Flash Memory

ECE 3080
Professor Doolittle

Nathan Fu

Memory: A Brief Overview

Two types:

- Volatile
  - **R**andom **A**ccess **M**emory (RAM) : faster, data lost when power switched off
- Nonvolatile
  - **R**ead-**O**nly **M**emory (ROM) : solid-state, retain information even when power switched off

Memory: A Brief Overview

Read-Only Memory categories:

- ROM : permanent data written in memory array
- PROM : user programmable data written to memory (**P**rogrammable **ROM**)
- EPROM : electrically programmed, erased via UV exposure (**E**rasable **PROM**)
- EEPROM : electrically erasable and programmable in-system (**E**lectrically **EPROM**)

Flash Memory Introduction

- A specific type of EEPROM with the ability to erase whole blocks (as opposed to bytes) in one erase cycle
- Used when larger amounts of data need to be stored without power
- Some things that use flash memory:
  - Computer BIOS
  - Memory cards
  - Solid-state disks (SSDs)

MOS Memory Cells

Common solutions to store charge:

- In “traps” in the insulator or at the interface between two dielectrics (usually silicon oxide/nitride). These devices are called MNOS (**M**etal-**N**itride-**O**xide-**S**ilicon) cells.
- In a conductive material layer surrounded by an insulator between the gate and the channel. These devices are called “floating gate” cells.

MNOS Devices

- Big breakthrough: SNOS (**S**ilicon-**N**itride-**O**xide-**S**ilicon)
  - Based on the use of LPCVD (**L**ow-**P**ressure **C**hemical **V**apor **D**eposition)
  - Inherent radiation hardness, able to adjust to application (slow programming-high retention or fast programming-low retention)
- Modern counterpart: SONOS
  - Higher memory window
  - Low power usage
  - High lifetime
- MNOS devices are hardly used nowadays; used in specific applications

Floating Gate Technology

- Stacked gate MOS transistor
  - First gate is *floating gate*
  - Second gate is *control gate* – acts as external gate of transistor
- Classes of FGs based on programming mechanisms

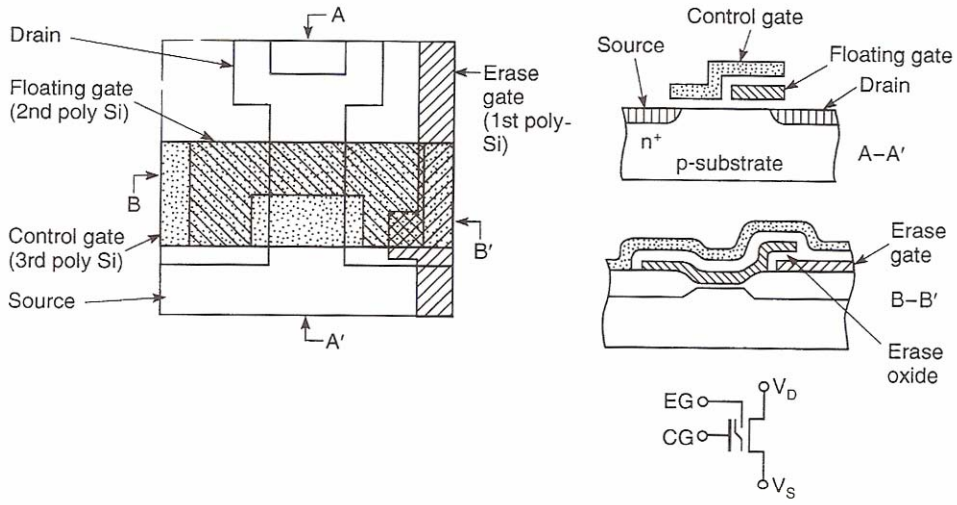


Figure 4.5 Top and cross-sectional view of the Flash EEPROM cell [4.3].

