Flash Memory Overview

Steven Swanson

Welcome to the Data Age

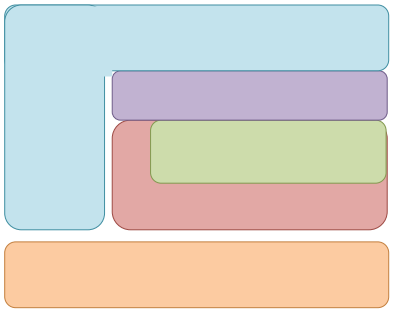
- The world processed 9 Zettabytes of data in 2008*
- Acquiring data is easy
- Extracting knowledge is hard
  - Storage performance is major bottleneck
  - Solid-state storage can help

*http://hmi.ucsd.edu

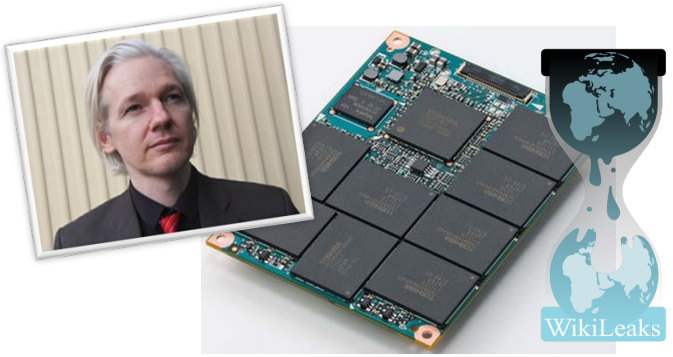
NVSL

Non-volatile Systems Laboratory

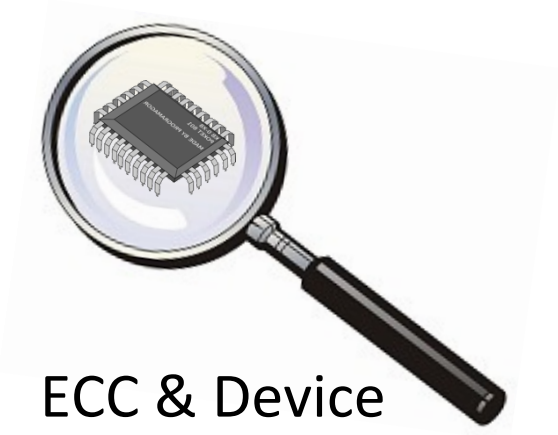
Hardware/Software Prototyping



Programming interfaces

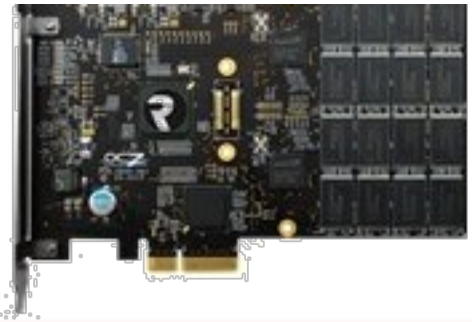


Data Security

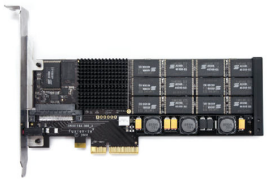
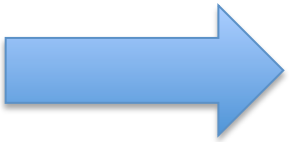


ECC & Device Characterization

The Flash Juggernaut

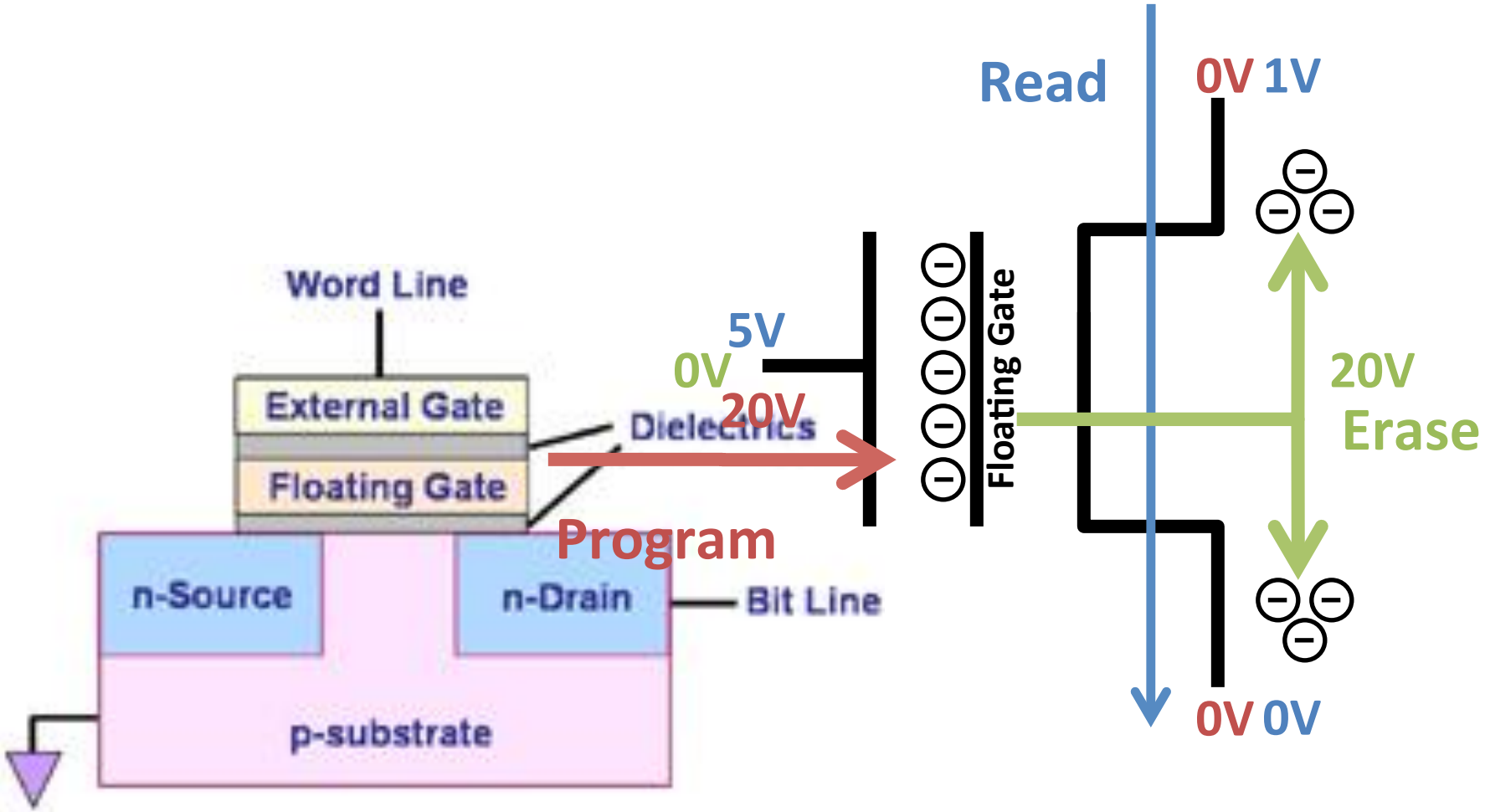


Flash is Fast!

