Demystifying the SSD

Its limitations, Usage and Benefits

Scott Stetzer
Director, Enterprise SSD

Scott Shadley
Sr. Technical Manager

SSD Marketing
STEC, Inc.

Agenda

Content focus will be:

1. Understanding NAND behaviors and limitations

2. Managing the SSD to ensure consistent performance and reliability

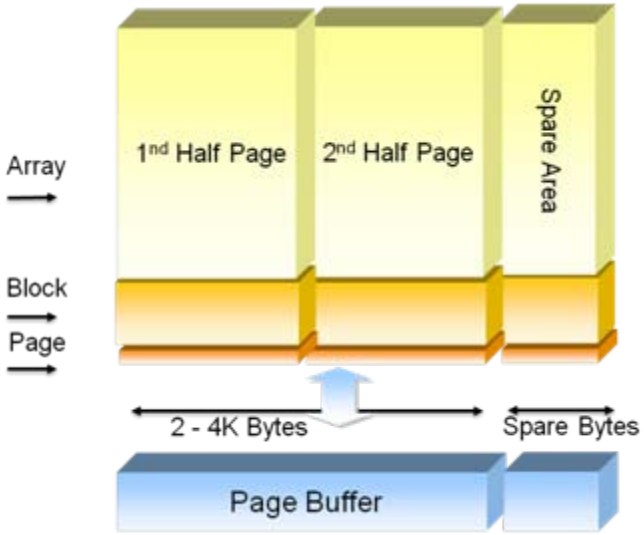
3. Understanding the notable difference between SSD grades and application usage.

NAND Behaviors

The Basics…

Blocks, Pages and Planes

- With a smallest writable size of 2-4KB SSDs have immediate media related complexity
- With the smaller “re-writable” size of a block, even more work must be done
- NAND while inexpensive and fast, can be very complex to work with in high performance systems



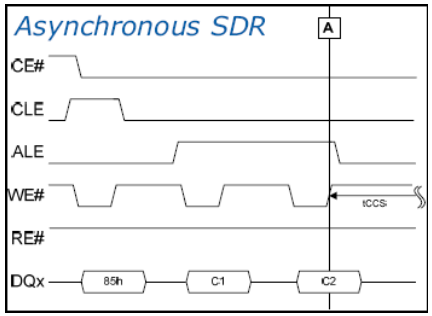
NAND Behaviors

The interface changes…

- Complexity increases with the move to new faster interfaces
- Running with old and new interfaces forces new controllers

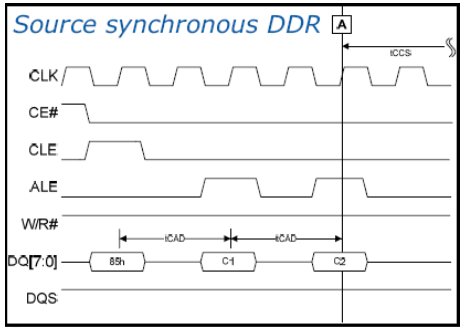
- Running Asynchronous NAND allows more liberal control

- Running the new Synchronous interfaces tightens windows



Async runs up to 50Mbs

Sync can run up to 200Mbs



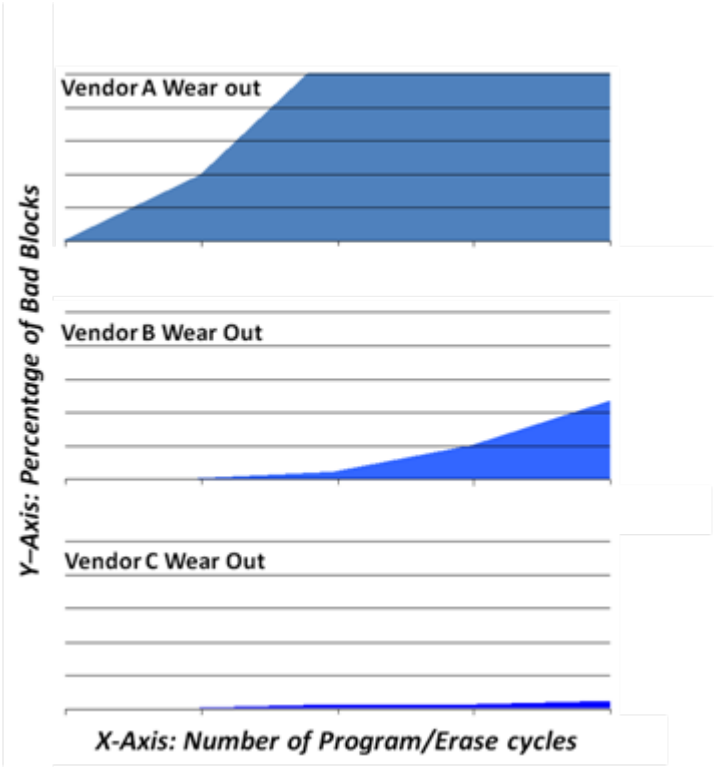
NAND Limitations

The Endurance numbers…

- Moving from HDD and mechanical issues to SSD with "hard" limits on writing can be very complex

- Not all vendors show the same wear levels on raw NAND

- As geometry shrinks so do Endurance and Reliability



NAND Limitations

Retaining Customer Data…

- Raw NAND retention is inversely proportional to cycles
- NAND media types also have different wear out factors
- How long is good enough for Enterprise SSDs??

