Vidyabhushan Mohan

Modeling The Physical Characteristics of NAND Flash Memory

Outline

- Introduction
- Motivation
- Overview of FlashPower and FENCE
- NAND Flash Memory Primer
- FlashPower – Modeling and Validation
- FlashEnduraNCE – Modeling and analysis
- Conclusion
- Future Work

Flash Memory

- Solid State Memory – Early 80s

Flash Memory

- Systems have different requirements
  - Power
  - Performance
  - Reliability
- Need – Flash Modeling tools.

Motivation - Flash Modeling tools

System level

+ Use as one single value.
+ Device independent analysis.
- Cannot isolate the effect of device parameters

Architecture level

+ Abstract device-level details.
+ Explore design space.
+ Parameterizable.
+ Use with other simulators.

Device level

+ Device level characterizations.
+ Device level optimizations.
- Cannot be directly translated to system performance.

Motivation - Why Power and Endurance?

- Power
  - Impacts capacity and performance
  - New hybrid memory systems
- Endurance
  - Impacts the deployment of NAND Flash
  - Simplistic endurance estimates

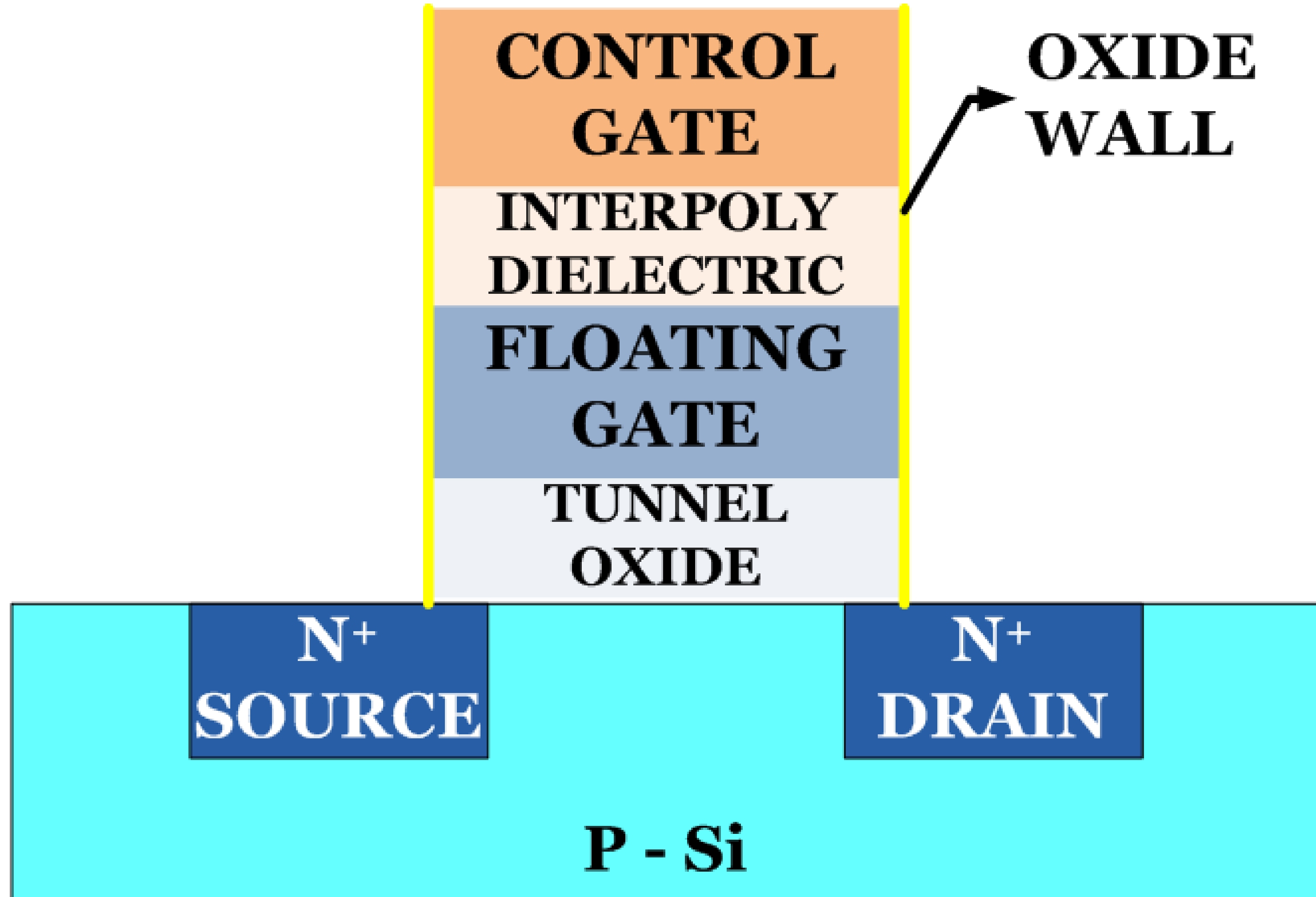
Overview of FlashPower and FENCE

- FlashPower
  - Model power consumption of a NAND Flash chip
  - Parameterized
  - Integrated with CACTI 5.3
- FlashEnduraNCE (FENCE)
  - Stress and Recovery Model
  - Quantify the benefit that recovery can provide

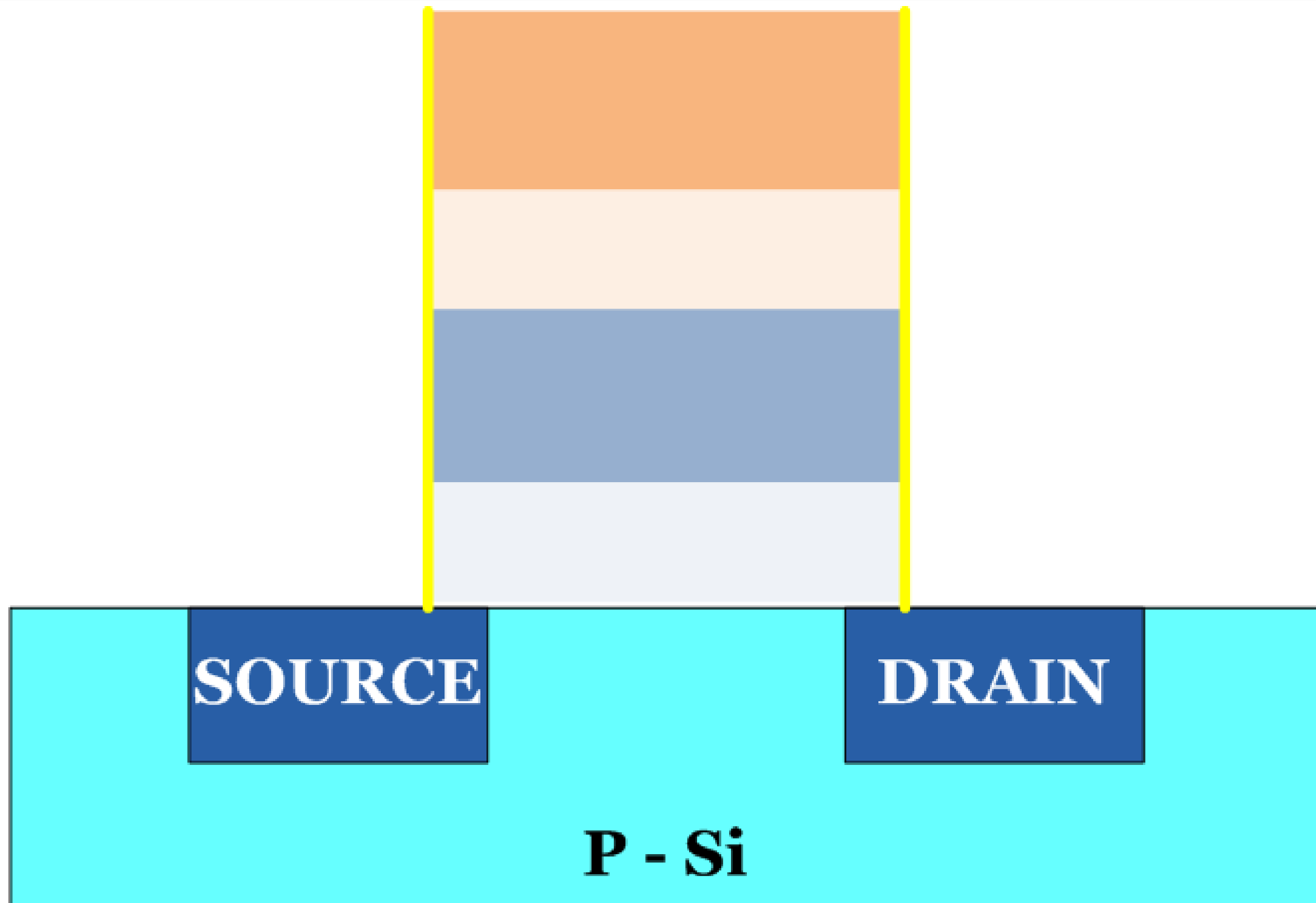
Outline

- Introduction
- Motivation
- Overview of FlashPower and FENCE
- NAND Flash Memory Primer
- FlashPower – Modeling and Validation
- FlashEnduraNCE – Modeling and analysis
- Future Work
- Conclusion

NAND Flash Memory Primer – A Floating Gate Transistor (FGT)

