EDITOR TOM COUGHLIN
Coughlin Associates; tom@tomcoughlin.com

New IEEE Media Sanitization Specification Enables Circular Economy for Storage

Jonmichael Hands, Chia Network
Tom Coughlin, Coughlin Associates

Modern media sanitization techniques can securely eliminate data on digital storage devices. This enables more effective efforts to reuse and recycle these devices, enabling a circular economy for data storage.

Digital Object Identifier 10.1109/MC.2022.3218364
Date of current version: 9 January 2023

Data growth has exploded, creating amazing opportunities and enabling quality of life improvements. The amount of data being created has far outpaced the amount of data being stored, with the International Data Corporation (IDC) forecasting that, in 2026, the massive 20.5 ZB of data being stored in the world will make up only about 10% of the total data generated that year (see Figure 1). This growth of stored data needs to be sustainable, with more companies than ever involved in the storage of digital data setting net-zero emission goals by 2030.

RAPID DATA GROWTH DEMANDS SUSTAINABLE PRACTICES

A modern high-capacity 3.5-in hard drive has an environmental footprint of 2.55 kg CO_2 emitted per terabyte per year.[2] One study estimated the embedded carbon from manufacturing solid-state drives (SSDs) to be as high as 0.16 kg CO_2 emitted

per gigabyte (or 160 kg CO_2 emitted per terabyte).[3] While more new semiconductor manufacturing is being moved to using primarily renewable energy, the overwhelming majority today is not. The combination of rapid data growth and the tremendous amounts of energy and water required to manufacture these storage devices make them a great target for a circular economy to extract the maximum amount of value.

> This growth of stored data needs to be sustainable, with more companies than ever involved in the storage of digital data setting net-zero emission goals by 2030.

OPPORTUNITY FOR A CIRCULAR ECONOMY FOR DIGITAL STORAGE

Circular economy principles keep material at the highest value state possible for as long as possible—employing reuse, sharing, repair, refurbishment, remanufacturing, and recycling to create a closed-loop system. This can minimize the use of physical resources as well as the creation of waste, pollution, and carbon emissions. This "circular economy" is compared to a "linear economy" in the following sections.

Linear economy for digital storage

In a linear economy for storage, the purchaser of the storage deploys it for three to five years and then moves it to the "end of life" by putting it through a shredder or other physical destruction device. This leaves a pile of raw materials, including rare earth minerals, that are valuable but incredibly inefficient to sort out and recycle.

Circular economy for digital storage

A circular economy for storage prioritizes the longest first use, extending life through repair and health monitoring. For SSDs and hard disk drives (HDDs), there are many built-in modern features to detect the health of the drive through self-monitoring, analysis, and reporting technology and other log pages that can identify issues early, and vendors often release firmware updates to improve device reliability across the lifetime of the drive. It is well known that storage devices can last much longer than a five-year warranty, with some old HDDs and some of the very first SSDs deployed still running today.

A second use of the device requires the removal of all prior customer and user data through a process called *data* or *storage sanitization*. The first user uses "purge media sanitization" techniques to securely prevent any unauthorized access to the data and then verifies by attempting to read data back and verifying that no user data are recoverable before the transfer of ownership. The second device user may be able to run it for another three to five years (or, for less demanding applications, even longer) before total device failure or wear-out. In an ideal model, artificial intelligence/machine learning can be used to estimate further and predict device failure timing so that drives can be safely sanitized right before the end of life. The drives can then be sent back to the manufacturer or processing facility to recover any component subassemblies without shredding or melting the drive components.

In HDDs, a particularly valuable subassembly is the actuator magnet assembly. This does not change very much generation over generation. This allows the rare earth minerals and components in these magnets to be separated out for the highest value recovery. After the magnet assemblies are removed, the other HDD raw materials can be broken down into resources to be recycled and put back into the production of new devices that contain recycled materials. Figure 2 illustrates the circular economy for storage devices, such as SSDs and HDDs.

