Sub-10 nm transistors

and what to do with them in HEP

A.M. - CERN / May 3rd, 2019

Quotation

by Otto G. Folberth and J. Hartmut Bleher†

1. Introduction

The sweeping conquest of industrial and domestic electronics by semiconductor technology which took place over the last 20 years has been equalled or surpassed by few technical developments throughout the history of mankind. A number of different factors had to coincide

MICROELECTRONICS JOURNAL, Vol. 9 No. 4 © 1979 Mackintosh Publications Ltd., Luton.

Topics

- Past predictions
  - How good are we at “forecasting (*imagining?*) the future”?

- Revisit some major innovations that allowed scaling to go-on for 50+ years

- Sub 10 nm devices are now on the market
  - … and few atomic layers are left
  - are we approaching the end of the road?

- Possible innovations allowing growth for another generation
  - (Several) new devices from IEDM
  - (One) innovative circuit from ISSCC (If interested, please see backup slides)

- What use to make of sub-10nm transistors in HEP

Past predictions (1971)

Solid-State Electronics, 1972, Vol. 15, pp. 819–829. Pergamon Press. Printed in Great Britain

FUNDAMENTAL LIMITATIONS IN MICROELECTRONICS—I. MOS TECHNOLOGY*

B. HOENEISEN and C. A. MEAD
California Institute of Technology, Pasadena, California 91109, U.S.A.

(*Received* 11 *August* 1971; *in revised form* 8 *November* 1971)

The minimum channel length of a 2V transistor is $\approx 0.4\ \mu m$. This length is a factor of 10 smaller than the channel of the smallest present day devices. The mask alignment tolerances required to manufacture such a device are within the capabilities of electron beam pattern generation techniques. Thus we can envision fully dynamic or complementary integrated silicon chips with up to $\approx 3 \times 10^7$ MOS transistors per cm^2, operating in the 10 to 30 MHz range, as shown in Fig. 1.

The foundation paper (1974)

IEEE JSSC, VOL. SC–9, NO. 5, Oct. 1974

Design of Ion-Implanted MOSFET's with Very Small Physical Dimensions

ROBERT H. DENNARD, MEMBER, IEEE, FRITZ H. GAENSSLEN, HWA-NIEN YU, MEMBER, IEEE, V. LEO RIDEOUT, MEMBER, IEEE, ERNEST BASSOUS, AND ANDRE R. LEBLANC, MEMBER, IEEE

Classic Paper

This paper considers the design, fabrication, and characterization of very small MOSFET switching devices suitable for digital integrated circuits using dimensions of the order of 1 μ. Scaling relationships are presented which show how a conventional MOSFET can be reduced in size. An improved small device structure is presented that uses ion implantation to provide shallow source and drain regions and a nonuniform substrate doping profile. One-dimensional models are used to predict the substrate doping profile and the corresponding threshold voltage versus source voltage characteristic. A two-dimensional current transport model is used to predict the relative degree of short-channel effects for different device parameter combinations. Polysilicon-gate MOSFET's with channel lengths as short as 0.5 μ were fabricated, and the device characteristics measured and compared with predicted values. The performance improvement expected from using these very small devices in highly miniaturized integrated circuits is projected.

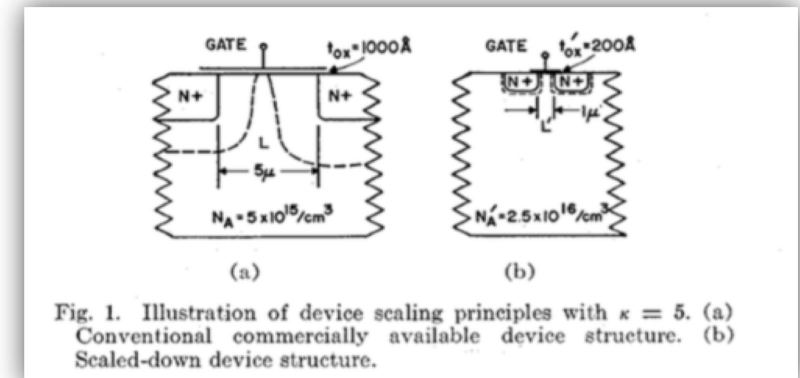


Fig. 1. Illustration of device scaling principles with $\kappa = 5$. (a) Conventional commercially available device structure. (b) Scaled-down device structure.

Past predictions (1989)

IEEE TRANSACTIONS ON ELECTRON DEVICES, VOL. 36, NO. 9, SEPTEMBER 1989

MOSFET Scaling Limits Determined by Subthreshold Conduction

JOSEPH M. PIMBLEY, MEMBER, IEEE, AND JAMES D. MEINDL, FELLOW, IEEE

mation and ultrathin (< 50 Å) gate insulators. With vanishingly small (< 50 Å) junction depth, a 30-Å gate oxide dielectric and a channel acceptor concentration of 2 × 10^{18} per cubic centimeter, one may achieve acceptably low subthreshold conduction at effective channel lengths down to 0.06 μm at an operating temperature of 300 K.

Past predictions (2001)

PROCEEDINGS OF THE IEEE, VOL. 89, NO. 3, MARCH 2001

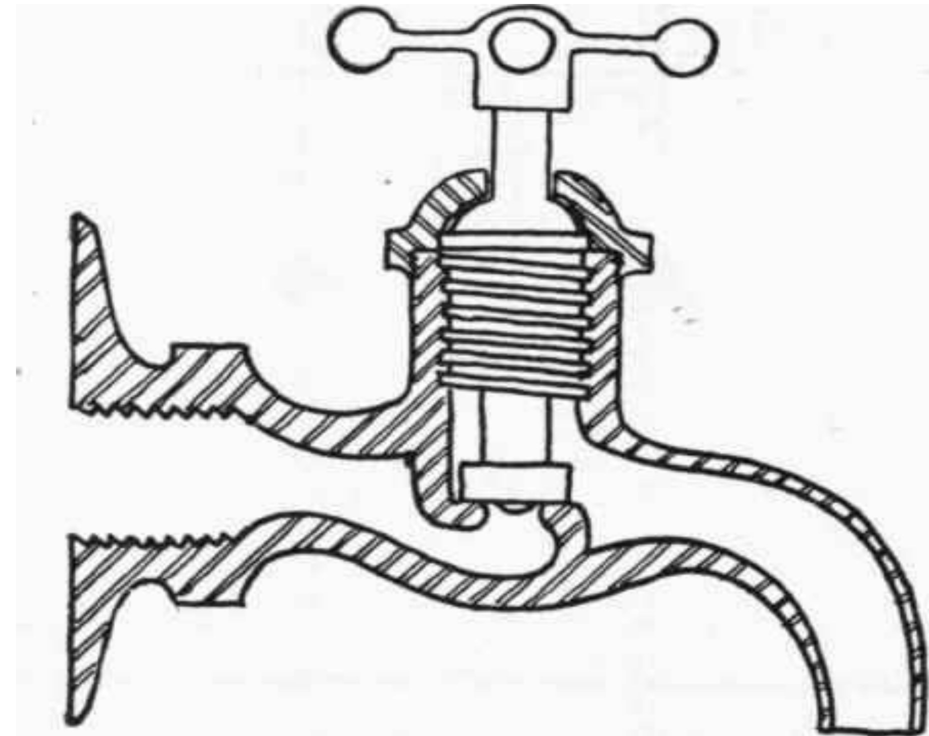
Device Scaling Limits of Si MOSFETs and Their Application Dependencies

DAVID J. FRANK, MEMBER, IEEE, ROBERT H. DENNARD, FELLOW, IEEE,
EDWARD NOWAK, MEMBER, IEEE, PAUL M. SOLOMON, FELLOW, IEEE, YUAN TAUR, FELLOW, IEEE,
AND HON-SUM PHILIP WONG, FELLOW, IEEE

The scale length theory that has been presented here provides a useful framework within which to understand the tradeoff between channel length and short channel effects. Using this theory in conjunction with the various limiting effects, we have projected that bulk-like CMOS should be extendible down to about 14-nm nominal channel length for high-performance logic and ∼35 nm for very low power applications, with intermediate applications falling in between.

What do we want from a transistor anyway? (sorry analog engineers…)

- A transistor (a digital transistor) is a device that “should” have the following characteristics:
  - works as a switch (on or off)
  - three terminals: an input, an output, a control
  - makes a “sharp” transition between the two states (open or closed) in a time as short as possible (i.e. carry charge quickly through it)
  - no leakage current when off (I_{on}/I_{off} >[>] 10^6)
  - … while delivering high current when on (drive strongly the load), $I_{on,min}$ ~ 1mA/um
  - control terminal induces a transition between the two states with a voltage drive (V_{tr}) as small as possible: $P = ½ C V_{dd}^2$ (today V_{tr} ~ 1/2 V_{dd})
  - control terminal should not be influenced by input/output terminal(s)
  - be physically small (otherwise other “parasitics” ruin the party)
  - must have complementary type (i.e. a second type which is turned on when the first is turned off using the same “control”).



- “Good analog” characteristics are desirable but by far not necessary or even important for the the majority of applications.
  In fact modern deep-submicron devices have “horrible” analog characteristics and analog designers have a hard time to achieve what was “easy” 20 years ago

Brief review of breakthroughs in the last 20 years

- Lithography
  - Computational lithography
  - Immersion lithography
  - EUV
  - …

- Strained Silicon

- High-K Metal Gate

- FinFET

Problem #1: velocity saturation

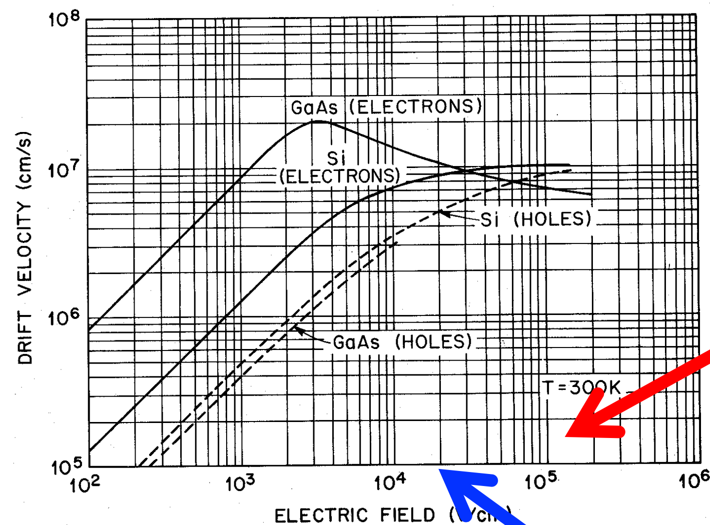


Fig. 23 Drift velocity versus electric field in GaAs and Si.[12, 13] Note at for *n*-type GaAs, there is a region of negative differential mobility.

Modern transistors operate here.

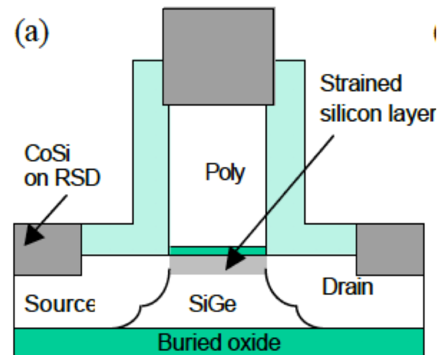
Si Detectors operate here

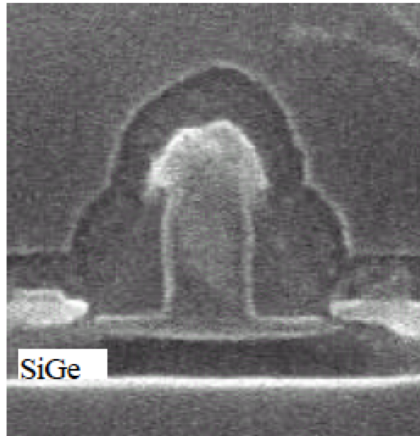
- Beyond a certain value of the electric field, electrons (and holes) are not further accelerated by an increase in the V_{DS} voltage

Faster carriers: Strained Silicon (1)

NMOS (strain)

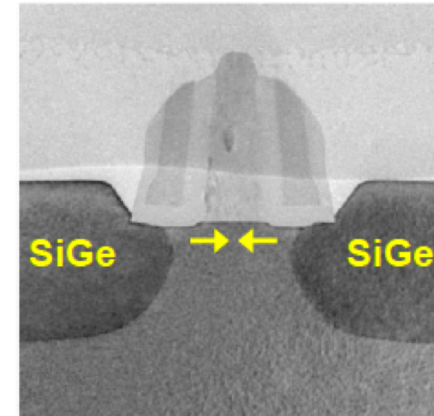
from B.H. Lee et al., IEDM 2002





PMOS (compress)

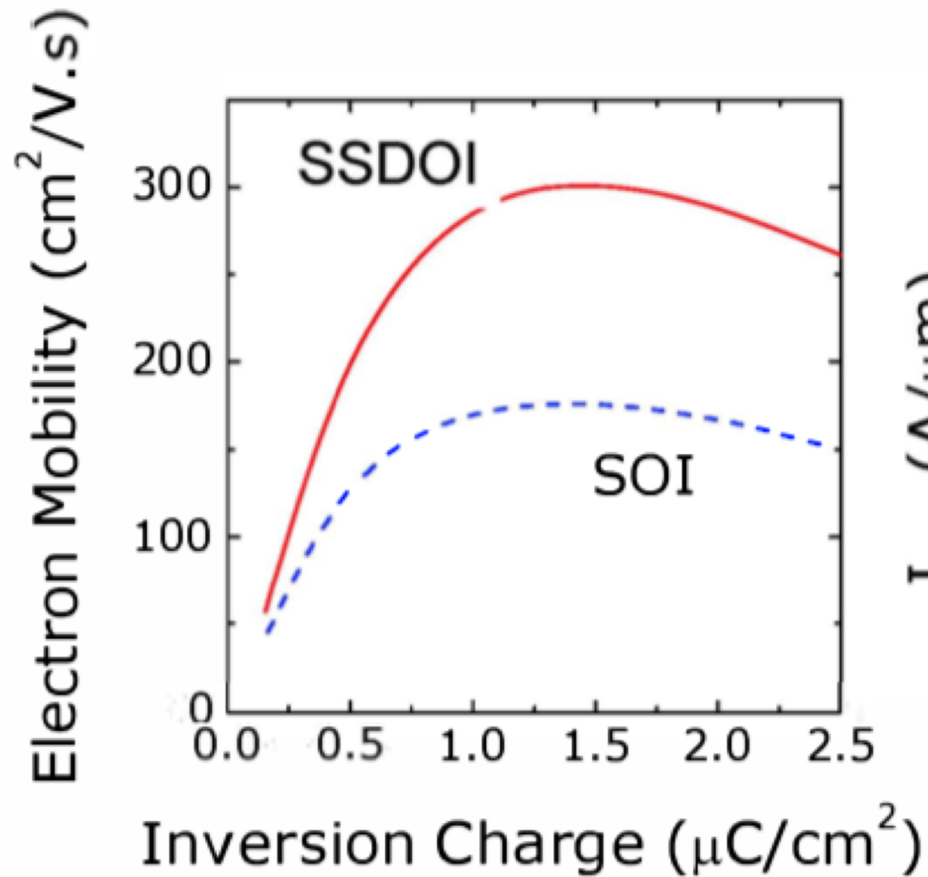
from T. Ghani et al., IEDM 2003

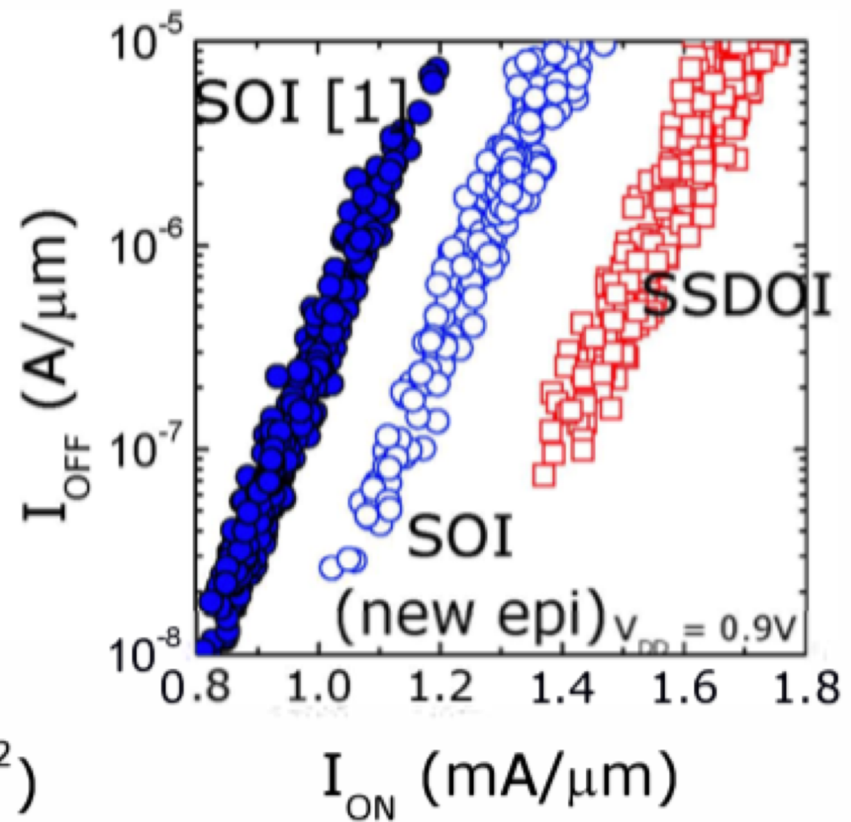


improvement. Dramatic (>50%) strain induced hole channel mobility improvement is demonstrated for our devices with 17% Ge composition. Fig. 2 shows significant improvement

Strained Silicon (2)

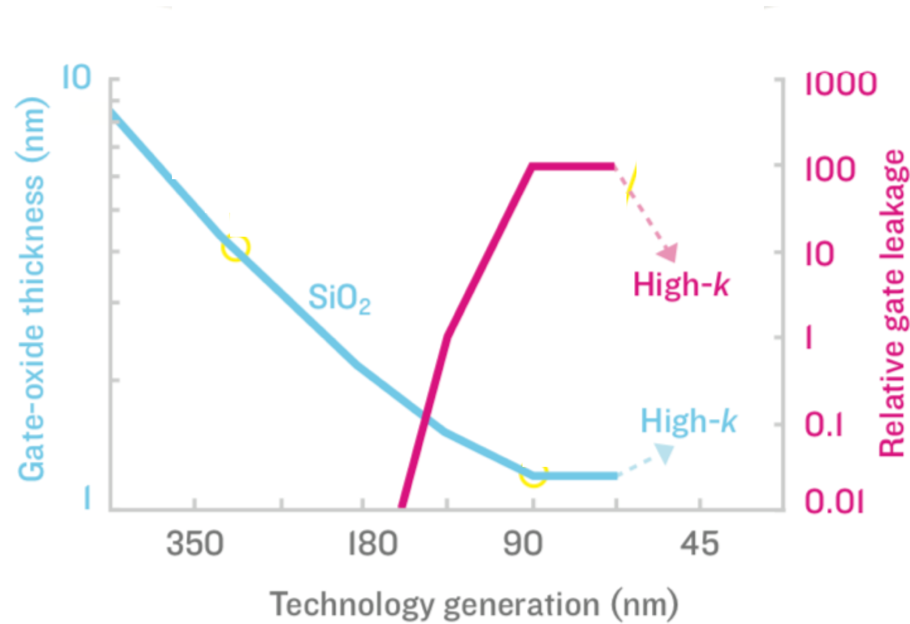
from A. Khakifirooz et al., Symp VLSI Tech 2012

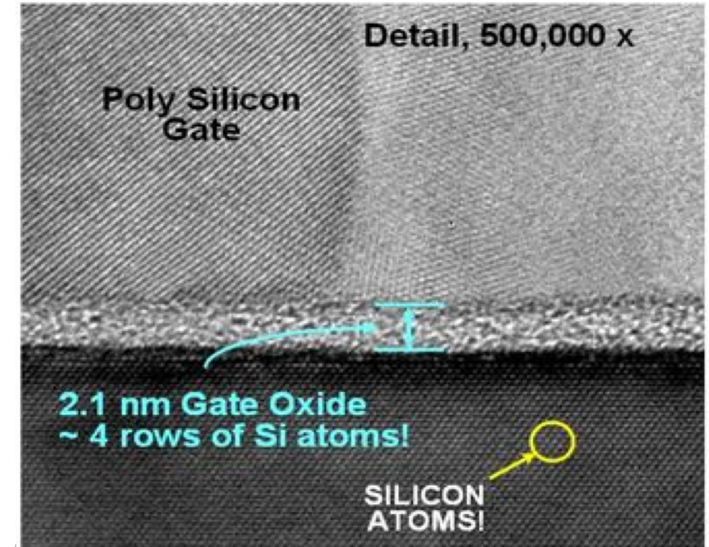




Problem #2: leakage through gate oxide

from M. Bohr et al., IEEE Spectrum Oct. 2007





Gate oxide in a 130nm technology

High-K gate dielectric

US006504214B1

(12) **United States Patent**
Yu et al.

