Using BIP to reinforce correctness of resource-constrained IoT applications
Alexios Lekidis, Emmanouela Stachtiari, Panagiotis Katsaros, Marius Bozga, Christos Georgiadis

To cite this version:

HAL Id: hal-01213705
https://hal.archives-ouvertes.fr/hal-01213705
Submitted on 8 Oct 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Using BIP to reinforce correctness of resource-constrained IoT applications

Alexios Lekidis *, Emmanouela Stachtiari †, Panagiotis Katsaros †, Marius Bozga * and Christos K. Georgiadis ‡

* Univ. Grenoble Alpes, VERIMAG, F-38000 Grenoble, France
** CNRS, VERIMAG, F-38000 Grenoble, France
† Department of Informatics, Aristotle University of Thessaloniki
54124 Thessaloniki, Greece
email: {emmastac, katsaros}@csd.auth.gr
‡ Department of Applied Informatics, University of Macedonia
54006 Thessaloniki, Greece
email: {geor}@uom.edu.gr

Abstract—IoT applications have either a sense-only or a sense-compute-actuate goal and they implement a capability to process and respond to multiple (external) events while performing computations. Existing IoT operating systems provide a versatile execution environment that adheres to the limitations of the interconnected resource-constrained devices. To reduce the development effort, applications are often built on top of RESTful web services, which can be shared and reused. However, the asynchronous communication between remote nodes is prone to event scheduling delays, which cannot be predicted and taken into account while programming the application. Moreover, to avoid long delays in message processing and communication due to packet collisions, the data transmission frequencies between the system’s nodes have to carefully chosen. In general, even when appropriate debugging tools and simulators are available, it is still a hard challenge to guarantee the required functional and non-functional properties at the application and system levels. To this end, we focus on IoT applications for the Contiki OS and we introduce a model-based rigorous analysis approach using the BIP component framework. At the application level, we verify qualitative properties regarding service responsiveness, whereas at the system level we can validate qualitative and quantitative properties using statistical model checking. We present results for an application scenario running on a distributed system infrastructure with nodes executing the Contiki OS.

I. INTRODUCTION

The main challenge in IoT system design is to implement a lightweight architecture with abstractions that provide an appropriate execution environment, while staying within the resource limitations of the interconnected devices. Such an environment can be based on existing IoT operating systems ([1], [2], [3]), which facilitate system integration by abstracting hardware and allowing control of the system’s nodes.

In this context, applications are implemented as event-driven systems with processes acting as event handlers that run to completion. Due to the resource limitations, all processes of a node usually share the same stack, under the condition that an event handler cannot block. When an event is destined for a process, the process is scheduled and the event - along with accompanying data - is delivered to the process through the activation of its event handler.

IoT operating systems decouple the applications’ design from the low-level kernel functions (CPU multiplexing and event scheduling). Thus, the development of IoT applications can proceed independently from their deployment, which has the advantages of programming at a higher-level, but opens a possibility for design errors at the overall system level. Depending on the way that IoT applications are eventually deployed in a distributed environment, they may have to handle and route many different types of events [4]. In general, it is hard to ensure seamless interactions between the system’s components given their high heterogeneity (different device types, measurement units and interaction modes).

Application development often relies on loosely coupled RESTful web services that may be shared and reused. Then, interconnected things are represented as abstract resources controlled by a server process and identified by a Universal Resource Identifier (URI). We can thus access sensors over the web in a request/reply or publish/subscribe manner using ordinary web clients. The resources are decoupled from the services and can be represented by means of various formats, such as XML or JSON. They are accessed and manipulated using the methods of the HTTP or the CoAP protocol. However, the use of web services is not straightforward, because many IoT applications require a multicast and asynchronous communication compared to the unicast and synchronous approach of typical Internet applications. Moreover, IoT application design is of vital importance as web services have to reside in battery operated devices, which are usually deployed in inaccessible or distant areas and thus cannot be frequently changed in case of increased energy consumption.

Special purpose tools for testing by simulation are needed to aid the developers throughout the application design and deployment on the IoT system ([5], [6]). However, the currently available simulation techniques and tools are generally time consuming and hard to use. Even when they are available, it is still a hard challenge to guarantee the required qualitative and quantitative correctness properties at the application and system levels. To this end, we advocate a model-based rigorous analysis approach using the BIP component framework [7].

The research leading to these results has been partially funded by the European ARTEMIS ARROWHEAD project
BIP is a formal language that has been proved successful in building executable models of mixed software/hardware systems. BIP models can be formally analyzed for guaranteeing important functional correctness properties and for evaluating the system’s performance. Validated models can then support the systematic generation of code for deploying the modeled applications on a distributed system, thus preserving the verified properties in the actual implementation.

Our approach is based on the development of faithful models for the IoT application and system levels. A RESTful application is modeled and analyzed for deadlock freedom and other safety and liveness properties, whereas at the system level we validate important non-functional properties using statistical model checking. To the best of our knowledge, these analyses are not supported by the tools used in IoT application development. Our concrete contributions are as follows:

- A BIP application level model is introduced for RESTful IoT applications. The interconnected things are represented by abstract resources controlled by a server [8]. The resources are accessed and manipulated using the CoAP protocol in client/server request/responses.
- At the system level, we present a BIP model for applications running on the Contiki OS. This model was initially constructed according to the existing communication standards for the supported protocol stack and it was later calibrated based on software-dependent runtime constraints of the Contiki OS.
- A RESTful service-based application scenario is analyzed with state space exploration and statistical model checking. We provide results for key functional and non-functional properties.
- We provide proof-of-concept results for studying error behaviors in the IoT system.

