ElectronicsForU Network : July Ed. of India Technology Week @Home 2020

Chips for GreenTech: What will Change?

Gaurav Gupta, Ph.D.

15-July-2020

https://www.linkedin.com/in/gaurav-gupta-1205399/
https://www.quora.com/profile/Gaurav-Gupta-1977

Outline

- Why and Why Now: Energy Crisis
- Traditional Approach: Scaling Nodes
- New Approach:

  Novel

  Materials Transistors Circuits Architectures
- Enabling New Applications
- Scope for India (Discussion)
- Testing (Discussion)
- Fab/Factory vs Outsource (Discussion)
- Volume (Local Market vs Global Market) (Discussion)

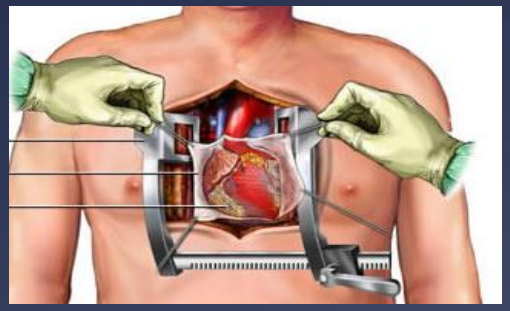
Why ?

Energy Crisis

200 Trillion-Watt-hr. consumed in 2015 by Data Servers [1] !

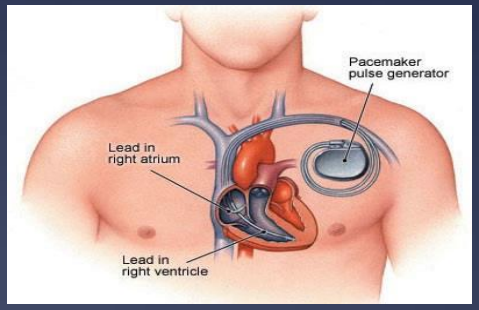
Save Energy. Save Earth

No need to repeatedly operate to Replace Implants



Save Lives & Money
(Bioelectronics)

Battery Drainage requires operation to replace implant



Explosion of devices & energy demand

Clean & Smart Environment

Objective

Traditional Computing Chips

Large Static + Dynamic Power Dissipation

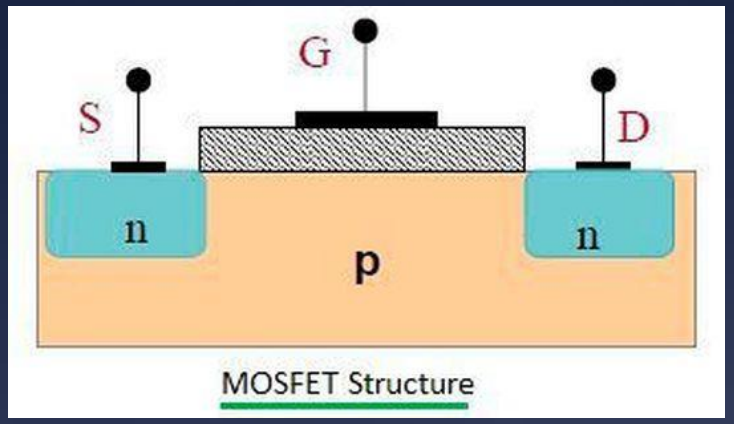
Future Computing Chips

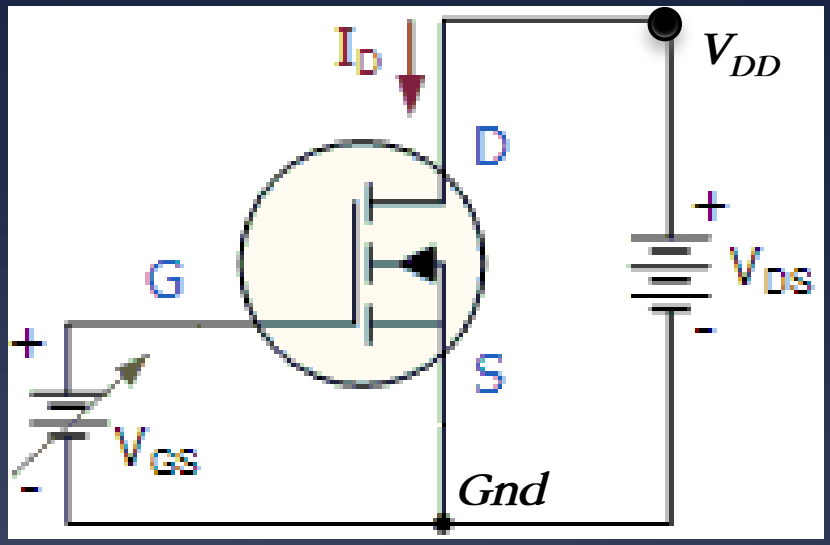
- Robust Transport
- Minimum leakage of current & energy (as heat)
- Task accomplished by consuming smaller amount of energy
- Zero Standby Power with much smaller break-even time and power

But, this is going to be "Very Very" Application Specific

Traditional Approach

Node Scaling





Scale:

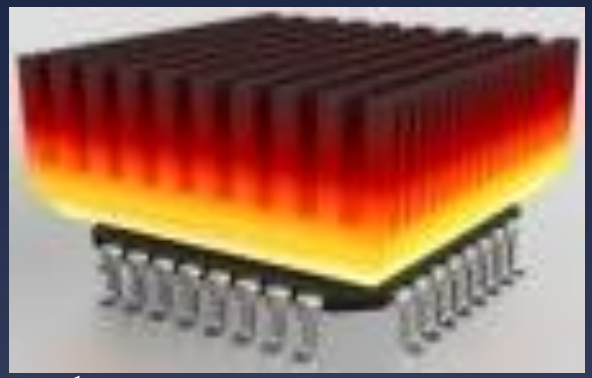
- Bias Voltage V_{DD}
- Dimensions – Channel Length, Fins
- High-K Gate Materials (HfO_2)
- Strained Transistors
- Clock Frequency
- Parallelism

Accomplish same task with lower energy/power consumption.
i.e. femto-Joules consumed for toggling 1-bit

Standard Trilemma

Silicon - scaling down of nodes automatically improved all three for many decades

Power



But now we have hit the wall !

$P \sim f.C.V^2$

Power

Quantum Tunneling

Speed

Chip

Speed

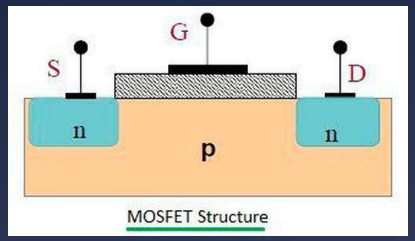
Area

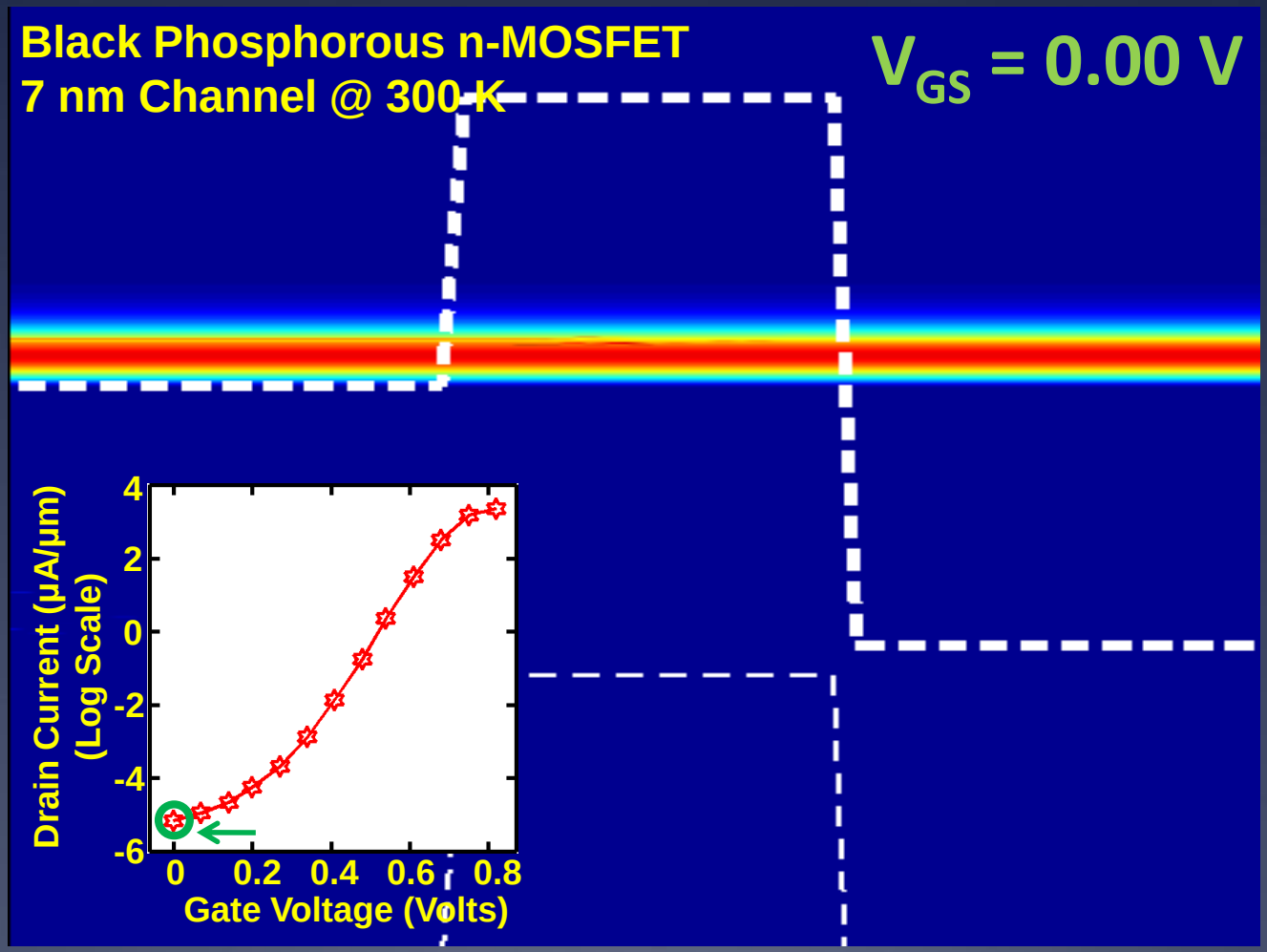
Area

Interconnect Delay

Quantum Transport Simulation of 7 nm Black Phosphorus Transistor

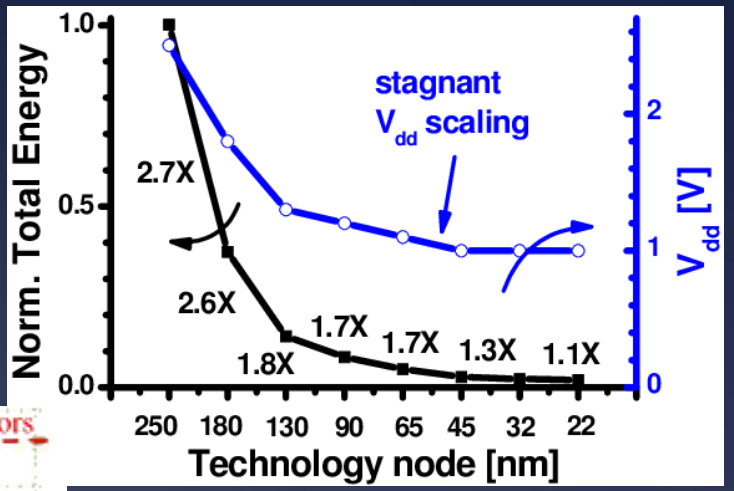
Observe the leakage through the bandgap



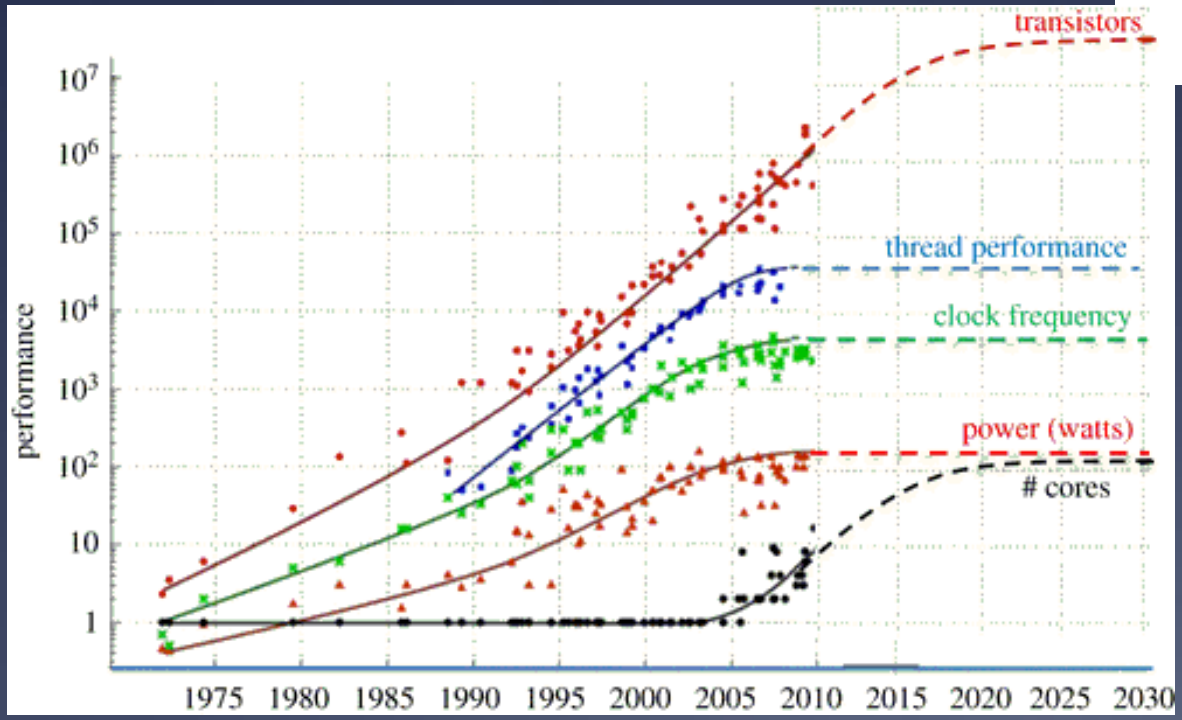


End of evolution of GreenTech Chips?

