OpenPGP Message Format
draft-ietf-openpgp-crypto-refresh-05

Abstract
This document specifies the message formats used in OpenPGP. OpenPGP
provides encryption with public-key or symmetric cryptographic
algorithms, digital signatures, compression and key management.

This document is maintained in order to publish all necessary
information needed to develop interoperable applications based on the
OpenPGP format. It is not a step-by-step cookbook for writing an
application. It describes only the format and methods needed to
read, check, generate, and write conforming packets crossing any
network. It does not deal with storage and implementation questions.
It does, however, discuss implementation issues necessary to avoid
security flaws.

This document obsoletes: RFC 4880 (OpenPGP), RFC 5581 (Camellia in
OpenPGP) and RFC 6637 (Elliptic Curves in OpenPGP).

Status of This Memo
This Internet-Draft is submitted in full conformance with the
provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering
Task Force (IETF). Note that other groups may also distribute
working documents as Internet-Drafts. The list of current Internet-
Drafts is at https://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months
and may be updated, replaced, or obsoleted by other documents at any
time. It is inappropriate to use Internet-Drafts as reference
material or to cite them other than as "work in progress."

This Internet-Draft will expire on 8 September 2022.
Copyright Notice

Copyright (c) 2022 IETF Trust and the persons identified as the
document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust’s Legal
license-info) in effect on the date of publication of this document.
Please review these documents carefully, as they describe your rights
and restrictions with respect to this document. Code Components
extracted from this document must include Revised BSD License text as
described in Section 4.e of the Trust Legal Provisions and are
provided without warranty as described in the Revised BSD License.

Table of Contents

1. Introduction .................................................. 7
  1.1. Terms ..................................................... 7
2. General functions ............................................. 8
  2.1. Confidentiality via Encryption .............................. 8
  2.2. Authentication via Digital Signature ....................... 9
  2.3. Compression .............................................. 10
  2.4. Conversion to Radix-64 .................................. 10
  2.5. Signature-Only Applications ............................... 10
3. Data Element Formats ......................................... 10
  3.1. Scalar Numbers .......................................... 10
  3.2. Multiprecision Integers ................................ 11
     3.2.1. Using MPIs to encode other data ..................... 11
  3.3. Key IDs .................................................. 11
  3.4. Text ..................................................... 12
  3.5. Time Fields .............................................. 12
  3.6. Keyrings ............................................... 12
  3.7. String-to-Key (S2K) Specifiers ............................ 12
     3.7.1. String-to-Key (S2K)Specifier Types ................ 12
        3.7.1.1. Simple S2K ..................................... 13
        3.7.1.2. Salted S2K .................................... 14
        3.7.1.3. Iterated and Salted S2K ....................... 14
        3.7.1.4. Argon2 ........................................ 15
     3.7.2. String-to-Key Usage ................................ 16
        3.7.2.1. Secret-Key Encryption .......................... 16
        3.7.2.2. Symmetric-Key Message Encryption ............ 17
4. Packet Syntax ................................................. 18
  4.1. Overview ................................................. 18
  4.2. Packet Headers .......................................... 18
     4.2.1. OpenPGP Format Packet Lengths .................... 19
        4.2.1.1. One-Octet Lengths ............................. 20
        4.2.1.2. Two-Octet Lengths ............................ 20
        4.2.1.3. Five-Octet Lengths ........................... 20
5.2.3.32. Issuer Fingerprint .......................... 50
5.2.3.33. Intended Recipient Fingerprint ............... 50
5.2.4. Computing Signatures ............................ 51
5.2.4.1. Subpacket Hints ............................... 52
5.3. Symmetric-Key Encrypted Session Key Packets (Tag 3) ... 53
5.3.1. v4 SKESK ........................................ 53
5.3.2. v5 SKESK ........................................ 54
5.4. One-Pass Signature Packets (Tag 4) ................ 55
5.5. Key Material Packet ................................ 56
5.5.1. Key Packet Variants .............................. 56
5.5.1.1. Public-Key Packet (Tag 6) ..................... 56
5.5.1.2. Public-Subkey Packet (Tag 14) ................ 57
5.5.1.3. Secret-Key Packet (Tag 5) .................... 57
5.5.1.4. Secret-Subkey Packet (Tag 7) ................ 57
5.5.2. Public-Key Packet Formats ........................ 57
5.5.3. Secret-Key Packet Formats ........................ 59
5.6. Algorithm-specific Parts of Keys .................... 61
5.6.1. Algorithm-Specific Part for RSA Keys ............. 62
5.6.2. Algorithm-Specific Part for DSA Keys ............. 62
5.6.3. Algorithm-Specific Part for Elgamal Keys ......... 62
5.6.4. Algorithm-Specific Part for ECDSA Keys .......... 63
5.6.5. Algorithm-Specific Part for EdDSA Keys .......... 63
5.6.6. Algorithm-Specific Part for ECDH Keys .......... 63
5.6.6.1. ECDH Secret Key Material .................... 64
5.7. Compressed Data Packet (Tag 8) ..................... 65
5.8. Symmetrically Encrypted Data Packet (Tag 9) ......... 66
5.9. Marker Packet (Tag 10) ............................ 67
5.10. Literal Data Packet (Tag 11) ....................... 67
5.10.1. Special Filename _CONSOLE (Deprecated) .......... 69
5.11. Trust Packet (Tag 12) ................................ 69
5.12. User ID Packet (Tag 13) ........................... 70
5.13. User Attribute Packet (Tag 17) .................... 70
5.13.1. The Image Attribute Subpacket .................. 71
5.14. Sym. Encrypted Integrity Protected Data Packet (Tag 18) ........................................ 71
5.14.3. EAX Mode ....................................... 76
5.14.4. OCB Mode ....................................... 76
5.14.5. GCM Mode ....................................... 76
5.15. Padding Packet (Tag 21) ........................... 76
6. Radix-64 Conversions .................................. 77
6.1. An Implementation of the CRC-24 in "C" ............... 78
6.2. Forming ASCII Armor ................................ 78
6.3. Encoding Binary in Radix-64 ........................ 81
6.4. Decoding Radix-64 .................................. 83
13.2. EC Point Wire Formats .......................... 109
13.2.1. SEC1 EC Point Wire Format .................. 109
13.2.2. Prefixed Native EC Point Wire Format ....... 110
13.2.3. Notes on EC Point Wire Formats ............. 110
13.3. EC Scalar Wire Formats ........................ 110
13.3.1. EC Octet String Wire Format ................ 111
13.3.2. Elliptic Curve Prefixed Octet String Wire Format .... 112
13.4. Key Derivation Function ........................ 112
13.5. EC DH Algorithm (ECDH) ......................... 113
14.1. PKCS#1 Encoding in OpenPGP .................... 116
14.1.1. EME-PKCS1-v1_5-ENCODE ..................... 116
14.1.2. EME-PKCS1-v1_5-DECODE ...................... 117
14.1.3. EMSA-PKCS1-v1_5 ............................ 118
14.2. Symmetric Algorithm Preferences ............... 119
14.2.1. Plaintext .................................... 119
14.3. Other Algorithm Preferences ................... 120
14.3.1. Compression Preferences ..................... 120
14.3.1.1. Uncompressed ............................... 120
14.3.2. Hash Algorithm Preferences ................. 120
14.4. RSA ............................................ 121
14.5. DSA ............................................ 121
14.6. Elgamal ....................................... 121
14.7. EdDSA ......................................... 122
14.8. Reserved Algorithm Numbers .................... 122
14.9. OpenPGP CFB Mode ................................ 122
14.10. Private or Experimental Parameters .......... 124
14.11. Meta-Considerations for Expansion .......... 124
15. Security Considerations .......................... 124
15.1. Avoiding Ciphertext Malleability ............... 128
15.2. Escrowed Revocation Signatures ................. 130
15.3. Random Number Generation and Seeding .......... 131
15.4. Traffic Analysis ................................ 131
16. Implementation Nits .............................. 132
17. References ....................................... 133
17.1. Normative References ........................... 133
17.2. Informative References .......................... 136
Appendix A. Test vectors ............................. 138
A.1. Sample EdDSA key ................................ 138
A.2. Sample EdDSA signature .......................... 138
A.3. Sample AEAD-EAX encryption and decryption ....... 139
A.3.1. Sample Parameters .............................. 139
A.3.2. Sample symmetric-key encrypted session key packet (v5) ......................... 139
A.3.3. Starting AEAD-EAX decryption of the session key ..... 140
A.3.4. Sample v2 SEIPD packet ........................ 140
A.3.5. Decryption of data .............................. 141
A.3.6. Complete AEAD-EAX encrypted packet sequence .... 142
1. Introduction

This document provides information on the message-exchange packet formats used by OpenPGP to provide encryption, decryption, signing, and key management functions. It is a revision of RFC 4880, "OpenPGP Message Format", which is a revision of RFC 2440, which itself replaces RFC 1991, "PGP Message Exchange Formats" [RFC1991] [RFC2440] [RFC4880].

This document obsoletes: RFC 4880 (OpenPGP), RFC 5581 (Camellia in OpenPGP) and RFC 6637 (Elliptic Curves in OpenPGP).

1.1. Terms

* OpenPGP - This is a term for security software that uses PGP 5 as a basis, formalized in this document.

* PGP - Pretty Good Privacy. PGP is a family of software systems developed by Philip R. Zimmermann from which OpenPGP is based.

* PGP 2 - This version of PGP has many variants; where necessary a more detailed version number is used here. PGP 2 uses only RSA, MD5, and IDEA for its cryptographic transforms. An informational RFC, RFC 1991, was written describing this version of PGP.
* PGP 5 - This version of PGP is formerly known as "PGP 3" in the community. It has new formats and corrects a number of problems in the PGP 2 design. It is referred to here as PGP 5 because that software was the first release of the "PGP 3" code base.

* GnuPG - GNU Privacy Guard, also called GPG. GnuPG is an OpenPGP implementation that avoids all encumbered algorithms. Consequently, early versions of GnuPG did not include RSA public keys.

"PGP", "Pretty Good", and "Pretty Good Privacy" are trademarks of PGP Corporation and are used with permission. The term "OpenPGP" refers to the protocol described in this and related documents.

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

The key words "PRIVATE USE", "SPECIFICATION REQUIRED", and "RFC REQUIRED" that appear in this document when used to describe namespace allocation are to be interpreted as described in [RFC8126].

2. General functions

OpenPGP provides data integrity services for messages and data files by using these core technologies:

* digital signatures
* encryption
* compression
* Radix-64 conversion

In addition, OpenPGP provides key management and certificate services, but many of these are beyond the scope of this document.

2.1. Confidentiality via Encryption

OpenPGP combines symmetric-key encryption and public-key encryption to provide confidentiality. When made confidential, first the object is encrypted using a symmetric encryption algorithm. Each symmetric key is used only once, for a single object. A new "session key" is generated as a random number for each object (sometimes referred to as a session). Since it is used only once, the session key is bound
to the message and transmitted with it. To protect the key, it is encrypted with the receiver’s public key. The sequence is as follows:

1. The sender creates a message.
2. The sending OpenPGP generates a random number to be used as a session key for this message only.
3. The session key is encrypted using each recipient’s public key. These "encrypted session keys" start the message.
4. The sending OpenPGP encrypts the message using the session key, which forms the remainder of the message.
5. The receiving OpenPGP decrypts the session key using the recipient’s private key.
6. The receiving OpenPGP decrypts the message using the session key. If the message was compressed, it will be decompressed.

With symmetric-key encryption, an object may be encrypted with a symmetric key derived from a passphrase (or other shared secret), or a two-stage mechanism similar to the public-key method described above in which a session key is itself encrypted with a symmetric algorithm keyed from a shared secret.

Both digital signature and confidentiality services may be applied to the same message. First, a signature is generated for the message and attached to the message. Then the message plus signature is encrypted using a symmetric session key. Finally, the session key is encrypted using public-key encryption and prefixed to the encrypted block.

2.2. Authentication via Digital Signature

The digital signature uses a hash code or message digest algorithm, and a public-key signature algorithm. The sequence is as follows:

1. The sender creates a message.
2. The sending software generates a hash code of the message.
3. The sending software generates a signature from the hash code using the sender’s private key.
4. The binary signature is attached to the message.
5. The receiving software keeps a copy of the message signature.

6. The receiving software generates a new hash code for the received message and verifies it using the message’s signature. If the verification is successful, the message is accepted as authentic.

2.3. Compression

If an implementation does not implement compression, its authors should be aware that most OpenPGP messages in the world are compressed. Thus, it may even be wise for a space-constrained implementation to implement decompression, but not compression.

2.4. Conversion to Radix-64

OpenPGP’s underlying native representation for encrypted messages, signature certificates, and keys is a stream of arbitrary octets. Some systems only permit the use of blocks consisting of seven-bit, printable text. For transporting OpenPGP’s native raw binary octets through channels that are not safe to raw binary data, a printable encoding of these binary octets is needed. OpenPGP provides the service of converting the raw 8-bit binary octet stream to a stream of printable ASCII characters, called Radix-64 encoding or ASCII Armor.

Implementations SHOULD provide Radix-64 conversions.

2.5. Signature-Only Applications

OpenPGP is designed for applications that use both encryption and signatures, but there are a number of problems that are solved by a signature-only implementation. Although this specification requires both encryption and signatures, it is reasonable for there to be subset implementations that are non-conformant only in that they omit encryption.

3. Data Element Formats

This section describes the data elements used by OpenPGP.

3.1. Scalar Numbers

Scalar numbers are unsigned and are always stored in big-endian format. Using n[k] to refer to the kth octet being interpreted, the value of a two-octet scalar is ((n[0] << 8) + n[1]). The value of a four-octet scalar is ((n[0] << 24) + (n[1] << 16) + (n[2] << 8) + n[3]).
3.2. Multiprecision Integers

Multiprecision integers (also called MPIs) are unsigned integers used to hold large integers such as the ones used in cryptographic calculations.

An MPI consists of two pieces: a two-octet scalar that is the length of the MPI in bits followed by a string of octets that contain the actual integer.

These octets form a big-endian number; a big-endian number can be made into an MPI by prefixing it with the appropriate length.

Examples:

(all numbers are in hexadecimal)

The string of octets [00 01 01] forms an MPI with the value 1. The string [00 09 01 FF] forms an MPI with the value of 511.

Additional rules:

The size of an MPI is ((MPI.length + 7) / 8) + 2 octets.

The length field of an MPI describes the length starting from its most significant non-zero bit. Thus, the MPI [00 02 01] is not formed correctly. It should be [00 01 01].

Unused bits of an MPI MUST be zero.

Also note that when an MPI is encrypted, the length refers to the plaintext MPI. It may be ill-formed in its ciphertext.

3.2.1. Using MPIs to encode other data

Note that MPIs are used in some places used to encode non-integer data, such as an elliptic curve point (see Section 13.2, or an octet string of known, fixed length (see Section 13.3). The wire representation is the same: two octets of length in bits counted from the first non-zero bit, followed by the smallest series of octets that can represent the value while stripping off any leading zero octets.

3.3. Key IDs

A Key ID is an eight-octet scalar that identifies a key. Implementations SHOULD NOT assume that Key IDs are unique. Section 12.2 describes how Key IDs are formed.
3.4. Text

Unless otherwise specified, the character set for text is the UTF-8 [RFC3629] encoding of Unicode [ISO10646].

3.5. Time Fields

A time field is an unsigned four-octet number containing the number of seconds elapsed since midnight, 1 January 1970 UTC.

3.6. Keyrings

A keyring is a collection of one or more keys in a file or database. Traditionally, a keyring is simply a sequential list of keys, but may be any suitable database. It is beyond the scope of this standard to discuss the details of keyrings or other databases.

3.7. String-to-Key (S2K) Specifiers

A string-to-key (S2K) specifier is used to convert a passphrase string into a symmetric-key encryption/decryption key. They are used in two places, currently: to encrypt the secret part of private keys in the private keyring, and to convert passphrases to encryption keys for symmetrically encrypted messages.

3.7.1. String-to-Key (S2K) Specifier Types

There are four types of S2K specifiers currently supported, and some reserved values:
### Table 1: S2K type registry

These are described in the subsections below.

#### 3.7.1.1. Simple S2K

This directly hashes the string to produce the key data. See below for how this hashing is done.

Octet 0: 0x00  
Octet 1: hash algorithm

Simple S2K hashes the passphrase to produce the session key. The manner in which this is done depends on the size of the session key (which will depend on the cipher used) and the size of the hash algorithm’s output. If the hash size is greater than the session key size, the high-order (leftmost) octets of the hash are used as the key.
If the hash size is less than the key size, multiple instances of the hash context are created --- enough to produce the required key data. These instances are preloaded with 0, 1, 2, ... octets of zeros (that is to say, the first instance has no preloading, the second gets preloaded with 1 octet of zero, the third is preloaded with two octets of zeros, and so forth).

As the data is hashed, it is given independently to each hash context. Since the contexts have been initialized differently, they will each produce different hash output. Once the passphrase is hashed, the output data from the multiple hashes is concatenated, first hash leftmost, to produce the key data, with any excess octets on the right discarded.

3.7.1.2. Salted S2K

This includes a "salt" value in the S2K specifier --- some arbitrary data --- that gets hashed along with the passphrase string, to help prevent dictionary attacks.

Octet 0:        0x01
Octet 1:        hash algorithm
Octets 2-9:     8-octet salt value

Salted S2K is exactly like Simple S2K, except that the input to the hash function(s) consists of the 8 octets of salt from the S2K specifier, followed by the passphrase.

3.7.1.3. Iterated and Salted S2K

This includes both a salt and an octet count. The salt is combined with the passphrase and the resulting value is hashed repeatedly. This further increases the amount of work an attacker must do to try dictionary attacks.

Octet 0:        0x03
Octet 1:        hash algorithm
Octets 2-9:     8-octet salt value
Octet 10:       count, a one-octet, coded value

The count is coded into a one-octet number using the following formula:

```c
#define EXPBIAS 6
count = ((Int32)16 + (c & 15)) << ((c >> 4) + EXPBIAS);
```

The above formula is in C, where "Int32" is a type for a 32-bit integer, and the variable "c" is the coded count, Octet 10.
Iterated-Salted S2K hashes the passphrase and salt data multiple times. The total number of octets to be hashed is specified in the encoded count in the S2K specifier. Note that the resulting count value is an octet count of how many octets will be hashed, not an iteration count.

Initially, one or more hash contexts are set up as with the other S2K algorithms, depending on how many octets of key data are needed. Then the salt, followed by the passphrase data, is repeatedly hashed until the number of octets specified by the octet count has been hashed. The one exception is that if the octet count is less than the size of the salt plus passphrase, the full salt plus passphrase will be hashed even though that is greater than the octet count. After the hashing is done, the data is unloaded from the hash context(s) as with the other S2K algorithms.

3.7.1.4. Argon2

This S2K method hashes the passphrase using Argon2, specified in [RFC9106]. This provides memory-hardness, further protecting the passphrase against brute-force attacks.

Octet  0:  0x04
Octets 1-16: 16-octet salt value
Octet  17: one-octet number of passes t
Octet  18: one-octet degree of parallelism p
Octet  19: one-octet exponent indicating the memory size m

The salt SHOULD be unique for each password.

The number of passes t and the degree of parallelism p MUST be non-zero.

The memory size m is 2**encoded_m kibibytes of RAM, where "encoded_m" is the encoded memory size in Octet 19. The encoded memory size MUST be a value from 3+ceil(log_2(p)) to 31, such that the decoded memory size m is a value from 8*p to 2**31. Note that memory-hardness size is indicated in kibibytes (KiB), not octets.

Argon2 is invoked with the passphrase as P, the salt as S, the values of t, p and m as described above, the required key size as the tag length T, 0x13 as the version v, and Argon2id as the type.

For the recommended values of t, p and m, see Section 4 of [RFC9106]. If the recommended value of m for a given application is not a power of 2, it is RECOMMENDED to round up to the next power of 2 if the resulting performance would be acceptable, and round down otherwise (keeping in mind that m must be at least 8*p).
As an example, with the first recommended option (t=1, p=4, m=2**21), the full S2K specifier would be:

```
04 XX XX XX XX XX XX XX XX XX XX XX XX XX XX XX XX
XX 01 04 15
```

(where XX represents a random octet of salt).

### 3.7.2. String-to-Key Usage

Simple S2K and Salted S2K specifiers can be brute-forced when used with a low-entropy string, such as those typically provided by users. In addition, the usage of Simple S2K can lead to key and IV reuse (see Section 5.3). Therefore, when generating S2K specifiers, implementations MUST NOT use Simple S2K, and SHOULD NOT use Salted S2K unless the implementation knows that the string is high-entropy (for example, it generated the string itself using a known-good source of randomness). It is RECOMMENDED that implementations use Argon2.

#### 3.7.2.1. Secret-Key Encryption

An S2K specifier can be stored in the secret keyring to specify how to convert the passphrase to a key that unlocks the secret data. Older versions of PGP just stored a symmetric cipher algorithm octet preceding the secret data or a zero to indicate that the secret data was unencrypted. The MD5 hash function was always used to convert the passphrase to a key for the specified cipher algorithm.

For compatibility, when an S2K specifier is used, the special value 253, 254, or 255 is stored in the position where the cipher algorithm octet would have been in the old data structure. This is then followed immediately by a one-octet algorithm identifier, and other fields relevant to the type of encryption used.

Therefore, the first octet of the secret key material describes how the secret key data is presented.

In the table below, check(x) means the "2-octet checksum" meaning the sum of all octets in x mod 65536.
<table>
<thead>
<tr>
<th>First octet</th>
<th>Next fields</th>
<th>Encryption</th>
<th>Generate?</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>-</td>
<td>cleartext</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>secrets</td>
<td>check(secrets)</td>
</tr>
<tr>
<td>Known symmetric cipher algo ID (see Section 9.3)</td>
<td>IV</td>
<td>CFB(MD5(password), secrets</td>
<td></td>
</tr>
<tr>
<td>253</td>
<td>cipher-algo, AEAD-mode, S2K-specifier, nonce</td>
<td>AEAD(S2K(password), secrets, pubkey)</td>
<td>Yes</td>
</tr>
<tr>
<td>254</td>
<td>cipher-algo, S2K-specifier, IV</td>
<td>CFB(S2K(password), secrets</td>
<td></td>
</tr>
<tr>
<td>255</td>
<td>cipher-algo, S2K-specifier, IV</td>
<td>CFB(S2K(password), secrets</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Secret Key protection details

Each row with "Generate?" marked as "No" is described for backward compatibility, and MUST NOT be generated.

An implementation MUST NOT create and MUST reject as malformed a secret key packet where the S2K usage octet is anything but 253 and the S2K specifier type is Argon2.

3.7.2.2. Symmetric-Key Message Encryption

OpenPGP can create a Symmetric-key Encrypted Session Key (ESK) packet at the front of a message. This is used to allow S2K specifiers to be used for the passphrase conversion or to create messages with a mix of symmetric-key ESKs and public-key ESKs. This allows a message to be decrypted either with a passphrase or a public-key pair.

PGP 2 always used IDEA with Simple string-to-key conversion when encrypting a message with a symmetric algorithm. See Section 5.8. This MUST NOT be generated, but MAY be consumed for backward-compatibility.
4. Packet Syntax

This section describes the packets used by OpenPGP.

4.1. Overview

An OpenPGP message is constructed from a number of records that are traditionally called packets. A packet is a chunk of data that has a tag specifying its meaning. An OpenPGP message, keyring, certificate, and so forth consists of a number of packets. Some of those packets may contain other OpenPGP packets (for example, a compressed data packet, when uncompressed, contains OpenPGP packets).

Each packet consists of a packet header, followed by the packet body. The packet header is of variable length.

When handling a stream of packets, the length information in each packet header is the canonical source of packet boundaries. An implementation handling a packet stream that wants to find the next packet MUST look for it at the precise offset indicated in the previous packet header.

Additionally, some packets contain internal length indicators (for example, a subfield within the packet). In the event that a subfield length indicator within a packet implies inclusion of octets outside the range indicated in the packet header, a parser MUST truncate the subfield at the octet boundary indicated in the packet header. Such a truncation renders the packet malformed and unusable. An implementation MUST NOT interpret octets outside the range indicated in the packet header as part of the contents of the packet.

4.2. Packet Headers

The first octet of the packet header is called the "Packet Tag". It determines the format of the header and denotes the packet contents. The remainder of the packet header is the length of the packet.

There are two packet formats, the (current) OpenPGP packet format specified by this document and its predecessors and the Legacy packet format as used by PGP 2.x implementations.

Note that the most significant bit is the leftmost bit, called bit 7. A mask for this bit is 0x80 in hexadecimal.
PTag 7 6 5 4 3 2 1 0

Bit 7 -- Always one
Bit 6 -- Always one (except for Legacy packet format)

The Legacy packet format MAY be used when consuming packets to facilitate interoperability with legacy implementations and accessing archived data. The Legacy packet format SHOULD NOT be used to generate new data, unless the recipient is known to only support the Legacy packet format.

An implementation that consumes and re-distributes pre-existing OpenPGP data (such as Transferable Public Keys) may encounter packets framed with the Legacy packet format. Such an implementation MAY either re-distribute these packets in their Legacy format, or transform them to the current OpenPGP packet format before re-distribution.

The current OpenPGP packet format packets contain:

Bits 5 to 0 -- packet tag

Legacy packet format packets contain:

Bits 5 to 2 -- packet tag
Bits 1 to 0 -- length-type

4.2.1. OpenPGP Format Packet Lengths

OpenPGP format packets have four possible ways of encoding length:

1. A one-octet Body Length header encodes packet lengths of up to 191 octets.

2. A two-octet Body Length header encodes packet lengths of 192 to 8383 octets.

3. A five-octet Body Length header encodes packet lengths of up to 4,294,967,295 (0xFFFFFFFF) octets in length. (This actually encodes a four-octet scalar number.)

4. When the length of the packet body is not known in advance by the issuer, Partial Body Length headers encode a packet of indeterminate length, effectively making it a stream.
4.2.1.1. One-Octet Lengths

A one-octet Body Length header encodes a length of 0 to 191 octets. This type of length header is recognized because the one octet value is less than 192. The body length is equal to:

\[
\text{bodyLen} = \text{1st}_\text{octet};
\]

4.2.1.2. Two-Octet Lengths

A two-octet Body Length header encodes a length of 192 to 8383 octets. It is recognized because its first octet is in the range 192 to 223. The body length is equal to:

\[
\text{bodyLen} = \left((\text{1st}_\text{octet} - 192) \ll 8\right) + \left(\text{2nd}_\text{octet}\right) + 192
\]

4.2.1.3. Five-Octet Lengths

A five-octet Body Length header consists of a single octet holding the value 255, followed by a four-octet scalar. The body length is equal to:

\[
\text{bodyLen} = \left(\text{2nd}_\text{octet} \ll 24\right) \mid \left(\text{3rd}_\text{octet} \ll 16\right) \mid \left(\text{4th}_\text{octet} \ll 8\right) \mid \text{5th}_\text{octet}
\]

This basic set of one, two, and five-octet lengths is also used internally to some packets.

4.2.1.4. Partial Body Lengths

A Partial Body Length header is one octet long and encodes the length of only part of the data packet. This length is a power of 2, from 1 to 1,073,741,824 (2 to the 30th power). It is recognized by its one octet value that is greater than or equal to 224, and less than 255. The Partial Body Length is equal to:

\[
\text{partialBodyLen} = 1 \ll (\text{1st}_\text{octet} \& 0x1F);
\]

Each Partial Body Length header is followed by a portion of the packet body data. The Partial Body Length header specifies this portion’s length. Another length header (one octet, two-octet, five-octet, or partial) follows that portion. The last length header in the packet MUST NOT be a Partial Body Length header. Partial Body Length headers may only be used for the non-final parts of the packet.

Note also that the last Body Length header can be a zero-length header.
An implementation MAY use Partial Body Lengths for data packets, be they literal, compressed, or encrypted. The first partial length MUST be at least 512 octets long. Partial Body Lengths MUST NOT be used for any other packet types.

4.2.2. Legacy Format Packet Lengths

The meaning of the length-type in Legacy format packets is:

0  The packet has a one-octet length. The header is 2 octets long.
1  The packet has a two-octet length. The header is 3 octets long.
2  The packet has a four-octet length. The header is 5 octets long.

3  The packet is of indeterminate length. The header is 1 octet long, and the implementation must determine how long the packet is. If the packet is in a file, this means that the packet extends until the end of the file. The OpenPGP format headers have a mechanism for precisely encoding data of indeterminate length. An implementation MUST NOT generate a Legacy format packet with indeterminate length. An implementation MAY interpret an indeterminate length Legacy format packet in order to deal with historic data, or data generated by a legacy system.

4.2.3. Packet Length Examples

These examples show ways that OpenPGP format packets might encode the packet lengths.

A packet with length 100 may have its length encoded in one octet: 0x64. This is followed by 100 octets of data.

A packet with length 1723 may have its length encoded in two octets: 0xC5, 0xFB. This header is followed by the 1723 octets of data.

A packet with length 100000 may have its length encoded in five octets: 0xFF, 0x00, 0x01, 0x86, 0xA0.

It might also be encoded in the following octet stream: 0xEF, first 32768 octets of data; 0xE1, next two octets of data; 0xE0, next one octet of data; 0xF0, next 65536 octets of data; 0xC5, 0xDD, last 1693 octets of data. This is just one possible encoding, and many variations are possible on the size of the Partial Body Length headers, as long as a regular Body Length header encodes the last portion of the data.
Please note that in all of these explanations, the total length of
the packet is the length of the header(s) plus the length of the
body.

