Binary Representation of Device Descriptions: CBOR versus RDF HDT

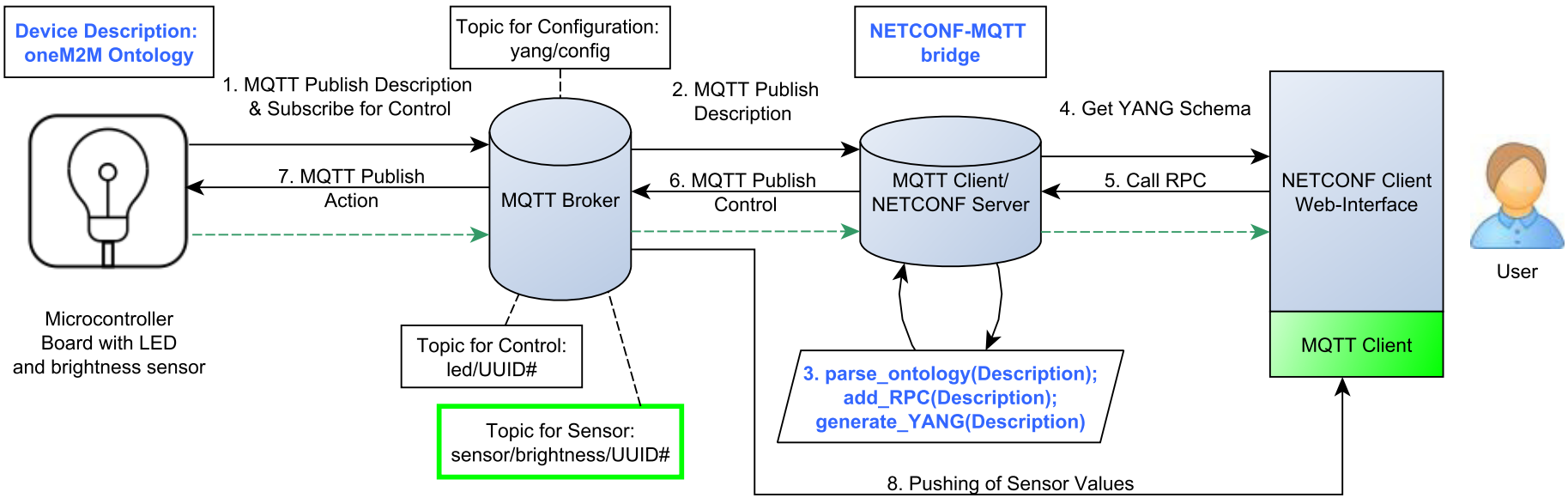
Kristina Sahlmann, Alexander Lindemann, Bettina Schnor

Hochschule für Technik und Wirtschaft Berlin
University of Applied Sciences

Universität Potsdam
Institut für Informatik und Computational Science
Professur Betriebssysteme und Verteilte Systeme

Fachgespräche Sensornetze 2018
Braunschweig, den 13.-14. September 2018

Extended System-Architecture with NETCONF-MQTT bridge



Challenge

Ontology-based Device Description
has a file size of ca. 19 kB using JSON-LD syntax.

Ontology for Device Descriptions

Example of ontology in JSON-LD Syntax

```
1 ...
2 "@id" : "https://www.cs.uni-potsdam.de/bs/research
    /myno#myDevice",
3 "http://www.onem2m.org/ontology/Base_Ontology/
    base_ontology#Device" ],
4 "http://www.onem2m.org/ontology/Base_Ontology/
    base_ontology#hasFunctionality"
5 : [ { "@id" : "https://www.cs.uni-potsdam.de/bs/
    research/myno#switchOff" } ],
6 ...
```

