LUKS2 On-Disk Format Specification

version 1.1.4, 2025-06-16

Milan Brož <gmazyland@gmail.com>

1 Introduction

LUKS2 is the second version of the *Linux Unified Key Setup* for disk encryption management. It is the follow-up of the LUKS1 [1, 2] format that extends capabilities of the on-disk format and removes some known problems and limitations. Most of the basic concepts of LUKS1 remain in place as designed in *New Methods in Hard Disk Encryption* [2] by Clemens Fruhwirth.

LUKS provides a generic key store on the dedicated area on a disk, with the ability to use multiple passphrases[1] to unlock a stored key. LUKS2 extends this concept for 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.

While the implementation of LUKS2 is intended to be used with Linux-based dm-crypt [3] disk encryption, it is a generic format.

1.1 Design Goals

The LUKS header provides metadata for a disk encryption setup. LUKS1 version [1] contains a binary header for storing necessary metadata (like encryption algorithms parameters) and eight keyslots for independent passphrases to unlock one volume key.
The LUKS2 format is designed to provide these features:

- Cover all possibilities of LUKS1.
- Support configurable memory-hard key-derivation algorithms.
- The new header can store additional metadata for external tools and expose an interface for regular updates.
- LUKS2 uses only a small binary header that can be easily used by automatic detection tools like *blkid* in Linux. The binary header is partially compatible with LUKS1, so legacy tools still recognize a partition as the LUKS type and can see device UUID. All other metadata are stored in the non-binary header.
- The header includes a checksum mechanism that detects data corruption and unintentional header mangling.
- LUKS header can be detached from a LUKS device and can be stored on a separate device or in a file. With the detached header, the encrypted device contains no visible and detectable metadata.

[1] LUKS can use a passphrase or a key file, both are processed identically.

- Metadata area is stored in two copies to allow for a possible recovery.[2] The recovery is transparent for most of the operations (device should recover automatically if at least one header is correct).
- Keyslot binary area is not duplicated (for security reasons), but the area is now allocated in higher device offset where a random data corruption should happen more rarely.
- A header can be upgraded in-place for most of existing LUKS1 devices.[3]
- Header store persistent flags that are used during activation.[4]
- The number of keyslots is limited only by the provided header area size.[5]
- Keyslots have priorities. Some keyslots can be marked for use only if explicitly specified (for example as a recovery keyslot).
- Metadata are stored in the JSON format that allows for future extensions without modifying binary structures. Such an extension is for example support for authenticated encryption or support for online data reencryption.
- All metadata are algorithm-agnostic and can be upgraded to new algorithms later without header structure changes.
- Volume key digest is no longer limited by 20 bytes (based on legacy SHA-1) as in the LUKS1 header.
- Keyslot can contain an unbound key (key not assigned to any encrypted data segment) that can be used for external applications. Size of an unbound key can be different from volume key size in other keyslots.
- The header contains a concept of *tokens* that are objects, assigned to keyslots, which contain metadata describing *where to get unlocking passphrase*. Tokens can be used for support of external key store mechanisms.

1.2 Security Goals

The primary goal for LUKS2 is to provide data confidentiality for user data (user-friendly access to encrypted data).

The secondary goal is to provide availability of the stored metadata in the case of partial and random metadata corruption (except for the encrypted key material where LUKS anti-forensic function [1] is applied).

The format also allows an easy way for backup of the whole LUKS2 header (including key material).[6]

Other goals or additional security measures can be achieved when specific cryptographic primitives or optional extensions are used. Examples of such extensions are data integrity protection (authenticated encryption) or online reencryption (to limit encryption keys lifetime, see Section 4.9).

[2] A common issue with LUKS1 is metadata corruption caused by a partitioning tool that does not recognize LUKS format.

[3] A configuration that does not use default on-disk alignment of user data offset could lack needed space for new metadata area.

[4] Example of a persistent flag is support for the TRIM operation. These flags should replace the need for flags in the external `/etc/crypttab` file.

[5] Reference implementation limits the number of keyslots to 32.

[6] See *luksHeaderBackup* and *luksHeaderRestore* commands in the reference implementation.

The LUKS2 on-disk format is designed to be used on a common off-the-shelf storage device that does not contain any trusted element. Metadata checksum and redundancy can prevent random metadata corruption only. It does not cover intentional modification. An attacker with physical access to the storage device can modify or corrupt part of the visible metadata or ciphertext data, mangle offsets, encryption parameters, or even replace the whole header with different metadata. Such an attack can lead to denying access or corruption in decrypted data. Some metadata attributes are intended to be modified without the physical access to user data (like removing keyslots or setting device labels).

The LUKS2 metadata can be stored in a detached form (separating user data and LUKS2 metadata header).

The token can contain user-defined metadata that are not part of the basic LUKS2 definition (in general, any JSON metadata can be stored as a token). Because the metadata in the LUKS2 header are visible (can be trivially extracted), stored metadata should not contain any fields that significantly help to attack the confidentiality of user data. Additional extensions should store sensitive attributes in encrypted form.

LUKS2 should not be used when you need to provide any form of plausible deniability. For more information, see LUKS/cryptsetup Frequently Asked Questions [4].

1.3 Reference Implementation

The LUKS2 format is currently implemented and fully supported in libcryptsetup [5] on Linux systems, together with the LUKS1 format. The implementation automatically detects version according to the binary header.

New LUKS2 devices can be created with `--type luks2` option. For example, `cryptsetup format --type luks2 <device>`.

2 LUKS2 On-Disk Format

The LUKS2 header is located at the beginning (sector 0) of the block device (for a detached header on a dedicated block device or in a file). The basic on-disk structure is illustrated in Figure 1.

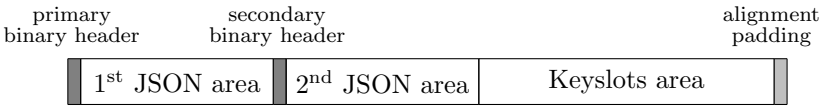

Figure 1: LUKS2 header on-disk structure.

The LUKS2 header contains three logical areas:

- binary structured header (one 4096-byte sector, only 512-bytes are used),
- area for metadata stored in the JSON format and
- keyslot area (keyslots binary data).