4.3. Packet Tags

The packet tag denotes what type of packet the body holds. Note that
Legacy format headers can only have tags less than 16, whereas
OpenPGP format headers can have tags as great as 63. The defined
tags (in decimal) are as follows:

<table>
<thead>
<tr>
<th>Tag</th>
<th>Packet Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Reserved - a packet tag MUST NOT have this value</td>
</tr>
<tr>
<td>1</td>
<td>Public-Key Encrypted Session Key Packet</td>
</tr>
<tr>
<td>2</td>
<td>Signature Packet</td>
</tr>
<tr>
<td>3</td>
<td>Symmetric-Key Encrypted Session Key Packet</td>
</tr>
<tr>
<td>4</td>
<td>One-Pass Signature Packet</td>
</tr>
<tr>
<td>5</td>
<td>Secret-Key Packet</td>
</tr>
<tr>
<td>6</td>
<td>Public-Key Packet</td>
</tr>
<tr>
<td>7</td>
<td>Secret-Subkey Packet</td>
</tr>
<tr>
<td>8</td>
<td>Compressed Data Packet</td>
</tr>
<tr>
<td>9</td>
<td>Symmetrically Encrypted Data Packet</td>
</tr>
<tr>
<td>10</td>
<td>Marker Packet</td>
</tr>
<tr>
<td>11</td>
<td>Literal Data Packet</td>
</tr>
<tr>
<td>12</td>
<td>Trust Packet</td>
</tr>
<tr>
<td>13</td>
<td>User ID Packet</td>
</tr>
<tr>
<td>14</td>
<td>Public-Subkey Packet</td>
</tr>
<tr>
<td>17</td>
<td>User Attribute Packet</td>
</tr>
<tr>
<td>18</td>
<td>Sym. Encrypted and Integrity Protected Data Packet</td>
</tr>
</tbody>
</table>
5. Packet Types

5.1. Public-Key Encrypted Session Key Packets (Tag 1)

Zero or more Public-Key Encrypted Session Key (PKESK) packets and/or Symmetric-Key Encrypted Session Key packets (Section 5.3) may precede an encryption container (that is, a Symmetrically Encrypted Integrity Protected Data packet or --- for historic data --- a Symmetrically Encrypted Data packet), which holds an encrypted message. The message is encrypted with the session key, and the session key is itself encrypted and stored in the Encrypted Session Key packet(s). The encryption container is preceded by one Public-Key Encrypted Session Key packet for each OpenPGP key to which the message is encrypted. The recipient of the message finds a session key that is encrypted to their public key, decrypts the session key, and then uses the session key to decrypt the message.

The body of this packet starts with a one-octet number giving the version number of the packet type. The currently defined versions are 3 and 5. The remainder of the packet depends on the version.

The versions differ in how they identify the recipient key, and in what they encode. The version of the PKESK packet must align with the version of the SEIPD packet (see Section 11.3.2.1).

5.1.1. v3 PKESK

A version 3 Public-Key Encrypted Session Key (PKESK) packet precedes a version 1 Symmetrically Encrypted Integrity Protected Data (v1 SEIPD, see Section 5.14.1) packet. In historic data, it is sometimes found preceding a deprecated Symmetrically Encrypted Data packet (SED, see Section 5.8). A v3 PKESK packet MUST NOT precede a v2 SEIPD packet (see Section 11.3.2.1).

The v3 PKESK packet consists of:
* A one-octet version number with value 3.

* An eight-octet number that gives the Key ID of the public key to which the session key is encrypted. If the session key is encrypted to a subkey, then the Key ID of this subkey is used here instead of the Key ID of the primary key. The Key ID may also be all zeros, for an "anonymous recipient" (see Section 5.1.6).

* A one-octet number giving the public-key algorithm used.

* A series of values comprising the encrypted session key. This is algorithm-specific and described below.

When creating a v3 PKESK packet, the session key is first prefixed with a one-octet algorithm identifier that specifies the symmetric encryption algorithm used to encrypt the following encryption container. Then a two-octet checksum is appended, which is equal to the sum of the preceding session key octets, not including the algorithm identifier, modulo 65536.

The resulting octet string (algorithm identifier, session key, and checksum) is encrypted according to the public-key algorithm used, as described below.

5.1.2. v5 PKESK

A version 5 Public-Key Encrypted Session Key (PKESK) packet precedes a version 2 Symmetrically Encrypted Integrity Protected Data (v2 SEIPD, see Section 5.14.2) packet. A v5 PKESK packet MUST NOT precede a v1 SEIPD packet or a deprecated Symmetrically Encrypted Data packet (see Section 11.3.2.1).

The v5 PKESK packet consists of:

* A one-octet version number with value 5.

* A one octet key version number and N octets of the fingerprint of the public key or subkey to which the session key is encrypted. Note that the length N of the fingerprint for a version 4 key is 20 octets; for a version 5 key N is 32. The key version number may also be zero, and the fingerprint omitted (that is, the length N is zero in this case), for an "anonymous recipient" (see Section 5.1.6).

* A one-octet number giving the public-key algorithm used.

* A series of values comprising the encrypted session key. This is algorithm-specific and described below.
When creating a V5 PKESK packet, the symmetric encryption algorithm identifier is not included. Before encrypting, a two-octet checksum is appended, which is equal to the sum of the preceding session key octets, modulo 65536.

The resulting octet string (session key and checksum) is encrypted according to the public-key algorithm used, as described below.

5.1.3. Algorithm Specific Fields for RSA encryption

* Multiprecision integer (MPI) of RSA-encrypted value \( m^e \mod n \).

The value "m" in the above formula is the plaintext value described above, encoded in the PKCS#1 block encoding EME-PKCS1-v1_5 described in Section 7.2.1 of [RFC8017] (see also Section 14.1). Note that when an implementation forms several PKESKs with one session key, forming a message that can be decrypted by several keys, the implementation MUST make a new PKCS#1 encoding for each key.

5.1.4. Algorithm Specific Fields for Elgamal encryption

* MPI of Elgamal (Diffie-Hellman) value \( g^k \mod p \).

* MPI of Elgamal (Diffie-Hellman) value \( m \cdot y^k \mod p \).

The value "m" in the above formula is the plaintext value described above, encoded in the PKCS#1 block encoding EME-PKCS1-v1_5 described in Section 7.2.1 of [RFC8017] (see also Section 14.1). Note that when an implementation forms several PKESKs with one session key, forming a message that can be decrypted by several keys, the implementation MUST make a new PKCS#1 encoding for each key.

5.1.5. Algorithm-Specific Fields for ECDH encryption

* MPI of an EC point representing an ephemeral public key, in the point format associated with the curve as specified in Section 9.2.

* A one-octet size, followed by a symmetric key encoded using the method described in Section 13.5.

5.1.6. Notes on PKESK

An implementation MAY accept or use a Key ID of all zeros, or a key version of zero and no key fingerprint, to hide the intended decryption key. In this case, the receiving implementation would try all available private keys, checking for a valid decrypted session key. This format helps reduce traffic analysis of messages.
5.2. Signature Packet (Tag 2)

A Signature packet describes a binding between some public key and some data. The most common signatures are a signature of a file or a block of text, and a signature that is a certification of a User ID.

Three versions of Signature packets are defined. Version 3 provides basic signature information, while versions 4 and 5 provide an expandable format with subpackets that can specify more information about the signature.

An implementation MUST generate a version 5 signature when signing with a version 5 key. An implementation MUST generate a version 4 signature when signing with a version 4 key. Implementations MUST NOT create version 3 signatures; they MAY accept version 3 signatures.

5.2.1. Signature Types

There are a number of possible meanings for a signature, which are indicated in a signature type octet in any given signature. Please note that the vagueness of these meanings is not a flaw, but a feature of the system. Because OpenPGP places final authority for validity upon the receiver of a signature, it may be that one signer’s casual act might be more rigorous than some other authority’s positive act. See Section 5.2.4 for detailed information on how to compute and verify signatures of each type.

These meanings are as follows:

0x00: Signature of a binary document.
   This means the signer owns it, created it, or certifies that it has not been modified.

0x01: Signature of a canonical text document.
   This means the signer owns it, created it, or certifies that it has not been modified. The signature is calculated over the text data with its line endings converted to <CR><LF>.

0x02: Standalone signature.
   This signature is a signature of only its own subpacket contents. It is calculated identically to a signature over a zero-length binary document. V3 standalone signatures MUST NOT be generated and MUST be ignored.

0x10: Generic certification of a User ID and Public-Key packet.
The issuer of this certification does not make any particular assertion as to how well the certifier has checked that the owner of the key is in fact the person described by the User ID.

0x11: Persona certification of a User ID and Public-Key packet.
The issuer of this certification has not done any verification of the claim that the owner of this key is the User ID specified.

0x12: Casual certification of a User ID and Public-Key packet.
The issuer of this certification has done some casual verification of the claim of identity.

0x13: Positive certification of a User ID and Public-Key packet.
The issuer of this certification has done substantial verification of the claim of identity.

Most OpenPGP implementations make their "key signatures" as 0x10 certifications. Some implementations can issue 0x11-0x13 certifications, but few differentiate between the types.

0x18: Subkey Binding Signature.
This signature is a statement by the top-level signing key that indicates that it owns the subkey. This signature is calculated directly on the primary key and subkey, and not on any User ID or other packets. A signature that binds a signing subkey MUST have an Embedded Signature subpacket in this binding signature that contains a 0x19 signature made by the signing subkey on the primary key and subkey.

0x19: Primary Key Binding Signature.
This signature is a statement by a signing subkey, indicating that it is owned by the primary key and subkey. This signature is calculated the same way as a 0x18 signature: directly on the primary key and subkey, and not on any User ID or other packets.

0x1F: Signature directly on a key.
This signature is calculated directly on a key. It binds the information in the Signature subpackets to the key, and is appropriate to be used for subpackets that provide information about the key, such as the Key Flags subpacket or (deprecated) Revocation Key. It is also appropriate for statements that non-self certifiers want to make about the key itself, rather than the binding between a key and a name.

0x20: Key revocation signature.
The signature is calculated directly on the key being revoked. A revoked key is not to be used. Only revocation signatures by the key being revoked, or by a (deprecated) Revocation Key, should be considered valid revocation signatures.

0x28: Subkey revocation signature.  
The signature is calculated directly on the subkey being revoked. A revoked subkey is not to be used. Only revocation signatures by the top-level signature key that is bound to this subkey, or by a (deprecated) Revocation Key, should be considered valid revocation signatures.

0x30: Certification revocation signature.  
This signature revokes an earlier User ID certification signature (signature class 0x10 through 0x13) or direct-key signature (0x1F). It should be issued by the same key that issued the revoked signature or by a (deprecated) Revocation Key. The signature is computed over the same data as the certificate that it revokes, and should have a later creation date than that certificate.

0x40: Timestamp signature.  
This signature is only meaningful for the timestamp contained in it.

0x50: Third-Party Confirmation signature.  
This signature is a signature over some other OpenPGP Signature packet(s). It is analogous to a notary seal on the signed data. A third-party signature SHOULD include Signature Target subpacket(s) to give easy identification. Note that we really do mean SHOULD. There are plausible uses for this (such as a blind party that only sees the signature, not the key or source document) that cannot include a target subpacket.

5.2.2. Version 3 Signature Packet Format

The body of a version 3 Signature Packet contains:

* One-octet version number (3).
* One-octet length of following hashed material. MUST be 5.
  - One-octet signature type.
  - Four-octet creation time.
* Eight-octet Key ID of signer.
* One-octet public-key algorithm.

* One-octet hash algorithm.

* Two-octet field holding left 16 bits of signed hash value.

* One or more multiprecision integers comprising the signature. This portion is algorithm specific, as described below.

The concatenation of the data to be signed, the signature type, and creation time from the Signature packet (5 additional octets) is hashed. The resulting hash value is used in the signature algorithm. The high 16 bits (first two octets) of the hash are included in the Signature packet to provide a way to reject some invalid signatures without performing a signature verification.

Algorithm-Specific Fields for RSA signatures:

* Multiprecision integer (MPI) of RSA signature value m**d mod n.

Algorithm-Specific Fields for DSA signatures:

* MPI of DSA value r.

* MPI of DSA value s.

The signature calculation is based on a hash of the signed data, as described above. The details of the calculation are different for DSA signatures than for RSA signatures.

With RSA signatures, the hash value is encoded using PKCS#1 encoding type EMSA-PKCS1-v1_5 as described in Section 9.2 of [RFC8017]. This requires inserting the hash value as an octet string into an ASN.1 structure. The object identifier for the type of hash being used is included in the structure. The hexadecimal representations for the currently defined hash algorithms are as follows:
<table>
<thead>
<tr>
<th>algorithm</th>
<th>hexadecimal representation</th>
</tr>
</thead>
<tbody>
<tr>
<td>MD5</td>
<td>0x2A, 0x86, 0x48, 0x86, 0xF7, 0x0D, 0x02, 0x05</td>
</tr>
<tr>
<td>RIPEMD-160</td>
<td>0x2B, 0x24, 0x03, 0x02, 0x01</td>
</tr>
<tr>
<td>SHA-1</td>
<td>0x2B, 0x0E, 0x03, 0x02, 0x1A</td>
</tr>
<tr>
<td>SHA224</td>
<td>0x2B, 0x0E, 0x03, 0x02, 0x04</td>
</tr>
<tr>
<td>SHA256</td>
<td>0x60, 0x86, 0x48, 0x01, 0x65, 0x03, 0x04, 0x02, 0x01</td>
</tr>
<tr>
<td>SHA384</td>
<td>0x60, 0x86, 0x48, 0x01, 0x65, 0x03, 0x04, 0x02, 0x02</td>
</tr>
<tr>
<td>SHA512</td>
<td>0x60, 0x86, 0x48, 0x01, 0x65, 0x03, 0x04, 0x02, 0x03</td>
</tr>
</tbody>
</table>

Table 4: Hash hexadecimal representations

The ASN.1 Object Identifiers (OIDs) are as follows:

<table>
<thead>
<tr>
<th>algorithm</th>
<th>OID</th>
</tr>
</thead>
<tbody>
<tr>
<td>MD5</td>
<td>1.2.840.113549.2.5</td>
</tr>
<tr>
<td>RIPEMD-160</td>
<td>1.3.36.3.2.1</td>
</tr>
<tr>
<td>SHA-1</td>
<td>1.3.14.3.2.26</td>
</tr>
<tr>
<td>SHA224</td>
<td>2.16.840.1.101.3.4.2.4</td>
</tr>
<tr>
<td>SHA256</td>
<td>2.16.840.1.101.3.4.2.1</td>
</tr>
<tr>
<td>SHA384</td>
<td>2.16.840.1.101.3.4.2.2</td>
</tr>
<tr>
<td>SHA512</td>
<td>2.16.840.1.101.3.4.2.3</td>
</tr>
</tbody>
</table>

Table 5: Hash OIDs

The full hash prefixes for these are as follows:
<table>
<thead>
<tr>
<th>algorithm</th>
<th>full hash prefix</th>
</tr>
</thead>
<tbody>
<tr>
<td>MD5</td>
<td>0x30, 0x20, 0x30, 0x0C, 0x06, 0x08, 0x2A, 0x86, 0x48, 0x86, 0xF7, 0x0D, 0x02, 0x05, 0x00, 0x04, 0x10</td>
</tr>
<tr>
<td>RIPEMD-160</td>
<td>0x30, 0x21, 0x30, 0x09, 0x06, 0x05, 0x2B, 0x03, 0x02, 0x01, 0x05, 0x00, 0x04, 0x14</td>
</tr>
<tr>
<td>SHA-1</td>
<td>0x30, 0x21, 0x30, 0x09, 0x06, 0x05, 0x2B, 0x0E, 0x03, 0x02, 0x1A, 0x05, 0x00, 0x04, 0x14</td>
</tr>
<tr>
<td>SHA224</td>
<td>0x30, 0x2D, 0x30, 0x0D, 0x06, 0x09, 0x60, 0x86, 0x48, 0x01, 0x65, 0x03, 0x04, 0x02, 0x04, 0x05, 0x00, 0x04, 0x1C</td>
</tr>
<tr>
<td>SHA256</td>
<td>0x30, 0x31, 0x30, 0x0D, 0x06, 0x09, 0x60, 0x86, 0x48, 0x01, 0x65, 0x03, 0x04, 0x02, 0x01, 0x05, 0x00, 0x04, 0x04, 0x20</td>
</tr>
<tr>
<td>SHA384</td>
<td>0x30, 0x41, 0x30, 0x0D, 0x06, 0x09, 0x60, 0x86, 0x48, 0x01, 0x65, 0x03, 0x04, 0x02, 0x02, 0x05, 0x00, 0x04, 0x04, 0x30</td>
</tr>
<tr>
<td>SHA512</td>
<td>0x30, 0x51, 0x30, 0x0D, 0x06, 0x09, 0x60, 0x86, 0x48, 0x01, 0x65, 0x03, 0x04, 0x02, 0x03, 0x05, 0x00, 0x04, 0x40</td>
</tr>
</tbody>
</table>

Table 6: Hash hexadecimal prefixes

DSA signatures MUST use hashes that are equal in size to the number of bits of q, the group generated by the DSA key's generator value.

If the output size of the chosen hash is larger than the number of bits of q, the hash result is truncated to fit by taking the number of leftmost bits equal to the number of bits of q. This (possibly truncated) hash function result is treated as a number and used directly in the DSA signature algorithm.

5.2.3. Version 4 and 5 Signature Packet Formats

The body of a V4 or V5 Signature packet contains:
* One-octet version number. This is 4 for V4 signatures and 5 for V5 signatures.
* One-octet signature type.
* One-octet public-key algorithm.
* One-octet hash algorithm.

* A scalar octet count for following hashed subpacket data. For a V4 signature, this is a two-octet field. For a V5 signature, this is a four-octet field. Note that this is the length in octets of all of the hashed subpackets; a pointer incremented by this number will skip over the hashed subpackets.

* Hashed subpacket data set (zero or more subpackets).

* A scalar octet count for the following unhashed subpacket data. For a V4 signature, this is a two-octet field. For a V5 signature, this is a four-octet field. Note that this is the length in octets of all of the unhashed subpackets; a pointer incremented by this number will skip over the unhashed subpackets.

* Unhashed subpacket data set (zero or more subpackets).

* Two-octet field holding the left 16 bits of the signed hash value.

* Only for V5 signatures, a 16 octet field containing random values used as salt.

* One or more multiprecision integers comprising the signature. This portion is algorithm specific:

5.2.3.1. Algorithm-Specific Fields for RSA signatures

* Multiprecision integer (MPI) of RSA signature value m**d mod n.

5.2.3.2. Algorithm-Specific Fields for DSA or ECDSA signatures

* MPI of DSA or ECDSA value r.
* MPI of DSA or ECDSA value s.

A version 3 signature MUST NOT be created and MUST NOT be used with ECDSA.

5.2.3.3. Algorithm-Specific Fields for EdDSA signatures
* Two MPI-encoded values, whose contents and formatting depend on the choice of curve used (see Section 9.2.1).

A version 3 signature MUST NOT be created and MUST NOT be used with EdDSA.

5.2.3.3.1. Algorithm-Specific Fields for Ed25519 signatures

The two MPIs for Ed25519 use octet strings $R$ and $S$ as described in [RFC8032].

* MPI of an EC point $R$, represented as a (non-prefixed) native (little-endian) octet string up to 32 octets.

* MPI of EdDSA value $S$, also in (non-prefixed) native little-endian format with a length up to 32 octets.

5.2.3.3.2. Algorithm-Specific Fields for Ed448 signatures

For Ed448 signatures, the native signature format is used as described in [RFC8032]. The two MPIs are composed as follows:

* The first MPI has a body of 58 octets: a prefix 0x40 octet, followed by 57 octets of the native signature.

* The second MPI is set to 0 (this is a placeholder, and is unused). Note that an MPI with a value of 0 is encoded on the wire as a pair of zero octets: 00 00.

5.2.3.4. Notes on Signatures

The concatenation of the data being signed and the signature data from the version number through the hashed subpacket data (inclusive) is hashed. The resulting hash value is what is signed. The high 16 bits (first two octets) of the hash are included in the Signature packet to provide a way to reject some invalid signatures without performing a signature verification.

There are two fields consisting of Signature subpackets. The first field is hashed with the rest of the signature data, while the second is unhashed. The second set of subpackets is not cryptographically protected by the signature and should include only advisory information.

The differences between a V4 and V5 signature are two-fold: first, a V5 signature increases the width of the size indicators for the signed data, making it more capable when signing large keys or messages. Second, the hash is salted with 128 bit of random data.
The algorithms for converting the hash function result to a signature are described in Section 5.2.4.

5.2.3.5. Signature Subpacket Specification

A subpacket data set consists of zero or more Signature subpackets. In Signature packets, the subpacket data set is preceded by a two-octet (for V4 signatures) or four-octet (for V5 signatures) scalar count of the length in octets of all the subpackets. A pointer incremented by this number will skip over the subpacket data set.

Each subpacket consists of a subpacket header and a body. The header consists of:

* the subpacket length (1, 2, or 5 octets),
* the subpacket type (1 octet),

and is followed by the subpacket-specific data.

The length includes the type octet but not this length. Its format is similar to the "new" format packet header lengths, but cannot have Partial Body Lengths. That is:

if the 1st octet < 192, then
    lengthOfLength = 1
    subpacketLen = 1st_octet

if the 1st octet >= 192 and < 255, then
    lengthOfLength = 2
    subpacketLen = ((1st_octet - 192) << 8) + (2nd_octet) + 192

if the 1st octet = 255, then
    lengthOfLength = 5
    subpacket length = [four-octet scalar starting at 2nd_octet]

The value of the subpacket type octet may be:

<table>
<thead>
<tr>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Reserved</td>
</tr>
<tr>
<td>1</td>
<td>Reserved</td>
</tr>
<tr>
<td>2</td>
<td>Signature Creation Time</td>
</tr>
<tr>
<td>3</td>
<td>Signature Expiration Time</td>
</tr>
<tr>
<td></td>
<td>Description</td>
</tr>
<tr>
<td>---</td>
<td>-------------------------------------------------</td>
</tr>
<tr>
<td>4</td>
<td>Exportable Certification</td>
</tr>
<tr>
<td>5</td>
<td>Trust Signature</td>
</tr>
<tr>
<td>6</td>
<td>Regular Expression</td>
</tr>
<tr>
<td>7</td>
<td>Revocable</td>
</tr>
<tr>
<td>8</td>
<td>Reserved</td>
</tr>
<tr>
<td>9</td>
<td>Key Expiration Time</td>
</tr>
<tr>
<td>10</td>
<td>Placeholder for backward compatibility</td>
</tr>
<tr>
<td>11</td>
<td>Preferred Symmetric Ciphers for v1 SEIPD</td>
</tr>
<tr>
<td>12</td>
<td>Revocation Key (deprecated)</td>
</tr>
<tr>
<td>13 to 15</td>
<td>Reserved</td>
</tr>
<tr>
<td>16</td>
<td>Issuer</td>
</tr>
<tr>
<td>17 to 19</td>
<td>Reserved</td>
</tr>
<tr>
<td>20</td>
<td>Notation Data</td>
</tr>
<tr>
<td>21</td>
<td>Preferred Hash Algorithms</td>
</tr>
<tr>
<td>22</td>
<td>Preferred Compression Algorithms</td>
</tr>
<tr>
<td>23</td>
<td>Key Server Preferences</td>
</tr>
<tr>
<td>24</td>
<td>Preferred Key Server</td>
</tr>
<tr>
<td>25</td>
<td>Primary User ID</td>
</tr>
<tr>
<td>26</td>
<td>Policy URI</td>
</tr>
<tr>
<td>27</td>
<td>Key Flags</td>
</tr>
<tr>
<td>28</td>
<td>Signer’s User ID</td>
</tr>
<tr>
<td>29</td>
<td>Reason for Revocation</td>
</tr>
<tr>
<td>30</td>
<td>Features</td>
</tr>
<tr>
<td>31</td>
<td>Signature Target</td>
</tr>
</tbody>
</table>
### Table 7: Subpacket type registry

<table>
<thead>
<tr>
<th>Subpacket Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>32</td>
<td>Embedded Signature</td>
</tr>
<tr>
<td>33</td>
<td>Issuer Fingerprint</td>
</tr>
<tr>
<td>34</td>
<td>Reserved</td>
</tr>
<tr>
<td>35</td>
<td>Intended Recipient Fingerprint</td>
</tr>
<tr>
<td>37</td>
<td>Reserved (Attested Certifications)</td>
</tr>
<tr>
<td>38</td>
<td>Reserved (Key Block)</td>
</tr>
<tr>
<td>39</td>
<td>Preferred AEAD Ciphersuites</td>
</tr>
<tr>
<td>100 to 110</td>
<td>Private or experimental</td>
</tr>
</tbody>
</table>

An implementation SHOULD ignore any subpacket of a type that it does not recognize.

Bit 7 of the subpacket type is the "critical" bit. If set, it denotes that the subpacket is one that is critical for the evaluator of the signature to recognize. If a subpacket is encountered that is marked critical but is unknown to the evaluating software, the evaluator SHOULD consider the signature to be in error.

An evaluator may "recognize" a subpacket, but not implement it. The purpose of the critical bit is to allow the signer to tell an evaluator that it would prefer a new, unknown feature to generate an error than be ignored.

Implementations SHOULD implement the four preferred algorithm subpackets (11, 21, 22, and 34), as well as the "Reason for Revocation" subpacket. Note, however, that if an implementation chooses not to implement some of the preferences, it is required to behave in a polite manner to respect the wishes of those users who do implement these preferences.
5.2.3.6. Signature Subpacket Types

A number of subpackets are currently defined. Some subpackets apply to the signature itself and some are attributes of the key. Subpackets that are found on a self-signature are placed on a certification made by the key itself. Note that a key may have more than one User ID, and thus may have more than one self-signature, and differing subpackets.

A subpacket may be found either in the hashed or unhashed subpacket sections of a signature. If a subpacket is not hashed, then the information in it cannot be considered definitive because it is not part of the signature proper.

5.2.3.7. Notes on Self-Signatures

A self-signature is a binding signature made by the key to which the signature refers. There are three types of self-signatures, the certification signatures (types 0x10–0x13), the direct-key signature (type 0x1F), and the subkey binding signature (type 0x18). A cryptographically-valid self-signature should be accepted from any primary key, regardless of what Key Flags (Section 5.2.3.26) apply to the primary key. In particular, a primary key does not need to have 0x01 set in the first octet of Key Flags order to make a valid self-signature.

For certification self-signatures, each User ID may have a self-signature, and thus different subpackets in those self-signatures. For subkey binding signatures, each subkey in fact has a self-signature. Subpackets that appear in a certification self-signature apply to the user name, and subpackets that appear in the subkey self-signature apply to the subkey. Lastly, subpackets on the direct-key signature apply to the entire key.

Implementing software should interpret a self-signature’s preference subpackets as narrowly as possible. For example, suppose a key has two user names, Alice and Bob. Suppose that Alice prefers the AEAD ciphersuite AES-256 with OCB, and Bob prefers Camellia-256 with GCM. If the software locates this key via Alice’s name, then the preferred AEAD ciphersuite is AES-256 with OCB; if software locates the key via Bob’s name, then the preferred algorithm is Camellia-256 with GCM. If the key is located by Key ID, the algorithm of the primary User ID of the key provides the preferred AEAD ciphersuite.

Revoking a self-signature or allowing it to expire has a semantic meaning that varies with the signature type. Revoking the self-signature on a User ID effectively retires that user name. The self-signature is a statement, "My name X is tied to my signing key K" and
is corroborated by other users’ certifications. If another user
revokes their certification, they are effectively saying that they no
longer believe that name and that key are tied together. Similarly,
if the users themselves revoke their self-signature, then the users
no longer go by that name, no longer have that email address, etc.
Revoking a binding signature effectively retires that subkey.
Revoking a direct-key signature cancels that signature. Please see
Section 5.2.3.28 for more relevant detail.

Since a self-signature contains important information about the key’s
use, an implementation SHOULD allow the user to rewrite the self-
signature, and important information in it, such as preferences and
key expiration.

It is good practice to verify that a self-signature imported into an
implementation doesn’t advertise features that the implementation
doesn’t support, rewriting the signature as appropriate.

An implementation that encounters multiple self-signatures on the
same object may resolve the ambiguity in any way it sees fit, but it
is RECOMMENDED that priority be given to the most recent self-
signature.

5.2.3.8. Signature Creation Time

(4-octet time field)

The time the signature was made.

MUST be present in the hashed area.

5.2.3.9. Issuer

(8-octet Key ID)

The OpenPGP Key ID of the key issuing the signature. If the version
of that key is greater than 4, this subpacket MUST NOT be included in
the signature.

5.2.3.10. Key Expiration Time

(4-octet time field)
The validity period of the key. This is the number of seconds after
the key creation time that the key expires. For a direct or
certification self-signature, the key creation time is that of the
primary key. For a subkey binding signature, the key creation time
is that of the subkey. If this is not present or has a value of
zero, the key never expires. This is found only on a self-signature.

5.2.3.11. Preferred Symmetric Ciphers for v1 SEIPD

(array of one-octet values)

A series of symmetric cipher algorithm identifiers indicating how the
keyholder prefers to receive version 1 Symmetrically Encrypted
Integrity Protected Data (Section 5.14.1). The subpacket body is an
ordered list of octets with the most preferred listed first. It is
assumed that only algorithms listed are supported by the recipient’s
software. Algorithm numbers are in Section 9.3. This is only found
on a self-signature.

When generating a v2 SEIPD packet, this preference list is not
relevant. See Section 5.2.3.12 instead.

5.2.3.12. Preferred AEAD Ciphersuites

(array of pairs of octets indicating Symmetric Cipher and AEAD
algorithms)

A series of paired algorithm identifiers indicating how the keyholder
prefers to receive version 2 Symmetrically Encrypted Integrity
Protected Data (Section 5.14.2). Each pair of octets indicates a
combination of a symmetric cipher and an AEAD mode that the key
holder prefers to use. The symmetric cipher identifier precedes the
AEAD identifier in each pair. The subpacket body is an ordered list
of pairs of octets with the most preferred algorithm combination
listed first.

It is assumed that only the combinations of algorithms listed are
supported by the recipient’s software, with the exception of the
mandatory-to-implement combination of AES-128 and OCB. If AES-128
and OCB are not found in the subpacket, it is implicitly listed at
the end.

AEAD algorithm numbers are listed in Section 9.6. Symmetric cipher
algorithm numbers are listed in Section 9.3.

For example, a subpacket with content of these six octets:

09 02 09 03 13 02
Indicates that the keyholder prefers to receive v2 SEIPD using AES-256 with OCB, then AES-256 with GCM, then Camellia-256 with OCB, and finally the implicit AES-128 with OCB.

