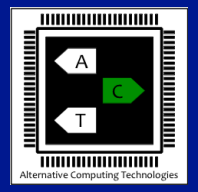
Microsoft Research Faculty Summit 2014

15TH ANNUAL

Microsoft Research
Faculty Summit
2014 15TH ANNUAL

Approximation: Enabling Perpetual Computing on the Edge

Hadi Esmaeilzadeh
Alternative Computing Technologies (ACT) Lab
School of Computer Science
Georgia Institute of Technology



Microsoft

Approximation: Enabling Perpetual Computing on the Edge

Microsoft Faculty Summit

Impossibly Small Devices

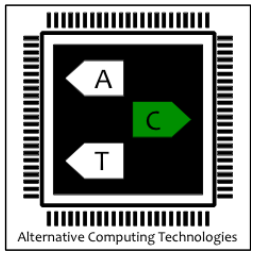
July 14, 2014

Hadi Esmaeilzadeh

Alternative Computing Technologies (**ACT**) Lab

School of Computer Science

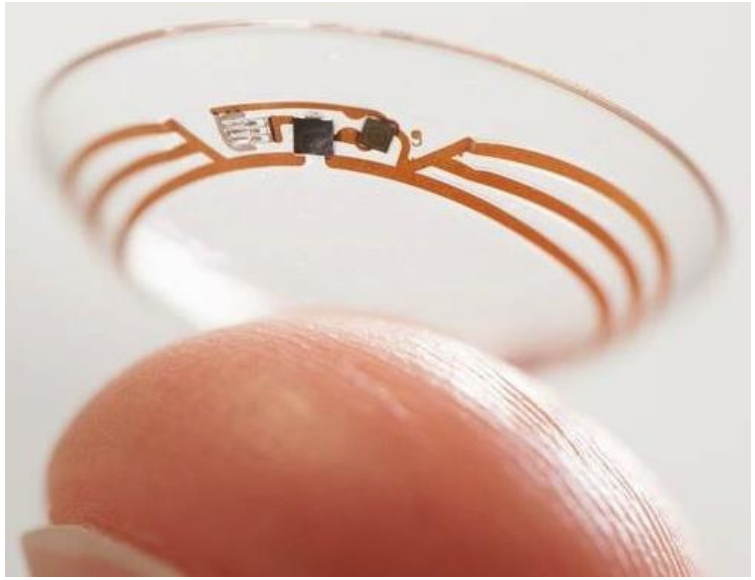
Georgia Institute of Technology

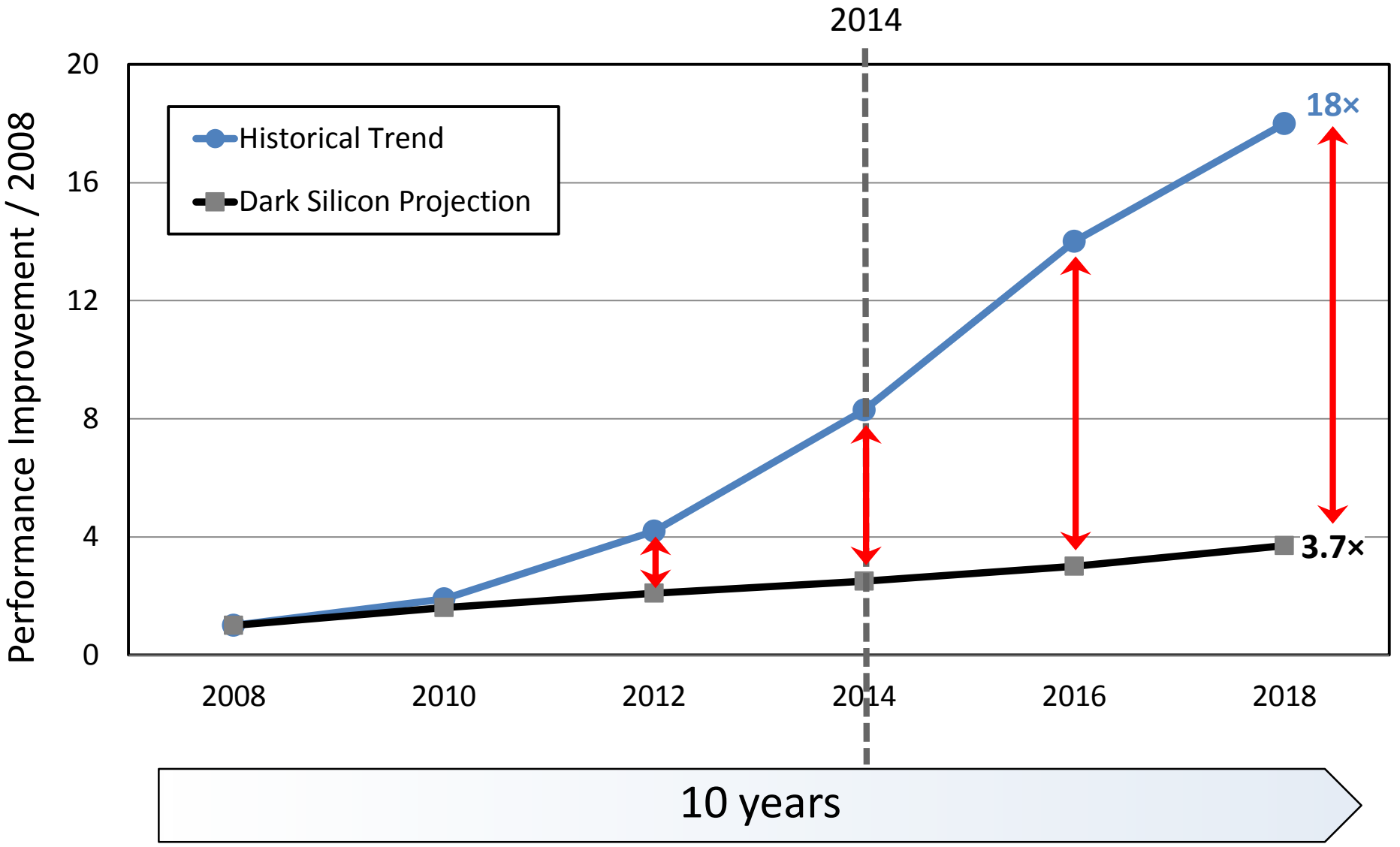


Time for a **perpetual** revolution

Mobile Computing

Perpetual Computing





Esmaeilzadeh, Belem, St. Amant, Sankaralingam, Burger, "Dark Silicon and the End of Multicore Scaling," ISCA 2011
IEEE Micro Top Picks 2012, Communications of the ACM Research Highlights 2013

Radical shift is necessary

Radical departures from conventional approaches are necessary

- Provide efficiency and performance that can enable the **perpetual revolution**
- Extract more performance and efficiency from silicon or any new technology that may dominate

