Design and Implementation of a Windows Kernel Driver for LUKS2-encrypted Volumes

I do not know yet whether I want to have a subtitle, have a placeholder for now

Max Ihlenfeldt

Universität Augsburg Lehrstuhl für Organic Computing Bachelorarbeit im Studiengang Informatik

Copyright © 2021 Max Ihlenfeldt

This document is licensed under the Creative Commons Attribution-ShareAlike 4.0 International Public License (CC BY-SA 4.0).

Abstract

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore e

Contents

1 Introduction	 		 	1
2 Background	 		 	2 2
2.1.1 On-Disk Format				2 4
2.1.3 Using an Unlocked Partition				6
2.2 Introduction to Windows Kernel Driver Development				8
2.2.1 Structure and Hierarchy of the Windows Operating Syst				8
2.2.2 The Windows Driver Model for Kernel Drivers				9
2.2.3 Communication Between Kernel and Userspace	 	•	 •	9
3 Related Work	 		 	10
3.1 Measuring Filesystem Driver Performance				10
3.2 Cryptographic Aspects of LUKS2	 			10
4 Other Approaches				11
4.1 Linux Kernel Implementation of LUKS2				11
4.2 VeraCrypt				11
4.3 BitLocker				11
5 Design and implementation of our approach				12
5.1 Failed Attempts				12
5.2 The Final WDM Driver				12
5.2.1 Architecture				12
5.2.2 Initialization and Configuration				12
5.2.3 De-/encrypting Reads and Writes				12
5.2.4 Handling Other Request Types				$\frac{12}{12}$
5.3 Security Considerations				12
6 Performance of Our Driver				12
6.1 First Experiments				
6.2 Final Experimental Setup				
6.3 Results				13
7 Discussion	 		 	14
8 Conclusion	 		 	15
List of Figures	 		 	16
List of Tables	 		 	17
References				18

1 Introduction

Explain use case etc.

Note that in this thesis the terms disk, drive, volume and partition are used somewhat loosely and probably mean roughly the same.

2 Background

2 Background

2.1 LUKS2 Disk Encryption

Linux Unified Key Setup 2, or short LUKS2, is the second version of a disk encryption standard. It provides a specification [1] for a on-disk format for storing the encryption metadata as well as the encrypted user data. Unlocking an encrypted disk is achieved by providing one of possibly multiple passphrases or keyfiles. The intended usage of LUKS2 is together with the Linux dm-crypt subsystem, but that is not mandatory¹.

The differences between the original LUKS and LUKS2 are minor. According to [1], LUKS2 adds "more flexible ways of storing metadata, redundant information to provide recovery in the case of corruption in a metadata area, and an interface to store externally managed metadata for integration with other tools." Practically, this means that LUKS2 has a different on-disk layout and, among other things, supports more password hashing algorithms (more precisely, password-based key derivation functions).

The reference implementation² is designed only for usage on Linux, which is why we developed a new Rust library for interacting with LUKS2 partitions. This is not a full equivalent, but only a cross-platform helper. Its task is to take care of all the cryptographic work needed before actually decrypting and encrypting data. Notably, it lacks the following features of the reference implementation:

- formatting new LUKS2 partitions,
- modifying or repairing existing LUKS2 partitions,
- converting a LUKS partition to a LUKS2 partition,
- actually mounting a LUKS2 partition for read/write usage (this is what our kernel driver and its userspace configuration tool is for).

Our library does provide access to the raw decrypted user data, but the practical use of this is very limited: the decrypted data is in the format of a filesystem, e.g. FAT32, btrfs, or Ext4. Therefore a filesystem driver is needed to actually access the stored files. One way of exposing the decrypted data to the system's filesystem drivers is by transparently decrypting the data directly in the kernel, which is what our driver does (see section 5).

2.1.1 On-Disk Format

Figure 1 shows the high-level layout of a LUKS2-encrypted disk.

The two binary headers have a size of exactly one sector, so that they are always written atomically. Only the first 512 bytes are actually used. The header marks the disk as following the LUKS2 specification, and contains metadata such as labels, a UUID, and a header checksum. The labels and UUID can be accessed using the blkid³ command-line tool and also be used in the udev⁴ Linux subsystem. For the detailed contents, see Figure 2. Figure 3 also contains an example hexdump of a binary header.