Note that support for version 2 of the Symmetrically Encrypted Integrity Protected Data packet (Section 5.14.2) in general is indicated by a Feature Flag (Section 5.2.3.29).

This subpacket is only found on a self-signature.

When generating a v1 SEIPD packet, this preference list is not relevant. See Section 5.2.3.11 instead.

5.2.3.13. Preferred Hash Algorithms

(array of one-octet values)

Message digest algorithm numbers that indicate which algorithms the key holder prefers to receive. Like the preferred AEAD ciphersuites, the list is ordered. Algorithm numbers are in Section 9.5. This is only found on a self-signature.

5.2.3.14. Preferred Compression Algorithms

(array of one-octet values)

Compression algorithm numbers that indicate which algorithms the key holder prefers to use. Like the preferred AEAD ciphersuites, the list is ordered. Algorithm numbers are in Section 9.4. A zero, or the absence of this subpacket, denotes that uncompressed data is preferred; the key holder’s software might have no compression software in that implementation. This is only found on a self-signature.

5.2.3.15. Signature Expiration Time

(4-octet time field)

The validity period of the signature. This is the number of seconds after the signature creation time that the signature expires. If this is not present or has a value of zero, it never expires.

5.2.3.16. Exportable Certification

(1 octet of exportability, 0 for not, 1 for exportable)
This subpacket denotes whether a certification signature is "exportable", to be used by other users than the signature's issuer. The packet body contains a Boolean flag indicating whether the signature is exportable. If this packet is not present, the certification is exportable; it is equivalent to a flag containing a 1.

Non-exportable, or "local", certifications are signatures made by a user to mark a key as valid within that user's implementation only.

Thus, when an implementation prepares a user's copy of a key for transport to another user (this is the process of "exporting" the key), any local certification signatures are deleted from the key.

The receiver of a transported key "imports" it, and likewise trims any local certifications. In normal operation, there won't be any, assuming the import is performed on an exported key. However, there are instances where this can reasonably happen. For example, if an implementation allows keys to be imported from a key database in addition to an exported key, then this situation can arise.

Some implementations do not represent the interest of a single user (for example, a key server). Such implementations always trim local certifications from any key they handle.

5.2.3.17. Revocable

(1 octet of revocability, 0 for not, 1 for revocable)

Signature's revocability status. The packet body contains a Boolean flag indicating whether the signature is revocable. Signatures that are not revocable have any later revocation signatures ignored. They represent a commitment by the signer that he cannot revoke his signature for the life of his key. If this packet is not present, the signature is revocable.

5.2.3.18. Trust Signature

(1 octet "level" (depth), 1 octet of trust amount)

Signer asserts that the key is not only valid but also trustworthy at the specified level. Level 0 has the same meaning as an ordinary validity signature. Level 1 means that the signed key is asserted to be a valid trusted introducer, with the 2nd octet of the body specifying the degree of trust. Level 2 means that the signed key is asserted to be trusted to issue level 1 trust signatures; that is, the signed key is a "meta introducer". Generally, a level n trust signature asserts that a key is trusted to issue level n-1 trust
signatures. The trust amount is in a range from 0-255, interpreted such that values less than 120 indicate partial trust and values of 120 or greater indicate complete trust. Implementations SHOULD emit values of 60 for partial trust and 120 for complete trust.

5.2.3.19. Regular Expression

(null-terminated regular expression)

Used in conjunction with trust Signature packets (of level > 0) to limit the scope of trust that is extended. Only signatures by the target key on User IDs that match the regular expression in the body of this packet have trust extended by the trust Signature subpacket. The regular expression uses the same syntax as the Henry Spencer’s "almost public domain" regular expression [REGEX] package. A description of the syntax is found in Section 8.

5.2.3.20. Revocation Key

(1 octet of class, 1 octet of public-key algorithm ID, 20 octets of V4 fingerprint)

This mechanism is deprecated. Applications MUST NOT generate such a subpacket.

An application that wants the functionality of delegating revocation SHOULD instead use an escrowed Revocation Signature. See Section 15.2 for more details.

The remainder of this section describes how some implementations attempt to interpret this deprecated subpacket.

This packet was intended to authorize the specified key to issue revocation signatures for this key. Class octet must have bit 0x80 set. If the bit 0x40 is set, then this means that the revocation information is sensitive. Other bits are for future expansion to other kinds of authorizations. This is only found on a direct-key self-signature (type 0x1f). The use on other types of self-signatures is unspecified.
If the "sensitive" flag is set, the keyholder feels this subpacket contains private trust information that describes a real-world sensitive relationship. If this flag is set, implementations SHOULD NOT export this signature to other users except in cases where the data needs to be available: when the signature is sent to the designated revoker, or when it is accompanied by a revocation signature from that revoker. Note that it may be appropriate to isolate this subpacket within a separate signature so that it is not combined with other subpackets that need to be exported.

5.2.3.21. Notation Data

(4 octets of flags, 2 octets of name length (M), 2 octets of value length (N), M octets of name data, N octets of value data)

This subpacket describes a "notation" on the signature that the issuer wishes to make. The notation has a name and a value, each of which are strings of octets. There may be more than one notation in a signature. Notations can be used for any extension the issuer of the signature cares to make. The "flags" field holds four octets of flags.

All undefined flags MUST be zero. Defined flags are as follows:

First octet:

<table>
<thead>
<tr>
<th>flag</th>
<th>shorthand</th>
<th>definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>0x80</td>
<td>human-readable</td>
<td>This note value is text.</td>
</tr>
</tbody>
</table>

Table 8: Notation flag registry (first octet)

Other octets: none.

Notation names are arbitrary strings encoded in UTF-8. They reside in two namespaces: The IETF namespace and the user namespace.

The IETF namespace is registered with IANA. These names MUST NOT contain the "@" character (0x40). This is a tag for the user namespace.

Names in the user namespace consist of a UTF-8 string tag followed by "@" followed by a DNS domain name. Note that the tag MUST NOT contain an "@" character. For example, the "sample" tag used by Example Corporation could be "sample@example.com".
Names in a user space are owned and controlled by the owners of that domain. Obviously, it’s bad form to create a new name in a DNS space that you don’t own.

Since the user namespace is in the form of an email address, implementers MAY wish to arrange for that address to reach a person who can be consulted about the use of the named tag. Note that due to UTF-8 encoding, not all valid user space name tags are valid email addresses.

If there is a critical notation, the criticality applies to that specific notation and not to notations in general.

5.2.3.22. Key Server Preferences

(N octets of flags)

This is a list of one-bit flags that indicate preferences that the key holder has about how the key is handled on a key server. All undefined flags MUST be zero.

First octet:

```
+======+===========+============================================+
| flag | shorthand | definition                                 |
+======+===========+============================================+
     | 0x80      | No-modify                                 |
     |          | The key holder requests that this key only |
     |          | be modified or updated by the key holder   |
     |          | or an administrator of the key server.     |
+======+===========+============================================+
```

Table 9: Key server preferences flag registry (first octet)

This is found only on a self-signature.

5.2.3.23. Preferred Key Server

(String)

This is a URI of a key server that the key holder prefers be used for updates. Note that keys with multiple User IDs can have a preferred key server for each User ID. Note also that since this is a URI, the key server can actually be a copy of the key retrieved by ftp, http, finger, etc.
5.2.3.24. Primary User ID

(1 octet, Boolean)

This is a flag in a User ID’s self-signature that states whether this User ID is the main User ID for this key. It is reasonable for an implementation to resolve ambiguities in preferences, etc. by referring to the primary User ID. If this flag is absent, its value is zero. If more than one User ID in a key is marked as primary, the implementation may resolve the ambiguity in any way it sees fit, but it is RECOMMENDED that priority be given to the User ID with the most recent self-signature.

When appearing on a self-signature on a User ID packet, this subpacket applies only to User ID packets. When appearing on a self-signature on a User Attribute packet, this subpacket applies only to User Attribute packets. That is to say, there are two different and independent "primaries" --- one for User IDs, and one for User Attributes.

5.2.3.25. Policy URI

(String)

This subpacket contains a URI of a document that describes the policy under which the signature was issued.

5.2.3.26. Key Flags

(N octets of flags)

This subpacket contains a list of binary flags that hold information about a key. It is a string of octets, and an implementation MUST NOT assume a fixed size. This is so it can grow over time. If a list is shorter than an implementation expects, the unstated flags are considered to be zero. The defined flags are as follows:

First octet:
<table>
<thead>
<tr>
<th>flag</th>
<th>definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>0x01</td>
<td>This key may be used to make User ID certifications (signature types 0x10-0x13) or direct key signatures (signature type 0x1F) over other keys.</td>
</tr>
<tr>
<td>0x02</td>
<td>This key may be used to sign data.</td>
</tr>
<tr>
<td>0x04</td>
<td>This key may be used to encrypt communications.</td>
</tr>
<tr>
<td>0x08</td>
<td>This key may be used to encrypt storage.</td>
</tr>
<tr>
<td>0x10</td>
<td>The private component of this key may have been split by a secret-sharing mechanism.</td>
</tr>
<tr>
<td>0x20</td>
<td>This key may be used for authentication.</td>
</tr>
<tr>
<td>0x80</td>
<td>The private component of this key may be in the possession of more than one person.</td>
</tr>
</tbody>
</table>

Table 10: Key flags registry (first octet)

Second octet:

<table>
<thead>
<tr>
<th>flag</th>
<th>definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>0x04</td>
<td>Reserved (ADSK).</td>
</tr>
<tr>
<td>0x08</td>
<td>Reserved (timestamping).</td>
</tr>
</tbody>
</table>

Table 11: Key flags registry (second octet)

Usage notes:

The flags in this packet may appear in self-signatures or in certification signatures. They mean different things depending on who is making the statement — for example, a certification signature that has the "sign data" flag is stating that the certification is for that use. On the other hand, the "communications encryption" flag in a self-signature is stating a preference that a given key be used for communications. Note however, that it is a thorny issue to determine what is "communications" and what is "storage". This decision is left wholly
up to the implementation; the authors of this document do not claim any special wisdom on the issue and realize that accepted opinion may change.

The "split key" (0x10) and "group key" (0x80) flags are placed on a self-signature only; they are meaningless on a certification signature. They SHOULD be placed only on a direct-key signature (type 0x1F) or a subkey signature (type 0x18), one that refers to the key the flag applies to.

5.2.3.27. Signer’s User ID

(String)

This subpacket allows a keyholder to state which User ID is responsible for the signing. Many keyholders use a single key for different purposes, such as business communications as well as personal communications. This subpacket allows such a keyholder to state which of their roles is making a signature.

This subpacket is not appropriate to use to refer to a User Attribute packet.

5.2.3.28. Reason for Revocation

(1 octet of revocation code, N octets of reason string)

This subpacket is used only in key revocation and certification revocation signatures. It describes the reason why the key or certificate was revoked.

The first octet contains a machine-readable code that denotes the reason for the revocation:
<table>
<thead>
<tr>
<th>Code</th>
<th>Reason</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>No reason specified (key revocations or cert revocations)</td>
</tr>
<tr>
<td>1</td>
<td>Key is superseded (key revocations)</td>
</tr>
<tr>
<td>2</td>
<td>Key material has been compromised (key revocations)</td>
</tr>
<tr>
<td>3</td>
<td>Key is retired and no longer used (key revocations)</td>
</tr>
<tr>
<td>32</td>
<td>User ID information is no longer valid (cert revocations)</td>
</tr>
<tr>
<td>100-110</td>
<td>Private Use</td>
</tr>
</tbody>
</table>

Table 12: Reasons for revocation

Following the revocation code is a string of octets that gives information about the Reason for Revocation in human-readable form (UTF-8). The string may be null (of zero length). The length of the subpacket is the length of the reason string plus one. An implementation SHOULD implement this subpacket, include it in all revocation signatures, and interpret revocations appropriately. There are important semantic differences between the reasons, and there are thus important reasons for revoking signatures.

If a key has been revoked because of a compromise, all signatures created by that key are suspect. However, if it was merely superseded or retired, old signatures are still valid. If the revoked signature is the self-signature for certifying a User ID, a revocation denotes that that user name is no longer in use. Such a revocation SHOULD include a 0x20 code.

Note that any signature may be revoked, including a certification on some other person’s key. There are many good reasons for revoking a certification signature, such as the case where the keyholder leaves the employ of a business with an email address. A revoked certification is no longer a part of validity calculations.
5.2.3.29. Features

(N octets of flags)

The Features subpacket denotes which advanced OpenPGP features a user’s implementation supports. This is so that as features are added to OpenPGP that cannot be backwards-compatible, a user can state that they can use that feature. The flags are single bits that indicate that a given feature is supported.

This subpacket is similar to a preferences subpacket, and only appears in a self-signature.

An implementation SHOULD NOT use a feature listed when sending to a user who does not state that they can use it.

Defined features are as follows:

First octet:

<table>
<thead>
<tr>
<th>Feature</th>
<th>Definition</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>0x01</td>
<td>Symmetrically Encrypted Integrity</td>
<td>Section 5.14.1</td>
</tr>
<tr>
<td></td>
<td>Protected Data packet version 1</td>
<td></td>
</tr>
<tr>
<td>0x02</td>
<td>Reserved</td>
<td></td>
</tr>
<tr>
<td>0x04</td>
<td>Reserved</td>
<td></td>
</tr>
<tr>
<td>0x08</td>
<td>Symmetrically Encrypted Integrity</td>
<td>Section 5.14.2</td>
</tr>
<tr>
<td></td>
<td>Protected Data packet version 2</td>
<td></td>
</tr>
</tbody>
</table>

Table 13: Features registry

If an implementation implements any of the defined features, it SHOULD implement the Features subpacket, too.

An implementation may freely infer features from other suitable implementation-dependent mechanisms.

See Section 15.1 for details about how to use the Features subpacket when generating encryption data.
5.2.3.30. Signature Target

(1 octet public-key algorithm, 1 octet hash algorithm, N octets hash)

This subpacket identifies a specific target signature to which a signature refers. For revocation signatures, this subpacket provides explicit designation of which signature is being revoked. For a third-party or timestamp signature, this designates what signature is signed. All arguments are an identifier of that target signature.

The N octets of hash data MUST be the size of the hash of the signature. For example, a target signature with a SHA-1 hash MUST have 20 octets of hash data.

5.2.3.31. Embedded Signature

(1 signature packet body)

This subpacket contains a complete Signature packet body as specified in Section 5.2. It is useful when one signature needs to refer to, or be incorporated in, another signature.

5.2.3.32. Issuer Fingerprint

(1 octet key version number, N octets of fingerprint)

The OpenPGP Key fingerprint of the key issuing the signature. This subpacket SHOULD be included in all signatures. If the version of the issuing key is 4 and an Issuer subpacket is also included in the signature, the key ID of the Issuer subpacket MUST match the low 64 bits of the fingerprint.

Note that the length N of the fingerprint for a version 4 key is 20 octets; for a version 5 key N is 32.

5.2.3.33. Intended Recipient Fingerprint

(1 octet key version number, N octets of fingerprint)

The OpenPGP Key fingerprint of the intended recipient primary key. If one or more subpackets of this type are included in a signature, it SHOULD be considered valid only in an encrypted context, where the key it was encrypted to is one of the indicated primary keys, or one of their subkeys. This can be used to prevent forwarding a signature outside of its intended, encrypted context.

Note that the length N of the fingerprint for a version 4 key is 20 octets; for a version 5 key N is 32.
5.2.4. Computing Signatures

All signatures are formed by producing a hash over the signature data, and then using the resulting hash in the signature algorithm.

When a V5 signature is made, the salt is hashed first.

For binary document signatures (type 0x00), the document data is hashed directly. For text document signatures (type 0x01), the document is canonicalized by converting line endings to <CR><LF>, and the resulting data is hashed.

When a V4 signature is made over a key, the hash data starts with the octet 0x99, followed by a two-octet length of the key, and then body of the key packet. When a V5 signature is made over a key, the hash data starts with the octet 0x9a, followed by a four-octet length of the key, and then body of the key packet.

A subkey binding signature (type 0x18) or primary key binding signature (type 0x19) then hashes the subkey using the same format as the main key (also using 0x99 or 0x9a as the first octet). Primary key revocation signatures (type 0x20) hash only the key being revoked. Subkey revocation signature (type 0x28) hash first the primary key and then the subkey being revoked.

A certification signature (type 0x10 through 0x13) hashes the User ID being bound to the key into the hash context after the above data. A V3 certification hashes the contents of the User ID or attribute packet packet, without any header. A V4 or V5 certification hashes the constant 0xB4 for User ID certifications or the constant 0xD1 for User Attribute certifications, followed by a four-octet number giving the length of the User ID or User Attribute data, and then the User ID or User Attribute data.

When a signature is made over a Signature packet (type 0x50, "Third-Party Confirmation signature"), the hash data starts with the octet 0x88, followed by the four-octet length of the signature, and then the body of the Signature packet. (Note that this is a Legacy packet header for a Signature packet with the length-of-length field set to zero.) The unhashed subpacket data of the Signature packet being hashed is not included in the hash, and the unhashed subpacket data length value is set to zero.

Once the data body is hashed, then a trailer is hashed. This trailer depends on the version of the signature.
* A V3 signature hashes five octets of the packet body, starting from the signature type field. This data is the signature type, followed by the four-octet signature time.

* A V4 or V5 signature hashes the packet body starting from its first field, the version number, through the end of the hashed subpacket data and a final extra trailer. Thus, the hashed fields are:
  - An octet indicating the signature version (0x04 for V4, 0x05 for V5),
  - the signature type,
  - the public-key algorithm,
  - the hash algorithm,
  - the hashed subpacket length,
  - the hashed subpacket body,
  - A second version octet (0x04 for V4, 0x05 for V5)
  - A single octet 0xFF,
  - A number representing the length of the hashed data from the Signature packet stopping right before the second version octet. For a V4 signature, this is a four-octet big-endian number, considered to be an unsigned integer modulo 2**32. For a V5 signature, this is an eight-octet big-endian number, considered to be an unsigned integer modulo 2**64.

After all this has been hashed in a single hash context, the resulting hash field is used in the signature algorithm and placed at the end of the Signature packet.

5.2.4.1. Subpacket Hints

It is certainly possible for a signature to contain conflicting information in subpackets. For example, a signature may contain multiple copies of a preference or multiple expiration times. In most cases, an implementation SHOULD use the last subpacket in the signature, but MAY use any conflict resolution scheme that makes more sense. Please note that we are intentionally leaving conflict resolution to the implementer; most conflicts are simply syntax errors, and the wishy-washy language here allows a receiver to be generous in what they accept, while putting pressure on a creator to
be stingy in what they generate.

Some apparent conflicts may actually make sense --- for example, suppose a keyholder has a V3 key and a V4 key that share the same RSA key material. Either of these keys can verify a signature created by the other, and it may be reasonable for a signature to contain an issuer subpacket for each key, as a way of explicitly tying those keys to the signature.

5.3. Symmetric-Key Encrypted Session Key Packets (Tag 3)

The Symmetric-Key Encrypted Session Key (SKESK) packet holds the symmetric-key encryption of a session key used to encrypt a message. Zero or more Public-Key Encrypted Session Key packets (Section 5.1) and/or Symmetric-Key Encrypted Session Key packets may precede a an encryption container (that is, a Symmetrically Encrypted Integrity Protected Data packet or --- for historic data --- a Symmetrically Encrypted Data packet) that holds an encrypted message. The message is encrypted with a session key, and the session key is itself encrypted and stored in the Encrypted Session Key packet(s).

If the encryption container is preceded by one or more Symmetric-Key Encrypted Session Key packets, each specifies a passphrase that may be used to decrypt the message. This allows a message to be encrypted to a number of public keys, and also to one or more passphrases.

The body of this packet starts with a one-octet number giving the version number of the packet type. The currently defined versions are 4 and 5. The remainder of the packet depends on the version.

The versions differ in how they encrypt the session key with the password, and in what they encode. The version of the SKESK packet must align with the version of the SEIPD packet (see Section 11.3.2.1).

5.3.1. v4 SKESK

A version 4 Symmetric-Key Encrypted Session Key (SKESK) packet precedes a version 1 Symmetrically Encrypted Integrity Protected Data (v1 SEIPD, see Section 5.14.1) packet. In historic data, it is sometimes found preceding a deprecated Symmetrically Encrypted Data packet (SED, see Section 5.8). A v4 SKESK packet MUST NOT precede a v2 SEIPD packet (see Section 11.3.2.1).

A version 4 Symmetric-Key Encrypted Session Key packet consists of:

* A one-octet version number with value 4.
* A one-octet number describing the symmetric algorithm used.

* A string-to-key (S2K) specifier. The length of the string-to-key specifier depends on its type (see Section 3.7.1).

* Optionally, the encrypted session key itself, which is decrypted with the string-to-key object.

If the encrypted session key is not present (which can be detected on the basis of packet length and S2K specifier size), then the S2K algorithm applied to the passphrase produces the session key for decrypting the message, using the symmetric cipher algorithm from the Symmetric-Key Encrypted Session Key packet.

If the encrypted session key is present, the result of applying the S2K algorithm to the passphrase is used to decrypt just that encrypted session key field, using CFB mode with an IV of all zeros. The decryption result consists of a one-octet algorithm identifier that specifies the symmetric-key encryption algorithm used to encrypt the following encryption container, followed by the session key octets themselves.

Note: because an all-zero IV is used for this decryption, the S2K specifier MUST use a salt value, either a Salted S2K, an Iterated-Salted S2K, or Argon2. The salt value will ensure that the decryption key is not repeated even if the passphrase is reused.

5.3.2. v5 SKESK

A version 5 Symmetric-Key Encrypted Session Key (SKESK) packet precedes a version 2 Symmetrically Encrypted Integrity Protected Data (v2 SEIPD, see Section 5.14.2) packet. A v5 SKESK packet MUST NOT precede a v1 SEIPD packet or a deprecated Symmetrically Encrypted Data packet (see Section 11.3.2.1).

A version 5 Symmetric-Key Encrypted Session Key packet consists of:

* A one-octet version number with value 5.

* A one-octet scalar octet count of the following 5 fields.

* A one-octet symmetric cipher algorithm identifier.

* A one-octet AEAD algorithm identifier.

* A one-octet scalar octet count of the following field.
* A string-to-key (S2K) specifier. The length of the string-to-key specifier depends on its type (see Section 3.7.1).

* A starting initialization vector of size specified by the AEAD algorithm.

* The encrypted session key itself.

* An authentication tag for the AEAD mode.

HKDF is used with SHA256 as hash algorithm, the key derived from S2K as Initial Keying Material (IKM), no salt, and the Packet Tag in new format encoding (bits 7 and 6 set, bits 5-0 carry the packet tag), the packet version, and the cipher-algo and AEAD-mode used to encrypt the key material, are used as info parameter. Then, the session key is encrypted using the resulting key, with the AEAD algorithm specified for version 2 of the Symmetrically Encrypted Integrity Protected Data packet. Note that no chunks are used and that there is only one authentication tag. The Packet Tag in OpenPGP format encoding (bits 7 and 6 set, bits 5-0 carry the packet tag), the packet version number, the cipher algorithm octet, and the AEAD algorithm octet are given as additional data. For example, the additional data used with AES-128 with OCB consists of the octets 0xC3, 0x05, 0x07, and 0x02.

5.4. One-Pass Signature Packets (Tag 4)

The One-Pass Signature packet precedes the signed data and contains enough information to allow the receiver to begin calculating any hashes needed to verify the signature. It allows the Signature packet to be placed at the end of the message, so that the signer can compute the entire signed message in one pass.

The body of this packet consists of:

* A one-octet version number. The currently defined versions are 3 and 5.

* A one-octet signature type. Signature types are described in Section 5.2.1.

* A one-octet number describing the hash algorithm used.

* A one-octet number describing the public-key algorithm used.

* Only for V5 packets, a 16 octet field containing random values used as salt. The value must match the salt field of the corresponding Signature packet.
* Only for V3 packets, an eight-octet number holding the Key ID of the signing key.

* Only for V5 packets, a one octet key version number and N octets of the fingerprint of the signing key. Note that the length N of the fingerprint for a version 5 key is 32.

* A one-octet number holding a flag showing whether the signature is nested. A zero value indicates that the next packet is another One-Pass Signature packet that describes another signature to be applied to the same message data.

When generating a one-pass signature, the OPS packet version MUST correspond to the version of the associated signature packet, except for the historical accident that v4 keys use a v3 one-pass signature packet (there is no v4 OPS):

<table>
<thead>
<tr>
<th>Signing key version</th>
<th>OPS packet version</th>
<th>Signature packet version</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>

Table 14: Versions of packets used in a one-pass signature

Note that if a message contains more than one one-pass signature, then the Signature packets bracket the message; that is, the first Signature packet after the message corresponds to the last one-pass packet and the final Signature packet corresponds to the first one-pass packet.

5.5. Key Material Packet

A key material packet contains all the information about a public or private key. There are four variants of this packet type, and two major versions. Consequently, this section is complex.

5.5.1. Key Packet Variants

5.5.1.1. Public-Key Packet (Tag 6)

A Public-Key packet starts a series of packets that forms an OpenPGP key (sometimes called an OpenPGP certificate).
5.5.1.2. Public-Subkey Packet (Tag 14)

A Public-Subkey packet (tag 14) has exactly the same format as a Public-Key packet, but denotes a subkey. One or more subkeys may be associated with a top-level key. By convention, the top-level key provides signature services, and the subkeys provide encryption services.

5.5.1.3. Secret-Key Packet (Tag 5)

A Secret-Key packet contains all the information that is found in a Public-Key packet, including the public-key material, but also includes the secret-key material after all the public-key fields.

5.5.1.4. Secret-Subkey Packet (Tag 7)

A Secret-Subkey packet (tag 7) is the subkey analog of the Secret Key packet and has exactly the same format.

5.5.2. Public-Key Packet Formats

There are three versions of key-material packets.

OpenPGP implementations SHOULD create keys with version 5 format. V4 keys are deprecated; an implementation SHOULD NOT generate a V4 key, but SHOULD accept it. V3 keys are deprecated; an implementation MUST NOT generate a V3 key, but MAY accept it. V2 keys are deprecated; an implementation MUST NOT generate a V2 key, but MAY accept it.

A version 3 public key or public-subkey packet contains:

* A one-octet version number (3).
* A four-octet number denoting the time that the key was created.
* A two-octet number denoting the time in days that this key is valid. If this number is zero, then it does not expire.
* A one-octet number denoting the public-key algorithm of this key.
* A series of multiprecision integers comprising the key material:
  - a multiprecision integer (MPI) of RSA public modulus n;
  - an MPI of RSA public encryption exponent e.
V3 keys are deprecated. They contain three weaknesses. First, it is relatively easy to construct a V3 key that has the same Key ID as any other key because the Key ID is simply the low 64 bits of the public modulus. Secondly, because the fingerprint of a V3 key hashes the key material, but not its length, there is an increased opportunity for fingerprint collisions. Third, there are weaknesses in the MD5 hash algorithm that make developers prefer other algorithms. See Section 12.2 for a fuller discussion of Key IDs and fingerprints.

V2 keys are identical to the deprecated V3 keys except for the version number.

The version 4 format is similar to the version 3 format except for the absence of a validity period. This has been moved to the Signature packet. In addition, fingerprints of version 4 keys are calculated differently from version 3 keys, as described in Section 12.

A version 4 packet contains:

* A one-octet version number (4).
* A four-octet number denoting the time that the key was created.
* A one-octet number denoting the public-key algorithm of this key.
* A series of values comprising the key material. This is algorithm-specific and described in Section 5.6.

The version 5 format is similar to the version 4 format except for the addition of a count for the key material. This count helps parsing secret key packets (which are an extension of the public key packet format) in the case of an unknown algorithm. In addition, fingerprints of version 5 keys are calculated differently from version 4 keys, as described in Section 12.

A version 5 packet contains:

* A one-octet version number (5).
* A four-octet number denoting the time that the key was created.
* A one-octet number denoting the public-key algorithm of this key.
* A four-octet scalar octet count for the following public key material.
* A series of values comprising the public key material. This is
  algorithm-specific and described in Section 5.6.

5.5.3. Secret-Key Packet Formats

The Secret-Key and Secret-Subkey packets contain all the data of the
Public-Key and Public-Subkey packets, with additional algorithm-
specific secret-key data appended, usually in encrypted form.

The packet contains:

* The fields of a Public-Key or Public-Subkey packet, as described
  above.

* One octet indicating string-to-key usage conventions. Zero
  indicates that the secret-key data is not encrypted. 255, 254, or
  253 indicates that a string-to-key specifier is being given. Any
  other value is a symmetric-key encryption algorithm identifier. A
  version 5 packet MUST NOT use the value 255.

* Only for a version 5 packet, a one-octet scalar octet count of the
  next 5 optional fields.

* [Optional] If string-to-key usage octet was 255, 254, or 253, a
  one-octet symmetric encryption algorithm.

* [Optional] If string-to-key usage octet was 253, a one-octet AEAD
  algorithm.

* [Optional] Only for a version 5 packet, and if string-to-key usage
  octet was 255, 254, or 253, an one-octet count of the following
  field.

* [Optional] If string-to-key usage octet was 255, 254, or 253, a
  string-to-key (S2K) specifier. The length of the string-to-key
  specifier depends on its type (see Section 3.7.1).

* [Optional] If string-to-key usage octet was 253 (that is, the
  secret data is AEAD-encrypted), an initialization vector (IV) of
  size specified by the AEAD algorithm (see Section 5.14.2), which
  is used as the nonce for the AEAD algorithm.

* [Optional] If string-to-key usage octet was 255, 254, or a cipher
  algorithm identifier (that is, the secret data is CFB-encrypted),
  an initialization vector (IV) of the same length as the cipher’s
  block size.
* Plain or encrypted multiprecision integers comprising the secret key data. This is algorithm-specific and described in Section 5.6. If the string-to-key usage octet is 253, then an AEAD authentication tag is part of that data. If the string-to-key usage octet is 254, a 20-octet SHA-1 hash of the plaintext of the algorithm-specific portion is appended to plaintext and encrypted with it. If the string-to-key usage octet is 255 or another nonzero value (that is, a symmetric-key encryption algorithm identifier), a two-octet checksum of the plaintext of the algorithm-specific portion (sum of all octets, mod 65536) is appended to plaintext and encrypted with it. (This is deprecated and SHOULD NOT be used, see below.)

