CEEDS
미래전자소자 연구센터
Center for Emerging Electronic Devices and Systems

US-KOREA NANO FORUM
SEOUL, SEPTEMBER 26, 2016

Brain Inspired Semiconductor Device Technology

Byoung Hun Lee

Director, Center for Emerging Electronic Materials and Systems (CEEDS)
School of Material Science and Engineering,
Gwangju Institute of Science and Technology (GIST)

Gwangju Institute of Science and Technology

QS World University Rankings
(Citations per Faculty, 2015/16)
2nd in the world

SCI papers
(Papers per faculty / 18 years, 1995-2016)
1st in the nation

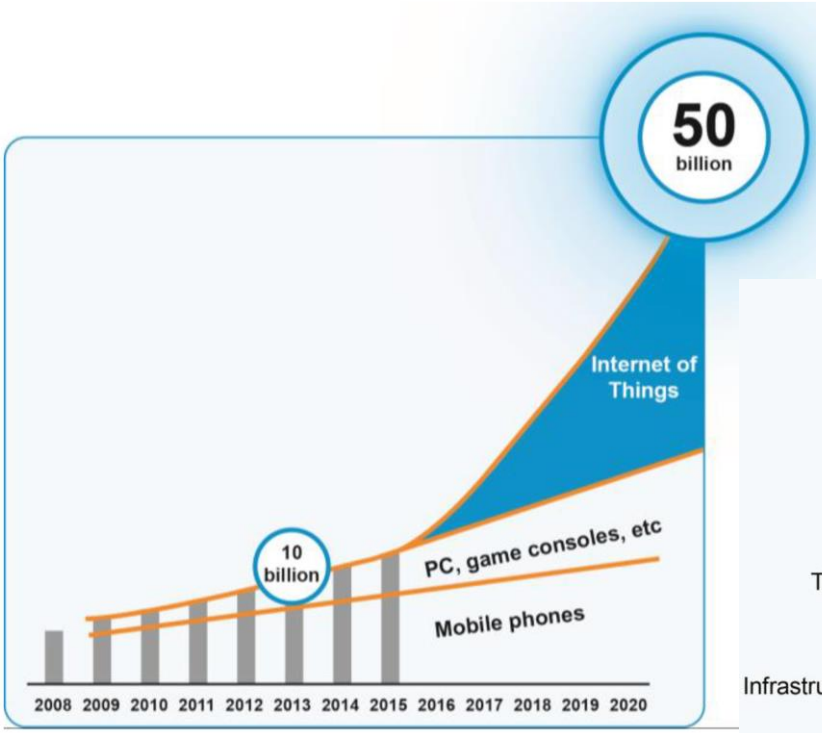
Patents, Research grant
(Oversea patent application)
1st in the nation

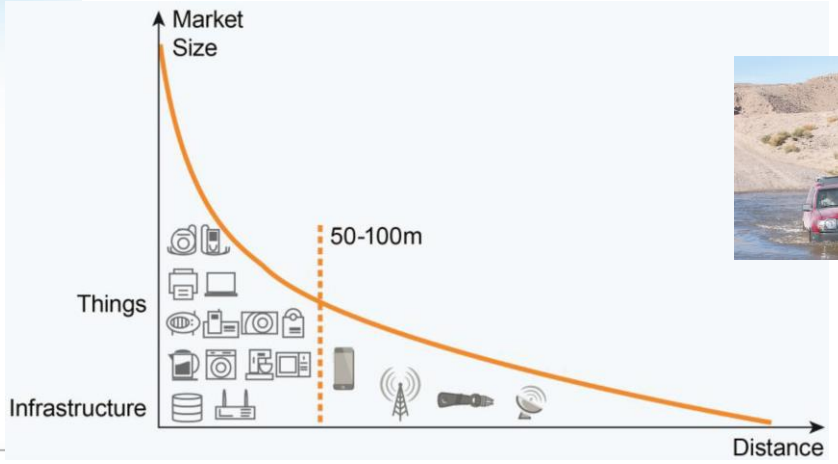
THE in 2015 (<50years)
33nd in the world

Outline

- Motivation: End of roadmap and energy crisis
- Motivation: Inspirations from brain functions
- Bio inspired semiconductor device technology
- Summary

Information processing needs in future

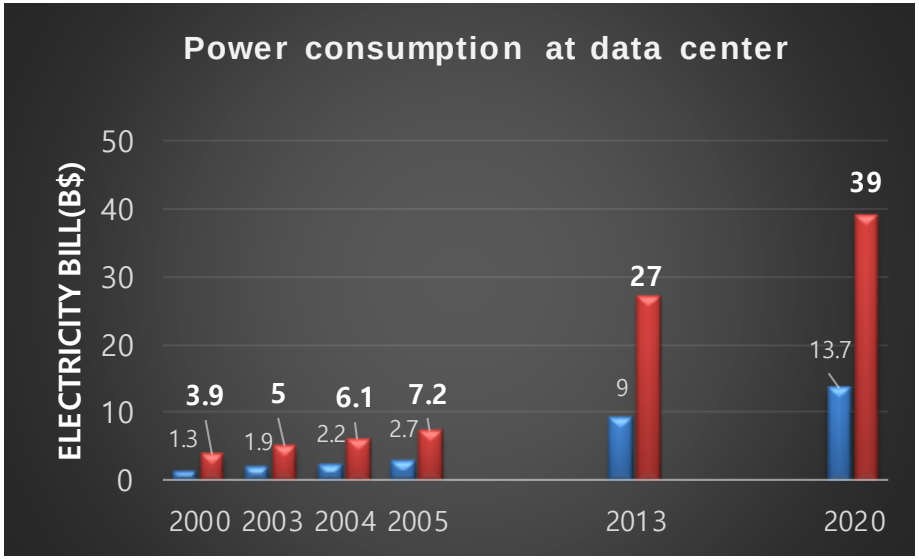




- The number of PC, cellphone etc double in every 5 years
- IoT will increase this rate more than double (and market size too)
- IoT may not require a breakthrough in technology, but a real breakthrough in nanoelectronics will be necessary to sustain the IoT based society

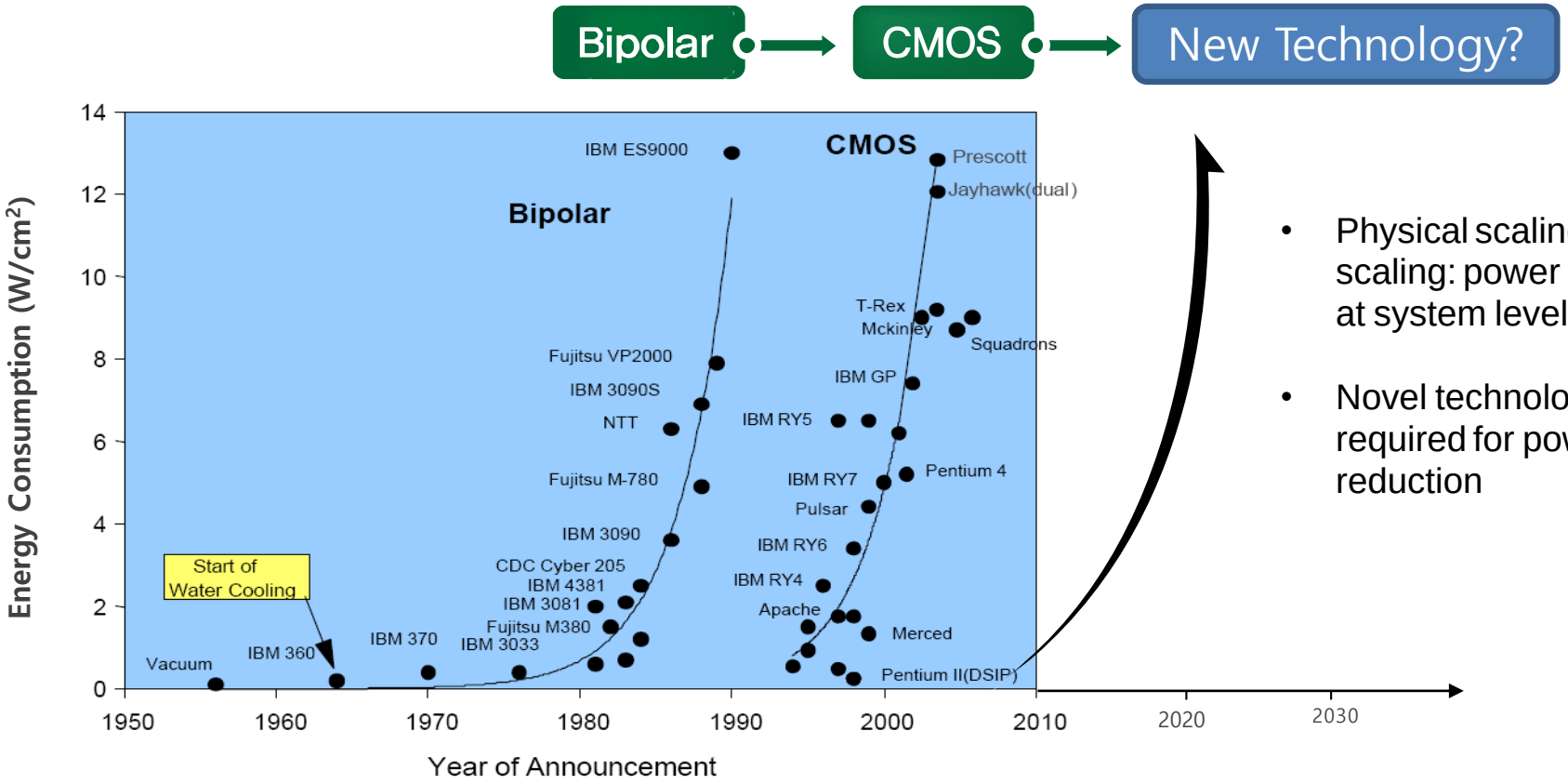
Mike Annderson, "Short-range Low Power Wireless Devices and Internet of Things (IoT)", Connect Blue Inc., Tech. reports, 2014.