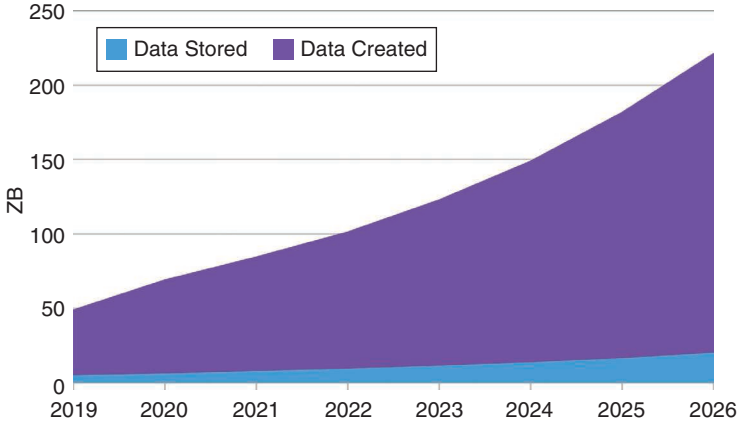

FIGURE 1. The IDC projection of data generated versus stored out to 2026.[1]

THE HISTORY

The need to remove old data from storage devices is not a new idea. The industry has come up with many terms to describe this: *data deletion*, *secure data removal*, *data shredding*, *data wiping*, *data overwriting*, *data erasure*, *data clearing*, and *data destruction*. Only one of these terms has a rigorous definition, though: *data sanitization*.

Storing, securing, and processing data in the information and communications technology (ICT) industry is fundamental to operating global businesses. Companies go to great lengths to secure their data and prevent confidential information from being made available to others. When a company is done using its ICT equipment, including the storage devices, it is essential to render the data inaccessible.

Ancient media sanitization specifications like U.S. Department of 5220.22-M date back to 1995[4] and were meant for old HDD technology, where the head positioning was not anywhere near as accurate as it is today. The 5220.22-M data sanitization process involved multiple-pass overwrites, with three passes being standard and seven passes used for an extended erase, consuming tremendous amounts of time and energy as well as accumulating media wear.

Even early SSD secure erase implementations had multiple-overwrite modes, which is catastrophic for SSD NAND endurance. Unfortunately, many data-erasing software companies still support the 5220.22-M specification, and customers unknowingly ask for them even after these sanitization methods have been deprecated.[5] Modern media sanitization specifications have made it clear that multiple erase passes are not required to securely prevent access to storage media data.

National Institute of Standards and Technology (NIST) SP 800-88 r1 is one of the most widely used media sanitization specifications because it was used and adopted by the U.S. federal government, with the last version published in December 2014.[6] Unfortunately, NIST SP 800-88 r1 is only a guideline—there are no compliance and conformance requirements, despite many companies claiming to have requirements based upon this document. It contains much text in common with ISO/International Electrotechnical Commission (IEC) 27040-2015[7] because they were written at the same time and by many of the same authors.

MODERN STANDARDS FOR MEDIA SANITIZATION

IEEE 2883-2022 is the latest standard for media sanitization.[8] It defines sanitization methods and techniques for the specific storage media type (HDD, SSD, optical, removable, etc.) and specifies interface-specific techniques (SATA, SAS, and NVMe). This specification can align the industry on terminology and modern techniques for media sanitization, targeting all logical and physical locations for data—including user data, old data, metadata, and overprovisioning. The three sanitization methods outlined in the specification are clear, purge, and destruct.

Other specifications for the ecosystem of media sanitization reference the work done in IEEE 2883-2022. ISO/IEC 27040 adds additional requirements for sanitization compliance, including identification of the form of storage, verifying the results, and documenting and producing evidence of storage sanitization.[7] The update to ISO/IEC DIS 27040 (second edition) is expected to be published in early 2023.