The sector containing the binary header is followed by the JSON area. This area arguably contains the metadata that is most relevant for decryption and encryption. Figure 4 contains an overview of the objects stored in JSON and their relationships. For

¹ As we show in this thesis, it is possible to make the combination of LUKS2 and Windows work.

² https://gitlab.com/cryptsetup/cryptsetup

³ https://linux.die.net/man/8/blkid

⁴ https://linux.die.net/man/8/udev

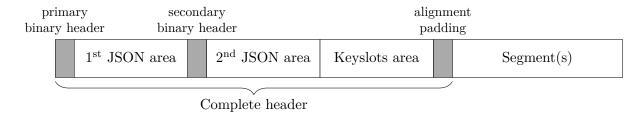


Figure 1: LUKS2 on-disk format (modified after [1]). The complete header consists of three areas: a binary header of exactly one 4096-byte sector, JSON metadata, and the binary keyslots data. A *keyslot* is an "encrypted area on disk that contains a key" [1]. For redundancy, the binary header and the JSON metadata are stored twice. After that follow one or areas containing encrypted user data. The specification calls these areas *segments*.

```
#define MAGIC_1ST "LUKS\xba\xbe"
\#define\ MAGIC\_2ND\ "SKUL\xba\xbe"
#define MAGIC_L
                  40
#define UUID L
                  48
#define LABEL L
#define SALT_L
#define CSUM_ALG_L 32
#define CSUM_L
struct luks2 hdr disk {
   char magic[MAGIC L];
                                   // MAGIC 1ST or MAGIC 2ND
   uint16_t version;
                                   // Version 2
   uint64_t hdr_size;
                                   // size including JSON area [bytes]
                                   // sequence ID, increased on update
   uint64_t seqid;
                                   // ASCII label or empty
    char
            label[LABEL_L];
    char
            csum_alg[CSUM_ALG_L]; // checksum algorithm, "sha256"
   uint8_t salt[SALT_L];
                                  // salt, unique for every header
                                  // UUID of device
    char
            uuid[UUID_L];
            subsystem[LABEL_L]; // owner subsystem label or empty
    char
   uint64_t hdr_offset;
                                   // offset from device start [bytes]
   char _padding[184];
uint8_t csum[CSUM_L];
                                   // must be zeroed
                                   // header checksum
           _padding4096[7*512];
                                   // Padding, must be zeroed
} __attribute__((packed));
```

Figure 2: LUKS2 binary header structure from [1]. Integers are stored in big-endian format, and all strings have to be null-terminated. The magic, version, and unid fields are also present in the LUKS1 binary header and were placed at the same offsets as there.

this thesis' brevity's sake, please refer to Chapter 3.1 in [1] for an example of a LUKS2 JSON area.

After the JSON area, the keyslots area is stored on the disk. This is space reserved for storing encrypted cryptographic keys. The metadata from the JSON keyslot objects describe the position of a key on the disk as well as information on how to decrypt it.

4 Background

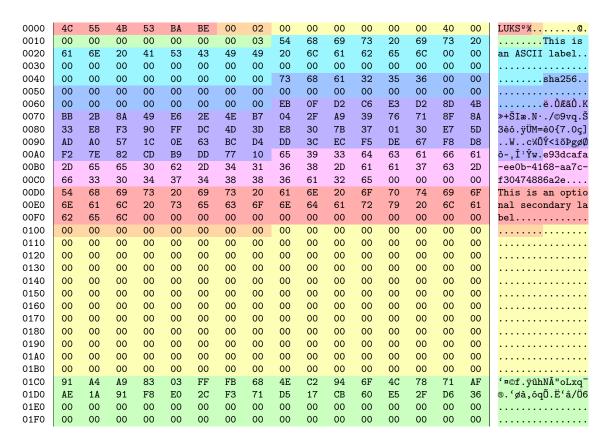


Figure 3: LUKS2 binary header example. The fields, as described in Figure 2, were coloured differently to be easily distinguishable. A similar header, although with different salt and hash, can be generated by executing fallocate -1 16M luks2.img && cryptsetup luksFormat --label 'This is an ASCII label' --subsystem 'This is an optional secondary label' --uuid e93dcafa-ee0b-4168-aa7c-f30474886a2e luks2.img in a Linux shell.

