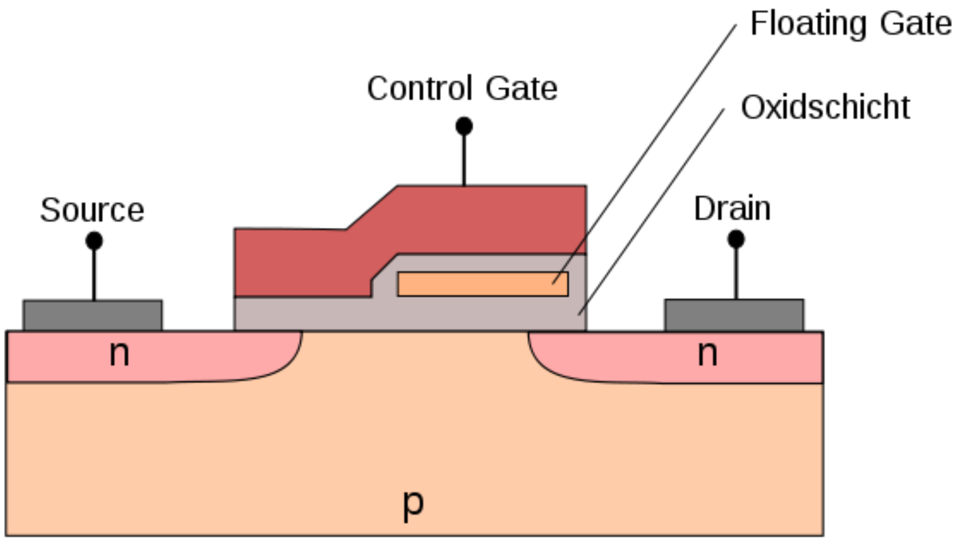
Scalus Winter School
Storage Systems

Flash Memory

André Brinkmann

Flash Memory

- Floating gate of a flash cell is electrically isolated
- Applying high voltages between source and drain accelerates electrons and improves probability that they float through isolator (in case of positive gate source voltage)
  - Fowler-Nordheim tunneling effect
  - Bits are persistently stored on the floating gate

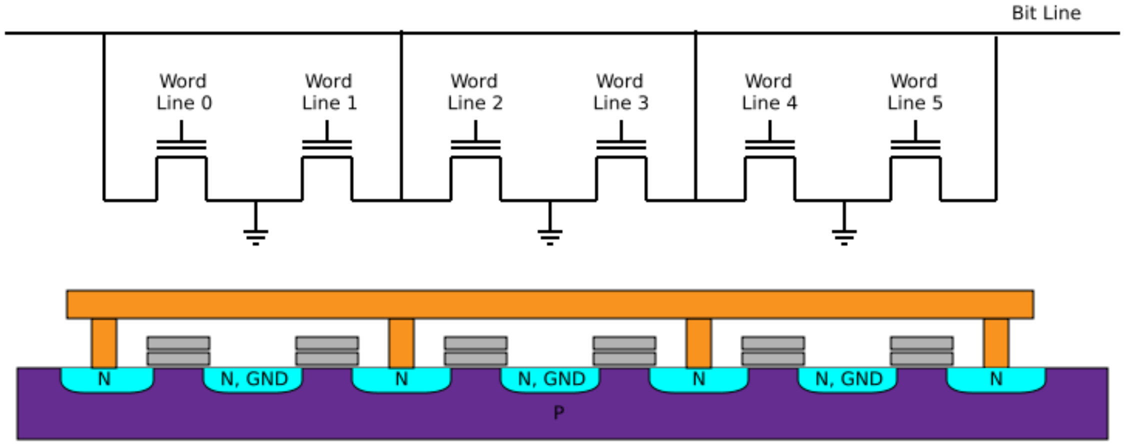


- It is only possible to write ‘0‘
- High negative voltage enables information erasure
- Erasure harms isolator
- max. of 1 Mio. erasure cycles

Slide based on www.wikipedia.de

NOR-Flash

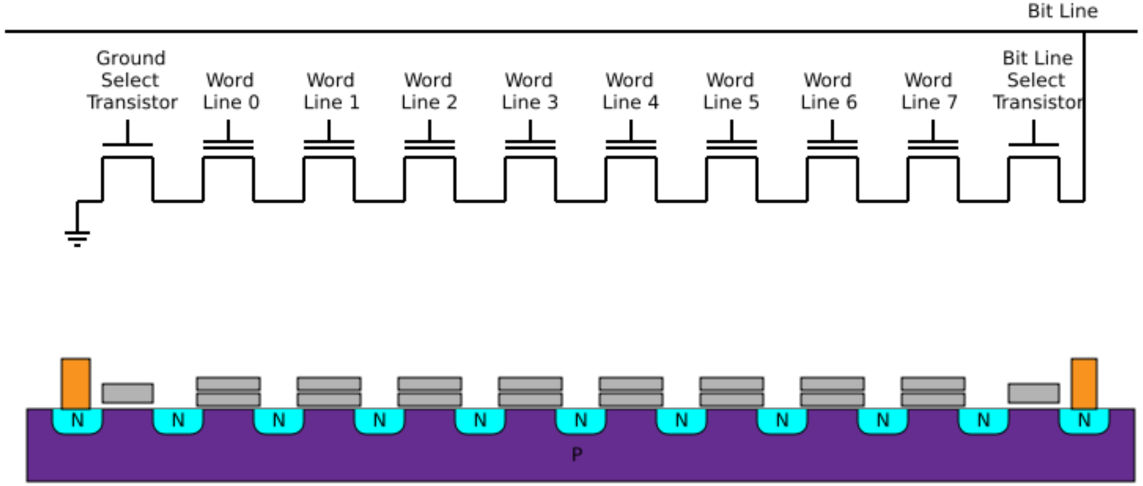
- Developed as a replacement for EPROMs and ROMs
- Addressing done via standard memory address bus
  - No “Glue Logic“ between controller and flash
  - Can be used to store code
- Bytes can be individually written (Transfer from ‘1‘ → ‘0‘)
- Only complete blocks can be erased (Reset all bits to ‘1‘)
- Block site between 64 Kbyte and 256 KByte
- Slow writing data
- Capacity smaller 1 GBit