* If the string-to-key usage octet is zero, then a two-octet checksum of the algorithm-specific portion (sum of all octets, mod 65536).

The details about storing algorithm-specific secrets above are summarized in Table 2.

Note that the version 5 packet format adds two count values to help parsing packets with unknown S2K or public key algorithms.

Secret MPI values can be encrypted using a passphrase. If a string-to-key specifier is given, that describes the algorithm for converting the passphrase to a key, else a simple MD5 hash of the passphrase is used. Implementations MUST use a string-to-key specifier; the simple hash is for backward compatibility and is deprecated, though implementations MAY continue to use existing private keys in the old format. The cipher for encrypting the MPIs is specified in the Secret-Key packet.

Encryption/decryption of the secret data is done using the key created from the passphrase and the initialization vector from the packet. If the string-to-key usage octet is not 253, CFB mode is used. A different mode is used with V3 keys (which are only RSA) than with other key formats. With V3 keys, the MPI bit count prefix (that is, the first two octets) is not encrypted. Only the MPI non-prefix data is encrypted. Furthermore, the CFB state is resynchronized at the beginning of each new MPI value, so that the CFB block boundary is aligned with the start of the MPI data.

With V4 and V5 keys, a simpler method is used. All secret MPI values are encrypted, including the MPI bitcount prefix.

If the string-to-key usage octet is 253, the key encryption key is derived using HKDF (see [RFC5869]) to provide key separation. HKDF is used with SHA256 as hash algorithm, the key derived from S2K as
Initial Keying Material (IKM), no salt, and the Packet Tag in OpenPGP format encoding (bits 7 and 6 set, bits 5-0 carry the packet tag), the packet version, and the cipher-algo and AEAD-mode used to encrypt the key material, are used as info parameter. Then, the encrypted MPI values are encrypted as one combined plaintext using one of the AEAD algorithms specified for version 2 of the Symmetrically Encrypted Integrity Protected Data packet. Note that no chunks are used and that there is only one authentication tag. As additional data, the Packet Tag in OpenPGP format encoding (bits 7 and 6 set, bits 5-0 carry the packet tag), followed by the public key packet fields, starting with the packet version number, are passed to the AEAD algorithm. For example, the additional data used with a Secret-Key Packet of version 4 consists of the octets 0xC5, 0x04, followed by four octets of creation time, one octet denoting the public-key algorithm, and the algorithm-specific public-key parameters. For a Secret-Subkey Packet, the first octet would be 0xC7. For a version 5 key packet, the second octet would be 0x05, and the four-octet octet count of the public key material would be included as well (see Section 5.5.2).

The two-octet checksum that follows the algorithm-specific portion is the algebraic sum, mod 65536, of the plaintext of all the algorithm-specific octets (including MPI prefix and data). With V3 keys, the checksum is stored in the clear. With V4 keys, the checksum is encrypted like the algorithm-specific data. This value is used to check that the passphrase was correct. However, this checksum is deprecated; an implementation SHOULD NOT use it, but should rather use the SHA-1 hash denoted with a usage octet of 254. The reason for this is that there are some attacks that involve undetectably modifying the secret key. If the string-to-key usage octet is 253 no checksum or SHA-1 hash is used but the authentication tag of the AEAD algorithm follows.

When decrypting the secret key material using any of these schemes (that is, where the usage octet is non-zero), the resulting cleartext octet stream MUST be well-formed. In particular, an implementation MUST NOT interpret octets beyond the unwrapped cleartext octet stream as part of any of the unwrapped MPI objects. Furthermore, an implementation MUST reject as unusable any secret key material whose cleartext length does not align with the lengths of the unwrapped MPI objects.

5.6. Algorithm-specific Parts of Keys

The public and secret key format specifies algorithm-specific parts of a key. The following sections describe them in detail.
5.6.1. Algorithm-Specific Part for RSA Keys

The public key is this series of multiprecision integers:
* MPI of RSA public modulus n;
* MPI of RSA public encryption exponent e.

The secret key is this series of multiprecision integers:
* MPI of RSA secret exponent d;
* MPI of RSA secret prime value p;
* MPI of RSA secret prime value q (p < q);
* MPI of u, the multiplicative inverse of p, mod q.

5.6.2. Algorithm-Specific Part for DSA Keys

The public key is this series of multiprecision integers:
* MPI of DSA prime p;
* MPI of DSA group order q (q is a prime divisor of p-1);
* MPI of DSA group generator g;
* MPI of DSA public-key value y (= g**x mod p where x is secret).

The secret key is this single multiprecision integer:
* MPI of DSA secret exponent x.

5.6.3. Algorithm-Specific Part for Elgamal Keys

The public key is this series of multiprecision integers:
* MPI of Elgamal prime p;
* MPI of Elgamal group generator g;
* MPI of Elgamal public key value y (= g**x mod p where x is secret).

The secret key is this single multiprecision integer:
* MPI of Elgamal secret exponent x.
5.6.4. Algorithm-Specific Part for ECDSA Keys

The public key is this series of values:
* A variable-length field containing a curve OID, which is formatted as follows:
  - A one-octet size of the following field; values 0 and 0xFF are reserved for future extensions,
  - The octets representing a curve OID (defined in Section 9.2);
* MPI of an EC point representing a public key.

The secret key is this single multiprecision integer:
* MPI of an integer representing the secret key, which is a scalar of the public EC point.

5.6.5. Algorithm-Specific Part for EdDSA Keys

The public key is this series of values:
* A variable-length field containing a curve OID, formatted as follows:
  - A one-octet size of the following field; values 0 and 0xFF are reserved for future extensions,
  - The octets representing a curve OID, defined in Section 9.2;
* An MPI of an EC point representing a public key Q in prefixed native form (see Section 13.2.2).

The secret key is this single multiprecision integer:
* An MPI-encoded octet string representing the native form of the secret key, in the curve-specific format described in Section 9.2.1.

See [RFC8032] for more details about the native octet strings.

5.6.6. Algorithm-Specific Part for ECDH Keys

The public key is this series of values:
* A variable-length field containing a curve OID, which is formatted as follows:
- A one-octet size of the following field; values 0 and 0xFF are reserved for future extensions,
- Octets representing a curve OID, defined in Section 9.2;

* MPI of an EC point representing a public key, in the point format associated with the curve as specified in Section 9.2.1

* A variable-length field containing KDF parameters, which is formatted as follows:
  - A one-octet size of the following fields; values 0 and 0xFF are reserved for future extensions,
  - A one-octet value 1, reserved for future extensions,
  - A one-octet hash function ID used with a KDF,
  - A one-octet algorithm ID for the symmetric algorithm used to wrap the symmetric key used for the message encryption; see Section 13.5 for details.

Observe that an ECDH public key is composed of the same sequence of fields that define an ECDSA key plus the KDF parameters field.

The secret key is this single multiprecision integer:

* An MPI representing the secret key, in the curve-specific format described in Section 9.2.1.

5.6.6.1. ECDH Secret Key Material

When curve P-256, P-384, or P-521 are used in ECDH, their secret keys are represented as a simple integer in standard MPI form. Other curves are presented on the wire differently (though still as a single MPI), as described below and in Section 9.2.1.

5.6.6.1.1. Curve25519 ECDH Secret Key Material

A Curve25519 secret key is stored as a standard integer in big-endian MPI form. Note that this form is in reverse octet order from the little-endian "native" form found in [RFC7748].
5.6.6.1.2.  X448 ECDH Secret Key Material

An X448 secret key is contained within its MPI as a prefixed octet string (see Section 13.3.2), which encapsulates the native secret key format found in [RFC7748]. The full wire format (as an MPI) will thus be the three octets 01 c7 40 followed by the full 56 octet native secret key.

When generating a new X448 secret key from 56 fully-random octets, the following pseudocode produces the MPI wire format:

```python
def X448_MPI_from_random(octet_list):
    prefixed_header = [0x01, 0xc7, 0x40]
    return prefixed_header ++ octet_list
```

5.7.  Compressed Data Packet (Tag 8)

The Compressed Data packet contains compressed data. Typically, this packet is found as the contents of an encrypted packet, or following a Signature or One-Pass Signature packet, and contains a literal data packet.

The body of this packet consists of:

* One octet that gives the algorithm used to compress the packet.

* Compressed data, which makes up the remainder of the packet.
A Compressed Data Packet’s body contains an block that compresses some set of packets. See Section 11 for details on how messages are formed.

ZIP-compressed packets are compressed with raw [RFC1951] DEFLATE blocks.

ZLIB-compressed packets are compressed with [RFC1950] ZLIB-style blocks.

BZip2-compressed packets are compressed using the BZip2 [BZ2] algorithm.

An implementation that generates a Compressed Data packet MUST use the non-legacy format for packet framing (see Section 4.2.1). It MUST NOT generate a Compressed Data packet with Legacy format (Section 4.2.2)

An implementation that deals with either historic data or data generated by legacy implementations MAY interpret Compressed Data packets that use the Legacy format for packet framing.

5.8. Symmetrically Encrypted Data Packet (Tag 9)

The Symmetrically Encrypted Data packet contains data encrypted with a symmetric-key algorithm. When it has been decrypted, it contains other packets (usually a literal data packet or compressed data packet, but in theory other Symmetrically Encrypted Data packets or sequences of packets that form whole OpenPGP messages).

This packet is obsolete. An implementation MUST NOT create this packet. An implementation MAY process such a packet but it MUST return a clear diagnostic that a non-integrity protected packet has been processed. The implementation SHOULD also return an error in this case and stop processing.

This packet format is impossible to handle safely in general because the ciphertext it provides is malleable. See Section 15.1 about selecting a better OpenPGP encryption container that does not have this flaw.

The body of this packet consists of:

* Encrypted data, the output of the selected symmetric-key cipher operating in OpenPGP’s variant of Cipher Feedback (CFB) mode.
The symmetric cipher used may be specified in a Public-Key or Symmetric-Key Encrypted Session Key packet that precedes the Symmetrically Encrypted Data packet. In that case, the cipher algorithm octet is prefixed to the session key before it is encrypted. If no packets of these types precede the encrypted data, the IDEA algorithm is used with the session key calculated as the MD5 hash of the passphrase, though this use is deprecated.

The data is encrypted in CFB mode, with a CFB shift size equal to the cipher’s block size. The Initial Vector (IV) is specified as all zeros. Instead of using an IV, OpenPGP prefixes a string of length equal to the block size of the cipher plus two to the data before it is encrypted. The first block-size octets (for example, 8 octets for a 64-bit block length) are random, and the following two octets are copies of the last two octets of the IV. For example, in an 8-octet block, octet 9 is a repeat of octet 7, and octet 10 is a repeat of octet 8. In a cipher of length 16, octet 17 is a repeat of octet 15 and octet 18 is a repeat of octet 16. As a pedantic clarification, in both these examples, we consider the first octet to be numbered 1.

After encrypting the first block-size-plus-two octets, the CFB state is resynchronized. The last block-size octets of ciphertext are passed through the cipher and the block boundary is reset.

The repetition of 16 bits in the random data prefixed to the message allows the receiver to immediately check whether the session key is incorrect. See Section 15 for hints on the proper use of this "quick check".

5.9. Marker Packet (Tag 10)

The body of this packet consists of:

* The three octets 0x50, 0x47, 0x50 (which spell "PGP" in UTF-8).

Such a packet MUST be ignored when received.

5.10. Literal Data Packet (Tag 11)

A Literal Data packet contains the body of a message; data that is not to be further interpreted.

The body of this packet consists of:

* A one-octet field that describes how the data is formatted.
If it is a b (0x62), then the Literal packet contains binary data. If it is a u (0x75), then the Literal packet contains UTF-8-encoded text data, and thus may need line ends converted to local form, or other text mode changes.

Older versions of OpenPGP used t (0x74) to indicate textual data, but did not specify the character encoding. Implementations SHOULD NOT emit this value. An implementation that receives a literal data packet with this value in the format field SHOULD interpret the packet data as UTF-8 encoded text, unless reliable (not attacker-controlled) context indicates a specific alternate text encoding. This mode is deprecated due to its ambiguity.

Early versions of PGP also defined a value of l as a ‘local’ mode for machine-local conversions. [RFC1991] incorrectly stated this local mode flag as 1 (ASCII numeral one). Both of these local modes are deprecated.

* File name as a string (one-octet length, followed by a file name). This may be a zero-length string. Commonly, if the source of the encrypted data is a file, this will be the name of the encrypted file. An implementation MAY consider the file name in the Literal packet to be a more authoritative name than the actual file name.

* A four-octet number that indicates a date associated with the literal data. Commonly, the date might be the modification date of a file, or the time the packet was created, or a zero that indicates no specific time.

* The remainder of the packet is literal data.

Text data is stored with <CR><LF> text endings (that is, network-normal line endings). These should be converted to native line endings by the receiving software.

Note that OpenPGP signatures do not include the formatting octet, the file name, and the date field of the literal packet in a signature hash and thus those fields are not protected against tampering in a signed document. A receiving implementation MUST NOT treat those fields as though they were cryptographically secured by the surrounding signature either when representing them to the user or acting on them.

Due to their inherent malleability, an implementation that generates a literal data packet SHOULD avoid storing any significant data in these fields. If the implementation is certain that the data is textual and is encoded with UTF-8 (for example, if it will follow this literal data packet with a signature packet of type 0x01 (see
Section 5.2.1), it MAY set the format octet to u. Otherwise, it
SHOULD set the format octet to b. It SHOULD set the filename to the
empty string (encoded as a single zero octet), and the timestamp to
zero (encoded as four zero octets).

An application that wishes to include such filesystem metadata within
a signature is advised to sign an encapsulated archive (for example,
[PAX]).

An implementation that generates a Literal Data packet MUST use the
OpenPGP format for packet framing (see Section 4.2.1). It MUST NOT
generate a Literal Data packet with Legacy format (Section 4.2.2)

An implementation that deals with either historic data or data
generated by legacy implementations MAY interpret Literal Data
packets that use the Legacy format for packet framing.

5.10.1. Special Filename _CONSOLE (Deprecated)

The Literal Data packet’s filename field has a historical special
case for the special name _CONSOLE. When the filename field is
_CONSOLE, the message is considered to be "for your eyes only". This
advises that the message data is unusually sensitive, and the
receiving program should process it more carefully, perhaps avoiding
storing the received data to disk, for example.

An OpenPGP deployment that generates literal data packets MUST NOT
depend on this indicator being honored in any particular way. It
cannot be enforced, and the field itself is not covered by any
cryptographic signature.

It is NOT RECOMMENDED to use this special filename in a newly-
generated literal data packet.

5.11. Trust Packet (Tag 12)

The Trust packet is used only within keyrings and is not normally
exported. Trust packets contain data that record the user’s
specifications of which key holders are trustworthy introducers,
along with other information that implementing software uses for
trust information. The format of Trust packets is defined by a given
implementation.

Trust packets SHOULD NOT be emitted to output streams that are
transferred to other users, and they SHOULD be ignored on any input
other than local keyring files.
5.12. User ID Packet (Tag 13)

A User ID packet consists of UTF-8 text that is intended to represent
the name and email address of the key holder. By convention, it
includes an [RFC2822] mail name-addr, but there are no restrictions
on its content. The packet length in the header specifies the length
of the User ID.

5.13. User Attribute Packet (Tag 17)

The User Attribute packet is a variation of the User ID packet. It
is capable of storing more types of data than the User ID packet,
which is limited to text. Like the User ID packet, a User Attribute
packet may be certified by the key owner ("self-signed") or any other
key owner who cares to certify it. Except as noted, a User Attribute
packet may be used anywhere that a User ID packet may be used.

While User Attribute packets are not a required part of the OpenPGP
standard, implementations SHOULD provide at least enough
compatibility to properly handle a certification signature on the
User Attribute packet. A simple way to do this is by treating the
User Attribute packet as a User ID packet with opaque contents, but
an implementation may use any method desired.

The User Attribute packet is made up of one or more attribute
subpackets. Each subpacket consists of a subpacket header and a
body. The header consists of:

* the subpacket length (1, 2, or 5 octets)
* the subpacket type (1 octet)

and is followed by the subpacket specific data.

The following table lists the currently known subpackets:

<table>
<thead>
<tr>
<th>Type</th>
<th>Attribute Subpacket</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Image Attribute Subpacket</td>
</tr>
<tr>
<td>100-110</td>
<td>Private/Experimental Use</td>
</tr>
</tbody>
</table>

Table 15: User Attribute type registry

An implementation SHOULD ignore any subpacket of a type that it does
not recognize.
5.13.1. The Image Attribute Subpacket

The Image Attribute subpacket is used to encode an image, presumably (but not required to be) that of the key owner.

The Image Attribute subpacket begins with an image header. The first two octets of the image header contain the length of the image header. Note that unlike other multi-octet numerical values in this document, due to a historical accident this value is encoded as a little-endian number. The image header length is followed by a single octet for the image header version. The only currently defined version of the image header is 1, which is a 16-octet image header. The first three octets of a version 1 image header are thus 0x10, 0x00, 0x01.

The fourth octet of a version 1 image header designates the encoding format of the image. The only currently defined encoding format is the value 1 to indicate JPEG. Image format types 100 through 110 are reserved for private or experimental use. The rest of the version 1 image header is made up of 12 reserved octets, all of which MUST be set to 0.

The rest of the image subpacket contains the image itself. As the only currently defined image type is JPEG, the image is encoded in the JPEG File Interchange Format (JFIF), a standard file format for JPEG images [JFIF].

An implementation MAY try to determine the type of an image by examination of the image data if it is unable to handle a particular version of the image header or if a specified encoding format value is not recognized.

5.14. Sym. Encrypted Integrity Protected Data Packet (Tag 18)

This packet contains integrity protected and encrypted data. When it has been decrypted, it will contain other packets forming an OpenPGP Message (see Section 11.3).

The first octet of this packet is always used to indicate the version number, but different versions contain differently-structured ciphertext. Version 1 of this packet contains data encrypted with a symmetric-key algorithm and protected against modification by the SHA-1 hash algorithm. This is a legacy OpenPGP mechanism that offers some protections against ciphertext malleability.
Version 2 of this packet contains data encrypted with an authenticated encryption and additional data (AEAD) construction. This offers a more cryptographically rigorous defense against ciphertext malleability, but may not be as widely supported yet. See Section 15.1 for more details on choosing between these formats.


A version 1 Symmetrically Encrypted Integrity Protected Data packet consists of:

* A one-octet version number with value 1.

* Encrypted data, the output of the selected symmetric-key cipher operating in Cipher Feedback mode with shift amount equal to the block size of the cipher (CFB-n where n is the block size).

The symmetric cipher used MUST be specified in a Public-Key or Symmetric-Key Encrypted Session Key packet that precedes the Symmetrically Encrypted Integrity Protected Data packet. In either case, the cipher algorithm octet is prefixed to the session key before it is encrypted.

The data is encrypted in CFB mode, with a CFB shift size equal to the cipher’s block size. The Initial Vector (IV) is specified as all zeros. Instead of using an IV, OpenPGP prefixes an octet string to the data before it is encrypted. The length of the octet string equals the block size of the cipher in octets, plus two. The first octets in the group, of length equal to the block size of the cipher, are random; the last two octets are each copies of their 2nd preceding octet. For example, with a cipher whose block size is 128 bits or 16 octets, the prefix data will contain 16 random octets, then two more octets, which are copies of the 15th and 16th octets, respectively. Unlike the Symmetrically Encrypted Data Packet, no special CFB resynchronization is done after encrypting this prefix data. See Section 14.9 for more details.

The repetition of 16 bits in the random data prefixed to the message allows the receiver to immediately check whether the session key is incorrect.
Two constant octets with the values 0xD3 and 0x14 are appended to the plaintext. Then, the plaintext of the data to be encrypted is passed through the SHA-1 hash function. The input to the hash function includes the prefix data described above; it includes all of the plaintext, including the trailing constant octets 0xD3, 0x14. The 20 octets of the SHA-1 hash are then appended to the plaintext (after the constant octets 0xD3, 0x14) and encrypted along with the plaintext using the same CFB context. This trailing checksum is known as the Modification Detection Code (MDC).

During decryption, the plaintext data should be hashed with SHA-1, including the prefix data as well as the trailing constant octets 0xD3, 0x14, but excluding the last 20 octets containing the SHA-1 hash. The computed SHA-1 hash is then compared with the last 20 octets of plaintext. A mismatch of the hash indicates that the message has been modified and MUST be treated as a security problem. Any failure SHOULD be reported to the user.

NON-NORMATIVE EXPLANATION

The Modification Detection Code (MDC) system, as the integrity protection mechanism of version 1 of the Symmetrically Encrypted Integrity Protected Data packet is called, was created to provide an integrity mechanism that is less strong than a signature, yet stronger than bare CFB encryption.

It is a limitation of CFB encryption that damage to the ciphertext will corrupt the affected cipher blocks and the block following. Additionally, if data is removed from the end of a CFB-encrypted block, that removal is undetectable. (Note also that CBC mode has a similar limitation, but data removed from the front of the block is undetectable.)

The obvious way to protect or authenticate an encrypted block is to digitally sign it. However, many people do not wish to habitually sign data, for a large number of reasons beyond the scope of this document. Suffice it to say that many people consider properties such as deniability to be as valuable as integrity.

OpenPGP addresses this desire to have more security than raw encryption and yet preserve deniability with the MDC system. An MDC is intentionally not a MAC. Its name was not selected by accident. It is analogous to a checksum.
Despite the fact that it is a relatively modest system, it has proved itself in the real world. It is an effective defense to several attacks that have surfaced since it has been created. It has met its modest goals admirably.

Consequently, because it is a modest security system, it has modest requirements on the hash function(s) it employs. It does not rely on a hash function being collision-free, it relies on a hash function being one-way. If a forger, Frank, wishes to send Alice a (digitally) unsigned message that says, "I’ve always secretly loved you, signed Bob", it is far easier for him to construct a new message than it is to modify anything intercepted from Bob. (Note also that if Bob wishes to communicate secretly with Alice, but without authentication or identification and with a threat model that includes forgers, he has a problem that transcends mere cryptography.)

Note also that unlike nearly every other OpenPGP subsystem, there are no parameters in the MDC system. It hard-defines SHA-1 as its hash function. This is not an accident. It is an intentional choice to avoid downgrade and cross-grade attacks while making a simple, fast system. (A downgrade attack would be an attack that replaced SHA2-256 with SHA-1, for example. A cross-grade attack would replace SHA-1 with another 160-bit hash, such as RIPE-MD/160, for example.)

However, no update will be needed because the MDC has been replaced by the AEAD encryption described in this document.


A version 2 Symmetrically Encrypted Integrity Protected Data packet consists of:

* A one-octet version number with value 2.
* A one-octet cipher algorithm.
* A one-octet AEAD algorithm.
* A one-octet chunk size.
* Thirty-two octets of salt. The salt is used to derive the message key and must be unique.
* Encrypted data, the output of the selected symmetric-key cipher operating in the given AEAD mode.
* A final, summary authentication tag for the AEAD mode.

The decrypted session key and the salt are used to derive an M-bit message key and N-64 bits used as initialization vector, where M is the key size of the symmetric algorithm and N is the nonce size of the AEAD algorithm. M + N - 64 bits are derived using HKDF (see [RFC5869]). The left-most M bits are used as symmetric algorithm key, the remaining N - 64 bits are used as initialization vector. HKDF is used with SHA256 as hash algorithm, the session key as Initial Keying Material (IKM), the salt as salt, and the Packet Tag in OpenPGP format encoding (bits 7 and 6 set, bits 5-0 carry the packet tag), version number, cipher algorithm octet, AEAD algorithm octet, and chunk size octet as info parameter.

The KDF mechanism provides key separation between cipher and AEAD algorithms. Furthermore, an implementation can securely reply to a message even if a recipients certificate is unknown by reusing the encrypted session key packets and replying with a different salt yielding a new, unique message key.

A v2 SEIPD packet consists of one or more chunks of data. The plaintext of each chunk is of a size specified using the chunk size octet using the method specified below.

The encrypted data consists of the encryption of each chunk of plaintext, followed immediately by the relevant authentication tag. If the last chunk of plaintext is smaller than the chunk size, the ciphertext for that data may be shorter; it is nevertheless followed by a full authentication tag.

For each chunk, the AEAD construction is given the Packet Tag in OpenPGP format encoding (bits 7 and 6 set, bits 5-0 carry the packet tag), version number, cipher algorithm octet, AEAD algorithm octet, and chunk size octet as additional data. For example, the additional data of the first chunk using EAX and AES-128 with a chunk size of $2^{16}$ octets consists of the octets 0xD2, 0x02, 0x01, and 0x10.

After the final chunk, the AEAD algorithm is used to produce a final authentication tag encrypting the empty string. This AEAD instance is given the additional data specified above, plus an eight-octet, big-endian value specifying the total number of plaintext octets encrypted. This allows detection of a truncated ciphertext.

The chunk size octet specifies the size of chunks using the following formula (in C), where c is the chunk size octet:

\[
\text{chunk\_size} = ((\text{uint64\_t})1 << (c + 6))
\]
An implementation MUST accept chunk size octets with values from 0 to 16. An implementation MUST NOT create data with a chunk size octet value larger than 16 (4 MiB chunks).

The nonce for AEAD mode consists of two parts. Let N be the size of the nonce. The left-most N - 64 bits are the initialization vector derived using HKDF. The right-most 64 bits are the chunk index as big-endian value. The index of the first chunk is zero.

5.14.3. EAX Mode

The EAX AEAD Algorithm used in this document is defined in [EAX].

The EAX algorithm can only use block ciphers with 16-octet blocks. The nonce is 16 octets long. EAX authentication tags are 16 octets long.

5.14.4. OCB Mode

The OCB AEAD Algorithm used in this document is defined in [RFC7253].

The OCB algorithm can only use block ciphers with 16-octet blocks. The nonce is 15 octets long. OCB authentication tags are 16 octets long.

5.14.5. GCM Mode

The GCM AEAD Algorithm used in this document is defined in [SP800-38D].

The GCM algorithm can only use block ciphers with 16-octet blocks. The nonce is 12 octets long. GCM authentication tags are 16 octets long.

5.15. Padding Packet (Tag 21)

The Padding packet contains random data, and can be used to defend against traffic analysis (see Section 15.4) on version 2 SEIPD messages (see Section 5.14.2) and Transferable Public Keys (see Section 11.1).

Such a packet MUST be ignored when received.

Its contents SHOULD be random octets to make the length obfuscation it provides more robust even when compressed.

An implementation adding padding to an OpenPGP stream SHOULD place such a packet:
At the end of a v5 Transferable Public Key that is transferred over an encrypted channel (see Section 11.1).

As the last packet of an Optionally Padded Message within a version 2 Symmetrically Encrypted Integrity Protected Data Packet (see Section 11.3.1).

An implementation MUST be able to process padding packets anywhere else in an OpenPGP stream, so that future revisions of this document may specify further locations for padding.

Policy about how large to make such a packet to defend against traffic analysis is beyond the scope of this document.

6. Radix-64 Conversions

As stated in the introduction, OpenPGP’s underlying native representation for objects is a stream of arbitrary octets, and some systems desire these objects to be immune to damage caused by character set translation, data conversions, etc.

In principle, any printable encoding scheme that met the requirements of the unsafe channel would suffice, since it would not change the underlying binary bit streams of the native OpenPGP data structures. The OpenPGP standard specifies one such printable encoding scheme to ensure interoperability.

OpenPGP’s Radix-64 encoding is composed of two parts: a base64 encoding of the binary data and an optional checksum. The base64 encoding is identical to the MIME base64 content-transfer-encoding [RFC2045].

The optional checksum is a 24-bit Cyclic Redundancy Check (CRC) converted to four characters of radix-64 encoding by the same MIME base64 transformation, preceded by an equal sign (=). The CRC is computed by using the generator 0x864CFB and an initialization of 0xB704CE. The accumulation is done on the data before it is converted to radix-64, rather than on the converted data. A sample implementation of this algorithm is in Section 6.1.

If present, the checksum with its leading equal sign MUST appear on the next line after the base64 encoded data.

Rationale for CRC-24: The size of 24 bits fits evenly into printable base64. The nonzero initialization can detect more errors than a zero initialization.
6.1. An Implementation of the CRC-24 in "C"

```c
#define CRC24_INIT 0xB704CEL
#define CRC24_GENERATOR 0x864CFBL

typedef unsigned long crc24;
crc24 crc_octets(unsigned char *octets, size_t len)
{
    crc24 crc = CRC24_INIT;
    int i;
    while (len--) {
        crc ^= (*octets++) << 16;
        for (i = 0; i < 8; i++) {
            crc <<= 1;
            if (crc & 0x1000000) {
                crc &= 0xffffff; /* Clear bit 25 to avoid overflow */
                crc ^= CRC24_GENERATOR;
            }
        }
    }
    return crc & 0xFFFFFFL;
}
```

6.2. Forming ASCII Armor

When OpenPGP encodes data into ASCII Armor, it puts specific headers around the Radix-64 encoded data, so OpenPGP can reconstruct the data later. An OpenPGP implementation MAY use ASCII armor to protect raw binary data. OpenPGP informs the user what kind of data is encoded in the ASCII armor through the use of the headers.

Concatenating the following data creates ASCII Armor:

* An Armor Header Line, appropriate for the type of data
* Armor Headers
* A blank (zero-length, or containing only whitespace) line
* The ASCII-Armored data
* An Armor Checksum
* The Armor Tail, which depends on the Armor Header Line
An Armor Header Line consists of the appropriate header line text surrounded by five (5) dashes (−, 0x2D) on either side of the header line text. The header line text is chosen based upon the type of data that is being encoded in Armor, and how it is being encoded. Header line texts include the following strings:

BEGIN PGP MESSAGE
   Used for signed, encrypted, or compressed files.

BEGIN PGP PUBLIC KEY BLOCK
   Used for armoring public keys.

