DRIVES WITH INTEGRITY
Designing for Zero Silent Data Corruption Tolerance

Vinod Ambrose, Roger Corell
Q2 2016

Legal Disclaimer

Intel may make changes to specifications and product descriptions at any time, without notice. Designers must not rely on the absence or characteristics of any features or instructions marked "reserved" or "undefined". Intel reserves these for future definition and shall have no responsibility whatsoever for conflicts or incompatibilities arising from future changes to them. The information here is subject to change without notice. Do not finalize a design with this information.

The products described in this document may contain design defects or errors known as errata which may cause the product to deviate from published specifications. Current characterized errata are available on request.

Intel technologies' features and benefits depend on system configuration and may require enabled hardware, software or service activation. Performance varies depending on system configuration. No computer system can be absolutely secure. Check with your system manufacturer or retailer or learn more at intel.com.

Intel disclaims all express and implied warranties, including without limitation, the implied warranties of merchantability, fitness for a particular purpose, and non-infringement, as well as any warranty arising from course of performance, course of dealing, or usage in trade.

Tests document performance of components on a particular test, in specific systems. Differences in hardware, software, or configuration will affect actual performance. Consult other sources of information to evaluate performance as you consider your purchase.

For more complete information about performance and benchmark results, visit http://www.intel.com/performance.

Intel does not control or audit the design or implementation of third party benchmark data or Web sites referenced in this document. Intel encourages all of its customers to visit the referenced Web sites or others where similar performance benchmark data are reported and confirm whether the referenced benchmark data are accurate and reflect performance of systems available for purchase.

Intel and the Intel logo are trademarks of Intel Corporation in the U.S. and other countries.

*Other names and brands may be claimed as the property of others.

Silent Data Corruption

"When data corruption goes undetected, it becomes silent and is a high risk for applications" *EnterpriseStorageForum.com*

Data corruption

Detected   **Undetected**

Uncorrected

Corrected

Silent Data Corruption

Block sizes for illustrative purposes only. Does not correlate to frequency of occurrence.

See appendix for footnote 1

SDC events may occur more frequently than perceived

11 **drives per 1000** can experience SDC in a year[1]

up to **10%** **catastrophic storage system failures** have been linked to SDC[1]

Silent Data Corruption can be costly

Businesses can be **impacted on multiple levels**

Incorrect operations
Loss of critical data
Loss of business
Significant downtime

Data center **downtime cost** on average

$7,900 per minute[2]

Actual cases demonstrate the impact

amazon

S3 cloud service **down for 36 hours**
http://cloudsecurity.org/blog/2008/06/25/a-question-of-integrity-to-md5-or-not-to-md5.html
http://www.dataspan.com/blog/post/beware-silent-data-corruption

f

10-15% photos **not available for 48 hours**
http://www.computerworld.com/s/article/9129263/

NETFLIX

55 distribution centers **down for 3 days**
http://www.pcmag.com/article2/0,2817,2328778,00.asp

* Names and brands of others may be claimed as property of others. See appendix for footnote 2.
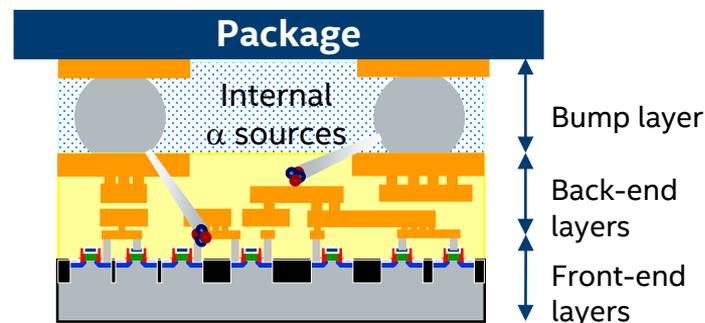
Silent Data Corruption has two main external causes

