Internet Engineering Task Force   A. Pelov, Ed.
Internet-Draft   Acklio
Intended status: Informational   L. Toutain, Ed.
Expires: January 5, 2016   Telecom Bretagne
                                                       Y. Delibie, Ed.
                                                              Kerlink
                                                         July 4, 2015

                    Constrained Signaling Over LR-WAN
                        draft-pelov-core-cosol-00

Abstract

   This document presents a new type of far-reaching, low-rate radio
   technologies and an extensible mechanism to operate these networks
   based on CoAP. The emerging Wide Area Networks based on them - Low-
   Rate WAN (LR-WAN) preset a particular set of constraints, which
   places them at the intersection of infrastructure networks, ultra-
   dense networks, delay-tolerant networks and low-power and lossy
   networks. The main objectives of LR-WAN signaling is to minimize the
   number of exchanged messages, minimize the size of each message in a
   secure and extensible manner. This document describes the use of the
   Constrained Application Protocol (CoAP) as the main signaling
   protocol for LR-WANs, over which minimal messages are exchanged
   allowing the full operation of the network, such as authentication,
   authorization, and management. The use of CoAP signaling provides a
   generic mechanism that can be applied to different LR-WAN
   technologies.

Table of Contents

[1](#). Introduction

   The goal of this document is to provide the necessary mechanisms to
   operate a Low-Rate Wide-Area Network (LR-WAN) by using IETF CoAP
   [[RFC7252](#)] as a core signaling protocol.

Far-Reaching, low-rate communication technologies (FARE) have emerged
in the past several years, and are the base for building Low-Rate
Wide-Area Networks (LR-WAN). LR-WANs have the following
characteristics:

o Work in narrow, license-free (ISM) bands with good propagation
   properties (< 1GHz)

o Low- to very-low throughput (1-200 kbps)

o Low-power operation (25 mW in Europe)

o Far-Reaching communication capabilities (up to 30 km with line-of-
   sight, several km in urban environment)

o Strong channel access restrictions (1% to 10% duty cycling)

o Infrastructure-based

o Star topology

LR-WANs are built on Far-Reaching Radio communication technologies
(FARE), which use advanced signal processing techniques and
combination of appropriate modulation and coding approaches to
provide the aforementioned radio characteristics.

The absence of license fees and the Far-Reaching connectivity allow
for an extremely competitive pricing of LR-WANs compared to other
networking technologies, e.g. cellular or mesh. LR-WANs are
sometimes referred to as LPWAN (Low-Power WAN), e.g. by Semtech
[LoRa]. Even though LR-WANs are extremely limited in terms of
network performance, they are enough for a wide class of
applications, among which [LTN001]:

o Metering (water, gas, electricity)

o Infrastructure networks (water, gas, electricity, roads,
   pipelines, drains)

o Environment/Smart City (waste management, air pollution monitoring
   and alerting, acoustic noise monitoring, public lighting
   management, parking management, self service bike rental, digital
   board monitoring, water pipe leakage monitoring)

o Environment/Country side (soil quality, livestock surveillance,
   cattle and pet monitoring, climate, irrigation)

o Remote monitoring (house, building)

o Industrial (water tank, asset tracking)

o Automotive (vehicle tracking, impact detection, pay as you drive,
   assistance request, ...)

o Logistics (goods tracking, conservation monitoring)

o Healthcare (patient monitoring, home medical equipment usage)

o House appliances (pet tracking, white goods, personal asset)

o Truck (tyre monitoring)

o Identification (authentication)

The IEEE is studying LR-WANs, but limited to the case of low-energy
critical infrastructure monitoring (LECIM), under the group IEEE
802.15.4k [IEEE.802-15.4k].

The combination of the above characteristics and the envisioned
applications define a new class of networks with the following unique
constraints:

o Potentially extremely high density (expected of up to 10k-100k+
   end-devices managed by a single radio antena)

o Coexistence of delay-tolerant and critical applications (metering
   and alarms)

o Low-power, low-throughput, lossy connectivity (use of ISM bands)

o Limited payload (100 bytes max, typically less than 50 bytes, 12
   bytes for UNB)

CoAP is a client-server protocol specialized for constrained networks
and devices. CoAP is highly optimized, extensible, standard
protocol, which in conjunction with the Concise Binary Object
Representation (CBOR) is the ideal candidate for the signaling
protocol of the control plane of an LR-WAN.

It can be used during all stages of the lifecycle of the network,
e.g. discovery, authentication, operation. Furthermore, this can be
achieved by following RESTful management paradigm, by using a
particular resource tree definition or adopting CoMI
[I-D.vanderstok-core-comi].