Slide based on www.wikipedia.de

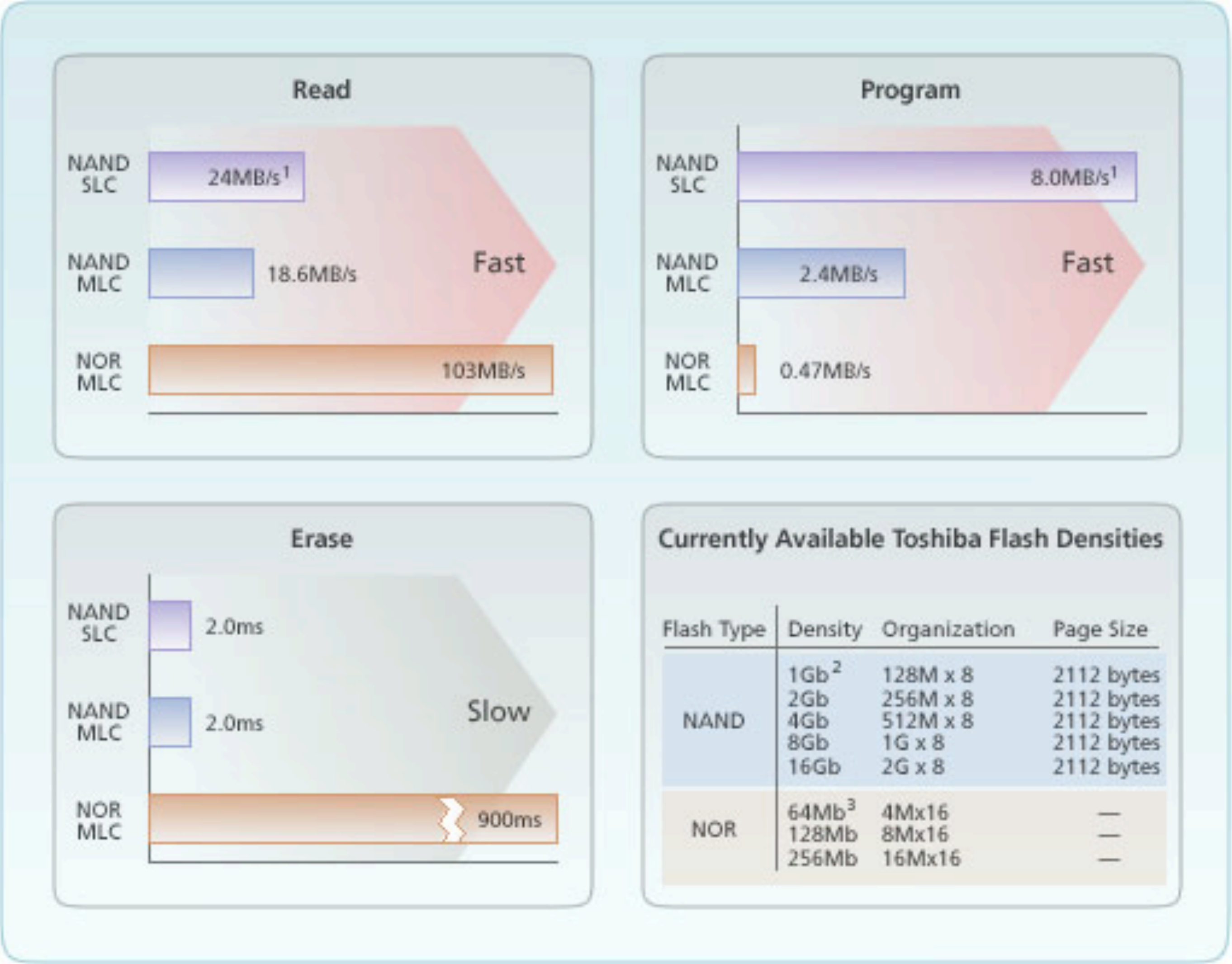
NAND-Flash

- NAND flash works page and block oriented
- Page has to be read sequentially and resembles NAND gates
- Page size between 512 Byte and 4 Kbyte
  - Additionally 12 to 16 bytes of checksum data
- Pages are combined into blocks
- Pages can only be written once without flashing the complete block
- Higher density than for NOR flash



Slide based on www.wikipedia.de

Comparison NOR/NAND



Currently Available Toshiba Flash Densities

Flash Type	Density	Organization	Page Size
NAND	1Gb[2]	128M x 8	2112 bytes
	2Gb	256M x 8	2112 bytes
	4Gb	512M x 8	2112 bytes
	8Gb	1G x 8	2112 bytes
	16Gb	2G x 8	2112 bytes
NOR	64Mb[3]	4Mx16	—
	128Mb	8Mx16	—
	256Mb	16Mx16	—