The rest of the paper develops along the following lines. Section II presents the current state of practice for the design of IoT applications. We specifically focus on the widely used Contiki OS, which is a typical case of the system architecture principles implemented in most IoT operating systems. Section III provides the necessary background on the BIP language and the supported statistical model checking functionality. Section IV describes in detail our BIP application and system level models. Section V presents the case study scenarios and the obtained state space exploration and statistical model checking results, including those illustrating the study of various error behaviors. Finally, the paper concludes with a critical view of our work’s contributions and a discussion on the future research and development prospects.

II. IoT APPLICATION DESIGN

We focus on the specific programming abstractions of the Contiki OS, which in general are in line with to the same principles behind the programming models of today’s IoT operating systems. However, our rigorous analysis approach is still applicable for other operating systems, if there are faithful models of their implemented programming abstractions.

The Contiki OS was initially introduced to support the development Wireless Sensor Network (WSN) applications [1], but it was later extended, in order to be integrated into the IoT. Two kinds of events are handled: (i) asynchronous events which are enqueued by the kernel and are dispatched later to the target process, and (ii) synchronous events that cause the target process to be scheduled immediately; execution control returns to the event posting process only when the target process has finished the event processing. A polling mechanism is also provided, which is usually used by processes that check for status updates of hardware devices. Polling is realized by scheduling high-priority events, which trigger calls of all processes having a poll handler, in the order of their priority.

An event scheduler dispatches events to running processes and periodically calls process poll handlers. Once an event has been scheduled, its event handler must run to completion since it cannot be preempted by the kernel.

```c
1  #define PROCESS(client, "Client Example")
2  AUTOSTART_PROCESSES(&client);
3
4  static struct uip_udp_conn *conn;
5  static struct etimer et;
6  const int period = 160; /* in sec. */

7  PROCESS_THREAD(client, ev, data) {
8      PROCESS_BEGIN();
9      PROCESS_POLLHANDLER(printf("Process polled\n");)
10     PROCESS_EXITHANDLER(conn_close();)
11     PROCESS_BEGIN();
12     SENSORS_ACTIVATE(button_sensor);
13     server1 = conn_new(); /* function call */
14     etimer_set(&et, period);
15     while(1) {
16         PROCESS_YIELD();
17         if (ev == tcpip_event & data == &button_sensor)
18             PROCESS_WAIT_EVENT_UNTIL
19             (ev == PROCESS_EVENT_TIMER);
20             /* function call for data request */
21             etimer_reset(&et);
22         } else if (ev == tcpip_event) {
23             /* function call to handle received data */
24         }
25     }
26     PROCESS_END();
27  }
```

Listing 1: Contiki client process

In typical Contiki applications, processes are implemented as lightweight threads, called protothreads, that do not have their own stack. A protothread consists of a C function and a single local continuation, i.e. a low-level mechanism to save and restore the context, when a blocking operation is invoked. Thus, when an event handler does not run to completion, the scheduling of other processes takes place. The protothread’s local continuation is set before each conditional blocking wait. If the wait is to be performed, an explicit return statement is executed and the control returns to the caller. Upon the next invocation of the protothread the local continuation that was previously set is resumed, the program jumps to the same blocking wait statement and re-evaluates the condition. The protothread’s execution continues, once it is allowed by the blocking condition. A local continuation is a snapshot of the current process state and its main difference from ordinary continuations is that the call history and values of local variables are not preserved. If some variables need to be saved across a blocking statement, this limitation can be sidestepped by declaring them as static local variables.

A typical Contiki application consists of processes defined in a C module. Listing 1 shows the code structure of a client process. In the PROCESS macro, the reference variable client and a string are assigned to the process, with the latter used in debugging. The AUTOSTART_PROCESSES macro (line 2) requests the process to be automatically started when

\footnote{The Contiki OS is open source software and its design is transparent to the development community.}
the module is booted. The process code is enclosed in a
PROCESS_THREAD macro, which defines the accessed vari-
ables, the handled event (ev) and its data. The process control
flow is included between the two macros PROCESS_BEGIN
and PROCESS_END, but handlers for the exit and poll events
will be enabled independently from the control flow and are
therefore placed before PROCESS_BEGIN.

```c
1 /* Getter. Returns the reading from light sensor */
2 RESOURCE(light, METHOD_GET, "light");
3 RESOURCE(leds_toggle, METHOD_POST | METHOD_PUT, "toggle");
4 PROCESS(server, "Rest Server Example");
5 PROCESS_THREAD(server, ev, data)
6 { /* Start rest engine process: handles
7    invocation of resource handlers */
8   /* Actuator. Toggles the red led */
9   void leds_toggle_handler(REQUEST* req, RESPONSE* resp)
10   { leds_toggle_handler(REQUEST* req, RESPONSE* resp)
11     /* Getter. Returns the reading from light sensor */
12     void light_handler(REQUEST* request, RESPONSE* response)
13       { uint16_t light_solar;
14         read_light_sensor(&light_solar); /* function call */
15         sprintf(temp, "%u;%u", light_photosynthetic, light_solar);
16         rest_set_header_content_type(response, TEXT_PLAIN);
17         rest_set_response_payload(response, temp, strlen(temp));
18       }
19   }
20  PROCESS(server, "Rest Server Example");
21   AUTOSTART_PROCESSES(&server);
22   PROCESS_THREAD(server, ev, data)
23   { PROCESS_BEGIN();
24     /* Start rest engine process: handles
25        invocation of resource handlers */
26     /* Start rest engine process: handles
27        invocation of resource handlers */
28     rest_init();
29     SENSORS_ACTIVATE(light_sensor);
30     rest_activate_resource(&resource_light);
31     rest_activate_resource(&resource_toggle);
32     PROCESS_END();
33   }
34 }
```

Listing 2: RESTful Contiki server

In the infinite loop (line 15), the process waits for and
processes events. First, it is blocked on line 16 until the receipt
of an event that triggers evaluation of conditions in lines 17
and 22. Depending on whether the event comes from the button
sensor or the network stack, the execution flow is diverted
accordingly. If both conditions are false, the process execution
returns to the PROCESS_YIELD macro and it is blocked. The process
is also blocked on the macro of line 18, until a timer fires an event. Then lines 20 and 21 are executed and the process
is blocked again on line 15. Finally, upon the arrival of an
exit event, the exit event handler is executed and the process
reaches the PROCESS_END macro. Process execution ends when the PROCESS_END macro is reached.