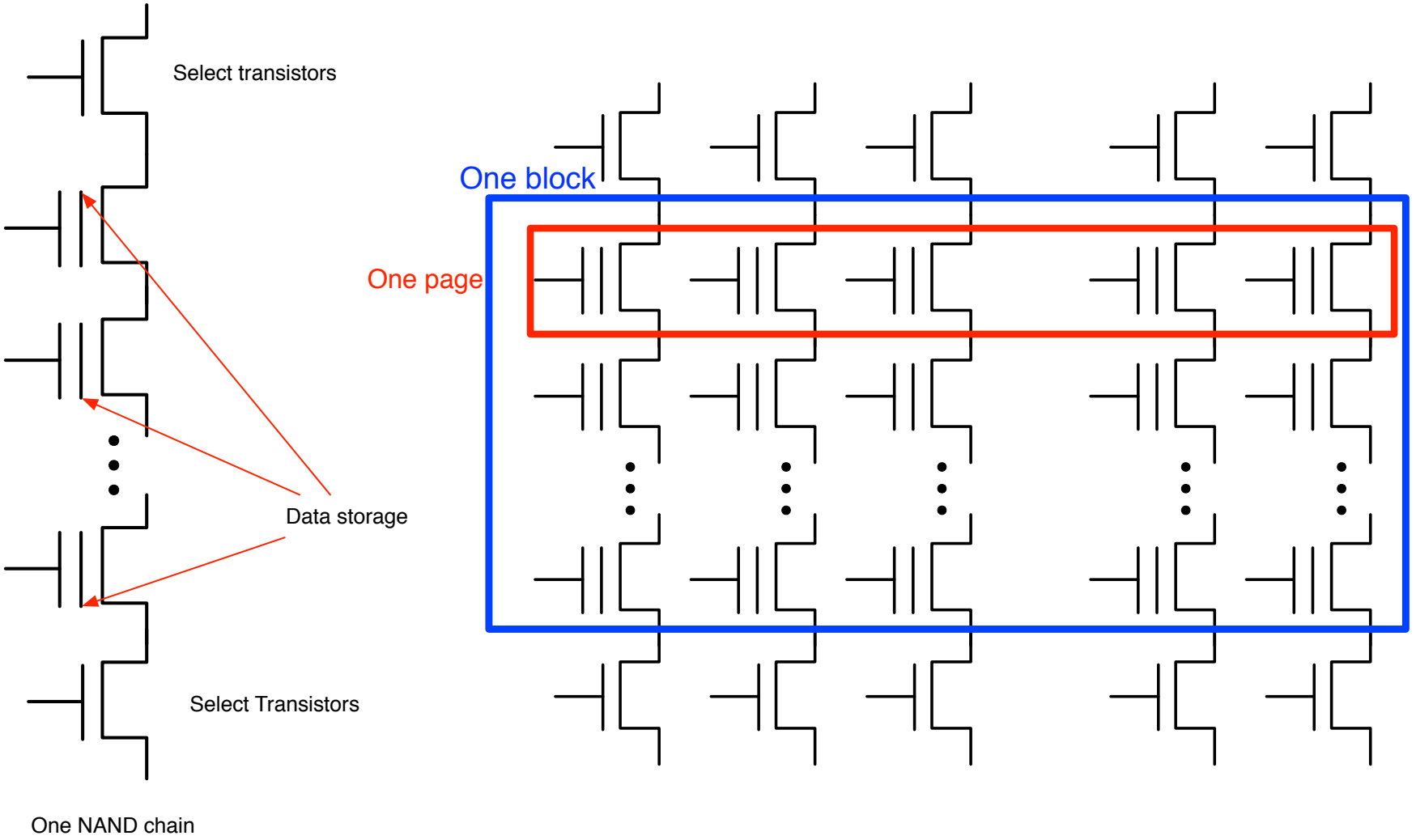
Hard Drives		PCIe-Flash 2007
		
Lat.: 7.1ms		68us
BW: 2.6MB/s		250MB/s
1x		104x
1x		96x

- Random 4KB Reads from user space

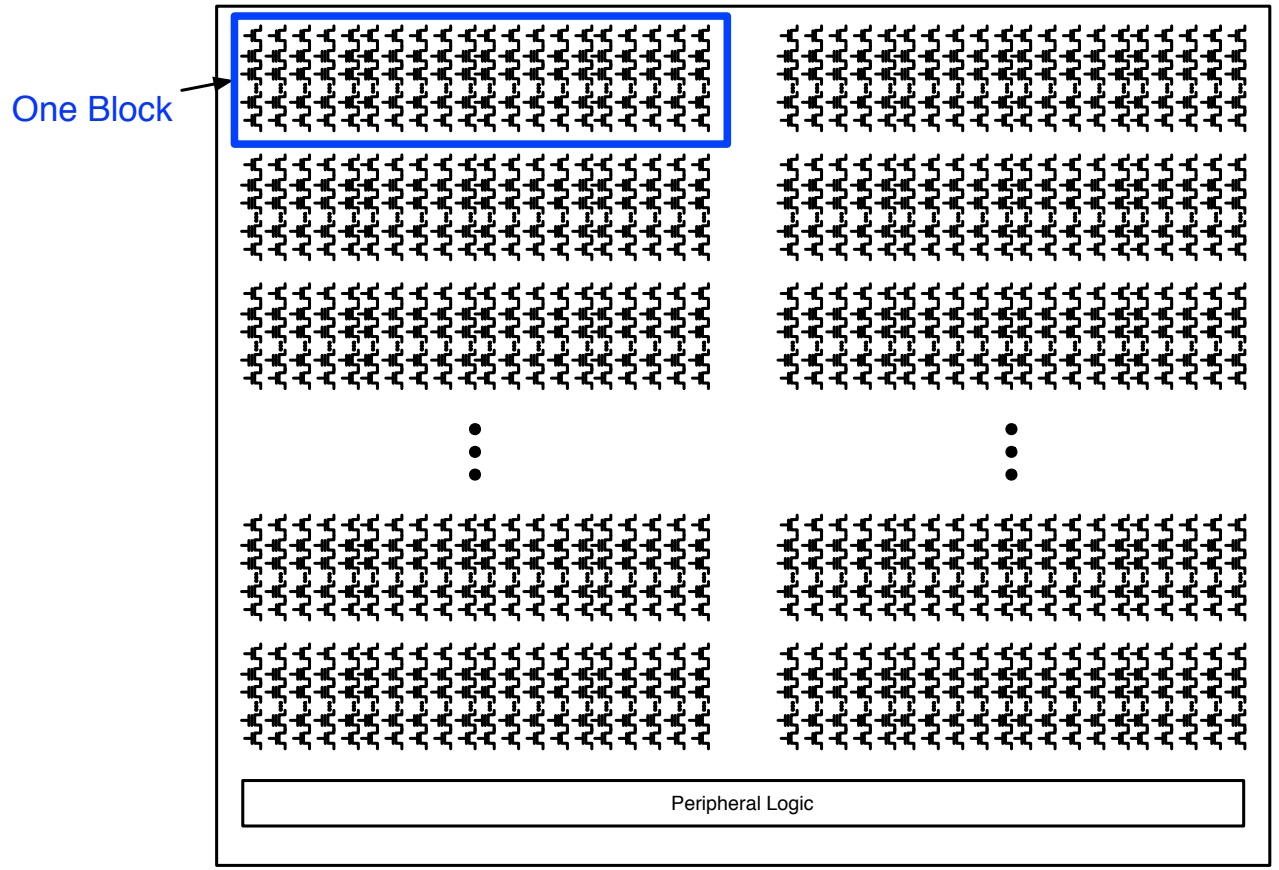
Flash Operations



Organizing Flash Cells into Chips

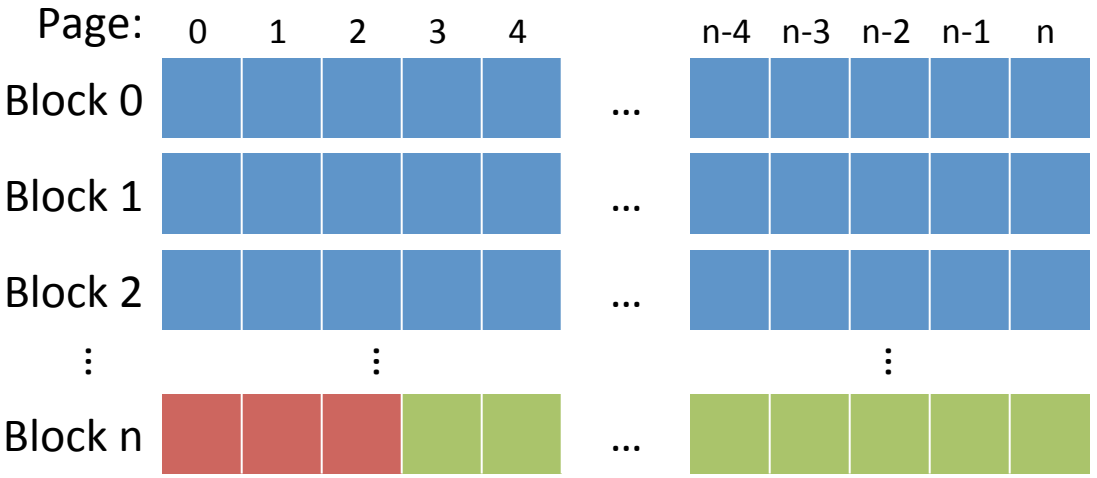


Organizing Flash Cells into Chips

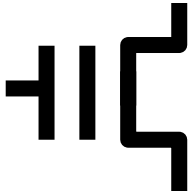


- ~16K blocks/chip
- ~16-64Gbits/chip