1.1. Requirements Language

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].

2. LR-WAN Technologies

2.1. FARE technologies

There are two classes of Far-Reaching radio Technologies, using
different radio modulation approaches:

o Ultra Narrow Band (UNB)

o Spread-spectrum (SS)

An example of UNB is the technology developed and promoted by SigFox
[SigFox]. Semtech LoRa [LoRa] uses a direct-sequence spread-spectrum
with orthogonal codes (OSSS).

Both approaches have their advantages and will coexist in the future,
as there are currently several operators, which deploy the two types
in the same areas.

2.2. Physical Layer Characteristics

At the physical layer, the important part is the possibility to
reconstruct the signal at long distances. The used ISM bands are
defined around the world (e.g. 868 MHz in Europe and 900 MHz in USA)
and require a 1% (or 10%) duty cycling, or alternatively - advanced
detection and channel reallocation techniques. In reality, all
deployed networks use the duty cycling limitation, with the following
distinction. There is one 100kHz band in which 10% duty cycling is
allowed, with a slightly more emission power. The rest of the bands
are limited at 1% duty cycling and very restricted power of emission
(e.g. 25 mW in Europe).

UNB LR-WANs make the distinction between Uplink and Downlink, first
depending on the modulation, and second with the 10% duty-cycling
channel been used for the Downlink. OSSS LR-WANs make no such
distinction, although for the operation of a network, an operator can
chose to use the same Uplink/Downlink channel separation.

Note that the 1% or 10% duty-cycle limitation counts for all trafic
originating from an electronic equipment, e.g. an antena managing
100k objects must obey the same limitation as an end-device, with all

frames emitted from the antena (data, acknowledgements) counting towards its quota.

2.2.1. Ultra Narrowband FARE radios

Ultra Narrowband (UNB) technologies generally possess the following physical layer characteristics [LTN003]:

o Uplink:

   * channelization mask 100kHz (600 kHz USA)

   * baud rate 100 bauds (600 bauds USA)

   * modulation BPSK

o Downlink:

   * channelization mask: dynamic selection

   * down link baud rate: 600 baud

   * modulation scheme: GFSK

   * downlink transmission power: 500 mW, 10% duty cycle

2.2.2. Spread-spectrum FARE radios

OSSS technologies possess the following physical layer characteristics [LTN003]:

o channelization mask: from 8 kHz to 500 kHz (depending on spreading factor)

o chip rate: 8 kcps up to 500 kcps

o data rate: 30-50 000 bps

o modulation scheme: equivalent to DSSS with orthogonal signaling

No particular distinction is made between the Uplink and the Downlink.

2.3. MAC Layer Characteristics

Several proprietary MAC frame formats exist for UNB and OSSS. However, they are designed to operate the network in a centralized, highly-vertically-integrated fashion. The only standard MAC frame

format is the IEEE 802.15.4k, which is based on the well-known IEEE
802.15.4 with the addition of a fragmentation sub-layer.

The channel access method is based on ALOHA, although it is up to the
network operator to chose if an appropriate Node-F polling should be
implemented.

3. CoSOL Architecture

3.1. General LR-WAN architecture

We can identify three types of entities in a typical LR-WAN. These
are:

o Node-F: far-reachable node, e.g. the end-point, object, device

o Node-R: radio relay, bridging the Far-Reaching radio technology to
   a different medium (often a LAN or cellular WAN)

o Node-G: gateway node, interconnection between the radio-relay node
   and the Internet

```
+--------+              +--------+            +--------+
| Node-F | <-- FARE --> | Node-R | <-- IP --> | Node-G |
+--------+              +--------+            +--------+
```
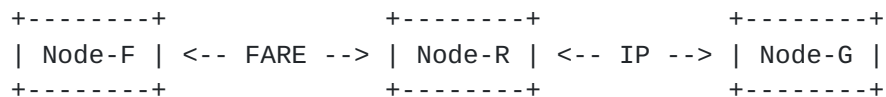
   General architecture of an LR-WAN. FARE radio technology is used
                only between the Node-F and the Node-R.

                             Figure 1

Of these, only Node-F and Node-R communicate through a FARE
technology. However, due to the extreme constraints of these
technologies, they are always behind a gateway (Node-G). Note, that
the Node-R and Node-G can be collocated, e.g. on a single hardware
equipment.

The Node-G is connected to the Internet and is assumed to have
sufficient computational resources to store a context for each of the
Node-Fs. The strong limitation here is the radio link.

