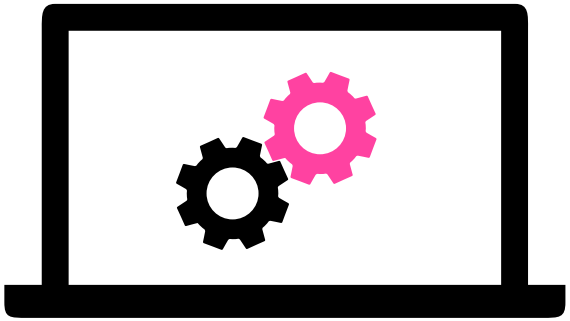
The Perspective of the

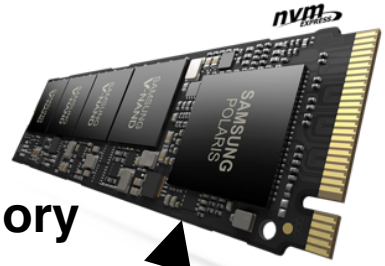


Computer Architect

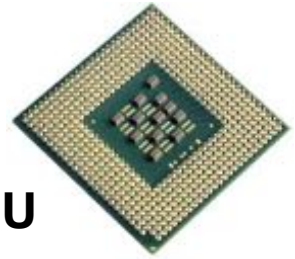
NOEL

CMOS Circuits

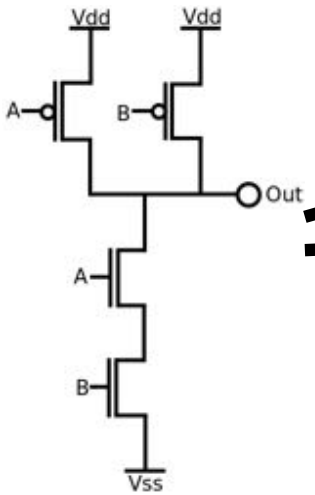
Transistors



memory



CPU



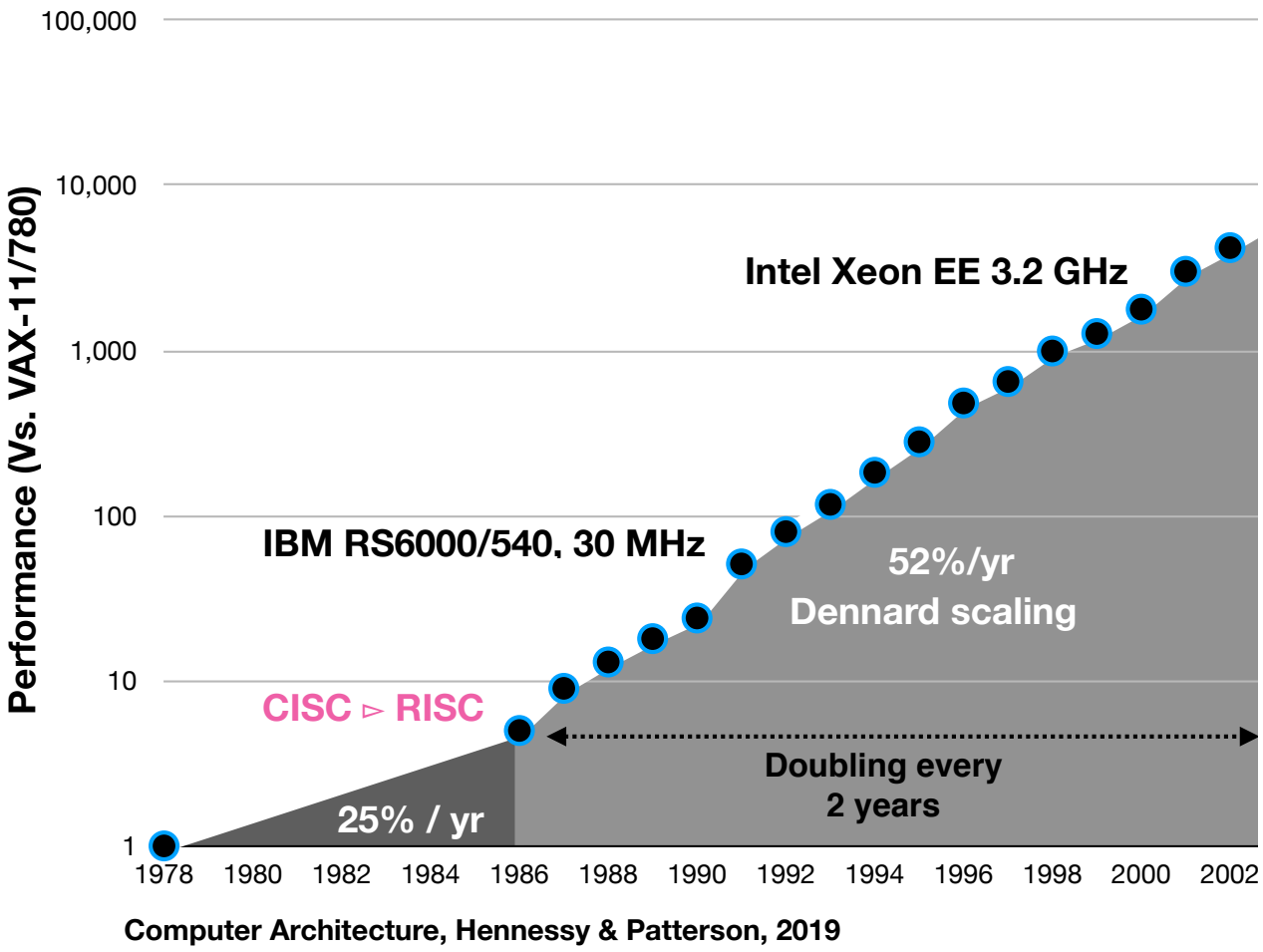
logical gates

float
MOSFET

ALU, control

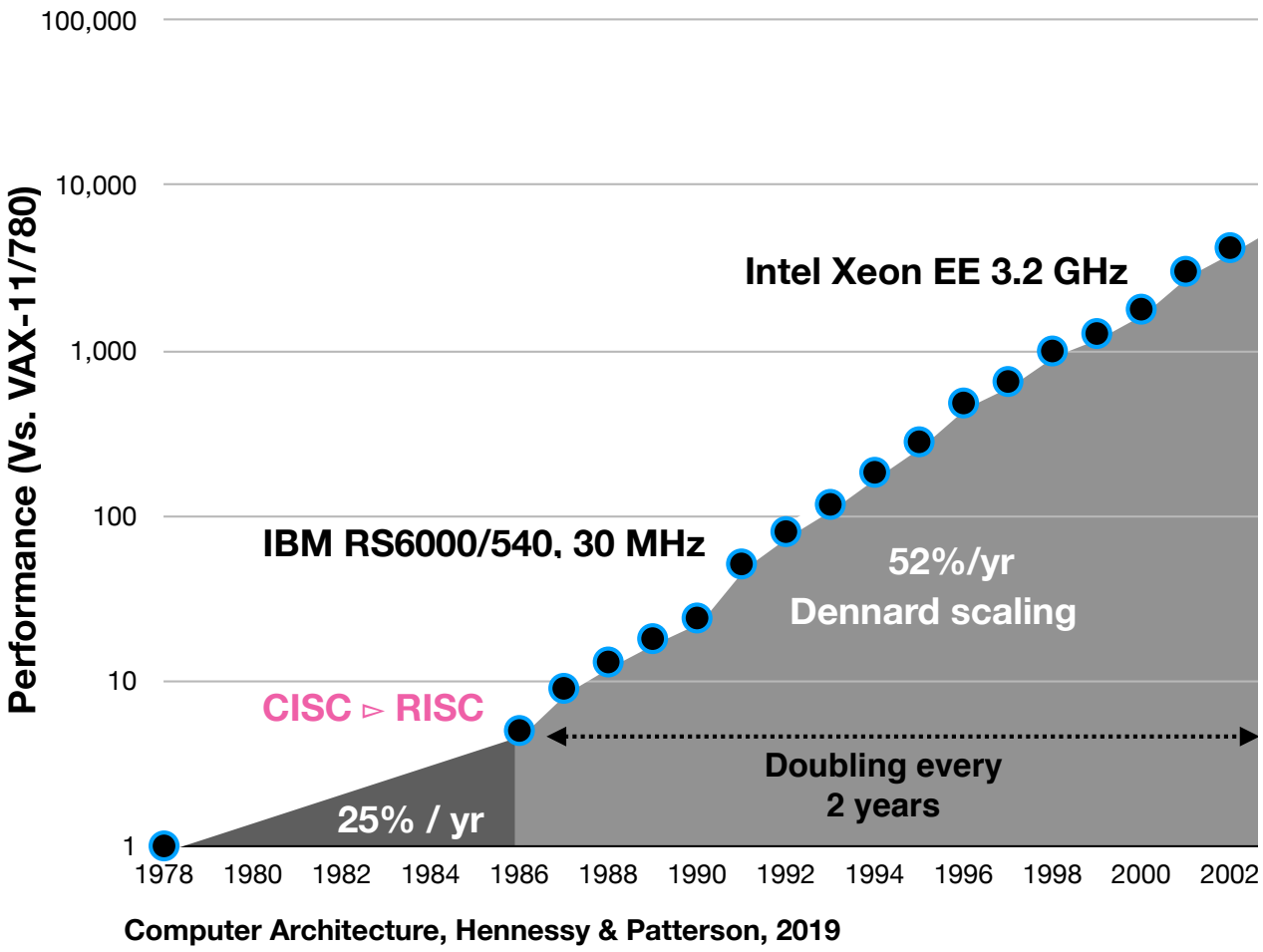
NOEL

Computing Performance



Computer Architecture, Hennessy & Patterson, 2019

Computing Performance



Computer Architecture, Hennessy & Patterson, 2019

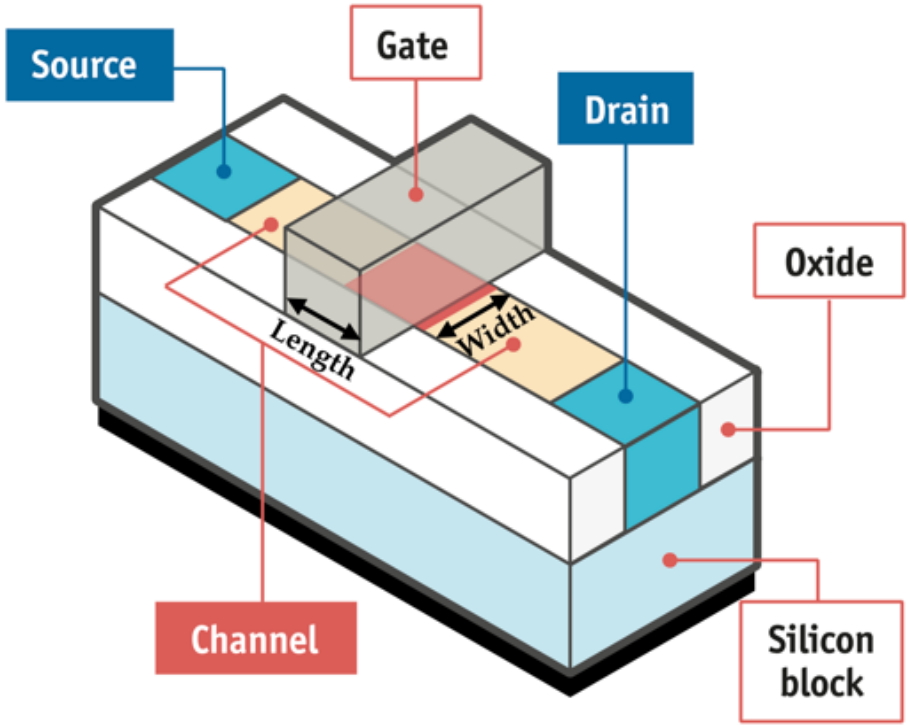
Apollo 11
was landed on the moon using a
computer that had
1,300 times less processing power
then the iPhone5

Two Strategies

Pack more transistors

Clock faster

The Electrons Highway



A transistor

Each lane is 5-nm wide, resulting in 6 lanes for 30 nm transistor

Highway of electrons

The Electrons Highway

dangling bonds at the silicon–silicon-dioxide interface

Accidents

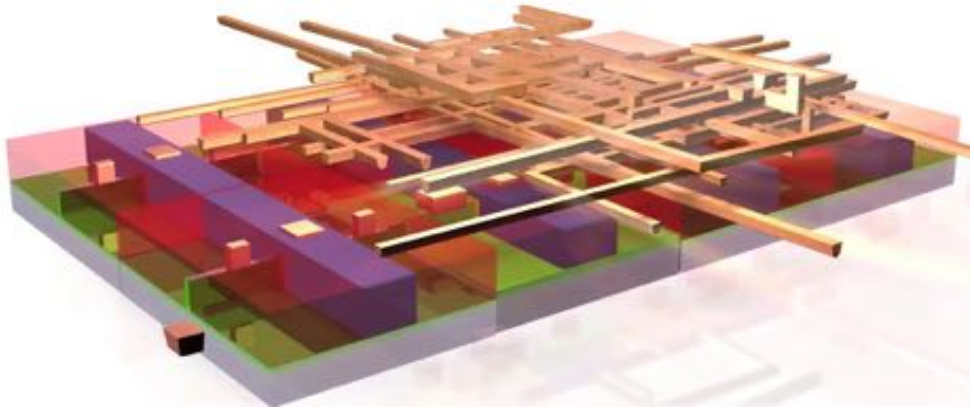
Given the the density of traps (5.4×10^{-4}/ nm^2) and the transistors' surface area (78.5 nm^2), each single-lane transistor will have 0.042 traps. Therefore, 4.2 percent of these nanotransistors will have one or more traps.

