



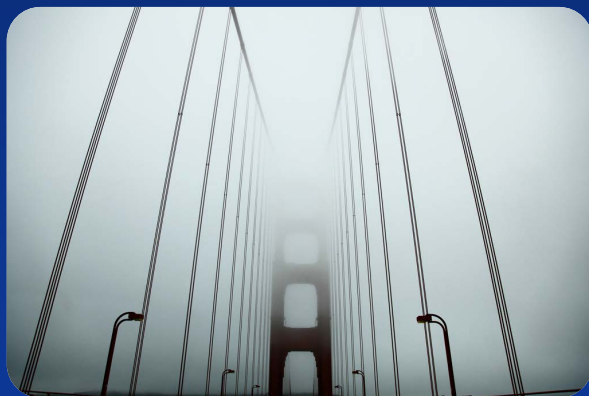
Industrial Embedded Technology for an Interconnected World

Can You See Inside Your SSD?


Scott Phillips

Director, Product Marketing at Virtium

Limited visibility into how SSD is functioning
Standard SMART data is limited and often cryptic



SMART Details

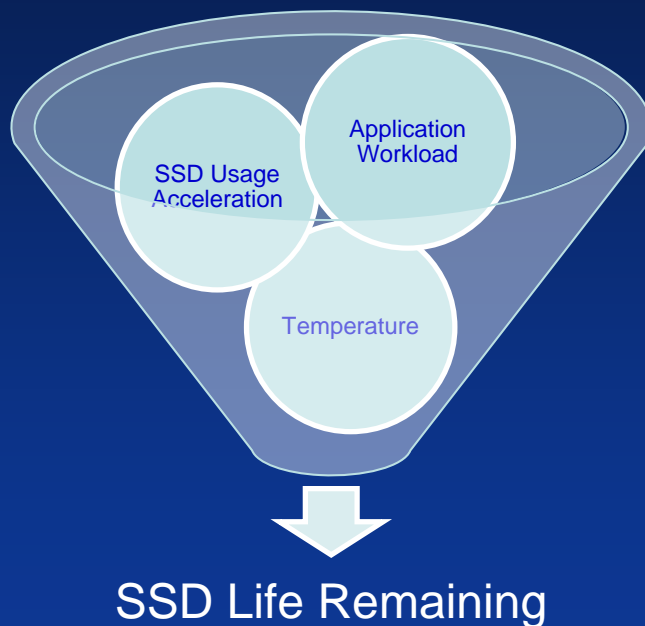


C:\ 111.64 GB

ID	Description	Raw	Normalized	Threshold	Action
05	Re-allocated Sector Count	0	100	0	Ready for use.
09	Power-On Hours Count	894834	0	0	Ready for use.
0C	Power Cycle Count	12	100	0	Ready for use.
B5	Program Fail Count	0	0	0	Ready for use.
B6	Erase Fail Count	0	0	0	Ready for use.
C0	Unsafe Shutdown Count	6	100	0	Ready for use.
E1	Host Writes	129.31 GB	100	0	Ready for use.
E8	Available Reserved Space	0	100	10	Ready for use.
E9	Media Wearout Indicator	0	100	0	Ready for use.
F1	Total LBAs Written	129.31 GB	100	0	Ready for use.
F2	Total LBAs Read	70.72 GB	100	0	Ready for use.
F9	Total NAND Writes	102	100	0	Ready for use.

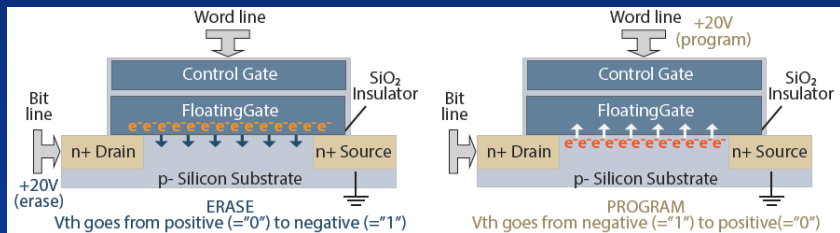
Managing NAND reliability and endurance is tricky