The binary and JSON areas are stored twice on the device (primary and secondary header) and under normal circumstances contain same functional metadata. The binary header size ensures that the binary header is always written to only one sector (atomic write). Binary data in the keyslots area is allocated on-demand. There is no redundancy in the binary keyslots area.

2.1 Binary Header

The binary header is intended for a quick scanning by *blkid* and contains a signature to detect LUKS device, basic information (labels), header size and metadata checksum. Binary header in the C structure is described in Figure 2.

All integer values are stored in the big-endian format. All strings are in the C format, and a valid header must have all strings terminated by the zero byte.

The primary binary header must be stored in sector 0 of the device. The secondary header starts immediately after the primary header JSON area (see *hdr_size* in primary header). To allow for an easy recovery, the secondary header must start at a fixed offset listed in Table 1.

Offset (hexa) [bytes]	**JSON area** [kB]
16384 (0x004000)	12
32768 (0x008000)	28
65536 (0x010000)	60
131072 (0x020000)	124
262144 (0x040000)	252
524288 (0x080000)	508
1048576 (0x100000)	1020
2097152 (0x200000)	2044
4194304 (0x400000)	4092

Table 1: Possible LUKS2 secondary header offsets and JSON area size.

The LUKS1 compatible fields (*magic*, *UUID*) are placed intentionally on the same offsets. The binary header contains these fields:

- **magic** contains the unique string (see C defines `MAGIC_1ST` for the primary header and `MAGIC_2ND` for the secondary header in Figure 2).
- **version** must be set to 2 for LUKS2.
- **hdr_size** contains the size of the header with the JSON data area. The offset and size of the secondary header must match this size. It is a prevention to rewrite of a header with a different JSON area size.
- **seqid** is a counter (sequential number) that is always increased when a new update of the header is written. The header with a higher *seqid* is more recent and is used for recovery (if there are primary and secondary headers with different seqid, the more recent one is automatically used).
- **label** is an optional label (similar to a filesystem label).
- **csum_alg** is a checksum algorithm. Metadata checksum covers both the binary data and the following JSON area and is calculated with the checksum field zeroed. By default, plain SHA-256 function is used as the checksum algorithm.
- **salt** is generated by an RNG and is different for every header (it differs on the primary and secondary header), even the backup header must contain a different salt. The salt is not used after the binary header is read, the main intention is to avoid deduplication of the header sector. The salt must be regenerated on every header repair (but not on a regular update).

```
#define MAGIC_1ST "LUKS\xba\xbe"
#define MAGIC_2ND "SKUL\xba\xbe"
#define MAGIC_L     6
#define UUID_L     40
#define LABEL_L    48
#define SALT_L     64
#define CSUM_ALG_L 32
#define CSUM_L     64

// All integers are stored as big-endian.
// Header structure must be exactly 4096 bytes.

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]
  uint64_t seqid;                // sequence ID, increased on update
  char     label[LABEL_L];       // ASCII label or empty
  char     csum_alg[CSUM_ALG_L]; // checksum algorithm, "sha256"
  uint8_t  salt[SALT_L];         // salt, unique for every header
  char     uuid[UUID_L];         // UUID of device
  char     subsystem[LABEL_L];   // owner subsystem label or empty
  uint64_t hdr_offset;           // offset from device start [bytes]
  char    _padding[184];         // must be zeroed
  uint8_t  csum[CSUM_L];         // header checksum
  char    _padding4096[7*512];   // Padding, must be zeroed
} __attribute__ ((packed));
```

Figure 2: LUKS2 binary header on-disk structure.

- **uuid** is device UUID with the same format as in LUKS1.
- **subsystem** is an optional secondary label.
- **hdr_offset** must match the physical header offset on the device (in bytes). If it does not match, the header is misplaced and must not be used. It is a prevention to partition resize or manipulation with the device start offset.
- **csum** contains a checksum calculated with the *csum_alg* algorithm. If the checksum algorithm tag is shorter than the *csum* field length, the rest of this field must be zeroed.

The rest of the binary header (including padding fields) must be zeroed. The *version*, *UUID*, *label* and *subsystem* fields are intended to be used in the *udev* database and *udev* triggered actions. For example, a system can manage all LUKS2 devices with a specific subsystem field automatically by some external tool. The *label* and *UUID* can be used the same way as a filesystem label.

2.2 JSON Area

The JSON area starts immediately after the binary header (end of JSON area must be aligned to 4096-byte sector offset). Size of JSON area is determined from binary header *hdr_size* field: $JSON\ area\ size = hdr\_size - 4096$.

The area contains metadata in JSON format [6]. The JSON metadata are stored in the area as a C string that must be terminated by the zero character. The unused remainder of the area must be empty and filled with zeroes. The header cannot store larger metadata than this fixed JSON area.

2.3 Keyslots Area

Keyslots area is a reserved space on the disk that can be allocated for a binary data from keyslots. There are stored encrypted keys referenced from keyslots metadata. The structure of stored keyslot binary data depends on the keyslot type. The *luks2* keyslots type uses the same LUKS1 binary structure. The *reencrypt* type (optional online reencryption extension, see section 4.9) uses keyslot allocated area to provide redundancy for data recovery in the case of unexpected reencryption interruption (crash on power fail or a similar situation).

The allocated area is defined in a keyslot by an *area* object that contains *offset* (from the device beginning) and *size* fields. Both fields must be validated to point to the keyslot area. Invalid values must be rejected.

3 LUKS2 JSON Metadata Format

The LUKS2 metadata allows defining objects that, according to the *type* field, defines a specific functionality. Objects that are not recognized byt the implementation are ignored, but metadata are still maintained inside the JSON metadata. Implementation must validate the JSON structure before updating the on-disk header.

The LUKS2 structure has 5 mandatory top-level objects (see Figure 3) as follows:

- **config** contains persistent header configuration attributes.
- **keyslots** are objects describing encrypted keys storage areas.
- **digests** are used to verify that keys decrypted from keyslots are correct.
- **segments** describe areas on disk that contain user encrypted data.
- **tokens** can optionally include additional metadata, bindings to other systems – *how to get a passphrase for the keyslot.*

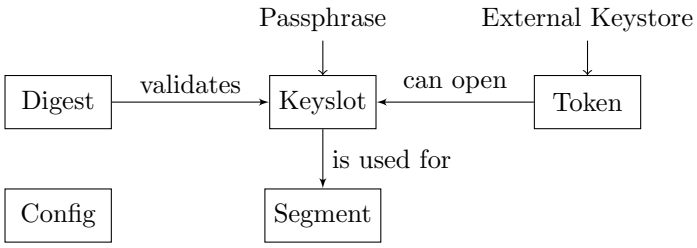

Figure 3: LUKS2 objects schema.

