Searching for the Milli-Volt Switch

Eli Yablonovitch,
Winton Inaugural Symposium on Energy Efficiency
Cambridge, United Kingdom
Oct. 1, 2012

Contra Costa-UC Berkeley-MIT-LATTC-Stanford-Tuskegee
In 2010, data centers accounted for ~1.3% of all electricity use worldwide, ~2% of all electricity use in the U.S.
Vision for 2020: Swarms of Electronics:

Infrastructural core

Driver for More of Moore’s Law

Mobile access

Driver for More Than Moore’s Law

Sensory swarm (trillions of devices)

Body Centered Networks (Medicine)
Power Usage Rising Faster Than Past Trend

• Because power consumption $\propto V_{dd}^2$ and $V_{dd}$ (operation voltage) scaling has slowed after 0.13μm node.

<table>
<thead>
<tr>
<th>Technology Node</th>
<th>0.25 μm</th>
<th>0.18 μm</th>
<th>0.13 μm</th>
<th>90 nm</th>
<th>65 nm</th>
<th>45 nm</th>
<th>32 nm</th>
<th>22 nm</th>
<th>16 nm</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V_{dd}$</td>
<td>2.5 V</td>
<td>1.8 V</td>
<td>1.3 V</td>
<td>1.2 V</td>
<td>1.1 V</td>
<td>1.0 V</td>
<td>0.9 V</td>
<td>0.8 V</td>
<td>0.7 V</td>
</tr>
</tbody>
</table>

High Performance ITRS Roadmap
What is the energy cost of reading out your flash memory?
Read the current going through a resistor, in the presence of noise:

\[(\Delta i)^2 = 2q \, i \times \Delta f \ldots \ldots \ldots \text{Shot Noise}\]

\[(\Delta i)^2 = \frac{4kT}{R} \times \Delta f \ldots \ldots \ldots \text{Johnson Noise}\]

Required voltage \( V = iR >> 2kT / q \sim 50\text{mVolts} \)

\[
\text{Signal – to – Noise Ratio} = \frac{i}{\sqrt{2q \, i \Delta f}} = \sqrt{\frac{i}{2q \, \Delta f}}
\]

\( i > 2q \times \Delta f \)

Required power \( iV > 2q \, \Delta f \times \frac{2kT}{q} = 4kT \times \Delta f \)

With a safety margin:

Energy Consumed \sim 40\, kT\, \text{per bit processed}
Units:

~40kT/bit of information

0.16 atto-Joules/bit of information

0.16 nano-Watts/Gbit/second

This is about $10^6$ times less energy than we are using today!
What will be the energy cost, per bit processed?

1. Logic  
   energy cost \(~40kT\) per bit processed

2. Storage  
   energy cost \(~40kT\) per bit processed

3. Communications  
   currently \ (>100,000kT\) per bit processed

.
There are many types of memory possible:
1. Flash
2. SRAM
3. Dram
4. Magnetic Spin
5. Nano-Electro-Chemical Cells
6. Nano-Electro-Mechanical NEMS
7. Memristor
8. Chalcogenide glass (phase change)
9. Carbon Nanotubes

Similarly, there are many ways to do logic.

But there are not many ways to communicate:

1. Microwaves (electrical)
2. Optical
Moore
You?
What is the energy cost for electrical communication?

\[ V_{\text{noise}}^2 = 4kT R \Delta f \]

\[ \frac{V_{\text{noise}}^2}{R} = 4kT \Delta f \]

Signal Energy \[ \geq \] Noise Power per bit \[ = 4kT \] per bit

All information processing costs \[ \sim 40kT \] per bit.

(for good Signal-to-Noise Ratio)

Great!

So what’s the problem?
The natural voltage range for wired communication is rather low:

\[
V_{\text{noise}}^2 = 4kT R \Delta f
\]

\[
V_{\text{noise}}^2 = 4kT R \frac{1}{RC}
\]

\[
V_{\text{noise}}^2 = 4kT \times \frac{1}{C}
\]

\[
V_{\text{noise}}^2 = \frac{4kT}{q} \times \frac{q}{C}
\]

\[
V_{\text{noise}} = \sqrt{\frac{4kT/q \times q/C}{100 \text{mVolts} \times 10 \mu\text{Volts}}}
\]

\[
V \approx 1 \text{ mVolt}
\]

The wire wants 1000 electrons at 1 mVolt each. (to fulfill the signal-to-noise requirement >1eV of energy)

The natural voltage range for a thermally activated switch like transistors is >>kT/q, eg. ~ 40kT/q or about ~1 Volt

Voltage Matching Crisis at the nano-scale!

If you ignore it the penalty will be

\[(1 \text{Volt}/1 \text{mVolt})^2 = 10^6\]

The thermally activated device wants at least one electron at ~1 Volt.
The New Switch has to Satisfy Three Specifications:

1. Steepness (or sensitivity)
   switches with only a few milli-volts
   60mV/decade ⇒ 1mV/decade

2. On/Off ratio.          10^6 : 1

3. Current Density or Conductance Density
   (for miniaturization)
   old spec at 1 Volt:     1 mAmp/micron
   our spec:               1 milli-mho/micron
A low-voltage technology, or an impedance matching device, needs to be invented/discovered at the Nano-scale:

- Transistor amplifier with steeper sub-threshold slope
- TFET's
- Negative Capacitance Gates

-- VO$_2$ metal-insulator transition
An amplifying transistor as a voltage matching device:

- Small voltage in
- Large voltage out

Amplification of weak signals has an energy cost!
Amplification of weak signals has a speed penalty!

\[
\ln(I) \quad \text{steeper sub-threshold slope}
\]

\[V_g \]
The Zener Diode:

$E_{Fv}$

$E_{Fc}$

Bias Voltage

Current

Diffusion current
The Esaki Diode:
The Backward Diode as a Switch:

The Backward Diode: These have been routinely made in Ge homo-junctions, since the 1960's.

