RELOAD/CoAP P2P Overlays for Network Coding Based Constrained Environments

Eman Al-Hawri, Noelia Correia, and Alvaro Barradas[✉]

CEOT, University of Algarve, 8005-139 Faro, Portugal
{esalhawri,ncorreia,abarra}@ualg.pt

Abstract. The Internet of Things will bring into the Internet all kinds of smart systems, which will be able to interact with each other. Therefore, applications relying on data sharing for collaboration will increase, and effective distributed solutions for data storage become necessary. This need led to the proposal of a CoAP Usage for RELOAD, a generic P2P protocol that accepts pluggable application layers (Usages). This allows P2P overlay networks to be built where constrained systems store their data and clients are able to retrieve it. Since many smart systems rely on wireless networks to communicate, where network coding can be used to reduce packet error rate, P2P overlays should be prepared to store data from network coding based networks. More specifically, encoding vectors and encoded data must be stored, and a decoding service is required. In this article, we propose a CoAP Usage extension so that network coding based constrained networks can use RELOAD/CoAP P2P distributed storage.

Keywords: P2P · RELOAD · CoAP Usage · Network coding

1 Introduction

Smart systems appear everywhere and this is expected to increase in the near future [1]. Such trend, combined with the potential for the Internet of Things (IoT) to connect tens of billions of objects to the Internet, will allow geographically distributed systems to collaborate and smart services to emerge. In such scenarios, peer-to-peer (P2P) networks may have a key role allowing federated peers to collaborate and improve their businesses.

In the context of federated constrained systems/networks some open standards become important. Within Internet Engineering Task Force (IETF), the Constrained RESTful Environments (CoRE) working group has focused on the development of Constraint Application Protocol (CoAP), a data messaging/transfer protocol providing a request/response interaction model between application endpoints [2, 3]. The "coap://" Uniform Resource Identifier (URI) scheme is used to identify CoAP resources, while "/.well-known" is defined as a default entry point for resource discovery. The Internet media type "application/link-format" is assigned for CoRE Link Format payloads [4, 5].

More recently, CoAP Usage for REsource LOcation And Discovery (RELOAD), a base protocol that provides a generic self-organizing P2P overlay network service, was proposed [6]. The use of pluggable application layers, called Usages, allows RELOAD to fit any purpose [7]. In a RELOAD/CoAP architecture the proxy nodes form a

distributed P2P overlay network to announce resources, allowing clients to discover them. More specifically, the overlay can be used: as a lookup service, to store existing resources, and as cache for data.

Many applications (e.g. industrial, environmental) require smart systems to communicate using wireless constrained networks. Such constrained environments usually have energy efficiency and end-to-end packet error rate concerns, which are competing goals (e.g. a packet may be sent through multiple paths to reduce error rate but this increases energy consumption). An elegant way of achieving a balance between these two goals is to use Network Coding (NC) [8]. When NC is used, nodes can perform some linear combination on the received packets and route encoded packets. At the destination, the decoding process can be applied to recover the original packets. Since linear combinations of packets are being done, the sink will be able to recover lost packets from a sufficient number of other arriving packets.

In this article, we propose an extension of CoAP Usage so that NC based constrained networks can benefit from RELOAD/CoAP P2P distributed storage. That is, although packets will travel from sources toward sinks/gateways, their final destination will be the P2P overlay where storage is done. When compared with the work in [6] a new data Kind structure is proposed for encoded data and encoding vectors to be stored. For the decoding process, required when original data packets are not received, a decoding service is required at the P2P overlay network. Such architecture allows constrained networks to reduce packet error rate while benefiting from an effective distributed solution for data storage. This work serves as a basis for future proposal of algorithms that determine the most effective routing trees, for packet forwarding toward sink/gateway nodes, and best placement of coding nodes.

The remainder of this article is organized as follows. Section 2 explains how smart systems can benefit from our proposal. Section 3 discusses network coding, while Sect. 4 presents the proposed architecture and extension of CoAP Usage. Finally, in Sect. 5 some conclusions are drawn.

2 Relationship to Smart Systems

A smart system is a collection of sensing, actuation, and controlling functions, allowing situations to be analyzed and decision making. Such systems mainly rely on sensor information sources, which might be geographically distributed [1]. With the increasing presence of smart systems in the Internet, applications relying on data sharing will increase and effective solutions for storage/discovery of data become necessary.

Although Cloud storage solutions have emerged, extra storage space always requires extra physical disks and processing, meaning that connecting hundreds of millions of sensors to the Cloud will be extremely demanding. Besides this, such places are known as storage places meaning that they are exposed to attacks. P2P approaches can overcome both these issues because participating peers/proxies also contribute to storage and processing, peers communicate directly posing no burden on a specific set of servers, and bandwidth bottlenecks are avoided. Encrypted communication tunnels can also be built between peers.