The Electrons Highway

Accidents

Going 3D

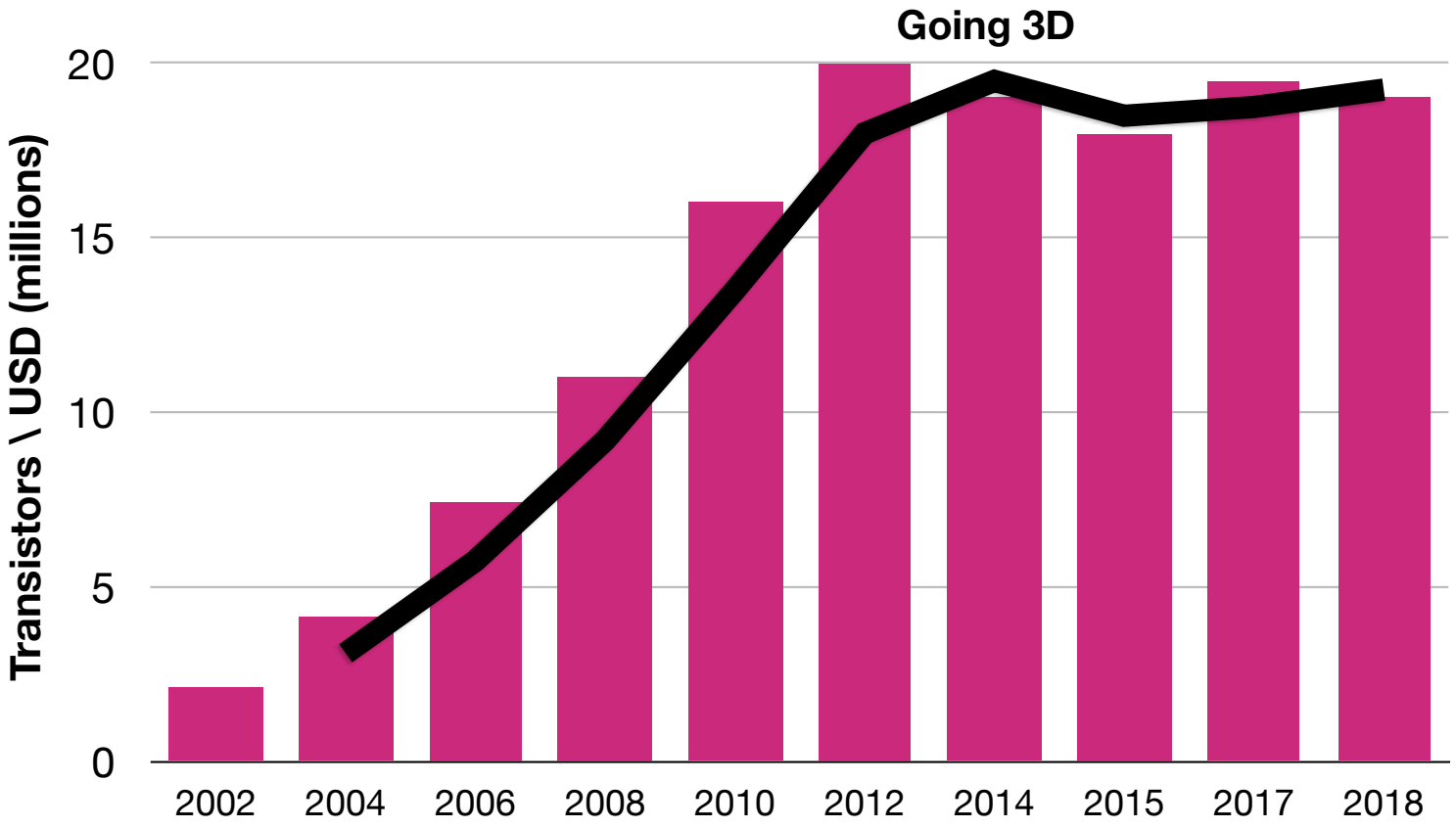
The Electrons Highway



3D transistor

Going 3D

Computing Performance



A Neuromorph's Prospectus Kwabena Boahen, IEEE, 2018

NOEL

Clocking Faster

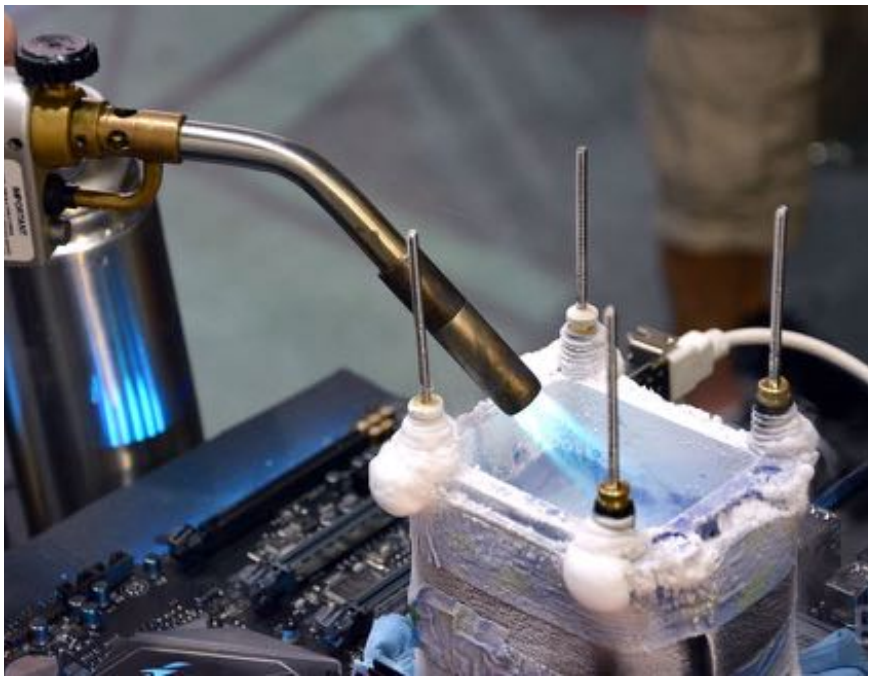
Over-clocking

Over-heating

$$\text{Power}_{\text{dynamic}} \propto 1/2 \times \text{Capacitive load} \times \text{Voltage}^2 \times \text{Frequency switched}$$

Given that heat must be dissipated from a chip that is about 1.5 cm on a side, we have reached the limit of what can be cooled by air.

Cooling Systems

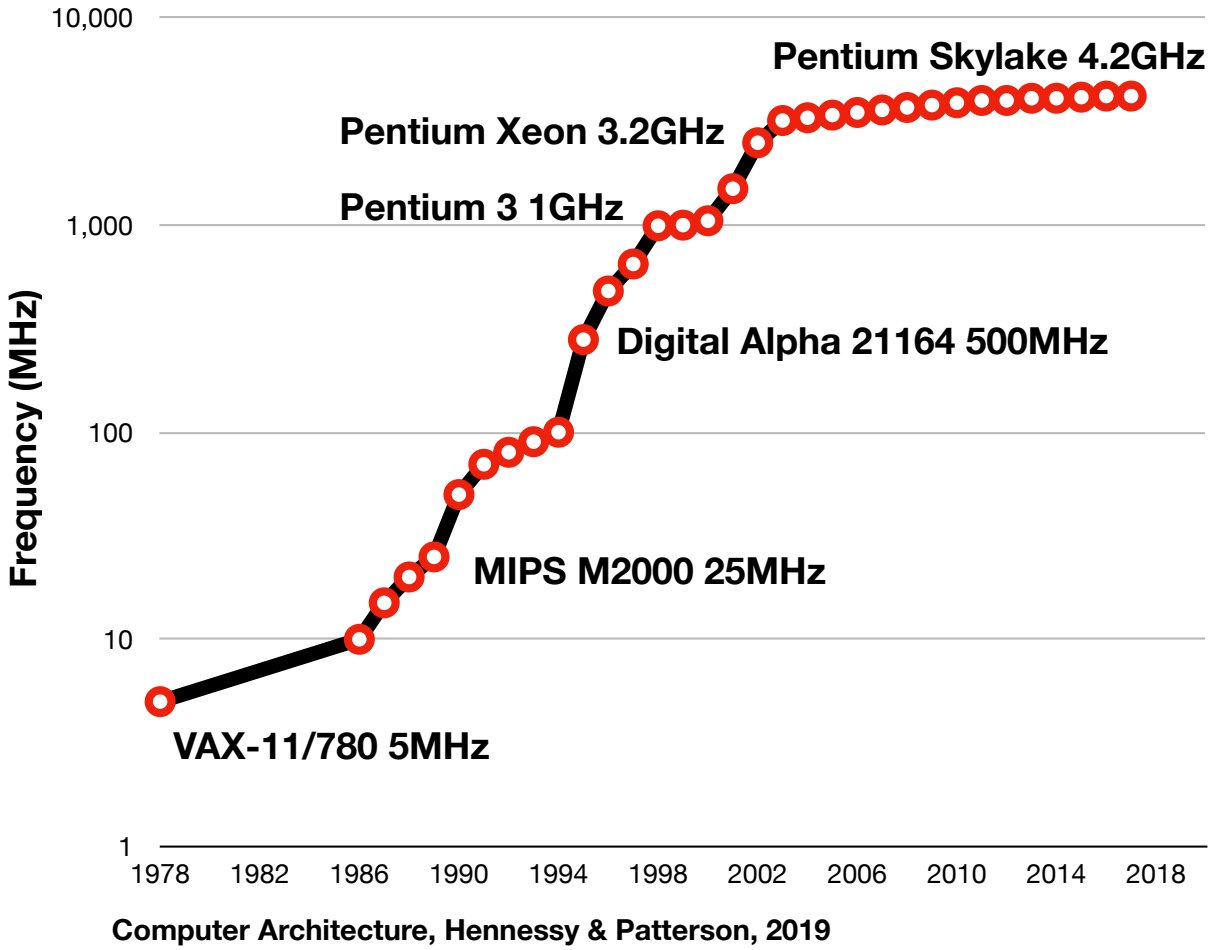


Chip

Cooler

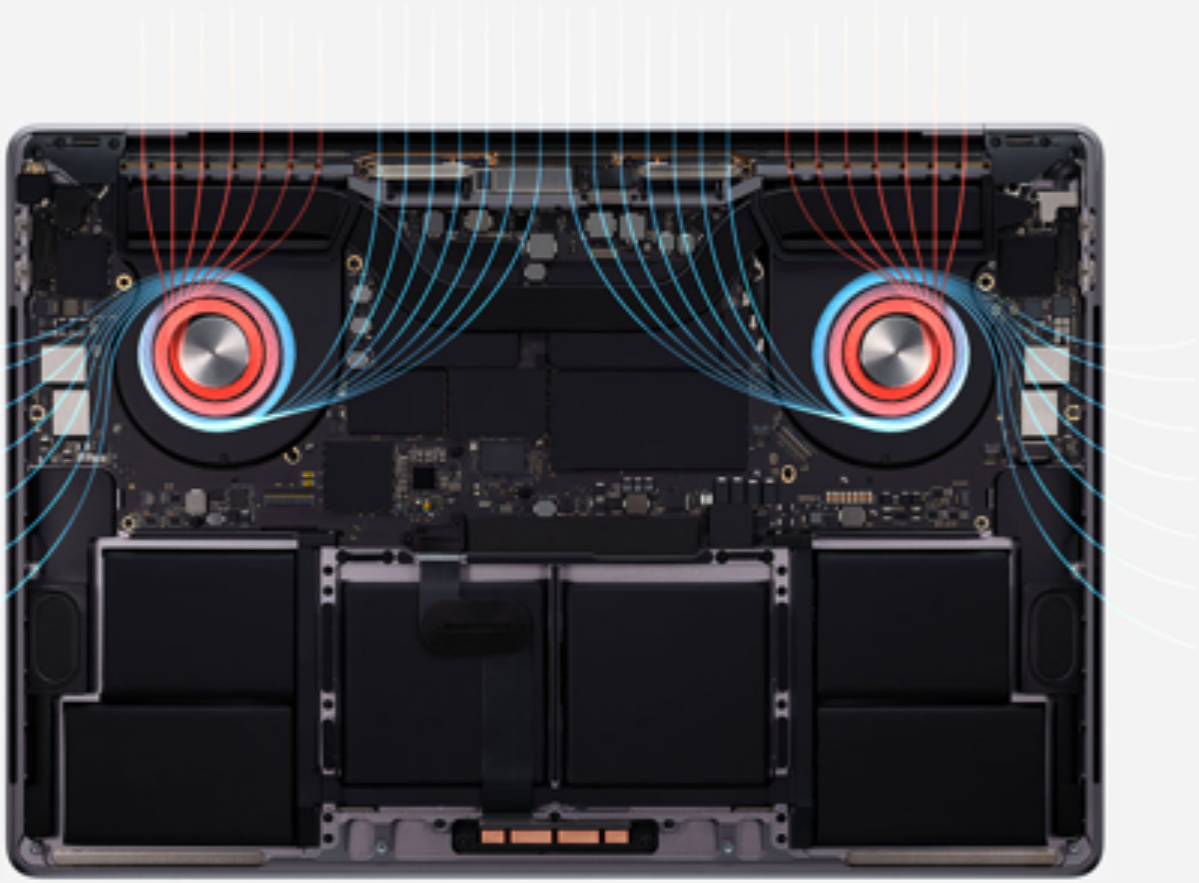
NOEL

Clocking Faster



Computer Architecture, Hennessy & Patterson, 2019

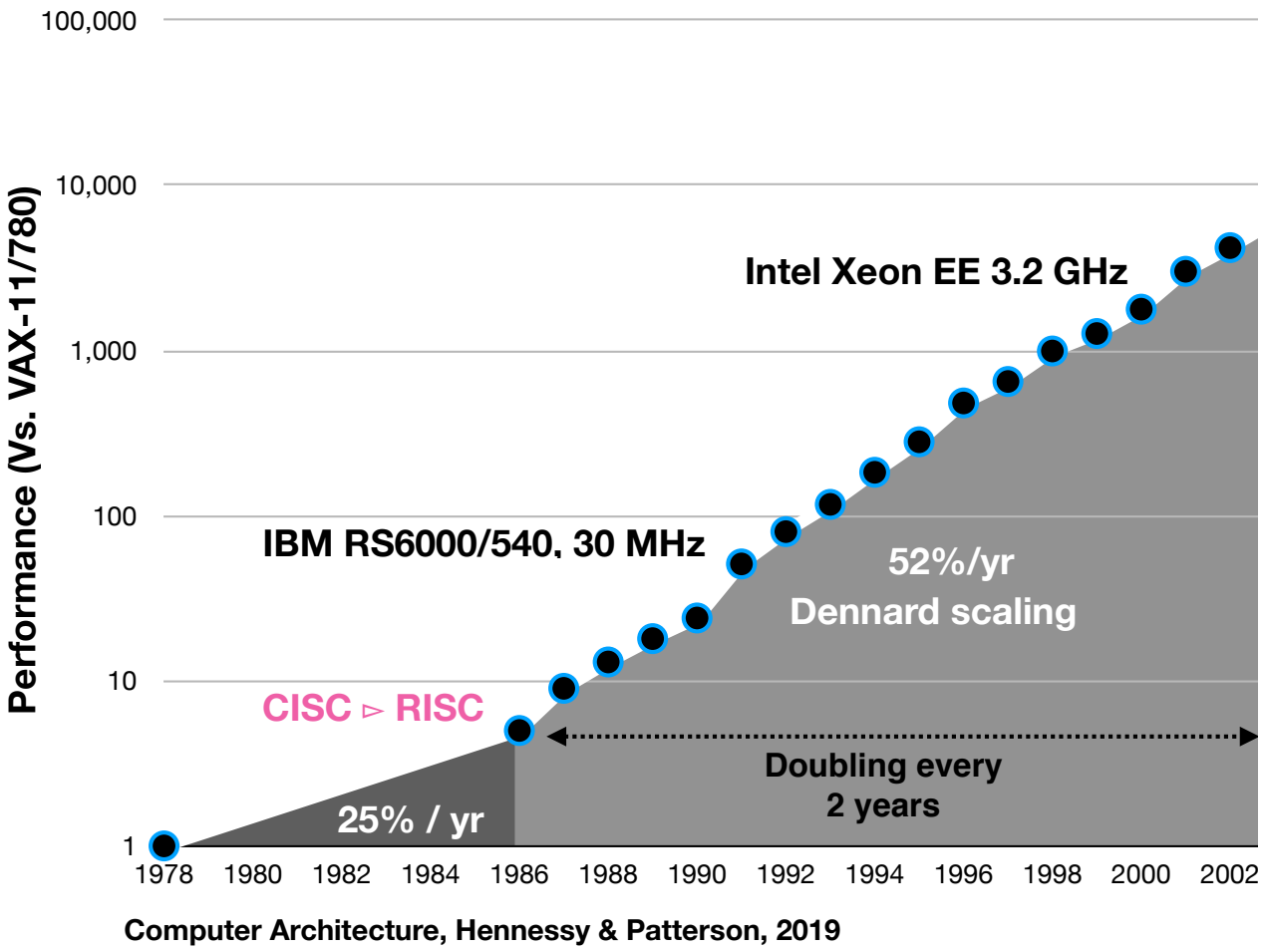
- Clock frequency growth has to slow down since we can't reduce voltage or increase power per chip.
- In correspondence to the period of slow performance improvement range, **since 2003**, clock speed has stagnated.



