Driving Fully-Adiabatic Logic Circuits Using Custom High-*Q* MEMS Resonators

Venkiteswaran Anantharam
ECE Dept.
venki1@ufl.edu

Maojiao He
ECE Dept.
maojiaoh@ufl.edu

Krishna Natarajan
ECE Dept.
krishnat@ufl.edu

Huikai Xie
ECE Dept.
hkxie@ece.ufl.edu
(352) 846-0441

Michael P. Frank
CISE and ECE Depts.
mpf@cise.ufl.edu
(352) 392-6888

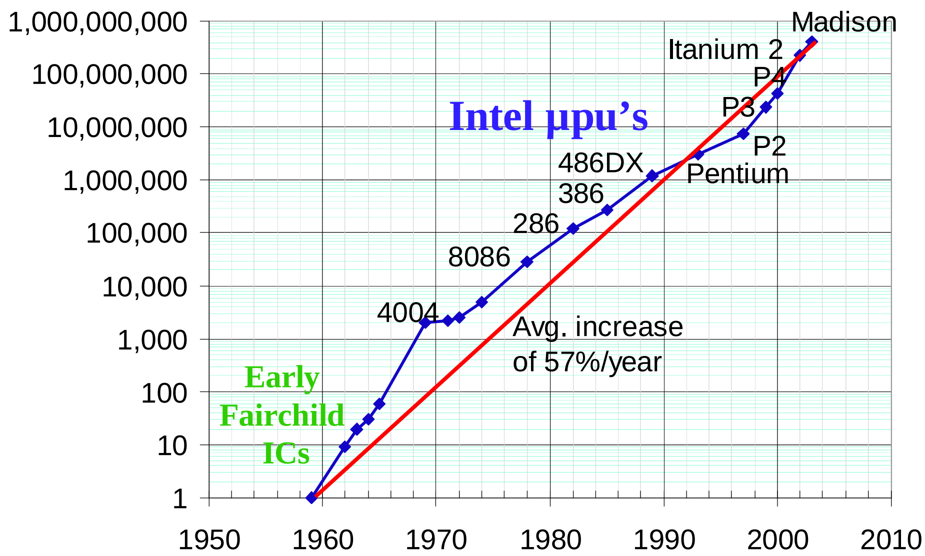
University of Florida, Gainesville, Florida

Presented at:
Workshop on Methodologies in Low-Power Design (MLPD ’04)
2004 Int’l Conf. on Embedded Systems and Applications (ESA ’04)
2004 Int’l Multiconf. in Computer Science & Computer Engineering
Las Vegas, Nevada, June 21-24, 2004

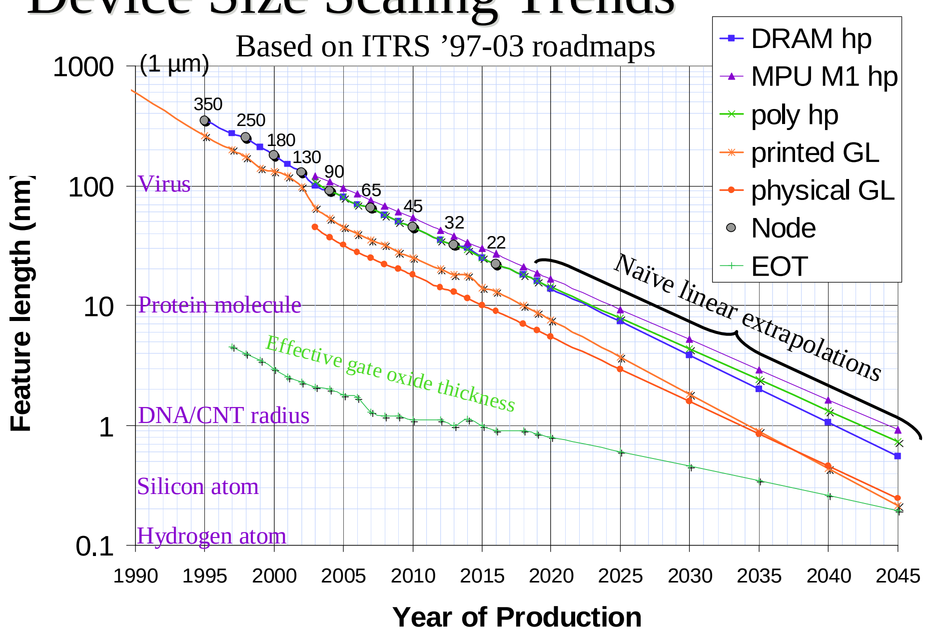
Abstract

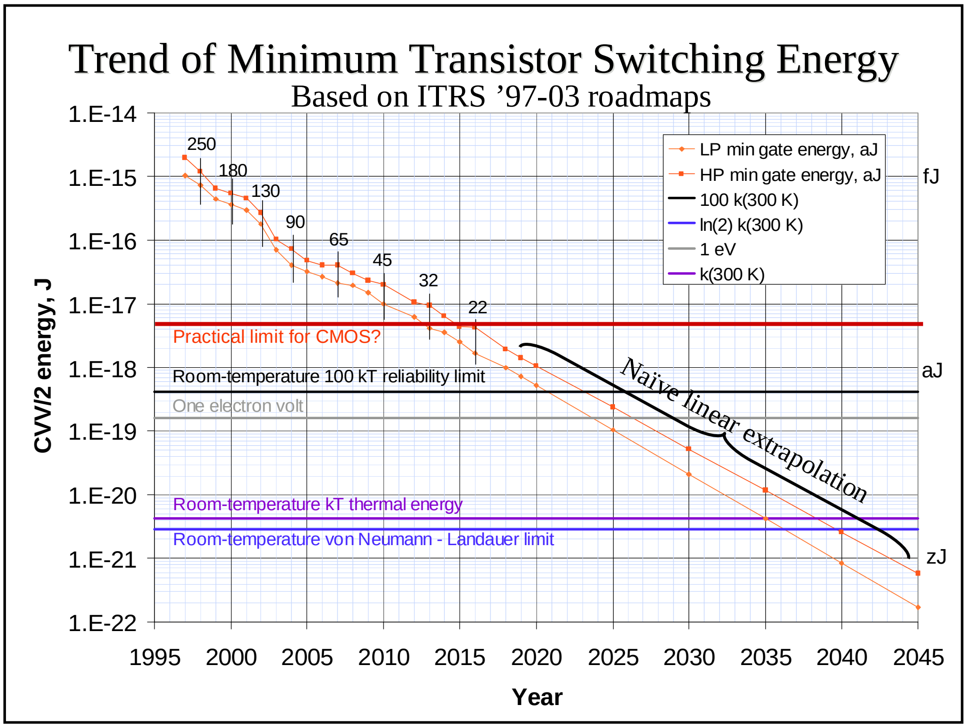
- Adiabatic techniques are absolutely required if we want to keep improving digital power-performance beyond the next decade or so.
  - Due to fundamental limits on the energy dissipation of conventional logic.
- Truly adiabatic (asymptotically non-dissipative) logic must be driven by quasi-trapezoidal (flat-topped) power-clock waveforms.
  - And these must be generated by a high-*Q* (quality) resonant component.
- Fortunately, present-day MEMS resonators can achieve very high *Q* factors, and fairly high (MHz-GHz or so) frequencies.
  - They are beginning to be used commercially today in RF applications.
- Our group is building custom MEMS resonators for driving adiabatic logic, in an integrated CMOS/MEMS process.
  - The resonators are designed to have the right characteristics to do the job.
- We describe some prototype resonator designs, and their important figures of merit (and demerit) that need to be optimized.
  - And we also discuss our specific design strategies for doing so.
- Our simulations in Coventorware & Cadence suggest we might obtain ~10× power-performance boosts vs. standard CMOS.
  - Predictions will be validated with a test chip to tape out July 26.

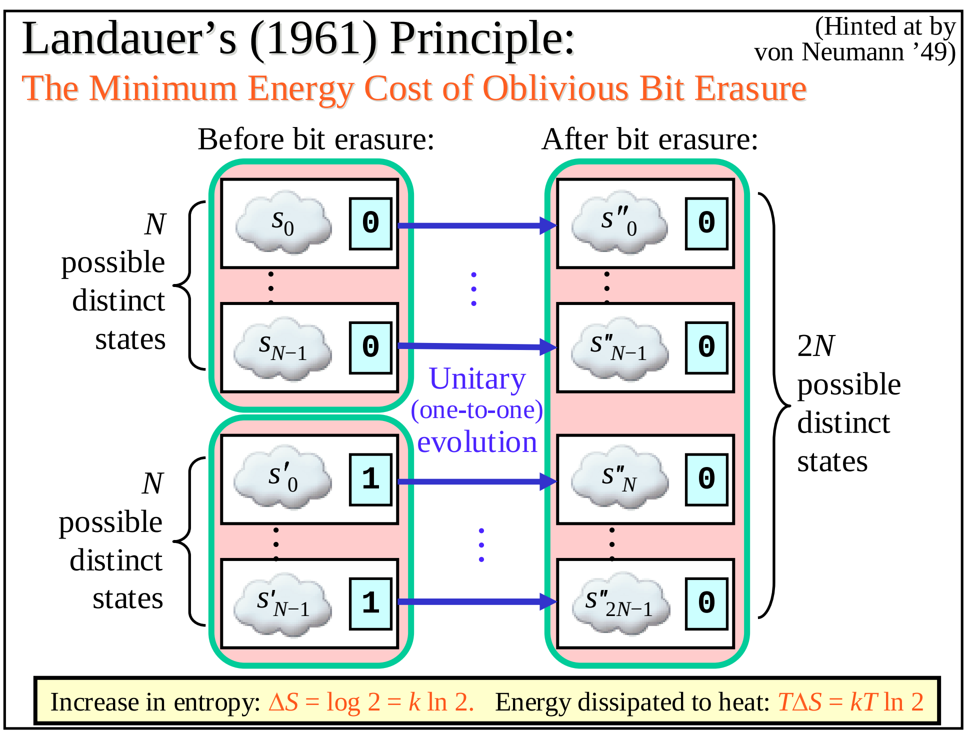
Moore's Law – Devices per IC



Device Size Scaling Trends




Trend of Minimum Transistor Switching Energy
Based on ITRS '97-03 roadmaps
CVV/2 energy, J
1.E-14
1.E-15
1.E-16
1.E-17
1.E-18
1.E-19
1.E-20
1.E-21
1.E-22
1995
2000
2005
2010
2015
2020
2025
2030
2035
2040
2045
Year
250
180
130
90
65
45
32
22
LP min gate energy, aJ
HP min gate energy, aJ
100 k(300 K)
ln(2) k(300 K)
1 eV
k(300 K)
fJ
aJ
zJ
Practical limit for CMOS?
Room-temperature 100 kT reliability limit
One electron volt
Room-temperature kT thermal energy
Room-temperature von Neumann - Landauer limit
Naïve linear extrapolation