Except top-level objects listed above, all JSON objects must have their names formatted as a string that represents a number in the decimal notation (unsigned integer) – for example *"0"*, *"1"* and must contain attribute *type*. According to the *type*, the implementation decides how to handle (or ignore) such an object. This notation allows mapping to LUKS1 API functions that use an integer as a reference to keyslots objects.

Binary data inside JSON (for example *salt*) are stored in the Base64 [7] encoding. JSON cannot store 64-bit integers directly, a value for an object that represents unsigned 64-bit integer (*offset* or *size*) is stored as a string in the

decimal notation and later converted to the 64-bit unsigned integer. Such an integer is referenced as *string-uint64* later.

3.1 LUKS2 JSON Example

The following example contains full JSON metadata from the reference implementation for a LUKS2 device that is encrypted with the AES-XTS cipher, contains two keyslots and one token. The token type is *keyring* (can be unlocked by a passphrase in the keyring) and it is bound to the second keyslot.

```
{
  "keyslots":{
    "0":{
      "type":"luks2",
      "key_size":32,
      "af":{
        "type":"luks1",
        "stripes":4000,
        "hash":"sha256"
      },
      "area":{
        "type":"raw",
        "encryption":"aes-xts-plain64",
        "key_size":32,
        "offset":"32768",
        "size":"131072"
      },
      "kdf":{
        "type":"argon2i",
        "time":4,
        "memory":235980,
        "cpus":2,
        "salt":"z6vz4xK7cjan92rDA5JF8O6Jk2HouVOO8DMB6GlztVk="
      }
    },
    "1":{
      "type":"luks2",
      "key_size":32,
      "af":{
        "type":"luks1",
        "stripes":4000,
        "hash":"sha256"
      },
      "area":{
        "type":"raw",
        "encryption":"aes-xts-plain64",
        "key_size":32,
        "offset":"163840",
        "size":"131072"
      },
      "kdf":{
        "type":"pbkdf2",
        "hash":"sha256",
        "iterations":1774240,
        "salt":"vWcwY3rx2fKpXW2Q6oSCNf8j5bvdJyEzB6BNXECGDsI="
      }
    }
  },
  "tokens":{
    "0":{
      "type":"luks2-keyring",
```

```
52          "keyslots":[
53            "1"
54          ],
55          "key_description":"MyKeyringKeyID"
56        }
57    },
58    "segments":{
59      "0":{
60        "type":"crypt",
61        "offset":"4194304",
62        "iv_tweak":"0",
63        "size":"dynamic",
64        "encryption":"aes-xts-plain64",
65        "sector_size":512
66      }
67    },
68    "digests":{
69      "0":{
70        "type":"pbkdf2",
71        "keyslots":[
72          "0",
73          "1"
74        ],
75        "segments":[
76          "0"
77        ],
78        "hash":"sha256",
79        "iterations":110890,
80        "salt":"G8gqtKhS96IbogHyJLO+t9kmjLkx+DM3HHJqQtgc2Dk=",
81        "digest":"C9JWko5m+oYmjg6R0t/98cGGzLr/4UaG3hImSJMivfc="
82      }
83    },
84    "config":{
85      "json_size":"12288",
86      "keyslots_size":"4161536",
87      "flags":[
88        "allow-discards"
89      ]
90    }
91  }
```

3.2 Keyslots Object

Keyslots object contains information about stored keys – areas, where binary keyslot data are located, encryption and anti-forensic function used, password-based key derivation function (PBKDF) and related parameters.

Every keyslot object must contain:

- **type** [string] the keyslot type.
- **key_size** [integer] the key size (in bytes) stored in keyslot.
- **area** [object] the allocated area in the binary keyslots area.
- **priority** [integer,optional] the keyslot priority. Here 0 means *ignore* (the slot should be used only if explicitly stated), 1 means *normal* priority and 2 means *high* priority (tried before *normal* priority).

3.2.1 Keyslot Type *luks2*

The *luks2* keyslot type uses the same logic as LUKS1 keyslot, but allows for per-keyslot algorithms (for example different PBKDF).
Only *raw* type is supported for the **area** field.

The *luks2* object must contain these additional fields:

- **kdf** [object] the PBKDF type and parameters used.
- **af** [object] the anti-forensic splitter [1] (only the *luks1* type is currently used).

3.2.2 Keyslot Type *reencrypt* (Optional Extension)

The *reencrypt* keyslot type is present when there is an online reencryption operation in progress (see section 4.9). Allocated binary area is used for redundancy recovery data (it does not contain a key).
The **key_size** field must be set to *1*. The **area** type must be *none*, *checksum*, *journal*, *datashift*, *datashift-journal* or *datashift-checksum*.

The *reencrypt* object must contain these additional fields:

- **mode** [string] the reencryption mode. Only *reencrypt*, *encrypt* and *decrypt* values are supported.
- **direction** [string] the reencryption direction. Only *forward* or *backward* values are supported.

3.2.3 Area Object

The area object describes the allocated space in the binary keyslots area and related metadata.

The *area* object contains these mandatory fields:

- **type** [string] the area type.
- **offset** [string-uint64] the offset from the device start to the beginning of the binary area (in bytes).
- **size** [string-uint64] the area size (in bytes).

Area type *raw* contains these additional fields:

- **encryption** [string] the area encryption algorithm, in dm-crypt notation (for example `aes-xts-plain64`).
- **key_size** [integer] the area encryption key size.

Area type *none* and *journal* (used only for reencryption optional extension) contain only mandatory fields.

Area type *checksum* (used only for reencryption optional extension) contains these additional fields:

- **hash** [string] The hash algorithm for the checksum resilience mode.
- **sector_size** [integer] The data unit size for digest checksum calculated with the **hash** algorithm.

