Alternative Memories: Non-Volatile Memories: Flash

Single Transistor Flash Memory Cell
What is DRAM?

- (Relatively) fast reads and (relatively) fast writes
- Unlimited number of writes
- Volatile - loses data storage without power
- Dynamic - loses data without periodic refresh
- Could be fabricated using similar materials and (relatively) similar silicon based process technologies as leading edge processors
Advantage/disadvantage of Alternatives:

- (Relatively) slower reads and (on some) really slow writes
- (Some) limited number of writes
- Non-Volatile - keeps data storage without power
- May require new materials and (relatively) different process technologies as leading edge processors
Flash: Basic Idea I

- Electrical charges are forced to tunnel through oxides and trapped in the floating gate.
- High voltage forces tunneling
- Trapped charges in floating gate then alters $V_t$
- Differences in $V_t$ of transistor then sensed as 0/1
- Limited number of write cycles
Charge Pump

"N" is the number of stages

Idealized charge pumping circuit
- Build up larger voltage for programming floating gate

V_{out} = (V_{dd} - V_t) \times N + V_{dd}
Flash: Basic Idea II

- Explicit program and erase cycles
Granularity

Write/Erase entire device

Bit by bit write/erase

Block write/erase
(Flash devices)
NOR Flash Array

Word Line

Word Line

Word Line

Word Line

Bit Line

Source Line
NAND versus NOR

**NOR**
- Better E/W Endurance (>100K vs >10K)
- Fast Read (~100ns)
- Slow Write (~10 us)
- Used for Code

**NAND**
- Smaller Cell Size (~40%)
- Slow Read (~1 us)
- “Fast” Write (~1 us)
- Used for Data
Multi (voltage) Level Cell

- Control Gate
- Floating Gate
- Tunnel Oxide
- Word Line
- Bit Line
- Drain
- Source
- Program
- Erase

- $V_{\text{ref}_0}$
- $V_{\text{ref}_1}$
- $V_{\text{ref}_2}$

- logic 00 range
- logic 01 range
- logic 10 range
- logic 11 range
Reads and Writes

- Reads are relatively straightforward
- Writes are complex
  - How long do we hold the reverse bias currents to “erase”?
  - Did the cells erase properly?
  - Did the write succeed?
    - If the write failed, recover, remap and re-write to another sector/block
Redundancy

16 spare

512 columns

Sense Amps

Mux
Comparison to DRAM

- Row/column/sense amp structure and access sequence is similar
- Controller logic/sequence has to be much more sophisticated
- Multiple banks per chip so “simultaneous” R/W can be performed.