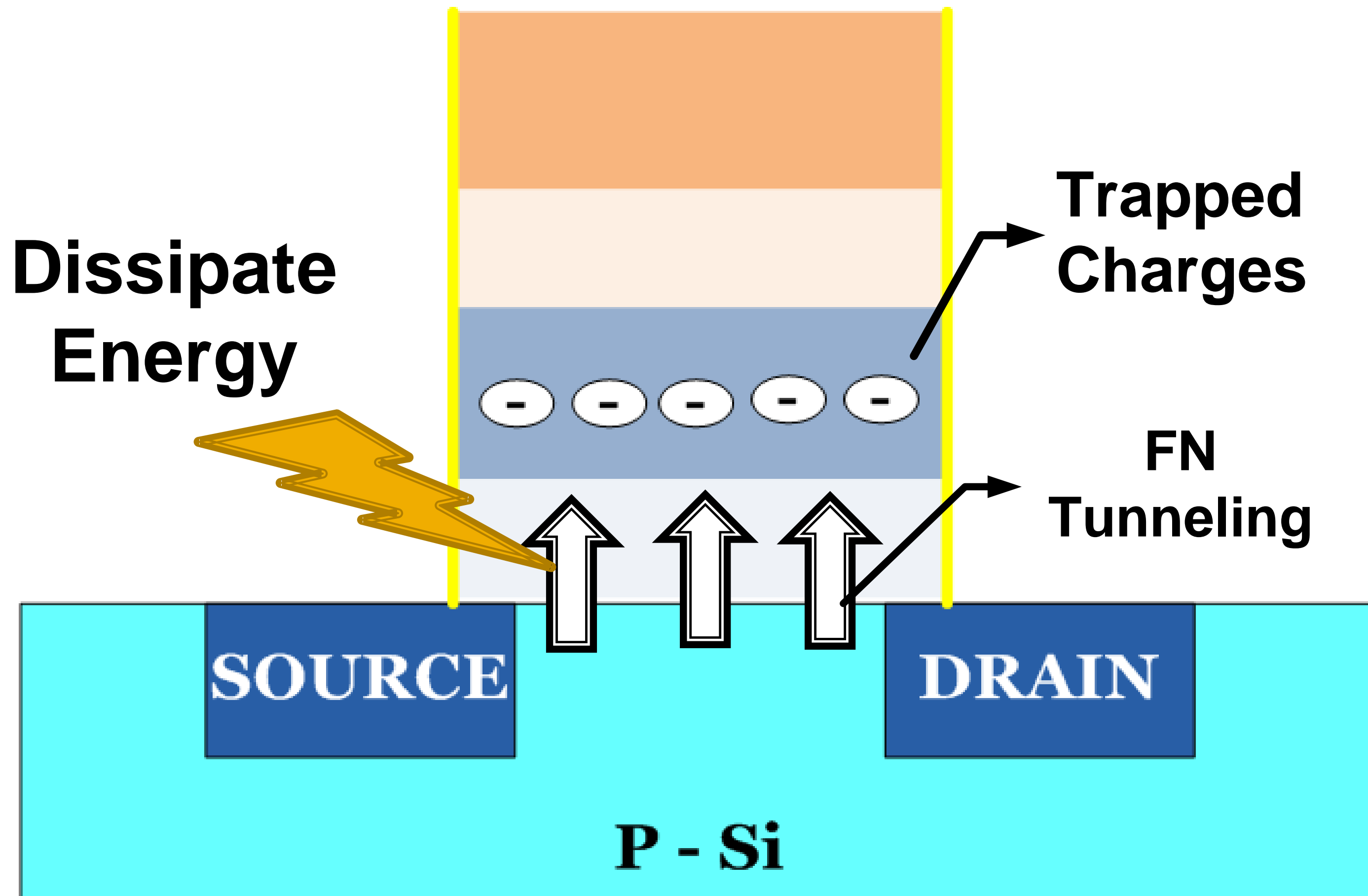


FGT-Default Value '1'



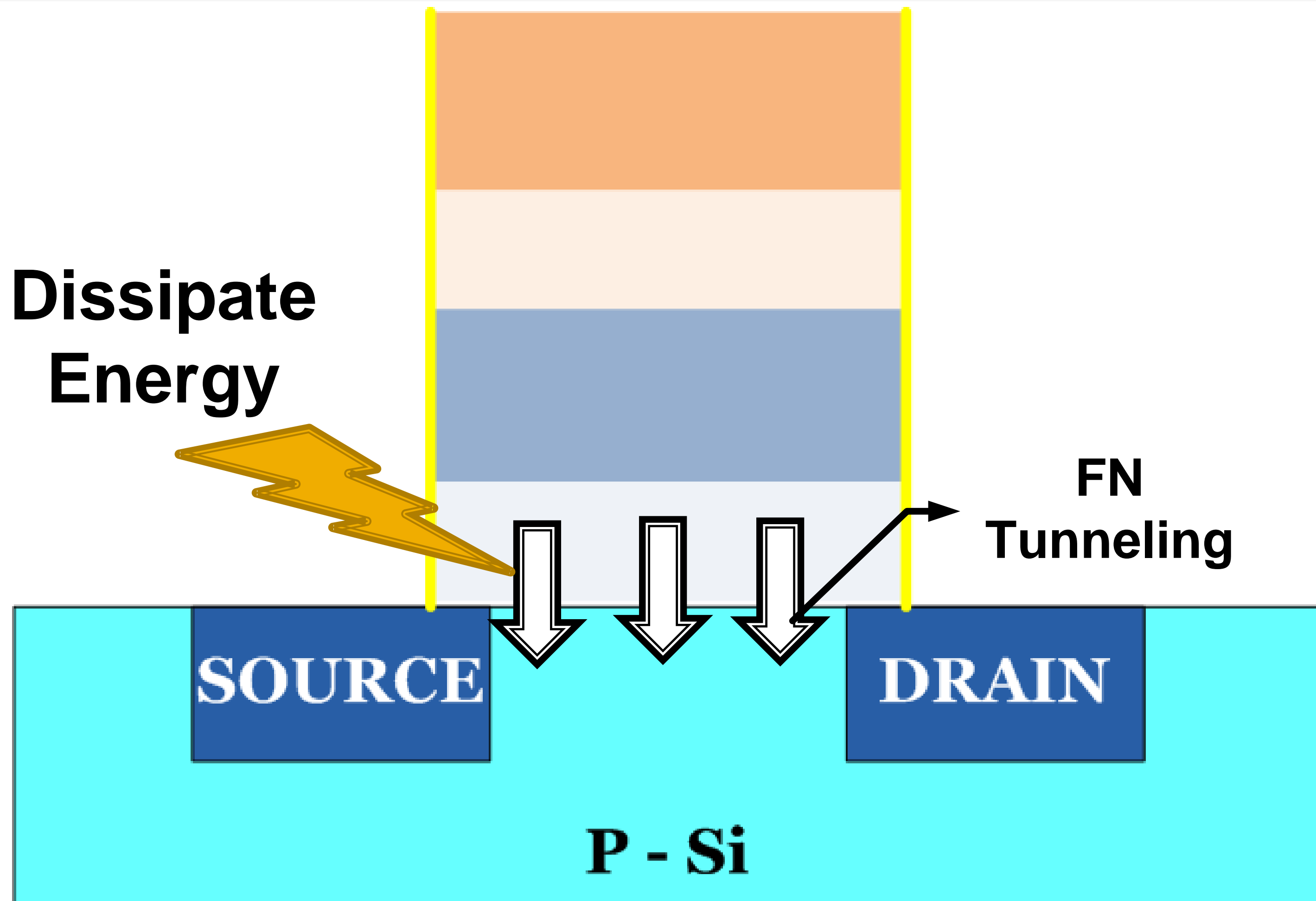
The threshold voltage of the FGT is small

FGT – Program 'o'



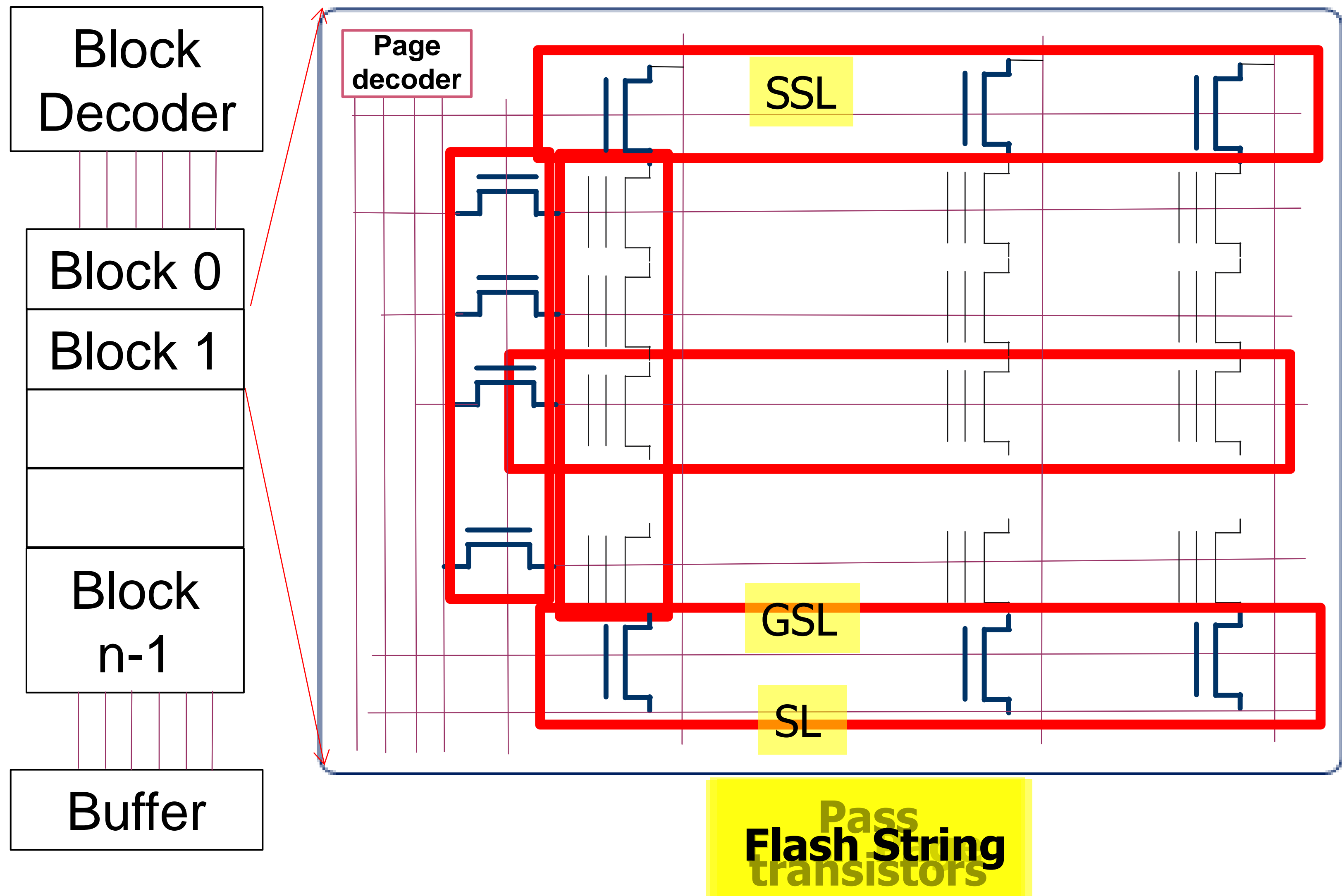
Increases the threshold voltage of the FGT

