Realizing Energy-Efficient Integrated Circuits with NEM Relays

Elad Alon (UC Berkeley)
in collaboration with
Tsu-Jae King Liu (UC Berkeley),
Vladimir Stojanovic (MIT),
Dejan Marković (UCLA),
Chanro Park, Rinus Lee, Wei-Yip Loh (SEMATECH)
Digital Computing Power Crisis

Power Density Prediction circa 2000

- Since ~2000 supply voltages ($V_{DD}$) stuck at ~1V
  - Leakage stops you from lowering threshold ($V_T$)

- Leads to very poor power scaling… 1kW chips?
Parallelism to the Rescue

- Parallelism allows slower, more efficient cores
  - While maintaining overall throughput

- Works well (if you can parallel program), but…
Leakage: Game Over for CMOS

- CMOS circuits have an absolute minimum energy
  - Need to balance leakage and dynamic components

- Parallelism doesn’t help if already at $E_{\text{min}}$
NEM Relays: The Next Savior?

- Mechanical relays don’t leak, turn on abruptly
  - Potential pathway to continued energy scaling
  - Relay $E_{\text{min}}$: $\sim1\text{aJ}$ (>10X better than 65nm CMOS)
- Device/circuit co-design critical

---

R. Nathanael et al., “4-Terminal Relay Technology for Complementary Logic,” *IEDM 2009*
Relay Structure and Operation

**ON:**
\[ |V_{gb}| > V_{pi} \text{ (pull-in)} \]

**OFF:**
\[ |V_{gb}| < V_{po} \text{ (pull-out)} \]

Poly-SiGe

Tungsten
NEM Relay as a Logic Element

- Mimics operation of CMOS transistors
  - Electrostatic actuation is ambipolar
- Unlike CMOS, non-inverting logic is possible
  - Switch state set only by gate-to-body voltage
Digital Circuit Design with NEM Relays

- **CMOS**: delay set by electrical time constant
  - Cascade simple gates to distribute fanout

- **Relays**: delay dominated by mechanical movement
  - So, want all to switch simultaneously
  - Implement logic as a single complex gate (1930’s)
Need to Compare at Block Level

- Single mechanical delay per block
  - Substantially mitigates perceived delay disadvantage

- Often fewer devices for same function
  - Comparable area despite larger individual devices
Example: 32-bit Relay Adder

- **Ripple carry configuration**
  - Cascade full adder cells to create larger complex gate

- **Stack of 32 relays, still a single mechanical delay**
Scaled Relay vs. CMOS Adders

- Compare vs. CMOS adder* in 90nm technology
- For similar area:
  - >9x lower E/op
  - >10x greater delay

Parallelism

- Can extend energy benefit up to GOPS throughput
  - As long as parallelism is available
Contact Resistance

- Low contact R not critical

- Good news for reliability…
Higher contact R, hard contact (W) improves reliability
- Limits power dissipation, material flow

Current endurance record: 65 billion cycles
- Theory/experiments predict >$10^{15}$ cycles @ 1V VDD
Circuit Demonstration Platform

Test Devices
9mm

8-bit adders
4-bit and 2-bit adders
2-bit accumulator
SRAM
Flip-flops
DRAMs
7:3 Compressor
4-bit DAC
4-bit ADCs
Oscillators

F. Chen et al., ISSCC 2010, M. Spencer et al., JSSC, Jan. 2011
Relay VLSI Design Infrastructure

- Verilog-A model & Logic Synthesis customized for relays
- Flow supports multiple device designs and foundries
Looking Forward: Need Advanced Materials

- Advanced materials crucial to solving remaining technology challenges
  - E.g., W contacts unstable due to oxidation
  - Sematech enabling exploration of Ru/RuO₂ contacts
Scaling Back to The Future?

1 µm litho (UCB)

120µm x 150µm

0.25 µm litho (Sematech)

20µm x 20µm
Conclusions

- Relay characteristics enable energy scaling beyond CMOS
  - Nearly ideal $I_{on}/I_{off}$
  - Need to adapt circuit design style

- Reliability improving
  - Circuit level insights critical (contact R)
  - Demonstrated simple, operational circuits

- Potential for 10X or more lower E/op than CMOS
  - Scaling, advanced materials critical
  - Next step: >10k relay µC demo with scaled devices
Acknowledgements

- **Circuit design**
  - Matthew Spencer, Patrick Kwong, Abhinav Gupta
  - Fred Chen, Hossein Fariborzi
  - Chengcheng Wang, Kevin Dwan

- **Device design**
  - Philip Chen, Louis Hutin, Jaeseok Jeon, Hei Kam, Rhesa Nathanael, Vincent Pott

- **Sponsors**
  - DARPA NEMS program
  - FCRP (C2S2, MSD)
  - Berkeley Wireless Research Center
  - MIT CICS
  - NSF