Information processing requires HUGE power!



- In 2005, 27M server consumed 14GW (7B$), ~ 14 nuclear power plant
- In 2015, 28B$ will be spent with ~4% of total electricity worldwide
- In 2030, more than 100B$ (180 nuclear power plant) will be necessary for data center alone with the accelerated usage of data for IoT
- Including PC, cellphone and other wearable, portable ICT equipment, the energy consumption will not be sustainable with current technology

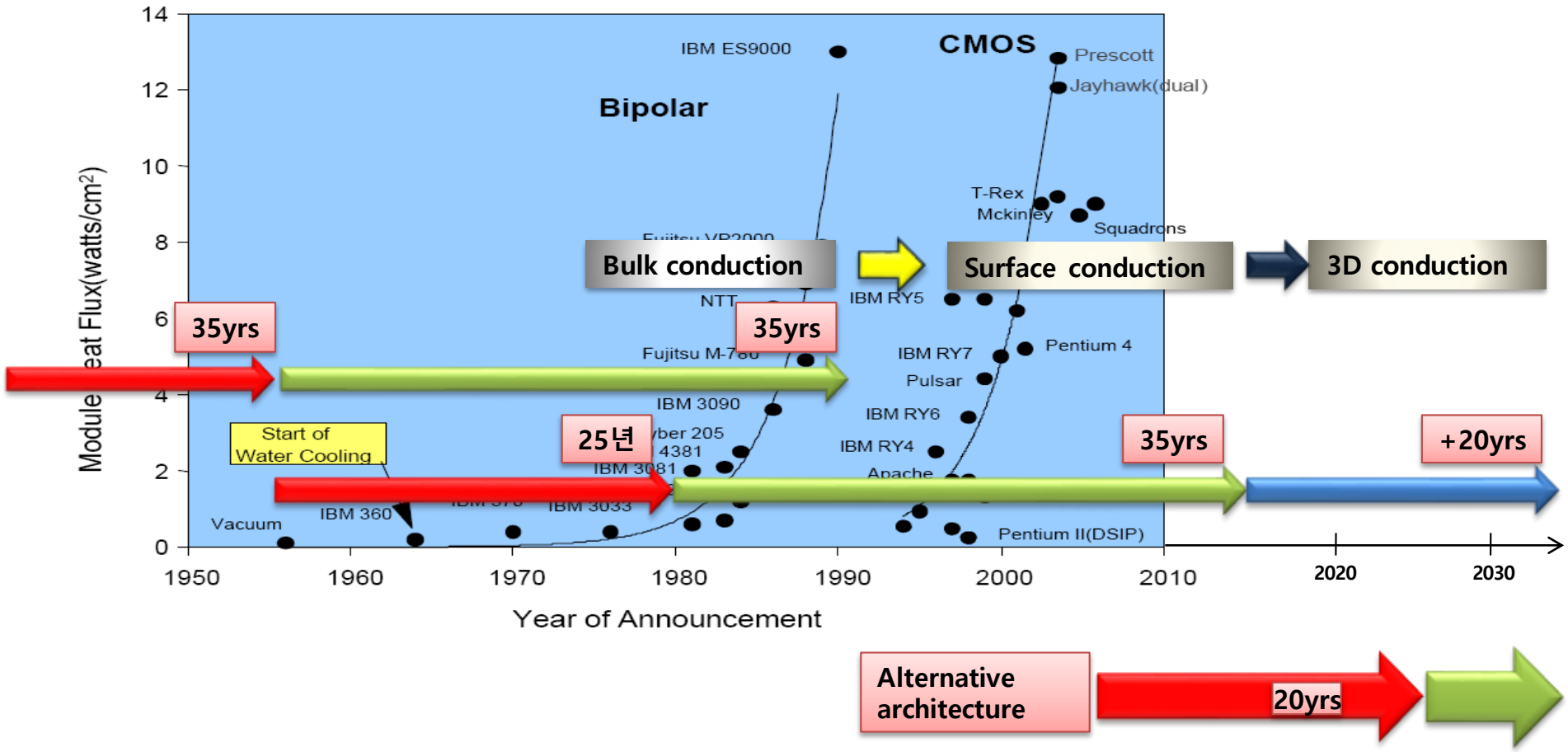
End of roadmap



- Physical scaling to power scaling: power reduction at system level
- Novel technology is required for power reduction

- CMOS technology was conceived in 1960.
- Do we have the new technology that can replace CMOS technology?

Alternative target of research



- Conservative goal: Use semiconductor baseline techology, Solve the power problem

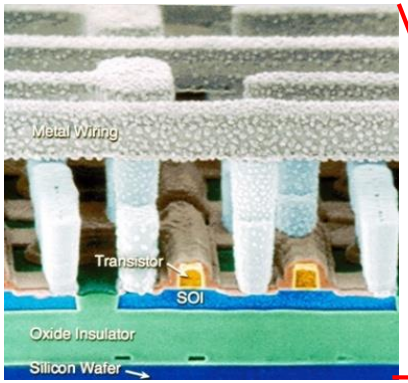
Short term solution: low power device technology

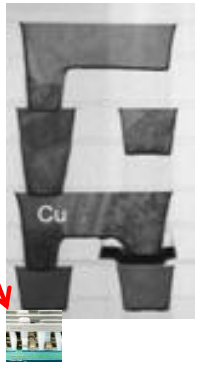
Multi V_{DD}
Dynamic V_{DD} scaling
V_{DD} scaling using finFET, NWFET
Low V_{DD} channel (III-V, tFET)

Multi-V_{th}
Dynamic V_{th} scaling
Low V_{th} channel (III-V, SiGe)

Multi core
External cooling
Hot spot management

$$P = \alpha C_{tot} V_{DD}^2 f + V_{DD} \cdot I_{leak} \left(e^{-V_{th}/(kT/q)} \right)$$



