward Digital Twin That Might Work
- A Backtracking Algorithmic Toolbox for Solving the Subgraph Isomorphism Problem
- AI Storm ... From Logical Inference and Chatbots to Signal Weighting, Entropy Pooling...: Future of AI in Marketing
- Efficient End-to-End Asynchronous Time-Series Modeling With Deep Learning to Predict Customer Attrition
- Mind Genomics With Big Data for Digital Marketing on the Internet
- Supercomputing in the Study and Stimulation of the Brain
- An Experimental Healthcare System: Essence and Challenges
- What Supercomputing Will Be Like in the Coming Years
- The Ultimate Data Flow for Ultimate Super Computers-on-a-Chip

## Reviews and Testimonials of 8 Nobel Laureates



- "I want to commend Dr. Veljko Milutinovic and Dr. Milos Kotlar for having completed this volume in the timely field of supercomputing in these difficult times, and I want to share their optimism regarding its use as a textbook around the world."
- Prof. Kurt Wüthrich, Nobel Laureate, Switzerland

"Our complex and fast-moving world meets big data issues that call for reliable, efficient analysis and prompt response. A most advanced approach to dealing with these issues is presented in this book where algorithms of Control Flow represent the host architecture of supercomputers and Data Flow represents the acceleration architecture. A comprehensive list of applications is illustrated in the book including natural language processing, medical research, customer-oriented studies, and many more."

– Prof. Dan Shechtman, Nobel Laureate, Israel





- "Supercomputers have become a ubiquitous instrument in many areas of science and technology. Very hard to imagine modern physics, biology or chemistry without the use of this versatile tool. The breakthroughs in the development of supercomputers expand the range of problems we can tackle. Supercomputers, as well as specialised computers will undoubtedly contribute significantly to the overall landscape of discoveries in many different disciplines in the future."
- Prof. Konstantin Novoselov, Nobel Laureate, National University of Singapore, Singapore– Prof. Kurt Wüthrich, Nobel Laureate, Switzerland

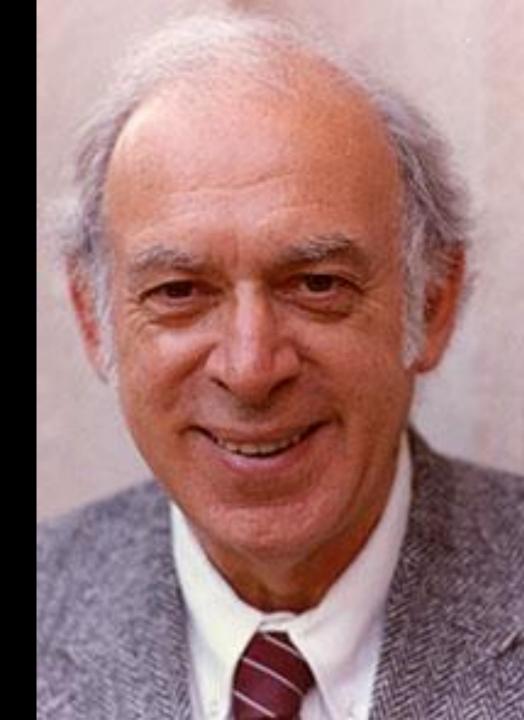
- "It's clear that computers can do anything, as the pioneers already recognised, and it's probably only a matter of time before they can perform any of the kinds of tasks we humans take for granted quicker than blinking. I guess this book represents a stage on this exciting and very important journey."
- – Prof. Tim Hunt, Nobel Laureate, UK





- "Computers have become essential tools for the pursuit of both experimental and theoretical physics, as well as synthetic chemistry, paleontology, the medical sciences, economics the social sciences and so much more. Science, being the most international of all endeavors, will be well served by this important book, which honors the golden anniversary of the creation of Montinegro's Academy of Science."
- Prof. Sheldon Glashow, Nobel Laureate, Boston University and Harvard University, USA

- "Humankind's continuing quests to uncover, understand, and utilize the secrets of nature have been greatly enhanced and will be further extended by the power of supercomputers."
- – Prof. Jerome I Friedman, Nobel Laureate, USA



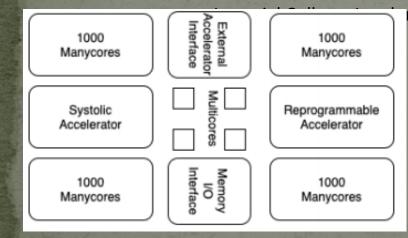


- "I wish to congratulate very warmly the Montenegrin Academy of Arts and Sciences on the occasion of its 50th Anniversary and wish the CANU a highly successful future. Science shapes the Future of Mankind."
- – Prof. Jean-Marie Lehn, Nobel Laureate, France

- "Aim high, stay grounded! These four crisp words of wisdom accompany the best wishes for the 50th anniversary of the Montenegrin Academy of Sciences and Arts."
- – Prof. Stefan Hell, Nobel Laureate, Germany



## Ultimate DataFlow SuperComputing for BigData DeepAnalytics



Estimated Transistor Count					
3.29 million					
0 million [17]					
illion [18]					
4 billion					
billion [19]					
billion [20]					
00 million					
00 million					
00 billion					
Interface to External Accelerators         <100 million					

that, for future 100 billion transistor chips, the most effective resources to include are those based on the dataflow principle. For some important applications, such resources bring significant speedups, that would fully justify the incorporation of additional 70 billion transistors. The speedups could be, in reality, from about 10x to about 100x, and the explanations follow in the rest of this article.

23

### **Major Sources of Inspiration**

A. Richard Feynman:

Impact of logic/arithmetic and memory/IO Compiler-generated execution graph

**B. Ilya Prigogine:** 

Impact of energy, entropy, order, and optimization Compiler-generated data separation

C. Daniel Kahneman:

Impact of approximate computing on precision Compiler-controlled approx computing

D.Tim Hunt:

Impact of system latency on precision Compiler-controlled system latency

## The Major Axiom of Optimal Computing

#### A. Whenever the **Technology** changes,

the Fundamental Paradigm of Computer Architecture has to change, too.

#### aSoG (not: FPGA)

 B. If several paradigms are available, the most suitable paradigm for adoption is the one most effective for modern Applications. BrontoData (not: ExaBigData)

Is the von Neumann Paradigm still the most effective one?

A. MultiCores?B. ManyCores?

### The Holy Trinity of Generalized Computing



## Architecture

## Technology

- Size

- Power

## The von Neumann Paradigm (1940s)

# $\lim_{i \to \infty} \left( \frac{TALU(i)}{TCOMM(i)} \right) \to \infty$

## **Optimal Solution: Finite Automata**

The Nobel Laureate Richard Feynman Observations

# $\lim_{i \to \infty} \left( \frac{TALU(i)}{TCOMM(i)} \right) \to 0(t \to \infty)$

Where is the technology now? A. Closer to 1940s? B. Closer to  $t \rightarrow \infty$ ?

## State of the Art in Technology Today The Power Challenge The Data Movement Challenge

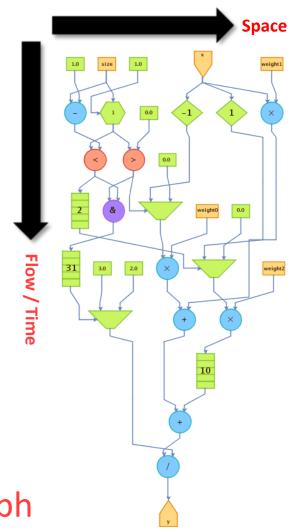
	2015	2020
Double precision FLOP	100pj	10pj
пеногу		

- Moving data off-chip will use 200x more energy than computing!
- Moving data in 1940s was using **1/60x** ...
- Conclusion: We are getting close to the Feynman Asymptote!
- Important: Power and speed could be traded!

# The Maxeler Technology Vision: MultiScale DataFlow

- Thinking in space rather than in time
- Difficult change in mindset to overcome
- □Transformation of data through flow over time
- Instructions are parallelized across the available space

**Optimal Solution: Execution Graph** 



## Comparing the Two Approaches

• The Von-Neumann paradigm resembles an old wall clock

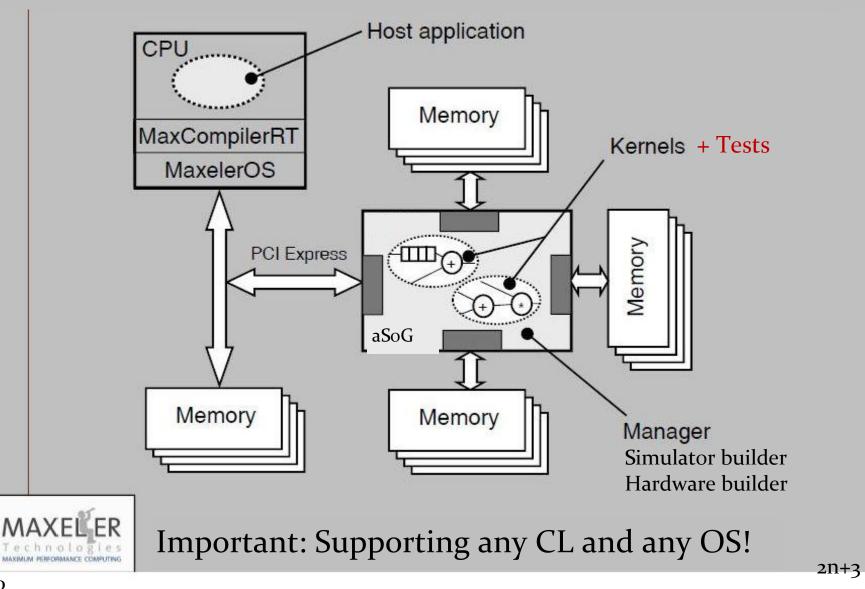


• The Feynman paradigm resembles lightning! Why?

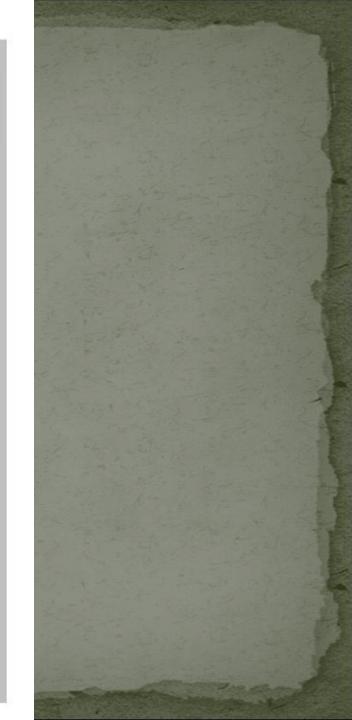
## **Programming the Two Paradigms**

von Neumann: The Program Moves Data Feynman: The Program Configures Hardware What moves data? External sources till input. Voltage difference through aSoG! Voltage difference moves the important stuff!

## The Maxeler Generic Architecture Application



11/60



## Why The Acceleration Approach?

Nobel Laureate Ilya Prigogine: Injecting Energy to Decrease Entropy!

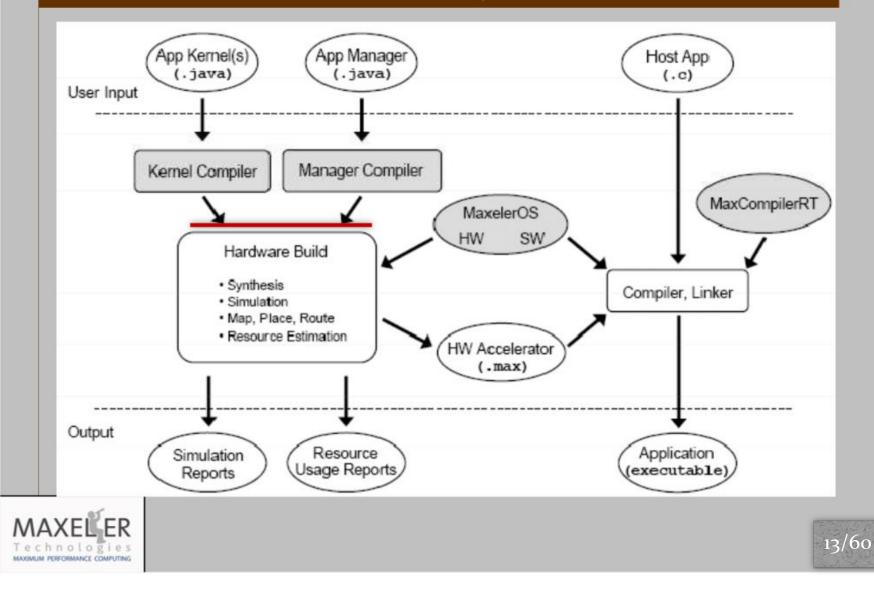
#### Corollary:

Burning energy to split spatial and temporal decreases the entropy of computing and enables the DataFlow compiler to create a maximally effective execution graph.

#### Final goal:

The execution graph with the minimal length of edges.

#### MaxCompiler



## Alliances Being Formed

Intel acquired AlteraQualcomm and IBM teaming up with Xilinx

However:

OpenCL(I)

C

Intel

Altera

C

OpenCL(A) MaxCompiler

Altera@MaxZ

C

X>>1

Y>>X

# Nano Accelerators

Invisible on the DataFlow Concept Level
Invisible to DataFlow Programmers
Visible to the MaxCompiler
The MaxCompiler knows how to utilize them

Best protected by two aSoG (now FPGA) protection levels and two Vendor (e.g., Maxeler) protection levels!

## **Publications of Interest for NanoAcceleration**

- 1. Flynn, M., Mencer, O., Milutinovic, V., Rakocevic, G., Stenstrom, P., Trobec, R., Valero, M., Moving from Petaflops (on Simple Benchmarks) to Petadata per Unit of Time and Power (On Sophisticated Benchmarks), **Communications of the ACM** (nano-acceleration), May 2013.
- 2. Trobec, R., Vasiljevic, R., Tomasevic, M., Milutinovic, V., Beiveide, M., Valero, M., Interconnection Networks for SuperComputing, ACM Computing Surveys (nano-acceleration), 2017.
- 3. Milutinovic, V., Tomasevic, M., Markovic, B., Tremblay, M., The Split Temporal/Spatial Cache: Initial Performance Analysis, Proceedings of the SCIzzL-5, Santa Clara, California, USA, March 26, 1996, pp 72-78.
- 4. Milutinovic, V., Tomasevic, M., Markovic, B., Tremblay, M., The Split Temporal/Spatial Cache: Initial Complexity Analysis, Proceedings of the SCIzzL-5, Santa Clara, California, USA, September, 1996.
- 5. Milutinovic, V., A Comparison of Suboptimal Detection Algorithms Applied to the Additive Mix of Orthogonal Sinusoidal Signals, IEEE Transactions on Communications, Vol. COM-36, No. 5, May 1988, pp. 538-543.
- 6. Milutinovic, V., Mapping of Neural Networks on the Honeycomb Architectures, **Proceedings of the IEEE**, Vol. 77, No 12, December 1989, pp. 1875-1878.

8

7. Helbig, W., Milutinovic, V., The RCA's DCFL E/D MESFET GaAs 32-bit Experimental RISC Machine, **IEEE Transactions on Computers**, vol. 36, No. 2, February 1989, pp. 263-274.

Inspired by: Hunt

In Turne FFGA Accelerator for Floating-Point Matrix Multiplication, IEE Computers & Digital Techniques (nano-acceleration), 2012, 6, (4), pp. 249-256. The IET 2014 Premium Award for Computing & Digital Techniques.

