



Faculteit Industriële  
Ingenieurswetenschappen



universiteit  
▶ hasselt      KU LEUVEN



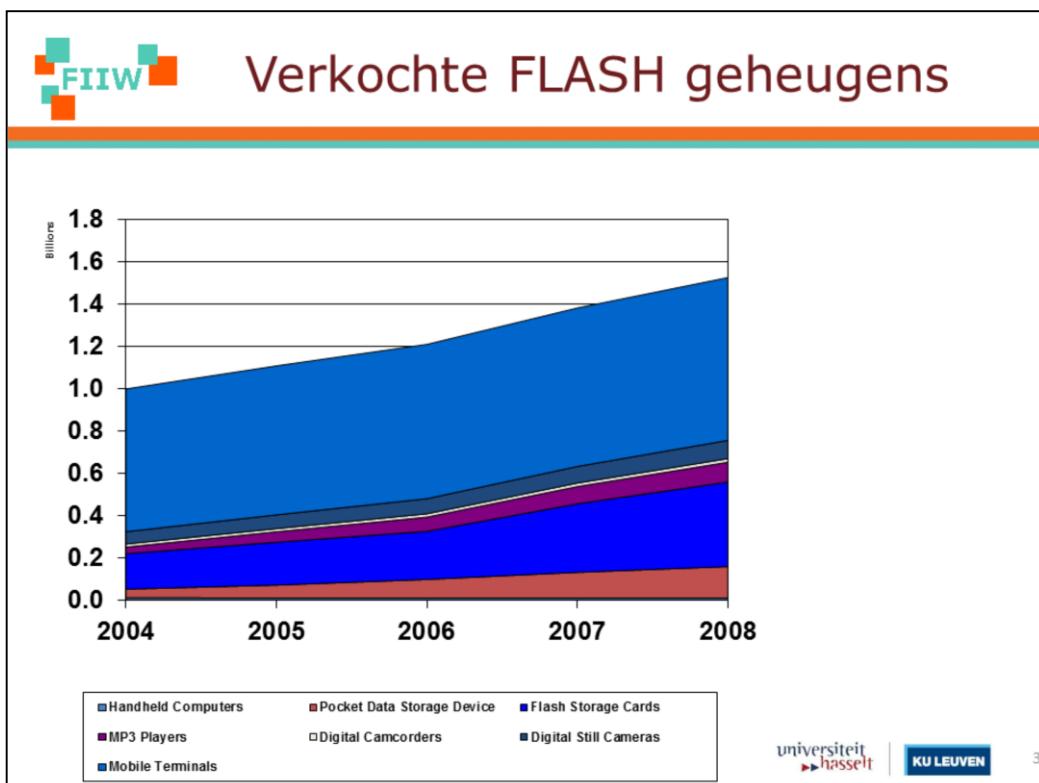
Jan Genoe  
jan.genoe@kuleuven.be

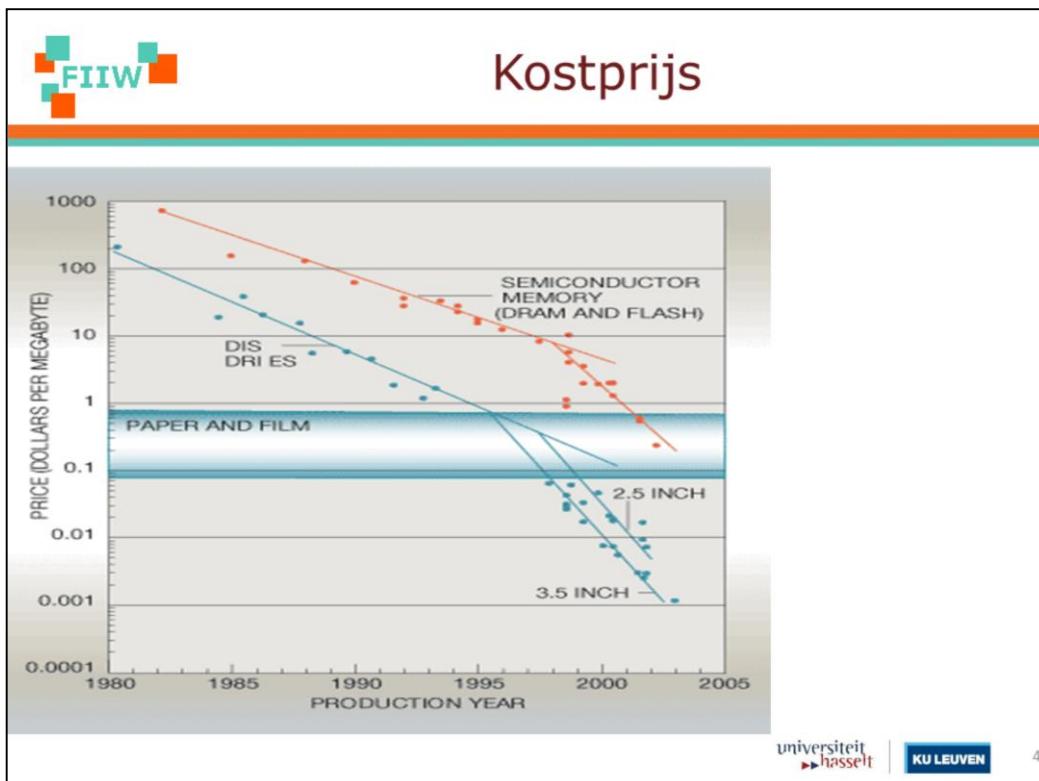
## Flash Memory

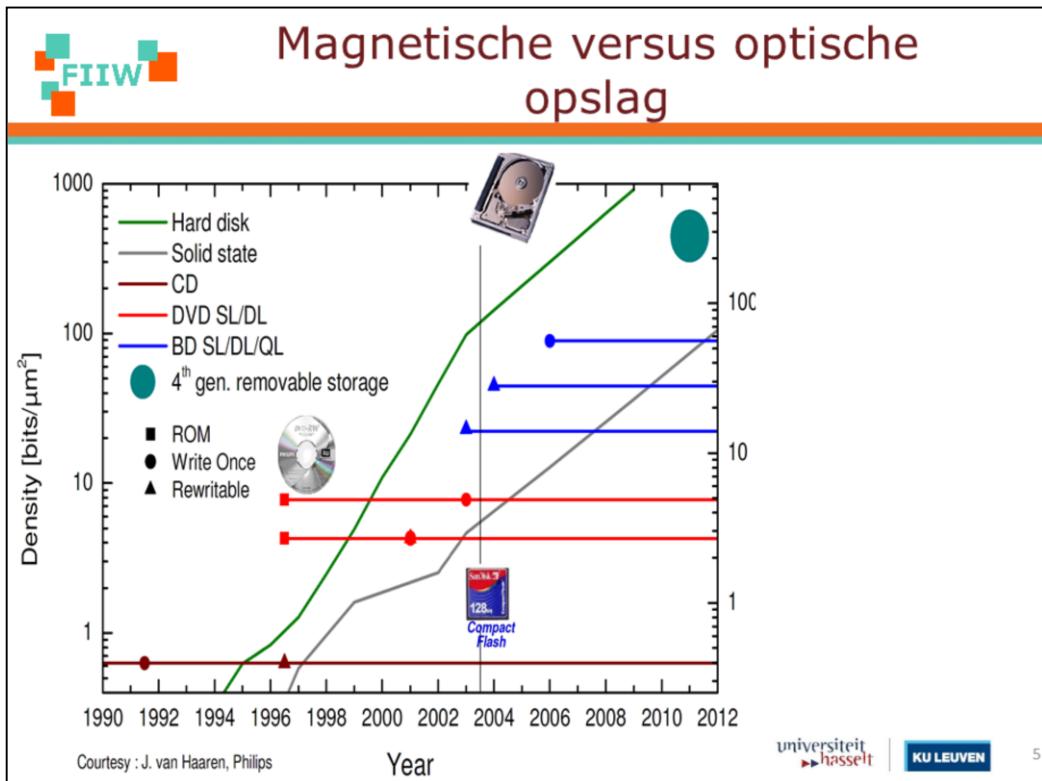


[1] [http://en.wikipedia.org/wiki/Flash\\_memory](http://en.wikipedia.org/wiki/Flash_memory)



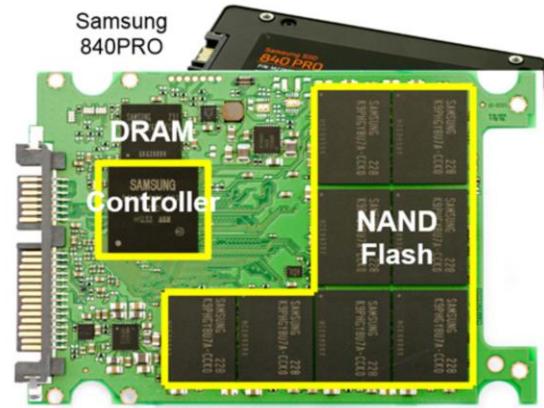
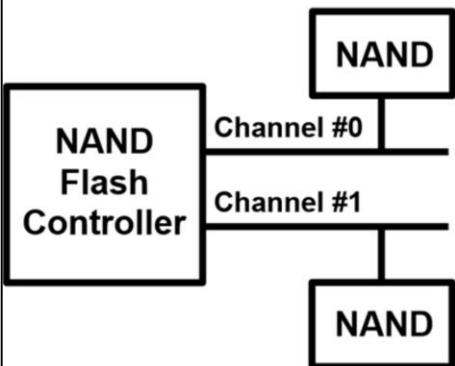






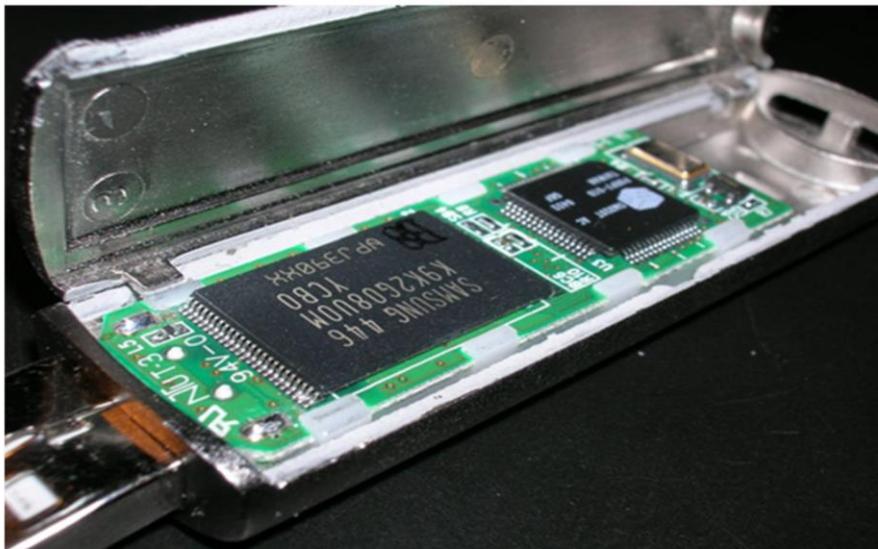


## Binnenkant SSD drive





## Binnenkant USB stick

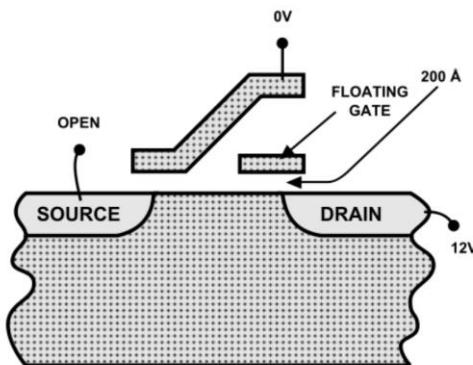