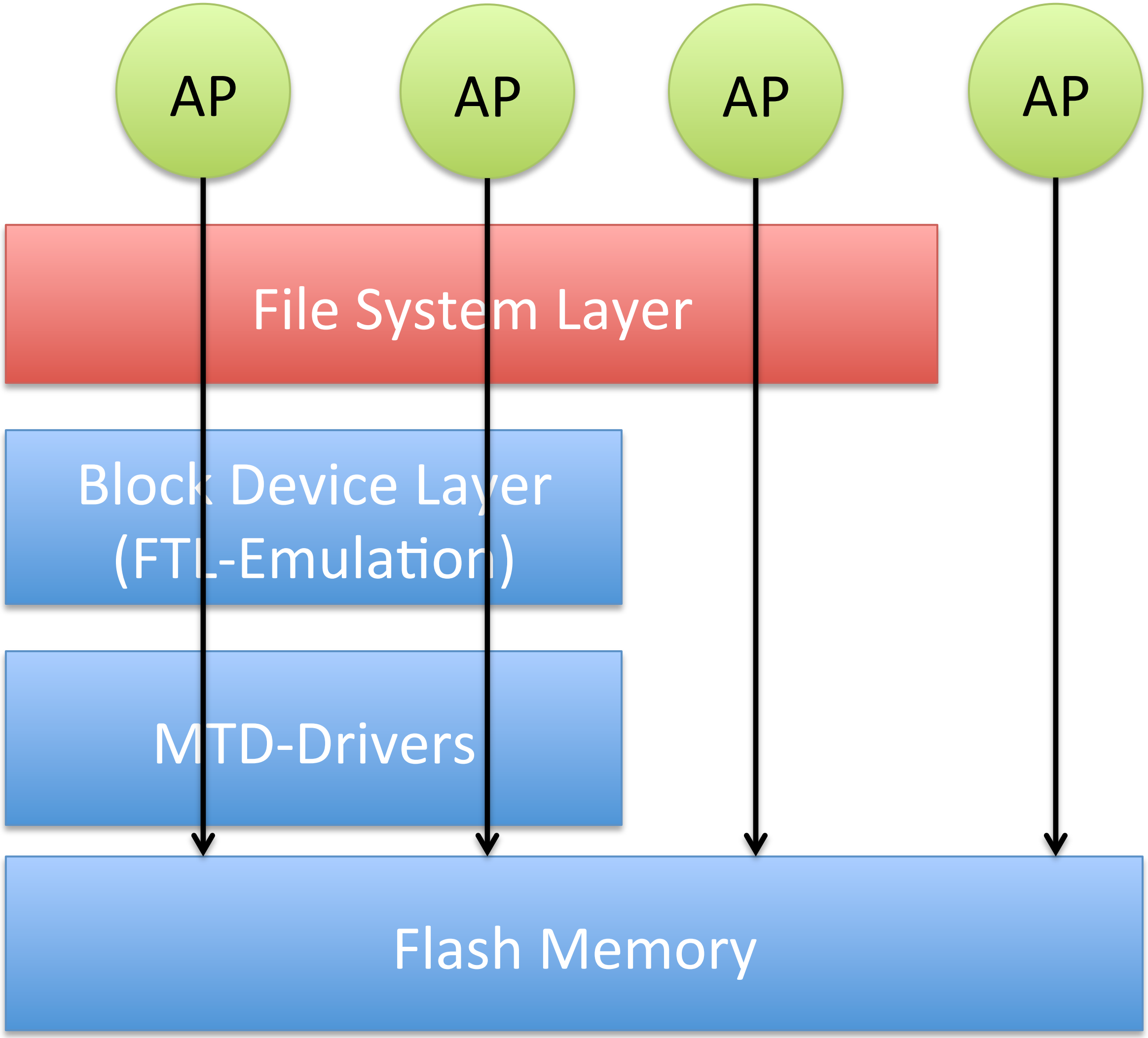
Slide based on a technology survey from Toshiba 2010

Challenges

- Good Performance
- Limited Cost per Unit
- Strong Demands in Reliability
- Access Frequencies
- Tight Coupling with Other Components
- Low Compatibility among Vendors

Slide based on lecture “Flash-Memory Storage Systems” by Tei-Wei Kuo

Issues – Architecture



FTL: Flash Translation Layer
MTD: Memory Technology Device

Slide based on lecture "Flash-Memory Storage Systems" by Tei-Wei Kuo

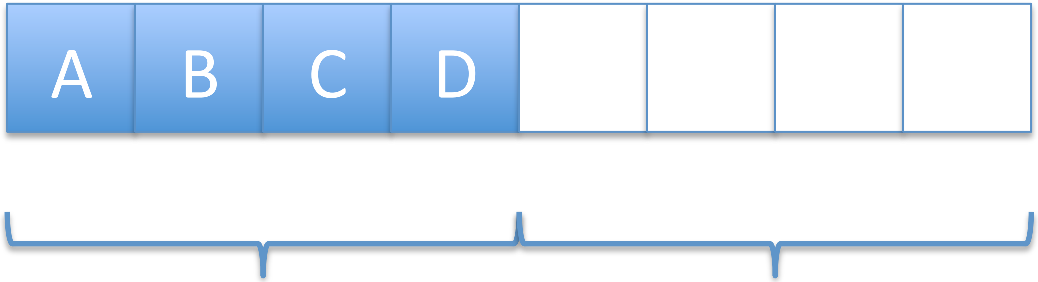
Issues: Flash Characteristics

- Write-Once
  - No writing on the same page unless its residing block is erased
  - Pages are classified into valid, invalid, and free pages
- Bulk-Erasing
  - Pages are erased in a block unit to recycle used but invalid pages
- Wear-Leveling
  - Each block has a limited lifetime in erasing counts