(10) **Patent No.:** **US 6,504,214 B1**
(45) **Date of Patent:** **Jan. 7, 2003**

(54) **MOSFET DEVICE HAVING HIGH-K DIELECTRIC LAYER**

(75) Inventors: **Bin Yu**, Cupertino, CA (US); **Qi Xiang**, San Jose, CA (US)

(73) Assignee: **Advanced Micro Devices, Inc.**, Sunnyvale, CA (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

(21) Appl. No.: **10/044,246**

(22) Filed: **Jan. 11, 2002**

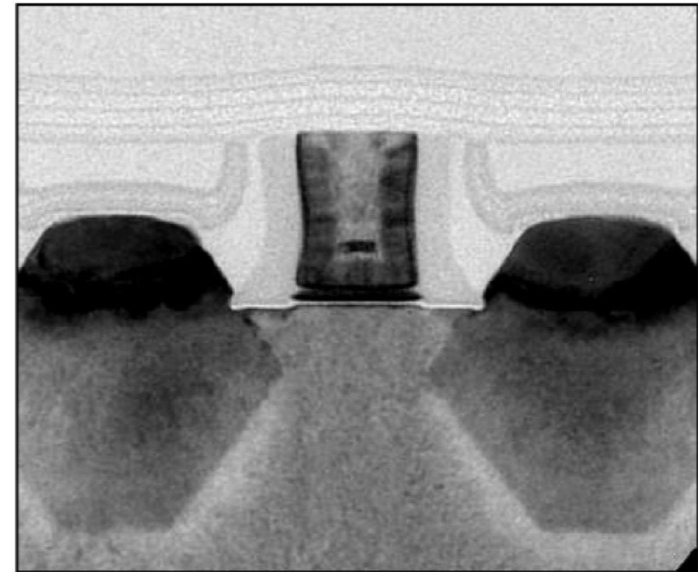
6,020,024	A	2/2000	Maiti et al.	
6,171,910	B1 *	1/2001	Hobbs et al.	438/275
6,210,999	B1	4/2001	Gardner et al.	
6,232,641	B1 *	5/2001	Miyano et al.	257/382
6,261,887	B1 *	7/2001	Rodder	438/218
6,300,202	B1 *	10/2001	Hobbs et al.	438/287
6,346,438	B1 *	2/2002	Yagishita et al.	438/197
6,407,435	B1 *	6/2002	Ma et al.	257/411
2002/0031909	A1 *	3/2002	Cabral et al.	438/655

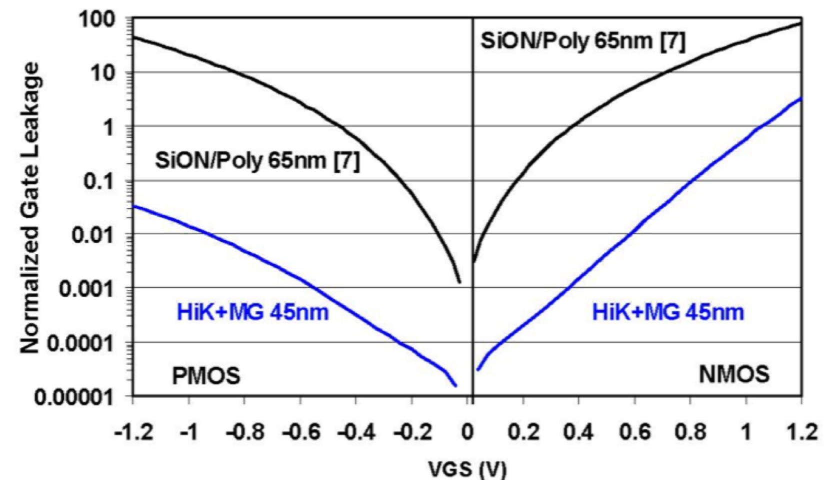
* cited by examiner

Primary Examiner—Carl Whitehead, Jr.
Assistant Examiner—Stephen W. Smoot
(74) *Attorney, Agent, or Firm*—Renner, Otto, Boisselle & Sklar, LLP

(57) **ABSTRACT**

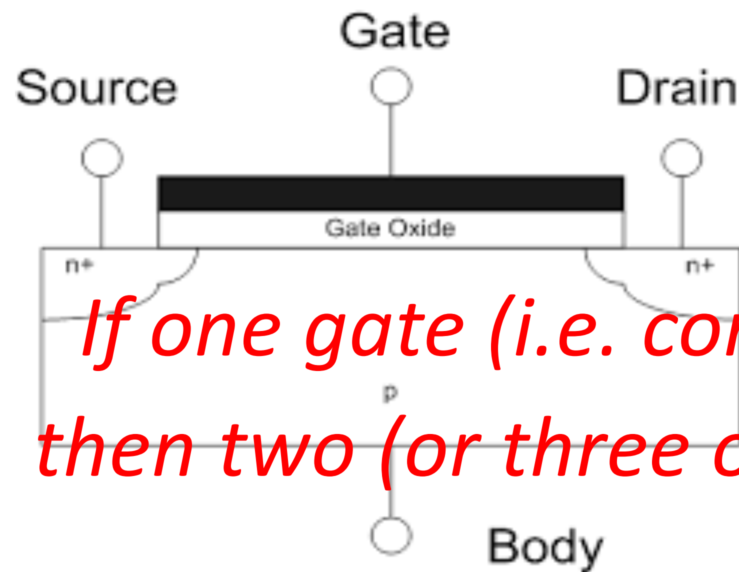
greater detail below. Although other materials can be selected for the gate dielectric **34**, hafnium oxide (e.g., HfO_2), zirconium oxide (e.g., ZrO_2), cerium oxide (e.g., CeO_2), aluminum oxide (e.g., Al_2O_3), titanium oxide (e.g., TiO_2), yttrium oxide (e.g., Y_2O_3) and barium strontium titanate (BST) are example suitable materials for the gate dielectric **34**. In addition, all binary and ternary metal oxides and ferroelectric materials having a K higher than, in one embodiment, about twenty (20) can be used for the gate dielectric **34**.



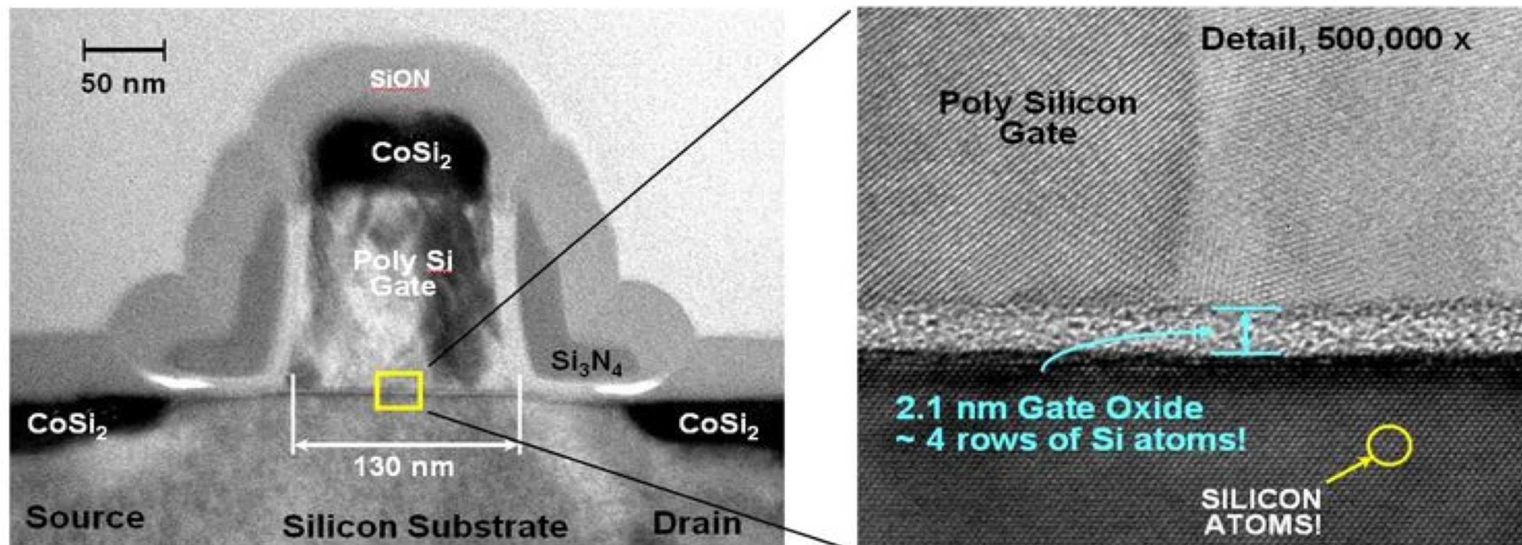


from K. Mistry et al., IEDM 2007

How to get to FINFETs



If one gate (i.e. control surface) is good, then two (or three or four) are even better



Toshiba 1983

Solid-State Electronics Vol. 27, Nos. 8/9, pp. 827–828, 1984
Printed in Great Britain
0038–1101/84 $3.00 + .00
© 1984 Pergamon Press Ltd.

CALCULATED THRESHOLD-VOLTAGE CHARACTERISTICS OF AN XMOS TRANSISTOR HAVING AN ADDITIONAL BOTTOM GATE

(*Received* 30 *May* 1983; *in revised form* 24 *August* 1983)

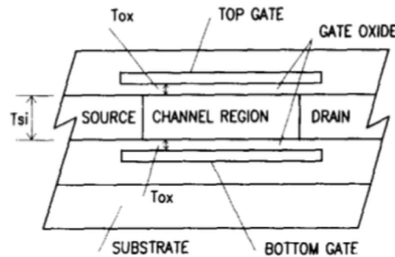


Fig. 1. Schematic cross-sectional structure of an XMOS transistor having an additional bottom gate which is symmetrically placed to a top gate with a channel region between them. "*X*" originates from Greek capital letter of xi as this structure resembles its shape.

Electronic Device Division,
Electrotechnical Laboratory,
Sakura-mura,
Ibaraki, 305,
Japan
T. SEKIGAWA and
Y. HAYASHI

IBM 1994

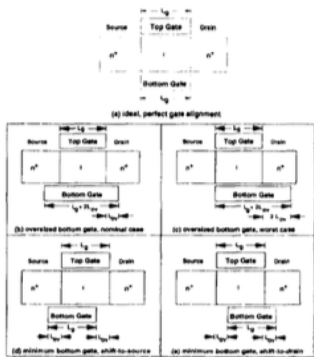


Figure 1: Schematic views of the double-gate SOI MOSFET's.

Hitachi 1991

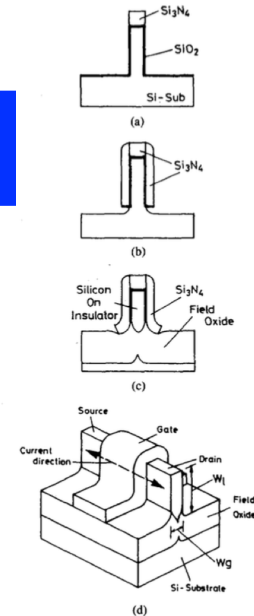
IEEE TRANSACTIONS ON ELECTRON DEVICES, VOL. 38, NO. 6, JUNE 1991 1419

Impact of the Vertical SOI "DELTA" Structure on Planar Device Technology

Digh Hisamoto, *Member, IEEE*, Toru Kaga, *Member, IEEE*, and Eiji Takeda, *Senior Member, IEEE*

***Abstract*—A fully depleted lean channel transistor (DELTA) with its gate incorporated into a new vertical ultra-thin SOI structure is presented. In the deep-submicrometer region, selective oxidation produces and isolates an ultra-thin SOI MOSFET that has high crystalline quality, as good as that of conventional bulk single-crystal devices. Experiments and three-dimensional simulations have shown that this new gate structure has effective channel control, and that the vertical ultra-thin SOI structure provides superior device characteristics: reduction in short-channel effects, minimized subthreshold swing, and high transconductance.**

is required. Moreover, it is evident that these structures are difficult to contact to the substrate, and thus suffer from a substrate floating effect.



(a)–(c) Process flow of selective oxidation. (d) Schematic cross section of DELTA.

Berkely 2000

2320 IEEE TRANSACTIONS ON ELECTRON DEVICES, VOL. 47, NO. 12, DECEMBER 2000

FinFET—A Self-Aligned Double-Gate MOSFET Scalable to 20 nm

Digh Hisamoto, *Member, IEEE*, Wen-Chin Lee, Jakub Kedzierski, Hideki Takeuchi, Kazuya Asano, *Member, IEEE*, Charles Kuo, Erik Anderson, Tsu-Jae King, Jeffrey Bokor, *Fellow, IEEE*, and Chenming Hu, *Fellow, IEEE*

***Abstract*—MOSFETs with gate length down to 17 nm are reported. To suppress the short channel effect, a novel self-aligned double-gate MOSFET, FinFET, is proposed. By using boron-doped $Si_{0.4}Ge_{0.6}$ as a gate material, the desired threshold voltage was achieved for the ultrathin body device. The quasiplanar nature of this new variant of the vertical double-gate MOSFETs can be fabricated relatively easily using the conventional planar MOSFET process technologies.**

***Index Terms*—Fully depleted SOI, MOSFET, poly SiGe, short-channel effect.**

I. INTRODUCTION

TO DEVELOP sub-50-nm MOSFETs, the double-gate structure has been widely studied. This is because

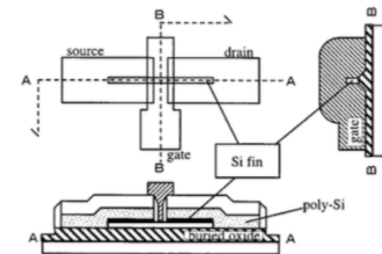
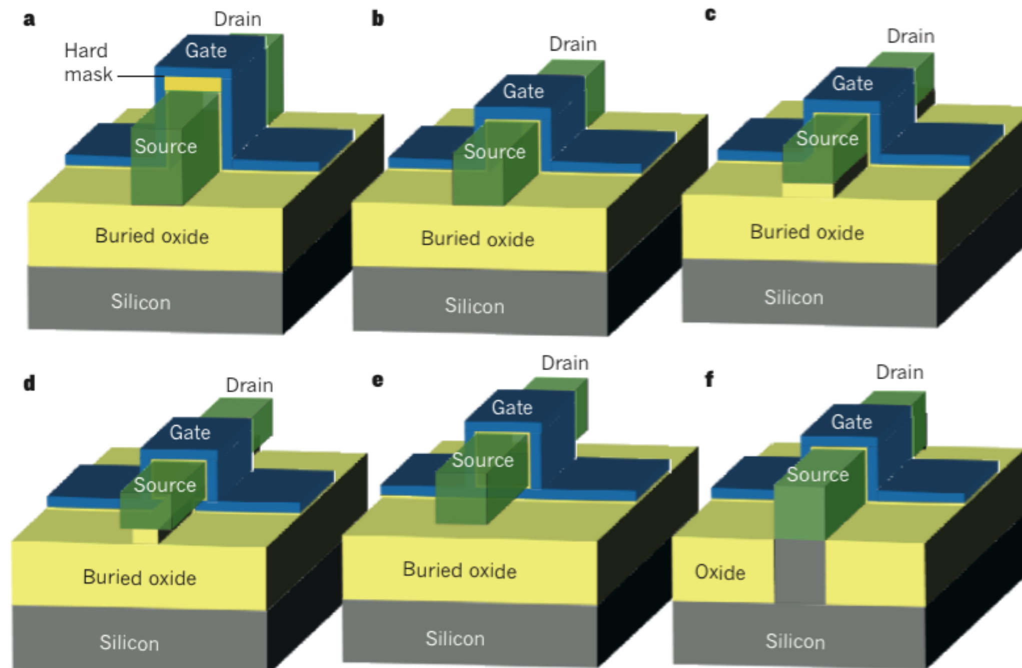


Fig. 1. FinFET typical layout and schematic cross sectional structures.