Exhausted the traditional scaling approaches.



R. G. Dreslinski et al., "Near Threshold Computing: Overcoming Performance Degradation from Aggressive Voltage Scaling", Workshop on Energy-Efficient Design, pp. 44-49, 2009



John Shalf, "The future of computing beyond Moore's Law", Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, vol. 378, issue 2166, 2020

Can we continue improving the energy and power efficiency of Chips ?

Of course Yes

Otherwise we won't be having this webinar

So let's see:

- **How**
- **Where does it lead us (if time permits)**
- **What sort of future they can unfold (if time permits)**

New Approach

Broad Classification

Microelectronics

Materials & their combinations

- **Low Power Segment**
- 2D Materials, Topological Insulators (TIs)
- **High-Power Segment**
- Wide-Bandgap Semiconductors

Transistors

- Tunneling Field Effect Transistors (TFETs)
- Van der Waals heterostructures based FETs
- Negative-Capacitance Field Effect Transistors (NC-FETs)
- Selectors

IC Design

Circuits

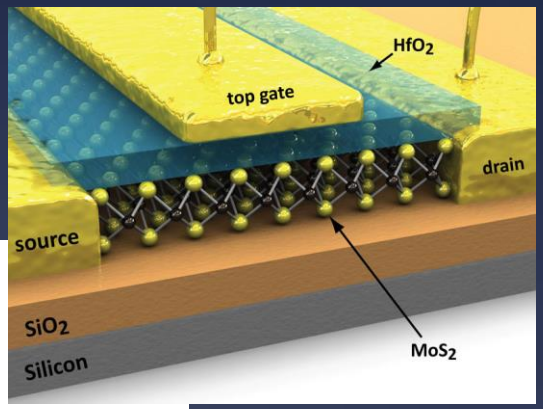
- Memory-in-Logic (new power-gating scheme)
- Adiabatic Circuits

Architectures

- Approximate Computing
- Neuromorphic Computing
- Quantum Computing

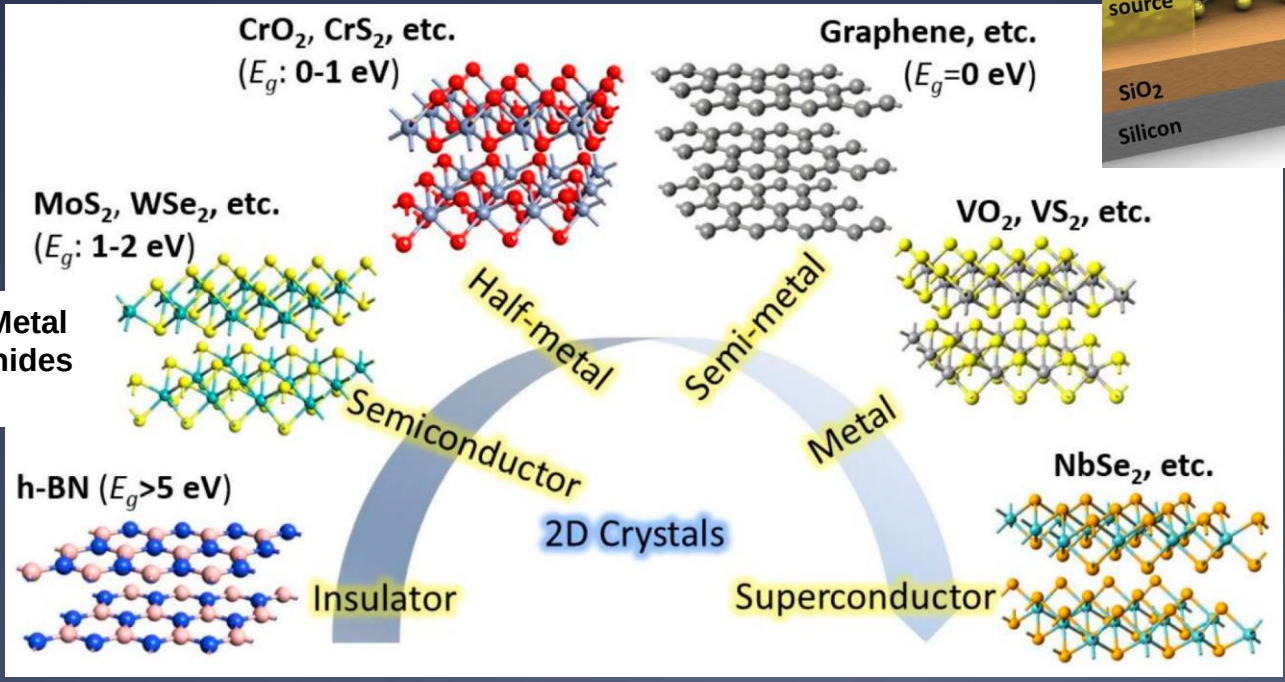
2D Materials

- Atoms in single layer
- Typically lesser scattering (absence of bonds in perpendicular direction):
  - Longer mean-free path
  - Higher mobility
- Robust electron transport



https://physicsworld.com/a/molybdenite-transistor-is-a-first/

Transition Metal Dichalcogenides (TMDs)

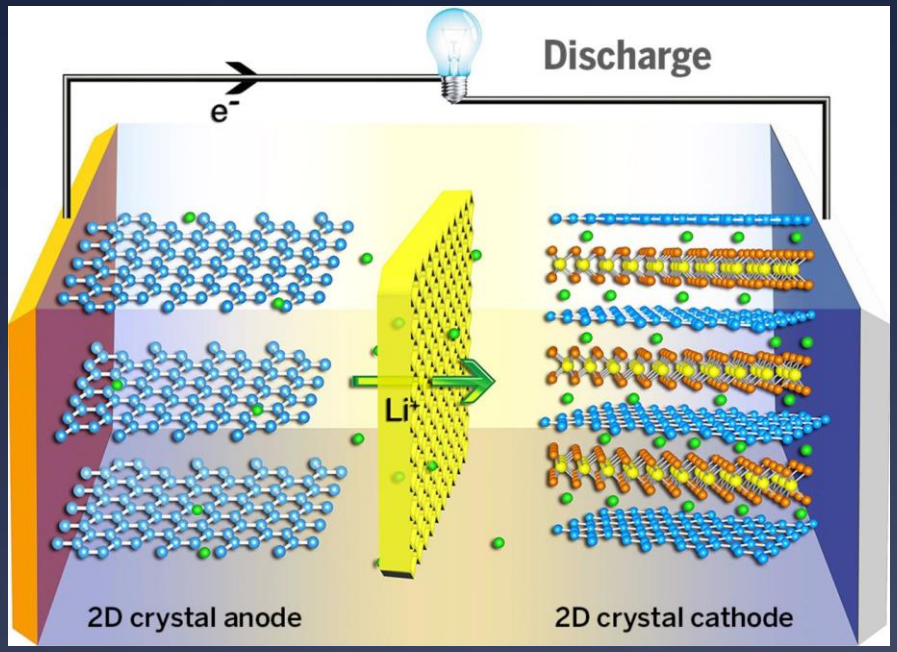


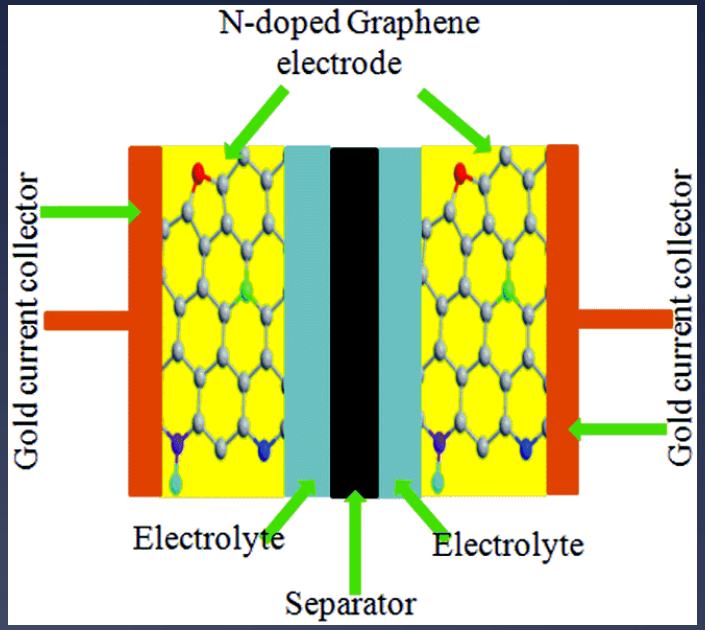
M. R. E. Tanjil et al., "Ångström-Scale, Atomically Thin 2D Materials for Corrosion Mitigation and Passivation", Coatings, 9(2), pp. 133, 2019

2D Material Applications

besides Computing

Future batteries with high energy density and fast charging

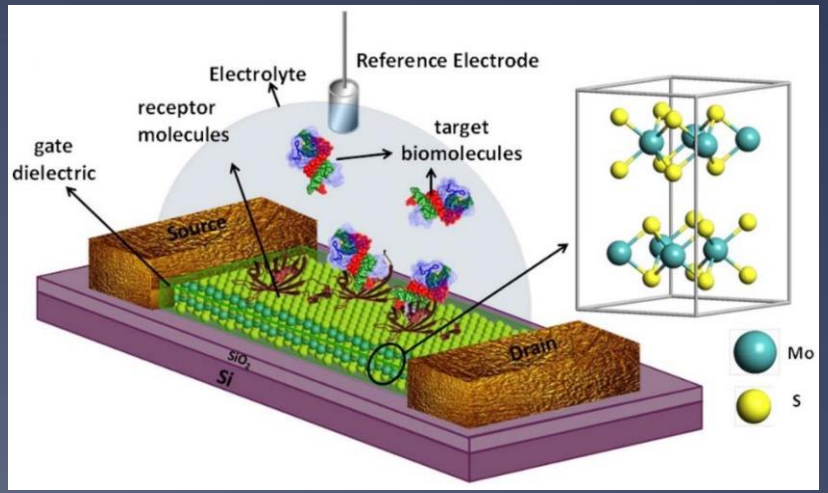




Top Left: F. Bonaccorso et al., "Graphene, related two-dimensional crystals, and hybrid systems for energy conversion and storage", Science, Vol. 347, Issue 6217, pp. 1246501, 2015 ; Top Right: E. Haque et al., "Nitrogen doped graphene via thermal treatment of composite solid precursors as a high performance supercapacitor", RSC Adv., vol. 5, pp. 30679-30686, 2015

Bio-Sensors

better SNR → lower resolution ADC → less energy consumption

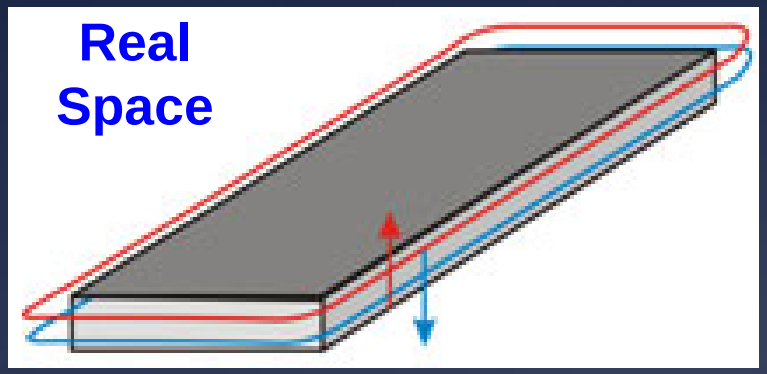


K. Shavanova et al., "Application of 2D Non-Graphene Materials and 2D Oxide Nanostructures for Biosensing Technology", Sensors 16(2), 223, 2016

Topological Insulator (TI)

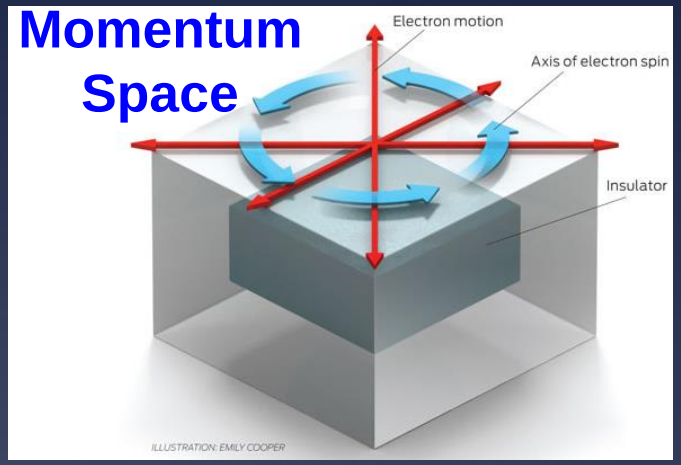
Nobel Prize 2016 Physics: topological phase transitions and topological phases of matter

https://www-ssrl.slac.stanford.edu/research/highlights_archive/topological_insulator.html



