FAMU-FSU College of Engineering

http://www.eng.fsu.edu/~mpf

# **Reversible Computing**: Its Promise and Challenges

**MARCO-FCRP/NCN Workshop on Nano-Scale Reversible Computing**

Materials, Structures, and Devices Focus Center

Massachusetts Institute of Technology

Monday, Feb. 14, 2005

# Interested Parties in Supercomputing Circles

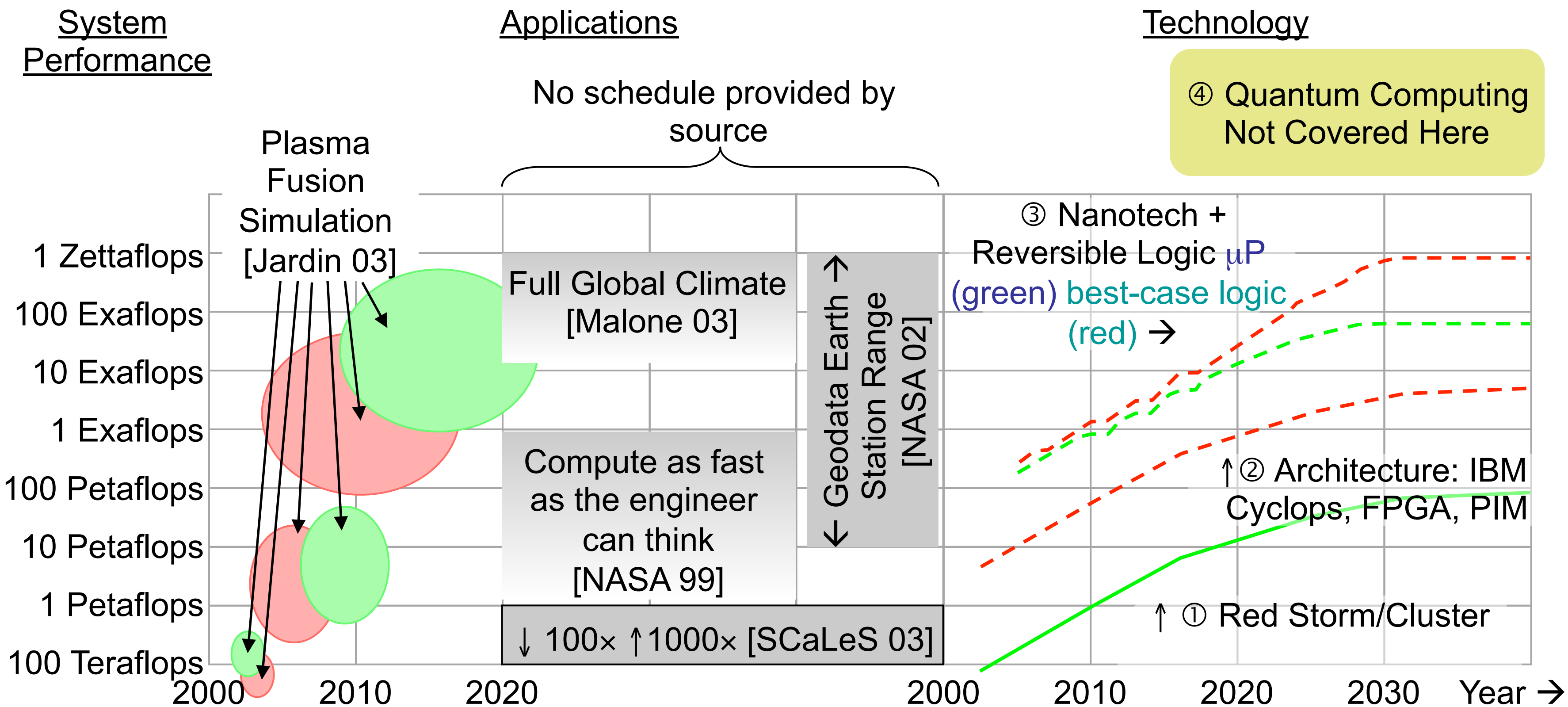
(Erik DeBene-dictis)

- Who is interested in reversible computing?
  - Well, for one thing, a growing number of influential people in the supercomputing community are…
- According to Erik DeBenedictis, these include:
  - Marc Snir
    - Ex IBM VP, architect of MPI, now heads UIUC CS dept.
    - Chair of NAS (NASA Advanced Supercomputing Division) committee on supercomputing
  - Horst Simon
    - VP-level role @ Lawrence Berkeley
    - Director of NERSC (Nat'l Energy Research Scientific Computing) center
    - Director of Computational Research Division

Sandia National Laboratories

# Applications and $100M Supercomputers

(Erik DeBene-dictis)