MLC MEDIA Retention				
Bake Hours @125C	RdVerify@ 4hrs	RdVerify@ 8hrs	RdVerify@ 12hrs	RdVerify@ 24hrs
Equivalent years @ 50C	~0.8 yrs	**~1.6 yrs @10k+ cycles**	~2.4 yrs	~4.8 yrs

MLC Media @ current node

SLC Media Retention			
Bake Hours @125C	RdVer@ 24hrs	RdVer@ 48 Hrs	RdVer@ 72 Hrs
Equivalent years @ 50C	~4.8 yrs	~9.5 yrs	**~14.3 yrs @ 100K+ cycles**

SLC Media @ current node

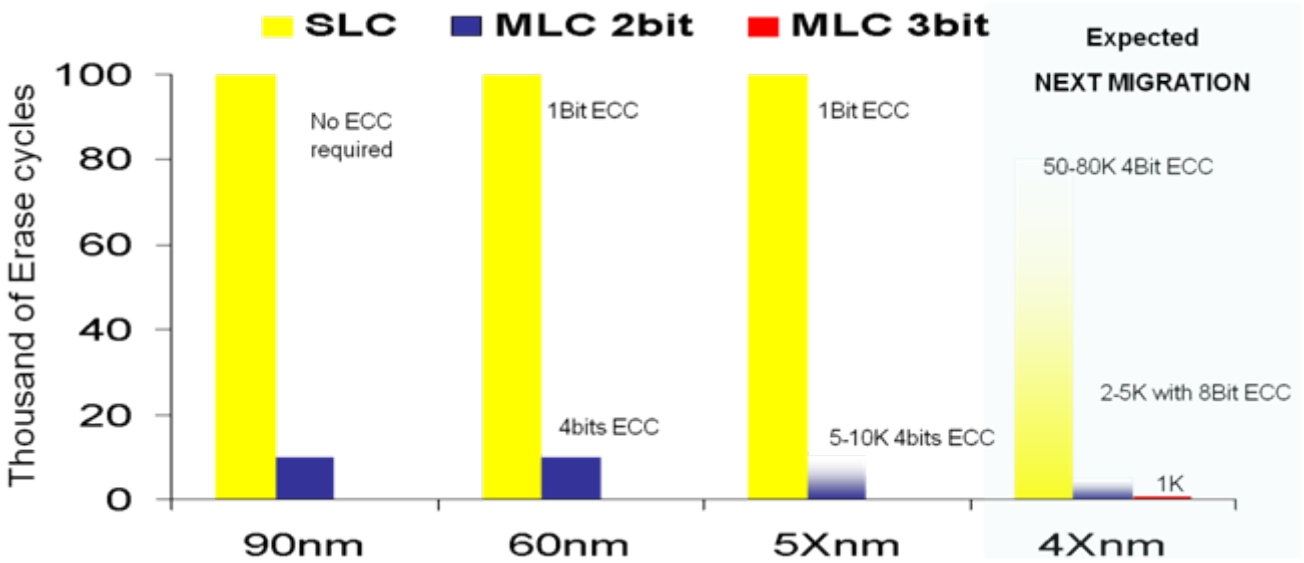
NAND Limitations

The nm equation…

Moving from generation to generation is not a simple task

Constraints imposed by the NAND increase controller needs

As geometry shrinks so do the NAND requirements



Agenda

Content focus will be:

1. Understanding NAND behaviors and limitations

2. Managing the SSD to ensure consistent performance and reliability

3. Understanding the notable difference between SSD grades and application usage.

Why is there so much difference

HDD performance has always been gated,

The fastest HDDs can only sustain about 350 IOPS

SSDs have opened the gates, and the current models can sustain over 40,000 IOPS

01011001100101001010010 1

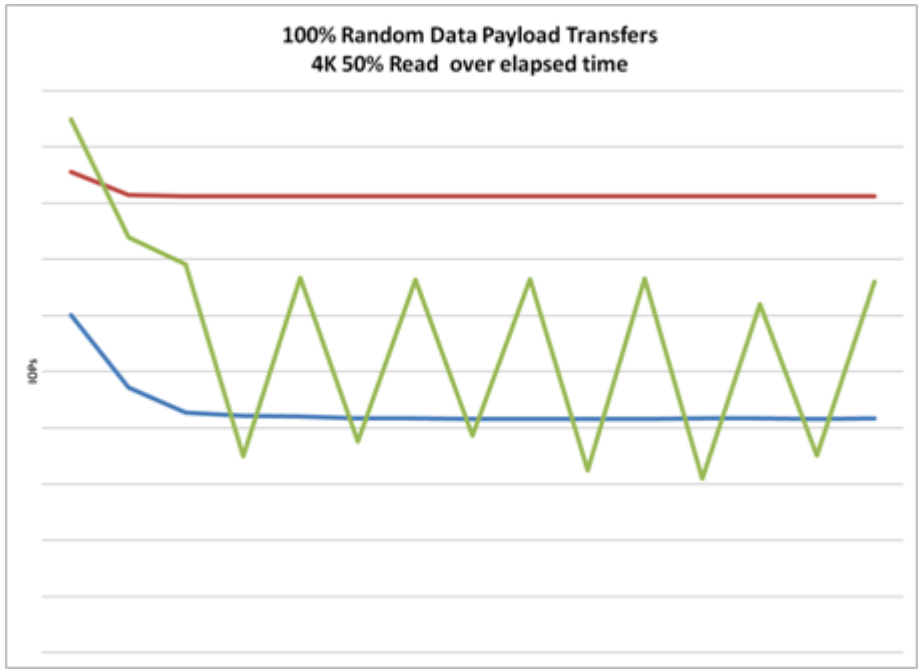
1100110101100110110110 10

Workload Modeling

Is there a model that works for SSDs

- The concern is how to make these valuable
- Which test do you run?
- Sequential or Random Data?
- How long do you run them?