In an actual deployment, a (limited) set of Node-Rs cover a large
area with a potentially very-high number of Node-Fs. A single Node-G
is capable of controlling all Node-Rs.

```
                      o
           o         o
            o           (((*)))-------\
          o               o           |
           o             o            |
        o      (((*)))---------------+------Node-G
          o       o                   |
           o           o              |
      o      (((*)))------------------+
               o                      |
          o     o  o          o       |
            o     (((*)))----------/
            o                 o
               o           o
```

```
                                    o       = Node-F
                                  (((*)))   = Node-R
```
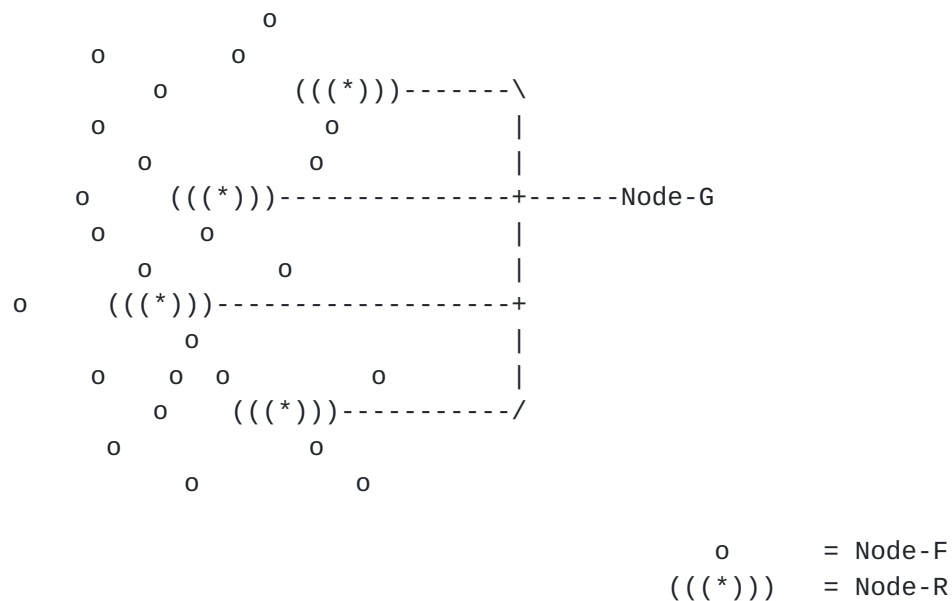
       An example coverage of an area with several Node-Rs. Note that a
              single Node-F may be covered by several Node-Rs.

                                  Figure 2

3.2. Node-F lifecycle

   Similar to other wireless infrastructure-based technologies, a Node-F
   can go through several stages:

   o Semi-Association

   o Network Discovery

   o Authentication

   o Association

   o Dissociation

   The Node-F state machine is then the following:

```
     +-------------------------------------------------------+
     |                                                       |
     V                                                       |
 Semi-associated ----------+                                 |
     |        ^            |                                 |
     |        |            |                                 |
     V        |            V                                 |
   Disconnected -> Network discovery -> Associated -> Authenticated
        ^                                              |
        |                                              |
        +----------------------------------------------+
```

                 Node-F connectivity state machine.

                           Figure 3
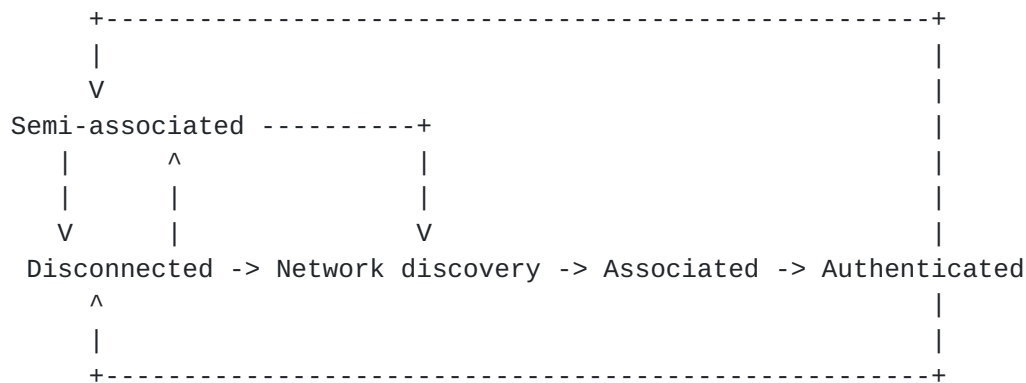
   The Node-F can be in Semi-Associated mode. Upon start, and depending
   on the application, a Node-F can use a state of uni-directional
   communication, where it is considered semi-associated to the network.
   In that state, the Node-F broadcasts frames, handled by the Node-G,
   but the network cannot join the Node-F on a regular basis. This is a
   degraded LR-WAN operating mode and if caution is not used, can lead
   to significant scalability and evolvability issues.