Flash Operations

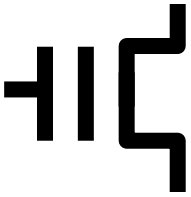


SLC: Single Level Cell



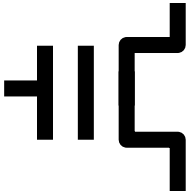
== 1 bit

MLC: Multi Level Cell



== 2 bits

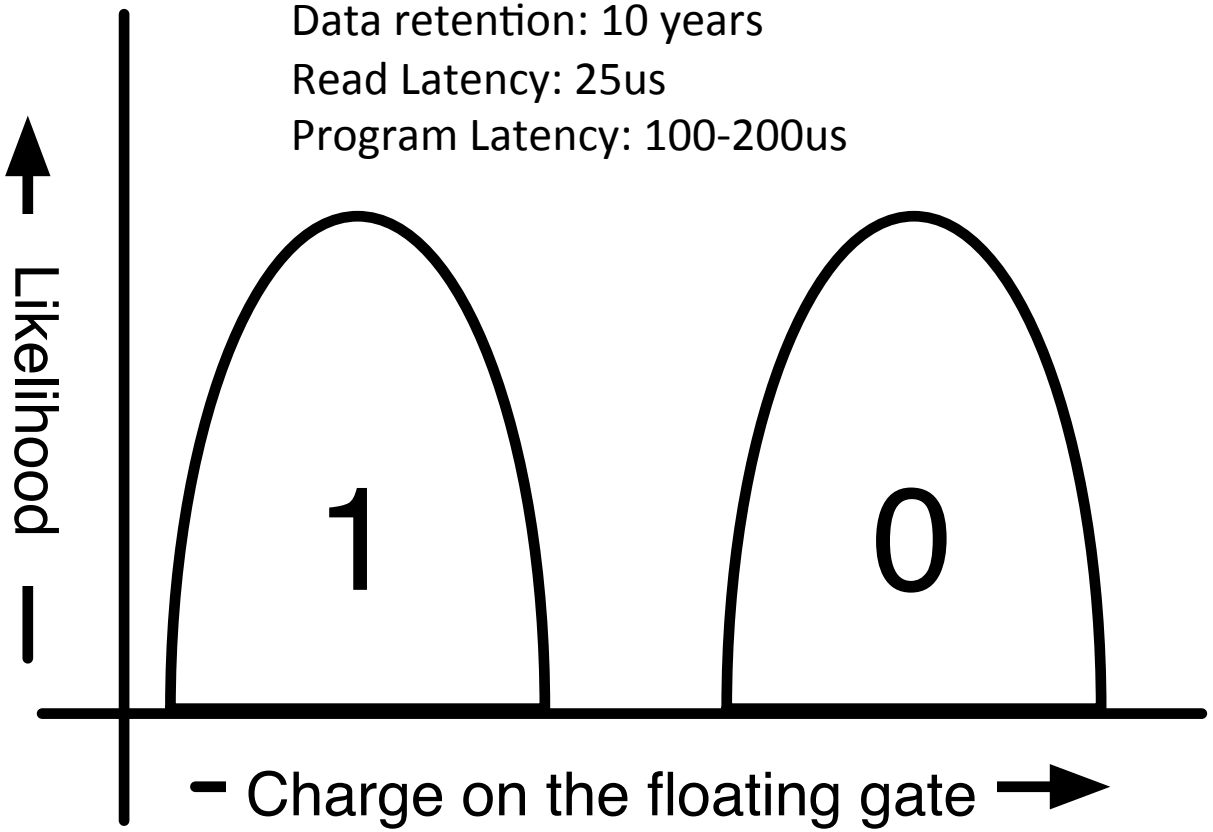
TLC: Triple Level Cell

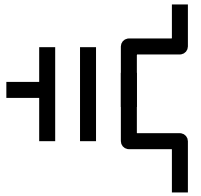


== 3 bits

Single-Level Cell

Endurance: 100,000 Cycles
Data retention: 10 years
Read Latency: 25us
Program Latency: 100-200us