- Application workload affects SSD usage acceleration (+/-), which impacts SSD available life
- Temperature affects reliability, endurance, and data retention

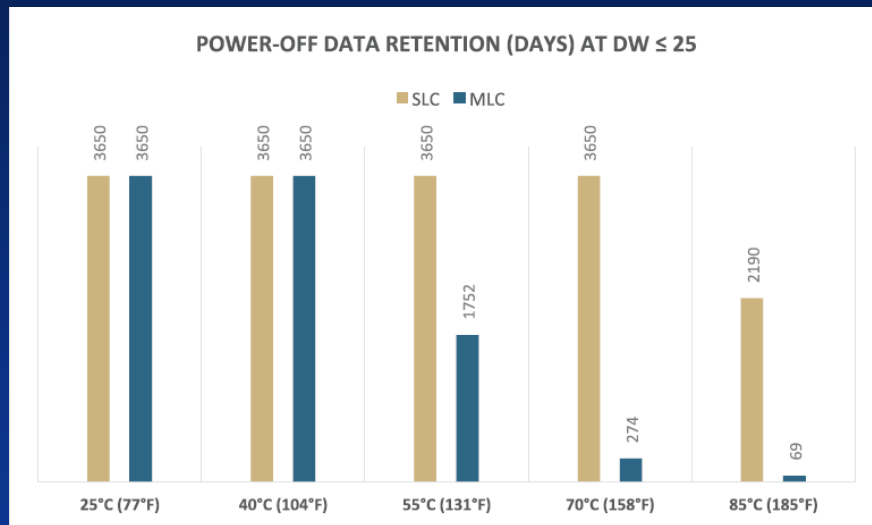


Why is it Happening?

Fundamental NAND architecture requires careful management



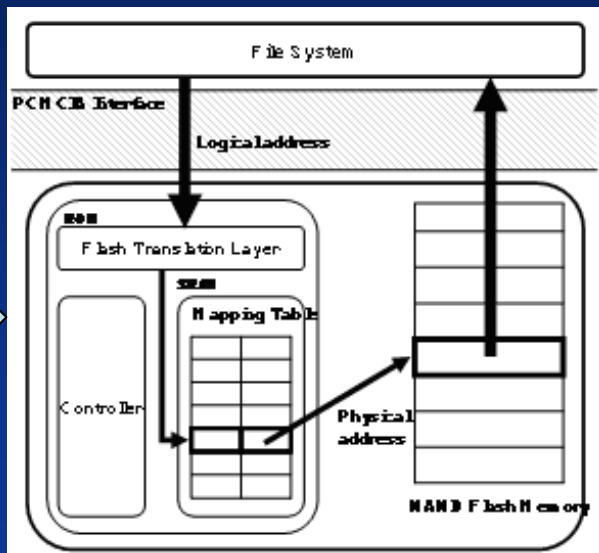
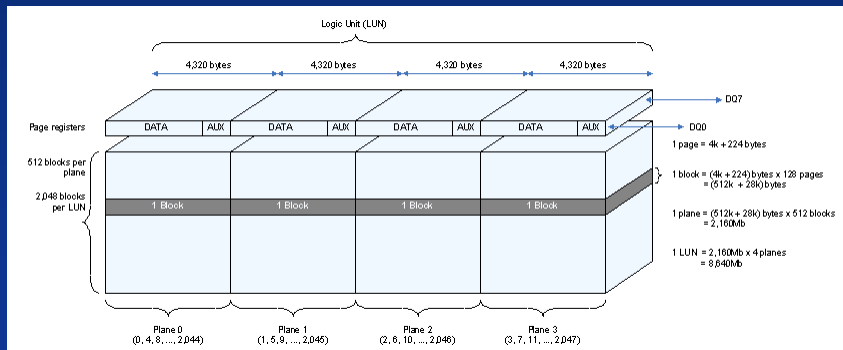
Planar NAND Flash Cell Structure



Temperature Effects on NAND Flash Data Retention

Why is it Happening?