See a live demo in Booth 511

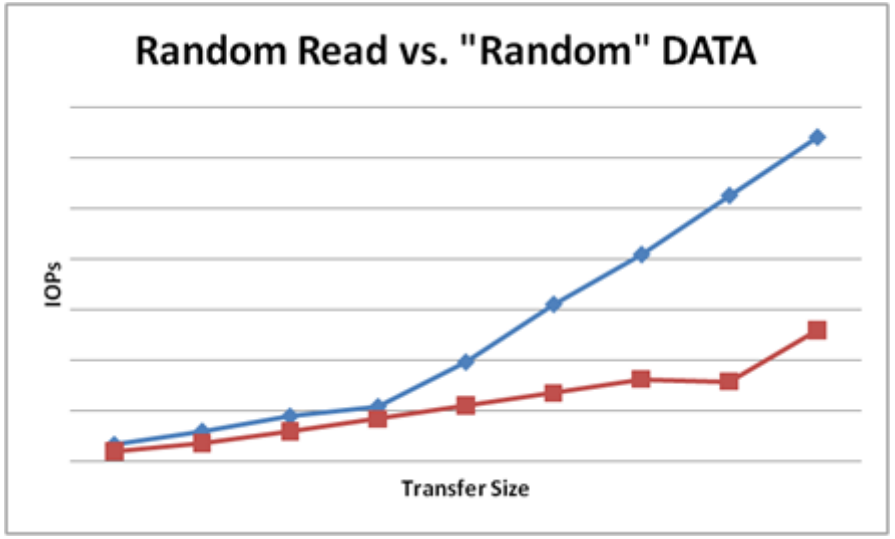


What is the Right Workload – the Program

IOMeter – The current choice of many companies

How do you use the tool the right way?

What is random about zeros?



Managing the SSD – Consistency

Access patterns affect the final number

SSD Bathtubs do exist

How big the tub is the real question

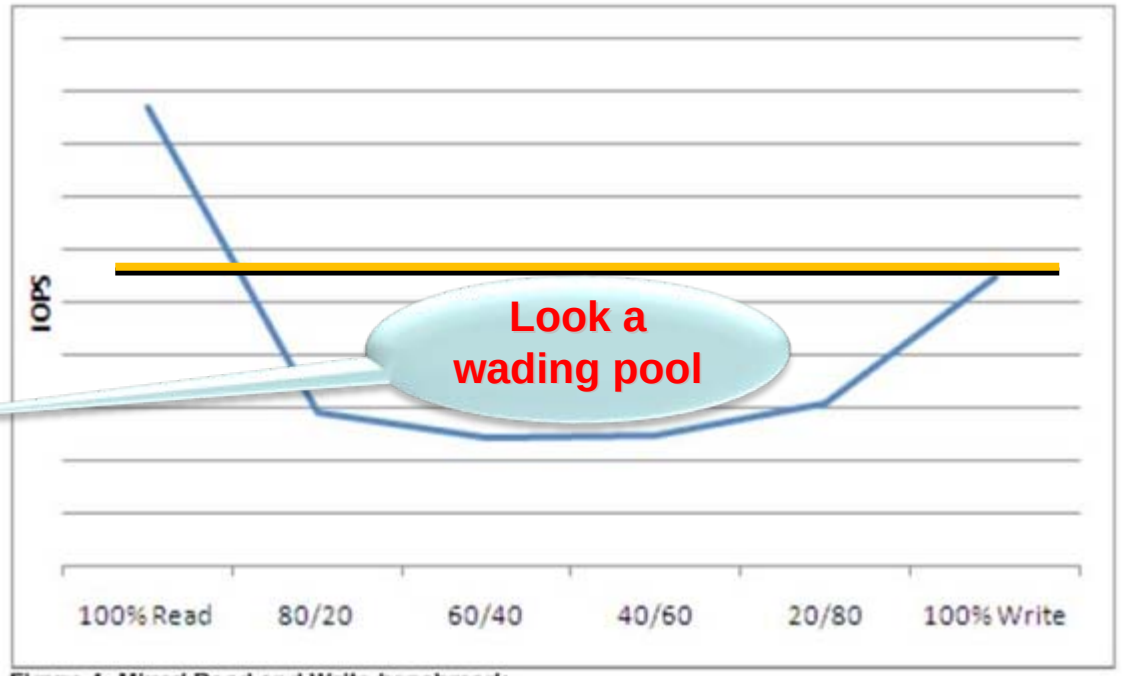


Figure 4. Mixed Read and Write benchmark

Managing the SSD – payload performance

Payload size from the host to the SSD affects top performance

Data Sheet numbers tend to focus on the end caps

What about pre-conditioning?

