

lwig  
Internet-Draft  
Intended status: Informational  
Expires: May 10, 2019

R. Struik  
Struik Security Consultancy  
November 06, 2018

Alternative Elliptic Curve Representations  
draft-ietf-lwig-curve-representations-01

Abstract

This document specifies how to represent Montgomery curves and (twisted) Edwards curves as curves in short-Weierstrass form and illustrates how this can be used to carry out elliptic curve computations using existing implementations of, e.g., ECDSA and ECDH using NIST prime curves.

Requirements Language

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 RFC 2119 [RFC2119].

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 May 10, 2019.

Copyright Notice

Copyright (c) 2018 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 Provisions Relating to IETF Documents

(<https://trustee.ietf.org/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 Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

## Table of Contents

1. Fostering Code Reuse with New Elliptic Curves . . . . .	3
2. Specification of Wei25519 . . . . .	3
3. Use of Representation Switches . . . . .	4
4. Examples . . . . .	5
4.1. Implementation of X25519 . . . . .	5
4.2. Implementation of Ed25519 . . . . .	5
4.3. Specification of ECDSA-SHA256-Wei25519 . . . . .	5
4.4. Other Uses . . . . .	6
5. Caveats . . . . .	6
6. Security Considerations . . . . .	7
7. Privacy Considerations . . . . .	8
8. IANA Considerations . . . . .	8
9. Acknowledgements . . . . .	8
10. References . . . . .	8
10.1. Normative References . . . . .	8
10.2. Informative References . . . . .	9
Appendix A. Some (non-Binary) Elliptic Curves . . . . .	10
A.1. Curves in short-Weierstrass Form . . . . .	10
A.2. Montgomery Curves . . . . .	10
A.3. Twisted Edwards Curves . . . . .	10
Appendix B. Elliptic Curve Nomenclature . . . . .	11
Appendix C. Elliptic Curve Group Operations . . . . .	11
C.1. Group Law for Weierstrass Curves . . . . .	11
C.2. Group Law for Montgomery Curves . . . . .	12
C.3. Group Law for Twisted Edwards Curves . . . . .	13
Appendix D. Relationship Between Curve Models . . . . .	14
D.1. Mapping between twisted Edwards Curves and Montgomery Curves . . . . .	14
D.2. Mapping between Montgomery Curves and Weierstrass Curves . . . . .	14
D.3. Mapping between twisted Edwards Curves and Weierstrass Curves . . . . .	15
Appendix E. Curve25519 and Cousins . . . . .	15
E.1. Curve Definition and Alternative Representations . . . . .	15
E.2. Switching between Alternative Representations . . . . .	16
E.3. Domain Parameters . . . . .	17
Appendix F. Further Mappings . . . . .	19
F.1. Isomorphic Mapping between Weierstrass Curves . . . . .	19
F.2. Isogenous Mapping between Weierstrass Curves . . . . .	20

Appendix G. Further Cousins of Curve25519 . . . . .	21
G.1. Further Alternative Representations . . . . .	21
G.2. Further Switching . . . . .	22
G.3. Further Domain Parameters . . . . .	23
Appendix H. Isogeny Details . . . . .	24
H.1. Isogeny Parameters . . . . .	24
H.1.1. Coefficients of $u(x)$ . . . . .	24
H.1.2. Coefficients of $v(x)$ . . . . .	26
H.1.3. Coefficients of $w(x)$ . . . . .	29
H.2. Dual Isogeny Parameters . . . . .	30
H.2.1. Coefficients of $u'(x)$ . . . . .	30
H.2.2. Coefficients of $v'(x)$ . . . . .	32
H.2.3. Coefficients of $w'(x)$ . . . . .	35
Author's Address . . . . .	36

## 1. Fostering Code Reuse with New Elliptic Curves

It is well-known that elliptic curves can be represented using different curve models. Recently, IETF standardized elliptic curves that are claimed to have better performance and improved robustness against "real world" attacks than curves represented in the traditional "short" Weierstrass model. This document specifies an alternative representation of points of Curve25519, a so-called Montgomery curve, and of points of Edwards25519, a so-called twisted Edwards curve, which are both specified in [RFC7748], as points of a specific so-called "short" Weierstrass curve, called Wei25519. We also define how to efficiently switch between these different representations.

Use of Wei25519 allows easy definition of new signature schemes and key agreement schemes already specified for traditional NIST prime curves, thereby allowing easy integration with existing specifications, such as NIST SP 800-56a [SP-800-56a], FIPS Pub 186-4 [FIPS-186-4], and ANSI X9.62-2005 [ANSI-X9.62], and fostering code reuse on platforms that already implement some of these schemes using elliptic curve arithmetic for curves in "short" Weierstrass form (see Appendix C.1).

## 2. Specification of Wei25519

For the specification of Wei25519 and its relationship to Curve25519 and Edwards25519, see Appendix E. For further details and background information on elliptic curves, we refer to the other appendices.

The use of Wei25519 allows reuse of existing generic code that implements short-Weierstrass curves, such as the NIST curve P256, to also implement the CFRG curves Curve25519 and Edwards25519. We also cater to reusing of existing code where some domain parameters may

have been hardcoded, thereby widening the scope of applicability. To this end, we specify the short-Weierstrass curves Wei25519.2 and Wei25519.-3, with hardcoded domain parameter  $a=2$  and  $a=-3 \pmod{p}$ , respectively; see Appendix G.

### 3. Use of Representation Switches

The curves Curve25519, Edwards25519, and Wei25519, as specified in Appendix E.3, are all isomorphic, with the transformations of Appendix E.2. These transformations map the specified base point of each of these curves to the specified base point of each of the other curves. Consequently, a public-key pair  $(k, R:=k*G)$  for any one of these curves corresponds, via these isomorphic mappings, to the public-key pair  $(k, R':=k*G')$  for each of these other curves (where  $G$  and  $G'$  are the corresponding base points of these curves). This observation extends to the case where one also considers curve Wei25519.2 (which has hardcoded domain parameter  $a=2$ ), as specified in Appendix G.3, since it is isomorphic to Wei25519, with the transformation of Appendix G.2, and, thereby, also isomorphic to Curve25519 and Edwards25519.

The curve Wei25519.-3 (which has hardcoded domain parameter  $a=-3 \pmod{p}$ ) is not isomorphic to the curve Wei25519, but is related in a slightly weaker sense: the curve Wei25519 is isogenous to the curve Wei25519.-3, where the mapping of Appendix G.2 is an isogeny of degree  $l=47$  that maps the specified base point  $G$  of Wei25519 to the specified base point  $G'$  of Wei25519.-3 and where the so-called dual isogeny (which maps Wei25519.-3 to Wei25519) has the same degree  $l=47$ , but does not map  $G'$  to  $G$ , but to a fixed multiple hereof, where this multiple is  $l=47$ . Consequently, a public-key pair  $(k, R:=k*G)$  for Wei25519 corresponds to the public-key pair  $(k, R':=k*G')$  for Wei25519.-3 (via the  $l$ -isogeny), whereas the public-key pair  $(k, R':=k*G')$  corresponds to the public-key pair  $(l*k, l*R=l*k*G)$  of Wei25519 (via the dual isogeny). (Note the extra scalar  $l=47$  here.)

Alternative curve representations can, therefore, be used in any cryptographic scheme that involves computations on public-private key pairs, where implementations may carry out computations on the corresponding object for the isomorphic or isogenous curve and convert the results back to the original curve (where, in case this involves an  $l$ -isogeny, one has to take into account the factor  $l$ ). This includes use with elliptic-curve based signature schemes and key agreement and key transport schemes.

## 4. Examples

### 4.1. Implementation of X25519

RFC 7748 [RFC7748] specifies the use of X25519, a co-factor Diffie-Hellman key agreement scheme, with instantiation by the Montgomery curve Curve25519. This key agreement scheme was already specified in Section 6.1.2.2 of NIST SP 800-56a [SP-800-56a] for elliptic curves in short Weierstrass form. Hence, one can implement X25519 using existing NIST routines by (1) representing a point of the Montgomery curve Curve25519 as a point of the Weierstrass curve Wei25519; (2) instantiating the co-factor Diffie-Hellman key agreement scheme of the NIST specification with the resulting point and Wei25519 domain parameters; (3) representing the key resulting from this scheme (which is a point of the curve Wei25519 in Weierstrass form) as a point of the Montgomery curve Curve25519. The representation change can be implemented via a simple wrapper and involves a single modular addition (see Appendix D.2). Using this method has the additional advantage that one can reuse the public-private key pair routines, domain parameter validation, and other checks that are already part of the NIST specifications. Note: at this point, it is unclear whether this implies that a FIPS-accredited module implementing co-factor Diffie-Hellman for, e.g., P-256 would also extend this accreditation to X25519.

### 4.2. Implementation of Ed25519

RFC 8032 [RFC8032] specifies Ed25519, a "full" Schnorr signature scheme, with instantiation by the twisted Edwards curve Edwards25519. One can implement the computation of the ephemeral key pair for Ed25519 using an existing Montgomery curve implementation by (1) generating a public-private key pair  $(k, R' := k * G')$  for Curve25519; (2) representing this public-private key as the pair  $(k, R := k * G)$  for Ed25519. As before, the representation change can be implemented via a simple wrapper. Note that the Montgomery ladder specified in Section 5 of RFC7748 [RFC7748] does not provide sufficient information to reconstruct  $R'$  (since it does not compute the y-coordinate of  $R'$ ). However, this deficiency can be remedied by using a slightly modified version of the Montgomery ladder that includes reconstruction of the y-coordinate of  $R' := k * G'$  at the end of hereof (which uses the v-coordinate  $G_v$  of the base point of Curve25519 as well). For details, see Appendix D.2.

### 4.3. Specification of ECDSA-SHA256-Wei25519

FIPS Pub 186-4 [FIPS-186-4] specifies the signature scheme ECDSA and can be instantiated not just with the NIST prime curves, but also with other Weierstrass curves (that satisfy additional cryptographic

criteria). In particular, one can instantiate this scheme with the Weierstrass curve Wei25519 and the hash function SHA-256, where an implementation may generate a public-private key pair for Wei25519 by (1) internally carrying out these computations on the Montgomery curve Curve25519, the twisted Edwards curve Edwards25519, or even the Weierstrass curve Wei25519.-3 (with hardcoded  $a=-3$  domain parameter); (2) representing the result as a key pair for the curve Wei25519. Note that, in either case, one can implement these schemes with the same representation conventions as used with existing NIST specifications, including bit/byte-ordering, compression functions, and the-like. This allows implementations of ECDSA with the hash function SHA-256 and with the NIST curve P-256 or with the curve Wei25519 specified in this draft to use the same implementation (instantiated with, respectively, the NIST P-256 elliptic curve domain parameters or with the domain parameters of curve Wei25519 specified in Appendix E).

#### 4.4. Other Uses

Any existing specification of cryptographic schemes using elliptic curves in Weierstrass form and that allows introduction of a new elliptic curve (here: Wei25519) is amenable to similar constructs, thus spawning "offspring" protocols, simply by instantiating these using the new curve in "short" Weierstrass form, thereby allowing code and/or specifications reuse and, for implementations that so desire, carrying out curve computations "under the hood" on Montgomery curve and twisted Edwards curve cousins hereof (where these exist). This would simply require definition of a new object identifier for any such envisioned "offspring" protocol. This could significantly simplify standardization of schemes and help keeping the resource and maintenance cost of implementations supporting algorithm agility [RFC7696] at bay.

#### 5. Caveats

The examples above illustrate how specifying the Weierstrass curve Wei25519 may facilitate reuse of existing code and may simplify standards development. However, the following caveats apply:

1. Unfriendly wire formats. The transformations between alternative curve representations can be implemented at negligible relative incremental cost if the curve points are represented as affine points. If a point is represented in compressed format, conversion usually requires a costly point decompression step. This is the case in [RFC7748], where the inputs to the co-factor Diffie-Hellman scheme X25519, as well as its output, are represented in x-coordinate-only format;

2. Unfriendly representation conventions. While elliptic curve computations are carried-out in a field  $GF(q)$  and, thereby, involve large integer arithmetic, these integers are represented as bit- and byte-strings. Here, [RFC8032] uses least-significant-byte (LSB)/least-significant-bit (lsb) conventions, whereas [RFC7748] uses LSB/most-significant-bit (msb) conventions, and where most other cryptographic specifications, including NIST SP800-56a [SP-800-56a], FIPS Pub 186-4 [FIPS-186-4], and ANSI X9.62-2005 [ANSI-X9.62] use MSB/msb conventions. Since each pair of conventions is different, this does necessitate bit/byte representation conversions;
3. Unfriendly domain parameters. All traditional NIST curves are Weierstrass curve with domain parameter  $a=-3$ , while all Brainpool curves [RFC5639] are isomorphic to a Weierstrass curve of this form. Thus, one can expect there to be existing Weierstrass implementations with a hardcoded  $a=-3$  domain parameter ("Jacobian-friendly"). For those implementations, including the curve Wei25519 as a potential vehicle for offering support for the CFRG curves Curve25519 and Edwards25519 is not possible, since not of the required form. Instead, one has to implement Curve25519.-3 and include code that implements the isogeny and dual isogeny from and to Wei25519. This isogeny has degree  $l=47$  and requires roughly 9kB of storage for isogeny and dual-isogeny computations (see the tables in Appendix H). Note that storage would have reduced to a single 32-byte table if the curve would have been generated so as to be isomorphic to a Weierstrass curve with hardcoded  $a=-3$  parameter (this corresponds to  $l=1$ ). Note: an example of such a curve is the Montgomery curve  $M_{\{A,B\}}$  over  $GF(p)$  with  $p=2^{255}-19$ ,  $A=-1410290$ , and  $B=1$  or (if one wants the base point to still have  $u$ -coordinate  $u=9$ )  $A=-3960846$ . In either case, the resulting curve has the same cryptographic properties as Curve25519, while being more "Jacobian-friendly".

## 6. Security Considerations

The different representations of elliptic curve points discussed in this document are all obtained using a publicly known transformation, which is either an isomorphism or a low-degree isogeny. It is well-known that an isomorphism maps elliptic curve points to equivalent mathematical objects and that the complexity of cryptographic problems (such as the discrete logarithm problem) of curves related via a low-degree isogeny are tightly related. Thus, the use of these techniques does not negatively impact cryptographic security.

As to implementation security, reusing existing high-quality code or generic implementations that have been carefully designed to withstand implementation attacks for one curve model may allow a more

economical way of development and maintenance than providing this same functionality for each curve model separately (if multiple curve models need to be supported) and, otherwise, may allow a more gradual migration path, where one may initially use existing and accredited chipsets that cater to the pre-dominant curve model used in practice for over 15 years.

## 7. Privacy Considerations

The transformations between different curve models described in this document are publicly known and, therefore, do not affect privacy provisions.

## 8. IANA Considerations

There is *\*currently\** no IANA action required for this document. New object identifiers would be required in case one wishes to specify one or more of the "offspring" protocols exemplified in Section 4.

## 9. Acknowledgements

Thanks to Nikolas Rosener for discussions surrounding implementation details of the techniques described in this document and to Phillip Hallam-Baker for triggering inclusion of verbiage on the use of Montgomery ladders with recovery of the y-coordinate.

## 10. References

### 10.1. Normative References

[ANSI-X9.62]

ANSI X9.62-2005, "Public Key Cryptography for the Financial Services Industry: The Elliptic Curve Digital Signature Algorithm (ECDSA)", American National Standard for Financial Services, Accredited Standards Committee X9, Inc, Anapolis, MD, 2005.

[FIPS-186-4]

FIPS 186-4, "Digital Signature Standard (DSS), Federal Information Processing Standards Publication 186-4", US Department of Commerce/National Institute of Standards and Technology, Gaithersburg, MD, July 2013.

[RFC2119]

Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <<https://www.rfc-editor.org/info/rfc2119>>.



- [RFC5639] Lochter, M. and J. Merkle, "Elliptic Curve Cryptography (ECC) Brainpool Standard Curves and Curve Generation", RFC 5639, DOI 10.17487/RFC5639, March 2010, <<https://www.rfc-editor.org/info/rfc5639>>.
- [RFC7696] Housley, R., "Guidelines for Cryptographic Algorithm Agility and Selecting Mandatory-to-Implement Algorithms", BCP 201, RFC 7696, DOI 10.17487/RFC7696, November 2015, <<https://www.rfc-editor.org/info/rfc7696>>.
- [RFC7748] Langley, A., Hamburg, M., and S. Turner, "Elliptic Curves for Security", RFC 7748, DOI 10.17487/RFC7748, January 2016, <<https://www.rfc-editor.org/info/rfc7748>>.
- [RFC8032] Josefsson, S. and I. Liusvaara, "Edwards-Curve Digital Signature Algorithm (EdDSA)", RFC 8032, DOI 10.17487/RFC8032, January 2017, <<https://www.rfc-editor.org/info/rfc8032>>.
- [SP-800-56a] NIST SP 800-56a, "Recommendation for Pair-Wise Key Establishment Schemes Using Discrete Log Cryptography, Revision 2", US Department of Commerce/National Institute of Standards and Technology, Gaithersburg, MD, June 2013.

## 10.2. Informative References

- [ECC] I.F. Blake, G. Seroussi, N.P. Smart, "Elliptic Curves in Cryptography", Cambridge University Press, Lecture Notes Series 265, July 1999.
- [ECC-Isogeny] E. Brier, M. Joye, "Fast Point Multiplication on Elliptic Curves through Isogenies", AAECC, Lecture Notes in Computer Science, Vol. 2643, New York: Springer-Verlag, 2003.
- [GECC] D. Hankerson, A.J. Menezes, S.A. Vanstone, "Guide to Elliptic Curve Cryptography", New York: Springer-Verlag, 2004.
- [HW-ECC] W.P. Liu, "How to Use the Kinets LTC ECC HW to Accelerate Curve25519 (version 7)", NXP, <https://community.nxp.com/docs/DOC-330199>, April 2017.

## Appendix A. Some (non-Binary) Elliptic Curves

### A.1. Curves in short-Weierstrass Form

Let  $GF(q)$  denote the finite field with  $q$  elements, where  $q$  is an odd prime power and where  $q$  is not divisible by three. Let  $W_{\{a,b\}}$  be the Weierstrass curve with defining equation  $y^2 = x^3 + a*x + b$ , where  $a$  and  $b$  are elements of  $GF(q)$  and where  $4*a^3 + 27*b^2$  is nonzero. The points of  $W_{\{a,b\}}$  are the ordered pairs  $(x, y)$  whose coordinates are elements of  $GF(q)$  and that satisfy the defining equation (the so-called affine points), together with the special point  $O$  (the so-called "point at infinity"). This set forms a group under addition, via the so-called "secant-and-tangent" rule, where the point at infinity serves as the identity element. See Appendix C.1 for details of the group operation.

### A.2. Montgomery Curves

Let  $GF(q)$  denote the finite field with  $q$  elements, where  $q$  is an odd prime power. Let  $M_{\{A,B\}}$  be the Montgomery curve with defining equation  $B*v^2 = u^3 + A*u^2 + u$ , where  $A$  and  $B$  are elements of  $GF(q)$  with  $A$  unequal to  $(+/-)2$  and with  $B$  nonzero. The points of  $M_{\{A,B\}}$  are the ordered pairs  $(u, v)$  whose coordinates are elements of  $GF(q)$  and that satisfy the defining equation (the so-called affine points), together with the special point  $O$  (the so-called "point at infinity"). This set forms a group under addition, via the so-called "secant-and-tangent" rule, where the point at infinity serves as the identity element. See Appendix C.2 for details of the group operation.

### A.3. Twisted Edwards Curves

Let  $GF(q)$  denote the finite field with  $q$  elements, where  $q$  is an odd prime power. Let  $E_{\{a,d\}}$  be the twisted Edwards curve with defining equation  $a*x^2 + y^2 = 1 + d*x^2*y^2$ , where  $a$  and  $d$  are distinct nonzero elements of  $GF(q)$ . The points of  $E_{\{a,d\}}$  are the ordered pairs  $(x, y)$  whose coordinates are elements of  $GF(q)$  and that satisfy the defining equation (the so-called affine points). It can be shown that this set forms a group under addition if  $a$  is a square in  $GF(q)$ , whereas  $d$  is not, where the point  $(0, 1)$  serves as the identity element. (Note that the identity element satisfies the defining equation.) See Appendix C.3 for details of the group operation.

An Edwards curve is a twisted Edwards curve with  $a=1$ .

## Appendix B. Elliptic Curve Nomenclature

Each curve defined in Appendix A forms a commutative group under addition. In Appendix C we specify the group laws, which depend on the curve model in question. For completeness, we here include some common elliptic curve nomenclature and basic properties (primarily so as to keep this document self-contained). These notions are mainly used in Appendix E and Appendix G and not essential for our exposition. This section can be skipped at first reading.

Any point  $P$  of a curve is a generator of the cyclic subgroup  $(P) := \{k \cdot P \mid k = 0, 1, 2, \dots\}$  of the curve. If  $(P)$  has cardinality  $l$ , then  $l$  is called the order of  $P$ . The order of a curve is the cardinality of the set of its points. A curve is cyclic if it is generated by some point of this curve. All curves of prime order are cyclic, while all curves of order  $|E| = h \cdot n$ , where  $n$  is a large prime number and where  $h$  is a small number (the so-called co-factor), have a large cyclic subgroup of prime order  $n$ . In this case, a generator of order  $n$  is called a base point, commonly denoted by  $G$ . A point of order dividing  $h$  is said to be in the small subgroup. For curves of prime order, this small subgroup is the singleton set, consisting of only the identity element.

If  $R$  is a point on a curve  $E$  that is also contained in  $(P)$ , there is a unique integer  $k$  in the interval  $[0, l-1]$  so that  $R = kP$ , where  $l$  is the order of  $P$ . This number is called the discrete logarithm of  $R$  to the base  $P$ . The discrete logarithm problem is the problem of finding the discrete logarithm of  $R$  to the base  $P$  for any two points  $P$  and  $R$  of the curve, if such a number exists.

A public-private key pair is an ordered pair  $(k, R := kG)$ , where  $G$  is a fixed base point of the curve. Here,  $k$  (the private key) is an integer in the interval  $[0, n-1]$ , where  $G$  has order  $n$ .

A quadratic twist of a curve  $E$  defined over a field  $GF(q)$  is a curve  $E'$  related to  $E$ , with cardinality  $|E| + |E'| = 2 \cdot (q+1)$ . If  $E$  is a curve in one of the curve models specified in this document, a quadratic twist of this curve can be expressed using the same curve model, although (naturally) with different curve parameters.

## Appendix C. Elliptic Curve Group Operations

### C.1. Group Law for Weierstrass Curves

For each point  $P$  of the Weierstrass curve  $W_{\{a,b\}}$ , the point at infinity  $O$  serves as identity element, i.e.,  $P + O = O + P = P$ .

For each affine point  $P := (x, y)$  of the Weierstrass curve  $W_{\{a,b\}}$ , the point  $-P$  is the point  $(x, -y)$  and one has  $P + (-P) = O$ .

Let  $P_1 := (x_1, y_1)$  and  $P_2 := (x_2, y_2)$  be distinct affine points of the Weierstrass curve  $W_{\{a,b\}}$  and let  $Q := P_1 + P_2$ , where  $Q$  is not the identity element. Then  $Q := (x, y)$ , where

$$x + x_1 + x_2 = \lambda^2 \text{ and } y + y_1 = \lambda(x_1 - x), \text{ where } \lambda = (y_2 - y_1)/(x_2 - x_1).$$

Let  $P := (x_1, y_1)$  be an affine point of the Weierstrass curve  $W_{\{a,b\}}$  and let  $Q := 2P$ , where  $Q$  is not the identity element. Then  $Q := (x, y)$ , where

$$x + 2x_1 = \lambda^2 \text{ and } y + y_1 = \lambda(x_1 - x), \text{ where } \lambda = (3x_1^2 + a)/(2y_1).$$

From the group law above it follows that if  $P = (x, y)$ ,  $P_1 = kP = (x_1, y_1)$ , and  $P_2 = (k+1)P = (x_2, y_2)$  are affine points of the Weierstrass curve  $W_{\{a,b\}}$  and if  $y$  is nonzero, then the  $y$ -coordinate of  $P_1$  can be expressed in terms of the  $x$ -coordinates of  $P$ ,  $P_1$ , and  $P_2$ , and the  $y$ -coordinate of  $P$ , as

$$y_1 = ((x \cdot x_1 + a) \cdot (x + x_1) + 2b - x_2 \cdot (x - x_1)^2) / (2y).$$

This property allows recovery of the  $y$ -coordinate of a point  $P_1 = kP$  that is computed via the so-called Montgomery ladder, where  $P$  is an affine point with nonzero  $y$ -coordinate. Further details are out of scope.

## C.2. Group Law for Montgomery Curves

For each point  $P$  of the Montgomery curve  $M_{\{A,B\}}$ , the point at infinity  $O$  serves as identity element, i.e.,  $P + O = O + P = P$ .

For each affine point  $P := (x, y)$  of the Montgomery curve  $M_{\{A,B\}}$ , the point  $-P$  is the point  $(x, -y)$  and one has  $P + (-P) = O$ .

Let  $P_1 := (x_1, y_1)$  and  $P_2 := (x_2, y_2)$  be distinct affine points of the Montgomery curve  $M_{\{A,B\}}$  and let  $Q := P_1 + P_2$ , where  $Q$  is not the identity element. Then  $Q := (x, y)$ , where

$$x + x_1 + x_2 = B\lambda^2 - A \text{ and } y + y_1 = \lambda(x_1 - x), \text{ where } \lambda = (y_2 - y_1)/(x_2 - x_1).$$

Let  $P := (x_1, y_1)$  be an affine point of the Montgomery curve  $M_{\{A,B\}}$  and let  $Q := 2P$ , where  $Q$  is not the identity element. Then  $Q := (x, y)$ , where

$x + 2*x_1 = B*\lambda^2 - A$  and  $y + y_1 = \lambda*(x_1 - x)$ , where  $\lambda = (3*x_1^2 + 2*A*x_1 + 1)/(2*B*y_1)$ .

From the group law above it follows that if  $P=(x, y)$ ,  $P_1=k*P=(x_1, y_1)$ , and  $P_2=(k+1)*P=(x_2, y_2)$  are affine points of the Montgomery curve  $M_{\{A,B\}}$  and if  $y$  is nonzero, then the  $y$ -coordinate of  $P_1$  can be expressed in terms of the  $x$ -coordinates of  $P$ ,  $P_1$ , and  $P_2$ , and the  $y$ -coordinate of  $P$ , as

$$y_1 = ((x*x_1 + 1)*(x + x_1 + 2*A) - 2*A - x_2*(x - x_1)^2) / (2*B*y).$$

This property allows recovery of the  $y$ -coordinate of a point  $P_1=k*P$  that is computed via the so-called Montgomery ladder, where  $P$  is an affine point with nonzero  $y$ -coordinate. Further details are out of scope.

### C.3. Group Law for Twisted Edwards Curves

Note: The group laws below hold for twisted Edwards curves  $E_{\{a,d\}}$  where  $a$  is a square in  $GF(q)$ , whereas  $d$  is not. In this case, the addition formulae below are defined for each pair of points, without exceptions. Generalizations of this group law to other twisted Edwards curves are out of scope.

For each point  $P$  of the twisted Edwards curve  $E_{\{a,d\}}$ , the point  $O=(0,1)$  serves as identity element, i.e.,  $P + O = O + P = P$ .

For each point  $P:=(x, y)$  of the twisted Edwards curve  $E_{\{a,d\}}$ , the point  $-P$  is the point  $(-x, y)$  and one has  $P + (-P) = O$ .

Let  $P_1:=(x_1, y_1)$  and  $P_2:=(x_2, y_2)$  be points of the twisted Edwards curve  $E_{\{a,d\}}$  and let  $Q:=P_1 + P_2$ . Then  $Q:=(x, y)$ , where

$$x = (x_1*y_2 + x_2*y_1)/(1 + d*x_1*x_2*y_1*y_2) \text{ and } y = (y_1*y_2 - a*x_1*x_2)/(1 - d*x_1*x_2*y_1*y_2).$$

Let  $P:=(x_1, y_1)$  be a point of the twisted Edwards curve  $E_{\{a,d\}}$  and let  $Q:=2*P$ . Then  $Q:=(x, y)$ , where

$$x = (2*x_1*y_1)/(1 + d*x_1^2*y_1^2) \text{ and } y = (y_1^2 - a*x_1^2)/(1 - d*x_1^2*y_1^2).$$

Note that one can use the formulae for point addition for implementing point doubling, taking inverses and adding the identity element as well (i.e., the point addition formulae are uniform and complete (subject to our Note above)).

## Appendix D. Relationship Between Curve Models

The non-binary curves specified in Appendix A are expressed in different curve models, viz. as curves in short-Weierstrass form, as Montgomery curves, or as twisted Edwards curves. These curve models are related, as follows.

### D.1. Mapping between twisted Edwards Curves and Montgomery Curves

One can map points of the Montgomery curve  $M_{\{A,B\}}$  to points of the twisted Edwards curve  $E_{\{a,d\}}$ , where  $a := (A+2)/B$  and  $d := (A-2)/B$  and, conversely, map points of the twisted Edwards curve  $E_{\{a,d\}}$  to points of the Montgomery curve  $M_{\{A,B\}}$ , where  $A := 2(a+d)/(a-d)$  and where  $B := 4/(a-d)$ . For twisted Edwards curves we consider (i.e., those where  $a$  is a square in  $GF(q)$ , whereas  $d$  is not), this defines a one-to-one correspondence, which – in fact – is an isomorphism between  $M_{\{A,B\}}$  and  $E_{\{a,d\}}$ , thereby showing that, e.g., the discrete logarithm problem in either curve model is equally hard.

For the Montgomery curves and twisted Edwards curves we consider, the mapping from  $M_{\{A,B\}}$  to  $E_{\{a,d\}}$  is defined by mapping the point at infinity  $O$  and the point  $(0, 0)$  of order two of  $M_{\{A,B\}}$  to, respectively, the point  $(0, 1)$  and the point  $(0, -1)$  of order two of  $E_{\{a,d\}}$ , while mapping each other point  $(u, v)$  of  $M_{\{A,B\}}$  to the point  $(x, y) := (u/v, (u-1)/(u+1))$  of  $E_{\{a,d\}}$ . The inverse mapping from  $E_{\{a,d\}}$  to  $M_{\{A,B\}}$  is defined by mapping the point  $(0, 1)$  and the point  $(0, -1)$  of order two of  $E_{\{a,d\}}$  to, respectively, the point at infinity  $O$  and the point  $(0, 0)$  of order two of  $M_{\{A,B\}}$ , while each other point  $(x, y)$  of  $E_{\{a,d\}}$  is mapped to the point  $(u, v) := ((1+y)/(1-y), (1+y)/((1-y)*x))$  of  $M_{\{A,B\}}$ .

Implementations may take advantage of this mapping to carry out elliptic curve group operations originally defined for a twisted Edwards curve on the corresponding Montgomery curve, or vice-versa, and translating the result back to the original curve, thereby potentially allowing code reuse.

### D.2. Mapping between Montgomery Curves and Weierstrass Curves

One can map points of the Montgomery curve  $M_{\{A,B\}}$  to points of the Weierstrass curve  $W_{\{a,b\}}$ , where  $a := (3-A^2)/(3*B^2)$  and  $b := (2*A^3-9*A)/(27*B^3)$ . This defines a one-to-one correspondence, which – in fact – is an isomorphism between  $M_{\{A,B\}}$  and  $W_{\{a,b\}}$ , thereby showing that, e.g., the discrete logarithm problem in either curve model is equally hard.

The mapping from  $M_{\{A,B\}}$  to  $W_{\{a,b\}}$  is defined by mapping the point at infinity  $O$  of  $M_{\{A,B\}}$  to the point at infinity  $O$  of  $W_{\{a,b\}}$ , while

mapping each other point  $(u, v)$  of  $M_{\{A,B\}}$  to the point  $(x, y) := (u/B + A/(3*B), v/B)$  of  $W_{\{a,b\}}$ . Note that not all Weierstrass curves can be injectively mapped to Montgomery curves, since the latter have a point of order two and the former may not. In particular, if a Weierstrass curve has prime order, such as is the case with the so-called "NIST curves", this inverse mapping is not defined.

This mapping can be used to implement elliptic curve group operations originally defined for a twisted Edwards curve or for a Montgomery curve using group operations on the corresponding elliptic curve in short-Weierstrass form and translating the result back to the original curve, thereby potentially allowing code reuse.

Note that implementations for elliptic curves with short-Weierstrass form that hard-code the domain parameter  $a$  to  $a = -3$  (which value is known to allow more efficient implementations) cannot always be used this way, since the curve  $W_{\{a,b\}}$  resulting from an isomorphic mapping cannot always be expressed as a Weierstrass curve with  $a = -3$  via a coordinate transformation. For more details, see Appendix F.

### D.3. Mapping between twisted Edwards Curves and Weierstrass Curves

One can map points of the twisted Edwards curve  $E_{\{a,d\}}$  to points of the Weierstrass curve  $W_{\{a,b\}}$ , via function composition, where one uses the isomorphic mapping between twisted Edwards curve and Montgomery curves of Appendix D.1 and the one between Montgomery and Weierstrass curves of Appendix D.2. Obviously, one can use function composition (now using the respective inverses) to realize the inverse of this mapping.

## Appendix E. Curve25519 and Cousins

### E.1. Curve Definition and Alternative Representations

The elliptic curve Curve25519 is the Montgomery curve  $M_{\{A,B\}}$  defined over the prime field  $GF(p)$ , with  $p := 2^{255} - 19$ , where  $A := 486662$  and  $B := 1$ . This curve has order  $h*n$ , where  $h=8$  and where  $n$  is a prime number. For this curve,  $A^2-4$  is not a square in  $GF(p)$ , whereas  $A+2$  is. The quadratic twist of this curve has order  $h_1*n_1$ , where  $h_1=4$  and where  $n_1$  is a prime number. For this curve, the base point is the point  $(G_u, G_v)$ , where  $G_u=9$  and where  $G_v$  is an odd integer in the interval  $[0, p-1]$ .

This curve has the same group structure as (is "isomorphic" to) the twisted Edwards curve  $E_{\{a,d\}}$  defined over  $GF(p)$ , with as base point the point  $(G_x, G_y)$ , where parameters are as specified in Appendix E.3. This curve is denoted as Edwards25519. For this curve, the parameter

$a$  is a square in  $\text{GF}(p)$ , whereas  $d$  is not, so the group laws of Appendix C.3 apply.

The curve is also isomorphic to the elliptic curve  $W_{\{a,b\}}$  in short-Weierstrass form defined over  $\text{GF}(p)$ , with as base point the point  $(Gx', Gy')$ , where parameters are as specified in Appendix E.3. This curve is denoted as Wei25519.

## E.2. Switching between Alternative Representations

Each affine point  $(u, v)$  of Curve25519 corresponds to the point  $(x, y) := (u + A/3, y)$  of Wei25519, while the point at infinity of Curve25519 corresponds to the point at infinity of Wei25519. (Here, we used the mapping of Appendix D.2.) Under this mapping, the base point  $(Gu, Gv)$  of Curve25519 corresponds to the base point  $(Gx', Gy')$  of Wei25519. The inverse mapping maps the affine point  $(x, y)$  of Wei25519 to  $(u, v) := (x - A/3, y)$  of Curve25519, while mapping the point at infinity of Wei25519 to the point at infinity of Curve25519. Note that this mapping involves a simple shift of the first coordinate and can be implemented via integer-only arithmetic as a shift of  $(p+A)/3$  for the isomorphic mapping and a shift of  $-(p+A)/3$  for its inverse, where  $\text{delta} = (p+A)/3$  is the element of  $\text{GF}(p)$  defined by

```
delta 19298681539552699237261830834781317975544997444273427339909597
      334652188435537
```

```
(=0x2aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaad2
451)
```

The curve Edwards25519 is isomorphic to the curve Curve25519, where the base point  $(Gu, Gv)$  of Curve25519 corresponds to the base point  $(Gx, Gy)$  of Edwards25519 and where the point at infinity and the point  $(0, 0)$  of order two of Curve25519 correspond to, respectively, the point  $(0, 1)$  and the point  $(0, -1)$  of order two of Edwards25519 and where each other point  $(u, v)$  of Curve25519 corresponds to the point  $(c*u/v, (u-1)/(u+1))$  of Edwards25519, where  $c$  is the element of  $\text{GF}(p)$  defined by

```
c  sqrt(-(A+2))
```

```
51042569399160536130206135233146329284152202253034631822681833788
666877215207
```

```
(=0x70d9120b 9f5ff944 2d84f723 fc03b081 3a5e2c2e b482e57d
3391fb55 00ba81e7)
```

(Here, we used the mapping of Appendix D.1.) The inverse mapping from Edwards25519 to Curve25519 is defined by mapping the point  $(0,$



1) and the point  $(0, -1)$  of order two of Edwards25519 to, respectively, the point at infinity and the point  $(0,0)$  of order two of Curve25519 and having each other point  $(x, y)$  of Edwards25519 correspond to the point  $((1 + y)/(1 - y), c*(1 + y)/((1-y)*x))$ .

The curve Edwards25519 is isomorphic to the Weierstrass curve Wei25519, where the base point  $(Gx, Gy)$  of Edwards25519 corresponds to the base point  $(Gx', Gy')$  of Wei25519 and where the identity element  $(0,1)$  and the point  $(0,-1)$  of order two of Edwards25519 correspond to, respectively, the point at infinity  $O$  and the point  $(A/3, 0)$  of order two of Wei25519 and where each other point  $(x, y)$  of Edwards25519 corresponds to the point  $(x', y') := ((1+y)/(1-y)+A/3, c*(1+y)/((1-y)*x))$  of Wei25519, where  $c$  was defined before. (Here, we used the mapping of Appendix D.3.) The inverse mapping from Wei25519 to Edwards25519 is defined by mapping the point at infinity  $O$  and the point  $(A/3, 0)$  of order two of Wei25519 to, respectively, the identity element  $(0,1)$  and the point  $(0,-1)$  of order two of Edwards25519 and having each other point  $(x, y)$  of Wei25519 correspond to the point  $(c*(3*x-A)/(3*y), (3*x-A-3)/(3*x-A+3))$ .

Note that these mappings can be easily realized in projective coordinates, using a few field multiplications only, thus allowing switching between alternative representations with negligible relative incremental cost.

### E.3. Domain Parameters

The parameters of the Montgomery curve and the corresponding isomorphic curves in twisted Edwards curve and short-Weierstrass form are as indicated below. Here, the domain parameters of the Montgomery curve Curve25519 and of the twisted Edwards curve Edwards25519 are as specified in RFC 7748; the domain parameters of Wei25519 are "new".

General parameters (for all curve models):

p  $2^{255}-19$

(=0x7ffffffff ffffffff ffffffff ffffffff ffffffff ffffffff  
ffffffff ffffffff)

h 8

n 72370055773322622139731865630429942408571163593799076060019509382  
85454250989

(= $2^{252}$  + 0x14def9de a2f79cd6 5812631a 5cf5d3ed)

h1 4

n1 14474011154664524427946373126085988481603263447650325797860494125  
407373907997

(=2^{253} - 0x29bdf3bd 45ef39ac b024c634 b9eba7e3)

Montgomery curve-specific parameters (for Curve25519):

A 486662

B 1

Gu 9 (=0x9)

Gv 14781619447589544791020593568409986887264606134616475288964881837  
755586237401

(=0x20ae19a1 b8a086b4 e01edd2c 7748d14c 923d4d7e 6d7c61b2  
29e9c5a2 7eced3d9)

Twisted Edwards curve-specific parameters (for Edwards25519):

a -1 (-0x01)

d -121665/121666

(=370957059346694393431380835087545651895421138798432190163887855  
33085940283555)

(=0x52036cee 2b6ffe73 8cc74079 7779e898 00700a4d 4141d8ab  
75eb4dca 135978a3)

Gx 15112221349535400772501151409588531511454012693041857206046113283  
949847762202

(=0x216936d3 cd6e53fe c0a4e231 fdd6dc5c 692cc760 9525a7b2  
c9562d60 8f25d51a)

Gy 4/5

(=463168356949264781694283940034751631413079938662562256157830336  
03165251855960)

(=0x66666666 66666666 66666666 66666666 66666666 66666666  
66666666 66666658)

Weierstrass curve-specific parameters (for Wei25519):

```

a  19298681539552699237261830834781317975544997444273427339909597334
    573241639236

    (=0x2aaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa
    aaaaaa98 4914a144)

b  55751746669818908907645289078257140818241103727901012315294400837
    956729358436

    (=0x7b425ed0 97b425ed 097b425e d097b425 ed097b42 5ed097b4
    260b5e9c 7710c864)

Gx' 19298681539552699237261830834781317975544997444273427339909597334
    652188435546

    (=0x2aaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa
    aaaaaaaa aaad245a)

Gy' 14781619447589544791020593568409986887264606134616475288964881837
    755586237401

    (=0x20ae19a1 b8a086b4 e01edd2c 7748d14c 923d4d7e 6d7c61b2
    29e9c5a2 7eced3d9)

```

## Appendix F. Further Mappings

The non-binary curves specified in Appendix A are expressed in different curve models, viz. as curves in short-Weierstrass form, as Montgomery curves, or as twisted Edwards curves. Within each curve model, further mappings exist that induce a mapping between elliptic curves within each curve model. This can be exploited to force some of the domain parameters to a value that allows a more efficient implementation of the addition formulae.

### F.1. Isomorphic Mapping between Weierstrass Curves

One can map points of the Weierstrass curve  $W_{\{a,b\}}$  to points of the Weierstrass curve  $W_{\{a',b'\}}$ , where  $a:=a'*s^4$  and  $b:=b'*s^6$  for some nonzero value  $s$  of the finite field  $GF(q)$ . This defines a one-to-one correspondence, which - in fact - is an isomorphism between  $W_{\{a,b\}}$  and  $W_{\{a',b'\}}$ , thereby showing that, e.g., the discrete logarithm problem in either curve model is equally hard.

The mapping from  $W_{\{a,b\}}$  to  $W_{\{a',b'\}}$  is defined by mapping the point at infinity  $O$  of  $W_{\{a,b\}}$  to the point at infinity  $O$  of  $W_{\{a',b'\}}$ , while mapping each other point  $(x, y)$  of  $W_{\{a,b\}}$  to the point  $(x', y') := (x*s^2, y*s^3)$  of  $W_{\{a',b'\}}$ . The inverse mapping from  $W_{\{a',b'\}}$  to  $W_{\{a,b\}}$  is defined by mapping the point at infinity  $O$  of  $W_{\{a',b'\}}$

to the point at infinity  $O$  of  $W_{\{a,b\}}$ , while mapping each other point  $(x', y')$  of  $W_{\{a',b'\}}$  to the point  $(x, y) := (x/s^2, y/s^3)$  of  $W_{\{a,b\}}$ .

Implementations may take advantage of this mapping to carry out elliptic curve group operations originally defined for a Weierstrass curve with a generic domain parameter  $a$  on a corresponding isomorphic Weierstrass curve with domain parameter  $a'$  that has a special form, which is known to allow for more efficient implementations of addition laws, and translating the result back to the original curve. In particular, it is known that such efficiency improvements exist if  $a' \equiv -3 \pmod{p}$  and one uses so-called Jacobian coordinates with a particular projective version of the addition laws of Appendix C.1. While not all Weierstrass curves can be put into this form, all traditional NIST curves have domain parameter  $a \equiv -3$ , while all Brainpool curves [RFC5639] are isomorphic to a Weierstrass curve of this form.

Note that implementations for elliptic curves with short-Weierstrass form that hard-code the domain parameter  $a$  to  $a \equiv -3$  cannot always be used this way, since the curve  $W_{\{a,b\}}$  cannot always be expressed in terms of a Weierstrass curve with  $a' \equiv -3$  via a coordinate transformation: this only holds if  $a'/a$  is a fourth power in  $GF(q)$  (see Section 3.1.5 of [GECC]). However, even in this case, one can still express the curve  $W_{\{a,b\}}$  as a Weierstrass curve with a small domain parameter value  $a'$ , thereby still allowing a more efficient implementation than with a general domain parameter value  $a$ .

## F.2. Isogenous Mapping between Weierstrass Curves

One can still map points of the Weierstrass curve  $W_{\{a,b\}}$  to points of the Weierstrass curve  $W_{\{a',b'\}}$ , where  $a' \equiv -3 \pmod{p}$ , even if  $a'/a$  is not a fourth power in  $GF(q)$ . In that case, this mapping cannot be an isomorphism (see Appendix F.1). Instead, the mapping is a so-called isogeny (or homomorphism). Since most elliptic curve operations process points of prime order or use so-called "co-factor multiplication", in practice the resulting mapping has similar properties as an isomorphism. In particular, one can still take advantage of this mapping to carry out elliptic curve group operations originally defined for a Weierstrass curve with domain parameter  $a$  unequal to  $-3 \pmod{p}$  on a corresponding isogenous Weierstrass curve with domain parameter  $a' \equiv -3 \pmod{p}$  and translating the result back to the original curve.

In this case, the mapping from  $W_{\{a,b\}}$  to  $W_{\{a',b'\}}$  is defined by mapping the point at infinity  $O$  of  $W_{\{a,b\}}$  to the point at infinity  $O$  of  $W_{\{a',b'\}}$ , while mapping each other point  $(x, y)$  of  $W_{\{a,b\}}$  to the point  $(x', y') := (u(x)/w(x)^2, y*v(x)/w(x)^3)$  of  $W_{\{a',b'\}}$ . Here,  $u(x)$ ,  $v(x)$ , and  $w(x)$  are polynomials that depend on the isogeny in

question. The inverse mapping from  $W_{\{a',b'\}}$  to  $W_{\{a,b\}}$  is again an isogeny and defined by mapping the point at infinity  $O$  of  $W_{\{a',b'\}}$  to the point at infinity  $O$  of  $W_{\{a,b\}}$ , while mapping each other point  $(x', y')$  of  $W_{\{a',b'\}}$  to the point  $(x, y) := (u'(x')/w'(x')^2, y' * v'(x')/w'(x')^3)$  of  $W_{\{a,b\}}$ , where -- again --  $u'(x')$ ,  $v'(x')$ , and  $w'(x')$  are polynomials that depend on the isogeny in question. These mappings have the property that their composition is not the identity mapping (as is the case with the isomorphic mappings discussed in Appendix F.1), but rather a fixed multiple hereof: if this multiple is  $l$  then the isogeny is called an isogeny of degree  $l$  (or  $l$ -isogeny) and  $u$ ,  $v$ , and  $w$  (and, similarly,  $u'$ ,  $v'$ , and  $w'$ ) are polynomials of degrees  $l$ ,  $3(l-1)/2$ , and  $(l-1)/2$ , respectively. Note that an isomorphism is simply an isogeny of degree  $l=1$ . Details of how to determine isogenies are outside scope of this document (for this, contact the author of this document).

Implementations may take advantage of this mapping to carry out elliptic curve group operations originally defined for a Weierstrass curve with a generic domain parameter  $a$  on a corresponding isogenous Weierstrass curve with domain parameter  $a'=-3 \pmod{p}$ , where one can use so-called Jacobian coordinates with a particular projective version of the addition laws of Appendix C.1. Since all traditional NIST curves have domain parameter  $a=-3$ , while all Brainpool curves [RFC5639] are isomorphic to a Weierstrass curve of this form, this allows taking advantage of existing implementations for these curves that may have a hardcoded  $a=-3 \pmod{p}$  domain parameter, provided one switches back and forth to this curve form using the isogenous mapping in question.

Note that isogenous mappings can be easily realized in projective coordinates and involves roughly  $3 \times l$  finite field multiplications, thus allowing switching between alternative representations at relative low incremental cost compared to that of elliptic curve scalar multiplications (provided the isogeny has low degree  $l$ ). Note, however, that this does require storage of the polynomial coefficients of the isogeny and dual isogeny involved. This illustrates that low-degree isogenies are to be preferred, since an  $l$ -isogeny (usually) requires storing roughly  $6 \times l$  elements of  $GF(q)$ . While there are many isogenies, we therefore only consider those with the desired property with lowest possible degree.

## Appendix G. Further Cousins of Curve25519

### G.1. Further Alternative Representations

The Weierstrass curve Wei25519 is isomorphic to the Weierstrass curve Wei25519.2 defined over  $GF(p)$ , with as base point the pair  $(Glx', Gly')$ , and isogenous to the Weierstrass curve Wei25519.-3

defined over  $\text{GF}(p)$ , with as base point the pair  $(G2x', G2y')$ , where parameters are as specified in Appendix G.3 and where the related mappings are as specified in Appendix G.2.

## G.2. Further Switching

Each affine point  $(x, y)$  of Wei25519 corresponds to the point  $(x', y') := (x \cdot s^2, y \cdot s^3)$  of Wei25519.2, where  $s$  is the element of  $\text{GF}(p)$  defined by

```
s    20343593038935618591794247374137143598394058341193943326473831977
    39407761440

    (=0x047f6814 6d568b44 7e4552ea a5ed633d 02d62964 a2b0a120
    5e7941e9 375de020),
```

while the point at infinity of Wei25519 corresponds to the point at infinity of Wei25519.2. (Here, we used the mapping of Appendix F.1.) Under this mapping, the base point  $(Gx', Gy')$  of Wei25519 corresponds to the base point  $(G1x', G1y')$  of Wei25519.2. The inverse mapping maps the affine point  $(x', y')$  of Wei25519.2 to  $(x, y) := (x'/s^2, y'/s^3)$  of Wei25519, while mapping the point at infinity of Wei25519.2 to the point at infinity of Wei25519. Note that this mapping (and its inverse) involves a modular multiplication of both coordinates with fixed constants  $s^2$  and  $s^3$  (respectively,  $1/s^2$  and  $1/s^3$ ), which can be precomputed.

Each affine point  $(x, y)$  of Wei25519 corresponds to the point  $(x', y') := (x1/t^2, y1/t^3)$  of Wei25519.-3, where  $(x1, y1) = (u(x)/w(x)^2, y \cdot v(x)/w(x)^3)$ , where  $u$ ,  $v$ , and  $w$  are the polynomials with coefficients in  $\text{GF}(p)$  as defined in Appendix H.1 and where  $t$  is the element of  $\text{GF}(p)$  defined by

```
t    26012855558634277483276064234565597076862996895623795164528458073
    435568115620

    (=0x3982c126 59ad1749 ab8bc495 bb1a9d64 c9deffc5 e7b8e601
    a5651992 07d48fa4),
```

while the point at infinity of Wei25519 corresponds to the point at infinity of Wei25519.-3. (Here, we used the isogenous mapping of Appendix F.2.) Under this isogenous mapping, the base point  $(Gx', Gy')$  of Wei25519 corresponds to the base point  $(G2x', G2y')$  of Wei25519.-3. The dual isogeny maps the affine point  $(x', y')$  of Wei25519.-3 to  $(x, y) := (u'(x1)/w'(x1)^2, y1 \cdot v'(x1)/w'(x1)^3)$  of Wei25519, where  $(x1, y1) = (x' \cdot t^2, y' \cdot t^3)$  and where  $u'$ ,  $v'$ , and  $w'$  are the polynomials with coefficients in  $\text{GF}(p)$  as defined in Appendix H.2, while mapping the point at infinity of Wei25519.-3 to

the point at infinity of Wei25519. Under this dual isogenous mapping, the base point  $(G2x', G2y')$  of Wei25519.-3 corresponds to a multiple of the base point  $(Gx', Gy')$  of Wei25519, where this multiple is  $l=47$  (the degree of the isogeny; see the description in Appendix F.1). Note that this isogenous map (and its dual) primarily involves the evaluation of three fixed polynomials involving the x-coordinate, which takes roughly 140 modular multiplications (or less than 5-10% relative incremental cost compared to the cost of an elliptic curve scalar multiplication).

### G.3. Further Domain Parameters

The parameters of the Weierstrass curve with  $a=2$  that is isomorphic with Wei25519 and the parameters of the Weierstrass curve with  $a=-3$  that is isogenous with Wei25519 are as indicated below. Both domain parameter sets can be exploited directly to derive more efficient point addition formulae, should an implementation facilitate this.

General parameters: same as for Wei25519 (see Appendix E.3)

Weierstrass curve-specific parameters (for Wei25519.2, i.e., with  $a=2$ ):

```
a    2 (=0x02)

b    12102640281269758552371076649779977768474709596484288167752775713
    178787220689

    (=0x1ac1da05 b55bc146 33bd39e4 7f94302e f19843dc f669916f
    6a5dfd01 65538cd1)

G1x' 107705531383684005184170201967961611367923681983263378231495026
    81097436401658

    (=0x17cfeac3 78aed661 318e8634 582275b6 d9ad4def 072ea193
    5ee3c4e8 7a940ffa)

G1y' 544305758615084056530986689844575286168071033325025775211614397
    7388639873869

    (=0x0c08a952 c55dfad6 2c4f13f1 a8f68dca dc5c331d 297a37b6
    f0d7fdcc 51e16b4d)
```

Weierstrass curve-specific parameters (for Wei25519.-3, i.e., with  $a=-3$ ):

```
a    -3
```

```
(=0x7fffffff ffffffff ffffffff ffffffff ffffffff ffffffff
ffffffff fffffffea)
```

```
b 29689592517550930188872794512874050362622433571298029721775200646
451501277098
```

```
(=0x41a3b6bf c668778e be2954a4 b1df36d1 485ecef1 ea614295
796e1022 40891faa)
```

```
G2x' 538371792299408724349427232574807773704511272123391981336972078
46219400243292
```

```
(=0x7706c37b 5a84128a 3884a5d7 1811f1b5 5da3230f fb17a8ab
0b32e48d 31a6685c)
```

```
G2y' 695480730911001844144020555292799703925148674228551417730708041
8460388229929
```

```
(=0x0f60480c 7a5c0e11 40340adc 79d6a2bf 0cb57ad0 49d025dc
38d80c77 985f0329)
```

## Appendix H. Isogeny Details

The isogeny and dual isogeny are both isogenies with degree  $l=47$ . Both are specified by a triple of polynomials  $u$ ,  $v$ , and  $w$  (resp.  $u'$ ,  $v'$ , and  $w'$ ) of degree 47, 69, and 23, respectively, with coefficients in  $\text{GF}(p)$ . The coefficients of each of these polynomials are specified in Appendix H.1 (for the isogeny) and in Appendix H.2 (for the dual isogeny). For each polynomial in variable  $x$ , the coefficients are tabulated as sequence of coefficients of  $x^0$ ,  $x^1$ ,  $x^2$ , ..., in hexadecimal format.

### H.1. Isogeny Parameters

#### H.1.1. Coefficients of $u(x)$

```
0 0x670ed14828b6f1791ceb3a9cc0edfe127dee8729c5a72ddf77bb1abaebbbba1e8
1 0x1135ca8bd5383cb3545402c8bce2ced14b45c29b241e4751b035f27524a9f932
2 0x3223806ff5f669c430efd74df8389f058d180e2fcffa5cdef3eacecdd2c34771
3 0x31b8fecf3f17a819c228517f6cd9814466c8c8bea2efccc47a29bfc14c364266
4 0x2541305c958c5a326f44efad2bec284e7abee840fadb08f2d994cd382fd8ce42
5 0x6e6f9c5792f3ff497f860f44a9c469cec42bd711526b733e10915be5b2dbd8c6
```



6 0x3e9ad2e5f594b9ce6b06d4565891d28a1be8790000b396ef0bf59215d6cabfde  
7 0x278448895d236403bbc161347d19c913e7df5f372732a823ed807ee1d30206be  
8 0x42f9d171ea8dc2f4a14ea46cc0ee54967175ecfe83a975137b753cb127c35060  
9 0x128e40efa2d3ccb51567e73bae91e7c31eac45700fa13ce5781cbe5ddc985648  
10 0x450e5086c065430b496d88952dd2d5f2c5102bc27074d4d1e98bfa47413e0645  
11 0x487ef93da70dfd44a4db8cb41542e33d1aa32237bdca3a59b3ce1c59585f253d  
12 0x33d209270026b1d2db96efb36cc2fa0a49be1307f49689022eab1892b010b785  
13 0x4732b5996a20ebc4d5c5e2375d3b6c4b700c681bd9904343a14a0555ef0ecd48  
14 0x64dc9e8272b9f5c6ad3470db543238386f42b18cb1c592cc6caf7893141b2107  
15 0x52bbacd1f85c61ef7eafd8da27260fa2821f7a961867ed449b283036508ac5c5  
16 0x320447ed91210985e2c401cfel93db1379424cf748f92fd61ab5cc356bc89a2  
17 0x23d23a49bbcdf8cf4c4ce8a4ff7dd87d1ad1970317686254d5b4d2ec050d019f  
18 0x1601fca063f0bbbf15f198b3c20e474c2170294fa981f73365732d2372b40cd4  
19 0x7bf3f93840035e9688cfff402cee204a17c0de9779fc33503537dd78021bf4c4  
20 0x311998ce59fb7e1cd6af591ece3e84dfcb1c330cbcf28c0349e37b9581452853  
21 0x7ae5e41acfd28a9add2216dfed34756575a19b16984c1f3847b694326dad7f99  
22 0x704957e279244a5b107a6c57bd0ab9afe5227b7c0be2052cd3513772a40efee7  
23 0x56b918b5a0c583cb763550f8f71481e57c13bdcef2e5cfc8091d0821266f233b  
24 0x677073fed43ab291e496f798fbcf217bac3f014e35d0c2fa07f041ae746a04d7  
25 0x22225388e76f9688c7d4053b50ba41d0d8b71a2f21da8353d98472243ef50170  
26 0x66930b3dffdd3995a2502cef790d78b091c875192d8074bb5d5639f736400555  
27 0x79eb677c5e36971e8d64d56ebc0dedb4e9b7dd2d7b01343ebbd4d358d376e490  
28 0x48a204c2ca6d8636e9994842605bd648b91b637844e38d6c7dd707edce8256e2  
29 0xfb3529b0d4b9ce2d70760f33e8ce997a58999718e9277caf48623d27ae6a788

30 0x4352604bffd0c7d7a9ed898a2c6e7cf2512ffb89407271ba1f2c2d0ead8cc5aa  
31 0x6667697b29785fb6f0bd5e04d828991a5fe525370216f347ec767a26e7aac936  
32 0x9fc950b083c56dbd989badf9887255e203c879f123a7cb28901e50aea6d64dc  
33 0x41e51b51b5caadd1c15436bbf37596a1d7288a5f495d6b5b1ae66f8b2942b31d  
34 0x73b59fec709aa1cabd429e981c6284822a8b7b07620c831ab41fd31d5cf7430  
35 0x67e9b88e9a1bfbcb2554107d67d814986f1b09c3107a060cba21c019a2d5dc848  
36 0x6881494a1066ca176c5e174713786040affb4268b19d2abf28ef4293429f89c1  
37 0x5f4d30502ff1e1ccd624e6f506569454ab771869d7483e26afc09dea0c5ccd3d  
38 0x2a814cfc5859bca51e539c159955cbe729a58978b52329575d09bc6c3bf97ad  
39 0x1313c8aaae20d6f4397f0d8b19e52cfcdf8d8e10fba144aec1778fd10ddf4e9c  
40 0x7008d38f434b98953a996d4cc79fcbe9f9502411dcdf92005f725cea7ce82ad47  
41 0x5a74d1296aaaa245ffb848f434531fa3ba9e5cb9098a7091d36c2777d4cf5a13  
42 0x4bd3b700606397083f8038177bdaa1ac6eddba0447537582723cae0fd29341a9  
43 0x573453fb2b093016f3368356c786519d54ed05f5372c01723b4da520597ec217  
44 0x77f5c605bdb3a30d7d9c8840fce38650910d4418eed707a212c8927f41c2c812  
45 0x16d6b9f7ff57ca32350057de1204cc6d69d4ef1b255dfef8080118e2fef6ace3  
46 0x34e8595832a4021f8b5744014c6b4f7da7df0d0329e8b6b4d44c8fadad6513b7  
47 0x1

#### H.1.2. Coefficients of $v(x)$

0 0xf9f5eb7134e6f8dafa30c45afa58d7bfc6d4e3ccbb5de87b562fd77403972b2  
1 0x36c2dcd9e88f0d2d517a15fc453a098bbbb5a05eb6e8da906fae418a4e1a13f7  
2 0xb40078302c24fa394a834880d5bf46732ca1b4894172fb7f775821276f558b3  
3 0x53dd8e2234573f7f3f7df11e90a7bdd7b75d807f9712f521d4fb18af59aa5f26  
4 0x6d4d7bb08de9061988a8cf6ff3beb10e933d4d2fbb8872d256a38c74c8c2ceda

5 0x71bfe5831b30e28cd0fbele9916ab2291c6beacc5af08e2c9165c632e61dd2f5  
6 0x7c524f4d17ff2ee88463da012fc12a5b67d7fb5bd0ab59f4bbf162d76be1c89c  
7 0x758183d5e07878d3364e3fd4c863a5dc1fe723f48c4ab4273fc034f5454d59a4  
8 0x1eb41ef2479444ecdccbc200f64bde53f434a02b6c3f485d32f14da6aa7700e1  
9 0x1490f3851f016cc3cf8a1e3c16a53317253d232ed425297531b560d70770315c  
10 0x9bc43131964e46d905c3489c9d465c3abbd26eab9371c10e429b36d4b86469c  
11 0x5f27c173d94c7a413a288348d3fc88daa0bcf5af8f436a47262050f240e9be3b  
12 0x1d20010ec741aaa393cd19f0133b35f067adab0d105babe75fe45c8ba2732ceb  
13 0x1b3c669ae49b86be2f0c946a9ff6c48e44740d7d9804146915747c3c025996a  
14 0x24c6090f79ec13e3ae454d8f0f98e0c30a8938180595f79602f2ba013b3c10db  
15 0x4650c5b5648c6c43ac75a2042048c699e44437929268661726e7182a31b1532f  
16 0x957a835fb8bac3360b5008790e4c1f3389589ba74c8e8bf648b856ba7f22ba5  
17 0x1cd1300bc534880f95c7885d8df04a82bd54ed3e904b0749e0e3f8cb3240c7c7  
18 0x760b486e0d3c6ee0833b34b64b7ebc846055d4d1e0beeb6aedd5132399ada0ea  
19 0x1c666846c63965ef7edf519d6ada738f2b676ae38ff1f4621533373931b3220e  
20 0x365055118b38d4bc0df86648044affea2ef33e9a392ad336444e7d15e45585d1  
21 0x736487bde4b555abfccd3ea7ddcda98eda0d7c879664117dee906a88bc551194  
22 0x70de05ab9520222a37c7a84c61eedff71cb50c5f6647fc2a5d6e0ff2305cea37  
23 0x59053f6cdf6517ab3fe4bd9c9271d1892f8cf353d8041b98409e1e341a01f8b5  
24 0x375db54ed12fe8df9a198ea40200e812c2660b7022681d7932d89faf7c6e88d  
25 0x2a070c31d1c1a064daf56c79a044bd1cd6d13f1ddb0ff039b03a6469aaa9ed77  
26 0x41482351e7f69a756a5a2c0b3fa0681c03c550341d0ca0f76c5b394db9d2de8d  
27 0x747ac1109c9e9368d94a302cb5a1d23fcc7f0fd8a574efb7ddcaa738297c407a  
28 0x45682f1f2aab6358247e364834e2181ad0448bb815c587675fb2fee5a2119064

29 0x148c5bf44870dfd307317f0a0e4a8c163940bee1d2f01455a2e658aa92c13620  
30 0x6add1361e56ffa2d2fbbddba284b35be5845aec8069fc28af009d53290a705ce  
31 0x6631614c617400dc00f2c55357f67a94268e7b5369b02e55d5db46c935be3af5  
32 0x17cfffba496c64bb89d91c8c082f4c288c3c87feabd6b08591fe5a92216c094637  
33 0x648ff88155969f54c955a1834ad227b93062bb191170dd8c4d759f79ad5da250  
34 0x73e50900b89e5f295052b97f9d0c9edb0fc7d97b7fa5e3cfeefe33dd6a9cb223  
35 0x6afcb2f2ffe6c08508477aa4956cbd3dc864257f5059685adf2c68d4f2338f00  
36 0x372fd49701954c1b8f00926a8cb4b157d4165b75d53fa0476716554bf101b74c  
37 0x334ed41325f3724ff8becbf2b3443fea6d30fa543d1ca13188aceb2bdaf5f4e  
38 0x70e629c95a94e8e1b3974acb25e18ba42f8d5991786f0931f650c283adfe82fd  
39 0x738a625f4c62d3d645f1274e09ab344e72d441f3c0e82989d3e21e19212f23f3  
40 0x7093737294b29f21522f5664a9941c9b476f75d443b647bd2c777040bcd12a6a  
41 0xa996bad5863d821ccb8b89fa329ddbe5317a46bcb32552db396bea933765436  
42 0x2da237e3741b75dd0264836e7ef634fc0bc36ab187ebc790591a77c257b06f53  
43 0x1902f3daa86fa4f430b57212924fdc9e40f09e809f3991a0b3a10ab186c50ee5  
44 0x12baffec1bf20c921afd3cdf67a7f1d87c00d5326a3e5c83841593c214dadcb1  
45 0x6460f5a68123cb9e7bc1289cd5023c0c9ccd2d98eea24484fb3825b59dcd09aa  
46 0x2c7d63a868ffc9f0fd034f821d84736c5bc33325ce98aba5f0d95fef6f230ec8  
47 0x756e0063349a702db7406984c285a9b6bfba48177950d4361d8efa77408dc860  
48 0x37f3e30032b21e0279738e0a2b689625447831a2ccf15c638672da9aa7255ae  
49 0x1107c0dbe15d6ca9e790768317a40bcf23c80f1841f03ca79dd3e3ef4ea1ae30  
50 0x61ff7f25721d6206041c59a788316b09e05135a2aad94d539c65daa68b302cc2  
51 0x5dbfe346cbd0d61b9a3b5c42ec0518d3ae81cabcc32245060d7b0cd982b8d071  
52 0x4b6595e8501e9ec3e75f46107d2fd76511764efca179f69196eb45c0aa6fade3

53 0x72d17a5aa7bd8a2540aa9b02d9605f2a714f44abfb4c35d518b7abc39b477870  
54 0x658d8c134bac37729ec40d27d50b637201abbf1ab4157316358953548c49cf22  
55 0x36ac53b9118581ace574d5a08f9647e6a916f92dda684a4dbc405e2646b0243f  
56 0x1917a98f387d1e323e84a0f02d53307b1dd949e1a27b0de14514f89d9c0ef4b6  
57 0x21573434fde7ce56e8777c79539479441942dba535ade8ecb77763f7eb05d797  
58 0xe0bf482dc40884719bea5503422b603f3a8edb582f52838caa6eaab6eeac7ef  
59 0x3b0471eb53bd83e14fbc13928fe1691820349a963be8f7e9815848a53d03f5eb  
60 0x1e92cb067b24a729c42d3abb7a1179c577970f0ab3e6b0ce8d66c5b8f7001262  
61 0x74ea885c1ebed6f74964262402432ef184c42884fceb2f8dba3a9d67a1344dd7  
62 0x433ebce2ce9b0dc314425cfc2b234614d3c34f2c9da9fff4fddddd1ce242d035b  
63 0x33ac69e6be858dde7b83a9ff6f11de443128b39cec6e410e8d3b570e405ff896  
64 0xdab71e2ae94e6530a501ed8cf3df26731dd1d41cd81578341e12dca3cb71aa3  
65 0x537f58d52d18ce5b1d5a6bd3a420e796e64173491ad43dd4d1083a7dcc7dd201  
66 0x49c2f6afa93fdcc4e0f8128a8b06da4c75049be14edf3e103821ab604c60f8ae  
67 0x10a333eabd6135aeaa3f5f5f7e73d102e4fd7e4bf0902fc55b00da235fa1ad08  
68 0xf5c86044bf6032f5102e601f2a0f73c7bce9384bedd120f3e72d78484179d9c  
69 0x1

#### H.1.3. Coefficients of $w(x)$

0 0x3da24d42421264f30939ff00203880f2b017eb3fecf8933ae61e18df8c8ba116  
1 0x457f20bc393cdc9a66848ce174e2fa41d77e6dbae05a317a1fb6e3ae78760f8  
2 0x7f608a2285c480d5c9592c435431fae94695beef79d770bb6d029c1d10a53295  
3 0x3832accc520a485100a0a1695792465142a5572bed1b2e50e1f8f662ac7289bb  
4 0x2df1b0559e31b328eb34beedd5e537c3f4d7b9befb0749f75d6d0d866d26fbaa  
5 0x25396820381d04015a9f655ddd41c74303ded05d54a7750e2f58006659adda28

6 0x6fa070a70ca2bc6d4d0795fb28d4990b2cc80cd72d48b603a8ac8c8268bef6a6  
7 0x27f488578357388b20fbc7503328e1d10de602b082b3c7b8ceb33c29fea7a0d2  
8 0x15776851a7cabcf84c632118306915c0c15c75068a47021968c7438d46076e6  
9 0x101565b08a9af015c172fb194b940a4df25c4fb1d85f72d153efc79131d45e8f  
10 0x196b0ffbf92f3229fealdac0d74591b905ccaab6b83f905ee813ee8449f8a62c  
11 0x1f55784691719f765f04ee9051ec95d5deb42ae45405a9d87833855a6d95a94  
12 0x628858f79cca86305739d084d365d5a9e56e51a4485d253ae3f2e4a379fa8aff  
13 0x4a842dcd943a80d1e6e1dab3622a8c4d390da1592d1e56d1c14c4d3f72dd01a5  
14 0xf3bfc9cb17a1125f94766a4097d0f1018963bc11cb7bc0c7a1d94d65e282477  
15 0x1c4bd70488c4882846500691fa7543b7ef694446d9c3e3b4707ea2c99383e53c  
16 0x2d7017e47b24b89b0528932c4ade43f09091b91db0072e6ebdc5e777cb215e35  
17 0x781d69243b6c86f59416f91f7decaca93eab9cdc36a184191810c56ed85e0fdc  
18 0x5f20526f4177357da40a18da054731d442ad2a5a4727322ba8ed10d32eca24fb  
19 0x33e4cab64ed8a00d8012104fe8f928e6173c428eff95bbbe569ea46126a4f3cd  
20 0x50555b6f07e308d33776922b6566829d122e19b25b7bbacbb0a4b1a7dc40192  
21 0x533fa4bf1e2a2aae2f979065fdbb5b667ede2f85543fddbba146aa3a4ef2d281  
22 0x5a742cac1952010fc5aba200a635a7bed3ef868194f45b5a6a2647d6d6b289d2  
23 0x1

## H.2. Dual Isogeny Parameters

### H.2.1. Coefficients of $u'(x)$

0 0xf0eddb584a20aaac8f1419efdd02a5cca77b21e4cfae78c49b5127d98bc5882  
1 0x7115e60d44a58630417df33dd45b8a546fa00b79fea3b2bdc449694bade87c0a  
2 0xb3f3a6f3c445c7dc1f91121275414e88c32ff3f367ba0edad4d75b7e7b94b65  
3 0x1eb31bb333d7048b87f2b3d4ec76d69035927b41c30274368649c87c52e1ab30

4 0x552c886c2044153e280832264066cce2a7da1127dc9720e2a380e9d37049ac64  
5 0x4504f27908db2e1f5840b74ae42445298755d9493141f5417c02f04d47797dda  
6 0x82c242cce1eb19698a4fa30b5affe64e5051c04ae8b52cb68d89ee85222e628  
7 0x480473406add76cf1d77661b3ff506c038d9cdd5ad6e1ea41969430bb876d223  
8 0x25f47bb506fba80c79d1763365fa9076d4c4cb6644f73ed37918074397e88588  
9 0x10f13ed36eab593fa20817f6bb70cac292e18d300498f6642e35cbdf772f0855  
10 0x7d28329d695fb3305620f83a58df1531e89a43c7b3151d16f3b60a8246c36ade  
11 0x2c5ec8c42b16dc6409bdd2c7b4ffe9d65d7209e886badbd5f865dec35e4ab4a  
12 0x7f4f33cd50255537e6cde15a4a327a5790c37e081802654b56c956434354e133  
13 0x7d30431a121d9240c761998cf83d228237e80c3ef5c7191ec9617208e0ab8cec  
14 0x4d2a7d6609610c1deed56425a4615b92f70a507e1079b2681d96a2b874cf0630  
15 0x74676df60a9906901d1dc316c639ff6ae0fcdb02b5571d4b83fc2eedcd2936a8  
16 0x22f8212219aca01410f06eb234ed53bd5b8fbe7c08652b8002bcd1ea3cdae387  
17 0x7edb04449565d7c566b934a87fadade5515f23bda1ce25daa19ffff0c6a5ccc2f  
18 0x106ef71aa3aa34e8ecf4c07a67d03f0949d7d015ef2c1e32eb698dd3bec5a18c  
19 0x17913eb705db126ac3172447bcd811a62744d505ad0eea94cfcfdde5ca7428  
20 0x2cc793e6d3b592dcf5472057a991ff1a5ab43b4680bb34c0f5faffc5307827c1  
21 0x6dafcc0b16f98300cddb5e0a7d7ff04a0e73ca558c54461781d5a5ccb1ea0122  
22 0x7e418891cf222c021b0ae5f5232b9c0dc8270d4925a13174a0f0ac5e7a4c8045  
23 0x76553bd26fecb019ead31142684789fea7754c2dc9ab9197c623f45d60749058  
24 0x693efb3f81086043656d81840902b6f3a9a4b0e8f2a5a5edf5ce1c7f50a3898e  
25 0x46c630eac2b86d36f18a061882b756917718a359f44752a5caf41be506788921  
26 0x1dcfa01773628753bc6f448ac11be8a3bffa0011b9284967629b827e064f614  
27 0x8430b5b97d49b0938d1f66ecb9d2043025c6eec624f8f02042b9621b2b5cb19

28 0x66f66a6669272d47d3ec1efea36ee01d4a54ed50e9ec84475f668a5a9850f9be  
29 0x539128823b5ef3e87e901ab22f06d518a9bad15f5d375b49fe1e893ab38b1345  
30 0x2bd01c49d6fff22c213a8688924c10bf29269388a69a08d7f326695b3c213931  
31 0x3f7bea1baeccea3980201dc40d67c26db0e3b15b5a19b6cdac6de477aa717ac1  
32 0x6e0a72d94867807f7150fcb1233062f911b46e2ad11a3eac3c6c4c91e0f4a3fa  
33 0x5963f3cc262253f56fc103e50217e7e5b823ae8e1617f9e11f4c9c595fbb5bf6  
34 0x41440b6fe787777bc7b63afac9f4a38ddadcebc3d72f8fc73835247ba05f3a1d  
35 0x66d185401c1d2d0b84fcf6758a6a985bf9695651271c08f4b69ce89175fb7b34  
36 0x2673fb8c65bc4fe41905381093429a2601c46a309c03077ca229bac7d6ccf239  
37 0x1ce4d895ee601918a080de353633c82b75a3f61e8247763767d146554dd2f862  
38 0x18efa6c72fa908347547a89028a44f79f22542baa588601f2b3ed25a5e56d27c  
39 0x53de362e2f8ff220f8921620a71e8faa1aa57f8886fcbb6808fa3a5560570543  
40 0xdc29a73b97f08aa8774911474e651130ed364e8d8cffd4a80dee633aacecc47  
41 0x4e7eb8584ae4de525389d1e9300fc4480b3d9c8a5a45ecfbe33311029d8f6b99  
42 0x6c3cba4aa9229550fa82e1cfaee4b02f2c0cb86f79e0d412b8e32b00b7959d80  
43 0x5a9d104ae585b94af68eeb16b1349776b601f97b7ce716701645b1a75b68dcf3  
44 0x754e014b5e87af035b3d5fe6fb49f4631e32549f6341c6693c5172a6388e273e  
45 0x6710d8265118e22eaceba09566c86f642ab42da58c435083a353eaa12d866c39  
46 0x6e88ac659ce146c369f8b24c3a49f8dca547827250cf7963a455851cfc4f8d22  
47 0x971eb5f253356cd1fde9fb21f4a4902aa5b8d804a2b57ba775dc130181ae2e8

#### H.2.2. Coefficients of $v'(x)$

0 0x43c9b67cc5b16e167b55f190db61e44d48d813a7112910f10e3fd8da85d61d3  
1 0x72046db07e0e7882ff3f0f38b54b45ca84153be47a7fd1dd8f6402e17c47966f  
2 0x1593d97b65a070b6b3f879fe3dc4d1ef03c0e781c997111d5c1748f956f1ffc0



3 0x54e5fec076b8779338432bdc5a449e36823a0a7c905fd37f232330b026a143a0  
4 0x46328dd9bc336e0873abd453db472468393333fbf2010c6ac283933216e98038  
5 0x25d0c64de1dfelc6d5f5f2d98ab637d8b39bcf0d886a23dabac18c80d7eb03ce  
6 0x3a175c46b2cd8e2b313dde2d5f3097b78114a6295f283cf58a33844b0c8d8b34  
7 0x5cf4e6f745bdd61181a7d1b4db31dc4c30c84957f63cdf163bee5e466a7a8d38  
8 0x639071c39b723eea51cfd870478331d60396b31f39a593ebdd9b1eb543875283  
9 0x7ea8f895dcd85fc6cb2b58793789bd9246e62fa7a8c7116936876f4d8dff869b  
10 0x503818acb535bcaacf8ad44a83c213a9ce83af7c937dc9b3e5b6efedc0a7428c  
11 0xe815373920ec3cbf3f8cae20d4389d367dc4398e01691244af90edc3e6d42b8  
12 0x7e4b23e1e0b739087f77910cc635a92a3dc184a791400cbceae056c19c853815  
13 0x145322201db4b5ec0a643229e07c0ab7c36e4274745689be2c19cfa8a702129d  
14 0xfde79514935d9b40f52e33429621a200acc092f6e5dec14b49e73f2f59c780d  
15 0x37517ac5c04dc48145a9d6e14803b8ce9cb6a5d01c6f0ad1b04ff3353d02d815  
16 0x58ae96b8eefe9e80f24d3b886932fe3c27aaea810fa189c702f93987c8c97854  
17 0x6f6402c90fa379096d5f436035bebc9d29302126e9b117887abfa7d4b3c5709a  
18 0x1dbdf2b9ec09a8defeb485cc16ea98d0d45c5b9877ff16bd04c0110d2f64961  
19 0x53c51706af523ab5b32291de6c6b1ee7c5cbd0a5b317218f917b12ff38421452  
20 0x1b1051c7aec7d37a349208e3950b679d14e39f979db4fcd7b50d7d27dc918650  
21 0x1547e8d36262d5434cfb029cdd29385353124c3c35b1423c6cca1f87910b305b  
22 0x198efe984efc817835e28f704d41e4583a1e2398f7ce14045c4575d0445c6ce7  
23 0x492276dfe9588ee5cd9f553d990f377935d721822ecd0333ce2eb1d4324d539c  
24 0x77bad5319bacd5ed99e1905ce2ae89294efa7ee1f74314e4095c618a4e580c9b  
25 0x2cb3d532b8eac41c61b683f7b02feb9c2761f8b4286a54c3c4b60dd8081a312e  
26 0x37d189ea60443e2fee9b7ba8a34ed79ff3883dcefc06592836d2a9dd2ee3656e

27 0x79a80f9a0e6b8ded17a3d6ccf71eb565e3704c3543b77d70bca854345e880aba  
28 0x47718530ef8e8c75f069acb2d9925c5537908e220b28c8a2859b856f46d5f8db  
29 0x7dc518f82b55a36b4fa084b05bf21e3efce481d278a9f5c6a49701e56dac01ec  
30 0x340a318dad4b8d348a0838659672792a0f00b7105881e6080a340f708a9c7f94  
31 0x55f04d9d8891636d4e9c808a1fa95ad0dae7a8492257b20448023aad3203278e  
32 0x39dc465d58259f9f70bb430d27e2f0ab384a550e1259655443e14bdecba85530  
33 0x757385464cff265379a1adfadfd6f6a03fa8a2278761d4889ab097eff4d1ac28  
34 0x4d575654dbe39778857f4e688cc657416ce524d54864ebe8995ba766efa7ca2b  
35 0x47adb6aecc1949f2dc9f01206cc23eb4a0c29585d475dd24dc463c5087809298  
36 0x30d39e8b0c451a8fcf3d2abab4b86ffa374265abbe77c5903db4c1be8cec7672  
37 0x28cf47b39112297f0daaaa621f8e777875adc26f35dec0ba475c2ee148562b41  
38 0x36199723cc59867e2e309fe9941cd33722c807bb2d0a06eeb41de93f1b93f2f5  
39 0x5cdeb1f2ee1c7d694bdd884cb1c5c22de206684e1cafb8d3adb9a33cb85e19a2  
40 0xf6e6b3fc54c2d25871011b1499bb0ef015c6d0da802ae7eccf1d8c3fb73856c  
41 0xc1422c98b672414344a9c05492b926f473f05033b9f85b8788b4bb9a080053c  
42 0x19a8527de35d4faacb00184e0423962247319703a815eecf355f143c2c18f17f  
43 0x7812dc3313e6cf093da4617f06062e8e8969d648dfe6b5c331bccd58eb428383  
44 0x61e537180c84c79e1fd2d4f9d386e1c4f0442247605b8d8904d122ee7ef9f7be  
45 0x544d8621d05540576cfc9b58a3dab19145332b88eb0b86f4c15567c37205adf9  
46 0x11be3ef96e6e07556356b51e2479436d9966b7b083892b390caec22a117aa48e  
47 0x205cda31289cf75ab0759c14c43cb30f7287969ea3dc0d5286a3853a4d403187  
48 0x48d8fc6934f4f0a99f0f2cc59010389e2a0b20d6909bfcf8d7d0249f360acdc  
49 0x42cecc6d9bdca6d382e97fcea46a79c3eda2853091a8f399a2252115bf9a1454  
50 0x117d41b24f2f69cb3270b359c181607931f62c56d070bbd14dc9e3f9ab1432e

51 0x7c51564c66f68e2ad4ce6ea0d68f920fafa375376709c606c88a0ed44207aa1e  
52 0x48f25191fc8ac7d9f21adf6df23b76ccbca9cb02b815acdbebfa3f4eddc71b34  
53 0x4fc21a62c4688de70e28ad3d5956633fc9833bc7be09dc7bc500b7fae1e1c9a8  
54 0x1f23f25be0912173c3ef98e1c9990205a69d0bf2303d201d27a5499247f06789  
55 0x3131495618a0ac4cb11a702f3f8bab66c4fa1066d0a741af3c92d5c246edd579  
56 0xd93fe40faa53913638e497328a1b47603cb062c7afc9e96278603f29fd11fd4  
57 0x6b348bc59e984c91d696d1e3c3cfae44021f06f74798c787c355437fb696093d  
58 0x65af00e73043edcb479620c8b48098b89809d577a4071c8e33e8678829138b8a  
59 0x5e62ffb032b2ddb06591f86a46a18effd5d6ecf3f129bb2bacfd51a3739a98b6  
60 0x62c974ef3593fc86f7d78883b8727a2f7359a282cbc0196948e7a793e60ce1a1  
61 0x204d708e3f500aad64283f753e7d9bab976aa42a4ca1ce5e9d2264639e8b1110  
62 0xa90f0059da81a012e9d0a756809fab2ce61cb45965d4d1513a06227783ee4ea  
63 0x39fa55971c9e833f61139c39e243d40869fd7e8a1417ee4e7719dd2dd242766f  
64 0x22677c1e659caa324f0c74a013921facf62d0d78f273563145cc1ddccfcc4421  
65 0x3468cf6df7e93f7ff1fe1dd7e180a89dec3ed4f72843b4ea8a8d780011a245b2  
66 0x68f75a0e2210f52a90704ed5f511918d1f6bcfcd26b462cc4975252369db6e9d  
67 0x6220c0699696e9bcab0fe3a80d437519bd2bdf3caef665e106b2dd47585ddd9f  
68 0x553ad47b129fb347992b576479b0a89f8d71f1196f83e5eaab5f533a1dd6f6d7  
69 0x239aef387e116ec8730fa15af053485ca707650d9f8917a75f22acf6213197df

#### H.2.3. Coefficients of $w'(x)$

0 0x6bd7f1fc5dd51b7d832848c180f019bcbdb101d4b3435230a79cc4f95c35e15e  
1 0x17413bb3ee505184a504e14419b8d7c8517a0d268f65b0d7f5b0ba68d6166dd0  
2 0x47f4471beed06e5e2b6d5569c20e30346bdba2921d9676603c58e55431572f90  
3 0x2af7eaafd04f6910a5b01cdb0c27dca09487f1cd1116b38db34563e7b0b414eb

4 0x57f0a593459732eef11d2e2f7085bf9adf534879ba56f7afd17c4a40d3d3477b  
5 0x4da04e912f145c8d1e5957e0a9e44cca83e74345b38583b70840bdfdbd0288ed  
6 0x7cc9c3a51a3767d9d37c6652c349adc09bfe477d99f249a2a7bc803c1c5f39ed  
7 0x425d7e58b8adf87eebf445b424ba308ee7880228921651995a7eab548180ad49  
8 0x48156db5c99248234c09f43fedf509005943d3d5f5d7422621617467b06d314f  
9 0xd837dbbd1af32d04e2699cb026399c1928472aa1a7f0a1d3afd24bc9923456a  
10 0x5b8806e0f924e67c1f207464a9d025758c078b43ddc0ea9afe9993641e5650be  
11 0x29c91284e5d14939a6c9bc848908bd9df1f8346c259bbd40f3ed65182f3a2f39  
12 0x25550b0f3bceef18a6bf4a46c45bf1b92f22a76d456bdfd19d07398c80b0f946  
13 0x495d289b1db16229d7d4630cb65d52500256547401f121a9b09fb8e82cf01953  
14 0x718c8c610ea7048a370eabfd9888c633ee31dd70f8bcc58361962bb08619963e  
15 0x55d8a5ceef588ab52a07fa6047d6045550a5c52c91cc8b6b82eeb033c8ca557d  
16 0x620b5a4974cc3395f96b2a0fa9e6454202ef2c00d82b0e6c534b3b1d20f9a572  
17 0x4991b763929b00241a1a9a68e00e90c5df087f90b3352c0f4d8094a51429524e  
18 0x18b6b49c5650fb82e36e25fd4eb6decfd40b46c37425e6597c7444a1b6afb4e  
19 0x6868305b4f40654460aad63af3cb9151ab67c775eaac5e5df90d3aea58dee141  
20 0x16bc90219a36063a22889db810730a8b719c267d538cd28fa7c0d04f124c8580  
21 0x3628f9cf1fbe3eb559854e3b1c06a4cd6a26906b4e2d2e70616a493bba2dc574  
22 0x64abcc6759f1ce1ab57d41e17c2633f717064e35a7233a6682f8cf8e9538afec  
23 0x1

## Author's Address

Rene Struik  
Struik Security Consultancy  
  
Email: rstruik.ext@gmail.com

lwig  
Internet-Draft  
Intended status: Informational  
Expires: July 25, 2022

R. Struik  
Struik Security Consultancy  
Jan 21, 2022

Alternative Elliptic Curve Representations  
draft-ietf-lwig-curve-representations-23

Abstract

This document specifies how to represent Montgomery curves and (twisted) Edwards curves as curves in short-Weierstrass form and illustrates how this can be used to carry out elliptic curve computations leveraging existing implementations and specifications of, e.g., ECDSA and ECDH using NIST prime curves. We also provide extensive background material that may be useful for implementers of elliptic curve cryptography.