Area type *datashift* and *datashift-journal* (used only for reencryption optional extension) contains this additional field:

- **shift_size** [string-uint64] The data shift (in bytes) performed during reencryption (shift direction is according to **direction** field).

Area type *datashift-checksum* (used only for reencryption optional extension) contains these additional fields:

- **hash** [string] The hash algorithm for the checksum resilience mode.
- **sector_size** [integer] The data unit size for digest checksum calculated with the **hash** algorithm.
- **shift_size** [string-uint64] The data shift (in bytes) performed during reencryption (shift direction is according to **direction** field).

3.2.4 Anti-Forensic Object

The LUKS1 AF splitter is no longer much effective on modern storage devices. The functionality is here mainly for compatibility reasons. In future, it will be probably replaced.

The *af* (anti-forensic splitter) object contains this madatory field:

- **type** [string] the anti-forensic function type.

AF type *luks1* (compatible with LUKS1 [1]) contains these additional fields:

- **stripes** [integer] the number of stripes, for historical reasons only the 4000 value is supported.
- **hash** [string] the hash algorithm used.

3.2.5 Key Derivation Object

The object describes PBKDF attributes used for the keyslot.

The *kdf* object mandatory fields are:

- **type** [string] the PBKDF type.
- **salt** [base64] the salt for PBKDF (binary data).

The *pbkdf2*[7] type (compatible with LUKS1) contains these additional fields:

- **hash** [string] the hash algorithm for the PBKDF2 (SHA-256).
- **iterations** [integer] the PBKDF2 iterations count.

The *argon2i* and *argon2id*[8] type contains these additional fields:

- **time** [integer] the time cost (in fact the iterations count for Argon2).
- **memory** [integer] the memory cost, in kilobytes. If not available, the keyslot cannot be unlocked.
- **cpus** [integer] the required number of threads (CPU cores number cost). If not available, unlocking will be slower.

3.3 Segments Object

Segments object contains a definition of areas on the disk containing user data (in LUKS1 mentioned as the user data payload). For a normal LUKS device, there is only one data segment present.

During the data reencryption, the data area is internally divided according to the new and the old key, but only one abstracted area should be presented

[7] PBKDF2 contains the time cost (iterations) that describes how many times PBKDF2 must iterate to derive the candidate key.

[8] Argon2 algorithms, here used as PBKDF, are memory-hard [8] and have three costs: time, memory required and number of threads (CPUs).

to the user. Multiple segments (sorted by id) must cover the whole data area without any gaps and must not overlap.

Optional hardware-encrypted segment types are described in Chapter 4.10.

The *segment* object contains these mandatory fields:

- **type** [string] the segment type.
- **offset** [string-uint64] the offset from the device start to the beginning of the segment (in bytes).
- **size** [string or string-uint64] the segment size (in bytes) or *dynamic* if the size of the underlying device should be used (dynamic resize).
- **flags** [array,optional] the array of string objects marking segment with additional information.

Unknown **flags** are ignored. For the optional reencryption extension, the **flags** array can contain *in-reencryption* (reencryption in progress for this segment), *backup-final* (expected state after reencryption), *backup-previous* (state before reencryption started) and *backup-moved-segment* (moved segment if data shift is specified). Segment marked with backup flags must be the last (if sorted by the object name).

3.3.1 Segment Type *linear*

This object represents metadata about plaintext (unencrypted) data segment. It is used only during the online reencryption process.

The *linear* object contains only mandatory fields defined above.

3.3.2 Segment Type *crypt*

This object represents metadata about encrypted data segment.

The *crypt* object must contain these additional fields:

- **iv_tweak** [string-uint64] the starting offset for the Initialization Vector (IV tweak).
- **encryption** [string] the segment encryption algorithm, in the dm-crypt notation (for example `aes-xts-plain64`).
- **sector_size** [integer] the sector size for segment (512, 1024, 2048 or 4096 bytes).
- **integrity** [object,optional] the LUKS2 user data integrity protection type.

User data integrity protection is an experimental feature [9]) and requires *dm-integrity* (or hardware inline tags) and *dm-crypt* drivers with integrity support.

The *integrity* object contains these fields:

- **type** [string] the integrity type (in the dm-crypt notation, for example *aead* or *hmac(sha256)*).
- **journal_encryption** [string] the encryption type for the *dm-integrity* journal (not implemented yet, use *none*).
- **journal_integrity** [string] the integrity protection type for the *dm-integrity* journal (not implemented yet, use *none*).
- **key_size** [integer,optional] the key size (in bytes) for integrity type. If not set, the default key size is used. Use only for non-standard key sizes.

3.4 Digests Object

The digests object is used to verify that a key decrypted from a keyslot is correct. Digests are assigned to keyslots and segments. If it is not assigned to a segment, then it is a digest for an unbound key. Every keyslot must have one assigned digest object.

The *digest* object contains these fields:

- **type** [string] the digest type (only the *pbkdf2* type compatible with LUKS1 is used).
- **keyslots** [array] the array of keyslot objects names that are assigned to the digest.
- **segments** [array] the array of segment objects names that are assigned to the digest.
- **salt** [base64] the binary salt for the digest.
- **digest** [base64] the binary digest data.

The *pbkdf2* digest (similar to a kdf object in keyslot) contains these fields:

- **hash** [string] the hash algorithm for PBKDF2 (SHA-256).
- **iterations** [integer] the PBKDF2 iterations count.

3.5 Config Object

The *config* object contains attributes that are global for the LUKS device.

It contains these fields:

- **json_size** [string-uint64] the JSON area size (in bytes). Must match the binary header.
- **keyslots_size** [string-uint64] the binary keyslot area size (in bytes). Must be aligned to 4096 bytes.
- **flags** [array, optional] the array of string objects with persistent flags for the device.
- **requirements** [array, optional] the array of string objects with additional required features for the LUKS device.

The **flags** array can contain device activation flags (various performance and compatibility settings). Unknown **flags** are ignored.

The reference implementation uses these flags:

- **allow-discards** allows TRIM (discards) on the active device.
- **same-cpu-crypt** compatibility performance flag for dm-crypt [3] to perform encryption using the same CPU that originated the request.
- **submit-from-crypt-cpus** compatibility performance flag for dm-crypt [3] to disable offloading write requests to a separate thread after encryption.
- **no-journal** disable data journalling for dm-integrity [10].
- **no-read-workqueue** compatibility performance flag for dm-crypt [3] to bypass dm-crypt read workqueue and process read requests synchronously.
- **no-write-workqueue** compatibility performance flag for dm-crypt [3] to bypass dm-crypt write workqueue and process write requests synchronously.

