Self-Encrypting Drives for Data-at-Rest Security

A Self-Encrypting Drive (SED) is a Solid-State Drive (SSD) or Hard Disk Drive (HDD) with built-in security to safeguard the confidentiality of data stored on the device. High-level, a SED consists of large amounts of media for data storage, has a controller chip for device management, and implements a security subsystem for added protection.

Primarily SEDs are devices with large amounts of non-volatile memory media. The media is typically electronic memory based on NAND gate technology in case of SSDs or magnetic recording media for HDDs (but other forms of non-volatile storage media do exist). A SED secures all data stored on its media against unauthorized disclosure using cryptography and access controls.

Besides media, another key component of an SED is a controller chip that runs a dedicated real-time operating system implementing much of the device's functionality. Besides facilitating data transport to and from the device (using one of several storage transport protocols as defined in standards such as NVMe, SATA, SAS), it also implements a high-speed Full Disk Encryption (FDE) engine. This FDE engine sits in the data path of the chip and encrypts or decrypts any data that moves to or from the media. This requires a sophisticated security infrastructure to ensure all data encryption keys are of high quality and always protected.

The final key element of a SED is its security subsystem, implemented as a combination of (dedicated) hardware and firmware. It manages security specific tasks such as key generation, wrapping, and destruction, device secure boot, data access controls, and more. Recent generations of SEDs may contain a root-of-trust that isolates all security sensitive operations and provides validated and measured boot. The security subsystem is based on a set of open standards from the Trusted Computing Group (TCG) and exposes a protocol interface through which the host system can manage the device. The main TCG standards governing SEDs are the Opal SSC and Enterprise SSC Storage Specifications, each referencing multiple other standards.

The SED provides two ways to maintain data confidentiality: instant secure erase for destruction and authentication & authorization to manage access. Data instant secure erase, or crypto erase, involves the destruction of keys used to encrypt data stored on media. Once a key is erased, cleartext data is no longer recoverable as the encrypted data (cipher text) cannot be decrypted without a valid key. Crypto Erase is governed by multiple standards, such as IEEE2883: 2022, NIST SP800-88rev1, ISO27040:2015, TCG Storage Opal or Enterprise SSC Specification, NVM Express, INCITS T10 SCSI, and INCITS T13 ATA.

Data Access Controls on a SED are enforced at the device level, requiring the end-user or a system level process to authenticate. A valid authentication and authorization will unlock data access until the next device power cycle or after sending a lock instruction to the device. Unless unlocked, a SED will reject any requests to read from or write data to the device. If data access controls are enabled, then only a successful authentication will release the encryption key for data encryption or decryption purposes.

Under the Hood

Critical to the security of SEDs are a correct implementation of an approved confidentiality algorithm and high-quality encryption keys. For data encryption most modern SEDs implement the NIST version of the Advanced Encryption Standard (AES)[1] algorithm with a key length of 256 bits and a NIST approved confidentiality mode, e.g., XTS-AES[2]. As for encryption keys, those are used for data encryption and to protect other keys or security sensitive information within the device. The security of encryption keys is determined by how they are generated: using an approved random bit generator[3] and multiple high-quality entropy sources. If not done correctly then the resulting poor-quality keys may endanger the data stored on media by enabling brute force attacks.

Any user data passing through a SED interface is encrypted or decrypted by the FDE engine with a Data Encryption Key (DEK) and e.g., the XTS-AES-256 symmetric key encryption algorithm. The host system can (optionally) partition the device's storage space into multiple addressable bands (also known as Ranges) to support secure data separation. Each Range is then assigned its own DEK. All DEKs are generated by the device and never leave the device, i.e., there is no formal mechanism to escrow a DEK. Each DEK is encrypted by a Key Encryption Key (KEK) using an approved key wrapping algorithm to ensure the cleartext DEK is never exposed within the device up to the moment it is used by the FDE engine. For the FDE engine to load the cleartext DEK it will need access to the KEK that protects the DEK. Although there are multiple ways to implement this, most architectures use a device unique KEK that is only accessible by the FDE engine (sometimes referred to as the master key or root key). This device-unique KEK protects the confidentiality of the DEK outside the FDE engine.