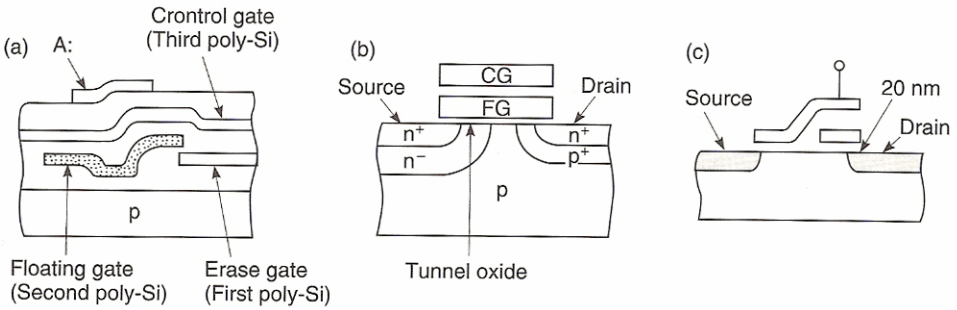


Figure 4.6 Stacked gate Flash EEPROM cell variations: (*a*) [4.3], (*b*) [4.7], and (*c*) [4.10].

Image courtesy of Gill, Lai

Floating Gate Technology

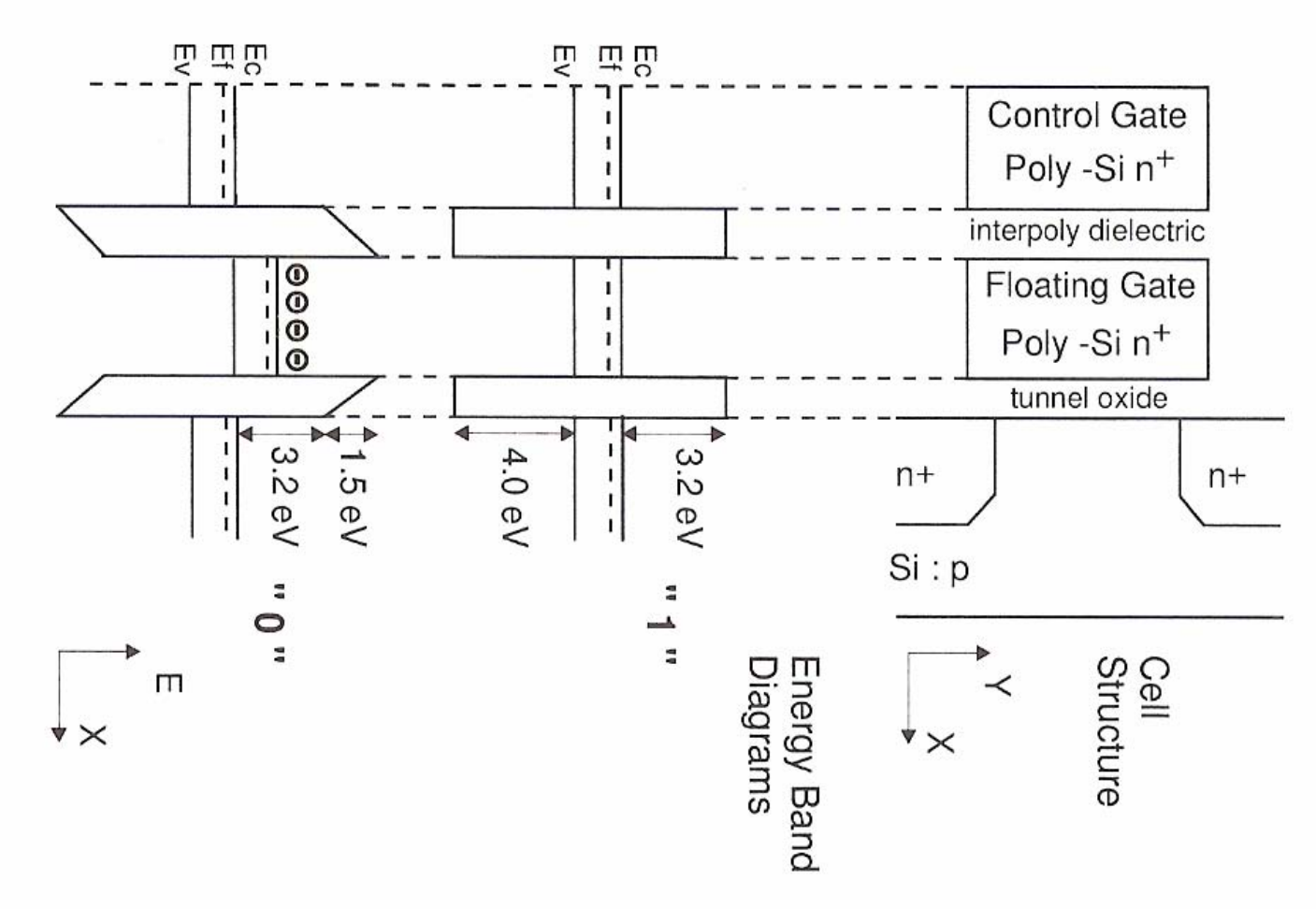
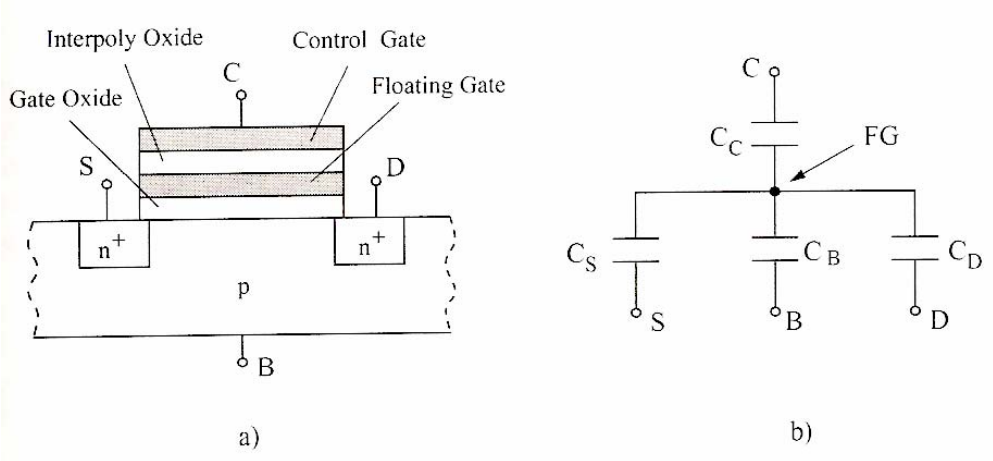