Ontology for Device Descriptions

Example of ontology in JSON-LD Syntax, optimized

```
1  { "@context": {
2      "oneM2M": "http://www.onem2m.org/ontology/
           Base_Ontology/base_ontology#",
3      "myno": "https://www.cs.uni-potsdam.de/bs/
           research/myno#" },
4    "@graph": [
5  {        "@id": "myno:myDevice",  ...  ],
6        "oneM2M:hasFunctionality": [
7          { "@id": "myno:funcSwitchOff" },   ...
```

Requirements for compression of ontology file

- small file size regarding 32 kiB RAM (and 512 kiB Flash memory)
- ontology data structures, particularly strings and structural elements like curly and square braces, should be compressed efficiently
- possibility to edit the file on a constrained device e.g. to change the UUID

Evaluation of binary compression

Two obvious formats

- CBOR
- RDF HDT

Concise Binary Object Representation (CBOR)

RFC 7049 [BH13] in 2013

Each byte is encoded as a major type (the high-order 3 bits) and additional information (the low-order 5 bits).

X X X	X X X X X
major type	additional information

CBOR Major Types

	Major Type
0	unsigned integer
1	negative integer
2	byte string
3	text string
4	array of data items
5	map of pairs of data items
6	optional semantic tagging of other major types
7	floating-point numbers and simple data types

CBOR Encoding Examples

datatype	value	major type	add. info.	hex	binary
integer	20	0	20	0x14	000 10100
string	CBOR	3	4 (str. length)	0x6443424F52	*

*011 00100 01000011 01000010 01001111 01010010

CBOR Evaluation Results

Better result was achieved through JSON-LD optimization

- introduce context prefix for namespaces
- remove unnecessary white spaces and tabs

Ontology file size in Bytes for JSON-LD and CBOR

JSON-LD	19,771
CBOR	16,126

18.43% space savings

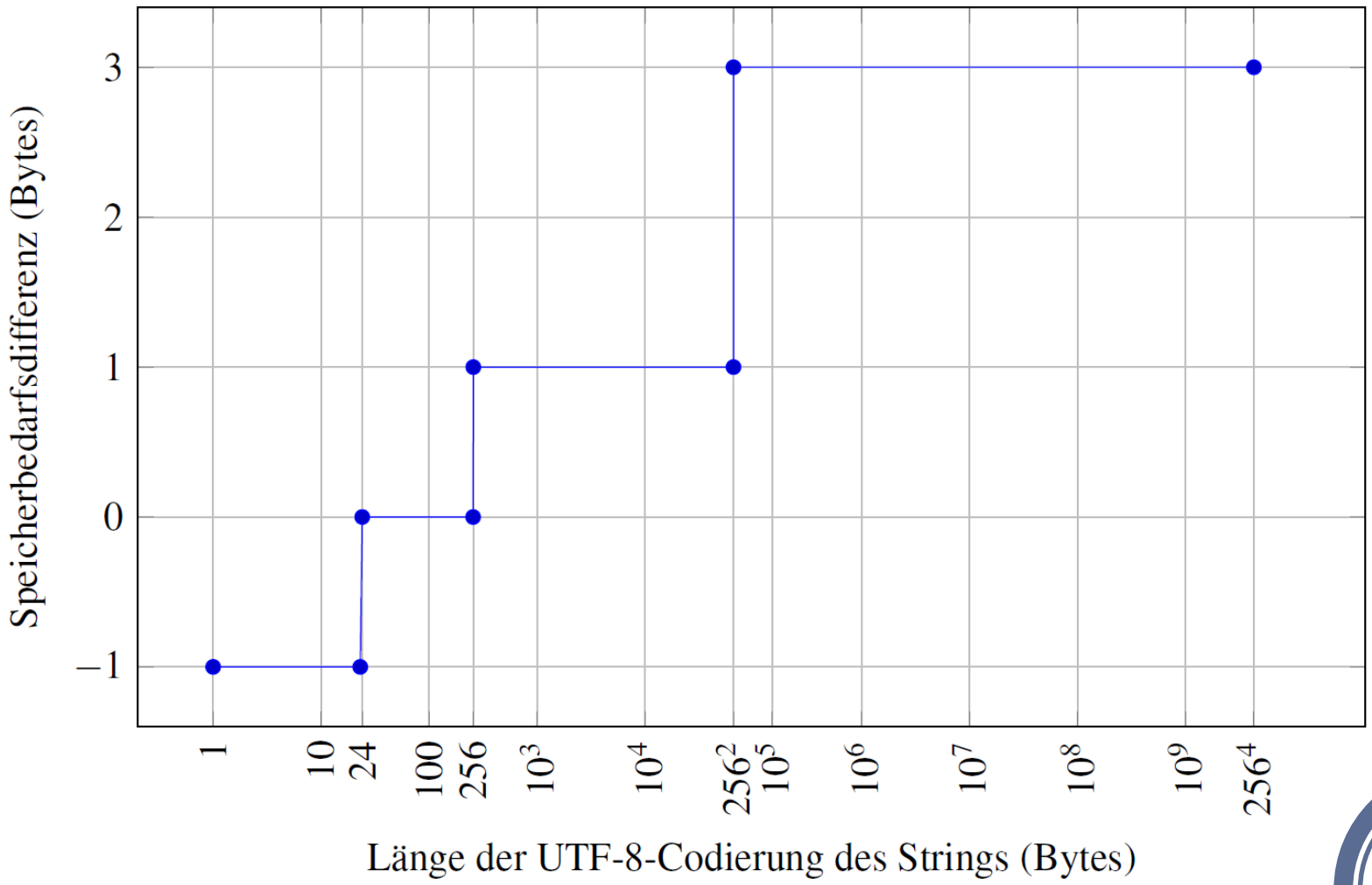
Ontology file size in Bytes for JSON-LD (opt.) and CBOR