FGT-Erase '1'

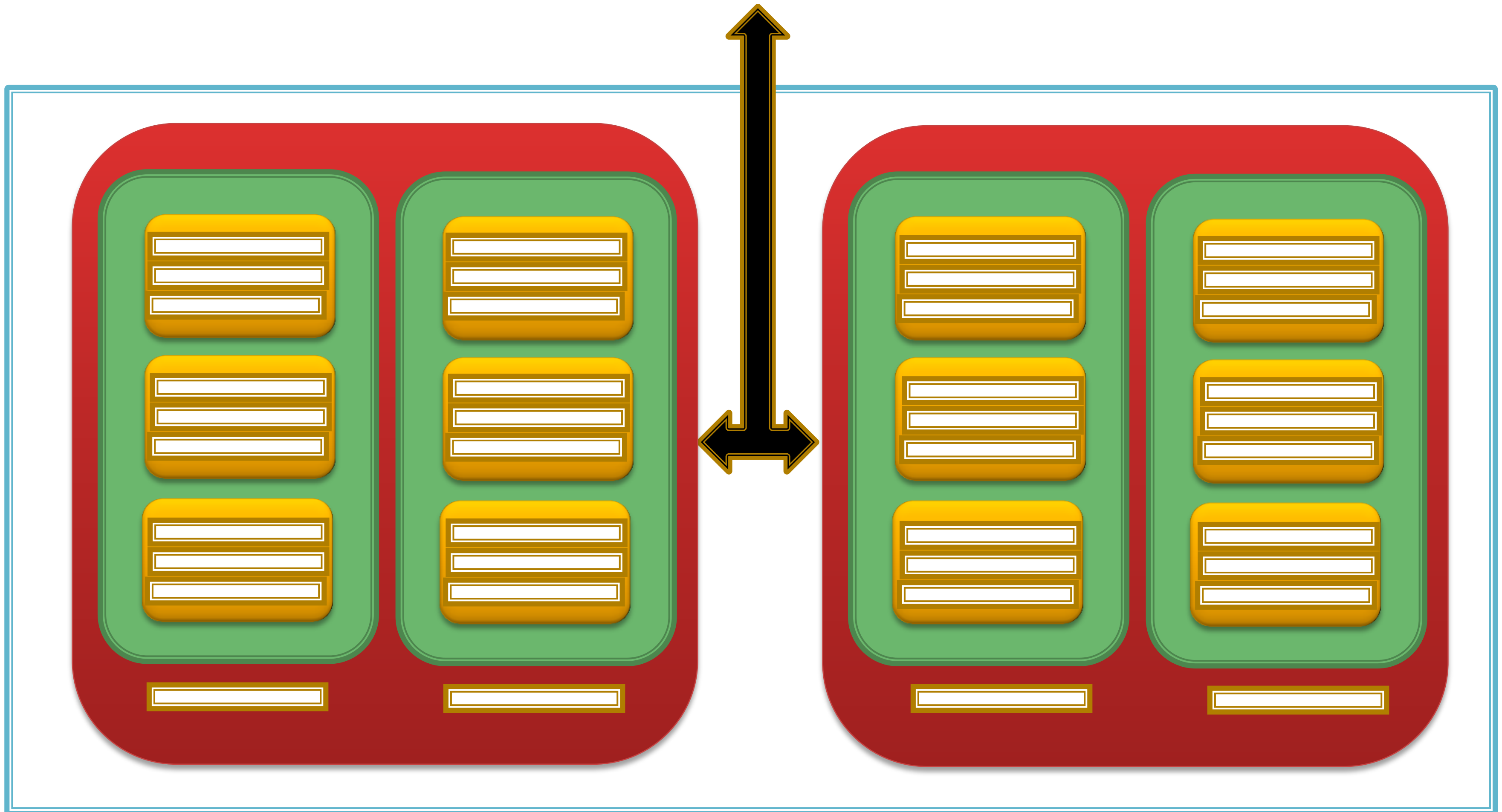


Decreases the threshold voltage of the FGT

NAND Flash Memory Primer – A Flash Plane



NAND Flash Memory Primer – A Flash Memory Chip

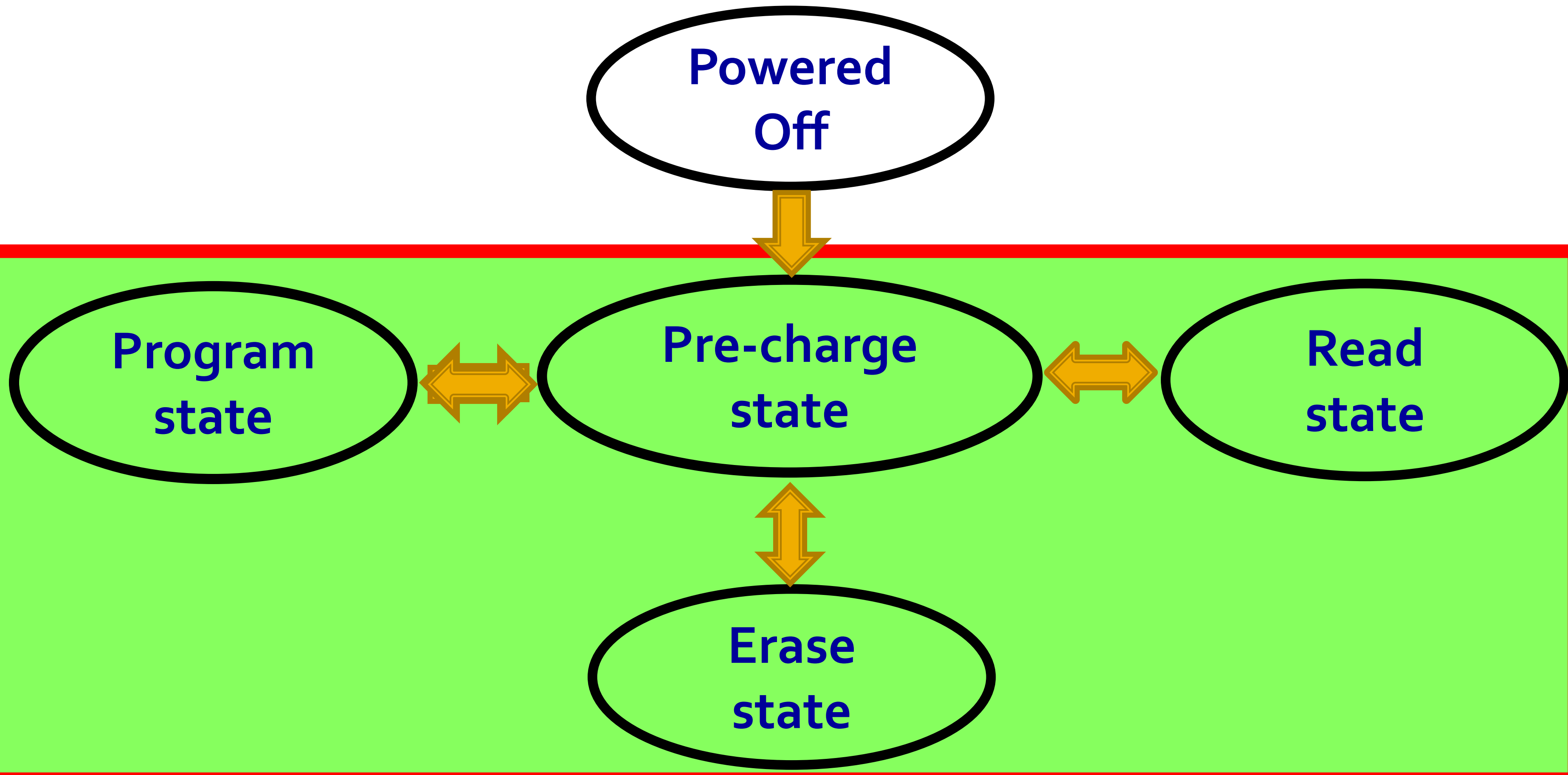


Dies=>Planes + Buffer=>Block=>Pages

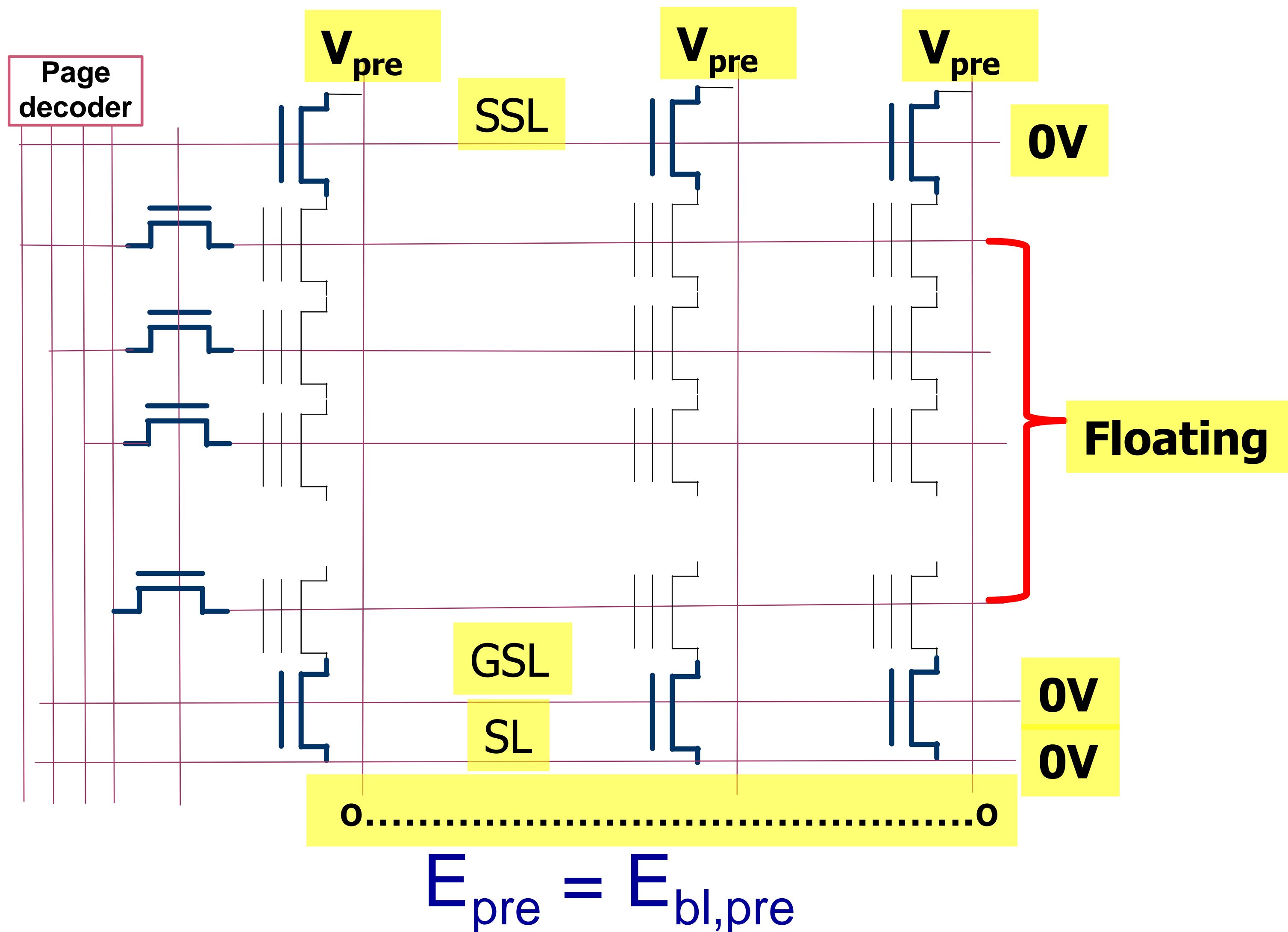
Outline

- Introduction
- Motivation
- Overview of FlashPower and FENCE
- NAND Flash Memory Primer
- FlashPower – Modeling and Validation
- FlashEnduraNCE – Modeling and analysis