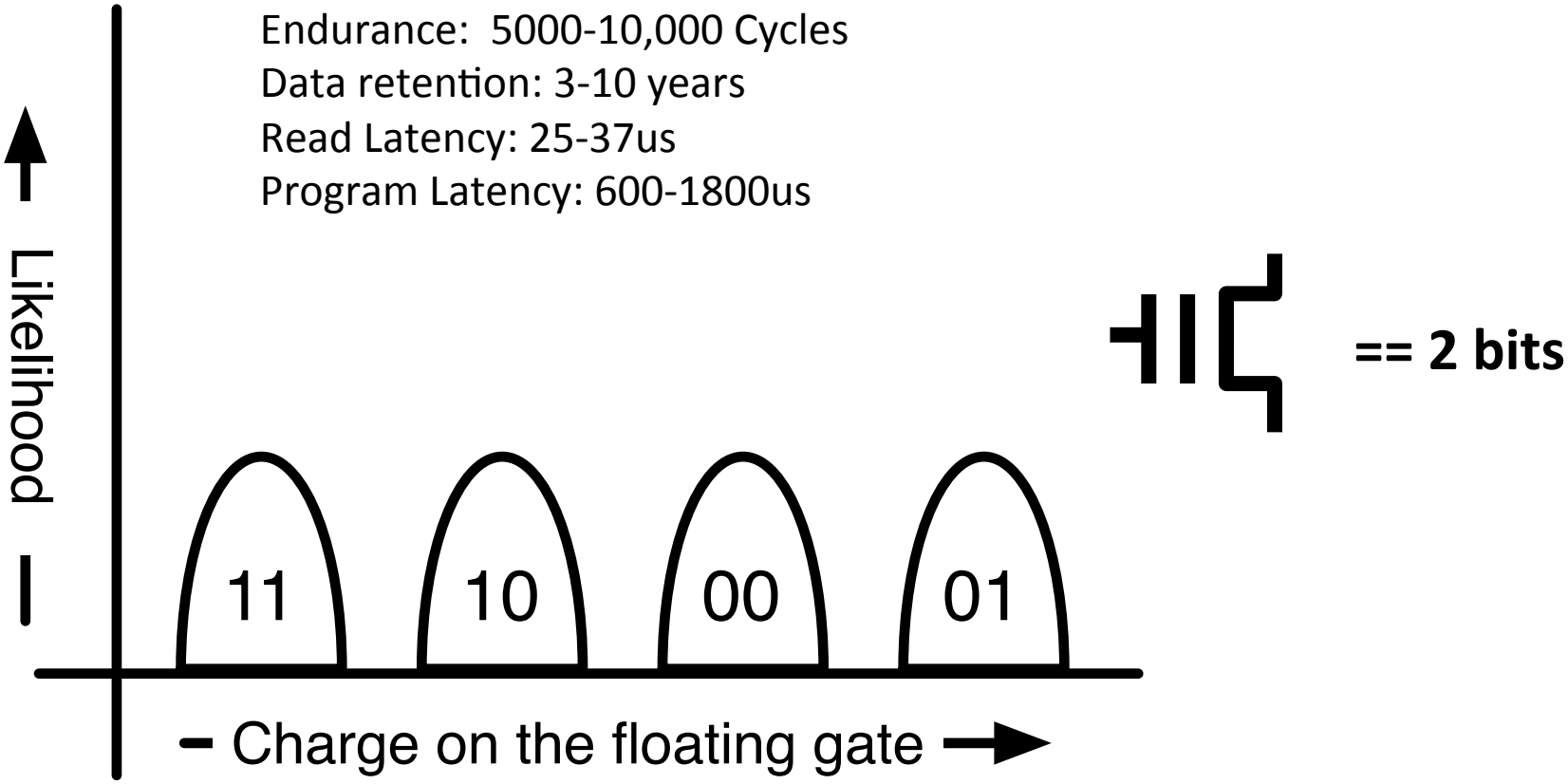




== 1 bit

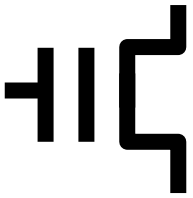
Multi-Level Cell (2 bits)

Endurance: 5000-10,000 Cycles
Data retention: 3-10 years
Read Latency: 25-37us
Program Latency: 600-1800us



Triple-level Cell (3bits)

Endurance: ~500-1000 Cycles
Data retention: 3 years
Read Time: 60-120us
Program Time: 500-6500us

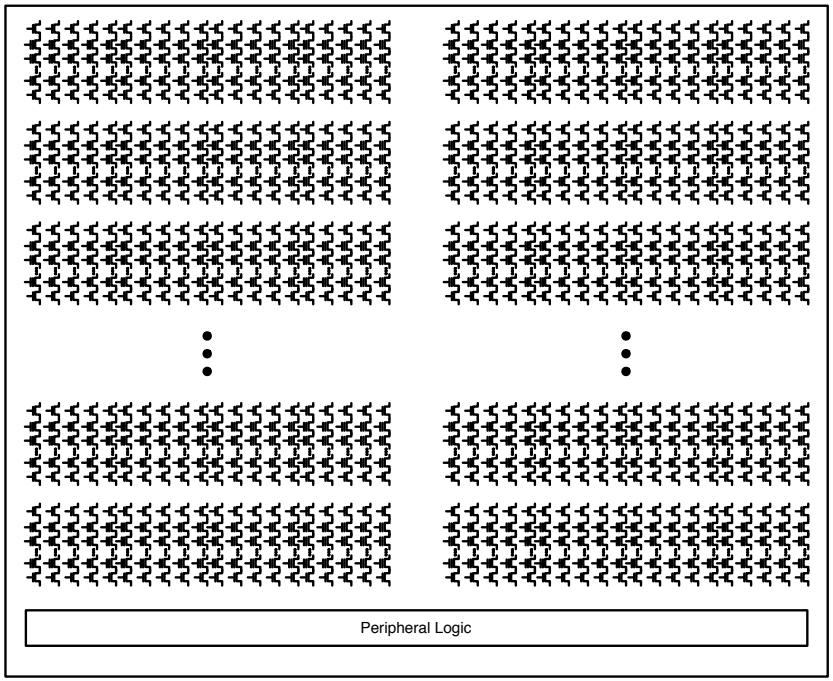


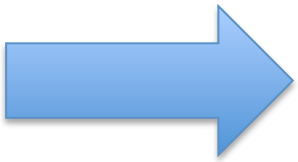
== 3 bits

Flash Failure Mechanisms

- Program/Erase (PE) Wear
  - Permanent damaged to the gate oxide at each flash cell
  - Caused by high program/erase voltages
  - Damage causes charge to leak off the floating gate
- Program disturb
  - Data corruption caused by interference from programming adjacent cells.
  - No permanent damage