BEGIN PGP PRIVATE KEY BLOCK
   Used for armoring private keys.

BEGIN PGP MESSAGE, PART X/Y
   Used for multi-part messages, where the armor is split amongst Y parts, and this is the Xth part out of Y.

BEGIN PGP MESSAGE, PART X
   Used for multi-part messages, where this is the Xth part of an unspecified number of parts. Requires the MESSAGE-ID Armor Header to be used.

BEGIN PGP SIGNATURE
   Used for detached signatures, OpenPGP/MIME signatures, and cleartext signatures.

Note that all these Armor Header Lines are to consist of a complete line. That is to say, there is always a line ending preceding the starting five dashes, and following the ending five dashes. The header lines, therefore, MUST start at the beginning of a line, and MUST NOT have text other than whitespace following them on the same line. These line endings are considered a part of the Armor Header Line for the purposes of determining the content they delimit. This is particularly important when computing a cleartext signature (see Section 7).

The Armor Headers are pairs of strings that can give the user or the receiving OpenPGP implementation some information about how to decode or use the message. The Armor Headers are a part of the armor, not a part of the message, and hence are not protected by any signatures applied to the message.
The format of an Armor Header is that of a key-value pair. A colon (: 0x38) and a single space (0x20) separate the key and value. OpenPGP should consider improperly formatted Armor Headers to be corruption of the ASCII Armor. Unknown keys should be reported to the user, but OpenPGP should continue to process the message.

Note that some transport methods are sensitive to line length. While there is a limit of 76 characters for the Radix-64 data (Section 6.3), there is no limit to the length of Armor Headers. Care should be taken that the Armor Headers are short enough to survive transport. One way to do this is to repeat an Armor Header Key multiple times with different values for each so that no one line is overly long.

Currently defined Armor Header Keys are as follows:

* "Version", which states the OpenPGP implementation and version used to encode the message. To minimize metadata, implementations SHOULD NOT emit this key and its corresponding value except for debugging purposes with explicit user consent.

* "Comment", a user-defined comment. OpenPGP defines all text to be in UTF-8. A comment may be any UTF-8 string. However, the whole point of armoring is to provide seven-bit-clean data. Consequently, if a comment has characters that are outside the US-ASCII range of UTF, they may very well not survive transport.

* "MessageID", a 32-character string of printable characters. The string must be the same for all parts of a multi-part message that uses the "PART X" Armor Header. MessageID strings should be unique enough that the recipient of the mail can associate all the parts of a message with each other. A good checksum or cryptographic hash function is sufficient.

The MessageID SHOULD NOT appear unless it is in a multi-part message. If it appears at all, it MUST be computed from the finished (encrypted, signed, etc.) message in a deterministic fashion, rather than contain a purely random value. This is to allow the legitimate recipient to determine that the MessageID cannot serve as a covert means of leaking cryptographic key information.

* "Hash", a comma-separated list of hash algorithms used in this message. This is used only in cleartext signed messages.

* "SaltedHash", a salt and hash algorithm used in this message. This is used only in cleartext signed messages that are followed by a v5 Signature.
* "Charset", a description of the character set that the plaintext is in. Please note that OpenPGP defines text to be in UTF-8. An implementation will get best results by translating into and out of UTF-8. However, there are many instances where this is easier said than done. Also, there are communities of users who have no need for UTF-8 because they are all happy with a character set like ISO Latin-5 or a Japanese character set. In such instances, an implementation MAY override the UTF-8 default by using this header key. An implementation MAY implement this key and any translations it cares to; an implementation MAY ignore it and assume all text is UTF-8.

The Armor Tail Line is composed in the same manner as the Armor Header Line, except the string "BEGIN" is replaced by the string "END".

6.3. Encoding Binary in Radix-64

The encoding process represents 24-bit groups of input bits as output strings of 4 encoded characters. Proceeding from left to right, a 24-bit input group is formed by concatenating three 8-bit input groups. These 24 bits are then treated as four concatenated 6-bit groups, each of which is translated into a single digit in the Radix-64 alphabet. When encoding a bit stream with the Radix-64 encoding, the bit stream must be presumed to be ordered with the most significant bit first. That is, the first bit in the stream will be the high-order bit in the first 8-bit octet, and the eighth bit will be the low-order bit in the first 8-bit octet, and so on.

```
first octet second octet third octet
7 6 5 4 3 2 1 07 6 5 4 3 2 1 07 6 5 4 3 2 1 0
5 4 3 2 1 05 4 3 2 1 05 4 3 2 1 05 4 3 2 1 0
1.index 2.index 3.index 4.index
```

Each 6-bit group is used as an index into an array of 64 printable characters from the table below. The character referenced by the index is placed in the output string.
Table 16: Encoding for Radix-64

<table>
<thead>
<tr>
<th>Value</th>
<th>Encoding</th>
<th>Value</th>
<th>Encoding</th>
<th>Value</th>
<th>Encoding</th>
<th>Value</th>
<th>Encoding</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>A</td>
<td>17</td>
<td>R</td>
<td>34</td>
<td>i</td>
<td>51</td>
<td>z</td>
</tr>
<tr>
<td>1</td>
<td>B</td>
<td>18</td>
<td>S</td>
<td>35</td>
<td>j</td>
<td>52</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>C</td>
<td>19</td>
<td>T</td>
<td>36</td>
<td>k</td>
<td>53</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>D</td>
<td>20</td>
<td>U</td>
<td>37</td>
<td>l</td>
<td>54</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>E</td>
<td>21</td>
<td>V</td>
<td>38</td>
<td>m</td>
<td>55</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>F</td>
<td>22</td>
<td>W</td>
<td>39</td>
<td>n</td>
<td>56</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>G</td>
<td>23</td>
<td>X</td>
<td>40</td>
<td>o</td>
<td>57</td>
<td>5</td>
</tr>
<tr>
<td>7</td>
<td>H</td>
<td>24</td>
<td>Y</td>
<td>41</td>
<td>p</td>
<td>58</td>
<td>6</td>
</tr>
<tr>
<td>8</td>
<td>I</td>
<td>25</td>
<td>Z</td>
<td>42</td>
<td>q</td>
<td>59</td>
<td>7</td>
</tr>
<tr>
<td>9</td>
<td>J</td>
<td>26</td>
<td>a</td>
<td>43</td>
<td>r</td>
<td>60</td>
<td>8</td>
</tr>
<tr>
<td>10</td>
<td>K</td>
<td>27</td>
<td>b</td>
<td>44</td>
<td>s</td>
<td>61</td>
<td>9</td>
</tr>
<tr>
<td>11</td>
<td>L</td>
<td>28</td>
<td>c</td>
<td>45</td>
<td>t</td>
<td>62</td>
<td>+</td>
</tr>
<tr>
<td>12</td>
<td>M</td>
<td>29</td>
<td>d</td>
<td>46</td>
<td>u</td>
<td>63</td>
<td>/</td>
</tr>
<tr>
<td>13</td>
<td>N</td>
<td>30</td>
<td>e</td>
<td>47</td>
<td>v</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>O</td>
<td>31</td>
<td>f</td>
<td>48</td>
<td>w</td>
<td>(pad)</td>
<td>=</td>
</tr>
<tr>
<td>15</td>
<td>P</td>
<td>32</td>
<td>g</td>
<td>49</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Q</td>
<td>33</td>
<td>h</td>
<td>50</td>
<td>y</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The encoded output stream must be represented in lines of no more than 76 characters each.

Special processing is performed if fewer than 24 bits are available at the end of the data being encoded. There are three possibilities:

1. The last data group has 24 bits (3 octets). No special processing is needed.
2. The last data group has 16 bits (2 octets). The first two 6-bit groups are processed as above. The third (incomplete) data group has two zero-value bits added to it, and is processed as above. A pad character (=) is added to the output.

3. The last data group has 8 bits (1 octet). The first 6-bit group is processed as above. The second (incomplete) data group has four zero-value bits added to it, and is processed as above. Two pad characters (=) are added to the output.

6.4. Decoding Radix-64

In Radix-64 data, characters other than those in the table, line breaks, and other white space probably indicate a transmission error, about which a warning message or even a message rejection might be appropriate under some circumstances. Decoding software must ignore all white space.

Because it is used only for padding at the end of the data, the occurrence of any "=" characters may be taken as evidence that the end of the data has been reached (without truncation in transit). No such assurance is possible, however, when the number of octets transmitted was a multiple of three and no "=" characters are present.

6.5. Examples of Radix-64
6.6. Example of an ASCII Armored Message

-----BEGIN PGP MESSAGE-----
yDgBO22WxBHv7O8X7O/jygAEzo156iUKiXmV+XmpCtmpqQUKiQrFqclFqUDBovzS
vBSFjNSiVHsuAA==
=njUN
-----END PGP MESSAGE-----

Note that this example has extra indenting; an actual armored message would have no leading whitespace.

7. Cleartext Signature Framework

It is desirable to be able to sign a textual octet stream without ASCII armoring the stream itself, so the signed text is still readable without special software. In order to bind a signature to such a cleartext, this framework is used, which follows the same basic format and restrictions as the ASCII armoring described in Section 6.2. (Note that this framework is not intended to be reversible. [RFC3156] defines another way to sign cleartext messages for environments that support MIME.)

The cleartext signed message consists of:
* The cleartext header -----BEGIN PGP SIGNED MESSAGE----- on a single line,

* If the message is signed using v3 or v4 Signatures, one or more "Hash" Armor Headers,

* If the message is signed using v5 Signatures, one or more "SaltedHash" Armor Headers,

* Exactly one empty line not included into the message digest,

* The dash-escaped cleartext that is included into the message digest,

* The ASCII armored signature(s) including the -----BEGIN PGP SIGNATURE----- Armor Header and Armor Tail Lines.

If the "Hash" Armor Header is given, the specified message digest algorithm(s) are used for the signature. If more than one message digest is used in the signature, the "Hash" armor header contains a comma-delimited list of used message digests.

If the "SaltedHash" Armor Header is given, the specified message digest algorithm and salt are used for a signature. The message digest name is followed by a colon (:) followed by 22 characters of Radix-64 encoded salt without padding. Note: The "SaltedHash" Armor Header contains digest algorithm and salt for a single signature; a second signature requires a second "SaltedHash" Armor Header.

Current message digest names are described with the algorithm IDs in Section 9.5.

An implementation SHOULD add a line break after the cleartext, but MAY omit it if the cleartext ends with a line break. This is for visual clarity.

7.1. Dash-Escaped Text

The cleartext content of the message must also be dash-escaped.

Dash-escaped cleartext is the ordinary cleartext where every line starting with a "-" (HYPHEN-MINUS, U+002D) is prefixed by the sequence "-" (HYPHEN-MINUS, U+002D) and " " (SPACE, U+0020). This prevents the parser from recognizing armor headers of the cleartext itself. An implementation MAY dash-escape any line, SHOULD dash-escape lines commencing "From" followed by a space, and MUST dash-escape any line commencing in a dash. The message digest is computed using the cleartext itself, not the dash-escaped form.
As with binary signatures on text documents, a cleartext signature is calculated on the text using canonical <CR><LF> line endings. The line ending (that is, the <CR><LF>) before the -----BEGIN PGP SIGNATURE----- line that terminates the signed text is not considered part of the signed text.

When reversing dash-escaping, an implementation MUST strip the string - if it occurs at the beginning of a line, and SHOULD warn on - and any character other than a space at the beginning of a line.

Also, any trailing whitespace --- spaces (0x20) and tabs (0x09) --- at the end of any line is removed when the cleartext signature is generated.

8. Regular Expressions

A regular expression is zero or more branches, separated by |. It matches anything that matches one of the branches.

A branch is zero or more pieces, concatenated. It matches a match for the first, followed by a match for the second, etc.

A piece is an atom possibly followed by *, +, or ?. An atom followed by * matches a sequence of 0 or more matches of the atom. An atom followed by + matches a sequence of 1 or more matches of the atom. An atom followed by ? matches a match of the atom, or the null string.

An atom is a regular expression in parentheses (matching a match for the regular expression), a range (see below), . (matching any single character), ^ (matching the null string at the beginning of the input string), $ (matching the null string at the end of the input string), \ followed by a single character (matching that character), or a single character with no other significance (matching that character).

A range is a sequence of characters enclosed in [ ]. It normally matches any single character from the sequence. If the sequence begins with ^, it matches any single character not from the rest of the sequence. If two characters in the sequence are separated by -, this is shorthand for the full list of ASCII characters between them (for example, [0-9] matches any decimal digit). To include a literal ] in the sequence, make it the first character (following a possible ^). To include a literal -, make it the first or last character.
9. Constants

This section describes the constants used in OpenPGP.

Note that these tables are not exhaustive lists; an implementation MAY implement an algorithm not on these lists, so long as the algorithm numbers are chosen from the private or experimental algorithm range.

See Section 14 for more discussion of the algorithms.

9.1. Public-Key Algorithms

<table>
<thead>
<tr>
<th>ID</th>
<th>Algorithm</th>
<th>Public Key Format</th>
<th>Secret Key Format</th>
<th>Signature Format</th>
<th>PKESK Format</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>RSA (Encrypt or Sign) [HAC]</td>
<td>MPI(n), MPI(e)</td>
<td>MPI(d), MPI(p),</td>
<td>MPI(m**d mod n)</td>
<td>[Section 5.2.3.1]</td>
</tr>
<tr>
<td></td>
<td>[Section 5.6.1]</td>
<td>[Section 5.6.1]</td>
<td>MPI(q), MPI(u)</td>
<td></td>
<td>[Section 5.1.3]</td>
</tr>
<tr>
<td>2</td>
<td>RSA Encrypt-Only [HAC]</td>
<td>MPI(n), MPI(e)</td>
<td>MPI(d), MPI(p),</td>
<td>N/A</td>
<td>[Section 5.1.3]</td>
</tr>
<tr>
<td></td>
<td>[Section 5.6.1]</td>
<td>[Section 5.6.1]</td>
<td>MPI(q), MPI(u)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>RSA Sign-Only [HAC]</td>
<td>MPI(n), MPI(e)</td>
<td>MPI(d), MPI(p),</td>
<td>MPI(m**d mod n)</td>
<td>[Section 5.1.3]</td>
</tr>
<tr>
<td></td>
<td>[Section 5.6.1]</td>
<td>[Section 5.6.1]</td>
<td>MPI(q), MPI(u)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Elgamal (Encrypt-Only) [ELGAMAL] [HAC]</td>
<td>MPI(p), MPI(g), MPI(y)</td>
<td>MPI(x)</td>
<td>N/A</td>
<td>[Section 5.6.3]</td>
</tr>
<tr>
<td></td>
<td>[Section 5.6.3]</td>
<td>[Section 5.6.3]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>DSA (Digital Signature Algorithm) [FIPS186] [HAC]</td>
<td>MPI(p), MPI(g), MPI(y)</td>
<td>MPI(x)</td>
<td>MPI(r),</td>
<td>[Section 5.2.3.2]</td>
</tr>
<tr>
<td></td>
<td>[Section 5.6.2]</td>
<td>[Section 5.6.2]</td>
<td></td>
<td>MPI(s)</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>ECDH public key algorithm</td>
<td>OID,</td>
<td>MPI(value in curve-</td>
<td>N/A</td>
<td>[Section 5.1.4]</td>
</tr>
<tr>
<td></td>
<td>[Section 5.6.3]</td>
<td>MPI(point)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>in curve-specific point format), KDFParams [see Section 9.2.1, Section 5.6.6]</td>
<td>specific format) [Section 9.2.1]</td>
<td>specific point format), size octet, encoded key [Section 9.2.1, Section 5.1.5, Section 13.5]</td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>ECDSA public key algorithm [FIPS186]</td>
<td>OID, MPI(point in SEC1 format) [Section 5.6.4]</td>
<td>MPI(value)</td>
<td>MPI(r), MPI(s) [Section 5.2.3.2]</td>
<td>N/A</td>
</tr>
<tr>
<td>20</td>
<td>Reserved (formerly Elgamal Encrypt or Sign)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>Reserved for Diffie-Hellman (X9.42, as defined for IETF-S/MIME)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>EdDSA [RFC8032]</td>
<td>OID, MPI(point in prefixed native format) [see Section 13.2.2, Section 5.6.5]</td>
<td>MPI(value in curve-specific format) [see Section 9.2.1]</td>
<td>MPI, MPI [see Section 9.2.1, Section 5.2.3.3]</td>
<td>N/A</td>
</tr>
<tr>
<td>23</td>
<td>Reserved (AEDH)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>Reserved (AEDSA)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Implementations MUST implement EdDSA (19) for signatures, and ECDH (18) for encryption. Implementations SHOULD implement RSA (1) for signatures and encryption.

RSA Encrypt-Only (2) and RSA Sign-Only (3) are deprecated and SHOULD NOT be generated, but may be interpreted. See Section 14.4. See Section 14.8 for notes on Elgamal Encrypt or Sign (20), and X9.42 (21). Implementations MAY implement any other algorithm.

Note that an implementation conforming to the previous version of this standard ([RFC4880]) have only DSA (17) and Elgamal (16) as its MUST-implement algorithms.

A compatible specification of ECDSA is given in [RFC6090] as "KT-I Signatures" and in [SEC1]; ECDH is defined in Section 13.5 of this document.

9.2. ECC Curves for OpenPGP

The parameter curve OID is an array of octets that define a named curve.

The table below specifies the exact sequence of octets for each named curve referenced in this document. It also specifies which public key algorithms the curve can be used with, as well as the size of expected elements in octets:
The "Field Size (fsize)" column represents the field size of the group in number of octets, rounded up, such that x or y coordinates for a point on the curve, native point representations, or scalars with high enough entropy for the curve can be represented in that many octets.

The sequence of octets in the third column is the result of applying the Distinguished Encoding Rules (DER) to the ASN.1 Object Identifier with subsequent truncation. The truncation removes the two fields of encoded Object Identifier. The first omitted field is one octet representing the Object Identifier tag, and the second omitted field is the length of the Object Identifier body. For example, the complete ASN.1 DER encoding for the NIST P-256 curve OID is "06 08 2A 86 48 CE 3D 03 01 07", from which the first entry in the table above is constructed by omitting the first two octets. Only the truncated sequence of octets is the valid representation of a curve OID.

Implementations MUST implement Ed25519 for use with EdDSA, and Curve25519 for use with ECDH. Implementations SHOULD implement Ed448 for use with EdDSA, and X448 for use with ECDH.
### 9.2.1. Curve-Specific Wire Formats

Some Elliptic Curve Public Key Algorithms use different conventions for specific fields depending on the curve in use. Each field is always formatted as an MPI, but with a curve-specific framing. This table summarizes those distinctions.

<table>
<thead>
<tr>
<th>Curve</th>
<th>ECDH Point Format</th>
<th>ECDH Secret Key MPI</th>
<th>EdDSA Secret Key MPI</th>
<th>EdDSA Signature first MPI</th>
<th>EdDSA Signature second MPI</th>
</tr>
</thead>
<tbody>
<tr>
<td>NIST P-256</td>
<td>SEC1</td>
<td>integer</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>NIST P-384</td>
<td>SEC1</td>
<td>integer</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>NIST P-521</td>
<td>SEC1</td>
<td>integer</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Ed25519</td>
<td>N/A</td>
<td>N/A</td>
<td>32 octets of secret</td>
<td>32 octets of R</td>
<td>32 octets of S</td>
</tr>
<tr>
<td>Ed448</td>
<td>N/A</td>
<td>N/A</td>
<td>prefixed 57 octets of secret</td>
<td>prefixed 114 octets of signature</td>
<td>0 [this is an unused placeholder]</td>
</tr>
<tr>
<td>Curve25519</td>
<td>prefixed native</td>
<td>integer (see Section 5.6.6.1.1)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>X448</td>
<td>prefixed native</td>
<td>prefixed 56 octets of secret</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>

**Table 19: Curve-specific wire formats**

For the native octet-string forms of EdDSA values, see [RFC8032].
For the native octet-string forms of ECDH secret scalars and points, see [RFC7748].
### 9.3. Symmetric-Key Algorithms

<table>
<thead>
<tr>
<th>ID</th>
<th>Algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Plaintext or unencrypted data</td>
</tr>
<tr>
<td>1</td>
<td>IDEA [IDEA]</td>
</tr>
<tr>
<td>2</td>
<td>TripleDES (DES-ED3, [SCHNEIER], [HAC] - 168 bit key derived from 192)</td>
</tr>
<tr>
<td>3</td>
<td>CAST5 (128 bit key, as per [RFC2144])</td>
</tr>
<tr>
<td>4</td>
<td>Blowfish (128 bit key, 16 rounds) [BLOWFISH]</td>
</tr>
<tr>
<td>5</td>
<td>Reserved</td>
</tr>
<tr>
<td>6</td>
<td>Reserved</td>
</tr>
<tr>
<td>7</td>
<td>AES with 128-bit key [AES]</td>
</tr>
<tr>
<td>8</td>
<td>AES with 192-bit key</td>
</tr>
<tr>
<td>9</td>
<td>AES with 256-bit key</td>
</tr>
<tr>
<td>10</td>
<td>Twofish with 256-bit key [TWOFISH]</td>
</tr>
<tr>
<td>11</td>
<td>Camellia with 128-bit key [RFC3713]</td>
</tr>
<tr>
<td>12</td>
<td>Camellia with 192-bit key</td>
</tr>
<tr>
<td>13</td>
<td>Camellia with 256-bit key</td>
</tr>
<tr>
<td>100 to 110</td>
<td>Private/Experimental algorithm</td>
</tr>
<tr>
<td>253, 254 and 255</td>
<td>Reserved to avoid collision with Secret Key Encryption (see Section 3.7.2.1 and Section 5.5.3)</td>
</tr>
</tbody>
</table>

Table 20: Symmetric-key algorithm registry
Implementations MUST implement AES-128. Implementations SHOULD implement AES-256. Implementations MUST NOT encrypt data with IDEA, TripleDES, or CAST5. Implementations MAY decrypt data that uses IDEA, TripleDES, or CAST5 for the sake of reading older messages or new messages from legacy clients. Implementations MAY implement any other algorithm.

9.4. Compression Algorithms

<table>
<thead>
<tr>
<th>ID</th>
<th>Algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Uncompressed</td>
</tr>
<tr>
<td>1</td>
<td>ZIP [RFC1951]</td>
</tr>
<tr>
<td>2</td>
<td>ZLIB [RFC1950]</td>
</tr>
<tr>
<td>3</td>
<td>BZip2 [BZ2]</td>
</tr>
<tr>
<td>100 to 110</td>
<td>Private/Experimental algorithm</td>
</tr>
</tbody>
</table>

Table 21: Compression algorithm registry

Implementations MUST implement uncompressed data. Implementations SHOULD implement ZLIB. For interoperability reasons implementations SHOULD be able to decompress using ZIP. Implementations MAY implement any other algorithm.

9.5. Hash Algorithms

<table>
<thead>
<tr>
<th>ID</th>
<th>Algorithm</th>
<th>Text Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>MD5 [HAC]</td>
<td>&quot;MD5&quot;</td>
</tr>
<tr>
<td>2</td>
<td>SHA-1 [FIPS180]</td>
<td>&quot;SHA1&quot;</td>
</tr>
<tr>
<td>3</td>
<td>RIPE-MD/160 [HAC]</td>
<td>&quot;RIPEMD160&quot;</td>
</tr>
<tr>
<td>4</td>
<td>Reserved</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Reserved</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Reserved</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Reserved</td>
<td></td>
</tr>
</tbody>
</table>
9.6. AEAD Algorithms

<table>
<thead>
<tr>
<th>ID</th>
<th>Algorithm</th>
<th>IV length (octets)</th>
<th>authentication tag length (octets)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>EAX [EAX]</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>2</td>
<td>OCB [RFC7253]</td>
<td>15</td>
<td>16</td>
</tr>
<tr>
<td>3</td>
<td>GCM [SP800-38D]</td>
<td>12</td>
<td>16</td>
</tr>
</tbody>
</table>

Table 22: Hash algorithm registry

Implementations MUST implement SHA2-256. Implementations SHOULD implement SHA2-384 and SHA2-512. Implementations MAY implement other algorithms. Implementations SHOULD NOT create messages which require the use of SHA-1 with the exception of computing version 4 key fingerprints and for purposes of the Modification Detection Code (MDC) in version 1 Symmetrically Encrypted Integrity Protected Data packets. Implementations MUST NOT generate signatures with MD5, SHA-1, or RIPE-MD/160. Implementations MUST NOT use MD5, SHA-1, or RIPE-MD/160 as a hash function in an ECDH KDF. Implementations MUST NOT validate any recent signature that depends on MD5, SHA-1, or RIPE-MD/160. Implementations SHOULD NOT validate any old signature that depends on MD5, SHA-1, or RIPE-MD/160 unless the signature's creation date predates known weakness of the algorithm used, and the implementation is confident that the message has been in the secure custody of the user the whole time.
Implementations MUST implement OCB. Implementations MAY implement EAX, GCM and other algorithms.

10. IANA Considerations

Because this document obsoletes [RFC4880], IANA is requested to update all registration information that references [RFC4880] to instead reference this RFC.

OpenPGP is highly parameterized, and consequently there are a number of considerations for allocating parameters for extensions. This section describes how IANA should look at extensions to the protocol as described in this document.

10.1. New String-to-Key Specifier Types

OpenPGP S2K specifiers contain a mechanism for new algorithms to turn a string into a key. This specification creates a registry of S2K specifier types. The registry includes the S2K type, the name of the S2K, and a reference to the defining specification. The initial values for this registry can be found in Section 3.7.1. Adding a new S2K specifier MUST be done through the SPECIFICATION REQUIRED method, as described in [RFC8126].

IANA should add a column "Generate?" to the S2K type registry, with initial values taken from Section 3.7.1.

10.2. New Packets

Major new features of OpenPGP are defined through new packet types. This specification creates a registry of packet types. The registry includes the packet type, the name of the packet, and a reference to the defining specification. The initial values for this registry can be found in Section 4.3. Adding a new packet type MUST be done through the RFC REQUIRED method, as described in [RFC8126].
10.2.1. User Attribute Types

The User Attribute packet permits an extensible mechanism for other types of certificate identification. This specification creates a registry of User Attribute types. The registry includes the User Attribute type, the name of the User Attribute, and a reference to the defining specification. The initial values for this registry can be found in Section 5.13. Adding a new User Attribute type MUST be done through the SPECIFICATION REQUIRED method, as described in [RFC8126].

10.2.1.1. Image Format Subpacket Types

Within User Attribute packets, there is an extensible mechanism for other types of image-based User Attributes. This specification creates a registry of Image Attribute subpacket types. The registry includes the Image Attribute subpacket type, the name of the Image Attribute subpacket, and a reference to the defining specification. The initial values for this registry can be found in Section 5.13.1. Adding a new Image Attribute subpacket type MUST be done through the SPECIFICATION REQUIRED method, as described in [RFC8126].

10.2.2. New Signature Subpackets

OpenPGP signatures contain a mechanism for signed (or unsigned) data to be added to them for a variety of purposes in the Signature subpackets as discussed in Section 5.2.3.5. This specification creates a registry of Signature subpacket types. The registry includes the Signature subpacket type, the name of the subpacket, and a reference to the defining specification. The initial values for this registry can be found in Section 5.2.3.5. Adding a new Signature subpacket MUST be done through the SPECIFICATION REQUIRED method, as described in [RFC8126].

10.2.2.1. Signature Notation Data Subpackets

OpenPGP signatures further contain a mechanism for extensions in signatures. These are the Notation Data subpackets, which contain a key/value pair. Notations contain a user space that is completely unmanaged and an IETF space.

This specification creates a registry of Signature Notation Data types. The registry includes the Signature Notation Data type, the name of the Signature Notation Data, its allowed values, and a reference to the defining specification. The initial values for this registry can be found in Section 5.2.3.21. Adding a new Signature Notation Data subpacket MUST be done through the SPECIFICATION REQUIRED method, as described in [RFC8126].
10.2.2.2. Signature Notation Data Subpacket Notation Flags

This specification creates a new registry of Signature Notation Data Subpacket Notation Flags. The registry includes the columns "Flag", "Description", "Security Recommended", "Interoperability Recommended", and "Reference". The initial values for this registry can be found in Section 5.2.3.21. Adding a new item MUST be done through the SPECIFICATION REQUIRED method, as described in [RFC8126].

10.2.2.3. Key Server Preference Extensions

OpenPGP signatures contain a mechanism for preferences to be specified about key servers. This specification creates a registry of key server preferences. The registry includes the key server preference, the name of the preference, and a reference to the defining specification. The initial values for this registry can be found in Section 5.2.3.22. Adding a new key server preference MUST be done through the SPECIFICATION REQUIRED method, as described in [RFC8126].

10.2.2.4. Key Flags Extensions

OpenPGP signatures contain a mechanism for flags to be specified about key usage. This specification creates a registry of key usage flags. The registry includes the key flags value, the name of the flag, and a reference to the defining specification. The initial values for this registry can be found in Section 5.2.3.26. Adding a new key usage flag MUST be done through the SPECIFICATION REQUIRED method, as described in [RFC8126].

10.2.2.5. Reason for Revocation Extensions

OpenPGP signatures contain a mechanism for flags to be specified about why a key was revoked. This specification creates a registry of "Reason for Revocation" flags. The registry includes the "Reason for Revocation" flags value, the name of the flag, and a reference to the defining specification. The initial values for this registry can be found in Section 5.2.3.28. Adding a new feature flag MUST be done through the SPECIFICATION REQUIRED method, as described in [RFC8126].

10.2.2.6. Implementation Features

OpenPGP signatures contain a mechanism for flags to be specified stating which optional features an implementation supports. This specification creates a registry of feature-implementation flags. The registry includes the feature-implementation flags value, the name of the flag, and a reference to the defining specification. The initial values for this registry can be found in Section 5.2.3.29.
Adding a new feature-implementation flag MUST be done through the SPECIFICATION REQUIRED method, as described in [RFC8126].

Also see Section 14.11 for more information about when feature flags are needed.