Mac Pro 16

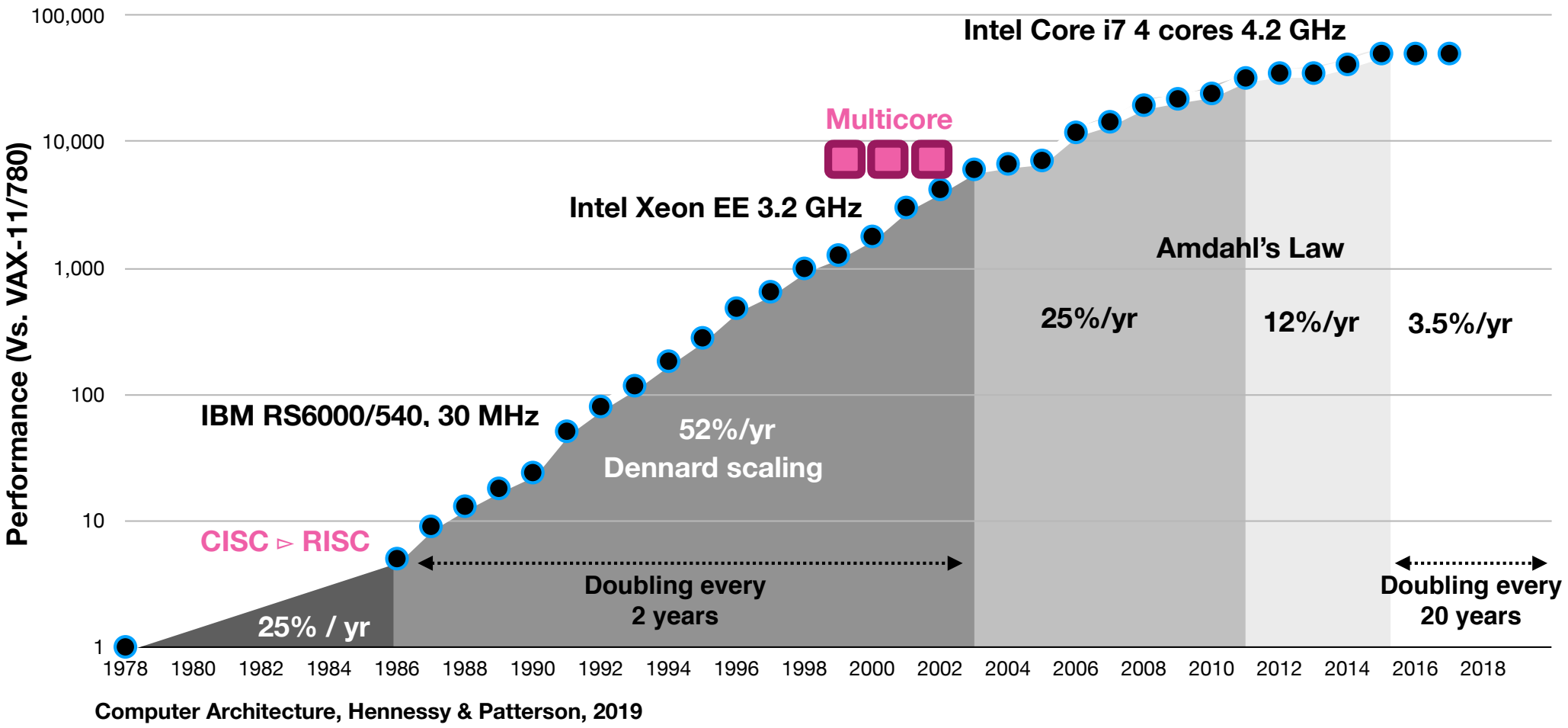
More advanced thermal architecture enables faster processing. The thermal architecture in MacBook Pro has been completely redesigned, featuring larger impellers with improved fan blades for optimal airflow and more heat-dispersing fins for more effective cooling. The resulting gain in cooling capacity allows MacBook Pro to deliver up to 12 watts more maximum sustained power.

Computing Performance



Computer Architecture, Hennessy & Patterson, 2019

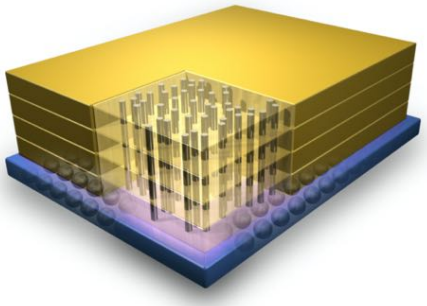
Computing Performance



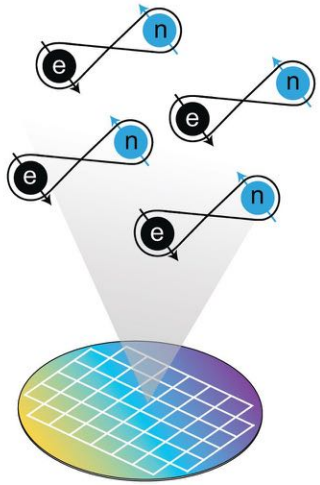
Computer Architecture, Hennessy & Patterson, 2019

NOEL

Emerging Computing Architectures

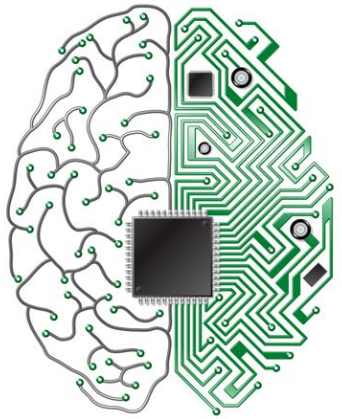


3D integrated circuits



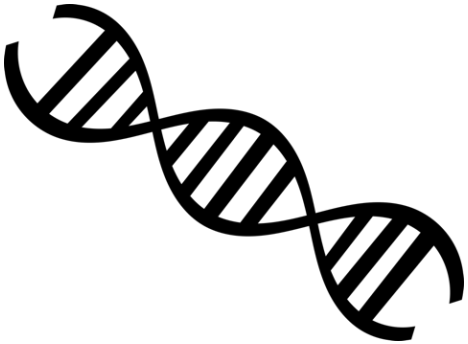
quantum computers

GPU computing

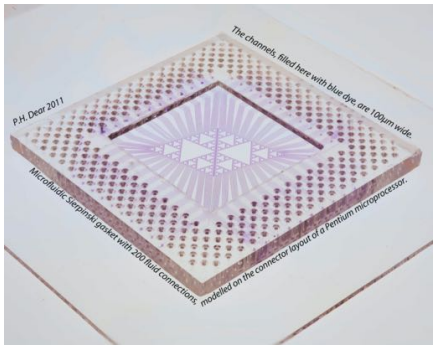


neuromorphic computing

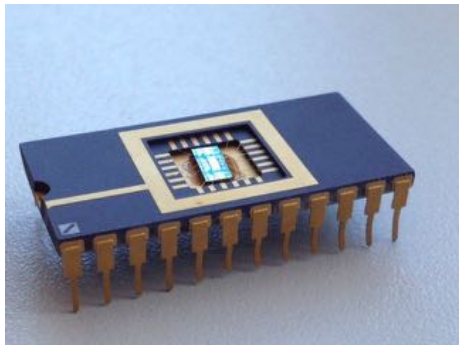
molecular computing



DNA computing



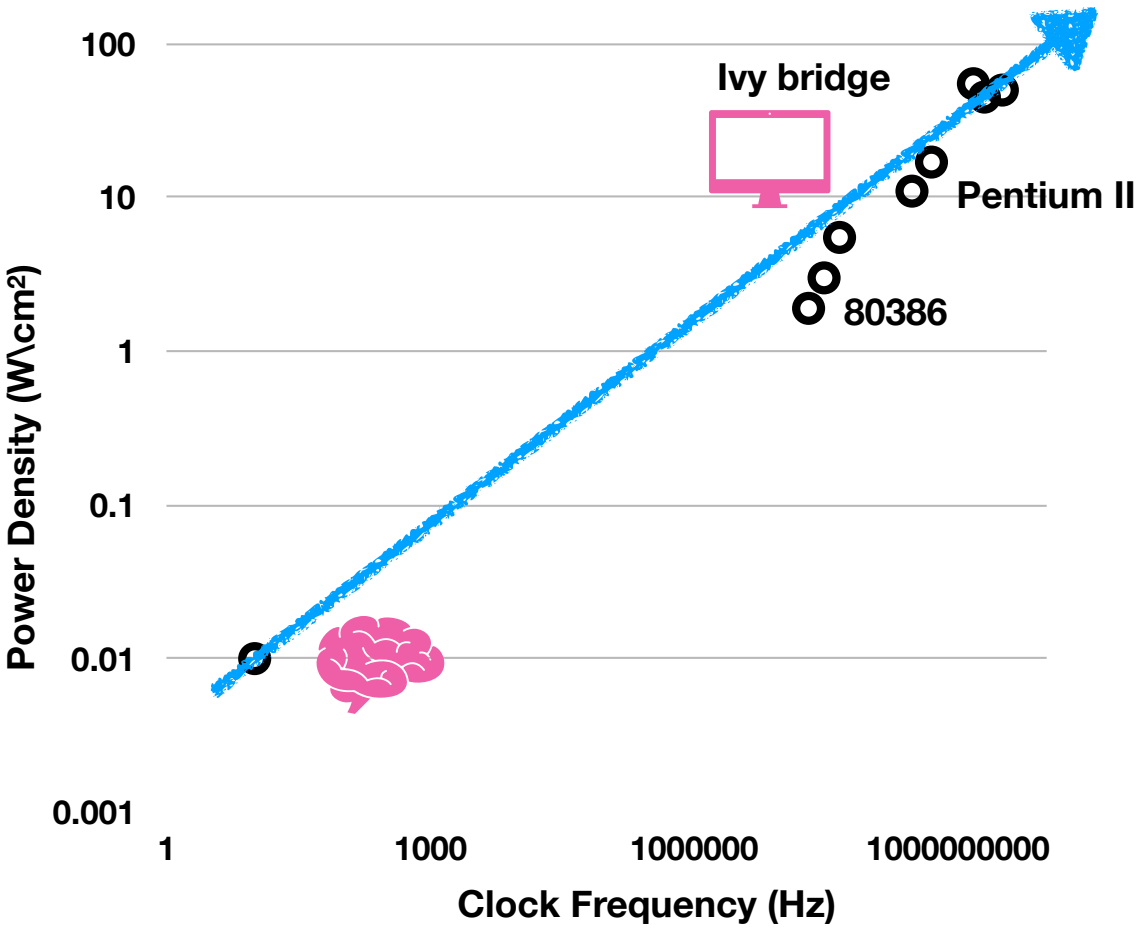
programmable microfluidics



new transistor technology

NOEL

Power - Clock



Merolla et. al. Science, 2014

It is postulated that the human **brain** operates at 1 exaFLOP.
…The Blue Gene operates at the PetaFLOP scale

- Task, navigational, and social intelligence
- < million neurons
- < milliwatt
- Ionic device physics
- Bulk mobility 10 million times lower than that of electronics.
- Many orders of magnitude more task-competent and power-efficient than current neuronal simulations or autonomous robots.

- DARPA Grand Challenge robotic cars
- Densely GPS-defined path
- Carrying over a kilowatt of sensing and computing power

NOEL

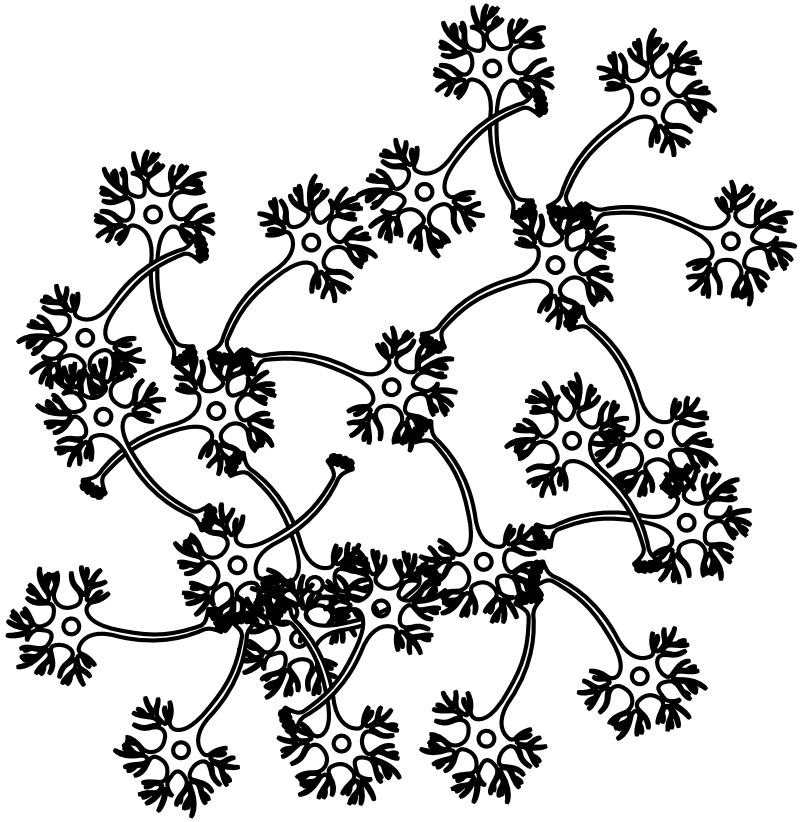
RESEARCH *Science* **360**, 1124–1126 (2018)

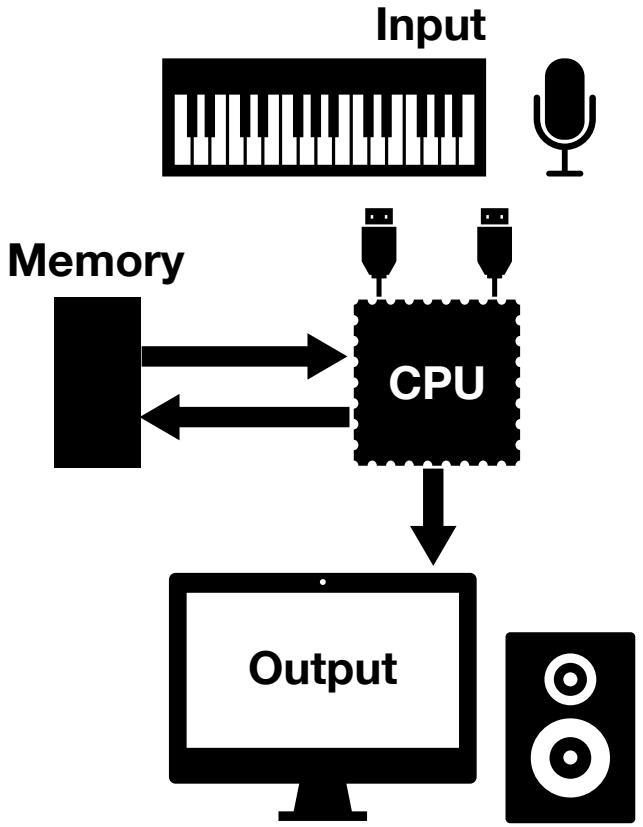
EVOLUTIONARY COGNITION

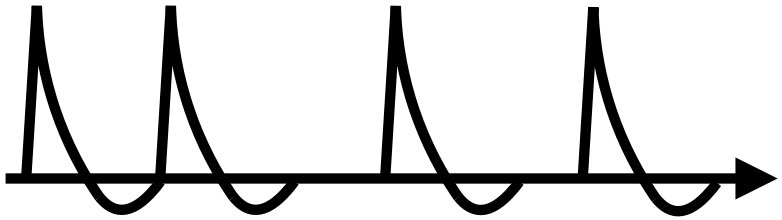
Numerical ordering of zero in honey bees

Scarlett R. Howard[1], Aurore Avarguès-Weber[2*], Jair E. Garcia[1*], Andrew D. Greentree[3], Adrian G. Dyer[1,4†]

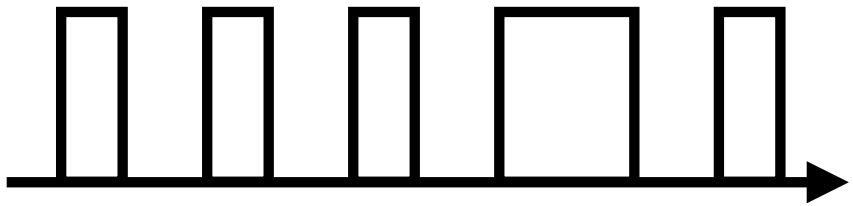
Some vertebrates demonstrate complex numerosity concepts—including addition, sequential ordering of numbers, or even the concept of zero—but whether an insect can develop an understanding for such concepts remains unknown. We trained individual honey bees to the numerical concepts of "greater than" or "less than" using stimuli containing one to six elemental features. Bees could subsequently extrapolate the concept of less than to order zero numerosity at the lower end of the numerical continuum. Bees demonstrated an understanding that parallels animals such as the African grey parrot, nonhuman primates, and even preschool children.