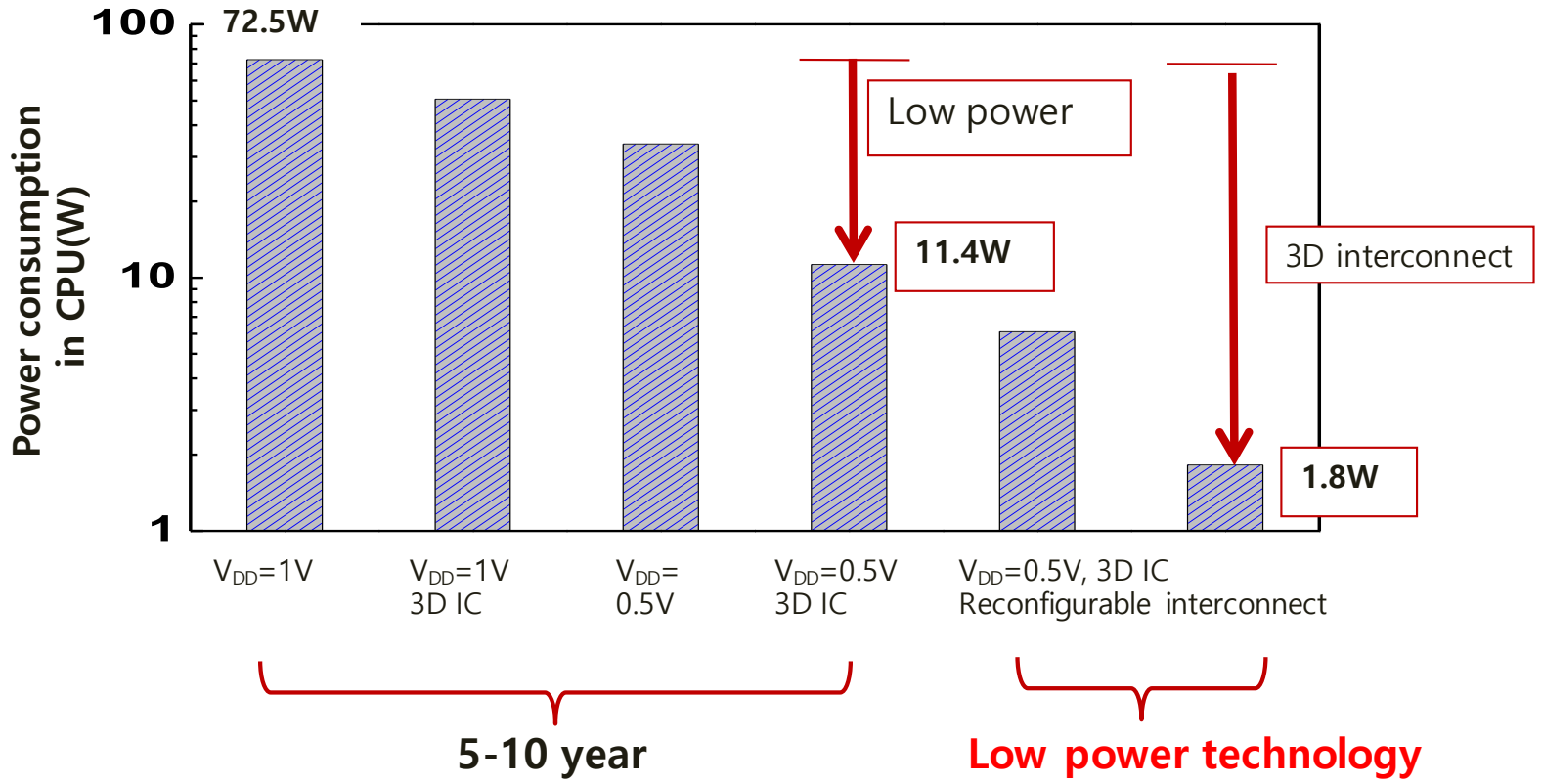


Low V_{DD}
Low leakage gate stack
Novel device (Underlapped SOI,..)
Hybrid device (NEMS-CMOS)

Clock gating
Sleep transistors
Fine granulity clock gating
Architectural switching reduction

- So far, we are trying old technologies to save power: low VDD, Design optimization

Low power device technology



- Low power device technology: power savings <1/10 ~ 1/100
- Extreme low power technology: Power savings <1/100 -1/1000

Summary

1. Problem statement
   - Future computing/ information processing needs far exceed the capacity of current energy supply
   - Key challenge is the total power consumption of system (performance might be a secondary issue with a parallel and cloud computing route)

2. Origin of problem
   - Power consumption in a semiconductor chip is reaching its limit
   - Rack of innovation: scaling for half century (no new device, architecture)

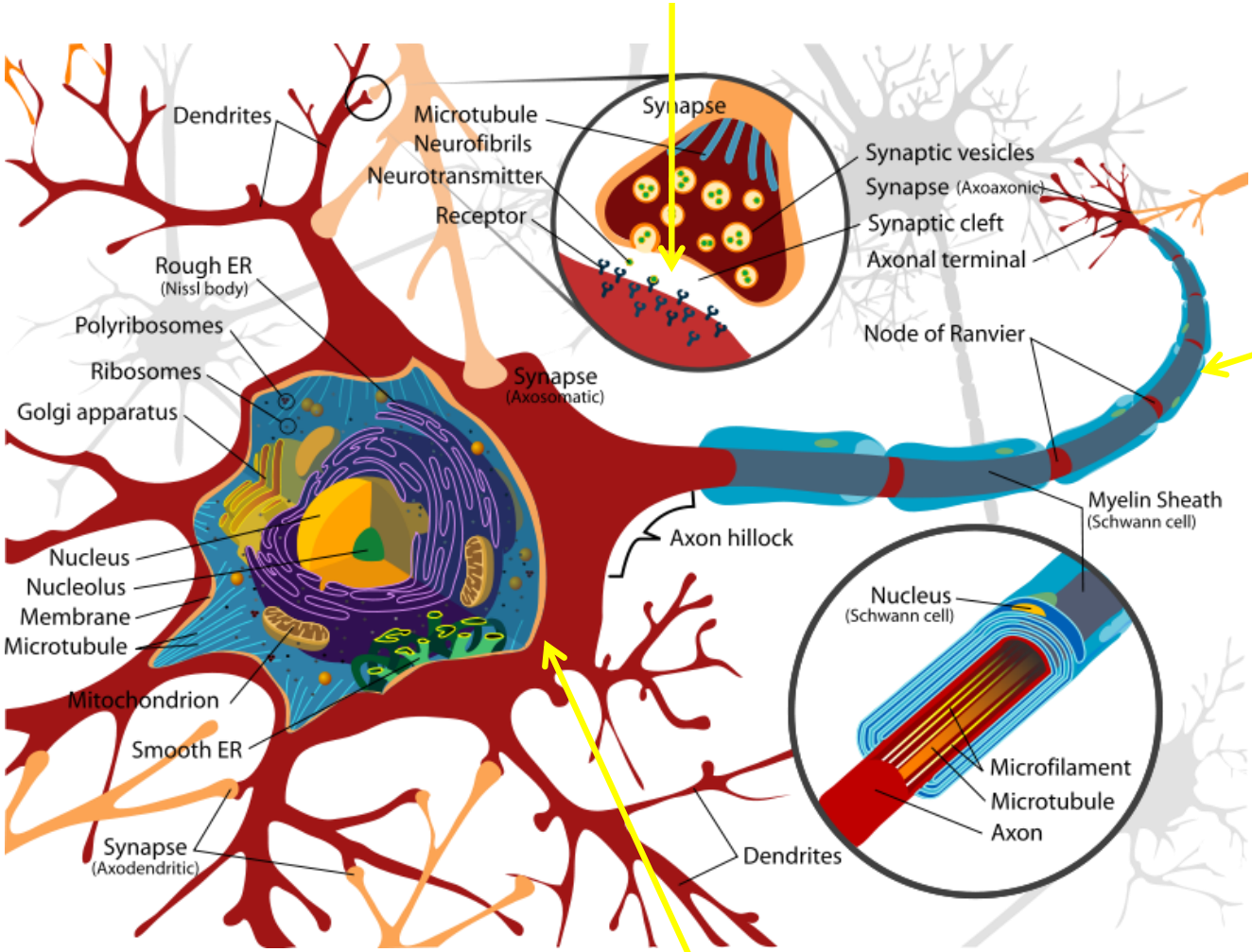
3. Constraints
   - There is not enough time to fix this problem and get back on track: implementation of new device and architecture may take more than 20 years easily
   - Need to solve the problems with existing tools and live with it till full solutions are available

Outline

- Motivation: End of roadmap and energy crisis
- Motivation: Inspirations from brain functions
- Bio inspired semiconductor device technology
- Summary