Landauer's (1961) Principle:
(Hinted at by von Neumann '49)
The Minimum Energy Cost of Oblivious Bit Erasure
Before bit erasure:
After bit erasure:
N possible distinct states
N possible distinct states
2N possible distinct states
s0 0
sN−1 0
s′0 1
s′N−1 1
s″0 0
s″N−1 0
s″N 0
s″2N−1 0
Unitary (one-to-one) evolution
Increase in entropy: ΔS = log 2 = k ln 2. Energy dissipated to heat: TΔS = kT ln 2

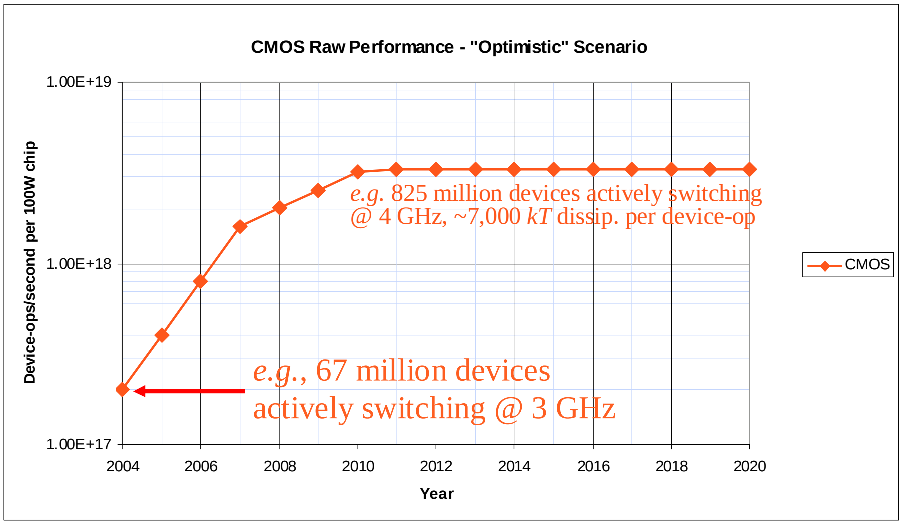
Reliability Bound on Bit Energy

- To reliably store (latch) a bit of data with less than 1 error in N repetitions requires that:
  - In the equilibrium microstate distribution, when latching, the number of accessible microstates interpreted as the correct stored bit value should be N times the number leading to the incorrect bit value.
    - ∴ There needs to be $\Delta E \gtrsim k_B T \ln N$ energy difference between states having correct and incorrect bit values, at time of latching, in a device (near equil.) at temperature T.
      - This follows directly from the Fermi-Dirac distribution!
    - If and when an energy of this same magnitude later gets *dissipated* by the device, this would lead to an characteristic entropy increase of $\Delta S = \log N = k_B \ln N$.
      - And free energy loss of $k_B T_{env} \ln N$, if environment is at T_{env}.

A Fairly Conventional "Optimistic" Technology Scenario for CMOS

- Suppose device lengths are cut in half every 3 years…
  - From 90 nm today down to 22 nm node in 2010 (then stop).
  - Node capacitances, gate delays also decrease accordingly…
- "Technology boosters" such as high-κ dielectrics & novel FET structures (FinFET, surround-gate, *etc.*) keep leakage power manageable, for a little while…
  - However, note the absolute minimum room-T subthreshold slope for FETs will remain 60 mV/decade! (= $(kT/q) \log 10$)
    - Assume this point is also reached by around 2007.
- Voltages then reach a minimum of ~0.5V in 2007.
  - Can't go lower while keeping on/off ratio above 10^8 level!
    - A minimum level chosen so as to keep leakage manageable
- Now, consider what all this implies about future chip performance, given a 100 W maximum power level…
  - Let max raw performance = 100 W / (½CV^2 gate energy)

Not much life left for standard CMOS…



Even if the leakage problem were solved, the ~100 *kT* limit for reliable switching is only another factor of 70 beyond this point!

Reversible Computing

- A *reversible* digital logic operation is:
  - Any operation that performs an invertible (one-to-one) transformation of the device's local digital state.
- Landauer's principle <u>only</u> limits the energy dissipation of ordinary *irreversible* (many-to-one) logic operations.
  - Reversible logic operations can dissipate much less energy,
    - Since they can be implemented in a thermodynamically reversible way.
- In 1973, Charles Bennett (IBM Research) showed how any desired computation can in fact be performed using *only* reversible operations (with basically no bit erasure).
  - This opened up the possibility of a vastly more energy-efficient alternative paradigm for digital computation.
- After 30 years of (sporadic) research, this idea is finally approaching the realm of practical implementability…
  - Making it happen is the goal of the RevComp project at UF.

Adiabatic Circuits

- Reversible logic can be implemented today using fairly ordinary voltage-coded CMOS VLSI circuits.
  - With a few changes to the logic-gate/circuit architecture.
- We avoid dissipating most of the circuit node energy when switching, by transferring charges in a nearly *adiabatic* (lit. “without flow of heat”) fashion.
  - *I.e.*, asymptotically thermodynamically reversible.
    - In the limit, as various low-level technology parameters are scaled.
- There are many designs for purported “adiabatic” circuits in the literature, but most of them contain fatal flaws and are not truly adiabatic.
  - Many past designers are unaware of (or accidentally failed to meet) all the requirements for true thermodynamic reversibility.

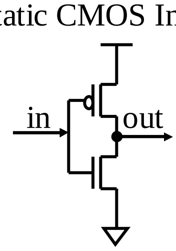
Conventional Logic is Irreversible

Even a simple NOT gate, as traditionally implemented…

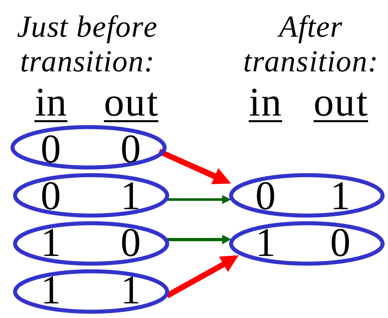
- Logic gate behavior (upon receiving new input):
  - Performs many-to-one transformation of local state!
  - ∴ required to dissipate $\gtrsim kT$, by Landauer principle
  - Incurs $\frac{1}{2}CV^2$ energy dissipation in 2 out of 4 cases.

Example:

Static CMOS Inverter:



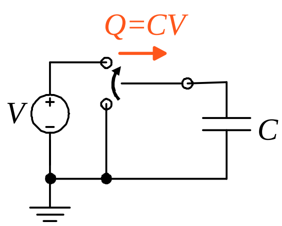
Transformation of local state:



Conventional *vs.* Adiabatic Charging

For charging a capacitive load C through a voltage swing V

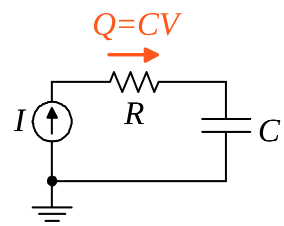
- Conventional charging:
  - Constant voltage source



  - Energy dissipated:

$$E_{\text{diss}} = \tfrac{1}{2} C V^2$$

- Ideal adiabatic charging:
  - Constant current source



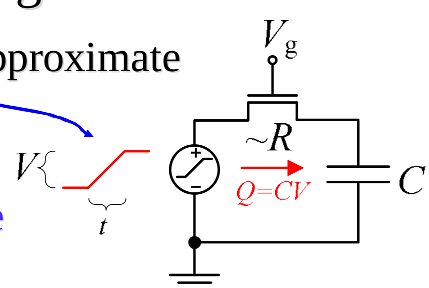
  - Energy dissipated:

$$E_{\text{diss}} = I^2 R t = \frac{Q^2 R}{t} = C V^2 \frac{RC}{t}$$

Note: Adiabatic beats conventional by advantage factor $A = t/2RC$.

Adiabatic Switching with MOSFETs

- Use a voltage ramp to approximate an ideal current source.
- Switch *conditionally*, if MOSFET gate voltage $V_g > V + V_T$ during ramp.



- Can discharge the load later using a similar ramp.
  - Either through the same path, or a different path.

$$t \gg RC \Rightarrow E_{\text{diss}} \to C V^2 \frac{RC}{t}$$

$$t \ll RC \Rightarrow E_{\text{diss}} \to \tfrac{1}{2} C V^2$$

Exact formula:

$$E_{\text{diss}} = s\left[1 + s\left(e^{-1/s} - 1\right)\right] C V^2$$

given *speed fraction* $s :\equiv RC/t$

Athas '96, Tzartzanis '98

Requirements for True Adiabatic Logic

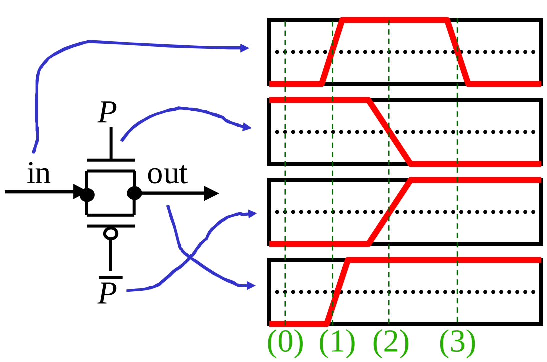
(See paper by Frank from MLPD ’03)

- Avoid passing current through diodes.
  - Crossing the “diode drop” leads to irreducible dissipation.
- Follow a “dry switching” discipline (in the relay lingo):
  - Never turn on a transistor when $V_{DS} \neq 0$.
  - Never turn off a transistor when $I_{DS} \neq 0$. ← Important but often neglected!
- Together these rules imply:
  - The logic design must be logically reversible
    - There is no way to erase information under these rules!
  - Transitions must be driven by a quasi-trapezoidal waveform
    - It must be generated resonantly, with high Q
- Of course, leakage power must also be kept manageable.
  - Because of this, the optimal design point will not necessarily use the smallest devices that can ever be manufactured!
    - Since the smallest devices may have insoluble problems with leakage.