- Future Work
- Conclusion

FlashPower - State Machine

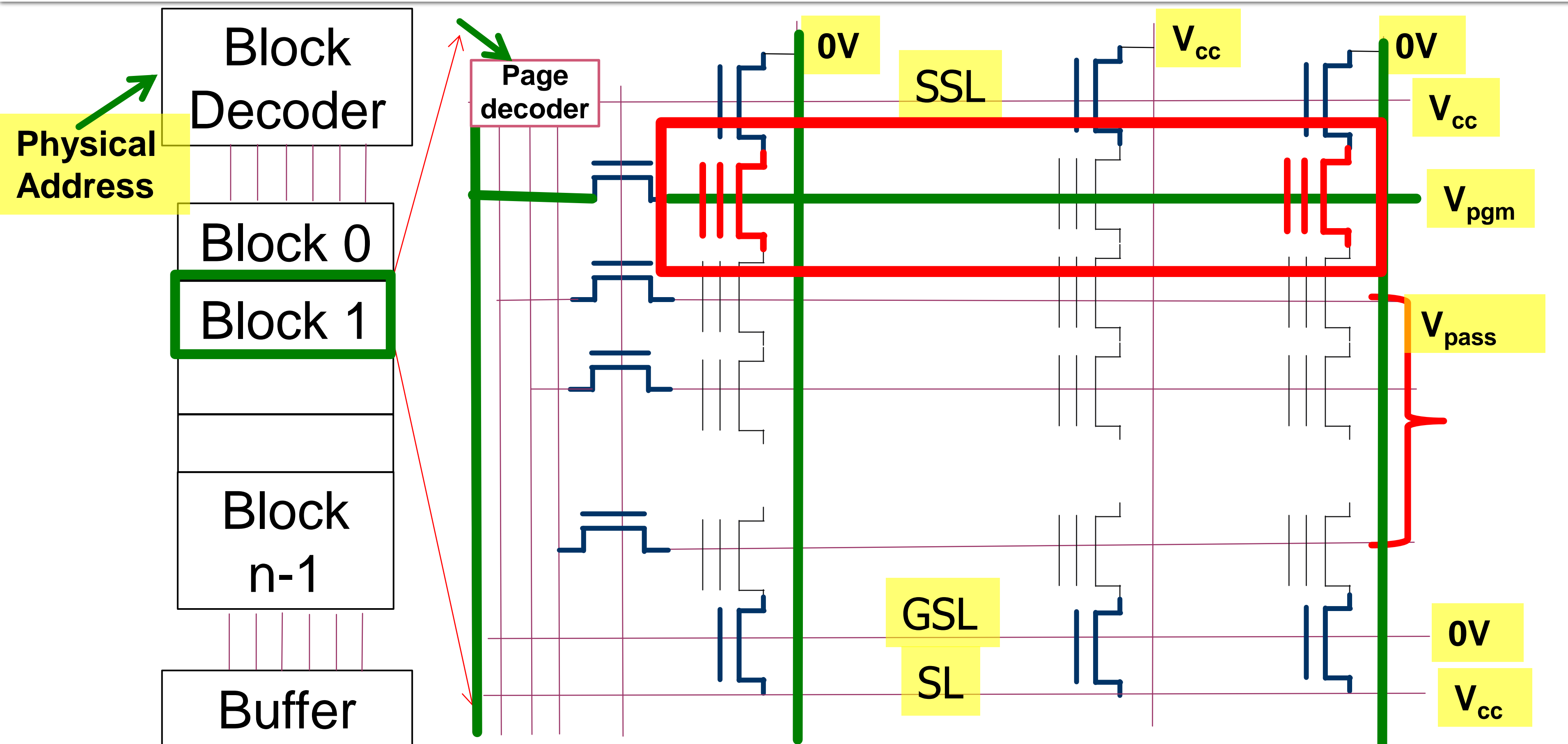


Pre-charge state



$$E_{pre} = E_{bl,pre}$$

Program Operation



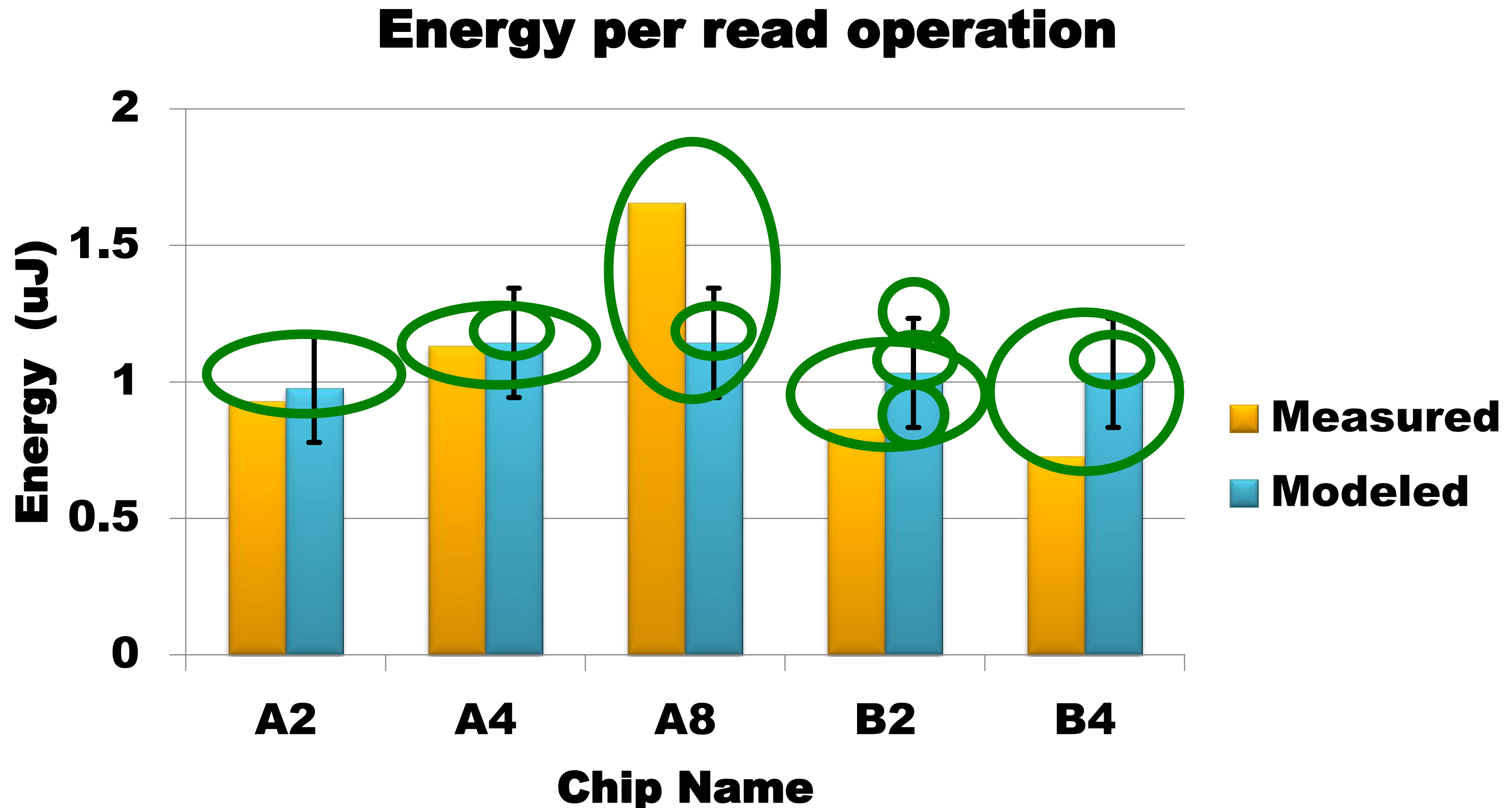
$$E_{program}=E_{iopins}+E_{buffer}+E_{dec,p}+N_{loop}*(E_{blip,p}+E_{bl,p}+E_{selected\text{-}page,p}+E_{unselected\text{-}page,p}+E_{sl}+E_{verify\text{-}program})+E_{p\text{-}pre}$$

Validation

- Flash Memory Characterization – Grupp et al, MICRO 2009.