universiteit  
▶ hasselt | KU LEUVEN

## Transistor

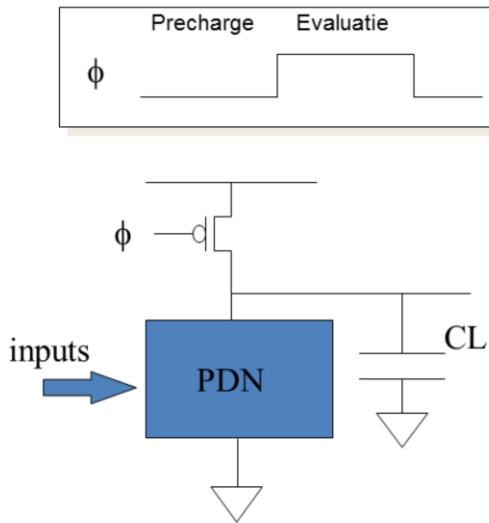
- VT verschuiving bepaalt het aan of af zijn van de transistor

Erasure Via Tunneling



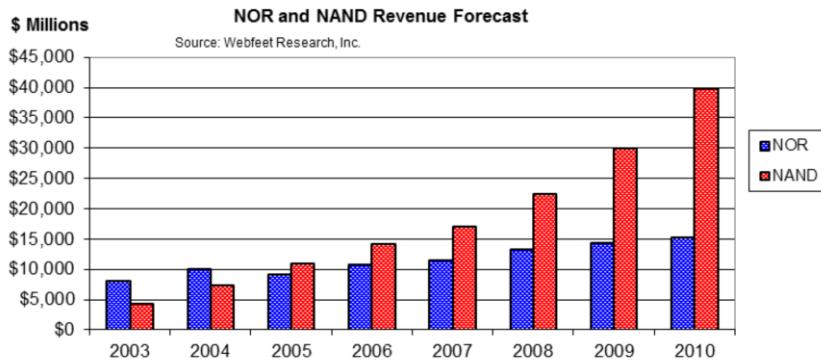
## Opladen van de bitlijnen

- Precharge fase ( $\phi = 0$ ):
  - De uitgang is opgeladen tot 1 ( $V_{DD}$ )
- Evaluatie fase ( $\phi = 1$ ):
  - De ingangen kunnen de uitgang naar 0 trekken.
  - Dit kan maar één keer. Gedurende de evaluatie fase is er geen overgang van 1 naar 0 meer toegelaten op de ingang