A Simple Reversible CMOS Latch

- Uses a single standard CMOS *transmission gate* (T-gate).
- Sequence of operation:

  (0) input level initially tied to latch ‘contents’ (output);
  (1) input changes gradually → output follows closely;
  (2) latch closes, charge is stored dynamically (node floats);
  (3) afterwards, the input signal can be removed.



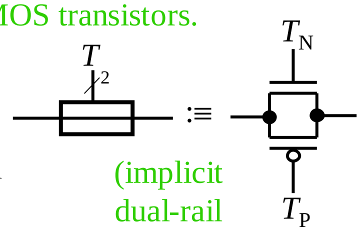
Before input:		*Input arrived:*		*Input removed:*	
in	out	in	out	in	out
0	0	0	0	0	0
		1	1	0	1

- Later, we can reversibly “unlatch” the data with an exactly time-reversed sequence of steps.

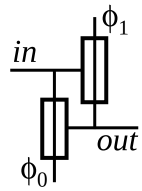
2LAL: 2-level Adiabatic Logic

A pipelined fully-adiabatic logic invented at UF (Spring 2000), implementable using ordinary CMOS transistors.

- Use simplified T-gate symbol:

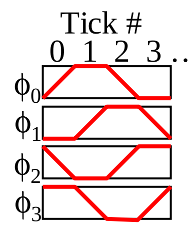


- Basic buffer element:
  - cross-coupled T-gates:
    - need 8 transistors to buffer 1 dual-rail signal

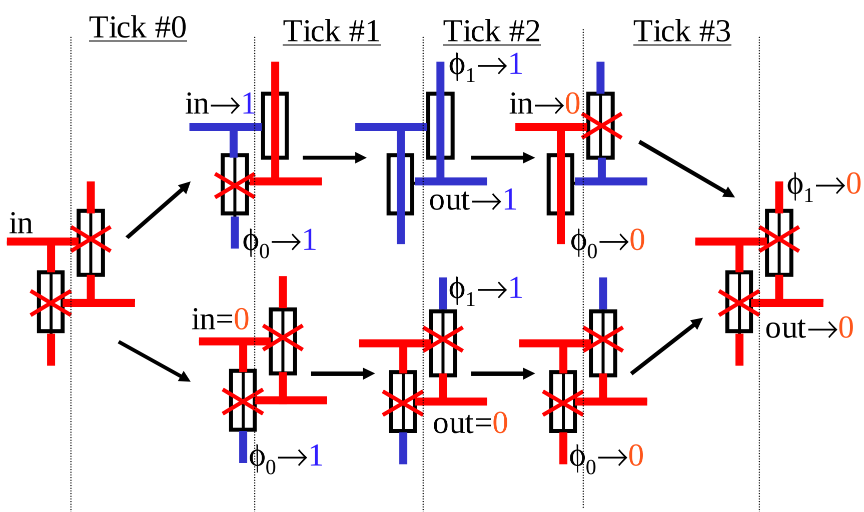


(implicit dual-rail encoding everywhere)

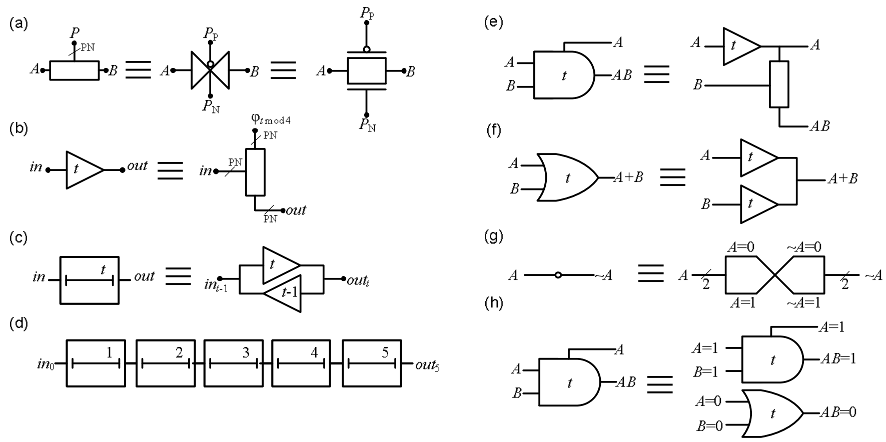
- Only 4 timing signals $\phi_{0\text{-}3}$ are needed. Only 4 ticks per cycle:
  - ϕ_i rises during ticks $t \equiv i \pmod 4$
  - ϕ_i falls during ticks $t \equiv i+2 \pmod 4$



2LAL Cycle of Operation



A Schematic Notation for 2LAL

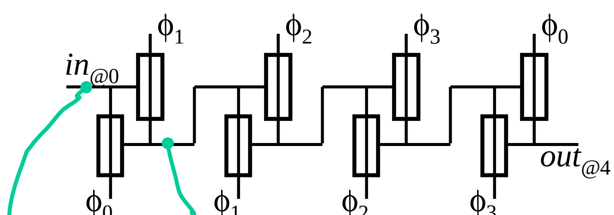


2LAL Shift Register Structure

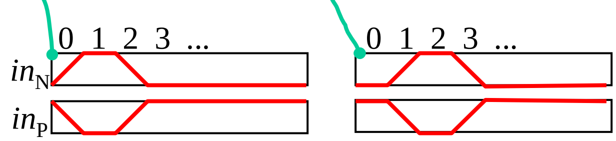
Animation:

2lal.swf

- 1-tick delay per logic stage:

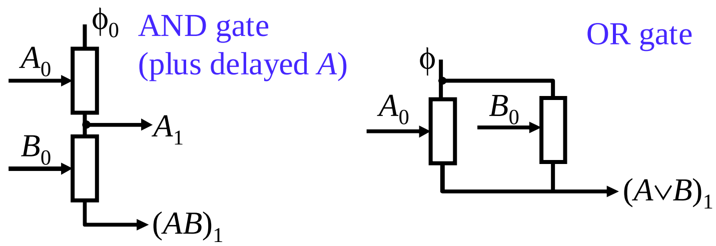


- Logic pulse timing and signal propagation:

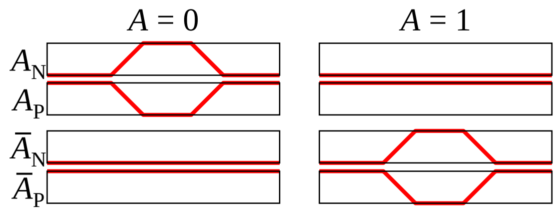


More Complex Logic Functions

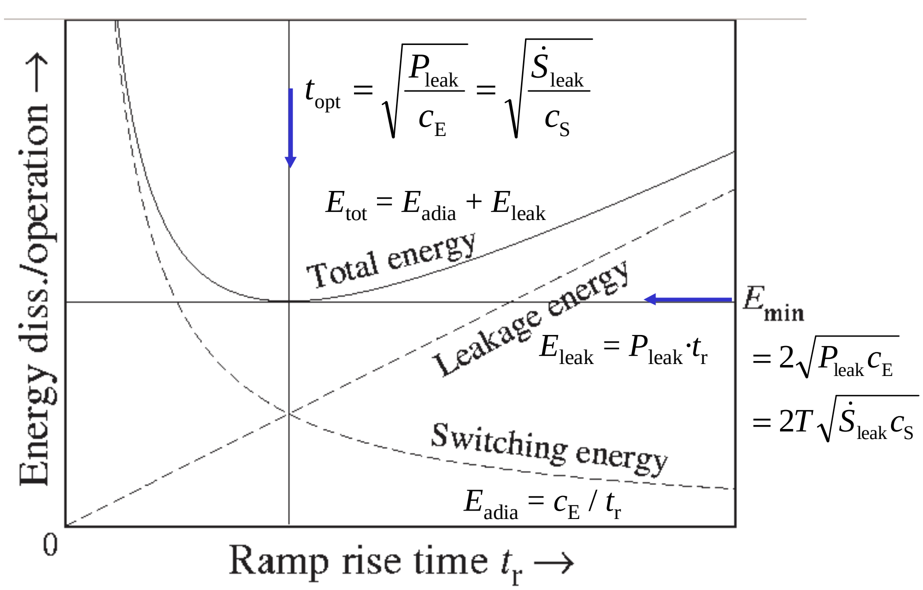
- Non-inverting multi-input Boolean functions:



- One way to do inverting functions in pipelined logic is to use a quad-rail logic encoding:
  - To invert, just swap the rails!
    - Zero-transistor "inverters."



Minimum Losses w. Leakage



UF CONFIDENTIAL – PATENT PENDING

MEMS Resonator Concept

A potential approach for efficiently driving adiabatic logic transitions

The Power Supply Problem

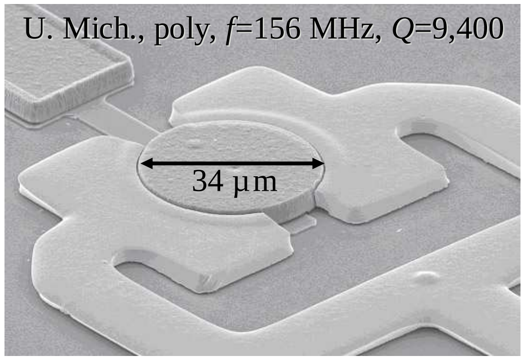
- In adiabatics, the factor of reduction in energy dissipated per switching event is limited to (at most) the *Q* factor of the clock/power supply.

$$Q_{\text{overall}} = (Q_{\text{logic}}^{-1} + Q_{\text{supply}}^{-1})^{-1}$$

- Electronic resonator designs typically have low *Q* factors, due to considerations such as:
  - Energy overhead of switching a clamping power MOSFET to limit the voltage swing of a sinusoidal *LC* oscillator.
  - Low coil count, substrate coupling in integrated inductors.
  - Unfavorable scaling of inductor *Q* with frequency.
- Our proposed solution:
  - Use electromechanical resonators instead!