Enable this new capability and meet its constraints

- Cost, power, energy, time, bandwidth, capacity, …

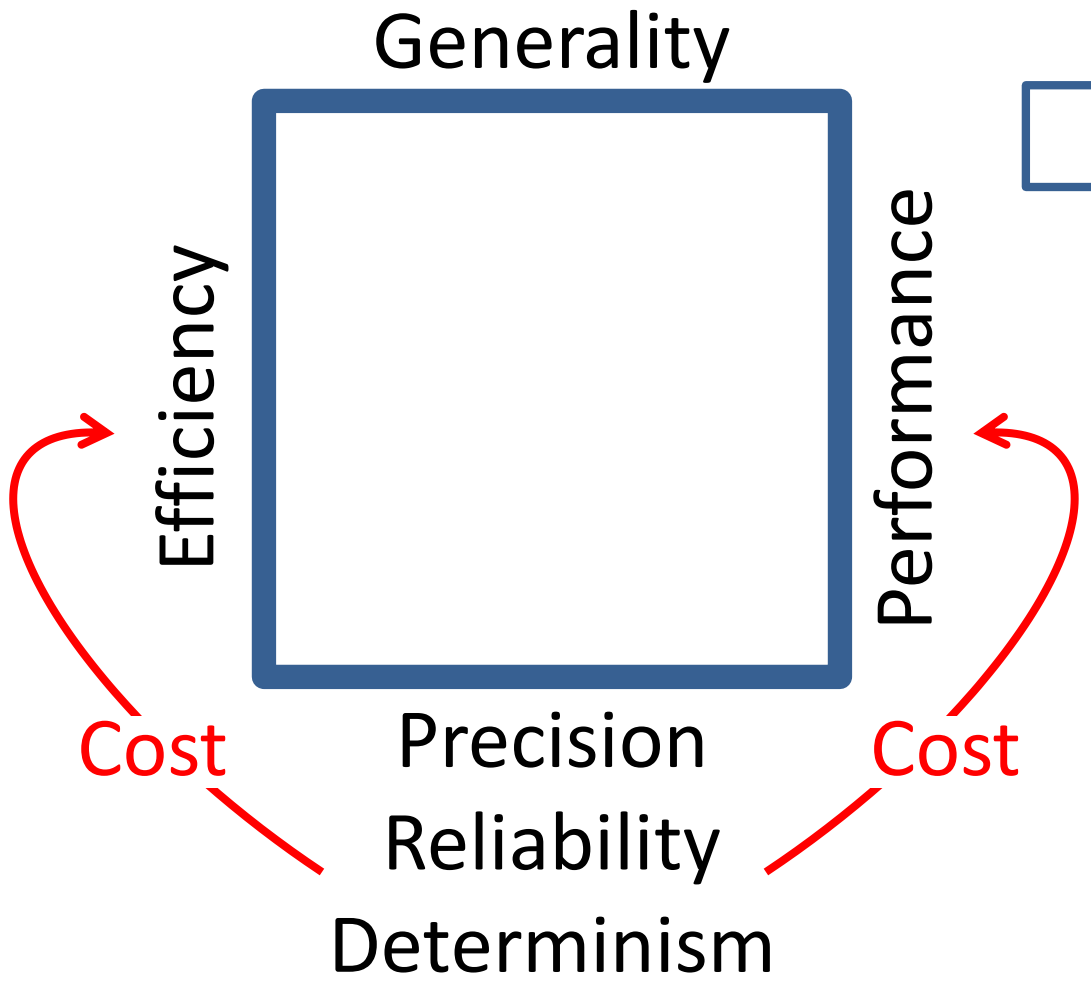
Approximate computing

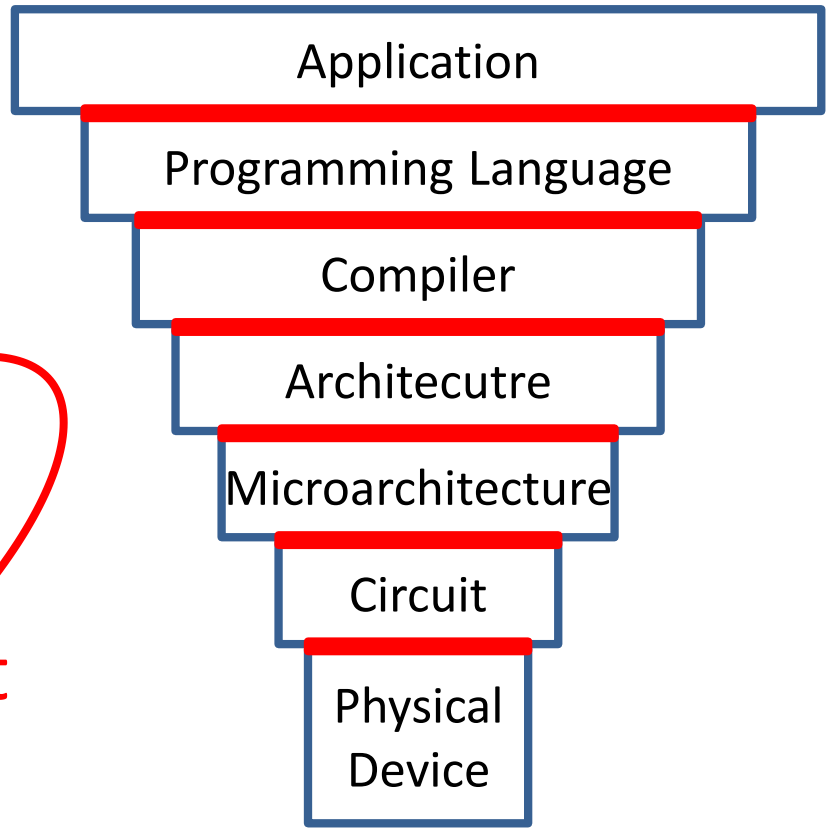
Embracing error

- Relax the abstraction of ***near-perfect*** accuracy in ***general-purpose*** computing/communication/storage
- Allow errors to happen during computation/communication/storage
  - Improve resource utilization efficiency
    - Energy, bandwidth, capacity, …
  - Improve performance
- Build ***acceptable*** systems from intentionally-made unreliable software and hardware components
- Avoid **overkill** and **worst-case** design