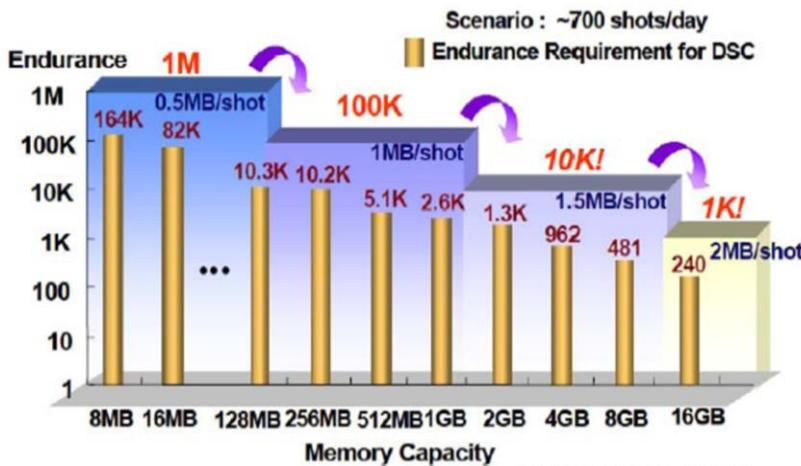
## Flash-gerichte NVM Markt



Flash Groeisnelheid: ~20% { NOR: ~11%      Code storage  
                          NAND: ~30%     Data storage

