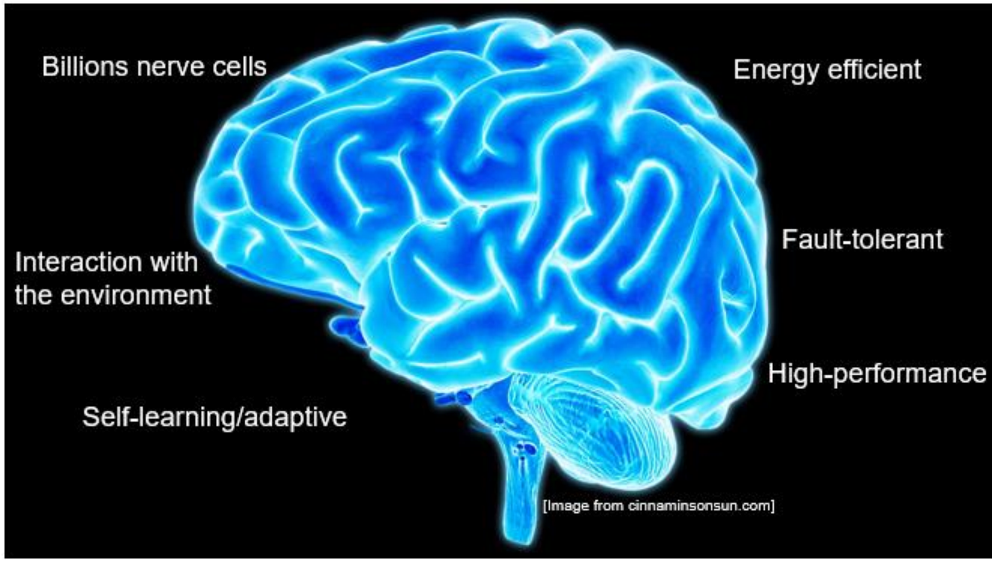
Neuro-inspired Computing Systems & Applications

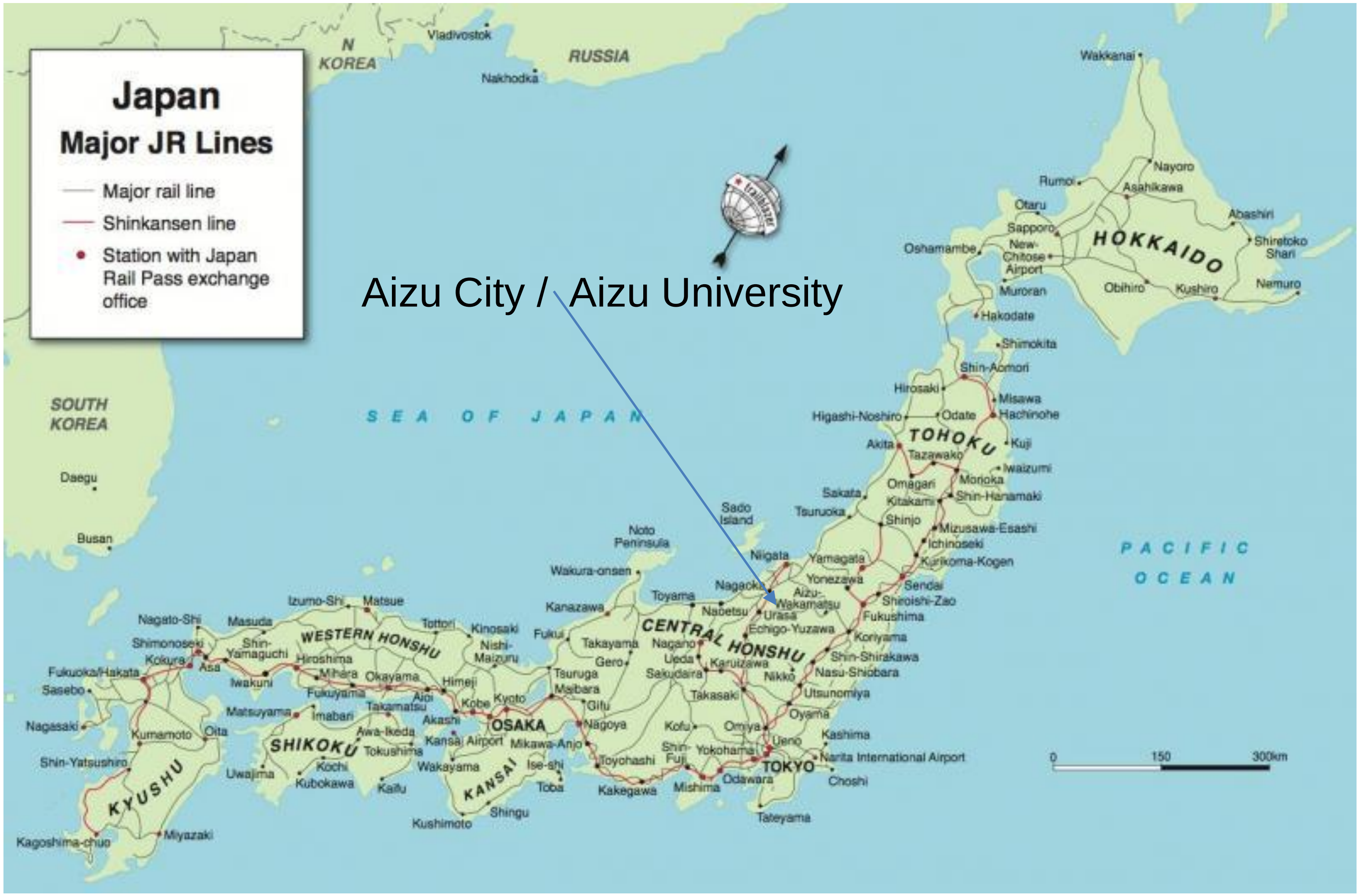
Ben Abdallah Abderazek

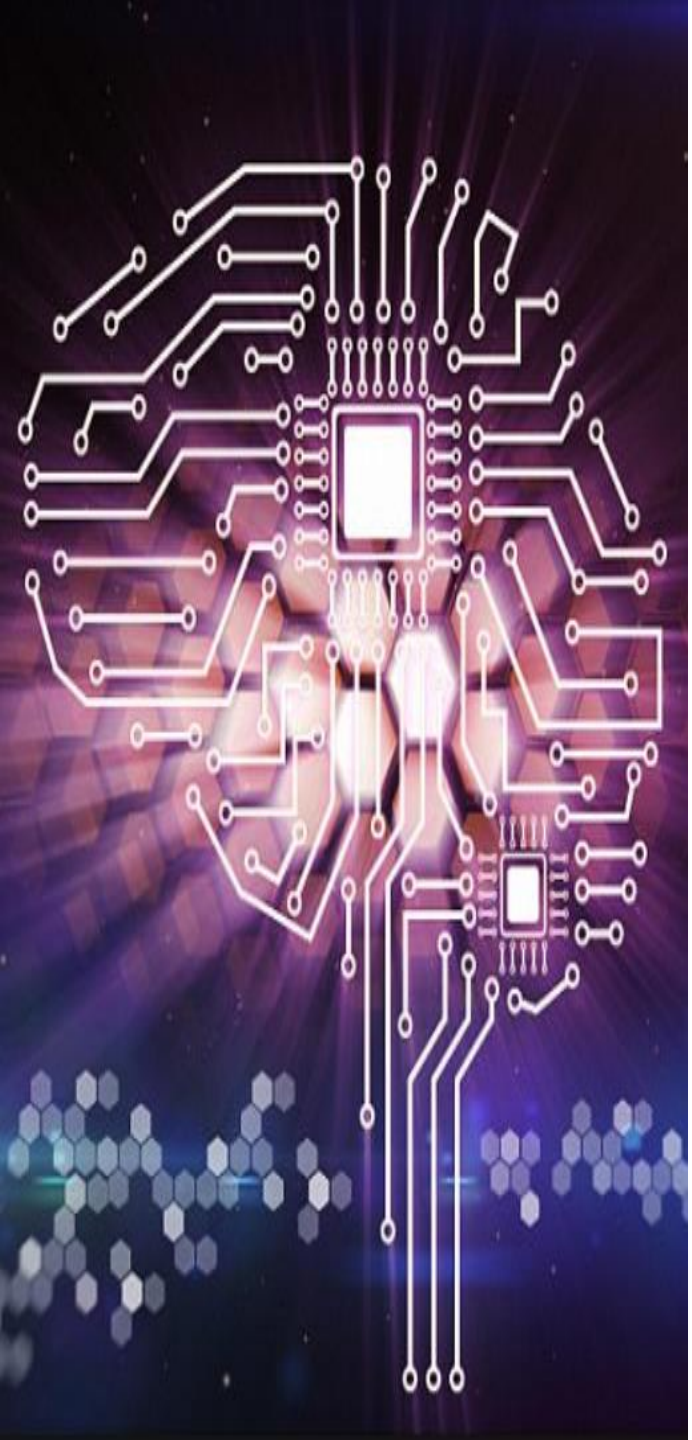
Adaptive Systems Laboratory

benab@u-aizu.ac.jp



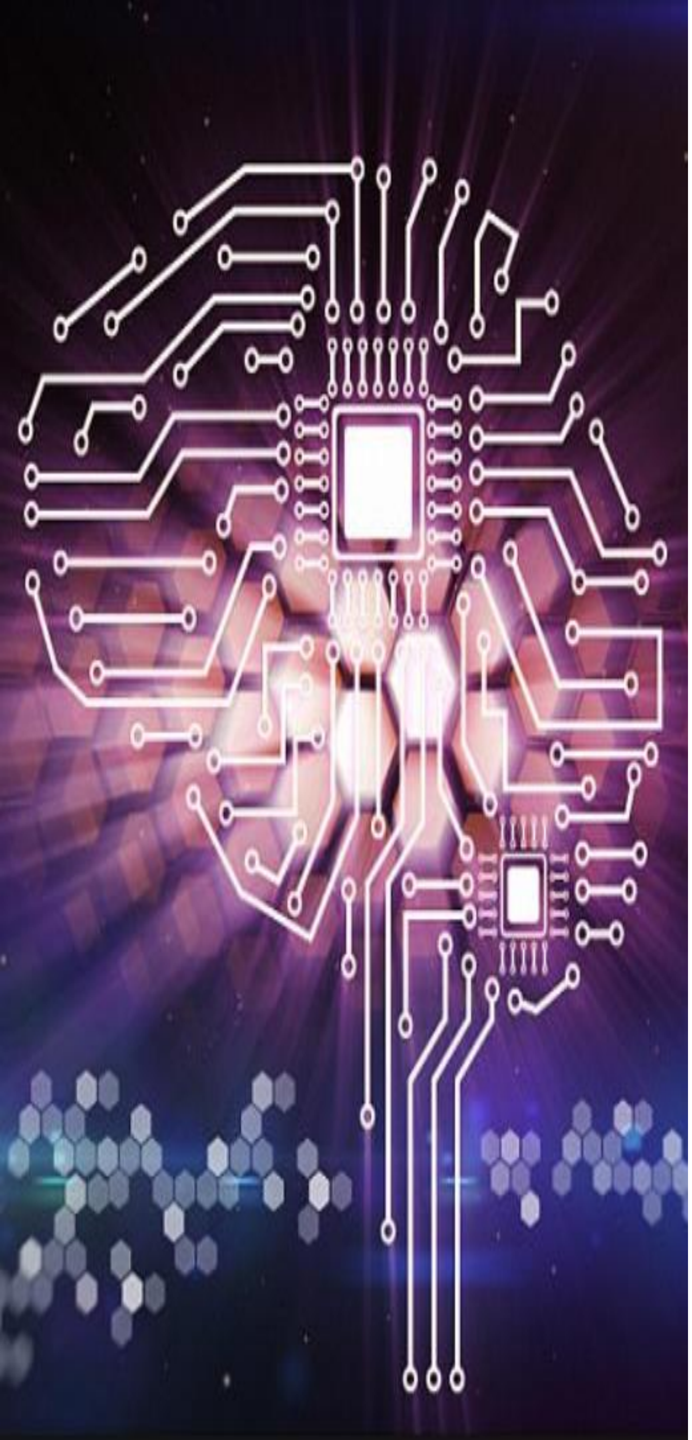
The University of Aizu





Outline

- Technology Transformation
- Neuron Modeling
- Neuro-inspired Systems/Chips
- Concluding Remarks

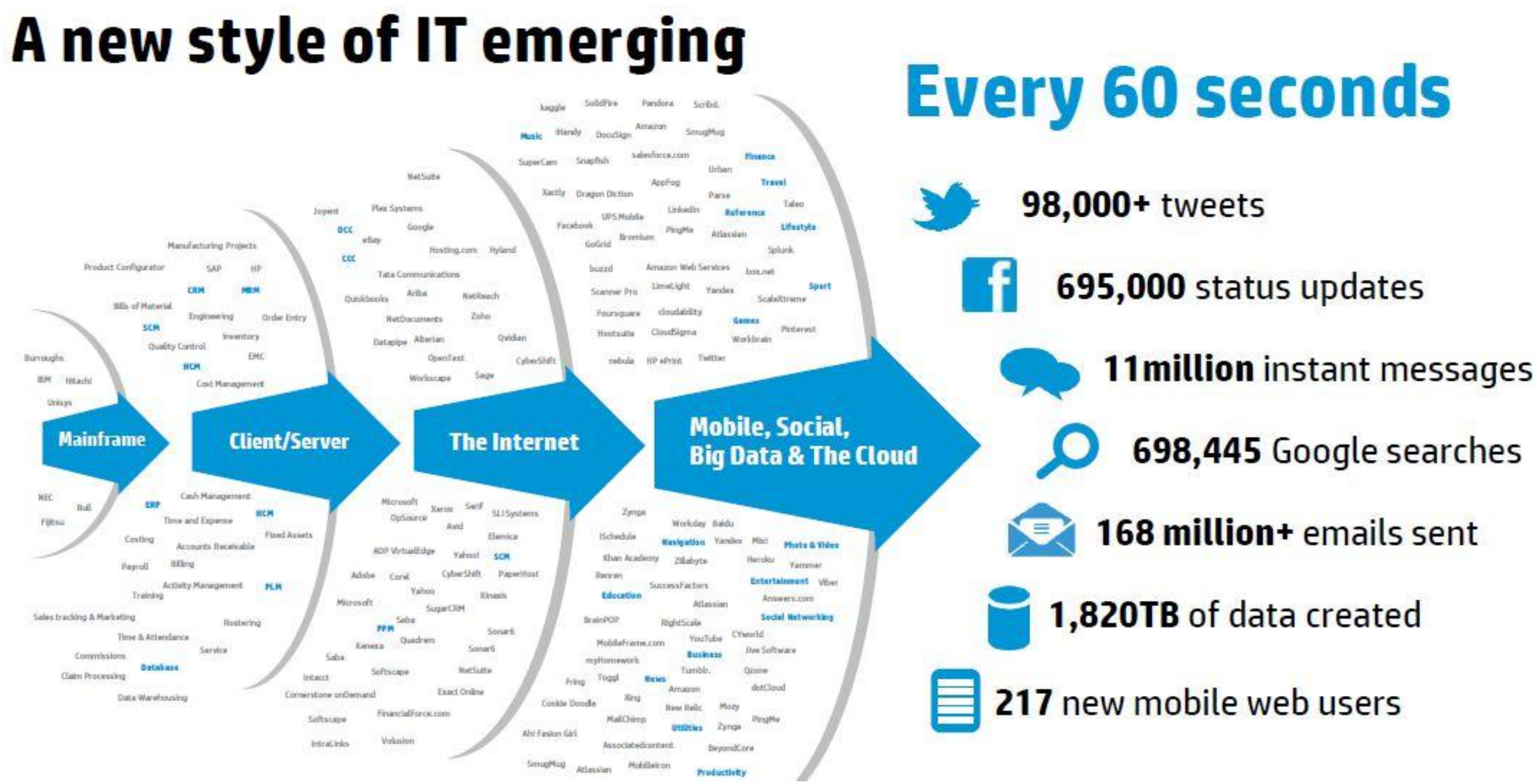


Outline

- Technology Transformation
- Neuron Modeling
- Neuro-inspired Systems/Chips
- Concluding Remarks