MEMS (& NEMS) Resonators

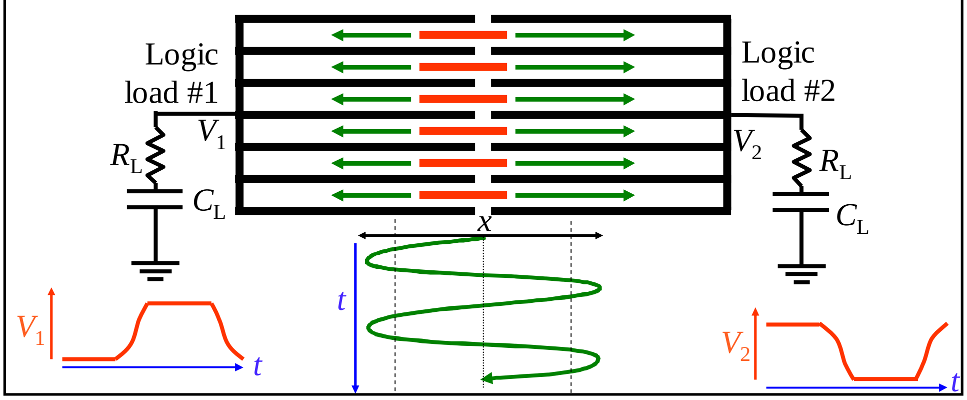
- State of the art of technology demonstrated in lab:
  - Frequencies up to the 100s of MHz, even GHz
  - *Q*'s >10,000 in vacuum, several thousand even in air!
- An important emerging technology being explored for use in RF filters, *etc*., in communications SoCs, *e.g*. for cellphones.



UF CONFIDENTIAL – PATENT PENDING

Original Concept

- Imagine a set of charged plates whose horizontal position oscillates between two sets of interdigitated fixed plates.
  - Structure forms a variable capacitor and voltage divider with the load.
- Capacitance changes substantially only when crossing border.
  - Produces nearly flat-topped (quasi-trapezoidal) output waveforms.
  - The two output signals have opposite phases (2 of the 4 φ's in 2LAL)

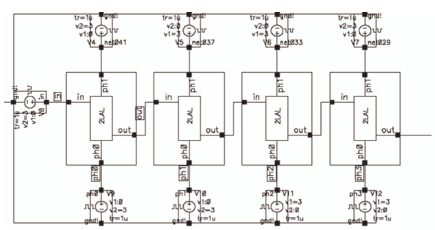


UF CONFIDENTIAL – PATENT PENDING

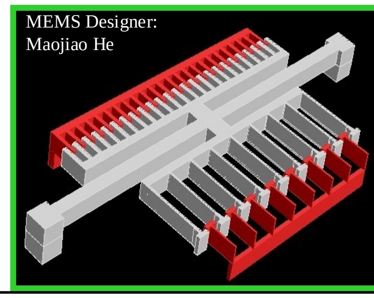
MEMS Resonant Power Supply for Ultra-Low-Power Adiabatic Circuits

A.k.a. The "AdiaMEMS" Project

- Part of CISE's Reversible & Quantum Computing group
  - Collab. with Huikai Xie (MEMS, ECE dept.) →
- **Goal:** Demonstrate orders-of-magnitude improvement in power-performance efficiency of digital CMOS circuits.
  - Based on reversible logic in adiabatic circuits powered by high-quality custom microelectromechanical resonators.
- **Funding:** $40K seed grant from SRC's Cross-Disciplinary Semiconductor Research (CSR) Program



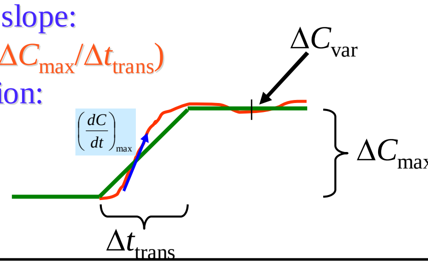
VLSI designer: Krishna Natarajan



UF CONFIDENTIAL – PATENT PENDING

Key Characteristics of Resonator

- **Goal:** Produce a near-ideal trapezoidal output voltage waveform resonantly, with high Q.
- To be optimized with logic: Resonant frequency f.
- Key resonator figures of merit:
  - Effective quality factor: $Q_{eff} = E_{trans}/E_{diss}$.
  - Area efficiency: $E_A = E_{trans}/A$.
- Key resonator figures of demerit:
  - Maximum relative transition slope:
    $s_{max} = (dC/dt)_{max} / (\Delta C_{max}/\Delta t_{trans})$
  - Fractional capacitance variation:
    $v_C = \Delta C_{var} / \Delta C_{max}$



UF CONFIDENTIAL – PATENT PENDING

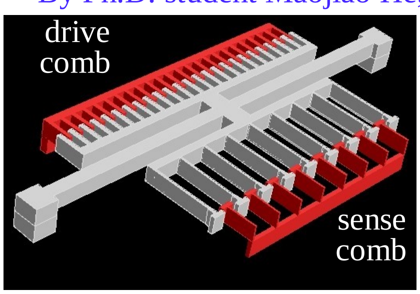
First MEMS Technology Tried

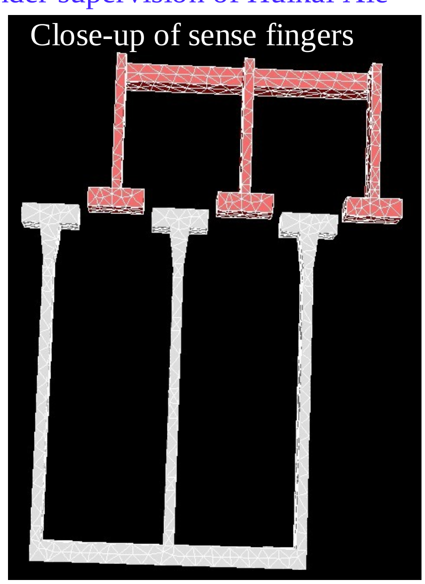
- MEMS process donated by Robert Bosch corp.
- It is a thin-film technology
  - We have since moved to a multi-layer, bulk single-crystal process which can be expected to do better.
- Integrated CMOS/MEMS devices will eventually be available in this process.
  - However our initial design was dual-die
    - CMOS side was not mature yet in this process
- Minimum etched structure width: $\lambda = 0.5$ μm
- Minimum etched gap size: $d = 0.1$ μm

UF CONFIDENTIAL – PATENT PENDING

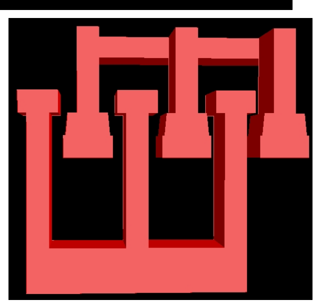
Some Early Resonator Designs

By Ph.D. student Maojiao He, under supervision of Huikai Xie



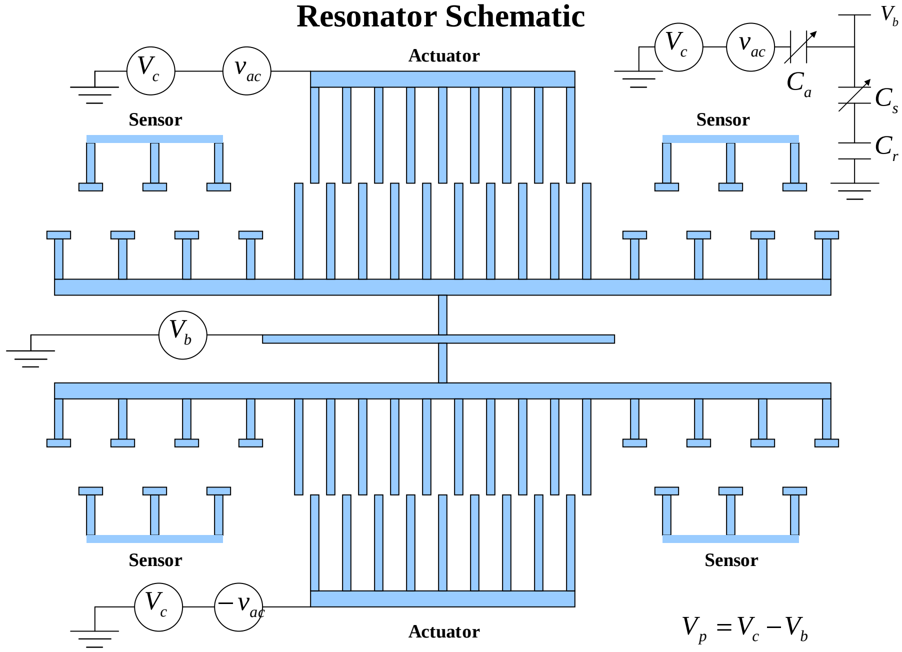


Another finger design



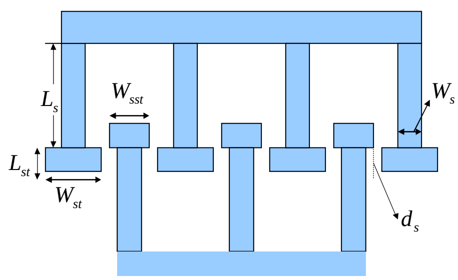
UF CONFIDENTIAL – PATENT PENDING

Resonator Schematic



UF CONFIDENTIAL – PATENT PENDING

Sensor Design



$d_s = d \qquad W_s = \lambda$

$L_{st} = \lambda \qquad X = 8L_{st}$

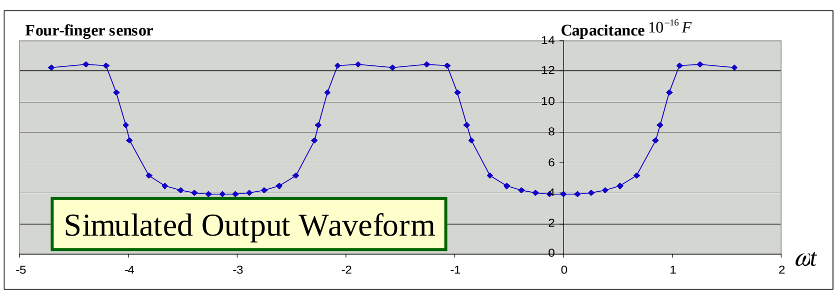
$W_{st} = W_s + 4d$

$W_{sst} = W_s + 2d$

$L_s >> d\ (L_s = 20\ d)$

$4C_{sf} \approx 8 \times 10^{-16}\ F$

(Early design w. thin fingers)