Making Disks out Flash Chips

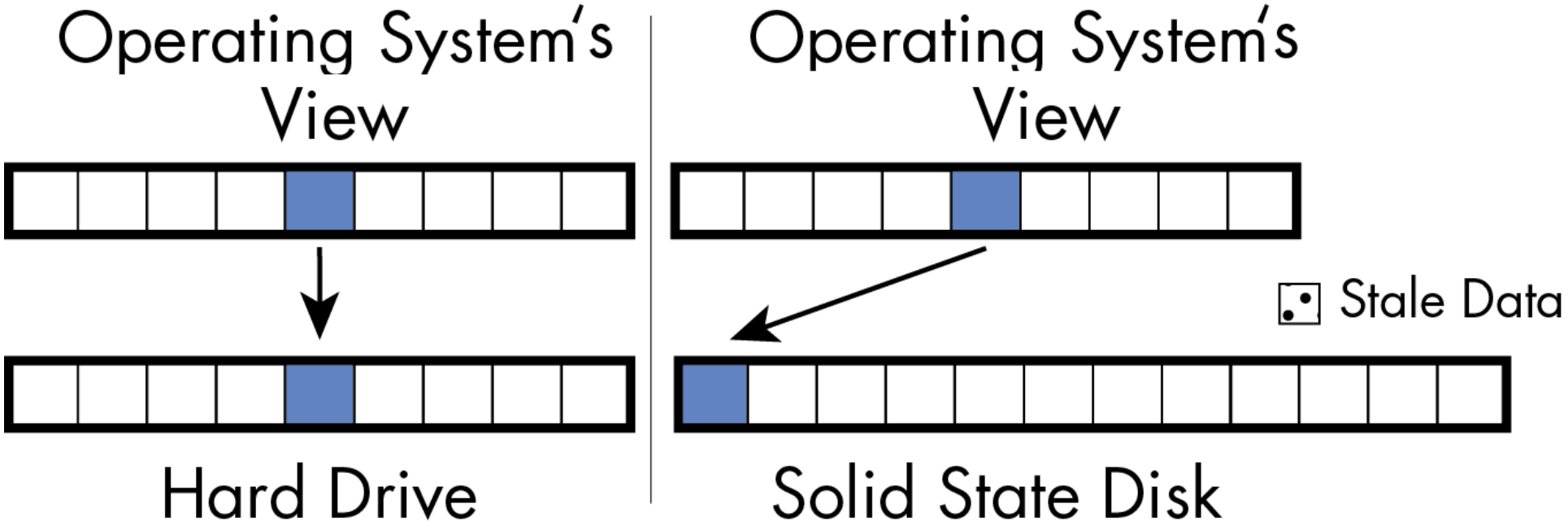




Read Pages
Write Pages
Erase Blocks
Hierarchical addresses
PE Wear

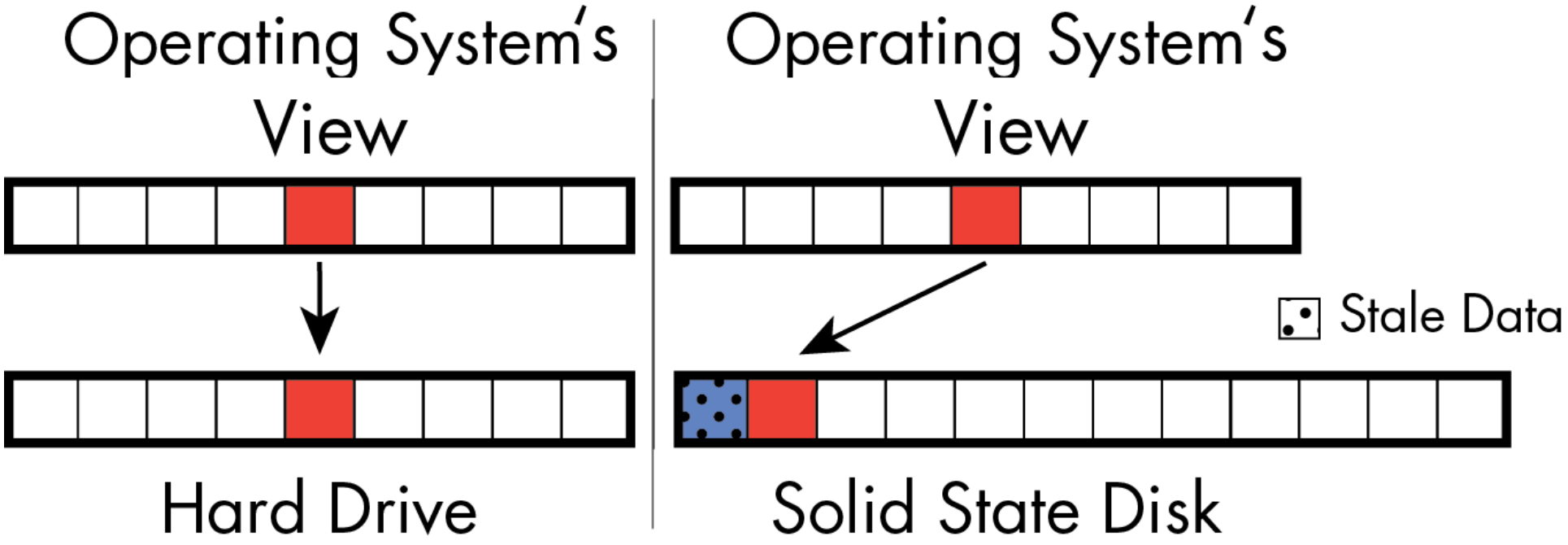
Read
Write
Flat address space
No wear limitations

Writing Data



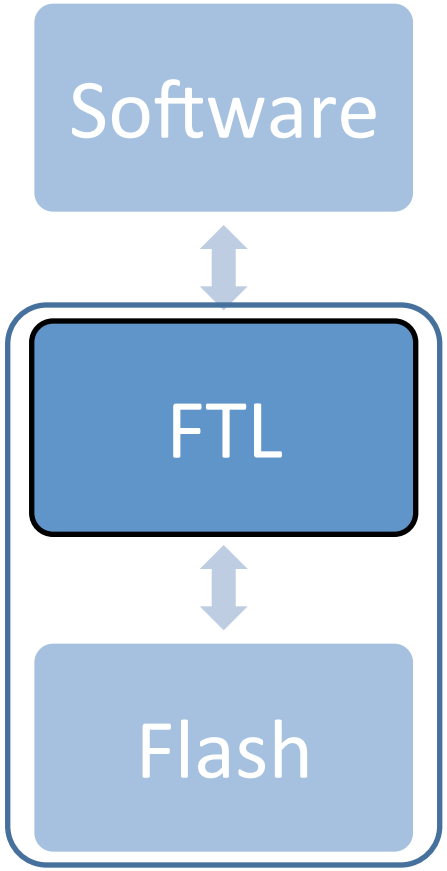
SSD Maintain a map between "virtual" logical block addresses and "physical" flash locations.