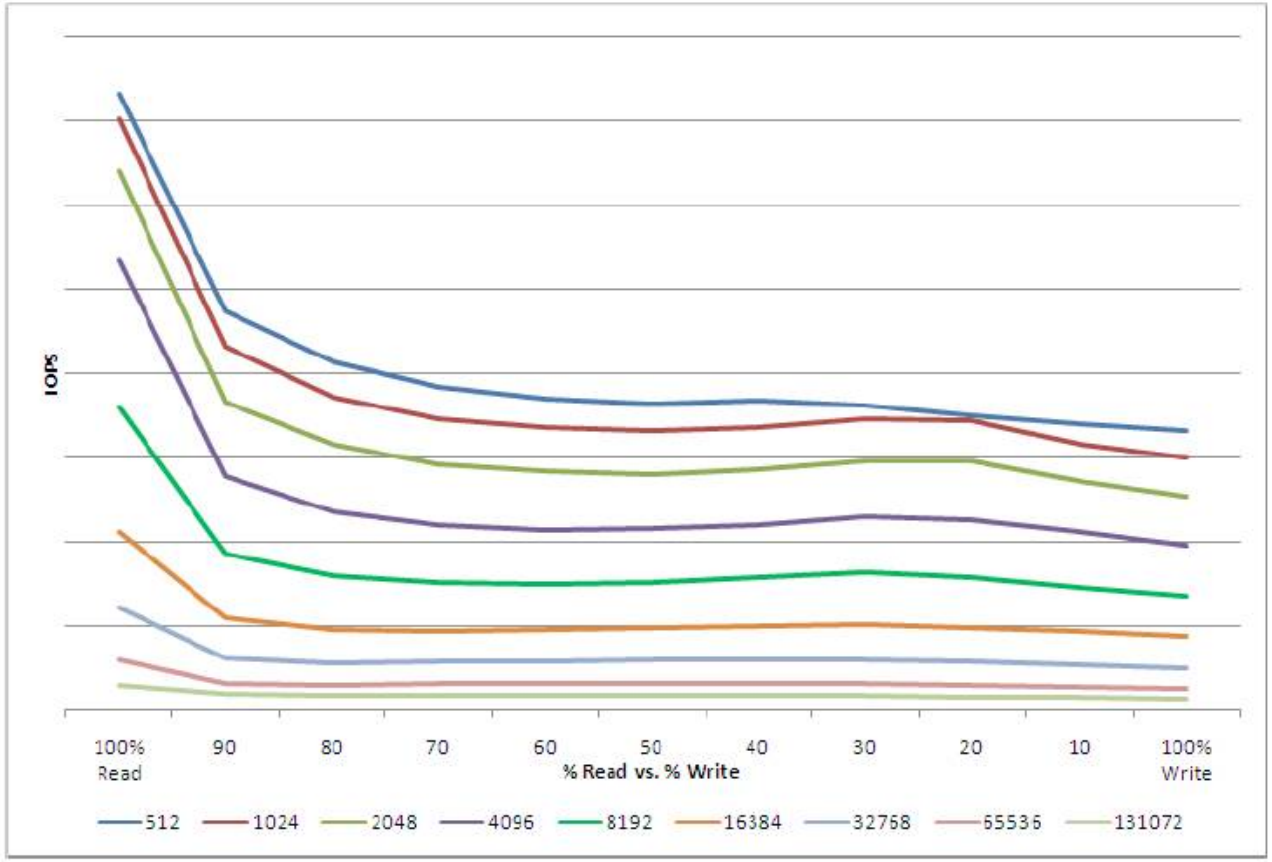


Figure 6. Benchmark test at varied block sizes and Queue Depth of 32

Managing the SSD – Pre-conditioning

Benchmarks can show great performance OOB

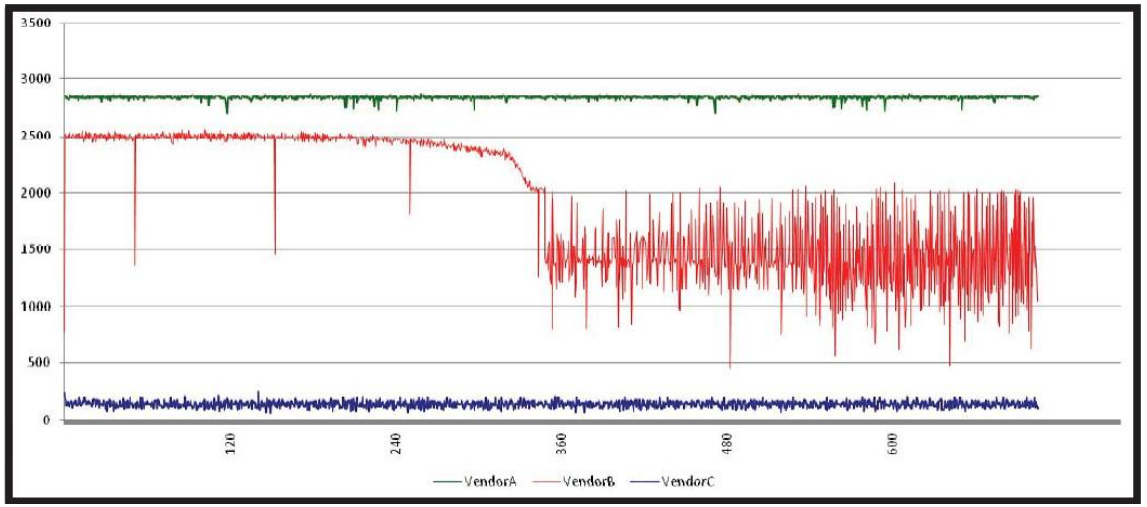
What happens to the numbers over time?

What is the right method of conditioning?

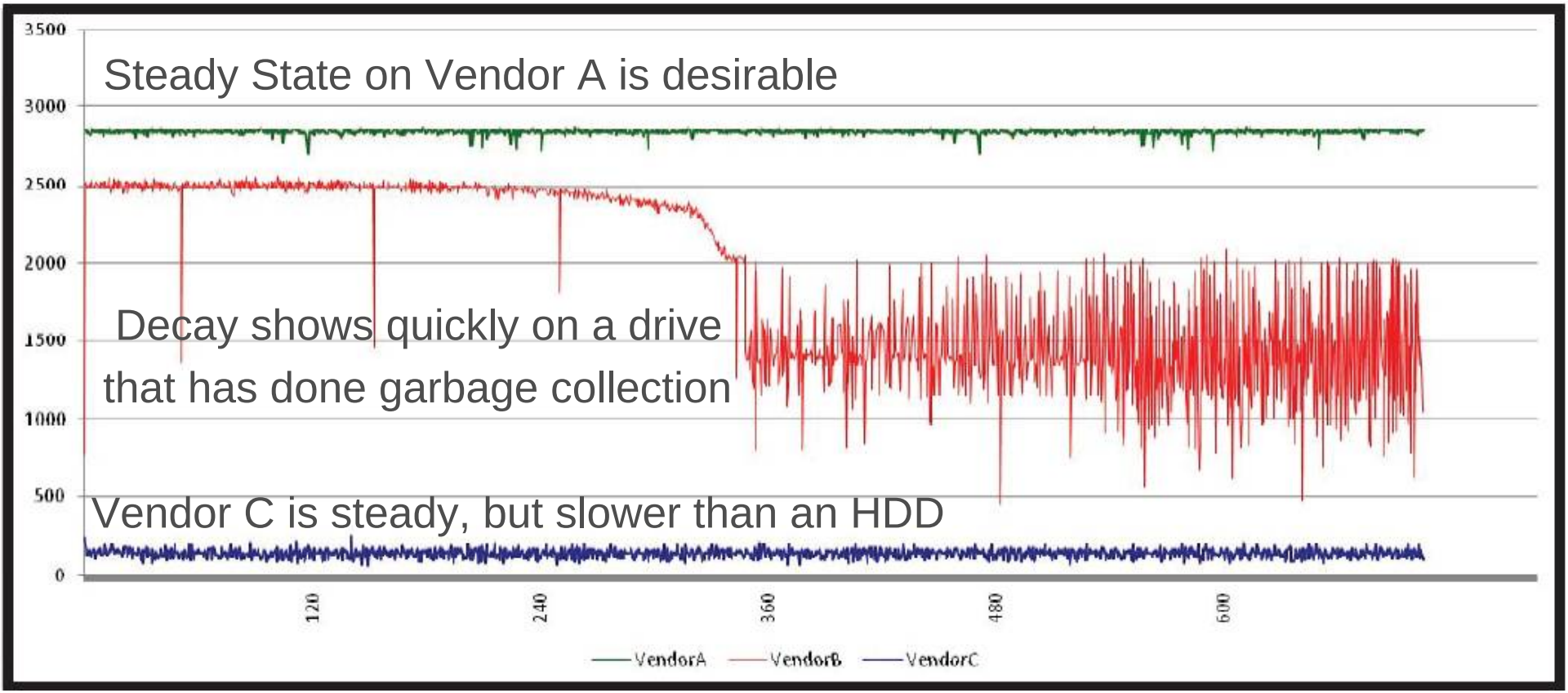
Methods in progress with the standards bodies