2D-TI (QSH Phase)
Spins flow on edges

http://spectrum.ieee.org/image/1876231



3D-TI – Spins flow on Surface

Electron Transport in Topological Insulators

Back-scattering Prohibited

unless:

- **Spin-Flip Mechanisms (Magnetic Impurities)**
- **Break Time-Reversal Symmetry (Magnetic-Field)**

Excellent Material for Transport and Electronic Devices

Hypothesis :

- **Perfect Transport**
- **No Heat Dissipation in the Channel**

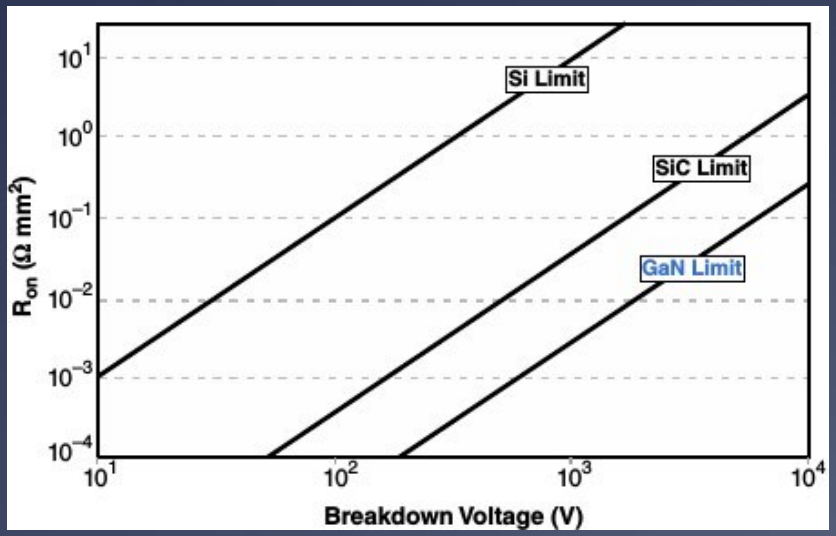
(proven for HgTe 2D-TI edge transport)

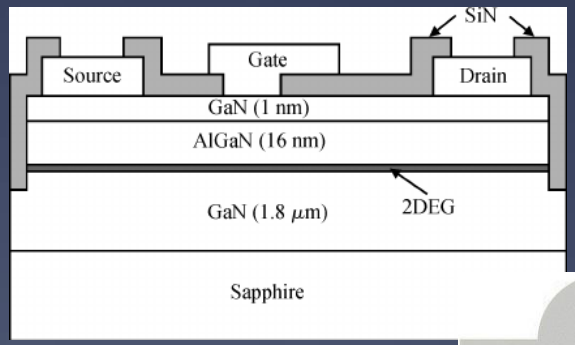
Wide-bandgap (WBG) Semiconductors

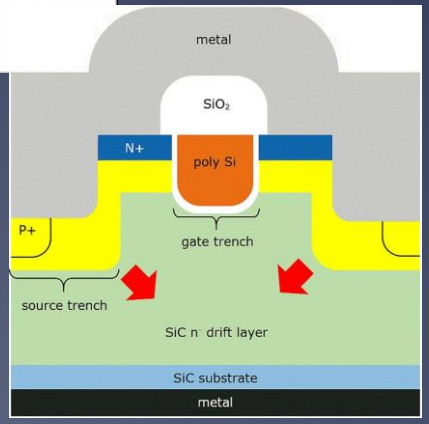
Gallium Nitride (GaN) and Silicon Carbide (SiC): Power Applications

Table 1.1 Material properties of Silicon, GaN, and SiC

Parameter		Silicon	GaN	SiC
Band Gap E_g	eV	1.12	3.39	3.26
Critical Field E_{Crit}	MV/cm	0.23	3.3	2.2
Electron Mobility μ_n	$cm^2/V \cdot s$	1400	1500	950
Permittivity ε_r		11.8	9	9.7
Thermal Conductivity λ	W/cm·K	1.5	1.3	3.8

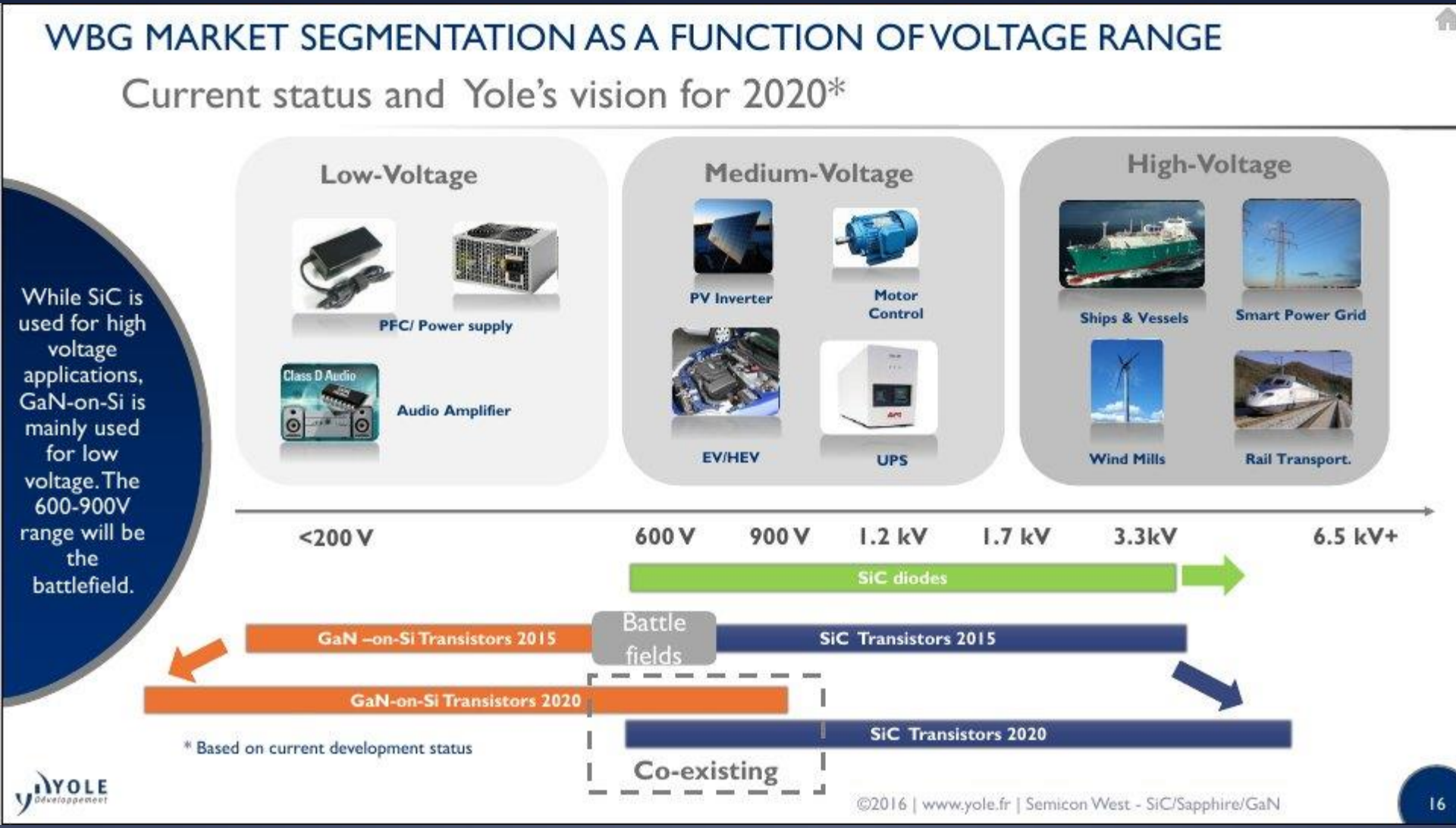






Both Table and R_{on} plot: Chap 1. GaN Transistors for Efficient Power Conversion by Alex Lidow

WBG Applications

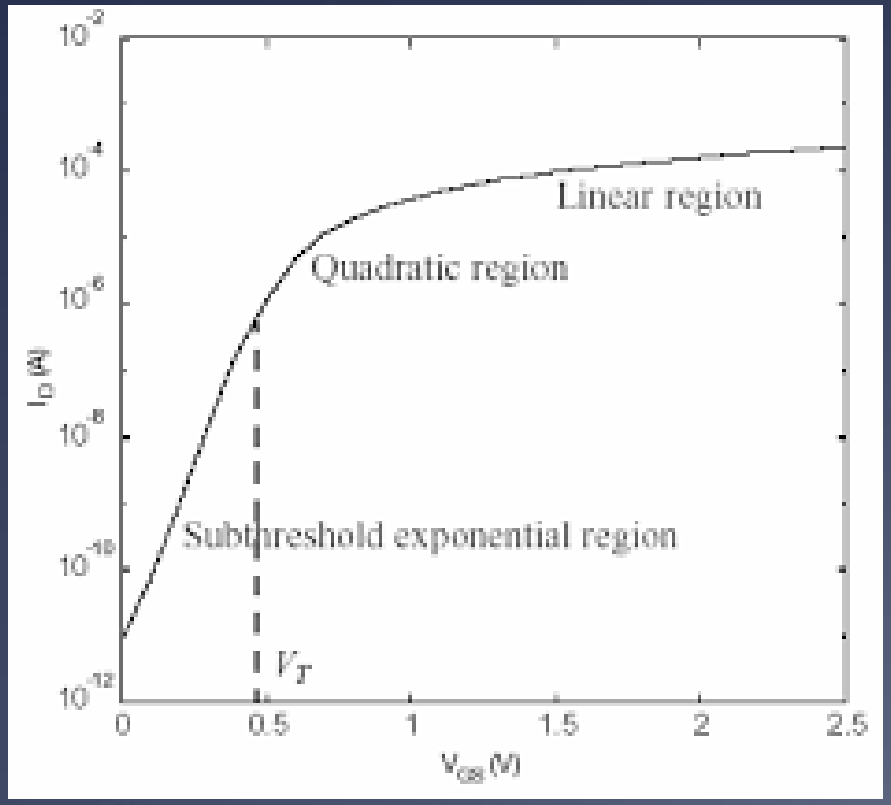


SiC, Sapphire, GaN... : what is the business evolution of the non-Silicon based semiconductor industry?, Yole Report, 2016

Sub-Threshold Slope (SS)

Within the available gate voltage swing, , low SS is required for:

- **high "switch-on current" (I_{ON}) for faster charging of load capacitors**
- **low switch-off current (I_{OFF}) for lower static dissipation**
- **MOSFETs ideal case 60 mV/decade @ 300 K (limited by thermionic emission over the barrier)**



$$SS = \ln(10)\frac{k_B T}{q}\left(1 + \frac{C_d}{C_{ox}}\right)$$

k_B is Boltzmann's constant

T is the temperature

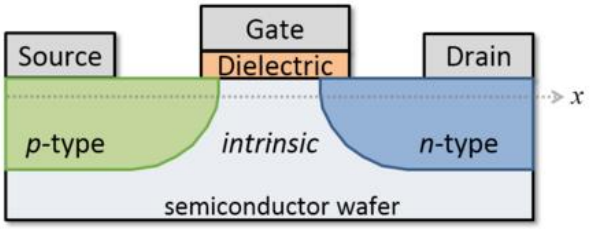
q the elementary charge

C_d the depletion layer capacitance

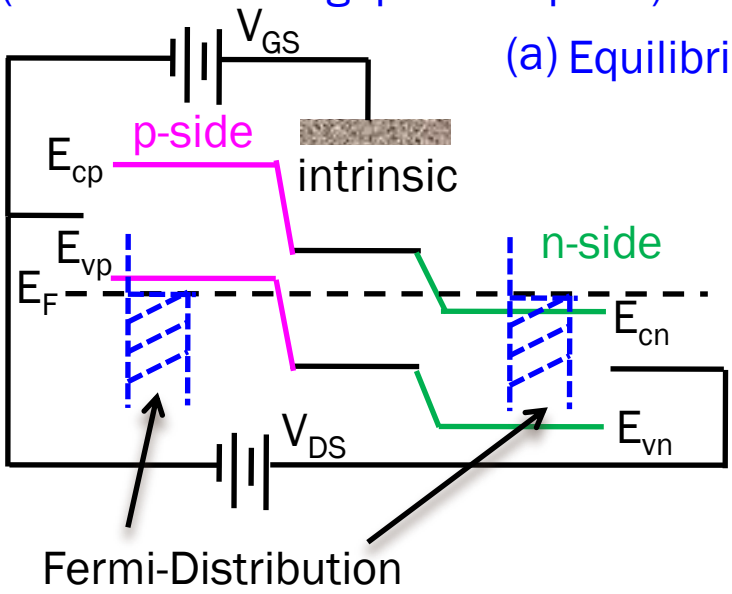
C_{ox} the gate-oxide capacitance.

https://inst.eecs.berkeley.edu/~ee40/fa03/lecture/lecture23.pdf

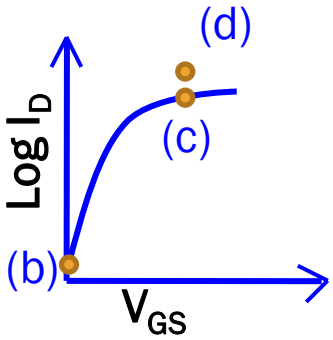
Tunneling FET (TFET)