Writing more data…



When you overwrite data, it goes to a new location.

Flash Translation Layer (FTL)



User

- Logical Block Address

Flash

- Write pages in order
- Erase/Write granularity
- Wears out

FTL

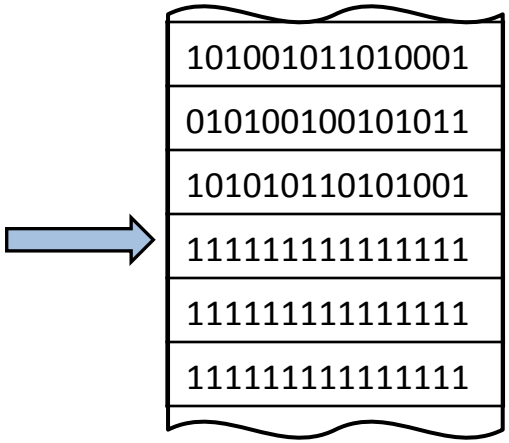
- Logical → Physical map
- Wear leveling
- Power cycle recovery

Centralized FTL State

Map

LBA	Physical Page Address	
0	Block 5	Page 7
2k	Block 27	Page 0
4k	Block 10	Page 2

Write Point

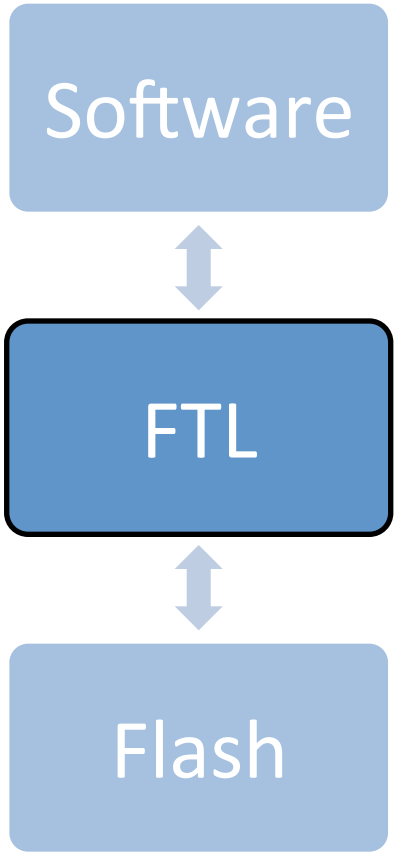


Next Sequence Number: 12

Block Info Table

Block	Erased	Erase Count	Valid Page Count	Sequence Number	Bad Block Indicator
0	False	3	15	5	False
1	True	7	0	-	False
2	False	0	4	9	False

Read



1. Read Data at LBA 2k

2. Map

LBA	Physical Page Address	
0	Block 5	Page 7
2k	Block 27	Page 0
4k	Block 10	Page 2

3. Flash Operation

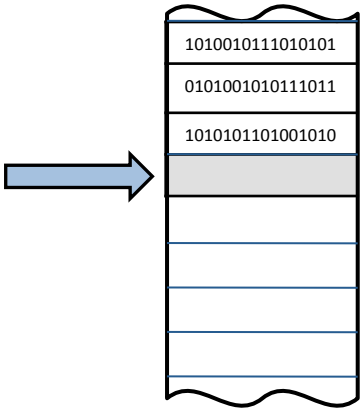
Write – Mid Block

Write 0101101011001010 **to LBA 2k**

Write Point = Block 2, Page 5

Map

LBA	Physical Page Address	
0	Block 5	Page 7
2k	Block 0	Page 0
4k	Block 10	Page 2



Next Sequence Number: 12

Block Info Table

Block	Erased	Erase Count	Valid Page Count	Sequence Number	Bad Block Indicator
0	False	3	15	5	False
1	True	7	0	-	False
2	False	0	4	9	False

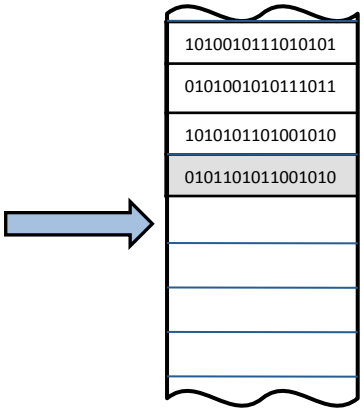
Write – Mid Block

Write 0101101011001010 **to LBA 2k**

Write Point = Block 2, ~~Page 5~~ Page 6

Map

LBA	Physical Page Address	
0	Block 5	Page 7
2k	Block ~~0~~ 2	Page ~~0~~ 5
4k	Block 10	Page 2



Next Sequence Number: 12

Block Info Table

Block	Erased	Erase Count	Valid Page Count	Sequence Number	Bad Block Indicator
0	False	3	~~15~~ 14	5	False
1	True	7	0	-	False
2	False	0	~~4~~ 5	9	False

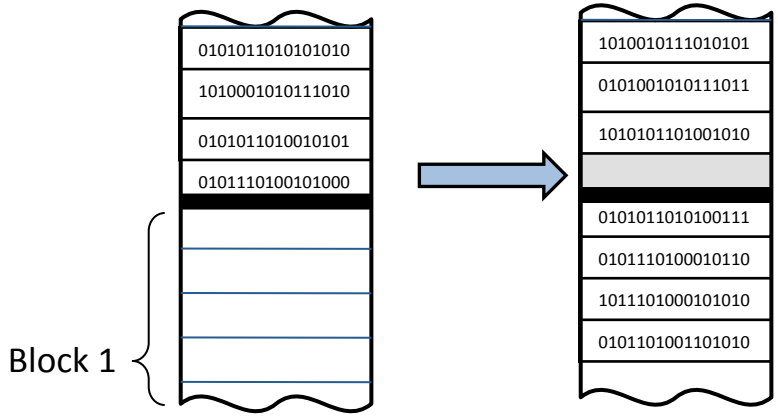
Write – Block Jump (1)

Write 0101001010100110 **to LBA 2k**

Write Point = Block 2, Page 63

Map

LBA	Physical Page Address	
0	Block 5	Page 7
2k	Block 0	Page 5
4k	Block 0	Page 2



Next Sequence Number: 12

Block Info Table

Block	Erased	Erase Count	Valid Page Count	Sequence Number	Bad Block Indicator
0	False	3	15	5	False
1	True	7	0	-	False
2	False	0	4	9	False