The **requirements** array can contain an array of additional features that are mandatory when manipulating with a LUKS device and metadata or that are required for proper device activation. If an implementation detects a string that it does not recognize, it must treat the whole metadata as read-only and must

avoid device activation. These requirement flags are used for future extensions to mark the header to be not backward compatible.

For the reference implementation [5], the *offline-reencrypt* flag is used that marks a device during offline reencryption to prevent an activation until the offline reencryption is finished.

For online reencryption in progress, the *online-reencrypt-v2* or *online-reencrypt-v3* flag is used (future extensions can increase the version due to compatibility reasons).

If OPAL extention is used, the *opal* requirement flag is used.

If per-sector inline hardware tags are used for authenticated encryption, the *inline-hw-tags* requirement flag is used. Note that with this requirement set, only *dm-crypt* is used and all *dm-integrity* options in metadata are ignored.

3.6 Tokens Object

A token is an object that can describe *how to get a passphrase* to unlock a particular keyslot. It can also contain additional user-defined JSON metadata.

The mandatory fields for every token are:

- **type** [string] the token type (tokens with *luks2-* prefix are reserved for the implementation internal use).
- **keyslots** [array] the array of keyslot objects names that are assigned to the token.

The rest of the JSON content is the particular token implementation and can contain arbitrary JSON structured data (implementation should provide an interface to the JSON metadata directly).

For example, the reference implementation of *luks2-keyring* token allows automatic activation of the device if the passphrase is preloaded into a keyring with the specified ID.

The *luks2-keyring* token type contains these fields:

- **type** [string] is set to the *luks2-keyring*.
- **keyslots** [array] is assigned to the specific keyslot(s).
- **key_description** [string] contains the ID of the keyring entry with a passphrase.

4 LUKS2 Operations

Basic operations of a LUKS2 device are the same as specified in LUKS1 [1]. Header update operations must be synchronized due to the redundancy of metadata and operations must be serialized to prevent concurrent processes from updating the metadata at the same time. The metadata update must be implemented in such a way that at least of one header (primary or secondary) is always valid to allow for a proper recovery in the case of a failure. These steps require an implementation of some high-level locking of metadata access.

4.1 Device Formatting

Initialization (formatting) of a LUKS2 device starts with generating basic metadata parameters, like *UUID* and writing both binary headers and basic metadata

structures. The JSON area must always contain valid LUKS2 top-level objects. The config object must be initialized to include proper area size parameters that match the binary header. If the LUKS2 header references a user data segment, that segment must be initialized with all mandatory parameters. The keyslots, digests and tokens can be empty in this step.

4.2 Keyslot Initialization

The next step is allocation of new keyslot and metadata and assignment to the key digests and segments. The key stored in the keyslot and salt for the keyslot should be generated using a cryptographically secure RNG.

The PBKDF cost parameters (iterations, memory, CPU cores) that are used to derive the keyslot unlocking key from a user passphrase must be either specified by the user, or it can be benchmarked according to user needs. See the reference cryptsetup implementation [5] as an example of this approach.

Once the key is generated, a new key digest is created, and a new keyslot object is allocated and assigned to the digest (and segment). The new keyslot contains the binary keyslot area allocated according to the stored key size.

The size of the binary allocated area is determined according to the key size and the anti-forensic (AF) splitter output (see section 2.4 in LUKS1 [1]).

LUKS2 keyslots can store different keys with different key sizes. The allocation of binary keyslot data depends on the order of creation. Keyslot positions are no longer fixed as in LUKS1.

The last step of keyslot initialization writes the encrypted key to the allocated binary keyslot area. A user passphrase and a salt are processed by the configured PBKDF. The PBKDF output key is used for the keyslot binary area encryption algorithm. The key is split using AF splitter and encrypted by the keyslot encryption algorithm.

4.3 Keyslot Content Retrieval

The user provided passphrase with the salt and parameters from the header metadata are processed through the PBKDF. The derived key is used to decrypt the binary keyslot area. The decrypted content is processed (merged) in the AF splitter. The assigned key digest is calculated with the recovered candidate key. If the calculated digest and the digest in metadata match, the recovered key is valid. If the digest does not match, the provided passphrase must be rejected.

4.4 Keyslot Revocation

To discard a keyslot, the binary area for the keyslot must be physically overwritten (to discard the stored data). After this step, the keyslot metadata object must be removed with all bindings to digests and segments.

Note that the key digest and its binding to the segment can remain in metadata (not assigned to any keyslots). If a user has the copy of the encryption key, the validity of the key can still be verified with this digest and the device can be later still activated.

4.5 Metadata Recovery

The replicated metadata allows for a full LUKS2 header recovery (except binary keyslot areas) if some part of headers become corrupted. A part of the recovery can be automated, but because this process can revert some intentional changes, a user interaction is suggested.

The automatic recovery should always update both copies to the more recent version (with higher *seqid*). The metadata handler should first try to load the primary header, then the secondary header. If one of the headers is more recent, the older header is updated. If the primary header is corrupted, a scan on several known offsets for the secondary header can be performed.

4.6 Mandatory Requirements

While the LUKS2 format is algorithm-agnostic, some algorithm implementations are crucial for internal function.

The cryptographic backend for LUKS2 must support these algorithms:

- **SHA-1** hash algorithm (for compatibility with old LUKS1 devices).
- **SHA-256** hash algorithm, used as the default checksum for the binary header and in the PBKDF2 digest.
- **PBKDF2** password-based key derivation (for digest and backward compatibility with LUKS1).
- **Argon2i and Argon2id** memory-hard key derivation functions for new LUKS2 keyslots.
- **AES-XTS** symmetric cipher for the default keyslot encryption and the default user data encryption.

4.7 Conversion from LUKS1

If an existing LUKS1 device header contains enough space for the LUKS2 metadata, then it can be converted in-place to the LUKS2 format. Reference implementation provides the `cryptsetup convert --type luks2` command.