Cosmic Rays

1. Protons and heavy ions originate from the sun and stars

2. They interact with atmosphere creating neutrons

3. Neutrons multiply quickly in a cascading reaction

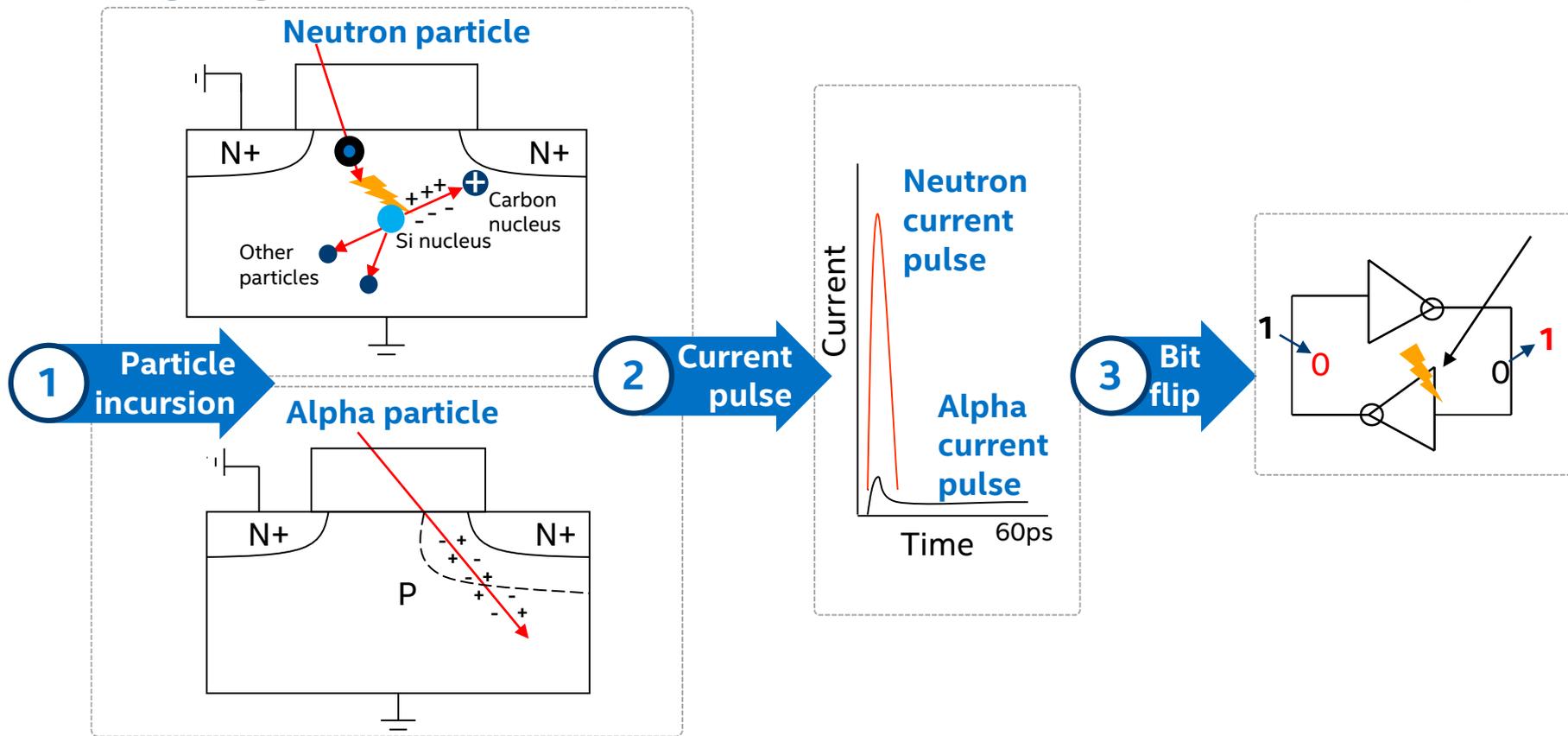4. At earths surface, ~10 neutrons/sec. pass through a person
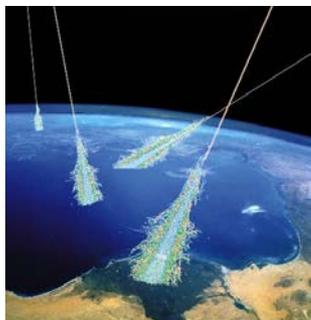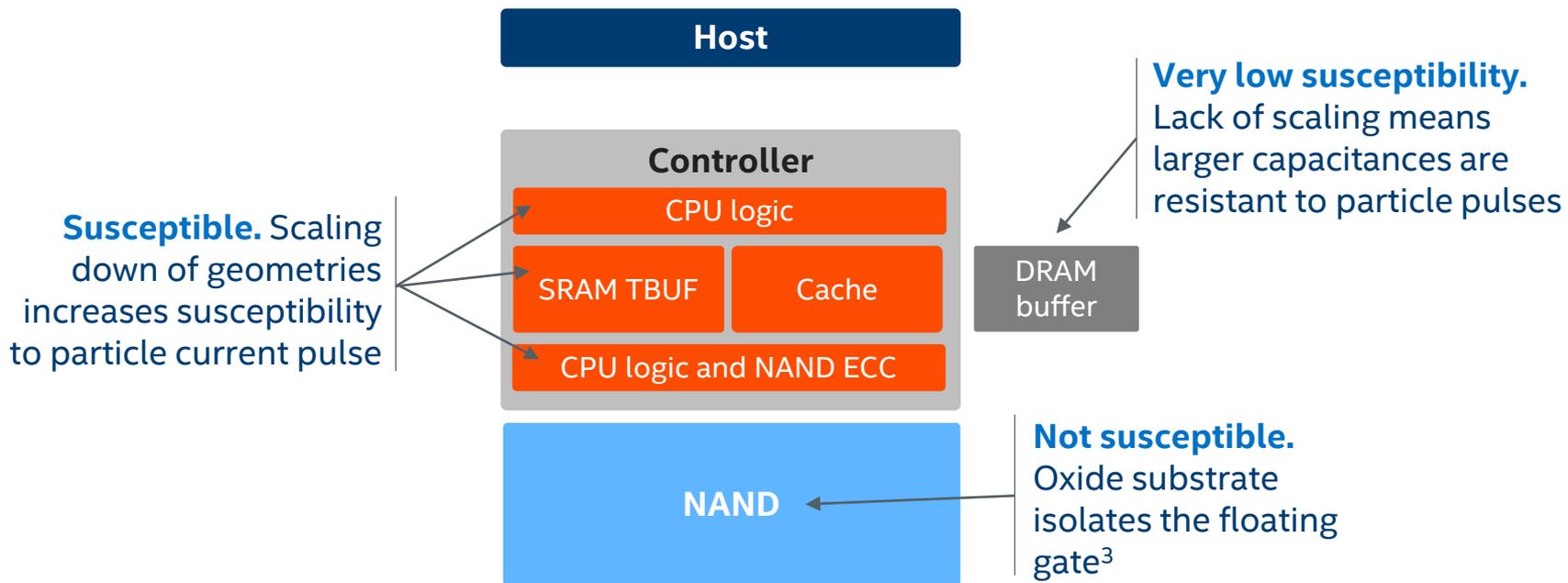