JSON-LD (opt.)	7,640
CBOR	**6,520**

14.65% space savings

String Encoding in CBOR comparing to JSON [Hof18]

Additional storage consumption for encoding of strings in CBOR comparing to JSON



Possibility to replace the UUID using placeholders

¢h@context¦dbasex;http://www.onem2m.org/ontology/Base_Ontology/base_ontology#cowlxRS
http://www.w3.org/2002/07/owl#crdfx+http://www.w3.org/1999/02/22-rdf-syntax-ns#drdfsx%http://www.w3.org/2000/01/rdf-schema#c
xsdx!http://www.w3.org/2001/XMLSchema#cynmxEMhttp://yang-netconf-mqtt#f@graph
≐¢c@idlbase:Commande@typeiowl:Class¢c@idxGSbase:ControllingFunctionalitye@typeiowl:Class¢c@idkbase:Devicee@typeiowl:Class¢c
@idxESCbase:MeasuringFunctionalitye@typeiowl:Class¢c@idnbase:Operatione@typeiowl:Class¢c@idsbase:OperationInpute@typeiowl:C
lass¢c@idtbase:OperationOutpute@typeiowl:Class¢c@idsbase:OperationStatee@typeiowl:Class¢c@idtbase:OutputDataPointe@typeiowl:
Class¢c@idlbase:Servicee@typeiowl:Class¢c@idrbase:ThingPropertye@typeiowl:Class¢c@idsbase:exposesCommande@typerowl:ObjectPro
perty¢c@idxEMbase:exposesFunctionalitye@typevowl:AnnotationProperty¢c@idobase:hasCommande@typerowl:ObjectProperty¢c@idxUSba
se:hasDataRestriction_patterne@typetowl:DatatypeProperty¢c@idubase:hasFunctionalitye@typevowl:AnnotationProperty£c@idmbase:h
asInpute@typerowl:ObjectPropertyjrdfs:range¡c@idrbase:ThingProperty¢c@idqbase:hasOperatione@typerowl:ObjectProperty¢c@idvbas
e:hasOperationStatee@typerowl:ObjectProperty¢c@idnbase:hasOutpute@typerowl:ObjectProperty¢c@idwbase:hasOutputDataPointe@type
rowl:ObjectProperty¢c@idobase:hasServicee@typerowl:ObjectProperty¢c@idrbase:hasSubServicee@typerowl:ObjectProperty¢c@idubase
:hasThingPropertye@typerowl:ObjectProperty¢c@idmbase:hasValuee@typetowl:DatatypeProperty£c@idxCAN
http://yang-netconf-mqtte@typelowl:Ontologykowl:imports¡c@
idxAhttp://www.onem2m.org/ontology/Base_Ontology/base_ontology-v0_9_0
£c@idsynm:YangDescriptione@typeiowl:Classordfs:subClassOf¡c@idrbase:ThingProperty£c@idlynm:cmdGreene@type,sowl:NamedIndividu
allbase:Commandmbase:hasInput¡c@idmynm:uuidInput£c@idjynm:cmdOffe@type,sowl:NamedIndividuallbase:Commandmbase:hasInput¡c@idm
ynm:uuidInput£c@idiynm:cmdOne@type,sowl:NamedIndividuallbase:Commandmbase:hasInput¡c@idmynm:uuidInput£c@idoynm:cmdSetColore@
type,sowl:NamedIndividuallbase:Commandmbase:hasInput,¡c@idnynm:colorInput¡c@idmynm:uuidInput¤c@idnynm:colorInpute@type,sowl:
NamedIndividualsbase:OperationInputxUSbase:hasDataRestriction_pattern…callegreenforangecredfyellowubase:hasThingProperty¡c@
idqynm:colorYangDesc£c@idqynm:colorYangDesce@type,sowl:NamedIndividualsynm:YangDescriptionmbase:hasValueocolor
parameter£c@idrynm:deviceCategorye@type,sowl:NamedIndividualrbase:ThingPropertymbase:hasValuehLED-LAMP£c@idnynm:deviceDesce@
type,sowl:NamedIndividualsynm:YangDescriptionmbase:hasValuexRSMQTT-Device identified by
UUID£c@idnynm:deviceUuide@type,sowl:NamedIndividualrbase:ThingPropertymbase:hasValuex#F97DF79-8A12-4F4F-8F69-6B8F3C2E78DD£c@
idrynm:funcDescBrighte@type,sowl:NamedIndividualsynm:YangDescriptionmbase:hasValuexCANrequest brightness
value£c@idqynm:funcDescColore@type,sowl:NamedIndividualsynm:YangDescriptionmbase:hasValuewset arbitrary LED
color£c@idqynm:funcDescGreene@type,sowl:NamedIndividualsynm:YangDescriptionmbase:hasValuexSUBSet the led color to
green£c@iduynm:funcDescSwitchOffe@type,sowl:NamedIndividualsynm:YangDescriptionmbase:hasValuetSwitches the led
off£c@idtynm:funcDescSwitchOne@type,sowl:NamedIndividualsynm:YangDescriptionmbase:hasValuesSwitches the led
on£c@idqynm:funcGetBrighte@type,sowl:NamedIndividualxESCbase:MeasuringFunctionalityubase:hasThingProperty¡c@idrynm:
Bright¤c@idpynm:funcSetColore@type,sowl:NamedIndividualxGSbase:ControllingFunctionalityobase:hasCommand¡c@idoynm:
rubase:hasThingProperty¡c@idqynm:funcDescColor¤c@idpynm:funcSetGreene@type,sowl:NamedIndividualxGSbase:Controlli
lityobase:hasCommand¡c@idlynm:cmdGreenubase:hasThingProperty¡c@idqynm:funcDescGreen¤c@idqynm:funcSwitchOffe@typ
ndividualxGSbase:ControllingFunctionalityobase:hasCommand¡c@idjynm:cmdOffubase:hasThingProperty¡c@iduynm:funcDe
c@idpynm:funcSwitchOne@type,sowl:NamedIndividualxGSbase:ControllingFunctionalityobase:hasCommand¡c@idiynm:cmdOn