Image courtesy of Pavan, Bez

Floating Gate Technology

Modeling the control gate, source, bulk, and drains as capacitances, FG potential, V_F is:



$$V_{FS} = \frac{C_C}{C_T}V_C + \frac{C_S}{C_T}V_S + \frac{C_D}{C_T}V_D + \frac{C_B}{C_T}V_B + \frac{Q}{C_T}$$

Where V_C, V_S, V_D, V_B are control gate, source, drain, bulk potentials respectively;
Q is the charge contained within the FG;
Total capacitance is $C_T = C_C + C_S + C_D + C_B$

Image courtesy of Olivo, Zanoni

Floating Gate Technology

If the source and bulk are both grounded and all potentials use the source as reference, equation becomes:

$$V_{FS} = \frac{C_C}{C_T} V_{CS} + \frac{C_D}{C_T} V_{DS} + \frac{Q}{C_T}$$

Further simplification can be made if the "coupling factor" and f are defined such that:

$$\alpha_C = \frac{C_C}{C_T} \qquad f = \frac{C_D}{C_C} \qquad \text{and} \qquad V_{FS} = \alpha_C \left(V_{CS} + f V_{DS} + \frac{Q}{C_C} \right)$$

Floating Gate Technology

Since FG devices depend on threshold voltage (potential, $V_{T_{FS}}$, that must be applied to the FG with $V_{DS} = 0$), we can rearrange the equation again to:

$$V_{T_{CS}} = \frac{1}{\alpha_C} V_{T_{FS}} - \frac{Q}{C_C}$$

$V_{T_{FS}}$ depends on the device technology and $V_{T_{CS}}$ varies with the charge within the FG. This is key to the understanding of the FG as a basic nonvolatile memory cell.