CoAP is a Web application transfer protocol that was developed for constrained systems to provide RESTful services. Although devices will be heterogeneous regarding their radio layer, CoAP is expected to become a common application layer protocol. For this reason, the CoAP Usage for RELOAD has been proposed, avoiding the use of centralized servers. The use of P2P overlays is particularly useful for power-constrained systems that are expected to explore sleeping modes for higher energy saving. This is so because clients can retrieve information directly from the overlay without performing either a connection to the proxy or the power-constrained system. Thus, RELOAD/CoAP P2P overlays are an effective response to many of the challenges that smart systems face.

Many smart systems rely on wireless constrained networks to communicate. Therefore, it makes sense to prepare P2P overlays to store data arriving from NC based wireless networks. More specifically, encoding vectors and encoded data must be stored, and a decoding service supplied. NC improves packet delivery rate and provides some degree of protection because an attacker cannot control the outcome of the decoding process without knowing all other coded packets [9]. These issues are critical for smart systems.

3 Network Coding

In network coding the nodes take several packets and combine them together for transmission, instead of simply relaying the data packets they receive [9]. This approach can be used to improve the throughput, efficiency, scalability, resilience to attacks and eavesdropping in networks. In the case of sensor Networks it can be used to reduce packet error rate [8].

When linear global coding is used the encoded packets are transmitted along the network, where encoding can be performed recursively. When packets reach their destination the decoding process can be applied on these packets for the original ones to be extracted. In the following discussion it is assumed that:

- Each packet consists of l bits, small packets are padded with 0's;
- s consecutive bits form a symbol over finite field F_{2^s};
- Each packet is a vector of $\frac{l}{s}$ symbols;
- The resulting linearly combined packets also have length l.

3.1 Encoding

Let us assume the original packets $(k^1, \ldots, k^n)$. An encoded packet ω will be a vector of $\frac{l}{s}$ symbols, $\omega = [\omega_1, \ldots, \omega_u, \ldots, \omega_{\frac{l}{s}}]$, where the uth symbol will be:

$$\omega_u = \sum\nolimits_{i=1}^{n} \alpha_i k_u^i \tag{1}$$

where coefficients $\alpha_1, \ldots, \alpha_n$ are in F_{2^s}, while k_u^i and ω_u are the uth symbols of k^i and ω, respectively. The encoding vector $\alpha = (\alpha_1, \ldots, \alpha_n)$ is used by recipients to decode the data. Since the encoding process can occur recursively through the network, any intermediate node can receive the encoded packets $\omega^1, \ldots, \omega^m$ and encode them again to generate a linear combination of these packets. Although a different set of coefficients is used to make the new linear combinations, $h_1, \ldots, h_m$, the corresponding encoding vector is not simply $h = (h_1, \ldots, h_m)$ since coefficients are with respect to the original packets $k^1, \ldots, k^n$. That is, the re-encoded packet $\tilde{\omega}$ ends up being tagged with an encoding vector whose elements are:

$$\tilde{\alpha}_i = \sum_{j=1}^{m} h_j \alpha_i^j \qquad (2)$$

3.2   Decoding

When a node receives m arrivals $(\alpha^1, \omega^1), \ldots, (\alpha^m, \omega^m)$, it needs to retrieve the original packets. This can be achieved by solving the system:

$$\omega^j = \sum_{i=1}^{n} \alpha_i^j k^i, \forall_j = 1, \ldots, m \qquad (3)$$

This linear system has m equations that need to be solved to extract the n unknowns (unknowns are k^i). To perform this, it requires that $m \geq n$ to be able to recover the original data, which means that the number of received packets should be at least equal to the number of the original packets.

4   Proposed P2P Architecture

4.1   Motivation

Network coding is an elegant way of finding a balance between energy efficiency and end-to-end packet error rate, which are competing goals. However, network coding based constrained wireless networks cannot benefit from RELOAD/CoAP P2P storage. For this reason we propose to extend CoAP Usage data Kind. A decoding service, to be provided by the P2P overlay, is required.

4.2   Overall Architecture

At the wireless section the nodes are organized in a way that packet flowing toward a sink/proxy node is ensured. Figure 1 shows the proposed architecture. Depending on factors like location, energy and functionality, nodes can be of type:

- Sensing: Data sources that forward their packets according to some routing table. Nodes can turn to sleep mode to reduce energy consumption.
- Encoding: Node able to perform encoding. Besides forwarding original packets, extra encoded packets are created by linearly combining received packets and packets overheard from neighbors. These might be required to have more energy, memory and processing capabilities than others.
- Relay: Work just as relays and forward any received packets, either original or encoded, towards the sink. Perform no encoding/decoding operations.
- Sink: Destination of data. It is connected to the RELOAD overlay network where packets can be stored and fetched by others.