NAND data structure requires even more management



Provide ability to see what is happening inside the SSD



- ✓ Monitor workload and effects on SSD (e.g. life remaining)
- ✓ View usage acceleration impacts
- ✓ Develop “Smarter” SMART attributes
- ✓ View temperature vs. errors



Virtium's vtView SSD Dashboard

Provide tools to
optimize and control
SSD operation



Optimize for Performance

- Firmware for specific application workloads

Optimize for Endurance

- Overprovisioning to lengthen available SSD life

Optimize for Reliability

- Built-in protection for high temp and power-loss

Optimize for Power

- Balance power and performance requirements

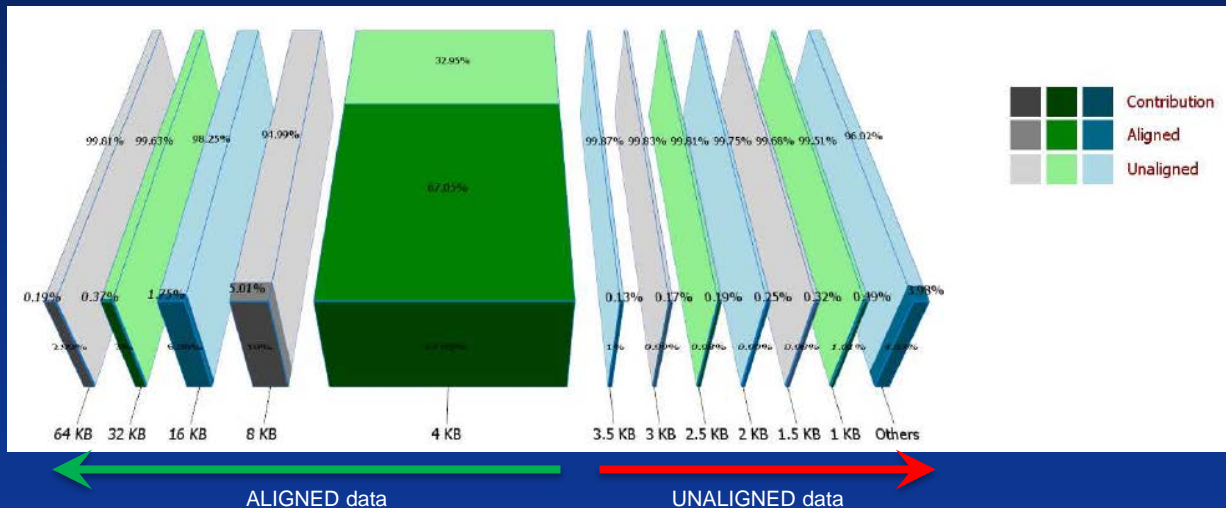
Proactive Notification

- Suggested maintenance, workload optimization

Data Alignment

- Understand the application's data transfer patterns
- Try to coalesce host data into larger files that are aligned on 4/8/16K page boundaries

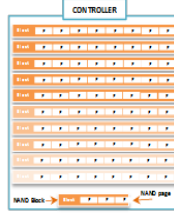

Smaller, unaligned data causes increase in SSD usage acceleration, which in turn increases latency and wears the SSD more rapidly