Requirements Language

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.

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 July 25, 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 Provisions Relating to IETF Documents (<https://trustee.ietf.org/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 Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

## Table of Contents

1. Fostering Code Reuse with New Elliptic Curves . . . . .	5
2. Specification of Wei25519 . . . . .	6
3. Use of Representation Switches . . . . .	7
4. Examples . . . . .	7
4.1. Implementation of X25519, Specification of ECDH25519 . .	8
4.2. Implementation of Ed25519 . . . . .	9
4.3. Specification of ECDSA25519 . . . . .	9
4.4. Other Uses (Wei448, ECDH448, ECDSA448, and Others) . . .	10
5. Caveats . . . . .	11
5.1. Wire Format . . . . .	11
5.2. Representation Conventions . . . . .	11
5.3. Domain Parameters . . . . .	12
6. Implementation Considerations . . . . .	13
7. Implementation Status . . . . .	14
8. Security Considerations . . . . .	15
9. Privacy Considerations . . . . .	16
10. Using Wei25519 and Wei448 with COSE and JOSE . . . . .	16
10.1. Using Wei25519 and Wei448 Keys with COSE and JOSE . . .	16
10.1.1. Encoding of Short-Weierstrass Curves with COSE . . .	17
10.1.2. Encoding of Short-Weierstrass Curves with JOSE . . .	18
10.2. Using ECDSA25519 and ECDSA448 with COSE and JOSE . . .	18
10.2.1. Encoding of ECDSA Instantiations with COSE . . . .	19
10.2.2. Encoding of ECDSA Instantiations with JOSE . . . .	20
10.3. Using ECDH25519 and ECDH448 with COSE and JOSE . . . .	21
10.3.1. Encoding of co-factor ECDH with COSE . . . . .	22
10.3.2. Encoding of co-factor ECDH with JOSE . . . . .	23
11. Using Wei25519 and Wei448 with PKIX and CMS . . . . .	23
11.1. Encoding of Short-Weierstrass Curves with PKIX . . . . .	23
11.2. Encoding of ECDSA Instantiations with PKIX . . . . .	23
11.3. Encoding of co-factor ECDH and Other Algorithms with PKIX . . . . .	24

11.4. Encoding of Elliptic-Curve-Based Algorithms with CMS . .	24
12. IANA Considerations . . . . .	24
12.1. OIDs for Use with PKIX and CMS . . . . .	25
12.2. COSE/JOSE IANA Considerations for Wei25519 . . . . .	25
12.2.1. COSE Elliptic Curves Registration . . . . .	25
12.2.2. COSE Algorithms Registration . . . . .	25
12.2.3. JOSE Elliptic Curves Registration . . . . .	26
12.2.4. JOSE Algorithms Registration (1/2) . . . . .	26
12.2.5. JOSE Algorithms Registration (2/2) . . . . .	27
12.3. COSE/JOSE IANA Considerations for Wei448 . . . . .	27
12.3.1. COSE Elliptic Curves Registration . . . . .	27
12.3.2. COSE Algorithms Registration (1/2) . . . . .	28
12.3.3. COSE Algorithms Registration (2/2) . . . . .	28
12.3.4. JOSE Elliptic Curves Registration . . . . .	28
12.3.5. JOSE Algorithms Registration (1/2) . . . . .	29
12.3.6. JOSE Algorithms Registration (2/2) . . . . .	29
13. Acknowledgements . . . . .	30
14. References . . . . .	30
14.1. Normative References . . . . .	30
14.2. Informative References . . . . .	33
Appendix A. Some (Non-Binary) Elliptic Curves . . . . .	35
A.1. Curves in Short-Weierstrass Form . . . . .	35
A.2. Montgomery Curves . . . . .	36
A.3. Twisted Edwards Curves . . . . .	36
Appendix B. Elliptic Curve Nomenclature and Finite Fields . . .	37
B.1. Elliptic Curve Nomenclature . . . . .	37
B.2. Finite Fields . . . . .	39
Appendix C. Elliptic Curve Group Operations . . . . .	40
C.1. Group Laws for Weierstrass Curves . . . . .	40
C.2. Group Laws for Montgomery Curves . . . . .	41
C.3. Group Laws for Twisted Edwards Curves . . . . .	42
Appendix D. Relationships Between Curve Models . . . . .	43
D.1. Mapping between Twisted Edwards Curves and Montgomery Curves . . . . .	43
D.2. Mapping between Montgomery Curves and Weierstrass Curves	44
D.3. Mapping between Twisted Edwards Curves and Weierstrass Curves . . . . .	45
Appendix E. Curve25519 and Cousins . . . . .	45
E.1. Curve Definition and Alternative Representations . . . .	45
E.2. Switching between Alternative Representations . . . . .	46
E.3. Domain Parameters . . . . .	47
Appendix F. Further Mappings . . . . .	49
F.1. Isomorphic Mapping between Twisted Edwards Curves . . . .	49
F.2. Isomorphic Mapping between Montgomery Curves . . . . .	50
F.3. Isomorphic Mapping between Weierstrass Curves . . . . .	51
F.4. Isogenous Mapping between Weierstrass Curves . . . . .	52
Appendix G. Further Cousins of Curve25519 . . . . .	53
G.1. Further Alternative Representations . . . . .	53

G.2.	Further Switching . . . . .	53
G.3.	Further Domain Parameters . . . . .	54
G.4.	Isogeny Details . . . . .	56
G.4.1.	Isogeny Parameters . . . . .	56
G.4.2.	Dual Isogeny Parameters . . . . .	62
Appendix H.	Point Compression . . . . .	68
H.1.	Point Compression for Weierstrass Curves . . . . .	69
H.2.	Point Compression for Montgomery Curves . . . . .	70
H.3.	Point Compression for Twisted Edwards Curves . . . . .	70
Appendix I.	Data Conversions . . . . .	71
I.1.	Strings and String Operations . . . . .	71
I.2.	Conversion between Bit Strings and Integers (BS2I, I2BS) . . . . .	72
I.3.	Conversion between Octet Strings and Integers (OS2I, I2OS) . . . . .	73
I.4.	Conversion between Octet Strings and Bit Strings (OS2BS, BS2OS) . . . . .	73
I.5.	Conversion between Field Elements and Octet Strings (FE2OS, OS2FE) . . . . .	74
I.6.	Conversion between Elements of $\mathbb{Z}_n$ and Octet Strings (ZnE2OS, OS2ZnE) . . . . .	74
I.7.	Ordering Conventions . . . . .	75
I.8.	Conversion Between Curve Points and Octet Strings . . . . .	76
Appendix J.	Representation Examples Curve25519 Family Members . . . . .	78
J.1.	Example with Curve25519 . . . . .	79
J.2.	Example with Edwards25519 . . . . .	81
J.3.	Example with Wei25519 . . . . .	83
J.4.	Example with Wei25519.2 . . . . .	86
J.5.	Example with Wei25519.-3 . . . . .	88
Appendix K.	Auxiliary Functions . . . . .	90
K.1.	Square Roots in $\text{GF}(q)$ . . . . .	90
K.1.1.	Square Roots in $\text{GF}(q)$ , where $q = 3 \pmod{4}$ . . . . .	91
K.1.2.	Square Roots in $\text{GF}(q)$ , where $q = 5 \pmod{8}$ . . . . .	91
K.2.	Inversion . . . . .	91
K.3.	Mappings to Curve Points . . . . .	92
K.3.1.	Mapping to Points of Weierstrass Curve . . . . .	92
K.3.2.	Mapping to Points of Montgomery Curve . . . . .	93
K.3.3.	Mapping to Points of Twisted Edwards Curve . . . . .	94
K.4.	Mappings to High-Order Curve Points . . . . .	95
K.4.1.	Mapping to High-Order Points of Weierstrass Curve . . . . .	95
K.4.2.	Mapping to High-Order Points of Montgomery Curve . . . . .	96
K.4.3.	Mapping to High-Order Points of Twisted Edwards Curve . . . . .	97
K.5.	Randomized Representation of Curve Points . . . . .	98
K.6.	Completing the Mappings to Curve Points . . . . .	99
Appendix L.	Curve secp256k1 and Friend . . . . .	102
L.1.	Curve Definition and Alternative Representation . . . . .	103
L.2.	Switching Between Representations . . . . .	103
L.3.	Domain Parameters . . . . .	103
L.4.	Isogeny Details . . . . .	105



L.4.1.	Isogeny Parameters . . . . .	105
L.4.2.	Dual Isogeny Parameters . . . . .	106
Appendix M.	Curve448 and Cousins . . . . .	106
M.1.	Curve Definition and Alternative Representations . . . . .	106
M.2.	Switching between Alternative Representations . . . . .	107
M.3.	Domain Parameters . . . . .	108
Appendix N.	Further Cousins of Curve448 . . . . .	111
N.1.	Further Alternative Representations . . . . .	111
N.2.	Further Switching . . . . .	111
N.3.	Further Domain Parameters . . . . .	114
N.4.	Isogeny Details . . . . .	116
N.4.1.	Isogeny Parameters . . . . .	116
N.4.2.	Dual Isogeny Parameters . . . . .	117
Appendix O.	Representation Examples Curve448 Family Members . . . . .	117
O.1.	Example with Curve448 . . . . .	118
O.2.	Example with Ed448 . . . . .	121
O.3.	Example with Wei448 . . . . .	124
O.4.	Example with Wei448.1 . . . . .	127
O.5.	Example with Wei448.-3 . . . . .	130
O.6.	Example with Edwards448 . . . . .	132
Appendix P.	Random Integers in $\mathbb{Z}_n$ . . . . .	135
P.1.	Conversion to Integers in $\mathbb{Z}_n$ via Modular Reduction . . . . .	136
P.2.	Conversion to Integers in $\mathbb{Z}_n$ via Scaling . . . . .	137
P.3.	Conversion to Integers in $\mathbb{Z}_n$ via the Discard Method . . . . .	138
Appendix Q.	ECDSA signatures . . . . .	138
Q.1.	ECDSA Signing Operation . . . . .	138
Q.2.	ECDSA Verification Operation . . . . .	139
Q.3.	Representation Examples ECDSA . . . . .	140
Q.3.1.	Example of ECDSA with Wei25519 and SHA-256 . . . . .	141
Q.3.2.	Example of ECDSA with Wei25519 and SHAKE128 . . . . .	143
Q.3.3.	Example of ECDSA with Wei448 and SHAKE256 . . . . .	145
Q.3.4.	Example of ECDSA with P-256 and SHA-256 . . . . .	147
Author's Address	. . . . .	150

## 1. Fostering Code Reuse with New Elliptic Curves

Elliptic curves can be represented using different curve models. Recently, IETF standardized elliptic curves that are claimed to have better performance and improved robustness against "real world" attacks than curves represented in the traditional short-Weierstrass curve model. These so-called CFRG curves [RFC7748] use the Montgomery curve model and the model of twisted Edwards curves.

In this document, we specify these curves using the traditional short-Weierstrass model and also define how to efficiently switch between representations in these different curve models. In particular, we specify Wei25519, which allows an alternative representation of points of Curve25519 (a Montgomery curve) and of

points of Edwards25519 (a twisted Edwards curve), as points of a corresponding short-Weierstrass curve. Similarly, we specify Wei448, which allows an alternative representation of points of Curve448 (a Montgomery curve) and of points of Ed448 (an Edwards curve), as points of a corresponding short-Weierstrass curve.

Use of Wei25519 and Wei448 allows easy definition of new instantiations of signature schemes and key agreement schemes already specified for traditional NIST prime curves, thereby allowing easy integration with existing specifications, such as NIST SP 800-56a [SP-800-56a], FIPS Pub 186-4 [FIPS-186-4], and ANSI X9.62-2005 [ANSI-X9.62], and fostering code reuse on platforms that already implement some of these schemes using elliptic curve arithmetic for curves in short-Weierstrass form (see Appendix C.1). To illustrate this, we specify how to use Wei25519 and Wei448 with co-factor ECDH and with ECDSA, thereby giving rise to the key agreement schemes ECDH25519 and ECDH448 and the signature schemes ECDSA25519 and ECDSA448. In all these cases, implementors may use the curve arithmetic for the curve model of their choosing (where they can efficiently switch between representations in different curve models, if required).

For ease of exposition, we consider Wei25519 first and introduce Wei448 simply as an illustration of how to create other "offspring" objects and protocols (see Section 4.4). We also provide extensive background material that we hope may be useful for implementors of elliptic curve cryptography or for cross-referencing with future specification work.

## 2. Specification of Wei25519

For the specification of Wei25519 and its relationship to Curve25519 and Edwards25519, see Appendix E. For further details and background information on elliptic curves, we refer to the other appendices.

The use of Wei25519 allows reuse of existing generic code that implements short-Weierstrass curves, such as the NIST curve P-256, to also implement the CFRG curves Curve25519 and Edwards25519. (Here, generic code refers to an implementation that does not depend on hardcoded domain parameters (see also Section 6).) We also cater to reusing of existing code where some domain parameters may have been hardcoded, thereby widening the scope of applicability. To this end, we specify the short-Weierstrass curves Wei25519.2 and Wei25519.-3, with hardcoded domain parameter  $a=2$  and  $a=-3 \pmod{p}$ , respectively; see Appendix G. (Here,  $p$  is the characteristic of the field over which these curves are defined.)

### 3. Use of Representation Switches

The curves Curve25519, Edwards25519, and Wei25519, as specified in Appendix E.3, are all isomorphic, with the transformations of Appendix E.2. These transformations map the specified base point of each of these curves to the specified base point of each of the other curves. Consequently, a public-private key pair  $(k, R := k * G)$  for any one of these curves corresponds, via these isomorphic mappings, to the public-private key pair  $(k, R' := k * G')$  for each of these other curves (where  $G$  and  $G'$  are the corresponding base points of these curves). This observation extends to the case where one also considers curve Wei25519.2 (which has hardcoded domain parameter  $a=2$ ), as specified in Appendix G.3, since it is isomorphic to Wei25519, with the transformation of Appendix G.2, and, thereby, also isomorphic to Curve25519 and Edwards25519.

The curve Wei25519.-3 (which has hardcoded domain parameter  $a=-3 \pmod{p}$ ) is not isomorphic to the curve Wei25519, but is related in a slightly weaker sense: the curve Wei25519 is isogenous to the curve Wei25519.-3, where the mapping of Appendix G.2 is an isogeny of degree  $l=47$  that maps the specified base point  $G$  of Wei25519 to the specified base point  $G'$  of Wei25519.-3 and where the so-called dual isogeny (which maps Wei25519.-3 to Wei25519) has the same degree  $l=47$ , but does not map  $G'$  to  $G$ , but to a fixed multiple hereof, where this multiple is  $l=47$ . Consequently, a public-private key pair  $(k, R := k * G)$  for Wei25519 corresponds to the public-private key pair  $(k, R' := k * G')$  for Wei25519.-3 (via the  $l$ -isogeny), whereas the public-private key pair  $(k, R' := k * G')$  corresponds to the public-private key pair  $(l * k, l * R = l * k * G)$  of Wei25519 (via the dual isogeny). (Note the extra scalar  $l=47$  here.)

Alternative curve representations can, therefore, be used in any cryptographic scheme that involves computations on public-private key pairs, where implementations may carry out computations on the corresponding object for the isomorphic or isogenous curve and convert the results back to the original curve (where, in case this involves an  $l$ -isogeny, one has to take into account the factor  $l$ ). This includes use with elliptic-curve based signature schemes and key agreement and key transport schemes.

For some examples of curve computations on each of the curves specified in Appendix E.3 and Appendix G.3, see Appendix J.

### 4. Examples

#### 4.1. Implementation of X25519, Specification of ECDH25519

RFC 7748 [RFC7748] specifies the use of X25519, a co-factor Diffie-Hellman key agreement scheme, with instantiation by the Montgomery curve Curve25519. This key agreement scheme was already specified in Section 6.1.2.2 of NIST SP 800-56a [SP-800-56a] for elliptic curves in short-Weierstrass form. Hence, one can implement X25519 using existing NIST routines by (1) representing a point of the Montgomery curve Curve25519 as a point of the Weierstrass curve Wei25519; (2) instantiating the co-factor Diffie-Hellman key agreement scheme of the NIST specification with the resulting point and Wei25519 domain parameters; (3) representing the key resulting from this scheme (which is a point of the curve Wei25519 in Weierstrass form) as a point of the Montgomery curve Curve25519. The representation change can be implemented via a simple wrapper and involves a single modular addition (see Appendix E.2). Using this method has the additional advantage that one can reuse the public-private key pair routines, domain parameter validation, and other checks that are already part of the NIST specifications.

A NIST-compliant version of the co-factor Diffie-Hellman key agreement scheme (denoted by ECDH25519) results if one keeps inputs (key contributions) and pre-output (shared key K) in the short-Weierstrass format (and, hence, does not perform Steps (1) and (3) above), where the actual output (shared secret Z) is the x-coordinate of K (if this is an affine point of the curve), represented as a fixed-size octet string in tight MSB/msb-order using the FE2OS mapping of Appendix I.5, and where the output is an error indicator otherwise (i.e., if K is the point at infinity O of the curve).

NOTE 1: A Montgomery version of the co-factor Diffie-Hellman key agreement scheme (denoted by X25519+) results by incorporating Steps (1), (2), and (3) above, i.e., where one keeps inputs (key contributions) and pre-output (shared key K) in the Montgomery curve format, as points of Curve25519, where one represents each affine point by only its x-coordinate, represented as a fixed-size octet string in tight LSB/msb-order using the FE2OS mapping and its reverse, the strict OS2FE mapping, of Appendix I.5, and where the actual output (shared secret Z) is the representation of the shared key K as defined above (if this is an affine point of the curve), and where the output is an error indicator otherwise (i.e., if K is the point at infinity O of the curve). The scheme X25519, as specified in [RFC7748], is a more lenient version of this X25519+ scheme, whereby one does not mandate rejection of shared keys in the small subgroup (which are instead represented as if these were the point (0,0) of order two), where one does not check whether a received key contribution is a point of Curve25519 rather than a point of a quadratic twist of this curve (for definitions of these terms, see

Appendix B.1), and where one uses the non-strict (rather than strict) OS2FE mapping (which, in this case, is always applied after setting the leftmost bit of the rightmost octet to zero). Moreover, with X25519, private keys are generated in the interval  $[2^{251}, 2^{252}-1]$  rather than in the interval  $[1, n-1]$  (the so-called "clamping") and one uses as base point  $G' := h \cdot G$ , where  $G$ ,  $n$ , and  $h$  are, respectively, the fixed base point, the order of the base point, and the co-factor of the curve in question.

NOTE 2: At this point, it is unclear whether a FIPS-accredited module implementing the co-factor Diffie-Hellman scheme with, e.g., P-256 would also extend this accreditation to the Montgomery versions X25519+ or X25519. (For cryptographic module validation program guidance, see, e.g., [FIPS-140-2].)

#### 4.2. Implementation of Ed25519

RFC 8032 [RFC8032] specifies Ed25519, a "full" Schnorr signature scheme, with instantiation by the twisted Edwards curve Edwards25519. One can implement the computation of the ephemeral key pair for Ed25519 using an existing Montgomery curve implementation by (1) generating a random public-private key pair  $(k, R' := k \cdot G')$  for Curve25519; (2) representing this public-private key as the pair  $(k, R := k \cdot G)$  for Ed25519. As before, the representation change can be implemented via a simple wrapper. Note that the Montgomery ladder specified in Section 5 of RFC7748 [RFC7748] does not provide sufficient information to reconstruct  $R' := (u, v)$  (since it does not compute the  $v$ -coordinate of  $R'$ ). However, this deficiency can be remedied by using a slightly modified version of the Montgomery ladder that includes reconstruction of the  $v$ -coordinate of  $R' := k \cdot G'$  at the end of the Montgomery ladder (which uses the  $v$ -coordinate of the base point  $G'$  of Curve25519 as well). For details, see Appendix C.2.

#### 4.3. Specification of ECDSA25519

FIPS Pub 186-4 [FIPS-186-4] specifies the signature scheme ECDSA and can be instantiated not just with the NIST prime curves, but also with other Weierstrass curves (that satisfy additional cryptographic criteria). In particular, one can instantiate this scheme with the Weierstrass curve Wei25519 and the hash function SHA-256 [FIPS-180-4], where an implementation may generate an ephemeral public-private key pair for Wei25519 by (1) internally carrying out these computations on the Montgomery curve Curve25519, the twisted Edwards curve Edwards25519, or even the Weierstrass curve Wei25519.-3 (with hardcoded  $a=-3$  domain parameter); (2) representing the result as a key pair for the curve Wei25519. Note that, in either case, one can implement these schemes with the same representation conventions

as used with existing NIST specifications, including bit/byte-ordering, compression functions, and the like. This allows generic implementations of ECDSA with the hash function SHA-256 and with the NIST curve P-256 or with the curve Wei25519 specified in this specification to reuse the same implementation (instantiated with, respectively, the NIST P-256 elliptic curve domain parameters or with the domain parameters of curve Wei25519 specified in Appendix E). We denote by ECDSA25519 the instantiation of ECDSA with SHA-256 and with curve Wei25519, where the signature  $(r,s)$  is represented as the right-concatenation of the integers  $r$  and  $s$  in the interval  $[1,n-1]$ , where  $n$  is the order of the base point of the curve in question, each represented as fixed-size octet strings in tight MSB/msb-order using the ZnE2OS mapping of Appendix I.6.

#### 4.4. Other Uses (Wei448, ECDH448, ECDSA448, and Others)

Any existing specification of cryptographic schemes using elliptic curves in Weierstrass form and that allows introduction of a new elliptic curve (here: Wei25519) is amenable to similar constructs, thus spawning "offspring" protocols, simply by instantiating these using the new curve in short-Weierstrass form, thereby allowing code and/or specifications reuse and, for implementations that so desire, carrying out curve computations "under the hood" on Montgomery curve and twisted Edwards curve cousins hereof (where these exist). This would simply require definition of a new object identifier for any such envisioned "offspring" protocol. This could significantly simplify standardization of schemes and help keeping at bay the resource and maintenance cost of implementations supporting algorithm agility [RFC7696].

We illustrate the construction of such offspring protocols for Curve448, another Montgomery curve recently standardized by IETF (see [RFC7748]). Similar to the case with Curve25519, one can represent points of this curve via different curve models, viz. as points of an Edwards curve (Ed448) or as points of a short-Weierstrass curve (Wei448). For the specification of Wei448 and its relationship to Curve448 and Ed448, see Appendix M. As with ECDH25519, one can now easily define a NIST-compliant version of co-factor Diffie-Hellman key agreement (denoted by ECDH448), by simply reusing the example of Section 4.1, but now using the short-Weierstrass curve Wei448, rather than Wei25519 (with the same representation and bit/byte-ordering conventions). Similarly, one can easily specify ECDSA with Wei448 and a suitable hash function, by simply reusing the example of Section 4.3, but now using the short-Weierstrass curve Wei448, rather than Wei25519, and picking as hash function SHAKE256 (see Section 6.3 of [FIPS-202]) with output size of  $d_0=512$  bits. We denote by ECDSA448 the resulting signature scheme (with the same representation and bit/byte-ordering conventions).

NOTE: A Montgomery version of the co-factor Diffie-Hellman key agreement scheme (denoted by X448+) results by reusing the description of X25519+ in Section 4.1, but now using the Montgomery curve Curve448, rather than Curve25519 (with the same checks and representation and bit/byte-ordering conventions). The scheme X448, as specified in [RFC7748], is a more lenient version of this X448+ scheme, whereby one does not mandate rejection of shared keys in the small subgroup (which are instead represented as if these were the point (0,0) of order two), nor checks whether a received key contribution is a point of Curve448 rather than a point of a quadratic twist of this curve, and where one uses the non-strict (rather than the strict) OS2FE mapping for converting octet strings to field elements. Moreover, with X448, private keys are generated in the interval  $[2^{445}, 2^{446}-1]$  rather than in the interval  $[1, n-1]$  (the so-called "clamping") and one uses as base point  $G' := h \cdot G$ , where  $G$ ,  $n$ , and  $h$  are, respectively, the fixed base point, the order of the base point, and the co-factor of the curve in question.

## 5. Caveats

The examples above illustrate how specifying the Weierstrass curve Wei25519 (or any curve in short-Weierstrass format, for that matter) may facilitate reuse of existing code and may simplify standards development. However, the following caveats apply:

### 5.1. Wire Format

The transformations between alternative curve representations can be implemented at negligible relative incremental cost if the curve points are represented as affine points. If a point is represented in compressed format, conversion usually requires a costly point decompression step. This is the case in [RFC7748], where the inputs to the co-factor Diffie-Hellman scheme X25519, as well as its output, are represented in u-coordinate-only format. This is also the case in [RFC8032], where the EdDSA signature includes the ephemeral signing key represented in compressed format (see Appendix H for details). Note that in the latter case compression is lossless, whereas it is lossy in the former case.

### 5.2. Representation Conventions

While elliptic curve computations are carried-out in a field  $GF(q)$  and, thereby, involve large integer arithmetic, these integers are represented as bit- and byte-strings. Here, [RFC8032] uses least-significant-byte (LSB)/least-significant-bit (lsb) conventions, whereas [RFC7748] uses LSB/most-significant-bit (msb) conventions, and where most other cryptographic specifications, including NIST SP800-56a [SP-800-56a], FIPS Pub 186-4 [FIPS-186-4], and ANSI

X9.62-2005 [ANSI-X9.62] use most-significant-byte (MSB)/msb conventions. Since each pair of conventions is different (see Appendix I for details and Appendix J for examples), this does necessitate bit/byte representation conversions.

### 5.3. Domain Parameters

All traditional NIST curves are Weierstrass curves with domain parameter  $a=-3$ , while all Brainpool curves [RFC5639] are isomorphic to a Weierstrass curve of this form. Thus, one can expect there to be existing Weierstrass implementations with a hardcoded  $a=-3$  domain parameter ("Jacobian-friendly"). For those implementations, including the curve Wei25519 as a potential vehicle for offering support for the CFRG curves Curve25519 and Edwards25519 is not possible, since it is not of the required form. Instead, one has to implement Wei25519.-3 and include code that implements the isogeny and dual isogeny from and to Wei25519. The lowest odd-degree isogeny has degree  $l=47$  and requires roughly 9kB of storage for isogeny and dual-isogeny computations (see the tables in Appendix G.4). Note that storage would have reduced to a single 64-byte table if only the Curve25519 curve would have been generated so as to be isomorphic to a Weierstrass curve with hardcoded  $a=-3$  parameter (this corresponds to  $l=1$ ).

NOTE 1: An example of a Montgomery curve defined over the same field as Curve25519 that is isomorphic to a Weierstrass curve with hardcoded  $a=-3$  parameter is the Montgomery curve  $M_{\{A,B\}}$  with  $B=1$  and  $A=-1410290$  (or, if one wants the base point to still have  $u$ -coordinate  $u=9$ , with  $B=1$  and  $A=-3960846$ ). In either case, the resulting curve has the same cryptographic properties as Curve25519 and the same performance (which relies on  $A$  being a 3-byte integer, as is the case with the domain parameter  $A=486662$  of Curve25519, and using the same special prime  $p=2^{255}-19$ ), while at the same time being "Jacobian-friendly" by design.

NOTE 2: While an implementation of Curve25519 via an isogenous Weierstrass curve with domain parameter  $a=-3$  requires a relatively large table (of size roughly 9kB), for a quadratic twist of Curve25519 (e.g., the Montgomery curve  $M_{\{A,B'\}}$  with  $A=486662$  and  $B'=2$ ) this implementation approach only requires a table of size less than 0.5kB (over 20x smaller), solely due to the fact that it is  $l$ -isogenous to a Weierstrass curve with  $a=-3$  parameter with relatively small parameter  $l=2$  (compared to  $l=47$ , as is the case with Curve25519 itself).



## 6. Implementation Considerations

The efficiency of elliptic curve arithmetic is primarily determined by the efficiency of its group operations (see Appendix C). Numerous optimized formulae exist, such as the use of so-called Montgomery ladders with Montgomery curves [Mont-Ladder] or with Weierstrass curves [Wei-Ladder], the use of hardcoded  $a=-3$  domain parameter for Weierstrass curves [ECC-Isogeny], and the use of hardcoded  $a=-1$  domain parameters for twisted Edwards curves [tEd-Formulas]. These all target reduction of the number of finite field operations (primarily, finite field multiplications and squarings). Other optimizations target more efficient modular reductions underlying these finite field operations, by specifying curves defined over a field  $\text{GF}(q)$ , where the field size  $q$  has a special form or a specific bit-length (typically, close to a multiple of a machine word). Depending on the implementation strategy, the bit-length of  $q$  may also facilitate reduced so-called "carry-effects" of integer arithmetic.

Most curves use a combination of these design philosophies. All NIST curves [FIPS-186-4] and Brainpool curves [RFC5639] are Weierstrass curves with  $a=-3$  domain parameter, thus facilitating more efficient elliptic curve group operations than with  $a \neq -3$  (via so-called Jacobian coordinates). The NIST curves and the Montgomery curve Curve25519 are defined over prime fields, where the prime number has a special form, whereas the Brainpool curves - by design - use a generic prime number. None of the NIST prime curves, nor the Brainpool curves, can be expressed as Montgomery or twisted Edwards curves, whereas - conversely - Montgomery curves and twisted curves can be expressed as Weierstrass curves.

While use of Wei25519 allows reuse of existing generic code that implements short-Weierstrass curves, such as the NIST curve P-256, to also implement the CFRG curves Curve25519 or Edwards25519, this obviously does not result in an implementation of these CFRG curves that exploits the specific structure of the underlying field or other specific domain parameters (since generic). Reuse of generic code, therefore, may result in a less computationally efficient curve implementation than would have been possible if the implementation had specifically targeted Curve25519 or Edwards25519 alone (with the overall cost differential estimated to be somewhere in the interval [1.00-1.25]). If existing generic code offers hardware support, however, the overall speed may still be larger, since less efficient formulae for curve arithmetic using Wei25519 curves compared to a direct implementation of Curve25519 or Edwards25519 arithmetic may be more than compensated for by faster implementations of the finite field arithmetic itself.

Overall, one should consider not just code reuse and computational efficiency, but also development and maintenance cost, and, e.g, the cost of providing effective implementation attack countermeasures (see also Section 8).

## 7. Implementation Status

[Note to the RFC Editor] Please remove this entire section before publication, as well as the reference to [RFC7942].

This section records the status of known implementations of the protocol defined by this specification at the time of posting of this Internet-Draft, and is based on a proposal described in [RFC7942]. The description of implementations in this section is intended to assist the IETF in its decision processes in progressing drafts to RFCs. Please note that the listing of any individual implementation here does not imply endorsement by the IETF. Furthermore, no effort has been spent to verify the information presented here that was supplied by IETF contributors. This is not intended as, and must not be construed to be, a catalog of available implementations or their features. Readers are advised to note that other implementations may exist.

According to [RFC7942], "this will allow reviewers and working groups to assign due consideration to documents that have the benefit of running code, which may serve as evidence of valuable experimentation and feedback that have made the implemented protocols more mature. It is up to the individual working groups to use this information as they see fit.

Nikolas Rosener evaluated the performance of switching between different curve models in his Master's thesis [Rosener]. For an implementation of Wei25519, see <<https://github.com/ncme/c25519>>. For support of this curve in tinydtls, see <<https://github.com/ncme/tinydtls>>.

ANSSI (the national cybersecurity agency of France) implemented the Ed25519 signature scheme using a generic ECC library for short-Weierstrass curves instantiated with the Wei25519 domain parameters, where this was motivated by the desire to both keep the library core mathematical foundations simple and keep the defense-in-depth (regarding software security and side-channels) focused on a rather limited part. For further details, see <<https://githubmemory.com/index.php/repo/ANSSI-FR/libecc>>.

According to <<https://community.nxp.com/docs/DOC-330199>>, an implementation of Wei25519 on the Kinets LTC ECC HW platform improves

the performance by over a factor ten compared to a stand-alone implementation of Curve25519 without hardware support.

The signature scheme ECDSA25519 (see Section 4.3) is supported in [RFC8928].

## 8. Security Considerations

The different representations of elliptic curve points discussed in this document are all obtained using a publicly known transformation, which is either an isomorphism or a low-degree isogeny. It is well-known that an isomorphism maps elliptic curve points to equivalent mathematical objects and that the complexity of cryptographic problems (such as the discrete logarithm problem) of curves related via a low-degree isogeny are tightly related. Thus, the use of these techniques does not negatively impact cryptographic security of elliptic curve operations.

As to implementation security, reusing existing high-quality code or generic implementations that have been carefully designed to withstand implementation attacks for one curve model may allow a more economical way of development and maintenance than providing this same functionality for each curve model separately (if multiple curve models need to be supported) and, otherwise, may allow a more gradual migration path, where one may initially use existing and accredited chipsets that cater to the pre-dominant curve model used in practice for over 15 years.

Elliptic curves are generally used as objects in a broader cryptographic scheme that may include processing steps that depend on the representation conventions used (such as with, e.g., key derivation following key establishment). These schemes should (obviously) unambiguously specify fixed representations of each input and output (e.g., representing each elliptic curve point always in short-Weierstrass form and in uncompressed tight MSB/msb format).

To prevent cross-protocol attacks, private keys SHOULD only be used with one cryptographic scheme.

Private keys MUST NOT be reused between Ed25519 (as specified in [RFC8032]) and ECDSA25519 (as specified in Section 4.3). Similarly, private keys MUST NOT be reused between Ed448 (as specified in [RFC8032]) and ECDSA448 (as specified in Section 4.4).

To prevent intra-protocol cross-instantiation attacks, ephemeral private keys MUST NOT be reused between instantiations of ECDSA25519 or of ECDSA448.

With ECDSA25519 and ECDSA448, the same private signature key MUST NOT be reused between application scenarios where message encoding and decoding rules vary, since this may jeopardize message unforgeability properties; see also the Note in Section 10.2.1. (In fact, this holds for any signature scheme, not just ECDSA.)

## 9. Privacy Considerations

The transformations between different curve models described in this document are publicly known and, therefore, do not affect privacy provisions.

Use of a public key in any protocol for which successful execution evidences knowledge of the corresponding private key implicitly indicates the entity holding this private key. Reuse of this public key with more than one protocol or more than one protocol instantiation may, therefore, allow traceability of this entity. It may also allow correlation of meta-data communicated with this common data element (e.g., different addressing information), even if an observer cannot technically verify the binding of this meta-data.

The randomized representation described in Appendix K.5 allows random curve points to be represented as random pairs of field elements, thereby assisting in obfuscating the presence of these curve points in some applications. For representations as random binary strings, see Appendix K.6.

## 10. Using Wei25519 and Wei448 with COSE and JOSE

This section defines algorithm encodings and representations enabling the use of the curves Wei25519 and Wei448 and their use with ECDH and ECDSA with JOSE [RFC7518] and COSE [RFC8152] messages.

All octet string encodings below use the MSB/msb-ordering conventions as defined in Appendix I.7. For CBOR representation details, we refer to [RFC8949]; for base64url encodings, we refer to [RFC4648].

### 10.1. Using Wei25519 and Wei448 Keys with COSE and JOSE

For Weierstrass curves, the representation of the point at infinity  $O$  is curve-specific (see Appendix H.1). For the short-Weierstrass curve Wei25519, we define  $O:=(-1,0)$ , whereas for Wei448, we define  $O:=(1,0)$ .

The encodings below specify the use of short-Weierstrass curves with COSE (see Section 10.1.1) and JOSE (see Section 10.1.2), where the encoding for a specific curve results by setting the "crv" parameter

to the unique name of the curve in question (i.e., "Wei25519" for the curve Wei25519 and "Wei448" for the curve Wei448).

#### 10.1.1. Encoding of Short-Weierstrass Curves with COSE

With COSE, points of short-Weierstrass curves are encoded using the "EC2" key type (Section 13.1.1 of [RFC8152]) or the "OKP" key type (Section 7.2 of [I-D.ietf-cose-rfc8152bis-algs]), which are instantiated by setting the "crv" parameter to the (unique) name of the curve in question and the "kty" parameter to "EC2" or "OKP", respectively, where key type-specific settings are as follows:

- a. With the "EC2" type, each affine point (X, Y) is encoded by setting the parameters "x" and "y" to the octet string representations of the elements X and Y, respectively, in tight MSB/msb-order, and converting each to a CBOR byte string. Each compressed point (X, t) is encoded by setting the parameter "x" to the octet representation of the element X, in tight MSB/msb-order, converted to a CBOR byte string, and by setting the parameter "y" to the CBOR false or CBOR true value, depending on whether, respectively,  $t=0$  or  $t=1$ . For representation details and for details on the reverse mappings, see Appendix I.8. (Note that for affine points of a curve defined over a prime field this representation is consistent with the "EC2" representation in Section 13.1.1 of [RFC8152].)
- b. With the "OKP" type, each point is encoded by setting the parameter "x" to the "squeezed" point representation of this point, in MSB/msb-order, and converting this to a CBOR byte string. For representation details and for details on the reverse mappings, see Appendix I.8. (Note that for affine points of a curve defined over a prime field this representation is consistent with the "OKP" representation in Section 7.2 of [I-D.ietf-cose-rfc8152bis-algs], which affords a curve-specific octet string encoding.)

In either case, if the point is a public key (i.e., the private key is well-defined), the parameter "d" encodes the corresponding private key, using the octet string representation, in tight MSB/msb-order, and converting this to a CBOR byte string (see Appendix I.6).

For curve points, the "crv" parameter and the parameters referenced with the applicable key type-specific settings above MUST be present in the structure, whereas the parameter "d" MUST NOT be present, while for private keys, the parameters "crv" and "d" MUST be present and the applicable key type-specific parameters of the corresponding public-key are RECOMMENDED to be present.

### 10.1.2. Encoding of Short-Weierstrass Curves with JOSE

With JOSE, points of short-Weierstrass curves are encoded using the "EC" key type (Section 6.2 of [RFC7518]) or the "OKP" key type (Section 2 of [RFC8037]), which are instantiated by setting the "crv" parameter to the (unique) name of the curve in question and the "kty" parameter to "EC" or "OKP", respectively, where key type-specific settings are as follows:

- a. With the "EC" type, each affine curve point (X, Y) is encoded by setting the parameters "x" and "y" to the octet string representations of the elements X and Y, respectively, in tight MSB/msb-order, and converting each using the base64url encoding. The point at infinity 0 is encoded as if this were an affine point. For representation details and details on the reverse mappings, see Appendix I.8. (Note that for affine points of a curve defined over a prime field this representation is consistent with the "EC" representation in Section 6.2 of [RFC7518].)
- b. With the "OKP" type, each curve point is encoded by setting the parameter "x" to the "squeezed" point representation of this point, in MSB/msb-order, and converting this using the base64url encoding. For representation details and for details on the reverse mappings, see Appendix I.8. (Note that for affine points of a curve defined over a prime field this representation is consistent with the "OKP" representation in Section 2 of [RFC8037], which affords a curve-specific octet string encoding.)

In either case, if the point is a public key (i.e., the private key is well-defined), the parameter "d" encodes the corresponding private key, using the octet string representation, in tight MSB/msb-order, and converting this using the base64url encoding (see Appendix I.6).

For curve points, the "crv" parameter and the parameters referenced with the applicable key type-specific settings above MUST be present in the structure, whereas the parameter "d" MUST NOT be present, while for private keys, the parameters "crv" and "d" MUST be present and the applicable key type-specific parameters of the corresponding public-key are RECOMMENDED to be present.

### 10.2. Using ECDSA25519 and ECDSA448 with COSE and JOSE

FIPS Pub 186-4 [FIPS-186-4] specifies the signature scheme ECDSA and can be instantiated with suitable combinations of elliptic curves in short-Weierstrass form and hash functions (that satisfy particular cryptographic criteria). While this completely specifies the

internal workings of the signing and signature verification operations, this does not uniquely specify the input/output formats:

- a. The signing operation takes as inputs a message  $m$  (represented as a bit string) and a private key  $d$  in the interval  $[1, n-1]$  and produces as output a signature, which is an ordered pair  $(r, s)$  of integers in the interval  $[1, n-1]$ , where  $n$  is the order of the base point of the curve in question;
- b. The signature verification operation takes as inputs a message  $m$ , a public key  $Q$ , and a signature  $(r, s)$  and produces as output the value "valid" or "invalid", depending upon whether the message was purportedly signed by a holder of the private key of the public-private key pair  $(d, Q)$  for the curve used with the signature scheme in question.

All inputs and outputs are uniquely determined by specifying the encodings of the message  $m$ , the private key  $d$ , the public key  $Q$ , the signature, and the values "valid" and "invalid".

The encodings below specify the use of instantiations of ECDSA with COSE (see Section 10.2.1) and JOSE (see Section 10.2.2), where the encoding for a specific ECDSA instantiation (i.e., with a specific short-Weierstrass curve and specific hash function) results by setting the "crv" parameter to the unique name of the underlying curve in question and the "alg" parameter to the unique name of the specific signature scheme instantiation. For JOSE, this is realized by setting the "alg" parameter to "ECDSA25519" for the ECDSA scheme defined in Section 4.3 and to "ECDSA448" for the scheme defined in Section 4.4. For COSE, this is realized by setting the "alg" parameter to "ES256" (short-hand for "ECDSA with SHA-256") for the ECDSA scheme defined in Section 4.3 and to "ECDSA with SHAKE256" for the scheme defined in Section 4.4. Note that, in the case of JOSE, the "alg" name uniquely defines the curve (and, thereby, implicitly the underlying "crv" parameter) and the underlying hash function, while in the case of COSE, the "alg" name uniquely defines the underlying hash function, but not the underlying curve.

#### 10.2.1. Encoding of ECDSA Instantiations with COSE

Instantiations of ECDSA used with COSE use the following encodings of inputs and outputs:

- a. The message  $m$  is the COSE Sig\_structure as specified in Section 4.4 of [RFC8152], converted to the CBOR byte string ToBeSigned in accordance with the Core Deterministic Encoding Requirements of Section 4.2.1 of [RFC8949]), converted to a bit string using the OS2BS mapping of Appendix I.4;

- b. The public key  $Q$  and the private key  $d$  are encoded as specified in Section 10.1.1, where the "crv" parameter is set to the unique name of the curve used with this particular instantiation of ECDSA;
- c. The Cose signature is encoded as the right-concatenation of the octet string representations of the coordinates of the signature pair  $(r, s)$ , in left-to-right order, where  $r$  and  $s$  are each represented as octet strings in tight MSB/msb-order using the ZnE2OS mapping of Appendix I.6, converted to a CBOR byte string. Note that, since we use a tight representation, this right-concatenated octet string has fixed size  $2 \cdot l$ , where the parameter  $l$  is uniquely defined by the set  $Z_n$  in question (where  $n$  is the (prime) order of the base point of the curve in question). The inverse mapping results by checking that the purported encoded signature (after CBOR decoding) has indeed size  $2 \cdot l$ , and by converting the left-side and right-side halves of this octet string (each of length  $l$ ) to, respectively, the integers  $r$  and  $s$  in  $Z_n$ , via the strict OS2ZnE mapping of Appendix I.6.

When using a COSE key for this algorithm, if the "alg" field is present, it MUST be set to the (unique) name of this particular instantiation of ECDSA and the "crv" parameter MUST be set to the (unique) name of the corresponding curve; if the "key\_ops" field is present, it MUST include "sign" when creating an ECDSA signature and it MUST include "verify" when verifying an ECDSA signature.

NOTE: Care should be taken that signers and verifiers do have a common understanding of message encoding rules, since otherwise signature verification may fail for messages with the same semantics. As an example, if there is ambiguity as to whether to represent the binary digit 0 as the integer 0 or as the CBOR false value (represented as the CBOR bit string b000\_00000 or b111\_10100, respectively), signing and signature verification may depend on different ToBeSigned strings and, thereby, may fail unexpectedly. This explains the (strong) requirement for deterministic encoding rules above and, thereby, the requirement for strong typing of any CBOR encodings used with signed messages. Further care should be taken that message decoding rules are always unambiguous, since otherwise the semantics of signed messages may not be clear or the unforgeability property of signatures may be jeopardized.

#### 10.2.2. Encoding of ECDSA Instantiations with JOSE

Instantiations of ECDSA used with JOSE use the following encodings of inputs and outputs:



- a. The message  $m$  is the JWS Signing Input as specified in [RFC7515], converted to a bit string, using the OS2BS mapping of Appendix I.4;
- b. The public key and the private key are encoded as specified in Section 10.1.2, where the "crv" parameter is set to the unique name of the curve used with this particular instantiation of ECDSA;
- c. The JWS signature is encoded as the right-concatenation of the octet string representations of the coordinates of the signature pair  $(r, s)$ , in left-to-right order, where  $r$  and  $s$  are each represented as octet strings in tight MSB/msb-order using the  $ZnE2OS$  mapping of Appendix I.6, converted using the base64url encoding. Note that, since we use a tight representation, this right-concatenated octet string has fixed size  $2 \cdot l$ , where the parameter  $l$  is uniquely defined by the set  $Z_n$  in question (where  $n$  is the (prime) order of the base point of the curve in question). The inverse mapping results by checking that the purported encoded signature (after base64url decoding) has indeed size  $2 \cdot l$ , and by converting the left-side and right-side halves of this octet string (each of length  $l$ ) to, respectively, the integers  $r$  and  $s$  in  $Z_n$ , via the strict  $OS2ZnE$  mapping of Appendix I.6.

When using a JOSE key for this algorithm, if the "alg" field is present, it MUST be set to the (unique) name of this particular instantiation of ECDSA and the "crv" parameter MUST be set to the (unique) name of the corresponding curve; if the "key\_ops" field is present, it MUST include "sign" when creating an ECDSA signature and it MUST include "verify" when verifying an ECDSA signature; if the JWK \_use\_ field is present, its value MUST be "sig".

### 10.3. Using ECDH25519 and ECDH448 with COSE and JOSE

Section 6.1.2.2 of NIST SP 800-56a [SP-800-56a] specifies the co-factor elliptic-curve Diffie-Hellman key agreement scheme (co-factor ECDH) and can be instantiated with a suitable elliptic curve in short-Weierstrass form (that satisfies particular cryptographic criteria). While this completely specifies the internal workings of the key agreement scheme in question, this does not uniquely specify the input/output formats:

- a. The co-factor Diffie-Hellman primitive (Section 5.7.1.2 of [SP-800-56a]) takes as inputs a private key  $d$  in the interval  $[1, n-1]$  from one of the parties and a point  $Q'$  obtained from the other party and produces the shared key  $K := h \cdot (d \cdot Q')$ , where  $h$  and  $n$  are, respectively, the co-factor and the order of the base

point of the curve in question and where  $Q'$  is a point of this curve. If this shared key  $K$  is the point at infinity  $O$  of the curve, the output is an error indicator;

- b. If the shared key  $K$  is an affine point of the curve, the output is the (raw) shared secret  $Z$ , which is the fixed-size octet representation of the x-coordinate of  $K$ , using the FE2OS mapping of Appendix I.5, represented in tight-MSB/msb-order (see Appendix I.7).

(NOTE: A subsequent key derivation function (kdf) takes as inputs the shared secret  $Z$  and side information OtherInfo and produces as output an octet string of DerivedKeyingMaterial, where details depend on the used kdf in question. This step is out of scope.)

The inputs and outputs are uniquely determined by specifying the encodings of private keys, curve points, and the error indicator for this key agreement scheme.

The encodings below specify the use of instantiations of ECDH with COSE (see Section 10.3.1) and JOSE (see Section 10.3.2), where the encoding for a specific co-factor ECDH instantiation (i.e., with a specific short-Weierstrass curve) results by setting the "crv" parameter to the unique name of the underlying curve in question and the "alg" parameter to the unique name of the specific key agreement scheme instantiation (i.e., "ECDH25519" for the co-factor ECDH scheme defined in Section 4.1 and "ECDH448" for the scheme defined in Section 4.4). Note that, in this case, the "alg" name uniquely defines the curve (and, thereby, implicitly the underlying "crv" parameter).

#### 10.3.1. Encoding of co-factor ECDH with COSE

Instantiations of co-factor ECDH used with COSE use the following encodings of inputs and outputs:

- a. Curve points and private keys are encoded as specified in Section 10.1.1, where the "crv" parameter is set to the unique name of the curve used with this particular instantiation of ECDH.

When using a COSE key for this algorithm, if the "alg" field is present, it MUST be set to the (unique) name of this particular instantiation of co-factor ECDH and the "crv" parameter MUST be set to the (unique) name of the corresponding curve; if the "key\_ops" field is present, it MUST include "derive shared secret" for the private key.

### 10.3.2. Encoding of co-factor ECDH with JOSE

Instantiations of co-factor ECDH used with JOSE use the following encodings of inputs and outputs:

- a. Curve points and private keys are encoded as specified in Section 10.1.2, where the "crv" parameter is set to the unique name of the curve used with this particular instantiation of ECDH.

When using a JOSE key for this algorithm, if the "alg" field is present, it MUST be set to the (unique) name of this particular instantiation of co-factor ECDH and the "crv" parameter MUST be set to the (unique) name of the corresponding curve; if the "key\_ops" field is present, it MUST include "derive shared secret" for the private key.

## 11. Using Wei25519 and Wei448 with PKIX and CMS

This section illustrates how to use the curves Wei25519 and Wei448 with ECDH and ECDSA with PKIX certificates (see [RFC5280] and [RFC5480]) and with CMS (see [RFC5652] and [RFC5753]).

### 11.1. Encoding of Short-Weierstrass Curves with PKIX

The namedCurve field in the ECPParameters field in the SubjectPublicKeyInfo structure [RFC5280] indicates the elliptic curve domain parameters for a specific curve, via a unique name of the curve in question (where these are the unique object identifiers id-Wei25519 for the curve Wei25519 and id-Wei448 for the curve Wei448).

Affine and compressed curve points are encoded using the "SEC1"-representation (see Note 2 of Appendix I.8), using the tight MSB/msb-ordering conventions. This is consistent with the representation in Section 2.2 of [RFC5480], after correcting for the error in [SEC1] (for the correction, see Note in Appendix H.1).

### 11.2. Encoding of ECDSA Instantiations with PKIX

ECDSA25519, as defined in Section 4.3, is the instantiation of ECDSA with SHA-256 and with curve Wei25519. With [RFC5480], ECDSA can be instantiated with suitable elliptic curves and hash functions. This allows support for ECDSA25519 by instantiating ECDSA with the curve Wei25519 and the hash function SHA-256, where curve Wei25519 is identified by its object identifier id-Wei25519 (see Section 11.1), where ECDSA with SHA-256 is identified by the object identifier id-ecdsa-with-SHA256 (see [RFC5480]), and where all other aspects are specified in [RFC5480].

ECDSA448, as defined in Section 4.4, is the instantiation of ECDSA with SHAKE256 with output size  $d=512$  bits and with curve Wei448. With [RFC5480], ECDSA can be instantiated with suitable elliptic curves and hash functions. This allows support for ECDSA448 by instantiating ECDSA with the curve Wei448 and the hash function SHAKE256 with output size of  $d=512$  bits, where curve Wei448 is identified by its object identifier `id-Wei448` (see Section 11.1), where ECDSA with SHAKE256 with output size of  $d=512$  bits is identified by the object identifier `id-ecdsa-with-shake256` (see [RFC8692]), and where all other aspects are specified in [RFC5480].

### 11.3. Encoding of co-factor ECDH and Other Algorithms with PKIX

With [RFC5480], the algorithm field in the `SubjectPublicKeyInfo` structure indicates the algorithm and the elliptic curve domain parameters for a specific curve, where that specification defines three algorithm identifiers (viz. `id-ecPublicKey`, `id-ecDH`, and `id-ecMQV`). Each of these algorithms can be instantiated with suitable elliptic curves, thereby allowing support for their use with the curves Wei25519 and Wei448, where these curves are identified by their unique object identifiers `id-Wei25519` and `id-Wei448`, respectively, (see Section 11.1) and where all other aspects are specified in [RFC5480].

### 11.4. Encoding of Elliptic-Curve-Based Algorithms with CMS

With [RFC5753], elliptic-curve based algorithms should use one of the elliptic curve domain parameters specified in [RFC5480], where the unique name of each such curve is identified by the object identifier of this curve defined in that document. Each of these algorithms can be instantiated with suitable elliptic curves, thereby allowing support for their use with the curves Wei25519 and Wei448, where these curves are identified by their unique object identifiers `id-Wei25519` and `id-Wei448`, respectively, (see Section 11.1) and where all other aspects are specified in [RFC5753].

## 12. IANA Considerations

Code points are requested for curves Wei25519 and Wei448 and their use with ECDSA and co-factor ECDH, using the representation conventions of this document.

New code points would be required in case one wishes to specify one or more other "offspring" protocols beyond those exemplified in Section 4.4. Specification hereof is, however, outside the scope of the current document.

### 12.1. OIDs for Use with PKIX and CMS

This section registers the following object identifiers for the curves introduced in this document:

- a. id-Wei25519 OBJECT IDENTIFIER ::= TBD (Requested value: {iso(1) identified-organization(3) thawte (101) 108 });
- b. id-Wei448 OBJECT IDENTIFIER ::= TBD (Requested value: {iso(1) identified-organization(3) thawte (101) 109 }).

For a description of how these are used with PKIX certificates and CMS, see Section 11.

### 12.2. COSE/JOSE IANA Considerations for Wei25519

#### 12.2.1. COSE Elliptic Curves Registration

This section registers the following value in the IANA "COSE Elliptic Curves" registry [IANA.COSE.Curves].

Name: Wei25519;

Value: TBD (Requested value: -1);

Key Type: EC2 or OKP;

Description: short-Weierstrass curve Wei25519;

Change Controller: IESG;

Reference: specified in Appendix E.3 of this specification; for encodings, see Section 10.1;

Recommended: Yes.

(Note that The "kty" value for Wei25519 may be "EC2" or "OKP".)

#### 12.2.2. COSE Algorithms Registration

This section registers the following value in the IANA "COSE Algorithms" registry [IANA.COSE.Algorithms].

Name: ECDH25519;

Value: TBD (Requested value: -24);

Description: NIST-compliant co-factor Diffie-Hellman w/ curve Wei25519 and key derivation function HKDF SHA256;

Change Controller: IESG;

Reference: specified in Section 4.1 of this specification; for encodings, see Section 10.3;

Recommended: Yes.

#### 12.2.3. JOSE Elliptic Curves Registration

This section registers the following value in the IANA "JSON Web Key Elliptic Curve" registry [IANA.JOSE.Curves].

Curve Name: Wei25519;

Curve Description: short-Weierstrass curve Wei25519;

JOSE Implementation Requirements: Optional;

Change Controller: IESG;

Reference: specified in Appendix E.3 of this specification; for encodings, see Section 10.1.

(Note that The "kty" value for Wei25519 may be "EC" or "OKP".)

#### 12.2.4. JOSE Algorithms Registration (1/2)

This section registers the following value in the IANA "JSON Web Signature and Encryption Algorithms" registry [IANA.JOSE.Algorithms].

Algorithm Name: ECDSA25519;

Algorithm Description: ECDSA using SHA-256 and curve Wei25519;

Algorithm Usage Locations: alg;

JOSE Implementation Requirements: Optional;

Change Controller: IESG;

Reference: specified in Section 4.3 of this specification; for encodings, see Section 10.2;

Algorithm Analysis Document(s): Section 4.3 of this specification.

#### 12.2.5. JOSE Algorithms Registration (2/2)

This section registers the following value in the IANA "JSON Web Signature and Encryption Algorithms" registry [IANA.JOSE.Algorithms].

Algorithm Name: ECDH25519;

Algorithm Description: NIST-compliant co-factor Diffie-Hellman w/  
curve Wei25519 and key derivation function HKDF SHA256;

Algorithm Usage Locations: alg;

JOSE Implementation Requirements: Optional;

Change Controller: IESG;

Reference: specified in Section 4.1 of this specification; for  
encodings, see Section 10.3;

Algorithm Analysis Document(s): Section 4.1 of this specification.

#### 12.3. COSE/JOSE IANA Considerations for Wei448

##### 12.3.1. COSE Elliptic Curves Registration

This section registers the following value in the IANA "COSE Elliptic Curves" registry [IANA.COSE.Curves].

Name: Wei448;

Value: TBD (Requested value: -2);

Key Type: EC2 or OKP;

Description: short-Weierstrass curve Wei448;

Change Controller: IESG;

Reference: specified in Appendix M.3 of this specification; for  
encodings, see Section 10.1;

Recommended: Yes.

(Note that The "kty" value for Wei448 may be "EC2" or "OKP".)

### 12.3.2. COSE Algorithms Registration (1/2)

This section registers the following value in the IANA "COSE Algorithms" registry [IANA.COSE.Algorithms].

Name: ECDSA with SHAKE256;

Value: TBD (Requested value: -48);

Description: ECDSA with SHAKE256;

Change Controller: IESG;

Reference: specified in Section 4.4 of this specification; for encodings, see Section 10.2;

Recommended: Yes.

### 12.3.3. COSE Algorithms Registration (2/2)

This section registers the following value in the IANA "COSE Algorithms" registry [IANA.COSE.Algorithms].

Name: ECDH448;

Value: TBD (Requested value: -49);

Description: NIST-compliant co-factor Diffie-Hellman w/ curve Wei448 and key derivation function HKDF SHA512;

Change Controller: IESG;

Reference: specified in Section 4.4 of this specification; for encodings, see Section 10.1; for key derivation, see Section 11.1 of [RFC8152];

Recommended: Yes.

### 12.3.4. JOSE Elliptic Curves Registration

This section registers the following value in the IANA "JSON Web Key Elliptic Curve" registry [IANA.JOSE.Curves].

Curve Name: Wei448;

Curve Description: short-Weierstrass curve Wei448;

JOSE Implementation Requirements: Optional;



Change Controller: IESG;

Reference: specified in Appendix M.3 of this specification; for encodings, see Section 10.1.

(Note that The "kty" value for Wei448 may be "EC" or "OKP".)

#### 12.3.5. JOSE Algorithms Registration (1/2)

This section registers the following value in the IANA "JSON Web Signature and Encryption Algorithms" registry [IANA.JOSE.Algorithms].

Algorithm Name: ECDSA448;

Algorithm Description: ECDSA using SHAKE256 and curve Wei448;

Algorithm Usage Locations: alg;

JOSE Implementation Requirements: Optional;

Change Controller: IESG;

Reference: specified in Section 4.4 of this specification; for encodings, see Section 10.2;

Algorithm Analysis Document(s): Section 4.4 of this specification.

#### 12.3.6. JOSE Algorithms Registration (2/2)

This section registers the following value in the IANA "JSON Web Signature and Encryption Algorithms" registry [IANA.JOSE.Algorithms].

Algorithm Name: ECDH448;

Algorithm Description: NIST-compliant co-factor Diffie-Hellman w/ curve Wei448;

Algorithm Usage Locations: alg;

JOSE Implementation Requirements: Optional;

Change Controller: IESG;

Reference: specified in Section 4.4 of this specification; for encodings, see Section 10.3;

Algorithm Analysis Document(s): Section 4.4 of this specification.

### 13. Acknowledgements

Thanks to Nikolas Rosener for discussions surrounding implementation details of the techniques described in this document and to Phillip Hallam-Baker for triggering inclusion of verbiage on the use of Montgomery ladders with recovery of the y-coordinate. Thanks to Stanislav Smyshlyaev, Vasily Nikolaev, and Benjamin Smith for their careful reviews.

### 14. References

#### 14.1. Normative References

[ANSI-X9.62]

ANSI X9.62-2005, "Public Key Cryptography for the Financial Services Industry: The Elliptic Curve Digital Signature Algorithm (ECDSA)", American National Standard for Financial Services, Accredited Standards Committee X9, Inc, Anapolis, MD, 2005.

[FIPS-180-4]

FIPS 180-4, "Secure Hash Standard (SHS), Federal Information Processing Standards Publication 180-4", US Department of Commerce/National Institute of Standards and Technology, Gaithersburg, MD, August 2015.

[FIPS-186-4]

FIPS 186-4, "Digital Signature Standard (DSS), Federal Information Processing Standards Publication 186-4", US Department of Commerce/National Institute of Standards and Technology, Gaithersburg, MD, July 2013.

[FIPS-202]

FIPS 202, "SHA-3 Standard: Permutation-Based Hash and Extendable-Output Functions, Federal Information Processing Standards Publication 202", US Department of Commerce/National Institute of Standards and Technology, Gaithersburg, MD, August 2015.

[I-D.ietf-cose-rfc8152bis-algs]

Schaad, J., "CBOR Object Signing and Encryption (COSE): Initial Algorithms", draft-ietf-cose-rfc8152bis-algs-12 (work in progress), September 2020.

[RFC0020]

Cerf, V., "ASCII format for network interchange", STD 80, RFC 20, DOI 10.17487/RFC0020, October 1969, <<https://www.rfc-editor.org/info/rfc20>>.

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <<https://www.rfc-editor.org/info/rfc2119>>.
- [RFC4648] Josefsson, S., "The Base16, Base32, and Base64 Data Encodings", RFC 4648, DOI 10.17487/RFC4648, October 2006, <<https://www.rfc-editor.org/info/rfc4648>>.
- [RFC5280] Cooper, D., Santesson, S., Farrell, S., Boeyen, S., Housley, R., and W. Polk, "Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile", RFC 5280, DOI 10.17487/RFC5280, May 2008, <<https://www.rfc-editor.org/info/rfc5280>>.
- [RFC5480] Turner, S., Brown, D., Yiu, K., Housley, R., and T. Polk, "Elliptic Curve Cryptography Subject Public Key Information", RFC 5480, DOI 10.17487/RFC5480, March 2009, <<https://www.rfc-editor.org/info/rfc5480>>.
- [RFC5639] Lochter, M. and J. Merkle, "Elliptic Curve Cryptography (ECC) Brainpool Standard Curves and Curve Generation", RFC 5639, DOI 10.17487/RFC5639, March 2010, <<https://www.rfc-editor.org/info/rfc5639>>.
- [RFC5652] Housley, R., "Cryptographic Message Syntax (CMS)", STD 70, RFC 5652, DOI 10.17487/RFC5652, September 2009, <<https://www.rfc-editor.org/info/rfc5652>>.
- [RFC5753] Turner, S. and D. Brown, "Use of Elliptic Curve Cryptography (ECC) Algorithms in Cryptographic Message Syntax (CMS)", RFC 5753, DOI 10.17487/RFC5753, January 2010, <<https://www.rfc-editor.org/info/rfc5753>>.
- [RFC7515] Jones, M., Bradley, J., and N. Sakimura, "JSON Web Signature (JWS)", RFC 7515, DOI 10.17487/RFC7515, May 2015, <<https://www.rfc-editor.org/info/rfc7515>>.
- [RFC7518] Jones, M., "JSON Web Algorithms (JWA)", RFC 7518, DOI 10.17487/RFC7518, May 2015, <<https://www.rfc-editor.org/info/rfc7518>>.
- [RFC7696] Housley, R., "Guidelines for Cryptographic Algorithm Agility and Selecting Mandatory-to-Implement Algorithms", BCP 201, RFC 7696, DOI 10.17487/RFC7696, November 2015, <<https://www.rfc-editor.org/info/rfc7696>>.

- [RFC7748] Langley, A., Hamburg, M., and S. Turner, "Elliptic Curves for Security", RFC 7748, DOI 10.17487/RFC7748, January 2016, <<https://www.rfc-editor.org/info/rfc7748>>.
- [RFC7942] Sheffer, Y. and A. Farrel, "Improving Awareness of Running Code: The Implementation Status Section", BCP 205, RFC 7942, DOI 10.17487/RFC7942, July 2016, <<https://www.rfc-editor.org/info/rfc7942>>.
- [RFC8032] Josefsson, S. and I. Liusvaara, "Edwards-Curve Digital Signature Algorithm (EdDSA)", RFC 8032, DOI 10.17487/RFC8032, January 2017, <<https://www.rfc-editor.org/info/rfc8032>>.
- [RFC8037] Liusvaara, I., "CFRG Elliptic Curve Diffie-Hellman (ECDH) and Signatures in JSON Object Signing and Encryption (JOSE)", RFC 8037, DOI 10.17487/RFC8037, January 2017, <<https://www.rfc-editor.org/info/rfc8037>>.
- [RFC8152] Schaad, J., "CBOR Object Signing and Encryption (COSE)", RFC 8152, DOI 10.17487/RFC8152, July 2017, <<https://www.rfc-editor.org/info/rfc8152>>.
- [RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <<https://www.rfc-editor.org/info/rfc8174>>.
- [RFC8692] Kampanakis, P. and Q. Dang, "Internet X.509 Public Key Infrastructure: Additional Algorithm Identifiers for RSASSA-PSS and ECDSA Using SHAKES", RFC 8692, DOI 10.17487/RFC8692, December 2019, <<https://www.rfc-editor.org/info/rfc8692>>.
- [RFC8949] Bormann, C. and P. Hoffman, "Concise Binary Object Representation (CBOR)", STD 94, RFC 8949, DOI 10.17487/RFC8949, December 2020, <<https://www.rfc-editor.org/info/rfc8949>>.
- [SEC1] SEC1, "SEC 1: Elliptic Curve Cryptography, Version 2.0", Standards for Efficient Cryptography, , June 2009.
- [SEC2] SEC2, "SEC 2: Elliptic Curve Cryptography, Version 2.0", Standards for Efficient Cryptography, , January 2010.

[SP-800-56a]

NIST SP 800-56a, "Recommendation for Pair-Wise Key Establishment Schemes Using Discrete Log Cryptography, Revision 3", US Department of Commerce/National Institute of Standards and Technology, Gaithersburg, MD, April 2018.

[SP-800-56c]

NIST SP 800-56c, "Recommendation for Key-Derivation Methods in Key-Establishment Schemes, Revision 1", US Department of Commerce/National Institute of Standards and Technology, Gaithersburg, MD, April 2018.

#### 14.2. Informative References

[comm-FIPS-186-5]

FIPS 186-5, "Public Comments Received on Draft FIPS Pub 186-5", US Department of Commerce/National Institute of Standards and Technology, Gaithersburg, MD, April 6, 2020.

[draft-FIPS-186-5]

FIPS 186-5, "Digital Signature Standard (DSS) (Draft)", US Department of Commerce/National Institute of Standards and Technology, Gaithersburg, MD, October 31, 2019.

[draft-NIST-800-186]

NIST SP 800-186, "Recommendations for Discrete Logarithm-Based Cryptography, Elliptic Curve Domain Parameters (Draft)", US Department of Commerce/National Institute of Standards and Technology, Gaithersburg, MD, October 31, 2019.

[ECC]

I.F. Blake, G. Seroussi, N.P. Smart, "Elliptic Curves in Cryptography", Cambridge University Press, Lecture Notes Series 265, July 1999.

[ECC-Isogeny]

E. Brier, M. Joye, "Fast Point Multiplication on Elliptic Curves through Isogenies", AAECC, Lecture Notes in Computer Science, Vol. 2643, New York: Springer-Verlag, 2003.

[FIPS-140-2]

FIPS 140-2, "Implementation Guidance for FIPS 140-2 and the Cryptographic Module Validation Program", US Department of Commerce/National Institute of Standards and Technology, Gaithersburg, MD, August 28, 2020.

- [GECC] D. Hankerson, A.J. Menezes, S.A. Vanstone, "Guide to Elliptic Curve Cryptography", New York: Springer-Verlag, 2004.
- [Handbook] A.J. Menezes, P. van Oorschot, S.A. Vanstone,, "Handbook of Applied Cryptography", Boca Raton: CRC Press, 1995.
- [IANA.COSE.Algorithms] IANA, "COSE Algorithms", IANA, <https://www.iana.org/assignments/cose/cose.xhtml#algorithms>.
- [IANA.COSE.Curves] IANA, "COSE Elliptic Curves", IANA, <https://www.iana.org/assignments/cose/cose.xhtml#elliptic-curves>.
- [IANA.JOSE.Algorithms] IANA, "JSON Web Signature and Encryption Algorithms", IANA, <https://www.iana.org/assignments/jose/jose.xhtml#web-signature-encryption-algorithms>.
- [IANA.JOSE.Curves] IANA, "JSON Web Key Elliptic Curve", IANA, <https://www.iana.org/assignments/jose/jose.xhtml#web-key-elliptic-curve>.
- [Mont-Ladder] P.L. Montgomery, "Speeding the Pollard and Elliptic Curve Methods of Factorization", Mathematics of Computation, Vol. 48, 1987.
- [RFC8928] Thubert, P., Ed., Sarikaya, B., Sethi, M., and R. Struik, "Address-Protected Neighbor Discovery for Low-Power and Lossy Networks", RFC 8928, DOI 10.17487/RFC8928, November 2020, <<https://www.rfc-editor.org/info/rfc8928>>.
- [Rosener] N. Rosener, "Evaluating the Performance of Transformations Between Curve Representations in Elliptic Curve Cryptography for Constrained Device Security", M.Sc. Universitat Bremen, August 2018.

- [SWUmap] E. Brier, J-S. Coron, Th. Icart, D. Madore, H. Randriam, M. Tibouchi, "Efficient Indifferentiable Hashing into Ordinary Elliptic Curves", CRYPTO 2010, Lecture Notes in Computer Science, Vol. 6223, New York: Springer-Verlag, 2010.
- [tEd] D.J. Bernstein, P. Birkner, M. Joye, T. Lange, C. Peters, "Twisted Edwards Curves", Africacrypt 2008, Lecture Notes in Computer Science, Vol. 5023, New York: Springer-Verlag, 2008.
- [tEd-Formulas] H. Hisil, K.K.H. Wong, G. Carter, E. Dawson, "Twisted Edwards Curves Revisited", ASIACRYPT 2008, Lecture Notes in Computer Science, Vol. 5350, New York: Springer-Verlag, 2008.
- [Tibouchi] M. Tibouchi, "Elligator Squared -- Uniform Points on Elliptic Curves of Prime Order as Uniform Random Strings", Financial Cryptography 2014, Lecture Notes in Computer Science, Vol. 8437, New York: Springer-Verlag, 2014.
- [Tibouchi-cleancut] T. Kim, M. Tibouchi, "Improved Elliptic Curve Hashing and Point Representation", DCC 2017, Des. Codes Cryptogr., Vol. 82, pp. 161-177, New York: Springer-Verlag, 2017.
- [Wei-Ladder] T. Izu, Ts. Takagi,, "A Fast Parallel Elliptic Curve Multiplication Resistant Against Side Channel Attacks", Centre for Applied Cryptographic Research, Corr 2002-03, 2002.

## Appendix A. Some (Non-Binary) Elliptic Curves

This section defines the three different curve models we consider, viz. short-Weierstrass curves, Montgomery curves, and twisted Edwards curves. For nomenclature, see Appendix B.

### A.1. Curves in Short-Weierstrass Form

Let  $GF(q)$  denote the finite field with  $q$  elements, where  $q$  is an odd prime power and where  $q$  is not divisible by three. Let  $W_{\{a,b\}}$  be the Weierstrass curve with defining equation  $Y^2 = X^3 + aX + b$ , where  $a$  and  $b$  are elements of  $GF(q)$  and where  $4a^3 + 27b^2$  is nonzero. The points of  $W_{\{a,b\}}$  are the ordered pairs  $(X, Y)$  whose coordinates are elements of  $GF(q)$  and that satisfy the defining

equation (the so-called affine points), together with the special point  $O$  (the so-called "point at infinity"). This set forms a group under addition, via the so-called "chord-and-tangent" rule, where the point at infinity serves as the identity element. See Appendix C.1 for details of the group operation.

A quadratic twist of  $W_{\{a,b\}}$  is a curve  $W_{\{a',b'\}}$  defined over the same field for which  $a' := a \cdot \gamma^2$  and  $b' := b \cdot \gamma^3$ , where  $\gamma$  is an element of  $GF(q)$  that is not a square in  $GF(q)$ .

## A.2. Montgomery Curves

Let  $GF(q)$  denote the finite field with  $q$  elements, where  $q$  is an odd prime power. Let  $M_{\{A,B\}}$  be the Montgomery curve with defining equation  $B \cdot v^2 = u^3 + A \cdot u^2 + u$ , where  $A$  and  $B$  are elements of  $GF(q)$  and where  $A$  is unequal to  $(\pm)2$  and where  $B$  is nonzero. The points of  $M_{\{A,B\}}$  are the ordered pairs  $(u, v)$  whose coordinates are elements of  $GF(q)$  and that satisfy the defining equation (the so-called affine points), together with the special point  $O$  (the so-called "point at infinity"). This set forms a group under addition, via the so-called "chord-and-tangent" rule, where the point at infinity serves as the identity element. See Appendix C.2 for details of the group operation.

A quadratic twist of  $M_{\{A,B\}}$  is a curve  $M_{\{A',B'\}}$  defined over the same field for which  $A' := A$  and  $B' := B \cdot \gamma$ , where  $\gamma$  is an element of  $GF(q)$  that is not a square in  $GF(q)$ .

## A.3. Twisted Edwards Curves

Let  $GF(q)$  denote the finite field with  $q$  elements, where  $q$  is an odd prime power. Let  $E_{\{a,d\}}$  be the twisted Edwards curve with defining equation  $a \cdot x^2 + y^2 = 1 + d \cdot x^2 \cdot y^2$ , where  $a$  and  $d$  are distinct nonzero elements of  $GF(q)$ . The points of  $E_{\{a,d\}}$  are the ordered pairs  $(x, y)$  whose coordinates are elements of  $GF(q)$  and that satisfy the defining equation (the so-called affine points). It can be shown that this set forms a group under addition if  $a$  is a square in  $GF(q)$ , whereas  $d$  is not, where the point  $O := (0, 1)$  serves as the identity element. (Note that the identity element satisfies the defining equation.) See Appendix C.3 for details of the group operation. (All curves  $E_{\{a,d\}}$  in this document are assumed to satisfy the condition on domain parameters  $a$  and  $d$  above and, thereby, satisfy the Note in that appendix.)

An Edwards curve is a twisted Edwards curve with  $a=1$ .



A quadratic twist of  $E_{\{a,d\}}$  is a curve  $E_{\{a',d'\}}$  defined over the same field for which  $a' := a \cdot \gamma$  and  $d' := d \cdot \gamma$ , where  $\gamma$  is an element of  $\text{GF}(q)$  that is not a square in  $\text{GF}(q)$ .

## Appendix B. Elliptic Curve Nomenclature and Finite Fields

This section provides brief background information on elliptic curves and finite fields that should be sufficient to understand constructions and examples in this document.

### B.1. Elliptic Curve Nomenclature

The set of points of each curve defined in Appendix A forms a commutative group under addition (denoted by '+'). In Appendix C we specify the group laws, which depend on the curve model in question. For completeness, we here include some common elliptic curve nomenclature and basic properties (primarily so as to keep this document self-contained). These notions are mainly used in Appendix E and Appendix G and not essential for our exposition. This section can be skipped at first reading.

Any point  $P$  of a curve  $E$  is a generator of the cyclic subgroup  $\langle P \rangle := \{k \cdot P \mid k = 0, 1, 2, \dots\}$  of the curve. (Here,  $k \cdot P$  denotes the sum of  $k$  copies of  $P$ , where  $0 \cdot P$  is the identity element  $O$  of the curve;  $k \cdot P$  is commonly referred to as scalar multiplication of  $P$  by  $k$ .) If  $\langle P \rangle$  has cardinality  $l$ , then  $l$  is called the order of  $P$  and  $l$  is the smallest positive integer so that  $l \cdot P = O$ . The order of curve  $E$  is the cardinality of the set of its points, commonly denoted by  $|E|$ . A curve is cyclic if it is generated by some point of this curve. All curves of prime order are cyclic, while all curves of order  $h \cdot n$ , where  $n$  is a large prime number and where  $h$  is a small number (the so-called co-factor), have a large cyclic subgroup of prime order  $n$ . In this case, a generator of order  $n$  is called a base point, commonly denoted by  $G$ , while a point of order dividing  $h$  is said to be in the small subgroup (or said to be a low-order point). For curves of prime order, this small subgroup is the singleton set, consisting of only the identity element  $O$ . A point that is not in the small subgroup is said to be a high-order point (since it has order at least  $n$ ). A point  $P$  of the curve is in the small subgroup if  $h \cdot P = O$  (and is a high-order point otherwise); this point  $P$  has order  $n$  if  $n \cdot P = O$  and if it is not the identity element  $O$ . (The latter order check is commonly called full public key validation.) The above definitions extend to curves with a relatively large co-factor, by defining  $n$  to be the size of its largest prime-order subgroup.

If  $R$  is a point of the curve that is also contained in  $\langle P \rangle$ , there is a unique integer  $k$  in the interval  $[0, l-1]$  so that  $R = k \cdot P$ , where  $l$  is the order of  $P$ . This number is called the discrete logarithm of  $R$  to

the base  $P$ . The discrete logarithm problem is the problem of finding the discrete logarithm of  $R$  to the base  $P$  for any two points  $P$  and  $R$  of the curve, if such a number exists.

Random points  $R$  of  $\langle P \rangle$ , where  $P$  has order  $l$ , can be computed by generating a random integer  $k$  in the interval  $[0, l-1]$  and by subsequently computing  $R := k * P$ , where  $R$  then has order  $l/\gcd(k, l)$ . In particular, if  $P$  is a high-order point (of curve  $E$  of order  $h * n$ ), then so is  $R$ , unless  $k$  is a multiple of  $n$  (in which case  $R$  is a low-order point). For methods for generating  $k$ , see Appendix P.

If  $P$  is a fixed base point  $G$  of the curve, the pair  $(k, R := k * G)$  is commonly called a public-private key pair, the integer  $k$  the private key, and the point  $R$  the corresponding public key. The private key  $k$  can be represented as an integer in the interval  $[0, n-1]$ , where  $G$  has order  $n$ . If this representation is nonzero,  $R$  has order  $n$ ; otherwise, it has order one and is the identity element  $O$  of the curve.

A curve  $E$  defined over the field  $GF(q)$  has order  $|E|$  relatively close to  $q$ . More precisely,  $|E| = q + 1 - t$  for some integer  $t$  (the so-called trace) with absolute value at most  $2 * \sqrt{q}$ . This is commonly referred to as the Hasse bound.

In this document, a quadratic twist of a curve  $E$  defined over a field  $GF(q)$  is a specific curve  $E'$  related to  $E$  defined over the same field, with cardinality  $|E'|$ , where  $|E| + |E'| = 2 * (q + 1)$ . If  $E$  is a curve in one of the curve models specified in this document, a quadratic twist  $E'$  of this curve can be expressed using the same curve model, although (naturally) with its own curve parameters (see Appendix A). Points that are points of both  $E$  and  $E'$  have order one or two. Two curves  $E_1$  and  $E_2$  defined over the field  $GF(q)$  are said to be isogenous if these have the same order and are said to be isomorphic if the defining equation of  $E_1$  can be transformed into the defining equation of  $E_2$  via a so-called admissible change of variables. Note that isomorphic curves have necessarily the same order and are, thus, a special case of isogenous curves. Isomorphic curves have the same group structure, whereas this is not necessarily the case for isogenous curves. Further details are out of scope.

Curves in short-Weierstrass form can have prime order, whereas Montgomery curves and twisted Edwards curves always have an order that is a multiple of four (and, thereby, a small subgroup of cardinality four).

An ordered pair  $(x, y)$  whose coordinates are elements of  $GF(q)$  can be associated with any ordered triple of the form  $[x * z : y * z : z]$ , where  $z$  is a nonzero element of  $GF(q)$ , and can be uniquely recovered from

such a representation. The latter representation is commonly called a representation in projective coordinates. Sometimes, yet other representations are useful (e.g., representation in Jacobian coordinates). Further details are out of scope.

The group laws in Appendix C are mostly expressed in terms of affine points, but can also be expressed in terms of the representation of these points in projective coordinates, thereby allowing clearing of denominators. The group laws may also involve non-affine points (such as the point at infinity  $O$  of a Weierstrass curve or of a Montgomery curve). Those can also be represented in projective coordinates. Further details are out of scope.

## B.2. Finite Fields

The field  $GF(q)$ , where  $q$  is a prime power, is defined as follows.

If  $q=p$  is a prime number, the field  $GF(p)$  consists of the integers in the interval  $[0, p-1]$  and two binary operations on this set: addition and multiplication modulo  $p$ . This field is commonly called a prime field. The additive and multiplicative identity elements are  $0$  and  $1$ , respectively.

If  $q=p^m$ , where  $p$  is a prime number and where  $m>0$ , the field  $GF(q)$  is defined in terms of an irreducible polynomial  $f(z)$  in  $z$  of degree  $m$  with coefficients in  $GF(p)$  (i.e.,  $f(z)$  cannot be written as the product of two polynomials in  $z$  of lower degree with coefficients in  $GF(p)$ ): in this case,  $GF(q)$  consists of the polynomials in  $z$  of degree smaller than  $m$  with coefficients in  $GF(p)$  and two binary operations on this set: polynomial addition and polynomial multiplication modulo the irreducible polynomial  $f(z)$ . By definition, each element  $x$  of  $GF(q)$  is a polynomial in  $z$  of degree smaller than  $m$  and can, therefore, be uniquely represented as a vector  $(x_{m-1}, x_{m-2}, \dots, x_1, x_0)$  of length  $m$  with coefficients in  $GF(p)$ , where  $x_i$  is the coefficient of  $z^i$  of polynomial  $x$ . Note that this representation depends on the irreducible polynomial  $f(z)$  of the field  $GF(p^m)$  in question (which is often fixed in practice). Note that  $GF(q)$  contains the prime field  $GF(p)$  as a subset. If  $m=1$ , the definitions of  $GF(p)$  and  $GF(p^1)$  above coincide, since each nonzero element of  $GF(p)$  can be viewed as a polynomial in  $z$  of degree zero. If  $m>1$  (i.e., if  $q$  is a strict prime power), then  $GF(q)$  is called a (nontrivial) extension field of  $GF(p)$ . The number  $p$  is called the characteristic of  $GF(q)$ .

Any nonzero element  $g$  of  $GF(q)$  is a generator of the cyclic multiplicative subgroup  $\langle g \rangle := \{g^k \mid k = 0, 1, 2, \dots\}$  of  $GF(q) \setminus \{0\}$ . (Here,  $g^k$  denotes the product of  $k$  copies of  $g$ , where  $g^0$  is the multiplicative identity element  $1$  of  $GF(q) \setminus \{0\}$ .) If  $\langle g \rangle$  has

cardinality  $l$ , then  $l$  is called the order of  $g$  and  $l$  is the smallest positive integer so that  $g^l=1$ . For each finite field  $GF(q)$ , the set  $GF(q)\setminus\{0\}$  forms a cyclic group, i.e., it is generated by some nonzero element hereof. Each such generator is called a primitive element of  $GF(q)$  and has order  $q-1$ . Each nonzero element of  $GF(q)$  has order dividing  $q-1$  (a property commonly referred to as Fermat's Little Theorem).

A field element  $y$  is called a square in  $GF(q)$  if it can be expressed as  $y=x^2$  for some  $x$  in  $GF(q)$ ; it is called a non-square in  $GF(q)$  otherwise. If  $y$  is a square in  $GF(q)$ , we denote by  $\text{sqrt}(y)$  one of its square roots (the other one being  $-\text{sqrt}(y)$ ). For methods for computing square roots in  $GF(q)$  – if these exist – and for computing inverses in  $GF(q)\setminus\{0\}$ , see Appendix K.1 and Appendix K.2, respectively. For methods for mapping a nonzero field element that is not a square in  $GF(q)$  to a point of a curve, see Appendix K.3 (or see Appendix K.4, if one wishes to always obtain a high-order point of the curve in question).

NOTE: The curves in Appendix E and Appendix G are all defined over a prime field  $GF(p)$ , thereby reducing all operations to simple modular integer arithmetic. Strictly speaking we could, therefore, have refrained from introducing extension fields. Nevertheless, we included the more general exposition, so as to accommodate potential introduction of new curves that are defined over a (nontrivial) extension field at some point in the future. This includes curves proposed for post-quantum isogeny-based schemes, which are defined over a quadratic extension field (i.e., where  $q=p^2$ ), and elliptic curves used with pairing-based cryptography. The exposition in either case is almost the same and now automatically yields, e.g., data conversion routines for any finite field object (see Appendix I). Readers not interested in this could simply view all fields as prime fields.

## Appendix C. Elliptic Curve Group Operations

This section specifies group operations for elliptic curves in short-Weierstrass form, for Montgomery curves, and for twisted Edwards curves.

### C.1. Group Laws for Weierstrass Curves

For each point  $P$  of the Weierstrass curve  $W_{\{a,b\}}$ , the point at infinity  $O$  serves as identity element, i.e.,  $P + O = O + P = P$ .

For each affine point  $P=(X, Y)$  of the Weierstrass curve  $W_{\{a,b\}}$ , the point  $-P$  is the point  $(X, -Y)$  and one has  $P + (-P) = O$  (i.e.,  $-P$  is the inverse of  $P$ ). For the point at infinity  $O$ , one has  $-O=O$ .