In the Contiki OS, it is possible to develop applications
based on loosely coupled RESTful web services that can
be shared and reused. Listing 2 shows the code structure of a
REST Contiki server comprising resource definitions,
and a process for activating the REST resources and the
REST engine (a number mediators between the application and
communication layers are available for handling the dispatch
of incoming messages and the delivery of their responses).

A resource definition includes a resource variable, the sup-
ported HTTP methods and a URL. Our example shows the
definitions of two resources, a light sensor (line 2) and a LED
toggle switch actuator (line 13). A function conventionally
named [resource_name]_handler is associated with
each resource. This function is invoked upon every single URL
request of the resource with one of the supported HTTP meth-
ods. The resource handlers enclose C code without protothread macros, since they do not implement independent processes.
The light sensor resource handler of our example is a getter
that reads the solar light indication and adds it to the response.
On the other hand, the leds toggle handler switches the red
LEDs on and off. The server process in line 23 initializes the
REST engine process and activates the defined resources, in
order to render them accessible. The REST engine implements
the receipt of URL requests, the invocation of the appropriate
handlers and the delivery of responses. Only one request is
handled at a time and therefore the concurrent execution of
resource handlers is not possible. However, another process
implementation could be used instead of the REST engine.

The programming of Contiki applications can proceed inde-
pendently from their deployment over the IoT system nodes.

The asynchronous communication between remote nodes is
prone to event scheduling delays, due to the execution of other
processes on the same nodes. These delays cannot be predicted
and taken into account, while programming the application.

The problem is further complicated because all processes in a
Contiki node share a single buffer for communication process-
ing. The delays for the encoding and decoding of messages are
hard to be predicted, since they depend on the node’s overall
CPU load. Moreover, there is high probability of a collision
occurrence, if two or more nodes access the communication
medium simultaneously. In this case, all the involved nodes
will back off the transmission for a random period of time and
will retry once this period has elapsed. To avoid long delays in
message processing and communication, the developers will
have to carefully choose the data transmission frequencies
between the system’s nodes.

Currently, the only way to cope with the discussed problems
is the analysis by simulation. For this purpose, the Contiki
developers have introduced Cooja [9], a flexible Java-based
simulation tool. Cooja is based on the lower level MPSim
platform emulator [6] that provides accurate information for a
system’s underlying hardware. It is thus possible to investigate
the system’s functional and non-functional behavior and to
inspect various performance aspects based on Contiki’s native
timeLine module (e.g. message buffer utilization, energy
consumption etc.). In the Cooja simulator, the user can se-
lect parameters for the network environment. Then, the IoT
application’s processes are allocated on the system’s node
representations, as imposed by the distribution of the inter-
connected devices. The functional behavior of the simulated
IoT system can be tested through the simulation of simple
execution scenarios. By simulating more realistic workloads,
it is then possible to inspect the system’s performance.

However, a system simulation can only provide partial
assurance of the system’s behavior and it is not adequate
for guaranteeing correctness properties for and application’s
functional and non-functional requirements.

III. BACKGROUND

A. The BIP component framework

BIP is an expressive component framework that allows build-
ing complex, hierarchically structured models from atomic
components characterized by their behavior and their inter-
faces. We provide the formal semantics of the BIP framework,
as is defined in [7] and earlier articles.

BIP atomic components are transition systems extended with
a set of ports and a set of variables. An atomic component \( C \) is defined as a tuple \((Q, X, P, T)\), where \( Q \) a set of control locations, \( X \) a set of variables, \( P \) a set of communication ports and \( T \) a set of transitions. Each transition \( \tau \) is of the form \((q, p, g, f, q')\) where \( q, q' \in Q \) are control locations, \( p \in P \) is a port, \( g \) is a guard and \( f \) is the update function of \( \tau \). \( g \) is a predicate defined over variables in \( X \) and \( f \) is a function (BIP can invoke functions written in C/C++ that computes new values for \( X \) according to their current values.

\[
\text{(i) Port use}
\]

\[
\text{(ii) Flat Connectors}
\]

\[
\text{(iii) Hierarchical Connectors}
\]

![Fig. 1: Flat and hierarchical BIP connectors](image)

In order to compose a set of \( n \) atomic components \( \{C_i = (Q_i, X_i, P_i, T_i)\}_{i=1}^n \), we assume that their respective sets of ports and variables are pairwise disjoint. We define the global set \( P = \bigcup_{i=1}^n P_i \) of ports. An interaction \( a \) is a triple \((P_a, G_a, F_a)\), where \( P_a \subseteq P \) is a set of ports, \( G_a \) is a guard and \( F_a \) is a data transfer function. By definition \( P_a \) contains at most one port from each component. We denote \( P_a = \{p_i\}_{i \in I} \) with \( I \subseteq \{1 \ldots n\} \) and \( p_i \in P_i \). \( G_a \) and \( F_a \) are defined on the variables of participating components, that is \( I \subseteq X_i \).