(*) Apparently a Japanese 1980 patent pre-dates all these publications but I have not been able to locate it.

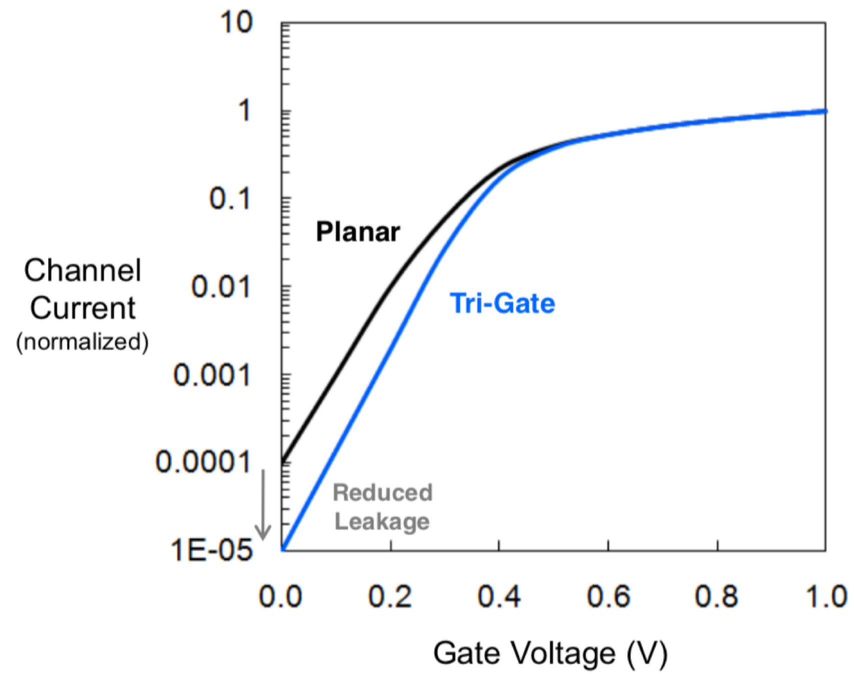
Multi-gate devices

- Intel: Tri-gate
- TSMC, GF, Samsung: Finfets



from: Ferain, "Multigate transistors as the future of classical metal–oxide–semiconductor field-effect transistors",

Subthreshold slope improvement

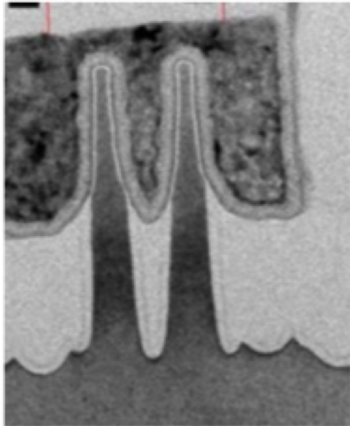


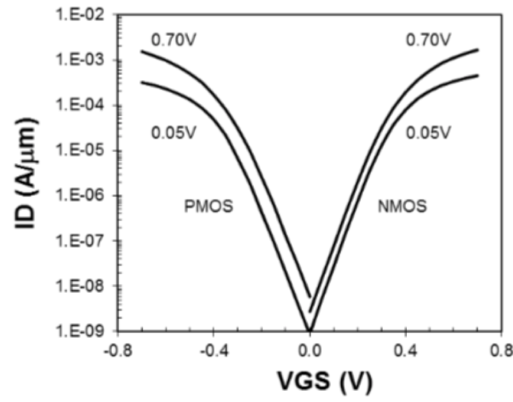
Intel @ 22 nm

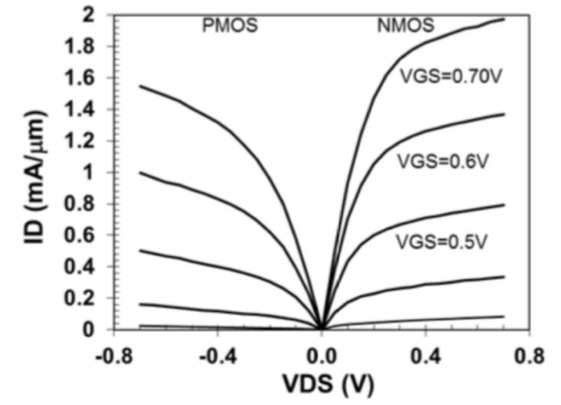
from: M. Bohr (Intel) , Intel's Revolutionary 22 nm Transistor Technology, May 2011

10 nm Finfet

from: A. Auth (Intel), A 10nm High Performance and Low Power CMOS Technology…, IEDM 2017



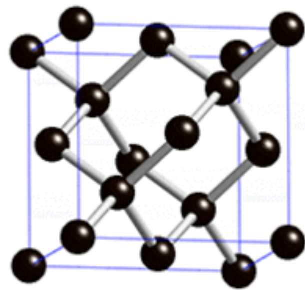


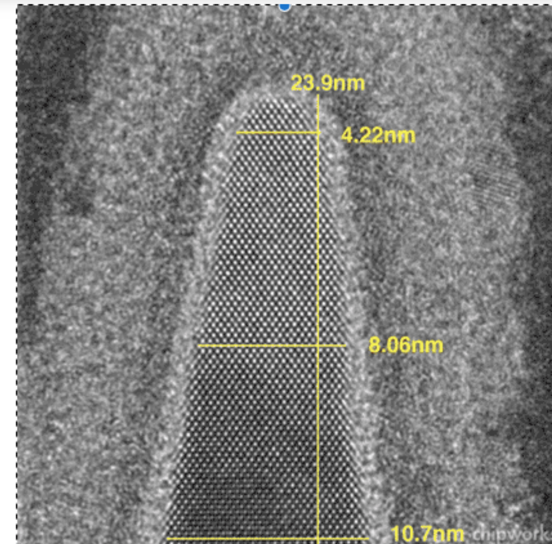


FinFET dimensions

From the ITRS 2011 report

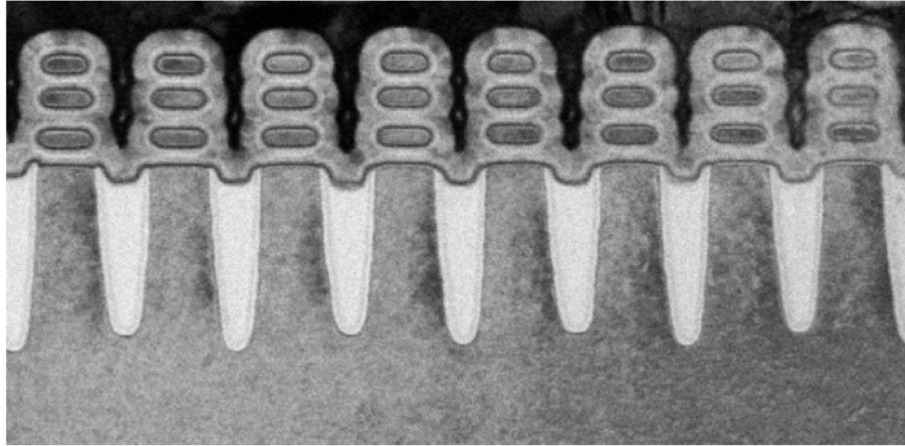
Year of Production	2013	2015	2017	2019	2021	2023	2025	2028
FinFET Fin Width (new) (nm)	**7.6**	**7.2**	**6.8**	**6.4**	**6.1**	**5.7**	**5.4**	**5.0**

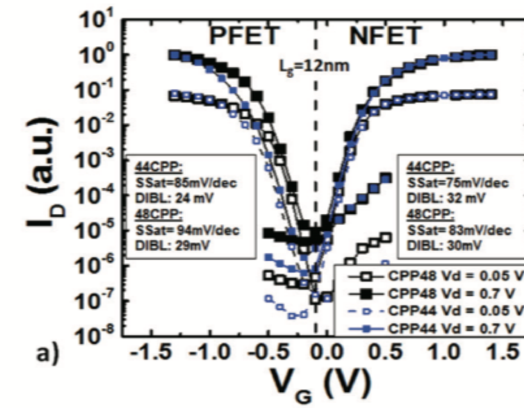




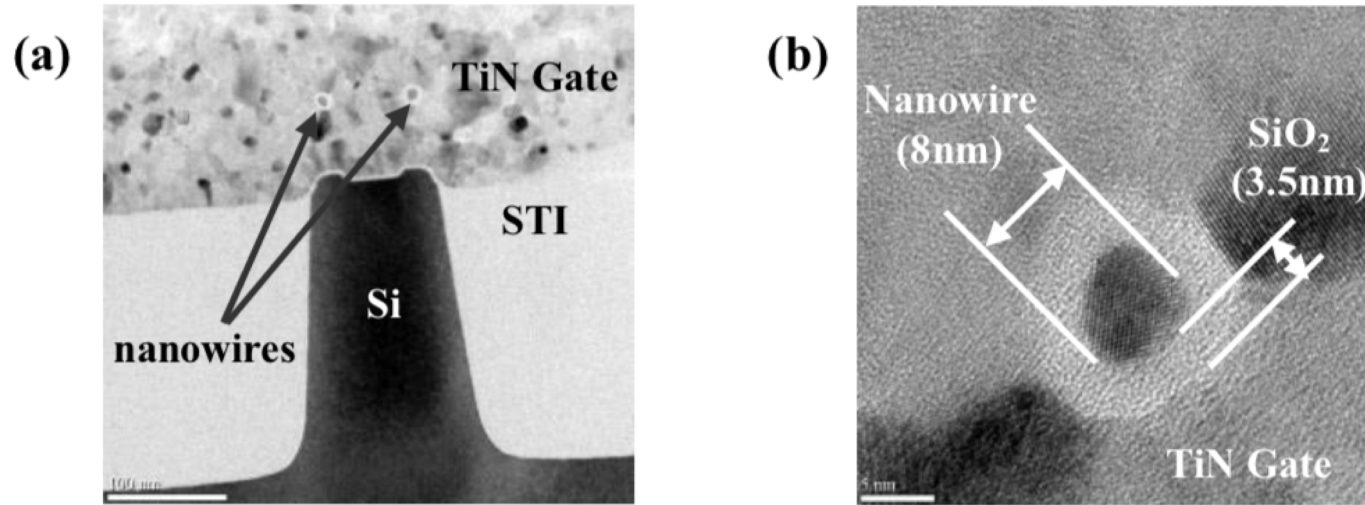
GAA nanowire transistor = FINFET++

From N. Loubet – IBM –
Stacked Nanosheet Gate-All-Around Transistor to Enable Scaling Beyond FinFET – VLSI Symp 2017





GAA nanowire transistor

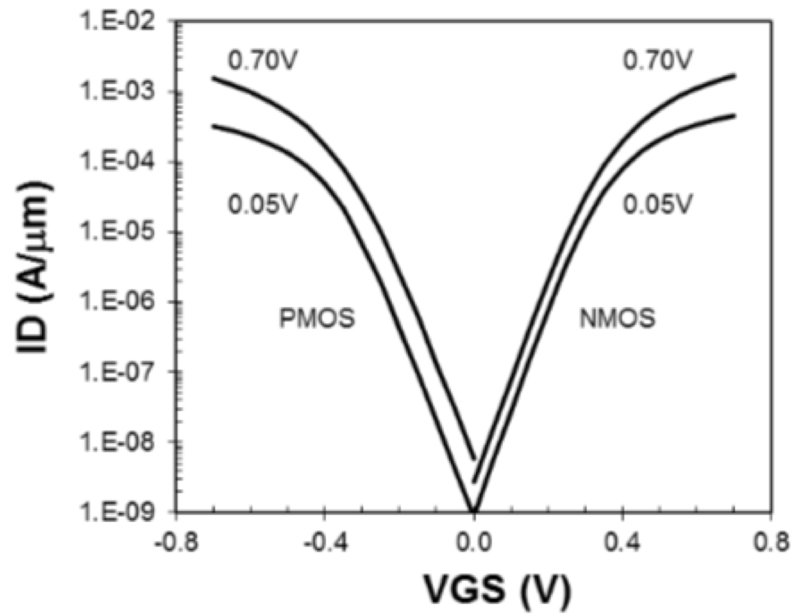


from: Kyoung Hwan Yeo et al., Gate-All-Around (GAA) Twin Silicon Nanowire MOSFET…, IEDM 2006

and yet…

https://en.wikichip.org/wiki/technology_node

Number of Foundries with a Cutting Edge Logic Fab										
SilTerra										
X-FAB										
Dongbu HiTek										
ADI	ADI									
Atmel	Atmel									
Rohm	Rohm									
Sanyo	Sanyo									
Mitsubishi	Mitsubishi									
ON	ON									
Hitachi	Hitachi									
Cypress	Cypress	Cypress								
Sony	Sony	Sony								
Infineon	Infineon	Infineon								
Sharp	Sharp	Sharp								
Freescale	Freescale	Freescale								
Renesas (NEC)	Renesas	Renesas	Renesas	Renesas						
SMIC	SMIC	SMIC	SMIC	SMIC						
Toshiba	Toshiba	Toshiba	Toshiba	Toshiba						
Fujitsu	Fujitsu	Fujitsu	Fujitsu	Fujitsu						
TI	TI	TI	TI	TI						
Panasonic	Panasonic	Panasonic	Panasonic	Panasonic	Panasonic					
STMicroelectronics	STM	STM	STM	STM	STM					
UMC	UMC	UMC	UMC	UMC	UMC					
IBM	IBM	IBM	IBM	IBM	IBM	IBM				
AMD	AMD	AMD	GlobalFoundries	GF	GF	GF	GF			
Samsung	Samsung	Samsung	Samsung	Samsung	Samsung	Samsung	Samsung	Samsung	Samsung	
TSMC	TSMC	TSMC	TSMC	TSMC	TSMC	TSMC	TSMC	TSMC	TSMC	
Intel	Intel	Intel	Intel	Intel	Intel	Intel	Intel	Intel	Intel	Future
180 nm	130 nm	90 nm	65 nm	45 nm/40 nm	32 nm/28 nm	22 nm/20 nm	16 nm/14 nm	10 nm	7 nm	5 nm



New Ideas

Negative Capacitance Transistors (Ferroelectric Transistors)

Tunneling Transistors

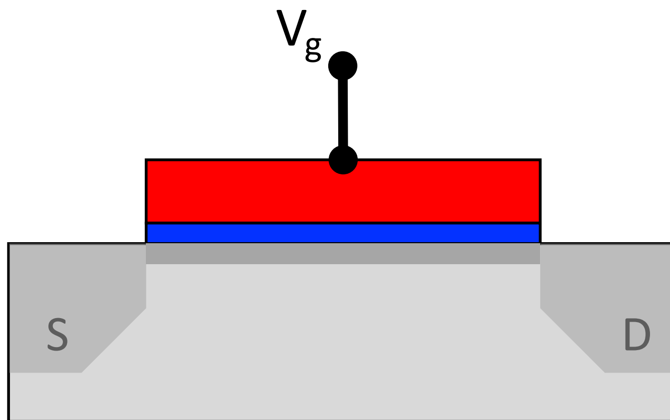
What does Boltzmann have to do with microelectronics?

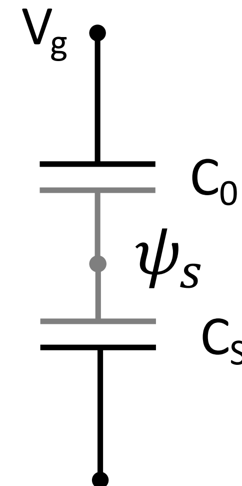
$$SS = \frac{\partial V_g}{\partial(logI_d)} = \frac{\partial V_g}{\partial \psi_s} * \frac{\partial \psi_s}{\partial(logI_d)}$$

$$\min\left(\frac{\partial \psi_s}{\partial(logI_d)}\right) = Ln(10) * \frac{k_B T}{q} \approx 60 \; \frac{mV}{decade}$$

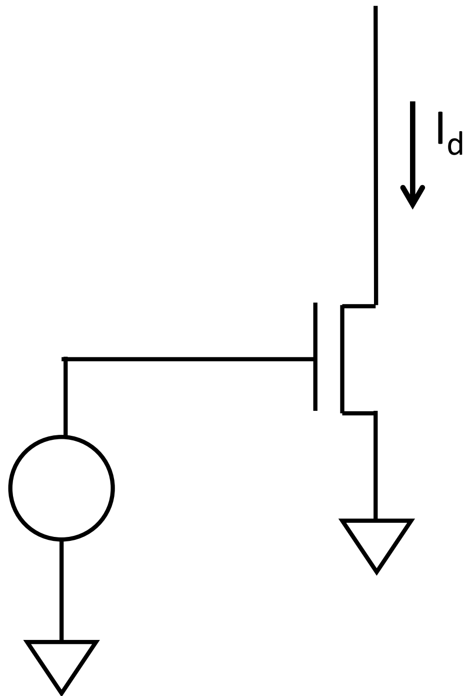
$$\frac{\partial V_g}{\partial \psi_s} = 1 + \frac{C_S}{C_o}$$

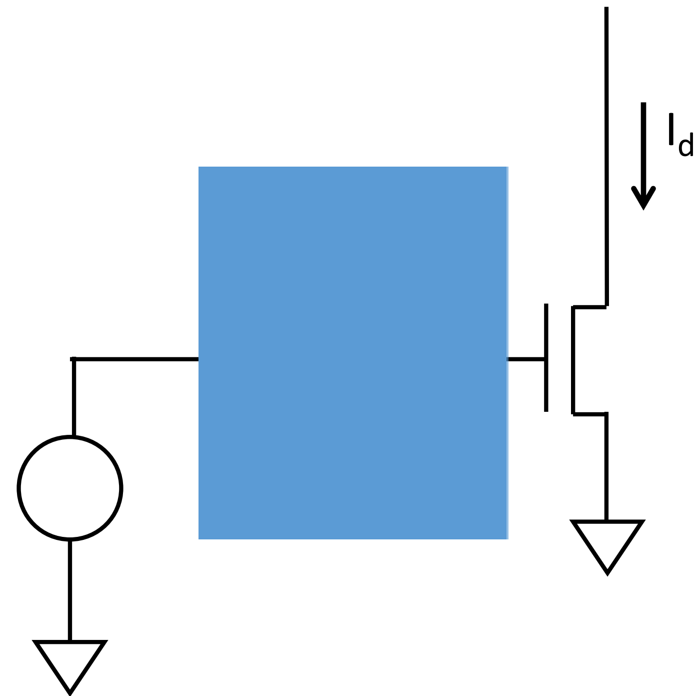
for a full derivation of this formula, see Y. Tsividis' book, page 208