Ontology Structure

The ontology defines so-called *Triples*:
Subject $\rightarrow$ ***Predicate*** $\rightarrow$ ***Object*** ***(SPO)***

An example for an *SPO triple* within IoT:
myDevice $\rightarrow$ ***hasFunctionality*** $\rightarrow$ ***funcSwitchOff***

RDF HDT

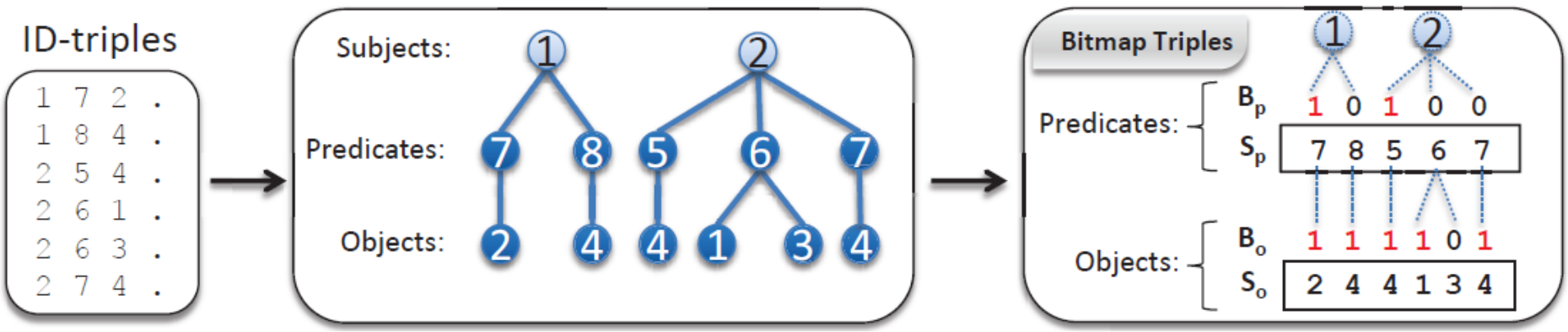
W3C Member Submission [FMPGP11] from 2011

HDT splits ontology file in three logical components:
Header, Dictionary and Triples

- Header holds meta data about the file using a plain RDF structure.
- The Dictionary is a catalog of triples for the used terms based on Front-Coding which is again based on words with similar prefix. Subjects, Predicates and Objects are subsumed in sections to avoid repeating. Each section is sorted lexicographically and then IDs are assigned to each term.
- RDF Triples are now tuples of three IDs.

RDF HDT

HDT proposes an encoding of Triples called Bitmap Triples (BT)