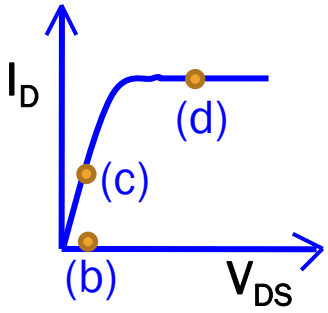


(Unbroken Bandgap Assumption)

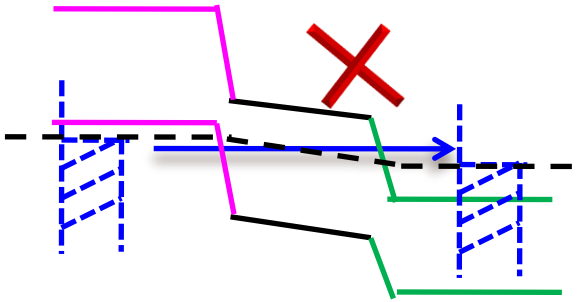


(a) Equilibrium



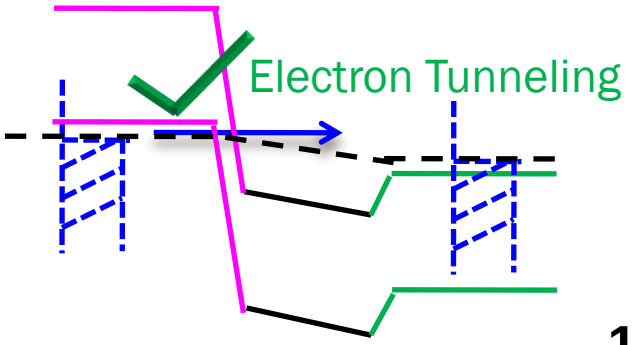


(b) Small Channel Bias



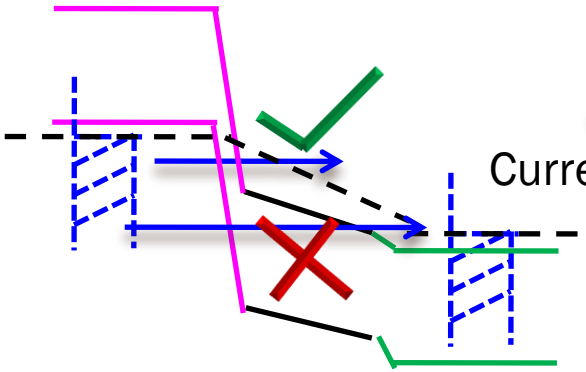
Tunneling Suppressed

(c) Apply Gate Bias

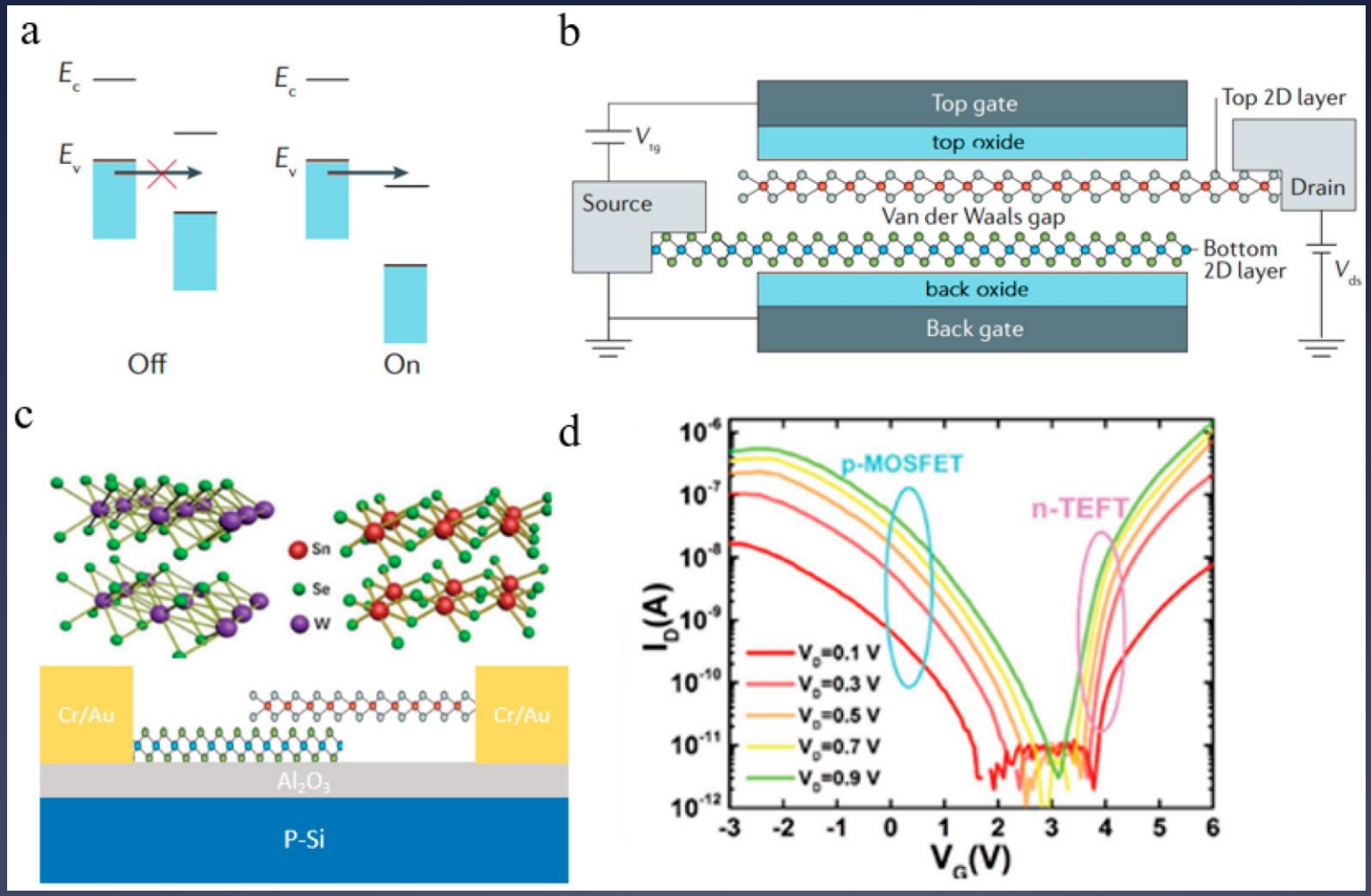


Current Saturation

(d) Large Channel Bias



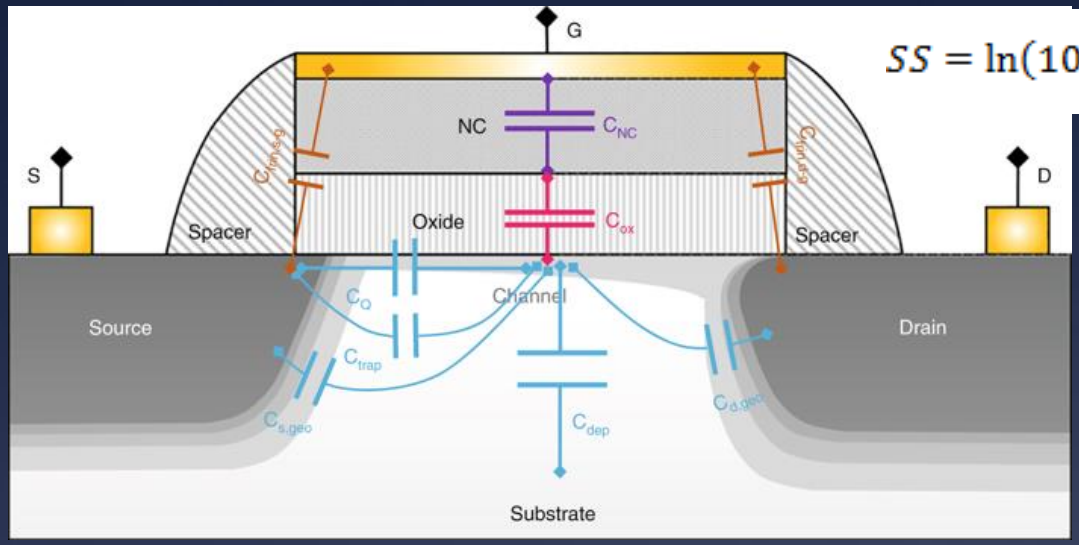
Van der Waals Heterostructure FETs



J. Li et al., "Van der Waals Heterostructure Based Field Effect Transistor Application", Crystals , 8(1), pp. 8, 2018

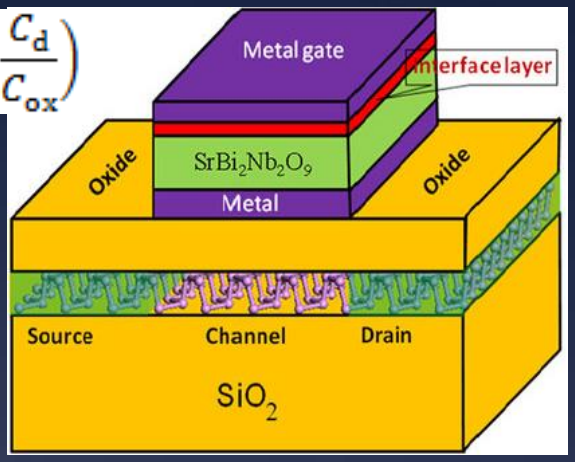
- Larger cross-section enables more current
- Robust transport in both atomic layers → lower heat dissipation

Negative Capacitance FET



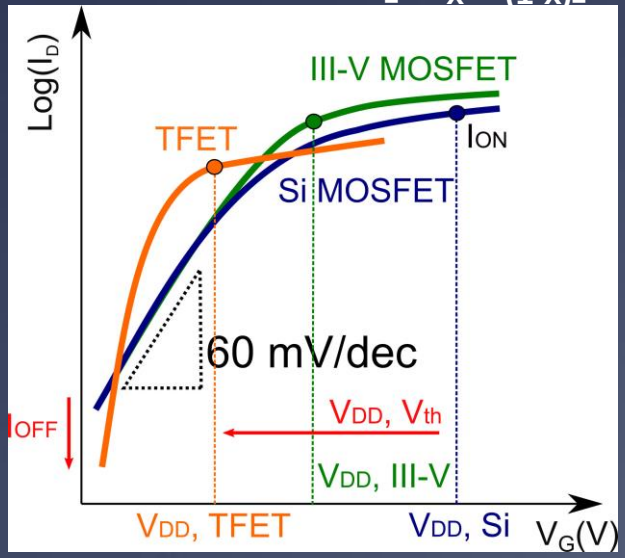
$$SS = \ln(10)\frac{k_B T}{q}\left(1 + \frac{C_d}{C_{ox}}\right)$$

https://www.nature.com/articles/s41928-020-0377-0

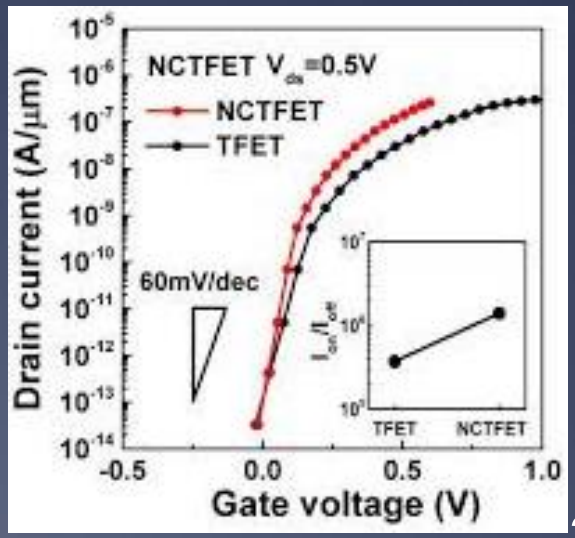


https://www.researchgate.net/publication/305663758_Negative_capacitance_transistors_with_monolayer_black_phosphorus

Negative Capacitance Gates: Ferroelectric Materials like $Pb[Zr_xTi_{(1-x)}]O_3$ (PZT), $Hf_{0.5}Zr_{0.5}O_2$ (HfZrO), $SrTiO_3$ (STO)