Dissipation in Resonator

Ways to minimize some major sources of dissipation:

- Air damping:
  - Vacuum packaging, small size, or optimize airflow
- Clamping losses to the substrate:
  - Locate support at a nodal point of vibration mode
  - Use impedance-mismatched supports to reflect energy back
- Thermoelastic dissipation (heat flow resulting from nonuniform strain):
  - Small size
  - Use stiff, high thermal conductivity materials (Si, diamond?)
  - Utilize modes with uniform compression/expansion
- Surface loss mechanisms:
  - Avoid layered structures (thin-film interfaces) at surfaces
- Intrinsic material losses:
  - Prefer single-crystal materials

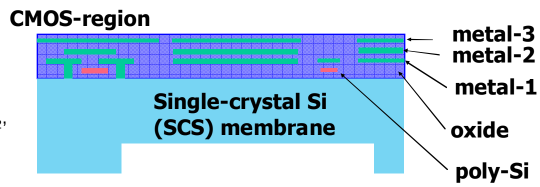
Status / Plans for Near Future

- Improved resonator designs afforded by a suitably modified post-CMOS process flow are being developed.
  - I will briefly review some aspects of the new process.
- A small prototype resonator design was taped out in a post-CMOS MEMS process (TSMC .35)
  - Parts were just received last week; are presently being etched.
- Process donation has been obtained from MOSIS for fabricating a integrated CMOS/MEMS test chip (~$20k).
  - Resonator driving a simple 2LAL shift register or adder pipeline
  - Tape-out for this chip is scheduled for July 26.
- Test the various parts separately, & together.
  - Characterize power dissipation using sensitive calorimetry techniques.

Post CMOS-MEMS Process (DRIE)

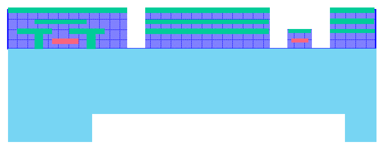
(a) Backside etch

STS: 12-sec etching
130-sccm SF_6, 13-sccm O_2, 23 mT, 600 W coil power, 12 W platen power;
8-sec passivation
85-sccm C_4F_8, 12 mT, 600 W coil power, 0 platen power.

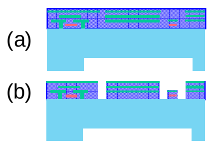


(b) Oxide etch

PlasmaTherm-790:
22.5-sccm CHF_3, 16-sccm O_2, 100 W, 125 mT for 125 minutes and then 100 mT for 10 minutes.

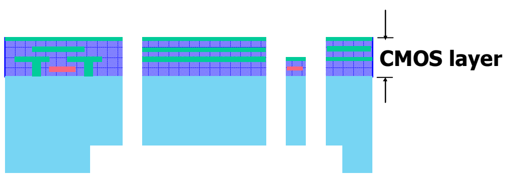


Post CMOS-MEMS Process (DRIE)



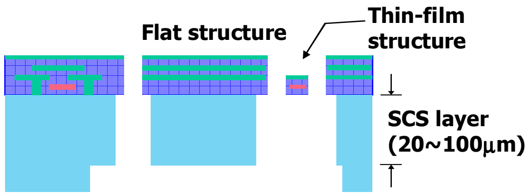
(c) Deep Si etch

STS: same as Step (a).



(d) Si undercut

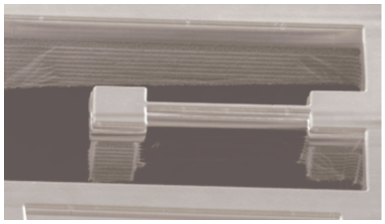
STS: 130-sccm SF_6, 13-sccm O_2, 23 mT, 600 W coil power, and 0 platen power.



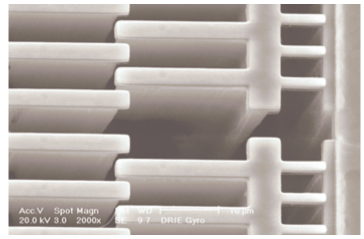
H. Xie et al, J. MEMS, Vol.11, no.2, 2002

Electrical Isolation of Silicon

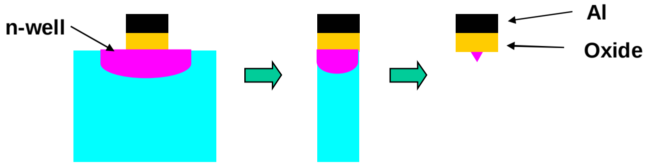
- Electrically isolated silicon island



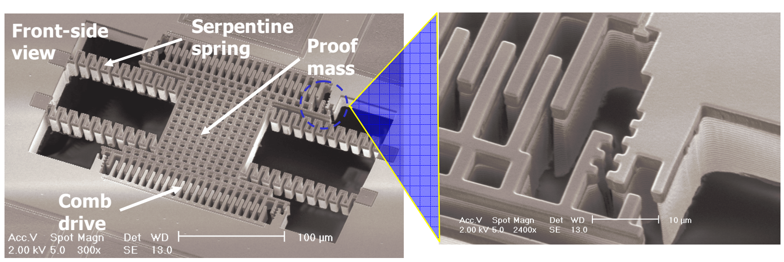
- Electrically isolated comb fingers

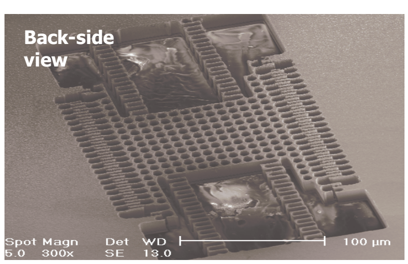


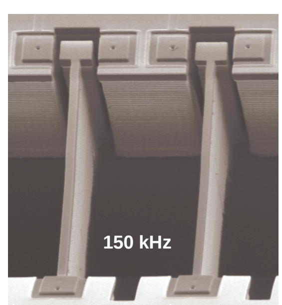
- Using n-well to improve undercut yield