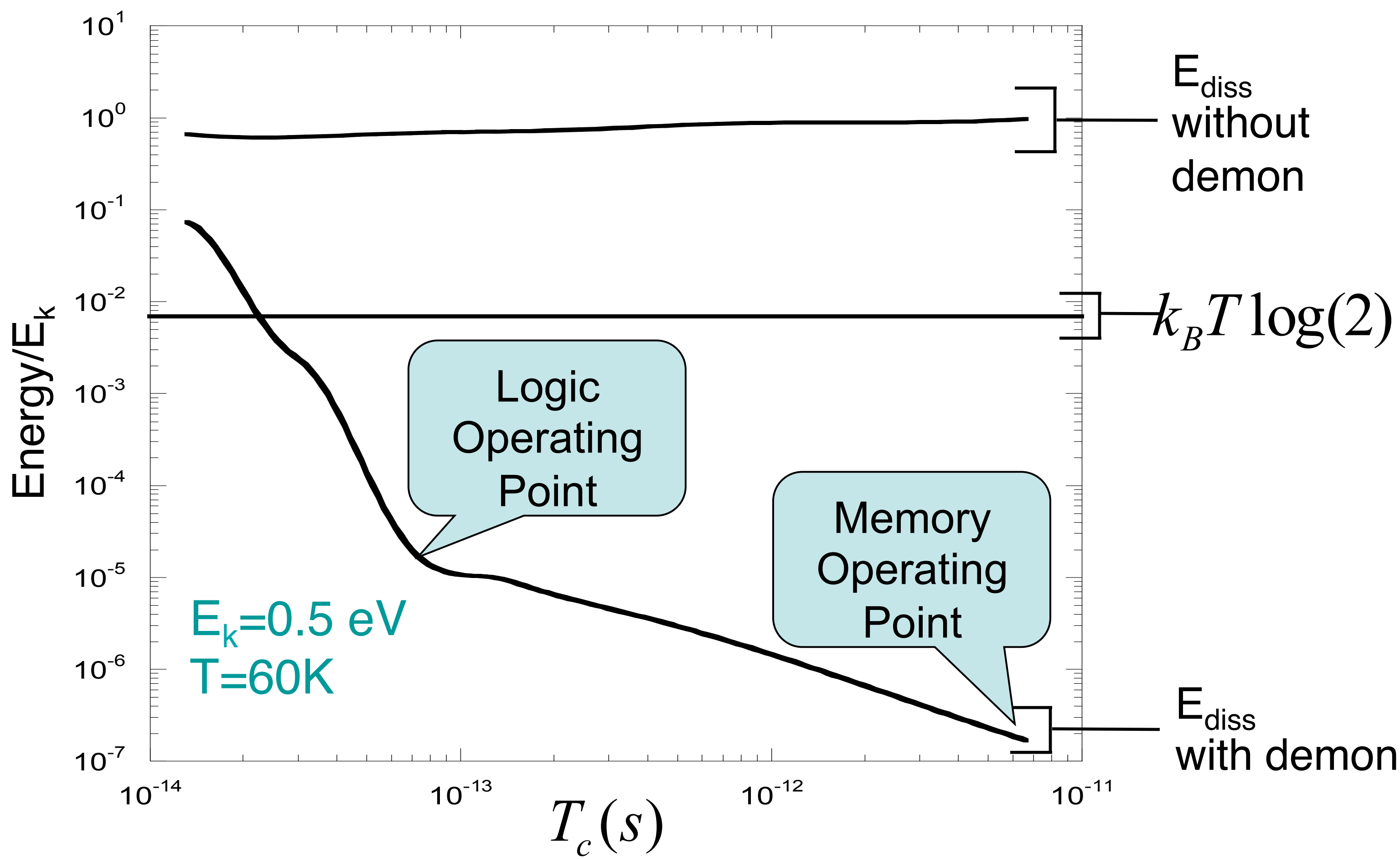



[Jardin 03] S.C. Jardin, "Plasma Science Contribution to the SCaLeS Report," Princeton Plasma Physics Laboratory, PPPL-3879 UC-70, available on Internet.
[Malone 03] Robert C. Malone, John B. Drake, Philip W. Jones, Douglas A. Rotman, "High-End Computing in Climate Modeling," contribution to SCaLeS report.
[NASA 99] R. T. Biedron, P. Mehrotra, M. L. Nelson, F. S. Preston, J. J. Rehder, J. L. Rogers, D. H. Rudy, J. Sobieski, and O. O. Storaasli, "Compute as Fast as the Engineers Can Think!" NASA/TM-1999-209715, available on Internet.
[NASA 02] NASA Goddard Space Flight Center, "Advanced Weather Prediction Technologies: NASA's Contribution to the Operational Agencies," available on Internet.
[SCaLeS 03] Workshop on the Science Case for Large-scale Simulation, June 24-25, proceedings on Internet a http://www.pnl.gov/scales/.
[DeBenedictis 04], Erik P. DeBenedictis, "Matching Supercomputing to Progress in Science," July 2004. Presentation at Lawrence Berkeley National Laboratory, also published as Sandia National Laboratories SAND report SAND2004-3333P. Sandia technical reports are available by going to http://www.sandia.gov and accessing the technical librar

Sandia National Laboratories

(Erik DeBene-dictis)

# Energy loss for erasing a single bit




From Craig Lent, *Maxwell's Demon* paper

(Erik DeBenedictis)

# RL for Computational Science and Defense

- Supercomputing drives science and defense
- Planning exercises indicate most ambitious problems reach 1 Zettaflops
- 1 Zettaflops simply exceeds Landauer's Limit
  - Assuming $100M budget
- Reversible logic could be a solution
- If reversible logic is to be practical, let's get started with engineering
- If reversible logic is not to be practical, we need to get started seeking solutions to science and defense needs in some other way

# What is Reversible Computing?

Hereafter: "**RC**"

- A working definition, for our purposes:
  - A computing process in which only a fraction $f \ll 1$ of a typical logic signal's energy gets dissipated to heat (on average) per digital manipulation of the signal.
    - Where "manipulations" most generally could include bit storage, communication, and logic operations.
- It is associated with the following further claims:
  - The value of $f$ has no fundamental, technology-independent lower bound greater than 0.
  - Further, the energy *dissipated* can be $\ll kT \ln 2$, even if the signal energy *itself* is well over $kT$.
- Let us now consider the status of the question:
  - Is RC (as characterized above) physically possible?
    - Secondary question: Can it also be economically viable?

# Status of Reversible Computing

- **Fact:** There is **no** valid, rigorous proof in the literature (assuming only rock-solid axioms of fundamental physics) that would justify our dismissing reversible computing as impossible.
  - At least, I haven't come across one in a decade of immersion in this field.
    - I have read and carefully studied hundreds of related papers.
  - Every purported "proof" of impossibility that I've seen (and there are quite a few) contains major flaws.
    - Either logical fallacies, or unjustified assumptions.
- But, also **fact:** A complete physical model of a good reversible computer (including all the relevant physics of all necessary subsystems) has not been described, quite yet... (Let alone a full working prototype.)
  - Current models/hardware come very close, but none are quite there yet…
- But, also **fact:** Reversible computing will be absolutely necessary if we are to circumvent the various near-term power-related performance limits.
  - Without it, industry's progress will stall much sooner, rather than much later.
- **My position:** We should pursue RC vigorously, at least until:
  - We find a rigorous impossibility proof that the scientific community agrees on…
    - A "beyond a reasonable doubt" criterion is required to "convict" RC of impossibility.
  - Or, until many serious, independent efforts try and fail to breach the *kT* barrier
    - "Inductive leap" argument. (Historical precedent: Perpetual motion machines.)

# Myths of Historical Fact

- **Widespread myth:** "von Neumann/Shannon proved that logic/communication requires fixed dissipation."
  - **No.** All we have from von Neumann on this is this brief, second-hand account of a lecture, which contains no rigorous proof, or even clearly stated assumptions.
- We **cannot infer** that von Neumann would not have fully endorsed reversible logic, if the concept behind it had been known to him.

> He then calculated the energy which is dissipated "per elementary act of information, that is, per elementary decision of a two-way alternative and per elementary transmittal of 1 unit of information." He did this for three cases: the thermodynamical minimum, the vacuum tube, and the neuron.
>
> In the third lecture he said that thermodynamical information is measured by the logarithm, to the base two, of the number of alternatives involved. The thermodynamical information in the case of two alternatives is thus one, "except that this is not the unit in which you measure energy. Entropy is energy only if you specify the temperature. So, running at low temperature you can say what energy should be dissipated." He then computed the thermodynamical minimum of energy per elementary act of information from the formula $kT \log_e N$ ergs, where $k$ is Boltzmann's constant ($1.4 \times 10^{-16}$ ergs per degree), $T$ is the temperature in absolute units, and $N$ is the number of alternatives. For a binary act $N = 2$, and taking the temperature to be about 300 degrees absolute, he obtained $3 \times 10^{-14}$ ergs for the thermodynamical minimum.