Inspired by: Feynman

Inspired by: Prigogine

Inspired by: Kahneman

## Feynman

THE INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS THE INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS ADDRESS CONTRACTOR OF THE INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS ADDRESS CONTRACTOR OF THE INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS ADDRESS CONTRACTOR OF THE INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS ADDRESS CONTRACTOR OF THE INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS ADDRESS CONTRACTOR OF THE INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS ADDRESS CONTRACTOR OF THE INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS ADDRESS CONTRACTOR OF THE INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS ADDRESS CONTRACTOR OF THE INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS ADDRESS CONTRACTOR OF THE INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS ADDRESS CONTRACTOR OF THE INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS ADDRESS CONTRACTOR OF THE INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS ADDRESS CONTRACTOR OF THE INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS ADDRESS CONTRACTOR OF THE INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS

# $I = \log_2(N_{Pe} + 1)$ $N_{kolol}(I) = (I - 1)(2^{l-1} + 2^{l-2})$ $N_{bus}(I) = (I - 2)(2^{l-2} + 2^{l-3})$ $N_{kile}(I) = (I - 1)(2^{l-1} + 2^{l-2}) - (I - 2)(2^{l-2} + 2^{l-3}) - (2^{l} - 1)$

$$\begin{split} & N_{pe}(l) = 2^{l} - 1 \\ & N_{total}(N_{pe}) = \left[ \log_{2}(N_{pe} + 1) - 1 \right] \left[ \frac{N_{pe} + 1}{2} + \frac{N_{pe} + 1}{2^{2}} \right] \\ & N_{total}(N_{pe}) = \left[ \log_{2}(N_{pe} + 1) - 1 \right] \left[ \frac{3N_{pe} + 3}{4} \right] \\ & N_{total}(N_{pe}) = \frac{3}{4} (N_{pe} + 1) \left[ \log_{2}(N_{pe} + 1) - 1 \right] \\ & N_{bas}(l) = \frac{3}{8} (N_{pe} + 1) \left[ \log_{2}(N_{pe} + 1) - 2 \right] \\ & N_{idle}(N_{pe}) = \frac{3}{4} (N_{pe} + 1) \left[ \log_{2}(N_{pe} + 1) - 1 \right] - \frac{3}{8} (N_{pe} + 1) \left[ \log_{2}(N_{pe} + 1) - 2 \right] \\ & N_{idle}(l) = \frac{3}{8} (N_{pe} + 1) \left[ \log_{2}(N_{pe} + 1) - 1 \right] - \frac{3}{8} (N_{pe} + 1) \left[ \log_{2}(N_{pe} + 1) - 1 \right] - \frac{3}{8} (N_{pe} + 1) \left[ \log_{2}(N_{pe} + 1) \right] \\ & N_{idle}(l) = \frac{3}{8} (N_{pe} + 1) \left[ \log_{2}(N_{pe} + 1) \right] - \frac{1}{8} (11N_{pe} + 3) \end{split}$$

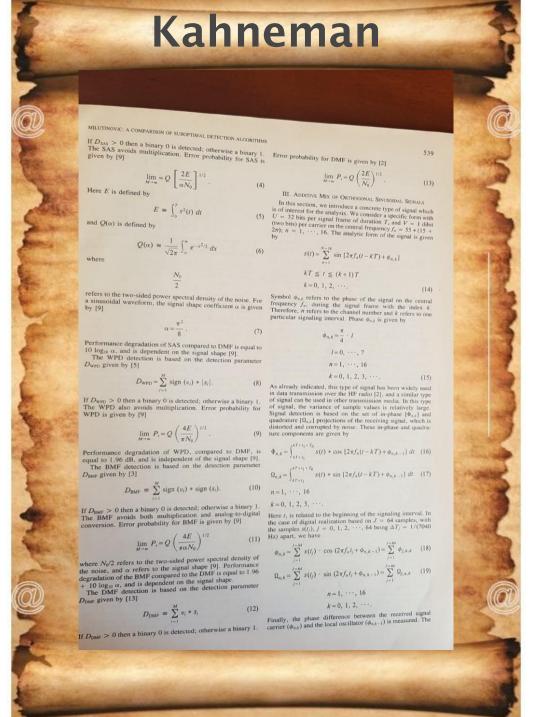
F.4.17.1

 $U(l) = \frac{N_{pe}(l) + N_{bus}(l)}{N_{total}(l)} = \frac{2^{l} - 1 + (l-2)(2^{l-2} + 2^{l-3})}{(l-1)(2^{l-1} + 2^{l-2})}$  $U(N_{pe}) = \frac{N_{pe} + \frac{3}{8} (N_{pe} + 1) \left[ \log_2 (N_{pe} + 1) - 2 \right]}{\frac{3}{4} (N_{pe} + 1) \left[ \log_2 (N_{pe} + 1) - 1 \right]}$ 

f.4.18.1

Prigogine TINOVIC et al.: MULTIMICROPROCESSOR ARCHITECTURE FOR REALTIME C COMPARISON OF PERFORMANCE OF FFT/SIMD AND DFT/ when and / FFT/SIMD DFT/MISD One proach on the  $2 \log N + wN$  $\max \{wN + 1, N + w\}$ input ity of (3N/2) - 2H arate IN inputs cN/2 + qc(N-2)of inp clN set in an FF  $\frac{(\log 4 lN) - 2l}{2 \log N + wN}$  $\log 4lN) - 2l$ for ir  $2l \max \{wN + 1, N + w\}$ still t be d Whe whet  $\sum_{i=0}^{\log N-1} c_i = \sum_{i=0}^{X-1} 2^i L + \sum_{i=X}^{\log N-1} N/2$ delay It for 1 SIM  $= L(2^{X} - 1) + \frac{N}{2} \log (N - X)$ tion the  $= (N/2)(\log 4L) - L.$ be cap The serial execution time of the FFT is then  $((N/2) \log as$ 4L - L) which is always smaller than LN, the serial ex-solution only of N and we ecution time of the DFT, for any positive values of N and wo L. Therefore, it is the best serial execution time. This statement presumes that the units of time are equally de-fined for the statement of time are equally de-Ined for both the FFT and DFT algorithms. As we men-ioned be the formed by the formed tioned before, this is not true in our analysis and the re in sulting and the re in our analysis and the re in sulting and the re in such as the re in sulting and the re in such as the re in suc Sulting error favors the FFT algorithm. Given that the im

/60



/60

## Hunt

MILUTINOVIĆ et al.: GaAs-BASED MICROPROCESSOR ARCHITECTURE FOR REAL-TIME In the case of bo looping, as indicated in Fig. 5, the In the cuse of the new value for the loop control variable is computation of the new value for the loop control variable is computation writing the new value back into the multiport followed by both writing the new value back into the multiport data memory, and testing the condition. The longer of these data memory will determine the total execution time.  $= \begin{cases} T_{\text{GOTO}} + T_{A/C}(K_{A/C} = 1, K_B; N_{\text{OPER}} = N_{\text{OPER}}^{\text{eff}}); \\ (N_{G/\text{RELA}} + N_{G/BR}) * T_G > T_{MM} \\ T_{\text{GOTO}} + T_{A/C}(K_{A/C} = 0; N_{\text{OPER}} = N_{\text{OPER}}^{\text{eff}}); \\ (N_{G/\text{RELA}} + N_{G/BR}) * T_G \le T_{MM} \end{cases}$  $T_{\rm DO} = \cdot$ (15)

where  $N_{G/RELA}$  is the number of gate delays for the unit which determines the value of the HLL relation, and  $N_{\text{OPER}}^{\text{eff}} = \max$  $\{(N_{OPER/I} + 1), N_{OPER/F}, (N_{OPER/S} + 1)\}$ . Symbols *I*, *F*, and S refer to INITIAL, FINAL, and STEP expressions in the primary control statement. By this we conclude the derivation of execution-time formulas for assignments and control constructs.

Now we concentrate on execution-time formulas for call/ Fi return and low-level I/O. We assume the same number of input and output ports of the multport data memory  $(N_{MM})$ , and certain number of input and output subroutine parameters  $(N_{\text{PARA}}; N_{\text{PARA}} \neq N_{MM})$ . Consequently,

## $T_{\text{CALL/RETURN}} = T_{\text{CALL}} + T_{\text{RETURN}} = 2$ \* $\left[T_{PM} + \left(\max\left\{1, \left\lceil\frac{N_{PARA}}{N_{MM}}\right\rceil\right\} - 1\right)\right]$ \* max { $T_F$ , 2 \* $T_{MM} + T_G$ } + max $[N_{G/BR} * T_G]$ , $[sgn (N_{PARA})$ (16) $* (2 * T_{MM} + T_G)]$



Special Acknowledgements to: Simon Aglionby, Georgi Gaydadjiev, Itay Greenspon, and Nemanja Trifunovic

## IF(2012)=3.80

Read Edit View history Search

## ACM Computing Surveys

From Wikipedia, the free encyclopedia

**ACM Computing Surveys** (CSUR) is a peer reviewed scientific journal published by the Association for Computing Machinery. The journal publishes survey articles and tutorials related to computer science and computing. It was founded in 1969; the first editor-in-chief was William S. Dorn.<sup>[1]</sup>

In ISI Journal Citation Reports, *ACM Computing Surveys* has the highest impact factor among all computer science journals.<sup>[2]</sup> In a 2008 ranking of computer science journals, *ACM Computing Surveys* received the highest rank "A\*".<sup>[3]</sup>

## See also [edit]

ACM Computing Reviews

## References [edit]

- 1. ^ Dorn, William S. (1969). "Editor's Preview...". ACM Computing Surveys. 1 (1): 2–5. doi:10.1145/356540.356542 2.
- 2. \* "Journal Citation Reports" . ISI Web of Knowledge. Retrieved 2009-10-03. "JCR Science Edition 2008"; subject categories "COMPUTER SCIENCE, ...".
- 3. "Journal Rankings" &. CORE: The Computing Research and Education Association of Australasia. July 2008. Archived & from the original on 29 March 2010. Retrieved 2010-03-19..

## External links [edit]

- ACM Computing Surveys home page ₽.
- ACM Computing Surveys
   in ACM Digital Library.
- ACM Computing Surveys
   in DBLP.



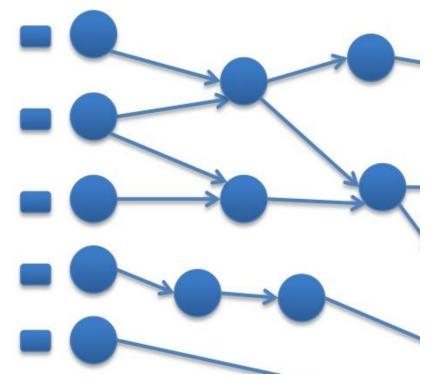
ACM Computing Surveys			
Abbreviated title (ISO 4)	ACM Comput. Surv.		
Discipline	Computer science		
Language	English		
Edited by	Sartaj K Sahni		
Publication details			
Publisher	ACM (United States)		
Publication history	1969-present		
Frequency	Quarterly		
Index	ing		
ISSN	0360-0300 & (print) 1557-7341 & (web)		
Links			
🔹 Journal homepage 🗗			
• Online access 🗗			
<ul> <li>Online archive</li></ul>			

Q

# ce: Feynman Enabled by Prigogine

- TALU possible at zero power (Arithmetic+Logic)
- TCOMM not possible at zero power (MEM+MPS)



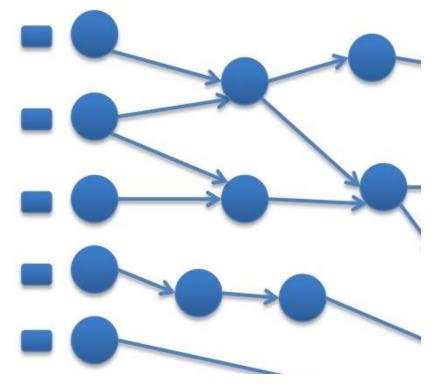


# Essence: Feynman

- TALU possible at zero power (Arithmetic+Logic)
- TCOMM not possible at zero power (MEM+MPS)



PM



## PM

# Essence: Feynman

- TALU possible at zero power (Arithmetic+Logic)
- TCOMM not possible at zero power (MEM+MPS)





# Programming the Maxeler Technology Generic Acceleration Architecture

MaxJ, the Maxeler Java,

a DSL acting as a SuperSet of classical Java: A. A vector of built-in domain-specific classes B. Two sets of variables: SW + HW

MaxJ is a SubSet of OpenSPL, created by the Imperial-Stanford-Tokyo-Tsinghua consortium.

Possible Future Mutations of OpenSPL: MaxPython and/or MaxR (lower Kolmogorov complexity) MaxHaskel and/or MaxScala (easier extension to approximate computing).

# Approximate Computing for Better Precision: Kahneman

Note: Small approximations in one domain may bring large benefits in another domain

Example: Weather forecast

A 15-bit computational precision (rather than the 64-bit precision) may decrease the forecast precision for only 2%, and at the same time, may increase the grid precision 25 times, and the forecast precision at grid intersections up to 10<sup>4</sup>.

49

Easily doable in DataFlow, difficult to do in ControlFlow.

# Delayed Decision for Better Precision: Hunt

Note: Small latencies in time domain may bring large benefits in precision domains

Example: Optimal utilization of internal DataFlow pipelines

Compiler optimizations create internal pipelines that experienced DataFlow programmers know how to utilize

## **BigDataAnalytics**

## Existing Maxeler-based publications:

20 [Size] 20 [Power] 20, 200 [Speedup] 20 [Precision]



## Ultimate aSoG-based future:

20-200 [Size] 20-200 [Power] 20, 200, 2000, 20000 [Speedup] 20+ [Precision]

51

Architecture

# Technology

Maxeler Dataflow Appliance

- Software Based Solution
- Dataflow Computing in the Datacentre



**The CPU** Conventional CPU cores and up to 6 DFEs with 288GB of RAM

## 1U: Good for Fog





**The Dataflow Appliance** Dense compute with 8 DFEs, 768GB of RAM and dynamic allocation of DFEs to CPU servers with zero-copy RDMA access



40U: Good for Cloud

**The Networking Appliance** Intel Xeon CPUs and 4 DFEs with direct links to up to twelve 40Gbit Ethernet connections





MicroMAX.5: Good for Dew (Edge Processing of IoT Data)





# **The Major Application Successes**

## • Finances:

- Credit derivatives
- Risk assessment
- Stability of economical systems
- Evaluation of econo-political mechanisms
- GeoPhysics:
  - Oil&Gas
  - Weather forecast
  - Astronomy
  - Climate changes
- Science:
  - Physics
  - Chemistry
  - Biology
  - Genomics
- Engineering: Synergy of all the Above (ML, etc...)

53

## J.P.Morgan

## Innovation in Investment Banking Technology Field Programmable Gate Arrays (FPGAs)

A Field Programmable Gate Array (FPGA) is a silicon chip containing a matrix of configurable logic blocks (CLBs) that are connected through programmable interconnects. By combining optimized use of available silicon with fine-grained parallelism, sustained acceleration improvements of over 300x can be achieved across a range of vanilla and complex mathematical models. The current work is the first time that FPGA technology has been employed at this scale to accelerate computational performance anywhere in the finance industry.

### **Power and Versatility**

- Can accelerate performance by between 100 and 1,000x across a range of mathematical models, with the ability to perform a task in less than a second
- Can be reprogrammed and precisely configured to compute exact algorithm(s) at the desired level of numerical accuracy required by any given application, unlike normal microprocessors whose design is fixed by the manufacturer
- Can be deeply pipelined to achieve maximum parallelism from arithmetic, algorithms and data streaming

### Key Business Challenges

- Reduce the execution time of existing applications to meet business and regulatory demands
- Decrease cost of running existing applications and developing new ones
   Provide fast, cost-effective extra computational capacity to address
- problems that are currently inextricable
- Achieve a step-change improvement in price-performance and end-to-end
   compute time across many applications

### Key Benefits (Business/Clients)

- Competitive advantage to valuation, execution, risk management and complex scenario analyses by speeding up existing applications
- Lower cost of existing applications as hardware costs can be reduced by a factor between 100 and 1,000
- Ability to perform previously difficult calculations, such as complex trading strategies or risk evaluations of global portfolio simulations.

# Technology Overview - Low clock speed chips - Maximal usage of available silicon resources - Acceleration through use of fine-grained parallelism - Reconfigurable hardware - Silicon configurable to fit algorithm

### LOB/Function(s) Impacted

- Credit & interest rates
   Equities & commodities
   Loan & mortgage modeling
   Finance & accounting
- High frequency trading Risk management & VaR

### Industry/External Recognition

- Used by Cisco in all routers Simulation of real and theoretical systems
- Geophysics for oil and gas exploration
   Astrophysics & hydrodynamics
- Defense for cryptography
- Video games
- Genotyping
- cunt Aburd

### Technology ignites our business. Be the spark.

### **Functionality Overview**

Double precision floating point-capable FPGAs became commercially available in 2002, but it was the arrival of the Virtex 5 and 6 series chips from market leader Xilinx that really provided the scale required for the development of production-grade accelerated solutions. Using FPGAs in high performance compute solutions provides distinct advantages over conventional CPU clusters.