Let  $P_1 := (X_1, Y_1)$  and  $P_2 := (X_2, Y_2)$  be distinct affine points of the Weierstrass curve  $W_{\{a,b\}}$  and let  $Q := P_1 + P_2$ , where  $Q$  is not the identity element. Then  $Q = (X, Y)$ , where

$$X + X_1 + X_2 = \lambda^2 \text{ and } Y + Y_1 = \lambda(X_1 - X), \text{ where}$$

$$\lambda := (Y_2 - Y_1) / (X_2 - X_1).$$

Let  $P := (X_1, Y_1)$  be an affine point of the Weierstrass curve  $W_{\{a,b\}}$  and let  $Q := 2P$ , where  $Q$  is not the identity element. Then  $Q = (X, Y)$ , where

$$X + 2X_1 = \lambda^2 \text{ and } Y + Y_1 = \lambda(X_1 - X), \text{ where}$$

$$\lambda := (3X_1^2 + a) / (2Y_1).$$

From the group laws above it follows that if  $P = (X, Y)$ ,  $P_1 = (X_1, Y_1)$ , and  $P_2 = (X_2, Y_2)$  are distinct affine points of the Weierstrass curve  $W_{\{a,b\}}$  with  $P_2 := P + P_1$  and if  $Y$  is nonzero, then the  $Y$ -coordinate of  $P_1$  can be expressed in terms of the  $X$ -coordinates of  $P$ ,  $P_1$ , and  $P_2$ , and the  $Y$ -coordinate of  $P$ , since

$$2Y \cdot Y_1 = (X \cdot X_1 + a) \cdot (X + X_1) + 2b - X_2 \cdot (X - X_1)^2.$$

This property allows recovery of the  $Y$ -coordinate of a point  $P_1 = kP$  that is computed via the so-called Montgomery ladder, where  $P$  is an affine point with nonzero  $Y$ -coordinate (i.e., it does not have order two). For future reference, note that the expression above uniquely determines the  $X$ -coordinate of  $P_2$  in terms of the  $X$ -coordinates of  $P$  and  $P_1$  and the product of their  $Y$ -coordinates. Further details are out of scope.

## C.2. Group Laws for Montgomery Curves

For each point  $P$  of the Montgomery curve  $M_{\{A,B\}}$ , the point at infinity  $O$  serves as identity element, i.e.,  $P + O = O + P = P$ .

For each affine point  $P := (u, v)$  of the Montgomery curve  $M_{\{A,B\}}$ , the point  $-P$  is the point  $(u, -v)$  and one has  $P + (-P) = O$  (i.e.,  $-P$  is the inverse of  $P$ ). For the point at infinity  $O$ , one has  $-O = O$ .

Let  $P_1 := (u_1, v_1)$  and  $P_2 := (u_2, v_2)$  be distinct affine points of the Montgomery curve  $M_{\{A,B\}}$  and let  $Q := P_1 + P_2$ , where  $Q$  is not the identity element. Then  $Q = (u, v)$ , where

$$u + u_1 + u_2 = B\lambda^2 - A \text{ and } v + v_1 = \lambda(u_1 - u), \text{ where}$$

$$\lambda := (v_2 - v_1) / (u_2 - u_1).$$

Let  $P := (u_1, v_1)$  be an affine point of the Montgomery curve  $M_{\{A,B\}}$  and let  $Q := 2P$ , where  $Q$  is not the identity element. Then  $Q = (u, v)$ , where

$$u + 2u_1 = B\lambda^2 - A \text{ and } v + v_1 = \lambda(u_1 - u), \text{ where}$$

$$\lambda := (3u_1^2 + 2Au_1 + 1) / (2Bv_1).$$

From the group laws above it follows that if  $P = (u, v)$ ,  $P_1 = (u_1, v_1)$ , and  $P_2 = (u_2, v_2)$  are distinct affine points of the Montgomery curve  $M_{\{A,B\}}$  with  $P_2 = P + P_1$  and if  $v$  is nonzero, then the  $v$ -coordinate of  $P_1$  can be expressed in terms of the  $u$ -coordinates of  $P$ ,  $P_1$ , and  $P_2$ , and the  $v$ -coordinate of  $P$ , since

$$2Bv^2v_1 = (u^2u_1 + 1)(u + u_1 + 2A) - 2A - u^2(u - u_1)^2.$$

This property allows recovery of the  $v$ -coordinate of a point  $P_1 = kP$  that is computed via the so-called Montgomery ladder, where  $P$  is an affine point with nonzero  $v$ -coordinate (i.e., it does not have order two). For future reference, note that the expression above uniquely determines the  $u$ -coordinate of  $P_2$  in terms of the  $u$ -coordinates of  $P$  and  $P_1$  and the product of their  $v$ -coordinates. Further details are out of scope.

### C.3. Group Laws for Twisted Edwards Curves

Note: The group laws below hold for twisted Edwards curves  $E_{\{a,d\}}$  where  $a$  is a square in  $GF(q)$ , whereas  $d$  is not. In this case, the addition formulae below are defined for each pair of points, without exceptions. Generalizations of this group law to other twisted Edwards curves are out of scope.

For each point  $P$  of the twisted Edwards curve  $E_{\{a,d\}}$ , the point  $O := (0, 1)$  serves as identity element, i.e.,  $P + O = O + P = P$ .

For each point  $P := (x, y)$  of the twisted Edwards curve  $E_{\{a,d\}}$ , the point  $-P$  is the point  $(-x, y)$  and one has  $P + (-P) = O$  (i.e.,  $-P$  is the inverse of  $P$ ).

Let  $P_1 := (x_1, y_1)$  and  $P_2 := (x_2, y_2)$  be points of the twisted Edwards curve  $E_{\{a,d\}}$  and let  $Q := P_1 + P_2$ . Then  $Q = (x, y)$ , where

$$x = (x_1y_2 + x_2y_1) / (1 + d x_1x_2y_1y_2) \text{ and}$$

$$y = (y_1y_2 - a x_1x_2) / (1 - d x_1x_2y_1y_2).$$

Let  $P := (x_1, y_1)$  be a point of the twisted Edwards curve  $E_{\{a,d\}}$  and let  $Q := 2P$ . Then  $Q = (x, y)$ , where

$$x = (2*x_1*y_1)/(1 + d*x_1^2*y_1^2) \text{ and}$$

$$y = (y_1^2 - a*x_1^2)/(1 - d*x_1^2*y_1^2).$$

Note that one can use the formulae for point addition for point doubling, taking inverses, and adding the identity element as well (i.e., the point addition formulae are uniform and complete (subject to our Note above)).

From the group laws above (subject to our Note above) it follows that if  $P=(x, y)$ ,  $P_1=(x_1, y_1)$ , and  $P_2=P+P_1=(x_2, y_2)$  are points of the twisted Edwards curve  $E_{\{a,d\}}$  with  $P_2:=P+P_1$  and if  $x$  is nonzero, then the  $x$ -coordinate of  $P_1$  can be expressed in terms of the  $y$ -coordinates of  $P$ ,  $P_1$ , and  $P_2$ , and the  $x$ -coordinate of  $P$ , since

$$x*x_1*(a-d*y*y_1*y_2)=y*y_1-y_2.$$

(Here, observe that  $a-d*y*y_1*y_2$  is nonzero per our Note above.) This property allows recovery of the  $x$ -coordinate of a point  $P_1=k*P$  that is computed via the so-called Montgomery ladder, where  $P$  is an affine point with nonzero  $x$ -coordinate (i.e., it does not have order one or two). For future reference, note that the group law (subject to our Note above) uniquely determines the  $y$ -coordinate of  $P_2$  in terms of the  $y$ -coordinates of  $P$  and  $P_1$  and the product of their  $x$ -coordinates. Further details are out of scope.

#### Appendix D. Relationships Between Curve Models

The non-binary curves specified in Appendix A are expressed in different curve models, viz. as curves in short-Weierstrass form, as Montgomery curves, or as twisted Edwards curves. These curve models are related, as follows.

##### D.1. Mapping between Twisted Edwards Curves and Montgomery Curves

One can map points of the Montgomery curve  $M_{\{A,B\}}$  to points of the twisted Edwards curve  $E_{\{a,d\}}$ , where  $a:=(A+2)/B$  and  $d:=(A-2)/B$  and, conversely, map points of the twisted Edwards curve  $E_{\{a,d\}}$  to points of the Montgomery curve  $M_{\{A,B\}}$ , where  $A:=2*(a+d)/(a-d)$  and where  $B:=4/(a-d)$ . For twisted Edwards curves we consider (i.e., those where  $a$  is a square in  $GF(q)$ , whereas  $d$  is not), this defines a one-to-one correspondence, which – in fact – is an isomorphism between  $M_{\{A,B\}}$  and  $E_{\{a,d\}}$ , thereby showing that, e.g., the discrete logarithm problem in either curve model is equally hard.

For the Montgomery curves and twisted Edwards curves we consider, the mapping from  $M_{\{A,B\}}$  to  $E_{\{a,d\}}$  is defined by mapping the point at infinity  $O$  and the point  $(0, 0)$  of order two of  $M_{\{A,B\}}$  to,

respectively, the point  $(0, 1)$  and the point  $(0, -1)$  of order two of  $E_{\{a,d\}}$ , while mapping each other point  $(u, v)$  of  $M_{\{A,B\}}$  to the point  $(x,y):=(u/v, (u-1)/(u+1))$  of  $E_{\{a,d\}}$ . (Note that this is well-defined, since neither  $(A-2)/B$  nor  $A^2-4$  are squares in  $GF(q)$ , so  $M_{\{A,B\}}$  has a single point of order two and no affine points  $(u,v)$  with  $u=-1$ .) The inverse mapping from  $E_{\{a,d\}}$  to  $M_{\{A,B\}}$  is defined by mapping the point  $(0, 1)$  and the point  $(0, -1)$  of order two of  $E_{\{a,d\}}$  to, respectively, the point at infinity  $O$  and the point  $(0, 0)$  of order two of  $M_{\{A,B\}}$ , while each other point  $(x, y)$  of  $E_{\{a,d\}}$  is mapped to the point  $(u,v):=((1+y)/(1-y), (1+y)/((1-y)*x))$  of  $M_{\{A,B\}}$ . (Note that this is well-defined, since for points  $(x,y)$  of  $E_{\{a,d\}}$ ,  $x=0$  only if  $y=(+/-)1$ .)

Implementations may take advantage of this mapping to carry out elliptic curve group operations originally defined for a twisted Edwards curve on the corresponding Montgomery curve, or vice-versa, and translating the result back to the original curve, thereby potentially allowing code reuse.

#### D.2. Mapping between Montgomery Curves and Weierstrass Curves

One can map points of the Montgomery curve  $M_{\{A,B\}}$  to points of the Weierstrass curve  $W_{\{a,b\}}$ , where  $a:=(3-A^2)/(3*B^2)$  and  $b:=(2*A^3-9*A)/(27*B^3)$ . This defines a one-to-one correspondence, which - in fact - is an isomorphism between  $M_{\{A,B\}}$  and  $W_{\{a,b\}}$ , thereby showing that, e.g., the discrete logarithm problem in either curve model is equally hard.

The mapping from  $M_{\{A,B\}}$  to  $W_{\{a,b\}}$  is defined by mapping the point at infinity  $O$  of  $M_{\{A,B\}}$  to the point at infinity  $O$  of  $W_{\{a,b\}}$ , while mapping each other point  $(u,v)$  of  $M_{\{A,B\}}$  to the point  $(X,Y):=((u+A/3)/B, v/B)$  of  $W_{\{a,b\}}$ .

Note that not all Weierstrass curves can be mapped to Montgomery curves, since the latter have a point of order two and the former may not. In particular, if a Weierstrass curve has prime order, such as is the case with the so-called NIST prime curves, this inverse mapping is not defined.

If the Weierstrass curve  $W_{\{a,b\}}$  has a point  $(\alpha, 0)$  of order two and  $c:=a+3*(\alpha)^2$  is a square in  $GF(q)$ , one can map points of this curve to points of the Montgomery curve  $M_{\{A,B\}}$ , where  $A:=3*\alpha/\gamma$  and  $B:=1/\gamma$  and where  $\gamma$  is any square root of  $c$ . In this case, the mapping from  $W_{\{a,b\}}$  to  $M_{\{A,B\}}$  is defined by mapping the point at infinity  $O$  of  $W_{\{a,b\}}$  to the point at infinity  $O$  of  $M_{\{A,B\}}$ , while mapping each other point  $(X,Y)$  of  $W_{\{a,b\}}$  to the point  $(u,v):=((X-\alpha)/\gamma, Y/\gamma)$  of  $M_{\{A,B\}}$ . As before, this defines a one-to-one correspondence, which - in fact - is an isomorphism



between  $W_{\{a,b\}}$  and  $M_{\{A,B\}}$ . It is easy to see that the mapping from  $W_{\{a,b\}}$  to  $M_{\{A,B\}}$  and that from  $M_{\{A,B\}}$  to  $W_{\{a,b\}}$  (if defined) are each other's inverse.

This mapping can be used to implement elliptic curve group operations originally defined for a twisted Edwards curve or for a Montgomery curve using group operations for the corresponding elliptic curve in short-Weierstrass form and translating the result back to the original curve, thereby potentially allowing code reuse.

Note that implementations for elliptic curves with short-Weierstrass form that hard-code the domain parameter  $a$  to  $a = -3$  (which value is known to allow more efficient implementations) cannot always be used this way, since the curve  $W_{\{a,b\}}$  resulting from an isomorphic mapping cannot always be expressed as a Weierstrass curve with  $a = -3$  via a coordinate transformation. For more details, see Appendix F.

### D.3. Mapping between Twisted Edwards Curves and Weierstrass Curves

One can map points of the twisted Edwards curve  $E_{\{a,d\}}$  to points of the Weierstrass curve  $W_{\{a,b\}}$ , via function composition, where one uses the isomorphic mapping between twisted Edwards curves and Montgomery curves of Appendix D.1 and the one between Montgomery and Weierstrass curves of Appendix D.2. Obviously, one can use function composition (now using the respective inverses - if these exist) to realize the inverse of this mapping.

## Appendix E. Curve25519 and Cousins

This section introduces curves related to Curve25519 and explains their relationships.

### E.1. Curve Definition and Alternative Representations

The elliptic curve Curve25519 is the Montgomery curve  $M_{\{A,B\}}$  defined over the prime field  $GF(p)$ , with  $p = 2^{255} - 19$ , where  $A = 486662$  and  $B = 1$ . This curve has order  $h \cdot n$ , where  $h = 8$  and where  $n$  is a prime number. For this curve,  $A^2 - 4$  is not a square in  $GF(p)$ , whereas  $A + 2$  is. The quadratic twist of this curve has order  $h_1 \cdot n_1$ , where  $h_1 = 4$  and where  $n_1$  is a prime number. For this curve, the base point is the point  $(G_u, G_v)$ , where  $G_u = 9$  and where  $G_v$  is an odd integer in the interval  $[0, p-1]$ .

This curve has the same group structure as (is "isomorphic" to) the twisted Edwards curve  $E_{\{a,d\}}$  defined over  $GF(p)$ , with as base point the point  $(G_x, G_y)$ , where parameters are as specified in Appendix E.3. This curve is denoted as Edwards25519. For this

curve, the parameter  $a$  is a square in  $\text{GF}(p)$ , whereas  $d$  is not, so the group laws of Appendix C.3 apply.

The curve is also isomorphic to the elliptic curve  $W_{\{a,b\}}$  in short-Weierstrass form defined over  $\text{GF}(p)$ , with as base point the point  $(GX, GY)$ , where parameters are as specified in Appendix E.3. This curve is denoted as Wei25519. For this curve, the parameter  $b$  is a square in  $\text{GF}(p)$ . (For future reference, we note that this curve has no affine points with  $x$ -coordinate  $-1$ .)

## E.2. Switching between Alternative Representations

Each affine point  $(u, v)$  of Curve25519 corresponds to the point  $(X, Y) := (u + A/3, v)$  of Wei25519, while the point at infinity of Curve25519 corresponds to the point at infinity of Wei25519. (Here, we used the mappings of Appendix D.2 and that  $B=1$ .) Under this mapping, the base point  $(Gu, Gv)$  of Curve25519 corresponds to the base point  $(GX, GY)$  of Wei25519. The inverse mapping maps the affine point  $(X, Y)$  of Wei25519 to  $(u, v) := (X - A/3, Y)$  of Curve25519, while mapping the point at infinity of Wei25519 to the point at infinity of Curve25519. Note that this mapping involves a simple shift of the first coordinate and can be implemented via integer-only arithmetic as a shift of  $\delta$  for the isomorphic mapping and a shift of  $-\delta$  for its inverse, where  $\delta := (p+A)/3$  is the integer defined by

```
delta 19298681539552699237261830834781317975544997444273427339909597
      334652188435537
```

```
(=0x2aaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa
aaaaaaaa aaad2451).
```

(Note that, depending on the implementation details of the field arithmetic, one may have to shift the result by  $+p$  or  $-p$  if this integer is not in the interval  $[0, p-1]$ .)

The curve Edwards25519 is isomorphic to the curve Curve25519, where the base point  $(Gu, Gv)$  of Curve25519 corresponds to the base point  $(Gx, Gy)$  of Edwards25519 and where the point at infinity and the point  $(0,0)$  of order two of Curve25519 correspond to, respectively, the point  $(0, 1)$  and the point  $(0, -1)$  of order two of Edwards25519 and where each other point  $(u, v)$  of Curve25519 corresponds to the point  $(c*u/v, (u-1)/(u+1))$  of Edwards25519, where  $c$  is the element of  $\text{GF}(p)$  defined by

```
c  sqrt(-(A+2)/B)
```

```
51042569399160536130206135233146329284152202253034631822681833788
666877215207
```

```
(=0x70d9120b 9f5ff944 2d84f723 fc03b081 3a5e2c2e b482e57d
3391fb55 00ba81e7).
```

(Here, we used the mapping of Appendix D.1 and normalized this using the mapping of Appendix F.1 (where the element  $s$  of that appendix is set to  $c$  above).) The inverse mapping from Edwards25519 to Curve25519 is defined by mapping the point  $(0, 1)$  and the point  $(0, -1)$  of order two of Edwards25519 to, respectively, the point at infinity and the point  $(0,0)$  of order two of Curve25519 and having each other point  $(x, y)$  of Edwards25519 correspond to the point  $((1 + y)/(1 - y), c*(1 + y)/((1-y)*x))$  of Curve25519.

The curve Edwards25519 is isomorphic to the Weierstrass curve Wei25519, where the base point  $(G_x, G_y)$  of Edwards25519 corresponds to the base point  $(G_X, G_Y)$  of Wei25519 and where the identity element  $(0,1)$  and the point  $(0,-1)$  of order two of Edwards25519 correspond to, respectively, the point at infinity  $O$  and the point  $(A/3, 0)$  of order two of Wei25519 and where each other point  $(x, y)$  of Edwards25519 corresponds to the point  $(X, Y) := ((1+y)/(1-y) + A/3, c*(1+y)/((1-y)*x))$  of Wei25519, where  $c$  was defined before. (Here, we used the mapping of Appendix D.3.) The inverse mapping from Wei25519 to Edwards25519 is defined by mapping the point at infinity  $O$  and the point  $(A/3, 0)$  of order two of Wei25519 to, respectively, the identity element  $(0,1)$  and the point  $(0,-1)$  of order two of Edwards25519 and having each other point  $(X, Y)$  of Wei25519 correspond to the point  $(c*(X-A/3)/Y, (X-A/3-1)/(X-A/3+1))$  of Edwards25519.

Note that these mappings can be easily realized if points are represented in projective coordinates, using a few field multiplications only, thus allowing switching between alternative curve representations with negligible relative incremental cost.

### E.3. Domain Parameters

The parameters of the Montgomery curve and the corresponding isomorphic curves in twisted Edwards curve and short-Weierstrass form are as indicated below. Here, the domain parameters of the Montgomery curve Curve25519 and of the twisted Edwards curve Edwards25519 are as specified in [RFC7748]; the domain parameters of Wei25519 are "new".

General parameters (for all curve models):

$p = 2^{255} - 19$

```
(=0x7fffffffff ffffffff ffffffff ffffffff ffffffff ffffffff
ffffffff ffffffff)
```

h 8

n 72370055773322622139731865630429942408571163593799076060019509382  
85454250989

(=2<sup>{252}</sup> + 0x14def9de a2f79cd6 5812631a 5cf5d3ed)

h1 4

n1 14474011154664524427946373126085988481603263447650325797860494125  
407373907997

(=2<sup>{253}</sup> - 0x29bdf3bd 45ef39ac b024c634 b9eba7e3)

Montgomery curve-specific parameters (for Curve25519):

A 486662 (=0x076d06)

B 1 (=0x01)

Gu 9 (=0x09)

Gv 14781619447589544791020593568409986887264606134616475288964881837  
755586237401

(=0x20ae19a1 b8a086b4 e01edd2c 7748d14c 923d4d7e 6d7c61b2  
29e9c5a2 7eced3d9)

Twisted Edwards curve-specific parameters (for Edwards25519):

a -1 (-0x01)

d -121665/121666 = - (A-2)/(A+2)

(=370957059346694393431380835087545651895421138798432190163887855  
33085940283555)

(=0x52036cee 2b6ffe73 8cc74079 7779e898 00700a4d 4141d8ab  
75eb4dca 135978a3)

Gx 15112221349535400772501151409588531511454012693041857206046113283  
949847762202

(=0x216936d3 cd6e53fe c0a4e231 fdd6dc5c 692cc760 9525a7b2  
c9562d60 8f25d51a)

Gy 4/5

```
(=463168356949264781694283940034751631413079938662562256157830336
03165251855960)
```

```
(=0x66666666 66666666 66666666 66666666 66666666 66666666
66666666 66666658)
```

Weierstrass curve-specific parameters (for Wei25519):

```
a 19298681539552699237261830834781317975544997444273427339909597334
573241639236
```

```
(=0x2aaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa
aaaaaa98 4914a144)
```

```
b 55751746669818908907645289078257140818241103727901012315294400837
956729358436
```

```
(=0x7b425ed0 97b425ed 097b425e d097b425 ed097b42 5ed097b4
260b5e9c 7710c864)
```

```
GX 19298681539552699237261830834781317975544997444273427339909597334
652188435546
```

```
(=0x2aaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa
aaaaaaaa aaad245a)
```

```
GY 14781619447589544791020593568409986887264606134616475288964881837
755586237401
```

```
(=0x20ae19a1 b8a086b4 e01edd2c 7748d14c 923d4d7e 6d7c61b2
29e9c5a2 7eced3d9)
```

## Appendix F. Further Mappings

The non-binary curves specified in Appendix A are expressed in different curve models, viz. as curves in short-Weierstrass form, as Montgomery curves, or as twisted Edwards curves. In Appendix D we already described relationships between these various curve models. Further mappings exist between elliptic curves within the same curve model. These can be exploited to force some of the domain parameters to specific values that allow for a more efficient implementation of the addition formulae.

### F.1. Isomorphic Mapping between Twisted Edwards Curves

One can map points of the twisted Edwards curve  $E_{\{a,d\}}$  to points of the twisted Edwards curve  $E_{\{a',d'\}}$ , where  $a:=a'*s^2$  and  $d:=d'*s^2$  for some nonzero element  $s$  of  $\text{GF}(q)$ . This defines a one-to-one

correspondence, which - in fact - is an isomorphism between  $E_{\{a,d\}}$  and  $E_{\{a',d'\}}$ .

The mapping from  $E_{\{a,d\}}$  to  $E_{\{a',d'\}}$  is defined by mapping the point  $(x,y)$  of  $E_{\{a,d\}}$  to the point  $(x', y') := (s*x, y)$  of  $E_{\{a',d'\}}$ . The inverse mapping from  $E_{\{a',d'\}}$  to  $E_{\{a,d\}}$  is defined by mapping the point  $(x', y')$  of  $E_{\{a',d'\}}$  to the point  $(x, y) := (x'/s, y')$  of  $E_{\{a,d\}}$ .

Implementations may take advantage of this mapping to carry out elliptic curve group operations originally defined for a twisted Edwards curve with generic domain parameters  $a$  and  $d$  on a corresponding isomorphic twisted Edwards curve with domain parameters  $a'$  and  $d'$  that have a more special form and that are known to allow for more efficient implementations of addition laws and translating the result back to the original curve. In particular, it is known that such efficiency improvements exist if  $a' := (+/-)1$  (see [tEd-Formulas]).

## F.2. Isomorphic Mapping between Montgomery Curves

One can map points of the Montgomery curve  $M_{\{A,B\}}$  to points of the Montgomery curve  $M_{\{A',B'\}}$ , where  $A := A'$  and  $B := B' * s^2$  for some nonzero element  $s$  of  $GF(q)$ . This defines a one-to-one correspondence, which - in fact - is an isomorphism between  $M_{\{A,B\}}$  and  $M_{\{A',B'\}}$ .

The mapping from  $M_{\{A,B\}}$  to  $M_{\{A',B'\}}$  is defined by mapping the point at infinity  $O$  of  $M_{\{A,B\}}$  to the point at infinity  $O$  of  $M_{\{A',B'\}}$ , while mapping each other point  $(u,v)$  of  $M_{\{A,B\}}$  to the point  $(u', v') := (u, s*v)$  of  $M_{\{A',B'\}}$ . The inverse mapping from  $M_{\{A',B'\}}$  to  $M_{\{A,B\}}$  is defined by mapping the point at infinity  $O$  of  $M_{\{A',B'\}}$  to the point at infinity  $O$  of  $M_{\{A,B\}}$ , while mapping each other point  $(u', v')$  of  $M_{\{A',B'\}}$  to the point  $(u,v) := (u', v'/s)$  of  $M_{\{A,B\}}$ .

One can also map points of the Montgomery curve  $M_{\{A,B\}}$  to points of the Montgomery curve  $M_{\{A',B'\}}$ , where  $A' := -A$  and  $B' := -B$ . This defines a one-to-one correspondence, which - in fact - is an isomorphism between  $M_{\{A,B\}}$  and  $M_{\{A',B'\}}$ .

In this case, the mapping from  $M_{\{A,B\}}$  to  $M_{\{A',B'\}}$  is defined by mapping the point at infinity  $O$  of  $M_{\{A,B\}}$  to the point at infinity  $O$  of  $M_{\{A',B'\}}$ , while mapping each other point  $(u,v)$  of  $M_{\{A,B\}}$  to the point  $(u', v') := (-u, v)$  of  $M_{\{A',B'\}}$ . The inverse mapping from  $M_{\{A',B'\}}$  to  $M_{\{A,B\}}$  is defined by mapping the point at infinity  $O$  of  $M_{\{A',B'\}}$  to the point at infinity  $O$  of  $M_{\{A,B\}}$ , while mapping each other point  $(u', v')$  of  $M_{\{A',B'\}}$  to the point  $(u,v) := (-u', v')$  of  $M_{\{A,B\}}$ .

Implementations may take advantage of these mappings to carry out elliptic curve groups operations originally defined for a Montgomery curve with generic domain parameters  $A$  and  $B$  on a corresponding isomorphic Montgomery curve with domain parameters  $A'$  and  $B'$  that have a more special form and that are known to allow for more efficient implementations of addition laws and translating the result back to the original curve. In particular, it is known that such efficiency improvements exist if  $B'$  assumes a small absolute value, such as  $B' := (+/-)1$ . (see [Mont-Ladder]).

### F.3. Isomorphic Mapping between Weierstrass Curves

One can map points of the Weierstrass curve  $W_{\{a,b\}}$  to points of the Weierstrass curve  $W_{\{a',b'\}}$ , where  $a' := a*s^4$  and  $b' := b*s^6$  for some nonzero element  $s$  of  $GF(q)$ . This defines a one-to-one correspondence, which - in fact - is an isomorphism between  $W_{\{a,b\}}$  and  $W_{\{a',b'\}}$ .

The mapping from  $W_{\{a,b\}}$  to  $W_{\{a',b'\}}$  is defined by mapping the point at infinity  $O$  of  $W_{\{a,b\}}$  to the point at infinity  $O$  of  $W_{\{a',b'\}}$ , while mapping each other point  $(X,Y)$  of  $W_{\{a,b\}}$  to the point  $(X',Y') := (X*s^2, Y*s^3)$  of  $W_{\{a',b'\}}$ . The inverse mapping from  $W_{\{a',b'\}}$  to  $W_{\{a,b\}}$  is defined by mapping the point at infinity  $O$  of  $W_{\{a',b'\}}$  to the point at infinity  $O$  of  $W_{\{a,b\}}$ , while mapping each other point  $(X', Y')$  of  $W_{\{a',b'\}}$  to the point  $(X,Y) := (X'/s^2, Y'/s^3)$  of  $W_{\{a,b\}}$ .

Implementations may take advantage of this mapping to carry out elliptic curve group operations originally defined for a Weierstrass curve with generic domain parameters  $a$  and  $b$  on a corresponding isomorphic Weierstrass curve with domain parameter  $a'$  and  $b'$  that have a more special form and that are known to allow for more efficient implementations of addition laws and translating the result back to the original curve. In particular, it is known that such efficiency improvements exist if  $a' = -3 \pmod{p}$ , where  $p$  is the characteristic of  $GF(q)$ , and one uses so-called Jacobian coordinates with a particular projective version of the addition laws of Appendix C.1. While not all Weierstrass curves can be put into this form, all traditional NIST curves have domain parameter  $a = -3$ , while all Brainpool curves [RFC5639] are isomorphic to a Weierstrass curve of this form via the above mapping.

Note that implementations for elliptic curves with short-Weierstrass form that hard-code the domain parameter  $a$  to  $a = -3$  cannot always be used this way, since the curve  $W_{\{a,b\}}$  cannot always be expressed in terms of a Weierstrass curve with  $a' = -3$  via a coordinate transformation: this only holds if  $a'/a$  is a fourth power in  $GF(q)$  (see Section 3.1.5 of [GECC]). However, even in this case, one can

still express the curve  $W_{\{a,b\}}$  as a Weierstrass curve with a small domain parameter value  $a'$ , thereby still allowing a more efficient implementation than with a general domain parameter value  $a$ .

#### F.4. Isogenous Mapping between Weierstrass Curves

One can still map points of the Weierstrass curve  $W_{\{a,b\}}$  to points of the Weierstrass curve  $W_{\{a',b'\}}$ , where  $a' := -3 \pmod{p}$  and where  $p$  is the characteristic of  $\text{GF}(q)$ , even if  $a'/a$  is not a fourth power in  $\text{GF}(q)$ . In that case, this mapping cannot be an isomorphism (see Appendix F.3). Instead, the mapping is a so-called isogeny (or homomorphism). Since most elliptic curve operations process points of prime order or use so-called "co-factor multiplication", in practice the resulting mapping has similar properties as an isomorphism. In particular, one can still take advantage of this mapping to carry out elliptic curve group operations originally defined for a Weierstrass curve with domain parameter  $a$  unequal to  $-3 \pmod{p}$  on a corresponding isogenous Weierstrass curve with domain parameter  $a' := -3 \pmod{p}$  and translating the result back to the original curve.

In this case, the mapping from  $W_{\{a,b\}}$  to  $W_{\{a',b'\}}$  is defined by mapping the point at infinity  $O$  of  $W_{\{a,b\}}$  to the point at infinity  $O$  of  $W_{\{a',b'\}}$ , while mapping each other point  $(X,Y)$  of  $W_{\{a,b\}}$  to the point  $(X',Y') := (u(X)/w(X)^2, Y*v(X)/w(X)^3)$  of  $W_{\{a',b'\}}$ . Here,  $u(X)$ ,  $v(X)$ , and  $w(X)$  are polynomials in  $X$  that depend on the isogeny in question, as do domain parameters  $a'$  and  $b'$ . The inverse mapping from  $W_{\{a',b'\}}$  to  $W_{\{a,b\}}$  is again an isogeny (called the dual isogeny) and defined by mapping the point at infinity  $O$  of  $W_{\{a',b'\}}$  to the point at infinity  $O$  of  $W_{\{a,b\}}$ , while mapping each other point  $(X', Y')$  of  $W_{\{a',b'\}}$  to the point  $(X,Y) := (u'(X')/w'(X')^2, Y'*v'(X')/w'(X')^3)$  of  $W_{\{a,b\}}$ , where -- again --  $u'(X')$ ,  $v'(X')$ , and  $w'(X')$  are polynomials in  $X'$  that depend on the isogeny in question. These mappings have the property that their composition is not the identity mapping (as was the case with the isomorphic mappings discussed in Appendix F.3), but rather a fixed multiple hereof: if this multiple is  $l$  then the isogeny is called an isogeny of degree  $l$  (or  $l$ -isogeny) and  $u$ ,  $v$ , and  $w$  (and, similarly,  $u'$ ,  $v'$ , and  $w'$ ) are polynomials of degrees  $l$ ,  $3*(l-1)/2$ , and  $(l-1)/2$ , respectively. Note that an isomorphism is simply an isogeny of degree  $l=1$ . Details of how to determine isogenies are out of the scope of this document. The above formulas assume that the isogeny has odd degree (i.e.,  $l$  is odd); detailed formulas for even-degree isogenies are similar, but out of scope.

Implementations may take advantage of this mapping to carry out elliptic curve group operations originally defined for a Weierstrass curve with a generic domain parameter  $a$  on a corresponding isogenous



Weierstrass curve with domain parameter  $a' = -3 \pmod{p}$ , where one can use so-called Jacobian coordinates with a particular projective version of the addition laws of Appendix C.1. Since all traditional NIST curves have domain parameter  $a = -3$ , while all Brainpool curves [RFC5639] are isomorphic to a Weierstrass curve of this form, this allows taking advantage of existing implementations for these curves that may have a hardcoded  $a = -3 \pmod{p}$  domain parameter, provided one switches back and forth to this curve form using the isogenous mapping in question.

Note that isogenous mappings can be easily realized using representations in projective coordinates and involves roughly  $3 \times 1$  finite field multiplications, thus allowing switching between alternative representations at relatively low incremental cost compared to that of elliptic curve scalar multiplications (provided the isogeny has low degree  $l$ ). Note, however, that this does require storage of the polynomial coefficients of the isogeny and dual isogeny involved. This illustrates that low-degree isogenies are to be preferred, since an  $l$ -isogeny (usually) requires storing roughly  $6 \times 1$  elements of  $\text{GF}(q)$ . While there are many isogenies, we therefore only consider those with the desired property with lowest possible degree.

## Appendix G. Further Cousins of Curve25519

This section introduces some further curves related to Curve25519 and explains their relationships.

### G.1. Further Alternative Representations

The Weierstrass curve Wei25519 is isomorphic to the Weierstrass curve Wei25519.2 defined over  $\text{GF}(p)$ , with as base point the pair  $(G2X, G2Y)$ , and isogenous to the Weierstrass curve Wei25519.-3 defined over  $\text{GF}(p)$ , with as base point the pair  $(G3X, G3Y)$ , where parameters are as specified in Appendix G.3 and where the related mappings are as specified in Appendix G.2.

### G.2. Further Switching

Each affine point  $(X, Y)$  of Wei25519 corresponds to the point  $(X', Y') := (X \cdot s^2, Y \cdot s^3)$  of Wei25519.2, where  $s$  is the element of  $\text{GF}(p)$  defined by

```
s    20343593038935618591794247374137143598394058341193943326473831977
    39407761440

    (=0x047f6814 6d568b44 7e4552ea a5ed633d 02d62964 a2b0a120
    5e7941e9 375de020),
```

while the point at infinity of Wei25519 corresponds to the point at infinity of Wei25519.2. (Here, we used the mapping of Appendix F.3.) Under this mapping, the base point  $(GX, GY)$  of Wei25519 corresponds to the base point  $(G2X, G2Y)$  of Wei25519.2. The inverse mapping maps the affine point  $(X', Y')$  of Wei25519.2 to  $(X, Y) := (X'/s^2, Y'/s^3)$  of Wei25519, while mapping the point at infinity  $O$  of Wei25519.2 to the point at infinity  $O$  of Wei25519. Note that this mapping (and its inverse) involves a modular multiplication of both coordinates with fixed constants  $s^2$  and  $s^3$  (respectively,  $1/s^2$  and  $1/s^3$ ), which can be precomputed.

Each affine point  $(X, Y)$  of Wei25519 corresponds to the point  $(X', Y') := (X1*t^2, Y1*t^3)$  of Wei25519.-3, where  $(X1, Y1) = (u(X)/w(X)^2, Y*v(X)/w(X)^3)$ , where  $u$ ,  $v$ , and  $w$  are the polynomials with coefficients in  $GF(p)$  as defined in Appendix G.4.1 and where  $t$  is the element of  $GF(p)$  defined by

t 35728133398289175649586938605660542688691615699169662967154525084  
644181596229

(=0x4efd6829 88ff8526 e189f712 5999550c e9ef729b ed1a7015  
73b1bab8 8bfcd845),

while the point at infinity of Wei25519 corresponds to the point at infinity of Wei25519.-3. (Here, we used the isogenous mapping of Appendix F.4.) Under this isogenous mapping, the base point  $(GX, GY)$  of Wei25519 corresponds to the base point  $(G3X, G3Y)$  of Wei25519.-3. The dual isogeny maps the affine point  $(X', Y')$  of Wei25519.-3 to the affine point  $(X, Y) := (u'(X1)/w'(X1)^2, Y1*v'(X1)/w'(X1)^3)$  of Wei25519, where  $(X1, Y1) = (X'/t^2, Y'/t^3)$  and where  $u'$ ,  $v'$ , and  $w'$  are the polynomials with coefficients in  $GF(p)$  as defined in Appendix G.4.2, while mapping the point at infinity  $O$  of Wei25519.-3 to the point at infinity  $O$  of Wei25519. Under this dual isogenous mapping, the base point  $(G3X, G3Y)$  of Wei25519.-3 corresponds to a multiple of the base point  $(GX, GY)$  of Wei25519, where this multiple is  $l=47$  (the degree of the isogeny; see the description in Appendix F.4). Note that this isogenous map (and its dual) primarily involves the evaluation of three fixed polynomials involving the  $x$ -coordinate, which takes roughly 140 modular multiplications (or less than 5-10% relative incremental cost compared to the cost of an elliptic curve scalar multiplication).

### G.3. Further Domain Parameters

The parameters of the Weierstrass curve with  $a=2$  that is isomorphic with Wei25519 and the parameters of the Weierstrass curve with  $a=-3$  that is isogenous with Wei25519 are as indicated below. Both domain

parameter sets can be exploited directly to derive more efficient point addition formulae, should an implementation facilitate this.

General parameters: same as for Wei25519 (see Appendix E.3)

Weierstrass curve-specific parameters (for Wei25519.2, i.e., with  $a=2$ ):

```
a    2 (=0x02)

b    12102640281269758552371076649779977768474709596484288167752775713
    178787220689

    (=0x1ac1da05 b55bc146 33bd39e4 7f94302e f19843dc f669916f
    6a5dfd01 65538cd1)

G2X 10770553138368400518417020196796161136792368198326337823149502681
    097436401658

    (=0x17cfeac3 78aed661 318e8634 582275b6 d9ad4def 072ea193
    5ee3c4e8 7a940ffa)

G2Y 54430575861508405653098668984457528616807103332502577521161439773
    88639873869

    (=0x0c08a952 c55dfad6 2c4f13f1 a8f68dca dc5c331d 297a37b6
    f0d7fdcc 51e16b4d)
```

Weierstrass curve-specific parameters (for Wei25519.-3, i.e., with  $a=-3$ ):

```
a    -3

    (=0x7fffffff ffffffff ffffffff ffffffff ffffffff ffffffff
    ffffffff fffffffea)

b    29689592517550930188872794512874050362622433571298029721775200646
    451501277098

    (=0x41a3b6bf c668778e be2954a4 b1df36d1 485ecef1 ea614295
    796e1022 40891faa)

G3X 53837179229940872434942723257480777370451127212339198133697207846
    219400243292

    (=0x7706c37b 5a84128a 3884a5d7 1811f1b5 5da3230f fb17a8ab
    0b32e48d 31a6685c)
```

G3Y 69548073091100184414402055529279970392514867422855141773070804184  
60388229929

(=0x0f60480c 7a5c0e11 40340adc 79d6a2bf 0cb57ad0 49d025dc  
38d80c77 985f0329)

#### G.4. Isogeny Details

The isogeny and dual isogeny are both isogenies with degree  $l=47$ . Both are specified by a triple of polynomials  $u$ ,  $v$ , and  $w$  (resp.  $u'$ ,  $v'$ , and  $w'$ ) of degree 47, 69, and 23, respectively, with coefficients in  $\text{GF}(p)$ . The coefficients of each of these polynomials are specified in Appendix G.4.1 (for the isogeny) and in Appendix G.4.2 (for the dual isogeny). For each polynomial in variable  $x$ , the coefficients are tabulated as the sequence of coefficients of  $x^0$ ,  $x^1$ ,  $x^2$ , ..., in hexadecimal format.

##### G.4.1. Isogeny Parameters

###### G.4.1.1. Coefficients of $u(x)$

0 0x670ed14828b6f1791ceb3a9cc0edfe127dee8729c5a72ddf77bb1abaebbbale8  
1 0x1135ca8bd5383cb3545402c8bce2ced14b45c29b241e4751b035f27524a9f932  
2 0x3223806ff5f669c430efd74df8389f058d180e2fcffa5cdef3eacecdd2c34771  
3 0x31b8fecf3f17a819c228517f6cd9814466c8c8bea2efccc47a29bfc14c364266  
4 0x2541305c958c5a326f44efad2bec284e7abee840fadb08f2d994cd382fd8ce42  
5 0x6e6f9c5792f3ff497f860f44a9c469cec42bd711526b733e10915be5b2dbd8c6  
6 0x3e9ad2e5f594b9ce6b06d4565891d28a1be8790000b396ef0bf59215d6cabfde  
7 0x278448895d236403bbc161347d19c913e7df5f372732a823ed807ee1d30206be  
8 0x42f9d171ea8dc2f4a14ea46cc0ee54967175ecfe83a975137b753cb127c35060  
9 0x128e40efa2d3ccb51567e73bae91e7c31eac45700fa13ce5781cbe5ddc985648  
10 0x450e5086c065430b496d88952dd2d5f2c5102bc27074d4d1e98bfa47413e0645  
11 0x487ef93da70dfd44a4db8cb41542e33d1aa32237bdca3a59b3ce1c59585f253d  
12 0x33d209270026b1d2db96efb36cc2fa0a49be1307f49689022eab1892b010b785  
13 0x4732b5996a20ebc4d5c5e2375d3b6c4b700c681bd9904343a14a0555ef0ecd48

14 0x64dc9e8272b9f5c6ad3470db543238386f42b18cb1c592cc6caf7893141b2107  
15 0x52bbacd1f85c61ef7eafd8da27260fa2821f7a961867ed449b283036508ac5c5  
16 0x320447ed91210985e2c401cfe1a93db1379424cf748f92fd61ab5cc356bc89a2  
17 0x23d23a49bbcdf8cf4c4ce8a4ff7dd87dlad1970317686254d5b4d2ec050d019f  
18 0x1601fca063f0bbbf15f198b3c20e474c2170294fa981f73365732d2372b40cd4  
19 0x7bf3f93840035e9688cfff402cee204a17c0de9779fc33503537dd78021bf4c4  
20 0x311998ce59fb7e1cd6af591ece3e84dfcb1c330cbcf28c0349e37b9581452853  
21 0x7ae5e41acfd28a9add2216dfed34756575a19b16984c1f3847b694326dad7f99  
22 0x704957e279244a5b107a6c57bd0ab9afe5227b7c0be2052cd3513772a40efee7  
23 0x56b918b5a0c583cb763550f8f71481e57c13bdcef2e5cfc8091d0821266f233b  
24 0x677073fed43ab291e496f798fbcf217bac3f014e35d0c2fa07f041ae746a04d7  
25 0x22225388e76f9688c7d4053b50ba41d0d8b71a2f21da8353d98472243ef50170  
26 0x66930b3dffd3995a2502cef790d78b091c875192d8074bb5d5639f736400555  
27 0x79eb677c5e36971e8d64d56ebc0dedb4e9b7dd2d7b01343ebbd4d358d376e490  
28 0x48a204c2ca6d8636e9994842605bd648b91b637844e38d6c7dd707edce8256e2  
29 0x0fb3529b0d4b9ce2d70760f33e8ce997a58999718e9277caf48623d27ae6a788  
30 0x4352604bffd0c7d7a9ed898a2c6e7cf2512ffb89407271ba1f2c2d0ead8cc5aa  
31 0x6667697b29785fb6f0bd5e04d828991a5fe525370216f347ec767a26e7aac936  
32 0x09fc950b083c56dbd989badf9887255e203c879f123a7cb28901e50aea6d64dc  
33 0x41e51b51b5caadd1c15436bbf37596a1d7288a5f495d6b5b1ae66f8b2942b31d  
34 0x073b59fec709aalcabd429e981c6284822a8b7b07620c831ab41fd31d5cf7430  
35 0x67e9b88e9a1bfbc2554107d67d814986f1b09c3107a060cba21c019a2d5dc848  
36 0x6881494a1066ca176c5e174713786040affb4268b19d2abf28ef4293429f89c1  
37 0x5f4d30502ff1e1ccd624e6f506569454ab771869d7483e26afc09dea0c5ccd3d

38 0x02a814cfc5859bca51e539c159955cbe729a58978b52329575d09bc6c3bf97ad  
39 0x1313c8aaae20d6f4397f0d8b19e52cfcdf8d8e10fba144aec1778fd10ddf4e9c  
40 0x7008d38f434b98953a996d4cc79fcbe9f9502411dcdf92005f725cea7ce82ad47  
41 0x5a74d1296aaaa245ffb848f434531fa3ba9e5cb9098a7091d36c2777d4cf5a13  
42 0x4bd3b700606397083f8038177bdaa1ac6eddba0447537582723cae0fd29341a9  
43 0x573453fb2b093016f3368356c786519d54ed05f5372c01723b4da520597ec217  
44 0x77f5c605bdb3a30d7d9c8840fce38650910d4418eed707a212c8927f41c2c812  
45 0x16d6b9f7ff57ca32350057de1204cc6d69d4ef1b255dfef8080118e2fef6ace3  
46 0x34e8595832a4021f8b5744014c6b4f7da7df0d0329e8b6b4d44c8fadad6513b7  
47 0x01

#### G.4.1.2. Coefficients of $v(x)$

0 0x0f9f5eb7134e6f8dafa30c45afa58d7bfc6d4e3ccbb5de87b562fd77403972b2  
1 0x36c2dcd9e88f0d2d517a15fc453a098bbbb5a05eb6e8da906fae418a4e1a13f7  
2 0x0b40078302c24fa394a834880d5bf46732ca1b4894172fb7f775821276f558b3  
3 0x53dd8e2234573f7f3f7df11e90a7bdd7b75d807f9712f521d4fb18af59aa5f26  
4 0x6d4d7bb08de9061988a8cf6ff3beb10e933d4d2fbb8872d256a38c74c8c2ceda  
5 0x71bfe5831b30e28cd0fbe1e9916ab2291c6beacc5af08e2c9165c632e61dd2f5  
6 0x7c524f4d17ff2ee88463da012fc12a5b67d7fb5bd0ab59f4bbf162d76be1c89c  
7 0x758183d5e07878d3364e3fd4c863a5dc1fe723f48c4ab4273fc034f5454d59a4  
8 0x1eb41ef2479444ecdccbc200f64bde53f434a02b6c3f485d32f14da6aa7700e1  
9 0x1490f3851f016cc3cf8a1e3c16a53317253d232ed425297531b560d70770315c  
10 0x09bc43131964e46d905c3489c9d465c3abbd26eab9371c10e429b36d4b86469c  
11 0x5f27c173d94c7a413a288348d3fc88daa0bcf5af8f436a47262050f240e9be3b  
12 0x1d20010ec741aaa393cd19f0133b35f067adab0d105babe75fe45c8ba2732ceb

13 0x01b3c669ae49b86be2f0c946a9ff6c48e44740d7d9804146915747c3c025996a  
14 0x24c6090f79ec13e3ae454d8f0f98e0c30a8938180595f79602f2ba013b3c10db  
15 0x4650c5b5648c6c43ac75a2042048c699e44437929268661726e7182a31b1532f  
16 0x0957a835fb8bac3360b5008790e4c1f3389589ba74c8e8bf648b856ba7f22ba5  
17 0x1cd1300bc534880f95c7885d8df04a82bd54ed3e904b0749e0e3f8cb3240c7c7  
18 0x760b486e0d3c6ee0833b34b64b7ebc846055d4d1e0beeb6aedd5132399ada0ea  
19 0x1c666846c63965ef7edf519d6ada738f2b676ae38ff1f4621533373931b3220e  
20 0x365055118b38d4bc0df86648044affea2ef33e9a392ad336444e7d15e45585d1  
21 0x736487bde4b555abfccd3ea7ddcda98eda0d7c879664117dee906a88bc551194  
22 0x70de05ab9520222a37c7a84c61eedff71cb50c5f6647fc2a5d6e0ff2305cea37  
23 0x59053f6cdf6517ab3fe4bd9c9271d1892f8cf353d8041b98409e1e341a01f8b5  
24 0x375db54ed12fe8df9a198ea40200e812c2660b7022681d7932d89fafe7c6e88d  
25 0x2a070c31dlc1a064daf56c79a044bd1cd6d13f1ddb0ff039b03a6469aaa9ed77  
26 0x41482351e7f69a756a5a2c0b3fa0681c03c550341d0ca0f76c5b394db9d2de8d  
27 0x747ac1109c9e9368d94a302cb5a1d23fcc7f0fd8a574efb7ddcaa738297c407a  
28 0x45682f1f2aab6358247e364834e2181ad0448bb815c587675fb2fee5a2119064  
29 0x148c5bf44870dfd307317f0a0e4a8c163940bee1d2f01455a2e658aa92c13620  
30 0x6add1361e56ffa2d2fbbddba284b35be5845aec8069fc28af009d53290a705ce  
31 0x6631614c617400dc00f2c55357f67a94268e7b5369b02e55d5db46c935be3af5  
32 0x17cffb496c64bb89d91c8c082f4c288c3c87feabd6b08591fe5a92216c094637  
33 0x648ff88155969f54c955a1834ad227b93062bb191170dd8c4d759f79ad5da250  
34 0x73e50900b89e5f295052b97f9d0c9edb0fc7d97b7fa5e3cfeefe33dd6a9cb223  
35 0x6afcb2f2ffe6c08508477aa4956cbd3dc864257f5059685adf2c68d4f2338f00  
36 0x372fd49701954c1b8f00926a8cb4b157d4165b75d53fa0476716554bf101b74c

37 0x0334ed41325f3724ff8becbf2b3443fea6d30fa543dlca13188aceb2bdaf5f4e  
38 0x70e629c95a94e8e1b3974acb25e18ba42f8d5991786f0931f650c283adfe82fd  
39 0x738a625f4c62d3d645f1274e09ab344e72d441f3c0e82989d3e21e19212f23f3  
40 0x7093737294b29f21522f5664a9941c9b476f75d443b647bd2c777040bcd12a6a  
41 0x0a996bad5863d821ccb8b89fa329ddbe5317a46bcb32552db396bea933765436  
42 0x2da237e3741b75dd0264836e7ef634fc0bc36ab187ebc790591a77c257b06f53  
43 0x1902f3daa86fa4f430b57212924fdc9e40f09e809f3991a0b3a10ab186c50ee5  
44 0x12baffec1bf20c921afd3cdf67a7f1d87c00d5326a3e5c83841593c214dadcb1  
45 0x6460f5a68123cb9e7bc1289cd5023c0c9ccd2d98eea24484fb3825b59dcd09aa  
46 0x2c7d63a868ffc9f0fd034f821d84736c5bc33325ce98aba5f0d95fef6f230ec8  
47 0x756e0063349a702db7406984c285a9b6bfba48177950d4361d8efa77408dc860  
48 0x037f3e30032b21e0279738e0a2b689625447831a2ccf15c638672da9aa7255ae  
49 0x1107c0dbe15d6ca9e790768317a40bcf23c80f1841f03ca79dd3e3ef4ea1ae30  
50 0x61ff7f25721d6206041c59a788316b09e05135a2aad94d539c65daa68b302cc2  
51 0x5dbfe346cbd0d61b9a3b5c42ec0518d3ae81cabcc32245060d7b0cd982b8d071  
52 0x4b6595e8501e9ec3e75f46107d2fd76511764efca179f69196eb45c0aa6fade3  
53 0x72d17a5aa7bd8a2540aa9b02d9605f2a714f44abfb4c35d518b7abc39b477870  
54 0x658d8c134bac37729ec40d27d50b637201abbf1ab4157316358953548c49cf22  
55 0x36ac53b9118581ace574d5a08f9647e6a916f92dda684a4dbc405e2646b0243f  
56 0x1917a98f387d1e323e84a0f02d53307b1dd949e1a27b0de14514f89d9c0ef4b6  
57 0x21573434fde7ce56e8777c79539479441942dba535ade8ecb77763f7eb05d797  
58 0x0e0bf482dc40884719bea5503422b603f3a8edb582f52838caa6eaab6eeac7ef  
59 0x3b0471eb53bd83e14fbc13928fe1691820349a963be8f7e9815848a53d03f5eb  
60 0x1e92cb067b24a729c42d3abb7a1179c577970f0ab3e6b0ce8d66c5b8f7001262



61 0x74ea885c1ebed6f74964262402432ef184c42884fceb2f8dba3a9d67a1344dd7  
62 0x433ebce2ce9b0dc314425cfc2b234614d3c34f2c9da9fff4fddddd1ce242d035b  
63 0x33ac69e6be858dde7b83a9ff6f11de443128b39cec6e410e8d3b570e405ff896  
64 0x0dab71e2ae94e6530a501ed8cf3df26731dd1d41cd81578341e12dca3cb71aa3  
65 0x537f58d52d18ce5b1d5a6bd3a420e796e64173491ad43dd4d1083a7dcc7dd201  
66 0x49c2f6afa93fdcc4e0f8128a8b06da4c75049be14edf3e103821ab604c60f8ae  
67 0x10a333eabd6135aeaa3f5f5f7e73d102e4fd7e4bf0902fc55b00da235fa1ad08  
68 0x0f5c86044bf6032f5102e601f2a0f73c7bce9384bedd120f3e72d78484179d9c  
69 0x01

#### G.4.1.3. Coefficients of $w(x)$

0 0x3da24d42421264f30939ff00203880f2b017eb3fecf8933ae61e18df8c8ba116  
1 0x0457f20bc393cdc9a66848ce174e2fa41d77e6dbae05a317a1fb6e3ae78760f8  
2 0x7f608a2285c480d5c9592c435431fae94695beef79d770bb6d029c1d10a53295  
3 0x3832accc520a485100a0a1695792465142a5572bed1b2e50e1f8f662ac7289bb  
4 0x2df1b0559e31b328eb34beedd5e537c3f4d7b9befb0749f75d6d0d866d26fbaa  
5 0x25396820381d04015a9f655ddd41c74303ded05d54a7750e2f58006659adda28  
6 0x6fa070a70ca2bc6d4d0795fb28d4990b2cc80cd72d48b603a8ac8c8268bef6a6  
7 0x27f488578357388b20fbc7503328e1d10de602b082b3c7b8ceb33c29fea7a0d2  
8 0x15776851a7cabcf84c632118306915c0c15c75068a47021968c7438d46076e6  
9 0x101565b08a9af015c172fb194b940a4df25c4fb1d85f72d153efc79131d45e8f  
10 0x196b0ffbf92f3229fealdac0d74591b905ccaab6b83f905ee813ee8449f8a62c  
11 0x01f55784691719f765f04ee9051ec95d5deb42ae45405a9d87833855a6d95a94  
12 0x628858f79cca86305739d084d365d5a9e56e51a4485d253ae3f2e4a379fa8aff  
13 0x4a842dcd943a80d1e6e1dab3622a8c4d390da1592d1e56d1c14c4d3f72dd01a5

14 0x0f3bfc9cb17a1125f94766a4097d0f1018963bc11cb7bc0c7a1d94d65e282477  
15 0x1c4bd70488c4882846500691fa7543b7ef694446d9c3e3b4707ea2c99383e53c  
16 0x2d7017e47b24b89b0528932c4ade43f09091b91db0072e6ebdc5e777cb215e35  
17 0x781d69243b6c86f59416f91f7decaca93eab9cdc36a184191810c56ed85e0fdc  
18 0x5f20526f4177357da40a18da054731d442ad2a5a4727322ba8ed10d32eca24fb  
19 0x33e4cab64ed8a00d8012104fe8f928e6173c428eff95bbbe569ea46126a4f3cd  
20 0x050555b6f07e308d33776922b6566829d122e19b25b7bbacbb0a4b1a7dc40192  
21 0x533fa4bf1e2a2aae2f979065fdbb5b667ede2f85543fddbba146aa3a4ef2d281  
22 0x5a742cac1952010fc5aba200a635a7bed3ef868194f45b5a6a2647d6d6b289d2  
23 0x01

#### G.4.2. Dual Isogeny Parameters

##### G.4.2.1. Coefficients of $u'(x)$

0 0x0f0eddb584a20aaac8f1419efdd02a5cca77b21e4cfae78c49b5127d98bc5882  
1 0x7115e60d44a58630417df33dd45b8a546fa00b79fea3b2bdc449694bade87c0a  
2 0x0b3f3a6f3c445c7dc1f91121275414e88c32ff3f367ba0edad4d75b7e7b94b65  
3 0x1eb31bb333d7048b87f2b3d4ec76d69035927b41c30274368649c87c52e1ab30  
4 0x552c886c2044153e280832264066cce2a7da1127dc9720e2a380e9d37049ac64  
5 0x4504f27908db2e1f5840b74ae42445298755d9493141f5417c02f04d47797dda  
6 0x082c242cce1eb19698a4fa30b5affe64e5051c04ae8b52cb68d89ee85222e628  
7 0x480473406add76cf1d77661b3ff506c038d9cdd5ad6e1ea41969430bb876d223  
8 0x25f47bb506fba80c79d1763365fa9076d4c4cb6644f73ed37918074397e88588  
9 0x10f13ed36eab593fa20817f6bb70cac292e18d300498f6642e35cbdf772f0855  
10 0x7d28329d695fb3305620f83a58df1531e89a43c7b3151d16f3b60a8246c36ade  
11 0x02c5ec8c42b16dc6409bdd2c7b4ffe9d65d7209e886badbd5f865dec35e4ab4a

12 0x7f4f33cd50255537e6cde15a4a327a5790c37e081802654b56c956434354e133  
13 0x7d30431a121d9240c761998cf83d228237e80c3ef5c7191ec9617208e0ab8cec  
14 0x4d2a7d6609610c1deed56425a4615b92f70a507e1079b2681d96a2b874cf0630  
15 0x74676df60a9906901d1dc316c639ff6ae0fcdb02b5571d4b83fc2eedcd2936a8  
16 0x22f8212219aca01410f06eb234ed53bd5b8fbe7c08652b8002bcd1ea3cdae387  
17 0x7edb04449565d7c566b934a87fadade5515f23bda1ce25daa19fff0c6a5ccc2f  
18 0x106ef71aa3aa34e8ecf4c07a67d03f0949d7d015ef2c1e32eb698dd3bec5a18c  
19 0x0017913eb705db126ac3172447bcd811a62744d505ad0eea94cfcfdde5ca7428  
20 0x2cc793e6d3b592dcf5472057a991ffa5ab43b4680bb34c0f5faffc5307827c1  
21 0x6dafcc0b16f98300cddb5e0a7d7ff04a0e73ca558c54461781d5a5ccb1ea0122  
22 0x7e418891cf222c021b0ae5f5232b9c0dc8270d4925a13174a0f0ac5e7a4c8045  
23 0x76553bd26fecb019ead31142684789fea7754c2dc9ab9197c623f45d60749058  
24 0x693efb3f81086043656d81840902b6f3a9a4b0e8f2a5a5edf5ce1c7f50a3898e  
25 0x46c630eac2b86d36f18a061882b756917718a359f44752a5caf41be506788921  
26 0x01dcfa01773628753bc6f448ac11be8a3bffa0011b9284967629b827e064f614  
27 0x08430b5b97d49b0938d1f66ecb9d2043025c6eec624f8f02042b9621b2b5cb19  
28 0x66f66a6669272d47d3ec1efea36ee01d4a54ed50e9ec84475f668a5a9850f9be  
29 0x539128823b5ef3e87e901ab22f06d518a9bad15f5d375b49fe1e893ab38b1345  
30 0x2bd01c49d6fff22c213a8688924c10bf29269388a69a08d7f326695b3c213931  
31 0x3f7bea1baeccea3980201dc40d67c26db0e3b15b5a19b6cdac6de477aa717ac1  
32 0x6e0a72d94867807f7150fcb1233062f911b46e2ad11a3eac3c6c4c91e0f4a3fa  
33 0x5963f3cc262253f56fc103e50217e7e5b823ae8e1617f9e11f4c9c595fbb5bf6  
34 0x41440b6fe787777bc7b63afac9f4a38ddadcebc3d72f8fc73835247ba05f3a1d  
35 0x66d185401c1d2d0b84fcf6758a6a985bf9695651271c08f4b69ce89175fb7b34

36 0x2673fb8c65bc4fe41905381093429a2601c46a309c03077ca229bac7d6ccf239  
37 0x1ce4d895ee601918a080de353633c82b75a3f61e8247763767d146554dd2f862  
38 0x18efa6c72fa908347547a89028a44f79f22542baa588601f2b3ed25a5e56d27c  
39 0x53de362e2f8ff220f8921620a71e8faa1aa57f8886fcbb6808fa3a5560570543  
40 0x0dc29a73b97f08aa8774911474e651130ed364e8d8cffd4a80dee633aacecc47  
41 0x4e7eb8584ae4de525389d1e9300fc4480b3d9c8a5a45ecfbe33311029d8f6b99  
42 0x6c3cba4aa9229550fa82e1cfaee4b02f2c0cb86f79e0d412b8e32b00b7959d80  
43 0x5a9d104ae585b94af68eeb16b1349776b601f97b7ce716701645b1a75b68dcf3  
44 0x754e014b5e87af035b3d5fe6fb49f4631e32549f6341c6693c5172a6388e273e  
45 0x6710d8265118e22eaceba09566c86f642ab42da58c435083a353eaa12d866c39  
46 0x6e88ac659ce146c369f8b24c3a49f8dca547827250cf7963a455851cfc4f8d22  
47 0x0971eb5f253356cd1fde9fb21f4a4902aa5b8d804a2b57ba775dc130181ae2e8

#### G.4.2.2. Coefficients of $v'(x)$

0 0x043c9b67cc5b16e167b55f190db61e44d48d813a7112910f10e3fd8da85d61d3  
1 0x72046db07e0e7882ff3f0f38b54b45ca84153be47a7fd1dd8f6402e17c47966f  
2 0x1593d97b65a070b6b3f879fe3dc4d1ef03c0e781c997111d5c1748f956f1ffc0  
3 0x54e5fec076b8779338432bdc5a449e36823a0a7c905fd37f232330b026a143a0  
4 0x46328dd9bc336e0873abd453db472468393333fbf2010c6ac283933216e98038  
5 0x25d0c64de1dfe1c6d5f5f2d98ab637d8b39bcf0d886a23dabac18c80d7eb03ce  
6 0x3a175c46b2cd8e2b313dde2d5f3097b78114a6295f283cf58a33844b0c8d8b34  
7 0x5cf4e6f745bdd61181a7d1b4db31dc4c30c84957f63cdf163bee5e466a7a8d38  
8 0x639071c39b723eea51cfd870478331d60396b31f39a593ebdd9b1eb543875283  
9 0x7ea8f895dcd85fc6cb2b58793789bd9246e62fa7a8c7116936876f4d8dff869b  
10 0x503818acb535bcaacf8ad44a83c213a9ce83af7c937dc9b3e5b6efedc0a7428c

11 0x0e815373920ec3cbf3f8cae20d4389d367dc4398e01691244af90edc3e6d42b8  
12 0x7e4b23e1e0b739087f77910cc635a92a3dc184a791400cbceae056c19c853815  
13 0x145322201db4b5ec0a643229e07c0ab7c36e4274745689be2c19cfa8a702129d  
14 0x0fde79514935d9b40f52e33429621a200acc092f6e5dec14b49e73f2f59c780d  
15 0x37517ac5c04dc48145a9d6e14803b8ce9cb6a5d01c6f0ad1b04ff3353d02d815  
16 0x58ae96b8eefe9e80f24d3b886932fe3c27aaea810fa189c702f93987c8c97854  
17 0x6f6402c90fa379096d5f436035bebc9d29302126e9b117887abfa7d4b3c5709a  
18 0x01dbdf2b9ec09a8defeb485cc16ea98d0d45c5b9877ff16bd04c0110d2f64961  
19 0x53c51706af523ab5b32291de6c6b1ee7c5cbd0a5b317218f917b12ff38421452  
20 0x1b1051c7aec7d37a349208e3950b679d14e39f979db4fcd7b50d7d27dc918650  
21 0x1547e8d36262d5434cfb029cdd29385353124c3c35b1423c6cca1f87910b305b  
22 0x198efe984efc817835e28f704d41e4583a1e2398f7ce14045c4575d0445c6ce7  
23 0x492276dfe9588ee5cd9f553d990f377935d721822ecd0333ce2eb1d4324d539c  
24 0x77bad5319bacd5ed99e1905ce2ae89294efa7ee1f74314e4095c618a4e580c9b  
25 0x2cb3d532b8eac41c61b683f7b02feb9c2761f8b4286a54c3c4b60dd8081a312e  
26 0x37d189ea60443e2fee9b7ba8a34ed79ff3883dcefc06592836d2a9dd2ee3656e  
27 0x79a80f9a0e6b8ded17a3d6ccf71eb565e3704c3543b77d70bca854345e880aba  
28 0x47718530ef8e8c75f069acb2d9925c5537908e220b28c8a2859b856f46d5f8db  
29 0x7dc518f82b55a36b4fa084b05bf21e3efce481d278a9f5c6a49701e56dac01ec  
30 0x340a318dad4b8d348a0838659672792a0f00b7105881e6080a340f708a9c7f94  
31 0x55f04d9d8891636d4e9c808a1fa95ad0dae7a8492257b20448023aad3203278e  
32 0x39dc465d58259f9f70bb430d27e2f0ab384a550e1259655443e14bdecba85530  
33 0x757385464cff265379a1adfadfd6f6a03fa8a2278761d4889ab097eff4d1ac28  
34 0x4d575654dbe39778857f4e688cc657416ce524d54864ebe8995ba766efa7ca2b

35 0x47adb6aecc1949f2dc9f01206cc23eb4a0c29585d475dd24dc463c5087809298  
36 0x30d39e8b0c451a8fcf3d2abab4b86ffa374265abb77c5903db4c1be8cec7672  
37 0x28cf47b39112297f0daaaa621f8e777875adc26f35dec0ba475c2ee148562b41  
38 0x36199723cc59867e2e309fe9941cd33722c807bb2d0a06eeb41de93f1b93f2f5  
39 0x5cdeb1f2ee1c7d694bdd884cb1c5c22de206684e1cafb8d3adb9a33cb85e19a2  
40 0x0f6e6b3fc54c2d25871011b1499bb0ef015c6d0da802ae7eccf1d8c3fb73856c  
41 0x0c1422c98b672414344a9c05492b926f473f05033b9f85b8788b4bb9a080053c  
42 0x19a8527de35d4faacb00184e0423962247319703a815eecf355f143c2c18f17f  
43 0x7812dc3313e6cf093da4617f06062e8e8969d648dfe6b5c331bccd58eb428383  
44 0x61e537180c84c79e1fd2d4f9d386e1c4f0442247605b8d8904d122ee7ef9f7be  
45 0x544d8621d05540576cfc9b58a3dab19145332b88eb0b86f4c15567c37205adf9  
46 0x11be3ef96e6e07556356b51e2479436d9966b7b083892b390caec22a117aa48e  
47 0x205cda31289cf75ab0759c14c43cb30f7287969ea3dc0d5286a3853a4d403187  
48 0x048d8fc6934f4f0a99f0f2cc59010389e2a0b20d6909bfcf8d7d0249f360acdc  
49 0x42cecc6d9bdca6d382e97fcea46a79c3eda2853091a8f399a2252115bf9a1454  
50 0x0117d41b24f2f69cb3270b359c181607931f62c56d070bbd14dc9e3f9ab1432e  
51 0x7c51564c66f68e2ad4ce6ea0d68f920fafa375376709c606c88a0ed44207aa1e  
52 0x48f25191fc8ac7d9f21adf6df23b76ccbca9cb02b815acdbefbfa3f4eddc71b34  
53 0x4fc21a62c4688de70e28ad3d5956633fc9833bc7be09dc7bc500b7fae1e1c9a8  
54 0x1f23f25be0912173c3ef98e1c9990205a69d0bf2303d201d27a5499247f06789  
55 0x3131495618a0ac4cb11a702f3f8bab66c4fa1066d0a741af3c92d5c246edd579  
56 0x0d93fe40faa53913638e497328a1b47603cb062c7afc9e96278603f29fd11fd4  
57 0x6b348bc59e984c91d696d1e3c3cfae44021f06f74798c787c355437fb696093d  
58 0x65af00e73043edcb479620c8b48098b89809d577a4071c8e33e8678829138b8a

59 0x5e62ffb032b2ddb06591f86a46a18effd5d6ecf3f129bb2bacfd51a3739a98b6  
60 0x62c974ef3593fc86f7d78883b8727a2f7359a282cbc0196948e7a793e60ce1a1  
61 0x204d708e3f500aad64283f753e7d9bab976aa42a4ca1ce5e9d2264639e8b1110  
62 0x0a90f0059da81a012e9d0a756809fab2ce61cb45965d4d1513a06227783ee4ea  
63 0x39fa55971c9e833f61139c39e243d40869fd7e8a1417ee4e7719dd2dd242766f  
64 0x22677c1e659caa324f0c74a013921facf62d0d78f273563145cc1ddccfcc4421  
65 0x3468cf6df7e93f7ff1fe1dd7e180a89dec3ed4f72843b4ea8a8d780011a245b2  
66 0x68f75a0e2210f52a90704ed5f511918d1f6bcfcd26b462cc4975252369db6e9d  
67 0x6220c0699696e9bcab0fe3a80d437519bd2bdf3caef665e106b2dd47585ddd9f  
68 0x553ad47b129fb347992b576479b0a89f8d71f1196f83e5eaab5f533a1dd6f6d7  
69 0x239aef387e116ec8730fa15af053485ca707650d9f8917a75f22acf6213197df

#### G.4.2.3. Coefficients of $w'(x)$

0 0x6bd7f1fc5dd51b7d832848c180f019bcbdb101d4b3435230a79cc4f95c35e15e  
1 0x17413bb3ee505184a504e14419b8d7c8517a0d268f65b0d7f5b0ba68d6166dd0  
2 0x47f4471beed06e5e2b6d5569c20e30346bdba2921d9676603c58e55431572f90  
3 0x2af7eaaafd04f6910a5b01cdb0c27dca09487f1cd1116b38db34563e7b0b414eb  
4 0x57f0a593459732eef11d2e2f7085bf9adf534879ba56f7afd17c4a40d3d3477b  
5 0x4da04e912f145c8d1e5957e0a9e44cca83e74345b38583b70840bdfdbd0288ed  
6 0x7cc9c3a51a3767d9d37c6652c349adc09bfe477d99f249a2a7bc803c1c5f39ed  
7 0x425d7e58b8adf87eebf445b424ba308ee7880228921651995a7eab548180ad49  
8 0x48156db5c99248234c09f43fedf509005943d3d5f5d7422621617467b06d314f  
9 0x0d837dbbd1af32d04e2699cb026399c1928472aa1a7f0a1d3afd24bc9923456a  
10 0x5b8806e0f924e67c1f207464a9d025758c078b43ddc0ea9afe9993641e5650be  
11 0x29c91284e5d14939a6c9bc848908bd9df1f8346c259bbd40f3ed65182f3a2f39

```
12 0x25550b0f3bceef18a6bf4a46c45bf1b92f22a76d456bdfd19d07398c80b0f946
13 0x495d289b1db16229d7d4630cb65d52500256547401f121a9b09fb8e82cf01953
14 0x718c8c610ea7048a370eabfd9888c633ee31dd70f8bcc58361962bb08619963e
15 0x55d8a5ceef588ab52a07fa6047d6045550a5c52c91cc8b6b82eeb033c8ca557d
16 0x620b5a4974cc3395f96b2a0fa9e6454202ef2c00d82b0e6c534b3b1d20f9a572
17 0x4991b763929b00241a1a9a68e00e90c5df087f90b3352c0f4d8094a51429524e
18 0x18b6b49c5650fb82e36e25fd4eb6decfdd40b46c37425e6597c7444a1b6afb4e
19 0x6868305b4f40654460aad63af3cb9151ab67c775eaac5e5df90d3aea58dee141
20 0x16bc90219a36063a22889db810730a8b719c267d538cd28fa7c0d04f124c8580
21 0x3628f9cf1fbe3eb559854e3b1c06a4cd6a26906b4e2d2e70616a493bba2dc574
22 0x64abcc6759f1ce1ab57d41e17c2633f717064e35a7233a6682f8cf8e9538afec
23 0x01
```

#### Appendix H. Point Compression

Point compression allows a shorter representation of affine points of an elliptic curve by exploiting algebraic relationships between the coordinate values based on the defining equation of the curve in question. Point decompression refers to the reverse process, where one tries and recover an affine point from its compressed representation and information on the domain parameters of the curve. Consequently, point compression followed by point decompression is the identity map.

The description below makes use of an auxiliary function (the parity function), which we first define for prime fields  $\text{GF}(p)$ , with  $p$  odd, and then extend to all fields  $\text{GF}(q)$ , where  $q$  is an odd prime power. We assume each finite field to be unambiguously defined and known from context.

Let  $y$  be a nonzero element of  $\text{GF}(q)$ . If  $q=p$  is an odd prime number,  $y$  and  $p-y$  can be uniquely represented as integers in the interval  $[1, p-1]$  and have odd sum  $p$ . Consequently, one can distinguish  $y$  from  $-y$  via the parity of this representation, i.e., via  $\text{par}(y) := y \pmod{2}$ . If  $q=p^m$ , where  $p$  is an odd prime number and where  $m>0$ , both  $y$  and  $-y$  can be uniquely represented as vectors of length  $m$ , with coefficients in  $\text{GF}(p)$  (see Appendix B.2). In this case, the leftmost



nonzero coordinate values of  $y$  and  $-y$  are in the same position and have representations in  $[1, p-1]$  with different parity. As a result, one can distinguish  $y$  from  $-y$  via the parity of the representation of this coordinate value. This extends the definition of the parity function to any odd-size field  $\text{GF}(q)$ , where one defines  $\text{par}(0) := 0$ . The value of the parity function is commonly called the parity bit.

#### H.1. Point Compression for Weierstrass Curves

If  $P := (X, Y)$  is an affine point of the Weierstrass curve  $W_{\{a,b\}}$  defined over the field  $\text{GF}(q)$ , then so is  $-P := (X, -Y)$ . Since the defining equation  $Y^2 = X^3 + aX + b$  has at most two solutions with fixed  $X$ -value, one can represent  $P$  by its  $X$ -coordinate and one bit of information that allows one to distinguish  $P$  from  $-P$ , i.e., one can represent  $P$  as the ordered pair  $\text{compr}(P) := (X, \text{par}(Y))$ . If  $P$  is a point of order two, one can uniquely represent  $P$  by its  $X$ -coordinate alone, since  $Y=0$  and has fixed parity. Conversely, given the ordered pair  $(X, t)$ , where  $X$  is an element of  $\text{GF}(q)$  and where  $t=0$  or  $t=1$ , and the domain parameters of the curve  $W_{\{a,b\}}$ , one can use the defining equation of the curve to try and determine candidate values for the  $Y$ -coordinate given  $X$ , by solving the quadratic equation  $Y^2 := \alpha$ , where  $\alpha := X^3 + aX + b$ . If  $\alpha$  is not a square in  $\text{GF}(q)$ , this equation does not have a solution in  $\text{GF}(q)$  and the ordered pair  $(X, t)$  does not correspond to a point of this curve. Otherwise, there are two solutions, viz.  $Y = \sqrt{\alpha}$  and  $-Y$ . If  $\alpha$  is a nonzero element of  $\text{GF}(q)$ , one can uniquely recover the  $Y$ -coordinate for which  $\text{par}(Y) := t$  and, thereby, the point  $P := (X, Y)$ . This is also the case if  $\alpha=0$  and  $t=0$ , in which case  $Y=0$  and the point  $P$  has order two. However, if  $\alpha=0$  and  $t=1$ , the ordered pair  $(X, t)$  does not correspond to the outcome of the point compression function.

NOTE: the procedure above corrects an error in the point decompression procedure for Weierstrass curves defined over the prime field  $\text{GF}(p)$  of [SEC1], which erroneously converts a purported compressed point for which  $\alpha=0$  and  $t=1$  (in the notation above), to the ordered pair  $(0, p)$ .

We extend the definition of the point compression function to all points of the curve  $W_{\{a,b\}}$ , by associating the (non-affine) point at infinity  $O$  with any ordered pair  $\text{compr}(O) := (X, 0)$ , where  $X$  is any element of  $\text{GF}(q)$  for which  $\alpha := X^3 + aX + b$  is not a square in  $\text{GF}(q)$ , and recover this point accordingly. In this case, the point at infinity  $O$  can be represented by any ordered pair  $(X, 0)$  of elements of  $\text{GF}(q)$  for which  $X^3 + aX + b$  is not a square in  $\text{GF}(q)$ . Note that this ordered pair does not satisfy the defining equation of the curve in question. An application may fix a specific suitable value of  $X$  or choose multiple such values and use this to encode additional information. Further details are out of scope.

## H.2. Point Compression for Montgomery Curves

If  $P := (u, v)$  is an affine point of the Montgomery curve  $M_{\{A,B\}}$  defined over the field  $GF(q)$ , then so is  $-P := (u, -v)$ . Since the defining equation  $B*v^2 = u^3 + A*u^2 + u$  has at most two solutions with fixed  $u$ -value, one can represent  $P$  by its  $u$ -coordinate and one bit of information that allows one to distinguish  $P$  from  $-P$ , i.e., one can represent  $P$  as the ordered pair  $\text{compr}(P) := (u, \text{par}(v))$ . If  $P$  is a point of order two, one can uniquely represent  $P$  by its  $u$ -coordinate alone, since  $v=0$  and has fixed parity. Conversely, given the ordered pair  $(u, t)$ , where  $u$  is an element of  $GF(q)$  and where  $t=0$  or  $t=1$ , and the domain parameters of the curve  $M_{\{A,B\}}$ , one can use the defining equation of the curve to try and determine candidate values for the  $v$ -coordinate given  $u$ , by solving the quadratic equation  $v^2 := \alpha$ , where  $\alpha := (u^3 + A*u^2 + u)/B$ . If  $\alpha$  is not a square in  $GF(q)$ , this equation does not have a solution in  $GF(q)$  and the ordered pair  $(u, t)$  does not correspond to a point of this curve. Otherwise, there are two solutions, viz.  $v = \sqrt{\alpha}$  and  $-v$ . If  $\alpha$  is a nonzero element of  $GF(q)$ , one can uniquely recover the  $v$ -coordinate for which  $\text{par}(v) := t$  and, thereby, the affine point  $P := (u, v)$ . This is also the case if  $\alpha=0$  and  $t=0$ , in which case  $v=0$  and the point  $P$  has order two. However, if  $\alpha=0$  and  $t=1$ , the ordered pair  $(u, t)$  does not correspond to the outcome of the point compression function.

We extend the definition of the point compression function to all points of the curve  $M_{\{A,B\}}$ , by associating the (non-affine) point at infinity  $O$  with the ordered pair  $\text{compr}(O) := (0, 1)$  and recover this point accordingly. (Note that this corresponds to the case  $\alpha=0$  and  $t=1$  above.) The point at infinity  $O$  can be represented by the ordered pair  $(0, 1)$  of elements of  $GF(q)$ . Note that this ordered pair does not satisfy the defining equation of the curve in question.

## H.3. Point Compression for Twisted Edwards Curves

If  $P := (x, y)$  is an affine point of the twisted Edwards curve  $E_{\{a,d\}}$  defined over the field  $GF(q)$ , then so is  $-P := (-x, y)$ . Since the defining equation  $a*x^2 + y^2 = 1 + d*x^2*y^2$  has at most two solutions with fixed  $y$ -value, one can represent  $P$  by its  $y$ -coordinate and one bit of information that allows one to distinguish  $P$  from  $-P$ , i.e., one can represent  $P$  as the ordered pair  $\text{compr}(P) := (\text{par}(x), y)$ . If  $P$  is a point of order one or two, one can uniquely represent  $P$  by its  $y$ -coordinate alone, since  $x=0$  and has fixed parity. Conversely, given the ordered pair  $(t, y)$ , where  $y$  is an element of  $GF(q)$  and where  $t=0$  or  $t=1$ , and the domain parameters of the curve  $E_{\{a,d\}}$ , one can use the defining equation of the curve to try and determine candidate values for the  $x$ -coordinate given  $y$ , by solving the quadratic equation  $x^2 := \alpha$ , where  $\alpha := (1 - y^2)/(a - d*y^2)$ .

(Here, observe that the denominator is nonzero for any point of  $E_{\{a,d\}}$ .) If  $\alpha$  is not a square in  $GF(q)$ , this equation does not have a solution in  $GF(q)$  and the ordered pair  $(t, y)$  does not correspond to a point of this curve. Otherwise, there are two solutions, viz.  $x=\sqrt{\alpha}$  and  $-x$ . If  $\alpha$  is a nonzero element of  $GF(q)$ , one can uniquely recover the  $x$ -coordinate for which  $\text{par}(x):=t$  and, thereby, the affine point  $P:=(x, y)$ . This is also the case if  $\alpha=0$  and  $t=0$ , in which case  $x=0$  and the point  $P$  has order one or two. However, if  $\alpha=0$  and  $t=1$ , the ordered pair  $(t, y)$  does not correspond to the outcome of the point compression function.

Note that the point compression function is defined for all points of the twisted Edwards curve  $E_{\{a,d\}}$ . Here, the identity element  $O:=(0,1)$  is associated with the compressed point  $\text{compr}(O):=(0,1)$ . (Note that this corresponds to the case  $\alpha=0$  and  $t=0$  above.)

We extend the definition of the compression function further, to also include a special marker element 'btm', by associating this marker element with the ordered pair  $\text{compr}(\text{btm}):=(1,1)$  and recover this marker element accordingly. (Note that this corresponds to the case  $\alpha=0$  and  $t=1$  above.) The marker element 'btm' can be represented by the ordered pair  $(1,1)$  of elements of  $GF(q)$ . Note that this ordered pair does not satisfy the defining equation of the curve in question.

## Appendix I. Data Conversions

This section introduces various data conversion routines that are useful when representing integers, finite field elements, and curve points as binary or octet strings.

### I.1. Strings and String Operations

The string over some alphabet  $S$  consisting of the symbols  $x_{\{1-1\}}$ ,  $x_{\{1-2\}}$ , ...,  $x_1$ ,  $x_0$  (each in  $S$ ), in this order, is denoted by  $\text{str}(x_{\{1-1\}}, x_{\{1-2\}}, \dots, x_1, x_0)$  (or simply as  $x_{\{1-1\}} x_{\{1-2\}} \dots x_1 x_0$ , if the individual symbols can be uniquely identified). The length of this string (over  $S$ ) is the number of symbols it contains (here: 1). The empty string is the (unique) string of length  $l=0$ . Strings are commonly indicated by surrounding these by double quotation marks.

The right-concatenation of two strings  $X$  and  $Y$  (defined over the same alphabet) is the string  $Z$  consisting of the symbols of  $X$  (in the same order) followed by the symbols of  $Y$  (in the same order). The length of the resulting string  $Z$  is the sum of the lengths of  $X$  and  $Y$ . This string operation is denoted by  $Z:=X||Y$ . The string  $X$  is called a prefix of  $Z$ ; the string  $Y$  a postfix of  $Z$ . The  $t$ -prefix of a string  $Z$

of length  $l$  is its unique prefix  $X$  of length  $t$ ; the  $t$ -postfix its unique postfix  $Y$  of length  $t$  (where in both cases  $t$  is an integer in the interval  $[0, l]$ ). One can define these notions as well if  $t$  is outside the interval  $[0, l]$  by stipulating that a  $t$ -prefix or  $t$ -postfix is the empty string if  $t$  is negative and that it is the entire string  $Z$  if  $t$  is larger than  $l$ . Sometimes, a  $t$ -prefix of a string  $Z$  is denoted by  $\text{trunc-left}(Z, t)$ ; a  $t$ -postfix by  $\text{trunc-right}(Z, t)$ . A string  $X$  is called a substring of  $Z$  if it is a prefix of some postfix of  $Z$ . The string resulting from prepending the string  $Y$  with  $X$  is the string  $X||Y$ . The symbols of a string  $Z$  of length  $l$  can be labelled from left to right, using consecutive integers in the interval  $[0, l)$  starting with zero, where each label identifies a symbol via its position in the string.

An octet (or byte) is an integer in the interval  $[0, 256)$ . An octet string is a string, where the alphabet is the set of all octets. A binary string (or bit string, for short) is a string, where the alphabet is the set  $\{0, 1\}$  of binary digits. A hex digit is an integer in the interval  $[0, 16)$ , with the convention to denote the integers 10, 11, 12, 13, 14, and 15 by the symbols 'a', 'b', 'c', 'd', 'e', and 'f', respectively. A hexadecimal string (or hex string, for short) is a string, where the alphabet is the set of all hex digits. Note that the length of a string is defined in terms of the underlying alphabet, as are the operations in the previous paragraph.

Note that an octet  $z$  can be uniquely represented in base 16 as the integer  $z := 16 \cdot z_1 + z_0$ , where  $z_1$  and  $z_0$  are hex digits, and, thereby, as the hexadecimal string  $z_1||z_0$  of length two. This allows a concise description of octet strings as hex strings (commonly indicated by the "0x"-prefix).

An ASCII character is a symbol of the so-called ASCII alphabet: the set of symbols corresponding to the set of integers in the interval  $[0, 128)$  according to the ASCII-table (see [RFC0020]). An ASCII string is a string, where the alphabet is the set of all ASCII characters. All ASCII characters corresponding to integers in the interval  $[33, 126]$  are single printable characters (and can therefore be uniquely identified in a printable ASCII string). There is a 1-1 correspondence between ASCII characters and integers in the interval  $[0, 128)$ , thereby allowing each ASCII character to be uniquely represented by an octet.

## I.2. Conversion between Bit Strings and Integers (BS2I, I2BS)

There is a 1-1 correspondence between bit strings of length  $l$  and integers in the interval  $[0, 2^l)$ , where the bit string  $X := \text{str}(x_{l-1}, x_{l-2}, \dots, x_1, x_0)$  corresponds to the integer

$x := x_{l-1} \cdot 2^{l-1} + x_{l-2} \cdot 2^{l-2} + \dots + x_1 \cdot 2 + x_0 \cdot 1$ . (If  $l=0$ , the empty bit string corresponds to the integer zero.) Note that while the mapping from bit strings to integers is uniquely defined, the inverse mapping from integers to bit strings is not, since any non-negative integer smaller than  $2^t$  can be represented as a bit string of length at least  $t$  (due to leading zero coefficients in base 2 representation). The latter representation is called tight if the bit string representation has minimal length (the so-called bit-length of the integer in question). This defines the mapping BS2I from bit strings to integers and the mapping I2BS( $x, l$ ) from non-negative integers smaller than  $2^l$  to bit strings of length  $l$ . Note that this also defines a 1-1 correspondence between bit strings of length four and hex digits, and the encoding of ASCII characters as bit strings of length eight (where the leftmost bit has the value zero), as stipulated in [RFC0020].