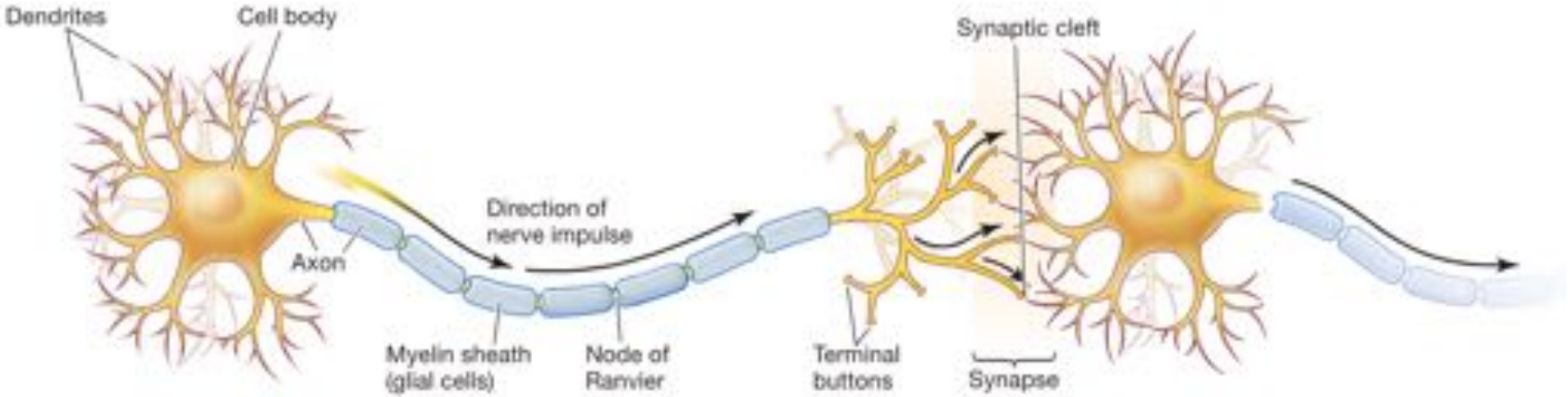
Neuromorphic Architecture



Von-Neuman Architecture

NOEL

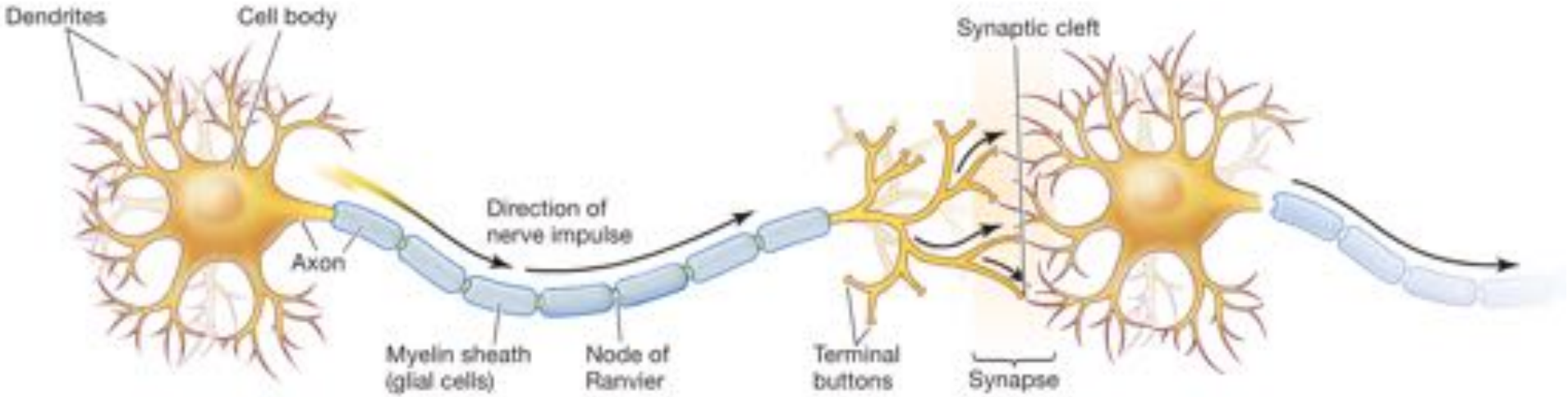
Power consumptions of elementary brain operations



1 fJ / ion channel

Ion channel opens and closes continually, passing a current of 0 / 5 pA. A pulse of voltage or ligand activation may cause channel's open probability to rise from ~0 to 0.2, decaying exponentially thereafter (at the population level) with a time constant of about 10 ms. Thus, the ion channel conducts an average of 1 pA across 0.1 V (~V_m) and consumes 1 fJ.

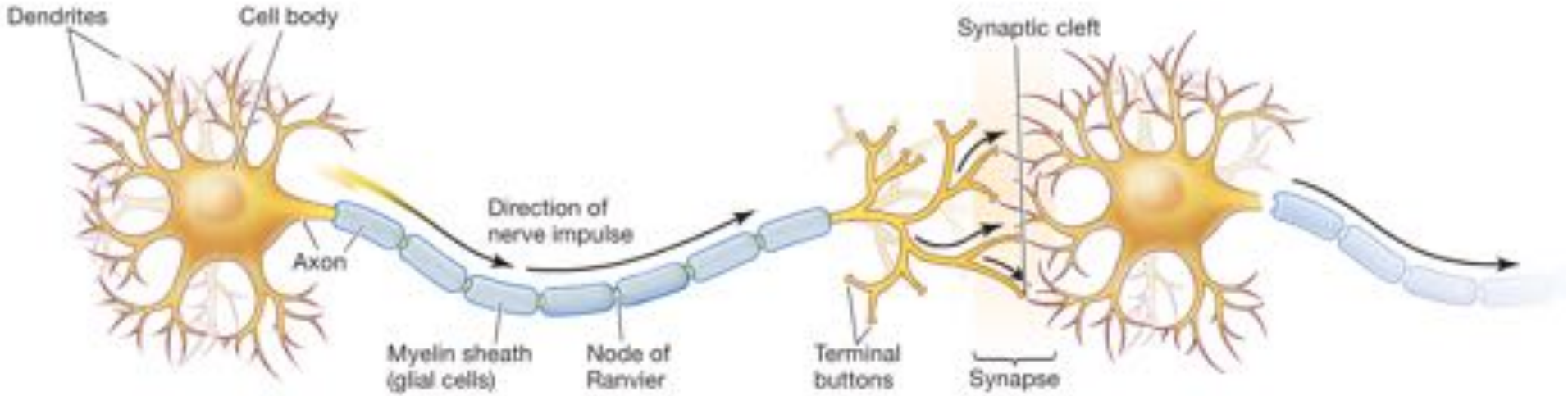
Power consumptions of elementary brain operations



1 fJ / ion channel
20 fJ / op

When communicating a spike, only 20 ion channels are opened per synapse. As spikes arrive infrequently (1 spike/s/synapse), and only 100 of the 10^4 synapses a neuron receives are active at any time, total passing current is 2 nA with a power consumption of 20 fJ/op.

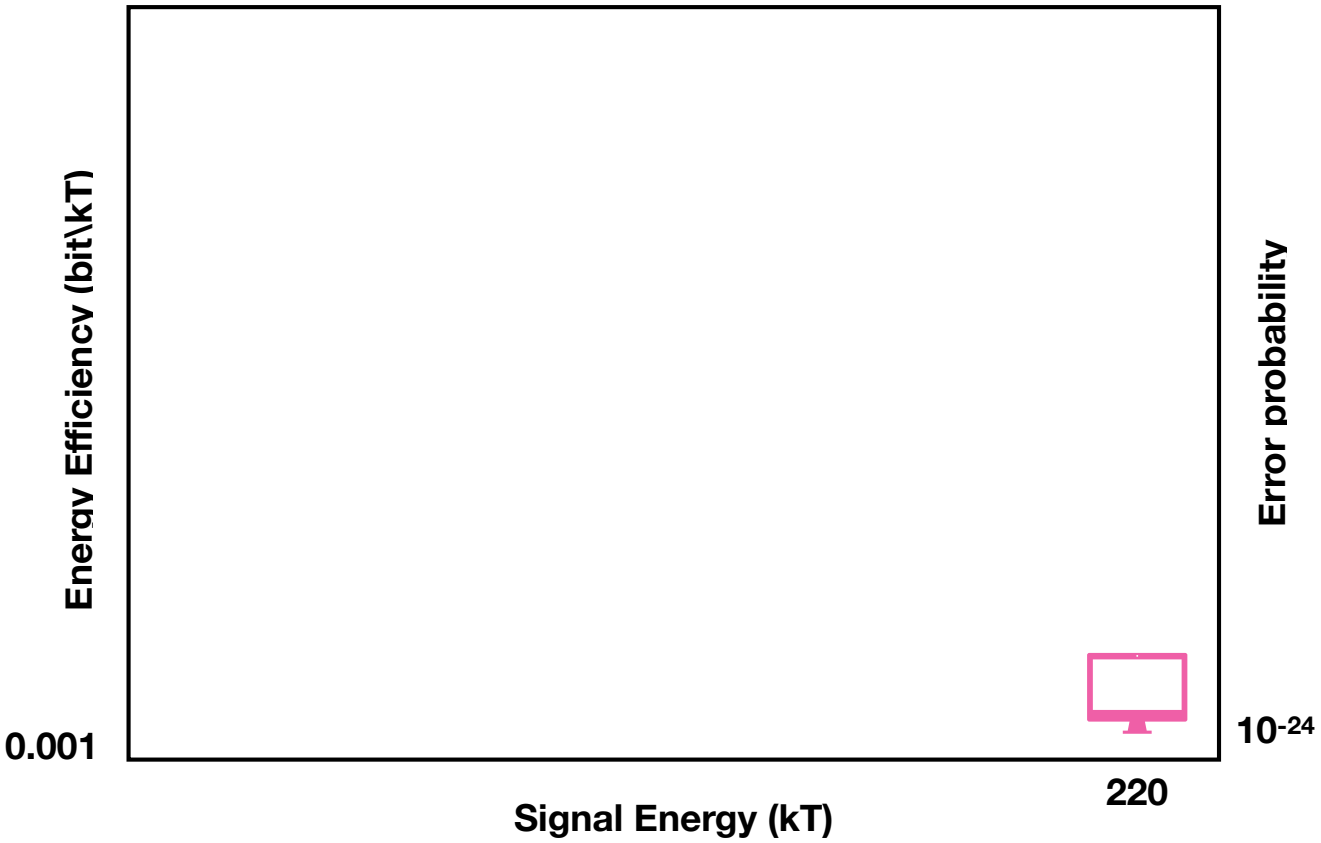
Power consumptions of elementary brain operations



1 fJ / ion channel
20 fJ / op
20 W / brain

To convey spikes from each of the brain's 10^{15} synapses once per second, 20W are therefore sufficient.

Energy - Efficiency - Error rate

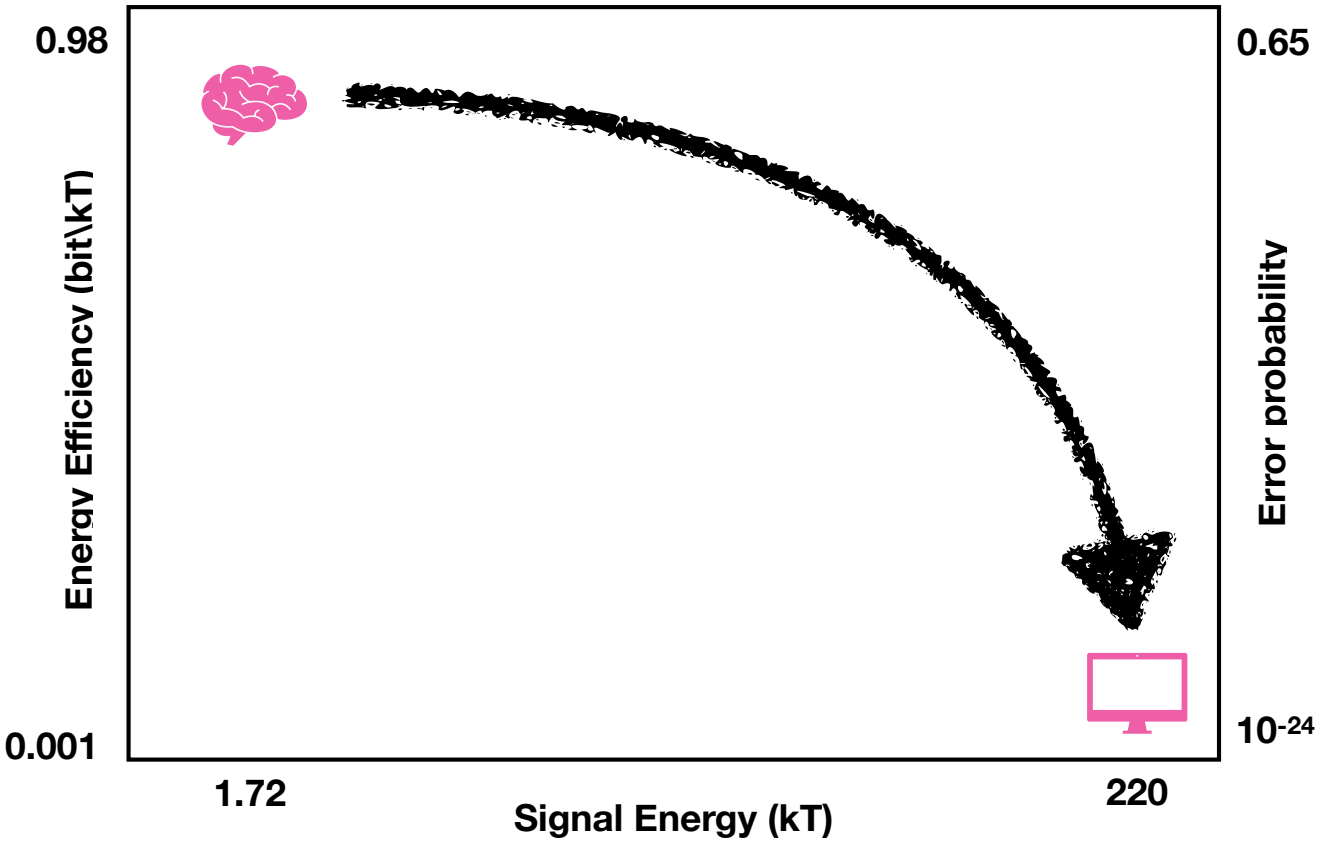


- Each signal the channel conveys carries a certain number of information bits (b).
- b number grows logarithmically (b = ½$\log_2$ (1 + E/kT)) with the ratio of signal energy (E) to noise energy (kT, for thermal noise).
- If signal energy decreases from 15kT to 3kT, number of bits drops from 2 to 1 and energy efficiency (b/(E + kT)) doubles, increasing from ⅛ to ¼ bits per kT.

A Neuromorph's Prospectus Kwabena Boahen, IEEE, 2018

A digital computer's binary signals are corrupted < 1/10 days, even though each of its billion circuits performs a billion operations per second. Such a low error rate ($p_{err} = 10^{-24}$), enforce signal energy E of > 220kT.