Overprovisioning

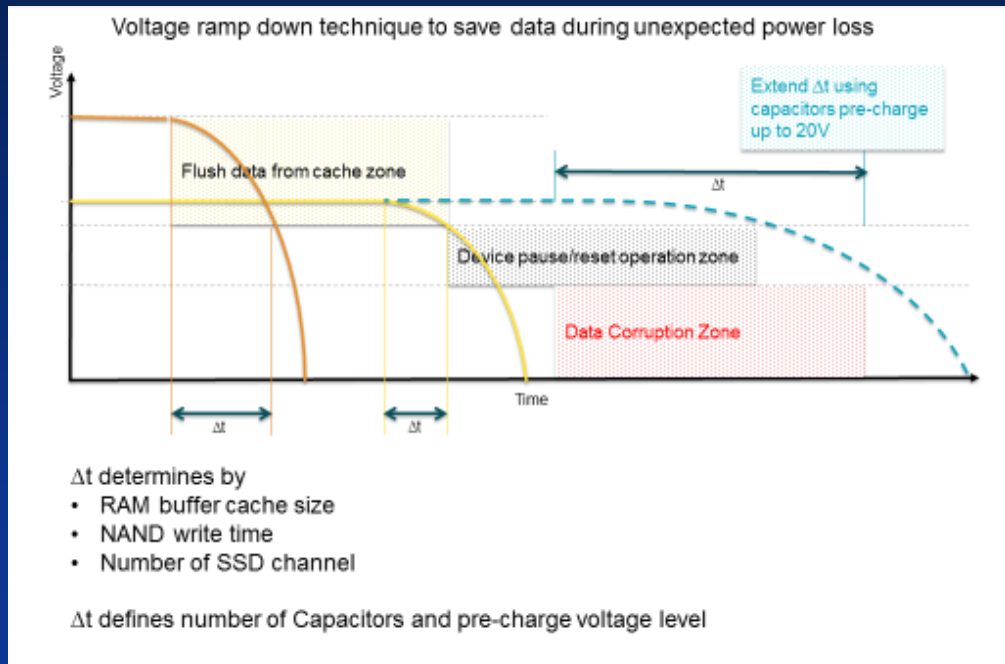
- Selecting the right SSD configuration requires knowledge of application workload
 - Smaller, random data benefits greatly
 - Larger, sequential data benefits marginally

Effects of over-provisioning on SSD endurance and performance

SSD NAND SPACE	CONTROLLER OPERATION	PERFORMANCE	ENDURANCE
Standard Operation			
	<p>Normal WRITE operation at steady state SSD</p> <pre> graph LR A[Host data] -- 1 --> B[SSD data] B -- 2 --> C{Block avail?} C -- 3 --> D[Locate block] D -- 4 --> E[Erase block] E -- 5 --> F[Write data to NAND] </pre> <ul style="list-style-type: none"> • Data in NAND flash must be erased before writing new data • Must erase entire block • 1 SLC block = 128 pages x 8KB/page = 1 MB (+ 128K ECC) • 1 MLC block = 256 pages x 16KB/page = 4 MB (+ 320K ECC) 	<p>Typical SLC NAND:</p> <p>Write 400µsec Erase 5000µsec TOTAL 5400µsec</p> <p>Typical MLC NAND:</p> <p>Write 1400µsec Erase 5000µsec TOTAL 6400µsec</p>	<p>TBW: Lifetime Terabytes Written</p> $TBW = \frac{Capacity \times P/E \text{ cycles}}{WAF \times WLF}$ <p>Example</p> <p>Capacity (Gigabytes) 128 GB P/E Program/Erase cycles (K) 100K WAF: Write Application Factor = Flash writes / Host writes 5 WLF: Wear Leveling Factor = Maximum P/E cycle / Average 2</p> <p>TBW = (16 GB x 100K P/E) / (5 x 2) = 1.28 TB</p>
Operating with Virtium Over Provisioning (OP)			
	<p>Re-WRITE operation updating 2 pages data on previous valid data block D</p> <pre> graph LR A[Host data] -- 1 --> B[SSD data] B -- 2 --> C{Pre-erased block avail?} C -- 3 --> D[Write data to NAND] </pre> <ul style="list-style-type: none"> • Avoid whole block erase cycle to update only two pages in the block • Map new data to pre-erased block in over provisioning (OP) area • Eliminates two operation cycles • During idle time (no host operation), perform garbage collection to manage available space and pre-erase OP blocks 	<p>Typical SLC NAND:</p> <p>Write 400µsec TOTAL 400µsec</p> <p>Typical MLC NAND:</p> <p>Write 1400µsec TOTAL 1400µsec</p> <p>Avoiding slow erase operation increases performance</p>	<p>Example</p> <p>WAF: Write Application Factor = Flash writes / Host writes 3</p> <p>TBW = (128 GB x 100K P/E) / (3 x 2) = 2.13 TB</p> <p>WAF drops from 5 to 3 – increases endurance from 1.28TBW to 2.13TBW</p>