Technology Transformation

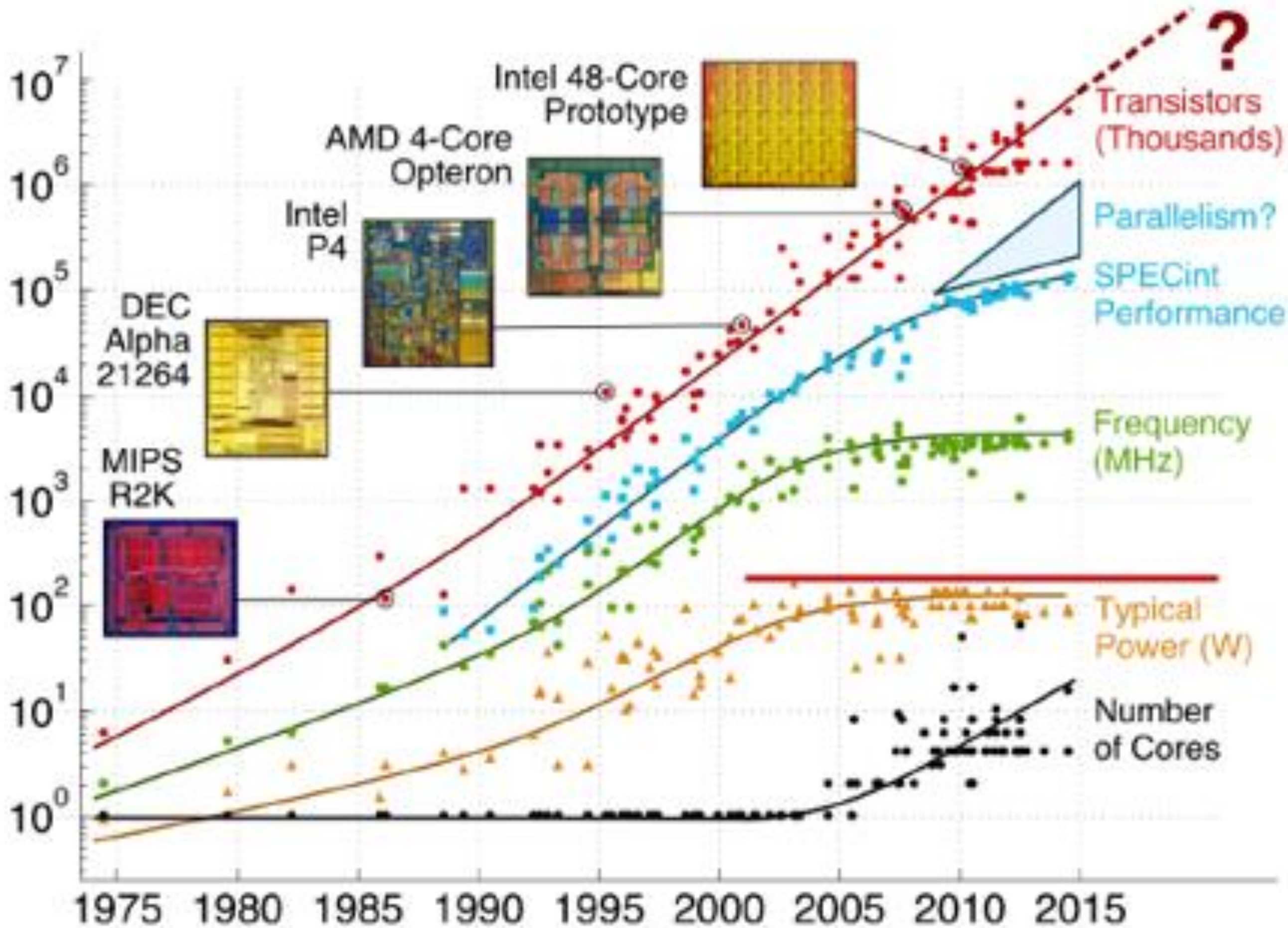
Massive amounts of data is generated.



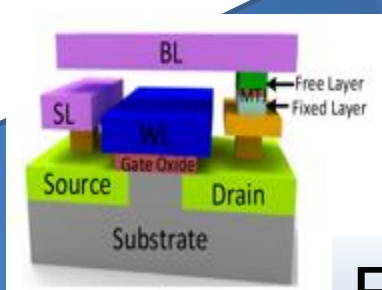
Source: https://practicalanalytics.files.wordpress.com/2012/10/newstyleofit.jpg

Technology Transformation

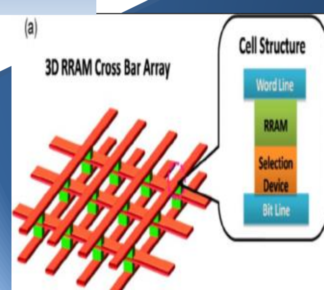
Constant Increase of the number of transistors/cores



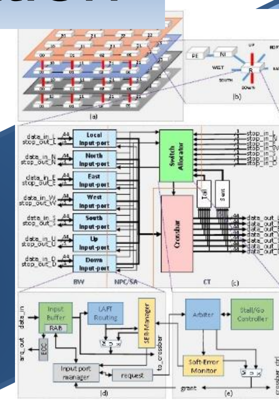
Technology Transformation



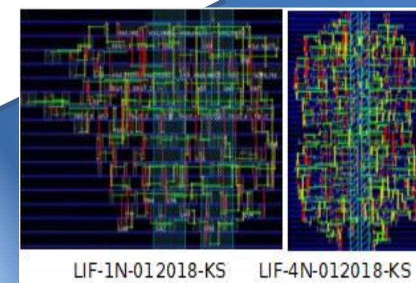
Emerging transistors



Emerging memories



3D Integration

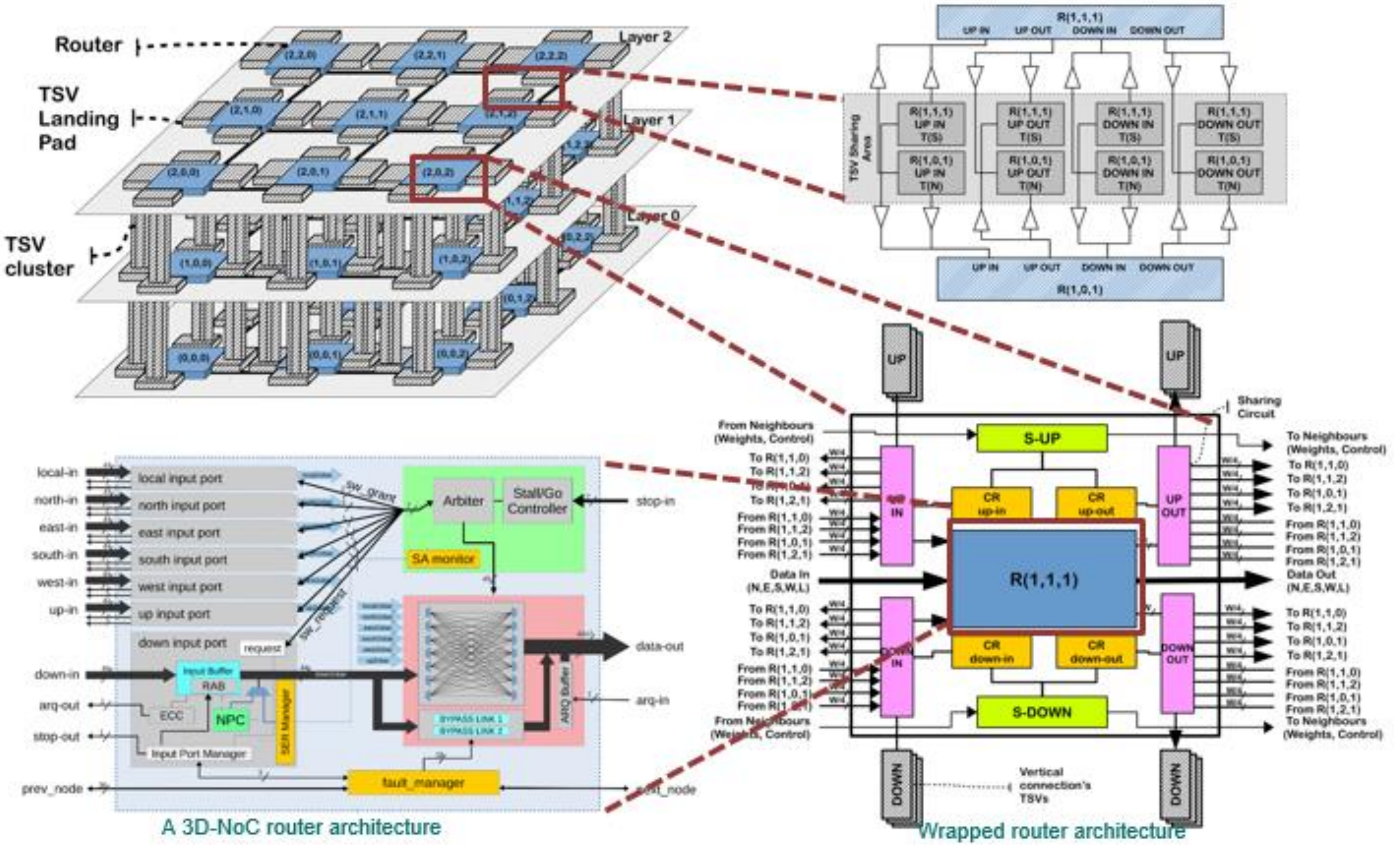


Special architectures

There are many emerging technologies

Technology Transformation

3D-NoC with TSV-cluster Defects Recovery



A 3D-NoC router architecture

Wrapped router architecture

Khanh N. Dang, Akram Ben Ahmed, Yuichi Okuyama, and Abderazek Ben Abdallah, "[Scalable Design Methodology and Online Algorithm for TSV-cluster Defects Recovery in Highly Reliable 3D-NoC Systems](#)", ***IEEE Transactions on Emerging Topics in Computing***, 2017 (in press). DOI: [10.1109/TETC.2017.2762407](#)

Technology Transformation
Robust Scalable NoC

RAF (Reliability Acceleration Factor), which represent the efficiency of the applied fault-tolerances, is given by the following equation:

$$\text{RAF} = \frac{\lambda_{original}}{\lambda_{FT}} = \frac{\mathbf{MTTF}_{FT}}{\mathbf{MTTF}_{original}} \geq 1 \qquad (1)$$

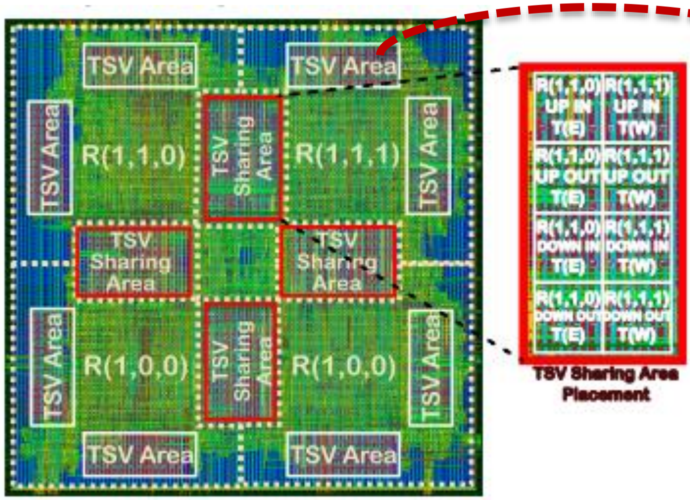
Where:

- λ is the fault rate and it is the inverse value of Mean Time to Failure (*MTTF*).
- $\mathbf{MTTF}_{original}$ is the MTTF of the original system.
- $\mathbf{MTTF}_{FT}$ is the MTTF of the fault-tolerant system.

Khanh N. Dang, Akram Ben Ahmed, Xuan-Tu Tran, Yuichi Okuyama, Abderazek Ben Abdallah, "A Comprehensive Reliability Assessment of Fault-Resilient Network-on-Chip Using Analytical Model", ***IEEE Transactions on Very Large Scale Integration (VLSI) Systems***, Vol. 25, Issue: 11, pp. 3099 – 3112, Nov. 2017. DOI:10.1109/TVLSI.2017.2736004

Technology Transformation
Robust Scalable NoC

Model		Area (μm^2)	Power (mW)			Speed (Mhz)
			Static	Dynamic	Total	
Baseline router [2]		18,873	5.1229	0.9429	6.0658	925.28
Proposal	Router	29,780	10.017	2.2574	12.3144	613.50
	Serialization	3,318	0.9877	0.2807	1.2684	-
	TSV Sharing	5,740	0.7863	0.2892	1.0300	-
	Total	38,838	11.7910	2.8273	14.6128	537.63



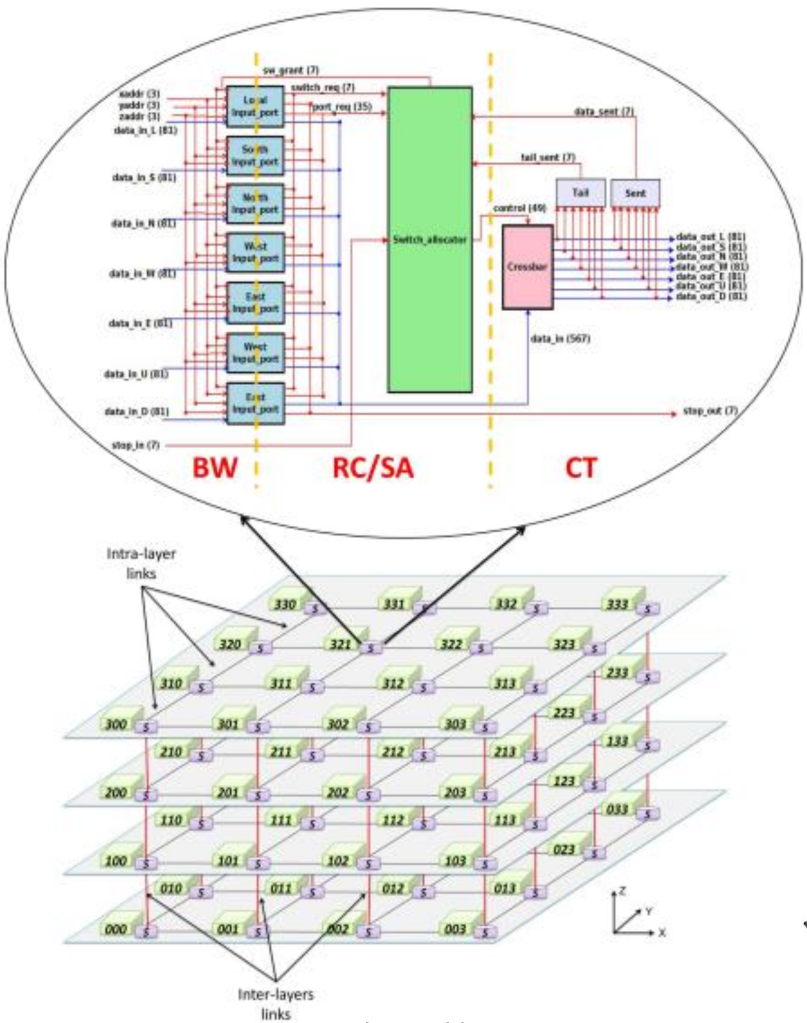
Single layer layout illustrating the TSV sharing areas (red boxes). The layout size is **865μm × 865μm.**

The sharing TSV area are the red boxes. Each sharing area has 8 clusters for 4 ports and 2 routers.