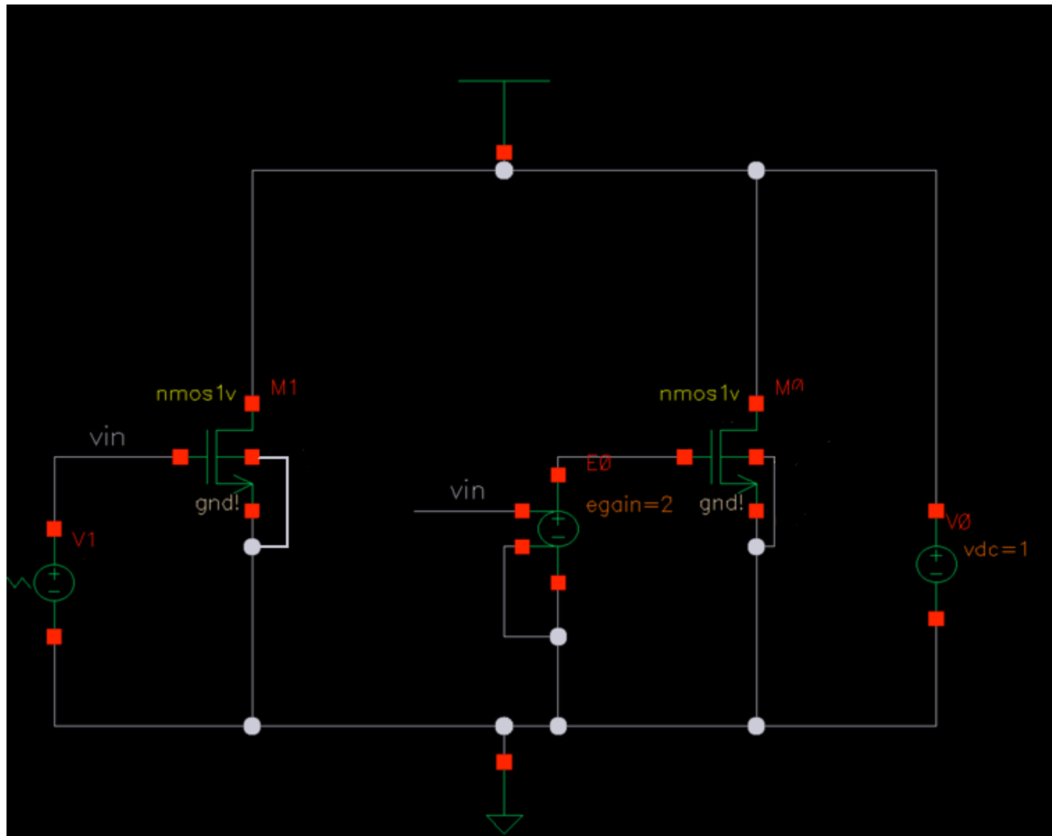


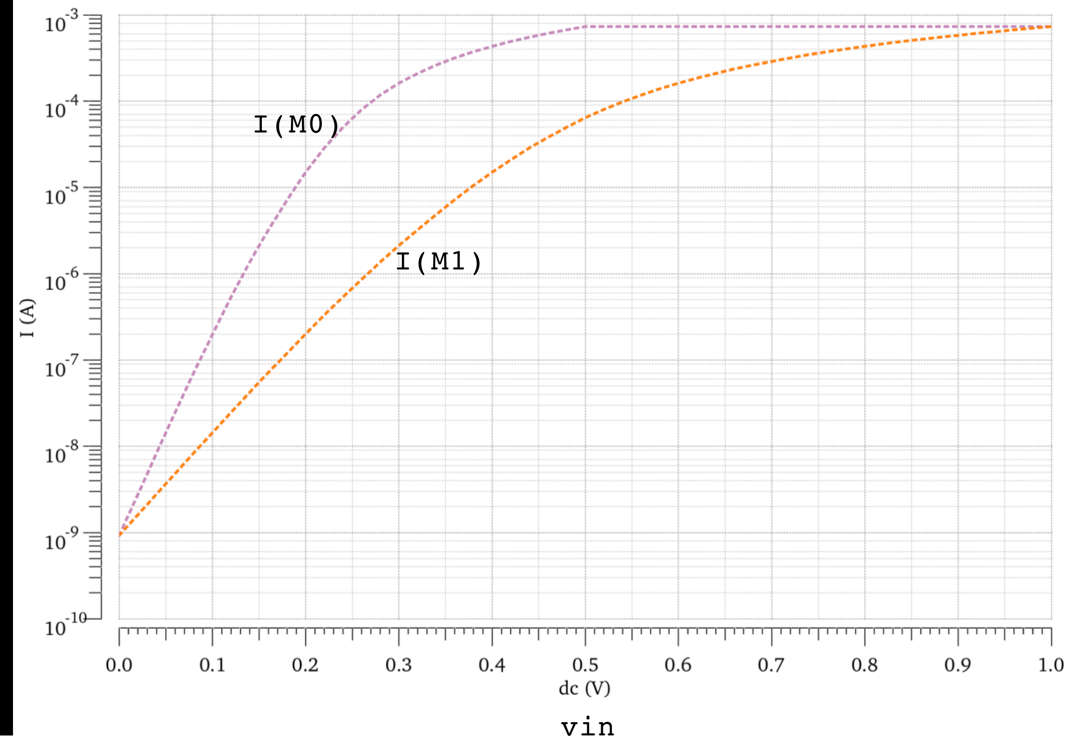
[A little Gedankenexperiment (1)]



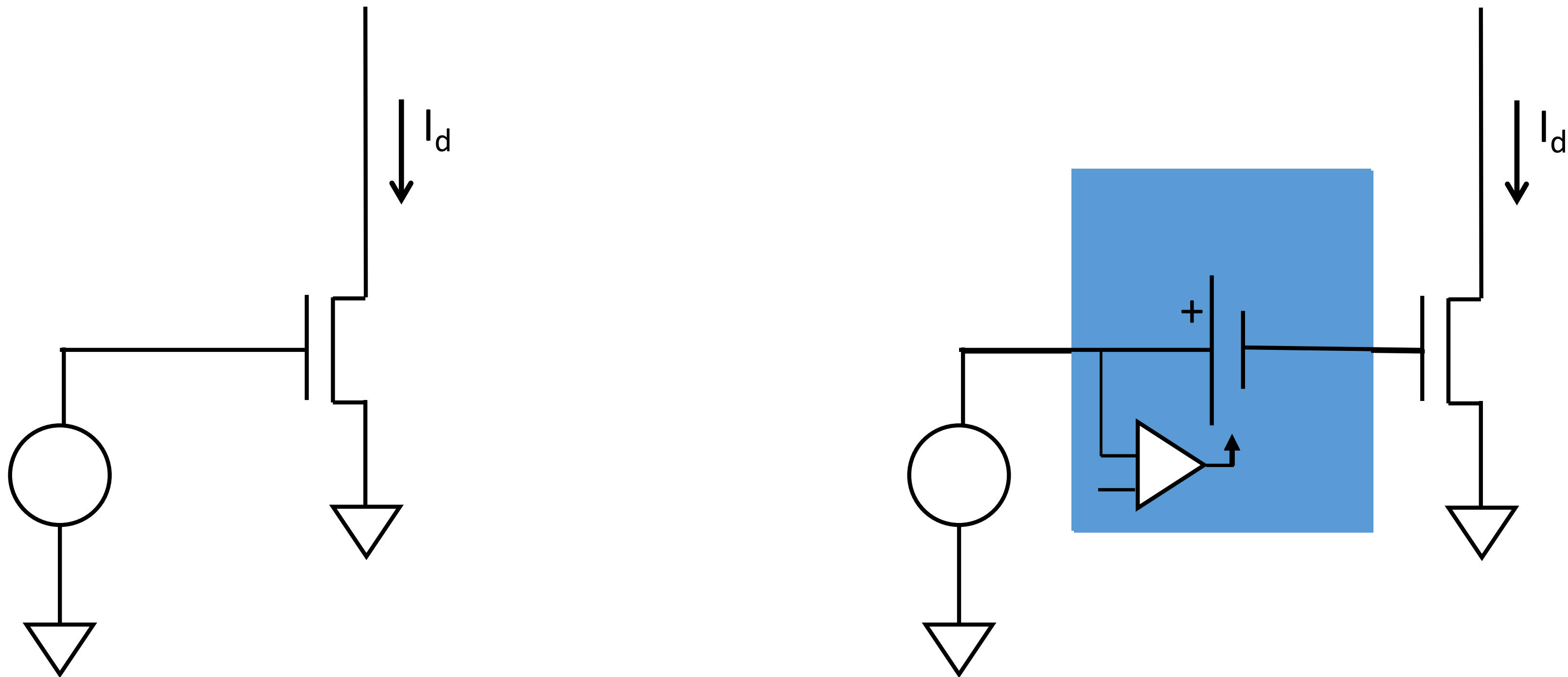


[A little Gedankenexperiment(2)]

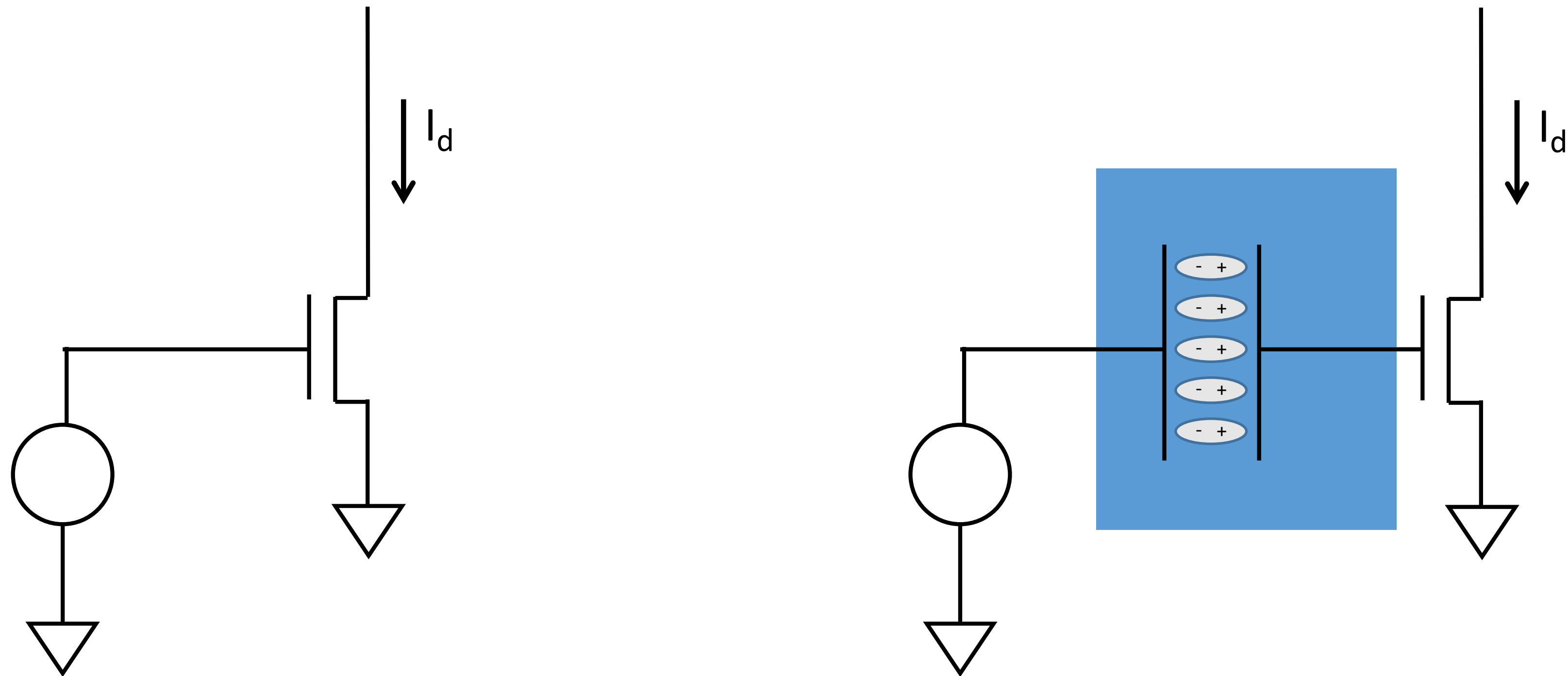




[A little Gedankenexperiment (3)]

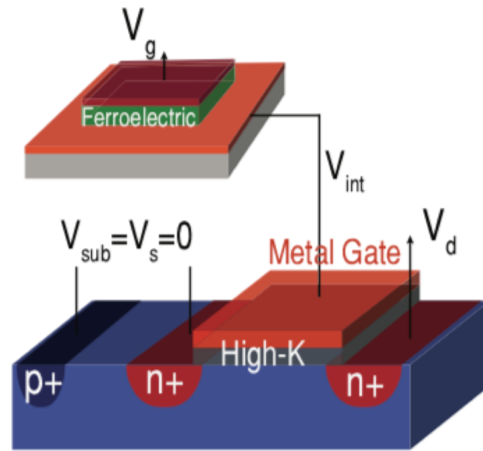


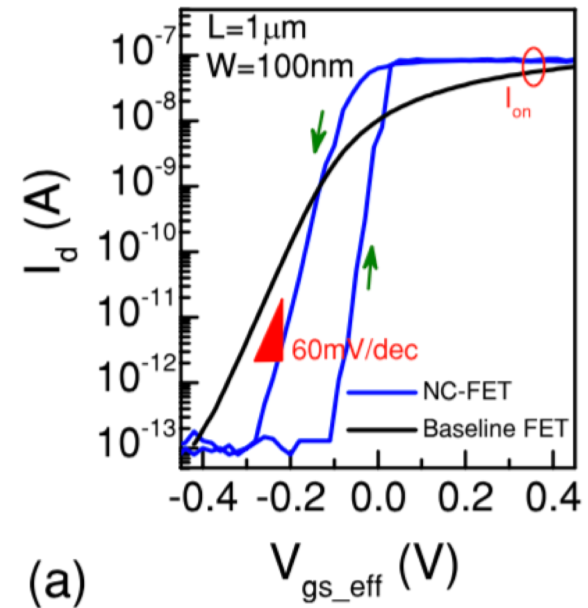
[A little Gedankenexperiment (4)]



A real exercise

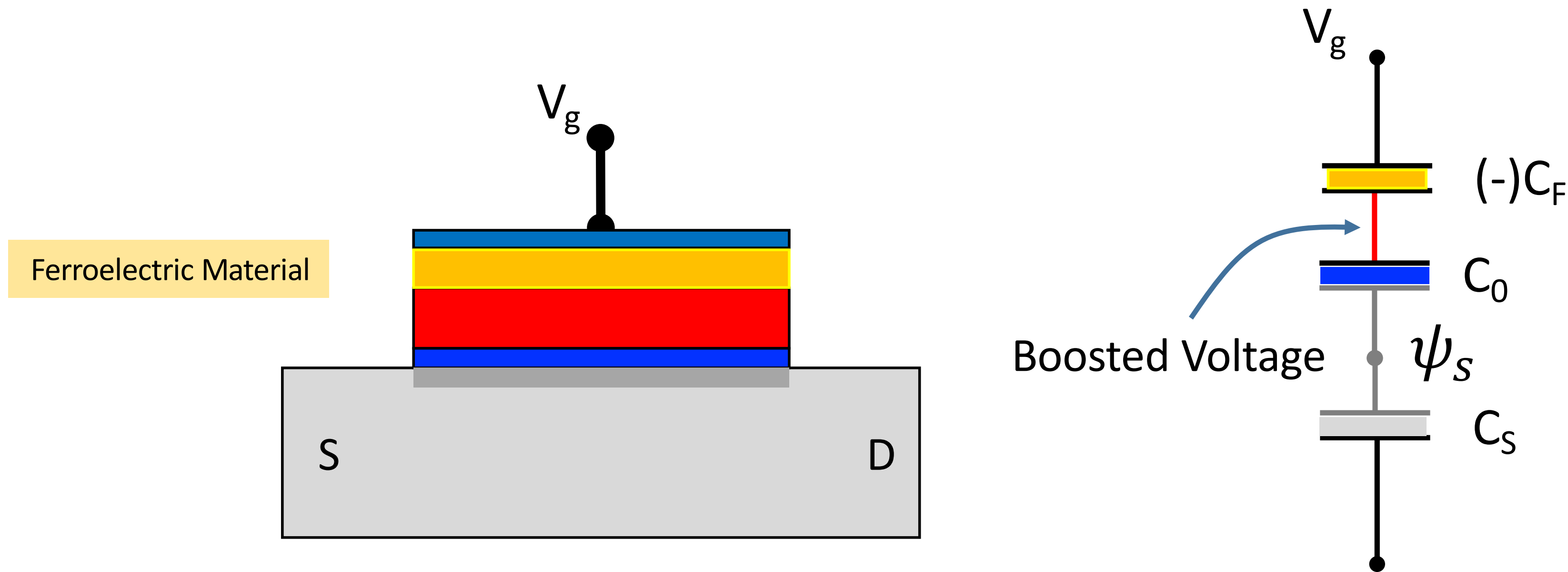
PZT (Lead-Zirconate-Titanate) capacitor





from: A. Saeidi et al, , "Negative Capacitance as Digital and Analog Performance Booster for Complementary MOS Transistors", EPFL, arXiv:1804.09622v2

NC gate stack



@IEDM 2018, Negative Capacitance devices occupied 3 sessions and are mentioned in 16 papers

Other non-Boltzmann limited devices

Tunneling Transistors: History

1952 PROCEEDINGS OF THE I.R.E. 1377

Junction Fieldistors*

O. M. STUETZER†, ASSOCIATE, IRE

Summary—A high-impedance, input-low impedance output amplifying device which utilizes surface conductivity control in the neighborhood of a *p-n* junction is described. The transconductances of the order of 1,000 micromhos can be reproduced at very low frequencies.

