Network Working Group   A. Keranen
Internet-Draft   Ericsson
Intended status: Standards Track   C. Bormann
Expires: May 7, 2020   Universitaet Bremen TZI
                                                  November 4, 2019

SenML Data Value Content-Format Indication
draft-ietf-core-senml-data-ct-01

Abstract

   The Sensor Measurement Lists (SenML) media type supports multiple
   types of values, from numbers to text strings and arbitrary binary
   data values. In order to simplify processing of the data values this
   document proposes to specify a new SenML field for indicating the
   Content-Format of the data.

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 7, 2020.

Copyright Notice

the Trust Legal Provisions and are provided without warranty as
described in the Simplified BSD License.

Table of Contents

1. Introduction

The Sensor Measurement Lists (SenML) media type [RFC8428] can be used
to send various different kinds of data. In the example given in
Figure 1, a temperature value, an indication whether a lock is open,
and a data value (with SenML field "vd") read from an NFC reader is
sent in a single SenML pack.

```
[
  {"bn":"urn:dev:ow:10e2073a01080063:","n":"temp","u":"Cel","v":7.1},
  {"n":"open","vb":false},
  {"n":"nfc-reader","vd":"aGkgCg"}
]
```

          Figure 1: SenML pack with unidentified binary data

The receiver is expected to know how to interpret the data in the
"vd" field based on the context, e.g., name of the data source and
out-of-band knowledge of the application. However, this context may
not always be easily available to entities processing the SenML pack.
To facilitate automatic interpretation it is useful to be able to
indicate an Internet media type and content-coding right in the SenML
Record. The CoAP Content-Format (Section 12.3 in [RFC7252]) provides
just this information; enclosing a Content-Format number (in this
case number 60 as defined for content-type application/cbor in
[RFC7049]) in the Record is illustrated in Figure 2. All registered
CoAP Content-Formats are listed in the Content-Formats subregistry of
the CoRE Parameters registry [IANA.core-parameters].

```
{"n":"nfc-reader", "vd":"gmNmb28YKg", "ct":"60"}
```

        Figure 2: SenML Record with binary data identified as CBOR

   In this example SenML Record the data value contains a string "foo"
   and a number 42 encoded in a CBOR [RFC7049] array. Since the example
   above uses the JSON format of SenML, the data value containing the
   binary CBOR value is base64-encoded. The data value after base64
   decoding is shown with CBOR diagnostic notation in Figure 3.

```
82           # array(2)
   63        # text(3)
      666F6F # "foo"
   18 2A     # unsigned(42)
```

        Figure 3: Example Data Value in CBOR diagnostic notation

2. Terminology

   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.

   Readers should also be familiar with the terms and concepts discussed
   in [RFC8428]. Awareness of terminology issues discussed in
   [I-D.bormann-core-media-content-type-format] can also be very
   helpful.

3. SenML Content-Format ("ct") Field

   When a SenML Record contains a Data Value field ("vd"), the Record
   MAY also include a Content-Format indication field. The Content-
   Format indication uses label "ct" and a string value with either a
   CoAP Content-Format identifier in decimal form with no leading zeros
   except for the value "0" itself (representing an unsigned integer in
   the range of 0-65535, similar to the CoRE Link Format [RFC6690] "ct"
   attribute) or with a string containing a Content-Type and optionally
   a Content-Coding (see below).

   The CoAP Content-Format identifier provides a simple and efficient
   way to indicate the type of the data. Since some Internet media
   types and their content coding and parameter alternatives do not have
   assigned CoAP Content-Format identifiers, using Content-Type and
   Content-Coding is also allowed. Both methods use a string value in
   the "ct" field to keep its data type consistent across uses. When

the "ct" field contains only digits, it is interpreted as a CoAP
Content-Format identifier.

To indicate that a Content-Coding is used with a Content-Type, the
Content-Coding value (e.g., "deflate" [RFC7230]) is appended to the
Content Type (media type and parameters, if any), separated by a "@"
sign. For example: "text/plain; charset=utf-8@deflate". If Content-
Coding is not specified with a Content-Type (no "@" sign is present
outside any media type parameters), the identity (i.e., no)
transformation is used.

4. SenML Base Content-Format ("bct") Field

The Base Content-Format Field, label "bct", provides a default value
for the Content-Format Field (label "ct") within its range. The
range of the base field includes the Record containing it, up to (but
not including) the next Record containing a "bct" field, if any, or
up to the end of the pack otherwise. Resolution (Section 4.6 of
[RFC8428]) of this base field is performed by adding its value with
the label "ct" to all Records in this range that carry a "vd" field
but do not already contain a Content-Format ("ct") field.