RDF HDT File Example

Header with Meta Data

```
$HDT.<http://purl.org/HDT/hdt#HDTv1>..v5$HDT.ntriples.length=1834;.Âr<file://mqtt-cap-json-v5-2-turtle-opt.owl>
<http://www.w3.org/1999/02/22-rdf-syntax-ns#type> <http://purl.org/HDT/hdt#Dataset>
..<file://mqtt-cap-json-v5-2-turtle-opt.owl> <http://www.w3.org/1999/02/22-rdf-syntax-ns#type>
<http://rdfs.org/ns/void#Dataset> ..<file://mqtt-cap-json-v5-2-turtle-opt.owl> <http://rdfs.org/ns/void#triples> "179"
..<file://mqtt-cap-json-v5-2-turtle-opt.owl> <http://rdfs.org/ns/void#properties> "19"
..<file://mqtt-cap-json-v5-2-turtle-opt.owl> <http://rdfs.org/ns/void#distinctSubjects> "61"
..<file://mqtt-cap-json-v5-2-turtle-opt.owl> <http://rdfs.org/ns/void#distinctObjects> "76"
..<file://mqtt-cap-json-v5-2-turtle-opt.owl> <http://purl.org/HDT/hdt#statisticalInformation> _:statistics
..<file://mqtt-cap-json-v5-2-turtle-opt.owl> <http://purl.org/HDT/hdt#publicationInformation> _:publicationInformation
..<file://mqtt-cap-json-v5-2-turtle-opt.owl> <http://purl.org/HDT/hdt#formatInformation> _:format ..._:format
<http://purl.org/HDT/hdt#dictionary> _:dictionary ..._:format <http://purl.org/HDT/hdt#triples> _:triples ..._:dictionary
<http://purl.org/dc/terms/format> <http://purl.org/HDT/hdt#dictionaryFour> ..._:dictionary
<http://purl.org/HDT/hdt#dictionarynumSharedSubjectObject> "41" ..._:dictionary <http://purl.org/HDT/hdt#dictionarymapping>
"1" ..._:dictionary <http://purl.org/HDT/hdt#dictionarysizeStrings> "3750" ..._:dictionary
<http://purl.org/HDT/hdt#dictionaryblockSize> "16" ..._:triples <http://purl.org/dc/terms/format>
<http://purl.org/HDT/hdt#triplesBitmap> ..._:triples <http://purl.org/HDT/hdt#triplesnumTriples> "179" ..._:triples
<http://purl.org/HDT/hdt#triplesOrder> "SPO" ..._:statistics <http://purl.org/HDT/hdt#originalSize> "9992" ..._:statistics
<http://purl.org/HDT/hdt#hdtSize> "4279" ..._:publicationInformation <http://purl.org/dc/terms/issued>
"2018-06-14T09:10:46+02:00"..
```