(*Theory of Self-Reproducing Automata*, p.66)

- As for Shannon, his papers explicitly address signal power transmitted, but *nowhere* say that this power must be dissipated.
  - E.g., see *Collected Papers*…

# Generalization Fallacy

- A common fallacy in the anti-RC literature:
  - "I tried a few ways to do it, and couldn't figure out how to get it to work. Therefore, it must be impossible."
- **<u>No.</u>** *Many* aspects of reversible computing were originally conjectured to be impossible, then were later found to be possible (and often rather easy).
  - I can show you quite a large list of such items! (Next slide)

# Some Doubts and Answers

| Some Claims Against Reversible Computing | Eventual Resolution of Claim |
|---|---|
| John von Neumann, 1949 – Offhandedly claims during a lecture that computing requires *kT* ln 2 dissipation per "elementary act of decision" (bit-operation). | No proof provided. Twelve years later, Rolf Landauer of IBM tries valiantly to prove it, but succeeds only for logically irreversible operations. |
| Rolf Landauer, 1961 – Proposes that the logically irreversible operations which necessarily cause dissipation are unavoidable. | Landauer's argument for unavoidability of logically irreversible operations was conclusively refuted by Bennett's 1973 paper. |
| Bennett's 1973 construction is criticized for using too much memory. | Bennett devises a more space-efficient version of the algorithm in 1989. |
| Bennett's models criticized by various parties for depending on random Brownian motion, and not making steady forward progress. | Fredkin and Toffoli at MIT, 1980, provide ballistic "billiard ball" model of reversible computing that makes steady progress. |
| Various parties note that Fredkin's original classical-mechanical billiard-ball model is chaotically unstable. | Zurek, 1984, shows that quantum models can avoid the chaotic instabilities. (Though there are workable classical ways to fix the problem also.) |
| Various parties propose that classical reversible logic principles won't work at the nanoscale, for unspecified or vaguely-stated reasons. | Drexler, 1980's, designs various mechanical nanoscale reversible logics and carefully analyzes their energy dissipation. |
| Carver Mead, CalTech, 1980 – Attempts to show that the *kT* bound is unavoidable in electronic devices, via a collection of counter-examples. | No general proof provided. Later he asked Feynman about the issue; in 1985 Feynman provided a quantum-mechanical model of reversible computing. |
| Various parties point out that Feynman's model only supports serial computation. | Margolus at MIT, 1990, demonstrates a parallel quantum model of reversible computing—but only with 1 dimension of parallelism. |
| People question whether the various theoretical models can be validated with a working electronic implementation. | Seitz and colleagues at CalTech, 1985, demonstrate working energy recovery circuits using adiabatic switching principles. |
| Seitz, 1985—Has some working circuits, unsure if arbitrary logic is possible. | Koller & Athas, Hall, and Merkle (1992) separately devise general reversible combinational logics. |
| Koller & Athas, 1992 – Conjecture reversible *sequential* feedback logic impossible. | Younis & Knight @MIT do reversible sequential, pipelineable circuits in 1993-94. |
| Some computer architects wonder whether the constraint of reversible logic leads to unreasonable design convolutions. | Vieri, Frank and coworkers at MIT, 1995-99, refute these qualms by demonstrating straightforward designs for fully-reversible, scalable gate arrays, microprocessors, and instruction sets. |
| Some computer science theorists suggest that the algorithmic overheads of reversible computing might outweigh their practical benefits. | Frank, 1997-2003, publishes a variety of rigorous theoretical analysis refuting these claims for the most general classes of applications. |
| Various parties point out that high-quality power supplies for adiabatic circuits seem difficult to build electronically. | Frank, 2000, suggests microscale/nanoscale electro-mechanical resonators for high-quality energy recovery with desired waveform shape and frequency. |
| Frank, 2002—Briefly wonders if synchronization of parallel reversible computation in 3 dimensions (not covered by Margolus) might not be possible. | Later that year, Frank devises a simple mechanical model showing that parallel reversible systems can indeed be synchronized locally in 3 dimensions. |

# Straw-Man Arguments

- A popular, but unsound, debating tactic:
  - Mis-characterize opponent's claims as being something other than what they really are,
    - something very easy to refute.
  - Then, easily knock down the "straw man" that one has conveniently set up for oneself,
    - instead of dismantling the opponent's true position.
      - which may be a lot more difficult to do!
  - Finally, pretend that one has proved something meaningful by this.
- The honest scientist/engineer must scrupulously avoid using such tactics.

# Why we should expect that RC will turn out to be possible

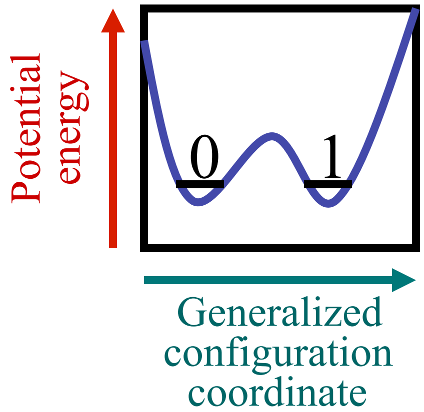
- Quantum theory tells us (quite rigorously) that <u>any</u> perfectly-closed system evolves unitarily (reversibly), with <u>no</u> entropy increase…
  - Given a detailed model where the underlying physics (*e.g.*, a Hamiltonian) is perfectly known, and all our knowledge about the state is tracked, without throwing any away.
- If the system is <u>not</u> perfectly closed, or the physics is <u>not</u> quite perfectly known, or if our dynamical model repeatedly <u>discards</u> knowledge about the state, then entropy will steadily increase at some rate, until equilibrium…
  - But, we can expect that as isolation setups become better, our physics becomes more accurate, and the state evolution is more faithfully tracked by the model, the <u>rate</u> of entropy increase can be suppressed to arbitrarily low levels.
    - Note that this is true independently of the initial state of the system!
- A reversible computer is then just a precisely-modeled physical system whose initial state happens to have been pre-arranged so that its dynamical trajectory will (by design) closely correspond to a desired computation.
  - With only a very small rate of accumulation of entropy (and also deterministic error)
    - → And only very low power input needed to remove the entropy / correct the errors
  - → There is nothing about a process of "computation" <u>per se</u> that makes it <u>less</u> predictable than general physical systems! (It can't be, it's just a special case.)
- Thus, any successful proof that RC is impossible would basically have to prove one of the following:
  - (a) Physical systems *cannot* be arbitrarily well isolated from outside disturbances
  - (b) A system's Hamiltonian *cannot* be determined with arbitrarily high precision