Given a set \( \gamma \) of interactions, a priority is defined as a strict partial order \( \gamma \subseteq \gamma \times \gamma \). We write \( a \gamma b \) for \((a, b) \in \gamma\), to express the fact that interaction \( a \) has lower priority than \( b \).

A composite component \( \gamma(C_1, \ldots, C_n) \) is defined by a set of components \( C_1, \ldots, C_n \), a set of interactions \( \gamma \) and a priority \( \gamma \subseteq \gamma \times \gamma \). If \( \gamma \) is the empty relation, we may omit \( \gamma \) and simply write \( \gamma(C_1, \ldots, C_n) \). A global state of \( \gamma(C_1, \ldots, C_n) \) where \( C_i = (Q_i, X_i, P_i, T_i) \) is defined by a pair \((q, v)\), where \( q = (q_1, \ldots, q_n) \) is a tuple of control locations such that \( q_i \in Q_i \), and \( v = (v_1, \ldots, v_n) \) is a tuple of valuations of variables such that \( v_i \in Val(X_i) = \{\sigma : X_i \rightarrow \{0, 1\}\} \), for all \( i = 1, \ldots, n \) and \( D \) being some universal data domain.

The behavior of a composite component \( \gamma(C_1, \ldots, C_n) \) is defined as a labelled transition system over the set \( S \) of global states of \( C \) and the transition relation with the following semantics: \( C \) can execute an interaction \( a \in \gamma \), iff (i) for each port \( p_i \in P_a \), the corresponding atomic component \( C_i \) allows a transition labelled by \( p_i \) (i.e. the corresponding guard \( g_i \) evaluates to true), and (ii) the guard \( G_a \) of the interaction evaluates to true. If these conditions hold true for an interaction \( a \) at state \((q, v)\), then \( a \) is enabled at that state. Execution of \( a \) modifies participating components’ variables by first applying the data transfer function \( F_a \) on variables of all interacting components and then the update function \( f_i \) for each interacting component. Components that do not participate in the interaction stay unchanged.

In BIP, interactions between components are specified by connectors. A connector defines a set of interactions based on the synchronization attributes of the connected ports (Fig. 1i), which may be either trigger or synchron:

- if all connected ports are synchrons, then synchronization is by rendezvous, i.e. the defined interaction may be executed only if all the connected components allow the transitions of those ports (Fig. 1ii).
- if a connector has one trigger, the synchronization is by broadcast, i.e. the interactions are all non-empty subsets of the connected ports with the trigger port (Fig. 1ii).

Connectors can export their ports for building hierarchies of connectors (Fig. 1iii), and can use data variables, in order to compute transfer functions associated with interactions. Computations take place iteratively either upwards (up) or downwards (down) through the connectors’ hierarchy levels, but computed values are not stored between the execution of two interactions (connectors are stateless).

SBIP [10] is a BIP extension that relies on a stochastic semantics, for the verification of large-scale systems by using Statistical Model Checking ([11], [12]). Existing tool support allows verifying various forms of qualitative and quantitative properties connected to non-functional requirements.

Figure 2 shows two atomic components from our IoT system model in BIP. On the left, we have the behavior of the communication Channel component that interacts through the ports beginT, busy, free and endT with the ProtStack MsgSender component on the right. The latter exhibits a stochastic behavior, due to the sampling from the uniform distribution for assigning value to the variable distVal.

![Fig. 2: Two atomic BIP components and their interactions](image)

### B. Statistical Model Checking (SMC)

SMC was recently proposed as a means to cope with the scalability issues in numerical methods that are typically used to check stochastic systems. Consider a system model \( M \) and a set of non-functional requirements \( R_1, \ldots, R_n \). Each requirement is formalized by a stochastic temporal property \( \phi \) written in the Probabilistic Bounded Linear Temporal Logic (PBLTL) [10]. SMC is then used to apply a series of simulation-based analyses in order to answer two questions: (1) qualitative: is the probability \( Pr_M(\phi) \) for \( M \) to satisfy \( \phi \) greater or equal to a threshold \( \theta \)? and (2) quantitative: what is the probability for \( M \) to satisfy \( \phi \)? Both of these questions can serve to decide a PBLTL property.

Quantitative analysis aims to compute the value of \( Pr_M(\phi) \). This depends on the existence of a counterexample to the negation of \( \phi (\neg \phi) \), for the threshold \( \theta \). It is of polynomial complexity and, depending on the system model \( M \) and the property \( \phi \), it may or may not terminate within a finite number of steps. In [12], the authors propose an estimation procedure to compute a value for \( \phi \), such that \(|\hat{p} - \phi| < \delta \) with confidence \( 1 - \alpha \), where \( \delta \) denotes the precision. This procedure is based on the Chernoff-Hoeffding bound [13].

Existing approaches for answering the qualitative question
are based on hypothesis testing [11]. If \( p = Pr_M(\phi) \), to decide if \( p \geq \theta \), we can test \( H: p \geq \theta \) against \( K: p < \theta \). Such a solution does not guarantee a correct result but it allows to bound the error probability. The strength \((\alpha, \beta)\) of a test is determined by parameters \( \alpha \) and \( \beta \), such that the probability of accepting \( K \) (resp. \( H \)) when \( H \) (resp. \( K \)) holds is less or equal to \( \alpha \) (resp. \( \beta \)). However, it is not possible for the two hypotheses to hold simultaneously and therefore the ideal performance of a test is not guaranteed. A solution to this problem is to relax the test by working with an indifference region \((p_1, p_0)\) with \( p_0 \geq p_1 \) \((p_0 - p_1 \) is the size of the region). In this context, we test the hypothesis \( H_0 : p \geq p_0 \) against \( H_1 : p \leq p_1 \) instead of \( H \) against \( K \). If the value of \( p \) is between \( p_1 \) and \( p_0 \) (the indifference region), then we say that the probability is sufficiently close to \( \theta \), so that we are indifferent with respect to which of the two hypotheses \( K \) or \( H \) is accepted.