   The Network Discovery can be reactive or proactive. The former is
   based on detecting beacon frames sent periodically by the network
   (e.g. Node-G). The latter is implemented by the Node-F broadcasting
   probe request frames, to which all appropriate Node-Gs must respond.

   Once a network has been discovered, the Node-F and the Node-G can
   perform mutual authentication.

   Upon authentication, the Node-G configures the necessary network
   parameters of the Node-F, which is henceforth associated to the
   network. The association request may be explicit or implicit, in
   which case after successful authentication the Node-F enters
   automatically the associated state. In this stage there is bi-
   directional communication between the Node-F and the Node-G.

   Finally, the Node-F may decide to dissociate from the network by
   sending an explicit request. Upon dissociation the Node-G may
   release all contexts related to the Node-F and re-association
   requires going through the authentication stage again. Node mobility
   is achieved by explicitly dissociating from the old Node-G and then
   authenticating to the new Node-G. Implicit dissociation is also
   possible upon the expiration of predefined timers, or in case of
   mobility optimization.

3.3. CoAP as Signaling Protocol for LR-WANs

   Use as CoAP for signaling is implemented as follows. The MAC,
   network and/or transport layers MUST provide a mechanism to
   differentiate user data from signaling data frames (e.g. by using
   separate MAC addresses, IP addresses and/or UDP-ports). Both the
   Node-G and the Node-F are running CoAP servers for implementing the
   control plane. Frames exchanged over the FARE radio interface and
   marked as "signaling data" are handled by the corresponding control
   plane CoAP servers. CoAP requests are thus used to keep a shared
   vision of the network and the node between the two. This is realized
   by a virtual, shared resource-tree as described in Section 3.4.

   The Node-G runs a (virtual) CoAP server for each Node-F. This server
   is identified with a DNS name, e.g. "node123.home.node-
   g.example.com", which can be used explicitly in the CoAP messages via
   the Proxy-Uri option if needed.

   Note, that the Node-R acts only as a transceiver and as such is
   transparent from protocol point of view. As such, the following
   management scheme applies:

```
   +--------+                           +--------+
   | Node-F | <-- LR-WAN constraints --> | Node-G |
   +--------+                           +--------+
```
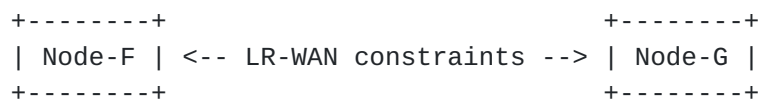
                  Node-F connectivity state machine.

                             Figure 4

3.3.1. Semi-Association

   When in a semi-associated state, a Node-F broadcasts its messages
   without performing network discovery, or association. If the Node-F
   is under the coverage of a Node-G, the Node-G will receive the
   broadcast, and forward the user data. The frames SHOULD be signed,
   so that they could be authenticated by the network. Layer 2
   acknowledgements MUST be used, and in some cases piggybacking on them
   can provoke the Node-F to associate to the network.

   The broadcast messages MUST include the necessary information to join
   the user data destination, and enough information for the Node-G to
   authenticate the message sender. This can be achieved through a
   Confirmable CoAP message, where the user data are POSTed to a well-
   known resource defined on the Node-G. DTLS with integrity check can
   be used, with long-lived keys negotiated by the Node-F and the
   network.

Even though an application can be implemented by using only simplex
association capabilities, there are huge negative consequences
related to scalability and evolvability in this case. For example, a
Node-F which periodically broadcasts information will occupy the
spectrum, even if there is no operator willing to accept its trafic.
In addition, no channel access management can be applied.

```
Node-F      Node-G
   |           |
   |           |
  +--------->|           Header: POST
   |   POST    |         Uri-Host: "destination.example.com"
   |           |         Uri-Path: "temp"
   |           |
   |           |
   |<---------+           Header: 2.01 Created
   |   2.01    |
   |           |
   |           |
```
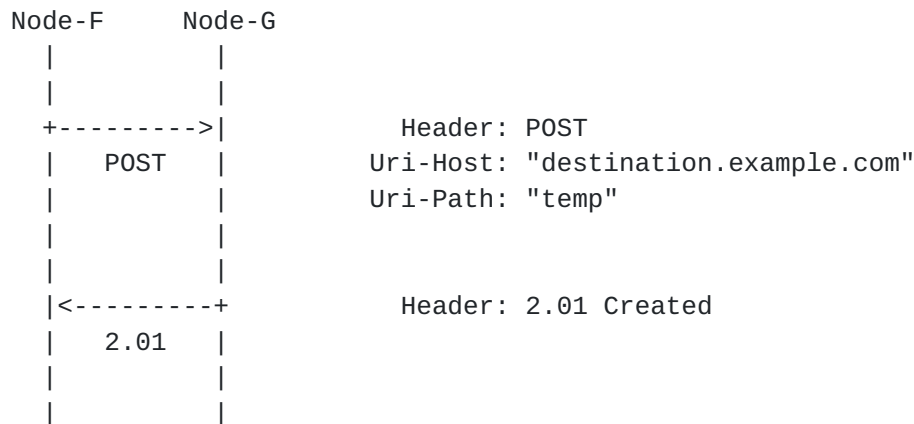
               Sending a message in a semi-associated state.