Khanh N. Dang, Akram Ben Ahmed, Xuan-Tu Tran, Yuichi Okuyama, Abderazek Ben Abdallah, "[A Comprehensive Reliability Assessment of Fault-Resilient Network-on-Chip Using Analytical Model]", *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 25, Issue: 11, pp. 3099 – 3112, Nov. 2017. DOI:10.1109/TVLSI.2017.2736004

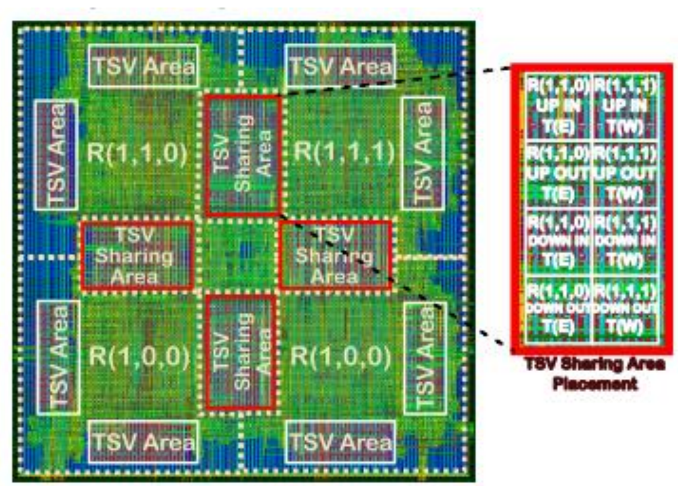
Technology Transformation

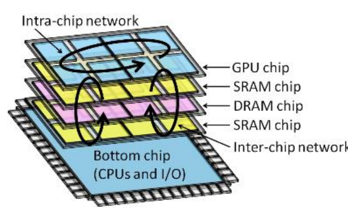
Robust Scalable NoC

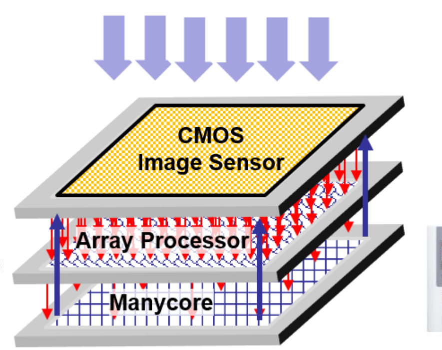


OASIS Network-on-Chip System

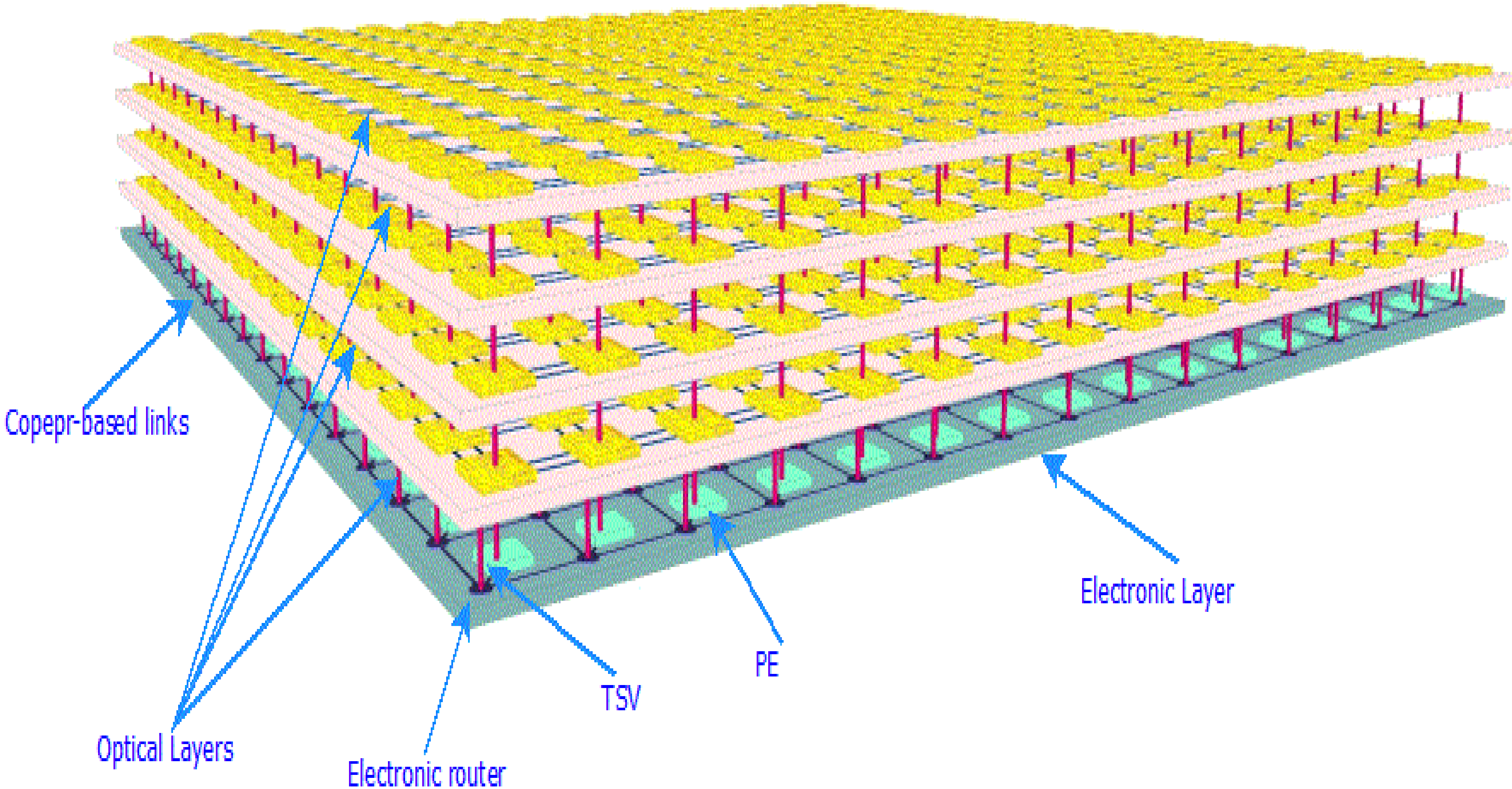
(7/14)







Technology Transformation
Hybrid Electro-Photonic NoC



3D-PHENIC System Architecture

Technology Transformation
Hybrid Electro-Photonic NoC

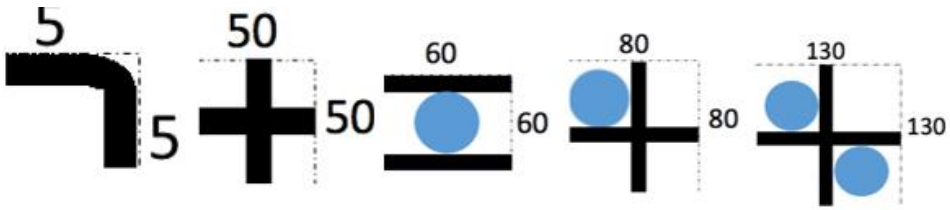
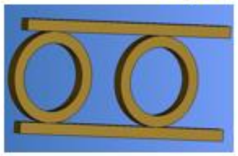


Fig. 20 Perpendicular Redundant MRs



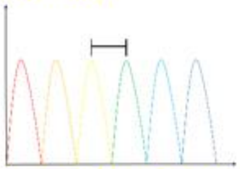
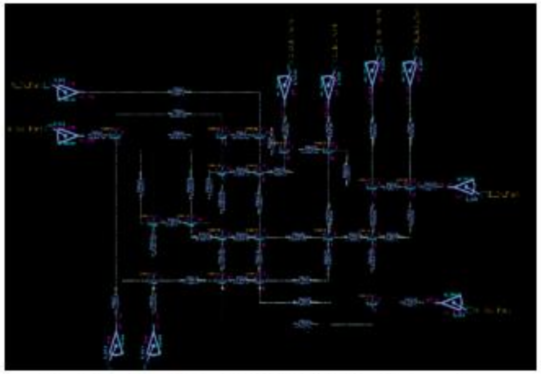


Fig. 19 Parallel Redundant MRs

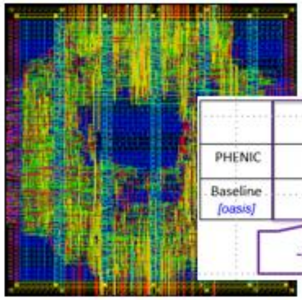
Fig. 21 Wavelength Spacing

$$LOSS_{router} = Loss_{bend} \times N_{bend} + Loss_{cross} \times N_{cross} + Loss_{MR_{OFF}} \times N_{MR_{OFF}} + Loss_{MR_{ON}} \times N_{MR_{ON}}$$

$$Delay_{router} = Delay_{bend} \times N_{bend} + Delay_{cross} \times N_{cross} + Delay_{MR_{OFF}} \times N_{MR_{OFF}} + Delay_{MR_{ON}} \times N_{MR_{ON}}$$



Model of a 5 Port FTTDOR Switch and a Wavelength Shifting Controller



	Area (um²)	Power (uW)
PHENIC	18130	210
Baseline [oasis]	28306	311
	-35%	-32%

PHENIC's electronic controller layout in 45 nm process

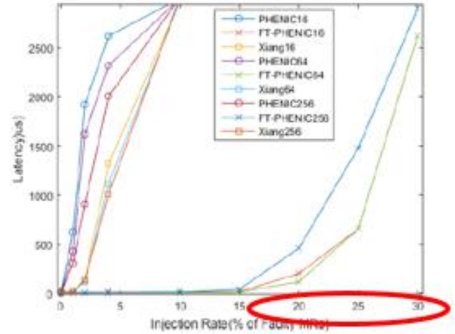


Fig. 10 Latency results of each system as faults are introduced.

Achraf Ben Ahmed, Tsutomu Yoshinaga, Abderazek Ben Abdallah, "[Scalable Photonic Networks-on-Chip Architecture Based on a Novel Wavelength-Shifting Mechanism]", ***IEEE Transactions on Emerging Topics in Computing***, 2017 (in press). DOI: 10.1109/TETC.2017.2737016

Technology Transformation

What is the issue with the current computing technology?

Technology Transformation

What is the issue with the current computing technology?

Scalability issue.

Technology Transformation

What does that mean ?

Technology Transformation

What does that mean ?

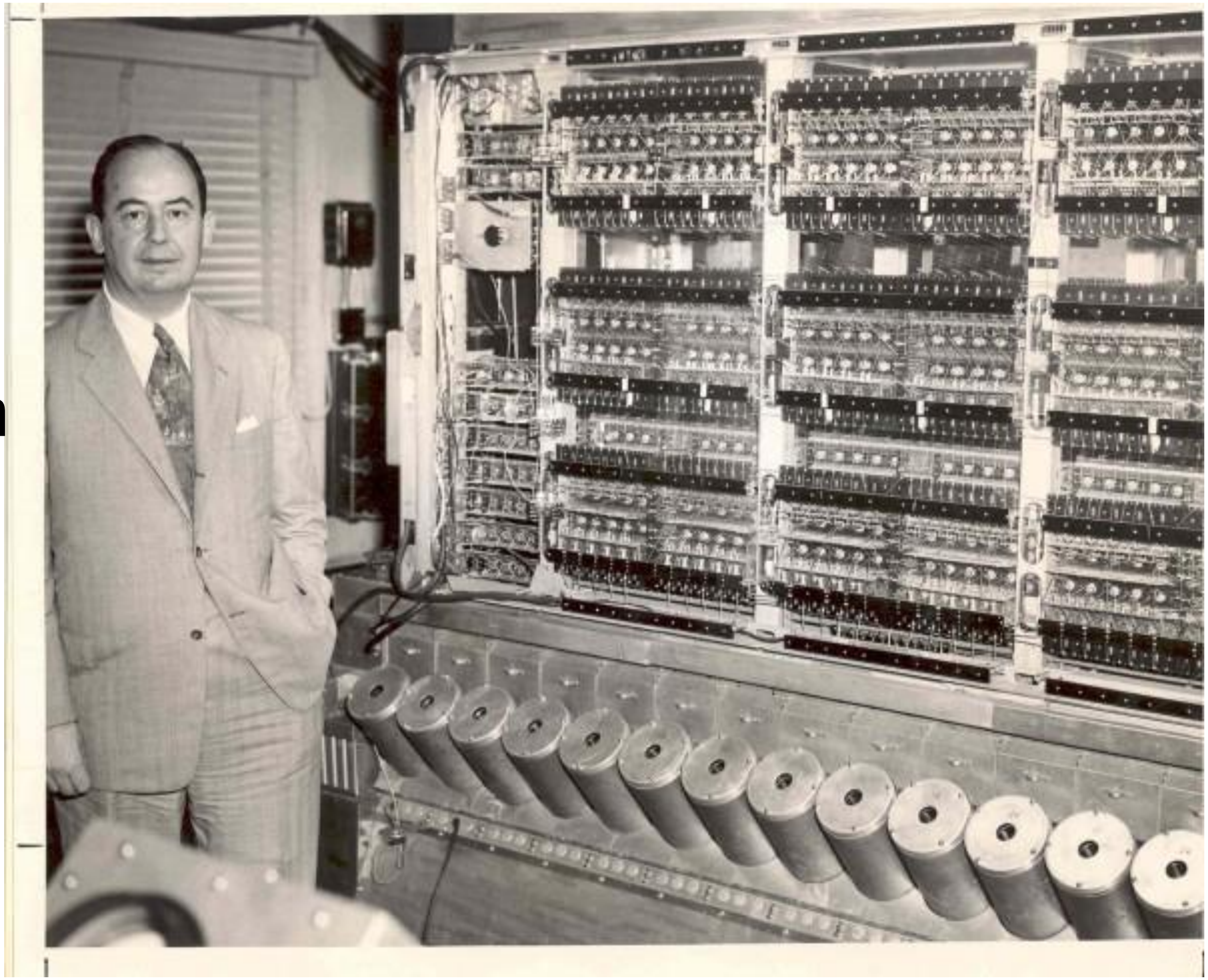
i. Transistor nbr doubles every year, but we cannot get energy to operate the whole chip - Dark Silicon.

ii. We double the number of transistors with smaller sizes, but we are producing much more heat in the same space.

iii. The speed of the chip increases, but the memory bandwidth does not keep-up.

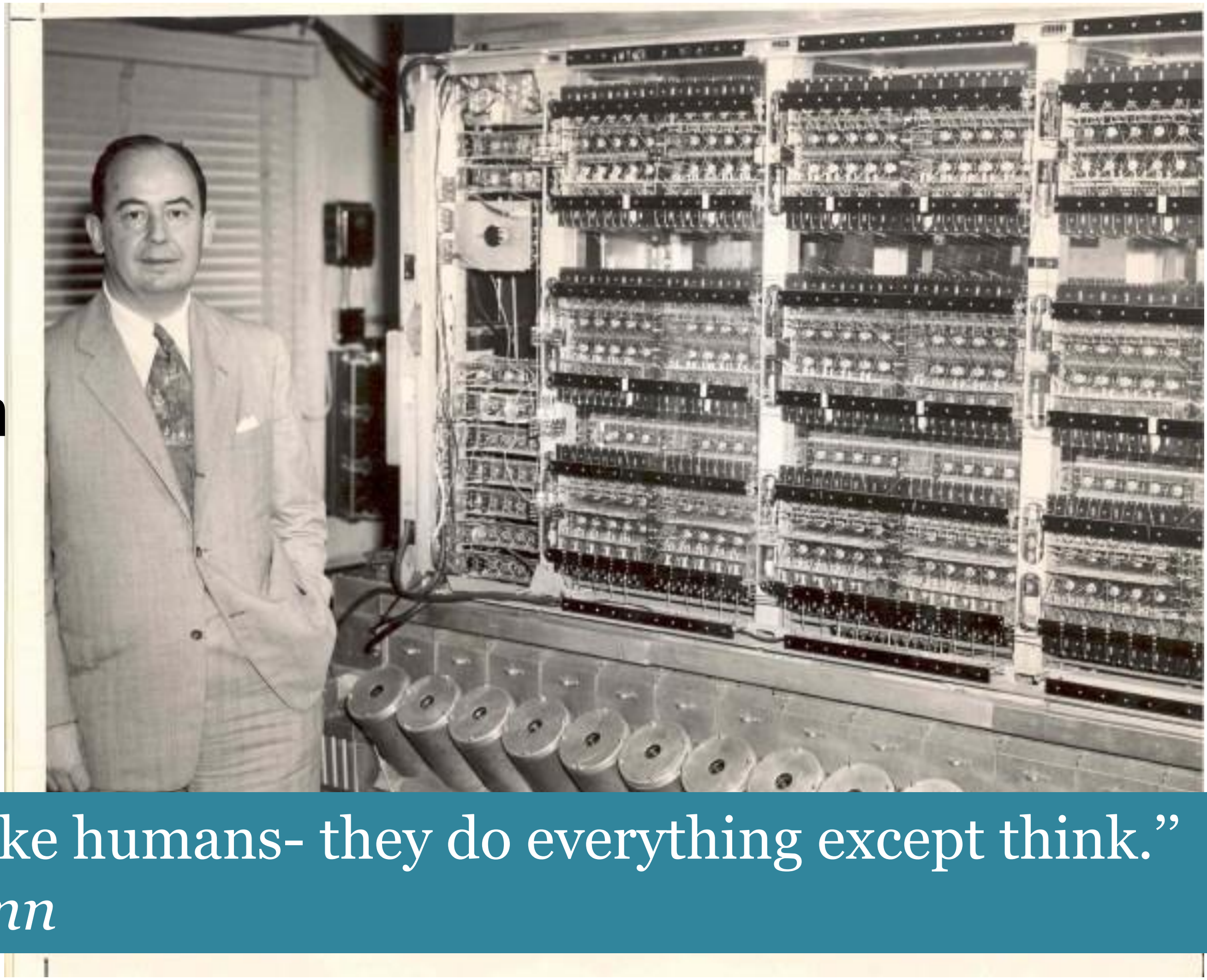
John von Neumann Machine

stored-program Computer.