Energy - Efficiency - Error rate



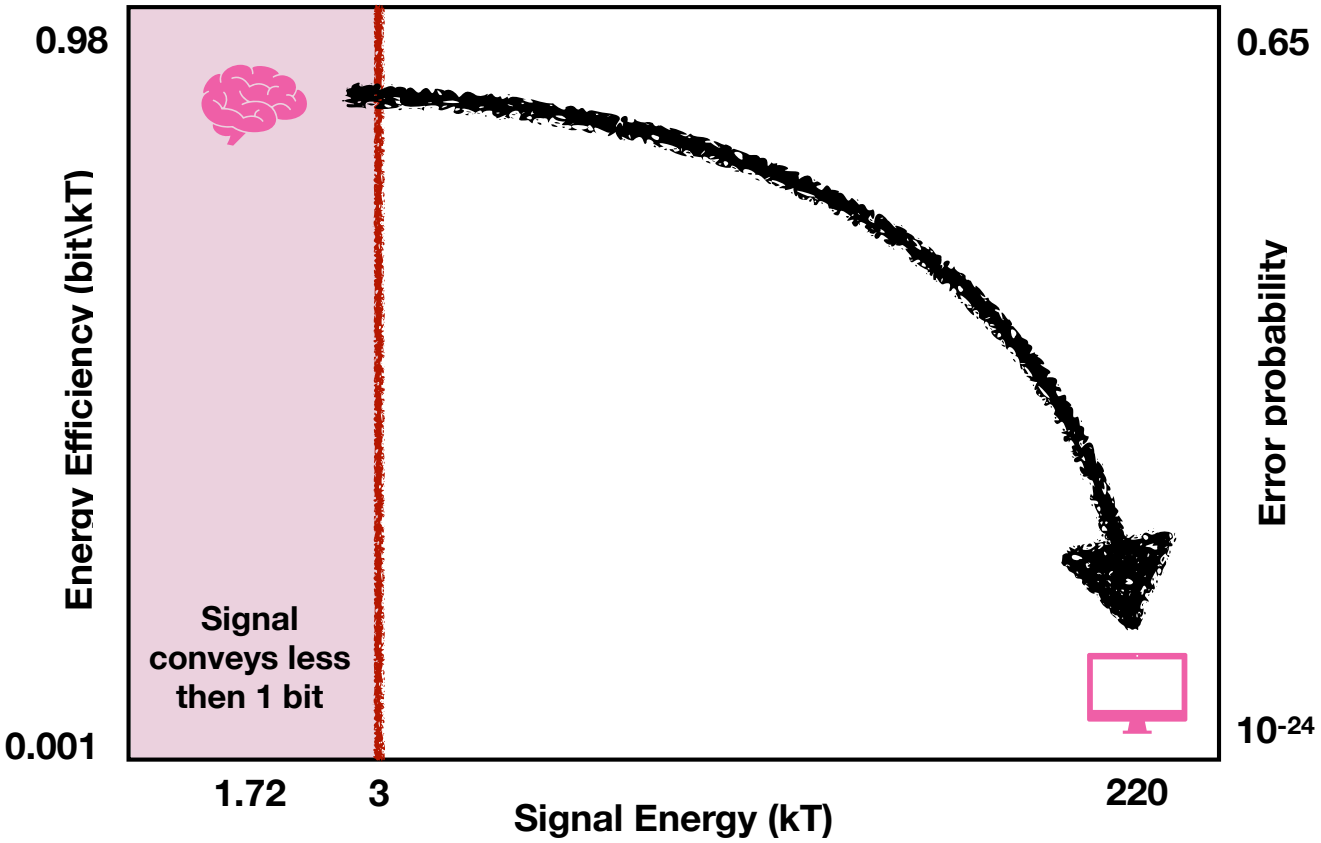
- Each signal the channel conveys carries a certain number of information bits (b).
- b number grows logarithmically (b = ½$\log_2$ (1 + E/kT)) with the ratio of signal energy (E) to noise energy (kT, for thermal noise).
- If signal energy decreases from 15kT to 3kT, number of bits drops from 2 to 1 and energy efficiency (b/(E + kT)) doubles, increasing from ⅛ to ¼ bits per kT.

A Neuromorph's Prospectus Kwabena Boahen, IEEE, 2018

A digital computer's binary signals are corrupted < 1/10 days, even though each of its billion circuits performs a billion operations per second. Such a low error rate ($p_{err} = 10^{-24}$), enforce signal energy E of > 220kT.

NOEL

Energy - Efficiency - Error rate



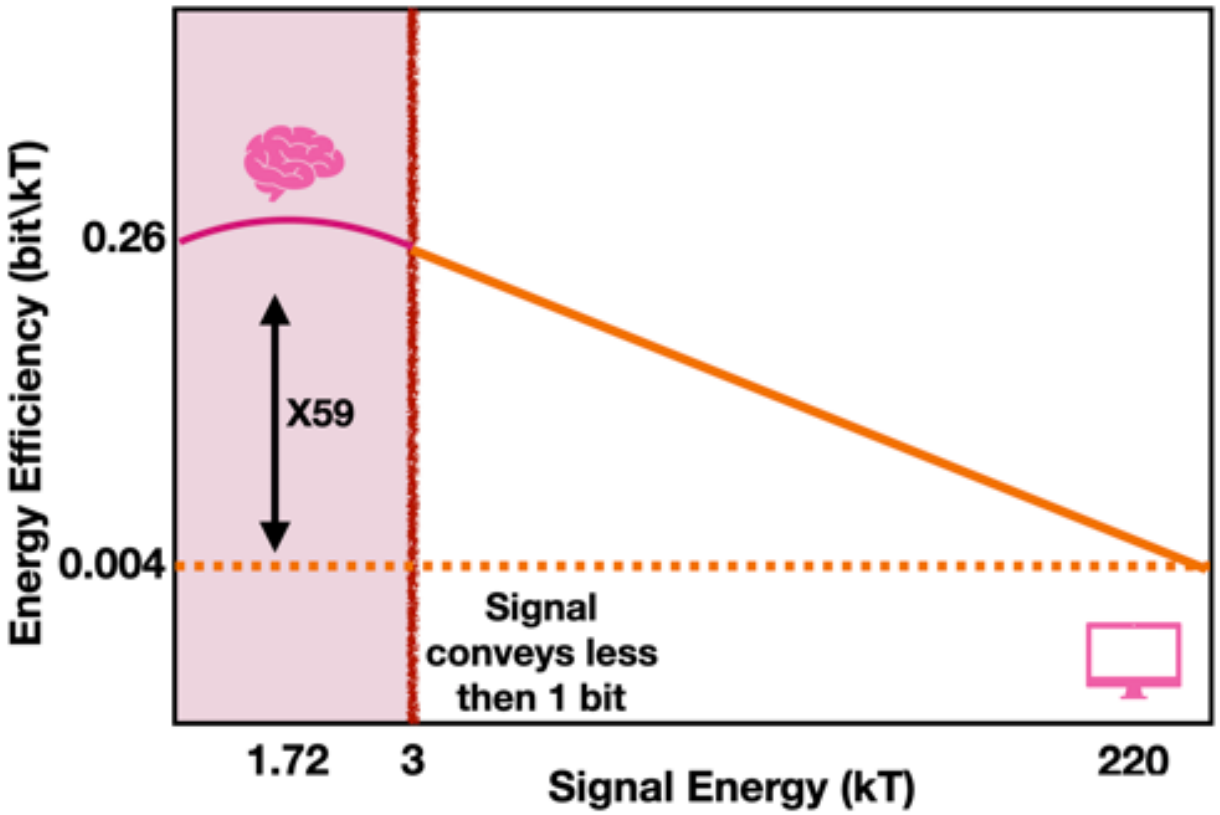
- Each signal the channel conveys carries a certain number of information bits (b).
- b number grows logarithmically (b = ½$\log_2$ (1 + E/kT)) with the ratio of signal energy (E) to noise energy (kT, for thermal noise).
- If signal energy decreases from 15kT to 3kT, number of bits drops from 2 to 1 and energy efficiency (b/(E + kT)) doubles, increasing from ⅛ to ¼ bits per kT.
- For E < 3kT, a signal conveys less than 1 bit (probabilistic) as noise frequently foils the signal (error probability). The brain operates in this regime.

A Neuromorph's Prospectus Kwabena Boahen, IEEE, 2018

A digital computer's binary signals are corrupted < 1/10 days, even though each of its billion circuits performs a billion operations per second. Such a low error rate ($p_{err} = 10^{-24}$), enforce signal energy E of > 220kT.

NOEL

Energy - Efficiency - Error rate

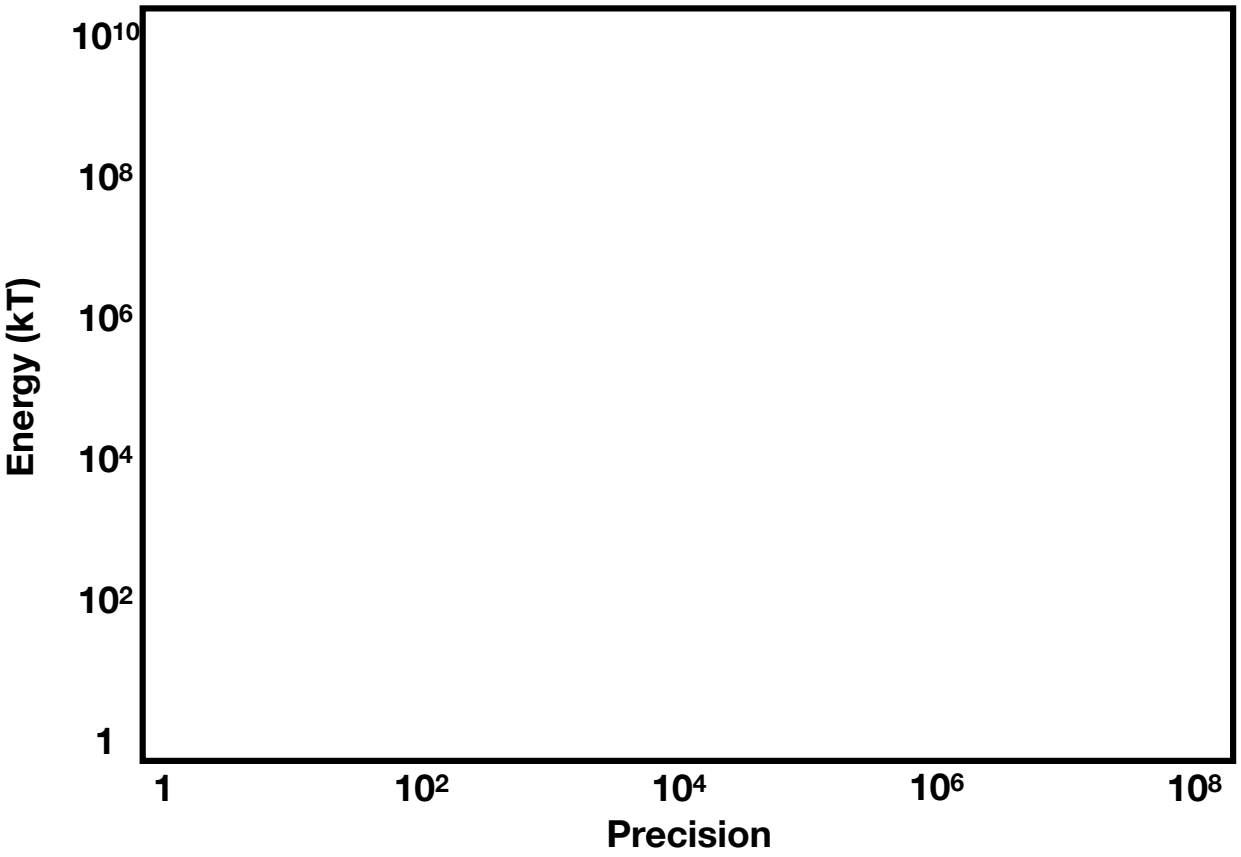


- Each signal the channel conveys carries a certain number of information bits (b).
- b number grows logarithmically (b = ½$\log_2$ (1 + E/kT)) with the ratio of signal energy (E) to noise energy (kT, for thermal noise).
- If signal energy decreases from 15kT to 3kT, number of bits drops from 2 to 1 and energy efficiency (b/(E + kT)) doubles, increasing from ⅛ to ¼ bits per kT.
- For E < 3kT, a signal conveys less than 1 bit (probabilistic) as noise frequently foils the signal (error probability). The brain operates in this regime.

A Neuromorph's Prospectus Kwabena Boahen, IEEE, 2018

A digital computer's binary signals are corrupted < 1/10 days, even though each of its billion circuits performs a billion operations per second. Such a low error rate ($p_{err} = 10^{-24}$), enforce signal energy E of > 220kT.

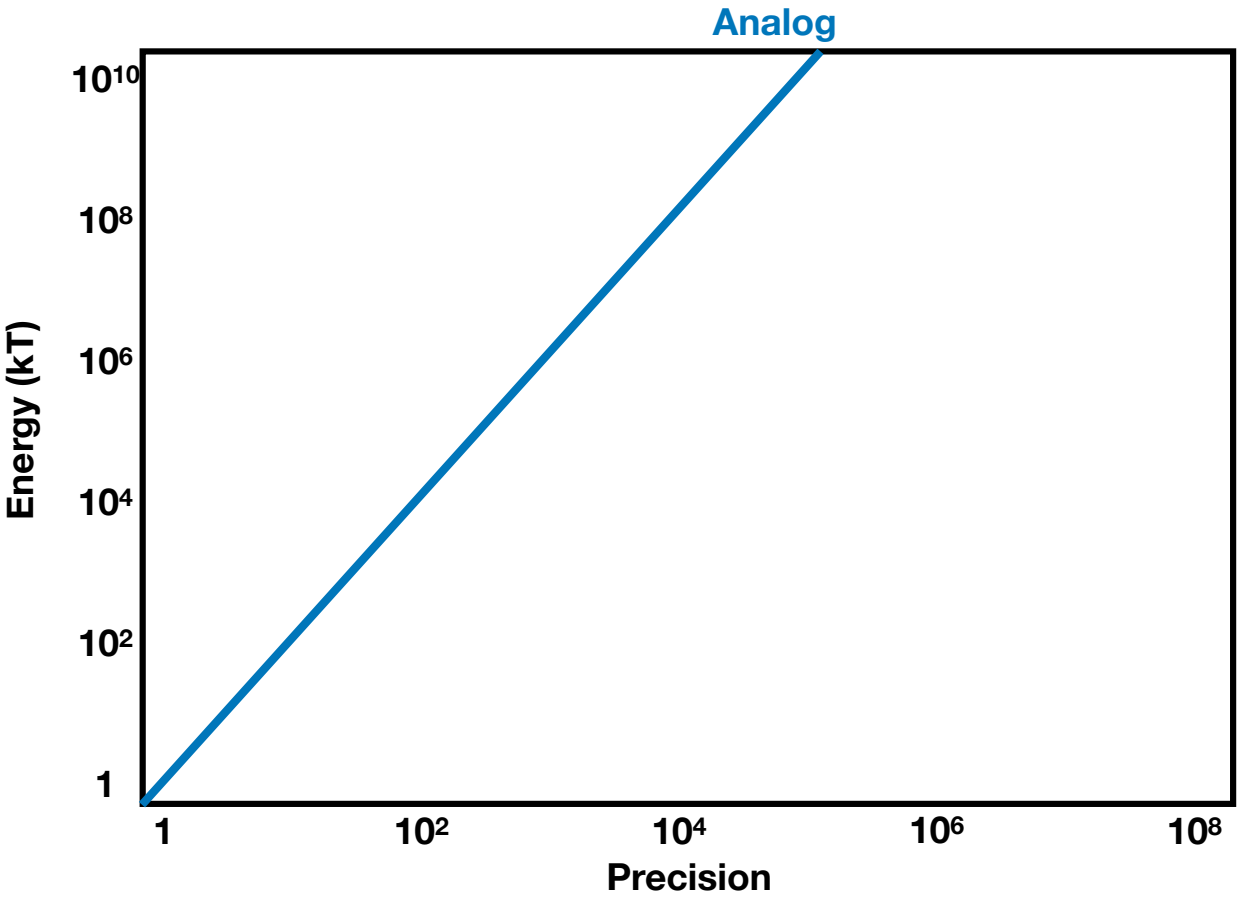
Where is Neuromorphic Computing is relevant?