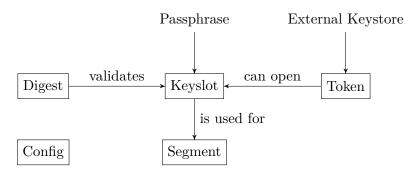


Figure 4: LUKS2 object schema from [1]. The most important objects are the following: keyslots, which describe the details of how cryptographic keys are stored and encrypted; digests, which can be used to verify that one has successfully extracted a key from a keyslot; and segments, which describe the disk areas where the encrypted user data is stored. Figure 1 shows where the areas described by the keyslot and segment objects actually lie on disk.

2.1.2 Unlocking a Partition

For simplicity, our LUKS2 Rust library does not support unlocking a keyslot using an external keystore defined by a token. Only unlocking via password is implemented. The library does however include support for different password-based key derivation functions

(PBKDFs), namely pbkdf2 with SHA-256, argon2i, and argon2id. These are all the PBKDF algorithms that are listed in the LUKS2 specification (see [1], Table 3). The default PBKDF used by LUKS2 is argon2i [2].

The LUKS2 specification allows for multiple segments in one partition. To make things easier, our driver only supports unlocking one segment. Therefore, in this thesis we may speak of unlocking a partition and mean unlocking one of the partition's segments.

To unlock a segment means to derive the cryptographic key that is needed for reading decrypted or writing encrypted data. This key is called the segment's master key.

LUKS2 uses a process called *anti-forensic splitting* to store the master key on the disk. This method was introduced in [3]. It is used to diffuse the key's bytes into a longer sequence of bytes that has the following property: if at least one bit of the diffused sequence is changed, the key cannot be recovered. This is achieved by a clever combination of XOR and a hash function. The motivation behind this to make it easier (or possible) to dispose of an old key in such a way that it cannot be recovered from the disk. This is because it is much more feasible to partially erase a long sequence of bytes than to completely erase a short sequence. Erasing here means to overwrite the data in such a way that it cannot be recovered, which is not as trivial as one might think.

[3] calls the operation that splits data anti-forensically AFsplit and the recovery operation AFmerge. We will adhere to this convention (with slight variations).

To necessitate the need of a password to recover the key, the data is also encrypted before it gets written to the disk. The encryption key is a hash of the password obtained by a PBKDF.

The properties of anti-forensic splitting can be used when the user wants to change the password: the master key is derived using the old password and then re-encrypted with the new password. The key as it was encrypted with the old password can then be destroyed.

[3] presents two templates for storing keys, TKS1 and TKS2. The difference is whether the key is encrypted before or after splitting it. LUKS and LUKS2 use TKS2, which is schematically explained in Figure 5. The hash function used by LUKS2 for anti-forensic splitting is SHA-256. Figure 6 shows the outline of an implementation of TKS2 in Rust.

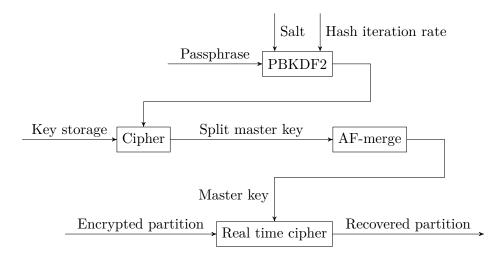


Figure 5: TKS2 scheme (modified after [3]).

The keyslots area on the disk is large enough to store multiple split and encrypted master keys. Thus one can configure a LUKS2 partition to be unlocked by different passwords. Unlocking then works as described in Figure 7.