John von Neumann Machine

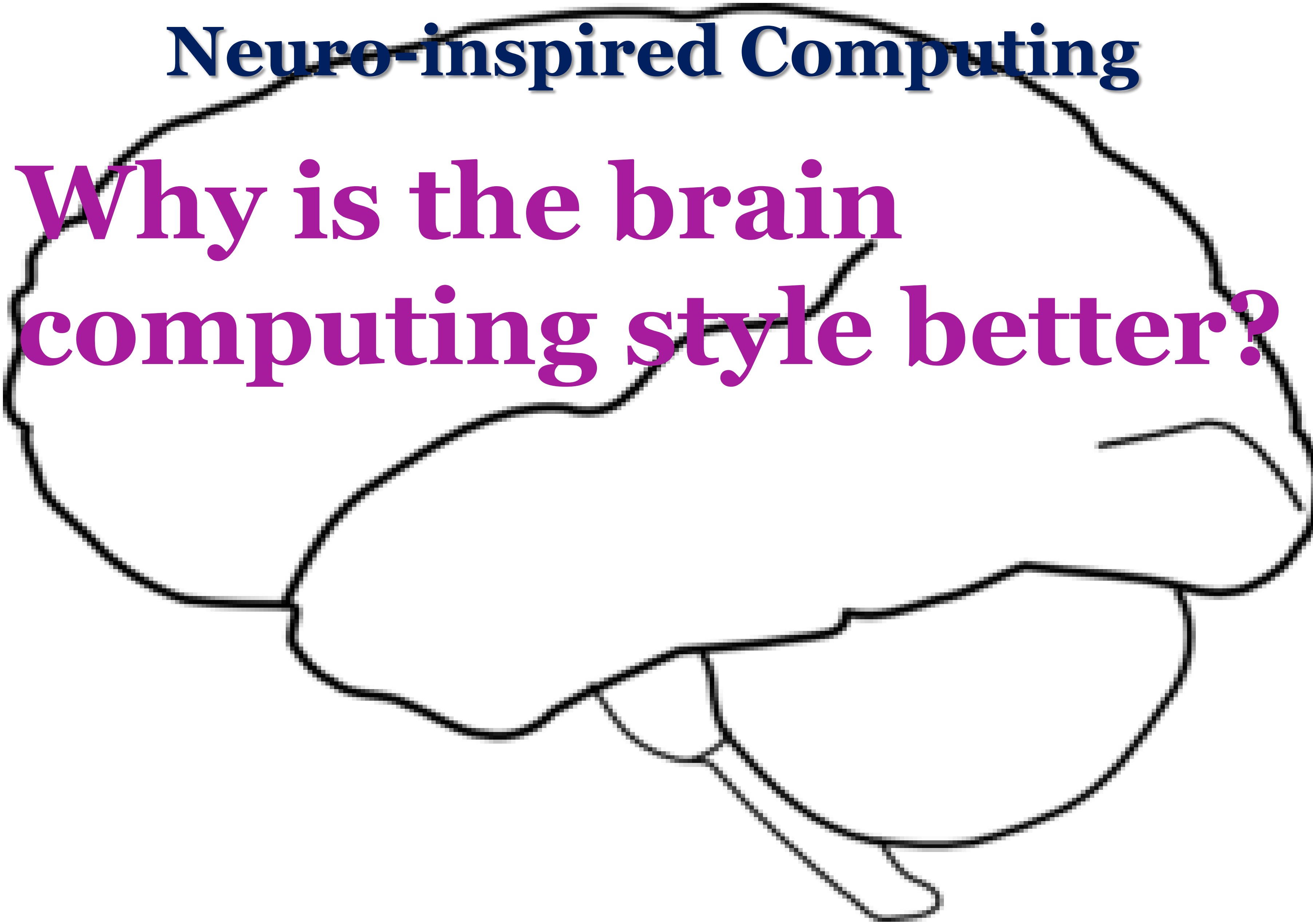
stored-program Computer.



"Computers are like humans- they do everything except think."
John von Neumann

Neuro-inspired Computing

Why is the brain computing style better?



Neuro-inspired Computing

Why is the brain computing style better?

BECAUSE

- ✓ Consumes low power - **~20W)**
- ✓ Fault tolerant -**brain continues to operate even when the circuit (neuron, neuroglia) is died)**
- ✓ Works in parallel **->10^6 parallelism vs. <10^1 for VN)**
- ✓ Faster than current computers - **i.e. simulation of a 5 s brain activity takes ~500 s on state-of-the- art supercomputer [US PTN 20160125287A1]**
- ✓ Learn and think - **needless to prove ☺**

Neuro-inspired Computing

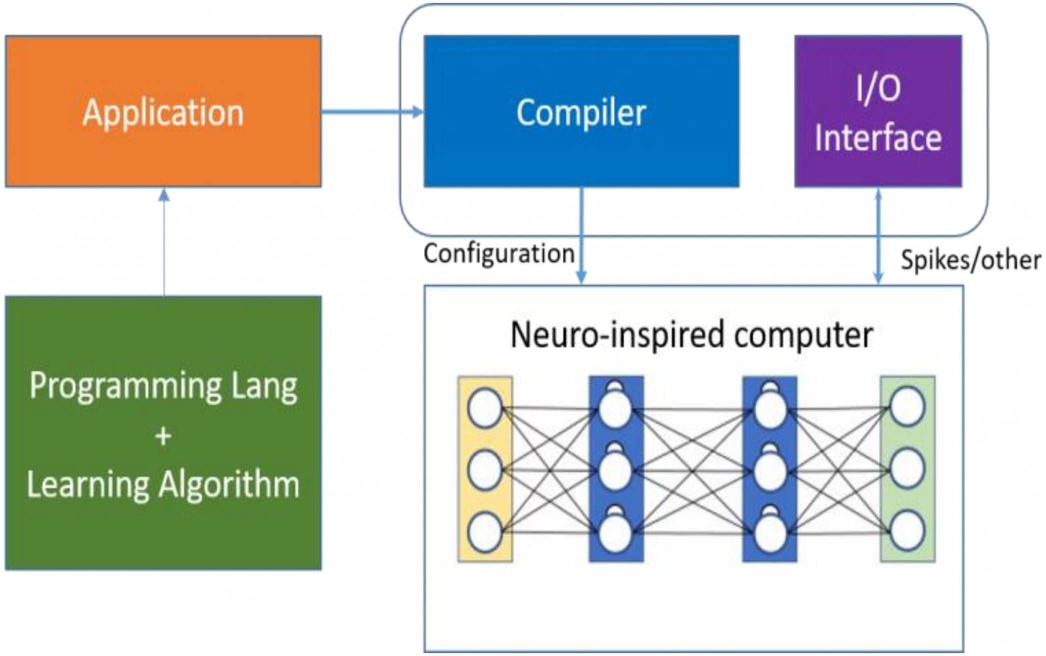
How do we design this new brain-like machine?

Neuro-inspired Computing

How do we design this new brain-like machine?

WE NEED

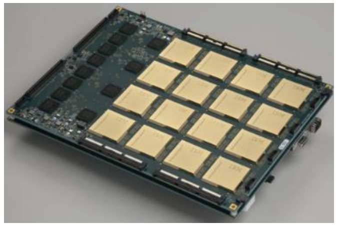
- ❖ New Software
  - Parallel programming abstraction
- ❖ <u>New Hardware</u>
  - Massively Parallel
  - Scalable connectivity
  - Low-powered cores



Type of Neuro-inspired Computing Systems

- **Neuromorphic Sensors** - electronic models of retinas and cochleas.
- **Smart sensors** – tracking chips, motion, pressor, auditory classifications and localization sensors.
- **Models of specific systems**: e.g. lamprey spinal cord for swimming, electric fish lateral line.
- **Pattern generators** – for locomotion or rhythmic behavior
- **Large-scale multi-core/chip systems** – for investigating models of neuronal computation and synaptic plasticity.

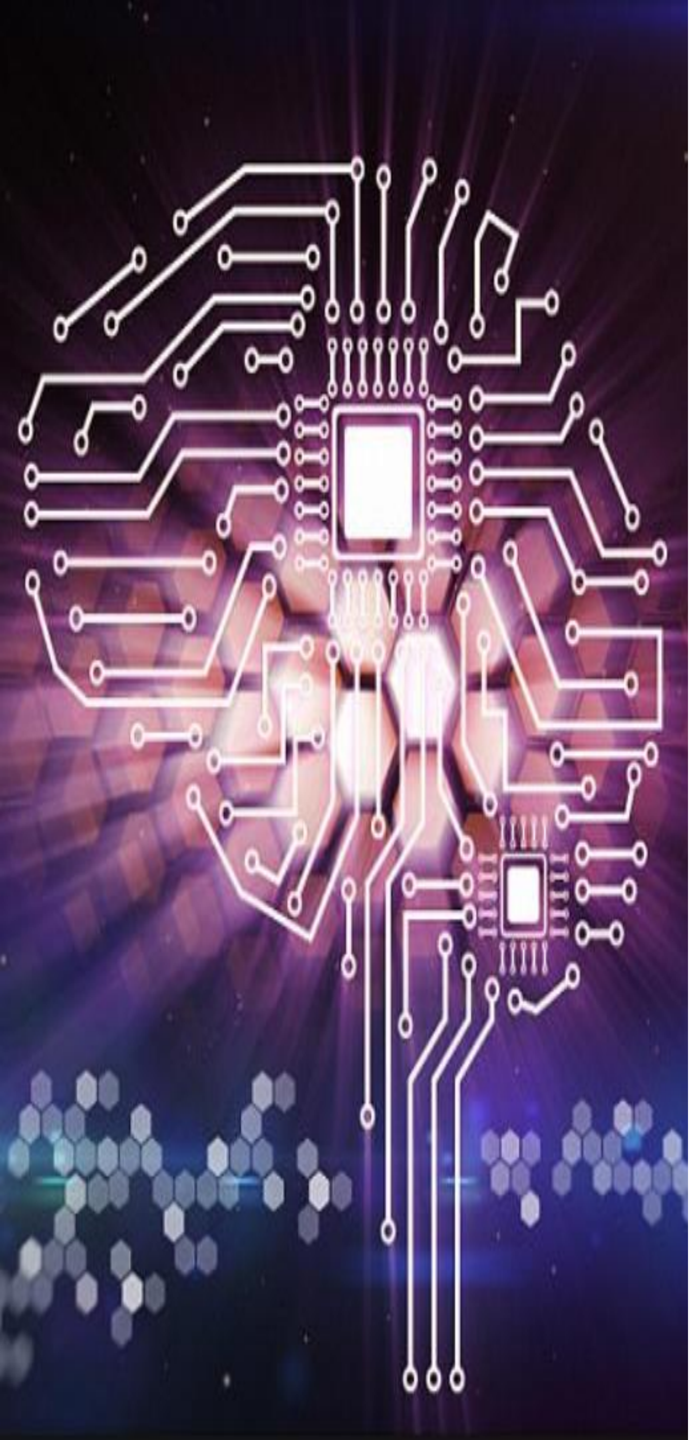
Neurogrid (Stanford)



TrueNorth (IBM)

Brainscales/HBP (Heidelberg, Lausanne)

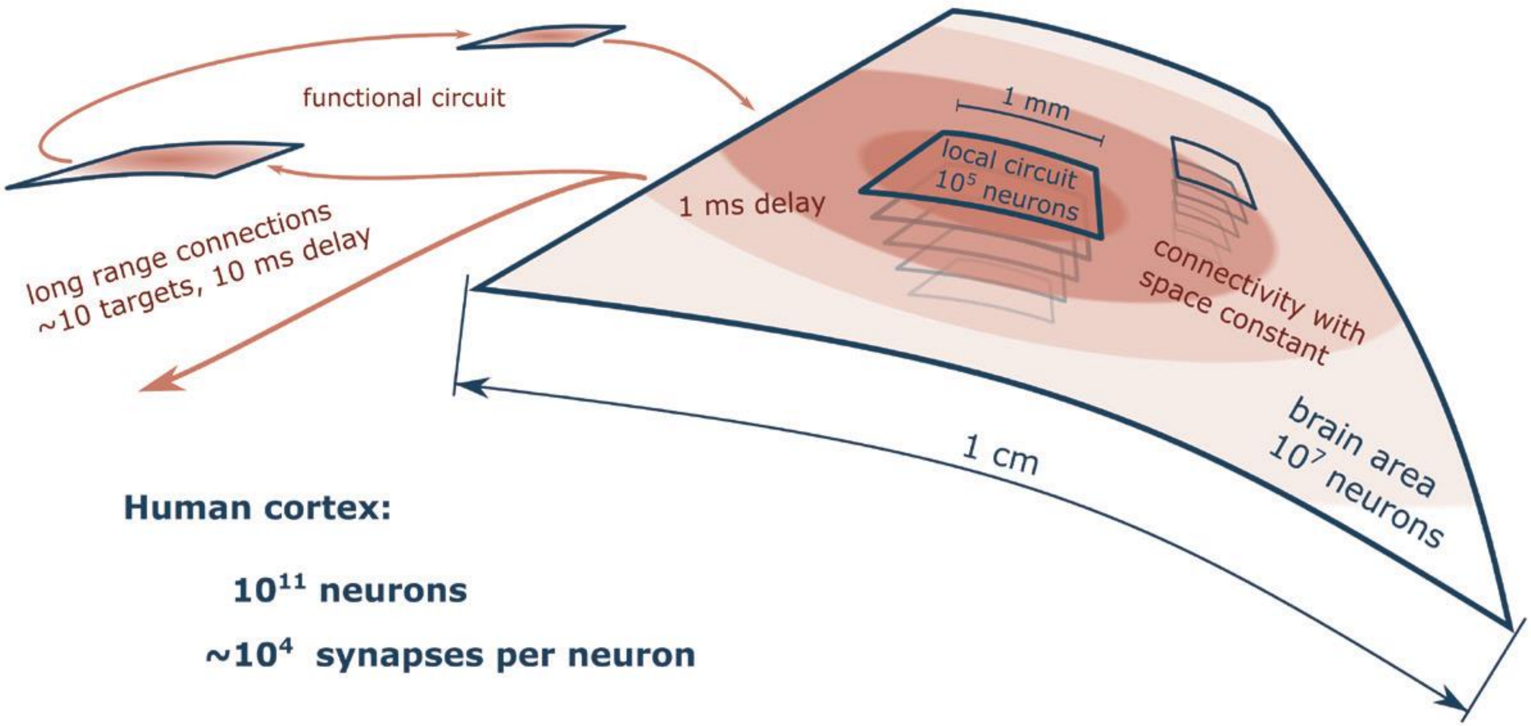
SpiNNaker (Manchester)