Slide based on lecture “Flash-Memory Storage Systems” by Tei-Wei Kuo

Issues: Flash Characteristics

- Example: „out-of-place update“



Live Pages

Free Pages

- Assume that we want to update pages A and B

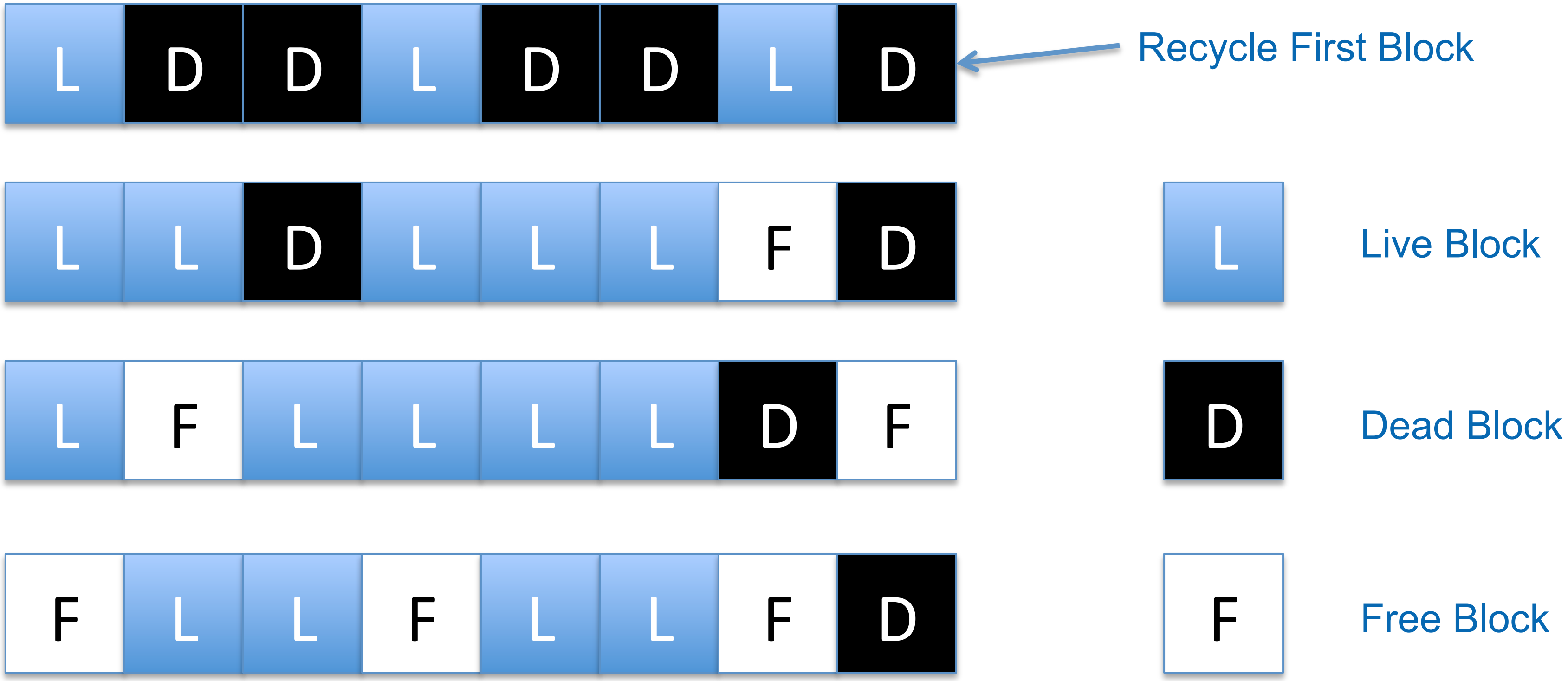
Issues: Flash Characteristics

- Example: „out-of-place update“

Dead Pages

Issues: Flash Characteristics

- Example 2: Garbage Collection



L	D	D	L	D	D	L	D
L	L	D	L	L	L	F	D
L	F	L	L	L	L	D	F
F	L	L	F	L	L	F	D

Recycle First Block

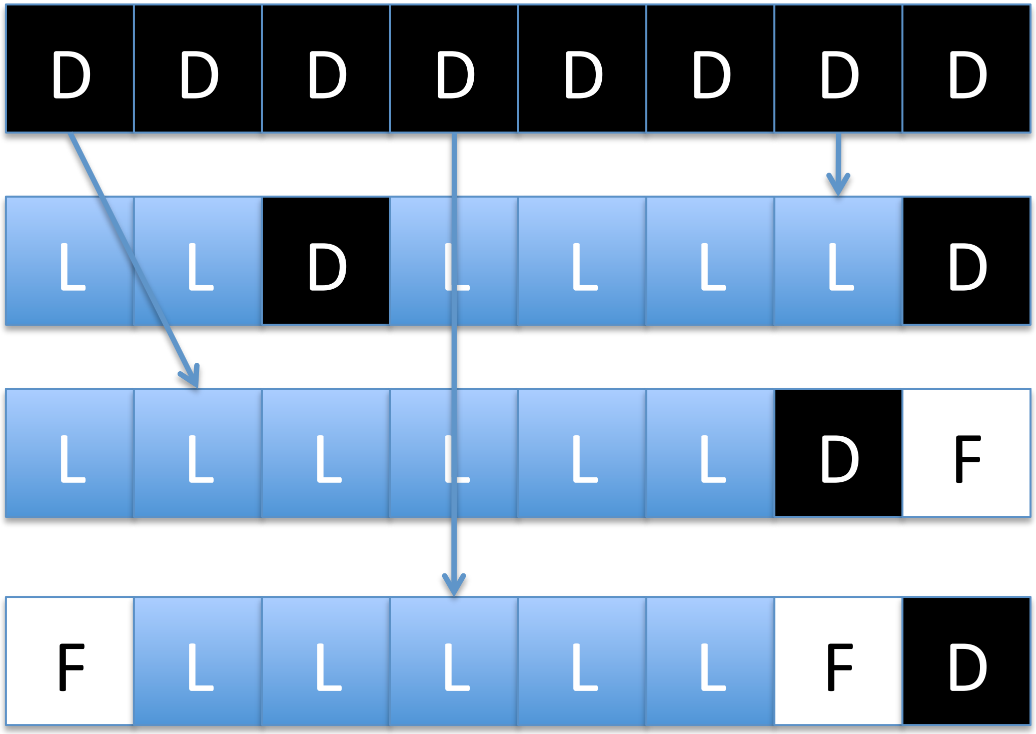
L Live Block

D Dead Block

F Free Block

Issues: Flash Characteristics

- Example 2: Garbage Collection



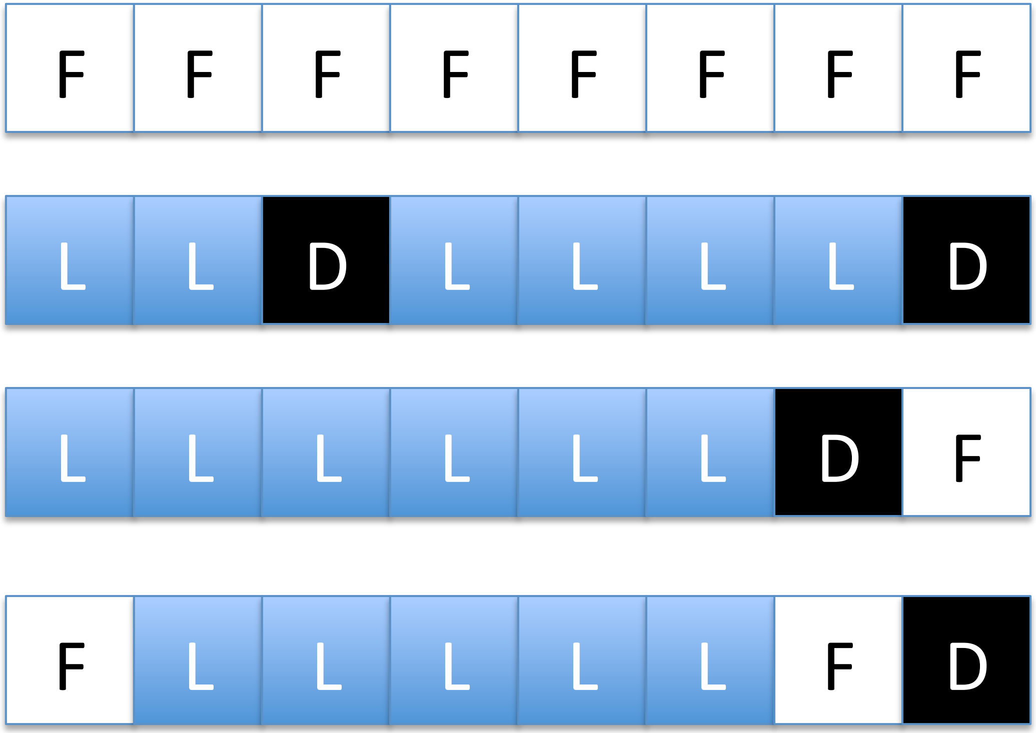
D	D	D	D	D	D	D	D
L	L	D	L	L	L	L	D
L	L	L	L	L	L	D	F
F	L	L	L	L	L	F	D

Copy Live Data somewhere else

Slide based on lecture "Flash-Memory Storage Systems" by Tei-Wei Kuo

Issues: Flash Characteristics

- Example 2: Garbage Collection



F	F	F	F	F	F	F	F
L	L	D	L	L	L	L	D
L	L	L	L	L	L	D	F
F	L	L	L	L	L	F	D

Erase Data Block

Overhead:
-Copying of Live Data
-Block Erasing

Issues: Flash Characteristics

- Example 3: Wear-Leveling

100	L	D	D	L	D	D	L	D
10	L	L	D	L	L	L	F	D
20	L	F	L	L	L	L	D	F
15	F	L	L	F	L	L	F	D

Erase Counter

Hot Data vs. Cold Data

- Hot data usually comes from
  - metadata of file-systems
    - Small. A piece of hot data is usually ≦ 2 sectors.
  - structured (or indexed) user files, etc.
- Cold data usually comes from
  - read-only (or WORM) files
    - E.g., bulk and sequential files that often have a number of sectors

Efficient Hot-Data Identification

Why Important?