Introduction

A SCHEMATIC DIAGRAM (Fig. 1) shows the essentials of the experimental device we shall discuss. A *p-n* junction is biased in the reverse direction by a current I_a. A control electrode is attached in close proximity to the junction. The distance d is of

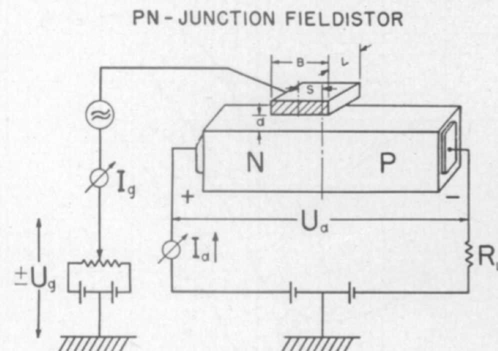
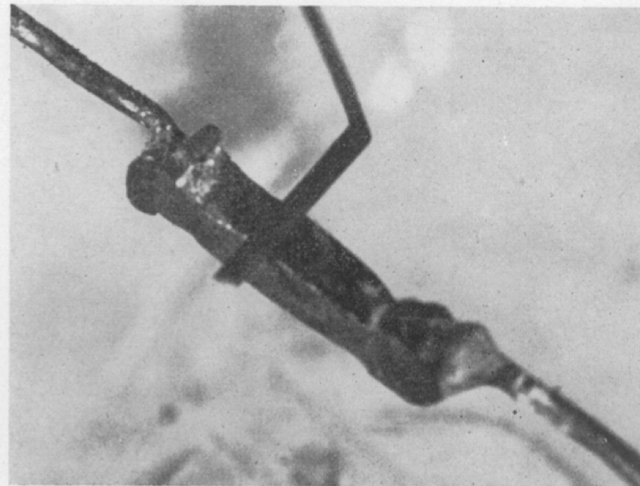


Fig. 1—Schematic diagram of arrangement.

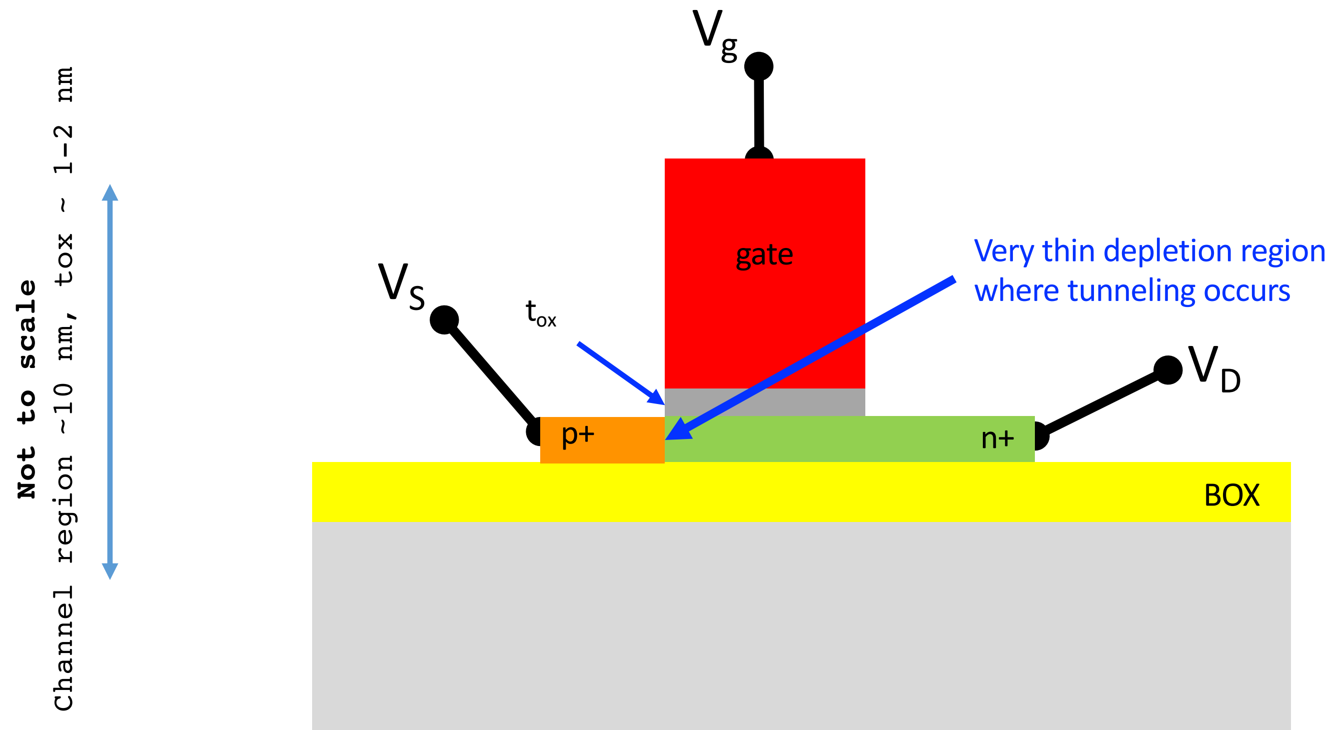
water free oils) increases the effect roughly in proportion with their dielectric constant.

However, if we introduce between surface and control electrode a liquid with a reasonably high polar moment, G_m changes sign and increases drastically. The output



(a)

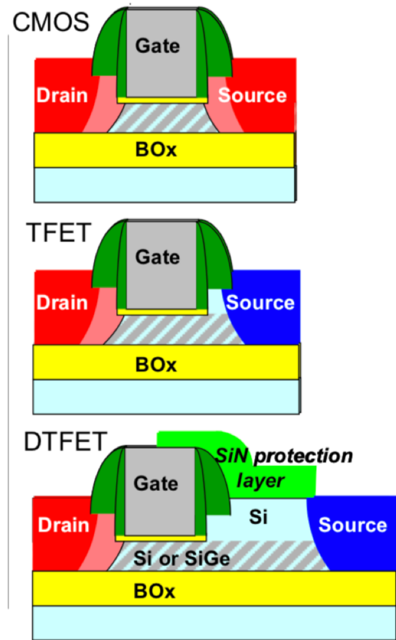
Tunneling Transistors: Principle

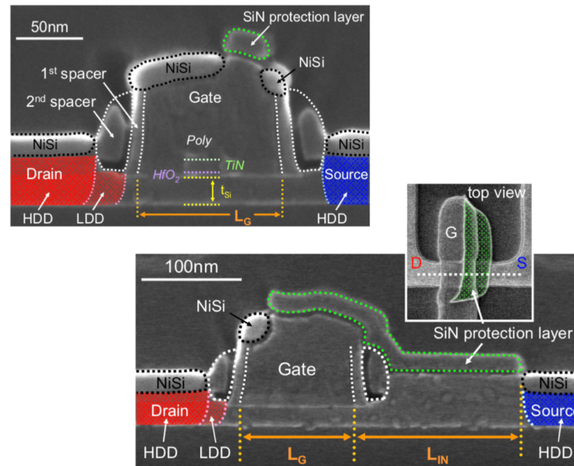


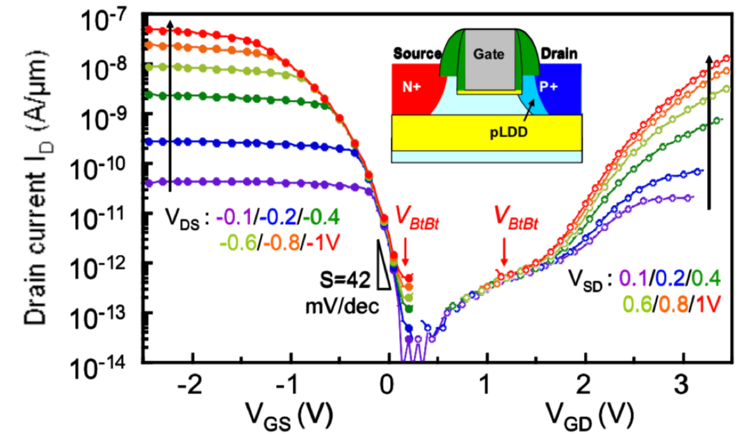
- Very highly doped S and D
- $V_S < V_D$

Tunneling Transistors: Example

from: F. Mayer et al., "Impact of SOI, Si1-xGexOI and GeOI substrates on CMOS compatible Tunnel FET performance", IEDM 2008







Summary: low supply voltage devices

- Currently with devices powered between 0.8 and 1.2 V, we use about 400-500 mV (~70-80 mV/dec * 6 dec) in just turning them on.

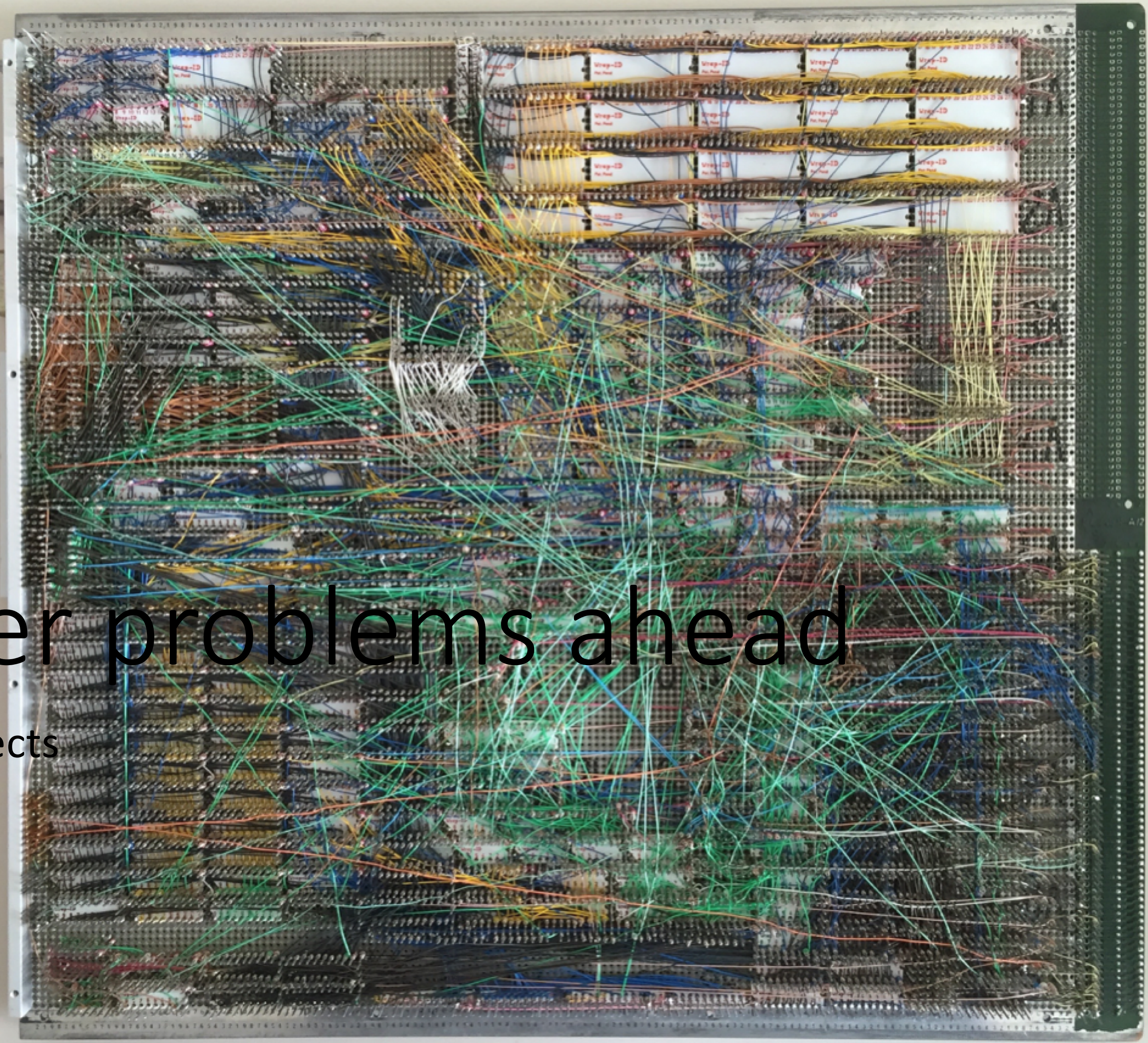
- If one could reduce reduce the “transition region” to about 50 mV (SS ~ 50/6 = 8-10 mV/dec)
  - Then it would be conceivable to have a digital logic supply at 150 mV therefore saving:

$$\text{Power saving: } (1.0)^2 / (0.15)^2 = 44 \text{ times}$$

(i.e. you could recharge your mobile phone less than once per month)

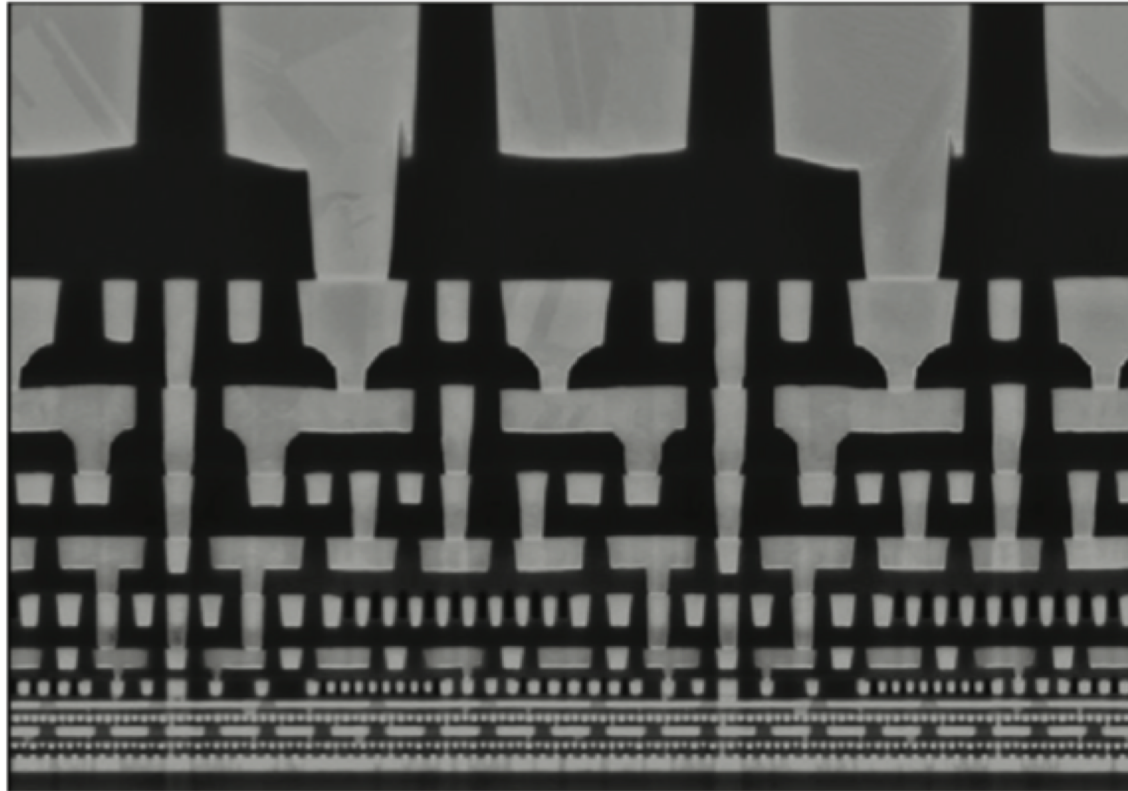
Other problems ahead

Interconnects



What is more important: Transistors or wires?

from K. Fischer – Intel - Low-k Interconnect Stack with multiple-layer gap…, IITC Conf. 2015



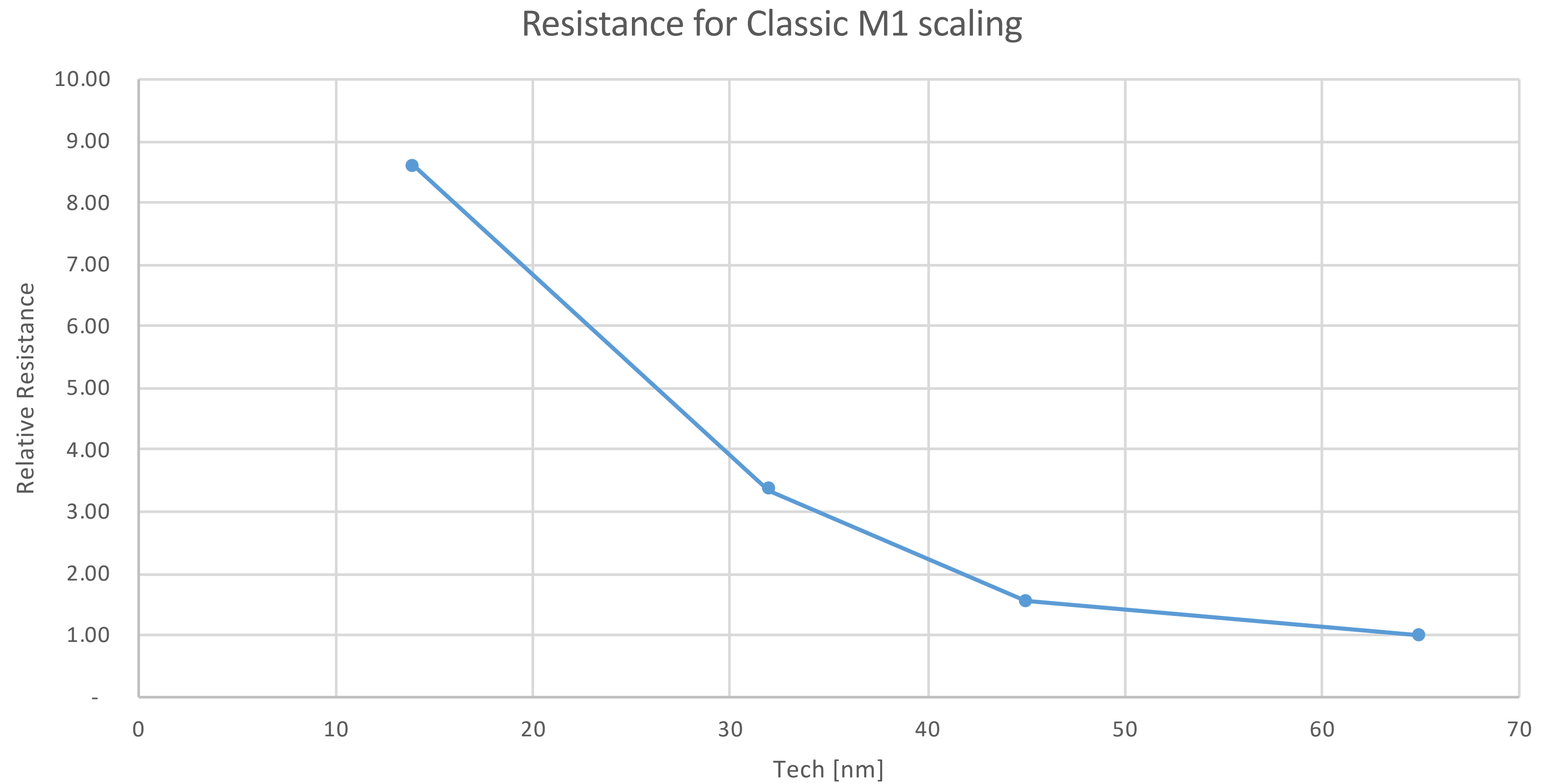
A 14 nm FINFET process metal stack

Metal	Pitch [nm]	Metal tickn.[nm]
0	56	40
1	81	42
2	73	40
3	76	37
4	80	75
..		
11	1000	1000