Outline

- Technology Transformation
- Neuron Modeling
- ASL Neuro-inspired Systems/Chips
- Concluding Remarks

Connectivity in Human Cortex



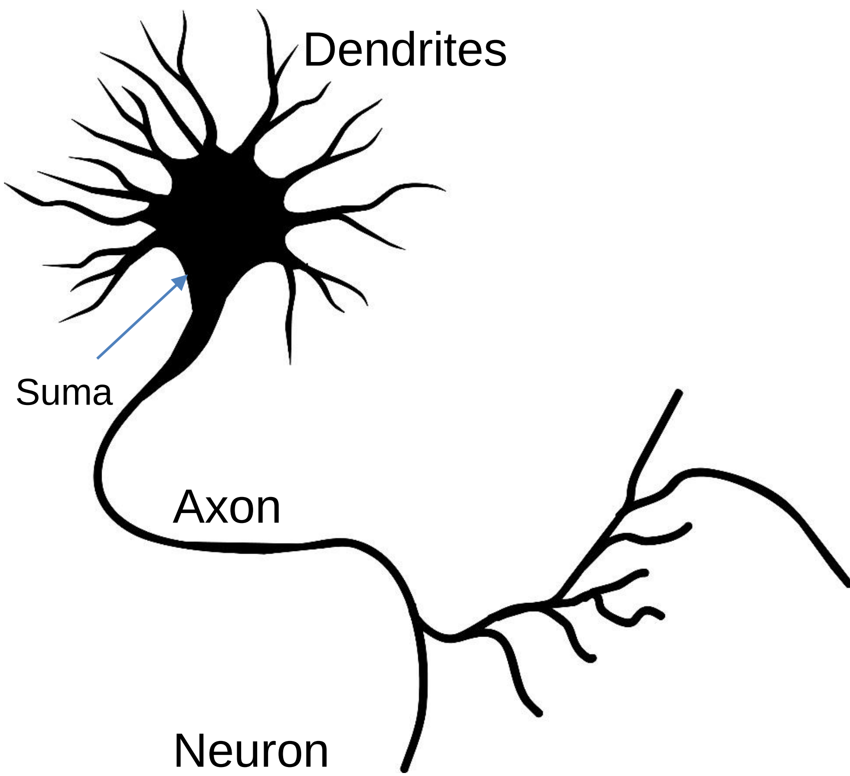
Desmann, 4th Biosupercomputing Symposium, 2012

There are three known level of connections in the Human Cortex:

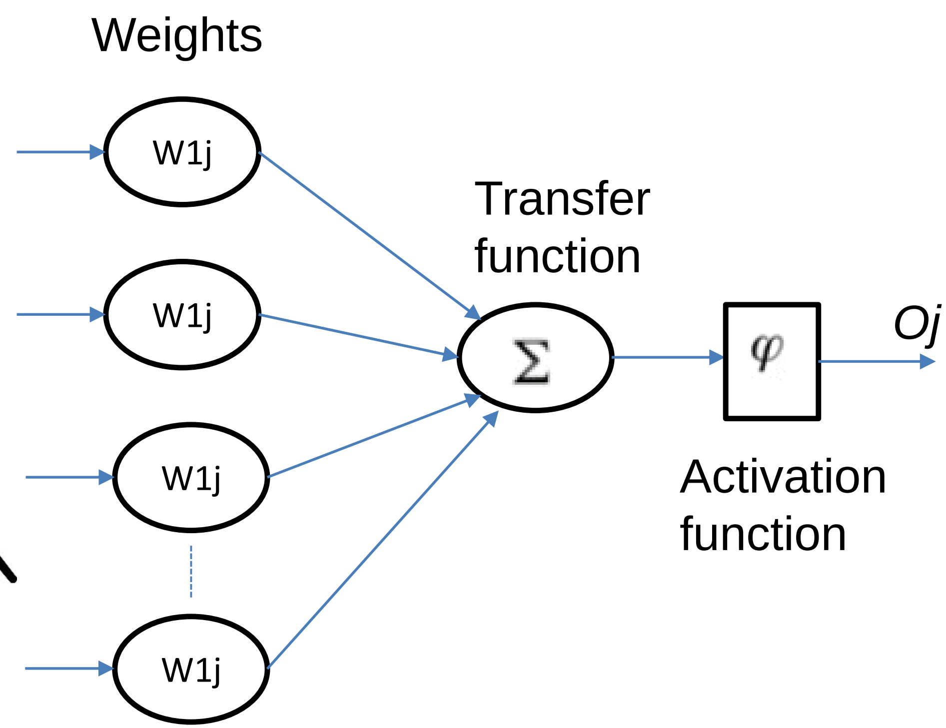
- connectivity of local microcircuit
- within-area connectivity with space constant
- long-range connections between areas

Neuron Modelling

Biology: Single Neuron



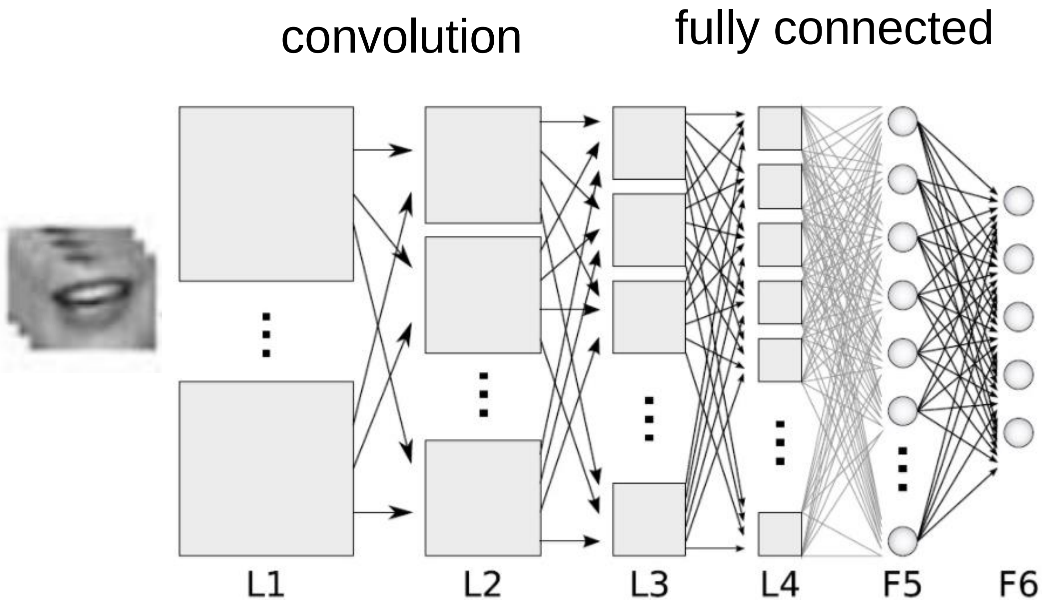
Machine Learning



Neuron Modelling

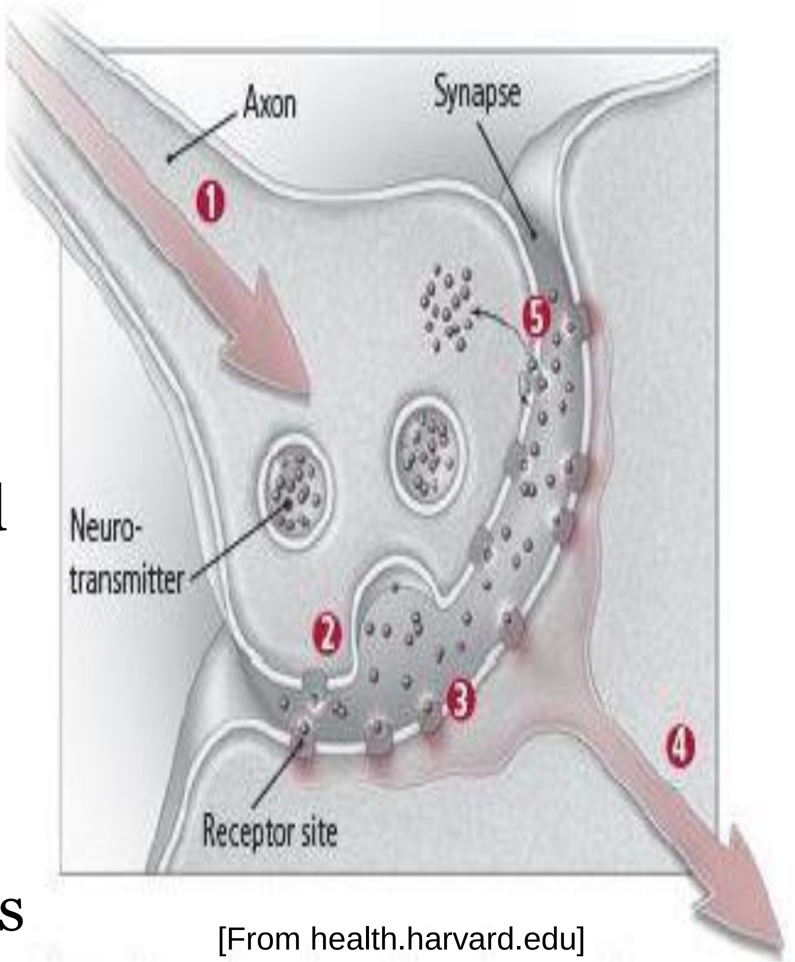
Biology: Tree Neurons

Machine Learning



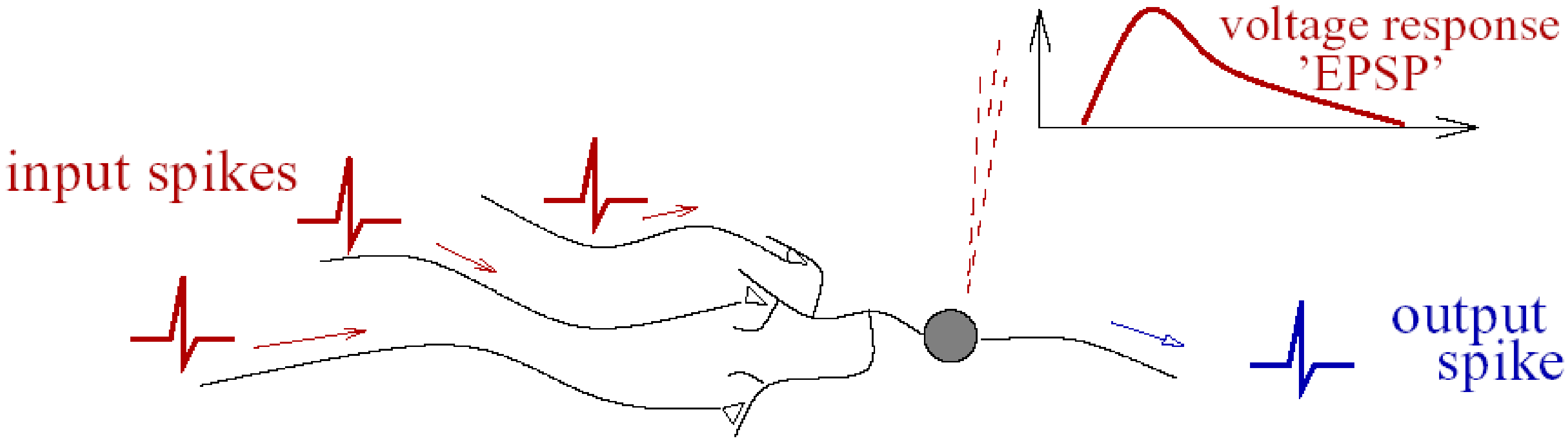
How neurons communicate?

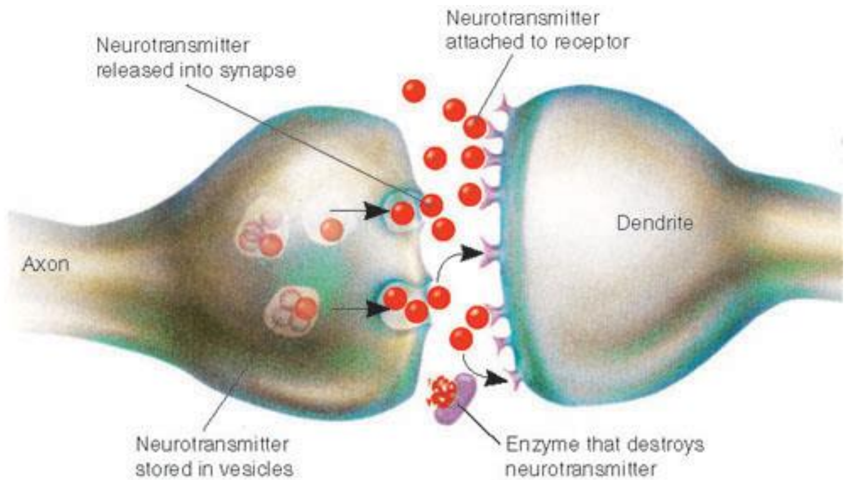
1. An electrical signal travels down the axon.
2. Chemical neurotransmitter molecules are released.
3. The neurotransmitter molecules bind to receptor sites.
4. The signal is picked up by the second neuron and is either passed along or halted.
5. The signal is also picked up by the first neuron, causing reuptake, the process by which the cell that released the neurotransmitter takes back some of the remaining molecules.



[From health.harvard.edu]

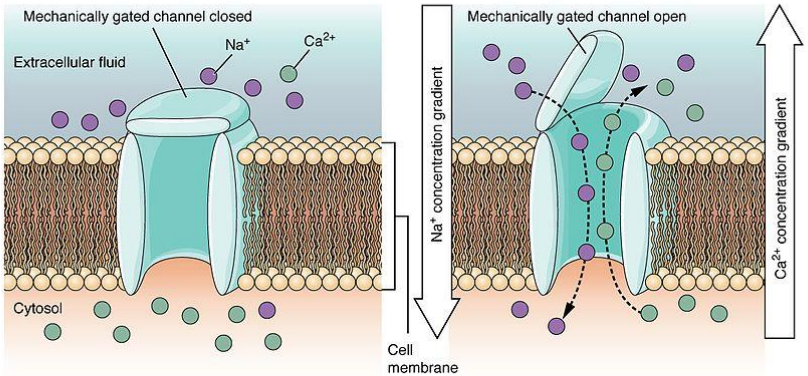
Spiking Neuron

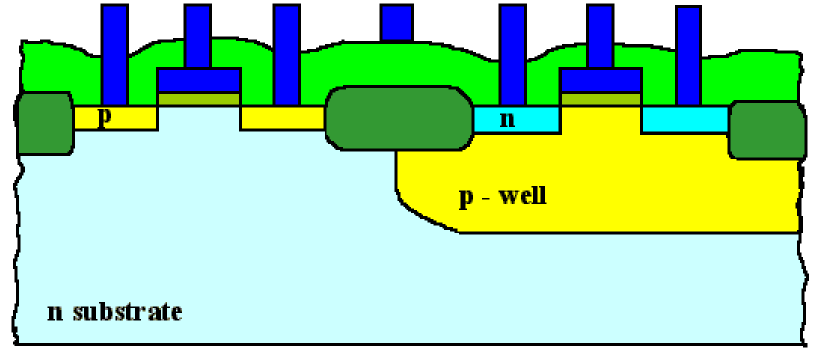




- Computing with precisely timed spikes is more powerful than with "rates". [W. Maass, 1999]

