

CoHaRT: Deterministic Transmission of Large Data Amounts using CoAP and Kad

Jan Skodzik, Peter Danielis, Vlado Altmann, Bjoern Konieczek,
Eike Bjoern Schweissguth, Frank Golatowski, Dirk Timmermann
University of Rostock
Institute of Applied Microelectronics and Computer Engineering
18051 Rostock, Germany, Tel./Fax: +49 381 498-7284 / -1187251
Email: jan.skodzik@uni-rostock.de

Abstract—In Industrial Ethernet environments, usually only small amounts of data are transmitted. These transmissions must be deterministic to meet hard real-time requirements. But prospectively, also the deterministic transmission of high amounts of data and data streams is required. In this paper, a system based on the modified Peer-to-Peer network Kad called HaRTKad combined with CoAP is presented to fulfill these requirements. The Kad network bases on UDP. Because of the limited size of UDP packets, only small amounts of data can be transmitted and interpreted. Therefore, CoAP was chosen as protocol on top of HaRTKad. A highlight of CoAP is its applicability on constrained devices like embedded systems in automation environments. It has a very low overhead, high simplicity, and is very effective and well suited for the intended domain. CoAP enables the transmission and interpretation of big data amounts using the blockwise transfer. Thereby, HaRTKad ensures the deterministic data transmission of the packets. The combination of CoAP and HaRTKad is called CoHaRT. Besides the described concept of CoHaRT, real measurements from a running prototype are presented to demonstrate the performance of the data transmission.

I. INTRODUCTION

Our environment is more and more affected by the rising number of connected devices. This development is also referred to as Internet of Things (IoT). Devices and machines begin to interact with growing intelligence and autonomously without human control at the same time. The communication between the devices is denoted as machine-to-machine (M2M) communication. The Constrained Application Protocol (CoAP) is a new standard, which presents a transfer protocol for constrained devices like embedded systems and enables a highly effective M2M communication [1]. These developments become more and more relevant also in the domain of automation, especially industrial automation. Existing solutions in the industrial automation use industrial Ethernet (IE) to allow for a total horizontal and vertical integration [2]. IE replaces the fieldbuses by using standard Ethernet technology to connect devices. But there are many unsolved issues like the avoidance of proprietary software and hardware and the most commonly used master/slave or client-server approaches, which often lack scalability, flexibility, self-organization, or reliability [3].

A previous development called HaRTKad is used to enable a deterministic data transmission for hard real-time scenarios based on the Peer-to-Peer (P2P) network Kad [4]. The resulting

system avoids the issues mentioned above as they are intrinsic features of a P2P technology. HaRTKad realizes the hard real-time behavior without any central instance and every node acts nearly autonomously by using a Time Division Multiplexed Access (TDMA) approach. Via unique IDs, the nodes determine their own time slots without a coordinating instance. Also, the need for the synchronization of all participants to realize the TDMA approach has been realized in a scalable manner by HaRTKad [5], [6].

However, the deterministic transmission of big amounts of data is not possible yet. Only data sizes of about 1000 bytes are possible to transmit as they fit into the payload of standard UDP packets. It is possible to send more data, but the problem is the reassembly of the data fragments. Prospectively, larger data amounts have to be transmitted reliably in automation environments. For example, in the case of data streams in cars, a deterministic behavior is required to ensure the timely data transmission [7]. The system should not be restricted by data sizes or different number of participants. It should scale in terms of data sizes and number of nodes. The presented system CoHaRT (CoAP HaRTKad) solves the problem of transmitting and interpreting big amounts of data by reassembling data fragments from the UDP packets within HaRTKad. CoAP as a light-weighted approach for resource usage offers the possibility of sending data blocks, which can be interpreted as one data set later. This approach has been realized by a prototype setup and measurements proof the high effectiveness of this approach.

The main contributions are:

- Description of the CoHaRT concept comprising CoAP and the hard real-time Kad-based approach HaRTKad.
- Presentation of a realized CoHaRT prototype.
- Performance analysis with a prototype setup to evaluate CoHaRT.