MEDIA SANITIZATION

Removing data is not trivial. All copies of the data must be located, all of the data must be classified for sensitivity and risk of the data being accessed by an unauthorized party, the data storage technologies that the data are located on (which are designed to guard against data loss) must be identified, and this data removal must be compliant with company or government policies.

> A circular economy for storage prioritizes the longest first use, extending life through repair and health monitoring.

The following is the definition of *sanitization*: a process or method to render access to target data on storage media infeasible for a given level of effort. Data sanitization targets all instances of stored data, across all copies, wherever the data reside. Storage

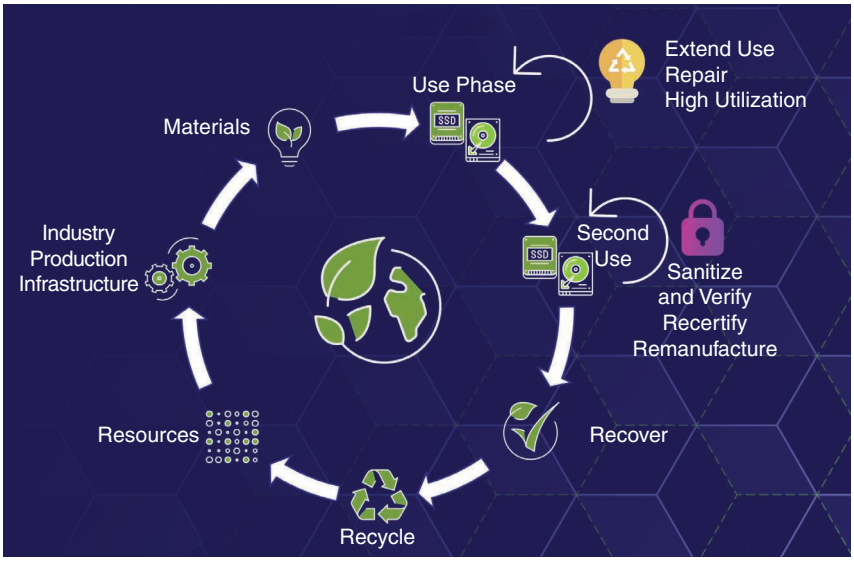

FIGURE 2. A circular economy for SSDs and HDDs.

sanitization focuses on ICT infrastructure that uses nonvolatile storage. Often, these data are abstracted away in a complex storage system that uses erasure coding, redundant array of independent disks, or other parity for data protection as well as other modern storage array features like compression, deduplication, and encryption. Logical sanitization targets these data on logical or virtual storage. Media sanitization targets the data on a specific storage device.

Figure 3 shows the three data sanitization methods discussed in the IEEE 2883-2022 specification. The following are more details about these three data sanitization methods.

- *Clear*: This is data sanitization using logical techniques on user data on all addressable storage locations for protection against simple, noninvasive data recovery techniques using the same host interface available to the user, for example, basic file recovery utilities that scan the drive logical blocks for data.

> It is well known that storage devices can last much longer than a five-year warranty, with some old HDDs and some of the very first SSDs deployed still running today.

- *Purge*: This is the most important sanitization method for enabling circularity for storage devices! Purge sanitization uses logical techniques or physical techniques that make the recovery of target data infeasible using state-of-the-art laboratory techniques but that preserve the storage media and the storage device in a potentially reusable state. Multiple purge methods can be used together to further decrease the probability of recovering data. Still, even a single one of these verified purge techniques is sufficient against state-of-the-art laboratory techniques (disassembly, electron microscopy, magnetic force microscopy, X-ray probing, etc.):
  - Sanitize purge cryptographic erase (CE) will change the media encryption key on a device, typically today encoded using Advanced Encryption Standard 256 (AES256), which is not only a secure way to sanitize a device but also happens in seconds.
  - Sanitize purge overwrite securely overwrites the storage media with various patterns that can be verified later. Overwrite can be used with hard drives that do not support CE.
  - Sanitize purge block erase can zero out the erase blocks on SSDs, the native way that NAND flash operates for erasing, and is able to complete in seconds to minutes for an entire SSD. Block erase can be used in conjunction with CE.
- *Destruct*: This is a method of sanitizing using physical techniques that make the recovery of the target data infeasible using state-of-the-art laboratory techniques and results in the subsequent inability to use the storage media for storage. Shredding and pulverizing, which were once approved methods of destruction, are not approved sanitization techniques today. This is because a small 2-mm-square shred of a disk platter can contain large amounts of recoverable data. In general, destruct should only be used if other sanitize methods fail or verification of sanitization fails. For example, if a drive is in a failed state and not responding to host commands, physical destruction may be the only way to make sure that