### **Operational Advantages**

- Significantly increases performance for two main types of applications: those based around highly complex mathematical models and those using simpler algorithms that can be massively parallelized
- Enables a dramatic increase in compute density per cubic meter by using FPGAs as computational accelerators
- Consumes around 1% of the power of a single CPU core

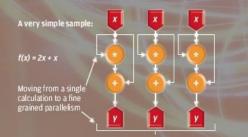
### Performance Improvements

- Performance improvements in the range 200-300x faster than the existing CPU cores used on the Compute BackBone (CBB) have been achieved in credit and interest rates hybrids businesses
- In equities, direct market access can run risk and loan stock at wire speed (3.5 micro secs) using a low-latency FPGA solution
- Benchmarked average throughput for J.P. Morgan's existing 40-node hybrid FPGA machine of 984MFlops/watt/cubic meter
- Potential standing at the top of the Green-500 ecological global supercomputer performance table

### FPGAs at Work

### Work

- An algorithm is implemented as a special configuration of a general purpose electric circuit
- Connections between prefabricated wires are programmable
- Function of calculating elements is itself programmable
   FPGAs are two dimensional matrix-structures of configurable logic blocks (CLBs) surrounded by input/output blocks that enable communication with the rest of the environment



## J.P.Morgan

### Development/Delivery

### Timeline

- Initial porting of an algorithm can vary from one to three months depending on complexity.
- Production capabilities then depend on the scale of the application and the scope and intensity of the testing and reconciliation cycle

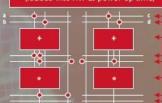
#### Partners

 Condon-based Applied Analytics group: includes three technology and business specialists with extensive experience in developing and delivering high performance solutions across a range of asset classes, models and lines of business
 Maxeler Technologies: external consultants trained in Imperial College, Stanford and MIT research labs

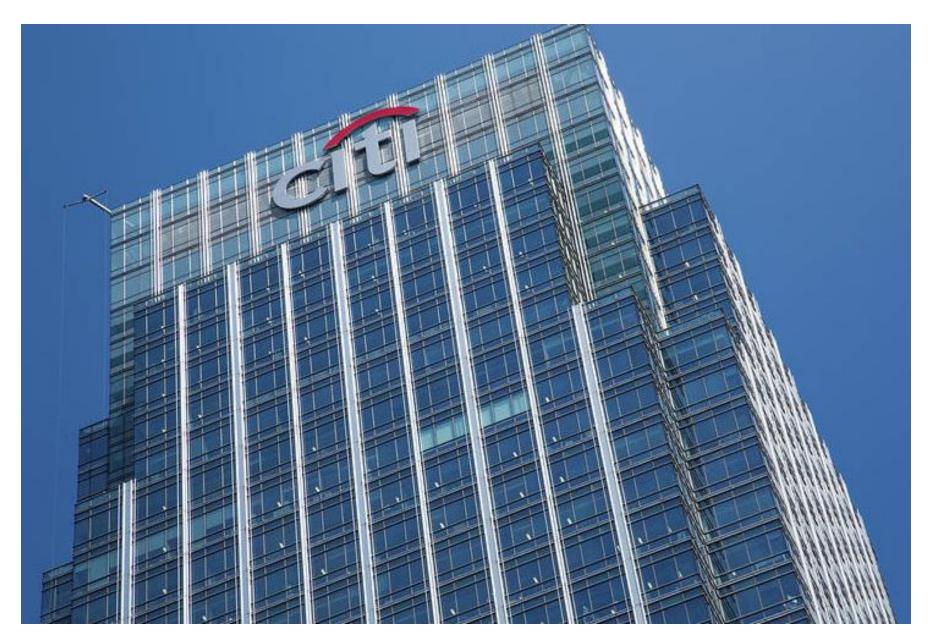
### A slightly more complex example:

e = (a+b)\*(c+d)

### Configuration Memory (loaded into HW at power up time)



Migrating algorithms from C++ to FPGAs involves doing a Fourier Transform from time domain execution to spatial domain execution in order to maximize computational throughput. It's a paradigm shift to stream computing that provides acceleration of up to 1,000x compared to an Intel CPU.



The know-how needed for deep security!



 $\leftarrow \rightarrow$ 

Designed for educational use only using Maxeler Technologies' curve construction methodology. This tool uses delayed data and displayed results are indicative representations only.

	DSF Pricing						
CME Ticker	Ticker	Price	Coupon	PV01	NPV	Implied Rate	Timestamp
T1UM4 2Y	CTPM4	100'057	0.750%	<mark>\$19.</mark> 97	\$179.69	0.6600%	4:00:03 PM CT 4/4/2014
F1UM4 5Y	CFPM4	100'115	2.000%	<mark>\$48.4</mark> 9	\$359.38	1.9259%	4:00:03 PM CT 4/4/2014
N1UM4 10Y	CNPM4	100'225	3.000%	\$90.16	\$703.12	2.9220%	4:00:03 PM CT 4/4/2014
B1UM4 30Y	CBPM4	102'270	3.750%	\$195.07	\$2,843.75	3.6042%	4:00:03 PM CT 4/4/2014
T1UU4 2Y	CTPU4	100'085	1.000%	<mark>\$1</mark> 9.93	\$265.62	0.8668 <mark>%</mark>	4:00:03 PM CT 4/4/2014
F1UU4 5Y	CFPU4	100'110	2.250%	<mark>\$</mark> 48.27	\$343.75	2.1788%	4:00:03 PM CT 4/4/2014
N1UU4 10Y	CNPU4	101'125	3.250%	\$89.55	\$1,390.62	3.0948%	4:00:03 PM CT 4/4/2014
B1UU4 30Y	CBPU4	106'020	4.000%	\$193.47	\$6,062.50	3.6868%	4:00:03 PM CT 4/4/2014

Please hover your mouse pointer over column titles and links for further information.

Quotes and analytics are updated every 15 minutes.

## (K) Analytics powered by Maxeler Technologies®

Instrument	CPU 1U-Node	Max 1U-Node	Comparison
European Swaptions	848,000	35,544,000	42x
American Options	38,400,000	720,000,000	19x
European Options	32,000,000	7,080,000,000	221x
Bermudan Swaptions	296	6,666	23x
Vanilla Swaps	176,000	32,800,000	186x
CDS	432,000	13,904,000	32x
CDS Bootstrap	14,000	872,000	62x

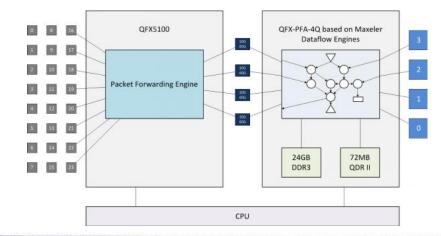


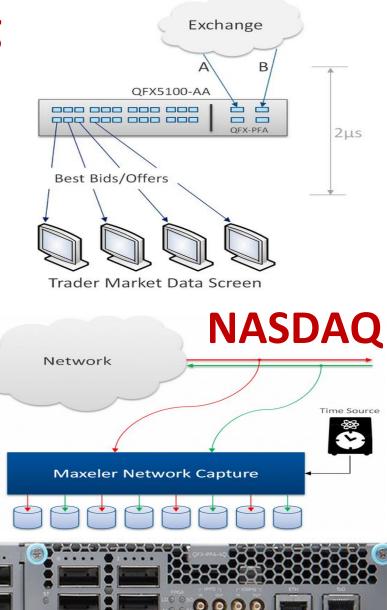
\$3

Ξ

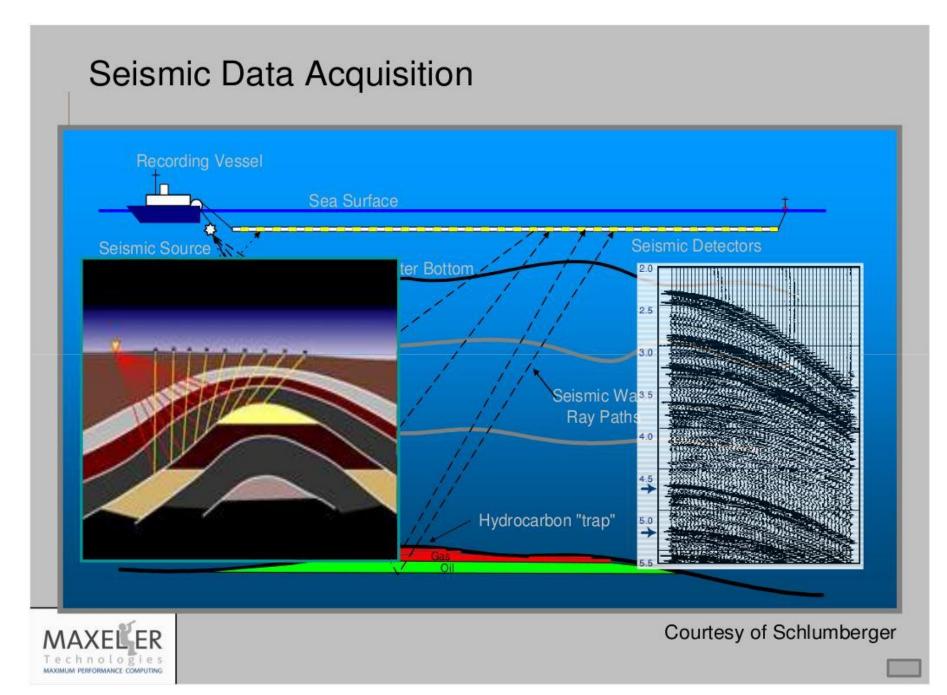
# **Juniper for Online Trading**



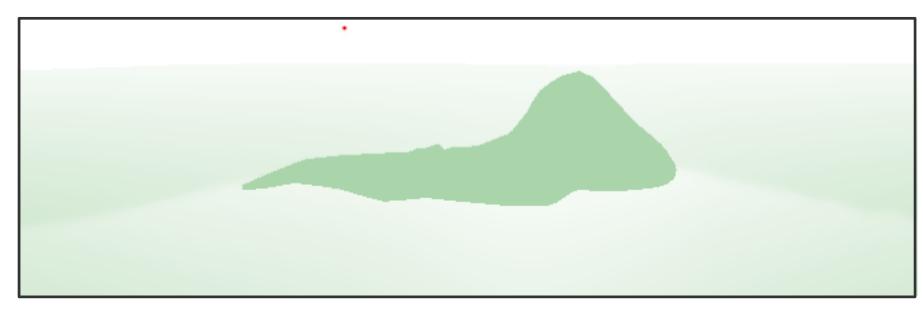








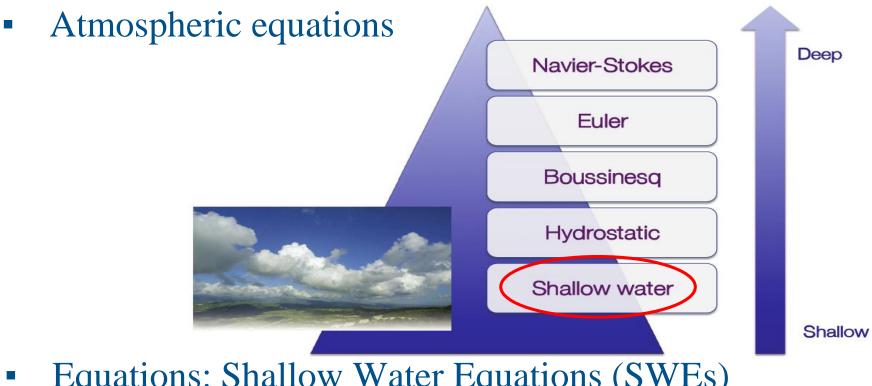
# Seismic Imaging



- Running on MaxNode servers
  - 8 parallel compute pipelines per chip
  - 10x less power: 150MHz vs 1.5GHz
  - 30x faster than microprocessors

**An Implementation of the Acoustic Wave Equation on FPGAs** T. Nemeth<sup>†</sup>, J. Stefani<sup>†</sup>, W. Liu<sup>†</sup>, R. Dimond<sup>‡</sup>, O. Pell<sup>‡</sup>, R.Ergas<sup>§</sup> <sup>†</sup>Chevron, <sup>‡</sup>Maxeler, <sup>§</sup>Formerly Chevron, SEG 2008

# Global Weather Simulation: Size is Relevant



Equations: Shallow Water Equations (SWEs)  $\frac{\partial Q}{\partial t} + \frac{1}{\Lambda} \frac{\partial (\Lambda F^{1})}{\partial x^{1}} + \frac{1}{\Lambda} \frac{\partial (\Lambda F^{1})}{\partial x^{2}} + S = 0$ 

London

ΛΛΛΧΕΪ

Omhutina

[L. Gan, H. Fu, W. Luk, C. Yang, W. Xue, X. Huang, Y. Zhang, and G. Yang, Accelerating solvers for global atmospheric equations through mixed-precision data flow engine, FPL2013] Tsinghua 60

Imperial College

## Weather Model – Performance Gain

Platform	Performance	Speedup
	()	
6-core CPU	4.66K	1
Tianhe-1A node	110.38K	23x
MaxWorkstation	468.1K	100x
MaxNode	1.54M	330x
Meshsize: 102	$24 \times 1024 \times 6$	14x

MaxNode speedup over Tianhe node: 14 times





## Weather Model -- Power Efficiency

Platform	Power Efficiency	Speedup
6-core CPU	20.71	1
Tianhe-1A node	306.6	14.8x
MaxWorkstation	2.52K	121.6x
MaxNode	3К	144.9x

Meshsize:  $1024 \times 1024 \times 6$ MaxNode is 9 times more power efficient

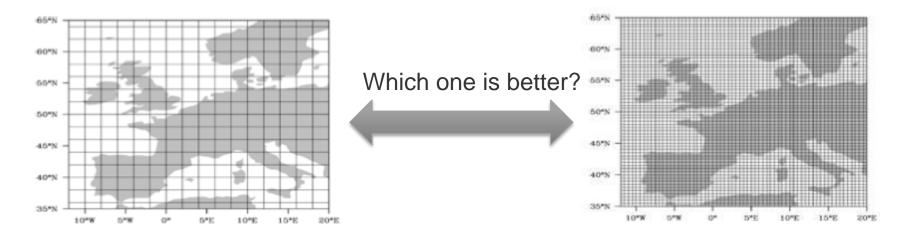






**9** x

## Weather and Climate Models: Precision



Finer grid and higher precision are obviously preferred but the computational requirements will increase → Power usage 2 \$\$

We use only 15 bits for 98% of the computation:

What about using reduced precision? (15 bits instead of 64 double precision FP)