JEDEC - JC64.8

SNIA - SSSI



Managing the SSD – Pre-conditioning



Managing the SSD – The terms of use

A closer Look part 1

<u>Wear Leveling</u>

What it really does

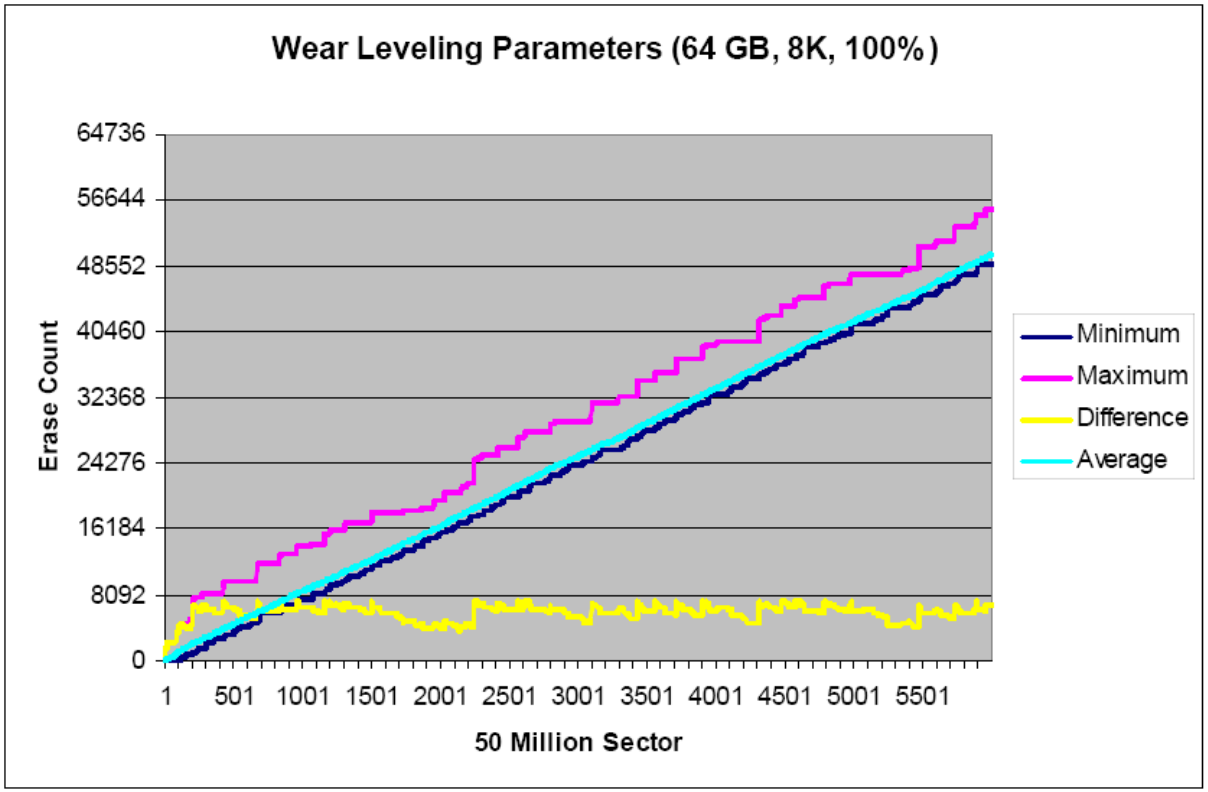
One example of data progression in the drive over time and region

Dirty Page	0-7	8-15	16-23	24-31	32-39	40-47	48-55	56-63
WL#0	29	22	16	11	7	4	2	1
WL#1	37	30	23	17	12	8	5	3
WL#2	44	38	31	24	18	13	9	6
WL#3	50	45	39	32	25	19	14	10
WL#4	55	51	46	40	33	26	20	15
WL#5	59	56	52	47	41	34	27	21
WL#6	62	60	57	53	48	42	35	28
WL#7	64	63	61	58	54	49	43	36