Floating Gate Technology

By choosing a "threshold shift" ($|Q/C_C|$), we can define two different device states: *erased* and *programmed*

Threshold voltages applied to control gate are:

$$V_{T_{CS}} = \frac{1}{\alpha_C} V_{T_{FS}} = V_{T_E}$$

$$V_{T_{CS}} = \frac{1}{\alpha_C} V_{T_{FS}} - \frac{Q}{C_C} = V_{T_P}$$

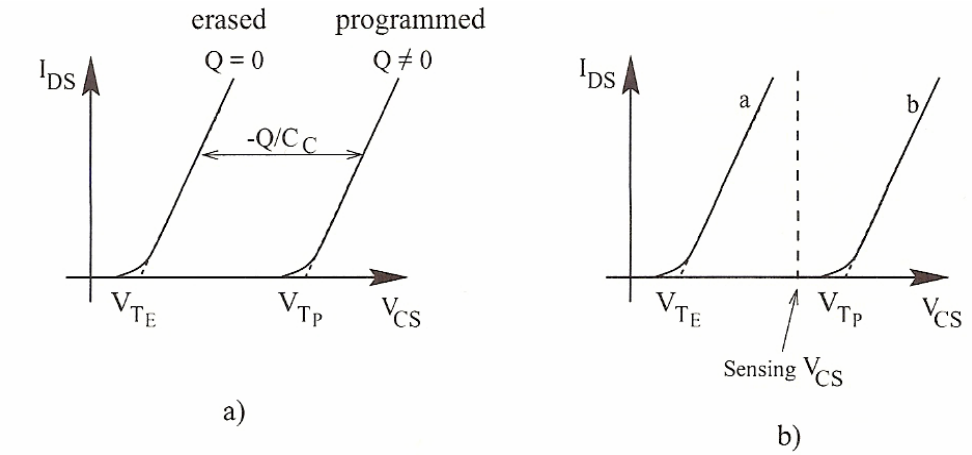


Figure 1.2 a) I-V trans-characteristics of a FG device for two different values of charge stored within the FG ($Q = 0$, and $Q < 0$), denoting two different states, respectively: erased and programmed; b) reading operation of a FG device: a suitable control gate voltage ($V_{T_E} < V_{CS} < V_{T_P}$) is applied to the device to determine whether it is conductive or not.

Device state is read by applying "sensing" voltage to control gate

- When FG I-V curve corresponds to *a*, $V_{CS} > V_{T_E}$ and device is ON
- When FG I-V curve corresponds to *b*, $V_{CS} < V_{T_P}$ and device is OFF

Image courtesy of Olivo, Zanoni

Charge Injection

- **C**hannel **H**ot-**E**lectron Programming (CHE)
- Source-Side Hot-Electron Programming
- Fowler-Nordheim Tunneling
  - Tunnel Programming
  - Tunnel Erase Through Thin Oxide

Channel Hot-Electron (CHE)

Hot Carriers : holes/electrons that have gained very high KE via a strong electric field

- High voltage is applied to drain and gate, causing high channel current and channel field to generate hot electrons
- High voltage on CG couples a voltage to the FG, attracting hot electrons to the FG

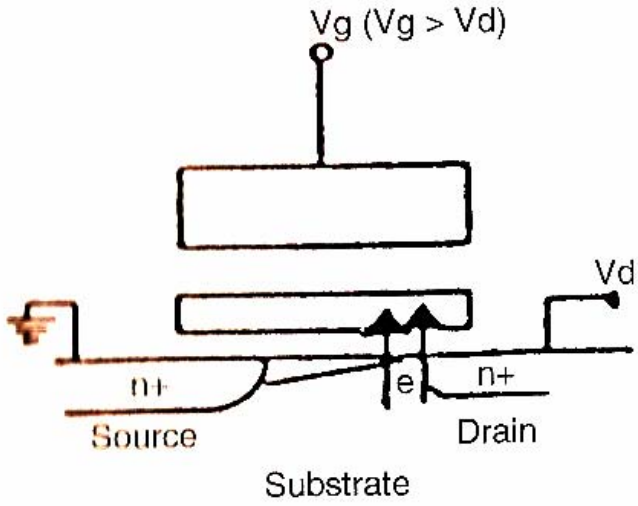


Figure 4.9 Schematic description of Flash EEPROM programming by CHE injection [4.14].

