IEEE Rebooting Computing: Motivation, Genesis, and Context

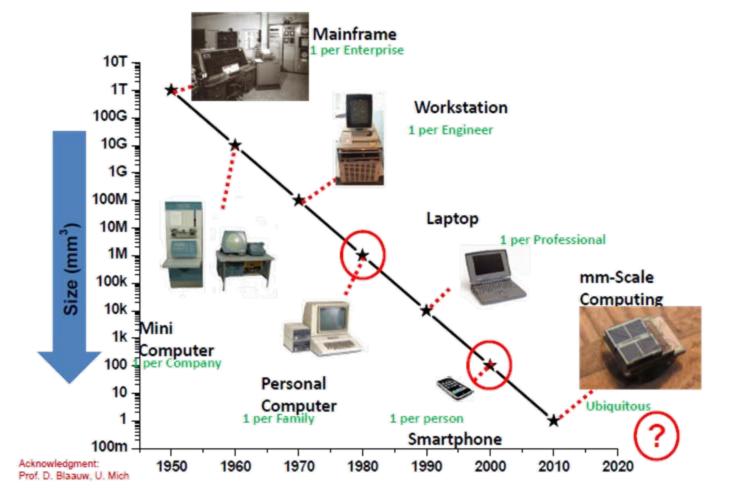
Erik DeBenedictis, presenter

Tom Conte, Bichlien Hoang, Alan Kadin, Yung-Hsiang Lu, Elie K. Track

> IEEE Future Directions Committee Rebooting Computing Working Group



The Past and Future of Computing



- Past exponential improvement due to Moore's Law scaling.
- Is there a Computer Roadmap for the Future?



2 11/2/2015 IEEE Rebooting Computing special session - ICCAD

The End of Moore's Law?

- CMOS transistors are still getting smaller but not faster and not cheaper
 - 11nm wasn't any better than 14nm, which was only marginally better than 22nm.
- Moore's Law really ends before ~2030 at 1.5nm
- 2x performance every 18 months <u>ended in 2005</u>
 Multicore didn't continue scaling limits of parallelism
- Not just logic, memory scaling ending as well
- Also critical power limitation:
 - "Dark Silicon" most of every chip must be turned off to avoid overheating





- Founded in Dec. 2012 by IEEE Future Directions
- Rethink Everything: "Soup to Nuts"
- Why IEEE? Pre-competitve, Inclusive, Worldwide
- 9 participating IEEE Societies and Councils





Rebooting Computing Committee

Co-Chairs:

- Elie Track IEEE CSC
- Tom Conte IEEE Computer Society
- Program Director: Bichlien Hoang (IEEE)

Members:

- Dan Allwood (MAG), Neal Anderson (NTC), David Atienza (CEDA)
- Jesse Beu (CS), Jonathan Candelaria (EDS), Erik DeBenedictis (CS)
- Paolo Gargini (ITRS), Glen Gulak (SSCS), Steve Hillenius (SRC)
- Scott Holmes (EDS), Subramanian Iyer (EDS), Alan Kadin (CSC)
- Arvind Kumar (EDS), Yung-Hsiang Lu (CS), David Mountain (EDS)
- <u>Oleg Mukhanov (CSC)</u>, Vojin G. Oklobdzijja (CASS)
- Angelos Stavrou (RS), William Tonti (RS), <u>Ian Young (SSCS)</u>





rebooting computing

Computing Discipline Model Computing Field Guide **Computing Ontology** CS K-12 Curriculum **Defining Computer Science** Foundations of Disciplinary Identity **Future Computing** Requirements Image of Computing **K-8 FUNdamentals** LabRats Multi-disciplinary Collaboration **MultiCore Open Artifact** -**Recombinant Systems Project-Problem Based** Learning **Recruiting CS Teachers Recruiting Women &** Minorities into CS Science of Computation Socially Relevant CS Tools for Fun and Beauty

The Magic and Beauty of Computer Science



A project sponsored by:



In cooperation with:





Silicon Valley, CA. From January 12-14th 2009

Peter Denning Frank J. Barrett Tim Bell **Geoff Brown** Vint Cerf **Robb** Cutler John Dunnion Ron Fry Susanne Hambrusch Susan Higgins Alan Kay Leonard Kleinrock Joshua Kroll Craig Martell Andrew McGettrick Jeff Moser Peter G. Neumann **Richard Snodgrass** Lawrence Snyder