Managing the SSD – The terms of use

A closer look part 2

Keys are to ensure the SSD does not amplify writes

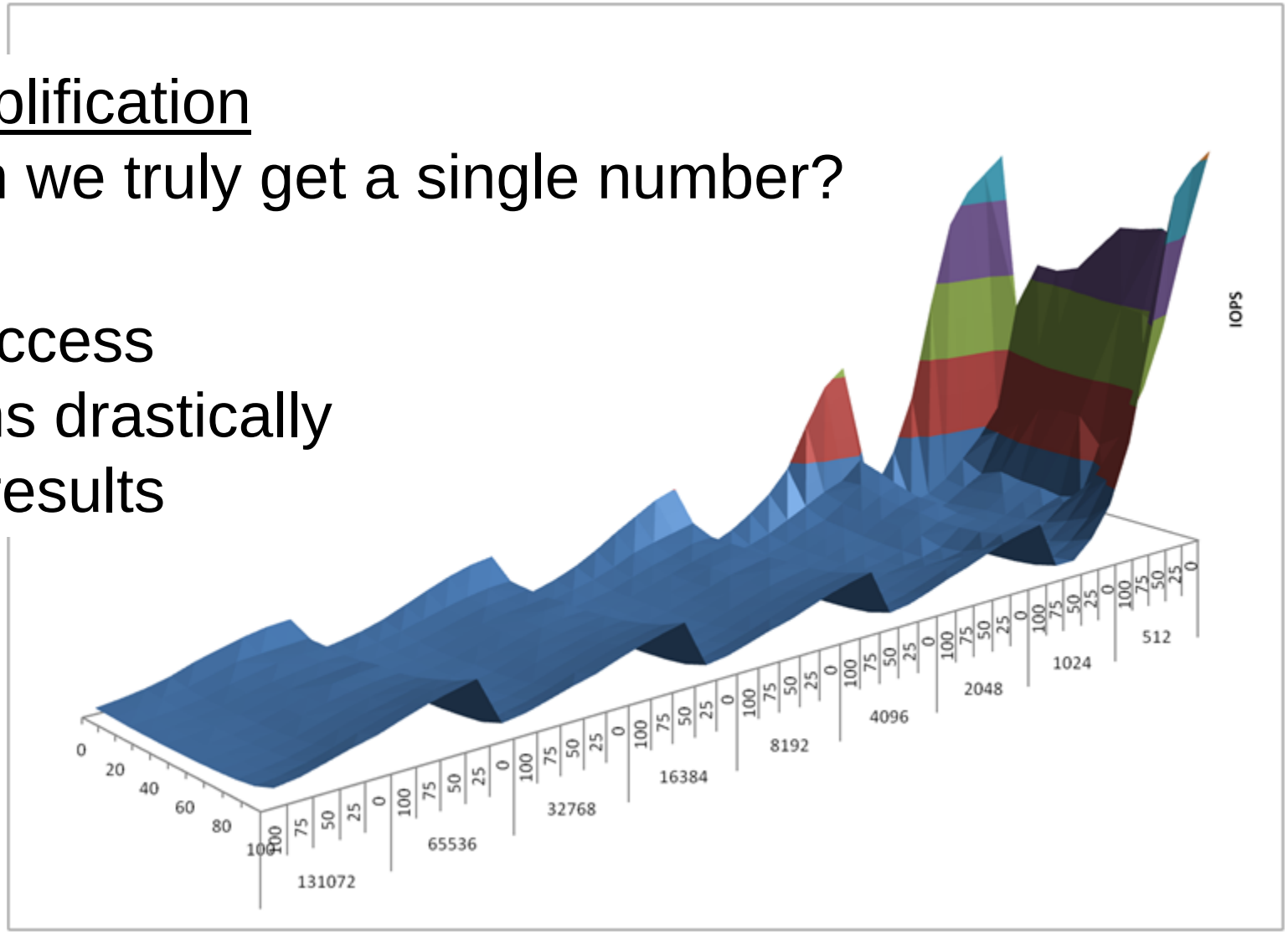


Managing the SSD – The terms of use

Write Amplification

Can we truly get a single number?

User access patterns drastically affect results



Agenda

Content focus will be:

1. Understanding NAND behaviors and limitations

2. Managing the SSD to ensure consistent performance and reliability

3. Understanding the notable difference between SSD grades and application usage.

What is done to make all this happen

SSD controllers are where the magic is

Let’s build an SSD based on all we have discussed and see where that drive belongs

Interface of choice?