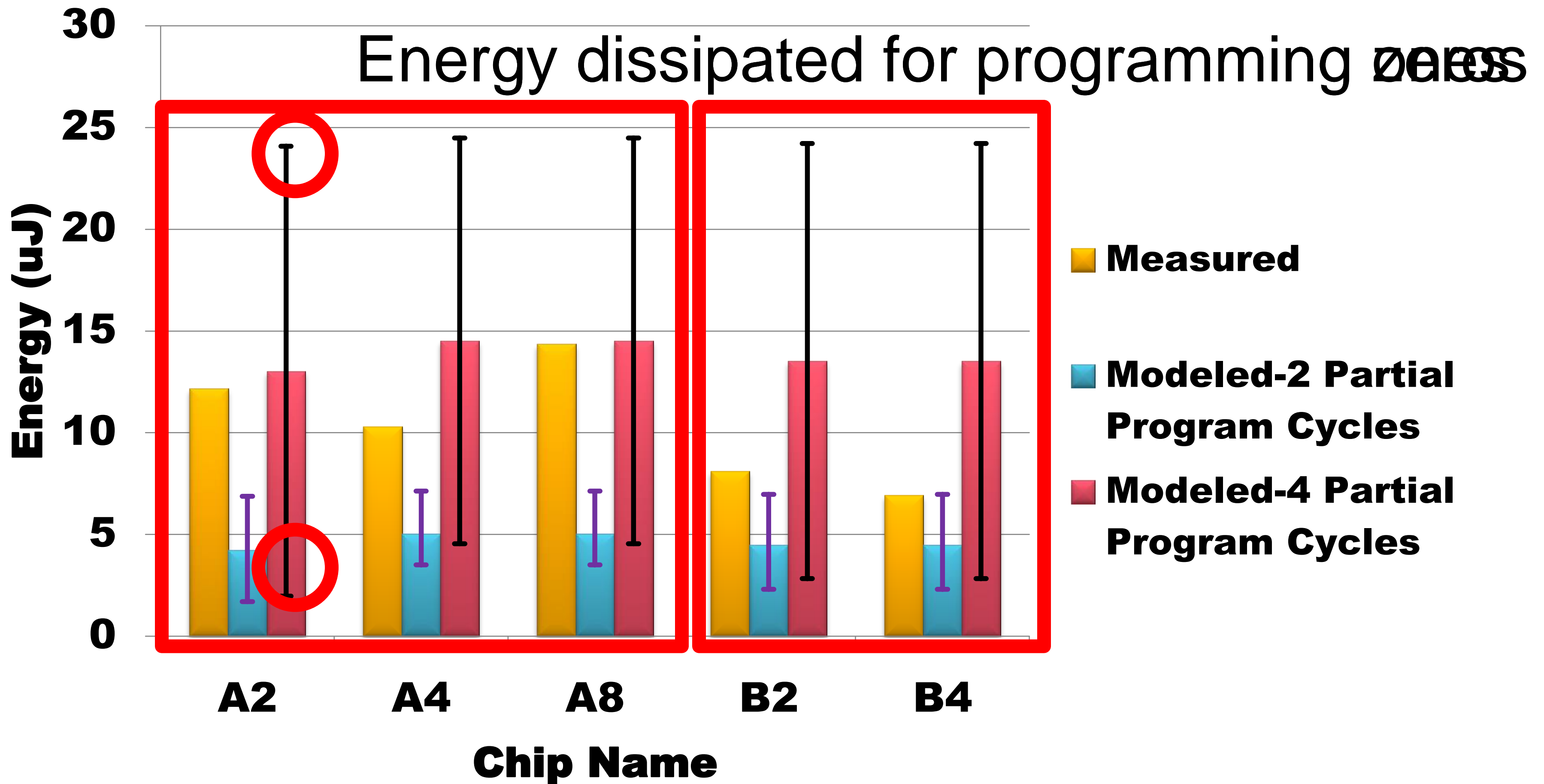
Chip Name	Capacity (Gb)	Page size (B)	Pages/ Block	Blocks/ plane	Planes /die	Dies
A2	2	2112	64	1024	2	1
A4	4	2112	64	4096	1	1
A8	8	2112	64	4096	2	1
B2	2	2112	64	2048	1	1
B4	4	2112	64	2048	2	1

Validation - Read



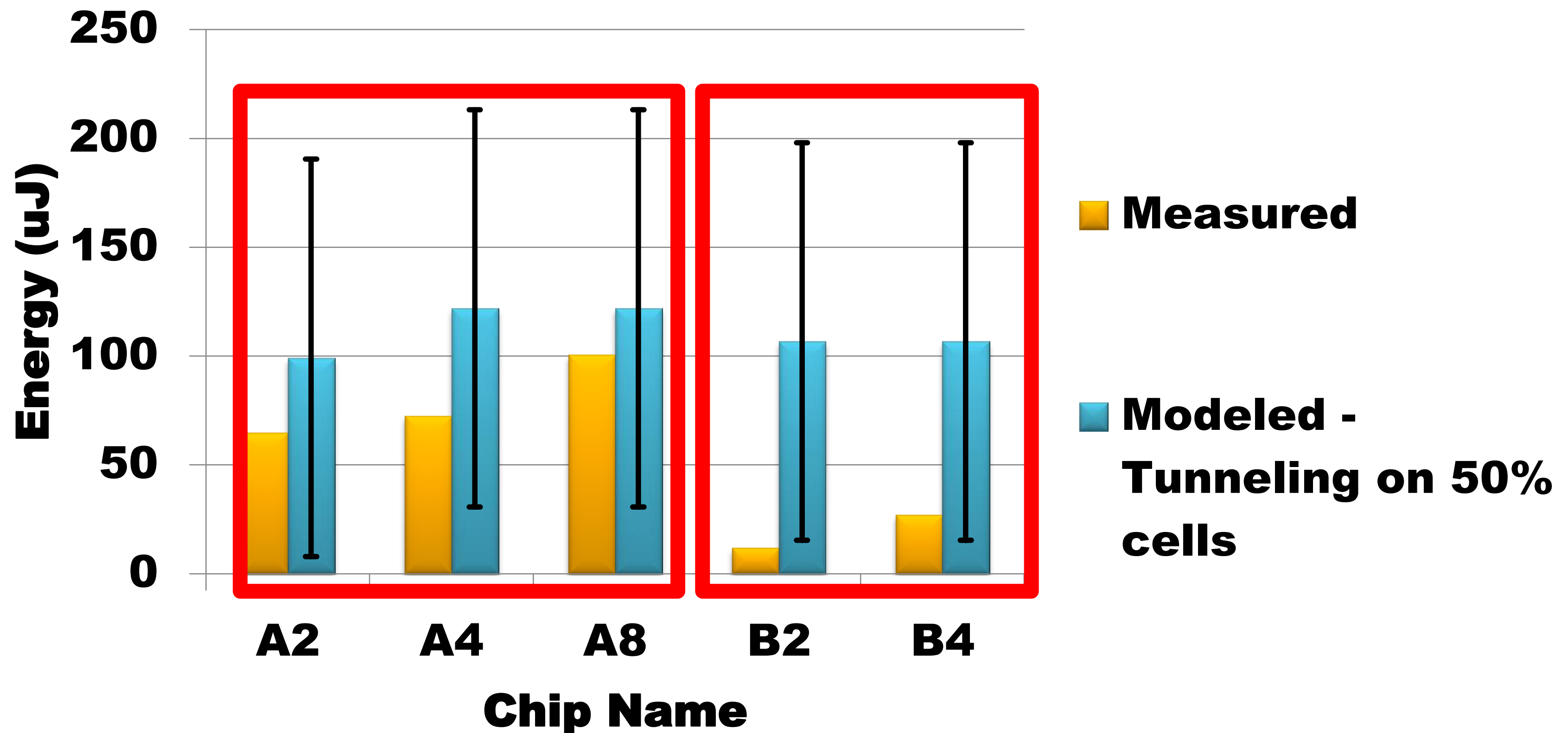
Error bars depict energy variation based on value of bits.

Validation- Program



Error bars depict energy variation based on value of bits.

Validation - Erase



Error bars depict energy variation based on value of bits.

Outline

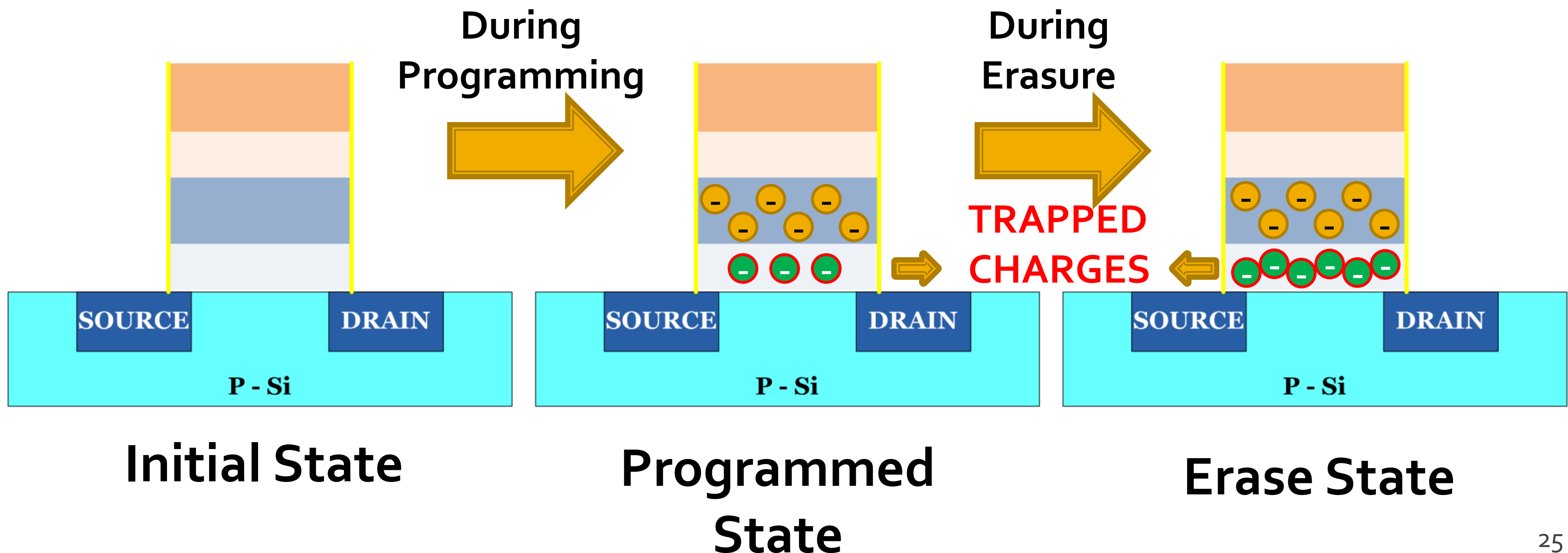
- Introduction
- Motivation
- Overview of FlashPower and FENCE
- NAND Flash Memory Primer
- FlashPower – Modeling and Validation
- FlashEnduraNCE – Modeling and analysis
- Future Work
- Conclusion