Avoiding overkill design

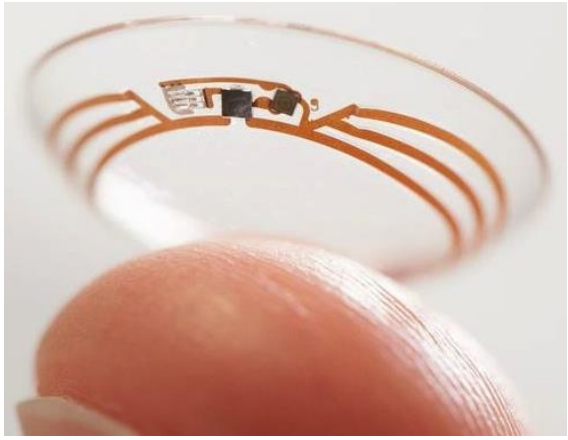
Approximate Computing





Data Analytic in the Cloud

Perpetual Edge Device

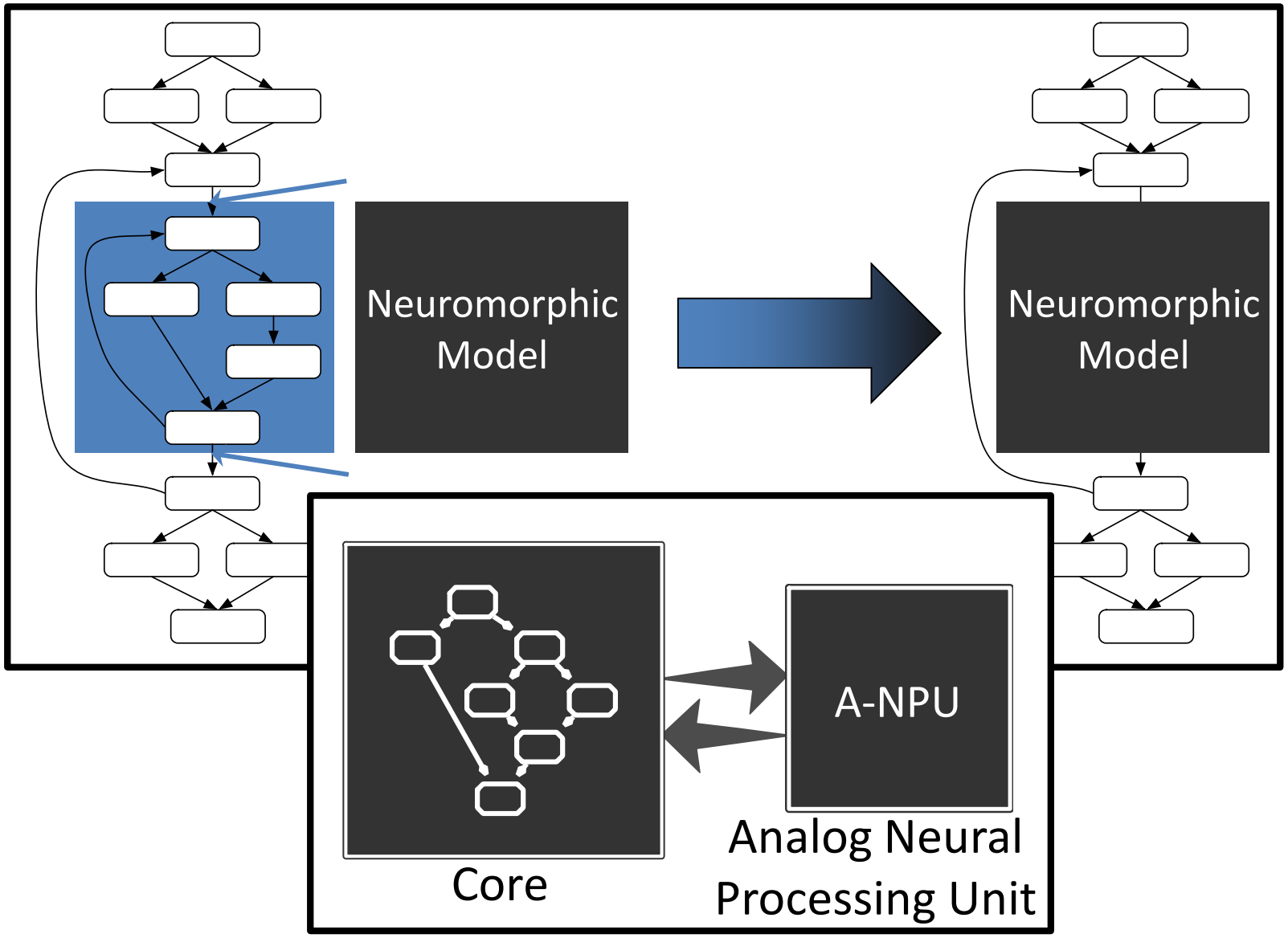


Approximate Acceleration

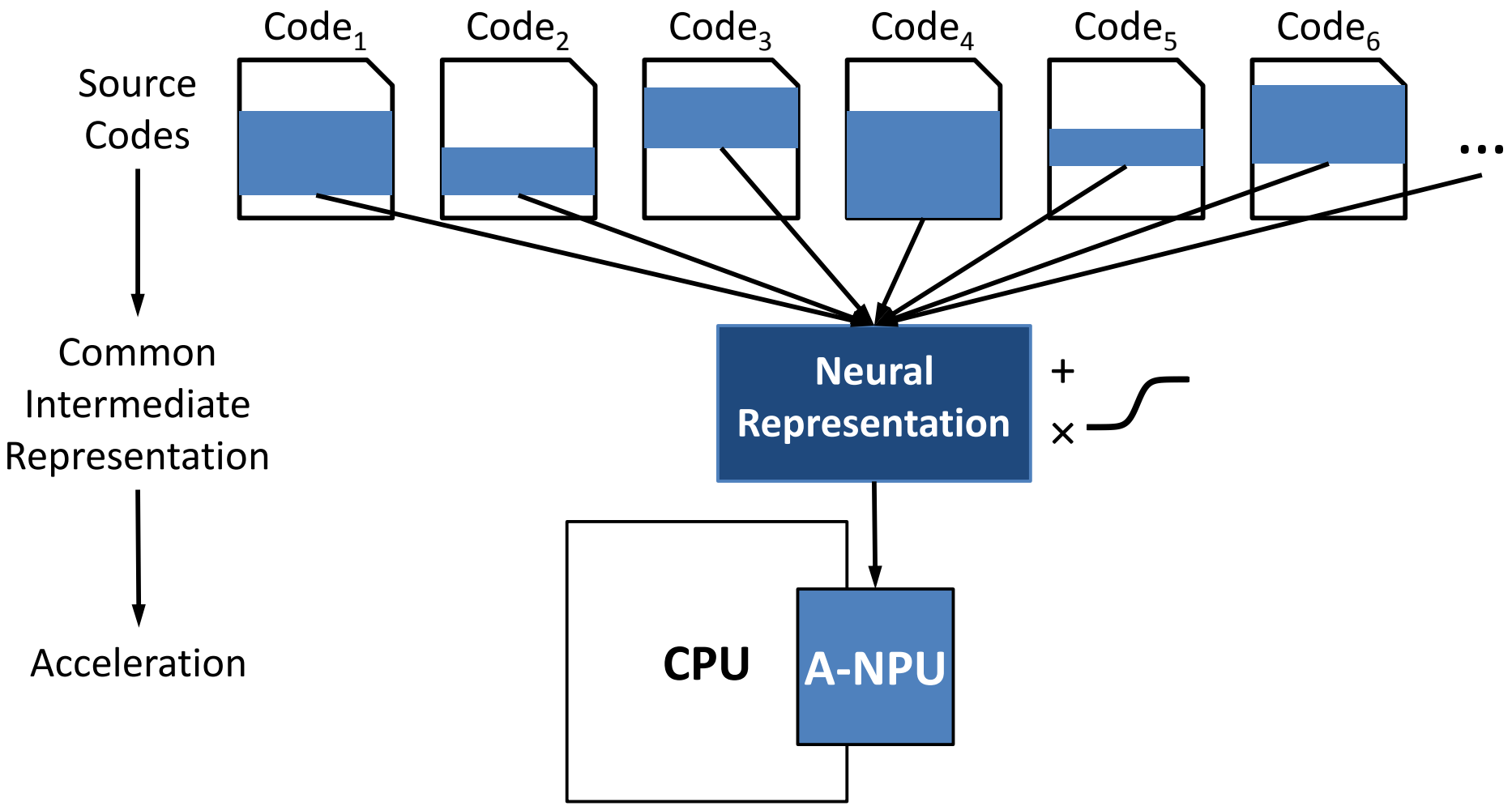
Bridging Neuromophic and von Neumann Computing