                             Figure 5

3.3.2. Network Discovery

A network can be discovered by a Node-F reactively or proactively.

Reactive network discovery is based on the detection of periodic
beacons emitted by the Node-G. The beacons are implemented with CoAP
messages with the No-Response option
[I-D.tcs-coap-no-response-option]. The Node-G POSTs its information
to a well-known resource, e.g. "/network/node-G/" or a resource alias
"/g" or CoMI YANG hash ID "/mg/GgQ".

```
Node-F      Node-G
   |           |
   |           |
   |<---------+           Header: POST
   |  POST /g |         Uri-Path: "g"
   |           |           [No-Responce]
   |           |
   |           |
```
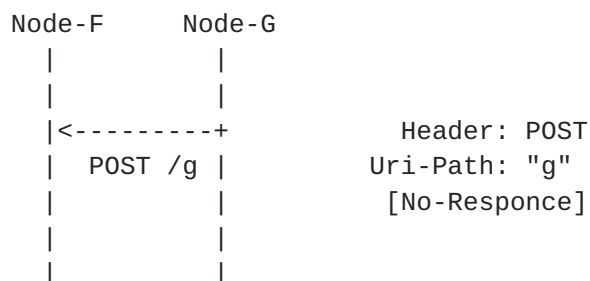
   Reactive network discovery. The Node-G sends periodically beacon
      messages, containing information pertinent to this network.

                             Figure 6

The CoAP POST request is processed at the Node-F. A resource is
created locally, with the representation, which provides the
appropriate network parameters, e.g. network ID, Node-G ID, and other
radio-related parameters, such as channel, beacon frequency and so
forth. This information allows the Node-F to begin the
authentication phase.

A Node-F may chose to proactively probe for the existence of network
coverage. In that case, it sends a Confirmable CoAP GET request to
obtain the information from a well-known resource, normally published
by the beacon messages, e.g. "/network/node-G/" or a resource alias
"/g" or CoMI YANG hash ID "/mg/GgQ".

```
Node-F     Node-G
   |          |
   |          |
  +--------->|           Header: GET
   |  GET /g  |         Uri-Path: "g"
   |          |           Accept: application/cbor
   |          |
   |          |
   |<---------+           Header: 2.05 Content
   |   2.05   |          Payload: ...
   |          |
```
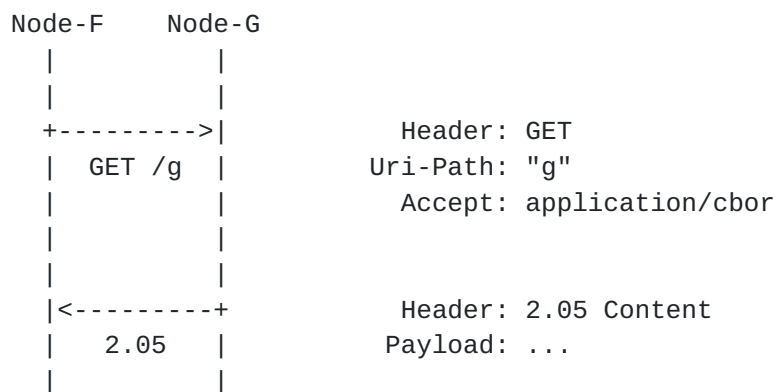
        Proactive network discovery. The Node-F request the information of
                        all surrounding Node-Gs.

                                Figure 7

Once the network is discovered, the Node-F has all necessary
information to start the authentication phase.