The remainder of this paper is organized as follow: In Section II, the related work is presented and previous works are briefly described. In the following Section III, the basics of Kad, CoAP, and the main concept are described before the prototype is presented in Section IV. The results and the evaluation of a prototype setup are presented in Section V. In terms of reliability, the usage of Erasure Resilient Codes (ERCs) is

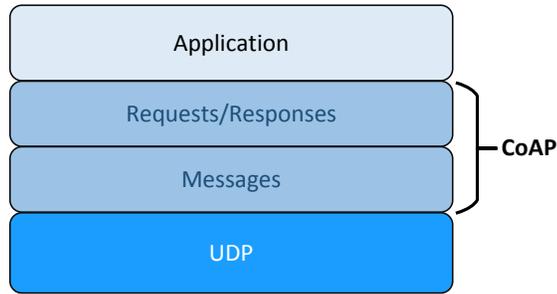


Fig. 1. Protocol stack of CoAP [1]

discussed in Section VI. Finally, the paper concludes in Section VII.

II. RELATED WORK

In IE environments, enabling a deterministic data transmission for hard real-time systems is mostly done via proprietary hardware or special protocols. Usually, a client-server or master/slave approach is chosen to manage the media access by performing a TDMA-based approach [3], [8]. De facto, there is no direct implementation, which also supports a scalable system in terms of the number of devices and a higher amount of data. In [4], a system called HaRTKad is presented, which is suitable for hard real-time environments without a central instance using the P2P network Kad. HaRTKad uses UDP packets to transmit data and is not able to reassemble the data fragments from the UDP payload in a standardized way. This problem can be addressed by higher level protocols like CoAP. Currently, CoAP is most commonly used in heavily resource constrained environments like sensor networks. However, recent research projects demonstrate the applicability of CoAP in medical or industrial applications. A deterministic timing behaviour of the communication between the interconnected devices is crucial in these areas. In [9], several medical sensors are interconnected in a wireless manner utilizing CoAP and 802.15.4. Although 802.15.4 offers reserved time slots to enable real-time communication, no hard real-time can be achieved due to the susceptibility of wireless communication to interferences. [10] specifies a system architecture where CoAP nodes are interconnected by multiple 802.15.4e subnets and a high speed backbone. The medium access is controlled by TDMA techniques. The susceptibility to interferences is reduced by utilizing a channel hopping algorithm. However, a centralized instance is needed to guarantee optimized traffic flow, which represents a single point of failure.

CoHaRT instead bases on wired communication. As it utilizes P2P technology, it does not rely on any central instance. Additionally, the combination of HaRTKad and CoAP enables a scalable hard real-time data transmission.

III. BASICS

HaRTKad: First, HaRTKad will be briefly described as it is the middleware ensuring the deterministic data transmission [4]. HaRTKad bases on the structured decentralized DHT-based P2P network Kad. Within Kad, every node has a unique

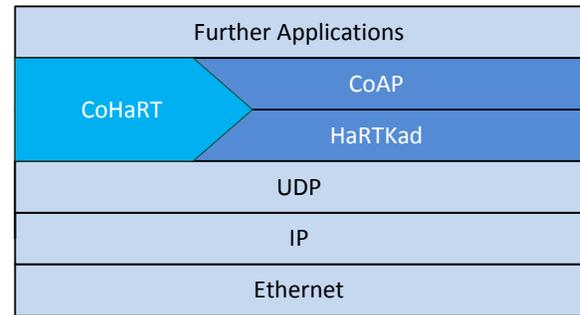


Fig. 2. Protocol stack of CoHaRT

ID. It is possible to create a relation between a node and data by using a distance metric. Kad uses the XOR metric to determine the distance between a node and data. Hash values, e.g., calculated by the Message Digest 5 algorithm are used as arguments for the hash function. If the distance is smaller than a given search tolerance a node becomes responsible for the data. Now, not only data is assigned to nodes by using the ID but also the time slots are. Every node has an assigned time slot, in which it is allowed to use the Ethernet medium to transmit UDP messages. This approach guarantees a deterministic data transmission behavior. As buffer overflows due to high congestion are avoided, no packets have to be dropped. An additional benefit is the searching for devices in the Kad network. In only $\log_2(\#Nodes)$ steps in the Kad network, the desired instance can be found only by storing a subset of routing information in a nodes memory.