10.2.3. New Packet Versions

The core OpenPGP packets all have version numbers, and can be revised by introducing a new version of an existing packet. This specification creates a registry of packet types. The registry includes the packet type, the number of the version, and a reference to the defining specification. The initial values for this registry can be found in Section 5. Adding a new packet version MUST be done through the RFC REQUIRED method, as described in [RFC8126].

10.3. New Algorithms

Section 9 lists the core algorithms that OpenPGP uses. Adding in a new algorithm is usually simple. For example, adding in a new symmetric cipher usually would not need anything more than allocating a constant for that cipher. If that cipher had other than a 64-bit or 128-bit block size, there might need to be additional documentation describing how OpenPGP-CFB mode would be adjusted. Similarly, when DSA was expanded from a maximum of 1024-bit public keys to 3072-bit public keys, the revision of FIPS 186 contained enough information itself to allow implementation. Changes to this document were made mainly for emphasis.

10.3.1. Public-Key Algorithms

OpenPGP specifies a number of public-key algorithms. This specification creates a registry of public-key algorithm identifiers. The registry includes the algorithm name, its key sizes and parameters, and a reference to the defining specification. The initial values for this registry can be found in Section 9.1. Adding a new public-key algorithm MUST be done through the SPECIFICATION REQUIRED method, as described in [RFC8126].

This document requests IANA register the following new public-key algorithm:
## 10.3.2. Symmetric-Key Algorithms

OpenPGP specifies a number of symmetric-key algorithms. This specification creates a registry of symmetric-key algorithm identifiers. The registry includes the algorithm name, its key sizes and block size, and a reference to the defining specification. The initial values for this registry can be found in Section 9.3. Adding a new symmetric-key algorithm MUST be done through the SPECIFICATION REQUIRED method, as described in [RFC8126].

### 10.3.3. Hash Algorithms

OpenPGP specifies a number of hash algorithms. This specification creates a registry of hash algorithm identifiers. The registry includes the algorithm name, a text representation of that name, its block size, an OID hash prefix, and a reference to the defining specification. The initial values for this registry can be found in Section 9.5 for the algorithm identifiers and text names, and Section 5.2.2 for the OIDs and expanded signature prefixes. Adding a new hash algorithm MUST be done through the SPECIFICATION REQUIRED method, as described in [RFC8126].

This document requests IANA register the following hash algorithms:

```
+----------+----------+----------------+
| ID       | Algorithm| Reference      |
|----------+----------+----------------|
| 12       | SHA3-256 | This doc       |
| 13       | Reserved |               |
| 14       | SHA3-512 | This doc       |
```

Table 25: New hash algorithms registered

---

[Note to RFC-Editor: Please remove the table above on publication.]
[Notes to RFC-Editor: Please remove the table above on publication.
It is desirable not to reuse old or reserved algorithms because some
existing tools might print a wrong description. The ID 13 has been
reserved so that the SHA3 algorithm IDs align nicely with their SHA2
countepts.]

10.3.4. Compression Algorithms

OpenPGP specifies a number of compression algorithms. This
specification creates a registry of compression algorithm
identifiers. The registry includes the algorithm name and a
reference to the defining specification. The initial values for this
registry can be found in Section 9.4. Adding a new compression key
algorithm MUST be done through the SPECIFICATION REQUIRED method, as
described in [RFC8126].

10.3.5. Elliptic Curve Algorithms

This document requests IANA add a registry of elliptic curves for use
in OpenPGP.

Each curve is identified on the wire by OID, and is acceptable for
use in certain OpenPGP public key algorithms. The table’s initial
headings and values can be found in Section 9.2. Adding a new
elliptic curve algorithm to OpenPGP MUST be done through the
SPECIFICATION REQUIRED method, as described in [RFC8126]. If the new
curve can be used for ECDH or EdDSA, it must also be added to the
"Curve-specific wire formats" table described in Section 9.2.1.

10.4. Elliptic Curve Point and Scalar Wire Formats

This document requests IANA add a registry of wire formats that
represent elliptic curve points. The table’s initial headings and
values can be found in Section 13.2. Adding a new EC point wire
format MUST be done through the SPECIFICATION REQUIRED method, as
described in [RFC8126].

This document also requests IANA add a registry of wire formats that
represent scalars for use with elliptic curve cryptography. The
table’s initial headings and values can be found in Section 13.3.
Adding a new EC scalar wire format MUST be done through the
SPECIFICATION REQUIRED method, as described in [RFC8126].
This document also requests that IANA add a registry mapping curve-specific MPI octet-string encoding conventions for ECDH and EdDSA. The table’s initial headings and values can be found in Section 9.2.1. Adding a new elliptic curve algorithm to OpenPGP MUST be done through the SPECIFICATION REQUIRED method, as described in [RFC8126], and requires adding an entry to this table if the curve is to be used with either EdDSA or ECDH.

10.5. Changes to existing registries

This document requests IANA add the following wire format columns to the OpenPGP public-key algorithm registry:

* Public Key Format
* Secret Key Format
* Signature Format
* PKESK Format

And populate them with the values found in Section 9.1.

11. Packet Composition

OpenPGP packets are assembled into sequences in order to create messages and to transfer keys. Not all possible packet sequences are meaningful and correct. This section describes the rules for how packets should be placed into sequences.

11.1. Transferable Public Keys

OpenPGP users may transfer public keys. The essential elements of a transferable public key are as follows:

* One Public-Key packet
* Zero or more revocation signatures
* Zero or more User ID packets
* After each User ID packet, zero or more Signature packets (certifications)
* Zero or more User Attribute packets
* After each User Attribute packet, zero or more Signature packets (certifications)
* Zero or more Subkey packets

* After each Subkey packet, one Signature packet, plus optionally a revocation

* An optional Padding packet

The Public-Key packet occurs first. Each of the following User ID packets provides the identity of the owner of this public key. If there are multiple User ID packets, this corresponds to multiple means of identifying the same unique individual user; for example, a user may have more than one email address, and construct a User ID for each one. A transferable public key SHOULD include at least one User ID packet unless storage requirements prohibit this.

Immediately following each User ID packet, there are zero or more Signature packets. Each Signature packet is calculated on the immediately preceding User ID packet and the initial Public-Key packet. The signature serves to certify the corresponding public key and User ID. In effect, the signer is testifying to his or her belief that this public key belongs to the user identified by this User ID.

Within the same section as the User ID packets, there are zero or more User Attribute packets. Like the User ID packets, a User Attribute packet is followed by zero or more Signature packets calculated on the immediately preceding User Attribute packet and the initial Public-Key packet.

User Attribute packets and User ID packets may be freely intermixed in this section, so long as the signatures that follow them are maintained on the proper User Attribute or User ID packet.

After the User ID packet or Attribute packet, there may be zero or more Subkey packets. In general, subkeys are provided in cases where the top-level public key is a signature-only key. However, any V4 or V5 key may have subkeys, and the subkeys may be encryption-only keys, signature-only keys, or general-purpose keys. V3 keys MUST NOT have subkeys.

Each Subkey packet MUST be followed by one Signature packet, which should be a subkey binding signature issued by the top-level key. For subkeys that can issue signatures, the subkey binding signature MUST contain an Embedded Signature subpacket with a primary key binding signature (0x19) issued by the subkey on the top-level key.
Subkey and Key packets may each be followed by a revocation Signature packet to indicate that the key is revoked. Revocation signatures are only accepted if they are issued by the key itself, or by a key that is authorized to issue revocations via a Revocation Key subpacket in a self-signature by the top-level key.

The optional trailing Padding packet is a mechanism to defend against traffic analysis (see Section 15.4). For maximum interoperability, if the Public-Key packet is a V4 key, the optional Padding packet SHOULD NOT be present unless the recipient has indicated that they are capable of ignoring it successfully. An implementation that is capable of receiving a transferable public key with a V5 Public-Key primary key MUST be able to accept (and ignore) the trailing optional Padding packet.

Transferable public-key packet sequences may be concatenated to allow transferring multiple public keys in one operation.

11.2. Transferable Secret Keys

OpenPGP users may transfer secret keys. The format of a transferable secret key is the same as a transferable public key except that secret-key and secret-subkey packets are used instead of the public key and public-subkey packets. Implementations SHOULD include self-signatures on any User IDs and subkeys, as this allows for a complete public key to be automatically extracted from the transferable secret key. Implementations MAY choose to omit the self-signatures, especially if a transferable public key accompanies the transferable secret key.

11.3. OpenPGP Messages

An OpenPGP message is a packet or sequence of packets that corresponds to the following grammatical rules (comma represents sequential composition, and vertical bar separates alternatives):


Compressed Message :-  Compressed Data Packet.

Literal Message :-  Literal Data Packet.

ESK :-  Public-Key Encrypted Session Key Packet | Symmetric-Key Encrypted Session Key Packet.

ESK Sequence :-  ESK | ESK Sequence, ESK.
Encrypted Data :-  Symmetrically Encrypted Data Packet | Symmetrically Encrypted Integrity Protected Data Packet

Encrypted Message :-  Encrypted Data | ESK Sequence, Encrypted Data.

One-Pass Signed Message :-  One-Pass Signature Packet, OpenPGP Message, Corresponding Signature Packet.

Signed Message :-  Signature Packet, OpenPGP Message | One-Pass Signed Message.

Optionally Padded Message :-  OpenPGP Message | OpenPGP Message, Padding Packet.

11.3.1. Unwrapping Encrypted and Compressed Messages

In addition to the above grammar, certain messages can be "unwrapped" to yield new messages. In particular:

* Decrypting a version 2 Symmetrically Encrypted and Integrity Protected Data packet must yield a valid Optionally Padded Message.

* Decrypting a version 1 Symmetrically Encrypted and Integrity Protected Data packet or --- for historic data --- a Symmetrically Encrypted Data packet must yield a valid OpenPGP Message.

* Decompressing a Compressed Data packet must also yield a valid OpenPGP Message.

When either such unwrapping is performed, the resulting stream of octets is parsed into a series OpenPGP packets like any other stream of octets. The packet boundaries found in the series of octets are expected to align with the length of the unwrapped octet stream. An implementation MUST NOT interpret octets beyond the boundaries of the unwrapped octet stream as part of any OpenPGP packet. If an implementation encounters a packet whose header length indicates that it would extend beyond the boundaries of the unwrapped octet stream, the implementation MUST reject that packet as malformed and unusable.

11.3.2. Additional Constraints on Packet Sequences

Note that some subtle combinations that are formally acceptable by this grammar are nonetheless unacceptable.
11.3.2.1. Packet Versions in Encrypted Messages

As noted above, an Encrypted Message is a sequence of zero or more PKESKs (Section 5.1) and SKESKs (Section 5.3), followed by an SEIPD (Section 5.14) payload. In some historic data, the payload may be a deprecated SED (Section 5.8) packet instead of SEIPD, though implementations MUST NOT generate SED packets (see Section 15.1). The versions of the preceding ESK packets within an Encrypted Message MUST align with the version of the payload SEIPD packet, as described in this section.

v3 PKESK and v4 SKESK packets both contain in their cleartext the symmetric cipher algorithm identifier in addition to the session key for the subsequent SEIPD packet. Since a v1 SEIPD does not contain a symmetric algorithm identifier, so all ESK packets preceding a v1 SEIPD payload MUST be either v3 PKESK or v4 SKESK.

On the other hand, the cleartext of the v5 ESK packets (either PKESK or SKESK) do not contain a symmetric cipher algorithm identifier, so they cannot be used in combination with a v1 SEIPD payload. The payload following any v5 PKESK or v5 SKESK packet MUST be a v2 SEIPD.

Additionally, to avoid potentially conflicting cipher algorithm identifiers, and for simplicity, implementations MUST NOT precede a v2 SEIPD payload with either v3 PKESK or v4 SKESK packets.

The acceptable versions of packets in an Encrypted Message are summarized in the following table:

<table>
<thead>
<tr>
<th>Version of Encrypted Data payload</th>
<th>Version of preceding Symmetric-Key ESK (if any)</th>
<th>Version of preceding Public-Key ESK (if any)</th>
</tr>
</thead>
<tbody>
<tr>
<td>v1 SEIPD</td>
<td>v4 SKESK</td>
<td>v3 PKESK</td>
</tr>
<tr>
<td>v2 SEIPD</td>
<td>v5 SKESK</td>
<td>v5 PKESK</td>
</tr>
</tbody>
</table>

Table 26: Encrypted Message Packet Version Alignment

An implementation processing an Encrypted Message MUST discard any preceding ESK packet with a version that does not align with the version of the payload.
11.4. Detached Signatures

Some OpenPGP applications use so-called "detached signatures". For example, a program bundle may contain a file, and with it a second file that is a detached signature of the first file. These detached signatures are simply a Signature packet stored separately from the data for which they are a signature.

12. Enhanced Key Formats

12.1. Key Structures

The format of an OpenPGP V3 key is as follows. Entries in square brackets are optional and ellipses indicate repetition.

RSA Public Key
  [Revocation Self Signature]
  User ID [Signature ...]
  [User ID [Signature ...] ...]

Each signature certifies the RSA public key and the preceding User ID. The RSA public key can have many User IDs and each User ID can have many signatures. V3 keys are deprecated. Implementations MUST NOT generate new V3 keys, but MAY continue to use existing ones.

The format of an OpenPGP V4 key that uses multiple public keys is similar except that the other keys are added to the end as "subkeys" of the primary key.

Primary-Key
  [Revocation Self Signature]
  [Direct Key Signature...]
  [User ID [Signature ...] ...]
  [User Attribute [Signature ...] ...]
  [[Subkey [Binding-Signature-Revocation ...]
    Subkey-Binding-Signature ...] ...]

A subkey always has at least one subkey binding signature after it that is issued using the primary key to tie the two keys together. These binding signatures may be in either V3 or V4 format, but SHOULD be V4. Subkeys that can issue signatures MUST have a V4 binding signature due to the REQUIRED embedded primary key binding signature.
In order to create self-signatures (see Section 5.2.3.7), the primary key MUST be an algorithm capable of making signatures (that is, not an encryption-only algorithm). The subkeys may be keys of any type. For example, there may be a single-key RSA key, an EdDSA primary key with an RSA encryption key, or an EdDSA primary key with an ECDH subkey, etc.

It is also possible to have a signature-only subkey. This permits a primary key that collects certifications (key signatures), but is used only for certifying subkeys that are used for encryption and signatures.

12.2. Key IDs and Fingerprints

For a V3 key, the eight-octet Key ID consists of the low 64 bits of the public modulus of the RSA key.

The fingerprint of a V3 key is formed by hashing the body (but not the two-octet length) of the MPIs that form the key material (public modulus n, followed by exponent e) with MD5. Note that both V3 keys and MD5 are deprecated.

A V4 fingerprint is the 160-bit SHA-1 hash of the octet 0x99, followed by the two-octet packet length, followed by the entire Public-Key packet starting with the version field. The Key ID is the low-order 64 bits of the fingerprint. Here are the fields of the hash material, with the example of an EdDSA key:

a.1) 0x99 (1 octet)

a.2) two-octet, big-endian scalar octet count of (b)-(e)

b) version number = 4 (1 octet);

c) timestamp of key creation (4 octets);

d) algorithm (1 octet): 22 = EdDSA (example);

e) Algorithm-specific fields.

Algorithm-Specific Fields for EdDSA keys (example):

e.1) A one-octet size of the following field;

e.2) The octets representing a curve OID, defined in Section 9.2;

e.3) An MPI of an EC point representing a public key Q in prefixed native form (see Section 13.2.2).
A V5 fingerprint is the 256-bit SHA2-256 hash of the octet 0x9A, followed by the four-octet packet length, followed by the entire Public-Key packet starting with the version field. The Key ID is the high-order 64 bits of the fingerprint. Here are the fields of the hash material, with the example of an EdDSA key:

a.1) 0x9A (1 octet)

a.2) four-octet scalar octet count of (b)-(f)

b) version number = 5 (1 octet);

c) timestamp of key creation (4 octets);

d) algorithm (1 octet): 22 = EdDSA (example);

e) four-octet scalar octet count for the following key material;

f) algorithm-specific fields.

Algorithm-Specific Fields for EdDSA keys (example):

f.1) A one-octet size of the following field;

f.2) The octets representing a curve OID, defined in Section 9.2;

f.3) An MPI of an EC point representing a public key Q in prefixed native form (see Section 13.2.2).

Note that it is possible for there to be collisions of Key IDs --- two different keys with the same Key ID. Note that there is a much smaller, but still non-zero, probability that two different keys have the same fingerprint.

Also note that if V3, V4, and V5 format keys share the same RSA key material, they will have different Key IDs as well as different fingerprints.

Finally, the Key ID and fingerprint of a subkey are calculated in the same way as for a primary key, including the 0x99 (V4 key) or 0x9A (V5 key) as the first octet (even though this is not a valid packet ID for a public subkey).

13. Elliptic Curve Cryptography

This section describes algorithms and parameters used with Elliptic Curve Cryptography (ECC) keys. A thorough introduction to ECC can be found in [KOBLITZ].
None of the ECC methods described in this document are allowed with deprecated V3 keys. Refer to [FIPS186], B.4.1, for the method to generate a uniformly distributed ECC private key.

13.1. Supported ECC Curves

This document references three named prime field curves defined in [FIPS186] as "Curve P-256", "Curve P-384", and "Curve P-521". These three [FIPS186] curves can be used with ECDSA and ECDH public key algorithms. Additionally, curve "Curve25519" and "Curve448" are referenced for use with Ed25519 and Ed448 (EdDSA signing, see [RFC8032]); and X25519 and X448 (ECDH encryption, see [RFC7748]).

The named curves are referenced as a sequence of octets in this document, called throughout, curve OID. Section 9.2 describes in detail how this sequence of octets is formed.

13.2. EC Point Wire Formats

A point on an elliptic curve will always be represented on the wire as an MPI. Each curve uses a specific point format for the data within the MPI itself. Each format uses a designated prefix octet to ensure that the high octet has at least one bit set to make the MPI a constant size.

<table>
<thead>
<tr>
<th>Name</th>
<th>Wire Format</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>SEC1</td>
<td>0x04</td>
<td></td>
</tr>
<tr>
<td>Prefixed native</td>
<td>0x40</td>
<td></td>
</tr>
</tbody>
</table>

Table 27: Elliptic Curve Point Wire Formats

13.2.1. SEC1 EC Point Wire Format

For a SEC1-encoded (uncompressed) point the content of the MPI is:

\[ B = 04 || x || y \]

where x and y are coordinates of the point \( P = (x, y) \), and each is encoded in the big-endian format and zero-padded to the adjusted underlying field size. The adjusted underlying field size is the underlying field size rounded up to the nearest 8-bit boundary, as noted in the "fsize" column in Section 9.2. This encoding is compatible with the definition given in [SEC1].
13.2.2. Prefixed Native EC Point Wire Format

For a custom compressed point the content of the MPI is:

\[ B = 40 \ || \ p \]

where \( p \) is the public key of the point encoded using the rules defined for the specified curve. This format is used for ECDH keys based on curves expressed in Montgomery form, and for points when using EdDSA.

13.2.3. Notes on EC Point Wire Formats

Given the above definitions, the exact size of the MPI payload for an encoded point is 515 bits for "Curve P-256", 771 for "Curve P-384", 1059 for "Curve P-521", 263 for both "Curve25519" and "Ed25519", 463 for "Ed448", and 455 for "X448". For example, the length of a EdDSA public key for the curve Ed25519 is 263 bits: 7 bits to represent the 0x40 prefix octet and 32 octets for the native value of the public key.

Even though the zero point, also called the point at infinity, may occur as a result of arithmetic operations on points of an elliptic curve, it SHALL NOT appear in data structures defined in this document.

Each particular curve uses a designated wire format for the point found in its public key or ECDH data structure. An implementation MUST NOT use a different wire format for a point than the wire format associated with the curve.

13.3. EC Scalar Wire Formats

Some non-curve values in elliptic curve cryptography (for example, secret keys and signature components) are not points on a curve, but are also encoded on the wire in OpenPGP as an MPI.

Because of different patterns of deployment, some curves treat these values as opaque bit strings with the high bit set, while others are treated as actual integers, encoded in the standard OpenPGP big-endian form. The choice of encoding is specific to the public key algorithm in use.
### Table 28: Elliptic Curve Scalar Encodings

#### 13.3.1. EC Octet String Wire Format

Some opaque strings of octets are represented on the wire as an MPI by simply stripping the leading zeros and counting the remaining bits. These strings are of known, fixed length. They are represented in this document as MPI(N octets of X) where N is the expected length in octets of the octet string.

For example, a five-octet opaque string (MPI(5 octets of X)) where X has the value 00 02 ee 19 00 would be represented on the wire as an MPI like so: 00 1a 02 ee 19 00.

To encode X to the wire format, we set the MPI’s two-octet bit counter to the value of the highest set bit (bit 26, or 0x001a), and do not transfer the leading all-zero octet to the wire.

To reverse the process, an implementation that knows this value has an expected length of 5 octets can take the following steps:

* ensure that the MPI’s two-octet bitcount is less than or equal to 40 (5 octets of 8 bits)
* allocate 5 octets, setting all to zero initially
* copy the MPI data octets (without the two count octets) into the lower octets of the allocated space
13.3.2. Elliptic Curve Prefixed Octet String Wire Format

Another way to ensure that a fixed-length bytestring is encoded simply to the wire while remaining in MPI format is to prefix the bytestring with a dedicated non-zero octet. This specification uses 0x40 as the prefix octet. This is represented in this standard as MPI(prefixed N octets of X), where N is the known bytestring length.

For example, a five-octet opaque string using MPI(prefixed 5 octets of X) where X has the value 00 02 ee 19 00 would be written to the wire form as: 00 2f 40 00 02 ee 19 00.

To encode the string, we prefix it with the octet 0x40 (whose 7th bit is set), then set the MPI’s two-octet bit counter to 47 (0x002f, 7 bits for the prefix octet and 40 bits for the string).

To decode the string from the wire, an implementation that knows that the variable is formed in this way can:

* ensure that the first three octets of the MPI (the two bit-count octets plus the prefix octet) are 00 2f 40, and
* use the remainder of the MPI directly off the wire.

Note that this is a similar approach to that used in the EC point encodings found in Section 13.2.2.

13.4. Key Derivation Function

A key derivation function (KDF) is necessary to implement EC encryption. The Concatenation Key Derivation Function (Approved Alternative 1) [SP800-56A] with the KDF hash function that is SHA2-256 [FIPS180] or stronger is REQUIRED.

For convenience, the synopsis of the encoding method is given below with significant simplifications attributable to the restricted choice of hash functions in this document. However, [SP800-56A] is the normative source of the definition.
// Implements KDF( X, oBits, Param );
// Input: point X = (x,y)
// oBits - the desired size of output
// hBits - the size of output of hash function Hash
// Param - octets representing the parameters
// Assumes that oBits <= hBits
// Convert the point X to the octet string:
// ZB' = 04 || x || y
// and extract the x portion from ZB'
ZB = x;
MB = Hash ( 00 || 00 || 00 || 01 || ZB || Param );
return oBits leftmost bits of MB.

Note that ZB in the KDF description above is the compact
representation of X as defined in Section 4.2 of [RFC6090].

13.5. EC DH Algorithm (ECDH)

The method is a combination of an ECC Diffie-Hellman method to
establish a shared secret, a key derivation method to process the
shared secret into a derived key, and a key wrapping method that uses
the derived key to protect a session key used to encrypt a message.

The One-Pass Diffie-Hellman method C(1, 1, ECC CDH) [SP800-56A] MUST
be implemented with the following restrictions: the ECC CDH primitive
employed by this method is modified to always assume the cofactor is
1, the KDF specified in Section 13.4 is used, and the KDF parameters
specified below are used.

The KDF parameters are encoded as a concatenation of the following 5
variable-length and fixed-length fields, which are compatible with
the definition of the OtherInfo bitstring [SP800-56A]:

* A variable-length field containing a curve OID, which is formatted
  as follows:
  - A one-octet size of the following field,
  - The octets representing a curve OID defined in Section 9.2;

* A one-octet public key algorithm ID defined in Section 9.1;

* A variable-length field containing KDF parameters, which are
  identical to the corresponding field in the ECDH public key, and
  are formatted as follows:
  - A one-octet size of the following fields; values 0 and 0xFF are
    reserved for future extensions,
- A one-octet value 0x01, reserved for future extensions,

- A one-octet hash function ID used with the KDF,

- A one-octet algorithm ID for the symmetric algorithm used to wrap the symmetric key for message encryption; see Section 13.5 for details;

* 20 octets representing the UTF-8 encoding of the string Anonymous Sender, which is the octet sequence 41 6E 6F 6E 79 6D 6F 75 73 20 53 65 6E 64 65 72 20 20 20 20;

* A variable-length field containing the fingerprint of the recipient encryption subkey or a primary key fingerprint identifying the key material that is needed for decryption. For version 4 keys, this field is 20 octets. For version 5 keys, this field is 32 octets.

The size in octets of the KDF parameters sequence, defined above, for encrypting to a v4 key is either 54 for curve P-256, 51 for curves P-384 and P-521, 56 for Curve25519, or 49 for X448. For encrypting to a v5 key, the size of the sequence is either 66 for curve P-256, 63 for curves P-384 and P-521, 68 for Curve25519, or 61 for X448.

The key wrapping method is described in [RFC3394]. The KDF produces a symmetric key that is used as a key-encryption key (KEK) as specified in [RFC3394]. Refer to Section 15 for the details regarding the choice of the KEK algorithm, which SHOULD be one of three AES algorithms. Key wrapping and unwrapping is performed with the default initial value of [RFC3394].

The input to the key wrapping method is the plaintext described in Section 5.1, "Public-Key Encrypted Session Key Packets (Tag 1)", padded using the method described in [PKCS5] to an 8-octet granularity.

For example, in a V4 Public-Key Encrypted Session Key packet, the following AES-256 session key, in which 32 octets are denoted from k0 to k31, is composed to form the following 40 octet sequence:

09 k0 k1 ... k31 s0 s1 05 05 05 05 05
The octets s0 and s1 above denote the checksum of the session key octets. This encoding allows the sender to obfuscate the size of the symmetric encryption key used to encrypt the data. For example, assuming that an AES algorithm is used for the session key, the sender MAY use 21, 13, and 5 octets of padding for AES-128, AES-192, and AES-256, respectively, to provide the same number of octets, 40 total, as an input to the key wrapping method.

In a V5 Public-Key Encrypted Session Key packet, the symmetric algorithm is not included, as described in Section 5.1. For example, an AES-256 session key would be composed as follows:

```
k0 k1 ... k31 s0 s1 06 06 06 06 06 06
```

The octets k0 to k31 above again denote the session key, and the octets s0 and s1 denote the checksum. In this case, assuming that an AES algorithm is used for the session key, the sender MAY use 22, 14, and 6 octets of padding for AES-128, AES-192, and AES-256, respectively, to provide the same number of octets, 40 total, as an input to the key wrapping method.

The output of the method consists of two fields. The first field is the MPI containing the ephemeral key used to establish the shared secret. The second field is composed of the following two subfields:

* One octet encoding the size in octets of the result of the key wrapping method; the value 255 is reserved for future extensions;

* Up to 254 octets representing the result of the key wrapping method, applied to the 8-octet padded session key, as described above.

Note that for session key sizes 128, 192, and 256 bits, the size of the result of the key wrapping method is, respectively, 32, 40, and 48 octets, unless size obfuscation is used.

For convenience, the synopsis of the encoding method is given below; however, this section, [SP800-56A], and [RFC3394] are the normative sources of the definition.

* Obtain the authenticated recipient public key R

* Generate an ephemeral key pair \(\{v, V=vG\}\)

* Compute the shared point \(S = vR\);

* \(m = \text{symm_alg_ID} \ || \ session \ key \ || \ checksum \ || \ pkcs5\_padding\);
* curve_OID_len = (octet)len(curve_OID);

* Param = curve_OID_len || curve_OID || public_key_alg_ID || 03 || 01 || KDF_hash_ID || KEK_alg_ID for AESKeyWrap || Anonymous Sender || recipient_fingerprint;

* Z_len = the key size for the KEK_alg_ID used with AESKeyWrap

* Compute Z = KDF( S, Z_len, Param );

* Compute C = AESKeyWrap( Z, m ) as per [RFC3394]

* VB = convert point V to the octet string

* Output (MPI(VB) || len(C) || C).

The decryption is the inverse of the method given. Note that the recipient obtains the shared secret by calculating

S = rV = rvG, where (r,R) is the recipient’s key pair.

Consistent with Section 5.14, AEAD encryption or a Modification Detection Code (MDC) MUST be used anytime the symmetric key is protected by ECDH.

14. Notes on Algorithms

14.1. PKCS#1 Encoding in OpenPGP

This standard makes use of the PKCS#1 functions EME-PKCS1-v1_5 and EMSA-PKCS1-v1_5. However, the calling conventions of these functions has changed in the past. To avoid potential confusion and interoperability problems, we are including local copies in this document, adapted from those in PKCS#1 v2.1 [RFC8017]. [RFC8017] should be treated as the ultimate authority on PKCS#1 for OpenPGP. Nonetheless, we believe that there is value in having a self-contained document that avoids problems in the future with needed changes in the conventions.

14.1.1. EME-PKCS1-v1_5-ENCODE

Input:

k = the length in octets of the key modulus.

M = message to be encoded, an octet string of length mLen, where mLen <= k - 11.
Output:

EM = encoded message, an octet string of length k.

Error: "message too long".

1. Length checking: If mLen > k - 11, output "message too long" and stop.

2. Generate an octet string PS of length k - mLen - 3 consisting of pseudo-randomly generated nonzero octets. The length of PS will be at least eight octets.

3. Concatenate PS, the message M, and other padding to form an encoded message EM of length k octets as

   EM = 0x00 || 0x02 || PS || 0x00 || M.

4. Output EM.

14.1.2. EME-PKCS1-v1_5-DECODE

Input:

EM = encoded message, an octet string

Output:

M = message, an octet string.

Error: "decryption error".

To decode an EME-PKCS1_v1_5 message, separate the encoded message EM into an octet string PS consisting of nonzero octets and a message M as follows

   EM = 0x00 || 0x02 || PS || 0x00 || M.