5. Mandatory to Understand Content-Format

If the Content-Format field needs to be understood by all processors
of the SenML Pack, the mandatory to understand versions of the
fields, "ct_" and "bct_", can be used. These fields have identical
semantics to the "ct" and "bct" fields respectively except that a
SenML processor that does not support this specification would reject
a SenML Pack with such fields and generate an error (see Section 4.4
of [RFC8428]).

Using the regular Content-Format indication enables to use this
extension in a backward compatible way to indicate information that
is not critical to be understood. The choice between the two methods
is application dependent.

If both a "ct_" field and a "ct" field are present in a resolved
Record (i.e., from fields in the Record or from base fields), the
"ct_" field overrides the "ct" field. Using both "ct" and "ct_" in
the same Record is NOT RECOMMENDED as it MAY be treated as an error
by the recipient.

6. Examples

The following examples are valid values for the "ct" and "bct" fields
(explanation/comments in parenthesis):

   o "60" (CoAP Content-Format for "application/cbor")

   o "0" (CoAP Content-Format for "text/plain" with parameter
      "charset=utf-8")

   o "application/json" (JSON Content-Type - equivalent to "50" CoAP
      Content-Format identifier)

   o "application/json@deflate" (JSON Content-Type with "deflate" as
      Content-Coding - equivalent to "11050" CoAP Content-Format
      identifier)

   o "text/csv" (Comma-Separated Values (CSV) [RFC4180] Content-Type)

   o "text/csv@gzip" (CSV with "gzip" as Content-Coding)

7. Security Considerations

   The indication of a media type in the data does not exempt a
   consuming application from properly checking its inputs. Also, the
   ability for an attacker to supply crafted SenML data that specify
   media types chosen by the attacker may expose vulnerabilities of
   handlers for these media types to the attacker.

8. IANA Considerations

   (Note to RFC Editor: Please replace all occurrences of "RFC-AAAA"
   with the RFC number of this specification and remove this note.)

   IANA is requested to assign new labels in the "SenML Labels"
   subregistry of the SenML registry [IANA.senml] (as defined in
   [RFC8428]) for the Content-Format indication as per Table 1:

   +---------------------+-------+-----------+---------+-----------+
   | Name | Label | JSON Type | XML Type | Reference |
   +---------------------+-------+-----------+---------+-----------+
Base Content-Format	bct	String	string	RFC-AAAA
Content-Format	ct	String	string	RFC-AAAA
   +---------------------+-------+-----------+---------+-----------+

             Table 1: IANA Registration for new SenML Labels

Acknowledgements

   The authors would like to thank Sergio Abreu for the discussions
   leading to the design of this extension and Isaac Rivera for reviews
   and feedback.

10. References

10.1. Normative References

[IANA.senml]
          IANA, "Sensor Measurement Lists (SenML)",
          <http://www.iana.org/assignments/senml>.

[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>.

[RFC7049] Bormann, C. and P. Hoffman, "Concise Binary Object
          Representation (CBOR)", RFC 7049, DOI 10.17487/RFC7049,
          October 2013, <https://www.rfc-editor.org/info/rfc7049>.

[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>.

[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>.

[RFC8428] Jennings, C., Shelby, Z., Arkko, J., Keranen, A., and C.
          Bormann, "Sensor Measurement Lists (SenML)", RFC 8428,
          DOI 10.17487/RFC8428, August 2018,
          <https://www.rfc-editor.org/info/rfc8428>.

10.2. Informative References

[I-D.bormann-core-media-content-type-format]
          Bormann, C., "On Media-Types, Content-Types, and related
          terminology", draft-bormann-core-media-content-type-
          format-01 (work in progress), July 2019.

[IANA.core-parameters]
          IANA, "Constrained RESTful Environments (CoRE)
          Parameters",
          <http://www.iana.org/assignments/core-parameters>.

[RFC4180] Shafranovich, Y., "Common Format and MIME Type for Comma-
          Separated Values (CSV) Files", RFC 4180,
          DOI 10.17487/RFC4180, October 2005,
          <https://www.rfc-editor.org/info/rfc4180>.

   [RFC6690]   Shelby, Z., "Constrained RESTful Environments (CoRE) Link
               Format", RFC 6690, DOI 10.17487/RFC6690, August 2012,
               <https://www.rfc-editor.org/info/rfc6690>.

   [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>.

Authors' Addresses

   Ari Keranen
   Ericsson
   Jorvas 02420
   Finland

   Email: ari.keranen@ericsson.com

   Carsten Bormann
   Universitaet Bremen TZI
   Postfach 330440
   Bremen D-28359
   Germany

   Phone: +49-421-218-63921
   Email: cabo@tzi.org