3.3.3. Association

Before being able to communicate, the Node-F must associate to the
network, and then eventually authenticate. The association phase
signals to the Node-G that there is a new device willing to
communicate with the network. This association SHOULD provide enough
information to allow the Node-G to start the authentication process.
For example, it may provide the AAA server, which could authenticate
the Node-F, or its EAP-Identity. Note, that the Node-F may elect to
mark the association message with the No-response option
[I-D.tcs-coap-no-response-option], waiting for the subsequent
authentication request from the Node-G.

```
   Node-F      Node-G
     |           |
     |           |
    +--------->|            Header: POST
    | POST /n  |          Uri-Path: "n"
    |           |           Payload: ...
    |           |
    |<---------+            Header: 2.01 Created
    |   2.01   |    Location-Path: "/n/n705"
    |           |
```
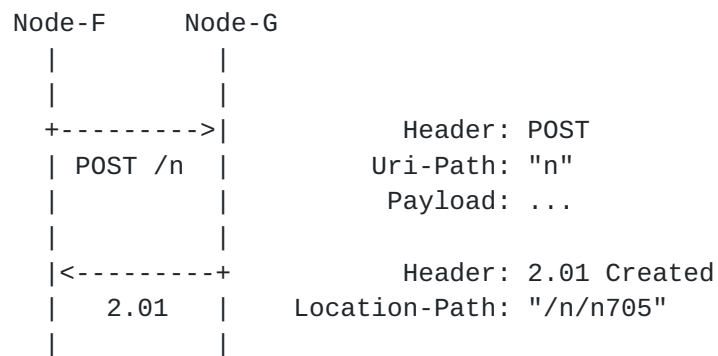
       Node-F associates to a network, by creating a corresponding resource
                       element on the Node-G.

                              Figure 8

3.3.4. Authentication

   The EAP-over-CoAP [I-D.garcia-core-security] specifies an approach
   to encapsulating EAP messages over CoAP. This allows to authenticate
   a Node-F, which wishes to join an LR-WAN, and negotiate the L2
   encryption keys, and DTLS keying material.

   As the Node-F has already associated to the Node-G, it is the Node-G
   that initiates the authentification request, by going directly to
   Step 1) of the EAP-over-CoAP specification.

```
    Node-F        Node-G
       |            |
       |            |
  1)   |<-----------+            Header: POST
       | POST /auth |        Uri-Path: "auth"
       |            |           [No-Responce]
       |            |
       |            |
  2)   +----------->|            Header: 2.01 Created
       | ACK /auth  |    Location-Path: "/auth/5"
       |            |
       |            |
       |            |
  3)   |<-----------+            Header: PUT
       | PUT /auth/5|        Uri-Path: "auth/5"
       |            |          Payload: EAP-PSK MSG 1
       |            |
       |            |
  4)   +----------->|            Header: 2.04 Changed
       | ACK /auth/5|        Payload: EAP-PSK MSG 2
       |            |
          ......
```
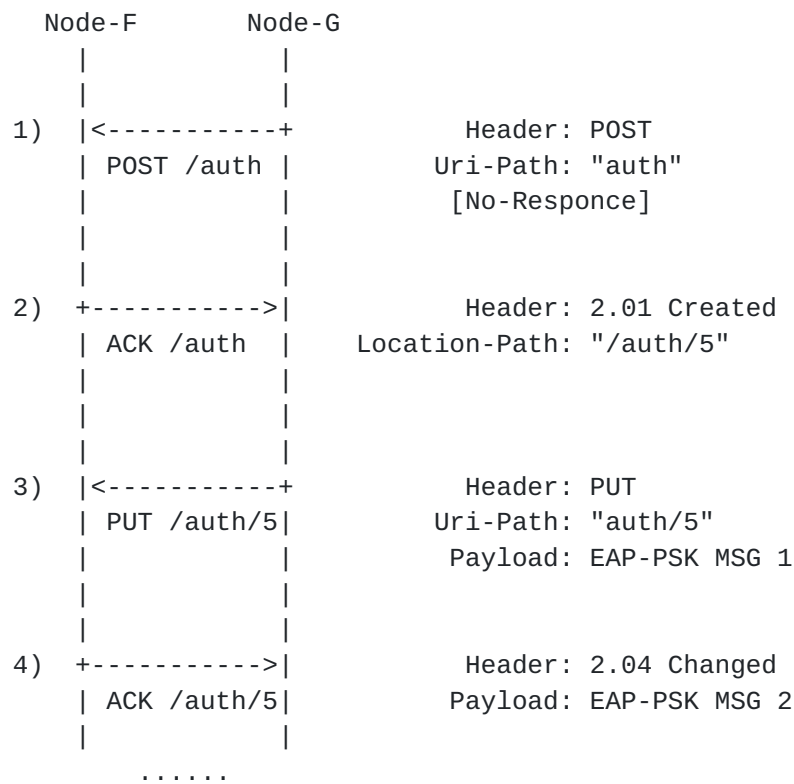
Node-F and Node-G perform mutual authentication following EAP-over-
CoAP.