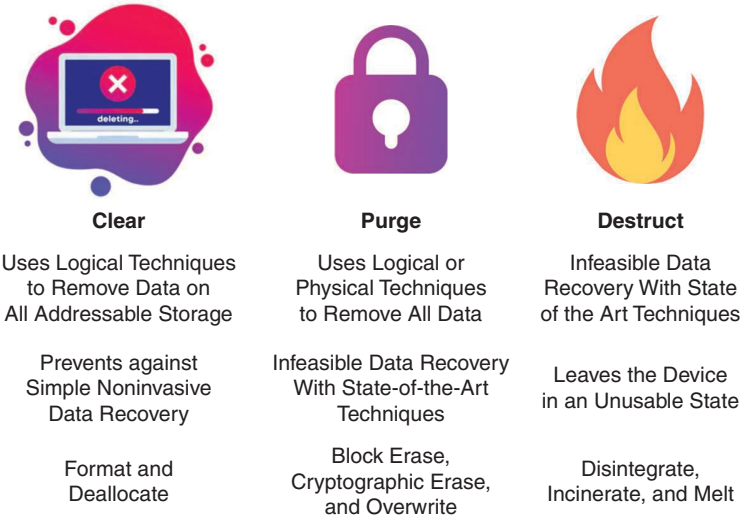

FIGURE 3. Three basic data sanitization methods. (Source: Open Compute Project[9]; used with permission.)

data cannot be recovered. The methods approved today for destruct sanitization are melting and incineration, which can take more energy and resources than a shredding machine. Destroying a storage device with these methods has the obvious security benefit of destroying data beyond any type of forensic recovery capability:

- *Melting*: destruct by changing the storage media from a solid to a liquid state, generally by the application of heat
- *Incineration*: destruct by burning a storage device completely.

The important element of purge sanitization is that it makes data recovery infeasible with state-of-the-art equipment. Note that data recovery companies often have access to special tools and methods that can be used to help customers recover data from their storage devices if they are not purged. These data recovery methods include the disassembly of the storage device and connecting various components, like the head disk assembly (containing the HDD heads on the voice coil actuator and the disk stack), to a different and operational circuit board. Even more extreme (and expensive) data recovery techniques include the use of electron microscopy; magnetic force microscopy with HDDs; and putting NAND die, extracted from an SSD, directly into a tester to read the raw pages.

Purge techniques prevent even the most sophisticated data recovery experts from obtaining any user data. NAND flash is known to provide high performance, low latency, and low power access to data. One of the drawbacks of NAND from an SSD perspective is that the cell has to be erased before it can be programmed. NAND writes in pages, often today around 16 kB, but NAND is erased in blocks—which can be hundreds of pages.

Modern SSDs have to deal with this by using spare memory areas or overprovisioning to do garbage collection of data during storage operations. For media sanitization, this is a nice feature of NAND flash. This is because multiple block erases can be executed in parallel quickly, leaving no recoverable data. This can be performed in seconds to a few minutes, much faster than overwriting all of the blocks with a random data pattern. During this sanitization, it is common to deallocate all of the logical blocks, commonly referred to as *TRIM*, to improve SSD endurance for the next use. SSDs do not need to be overwritten with data to perform a data purge because of this useful function of all NAND flash technology.