Physical structure of neuron and its functions

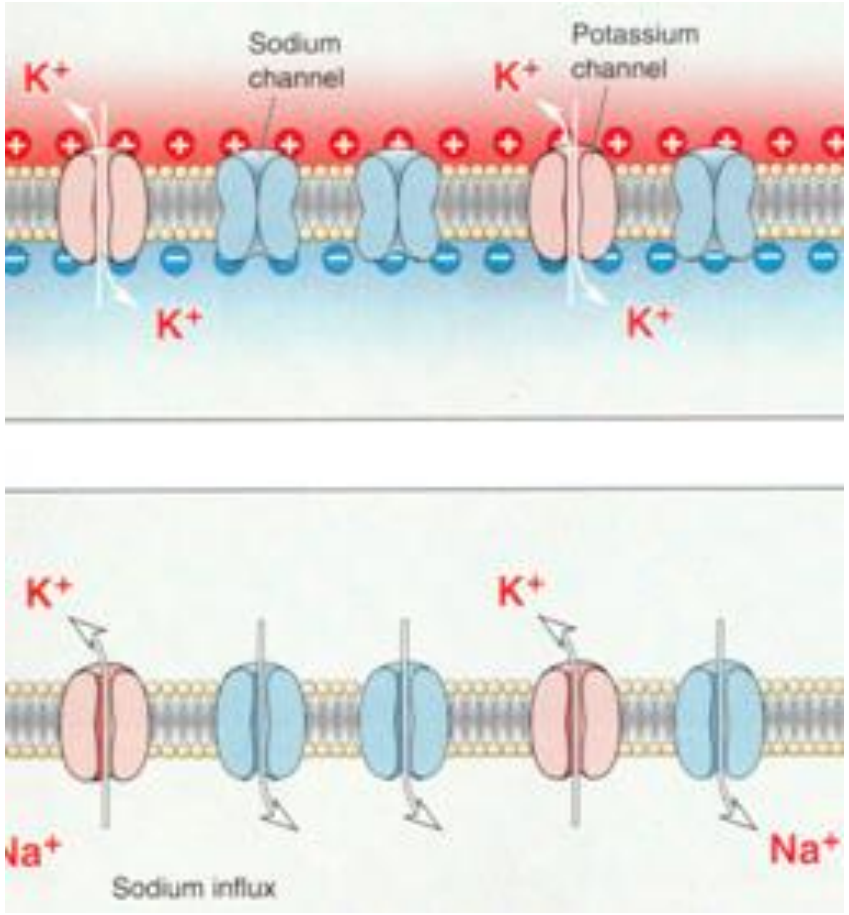
Adaptation, learning, STDP" 1sec- years



Temporal coding,
Signal propagation:
10-100msec

Membrane dynamics, action potential
generation: 1msec

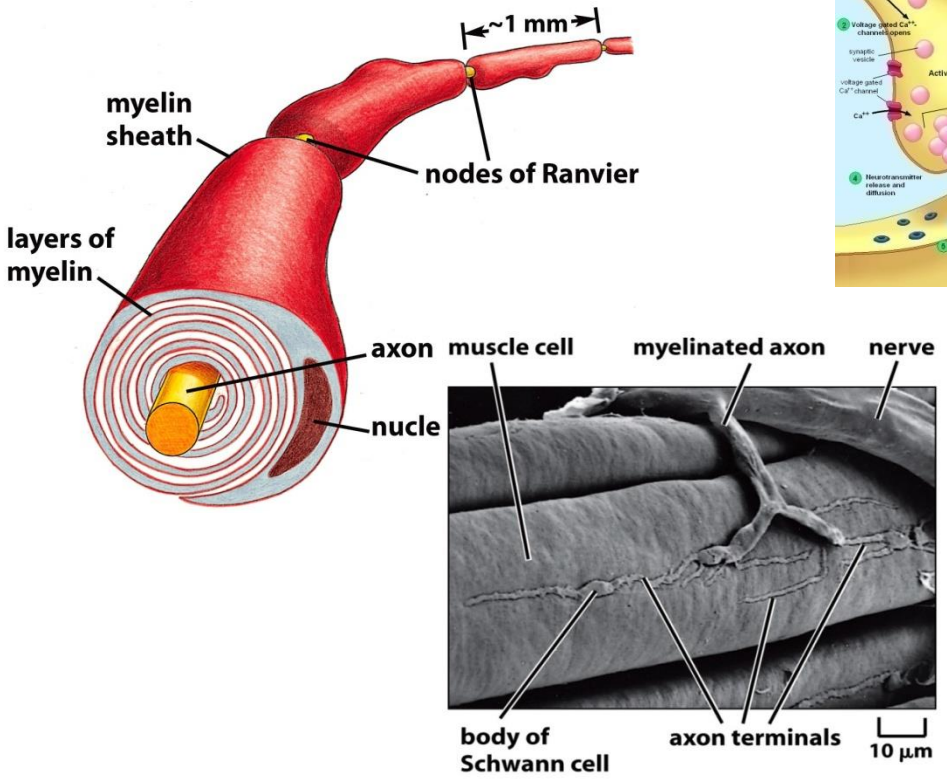
Electronics device vs chemical device

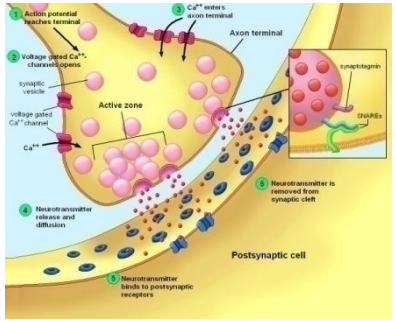


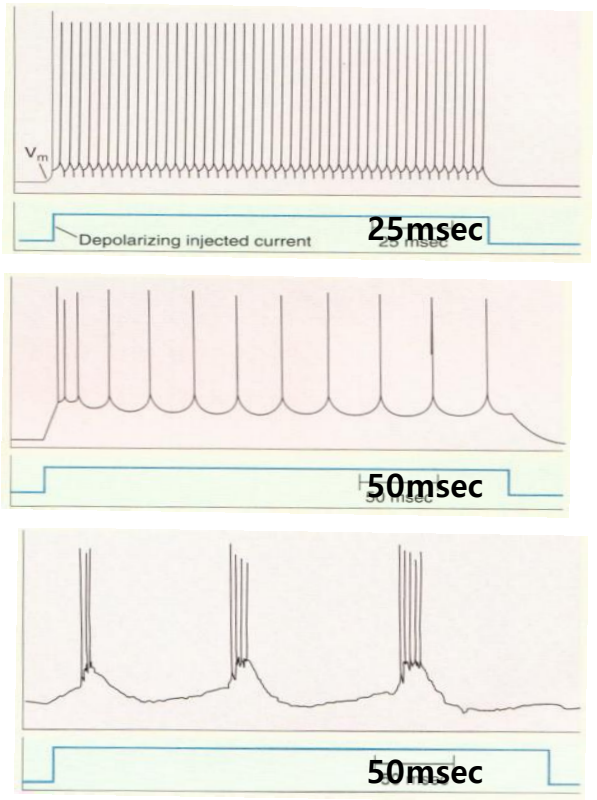
- Deliver the concentration difference of K+, Na+
- Action potential ~ 80 mV → Extreme low voltage operation
  → Noise problem
  → **Multiple signal input/ integration**
- Spatial and temporal multiplexing → Active sharing of the interconnect

- Chemical computing, extremely low operation voltage (<100mV)