If the first octet of EM does not have hexadecimal value 0x00, if the second octet of EM does not have hexadecimal value 0x02, if there is no octet with hexadecimal value 0x00 to separate PS from M, or if the length of PS is less than 8 octets, output "decryption error" and stop. See also the security note in Section 15 regarding differences in reporting between a decryption error and a padding error.
14.1.3. EMSA-PKCS1-v1_5

This encoding method is deterministic and only has an encoding operation.

Option:

Hash - a hash function in which hLen denotes the length in octets of the hash function output.

Input:

M = message to be encoded.

emLen = intended length in octets of the encoded message, at least tLen + 11, where tLen is the octet length of the DER encoding T of a certain value computed during the encoding operation.

Output:

EM = encoded message, an octet string of length emLen.

Errors: "message too long"; "intended encoded message length too short".

Steps:

1. Apply the hash function to the message M to produce a hash value H:

   H = Hash(M).

   If the hash function outputs "message too long," output "message too long" and stop.

2. Using the list in Section 5.2.2, produce an ASN.1 DER value for the hash function used. Let T be the full hash prefix from the list, and let tLen be the length in octets of T.

3. If emLen < tLen + 11, output "intended encoded message length too short" and stop.

4. Generate an octet string PS consisting of emLen - tLen - 3 octets with hexadecimal value 0xFF. The length of PS will be at least 8 octets.

5. Concatenate PS, the hash prefix T, and other padding to form the encoded message EM as
EM = 0x00 || 0x01 || PS || 0x00 || T.

6. Output EM.

14.2. Symmetric Algorithm Preferences

The symmetric algorithm preference is an ordered list of algorithms that the keyholder accepts. Since it is found on a self-signature, it is possible that a keyholder may have multiple, different preferences. For example, Alice may have AES-128 only specified for "alice@work.com" but Camellia-256, Twofish, and AES-128 specified for "alice@home.org". Note that it is also possible for preferences to be in a subkey’s binding signature.

Since AES-128 is the MUST-implement algorithm, if it is not explicitly in the list, it is tacitly at the end. However, it is good form to place it there explicitly. Note also that if an implementation does not implement the preference, then it is implicitly an AES-128-only implementation. Note further that implementations conforming to previous versions of this standard [RFC4880] have TripleDES as its only MUST-implement algorithm.

An implementation MUST NOT use a symmetric algorithm that is not in the recipient’s preference list. When encrypting to more than one recipient, the implementation finds a suitable algorithm by taking the intersection of the preferences of the recipients. Note that the MUST-implement algorithm, AES-128, ensures that the intersection is not null. The implementation may use any mechanism to pick an algorithm in the intersection.

If an implementation can decrypt a message that a keyholder doesn’t have in their preferences, the implementation SHOULD decrypt the message anyway, but MUST warn the keyholder that the protocol has been violated. For example, suppose that Alice, above, has software that implements all algorithms in this specification. Nonetheless, she prefers subsets for work or home. If she is sent a message encrypted with IDEA, which is not in her preferences, the software warns her that someone sent her an IDEA-encrypted message, but it would ideally decrypt it anyway.

14.2.1. Plaintext

Algorithm 0, "plaintext", may only be used to denote secret keys that are stored in the clear. Implementations MUST NOT use plaintext in encrypted data packets; they must use Literal Data packets to encode unencrypted literal data.
14.3. Other Algorithm Preferences

Other algorithm preferences work similarly to the symmetric algorithm preference, in that they specify which algorithms the keyholder accepts. There are two interesting cases that other comments need to be made about, though, the compression preferences and the hash preferences.

14.3.1. Compression Preferences

Like the algorithm preferences, an implementation MUST NOT use an algorithm that is not in the preference vector. If Uncompressed (0) is not explicitly in the list, it is tacitly at the end. That is, uncompressed messages may always be sent.

Note that earlier implementations may assume that the absence of compression preferences means that [ZIP(1), Uncompressed(0)] are preferred, and default to ZIP compression. Therefore, an implementation that prefers uncompressed data SHOULD explicitly state this in the preferred compression algorithms.

14.3.1.1. Uncompressed

Algorithm 0, "uncompressed", may only be used to denote a preference for uncompressed data. Implementations MUST NOT use uncompressed in Compressed Data packets; they must use Literal Data packets to encode uncompressed literal data.

14.3.2. Hash Algorithm Preferences

Typically, the choice of a hash algorithm is something the signer does, rather than the verifier, because a signer rarely knows who is going to be verifying the signature. This preference, though, allows a protocol based upon digital signatures ease in negotiation.

Thus, if Alice is authenticating herself to Bob with a signature, it makes sense for her to use a hash algorithm that Bob’s software uses. This preference allows Bob to state in his key which algorithms Alice may use.

Since SHA2-256 is the MUST-implement hash algorithm, if it is not explicitly in the list, it is tacitly at the end. However, it is good form to place it there explicitly.
14.4. RSA

There are algorithm types for RSA Sign-Only, and RSA Encrypt-Only keys. These types are deprecated. The "key flags" subpacket in a signature is a much better way to express the same idea, and generalizes it to all algorithms. An implementation SHOULD NOT create such a key, but MAY interpret it.

An implementation SHOULD NOT implement RSA keys of size less than 1024 bits.

14.5. DSA

An implementation SHOULD NOT implement DSA keys of size less than 1024 bits. It MUST NOT implement a DSA key with a q size of less than 160 bits. DSA keys MUST also be a multiple of 64 bits, and the q size MUST be a multiple of 8 bits. The Digital Signature Standard (DSS) [FIPS186] specifies that DSA be used in one of the following ways:

* 1024-bit key, 160-bit q, SHA-1, SHA2-224, SHA2-256, SHA2-384, or SHA2-512 hash

* 2048-bit key, 224-bit q, SHA2-224, SHA2-256, SHA2-384, or SHA2-512 hash

* 2048-bit key, 256-bit q, SHA2-256, SHA2-384, or SHA2-512 hash

* 3072-bit key, 256-bit q, SHA2-256, SHA2-384, or SHA2-512 hash

The above key and q size pairs were chosen to best balance the strength of the key with the strength of the hash. Implementations SHOULD use one of the above key and q size pairs when generating DSA keys. If DSS compliance is desired, one of the specified SHA hashes must be used as well. [FIPS186] is the ultimate authority on DSS, and should be consulted for all questions of DSS compliance.

Note that earlier versions of this standard only allowed a 160-bit q with no truncation allowed, so earlier implementations may not be able to handle signatures with a different q size or a truncated hash.

14.6. Elgamal

An implementation SHOULD NOT implement Elgamal keys of size less than 1024 bits.
14.7. EdDSA

Although the EdDSA algorithm allows arbitrary data as input, its use with OpenPGP requires that a digest of the message is used as input (pre-hashed). See Section 5.2.4 for details. Truncation of the resulting digest is never applied; the resulting digest value is used verbatim as input to the EdDSA algorithm.

For clarity: while [RFC8032] describes different variants of EdDSA, OpenPGP uses the "pure" variant (PureEdDSA). The hashing that happens with OpenPGP is done as part of the standard OpenPGP signature process, and that hash itself is fed as the input message to the PureEdDSA algorithm.

As specified in [RFC8032], Ed448 also expects a "context string". In OpenPGP, Ed448 is used with the empty string as a context string.

14.8. Reserved Algorithm Numbers

A number of algorithm IDs have been reserved for algorithms that would be useful to use in an OpenPGP implementation, yet there are issues that prevent an implementer from actually implementing the algorithm. These are marked in Section 9.1 as "reserved for".

The reserved public-key algorithm X9.42 (21) does not have the necessary parameters, parameter order, or semantics defined. The same is currently true for reserved public-key algorithms AEDH (23) and AEDSA (24).

Previous versions of OpenPGP permitted Elgamal [ELGAMAL] signatures with a public-key identifier of 20. These are no longer permitted. An implementation MUST NOT generate such keys. An implementation MUST NOT generate Elgamal signatures. See [BLEICHENBACHER].

14.9. OpenPGP CFB Mode

When using a version 1 Symmetrically Encrypted Integrity Protected Data packet (Section 5.14.1) or --- for historic data --- a Symmetrically Encrypted Data packet (Section 5.8), OpenPGP does symmetric encryption using a variant of Cipher Feedback mode (CFB mode). This section describes the procedure it uses in detail. This mode is what is used for Symmetrically Encrypted Integrity Protected Data Packets (and the dangerously malleable --- and deprecated --- Symmetrically Encrypted Data Packets). Some mechanisms for encrypting secret-key material also use CFB mode, as described in Section 3.7.2.1.
In the description below, the value BS is the block size in octets of the cipher. Most ciphers have a block size of 8 octets. The AES and Twofish have a block size of 16 octets. Also note that the description below assumes that the IV and CFB arrays start with an index of 1 (unlike the C language, which assumes arrays start with a zero index).

OpenPGP CFB mode uses an initialization vector (IV) of all zeros, and prefixes the plaintext with BS+2 octets of random data, such that octets BS+1 and BS+2 match octets BS-1 and BS. It does a CFB resynchronization after encrypting those BS+2 octets.

Thus, for an algorithm that has a block size of 8 octets (64 bits), the IV is 10 octets long and octets 7 and 8 of the IV are the same as octets 9 and 10. For an algorithm with a block size of 16 octets (128 bits), the IV is 18 octets long, and octets 17 and 18 replicate octets 15 and 16. Those extra two octets are an easy check for a correct key.

Step by step, here is the procedure:

1. The feedback register (FR) is set to the IV, which is all zeros.
2. FR is encrypted to produce FRE (FR Encrypted). This is the encryption of an all-zero value.
3. FRE is xored with the first BS octets of random data prefixed to the plaintext to produce C[1] through C[BS], the first BS octets of ciphertext.
4. FR is loaded with C[1] through C[BS].
5. FR is encrypted to produce FRE, the encryption of the first BS octets of ciphertext.
6. The left two octets of FRE get xored with the next two octets of data that were prefixed to the plaintext. This produces C[BS+1] and C[BS+2], the next two octets of ciphertext.
7. (The resynchronization step) FR is loaded with C[3] through C[BS+2].
8. FR is encrypted to produce FRE.
9. FRE is xored with the first BS octets of the given plaintext, now that we have finished encrypting the BS+2 octets of prefixed data. This produces C[BS+3] through C[BS+(BS+2)], the next BS octets of ciphertext.
10. FR is loaded with C[BS+3] to C[BS + (BS+2)] (which is C11-C18 for an 8-octet block).

11. FR is encrypted to produce FRE.

12. FRE is xored with the next BS octets of plaintext, to produce the next BS octets of ciphertext. These are loaded into FR, and the process is repeated until the plaintext is used up.

14.10. Private or Experimental Parameters

S2K specifiers, Signature subpacket types, User Attribute types, image format types, and algorithms described in Section 9 all reserve the range 100 to 110 for private and experimental use. Packet types reserve the range 60 to 63 for private and experimental use. These are intentionally managed with the PRIVATE USE method, as described in [RFC8126].

However, implementations need to be careful with these and promote them to full IANA-managed parameters when they grow beyond the original, limited system.

14.11. Meta-Considerations for Expansion

If OpenPGP is extended in a way that is not backwards-compatible, meaning that old implementations will not gracefully handle their absence of a new feature, the extension proposal can be declared in the key holder’s self-signature as part of the Features signature subpacket.

We cannot state definitively what extensions will not be upwards-compatible, but typically new algorithms are upwards-compatible, whereas new packets are not.

If an extension proposal does not update the Features system, it SHOULD include an explanation of why this is unnecessary. If the proposal contains neither an extension to the Features system nor an explanation of why such an extension is unnecessary, the proposal SHOULD be rejected.

15. Security Considerations

* As with any technology involving cryptography, you should check the current literature to determine if any algorithms used here have been found to be vulnerable to attack.
* This specification uses Public-Key Cryptography technologies. It is assumed that the private key portion of a public-private key pair is controlled and secured by the proper party or parties.

* The MD5 hash algorithm has been found to have weaknesses, with collisions found in a number of cases. MD5 is deprecated for use in OpenPGP. Implementations MUST NOT generate new signatures using MD5 as a hash function. They MAY continue to consider old signatures that used MD5 as valid.

* SHA2-224 and SHA2-384 require the same work as SHA2-256 and SHA2-512, respectively. In general, there are few reasons to use them outside of DSS compatibility. You need a situation where one needs more security than smaller hashes, but does not want to have the full 256-bit or 512-bit data length.

* Many security protocol designers think that it is a bad idea to use a single key for both privacy (encryption) and integrity (signatures). In fact, this was one of the motivating forces behind the V4 key format with separate signature and encryption keys. If you as an implementer promote dual-use keys, you should at least be aware of this controversy.

* The DSA algorithm will work with any hash, but is sensitive to the quality of the hash algorithm. Verifiers should be aware that even if the signer used a strong hash, an attacker could have modified the signature to use a weak one. Only signatures using acceptably strong hash algorithms should be accepted as valid.

* As OpenPGP combines many different asymmetric, symmetric, and hash algorithms, each with different measures of strength, care should be taken that the weakest element of an OpenPGP message is still sufficiently strong for the purpose at hand. While consensus about the strength of a given algorithm may evolve, NIST Special Publication 800-57 [SP800-57] recommends the following list of equivalent strengths:
<table>
<thead>
<tr>
<th>Asymmetric key size</th>
<th>Hash size</th>
<th>Symmetric key size</th>
</tr>
</thead>
<tbody>
<tr>
<td>1024</td>
<td>160</td>
<td>80</td>
</tr>
<tr>
<td>2048</td>
<td>224</td>
<td>112</td>
</tr>
<tr>
<td>3072</td>
<td>256</td>
<td>128</td>
</tr>
<tr>
<td>7680</td>
<td>384</td>
<td>192</td>
</tr>
<tr>
<td>15360</td>
<td>512</td>
<td>256</td>
</tr>
</tbody>
</table>

Table 29: Key length equivalences

* There is a somewhat-related potential security problem in signatures. If an attacker can find a message that hashes to the same hash with a different algorithm, a bogus signature structure can be constructed that evaluates correctly.

For example, suppose Alice DSA signs message M using hash algorithm H. Suppose that Mallet finds a message M’ that has the same hash value as M with H’. Mallet can then construct a signature block that verifies as Alice’s signature of M’ with H’. However, this would also constitute a weakness in either H or H’ or both. Should this ever occur, a revision will have to be made to this document to revise the allowed hash algorithms.

* If you are building an authentication system, the recipient may specify a preferred signing algorithm. However, the signer would be foolish to use a weak algorithm simply because the recipient requests it.

* Some of the encryption algorithms mentioned in this document have been analyzed less than others. For example, although CAST5 is presently considered strong, it has been analyzed less than TripleDES. Other algorithms may have other controversies surrounding them.
In late summer 2002, Jallad, Katz, and Schneier published an interesting attack on older versions of the OpenPGP protocol and some of its implementations [JKS02]. In this attack, the attacker modifies a message and sends it to a user who then returns the erroneously decrypted message to the attacker. The attacker is thus using the user as a random oracle, and can often decrypt the message. This attack is a particular form of ciphertext malleability. See Section 15.1 for information on how to defend against such an attack using more recent versions of OpenPGP.

PKCS#1 has been found to be vulnerable to attacks in which a system that reports errors in padding differently from errors in decryption becomes a random oracle that can leak the private key in mere millions of queries. Implementations must be aware of this attack and prevent it from happening. The simplest solution is to report a single error code for all variants of decryption errors so as not to leak information to an attacker.

Some technologies mentioned here may be subject to government control in some countries.

In winter 2005, Serge Mister and Robert Zuccherato from Entrust released a paper describing a way that the "quick check" in OpenPGP CFB mode can be used with a random oracle to decrypt two octets of every cipher block [MZ05]. They recommend as prevention not using the quick check at all.

Many implementers have taken this advice to heart for any data that is symmetrically encrypted and for which the session key is public-key encrypted. In this case, the quick check is not needed as the public-key encryption of the session key should guarantee that it is the right session key. In other cases, the implementation should use the quick check with care.

On the one hand, there is a danger to using it if there is a random oracle that can leak information to an attacker. In plainer language, there is a danger to using the quick check if timing information about the check can be exposed to an attacker, particularly via an automated service that allows rapidly repeated queries.

On the other hand, it is inconvenient to the user to be informed that they typed in the wrong passphrase only after a petabyte of data is decrypted. There are many cases in cryptographic engineering where the implementer must use care and wisdom, and this is one.
* An implementation SHOULD only use an AES algorithm as a KEK algorithm, since backward compatibility of the ECDH format is not a concern. The KEK algorithm is only used within the scope of a Public-Key Encrypted Session Key Packet, which represents an ECDH key recipient of a message. Compare this with the algorithm used for the session key of the message, which MAY be different from a KEK algorithm.

Side channel attacks are a concern when a compliant application’s use of the OpenPGP format can be modeled by a decryption or signing oracle, for example, when an application is a network service performing decryption to unauthenticated remote users. ECC scalar multiplication operations used in ECDSA and ECDH are vulnerable to side channel attacks. Countermeasures can often be taken at the higher protocol level, such as limiting the number of allowed failures or time-blinding of the operations associated with each network interface. Mitigations at the scalar multiplication level seek to eliminate any measurable distinction between the ECC point addition and doubling operations.

* V5 signatures include a 128 bit salt that is hashed first. This makes OpenPGP signatures non-deterministic and protects against a broad class of attacks that depend on creating a signature over a predictable message. Hashing the salt first means that there is no attacker controlled hashed prefix. An example of this kind of attack is described in the paper SHA-1 Is A Shambles (see [SHAMBLES]), which leverages a chosen prefix collision attack against SHA-1.

15.1. Avoiding Ciphertext Malleability

If ciphertext can be modified by an attacker but still subsequently decrypted to some new plaintext, it is considered "malleable". A number of attacks can arise in any cryptosystem that uses malleable encryption, so modern OpenPGP offers mechanisms to defend against it. However, legacy OpenPGP data may have been created before these mechanisms were available. Because OpenPGP implementations deal with historic stored data, they may encounter malleable ciphertexts.

When an OpenPGP implementation discovers that it is decrypting data that appears to be malleable, it MUST indicate a clear error message that the integrity of the message is suspect, SHOULD NOT release decrypted data to the user, and SHOULD halt with an error. An implementation that encounters malleable ciphertext MAY choose to release cleartext to the user if it is known to be dealing with historic archived legacy data, and the user is aware of the risks.
Any of the following OpenPGP data elements indicate that malleable ciphertext is present:

* all Symmetrically Encrypted Data packets (Section 5.8).

* within any encrypted container, any Compressed Data packet (Section 5.7) where there is a decompression failure.

* any version 1 Symmetrically Encrypted Integrity Protected Data packet (Section 5.14.1) where the internal Modification Detection Code does not validate.

* any version 2 Symmetrically Encrypted Integrity Protected Data packet (Section 5.14.2) where the authentication tag of any chunk fails, or where there is no final zero-octet chunk.

* any Secret Key packet with encrypted secret key material (Section 3.7.2.1) where there is an integrity failure, based on the value of the secret key protection octet:
  - value 255 or raw cipher algorithm: where the trailing 2-octet checksum does not match.
  - value 254: where the SHA1 checksum is mismatched.
  - value 253: where the AEAD authentication tag is invalid.

To avoid these circumstances, an implementation that generates OpenPGP encrypted data SHOULD select the encrypted container format with the most robust protections that can be handled by the intended recipients. In particular:

* The SED packet is deprecated, and MUST NOT be generated.

* When encrypting to one or more public keys:
  - all recipient keys indicate support for version 2 of the Symmetrically Encrypted Integrity Protected Data packet in their Features subpacket (Section 5.2.3.29), or are v5 keys without a Features subpacket, or the implementation can otherwise infer that all recipients support v2 SEIPD packets, the implementation MUST encrypt using a v2 SEIPD packet.
  - if one of the recipients does not support v2 SEIPD packets, then the message generator MAY use a v1 SEIPD packet instead.
Password-protected secret key material in a V5 Secret Key or V5 Secret Subkey packet SHOULD be protected with AEAD encryption (S2K usage octet 253) unless it will be transferred to an implementation that is known to not support AEAD.

Implementers should implement AEAD (v2 SEIPD and S2K usage octet 253) promptly and encourage its spread.

Users should migrate to AEAD with all due speed.

15.2. Escrowed Revocation Signatures

A keyholder Alice may wish to designate a third party to be able to revoke Alice’s own key.

The preferred way for her to do this is produce a specific Revocation Signature (signature types 0x20, 0x28, or 0x30) and distribute it securely to her preferred revoker who can hold it in escrow. The preferred revoker can then publish the escrowed Revocation Signature at whatever time is deemed appropriate, rather than generating a revocation signature themselves.

There are multiple advantages of using an escrowed Revocation Signature over the deprecated Revocation Key subpacket (Section 5.2.3.20):

* The keyholder can constrain what types of revocation the preferred revoker can issue, by only escrowing those specific signatures.

* There is no public/visible linkage between the keyholder and the preferred revoker.

* Third parties can verify the revocation without needing to find the key of the preferred revoker.

* The preferred revoker doesn’t even need to have a public OpenPGP key if some other secure transport is possible between them and the keyholder.

* Implementation support for enforcing a revocation from an authorized Revocation Key subpacket is uneven and unreliable.

* If the fingerprint mechanism suffers a cryptanalytic flaw, the escrowed Revocation Signature is not affected.
A Revocation Signature may also be split up into shares and distributed among multiple parties, requiring some subset of those parties to collaborate before the escrowed Revocation Signature is recreated.

15.3. Random Number Generation and Seeding

OpenPGP requires a cryptographically secure pseudorandom number generator (CSPRNG). In most cases, the operating system provides an appropriate facility such as a getrandom() syscall, which should be used absent other (for example, performance) concerns. It is RECOMMENDED to use an existing CSPRNG implementation in preference to crafting a new one. Many adequate cryptographic libraries are already available under favorable license terms. Should those prove unsatisfactory, [RFC4086] provides guidance on the generation of random values.

OpenPGP uses random data with three different levels of visibility:

* in publicly-visible fields such as nonces, IVs, public padding material, or salts,

* in shared-secret values, such as session keys for encrypted data or padding material within an encrypted packet, and

* in entirely private data, such as asymmetric key generation.

With a properly functioning CSPRNG, this does not present a security problem, as it is not feasible to determine the CSPRNG state from its output. However, with a broken CSPRNG, it may be possible for an attacker to use visible output to determine the CSPRNG internal state and thereby predict less-visible data like keying material, as documented in [CHECKOWAY].

An implementation can provide extra security against this form of attack by using separate CSPRNGs to generate random data with different levels of visibility.

15.4. Traffic Analysis

When sending OpenPGP data through the network, the size of the data may leak information to an attacker. There are circumstances where such a leak could be unacceptable from a security perspective.
For example, if possible cleartext messages for a given protocol are known to be either yes (three octets) and no (two octets) and the messages are sent within a Symmetrically-Encrypted Integrity Protected Data packet, the length of the encrypted message will reveal the contents of the cleartext.

In another example, sending an OpenPGP Transferable Public Key over an encrypted network connection might reveal the length of the certificate. Since the length of an OpenPGP certificate varies based on the content, an external observer interested in metadata (who is trying to contact who) may be able to guess the identity of the certificate sent, if its length is unique.

In both cases, an implementation can adjust the size of the compound structure by including a Padding packet (see Section 5.15).

16. Implementation Nits

This section is a collection of comments to help an implementer, particularly with an eye to backward compatibility. Often the differences are small, but small differences are frequently more vexing than large differences. Thus, this is a non-comprehensive list of potential problems and gotchas for a developer who is trying to be backward-compatible.

* There are many ways possible for two keys to have the same key material, but different fingerprints (and thus Key IDs). For example, since a V4 fingerprint is constructed by hashing the key creation time along with other things, two V4 keys created at different times, yet with the same key material will have different fingerprints.

* OpenPGP does not put limits on the size of public keys. However, larger keys are not necessarily better keys. Larger keys take more computation time to use, and this can quickly become impractical. Different OpenPGP implementations may also use different upper bounds for public key sizes, and so care should be taken when choosing sizes to maintain interoperability.

* ASCII armor is an optional feature of OpenPGP. The OpenPGP working group strives for a minimal set of mandatory-to-implement features, and since there could be useful implementations that only use binary object formats, this is not a "MUST" feature for an implementation. For example, an implementation that is using OpenPGP as a mechanism for file signatures may find ASCII armor unnecessary. OpenPGP permits an implementation to declare what features it does and does not support, but ASCII armor is not one of these. Since most implementations allow binary and armored
objects to be used indiscriminately, an implementation that does not implement ASCII armor may find itself with compatibility issues with general-purpose implementations. Moreover, implementations of OpenPGP-MIME [RFC3156] already have a requirement for ASCII armor so those implementations will necessarily have support.

* What this document calls Legacy packet format Section 4.2.2 is what older documents called the "old packet format". It is the packet format of the legacy PGP 2 implementation. Older RFCs called the current OpenPGP packet format Section 4.2.1 the "new packet format".

17. References

17.1. Normative References


17.2. Informative References


Appendix A. Test vectors

To help implementing this specification a non-normative example for the EdDSA algorithm is given.

A.1. Sample EdDSA key

The secret key used for this example is:

D: 1a8b1ff05ded48e18bf50166c664ab023ea70003d78d9e41f5758a91d850f8d2

Note that this is the raw secret key used as input to the EdDSA signing operation. The key was created on 2014-08-19 14:28:27 and thus the fingerprint of the OpenPGP key is:

C959 BDBA FA32 A2F8 9A15 3B67 8CFD E121 9796 5A9A

The algorithm specific input parameters without the MPI length headers are:

oid: 2b06010401da470f01
q: 403f098994bd916ed4053197934e4a87c80733a1280d62f8010992e43ee3b2406

The entire public key packet is thus:

98 33 04 53 f3 5f 0b 16 09 2b 06 01 04 01 da 47
0f 01 01 07 40 3f 09 89 94 bd d9 16 ed 40 53 19
79 34 e4 a8 7c 80 73 3a 12 80 d6 2f 80 10 99 2e
43 ee 3b 24 06

A.2. Sample EdDSA signature

The signature is created using the sample key over the input data "OpenPGP" on 2015-09-16 12:24:53 and thus the input to the hash function is:

m: 4f70656e504750040016080006050255f95f9504ff0000000c
Using the SHA2-256 hash algorithm yields the digest:

d: f6220a3f757814f4c2176ffbb68b00249cd4ccdc059c4b34ad871f30b1740280

Which is fed into the EdDSA signature function and yields this signature:

r: 56f90cca98e2102637bd983fdbl6c131dfd27ed82bf4dde5606e0d756aed3366
s: d09c4fa11527f038e0f57f2201d82f2ea2c9033265fa6ceb489e854bae61b404

The entire signature packet is thus:

```
88 5e 04 00 16 08 00 06 05 02 55 f9 5f 95 00 0a
09 10 8c fd e1 21 97 96 5a 9a f6 22 01 00 56 f9
0c ca 98 e2 10 26 37 bd 98 3f db 16 c1 31 df d2
7e d8 2b f4 dd e5 60 6e 0d 75 6a ed 33 66 01 00
d0 9c 4f a1 15 27 f0 38 e0 f5 7f 22 01 d8 2f 2e
a2 c9 03 32 65 fa 6c eb 48 9e 85 4b ae 61 b4 04
```

A.3. Sample AEAD-EAX encryption and decryption

This example encrypts the cleartext string Hello, world! with the password password, using AES-128 with AEAD-EAX encryption.

A.3.1. Sample Parameters

S2K:

Iterated and Salted S2K

Iterations:

65011712 (255), SHA2-256

Salt:

a5 ae 57 9d 1f c5 d8 2b

A.3.2. Sample symmetric-key encrypted session key packet (v5)

Packet header:

c3 40

Version, algorithms, S2K fields:
A.3.3. Starting AEAD-EAX decryption of the session key

The derived key is:

```
15 49 67 e5 90 aa 1f 92 3e 1c 0a c6 4c 88 f2 3d
```

HKDF info:
```
c3 05 07 01
```

HKDF output:
```
74 f0 46 03 63 a7 00 76 db 08 c4 92 ab f2 95 52
```

Authenticated Data:
```
c3 05 07 01
```

Nonce:
```
69 22 4f 91 99 93 b3 50 6f a3 b5 9a 6a 73 cf f8
```

Decrypted session key:
```
38 81 ba fe 98 54 12 45 9b 86 c3 6f 98 cb 9a 5e
```

A.3.4. Sample v2 SEIPD packet

Packet header:
```
d2 69
```

Version, AES-128, EAX, Chunk size octet:
```
02 07 01 06
```
Salt:
9f f9 0e 3b 32 19 64 f3 a4 29 13 c8 dc c6 61 93
25 01 52 27 ef b7 ea ea a4 9f 04 c2 e6 74 17 5d

Chunk #0 encrypted data:
4a 3d 22 6e d6 af cb 9c a9 ac 12 2c 14 70 e1 1c
63 d4 c0 ab 24 1c 6a 93 8a d4 8b f9 9a 5a 99 b9
0b ba 83 25 de

Chunk #0 authentication tag:
61 04 75 40 25 8a b7 99 9a 95 ad 05 1d da 96 eb

Final (zero-sized chunk #1) authentication tag:
15 43 1d fe f5 f5 e2 25 5c a7 82 61 54 6e 33 9a

A.3.5. Decryption of data

Starting AEAD-EAX decryption of data, using the session key.

HKDF info:
   d2 02 07 01 06

HKDF output:
   b5 04 22 ac 1c 26 be 9d dd 83 1d 5b bb 36 b6 4f
   78 b8 33 f2 e9 4a 60 c0

Message key:
   b5 04 22 ac 1c 26 be 9d dd 83 1d 5b bb 36 b6 4f

Initialization vector:
   78 b8 33 f2 e9 4a 60 c0

Chunk #0:

Nonce:
   78 b8 33 f2 e9 4a 60 c0 00 00 00 00 00 00 00 00

Additional authenticated data:
Decrypted chunk #0.