# Bistable Potential-Energy Wells

## A Technology-Independent Model of Digital Devices

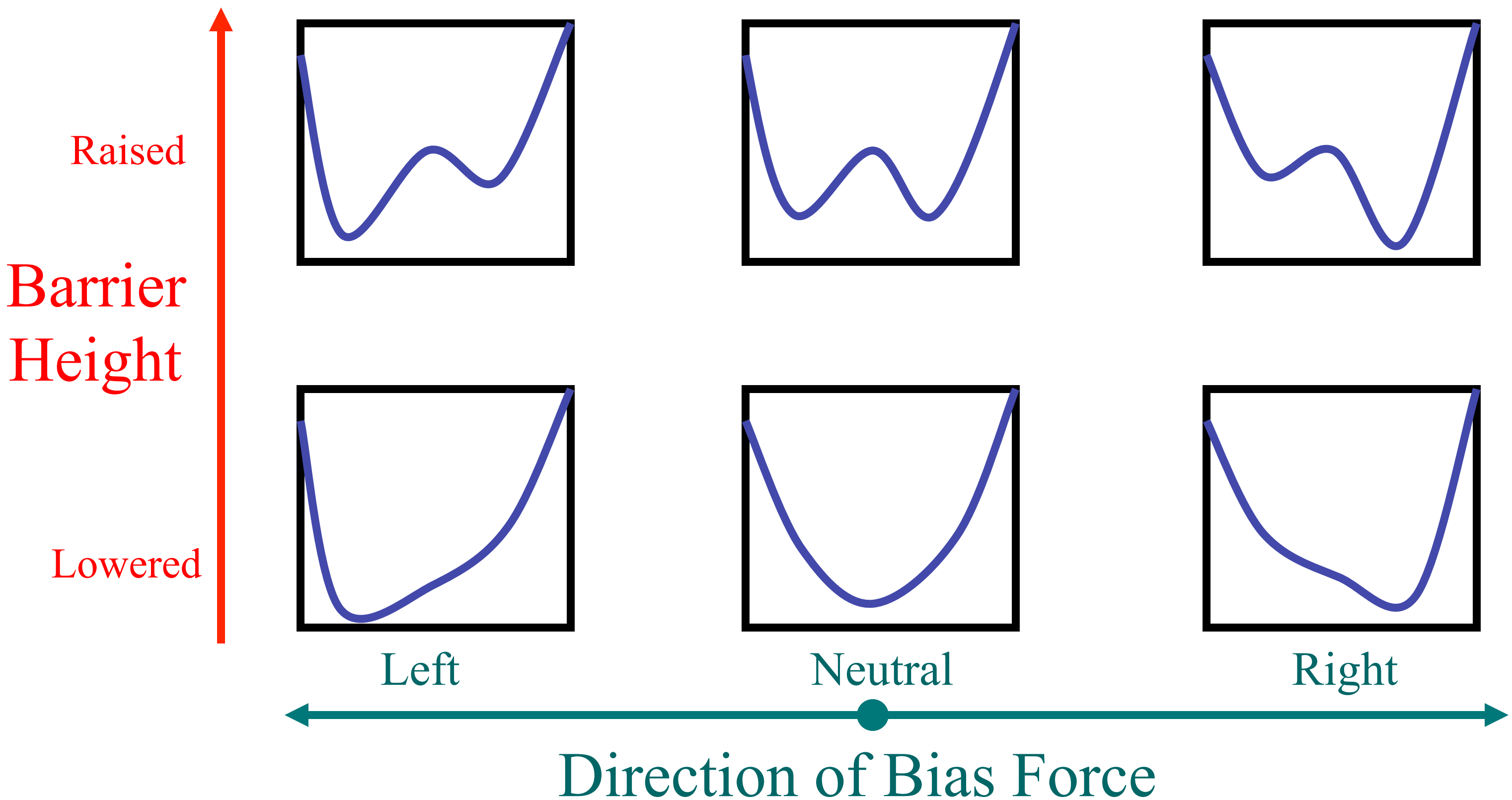
## (Based on Landauer ’61)

- Consider any system having an (adjustable) potential energy surface (PES) in its configuration space.
  - The PES should have at least two local minima (or *wells*)
  - Therefore the system is *bistable*
    - It has two stable (or at least metastable) configurations
      - Located at well bottoms
- The two stable states form a natural bit.
  - One state can represent 0, the other 1.
    - This picture can also be easily generalized to larger numbers of stable states.
- Consider now a PES having two adjustable parameters:
  - (1) “Height” (energy) of the potential energy barrier between wells, relative to well bottoms
  - (2) Relative height of the left and right states in the well (call this “bias”)



# Possible Parameter Settings

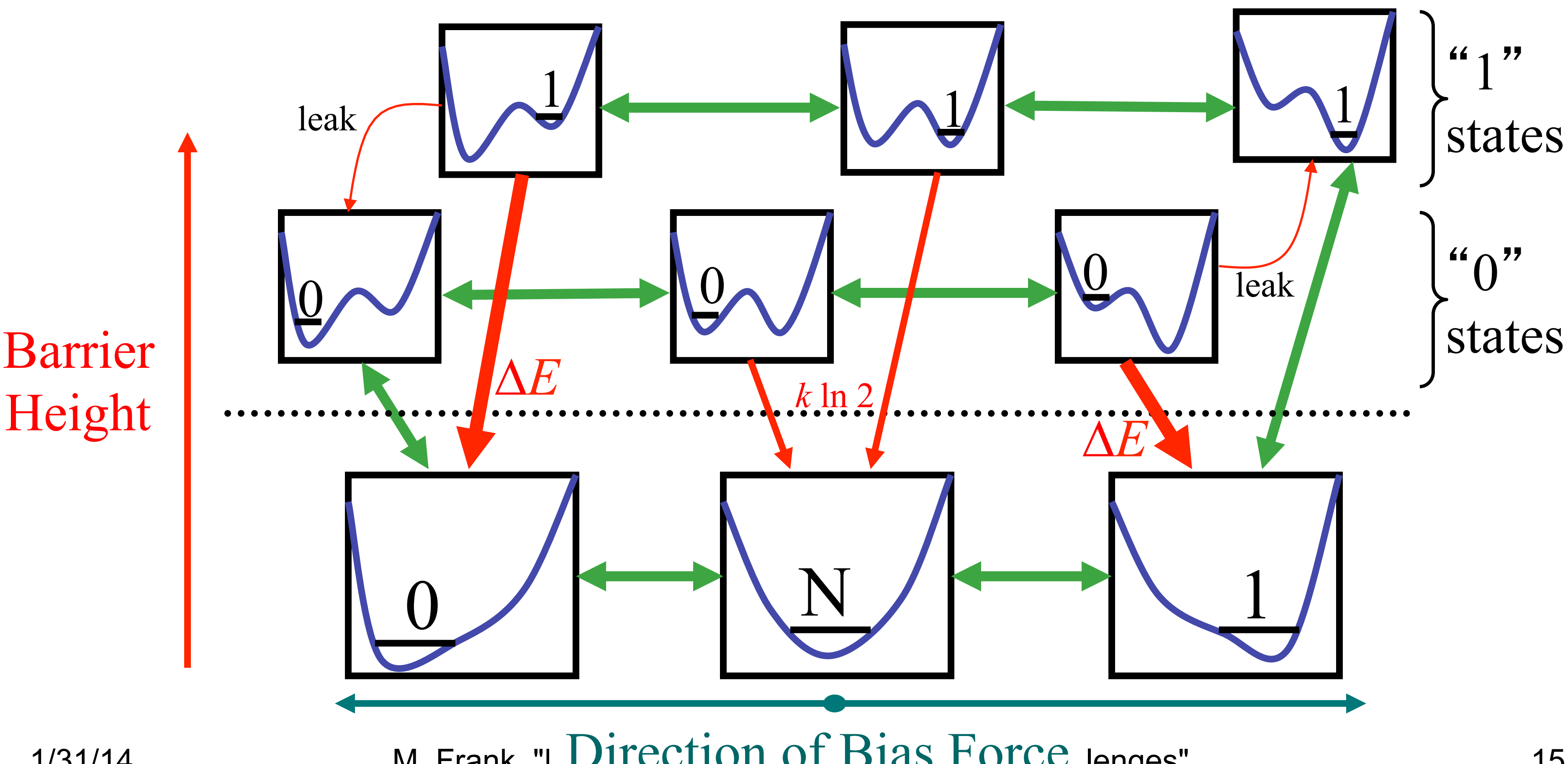
- In the following slides, we will distinguish six qualitatively different settings of the well parameters, as shown below…