LUKS1 [name]	**128-bit key** [sectors]	**256-bit key** [sectors]	**512-bit key** [sectors]
Header	0	0	0
Keyslot 0	8	8	8
Keyslot 1	136	264	512
Keyslot 2	264	520	1016
Keyslot 3	392	776	1520
Keyslot 4	520	1032	2024
Keyslot 5	648	1288	2528
Keyslot 6	776	1544	3032
Keyslot 7	904	1800	3536
Padding	1032	2056	4040
Data offset	2048	4096	4096
Unused sectors	1016	2040	56

Table 2: Offsets (in 512-byte sectors) of common LUKS1 headers.

For reference, Table 2 contains offsets of LUKS1 keyslots that can be converted to LUKS2 in-place. The resulting LUKS2 header has 12kB JSON area in all these cases. Note that the binary keyslot area is directly copied to the proper position, there is no recovery possible if the convert operation fails. Schema of area locations during conversion is illustrated in Figure 4.

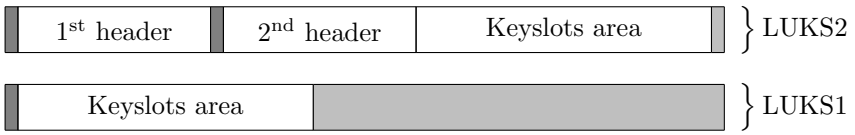

Figure 4: LUKS1/LUKS2 areas placement.

If a LUKS2 header uses compatible options with LUKS1 (PBKDF2, no integrity protection, no tokens, no unbound keys) then it can also be converted back to the LUKS1 header in-place with the `cryptsetup convert --type luks1` command.

4.8 Algorithm Definition Examples

The LUKS2 specification supports all algorithms that are provided by the cryptographic backend (in the Linux case by kernel dm-crypt and userspace cryptographic library). Figures 3 and 4 list few examples of symmetric ciphers for data encryption and PBKDF algorithms.

Algorithm in LUKS2 notation	**Description**
pbkdf2	PBKDF2 with HMAC-SHA256 [11]
argon2i	Argon2i as PBKDF (data independent) [8]
argon2id	Argon2id as PBKDF (combined mode) [8]

Table 3: LUKS2 PBKDF algorithms.

Algorithm in dm-crypt [3] notation	**Description**
aes-xts-plain64	AES in XTS mode with sequential IV [12, 13]
aes-cbc:essiv:sha256	AES in CBC mode with ESSIV IV [3, 12]
serpent-xts-plain64	Serpent cipher with sequential IV [14]
twofish-xts-plain64	Twofish cipher with sequential IV [15]
aegis128-random	AEGIS (128-bit) with random IV, AEAD [16]

Table 4: LUKS2 encryption algorithms examples.

AEAD algorithms[9] are experimental and require dm-integrity [10] (or hardware inline tags) support.

[9] AEGIS with 256-bit key and MORUS algorithms are no longer supported; these algorithms were removed from Linux kernel.

4.9 Online Reencryption (Optional Extension)

The online reencryption is a process of encryption key change while the data device is available for use during the whole process. The reencryption process is initiated by the user and runs in the background. During the reencryption LUKS2 metadata are continuously updated to cover old and new encryption keys and related data segments. Reencryption can run forward or backward. Special cases are encryption of not-yet-encrypted devices (encrypting device with plaintext) and decryption (removing the encryption layer). The LUKS2 online reencryption is an optional extension of the format; metadata before and after reencryption are fully compatible with the basic LUKS2 format.

This section describes on-disk metadata as used in the reference implementation in cryptsetup [5] project. Implementation is based on the Linux kernel device-mapper subsystem that allows mapping table reconfigurations during normal operations [17]. If there is a reencryption in progress, JSON metadata must contain proper requirement flag in *config* JSON object to prevent metadata manipulations on systems where reencryption extension is not implemented.

Because reencryption uses blocks (typically of megabytes in size) that cannot be written atomically, reencryption provides a resilience mechanism that triggers data recovery if reencryption is unexpectedly interrupted (power failure or a similar event). The data recovery requires additional storage for data. This storage is temporarily allocated in the LUKS2 header binary keyslot area. The resilience type then specifies how data recovery is performed. Schema of reencryption areas is illustrated in Figure 5.

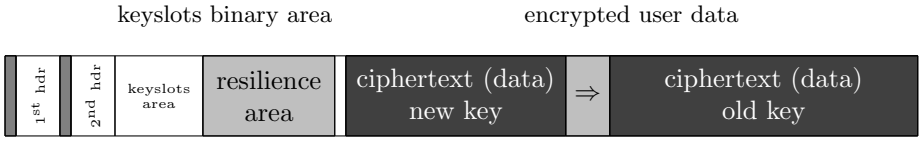

Figure 5: LUKS2 reencryption schema (in forward direction).

Resilience area type (data recovery type) for online reencryption can be:

- **none** No data recovery at all. Metadata are updated only when the reencryption is finished or after a correct interruption (TERM signal). In the case of a crash, no recovery is possible. The allocated resilience area is not used and is only reserved.
- **checksum** The resilience area contains array of checksums of **sector_size** blocks on reencrypted area, calculated with specified **hash** algorithm. The recovery code then compares checksums, and if the checksum is correct (the block was not yet reencrypted), the block is reencrypted. The **sector_size** must be an atomic unit for the underlying storage (it contains either old or new data).
- **journal** The resilience area contains the old copy of the data for the whole reencryption segment. In the case of a crash, these data are reencrypted again to the original position.
- **datashift** A special mode when ciphertext is moved. Reencryption then writes data to non-overlapping blocks, so no data resilience is needed.

- **datashift-checksum** A special mode when plaintext is produced during decryption from encrypted device. This mode combines *checksum* with *datashift* area resilience (checksum is utilized only for the last moved segment).
- **datashift-journal** A special mode when plaintext is produced during decryption from encrypted device. This mode combines *journal* with *datashift* area resilience (journal is utilized only for the last moved segment).

4.9.1 Reencryption Steps

The running reencryption can be described as a sequence of these repeated steps:

1. **Prepare JSON metadata** (segment *in-reencryption* flag is set).
2. **Configure overlay device-mapper devices**.
3. **Store resilience data** to the reserved resilience area.
4. **Write new block** to the final storage media location.
5. **Update JSON metadata** (segment *in-reencryption* flag is removed).
6. **Move segment offset.**
7. **Repeat** until the final block.
8. **Remove reencryption JSON metadata** and wipe the resilience (and possible unused datashift) area.