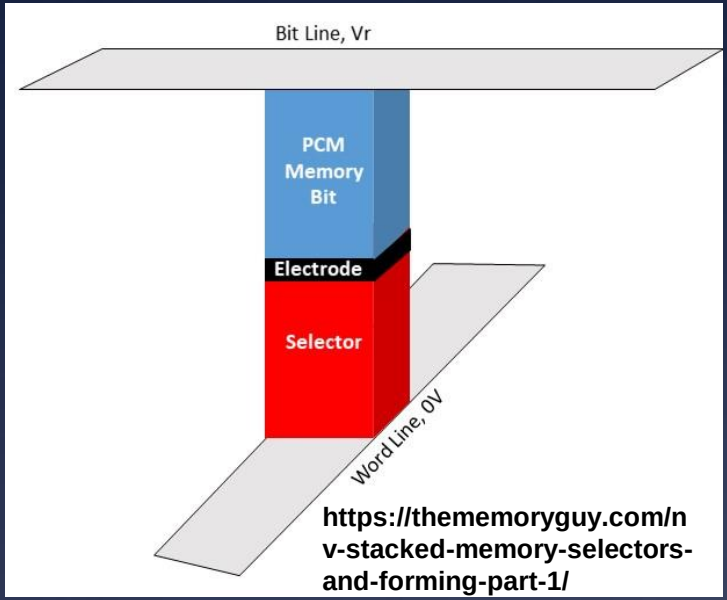


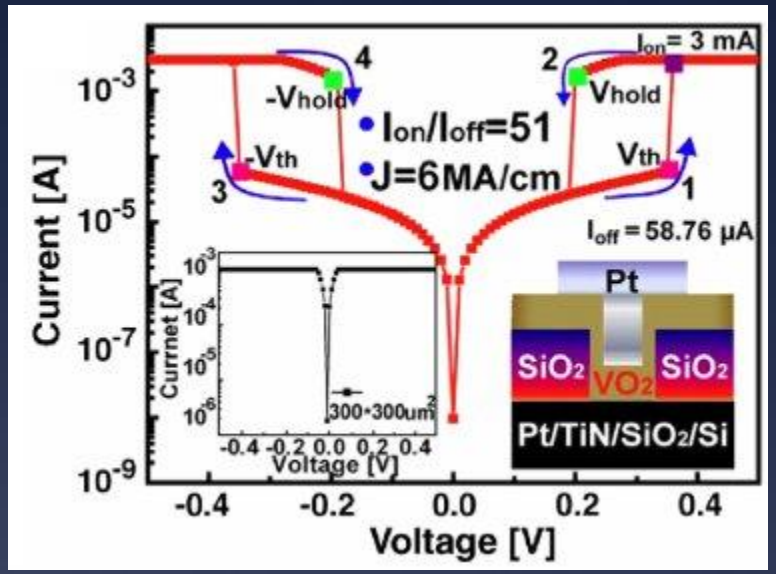
https://www.zurich.ibm.com/st/nanophotonics/tunneling.html



https://e3s-center.berkeley.edu/wp-content/uploads/2017/09/2-1-1_Kobayashi.pdf

Selectors

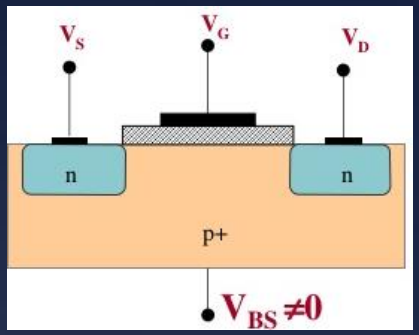




G. W. Burr et al., "Access devices for 3D crosspoint memory", Journal of Vacuum Science & Technology B 32, 040802 (2014)

- Two-Terminal Device (multi-layer stack of materials) to replace MOSFET especially for 3D-Memories
- On-Off like MOSFET
- On vs Off depends on direction of current flow through them
- MOSFET can only be grown as part of FEOL, but these can be part of BEOL

Low Power Schemes



1. Body Voltage & Multi-Threshold
2. Reduce/Multi V_{DD}
3. Reduce/Multi Freq.
4. Clock Gating
5. Store-in on-chip memory:- a. Volatile b. **Non-Volatile**
6. Retention Registers:- a. Volatile b. **Non-Volatile**
7. Complete Power-Off (Power-Gating)

$$V_T = V_{FB} + 2\varphi_b + \gamma_N \sqrt{(2\varphi_b - V_{BS})}$$

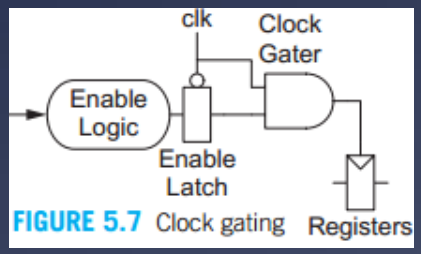


FIGURE 5.7 Clock gating

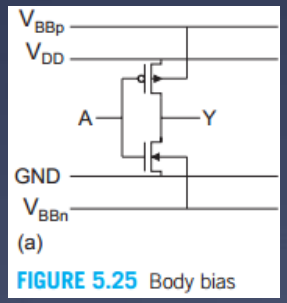
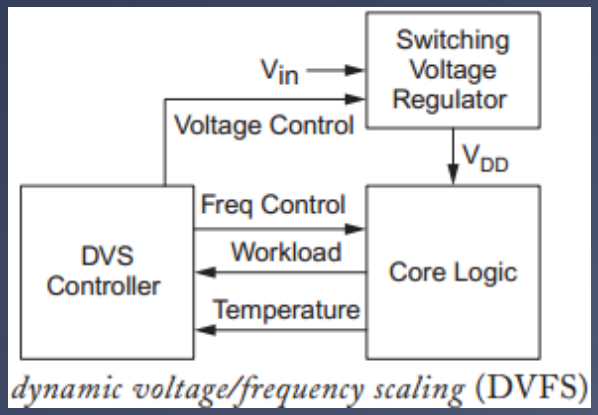


FIGURE 5.25 Body bias



dynamic voltage/frequency scaling (DVFS)

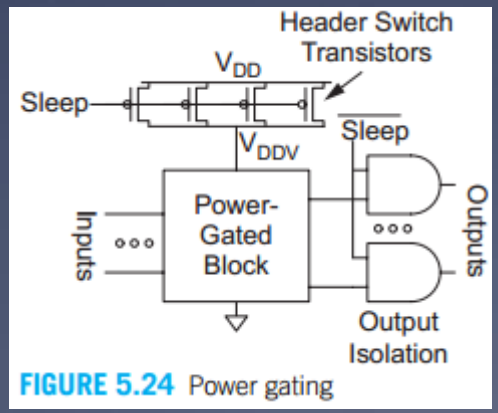
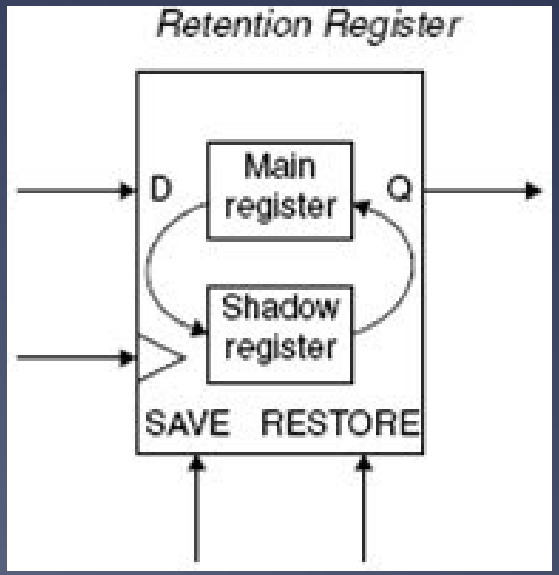
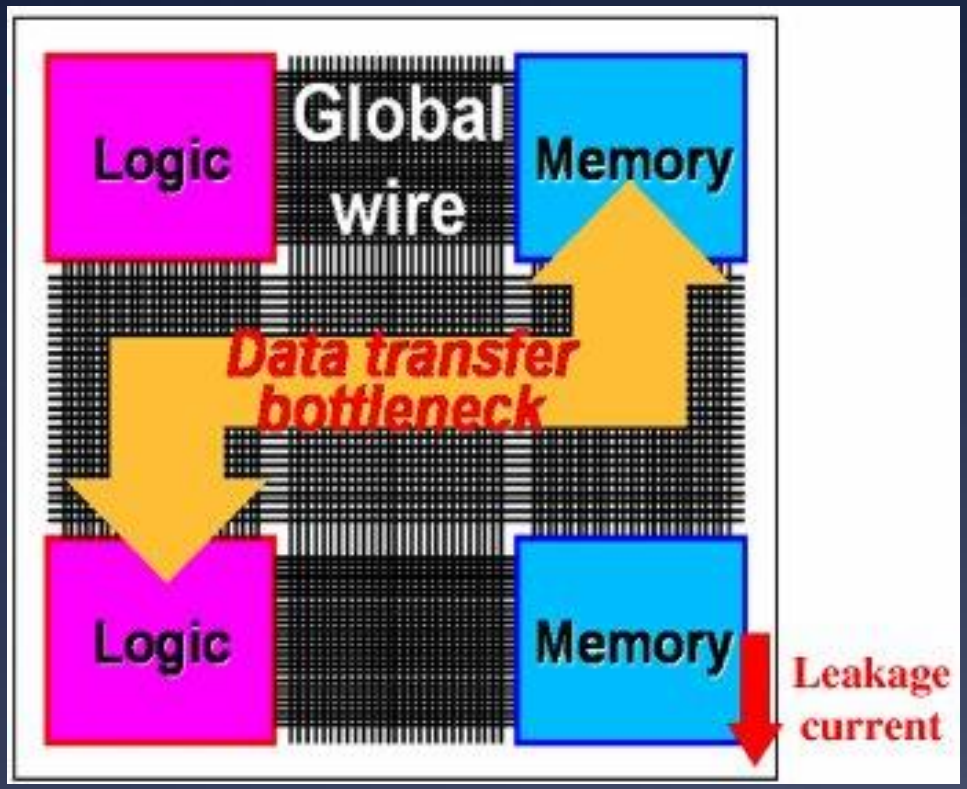
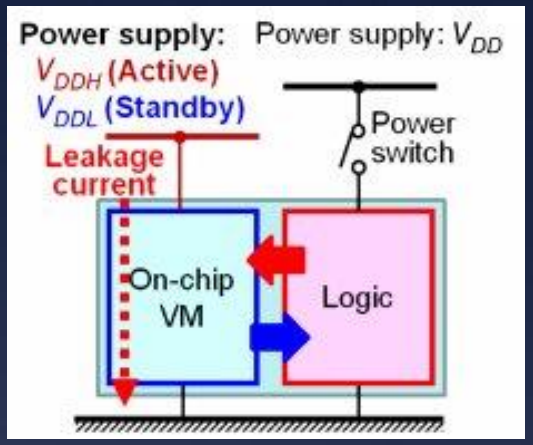