CoAP: CoAP uses HaRTKad as middleware to solve the problem of sending and interpreting big data amounts. CoAP is a RESTful web transfer protocol specially designed for the use in resource constrained environments. It bases on the client-server principle whereby a client sends a request to a server to either retrieve data from or send data to the server. The protocol stack of CoAP is depicted in Figure 1.

CoAP usually bases on UDP communication but is not limited to it [1]. Furthermore, the use of TCP as transport protocol is omitted because it is not real-time capable as it automatically resends lost data packets. Additionally, TCP uses a streaming-based approach to transfer large amounts of data. In combination with HaRTKad, the transmission must be interruptable even for a longer time, which is only possible with extensive modifications of the TCP stack and would lead to a proprietary solution. As UDP only transmits single packets, the complete data transmission can be halted at any time. But UDP does not allow the reassembly of huge data sets. Consequently, CoAP is used.

The communication on top of UDP bases on two layers. First, there is the message layer. Inside the message layer, messages can be declared as confirmable (CON), non-confirmable (NON), acknowledgement (ACK), or reset message (RST). A reliable communication is achieved by CON messages similar to a TCP transmission. After having sent a CON message, an ACK message is expected to confirm the successful transmission.

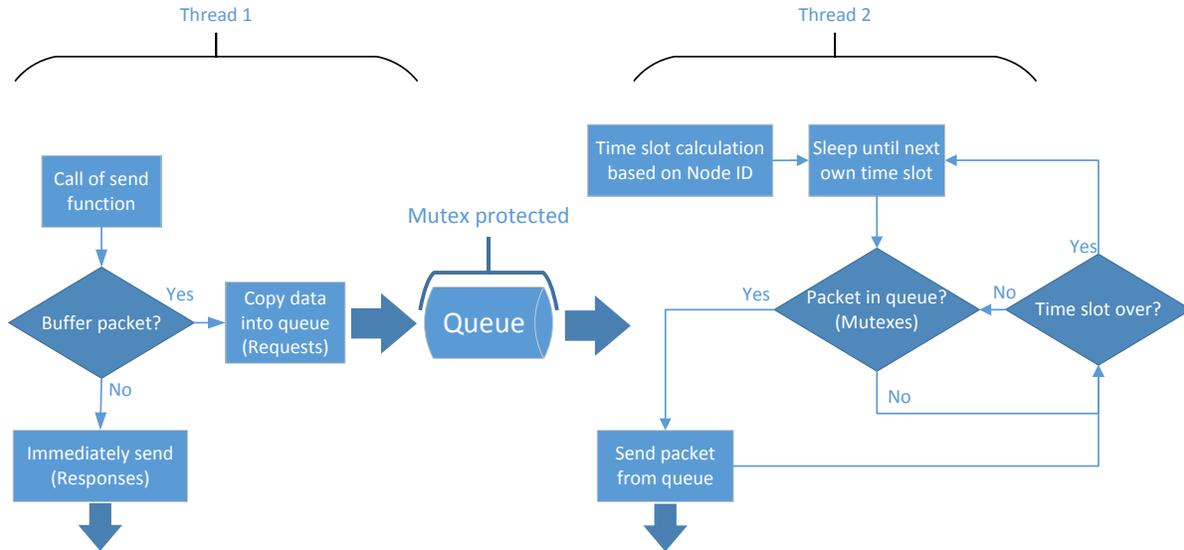


Fig. 3. Buffering the data using a queue

The assigning of the ACK message to a CON message is done by a message ID. If the sender of a CON message does not get the acknowledgement within a timeout, it sends the message again. In a hard real-time system, this behavior is not suitable in two regards. First, a retransmission violates the hard real-time condition as it is not a deterministic behavior. Second, it is usually not desired to receive a retransmitted message as the value of the information is zero if the information is too old. It only causes unnecessary network traffic. To avoid the repeated transmission of the information, NON messages are used as they do not expect an ACK message as confirmation. RST messages are used as a response if an instance cannot interpret a request. The request/response layer is above the message layer to enable the request/response principle between the network participants. In this layer, the basic methods are performed, e.g., a GET or PUT method.

CoHaRT: The resulting CoHaRT system is built as the combination of HaRTKad and CoAP and utilizes their benefits. Figure 2 shows the basic protocol stack of CoHaRT. CoHaRT completely bases on standard Ethernet like HaRTKad does.

The stack is nearly the same as it is for HaRTKad. But on top of HaRTKad, the CoAP layer is added to enable the transmission of big data. CoHaRT can also be a middleware for other applications like HaRTKad. In such a way, CoHaRT is able to deliver data streams to any endpoint in the Kad-based network. Additionally, from another point of view functionality of CoAP becomes hard real-time capable. However, not the complete CoAP specification was implemented as it was sufficient to implement the functionality needed for the data transmission to determine the performance.