It is assumed that encoding nodes propagate linear combinations of original packets and packets overheard from neighbors, besides original packets. Sinks/gateways store original and encoded packets at the overlay.
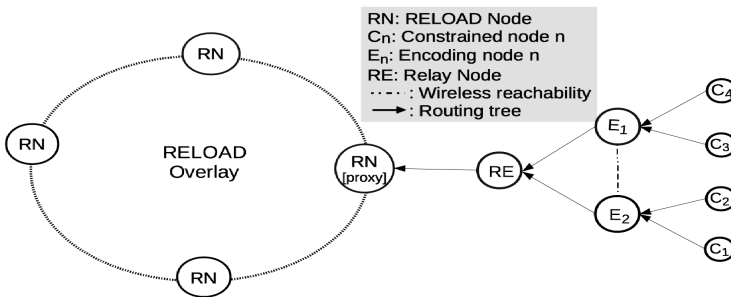

Fig. 1. Overall architecture.

4.3   Storing Data and Standard Extensions Needed

4.3.1   Data Kinds

The data Kinds defined by CoAP Usage include the CoAP-REGISTRATION, to announce resources available, and CoAP-CACHING for the storage of sensor measurements [6]. Although the CoAP protocol itself supports the caching of sensor measurements to reduce the response time and network bandwidth consumption for future requests (see [2]), the additional caching mechanism of CoAP Usage allows such data to be stored in the P2P overlay, improving even more the response time and bandwidth use because nodes can go to sleep mode for longer periods of time.

Regarding CoAP-CACHING, the possibility of storing proxy data and sensor data directly are both considered. Their resource names will be their CoAP URIs (e.g. "coap://overlay-1.com/proxy-1/", "coap://overlay-1.com/proxy-1/sensor-temp"). That is, the proxy data structure supports data from multiple sensor nodes forwarding their data to a proxy, while the sensor data structure stores measurements of a specific sensor

node. Listing 1 shows the data stored for proxy with NodeID "9996172" and URI "coap://overlay-1.com/proxy-1", where "mt" is the measurement time of data, "ttl" is the time-to-live, and "v" is the measurement value.

Listing 1. Storage of proxy data.

```
Resource-ID=h("coap://overlay-1.com/proxy-1/")
KEY=9996172,
VALUE=[
  </sensor-1/>; {mt=100000;ttl=10000;v=38};
           {mt=100055;ttl=456990;v=42};
           {mt=134000;ttl=234000;v=30}
  </sensor-2/>; {mt=100000;ttl=10000;v=40};
                        {mt=400000;ttl=17000;v=25}]
```

The "h(…)" is the hash over the URI, because the key space for storage in the overlay must be numeric. Listing 2 shows a direct storage of data from sensor with URI "coap://overlay-1.com/proxy-1/sensor-1". Proxy with NodeID "9996172" is the one responsible for the sensor-1.

Listing 2. Storage of sensor data.

```
Resource-ID=h("coap://overlay-1.com/proxy-1/sensor-1")
KEY=9996172,
VALUE=[
  {mt=100000;ttl=10000;v=38};
  {mt=100055;ttl=456990;v=42};
  {mt=134000;ttl=234000;v=30}]
```

4.3.2   Requirements Regarding Network Coding

Network coding based constrained networks forward original and encoded packets toward the proxy participating as peers in the P2P overlay network. Therefore, for network coding based constrained networks to benefit from such a distributed storage, the data Kind must also allow the storage of:

- *Encoding vectors:* Required for the decoding. For deterministic network coding these may be stored at the overlay at peer registration time (done once). A proxy stores encoding vectors of sensors for which it is responsible for.
- *Encoded packets:* Since decoding may not be done if a sufficient number of encoded packets has not arrived, storage of encoded packets is needed.

Therefore, the current CoAP Usage data Kind must be extended. Listing 3 shows an example when using such extended data Kind. The first value regards to the current data Kind structure and second value includes predefined encoding vectors and arriving encoded data.