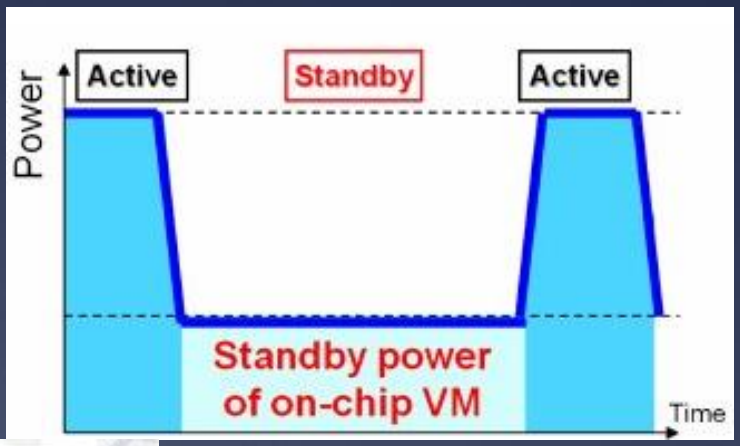
FIGURE 5.24 Power gating

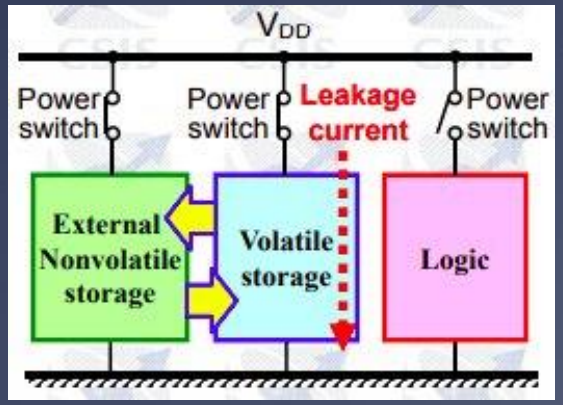


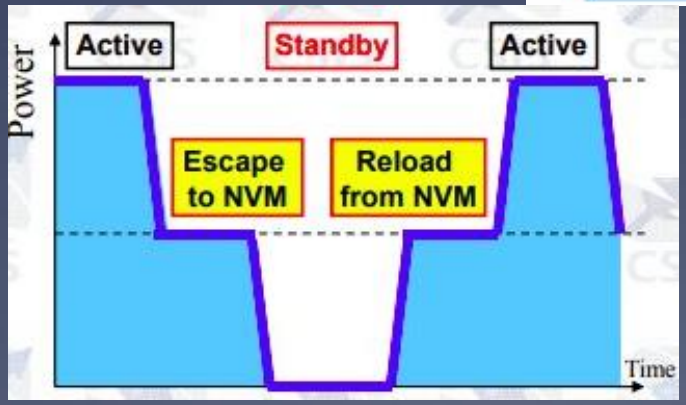
Power Gating





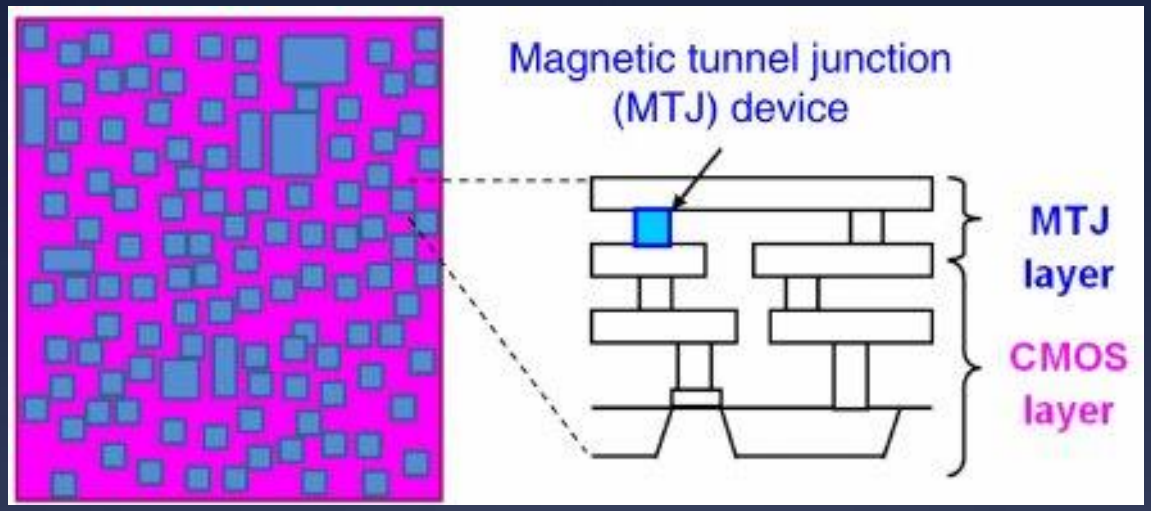


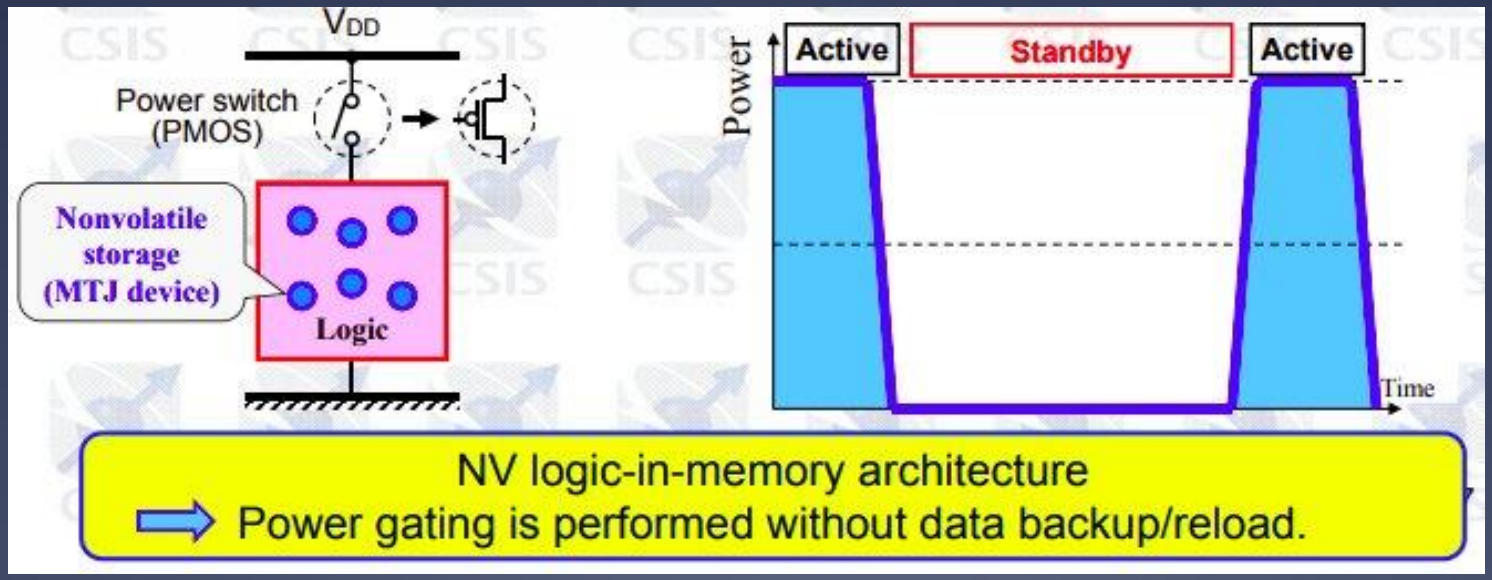




S. Matsunaga et al., "MTJ-based Nonvolatile Logic-in-Memory Architecture and Its Application", 11th Non-volatile Memory Technology Symposium, Shanghai, China, Nov. 9, 2012 ; T. Hanyu et al., "Beyond MRAM : Nonvolatile Logic-in-Memory VLSI", Chapter 7 of Introduction to Magnetic Random-Access Memory

Memory-in-Logic

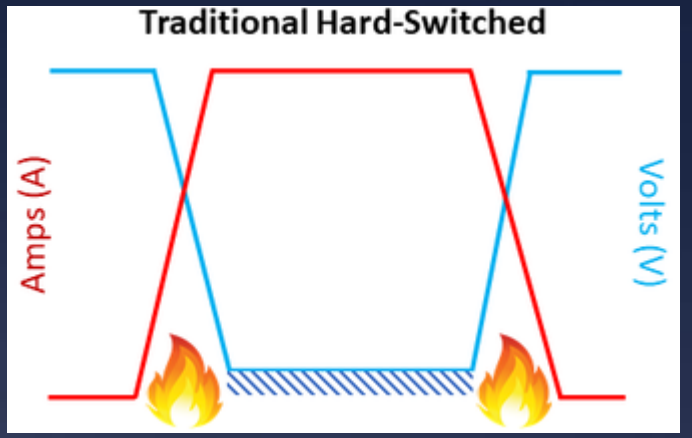


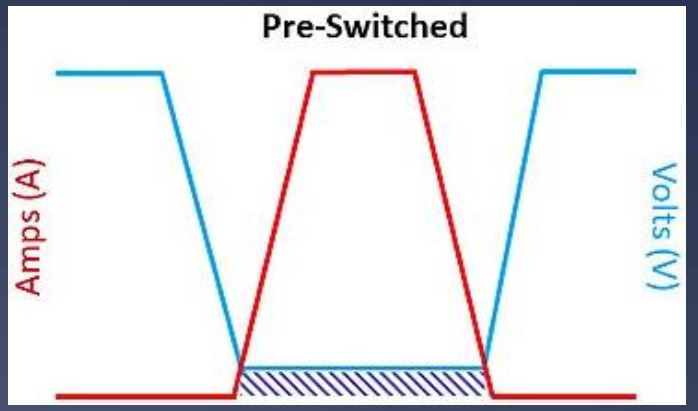


S. Matsunaga et al., "MTJ-based Nonvolatile Logic-in-Memory Architecture and Its Application", 11th Non-volatile Memory Technology Symposium, Shanghai, China, Nov. 9, 2012 ; T. Hanyu et al., "Beyond MRAM : Nonvolatile Logic-in-Memory VLSI", Chapter 7 of Introduction to Magnetic Random-Access Memory

Adiabatic Circuits

1. Never turn on a transistor when there is a voltage potential between the source and drain.
2. Never turn off a transistor when current is flowing through it.
3. recovering or recycling energy in the form of electric charge → power supplies of adiabatic logic circuits have also used circuit elements capable of storing energy.
4. Concept (1) and (2) also used in SMPS (Switched Mode Power Supplies). WBG semiconductors make this process very efficient.



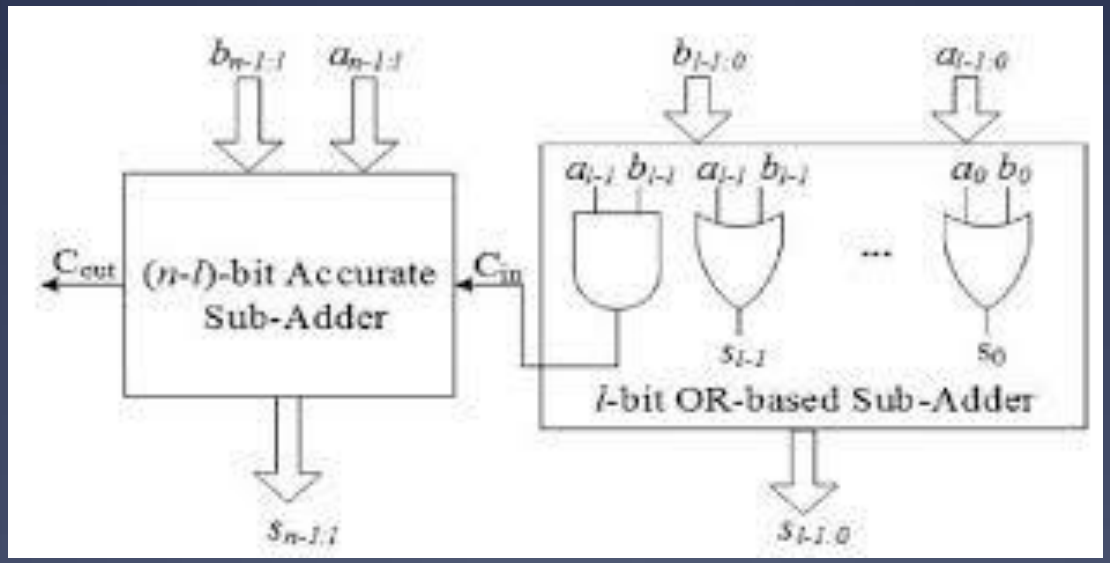


https://www.pre-switch.com/single-post/2018/05/28/Hard-Switching-Soft-Switching-Pre-Switching

Approximate Computing

1. Acceptably inaccurate result rather than a guaranteed accurate result.
2. Good for applications like Search engine, Machine Learning, Scientific Computing
3. Google using this approach in their Tensor Processing Units (TPUs)

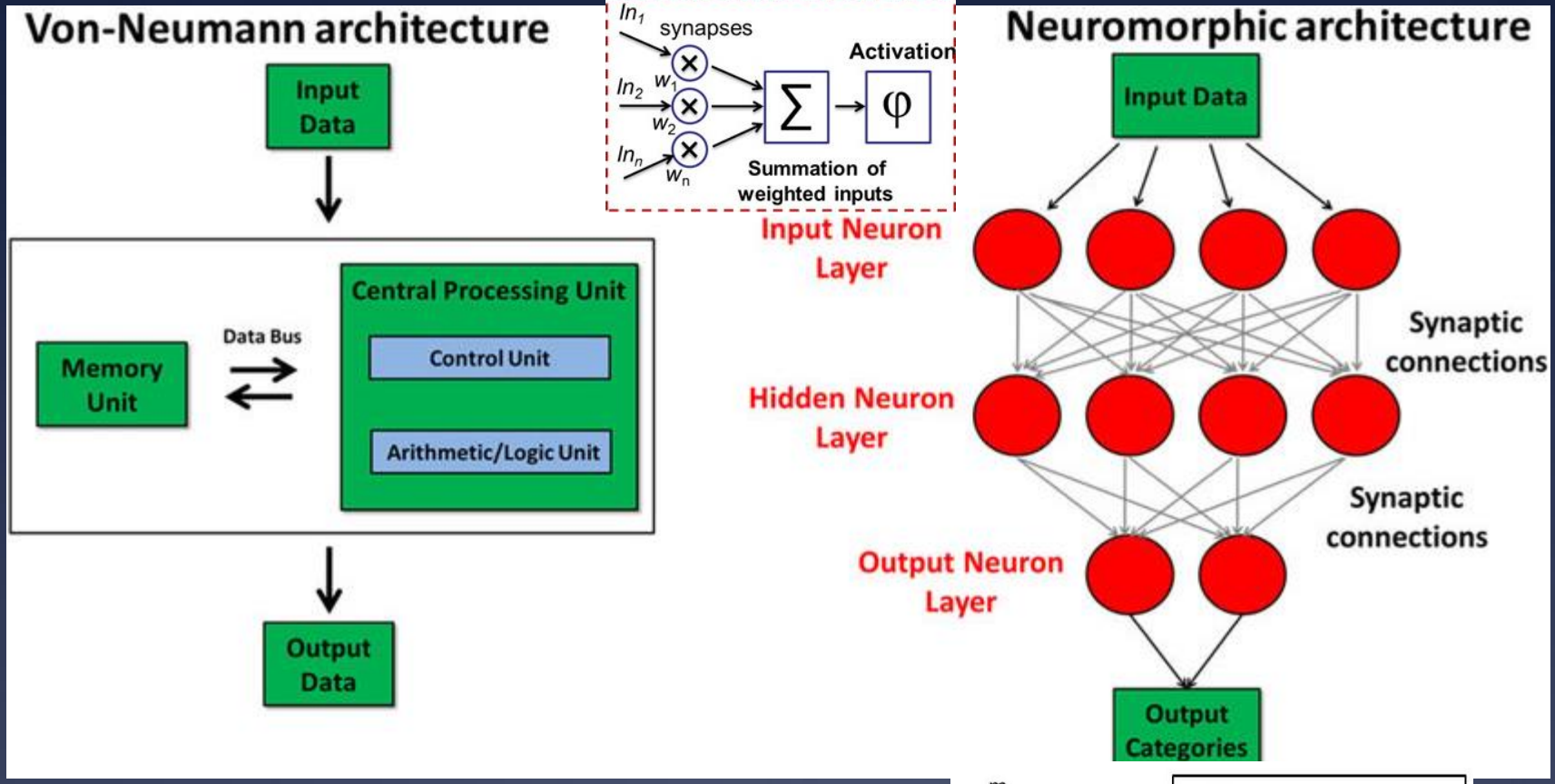
Approx. Adder Circuit

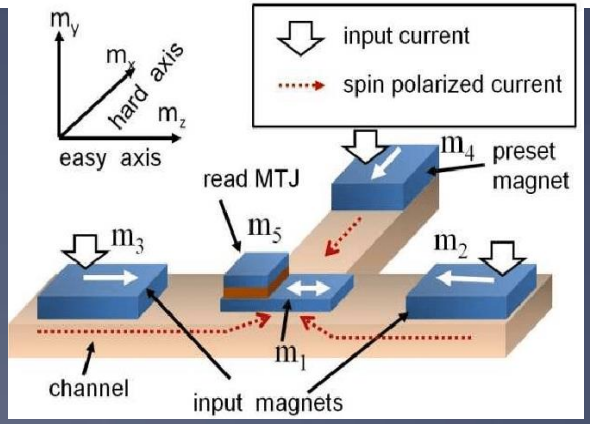


https://www.irjet.net/archives/V5/i10/IRJET-V5I10365.pdf

1. **Approximate Circuit**: reduced hardware. Save energy & area.
2. **Approximate Storage**: Truncate low bits while storing