63

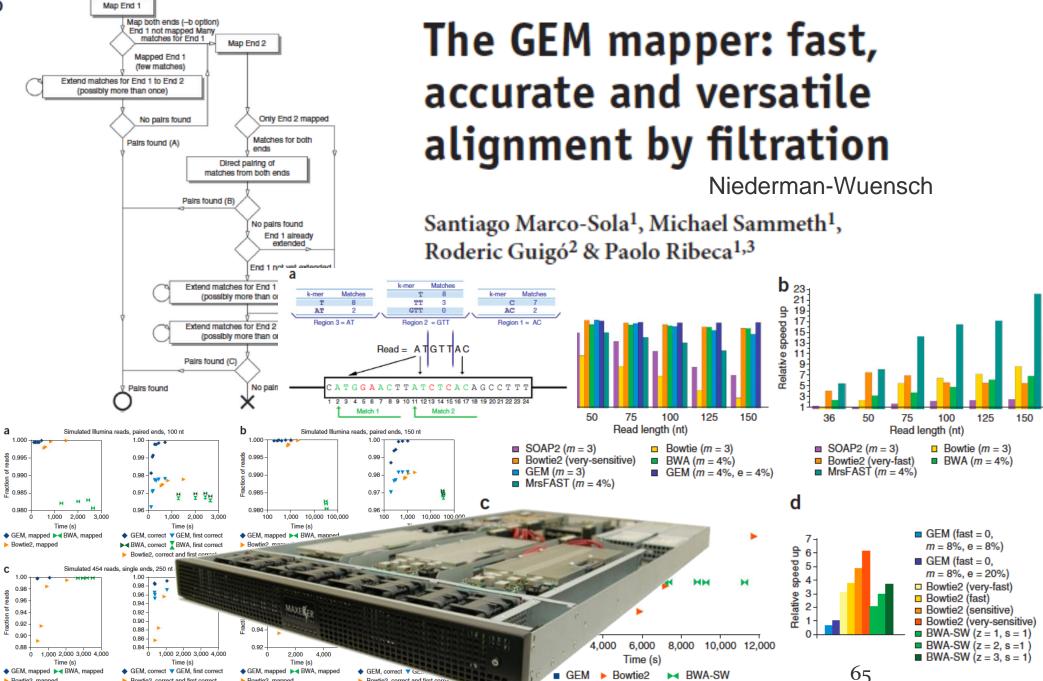


# Maxeler Running Smith Waterman

📕 Smith Wat	terman Demo - Maxeler Technologies
GCA AGA GAT AAT TGT	Query : UniRef50_F2T2I7 Histone-lysine N-methyltransferase n=8 Tax=E(1280)Best scores : splQ1DR06[SET1_COCIM Histone-lysine N-methyltransferase, H3(1271)4077splQ2UMH3[SET1_ASPOR Histone-lysine N-methyltransferase, H3(1229)3849splQ2UMH3[SET1_ASPFU Histone-lysine N-methyltransferase, H3(1220)3683splQ5B0Y5[SET1_EMENI Histone-lysine N-methyltransferase, H3(1220)3683splQ8X0S9[SET1_NEUCR Histone-lysine N-methyltransferase, H3(1220)3683splQ4WNH6[SET1_GIBZE Histone-lysine N-methyltransferase, H3(1252)2150splQ2GWF3[SET1_CHAGB Histone-lysine N-methyltransferase, H3(1252)2150splQ2GWF3[SET1_CHAGB Histone-lysine N-methyltransferase, H3(1076)2089splQ6CEK6[SET1_YARLI Histone-lysine N-methyltransferase, H3(1088)985splQ5ABG1[SET1_CANAL Histone-lysine N-methyltransferase, H3(1170)938splQ5ABG1[SET1_CANAL Histone-lysine N-methyltransferase, H3(1040)893
Ala       Arg       Asp       Asn       Cys       •         1       2       3       4       5         uniprot_sprotfasta       •         Number of sequences : 532224       •         Number of residues : 188726448       •         Scoring matrix :	Best alignment : MSRASAGFADFFPTAPSVLQKKRSSKAAQDRPKGKLKHDDDPQSSNPAPTAATAAVTVTGVGVPGAEEGGASDNNTNSDV MSRAPAGFADFFPTAPSVLQKKRS-KAAQDR-HAANTPKAADPLPNLGLSS-TPDIK-GGVGTSAD- HNNINSNNNNKNNSSSHTNINSNTQFDESAGAVARGDVNITPGDANGVGSSSSTSTGSS-VFSASILPQPGLTTSNGITH -NPVRAVGE-RSAE-TT-LALGDTNG-ATSSSSLSTGSSGFFSASA-P-PGVAKPNGISS PHALTPLINTDSSPSCKIASPSQKS-IA-ATGEIVPTSRFVDDIK-ATITPLQTPPTPRIQARPAGNAPKGYKLTYDPD C-ALTPLINTDSSPPCKIESPLGSKSGSTDAAPQLAPTCEAHGGPEPVTITPLHTPPTPRVQARPANSEVKGHKITYDPD
BLOSUM62	LERK-PLTKEKRRKPQYEVFDTTED-EAPPADPRIAIANYTRGAGCKQKTKYRPAPYILRPWPYDPATSVGPGPPTQIVV LDRKFP- <mark>SKARR</mark> RKPQYETFGVD <mark>DEKD</mark> PPCDPRMAIANYTRGAACKQKTKYRPTPYILRPWAYDPTTSVGPGPPTQIVV
Open Gap Penalty	TGYDPLTPLAPISALFSSFGDIAEIKNRTDPNTGRFLGVCSIRYKDSRMFRGGGPLLAAQAARRAYLECKKEQRIGVRRI TGFDPLTPIAAISALFSSFGDIGEINNRTDPMTGRFLGVCSIKYKDSRAFRGGISLSASQAVRRAYLECKKEQRIGTRRI QVSLDRDGVVSDRLVARIIGSOR-QQEP-PPLVME-E-KMKSE-EQDNLPPPTAPKGPS-RKPNM RVELDRNGVVSGRMVAKLITAQKAEFPSLEESRKESVGDNDNRLPIGDGAKKDNEQSKDNLPPSTAPKGPSGRSSLHPSL LIPEGPRATMMKPPAPSLIEETPILDQIKRDPYIFIAHCYVPVLSTTIPHLERRLKLFNWKSVRCDKTGYYIIFDNSRRG LAPDGPRA-VLKSPVPSRIEETPILQQIKRDPYIFIAHCYVPVLSTTVPHLERRLKLYDWKAVRCDKTGYYIIFENSRRG
Stop Compute Stopping computation – please wait	Performance: 812.0759 GCUPS







b

Bowtie2, correct and first correct

BWA-SW, correct and first correct

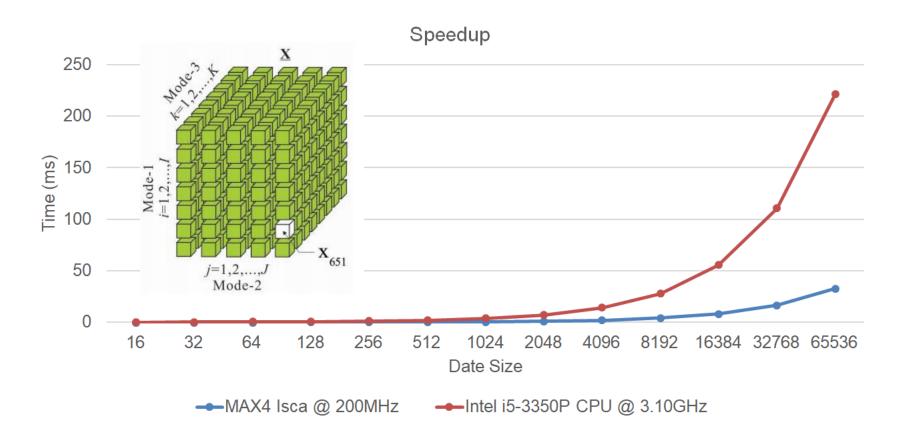
Bowtie2, mapped

Bowtie2, correct and first correct

▶ ■ BWA-SW, correct and first correct

Bowtie2, mapped

# Analysis of the Tensor Calculus Operations on DataFlow (PhD Thesis by Miloš Kotlar, on DataFlow-based Machine Learning)

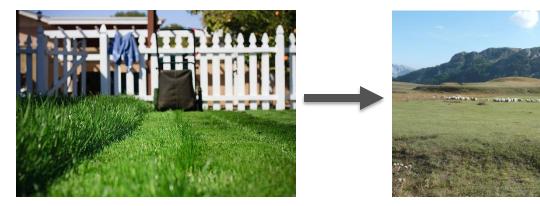


The speedup of **6.75x** achieved as early as for KiloData (Perceptron), with **10x** less on-chip transistors and the power savings of **4.6x** 

Conditions for the Y-Chart-Based "Kernelization" of Loops @ML

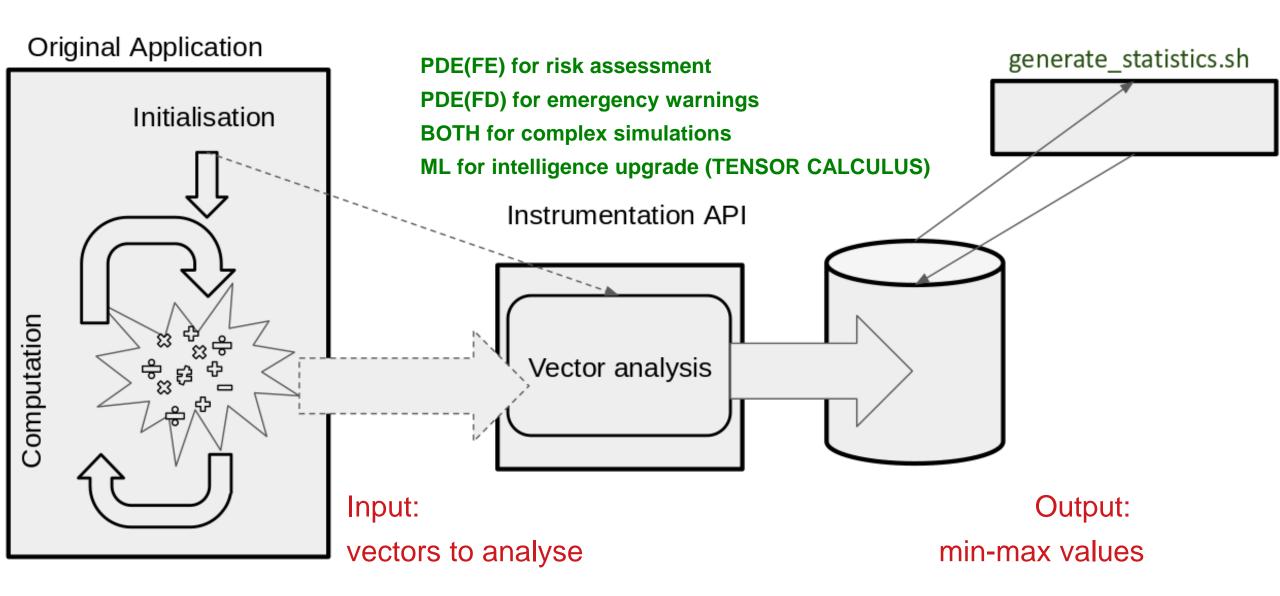
(PhD Thesis by Nenad Korolija, on the Mapping of Algorithms onto DataFlow)

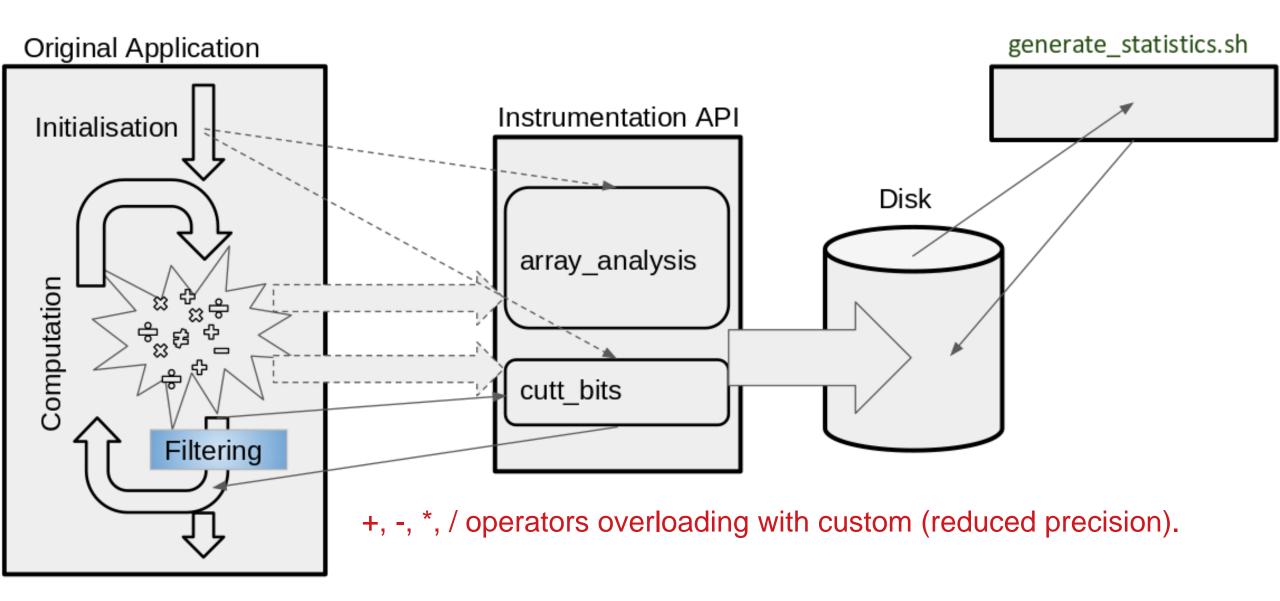
1.	BigData (RAM vs. STREAM)	O(n <sup>2</sup> )
2.	Code reusability (WORO vs. WORM)	+
3.	Overall application tolerance to latency	+
4.	Over 95% of run time in loops	++
5.	Reusability of the data in loops	++
6.	Potential for utilization of pipes	O(n)

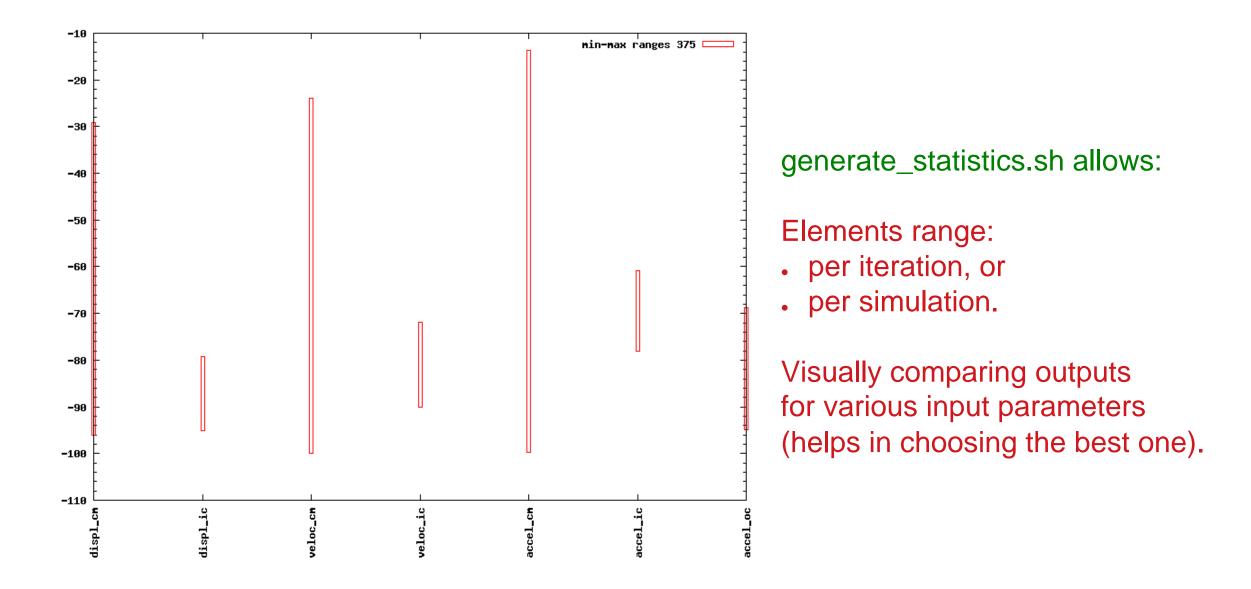


## **Essentials for speedup:** algorithmic modifications, pipeline utilization, data choreography, decision making on precision