IV . RESTful APPLICATION AND CONTIKI SYSTEM MODELS

Our verification approach is based on a BIP model for the Contiki IoT system architecture. As it is shown in Fig. 3, each Contiki node is represented by BIP models at three levels, namely the REST module allocated to the node, the Contiki OS and the model for the protocol stack, which allows communication through the network channel. The BIP components for the REST modules comprise the IoT application level model, and the IoT System Model integrates them with the OS and protocol stack components.

For an access point to a wireless communication medium that operates in a frequency at the 2.4 GHz band, the data rate is equal to 250 kbps, therefore from Equation 1 the symbol period is equal to 16\( \mu \)s. Thus, our timing abstraction ignores delays smaller than the inverse of this data rate, which is 4\( \mu \)s. However, this abstraction allows a much more fine-grained timing analysis compared to the one supported by the Cooja simulator, which is in the ms scale.

In our SBI model, we have integrated values of important parameters in order to model faithfully software-dependent runtime constraints of the Contiki OS. These parameters are: (i) the time needed for compression and decompression of the packets IP headers according to the HC1/HC2 encoding mechanisms [14], (ii) the pre and post buffering taking place for each packet transmission. The values for these parameters were obtained by measuring the duration from the beginning till the end of the corresponding executable block of code within the Contiki OS. We note that the actual parameter values differ from system to system, since they mainly depend on the available computational resources.

A detailed definition of BIP model’s structure (Figure 3) is given by the following grammar:

\[
\text{(SystemModel)} ::= \langle \text{AppModel} \rangle \langle \text{ConKernel} \rangle
\]

\[
\text{(AppModel)} ::= \langle \text{RestModule} \rangle *
\]

\[
\text{(RestModule)} ::= \langle \text{Process} \rangle * \langle \text{Resource} \rangle \langle \text{ResHandler} \rangle *
\]

\[
\text{(ConKernel)} ::= \langle \text{OS} \rangle + \langle \text{Network} \rangle
\]

\[
\text{(OS)} ::= \langle \text{Scheduler} \rangle \langle \text{Timer} \rangle \langle \text{CommHandler} \rangle
\]

\[
\text{(Network)} ::= \langle \text{ProtStack} \rangle + \langle \text{Channel} \rangle
\]

\[
\text{(ProtStack)} ::= \langle \text{MsgSender} \rangle \langle \text{MsgReceiver} \rangle
\]

The BIP model comprises two layers, namely the RESTful Application Model (AppModel) and the Contiki Kernel (ConKernel), which represent respectively, the application logic and the Contiki platform. AppModel consists of several RestModule composite components that are deployed in different Contiki nodes. Each RestModule includes a number of process components (P1 to Pk) and optionally a set of resource components (R1 to Rn), representing REST resources, and resource handler components (Hi to Hj), which manipulate the resources. The ConKernel includes the composite components for the OS of every node and the Network composite component, with components for the protocol stack of each node and the communication channel.

A. RESTful Application Model

AppModel consists of a number of RestModule components that are hosted in the system’s nodes. Fig. 4 shows the structure of the RestModule composite component and its interactions with the OS component. For simplicity, we show only the interactions of the OS with one process of the RestModule component, but the same interactions are applied for several processes. Every process is represented by an atomic component, whose behavior depends on the application’s logic.

Each component for a process interact with the ConKernel model through the ports shown in Fig. 4. The process is called (called port) when an event destined for the process is scheduled. After handling the event, it yields execution (yield port). Synchronous or asynchronous events may be posted that are handled by other processes using the postSyn and postAsyn ports. If a synchronous event has been posted, the process resumes execution (resume port) upon the end of event handling. Each process may also request polling for itself or other processes. It can set deadlines using timers (setTimer...
port) or send a message (sndMsg port). ConKernel informs the process about the completion of setting a timer or transferring a message (timerSet, msgSnt ports). When a process finishes its execution the end port is enabled.

Fig. 5 shows the structure of a RestModule component with a server. The server process is represented by an atomic component with ports for calling a resource handler (callHandler) and for getting the response (handlerEnd). Every resource handler interacts with one resource, in order to submit a query (sndQry port) and get its result (getRslt port). Multiple resource handlers are allowed to interact with the same resource.

![Fig. 5: The RestModule for a Server](image)

**B. Contiki Kernel Model**

The ConKernel model consists of the OS and the Network composite components. The former models the behavior of the Contiki kernel [1] regarding the scheduling and the event-based interprocess communication. The latter represents the Contiki network module and therefore models the entire protocol stack.