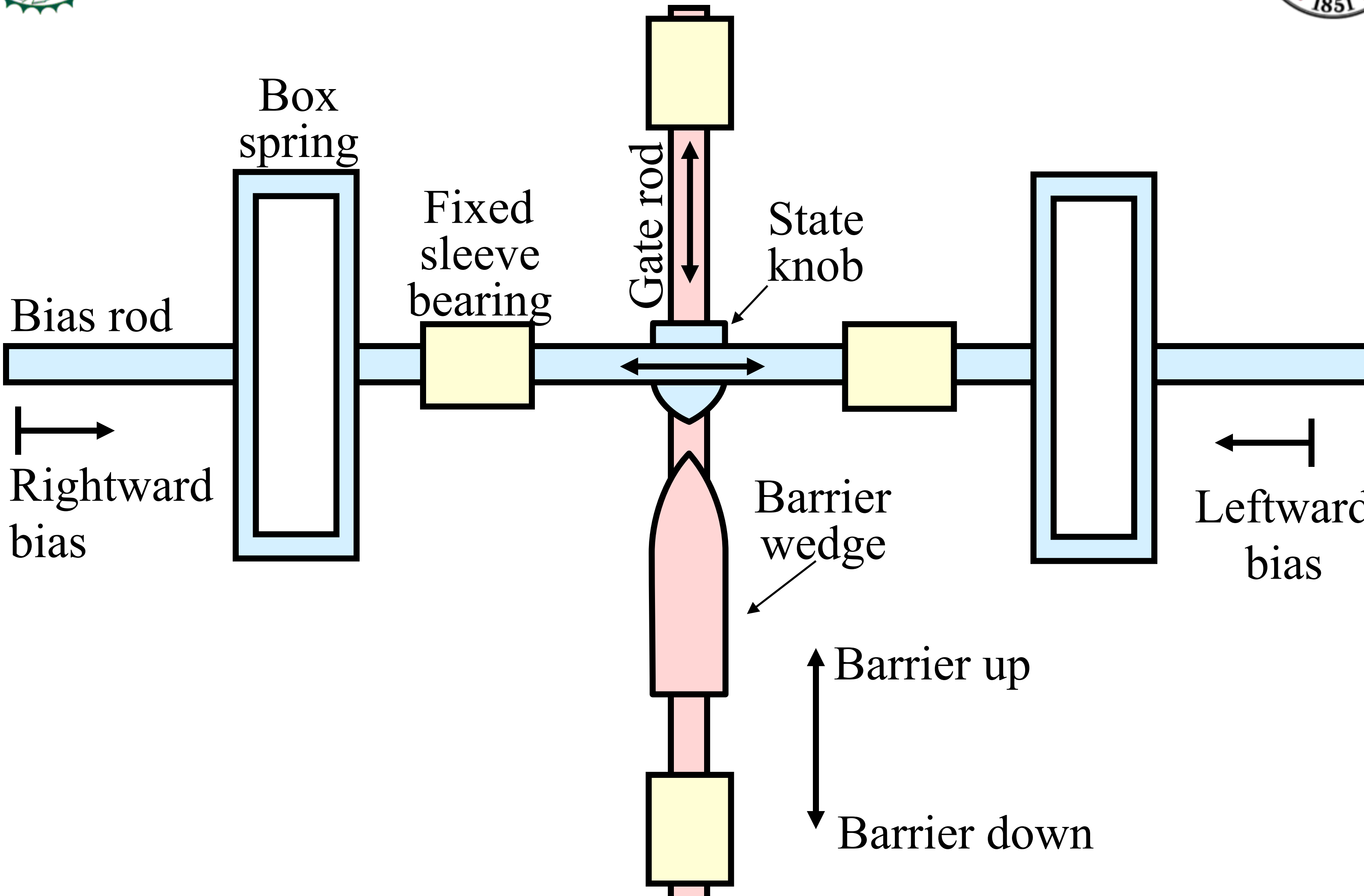
# Possible Well Transitions

- Catalog of all the possible transitions in the bistable wells, adiabatic & not...
  - We can characterize a wide variety of digital logic and memory styles in terms of how their operation corresponds to subgraphs of this diagram.

(Ignoring superposition states.)