Metal/Gate ratio in different technologies

Tech [nm]	M1 Width/Space	Ratio M1 Width/Lgate
250	320/320	1.28
130	160/160	1.53
65	90/90	1.5
28	50/50	1.7
14 (Finfet)	28/28	2

Wire resistance in “classical” scaling



Ultra-scaled metals

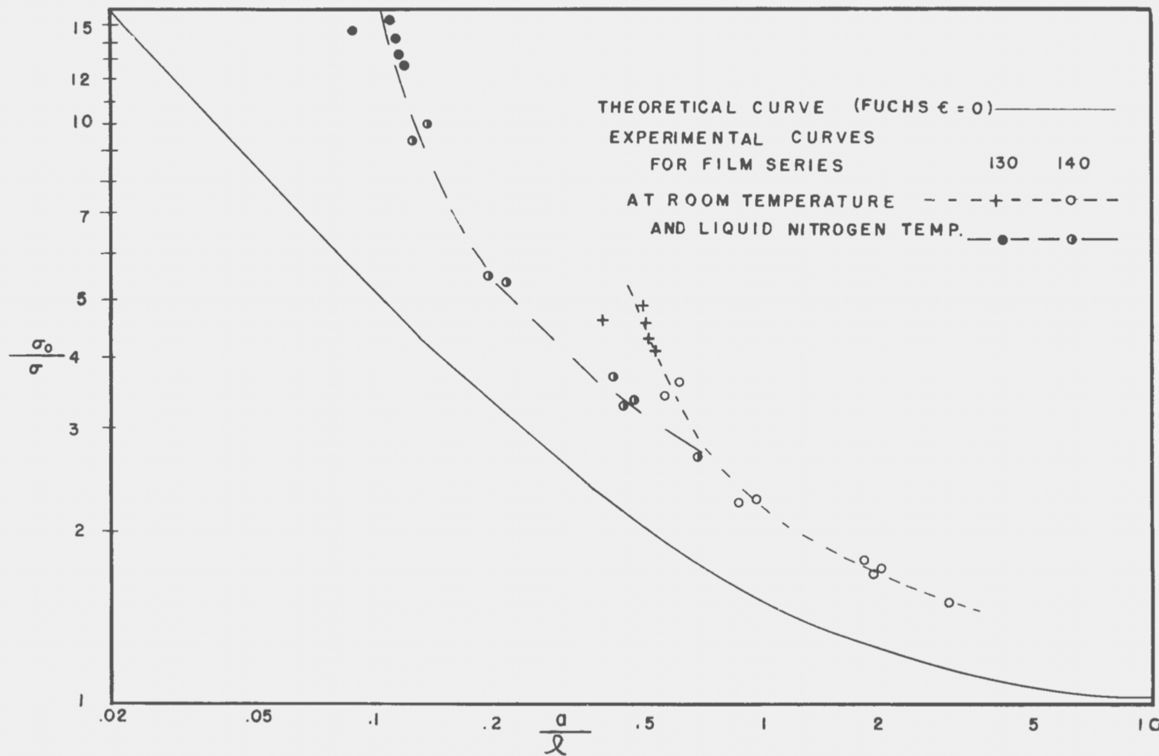
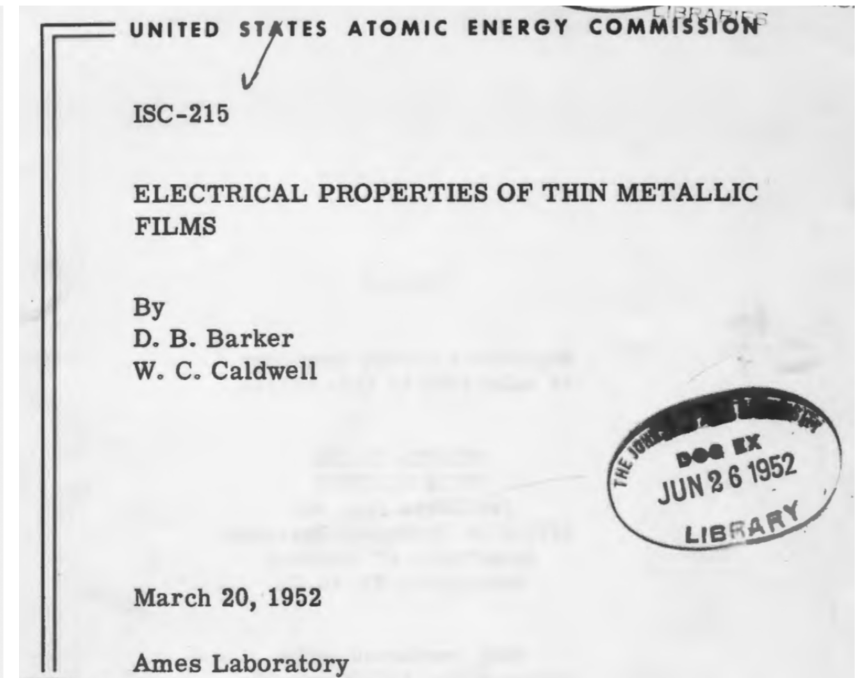


FIG. I. ELECTRICAL CONDUCTIVITY OF THIN SILVER FILMS.


UNITED STATES ATOMIC ENERGY COMMISSION

ISC-215

ELECTRICAL PROPERTIES OF THIN METALLIC FILMS

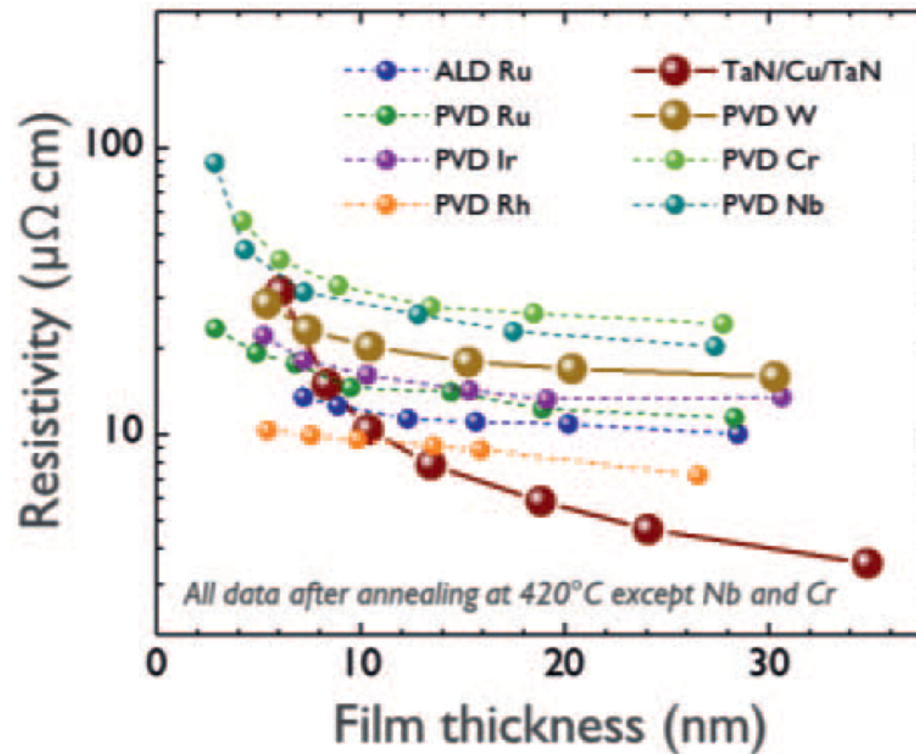
By
D. B. Barker
W. C. Caldwell

DOC EX
JUN 26 1952
LIBRARY

March 20, 1952

Ames Laboratory

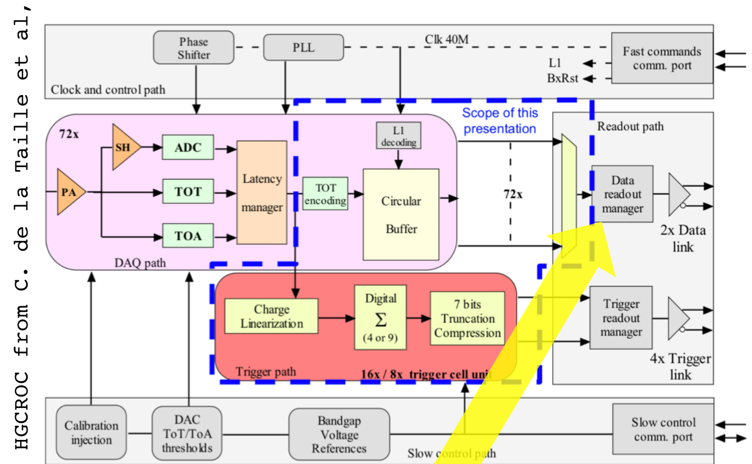
Ultra-scaled metals (2)



from K. Croes, IEDM 2018

What’s next for HEP

CMS-HGCROC BE Architecture



78 x Channels (ADC,TOA,TOT)

L1A
Memory
512 x [78x30] bit

On-chip Event Builder

On-chip Event Builder

...

...

On-chip Event Builder

Serializer A

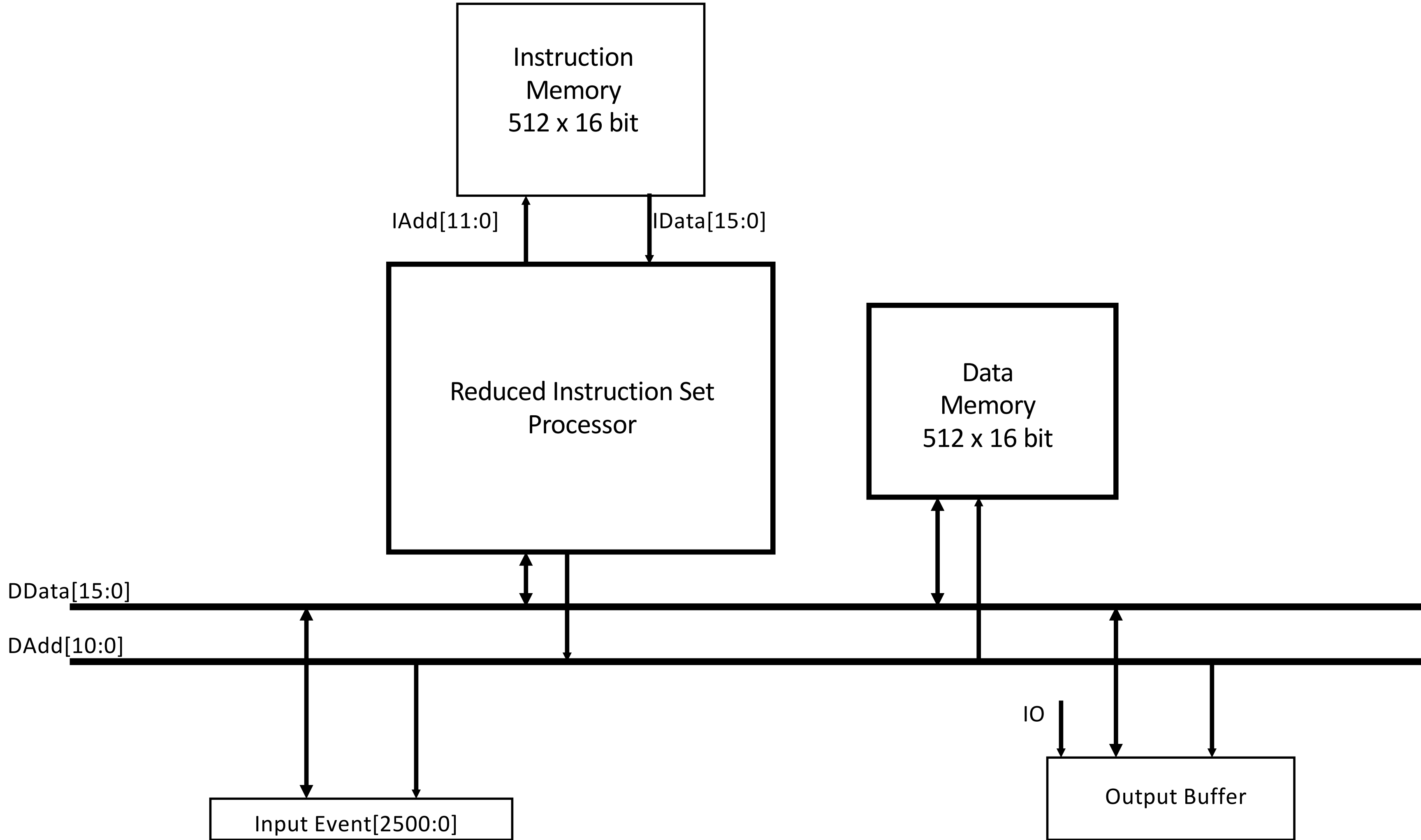
Serializer B

Event Dispatcher

L1A

Serializer Arbiter

Event Builder Architecture for HGCROC



Power comparison: 16-bit RISC for FE electronics

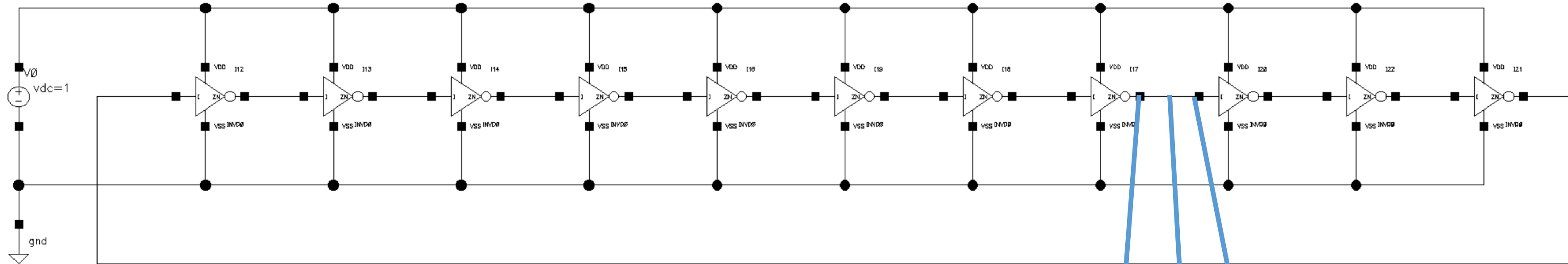
Processors + I/D memories

Full Verilog, synthesized and routed with Innovus tool, no optimization

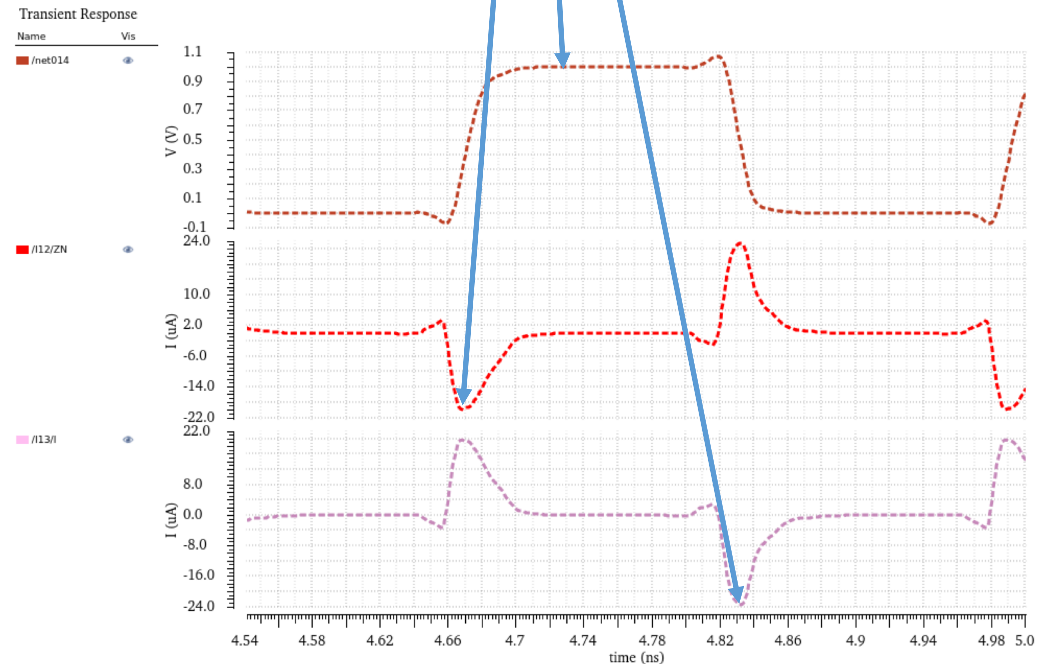
	130 nm	65 m	28 nm (estimate)	7 nm FF (from 1,2)	5 nm (speculative NC)[$]
f_{osc} [MHz]	300	300	300	300	300
V_{dd} [V]	1.2	1.0	0.9	0.75	(.25)
N_cells	62325	65428			
Area [um^2]	1,501,000	454,000			
P_{total} [mW] (one processor)	102	40.9	(14)	(2.2)	(0.24)
P/channel [{*}] [mW]	13	5	(1.7)	(0.26)	(0.03)

{*} Assuming 10 processors per FE chip
{$} Assuming only Vdd scaling from 7 nm FF
{1} Shien-Yang Wu et al, IEDM 2013
{2} Shien-Yang Wu et al, IEDM 2016

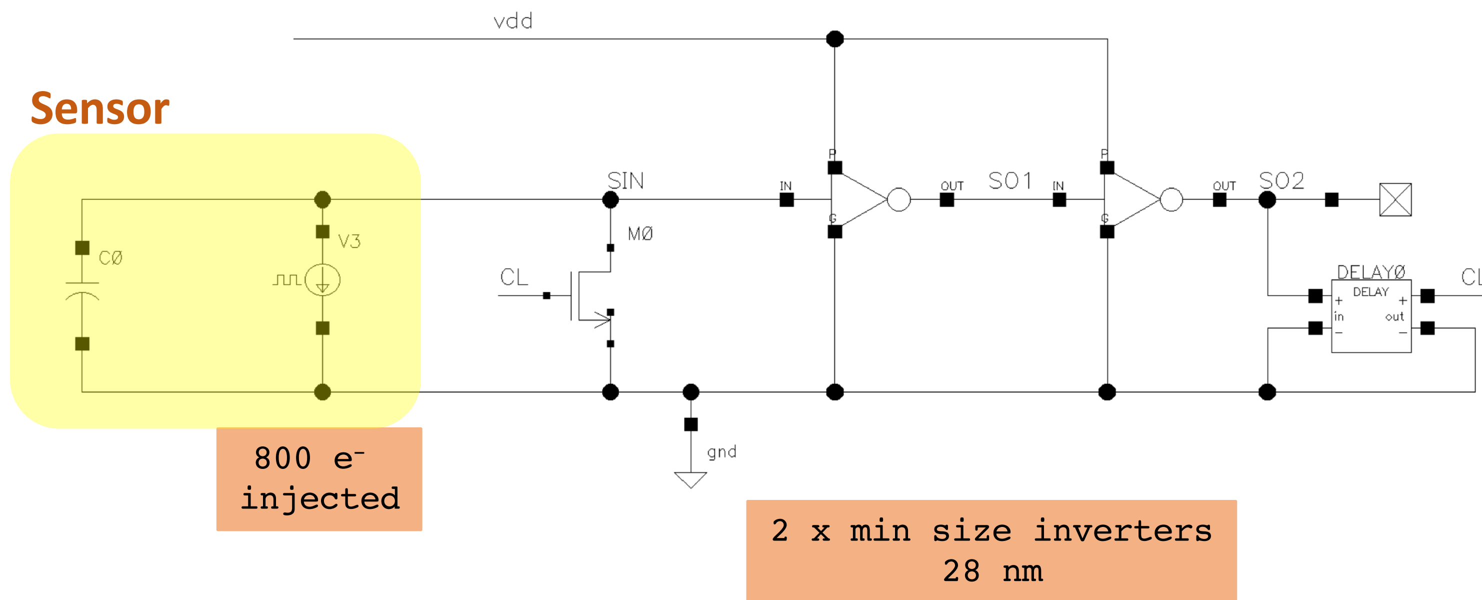
How many electrons are needed to switch a logic gate ?