Slow, but powerful signal processing

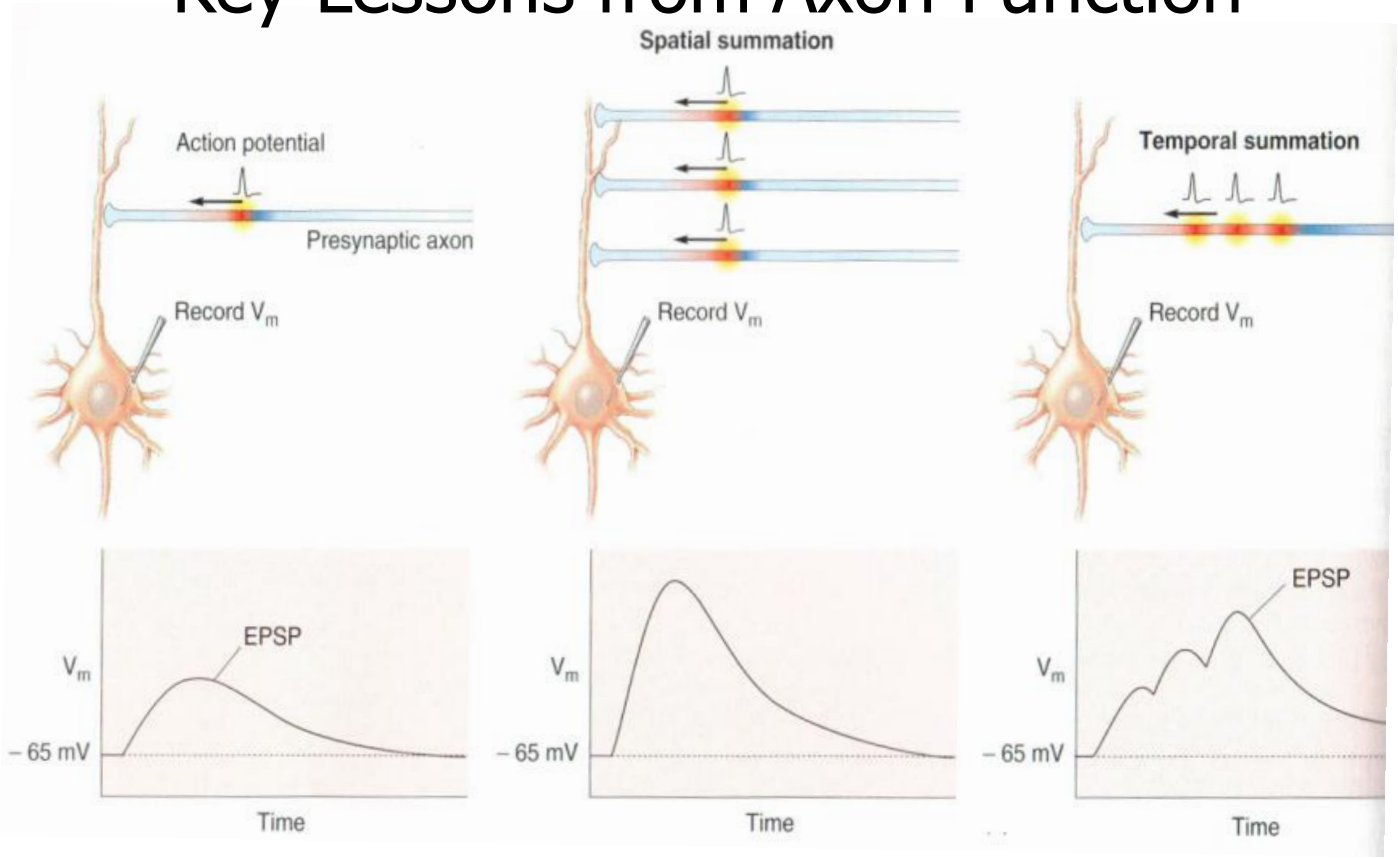






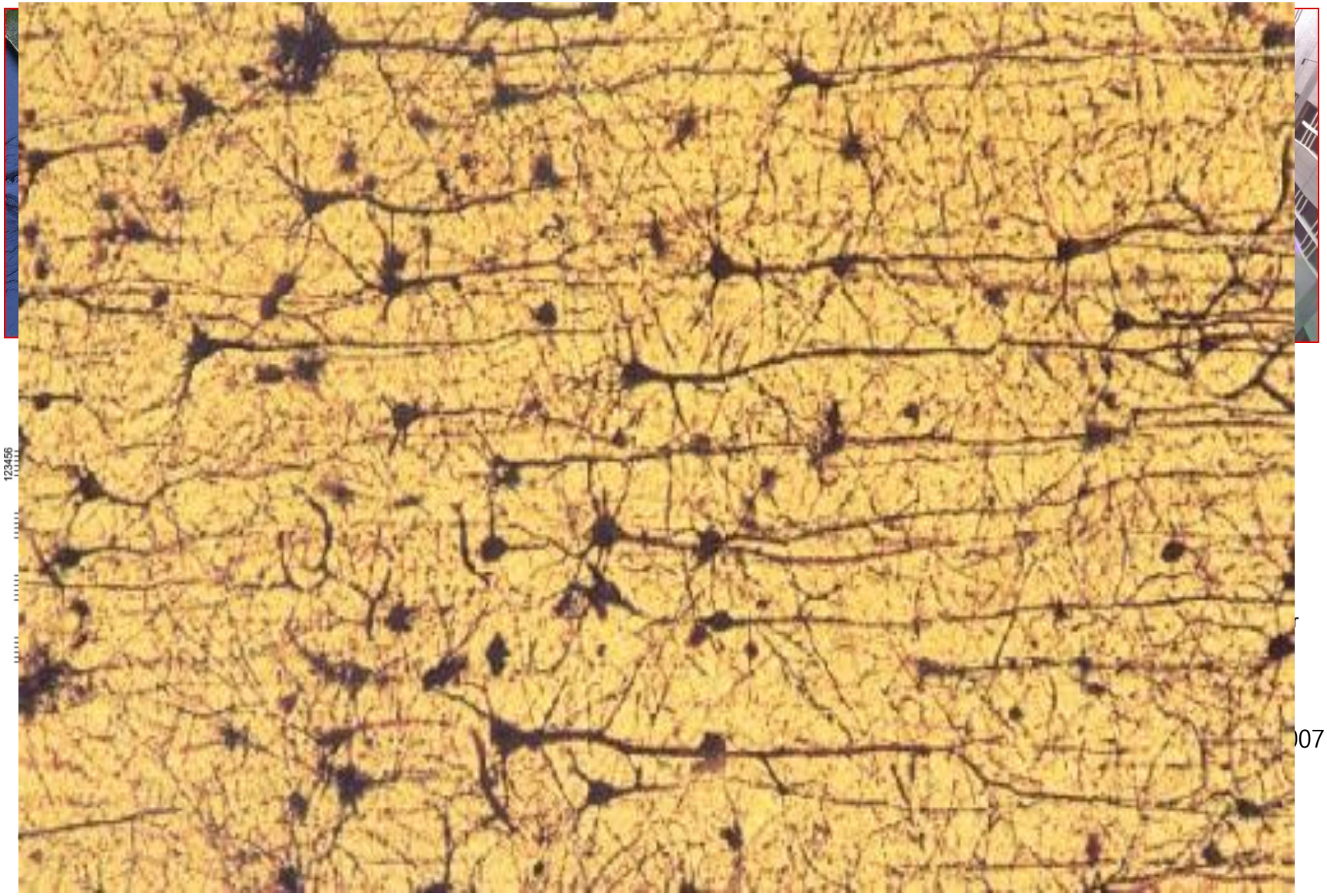
- Human: 10m/sec (0.2 sec from head to foot), multiple signal/ parallel processing , distributed computing
- Semiconductor : 20nsec from head to foot, 1 signal at a time, central processing
- Stress time dependent plasticity (STDP): logic/memory function

Key Lessons from Axon Function



- Temporal/Spatial summation → Logic/memory operation at interconnect
- Time multiplexing → Dynamically reconfigurable interconnect
- Extreme fan out (order of 10^4 vs 10^1): Reconfigurable, multifunctional circuit
- Temporal/spatial synchronization: Noise/defect tolerant signal processing
- Plasticity and multiplexing : reconfigurable architecture

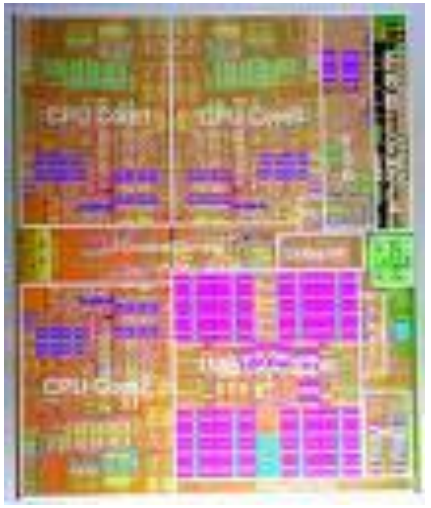
Vertically stacked architecture



Nature's way of power management

$$Power \propto \alpha C_{tot} V_{DD}^2 f + V_{DD} \cdot I_{leak} \left(e^{-V_{th}/(kT/q)} \right)$$

- Slow operation (f)
- Vertical stacking (C_{tot})
- Chemical computing (analog computing, I_{leak}, V_{DD})
- Multi valued logic (V_{DD}, C_{tot})
- Logic and memory functions are mixed
- Distributed memory (local memory): multicore



- **Calculator vs information processor**

Outline

- Motivation: End of roadmap and energy crisis
- Motivation: Inspirations from brain functions
- Bio inspired semiconductor device technology
- Summary

Inspirations to semiconductor technology

- **Low C_{tot}**
  - Device count reduction : CMOS-hybrid devices
  - Dynamically reconfigurable 3D integrated logic
  - Interconnect length reduction : Vertical Integration, multi value, multiband, multi-flexing, reconfigurable interconnect
  - Non capacitive coupling: Optical interconnect and I/O, Magnetic quantum cellular automata, Molecular wire
- **Low Frequency:** multicore, neuromorphic architecture, spatial multiplexing,
- **Low V_{DD} :** Noise / fault tolerant computing device/architecture

Bio inspired architecture is more than enough to take us for next a few decades of new electronics!!