Mission of Rebooting Computing

Reexamine the foundations of computing

- Logic and memory devices, micro- and macro-architectures
- Both Software and Hardware Systems
- Sponsor Meetings, Publications, Competitions
 - Series of RC Summits
 - Special Dec. issue of IEEE Computer Magazine on RC
 - Low-Power Image Recognition Challenge (LPIRC)
- Team with other organizations to promote goals
 - ITRS International Technology Roadmap for Semiconductors
 - Government National Strategic Computing Initiative, etc.

Develop new Computing Roadmaps and Standards



RC on the Internet

IEEE RC Web Portal

http://rebootingcomputing.ieee.org/

- Features the latest news, articles, and videos related to RC
- Includes links to all RC Meeting Reports
- Updated once or twice per month

IEEE RC Technical Community

- Join for free, even if not IEEE member
- RC Social Media
 - Twitter: <u>https://twitter.com/IEEERebootComp</u>
 - Facebook: <u>https://www.facebook.com/IEEE-Rebooting-Computing-1599003380346914</u>
 - LinkedIn: <u>https://www.linkedin.com/grp/home?gid=8124309</u>
 - Blog: <u>http://rebootingcomputing-ieee.blogspot.com/</u>



IEEE Rebooting Computing

- RCS 1: Dec. 2013 Washington, DC
 - Three Pillars of Future Computing Energy Efficiency, Security, Applications/HCI



Low-Power Image Recognition Challenge

Competition is an efficient way to assess the state of the art in the field.



Space X Prize



DARPA Autonomous Vehicle

MORE IN NEXT TALK



RC Special Issue of Computer Magazine

- Special issue on RC in IEEE Computer Magazine (flagship monthly magazine of IEEE Computer Society)
 - Guest editors from RC Committee:
 Erik DeBenedictis, Tom Conte, and Elie Track, Tom Conte
 - Directed at general audience of computer engineers and scientists
 - Coming out in early December 2015

Sneak Peek of Articles in Issue

- Computing beyond Moore's Law, J. Shalf (LBL)
- Memristor-Based Gate Networks, C. Krieger, et al. (LPS)
- Memory-Based Computing, K. Bresniker et al (HP Labs)
- Nanoscale Intelligent Architectures, Khasanvis et al (Umass)
- Nanotube-Based 3DIntegration, M. Sabry Ali et al (Stanford)
- Superconducting Computing, D.S. Holmes, A. Kadin, & M. Johnson



ITRS Collaboration

International Technology Roadmap for Semiconductors

 Sponsored by international semiconductor industry to plan future trends in IC fabrication following Moore's Law

Joint Agreement between IEEE and ITRS in 2015

- Coordinated efforts on planning meetings, reports, and roadmaps.
- Joint meetings at Stanford Univ. in Feb. and Jul. 2015.



ITRS 2.0 <u>http://www.itrs2.net/</u>

 New program of ITRS, with new broader focus on both emerging research devices and on system applications.



Upcoming RCS 4

- Roadmapping the Future of Computing:
 - Discovering How We May Compute
- 9-11 Dec. 2015, Washington Hilton, Washington DC
 - Follows the end of International Electronic Devices Meeting (<u>IEDM 2015</u>)
- A concise description of the major findings to date from the RC initiative;
- An overview of major national and international initiatives;
- A detailed presentation applying the framework developed to emerging engines of computing;
- A poster session (*contributions solicited*) providing additional details on current research efforts;
- In-depth discussions to focus the momentum towards creating a *Roadmap of Future Computing*.