### I.3. Conversion between Octet Strings and Integers (OS2I, I2OS)

There is a 1-1 correspondence between octet strings of length  $l$  and integers in the interval  $[0, 256^l)$ , where the octet string  $X := \text{str}(X_{l-1}, X_{l-2}, \dots, X_1, X_0)$  corresponds to the integer  $x := X_{l-1} \cdot 256^{l-1} + X_{l-2} \cdot 256^{l-2} + \dots + X_1 \cdot 256 + X_0 \cdot 1$ . (If  $l=0$ , the empty string corresponds to the integer zero.) Note that while the mapping from octet strings to integers is uniquely defined, the inverse mapping from integers to octet strings is not, since any non-negative integer smaller than  $256^t$  can be represented as an octet string of length at least  $t$  (due to leading zero coefficients in base 256 representation). The latter representation is called tight if the octet string representation has minimal length (the so-called byte-length of the integer in question). This defines the mapping OS2I from octet strings to integers and the mapping I2OS( $x, l$ ) from non-negative integers smaller than  $256^l$  to octet strings of length  $l$ .

### I.4. Conversion between Octet Strings and Bit Strings (OS2BS, BS2OS)

There is a 1-1 correspondence between octet strings of length  $l$  and bit strings of length  $8 \cdot l$ , where the octet string  $X := \text{str}(X_{l-1}, X_{l-2}, \dots, X_1, X_0)$  corresponds to the right-concatenation of the 8-bit strings  $x_{l-1}, x_{l-2}, \dots, x_1, x_0$ , where each octet  $X_i$  corresponds to the 8-bit string  $x_i$  according to the mapping of Appendix I.2 above. Note that the mapping from octet strings to bit strings is uniquely defined and so is the inverse mapping from bit strings to octet strings, if one prepends each bit string with the smallest number of 0 bits so as to result in a bit string of length divisible by eight (i.e., one uses pre-padding). This defines the mapping OS2BS from octet strings to bit strings and the corresponding mapping BS2OS from bit strings to octet strings. When we refer to a

specific bit position in an octet string, this indicates the corresponding position, when this octet string is viewed as a bit string using the OS2BS mapping above.

#### I.5. Conversion between Field Elements and Octet Strings (FE2OS, OS2FE)

There is a 1-1 correspondence between elements of the fixed finite field  $GF(q)$ , where  $q=p^m$ , where  $p$  is a prime number and where  $m>0$ , and vectors of length  $m$ , with coefficients in  $GF(p)$ , where each element  $x$  of  $GF(q)$  is a vector  $(x_{m-1}, x_{m-2}, \dots, x_1, x_0)$  according to the conventions of Appendix B.2. In this case, this field element can be uniquely represented by the right-concatenation of the octet strings  $X_{m-1}, X_{m-2}, \dots, X_1, X_0$ , where each octet string  $X_i$  corresponds to the integer  $x_i$  in the interval  $[0, p-1]$  according to the mapping of Appendix I.3 above. Note that both the mapping from field elements to octet strings and the inverse mapping from octet strings to field elements are only uniquely defined if each octet string  $X_i$  has the same fixed size (e.g., the smallest integer  $l$  so that  $256^l \geq p$ ) and if all integers are reduced modulo  $p$ . If so, the latter representation is called tight if  $l$  is minimal so that  $256^l \geq p$ . This defines the mapping  $FE2OS(x, l)$  from field elements to octet strings and the mapping  $OS2FE(X, l)$  from octet strings to field elements, where the underlying field is implicit and assumed to be known from context. In this case, the octet string has length  $l \cdot m$ . (Observe that with tight representations, the parameter  $l$  is uniquely defined by the characteristic  $p$  of the field  $GF(q)$  in question.) The  $OS2FE(X, l)$  mapping is called strict if it operates as the  $OS2FE(X, l)$  function, except that it fails whenever it would require at least one modular reduction. Notice that the tight  $FE2OS$  mapping followed by the strict  $OS2FE$  mapping is the identity map (and, hence,  $OS2FE$  never fails in this case).

#### I.6. Conversion between Elements of $Z_n$ and Octet Strings ( $ZnE2OS$ , $OS2ZnE$ )

There is a 1-1 correspondence between elements of the set  $Z_n$  of integers modulo  $n$  and integers in the interval  $[0, n)$ , where each element  $x$  of  $Z_n$  is uniquely represented by the integer  $x \bmod n$ . In this case,  $x \bmod n$  can be uniquely represented by the octet string  $X$  according to the mapping of Appendix I.3 above. Note that both the mapping from elements of  $Z_n$  to octet strings and the inverse mapping from octet strings to elements of  $Z_n$  are only uniquely defined if the octet string has a fixed size (e.g., the smallest integer  $l$  so that  $256^l \geq n$ ) and if all integers are first reduced modulo  $n$ . If so, the latter representation is called tight if  $l$  is minimal so that  $256^l \geq n$ . This defines the mapping  $ZnE2OS(x, l)$  from elements of  $Z_n$  to octet strings and the mapping  $OS2ZnE(X, l)$  from octet strings

to elements of  $Z_n$ , where the underlying modulus  $n$  is implicit and assumed to be known from context. In this case, the octet string has length  $l$ . (Observe that with tight representations, the parameter  $l$  is uniquely defined by the parameter  $n$  in question.) The  $OS2ZnE(X, l)$  mapping is called strict if it operates as the  $OS2ZnE(X, l)$  function, except that it fails whenever it would require at least one modular reduction. Notice that the tight  $ZnE2OS$  mapping followed by the strict  $OS2ZnE$  mapping is the identity map (and, hence,  $ZnE2OS$  never fails in this case).

Note that if  $n$  is a prime number  $p$ , the conversions  $ZnE2OS$  and  $FE2OS$  are consistent, as are  $OS2ZnE$  and  $OS2FE$ . This is, however, no longer the case if  $n$  is a strict prime power.

The conversion rules for composite (i.e., non-prime)  $n$  values may be useful, e.g., when encoding RSA parameters (or elements of any other non-prime size set  $Z_n$ , for that matter).

#### I.7. Ordering Conventions

One can consider various representation functions, depending on bit-ordering and octet-ordering conventions.

The description below makes use of an auxiliary function (the reversion function), where the reverse of the string  $X := \text{str}(x_{l-1}, x_{l-2}, \dots, x_1, x_0)$  is defined to be the string  $X' := \text{rev}(X) := \text{str}(x_0, x_1, \dots, x_{l-2}, x_{l-1})$ . Below, we use this reversion function with binary and octet strings.

We now describe representations in most-significant-bit first (msb) or least-significant-bit first (lsb) order and those in most-significant-byte first (MSB) or least-significant-byte first (LSB) order.

One distinguishes the following octet-string representations of integers and field elements:

1. MSB, msb: represent field elements and integers as above, yielding the octet string  $\text{str}(X_{l-1}, X_{l-2}, \dots, X_1, X_0)$ .
2. MSB, lsb: reverse the bit-order of each octet, viewed as 8-bit string, yielding the octet string  $\text{str}((\text{rev}(X_{l-1})), \text{rev}(X_{l-2}), \dots, \text{rev}(X_1), \text{rev}(X_0))$ .
3. LSB, lsb: reverse the octet string and bit-order of each octet, yielding the octet string  $\text{str}(\text{rev}(X_{l-1}), \text{rev}(X_{l-2}), \dots, \text{rev}(X_1), \text{rev}(X_0))$ .

4. LSB, msb: reverse the octet string, yielding the octet string  $\text{str}(X_{\{0\}}, X_{\{1\}}, \dots, X_{\{l-2\}}, X_{\{l-1\}})$ .

Thus, the 2-octet string "07e3" represents the integer 2019 (=0x07e3) in MSB/msb order, the integer 57,543 (0xe0c7) in MSB/lsb order, the integer 51,168 (0xc7e0) in LSB/lsb order, and the integer 58,119 (=0xe307) in LSB/msb order.

Note that, with the above data conversions, there is still some ambiguity as to how to represent an integer or a field element as a bit string or octet string (due to leading zeros). However, tight representations (as defined above) are non-ambiguous. (Note, in particular, that tightness implies that elements of  $\text{GF}(q)$  are always uniquely represented.)

#### I.8. Conversion Between Curve Points and Octet Strings

For each of the curve models we consider, each affine point is an ordered pair  $(X, Y)$  whose coordinates are elements of a finite field  $\text{GF}(q)$  and that satisfy the defining equation for the curve in question. Each compressed point is an ordered pair  $(X, t)$  (for Weierstrass curves and Montgomery curves) or  $(t, X)$  (for twisted Edwards curves) where  $X$  is an element of  $\text{GF}(q)$  and where  $t$  is an element of  $\{0,1\}$  (see Appendix H).

The affine point  $(X, Y)$  is represented by the ordered pair whose coordinates are the octet string representations of the elements  $X$  and  $Y$  of  $\text{GF}(q)$ , respectively, using the tight FE2OS mapping of Appendix I.5. Note that, since we use a tight representation, this results in a pair of octet strings (each of length  $l*m$ ), where the parameters  $l$  and  $m$  are uniquely defined by the field  $\text{GF}(q)$  in question. The inverse mapping results by converting the first and second coordinate of this pair (each an octet string of length  $l*m$ ) to, respectively, the elements  $X$  and  $Y$  of  $\text{GF}(q)$  via the strict OS2FE mapping of Appendix I.5. Note that if it is not a priori known whether the input to this inverse mapping actually represents an affine curve point, one should check that the output  $(X,Y)$  -- if defined -- is indeed an affine point of the curve in question, where this operation fails if this is not the case. (This check involves simply checking whether the ordered pair  $(X,Y)$  satisfies the defining equation for this curve.)

The compressed point  $(X, t)$  or  $(t, X)$  is represented by the ordered pair whose coordinates are the octet string representations of the parity bit  $t$  in  $\{0,1\}$  and the element  $X$  of  $\text{GF}(q)$ , respectively, using the tight FE2OS mapping of Appendix I.5. Note that, since we use tight representations, this results in an ordered pair of octet strings (of length 1 and  $l*m$ , respectively), where the parameters  $l$



and  $m$  are uniquely defined by the field  $GF(q)$  in question. The inverse mapping results by converting the first and second coordinate of this pair (each an octet string, of length  $l$  and  $l*m$ , respectively) to, respectively, the element  $t$  of  $\{0,1\}$  and the element  $X$  of  $GF(q)$  via the strict OS2FE mapping of Appendix I.5, and representing this as the compressed point  $(X, t)$  or  $(t, X)$  according to the curve model in question. Note that if it is not a priori known whether the input to this inverse mapping actually represents a compressed curve point, one should check that the output  $(X, t)$  or  $(t, X)$  -- if defined -- is indeed a compressed point of the curve in question, using the point decompression process for this curve (see Appendix H), where this operation fails if this is not the case. (This check does include checking whether an element is a square in  $GF(q)$ , but does not require actually computing square roots (see also the Note in Appendix K.1).)

NOTE 1: The representations of affine and compressed points above are as ordered pairs of octet strings. In practice, one often represents these as octet strings instead, via right-concatenation of its coordinates (in left-to-right order). Since each coordinate has known length, this operation is reversible. When appropriate, we refer to the latter as the octet (rather than the pair) representation of a point.

NOTE 2: The octet representation of compressed points above identifies the parity bit  $t$  of the curve point in question via the 1-octet representations of the integers 0 and 1. Obviously, other 1-1 mappings are also possible. As an example, with [SEC1], the parity bit  $t$  is represented by 0x02 or 0x03 depending on whether  $t=0$  or  $t=1$ , respectively. The same [SEC1] specification represents affine points as above (as octet string), but prepends this with the 1-octet prefix 0x04, and represents the identity element of the curve as the 1-octet string 0x00. This variable-size point representation has the property that its 1-octet prefix identifies whether it encodes an affine curve point, a compressed point (including parity bit), or the identity element, while the remainder of this representation uniquely determines the curve point's value. While the description in [SEC1] only applies to Weierstrass curves, the description above applies to each of the curve models we consider (i.e., these apply to Montgomery curves and twisted Edwards curves as well) and also applies to curves defined over extension fields. Collectively, we simply refer to this as the "SEC1" point representation.

Note that elements of a prime field  $GF(p)$ , where  $p$  is a 255-bit prime number, have a tight representation as a 32-octet string, where a fixed bit position is always set to zero. (This is the leftmost bit position of this octet string if one follows the MSB/msb

representation conventions.) This allows the parity bit of a compressed point (see Appendix H) to be encoded in this bit position and, thereby, allows a compressed point and an element of  $GF(p)$  to be represented by an octet string of the same length. This is called the "squeezed" point representation. (We will use this squeezed representation in Appendix J.) Obviously, other representations (e.g., those of elements of  $Z_n$ ) may also have fixed bit values in certain positions, which can be used to squeeze-in additional information. Further details are out of scope.

Notice that elements of a prime field  $GF(p)$ , where  $p$  is a prime number with bit-length  $m$  divisible by eight, have a tight representation as an  $(m/8)$ -octet string, but do not have a bit position that is always set to zero. Thus, in this case, one cannot represent a compressed point as an octet string of the same length as an element of  $GF(p)$ . However, one can still encode this as an octet string of length  $(m/8)+1$  (see Note 1 above). If one uses right-concatenation as in Note 1, but (for historical reasons) represents the parity bit  $t$  of the compressed point in question by  $0x00$  or  $0x80$  depending on whether  $t=0$  or  $t=1$ , respectively, this is again called the "squeezed" representation (despite this being somewhat a misnomer, since each point is now represented as an octet string that is one octet longer than the tight representation of elements of  $GF(p)$ ). Notice that this representation corresponds to the compressed point representation of Appendix I.8 as octet string, but with the bit-ordering in the 1-octet representation of  $t$  reversed. (Note that this puts the parity bit  $t$  in the leftmost bit position of the octet string if one follows the MSB/msb representation conventions.) We will use this squeezed representation in Appendix O.

#### Appendix J. Representation Examples Curve25519 Family Members

We present some examples of computations using the curves introduced in Appendix E and Appendix G of this document. In each case, we indicate the values of  $P$ ,  $k \cdot P$ , and  $(k+1) \cdot P$ , where  $P$  is a fixed multiple (here: 2019) of the base point of the curve in question and where the private key  $k$  is the integer

```
k  45467544759954639344191351164156560595299236761702065033670739677
    691372543056

    (=0x6485b7e6 cd83e5c2 0d5dbfe4 f915494d 9cf5c65d 778c32c3
    c08d5abd 15e29c50).
```

In the examples below, each curve point is represented using the "squeezed" point representation (see Appendix I.8), whereby each point is represented as a 32-octet string, where the ordering convention (see Appendix I.7) depends on the underlying curve model

in question. Here, points of a Weierstrass curve are represented in tight MSB/msb-order, points of a Montgomery curve in tight LSB/msb-order, and points of a twisted Edwards curve in tight LSB/lwb-order. For points that are a public key, the corresponding private keys are represented as 32-octet strings, using the same (tight) ordering conventions as with the public keys. For affine points, we also give the tight representation of each of its coordinates, using the same ordering conventions as used with the squeezed point representation. For further details, see the examples themselves.

#### J.1. Example with Curve25519

$P_m = (u, v)$ ,  $k * P_m = (u_1, v_1)$ , and  $(k+1) * P_m = (u_2, v_2)$  with Curve25519:

```

u  53025657538808013645618620393754461319535915376830819974982289332
    088255623750

    (=0x753b7566 df35d574 4734142c 9abf931c ea290160 aa75853c
    7f972467 b7f13246).

v  53327798092436462013048370302019946300826511459161905709144645521
    233690313086

    (=0x75e676ce deee3b3c 12942357 22f1d884 ac06de07 330fb07b
    ae35ca26 df75417e).

u1 42039618818474335439333192910143029294450651736166602435248528442
    691717668056

    (=0x5cf194be f0bdd6d6 be58e18a 8f16740a ec25f4b0 67f7980a
    23bb6468 88bb9cd8).

v1 76981661982917351630937517222412729130882368858134322156485762195
    67913357634

    (=0x110501f6 1dff511e d6c4e9b9 bfd5acbe 8bf043b8 c3e381dd
    f5771306 479ad142).

u2 34175116482377882355440137752573651838273760818624557524643126101
    82464621878

    (=0x078e3e38 41c3e0d0 373e5454 ecffae33 2798b10a 55c72117
    62629f97 f1394d36).

v2 43046985853631671610553834968785204191967171967937842531656254539
    962663994648

```

```
(=0x5f2bbb06 f7ec5953 2c2a1a62 21124585 1d2682e0 cc37307e
fbc17f7f 7fda8518).
```

As suggested in Appendix C.2, the  $v$ -coordinate of  $k \cdot P_m$  can be indirectly computed from the  $u$ -coordinates of  $P_m$ ,  $k \cdot P_m$ , and  $(k+1) \cdot P_m$ , and the  $v$ -coordinate of  $P_m$ , which allows computation of the entire point  $k \cdot P_m$  (and not just its  $u$ -coordinate) if  $k \cdot P_m$  is computed using the Montgomery ladder (as, e.g., [RFC7748] recommends), since that algorithm computes both  $u_1$  and  $u_2$  and the  $v$ -coordinate of the point  $P_m$  may be available from context.

The representation of  $k$  and the compressed representations of  $P_m$  and  $k \cdot P_m$  in tight LSB/msb-order are given by

```
repr(k)      0x509ce215 bd5a8dc0 c3328c77 5dc6f59c 4d4915f9 e4bf5d0d
              c2e583cd e6b78564

repr(Pm)     0x4632f1b7 6724977f 3c8575aa 600129ea 1c93bf9a 2c143447
              74d535df 66753b75;

repr(k*Pm)   0xd89cbb88 6864bb23 0a98f767 b0f425ec 0a74168f 8ae158be
              d6d6bdf0 be94f15c,
```

where the leftmost bit of the rightmost octet indicates the parity of the  $v$ -coordinate of the point of Curve25519 in question (which, in this case, are both zero, since  $v$  and  $v_1$  are even). See Appendix H.2 and Appendix I for further detail on (squeezed) point compression.

The scalar representation and (squeezed) point representation illustrated above are consistent with the representations specified in [RFC7748], except that in [RFC7748] only an affine point's  $u$ -coordinate is represented (i.e., the  $v$ -coordinate of any point is always implicitly assumed to have an even value) and that the representation of the point at infinity is not specified. Another difference is that [RFC7748] allows non-unique representations of some elements of  $GF(p)$ , whereas our representation conventions do not (since tight).

A randomized representation ( $t_1$ ,  $t_2$ ) of the point  $k \cdot P_m$  in tight LSB/msb order is given by

```
t1          409531317901122685707535715924445398426503483189854716584
              37762538294289253464

              (=0x5844b232 8c4586dc 62f593c5 599c2a8c e61ba893 bb052de6
              77510a42 b3a68a5a)
```

t2 451856098332889407421278004628150814449259902023388533929  
08848927625430980881

(=0x11598452 e65138dc ce948d7e d8f46a18 b640722c 8e170957  
751b7729 1b26e663),

where this representation is defined in Appendix K.5 and uses the mapping of Appendix K.3.2 with the default square root function.

This representation can also be expressed in tight LSB/msb order as the pair ((u1,s1),(u2,s2)), where (s1,s2):=(0,0) and where

u1 545187339829846945538068364048581821018455714632595988990  
2000416117254237099

(=0xabab17e4 f1dbafb1 ede0c4b3 bedb7734 9c85f2a7 917c5edf  
ad4bd96a a7a60d0c)

u2 236263468848031270223854046645772980064576816578949344957  
7618817248044779847

(=0x47099c3e 9b5cc8fe eaac5db0 6fb413fa b3ef4516 7bfc4dc4b  
8368f22e 2f343905),

where this uses the default completed mapping defined in Appendix K.6 and the mapping of Appendix K.4.2 (with the default square root function).

## J.2. Example with Edwards25519

$Pe=(x, y)$ ,  $k*Pe=(x1, y1)$ , and  $(k+1)*Pe=(x2, y2)$  with Edwards25519:

x 25301662348702136092602268236183361085863932475593120475382959053  
365387223252

(=0x37f03bc0 1070ed12 d3218f8b ba1abb74 fd6b94eb 62033d09  
83851e21 d6a460d4).

y 54434749145175762798550436656748568411099702168121592090608501578  
942019473360

(=0x7858f9e7 6774ed8e 23d614d2 36715fc7 56813b02 9aa13c18  
960705c5 b3a30fd0).

x1 42966967796585460733861724865699548279978730460766025087444502812  
416557284873

(=0x5efe7124 465b5bdb b364bb3e e4f106e2 18d59b36 48f4fe83  
c11afc91 785d7e09).

y1 46006463385134057167371782068441558951541960707376246310705917936  
352255317084

(=0x65b6bc49 985badaf bc5fdd96 fb189502 35d5effd 540b439d  
60508827 80bc945c).

x2 42629294840915692510487991904657367226900127896202625319538173473  
104931719808

(=0x5e3f536a 3be2364a 1fa775a3 5f8f65ae 93f4a89d 81a04a2e  
87783748 00120a80).

y2 29739282897206659585364020239089516293417836047563355347155817358  
737209129078

(=0x41bfd66e 64bdd801 c581a720 f48172a8 187445fa 350924a2  
c92c791e 38d57876).

The representation of  $k$  and the compressed representations of  $P_e$  and  $k*P_e$  in tight LSB/lsb-order are given by

repr( $k$ )       =0x0a3947a8 bd5ab103 c34c31ee ba63af39 b292a89f 27fdbab0  
43a7c1b3 67eda126;

repr( $P_e$ )       =0x0bf0c5cd a3a0e069 183c8559 40dc816a e3fa8e6c 4b286bc4  
71b72ee6 e79flale;

repr( $k*P_e$ )    =0x3a293d01 e4110a06 b9c2d02a bff7abac 40a918df 69bbfa3d  
f5b5da19 923d6da7,

where the rightmost bit of the rightmost octet indicates the parity of the  $x$ -coordinate of the point of Edwards25519 in question (which, in this case, are zero and one, respectively, since  $x$  is even and  $x_1$  is odd). See Appendix H.3 and Appendix I for further detail on (squeezed) point compression.

The scalar representation and (squeezed) point representation illustrated above are fully consistent with the representations specified in [RFC8032]. Note that, contrary to [RFC7748], [RFC8032] requires unique representations of all elements of  $GF(p)$ .

A randomized representation ( $t_1, t_2$ ) of the point  $k*P_e$  in tight LSB/lsb order is given by

```

t1      577913017083163641949634219017190182170288776648725395935
        97750427519399254040

        (=0x181a32c5 10e06dbc ea321882 f3519055 535e289e 8faac654
        82e26f61 aded23fe)

t2      454881407940919718426608573125377401686255068210624245884
        05479716220480287974

        (=0x672e36c5 ae353073 cdfac343 e8297b05 1b010d0f 5b1016db
        dd4baf54 28068926),

```

where this representation is defined in Appendix K.5 and uses the mapping of Appendix K.3.3 with the default square root function and underlying isomorphic mapping between Edwards25519 and Curve25519 of Appendix E.2.

This representation can also be expressed in tight LSB/lsb order as the pair  $((u_1, s_1), (u_2, s_2))$ , where  $(s_1, s_2) := (0, 1)$  and where

```

u1      224462652213914013165861386626523724285418072774741333590
        46191305234585192644

        (=0x2311ee45 c788a81b 7fcd7ae1 c6982d7b 537011fd d49e2eb4
        62b9c08c 5344058c)

u2      103951215490226901552766901992808623194604650181530822362
        9026508474142603215

        (=0xf3ed475b fd95335c 3a0ceb7e 319f8d3c cc651d5b 17eb4439
        e3b25693 0bea3240),

```

where this uses the default completed mapping defined in Appendix K.6 and the mapping of Appendix K.4.3 (with the default square root function).

### J.3. Example with Wei25519

$Pw = (X, Y)$ ,  $k * Pw = (X_1, Y_1)$ , and  $(k+1) * Pw = (X_2, Y_2)$  with Wei25519:

```

X      14428294459702615171094958724191825368445920488283965295163094662
        783879239338

        (=0x1fe62011 89e0801e f1debed7 456a3dc7 94d3ac0b 55202fe7
        2a41cf12 629e56aa).

Y      53327798092436462013048370302019946300826511459161905709144645521
        233690313086

```

(=0x75e676ce deee3b3c 12942357 22f1d884 ac06de07 330fb07b  
ae35ca26 df75417e).

X1 34422557393689369648095312405803933433606568476197477554293337733  
87341283644

(=0x079c3f69 9b688181 69038c35 39c11eb5 96d09f5b 12a242b4  
ce660f13 3368c13c).

Y1 76981661982917351630937517222412729130882368858134322156485762195  
67913357634

(=0x110501f6 1dff511e d6c4e9b9 bfd5acbe 8bf043b8 c3e381dd  
f5771306 479ad142).

X2 22716193187790487472805844610038683159372373526135883092373909944  
834653057415

(=0x3238e8e2 ec6e8b7a e1e8feff 97aa58dd d2435bb5 0071cbc2  
0d0d4a42 9be67187).

Y2 43046985853631671610553834968785204191967171967937842531656254539  
962663994648

(=0x5f2bbb06 f7ec5953 2c2a1a62 21124585 1d2682e0 cc37307e  
fbc17f7f 7fda8518).

The representation of  $k$  and the compressed representations of  $P_w$  and  $k \cdot P_w$  in tight MSB/msb-order are given by

repr( $k$ ) =0x6485b7e6 cd83e5c2 0d5dbfe4 f915494d 9cf5c65d 778c32c3  
c08d5abd 15e29c50;

repr( $P_w$ ) =0x1fe62011 89e0801e f1debed7 456a3dc7 94d3ac0b 55202fe7  
2a41cf12 629e56aa;

repr( $k \cdot P_w$ ) =0x079c3f69 9b688181 69038c35 39c11eb5 96d09f5b 12a242b4  
ce660f13 3368c13c,

where the leftmost bit of the leftmost octet indicates the parity of the  $Y$ -coordinate of the point of Wei25519 in question (which, in this case, are both zero, since  $Y$  and  $Y_1$  are even). See Appendix H.1 and Appendix I for further detail on (squeezed) point compression.

The scalar representation is consistent with the representations specified in [SEC1]; the (squeezed) point representation illustrated above is "new". For completeness, we include a SEC1-consistent



representation of the point  $P_w$  in affine format and in compressed format below.

The SECl-compliant affine representation of the point  $P_w$  in tight MSB/msb-order is given by

```
aff(Pw)      =0x04 1fe62011 89e0801e fldebed7 456a3dc7 94d3ac0b
              55202fe7 2a41cf12 629e56aa

              75e676ce deee3b3c 12942357 22fld884 ac06de07 330fb07b
              ae35ca26 df75417e,
```

whereas the SECl-compliant compressed representation of the point  $P_w$  in tight MSB/msb-order is given by

```
compr(Pw)    =0x02 1fe62011 89e0801e fldebed7 456a3dc7 94d3ac0b
              55202fe7 2a41cf12 629e56aa;
```

The SECl-compliant uncompressed format  $\text{aff}(P_w)$  of an affine point  $P_w$  corresponds to the right-concatenation of its X- and Y-coordinates, each in tight MSB/msb-order, prepended by the string 0x04, where the reverse procedure is uniquely defined, since elements of  $\text{GF}(p)$  have a unique fixed-size representation. The (squeezed) compressed format  $\text{repr}(P_w)$  corresponds to the SECl-compliant compressed format by extracting the parity bit  $t$  from the leftmost bit of the leftmost octet of  $\text{repr}(P_w)$ , replacing the bit position by the value zero, and prepending the octet string with 0x02 or 0x03, depending on whether  $t=0$  or  $t=1$ , respectively, where the reverse procedure is uniquely defined, since  $\text{GF}(p)$  is a 255-bit prime field. For further details, see [SECl]. Note that, due to the bit-length of the prime  $p$ , the squeezed compressed format  $\text{repr}(P_w)$  is one octet shorter than the SECl-compliant compressed format  $\text{compr}(P_w)$ .

A randomized representation  $(t_1, t_2)$  of the point  $k \cdot P_w$  in tight MSB/msb order is given by

```
t1           446363445988889734093446280484122107283059206243307955388
              84223152228795899590

              (=0x62af4697 4dd469ac 96c64809 c16c8517 b6a0cee5 40ba0e2e
              6dd2b36a fcc75ec6)

t2           213890166610228613105792710708385961712211281744756216061
              11930888059603107561

              (=0x2f49c121 8fed7912 031157ee ae066507 a972320b 6180e267
              4025b006 2e67bee9),
```

where this representation is defined in Appendix K.5 and uses the mapping of Appendix K.3.1 with the default square root function.

This representation can also be expressed in tight MSB/msb order as the pair  $((u_1, s_1), (u_2, s_2))$ , where  $(s_1, s_2) := (1, 0)$  and where

```
u1      520092833970966289810117689157951302936446424265230088162
        65117106436465991934

        (=0x72fc3612 b18d2644 c2a85b3b dd66cd58 07ebf07b 2131b77d
        6d7579da 5efba0fe)

u2      134005949856425653115405838878115551263976839535650697250
        78991786686428785368

        (=0x1da077cd 6fa87515 731029a8 bd88da6a 34e38b83 51191edf
        8a3b92d7 ba24aad8),
```

where this uses the default completed mapping defined in Appendix K.6 and the mapping of Appendix K.4.1 (with the default square root function).

#### J.4. Example with Wei25519.2

$Pw2 = (X, Y)$ ,  $k * Pw2 = (X1, Y1)$ , and  $(k+1) * Pw2 = (X2, Y2)$  with Wei25519.2:

```
X      17830493209951148331008014701079988862634531394137235438571836389
        227198459763

        (=0x276bb396 d766b695 bfe60ab1 3c0260dd c09f5bcf 7b3ca47c
        f21c8672 d1ecaf73).

Y      21064492012933896105338241940477778461866060481408222122979836206
        137075789640

        (=0x2e921479 5ad47af7 784831de 572ed8e9 7e20e137 cc67378c
        184ca19f f9136f48).

X1     65470988951686461979789632362377759464688342154017353834939203791
        39281908968

        (=0x0e7986d2 e94354ab 8abd8806 3154536a 4dcf8e6e 65557183
        e242192d 3b87f4e8).

Y1     51489590494292183562535790579480033229043271539297275888817125227
        35262330110
```

(=0x0b623521 c1ff84bc 1522ff26 3376796d be77fcad 1fcabc28  
98f1be85 d7576cfe).

X2 83741788501517200942826153677682120998854086551751663061374935388  
3494226693

(=0x01d9f633 b2ac2606 9e6e93f7 6917446c 2b27c16f 729121d7  
709c0a58 00ef9b05).

Y2 42567334190622848157611574766896093933050043101247319937794684825  
168161540336

(=0x5e1c41e1 fb74e41b 3a19ce50 e1b2caf7 7cabcb3 0c1c1474  
a4fd13e6 6c4c08f0).

The representation of  $k$  and the compressed representations of  $Pw2$  and  $k*Pw2$  in tight MSB/msb-order are given by

repr( $k$ ) =0x6485b7e6 cd83e5c2 0d5dbfe4 f915494d 9cf5c65d 778c32c3  
c08d5abd 15e29c50;

repr( $Pw2$ ) =0x276bb396 d766b695 bfe60ab1 3c0260dd c09f5bcf 7b3ca47c  
f21c8672 d1ecaf73;

repr( $k*Pw2$ ) =0x0e7986d2 e94354ab 8abd8806 3154536a 4dcf8e6e 65557183  
e242192d 3b87f4e8,

where the leftmost bit of the leftmost octet indicates the parity of the Y-coordinate of the point of Wei25519.2 in question (which, in this case, are both zero, since  $Y$  and  $Y1$  are even). See Appendix Appendix H.1 and Appendix I for further detail on (squeezed) point compression.

A randomized representation ( $t1$ ,  $t2$ ) of the point  $k*Pw2$  in tight MSB/msb order is given by

$t1$  416669672354928148679758598803660112405431159793278161879  
36189858804289581274

(=0x5cleaaef 80f9d4af 33c119fc c99acd58 f81e7d69 999c7048  
e4043a77 87a930da)

$t2$  361115271162391608083096560179337391059615651279123199921  
18531180247832114098

(=0x4fd66668 e7174775 de44c852 92df8cfe b9832ef8 2570b3b8  
fe5ec21a b2d4b3b2),

where this representation is defined in Appendix K.5 and uses the mapping of Appendix K.3.1 with the default square root function.

This representation can also be expressed in tight MSB/msb order as the pair  $((u_1, s_1), (u_2, s_2))$ , where  $(s_1, s_2) := (1, 0)$  and where

```
u1      138215499313862453472915174740765454800858627563772726738
        62176256261157017834

        (=0x1e8eb854 2ce139f7 fdbf2059 ac257c89 d7e2e5fe 9c4b97e6
        7656d42c 590bd8ea)

u2      528750192685398685104289021251049791405104665681275304080
        7706116783659458600

        (=0x0bb09eba b0470a84 0ce1ba90 0aeab208 7e8d4760 1309d7af
        e3712e1f 2232a028),
```

where this uses the default completed mapping defined in Appendix K.6 and the mapping of Appendix K.4.1 (with the default square root function).

#### J.5. Example with Wei25519.-3

$Pw3 = (X, Y)$ ,  $k * Pw3 = (X1, Y1)$ , and  $(k+1) * Pw3 = (X2, Y2)$  with Wei25519.-3:

```
X      14780197759513083469009623947734627174363231692126610860256057394
        455099634096

        (=0x20ad4ba4 612f0586 221787b0 d01ba46c d1d8cd5a 0348ef00
        eb4c9272 03ca71b0).

Y      45596733430378470319805536538617129933663237960146030424392249401
        952949482817

        (=0x64ced628 e982648e 4bfcf30c 71c4d267 ba48b0ce fee20062
        b43ef4c9 73f7b541).

X1     47362979975244556396292400751828272600887612546997532158738958926
        60745725532

        (=0x0a78a650 a39995ef dcf4de88 940d4ce9 5b2ca35c c5d70e06
        63b8455e 2e04e65c).

Y1     30318112837157047703426636957515037640997356617656007157255559136
        153389790354
```

(=0x4307719a 20d08741 58d5889e 8c8ec27e 246b0342 55f8fd62  
dbc9ca09 e79c7492)).

X2 23778942085873786433506063022059853212880296499622328201295446580  
293591664363

(=0x3492677e 6ae9d1c3 e08f908b 61033f3d 4e8322c9 fba6da81  
2c95b067 9b1486eb)).

Y2 44846366394651736248316749170687053272682847823018287439056537991  
969511150494

(=0x632624d4 ab94c83a 796511c0 5f5412a3 876e56d2 ed18eca3  
21b95bef 7bf9939e)).

The representation of  $k$  and the compressed representations of  $Pw3$  and  $k*Pw3$  in tight MSB/msb-order are given by

repr( $k$ ) =0x6485b7e6 cd83e5c2 0d5dbfe4 f915494d 9cf5c65d 778c32c3  
c08d5abd 15e29c50;

repr( $Pw3$ ) =0xa0ad4ba4 612f0586 221787b0 d01ba46c d1d8cd5a 0348ef00  
eb4c9272 03ca71b0;

repr( $k*Pw3$ ) =0x0a78a650 a39995ef dcf4de88 940d4ce9 5b2ca35c c5d70e06  
63b8455e 2e04e65c,

where the leftmost bit of the leftmost octet indicates the parity of the Y-coordinate of the point of Wei25519.-3 in question (which, in this case, are one and zero, respectively, since  $Y$  is odd and  $Y1$  is even). See Appendix H.1 and Appendix I for further detail on (squeezed) point compression.

A randomized representation ( $t1$ ,  $t2$ ) of the point  $k*Pw3$  in tight MSB/msb order is given by

$t1$  573714937613596601525680684642155667097217474964816246889  
88981227297409008259

(=0x7ed71d5f 566d2259 99bdb404 bfb9d6cf d2e86ccb 1894d4a6  
c75e3c69 e5eb0283)

$t2$  269945781324580189815142015663892935722419453863927287235  
57891665397640090729

(=0x3bae63c8 70f60de0 c2e35f94 d24220f1 bb6efd00 37625869  
f84923de ff4c5469),

where this representation is defined in Appendix K.5 and uses the mapping of Appendix K.3.1 with the default square root function.

This representation can also be expressed in tight MSB/msb order as the pair  $((u_1, s_1), (u_2, s_2))$ , where  $(s_1, s_2) := (1, 1)$  and where

```
u1      273592510979600674027837477146355037032732195078153389134
        81162438438522584713

        (=0x3c7cc990 81eed784 9ca746d7 c479a902 ce9de65f 1150e7b9
        c87d08d2 9785fe89)

u2      271488765024747755704729103260177059745349171282146823458
        00069381584030663589

        (=0x3c05b835 1283fca7 705eba74 1e6b853e db3ed5dc d1891daa
        c1643d8d d63a03a5),
```

where this uses the default completed mapping defined in Appendix K.6 and the mapping of Appendix K.4.1 (with the default square root function).

## Appendix K. Auxiliary Functions

This section illustrates how one could implement common routines, such as taking square roots and inverses in finite fields, and how to map field elements to curve points and to curve points that avoid outlier points in the small subgroup.

### K.1. Square Roots in $\text{GF}(q)$

Square roots are easy to compute in  $\text{GF}(q)$  if  $q = 3 \pmod{4}$  (see Appendix K.1.1) or if  $q = 5 \pmod{8}$  (see Appendix K.1.2). Details on how to compute square roots for other values of  $q$  are out of scope. If square roots are easy to compute in  $\text{GF}(q)$ , then so are these in  $\text{GF}(q^2)$ .

NOTE: If one wishes to check whether an element is a square in  $\text{GF}(q)$ , rather than actually compute square roots, more efficient methods can be used. As an example, if  $\text{GF}(q)$  is a prime field (i.e.,  $q=p$ ), one can efficiently check whether an element  $y$  is a square in  $\text{GF}(p)$  by computing its Legendre symbol  $(y/p)$  (see Section 2.4.5 of [Handbook]). Details on how to efficiently check whether an element is a square in  $\text{GF}(q)$  for other values of  $q$  are out of scope. If checking whether an element is a square is easy in  $\text{GF}(q)$ , then so it is in  $\text{GF}(q^2)$ .

K.1.1. Square Roots in  $GF(q)$ , where  $q = 3 \pmod{4}$ 

If  $y$  is a nonzero element of  $GF(q)$  and  $z := y^{\{(q-3)/4\}}$ , then  $y$  is a square in  $GF(q)$  only if  $y*z^2=1$ .

- a. If  $y*z^2=+1$ ,  $z$  is a square root of  $1/y$  and  $y*z$  is a square root of  $y$  in  $GF(q)$ ;
- b. If  $y*z^2=-1$ ,  $z$  is a square root of  $-1/y$  and  $y*z$  is a square root of  $-y$  in  $GF(q)$ .

(Note that the field element  $-1$  is a non-square in  $GF(q)$ .)

K.1.2. Square Roots in  $GF(q)$ , where  $q = 5 \pmod{8}$ 

If  $y$  is a nonzero element of  $GF(q)$  and  $z := y^{\{(q-5)/8\}}$ , then  $y$  is a square in  $GF(q)$  only if  $y^2*z^4=1$ .

- a. If  $y*z^2=+1$ ,  $z$  is a square root of  $1/y$  and  $y*z$  is a square root of  $y$  in  $GF(q)$ ;
- b. If  $y*z^2=-1$ ,  $i*z$  is a square root of  $1/y$  and  $i*y*z$  is a square root of  $y$  in  $GF(q)$ ;
- c. If  $y*z^2=+i$ ,  $z$  is a square root of  $i/y$  and  $y*z$  is a square root of  $i*y$  in  $GF(q)$ ;
- d. If  $y*z^2=-i$ ,  $z$  is a square root of  $-i/y$  and  $i*y*z$  is a square root of  $i*y$  in  $GF(q)$ .

Here,  $i$  is an element of  $GF(q)$  for which  $i^2=-1$  (e.g.,  $i:=2^{\{(q-1)/4\}}$ ). This field element is a non-square in  $GF(q)$  and can be precomputed.

## K.2. Inversion

If  $y$  is an integer and  $\gcd(y,n)=1$ , one can efficiently compute  $1/y \pmod{n}$  via the extended Euclidean Algorithm (see Section 2.2.5 of [GECC]). One can use this algorithm as well to compute the inverse of a nonzero element  $y$  of a prime field  $GF(p)$ , since  $\gcd(y,p)=1$ .

The inverse of a nonzero element  $y$  of  $GF(q)$  can be computed as

$$1/y := y^{\{q-2\}} \text{ (since } y^{\{q-1\}}=1 \text{ by Fermat's Little Theorem).}$$

If inverses are easy to compute in  $GF(q)$ , then so are these in  $GF(q^2)$ . Further details are out of scope.

The inverses of two nonzero elements  $y_1$  and  $y_2$  of  $\text{GF}(q)$  can be computed by first computing the inverse  $z$  of  $y_1*y_2$  and by subsequently computing  $y_2*z:=1/y_1$  and  $y_1*z:=1/y_2$ .

NOTE 1: This method can be used to compute the inverse of a nonzero element  $y$  of  $\text{GF}(q)$  indirectly, as  $\text{lambda}*(\text{lambda}*y)^{-1}$ , where  $\text{lambda}$  is a random nonzero element of  $\text{GF}(q)$ . This yields an inversion routine (commonly called "blinded inversion") where the inversion operation itself does not leak information on  $y$ .

NOTE 2: This method can also be used to compute the inverse and a square root, respectively, of two nonzero elements  $x$  and  $y$  of  $\text{GF}(q)$  (where  $y$  is a square in  $\text{GF}(q)$ ) by first computing a square root  $z$  of  $1/(y*x^2)$  (see Appendix K.1) and by subsequently computing a square root of  $y$  as  $x*y*z$  and the inverse of  $x$  as  $x*y*z^2$ .

### K.3. Mappings to Curve Points

One can map elements of  $\text{GF}(q)$  that are not a square in  $\text{GF}(q)$  to points of a Weierstrass curve (see Appendix K.3.1), to points of a Montgomery curve (see Appendix K.3.2), or to points of a twisted Edwards curve (see Appendix K.3.3), under some mild conditions on the domain parameters. Full details on mappings that apply if these conditions are not satisfied are out of scope.

#### K.3.1. Mapping to Points of Weierstrass Curve

The description below assumes that the domain parameters  $a$  and  $b$  of the Weierstrass curve  $W_{\{a,b\}}$  are nonzero. For ease of exposition, we define  $f(z):=z^3+a*z+b$ . (Note that for an affine point  $(X,Y)$  of  $W_{\{a,b\}}$  one has  $Y^2=f(X)$ .)

If  $t$  is an element of  $\text{GF}(q)$  that is not a square in  $\text{GF}(q)$  and that is unequal to  $-1$ , then the element  $X:=(-b/a)*(1+1/(t+t^2))$  is the unique solution of the equation  $f(t*X)=t^3*f(X)$  and is nonzero. Consequently, either  $X$  or  $X':=t*X$  is the  $x$ -coordinate of an affine point of  $W_{\{a,b\}}$ , depending on whether  $f(X)$  is a square in  $\text{GF}(q)$ .

- a. If  $f(X)$  is a square in  $\text{GF}(q)$  and  $Y:=\text{sqrt}(f(X))$ , then  $t$  is mapped to the point  $P(t):=(X, Y)$ ;
- b. If  $f(X)$  is not a square in  $\text{GF}(q)$  and  $Y':=\text{sqrt}(f(X'))$ , then  $t$  is mapped to the point  $P(t):=(X', -Y')$ .

Formally, this mapping is not properly defined, since a nonzero square  $y:=x^2$  in  $\text{GF}(q)$  has two solutions, viz.  $x$  and  $-x$ ; it is properly defined, however, if one designates for each element of  $\text{GF}(q)$  that is a square in  $\text{GF}(q)$  precisely one square root as "the"



square root of this element. Note that always picking the square root with zero parity (see Appendix H) satisfies this condition (henceforth called the default square root function).

If  $-1$  is not a square in  $\text{GF}(q)$ , this element is mapped to the point at infinity  $O$  of  $W_{\{a,b\}}$ .

The set of points of  $W_{\{a,b\}}$  that arises this way has size roughly  $3/8$  of the order of the curve and each such point arises as image of one or two  $t$  values. Further details are out of scope.

NOTE 1: If  $-1$  is not a square in  $\text{GF}(q)$ , the mapping above yields the point at infinity for  $t=-1$ . One can modify this mapping, by mapping the element  $-1$  to any suitable point  $P_0$  of  $W_{\{a,b\}}$  (e.g., its base point  $G$  or any other affine point) and leaving the remainder of the mapping the same. Suitability of such a modification is application-specific. Details are out of scope.

NOTE 2: The description above assumes that the domain parameters  $a$  and  $b$  of the Weierstrass curve  $W_{\{a,b\}}$  are nonzero. If this is not the case, one can often find an isogenous curve  $W_{\{a',b'\}}$  for which the domain parameters  $a'$  and  $b'$  are nonzero. If so, one can map elements of  $\text{GF}(q)$  that are not a square in  $\text{GF}(q)$  to points of  $W_{\{a,b\}}$  via function composition, where one uses the mapping above to arrive at a point of  $W_{\{a',b'\}}$  and where one subsequently uses the dual isogeny from  $W_{\{a',b'\}}$  to  $W_{\{a,b\}}$  to arrive at a point of  $W_{\{a,b\}}$ . As an example, one can show that if  $a$  is zero and if  $-4*b$  is a cube in  $\text{GF}(q)$  (such as is the case with, e.g., the "BitCoin" curve secp256k1 [SEC2]), this curve is 3-isogenous to a curve with this property and the strategy above applies (for an example with secp256k1, see Appendix L). Further details are out of scope.

### K.3.2. Mapping to Points of Montgomery Curve

The description below assumes that the domain parameter  $A$  of the Montgomery curve  $M_{\{A,B\}}$  is nonzero. For ease of exposition, we define  $f(z) := z^3 + A*z^2 + z$ . (Note that for an affine point  $(u,v)$  of  $M_{\{A,B\}}$  one has  $B*v^2 = f(u)$ .)

If  $t$  is an element of  $\text{GF}(q)$  that is not a square in  $\text{GF}(q)$  and that is unequal to  $-1$ , then the element  $u := -(1+1/t)/A$  is the unique nonzero solution of the equation  $f(t*u) = t^3*f(u)$ . Consequently, either  $u$  or  $u' := t*u$  is the  $u$ -coordinate of an affine point of  $M_{\{A,B\}}$ , depending on whether  $f(u)/B$  is a square in  $\text{GF}(q)$ .

- a. If  $f(u)/B$  is a square in  $\text{GF}(q)$  and  $v := \text{sqrt}(f(u)/B)$ , then  $t$  is mapped to the point  $P(t) := (u, v)$ ;

- b. If  $f(u)/B$  is not a square in  $GF(q)$  and  $v' := \sqrt{f(u')/B}$ , then  $t$  is mapped to the point  $P(t) := (u', -v')$ .

As before, formally, this mapping is not properly defined, since a nonzero square  $y = x^2$  in  $GF(q)$  has two solutions, viz.  $x$  and  $-x$ ; it is properly defined, however, if one designates for each element of  $GF(q)$  that is a square in  $GF(q)$  precisely one square root as "the" square root of this element. Note that always picking the square root with zero parity (see Appendix H) satisfies this condition (henceforth called the default square root function).

If  $-1$  is not a square in  $GF(q)$ , this element is mapped to the point at infinity  $O$  of  $M_{\{A,B\}}$ .

The set of points of  $M_{\{A,B\}}$  that arises this way has size roughly  $1/2$  of the order of the curve and each such point arises as image of precisely one  $t$  value. Further details are out of scope.

NOTE 1: If  $-1$  is not a square in  $GF(q)$ , the mapping above yields the point at infinity for  $t = -1$ . One can modify this mapping, by mapping the element  $-1$  to any suitable point  $P_0$  of  $M_{\{a,b\}}$  (e.g., its base point  $G$  or any other affine point) and leaving the remainder of the mapping the same. Suitability of such a modification is application-specific. Details are out of scope.

NOTE 2: The description above assumes that the domain parameter  $A$  of the Montgomery curve  $M_{\{A,B\}}$  is nonzero. If this is not the case, the curve is a Weierstrass curve for which the domain parameter  $b$  is zero and Note 2 of Appendix K.3.1 applies. If  $q \equiv 3 \pmod{4}$ , an even simpler approach is possible, where one modifies the construction above and simply takes  $u := t$  and  $u' := -t$  (which works, since  $-1$  is not a square in  $GF(q)$  and  $f(-t) = -f(t)$ ). In this case, this construction can be extended to all elements  $t$  of  $GF(q)$  and, if so, yields a 1-1 mapping between  $GF(q)$  and all affine curve points.

### K.3.3. Mapping to Points of Twisted Edwards Curve

One can map elements of  $GF(q)$  that are not a square in  $GF(q)$  to points of the twisted Edwards curve  $E_{\{a,d\}}$  via function composition, where one uses the mapping of Appendix K.3.1 to arrive at a point of the Weierstrass curve  $W_{\{a,b\}}$  and where one subsequently uses the isomorphic mapping between twisted Edwards curves and Weierstrass curves of Appendix D.3 to arrive at a point of  $E_{\{a,d\}}$ . Another mapping is obtained by function composition, where one instead uses the mapping of Appendix K.3.2 to arrive at a point of the Montgomery curve  $M_{\{A,B\}}$  and where one subsequently uses the isomorphic mapping between twisted Edwards curves and Montgomery curves of Appendix D.1 to arrive at a point of  $E_{\{a,d\}}$ . Obviously, one can use function

composition (now using the respective pre-images - if these exist) to realize the pre-images of either mapping.

#### K.4. Mappings to High-Order Curve Points

Appendix K.3 described how one can map elements of  $GF(q)$  that are not a square in  $GF(q)$  to points of a Weierstrass curve, to points of a Montgomery curve, or to points of a twisted Edwards curve, under some mild conditions on the domain parameters. Below, we use the mappings of that appendix and the parity function `par(.)` specified in Appendix H to construct mappings to high-order curve points only (i.e., mappings that avoid points in the small subgroup, see Appendix B.1). We consider mappings to high-order points of a Weierstrass curve (see Appendix K.4.1), to high-order points of a Montgomery curve (see Appendix K.4.2), and to high-order points of a twisted Edwards curve (see Appendix K.4.3). As before, full details on mappings that apply if the mild conditions on the domain parameters are not satisfied are out of scope.

##### K.4.1. Mapping to High-Order Points of Weierstrass Curve

The description below assumes that the domain parameters  $a$  and  $b$  of the Weierstrass curve  $W_{\{a,b\}}$  are nonzero. For ease of exposition, we define  $f(z) := z^3 + a*z + b$ . (Note that for an affine point  $(X, Y)$  of  $W_{\{a,b\}}$  one has  $Y^2 = f(X)$ .)

If  $t$  is an element of  $GF(q)$  that is not a square in  $GF(q)$  and that is unequal to  $-1$ , the mapping of Appendix K.3.1 yields an affine point  $P(t) := (X, Y)$  of  $W_{\{a,b\}}$ . Let  $P_0 := (X_0, Y_0)$  be a fixed affine point of  $W_{\{a,b\}}$  for which neither  $P_0$ ,  $P_0 + P(t)$ , nor  $P_0 - P(t)$  is in the small subgroup of  $W_{\{a,b\}}$  (for any non-square element  $t \neq -1$  of  $GF(q)$ ). (Note that this implies that  $P_0$  and  $P(t)$  are distinct affine points of the curve and that these are not each other's inverse.) For binary digit  $s$ , the point  $Q(t, s)$  is now defined as follows:

- a. If  $\text{par}(Y_0 * Y) = s$ , then the pair  $(t, s)$  is mapped to the point  $Q(t, s) := P_0 + P(t)$ ;
- b. If  $\text{par}(Y_0 * Y) \neq s$ , then the pair  $(t, s)$  is mapped to the point  $Q(t, s) := P_0 - P(t)$ .

Note that this mapping is properly defined as long as the fixed point  $P_0$  (the so-called "curve offset") alluded to above indeed exists. In cases of practical interest that we are aware of, this is indeed the case (see, e.g., Table 1).

If  $-1$  is not a square in  $GF(q)$ , the pair  $(-1, s)$  is mapped to the affine point  $P_0$  of  $W_{\{a,b\}}$  (irrespective of the value of  $s$ ).

The set of points of  $W_{\{a,b\}}$  that arises this way has size roughly  $3/8$  of the order of the curve and each such point arises as image of up to four values of the pair  $(t,s)$ . Further details are out of scope.

From the group law for Weierstrass curves (see Appendix C.1) it follows that one can express the coordinates of  $Q(t,s)$ , with  $t \neq -1$ , in terms of the  $X$ -coordinates of  $P_0$  and  $P(t)$  and the product of their  $Y$ -coordinates. (Here, observe that  $Y_0 \cdot Y$  is a square root of  $f(X_0) \cdot f(X)$ .) Thus,  $Q(t,s)$  can be computed without the need to fully compute  $P(t)$ .

#### K.4.2. Mapping to High-Order Points of Montgomery Curve

The description below assumes that the domain parameters  $A$  and  $B$  of the Montgomery curve  $M_{\{A,B\}}$  are nonzero. For ease of exposition, we define  $f(z) := z^3 + A \cdot z^2 + z$ . (Note that for an affine point  $(u,v)$  of  $M_{\{A,B\}}$  one has  $B \cdot v^2 = f(u)$ .)

If  $t$  is an element of  $GF(q)$  that is not a square in  $GF(q)$  and that is unequal to  $-1$ , the mapping of Appendix K.3.2 yields an affine point  $P(t) := (u, v)$  of  $M_{\{A,B\}}$ . Let  $P_0 := (u_0, v_0)$  be a fixed affine point of  $M_{\{A,B\}}$  for which neither  $P_0$ ,  $P_0 + P(t)$ , nor  $P_0 - P(t)$  is in the small subgroup of  $M_{\{A,B\}}$  (for any non-square element  $t \neq -1$  of  $GF(q)$ ). (Note that this implies that  $P_0$  and  $P(t)$  are distinct affine points of the curve and that these are not each other's inverse.) For binary digit  $s$ , the point  $Q(t,s)$  is now defined as follows:

- a. If  $\text{par}(B \cdot v_0 \cdot v) = s$ , then the pair  $(t,s)$  is mapped to the point  $Q(t,s) := P_0 + P(t)$ ;
- b. If  $\text{par}(B \cdot v_0 \cdot v) \neq s$ , then the pair  $(t,s)$  is mapped to the point  $Q(t,s) := P_0 - P(t)$ .

Note that this mapping is properly defined as long as the fixed point  $P_0$  (the so-called "curve offset") alluded to above indeed exists. In cases of practical interest that we are aware of, this is indeed the case (see, e.g., Table 1).

If  $-1$  is not a square in  $GF(q)$ , the pair  $(-1,s)$  is mapped to the affine point  $P_0$  of  $M_{\{A,B\}}$  (irrespective of the value of  $s$ ).

The set of points of  $M_{\{A,B\}}$  that arises this way has size roughly  $1/2$  of the order of the curve and each such point arises as image of up to two values of the pair  $(t,s)$ . Further details are out of scope.

From the group law for Montgomery curves (see Appendix C.2) it follows that one can express the coordinates of  $Q(t,s)$ , with  $t \neq -1$ , in terms of the  $u$ -coordinates of  $P_0$  and  $P(t)$  and the product of their  $v$ -coordinates. (Here, observe that  $B \cdot v_0 \cdot v$  is a square root of  $f(u_0) \cdot f(u)$ .) Thus,  $Q(t,s)$  can be computed without the need to fully compute  $P(t)$ .

Curve	Fixed curve offset $P_0$	Non-Square
NIST P-224 [FIPS-186-4]	Base point $(G_x, G_y)$	11
NIST P-256 [FIPS-186-4]	$P_0 = (0, y)$ , $y$ even	-1
NIST P-384 [FIPS-186-4]	$P_0 = (0, y)$ , $y$ even	-1
NIST P-521 [FIPS-186-4]	$P_0 = (0, y)$ , $y$ even	-1
brainpoolP224r1 [RFC5639]	Base point $(G_x, G_y)$	-1
brainpoolP256r1 [RFC5639]	Base point $(G_x, G_y)$	-1
brainpoolP320r1 [RFC5639]	Base point $(G_x, G_y)$	-1
brainpoolP384r1 [RFC5639]	Base point $(G_x, G_y)$	-1
brainpoolP512r1 [RFC5639]	$P_0 = (3, y)$ , $y$ even	-1
Curve25519 [RFC7748]	$P_0 = (90, v)$ , $v$ even	2
Wei25519 [Appendix E.3]	$P_0 = (3, y)$ , $y$ even	2
Wei25519.2 [Appendix G.3]	$P_0 = (244, y)$ , $y$ even	2
Wei25519.-3 [Appendix G.3]	$P_0 = (41, y)$ , $y$ even	2
Curve448 [RFC7748]	$P_0 = (50, v)$ , $v$ even	-1
Wei448 [Appendix M.3]	$P_0 = (18, y)$ , $y$ even	-1
Wei448.1 [Appendix N.3]	$P_0 = (10, y)$ , $y$ even	-1
Wei448.-3 [Appendix N.3]	$P_0 = (8, y)$ , $y$ even	-1
secp256k1.m [Appendix L.3]	$P_0 = (0, y)$ , $y$ even	-1

Table 1: Fixed curve offsets for mappings that avoid low-order points, for some curves of practical interest, including listing of fixed non-square elements of their underlying finite fields.

#### K.4.3. Mapping to High-Order Points of Twisted Edwards Curve

One can map elements of  $\text{GF}(q)$  that are not a square in  $\text{GF}(q)$  to points of the twisted Edwards curve  $E_{\{a,d\}}$  via function composition, where one uses the mapping of Appendix K.4.1 to arrive at a point of the Weierstrass curve  $W_{\{a,b\}}$  that is not in the small subgroup and where one subsequently uses the isomorphic mapping between twisted Edwards curves and Weierstrass curves of Appendix D.3 to arrive at a point of  $E_{\{a,d\}}$  with this property. Another mapping is obtained by function composition, where one instead uses the mapping of Appendix K.4.2 to arrive at a point of the Montgomery curve  $M_{\{A,B\}}$  that does not have low order and where one subsequently uses the isomorphic mapping between twisted Edwards curves and Montgomery curves of Appendix D.1 to arrive at a point of  $E_{\{a,d\}}$  with this

property. Obviously, one can use function composition (now using the respective pre-images - if these exist) to realize the pre-images of either mapping.

#### K.5. Randomized Representation of Curve Points

The mappings of Appendix K.3 allow one to represent a curve point  $Q$  as a specific element  $t$  of  $\text{GF}(q)$ , provided this point arises as a point in the range of the mapping at hand. For Montgomery curves and twisted Edwards curves, this covers roughly half of the curve points; for Weierstrass curves, roughly  $3/8$  of the curve points. One can extend the mappings above, by mapping a pair  $(t_1, t_2)$  of inputs to the point  $Q := P_2(t_1, t_2) := P(t_1) + P(t_2)$ . In this case, each curve point has roughly  $q/4$  representations as an ordered pair  $(t_1, t_2)$  on average. In fact, one can show that if the input pairs are generated uniformly at random, then the corresponding curve points follow a distribution that is also (statistically indistinguishable from) a uniform distribution, and vice-versa. Here, each pair  $(t_1, t_2)$  deterministically yields a curve point, whereas for each curve point  $Q$ , a randomized algorithm yields an ordered pair  $(t_1, t_2)$  of pre-images of  $Q$ , where the expected number of randomized pre-images one has to try is small (four if one uses the mapping of Appendix K.3.1; two if one uses the mapping of Appendix K.3.2). For further details, see Algorithm 1 of [Tibouchi].

Similar properties hold if one uses the mappings of Appendix K.4 (rather than those of Appendix K.3): in this case, the mapping allows one to represent a curve point  $Q$  as a specific element  $(t, s)$  of  $\text{GF}(q) \times \{0, 1\}$ , provided this point arises as a point in the range of the mapping at hand. For Montgomery curves and twisted Edwards curves, this covers roughly half of the curve points; for Weierstrass curves, roughly  $3/8$  of the curve points. One can extend the mappings above, by mapping a pair  $((t_1, s_1), (t_2, s_2))$  of inputs to the point  $Q := Q_2((t_1, s_1), (t_2, s_2)) := Q(t_1, s_1) - Q(t_2, s_2)$ . In this case, each curve point has roughly  $q$  representations as an ordered pair  $((t_1, s_1), (t_2, s_2))$  on average. In fact, one can show that if the input pairs are generated uniformly at random, then the corresponding curve points follow a distribution that is also (statistically indistinguishable from) a uniform distribution, and vice-versa. Here, each pair  $((t_1, s_1), (t_2, s_2))$  deterministically yields a curve point, whereas for each curve point  $Q$ , a randomized algorithm yields an ordered pair  $((t_1, s_1), (t_2, s_2))$  of pre-images of  $Q$ , where the expected number of randomized pre-images one has to try is small (four if one uses the mapping of Appendix K.4.1; two if one uses the mapping of Appendix K.4.2). Further details are out of scope.

NOTE 1: The main difference between the two constructions above is that the first construction uses the mappings to curve points

described in Appendix K.3, while the second construction uses the mappings to high-order curve points described in Appendix K.4. Note that  $Q2((t1,s1), (t2,s2))$  assumes all values  $(+/-) P(t1) (+/-) P(t2)$  if one considers all possible values for the binary digits  $s1$  and  $s2$ . (This, thereby, includes the value  $P2(t1, t2)$ .)

NOTE 2: The results on the statistical distributions mentioned above still hold in practice if one makes a few localized changes to the constructions. In particular, these are independent of the specific choices for the point  $P0$  (used with input  $-1$  with the mappings of Appendix K.3, if applicable, respectively, used with the mappings of Appendix K.4) and also still hold if one re-defines the mappings  $P2$  or  $Q2$  locally so as to avoid points in the small subgroup.

#### K.6. Completing the Mappings to Curve Points

The mappings of Appendix K.4 operate on input pairs  $(t, s)$ , where  $t$  is an element of  $GF(q)$  that is not a square in  $GF(q)$  and where  $s$  is a binary digit from the set  $\{0,1\}$ . One can use these mappings to produce mappings that operate on input pairs  $(u, s)$ , where  $u$  is any nonzero element of  $GF(q)$ , via function composition, where one first maps the pair  $(u,s)$  to the pair  $(t,s):=(\delta*u^2,s)$ , where  $\delta$  is a fixed element of  $GF(q)$  that is not a square in  $GF(q)$ , and where one subsequently applies any of forementioned mappings to the resulting pair to yield a point of the curve in question. The resulting mapping to high-order curve points can be extended further to one that operates on all elements of  $GF(q)\times\{0,1\}$  by mapping each input  $(u, s)$  with  $u=0$  to any fixed high-order point  $P1$  of the curve in question. The resulting mapping is uniquely defined after fixing the curve offset  $P0$  (used with the mappings of Appendix K.4), the high-order point  $P1$  (used for inputs with  $u=0$  above), and the non-square element  $\delta$  of  $GF(q)$  (used for nonzero inputs  $u$  above).

For the mappings of Appendix K.3, one can use a similar function composition, where one simply drops the binary digit  $s$  and maps  $0$  to the point at infinity or any other suitable curve point  $P1$ . As before, the resulting mapping is uniquely defined after fixing the point  $P0$  (for input  $-1$  with the mappings of Appendix K.3, if applicable), the point  $P1$  (used for input  $u=0$  above), and the non-square element  $\delta$  of  $GF(q)$  (used for nonzero inputs  $u$  above). Further details are out of scope.

Similarly, one can use the completed mappings above to map a pair  $((u1,s1), (u2,s2))$  of elements of  $GF(q)\times\{0,1\}$  to a point of a curve, via function composition, where, in the first case, one first maps the pair  $((u1,s1), (u2,s2))$  to the pair  $((t1,s1), (t2,s2)):=((\delta*u1^2, s1), (\delta*u2^2, s2))$  and subsequently computes  $Q2compl((t1,s1), (t2,s2)):=Qcompl(t1,s1) - Qcompl(t2,s2)$ ,

where  $Q_{\text{compl}}(t,s) := Q(t,s)$  if  $t$  is nonzero and where  $Q_{\text{compl}}(0,s) := P_0$  otherwise (irrespective of the value of  $s$ ). In the second case, one first maps the pair  $(u_1, u_2)$  to the pair  $(t_1, t_2) := (\delta u_1^2, \delta u_2^2)$  and subsequently computes  $P_{2\text{compl}}(t_1, t_2) := P_{\text{compl}}(t_1) + P_{\text{compl}}(t_2)$ , where  $P_{\text{compl}}(t) := P(t)$  if  $t$  is nonzero and where  $P_{\text{compl}}(0) := P_1$  otherwise. In either case, again, the resulting mapping is uniquely defined after fixing the points  $P_0$  and  $P_1$  and the non-square element  $\delta$  of  $\text{GF}(q)$ .

NOTE 1: Each of the above mappings is fully and unambiguously defined by the triple  $(P_0, P_1, \delta)$ . One can locally change this mapping so as to avoid points in the small subgroup, should these otherwise occur, e.g., by setting any such re-defined image to any fixed high-order point  $P_2$  of the curve in question. In this case, the corresponding mapping is uniquely defined by the quadruple  $(P_0, P_1, P_2, \delta)$  and -- in practice -- has the same statistical distribution properties as the original mapping (see NOTE 2 of Appendix K.5). For each curve in Table 1, these completed mappings are uniquely defined by the mentioned fixed curve offset  $P_0$  and non-square element  $\delta$  of  $\text{GF}(q)$ , if one defines  $P_2 := P_1 := P_0$  (henceforth called the default completed mappings).

NOTE 2: For elliptic curves defined over prime fields (i.e.,  $q=p$ ) one can relax the completed mappings above and show that the statistical properties for randomized representations still hold if  $u_1$  is a random element of a sufficiently large interval in  $\text{GF}(p)$  and if  $u_2$  is a random element of a sufficiently large subset of  $\text{GF}(p)$  (see, e.g., [Tibouchi-cleancut]). This allows generating  $u_1$  and  $u_2$ , e.g., each as random bit strings of length  $m-1$ , where  $m$  is the bit-length of  $p$ , thereby allowing the pair  $(u_1, u_2)$  -- a random  $(2*m-2)$ -bit string -- to be used unaltered in this construction, without the need to carry out a reduction modulo  $p$  first. Table 2 illustrates how this can be used to realize randomized representations and completed mappings for each curve in Table 1, where these randomized bit strings have the same byte-length as the (tight) representation of affine curve points. (Here, the field elements  $u_1$  and  $u_2$  are obtained from their bit string representations using the BS2OS mapping of Appendix I.4 and the (non-strict) OS2FE mapping of Appendix I.5.) For each curve in Table 2, we refer to this version of the default completed mapping as being the "clean-cut" default completed mapping.



Curve	left-side	right-side
NIST P-224 [FIPS-186-4]	{u1:224}	{s1:1, s2:1, u2:222}
NIST P-256 [FIPS-186-4]	{s1:1, u1:255}	{s2:1, u2:255}
NIST P-384 [FIPS-186-4]	{u1:384}	{s1:1, s2:1, u2:382}
NIST P-521 [FIPS-186-4]	{s1:1, u1:527}	{s2:1, u2:527}
brainpoolP224r1 [RFC5639]	{s1:1, u1:223}	{s2:1, u2:223}
brainpoolP256r1 [RFC5639]	{s1:1, u1:255}	{s2:1, u2:255}
brainpoolP320r1 [RFC5639]	{s1:1, u1:319}	{s2:1, u2:319}
brainpoolP384r1 [RFC5639]	{s1:1, u1:383}	{s2:1, u2:383}
brainpoolP512r1 [RFC5639]	{s1:1, u1:511}	{s2:1, u2:511}
Curve25519 [RFC7748]	{s1:1, u1:255}	{s2:1, u2:255}
Wei25519 [Appendix E.3]	{s1:1, u1:255}	{s2:1, u2:255}
Wei25519.2 [Appendix G.3]	{s1:1, u1:255}	{s2:1, u2:255}
Wei25519.-3 [Appendix G.3]	{s1:1, u1:255}	{s2:1, u2:255}
Curve448 [RFC7748]	{u1:448}	{s1:1, s2:1, u2:446}
Wei448 [Appendix M.3]	{u1:448}	{s1:1, s2:1, u2:446}
Wei448.1 [Appendix N.3]	{u1:448}	{s1:1, s2:1, u2:446}
Wei448.-3 [Appendix N.3]	{u1:448}	{s1:1, s2:1, u2:446}
secp256k1.m [Appendix L.3]	{u1:256}	{s1:1, s2:1, u2:254}

Table 2: Randomized representation of curve points, for some curves of practical interest, including curve-specific relative ordering and bit-length of substrings representing the tuple  $((u1, s1), (u2, s2))$ , resulting in the bit string left-side || right-side. (Tailored towards avoiding modular reductions in mappings to curve points.)

Table 3 shows an alternative arrangement, tailored towards optimizing the efficiency of computing randomized representations of curve points (see Appendix K.5), rather than towards avoiding modular reductions in the mappings to curve points. (Here, we used randomized representations of elements of  $GF(p)$ , when appropriate, and the bias upper bound  $2^{-64}$  from Table 4.) For each curve in Table 3, we refer to this version of the default completed mapping as being the "point-randomization-optimized" default completed mapping (where both versions coincide if the prime number  $p$  is relatively close to a power of two). (Here, the field elements  $u1$  and  $u2$  are obtained from their bit string representations using the BS2OS

mapping of Appendix I.4 and the (non-strict) OS2FE mapping of Appendix I.5.) Suitability of each of these completed mappings is application-specific (and also depends on the maximum bias one can tolerate). Further details are out of scope of this document.

Curve	left-side	right-side
NIST P-224 [FIPS-186-4]	{u1:224}	{s1:1, s2:1, u2:222}
NIST P-256 [FIPS-186-4]	{u1:288}	{s1:1, s2:1, u2:222}
NIST P-384 [FIPS-186-4]	{u1:384}	{s1:1, s2:1, u2:382}
NIST P-521 [FIPS-186-4]	{s1:1, u1:527}	{s2:1, u2:527}
brainpoolP224r1 [RFC5639]	{s1:1, u1:287}	{s2:1, u2:159}
brainpoolP256r1 [RFC5639]	{s1:1, u1:319}	{s2:1, u2:191}
brainpoolP320r1 [RFC5639]	{s1:1, u1:383}	{s2:1, u2:255}
brainpoolP384r1 [RFC5639]	{s1:1, u1:447}	{s2:1, u2:319}
brainpoolP512r1 [RFC5639]	{s1:1, u1:575}	{s2:1, u2:447}
Curve25519 [RFC7748]	{s1:1, u1:255}	{s2:1, u2:255}
Wei25519 [Appendix E.3]	{s1:1, u1:255}	{s2:1, u2:255}
Wei25519.2 [Appendix G.3]	{s1:1, u1:255}	{s2:1, u2:255}
Wei25519.-3 [Appendix G.3]	{s1:1, u1:255}	{s2:1, u2:255}
Curve448 [RFC7748]	{u1:448}	{s1:1, s2:1, u2:446}
Wei448 [Appendix M.3]	{u1:448}	{s1:1, s2:1, u2:446}
Wei448.1 [Appendix N.3]	{u1:448}	{s1:1, s2:1, u2:446}
Wei448.-3 [Appendix N.3]	{u1:448}	{s1:1, s2:1, u2:446}
secp256k1.m [Appendix L.3]	{u1:256}	{s1:1, s2:1, u2:254}

Table 3: Randomized representation of curve points, for some curves of practical interest, including curve-specific relative ordering and bit-length of substrings representing the tuple  $((u1,s1),(u2,s2))$ , resulting in the bit string left-side || right-side. (Tailored towards efficient computation of randomized representations of curve points.)

#### Appendix L. Curve secp256k1 and Friend

This section illustrates how isogenies can be used to yield curves with specific properties (here, illustrated for the "BitCoin" curve secp256k1).

### L.1. Curve Definition and Alternative Representation

The elliptic curve `secp256k1` is the Weierstrass curve  $W_{\{a,b\}}$  defined over the prime field  $GF(p)$ , with  $p:=2^{256}-2^{32}-2^9-2^8-2^7-2^6-2^4-1$ , where  $a:=0$  and  $b:=7$ . This curve has order  $h*n$ , where  $h=1$  and where  $n$  is a prime number. For this curve, domain parameter  $a$  is zero, whereas  $b$  is not. The quadratic twist of this curve has order  $h1*n1$ , where  $h1$  is a 37-bit integer and where  $n1$  is a prime number. For this curve, the base point is the point  $(GX, GY)$ .

The curve `secp256k1` is 3-isogenous to the Weierstrass curve `secp256k1.m` defined over  $GF(p)$ , which has nonzero domain parameters  $a$  and  $b$  and has as base point the pair  $(GmX, GmY)$ , where parameters are as specified in Appendix L.3 and where the related mappings are as specified in Appendix L.2.

### L.2. Switching Between Representations

Each affine point  $(X,Y)$  of `secp256k1` corresponds to the point  $(X',Y') := (u(X)/w(X)^2, Y*v(X)/w(X)^3)$  of `secp256k1.m`, where  $u$ ,  $v$ , and  $w$  are the polynomials with coefficients in  $GF(p)$  as defined in Appendix L.4.1, while the point at infinity of `secp256k1` corresponds to the point at infinity of `secp256k1.m`. Under this isogenous mapping, the base point  $(GX,GY)$  of `secp256k1` corresponds to the base point  $(GmX,GmY)$  of `secp256k1.m`. The dual isogeny maps the affine point  $(X',Y')$  of `secp256k1.m` to the affine point  $(X,Y) := (u'(X')/w'(X')^2, Y'*v'(X')/w'(X')^3)$  of `secp256k1`, where  $u'$ ,  $v'$ , and  $w'$  are the polynomials with coefficients in  $GF(p)$  as defined in Appendix L.4.2, while mapping the point at infinity  $O$  of `secp256k1.m` to the point at infinity  $O$  of `secp256k1`. Under this dual isogenous mapping, the base point  $(GmX, GmY)$  of `secp256k1.m` corresponds to a multiple of the base point  $(GX, GY)$  of `secp256k1`, where this multiple is  $l=3$  (the degree of the isogeny; see the description in Appendix F.4). Note that this isogenous map (and its dual) primarily involves the evaluation of three fixed polynomials involving the  $x$ -coordinate, which takes roughly 10 modular multiplications (or less than 1% relative incremental cost compared to the cost of an elliptic curve scalar multiplication).

### L.3. Domain Parameters

The parameters of the curve `secp256k1` and the corresponding 3-isogenous curve `secp256k1.m` are as indicated below. Here, the domain parameters of the curve `secp256k1` are as specified in [SEC2]; the domain parameters of `secp256k1.m` are "new".

General parameters (for all curves):

p 2<sup>256</sup>-2<sup>32</sup>-2<sup>9</sup>-2<sup>8</sup>-2<sup>7</sup>-2<sup>6</sup>-2<sup>4</sup>-1  
(=0xffffffff ffffffff ffffffff ffffffff ffffffff ffffffff  
fffffffe fffffffc2f)  
h 1  
n 11579208923731619542357098500868790785283756427907490438260516314  
1518161494337  
(=0xffffffff ffffffff ffffffff fffffffe baaedce6 af48a03b  
bfd25e8c d0364141)  
h1 23479460174521 (=0x1a 9bfcab89)  
n1 10131766773001318469008702396060356387381009972480920692566974370  
31  
(=0x099ee564 ea5d84f5 08913936 a761b0d5 d792a426 a7779817  
ae2f5b67)

Weierstrass curve-specific parameters (for secp256k1):

a 0 (=0x00)  
b 7 (=0x07)  
GX 55066263022277343669578718895168534326250603453777594175500187360  
389116729240  
(=0x79be667e f9dcbbac 55a06295 ce870b07 029bfcdb 2dce28d9  
59f2815b 16f81798)  
GY 32670510020758816978083085130507043184471273380659243275938904335  
757337482424  
(=0x483ada77 26a3c465 5da4fbfc 0e1108a8 fd17b448 a6855419  
9c47d08f fb10d4b8)

Weierstrass curve-specific parameters (for secp256k1.m):

a 93991599167772749909245591943117186381494883464374162770646538702  
960816911535  
(=0xcfc5c21 75e2ef7d ccdce737 770b7381 5a2f13c5 09035ca2  
54a14ac9 f08974af)  
b 1771 (=0x06eb)

GmX 26591621185618668069038227574782692264471832498547635565821216767  
730887659845

(=0x3aca5300 959fa1d0 baf78dcf f77a616f 395e586d 67aced0a  
88798129 0c279145)

GmY 67622516283223102233819216063319565850973524550533340939716651159  
860372686848

(=0x9580fce5 3a170f4f b744579f f3d62086 12cd6a23 3e2de237  
f976c6a7 8611c800)

#### L.4. Isogeny Details

The isogeny and dual isogeny are both isogenies with degree  $l=3$ . Both are specified by a triple of polynomials  $u$ ,  $v$ , and  $w$  (resp.  $u'$ ,  $v'$ , and  $w'$ ) of degree 3, 3, and 1, respectively, with coefficients in  $\text{GF}(p)$ . The coefficients of each of these polynomials are specified in Appendix L.4.1 (for the isogeny) and in Appendix L.4.2 (for the dual isogeny). For each polynomial in variable  $x$ , the coefficients are tabulated as the sequence of coefficients of  $x^0$ ,  $x^1$ ,  $x^2$ , ..., in hexadecimal format.

##### L.4.1. Isogeny Parameters

###### L.4.1.1. Coefficients of $u(x)$

0 0x54  
  
1 0xa4d89db3ed06c81e6143ec2eca9f761d8d17260dc229e1da1f73f714506872a9  
  
2 0xcc58ffccbd9febb4a66222c7d1311d988d88c0624bcd68ec4c758a8e67dfd99b  
  
3 0x01

###### L.4.1.2. Coefficients of $v(x)$

0 0x1c  
  
1 0x94c7bc69befd17f2fae2e3ebf24df1f355d181fa1a8056103ba9baad4b40f029  
  
2 0xb2857fb31c6fe18ef993342bb9c9ac64d44d209371b41d6272b04fd61bcfc851  
  
3 0x01

L.4.1.3. Coefficients of  $w(x)$ 

```

0  0xe62c7fe65ecff5da53311163e8988ecc46c4603125e6b476263ac546b3efae5
1  0x01

```

## L.4.2. Dual Isogeny Parameters

L.4.2.1. Coefficients of  $u'(x)$ 

```

0  0x8e38e38e38e38e38e38e38e38e38e38e38e38e38e38e38e38e38e38e38e38e38e38daaaaaa8c7
1  0x44cd5cd7ce55a801725891578f7356bd936355fd0e2f538797cecff7a37244
2  0x668d0011162006c3c889f4680f9a4b77d0d26a89e6bb87b13bd8d1cfdd600a41
3  0x8e38e38e38e38e38e38e38e38e38e38e38e38e38e38e38e38e38e38e38e38e38e38daaaaaa88c

```

L.4.2.2. Coefficients of  $v'(x)$ 

```

0  0x4bda12f684bda12f684bda12f684bda12f684bda12f684bda12f684b8e38e23c
1  0x519ba9c1f48f68054def6a410f0fa6e8b71c6c3b4a8958324681f6508c01fada
2  0xb34680088b100361e444fa3407cd25bbe8693544f35dc3d89dec68e76eb00338
3  0x2f684bda12f684bda12f684bda12f684bda12f684bda12f684bda12f38e38d84

```

L.4.2.3. Coefficients of  $w'(x)$ 

```

0  0x4d7a804ce3901e71066ccbd44636539b2bb2df6c8e4be29d8d4fb028e43033de
1  0x01

```

## Appendix M. Curve448 and Cousins

This section introduces curves related to Curve448 and explains their relationships.

## M.1. Curve Definition and Alternative Representations

The elliptic curve Curve448 is the Montgomery curve  $M_{\{A,B\}}$  defined over the prime field  $\text{GF}(p)$ , with  $p := 2^{448} - 2^{224} - 1$ , where  $A := 156326$  and  $B := 1$ . This curve has order  $h \cdot n$ , where  $h = 4$  and where  $n$  is a prime number. For this curve,  $A^2 - 4$  is not a square in  $\text{GF}(p)$ , whereas  $A - 2$  is. The quadratic twist of this curve has order  $h_1 \cdot n_1$ , where  $h_1 = 4$  and where  $n_1$  is a prime number. For this curve, the base

point is the point  $(G_u, G_v)$ , where  $G_u=5$  and where  $G_v$  is an even integer in the interval  $[0, p-1]$ .

This curve has the same group structure as (is "isomorphic" to) the twisted Edwards curve  $E_{\{a,d\}}$  defined over  $GF(p)$ , with as base point the point  $(G_x, G_y)$ , where parameters are as specified in Appendix M.3. This curve is denoted as Ed448. For this curve, the parameter  $a$  is a square in  $GF(p)$ , whereas  $d$  is not, so the group laws of Appendix C.3 apply.

The curve is also isomorphic to the elliptic curve  $W_{\{a,b\}}$  in short-Weierstrass form defined over  $GF(p)$ , with as base point the point  $(G_x, G_y)$ , where parameters are as specified in Appendix M.3. This curve is denoted as Wei448.

## M.2. Switching between Alternative Representations

Each affine point  $(u, v)$  of Curve448 corresponds to the point  $(X, Y):=(u + A/3, v)$  of Wei448, while the point at infinity of Curve448 corresponds to the point at infinity of Wei448. (Here, we used the mappings of Appendix D.2 and that  $B=1$ .) Under this mapping, the base point  $(G_u, G_v)$  of Curve448 corresponds to the base point  $(G_x, G_y)$  of Wei448. The inverse mapping maps the affine point  $(X, Y)$  of Wei448 to  $(u, v):=(X - A/3, Y)$  of Curve448, while mapping the point at infinity of Wei448 to the point at infinity of Curve448. Note that this mapping involves a simple shift of the first coordinate and can be implemented via integer-only arithmetic as a shift of  $-\delta$  for the isomorphic mapping and a shift of  $\delta$  for its inverse, where  $\delta:=(p-A)/3$  is the integer defined by

```
delta 24227957476520229684977460262933484478454712022910602009383006
      63935374427222435908954654612328921819766962948206145457870178326
      72736371
```

```
(=0x55555555 55555555 55555555 55555555 55555555 55555555
55555554 ffffffff ffffffff ffffffff ffffffff ffffffff ffffffff
ffff3473).
```

(Note that, depending on the implementation details of the field arithmetic, one may have to shift the result by  $+p$  or  $-p$  if this integer is not in the interval  $[0, p-1]$ .)

The curve Ed448 is isomorphic to the curve Curve448, where the base point  $(G_u, G_v)$  of Curve448 corresponds to the base point  $(G_x, G_y)$  of Ed448 and where the point at infinity and the point  $(0,0)$  of order two of Curve448 correspond to, respectively, the point  $(0, 1)$  and the point  $(0, -1)$  of order two of Ed448 and where each other point  $(u, v)$

of Curve448 corresponds to the point  $(c*u/v, (u+1)/(u-1))$  of Ed448, where  $c$  is the element of  $GF(p)$  defined by

```
c  sqrt((A-2)/B)

19788846729546443953835400975385803825683515259105980214819977919
60874042320025157136042631277930307478554244641856917664538448351
92428

(=0x45b2c5f7 d649eed0 77ed1ae4 5f44d541 43e34f71 4b71aa96
c945af01 2d182975 0734cde9 faddbda4 c066f7ed 54419ca5 2c85dele
8aae4e6c).
```

(Here, we used the mapping of Appendix D.1 and normalized this using the mapping of Appendix F.1 (where the element  $s$  of that appendix is set to  $c$  above).) The inverse mapping from Ed448 to Curve448 is defined by mapping the point  $(0, 1)$  and the point  $(0, -1)$  of order two of Ed448 to, respectively, the point at infinity and the point  $(0,0)$  of order two of Curve448 and having each other point  $(x, y)$  of Ed448 correspond to the point  $((y + 1)/(y - 1), c*(y + 1)/((y-1)*x))$  of Curve448.

The curve Ed448 is isomorphic to the Weierstrass curve Wei448, where the base point  $(Gx, Gy)$  of Ed448 corresponds to the base point  $(GX, GY)$  of Wei448 and where the identity element  $(0,1)$  and the point  $(0,-1)$  of order two of Ed448 correspond to, respectively, the point at infinity  $O$  and the point  $(A/3, 0)$  of order two of Wei448 and where each other point  $(x, y)$  of Ed448 corresponds to the point  $(X, Y) := ((y+1)/(y-1)+A/3, c*(y+1)/((y-1)*x))$  of Wei448, where  $c$  was defined before. (Here, we used the mapping of Appendix D.3.) The inverse mapping from Wei448 to Ed448 is defined by mapping the point at infinity  $O$  and the point  $(A/3, 0)$  of order two of Wei448 to, respectively, the identity element  $(0,1)$  and the point  $(0,-1)$  of order two of Ed448 and having each other point  $(X, Y)$  of Wei448 correspond to the point  $(c*(X-A/3)/Y, (X-A/3+1)/(X-A/3-1))$  of Ed448.

Note that these mappings can be easily realized if points are represented in projective coordinates, using a few field multiplications only, thus allowing switching between alternative curve representations with negligible relative incremental cost.

### M.3. Domain Parameters

The parameters of the Montgomery curve and the corresponding isomorphic curves in twisted Edwards curve and short-Weierstrass form are as indicated below. Here, the domain parameters of the Montgomery curve Curve448 and of the twisted Edwards curve Ed448 are as specified in [RFC7748]; the domain parameters of Wei448 are "new".



IMPORTANT NOTE: the supposed base point of Ed448 specified in [RFC7748] is incorrect, since it has order  $2^n$ , and - in the notation below - that point is the point  $(Gx, -Gy) = -(Gx, Gy) + (0, -1)$ . The birational map in that document is also incorrect.