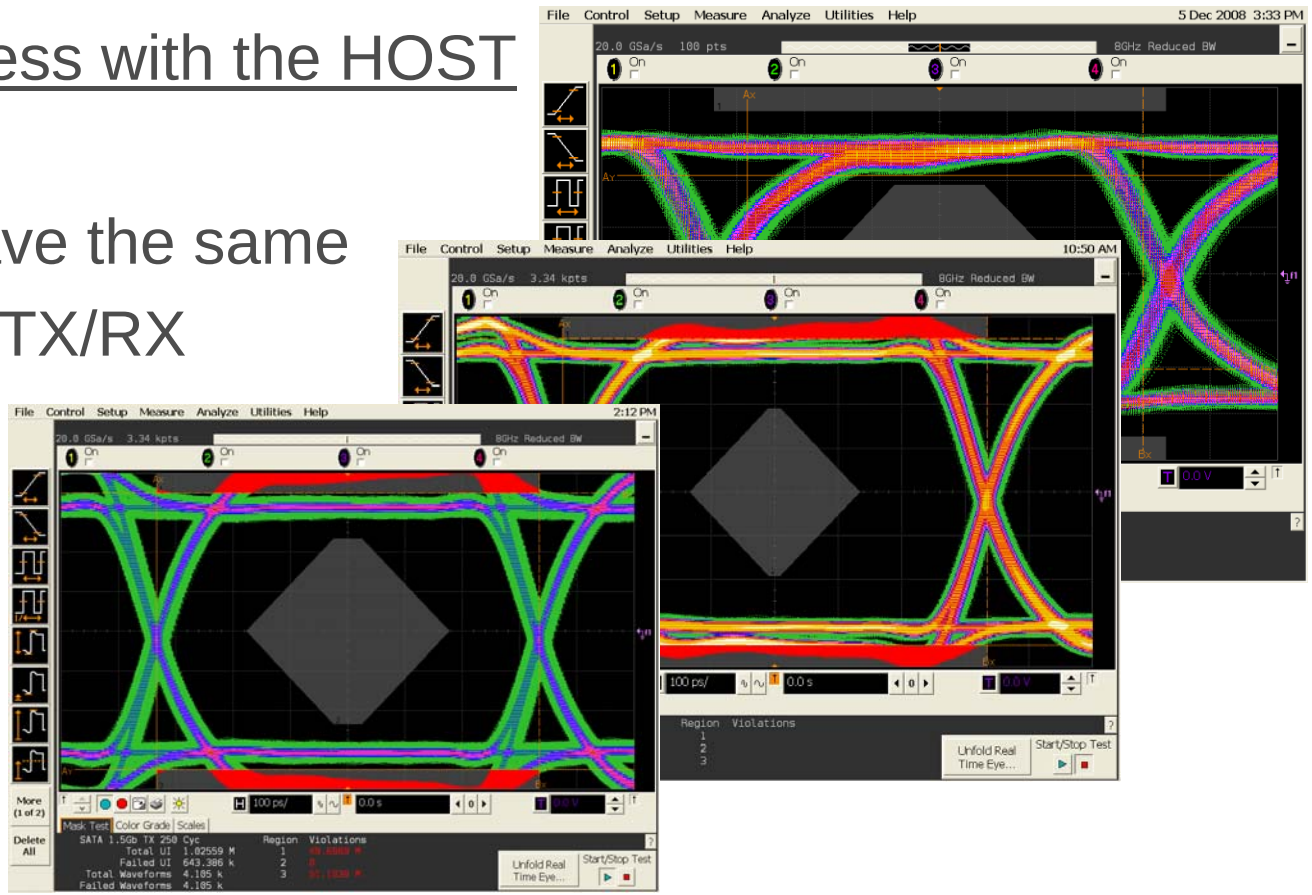
Why does it matter?

Now let’s go inside the controller

What are the parts that make SSD enterprise grade?

<u>Interface Robustness with the HOST</u>

Not all systems have the same level of control on TX/RX



After the interface comes the media

Why does media affect the drive performance?

SLC constraints

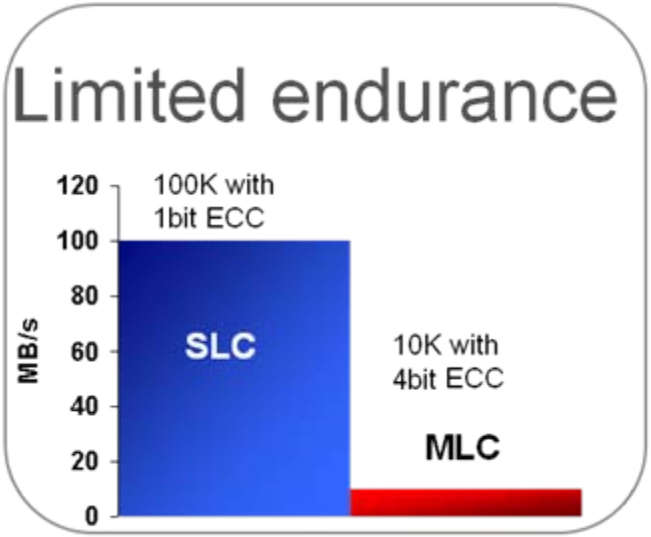
- less ECC, more NAND, more cost, more cycles available

MLC constraints

- More ECC, “less” NAND, less cost fewer cycles available

Generational constraints

- Each die shrink adds complexity



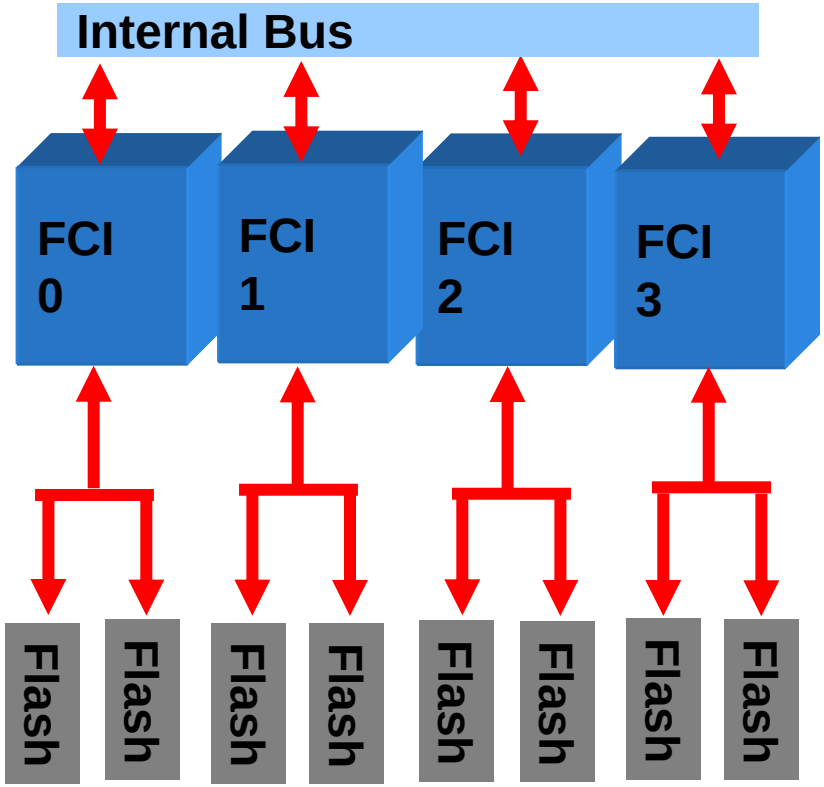
Non Enterprise SSDs Channel Management

Shared FCI (Flash Control Interface)