Unleashing the Beast

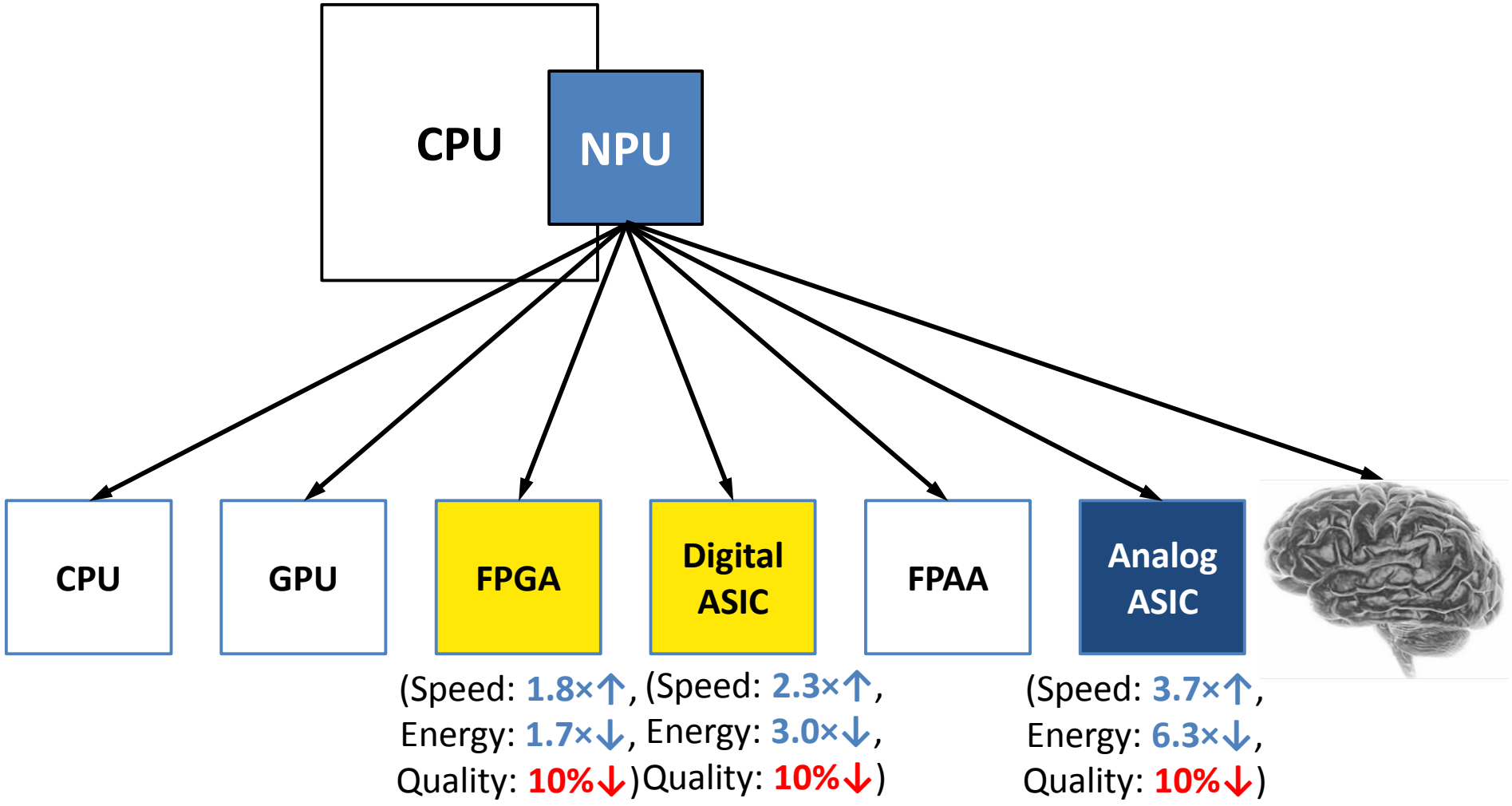
Neural Algorithmic Transformation



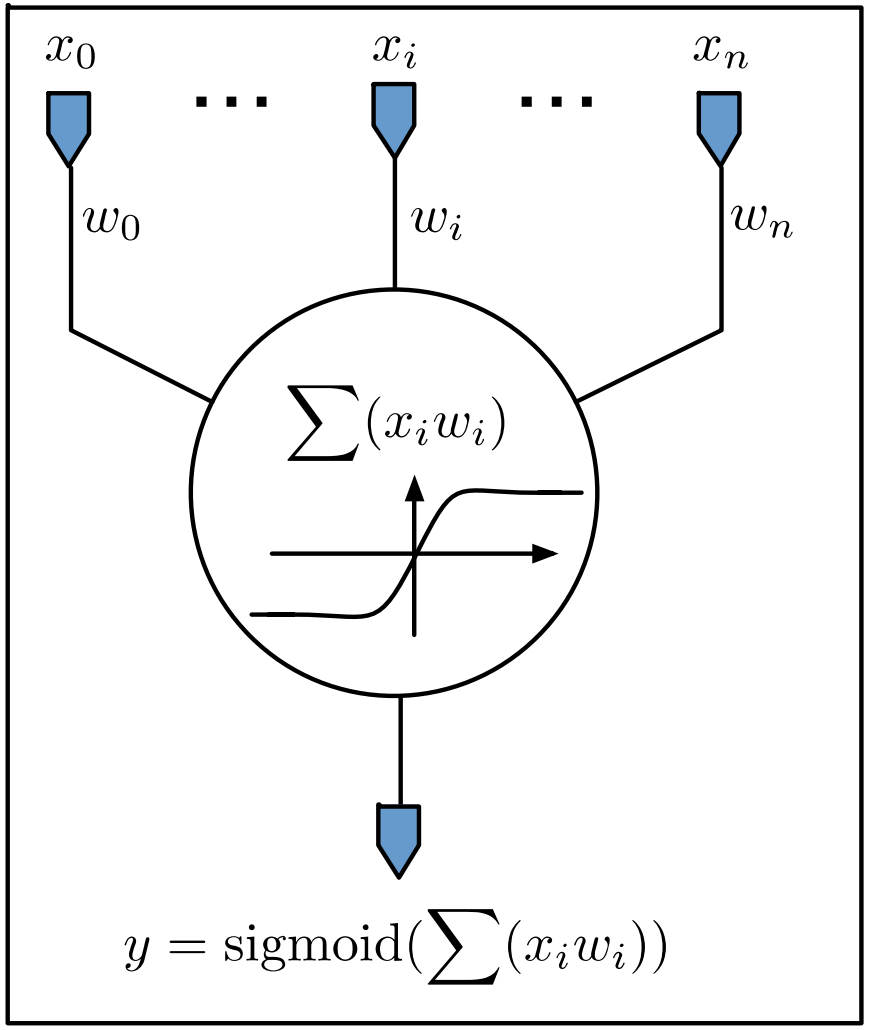
Neuromorphic Model as an Intermediate Representation