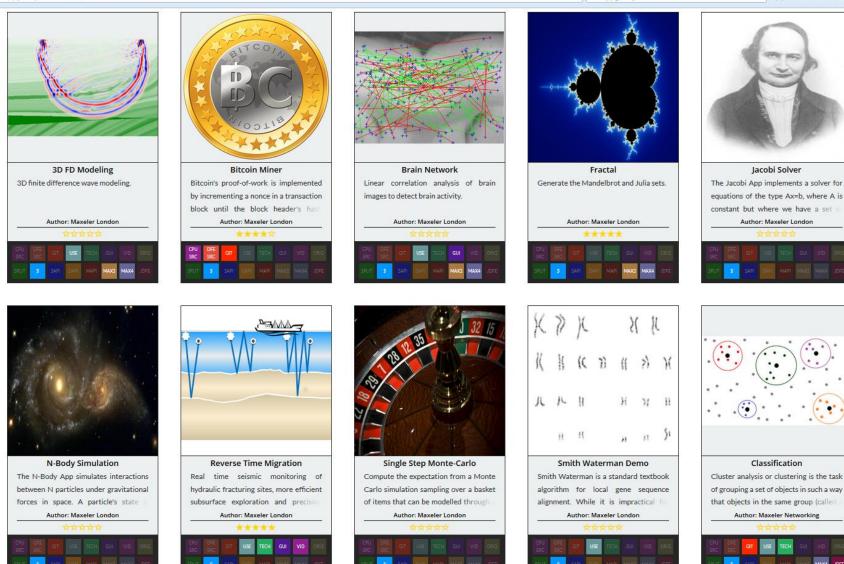






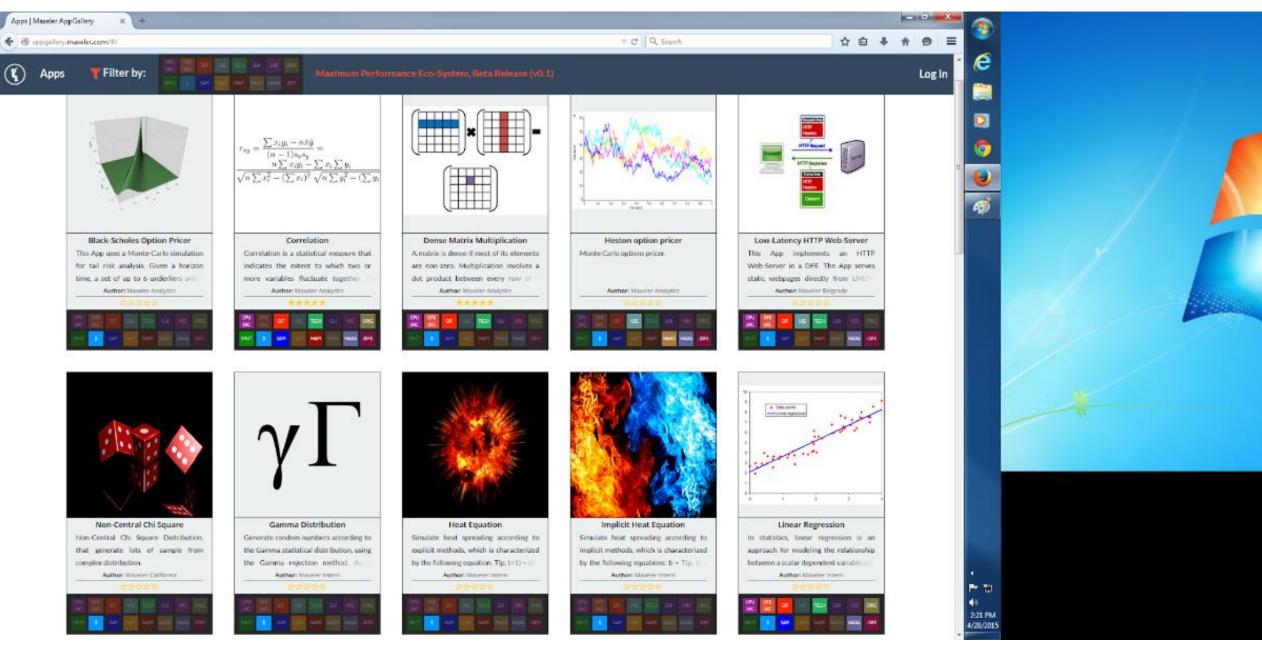
(appgallery.maxeler.com/#/

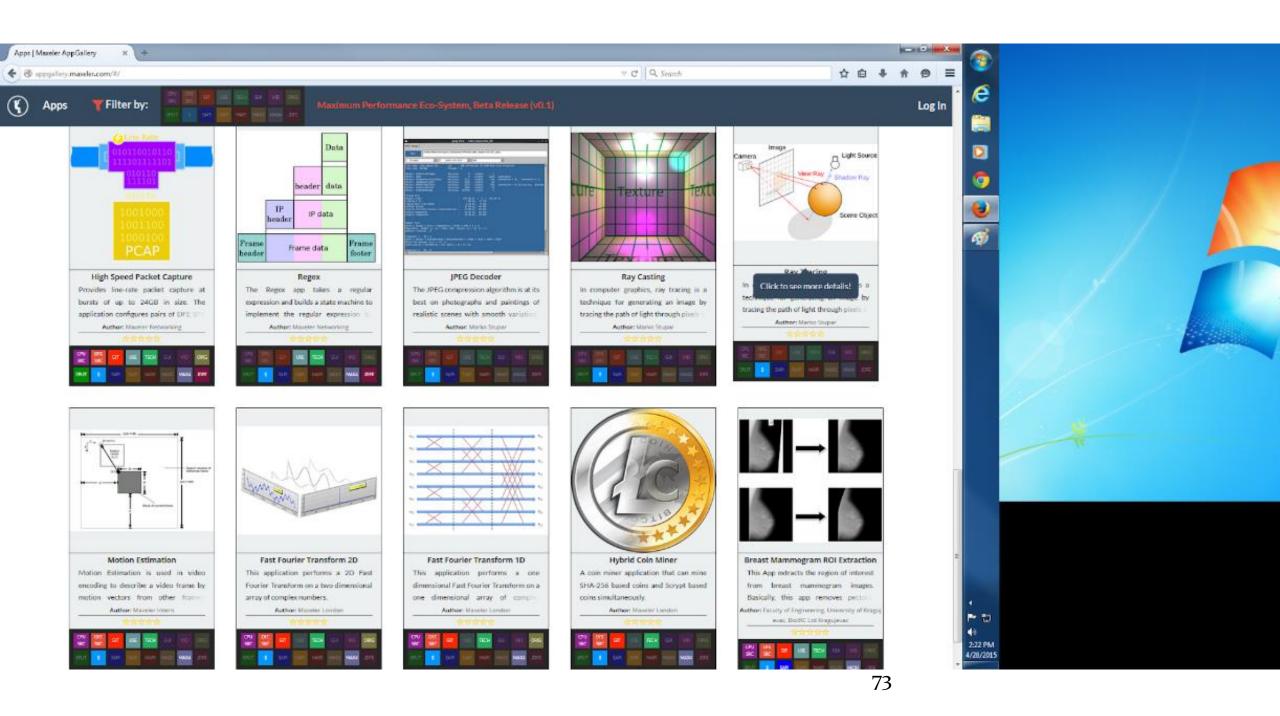
→ ☆ 自 🕹 ⋒ 😕 🗄

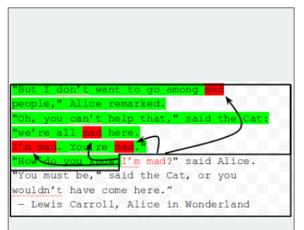


# appgallery.maxeler.com

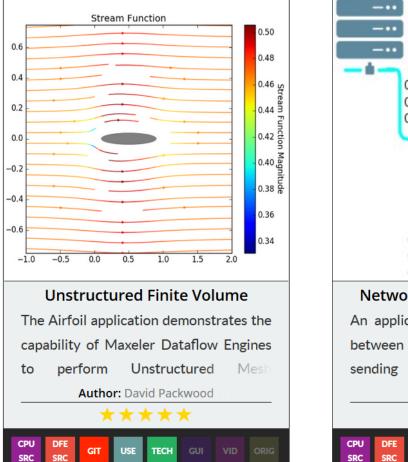
http://www.mi.sanu.ac.rs/~appgallery.maxeler/

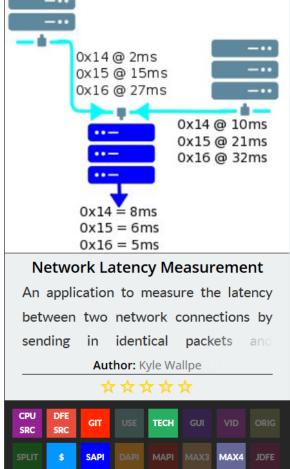












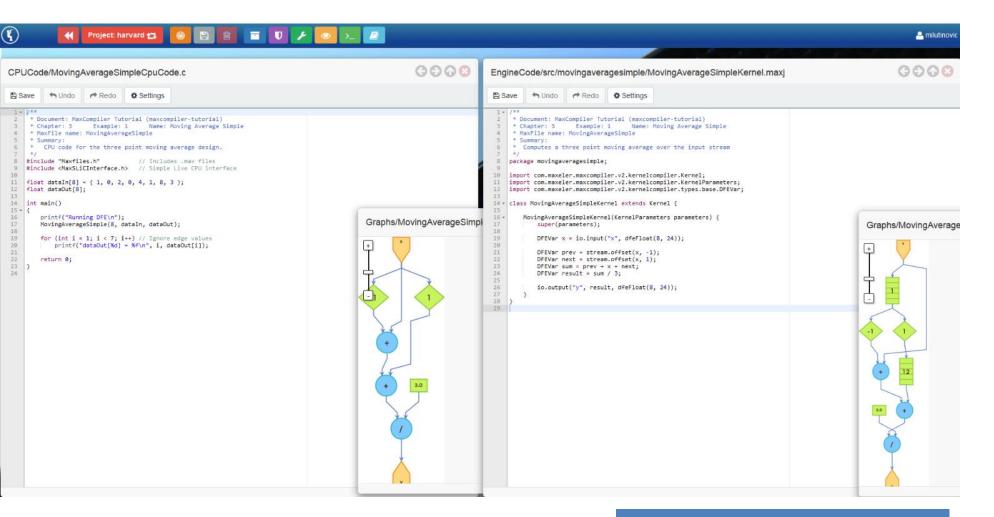
# appgallery.maxeler.com

DAPI MAPI MAX3

MAX4

http://www.mi.sanu.ac.rs/~appgallery.maxeler/

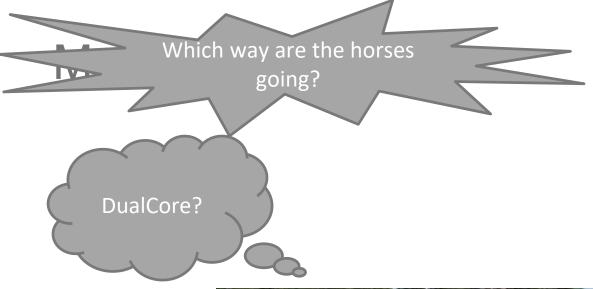




webide.maxeler.com https://maxeler.mi.sanu.ac.rs











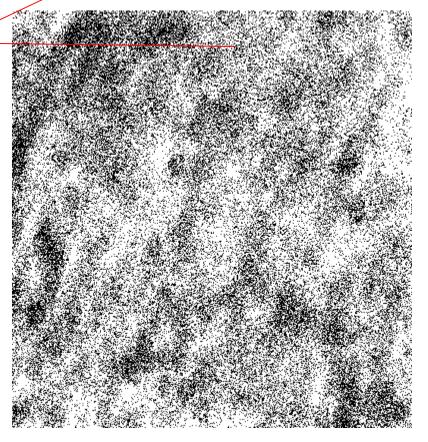
## ManyCore

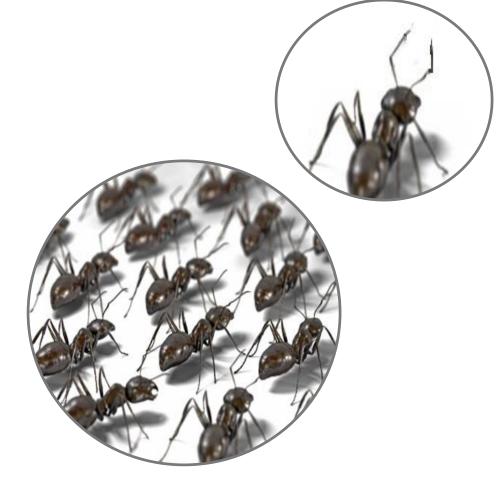
 Is it possible to use 2000 chicken instead of two horses?



• What is better, real and anecdotic?



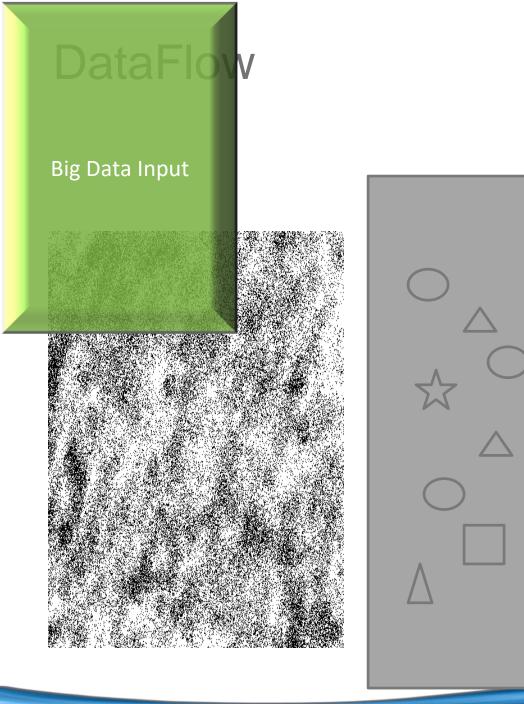




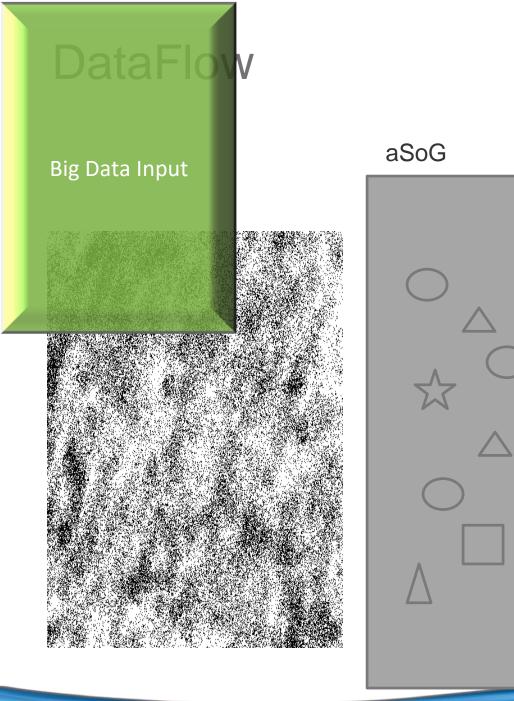
How about 2 000 000 ants?

81

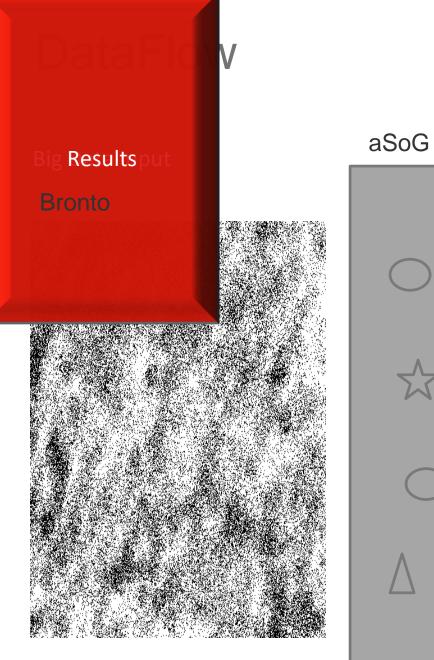














Marmelade

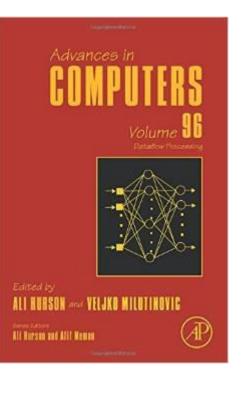


## An Edited Book Covering the Applications

- http://www.amazon.com/Dataflow-Processing-Volume-Advances-Computers/dp/0128021349
- http://www.elsevier.com/books/dataflow-processing/milutinovic/978-0-12-802134-7

## Indexed by: WoS (SCI)

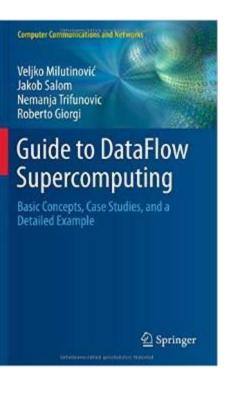
Contributions welcome for the follow-ups: Vol. 102 + Vol. 104 + etc...



# An Original Book Covering the Essence

http://www.amazon.com/Guide-DataFlow-Supercomputing-Concepts-Communications/dp/3319162284

http://www.springer.com/gp/book/9783319162287



The first source to use the term the Feynman Paradigm in contrast with the Von Neumann Paradigm



#### CLOUD // SOFTWARE AS A SERVICE



#### Google I/O: Hello Dataflow, Goodbye MapReduce

Google introduces Dataflow to handle streams and batches of big data, replacing MapReduce and challenging other public cloud services.



🔽 🔊 🖂

COMMENTS

COMMENT NOW

**4** 

Google I/O this year was overwhelmingly dominated by consumer technology, the end user interface, and extension of the Android Connect Directly universe into a new class of mobile devices. the computer you wear on your wrist.

At the same time, there were one or two

enterprise-scale data handling and cloud



Hadoop Jobs: 9 Ways To Get Hired

(Click image for larger view and slideshow.)

computing gems scattered among all the end user announcements.



#### Alibaba recently did the same!

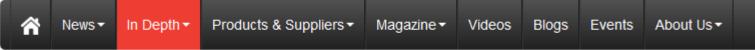


Share

15

🖤 Comment

### Intel says logic is faster than GPUs



#### Home Technology Article Alibaba recently claimed the same!

in Share 622

Intel's Programmable Systems Group takes its first step towards FPGA based system in package portfolio

34

8 Share 1

Speaking in 2012, Danny Biran – then Altera's senior VP for corporate strategy – said he saw a time when the company would be offering 'standard products' – devices featuring an FPGA, with different dice integrated in the package. "It's also possible these devices may integrate customer specific circuits if the business case is good enough," he noted.

🔰 Tweet

There was a lot going on behind the scenes then; already, Altera was talking with Intel about using its foundry service to build 'Generation 10' devices, eventually being acquired by Intel in 2015.