Figure 4 shows how NAND is addressed by die, planes, erase blocks, and pages. SSDs have controllers that can perform operations simultaneously across many different channels, making the purge block erase happen quickly in parallel across many blocks and across many dice.

Self-encrypting drives (SEDs) can encrypt data at rest. Most modern SSDs use an AES256 engine to encrypt all data going to the NAND, regardless of a user or host operating the system setting a user password. Most vendors will classify a drive as an SED if it supports the encryption of data at rest utilizing the Trusted Computing Group Opal specification to set user passwords and locking ranges for the encrypted data.[11] The fact that most drives use the AES256 engine at all times makes purge sanitization using CE practical by only requiring the drive to sanitize the media encryption key. CE can be performed in seconds and can also be easily verified by reading back data and ensuring that they are random and not the data that were present before the sanitize command. There are considerations for CE being a

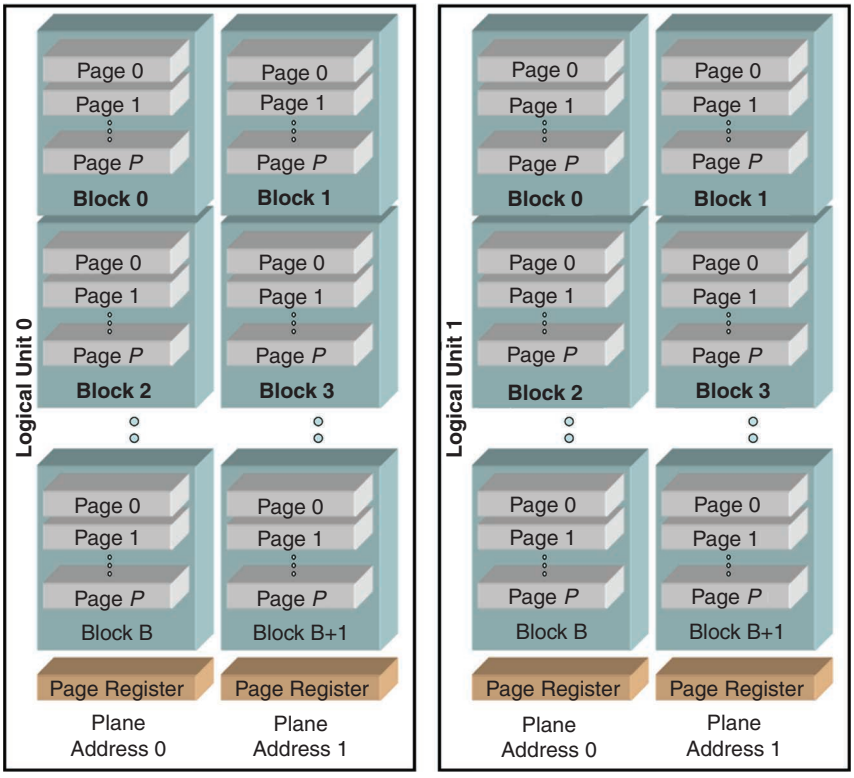

FIGURE 4. The nand flash organization: JEDEC ONFI 5.1 specification. JEDEC: Joint Electron Device Engineering Council. (Source: ONFI[10]; used with permission.)

supported purge sanitize method that are outlined in IEEE 2883-2022 Table B.1,[8] including key generation, media encryption, and key wrapping.

Modern cryptography using AES256 is trusted in the cloud and enterprise systems for data encryption. In a CE, sanitization can be performed in seconds by destroying a media encryption key, leaving all of the data on the drive encrypted. The effort needed to decrypt the data is dependent on the entropy and strength of the encryption algorithm. The NIST considers AES256[12] quantum safe today and in decades to come,[13] as it is likely that Grover's algorithm may provide little advantage to brute-force attacks. "Steal now and decrypt later" is an argument against CE but should not be used to dissuade the use of purge sanitization using CE based on the current consensus of AES256 and quantum computing experts.