As it is shown in Fig. 6, the OS composite component interacts with (i) the RESTModule component, in order to receive and handle incoming events and (ii) the Network component for the data exchange events. It consists of the Scheduler, the Timer and the CommHandler atomic components. The Scheduler component manages the incoming events from the processes of the RESTModule component. We distinguish four types of events, namely: (1) initialization (INIT event); (2) (a)synchronous event posting; (3) polling (Poll event); (4) yielding (YIELD event); and (5) exiting (EXIT or EXITED events). Events may be either synchronous and asynchronous, and are stored respectively in a stack and in a FIFO queue. The Scheduler checks periodically for the presence of incoming events, in the event queues (period $T_{scheduler}$). In sensor applications, this period affects the sensors’ battery life: if the system operates in battery-saving mode, it wakes up to check for new events, when the period expires. The presence of an event, either triggers immediate handling - as for the synchronous events (postSyn port) - or the event is deferred - as for the asynchronous events (postAsyn port). The Timer component models a timer process of the Contiki OS. It receives the incoming timer requests (setTimer port) of each RESTModule process and stores them as well in a stack. Each timer request includes a timer mode allowing to set, reset, restart or stop a particular timer. Finally, the CommHandler component models the TCP/IP process of the Contiki OS. It interacts with the Network component to either transmit the incoming frames or deliver the received frames to the process components of the RESTModule component. Thus, it uses one transmission and one reception buffer, called TxBuffer and RxBuffer respectively.

The Network composite component is comprised by the ProtStack and the Channel components. ProtStack consists of the MsgSender and MsgReceiver atomic components. The former models data transmission and the latter data reception. Both components represent data exchange through application layer messages of the CoAP or the HTTP protocol. Nevertheless, the transmission/reception mechanisms are handled by the underlying 6LoWPAN protocol of the Contiki protocol stack [15]. According to this protocol, the model uses the unslotted CSMA/CA mode of the IEEE 802.15.4 standard [16] (in the MAC layer) and applies an adaptation layer, in order to transport IPv6 packets over a Wireless Personal Network (WPAN). The Channel atomic component models the behavior of a communication medium, in order to receive requests for data transmission from network nodes, provide information regarding pending network transmissions, as well as resolve deterministically collisions that may arise from the simultaneous transmission attempt of multiple nodes. Fig. 2 illustrates the behavior and interactions of an abstract model for the ProtStack MsgSender and the Channel components. The Channel...
receives requests for data transmission through the interaction involving beginT ports of the two components. It approves one request for data transmission at a time and accordingly moves to the transmitting state l1 through the sending port. As long as it remains in l1, any following data transmissions from the other ProtStack MsgSender components will be halted through an interaction involving their busy ports. When the ongoing transmission ends, both components interact through the endT port and subsequently the Channel notifies all the components who requested access for data transmission that they can repeat their request (free port).

We experimented with a fault model to inject behavior related to extensive bandwidth loss. This allowed us to analyze the fault’s impact on the tested application scenario shown in Section V. A high-level view of the injected behavior is shown by the FaultHandler component in Fig. 7. We distinguish the NORMAL and LOSS states, which represent respectively the successfully transmitted and the delayed or lost packets. FaultHandler receives all transmitted packets through the recv port and decides based on its state, if they will be delivered to their destination. FaultHandler remains in each state, as long as the number of consecutive successful or delayed/lost packet transmissions is positive. This number is chosen by two probabilistic distributions, \( \lambda_{\text{success}} \) and \( \lambda_{\text{loss}} \), obtained from the analysis of debugging traces of a deployment on dedicated hardware platforms [17].

V. CASE STUDY

We introduce an IoT application example consisting of one client and multiple REST server nodes that communicate using the CoAP protocol. A basic interaction scenario is when the client periodically sends unicast GET requests sequentially to two servers, which process them and accordingly reply. The client waits for each response in a certain time frame, before sending again the request. Each node transmits additional CoAP acknowledgments to signal a message receipt.

We introduce an IoT application example consisting of one client and multiple REST server nodes that communicate using the CoAP protocol. A basic interaction scenario is when the client periodically sends unicast GET requests sequentially to two servers, which process them and accordingly reply. The client waits for each response in a certain time frame, before sending again the request. Each node transmits additional CoAP acknowledgments to signal a message receipt.

![Fig. 7: FaultHandler component](image)

![Fig. 8: Case-study client process](image)

TABLE I: Parameters of the modeled protocol stack

<table>
<thead>
<tr>
<th>Model parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>aTurnaroundTime</td>
<td>12 * symbolPeriod</td>
</tr>
<tr>
<td>aUnitBackoffPeriod</td>
<td>12 * symbolPeriod</td>
</tr>
<tr>
<td>aMaxFrameRetries</td>
<td>3</td>
</tr>
<tr>
<td>macMinBE</td>
<td>3</td>
</tr>
<tr>
<td>macAckWaitDuration</td>
<td>54 * symbolPeriod</td>
</tr>
<tr>
<td>macMaxCSMABackoffs</td>
<td>0-5 (default 4)</td>
</tr>
<tr>
<td>CCA duration</td>
<td>20 * symbolPeriod</td>
</tr>
<tr>
<td>symbolPeriod</td>
<td>20 + symbolPeriod</td>
</tr>
<tr>
<td>SIFS</td>
<td>40 + aUnitBackoff</td>
</tr>
<tr>
<td>aMaxFrameRetries</td>
<td>3</td>
</tr>
<tr>
<td>macMinBE</td>
<td>3</td>
</tr>
<tr>
<td>macAckWaitDuration</td>
<td>54 * symbolPeriod</td>
</tr>
<tr>
<td>macMaxCSMABackoffs</td>
<td>0-5 (default 4)</td>
</tr>
<tr>
<td>symbolPeriod</td>
<td>20 + symbolPeriod</td>
</tr>
</tbody>
</table>

ConKernel is parametrized through the Network component. We distinguish two types of model parameters (Table I). Those who are fixed and those which can be modified, in order to analyze performance aspects of the protocol stack. In particular, the second type of parameters may concern the exponential backoff mechanism of the IEEE 802.15.4 standard (used in the MAC layer of the protocol stack) or the timeout for packet reception. The network throughput, as well as the number of collisions in the channel are strongly influenced by such parameters depending on the deployed application. Example parameters of the second type are the macMinBE, macMaxBE, macMaxCSMABackoffs and the macMaxFrameRetries.

