Stash in a Flash

Aviad Zuck
Technion–Israel Institute of Technology
aviadzuc@cs.technion.ac.il

Yue Li
California Institute of Technology
yli@caltech.edu

Jehoshua Bruck
California Institute of Technology
bruck@caltech.edu

Donald E. Porter
University of North Carolina at Chapel Hill
porter@cs.unc.edu

Dan Tsafrir
Technion–Israel Institute of Technology and VMware Research Group
dan@cs.technion.ac.il

The ability to successfully hide data is becoming increasingly important for modern computer users, who often store private and sensitive data on their personal devices. Although encryption can hide data contents, encryption alone cannot hide the presence of encrypted data. Instead, we would like to hide data with plausible deniability so that even potent adversaries cant tell if system is even hiding data.

This work presents a new approach for hiding sensitive data within a larger data set on NAND flash devices. From a technical perspective, flash is well-suited for data hiding because it offers high-density, fast random access, and non-volatile storage, but with an abundance of internal randomness that is typically masked by on-device firmware. On the practical side, flash is ubiquitous in embedded systems, mobile phones, USB thumb drives, and in solid-state disks (SSDs) on personal laptops—precisely the type of devices that are most likely to be lost, stolen, or confiscated. SSDs are also significant in data centers and servers, which could also be the subject of search or seizure.

This paper presents a new technique to hide data in flash by manipulating the voltage level of pseudo-randomly-selected flash cells to encode two bits (rather than one) in the cell. In this model, we have one "public" bit interpreted using an SLC-style encoding, and extract a private bit using an MLC-style encoding. The locations of cells that encode hidden data is based on a secret key known only to the hiding user.

Intuitively, this technique requires that the voltage level in a cell encoding data must be (1) not statistically distinguishable from a cell only storing public data, and (2) the user must be able to reliably read the hidden data from this cell. Our key insight is that there is a wide enough variation in the range of voltage levels in a typical flash device to obscure the presence of fine-grained changes to a small fraction of the cells, and that the variation is wide enough to support reliably re-reading hidden data.

We store hidden data by transparently increasing the densities of select flash cells, but without creating a detectable deviation in the overall cell voltage distribution. Our data hiding scheme, called VolTage-HIde (VT-HI) [1], effectively mimics the incremental storage technique internally employed by flash vendors using only standard and widely-supported commands. Consequently, VT-HI is feasible in existing flash-based devices without any hardware modifications. Compared to prior work, our technique provides 24x and 50x higher encoding and decoding throughput and doubles the capacity, while being 37x more power efficient. VT-HI allows users to rewrite data and read it multiple times without damaging the integrity of public data, and without decreasing flash lifetime.

REFERENCES

[1] Aviad Zuck, Yue Li, Jehoshua Bruck, Donald E. Porter, and Dan Tsafrir. 2018. Stash in a Flash. In *16th USENIX Conference on File and Storage Technologies (FAST)*.

SYSTOR '18, June 4–7, 2018, HAIFA, Israel

ACM ISBN 978-1-4503-5849-1/18/06…$15.00
https://doi.org/10.1145/3211890.3211906