Alpha particles

1. Trace radioactive elements exist in some materials

2. Even pure leads can generate 1 particle/cm^2/khr.

Charge generated by particles can cause a bit flip

SSD susceptibilities to particle incursions

Host

Controller

CPU logic

SRAM TBUF | Cache

CPU logic and NAND ECC

DRAM buffer

NAND

Susceptible. Scaling down of geometries increases susceptibility to particle current pulse

Very low susceptibility. Lack of scaling means larger capacitances are resistant to particle pulses

Not susceptible. Oxide substrate isolates the floating gate[3]

*See appendix for footnote 3.

Particle incursion behaviors

Correct behavior

Host

X

Controller

CPU logic

X SRAM

Cache

CPU logic and NAND ECC

DRAM

X **NAND**

- Written data 'X' goes through logic and SRAM
- Controller picks NAND location and writes
- Read is simply reverse order

Incorrect behaviors

Host

X   +

Controller

CPU logic

① SRAM
X→Y

Cache
+ → −

②

CPU logic and NAND ECC

DRAM

③

Y   −   ?

NAND

Multiple modes

① Bit flip in transit if upstream of NAND ECC

② Bit flip in instruction cache. Controller executes wrong instructions

③ Flip in CPU logic. Can cause
- Firmware miss-execution
- Controller hang

Designing for data integrity. Beyond the basics

Basics

Host

③

Controller

② CPU logic

SRAM   Cache   ① DRAM

CPU logic and NAND ECC

NAND

① ECC or parity on RAMs

② End-to-End data protection

③ Host-applied CRC check
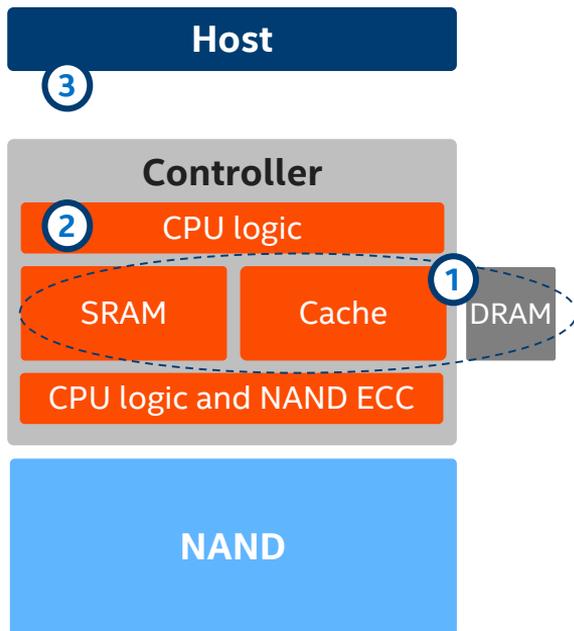
Beyond

① **Interleave** to reduce vulnerability to multi-bit errors

Interleaved W1 + W2								Parity Bit	
W1B1	W2B1	W1B2	W2B2	W1B3	W2B3	W1B4	W2B4	PB1	PB2
0	1	0	1	1	0	1	0	0	0

Interleaved W1 + W2								Parity Bit	
W1B1	W2B1	W1B2	W2B2	W1B3	W2B3	W1B4	W2B4	PB1	PB2
1	0	0	1	1	0	1	0	1	1

Particle caused multi-bit error

With interleaving, error detected

② **Protect all critical storage arrays** within controller

③ If unsure, **brick drive**   **~0.002%** SDC/year **prevented**[3]
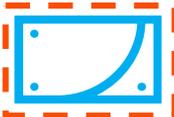
Validating SDC requirements

Traditional test methods do not go far enough

≤1e-22
Requirements

✓

1e-18
traditional test (RDT) limit[6]

Intel exposes SSDs to **high intensity particle beams** detecting to **requirements**

Los Alamos Neutron Science Center

SSDs

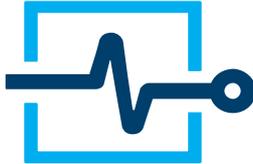Accelerator beam

Neutron particle testing

Intel

Alpha particle testing

Test conditions are up to **10,000,000 worse than real world**[5]

The Intel advantage

Designing for zero tolerance

Testing beyond traditional methods

Delivering unparalleled results

≥100X

more reliable preventing Silent Data Corruption[4]

Learn more here

* See appendix for footnote 4

Drives with Integrity.

FOOTNOTES AND DISCLAIMERS

1. **SDC frequency** - Source - NEC, "Silent Data Corruption in Disk Arrays". https://www.necam.com/docs/?id=54157ff5-5de8-4966-a99d-341cf2cb27d

2. **Cost of downtime** - Source – Emerson Network Power. http://www.emersonnetworkpower.com/en-US/About/NewsRoom/NewsReleases/Pages/Emerson-Network-Power-Study-Says-Unplanned-Data-Center-Outages-Cost-Companies-Nearly-9000-Per-Minute-.aspx

3. **No susceptibility in NAND and, firmware bricking of drive** – Source – Intel. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7289417 (recommend opening in Chrome browser)

4. **Silent Data Corruption.** Source - Intel. Test performed on Intel® SSD S3x00 drives, Samsung PM853T and SM843T, Micron P400e, Seagate 600 Pro and SanDisk Lightening drives. Drives were exposed to increasing amounts of radiation. After a drive "hang", a power cycle was performed to determine whether the drive would re-boot. If a drive re-booted it was read, and data was compared to the tester's master copy of the up-to-date data that the drive was expected to contain based on writes the drive had acknowledged as completed prior to the "hang" event. If the drive returned data that differed from the expected data, it was recorded as failing for silent errors. The annual rate of silent errors was projected from the rate during accelerated testing divided by the acceleration of the beam (see JEDEC standard JESD89A).

5. **Test conditions harsher than real world.** Source – Los Alamos Neutron Science Center http://wnr.lanl.gov/_assets/flight_paths/4FP30L_about.php

6. **RDT test limit.** Source – Intel. Typical Reliability Demonstration Test involve testing 1000 drives for 1000 hours to levels down to 10^{-18}.