## Aantal schrijf-wis cycli

### System Requirement (Reliability)



[2] K. Kim et al., NVSMW, pp.9-11, 2006

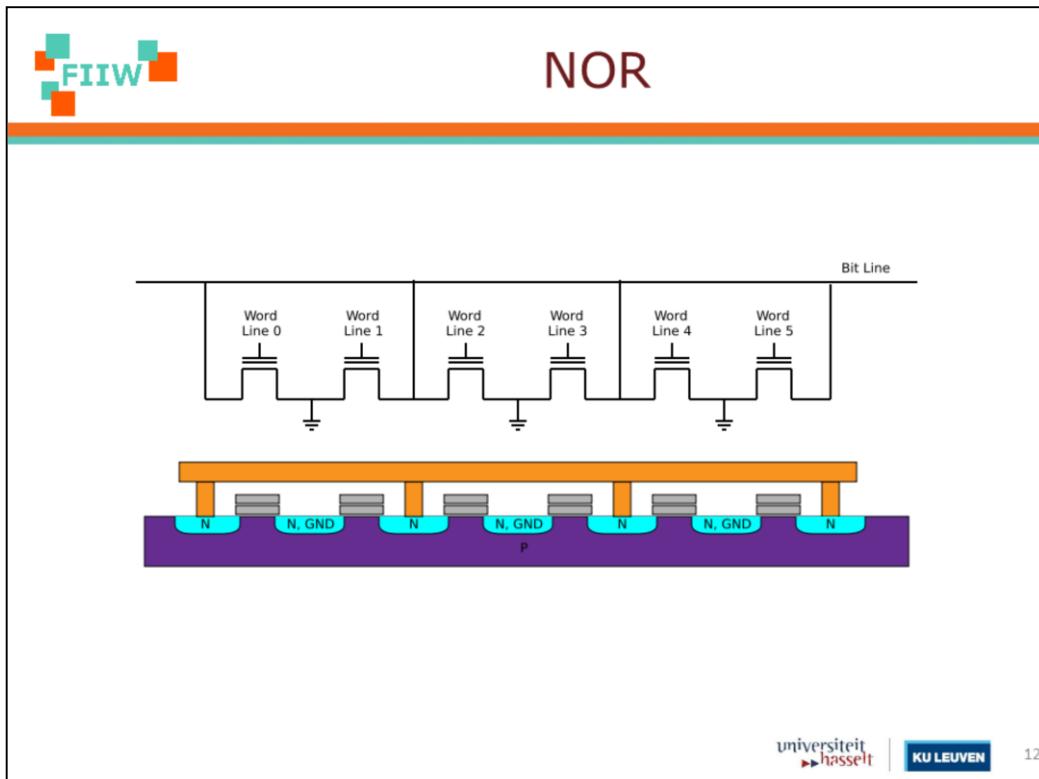
Jan Genoe: Geheugen

universiteit

hasselt

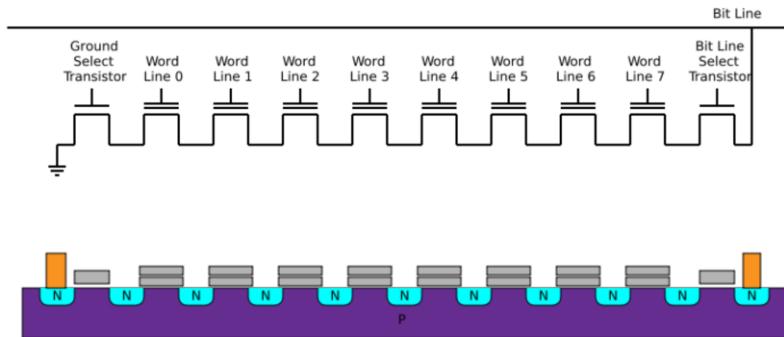
KU LEUVEN

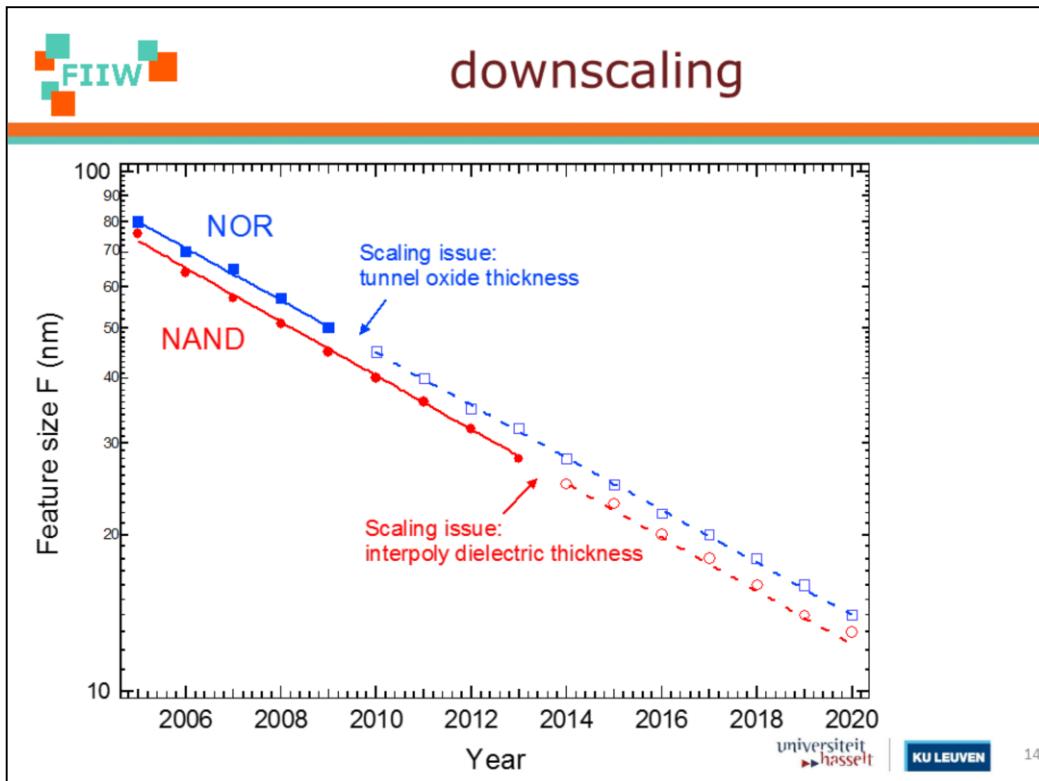
11





# NAND







## Non-Volatile Memory Produkten

### Code



Mobile



Consumer Electronics



Networking

Read	Writes	Density	Reliability
Fast Random	Medium	Small – Medium	No bad bits

### Data



Cards



MP3

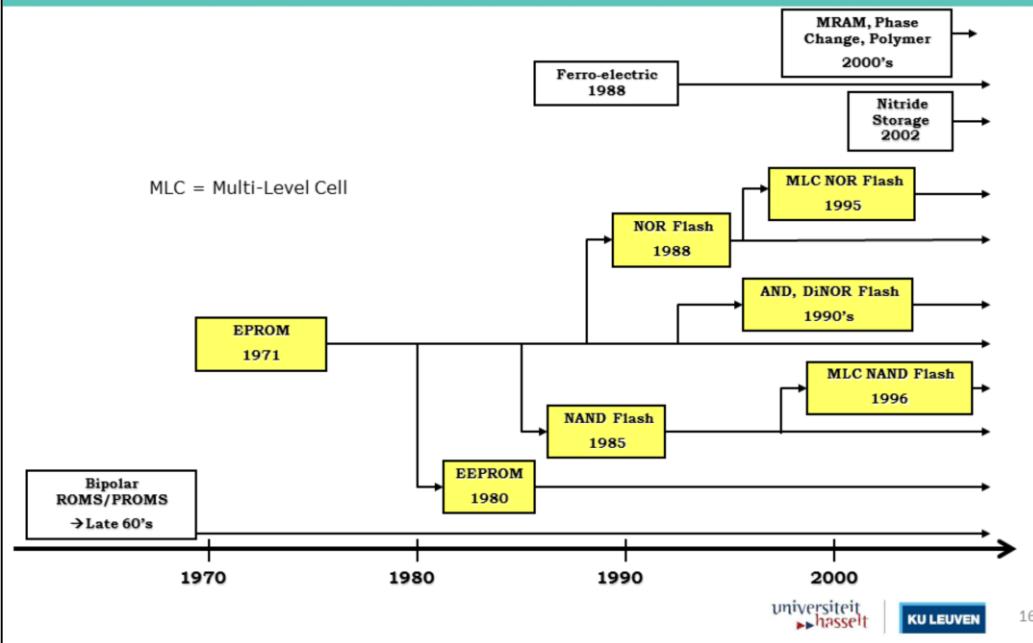


USB Drives

Read	Writes	Density	Reliability
Fast Sequential	Fast	Large	Bad bits allowed



## Non-Volatile Memory geschiedenis





## Non-Volatile Memory Terminologie

- **Program:** Opslag van lading op de floating gate
- **Erase:** Wissen van lading op de floating gate
- **Data Retention:** Hoelang het geheugen zijn data kan behouden
- **Endurance:** Het aantal Program/Erase Cycles dat het geheugen aankan
- **Disturbs:** Een mechanisme dat de opgeslagen data verstoort



## Technologie vergelijking

	Productie			onderzoek		
	NOR Flash	NAND Flash	Nitride	Phase Change	MRAM	FeRam
<b>Cost</b> •Cell Size	$10\lambda^2$	$6\lambda^2$	$6\lambda^2$	$8-15\lambda^2$	$8-15\lambda^2$	$10-20\lambda^2$
<b>Read Characteristics</b> •Cell Read Latency •Cell Read Bandwidth (Array Attribute)	10's ns 100's cells	10's us 1000's cells	10's ns 100's cells	10's ns 10's-100's cells	10's ns 100's cells	10's ns 100's cells
<b>Write Characteristics</b> •Cell Write Time •Cell Write Bandwidth (Array Attribute)	100's ns 10's cells	100's us 1000's cells	100's ns 10's cells	10's ns 10's cells	10's ns 10's cells	10's ns 10's cells

–  $\lambda$  stelt het kleinst mogelijke detail in de technologie voor.

- De mogelijkheden van het lithographie process bepalen  $\lambda$
- Voorbeeld : 0.12u lithographie,  $10\lambda^2$  cell grootte geeft een cell oppervlakte van  $0.144u^2$



## Nor versus Nand specificaties

	READ		PROGRAM		ERASE
	Random toegang	Throughput	Random toegang	Throughput	Throughput
Nor	98 ns	266 MB/s (x16)	500 us (512 bytes)	1 MB/s	0.128 MB/s (1 s per block)
Nand	25 us (1 <sup>st</sup> byte)	37 MB/s (x16)	300 us (2112 bytes)	5 MB/s	64 MB/s (2 ms per block)

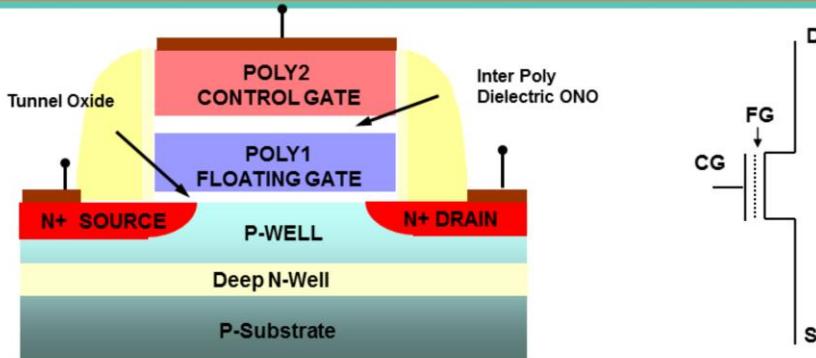
- Multi-Level Cells – Significante Kost Reductie, maar niet gemakkelijk