Figure 9

Upon the end of the authentication phase, a Master Shared Key (MSK)
is known by the Node-F and the Node-G, and is used to generate DTLS
encryption or integrity keys. Further communications should be
encrypted/signed with the freshly derived keys.

3.3.5. Dissociation

If the Node-F wishes to deregister from the network, it could do so
by deleting the context created upon association:

```
   Node-F           Node-G
     |                 |
     |                 |
     +--------------->|          Header: POST
     | DELETE /n/n705 |       Uri-Path: "n/n705"
     |                 |
     |                 |
     |<--------------+          Header: 2.02 Deleted
     |      2.02       |
     |                 |
```
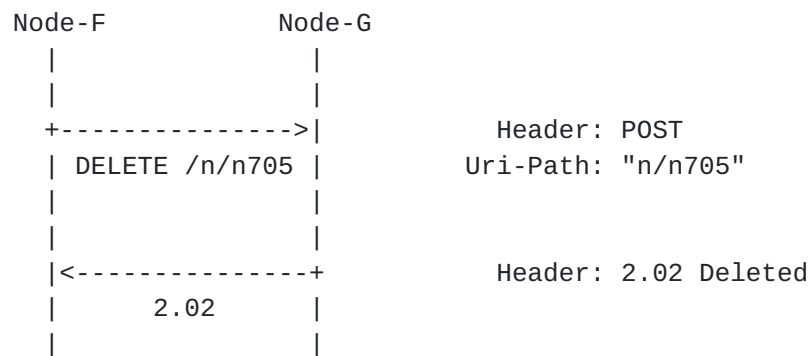
                 Node-F dissociates from the network by deleting its associated
                                    resources.

                                    Figure 10

3.4. Shared resource tree

   The Node-F and Node-G have to use any opportunity to save trafic.
   This can be handled by having a shared context on both devices, which
   is updated in an asynchronous fashion. In a RESTful approach, the
   shared context is a resource tree, synchronized with CoAP messages.
   Note, that this only concerns the control plane, responsible for
   managing the devices. The data plane is independent and can use any
   communication pattern, which fits the radio limitations.

   The shared resource tree can be structured freely, but will generally
   include the radio parameters of the Node-F and Node-G, their
   identities, authentication results, encryption/integrity preferences
   and parameters, compression methods, etc. It will can also include
   trafic shaping settings, restrictions, counters, and so forth. The
   resource tree can follow a structure defined with YANG.

   For example, for a typical OSSS installation, the following
   parameters should be specified:

   o Node-R beacon channels

   o Node-F response channel

   o Node-F response spreading factor

   o Node-F response coding rate

   o Node-F fall-back (default) channel
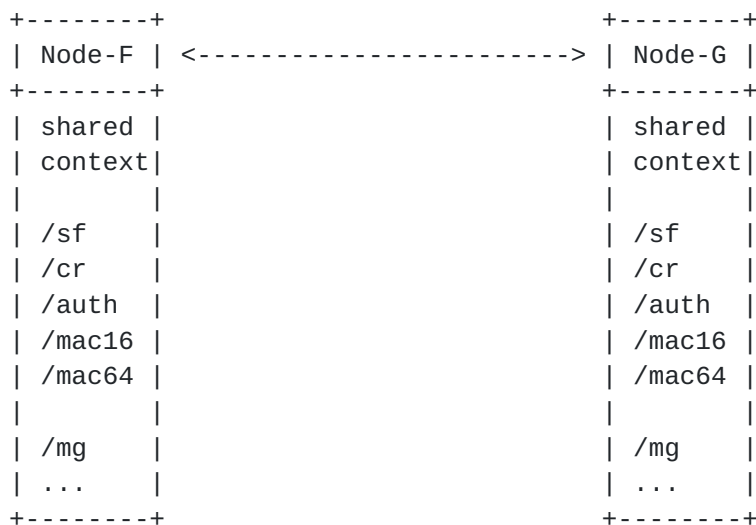
   o Node-F fall-back (default) spreading factor

o Node-F fall-back (default) coding rate

o ...

Upon authentication, the two nodes establish an authenticated
connection. Each of the resources can then be accessed in read-only,
read-write, or write-only mode. Access is performed with CoAPs GET,
PUT, POST and DELETE methods.