DRIE CMOS-MEMS Resonators





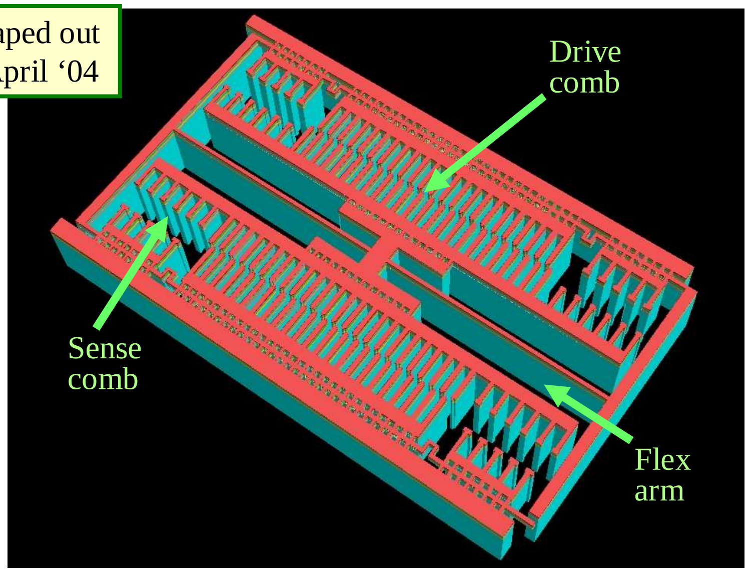


Resonators

UF CONFIDENTIAL – PATENT PENDING

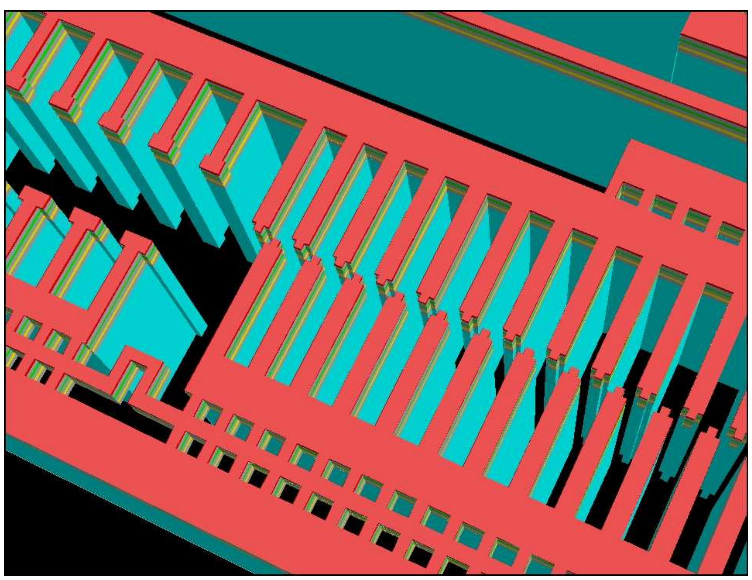
Post-TSMC35 AdiaMEMS Resonator

Taped out
April ‘04



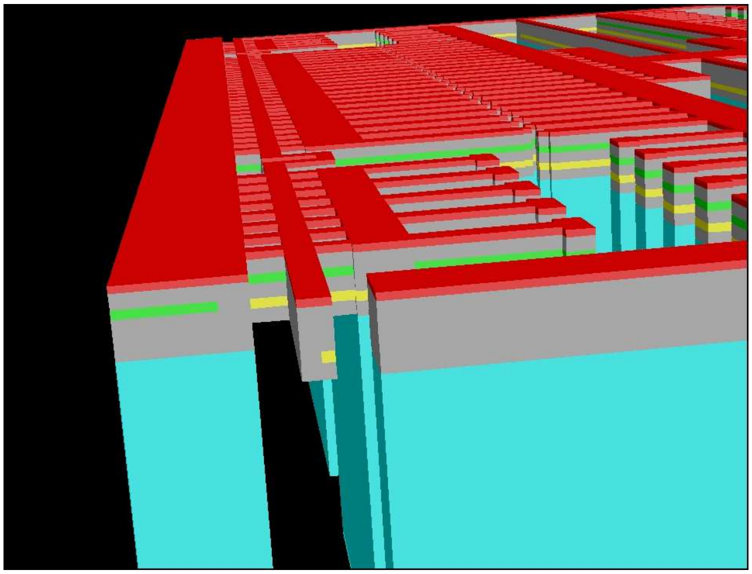
UF CONFIDENTIAL – PATENT PENDING

Close-Up View, Drive/Sense Combs



UF CONFIDENTIAL – PATENT PENDING

Side View, Showing Si Undercut



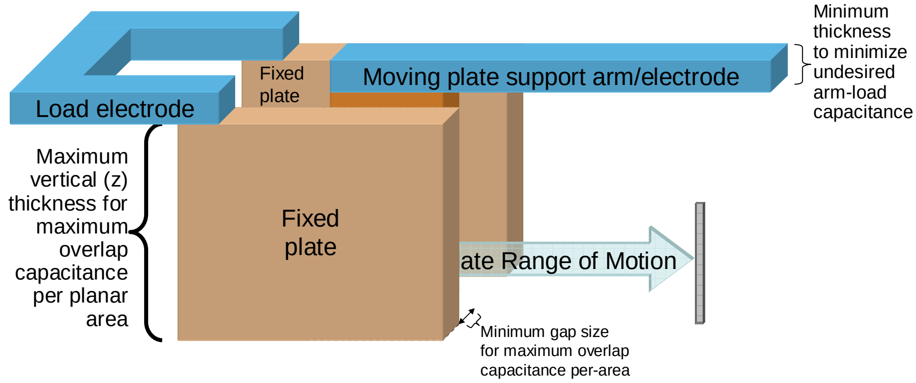
UF CONFIDENTIAL – PATENT PENDING

New Comb Finger Shape Concepts

For improved waveform shape and area efficiency

UF CONFIDENTIAL – PATENT PENDING

New Comb Finger Shape I

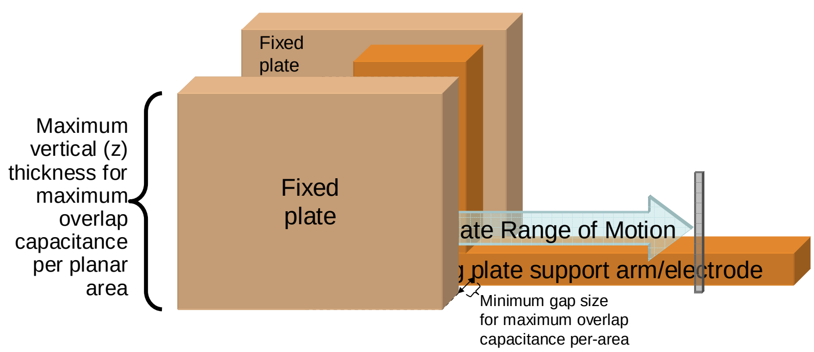


Note that the new configuration increases the magnitude of the capacitance variation while reducing the magnitude of departures from the desired trapezoidal wave shape.

Color key: Metal/oxide layers; Silicon substrate material

UF CONFIDENTIAL – PATENT PENDING

New Comb Finger Shape II

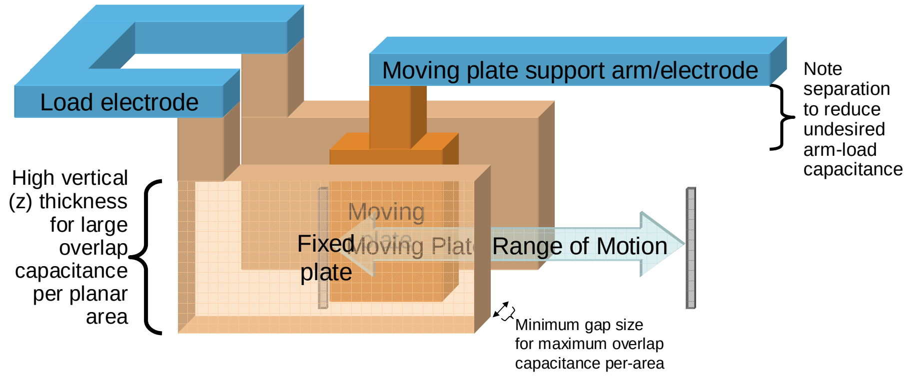


Note that the new configuration increases the magnitude of the capacitance variation while reducing the magnitude of departures from the desired trapezoidal wave shape.
In addition, the structures are made of silicon

Color key: Metal/oxide layers; Silicon substrate material

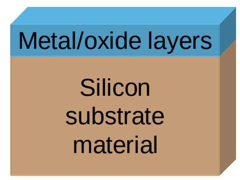
UF CONFIDENTIAL – PATENT PENDING

New Comb Finger Shape III



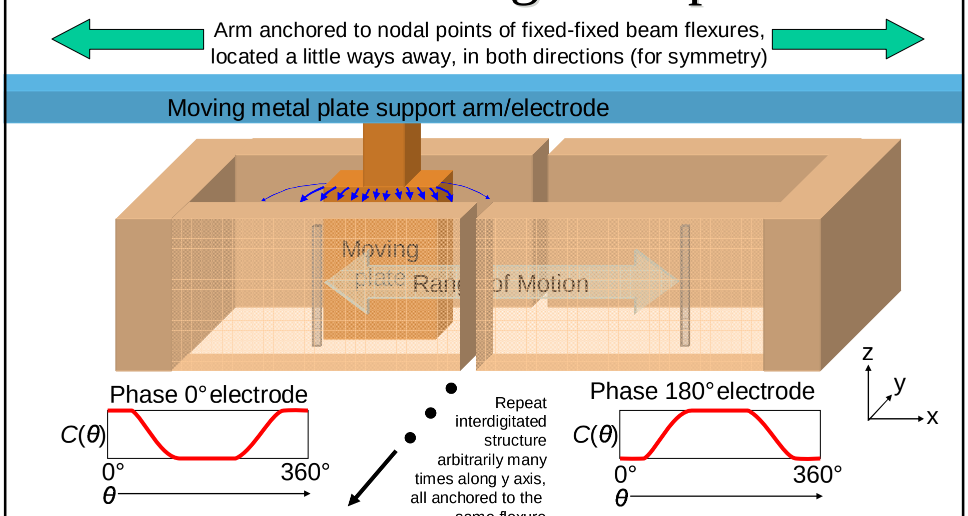
Note that the new configuration increases the magnitude of the capacitance variation while reducing the magnitude of departures from the desired trapezoidal wave shape.

Color key:



UF CONFIDENTIAL – PATENT PENDING

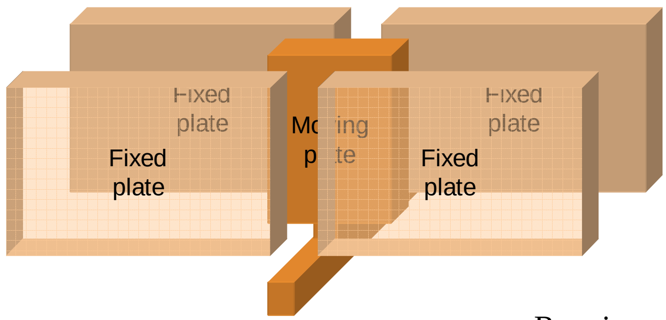
New Comb Finger Shape IV



Or, if we can do the structure on the previous slide, then why not this one too? Or, will there be a problem etching the intervening silicon out from in between the metal/oxide layers and the bulk substrate?

UF CONFIDENTIAL – PATENT PENDING

New Comb Finger Shape V

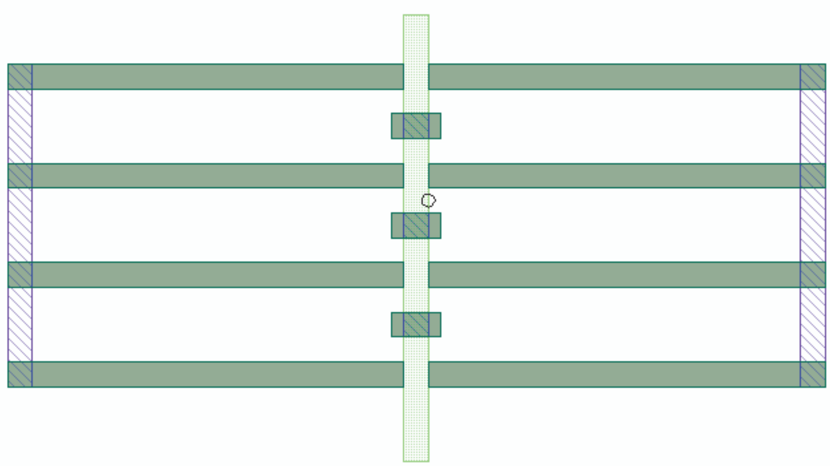


In this design, the plates are attached directly to a supprt arm which extends in the y direction instead of x. This arm can be the flexure, or it can be attached to a surrounding frame anchored to a flexure. Note that in the initial position, at all points, we only need etch from top and/or bottom, with no undercuts. Also, the flexure can be single-crystal Si.

Requires accurate, variable-depth backside etch (not presently available).