General parameters (for all curve models):

p  $2^{448} - 2^{224} - 1$

(=0xffffffff ffffffff ffffffff ffffffff ffffffff ffffffff  
 ffffffff ffffffff ffffffff ffffffff ffffffff ffffffff ffffffff  
 ffffffff)

h 4

n 18170968107390172263733095197200113358841034017182951507037254979  
 51460039615395857161957552916923759633102937090916623047737558596  
 49779

(= $2^{446}$  - 0x8335dc16 3bb124b6 5129c96f de933d8d 723a70aa  
 dc873d6d 54a7bb0d)

h1 4

n1 18170968107390172263733095197200113358841034017182951507037254979  
 51601601218258006270024365576458970017341485218301563757529931495  
 32941

(= $2^{446}$  + 0x0335dc16 3bb124b6 5129c96f de933d8d 723a70aa  
 dc873d6d 54a7bb0d)

Montgomery curve-specific parameters (for Curve448):

A 156326 (=0x0262a6)

B 1 (=0x01)

Gu 5 (=0x05)

Gv 35529392678556817526412750206378333480897639938771427183188089843  
 51690887869674100029326737658645509101427741472681058389855952906  
 06362

(=0x7d235d12 95f5b1f6 6c98ab6e 58326fce cbae5d34 f55545d0  
 60f75dc2 8df3f6ed b8027e23 46430d21 1312c4b1 50677af7 6fd7223d  
 457b5b1a)

Edwards curve-specific parameters (for Ed448):

a 1 (0x01)

d  $39082/39081 = (A+2)/(A-2)$

(=611975850744529176160423220965553317543219696871016626328968936  
41508786004263647489178559928366602041476867897998937814706546281  
5545017)

(=0xd78b4bdc 7f0daf19 f24f38c2 9373a2cc ad461572 42a50f37  
809b1da3 412a12e7 9ccc9c81 264cfe9a d0809970 58fb61c4 243cc32d  
baa156b9)

Gx 34539749303972951637400860415053741026665526007518329021640697028  
16456950736723444304817877593406332217083915834240417889241245677  
00732

(=0x79a70b2b 70400553 ae7c9df4 16c792c6 1128751a c9296924  
0c25a07d 728bdc93 e21f7787 ed697224 9de732f3 8496cd11 69871309  
3e9c04fc)

Gy 3/2

36341936214780344527466190394400226717682068034365903014074509959  
03061640833653863431981918493382729650444422309218186805267490091  
82721

(=0x7ffffffff ffffffff ffffffff ffffffff ffffffff ffffffff  
ffffffff 80000000 00000000 00000000 00000000 00000000 00000000  
00000001)

Weierstrass curve-specific parameters (for Wei448):

a 48455914953040459369954920525866968956909424045821204018766013278  
70748854444871817909309224657843639533925896412290915740356571996  
37535

(=0xaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa  
aaaaaaaa9 ffffffff ffffffff ffffffff ffffffff ffffffff fffffffe  
1a76d41f)

b 26919952751689144094419400292148316087171902247678446677092229599  
28193808024928787727394013698802021963292164673494953191916856645  
13904

(=0x5ed097b4 25ed097b 425ed097 b425ed09 7b425ed0 97b425ed  
097b425e 71c71c71 c71c71c7 1c71c71c 71c71c71 c71c71c7 1c72c87b  
7cc69f70)

GX 48455914953040459369954920525866968956909424045821204018766013278  
70748854444871817909309224657843639533925896412290915740356653456  
29073

(=0xaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa  
aaaaaaaa 00000000 00000000 00000000 00000000 00000000 00000000  
0000cb91)

GY 35529392678556817526412750206378333480897639938771427183188089843  
51690887869674100029326737658645509101427741472681058389855952906  
06362

(=0x7d235d12 95f5b1f6 6c98ab6e 58326fce cbae5d34 f55545d0  
60f75dc2 8df3f6ed b8027e23 46430d21 1312c4b1 50677af7 6fd7223d  
457b5b1a)

## Appendix N. Further Cousins of Curve448

This section introduces some further curves related to Curve448 and explains their relationships.

### N.1. Further Alternative Representations

The Weierstrass curve Wei448 is isomorphic to the Weierstrass curve Wei448.1 defined over  $\text{GF}(p)$ , with as base point the pair  $(G1X, G1Y)$ , and isogenous to the Weierstrass curve Wei448.-3 defined over  $\text{GF}(p)$ , with as base point the pair  $(G3X, G3Y)$ , where parameters are as specified in Appendix N.3 and where the related mappings are as specified in Appendix N.2.

The Edwards curve Ed448 is isogenous to the Edwards curve Edwards448 defined over  $\text{GF}(p)$ , with as base point the pair  $(G1x, G1y)$ , where parameters are as specified in Appendix N.3 and where the related mappings are as specified in Appendix N.2. For this curve, the domain parameter  $a$  is a square in  $\text{GF}(p)$ , whereas  $dl$  is not, so the group laws of Appendix C.3 apply.

### N.2. Further Switching

Each affine point  $(X, Y)$  of Wei448 corresponds to the point  $(X', Y') := (X*s^2, Y*s^3)$  of Wei448.1, where  $s$  is the element of  $\text{GF}(p)$  defined by

s 52322274343677442779379520589771028818568404587729117919590511061  
93509510238347880134473888687471465216641846232724641298954890800  
00881

```
(=0xb848cd01 981d2f83 f2829b42 eb86914e 88f44c9d 05dcdbdf
dbdd1e56 c4674bc8 d6d90d91 862a38f5 ca797ca7 f21c05cf a7ac32bf
d2ca0171),
```

while the point at infinity of Wei448 corresponds to the point at infinity of Wei448.1. (Here, we used the mapping of Appendix F.3.) Under this mapping, the base point  $(GX, GY)$  of Wei448 corresponds to the base point  $(G1X, G1Y)$  of Wei448.1. The inverse mapping maps the affine point  $(X', Y')$  of Wei448.1 to  $(X, Y) := (X'/s^2, Y'/s^3)$  of Wei448, while mapping the point at infinity  $O$  of Wei448.1 to the point at infinity  $O$  of Wei448. Note that this mapping (and its inverse) involves a modular multiplication of both coordinates with fixed constants  $s^2$  and  $s^3$  (respectively,  $1/s^2$  and  $1/s^3$ ), which can be precomputed.

The point at infinity and the point  $(A/3, 0)$  of order two of Wei448 both correspond to the point at infinity of Wei448.-3, while each other point  $(X, Y)$  of Wei448 corresponds to the point  $(X', Y') := (X1*t^2, Y1*t^3)$  of Wei448.-3, where  $(X1, Y1) = (u(X)/w(X), Y*v(X)/w(X)^2)$ , where  $u$ ,  $v$ , and  $w$  are the polynomials with coefficients in  $GF(p)$  as defined in Appendix N.4.1 and where  $t$  is the element of  $GF(p)$  defined by

```
t 23579450751475691430882365546539966269774125426758968522698856022
13378944265540874438945283200254318223329383397068961863760712339
07365
```

```
(=0x530c9a1d 7cf071d0 9646b83d b246626b 4e57ba5d 6a791bef
76197254 3209dc5c 20d81498 d5ab8d7a 2fb22507 ca68c040 a6c82eb3
b6c7aaa5).
```

(Here, we used the isogenous mapping of Appendix F.4.) Under this isogenous mapping, the base point  $(GX, GY)$  of Wei448 corresponds to the base point  $(G3X, G3Y)$  of Wei448.-3. The dual isogeny maps the point at infinity  $O$  and the point  $(\tau, 0)$  of order two of Wei448.-3, where  $\tau$  is the element of  $GF(p)$  defined by

```
tau 42178595713080601145580616893463205889346047807394283240821661315
01870168726890624132409486822657385666418069563147259152341712826
86207
```

```
(=0x948eabcf 057e0d55 9c372c98 075ddacf 6f3d19bc 514e5d23
248d685b 75f97a10 36696aaf 61c02d8e 3da778c3 8d9fda05 54c9258b
3c0e80ff),
```

to the point at infinity  $O$  of Wei448, while mapping each other point  $(X', Y')$  of Wei448.-3 to the affine point  $(X, Y) := (u'(X1)/w'(X1), Y1*v'(X1)/w'(X1)^2)$  of Wei448, where

$(X1, Y1) = (X'/t^2, Y'/t^3)$  and where  $u'$ ,  $v'$ , and  $w'$  are the polynomials with coefficients in  $GF(p)$  as defined in Appendix N.4.2. Under this dual isogenous mapping, the base point  $(G3X, G3Y)$  of Wei448.-3 corresponds to a multiple of the base point  $(GX, GY)$  of Wei448, where this multiple is  $l=2$  (the degree of the isogeny; see the description in Appendix F.4). Note that this isogenous map (and its dual) primarily involves the evaluation of three fixed polynomials involving the  $x$ -coordinate, which takes only a few modular multiplications (less than 0.5% relative incremental cost compared to the cost of an elliptic curve scalar multiplication).

Each point  $(x1, y1)$  of Edwards448 with nonzero coordinates corresponds to the point  $(x, y)$  of Ed448, where

$$x = c * x1 * y1 / (1 - d1 * x1^2 * y1^2) = c * x1 * y1 / (2 - x1^2 - y1^2) \text{ and}$$

$$y = (1 + d1 * x1^2 * y1^2) / (y1^2 - x1^2) = -(x1^2 + y1^2) / (x1^2 - y1^2),$$

while each other point (i.e., a point of order 1, 2, or 4) corresponds to the identity element  $(0, 1)$  of Ed448. (Here, we used the 4-isogenous mapping of Appendix F.4). Under this isogenous mapping, the base point  $(G1x, G1y)$  of Edwards448 corresponds to the base point  $(Gx, Gy)$  of Ed448. The dual isogeny maps each point  $(x, y)$  of Ed448 to the point  $(x1, y1)$  of Edwards448, where

$$x1 = (4 * x * y / c) / (y^2 - x^2) \text{ and}$$

$$y1 = (1 - d * x^2 * y^2) / (1 + d * x^2 * y^2) = (2 - x^2 - y^2) / (x^2 + y^2).$$

Under this dual isogenous mapping, the base point  $(Gx, Gy)$  of Ed448 corresponds to a multiple of the base point  $(G1x, G1y)$  of Edwards448, where this multiple is  $l=4$  (the degree of the isogeny; see the description in Appendix F.4). Note that this isogenous map (and its dual) primarily involves the evaluation of three fixed polynomials, which takes only a few multiplications (less than 0.5% relative incremental cost compared to the cost of an elliptic curve scalar multiplication).

Each point  $(x1, y1)$  of Edwards448 with nonzero coordinates corresponds to the point  $(u, v)$  of Curve448, where

$$u = y1^2 / x1^2 \text{ and } v = y1 * (2 - x1^2 - y1^2) / x1^3,$$

while each other point (i.e., a point of order 1, 2, or 4) corresponds to the point at infinity of Curve448. Under this isogenous mapping, the base point  $(G1x, G1y)$  of Edwards448 corresponds to the base point  $(Gu, Gv)$  of Curve448. The dual isogeny maps both the point at infinity and the point  $(0, 0)$  of order two of

Curve448 to the identity element (0,1) of Edwards448, while each other point (u,v) of Curve448 corresponds to the point (x1,y1) of Edwards448, where

$$x1 = 4*(u^2-1)*v/((u^2-1)^2+4*v^2) \text{ and}$$

$$y1 = u*((u^2-1)^2-4*v^2)/(2*(u^2+1)*v^2-u*(u^2-1)^2).$$

Under this dual isogenous mapping, the base point (Gu, Gv) of Curve448 corresponds to a multiple of the base point (Glx, Gly) of Edwards448, where this multiple is l=4 (the degree of the isogeny; see above).

### N.3. Further Domain Parameters

The parameters of the Weierstrass curve with a=1 that is isomorphic with Wei448 and the parameters of the Weierstrass curve with a=-3 that is isogenous with Wei448 are as indicated below. Both domain parameter sets can be exploited directly to derive more efficient point addition formulae, should an implementation facilitate this. The domain parameters of the Edwards curve Edwards448 are as specified in [RFC7748].

General parameters: same as for Wei448 (see Appendix M.3)

Weierstrass curve-specific parameters (for Wei448.1, i.e., with a=1):

```
a  1 (=0x01)

b  65961281701807170531944804985907990287225248056560036392380945951
   38183088507635437786021044927715119224497407914895790669345268896
   52743

   (=0xe8528596 bfbcbac9 7ebdbe4e 9683e25c 73a5ff37 6c4cd400
   5a75c425 8e3eb05a 9f6f8c24 24cb5aa9 0dcf9fa4 cab6691d 5530347c
   28437207)

G1X 19236211982508211644805033459306273038523230481309141518540414163
     72091186292458482231912460243257247478684005448999746809691007995
     9723

     (=0x06c672d5 b5bae33b 010fa210 9de7937a 95db8ffc 043c507f
     5e0d07a1 25382eaf 13f5fc3b 75db2614 6e6d002f d8364ed6 c9bc8fbf
     bbda22ab)

G1Y 30319443056877169804488072384563064288675576234196773667920807567
     7917792785875562195875622206632465988308466319556948821775845861
     64158
```

```
(=0x6ac9c53c 767cd3ae cbf904a1 2923502f 115355d1 6ae8911c
5c92f612 aa854455 d1e6d29f 4db4ddea 519a174f c0dd2505 ec3328ba
250a07be)
```

Weierstrass curve-specific parameters (for Wei448.-3, i.e., with a=-3):

a -3

```
(=0xffffffff ffffffff ffffffff ffffffff ffffffff ffffffff
fffffffe ffffffff ffffffff ffffffff ffffffff ffffffff ffffffff
fffffffc)
```

b 69993768681000150084833669961900533067383335592494498709534693464  
91314250731583068774689950893229681024927315747794587331422088592  
54465

```
(=0xf686723d 80e29d06 2d00a9f1 3305b698 85790019 cca78035
9dac226b efb1ae21 125397dd 16f255b0 cc5d18e5 43582a1c af90dfe2
c0aaec1)
```

G3X 40677474994869876470916133424311516856662407970799424837841348421  
87696274665113140719001227030116551378877280368526334985627104680  
88795

```
(=0x8f452c6b dc3265dd 580b2638 59a02b20 198cc020 1dd7fba1
8b431694 4a936052 fb4e4a41 93d01fa5 5fb5c732 7393208b 8170f3f2
be78d3db)
```

G3Y 54594210970205994927260789585006437115117066846498189378285031510  
90310290468347714929366106635470978666795512446629051235704504868  
06147

```
(=0xc0494f90 461db11c 35fb7646 8349399a ae230351 11330cce
b7473244 ab63c955 cf6ec02f 2656b439 44b19f4b 52eef12e 73026bbc
84444683)
```

Edwards curve-specific parameters (for Edwards448):

a 1 (0x01)

d1 -39081 = -(A-2)/4

```
(=726838724295606890549323807888004534353641360687318060281490199
18061232816673077268639638369867654593008888446184363736105349801
8326358)
```

```
(=0xffffffff ffffffff ffffffff ffffffff ffffffff ffffffff
fffffffe ffffffff ffffffff ffffffff ffffffff ffffffff ffffffff
ffff6756)
```

G1x 22458004029592430018760433409989603624678964163256413424612546168  
69504154674060329090291928693579532825780320751464461736746026352  
47710

```
(=0x4f1970c6 6bed0ded 221d15a6 22bf36da 9e146570 470f1767
ea6de324 a3d3a464 12ae1af7 2ab66511 433b80e1 8b00938e 2626a82b
c70cc05e)
```

Gly 29881921007848149267601793044393067343754404015408024209592824137  
23315061898358760035368786554187847339823032335034625005315450628  
32660

```
(=0x693f4671 6eb6bc24 88762037 56c9c762 4bea7373 6ca39840
87789c1e 05a0c2d7 3ad3ff1c e67c39c4 fdbd132c 4ed7c8ad 9808795b
f230fa14)
```

#### N.4. Isogeny Details

The isogeny and dual isogeny are both isogenies with degree  $l=2$ . Both are specified by a triple of polynomials  $u$ ,  $v$ , and  $w$  (resp.  $u'$ ,  $v'$ , and  $w'$ ) of degree 2, 2, and 1, respectively, with coefficients in  $\text{GF}(p)$ . The coefficients of each of these polynomials are specified in Appendix N.4.1 (for the isogeny) and in Appendix N.4.2 (for the dual isogeny). For each polynomial in variable  $x$ , the coefficients are tabulated as the sequence of coefficients of  $x^0$ ,  $x^1$ ,  $x^2$ , ..., in hexadecimal format.

#### N.4.1. Isogeny Parameters

#### N.4.1.1. Coefficients of $u(x)$

[illegible]

#### N.4.1.2. Coefficients of $v(x)$

[illegible]





multiple (here: 2019) of the base point of the curve in question and where the private key  $k$  is the integer

```
k  62662039304523906689788124833289384446202946474440057655160773695
    63756342505410402166230018620066482794080866641616932013327623579
    01952

    (=0xdcbb3bbb9 e42d7aca fe62052d 902123c7 0872b984 4c1e199f
    7c5d37bd 1171102b c20a6352 d9c91886 29b685de 51441e84 3afe2665
    5251aa80).
```

In the examples below, each curve point is represented using the compressed point representation (see Appendix I.8), but where (for historical reasons) the parity bit  $t$  of the compressed curve point in question is represented by  $0x00$  or  $0x80$  depending on whether  $t=0$  or  $t=1$ , respectively. Notice that this representation corresponds to the compressed point representation of Appendix I.8, but with the bit-ordering in the 1-octet representation of  $t$  reversed. (Note that this puts the parity bit  $t$  in the leftmost bit position of the octet string if one follows the MSB/msb representation conventions.) For simplicity, this representation is again called the "squeezed" representation, although each point is now represented as a 57-octet string and, thereby, one octet longer than the tight representation of elements of  $GF(p)$ . As before, the ordering convention (see Appendix I.7) depends on the underlying curve model in question. Here, points of a Weierstrass curve are represented in tight MSB/msb-order, points of a Montgomery curve in tight LSB/msb-order, and points of a twisted Edwards curve in tight LSB/lsh-order. For points that are a public key, the corresponding private keys are represented as 56-octet strings, using the same ordering conventions as with the public keys. For affine points, we also give the tight representation of each of its coordinates (as 56-octet strings), using the same ordering conventions as used with the squeezed point representation. For further details, see the examples themselves.

#### O.1. Example with Curve448

$P_m=(u, v)$ ,  $k \cdot P_m=(u_1, v_1)$ , and  $(k+1) \cdot P_m=(u_2, v_2)$  with Curve448:

```
u  53298594738299085772373536080133483236673782578895339676785179923
    90764298300090102709453866054695061082746243636045110750296444932
    27715

    (=0xbbbb91ba3 b0ef74c3 214394b4 d8f0d32d c4a92193 5f573009
    39fd86a3 8d54be2a 4d63380b 692381bb ed7339fd dca7b0cd a80166fe
    18c086c3).
```

v 30578727850066757341435137807347775064915058999485530015946871157  
86794631407274936870618580714107931661730999350222644894729285604  
97149  
  
(=0x6bb38e82 8d52337f 6f0395ef dc16c776 52162f5e 309112ae  
fc7401bf 0cfb0499 ebled555 bf507ebc c33b4753 2d6dc6c5 d68dealc  
cle4clfd).

u1 64579461799301726935877447646800238443923683299745374127971411973  
1251529516179188974322804922279188968365877164188075095074418806  
82513  
  
(=0xe37497bf 9f704689 54ec6537 cbbe91d0 3ffcdcdb 8b707253  
a2212cdb e020ba9a 0bf65a1d 5d9a128a f85c63a2 79a00139 7aca56db  
15335011).

v1 55735504615964066386264989698774850924544182484936624265048483231  
35693859362627880184586282439234602798023594054611737412667543758  
11547  
  
(=0xc44e5e0f 2c254d23 1dc082db 77175e8c fd37793c 22ebe200  
77905a5f 750b3c9f 4a95d4d5 4e1a1e54 d2d31689 4249252d 0c8b1c45  
1c1481db).

u2 32685564331119553171673802371596819258307818641496728161547328225  
07595618587323619256769558535630624960575212644680149034661008254  
8876  
  
(=0x0b831eca 9c6215b0 5d830361 4013732f 7a9dd07f ebb9441e  
49129264 eb724f44 dc53671c ffabb9ee 0c02aa74 b083cd82 a821a4cf  
6f6d8c8c).

v2 57682103223585233918507344950062950306770296215271320612937204938  
77499282103483092990510136901415757273082719665657484294344333591  
20741  
  
(=0xcb2988a4 6e37f9a9 a7a1255b 2fd2eea9 82308e7c eb8e18b8  
2252175f fd416a10 5984c6b8 36470e48 31879293 8f6139c6 f96164cb  
14010965).

As suggested in Appendix C.2, the  $v$ -coordinate of  $k \cdot P_m$  can be indirectly computed from the  $u$ -coordinates of  $P_m$ ,  $k \cdot P_m$ , and  $(k+1) \cdot P_m$ , and the  $v$ -coordinate of  $P_m$ , which allows computation of the entire point  $k \cdot P_m$  (and not just its  $u$ -coordinate) if  $k \cdot P_m$  is computed using the Montgomery ladder (as, e.g., [RFC7748] recommends), since that algorithm computes both  $u_1$  and  $u_2$  and the  $v$ -coordinate of the point  $P_m$  may be available from context.

The representation of  $k$  and the compressed representations of  $P_m$  and  $k*P_m$  in tight LSB/msb-order are given by

```
repr(k)      0x80aa5152 6526fe3a 841e4451 de85b629 8618c9d9 52630ac2
              2b107111 bd375d7c 9f191e4c 84b97208 c7232190 2d0562fe
              ca7a2de4 b9bbb3dc;

repr(Pm)      0xc386c018 fe6601a8 cdb0a7dc fd3973ed bb812369 0b38634d
              2abe548d a386fd39 0930575f 9321a9c4 2dd3f0d8 b4944321
              c374efb0 a31bb9bb 80;

repr(k*Pm)     0x11503315 db56ca7a 3901a079 a2635cf8 8a129a5d 1d5af60b
              9aba20e0 db2c21a2 5372708b dbdcfc3f d091becb 3765ec54
              8946709f bf9774e3 80,
```

where the leftmost bit of the rightmost octet indicates the parity of the  $v$ -coordinate of the point of Curve448 in question (which, in this case, are both one, since  $v$  and  $v_1$  are odd). See Appendix H.2 and Appendix I for further detail on (squeezed) point compression.

The scalar representation and (squeezed) point representation illustrated above are consistent with the representations specified in [RFC7748], except that in [RFC7748] only an affine point's  $u$ -coordinate is represented (i.e., the  $v$ -coordinate of any point is always implicitly assumed to have an even value) and that the representation of the point at infinity is not specified. (Note that due to the bit-length of the prime  $p$ , the lossless representation requires an additional octet compared to the lossy representation without  $v$ -coordinate.) Another difference is that [RFC7748] allows non-unique representations of some elements of  $GF(p)$ , whereas our representation conventions do not (since tight).

A randomized representation  $(t_1, t_2)$  of the point  $k*P_m$  in tight LSB/msb order is given by

```
t1      642695971489808425948939115432957219707501931105169269237
        122551860533279049805112466411050091592893048844749561382
        909707113070546618079

        (=0xdf86cb83 aelca6e6 da6afbaf afbb2fc0 606a136f 80eea078
        c868a5d7 7e638d09 99518385 65250cf1 9c034f96 1fa28f54
        f3016600 68335de2)

t2      569275737967591640709387827593956375775147481657775744720
        460881642951497067363381071471046477130052706607411985560
        522861593611384288817
```

```
(=0x3176361c 580a7bcd d7880d84 aba10bc6 57010328 afb728cc
2016461b 246bef46 0eb4bb04 8c1a3616 c3f74a56 3cc1790f
6472256b ca3481c8),
```

where this representation is defined in Appendix K.5 and uses the mapping of Appendix K.3.2 with the default square root function.

This representation can also be expressed in tight LSB/msb order as the pair  $((u_1, s_1), (u_2, s_2))$ , where  $(s_1, s_2) := (0, 1)$  and where

```
u1      136243399181827781288243566840664309780937553734476297986
555794212826774821697384612603539068963961668560923975117
429548012444908081181
```

```
(=0x1d50f12f 9d4a9f5b c49d8a59 0403a454 e9ab4208 ccde0595
11d72af1 f44cefe4 0c743579 6502c443 730c55e9 2981fce1
f172d988 fa7efc2f)
```

```
u2      316511454563659405723248762668968632925539726790750815582
281632434431599609191814743278306058750675581434472930261
478756904493088717708
```

```
(=0x8c3b4bf7 5ff5eaf5 4df2119b a413785d 73059f0b aa677e16
e6eb7cf6 1e066961 e54e4b52 6ae528b1 d3c8cf8e aa3a7df0
3b7a9a0d bb827a6f),
```

where this uses the default completed mapping defined in Appendix K.6 and the mapping of Appendix K.4.2 (with the default square root function).

## O.2. Example with Ed448

$Pe = (x, y)$ ,  $k * Pe = (x_1, y_1)$ , and  $(k+1) * Pe = (x_2, y_2)$  with Ed448:

```
x      12711234107145442394649604543297947887906244696692372551963816418
93066253979844478364753304240794498368174540810674220788120782656
62747
```

```
(=0x2cc52fd1 6370554f 00c0f73f 64bda240 f5950177 d9033f6d
74acd12d 68c79a51 315f556f 240973f9 e5f71ed7 9314ee9d c87f0b1b
bcc0fd1b).
```

```
y      69251010954633529003803699627438795111055087299023774963200632446
22677618700964599963790149315020469517869703738619380660774687159
85238
```

```
(=0xf3e8bb95 c9675fd0 0c388fc5 e96cfbc7 3c19d945 76849979
34c4ab60 73c4a763 c2a89bac d3879838 f4de11a3 3a4710c2 396dea1d
cc012956).
```

```
x1 69268794439088733926883958090256942256857349796922332363888137509
71700910417786272464007666020220956482611896297610096130434552586
39205
```

```
(=0xf3f8c472 ca2e730b 05cc9092 f9d40956 029113e3 e92c2d55
76406db2 c2903721 62f43371 1c0ec80c f8d7222d 1d701467 9da18531
0fb5bb65).
```

```
y1 50516707418203531159001223293623288296803299598968490915066154362
78541820739332329525138363312119838075438487384161435963107103409
09734
```

```
(=0xbleccbfc 5f5f92e8 d9129d14 b721c524 96fc1b1f a4c17c5f
e4979b0c 763f34ba 91299376 d2499220 19b05f56 c3bb6b5d ac988271
287d7aa6).
```

```
x2 67287262124444231243108222498849910362455590990935326363062127166
04126947894055981270997819628982374416022607672923451356182938105
87868
```

```
(=0xecfela4f a4cd7e2f 19afc16 1ce2198f 0a850beb 41afa209
94741609 5b1a858a 8e9548f5 011d188e d50484d3 119103f6 8bcd5ba2
a6e3e8dc).
```

```
y2 13744276256057290540518554008940700979716578667786691114397525367
92684542875757407063179870154307882588988293167000249160114881659
30341
```

```
(=0x3068a338 4016ebfd a229ac73 b5c30bba ff67e183 71d1185f
19dfbbe 28478baf 9034ebad 51407f01 35162743 c2c234bc 2d484c13
552ea565).
```

The representation of  $k$  and the compressed representations of  $P_e$  and  $k \cdot P_e$  in tight LSB/lsb-order are given by

```
repr(k)      =0x01558a4a a6647f5c 2178228a 7ba16d94 6118939b 4ac65043
              d4088e88 bdecba3e f9987832 219d4e10 e3c48409 b4a0467f
              535eb427 9ddcd3b;
```

```
repr(Pe)     =0x6a948033 b857b69c 4308e25c c5887b2f 1c19e1cb 35d91543
              c6e523ce 06d5232c 9e99216e a29b983c e3df3697 a3f11c30
              0bfae693 a9dd17cf 01;
```

```
repr(k*Pe) =0x655ebe14 8e411935 bad6ddc3 6afa0d98 0449924b 6ec99489
5d2cfc6e 30d9e927 fa3e8325 f8d83f69 24a384ed 28b9489b
1749fafa 3fd3378d 01,
```

where the rightmost bit of the rightmost octet indicates the parity of the x-coordinate of the point of Ed448 in question (which, in this case, are both one, since x and x1 are odd). See Appendix H.3 and Appendix I for further detail on (squeezed) point compression.

The scalar representation and (squeezed) point representation illustrated above are fully consistent with the representations specified in [RFC8032]. Note that, contrary to [RFC7748], [RFC8032] requires unique representations of all elements of  $GF(p)$ .

A randomized representation (t1, t2) of the point  $k*Pe$  in tight LSB/lsb order is given by

```
t1      397357047759003459380102071532091085834125520561197668989
747600577137881485970346806080038194336473483709104865191
806326006691504231547
```

```
(=0xde295d0e 5efceb9b f43967ca be45a54b alf75bdd a4b1b1b3
b24a8d1d f2056329 e506867e c968aa8b 866017e4 f0cbc343
2cf8e7fa 0b202fd1)
```

```
t2      711800301530600330791068062467600183663589340593884950808
136091389056251997893995894309660827763434071897306280320
151044063120296064809
```

```
(=0x94ecb72a 069a5322 e62d9357 c49d5664 1c351611 d1f361a8
cbb8a12c f410e821 4fbe8e02 8d85d404 399b4c7c 5a6a72ce
deef7b08 96302d5f),
```

where this representation is defined in Appendix K.5 and uses the mapping of Appendix K.3.3 with the default square root function and underlying isomorphic mapping between Ed448 and Curve448 of Appendix M.2.

This representation can also be expressed in tight LSB/lsb order as the pair ((u1,s1),(u2,s2)), where (s1,s2):=(1,1) and where

```
u1      799430080555285542466583392114886786202374259081179178887
990338902005327496428208435321295787094454554911799066625
85567756287085693163
```

```
(=0xd713005d bece883b de9e7077 e0084c74 e3f8ccf3 dcd9af2
2db99b77 5a9c3de7 c8d14433 634cee63 531d3d85 0637c24d
a28691a3 ac041438)
```

u2 273728972604711260959662149917071768586371733548553856048  
628325847723030459670661529224890730701519431099205367639  
437006368499972842925  
  
(=0xb5a46d1d be03f21b a4070e3c 51e42a50 1de9a4e6 3155b58c  
41dbdaed d5089539 cf69bbc8 78f3809d 5630ab65 c250e49b  
3a91a31d 067f1606),

where this uses the default completed mapping defined in Appendix K.6  
and the mapping of Appendix K.4.3 (with the default square root  
function).

### O.3. Example with Wei448

$Pw = (X, Y)$ ,  $k * Pw = (X1, Y1)$ , and  $(k+1) * Pw = (X2, Y2)$  with Wei448:

X 29070637261778856087396075817199998758219070555984737667402173284  
55389871077654193754799253725773241315783295429899652880118118204  
91344  
  
(=0x6663c64e 5b9a1f6d cbee3f5f 839b7dd8 6f53cc3e 0a01dab3  
e4a8314e 8d54be2a 4d63380b 692381bb ed7339fd dca7b0cd a80166fe  
18c15250).

Y 30578727850066757341435137807347775064915058999485530015946871157  
86794631407274936870618580714107931661730999350222644894729285604  
97149  
  
(=0x6bb38e82 8d52337f 6f0395ef dc16c776 52162f5e 309112ae  
fc7401bf 0cfb0499 ebled555 bf507ebc c33b4753 2d6dc6c5 d68dealc  
cle4clfd).

X1 40351504322781497250899987383866753965468971276834772118588405333  
77140867939355980788573436893357369201402928958042617224896092079  
46142  
  
(=0x8elf426a 4a1af133 ff970fe2 76693c7a eaa78786 361b1cfe  
4ccbd786 e020ba9a 0bf65a1d 5d9a128a f85c63a2 79a00139 7aca56db  
15341b9e).

Y1 55735504615964066386264989698774850924544182484936624265048483231  
35693859362627880184586282439234602798023594054611737412667543758  
11547  
  
(=0xc44e5e0f 2c254d23 ldc082db 77175e8c fd37793c 22ebe200  
77905a5f 750b3c9f 4a95d4d5 4e1a1e54 d2d31689 4249252d 0c8b1c45  
1c1481db).



X2 51724471386152414687122300763026650882740205909970876834920746101  
21508416303604179834986180511406702029983417676758930643822754281  
77944

(=0xb62dc975 470cc05b 082dae0b eabeldda 25487b2a 9663eec8  
f3bd3d0e eb724f44 dc53671c ffabb9ee 0c02aa74 b083cd82 a821a4cf  
6f6e5818).

Y2 57682103223585233918507344950062950306770296215271320612937204938  
77499282103483092990510136901415757273082719665657484294344333591  
20741

(=0xcb2988a4 6e37f9a9 a7a1255b 2fd2eea9 82308e7c eb8e18b8  
2252175f fd416a10 5984c6b8 36470e48 31879293 8f6139c6 f96164cb  
14010965).

The representation of  $k$  and the compressed representations of  $P_w$  and  $k*P_w$  in tight MSB/msb-order are given by

repr( $k$ ) =0xdc3bbb9 e42d7aca fe62052d 902123c7 0872b984 4c1e199f  
7c5d37bd 1171102b c20a6352 d9c91886 29b685de 51441e84  
3afe2665 5251aa80;

repr( $P_w$ ) =0x80 6663c64e 5b9a1f6d cbee3f5f 839b7dd8 6f53cc3e  
0a01dab3 e4a8314e 8d54be2a 4d63380b 692381bb ed7339fd  
dca7b0cd a80166fe 18c15250;

repr( $k*P_w$ ) =0x80 8elf426a 4a1af133 ff970fe2 76693c7a eaa78786  
361blcfe 4ccbd786 e020ba9a 0bf65ald 5d9a128a f85c63a2  
79a00139 7aca56db 15341b9e,

where the leftmost bit of the leftmost octet indicates the parity of the Y-coordinate of the point of Wei448 in question (which, in this case, are both one, since  $Y$  and  $Y_1$  are odd). See Appendix H.1 and Appendix I for further detail on (squeezed) point compression.

The scalar representation is consistent with the representations specified in [SEC1]; the (squeezed) point representation illustrated above is "new". For completeness, we include a SEC1-consistent representation of the point  $P_w$  in affine format and in compressed format below.

The SEC1-compliant affine representation of the point  $P_w$  in tight MSB/msb-order is given by

aff( $P_w$ ) =0x6663c64e 5b9a1f6d cbee3f5f 839b7dd8 6f53cc3e 0a01dab3  
e4a8314e 8d54be2a 4d63380b 692381bb ed7339fd dca7b0cd  
a80166fe 18c15250

```
6bb38e82 8d52337f 6f0395ef dc16c776 52162f5e 309112ae
fc7401bf 0cfb0499 eb1ed555 bf507ebc c33b4753 2d6dc6c5
d68dealc cle4clfd,
```

whereas the SECl-compliant compressed representation of the point  $P_w$  in tight MSB/msb-order is given by

```
compr(Pw)  =0x03 6663c64e 5b9a1f6d cbee3f5f 839b7dd8 6f53cc3e
0a01dab3 e4a8314e 8d54be2a 4d63380b 692381bb ed7339fd
dca7b0cd a80166fe 18c15250.
```

The SECl-compliant uncompressed format  $\text{aff}(P_w)$  of an affine point  $P_w$  corresponds to the right-concatenation of its X- and Y-coordinates, each in tight MSB/msb-order, prepended by the string 0x04, where the reverse procedure is uniquely defined, since elements of  $\text{GF}(p)$  have a unique fixed-size representation. The (squeezed) compressed format  $\text{repr}(P_w)$  corresponds to the SECl-compliant compressed format by extracting the parity bit  $t$  from the leftmost bit of the leftmost octet of  $\text{repr}(P_w)$ , and replacing this leftmost octet with 0x02 or 0x03, depending on whether  $t=0$  or  $t=1$ , respectively, where the reverse procedure is uniquely defined. For further details, see [SECl]. Note that, due to the bit-length of the prime  $p$ , the squeezed compressed format  $\text{repr}(P_w)$  and the SECl-compliant compressed format  $\text{compr}(P_w)$  have the same size.

A randomized representation  $(t_1, t_2)$  of the point  $k \cdot P_w$  in tight MSB/msb order is given by

```
t1      655783099225353926682910498535559663266263823350679216116
172951494291735730803127024621397533084891460609898061397
896825551162064841608

(=0xe6f93655 2765628b accfe61c 7dc6a594 e06fb243 70195ded
74d88a53 fdcdc2e8 077e0eff 62fa6a80 fa26b499 1f8796f5
21f2f03b f7e92b88)

t2      357918241879339174086992006475988394618511927120788596330
507910466738735762660894972854331591097934354210992993787
402433561014235472657

(=0x7e0ffcaf 7add27bc bb723629 95fdedd0 8769f676 78d953bc
0d38f4f6 d63a59dc 00f2d55a a4db7dab 16364503 591edcb1
e095a577 43dea311),
```

where this representation is defined in Appendix K.5 and uses the mapping of Appendix K.3.1 with the default square root function.

This representation can also be expressed in tight MSB/msb order as the pair  $((u_1, s_1), (u_2, s_2))$ , where  $(s_1, s_2) := (0, 0)$  and where

```

u1      276116573473684049599673971142041943002546018725744504858
        999210132924481156665376801365226215725437541502686055399
        974543995300346621026

        (=0x6140460c 1860a8cb 7c8ab942 b9509a84 95b4093c 95be5c8b
        df46e24c 069fe28a a23e4bfc 5bc29543 ee9ff503 febb80c8
        eb207253 8d7c6c62)

u2      128692595060487759871442054704123965938223087241863768179
        405512569340496286539849938457727539660932642464491037369
        291713756051590336193

        (=0x2d53abf2 370638a2 c2d38efe 718d0189 18d15d15 f132741b
        34405174 97fc0884 0c6be3a5 d9c201b9 cb0c3637 2674078e
        59ac8cd7 4f9fcec1),

```

where this uses the default completed mapping defined in Appendix K.6 and the mapping of Appendix K.4.1 (with the default square root function).

#### O.4. Example with Wei448.1

$Pw_1 = (X, Y)$ ,  $k * Pw_1 = (X_1, Y_1)$ , and  $(k+1) * Pw_1 = (X_2, Y_2)$  with Wei448.1:

```

X      41414505267302962826496323862800346730148184600706317030200831678
        13123337737005257876668389910719145841028692415431602235556184165
        13314

        (=0x91ddb90a 3c19f561 21de39ad a8c6bb00 579a6d2d 9ff6b810
        b109bf41 6e4e6227 0fc34010 be9ec68e 5ca11111 bc99e998 cff0f6db
        f4225122).

Y      21678703524693091005728527221124083240889481089231739678311939020
        43874709051080711177237887514058399787606848450432099149433728340
        08081

        (=0x4c5ac727 121de1f1 be917280 829a6d4c 9f615e3a 879a7dfd
        50f8bdcc 75d5856b 1d01ffaa 44e5ba0a ed0e341d 9383e15a 6cd48db8
        cle26c11).

X1     21211734920525001827254082557112140340208109740004519264558098189
        40985376833176210029490012696175276046779431389727351279961384020
        21113

```

```
(=0x4ab5bb5f ca80119b 6280f5d1 aec51745 23ab57ab 4d617195
38f453dd 2e8d9b66 a5417d1b ed0cee3d 4d6c84ca abdald41 b7a805dc
cbaefef9).
```

Y1 14152482531219571027190110620355502977165146571026919001455348108  
06142769037926777863731011790633441497896003632149582109867558046  
16181

```
(=0x31d8b337 09272016_25d2f9d6 cb0e2396 b7088c79 ffc8571f
6dc9bfe9 9e0783d4 1f684439 c02981f1 83f6696d 9c0377c9 431b8186
f503d5f5).
```

X2 30394319241133688143587947164786865078477223372122681434460686381  
23744153597949961703624604448300529949032402198106459850911229168  
46262

```
(=0x6b0d487d cba3633b_034f65a5 bbd1c8eb 1b6dcb1f 8d787db1
a581c08d ad23cbcd 6faa39d5 36731645 fd2fd6c0 03367bff 9093d29d
550d6ab6).
```

Y2 18866191129065867707969757296934620738822864945913956797432892866  
18725386530370846638505587040510045280940919798896557156654042590  
85719

```
(=0x4272da7b 7ad66918 144ae679 3811eb6b 2124b02f 42fd51f2
34e6f3ea 6285d40e 43cf726f 585b7e74 c4448acb b0c3ab89 d5a55678
c4622d97).
```

The representation of  $k$  and the compressed representations of  $Pw1$  and  $k*Pw1$  in tight MSB/msb-order are given by

```
repr(k)      =0xdc33bbb9 e42d7aca fe62052d 902123c7 0872b984 4c1e199f
              7c5d37bd 1171102b c20a6352 d9c91886 29b685de 51441e84
              3afe2665 5251aa80;
```

```
repr(Pw1)    =0x80 0x91ddb90a 3c19f561 21de39ad a8c6bb00 579a6d2d
              9ff6b810 b109bf41 6e4e6227 0fc34010 be9ec68e 5ca11111
              bc99e998 cff0f6db f4225122;
```

```
repr(k*Pw1)  =0x80 0x4ab5bb5f ca80119b 6280f5d1 aec51745 23ab57ab
              4d617195 38f453dd 2e8d9b66 a5417d1b ed0cee3d 4d6c84ca
              abdald41 b7a805dc cbaefef9,
```

where the leftmost bit of the leftmost octet indicates the parity of the Y-coordinate of the point of Wei448.1 in question (which, in this case, are both one, since Y and Y1 are odd). See Appendix H.1 and Appendix I for further detail on (squeezed) point compression.

A randomized representation  $(t1, t2)$  of the point  $k \cdot Pw1$  in tight MSB/msb order is given by

```
t1      303494474566270819668963081208440311422386279248346372989
        800906749888679443057479207554461646083343330145746687567
        323228377891922156528

        (=0x6ae4d2fc 57e63e5e bfdc44e6 5148d1bd b30b7c7b 2ca2a66a
        8a2bea6c 69113c79 7a4d6d0f 3c89b06a 3883ab2c e7d73f42
        24c82419 391e9bf0)

t2      637873534161581517938168102871523640780662020357386089328
        144426836947858617075256828298188817117945599296940030103
        858866119361786506090

        (=0xe0aa61c1 213a19b4 a9fddbb3 4c1377d0 4cd1fb84 017a1719
        e57b243b 31b13406 d5d77138 23c5a1b8 4fe271a5 2e53c98f
        900f2900 d1e76b6a),
```

where this representation is defined in Appendix K.5 and uses the mapping of Appendix K.3.1 with the default square root function.

This representation can also be expressed in tight MSB/msb order as the pair  $((u1, s1), (u2, s2))$ , where  $(s1, s2) := (1, 0)$  and where

```
u1      258036413119309433113527846878476684681744445436114935036
        372455666259396397921645423893888406553811930237985641251
        551672383206550397837

        (=0x5ae20fb6 5cafa07a 40421568 72419f49 dc31cbe9 766806f6
        8b1dbd7f 628c8ecf 10577848 e2e87ac2 fead0f09 6726ee34
        c2ed465f 5b7be38d)

u2      193962140052429320576140519455776109491178991023347646634
        723564200925012444187815484406230413980100291233975929881
        580671116555136082409

        (=0x4450c0ba ba9ee42a 4723b3b4 dbe7613f a78a2feb ee01752f
        9f8f51d6 41476eb8 041c9d87 d1b6df7b 9c6b48ad 2cdf4c20
        02d22f0c fbf521e9),
```

where this uses the default completed mapping defined in Appendix K.6 and the mapping of Appendix K.4.1 (with the default square root function).

## O.5. Example with Wei448.-3

$Pw3 = (X, Y)$ ,  $k * Pw3 = (X1, Y1)$ , and  $(k+1) * Pw3 = (X2, Y2)$  with Wei448.-3:

X 54121793865726175505902038600562190720650456678500106168173285986  
99999531708218763586616425010404811083912084906688745035466757984  
48968

(=0xbe9f5a23 51709e13 d5ad50c2 a27be8ee 1b051970 2580d5c3  
c2de7f75 3010635e d89ef547 8b67dc54 16d63c5b 1cc1116f dd453515  
71b39b48).

Y 14962282101304548030627835311887275833718070818965306362006934455  
59168773381983445709256615887526455657034051121085622763637035580  
12661

(=0x34b2dcc4 92d6a940 e6249c14 122d0ba4 5dc040e9 3f060d8f  
a65fa300 eb3cc969 25188b59 2d31039c f7a8e14a 48320a32 efe9b42b  
986afef5).

X1 18808295916646645825216065847266150404062470629833854840155953858  
63091795696773741607659794828181692381790403935750135247605982648  
6547

(=0x069fdd7c 2ec1ecbf d3cd0e27 1e8110c6 d2e478f2 aa393928  
64a5511e da0b8dc7 3834fd57 b5ef8527 361a8176 c6da44ee 63701c0c  
f49d7d13).

Y1 12212945244064471634326466576257313927639904273911210953487761656  
77684161144865373513143868308041748047828401098060667767703779846  
85920

(=0x2b03e68e b61581c4 9f977443 3e1ddc63 976f8f1d cdb185ee  
9c53328d b425973d 359bbc09 468645c4 0996a2c7 fda561be acb4d0b5  
745ab760).

X2 58672976485086436102048679093716482249296622848351051568512020319  
97872083950108489407370832733527154843728068195507632886574086695  
12670

(=0xcea6f66e e741e7b3 ee50acd4 bd6eachf 821fab72 bf5fe85b  
8f614af9 04aff677 15e820b9 e4bcc159 f67a97f3 2c176d2c d9b7cdeb  
f753f3de).

Y2 63661899992109030051219177516378471383513217472497460517936503629  
79522840238080543318627428149249774773108009447466292682661818280  
41265

```
(=0xe0394408 ed2b4efb b6b6ac7e bc815516 fdf31a6e d32db3f9
54cd8ac1 c7ddf0cc e7507688 a70f219a 57eef863 49003560 66747ca3
00105a31).
```

The representation of  $k$  and the compressed representations of  $Pw3$  and  $k*Pw3$  in tight MSB/msb-order are given by

```
repr(k)      =0xdc33bbb9 e42d7aca fe62052d 902123c7 0872b984 4c1e199f
              7c5d37bd 1171102b c20a6352 d9c91886 29b685de 51441e84
              3afe2665 5251aa80;

repr(Pw3)     =0x80 be9f5a23 51709e13 d5ad50c2 a27be8ee 1b051970
              2580d5c3 c2de7f75 3010635e d89ef547 8b67dc54 16d63c5b
              1cc1116f dd453515 71b39b48;

repr(k*Pw3)   =0x00 069fdd7c 2ec1ecbf d3cd0e27 1e8110c6 d2e478f2
              aa393928 64a5511e da0b8dc7 3834fd57 b5ef8527 361a8176
              c6da44ee 63701c0c f49d7d13,
```

where the leftmost bit of the leftmost octet indicates the parity of the Y-coordinate of the point of Wei448.-3 in question (which, in this case, are one and zero, respectively, since Y is odd and Y1 is even). See Appendix H.1 and Appendix I for further detail on (squeezed) point compression.

A randomized representation ( $t1$ ,  $t2$ ) of the point  $k*Pw3$  in tight MSB/msb order is given by

```
t1           450833060883286904091316612794941178576639837300736625958
              696097131313213727115363096930063001237631586932727905179
              306828042642854311987

              (=0x9ec9ba07 3fb2bb5e 9dbee995 067ce094 63601ecd 325f0930
              aea79cb8 745fa71d 4caa37ee f04fab67 ab2de747 4ac0a025
              830f4828 429cf833)

t2           339205723274519707955026734148022275762579914421865223818
              363622725164496136165251928391223173879522521195772276587
              373445978123589677750

              (=0x7778c1f9 9d900633 d161d7ea a963ddad e9101d3f f4f04710
              623d2a51 6ca10133 3db9ccc3 86df9271 fbb72740 77f79dd1
              9aed0bfb e3bc72b6),
```

where this representation is defined in Appendix K.5 and uses the mapping of Appendix K.3.1 with the default square root function.

This representation can also be expressed in tight MSB/msb order as the pair  $((u_1, s_1), (u_2, s_2))$ , where  $(s_1, s_2) := (0, 1)$  and where

```

u1      589255274721777493669102139212346422449226408440608788354
        266603544786997157375671957901717836941301424106139118763
        92799989153446639329

        (=0x14c11156 85eab1a5 f6c00d37 a3f6bd73 fe403dd1 31e337e7
        15927c25 0264a8f8 d2cd661e b5138468 92a3b91d 09284398
        17c2e361 96fa36e1)

u2      213991023129828413030692573508989139610229330687681826719
        574082317313789459478773972345123463766002343322541837566
        496527438452046182709

        (=0x4b5eac5a 3632b273 012a1050 7762eba4 8df1ccad 16dd9e6f
        d68e57a9 89de5a0c 1eda0951 e4f3de0e 39f5c37b 2f8f04d5
        52c093d8 fb983935),

```

where this uses the default completed mapping defined in Appendix K.6 and the mapping of Appendix K.4.1 (with the default square root function).

#### O.6. Example with Edwards448

$Pe_1 = (x, y)$ ,  $k * Pe_1 = (x_1, y_1)$ , and  $(k+1) * Pe_1 = (x_2, y_2)$  with Edwards448:

```

x      70320395893028961673046639985409870226249442701760956079298956688
        26896600999421897751877804946848997852325361659665744287620719558
        67733

        (=0xf7acf3ca b79b29c2 aa44863d 9edaeca4 8c90ad84 e460df42
        7dd9ab59 1bd8a844 07cb3419 59309b33 1e22bfa1 a2d37e10 e2e42a1f
        170f0855).

y      70628706854857281648863291487942166052137991441320055237644304464
        58787938273165391464653528929699350754224243613996187734424074211
        98773

        (=0xf8c2f181 3bceee8e 085ecd70 d1b6aa4c ea9b95bd 8f36ab44
        c79e9124 1ea625b7 f9f5ec57 89cc5af2 a2eb255a b252b874 509dc0d9
        685841b5).

x1     38125875041649701211705790554244713713134918749445854542272999596
        74058986304488795258334978838809456257721496105769894880185657328
        40277

```



```
(=0x864880b9 e1900c68 ba4a545a 6fe2b161 62dcc3b9 fa218e4b
feba9828 5cee5193 f2c989f6 c3b94eb6 2914dce7 b4818e4d 8fc8d51f
05a13355).
```

```
y1 11060653846610182753991162627427631707898421166839907726978369444
53337541552746428662176632660036639406375548888849623833963458813
1154
```

```
(=0x03e54af3 7f4cf5e6 5f1e2acd 5c4a4554 76adc652 b198ab2a
719e5aa9 ee749871 0193da82 ab6d000b f55836b1 0615653f 69514297
f4459f52).
```

```
x2 15620503788413497044804517304021524439062374489822547728508337937
50606335270276724725939683726318058744384611584731365019896485812
8760
```

```
(=0x05806f71 95e85352 ef3960ac 1ff9cf6c 3c99e0ee 2e75edfc
a133cafc 4a4b5fbf e4339859 c5fa123b 70ad2faf 7584ab9d 264540e7
7d560978).
```

```
y2 40019917514121727463122190125689377890703570698337158159153510836
68442386516751945577468473801561261386285902585868517988506010293
44096
```

```
(=0x8cf44811 3cec6e07 d1bbe9f5 4062075c 6fec0ac5 31272dce
1f446aeb d895373d e312c18d 6a345755 2861e014 0cc23158 a46ace4c
9ca21b60).
```

The representation of  $k$  and the compressed representations of  $Pe1$  and  $k*Pe1$  in tight LSB/lsb-order are given by

```
repr(k)      =0x01558a4a a6647f5c 2178228a 7ba16d94 6118939b 4ac65043
d4088e88 bdecha3e f9987832 219d4e10 e3c48409 b4a0467f
535eb427 9dddc3b;
```

```
repr(Pe1)    =0xad821a16 9b03b90a 2e1d4a4d 5aa4d745 4f5a3391 ea37af9f
eda46578 248979e3 22d56cf1 bda9d957 32556d8b 0eb37a10
717773dc 818f431f 01;
```

```
repr(k*Pe1)  =0x4af9a22f e9428a96 fca6a860 8d6claaf d000b6d5 415bc980
8e192e77 955a798e 54d5198d 4a63b56e 2aa2523a b35478fa
67af32fe cf52a7c0 01,
```

where the rightmost bit of the rightmost octet indicates the parity of the  $x$ -coordinate of the point of Edwards448 in question (which, in this case, are both one, since  $x$  and  $x1$  are odd). See Appendix H.3 and Appendix I for further detail on (squeezed) point compression.

The scalar representation and (squeezed) point representation illustrated above are fully consistent with the representations specified in [RFC8032]. Note that, contrary to [RFC7748], [RFC8032] requires unique representations of all elements of  $GF(p)$ .

A randomized representation  $(t_1, t_2)$  of the point  $k \cdot P_{el}$  in tight LSB/lsb order is given by

```
t1      125390048858887400104074787879402833851854739339836093733
        734638776755983021034212058415891288350265701101219981698
        849086128138510420407

        (=0xed921f3d 6ea4e452 dd06e783 782cbeb3 c5847a79 d9e6b993
        bd387cf5 feeddafa af8c038d f2732362 92724d37 273eedfc
        f2ab2499 98a79434)

t2      493324858478481242405018423865550638507715454654135514168
        842560149827360763382889199963980056979895918545280883247
        787003997982869314731

        (=0xd53a5125 193b6ab9 8db48161 20fb4865 02cf0546 3b48d8a6
        514af28f 43c026cb 0f2ff3d5 e558bb03 4b833cd1 1ca710cc
        9bf0c2a3 351083b5),
```

where this representation is defined in Appendix K.5 and uses the mapping of Appendix K.3.3 with the default square root function and underlying 4-isogenous mapping between Edwards448 and Curve448 of Appendix N.2.

This representation can also be expressed in tight LSB/lsb order as the pair  $((u_1, s_1), (u_2, s_2))$ , where  $(s_1, s_2) := (0, 0)$  and where

```
u1      135993582308059710871118067705651831584215992415511174727
        255533641033816319052989477276487981998957706382391254504
        484510842833065141388

        (=0x31276f5f d399d1cd 5d18c46a eba5388f 93effaf7 9574b23b
        ce34ba45 5050c160 477ae803 9c3112be 596281a7 b7ae4da6
        e9dd7688 191fa7f4)

u2      300725936379847215929002275525633229576034707671620463143
        626393832660436027759737097637786753095880885199368686863
        187789449179730426477

        (=0xb65c5ee8 597b5b55 a87e266f b9c1f5cb 5d224ec3 8fb22f32
        b0378e70 47ecc389 9585b06e 7fb4f70b 38a3b453 ab5c03d8
        37b5093b 9a4cd796),
```

where this uses the default completed mapping defined in Appendix K.6 and the mapping of Appendix K.4.3 (with the default square root function) and underlying 4-isogenous mapping between Edwards448 and Curve448 of Appendix N.2.

#### Appendix P. Random Integers in $\mathbb{Z}_n$

Any probability distribution on the interval  $[0, N-1]$  can be converted to a probability distribution on  $[0, n-1]$ , via a suitable function that maps inputs from the source distribution  $[0, N-1]$  to values in the interval  $[0, n-1]$ . We consider three such functions, each with the property that if the source distribution on  $[0, N-1]$  is statistically close to the uniform distribution, then so is the output distribution on  $[0, n-1]$ . (Here, we assume  $n$  and  $N$  to be integers of cryptographic interest, so large.) In practical applications, one can use these functions to convert the output of a cryptographically strong random bit generator (where  $N$  is a power of two and after conversion of the random bit string to an integer via the BS2I mapping of Appendix I.2) to a pseudo-random integer in the interval  $[0, n-1]$ , where the bias is small if  $N$  is suitably picked.

We consider mappings that convert an output of the source distribution to an integer in the interval  $[0, n-1]$  via modular reduction (Appendix P.1), via scaling (Appendix P.2), or via a membership test (Appendix P.3). For suitably picked  $N$  values and not too poor source distributions, the first two mappings never fail and any bias introduced by the conversion process can be made negligible in practice, while the third mapping (if it does not fail) inflates the bias by a small factor only in practice. (For details, see the remarks following each of the mappings below.)

NOTE: Each of the mappings below may yield a zero output value. One can modify each such mapping to always yield nonzero outputs, by setting output  $x$  to 1 if the original mapping would yield  $x=0$  for a specific input  $y$  and leaving the mapping the same otherwise (henceforth called the modified conversion function). This modification has negligible impact on the bias and does yield a conversion function to integers in the interval  $[1, n-1]$ . A similar remark applies if  $n=h*n_1$ , where  $h$  is a small integer: in that case, one can locally modify each mapping to always yield outputs in the interval  $[0, h*n_1-1]$  that are not divisible by  $n_1$ , simply by setting output  $x$  to  $x+1$  if the original mapping would otherwise yield  $x=0 \pmod{n_1}$ . (Notice that both modifications coincide if  $h=1$ .) These modifications may be useful if one wishes to generate integers in an interval of size  $n$  and where one wishes to avoid specific output values (e.g., if one wishes to generate high-order points of a curve of order  $h*n_1$ , with co-factor  $h$  (see Appendix B.1)). For simplicity,

we again refer to this as "the" modified conversion function (or h-modified conversion function, if h is not clear from context).

#### P.1. Conversion to Integers in $\mathbb{Z}_n$ via Modular Reduction

This function maps each integer  $y$  in the interval  $[0, N-1]$  to its remainder modulo  $n$ , i.e.,  $y$  is mapped to  $x := y \pmod{n}$ .

One can show that the bias introduced by this conversion function is at most  $\epsilon := 2\rho(1-\rho)/(N/n)$ , where  $r := N \pmod{n}$  and where  $\rho := r/n$ . Details are out of scope.

Note that if  $n$  does not divide  $N$ , this invariably introduces some bias, no matter the quality of the source distribution. In particular, the statistical distance of the distribution on  $\mathbb{Z}_n$  can be much larger than the statistical distance of the source distribution on  $\mathbb{Z}_N$ , since the bias introduced by the modular reduction step may be significantly larger than the bias of the source distribution on  $\mathbb{Z}_N$  if the value  $\rho$  above is not close to zero or one and if  $n/N$  is not sufficiently small. The maximum bias is, however, easy to determine from  $n$  and  $N$ . In particular, if the bit-length of  $N$  is sufficiently larger than that of  $n$ , the bias introduced by the modular reduction operation is negligible in practice. The same holds if  $N$  is close to a multiple of  $n$  (e.g., if  $n$  is close to a power of two and the input distribution is generated by a high-quality random bit generator with outputs of fixed bit-length).

Note: In practice, one does not determine the maximum bias  $\epsilon$  from  $n$  and  $N$ , but rather specifies a required upper bound (usually set to a value at most  $2^{-64}$ ) for  $\epsilon$  and subsequently determines the minimal value of  $N$  (where  $N$  is a power of two) for which this upper bound indeed applies, as a function of  $n$ . Table 4 illustrates this for several curves of practical interest.

Curve	$\text{eps0}=2^{\{-64\}}$	$\text{eps0}=2^{\{-100\}}$
NIST P-224 [FIPS-186-4]	224	224
NIST P-256 [FIPS-186-4]	288	352
NIST P-384 [FIPS-186-4]	384	384
NIST P-521 [FIPS-186-4]	521	521
brainpoolP224r1 [RFC5639]	287	323
brainpoolP256r1 [RFC5639]	319	354
brainpoolP320r1 [RFC5639]	379	417
brainpoolP384r1 [RFC5639]	445	482
brainpoolP512r1 [RFC5639]	575	608
Curve25519 [RFC7748]	252	252
Wei25519 [Appendix E.3]	252	252
Wei25519.2 [Appendix G.3]	252	252
Wei25519.-3 [Appendix G.3]	252	252
Curve448 [RFC7748]	446	446
Wei448 [Appendix M.3]	446	446
Wei448.1 [Appendix N.3]	446	446
Wei448.-3 [Appendix N.3]	446	446
secp256k1.m [Appendix L.3]	256	256

Table 4: Minimum value of  $m$  for which the bias (epsilon) introduced by converting integers in  $\mathbb{Z}_N$ , where  $N:=2^m$ , to integers in  $\mathbb{Z}_n$  via modular reduction or via scaling is lower than the indicated  $\text{eps0}$  value, for some curves of practical interest (where  $n$  is the order of the base point of the curve in question).

## P.2. Conversion to Integers in $\mathbb{Z}_n$ via Scaling

This function maps each integer  $y$  in the interval  $[0, N-1]$  to the integer  $x := \text{floor}(n*y/N)$ , where the floor function rounds real numbers downwards to an integer (i.e.,  $\text{floor}(z)$  is the unique integer  $i$  for which  $z$  is an element of the interval  $[i, i+1)$  of real numbers).

One can show that the bias introduced by this conversion function is at most  $\text{epsilon} := 2*\rho*(1-\rho)/(N/n)$ , where  $r := N \pmod n$  and where  $\rho := r/n$ . Details are out of scope.

The same remarks as in Appendix P.1 apply.

Note: this mapping corresponds to interpolation on the line with endpoints  $(0,0)$  and  $(N,n)$ , where values are truncated to integers. The division operation in this conversion function reduces to a binary string truncation operation if  $N$  is a power of two (which is often the case in practice). See also [comm-FIPS-186-5], pp. 80-82.

### P.3. Conversion to Integers in $Z_n$ via the Discard Method

This function (defined for  $N$  at least  $n$ ) is the identity map on the interval  $[0, n-1]$  and fails for each integer  $y$  outside this interval.

One can show that the statistical distance of the distribution on  $Z_n$  is at most roughly  $N/n$  times as large as the statistical distance of the source distribution on  $Z_N$  (if the latter is relatively negligible compared to  $n/N$ ). Details are out of scope.

Note that, under the above conditions, if  $N=2^m$  and if  $n$  has bit-length  $m$ , this conversion function fails with probability  $1 - n/N$  (which is at most  $1/2$ ) and, if it succeeds, does not inflate the statistical distance by more than (roughly) a factor two.

### Appendix Q. ECDSA signatures

The ECDSA signature scheme is specified in FIPS Pub 186-4 [FIPS-186-4], ANSI X9.62-2005 [ANSI-X9.62], SEC 1 [SEC1], and many other standards and can be instantiated with suitable combinations of short-Weierstrass curves and hash functions (that satisfy particular cryptographic criteria). Despite its wide-spread use, some details seem less well-understood. We, therefore, provide a concise specification of ECDSA (for short-Weierstrass curves defined over a prime field  $GF(p)$ ) and give some examples of ECDSA computations where the underlying short-Weierstrass curve has co-factor  $h>1$  and where the bit-length of the domain parameter  $n$  differs from the digest size of the used hash function, illustrated with the curves Wei25519 and Wei448 introduced in this document. Our description is consistent with all forementioned standards.

The signing operation takes as inputs a message  $m$  (represented as a bit string) and a private key  $d$  in the interval  $[1, n-1]$  and produces as output a signature, which is an ordered pair  $(r, s)$  of integers in the interval  $[1, n-1]$ , where  $n$  is the order of the base point  $G$  of the curve in question. The signature verification operation takes as inputs a message  $m$ , a public key  $Q$ , and a signature  $(r, s)$  and produces as output the value "valid" or "invalid", depending upon whether the message was purportedly signed by a holder of the private key of the public-private key pair  $(d, Q)$  for the curve used with the signature scheme. Full details are provided below, where we denote the applicable hash function by  $H$ .

#### Q.1. ECDSA Signing Operation

The signing operation involves the following steps:

- a. Generate a random ephemeral public-private key pair  $(k, R:=k*G)$ , by generating a random integer  $k$  in the interval  $[1, n-1]$  and computing  $R:=k*G$  (see, e.g., Appendix B.1);
- b. Compute  $k_1 := (1/k) \pmod n$  (see, e.g., NOTE 1 of Appendix K.2);
- c. Set  $x_R$  to the x-coordinate of the (affine) point  $R$ , convert this element of the field  $GF(p)$  to the integer  $r_0$  in the interval  $[0, p-1]$ , and set  $r := r_0 \pmod n$ , where  $x_R$  is converted to  $r_0$  by subsequently using the FE2OS and OS2I mappings of Appendix I.5 and Appendix I.3, respectively;
- d. Compute the hash value  $E := H(m)$  according to the applicable hash function  $H$ , where  $E$  is a bit string of length  $\text{hashlen}$  (the digest size of  $H$ );
- e. Represent  $E$  as the integer  $e$  in the interval  $[0, 2^l-1]$ , where  $e$  is the integer representation of the  $l$ -prefix of  $E$ , using the BS2I mapping of Appendix I.2, and where  $l$  is the bit-length of  $n$ . (For a definition of the  $l$ -prefix, see Appendix I.1);
- f. Compute  $s := k_1 * (e + r * d) \pmod n$ . Securely destroy  $k$  and  $k_1$ ;
- g. Return to the first step (Step a) if  $r$  and  $s$  are not both integers in the interval  $[1, n-1]$ ;
- h. Output the ordered pair  $(r, s)$  as the signature.

## Q.2. ECDSA Verification Operation

The verification operation involves the following steps:

- a. Check that the purported signer's public key  $Q$  is a point of the curve in question of order  $n$  (and output "reject" if this is not the case);
- b. Check that the coordinates of the purported signature  $(r, s)$  are both integers in the interval  $[1, n-1]$  (and output "reject" if this is not the case);
- c. Compute the hash value  $E := H(m)$  according to the applicable hash function  $H$ , where  $E$  is a bit string of length  $\text{hashlen}$  (the digest size of  $H$ );
- d. Represent  $E$  as the integer  $e$  in the interval  $[0, 2^l-1]$ , where  $e$  is the integer representation of the  $l$ -prefix of  $E$ , using the BS2I mapping of Appendix I.2, and where  $l$  is the bit-length of  $n$ . (For a definition of the  $l$ -prefix, see Appendix I.1);

- e. Compute  $s_1 := (1/s) \pmod n$  (see, e.g., Appendix K.2); compute  $u := e*s_1 \pmod n$  and  $v := r*s_1 \pmod n$ ;
- f. Compute the point  $R' := u*G + v*Q$ . Check whether  $R'$  is the identity element  $O$  of the curve (and output "reject" if this is the case);
- g. Set  $xR'$  to the x-coordinate of the (affine) point  $R'$ , convert this element of the field  $GF(p)$  to the integer  $r_0'$  in the interval  $[0, p-1]$ , and set  $r' := r_0' \pmod n$ , where  $xR'$  is converted to  $r_0'$  by subsequently using the FE2OS and OS2I mappings of Appendix I.5 and Appendix I.3, respectively;
- h. Output "accept" if  $r' = r$ ; output "reject" otherwise.

NOTE 1: For prime-order curves,  $r$  generally uniquely represents the x-coordinate of  $R$  (since, by the Hasse bound,  $|E| = n$  is relatively close to  $p$ ). For curves with co-factor  $h > 1$ , this result holds only if one would know  $r_0 \pmod{h*n}$ , rather than  $r := r_0 \pmod n$ .

NOTE 2: If an ECDSA signature  $(r, s)$  is valid for a particular message  $m$  and public key  $Q$ , then so is  $(r, -s)$  -- the so-called malleability. Note that this corresponds to changing the ephemeral signing key pair  $(k, R)$  in the first step of the signing operation to  $(-k, -R)$ , where the y-coordinates of  $R := (x_R, y_R)$  and  $-R = (x_R, -y_R)$  have different parity (see Appendix H). Since any party (not just the signer) can recompute the ephemeral signing key  $R'$  from a valid signature, since  $R' := (1/s)(e*G + r*Q)$ , this implies that any party can retroactively put the ECDSA signature in a form where the y-coordinate of the ephemeral signing key  $R$  has a fixed parity. This observation can be used to put ECDSA signatures in a form that generally allows unique and efficient recovery of  $R$  from  $r$  for prime-order curves (due to NOTE 1 above) and more efficient signature verification methods. Further details are out of scope.

### Q.3. Representation Examples ECDSA

We present some examples of ECDSA computations, when used with curve Wei25519 and SHA-256 (see Appendix Q.3.1), with Wei25519 and SHAKE128 with output size  $d_0 = 256$  (see Appendix Q.3.2), and with Wei448 and SHAKE256 with output size  $d_0 = 512$  (see Appendix Q.3.3). In each case, we indicate the signer's public key  $Q := d*G$ , the ephemeral signing key  $R := k*G$ , the message  $m$  that is signed, and some intermediate values in the ECDSA signing operation resulting in signature  $(r, s)$ . We write  $R := (x_R, y_R)$  and  $Q := (x_Q, y_Q)$ , and include the ASCII string corresponding to message  $m$ . Note that the domain parameter  $n$  of curve Wei25519 has bit-length  $l := 253$ , whereas the corresponding domain parameter for Wei448 has bit-length  $l := 446$  (which, in either case, differs from the digest size of the underlying hash function).



For completeness, we also provide an example of ECDSA computations, when used with NIST curve P-256 (see [FIPS-186-4]) and hash function SHA-256 (see [FIPS-180-4]), where the domain parameter  $n$  of the curve in question has bit-length  $l:=256$ , which is equal to the digest size of the underlying hash function (see Appendix Q.3.4). For details of the encoding of ASCII strings as bit strings, see Appendix I.4.

Q.3.1. Example of ECDSA with Wei25519 and SHA-256

d 47941274660029138864396347947568908774951195017212284524777080461  
79444885588

(=0x0a996146 d73d096f 6a606ad8 72e11b12 ce973033 524591c3  
ebcc630d b6368854).

xQ 34422557393689369648095312405803933433606568476197477554293337733  
87341283644

(=0x079c3f69 9b688181 69038c35 39c11eb5 96d09f5b 12a242b4  
ce660f13 3368c13c).

yQ 76981661982917351630937517222412729130882368858134322156485762195  
67913357634

(=0x110501f6 1dff511e d6c4e9b9 bfd5acbe 8bf043b8 c3e381dd  
f5771306 479ad142).

k 17426547602876470587191777825317027698752636279275919375559360929  
53735113209

(=0x03da4ec1 8dc83b53 5ab8857c bbd289ae 40e6d25b ba52923c  
e6b217a0 348ca9f9).

xR 38236544880946097675798638032669186189501319930946799635186226253  
710117141679

(=0x54891e12 88cf078e f3f1444c c1919e30 67eb5dd6 1c6f45d1  
94b9c0e1 192d7caf).

yR 24120175139256121256267158437786975197587143475570212981221664791  
614551611968

(=0x3553890b d265d561 032e2daa 10b9820c 4845dbf8 f6b4f432  
08f5df99 c375da40).

r 20515169942847866059327052174542149852157381340472616051764715622  
82845886734

(=0x04891e12 88cf078e f3f1444c c1919e2f ff907c7c ed9935a1 dc5dd15d 4860590e).

1/k 41122695303709273156068243481808769134600808188172269288861174824 34446546266

(=0x0917764a 5a76024b e9608472 bfec99be 0cffacbe 0a5a6805 0e4e75bc 36a0d55a).

m "example ECDSA w/ Wei25519 and SHA-256"

(=0x65 78616d70 6c652045 43445341 20772f20 57656932 35353139 20616e64 20534841 2d323536).

E 10340924651306471157182528854495725311608440786255119926874295925 4624066081637

(=0xe49f8f34 0ac7fd87 1ca6c035 1ac83b97 2ec4711e f4a79d37 214b6b94 c6f41365).

e 12926155814133088946478161068119656639510550982818899908592869906 828008260204

(=0x1c93f1e6 8158ffb0 e394d806 a3590772 e5d88e23 de94f3a6 e4296d72 98de826c).

s 18145968192643101430203980459406244543409512911444833316246990876 74236833451

(=0x04030680 d490837e 0b50800d 5052feb3 8181da43 f14fea65 d75fff8e 095d8eab).

The ECDSA signature (r,s) can be represented uniquely as the 64-octet string

0x04891e12 88cf078e f3f1444c c1919e2f ff907c7c ed9935a1 dc5dd15d 4860590e

04030680 d490837e 0b50800d 5052feb3 8181da43 f14fea65 d75fff8e 095d8eab,

where this string is the right-concatenation of the integers r and s, each represented as fixed-size octet string in tight MSB/msb-order using the ZnE2OS mapping of Appendix I.6. Since an ECDSA signature (r, s) is valid only if the ECDSA signature (r,-s) is, one can alternatively use the representation

0x04891e12 88cf078e f3f1444c c1919e2f ff907c7c ed9935a1 dc5dd15d  
4860590e

0bfcf97f 2b6f7c81 f4af7ff2 afad014c 935dlf9a bla7b270 80b2638c  
53984542,

with the same representation conventions.

#### Q.3.2. Example of ECDSA with Wei25519 and SHAKE128

d 50032130580855419870069268521079636534051105694026315073511374709  
23129445444

(=0x0b0fb7de 7b857528 c16cc691 f91acb6a e6f83700 c2257210  
d9ce4a66 540f5c44).

xQ 49674872575618115649605301860097524739691386255387989689284412105  
715250815836

(=0x6dd2fb44 ebc47199 0558875c 338b32a0 01c04e5e 54b0239f  
931ba404 43fee35c).

yQ 19668752079014976246249662506722644231308019872013845936101364656  
882653051514

(=0x2b7c1e81 e0d7311a 7e73c581 ac8d7478 f5d8402e a25ecf03  
2fcf49b3 ebe3ba7a).

k 67458228593538039868031175183537823353427877783158546151245140204  
51058711301

(=0x0eea001c 69e39d65 a93a736f 51dab17d 3c89d712 67b95dba  
28f43e6c 6d73fb05).

xR 22710793528316744414502819712682283876956423576126122262984645007  
656889457787

(=0x3235da86 6c184868 db1060f4 c57414ba f9dd8bbf af94eb8e  
65a26fa8 146d9c7b).

yR 48228386115947942380117850340406514077008333836380715701663219971  
594920954196

(=0x6aa04c98 30a51d5a 226fc67b 6ec00aa4 66eae465 432825e3  
c8da192d 330c8954).

r 99977679631995777258326002355330115438507449798639944497879219280  
0526704820

(=0x0235da86 6c184868 db1060f4 c57414ba bb409e23 c6ae150b 5d6b4658 fd8c20b4).

1/k 16237902548817115200666748510759761693156732885271500846541777492 82633956147

(=0x03970860 022244d0 1cee5f2e 973372d7 2000b51d 2d75731c 0e27428a 7e723b33).

m "example ECDSA w/ Wei25519 and SHAKE128"

(=0x6578 616d706c 65204543 44534120 772f2057 65693235 35313920 616e6420 5348414b 45313238).

E 37558481186175098606278970911021056916472038320089875122503537502 754552388537

(=0x530958d6 432ba571 6a20fd9b d3592234 943da1d6 57f55f07 2ab01860 3f9f9bb9).

e 46948101482718873257848713638776321145590047900112343903129421878 44319048567

(=0x0a612b1a c86574ae 2d441fb3 7a6b2446 9287b43a cafeabe0 e556030c 07f3f377).

s 60979034974035506260645462098255401877898928730177415844489376261 15704732698

(=0x0d7b4a83 96b37670 d6ef4ac7 6ce69d43 a65859de 4ecbe649 f56a7a1f 7fb5bcla).

The ECDSA signature (r,s) can be represented uniquely as the 64-octet string

0x0235da86 6c184868 db1060f4 c57414ba bb409e23 c6ae150b 5d6b4658 fd8c20b4

0d7b4a83 96b37670 d6ef4ac7 6ce69d43 a65859de 4ecbe649 f56a7a1f 7fb5bcla,

where this string is the right-concatenation of the integers r and s, each represented as fixed-size octet string in tight MSB/msb-order using the ZnE2OS mapping of Appendix I.6. Since an ECDSA signature (r, s) is valid only if the ECDSA signature (r,-s) is, one can alternatively use the representation

0x0235da86 6c184868 db1060f4 c57414ba bb409e23 c6ae150b 5d6b4658  
fd8c20b4

0284b57c 694c898f 2910b538 931962bc 6e86a000 542bb68c 62a7e8fa  
dd4017d3,

with the same representation conventions.

### Q.3.3. Example of ECDSA with Wei448 and SHAKE256

d 83773921833883065724152755040779926324701042667680137762329241115  
92597160376444120699241862910141955866217224630560765595890572227  
9690

(=0x1d818b12 92af6ef4 3f0ed657 b55d2ab7 a0cd1e64 516414d1  
d32ea610 dd6dddbe af65bc96 df648e6d fac1b907 6588b37e 984d5860  
7390970a).

xQ 40351504322781497250899987383866753965468971276834772118588405333  
77140867939355980788573436893357369201402928958042617224896092079  
46142

(=0x8elf426a 4a1af133 ff970fe2 76693c7a eaa78786 361b1cfe  
4ccbd786 e020ba9a 0bf65ald 5d9a128a f85c63a2 79a00139 7aca56db  
15341b9e).

yQ 55735504615964066386264989698774850924544182484936624265048483231  
35693859362627880184586282439234602798023594054611737412667543758  
11547

(=0xc44e5e0f 2c254d23 1dc082db 77175e8c fd37793c 22ebe200  
77905a5f 750b3c9f 4a95d4d5 4e1a1e54 d2d31689 4249252d 0c8b1c45  
1c1481db).

k 56463034235306169014882307562036113095966844917631298686749571574  
22895909756933115614724351575144190884397720504249121444938140865  
3424

(=0x13e308f8 2f7eb169 78a86240 a2087c59 38ad954c 5a725311  
00e2738b 93f87064 06846d1b 0348c213 5cd8f9db 21cbf970 6b70fa40  
29364070).

xR 46421117529223435940590399200091023258880155395346929342228475577  
87411917154572694868891187346300643187653728654052509827159201295  
60118

(=0xa37ff9a3 9734a3dc 9ccf72b7 cb3b8e5e 20d4e1f1 655c973a  
c72e4aa4 6f139529 84b1cd37 2524bf09 c4e38684 5c88cc79 e8e19242  
42398e36)).

yR 48450878695819342796480063527087959345962966106444727188216313803  
37436540801561730584163096514114057681225129685101546366763700225  
61560

(=0xaaa6202c df8711b2 6e5a8802 6c5d86b3 2f320d89 8f48a809  
40818982 bb74e0cc 7b884f20 aad090fb 90c4c93f fd84ed56 c03451d8  
84fc7718)).

r 10079181314443091413124208805690796541198087360981026328153965618  
84491837923780980544976081512453123921447854472219263731684084102  
60560

(=0x237ff9a3 9734a3dc 9ccf72b7 cb3b8e5e 20d4e1f1 655c973a  
c72e4aa5 757f4d55 fc1416a3 c77851e9 820a019f 40fdadcf a1f00dlc  
eb890450)).

1/k 13511362508598651506450197334516130806445911047753884276726477993  
82054003440714897722657048821186399503939251111689038388764827779  
24830

(=0x2f96a107 4a355722 1f20fd90 aed12db3 83b3c32f 593079f4  
779e2942 3ad2b5e6 0ea15bdc a57e5827 04ed1f09 e42b8352 68428208  
502444de)).

m "example ECDSA w/ Wei448 and SHAKE256"

(=0x6578616d 706c6520 45434453 4120772f 20576569 34343820  
616e6420 5348414b 45323536)).

E 70518111636481318756745253634149479528712660170653218748970032198  
09200302878838207999610865248047293428251821985248449616335015355  
074462070358583196774165

(=0x86a488f3 62da4be6 147ef640 34ad43ca 3fde6613 456cc034  
55555f34 c34778b9 02f05145 62e2113c f4894daf 4446bb10 636b43df  
fa5e2434 b1262bee 420fef15)).

e 95569862294810470138539721873942452276898897320233558414699702549  
60123138804500061287999759232209252819044656295606728178066107449  
5757

(=0x21a9223c d8b692f9 851fbd90 0d2b50f2 8ff79984 d15b300d  
155557cd 30dlde2e 40bc1451 58b8844f 3d22536b dl11aec4 18dad0f7  
fe97890d)).

s 15256839162738057100463520332129958798673106244466764276882516437  
65316731861037119394561858696609193031230290965981228906529672536  
92957

(=0x35bc73be 7e3fd7a1 de9fdb4f be96cbcd 0bb9beec f8286d04  
026f3440 d6e68e25 9916ea2b c4407e9a 83ecf91a 8473c1a1 cda742d0  
dab5d21d).

The ECDSA signature (r,s) can be represented uniquely as the  
112-octet string

0x237ff9a3 9734a3dc 9ccf72b7 cb3b8e5e 20d4e1f1 655c973a c72e4aa5  
757f4d55 fc1416a3 c77851e9 820a019f 40fdadcf a1f00dlc eb890450

35bc73be 7e3fd7a1 de9fdb4f be96cbcd 0bb9beec f8286d04 026f3440  
d6e68e25 9916ea2b c4407e9a 83ecf91a 8473c1a1 cda742d0 dab5d21d,

where this string is the right-concatenation of the integers r and s,  
each represented as fixed-size octet string in tight MSB/msb-order  
using the ZnE20S mapping of Appendix I.6. Since an ECDSA signature  
(r, s) is valid only if the ECDSA signature (r,-s) is, one can  
alternatively use the representation

0x237ff9a3 9734a3dc 9ccf72b7 cb3b8e5e 20d4e1f1 655c973a c72e4aa5  
757f4d55 fc1416a3 c77851e9 820a019f 40fdadcf a1f00dlc eb890450

0a438c41 81c0285e 216024b0 41693432 f4464113 07d792fb fd90cbbe  
a5e395c4 2b37f11d ea95b7f5 9d7fc958 0951cdb3 55d17fc1 d0a272d6,

with the same representation conventions.

#### Q.3.4. Example of ECDSA with P-256 and SHA-256

d 64502400493437371358766275827725703314178640739253280897215993954  
599262549170

(=0x8e9b109e 719098bf 980487df 1f5d77e9 cb29606e bed2263b  
5f57c213 df84f4b2).

xQ 57807358241436249728379122087876380298924820027722995515715270765  
240753673285

(=0x7fcdce27 70f6c45d 4183cbee 6fdb4b7b 58073335 7be9ef13  
bacf6e3c 7bd15445).

yQ 90436541859143682268950424386863654389577770182238183823381687388  
274600502701

- (=0xc7f144cd 1bbd9b7e 872cdfed b9eeb9f4 b3695d6e a90b24ad 8a462328 8588e5ad) .
- k 10917316901614856459075743461503453401956315994436627614714069472 2195121475236
- (=0xf15dd2ec 0clf92a1 0b60543c 20ccc85a 6bc502fc c8dlfa0f cc4e0efd 7e5b8ea4) .
- xR 67018809247931167566425050558236675606398890455759429072257638117 88903589477
- (=0x0ed12153 79636c48 3c2f7f15 5807d402 a3b22803 3af97c7e 17819ac3 169ea665) .
- yR 83689992559222976435347984539787632719385922964918301228608860125 842192869397
- (=0xb906db6f 843c3944 597cf51b b804f66b f66690df 75f9e046 d5d4f548 870a2815) .
- r 67018809247931167566425050558236675606398890455759429072257638117 88903589477
- (=0x0ed12153 79636c48 3c2f7f15 5807d402 a3b22803 3af97c7e 17819ac3 169ea665) .
- 1/k 11325557121201586890872285956108022533871237299549226886458455251 7263129747627
- (=0xfa6461b5 646b7d69 893c8e10 96376336 078815e5 9c225d70 f677307c 653ea0ab) .
- m "eyJhbGciOiJFUzI1NiJ9.eyJpc3MiOiJqb2UiLA0KICJleHAiOiJzMDA4MTkzODAsdQogImh0dHA6Ly9leGFtcGxlImlnVbS9pc19yb290Ijp0cnVlfQ"
- (=0x65794a 68624763 694f694a 46557a49 314e694a 392e6579 4a706333 4d694f69 4a716232 55694c41 304b4943 4a6c6548 41694f6a 457a4d44 41344d54 6b7a4f44 41734451 6f67496d 68306448 41364c79 396c6547 46746347 786c4c6d 4e766253 39706331 39796232 3930496a 7030636e 566c6651) .
- E 15276956252456799293212411937089506709811039436516676830575291628 425946659374
- (=0x21c67368 f436577f 447f8051 62ca13b8 0d046a3f e467247e 65ea477a a750fa2e) .



e 15276956252456799293212411937089506709811039436516676830575291628  
425946659374

(=0x21c67368 f436577f 447f8051 62ca13b8 0d046a3f e467247e  
65ea477a a750fa2e).

s 89123353657093021477366684784932901580138243670089627582817239001  
914975409365

(=0xc50a07d3 8c3c70e5 d8f12daf 084a5480 a66590c5 f293509a  
8f3f7f8a 83a354d5).

The ECDSA signature (r,s) can be represented uniquely as the 64-octet string

0x0ed12153 79636c48 3c2f7f15 5807d402 a3b22803 3af97c7e 17819ac3  
169ea665

c50a07d3 8c3c70e5 d8f12daf 084a5480 a66590c5 f293509a 8f3f7f8a  
83a354d5,

where this string is the right-concatenation of the integers r and s, each represented as fixed-size octet string in tight MSB/msb-order using the ZnE2OS mapping of Appendix I.6. Since an ECDSA signature (r, s) is valid only if the ECDSA signature (r,-s) is, one can alternatively use the representation

0x0ed12153 79636c48 3c2f7f15 5807d402 a3b22803 3af97c7e 17819ac3  
169ea665

3af5f82b 73c38f1b 270ed250 f7b5ab7f 168169e7 b4844dea 647a4b38  
78bfd07c,

with the same representation conventions.

NOTE: The example above corresponds to the JSON Web Signature example of Appendix A.3 of [RFC7515], where the base64url encoding of the public-private key pair (d, Q:=d\*G) above is given by

d "jpsQnnGQmL-YBIffH1l36cspYG6-0iY7XlfCE9-E9LI";

xQ "f83OJ3D2xF1Bg8vub9tLe1gHMzV76e8Tus9uPHvRVEU";

yQ "x\_FEzRu9m36HLN\_tue659LNpXW6pCyStikYjKIWI5a0",

where the base64url encoding of the ECDSA signature (r,s) is given by

```
"DtEhU3ljbEg8L38VWafUAqOyKAM6-Xx-  
F4GawxaepmXFCgfTjDxw5djxLa8ISlSApmWQxfKTUJqPP3-Kg6NU1Q",
```

and where the base64url encoding of the alternative ECDSA signature (r,-s) is given by

```
"DtEhU3ljbEg8L38VWafUAqOyKAM6-Xx-  
F4GawxaepmU69fgrc8OPGyc00lD3tat_FoFp57SETepkeks4eL_QfA".
```

Here, the JSON Web Signature is represented as the string m.sig, where the message field m and the signature field sig are separated by the "." symbol, where m assumes the value in our example and where the signature assumes the base64url encoding of one of the ECDSA signatures (r,s) or (r,-s), as indicated above. For precise details regarding JWS encodings, we refer to [RFC7515].

