

CBOR

Design and Overview

Paul Hoffman
Carsten Bormann

Why

- A new binary object format with different design goals from current formats
- Extremely **small code size** for both encoder and parser
 - Motivated by work on networks with constrained nodes
- Design objectives stated up front in the document
- Useful to any protocol or application that likes the design goals

What

- draft-bormann-cbor
- Concise Binary Object Representation: CBOR^{*)}
- Now in IETF Last Call, comments are certainly still welcome

*) Paul, trying for a pronounceable acronym, came up with this name without consciously realizing the obvious link to Carsten's name.



“Sea Boar”

Design goals (1 of 3)

1. The representation must be able to unambiguously encode most **common data formats** (such as JSON-like data) used in Internet standards.
2. The **code** for an encoder or parser must be able to be **compact** in order to support systems with very limited memory and processor power and instruction sets.

Design goals (2 of 3)

3. Data must be able to be parsed **without a schema description**.

4. The serialization must be reasonably compact, but **data compactness is secondary to code compactness** for the encoder and parser.

Design goals (3 of 3)

5. The format must be applicable to both **constrained nodes** and **high-volume applications**.

6. The format must support all **JSON** data types for conversion to and from JSON.

7. The format must be **extensible**, with the extended data being able to be parsed by earlier parsers.

Very quick overview of the format

- Initial byte contains both information about the **major type** (the high-order 3 bits) and **additional information** (the low-order 5 bits)
- Eight major types:
 - unsigned and negative **integers**
 - **byte** strings, **UTF-8** strings
 - **arrays**, **maps**
 - optional **tagging** and **simple types** (floating point, Booleans, etc.)

Semantics of additional information

- If the additional information is less than 24, it is directly used as a small unsigned integer
- If the additional information is 24 to 27, additional bytes immediately follow for a 1-, 2-, 4- or 8-byte unsigned integer

Some features

- IEEE floats in 16, 32, and 64 bits
- Can use indefinite lengths for streaming
- Optional semantic tagging for things like dates, arbitrary-length bignums, and decimal fractions

What's in the document

- Design objectives in the Introduction
- Complete description of the format
- Lots of worked-out examples, pseudo-code
- Discussion of using CBOR in protocols and applications
 - Allows them to define subsets of the format and rules for what is and is not allowed
- Rules for later expansion

History

- Lots of good feedback on appsawg mailing list
- Incorporated many suggestions
- Now: in IETF Last Call
- Later: we'll see who adopts it