A. Using the CoAP blockwise transfer

The reassembly of the data fragments from the UDP packets to interpret them as huge data is currently not possible. As a solution, CoAP is used. Since CoAP is designed for low power

lossy networks with low bandwidth, it offers a simple stop and wait approach called blockwise transfer [11] to transmit large amounts of payload. Hereby, the server divides the payload into multiple blocks of data. Upon a request, the server only sends the first block of the payload to the client. To inform the client that only a small fraction of the actual payload was sent, the server sets the block option in the CoAP header of the response. This option indicates the number of the sent block (Block), the block size (Size), and whether there are more blocks to follow (M). When a client notices a block option in a response, it buffers the received data and sends a request for the next data block also using the block option. As specified in [11], the block size can vary between 16 bytes and 1024 bytes per block. We use a block size of 1024 bytes to put as much payload as possible into a single UDP packet while ensuring that the MTU of UDP is not exceeded. CoAP blockwise transfer describes two types of block options: block1 and block2. The usage of the block options depends on the scenario. If a user or device wants to perform a PUT request to store data on another instance, it has to use the block1 option. Contrary, the block2 option must be used when data is retrieved from a remote device through a GET request. In our test application, every node acts as client and server on the CoAP level. In order to allow all nodes to store data on and retrieve data from remote nodes, we implemented both block option types. Figure 4 shows the structure of a CoAP packet using the block option. Furthermore, the header includes several options that are used to address a certain resource on a certain device. The Uri-Host option contains the 32 bit IPv4 address of the device hosting the desired resource. The port of the remote device, through which a connection can be established, is given by the Uri-Port option as a 16 bit unsigned integer. The default CoAP port is 5683. The Uri-Path option contains the path to the desired resource on the remote device. The length of this option can vary between 0 and 255 bytes. To allow

Byte 1		Byte 2		Byte 3		Byte 4	
Ver	T	TKL: 4		Code		Message ID	
Token							
3		4		Uri-Host: 32Bit Host IP Address			
7		2		Uri-Port: 5683			
11		11		Uri-Path: client.test			
13		1	14		#Block	M	Size 11111111
Payload ...							

Fig. 4. CoAP header including the block option

longer paths, multiple Uri-Path options can be used whereby each option contains one segment of the path. However, in our experiment we used a single Uri-Path option with the 11 byte string value "client.test".