A Neuromorph's Prospectus Kwabena Boahen, IEEE, 2018

NOEL

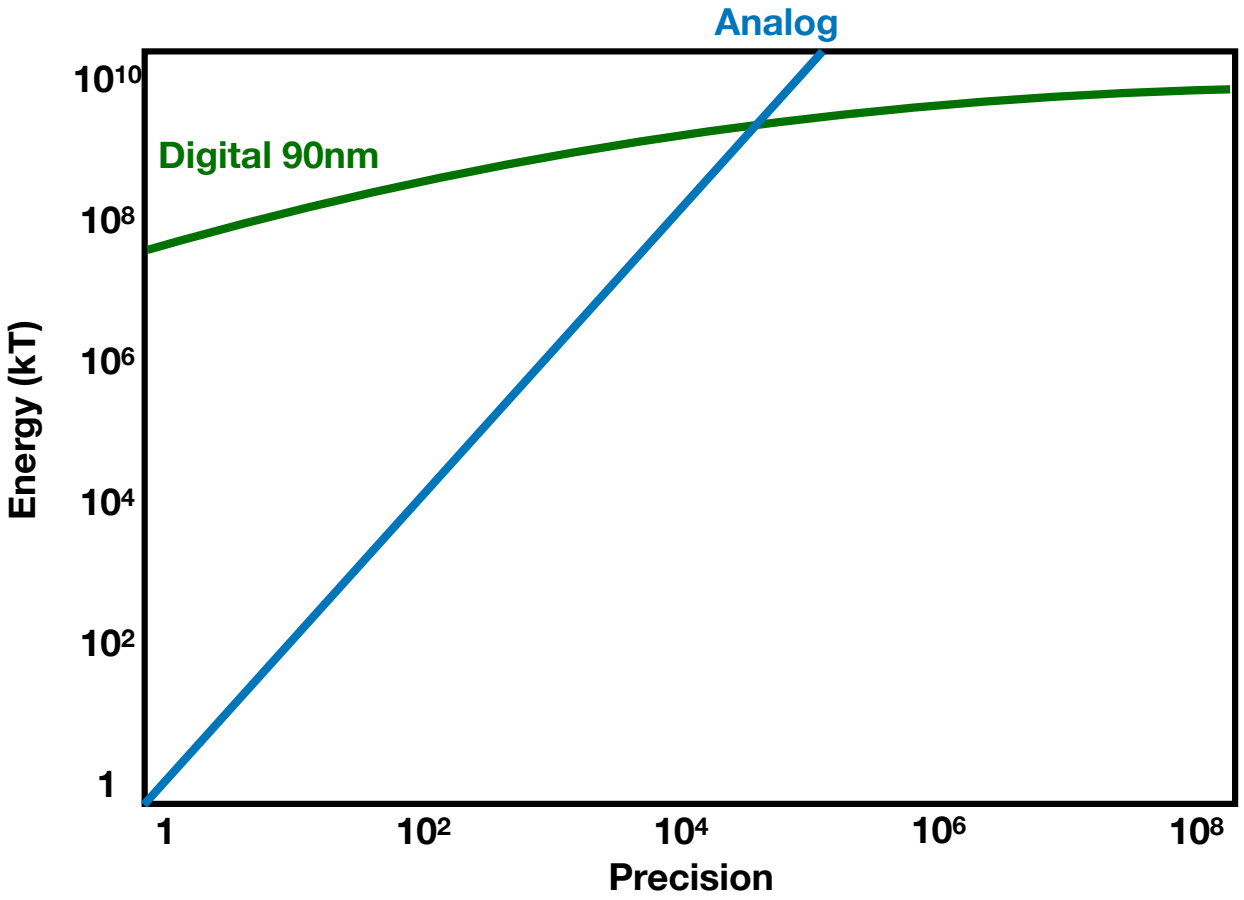
Where is Neuromorphic Computing is relevant?



A Neuromorph's Prospectus Kwabena Boahen, IEEE, 2018

- Energy consumed to generate an analog voltage signal scales quadratically with its amplitude.

Where is Neuromorphic Computing is relevant?

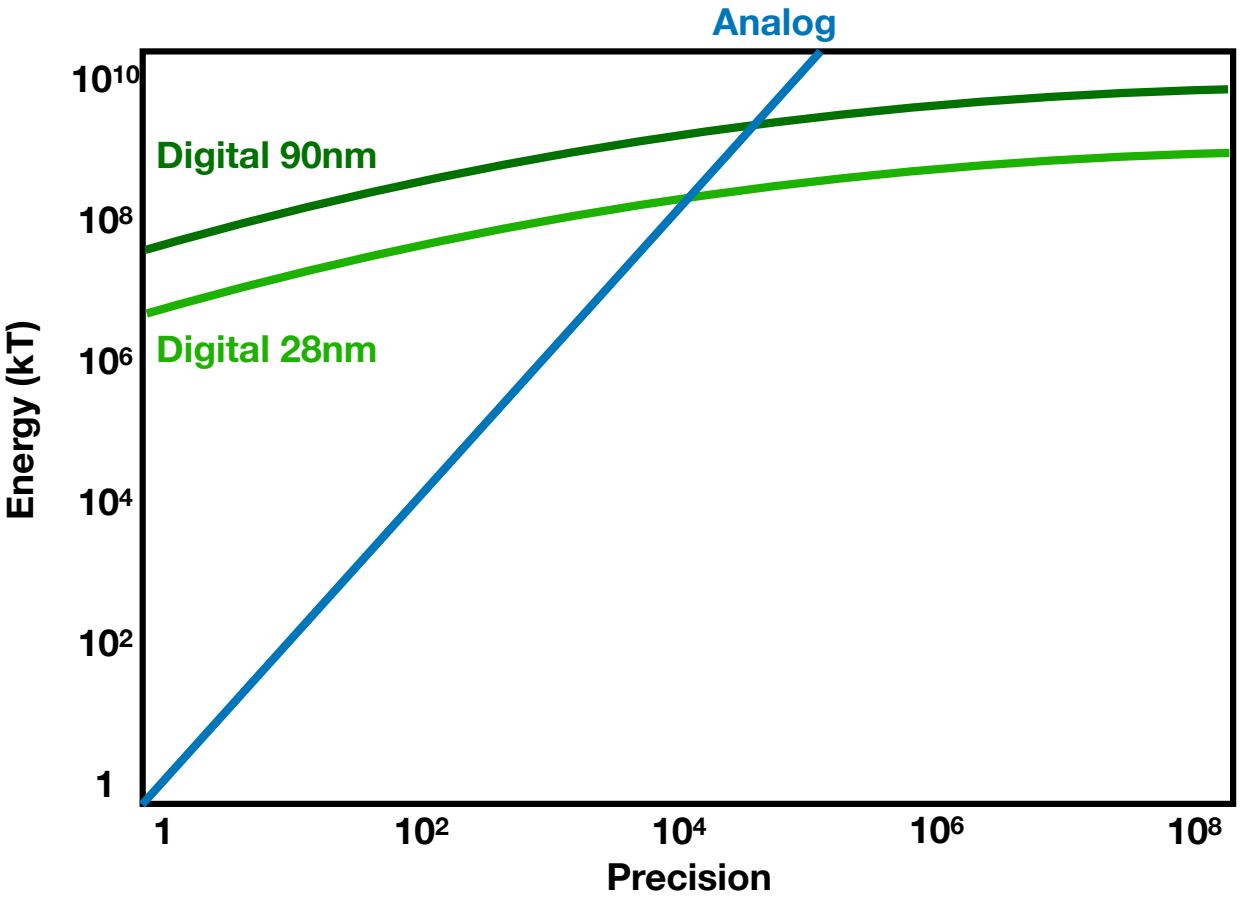


A Neuromorph's Prospectus Kwabena Boahen, IEEE, 2018

- Energy consumed to generate an analog voltage signal scales quadratically with its amplitude.
- Digital signals scales logarithmically

NOEL

Where is Neuromorphic Computing is relevant?

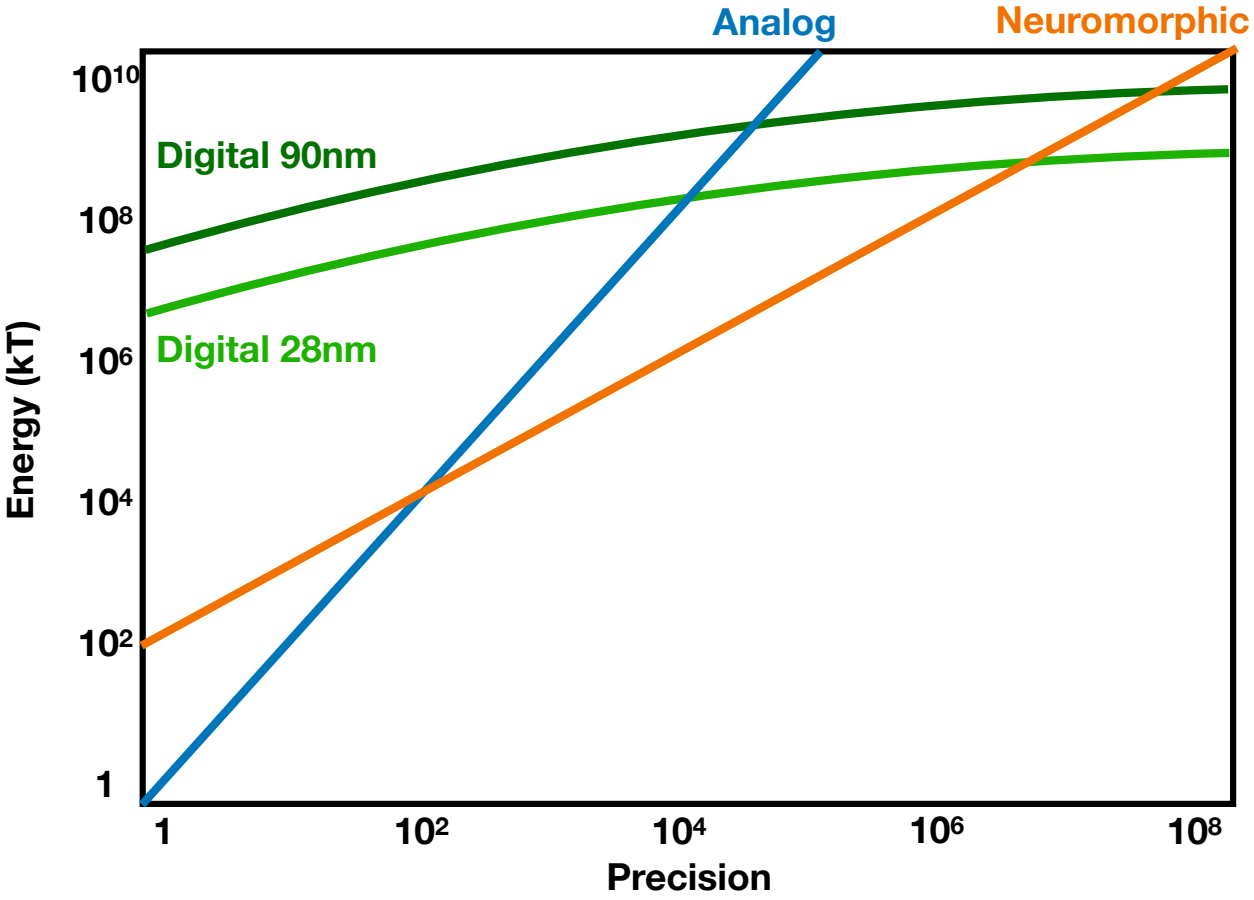


A Neuromorph's Prospectus Kwabena Boahen, IEEE, 2018

- Energy consumed to generate an analog voltage signal scales quadratically with its amplitude.
- Digital signals scales logarithmically
- The crossover point has migrated to the left over the years (with miniaturization) - favoring digital over analog computation for more and more applications

NOEL

Where is Neuromorphic Computing is relevant?

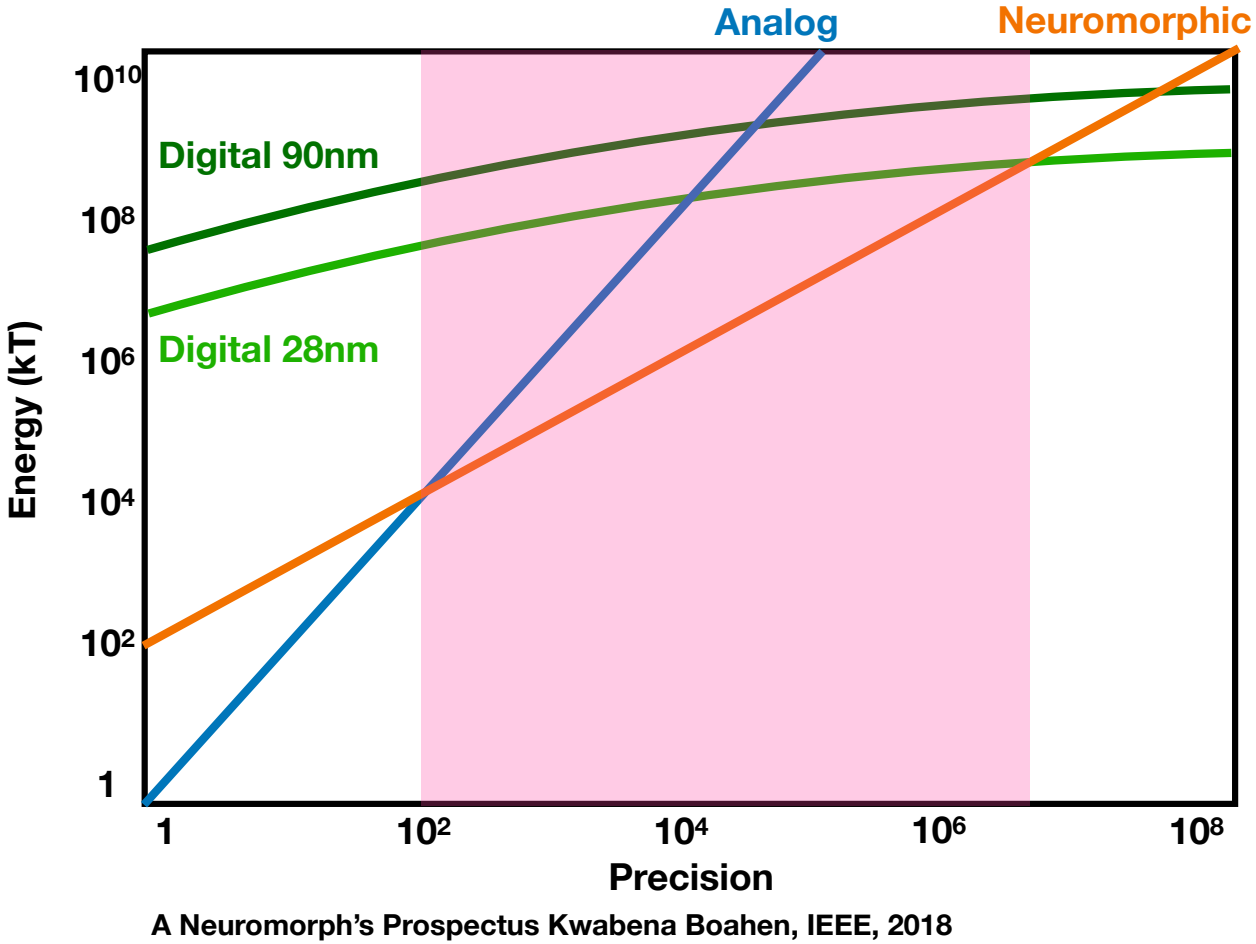


A Neuromorph's Prospectus Kwabena Boahen, IEEE, 2018

- Energy consumed to generate an analog voltage signal scales quadratically with its amplitude.
- Digital signals scales logarithmically
- The crossover point has migrated to the left over the years (with miniaturization) - favoring digital over analog computation for more and more applications
- Most neuromorphic architectures aim to **mix analog-digital design**

NOEL

Where is Neuromorphic Computing is relevant?



A Neuromorph's Prospectus Kwabena Boahen, IEEE, 2018

- Energy consumed to generate an analog voltage signal scales quadratically with its amplitude.
- Digital signals scales logarithmically
- The crossover point has migrated to the left over the years (with miniaturization) - favoring digital over analog computation for more and more applications
- Most neuromorphic architectures aim to **mix analog-digital design** to achieve best performance across five-decade precision range.

NOEL

IBM BlueGene Supercomputer

- Reconstruction and simulation were executed on supercomputers.
- BG can reach operating speeds in the petaFLOPS (10^{15}) range (Floating point operations per second)
- 65,536 cores
- 65 TB of RAM
- To simulate 0.01mm^3 of 31,000 neurons

NOEL

We need to redesign computers

NOEL

A silicon neuron

Misha Mahowald* & Rodney Douglas†‡

* Computation and Neural Systems Laboratory, California Institute of Technology, Pasadena, California 91125, USA
† MRC Anatomical Neuropharmacology Unit, University of Oxford, Oxford OX1 3TH, UK

BY combining neurophysiological principles with silicon engineering, we have produced an analog integrated circuit with the functional characteristics of real nerve cells. Because the physics underlying the conductivity of silicon devices and biological membranes is similar, the 'silicon neuron' is able to emulate efficiently the ion currents that cause nerve impulses and control the dynamics of their discharge. It operates in real-time and consumes little power, and many 'neurons' can be fabricated on a single silicon chip. The silicon neuron represents a step towards constructing artificial nervous systems that use more realistic principles of neural computation than do existing electronic neural networks.