National Initiatives

National Strategic Computing initiative

- <u>https://www.whitehouse.gov/the-press-</u> office/2015/07/29/executive-order-creating-national-strategiccomputing-initiative</u>
- Administered by Office of Science & Technology Policy (OSTP)
- Nanotech-Inspired Grand Challenge for Future Computing
 - <u>http://www.nano.gov/futurecomputing</u>

RC and NSCI are largely in-synch

- NSCI officials attended earlier RC Summits
- RC plans to offer additional assistance as NSCI develops
- Sensible Machines: Above and Beyond Exascale Computing
 - Preliminary White Paper available online at <u>http://rebootingcomputing.ieee.org/archived-articles-and-videos/general/sensible-machine</u>



14—

A Brief History of NSCI and Grand Challenge

- IEEE Rebooting Computing Summit 1, December 2013, presentation by Rob Leland on NSCI
- OSTP RFI: "Nanotechnology-Inspired Grand Challenges for the Next Decade" June 17, 2015
 - Reminds people of NNI program from the 1990s
- Submitted a response to RFI entitled "Sensible Machines" June 24
- Executive Order: National Strategic Computing Initiative July 29
 - An exascale program + fundamental science component for the future
- Response from OSTP: Sensible Machines shortlisted, asked to 'develop a program' – July 30
- NSCI workshop Oct. 20
 - Tom Kalil announces a new Grand Challenge

"to develop transformational computing capabilities by combining innovations in multiple scientific disciplines"

 Most of 2-day workshop on Exascale computers, both 1 Exaflops and future path

Exascale Computing and Energy

Enormous power consumption in data centers and supercomputers

Different applications (big data vs. scientific calculations), but similar large parallel machines with up to ~ 10 MW power Electrical power costs ~ \$1M/year for 1 MW of power.

Earlier 2010 target: PetaFLOPS/s computer

10¹⁵ floating-point operations per second

Equivalent to 10⁶ parallel processors at 1 GHz.

Largest present supercomputers ~ 30 PFLOPS

New 2020 target: ExaFLOPS/s computer

10¹⁸ FLOPS = 1000 PFLOPS 10⁹ CMOS processors at 1 GHz? Too much power! A different approach is needed! Major R&D effort into alternatives required.



Superconducting Supercomputing?



	Supercomputer Titan at ORNL - #2 of Top500	Superconducting Supercomputer	
Performance	17.6 PFLOP/s (#2 in world*)	20 PFLOP/s	~1x
Memory	710 TB (0.04 B/FLOPS)	5 PB (0.25 B/FLOPS)	7x
Power	8,200 kW avg. (not included: cooling, storage memory)	80 kW total power (includes cooling)	0.01x
Space	4,350 ft ² (404 m ² , not including cooling)	~200 ft ² (includes cooling)	0.05x
Cooling	additional power, space and infrastructure required	All cooling shown	

Courtesy of M. Manheimer

INTELLIGENCE ADVANCED RESEARCH PROJECTS ACTIVITY (IARPA)

#1 in TOP500, 2012-11 (17.6 PFLOP/s)







The "Sensible Machine" response to OSTP RFI

The central thesis of this white paper is that although our present understanding of brains is limited, we know enough now to design and build circuits that can accelerate certain computational tasks; and as we learn more about how brains communicate and process information, we will be able to harness that understanding to create a new exponential growth path for computing technology. – Stan Williams)

I believe this statement much more now than when I first wrote it.



A Nanotechnology-Inspired Grand Challenge for Future Computing

Create a new type of computer that can proactively interpret and learn from data, solve unfamiliar problems using what it has learned, and operate with the energy efficiency of the human brain.

Additional detail:

While it continues to be a national priority to advance conventional digital computing—which has been the engine of the information technology revolution—current technology falls far short of the human brain in terms of both the brain's sensing and problemsolving abilities and its low power consumption. Many experts predict that fundamental physical limitations will prevent transistor technology from ever matching these twin characteristics. We are therefore challenging the nanotechnology and computer science communities to look beyond the decades-old approach to computing based on the Von Neumann architecture as implemented with transistor-based processors, and chart a new path that will continue the rapid pace of innovation beyond the next decade.

There are growing problems facing the Nation that the new computing capabilities envisioned in this challenge might address, from delivering individualized treatments for disease, to allowing advanced robots to work safely alongside people, to proactively identifying and blocking cyber intrusions. To meet this challenge, major breakthroughs are needed not only in the basic devices that store and process information and the amount of energy they require, but in the way a computer analyzes images, sounds, and patterns; interprets and learns from data; and identifies and solves problems.