IV. THE PROTOTYPE

The used prototype bases on the ZedBoard [12]. It is an embedded system containing an ARM CPU running at 667 MHz. FreeRTOS is used as real-time operating system [13]. Within FreeRTOS, lwIP is used as TCP/IP stack, which enables the UDP communication [14]. This results in a platform with hard real-time capabilities. The performance has to be profiled by analyzing the performed interactions. First performance results of HaRTKad are presented in [4]. These results are generated from an interaction within one time slot. As we want to transmit bigger data over more than one time slot, it becomes necessary to implement a mechanism, which controls the sending of the UDP packets. The transmission of data over several time slots is depicted in Figure 3. The trigger for sending data can be carried out by different threads of an application. The first decision, which has to be made, is if the new packet is a request or response.

If it is a response, the packet can be sent immediately as it is an answer to a request from another instance, which occupies the current time slot and expects an answer within this time slot. However, requests are initial inquiries of a node, which have to be sent only in the own time slot. Therefore, these requests must be copied into the sending queue as an immediate transmission is not possible. This procedure is performed by thread 1 in Figure 3. Thread 2 describes the readout of the queue depending on the own time slot. First of all, the own time slot is calculated depending on the own unique ID (see [4] for more details). The instance waits until the own time slot is reached. If the own time slot is reached thread 2 checks if there is a packet in the queue to be sent. An existing packet in the queue is sent immediately and deleted from the queue. Afterwards, it is checked if the own time slot is over. If there is enough time left another potential packet is transmitted. The process of checking the queue for other packets in the time slot and accessing the queue is done by blocking mutexes. The blocking mutexes allow for waiting for an event until a given timeout and do not stress the CPU with unnecessary load. When the time slot is over, thread 2 sleeps until the next cycle and the own time slot is active again. The advantage of

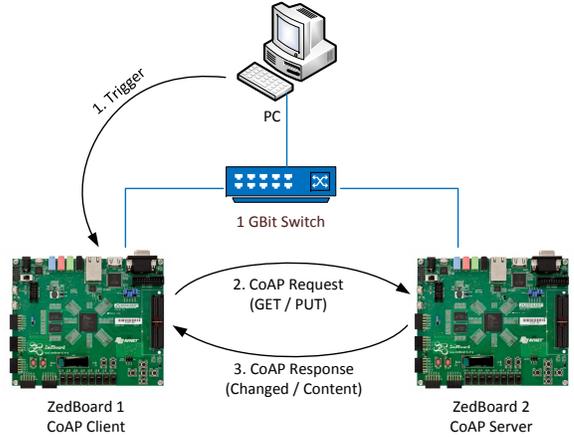


Fig. 5. Prototype setup

this method is that polling is avoided. The complete process is performed periodically and allows the transmission of big amounts of data over several time slots.

V. RESULTS AND EVALUATION

A setup of two CoHaRT prototypes, a 1 GBit switch, and a triggering PC have been used to determine the performance of CoHaRT. The triggering PC sends the commands to an instance performing the prototype. Now, the prototype starts with the interaction. The channel utilization is considered as the main criterion for the performance and is therefore determined for all test scenarios as this also indicates the maximum allowed number of devices. There are three different scenarios, which are presented and evaluated. The prototype setup is depicted in Figure 5.

A. Scenario 1: Dummy UDP packets

First of all, it was tested how efficient the queue mechanism performs. Dummy packages with a payload size of 1000 bytes have been sent from one ZedBoard to another. These results are independent from CoAP. In Figure 6, it can be seen that the utilization nearly constantly achieves 45%. The utilization is independent from the number of time slots and also from the size of the time slots. Only if a very small size for the time slots is chosen the utilization slightly increases. This happens due to the fact that a transmission is started in the own time slot but proceeds in the following time slot. As a result, a higher utilization is achieved by violating the time slot borders. The results proof the high performance of the queuing mechanism. The system also containing the HaRTKad implementation and the Ethernet part is not a limiting factor in the test.