The most frequently accessed resource tree elements should have short
aliases, in order to have short URIs. If the management server is
independent from the application servers, using a single- or double-
character abbreviation under the root tree is recommended.
Alternatively, the use of CoMI [I-D.vanderstok-core-comi] is
recommended if YANG representation is available.

For example:

/radio/interace/lora/lora1/spreading_factor -> /sf
/radio/interace/lora/lora1/coding_rate -> /cr

```
     +--------+                              +--------+
     | Node-F | <----------------------> | Node-G |
     +--------+                              +--------+
     | shared |                              | shared |
     | context|                              | context|
     |        |                              |        |
     | /sf    |                              | /sf    |
     | /cr    |                              | /cr    |
     | /auth  |                              | /auth  |
     | /mac16 |                              | /mac16 |
     | /mac64 |                              | /mac64 |
     |        |                              |        |
     | /mg    |                              | /mg    |
     | ...    |                              | ...    |
     +--------+                              +--------+
```

Node-F and Node-G have a shared context. Upon modification (e.g. the
operator changes the spreading factor /sf of the Node-F at the Node-
G), the Node-G will update the value on the Node-F with a CoaP PUT or
        a CoAP GET OBSERVE [I-D.ietf-core-observe] message.

Figure 11

4. Acknowledgements

5. IANA Considerations

   This memo includes no request to IANA.

6. Security Considerations

   All drafts are required to have a security considerations section.
   See RFC 3552 [RFC3552] for a guide.

7. References

7.1. Normative References

   [I-D.garcia-core-security]
             Garcia-Morchon, O., Kumar, S., Keoh, S., Hummen, R., and
             R. Struik, "Security Considerations in the IP-based
             Internet of Things", draft-garcia-core-security-06 (work
             in progress), September 2013.

   [I-D.ietf-core-observe]
             Hartke, K., "Observing Resources in CoAP", draft-ietf-
             core-observe-16 (work in progress), December 2014.

   [I-D.tcs-coap-no-response-option]
             Bhattacharyya, A., Bandyopadhyay, S., and A. Pal, "CoAP
             option for no server-response", draft-tcs-coap-no-
             response-option-11 (work in progress), June 2015.

   [I-D.vanderstok-core-comi]
             Stok, P., Greevenbosch, B., Bierman, A., Schoenwaelder,
             J., and A. Sehgal, "CoAP Management Interface", draft-
             vanderstok-core-comi-06 (work in progress), February 2015.

   [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate
             Requirement Levels", BCP 14, RFC 2119, March 1997.

   [RFC7252] Shelby, Z., Hartke, K., and C. Bormann, "The Constrained
             Application Protocol (CoAP)", RFC 7252, June 2014.

7.2. Informative References

   [IEEE.802-15.4k]
              Institute of Electrical and Electronics Engineers, "Low-
              Rate Wireless Personal Area Networks (LR-WPANs) -
              Amendment 5: Physical Layer Specifications for Low Energy,
              Critical Infrastructure Monitoring Networks., IEEE
              802.15.4k", IEEE Standard 802.15.4, 2013.

   [LoRa]   Semtech, "https://web.archive.org/web/20150510011904/
              https://www.semtech.com/wireless-rf/lora.html", May 2015.

   [LTN001]   European Telecommunications Standards Institute, "Low
              Throughput Networks (LTN); Use Cases for Low Throughput
              Networks, ETSI GS LTN 001", IEEE ETSI GS LTN 001, 2014.

   [LTN003]   European Telecommunications Standards Institute, "Low
              Throughput Networks (LTN); Protocols and Interfaces, ETSI
              GS LTN 003", IEEE ETSI GS LTN 003, 2014.

   [RFC3552] Rescorla, E. and B. Korver, "Guidelines for Writing RFC
              Text on Security Considerations", BCP 72, RFC 3552, July
              2003.

   [SigFox]   SigFox, "https://web.archive.org/web/20150628225901/
              http://www.sigfox.com/en/#!/technology", June 2015.

Authors' Addresses

   Alexander Pelov (editor)
   Acklio
   2 Rue de la Chataigneraie
   Cesson-Sevigne, Bretagne 35510
   FR

   Phone: +33299127004
   Email: a@ackl.io

   Laurent Toutain (editor)
   Telecom Bretagne
   2 Rue de la Chataigneraie
   Cesson-Sevigne, Bretagne 35510
   FR

   Phone: +33299127026
   Email: laurent.toutain@telecom-bretagne.eu

Yannick Delibie (editor)
Kerlink
1 rue Jacqueline Auriol
Thorigne-Fouillard, Bretagne 35235
FR

Phone: +33299122900
Email: yannick.delibie@kerlink.fr