A. Verification of functional requirements

As mentioned in Section II, the development of Contiki applications is an error-prone procedure. Scheduling for many processes may cause unpredictable delays to their event consumption. Thus, a client process can set a deadline for an expected response, or else resends the request. The deadline will have to be tuned such that it is feasible for the process to receive the response, in order to avoid sending redundant requests. The analyzed functional and non-functional requirements are relevant for most IoT applications that involve a number of client and server processes. The functional requirements are:

**FR1** The client eventually runs to completion (reaches `PROCESS_END`).

**FR2** The client eventually runs to completion (reaches `PROCESS_END`).

**FR3** The node eventually transmits the requests sent by the client process.

**FR4** The client eventually receives responses for the requests transmitted by its node.

**FR5** All messages for a client transmitted in its node are eventually received by the client process.

FR1 is a consequence of the fact that our system model terminates, since all clients must terminate before the system does. For FR2, we consider a client request as redundant, if it
causes the same response with the one for a previous request. Every such response may or may not reach the client process and this is monitored by the observer automaton in Fig. 9. FR3 requires that the client will transmit successfully all the sent requests. FR4 requires that the client will eventually obtain a response for each transmitted request. Finally, FR5 specifies that if a response has been received by the client’s node, the client process will obtain it. For each of the requirements FR3 to FR5 an observer automaton property was derived.

The observer for the FR2 (Fig. 9) starts its execution when the client process sends a message (sndMsg port). After having received an ack (recvAck port), the observer waits in state s2. If the transfer of the request or the response fails (endSnd port), the observer returns to its initial state. Alternatively, if the process sends a new message before having obtained a response (getMsg port), which has arrived at the node (recv port) or before the server has completed the response dispatch (endSnd port), then state Err is reached respectively from s3 or s2, and the property is falsified.

We tested three different values for the client’s timer: 14.5, 17.5 and 240 ms. The verification results are shown in the Table II. FR2 and FR5 are violated for 14.5 ms, but they are satisfied for the higher values. FR3 is violated for the first two timers, because a small period of sending client requests, incurs the possibility that the node might not be able to transfer a request due to a collision. For the same reason, FR4 is violated for the first two timers. From the verification of FR5, we can only infer that FR3 fails, because a server might fail to transfer the response, due to a collision. However, both FR3 and FR4 hold for the highest timer value, which allows for a safe margin between consecutive requests in order to avoid collisions.

<table>
<thead>
<tr>
<th>timer (ms)</th>
<th>F2</th>
<th>F3</th>
<th>F4</th>
<th>F5 state space</th>
</tr>
</thead>
<tbody>
<tr>
<td>14.5</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X 633.946</td>
</tr>
<tr>
<td>17.5</td>
<td>✓</td>
<td>X</td>
<td>✓</td>
<td>✓ 787.297</td>
</tr>
<tr>
<td>240.0</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓ 48.758</td>
</tr>
</tbody>
</table>

TABLE II: Verification results

B. Analysis of non-functional requirements

The overall performance of an IoT application deployed in a distributed and resource-constrained environment is connected to key non-functional requirements, such as the following:

**NFR1** Memory saving by properly sizing the message buffers in each system node. Such buffers are used in Contiki for the communication through the protocol stack.

**NFR2** Avoidance of overflow in the asynchronous event queue (FIFO) of each node.

**NFR3** Relatively low collision rate in the communication medium, in order to avoid large communication latencies, which have a strong impact to the network performance and may increase the probability of packet losses.

C. Experiments

We conducted two sets of experiments. The first one used the generated code, which was simulated in the Cooja environment as well as the IoT System Model in SBIP, in order to evaluate and compare their performance. Both simulations were focused in the response time for the reply of a Server node to the Client’s request. That is the total time elapsed for the end-to-end transmission, starting from the initialization of the message in the Server until it is correctly received by the Client. As shown in Figure 10 the obtained results were similar for both simulations, however the SBIP System Model had a slightly larger variability providing greater or smaller values than the observed outliers of the Cooja simulator. This is due to its additional accuracy, which enables a better fine-grained simulation in comparison with the Cooja environment.

![Fig. 10: SBIP/Cooja response times for each Server (in ms)](image-url)

The second set was focused in the injection in the IoT system model of a realistic representation for the loss of bandwidth through probabilistic distributions, which results in consecutive packet losses or out-of-order delivery in the system. The Cooja environment also allows the injection of packet losses in the system, through its UDGM - Distance Loss mode. However, since this behavior is based on user-provided simulation values for parameters as the SUCCESS_RATIO_TX and SUCCESS_RATIO_RX, it cannot reflect the reality accurately. Therefore, our analysis exceeds the simulation capabilities of the Cooja environment. The main reason for analyzing packet losses is that once each network node transmits a packet, it waits for the acknowledgment in a certain time frame, namely the macAckWaitDuration parameter of the IEEE 802.15.4 protocol (provided in Table I), before trying a retransmission. Therefore, if this time expires, the following retransmission will increase the packet transmission frequency, leading eventually in higher probability of collisions in the system. Indeed, Figure 11 illustrates such a behavior obtained from the simulation of the SBIP model with the addition of the Fault Model in BIP, presented in Section IV. In particular, the response time for the reply of Server 1 to the Client’s request can be increased up to 23.8 ms, due the presence of collisions in the communication medium.