In this case, only two nodes are allowed to communicate in parallel per time slot if we use 1 GBit connections. As the worst case, we assume that we do not have any knowledge about the network topology. Therefore, we must consider that there could be a point in the system, to which all the traffic must be passed. Thus, we are not allowed to increase the time utilization and therefore the time slot utilization higher than the bandwidth of our medium.

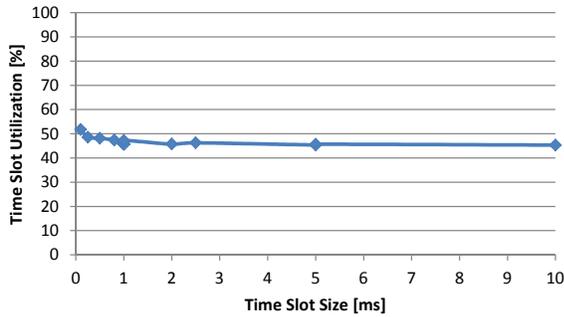


Fig. 6. Time slot utilization of the queue mechanism

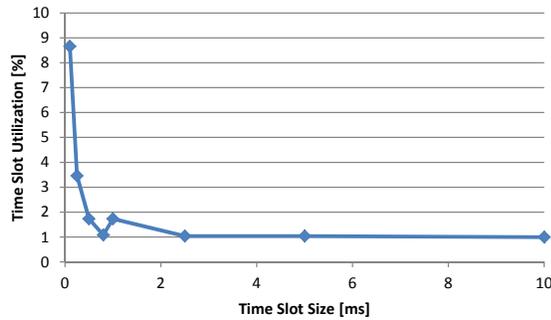


Fig. 7. Time slot utilization using CoAP without block options

B. Scenario 2: Using the CoAP protocol

In the second test scenario, the data is transmitted from one CoAP client to a CoAP server. Both instances are executed on the ZedBoards. However, both a CoAP client and server can exist on one device and therefore this does not represent a disadvantage. The number of time slots has been set to 10. Again, the size of the time slots has been varied. Only independent data, which fits into the payload of a single UDP packet, has been used. Therefore, only plain CoAP without the blockwise transfer has been used to interact between the devices. The result is apparent from Figure 7.

The utilization of the time slots approaches 1% with increasing time slot length. Thereby, it increases for smaller time slot lengths. This behavior was expected as the response to a request does not fit into the same slot for very small time slots. The time slot utilization increases with bigger time slots and approaches the value of 1%. In contrast, for small time slots the utilization increases significantly. This was expected due to the fact that the responses do not fit into the same time slot like the request. But for overall performance determination, the request and response are considered and thus leads to an increased calculated time slot utilization. It also has to be considered that the time for searching for an interaction partner is not included in the timings. They should be performed only once within one time slot. Typical values are 550 μ s up to a few milliseconds [4].

When using small time slots, it could happen that responses do not fit into the same time slot like the corresponding request. With the initial answer strategy, they are transmitted directly

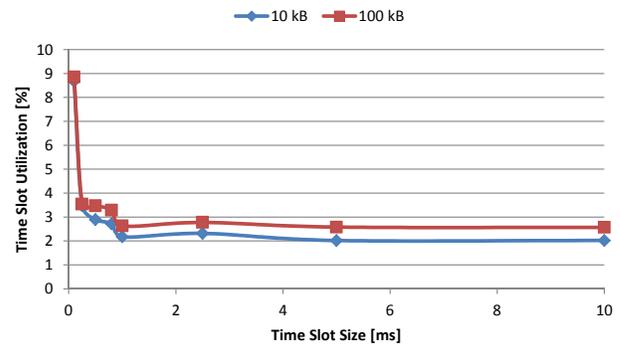


Fig. 8. Time slot utilization using CoAP with block options

from the requested communication partner. For example, a CoAP client performs a PUT operation on a CoAP server. First, it sends a request within its time slot. The CoAP server receives the request and generates a response. As the time slot of the CoAP client is very small, the response does not fit into the same time slot. Therefore, the answer could be transmitted in a time slot of a participant, which is not involved in the interaction process of the other two devices. This makes it hard to control the communication. A solution could be to also send responses via the queue. In the example, the CoAP server would wait for his own time slot and then transmits the response to the CoAP client.