- 2 maal zoveel data op ongeveer de gelijke oppervlakte

Block Size = 128KB



## Flash Memory Device – NAND & NOR



### Stacked Gate NMOS Transistor

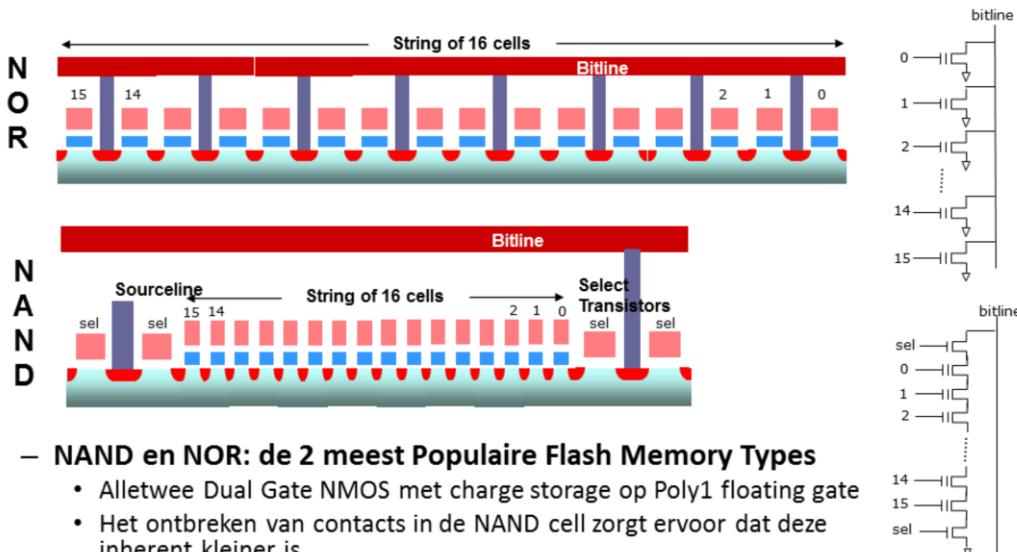
- Poly1 Floating Gate voor ladingsopslag
- Poly2 Control Gate om de transistor te kunnen aansturen
- Tunnel-oxide als Gate oxide
- Oxide-Nitride-Oxide (ONO) voor het inter Poly Dielectric
- Source/Drain Junkties optimaal voor Program/Erase/Leakage

KU LEUVEN

20



## NAND vs. NOR Cross-sections

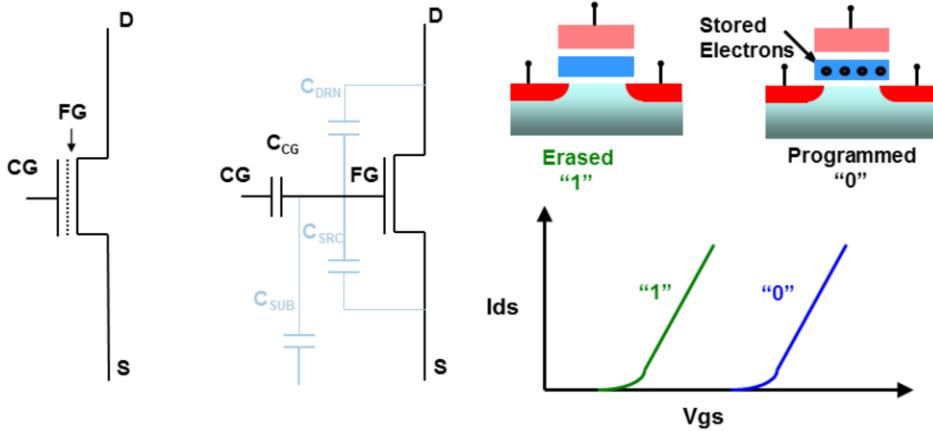


### – NAND en NOR: de 2 meest Populaire Flash Memory Types

- Alleenweer Dual Gate NMOS met charge storage op Poly1 floating gate
- Het ontbreken van contacts in de NAND cell zorgt ervoor dat deze inherent kleiner is.



## Flash Memory Device – werking



- **Programming** = Elektronen opslaan op de FG = hoge V<sub>t</sub>
- **Erasing** = Elektronen weghalen van de FG = lage V<sub>t</sub>
- **Threshold Voltage verschuiving** =  $\Delta Q_{FG} / C_{CG}$

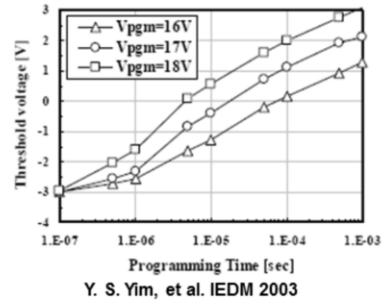
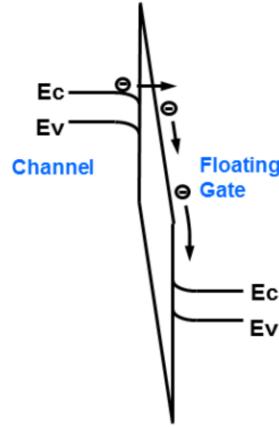
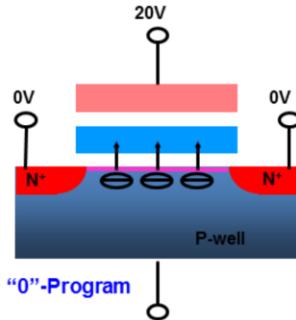
universiteit  
► Hasselt

KU LEUVEN

22



## Nand Flash Programming – FN Tunneling



- **Tunnel Programming** door een hoge positieve spanning aan te leggen op de top-gate (typisch 3 tot 4 maal de voedingsspanning)
- Program tijd  $\sim$ 300us (zeer lang in elektronica termen)
- Program stroom  $\sim$  lage tunnel stroom. Deze lage stroom laat een hoog parallelisme toe.

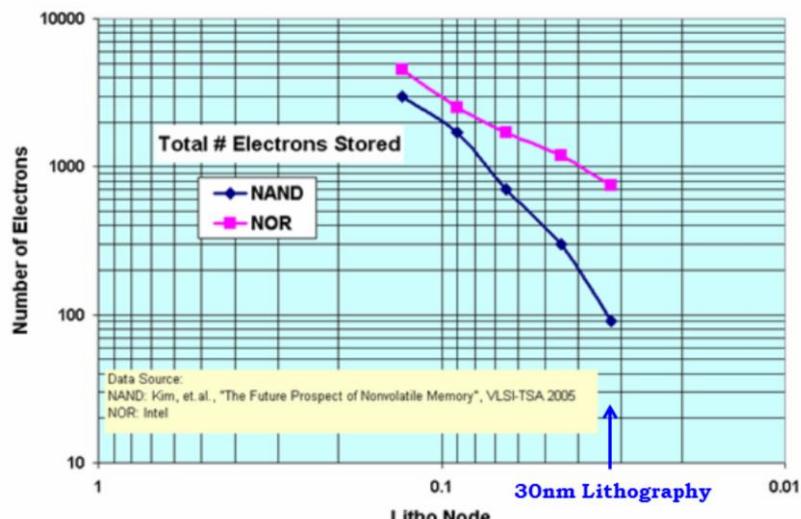
**NOR Flash Programming – Channel Hot Electron**

– **Channel Hot Electron Programming** - Gate voltage inverts channel; drain voltage accelerates electrons towards drain; gate voltage pulls them to the floating gate

- In Lucky Electron Model, electron crosses channel without collision, gaining  $> 3.2\text{eV}$ , hits Si atom, bounces over barrier
- Program Time  $\sim 0.5\text{-}1\text{ms}$ . Program current  $\sim 50\text{mA/cell}$