Flash EnduraNCE (FENCE)

- Stress Model
- Recovery Model
- When do the failures occur?
- Impact of detrapping on endurance
- Analysis of SSD level Endurance

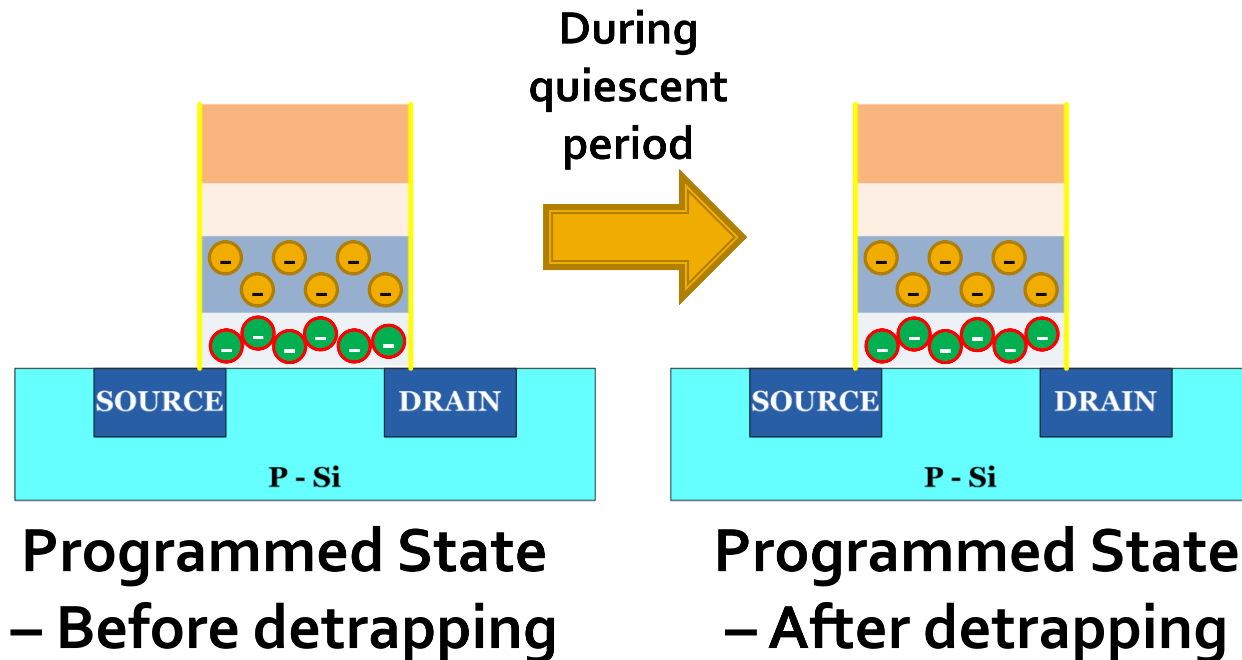
FENCE – Stress

- Stress Model
  - Program and Erase operations are stress events
  - Results in charges trapped in tunnel oxide.



FENCE – Recovery

- Recovery model
  - Quiescent period between stresses
  - Some charges get **detrapped**

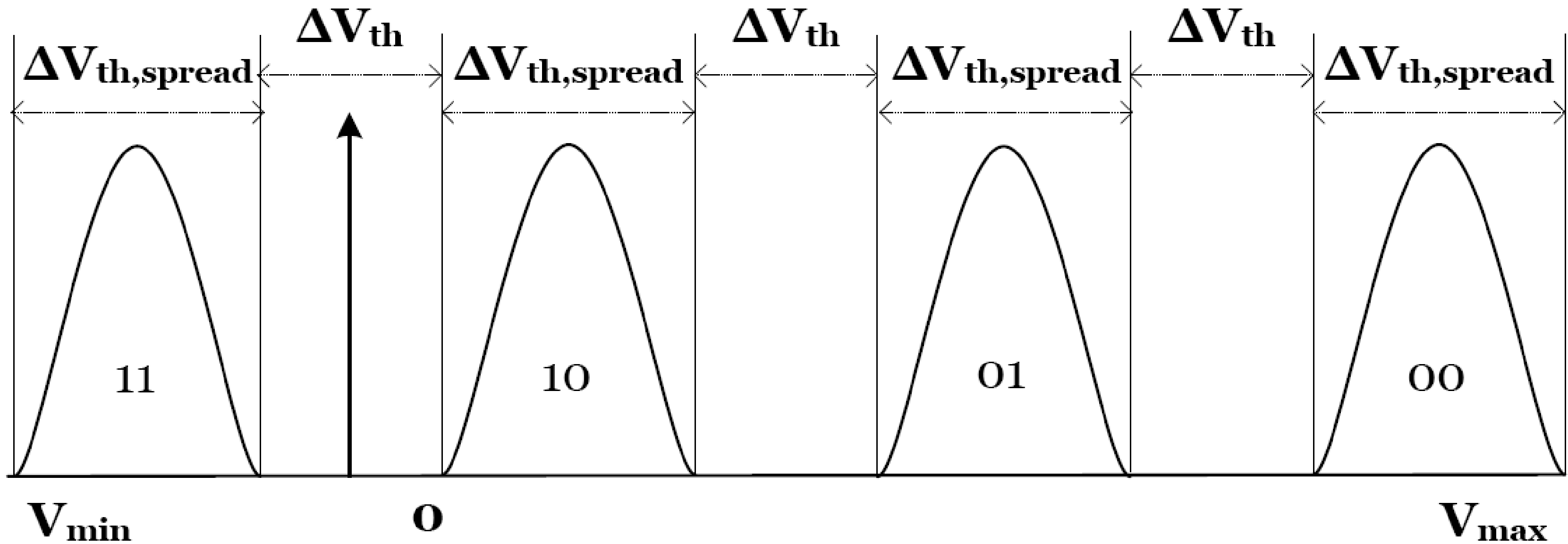


FENCE – Stress and Recovery Model

- Increase in threshold voltage due to stress - $\Delta V_{th,s} = C_1 * cycle^{0.62} + C_2 * cycle^{0.30}$
- Decrease in threshold voltage due to recovery - $\Delta V_{th,r} = C_3 * \ln(\Delta V_{th,s}) * \ln(t)$
- Effective increase in threshold voltage - $\delta V_{th} = \Delta V_{th,s} - \Delta V_{th,r}$

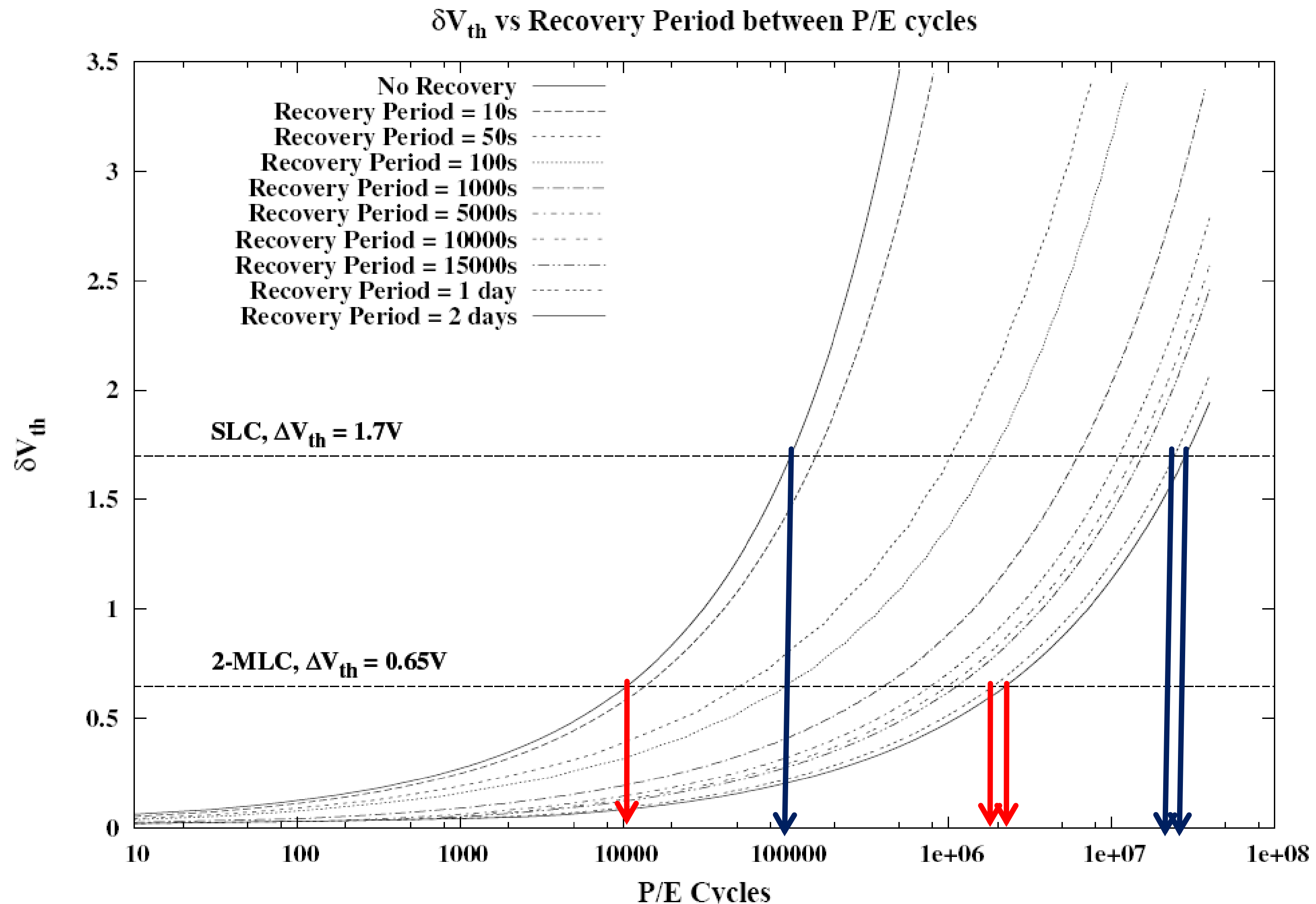
When do failures occur?