Image courtesy of Gill, Lai

Fowler-Nordheim Tunneling

- Basis of FLOTOX (**FLO**ating gate **T**hin **OX**ide)
- Writing – High voltage is applied to CG with drain at low bias. CG voltage couples a voltage on the FG, allowing electrons to tunnel from the drain to the FG through a thin oxide (~8-10 nm)
- Erasing – Drain raised to high voltage, CG grounded, causing electrons to tunnel back into the drain

FLOTOX

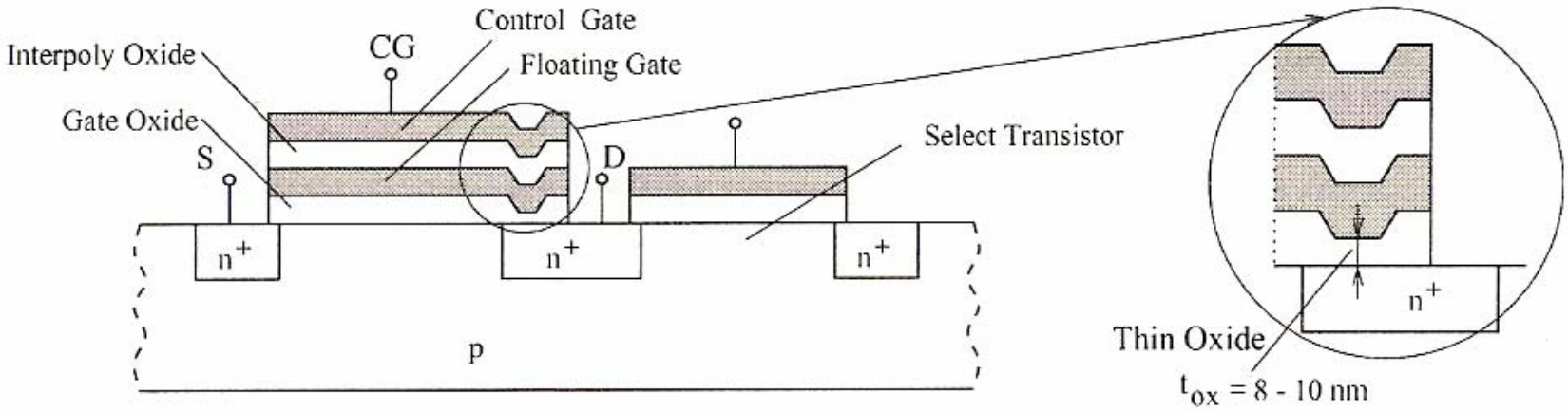


Figure 1.3 Schematic section of a FLOTOX cell including the select transistor.

- Select transistor controls drain bias

Image courtesy of Olivo, Zanoni

Flash Cell Operations

Three stages:

- Programming
  - CHE
  - F-N tunneling
- Reading
  - FG I-V curve
- Erasing
  - F-N tunneling

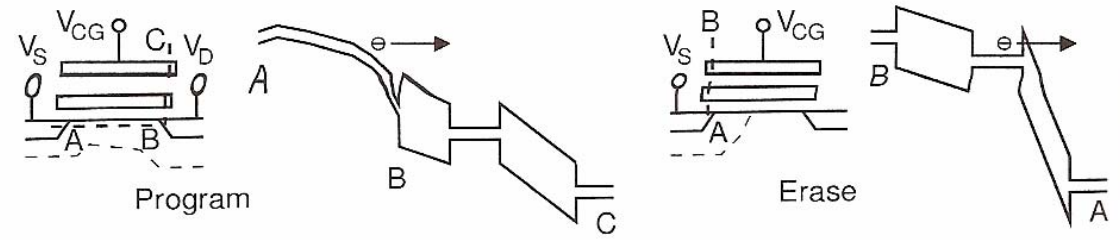


Figure 4.1 Schematic representation of programming and erasing mechanisms for the industry standard Flash cell.