#### Author's Address

Rene Struik  
Struik Security Consultancy

Email: rstruik.ext@gmail.com

LWIG  
Internet-Draft  
Intended status: Informational  
Expires: August 25, 2019

R. Jadhav, Ed.  
R. Sahoo  
Huawei  
S. Duquennoy  
Inria  
J. Eriksson  
Yanzi Networks  
February 21, 2019

Neighbor Management Policy for 6LoWPAN  
draft-ietf-lwig-nbr-mgmt-policy-03

Abstract

This document describes the problems associated with neighbor cache management in multihop networks involving resource-constrained devices. Thereafter, it also presents a sample neighbor management policy that allows efficient cache management in multihop LLNs (low-power and lossy networks such as LoWPAN) with resource-constrained devices.

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 August 25, 2019.

Copyright Notice

Copyright (c) 2019 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 Provisions Relating to IETF Documents (<https://trustee.ietf.org/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 Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

## Table of Contents

1. Introduction . . . . .	2
1.1. Requirements Language and Terminology . . . . .	4
2. Neighbor Management . . . . .	5
2.1. Significance of Neighbor management policy . . . . .	5
2.2. Trivial neighbor management policies . . . . .	6
2.3. Lifecycle of a NCE . . . . .	7
2.3.1. NCE Insertion . . . . .	7
2.3.2. NCE Deletion . . . . .	10
2.3.3. NCE Eviction . . . . .	11
2.3.3.1. Eviction for directly connected routing entries . . . . .	11
2.3.4. NCE Reinforcement . . . . .	12
2.4. Requirements of a good neighbor management policy . . . . .	12
2.5. Approaches to neighbor management policy . . . . .	13
2.5.1. Reactive Approach . . . . .	13
2.5.2. Proactive Approach . . . . .	13
3. Reservation based Neighbor Management Policy . . . . .	14
3.1. Limitations of such a policy . . . . .	16
4. Acknowledgements . . . . .	16
5. IANA Considerations . . . . .	17
6. Security Considerations . . . . .	17
7. References . . . . .	17
7.1. Normative References . . . . .	17
7.2. Informative References . . . . .	18
Appendix A. Performance Result . . . . .	19
Authors' Addresses . . . . .	20

## 1. Introduction

In a wireless multihop LLN, the node densities (maximum nodes reachable on the same hop) may vary significantly depending upon deployments and scenarios. Examples of such networks is LoWPAN [REF] networks. While there is some policy control possible with regards to the network size in terms of maximum number of devices connected, it is especially difficult to set a figure on what will be the maximum node density given a deployment. For e.g. A network can put an upper limit on max 1000 devices but it is impossible to state what the node density will be in this 1000 node network.

A neighbor cache is used for populating neighboring one-hop connected nodes information such as MAC address, link local IP address and

other reachability state information. Node density has direct implications on the neighbor cache and in constrained network scenario the size of the neighbor cache will be limited. Thus there are chances that a node may not be able to fit all the neighboring nodes in its cache in which case it has to prioritize entries and thus needs a neighbor management policy.

This draft presents problems related to neighbor management policies by considering a security-enabled multi-hop 6lo network. This document considers RPL [RFC6550] as a routing protocol and PANA (EAP-PANA) [RFC5191] as a network access protocol. For RPL, both the storing and non-storing mode of operations are considered. We also provide a sample neighbor management policy which can be used in such networks and its limitations. The aim of such a policy is to retain set of neighbor cache entries with high quality links such that routing adjacencies are stabilized.

The problem statement and the proposed solution described is also relevant to other multihop constrained scenarios such as 6TiSCH [I-D.ietf-6tisch-architecture]. [I-D.ietf-6tisch-minimal-security] talks about a pledge (new joinee) node authenticating via a Join Proxy. The considerations mentioned in this document are applicable for such networks as well.

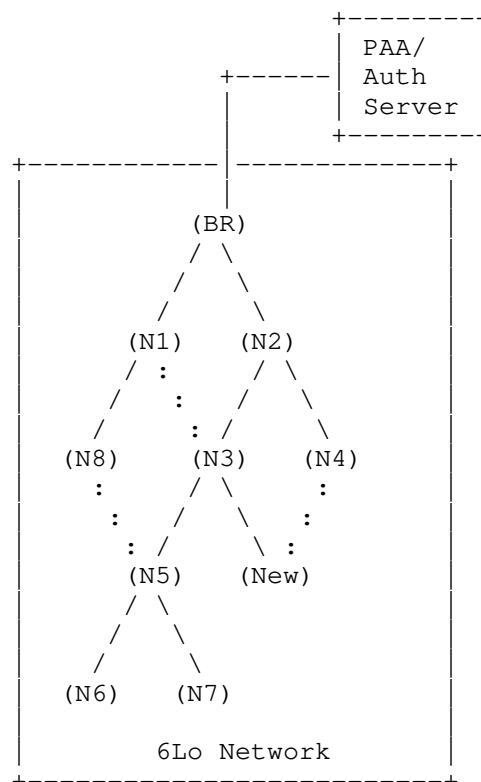


Figure 1: Sample Topology

### 1.1. Requirements Language and Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

NDP: Neighbor Discovery protocol [REF]

NS: Neighbor Solicitation

NA: Neighbor Advertisement

LLN: Low Power and Lossy Networks

RPL: Routing Protocol for LLNs [REF]

DAO: DODAG Advertisement Object

DIO: DODAG Information Object

ARO: Address Registration Option defined as part of IPv6 NDP

PaC (PANA Client): New joining node which is yet to be authenticated.

PRE (PANA Relay Element): An already authenticated and network joined node which is willing to act as a relay element for PaCs to complete their authentication procedure on multi-hop networks. [RFC6345] describes the details of PRE.

PAA (PANA Auth Agent): Auth server which hosts the credentials database. PaC will handshake with PAA to complete authentication procedure.

PCI: PANA Client Initiation is the first message sent by the PaC which initiates the authentication procedure

Routing Child: A downstream node who is part of the routing table of the parent. For e.g. in the sample topology above N5 is the directly connected routing child for N3. N6 and N7 are also part of N3 routing table, they are routing child nodes but not directly connected. For N6 and N7 the document might alternatively use a term grand-child.

Routing Parent: In Figure 1, N1 and N2 are possible routing parents for N3.

Neighbor Cache Entry (NCE): A neighbor entry managed on behalf of directly connected peer.

This document also uses terminology described in [RFC6550] and [RFC6775].

## 2. Neighbor Management

### 2.1. Significance of Neighbor management policy

Multihop mesh networks present unique challenges to neighbor management especially with resource constrained nodes. In cases where the node density is higher than the neighbor cache size, the entries have to be prioritized. [Woo\_et\_al] and [Dawans\_et\_al] talk about prioritization of neighbor entries by using link quality estimation techniques. But prioritization alone may not necessarily be optimal in all cases. The reason or function why neighbor entry was added also needs to be taken in consideration. For example, evicting a routing direct child might have a ripple effect in turn impacting all the sub-children as well.

In case of key management protocols deployed above MAC layer in multihop network, the neighbor management kicks in early even before the routing adjacencies are established. Since a new joining node needs to discover/attach to a relay element for completing its authentication procedure, the neighbor cache entries have to be appropriately populated both on a PaC and on the PRE. If a neighbor entry whose authentication is in progress is evicted, it will negatively impact the authentication procedure.

Another important consideration is that with increased node density, the prioritization based on link estimation parameters might not help since there might be more well connected peers. In dense deployments the number of directly attached neighbors with good quality links might still be higher than the max entries in neighbor cache size.

## 2.2. Trivial neighbor management policies

This section investigates policies which are used by most of the current operating systems for constrained nodes. While such policies are trivial to implement they may not be able to deal with the constrained network scenario. Note that such policies can still be used if it is known apriori that the neighbor cache can hold entries for maximum node density.

- a. First Come First Serve (FCFS) policy
- b. Least Recently Used (LRU) policy

The primary distinction between these policies is how it treats a new entry when the neighbor cache is full. In case of FCFS policy, the new entry is simply rejected while with LRU, the new entry replaces the least recently used entry.

RPL works by initiating a downstream multicast DIO to establish upstream network path. Subsequently DAO messages might be sent by the nodes to establish downstream paths to the nodes. Thus the network is flooded with multicast DIO messages initially and similarly there are chances that the same node is ended up been selected as a preferred parent by most of the child nodes and thus receives a DAO message from all these child nodes. Note that once a node establishes a parent entry or a routing entry on behalf of a directly connected node then it has to also provision a neighbor cache entry for it for subsequent unicast traffic.

In case of FCFS policy, a node might end up hosting all the neighbor entries based on DIO or DAO messages. Once the cache is full all the subsequent attempts to add an NCE will fail.



In case of LRU policy, a node might end up churning lot of neighbor entries because once the cache gets full and there is a request for new entry, it would result in evicting the least recently used (but active) entry. If at later point of time, there is a traffic for the evicted entry then the old entry has to be reinstated using IPv6 NDP procedure. This would mean reinstating the entry by evicting another least recently used entry. If the node density is very high, then this churn would be substantially high to extent that it would disrupt any routing adjacencies to be established in the network in a stable way.

## 2.3. Lifecycle of a NCE

### 2.3.1. NCE Insertion

IPv6 NDP [RFC6775] defines signaling involved in resolving the IPv6 addresses to its corresponding MAC addresses which gets populated in the neighbor cache. In case of constrained network, it is desired that such control traffic is minimized and thus the neighbor cache entries are populated as part of existing messaging. One example would be when the node receives a DAO message from its immediate child node, it not only makes an addition to the routing table but also creates a neighbor cache entry for the node. Thus it eliminates need for additional IPv6 NDP NS/NA messaging involved to resolve MAC address. Similar heuristic is used to add neighbor entries in other cases as well. Section 10.3.2 of [RFC6775] describes update and addition of such NCEs based on routing information packets.

Following are the possible signaling scenarios in which case a neighbor entry may get added.

**Node Joining procedure:** A new joinee node discovers a relay element to initiate its auth procedure. At the end of the discovery phase the new joinee node would have known the link local IP address of the relay element. The joinee node will send an unsecured-NS to the relay element to solicit its NA. The PRE may send a NA with the suitable status code as defined in section 6.5.3 of [RFC6775].

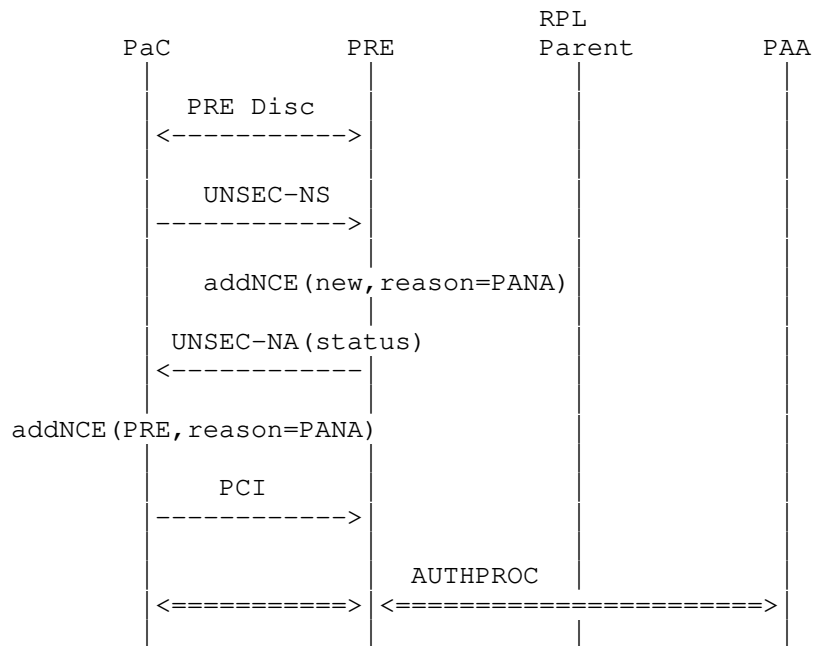


Figure 2: NCE creation between PaC and PRE during relay discovery process

Relay element does not hold any state information on behalf of the new joinee node except for its neighbor cache entry. Thus in the Figure 1 the new joinee node may select node N3 as its PRE, in which case N3 has to add a neighbor entry on behalf of the new joinee node.

Post authentication the node enters into network discovery phase. The node selects one or more of its neighboring peer as its preferred parent based on the DIO received from these peers. Note that the node's selected relay element and its preferred parent may not be same. The preferred parent serves as a default router node to which all its upstream traffic is directed. Thus an NCE on behalf of preferred parent needs to be added. In Figure 1 node N5 selects N3 as its preferred parent. N5 needs to add neighbor entry on behalf of N3 which is its directly connected RPL preferred parent.

In case of RPL storing MOP (mode of operation), the node may send a DAO message containing its reachability information to its preferred parent. The parent node in turn may pass this information upstream to its parent by generating a DAO retaining the child node's reachability information, establishing a downstream routing path towards the node who originated the DAO. The preferred parent has to maintain a neighbor entry on behalf of the directly connected child

node. For example, in the Figure 1, node N3 needs to maintain a neighbor entry on behalf of N5 which is its directly connected child node. Nodes N6 and N7 are grand-child nodes for node N3 for whom no neighbor entry is required.

As mentioned in Section 10.3.2 of [RFC6775], the NCEs on parent and child can be added directly as a result of RPL DIO/DAO signalling without any explicit NS/NA messaging.

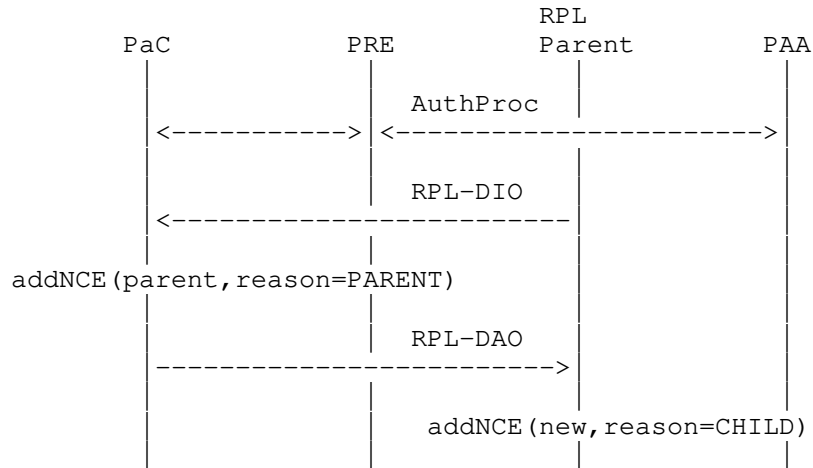


Figure 3: NCE creation call Flow for RPL storing MOP

In case of non-storing MOP, the parent node needs to know the global IPv6 address of the immediate child nodes. This is needed since the source routing header carries the global addresses and thus the NCE of the child node should contain the global address. Secondly, the RPL DAO is addressed directly to the root node in case of non-storing mode. Thus RPL messaging cannot be used for creating NCE entries on parent and child, unlike storing MOP. The child node may send a secure unicast NS with ARO option containing its global address to be registered on the parent node. The child node can still use RPL DIO to create an NCE on behalf of the parent node.

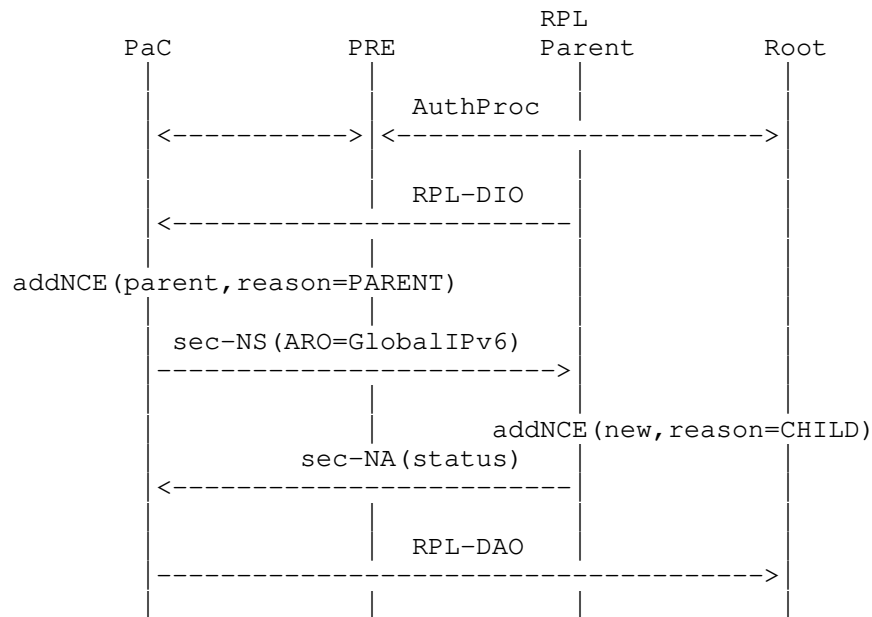


Figure 4: NCE creation call Flow for non-storing MOP

This document expects the neighbor management policy to remember the reason why the neighbor entry is inserted. Secondly, the router may remember whether the NS received was secured or unsecured and accordingly use it to prioritize eviction entries. As described in the next sections, this reason will help the policy to prioritize the entries in case an eviction is required.

### 2.3.2. NCE Deletion

It is imperative that an unwanted neighbor entry be removed as soon as possible. This section talks about different cases in which neighbor entry can be deleted.

**Route Invalidation:** In case of storing MOP, when the child node decides to switch its preferred parent, the RPL specifications allows the node to send a no-path DAO message to invalidate the route along the previous path(s). A directly connected parent node can use this message to clear the NCE. While the entry can be immediately cleared, usually the implementations choose to wait a small amount of time before clearing the entry. This is to avoid any impact on the in-transit traffic. Thus this also establishes the importance of route invalidation to achieve optimized neighbor cache utilization.

Efficient neighbor cache management depends upon efficient route invalidation since the neighbor cache entries are associated with routing entries. With regards to RPL, issues with the route invalidation has been highlighted in [I-D.ietf-roll-efficient-npdao]. [I-D.ietf-roll-efficient-npdao] also defines a new mechanism for improved route invalidation in storing MOP which helps optimized cleanup of neighbor cache entries.

In case of non-storing mode, the no-path DAO cannot be not employed since the previous parent does not having any routing information to be invalidated. But the previous parent may still contain the NCE on behalf of the child node. This document recommends use of [RFC6775] section 6.5.3. which allows sending a zero lifetime ARO option in NS for deregistering the corresponding neighbor entry.

[RFC6775], ND optimizations for 6LoWPANs, section 5.5.3. talks about deleting the entries in case the NUD (neighbor unreachability detection) fails either due to no response to NS messages or due to failure response. NCEs in such cases should be deleted. An example where NUD NS would fail because of no response is the case where the child node switches its parent due to link unavailability. The parent in such a case would not receive the no-path DAO message or any other traffic from the child node. Thus on NCE lifetime expiry, the parent node would send NS which would fail with no response, thus triggering entry deletion.

#### 2.3.3. NCE Eviction

The eviction rules have a major impact on the neighbor management policy. Eviction rules are used when the policy has to forcibly remove an active neighbor entry from the cache to make space for the new (hopefully higher priority) entry. The eviction policy may take into account several considerations such as the reason why the entry was made, is the entry in active use currently, how good (for e.g., based on link estimation) the entry currently is.

##### 2.3.3.1. Eviction for directly connected routing entries

This section talks about implications of an eviction in which a parent node decides of evicting a directly connected routing child NCE. In the sample topology Figure 1, lets assume N3 needs to evict N5 from its neighbor cache. In case of RPL's storing MOP, eviction of directly connected routing child NCE also has impact on all the sub-children. Thus not only will it result in impacting N5 but also nodes N6 and N7. It is important to note that such an eviction has less impact on RPL's non-storing MOP i.e. in case of non-storing mode N5 might end up selecting alternate parent N8 and does not result in any additional control overhead for node N6 and N7.

Thus RPL's non-storing MOP provides additional eviction flexibility for a neighbor management policy in terms evicting directly connected child entries.

#### 2.3.4. NCE Reinforcement

It is expected that the latest reachability state and metric information be maintained in context to the NCE. With wireless networks, the neighbor cache entries prioritization may change over a period of time especially the link quality estimation parameters or the routing metrics. Reinforcement refers to updating the parameters in context to the NCEs which helps in prioritizing the entries when it comes to handling eviction. In wireless networks, on reception of incoming packet, the receiver node's physical and MAC layer may derive certain signal reception parameters (such as RSSI, LQI) which can be considered for reinforcement purpose if the corresponding transmitter/source entry in neighbor cache is found. It should be noted that the signal quality parameters may have high variance in 6Lo networks and thus statistical techniques (such as weighted averaging) are usually employed for deciding about a link quality over a period of time. Reinforcement can be achieved using one or more of the following techniques:

**Passive Monitoring:** Reinforcing the quality parameters using packets received from the source. TrickleD DIO, periodic beacons, application traffic etc can be used for such monitoring.

**Active Probing:** A node may select subset of entries for active probing wherein it sends a message to the neighbor entry's target and can expect a response message back. An example of such probing is [CONTIKI] where unicast DIS is sent to solicit a unicast DIO without impacting the trickle timers. Though it adds a control overhead on the link, periodic probing can help to ascertain connectivity in the absence of any other traffic from the neighboring node.

#### 2.4. Requirements of a good neighbor management policy

**Route Stability:** Stable NCEs will result in stable routing adjacencies. Thus it is important to avoid unnecessary NCE churn for routing path stability.

**Control overhead:** A neighbor management policy may have to use signalling messages for policy handling (such as rejection of NCE). It is required that such overhead be kept as low as possible.

Scalability: Scalability with respect to large and uneven node densities in the network.

## 2.5. Approaches to neighbor management policy

Neighbor management policy depends upon the neighbor cache space availability and the same can be advertised proactively or can be handled reactively.

### 2.5.1. Reactive Approach

In this approach, the nodes select their RPL parent or the relay element purely based on link metrics and subsequently when they try to allocate their NCE in the target node, it may fail due to unavailability of the cache space. The failure can be communicated depending upon the signaling involved:

NS failure: Section 6.5.3 of [RFC6775] defines a procedure for NS failure handling in case the router's neighbor cache is full. It results in a unicast NA with ARO status field set to two.

DAO NACK: Section 9.3 of RPL [RFC6550] specifies on how can the parent node react to DAOs from child. In case the parent could not make a NCE on behalf of the child node, a negative ACK with status (between 127-255) should be sent to the child node. The natural reaction of the child node would be to switch to an alternate parent.

PANA Failure: PaC's auth session starts with a PaC discovering a PRE. The discovery procedure is not standardized and can be based upon various factors including signal strength of discovery messages from PRE. Post discovery, the PaC needs to send an unsecured unicast NS message with an ARO containing its link-local IPv6 address. NS helps to determine whether the PRE can allocate an NCE for the PaC. PRE accordingly sends a NA response with appropriate status field.

### 2.5.2. Proactive Approach

Neighbor cache availability could be proactively advertised by the parent nodes in the DIO messages and in the PRE discovery messages. A child RPL node may additionally use this information from DIO as part of parent selection process.

[I-D.richardson-6tisch-roll-enrollment-priority] defines a signalling change in DIO messages to inform child nodes of the priority of the 6LR. The priority field signals whether the 6LR is ready and has enough resources to serve new child nodes. If 6LR's neighbor cache

gets full it can set the min priority to 0x7f(127) to stop the joining process via it. When a LR or leaf node receives DIO from a parent LR with min priority set to 127 the below actions

1. If its not yet joined the DODAG don't choose this LR as preferred parent to join the DODAG.
2. If the node is already in the DODAG and this LR is not in the parent list don't add it to parent list.
3. If node is already in the DODAG and the LR is in its standby parent list remove the LR from its standby parent list.

In case of new joinee node, the node may use PRE discovery messages with space availability information to select an appropriate PRE. Proactive signaling of neighbor cache space availability will help the nodes to select the parent node or relay node such that the failure signaling due to cache full event can be reduced.

Currently there is no standard way of signaling such neighbor cache space availability information. RPL's DIO messages carry metric information and can be augmented with neighbor cache space as an additional metric. In case of PRE discovery however there is no standard way of defining this information since the PRE discovery procedure itself is not standardized.

In a wireless or shared bus network, a multicast DIO metric advertisement may reach several child nodes eventually everyone responding by selecting the same parent node causing neighbor cache to be exhausted. Thus the failure handling approaches defined in the Reactive Approach section applies here as well. But importantly the failure signaling will be significantly reduced because of proactive advertisement.

### 3. Reservation based Neighbor Management Policy

This section defines a sample neighbor management policy, with the primary objective to reduce NCE churn and to ensure stability of routing adjacencies. The scheme uses a reservation based policy to reserve NCEs for:



NCE Entry for	MAX count	Reason
Routing Parent	MAX_ROUTING_PARENT_NCE_NUM	PARENT
Routing child	MAX_ROUTING_CHILD_NCE_NUM	CHILD
Others such as pre-auth sessions	MAX_OTHER_NCE_NUM	OTHER

Table 1: Neighbor Cache Entry reservation

Table 1 denotes that the NCEs are reserved depending upon the reason for its addition. MAX\_ROUTING\_PARENT\_NCE\_NUM specifies maximum number of parent entries a node should allow. MAX\_ROUTING\_CHILD\_NCE\_NUM specifies maximum number of child entries that are allowed after which the node SHOULD decline the DAO signalling. MAX\_OTHER\_NCE\_NUM specifies any other entries that might be created. Pre-auth session entries are usually short-lived and should be considered part of this category.

Note that reservation policy depends upon identification of the reason behind making an NCE . In case of pre-auth sessions, the corresponding NCE is created based on unsecured NS/NA. In case of storing MOP, CHILD NCEs are created either based on DAO (as shown in Figure 3) or based on secured NS/NA messaging (as shown in Figure 4). In case of non-storing MOP, a secured NS/NA messaging as shown in Figure 4 needs to be used.

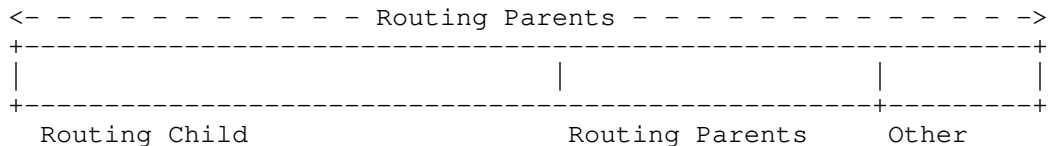


Figure 5: Reservation of NCEs in neighbor table

As shown in the Figure 5, the neighbor cache is partitioned into different entry types. The routing parents can possibly occupy any entry type if found vacant since in case an eviction is sought the non-preferred routing parent could be evicted without much impact on the functioning or on the control traffic. The eviction could be done based on reasons specified in Section 2.3.3.

Routing Child entries are made in context to directly connected peers and these entries are not deleted unless they are unreachabe or there is any reason for the parent node to believe that it is no longer the

preferred parent for the child node. Deletion may happen based on reasons mentioned in Section 2.3.2.

Other entries (OTHER) may be made in response to temporary requirement of making an NCE. One such case is the pre authentication phase where in the relay node makes an entry of the PaC temporarily till the time the authentication phase is completed. The NCE made thus is garbage collected at the end of the lifetime. Also an implementation may choose to keep a lower lifetime for such NCEs depending upon the time taken to complete the authentication process.

### 3.1. Limitations of such a policy

The reservation based policy mentioned in this section may result in sub-optimal path selection due to lack of NCE resource on the parent nodes. Also the restriction of maximum pre-auth sessions in the form of MAX\_OTHER\_NCE\_NUM limits the maximum relay sessions that can be supported on the relay node.

The reservation policy allows the parent node to reject the child node's DAO or NS. But the child node cannot remember this rejection and may reattempt the same parent after some time depending upon triggers such as reception of DIO from the same parent who rejected it previously. One of the only way to stop the child node from reattempting such parent selection would be to also include a proactive approach wherein the parent node signals its resource availability in the DIO message as mentioned in Section 2.5.2. Such a scheme of signalling parent node's resource availability is currently not standardized.

RPL's storing MOP imposes additional restrictions. One such case is where a child node may have a given parent node as its only parent and that parent node's NCE are all used up. In such a case, the child node would keep on retrying and failing to send a DAO through the parent node. Ideally the parent node could have evicted a least used child node or a child node who has an alternate parent available. Evicting such a child node is a complex process and may increase the control overhead as described in Section 2.3.3.1. Thus the reservation based policy requires that the minimum node density is sufficiently high so that every child finds a parent node in its vicinity with enough resources.

## 4. Acknowledgements

Thanks to Mohit Sethi for the review and feedback. Thanks to Malisa Vucinic for pointing towards security aspects of un-authenticated nodes neighbor cache management.

## 5. IANA Considerations

This memo includes no request to IANA.

## 6. Security Considerations

The Join Relay or the PANA Relay Element allows the un-authenticated nodes to join the network. Since the NS/NA signaling is unencrypted, it is possible that a malicious node may try to spoof and occupy all the entries. This draft explains the use of reservation based policy for managing such entries. Following points try to reduce the impact of such attacks:

- a. Only a subset of entries are reserved for un-authenticated nodes doing plain-text NS/NA.
- b. It is recommended that NCE timeout be configured a lower value to evict such entries as soon as possible. New joining nodes are supposed to use these entries and thus are only needed during the time the authentication is in progress. Thus a short-duration NCE timeout will help reduce the impact of DoS attacks.

## 7. References

### 7.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <<https://www.rfc-editor.org/info/rfc2119>>.
- [RFC6550] Winter, T., Ed., Thubert, P., Ed., Brandt, A., Hui, J., Kelsey, R., Levis, P., Pister, K., Struik, R., Vasseur, JP., and R. Alexander, "RPL: IPv6 Routing Protocol for Low-Power and Lossy Networks", RFC 6550, DOI 10.17487/RFC6550, March 2012, <<https://www.rfc-editor.org/info/rfc6550>>.
- [RFC6775] Shelby, Z., Ed., Chakrabarti, S., Nordmark, E., and C. Bormann, "Neighbor Discovery Optimization for IPv6 over Low-Power Wireless Personal Area Networks (6LoWPANs)", RFC 6775, DOI 10.17487/RFC6775, November 2012, <<https://www.rfc-editor.org/info/rfc6775>>.

## 7.2. Informative References

- [CONTIKI]   Thingsquare, "Contiki: The Open Source OS for IoT", 2012, <<http://www.contiki-os.org>>.
- [Dawans\_et\_al]   Dawans, S., Duquennoy, S., and O. Bonaventure, "On Link Estimation in Dense RPL Deployments", 2012.
- [I-D.ietf-6tisch-architecture]   Thubert, P., "An Architecture for IPv6 over the TSCH mode of IEEE 802.15.4", draft-ietf-6tisch-architecture-19 (work in progress), December 2018.
- [I-D.ietf-6tisch-minimal-security]   Vucinic, M., Simon, J., Pister, K., and M. Richardson, "Minimal Security Framework for 6TiSCH", draft-ietf-6tisch-minimal-security-09 (work in progress), November 2018.
- [I-D.ietf-roll-efficient-npdao]   Jadhav, R., Thubert, P., Sahoo, R., and Z. Cao, "Efficient Route Invalidation", draft-ietf-roll-efficient-npdao-09 (work in progress), October 2018.
- [I-D.richardson-6tisch-roll-enrollment-priority]   Richardson, M., "Enabling secure network enrollment in RPL networks", draft-richardson-6tisch-roll-enrollment-priority-02 (work in progress), February 2019.
- [LWIP]   "lwIP: A Lightweight TCP/IP stack", <<https://savannah.nongnu.org/projects/lwip/>>.
- [RFC5191]   Forsberg, D., Ohba, Y., Ed., Patil, B., Tschofenig, H., and A. Yegin, "Protocol for Carrying Authentication for Network Access (PANA)", RFC 5191, DOI 10.17487/RFC5191, May 2008, <<https://www.rfc-editor.org/info/rfc5191>>.
- [RFC6345]   Duffy, P., Chakrabarti, S., Cragie, R., Ohba, Y., Ed., and A. Yegin, "Protocol for Carrying Authentication for Network Access (PANA) Relay Element", RFC 6345, DOI 10.17487/RFC6345, August 2011, <<https://www.rfc-editor.org/info/rfc6345>>.
- [WHITEFIELD]   "Whitefield Framework", <<https://github.com/whitefield-framework/whitefield>>.

[Woo\_et\_al]

Woo, A., Tong, T., and D. Culler, "Taming the Underlying Challenges of Reliable Multihop Routing in Sensor Networks", 2003.

#### Appendix A. Performance Result

This appendix provides the details of the performance evaluation of this draft.

**Setup:** To perform the test [WHITEFIELD] framework was used with LwIP [LWIP] as the network stack. A version of this draft is implemented in the lwip to provide efficient neighbor cache management policy that will work with relatively lower neighbor cache size. RPL is used as routing protocol to form mesh network. The available neighbor table size was split 60%, 30% and 10% among direct child entries, parent entries and others respectively.

**Topology:** Test was performed with a 64 nodes network including the border router. Grid topology based network was used and all the nodes were in close range with each other simulating a dense condition. 802.15.4 in 2.4GHz range with single channel and un-slotted CSMA wireless RF was used.

**Steps:** Experiment has been conducted with different neighbor cache sizes 10, 20 and 40. For each NC size we have collected sample readings for packet delivery rate by enabling and disabling the new neighbor Cache Management Policy.

**Data transmission frequency:** Each node in the network sends 104 bytes (IPv6 Header + RPI + UDP + Data) of UDP request to BR at each 10 second interval. udp Server running at BR process these requests and sends the response back , which is also of same size 104 bytes. A duration of 2 minutes delay is added, for network to get stable, before nodes starts sending request messages at 10 sec interval. To calculate PDR one request and response pair is considered as one successful transaction.

#### Packet Delivery Rate Performance

Neighbor Cache Size	PDR With New policy	PDR Without New policy
10	96.3	7.8
20	97.5	31.3
40	98.7	98.6

Table 2: Packet delivery rate

## Authors' Addresses

Rahul Arvind Jadhav (editor)  
 Huawei  
 Kundalahalli Village, Whitefield,  
 Bangalore, Karnataka 560037  
 India

Phone: +91-080-49160700  
 Email: rahul.ietf@gmail.com

Rabi Narayan Sahoo  
 Huawei  
 Kundalahalli Village, Whitefield,  
 Bangalore, Karnataka 560037  
 India

Phone: +91-080-49160700  
 Email: rabinarayans@huawei.com

Simon Duquennoy  
 Inria  
 40 Avenue Halley  
 Building A  
 Villeneuve d'Ascq  
 France

Phone: +33 768227731  
 Email: simon.duquennoy@inria.fr

Joakim Eriksson  
 Yanzi Networks

Email: joakime@sics.se

LWIG Working Group  
Internet-Draft  
Intended status: Informational  
Expires: September 12, 2019

J. Mattsson  
F. Palombini  
Ericsson AB  
March 11, 2019

Comparison of CoAP Security Protocols  
draft-ietf-lwig-security-protocol-comparison-03

Abstract

This document analyzes and compares the sizes of key exchange flights and the per-packet message size overheads when using different security protocols to secure CoAP. The analyzed security protocols are DTLS 1.2, DTLS 1.3, TLS 1.2, TLS 1.3, EDHOC, OSCORE, and Group OSCORE. The DTLS and TLS record layers are analyzed with and without 6LoWPAN-GHC compression. DTLS is analyzed with and without Connection ID.

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 September 12, 2019.

Copyright Notice

Copyright (c) 2019 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 Provisions Relating to IETF Documents (<https://trustee.ietf.org/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 Simplified BSD License text as described in Section 4.e of

the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

## Table of Contents

1. Introduction . . . . .	3
2. Overhead of Key Exchange Protocols . . . . .	3
2.1. Summary . . . . .	4
2.2. DTLS 1.3 . . . . .	5
2.2.1. Message Sizes RPK + ECDHE . . . . .	5
2.2.2. Message Sizes PSK + ECDHE . . . . .	10
2.2.3. Message Sizes PSK . . . . .	11
2.2.4. Cached Information . . . . .	12
2.2.5. Resumption . . . . .	13
2.2.6. Without Connection ID . . . . .	14
2.2.7. DTLS Raw Public Keys . . . . .	15
2.3. TLS 1.3 . . . . .	16
2.3.1. Message Sizes RPK + ECDHE . . . . .	16
2.3.2. Message Sizes PSK + ECDHE . . . . .	22
2.3.3. Message Sizes PSK . . . . .	23
2.4. EDHOC . . . . .	24
2.4.1. Message Sizes RPK . . . . .	24
2.4.2. Message Sizes Certificates . . . . .	26
2.4.3. Message Sizes PSK . . . . .	26
2.4.4. message_1 . . . . .	26
2.4.5. message_2 . . . . .	26
2.4.6. message_3 . . . . .	27
2.4.7. Summary . . . . .	27
2.5. Conclusion . . . . .	27
3. Overhead for Protection of Application Data . . . . .	28
3.1. Summary . . . . .	28
3.2. DTLS 1.2 . . . . .	30
3.2.1. DTLS 1.2 . . . . .	30
3.2.2. DTLS 1.2 with 6LoWPAN-GHC . . . . .	30
3.2.3. DTLS 1.2 with Connection ID . . . . .	31
3.2.4. DTLS 1.2 with Connection ID and 6LoWPAN-GHC . . . . .	32
3.3. DTLS 1.3 . . . . .	32
3.3.1. DTLS 1.3 . . . . .	32
3.3.2. DTLS 1.3 with 6LoWPAN-GHC . . . . .	33
3.3.3. DTLS 1.3 with Connection ID . . . . .	33
3.3.4. DTLS 1.3 with Connection ID and 6LoWPAN-GHC . . . . .	34
3.4. TLS 1.2 . . . . .	34
3.4.1. TLS 1.2 . . . . .	34
3.4.2. TLS 1.2 with 6LoWPAN-GHC . . . . .	35
3.5. TLS 1.3 . . . . .	35
3.5.1. TLS 1.3 . . . . .	35
3.5.2. TLS 1.3 with 6LoWPAN-GHC . . . . .	36
3.6. OSCORE . . . . .	36



3.7. Group OSCORE . . . . .	38
3.8. Conclusion . . . . .	38
4. Security Considerations . . . . .	39
5. IANA Considerations . . . . .	39
6. Informative References . . . . .	39
Acknowledgments . . . . .	41
Authors' Addresses . . . . .	41

## 1. Introduction

This document analyzes and compares the sizes of key exchange flights and the per-packet message size overheads when using different security protocols to secure CoAP over UDP [RFC7252] and TCP [RFC8323]. The analyzed security protocols are DTLS 1.2 [RFC6347], DTLS 1.3 [I-D.ietf-tls-dtls13], TLS 1.2 [RFC5246], TLS 1.3 [RFC8446], EDHOC [I-D.selander-ace-cose-ecdh], OSCORE [I-D.ietf-core-object-security], and Group OSCORE [I-D.ietf-core-oscore-groupcomm].

The DTLS and TLS record layers are analyzed with and without 6LoWPAN-GHC compression. DTLS is analyzed with and without Connection ID [I-D.ietf-tls-dtls-connection-id]. Readers are expected to be familiar with some of the terms described in RFC 7925 [RFC7925], such as ICV. Section 2 compares the overhead of key exchange, while Section 3 covers the overhead for protection of application data.

## 2. Overhead of Key Exchange Protocols

This section analyzes and compares the sizes of key exchange flights for different protocols.

To enable a fair comparison between protocols, the following assumptions are made:

- o All the overhead calculations in this section use AES-CCM with a tag length of 8 bytes (e.g. AES\_128\_CCM\_8 or AES-CCM-16-64-128).
- o A minimum number of algorithms and cipher suites is offered. The algorithm used/offered are Curve25519, ECDSA with P-256, AES-CCM\_8, SHA-256.
- o The length of key identifiers are 1 byte.
- o The length of connection identifiers are 1 byte.
- o DTLS RPK makes use of point compression, which saves 32 bytes.
- o DTLS handshake message fragmentation is not considered.

- o Only the DTLS mandatory extensions are considered, except for Connection ID.

Section 2.1 gives a short summary of the message overhead based on different parameters and some assumptions. The following sections detail the assumptions and the calculations.

## 2.1. Summary

The DTLS overhead is dependent on the parameter Connection ID. The following overheads apply for all Connection IDs of the same length, when Connection ID is used.

The EDHOC overhead is dependent on the key identifiers included. The following overheads apply for Sender IDs of the same length.

All the overhead are dependent on the tag length. The following overheads apply for tags of the same length.

Figure 1 compares the message sizes of EDHOC [I-D.selander-ace-cose-ecdhe] with the DTLS 1.3 [I-D.ietf-tls-dtls13] and TLS 1.3 [RFC8446] handshakes with connection ID.

Flight	#1	#2	#3	Total
DTLS 1.3 RPK + ECDHE	150	373	213	736
DTLS 1.3 Cached X.509/RPK + ECDHE	182	347	213	742
DTLS 1.3 PSK + ECDHE	184	190	57	431
DTLS 1.3 PSK	134	150	57	341
EDHOC RPK + ECDHE	39	114	80	233
EDHOC PSK + ECDHE	41	45	11	97

Figure 1: Comparison of message sizes in bytes with Connection ID

Figure 2 compares of message sizes of DTLS 1.3 [I-D.ietf-tls-dtls13] and TLS 1.3 [RFC8446] handshakes without connection ID.

Flight	#1	#2	#3	Total
DTLS 1.3 RPK + ECDHE	144	364	212	722
DTLS 1.3 PSK + ECDHE	178	183	56	417
DTLS 1.3 PSK	128	143	56	327
TLS 1.3 RPK + ECDHE	129	322	194	645
TLS 1.3 PSK + ECDHE	163	157	50	370
TLS 1.3 PSK	113	117	50	280

Figure 2: Comparison of message sizes in bytes without Connection ID

The details of the message size calculations are given in the following sections.

## 2.2. DTLS 1.3

This section gives an estimate of the message sizes of DTLS 1.3 with different authentication methods. Note that the examples in this section are not test vectors, the cryptographic parts are just replaced with byte strings of the same length, while other fixed length fields are replaced with arbitrary strings or omitted, in which case their length is indicated. Values that are not arbitrary are given in hexadecimal.

### 2.2.1. Message Sizes RPK + ECDHE

In this section, a Connection ID of 1 byte is used.

#### 2.2.1.1. flight\_1

Record Header - DTLSPlaintext (13 bytes):

16 fe fd EE EE SS SS SS SS SS SS LL LL

Handshake Header - Client Hello (10 bytes):

01 LL LL LL SS SS 00 00 00 LL LL LL

Legacy Version (2 bytes):

fe fd

Client Random (32 bytes):

00 01 02 03 04 05 06 07 08 09 0a 0b 0c 0d 0e 0f 10 11 12 13 14 15  
16 17 18 19 1a 1b 1c 1d 1e 1f

Legacy Session ID (1 bytes):

00

Legacy Cookie (1 bytes):  
00

Cipher Suites (TLS\_AES\_128\_CCM\_8\_SHA256) (4 bytes):  
00 02 13 05

Compression Methods (null) (2 bytes):  
01 00

Extensions Length (2 bytes):  
LL LL

Extension - Supported Groups (x25519) (8 bytes):  
00 0a 00 04 00 02 00 1d

Extension - Signature Algorithms (ecdsa\_secp256r1\_sha256)  
(8 bytes):  
00 0d 00 04 00 02 08 07

Extension - Key Share (42 bytes):  
00 33 00 26 00 24 00 1d 00 20  
00 01 02 03 04 05 06 07 08 09 0a 0b 0c 0d 0e 0f 10 11 12 13 14 15  
16 17 18 19 1a 1b 1c 1d 1e 1f

Extension - Supported Versions (1.3) (7 bytes):  
00 2b 00 03 02 03 04

Extension - Client Certificate Type (Raw Public Key) (6 bytes):  
00 13 00 01 01 02

Extension - Server Certificate Type (Raw Public Key) (6 bytes):  
00 14 00 01 01 02

Extension - Connection Identifier (43) (6 bytes):  
XX XX 00 02 01 42

13 + 10 + 2 + 32 + 1 + 1 + 4 + 2 + 2 + 8 + 8 + 42 + 7 + 6 + 6 + 6 = 150  
bytes

DTLS 1.3 RPK + ECDHE flight\_1 gives 150 bytes of overhead.

#### 2.2.1.2. flight\_2

Record Header - DTLSPlaintext (13 bytes):  
16 fe fd EE EE SS SS SS SS SS SS LL LL

Handshake Header - Server Hello (10 bytes):  
02 LL LL LL SS SS 00 00 00 LL LL LL

Legacy Version (2 bytes):  
fe fd

Server Random (32 bytes):  
00 01 02 03 04 05 06 07 08 09 0a 0b 0c 0d 0e 0f 10 11 12 13 14 15  
16 17 18 19 1a 1b 1c 1d 1e 1f

Legacy Session ID (1 bytes):  
00

Cipher Suite (TLS\_AES\_128\_CCM\_8\_SHA256) (2 bytes):  
13 05

Compression Method (null) (1 bytes):  
00

Extensions Length (2 bytes):  
LL LL

Extension - Key Share (40 bytes):  
00 33 00 24 00 1d 00 20  
00 01 02 03 04 05 06 07 08 09 0a 0b 0c 0d 0e 0f 10 11 12 13 14 15  
16 17 18 19 1a 1b 1c 1d 1e 1f

Extension - Supported Versions (1.3) (6 bytes):  
00 2b 00 02 03 04

Extension - Connection Identifier (43) (6 bytes):  
XX XX 00 02 01 43

Record Header - DTLS Ciphertext, Full (6 bytes):  
HH ES SS 43 LL LL

Handshake Header - Encrypted Extensions (10 bytes):  
08 LL LL LL SS SS 00 00 00 LL LL LL

Extensions Length (2 bytes):  
LL LL

Extension - Client Certificate Type (Raw Public Key) (6 bytes):  
00 13 00 01 01 02

Extension - Server Certificate Type (Raw Public Key) (6 bytes):  
00 14 00 01 01 02

Handshake Header - Certificate Request (10 bytes):  
0d LL LL LL SS SS 00 00 00 LL LL LL

Request Context (1 bytes):  
00

Extensions Length (2 bytes):  
LL LL

Extension - Signature Algorithms (ecdsa\_secp256r1\_sha256)  
(8 bytes):  
00 0d 00 04 00 02 08 07

Handshake Header - Certificate (10 bytes):  
0b LL LL LL SS SS 00 00 00 LL LL LL

Request Context (1 bytes):  
00

Certificate List Length (3 bytes):  
LL LL LL

Certificate Length (3 bytes):  
LL LL LL

Certificate (59 bytes) // Point compression  
....

Certificate Extensions (2 bytes):  
00 00

Handshake Header - Certificate Verify (10 bytes):  
0f LL LL LL SS SS 00 00 00 LL LL LL

Signature (68 bytes):  
ZZ ZZ 00 40 ....

Handshake Header - Finished (10 bytes):  
14 LL LL LL SS SS 00 00 00 LL LL LL

Verify Data (32 bytes):  
00 01 02 03 04 05 06 07 08 09 0a 0b 0c 0d 0e 0f 10 11 12 13 14 15  
16 17 18 19 1a 1b 1c 1d 1e 1f

Record Type (1 byte):  
16

Auth Tag (8 bytes):  
e0 8b 0e 45 5a 35 0a e5

13 + 102 + 6 + 24 + 21 + 78 + 78 + 42 + 1 + 8 = 373 bytes

DTLS 1.3 RPK + ECDHE flight\_2 gives 373 bytes of overhead.

#### 2.2.1.3. flight\_3

Record Header (6 bytes) // DTLSCiphertext, Full:  
ZZ ES SS 42 LL LL

Handshake Header - Certificate (10 bytes):  
0b LL LL LL SS SS XX XX XX LL LL LL

Request Context (1 bytes):  
00

Certificate List Length (3 bytes):  
LL LL LL

Certificate Length (3 bytes):  
LL LL LL

Certificate (59 bytes) // Point compression  
....

Certificate Extensions (2 bytes):  
00 00

Handshake Header - Certificate Verify (10 bytes):  
0f LL LL LL SS SS 00 00 00 LL LL LL

Signature (68 bytes):  
04 03 LL LL //ecdsa\_secp256r1\_sha256  
00 01 02 03 04 05 06 07 08 09 0a 0b 0c 0d 0e 0f 10 11 12 13 14 15  
16 17 18 19 1a 1b 1c 1d 1e 1f 00 01 02 03 04 05 06 07 08 09 0a 0b  
0c 0d 0e 0f 10 11 12 13 14 15 16 17 18 19 1a 1b 1c 1d 1e 1f

Handshake Header - Finished (10 bytes):  
14 LL LL LL SS SS 00 00 00 LL LL LL

Verify Data (32 bytes) // SHA-256:  
00 01 02 03 04 05 06 07 08 09 0a 0b 0c 0d 0e 0f 10 11 12 13 14 15  
16 17 18 19 1a 1b 1c 1d 1e 1f

Record Type (1 byte):  
16

Auth Tag (8 bytes) // AES-CCM\_8:  
00 01 02 03 04 05 06 07

6 + 78 + 78 + 42 + 1 + 8 = 213 bytes

DTLS 1.3 RPK + ECDHE flight\_2 gives 213 bytes of overhead.

## 2.2.2. Message Sizes PSK + ECDHE

### 2.2.2.1. flight\_1

The differences in overhead compared to Section 2.2.1.1 are:

The following is added:

+ Extension - PSK Key Exchange Modes (6 bytes):

00 2d 00 02 01 01

+ Extension - Pre Shared Key (48 bytes):

00 29 00 2F

00 0a 00 01 ID 00 00 00 00

00 21 20 00 01 02 03 04 05 06 07 08 09 0a 0b 0c 0d 0e 0f 10 11 12 13  
14 15 16 17 18 19 1a 1b 1c 1d 1e 1f

The following is removed:

- Extension - Signature Algorithms (ecdsa\_secp256r1\_sha256) (8 bytes)

- Extension - Client Certificate Type (Raw Public Key) (6 bytes)

- Extension - Server Certificate Type (Raw Public Key) (6 bytes)

In total:

$150 + 6 + 48 - 8 - 6 - 6 = 184$  bytes

DTLS 1.3 PSK + ECDHE flight\_1 gives 184 bytes of overhead.

### 2.2.2.2. flight\_2

The differences in overhead compared to Section 2.2.1.2 are:

The following is added:

+ Extension - Pre Shared Key (6 bytes)

00 29 00 02 00 00

The following is removed:



- Handshake Message Certificate (78 bytes)
- Handshake Message CertificateVerify (78 bytes)
- Handshake Message CertificateRequest (21 bytes)
- Extension - Client Certificate Type (Raw Public Key) (6 bytes)
- Extension - Server Certificate Type (Raw Public Key) (6 bytes)

In total:

$373 - 78 - 78 - 21 - 6 - 6 + 6 = 190$  bytes

DTLS 1.3 PSK + ECDHE flight\_2 gives 190 bytes of overhead.

#### 2.2.2.3. flight\_3

The differences in overhead compared to Section 2.2.1.3 are:

The following is removed:

- Handshake Message Certificate (78 bytes)
- Handshake Message Certificate Verify (78 bytes)

In total:

$213 - 78 - 78 = 57$  bytes

DTLS 1.3 PSK + ECDHE flight\_3 gives 57 bytes of overhead.

#### 2.2.3. Message Sizes PSK

##### 2.2.3.1. flight\_1

The differences in overhead compared to Section 2.2.2.1 are:

The following is removed:

- Extension - Supported Groups (x25519) (8 bytes)
- Extension - Key Share (42 bytes)

In total:

$184 - 8 - 42 = 134$  bytes

DTLS 1.3 PSK flight\_1 gives 134 bytes of overhead.

#### 2.2.3.2. flight\_2

The differences in overhead compared to Section 2.2.2.2 are:

The following is removed:

- Extension - Key Share (40 bytes)

In total:

$190 - 40 = 150$  bytes

DTLS 1.3 PSK flight\_2 gives 150 bytes of overhead.

#### 2.2.3.3. flight\_3

There are no differences in overhead compared to Section 2.2.2.3.

DTLS 1.3 PSK flight\_3 gives 57 bytes of overhead.

#### 2.2.4. Cached Information

In this section, we consider the effect of [RFC7924] on the message size overhead.

Cached information together with server X.509 can be used to move bytes from flight #2 to flight #1 (cached RPK increases the number of bytes compared to cached X.509).

The differences compared to Section 2.2.1 are the following.

For the flight #1, the following is added:

+ Extension - Client Cached Information (39 bytes):

```
00 19 LL LL LL LL
01 00 01 02 03 04 05 06 07 08 09 0a 0b 0c 0d 0e 0f 10 11 12 13 14 15
16 17 18 19 1a 1b 1c 1d 1e 1f
```

And the following is removed:

- Extension - Server Certificate Type (Raw Public Key) (6 bytes)

Giving a total of:

$150 + 33 = 183$  bytes

For the flight #2, the following is added:

+ Extension - Server Cashed Information (7 bytes):  
00 19 LL LL LL LL 01

And the following is removed:

- Extension - Server Certificate Type (Raw Public Key) (6 bytes)
- Server Certificate (59 bytes -> 32 bytes)

Giving a total of:

$373 - 26 = 347$  bytes

A summary of the calculation is given in Figure 3.

Flight	#1	#2	#3	Total
DTLS 1.3 Cached X.509/RPK + ECDHE	183	347	213	743
DTLS 1.3 RPK + ECDHE	150	373	213	736

Figure 3: Comparison of message sizes in bytes for DTLS 1.3 RPK + ECDH with and without cached X.509

#### 2.2.5. Resumption

To enable resumption, a 4th flight (New Session Ticket) is added to the PSK handshake.

Record Header - DTLS Ciphertext, Full (6 bytes):  
 HH ES SS 43 LL LL

Handshake Header - New Session Ticket (10 bytes):  
 04 LL LL LL SS SS 00 00 00 LL LL LL

Ticket Lifetime (4 bytes):  
 00 01 02 03

Ticket Age Add (4 bytes):  
 00 01 02 03

Ticket Nonce (2 bytes):  
 01 00

Ticket (6 bytes):  
 00 04 ID ID ID ID

Extensions (2 bytes):  
 00 00

Auth Tag (8 bytes) // AES-CCM\_8:  
 00 01 02 03 04 05 06 07

$6 + 10 + 4 + 4 + 2 + 6 + 2 + 8 = 42$  bytes

The initial handshake when resumption is enabled is just a PSK handshake with  $134 + 150 + 57 + 42 = 383$  bytes.

#### 2.2.6. Without Connection ID

Without a Connection ID the DTLS 1.3 flight sizes changes as follows.

DTLS 1.3 Flight #1: -6 bytes  
 DTLS 1.3 Flight #2: -7 bytes  
 DTLS 1.3 Flight #3: -1 byte

Flight	#1	#2	#3	Total
DTLS 1.3 RPK + ECDHE (no cid)	144	364	212	722
DTLS 1.3 PSK + ECDHE (no cid)	178	183	56	417
DTLS 1.3 PSK (no cid)	128	143	56	327

Figure 4: Comparison of message sizes in bytes for DTLS 1.3 without Connection ID

## 2.2.7.    DTLS Raw Public Keys

TODO

## 2.2.7.1.    SubjectPublicKeyInfo without point compression

```
0x30 // Sequence
0x59 // Size 89

0x30 // Sequence
0x13 // Size 19
0x06 0x07 0x2A 0x86 0x48 0xCE 0x3D 0x02 0x01
      // OID 1.2.840.10045.2.1 (ecPublicKey)
0x06 0x08 0x2A 0x86 0x48 0xCE 0x3D 0x03 0x01 0x07
      // OID 1.2.840.10045.3.1.7 (secp256r1)

0x03 // Bit string
0x42 // Size 66
0x00 // Unused bits 0
0x04 // Uncompressed
..... 64 bytes X and Y

Total of 91 bytes
```

## 2.2.7.2.    SubjectPublicKeyInfo with point compression

```
0x30 // Sequence
0x59 // Size 89

0x30 // Sequence
0x13 // Size 19
0x06 0x07 0x2A 0x86 0x48 0xCE 0x3D 0x02 0x01
      // OID 1.2.840.10045.2.1 (ecPublicKey)
0x06 0x08 0x2A 0x86 0x48 0xCE 0x3D 0x03 0x01 0x07
      // OID 1.2.840.10045.3.1.7 (secp256r1)

0x03 // Bit string
0x42 // Size 66
0x00 // Unused bits 0
0x03 // Compressed
..... 32 bytes X

Total of 59 bytes
```

### 2.3. TLS 1.3

In this section, the message sizes are calculated for TLS 1.3. The major changes compared to DTLS 1.3 are that the record header is smaller, the handshake headers is smaller, and that Connection ID is not supported. Recently, additional work has taken shape with the goal to further reduce overhead for TLS 1.3 (see [I-D.schaad-ace-tls-cbor-handshake] ).

#### TLS Assumptions:

- o Minimum number of algorithms and cipher suites offered
- o Curve25519, ECDSA with P-256, AES-CCM\_8, SHA-256
- o Length of key identifiers: 1 bytes
- o TLS RPK with point compression (saves 32 bytes)
- o Only mandatory TLS extensions

For the PSK calculations, [Ulfheim-TLS13] was a useful resource, while for RPK calculations we followed the work of [IoT-Cert].

#### 2.3.1. Message Sizes RPK + ECDHE

##### 2.3.1.1. flight\_1

Record Header - TLSPlaintext (5 bytes):

16 03 03 LL LL

Handshake Header - Client Hello (4 bytes):

01 LL LL LL

Legacy Version (2 bytes):

03 03

Client Random (32 bytes):

00 01 02 03 04 05 06 07 08 09 0a 0b 0c 0d 0e 0f 10 11 12 13 14 15  
16 17 18 19 1a 1b 1c 1d 1e 1f

Legacy Session ID (1 bytes):

00

Cipher Suites (TLS\_AES\_128\_CCM\_8\_SHA256) (4 bytes):

00 02 13 05

Compression Methods (null) (2 bytes):

01 00

Extensions Length (2 bytes):

LL LL

Extension - Supported Groups (x25519) (8 bytes):

00 0a 00 04 00 02 00 1d

Extension - Signature Algorithms (ecdsa\_secp256r1\_sha256) (8 bytes):

00 0d 00 04 00 02 08 07

Extension - Key Share (42 bytes):

00 33 00 26 00 24 00 1d 00 20  
00 01 02 03 04 05 06 07 08 09 0a 0b 0c 0d 0e 0f 10 11 12 13 14 15  
16 17 18 19 1a 1b 1c 1d 1e 1f

Extension - Supported Versions (1.3) (7 bytes):

00 2b 00 03 02 03 04

Extension - Client Certificate Type (Raw Public Key) (6 bytes):

00 13 00 01 01 02

Extension - Server Certificate Type (Raw Public Key) (6 bytes):

00 14 00 01 01 02

5 + 4 + 2 + 32 + 1 + 4 + 2 + 2 + 8 + 8 + 42 + 7 + 6 + 6 = 129 bytes

TLS 1.3 RPK + ECDHE flight\_1 gives 129 bytes of overhead.

## 2.3.1.2.    flight\_2

Record Header - TLSPlaintext (5 bytes):

16 03 03 LL LL

Handshake Header - Server Hello (4 bytes):

02 LL LL LL

Legacy Version (2 bytes):

fe fd

Server Random (32 bytes):

00 01 02 03 04 05 06 07 08 09 0a 0b 0c 0d 0e 0f 10 11 12 13 14 15  
16 17 18 19 1a 1b 1c 1d 1e 1f

Legacy Session ID (1 bytes):

00

Cipher Suite (TLS\_AES\_128\_CCM\_8\_SHA256) (2 bytes):

13 05

Compression Method (null) (1 bytes):

00

Extensions Length (2 bytes):

LL LL

Extension - Key Share (40 bytes):

00 33 00 24 00 1d 00 20  
00 01 02 03 04 05 06 07 08 09 0a 0b 0c 0d 0e 0f 10 11 12 13 14 15  
16 17 18 19 1a 1b 1c 1d 1e 1f

Extension - Supported Versions (1.3) (6 bytes):

00 2b 00 02 03 04

Record Header - TLSCiphertext (5 bytes):

17 03 03 LL LL

Handshake Header - Encrypted Extensions (4 bytes):

08 LL LL LL

Extensions Length (2 bytes):

LL LL

Extension - Client Certificate Type (Raw Public Key) (6 bytes):

00 13 00 01 01 02

Extension - Server Certificate Type (Raw Public Key) (6 bytes):



00 14 00 01 01 02

Handshake Header - Certificate Request (4 bytes):

0d LL LL LL

Request Context (1 bytes):

00

Extensions Length (2 bytes):

LL LL

Extension - Signature Algorithms (ecdsa\_secp256r1\_sha256) (8 bytes):

00 0d 00 04 00 02 08 07

Handshake Header - Certificate (4 bytes):

0b LL LL LL

Request Context (1 bytes):

00

Certificate List Length (3 bytes):

LL LL LL

Certificate Length (3 bytes):

LL LL LL

Certificate (59 bytes) // Point compression

....

Certificate Extensions (2 bytes):

00 00

Handshake Header - Certificate Verify (4 bytes):

0f LL LL LL

Signature (68 bytes):

ZZ ZZ 00 40 ....

Handshake Header - Finished (4 bytes):

14 LL LL LL

Verify Data (32 bytes):

00 01 02 03 04 05 06 07 08 09 0a 0b 0c 0d 0e 0f 10 11 12 13 14 15  
16 17 18 19 1a 1b 1c 1d 1e 1f

Record Type (1 byte):

16

Auth Tag (8 bytes):  
e0 8b 0e 45 5a 35 0a e5

$5 + 90 + 5 + 18 + 15 + 72 + 72 + 36 + 1 + 8 = 322$  bytes

TLS 1.3 RPK + ECDHE flight\_2 gives 322 bytes of overhead.

2.3.1.3. flight\_3

Record Header - TLSCiphertext (5 bytes):

17 03 03 LL LL

Handshake Header - Certificate (4 bytes):

0b LL LL LL

Request Context (1 bytes):

00

Certificate List Length (3 bytes):

LL LL LL

Certificate Length (3 bytes):

LL LL LL

Certificate (59 bytes) // Point compression

....

Certificate Extensions (2 bytes):

00 00

Handshake Header - Certificate Verify (4 bytes):

0f LL LL LL

Signature (68 bytes):

04 03 LL LL //ecdsa\_secp256r1\_sha256

00 01 02 03 04 05 06 07 08 09 0a 0b 0c 0d 0e 0f 10 11 12 13 14 15

16 17 18 19 1a 1b 1c 1d 1e 1f 00 01 02 03 04 05 06 07 08 09 0a 0b

0c 0d 0e 0f 10 11 12 13 14 15 16 17 18 19 1a 1b 1c 1d 1e 1f

Handshake Header - Finished (4 bytes):

14 LL LL LL

Verify Data (32 bytes) // SHA-256:

00 01 02 03 04 05 06 07 08 09 0a 0b 0c 0d 0e 0f 10 11 12 13 14 15

16 17 18 19 1a 1b 1c 1d 1e 1f

Record Type (1 byte)

16

Auth Tag (8 bytes) // AES-CCM\_8:

00 01 02 03 04 05 06 07

5 + 72 + 72 + 36 + 1 + 8 = 194 bytes

TLS 1.3 RPK + ECDHE flight\_3 gives 194 bytes of overhead.

### 2.3.2. Message Sizes PSK + ECDHE

#### 2.3.2.1. flight\_1

The differences in overhead compared to Section 2.3.1.3 are:

The following is added:

+ Extension - PSK Key Exchange Modes (6 bytes):

00 2d 00 02 01 01

+ Extension - Pre Shared Key (48 bytes):

00 29 00 2f

00 0a 00 01 ID 00 00 00 00

00 21 20 00 01 02 03 04 05 06 07 08 09 0a 0b 0c 0d 0e 0f 10 11 12 13

14 15 16 17 18 19 1a 1b 1c 1d 1e 1f

The following is removed:

- Extension - Signature Algorithms (ecdsa\_secp256r1\_sha256) (8 bytes)

- Extension - Client Certificate Type (Raw Public Key) (6 bytes)

- Extension - Server Certificate Type (Raw Public Key) (6 bytes)

In total:

$129 + 6 + 48 - 8 - 6 - 6 = 163$  bytes

TLS 1.3 PSK + ECDHE flight\_1 gives 166 bytes of overhead.

#### 2.3.2.2. flight\_2

The differences in overhead compared to Section 2.3.1.2 are:

The following is added:

+ Extension - Pre Shared Key (6 bytes)

00 29 00 02 00 00

The following is removed:

- Handshake Message Certificate (72 bytes)
- Handshake Message CertificateVerify (72 bytes)
- Handshake Message CertificateRequest (15 bytes)
- Extension - Client Certificate Type (Raw Public Key) (6 bytes)
- Extension - Server Certificate Type (Raw Public Key) (6 bytes)

In total:

$322 - 72 - 72 - 15 - 6 - 6 + 6 = 157$  bytes

TLS 1.3 PSK + ECDHE flight\_2 gives 157 bytes of overhead.

#### 2.3.2.3. flight\_3

The differences in overhead compared to Section 2.3.1.3 are:

The following is removed:

- Handshake Message Certificate (72 bytes)
- Handshake Message Certificate Verify (72 bytes)

In total:

$194 - 72 - 72 = 50$  bytes

TLS 1.3 PSK + ECDHE flight\_3 gives 50 bytes of overhead.

#### 2.3.3. Message Sizes PSK

##### 2.3.3.1. flight\_1

The differences in overhead compared to Section 2.3.2.1 are:

The following is removed:

- Extension - Supported Groups (x25519) (8 bytes)
- Extension - Key Share (42 bytes)

In total:

$163 - 8 - 42 = 113$  bytes

TLS 1.3 PSK flight\_1 gives 116 bytes of overhead.

#### 2.3.3.2. flight\_2

The differences in overhead compared to Section 2.3.2.2 are:

The following is removed:

- Extension - Key Share (40 bytes)

In total:

$157 - 40 = 117$  bytes

TLS 1.3 PSK flight\_2 gives 117 bytes of overhead.

#### 2.3.3.3. flight\_3

There are no differences in overhead compared to Section 2.3.2.3.

TLS 1.3 PSK flight\_3 gives 57 bytes of overhead.

### 2.4. EDHOC

This section gives an estimate of the message sizes of EDHOC with different authentication methods. Note that the examples in this section are not test vectors, the cryptographic parts are just replaced with byte strings of the same length. All examples are given in CBOR diagnostic notation and hexadecimal.

#### 2.4.1. Message Sizes RPK

##### 2.4.1.1. message\_1

```
message_1 = (  
  1,  
  0,  
  h'000102030405060708090a0b0c0d0e0f101112131415161718191a1b1c1d  
    1e1f',  
  h'c3'  
)
```

message\_1 (38 bytes):

```
01 00 58 20 00 01 02 03 04 05 06 07 08 09 0A 0B 0C 0D 0E 0F  
10 11 12 13 14 15 16 17 18 19 1A 1B 1C 1D 1E 1F 41 C3
```

## 2.4.1.2. message\_2

```
plaintext = <<
  h'al',
  h'000102030405060708090a0b0c0d0e0f101112131415161718191a1b1c1d
  1e1f202122232425262728292a2b2c2d2e2f303132333435363738393a3b
  3c3d3e3f'
>>
```

The header map { 4 : h'al' } is encoded as the two bytes h'al'. The length of plaintext is 68 bytes so assuming a 64-bit MAC value the length of ciphertext is 76 bytes.

```
message_2 = (
  h'000102030405060708090a0b0c0d0e0f101112131415161718191a1b1c1d
  1e1f',
  h'c4',
  h'000102030405060708090a0b0c0d0e0f101112131415161718191a1b1c1d
  1e1f202122232425262728292a2b2c2d2e2f303132333435363738393a3b
  3c3d3e3f404142434445464748494a4b'
)
```

message\_2 (114 bytes):

```
58 20 00 01 02 03 04 05 06 07 08 09 0A 0B 0C 0D 0E 0F 10 11
12 13 14 15 16 17 18 19 1A 1B 1C 1D 1E 1F 41 C4 58 51 00 01
02 03 04 05 06 07 08 09 0A 0B 0C 0D 0E 0F 10 11 12 13 14 15
16 17 18 19 1A 1B 1C 1D 1E 1F 20 21 22 23 24 25 26 27 28 29
2A 2B 2C 2D 2E 2F 30 31 32 33 34 35 36 37 38 39 3A 3B 3C 3D
3E 3F 40 41 42 43 44 45 46 47 48 49 4A 4B
```

## 2.4.1.3. message\_3

The plaintext and ciphertext in message\_3 are assumed to be of equal sizes as in message\_2.

```
message_3 = (
  h'c4',
  h'000102030405060708090a0b0c0d0e0f101112131415161718191a1b1c1d
  1e1f202122232425262728292a2b2c2d2e2f303132333435363738393a3b
  3c3d3e3f404142434445464748494a4b'
)
```

message\_3 (80 bytes):

```
41 C4 58 51 00 01 02 03 04 05 06 07 08 09 0A 0B 0C 0D 0E 0F
10 11 12 13 14 15 16 17 18 19 1A 1B 1C 1D 1E 1F 20 21 22 23
24 25 26 27 28 29 2A 2B 2C 2D 2E 2F 30 31 32 33 34 35 36 37
38 39 3A 3B 3C 3D 3E 3F 40 41 42 43 44 45 46 47 48 49 4A 4B
```

#### 2.4.2. Message Sizes Certificates

When the certificates are distributed out-of-band and identified with the x5t header parameter and a SHA256/64 hash value, the header map will be 13 bytes (assuming labels in the range -24...23).

```
{ TDB1 : [ TDB6, h'0001020304050607' ] }
```

When the certificates are identified with the x5chain header parameter, the message sizes depends on the size of the (truncated) certificate chains. The header map will be 3 bytes + the size of the certificate chain (assuming a label in the range -24...23).

```
{ TDB3 : h'0001020304050607...' }
```

#### 2.4.3. Message Sizes PSK

##### 2.4.4. message\_1

```
message_1 = (
    4,
    0,
    h'000102030405060708090a0b0c0d0e0f101112131415161718191a1b1c1d
    1e1f',
    h'c3',
    h'a2'
)
```

message\_1 (40 bytes):

```
04 00 58 20 00 01 02 03 04 05 06 07 08 09 0A 0B 0C 0D 0E 0F
10 11 12 13 14 15 16 17 18 19 1A 1B 1C 1D 1E 1F 41 C3 41 A2
```

##### 2.4.5. message\_2

Assuming a 0 byte plaintext and a 64-bit MAC value the ciphertext is 8 bytes

```
message_2 = (
    h'000102030405060708090a0b0c0d0e0f101112131415161718191a1b1c1d
    1e1f',
    h'c4',
    h'0001020304050607'
)
```

message\_2 (45 bytes):

```
58 20 00 01 02 03 04 05 06 07 08 09 0A 0B 0C 0D 0E 0F 10 11
12 13 14 15 16 17 18 19 1A 1B 1C 1D 1E 1F 41 C4 48 61 62 63
64 65 66 67 68
```



## 2.4.6. message\_3

The plaintext and ciphertext in message\_3 are assumed to be of equal sizes as in message\_2.

```
message_3 = (
    h'c4',
    h'0001020304050607'
)
```

```
message_3 (11 bytes):
41 C4 48 00 01 02 03 04 05 06 07
```

## 2.4.7. Summary

The previous examples of typical message sizes are summarized in Figure 5.

	PSK	RPK	x5t	x5chain
message_1	40	38	38	38
message_2	45	114	126	116 + Certificate chain
message_3	11	80	91	81 + Certificate chain
Total	96	232	255	235 + Certificate chains

Figure 5: Typical message sizes in bytes

## 2.5. Conclusion

To do a fair comparison, one has to choose a specific deployment and look at the topology, the whole protocol stack, frame sizes (e.g. 51 or 128 bytes), how and where in the protocol stack fragmentation is done, and the expected packet loss. Note that the number of byte in each frame that is available for the key exchange protocol may depend on the underlying protocol layers as well as the number of hops in multi-hop networks. The packet loss depends may depend on how many other devices that are transmitting at the same time, and may increase during network formation. The total overhead will be larger due to mechanisms for fragmentation, retransmission, and packet ordering. The overhead of fragmentation is roughly proportional to the number of fragments, while the expected overhead due to retransmission in noisy environments is a superlinear function of the flight sizes.

### 3. Overhead for Protection of Application Data

To enable comparison, all the overhead calculations in this section use AES-CCM with a tag length of 8 bytes (e.g. AES\_128\_CCM\_8 or AES-CCM-16-64), a plaintext of 6 bytes, and the sequence number '05'. This follows the example in [RFC7400], Figure 16.

Note that the compressed overhead calculations for DLTS 1.2, DTLS 1.3, TLS 1.2 and TLS 1.3 are dependent on the parameters epoch, sequence number, and length, and all the overhead calculations are dependent on the parameter Connection ID when used. Note that the OSCORE overhead calculations are dependent on the CoAP option numbers, as well as the length of the OSCORE parameters Sender ID and Sequence Number. The following calculations are only examples.

Section 3.1 gives a short summary of the message overhead based on different parameters and some assumptions. The following sections detail the assumptions and the calculations.

#### 3.1. Summary

The DTLS overhead is dependent on the parameter Connection ID. The following overheads apply for all Connection IDs with the same length.

The compression overhead (GHC) is dependent on the parameters epoch, sequence number, Connection ID, and length (where applicable). The following overheads should be representative for sequence numbers and Connection IDs with the same length.

The OSCORE overhead is dependent on the included CoAP Option numbers as well as the length of the OSCORE parameters Sender ID and sequence number. The following overheads apply for all sequence numbers and Sender IDs with the same length.

Sequence Number	'05'	'1005'	'100005'
DTLS 1.2	29	29	29
DTLS 1.3	11	12	12
DTLS 1.2 (GHC)	16	16	16
DTLS 1.3 (GHC)	12	13	13
TLS 1.2	21	21	21
TLS 1.3	14	14	14
TLS 1.2 (GHC)	17	18	19
TLS 1.3 (GHC)	15	16	17
OSCORE request	13	14	15
OSCORE response	11	11	11

Figure 6: Overhead in bytes as a function of sequence number  
(Connection/Sender ID = '')

Connection/Sender ID	' '	'42'	'4002'
DTLS 1.2	29	30	31
DTLS 1.3	11	12	13
DTLS 1.2 (GHC)	16	17	18
DTLS 1.3 (GHC)	12	13	14
OSCORE request	13	14	15
OSCORE response	11	11	11

Figure 7: Overhead in bytes as a function of Connection/Sender ID  
(Sequence Number = '05')

Protocol	Overhead	Overhead (GHC)
DTLS 1.2	21	8
DTLS 1.3	3	4
TLS 1.2	13	9
TLS 1.3	6	7
OSCORE request	5	
OSCORE response	3	

Figure 8: Overhead (excluding ICV) in bytes  
(Connection/Sender ID = '', Sequence Number = '05')

### 3.2. DTLS 1.2

#### 3.2.1. DTLS 1.2

This section analyzes the overhead of DTLS 1.2 [RFC6347]. The nonce follow the strict profiling given in [RFC7925]. This example is taken directly from [RFC7400], Figure 16.

DTLS 1.2 record layer (35 bytes, 29 bytes overhead):

```
17 fe fd 00 01 00 00 00 00 00 05 00 16 00 01 00
00 00 00 00 05 ae a0 15 56 67 92 4d ff 8a 24 e4
cb 35 b9
```

Content type:

17

Version:

fe fd

Epoch:

00 01

Sequence number:

00 00 00 00 00 05

Length:

00 16

Nonce:

00 01 00 00 00 00 00 05

Ciphertext:

ae a0 15 56 67 92

ICV:

4d ff 8a 24 e4 cb 35 b9

DTLS 1.2 gives 29 bytes overhead.

#### 3.2.2. DTLS 1.2 with 6LoWPAN-GHC

This section analyzes the overhead of DTLS 1.2 [RFC6347] when compressed with 6LoWPAN-GHC [RFC7400]. The compression was done with [OlegHahm-ghc].

Note that the sequence number '01' used in [RFC7400], Figure 15 gives an exceptionally small overhead that is not representative.

Note that this header compression is not available when DTLS is used over transports that do not use 6LoWPAN together with 6LoWPAN-GHC.

Compressed DTLS 1.2 record layer (22 bytes, 16 bytes overhead):

```
b0 c3 03 05 00 16 f2 0e ae a0 15 56 67 92 4d ff
8a 24 e4 cb 35 b9
```

Compressed DTLS 1.2 record layer header and nonce:

```
b0 c3 03 05 00 16 f2 0e
```

Ciphertext:

```
ae a0 15 56 67 92
```

ICV:

```
4d ff 8a 24 e4 cb 35 b9
```

When compressed with 6LoWPAN-GHC, DTLS 1.2 with the above parameters (epoch, sequence number, length) gives 16 bytes overhead.

### 3.2.3. DTLS 1.2 with Connection ID

This section analyzes the overhead of DTLS 1.2 [RFC6347] with Connection ID [I-D.ietf-tls-dtls-connection-id]. The overhead calculations in this section uses Connection ID = '42'. DTLS record layer with a Connection ID = '' (the empty string) is equal to DTLS without Connection ID.

DTLS 1.2 record layer (36 bytes, 30 bytes overhead):

```
17 fe fd 00 01 00 00 00 00 00 05 42 00 16 00 01
00 00 00 00 00 05 ae a0 15 56 67 92 4d ff 8a 24
e4 cb 35 b9
```

Content type:

```
17
```

Version:

```
fe fd
```

Epoch:

```
00 01
```

Sequence number:

```
00 00 00 00 00 05
```

Connection ID:

```
42
```

Length:

```
00 16
```

Nonce:

```
00 01 00 00 00 00 00 05
```

Ciphertext:

```
ae a0 15 56 67 92
```

ICV:

```
4d ff 8a 24 e4 cb 35 b9
```

DTLS 1.2 with Connection ID gives 30 bytes overhead.

#### 3.2.4. DTLS 1.2 with Connection ID and 6LoWPAN-GHC

This section analyzes the overhead of DTLS 1.2 [RFC6347] with Connection ID [I-D.ietf-tls-dtls-connection-id] when compressed with 6LoWPAN-GHC [RFC7400] [OlegHahm-ghc].

Note that the sequence number '01' used in [RFC7400], Figure 15 gives an exceptionally small overhead that is not representative.

Note that this header compression is not available when DTLS is used over transports that do not use 6LoWPAN together with 6LoWPAN-GHC.

Compressed DTLS 1.2 record layer (23 bytes, 17 bytes overhead):  
b0 c3 04 05 42 00 16 f2 0e ae a0 15 56 67 92 4d  
ff 8a 24 e4 cb 35 b9

Compressed DTLS 1.2 record layer header and nonce:  
b0 c3 04 05 42 00 16 f2 0e  
Ciphertext:  
ae a0 15 56 67 92  
ICV:  
4d ff 8a 24 e4 cb 35 b9

When compressed with 6LoWPAN-GHC, DTLS 1.2 with the above parameters (epoch, sequence number, Connection ID, length) gives 17 bytes overhead.

#### 3.3. DTLS 1.3

##### 3.3.1. DTLS 1.3

This section analyzes the overhead of DTLS 1.3 [I-D.ietf-tls-dtls13]. The changes compared to DTLS 1.2 are: omission of version number, merging of epoch into the first byte containing signalling bits, optional omission of length, reduction of sequence number into a 1 or 2-bytes field.

Only the minimal header format for DTLS 1.3 is analyzed (see Figure 4 of [I-D.ietf-tls-dtls13]). The minimal header format omits the length field and only a 1-byte field is used to carry the 8 low order bits of the sequence number.

DTLS 1.3 record layer (17 bytes, 11 bytes overhead):  
21 05 ae a0 15 56 67 92 ec 4d ff 8a 24 e4 cb 35 b9

First byte (including epoch):

21

Sequence number:

05

Ciphertext (including encrypted content type):

ae a0 15 56 67 92 ec

ICV:

4d ff 8a 24 e4 cb 35 b9

DTLS 1.3 gives 11 bytes overhead.

### 3.3.2. DTLS 1.3 with 6LoWPAN-GHC

This section analyzes the overhead of DTLS 1.3 [I-D.ietf-tls-dtls13] when compressed with 6LoWPAN-GHC [RFC7400] [OlegHahm-ghc].

Note that this header compression is not available when DTLS is used over transports that do not use 6LoWPAN together with 6LoWPAN-GHC.

Compressed DTLS 1.3 record layer (18 bytes, 12 bytes overhead):

11 21 05 ae a0 15 56 67 92 ec 4d ff 8a 24 e4 cb  
35 b9

Compressed DTLS 1.3 record layer header and nonce:

11 21 05

Ciphertext (including encrypted content type):

ae a0 15 56 67 92 ec

ICV:

4d ff 8a 24 e4 cb 35 b9

When compressed with 6LoWPAN-GHC, DTLS 1.3 with the above parameters (epoch, sequence number, no length) gives 12 bytes overhead.

### 3.3.3. DTLS 1.3 with Connection ID

This section analyzes the overhead of DTLS 1.3 [I-D.ietf-tls-dtls13] with Connection ID [I-D.ietf-tls-dtls-connection-id].

In this example, the length field is omitted, and the 1-byte field is used for the sequence number. The minimal DTLSCiphertext structure is used (see Figure 4 of [I-D.ietf-tls-dtls13]), with the addition of the Connection ID field.

DTLS 1.3 record layer (18 bytes, 12 bytes overhead):  
31 42 05 ae a0 15 56 67 92 ec 4d ff 8a 24 e4 cb 35 b9

First byte (including epoch):

31

Connection ID:

42

Sequence number:

05

Ciphertext (including encrypted content type):

ae a0 15 56 67 92 ec

ICV:

4d ff 8a 24 e4 cb 35 b9

DTLS 1.3 with Connection ID gives 12 bytes overhead.

#### 3.3.4. DTLS 1.3 with Connection ID and 6LoWPAN-GHC

This section analyzes the overhead of DTLS 1.3 [I-D.ietf-tls-dtls13] with Connection ID [I-D.ietf-tls-dtls-connection-id] when compressed with 6LoWPAN-GHC [RFC7400] [OlegHahm-ghc].

Note that this header compression is not available when DTLS is used over transports that do not use 6LoWPAN together with 6LoWPAN-GHC.

Compressed DTLS 1.3 record layer (19 bytes, 13 bytes overhead):

12 31 05 42 ae a0 15 56 67 92 ec 4d ff 8a 24 e4  
cb 35 b9

Compressed DTLS 1.3 record layer header and nonce:

12 31 05 42

Ciphertext (including encrypted content type):

ae a0 15 56 67 92 ec

ICV:

4d ff 8a 24 e4 cb 35 b9

When compressed with 6LoWPAN-GHC, DTLS 1.3 with the above parameters (epoch, sequence number, Connection ID, no length) gives 13 bytes overhead.

#### 3.4. TLS 1.2

##### 3.4.1. TLS 1.2

This section analyzes the overhead of TLS 1.2 [RFC5246]. The changes compared to DTLS 1.2 is that the TLS 1.2 record layer does not have epoch and sequence number, and that the version is different.



TLS 1.2 Record Layer (27 bytes, 21 bytes overhead):  
17 03 03 00 16 00 00 00 00 00 00 05 ae a0 15  
56 67 92 4d ff 8a 24 e4 cb 35 b9

Content type:

17

Version:

03 03

Length:

00 16

Nonce:

00 00 00 00 00 00 00 05

Ciphertext:

ae a0 15 56 67 92

ICV:

4d ff 8a 24 e4 cb 35 b9

TLS 1.2 gives 21 bytes overhead.

#### 3.4.2. TLS 1.2 with 6LoWPAN-GHC

This section analyzes the overhead of TLS 1.2 [RFC5246] when compressed with 6LoWPAN-GHC [RFC7400] [OlegHahm-ghc].

Note that this header compression is not available when TLS is used over transports that do not use 6LoWPAN together with 6LoWPAN-GHC.

Compressed TLS 1.2 record layer (23 bytes, 17 bytes overhead):  
05 17 03 03 00 16 85 0f 05 ae a0 15 56 67 92 4d  
ff 8a 24 e4 cb 35 b9

Compressed TLS 1.2 record layer header and nonce:

05 17 03 03 00 16 85 0f 05

Ciphertext:

ae a0 15 56 67 92

ICV:

4d ff 8a 24 e4 cb 35 b9

When compressed with 6LoWPAN-GHC, TLS 1.2 with the above parameters (epoch, sequence number, length) gives 17 bytes overhead.

#### 3.5. TLS 1.3

##### 3.5.1. TLS 1.3

This section analyzes the overhead of TLS 1.3 [RFC8446]. The change compared to TLS 1.2 is that the TLS 1.3 record layer uses a different version.

TLS 1.3 Record Layer (20 bytes, 14 bytes overhead):  
17 03 03 00 16 ae a0 15 56 67 92 ec 4d ff 8a 24  
e4 cb 35 b9

Content type:

17

Legacy version:

03 03

Length:

00 0f

Ciphertext (including encrypted content type):

ae a0 15 56 67 92 ec

ICV:

4d ff 8a 24 e4 cb 35 b9

TLS 1.3 gives 14 bytes overhead.

### 3.5.2. TLS 1.3 with 6LoWPAN-GHC

This section analyzes the overhead of TLS 1.3 [RFC8446] when compressed with 6LoWPAN-GHC [RFC7400] [OlegHahm-ghc].

Note that this header compression is not available when TLS is used over transports that do not use 6LoWPAN together with 6LoWPAN-GHC.

Compressed TLS 1.3 record layer (21 bytes, 15 bytes overhead):  
14 17 03 03 00 0f ae a0 15 56 67 92 ec 4d ff 8a  
24 e4 cb 35 b9

Compressed TLS 1.3 record layer header and nonce:

14 17 03 03 00 0f

Ciphertext (including encrypted content type):

ae a0 15 56 67 92 ec

ICV:

4d ff 8a 24 e4 cb 35 b9

When compressed with 6LoWPAN-GHC, TLS 1.3 with the above parameters (epoch, sequence number, length) gives 15 bytes overhead.