Literal data packet with the string contents Hello, world!:

```
cb 13 62 00 00 00 00 00 48 65 6c 6c 6f 2c 20 77
6f 72 6c 64 21
```

Padding packet:

```
d5 0e ae 5b f0 cd 67 05 50 03 55 81 6c b0 c8 ff
```

Authenticating final tag:

Final nonce:

```
78 b8 33 f2 e9 4a 60 c0 00 00 00 00 00 00 00 01
```

Final additional authenticated data:

```
d2 02 07 01 06 00 00 00 00 00 00 00 25
```

A.3.6. Complete AEAD-EAX encrypted packet sequence

```
-----BEGIN PGP MESSAGE-----
w0AFHgcBCwMIpa5XnR/P2Cv/aSJPkZmTs1Bvo7WaanPP+Np0a4jjV+iuV0uH4dcF
ddcvYCMpkFl+mlkJSSJAAa+HD0mkCBwEGn/kO0zIZZPOkKRPI3MZhkyUBUIfvt+rq
pJ8EwuZOF11KPSJuig/LnKmsEiwUcOEcY9IAqyQcapOKl1vSmlqZuQu6gyXeYQR1
QCWKtSWala0PHdqW6xVDHf719e11XKeCYVRuM5o= =wG7F
-----END PGP MESSAGE-----
```

A.4. Sample AEAD-OCB encryption and decryption

This example encrypts the cleartext string Hello, world! with the password password, using AES-128 with AEAD-OCB encryption.

A.4.1. Sample Parameters

S2K:

Iterated and Salted S2K

Iterations:

```
65011712 (255), SHA2-256
```
Salt:
56 a2 98 d2 f5 e3 64 53

A.4.2. Sample symmetric-key encrypted session key packet (v5)

Packet header:

   c3 3f

Version, algorithms, S2K fields:
05 1d 07 02 0b 03 08 56 a2 98 d2 f5 e3 64 53 ff
cf cc

Nonce:

   cf cc 5c 11 66 4e db 9d b4 25 90 d7 dc 46 b0

Encrypted session key and AEAD tag:

   78 c5 c0 41 9c c5 1b 3a 46 87 cb 32 e5 b7 03 1c
e7 c6 69 75 76 5b 5c 21 d9 2a ef 4c c0 5c 3f ea

A.4.3. Starting AEAD-EAX decryption of the session key

The derived key is:

   e8 0d e2 43 a3 62 d9 3b 9d c6 07 ed e9 6a 73 56

HKDF info:

   c3 05 07 02

HKDF output:

   20 62 fb 76 31 ef be f4 df 81 67 ce d7 f3 a4 64

Authenticated Data:

   c3 05 07 02

Nonce:

   cf cc 5c 11 66 4e db 9d b4 25 90 d7 dc 46 b0

Decrypted session key:
A.4.4. Sample v2 SEIPD packet

Packet header:

d2 69

Version, AES-128, EAX, Chunk size octet:

02 07 02 06

Salt:

20 a6 61 f7 31 fc 9a 30 32 b5 62 33 26 02 7e 3a
5d 8d b5 74 8e be ff 0b 0c 59 10 d0 9e cd d6 41

Chunk #0 encrypted data:

ff 9f d3 85 62 75 80 35 bc 49 75 4c e1 bf 3f ff
a7 da d0 a3 b8 10 4f 51 33 cf 42 a4 10 0a 83 ee
f4 ca 1b 48 01

Chunk #0 authentication tag:

a8 84 6b f4 2b cd a7 c8 ce 9d 65 e2 12 f3 01 cb

Final (zero-sized chunk #1) authentication tag:

cd 98 fd ca de 69 4a 87 7a d4 24 73 23 f6 e8 57

A.4.5. Decryption of data

Starting AEAD-OCB decryption of data, using the session key.

HKDF info:

d2 02 07 02 06

HKDF output:

71 66 2a 11 ee 5b 4e 08 14 4e 6d e8 83 a0 09 99
eb de 12 bb 57 0d cf

Message key:

71 66 2a 11 ee 5b 4e 08 14 4e 6d e8 83 a0 09 99
A.4.6. Complete AEAD-EAX encrypted packet sequence

------BEGIN PGP MESSAGE------

wz8FHQcCCwMIVqKY0vXjZFP/z8xcEWZ02520JZDX3EaweMXAQZzFGzpGh8sy5bcD
HOfGaXV2W1wh2SrVTMBCPs+SaQlHAgYgpmH3MfyaMDK1YjMmA46XY21dI6+/wSM
WRDQns3wQf+fO4VidYA1vEl1TOG/P/+n2tCjuBBPUPFPqQCoPu9MobSAGohGv0
K82nym6dZeIS8wHLzZj9yt5pSod61CRzI/boVw==
=K/pk
------END PGP MESSAGE------
A.5. Sample AEAD-GCM encryption and decryption

This example encrypts the cleartext string Hello, world! with the password password, using AES-128 with AEAD-GCM encryption.

A.5.1. Sample Parameters

S2K:

Iterated and Salted S2K

Iterations:

65011712 (255), SHA2-256

Salt:

e9 d3 97 85 b2 07 00 08

A.5.2. Sample symmetric-key encrypted session key packet (v5)

Packet header:

c3 3c

Version, algorithms, S2K fields:

05 1a 07 03 0b 03 08 e9 d3 97 85 b2 07 00 08 ff
b4 2e

Nonce:

b4 2e 7c 48 3e f4 88 44 57 cb 37 26

Encrypted session key and AEAD tag:

0c 0c 4b f3 f2 cd 6c b7 b6 e3 8b 5b f3 34 67 c1
c7 19 44 dd 59 03 46 66 2f 5a de 61 ff 84 bc e0

A.5.3. Starting AEAD-EAX decryption of the session key

The derived key is:

25 02 81 71 5b ba 78 28 ef 71 ef 64 c4 78 47 53

HKDF info:

c3 05 07 03
HKDF output:
  de ec e5 81 8b c0 aa b9 0f 8a fb 02 fa 00 cd 13

Authenticated Data:
  c3 05 07 03

Nonce:
  b4 2e 7c 48 3e f4 88 44 57 cb 37 26

Decrypted session key:
  19 36 fc 85 68 98 02 74 bb 90 0d 83 19 36 0c 77

A.5.4. Sample v2 SEIPD packet

Packet header:
  d2 69

Version, AES-128, EAX, Chunk size octet:
  02 07 03 06

Salt:
  fc b9 44 90 bc b9 8b bd c9 d1 06 c6 09 02 66 94
  0f 72 e8 9e dc 21 b5 59 6b 15 76 b1 01 ed 0f 9f

Chunk #0 encrypted data:
  fc 6f c6 d6 5b bf d2 4d cd 07 90 96 6e 6d 1e 85
  a3 00 53 78 4c b1 d8 b6 a0 69 9e f1 21 55 a7 b2
  ad 62 58 53 1b

Chunk #0 authentication tag:
  57 65 1f d7 77 79 12 fa 95 e3 5d 9b 40 21 6f 69

Final (zero-sized chunk #1) authentication tag:
  a4 c2 48 db 28 ff 43 31 f1 63 29 07 39 9e 6f f9
A.5.5. Decryption of data

Starting AEAD-GCM decryption of data, using the session key.

HKDF info:
   d2 02 07 03 06

HKDF output:
   ea 14 38 80 3c b8 a4 77 40 ce 9b 54 c3 38 77 8d
   4d 2b dc 2b

Message key:
   ea 14 38 80 3c b8 a4 77 40 ce 9b 54 c3 38 77 8d

Initialization vector:
   4d 2b dc 2b

Chunk #0:

Nonce:
   4d 2b dc 2b 00 00 00 00 00 00 00 00

Additional authenticated data:
   d2 02 07 03 06

Decrypted chunk #0.

Literal data packet with the string contents Hello, world!:

   cb 13 62 00 00 00 00 00 48 65 6c 6c 6f 2c 20 77
   6f 72 6c 64 21

Padding packet:
   d5 0e 1c e2 26 9a 9e dd ef 81 03 21 72 b7 ed 7c

Authenticating final tag:

Final nonce:
   4d 2b dc 2b 00 00 00 00 00 00 00 01
Final additional authenticated data:

d2 02 07 03 06 00 00 00 00 00 00 00 25

A.5.6. Complete AEAD-EAX encrypted packet sequence

-----BEGIN PGP MESSAGE-----
wzwFGgcDCwMI6dOXhbIHAaj/tC58SD70iERXyzcmDAxL8/LNble244tb8zRnwccZ
RN1ZA0ZmL1reYf+EvdODsQIHAw8usUQVlMvcnRBSyJamaUD3lontwhtV1rFXax
Ae0pnxvxvtZbv9JNzQeQlm5tHoWjAFN4TLHYtqBpvnEhVaeyrWJYxtytXZR/Xd3kS
+PXjXztAIW99ppMJ12yyj/QzhXyYkHOZ5v+Q==
=C1Be
-----END PGP MESSAGE-----

A.6. Sample message encrypted using Argon2

These messages are the literal data "Hello, world!" encrypted using
Argon2 and the passphrase "password", using different session key
sizes. In all cases, the Argon2 parameters are t = 1, p = 4, and m =
21.

AES-128:

-----BEGIN PGP MESSAGE-----
Comment: Encrypted using AES with 128-bit key
Comment: Session key: 01FE16BBACFD1E7B78EF3B865187374F
wyceBwSUCvJg8j/1eUHUtA7N/zE2AQWn1L8rSLPP5Vqsunio+ECxHSPgGYK+y
YJz4ujF+DD1dB005NQRQXt/KJIf4m4mO1kyc/ujLbpnlJZMNq3o79gxBdtD0zH
XfA3pqV4mTzF
=uiks
-----END PGP MESSAGE-----

AES-192:

-----BEGIN PGP MESSAGE-----
Comment: Encrypted using AES with 192-bit key
Comment: Session key: 27006DAE68E509022CE45A14E569E91001C2955AF8D7FE194
wy8ECATHkxHFTJRzGk13K0NH4UP4AQQVHzLJ2va3FG8/pmp1Pd/H/mdoVS5VBLLw
F9I/adb1s56GYPY1ZjCVVKh+2bnq02s33AJJoyBexBI4QKATFRkbez2gldJldRys
LVg77MwWfgl2n/d572Wc1AM=
=n8Ma
-----END PGP MESSAGE-----

AES-256:

Koch & Wouters Expires 8 September 2022 [Page 149]
Comment: Encrypted using AES with 256-bit key
Comment: Session key: BBEDA55B9AAE63DAC45D4F49D89DACF4AF37FEFC13BAB2F1F8E18FB7458 0D8B0

wzcECQS4eJUgIG/3mcaILEJFpmj8AQVQv2917KtagdClm9UaQ/Z6M/5rok1SGpGu
623YmaKezGj8o4B+KuigTdo87X1wrup7l0wJypZls21Uwd67m9koF60eefH/K
95D1usliXOE8ayQJQmZrjf6Itv9PWWjMQ==

-----END PGP MESSAGE-----

Appendix B. Acknowledgements

This memo also draws on much previous work from a number of other authors, including: Derek Atkins, Charles Breed, Dave Del Torto, Marc Dyksterhouse, Gail Haspert, Gene Hoffman, Paul Hoffman, Ben Laurie, Raph Levien, Colin Plumb, Will Price, David Shaw, William Stallings, Mark Weaver, and Philip R. Zimmermann.

Appendix C. Document Workflow

This document is built from markdown using ruby-kramdown-rfc2629 (https://rubygems.org/gems/kramdown-rfc2629), and tracked using git (https://git-scm.com/). The markdown source under development can be found in the file crypto-refresh.md in the main branch of the git repository (https://gitlab.com/openpgp-wg/rfc4880bis). Discussion of this document should take place on the openpgp@ietf.org mailing list (https://www.ietf.org/mailman/listinfo/openpgp).

A non-substantive editorial nit can be submitted directly as a merge request (https://gitlab.com/openpgp-wg/rfc4880bis/-/merge_requests/new). A substantive proposed edit may also be submitted as a merge request, but should simultaneously be sent to the mailing list for discussion.

An open problem can be recorded and tracked as an issue (https://gitlab.com/openpgp-wg/rfc4880bis/-/issues) in the gitlab issue tracker, but discussion of the issue should take place on the mailing list.

[Note to RFC-Editor: Please remove this section on publication.]

Authors’ Addresses

Werner Koch (editor)
GnuPG e.V.
Rochusstr. 44
40479 Duesseldorf
Germany
Abstract

End-to-end encryption (E2EE) is an application of cryptography in communications systems between endpoints. E2EE systems are unique in providing features of confidentiality, integrity and authenticity for users. Improvements to E2EE strive to maximise the system's security while balancing usability and availability. Users of E2EE communications expect trustworthy providers of secure implementations to respect and protect their right to whisper.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on 8 September 2022.

Copyright Notice

Copyright (c) 2022 IETF Trust and the persons identified as the document authors. All rights reserved.
This document defines end-to-end encryption (E2EE) using three different dimensions that together comprise a full definition of E2EE, which can be applied in a variety of contexts.

The first is a formal definition that draws on the basic understanding of end points and cryptography. The second looks at E2EE systems from a design perspective, both its fundamental features and the direction of travel towards improving those features. Lastly we consider the expectations of the user of E2EE systems.
These dimensions taken as a whole comprise a generally comprehensible picture of consensus at the IETF as to what is end-to-end encryption, irrespective of application, from messaging to video conferencing, and between any number of end points.

2. Formal definition of end-to-end encryption

An end-to-end encrypted communications system, irrespective of the content or the specific methods employed, relies on two important and rigorous technical concepts: The end-to-end principle and what defines an end, according to the IETF because of its importance to internet protocols; and encryption, an application of cryptography and the primary means employed by the IETF to secure internet protocols. In the tradition of cryptography it’s also possible to achieve a succinct definition of end-to-end encrypted security.

2.1. End point

Intuitively, an "end" either sends messages or receives them, usually both; other systems on the path are just that - other systems.

It is, however, not trivial to establish the definition of an end point in isolation, because its existence inherently depends on at least one other entity in a communications system. That is why we will now move directly into an analysis of the end-to-end principle, which introduces nuance, described in the following sub-section.

However despite the nuance for engineers, it is now widely accepted that the communication system itself begins and ends with the user [RFC8890]. We imagine people (through an application’s user interface, or user agent) as components in a subsystem’s design. An important exception to this in E2EE systems might be the use of public key infrastructure where a third party is often used in the authentication phase to enhance the larger system’s trust model. Responsible use of of public key infrastructure is required in such cases, such that the E2EE system does not admit third parties under the user’s identity.

We cannot equate user agent and user, yet we also cannot fully separate them. As user-agent computing becomes more complex and often more proprietary, the user agent becomes less of an "advocate" for the best interests of the user. This is why we focus in a later section on the E2EE system being able to fulfill user expectations.
2.2. End-to-end principle

We need first to answer "What constitutes an end?", which is an important question in any review of the End-to-End Principle [RFC3724]. However the notion of an end point is more fully defined within the principle of end-to-end communications.

In 1984 the "end-to-end argument" was introduced [saltzer] as a design principle that helps guide placement of functions among the modules of a distributed computer system. It suggests that functions placed at low levels of a system may be redundant or of little value when compared with the cost of providing them at that low level. It is used to design around questions about which parts of the system should make which decisions, and as such the identity of the actual "speaker" or "end" may be less obvious than it appears. The communication described by Saltzer is between communicating processes, which may or may not be on the same physical machine, and may be implemented in various ways. For example, a BGP speaker is often implemented as a process that manages the Routing Information Base (RIB) and communicates with other BGP speakers using an operating system service that implements TCP. The RIB manager might find itself searching the RIB for prefixes that should be advertised to a peer, and performing "writes" to TCP for each one. TCP in this context often implements a variant of the algorithm described in RFC 868 (the "Nagle algorithm"), which accumulates writes in a buffer until there is no data in flight between the communicants, and then sends it – which might happen several times during a single search by the RIB manager. In that sense, the RIB manager might be thought of as the "end", because it decides what should be communicated, or TCP might be the "end", because it actually sends the TCP Segment, detects errors if they occur, retransmits it if necessary, and ultimately decides that the segment has been successfully transferred.
Another important question is "what statement exactly summarizes the end-to-end principle?". Saltzer answered this in two ways, the first of which is that the service implementing the transaction is most correct if it implements the intent of the application that sent it, which would be to move the message toward the destination address in the relevant IP header. Saltzer’s more thorough treatment, however, deals with end cases that come up in implementation: "Examples discussed in the paper", according to the abstract, "include bit error recovery, security using encryption, duplicate message suppression, recovery from system crashes, and delivery acknowledgement." It also notes that there is occasionally a rationale for ignoring the end-to-end arguments for the purposes of optimization. There may be other user expectations or design features, some explained below, which need to be balanced with the end-to-end argument.

More concisely, suppose that an end user is the end identity. An E2EE system may run between potential end points at different network layers within the end identity’s possession. These end points may then be considered acceptable sub-identities provided that no path between the end identity and sub-identity is accessible by any third party. There are quite a number of examples of common situations where tunnels are used and this does not apply. For instance, the examples below all provide encryption by which data is turned into clear text in locations that are not under control of the end user:

* The common VPNS business model whereby a TLS or an IPsec tunnel terminates at the service provider’s server and is subsequently forwarded to its destination elsewhere in unencrypted form;

* Email transport whereby an unencrypted message is traverses from sending mail user agent, between various mail transfer agents, and finally to the a receiving mail user agent, all over TLS protected connections;

* The encrypted connection of last mile connections such as those in 4G LTE;

This definition of end points accounts for potentially several devices owned by a user, and various application-specific forwarding or delivery options among them. It also accounts for E2EE systems running at different network layers. Regardless of the sub-identities allowed, the definition is contingent on that all end sub-identities are under the end identity’s control and no third party (or their sub-identities, e.g. system components under third-party control) can access the end sub-identities nor links between the sub-identity and end identity. This creates a tree hierarchy with the end user as the root at the top, and all potential end points being
under their direct control, without third party access. As an example, decryption at organizational network router before message forwarding (encrypted or unencrypted) to the end identity does not constitute E2EE. However, E2EE to a user’s personal device and subsequent E2EE message forwarding to another one of the user’s personal devices (without access available to any third party at any link or on device) maintains E2EE data possession for the user.

2.3. Encryption

From draft-dkg-hrpc-glossary-00, encryption is fundamental to the end-to-end principle. "End-to-End : The principal of extending characteristics of a protocol or system as far as possible within the system. For example, end-to-end instant message encryption would conceal communication content from one user’s instant messaging application through any intermediate devices and servers all the way to the recipient’s instant messaging application. If the message was decrypted at any intermediate point—for example at a service provider—then the property of end-to-end encryption would not be present."[dkg] Note that this only talks about the contents of the communication and not the metadata generated from it.

The way to achieve a truly end-to-end communications system is indeed to encrypt the content of the data exchanged between the endpoints, e.g. sender(s) and receiver(s). The more common end-to-end technique for encrypting uses a double-ratchet algorithm with an authenticated encryption scheme, present in many modern messenger applications such as those considered in the IETF Messaging Layer Security working group, whose charter is to create a document that satisfies the need for several Internet applications for group key establishment and message protection protocols [mls]. OpenPGP, mostly used for email, uses a different technique to achieve encryption. It is also chartered in the IETF to create a specification that covers object encryption, object signing, and identity certification [openpgp]. Both protocols rely on the use of asymmetric and symmetric encryption, and exchange public keys with amongst end points.

There are dozens of documents in the RFC Series that fundamentally and technically define encryption schemes. Perhaps interesting work to be done would be to survey all existing documents of this kind to define, in aggregate, their common features. The point is, the IETF has clear mandate and demonstrated expertise in defining the specifics of encrypted communications of the internet.

While encryption is fundamental to the end-to-end principle, it does not stand alone. As in the history of all security, authentication and data integrity properties are also linked, and contributed to the end-to-end nature of E2EE. Permission of data manipulation or
pseudo-identities for third parties to allow access under the user’s identity are against the intention of E2EE. Thus, endpoint authenticity must be established as (sub-)identities of the end user, and end-to-end integrity must also be maintained by the system. There is considerable system design flexibility available in entity authentication mechanisms and data authentication that still meet this requirement.

2.4. Succinct definition of end-to-end security

A succinct definition for end-to-end security can describe the security of the system by the probability of an adversary’s success in breaking the system. Example snippet:

The adversary successfully subverts an end-to-end encrypted system if it can succeed in either of the following: 1) the adversary can produce the participant’s local state (meaning the adversary has learned the contents of participant’s messages), or 2) the states of conversation participants do not match (meaning that the adversary has influenced their communication in some way). To prevent the adversary from trivially winning, we do not allow the adversary to compromise the participants’ local state.

We can say that a system is end-to-end secure if the adversary has negligible probability of success in either of these two scenarios [komlo].

3. End-to-end encrypted systems design

When looking at E2EE systems from a design perspective, the first consideration is the list of fundamental features that distinguish an E2EE system from one that does not employ E2EE. Secondly one must consider the direction of travel for improving the features of E2EE systems. In other words, what challenges are the designers, developers and implementers of E2EE systems facing?

The features and challenges listed below are framed holistically rather than from the perspective of their design, development, implementation or use.

3.1. Features

Defining a technology can also be done by inspecting what it does, or is meant to do, in the form of features. The features of end-to-end encryption from an implementation perspective can be inspected across several important categories: 1) the necessary features of E2EE of authenticity, confidentiality, and integrity, whereas features of 2) availability, deniability, forward secrecy, and post-compromise
security are enhancements to E2EE systems.

3.1.1. Necessary features

Authenticity  A system provides message authenticity if the recipient and sender agree on each other’s identities and the contents of their communications.

Confidentiality  A system provides message confidentiality if only the sender and intended recipient(s) can read the message plaintext, i.e. messages are encrypted by the sender such that only the intended recipient(s) can decrypt them.

Integrity  A system provides message integrity when it guarantees that messages that have been modified in transit can be detected reliably, i.e. a recipient is assured that a message cannot be undetectably modified in any way.

3.1.2. Optional/desirable features

Availability  A system provides high availability if the user is able to get to the message when they so desire and potentially from more than one device, i.e. a message arrives to a recipient even if they have been offline for a long time.

Deniability  Deniability ensures that anyone with a record of the transcript, including message recipients, cannot cryptographically prove to others that a particular participant of a communication authored the message. As demonstrated by the Signal and OTR protocols, this optional property must exist in conjunction with the necessary property of message authenticity, i.e. participants in a communication must be assured that they are communicating with the intended parties but this assurance cannot be proof to any other parties.

Forward secrecy  Forward secrecy is a security property that prevents attackers from decrypting encrypted data they have previously captured over a communication channel before the time of compromise, even if they have compromised one of the endpoints. Forward secrecy is usually achieved by updating the encryption/decryption keys, and older ones are deleted periodically.

Post-compromise security  Post-compromise security is a security
property that seeks to guarantee a way to recover from an end-
point compromise (and consequently that communication sent post-
compromise is protected with the same security properties that
existed before the compromise). It is usually achieved by adding
ephemeral key exchanges to the derivation of encryption/decryption
keys.

Metadata obfuscation Steps should be taken to minimize metadata such
as user obfuscating IP addresses, reducing non-routing metadata,
and avoiding extraneous message headers can enhance the
confidentiality and security features of E2EE systems.

3.2. Challenges

Earlier we defined end-to-end encryption using formal definitions
assumed by internet protocol implementations. Also because "the IETF
is a place for state-of-the-art producing high quality, relevant
technical documents that influence the way people design, use, and
manage the Internet" we can be confident that current deployments of
end-to-end encrypted technologies in the IETF indicate the cutting
edge of their developments, yet another way to define what is, or
ideally should be, how a technology is defined.

Below is an exhaustive, yet vaguely summarised, list of the
challenges currently faced by protocol designers of end-to-end
encrypted systems. In other words, in order to realise the goals of
end-to-end encrypted systems, both for users and implementers (see
previous section), these problems must be tackled. Problems that
fall outside of this list are likely 1) unnecessary feature requests
that negligibly, or do nothing to, achieve the aims of end-to-end
encrypted systems or are 2) in some way antithetical to the goals of
end-to-end encrypted systems.

Public key verification is very difficult for users to manage.
Authentication of the two ends is required for confidential
conversations. Therefore solving the problem of verification of
public keys is a major concern for any end-to-end encrypted system
design. Some applications bind together the account identity and the
key, and leave users to establish a trust relationship between them,
assisted by public key fingerprint information.

Users want to smoothly switch application use between devices, but
this comes at a cost to the security of user data. Thus, there is a
problem of availability in end-to-end encrypted systems because the
account identity’s private key is generated by and stored on the end-
user’s original device and to move the private key to another device
compromises the security of one of the end-points of the system.
Existing protocols are vulnerable to meta-data analysis, even though meta-data is often much more sensitive than content. Meta-data is plaintext information that travels across the wire and includes delivery-relevant details that central servers need such as the account identity of end-points, timestamps, message size. Meta-data is difficult to obfuscate efficiently.

Users need to communicate in groups, but this presents major problems of scale for end-to-end encryption systems that rely on public key cryptography.

The whole of a user’s data should remain secure if only one message is compromised. However, for encrypted communication, you must currently choose between forward secrecy or the ability to communicate asynchronously. This presents a problem for application design that uses end-to-end encryption for asynchronous messaging over email, RCS, etc.

Users of E2EE systems should be able to communicate with any medium of their choice, from text to large files, however there is often a resource problem because there are no open protocols to allow users to securely share the same resource in an end-to-end encrypted system. Client-side, e.g. end-point, activities like URL unfurling scanning.

Usability considerations are sometimes in conflict with security considerations, such as message read status, typing indicators, URL/link previews.

Deployment is notoriously challenging for any software application where maintenance and updates can be particularly disastrous for obsolete cryptographic libraries.

4. End-user expectations

While the formal definition and properties of an E2EE system relate to communication security, they do not draw from a comprehensive threat model or speak to what users expect from E2EE communication. It is in this context that some E2EE designs and architectures may ultimately run contrary to user expectations of E2EE systems [GEC-EU]. Although some system designs do not directly violate "the math" of encryption algorithms, they do so by implicating and weakening other important aspects of an E2EE _system_.

4.1. A conversation is confidential

Users talking to one another in an E2EE system should be the only ones that know what they are talking about [RFC7624]. People have the right to data privacy as defined in international human rights law and within the right to free expression and to hold opinions is inferred the right to whisper, whether or not they are using digital communications or walking through a field.

4.2. Providers are trustworthy

While "trustworthy" can be rigourously defined from an engineering perspective, for the purposes of this document we choose a definition of Trustworthy inspired by an internal workshop by Internet Society staff:

Trustworthy  A system is completely trustworthy if and only if it is completely resilient, reliable, accountable, and secure in a way that consistently meets users’ expectations. The opposite of trustworthy is untrustworthy.

This definition is complete in its positive and negative aspects: what it is, e.g. "Worthy of confidence" and what it is not, e.g. in RFC 7258: "behavior that subverts the intent of communicating parties without the agreement of those parties" [RFC7258].

Therefore, a trustworthy end-to-end encrypted communication system is the set of functions needed by two or more parties to communicate among each other in a confidential and authenticated fashion without any third party having access to the content of that communication where the functions that offer the confidentiality and authenticity are trustworthy.

4.3. Access by a third-party is impossible

No matter the specifics, any methods used to access to the content of the messages by a third party would violate a user’s expectations of E2EE messaging. "[T]hese access methods scan message contents on the user’s [device]", which are then "scanned for matches against a database of prohibited content before, and sometimes after, the message is sent to the recipient" [GEC-EU]. Third party access also covers cases without scanning - namely, it should not be possible for any third-party end point to access the data regardless of reason.
If a method makes private communication, intended to be sent over an encrypted channel between end points, available to parties other than the sender and intended recipient(s), without formally interfering with channel confidentiality, that method violates the understood expectation of that security property.

4.4. Pattern inference is minimised

Analyses such as traffic fingerprinting or other (encrypted or unencrypted) data analysis techniques should be considered outside the scope of an E2EE system’s goals of providing secure communications to end users.

Such methods of analyses, outside of or as part of E2EE system design, allow third parties to draw inferences from communication that was intended to be confidential. "By allowing private user data to be scanned via direct access by servers and their providers," the use of these methods should be considered an affront to "the privacy expectations of users of end-to-end encrypted communication systems" [GEC-EU].

Not only should an E2EE system value user data privacy by not enabling pattern inference, it should actively be attempting to solve issues of metadata and traceability (enhanced metadata) through further innovation that stays ahead of advances in these techniques.

4.5. The E2EE system is not compromised

RFC 3552 talks about the Internet Threat model such as the assumption that the user can expect any communications systems, but perhaps especially E2EE systems, to not be intentionally compromised [RFC3552]. Intentional compromises of E2EE systems are often referred to as "backdoors" but are often presented as additional design features under terms like "key escrow." Users of E2EE systems would not expect a front, back or side door entrance into their confidential conversations and would expect a provider to actively resist - technically and legally - compromise through these means.

5. Conclusions

From messaging to video conferencing, there are many competing features in an E2EE system that is secure and usable. The most well designed system cannot meet the expectations of every user, nor does an ideal system exist from any dimension. E2EE is a technology that is constantly improving to achieve the ideal as defined in this document.
Features and functionalities of E2EE systems should be developed and improved in service of end user expectations for privacy preserving communications.

6. Acknowledgements

Fred Baker, Stephen Farrell, Richard Barnes, Olaf Kolkman all contributed to the early strategic thinking of this document and whether it would be useful to the IETF community.

The folks at Riseup and the LEAP Encryption Access Project have articulated brilliantly the hardest parts of end-to-end encryption systems that serve the end users’ right to whisper.

Ryan Polk at the Internet Society has energy to spare when it comes to organising meaningful contributions, like this one, for the technical advisors of the Global Encryption Coalition.

Adrian Farrel, are acknowledged for their review, comments, or questions that lead to improvement of this document.

7. Security Considerations

This document does not specify new protocols and therefore does not bring up technical security considerations.

Because some policy decisions may affect the security of the Internet, a clear and shared definition of end to end encrypted communication is important in policy related discussions. This document aims to provide that clarity.

8. IANA Considerations

This document has no actions for IANA.

9. Informative References


Authors’ Addresses

Mallory Knodel
CDT
Email: mknodel@cdt.org