RDF HDT File Example

Dictionary


```
$HDT.<http://purl.org/HDT/hdt#dictionaryFour>.mapping=1;sizeStrings=6234;..à.©+„.)..„|..dÛŠ$Î[™http://www.onem2m.org/ontology/
Base_Ontology/base_ontology#Command.¾ntrollingFunctionality.»Device.»MeasuringFunctionality.»Operation.ÄInput.ÄState.↳utputDat
aPoint.»Service.»ThingProperty.‡yang-netconf-mqtt#YangDescription.™cmdGreen.œOff..n.œSetColor.šolorInput.http://yang-netconf-m
qtt#colorYangDesc.™deviceCategory.ŸDesc.ŸUuid.™funcDescBright.¡Color.¡Green.¡SwitchOff.¨n..GetBright..SetColor. Green.žwitchOf
f.¤n.™opDescState.›MqttGreen.http://yang-netconf-mqtt#opMqttOff. n.ŸSetColor.›State.šutDpBrightness.™servBrightness..Netconf.™
uuidInput..YangDesc.-ÝÜ
."I,....ƒV..&.4w]Íhttp://www.onem2m.org/ontology/Base_Ontology/base_ontology#OperationOutput.»exposesCommand.ÂFunctionality.»h
asCommand.¾DataRestriction_pattern.¾Functionality.¾Input.¾Operation.ÇState.¿utput.ÄDataPoint.¾Service.¿ubService.¾ThingPropert
y.¾Value.‡yang-netconf-mqtt.http://yang-netconf-mqtt#mqttMethod..Topic.šyDevice.™paramYangDesc.,õOœ.“w,.↳..ƒV.XÞ.»ÃÊ=http://ww
w.onem2m.org/ontology/Base_Ontology/base_ontology#exposesCommand.ÂFunctionality.»hasCommand.¾DataRestriction_pattern.¾Function
ality.¾Input.¾Operation.ÇState.¿utputDataPoint.¾Service.¿ubService.¾ThingProperty.¾Value.‹w3.org/1999/02/22-rdf-syntax-ns#type
.'2000/01/rdf-schema#range.¥subClassOf.http://www.w3.org/2002/07/owl#imports.‡yang-netconf-mqtt#mqttMethod..Topic.k²€i.£.Ž.É..
„i.↑°.~Z./W"ERROR"^^<http://www.w3.org/2001/XMLSchema#string>..F97DF79-8A12-4F4F-8F69-6B8F3C2E78DD"^^<http://www.w3.org/2001/X
MLSchema#string>..GREEN"^^<http://www.w3.org/2001/XMLSchema#string>..LED-LAMP"^^<http://www.w3.org/2001/XMLSchema#string>..MQT
T-Device identified by
UUID"^^<http://www.w3.org/2001/XMLSchema#string>..NOOP"^^<http://www.w3.org/2001/XMLSchema#string>..OFF"^^<http://www.w3.org/2
001/XMLSchema#string>.,K"^^<http://www.w3.org/2001/XMLSchema#string>.,N"^^<http://www.w3.org/2001/XMLSchema#string>..Set the
led color to green"^^<http://www.w3.org/2001/XMLSchema#string>.,witches the led