- Failure when $\delta V_{th} = \Delta V_{th}$



- ΔVth = 1.7V for SLC and 0.65V for MLC

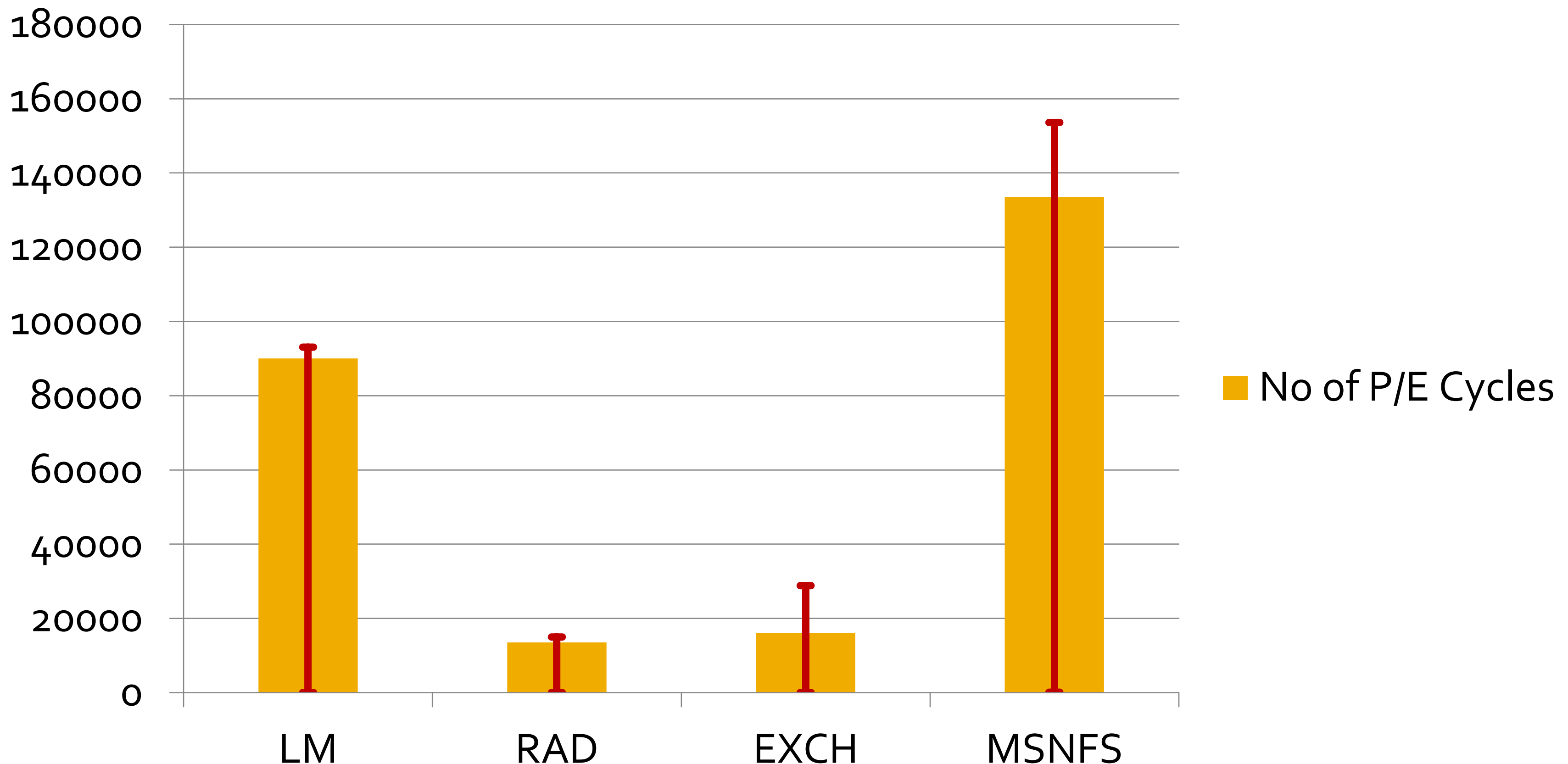
Impact of detrapping on endurance



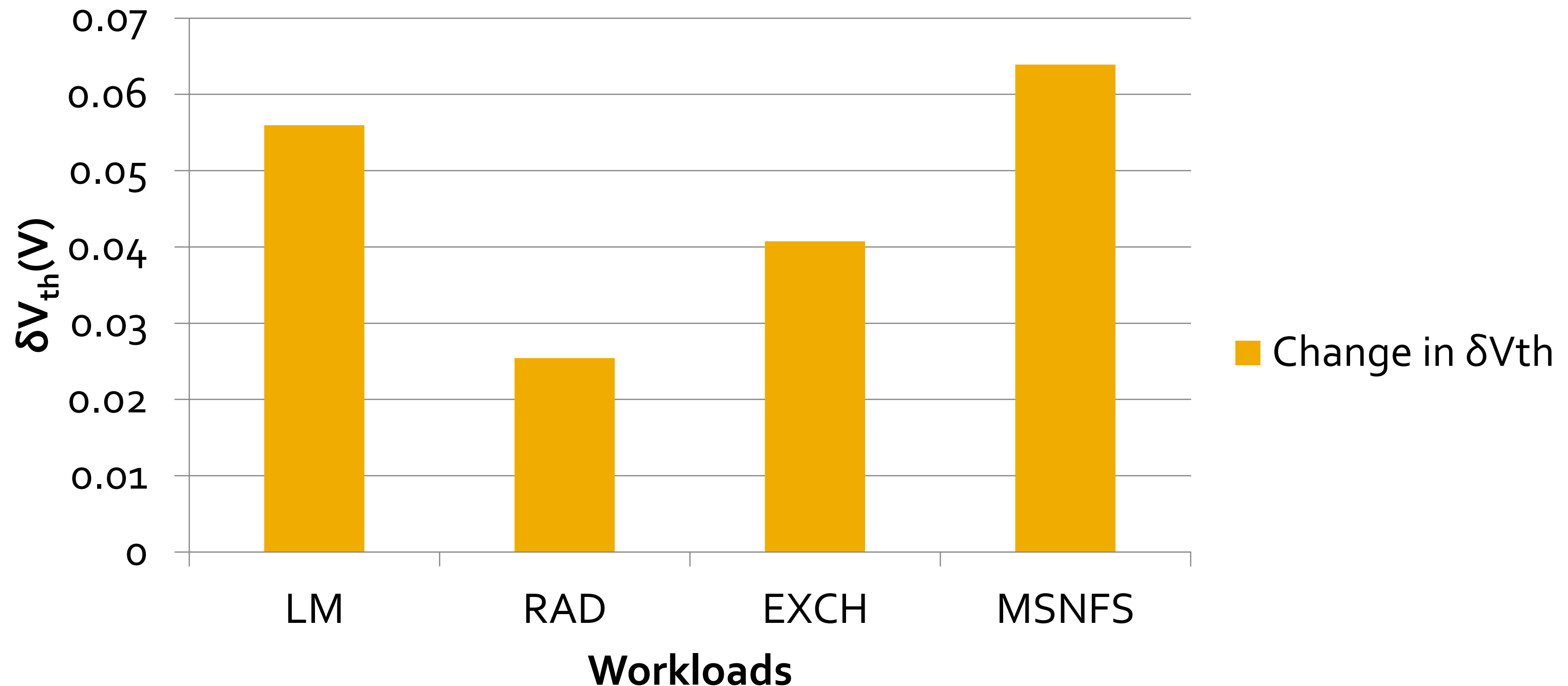
Analysis of SSD-level Endurance

- 32 GB Enterprise SSD is modeled
  - Like Intel-X25E
- Disksim –Storage System Simulator
- Enterprise workloads – from Microsoft
  - LM – Live Maps
  - RAD – Radius Authentication Server Workload
  - EXCH – Mail Exchange Server Workload
  - MSNFS – MSN File System Workload

Analysis of SSD-level Endurance



Analysis of SSD-level Endurance



Conclusion

- Contribution: Architecture level modeling tools
- FlashPower
  - Measure energy dissipated by NAND Flash chips
  - Results validated with 5 chips
  - Insights into relationship between value of bits and energy dissipation
- FENCE
  - Stress and recovery model
  - Quantify the impact of detrapping on Endurance
  - Analyze SSD-level Endurance

Future Work

- FlashPower
  - Extension to support MLC flash memory
- Validate FENCE with real chip measurements
- Analyze the impact of temperature on endurance
- Comprehensive analysis of flash memory reliability by studying
  - Bit Error rates
  - Data retention

Publications

- Vidyabhushan Mohan, Sudhanva Gurumurthi, Mircea R. Stan. *FlashPower*: A Detailed Power Model for NAND Flash Memory. Design Automation and Test in Europe (DATE) '10.
- Vidyabhushan Mohan, Taniya Siddiqua, Sudhanva Gurumurthi and Mircea R. Stan. How I Learned to Stop Worrying and Love Flash Endurance. HotStorage'10. **Under Submission.**

Acknowledgments

- Sudhanva Gurumurthi
- Mircea R. Stan
- Taniya Siddiqua
- Joe E. Brewer – University of Florida
- Koji Sakui – Intel, Tohoku University
- Laura Grupp – University of California, San Diego
- Steve Swanson – University of California, San Diego

Questions?