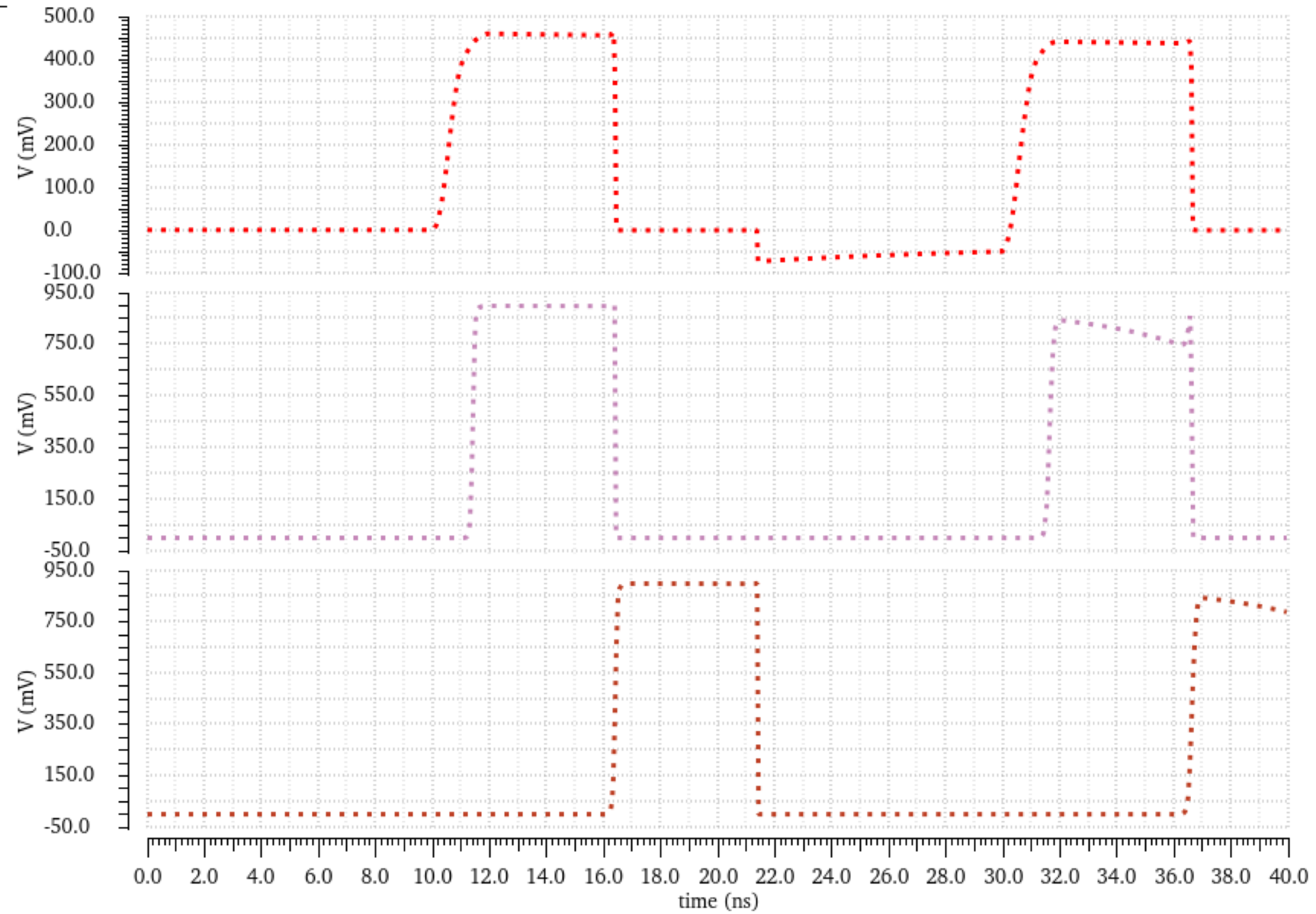
- **65 nm: ~ 2500 e$^-$**
- **28 nm: ~ 850 e$^-$**



Digital Amplifier for small cell Si sensor



Digital Amplifier (2)



And what about 5 nm?

- A 5 nm “transistor” switches with < ~100 e^- input signal
- That is the signal produced by a MIP particle in about 1um of silicon

Significant issues still exist in the integration of an appropriate sensor with very low parasitic capacitances (intrinsic and extrinsic), but from the point of view of the sensing electronics, it may well be possible to design a pixelized detector with sufficiently small cells to be read out entirely by simple inverters.

Take home message

- ”Brute-force” growth (was called “scaling”) is being replaced by “more-sweat” growth
  - More sweat also implies more investments, much more investments (especially human)!
- But lots of opportunities are still open for creative designers.
- Much functionality can still be added to detectors for physics
  - The impact of “digital” is still very small in HEP, replace “quantity” of data with “quality” of data
  - Beware, gain in analog may even be < 1
- More exotic technologies (TSVs, wafer stacking, adv. packaging...) may become available also for low-volume, but history teaches that one should bet on mainstream opportunities
- New engineering “structures” ~~may be~~ mandatory to exploit the above!

Thank you

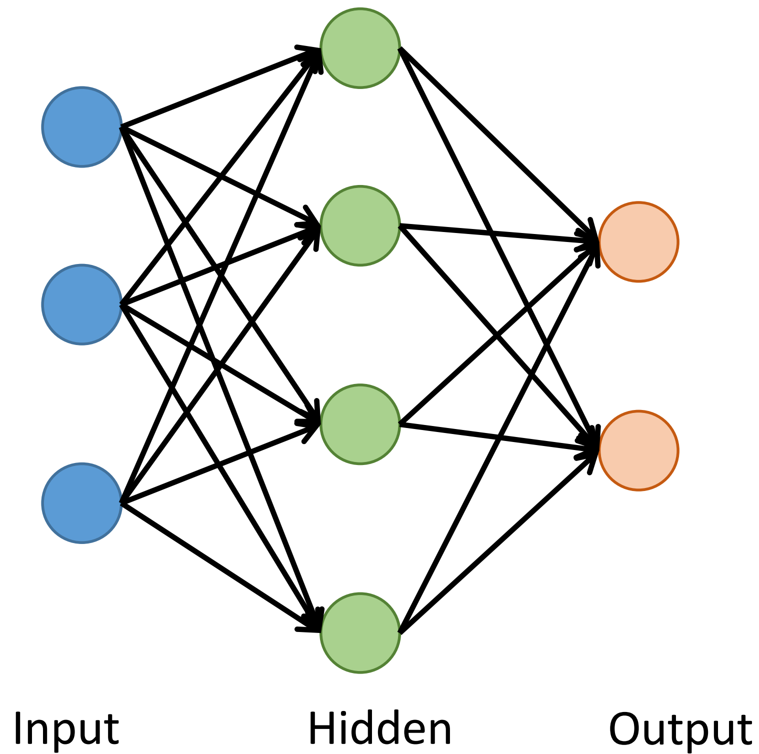
Recent Circuit Trends

- ❑ Complex Architectures
  - ❑ In memory computing

Complex Architectures

- Where to use *Intelligent Machines* ?
  - Self Driving Cars
  - Data (Image/Video) Filtering (i.e. pattern matching)
  - Robotics
  - Medical (diagnostics, imaging etc.)
- Good for Technology:
  - It requires huge computational resources (unlike IOT…)
  - People (at least some) will be ready to pay more for new powerful hardware and this can support the development of more advanced techs *("virtuous circle")*
- Useful for HEP?
  - Hopefully finally something can be done for new forms of event filtering, tracking and pattern recognition

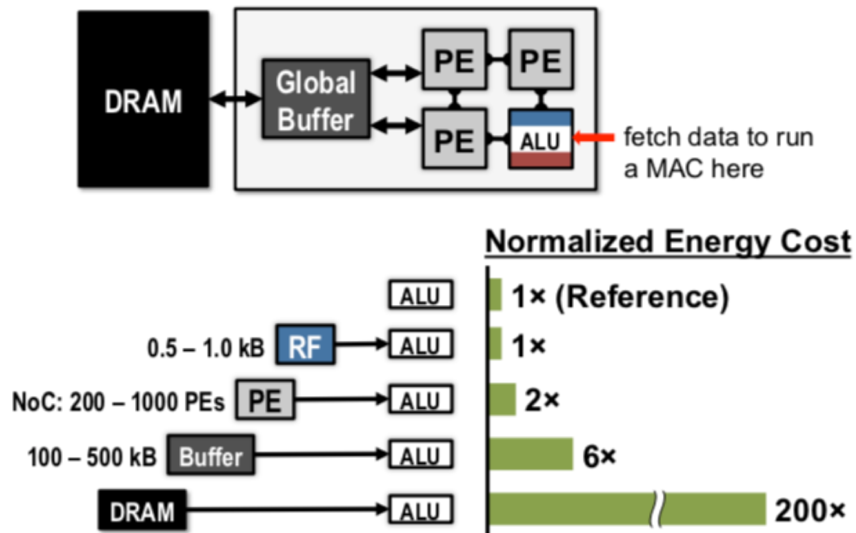
NN basics



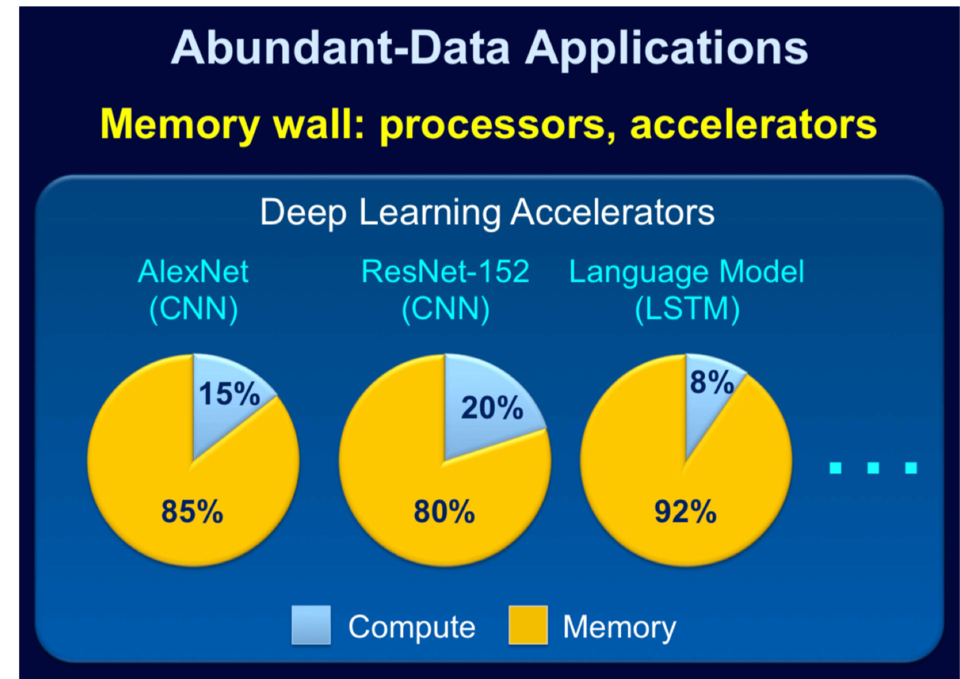
$$o_i(t) = \sum_k w_{ki}\, o_k(t)$$

$$\sigma(o_i) \begin{cases} 1 \; if \; o_i > thr \\ 0 \; if \; o_i < thr \end{cases}$$

Memory access cost

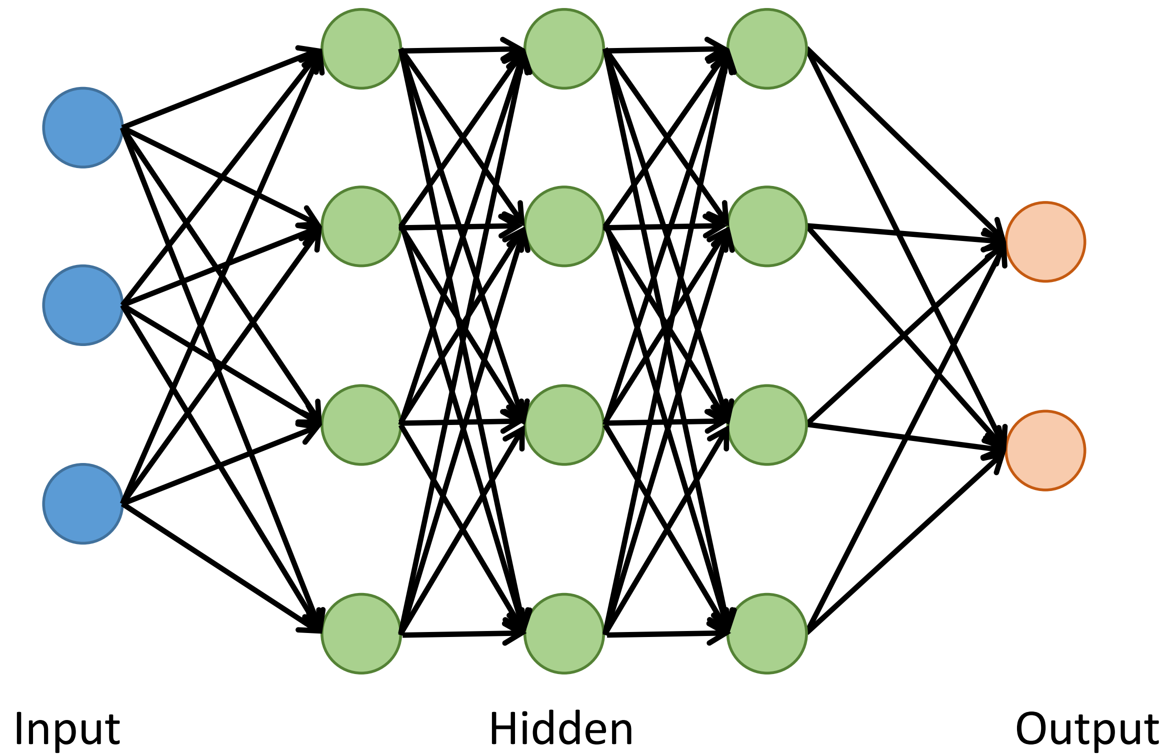


from V. Sze et al.: Efficient Processing of Deep Neural Networks: A Tutorial and Survey, ArXiv:1703.09039v2



from S. Mitra: Abundant Data Computing

DNN basics



$$o_i(t) = \sum_k w_{ki}\, o_k(t)$$

$$\sigma(o_i) \begin{cases} 1 \; if \; o_i > thr \\ 0 \; if \; o_i < thr \end{cases}$$

Complex Architectures

- Keywords: *Neural Networks, Machine Learning, Neuromorphic Computing, Convolutional NN, Deep Learning, etc. etc.*
- Impressive hardware @ ISSCC:
  - 7.1 - Samsung - 11.5 TOPS, 1024 MACs, 8 nm, for mobile
  - 7.2 – Toshiba, 20.5 TOPS, + 2 ARM cores + 4 DSPs, 16 nm, for automotive
  - 7.3 – Univ. Michigan, 0.88 TOPS, 240 mW, 28 nm
  - 7.5 – Tsinghua Univ., 14.9 TOPS, 330 mW, 65 nm

A 2.1TFLOPS/W Mobile Deep RL Accelerator with Transposable PE Array and Experience Compression
KAIST, Korea