Midterm: Bioinspired Extreme Low Power Technology

Neuromorphic

Multi-valued logic

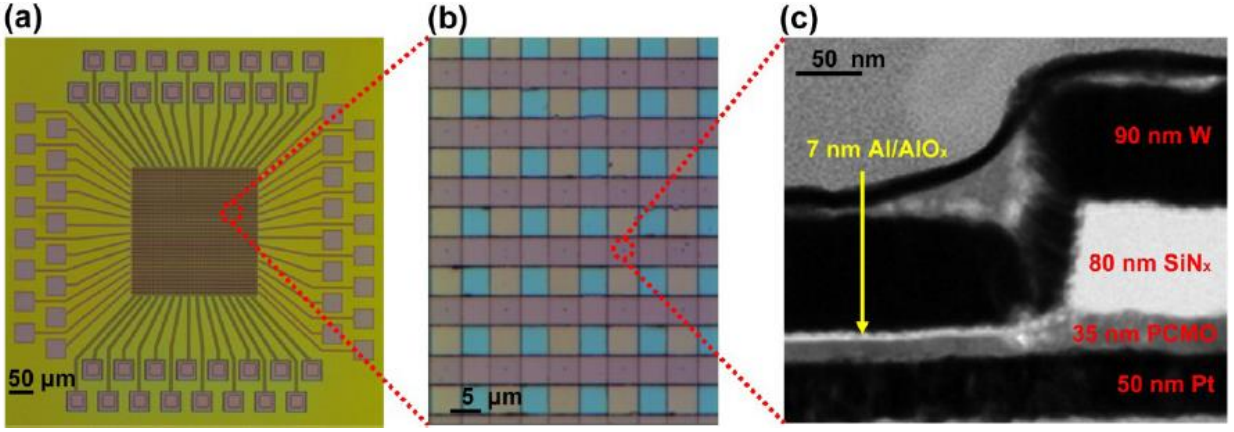
Reconfigurable logic

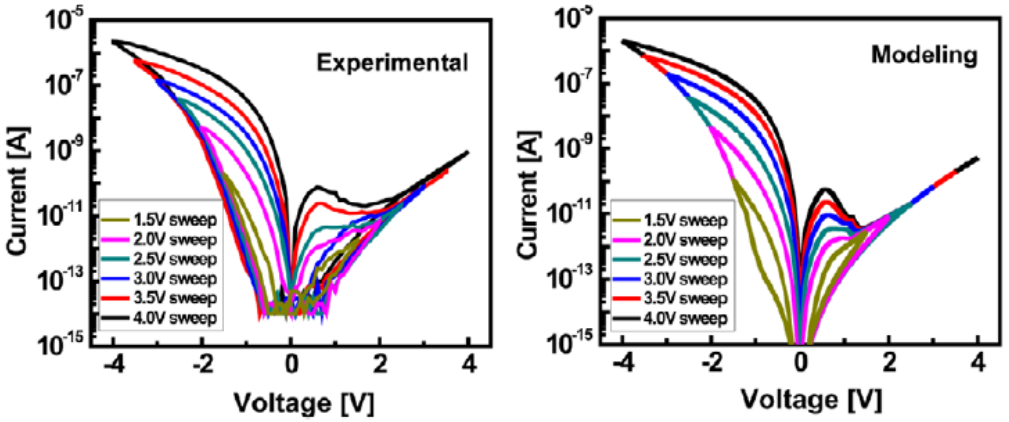
Monolithic 3D integration

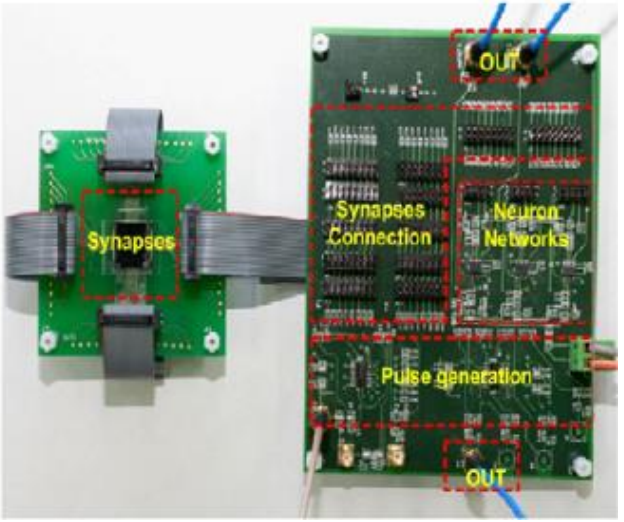
- Same functionality with smaller number of device
- Reduction of interconnect length with higher information density

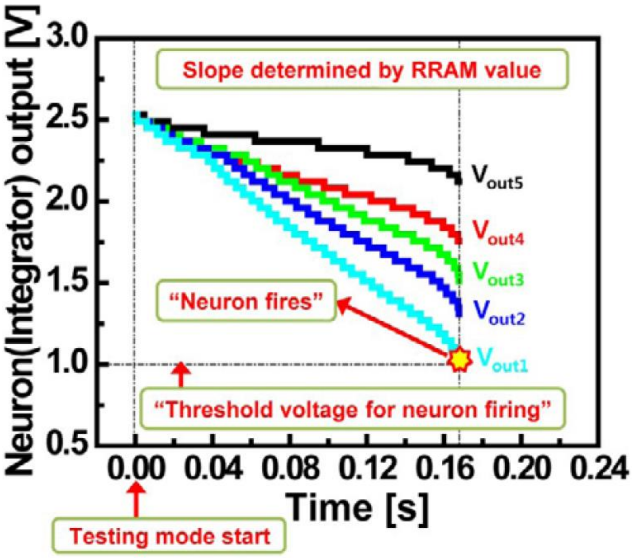
IOP PUBLISHING NANOTECHNOLOGY
Nanotechnology **24** (2013) 384009 (6pp) doi:10.1088/0957-4484/24/38/384009

Nanoscale RRAM-based synaptic electronics: toward a neuromorphic computing device

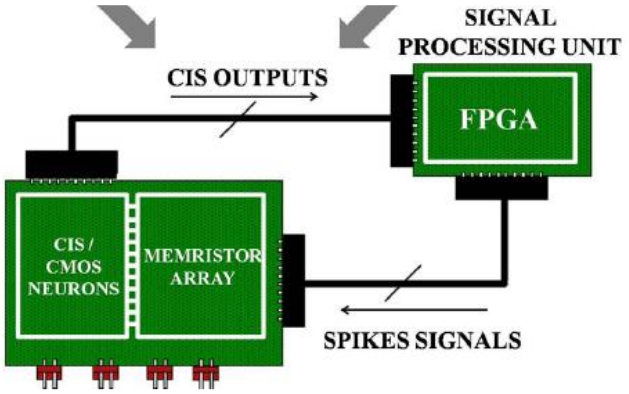


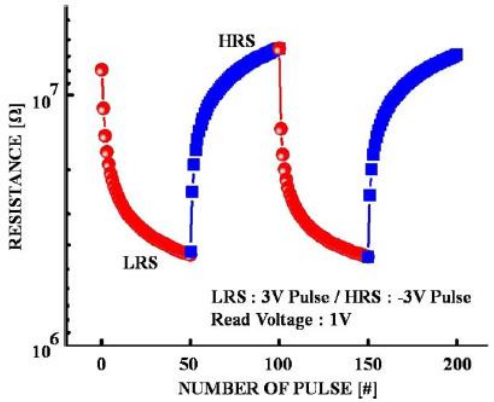


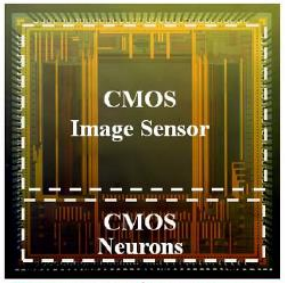


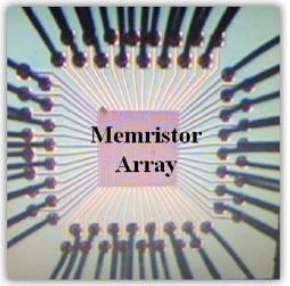


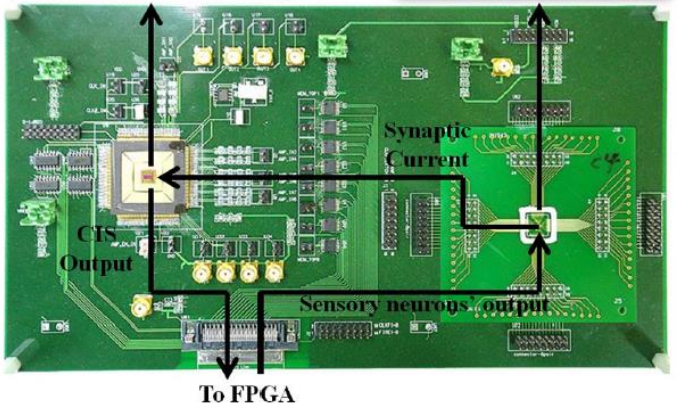
Neuromorphic Hardware System for Visual Pattern Recognition With Memristor Array and CMOS Neuron