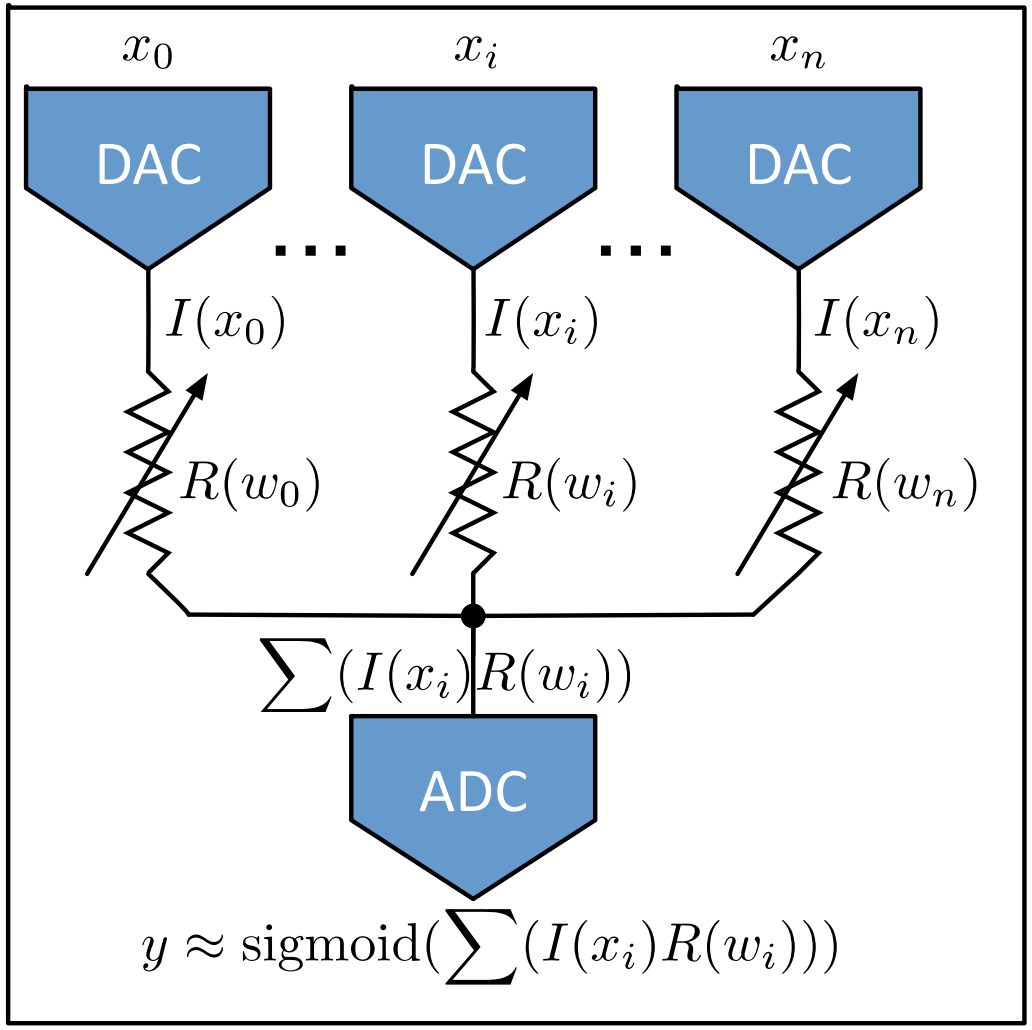


NPU design alternatives



Analog Circuits for Neural Computation





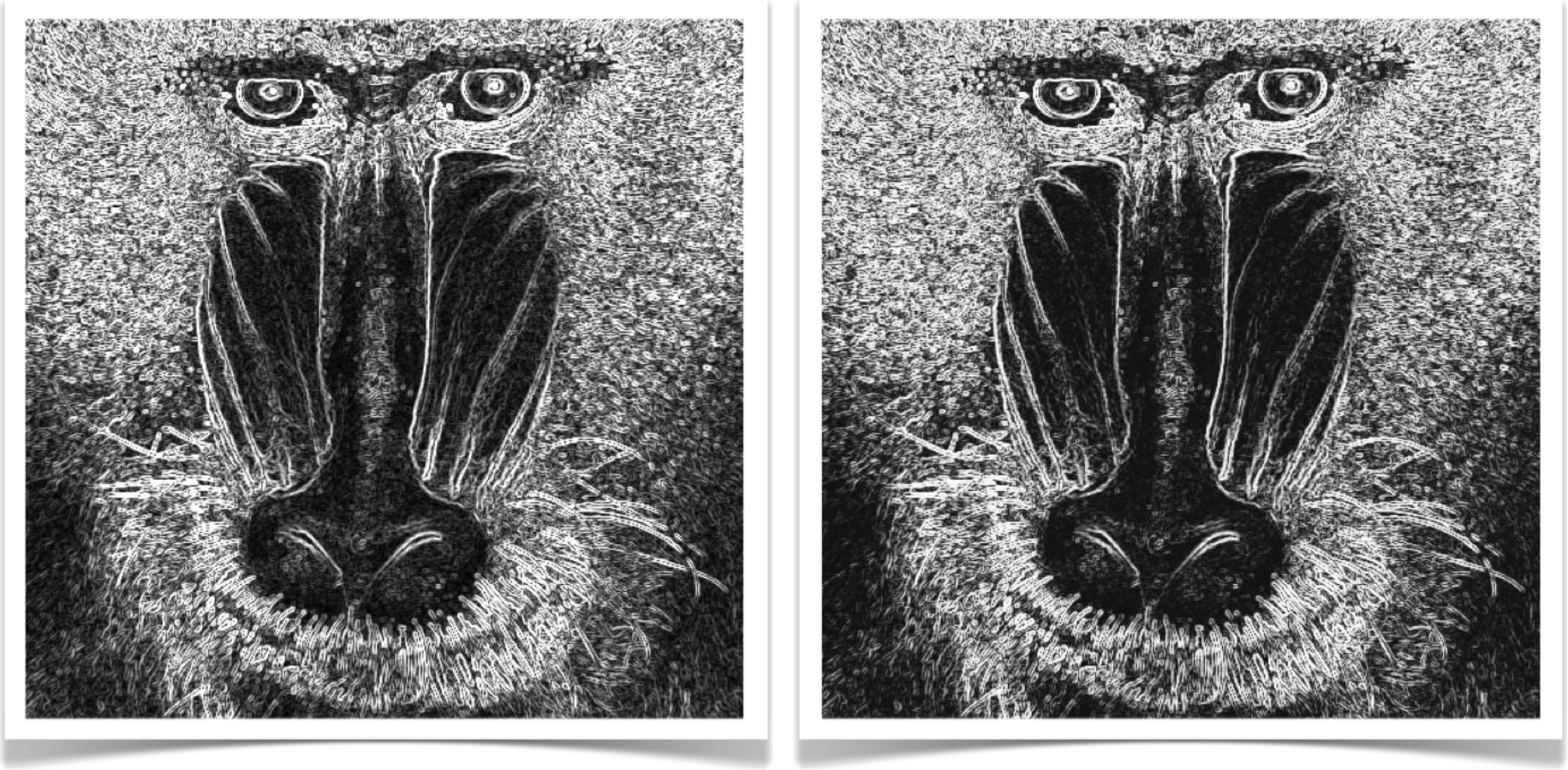
A-NPU acceleration: a fair bargain

23× improvement in energy-delay product

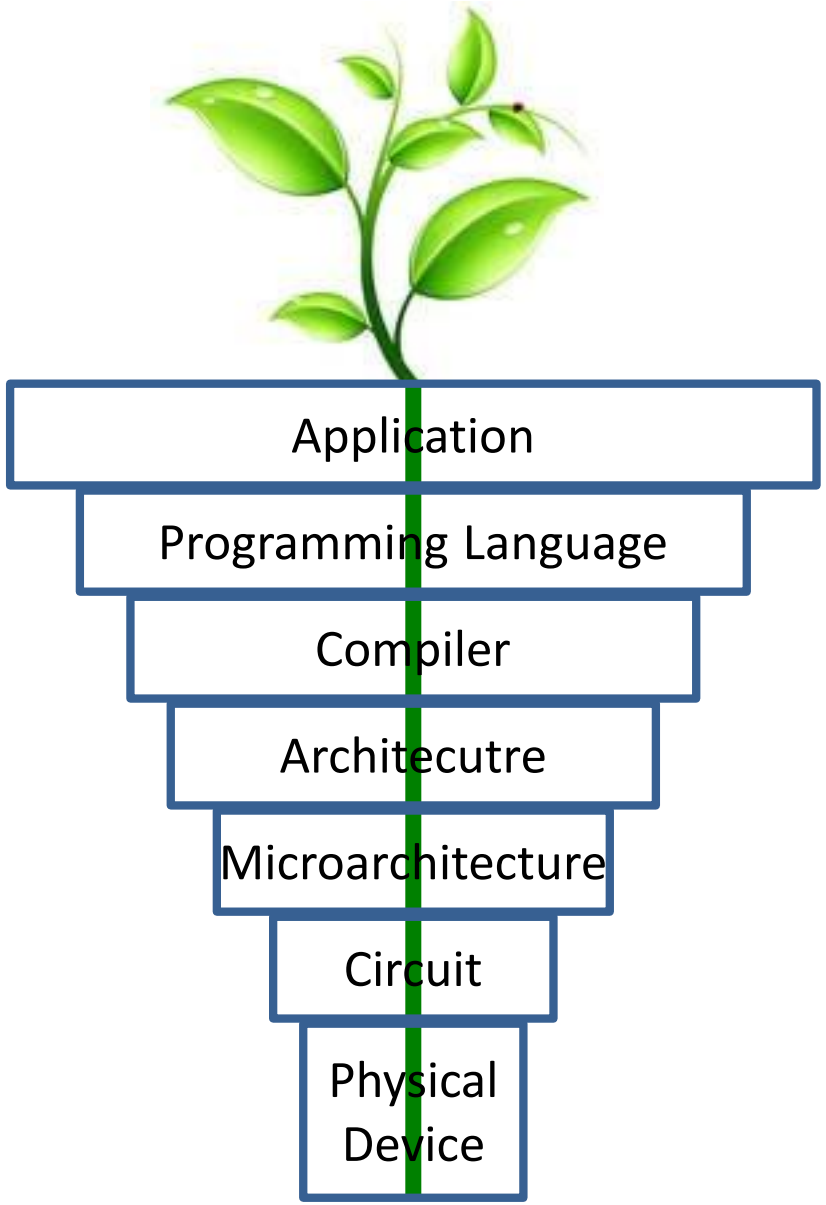
- 3.7× speedup
- 6.3× energy reduction

Less than 10% quality loss

Approximate computing versus conventional computing

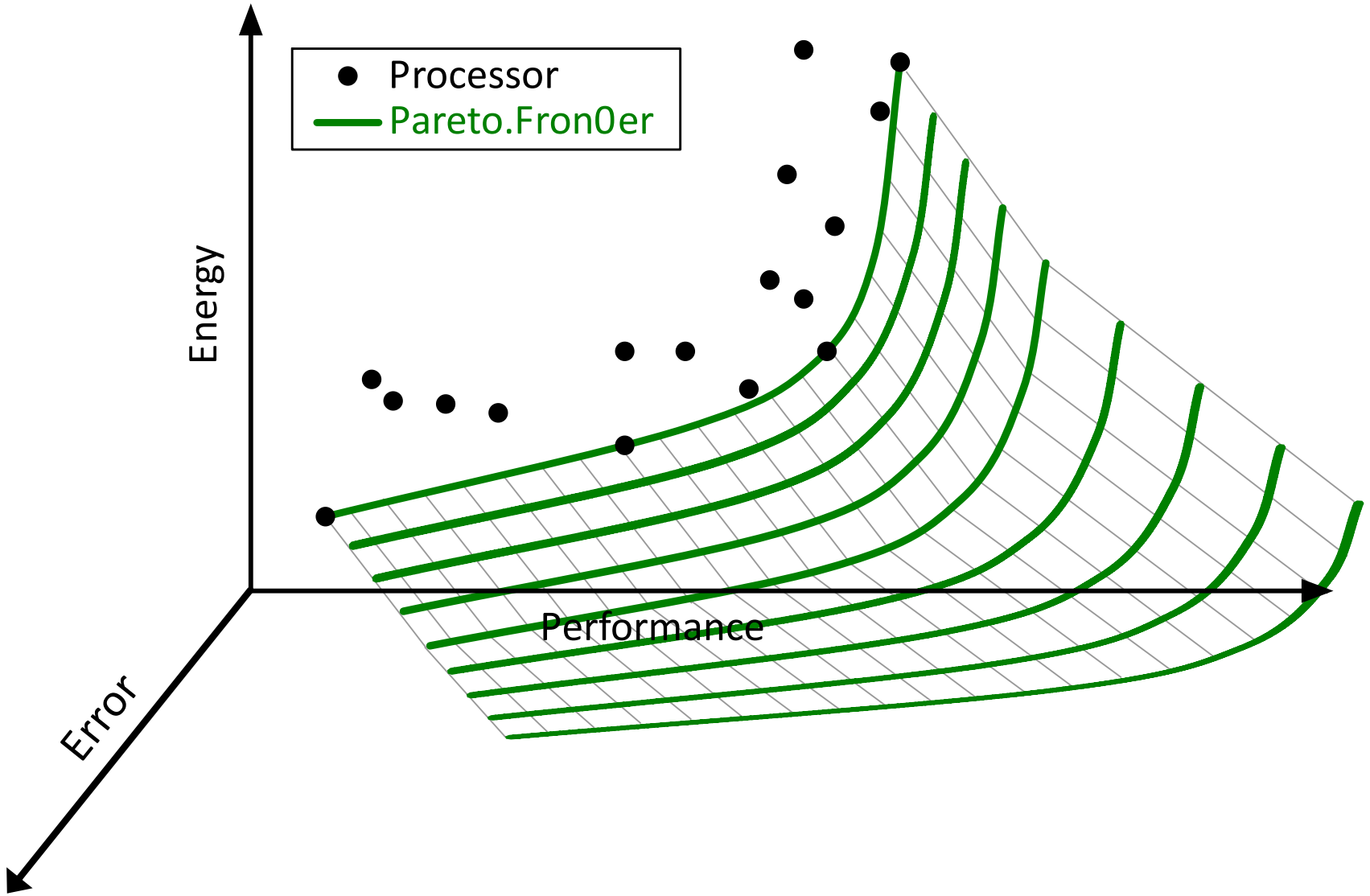


Vertical approach is essential



Embracing error

Enabling **perpetual** computing



Mahmoud Farshchian, “The Fifth Day of Creation,” 1990

Constraint-centric programming

```
float grayscale(float red, float green, float blue) {
    float  luminance =
        red * 0.30 + green * 0.59 + blue * 0.11;

    constrain error[luminance] < 10%;
                energy < 50%

    return luminance;
}
```

Save the planet and return
your name badge before you
leave (on Tuesday)

Microsoft Privacy Policy statement applies to all information collected. Read at research.microsoft.com

© 2014 Microsoft Corporation. All rights reserved. Microsoft, Windows, Windows Vista and other product names are or may be registered trademarks and/or trademarks in the U.S. and/or other countries.
The information herein is for informational purposes only and represents the current view of Microsoft Corporation as of the date of this presentation. Because Microsoft must respond to changing market conditions, it should not be interpreted to be a commitment on the part of Microsoft, and Microsoft cannot guarantee the accuracy of any information provided after the date of this presentation. MICROSOFT MAKES NO WARRANTIES, EXPRESS, IMPLIED OR STATUTORY, AS TO THE INFORMATION IN THIS PRESENTATION.