### 3.6. OSCORE

This section analyzes the overhead of OSCORE [I-D.ietf-core-object-security].

The below calculation Option Delta = '9', Sender ID = '' (empty string), and Sequence Number = '05', and is only an example. Note that Sender ID = '' (empty string) can only be used by one client per server.

OSCORE request (19 bytes, 13 bytes overhead):  
92 09 05  
ff ec ae a0 15 56 67 92 4d ff 8a 24 e4 cb 35 b9

CoAP option delta and length:  
92  
Option value (flag byte and sequence number):  
09 05  
Payload marker:  
ff  
Ciphertext (including encrypted code):  
ec ae a0 15 56 67 92  
ICV:  
4d ff 8a 24 e4 cb 35 b9

The below calculation Option Delta = '9', Sender ID = '42', and Sequence Number = '05', and is only an example.

OSCORE request (20 bytes, 14 bytes overhead):  
93 09 05 42  
ff ec ae a0 15 56 67 92 4d ff 8a 24 e4 cb 35 b9

CoAP option delta and length:  
93  
Option Value (flag byte, sequence number, and Sender ID):  
09 05 42  
Payload marker:  
ff  
Ciphertext (including encrypted code):  
ec ae a0 15 56 67 92  
ICV:  
4d ff 8a 24 e4 cb 35 b9

The below calculation uses Option Delta = '9'.

OSCORE response (17 bytes, 11 bytes overhead):  
90  
ff ec ae a0 15 56 67 92 4d ff 8a 24 e4 cb 35 b9

CoAP delta and option length:  
90

Option value:  
-

Payload marker:  
ff

Ciphertext (including encrypted code):  
ec ae a0 15 56 67 92

ICV:  
4d ff 8a 24 e4 cb 35 b9

OSCORE with the above parameters gives 13-14 bytes overhead for requests and 11 bytes overhead for responses.

Unlike DTLS and TLS, OSCORE has much smaller overhead for responses than requests.

### 3.7. Group OSCORE

This section analyzes the overhead of Group OSCORE [I-D.ietf-core-oscore-groupcomm].

TODO

### 3.8. Conclusion

DTLS 1.2 has quite a large overhead as it uses an explicit sequence number and an explicit nonce. TLS 1.2 has significantly less (but not small) overhead. TLS 1.3 has quite a small overhead. OSCORE and DTLS 1.3 (using the minimal structure) format have very small overhead.

The Generic Header Compression (6LoWPAN-GHC) can in addition to DTLS 1.2 handle TLS 1.2, and DTLS 1.2 with Connection ID. The Generic Header Compression (6LoWPAN-GHC) works very well for Connection ID and the overhead seems to increase exactly with the length of the Connection ID (which is optimal). The compression of TLS 1.2 is not as good as the compression of DTLS 1.2 (as the static dictionary only contains the DTLS 1.2 version number). Similar compression levels as for DTLS could be achieved also for TLS 1.2, but this would require different static dictionaries. For TLS 1.3 and DTLS 1.3, GHC increases the overhead. The 6LoWPAN-GHC header compression is not available when (D)TLS is used over transports that do not use 6LoWPAN together with 6LoWPAN-GHC.

New security protocols like OSCORE, TLS 1.3, and DTLS 1.3 have much lower overhead than DTLS 1.2 and TLS 1.2. The overhead is even smaller than DTLS 1.2 and TLS 1.2 over 6LoWPAN with compression, and therefore the small overhead is achieved even on deployments without 6LoWPAN or 6LoWPAN without compression. OSCORE is lightweight because it makes use of CoAP, CBOR, and COSE, which were designed to have as low overhead as possible.

Note that the compared protocols have slightly different use cases. TLS and DTLS are designed for the transport layer and are terminated in CoAP proxies. OSCORE is designed for the application layer and protects information end-to-end between the CoAP client and the CoAP server. Group OSCORE is designed for group communication and protects information between a CoAP client and any number of CoAP servers.

#### 4. Security Considerations

This document is purely informational.

#### 5. IANA Considerations

This document has no actions for IANA.

#### 6. Informative References

[I-D.ietf-core-object-security]

Selander, G., Mattsson, J., Palombini, F., and L. Seitz, "Object Security for Constrained RESTful Environments (OSCORE)", draft-ietf-core-object-security-15 (work in progress), August 2018.

[I-D.ietf-core-oscore-groupcomm]

Tiloca, M., Selander, G., Palombini, F., and J. Park, "Group OSCORE - Secure Group Communication for CoAP", draft-ietf-core-oscore-groupcomm-03 (work in progress), October 2018.

[I-D.ietf-tls-dtls-connection-id]

Rescorla, E., Tschofenig, H., Fossati, T., and T. Gondrom, "Connection Identifiers for DTLS 1.2", draft-ietf-tls-dtls-connection-id-02 (work in progress), October 2018.

[I-D.ietf-tls-dtls13]

Rescorla, E., Tschofenig, H., and N. Modadugu, "The Datagram Transport Layer Security (DTLS) Protocol Version 1.3", draft-ietf-tls-dtls13-30 (work in progress), November 2018.

- [I-D.schaad-ace-tls-cbor-handshake]  
Schaad, J., "TLS Handshake in CBOR", draft-schaad-ace-tls-cbor-handshake-00 (work in progress), March 2019.
- [I-D.selander-ace-cose-ecdhe]  
Selander, G., Mattsson, J., and F. Palombini, "Ephemeral Diffie-Hellman Over COSE (EDHOC)", draft-selander-ace-cose-ecdhe-12 (work in progress), February 2019.
- [IoT-Cert]  
Forsby, F., "Digital Certificates for the Internet of Things", June 2017, <<https://kth.diva-portal.org/smash/get/diva2:1153958/FULLTEXT01.pdf>>.
- [OlegHahm-ghc]  
Hahm, O., "Generic Header Compression", July 2016, <<https://github.com/OlegHahm/ghc>>.
- [RFC5246] Dierks, T. and E. Rescorla, "The Transport Layer Security (TLS) Protocol Version 1.2", RFC 5246, DOI 10.17487/RFC5246, August 2008, <<https://www.rfc-editor.org/info/rfc5246>>.
- [RFC6347] Rescorla, E. and N. Modadugu, "Datagram Transport Layer Security Version 1.2", RFC 6347, DOI 10.17487/RFC6347, January 2012, <<https://www.rfc-editor.org/info/rfc6347>>.
- [RFC7252] Shelby, Z., Hartke, K., and C. Bormann, "The Constrained Application Protocol (CoAP)", RFC 7252, DOI 10.17487/RFC7252, June 2014, <<https://www.rfc-editor.org/info/rfc7252>>.
- [RFC7400] Bormann, C., "6LoWPAN-GHC: Generic Header Compression for IPv6 over Low-Power Wireless Personal Area Networks (6LoWPANs)", RFC 7400, DOI 10.17487/RFC7400, November 2014, <<https://www.rfc-editor.org/info/rfc7400>>.
- [RFC7924] Santesson, S. and H. Tschofenig, "Transport Layer Security (TLS) Cached Information Extension", RFC 7924, DOI 10.17487/RFC7924, July 2016, <<https://www.rfc-editor.org/info/rfc7924>>.
- [RFC7925] Tschofenig, H., Ed. and T. Fossati, "Transport Layer Security (TLS) / Datagram Transport Layer Security (DTLS) Profiles for the Internet of Things", RFC 7925, DOI 10.17487/RFC7925, July 2016, <<https://www.rfc-editor.org/info/rfc7925>>.

- [RFC8323] Bormann, C., Lemay, S., Tschofenig, H., Hartke, K., Silverajan, B., and B. Raymor, Ed., "CoAP (Constrained Application Protocol) over TCP, TLS, and WebSockets", RFC 8323, DOI 10.17487/RFC8323, February 2018, <<https://www.rfc-editor.org/info/rfc8323>>.
- [RFC8446] Rescorla, E., "The Transport Layer Security (TLS) Protocol Version 1.3", RFC 8446, DOI 10.17487/RFC8446, August 2018, <<https://www.rfc-editor.org/info/rfc8446>>.
- [Ulfheim-TLS13] Driscoll, M., "Every Byte Explained The Illustrated TLS 1.3 Connection", March 2018, <<https://tls13.ulfheim.net>>.

#### Acknowledgments

The authors want to thank Ari Keraenen, Carsten Bormann, Goeran Selander, and Hannes Tschofenig for comments and suggestions on previous versions of the draft.

All 6LoWPAN-GHC compression was done with [OlegHahm-ghc].

#### Authors' Addresses

John Mattsson  
Ericsson AB

Email: [john.mattsson@ericsson.com](mailto:john.mattsson@ericsson.com)

Francesca Palombini  
Ericsson AB

Email: [francesca.palombini@ericsson.com](mailto:francesca.palombini@ericsson.com)

LWIG Working Group  
Internet-Draft  
Intended status: Informational  
Expires: May 6, 2021

J. Mattsson  
F. Palombini  
Ericsson AB  
M. Vucinic  
INRIA  
November 02, 2020

Comparison of CoAP Security Protocols  
draft-ietf-lwig-security-protocol-comparison-05

Abstract

This document analyzes and compares the sizes of key exchange flights and the per-packet message size overheads when using different security protocols to secure CoAP. The analyzed security protocols are DTLS 1.2, DTLS 1.3, TLS 1.2, TLS 1.3, EDHOC, OSCORE, and Group OSCORE. The DTLS and TLS record layers are analyzed with and without 6LoWPAN-GHC compression. DTLS is analyzed with and without Connection ID.

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 May 6, 2021.

Copyright Notice

Copyright (c) 2020 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 Provisions Relating to IETF Documents (<https://trustee.ietf.org/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 Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

## Table of Contents

1. Introduction . . . . .	3
2. Overhead of Key Exchange Protocols . . . . .	3
2.1. Summary . . . . .	4
2.2. DTLS 1.3 . . . . .	5
2.2.1. Message Sizes RPK + ECDHE . . . . .	5
2.2.2. Message Sizes PSK + ECDHE . . . . .	10
2.2.3. Message Sizes PSK . . . . .	11
2.2.4. Cached Information . . . . .	12
2.2.5. Resumption . . . . .	13
2.2.6. Without Connection ID . . . . .	14
2.2.7. DTLS Raw Public Keys . . . . .	15
2.3. TLS 1.3 . . . . .	16
2.3.1. Message Sizes RPK + ECDHE . . . . .	16
2.3.2. Message Sizes PSK + ECDHE . . . . .	22
2.3.3. Message Sizes PSK . . . . .	23
2.4. EDHOC . . . . .	24
2.4.1. Message Sizes RPK . . . . .	24
2.4.2. Summary . . . . .	25
2.5. Conclusion . . . . .	25
3. Overhead for Protection of Application Data . . . . .	26
3.1. Summary . . . . .	26
3.2. DTLS 1.2 . . . . .	28
3.2.1. DTLS 1.2 . . . . .	28
3.2.2. DTLS 1.2 with 6LoWPAN-GHC . . . . .	28
3.2.3. DTLS 1.2 with Connection ID . . . . .	29
3.2.4. DTLS 1.2 with Connection ID and 6LoWPAN-GHC . . . . .	30
3.3. DTLS 1.3 . . . . .	30
3.3.1. DTLS 1.3 . . . . .	30
3.3.2. DTLS 1.3 with 6LoWPAN-GHC . . . . .	31
3.3.3. DTLS 1.3 with Connection ID . . . . .	31
3.3.4. DTLS 1.3 with Connection ID and 6LoWPAN-GHC . . . . .	32
3.4. TLS 1.2 . . . . .	32
3.4.1. TLS 1.2 . . . . .	32
3.4.2. TLS 1.2 with 6LoWPAN-GHC . . . . .	33
3.5. TLS 1.3 . . . . .	33
3.5.1. TLS 1.3 . . . . .	33
3.5.2. TLS 1.3 with 6LoWPAN-GHC . . . . .	34
3.6. OSCORE . . . . .	34
3.7. Group OSCORE . . . . .	36
3.8. Conclusion . . . . .	36
4. Security Considerations . . . . .	37

5. IANA Considerations . . . . .	37
6. Informative References . . . . .	37
Acknowledgments . . . . .	39
Authors' Addresses . . . . .	39

## 1. Introduction

This document analyzes and compares the sizes of key exchange flights and the per-packet message size overheads when using different security protocols to secure CoAP over UDP [RFC7252] and TCP [RFC8323]. The analyzed security protocols are DTLS 1.2 [RFC6347], DTLS 1.3 [I-D.ietf-tls-dtls13], TLS 1.2 [RFC5246], TLS 1.3 [RFC8446], EDHOC [I-D.ietf-lake-edhoc], OSCORE [RFC8613], and Group OSCORE [I-D.ietf-core-oscore-groupcomm].

The DTLS and TLS record layers are analyzed with and without 6LoWPAN-GHC compression. DTLS is analyzed with and without Connection ID [I-D.ietf-tls-dtls-connection-id]. Readers are expected to be familiar with some of the terms described in RFC 7925 [RFC7925], such as ICV. Section 2 compares the overhead of key exchange, while Section 3 covers the overhead for protection of application data.

## 2. Overhead of Key Exchange Protocols

This section analyzes and compares the sizes of key exchange flights for different protocols.

To enable a fair comparison between protocols, the following assumptions are made:

- o All the overhead calculations in this section use AES-CCM with a tag length of 8 bytes (e.g. AES\_128\_CCM\_8 or AES-CCM-16-64-128).
- o A minimum number of algorithms and cipher suites is offered. The algorithm used/offered are Curve25519, ECDSA with P-256, AES-CCM\_8, SHA-256.
- o The length of key identifiers are 1 byte.
- o The length of connection identifiers are 1 byte.
- o DTLS RPK makes use of point compression, which saves 32 bytes.
- o DTLS handshake message fragmentation is not considered.
- o Only the DTLS mandatory extensions are considered, except for Connection ID.

Section 2.1 gives a short summary of the message overhead based on different parameters and some assumptions. The following sections detail the assumptions and the calculations.

## 2.1. Summary

The DTLS overhead is dependent on the parameter Connection ID. The following overheads apply for all Connection IDs of the same length, when Connection ID is used.

The EDHOC overhead is dependent on the key identifiers included. The following overheads apply for Sender IDs of the same length.

All the overhead are dependent on the tag length. The following overheads apply for tags of the same length.

Figure 1 compares the message sizes of EDHOC [I-D.ietf-lake-edhoc] with the DTLS 1.3 [I-D.ietf-tls-dtls13] and TLS 1.3 [RFC8446] handshakes with connection ID.

Flight	#1	#2	#3	Total
DTLS 1.3 RPK + ECDHE	150	373	213	736
DTLS 1.3 Cached X.509/RPK + ECDHE	182	347	213	742
DTLS 1.3 PSK + ECDHE	184	190	57	431
DTLS 1.3 PSK	134	150	57	341
EDHOC RPK + ECDHE	37	46	20	103
EDHOC X.509 + ECDHE	37	117	91	245

Figure 1: Comparison of message sizes in bytes with Connection ID

Figure 2 compares of message sizes of DTLS 1.3 [I-D.ietf-tls-dtls13] and TLS 1.3 [RFC8446] handshakes without connection ID.

Flight	#1	#2	#3	Total
DTLS 1.3 RPK + ECDHE	144	364	212	722
DTLS 1.3 PSK + ECDHE	178	183	56	417
DTLS 1.3 PSK	128	143	56	327
TLS 1.3 RPK + ECDHE	129	322	194	645
TLS 1.3 PSK + ECDHE	163	157	50	370
TLS 1.3 PSK	113	117	50	280

Figure 2: Comparison of message sizes in bytes without Connection ID

The details of the message size calculations are given in the following sections.

## 2.2. DTLS 1.3

This section gives an estimate of the message sizes of DTLS 1.3 with different authentication methods. Note that the examples in this section are not test vectors, the cryptographic parts are just replaced with byte strings of the same length, while other fixed length fields are replaced with arbitrary strings or omitted, in which case their length is indicated. Values that are not arbitrary are given in hexadecimal.

### 2.2.1. Message Sizes RPK + ECDHE

In this section, a Connection ID of 1 byte is used.

#### 2.2.1.1. flight\_1

Record Header - DTLSPlaintext (13 bytes):

16 fe fd EE EE SS SS SS SS SS SS LL LL

Handshake Header - Client Hello (10 bytes):

01 LL LL LL SS SS 00 00 00 LL LL LL

Legacy Version (2 bytes):

fe fd

Client Random (32 bytes):

00 01 02 03 04 05 06 07 08 09 0a 0b 0c 0d 0e 0f 10 11 12 13 14 15  
16 17 18 19 1a 1b 1c 1d 1e 1f

Legacy Session ID (1 bytes):

00

Legacy Cookie (1 bytes):  
00

Cipher Suites (TLS\_AES\_128\_CCM\_8\_SHA256) (4 bytes):  
00 02 13 05

Compression Methods (null) (2 bytes):  
01 00

Extensions Length (2 bytes):  
LL LL

Extension - Supported Groups (x25519) (8 bytes):  
00 0a 00 04 00 02 00 1d

Extension - Signature Algorithms (ecdsa\_secp256r1\_sha256)  
(8 bytes):  
00 0d 00 04 00 02 08 07

Extension - Key Share (42 bytes):  
00 33 00 26 00 24 00 1d 00 20  
00 01 02 03 04 05 06 07 08 09 0a 0b 0c 0d 0e 0f 10 11 12 13 14 15  
16 17 18 19 1a 1b 1c 1d 1e 1f

Extension - Supported Versions (1.3) (7 bytes):  
00 2b 00 03 02 03 04

Extension - Client Certificate Type (Raw Public Key) (6 bytes):  
00 13 00 01 01 02

Extension - Server Certificate Type (Raw Public Key) (6 bytes):  
00 14 00 01 01 02

Extension - Connection Identifier (43) (6 bytes):  
XX XX 00 02 01 42

13 + 10 + 2 + 32 + 1 + 1 + 4 + 2 + 2 + 8 + 8 + 42 + 7 + 6 + 6 + 6 = 150  
bytes

DTLS 1.3 RPK + ECDHE flight\_1 gives 150 bytes of overhead.

#### 2.2.1.2. flight\_2

Record Header - DTLSPlaintext (13 bytes):  
16 fe fd EE EE SS SS SS SS SS SS LL LL

Handshake Header - Server Hello (10 bytes):  
02 LL LL LL SS SS 00 00 00 LL LL LL

Legacy Version (2 bytes):  
fe fd

Server Random (32 bytes):  
00 01 02 03 04 05 06 07 08 09 0a 0b 0c 0d 0e 0f 10 11 12 13 14 15  
16 17 18 19 1a 1b 1c 1d 1e 1f

Legacy Session ID (1 bytes):  
00

Cipher Suite (TLS\_AES\_128\_CCM\_8\_SHA256) (2 bytes):  
13 05

Compression Method (null) (1 bytes):  
00

Extensions Length (2 bytes):  
LL LL

Extension - Key Share (40 bytes):  
00 33 00 24 00 1d 00 20  
00 01 02 03 04 05 06 07 08 09 0a 0b 0c 0d 0e 0f 10 11 12 13 14 15  
16 17 18 19 1a 1b 1c 1d 1e 1f

Extension - Supported Versions (1.3) (6 bytes):  
00 2b 00 02 03 04

Extension - Connection Identifier (43) (6 bytes):  
XX XX 00 02 01 43

Record Header - DTLS Ciphertext, Full (6 bytes):  
HH ES SS 43 LL LL

Handshake Header - Encrypted Extensions (10 bytes):  
08 LL LL LL SS SS 00 00 00 LL LL LL

Extensions Length (2 bytes):  
LL LL

Extension - Client Certificate Type (Raw Public Key) (6 bytes):  
00 13 00 01 01 02

Extension - Server Certificate Type (Raw Public Key) (6 bytes):  
00 14 00 01 01 02

Handshake Header - Certificate Request (10 bytes):  
0d LL LL LL SS SS 00 00 00 LL LL LL

Request Context (1 bytes):  
00

Extensions Length (2 bytes):  
LL LL

Extension - Signature Algorithms (ecdsa\_secp256r1\_sha256)  
(8 bytes):  
00 0d 00 04 00 02 08 07

Handshake Header - Certificate (10 bytes):  
0b LL LL LL SS SS 00 00 00 LL LL LL

Request Context (1 bytes):  
00

Certificate List Length (3 bytes):  
LL LL LL

Certificate Length (3 bytes):  
LL LL LL

Certificate (59 bytes) // Point compression  
....

Certificate Extensions (2 bytes):  
00 00

Handshake Header - Certificate Verify (10 bytes):  
0f LL LL LL SS SS 00 00 00 LL LL LL

Signature (68 bytes):  
ZZ ZZ 00 40 ....

Handshake Header - Finished (10 bytes):  
14 LL LL LL SS SS 00 00 00 LL LL LL

Verify Data (32 bytes):  
00 01 02 03 04 05 06 07 08 09 0a 0b 0c 0d 0e 0f 10 11 12 13 14 15  
16 17 18 19 1a 1b 1c 1d 1e 1f

Record Type (1 byte):  
16

Auth Tag (8 bytes):  
e0 8b 0e 45 5a 35 0a e5

13 + 102 + 6 + 24 + 21 + 78 + 78 + 42 + 1 + 8 = 373 bytes

DTLS 1.3 RPK + ECDHE flight\_2 gives 373 bytes of overhead.

### 2.2.1.3. flight\_3

Record Header (6 bytes) // DTLS Ciphertext, Full:  
ZZ ES SS 42 LL LL

Handshake Header - Certificate (10 bytes):  
0b LL LL LL SS SS XX XX XX LL LL LL

Request Context (1 bytes):  
00

Certificate List Length (3 bytes):  
LL LL LL

Certificate Length (3 bytes):  
LL LL LL

Certificate (59 bytes) // Point compression  
....

Certificate Extensions (2 bytes):  
00 00

Handshake Header - Certificate Verify (10 bytes):  
0f LL LL LL SS SS 00 00 00 LL LL LL

Signature (68 bytes):  
04 03 LL LL //ecdsa\_secp256r1\_sha256  
00 01 02 03 04 05 06 07 08 09 0a 0b 0c 0d 0e 0f 10 11 12 13 14 15  
16 17 18 19 1a 1b 1c 1d 1e 1f 00 01 02 03 04 05 06 07 08 09 0a 0b  
0c 0d 0e 0f 10 11 12 13 14 15 16 17 18 19 1a 1b 1c 1d 1e 1f

Handshake Header - Finished (10 bytes):  
14 LL LL LL SS SS 00 00 00 LL LL LL

Verify Data (32 bytes) // SHA-256:  
00 01 02 03 04 05 06 07 08 09 0a 0b 0c 0d 0e 0f 10 11 12 13 14 15  
16 17 18 19 1a 1b 1c 1d 1e 1f

Record Type (1 byte):  
16

Auth Tag (8 bytes) // AES-CCM\_8:  
00 01 02 03 04 05 06 07

6 + 78 + 78 + 42 + 1 + 8 = 213 bytes



DTLS 1.3 RPK + ECDHE flight\_2 gives 213 bytes of overhead.

#### 2.2.2. Message Sizes PSK + ECDHE

##### 2.2.2.1. flight\_1

The differences in overhead compared to Section 2.2.1.1 are:

The following is added:

+ Extension - PSK Key Exchange Modes (6 bytes):

00 2d 00 02 01 01

+ Extension - Pre Shared Key (48 bytes):

00 29 00 2F

00 0a 00 01 ID 00 00 00 00

00 21 20 00 01 02 03 04 05 06 07 08 09 0a 0b 0c 0d 0e 0f 10 11 12 13

14 15 16 17 18 19 1a 1b 1c 1d 1e 1f

The following is removed:

- Extension - Signature Algorithms (ecdsa\_secp256r1\_sha256) (8 bytes)

- Extension - Client Certificate Type (Raw Public Key) (6 bytes)

- Extension - Server Certificate Type (Raw Public Key) (6 bytes)

In total:

$150 + 6 + 48 - 8 - 6 - 6 = 184$  bytes

DTLS 1.3 PSK + ECDHE flight\_1 gives 184 bytes of overhead.

##### 2.2.2.2. flight\_2

The differences in overhead compared to Section 2.2.1.2 are:

The following is added:

+ Extension - Pre Shared Key (6 bytes)

00 29 00 02 00 00

The following is removed:

- Handshake Message Certificate (78 bytes)
- Handshake Message CertificateVerify (78 bytes)
- Handshake Message CertificateRequest (21 bytes)
- Extension - Client Certificate Type (Raw Public Key) (6 bytes)
- Extension - Server Certificate Type (Raw Public Key) (6 bytes)

In total:

$373 - 78 - 78 - 21 - 6 - 6 + 6 = 190$  bytes

DTLS 1.3 PSK + ECDHE flight\_2 gives 190 bytes of overhead.

#### 2.2.2.3. flight\_3

The differences in overhead compared to Section 2.2.1.3 are:

The following is removed:

- Handshake Message Certificate (78 bytes)
- Handshake Message Certificate Verify (78 bytes)

In total:

$213 - 78 - 78 = 57$  bytes

DTLS 1.3 PSK + ECDHE flight\_3 gives 57 bytes of overhead.

#### 2.2.3. Message Sizes PSK

##### 2.2.3.1. flight\_1

The differences in overhead compared to Section 2.2.2.1 are:

The following is removed:

- Extension - Supported Groups (x25519) (8 bytes)
- Extension - Key Share (42 bytes)

In total:

$184 - 8 - 42 = 134$  bytes

DTLS 1.3 PSK flight\_1 gives 134 bytes of overhead.

#### 2.2.3.2. flight\_2

The differences in overhead compared to Section 2.2.2.2 are:

The following is removed:

- Extension - Key Share (40 bytes)

In total:

$190 - 40 = 150$  bytes

DTLS 1.3 PSK flight\_2 gives 150 bytes of overhead.

#### 2.2.3.3. flight\_3

There are no differences in overhead compared to Section 2.2.2.3.

DTLS 1.3 PSK flight\_3 gives 57 bytes of overhead.

#### 2.2.4. Cached Information

In this section, we consider the effect of [RFC7924] on the message size overhead.

Cached information together with server X.509 can be used to move bytes from flight #2 to flight #1 (cached RPK increases the number of bytes compared to cached X.509).

The differences compared to Section 2.2.1 are the following.

For the flight #1, the following is added:

+ Extension - Client Cached Information (39 bytes):  
00 19 LL LL LL LL  
01 00 01 02 03 04 05 06 07 08 09 0a 0b 0c 0d 0e 0f 10 11 12 13 14 15  
16 17 18 19 1a 1b 1c 1d 1e 1f

And the following is removed:

- Extension - Server Certificate Type (Raw Public Key) (6 bytes)

Giving a total of:

$150 + 33 = 183$  bytes

For the flight #2, the following is added:

+ Extension - Server Cashed Information (7 bytes):  
00 19 LL LL LL LL 01

And the following is removed:

- Extension - Server Certificate Type (Raw Public Key) (6 bytes)
- Server Certificate (59 bytes -> 32 bytes)

Giving a total of:

$373 - 26 = 347$  bytes

A summary of the calculation is given in Figure 3.

Flight	#1	#2	#3	Total
DTLS 1.3 Cached X.509/RPK + ECDHE	183	347	213	743
DTLS 1.3 RPK + ECDHE	150	373	213	736

Figure 3: Comparison of message sizes in bytes for DTLS 1.3 RPK + ECDH with and without cached X.509

#### 2.2.5. Resumption

To enable resumption, a 4th flight (New Session Ticket) is added to the PSK handshake.

Record Header - DTLS Ciphertext, Full (6 bytes):  
 HH ES SS 43 LL LL

Handshake Header - New Session Ticket (10 bytes):  
 04 LL LL LL SS SS 00 00 00 LL LL LL

Ticket Lifetime (4 bytes):  
 00 01 02 03

Ticket Age Add (4 bytes):  
 00 01 02 03

Ticket Nonce (2 bytes):  
 01 00

Ticket (6 bytes):  
 00 04 ID ID ID ID

Extensions (2 bytes):  
 00 00

Auth Tag (8 bytes) // AES-CCM\_8:  
 00 01 02 03 04 05 06 07

$6 + 10 + 4 + 4 + 2 + 6 + 2 + 8 = 42$  bytes

The initial handshake when resumption is enabled is just a PSK handshake with  $134 + 150 + 57 + 42 = 383$  bytes.

#### 2.2.6. Without Connection ID

Without a Connection ID the DTLS 1.3 flight sizes changes as follows.

DTLS 1.3 Flight #1:    -6 bytes  
 DTLS 1.3 Flight #2:    -7 bytes  
 DTLS 1.3 Flight #3:    -1 byte

Flight	#1	#2	#3	Total
DTLS 1.3 RPK + ECDHE (no cid)	144	364	212	722
DTLS 1.3 PSK + ECDHE (no cid)	178	183	56	417
DTLS 1.3 PSK (no cid)	128	143	56	327

Figure 4: Comparison of message sizes in bytes for DTLS 1.3 without Connection ID

## 2.2.7. DTLS Raw Public Keys

TODO

## 2.2.7.1. SubjectPublicKeyInfo without point compression

0x30 // Sequence

0x59 // Size 89

0x30 // Sequence

0x13 // Size 19

0x06 0x07 0x2A 0x86 0x48 0xCE 0x3D 0x02 0x01

// OID 1.2.840.10045.2.1 (ecPublicKey)

0x06 0x08 0x2A 0x86 0x48 0xCE 0x3D 0x03 0x01 0x07

// OID 1.2.840.10045.3.1.7 (secp256r1)

0x03 // Bit string

0x42 // Size 66

0x00 // Unused bits 0

0x04 // Uncompressed

..... 64 bytes X and Y

Total of 91 bytes

## 2.2.7.2. SubjectPublicKeyInfo with point compression

0x30 // Sequence

0x59 // Size 89

0x30 // Sequence

0x13 // Size 19

0x06 0x07 0x2A 0x86 0x48 0xCE 0x3D 0x02 0x01

// OID 1.2.840.10045.2.1 (ecPublicKey)

0x06 0x08 0x2A 0x86 0x48 0xCE 0x3D 0x03 0x01 0x07

// OID 1.2.840.10045.3.1.7 (secp256r1)

0x03 // Bit string

0x42 // Size 66

0x00 // Unused bits 0

0x03 // Compressed

..... 32 bytes X

Total of 59 bytes

### 2.3. TLS 1.3

In this section, the message sizes are calculated for TLS 1.3. The major changes compared to DTLS 1.3 are that the record header is smaller, the handshake headers is smaller, and that Connection ID is not supported. Recently, additional work has taken shape with the goal to further reduce overhead for TLS 1.3 (see [I-D.rescorla-tls-ctls]).

#### TLS Assumptions:

- o Minimum number of algorithms and cipher suites offered
- o Curve25519, ECDSA with P-256, AES-CCM\_8, SHA-256
- o Length of key identifiers: 1 bytes
- o TLS RPK with point compression (saves 32 bytes)
- o Only mandatory TLS extensions

For the PSK calculations, [Ulfheim-TLS13] was a useful resource, while for RPK calculations we followed the work of [IoT-Cert].

#### 2.3.1. Message Sizes RPK + ECDHE

##### 2.3.1.1. flight\_1

Record Header - TLSPlaintext (5 bytes):

16 03 03 LL LL

Handshake Header - Client Hello (4 bytes):

01 LL LL LL

Legacy Version (2 bytes):

03 03

Client Random (32 bytes):

00 01 02 03 04 05 06 07 08 09 0a 0b 0c 0d 0e 0f 10 11 12 13 14 15  
16 17 18 19 1a 1b 1c 1d 1e 1f

Legacy Session ID (1 bytes):

00

Cipher Suites (TLS\_AES\_128\_CCM\_8\_SHA256) (4 bytes):

00 02 13 05

Compression Methods (null) (2 bytes):

01 00

Extensions Length (2 bytes):

LL LL

Extension - Supported Groups (x25519) (8 bytes):

00 0a 00 04 00 02 00 1d

Extension - Signature Algorithms(ecdsa\_secp256r1\_sha256) (8 bytes):

00 0d 00 04 00 02 08 07

Extension - Key Share (42 bytes):

00 33 00 26 00 24 00 1d 00 20  
00 01 02 03 04 05 06 07 08 09 0a 0b 0c 0d 0e 0f 10 11 12 13 14 15  
16 17 18 19 1a 1b 1c 1d 1e 1f

Extension - Supported Versions (1.3) (7 bytes):

00 2b 00 03 02 03 04

Extension - Client Certificate Type (Raw Public Key) (6 bytes):

00 13 00 01 01 02

Extension - Server Certificate Type (Raw Public Key) (6 bytes):

00 14 00 01 01 02

5 + 4 + 2 + 32 + 1 + 4 + 2 + 2 + 8 + 8 + 42 + 7 + 6 + 6 = 129 bytes

TLS 1.3 RPK + ECDHE flight\_1 gives 129 bytes of overhead.



2.3.1.2.    flight\_2

Record Header - TLSPlaintext (5 bytes):  
16 03 03 LL LL

Handshake Header - Server Hello (4 bytes):  
02 LL LL LL

Legacy Version (2 bytes):  
fe fd

Server Random (32 bytes):  
00 01 02 03 04 05 06 07 08 09 0a 0b 0c 0d 0e 0f 10 11 12 13 14 15  
16 17 18 19 1a 1b 1c 1d 1e 1f

Legacy Session ID (1 bytes):  
00

Cipher Suite (TLS\_AES\_128\_CCM\_8\_SHA256) (2 bytes):  
13 05

Compression Method (null) (1 bytes):  
00

Extensions Length (2 bytes):  
LL LL

Extension - Key Share (40 bytes):  
00 33 00 24 00 1d 00 20  
00 01 02 03 04 05 06 07 08 09 0a 0b 0c 0d 0e 0f 10 11 12 13 14 15  
16 17 18 19 1a 1b 1c 1d 1e 1f

Extension - Supported Versions (1.3) (6 bytes):  
00 2b 00 02 03 04

Record Header - TLSCiphertext (5 bytes):  
17 03 03 LL LL

Handshake Header - Encrypted Extensions (4 bytes):  
08 LL LL LL

Extensions Length (2 bytes):  
LL LL

Extension - Client Certificate Type (Raw Public Key) (6 bytes):  
00 13 00 01 01 02

Extension - Server Certificate Type (Raw Public Key) (6 bytes):

00 14 00 01 01 02

Handshake Header - Certificate Request (4 bytes):  
0d LL LL LL

Request Context (1 bytes):  
00

Extensions Length (2 bytes):  
LL LL

Extension - Signature Algorithms(ecdsa\_secp256r1\_sha256) (8 bytes):  
00 0d 00 04 00 02 08 07

Handshake Header - Certificate (4 bytes):  
0b LL LL LL

Request Context (1 bytes):  
00

Certificate List Length (3 bytes):  
LL LL LL

Certificate Length (3 bytes):  
LL LL LL

Certificate (59 bytes) // Point compression  
....

Certificate Extensions (2 bytes):  
00 00

Handshake Header - Certificate Verify (4 bytes):  
0f LL LL LL

Signature (68 bytes):  
ZZ ZZ 00 40 ....

Handshake Header - Finished (4 bytes):  
14 LL LL LL

Verify Data (32 bytes):  
00 01 02 03 04 05 06 07 08 09 0a 0b 0c 0d 0e 0f 10 11 12 13 14 15  
16 17 18 19 1a 1b 1c 1d 1e 1f

Record Type (1 byte):  
16

Auth Tag (8 bytes):

e0 8b 0e 45 5a 35 0a e5

$5 + 90 + 5 + 18 + 15 + 72 + 72 + 36 + 1 + 8 = 322$  bytes

TLS 1.3 RPK + ECDHE flight\_2 gives 322 bytes of overhead.

2.3.1.3. flight\_3

Record Header - TLSCiphertext (5 bytes):

17 03 03 LL LL

Handshake Header - Certificate (4 bytes):

0b LL LL LL

Request Context (1 bytes):

00

Certificate List Length (3 bytes):

LL LL LL

Certificate Length (3 bytes):

LL LL LL

Certificate (59 bytes) // Point compression

....

Certificate Extensions (2 bytes):

00 00

Handshake Header - Certificate Verify (4 bytes):

0f LL LL LL

Signature (68 bytes):

04 03 LL LL //ecdsa\_secp256r1\_sha256

00 01 02 03 04 05 06 07 08 09 0a 0b 0c 0d 0e 0f 10 11 12 13 14 15  
16 17 18 19 1a 1b 1c 1d 1e 1f 00 01 02 03 04 05 06 07 08 09 0a 0b  
0c 0d 0e 0f 10 11 12 13 14 15 16 17 18 19 1a 1b 1c 1d 1e 1f

Handshake Header - Finished (4 bytes):

14 LL LL LL

Verify Data (32 bytes) // SHA-256:

00 01 02 03 04 05 06 07 08 09 0a 0b 0c 0d 0e 0f 10 11 12 13 14 15  
16 17 18 19 1a 1b 1c 1d 1e 1f

Record Type (1 byte)

16

Auth Tag (8 bytes) // AES-CCM\_8:

00 01 02 03 04 05 06 07

5 + 72 + 72 + 36 + 1 + 8 = 194 bytes

TLS 1.3 RPK + ECDHE flight\_3 gives 194 bytes of overhead.

### 2.3.2. Message Sizes PSK + ECDHE

#### 2.3.2.1. flight\_1

The differences in overhead compared to Section 2.3.1.3 are:

The following is added:

+ Extension - PSK Key Exchange Modes (6 bytes):

00 2d 00 02 01 01

+ Extension - Pre Shared Key (48 bytes):

00 29 00 2F

00 0a 00 01 ID 00 00 00 00

00 21 20 00 01 02 03 04 05 06 07 08 09 0a 0b 0c 0d 0e 0f 10 11 12 13

14 15 16 17 18 19 1a 1b 1c 1d 1e 1f

The following is removed:

- Extension - Signature Algorithms (ecdsa\_secp256r1\_sha256) (8 bytes)

- Extension - Client Certificate Type (Raw Public Key) (6 bytes)

- Extension - Server Certificate Type (Raw Public Key) (6 bytes)

In total:

$129 + 6 + 48 - 8 - 6 - 6 = 163$  bytes

TLS 1.3 PSK + ECDHE flight\_1 gives 166 bytes of overhead.

#### 2.3.2.2. flight\_2

The differences in overhead compared to Section 2.3.1.2 are:

The following is added:

+ Extension - Pre Shared Key (6 bytes)

00 29 00 02 00 00

The following is removed:

- Handshake Message Certificate (72 bytes)
- Handshake Message CertificateVerify (72 bytes)
- Handshake Message CertificateRequest (15 bytes)
- Extension - Client Certificate Type (Raw Public Key) (6 bytes)
- Extension - Server Certificate Type (Raw Public Key) (6 bytes)

In total:

$322 - 72 - 72 - 15 - 6 - 6 + 6 = 157$  bytes

TLS 1.3 PSK + ECDHE flight\_2 gives 157 bytes of overhead.

#### 2.3.2.3. flight\_3

The differences in overhead compared to Section 2.3.1.3 are:

The following is removed:

- Handshake Message Certificate (72 bytes)
- Handshake Message Certificate Verify (72 bytes)

In total:

$194 - 72 - 72 = 50$  bytes

TLS 1.3 PSK + ECDHE flight\_3 gives 50 bytes of overhead.

#### 2.3.3. Message Sizes PSK

##### 2.3.3.1. flight\_1

The differences in overhead compared to Section 2.3.2.1 are:

The following is removed:

- Extension - Supported Groups (x25519) (8 bytes)
- Extension - Key Share (42 bytes)

In total:

$163 - 8 - 42 = 113$  bytes

TLS 1.3 PSK flight\_1 gives 116 bytes of overhead.

#### 2.3.3.2. flight\_2

The differences in overhead compared to Section 2.3.2.2 are:

The following is removed:

- Extension - Key Share (40 bytes)

In total:

157 - 40 = 117 bytes

TLS 1.3 PSK flight\_2 gives 117 bytes of overhead.

#### 2.3.3.3. flight\_3

There are no differences in overhead compared to Section 2.3.2.3.

TLS 1.3 PSK flight\_3 gives 57 bytes of overhead.

### 2.4. EDHOC

This section gives an estimate of the message sizes of EDHOC with authenticated with static Diffie-Hellman keys. All examples are given in CBOR diagnostic notation and hexadecimal, and are based on the test vectors in Appendix B.2 of [I-D.ietf-lake-edhoc].

#### 2.4.1. Message Sizes RPK

##### 2.4.1.1. message\_1

```
message_1 = (  
  13,  
  0,  
  h'8D3EF56D1B750A4351D68AC250A0E883790EFC80A538A444EE9E2B57E244  
    1A7C',  
  -2  
)
```

```
message_1 (37 bytes):  
0d 00 58 20 8d 3e f5 6d 1b 75 0a 43 51 d6 8a c2 50 a0 e8 83  
79 0e fc 80 a5 38 a4 44 ee 9e 2b 57 e2 44 1a 7c 21
```

#### 2.4.1.2. message\_2

```
message_2 = (
    h'52FBA0BDC8D953DD86CE1AB2FD7C05A4658C7C30AFDBFC3301047069451B
    AF35',
    8,
    h'DCF6FE9C524C22454DEB'
)
```

```
message_2 (46 bytes):
58 20 52 fb a0 bd c8 d9 53 dd 86 ce 1a b2 fd 7c 05 a4 65 8c
7c 30 af db fc 33 01 04 70 69 45 1b af 35 08 4a dc f6 fe 9c
52 4c 22 45 4d eb
```

#### 2.4.1.3. message\_3

```
message_3 = (
    8,
    h'53C3991999A5FFB86921E99B607C067770E0'
)
```

```
message_3 (20 bytes):
08 52 53 c3 99 19 99 a5 ff b8 69 21 e9 9b 60 7c 06 77 70 e0
```

#### 2.4.2. Summary

The typical message sizes for the previous example and for an example of EDHOC authenticated with signature keys and X.509 certificates based on Appendix B.1 of [I-D.ietf-lake-edhoc] are summarized in Figure 5.

	RPK	x5t
message_1	37	37
message_2	46	117
message_3	20	91
Total	103	245

Figure 5: Typical message sizes in bytes

#### 2.5. Conclusion

To do a fair comparison, one has to choose a specific deployment and look at the topology, the whole protocol stack, frame sizes (e.g. 51 or 128 bytes), how and where in the protocol stack fragmentation is



done, and the expected packet loss. Note that the number of bytes in each frame that is available for the key exchange protocol may depend on the underlying protocol layers as well as on the number of hops in multi-hop networks. The packet loss may depend on how many other devices are transmitting at the same time, and may increase during network formation. The total overhead will be larger due to mechanisms for fragmentation, retransmission, and packet ordering. The overhead of fragmentation is roughly proportional to the number of fragments, while the expected overhead due to retransmission in noisy environments is a superlinear function of the flight sizes.

### 3. Overhead for Protection of Application Data

To enable comparison, all the overhead calculations in this section use AES-CCM with a tag length of 8 bytes (e.g. AES\_128\_CCM\_8 or AES-CCM-16-64), a plaintext of 6 bytes, and the sequence number '05'. This follows the example in [RFC7400], Figure 16.

Note that the compressed overhead calculations for DTLS 1.2, DTLS 1.3, TLS 1.2 and TLS 1.3 are dependent on the parameters epoch, sequence number, and length, and all the overhead calculations are dependent on the parameter Connection ID when used. Note that the OSCORE overhead calculations are dependent on the CoAP option numbers, as well as the length of the OSCORE parameters Sender ID and Sequence Number. The following calculations are only examples.

Section 3.1 gives a short summary of the message overhead based on different parameters and some assumptions. The following sections detail the assumptions and the calculations.

#### 3.1. Summary

The DTLS overhead is dependent on the parameter Connection ID. The following overheads apply for all Connection IDs with the same length.

The compression overhead (GHC) is dependent on the parameters epoch, sequence number, Connection ID, and length (where applicable). The following overheads should be representative for sequence numbers and Connection IDs with the same length.

The OSCORE overhead is dependent on the included CoAP Option numbers as well as the length of the OSCORE parameters Sender ID and sequence number. The following overheads apply for all sequence numbers and Sender IDs with the same length.

Sequence Number	'05'	'1005'	'100005'
DTLS 1.2	29	29	29
DTLS 1.3	11	12	12
DTLS 1.2 (GHC)	16	16	16
DTLS 1.3 (GHC)	12	13	13
TLS 1.2	21	21	21
TLS 1.3	14	14	14
TLS 1.2 (GHC)	17	18	19
TLS 1.3 (GHC)	15	16	17
OSCORE request	13	14	15
OSCORE response	11	11	11

Figure 6: Overhead in bytes as a function of sequence number  
(Connection/Sender ID = '')

Connection/Sender ID	''	'42'	'4002'
DTLS 1.2	29	30	31
DTLS 1.3	11	12	13
DTLS 1.2 (GHC)	16	17	18
DTLS 1.3 (GHC)	12	13	14
OSCORE request	13	14	15
OSCORE response	11	11	11

Figure 7: Overhead in bytes as a function of Connection/Sender ID  
(Sequence Number = '05')

Protocol	Overhead	Overhead (GHC)
DTLS 1.2	21	8
DTLS 1.3	3	4
TLS 1.2	13	9
TLS 1.3	6	7
OSCORE request	5	
OSCORE response	3	

Figure 8: Overhead (excluding ICV) in bytes  
(Connection/Sender ID = '', Sequence Number = '05')

### 3.2. DTLS 1.2

#### 3.2.1. DTLS 1.2

This section analyzes the overhead of DTLS 1.2 [RFC6347]. The nonce follow the strict profiling given in [RFC7925]. This example is taken directly from [RFC7400], Figure 16.

DTLS 1.2 record layer (35 bytes, 29 bytes overhead):

```
17 fe fd 00 01 00 00 00 00 00 05 00 16 00 01 00
00 00 00 00 05 ae a0 15 56 67 92 4d ff 8a 24 e4
cb 35 b9
```

Content type:

17

Version:

fe fd

Epoch:

00 01

Sequence number:

00 00 00 00 00 05

Length:

00 16

Nonce:

00 01 00 00 00 00 00 05

Ciphertext:

ae a0 15 56 67 92

ICV:

4d ff 8a 24 e4 cb 35 b9

DTLS 1.2 gives 29 bytes overhead.

#### 3.2.2. DTLS 1.2 with 6LoWPAN-GHC

This section analyzes the overhead of DTLS 1.2 [RFC6347] when compressed with 6LoWPAN-GHC [RFC7400]. The compression was done with [OlegHahm-ghc].

Note that the sequence number '01' used in [RFC7400], Figure 15 gives an exceptionally small overhead that is not representative.

Note that this header compression is not available when DTLS is used over transports that do not use 6LoWPAN together with 6LoWPAN-GHC.

Compressed DTLS 1.2 record layer (22 bytes, 16 bytes overhead):  
b0 c3 03 05 00 16 f2 0e ae a0 15 56 67 92 4d ff  
8a 24 e4 cb 35 b9

Compressed DTLS 1.2 record layer header and nonce:  
b0 c3 03 05 00 16 f2 0e  
Ciphertext:  
ae a0 15 56 67 92  
ICV:  
4d ff 8a 24 e4 cb 35 b9

When compressed with 6LoWPAN-GHC, DTLS 1.2 with the above parameters (epoch, sequence number, length) gives 16 bytes overhead.

### 3.2.3. DTLS 1.2 with Connection ID

This section analyzes the overhead of DTLS 1.2 [RFC6347] with Connection ID [I-D.ietf-tls-dtls-connection-id]. The overhead calculations in this section uses Connection ID = '42'. DTLS record layer with a Connection ID = '' (the empty string) is equal to DTLS without Connection ID.

DTLS 1.2 record layer (36 bytes, 30 bytes overhead):  
17 fe fd 00 01 00 00 00 00 00 05 42 00 16 00 01  
00 00 00 00 00 05 ae a0 15 56 67 92 4d ff 8a 24  
e4 cb 35 b9

Content type:  
17  
Version:  
fe fd  
Epoch:  
00 01  
Sequence number:  
00 00 00 00 00 05  
Connection ID:  
42  
Length:  
00 16  
Nonce:  
00 01 00 00 00 00 00 05  
Ciphertext:  
ae a0 15 56 67 92  
ICV:  
4d ff 8a 24 e4 cb 35 b9

DTLS 1.2 with Connection ID gives 30 bytes overhead.

#### 3.2.4. DTLS 1.2 with Connection ID and 6LoWPAN-GHC

This section analyzes the overhead of DTLS 1.2 [RFC6347] with Connection ID [I-D.ietf-tls-dtls-connection-id] when compressed with 6LoWPAN-GHC [RFC7400] [OlegHahm-ghc].

Note that the sequence number '01' used in [RFC7400], Figure 15 gives an exceptionally small overhead that is not representative.

Note that this header compression is not available when DTLS is used over transports that do not use 6LoWPAN together with 6LoWPAN-GHC.

Compressed DTLS 1.2 record layer (23 bytes, 17 bytes overhead):

```
b0 c3 04 05 42 00 16 f2 0e ae a0 15 56 67 92 4d
ff 8a 24 e4 cb 35 b9
```

Compressed DTLS 1.2 record layer header and nonce:

```
b0 c3 04 05 42 00 16 f2 0e
```

Ciphertext:

```
ae a0 15 56 67 92
```

ICV:

```
4d ff 8a 24 e4 cb 35 b9
```

When compressed with 6LoWPAN-GHC, DTLS 1.2 with the above parameters (epoch, sequence number, Connection ID, length) gives 17 bytes overhead.

#### 3.3. DTLS 1.3

##### 3.3.1. DTLS 1.3

This section analyzes the overhead of DTLS 1.3 [I-D.ietf-tls-dtls13]. The changes compared to DTLS 1.2 are: omission of version number, merging of epoch into the first byte containing signalling bits, optional omission of length, reduction of sequence number into a 1 or 2-bytes field.

Only the minimal header format for DTLS 1.3 is analyzed (see Figure 4 of [I-D.ietf-tls-dtls13]). The minimal header format omit the length field and only a 1-byte field is used to carry the 8 low order bits of the sequence number

DTLS 1.3 record layer (17 bytes, 11 bytes overhead):  
21 05 ae a0 15 56 67 92 ec 4d ff 8a 24 e4 cb 35 b9

First byte (including epoch):

21

Sequence number:

05

Ciphertext (including encrypted content type):

ae a0 15 56 67 92 ec

ICV:

4d ff 8a 24 e4 cb 35 b9

DTLS 1.3 gives 11 bytes overhead.

### 3.3.2. DTLS 1.3 with 6LoWPAN-GHC

This section analyzes the overhead of DTLS 1.3 [I-D.ietf-tls-dtls13] when compressed with 6LoWPAN-GHC [RFC7400] [OlegHahm-ghc].

Note that this header compression is not available when DTLS is used over transports that do not use 6LoWPAN together with 6LoWPAN-GHC.

Compressed DTLS 1.3 record layer (18 bytes, 12 bytes overhead):

11 21 05 ae a0 15 56 67 92 ec 4d ff 8a 24 e4 cb  
35 b9

Compressed DTLS 1.3 record layer header and nonce:

11 21 05

Ciphertext (including encrypted content type):

ae a0 15 56 67 92 ec

ICV:

4d ff 8a 24 e4 cb 35 b9

When compressed with 6LoWPAN-GHC, DTLS 1.3 with the above parameters (epoch, sequence number, no length) gives 12 bytes overhead.

### 3.3.3. DTLS 1.3 with Connection ID

This section analyzes the overhead of DTLS 1.3 [I-D.ietf-tls-dtls13] with Connection ID [I-D.ietf-tls-dtls-connection-id].

In this example, the length field is omitted, and the 1-byte field is used for the sequence number. The minimal DTLSCiphertext structure is used (see Figure 4 of [I-D.ietf-tls-dtls13]), with the addition of the Connection ID field.

DTLS 1.3 record layer (18 bytes, 12 bytes overhead):  
31 42 05 ae a0 15 56 67 92 ec 4d ff 8a 24 e4 cb 35 b9

First byte (including epoch):

31

Connection ID:

42

Sequence number:

05

Ciphertext (including encrypted content type):

ae a0 15 56 67 92 ec

ICV:

4d ff 8a 24 e4 cb 35 b9

DTLS 1.3 with Connection ID gives 12 bytes overhead.

#### 3.3.4. DTLS 1.3 with Connection ID and 6LoWPAN-GHC

This section analyzes the overhead of DTLS 1.3 [I-D.ietf-tls-dtls13] with Connection ID [I-D.ietf-tls-dtls-connection-id] when compressed with 6LoWPAN-GHC [RFC7400] [OlegHahm-ghc].

Note that this header compression is not available when DTLS is used over transports that do not use 6LoWPAN together with 6LoWPAN-GHC.

Compressed DTLS 1.3 record layer (19 bytes, 13 bytes overhead):

12 31 05 42 ae a0 15 56 67 92 ec 4d ff 8a 24 e4  
cb 35 b9

Compressed DTLS 1.3 record layer header and nonce:

12 31 05 42

Ciphertext (including encrypted content type):

ae a0 15 56 67 92 ec

ICV:

4d ff 8a 24 e4 cb 35 b9

When compressed with 6LoWPAN-GHC, DTLS 1.3 with the above parameters (epoch, sequence number, Connection ID, no length) gives 13 bytes overhead.

#### 3.4. TLS 1.2

##### 3.4.1. TLS 1.2

This section analyzes the overhead of TLS 1.2 [RFC5246]. The changes compared to DTLS 1.2 is that the TLS 1.2 record layer does not have epoch and sequence number, and that the version is different.

TLS 1.2 Record Layer (27 bytes, 21 bytes overhead):

17 03 03 00 16 00 00 00 00 00 00 00 05 ae a0 15

56 67 92 4d ff 8a 24 e4 cb 35 b9

Content type:

17

Version:

03 03

Length:

00 16

Nonce:

00 00 00 00 00 00 00 05

Ciphertext:

ae a0 15 56 67 92

ICV:

4d ff 8a 24 e4 cb 35 b9

TLS 1.2 gives 21 bytes overhead.

#### 3.4.2. TLS 1.2 with 6LoWPAN-GHC

This section analyzes the overhead of TLS 1.2 [RFC5246] when compressed with 6LoWPAN-GHC [RFC7400] [OlegHahm-ghc].

Note that this header compression is not available when TLS is used over transports that do not use 6LoWPAN together with 6LoWPAN-GHC.

Compressed TLS 1.2 record layer (23 bytes, 17 bytes overhead):

05 17 03 03 00 16 85 0f 05 ae a0 15 56 67 92 4d

ff 8a 24 e4 cb 35 b9

Compressed TLS 1.2 record layer header and nonce:

05 17 03 03 00 16 85 0f 05

Ciphertext:

ae a0 15 56 67 92

ICV:

4d ff 8a 24 e4 cb 35 b9

When compressed with 6LoWPAN-GHC, TLS 1.2 with the above parameters (epoch, sequence number, length) gives 17 bytes overhead.

#### 3.5. TLS 1.3

##### 3.5.1. TLS 1.3

This section analyzes the overhead of TLS 1.3 [RFC8446]. The change compared to TLS 1.2 is that the TLS 1.3 record layer uses a different version.



TLS 1.3 Record Layer (20 bytes, 14 bytes overhead):  
17 03 03 00 16 ae a0 15 56 67 92 ec 4d ff 8a 24  
e4 cb 35 b9

Content type:

17

Legacy version:

03 03

Length:

00 0f

Ciphertext (including encrypted content type):

ae a0 15 56 67 92 ec

ICV:

4d ff 8a 24 e4 cb 35 b9

TLS 1.3 gives 14 bytes overhead.

### 3.5.2. TLS 1.3 with 6LoWPAN-GHC

This section analyzes the overhead of TLS 1.3 [RFC8446] when compressed with 6LoWPAN-GHC [RFC7400] [OlegHahm-ghc].

Note that this header compression is not available when TLS is used over transports that do not use 6LoWPAN together with 6LoWPAN-GHC.

Compressed TLS 1.3 record layer (21 bytes, 15 bytes overhead):

14 17 03 03 00 0f ae a0 15 56 67 92 ec 4d ff 8a  
24 e4 cb 35 b9

Compressed TLS 1.3 record layer header and nonce:

14 17 03 03 00 0f

Ciphertext (including encrypted content type):

ae a0 15 56 67 92 ec

ICV:

4d ff 8a 24 e4 cb 35 b9

When compressed with 6LoWPAN-GHC, TLS 1.3 with the above parameters (epoch, sequence number, length) gives 15 bytes overhead.

### 3.6. OSCORE

This section analyzes the overhead of OSCORE [RFC8613].

The below calculation Option Delta = '9', Sender ID = '' (empty string), and Sequence Number = '05', and is only an example. Note that Sender ID = '' (empty string) can only be used by one client per server.

OSCORE request (19 bytes, 13 bytes overhead):

92 09 05

ff ec ae a0 15 56 67 92 4d ff 8a 24 e4 cb 35 b9

CoAP option delta and length:

92

Option value (flag byte and sequence number):

09 05

Payload marker:

ff

Ciphertext (including encrypted code):

ec ae a0 15 56 67 92

ICV:

4d ff 8a 24 e4 cb 35 b9

The below calculation Option Delta = '9', Sender ID = '42', and Sequence Number = '05', and is only an example.

OSCORE request (20 bytes, 14 bytes overhead):

93 09 05 42

ff ec ae a0 15 56 67 92 4d ff 8a 24 e4 cb 35 b9

CoAP option delta and length:

93

Option Value (flag byte, sequence number, and Sender ID):

09 05 42

Payload marker:

ff

Ciphertext (including encrypted code):

ec ae a0 15 56 67 92

ICV:

4d ff 8a 24 e4 cb 35 b9

The below calculation uses Option Delta = '9'.

OSCORE response (17 bytes, 11 bytes overhead):

90

ff ec ae a0 15 56 67 92 4d ff 8a 24 e4 cb 35 b9

CoAP delta and option length:

90

Option value:

-

Payload marker:

ff

Ciphertext (including encrypted code):

ec ae a0 15 56 67 92

ICV:

4d ff 8a 24 e4 cb 35 b9

OSCORE with the above parameters gives 13-14 bytes overhead for requests and 11 bytes overhead for responses.

Unlike DTLS and TLS, OSCORE has much smaller overhead for responses than requests.

### 3.7. Group OSCORE

This section analyzes the overhead of Group OSCORE [I-D.ietf-core-oscore-groupcomm].

TODO

### 3.8. Conclusion

DTLS 1.2 has quite a large overhead as it uses an explicit sequence number and an explicit nonce. TLS 1.2 has significantly less (but not small) overhead. TLS 1.3 has quite a small overhead. OSCORE and DTLS 1.3 (using the minimal structure) format have very small overhead.

The Generic Header Compression (6LoWPAN-GHC) can in addition to DTLS 1.2 handle TLS 1.2, and DTLS 1.2 with Connection ID. The Generic Header Compression (6LoWPAN-GHC) works very well for Connection ID and the overhead seems to increase exactly with the length of the Connection ID (which is optimal). The compression of TLS 1.2 is not as good as the compression of DTLS 1.2 (as the static dictionary only contains the DTLS 1.2 version number). Similar compression levels as for DTLS could be achieved also for TLS 1.2, but this would require different static dictionaries. For TLS 1.3 and DTLS 1.3, GHC increases the overhead. The 6LoWPAN-GHC header compression is not available when (D)TLS is used over transports that do not use 6LoWPAN together with 6LoWPAN-GHC.

New security protocols like OSCORE, TLS 1.3, and DTLS 1.3 have much lower overhead than DTLS 1.2 and TLS 1.2. The overhead is even smaller than DTLS 1.2 and TLS 1.2 over 6LoWPAN with compression, and therefore the small overhead is achieved even on deployments without 6LoWPAN or 6LoWPAN without compression. OSCORE is lightweight because it makes use of CoAP, CBOR, and COSE, which were designed to have as low overhead as possible.

Note that the compared protocols have slightly different use cases. TLS and DTLS are designed for the transport layer and are terminated in CoAP proxies. OSCORE is designed for the application layer and protects information end-to-end between the CoAP client and the CoAP server. Group OSCORE is designed for group communication and protects information between a CoAP client and any number of CoAP servers.

#### 4. Security Considerations

This document is purely informational.

#### 5. IANA Considerations

This document has no actions for IANA.

#### 6. Informative References

[I-D.ietf-core-oscore-groupcomm]

Tiloca, M., Selander, G., Palombini, F., and J. Park,  
"Group OSCORE - Secure Group Communication for CoAP",  
draft-ietf-core-oscore-groupcomm-09 (work in progress),  
June 2020.

[I-D.ietf-lake-edhoc]

Selander, G., Mattsson, J., and F. Palombini, "Ephemeral  
Diffie-Hellman Over COSE (EDHOC)", draft-ietf-lake-  
edhoc-01 (work in progress), August 2020.

[I-D.ietf-tls-dtls-connection-id]

Rescorla, E., Tschofenig, H., and T. Fossati, "Connection  
Identifiers for DTLS 1.2", draft-ietf-tls-dtls-connection-  
id-07 (work in progress), October 2019.

[I-D.ietf-tls-dtls13]

Rescorla, E., Tschofenig, H., and N. Modadugu, "The  
Datagram Transport Layer Security (DTLS) Protocol Version  
1.3", draft-ietf-tls-dtls13-38 (work in progress), May  
2020.

- [I-D.rescorla-tls-ctls]  
Rescorla, E., Barnes, R., and H. Tschofenig, "Compact TLS 1.3", draft-rescorla-tls-ctls-04 (work in progress), March 2020.
- [IoT-Cert]  
Forsby, F., "Digital Certificates for the Internet of Things", June 2017, <<https://kth.diva-portal.org/smash/get/diva2:1153958/FULLTEXT01.pdf>>.
- [OlegHahm-ghc]  
Hahm, O., "Generic Header Compression", July 2016, <<https://github.com/OlegHahm/ghc>>.
- [RFC5246] Dierks, T. and E. Rescorla, "The Transport Layer Security (TLS) Protocol Version 1.2", RFC 5246, DOI 10.17487/RFC5246, August 2008, <<https://www.rfc-editor.org/info/rfc5246>>.
- [RFC6347] Rescorla, E. and N. Modadugu, "Datagram Transport Layer Security Version 1.2", RFC 6347, DOI 10.17487/RFC6347, January 2012, <<https://www.rfc-editor.org/info/rfc6347>>.
- [RFC7252] Shelby, Z., Hartke, K., and C. Bormann, "The Constrained Application Protocol (CoAP)", RFC 7252, DOI 10.17487/RFC7252, June 2014, <<https://www.rfc-editor.org/info/rfc7252>>.
- [RFC7400] Bormann, C., "6LoWPAN-GHC: Generic Header Compression for IPv6 over Low-Power Wireless Personal Area Networks (6LoWPANs)", RFC 7400, DOI 10.17487/RFC7400, November 2014, <<https://www.rfc-editor.org/info/rfc7400>>.
- [RFC7924] Santesson, S. and H. Tschofenig, "Transport Layer Security (TLS) Cached Information Extension", RFC 7924, DOI 10.17487/RFC7924, July 2016, <<https://www.rfc-editor.org/info/rfc7924>>.
- [RFC7925] Tschofenig, H., Ed. and T. Fossati, "Transport Layer Security (TLS) / Datagram Transport Layer Security (DTLS) Profiles for the Internet of Things", RFC 7925, DOI 10.17487/RFC7925, July 2016, <<https://www.rfc-editor.org/info/rfc7925>>.

- [RFC8323] Bormann, C., Lemay, S., Tschofenig, H., Hartke, K., Silverajan, B., and B. Raymor, Ed., "CoAP (Constrained Application Protocol) over TCP, TLS, and WebSockets", RFC 8323, DOI 10.17487/RFC8323, February 2018, <<https://www.rfc-editor.org/info/rfc8323>>.
- [RFC8446] Rescorla, E., "The Transport Layer Security (TLS) Protocol Version 1.3", RFC 8446, DOI 10.17487/RFC8446, August 2018, <<https://www.rfc-editor.org/info/rfc8446>>.
- [RFC8613] Selander, G., Mattsson, J., Palombini, F., and L. Seitz, "Object Security for Constrained RESTful Environments (OSCORE)", RFC 8613, DOI 10.17487/RFC8613, July 2019, <<https://www.rfc-editor.org/info/rfc8613>>.
- [Ulfheim-TLS13]  
Driscoll, M., "Every Byte Explained The Illustrated TLS 1.3 Connection", March 2018, <<https://tls13.ulfheim.net>>.

#### Acknowledgments

The authors want to thank Ari Keraenen, Carsten Bormann, Goeran Selander, and Hannes Tschofenig for comments and suggestions on previous versions of the draft.

All 6LoWPAN-GHC compression was done with [OlegHahm-ghc].

#### Authors' Addresses

John Preuss Mattsson  
Ericsson AB

Email: [john.mattsson@ericsson.com](mailto:john.mattsson@ericsson.com)

Francesca Palombini  
Ericsson AB

Email: [francesca.palombini@ericsson.com](mailto:francesca.palombini@ericsson.com)

Malisa Vucinic  
INRIA

Email: [malisa.vucinic@inria.fr](mailto:malisa.vucinic@inria.fr)

LWIG Working Group  
Internet-Draft  
Intended status: Informational  
Expires: September 10, 2019

C. Gomez  
UPC  
J. Crowcroft  
University of Cambridge  
M. Scharf  
Hochschule Esslingen  
March 9, 2019

TCP Usage Guidance in the Internet of Things (IoT)  
draft-ietf-lwig-tcp-constrained-node-networks-05

Abstract

This document provides guidance on how to implement and use the Transmission Control Protocol (TCP) in Constrained-Node Networks (CNNs), which are a characteristic of the Internet of Things (IoT). Such environments require a lightweight TCP implementation and may not make use of optional functionality. This document explains a number of known and deployed techniques to simplify a TCP stack as well as corresponding tradeoffs. The objective is to help embedded developers with decisions on which TCP features to use.

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 September 10, 2019.

Copyright Notice

Copyright (c) 2019 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 Provisions Relating to IETF Documents (<https://trustee.ietf.org/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 Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

## Table of Contents

1. Introduction . . . . .	3
2. Conventions used in this document . . . . .	4
3. Characteristics of CNNs relevant for TCP . . . . .	4
3.1. Network and link properties . . . . .	4
3.2. Usage scenarios . . . . .	5
3.3. Communication and traffic patterns . . . . .	6
4. TCP implementation and configuration in CNNs . . . . .	6
4.1. Path properties . . . . .	6
4.1.1. Maximum Segment Size (MSS) . . . . .	7
4.1.2. Explicit Congestion Notification (ECN) . . . . .	7
4.1.3. Explicit loss notifications . . . . .	8
4.2. TCP guidance for single-MSS windows and buffers . . . . .	9
4.2.1. Single-MSS stacks - benefits and issues . . . . .	9
4.2.2. TCP options for single-MSS stacks . . . . .	9
4.2.3. Delayed Acknowledgments for single-MSS stacks . . . . .	10
4.2.4. RTO estimation for single-MSS stacks . . . . .	10
4.3. General recommendations for TCP in CNNs . . . . .	11
4.3.1. Loss recovery and congestion/flow control . . . . .	11
4.3.1.1. Selective Acknowledgments (SACK) . . . . .	11
4.3.2. Delayed Acknowledgments . . . . .	12
5. TCP usage recommendations in CNNs . . . . .	12
5.1. TCP connection initiation . . . . .	12
5.2. Number of concurrent connections . . . . .	12
5.3. TCP connection lifetime . . . . .	13
6. Security Considerations . . . . .	15
7. Acknowledgments . . . . .	15
8. Annex. TCP implementations for constrained devices . . . . .	16
8.1. uIP . . . . .	16
8.2. lwIP . . . . .	16
8.3. RIOT . . . . .	17
8.4. TinyOS . . . . .	17
8.5. FreeRTOS . . . . .	18
8.6. uC/OS . . . . .	18
8.7. Summary . . . . .	18
9. Annex. Changes compared to previous versions . . . . .	20
9.1. Changes between -00 and -01 . . . . .	20
9.2. Changes between -01 and -02 . . . . .	20
9.3. Changes between -02 and -03 . . . . .	20
9.4. Changes between -03 and -04 . . . . .	21



9.5. Changes between -04 and -05 . . . . .	21
10. References . . . . .	21
10.1. Normative References . . . . .	21
10.2. Informative References . . . . .	23
Authors' Addresses . . . . .	26

## 1. Introduction

The Internet Protocol suite is being used for connecting Constrained-Node Networks (CNNs) to the Internet, enabling the so-called Internet of Things (IoT) [RFC7228]. In order to meet the requirements that stem from CNNs, the IETF has produced a suite of new protocols specifically designed for such environments (see e.g. [RFC8352]). New IETF protocol stack components include the IPv6 over Low-power Wireless Personal Area Networks (6LoWPAN) adaptation layer, the IPv6 Routing Protocol for Low-power and lossy networks (RPL) routing protocol, and the Constrained Application Protocol (CoAP).

As of the writing, the main current transport layer protocols in IP-based IoT scenarios are UDP and TCP. However, TCP has been criticized (often, unfairly) as a protocol for the IoT. In fact, some TCP features are not optimal for IoT scenarios, such as relatively long header size, unsuitability for multicast, and always-confirmed data delivery. However, many typical claims on TCP unsuitability for IoT (e.g. a high complexity, connection-oriented approach incompatibility with radio duty-cycling, and spurious congestion control activation in wireless links) are not valid, can be solved, or are also found in well accepted IoT end-to-end reliability mechanisms (see [IntComp] for a detailed analysis).

At the application layer, CoAP was developed over UDP [RFC7252]. However, the integration of some CoAP deployments with existing infrastructure is being challenged by middleboxes such as firewalls, which may limit and even block UDP-based communications. This the main reason why a CoAP over TCP specification has been developed [RFC8323].

Other application layer protocols not specifically designed for CNNs are also being considered for the IoT space. Some examples include HTTP/2 and even HTTP/1.1, both of which run over TCP by default [RFC7230] [RFC7540], and the Extensible Messaging and Presence Protocol (XMPP) [RFC6120]. TCP is also used by non-IETF application-layer protocols in the IoT space such as the Message Queue Telemetry Transport (MQTT) and its lightweight variants.

TCP is a sophisticated transport protocol that includes optional functionality (e.g. TCP options) that may improve performance in some environments. However, many optional TCP extensions require

complex logic inside the TCP stack and increase the codesize and the RAM requirements. Many TCP extensions are not required for interoperability with other standard-compliant TCP endpoints. Given the limited resources on constrained devices, careful "tuning" of the TCP implementation can make an implementation more lightweight.

This document provides guidance on how to implement and use TCP in CNNs. The overarching goal is to offer simple measures to allow for lightweight TCP implementation and suitable operation in such environments. A TCP implementation following the guidance in this document is intended to be compatible with a TCP endpoint that is compliant to the TCP standards, albeit possibly with a lower performance. This implies that such a TCP client would always be able to connect with a standard-compliant TCP server, and a corresponding TCP server would always be able to connect with a standard-compliant TCP client.

This document assumes that the reader is familiar with TCP. A comprehensive survey of the TCP standards can be found in [RFC7414]. Similar guidance regarding the use of TCP in special environments has been published before, e.g., for cellular wireless networks [RFC3481].

## 2. Conventions used in this document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

## 3. Characteristics of CNNs relevant for TCP

### 3.1. Network and link properties

CNNs are defined in [RFC7228] as networks whose characteristics are influenced by being composed of a significant portion of constrained nodes. The latter are characterized by significant limitations on processing, memory, and energy resources, among others [RFC7228]. The first two dimensions pose constraints on the complexity and on the memory footprint of the protocols that constrained nodes can support. The latter requires techniques to save energy, such as radio duty-cycling in wireless devices [RFC8352], as well as minimization of the number of messages transmitted/received (and their size).

[RFC7228] lists typical network constraints in CNN, including low achievable bitrate/throughput, high packet loss and high variability of packet loss, highly asymmetric link characteristics, severe penalties for using larger packets, limits on reachability over time,

etc. CNN may use wireless or wired technologies (e.g., Power Line Communication), and the transmission rates are typically low (e.g. below 1 Mbps).

For use of TCP, one challenge is that not all technologies in CNN may be aligned with typical Internet subnetwork design principles [RFC3819]. For instance, constrained nodes often use physical/link layer technologies that have been characterized as 'lossy', i.e., exhibit a relatively high bit error rate. Dealing with corruption loss is one of the open issues in the Internet [RFC6077].

### 3.2. Usage scenarios

There are different deployment and usage scenarios for CNNs. Some CNNs follow the star topology, whereby one or several hosts are linked to a central device that acts as a router connecting the CNN to the Internet. CNNs may also follow the multihop topology [RFC6606]. One key use case for the use of TCP is a model where constrained devices connect to unconstrained servers in the Internet. But it is also possible that both TCP endpoints run on constrained devices.

In constrained environments, there can be different types of devices [RFC7228]. For example, there can be devices with single combined send/receive buffer, devices with a separate send and receive buffer, or devices with a pool of multiple send/receive buffers. In the latter case, it is possible that buffers also be shared for other protocols.

When a CNN comprising one or more constrained devices and an unconstrained device communicate over the Internet using TCP, the communication possibly has to traverse a middlebox (e.g. a firewall, NAT, etc.). Figure 1 illustrates such scenario. Note that the scenario is asymmetric, as the unconstrained device will typically not suffer the severe constraints of the constrained device. The unconstrained device is expected to be mains-powered, to have high amount of memory and processing power, and to be connected to a resource-rich network.

Assuming that a majority of constrained devices will correspond to sensor nodes, the amount of data traffic sent by constrained devices (e.g. sensor node measurements) is expected to be higher than the amount of data traffic in the opposite direction. Nevertheless, constrained devices may receive requests (to which they may respond), commands (for configuration purposes and for constrained devices including actuators) and relatively infrequent firmware/software updates.

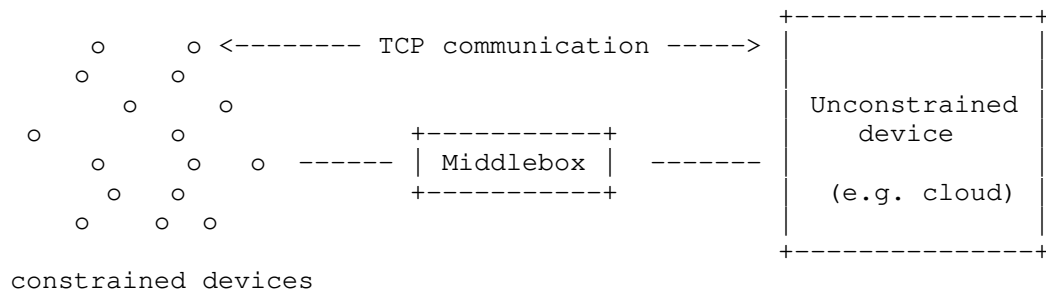


Figure 1: TCP communication between a constrained device and an unconstrained device, traversing a middlebox.

### 3.3. Communication and traffic patterns

IoT applications are characterized by a number of different communication patterns. The following non-comprehensive list explains some typical examples:

- o Unidirectional transfers: An IoT device (e.g. a sensor) can send (repeatedly) updates to the other endpoint. Not in every case there is a need for an application response back to the IoT device.
- o Request-response patterns: An IoT device receiving a request from the other endpoint, which triggers a response from the IoT device.
- o Bulk data transfers: A typical example for a long file transfer would be an IoT device firmware update.

A typical communication pattern is that a constrained device communicates with an unconstrained device (cf. Figure 1). But it is also possible that constrained devices communicate amongst themselves.

## 4. TCP implementation and configuration in CNNs

This section explains how a TCP stack can deal with typical constraints in CNN. The guidance in this section relates to the TCP implementation and its configuration.

### 4.1. Path properties

#### 4.1.1.1. Maximum Segment Size (MSS)

For the sake of lightweight implementation and operation, unless applications require handling large data units (i.e. leading to an IPv6 datagram size greater than 1280 bytes), it may be desirable to limit the MTU to 1280 bytes in order to avoid the need to support Path MTU Discovery [RFC8201].

An IPv6 datagram size exceeding 1280 bytes can be avoided by setting the TCP MSS not larger than 1220 bytes. (Note: IP version 6 is assumed.)

Note that setting the MTU to 1280 bytes is possible for link layer technologies in the CNN space, even if some of them are characterized by a short data unit payload size, e.g. up to a few tens or hundreds of bytes. For example, the maximum frame size in IEEE 802.15.4 is 127 bytes. 6LoWPAN defined an adaptation layer to support IPv6 over IEEE 802.15.4 networks. The adaptation layer includes a fragmentation mechanism, since IPv6 requires the layer below to support an MTU of 1280 bytes [RFC2460], while IEEE 802.15.4 lacked fragmentation mechanisms. 6LoWPAN defines an IEEE 802.15.4 link MTU of 1280 bytes [RFC4944]. Other technologies, such as Bluetooth LE [RFC7668], ITU-T G.9959 [RFC7428] or DECT-ULE [RFC8105], also use 6LoWPAN-based adaptation layers in order to enable IPv6 support. These technologies do support link layer fragmentation. By exploiting this functionality, the adaptation layers that enable IPv6 over such technologies also define an MTU of 1280 bytes.

On the other hand, there exist technologies also used in the CNN space, such as Master Slave / Token Passing (TP) [RFC8163], Narrowband IoT (NB-IoT) [RFC8376] or IEEE 802.11ah [I-D.delcarpio-6lo-wlanah], that do not suffer the same degree of frame size limitations as the technologies mentioned above. The MTU for MS/TP is recommended to be 1500 bytes [RFC8163], the MTU in NB-IoT is 1600 bytes, and the maximum frame payload size for IEEE 802.11ah is 7991 bytes.

Finally, note that using larger MSS (to a suitable extent) may be beneficial, especially when transferring large payloads, as it reduces the number of packets (and packet headers) required for a given payload.

#### 4.1.2. Explicit Congestion Notification (ECN)

Explicit Congestion Notification (ECN) [RFC3168] ECN allows a router to signal in the IP header of a packet that congestion is arising, for example when a queue size reaches a certain threshold. An ECN-enabled TCP receiver will echo back the congestion signal to the TCP

sender by setting a flag in its next TCP ACK. The sender triggers congestion control measures as if a packet loss had happened.

The document [RFC8087] outlines the principal gains in terms of increased throughput, reduced delay, and other benefits when ECN is used over a network path that includes equipment that supports Congestion Experienced (CE) marking. In the context of CNNs, a remarkable feature of ECN is that congestion can be signalled without incurring packet drops (which will lead to retransmissions and consumption of limited resources such as energy and bandwidth).

ECN can further reduce packet losses since congestion control measures can be applied earlier [RFC2884]. Less lost packets implies that the number of retransmitted segments decreases, which is particularly beneficial in CNNs, where energy and bandwidth resources are typically limited. Also, it makes sense to try to avoid packet drops for transactional workloads with small data sizes, which are typical for CNNs. In such traffic patterns, it is more difficult to detect packet loss without retransmission timeouts (e.g., as there may be no three duplicate ACKs). Any retransmission timeout slows down the data transfer significantly. In addition, if the constrained device uses power saving techniques, a retransmission timeout will incur a wake-up action, in contrast to ACK clock-triggered sending. When the congestion window of a TCP sender has a size of one segment, the TCP sender resets the retransmit timer, and the sender will only be able to send a new packet when the retransmit timer expires [RFC3168]. Effectively, the TCP sender reduces at that moment its sending rate from 1 segment per Round Trip Time (RTT) to 1 segment per RTO, which can result in a very low throughput. In addition to better throughput, ECN can also help reducing latency and jitter.

ECN can be incrementally deployed in the Internet. Guidance on configuration and usage of ECN is provided in [RFC7567]. Given the benefits, more and more TCP stacks in the Internet support ECN, and it specifically makes sense to leverage ECN in controlled environments such as CNNs. Note, however, that supporting ECN increases implementation complexity.

#### 4.1.3. Explicit loss notifications

There has been a significant body of research on solutions capable of explicitly indicating whether a TCP segment loss is due to corruption, in order to avoid activation of congestion control mechanisms [ETEN] [RFC2757]. While such solutions may provide significant improvement, they have not been widely deployed and remain as experimental work. In fact, as of today, the IETF has not standardized any such solution.

## 4.2. TCP guidance for single-MSS windows and buffers

This section discusses TCP stacks that focus on transferring a single MSS. More general guidance is provided in Section 4.3.

### 4.2.1. Single-MSS stacks - benefits and issues

A TCP stack can reduce the RAM requirements by advertising a TCP window size of one MSS, and also transmit at most one MSS of unacknowledged data. In that case, both congestion and flow control implementation is quite simple. Such a small receive and send window may be sufficient for simple message exchanges in the CNN space. However, only using a window of one MSS can significantly affect performance. A stop-and-wait operation results in low throughput for transfers that exceed the lengths of one MSS, e.g., a firmware download.

If CoAP is used over TCP with the default setting for NSTART in [RFC7252], a CoAP endpoint is not allowed to send a new message to a destination until a response for the previous message sent to that destination has been received. This is equivalent to an application-layer window size of 1. For this use of CoAP, a maximum TCP window of one MSS will be sufficient.

### 4.2.2. TCP options for single-MSS stacks

A TCP implementation needs to support options 0, 1 and 2 [RFC0793]. These options are sufficient for interoperability with a standard-compliant TCP endpoint, albeit many TCP stacks support additional options and can negotiate their use.

A TCP implementation for a constrained device that uses a single-MSS TCP receive or transmit window size may not benefit from supporting the following TCP options: Window scale [RFC7323], TCP Timestamps [RFC7323], Selective Acknowledgments (SACK) and SACK-Permitted [RFC2018]. Also other TCP options may not be required on a constrained device with a very lightweight implementation. With regard to the Window scale option, note that it is only useful if a window size greater than 64 kB is needed.

One potentially relevant TCP option in the context of CNNs is TCP Fast Open (TFO) [RFC7413]. As described in Section 5.3, TFO can be used to address the problem of traversing middleboxes that perform early filter state record deletion.

#### 4.2.3. Delayed Acknowledgments for single-MSS stacks

TCP Delayed Acknowledgments are meant to reduce the number of ACKs sent within a TCP connection, thus reducing network overhead, but they may increase the time until a sender may receive an ACK. In general, usefulness of Delayed ACKs depends heavily on the usage scenario. There can be interactions with stacks that use single-MSS windows.

A device that advertises a single-MSS receive window should avoid use of Delayed ACKs in order to avoid contributing unnecessary delay (of up to 500 ms) to the RTT [RFC5681], which limits the throughput and can increase the data delivery time.

A device that can send at most one MSS of data is significantly affected if the receiver uses Delayed ACKs, e.g., if a TCP server or receiver is outside the CNN. One known workaround is to split the data to be sent into two segments of smaller size. A standard compliant TCP receiver will then immediately acknowledge the second segment, which can improve throughput. This "split hack" works if the TCP receiver uses Delayed ACKs, but the downside is the overhead of sending two IP packets instead of one.

Similar issues happen when the sender uses the Nagle algorithm. Disabling the algorithm will not have impact if the sender can only handle stop-and-wait operation.

#### 4.2.4. RTO estimation for single-MSS stacks

The Retransmission Timeout (RTO) estimation is one of the fundamental TCP algorithms. There is a fundamental trade-off: A short, aggressive RTO behavior reduces wait time before retransmissions, but it also increases the probability of spurious timeouts. The latter lead to unnecessary waste of potentially scarce resources in CNNs such as energy and bandwidth. In contrast, a conservative timeout can result in long error recovery times and thus needlessly delay data delivery.

[RFC6298] describes the standard TCP RTO algorithm. If a TCP sender uses very small window size, and it cannot use Fast Retransmit/Fast Recovery or SACK, the Retransmission Timeout (RTO) algorithm has a larger impact on performance than for a more powerful TCP stack. In that case, RTO algorithm tuning may be considered, although careful assessment of possible drawbacks is recommended [I-D.ietf-tcpm-rto-consider].



As an example, an adaptive RTO algorithm for CoAP over UDP has been defined [I-D.ietf-core-cocoa] that has been found to perform well in CNN scenarios [Commag].

#### 4.3. General recommendations for TCP in CNNs

This section summarizes some widely used techniques to improve TCP, with a focus on their use in CNNs. The TCP extensions discussed here are useful in a wide range of network scenarios, including CNNs. This section is not comprehensive. A comprehensive survey of TCP extensions is published in [RFC7414].

##### 4.3.1. Loss recovery and congestion/flow control

Devices that have enough memory to allow larger TCP window size can leverage a more efficient loss recovery using Fast Retransmit and Fast Recovery [RFC5681], at the expense of slightly greater complexity and TCB size. Assuming that Delayed ACKs are used by the receiver, the mentioned algorithms work efficiently for window sizes of at least 5 MSS: If in a given TCP transmission of segments 1,2,3,4,5, and 6 the segment 2 gets lost, the sender should get an ACK for segment 1 when 3 arrives and duplicate acknowledgements when 4, 5, and 6 arrive. It will retransmit segment 2 when the third duplicate ACK arrives. In order to have segment 2, 3, 4, 5, and 6 sent, the window has to be at least five. With an MSS of 1220 byte, a buffer of the size of 5 MSS would require 6100 bytes.

For bulk data transfers further TCP improvements may also be useful, such as limited transmit [RFC3042].

##### 4.3.1.1. Selective Acknowledgments (SACK)

If a device with less severe memory and processing constraints can afford advertising a TCP window size of several MSS, it makes sense to support the SACK option to improve performance. SACK allows a data receiver to inform the data sender of non-contiguous data blocks received, thus a sender (having previously sent the SACK-Permitted option) can avoid performing unnecessary retransmissions, saving energy and bandwidth, as well as reducing latency. SACK is particularly useful for bulk data transfers. The receiver supporting SACK will need to manage the reception of possible out-of-order received segments, requiring sufficient buffer space. SACK adds  $8*n+2$  bytes to the TCP header, where  $n$  denotes the number of data blocks received, up to 4 blocks. For a low number of out-of-order segments, the header overhead penalty of SACK is compensated by avoiding unnecessary retransmissions.

#### 4.3.2. Delayed Acknowledgments

For certain traffic patterns, Delayed ACKs may have a detrimental effect, as already noted in Section 4.2.3. Advanced TCP stacks may use heuristics to determine the maximum delay for an ACK. For CNNs, the recommendation depends on the expected communication patterns.

If a stack is able to deal with more than one MSS of data, it may make sense to use a small timeout or disable delayed ACKs when traffic over a CNN is expected to mostly be small messages with a size typically below one MSS. For request-response traffic between a constrained device and a peer (e.g. backend infrastructure) that uses delayed ACKs, the maximum ACK rate of the peer will be typically of one ACK every 200 ms (or even lower). If in such conditions the peer device is administered by the same entity managing the constrained device, it is recommended to disable delayed ACKs at the peer side.