- Wear-Leveling
  - Pages that contain hot data could turn into dead pages very quickly
  - Blocks with dead pages are usually chosen for erasing
  - **Hot data should be written to blocks with smaller erase counts**
- Erase Efficiency (i.e., effective free pages reclaimed from garbage collection.)
  - **Mixture of hot data and non-hot data in blocks might deteriorate the efficiency of erase operations**

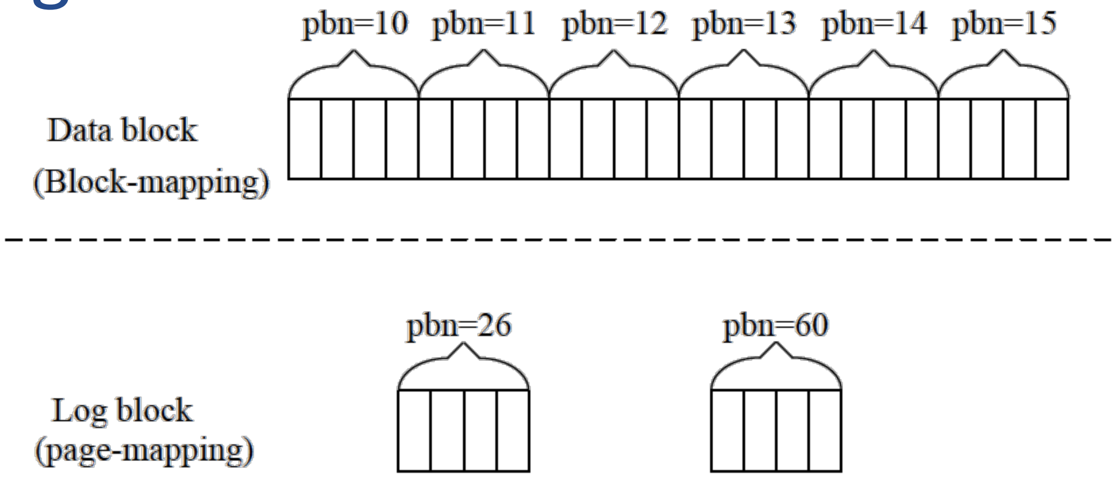
Flash Translation Layer (FTL)

- A software layer that allows the flash memory to look like a HDD
  - Address mapping: logical to physical
  - Garbage collection & power-off recovery
  - Wear-leveling & bad block management
- Popular FTL algorithms
  - FMAX, BAST, FAST, Super block, LAST
  - DFTL, DAC, etc...

Slide based on presentation "FASTerFTL for Enterprise-Class Flash Memory SSDs" by Lim, Lee und Moon

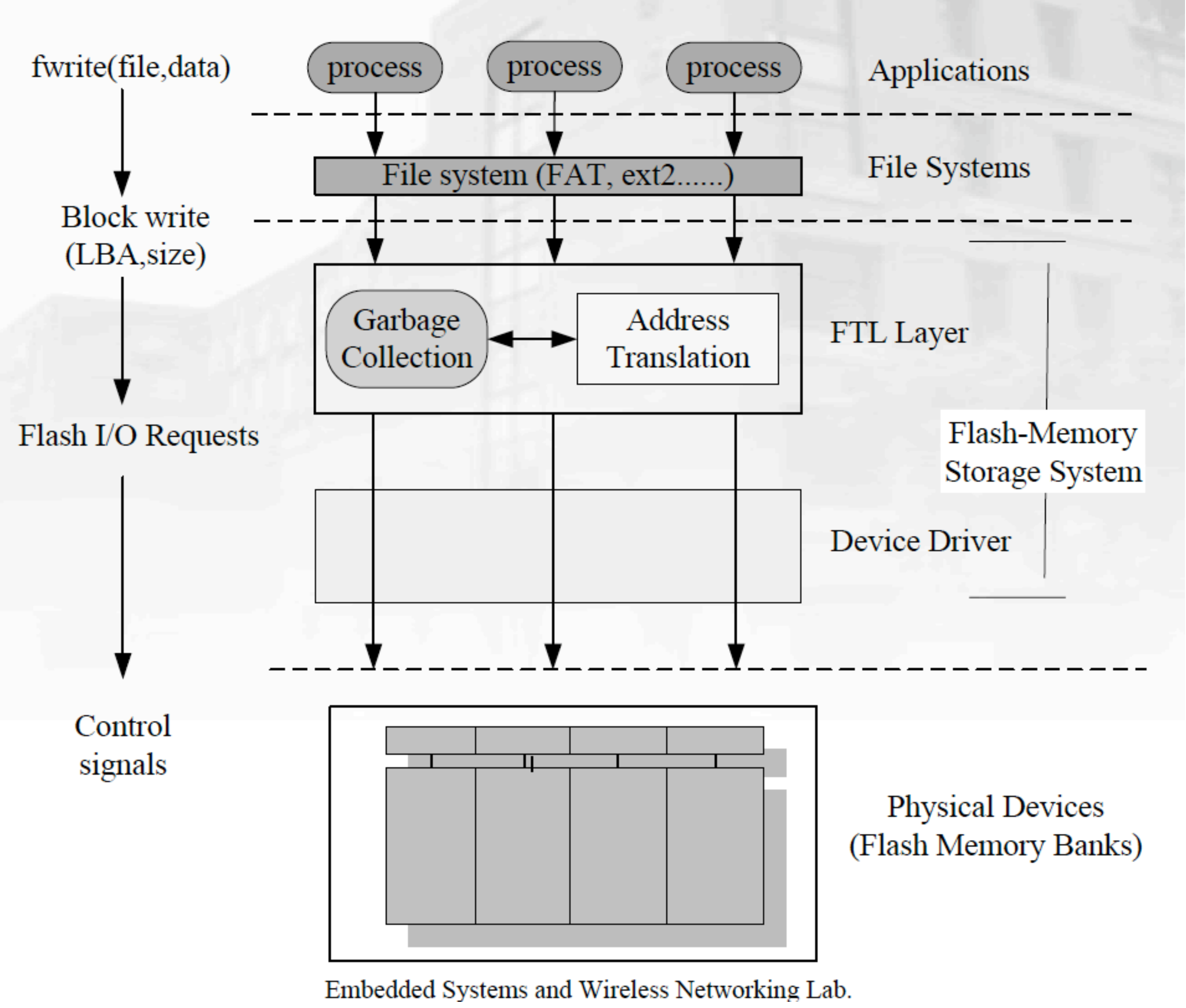
Flash Translation Layer (FTL)