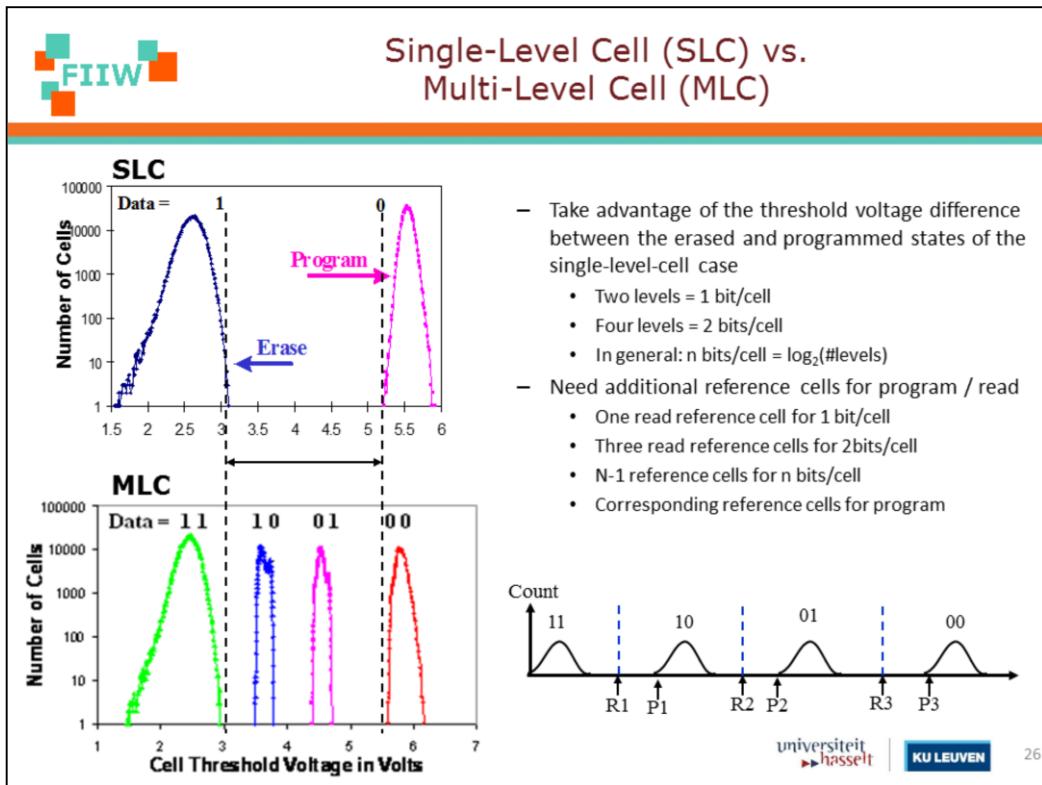
universiteit  
▶ hasselt | KU LEUVEN | 24

## Floating Gate Elektronen in functie van Lithografie evolutie

universiteit  
▶ hasselt

KU LEUVEN

25





# Multi-Level Cell Design

- **Why do MLC? Cost**

- Effectively cuts cell area per bit in half
- Provides the same cost improvement from an array area perspective as a litho generation

- **Three Key MLC Considerations**

- **Precise Charge Placement (Programming)**

- Cell programming must be accurately controlled, which requires a detailed understanding of cell physics, voltage control and timing
  - Precision voltage generation for stable wordline and drain voltage

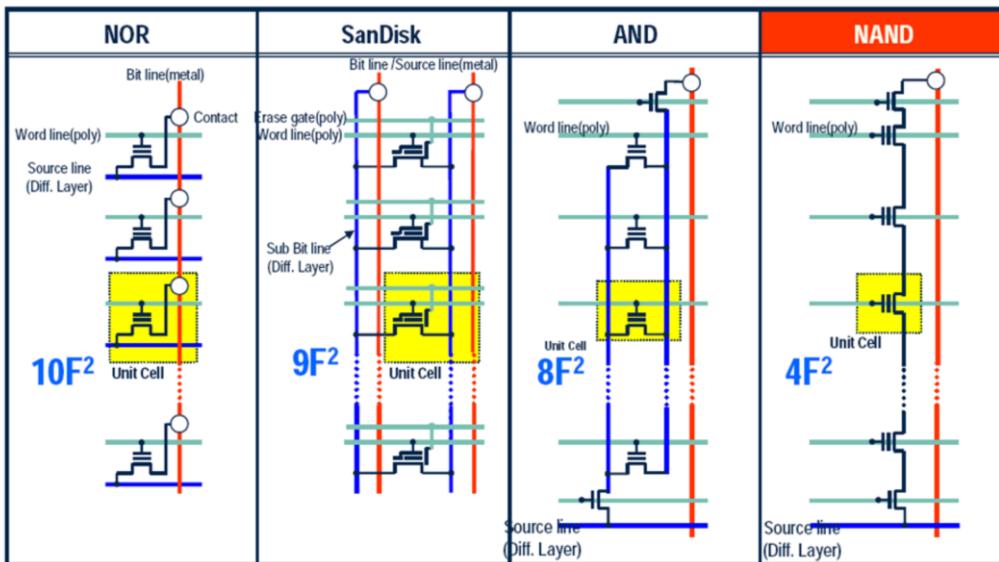
- **Precise Charge Sensing (Read)**

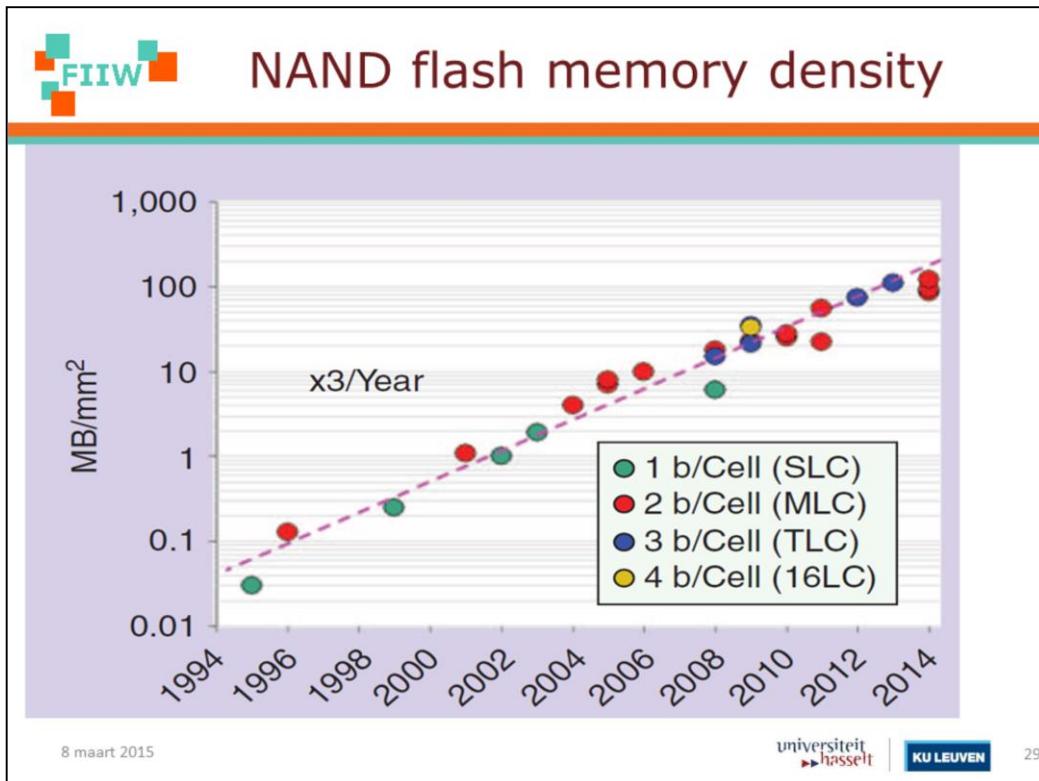
- MLC read operation is an analog to digital conversion of the charge stored in the cell
  - Device and capacitance matching, Collapsing sources of variation, Precision wordline and drain voltage generation, Low current sensing