For big time slots, this effect can be easily avoided. If the instance checks whether there is enough time left for a message to be sent it already adds an offset, which represents an estimated value for receiving the response. Now, the response should be transmitted in the same time slot. A queuing of the response would result in a very bad behavior as it is very ineffective for big time slots because an instance has to wait for the response for a long time.

C. Scenario 3: Using CoAP blockwise transfer

The last scenario also comprises the CoAP blockwise transfer by using the block option 1. The result is depicted in Figure 8. With an increasing number of time slots and therefore bigger cycle time, the effective data rate per node and per cycle is reduced.

However, the data rate per time slot remains stable. Furthermore, in Figure 8 it can be seen that the achieved time slot utilization can be increased compared to the second scenario without CoAP blockwise transfer. Two file sizes have been transmitted. First, files of the size of 10 KB and secondly files of the size of 100 KB have been transmitted by performing PUT operations between a CoAP client and server. Compared to the second scenario, the time slot utilization could be increased by up to the factor of 2 by transmitting 10 KB files. With a file size of 100 KB, the time slot utilization can be slightly increased compared to the 10 KB file size. This is due to the fact that only the first packet for the file transmission needs to be created. For the following parts, the same packet can be re-used with minor changes in the CoAP header and the exchange

ATTRIBUTE	SZENARIO 1	SZENARIO 2	SZENARIO 3
Description	UDP dummy packets	CoAP without block option	CoAP (10 KB/ 100 KB) with block option
Data rate per node [KB/s]	1142.75	43.25	54.33/65.72
Time slot utilization [%]	45.71%	1.73	2.02/2.62
Max. nodes per time slot	approx. 2	approx. 57	approx. 49/39

TABLE I
SUMMARY OF THE CoHART PERFORMANCE USING 50 TIME SLOTS PER CYCLE AND A TIME SLOT SIZE OF 1 MS

of the payload. With increasing file size, the number of new created packets becomes negligible compared to the increasing number of re-used packets. Therefore, the improvement in the performance with bigger files sizes becomes smaller. Again, the effect of a significantly higher time slot utilization is visible if the time slot sizes are chosen very small. It is the same effect like described for the second scenario.

VI. USE OF REED-SOLOMON CODES

In IE environments, it is usually assumed that data is transmitted successfully as the media access is controlled. We have also explored the possibility of increasing the data transmission reliability by using Erasure Resilient Codes (ERC). Reed Solomon (RS) Codes have been chosen as a realization of ERC. If we assume a data volume of 1000 bytes, the time for encoding equates to 430 ms and 620 ms for decoding. This is a quite huge time for a small amount of data. If the encoding and decoding time is a non time critical aspect, it could be used to increase the reliability of the data. However, due to the intended high performance, we renounce the usage of RS codes if new data appears frequently and has to be processed fast. Instead of RS codes, either data replication can be used or the reliability of the physical implementation of the network should be improved. A third opportunity is to improve the RS code performance by using hardware support. However, this would lead to a proprietary solution and is therefore not favored by us as we want to create a non proprietary solution.

A. Estimation of number of nodes

If only the data transmission is considered and the search within the Kad network is already performed, it is possible to determine the number of supported nodes via the available bandwidth. For all scenarios, 1 GBit connections between all devices are considered. Additionally, the number of time slots per cycle is set to 50 and the time slot size is set to 1 ms. With these values, the difference between the scenarios in Table I can be recognized. This results in a cycle time of 50 ms, in which every node is able to communicate. The mentioned data rate linearly depends on the number of time slots per cycle. As an example, in scenario 3 for 100 KB files the system is able to handle 1950 nodes assuring a deterministic data rate of 65.72 KB/s. The scenarios are only a demonstration and for a desired system the number and size of the time slots have to be adjusted depending on the given requirements.