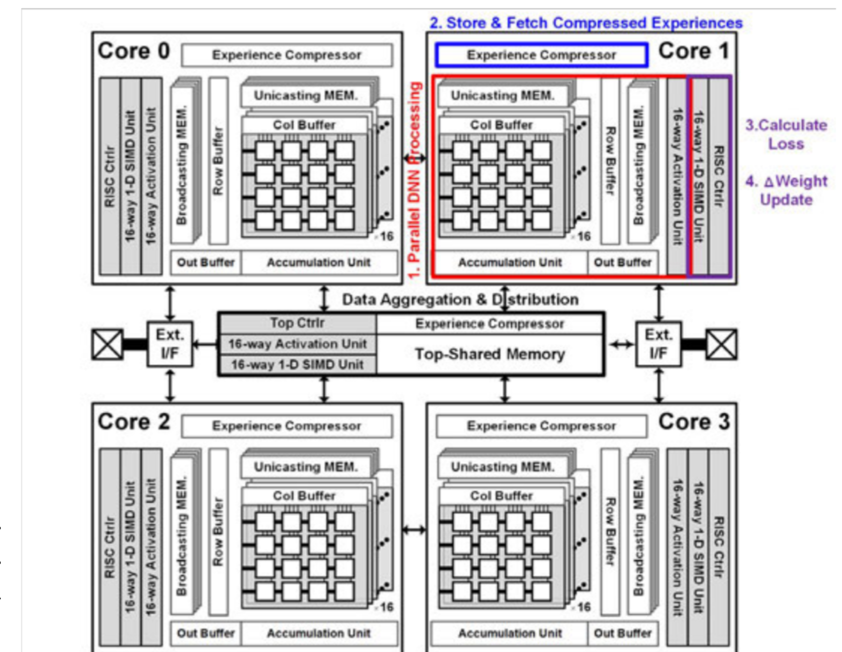
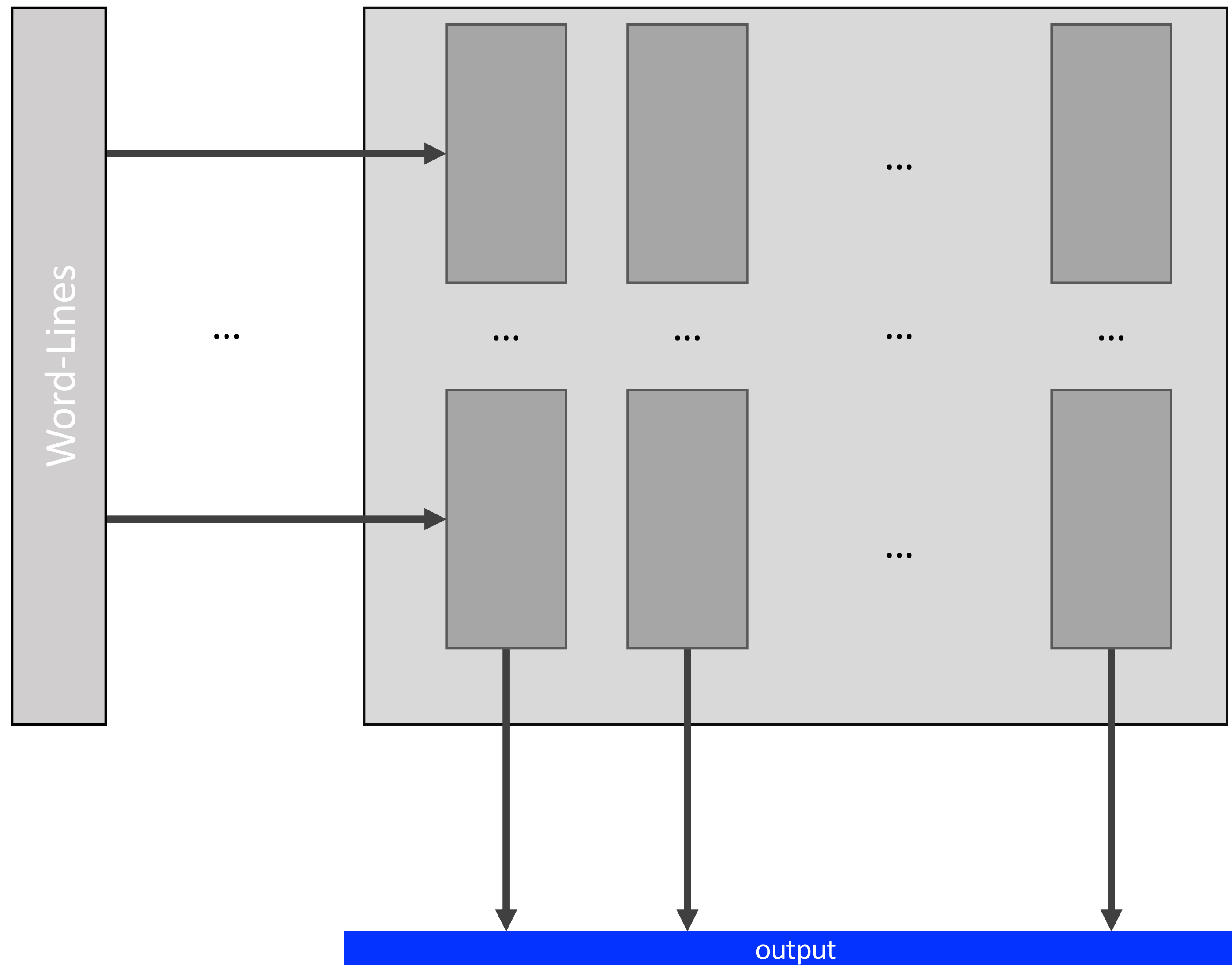
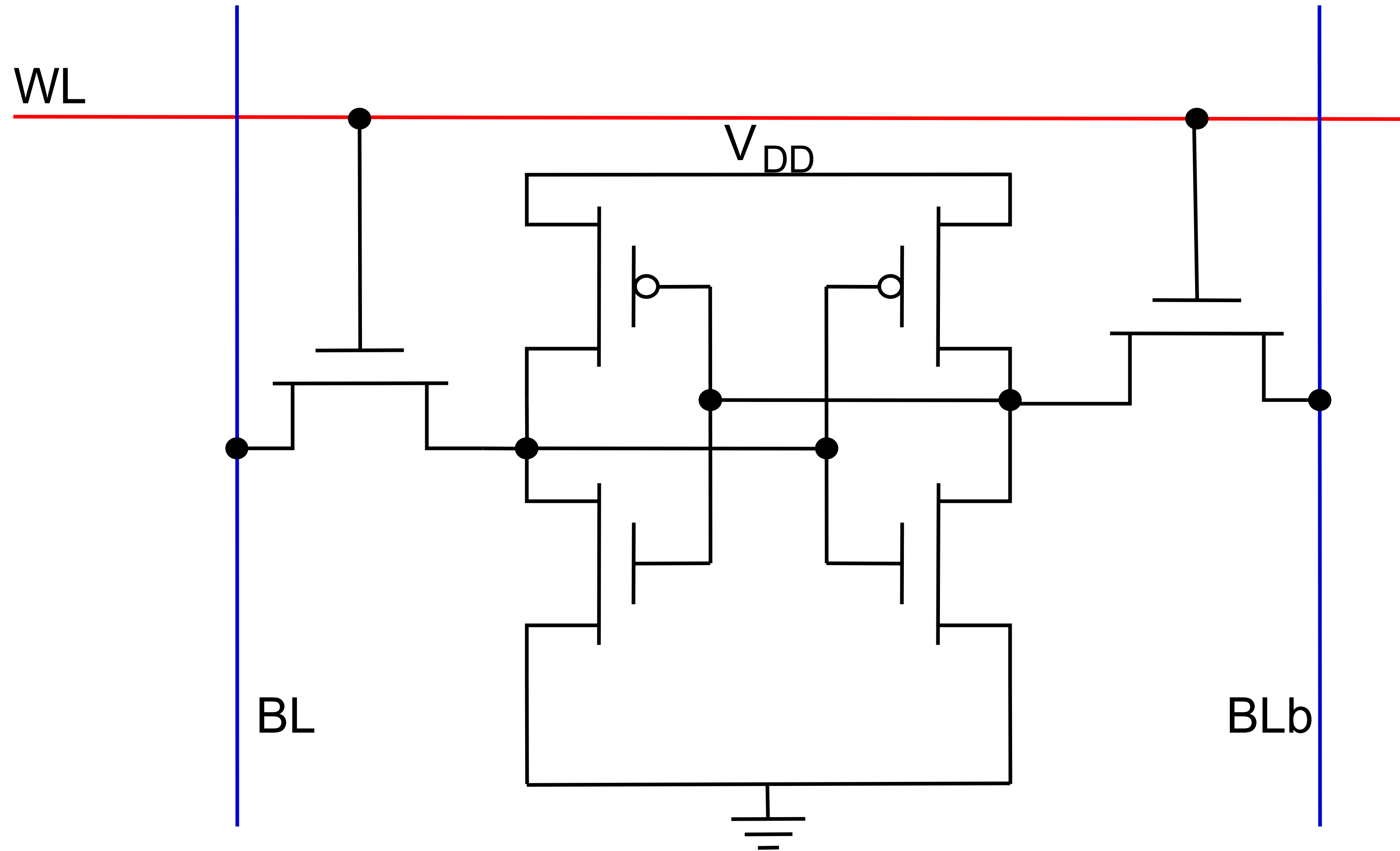


Figure 7.4.2: Overall architecture.

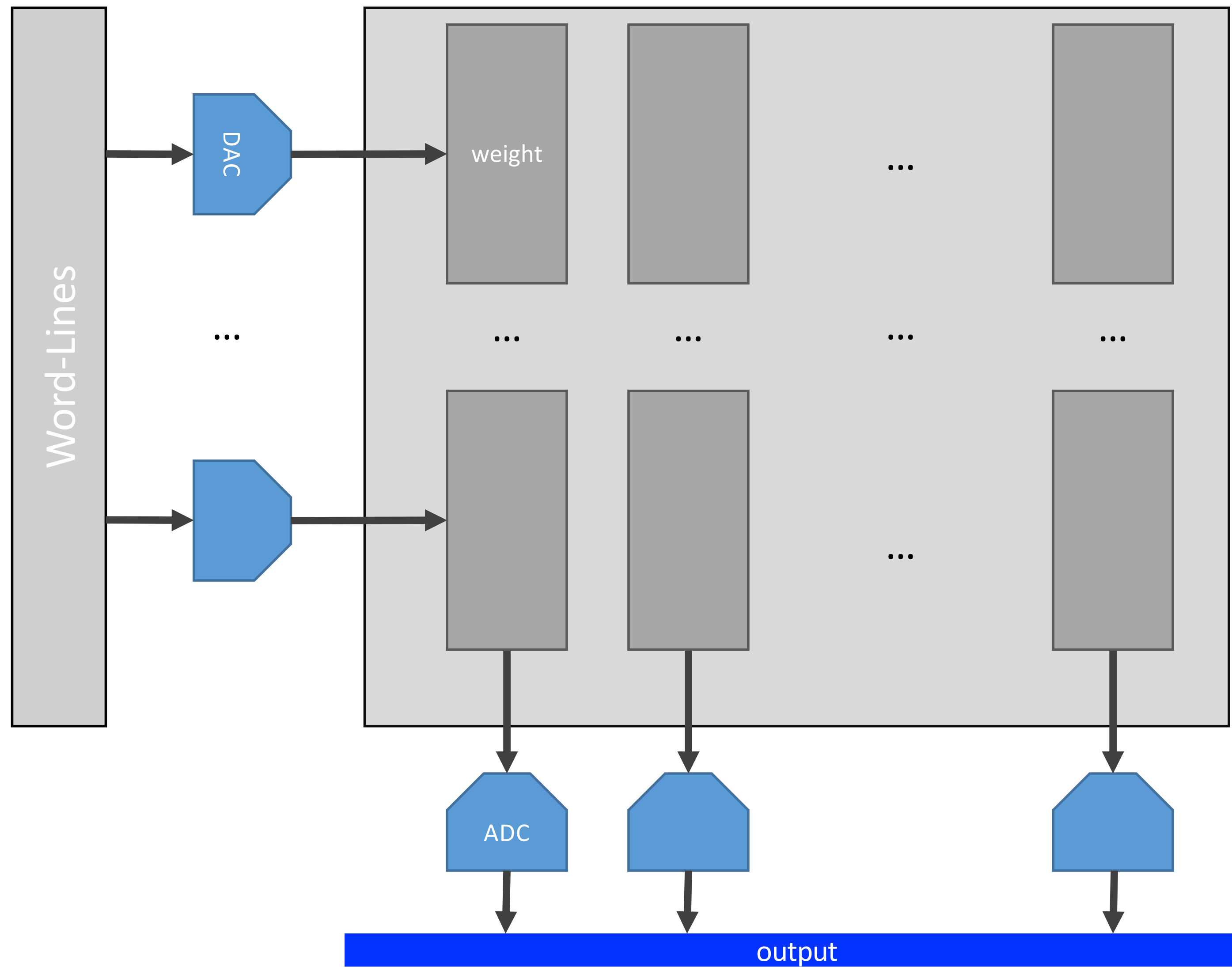
NN-like computation in memory (0)



NN-like computation in memory (1)



NN-like computation in memory (2)



NN-like computation in memory (3)

V_{WL}

V_{DD}

I_{BLb}

1

0

I_{BL}

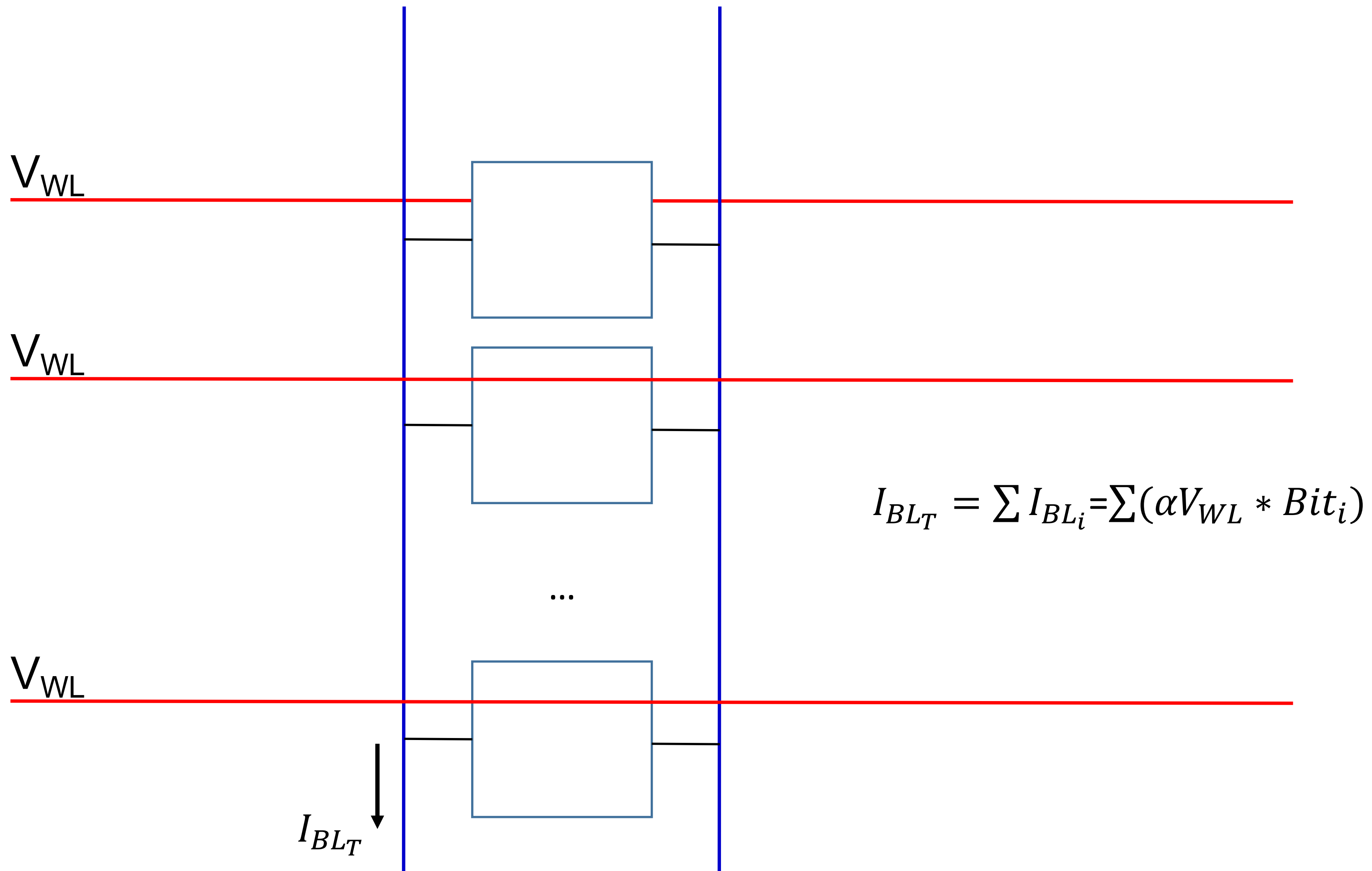
I_{BL}

$$I_{BL} \propto f(V_{WL}, bit)$$

BL

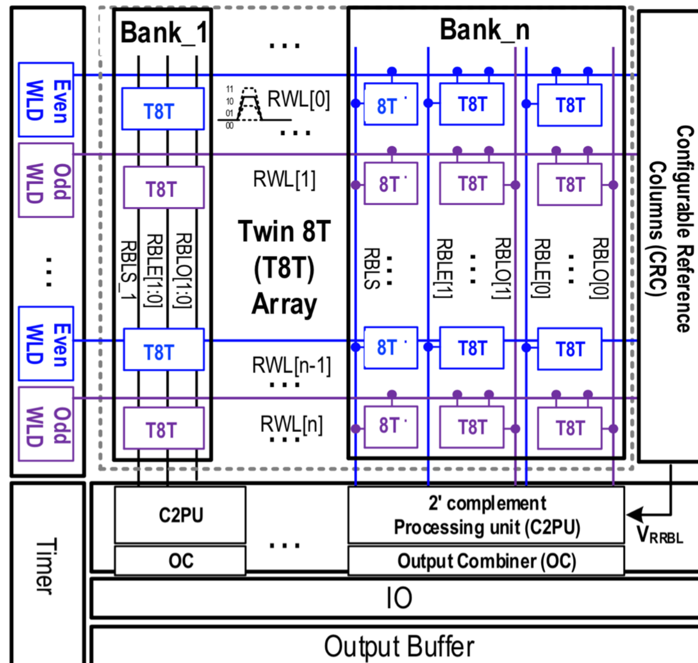
BLb

NN-like computation in memory (4)



Computing in Memory for NN

A Twin-8T SRAM Computation-In-Memory Macro for Multiple-bit CNN-Based Machine Learning, Xin Si et al., ISSCC 2019



Input precision(bit)	1	2	4
Weight precision(bit)	2	5	5
Output precision(bit)	3	5	7
Inference time(ns)	3.2	5.0	10.2
Macro Energy(pJ)	2.5	4.8	9.8
Macro Energy efficiency (TOPS/W)	72.1	37.5	18.37
Measured accuracy MNIST	99.02%	99.18%	99.52%
Measured accuracy CIFAR10	85.56%	90.2%	90.42%