Image courtesy of Pavan, Bez

Erasing

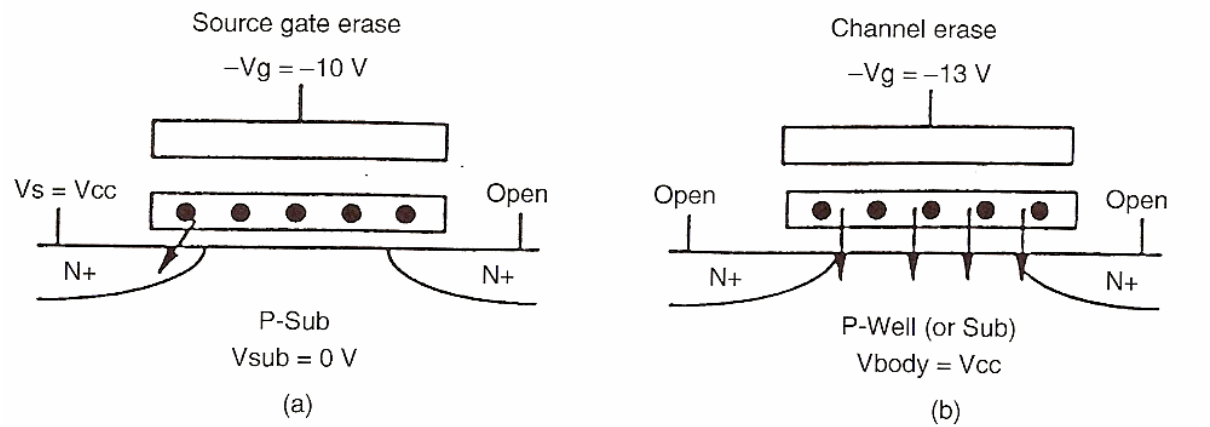


Figure 4.13 Cross-sectional drawing of (*a*) negative gate, floating gate to source erase, (*b*) negative gate, floating gate to channel erase [4.104].

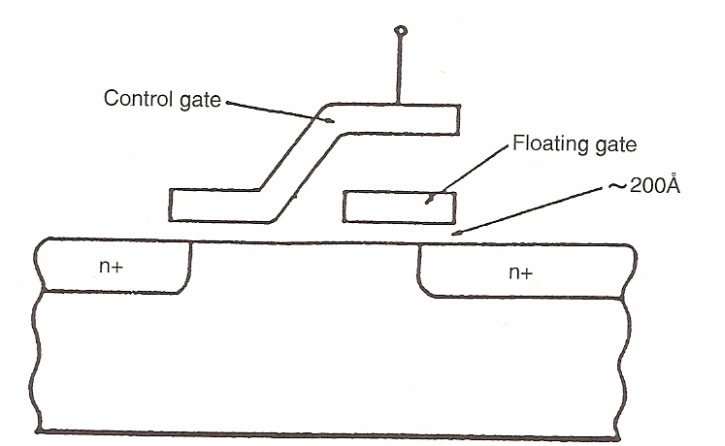


Figure 4.12 Drain erase Flash memory cell [4.10, 4.11, 4.87].

Images courtesy of Olivo, Zanoni

Flash Memory Architecture

Commonly used today:

- NAND
- NOR (Common Ground)