UF CONFIDENTIAL – PATENT PENDING

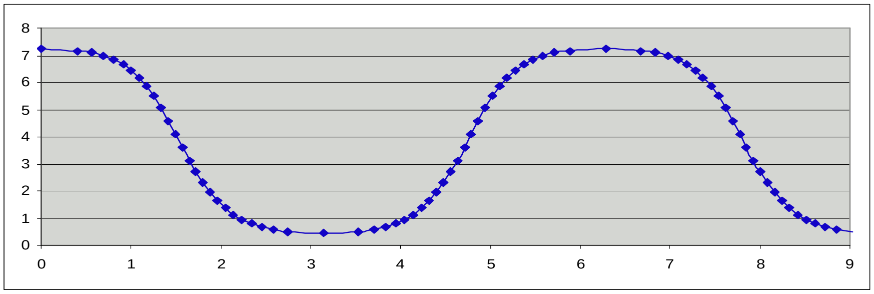
New finger: One Candidate Layout

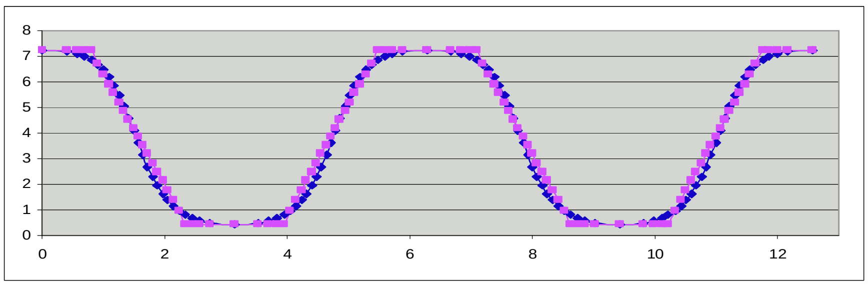


Coventor

UF CONFIDENTIAL – PATENT PENDING

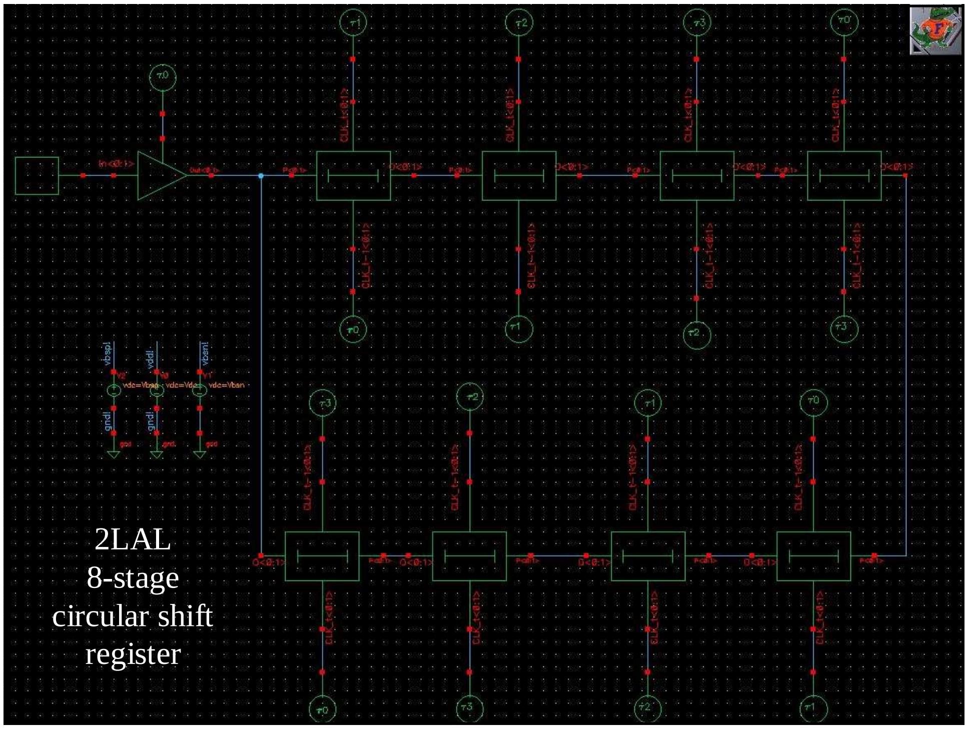
New finger simulation results



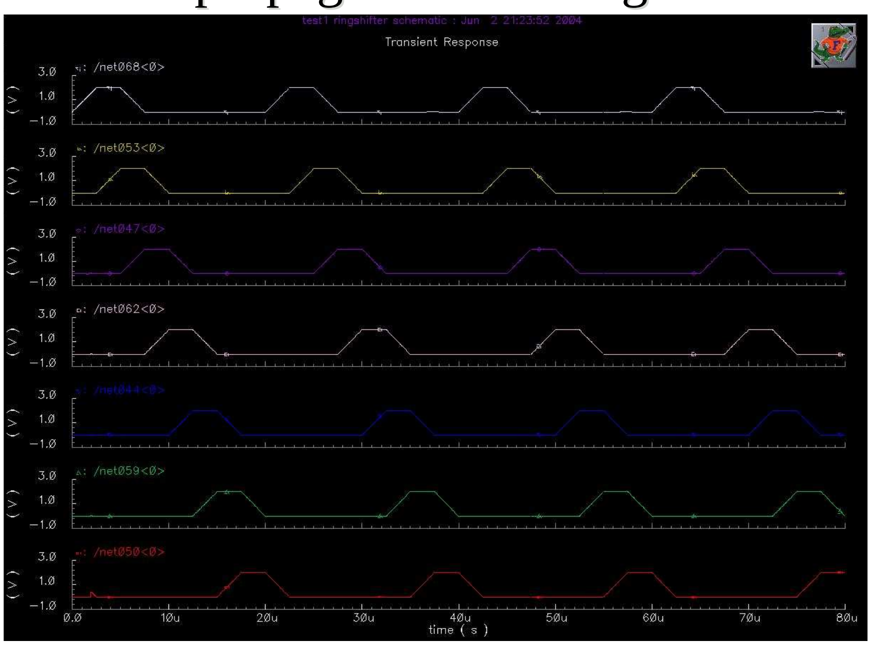


Cadence simulation results

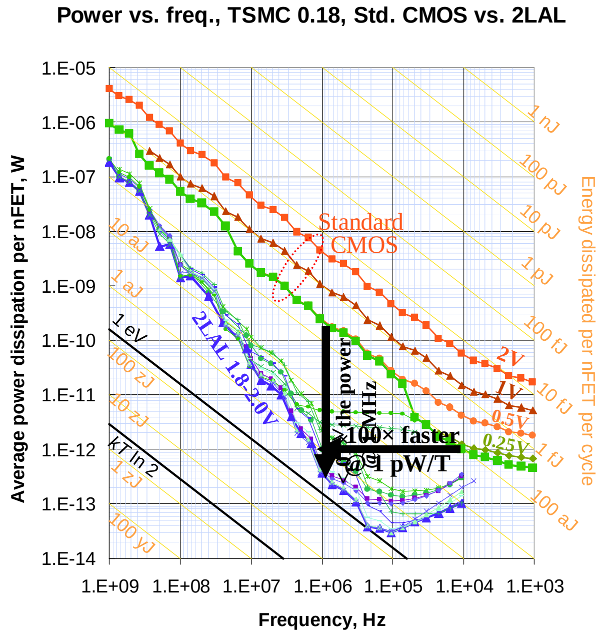
Work by AdiaMEMS project students:
Krishna Natarajan
Venkiteswaran Anantharam
(UF ECE Dept., under supervision of
Dr. Frank, CISE/ECE)



Pulse propagation in 8-stage circuit



Simulation Results from Cadence



Assumptions & caveats:

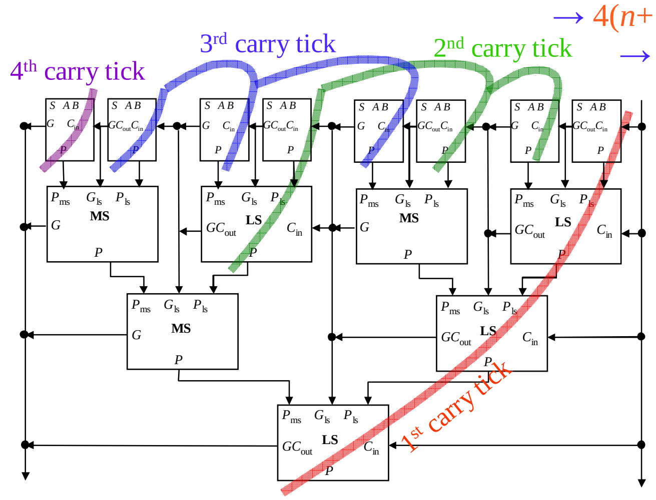
- Assumes ideal trapezoidal power/clock waveform.
- Minimum-sized devices, 2λ×3λ
  - * .18 µm (L) × .24 µm (W)
- nFET data is shown
  - * pFETs data is very similar
- Various body biases tried
  - * Higher V_{th} suppresses leakage
- Room temperature operation.
- Interconnect parasitics have not yet been included.
- Activity factor (transitions per device-cycle) is 1 for CMOS, 0.5 for 2LAL in this graph.
- Hardware overhead from fully-adiabatic design style is not yet reflected
  - * ≥2× transistor-tick hardware overhead in known reversible CMOS design styles

O(log *n*)-time carry-skip adder

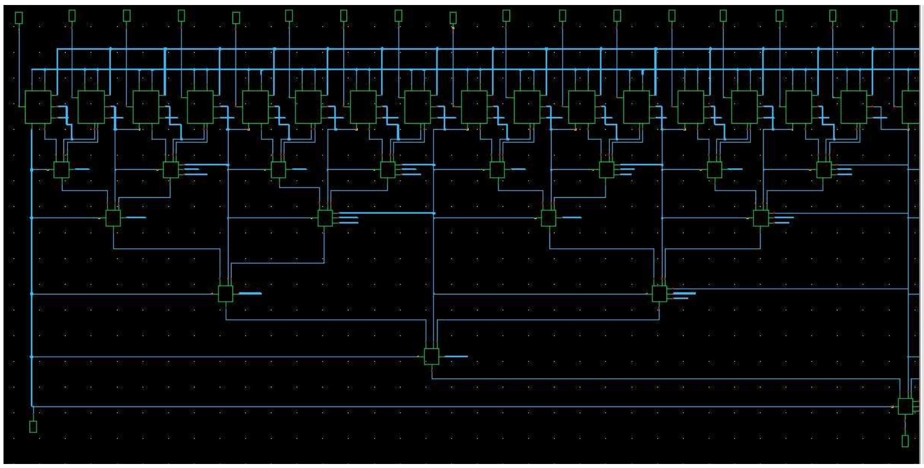
(8 bit segment shown)

With this structure, we can do a 2^n-bit add in $2(n+1)$ logic levels → $4(n+1)$ reversible ticks → $n+1$ clock cycles.