Backup Slides

FENCE – Stress Model

- Increase in threshold voltage due to stress - $\Delta V_{th,s}$
- $\Delta V_{th,s} = (\Delta N_{it} + \Delta N_{it}) * q / C_{ox}$
- $\Delta N_{it} = 0.08 * cycle^{0.62}$
- $\Delta N_{ot} = 5 * cycle^{0.30}$

FENCE – Recovery Model

- Decrease in threshold voltage due to recovery - $\Delta V_{th,r}$
- $\Delta V_{th,r} = 0.6 * \ln(\Delta V_{th,s}) * \ln(t)$
- **FENCE – Endurance Model**
- $\delta V_{th} = \Delta V_{th,s} - \Delta V_{th,r}$

Impact of Charge Detrapping on Endurance

Recovery Period	SLC, $\Delta V_{th} = 1.7V$		2-bit MLC, $\Delta V_{th} = 0.65V$	
	P/E Cycles	**Endurance Increase**	**P/E Cycles**	**Endurance Increase**
No recovery	107535	1x	10652	1x
10 seconds	153186	1.4x	13749	1.3x
50 seconds	1028724	9.6x	52444	4.9x
100 seconds	1837530	17.7x	99913	9.3x
1000 seconds	6214983	57.8x	403082	37.8x
5000 seconds	11093823	103x	780723	73.3x
10000 seconds	13753999	127x	990014	92.9x
15000 seconds	15497892	144x	1129379	106x
1 day	24274492	225x	1879352	176x
2 days	28487539	264x	2247910	211x

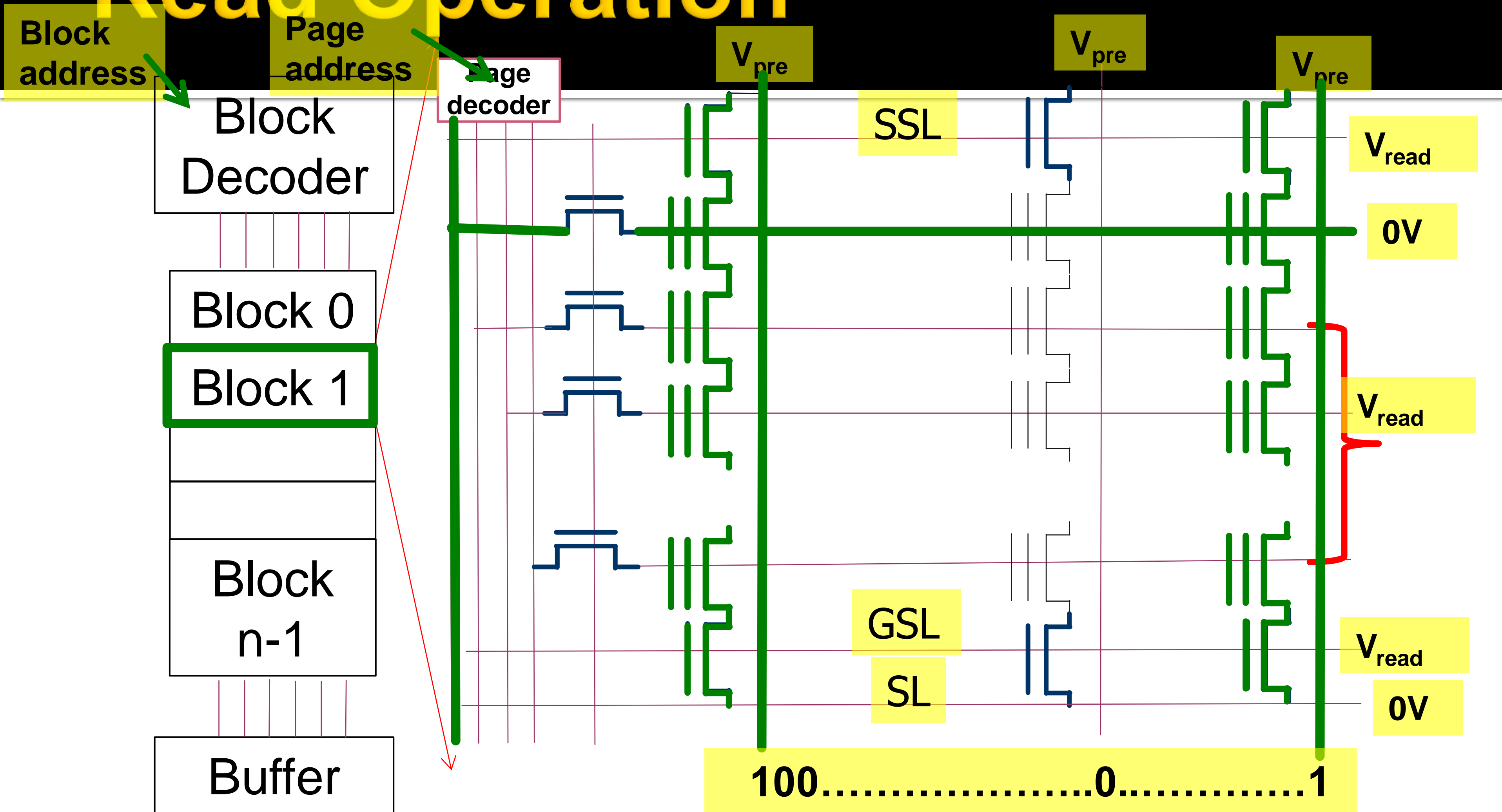
Table 5.1: Endurance limits with charge detrapping.

Recovery time distribution

Benchmarks	Recovery Period (seconds)			
	[1k-5k)	[5k-10k)	[10k-15k)	[15k-20k)
LM	100	0	0	0
RAD	0.001	0.0023	99.9957	0.001
EXCH	0.002	99.857	0.141	0
MSNFS	100	0	0	0

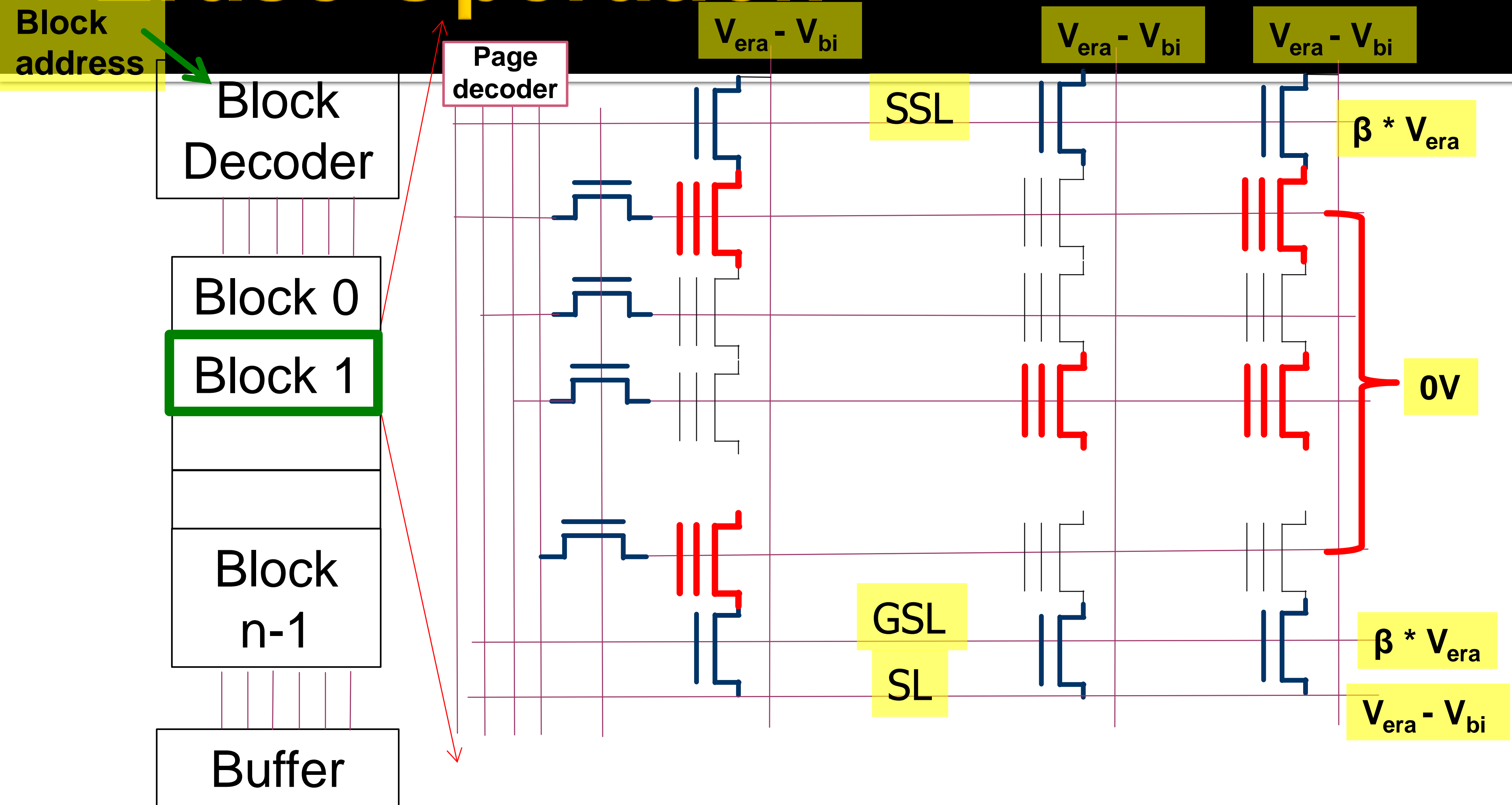
Table 6.2: Recovery time distributions. $[X, y)$ indicates recovery periods of duration t where $X \leq t < y$.

Read Operation



$$E_{read}=E_{dec,r}+E_{selected\text{-}page,r}+E_{unselected\text{-}page,r}+E_{sl}+E_{bl\text{-}1,r}+E_{buffer}+E_{iopins}+E_{r\text{-}pre}$$

Erase Operation



$$E_{erase} = E_{dec,erase} + E_{sl} + E_{well} + E_{bl,e} + N_{bits\text{-}0} * E_{tunnel,mc} + E_{verify\text{-}erase} + E_{r\text{-}pre}$$

Device Specific Parameters - Assumptions

- Bias voltages for read, program and erase
- Doping level of NAND block P-well and N-well
- Number of partial program cycles
- Capacitance of I/O pins – 5pF

Reference: J.E. Brewer and M. Gill, editors. *Nonvolatile Memory Technologies with Emphasis on Flash. IEEE Press, 2008*.

Inputs to Flash Power

- Required Inputs – represents micro-architecture of a Flash chip
  - Feature size of FGTs
  - Page size – data area + spare area
  - Pages per block
  - Blocks per plane – can be 2 dimensional
  - Operating voltage of the chip
- Optional Inputs
  - Read, program and erase and pre-charge voltages

NAND Flash Memory – Erase