- **Stable Charge Storage**

- Leakage rate needs to be less than one electron per day

## Oppervlak vergelijking





- [1] S. Narendra, L. Fujino, and K. Smith, "Through the Looking Glass Continued (III): Update to Trends in Solid-State Circuits and Systems from ISSCC 2014 [ISSCC Trends]," *IEEE Solid-State Circuits Mag.* 6, pp. 49–53, Mar. 2014.

2D NAND

128Gb

Image not to scale

Intel 3D NAND

256Gb

Image not to scale

32 Tiers

8 maart 2015

universiteit  
hasselt | KU LEUVEN

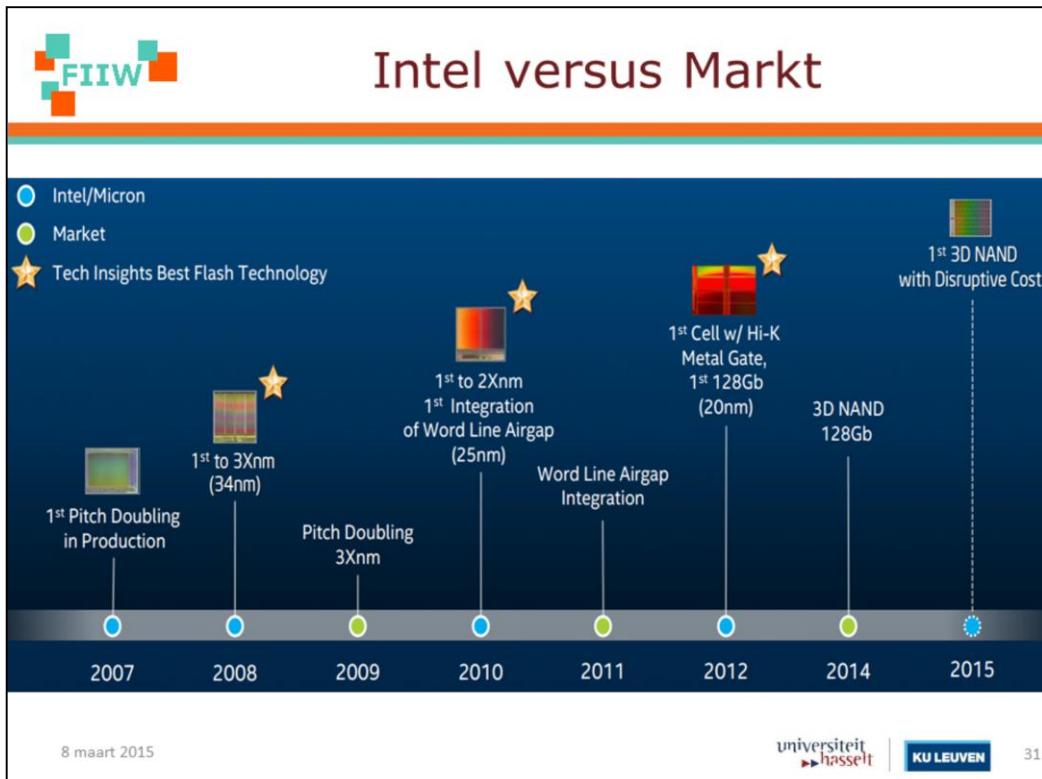
30

Intel komt in 2015 met ssd's op basis van 3d-nand-geheugen die een 'ontwrichtende prijs' zullen hebben. Verder maakt 3d-nand binnen enkele jaren ssd's met een opslagcapaciteit van 10TB mogelijk, volgens het bedrijf.

Het 3d-nand dat Intel met Micron maakt, bestaat uit 32 lagen en biedt een capaciteit van 256Gb op een enkele mlc-die. Volgens de fabrikant maakt dit 1TB-capaciteiten in een 2mm-package mogelijk en zorgt het voor een 'doorbraak op gebied van kosten', maar Intel noemt geen concrete prijzen. Uiteindelijk zou het 3d-nand ssd's met grotere opslagcapaciteit dan 10TB mogelijk maken en dit zou al binnen enkele jaren het geval kunnen zijn. Dat maakte Rob Crooke, general manager van Intels Non-Volatile Memory Solutions Group bekend tijdens een investeerderspresentatie. Bij tlc-geheugen zou de capaciteit op 385Gb per die uitkomen.

Intel heeft hoge verwachtingen van de ssd-markt. Momenteel maakt Samsung 3d-nand op basis van 32 lagen, waarmee een capaciteit van 86Gb op een mlc-die en 128Gb bij tlc gehaald wordt. In 2015 moet een nieuwe generatie van het zogenoemde 3d-v-nand verschijnen waarmee Samsung mogelijk de concurrentie met Intel kan aangaan. Toshiba en Sandisk zijn een fabiek aan het

bouwen voor de productie van 3d-nandgeheugen, maar ssd's met dit type geheugen van die fabrikanten worden pas in 2016 verwacht.