Electronic devise vs chemical device

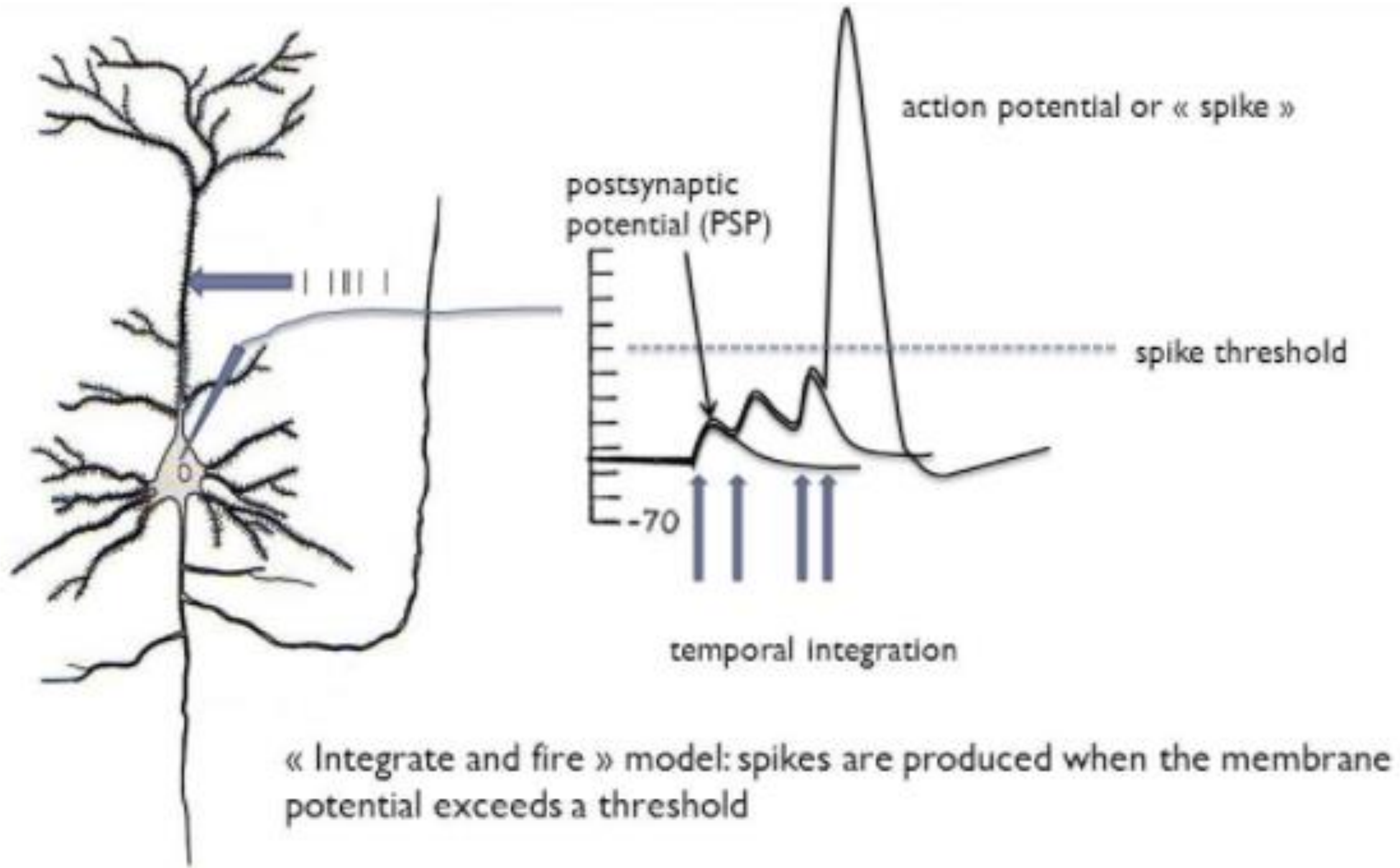




- Deliver the concentration difference of K+,Na+
- Action potential ~ 80 mV
  - ➢ Extreme low voltage operation
  - ➢ Noise problem
  - ➢ Multiple signal input/ integration

- Spatial and temporal multiplexing → Active sharing of the interconnect
- Chemical computing, extremely low operation voltage (<100mV) → Low power

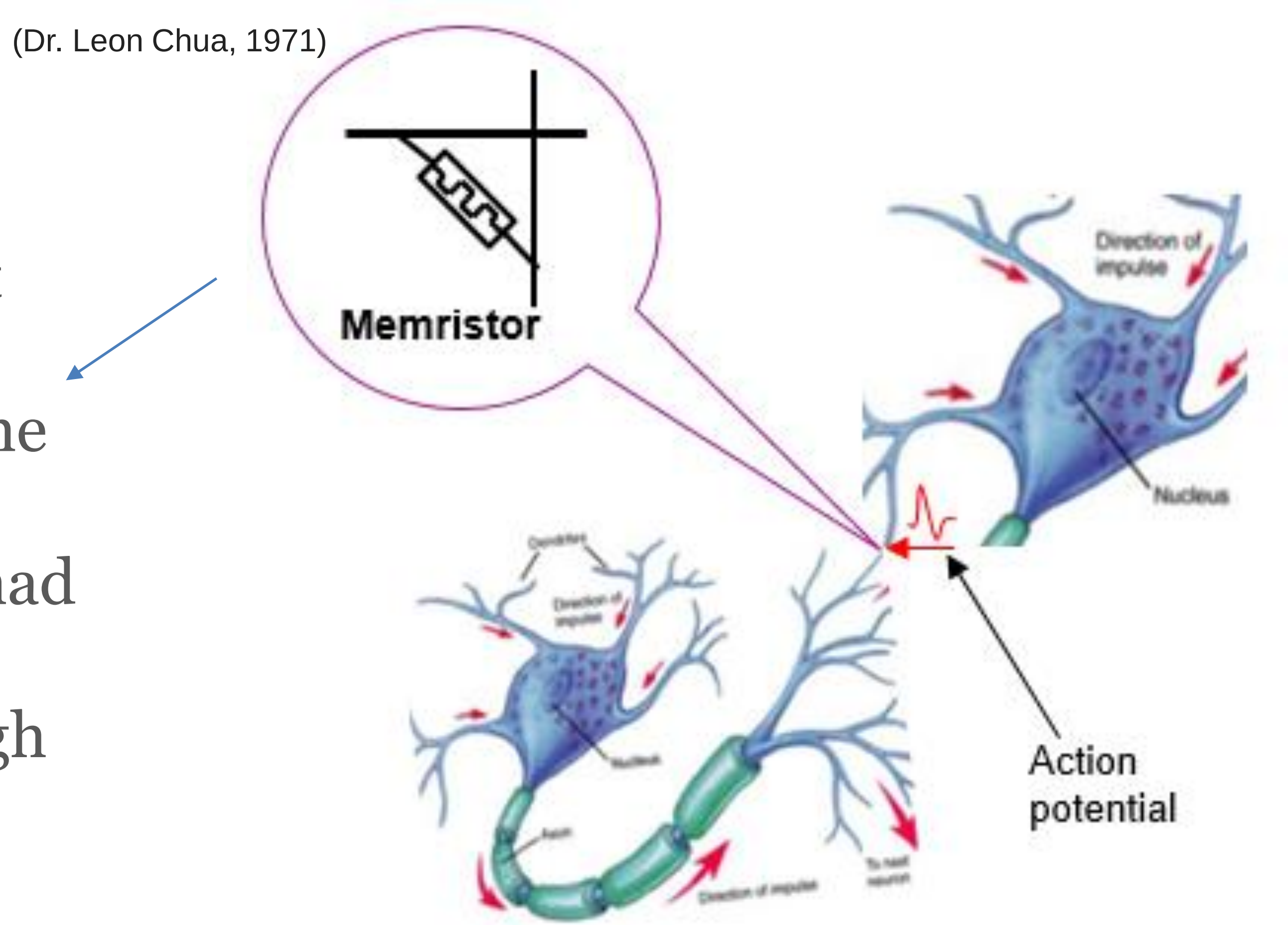
Fundamental interactions



« Integrate and fire » model: spikes are produced when the membrane potential exceeds a threshold

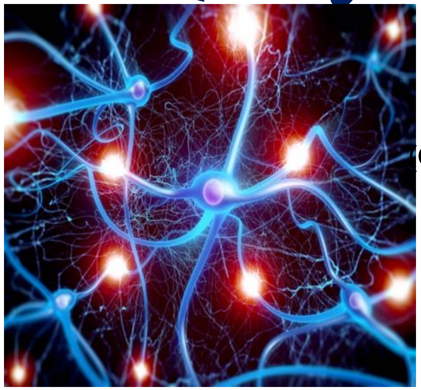
Action Potential (Synapse) Storage

The electrical resistor is not constant but depends on the history of current that had previously flowed through the device.

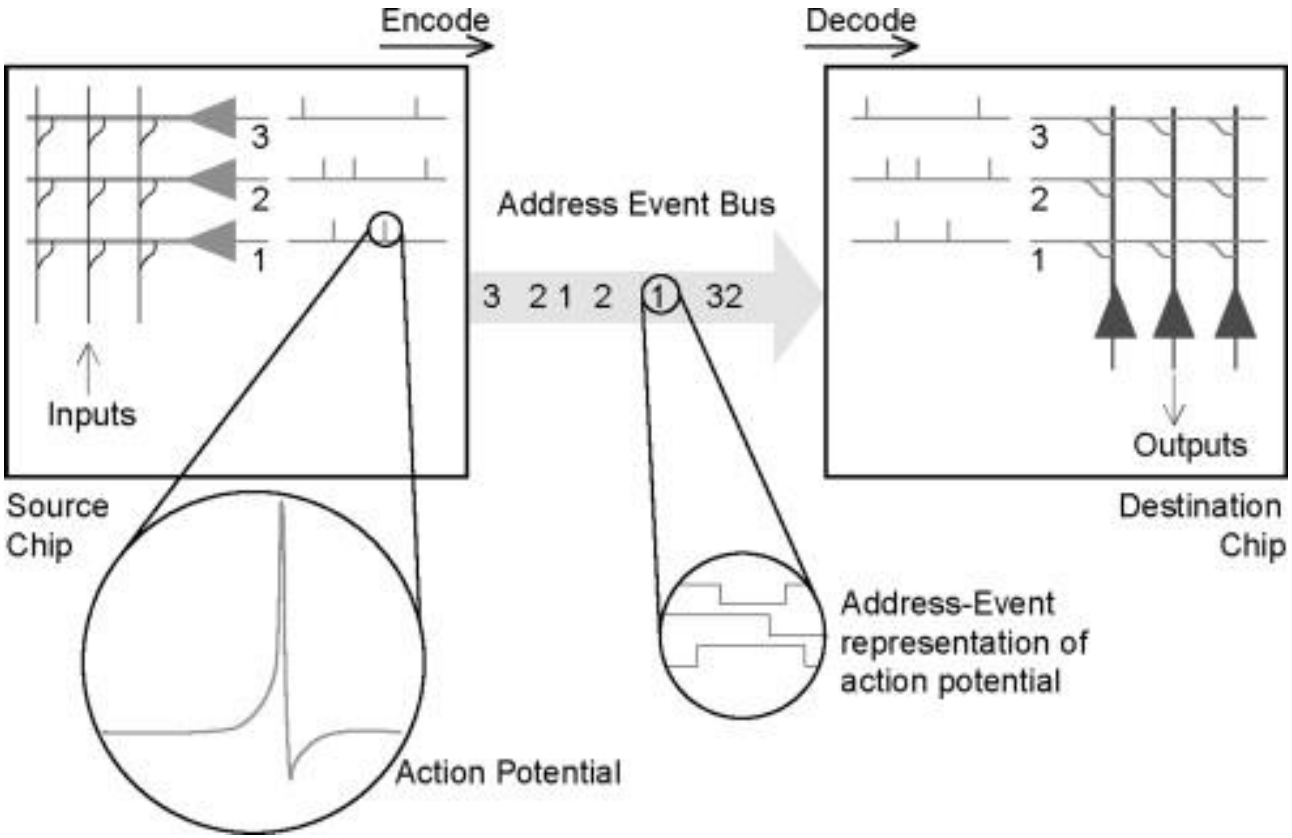


❖ Voltage **pulses** can be applied to a **memristor** to change its **resistance**, just as **spikes** can be applied to a **synapse** to change its **weight**.

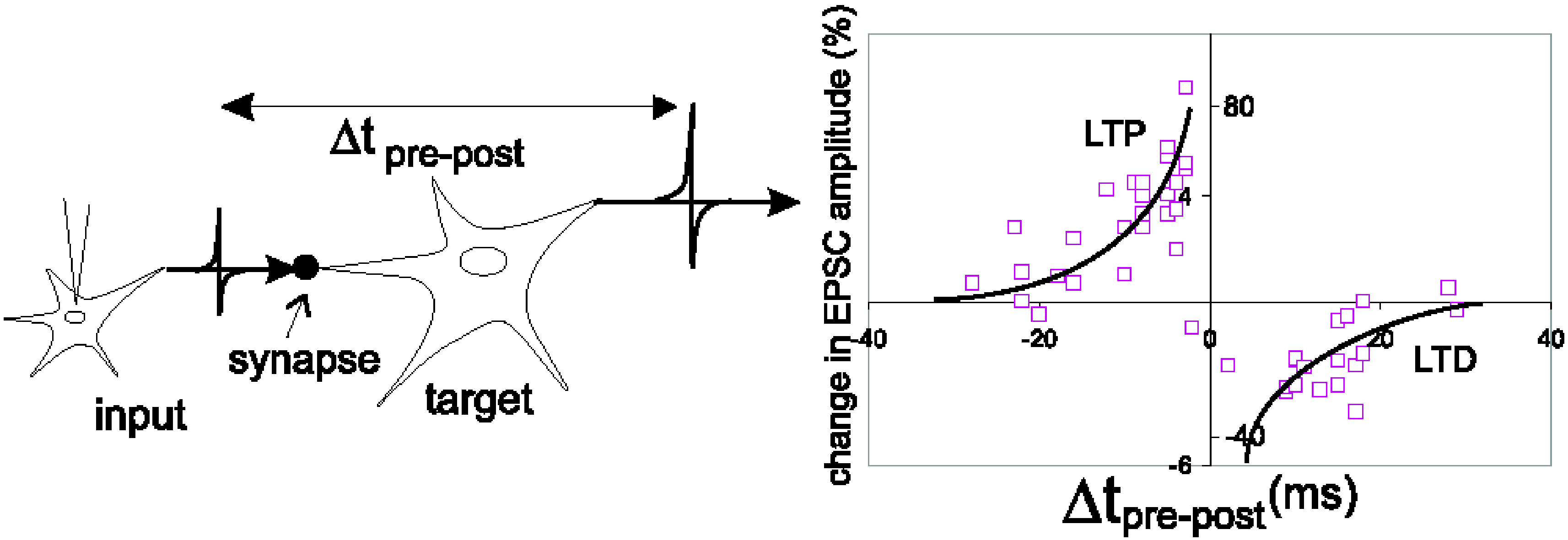
Wiring via AER (Asynchronous)



(Courtesy: iStock/Henrik5000)

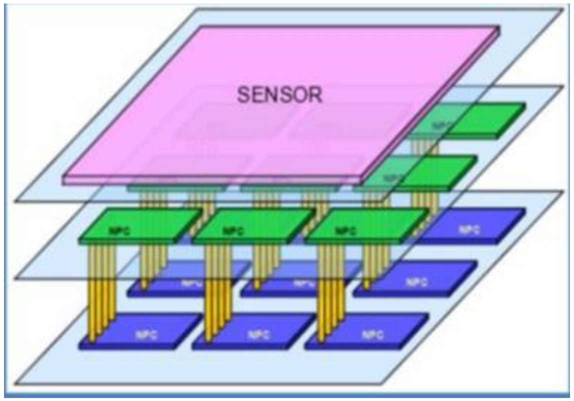


Spike-timing-dependent plasticity (STDP)