In contrast, Delayed ACKs allow to reduce the number of ACKs in bulk transfer type of traffic, e.g. for firmware/software updates or for transferring larger data units containing a batch of sensor readings.

Note that, in many scenarios, the peer that a constrained device communicates with will be a general purpose system that communicates with both constrained and unconstrained devices. Since delayed ACKs are often configured through system-wide parameters, delayed ACKs behavior at the peer will be the same regardless of the nature of the endpoints it talks to. Such a peer will typically have delayed ACKs enabled.

### 5. TCP usage recommendations in CNNs

This section discusses how a TCP stack can be used by applications that are developed for CNN scenarios. These remarks are by and large independent of how TCP is exactly implemented.

#### 5.1. TCP connection initiation

In the constrained device to unconstrained device scenario illustrated above, a TCP connection is typically initiated by the constrained device, in order for this device to support possible sleep periods to save energy.

#### 5.2. Number of concurrent connections

TCP endpoints with a small amount of RAM may only support a small number of connections. Each TCP connection requires storing a number of variables in the Transmission Control Block (TCB). Depending on the internal TCP implementation, each connection may result in

further memory overhead, and connections may compete for scarce resources (e.g. further memory overhead for send and receive buffers, etc).

A careful application design may try to keep the number of concurrent connections as small as possible. A client can for instance limit the number of simultaneous open connections that it maintains to a given server. Multiple connections could for instance be used to avoid the "head-of-line blocking" problem in an application transfer. However, in addition to consuming resources, using multiple connections can also cause undesirable side effects in congested networks. For example, the HTTP/1.1 specification encourages clients to be conservative when opening multiple connections [RFC7230]. Furthermore, each new connection will start with a 3-way handshake, therefore increasing message overhead.

Being conservative when opening multiple TCP connections is of particular importance in Constrained-Node Networks.

### 5.3. TCP connection lifetime

In order to minimize message overhead, it makes sense to keep a TCP connection open as long as the two TCP endpoints have more data to send. If applications exchange data rather infrequently, i.e., if TCP connections would stay idle for a long time, the idle time can result in problems. For instance, certain middleboxes such as firewalls or NAT devices are known to delete state records after an inactivity interval typically in the order of a few minutes [RFC6092]. The timeout duration used by a middlebox implementation may not be known to the TCP endpoints.

In CNNs, such middleboxes may e.g. be present at the boundary between the CNN and other networks. If the middlebox can be optimized for CNN use cases, it makes sense to increase the initial value for filter state inactivity timers to avoid problems with idle connections. Apart from that, this problem can be dealt with by different connection handling strategies, each having pros and cons.

One approach for infrequent data transfer is to use short-lived TCP connections. Instead of trying to maintain a TCP connection for long time, possibly short-lived connections can be opened between two endpoints, which are closed if no more data needs to be exchanged. For use cases that can cope with the additional messages and the latency resulting from starting new connections, it is recommended to use a sequence of short-lived connections, instead of maintaining a single long-lived connection.

The message and latency overhead that stems from using a sequence of short-lived connections could be reduced by TCP Fast Open (TFO) [RFC7413], which is an experimental TCP extension, at the expense of increased implementation complexity and increased TCP Control Block (TCB) size. TFO allows data to be carried in SYN (and SYN-ACK) segments, and to be consumed immediately by the receiving endpoint. This reduces the message and latency overhead compared to the traditional three-way handshake to establish a TCP connection. For security reasons, the connection initiator has to request a TFO cookie from the other endpoint. The cookie, with a size of 4 or 16 bytes, is then included in SYN packets of subsequent connections. The cookie needs to be refreshed (and obtained by the client) after a certain amount of time. Nevertheless, TFO is more efficient than frequently opening new TCP connections with the traditional three-way handshake, as long as the cookie can be reused in subsequent connections. However, as stated in RFC 7413, TFO deviates from the standard TCP semantics, since the data in the SYN could be replayed to an application in some rare circumstances. Applications should not use TFO unless they can tolerate this issue, e.g., by using Transport Layer Security (TLS) [RFC7413]. A comprehensive discussion on TFO can be found at RFC 7413.

Another approach is to use long-lived TCP connections with application-layer heartbeat messages. Various application protocols support such heartbeat messages. Periodic heartbeats requires transmission of packets, but they also allow aliveness checks at application level. In addition, they can prevent early filter state record deletion in middleboxes. In general, it makes sense realize aliveness checks at the highest protocol layer possible that is meaningful to the application, in order to maximize the depth of the aliveness check. In addition, timely detection of a dead peer may allow savings in terms of TCB memory use.

A TCP implementation may also be able to send "keep-alive" segments to test a TCP connection. According to [RFC1122], "keep-alives" are an optional TCP mechanism that is turned off by default, i.e., an application must explicitly enable it for a TCP connection. The interval between "keep-alive" messages must be configurable and it must default to no less than two hours. With this large timeout, TCP keep-alive messages are not very useful to avoid deletion of filter state records in middleboxes such as firewalls. However, sending TCP keep-alive probes more frequently risks draining power on energy-constrained devices.

## 6. Security Considerations

Best current practise for securing TCP and TCP-based communication also applies to CNN. As example, use of Transport Layer Security (TLS) is strongly recommended if it is applicable.

There are also TCP options which can improve TCP security. One example is the TCP Authentication Option (TCP-AO) [RFC5925]. However, this option adds overhead and complexity. TCP-AO typically has a size of 16-20 bytes.

For the mechanisms discussed in this document, the corresponding considerations apply. For instance, if TFO is used, the security considerations of [RFC7413] apply.

Constrained devices are expected to support smaller TCP window sizes than less limited devices. In such conditions, segment retransmission triggered by RTO expiration is expected to be relatively frequent, due to lack of (enough) duplicate ACKs, especially when a constrained device uses a single-MSS window size. For this reason, constrained devices running TCP may appear as particularly appealing victims of the so-called "shrew" Denial of Service (DoS) attack [shrew], whereby one or more sources generate a packet spike targetted to coincide with consecutive RTO-expiration-triggered retry attempts of a victim node. Note that the attack may be performed by Internet-connected devices, including constrained devices in the same CNN as the victim, as well as remote ones. Mitigation techniques include RTO randomization and attack blocking by routers able to detect shrew attacks based on their traffic pattern.

## 7. Acknowledgments

Carles Gomez has been funded in part by the Spanish Government (Ministerio de Educacion, Cultura y Deporte) through the Jose Castillejo grants CAS15/00336 and CAS18/00170, and by European Regional Development Fund (ERDF) and the Spanish Government through project TEC2016-79988-P, AEI/FEDER, UE. Part of his contribution to this work has been carried out during his stays as a visiting scholar at the Computer Laboratory of the University of Cambridge.

The authors appreciate the feedback received for this document. The following folks provided comments that helped improve the document: Carsten Bormann, Zhen Cao, Wei Genyu, Ari Keranen, Abhijan Bhattacharyya, Andres Arcia-Moret, Yoshifumi Nishida, Joe Touch, Fred Baker, Nik Sultana, Kerry Lynn, Erik Nordmark, Markku Kojo, Hannes Tschofenig, David Black, Yoshifumi Nishida, Ilpo Jarvinen and Emmanuel Baccelli. Simon Brummer provided details, and kindly

performed RAM and ROM usage measurements, on the RIOT TCP implementation. Xavi Vilajosana provided details on the OpenWSN TCP implementation. Rahul Jadhav kindly performed code size measurements on the Contiki-NG and lwIP 2.1.2 TCP implementations. He also provided details on the uIP TCP implementation.

## 8. Annex. TCP implementations for constrained devices

This section overviews the main features of TCP implementations for constrained devices. The survey is limited to open source stacks with small footprint. It is not meant to be all-encompassing. For more powerful embedded systems (e.g., with 32-bit processors), there are further stacks that comprehensively implement TCP. On the other hand, please be aware that this Annex is based on information available as of the writing.

### 8.1. uIP

uIP is a TCP/IP stack, targetted for 8 and 16-bit microcontrollers, which pioneered TCP/IP implementations for constrained devices. uIP has been deployed with Contiki and the Arduino Ethernet shield. A code size of ~5 kB (which comprises checksumming, IP, ICMP and TCP) has been reported for uIP [Dunk].

uIP uses the same global buffer for both incoming and outgoing traffic, which has a size of a single packet. In case of a retransmission, an application must be able to reproduce the same user data that had been transmitted. Multiple connections are supported, but need to share the global buffer.

The MSS is announced via the MSS option on connection establishment and the receive window size (of one MSS) is not modified during a connection. Stop-and-wait operation is used for sending data. Among other optimizations, this allows to avoid sliding window operations, which use 32-bit arithmetic extensively and are expensive on 8-bit CPUs.

Contiki uses the "split hack" technique (see Section 4.2.3) to avoid Delayed ACKs for senders using a single segment.

The code size of the TCP implementation in Contiki-NG has been measured to be of 3.2 kB on CC2538DK, cross-compiling on Linux.

### 8.2. lwIP

lwIP is a TCP/IP stack, targetted for 8- and 16-bit microcontrollers. lwIP has a total code size of ~14 kB to ~22 kB (which comprises

memory management, checksumming, network interfaces, IP, ICMP and TCP), and a TCP code size of ~9 kB to ~14 kB [Dunk].

In contrast with uIP, lwIP decouples applications from the network stack. lwIP supports a TCP transmission window greater than a single segment, as well as buffering of incoming and outgoing data. Other implemented mechanisms comprise slow start, congestion avoidance, fast retransmit and fast recovery. SACK and Window Scale support has been recently added to lwIP.

### 8.3. RIOT

The RIOT TCP implementation (called GNRC TCP) has been designed for Class 1 devices [RFC 7228]. The main target platforms are 8- and 16-bit microcontrollers, with 32-bit platforms also supported. GNRC TCP offers a similar function set as uIP, but it provides and maintains an independent receive buffer for each connection. In contrast to uIP, retransmission is also handled by GNRC TCP. GNRC TCP uses a single-MSS window size, which simplifies the implementation. The application programmer does not need to know anything about the TCP internals, therefore GNRC TCP can be seen as a user-friendly uIP TCP implementation.

The MSS is set on connections establishment and cannot be changed during connection lifetime. GNRC TCP allows multiple connections in parallel, but each TCB must be allocated somewhere in the system. By default there is only enough memory allocated for a single TCP connection, but it can be increased at compile time if the user needs multiple parallel connections.

The RIOT TCP implementation offers an optional POSIX socket wrapper that enables POSIX compliance, if needed.

Further details on RIOT and GNRC can be found in the literature [RIOT], [GNRC].

### 8.4. TinyOS

TinyOS was important as platform for early constrained devices. TinyOS has an experimental TCP stack that uses a simple nonblocking library-based implementation of TCP, which provides a subset of the socket interface primitives. The application is responsible for buffering. The TCP library does not do any receive-side buffering. Instead, it will immediately dispatch new, in-order data to the application and otherwise drop the segment. A send buffer is provided by the application. Multiple TCP connections are possible. Recently there has been little further work on the stack.

#### 8.5. FreeRTOS

FreeRTOS is a real-time operating system kernel for embedded devices that is supported by 16- and 32-bit microprocessors. Its TCP implementation is based on multiple-segment window size, although a 'Tiny-TCP' option, which is a single-MSS variant, can be enabled. Delayed ACKs are supported, with a 20-ms Delayed ACK timer as a technique intended 'to gain performance'.

#### 8.6. uC/OS

uC/OS is a real-time operating system kernel for embedded devices, which is maintained by Micrium. uC/OS is intended for 8-, 16- and 32-bit microprocessors. The uC/OS TCP implementation supports a multiple-segment window size.

#### 8.7. Summary



		uIP	lwIP orig	lwIP 2.1	RIOT	TinyOS	FreeRTOS	uC/OS
Memory	Code size (kB)	<5 (a)	~9 to ~14 (T1)	38 (T4)	<7 (T3)	N/A	<9.2 (T2)	N/A
TCP features	Single-Segm.	Yes	No	No	Yes	No	No	No
	Slow start	No	Yes	Yes	No	Yes	No	Yes
	Fast rec/retr	No	Yes	Yes	No	Yes	No	Yes
	Keep-alive	No	No	Yes	No	No	Yes	Yes
	Win. Scale	No	No	Yes	No	No	Yes	No
	TCP timest.	No	No	Yes	No	No	Yes	No
	SACK	No	No	Yes	No	No	Yes	No
	Del. ACKs	No	Yes	Yes	No	No	Yes	Yes
	Socket	No	No	Optional	(I)	Subset	Yes	Yes
	Concur. Conn.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
TLS supported		No	No	Yes	Yes	Yes	Yes	Yes

(T1) = TCP-only, on x86 and AVR platforms

(T2) = TCP-only, on ARM Cortex-M platform

(T3) = TCP-only, on ARM Cortex-M0+ platform (NOTE: RAM usage for the same platform

is ~2.5 kB for one TCP connection plus ~1.2 kB for each additional connection)

(T4) = TCP-only, on CC2538DK, cross-compiling on Linux

(a) = includes IP, ICMP and TCP on x86 and AVR platforms. The Contiki-NG TCP implementation has a code size of 3.2 kB on CC2538DK, cross-compiling on Linux

(I) = optional POSIX socket wrapper which enables POSIX compliance if needed

Mult. = Multiple

N/A = Not Available

Figure 2: Summary of TCP features for different lightweight TCP implementations. None of the implementations considered in this Annex support ECN or TFO.

## 9. Annex. Changes compared to previous versions

RFC Editor: To be removed prior to publication

### 9.1. Changes between -00 and -01

- o Changed title and abstract
- o Clarification that communication with standard-compliant TCP endpoints is required, based on feedback from Joe Touch
- o Additional discussion on communication patterns
- o Numerous changes to address a comprehensive review from Hannes Tschofenig
- o Reworded security considerations
- o Additional references and better distinction between normative and informative entries
- o Feedback from Rahul Jadhav on the uIP TCP implementation
- o Basic data for the TinyOS TCP implementation added, based on source code analysis

### 9.2. Changes between -01 and -02

- o Added text to the Introduction section, and a reference, on traditional bad perception of TCP for IoT
- o Added sections on FreeRTOS and uC/OS
- o Updated TinyOS section
- o Updated summary table
- o Reorganized Section 4 (single-MSS vs multiple-MSS window size), some content now also in new Section 5

### 9.3. Changes between -02 and -03

- o Rewording to better explain the benefit of ECN
- o Additional context information on the surveyed implementations
- o Added details, but removed "Data size" row, in the summary table

- o Added discussion on shrew attacks

#### 9.4. Changes between -03 and -04

- o Addressing the remaining TODOs
- o Alignment of the wording on TCP "keep-alives" with related discussions in the IETF transport area
- o Added further discussion on delayed ACKs
- o Removed OpenWSN subsection from the Annex

#### 9.5. Changes between -04 and -05

- o Addressing comments by Yoshifumi Nishida
- o Removed mentioning MD5 as an example (comment by David Black)
- o Added memory footprint details of TCP implementations (Contiki-NG and lwIP 2.1.2) provided by Rahul Jadhav in the Annex
- o Addressed comments by Ilpo Jarvinen throughout the whole document
- o Improved the RIOT section in the Annex, based on feedback from Emmanuel Baccelli

### 10. References

#### 10.1. Normative References

- [RFC0793] Postel, J., "Transmission Control Protocol", STD 7, RFC 793, DOI 10.17487/RFC0793, September 1981, <<https://www.rfc-editor.org/info/rfc793>>.
- [RFC1122] Braden, R., Ed., "Requirements for Internet Hosts - Communication Layers", STD 3, RFC 1122, DOI 10.17487/RFC1122, October 1989, <<https://www.rfc-editor.org/info/rfc1122>>.
- [RFC2018] Mathis, M., Mahdavi, J., Floyd, S., and A. Romanow, "TCP Selective Acknowledgment Options", RFC 2018, DOI 10.17487/RFC2018, October 1996, <<https://www.rfc-editor.org/info/rfc2018>>.

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <<https://www.rfc-editor.org/info/rfc2119>>.
- [RFC2460] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", RFC 2460, DOI 10.17487/RFC2460, December 1998, <<https://www.rfc-editor.org/info/rfc2460>>.
- [RFC3042] Allman, M., Balakrishnan, H., and S. Floyd, "Enhancing TCP's Loss Recovery Using Limited Transmit", RFC 3042, DOI 10.17487/RFC3042, January 2001, <<https://www.rfc-editor.org/info/rfc3042>>.
- [RFC3168] Ramakrishnan, K., Floyd, S., and D. Black, "The Addition of Explicit Congestion Notification (ECN) to IP", RFC 3168, DOI 10.17487/RFC3168, September 2001, <<https://www.rfc-editor.org/info/rfc3168>>.
- [RFC3819] Karn, P., Ed., Bormann, C., Fairhurst, G., Grossman, D., Ludwig, R., Mahdavi, J., Montenegro, G., Touch, J., and L. Wood, "Advice for Internet Subnetwork Designers", BCP 89, RFC 3819, DOI 10.17487/RFC3819, July 2004, <<https://www.rfc-editor.org/info/rfc3819>>.
- [RFC5681] Allman, M., Paxson, V., and E. Blanton, "TCP Congestion Control", RFC 5681, DOI 10.17487/RFC5681, September 2009, <<https://www.rfc-editor.org/info/rfc5681>>.
- [RFC5925] Touch, J., Mankin, A., and R. Bonica, "The TCP Authentication Option", RFC 5925, DOI 10.17487/RFC5925, June 2010, <<https://www.rfc-editor.org/info/rfc5925>>.
- [RFC6298] Paxson, V., Allman, M., Chu, J., and M. Sargent, "Computing TCP's Retransmission Timer", RFC 6298, DOI 10.17487/RFC6298, June 2011, <<https://www.rfc-editor.org/info/rfc6298>>.
- [RFC7228] Bormann, C., Ersue, M., and A. Keranen, "Terminology for Constrained-Node Networks", RFC 7228, DOI 10.17487/RFC7228, May 2014, <<https://www.rfc-editor.org/info/rfc7228>>.
- [RFC7323] Borman, D., Braden, B., Jacobson, V., and R. Scheffenegger, Ed., "TCP Extensions for High Performance", RFC 7323, DOI 10.17487/RFC7323, September 2014, <<https://www.rfc-editor.org/info/rfc7323>>.

- [RFC7413] Cheng, Y., Chu, J., Radhakrishnan, S., and A. Jain, "TCP Fast Open", RFC 7413, DOI 10.17487/RFC7413, December 2014, <<https://www.rfc-editor.org/info/rfc7413>>.

## 10.2. Informative References

- [Commag] A. Betzler, C. Gomez, I. Demirkol, J. Paradells, "CoAP Congestion Control for the Internet of Things", IEEE Communications Magazine, June 2016.
- [Dunk] A. Dunkels, "Full TCP/IP for 8-Bit Architectures", 2003.
- [ETEN] R. Krishnan et al, "Explicit transport error notification (ETEN) for error-prone wireless and satellite networks", Computer Networks 2004.
- [GNRC] M. Lenders et al., "Connecting the World of Embedded Mobiles: The RIOTApproach to Ubiquitous Networking for the IoT", 2018.
- [I-D.delcarpio-6lo-wlanah]  
Vega, L., Robles, I., and R. Morabito, "IPv6 over 802.11ah", draft-delcarpio-6lo-wlanah-01 (work in progress), October 2015.
- [I-D.ietf-core-cocoa]  
Bormann, C., Betzler, A., Gomez, C., and I. Demirkol, "CoAP Simple Congestion Control/Advanced", draft-ietf-core-cocoa-03 (work in progress), February 2018.
- [I-D.ietf-tcpm-rto-consider]  
Allman, M., "Retransmission Timeout Requirements", draft-ietf-tcpm-rto-consider-08 (work in progress), February 2019.
- [IntComp] C. Gomez, A. Arcia-Moret, J. Crowcroft, "TCP in the Internet of Things: from ostracism to prominence", IEEE Internet Computing, January-February 2018.
- [RFC2757] Montenegro, G., Dawkins, S., Kojo, M., Magret, V., and N. Vaidya, "Long Thin Networks", RFC 2757, DOI 10.17487/RFC2757, January 2000, <<https://www.rfc-editor.org/info/rfc2757>>.
- [RFC2884] Hadi Salim, J. and U. Ahmed, "Performance Evaluation of Explicit Congestion Notification (ECN) in IP Networks", RFC 2884, DOI 10.17487/RFC2884, July 2000, <<https://www.rfc-editor.org/info/rfc2884>>.

- [RFC3481] Inamura, H., Ed., Montenegro, G., Ed., Ludwig, R., Gurtov, A., and F. Khafizov, "TCP over Second (2.5G) and Third (3G) Generation Wireless Networks", BCP 71, RFC 3481, DOI 10.17487/RFC3481, February 2003, <<https://www.rfc-editor.org/info/rfc3481>>.
- [RFC4944] Montenegro, G., Kushalnagar, N., Hui, J., and D. Culler, "Transmission of IPv6 Packets over IEEE 802.15.4 Networks", RFC 4944, DOI 10.17487/RFC4944, September 2007, <<https://www.rfc-editor.org/info/rfc4944>>.
- [RFC6077] Papadimitriou, D., Ed., Welzl, M., Scharf, M., and B. Briscoe, "Open Research Issues in Internet Congestion Control", RFC 6077, DOI 10.17487/RFC6077, February 2011, <<https://www.rfc-editor.org/info/rfc6077>>.
- [RFC6092] Woodyatt, J., Ed., "Recommended Simple Security Capabilities in Customer Premises Equipment (CPE) for Providing Residential IPv6 Internet Service", RFC 6092, DOI 10.17487/RFC6092, January 2011, <<https://www.rfc-editor.org/info/rfc6092>>.
- [RFC6120] Saint-Andre, P., "Extensible Messaging and Presence Protocol (XMPP): Core", RFC 6120, DOI 10.17487/RFC6120, March 2011, <<https://www.rfc-editor.org/info/rfc6120>>.
- [RFC6606] Kim, E., Kaspar, D., Gomez, C., and C. Bormann, "Problem Statement and Requirements for IPv6 over Low-Power Wireless Personal Area Network (6LoWPAN) Routing", RFC 6606, DOI 10.17487/RFC6606, May 2012, <<https://www.rfc-editor.org/info/rfc6606>>.
- [RFC7230] Fielding, R., Ed. and J. Reschke, Ed., "Hypertext Transfer Protocol (HTTP/1.1): Message Syntax and Routing", RFC 7230, DOI 10.17487/RFC7230, June 2014, <<https://www.rfc-editor.org/info/rfc7230>>.
- [RFC7252] Shelby, Z., Hartke, K., and C. Bormann, "The Constrained Application Protocol (CoAP)", RFC 7252, DOI 10.17487/RFC7252, June 2014, <<https://www.rfc-editor.org/info/rfc7252>>.
- [RFC7414] Duke, M., Braden, R., Eddy, W., Blanton, E., and A. Zimmermann, "A Roadmap for Transmission Control Protocol (TCP) Specification Documents", RFC 7414, DOI 10.17487/RFC7414, February 2015, <<https://www.rfc-editor.org/info/rfc7414>>.

- [RFC7428] Brandt, A. and J. Buron, "Transmission of IPv6 Packets over ITU-T G.9959 Networks", RFC 7428, DOI 10.17487/RFC7428, February 2015, <<https://www.rfc-editor.org/info/rfc7428>>.
- [RFC7540] Belshe, M., Peon, R., and M. Thomson, Ed., "Hypertext Transfer Protocol Version 2 (HTTP/2)", RFC 7540, DOI 10.17487/RFC7540, May 2015, <<https://www.rfc-editor.org/info/rfc7540>>.
- [RFC7567] Baker, F., Ed. and G. Fairhurst, Ed., "IETF Recommendations Regarding Active Queue Management", BCP 197, RFC 7567, DOI 10.17487/RFC7567, July 2015, <<https://www.rfc-editor.org/info/rfc7567>>.
- [RFC7668] Nieminen, J., Savolainen, T., Isomaki, M., Patil, B., Shelby, Z., and C. Gomez, "IPv6 over BLUETOOTH(R) Low Energy", RFC 7668, DOI 10.17487/RFC7668, October 2015, <<https://www.rfc-editor.org/info/rfc7668>>.
- [RFC8087] Fairhurst, G. and M. Welzl, "The Benefits of Using Explicit Congestion Notification (ECN)", RFC 8087, DOI 10.17487/RFC8087, March 2017, <<https://www.rfc-editor.org/info/rfc8087>>.
- [RFC8105] Mariager, P., Petersen, J., Ed., Shelby, Z., Van de Logt, M., and D. Barthel, "Transmission of IPv6 Packets over Digital Enhanced Cordless Telecommunications (DECT) Ultra Low Energy (ULE)", RFC 8105, DOI 10.17487/RFC8105, May 2017, <<https://www.rfc-editor.org/info/rfc8105>>.
- [RFC8163] Lynn, K., Ed., Martocci, J., Neilson, C., and S. Donaldson, "Transmission of IPv6 over Master-Slave/Token-Passing (MS/TP) Networks", RFC 8163, DOI 10.17487/RFC8163, May 2017, <<https://www.rfc-editor.org/info/rfc8163>>.
- [RFC8201] McCann, J., Deering, S., Mogul, J., and R. Hinden, Ed., "Path MTU Discovery for IP version 6", STD 87, RFC 8201, DOI 10.17487/RFC8201, July 2017, <<https://www.rfc-editor.org/info/rfc8201>>.
- [RFC8323] Bormann, C., Lemay, S., Tschofenig, H., Hartke, K., Silverajan, B., and B. Raymor, Ed., "CoAP (Constrained Application Protocol) over TCP, TLS, and WebSockets", RFC 8323, DOI 10.17487/RFC8323, February 2018, <<https://www.rfc-editor.org/info/rfc8323>>.

- [RFC8352] Gomez, C., Kovatsch, M., Tian, H., and Z. Cao, Ed., "Energy-Efficient Features of Internet of Things Protocols", RFC 8352, DOI 10.17487/RFC8352, April 2018, <<https://www.rfc-editor.org/info/rfc8352>>.
- [RFC8376] Farrell, S., Ed., "Low-Power Wide Area Network (LPWAN) Overview", RFC 8376, DOI 10.17487/RFC8376, May 2018, <<https://www.rfc-editor.org/info/rfc8376>>.
- [RIOT] E. Baccelli et al., "RIOT: an Open Source Operating System for Low-end Embedded Devices in the IoT", 2018.
- [shrew] A. Kuzmanovic, E. Knightly, "Low-Rate TCP-Targeted Denial of Service Attacks", SIGCOMM'03 2003.

## Authors' Addresses

Carles Gomez  
UPC  
C/Esteve Terradas, 7  
Castelldefels 08860  
Spain  
  
Email: carlesgo@entel.upc.edu

Jon Crowcroft  
University of Cambridge  
JJ Thomson Avenue  
Cambridge, CB3 0FD  
United Kingdom  
  
Email: jon.crowcroft@cl.cam.ac.uk

Michael Scharf  
Hochschule Esslingen  
Flandernstr. 101  
Esslingen 73732  
Germany  
  
Email: michael.scharf@hs-esslingen.de



LWIG Working Group  
Internet-Draft  
Intended status: Informational  
Expires: May 3, 2021

C. Gomez  
UPC  
J. Crowcroft  
University of Cambridge  
M. Scharf  
Hochschule Esslingen  
October 30, 2020

TCP Usage Guidance in the Internet of Things (IoT)  
draft-ietf-lwig-tcp-constrained-node-networks-13

## Abstract

This document provides guidance on how to implement and use the Transmission Control Protocol (TCP) in Constrained-Node Networks (CNs), which are a characteristic of the Internet of Things (IoT). Such environments require a lightweight TCP implementation and may not make use of optional functionality. This document explains a number of known and deployed techniques to simplify a TCP stack as well as corresponding tradeoffs. The objective is to help embedded developers with decisions on which TCP features to use.

## 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 May 3, 2021.

## Copyright Notice

Copyright (c) 2020 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 Provisions Relating to IETF Documents (<https://trustee.ietf.org/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 Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

## Table of Contents

1. Introduction . . . . .	3
2. Characteristics of CNNs relevant for TCP . . . . .	4
2.1. Network and link properties . . . . .	4
2.2. Usage scenarios . . . . .	5
2.3. Communication and traffic patterns . . . . .	6
3. TCP implementation and configuration in CNNs . . . . .	6
3.1. Addressing path properties . . . . .	7
3.1.1. Maximum Segment Size (MSS) . . . . .	7
3.1.2. Explicit Congestion Notification (ECN) . . . . .	8
3.1.3. Explicit loss notifications . . . . .	9
3.2. TCP guidance for single-MSS stacks . . . . .	9
3.2.1. Single-MSS stacks - benefits and issues . . . . .	9
3.2.2. TCP options for single-MSS stacks . . . . .	10
3.2.3. Delayed Acknowledgments for single-MSS stacks . . . . .	10
3.2.4. RTO calculation for single-MSS stacks . . . . .	11
3.3. General recommendations for TCP in CNNs . . . . .	12
3.3.1. Loss recovery and congestion/flow control . . . . .	12
3.3.1.1. Selective Acknowledgments (SACK) . . . . .	13
3.3.2. Delayed Acknowledgments . . . . .	13
3.3.3. Initial Window . . . . .	14
4. TCP usage recommendations in CNNs . . . . .	14
4.1. TCP connection initiation . . . . .	14
4.2. Number of concurrent connections . . . . .	15
4.3. TCP connection lifetime . . . . .	15
5. Security Considerations . . . . .	17
6. Acknowledgments . . . . .	18
7. Annex. TCP implementations for constrained devices . . . . .	18
7.1. uIP . . . . .	19
7.2. lwIP . . . . .	19
7.3. RIOT . . . . .	19
7.4. TinyOS . . . . .	20
7.5. FreeRTOS . . . . .	20
7.6. uC/OS . . . . .	20
7.7. Summary . . . . .	21
8. Annex. Changes compared to previous versions . . . . .	22
8.1. Changes between -00 and -01 . . . . .	22
8.2. Changes between -01 and -02 . . . . .	22
8.3. Changes between -02 and -03 . . . . .	22
8.4. Changes between -03 and -04 . . . . .	23

8.5.	Changes between -04 and -05 . . . . .	23
8.6.	Changes between -05 and -06 . . . . .	23
8.7.	Changes between -06 and -07 . . . . .	23
8.8.	Changes between -07 and -08 . . . . .	23
8.9.	Changes between -08 and -09 . . . . .	23
8.10.	Changes between -09 and -10 . . . . .	24
8.11.	Changes between -10 and -11 . . . . .	24
8.12.	Changes between -11 and -12 . . . . .	24
8.13.	Changes between -12 and -13 . . . . .	24
9.	References . . . . .	24
9.1.	Normative References . . . . .	24
9.2.	Informative References . . . . .	25
	Authors' Addresses . . . . .	30

## 1. Introduction

The Internet Protocol suite is being used for connecting Constrained-Node Networks (CNs) to the Internet, enabling the so-called Internet of Things (IoT) [RFC7228]. In order to meet the requirements that stem from CNs, the IETF has produced a suite of new protocols specifically designed for such environments (see e.g. [RFC8352]). New IETF protocol stack components include the IPv6 over Low-power Wireless Personal Area Networks (6LoWPAN) adaptation layer [RFC4944][RFC6282][RFC6775], the IPv6 Routing Protocol for Low-power and lossy networks (RPL) routing protocol [RFC6550], and the Constrained Application Protocol (CoAP) [RFC7252].

As of the writing, the main current transport layer protocols in IP-based IoT scenarios are UDP and TCP. TCP has been criticized, often unfairly, as a protocol that is unsuitable for the IoT. It is true that some TCP features, such as relatively long header size, unsuitability for multicast, and always-confirmed data delivery, are not optimal for IoT scenarios. However, many typical claims on TCP unsuitability for IoT (e.g. a high complexity, connection-oriented approach incompatibility with radio duty-cycling, and spurious congestion control activation in wireless links) are not valid, can be solved, or are also found in well accepted IoT end-to-end reliability mechanisms (see [IntComp] for a detailed analysis).

At the application layer, CoAP was developed over UDP [RFC7252]. However, the integration of some CoAP deployments with existing infrastructure is being challenged by middleboxes such as firewalls, which may limit and even block UDP-based communications. This is the main reason why a CoAP over TCP specification has been developed [RFC8323].

Other application layer protocols not specifically designed for CNs are also being considered for the IoT space. Some examples include

HTTP/2 and even HTTP/1.1, both of which run over TCP by default [RFC7230] [RFC7540], and the Extensible Messaging and Presence Protocol (XMPP) [RFC6120]. TCP is also used by non-IETF application-layer protocols in the IoT space such as the Message Queuing Telemetry Transport (MQTT) [MQTT] and its lightweight variants.

TCP is a sophisticated transport protocol that includes optional functionality (e.g. TCP options) that may improve performance in some environments. However, many optional TCP extensions require complex logic inside the TCP stack and increase the code size and the memory requirements. Many TCP extensions are not required for interoperability with other standard-compliant TCP endpoints. Given the limited resources on constrained devices, careful selection of optional TCP features can make an implementation more lightweight.

This document provides guidance on how to implement and configure TCP, as well as on how TCP is advisable to be used by applications, in CNNs. The overarching goal is to offer simple measures to allow for lightweight TCP implementation and suitable operation in such environments. A TCP implementation following the guidance in this document is intended to be compatible with a TCP endpoint that is compliant to the TCP standards, albeit possibly with a lower performance. This implies that such a TCP client would always be able to connect with a standard-compliant TCP server, and a corresponding TCP server would always be able to connect with a standard-compliant TCP client.

This document assumes that the reader is familiar with TCP. A comprehensive survey of the TCP standards can be found in [RFC7414]. Similar guidance regarding the use of TCP in special environments has been published before, e.g., for cellular wireless networks [RFC3481].

## 2. Characteristics of CNNs relevant for TCP

### 2.1. Network and link properties

CNNs are defined in [RFC7228] as networks whose characteristics are influenced by being composed of a significant portion of constrained nodes. The latter are characterized by significant limitations on processing, memory, and energy resources, among others [RFC7228]. The first two dimensions pose constraints on the complexity and on the memory footprint of the protocols that constrained nodes can support. The latter requires techniques to save energy, such as radio duty-cycling in wireless devices [RFC8352], as well as minimization of the number of messages transmitted/received (and their size).

[RFC7228] lists typical network constraints in CNN, including low achievable bitrate/throughput, high packet loss and high variability of packet loss, highly asymmetric link characteristics, severe penalties for using larger packets, limits on reachability over time, etc. CNN may use wireless or wired technologies (e.g., Power Line Communication), and the transmission rates are typically low (e.g. below 1 Mbps).

For use of TCP, one challenge is that not all technologies in CNN may be aligned with typical Internet subnetwork design principles [RFC3819]. For instance, constrained nodes often use physical/link layer technologies that have been characterized as 'lossy', i.e., exhibit a relatively high bit error rate. Dealing with corruption loss is one of the open issues in the Internet [RFC6077].

## 2.2. Usage scenarios

There are different deployment and usage scenarios for CNNs. Some CNNs follow the star topology, whereby one or several hosts are linked to a central device that acts as a router connecting the CNN to the Internet. Alternatively, CNNs may also follow the multihop topology [RFC6606].

In constrained environments, there can be different types of devices [RFC7228]. For example, there can be devices with single combined send/receive buffer, devices with a separate send and receive buffer, or devices with a pool of multiple send/receive buffers. In the latter case, it is possible that buffers are also shared for other protocols.

One key use case for TCP in CNNs is a model where constrained devices connect to unconstrained servers in the Internet. But it is also possible that both TCP endpoints run on constrained devices. In the first case, communication possibly has to traverse a middlebox (e.g. a firewall, NAT, etc.). Figure 1 illustrates such a scenario. Note that the scenario is asymmetric, as the unconstrained device will typically not suffer the severe constraints of the constrained device. The unconstrained device is expected to be mains-powered, to have high amount of memory and processing power, and to be connected to a resource-rich network.

Assuming that a majority of constrained devices will correspond to sensor nodes, the amount of data traffic sent by constrained devices (e.g. sensor node measurements) is expected to be higher than the amount of data traffic in the opposite direction. Nevertheless, constrained devices may receive requests (to which they may respond), commands (for configuration purposes and for constrained devices

including actuators) and relatively infrequent firmware/software updates.

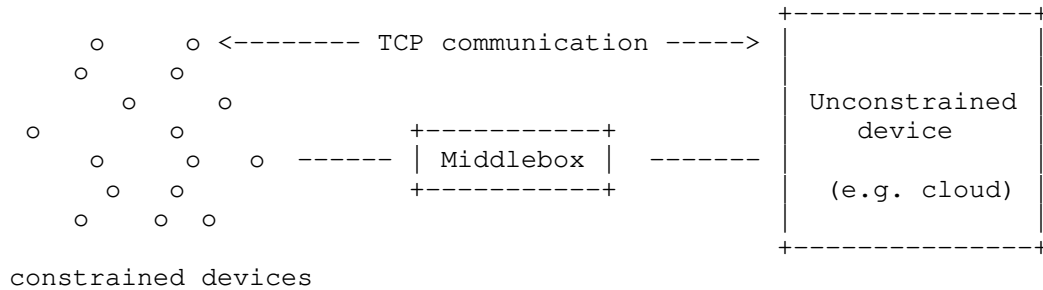


Figure 1: TCP communication between a constrained device and an unconstrained device, traversing a middlebox.

### 2.3. Communication and traffic patterns

IoT applications are characterized by a number of different communication patterns. The following non-comprehensive list explains some typical examples:

- o Unidirectional transfers: An IoT device (e.g. a sensor) can send (repeatedly) updates to the other endpoint. There is not always a need for an application response back to the IoT device.
- o Request-response patterns: An IoT device receiving a request from the other endpoint, which triggers a response from the IoT device.
- o Bulk data transfers: A typical example for a long file transfer would be an IoT device firmware update.

A typical communication pattern is that a constrained device communicates with an unconstrained device (cf. Figure 1). But it is also possible that constrained devices communicate amongst themselves.

### 3. TCP implementation and configuration in CNNs

This section explains how a TCP stack can deal with typical constraints in CNN. The guidance in this section relates to the TCP implementation and its configuration.

### 3.1. Addressing path properties

#### 3.1.1. Maximum Segment Size (MSS)

Assuming that IPv6 is used, and for the sake of lightweight implementation and operation, unless applications require handling large data units (i.e. leading to an IPv6 datagram size greater than 1280 bytes), it may be desirable to limit the IP datagram size to 1280 bytes in order to avoid the need to support Path MTU Discovery [RFC8201]. In addition, an IP datagram size of 1280 bytes avoids incurring IPv6-layer fragmentation [RFC8900].

An IPv6 datagram size exceeding 1280 bytes can be avoided by setting the TCP MSS not larger than 1220 bytes. Note that it is already a requirement that TCP implementations consume payload space instead of increasing datagram size when including IP or TCP options in an IP packet to be sent [RFC6691]. Therefore, it is not required to advertise an MSS smaller than 1220 bytes in order to accommodate TCP options.

Note that setting the MTU to 1280 bytes is possible for link layer technologies in the CNN space, even if some of them are characterized by a short data unit payload size, e.g. up to a few tens or hundreds of bytes. For example, the maximum frame size in IEEE 802.15.4 is 127 bytes. 6LoWPAN defined an adaptation layer to support IPv6 over IEEE 802.15.4 networks. The adaptation layer includes a fragmentation mechanism, since IPv6 requires the layer below to support an MTU of 1280 bytes [RFC8200], while IEEE 802.15.4 lacked fragmentation mechanisms. 6LoWPAN defines an IEEE 802.15.4 link MTU of 1280 bytes [RFC4944]. Other technologies, such as Bluetooth LE [RFC7668], ITU-T G.9959 [RFC7428] or DECT-ULE [RFC8105], also use 6LoWPAN-based adaptation layers in order to enable IPv6 support. These technologies do support link layer fragmentation. By exploiting this functionality, the adaptation layers that enable IPv6 over such technologies also define an MTU of 1280 bytes.

On the other hand, there exist technologies also used in the CNN space, such as Master Slave / Token Passing (TP) [RFC8163], Narrowband IoT (NB-IoT) [RFC8376] or IEEE 802.11ah [I-D.delcarpio-6lo-wlanah], that do not suffer the same degree of frame size limitations as the technologies mentioned above. The MTU for MS/TP is recommended to be 1500 bytes [RFC8163], the MTU in NB-IoT is 1600 bytes, and the maximum frame payload size for IEEE 802.11ah is 7991 bytes.

Using larger MSS (to a suitable extent) may be beneficial in some scenarios, especially when transferring large payloads, as it reduces the number of packets (and packet headers) required for a given

payload. However, the characteristics of the constrained network need to be considered. In particular, in a lossy network where unreliable fragment delivery is used, the amount of data that TCP unnecessarily retransmits due to fragment loss increases (and throughput decreases) quickly with the MSS. This happens because the loss of a fragment leads to the loss of the whole fragmented packet being transmitted. Unnecessary data retransmission is particularly harmful in CNNs due to the resource constraints of such environments. Note that, while the original 6LoWPAN fragmentation mechanism [RFC4944] does not offer reliable fragment delivery, fragment recovery functionality for 6LoWPAN or 6Lo environments is being standardized as of the writing [I-D.ietf-6lo-fragment-recovery].

### 3.1.2. Explicit Congestion Notification (ECN)

Explicit Congestion Notification (ECN) [RFC3168] ECN allows a router to signal in the IP header of a packet that congestion is arising, for example when a queue size reaches a certain threshold. An ECN-enabled TCP receiver will echo back the congestion signal to the TCP sender by setting a flag in its next TCP ACK. The sender triggers congestion control measures as if a packet loss had happened.

The document [RFC8087] outlines the principal gains in terms of increased throughput, reduced delay, and other benefits when ECN is used over a network path that includes equipment that supports Congestion Experienced (CE) marking. In the context of CNNs, a remarkable feature of ECN is that congestion can be signalled without incurring packet drops (which will lead to retransmissions and consumption of limited resources such as energy and bandwidth).

ECN can further reduce packet losses since congestion control measures can be applied earlier [RFC2884]. Fewer lost packets implies that the number of retransmitted segments decreases, which is particularly beneficial in CNNs, where energy and bandwidth resources are typically limited. Also, it makes sense to try to avoid packet drops for transactional workloads with small data sizes, which are typical for CNNs. In such traffic patterns, it is more difficult and often impossible to detect packet loss without retransmission timeouts (e.g., as there may be no three duplicate ACKs). Any retransmission timeout slows down the data transfer significantly. In addition, if the constrained device uses power saving techniques, a retransmission timeout will incur a wake-up action, in contrast to ACK clock- triggered sending. When the congestion window of a TCP sender has a size of one segment and a TCP ACK with an ECN signal (ECE flag) arrives at the TCP sender, the TCP sender resets the retransmit timer, and the sender will only be able to send a new packet when the retransmit timer expires. Effectively, the TCP sender reduces at that moment its sending rate from 1 segment per



Round Trip Time (RTT) to 1 segment per Retransmission Timeout (RTO) and reduces the sending rate further on each ECN signal received in subsequent TCP ACKs. Otherwise, if an ECN signal is not present in a subsequent TCP ACK the TCP sender resumes the normal ACK-clocked transmission of segments [RFC3168].

ECN can be incrementally deployed in the Internet. Guidance on configuration and usage of ECN is provided in [RFC7567]. Given the benefits, more and more TCP stacks in the Internet support ECN, and it specifically makes sense to leverage ECN in controlled environments such as CNNs. As of the writing, there is on-going work to extend the types of TCP packets that are ECN-capable, including pure ACKs [I-D.ietf-tcpm-generalized-ecn]. Such a feature may further increase the benefits of ECN in CNN environments. Note, however, that supporting ECN increases implementation complexity.

### 3.1.3. Explicit loss notifications

There has been a significant body of research on solutions capable of explicitly indicating whether a TCP segment loss is due to corruption, in order to avoid activation of congestion control mechanisms [ETEN] [RFC2757]. While such solutions may provide significant improvement, they have not been widely deployed and remain as experimental work. In fact, as of today, the IETF has not standardized any such solution.

## 3.2. TCP guidance for single-MSS stacks

This section discusses TCP stacks that allow transferring a single MSS. More general guidance is provided in Section 3.3.

### 3.2.1. Single-MSS stacks - benefits and issues

A TCP stack can reduce the memory requirements by advertising a TCP window size of one MSS, and also transmit at most one MSS of unacknowledged data. In that case, both congestion and flow control implementation are quite simple. Such a small receive and send window may be sufficient for simple message exchanges in the CNN space. However, only using a window of one MSS can significantly affect performance. A stop-and-wait operation results in low throughput for transfers that exceed the length of one MSS, e.g., a firmware download. Furthermore, a single-MSS solution relies solely on timer-based loss recovery, therefore missing the performance gain of Fast Retransmit and Fast Recovery (which require a larger window size, see Section 3.3.1).

If CoAP is used over TCP with the default setting for NSTART in [RFC7252], a CoAP endpoint is not allowed to send a new message to a

destination until a response for the previous message sent to that destination has been received. This is equivalent to an application-layer window size of 1 data unit. For this use of CoAP, a maximum TCP window of one MSS may be sufficient, as long as the CoAP message size does not exceed one MSS. An exception in CoAP over TCP, though, is the Capabilities and Settings Message (CSM) that must be sent at the start of the TCP connection. The first application message carrying user data is allowed to be sent immediately after the CSM message. If the sum of the CSM size plus the application message size exceeds the MSS, a sender using a single-MSS stack will need to wait for the ACK confirming the CSM before sending the application message.

### 3.2.2. TCP options for single-MSS stacks

A TCP implementation needs to support, at a minimum, TCP options 2, 1 and 0. These are, respectively, the Maximum Segment Size (MSS) option, the No-Operation option, and the End Of Option List marker [RFC0793]. None of these are a substantial burden to support. These options are sufficient for interoperability with a standard-compliant TCP endpoint, albeit many TCP stacks support additional options and can negotiate their use. A TCP implementation is permitted to silently ignore all other TCP options.

A TCP implementation for a constrained device that uses a single-MSS TCP receive or transmit window size may not benefit from supporting the following TCP options: Window scale [RFC7323], TCP Timestamps [RFC7323], Selective Acknowledgments (SACK) and SACK-Permitted [RFC2018]. Also other TCP options may not be required on a constrained device with a very lightweight implementation. With regard to the Window scale option, note that it is only useful if a window size greater than 64 kB is needed.

Note that a TCP sender can benefit from the TCP Timestamps option [RFC7323] in detecting spurious RTOs. The latter are quite likely to occur in CNN scenarios due to a number of reasons (e.g. route changes in a multihop scenario, link layer retries, etc.). The header overhead incurred by the Timestamps option (of up to 12 bytes) needs to be taken into account.

### 3.2.3. Delayed Acknowledgments for single-MSS stacks

TCP Delayed Acknowledgments are meant to reduce the number of ACKs sent within a TCP connection, thus reducing network overhead, but they may increase the time until a sender may receive an ACK. In general, usefulness of Delayed ACKs depends heavily on the usage scenario (see Section 3.3.2). There can be interactions with single-MSS stacks.

When traffic is unidirectional, if the sender can send at most one MSS of data or the receiver advertises a receive window not greater than the MSS, Delayed ACKs may unnecessarily contribute delay (up to 500 ms) to the RTT [RFC5681], which limits the throughput and can increase data delivery time. Note that, in some cases, it may not be possible to disable Delayed ACKs. One known workaround is to split the data to be sent into two segments of smaller size. A standard compliant TCP receiver may immediately acknowledge the second MSS of data, which can improve throughput. However, this 'split hack' may not always work since a TCP receiver is required to acknowledge every second full-sized segment, but not two consecutive small segments. The overhead of sending two IP packets instead of one is another downside of the 'split hack'.

Similar issues may happen when the sender uses the Nagle algorithm, since the sender may need to wait for an unnecessarily delayed ACK to send a new segment. Disabling the algorithm will not have impact if the sender can only handle stop-and-wait operation at the TCP level.

For request-response traffic, when the receiver uses Delayed ACKs, a response to a data message can piggyback an ACK, as long as the latter is sent before the Delayed ACK timer expires, thus avoiding unnecessary ACKs without payload. Disabling Delayed ACKs at the request sender allows an immediate ACK for the data segment carrying the response.

#### 3.2.4. RTO calculation for single-MSS stacks

The RTO calculation is one of the fundamental TCP algorithms [RFC6298]. There is a fundamental trade-off: A short, aggressive RTO behavior reduces wait time before retransmissions, but it also increases the probability of spurious timeouts. The latter lead to unnecessary waste of potentially scarce resources in CNNs such as energy and bandwidth. In contrast, a conservative timeout can result in long error recovery times and thus needlessly delay data delivery.

If a TCP sender uses a very small window size, and it cannot benefit from Fast Retransmit/Fast Recovery or SACK, the RTO algorithm has a large impact on performance. In that case, RTO algorithm tuning may be considered, although careful assessment of possible drawbacks is recommended [I-D.ietf-tcpm-rto-consider].

As an example, adaptive RTO algorithms defined for CoAP over UDP have been found to perform well in CNN scenarios [Commag] [I-D.ietf-core-fasor].

### 3.3. General recommendations for TCP in CNNs

This section summarizes some widely used techniques to improve TCP, with a focus on their use in CNNs. The TCP extensions discussed here are useful in a wide range of network scenarios, including CNNs. This section is not comprehensive. A comprehensive survey of TCP extensions is published in [RFC7414].

#### 3.3.1. Loss recovery and congestion/flow control

Devices that have enough memory to allow a larger (i.e. more than 3 MSS of data) TCP window size can leverage a more efficient loss recovery than the timer-based approach used for smaller TCP window size (see Section 3.2.1) by using Fast Retransmit and Fast Recovery [RFC5681], at the expense of slightly greater complexity and Transmission Control Block (TCB) size. Assuming that Delayed ACKs are used by the receiver, a window size of up to 5 MSS is required for Fast Retransmit and Fast Recovery to work efficiently: If in a given TCP transmission of full-sized segments 1, 2, 3, 4, and 5, segment 2 gets lost, and the ACK for segment 1 is held by the Delayed ACK timer, then the sender should get an ACK for segment 1 when 3 arrives and duplicate ACKs when segments 4, 5, and 6 arrive. It will retransmit segment 2 when the third duplicate ACK arrives. In order to have segments 2, 3, 4, 5, and 6 sent, the window has to be of at least 5 MSS. With an MSS of 1220 bytes, a buffer of a size of 5 MSS would require 6100 bytes.

The example in the previous paragraph did not use a further TCP improvement such as Limited Transmit [RFC3042]. The latter may also be useful for any transfer that has more than one segment in flight. Small transfers tend to benefit more from Limited Transmit, because they are more likely to not receive enough duplicate ACKs. Assuming the example in the previous paragraph, Limited Transmit allows sending 5 MSS with a congestion window (cwnd) of 3 segments, plus two additional segments for the first two duplicate ACKs. With Limited Transmit, even a cwnd of 2 segments allows sending 5 MSS, at the expense of additional delay contributed by the Delayed ACK timer for the ACK that confirms segment 1.

When a multiple-segment window is used, the receiver will need to manage the reception of possible out-of-order received segments, requiring sufficient buffer space. Note that even when a 1-MSS window is used, out-of-order arrival should also be managed, as the sender may send multiple sub-MSS packets that fit in the window. (On the other hand, the receiver is free to simply drop out-of-order segments, thus forcing retransmissions).

#### 3.3.1.1. Selective Acknowledgments (SACK)

If a device with less severe memory and processing constraints can afford advertising a TCP window size of several MSS, it makes sense to support the SACK option to improve performance. SACK allows a data receiver to inform the data sender of non-contiguous data blocks received, thus a sender (having previously sent the SACK-Permitted option) can avoid performing unnecessary retransmissions, saving energy and bandwidth, as well as reducing latency. In addition, SACK often allows for faster loss recovery when there is more than one lost segment in a window of data, since SACK recovery may complete with less RTTs. SACK is particularly useful for bulk data transfers. A receiver supporting SACK will need to keep track of the data blocks that need to be received. The sender will also need to keep track of which data segments need to be resent after learning which data blocks are missing at the receiver. SACK adds  $8 \cdot n + 2$  bytes to the TCP header, where  $n$  denotes the number of data blocks received, up to 4 blocks. For a low number of out-of-order segments, the header overhead penalty of SACK is compensated by avoiding unnecessary retransmissions. When the sender discovers the data blocks that have already been received, it needs to also store the necessary state to avoid unnecessary retransmission of data segments that have already been received.

#### 3.3.2. Delayed Acknowledgments

For certain traffic patterns, Delayed ACKs may have a detrimental effect, as already noted in Section 3.2.3. Advanced TCP stacks may use heuristics to determine the maximum delay for an ACK. For CNNs, the recommendation depends on the expected communication patterns.

When traffic over a CNN is expected to mostly be unidirectional messages with a size typically up to one MSS, and the time between two consecutive message transmissions is greater than the Delayed ACK timeout, it may make sense to use a smaller timeout or disable Delayed ACKs at the receiver. This avoids incurring additional delay, as well as the energy consumption of the sender (which might e.g. keep its radio interface in receive mode) during that time. Note that disabling Delayed ACKs may only be possible if the peer device is administered by the same entity managing the constrained device. For request-response traffic, enabling Delayed ACKs is recommended at the server end, in order to allow combining a response with the ACK into a single segment, thus increasing efficiency. In addition, if a client issues requests infrequently, disabling Delayed ACKs at the client allows an immediate ACK for the data segment carrying the response.

In contrast, Delayed ACKs allow to reduce the number of ACKs in bulk transfer type of traffic, e.g. for firmware/software updates or for transferring larger data units containing a batch of sensor readings.

Note that, in many scenarios, the peer that a constrained device communicates with will be a general purpose system that communicates with both constrained and unconstrained devices. Since delayed ACKs are often configured through system-wide parameters, delayed ACKs behavior at the peer will be the same regardless of the nature of the endpoints it talks to. Such a peer will typically have delayed ACKs enabled.

### 3.3.3. Initial Window

RFC 5681 specifies a TCP Initial Window (IW) of roughly 4 kB [RFC5681]. Subsequently, RFC 6928 defined an experimental new value for the IW, which in practice will result in an IW of 10 MSS [RFC6928]. The latter is nowadays used in many TCP implementations.

Note that a 10-MSS IW was recommended for resource-rich environments (e.g. broadband environments), which are significantly different from CNNs. In CNNs, many application layer data units are relatively small (e.g. below one MSS). However, larger objects (e.g. large files containing sensor readings, firmware updates, etc.) may also need to be transferred in CNNs. If such a large object is transferred in CNNs, with an IW setting of 10 MSS, there is significant buffer overflow risk, since many CNN devices support network or radio buffers of a size smaller than 10 MSS. In order to avoid such problem, in CNNs the IW needs to be carefully set, based on device and network resource constraints. In many cases, a safe IW setting will be smaller than 10 MSS.

## 4. TCP usage recommendations in CNNs

This section discusses how TCP can be used by applications that are developed for CNN scenarios. These remarks are by and large independent of how TCP is exactly implemented.

### 4.1. TCP connection initiation

In the constrained device to unconstrained device scenario illustrated above, a TCP connection is typically initiated by the constrained device, in order for this device to support possible sleep periods to save energy.

#### 4.2. Number of concurrent connections

TCP endpoints with a small amount of memory may only support a small number of connections. Each TCP connection requires storing a number of variables in the TCB. Depending on the internal TCP implementation, each connection may result in further memory overhead, and connections may compete for scarce resources (e.g. further memory overhead for send and receive buffers, etc).

A careful application design may try to keep the number of concurrent connections as small as possible. A client can for instance limit the number of simultaneous open connections that it maintains to a given server. Multiple connections could for instance be used to avoid the "head-of-line blocking" problem in an application transfer. However, in addition to consuming resources, using multiple connections can also cause undesirable side effects in congested networks. For example, the HTTP/1.1 specification encourages clients to be conservative when opening multiple connections [RFC7230]. Furthermore, each new connection will start with a 3-way handshake, therefore increasing message overhead.

Being conservative when opening multiple TCP connections is of particular importance in Constrained-Node Networks.

#### 4.3. TCP connection lifetime

In order to minimize message overhead, it makes sense to keep a TCP connection open as long as the two TCP endpoints have more data to send. If applications exchange data rather infrequently, i.e., if TCP connections would stay idle for a long time, the idle time can result in problems. For instance, certain middleboxes such as firewalls or NAT devices are known to delete state records after an inactivity interval. RFC 5382 specifies a minimum value for such interval of 124 minutes. Measurement studies have reported that TCP NAT binding timeouts are highly variable across devices, with a median around 60 minutes, the shortest timeout being around 2 minutes, and more than 50% of the devices with a timeout shorter than the aforementioned minimum timeout of 124 minutes [HomeGateway]. The timeout duration used by a middlebox implementation may not be known to the TCP endpoints.

In CNNs, such middleboxes may e.g. be present at the boundary between the CNN and other networks. If the middlebox can be optimized for CNN use cases, it makes sense to increase the initial value for filter state inactivity timers to avoid problems with idle connections. Apart from that, this problem can be dealt with by different connection handling strategies, each having pros and cons.

One approach for infrequent data transfer is to use short-lived TCP connections. Instead of trying to maintain a TCP connection for a long time, possibly short-lived connections can be opened between two endpoints, which are closed if no more data needs to be exchanged. For use cases that can cope with the additional messages and the latency resulting from starting new connections, it is recommended to use a sequence of short-lived connections, instead of maintaining a single long-lived connection.

The message and latency overhead that stems from using a sequence of short-lived connections could be reduced by TCP Fast Open (TFO) [RFC7413], which is an experimental TCP extension, at the expense of increased implementation complexity and increased TCP Control Block (TCB) size. TFO allows data to be carried in SYN (and SYN-ACK) segments, and to be consumed immediately by the receiving endpoint. This reduces the message and latency overhead compared to the traditional three-way handshake to establish a TCP connection. For security reasons, the connection initiator has to request a TFO cookie from the other endpoint. The cookie, with a size of 4 or 16 bytes, is then included in SYN packets of subsequent connections. The cookie needs to be refreshed (and obtained by the client) after a certain amount of time. While a given cookie is used for multiple connections between the same two endpoints, the latter may become vulnerable to privacy threats. In addition, a valid cookie may be stolen from a compromised host and may be used to perform SYN flood attacks, as well as amplified reflection attacks to victim hosts (see Section 5 of RFC 7413). Nevertheless, TFO is more efficient than frequently opening new TCP connections with the traditional three-way handshake, as long as the cookie can be reused in subsequent connections. However, as stated in RFC 7413, TFO deviates from the standard TCP semantics, since the data in the SYN could be replayed to an application in some rare circumstances. Applications should not use TFO unless they can tolerate this issue, e.g., by using Transport Layer Security (TLS) [RFC7413]. A comprehensive discussion on TFO can be found at RFC 7413.

Another approach is to use long-lived TCP connections with application-layer heartbeat messages. Various application protocols support such heartbeat messages (e.g. CoAP over TCP [RFC8323]). Periodic application-layer heartbeats can prevent early filter state record deletion in middleboxes. If the TCP binding timeout for a middlebox to be traversed by a given connection is known, middlebox filter state deletion will be avoided if the heartbeat period is lower than the middlebox TCP binding timeout. Otherwise, the implementer needs to take into account that middlebox TCP binding timeouts fall in a wide range of possible values [HomeGateway], and it may be hard to find a proper heartbeat period for application-layer heartbeat messages.



One specific advantage of Heartbeat messages is that they also allow aliveness checks at the application level. In general, it makes sense to realize aliveness checks at the highest protocol layer possible that is meaningful to the application, in order to maximize the depth of the aliveness check. In addition, timely detection of a dead peer may allow savings in terms of TCB memory use. However, the transmission of heartbeat messages consumes resources. This aspect needs to be assessed carefully, considering the characteristics of each specific CNN.

A TCP implementation may also be able to send "keep-alive" segments to test a TCP connection. According to [RFC1122], "keep-alives" are an optional TCP mechanism that is turned off by default, i.e., an application must explicitly enable it for a TCP connection. The interval between "keep-alive" messages must be configurable and it must default to no less than two hours. With this large timeout, TCP keep-alive messages might not always be useful to avoid deletion of filter state records in some middleboxes. However, sending TCP keep-alive probes more frequently risks draining power on energy-constrained devices.

## 5. Security Considerations

Best current practice for securing TCP and TCP-based communication also applies to CNN. As example, use of Transport Layer Security (TLS) [RFC8446] is strongly recommended if it is applicable. However, note that TLS protects only the contents of the data segments.

There are TCP options which can actually protect the transport layer. One example is the TCP Authentication Option (TCP-AO) [RFC5925]. However, this option adds overhead and complexity. TCP-AO typically has a size of 16-20 bytes. An implementer needs to assess the trade-off between security and performance when using TCP-AO, considering the characteristics (in terms of energy, bandwidth and computational power) of the environment where TCP will be used.

For the mechanisms discussed in this document, the corresponding considerations apply. For instance, if TFO is used, the security considerations of [RFC7413] apply.

Constrained devices are expected to support smaller TCP window sizes than less limited devices. In such conditions, segment retransmission triggered by RTO expiration is expected to be relatively frequent, due to lack of (enough) duplicate ACKs, especially when a constrained device uses a single-MSS implementation. For this reason, constrained devices running TCP may appear as particularly appealing victims of the so-called "shrew"

Denial of Service (DoS) attack [shrew], whereby one or more sources generate a packet spike targeted to coincide with consecutive RTO-expiration-triggered retry attempts of a victim node. Note that the attack may be performed by Internet-connected devices, including constrained devices in the same CNN as the victim, as well as remote ones. Mitigation techniques include RTO randomization and attack blocking by routers able to detect shrew attacks based on their traffic pattern.

## 6. Acknowledgments

Carles Gomez has been funded in part by the Spanish Government (Ministerio de Educacion, Cultura y Deporte) through the Jose Castillejo grants CAS15/00336 and CAS18/00170, and by European Regional Development Fund (ERDF) and the Spanish Government through projects TEC2016-79988-P, PID2019-106808RA-I00, AEI/FEDER, UE, and by Generalitat de Catalunya Grant 2017 SGR 376. Part of his contribution to this work has been carried out during his stays as a visiting scholar at the Computer Laboratory of the University of Cambridge.

The authors appreciate the feedback received for this document. The following folks provided comments that helped improve the document: Carsten Bormann, Zhen Cao, Wei Genyu, Ari Keranen, Abhijan Bhattacharyya, Andres Arcia-Moret, Yoshifumi Nishida, Joe Touch, Fred Baker, Nik Sultana, Kerry Lynn, Erik Nordmark, Markku Kojo, Hannes Tschofenig, David Black, Yoshifumi Nishida, Ilpo Jarvinen, Emmanuel Baccelli, Stuart Cheshire, Gorrry Fairhurst, Ingemar Johansson, Ted Lemon, and Michael Tuexen. Simon Brummer provided details, and kindly performed RAM and ROM usage measurements, on the RIOT TCP implementation. Xavi Vilajosana provided details on the OpenWSN TCP implementation. Rahul Jadhav kindly performed code size measurements on the Contiki-NG and lwIP 2.1.2 TCP implementations. He also provided details on the uIP TCP implementation.

## 7. Annex. TCP implementations for constrained devices

This section overviews the main features of TCP implementations for constrained devices. The survey is limited to open source stacks with small footprint. It is not meant to be all-encompassing. For more powerful embedded systems (e.g., with 32-bit processors), there are further stacks that comprehensively implement TCP. On the other hand, please be aware that this Annex is based on information available as of the writing.

### 7.1. uIP

uIP is a TCP/IP stack, targetted for 8 and 16-bit microcontrollers, which pioneered TCP/IP implementations for constrained devices. uIP has been deployed with Contiki and the Arduino Ethernet shield. A code size of ~5 kB (which comprises checksumming, IPv4, ICMP and TCP) has been reported for uIP [Dunk]. Later versions of uIP implement IPv6 as well.

uIP uses the same global buffer for both incoming and outgoing traffic, which has a size of a single packet. In case of a retransmission, an application must be able to reproduce the same user data that had been transmitted. Multiple connections are supported, but need to share the global buffer.

The MSS is announced via the MSS option on connection establishment and the receive window size (of one MSS) is not modified during a connection. Stop-and-wait operation is used for sending data. Among other optimizations, this allows to avoid sliding window operations, which use 32-bit arithmetic extensively and are expensive on 8-bit CPUs.

Contiki uses the "split hack" technique (see Section 3.2.3) to avoid Delayed ACKs for senders using a single segment.

The code size of the TCP implementation in Contiki-NG has been measured to be of 3.2 kB on CC2538DK, cross-compiling on Linux.

### 7.2. lwIP

lwIP is a TCP/IP stack, targetted for 8- and 16-bit microcontrollers. lwIP has a total code size of ~14 kB to ~22 kB (which comprises memory management, checksumming, network interfaces, IPv4, ICMP and TCP), and a TCP code size of ~9 kB to ~14 kB [Dunk]. Both IPv4 and IPv6 are supported in lwIP since v2.0.0.

In contrast with uIP, lwIP decouples applications from the network stack. lwIP supports a TCP transmission window greater than a single segment, as well as buffering of incoming and outgoing data. Other implemented mechanisms comprise slow start, congestion avoidance, fast retransmit and fast recovery. SACK and Window Scale support has been recently added to lwIP.

### 7.3. RIOT

The RIOT TCP implementation (called GNRC TCP) has been designed for Class 1 devices [RFC 7228]. The main target platforms are 8- and 16-bit microcontrollers, with 32-bit platforms also supported. GNRC

TCP offers a similar function set as uIP, but it provides and maintains an independent receive buffer for each connection. In contrast to uIP, retransmission is also handled by GNRC TCP. For simplicity, GNRC TCP uses a single-MSS implementation. The application programmer does not need to know anything about the TCP internals, therefore GNRC TCP can be seen as a user-friendly uIP TCP implementation.

The MSS is set on connections establishment and cannot be changed during connection lifetime. GNRC TCP allows multiple connections in parallel, but each TCB must be allocated somewhere in the system. By default there is only enough memory allocated for a single TCP connection, but it can be increased at compile time if the user needs multiple parallel connections.

The RIOT TCP implementation offers an optional POSIX socket wrapper that enables POSIX compliance, if needed.

Further details on RIOT and GNRC can be found in the literature [RIOT], [GNRC].

#### 7.4. TinyOS

TinyOS was important as a platform for early constrained devices. TinyOS has an experimental TCP stack that uses a simple nonblocking library-based implementation of TCP, which provides a subset of the socket interface primitives. The application is responsible for buffering. The TCP library does not do any receive-side buffering. Instead, it will immediately dispatch new, in-order data to the application and otherwise drop the segment. A send buffer is provided by the application. Multiple TCP connections are possible. Recently there has been little further work on the stack.

#### 7.5. FreeRTOS

FreeRTOS is a real-time operating system kernel for embedded devices that is supported by 16- and 32-bit microprocessors. Its TCP implementation is based on multiple-segment window size, although a 'Tiny-TCP' option, which is a single-MSS variant, can be enabled. Delayed ACKs are supported, with a 20-ms Delayed ACK timer as a technique intended 'to gain performance'.

#### 7.6. uC/OS

uC/OS is a real-time operating system kernel for embedded devices, which is maintained by Micrium. uC/OS is intended for 8-, 16- and 32-bit microprocessors. The uC/OS TCP implementation supports a multiple-segment window size.

## 7.7. Summary

		uIP	lwIP orig	lwIP 2.1	RIOT	TinyOS	FreeRTOS	uC/OS
Memory	Code size (kB)	<5 (a)	~9 to ~14 (T1)	38 (T4)	<7 (T3)	N/A	<9.2 (T2)	N/A
TCP features	Single-Segm.	Yes	No	No	Yes	No	No	No
	Slow start	No	Yes	Yes	No	Yes	No	Yes
	Fast rec/retr	No	Yes	Yes	No	Yes	No	Yes
	Keep-alive	No	No	Yes	No	No	Yes	Yes
	Win. Scale	No	No	Yes	No	No	Yes	No
	TCP timest.	No	No	Yes	No	No	Yes	No
	SACK	No	No	Yes	No	No	Yes	No
	Del. ACKs	No	Yes	Yes	No	No	Yes	Yes
	Socket	No	No	Optional	(I)	Subset	Yes	Yes
	Concur. Conn.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
TLS supported		No	No	Yes	Yes	Yes	Yes	Yes

(T1) = TCP-only, on x86 and AVR platforms

(T2) = TCP-only, on ARM Cortex-M platform

(T3) = TCP-only, on ARM Cortex-M0+ platform (NOTE: RAM usage for the same platform

is ~2.5 kB for one TCP connection plus ~1.2 kB for each additional connection)

(T4) = TCP-only, on CC2538DK, cross-compiling on Linux

(a) = includes IP, ICMP and TCP on x86 and AVR platforms. The Contiki-NG TCP implementation has a code size of 3.2 kB on CC2538DK, cross-compiling on Linux

(I) = optional POSIX socket wrapper which enables POSIX compliance if needed

Mult. = Multiple

N/A = Not Available

Figure 2: Summary of TCP features for different lightweight TCP implementations. None of the implementations considered in this Annex support ECN or TFO.

## 8. Annex. Changes compared to previous versions

RFC Editor: To be removed prior to publication

### 8.1. Changes between -00 and -01

- o Changed title and abstract
- o Clarification that communication with standard-compliant TCP endpoints is required, based on feedback from Joe Touch
- o Additional discussion on communication patterns
- o Numerous changes to address a comprehensive review from Hannes Tschofenig
- o Reworded security considerations
- o Additional references and better distinction between normative and informative entries
- o Feedback from Rahul Jadhav on the uIP TCP implementation
- o Basic data for the TinyOS TCP implementation added, based on source code analysis

### 8.2. Changes between -01 and -02

- o Added text to the Introduction section, and a reference, on traditional bad perception of TCP for IoT
- o Added sections on FreeRTOS and uC/OS
- o Updated TinyOS section
- o Updated summary table
- o Reorganized Section 4 (single-MSS vs multiple-MSS window size), some content now also in new Section 5

### 8.3. Changes between -02 and -03

- o Rewording to better explain the benefit of ECN
- o Additional context information on the surveyed implementations
- o Added details, but removed "Data size" row, in the summary table

- o Added discussion on shrew attacks
- 8.4. Changes between -03 and -04
- o Addressing the remaining TODOs
  - o Alignment of the wording on TCP "keep-alives" with related discussions in the IETF transport area
  - o Added further discussion on delayed ACKs
  - o Removed OpenWSN section from the Annex
- 8.5. Changes between -04 and -05
- o Addressing comments by Yoshifumi Nishida
  - o Removed mentioning MD5 as an example (comment by David Black)
  - o Added memory footprint details of TCP implementations (Contiki-NG and lwIP 2.1.2) provided by Rahul Jadhav in the Annex
  - o Addressed comments by Ilpo Jarvinen throughout the whole document
  - o Improved the RIOT section in the Annex, based on feedback from Emmanuel Baccelli
- 8.6. Changes between -05 and -06
- o Incorporated suggestions by Stuart Cheshire
- 8.7. Changes between -06 and -07
- o Addressed comments by Gorrry Fairhurst
- 8.8. Changes between -07 and -08
- o Addressed WGLC comments by Ilpo Jarvinen, Markku Kojo and Ingemar Johansson throughout the document, including the addition of a new section on Initial Window considerations.
- 8.9. Changes between -08 and -09
- o Addressed second round of comments by Ilpo Jarvinen and Markku Kojo, based on the previous draft update.

## 8.10. Changes between -09 and -10

- o Addressed comments by Erik Kline.
- o Addressed a comment by Markku Kojo on advice given in RFC 6691.

## 8.11. Changes between -10 and -11

- o Addressed a comment by Ted Lemon on MSS advice.

## 8.12. Changes between -11 and -12

- o Addressed comments from IESG and various directorates.

## 8.13. Changes between -12 and -13

- o Fixed two typos.
- o Addressed a comment by Barry Leiba.

## 9. References

## 9.1. Normative References

- [RFC0793] Postel, J., "Transmission Control Protocol", STD 7, RFC 793, DOI 10.17487/RFC0793, September 1981, <<https://www.rfc-editor.org/info/rfc793>>.
- [RFC1122] Braden, R., Ed., "Requirements for Internet Hosts - Communication Layers", STD 3, RFC 1122, DOI 10.17487/RFC1122, October 1989, <<https://www.rfc-editor.org/info/rfc1122>>.
- [RFC2018] Mathis, M., Mahdavi, J., Floyd, S., and A. Romanow, "TCP Selective Acknowledgment Options", RFC 2018, DOI 10.17487/RFC2018, October 1996, <<https://www.rfc-editor.org/info/rfc2018>>.
- [RFC3042] Allman, M., Balakrishnan, H., and S. Floyd, "Enhancing TCP's Loss Recovery Using Limited Transmit", RFC 3042, DOI 10.17487/RFC3042, January 2001, <<https://www.rfc-editor.org/info/rfc3042>>.
- [RFC3168] Ramakrishnan, K., Floyd, S., and D. Black, "The Addition of Explicit Congestion Notification (ECN) to IP", RFC 3168, DOI 10.17487/RFC3168, September 2001, <<https://www.rfc-editor.org/info/rfc3168>>.



- [RFC5681] Allman, M., Paxson, V., and E. Blanton, "TCP Congestion Control", RFC 5681, DOI 10.17487/RFC5681, September 2009, <<https://www.rfc-editor.org/info/rfc5681>>.
- [RFC6298] Paxson, V., Allman, M., Chu, J., and M. Sargent, "Computing TCP's Retransmission Timer", RFC 6298, DOI 10.17487/RFC6298, June 2011, <<https://www.rfc-editor.org/info/rfc6298>>.
- [RFC6691] Borman, D., "TCP Options and Maximum Segment Size (MSS)", RFC 6691, DOI 10.17487/RFC6691, July 2012, <<https://www.rfc-editor.org/info/rfc6691>>.
- [RFC6928] Chu, J., Dukkkipati, N., Cheng, Y., and M. Mathis, "Increasing TCP's Initial Window", RFC 6928, DOI 10.17487/RFC6928, April 2013, <<https://www.rfc-editor.org/info/rfc6928>>.
- [RFC7228] Bormann, C., Ersue, M., and A. Keranen, "Terminology for Constrained-Node Networks", RFC 7228, DOI 10.17487/RFC7228, May 2014, <<https://www.rfc-editor.org/info/rfc7228>>.
- [RFC7323] Borman, D., Braden, B., Jacobson, V., and R. Scheffenegger, Ed., "TCP Extensions for High Performance", RFC 7323, DOI 10.17487/RFC7323, September 2014, <<https://www.rfc-editor.org/info/rfc7323>>.
- [RFC7413] Cheng, Y., Chu, J., Radhakrishnan, S., and A. Jain, "TCP Fast Open", RFC 7413, DOI 10.17487/RFC7413, December 2014, <<https://www.rfc-editor.org/info/rfc7413>>.
- [RFC7567] Baker, F., Ed. and G. Fairhurst, Ed., "IETF Recommendations Regarding Active Queue Management", BCP 197, RFC 7567, DOI 10.17487/RFC7567, July 2015, <<https://www.rfc-editor.org/info/rfc7567>>.
- [RFC8200] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", STD 86, RFC 8200, DOI 10.17487/RFC8200, July 2017, <<https://www.rfc-editor.org/info/rfc8200>>.

## 9.2. Informative References

- [Commag] A. Betzler, C. Gomez, I. Demirkol, J. Paradells, "CoAP Congestion Control for the Internet of Things", IEEE Communications Magazine, June 2016.

- [Dunk] A. Dunkels, "Full TCP/IP for 8-Bit Architectures", 2003.
- [ETEN] R. Krishnan et al, "Explicit transport error notification (ETEN) for error-prone wireless and satellite networks", Computer Networks 2004.
- [GNRC] M. Lenders et al., "Connecting the World of Embedded Mobiles: The RIOT Approach to Ubiquitous Networking for the IoT", 2018.
- [HomeGateway] Haetoeinen, S., Nyrhinen, A., Eggert, L., Strowes, S., Sarolahti, P., and M. Kojo, "An Experimental Study of Home Gateway Characteristics", Proceedings of the 10th ACM SIGCOMM conference on Internet measurement 2010.
- [I-D.delcarpio-6lo-wlanah] Vega, L., Robles, I., and R. Morabito, "IPv6 over 802.11ah", draft-delcarpio-6lo-wlanah-01 (work in progress), October 2015.
- [I-D.ietf-6lo-fragment-recovery] Thubert, P., "6LoWPAN Selective Fragment Recovery", draft-ietf-6lo-fragment-recovery-21 (work in progress), March 2020.
- [I-D.ietf-core-fasor] Jarvinen, I., Kojo, M., Raitahila, I., and Z. Cao, "Fast-Slow Retransmission Timeout and Congestion Control Algorithm for CoAP", draft-ietf-core-fasor-01 (work in progress), October 2020.
- [I-D.ietf-tcpm-generalized-ecn] Bagnulo, M. and B. Briscoe, "ECN++: Adding Explicit Congestion Notification (ECN) to TCP Control Packets", draft-ietf-tcpm-generalized-ecn-05 (work in progress), November 2019.
- [I-D.ietf-tcpm-rto-consider] Allman, M., "Requirements for Time-Based Loss Detection", draft-ietf-tcpm-rto-consider-17 (work in progress), July 2020.
- [IntComp] C. Gomez, A. Arcia-Moret, J. Crowcroft, "TCP in the Internet of Things: from ostracism to prominence", IEEE Internet Computing, January-February 2018.

- [MQTT] ISO/IEC 20922:2016, "Message Queuing Telemetry Transport (MQTT) v3.1.1", 2016.
- [RFC2757] Montenegro, G., Dawkins, S., Kojo, M., Magret, V., and N. Vaidya, "Long Thin Networks", RFC 2757, DOI 10.17487/RFC2757, January 2000, <<https://www.rfc-editor.org/info/rfc2757>>.
- [RFC2884] Hadi Salim, J. and U. Ahmed, "Performance Evaluation of Explicit Congestion Notification (ECN) in IP Networks", RFC 2884, DOI 10.17487/RFC2884, July 2000, <<https://www.rfc-editor.org/info/rfc2884>>.
- [RFC3481] Inamura, H., Ed., Montenegro, G., Ed., Ludwig, R., Gurtov, A., and F. Khafizov, "TCP over Second (2.5G) and Third (3G) Generation Wireless Networks", BCP 71, RFC 3481, DOI 10.17487/RFC3481, February 2003, <<https://www.rfc-editor.org/info/rfc3481>>.
- [RFC3819] Karn, P., Ed., Bormann, C., Fairhurst, G., Grossman, D., Ludwig, R., Mahdavi, J., Montenegro, G., Touch, J., and L. Wood, "Advice for Internet Subnetwork Designers", BCP 89, RFC 3819, DOI 10.17487/RFC3819, July 2004, <<https://www.rfc-editor.org/info/rfc3819>>.
- [RFC4944] Montenegro, G., Kushalnagar, N., Hui, J., and D. Culler, "Transmission of IPv6 Packets over IEEE 802.15.4 Networks", RFC 4944, DOI 10.17487/RFC4944, September 2007, <<https://www.rfc-editor.org/info/rfc4944>>.
- [RFC5925] Touch, J., Mankin, A., and R. Bonica, "The TCP Authentication Option", RFC 5925, DOI 10.17487/RFC5925, June 2010, <<https://www.rfc-editor.org/info/rfc5925>>.
- [RFC6077] Papadimitriou, D., Ed., Welzl, M., Scharf, M., and B. Briscoe, "Open Research Issues in Internet Congestion Control", RFC 6077, DOI 10.17487/RFC6077, February 2011, <<https://www.rfc-editor.org/info/rfc6077>>.
- [RFC6120] Saint-Andre, P., "Extensible Messaging and Presence Protocol (XMPP): Core", RFC 6120, DOI 10.17487/RFC6120, March 2011, <<https://www.rfc-editor.org/info/rfc6120>>.
- [RFC6282] Hui, J., Ed. and P. Thubert, "Compression Format for IPv6 Datagrams over IEEE 802.15.4-Based Networks", RFC 6282, DOI 10.17487/RFC6282, September 2011, <<https://www.rfc-editor.org/info/rfc6282>>.

- [RFC6550] Winter, T., Ed., Thubert, P., Ed., Brandt, A., Hui, J., Kelsey, R., Levis, P., Pister, K., Struik, R., Vasseur, JP., and R. Alexander, "RPL: IPv6 Routing Protocol for Low-Power and Lossy Networks", RFC 6550, DOI 10.17487/RFC6550, March 2012, <<https://www.rfc-editor.org/info/rfc6550>>.
- [RFC6606] Kim, E., Kaspar, D., Gomez, C., and C. Bormann, "Problem Statement and Requirements for IPv6 over Low-Power Wireless Personal Area Network (6LoWPAN) Routing", RFC 6606, DOI 10.17487/RFC6606, May 2012, <<https://www.rfc-editor.org/info/rfc6606>>.
- [RFC6775] Shelby, Z., Ed., Chakrabarti, S., Nordmark, E., and C. Bormann, "Neighbor Discovery Optimization for IPv6 over Low-Power Wireless Personal Area Networks (6LoWPANs)", RFC 6775, DOI 10.17487/RFC6775, November 2012, <<https://www.rfc-editor.org/info/rfc6775>>.
- [RFC7230] Fielding, R., Ed. and J. Reschke, Ed., "Hypertext Transfer Protocol (HTTP/1.1): Message Syntax and Routing", RFC 7230, DOI 10.17487/RFC7230, June 2014, <<https://www.rfc-editor.org/info/rfc7230>>.
- [RFC7252] Shelby, Z., Hartke, K., and C. Bormann, "The Constrained Application Protocol (CoAP)", RFC 7252, DOI 10.17487/RFC7252, June 2014, <<https://www.rfc-editor.org/info/rfc7252>>.
- [RFC7414] Duke, M., Braden, R., Eddy, W., Blanton, E., and A. Zimmermann, "A Roadmap for Transmission Control Protocol (TCP) Specification Documents", RFC 7414, DOI 10.17487/RFC7414, February 2015, <<https://www.rfc-editor.org/info/rfc7414>>.
- [RFC7428] Brandt, A. and J. Buron, "Transmission of IPv6 Packets over ITU-T G.9959 Networks", RFC 7428, DOI 10.17487/RFC7428, February 2015, <<https://www.rfc-editor.org/info/rfc7428>>.
- [RFC7540] Belshe, M., Peon, R., and M. Thomson, Ed., "Hypertext Transfer Protocol Version 2 (HTTP/2)", RFC 7540, DOI 10.17487/RFC7540, May 2015, <<https://www.rfc-editor.org/info/rfc7540>>.

- [RFC7668] Nieminen, J., Savolainen, T., Isomaki, M., Patil, B., Shelby, Z., and C. Gomez, "IPv6 over BLUETOOTH(R) Low Energy", RFC 7668, DOI 10.17487/RFC7668, October 2015, <<https://www.rfc-editor.org/info/rfc7668>>.
- [RFC8087] Fairhurst, G. and M. Welzl, "The Benefits of Using Explicit Congestion Notification (ECN)", RFC 8087, DOI 10.17487/RFC8087, March 2017, <<https://www.rfc-editor.org/info/rfc8087>>.
- [RFC8105] Mariager, P., Petersen, J., Ed., Shelby, Z., Van de Logt, M., and D. Barthel, "Transmission of IPv6 Packets over Digital Enhanced Cordless Telecommunications (DECT) Ultra Low Energy (ULE)", RFC 8105, DOI 10.17487/RFC8105, May 2017, <<https://www.rfc-editor.org/info/rfc8105>>.
- [RFC8163] Lynn, K., Ed., Martocci, J., Neilson, C., and S. Donaldson, "Transmission of IPv6 over Master-Slave/Token-Passing (MS/TP) Networks", RFC 8163, DOI 10.17487/RFC8163, May 2017, <<https://www.rfc-editor.org/info/rfc8163>>.
- [RFC8201] McCann, J., Deering, S., Mogul, J., and R. Hinden, Ed., "Path MTU Discovery for IP version 6", STD 87, RFC 8201, DOI 10.17487/RFC8201, July 2017, <<https://www.rfc-editor.org/info/rfc8201>>.
- [RFC8323] Bormann, C., Lemay, S., Tschofenig, H., Hartke, K., Silverajan, B., and B. Raymor, Ed., "CoAP (Constrained Application Protocol) over TCP, TLS, and WebSockets", RFC 8323, DOI 10.17487/RFC8323, February 2018, <<https://www.rfc-editor.org/info/rfc8323>>.
- [RFC8352] Gomez, C., Kovatsch, M., Tian, H., and Z. Cao, Ed., "Energy-Efficient Features of Internet of Things Protocols", RFC 8352, DOI 10.17487/RFC8352, April 2018, <<https://www.rfc-editor.org/info/rfc8352>>.
- [RFC8376] Farrell, S., Ed., "Low-Power Wide Area Network (LPWAN) Overview", RFC 8376, DOI 10.17487/RFC8376, May 2018, <<https://www.rfc-editor.org/info/rfc8376>>.
- [RFC8446] Rescorla, E., "The Transport Layer Security (TLS) Protocol Version 1.3", RFC 8446, DOI 10.17487/RFC8446, August 2018, <<https://www.rfc-editor.org/info/rfc8446>>.

- [RFC8900] Bonica, R., Baker, F., Huston, G., Hinden, R., Troan, O., and F. Gont, "IP Fragmentation Considered Fragile", BCP 230, RFC 8900, DOI 10.17487/RFC8900, September 2020, <<https://www.rfc-editor.org/info/rfc8900>>.
- [RIOT] E. Baccelli et al., "RIOT: an Open Source Operating System for Low-end Embedded Devices in the IoT", 2018.
- [shrew] A. Kuzmanovic, E. Knightly, "Low-Rate TCP-Targeted Denial of Service Attacks", SIGCOMM'03 2003.

## Authors' Addresses

Carles Gomez  
UPC  
C/Esteve Terradas, 7  
Castelldefels 08860  
Spain  
  
Email: carlesgo@entel.upc.edu

Jon Crowcroft  
University of Cambridge  
JJ Thomson Avenue  
Cambridge, CB3 0FD  
United Kingdom  
  
Email: jon.crowcroft@cl.cam.ac.uk

Michael Scharf  
Hochschule Esslingen  
Flandernstr. 101  
Esslingen 73732  
Germany  
  
Email: michael.scharf@hs-esslingen.de