Neuromorphic Computing

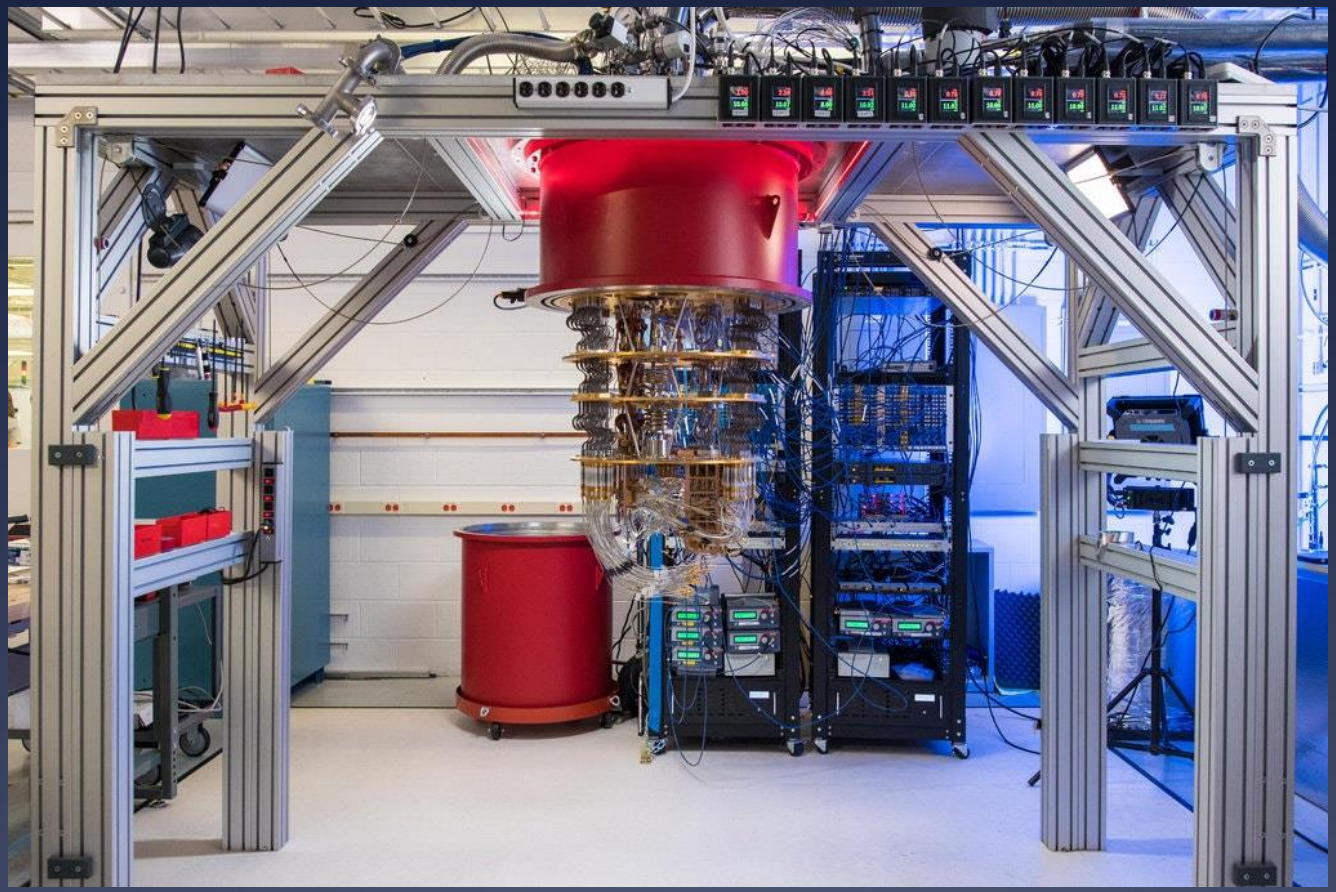




- Specifically suited for Image/Video/Text Identification type of problems.
- Non-Volatile Memory based implementation.

Quantum Computing

Google's Quantum Computer



https://www.nytimes.com/2019/10/23/technology/quantum-computing-google.html

- Very specific set of problems like Cryptography, Quantum Simulation for chemistry and drug discovery.
- NASA had bought annealing based Quantum Computer from D-Wave.

Summary

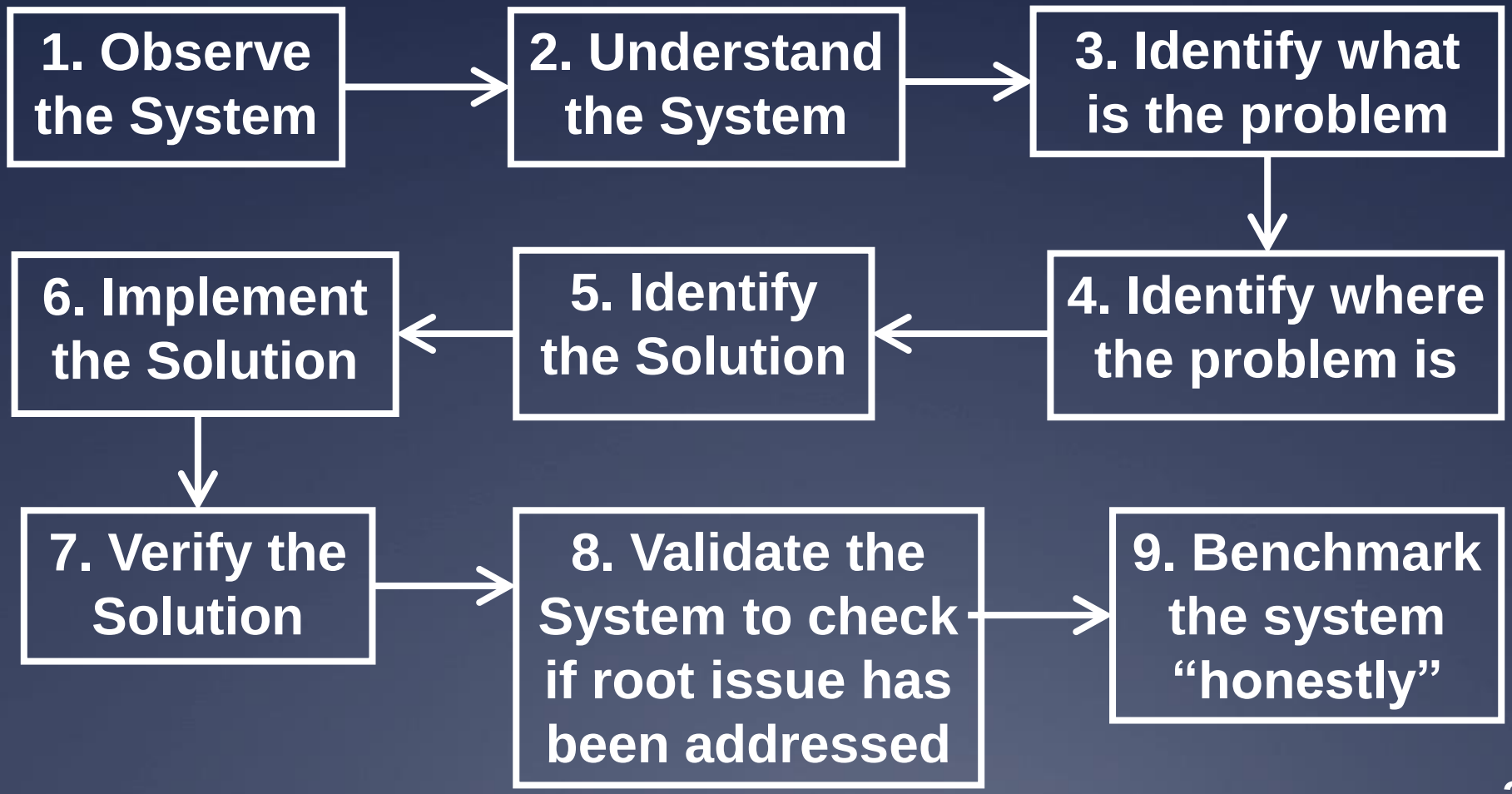
- We are at a cusp of massive change when new generation and type of GreenTech Chips shall enter our lives.
- GreenTech Chips shall see changes which range across materials, devices, circuits and architectures.
- As a designer of a chip or a system around it or an application over it, its suggested to go for first-principle thinking.
- These new solutions are very unlike previous ones and there is no concept of general or universal solution. They are not drop-down replacement of previous technology. So start by understanding the NEED, then SYSTEM, then SOLUTION, and finally do VERIFY and BENCHMARK it.

Back-Up

New Applications

Suggested Method

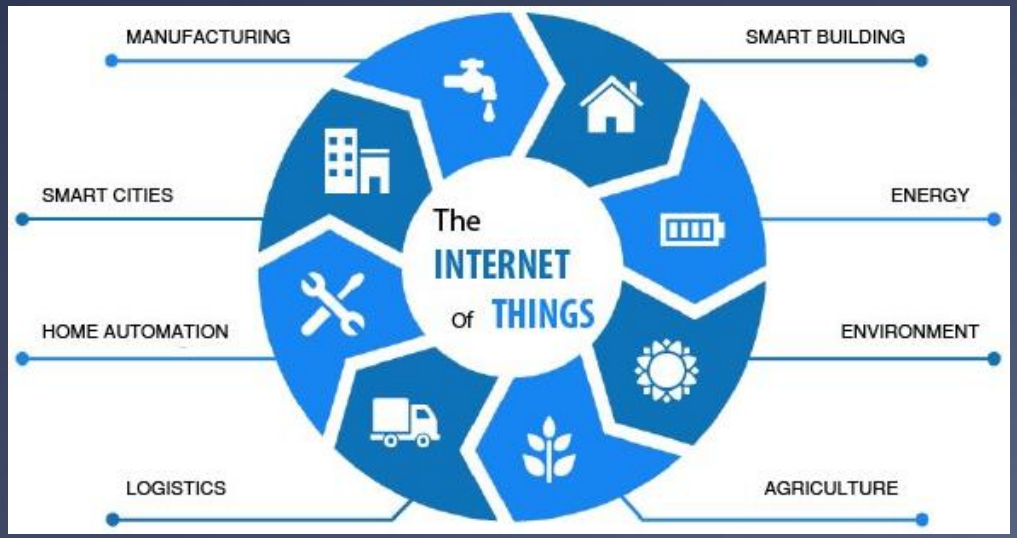
to check if one or more of “New Approach”
will actually solve a problem with a GreenTech solution,
or is it just going to be a fad or over-design.

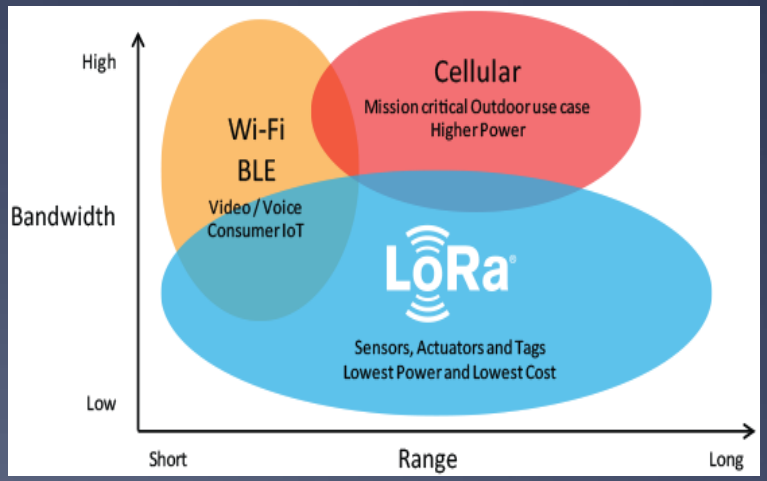


Internet-of-Things (IoT)

Low Battery Drainage, Extreme Low-Power, Energy Harvesting & Bio-Sensing Applications

- Low-Power Wireless [LoRa, LPWAN, FM, BLE], [WiFi ?]
- Analog vs Digital Computing
- Data Analytics [How Frequently ? Edge Vs Cloud ?]
- Security [Really required ? H/w vs S/w Encryption ? Cost ?]