6 Background

```
fn decrypt_keyslot(
1
        password: &[u8], keyslot: &LuksKeyslot, json: &LuksJson, /* ... */
2
    ) -> Result<Vec<u8>, LuksError> {
3
        let mut k = vec![
4
            0; keyslot.key_size() as usize * keyslot.af.stripes() as usize
5
6
        // read keyslot data from disk into k...
7
8
        let mut pw_hash = vec![0; area.key_size() as usize];
9
        match keyslot.kdf() {
10
            // hash into pw_hash using pbkdf2, argon2i, or argon2id...
11
        }
12
13
        // decrypt keyslot area using the password hash as key
14
        match area.key_size() {
15
            32 => {
16
                 let key1 = Aes128::new varkey(&pw hash[..16]).unwrap();
17
                 let key2 = Aes128::new_varkey(&pw_hash[16..]).unwrap();
18
                 let xts = Xts128::<Aes128>::new(key1, key2);
                 xts.decrypt_area(&mut k, sector_size, 0, get_tweak_default);
20
            },
21
             // 64 byte key uses AES256 instead...
22
        }
23
24
         // merge and hash master key
25
        let master_key = af::merge(
26
            &k, keyslot.key_size() as usize, af.stripes() as usize
27
        );
28
        let digest_actual = base64::decode(json.digests[&0].digest())?;
29
        let mut digest_computed = vec![0; digest_actual.len()];
30
        let salt = base64::decode(json.digests[&0].salt())?;
31
        pbkdf2:: pbkdf2:: < Hmac < Sha256 >> (
32
            &master_key, &salt, json.digests[&0].iterations(), &mut digest_computed
33
        );
34
35
        // compare digests
36
        if digest_computed == digest_actual {
37
            Ok(master_key)
39
        } else {
            Err(LuksError::InvalidPassword)
40
        }
41
42
```

Figure 6: LUKS2 master key decryption in Rust. Some values are hardcoded: only the digest with index 0 is used (lines 29, 31, 33), and it is assumed that the digest algorithm is always pbkdf2 with SHA-256 (line 32). The latter is compliant with the specification, which lists this digest algorithm as the only option, but not optimal in the sense of input validation.

2.1.3 Using an Unlocked Partition

After a partition has been unlocked, i.e. after the master key of one of its segments has been decrypted, the partition is ready to be read from and written to. This happens using what Figure 5 calls a real time cipher. LUKS2 supports different encryption algorithms for this purpose, see Figure 8 for a selection of them. Our driver only supports the default aes-xts-plain64 encryption. Therefore we will focus on that in this section.

- 1. The user supplies a password.
- 2. One of the available keyslots is selected.
- 3. Using the password and keyslot, the master key is decrypted as described above.
- 4. The derived master key is hashed and the result compared to the corresponding digest. Which digest and what hash parameters to used is defined in the JSON section.
- 5. If the digests match, the master key has been successfully decrypted. Else go to step 2 and select a keyslot that has not been used yet.
- 6. If the master key could not be decrypted with all available keyslots, the supplied password was not correct.

Figure 7: LUKS2 master key decryption with multiple available keyslots. LUKS2 allows defining priorities that govern the order in which the available keyslots are tried. For a more detailed pseudocode see Figure 5 in [4].

${f Algorithm}$	Description		
in dm-crypt notation	Description		
aes-xts-plain64	AES in XTS mode with sequential IV		
aes-cbc-essiv:sha256	AES in CBC mode with ESSIV IV		
serpent-xts-plain64	Serpent cipher with sequential IV		
twofish-xts-plain64	Twofish cipher with sequential IV		

Figure 8: Selection of LUKS2 encryption algorithms (modified after [1]). The dm-crypt notation follows this format: cipher[:keycount]-chainmode-ivmode[:ivopts] [5]. See [6] for the CBC mode and the Serpent and Twofish ciphers, and [3] for the ESSIV IV mode.

The AES encryption algorithm is a block cipher for processing 128 bit data blocks [7]. To encrypt data longer than one block, a block cipher mode is needed [6]. These are encryption functions that build on a existing block cipher. When using aes-xts-plain64, LUKS2 uses the XTS block cipher mode. The defining IEEE standard [8] describes it as follows: "XTS-AES is a tweakable block cipher that acts on data units of 128 b[its] or more and uses the AES block cipher as a subroutine. The key material for XTS-AES consists of a data encryption key (used by the AES block cipher) as well as a 'tweak key' that is used to incorporate the logical position of the data block into the encryption." This tweak key is called the *initialization vector*, or IV, in the context of LUKS2. This is what the "plain64" in aes-xts-plain64 means: "the initial vector is the 64-bit little-endian version of the sector number, padded with zeros if necessary" [5]. This sector number is relative to the first sector of the segment, i.e. the first sector uses an IV of 0.

The need for an IV arises from a critical problem that occurs when encrypting each block separately with the same key: if some unencrypted blocks are identical, then so will be their encrypted counterparts. This can lead to leaked information about structure and contents of the plaintext [6].

All this theory may sound complicated, but in subsubsection 5.2.3 we will see that the

8 Background

practical usage is quite simple⁵.