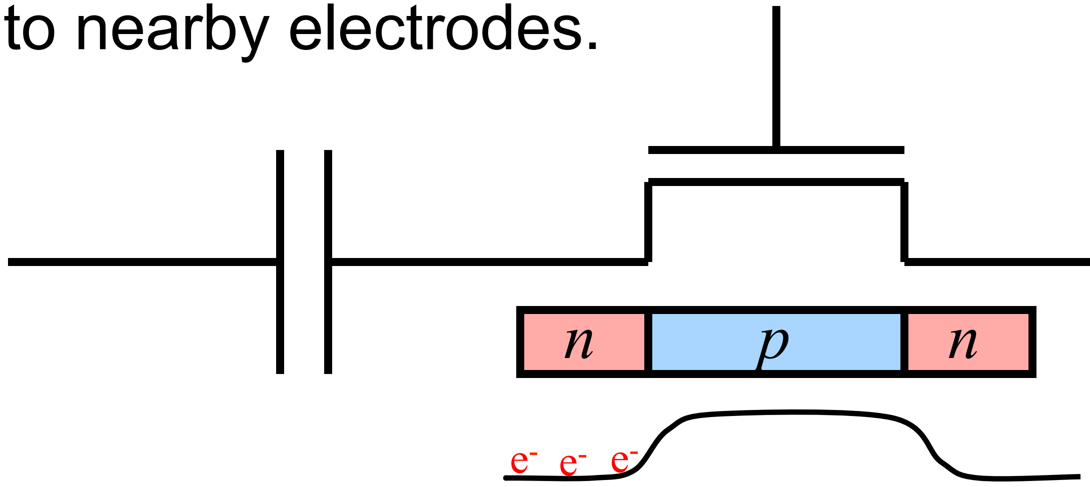


# Simple Mechanical Model



# MOSFET Implementation

- The logical state is in the location of a charge packet (excess of electrons) on either side terminal of a FET.
  - The charge packet might even consist of just a single excess electron in a sufficiently small (nanoscale) logic node.
- The potential energy barrier is provided by the built-in voltage across the PN junctions in the FET.
  - The barrier height is lowered when the device is turned on by adjusting the voltage on the gate electrode.
- Bias forces can be provided by (*e.g.*) capacitive coupling to nearby electrodes.