4.9.2 Reencryption Keyslot Protection

Online reencryption is a very complex operation, and with metadata visible on the device (see Section 1.2) and with the physical access to the device, it creates a new attack vector. The whole reencryption operation should be taken as a temporary state. A user should not start reencryption and keep partial metadata in place without finishing reencryption. During reencryption, keyslot manipulation operations are not available.

To protect intended operation (*encrypt*, *decrypt* or *reencrypt*), the *reencrypt* keyslot must be linked to the additional *digest* object. The digest is calculated from available encryption keys and critical reencryption metadata (*reencrypt* keyslot fields and backup segments). This digest must be validated on encrypted device activation and before all manipulation with metadata (for example, restarting reencryption after a crash).

The schema of metadata used for reencryption digest validation is detailed in Figure 6. Version prefix is always two bytes copied from reencryption flag ('v2', ASCII bytes 0x76, 0x32 or 'v3', ASCII bytes 0x76, 0x33, see Section 3.5). All integers are presented as unsigned and converted to big-endian byte order. String content is taken from JSON string objects defined in Section 3. Strings are processed without trailing null characters.

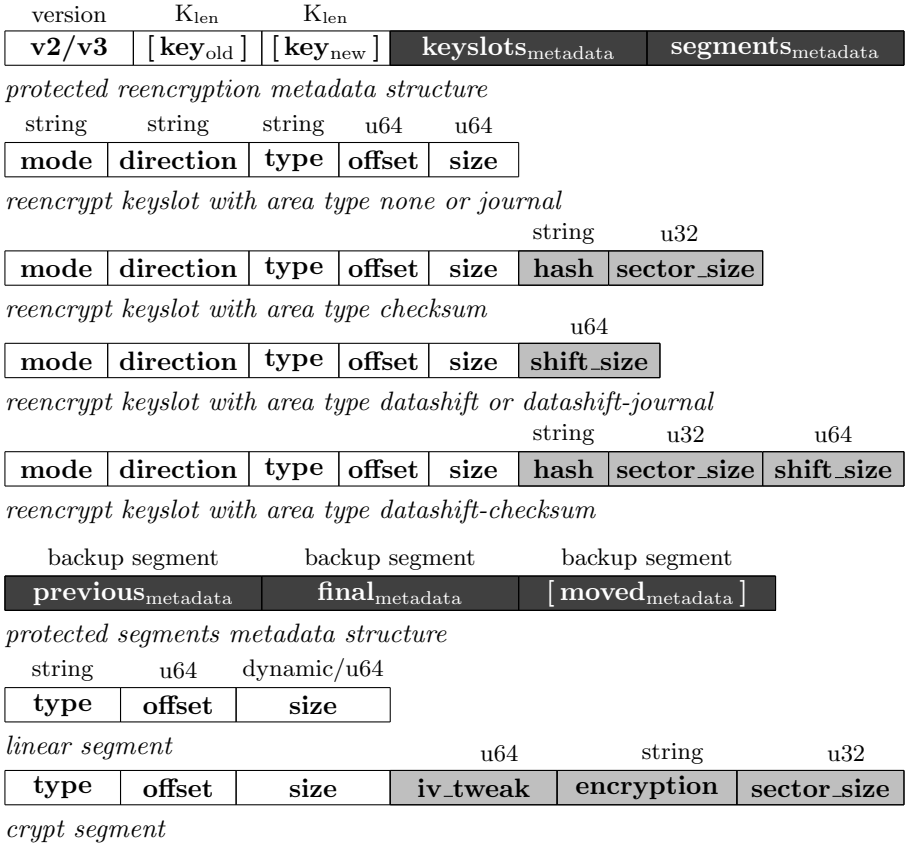

Figure 6: Reencryption digest calculation (binary input metadata).

Note that for *encryption* and *decryption* mode, the whole device must be treated as **unencrypted** – there are no quarantees of confidentiality as part of the device contains plaintext.

4.10 OPAL Hardware Encryption (Optional Extension)

OPAL extension allows the use of hardware encryption provided by OPAL2 subsystem [18] for self-encrypting drives (SED). LUKS2 uses a specific OPAL locking range, not the full device encryption, as the LUKS2 metadata must remain visible.

New segment type *hw-opal* defines hardware-only encryption, and segment type *hw-opal-crypt* defines combined encryption (software encryption above hardware). Additional stored fields are used directly in communication with the underlying OPAL device. If these segment types are used, the requirement flag *opal* must be set in the config section (see Chapter 3.5).

You need an administrator password for the OPAL device for the OPAL locking range configuration. This password is not stored in LUKS2 metadata. Configured OPAL locking range is then unlocked by the independent key stored in the LUKS2 keyslot.

4.10.1 Security Notes to OPAL Encryption

Hardware encryption changes the LUKS2 security model, as you have to trust your storage hardware. It can be combined with software encryption as an additional encryption layer or in hardware-only encryption. If OPAL encryption is used, the LUKS2 metadata is bound to the specific hardware device.

4.10.2 Segment Type *hw-opal*

For mandatory segment type fields see Chapter 3.3. The *hw-opal* object represents metadata for hardware-only encrypted data segment and must contain these additional fields:

- **opal_segment_number** [integer] the OPAL locking range identifier.
- **opal_key_size** [integer] the size of unlocking key (in bytes) for OPAL locking range. Usually set to 32 bytes for compatibility reasons. This key is stored in linked keyslots.
- **opal_segment_size** [string-uint64] the segment size (in bytes) of the underlying OPAL locking range. In normal configuration, it should be the same value as **size** field.

4.10.3 Segment Type *hw-opal-crypt*

The segment type *hw-opal-crypt* is the combination of *crypt* and *hw-opal* segment types. It describes combined software and hardware encryption (*dm-crypt* encryption stacked over OPAL hardware encrypted locking range).

The *hw-opal-crypt* object must contain all mandatory fields from *crypt* and *hw-opal* segment type. The linked keyslot object then stores the combined key where the first part is the key for OPAL locking range (of size defined by segment field **opal_key_size**) and the second part is the key for *dm-crypt*. These keys should be independent.