The combination of using SEDs to ensure data are encrypted throughout the life of the drive and using multiple-purge media sanitization methods like crypto and block erase can reduce the probability of recovering any user data down to zero, with verification after the sanitization to ensure that no old user data can be read back. The industry is working to create widespread use of these methods so that all storage devices can be safely reused to enable a circular economy and keep healthy devices running in the field longer.

Companies have zero tolerance for risk in handling customer data (and often face severe financial consequences if these data fall into the wrong hands). The media sanitization methods discussed also require additional verification, such as reading the user data back again. There is ongoing development to create standards on the conformance, compliance, verification, and validation of these techniques so that users can be sure there are no data on a device that can be recovered after a purge media sanitization. C

ACKNOWLEDGMENT

Tom Coughlin is the corresponding author.

REFERENCES

1. "Worldwide Global StorageSphere forecast, 2022–2026: An installed base of 7.9ZB of storage capacity in 2021 came at a cost of $370 billion — Is it enough?" IDC, Needham, MA, USA, Doc #US49051122, May 2022.
2. "EXOS X18 Sustainability Report." Seagate. Accessed: Oct. 15, 2022. [Online]. Available: https://www.seagate.com/global-citizenship/product-sustainability/exos-x18-sustainability-report/
3. S. Tanno and P. Nair, "The dirty secret of SSDs: Embedded carbon," 2022. [Online]. Available: https://arxiv.org/pdf/2207.10793.pdf
4. "DoD 5220.22-M National Industrial Security Program Operating Manual (NISPOM) January 1995," U.S. Department of Defense - Department of Energy - Nuclear Regulatory Commission - Central Intelligence Agency, Washington, DC, USA, 1997. Accessed: Jul. 1997. [Online]. Available: https://www.dami.army.pentagon.mil/site/IndustSec/docs/DoD%20522022-m.pdf
5. "Everything you need to know about the DoD 5220.22-M disk wiping standard & its applications today." Blancco. Accessed: Oct. 15, 2022. [Online]. Available: https://www.blancco.com/resources/blog-dod-5220-22-m-wiping-standard-method/
6. "Guidelines for media sanitization," National Institute Standards and Technology, Gaithersburg, MD, USA, NIST Special Publication 800-88 Revision 1, 2014. [Online]. Available: https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-88r1.pdf
7. *Information Technology — Security Techniques — Storage Security*, ISO/IEC 27040, International Organization for Standardization, Geneva, Switzerland, 2015. [Online]. Available: https://www.iso.org/standard/44404.html
8. *IEEE Standard for Sanitizing Storage*, IEEE Standard 2883-2022. [Online]. Available: https://standards.ieee.org/ieee/2883/10277/
9. "White paper: Data sanitization for the circular economy," Open Compute Project, Revision 1.0, Version 1.0, Jul. 2022. [Online]. Available: https://www.opencompute.org/documents/data-sanitization-for-the-circular-economy-1-pdf
10. "Specification," Open NAND Flash Interface, Version 5.1, May 3, 2022. [Online]. Available: https://www.onfi.org/specifications
11. "TCG storage security subsystem class: Opal specification," Trusted Computing Group, Beaverton, OR, USA, 2009. [Online]. Available: https://trustedcomputinggroup.org/resource/storage-work-group-storage-security-subsystem-class-opal/
12. "AES-256 joins the quantum resistance." Fierce Electronics. Accessed: Oct. 15, 2022. [Online]. Available: https://www.fierceelectronics.com/electronics/aes-256-joins-quantum-resistance
13. "Post-quantum cryptography PQC," National Institute Standards and Technology, Gaithersburg, MD, USA, 2022. [Online]. Available: https://csrc.nist.gov/Projects/post-quantum-cryptography/faqs

JONMICHAEL HANDS is with the Chia Network, San Francisco, CA 94104 USA. Contact him at jm@chia.net.

TOM COUGHLIN is president of Coughlin Associates, San Jose, CA 95124 USA. He is a Fellow of IEEE. Contact him at tom@tomcoughlin.com.