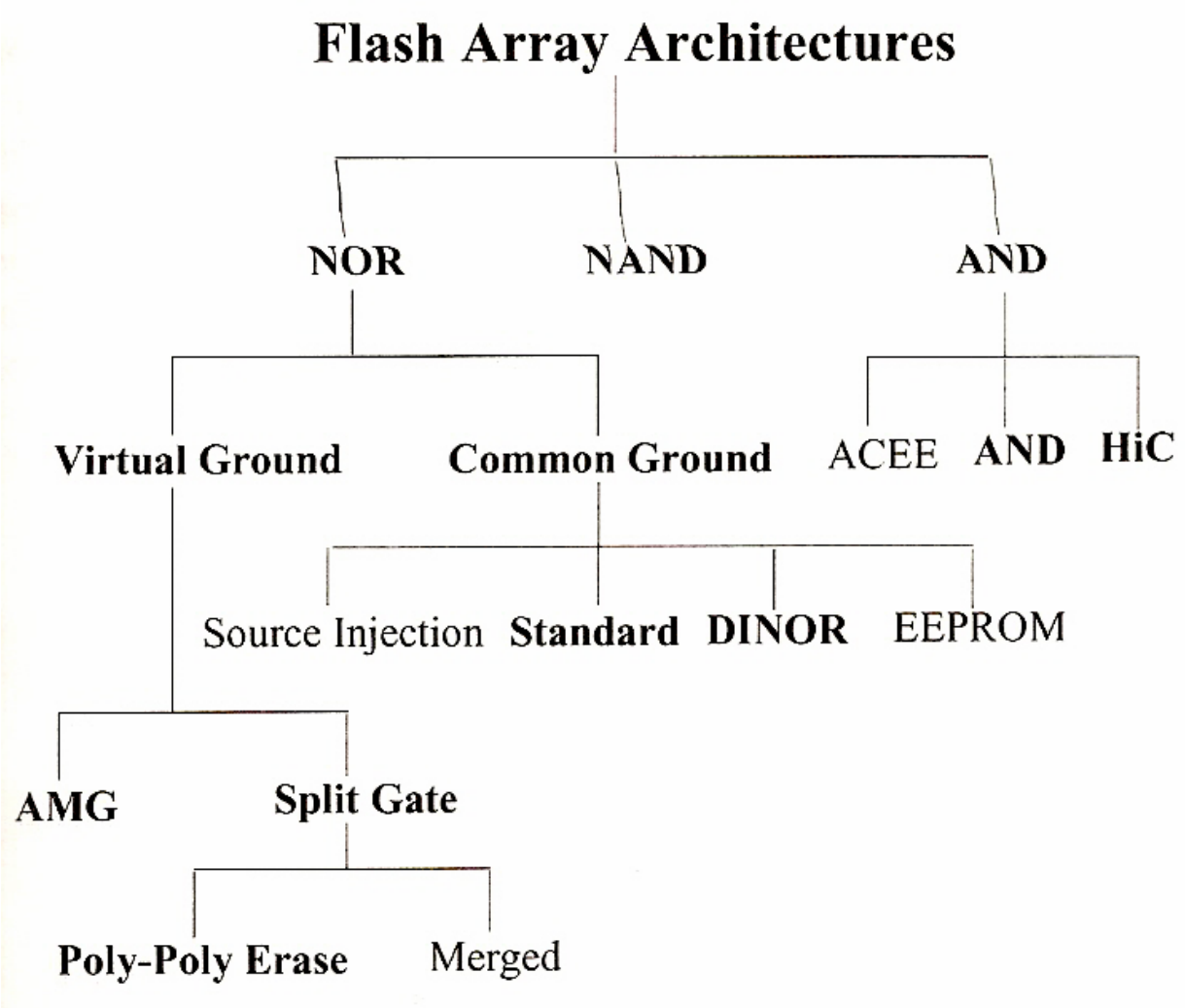


Image courtesy of Olivo, Zanoni

NAND vs. NOR

- NAND
  - High density
  - Medium read speed
  - High write, erase speed
  - Indirect access (I/O-like)
  - Less power consumption for write-intensive apps

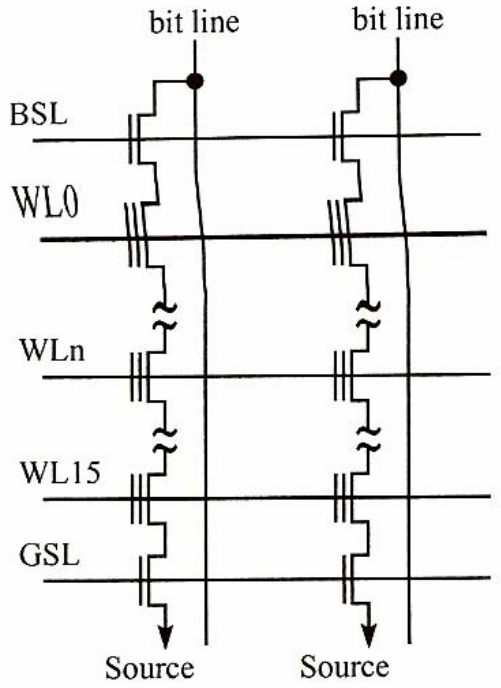


Figure 1.13 Basic structure of a NAND architecture.