Each FCI must support two unique NAND elements with shared data bus

Results:

- Performance degradation over life
- Faster elimination of spare blocks up to 4% instead of 2%
- Overall lifetime is reduced



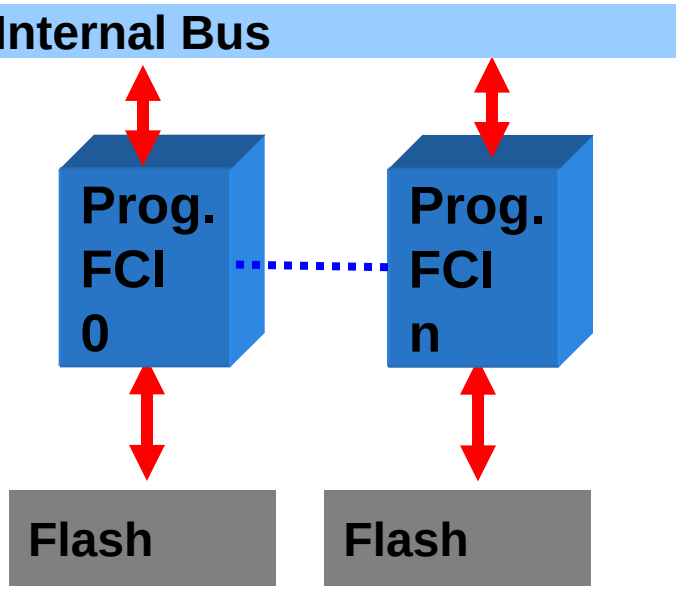
STEC Enterprise Full Channel Independence

Independent FCI (Flash Control Interface)

Each FCI supports its own Core and ECC

Each FCI runs independent of any other Flash
Data Wear is mitigated one block at a time

Results:
Performance sustained over life
Maximized use of Spare Blocks
Overall lifetime is enhanced
Platform independent



Now let's go inside the controller

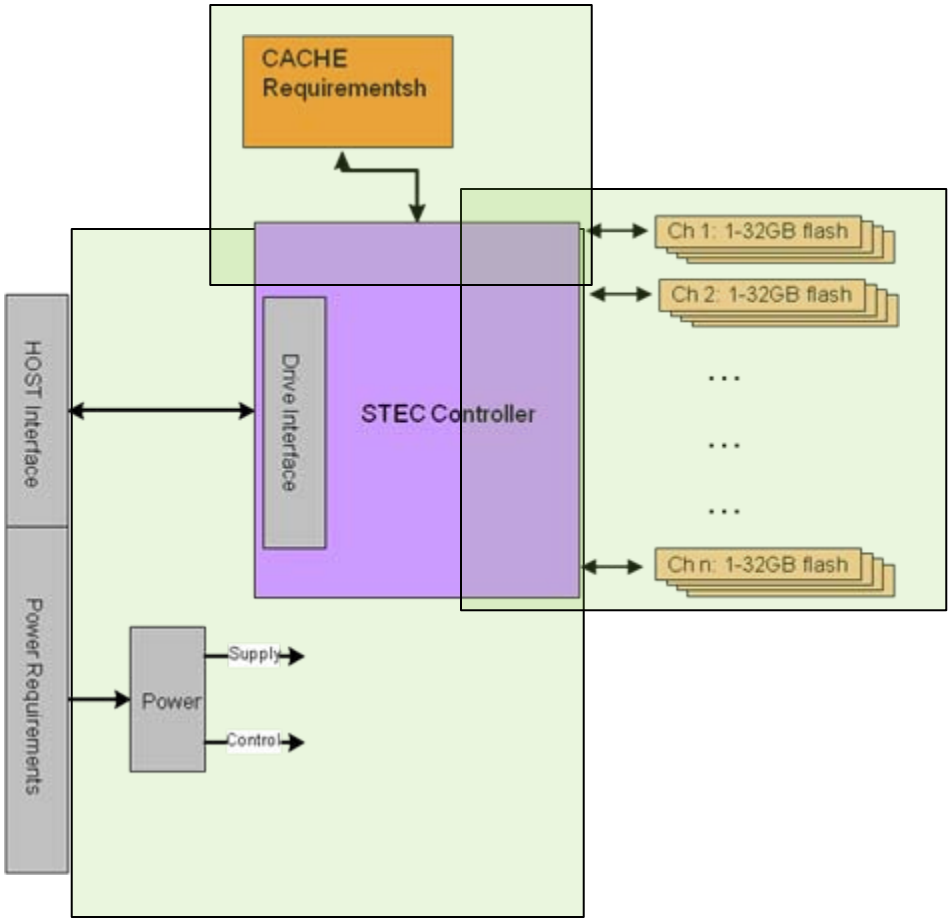
What are the parts that make SSD enterprise grade?

We have the Interface

We have the media control

What about the Data Paths?

Protecting the data in NAND is not enough, it must be protected everywhere in the drive



What about Statistics – Predicting Life

Just a “snapshot” of the available S.M.A.R.T attributes for use

None are as complete as the STEC attribute set to date.

Attribute	Flags	Type
Program Fail Count	0x32	Advisory
Reallocated Sector Count	0x32	Advisory
Erase/Program Cycle Count	0x32	Advisory
Wear Leveling Count	0x32	Advisory

Comprehensive Endurance monitoring

STEC monitors and tracks in all drives

- Erase activity counter
- Error conditions and events

SSD Classification

Building the SSD requires more than just the blocks

There are many pieces that must be assembled correctly to create the right SSD for the right application

Consumer-based SSDs will not successfully maintain an Enterprise workload environment

Wrap up Agenda Review

Content focus was:

1. Understanding NAND behaviors and limitations

2. Managing the SSD to ensure consistent performance and reliability

3. Understanding the notable difference between SSD grades and application usage.