[Diagram of energy levels and bias voltage-current relationship]
The sub-threshold slope for tunneling depends on the steepness of the band-edges.

The Backward Diode as a Switch:
2 Ways to Obtain Steepness:

- Modulate the Tunneling Barrier:

- Density of States Switch
  The sub-threshold slope for tunneling depends on the steepness of the band-edges:
$I \propto V_{OL}^{3/2}$

$I \propto \sqrt{V_{OL}}$

$I \propto \text{Constant}$
Type III band alignment

Idealized structure

- $E_F$ (nm)
- $E$ (eV)
- $n$-GaSb
- $p$-Al$_{0.3}$Ga$_{0.7}$Sb
- InAs
- AlSb
- GaSb
- n-GaSb

$E_C$, $E_V$
\( \gamma = 2.34 \text{ meV} \)

\( E_Z = 50 \text{ meV} \)

\( m^* = 0.1 \)

\( T_{\text{device}} = 2.16\% \)

\( L_X = 32 \text{ nm} \)

\( L_Z = 8.672 \text{ nm} \)
Switching Principle:

Conduction band

Valence band
Switching Principle:

Conduction band

Valence band
What could go wrong?

1. quantum-mechanical level repulsion:

levels never line up!
Evolution of the Tunnel Switch 2010-2012:

Homojunction Backward Diode

2d-2d pn Hetero-junction

Heterojunction

BiLayer Switch
The Bi-Layer pn-junction or the Bi-Layer Tunneling Field Effect Transistor
Drop-In Technology:
drop in to existing FinFET production line.

Performance depends on precision of Work Function control, and thickness precision of the dual channel layer.

Instead of n-channel and p-channel, make n and p on opposite sides of the same fin.

Performance depends on precision of Work Function control, and thickness precision of the dual channel layer.
Due to Capacitive Voltage divider, the gate efficiency is poor, ~15% for a silicon fin.
Quantum Capacitance impels small Effective Mass:

\[
\frac{N}{A} = \frac{\varepsilon_r \varepsilon_o}{qd} V
\]

\[
\frac{N}{A} = \frac{m^*}{\pi \hbar^2} qV
\]

Respectable gate efficiency requires:

\[
\frac{qm^*}{\pi \hbar^2} < \frac{\varepsilon_r \varepsilon_o}{qd^2}
\]

Respectable gate efficiency requires \( m^* < 0.1 m_o \)
Try InAs, effective mass is lower, density of states is lower, and $C_{\text{quantum}}$ is lower.

The lower n-channel carrier density makes it easier to swing the energy level.

Lower effective mass—easier tunneling.

We need $m_{\text{eff}} < 0.1$
InAs Band Diagram at Turn-on

No Doping!

For Lab experiments: use Electric Field Induced pn junction. For production use: Work Function induce pn junction.
InAs Asymmetric 15 nm Body

Ground the Valence Band Channel for better gate control on the n-channel

1 nm oxides
100 nm long channel length
Ground heavy hole band (heavy in the tunneling direction)
2D Nanomembranes for Novel Tunneling (A. Javey)

Layered Semiconductors

- TMDC (e.g. WSe$_2$, MoS$_2$)
- III-VI (e.g. GaSe)

III-V on Insulator (XOI)
High Performance InAs XOI n-MOSFETs

Electron Mobility: 1000-7000 cm²/Vs
SS ~ 75 mV/decade
$R_c \sim 80 \ \Omega \mu m$

Kuni Takei, et al, APL, 2011
First demonstration of a diode based on a van der Waals heterojunctions.

Clear rectifying behavior is observed.

Materials Approach:

Van der Waals 2D membranes:
- Removes lattice mismatch constraints
- Mix and Match: A wide range of heterojunctions is available
- Atomically abrupt interfaces
Scientific Research

many analog applications:

Sensors
Photodetectors
radio mixers

Backward Diode

Scientific Research:

MATERIALS:

InAs/GaSb
InAs

Fixed Layer Thickness:

MoS$_2$, WSe$_2$

Layered Chalcogenides

TFET

Roadmap:
What keeps me up at night:

Band edges are simply not sharper than $\sim kT/3q$, allowing us to pick up only a factor $\sim 3$ improvement.

What doesn’t worry me:

Manufacturability and Yield. If we can demonstrate individual high-performing devices, then a large international effort will become directed toward these problems.
A low-voltage technology, or an impedance matching device, needs to be invented/discovered at the Nano-scale:

- Transistor amplifier with steeper sub-threshold slope
- TFET's
- Negative Capacitance Gates

--VO₂ metal-insulator transition

**Cryo-Electronics**

kT/q~q/C

**Transistor Amplifier**

- Photo-diode
- Nano-transformer

**Giant Magneto-Resistance Spintronics**

- Mem's switch

**Cryo-Electronics**

K/T/q~q/C

**Electro-Chemical Switch**