2.2 Introduction to Windows Kernel Driver Development

This section gives an introduction on the development of Windows kernel drivers and related important concepts.

2.2.1 Structure and Hierarchy of the Windows Operating System

First, we will introduce some basic concepts of the Windows operating system that will be relevant in the following sections. Please note that we will focus on Windows 10 running on the x64 architecture, as that is the target of our driver. Some details may be different in other versions of Windows, though most information should still be valid.

- kernel vs user mode (page 23)
 - syscalls (explanation from page 7?)
 - OS model (page 46)
 - Executive (page 72)
 - Kernel (page 75)
- Architecture overview (Figure 2-1 and or Figure 2-4)
- HAL (page 79) (briefly)
- registry (page 32)
- mention Windows Sysinternals (esp. ProcMon, WinObj)
- mention/cite Microsoft Developer Documentation
- Windows SDK/API (Win32), Windows Driver Kit (page 43)
- DLLs?
- processes? (page 8) threads? (page 18)
- WinDbg (or in a later (sub)section) (page 38)
- virtual memory, paging (page 21)
- Driver Verifier?
- Test Mode (page 24)
- Objects and handles (page 30)
- Table 1-4 in [9] contains a list of useful tools
- Table 2-5 lists commonly used function prefixes

⁵ The implementation of cryptographic algorithms of course remains non-trivial. Implementing AES itself has been made much easier using the AES-NI CPU instruction set, though.

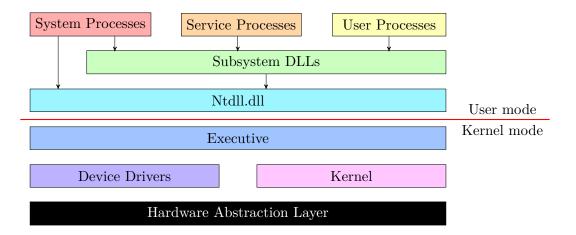


Figure 9: Simplified Windows architecture (modified after [9]).

2.2.2 The Windows Driver Model for Kernel Drivers

Also explain how it gets loaded (if not done already) page 83

• IRPs and other IO stuff

2.2.3 Communication Between Kernel and Userspace

Via ports

10 Related Work

- 3 Related Work
- 3.1 Measuring Filesystem Driver Performance
- 3.2 Cryptographic Aspects of LUKS2

[3]

Search for more papers, e.g. attacks against LUKS? e.g. [10]

4 Other Approaches

4.1 Linux Kernel Implementation of LUKS2

https://web.archive.org/web/20210728082434/https://blog.cloudflare.com/speeding-up-li

- 4.2 VeraCrypt
- 4.3 BitLocker

[11] and [12] and [13] and [14]

Use Ghidra and look at some of the code of fvevol.sys?

5 Design and implementation of our approach

5.1 Failed Attempts

FilterManager framework

Mention KMDF / UMDF and why we didn't use that if not already done in earlier section

5.2 The Final WDM Driver

Why WDM?

- 5.2.1 Architecture
- 5.2.2 Initialization and Configuration

luks2filterstart.exe

5.2.3 De-/encrypting Reads and Writes

custom AES implementation

make sure this chapter and subsubsection 2.1.3 are not too similar

LUKS2 supports many encryption algorithms (see [1], Table 4), but luks2flt only supports aes-xts-plain64.

```
VOID
EncryptWriteBuffer(
   PUINT8 Buffer,
   PLUKS2_VOLUME_INFO VolInfo,
   PLUKS2_VOLUME_CRYPTO CryptoInfo,
   UINT64 OrigByteOffset,
   UINT64 Length
)
   UINT64 Sector = OrigByteOffset / VolInfo->SectorSize;
   UINT64 Offset = 0;
   UINT8 Tweak[16];
    while (Offset < Length) {
        ToLeBytes(Sector, Tweak);
        CryptoInfo->Encrypt(
            &CryptoInfo->Xts, Buffer + Offset,
            VolInfo->SectorSize, Tweak
        Offset += VolInfo->SectorSize;
        Sector += 1;
    }
```

5.2.4 Handling Other Request Types

5.3 Security Considerations

How does cryptsetup send the master key to dm-crypt

6 Performance of Our Driver

6.1 First Experiments

6.2 Final Experimental Setup

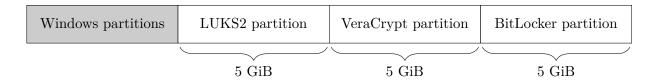


Figure 10: Disk layout of the real hardware SSD.

