

# Selecting a Data Format for an Endpoint Information Data Model

SACM Virtual Interim Meeting

06/15/2016

# Agenda

- Considerations (pertaining to data formats)
- Introductions
  - CBOR
  - JSON
  - XML
- Detailed Observations
- Next Steps

# Considerations

- Meets SACM Requirements <sup>1</sup>
- Supports SACM IM <sup>2</sup>
- Easy to document/understand examples?
- PROs and CONS with respect to SACM
  - Technical Features?
  - Compatibility Needs?
  - Others?

For example, some may complain that JSON doesn't do comments. But this isn't relevant unless SACM data formats need to have comments.

1. <https://datatracker.ietf.org/doc/draft-ietf-sacm-requirements/>  
2. <https://datatracker.ietf.org/doc/draft-ietf-sacm-information-model/>

# SACM Requirement Considerations

- SACM requirements which could affect choosing a data format
  - DM-003 Search Flexibility
  - DM-006 Data Cardinality
  - DM-016 Transport Agnostic
  - *There may be others...*
- Do we want a data model which supports relevant SHOULD requirements as well MUST requirements?
- Do we want to consider associated schema languages?

# Concise Binary Object Representation (CBOR)<sup>1</sup>

- CBOR is a compact, binary data format
- Potential PROs
  - Small footprint for encoding/decoding software
  - Data compactness
  - CBOR Data Definition Language (CDDL)<sup>2</sup> provides a way to express structures
- Potential CONS
  - Encoder/Decoder software availability?
  - Not human readable (i.e., more than a text editor is required)

1. <http://tools.ietf.org/html/rfc7049>

2. <https://datatracker.ietf.org/doc/draft-greevenbosch-appsawg-cbor-cddl/>

# JavaScript Object Notation (JSON)<sup>1</sup>

- JSON is a simple, text-based data interchange format
- Potential PROs
  - Simple
  - Rapid adoption taking place
- Potential CONs
  - Ambiguous data item ordering and name uniqueness (explained and addressed in I-JSON<sup>2</sup>)
  - Status of JSON Schema?

1. <https://tools.ietf.org/html/rfc7159>

2. <https://tools.ietf.org/html/rfc7493>

# Extensible Markup Language (XML)<sup>1</sup>

- XML is a text-based markup language for exchanging data
- Potential PROs
  - Established
  - Variety of standardized (and optional) capabilities
  - Namespaces
- Potential CONs
  - Verbose
  - Complex
  - Namespaces

1. <http://www.w3.org/XML/>

# Relationship Between CBOR and JSON

- Underlying data model is an extension of the JSON data model
  - All JSON types map directly to CBOR
  - Some CBOR types do not have an analog in JSON
- An objective of CBOR is to support all JSON data types for conversion to and from JSON
  - JSON to CBOR conversion seems straightforward. The major question is what binary number representation(s) to use for translated numeric values.
  - CBOR to JSON conversion is more complex. Non-normative guidelines are given in the CBOR spec
- Designing the data in JSON, and using CBOR as a compact, on-the-wire format, may be a useful strategy
  - CDDL may be leveraged to design the data structures we care about

# Constraints on JSON Objects

- JSON Data Interchange Format (ECMA-404)<sup>1</sup> does not specify:
  - Whether or not the order of object members is significant
  - Uniqueness requirements for object member names (i.e., are duplicate names allowed?)
- JSON Data Interchange Format (RFC 7149)<sup>2</sup>
  - “JSON parsing libraries have been observed to differ as to whether or not they make the ordering of object members visible to calling software.”
  - “When the names within an object are not unique, the behavior of software that receives such an object is unpredictable.”
- Internet JSON (I-JSON) Message Format (RFC 7493)<sup>3</sup>
  - “The order of object members in an I-JSON message does not change the meaning of an I-JSON message.”
  - Objects in I-JSON messages **MUST NOT** have members with duplicate names.”

1. [www.ecma-international.org/publications/files/ECMA-ST/ECMA-404.pdf](http://www.ecma-international.org/publications/files/ECMA-ST/ECMA-404.pdf)

2. <https://tools.ietf.org/html/rfc7159>

3. <https://tools.ietf.org/html/rfc7493>

# Impact of JSON Object Constraints on IM Data Models

- JSON Objects shouldn't be used to represent:
  - Ordered lists
  - Lists which use the same data item more than once
- An alternative is to employ JSON arrays to represent SACM lists
  - Order of array components is significant (spec is clear on that)
  - SACM data item names can be treated as data in JSON rather than as a data item

# Constraints on XML Content Models

- There is mixed/limited support for open content models
- A validating XML processor is supposed to fail when encountering an invalid XML instance
- DTD has no support for open content
- In XML Schema, all elements are considered closed unless special data-types or constructs are used (e.g., `xs:any`, `xs:opencontent`)

# Impact of XML Content Model on IM Data Models

- Since open content cannot be assumed, neither can extensibility of data elements
- Information element extensibility must be explicitly accounted for in the data model design
- For instance, an XML schema would need to specify when/how extensibility is permitted

# Next Steps

- Perform a detailed analysis of the SACM Requirements and how they influence the selection of a data format
- Work towards selecting a data format to develop an endpoint information data model

# References

# Concise Binary Object Representation (CBOR)

- CBOR site (<http://cbor.io/>)
- RFC-7049 (<http://tools.ietf.org/html/rfc7049>)
- CDDL (<https://datatracker.ietf.org/doc/draft-greevenbosch-appsawg-cbor-cddl/>)

# JavaScript Object Notation (JSON)

- JSON Site (<http://www.json.org/>)
- ECMA-404 (<http://www.ecma-international.org/publications/standards/Ecma-404.htm>)
- RFC-7159 (<https://tools.ietf.org/html/rfc7159>)
- RFC-7493 (<https://tools.ietf.org/html/rfc7493>)
- JSON Schema Site (<http://json-schema.org/>)
- JSON Schema in IETF (<https://tools.ietf.org/html/draft-zyp-json-schema-04>)

# Extensible Markup Language (XML)

- W3C XML Site (<http://www.w3.org/XML/>)
- XML 1.0 (<http://www.w3.org/TR/2008/REC-xml-20081126/>)
- XML Schema 1.1 (<http://www.w3.org/XML/Schema>)
- Namespaces in XML 1.0 (<http://www.w3.org/TR/2006/REC-xml-names-20060816/>)
- Namespaces in XML 1.1 (<http://www.w3.org/TR/2006/REC-xml-names11-20060816/>)