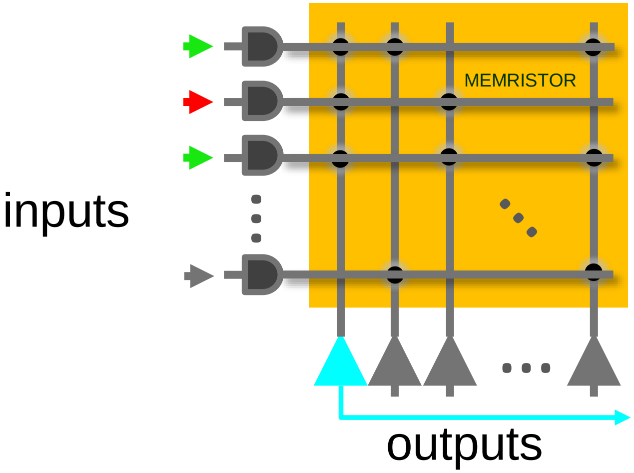


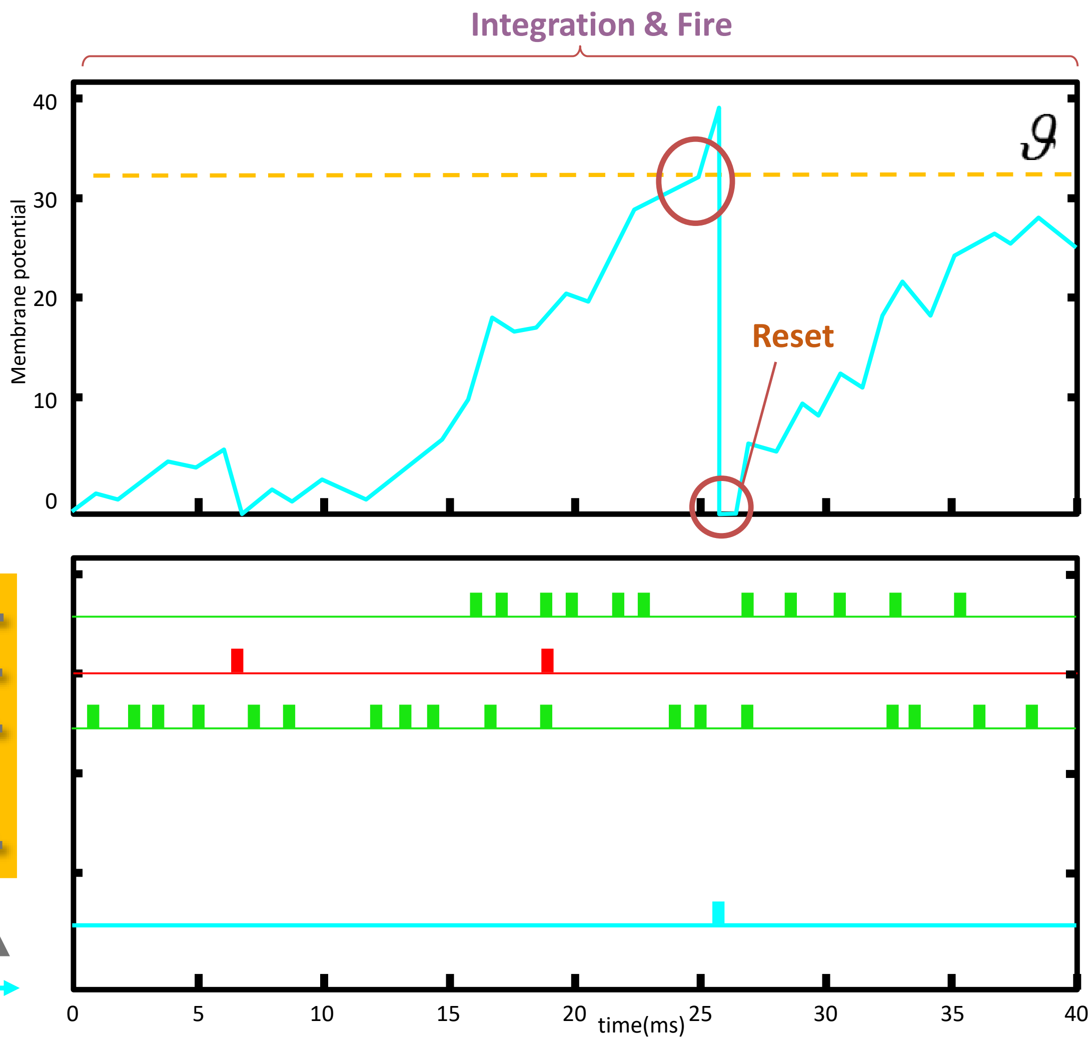
- Adjusts the strength of connections between neurons in the brain.
  - ✓ Adjusts the connection strengths based on the relative timing of a particular neuron's output and input action potentials.

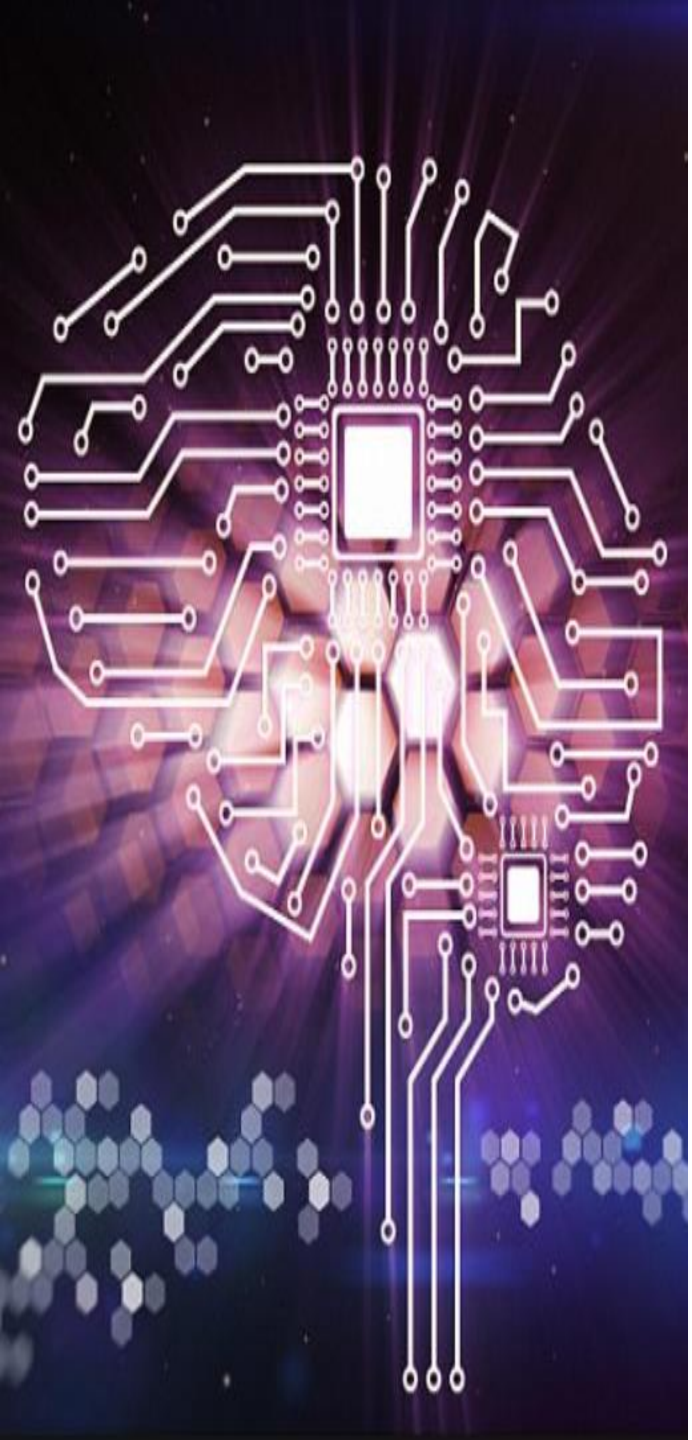
NASH: Neuro-inspired ArchitectureS in Hardware



NASH: Neuro-inspired ArchitectureS in Hardware





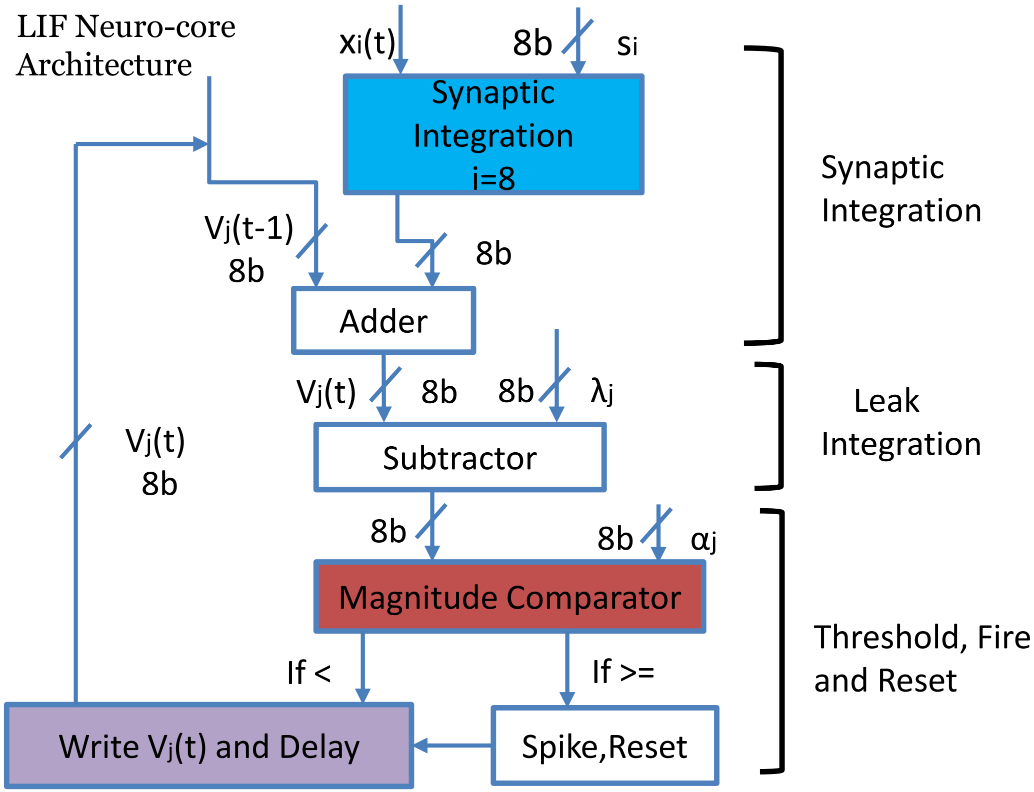


Outline

- Technology Transformation
- Neuron Modeling
- ASL Neuro-inspired Systems/Chips
- Concluding Remarks

LIF Neuro-core for NASH System

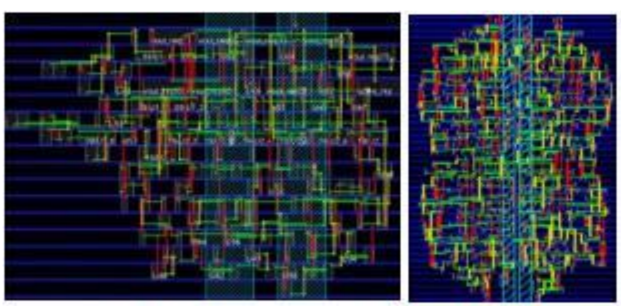
LIF Neuro-core Architecture



- Xi(t) – Spike input to the synapse
- Si – synaptic weight
- Vj(t) – Membrane potential
- αj – Neuron threshold
- Λj – Leak value

Table 1: Area Evaluation

Item	NC-1N	NC-4N
Cell Internal Power	6.9680 µW	20.5040 µW
Net Switching Power	4.8271 µW	14.8272 µW
Total Dynamic Power	11.7950 µW	35.3312 µW
Cell Leakage Power	4.6943 µW	14.3147 µW



LIF-1N-012018-KS LIF-4N-012018-KS

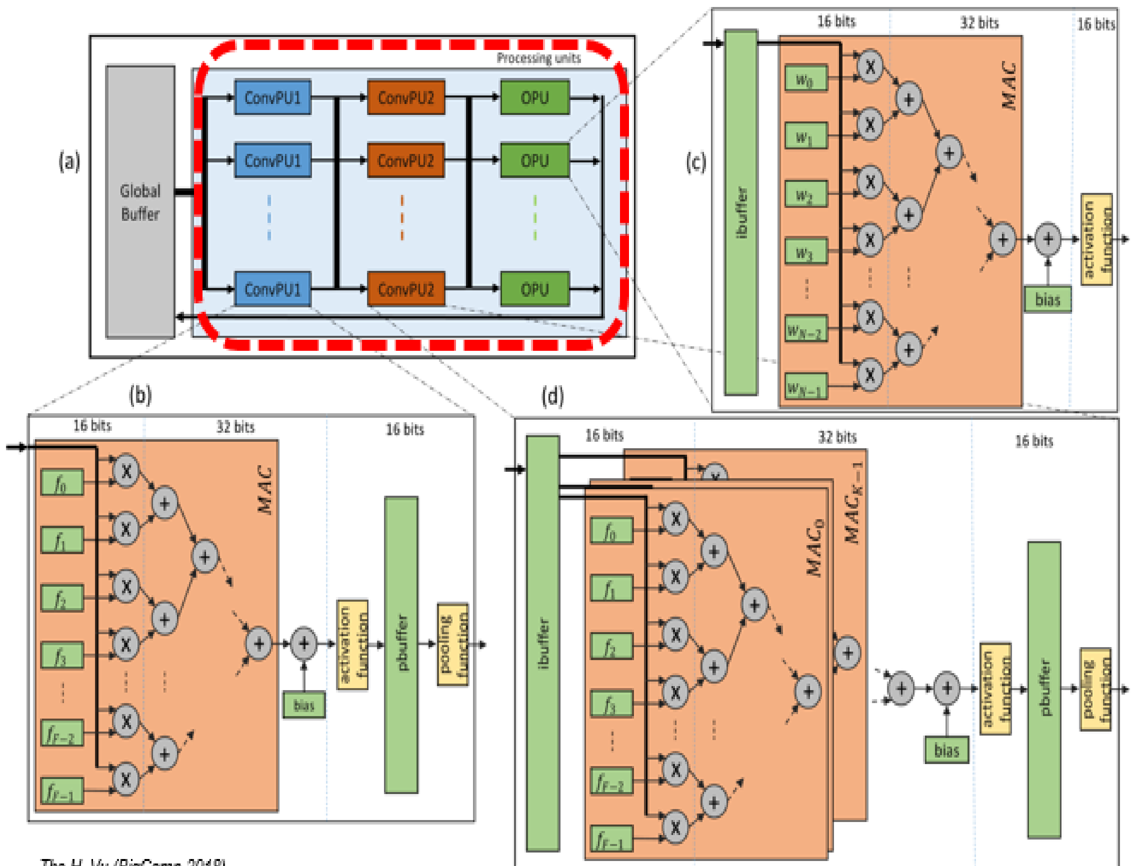
Placement of LIF-1N (Left) and LIF-4N (right)

Table 1: Power Evaluation

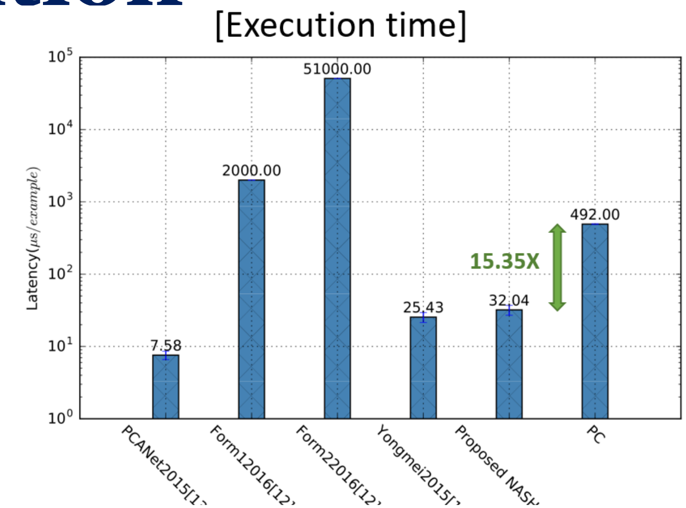
Item	NC-1N	NC-4N
Combinational Area	186.998 µm	562.856001 µm
Non-Comb Area	47.88002 µm	213.864000 µm
Total Cell Area	234.878002 µm	776.720001 µm