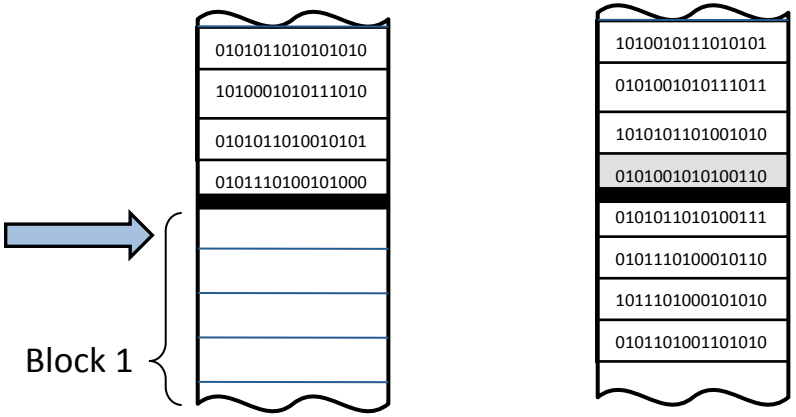
Write – Block Jump (1)

Write 0101001010100110 **to LBA 2k**

Write Point = ~~**Block 2, Page 63**~~
Block 1, Page 0

Map

LBA	Physical Page Address	
0	Block 5	Page 7
2k	Block ~~0~~ 2	Page ~~5~~ 63
4k	Block 0	Page 2



Next Sequence Number: 12

Block Info Table

Block	Erased	Erase Count	Valid Page Count	Sequence Number	Bad Block Indicator
0	False	3	~~15~~ 14	5	False
1	True	7	0	-	False
2	False	0	~~4~~ 5	9	False

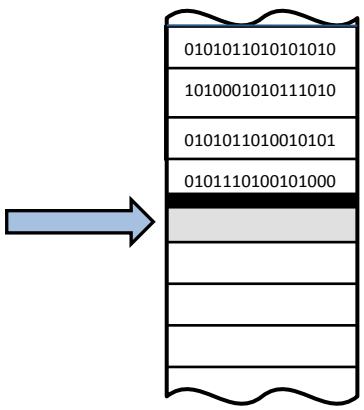
Write – Block Jump (2)

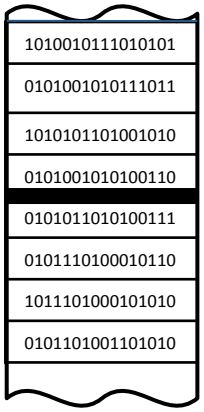
Write 1101000101101001 **to LBA 4k**

Write Point = Block 1, Page 0

Map

LBA	Physical Page Address	
0	Block 5	Page 7
2k	Block 2	Page 63
4k	Block 0	Page 2





Next Sequence Number: 12

Block Info Table

Block	Erased	Erase Count	Valid Page Count	Sequence Number	Bad Block Indicator
0	False	3	14	5	False
1	True	7	0	-	False
2	False	0	5	9	False

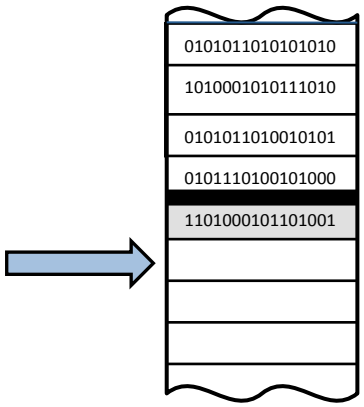
Write – Block Jump (2)

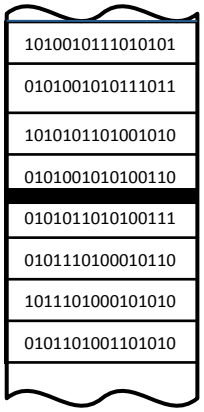
Write 1101000101101001 **to LBA 4k**

Write Point = Block 1, ~~Page 0~~ Page 1

Map

LBA	Physical Page Address	
0	Block 5	Page 7
2k	Block 2	Page 63
4k	Block ~~0~~ 1	Page ~~2~~ 0





Next Sequence Number: ~~12~~ 13

Block Info Table

Block	Erased	Erase Count	Valid Page Count	Sequence Number	Bad Block Indicator
0	False	3	~~14~~ 13	5	False
1	~~T~~ F	7	~~0~~ 1	12	False
2	False	0	5	9	False

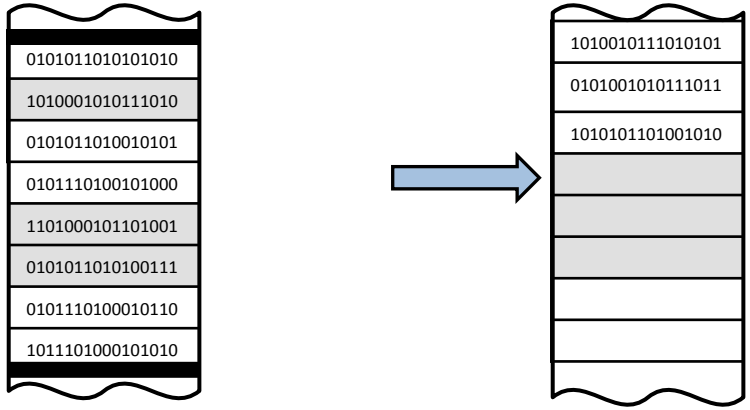
Erase

Block Info Table

Block	Erased	Erase Count	Valid Page Count	Sequence Number	Bad Block Indicator
0	False	3	13	5	False
1	False	7	1	12	False
2	False	0	3	9	False

Move Valid Pages

Block 2

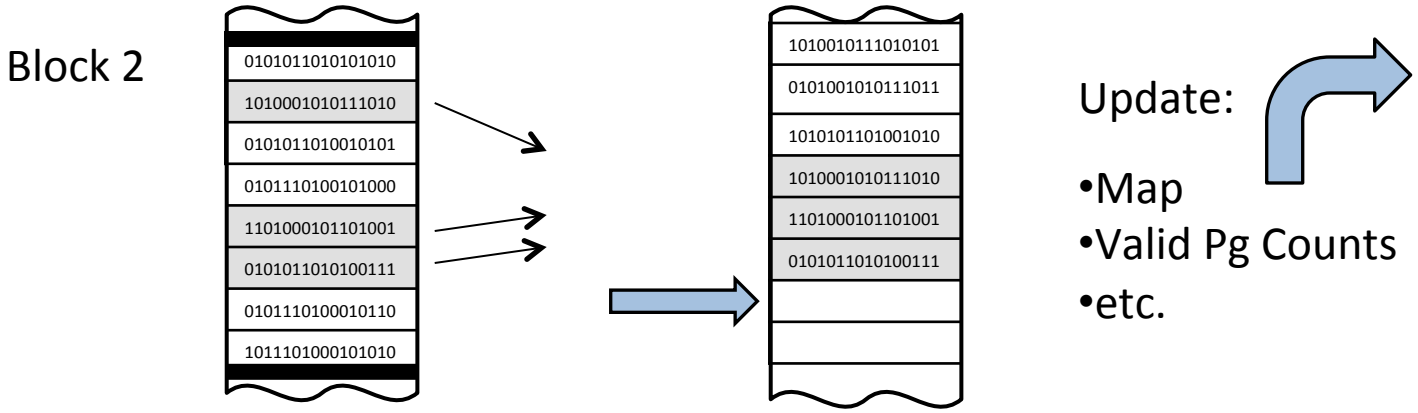


Erase

Block Info Table

Block	Erased	Erase Count	Valid Page Count	Sequence Number	Bad Block Indicator
0	False	3	13	5	False
1	False	7	1	12	False
2	False	0	~~3~~ 0	9	False