- Page-mapping FTL can map a logical page to any physical page
- Block-mapping FTL can only map a logical page to a fixed offset of a block
- Hybrid FTL includes log blocks



Slide based on presentation “Achieving Page-Mapping FTL Performance at Block-Mapping FTL Cost by Hiding Address Translations” by Hu et. al

System Architecture



Embedded Systems and Wireless Networking Lab.

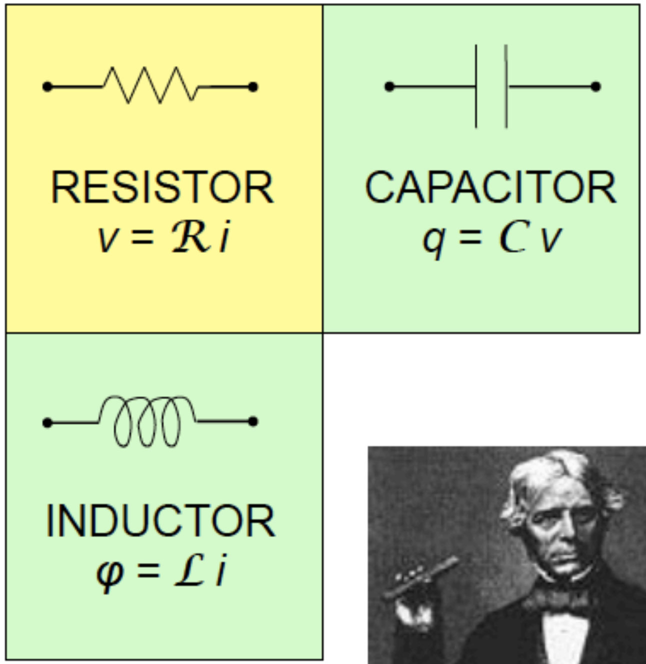
Observations

- Write throughput drops significantly after garbage collection starts
- The capacity of flash-memory storage systems increases very quickly such that also memory space requirements for address translation growth quickly
- Reliability becomes more and more critical when the manufacturing capacity increases
- Significant increment of flash-memory access numbers

Slide based on lecture “Flash-Memory Storage Systems” by Tei-Wei Kuo

3 passive linear circuit elements

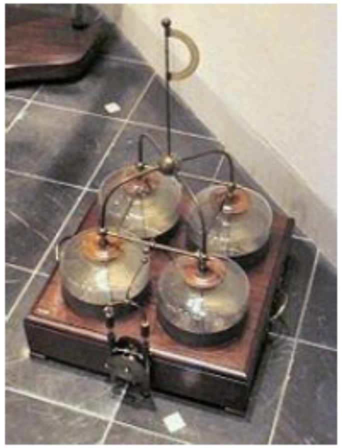
Resistor – 1827
Georg Ohm



RESISTOR
$v = \mathcal{R}\, i$

CAPACITOR
$q = C\, v$

INDUCTOR
$\varphi = \mathcal{L}\, i$



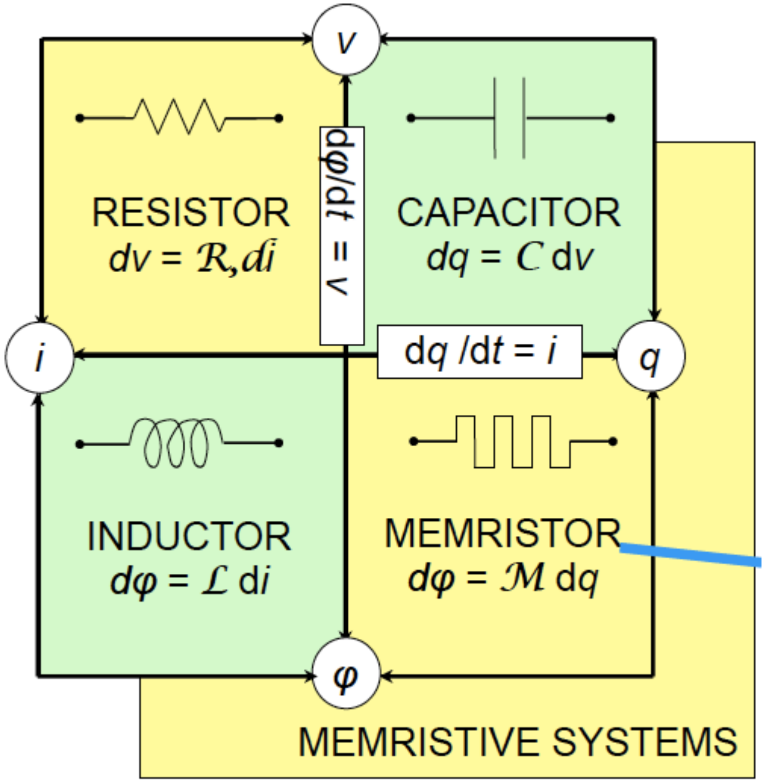
Capacitor - 1745
Volta / von Kleist & van Musschenbroek
Benjamin Franklin