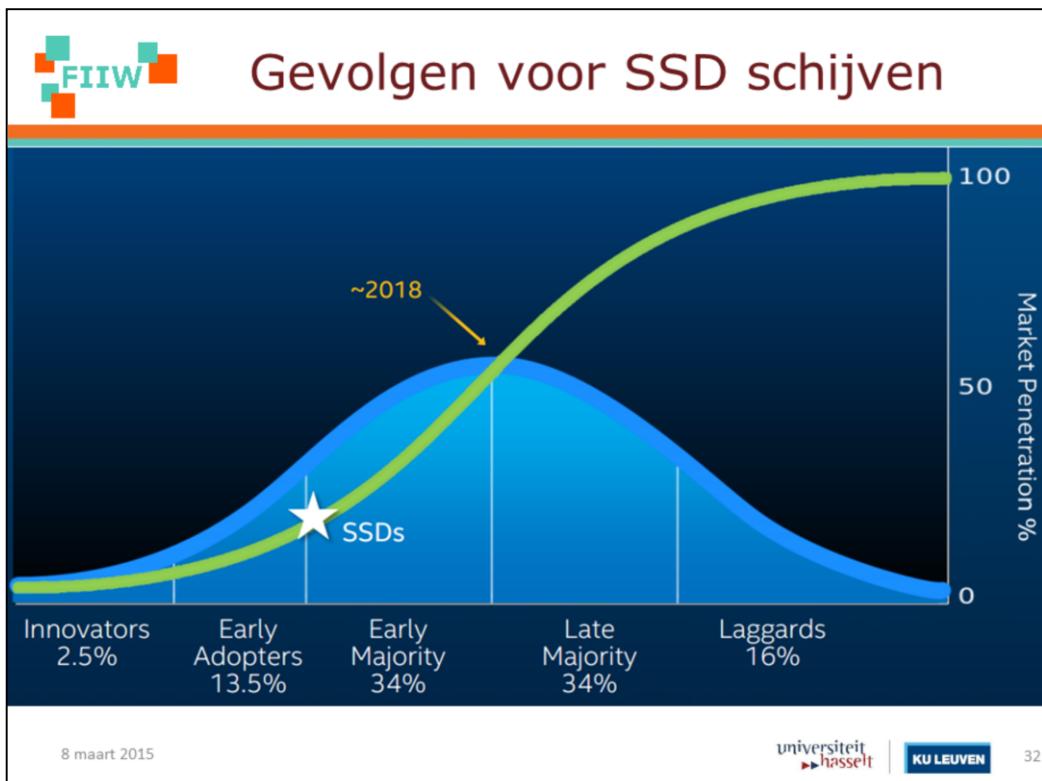


8 maart 2015

universiteit  
▶ hasseLT

KU LEUVEN

31



**FIIW**

## Samsung: Vertical NAND

### Limitations in planar NAND

Difficulty

Design rule

Interference Coupling

No. of electrons  
V<sub>T</sub> window, etc  
Process window

### Improvement through V-NAND

GAA structure → Lower coupling → 1-step high-speed PGM

Planar (Higher Coupling) VS. V-NAND (Lower Coupling)

Erase

1<sup>st</sup> PGM

2<sup>nd</sup> PGM

Channel

Charge trap

WL

universiteit  
hasseLT | KU LEUVEN

GAA: Gate All Around

[1] Kinam Kim, President, Samsung Electronics, Kiheung, Korea, Silicon Technologies and Solutions for The Data-Driven World, ISSCC 2015.

Wanneer je hoge koppeling hebt, heb je meerdere stappen nodig om meerdere bits in een cel te programmeren.



## Hogere densiteit in V-NAND

Higher density by stacking of cells in 3D

Planar      V-NAND

V-NAND is able to better support multi-bit technology

Planar (4bit)      V-NAND (4bit)

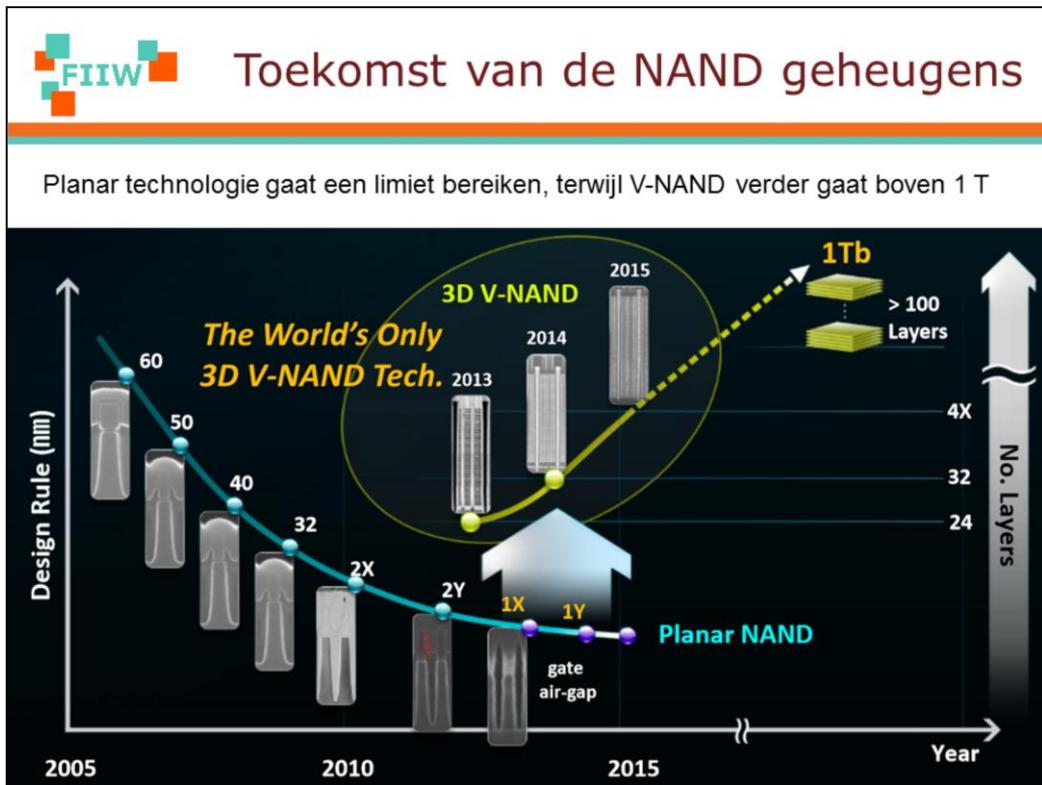
t<sub>PROG</sub>

bit	Planar (ms)	V-NAND (ms)
1bit	~10	~10
2bit	~20	~15
3bit	~35	~25

universiteit  
hasselt | KU LEUVEN

This slide compares Planar and V-NAND technologies. It shows that V-NAND allows for higher density by stacking cells in 3D. It also highlights that V-NAND is better suited for supporting multi-bit technology, specifically 4-bit, as evidenced by its superior programming time performance compared to Planar technology across 1-bit, 2-bit, and 3-bit operations.

[1] Silicon Technologies and Solutions for The Data-Driven World,  
ISSCC 2015, Keynote Samsung

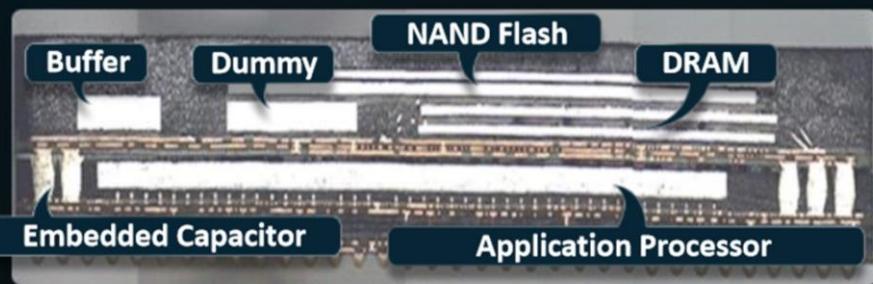


[1] Silicon Technologies and Solutions for The Data-Driven World,  
ISSCC 2015, Keynote Samsung



## NAND in PoP (Package-on-Package)

- | AP : Octa-core CPU
- | Memory : 4GB LPDDR4 DRAM  
64GB NAND Flash
- | Size : 14 x 14 x 1.5mm<sup>3</sup>



(voor gebruik in mobile platformen)

universiteit  
▶ hasselt

KU LEUVEN