Move Valid Pages



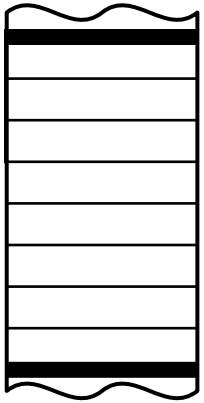
Erase

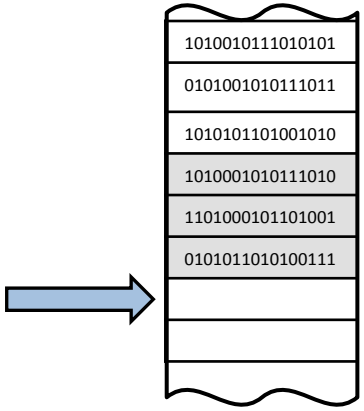
Block Info Table

Block	Erased	Erase Count	Valid Page Count	Sequence Number	Bad Block Indicator
0	False	3	13	5	False
1	False	7	1	12	False
2	~~F~~ T	~~0~~1	0	-	False

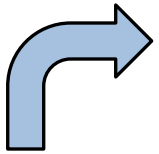
Move Valid Pages

Block 2





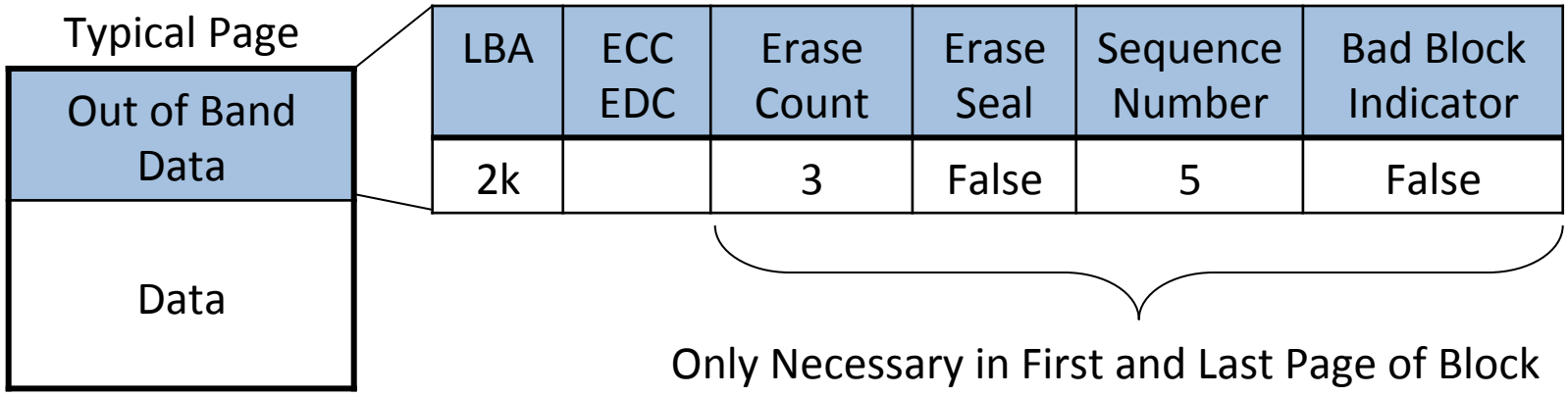
Update:



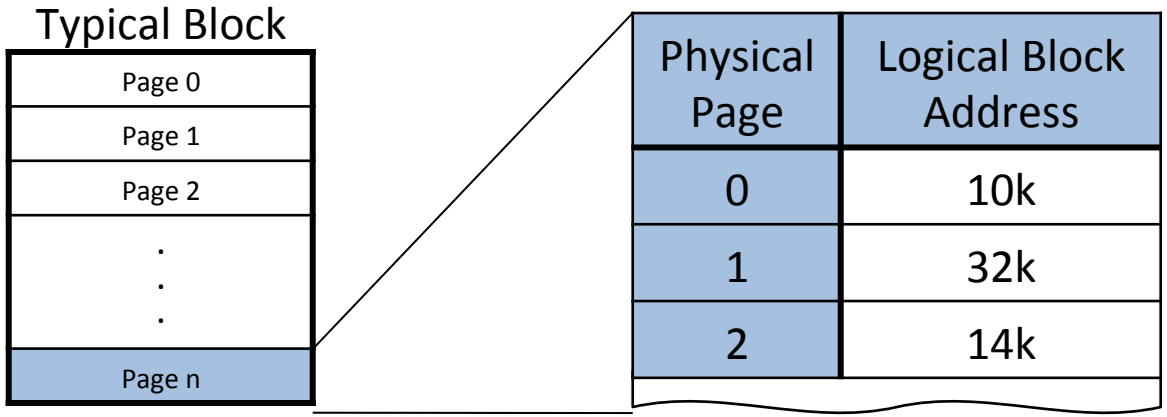
- Map
- Valid Pg Counts
- etc.

Distributed FTL State

Metadata



Summary Page



Power Cycle

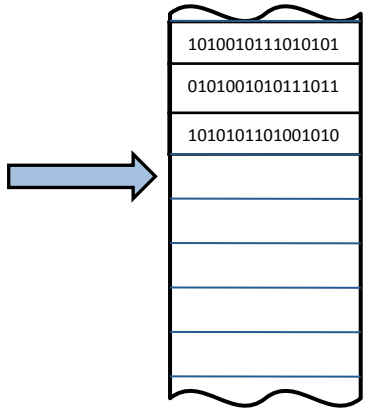
Scan each block:

1. Summary page
2. First Page
3. All Pages

Map

LBA	Physical Page Address	
0	Block 5	Page 7
2k	Block 27	Page 0
4k	Block 10	Page 2

Write Point



Block Info Table

Next Sequence Number: 12

Block	Erased	Erase Count	Valid Page Count	Sequence Number	Bad Block Indicator
0	False	3	15	5	False
1	True	7	0	0	False
2	False	0	4	9	False