Inductor – 1831
Michael Faraday
Joseph Henry

Chua generalizes circuit theory to nonlinear systems

- L. O. Chua, “Memristor - the missing circuit element,” IEEE Trans. Circuit Theory 18, 507–519 (1971)
- L. O. Chua and S. M. Kang, "Memristive devices and systems," *Proc. IEEE, 64 (2), 209-23 (1976)*
- Memristor "remembers" changes in the curren its resistance

- D. Strukov, G. Snider, D. Stewart & S. Williams: *The missing memristor found.* In: *Nature.* 453, 2008, S. 80-83



RESISTOR
$dv = \mathcal{R}\, di$

CAPACITOR
$dq = \mathcal{C}\, dv$

INDUCTOR
$d\varphi = \mathcal{L}\, di$

MEMRISTOR
$d\varphi = \mathcal{M}\, dq$

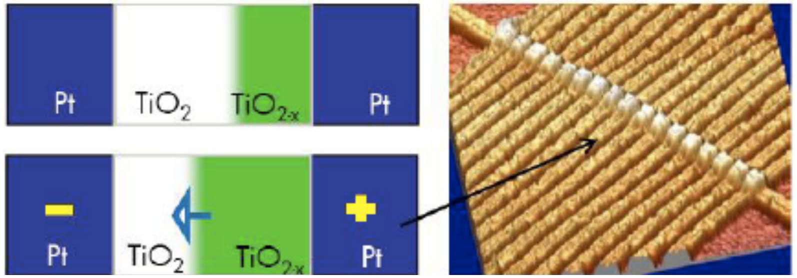
$d\varphi/dt = v$

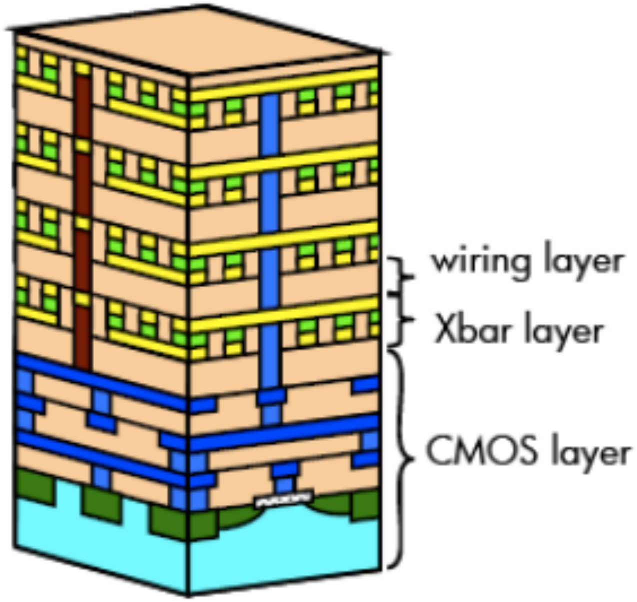
$dq/dt = i$

MEMRISTIVE SYSTEMS

Memristors

- Advantages
  - – Fast, cheap, dense, & low energy
  - – Can be fabbed with CMOS logic & 3D layers on single die
- Challenges
  - – Understand wear out
  - – Moving from the lab to the fab





Slide based on presentation “STORAGE RESEARCH@ HP Labs” by J. Wylie

Phase-change memory (PRAM)

- PRAM based on unique behavior of chalcogenide glass
- Each memory cell contains a material that has two phases with very different electrical properties
- An „amorphous phase“ exhibits high resistivity, while a „crystalline phase“ has much lower resistivity
- Reading the bit value stored in a cell consists of sensing its resistivity (a fast, low-power operation)

Slide based on the presentation “A Content-Aware Block Placement Algorithm for Reducing PRAM Storage Bit Writes” by B. Wongchaowart and Wikipedia

PRAM

- Individual PRAM cells can be programmed independently of other cells
- When overwriting data stored in PRAM, only cells whose current value differs from the new value to be written need to be updated
- Degrades much more slowly than Flash
- Challenges:
  - requirement of high programming current density
  - long-term resistance and threshold voltage drift

Slide based on the presentation "A Content-Aware Block Placement Algorithm for Reducing PRAM Storage Bit Writes" by B. Wongchaowart and Wikipedia

PRAM

- In order to change the bit value stored in a PRAM cell, the phase-change material must be brought into a different phase by heating
- Heating the phase-change material to its crystallization temperature for a sufficiently long period of time causes it to get into its crystalline state
- Heating it to a yet higher temperature for a short period of time makes the material amorphous
- Both of these operations require high-power current pulses (relative to the read operation)

Slide based on the presentation “A Content-Aware Block Placement Algorithm for Reducing PRAM Storage Bit Writes” by B. Wongchaowart and Wikipedia