6.3 Results

For testing purposes, we tried optimizing the performance by restricting support to AES256-XTS. This enabled removing some if-else constructs that dispatched de-/encryption functions based on whether AES128 or AES256 was used. Even though these conditionals were located in a performance critical section, we saw no speed improvements. Our theory for why this was the case is the following: our driver was only ever used for one LUKS2 partition and therefore always took the same path through the if-else (either always AES128 or always AES256). This trained the CPU's branch prediction on this one specific path. Thus, after a short training phase, the CPU always speculatively executed the correct path, resulting in the same performance as without the if-else.

14 Discussion

7 Discussion

8 Conclusion

16 List of Figures

List of Figures

1	LUKS2 on-disk format
2	LUKS2 binary header structure
3	LUKS2 binary header example
4	LUKS2 object schema
5	TKS2 scheme
6	LUKS2 master key decryption in Rust
7	LUKS2 master key decryption with multiple available keyslots
8	Selection of LUKS2 encryption algorithms
9	Simplified Windows architecture
10	Disk layout of the real hardware SSD

List of Tables 17

List of Tables

18 References

References

[1] M. Broz, LUKS2 On-Disk Format Specification Version 1.0.0, 2018, visited on 2021-07-31. [Online]. Available: https://gitlab.com/cryptsetup/LUKS2-docs/-/raw/861197a9de9cba9cc3231ad15da858c9f88b0252/luks2_doc_wip.pdf

- [2] Frequently Asked Questions Cryptsetup/LUKS, 2020, visited on 2021-08-10. [Online]. Available: https://gitlab.com/cryptsetup/cryptsetup/-/wikis/FrequentlyAskedQuestions?version_id=6297db166f8ae73c6a616ba874a12f5d43d37fd9
- [3] C. Fruhwirth, "New methods in hard disk encryption," Ph.D. dissertation, Vienna University of Technology, 2005.
- [4] C. Fruwirth, LUKS1 On-Disk Format Specification Version 1.2.3, 2018, visited on 2021-07-31. [Online]. Available: https://mirrors.edge.kernel.org/pub/linux/utils/cryptsetup/LUKS_docs/on-disk-format.pdf
- [5] dm-crypt: Linux device-mapper crypto target, 2020, visited on 2021-08-06. [Online]. Available: https://gitlab.com/cryptsetup/cryptsetup/-/wikis/DMCrypt? version_id=ee336a2c1c7ce51d1b09c10472bc777fa1aa18cd
- [6] N. Ferguson, B. Schneier, and T. Kohno, Cryptography Engineering Design Principles and Practical Applications. Wiley, 2010.
- [7] National Institute of Standards and Technology, "Advanced Encryption Standard (AES)," U.S. Department of Commerce, Tech. Rep., 2001.
- [8] IEEE Standard for Cryptographic Protection of Data on Block-Oriented Storage Devices, IEEE Std. 1619-2018 (Revision of IEEE Std. 1619-2007), 2019.
- [9] P. Yosifovich, D. A. Solomon, and A. Ionescu, Windows Internals, Part 1: System architecture, processes, threads, memory management, and more. Microsoft Press, 2017.
- [10] V. Polášek, "Argon2 security margin for disk encryption passwords," Master's thesis, Masaryk University, 2019.
- [11] J. D. Kornblum, "Implementing BitLocker drive encryption for forensic analysis," *Digital Investigation*, vol. 5, no. 3-4, pp. 75–84, 2009.
- [12] S. G. Lewis and T. Palumbo, "BitLocker full-disk encryption: Four years later," in *Proceedings of the 2018 ACM SIGUCCS Annual Conference*, 2018, pp. 147–150.
- [13] S. Türpe, A. Poller, J. Steffan, J.-P. Stotz, and J. Trukenmüller, "Attacking the Bit-Locker boot process," in *International Conference on Trusted Computing*. Springer, 2009, pp. 183–196.
- [14] C. Tan, L. Zhang, and L. Bao, "A deep exploration of BitLocker encryption and security analysis," in 2020 IEEE 20th International Conference on Communication Technology (ICCT). IEEE, 2020, pp. 1070–1074.