Now the first fruit of that work has appeared in the form of





share < 994</p>

Jordan Inkeles, Altera's director of product marketing for high end FPGAs

Stratix 10 MX. Designed to meet the needs of those developing high end communications systems, the device integrates 8 stacked memory dice alongside an FPGA die, providing users with a memory bandwidth of up to 1Tbyte/s. 28 June 2016 8

## QoL

Maxeler is one of the Top 10 HPC projects to impact QoL in the World :)

Scientific Computing [www.scientificcomputing.com/articles/2014/11]

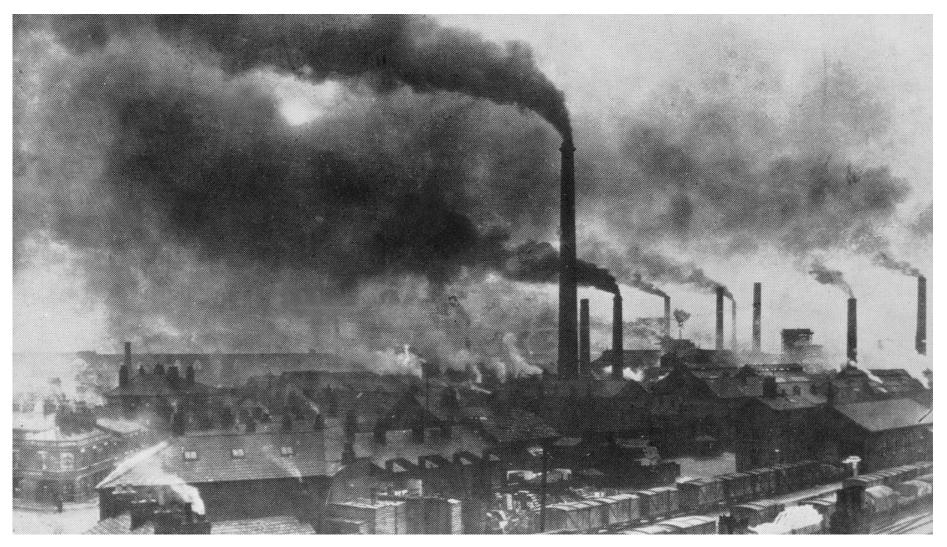
by

Don Johnson

of

Lawrence Livermore National Labs [editor@ScientificComputing.com]

# How About QoL?



# DataFlow







# Essence of the Paradigm:

For Big Data algorithms and for the same hardware price as before, achieving:

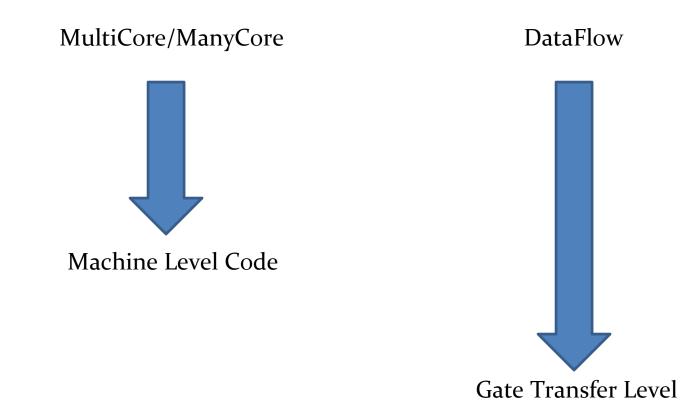
a) speed-up, 20-200

b) monthly electricity bills, reduced 20 timesc) size, 20 times smallerd) precision, X times better

The major issues of engineering are: design cost and design complexity. Remember, economy has its own rules: production count and market demand!

# Why is DataFlow so Much Faster?

• Factor: 20 to 200



# Why are Electricity Bills so Small?

• Factor: 20

MultiCore/ManyCore

DataFlow

# 

 $P = k f U^2$ 



Data Processing

## the Cubic Foot so Small?

MultiCore/ManyCore

DataFlow

# Why is the Precision Better?

• Factor: X

	Μ								•											
Ν	Bits	18	20	22	24	26	28	30	32	34	36	38	40	42	44	46	48	50	52	54
	18	1	1	1	1	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3
	20	1	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	3
	22	1	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	3
	24	1	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	3
	26	2	2	2	2	4	4	4	4	4	4	4	4	4	6	6	6	6	6	6
	28	2	2	2	2	4	4	4	4	4	4	4	4	4	6	6	6	6	6	6
	30	2	2	2	2	4	4	4	4	4	4	4	4	4	6	6	6	6	6	6
	32	2	2	2	2	4	4	4	4	4	4	4	4	4	6	6	6	6	6	6
	34	2	2	2	2	4	4	4	4	4	4	4	4	4	6	6	6	6	6	6
	36	2	3	3	3	4	4	4	4	4	5	5	5	5	6	6	6	6	6	7
	38	2	3	3	3	4	4	4	4	4	5	5	5	5	6	6	6	6	6	7
- ↓	· 40	2	3	3	3	4	4	4	4	4	5	5	5	5	6	6	6	6	6	7
	42	2	3	3	3	4	4	4	4	4	5	5	5	5	6	6	6	6	6	7
	44	3	3	3	3	6	6	6	6	6	6	6	6	6	9	9	9	9	9	9
	46	3	3	3	3	6	6	6	6	6	6	6	6	6	9	9	9	9	9	9
	48	3	3	3	3	6	6	6	6	6	6	6	6	6	9	9	9	9	9	9
	50	3	3	3	3	6	6	6	6	6	6	6	6	6	9	9	9	9	9	9
	52	3	3	3	3	6	6	6	6	6	6	6	6	6	9	9	9	9	9	9
	54	3	4	4	4	6	6	6	6	6	7	7	7	7	9	9	9	9	9	10

Endorsed by Jerome Friedman

Special thanks to: Jerome Friedman, Dan Shechtman, Tim Hunt, and Sheldon Glashow





US 20180189063A1

(19) United States

#### (12) Patent Application Publication FLEMING et al. (10) Pub. No.: US 2018/0189063 A1 (43) Pub. Date: Jul. 5, 2018

- (54) PROCESSORS, METHODS, AND SYSTEMS WITH A CONFIGURABLE SPATIAL ACCELERATOR
- (52) U.S. Cl. CPC ........ G06F 9/3016 (2013.01); G06F 13/4221
- Intel Corporation, Santa Clara, CA (57)

#### ABSTRACT

(2013.01)

Systems, methods, and apparatuses relating to a configurable spatial accelerator are described. In one embodiment, a processor includes a core with a decoder to decode an instruction into a decoded instruction and an execution unit to execute the decoded instruction to perform a first operation; a plurality of processing elements; and an interconnect network between the plurality of processing elements to receive an input of a dataflow graph comprising a plurality of nodes, wherein the dataflow graph is to be overlaid into the interconnect network and the plurality of processing elements with each node represented as a dataflow operator in the plurality of processing elements, and the plurality of processing elements are to perform a second operation by a respective, incoming operand set arriving at each of the dataflow operators of the plurality of processing elements.

- (71) Applicant: Intel Corporation, Santa Clara, CA (US)
- (72) Inventors: KERMIN FLEMING, Hudson, MA (US); KENT D. GLOSSOP, Merrimack, NH (US); SIMON C. STEELY, Jr., Hudson, NH (US)
- (21) Appl. No.: 15/396,395
- (22) Filed: Dec. 30, 2016

#### Publication Classification

- (51) Int. Cl.
  - G06F 9/30 (2006.01) G06F 13/42 (2006.01)

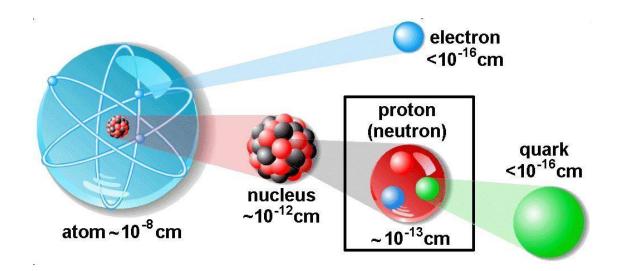
99

#### BQCD on a Maxeler Dataflow Computer

# MAXIMUM PERFORMANCE COMPUTING



#### **BQCD on a Dataflow Computer**



Quantum Chromodynamics (QCD): models interactions of subatomic particles

- ◆ Lattice QCD (LQCD): its discretisation, suitable for numerical computation
- Berlin QCD (BQCD): most popular implementation of the LQCD algorithm
- Conjugate Gradient (CG): Majority of the compute time (benchmark: 68%)
  - CG iteratively solves linear algebra problem of form Mx = b
  - Operator **M** contains Wilson-dslash and Clover operators

• PROJECT TARGET 40x speedup of CG part of BQCD, followed by speedup of the entire application by 20x comparing same size boxes Dataflow vs BlueGene/Q



#### **Maxeler QCD - Deployment**

# MaxelerQCD solutionisdeployedatJülichSupercomputing Center, running on a MaxelerDataflow system.

2 racks of Jülich BlueGene/Q	On-premise Maxeler Dataflow	Factor	
machine	system: scale to 1PF equivalent		
6.75 m <sup>3</sup>	0.87 m <sup>3</sup>	7.76	
1576.60 s	689 s	2.29	
169.6 kWh	4.42 kWh	38.4	
10,642.05 m <sup>3</sup> s	599.43 m <sup>3</sup> s	17.8	
	machine         6.75 m³         1576.60 s         169.6 kWh	machinesystem: scale to 1PF equivalent6.75 m³0.87 m³1576.60 s689 s169.6 kWh4.42 kWh	

10 2





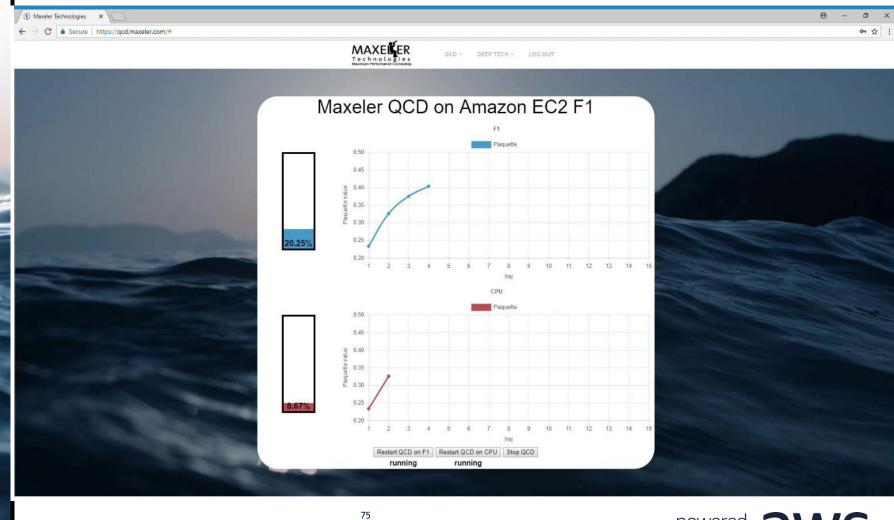








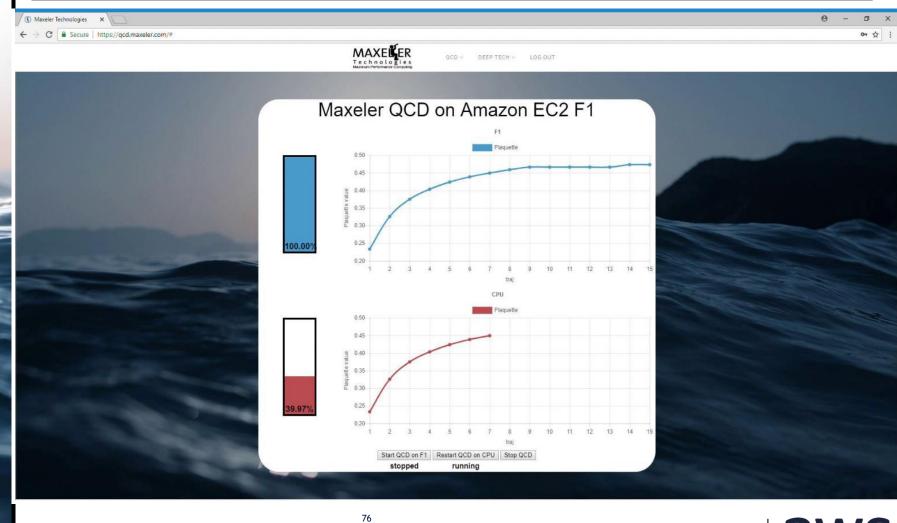
#### QCD Demo







## QCD Demo





105

#### MAXELLER Technologies Maximum Performance Computing

#### miniMAX5 Edge of IoT Platform





Dimensions	5.7 in (144mm) Wide x 5.7 in Deep (144mm) x 2.2 in High (57mm). excluding power supply								
Form factor	Desktop enclosure, fanless design. Wall or rail mount options available 34 oz (950g), excluding power supply								
Weight									
Power Supply	Separate wall plug unit, providing 60W of USB-PD power from 100-240V, 50-60Hz mains								
	Ethernet	1GbE or 10GbE (copper or fibre)	SFP+ Cage						
Input and Output (Standard Ports)	USB-C	Power input over USB-PD (Min 15V 3A supply required) USB-3 SuperSpeed II I/O on same connector supporting DisplayPort Alternate mode							
	Management LAN	1GbE	RJ45						
Input and Output	USB	Dual USB-3 Type A ports							
(Optional Ports)	Video Output	HDMI Type A							
	CPU	AMD 3rd Generation R- or G-Series - choose from - Quad Core Merlin Falcon RX-416GD @16GHz - Dual Core Brown Falcon GX-217I @17GHz	Other SBC options available on request						
Controlflow	Memory	2x 4GB DDR4-2133 SODIMM, total 8GB	Higher or lower capacities as required						
Engine	Storage	64GB Solid State Memory							
	Operating System	Linux - CentOS 7	Other OS option: available on request						
Dataflow	FPGA	Xilinx Kintex Ultrascale Plus series - choose from - KU5P (217K LUTS, 544 BRAMs, 1,824 DSPs) - KU3P (163K LUTS, 408 BRAMs, 1,368 DSPs) - KU11P (296K LUTS, 680 BRAMs, 2,928 DSPs)	KU5P fitted as standard						
Engine	Memory	1x 16GB DDR4-2400 SODIMM	Or 8GB or 32GB						

subEdge: microMAX

© Maxeler Technologies

www.maxeler.com

Purdue, IU. MIT, Harvard, Boston, NEU, Dartmouth, U of Massachusetts at Amherst, USC, UCLA, Columbia, NYU, Princeton, NJIT, CMU, Temple U, UIUC, Michigan, Wisconsin, Minnesota, FAU, FIU, Miami, Central Florida, U of Alabama, U of Kentucky, GeorgiaTech, Ohio State, Imperial, King's, Manchester, Huddersfield, Cambridge, Oxford, Dublin, Cork, Cardiff, Edinburgh, EPFL, ETH, TUWIEN, UNIWIE, Graz, Linz, Karlsruhe, Stuttgart, Bonn, Frankfurt, Heidelberg, Aachen, Darmstadt, Dortmund, KTH, Uppsala, Karlskrona, Karlstad, Napoli, Salerno, Siena, Pisa, Barcelona, Madric, Valencia, Oviedo, Ankara, Bogazici, Koc, Istanbul, Technion, Haifa, BerSheba, Eilat, Belgrade, Podgorica, Koper, Ljubljana, Maribor, Nova Gorica, etc, etc. Also at the World Bank in Washington DC, IMF, the Telenor Bank of Norway, the Raiffeisen Bank of Austria, Brookhaven National Laboratory, Lawrence Livermore National Laboratory, IBM TJ Watson, HP Encore Labs, Intel Oregon, Qualcomm VP, NCR, RCA, Fairchild,

Honeywell, Yahoo NY, Google CA, Microsoft, Finsoft, ABB Zurich, Oracle Zurich, and many other industrial labs, as well as at Tsinghua University, Shandong, NIS of Singapore, NTU of Singapore, Tokyo, Sendai, Seoul, Pusan, Sydney University of Technology, University of Sydney, Hobart, Auckland Toronto Montreal Durange Monterey Tech Cuerney and UNAM





#### Design of Systolic Arrays: An SIMD MultiMicroprocessor for DARPA

#### GaAs Systolic Array Based on 4096 Node Processor Elements

Adaptive signal processing is of crucial importance for advanced radar and communications Systems. In order to achieve real time throughput and latencies, one is forced to use advanced semiconductor technologies (e.g., gallium arsenide, or similar) and advanced parallel architectures (e.g., systolic arrays, or similar).

The systolic array described here was designed to support two important applications : (a) adaptive antenna array beamforming, and (b) adaptive Doppler spectral filtering. In both cases, in theory, the system output is calculated as the product of the signal vector  $\mathbf{x}$  (complex N-dimensional vector) and the weight vector  $\mathbf{w}$  (optimal N-dimensional vector).