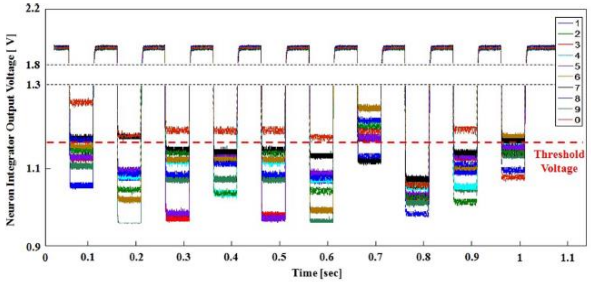


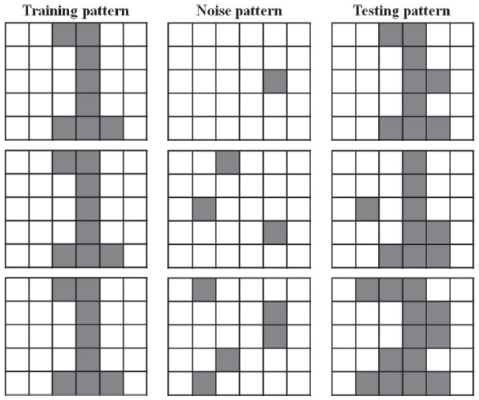












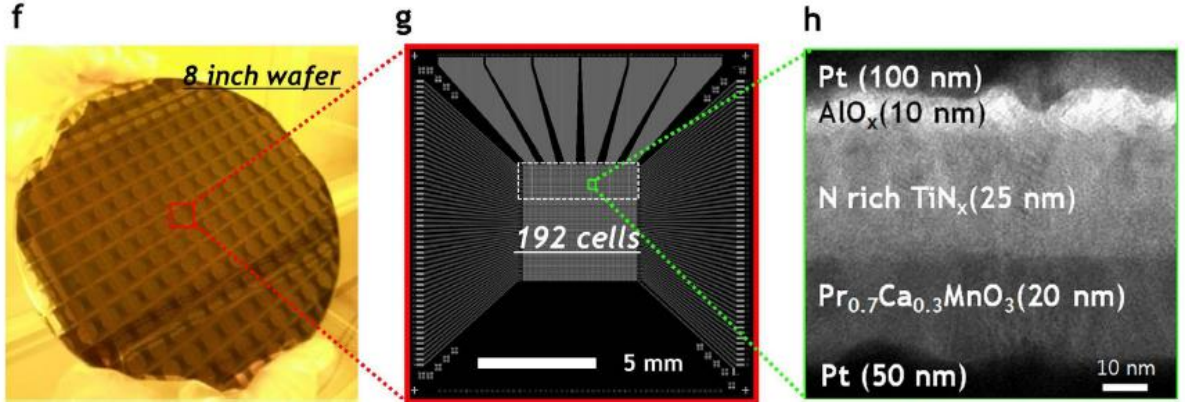
Noise level (# of noise pixel /30)	Correct recognition rate
0%	100%
3%	95%
10%	85%
16%	55%

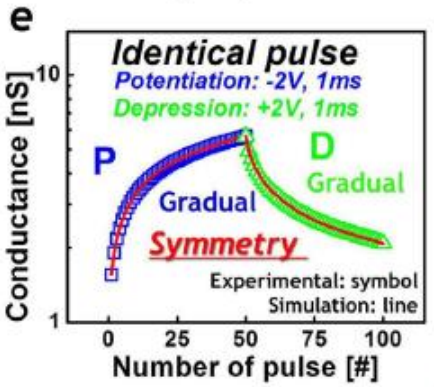
M. Chu et al. Industrial Electronics, 2014