In some SED implementations the KEK encrypted DEK is also wrapped by another key (the Transfer EK, or TEK) in support of other uses cases such as multi-user access or secure DEK destruction. For reliability purposes the device will manage several copies of a TEK-encrypted-KEK-encrypted-DEK and store those in multiple physical locations on the media. In the event of a crypto erase the device will destroy the TEK, thus rendering any copy of the DEK protected by that TEK unusable.[4]

The last layer of protection in the key hierarchy comes from the access control authentication credential, or Authentication Key (AK). This credential is managed outside the device and authenticates an end-user or system process. It is used to unlock or lock data access, invoke a crypto erase, or manage the device's access control policy. Currently, AKs are passphrases up to 32-bytes in length. They are set during device provisioning or as part of a credential rotation event and can be derived from multiple authentication factors. The AK is used to protect the TEK (by creating an intermediate key via an approved derivation mechanism) and is never stored (in cleartext) on the device. This avoids an off-line attack that involves searching the media for the actual AK.

For a host system to read from or write data to the SED, internally the key hierarchy is unwrapped to get to the cleartext DEK. It all starts with sending the AK from the host system to the device. After a

[1] As defined in NIST FIPS PUB 197.

[2] As defined in NIST SP800-38E.

[3] The NIST SP800-90 series of specifications define algorithms for random bit generation, entropy sources used in random bit generation, and a construction for implementing random bit generators.

[4] It should be noted that SED manufacturers may choose to implement different variants of the key hierarchy. The one outlined in this paper is just one approach, but others exist.

successful authentication with the AK, the TEK is decrypted and then used to unwrap the KEK-encrypted-DEK. The KEK-encrypted-DEK is then loaded into the FDE engine, after which the KEK is used to decrypt the DEK. Finally, the DEK is expanded in the FDE algorithm hardware, ready to encrypt or decrypt any data.

Besides encryption keys and AKs used to protect DEKs, a TCG compliant SED has additional credentials that protect other device functionality. These credentials or keys fall into three categories: device managed, manufacturer managed, and user managed. The number and kind of credentials differ per device manufacturer and (optional) features implemented by the SED[5]. It is important that each party (device, manufacturer, and end-user) properly manages these credentials for the ensure secure deployment of SEDs.

Another important aspect of SED data security is device firmware security. Keys and user data are at risk when rogue firmware that circumvent the SED's security runs on the device. To mitigate the risk of rogue firmware the SED implements firmware validation at device boot and during a firmware update. This requires the manufacturer to electronically sign the firmware binary (using an approved digital signature scheme) and for the device to validate that signature when applicable. Additionally, the firmware can be encrypted to protect its confidentiality.

Independent Evaluations

For end-users of SEDs it is important to know that the security infrastructure implementation is done correctly. Mistakes are easily made given the complexity of the security infrastructure of a SED and all the cryptography. Trusting the manufacturer to do it right is not always enough; that is where independent or 3rd-party evaluations come into the picture.

Security evaluation schemes define requirements for security functionality. Independent and accredited laboratories validate the security implementation against the requirements of a scheme. There are at least three evaluation schemes that apply to SEDs: NIST's Cryptographic Algorithm Validation Program (CAVP), the Cryptographic Module Validation Program (CMVP, also known as FIPS140), and an ISO based Common Criteria evaluation.

Implementations of cryptographic algorithms and components are evaluated by CAVP. Algorithm specific test vectors are used to verify the implementation. Each individual algorithm implementation is awarded a CAVP certificate and listed on the NIST Algorithm Validation Page after successfully passing the test. CAVP algorithm certificates are a pre-requisite for both a NIST FIPS140 certification and a NIAP Common Criteria evaluation.

CMVP is a joint US and Canadian evaluation program of cryptographic modules against the latest revision of the NIST FIPS140 standard, currently revision 3 or FIPS140-3. Products are tested by one of the CMVP accredited testing laboratories to validate its compliance with NIST FIPS140. Any product that successfully passes the evaluation will obtain a certificate and is listed on the CMVP Validated Modules page.

[5] They may include e.g. a manufacturer authentication credential for diagnostics access controls, firmware signature verification credentials, a physical ownership credential (PSID), device identity and attestation credentials, a device ownership credential (SID), or secure messaging keys.

As a member of the international Common Criteria Recognition Agreement, the National Information Assurance Partnership (NIAP) manages the Common Criteria evaluation scheme and the approved Common Criteria Protection Profiles for the United States. NIAP has adopted multiple approved Protection Profiles, including the "collaborative Protection Profile for Full Drive Encryption," used to evaluate the security compliance of Self Encrypting Drives. The evaluation is done by one of the Common Criteria accredited laboratories and certified products are listed on NIAP's Product Compliant List (PCL).

Depending on the use case, one or more evaluations are recommended to ensure the end user receives a SED that meets basic security requirements. Although passing one or more evaluations does not guarantee ultimate security, it does lower the risk of improper algorithm implementations and certain avoidable security mistakes.