Complex vector x is obtained by multiplying N input samples with the corresponding window weighting function consisting of N discrete values. Optimal vector  $\mathbf{w}$  is obtained as:

$$w = R^{-1}s^* = M^{-1}v^*.$$

# Symbol **R** refers to the N-by-N inverse convariance matrix of the signal with the (i,j) -th component defined as:

and symbol refers to N-dimensional vector which defines the antenna direction (in the case of adaptive antena beamforming) or Doppler peak (in the case of adaptive Doppler spectral filtering).

## Symbols M and V represent scaled values of R and S, respectively. In practice, the scaled values M and V may be easier to obtain, and consequently the remaining explanation is adjusted.

The core of the processing algorithm is the inversion of a N-by-N matrix in real time. This problem can be solved in a number of alternative ways which are computationally less complex. The one chosen here includes the operations explained in Figure Y1a.

Positive semi definite matrix **M** can be defined as:  $\mathbf{M} = \mathbf{U} \mathbf{D} \mathbf{U}^{\mathsf{T}}$ 

Matrices U and **D** are defined using the formula:

which is recursively updated using the formula:

$$U_{K}D_{K}U_{K}^{T} = U_{K-1}D_{K-1}U_{K-1}^{T} + x_{K}bx_{K}^{T}.$$

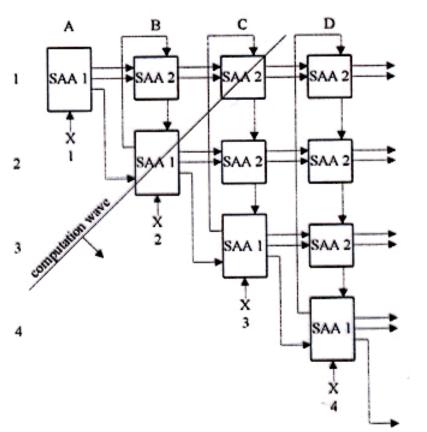
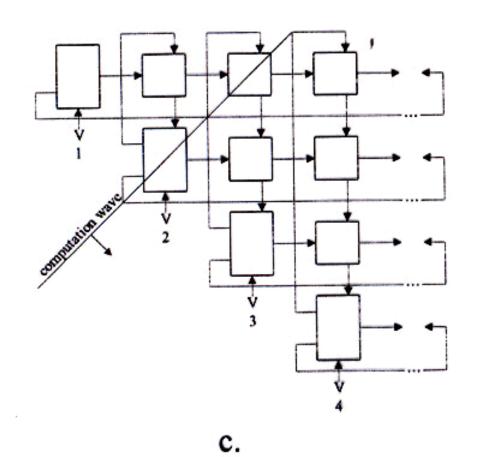
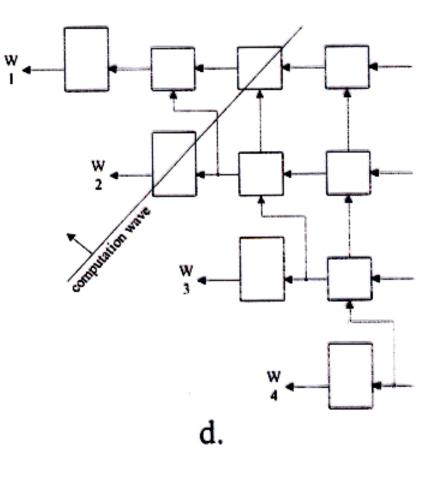


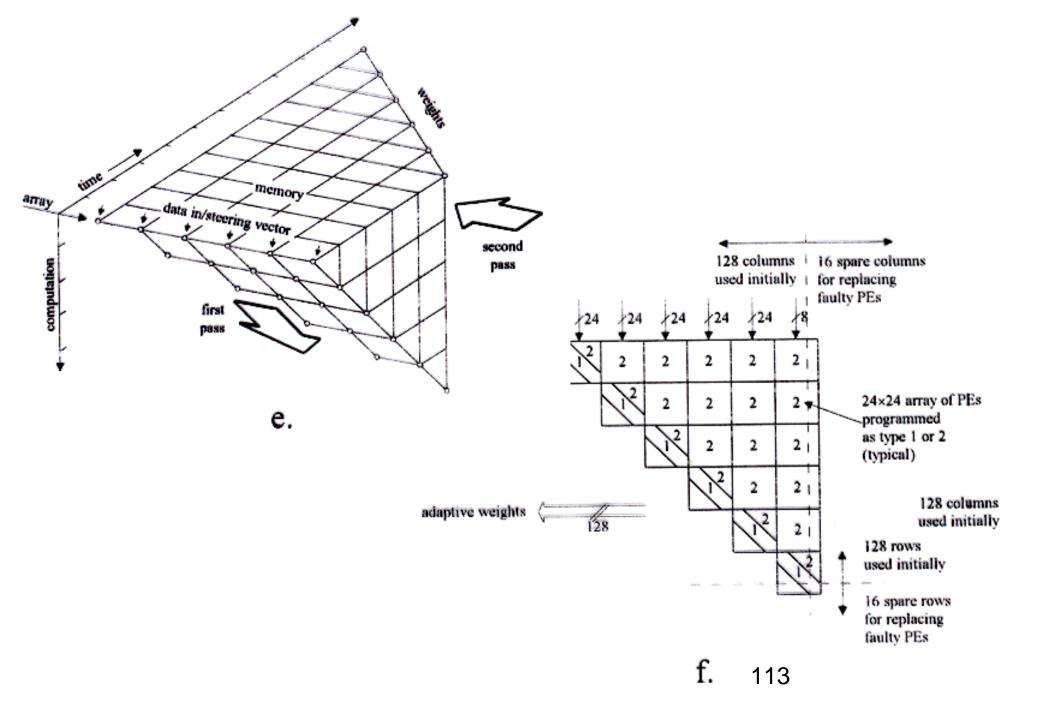
Figure Y1: Basic Operational Structure

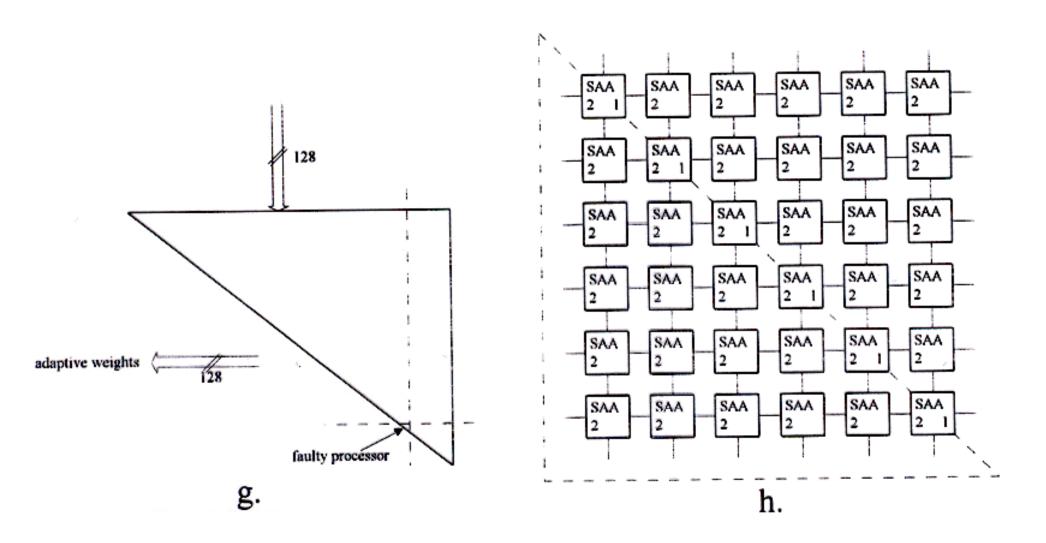
Legend:

- b. SAA1 – Cells involved in root covariance update, and the first step of back substitution;
- SAA2 Cells involved in root covariance update, and in both steps of back substitution;
- U Lower triangular matrix with unit diagonal elements
- D Diagonal matrix with positive or zero diagonal elements
- b A scalar initially set to 1
- K Iteration count.









#### Veljko Milutinović

#### VLSI for SuperComputing: From Applications and Algorithms to Masks and Chips

#1?

#1? Qualcomm

- #1? Qualcomm
- **#2**?

- #1? Qualcomm
- #2? Intel

- #1? Qualcomm
- #2? Intel
- **#**3?

- #1? Qualcomm
- #2? Intel
- #3? TSM (Taiwan Semiconductor Manufacturing)

- #1? Qualcomm
- #2? Intel
- #3? TSM
- The Qualcomm VP/TD started from this course (first as my PhD student and later as the course TA)

- #1? Qualcomm
- #2? Intel
- #3? TSM
- The Qualcomm VP/TD started from this course (first as my PhD student and later as the course TA)
- An Intel VP/TD took the IR4RVL in-a-nut-shell (at my past IEEE/ACM tutorial)

- #1? Qualcomm
- #2? Intel
- #3? TSM
- The Qualcomm VP/TD started from this course (first as my PhD student and later as the course TA)
- An Intel VP/TD took the IR4RVL in-a-nut-shell (at my past IEEE/ACM-HICSS conference tutorial)
- Maybe,

a Next Gen VP/TD comes from ETF 🔊

- #1? Qualcomm
- #2? Intel
- #3? TSM
- The Qualcomm VP/TD started from this course (first as my PhD student and later as the course TA)
- An Intel VP/TD took the IR4RVL in-a-nut-shell (at my past IEEE/ACM-HICSS conference tutorial)
- Maybe, a next TSM VP/TD comes from ETF or MF

- #1? Qualcomm
- #2? Intel
- #3? TSM
- The Qualcomm VP/TD started from this course (first as my PhD student and later as the course TA)
- An Intel VP/TD took the IR4RVL in-a-nut-shell (at my past IEEE/ACM-HICSS conference tutorial)
- Maybe,

a next TSM VP/TD comes from ETF (3) or MF (3) or FFH (3)

- #1? Qualcomm
- #2? Intel
- #3? TSM
- The Qualcomm VP/TD started from this course (first as my PhD student and later as the course TA)
- An Intel VP/TD took the IR4RVL in-a-nut-shell (at my past IEEE/ACM-HICSS conference tutorial)
- Maybe,

a next TSM VP/TD comes from ETF or MF (3) or FFH (3) or FON

## Who works here?



The Holistic Foundry (R&DFab) in VLSI for SuperComputing

 Phase#1: From Applications to Algorithms Phase#2: From Algorithms to Masks
 Phase#3: From Masks to Chips
 Phase#4: From Chips to Applications The Holistic Foundry (R&DFab) in VLSI for SuperComputing

 Phase#1: From Applications to Algorithms Phase#2: From Algorithms to Masks Phase#3: From Masks to Chips

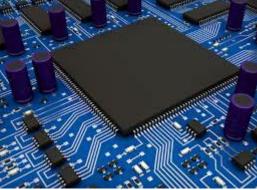
 Verification is crucial in each one of these phases, and related teaching is done in coop with ELSYS or HDLDH! The Holistic Foundry (R&DFab) in VLSI for SuperComputing

- Phase#1: From Applications to Algorithms Phase#2: From Algorithms to Masks Phase#3: From Masks to Chips Phase#4: From Chips to Applications
- Verification is crucial in each one of these phases, and related teaching is done in coop with ELSYS or HDLDH!
- Management issues of importance for an R&D Fab are covered in the accompanying course: IR4USP (including 12 related homework assignments)!

## Contents: From Algorithms to Masks

Part#1: VLSI for ControlFlow SuperComputing
Part#2: VLSI for DataFlow SuperComputing
Part#3: VLSI for WirelessFlow SuperComputing
Part#4: VLSI for EnergyFlow SuperComputing

(QMOC)



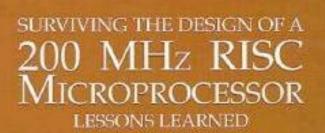
#### VLSI for ControlFlow SuperComputing

ManyCore Systems:

- Enabler Technology: VHDL vs Verilog (0.5 weeks)
- Design and Programming of a 200MHz RISC Microprocessor (2.5 weeks) + HW#1

MultiCore Systems:

- Enabler Technology: Verification by Elsys (2 weeks) + Lab#1
- Design of MicroProcessor and MultiMicroProcessor Systems by Wiley (1 week)

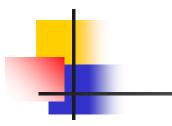




Veljko Milutinović Foreword by Michael Flynn

Constant Source

THE OTHER DESIGNATION OF THE OTHER OF THE OT



SURVIVING THE DESIGN OF MICROPROCESSOR AND MULTIMICROPROCESSOR SYSTEMS LESSONS LEARNED



Veljko Milutinović Foreword by Michael J. Flynn

www.

Wiley Series on Parallel and Distributed Computing Albert Y. Zomaya, Series Editor

#### VLSI for DataFlow SuperComputing

FineGrain DataFlow:

- Enabler Technology: Altera vs Xilinx (0.5 weeks)
- Design and Programming of the 200MHz Maxeler Machine (3.5 weeks) + HW#2

SystolicArray DataFlow:

- Enabler Technology: Systolic Array Architectures (0.5 weeks)
- Design of DARPA Systolic Architectures (0.5 weeks) + Lab#2

#### On Flexible DataFlow:

#### Advances in Computer Architecture (North Holland) by Veljko M. Milutinovic with a contribution from John Hennessy



#### On Fixed DataFlow:

 High-Level Language Computer Architecture (Elsevier Computer Science Press) by Veljko M. Milutinovic with a contribution from Michael Flynn



#### VLSI for WirelessFlow SuperComputing

WSNs: Part#1

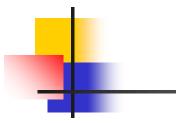
- Hardware (0.25 weeks)
- Software (0.25 weeks)

WSNs: Part#2

- Systems (SUN+Slimmer) (0.25 weeks)
- Applications (UbiComputing@WSN+DataMining@WSN) (0.25 weeks)

The stress is on integration of WSNs, IoT, Ethernet, and Internet:

Z. Babovic et al., "Web Performance Evaluation of Internet of Things Applications," IEEE Access 2016.



#### Click to LOOK INSIDE!



ULIANA GARREDYSKA - SREJAN KIKO VELKO MEUTINOVIC - NWA STOJMENOVIC ROVAN TROBEC CERT

#### Application and Multidisciplinary Aspects of Wireless Sensor Networks

Concepts, Integration, and Case Studies

Springer

COMPUTER COMMUNICATIONS AND NETWORKS



Goran Rakocevic · Tijana Djukic Nenad Filipovic · Veljko Milutinović Editors

Computational Medicine in Data Mining and Modeling

#### VLSI for Quantum, Molecular, Optical, and Chemical SuperComputing

**Basics:** 

- Hardware (N weeks)
- Software (N weeks)

Advances:

- Systems (N weeks)
- Applications (N weeks)

# Quantum Computing

#### A quantum computer uses qubits to run multidimensional quantum algorithms.

- Their processing power increases exponentially as qubits are added.
- A classical processor uses bits to operate various programs.

Their power increases linearly as more bits are added.

#### Potentials:

 Quantum computing is 158 million times faster than the most sophisticated supercomputer we have in the world today.

 It is a device that could do in four minutes what it would take a traditional supercomputer 10,000 years to accomplish.

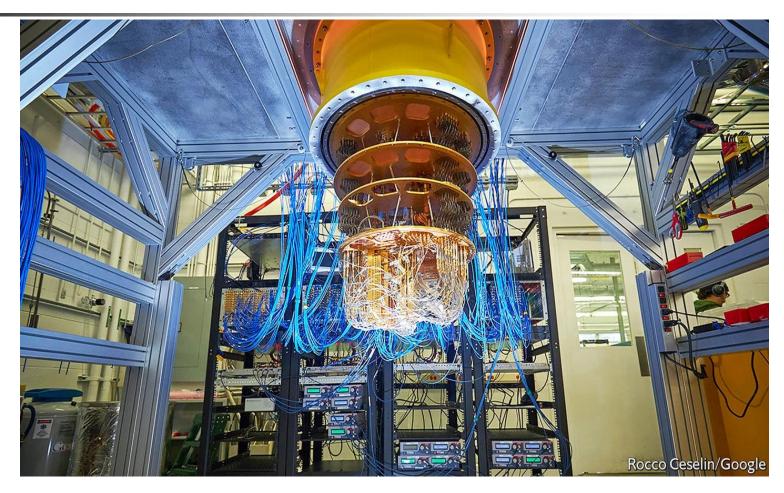




The First Eight Companies in Quantum Computing:

- Atom Computing
- Xanadu
- IBM
- ColdQuanta
- Zapata Computing
- Azure Quantum
- D-Wave
- Strangeworks

### Beginnings:



#### The Best Quantum Computing Stocks:

- IBM (NYSE: IBM)
- Alphabet (Nasdaq: GOOG, Nasdaq: GOOGL)
- Intel (Nasdaq: INTC)
- Microsoft (Nasdaq: MSFT)
- Amazon (Nasdaq: AMZN)
- Quantum Computing (OTC: QUBT)

### Libraries:

 TensorFlow Quantum (TFQ) is a <u>quantum machine learning</u> library for rapid prototyping of hybrid quantum-classical ML models.