off"^^<http://www.w3.org/2001/XMLSchema#string>.“n"^^<http://www.w3.org/2001/XMLSchema#string>..Target UUID for
request"^^<http://www.w3.org/2001/XMLSchema#string>..all"^^<http://www.w3.org/2001/XMLSchema#string>..color
parameter"^^<http://www.w3.org/2001/XMLSchema#string>.†"^^<http://www.w3.org/2001/XMLSchema#string>."error"^^<http://www.w3.or
g/2001/XMLSchema#string>..green"^^<http://www.w3.org/2001/XMLSchema#string>..led"^^<http://www.w3.org/2001/XMLSchema#string>..
nothing to do"^^<http://www.w3.org/2001/XMLSchema#string>..orange"^^<http://www.w3.org/2001/XMLSchema#string>..parameter
value"^^<http://www.w3.org/2001/XMLSchema#string>..red"^^<http://www.w3.org/2001/XMLSchema#string>.ƒquest brightness
value"^^<http://www.w3.org/2001/XMLSchema#string>..sensor/brightness/led01"^^<http://www.w3.org/2001/XMLSchema#string>.ƒt
arbitrary LED
color"^^<http://www.w3.org/2001/XMLSchema#string>.,uccessful"^^<http://www.w3.org/2001/XMLSchema#string>..yellow"^^<ht
.w3.org/2001/XMLSchema#string>.€http://www.onem2m.org/ontology/Base_Ontology/base_ontology-v0_9_0.‹w3.org/2002/07/o
onProperty.žClass.žDatatypeProperty.http://www.w3.org/2002/07/owl#NamedIndividual.žObjectProperty.Ÿntology./h0.
```


RDF HDT File Example

Triples with IDs

$HDT.<http://purl.org/HDT/hdt#triplesBitmap>.order=1;.Yé.üïÿ«JUU%IA..)'þþ;._«-Ñ.3.#ÿßVaÛ¶mÛÝ]z¿ßÛY/Änÿÿ.¦.5ÿ¬x..üÐÎ9çœsÎ9hœqÆ.
GK®¹æšk®¹æškŽ¹ÁÜ`n07#..ää..99ƒANÎ`.""0ç¦HNœå.sÍ9çœsî9çœsÎ¹è\QÌ5.±.%6..3..HHHHHHHHHHH.(.J(.J(.J.(.J7;>@E..J8.J-.J..J+.JA.JC.J3.
J4.J5.J..J...J...J...J...J:=D.J.($.J,<.($.J0<.($.J2<..($.J9<*/1..J.JB.%.J.....
!"#&.J.).J6.JHKGKIGK.KKKKKKKILFII.....'....J?.J]¨õU

RDF HDT Evaluation Results

Optimization through redundancy avoidance in RDF HDT

Ontology file size in Bytes for Turtle and RDF HDT

Turtle	20,385
HDT	5,568

72.68% space savings

Ontology file size in Bytes for Turtle (opt.) and RDF HDT

Turtle (opt.)	9,992
HDT	5,513

44.82% space savings

RDF HDT Evaluation Results

Better string syntax in N-Triples through elimination of the data type `http://www.w3.org/2001/XMLSchema#string` at every string