Glossary

AEAD Authenticated Encryption with Additional Data.

AF Anti-Forensic splitter defined for LUKS1. [1, 19]

Base64 Binary to text encoding scheme. [7]

blkid Utility to locate and print block device attributes.

dm-crypt Linux device-mapper crypto target. [3]

dm-integrity Linux device-mapper integrity target [10].

IV Initialization Vector for an encryption mode that tweaks encryption.

JSON JavaScript Object Notation (data-interchange format). [6]

Keyslot Encrypted area on disk that contains a key.

Length-preserving encryption Symmetric encryption where plaintext and ciphertext have the same size.

libcryptsetup Library implementing LUKS1 and LUKS2. [5]

Metadata locking A way how to serialize access to on-disk metadata updates.

OPAL Subsystem used for data encryption in self-encrypting drives. [18]

OPAL locking range Encrypted area on OPAL disk that can be independently locked.

PBKDF Password-Based Key Derivation Function.

RNG Cryptographically strong Random Number Generator.

Sector Atomic unit for block device (disk). Typical sector size is 4096 bytes.

TRIM Command that informs a block device that area of the disk is unused and can be discarded.

udev Device manager for Linux kernel implemented in userspace.

UUID Universally Unique IDentifier (of a block device).

Volume Key The key used for data encryption on disk. Sometimes called as Media Encryption Key (MEK).

References

[1] LUKS1 On-Disk Format Specification, Version 1.2.3, 2018. https://gitlab.com/cryptsetup/cryptsetup/wikis/Specification.

[2] Clemens Fruhwirth. *New methods in hard disk encryption.* PhD thesis, Institute for Computer Languages Theory and Logic Group, Vienna University of Technology, 2005. http://clemens.endorphin.org/nmihde/nmihde-A4-os.pdf.

[3] dm-crypt: Linux device-mapper crypto target, 2022. https://gitlab.com/cryptsetup/cryptsetup/wikis/DMCrypt.

[4] Arno Wagner and Milan Brož. Frequently Asked Questions Cryptsetup/LUKS, 2022. https://gitlab.com/cryptsetup/cryptsetup/wikis/FrequentlyAskedQuestions.

[5] Cryptsetup and LUKS, 2022. https://gitlab.com/cryptsetup/cryptsetup.

[6] The JSON Data Interchange Format. Technical Report Standard ECMA-404, 1st edition, ECMA, 2013. http://www.ecma-international.org/publications/files/ECMA-ST/ECMA-404.pdf.

[7] Simon Josefsson. The Base16, Base32, and Base64 Data Encodings. RFC 4648, 2006. https://www.ietf.org/rfc/rfc4648.txt.

[8] Alex Biryukov, Daniel Dinu, and Dmitry Khovratovich. Argon2: the memory-hard function for password hashing and other applications, 2017. https://www.cryptolux.org/index.php/Argon2.

[9] Milan Brož, Mikuláš Patočka, and Vashek Matyáš. Practical Cryptographic Data Integrity Protection with Full Disk Encryption Extended Version, 2018. https://arxiv.org/abs/1807.00309.

[10] dm-integrity: Linux device-mapper integrity target, 2022. https://gitlab.com/cryptsetup/cryptsetup/wikis/DMIntegrity.

[11] Burt Kaliski. PKCS #5: Password-Based Cryptography Specification Version 2.0. RFC 2898 (Informational), 2000. https://www.ietf.org/rfc/rfc2898.txt.

[12] FIPS Publication 197, The Advanced Encryption Standard (AES), 2001. https://nvlpubs.nist.gov/nistpubs/fips/nist.fips.197.pdf.

[13] Morris J. Dworkin. SP 800-38E. Recommendation for Block Cipher Modes of Operation: The XTS-AES Mode for Confidentiality on Storage Devices, 2010. NIST, https://nvlpubs.nist.gov/nistpubs/legacy/sp/nistspecialpublication800-38e.pdf.

[14] Ross Anderson, Eli Biham, and Lars Knudsen. Serpent: A Proposal for the Advanced Encryption Standard. https://www.cl.cam.ac.uk/~rja14/serpent.html.

[15] Bruce Schneier, John Kelsey, Doug Whiting, David Wagner, Chris Hall, and Niels Ferguson. *The Twofish Encryption Algorithm: A 128-bit Block Cipher.* John Wiley & Sons, Inc., 1999. https://www.schneier.com/academic/twofish/.

[16] Hongjun Wu and Bart Preneel. AEGIS, A Fast Authenticated Encryption Algorithm (v1.1). Technical report, 2016. https://competitions.cr.yp.to/round3/aegisv11.pdf.

[17] Ondřej Kozina. Online disk reencryption with LUKS2, 2019. https://okozina.fedorapeople.org/online-disk-reencryption-with-luks2.pdf.

[18] TCG Storage Security Subsystem Class: Opal. Technical Report version 2.02, Trusted Computing Group, 2022. https://trustedcomputinggroup.org/resource/storage-work-group-storage-security-subsystem-class-opal.

[19] Clemens Fruhwirth. TKS1, An anti-forensic, two level, and iterated key setup scheme, 2004. https://www.kernel.org/pub/linux/utils/cryptsetup/LUKS_docs/TKS1-draft.pdf.

Document History

Version	Date	Author
1.0.0	2018-08-02	Milan Broz <gmazyland@gmail.com>
Initial revision for LUKS2. *Not a final version, work in progress.*		
1.0.0	2018-10-23	Milan Broz <gmazyland@gmail.com>
Added segment flags.		
1.1.0	2022-01-10	Milan Broz <gmazyland@gmail.com>
Added reencryption, policy. Updates.		
1.1.1	2022-07-12	Milan Broz <gmazyland@gmail.com>
Added datashift-checksum/journal reencryption resilience area type.		
1.1.2	2023-12-17	Milan Broz <gmazyland@gmail.com>
Added OPAL hardware encryption segment extension.		
1.1.3	2024-02-17	Tobias Rosenkranz
Corrected error in reencryption resilience area type.		
1.1.4	2025-06-16	Milan Broz <gmazyland@gmail.com>
Added inline hw tags extension.		
Added optional integrity key size for crypt segment.		