Nature; Dec 19, 1991; 354, 6354

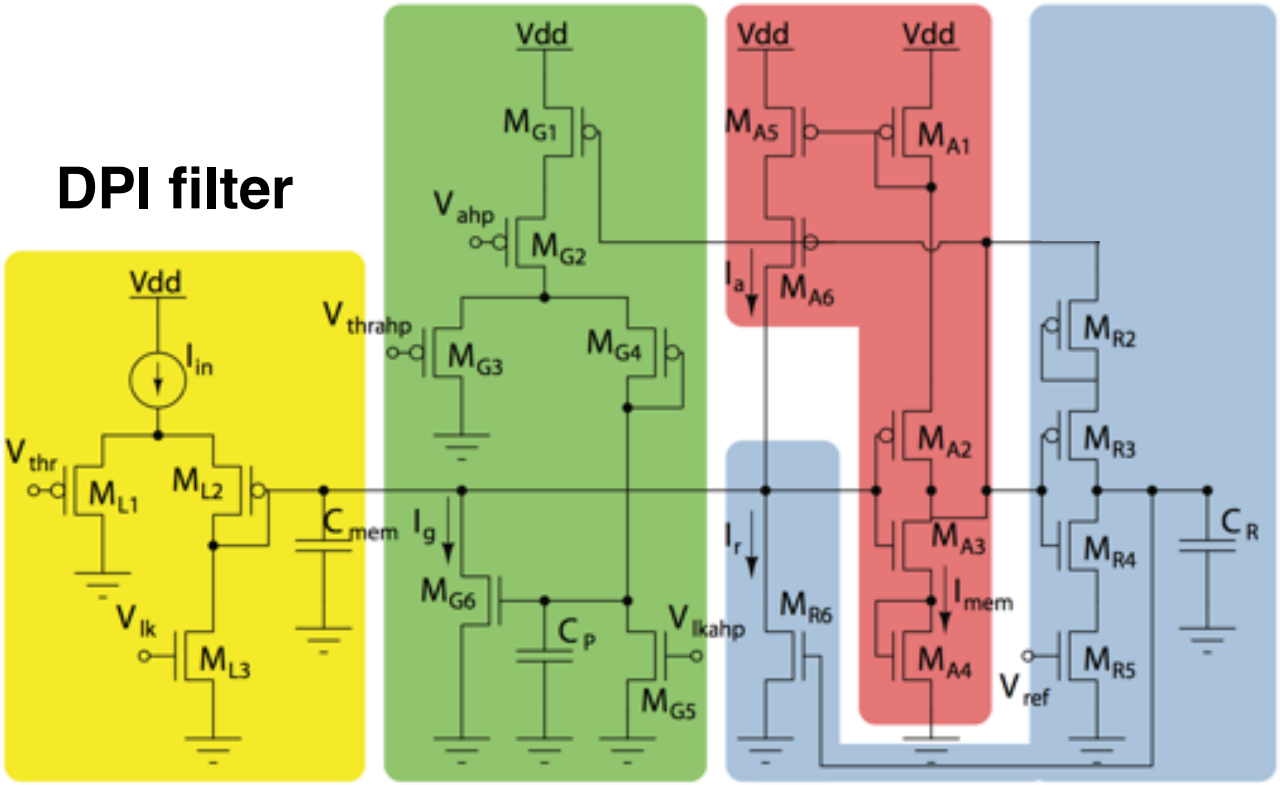
NOEL

spike-frequency adaptation

spike event generating amplifier

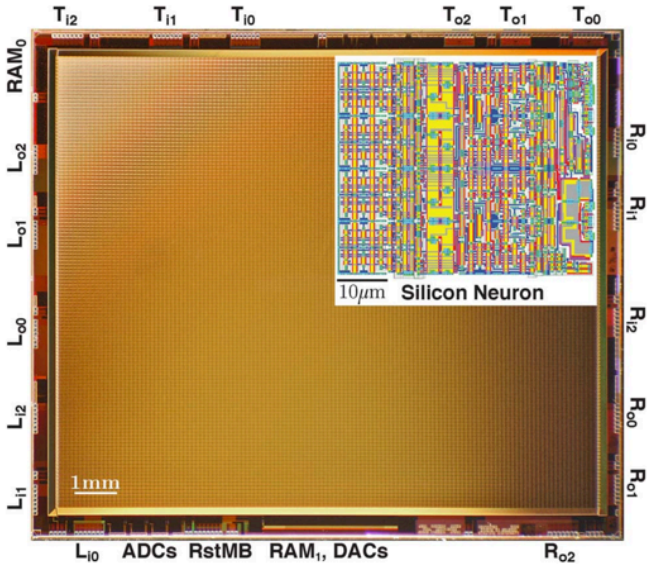
spike reset circuit, refractory period

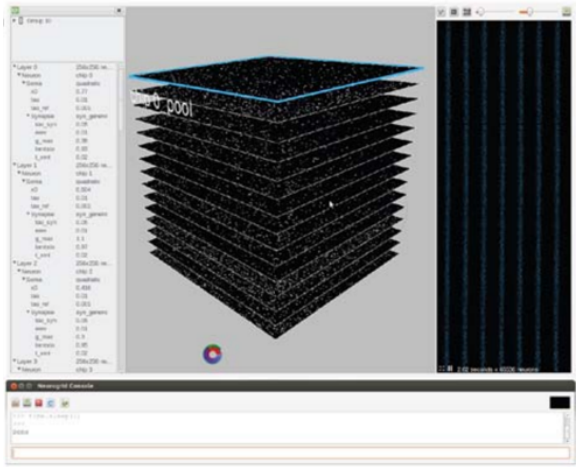
DPI filter

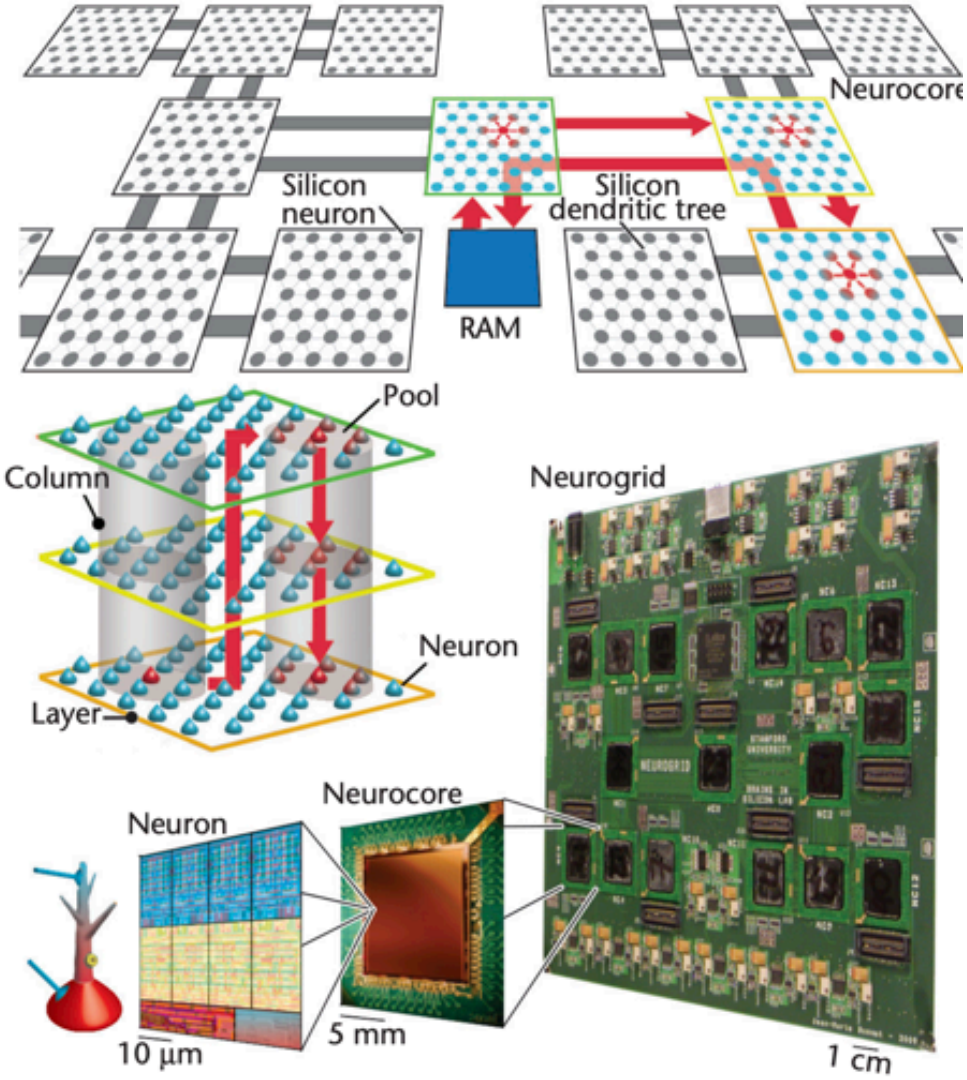


The DPI neuron circuit

Stanford's Neurogrid

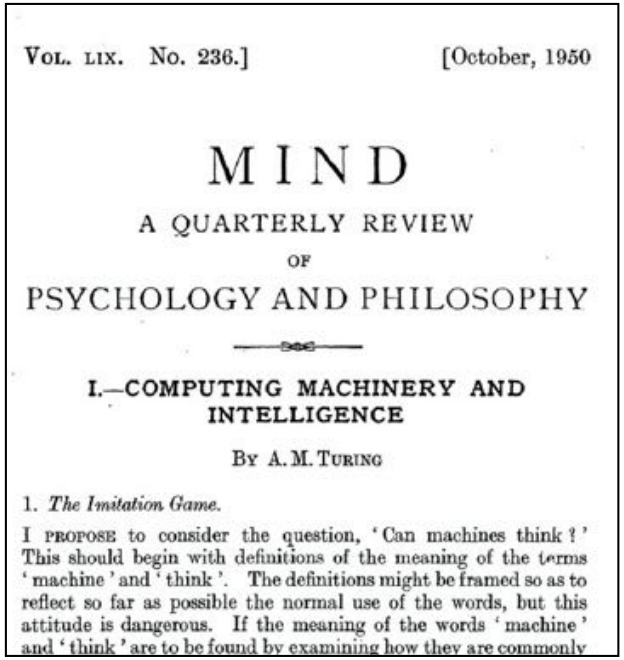






A long standing **dream**

A **long-standing dream** has been to **harness neuroscientific insights** to build a **versatile computer** that is: **efficient** in terms of **energy** and **space**;



VOL. LIX. No. 236.] [October, 1950

MIND

A QUARTERLY REVIEW

OF

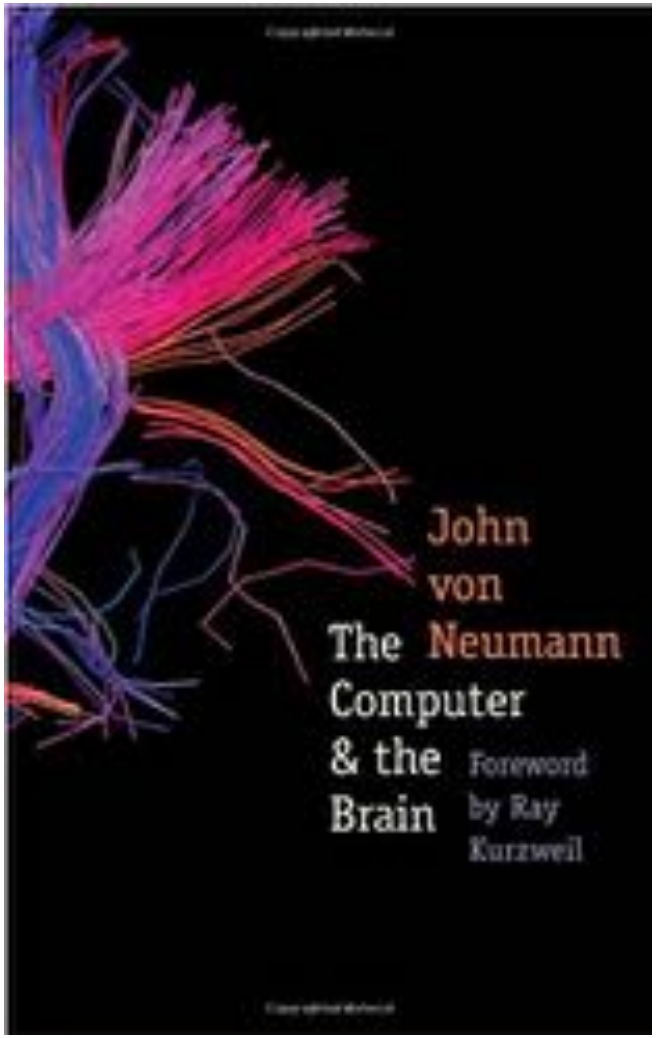
PSYCHOLOGY AND PHILOSOPHY

I.—COMPUTING MACHINERY AND INTELLIGENCE

BY A. M. TURING

1. *The Imitation Game.*

I PROPOSE to consider the question, 'Can machines think?' This should begin with definitions of the meaning of the terms 'machine' and 'think'. The definitions might be framed so as to reflect so far as possible the normal use of the words, but this attitude is dangerous. If the meaning of the words 'machine' and 'think' are to be found by examining how they are commonly

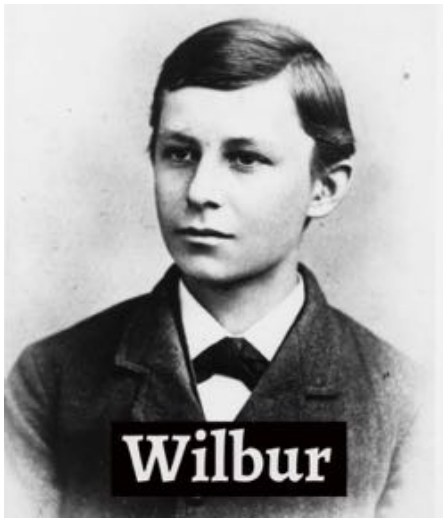


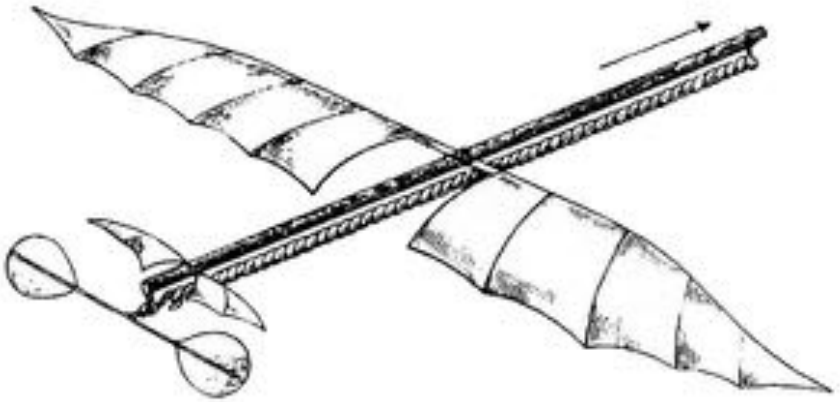

DREAMS

DREAMS

Road to First Flight

"When we **were children**, father brought home to us a small toy, actuated by a rubber spring, which would lift it self into the air… we could not understand that there was anything about a bird that would enable it to fly, that could not be built on a larger scale and used by man"
(Orville Wright, 1878).