Ontology file size in Bytes for N-Triples

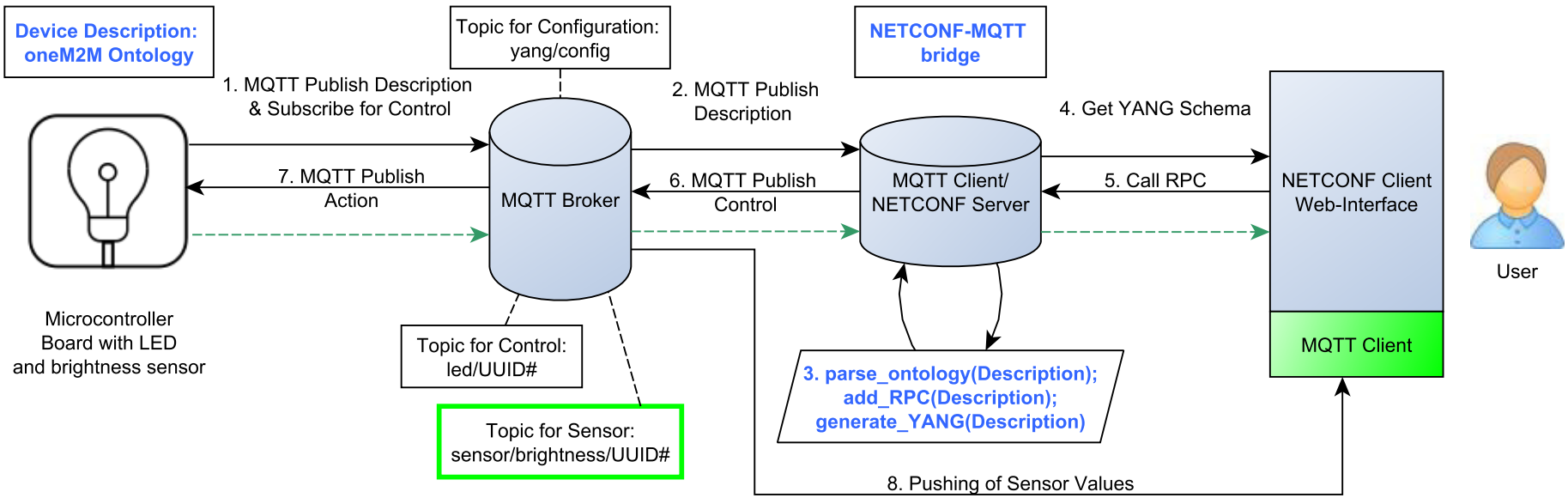
N-Triples	27,025
HDT	4,307

84.06% space savings

Conclusion

- RDF HDT shows better results for our use case
- RDF HDT has space savings (up to 84 %) because of consideration the data structure of ontology
- CBOR has poor compression (less than 15 %) because it does not compress long strings
- CBOR is still suitable for sensor data (integer values etc.)
- possible to edit the files on constrained devices e.g. to change the UUID

Demo: An Ontology-based NETCONF-MQTT Bridge for Sensor Devices in the IoT



Literatur


C. Bormann and P. Hoffman.
Concise Binary Object Representation (CBOR).
RFC 7049, IETF, October 2013.


Javier D. Fernández, Miguel A. Martínez-Prieto, Claudio Gutierrez, and Axel Polleres.
Binary RDF Representation for Publication and Exchange (HDT), 2011.
Available from: `https://www.w3.org/Submission/HDT/`.


Christian Hoffmann.
Effiziente Datendarstellung mittels CBOR für das Internet of Things.
Bachelor's Thesis, University of Potsdam, Germany, May 2018.