VII. CONCLUSION

In this paper, we have shown that the combination of a Peer-to-Peer based hard real-time capable system called HaRTKad

and CoAP leads to a system called CoHaRT, which allows the deterministic transmission of high amounts of data. The underlying HaRTKad features high scalability, flexibility, self-organisation, and reliability while CoAP offers the possibility of transferring and reassembling big data amounts in a standardized way. CoAP therefore enhances the functional scope of HaRTKad. It depends on the target application how to parametrize the system. For future work, we will investigate the possibility of utilizing Software Defined Networks (SDNs) to allow for more parallel and low jitter communication. Furthermore, a more flexible transmitting mechanism will be focused to solve the issue of the standard stop-and-go-like CoAP blockwise transmission.

REFERENCES

- [1] Z. Shelby, K. Hartke, and C. Bormann, *RFC 7252: The Constrained Application Protocol (CoAP)*, IETF Std., Juni 2014. [Online]. Available: <https://tools.ietf.org/html/rfc7252>
- [2] T. Sauter, "Integration aspects in automation - a technology survey," in *10th IEEE Conference on Emerging Technologies and Factory Automation, 2005. ETFA 2005.*, vol. 2, 2005, pp. 255–263.
- [3] P. Danielis, J. Skodzik, V. Altmann, E. B. Schweissguth, F. Golasowski, D. Timmermann, and J. Schacht, "Survey on real-time communication via ethernet in industrial automation environments," in *19th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA'2014)*, 2014.
- [4] J. Skodzik, P. Danielis, V. Altmann, and D. Timmermann, "Hartkad: A hard real-time kademia approach," in *11th IEEE Consumer Communications & Networking Conference (CCNC)*, 2014, pp. 566–571.
- [5] —, "Time Synchronization in the DHT-based P2P Network Kad for Real-Time Automation Scenarios," in *2nd IEEE WoWMoM Workshop on the Internet of Things: Smart Objects and Services (IoT-SoS)*, 2013.
- [6] J. Skodzik, V. Altmann, P. Danielis, and D. Timmermann, "A kad prototype for time synchronization in real-time automation scenarios," in *World Telecommunications Congress (WTC) 2014*, 2014.
- [7] T. Steinbach, H.-T. Lim, F. Korf, T. Schmidt, D. Herrscher, and A. Wolisz, "Tomorrow's in-car interconnect? a competitive evaluation of ieee 802.1 avb and time-triggered ethernet (as6802)," in *2012 IEEE Vehicular Technology Conference (VTC Fall)*, 2012, pp. 1–5.
- [8] M. Felsler, "Real Time Ethernet: Standardization and implementations," in *Industrial Electronics (ISIE), 2010 IEEE International Symposium on*, 2010, pp. 3766–3771.
- [9] H. Khattak, M. Ruta, and E. Di Sciascio, "Coap-based healthcare sensor networks: A survey," in *Applied Sciences and Technology (IBCAST), 2014 11th International Bhurban Conference on*, Jan 2014, pp. 499–503.
- [10] P. Thubert, T. Watteyne, M. Palattella, X. Vilajosana, and Q. Wang, "Ietf 6tsch: Combining ipv6 connectivity with industrial performance," in *Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS), 2013 Seventh International Conference on*, July 2013, pp. 541–546.
- [11] C. Bormann and Z. Shelby, *Blockwise transfers in CoAP: draft-ietf-core-block-14*, IETF Std., 2013. [Online]. Available: <http://tools.ietf.org/html/draft-ietf-core-block-14>
- [12] Avnet. [Online]. Available: <http://www.zedboard.org/>
- [13] Real Time Engineers Ltd. [Online]. Available: <http://www.freertos.org/>
- [14] I. Free Software Foundation, "lwip - a lightweight tcp/ip stack," 2014. [Online]. Available: <http://savannah.nongnu.org/projects/lwip/>