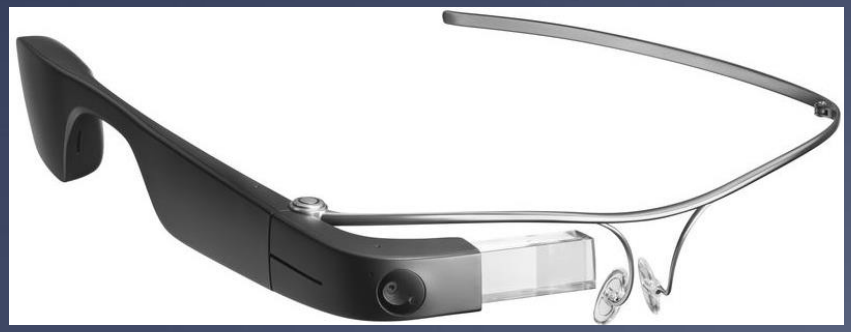
LoRa: short for “Long Range”

Wearables

- Consumer Electronics (an unmet “ Need ” ?)

- Specialized Applications (Wearable Elec. + e-Textiles)

Examples: Patients, healthcare worker, fire fighters, police/soldiers, mining workers, oil rig platform workers, construction workers, painters, Roofers, electricians, electrical power-line workers, fishers and related fishing workers, underwater welders, jewelry design



Heavy Drainage Battery Applications

Drones, Robots, Electric Vehicles, Nano-Satellites

- Understanding the "Need" (Technology, Market, Customers), e.g.:
  - "Coal-fired plants generate 72% of India's electricity" [13]
  - Norway's 95% electricity is hydropower [14]
  - Solving energy crisis and carbon foot-print

    → Case for EVs for India vs Norway ?
- This shall help decide upon the technology, e.g. for driving motor or charging battery, e.g.:
  - Power-MOSFET vs GaN HEMT
  - SiC FET vs IGBT
  - Switching Frequency & EMI (Electromagnetic Interference)
  - Thermal Constraints & Cooling requirements
  - System Size, Cost & Performance (True Bottleneck ?)

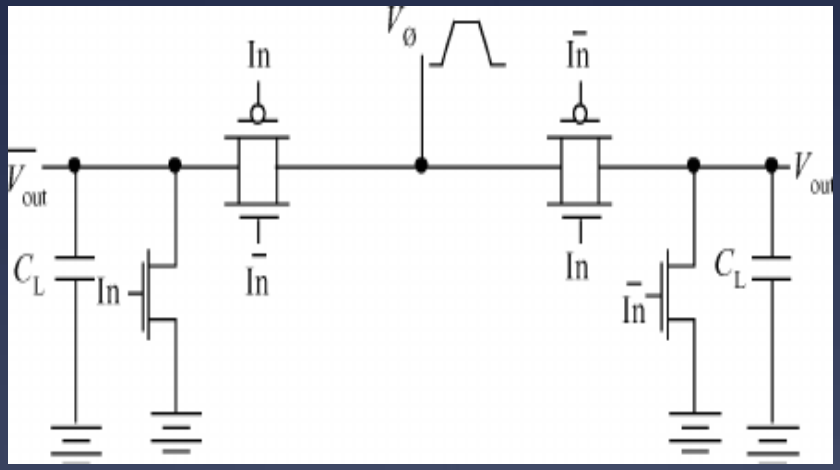
References (1/2)

[1] A. S. G. Andrae et al., “On Global Electricity Usage of Communication Technology: Trends to 2030”, Challenges 6(1), pp. 117-157, 2015.
[2] Data Centre Image: https://gcn.com/articles/2013/03/13/-/media/GIG/GCN/Redesign/Generic/datacenterheat.png
[3] Green-Earth-Leaf Image: http://www.viralsinhchauhan.in/wp-content/uploads/2019/07/Green-Planet.jpg
[4] Open Heart Surgery Image: Point 7: https://www.healthcarebusinesstech.com/the-10-most-expensive-medical-procedures/
[5] Pacemaker Image: https://www.h2hcardiaccenter.com/services/artificial-cardiac-pacemaker-implantation.php
[6] Pollution Image: https://www.theguardian.com/cities/2019/mar/05/india-home-to-22-of-worlds-30-most-polluted-cities-greenpeace-says
[7] IoT Chip Image: https://www.forbes.com/sites/louiscolumbus/2020/05/25/the-top-20-iot-startups-to-watch-in-2020/#1a0897bf7697
[8] Electric Vehicle Image: https://thebluecircle.co/2020/04/16/5-electric-vehicle-startups-making-india-clean-and-green/
[9] Chip with Heat Sink Image: https://lh3.googleusercontent.com/proxy/l8CTCdDd00lG6A0hSuFda5-ONvgy4TLwtv2nVzsvabV-kZlKORwNu5b-5VG64MA7mWxQ5NhNaMMN-kDANBFPABWMTI0CRAmzefethmOXH_YVEFfZrgc_WI0o8OZXzZGh1g
[10] Laptop Terminal loading Image: https://media.techpp.com/wp-content/uploads/2013/08/slow-computer.jpg
[11] Data Server Racks Image: https://www.engadget.com/2007-08-11-sonys-warhawk-server-farm-is-made-of-ps3s.html
[12] MOSFET Structure Image: https://www.power-and-beyond.com/mosfet-vs-bjt-whats-the-difference-a-909006/
[13] https://economictimes.indiatimes.com/industry/energy/power/india-will-not-be-able-to-achieve-its-renewable-energy-targets-anytime-soon/articleshow/69286279.cms
[14] https://en.wikipedia.org/wiki/Renewable_energy_in_Norway#Hydroelectric_power
[15] Neuromorphic Summation Image: https://dfan.engineering.asu.edu/neuromorphic-computing/

References (2/2)

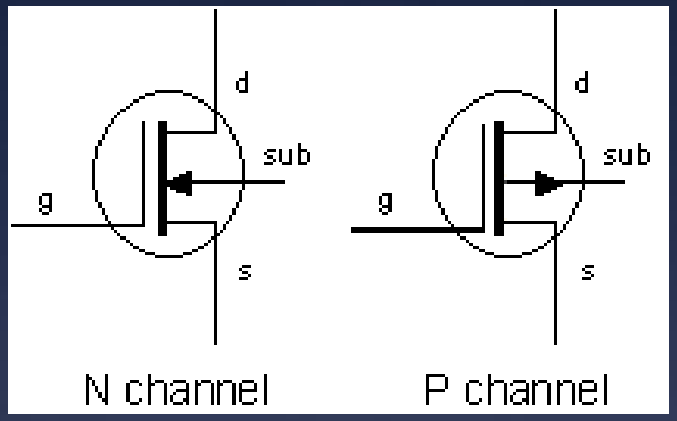
[16] Von-Neumann vs Neuromorphic Image: https://www.researchgate.net/publication/329413704_Challenges_in_materials_and_devices_for_Resistive-Switching-based_Neuromorphic_Computing
[17] Spintronic Synapse Image: https://www.researchgate.net/publication/225297114_Ultra_Low_Energy_Analog_Image_Processing_Using_Spin_Neurons
[18] IoT Image: https://softmedialab.com/blog/how-to-develop-an-iot-app/
[19] LoRa Image: https://www.semtech.com/lora
[20] Pebble Watch Image: https://www.amazon.com/Pebble-Technology-Corp-301BL-Smartwatch/dp/B00BKEQBI0
[21]Sony's Personal Aircon Image: https://www.designboom.com/technology/sony-reon-pocket-wearable-air-conditioner-07-26-2019/
[22]Google Glass Image: https://www.theverge.com/2020/2/4/21121472/google-glass-2-enterprise-edition-for-sale-directly-online
[23] E-textile: Hand with circuit Image: https://www.youtube.com/watch?v=cMcQeID6b40
[24] GaN HEMT Image: https://www.researchgate.net/publication/260295265_Breakdown_voltage_and_current_collapse_of_F-plasma_treated_AlGaNGaN_HEMTs
[25] TFET Image: https://www.wikiwand.com/en/Tunnel_field-effect_transistor
[26] Subthreshold Slope (SS) equation: https://www.zurich.ibm.com/st/nanophotonics/tunneling.html

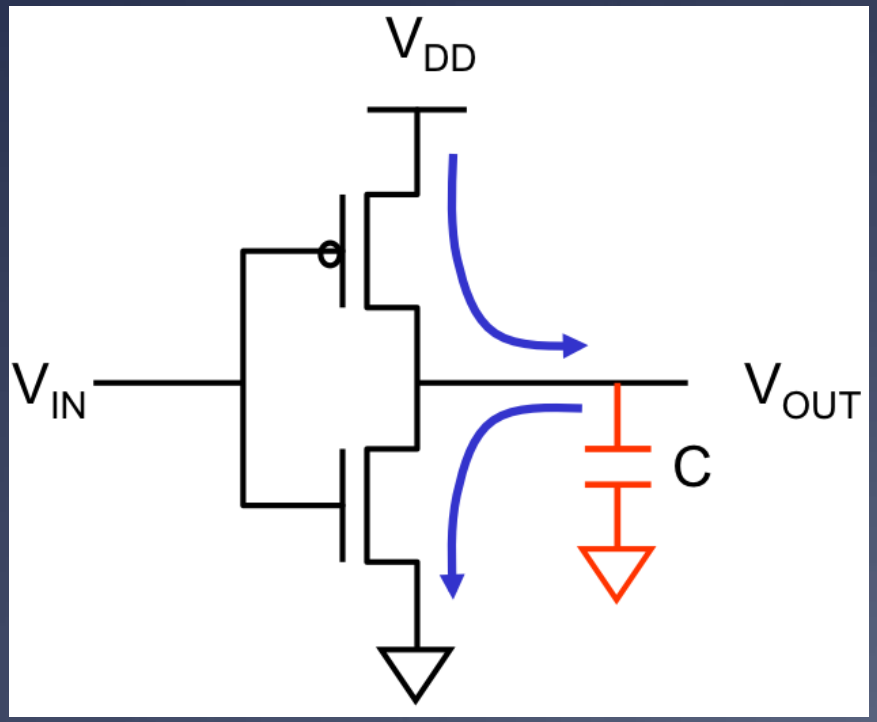
Adiabatic Adder

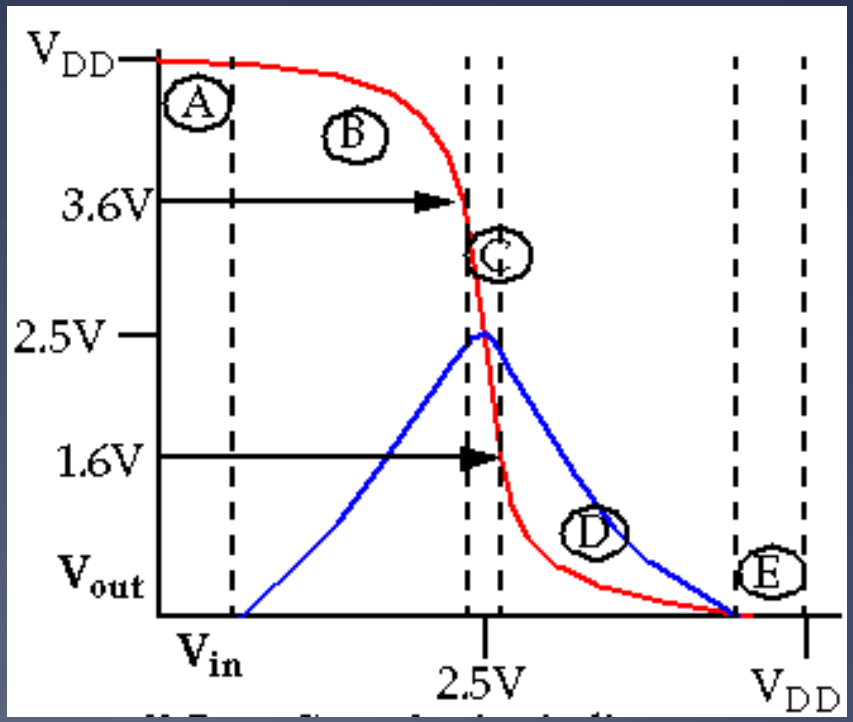


https://www.researchgate.net/publication/258308944_Scaling_trends_in_energy_recovery_logic_An_analytical_approach

Analog, Logic & Memory







How is Power Dissipated ?

$$P_{total} = P_{dynamic} + P_{static}$$

Power dissipation in CMOS circuits comes from two components:

- Dynamic dissipation due to $P_{dynamic} = P_{switching} + P_{short\ circuit}$
  - charging and discharging load capacitances as gates switch
  - "short-circuit" current while both pMOS and nMOS stacks are partially ON
- Static dissipation due to $P_{static} = \left(I_{sub} + I_{gate} + I_{junct} + I_{contention}\right)V_{DD}$
  - subthreshold leakage through OFF transistors
  - gate leakage through gate dielectric
  - junction leakage from source/drain diffusions
  - contention current in ratioed circuits (see Section 9.2.2)

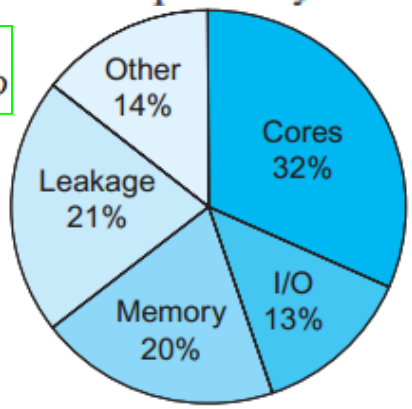


FIGURE 5.6 Power in Niagra2

Active Power → power consumed while the chip is doing useful work

Standby Power → power consumed while the chip is idle. If clocks are stopped and ratioed circuits are disabled, the **standby power is set by leakage**.

Sleep Mode → supplies to unneeded circuits are turned off to eliminate leakage → drastically reduces the **sleep power** required, but the chip requires time and energy to wake up → sleeping is only viable if the chip will idle for long enough.