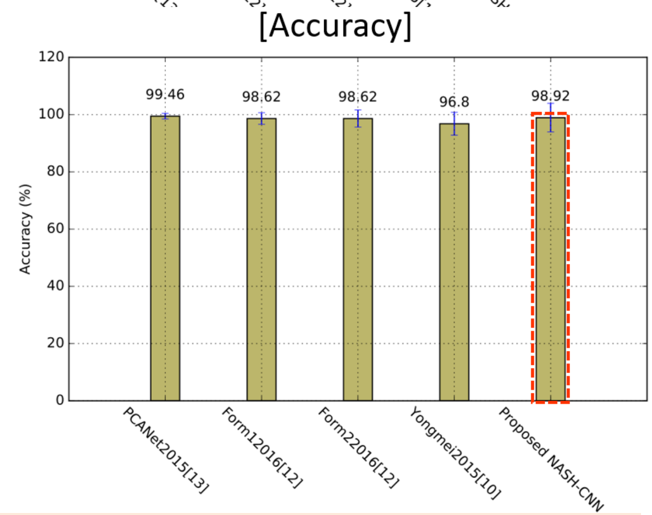
Kanta Suzuki, Yuichi Okuyama, Abderazek Ben Abdallah, "Hardware Design of a Leaky Integrate and Fire Neuron Core Towards the Design of a Low-power Neuro-inspired Spike-based Multicore SoC", Proc. Of IPSJ, 2018

Application I
Neuro-inspired Hardware System for Image Recognition



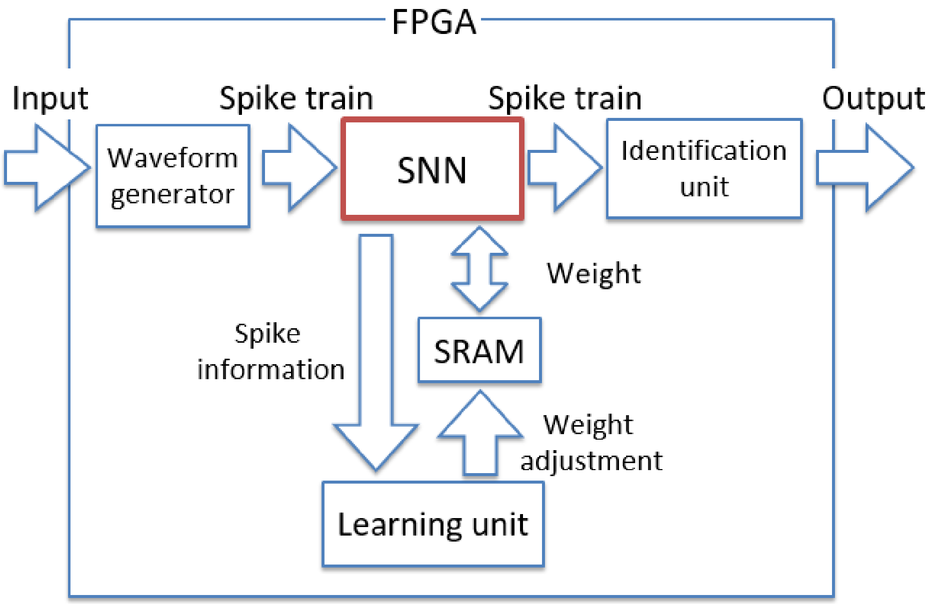
The H. Vu (BigComp-2018)





The H. Vu, Ryunosuke Murakami, Yuichi Okuyama, Abderazek Ben Abdallah, "Efficient Optimization and Hardware Acceleration of CNNs towards the Design of a Scalable Neuro-inspired Architecture in Hardware", Proc. of the IEEE International Conference on Big Data and Smart Computing (BigComp-2018), January 15-18, 2018

Application II
Neuro-inspired Hardware System for Autonomous Vehicles



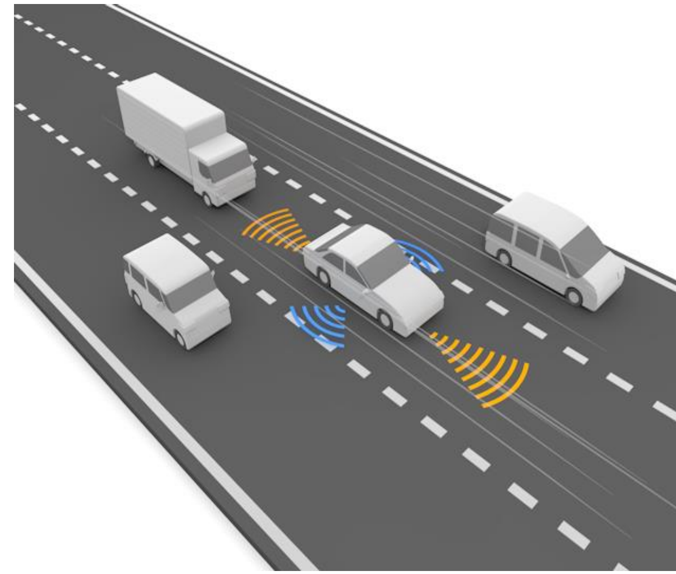
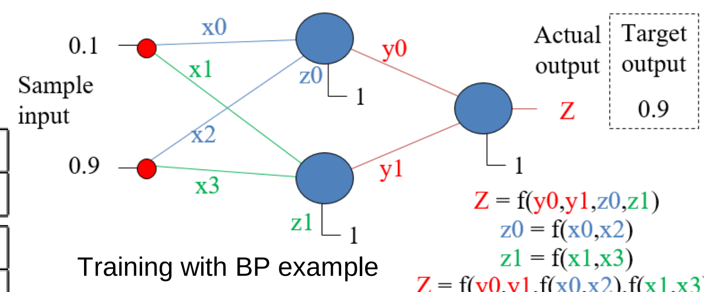


Figure 6: Spiking Neural Network System for Traffic-Light Recognition in Autonomous Vehicles.

Table 1 : ANN Performance Evaluation

ALUs	Registers	Pins	Fmax
10,989 (33%)	5,814 (18%)	432 (89%)	76.02 MHz
Memory	**DSP Block**	**Power Consumption**	
4,956 (1%)	54 (77%)	286.84 mW	

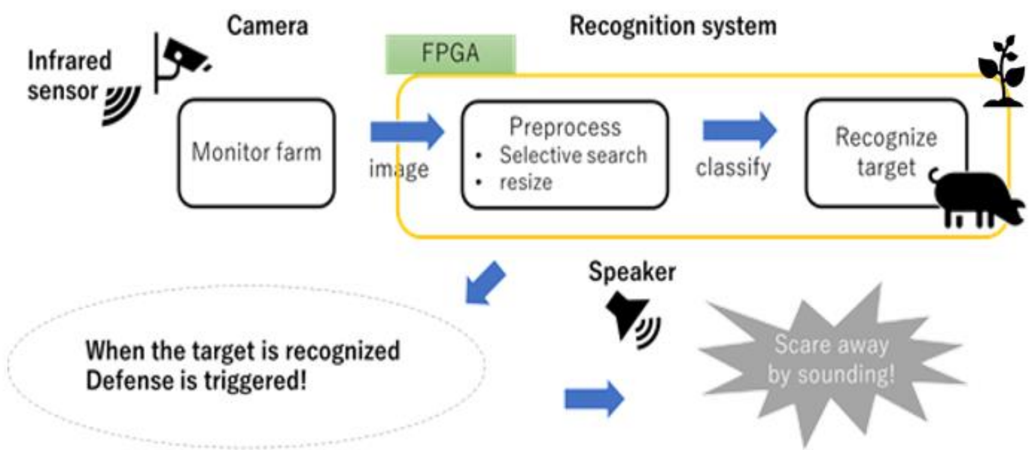


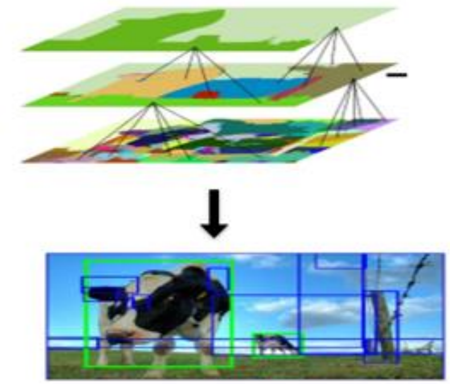
Training with BP example

Yuji Murakami, Yuichi Okuyama, Abderazek Ben Abdallah, "SRAM Based Neural Network System for Traffic-Light Recognition in Autonomous Vehicles", Information Processing Society Tohoku Branch Conference, Feb. 10, 2018

Application III

Neuro-inspired Hardware System for Visual Pattern Recognition in FARM Monitoring





出典:「Rich feature hierarchies for accurate object detection and semantic segmentation」

Fig 4. System overview: OASIS FMS-1

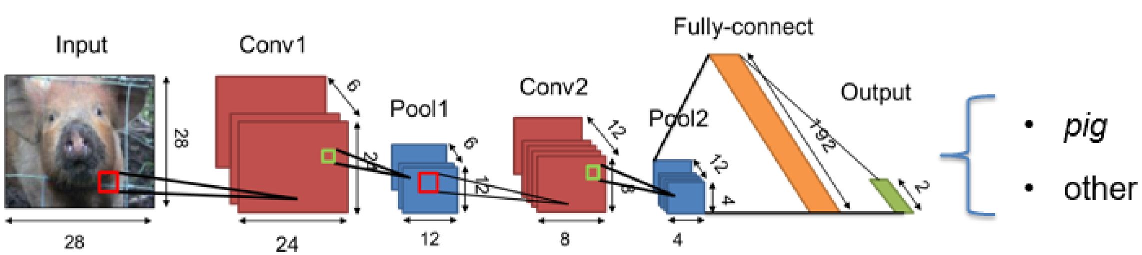


Fig 3. CNN example

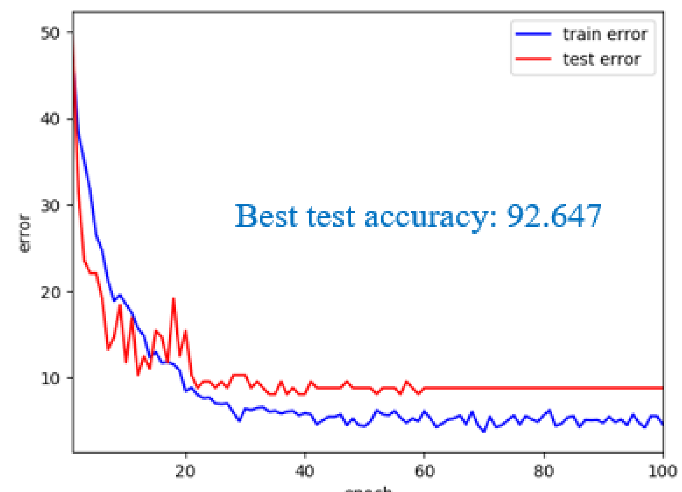


Fig 6. Learning result

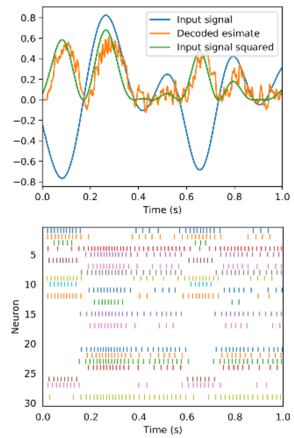
Ryunosuke Murakami, Yuichi Okuyama, Abderazek Ben Abdallah, "Animal Recognition and Identification with Deep Convolutional Neural Networks for Farm Monitoring", Information Processing Society Tohoku Branch Conference, Feb. 10, 2018

Application IV
Brain-inspired Drone Control with BCI

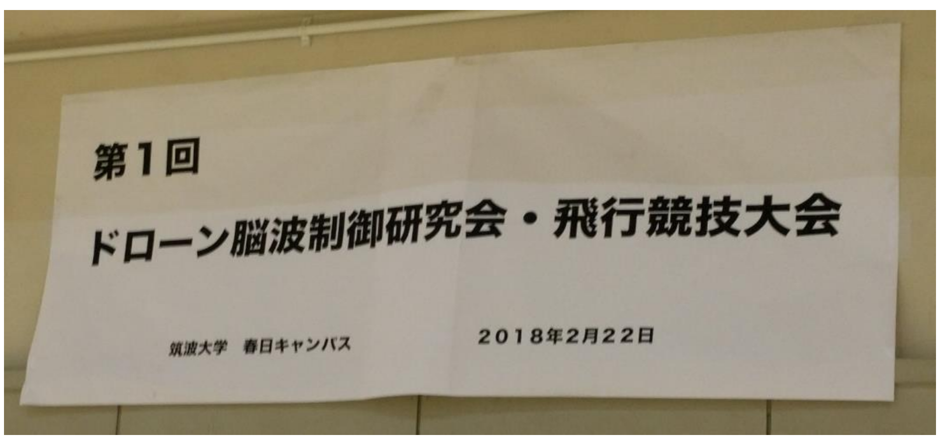
Flight navigation logic (in C++/Python)

Emotiv community SDK	Parrot ARDrone SDK	Other libraries
Facial command Mental command Raw EEG Gyro	Drone control Video inputs Sensor inputs	Spiking neurons Neural networks SLAM (mapping)

Brain to Brain drone system



Numerical computation with SNNs



Conclusion & References

1. Khanh N. Dang, Akram Ben Ahmed, Yuichi Okuyama, and Abderazek Ben Abdallah, ”Scalable Design Methodology and Online Algorithm for TSV-cluster Defects Recovery in Highly Reliable 3D-NoC Systems”, IEEE Transactions on Emerging Topics in Computing, 2017 (in press). DOI: 10.1109/TETC.2017.2762407
2. Achraf Ben Ahmed, Tsutomu Yoshinaga, Abderazek Ben Abdallah, “Scalable Photonic Networks-on-Chip Architecture Based on a Novel Wavelength-Shifting Mechanism”, IEEE Transactions on Emerging Topics in Computing, 2017 (in press). DOI: 10.1109/TETC.2017.2737016
3. Khanh N. Dang, Akram Ben Ahmed, Xuan-Tu Tran, Yuichi Okuyama, Abderazek Ben Abdallah, ”A Comprehensive Reliability Assessment of Fault-Resilient Network-on-Chip Using Analytical Model”, IEEE Transactions on Very Large Scale Integration (VLSI) Systems, Vol. 25, Issue: 11, pp. 3099 – 3112, Nov. 2017. DOI:10.1109/TVLSI.2017.2736004
4. The H. Vu, Ryunosuke Murakami, Yuichi Okuyama, Abderazek Ben Abdallah, ”Efficient Optimization and Hardware Acceleration of CNNs towards the Design of a Scalable Neuro-inspired Architecture in Hardware”, Proc. of the IEEE International Conference on Big Data and Smart Computing (BigComp-2018), January 15-18, 2018.
5. Ryunosuke Murakami, Yuichi Okuyama, Abderazek Ben Abdallah, ”Animal Recognition and Identification with Deep Convolutional Neural Networks for Farm Monitoring”, Information Processing Society Tohoku Branch Conference, Feb. 10, 2018
6. Yuji Murakami, Yuichi Okuyama, Abderazek Ben Abdallah, ”SRAM Based Neural Network System for Traffic-Light Recognition in Autonomous Vehicles”, Information Processing Society Tohoku Branch Conference, Feb. 10, 2018.
7. Kanta Suzuki, Yuichi Okuyama, Abderazek Ben Abdallah, ”Hardware Design of a Leaky Integrate and Fire Neuron Core Towards the Design of a Low-power Neuro-inspired Spike-based Multicore SoC”, Information Processing Society Tohoku Branch Conference, Feb. 10, 2018.

Thank you!

Ben Abdallah Abderazek

***A**daptive **S**ystems **L**aboratory*

benab@u-aizu.ac.jp

to Advance Knowledge for Humanity