Listing 3. Extended sensor data.

```
Resource-ID=h("coap://overlay-1.com/proxy-1/")
KEY=9996172,
VALUE=[
  </sensor-1/>; {mt=100000;ttl=10000;v=38};
        {mt=100055;ttl=456990;v=42};
        {mt=134000;ttl=234000;v=30}
  </sensor-2/>; {mt=100000;ttl=10000;v=40};
        {mt=400000;ttl=17000;v=25}]
ENCODING=[
  {enVector1;enVector2;...};
  {1stcodedData;2ndcodedData;...}]
```

4.3.3   Storing/Fetching Packets of Network Coding Based Constrained Network

Figure 2 illustrates how storage of packets would be done. In this example, encoding nodes E_1, E_2 receive packets from their children, namely E_1 receives k_1, k_2, and E_2 receives k_3, k_4. Moreover, these encoding nodes can hear each other. In this case, the linear combinations of packets from children plus packets they have heard is done $(k_1 + k_2 + k_3 + k_4)$. When a single packet gets lost, it can be retrieved when the decoding is performed. Please note that encoding can be performed at different areas of the wireless section, and nodes may forward packets toward different proxies/gateways. Therefore, decoding might require encoded packets stored by different proxies. Therefore, a relevant research issue is how to build routing trees that increase the effectiveness of encoding. Also, other encoding scenarios are possible (e.g. E_1 and E_2 could forward some overheard packets so that multiple packets can be recovered) with different impacts on bandwidth usage.

When a client needs to fetch sensor data from the overlay network, such data can be directly available (original data) or there might be a need to perform the decoding process, using stored encoded data and encoding vectors, in order to extract the original data. This decoding would be accomplished by RELOAD nodes able to provide such
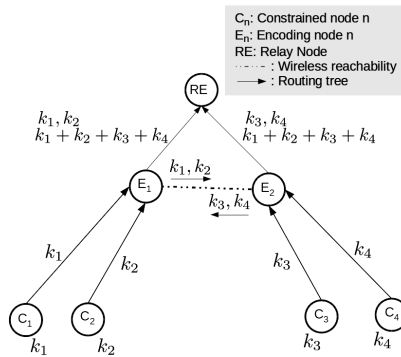

Fig. 2. Example of network coding at constrained network.

decoding service. Figure 3 illustrates a scenario where the decoding service is requested. A fetch request by some client may involve fetching original data and encoded data, which also requires fetching the encoding vectors, so that the highest amount of sensor measurements is returned to the client.

Fig. 3. Decoding at RELOAD/CoAP P2P overlay network.

Since peers face the problem of finding the peers providing the decoding service, the *Recursive Distributed Rendezvous (*ReDiR) service discovery mechanism can be used [10]. ReDiR ensures that the load is distributed among the nodes providing the service.

4.3.4 CoAP Option

CoAP allows options to be included in a message [2]. Each option instance specifies the option number, length of the option value, and the option value itself. For the proxy to be able to differentiate the type of payload at CoAP packets, which may be of original or encoded type, the CoAP option numbers registry should include an entry referring to encoded data. A CoAP packet carrying encoded data must carry such option.

5   Conclusions and Future Work

This article discusses how to federate network coding based constrained networks through RELOAD and CoAP Usage. The proposal is for network coding based constrained networks to be able to store encoding vectors and encoded data at the P2P overlay network, which requires an extension of CoAP Usage data Kind. This provides a scalable and efficient way of discovering cached sensor data of geographically dispersed sensor networks, allowing large scale applications to emerge, while taking advantage of network coding at the wireless section. This architecture serves as a basis

for future proposal of algorithms that determine the most effective routing trees and placement of coding nodes.

References

1. Jones, A., Subrahmanian, E., Hamins, A., Grant, C.: Humans' critical role in smart systems: a smart firefighting example. IEEE Internet Comput. **19**(3), 28–31 (2015)
2. Shelby, Z., et al.: The Constrained Application Protocol (CoAP). RFC 7252 (2014)
3. Correia, N., Schütz, G., Mazayev, A., Martins, J., Barradas, A.: An energy aware resource design model for constrained networks. IEEE Commun. Lett. **20**(8), 1631–1634 (2016)
4. Shelby, Z.: Constrained RESTful environments (CoRE) link format. RFC 6690 (2012)
5. Nottingham, M., Hammer-Lahav, E.: Defining well-known Uniform Resource Identifiers (URIs). RFC 5785 (2010)
6. Jimenez, J., et al.: A Constrained Application Protocol (CoAP) usage for REsource LOcation And Discovery (RELOAD). RFC 7650 (2015)
7. Jennings, C., et al.: REsource LOcation And Discovery (RELOAD) base protocol. RFC 6940 (2014)
8. Keller, L., Atsan, E., Argyraki, K., Fragouli, C.: SenseCode: network coding for reliable networks. ACM Trans. Sens. Netw. **9**(2), 25 (2013)
9. Fragouli, C., Le Boudec, J., Widmer, J.: Network coding: an instant primer. ACM SIGCOMM Comput. Commun. Rev. **36**(1), 63–68 (2006)
10. Maenpaa, J., Camarillo, G.: Service discovery usage for REsource LOcation And Discovery (RELOAD). RFC 7374 (2014)