Interfacing for control-flow.

#### **Tensor Flow Quantum:**

- TensorFlow Quantum focuses on *quantum data* and building *hybrid quantum-classical models*.
- It integrates quantum computing algorithms and logic designed in <u>Cirq</u>, and provides quantum computing primitives compatible with existing TensorFlow APIs, along with high-performance quantum circuit simulators.

### A Programming Example:

# A hybrid quantum-classical model. model = tf.keras.Sequential

([

- # Quantum circuit data comes in inside of tensors. tf.keras.Input(shape=(), dtype=tf.dtypes.string),
- # Parametrized Quantum Circuit (PQC) provides output
- # data from the input circuits run on a quantum computer. tfq.layers.PQC(my\_circuit, [cirq.Z(q1), cirq.X(q0)]),
- # Output data from quantum computer passed through model.

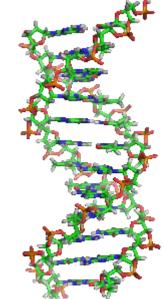
tf.keras.layers.Dense(50)

])

# **Molecular Computing**

 Molecular computing is a branch of computing that uses DNA, biochemistry, and molecular biology hardware, instead of traditional silicon-based computer technologies, to achieve dense computing!

Applications for dense computing are many!



#### Who invented DNA computing?

#### Leonard Adleman,

professor of computer science and molecular biology at the University of Southern California, USA, who pioneered the field when he built the first DNA based computer.

#### L. M. Adleman, Science 266, 1021–102; 1994.

# Essence:

With DNA,

the way the molecules can be triggered to bind with each other can be used to create a circuit of logic gates in test tubes.

Compatible with nature based computing!

### Techniques:

Many gates can be combined in a circuit:

 Each output DNA will bind to the next logic gate until some predictable terminal output strand is liberated,

to produce an intermediate or final result.

# N.B.:

- In *DNA computing*, information is represented using the four-character genetic alphabet: A [adenine], G [guanine], C [cytosine], and T [thymine].
- The core advantage of molecular computing is its potential to pack vastly more circuitry onto a microchip than silicon will ever be capable of—and to do it cheaply.

### Miniaturization:

- Molecules are only a few nanometers in size.
- Making it possible to manufacture chips that contain billions, even trillions, of switches and components.

#### Can DNA be programmed?

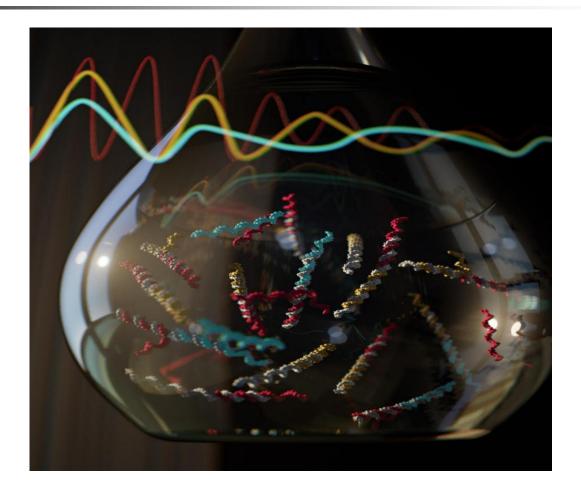
 Researchers at The University of Texas at Austin have programmed DNA molecules to follow specific instructions to create sophisticated molecular machines that could be capable of communication, signal processing, problem-solving, decision-making, and control of motion in living cells.

Libraries under construction at MIT.

#### How does DNA computing work?

- In one method, called DNA strand displacement, the input of DNA that binds to a DNA logic gate displaces a strand of DNA that serves as the output.
- Many gates can be combined in a circuit: Each output DNA will bind to the next logic gate until some predictable terminal output strand is liberated.

### An Experiment:



### Flexibility:

With the flexible molecular algorithms on the rise, one might be able to assemble a complex entity on the nanoscale with the reprogrammable tile set.
Flexibility is the issue for a great number of applicational.

for a great number of applications!

#### Do DNA computers exist?

- The DNA Computing Technology is in the research phase.
- DNA computers can't be found at your local electronics store yet.
- The technology is still in development; it didn't even exist as a concept a decade ago.



# **Optical Computing**

 Optical computing or photonic computing uses light waves produced by lasers or incoherent sources for data processing, data storage or data communication for computing.
 For decades, photons have shown promise to enable a higher <u>bandwidth</u>

than the <u>electrons</u> used in conventional computers.

#### Essence:

"Computation-by-propagation, where the computation takes place as the wave propagates through a medium, can perform computation at the speed of light!"
Speed of light = 300.000km/sec

### Applications:

 Application-specific devices, such as <u>synthetic-aperture radar</u> (SAR) and <u>optical correlators</u>, have been designed to use the principles of optical computing.

 Correlators can be used, for example, to detect and track objects, and to classify serial time-domain optical data.

### Techniques:

 Techniques have been developed that can perform continuous Fourier Transform optically, by utilizing the natural Fourier transforming property of lenses.

 The input is encoded using a <u>liquid crystal spatial light modulator</u>. Artificial Neural Network with Optical Components:

 Early optical neural networks used a photorefractive Volume Hologram to interconnect arrays of input neurons to arrays of output.

 With synaptic weights in proportion to the multiplexed hologram's strength.

#### **Optical Neural Networks:**

 Some artificial neural networks that have been implemented as optical neural networks include the <u>Hopfield neural network</u>.

#### Also the Kohonen map with liquid crystal spatial light modulators.

#### Do Photonic Chips Exist?

Now scientists have developed a deep neural network on a photonic microchip that can classify images in less than a nanosecond, roughly the same amount of time as a single tick of the kind of clocks found in state-of-the-art electronics.

A lot more effective compared to microprocessor based design!

# An Application:

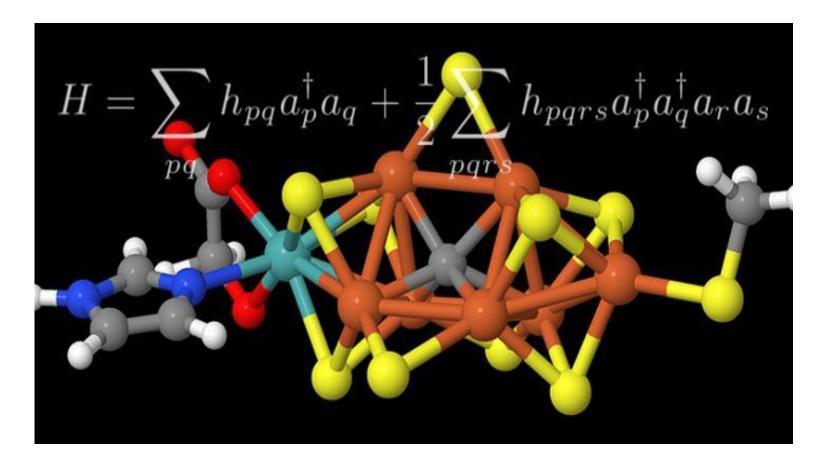


# Chemical Computing

 Computing with real molecules like programming electronic devices, but using principles taken from chemistry and appropriate chemical processes.

Effective for simulations in chemistry and physical chemistry!

### An Example:



# A chemical computer, also called a reaction-diffusion computer:

- Belousov–Zhabotinsky (BZ) computer, or gooware computer,
  - is an <u>unconventional computer</u> based on a semi-solid chemical "soup" where data are represented by varying concentrations of chemicals.
- The computations are performed by naturally occurring <u>chemical reactions</u>.

# Origins:

 Originally chemical reactions were seen as a simple move towards a stable equilibrium which was not very promising for computation.
 This was changed by a discovery made by <u>Boris Belousov</u>, a <u>Soviet</u> scientist, in the 1950s.

#### Belousov:

 Belousov created a <u>chemical reaction</u> between different salts and acids that swing back and forth between being yellow and clear.

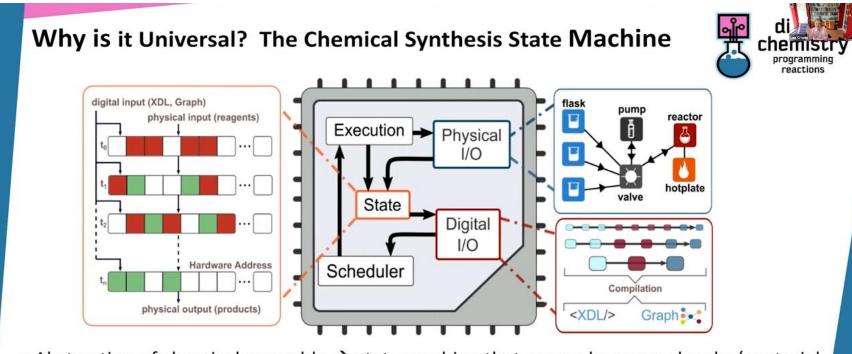
 This is because the concentration of the different components changes up and down in a cyclic way.



 Andrew Adamatzky at the <u>University of the West of England</u> has demonstrated simple logic gates using <u>reaction-diffusion</u> processes.

An important step towards programmability.

### An Implementation:



- Abstraction of chemical assembly  $\rightarrow$  state machine that can make any molecule / material
- Inputs are digital and physical → Outputs are physical

*Chemputation* is the process of running XDL code reliably on <u>any</u> compatible hardware

c.f. Computation - running programs on a digital computer

#### **European Projects**

#### ESF:

- RoMoL: Riding on Moore's Law (0 weeks)
- HiPeac: Parallel Programming Models (0 weeks)

FP7/H20:

- FP7: ProSense (0 weeks)
- FP7: BalCon (0 weeks)

#### **USA Projects**

Nature Based Construction and Computing:

- Purdue (X weeks)
- Indiana University (X weeks)
- Nature Based Media and Computing:
  - MIT (X weeks)
  - Harvard (X weeks)

#### **Example Algorithms for Practical Implementations**

#### Engineering:

- Computer Engineering (M weeks)
- Financial Engineering (M weeks)

#### Science:

- Physical Chemistry (M weeks)
- Computer Science (M weeks)

## SOME PREVIOUS OFFERINGS OF THIS TECH COURSE

- Purdue, Indiana;
- MIT, Harvard;
- Imperial, Kings;
- ETH, EPFL;
- UNIWIE, TUWIEN;
- Siena, Salerno, Barcelona, Madrid;
- Ljubljana, Koper, Zagreb, Rijeka, Podgorica, UBG;
- Technion, Jerusalem
- Bogazici, Koc;
- Tsinghua, Shandong.

### BOTTOM LINE: BRINGING ADVANCED INDUSTRIAL EXPERIENCE INTO THE CLASSROOM

 $\leftarrow \rightarrow C$   $\triangleq$  https://www.google.rs

Google DARPA's first 200MHz GaAs Microprocessor - a decade before Intel



Притисните Enter да бисте претражили.

#### BOTTOM LINE: BRINGING ADVANCED INDUSTRIAL EXPERIENCE INTO THE CLASSROOM

← → C 🔒 https://www.google.rs

Google

DARPA's first 200MHz GaAs Microprocessor - a decade before Intel

💷 🌷 🔍

Притисните Enter да бисте претражили.

#### Milutinovic, Veliko (Serbia)

www.balcon-project.eu > ... > Serbia \* Преведи ову страницу

... the **first GaAs microprocessor** in the world, agency **DARPA** project Star Wars, ... project has realized processor speed of **200MHz** about a **decade before Intel**, ...

#### BOTTOM LINE: BRINGING ADVANCED INDUSTRIAL EXPERIENCE INTO THE CLASSROOM

- → C 🔒 https://www.google.rs

Google

DARPA's first 200MHz GaAs Microprocessor - a decade before Intel

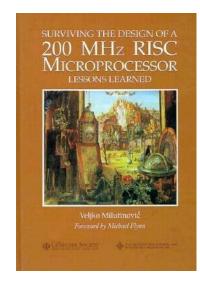
💷 🌵 🔍

Притисните Enter да бисте претражили.

#### Milutinovic, Veliko (Serbia)

www.balcon-project.eu > ... > Serbia \* Преведи ову страницу

... the first GaAs microprocessor in the world, agency DARPA project Star Wars, ... project has realized processor speed of 200MHz about a decade before Intel, ...



### BOTTOM LINE: BRINGING ADVANCED INDUSTRIAL EXPERIENCE INTO THE CLASSROOM (2)



MAXELER - today's fastest dataflow supercomputer for oil and gas industry 🛛 📖 🌷

ni and gas moustry

Q

Притисните Enter да бисте претражили.

#### BOTTOM LINE: BRINGING ADVANCED INDUSTRIAL EXPERIENCE INTO THE CLASSROOM (2)



MAXELER - today's fastest dataflow supercomputer for oil and gas industry 🛛 📖 🌷

Q

Притисните Enter да бисте претражили.

<u>HPCwire: Maxeler Launches MPC-X Series Dataflow Engines</u> www.hpcwire.com/.../maxeler launches mpc-x ... • Преведи ову страницу

21.03.2012. - Market Watch; Events ... "At Maxeler we are excited to offer the fastest computers on the planet ... in Oil and Gas exploration and in a range of other application areas. ...

#### BOTTOM LINE: BRINGING ADVANCED INDUSTRIAL EXPERIENCE INTO THE CLASSROOM (2)

Google

MAXELER - today's fastest dataflow supercomputer for oil and gas industry 🛛 📖 🌷

Q

Притисните Enter да бисте претражили.

HPCwire: Maxeler Launches MPC-X Series Dataflow Engines

www.hpcwire.com/.../maxeler\_launches\_mpc-x\_... • Преведи ову страницу 21.03.2012. - Market Watch; Events ... "At Maxeler we are excited to offer the fastest computers on the planet ... in Oil and Gas exploration and in a range of other application areas. ...



## BOTTOM LINE: BRINGING ADVANCED INDUSTRIAL EXPERIENCE INTO THE CLASSROOM (3)



Ericsson - the ProSense project

Притисните Enter да бисте претражили.

**•** 

Q

### BOTTOM LINE: BRINGING ADVANCED INDUSTRIAL EXPERIENCE INTO THE CLASSROOM (3)

Google

Ericsson - the ProSense project

🔳 🌷 🔍

Притисните Enter да бисте претражили.

PPT Authors - kondor.etf.rs

home.etf.rs/~vm/Belgrade%20overview.ppt ▼ Преведи ову страницу ProSense. 3 /30. ProSense. Project Team. Director for EU: Dr. Srđan Krčo, Ericsson, Ireland. Director for Serbia: Prof. Dr. Veljko Milutinović, UB. Team members.

### BOTTOM LINE: BRINGING ADVANCED INDUSTRIAL EXPERIENCE INTO THE CLASSROOM (3)

Google

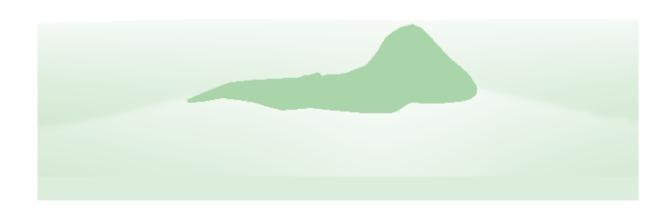
Ericsson - the ProSense project

🔳 🏮 🔍

Притисните Enter да бисте претражили.

IPPT Authors - kondor.etf.rs home.etf.rs/~vm/Belgrade%20overview.ppt ▼ Преведи ову страницу ProSense. 3 /30. ProSense. Project Team. Director for EU: Dr. Srđan Krčo, Ericsson, Ireland. Director for Serbia: Prof. Dr. Veljko Milutinović, UB. Team members.

#### Wireless Sensor Networks: ApplicationDesign and DataMining



# SUGGESTED READINGS:

• Z. Babovic, V. Milutinovic, "Novel System Architectures for Semantic-Based Integration of Sensor Networks," Elsevier, Advances in Computers, 2013.

- Z. Babovic, "DataFlow systems: From their origins to future applications in data analytics, deep learning, and the Internet of Things," In V. Milutinovic et al. "DataFlow Supercomputing Essentials," Springer, 2017.
- Special Issues of Elsevier Advances in Computers 2024.
- Special Issues of Springer Journal of Big Data, 2024.
- IPSI TIR (Transactions on Internet Research)
- IPSI TAR (Transactions on Advanced Research)