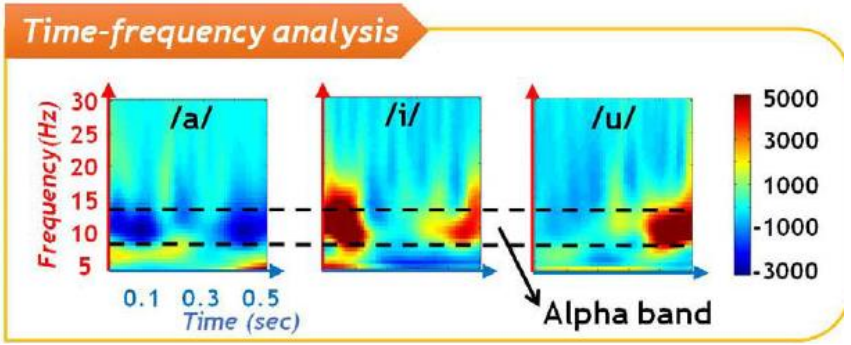
SCIENTIFIC REPORTS

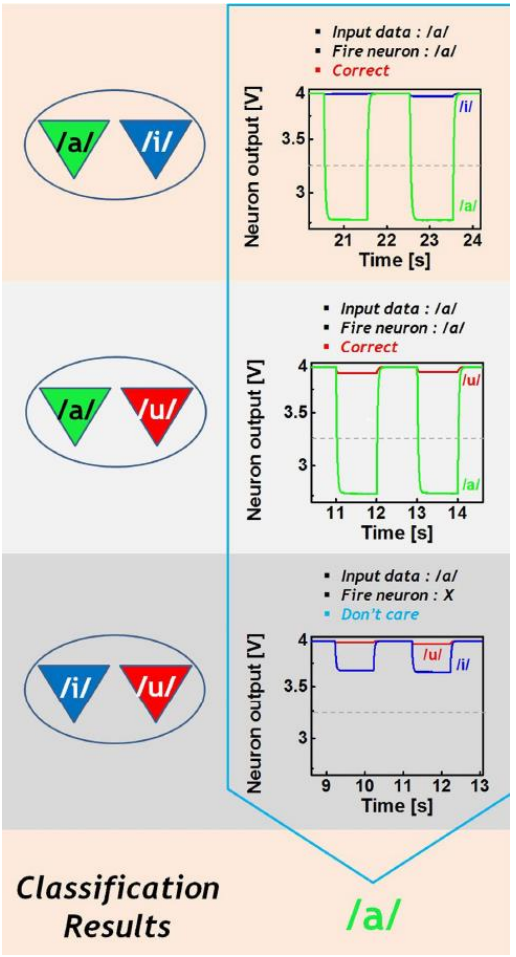
OPEN

Electronic system with memristive synapses for pattern recognition







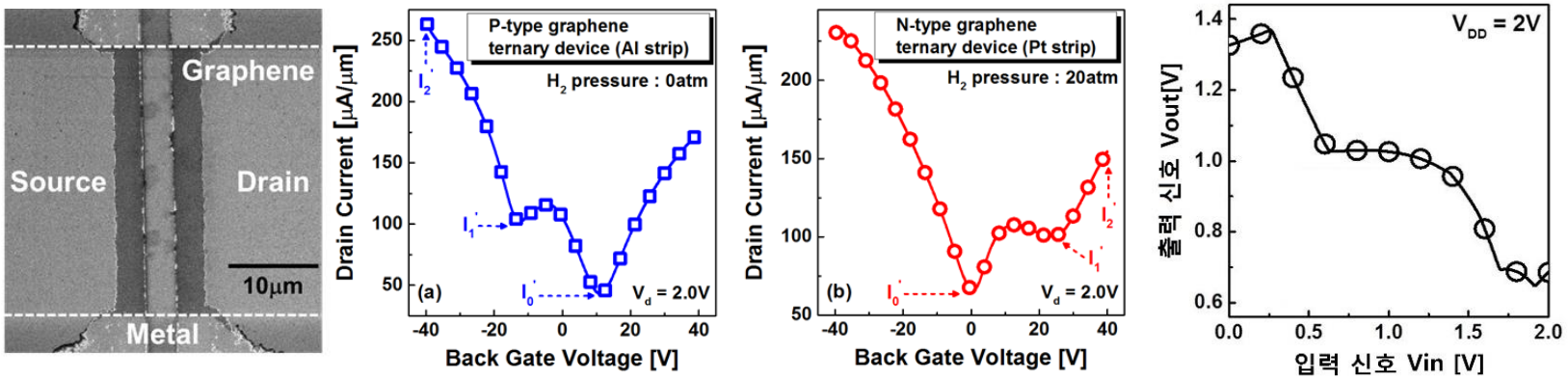


- Recognition of brain wave pattern responding to certain vowel

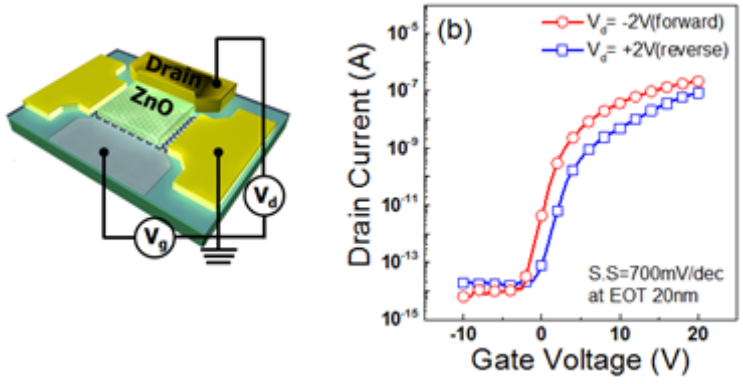
S. Park et al. SR, 2015

Ternary logic device

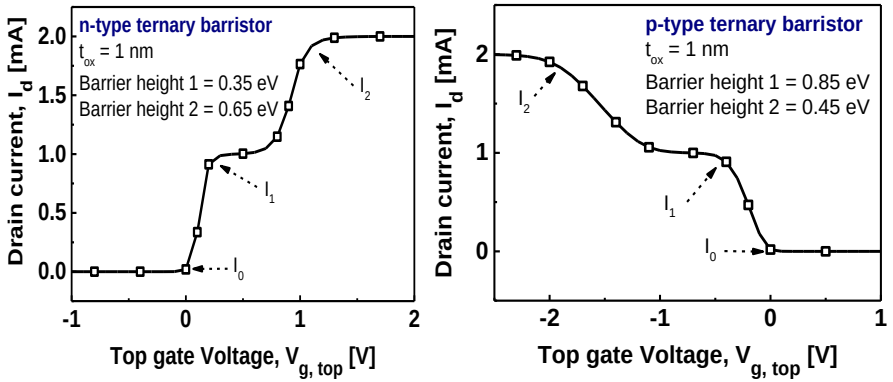
- Ternary logic device using graphene FET with PNP or NPN channel



- Graphene FET with high on-off ratio



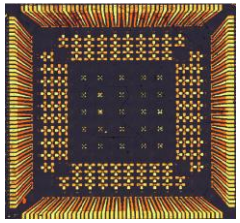
- Complimentary ternary barristor



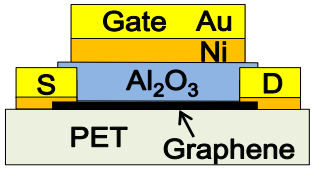
- Project to demonstrate mid-scale ALU has been launched in 2016

Y.J.Kim, in press, 2016, H.Hwang, in review, 2016

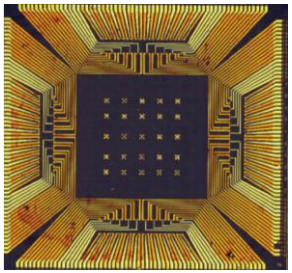
Examples of 3D stacking of 2D devices



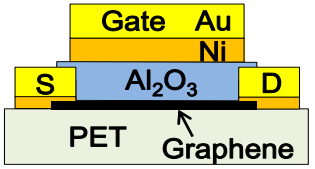
Upper level



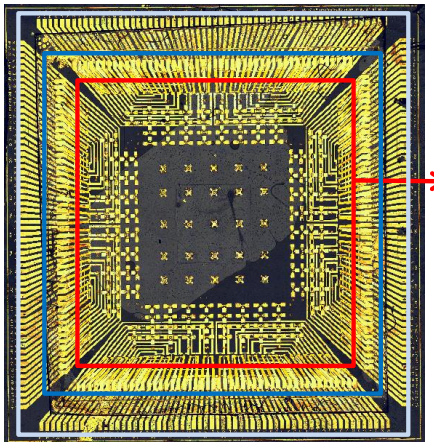
< Cross section view >

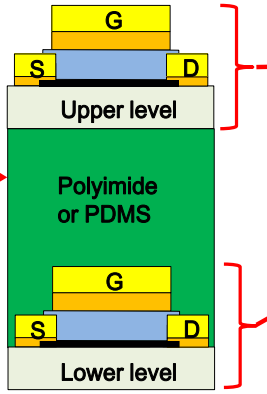


Lower level



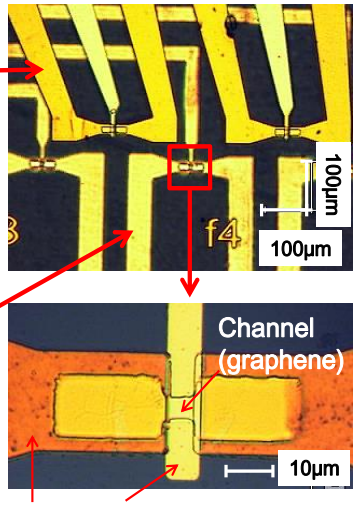
< Cross section view >



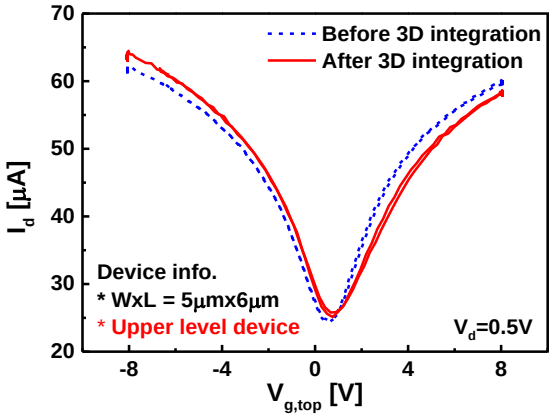


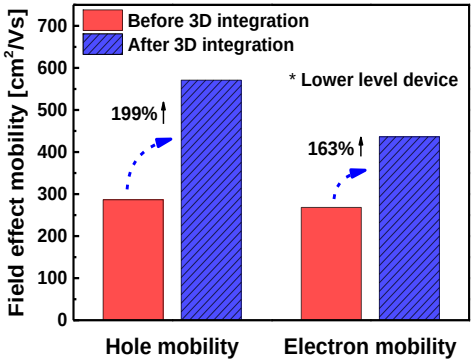
< Device area stacking >

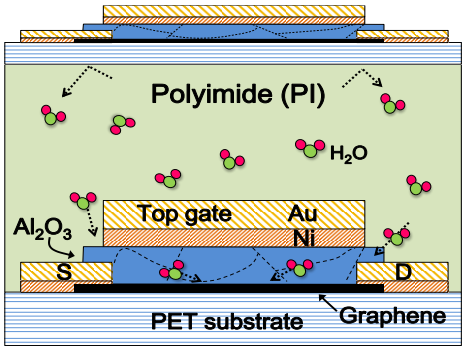
Device area — Upper level interconnect area — Lower level interconnect area

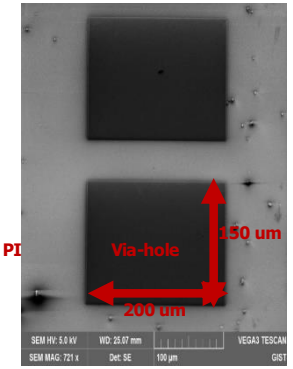


S/D, top gate (Ni/Au)









- The first demonstration of 2D layer stacking

Outline

- Motivation: End of roadmap and energy crisis
- Motivation: Inspirations from brain functions
- Bio inspired semiconductor device technology
- Summary

Summary

- New approach to adapt the novel devices into silicon fab is proposed
- Principles inspired by brain functions will be applied to minimize the power consumption of silicon chip
- To minimize the performance loss, new technology is better to be used in BEOL structure with a goal of ;
  - Simple circuit
  - Low leakage path
  - Shorter information travel
  - Replace some of FEOL block with power efficient circuits
- Key elements of this technology are being developed
  - Dynamically reconfigurable system
  - Ternary logic devices and circuits
  - Neuromophic circuit

Korean roadmap for ELP System Technology (draft)

	goals	Power savings	18	19	20	21	22	23	24	25	26	27
Power saving (product)		**1/500**							1/100		1/500	
Power saving (prototype		**1/1000**	1		1/10		1/100		1/500		1/1000	
Platform program	Extreme low V_{DD}	**1/20**	phase I			phase II			phase III			
	Monolithic 3D	**1/2**	phase I			phase II			phase III			
	Reconfigurable	**1/100**	phase I			phase II		phase III				
	Multi-value logic	**1/100**	phase I			phase II					phase III	
	Logic in memory	**1/10**	phase I			phase II				phase III		
	Optical interconnect	**1/10**	phase I					phase II		phase III		
	Neuromorphic	**1/20**	phase I				phase II			phase III		
	Process/Integration	-	phase I							phase III		
	Platform	-	phase I					phase II			phase III	
	International collaboration	-	**R&D network**					**R&D consortium**				
System program	CAD Infra	-	phase I				phase II		phase III			
	Novel architecture	-	phase I				phase II		phase III			
	Extreme low VDD	**1/10**	phase I				phase II		phase III			
	Neuromorphic	**1/100**	phase I				phase II			phase III		
	mW processor	**1/1000**	phase I					phase II			phase III	

Gwangju Institute of
Science and Technology