Many of these breakthroughs will require new kinds of nanoscale devices and materials integrated into three-dimensional systems and may take a decade or more to achieve. These nanotechnology innovations will have to be developed in close coordination with new computer architectures, and will likely be informed by our growing understanding of the brain—a remarkable, fault-tolerant system that consumes less power than an incandescent light bulb.

Key words: learn, interpret, data, "energy efficiency of the human brain" Key words absent: Neuromorphic, intelligence



OSTP Nanotechnology-Inspired Grand Challenge: Sensible Machines (extended version 2.5)

R. Stanley Williams Hewlett-Packard Laboratories

Erik P. DeBenedictis Sandia National Laboratories

October 12, 2015

URLs for further information

• White House announcement of Grand Challenge:

https://www.whitehouse.gov/blog/2015/10/15/nanotechnology-inspired-grand-challenge-future-computing

• nano.gov grand challenges portal:

http://www.nano.gov/grandchallenges

• IEEE Rebooting Computing Website:

http://rebootingcomputing.ieee.org/archived-articles-and-videos/general/sensible-machine

• Sensible Computer White Paper:

http://rebootingcomputing.ieee.org/images/files/pdf/SensibleMachines_v2.5_N_IEEE.pdf



Structure of a US Neuromorphic Computing Program

1. Connect Theory of Computation with Neuroscience and Nonlinear Dynamics

e.g. Boolean, CNN, Baysian Inference, Energy-Based Models, Markov Chains

- 2. Architecture of the Brain and Relation to Computing and Learning Theories of Mind: Albus, Eliasmith, Grossberg, Mead, many others
- 3. Simulation of Computational Models and Systems
- 4. System Software, Algorithms & Apps Make it Programmable/Adaptable
- 5. Chip Design System-on-Chip: Accelerators, Learning and Controllers

Compatible with standard processors, memory and data bus

6. Chip Processing and Integration – Full Service Back End of Line on CMOS

DoE Nanoscale Science Research Centers (NSRCs) - e.g. CINT

7. Devices and Materials – in situ and in operando test and measurement



HEWLETT PACKARD LABS PRESENTS

The Chua Lectures: From Memristors and Cellular Nonlinear Networks to the Edge of Chaos

Sept 8 – Nov 24



Hewlett Packard Labs is kicking off an exciting 12-part lecture series with the world-renowned <u>Professor Leon Chua</u> – accomplished IEEE Fellow, Professor of Electrical Engineering and Computer Sciences at UC Berkeley, and a pioneer in neural network and Memristor research.

Over the course of the 12 weekly lectures, Professor Chua will offer a peek into his life's work exploring distinct research areas which have emerged from highly nonlinear and dynamical phenomena including: Memristors, Cellular Nonlinear Networks (CNN), The Local Activity Principle and the Edge of Chaos.

Don't miss this rare opportunity to hear from one of the greatest thought leaders of our industry. 'Linearize then analyze' is no longer valid for understanding nanodevices or neurons – a new mathematical theory of electronics is needed, and was developed 35 years ago!

Event details:



What: Chua Lecture Series When: Every Tuesday starting September 8 through November 24, 10:30 a.m. – 12:00 p.m. Pacific Time How to attend: <u>Register here</u> – Attend in person or on the web.

•We highly recommend that you take the opportunity to hear the Professor Chua lectures in person at Building 20 Auditorium in Palo Alto.

Register for upcoming lecture and scroll down the registration page to sign up for the entire lecture series. This is open to everyone so feel free to share this event with your colleagues, friends and social networks!



Future of Rebooting Computing

New RC Meetings and Workshops

- Series of RC Meetings, dedicated RC Conference(s), special sessions
- Continue Low-Power Image Recognition Challenge (LPIRC), etc.

Broaden Collaborations

- Work with leaders in government on promoting R&D agenda for future computing.
- Strengthen interactions with computer industry, consortia, academia, and other organizations to promote future computing.

Develop IEEE/ITRS Computing Roadmap

Annual online reports with projections and recommendations

Develop new IEEE Standards

– Benchmarks for new classes of Computer Systems.





Web portal <u>http://rebootingcomputing.ieee.org</u>



11/2/2015 IEEE Rebooting Computing special session - ICCAD