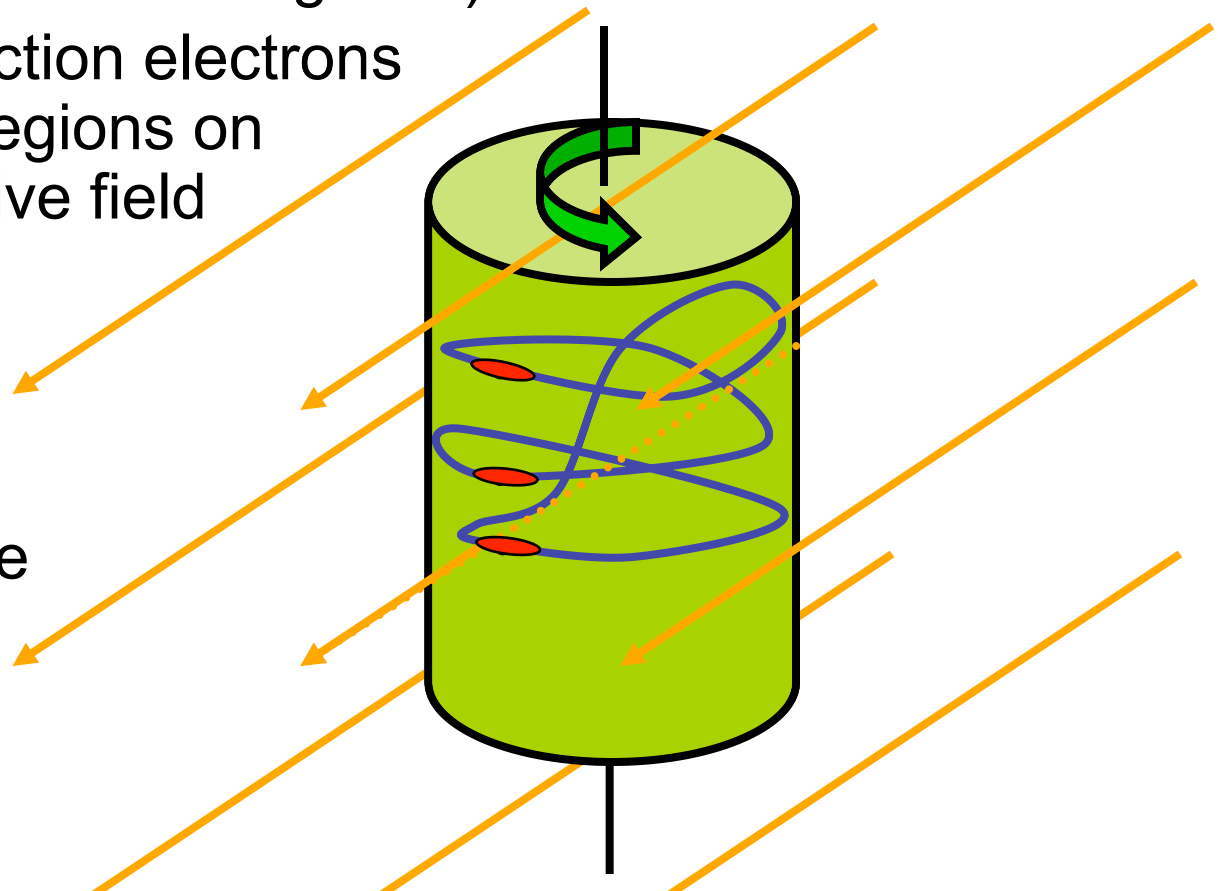


# Helical Logic

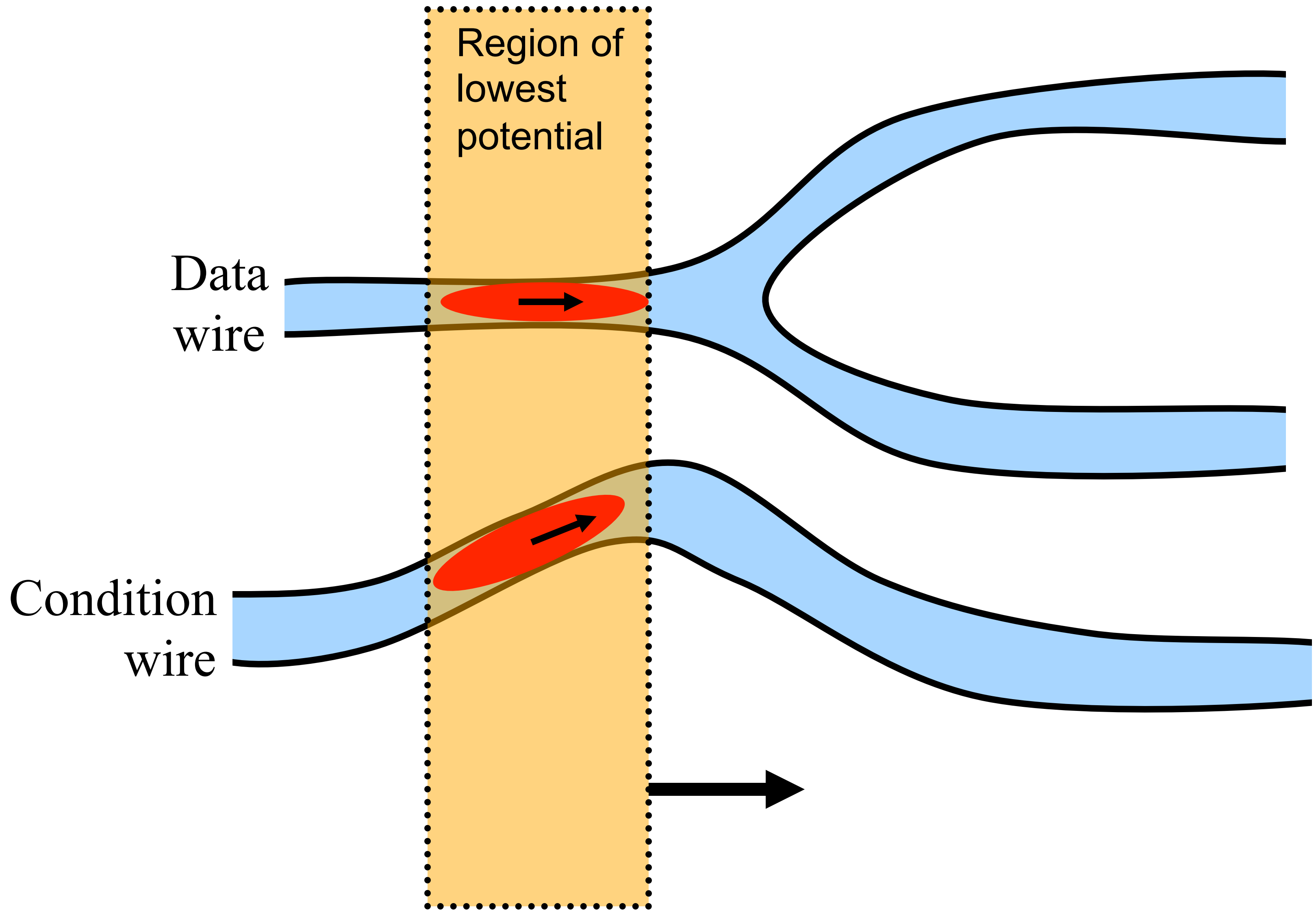
- A proposal by R. Merkle & K. E. Drexler,
  - Published in *Nanotechnology* journal, 1996
- Shows that we can do reversible logic using wires only!
  - Other structures are not needed...
- The wires *are* the devices!
  - Uses simple Coulombic repulsion between small packets of electrons to do logic
- Scales to single electrons, and nanoscale wires…
  - Can also use resistance-free “wires” consisting of vacuum waveguides
- Globally clocked...
  - by rotation of wiring relative to a global electrostatic field
- Can be used reversibly… $10^{-27}$ J/op @ 10 GHz!
  - This at low temps (1K), but still much more energy-efficient than FETs
    - Even when overhead of cooling is accounted for.

# HL: Overall Physical Structure

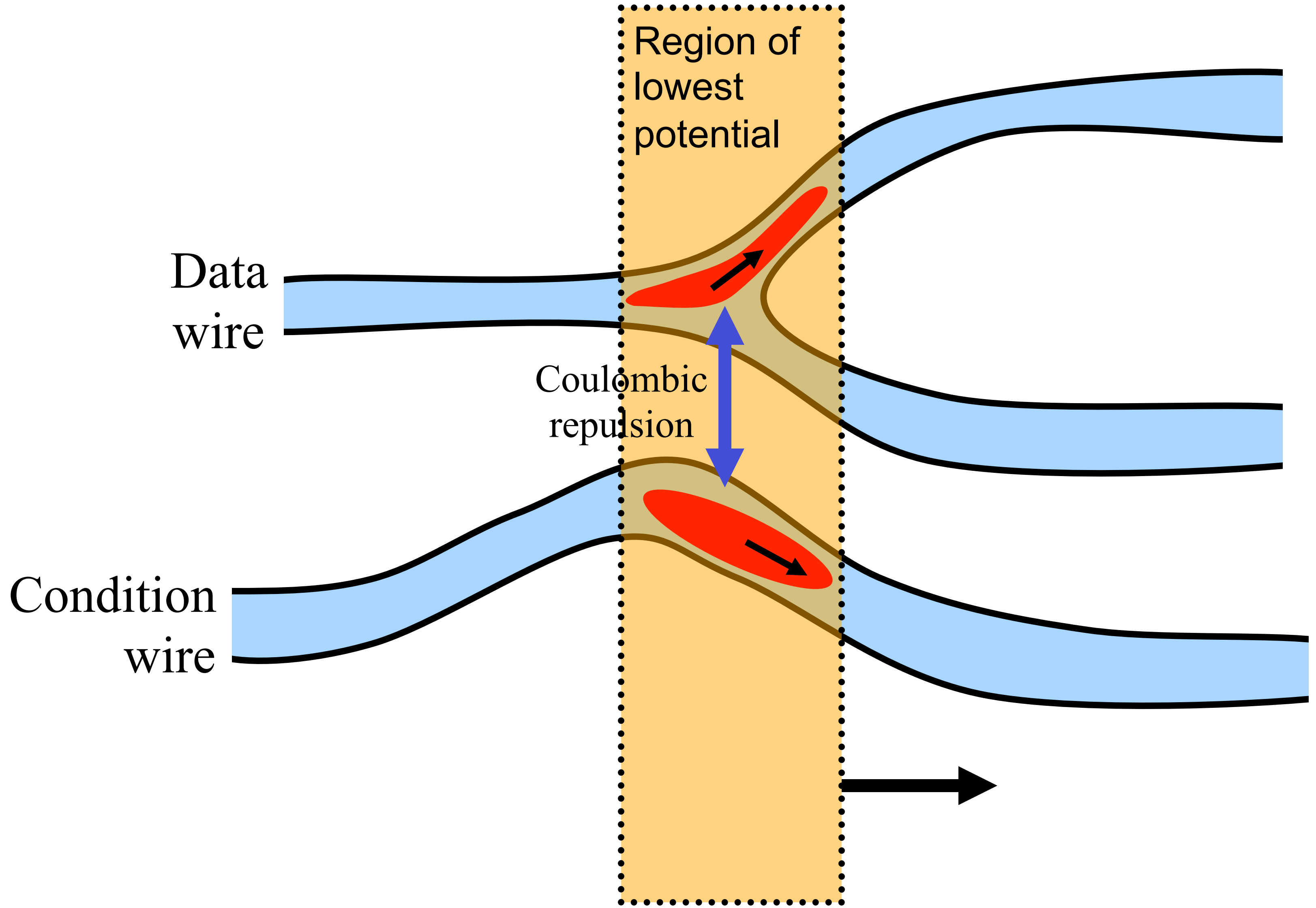
- Consider a cylinder of low-κ insulating material (*e.g.*, glass), containing embedded coils of wire (electron waveguides), rotating on its axis in a static, flat electric field (or, unmoving in a rotating field).
- An excess of conduction electrons will be attracted to regions on wire closest to positive field direction.
- These electron packets follow the field along as it rotates relative to the cylinder.
- Next slide: Logic!