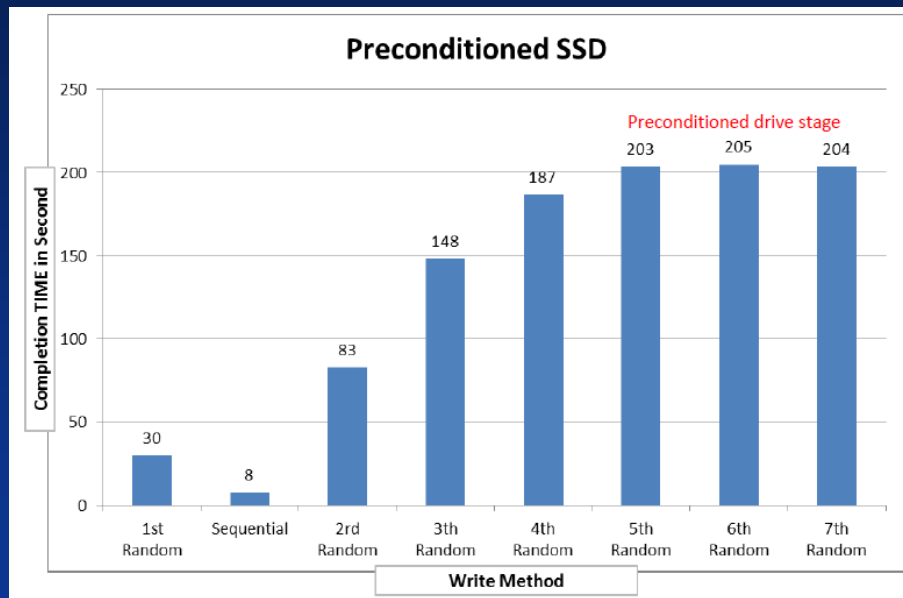
Data protection during power loss

- In-flight data as well as firmware metadata at risk during power loss



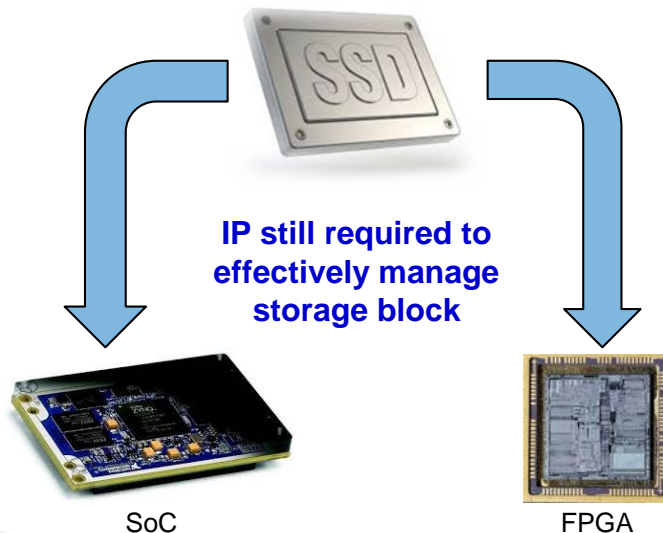
Preconditioning

- Shows SSD's true “steady-state” performance
- **Goal:** Write to drives until a steady-state write pattern is achieved
- **Methodology:** Multiple full drive capacity writes with random, 4K, unaligned data



* Refer to JEDEC for published SSD benchmarking guidelines

“Software-defined storage” is coming



Will You be Ready?

Align storage needs
with value-added
capabilities



- Not enough to rely on “integrators” providing lowest cost
- Align with suppliers providing expertise and value-add technology:
 - Firmware development and flash management expertise
 - Software tools to monitor/manage SSD
 - Flexibility to adapt SSD to application workloads and environmental conditions

It's a Wrap!

For more information, go to:

<http://www.virtium.com/resources/product-collateral/>



Industrial Embedded Technology for an Interconnected World