Image courtesy of Pavan, Bez

NAND vs. NOR

- NOR
  - Bootable
  - Executable
  - Low read latency

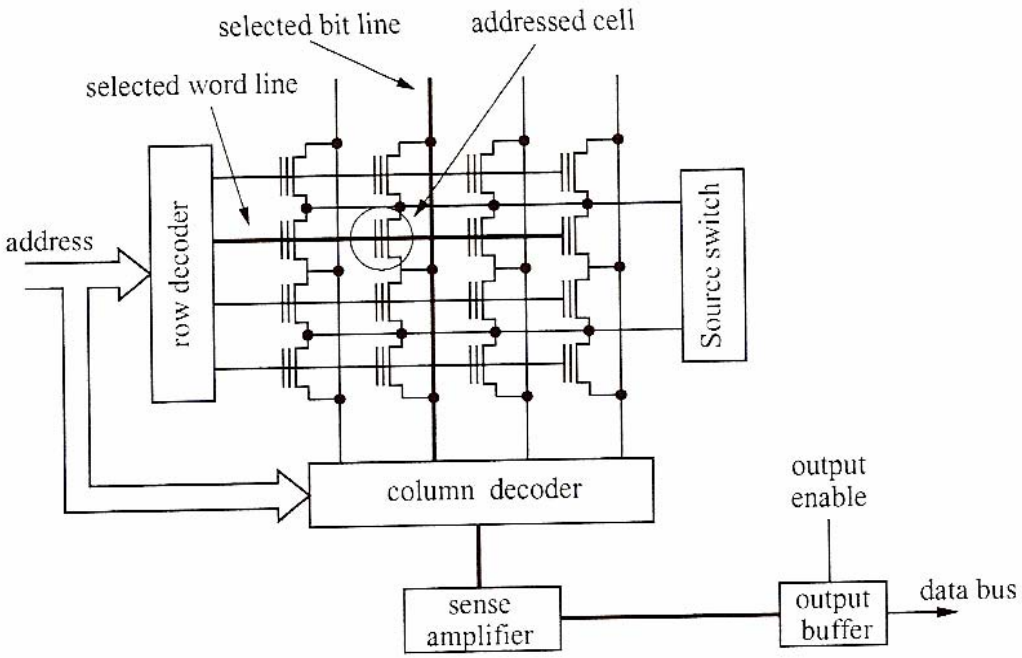


Figure 1.9 Schematic structure of the read path in a NOR organization. Only one bit at a time is here considered as addressable.

Image courtesy of Pavan, Bez

NAND vs. NOR

	SLC NAND Flash (x8)	MLC NAND Flash (x8)	MLC NOR Flash (x16)
Density	512 Mbits[1] – 4 Gbits[2]	1Gbit to 16Gbit	16Mbit to 1Gbit
Read Speed	24 MB/s[3]	18.6 MB/s	103MB/s
Write Speed	8.0 MB/s	2.4 MB/s	0.47 MB/s
Erase Time	2.0 mSec	2.0mSec	900mSec
Interface	I/O – indirect access	I/O – indirect access	Random access
Application	Program/Data mass storage	Program/Data mass storage	eXecuteInPlace

Figure 4: NAND and NOR Flash Operating Specifications

Taken from a 2006 data sheet

Image courtesy of Toshiba

The Future

- Mainstream Solid-State Disks?
  - A number of companies (ASUS, Apple Inc., Lenovo) are starting to use SSDs as main storage units, replacing the common hard disk drive
    - SSDs do not contain moving parts, and therefore, are “shock” resistant
    - SSDs require less power (up to 50%) and generate less heat than a regular HDD
    - Can be made much smaller and lighter than a HDD
    - Price, which used to be a major factor, is going down for flash memory (1GB of flash memory can go for ~$10)
- PCM? (**P**hase-**C**hange **M**emory)
  - AKA PRAM, PCRAM