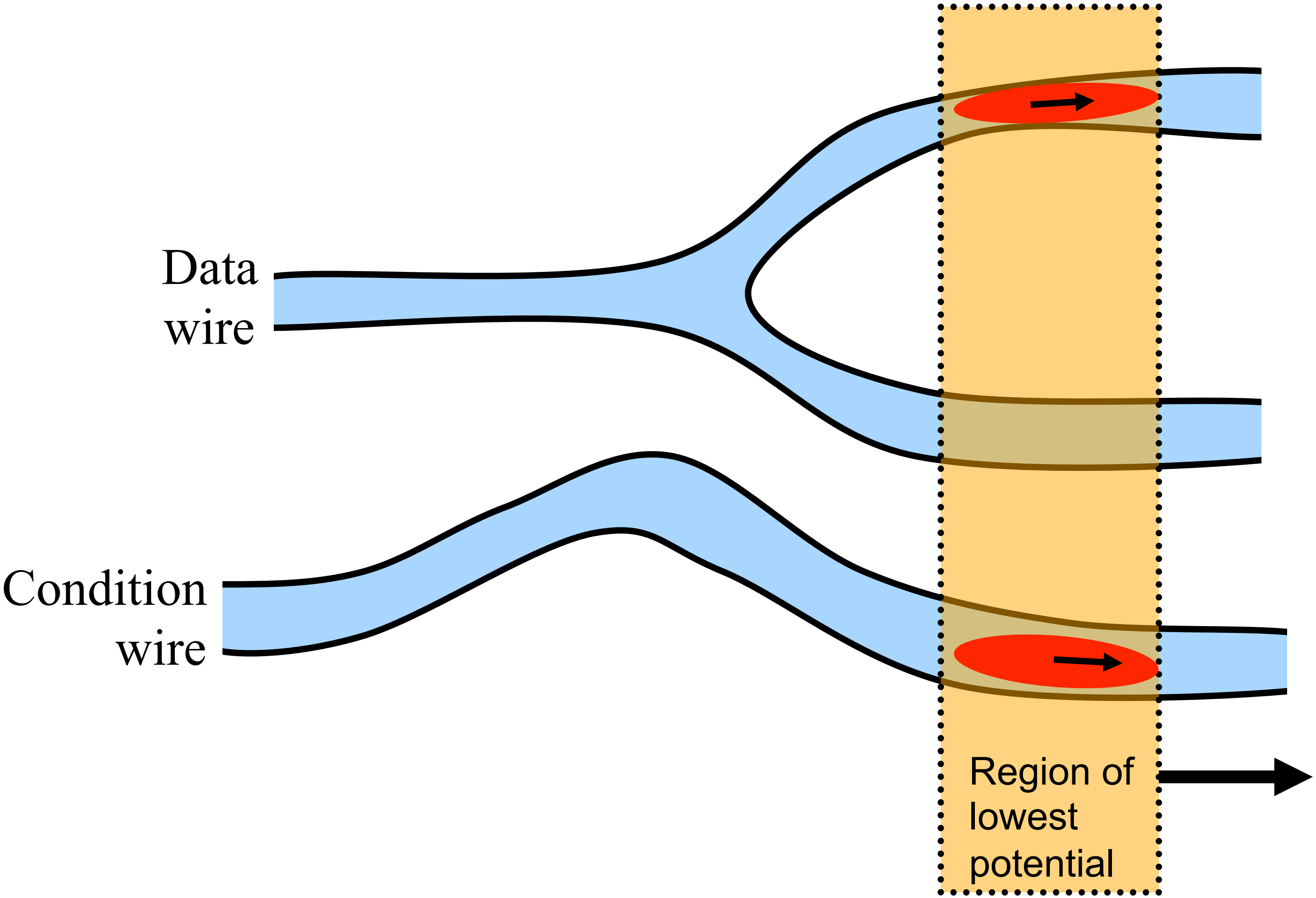
# Switch gate operation: 1 of 3



# Switch gate operation: 2 of 3



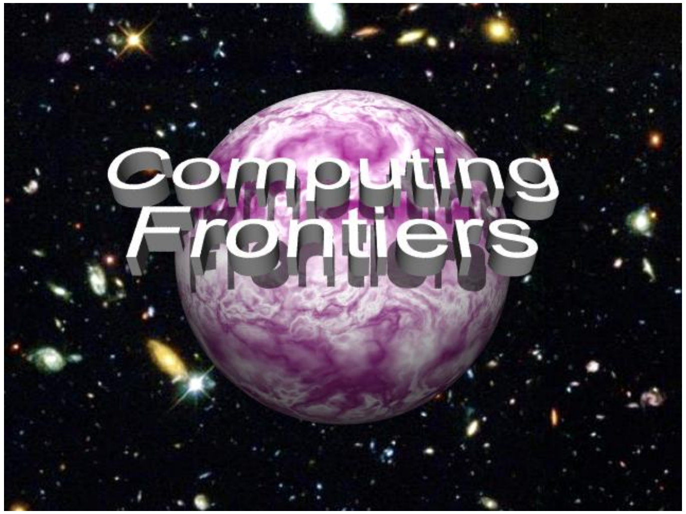
# Switch gate operation: 3 of 3



# Some Open Questions

- Here are some valid questions to ask about RC:
  - How exactly to design a resonant element w. high effective $Q$ to drive & synchronize reversible logic transitions?
    - while also avoiding undesired data-dependent back-action of the logic on the resonator
      - "dirtying" the state of the resonator → costs energy to correct
  - How to build a cheap reversible device with a very low adiabatic energy coefficient $c_E = E_{diss}/f_{op}$?
    - Low energy dissipation per op, at high frequency…
      - Requires low device $R$, and/or low $C$'s and high $V$'s.
    - Many device ideas are presently being explored for this…
  - How to optimize the logical architecture of reversible circuits and algorithms for best system-level cost-performance?
    - Great progress on this has already been made by numerous computer science theorists…
- We should view all of these as engineering problems to be solved, not as reasons to give up on reversible computing!

# The 1st International Workshop on Reversible Computing (RC' 05)




- A special session in the ACM *Computing Frontiers* conference (CF' 05).
  - To be held in Ischia, Italy, May 4-6, 2005.
- Speakers include:
  - Averin, Bennett, DeBenedictis, Forsberg, Frank, Fredkin, Frost, Semenov, Toffoli, Vitanyi, Williams… (& others)
- Handouts about the workshop are available here...
  - Attendees & sponsors are sought.
- Workshop website:
  - http://www.eng.fsu.edu/~mpf/CF05/RC05.htm