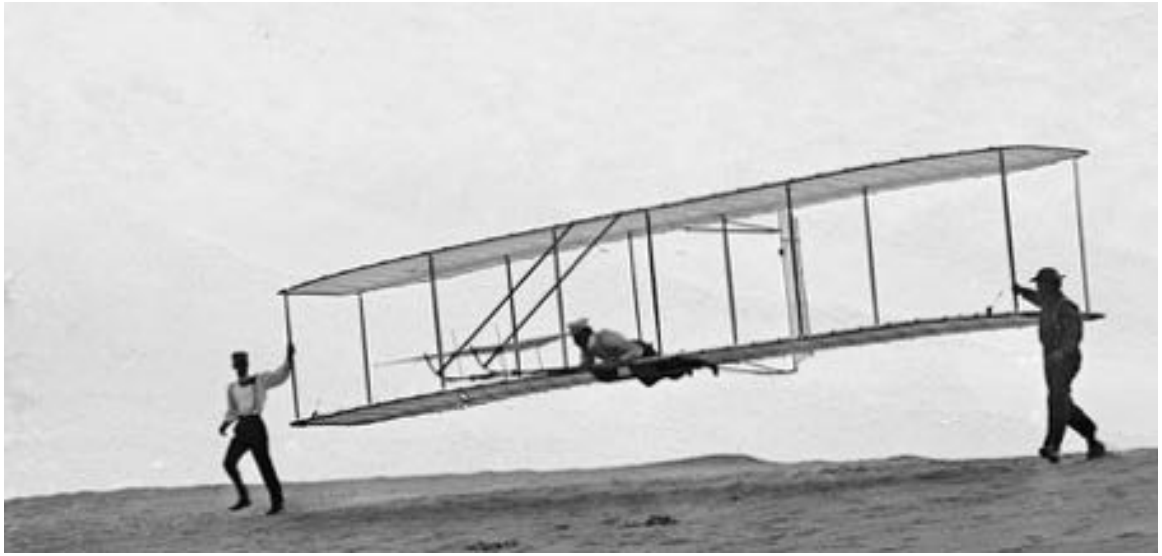




NOEL

DREAMS

1903 First Flight



Dayton Boys Solve Problem

Wilbur and Orville Wright Successfully Operate a Flying Machine in North Carolina--Description of Craft.

Dec 18, 1903 Dayton Herald

Bishop Milton Wright of this city has received a telegram from his sons Wilbur and Orville Wright, who are at Kitty Hawk, N. C., their fourth autumn, experimenting in gliding through the air on aeroplanes of their own make, and regulated by devices of their own invention, saying that they have had gratifying success with their true flying machine built by them this year

street: We have made four successful flights this morning, all against a 21-mile wind. We started from the level with engine powe ralone. Our average speed through the air was 31 miles. Our longest time in the air was 57 seconds. ORVILLE WRIGHT.'

By 'speed of 21 miles' is meant 10 miles an hour against a 21-mile wind.

A previous telegram from Wilbur,

NOEL

DREAMS

1903: First Flight

1903 - 1908: 12 Pilots

NOEL

DREAMS

1903: First Flight

1903 - 1908: 12 Pilots

1908 - 1912: 39 countries, 100's of airplanes, 1000's of pilots

DREAMS

1961 First human in space (USSR)

NOEL

DREAMS

1969 First human on the moon

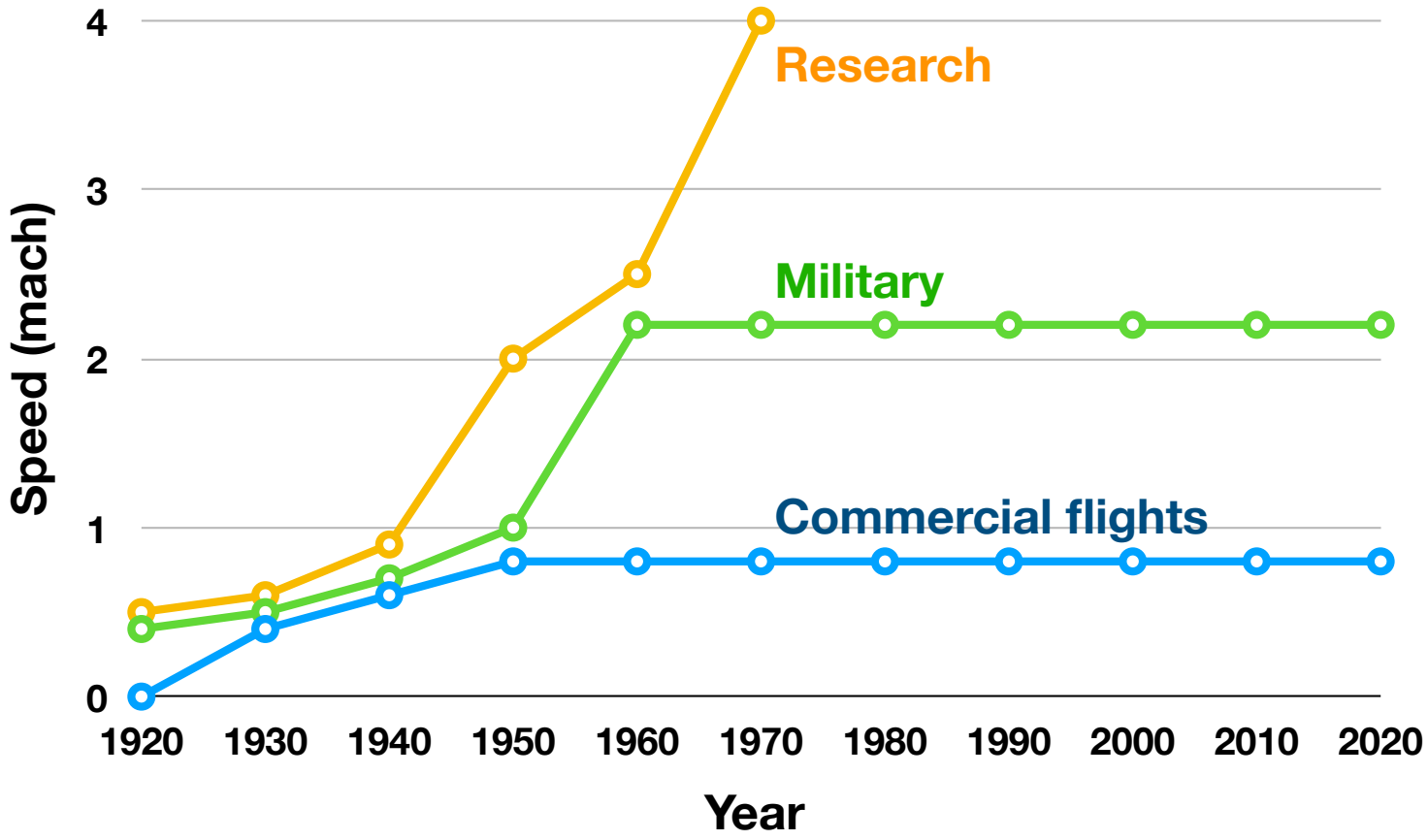
NOEL

DREAMS

1972 Last manned mission to the moon

NOEL

DREAMS


Speed (mach)
4
3
2
1
0
Research
Military
Commercial flights
1920
1930
1940
1950
1960
1970
1980
1990
2000
2010
2020
Year

NOEL

DREAMS

2018 First commercial travel to space

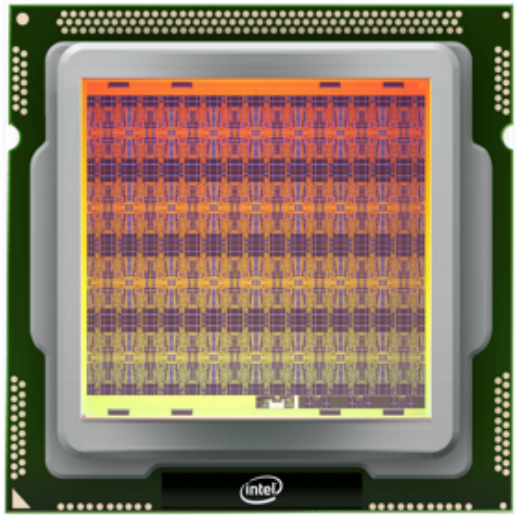
NOEL

DREAMS

NOEL

DREAMS

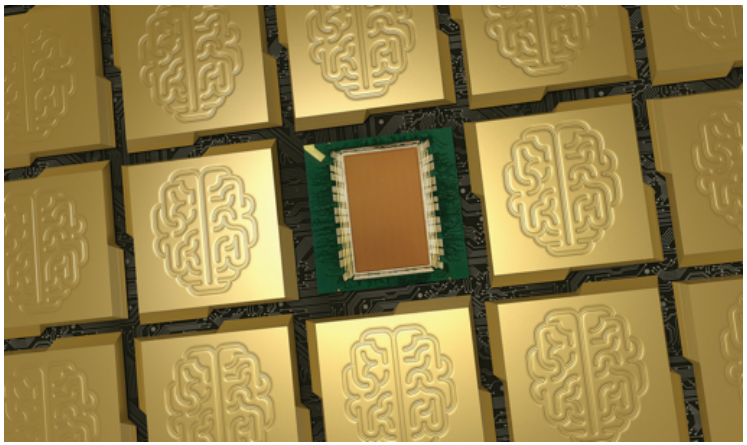
NOEL



Loihi: A Neuromorphic Manycore Processor with On-Chip Learning

IBM'S NEW BRAIN

The TrueNorth neuromorphic chip takes a big step toward using the human brain's architecture to reduce computing's power consumption

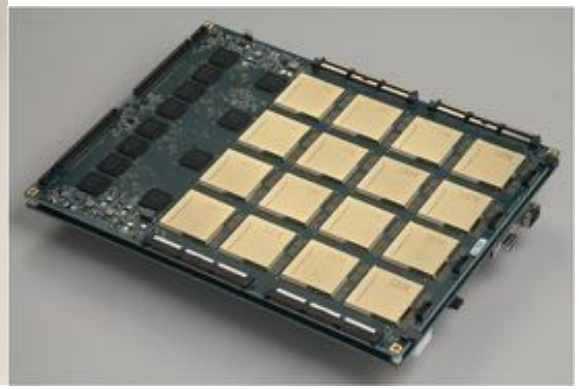


Creating a new paradigm of technology

Driven by its R&D institute, Samsung is set to lead the next century of technological innovation

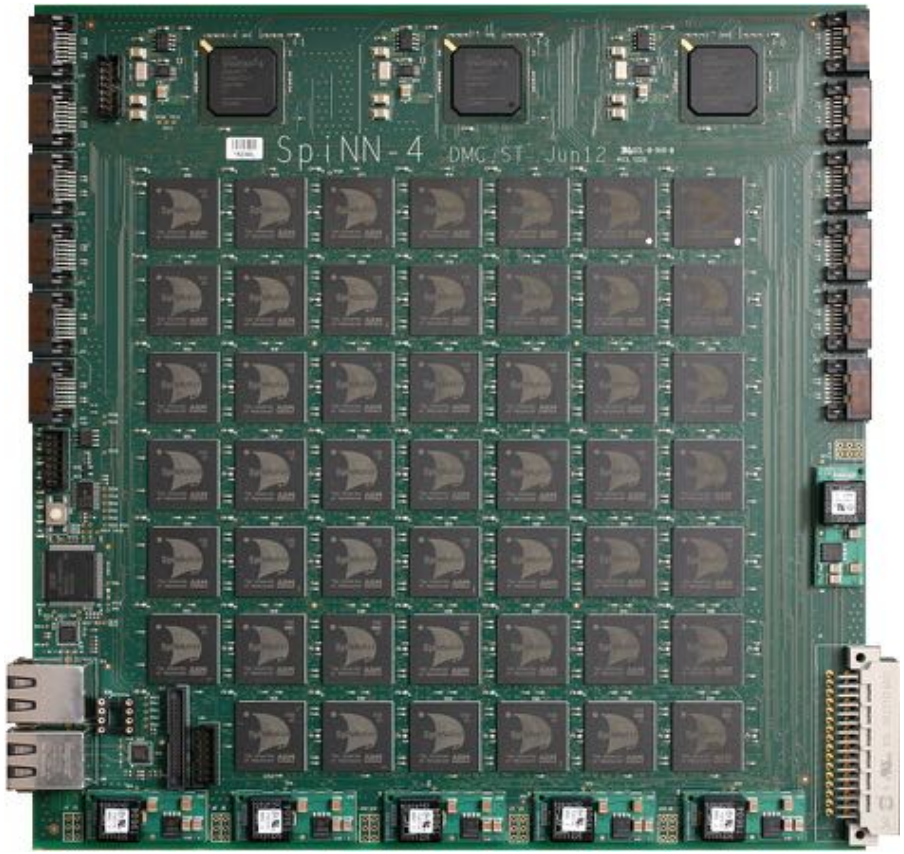
SAIT is also working on an artificial electronic brain that copies biological synaptic organization to create an unprecedented neuromorphic electronic platform that exhibits the unique capabilities of the human brain.

IBM TrueNorth



NOEL

The SpiNNaker



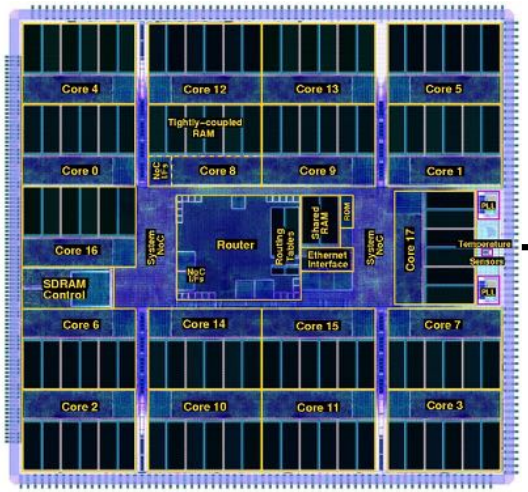
- Designed at the The University of Manchester.
- A million-core computing engine whose flagship goal is to be able to simulate the behaviour of aggregates of up to a billion neurons in real time.
- AER is implemented using packet-switched communication and multicast routing.

NOEL

18 cores / chip

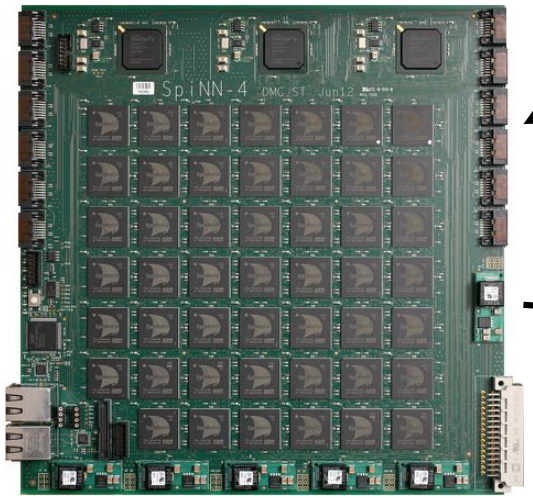
72 cores

1,000 neurons / core



Drosophila scale

pond snail scale



864 cores

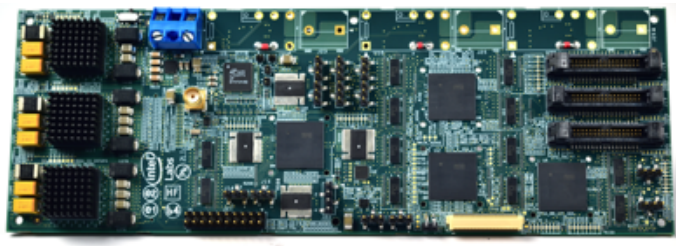
frog scale

20,000 cores / rack

mouse scale

500,000 cores

Intel Loihi



IPhone11 Pro max

"The powerful Apple-designed **A13 Bionic chip** provides unparalleled performance for every task while enabling an **unprecedented leap in battery life**… the A13 Bionic chip sets a new bar for smartphone performance and **power efficiency**… handle and features up to 20 percent faster than CPU and GPU. A13 Bionic is built for machine learning, with a **Neural Engine** for real-time photo and video analysis, and new Machine Learning Accelerators that allow the CPU to deliver more than **1 trillion operations** per second"

NEL
NEURO-BIOMORPHIC ENGINEERING LAB