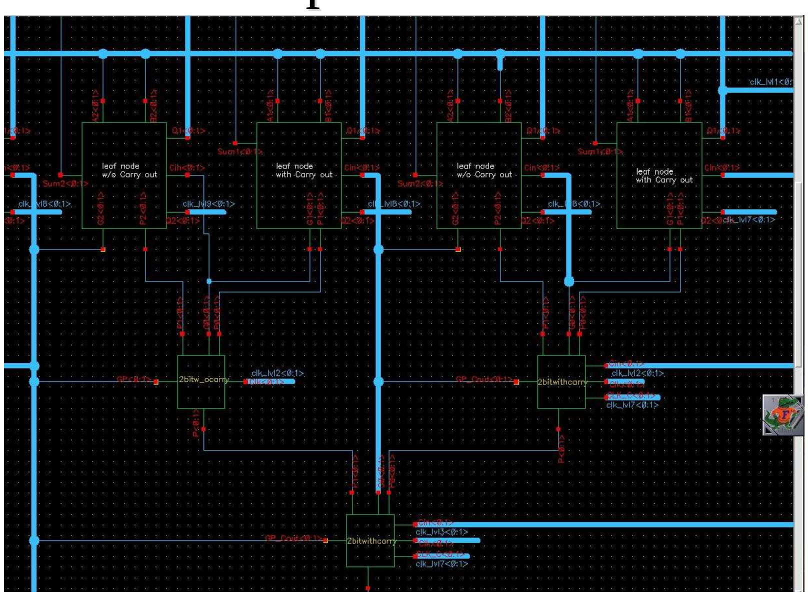
Hardware overhead is <2× regular ripple-carry.



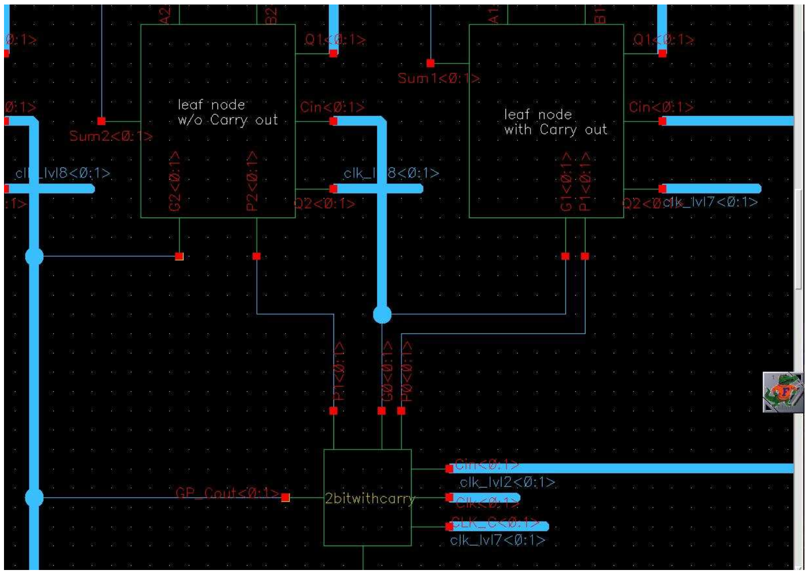
Adder Schematic – High 16 Bits



Close-Up of 4-bit Section

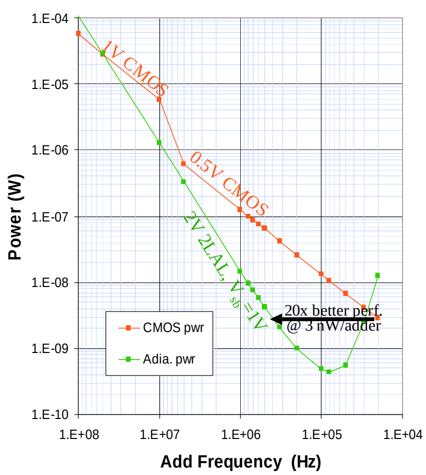


Two Bits

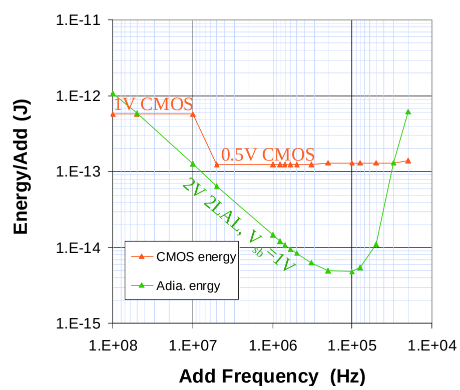


32-bit Adder Simulation Results

32-bit adder power vs. frequency



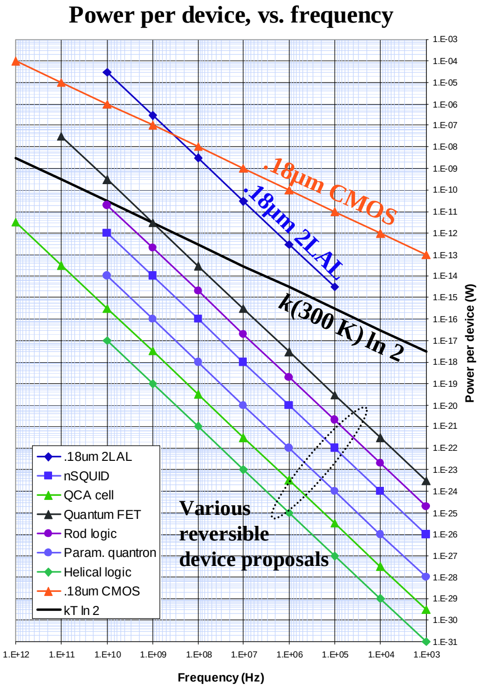
32-bit adder energy vs. frequency



(All results normalized to a throughput level of 1 add/cycle)

Plenty of Room for Device Improvement

- Recall, irreversible device technology has at most ~3-4 orders of magnitude of power-performance improvements remaining.
  - And then, the firm $kT \ln 2$ limit is encountered.
- But, a wide variety of proposed reversible device technologies have been analyzed by physicists.
  - With theoretical power-performance up to 10-12 orders of magnitude better than today's CMOS!
    - Ultimate limits are unclear.

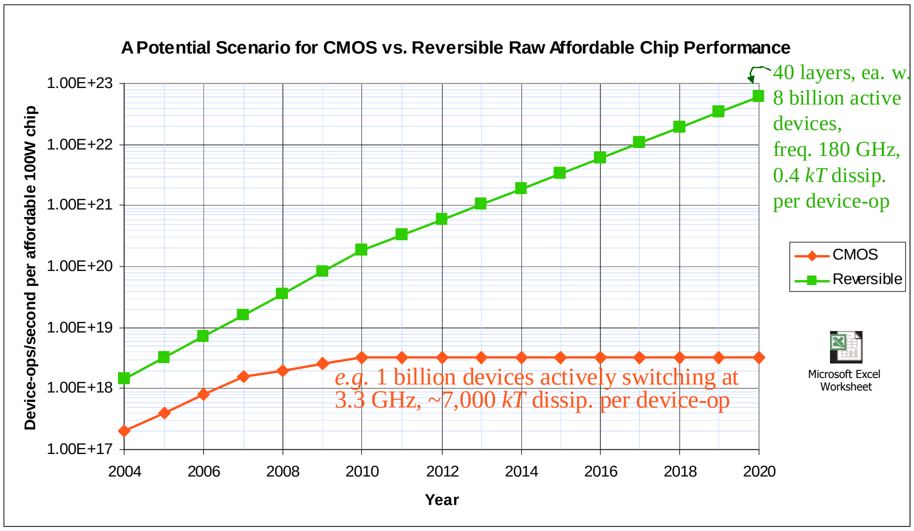


A Potential Scaling Scenario for Reversible Computing Technology

Make same assumptions as previously, except:

- Assume *energy coefficient* (energy diss. / freq.) of reversible technology continues declining at historical rate of 16× / 3 years, through 2020.
  - For adiabatic CMOS, $c_E = CV^2RC = C^2V^2R$.
    - This has been going as $\sim\ell^4$ under constant-field scaling.
  - But, requires new devices after CMOS scaling stops.
    - However, many candidates are waiting in the wings…
- Assume number of affordable *layers* of active circuitry per chip (or per package, *e.g.*, stacked dies) doubles every 3 years, through 2020.
  - Competitive pressures will tend to ensure this will happen, esp. if device-size scaling stops, as assumed.

Result of Scenario

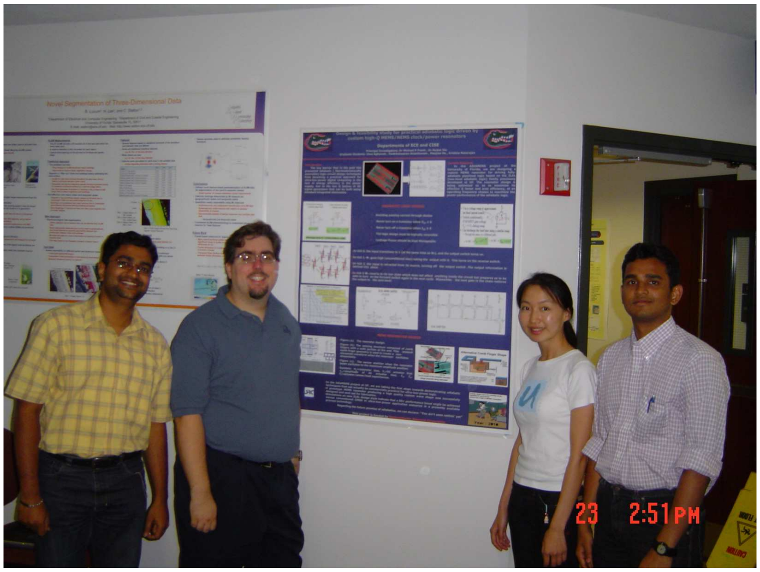


Note that by 2020, there could be a factor of 20,000× difference in *raw* performance per 100W package. (*E.g.*, a 100× overhead factor from reversible design could be absorbed while still showing a 200× boost in performance!)

Conclusions

- Standard CMOS is approaching imminent limits on raw performance per unit power consumed.
  - Due to various lower bounds on the energy dissipated by conventional irreversible switching.
- *Only* mostly-reversible logic architectures have the potential to bypass <u>all</u> of the known energy limits!
  - Via migration to an increasingly adiabatic, ballistic mode of operation, and an increasingly reversible logic design.
    - With increasingly high-*Q* energy transfers during logic.
- UF's AdiaMEMS project offers key techniques for near-term reversible computing in CMOS/MEMS.
  - Potentially viable technology for ultra-low-power products.
- Digital circuit architectures that are designed in a mostly-reversible logic style will be the *only* ones that can be easily ported to future ultra-high-performance reversible logic-device nanotechnologies.
  - We need to start paying more attention to these issues!

AdiaMEMS Project Members – Thanks!



Left to Right: Venki, Mike, Maojiao, Krishna, & Huikai