The construction of the SBIP System Model also allowed the analysis and validation of the aforementioned non-functional requirements for the case study. Therefore, we accordingly describe them with stochastic temporal properties using the Probabilistic Bounded Linear Temporal Logic (PBLTL) for-
malism [10] and present their evaluation results through the SBIP model checking tool.

**NFR1.** In order to satisfy this requirement several properties can be considered. Such properties are 1) \( \phi_1 = (size(TxBuffer) < A \) and 2) \( \phi_2 = (size(RxBuffer) < B) \), where \( size(TxBuffer) \) and \( size(RxBuffer) \) indicate the size of the transmission and reception buffer of the protocol stack respectively. \( B \) and \( A \) are fixed non-negative numbers representing a bound for the size of the two buffers. For the sake of brevity in this experiment we focus on \( \phi_2 \), namely the reception buffer of the CommHandler component, however this analysis can be even conducted for the transmission buffer. In this experiment we tried to estimate the value of \( B \) as it varies according to the value of the period that the Scheduler uses to check for the presence of incoming events in the event queues (\( p_{scheduler} \)). This period determines the frequency with which the Scheduler posts events concerning the communication through the protocol stack to the CommHandler component. If this frequency is increased, the number of transmitted packets over the network is equally increased. Therefore, more packets are received in the reception buffer of the CommHandler component. In particular, we have experimented with different values for \( p_{scheduler} \), such as \( p_1 = 0.1\text{ms} \), \( p_2 = 10\text{ms} \) and \( p_3 = 1s \). For \( p_{scheduler} = p_1 \), as illustrated from Figure 12, \( P(\phi_2) = 1 \) for \( A \) equal to 1. If \( p_{scheduler} \) is increased and is equal to \( p_2 \), \( \phi_2 = 1 \) for \( B \) equal to 5. In the worst-case scenario (\( p_3 \)), \( p_{scheduler} \) was equal to the packet transmission period (1s in the specific case study) and the value of \( A \) has to be 10, in order to guarantee that the \( \phi_2 \) always holds. The size of the reception buffer can be adjusted by the parameter \( MAX\_NUM\_QUEUED\_PACKETS \) of the Contiki OS, found specifically in the core module of the Contiki kernel. This parameter corresponds to the value of \( A \) in our analysis and is initially equal to 2.

**NFR2.** We have accordingly expressed the second requirement as: \( \phi_3 = (size(AsynFIFO) < MAX) \), where \( size(AsynFIFO) \) indicates the size of the asynchronous event queue in the model. For this experiment we have considered the value of \( MAX \) equal to 10. As for the particular case study we only consider a limited number of asynchronous events, we have evaluated that this property holds always (\( \phi_3 = 1 \)) for the chosen value of \( MAX \).

**NFR3.** This requirement is expressed as the property: \( \phi_4 = (NC \leq 1) \), where \( NC \) indicates the number of successive retransmissions following the occurrence of a collision in the model. In order to evaluate this property we have conducted two sets of experiments the first one only on the IoT System Model and the second with the addition of the Fault Model. We have tested the property \( \phi_4 \) for the first experiment in a large number of communication cycles and evaluated it as \( P(\phi_4) = 1 \), meaning that no collisions are present in the communication medium. However, for the second experiment the impact the same property was evaluated as \( P(\phi_4) = 0.55 \), as the extensive loss of bandwidth increased the packet transmissions and hence the number of collisions in the SBIP model. This is also illustrated in Figure 11.

**VI. RELATED WORK**

The notion of the IoT has received wide attention during the past few years. Synergetic activities need to be conducted in different research areas (such as informatics, telecommunications, electronics and social science) for achieving significant contributions to the advance of the IoT. Nevertheless, developing applications for the IoT can be challenging, due to functional and non-functional constraints, such as service interoperations and scarce resources. An important direction is the Cloud centric vision for worldwide implementation of IoT. Cloud receives data from the ubiquitous sensors, analyzes it and provides the user with easy to understand web based visualization. Developing IoT applications using low-level Cloud programming models and interfaces is complex. To overcome this, authors in [18] underline the necessity of an IoT application specific framework for rapid creation of applications and their deployment on Cloud infrastructures.

A model-based design methodology for the development of IoT applications in WSN has been presented in [19]. The proposed framework utilizes graphical tools of the Matlab environment to model and simulate the behavior of an application’s components in a high-level. While it is a promising approach which enables the generation of network simulations that can be executed in Contiki, it lacks support for exhaustive validation of functional and non-functional requirements. IoT applications have undoubtedly a Sense-Compute-Control (SCC) character, and thus the model-driven development framework for SCC applications, presented in [20] is worth mentioning. It separates SCC application development into different concerns and integrates a set of high-level languages...
to specify them. This approach is supported by automation techniques at different phases of SCC application development and allows an iterative development to handle evolutions in different concerns. Still, this approach is not capable to support extensive validation of the applicable requirements. An approach in model-based development, closer to ours is the modeling of system and application behaviors utilizing Finite State Machines (FSM). In [21] authors introduce an FSM approach to enable a model-driven development approach for service-based IoT applications. Through the utilization of FSM, the application logic can be described, which in sequence can be evaluated though a developed GUI tool.

Alternative WSN operating systems which were extended for IoT applications include the RESTful API implementation for TinyOS [3] and the Linux environment. However, considerable limitations exist for those two operation systems: Linux consists of a monolithic kernel, which lacks modularity and often results in a complex structure that is hard to understand, especially for large-scale systems. TinyOS on the other side, uses an event driven code style, thus it cannot support multithreading. An emerging under-development IoT operating system for such applications is RIOT [2], which is similar to Contiki and includes further optimizations in the resource usage. Nevertheless, as the development is still ongoing, it doesn’t yet include an implementation for CoAP and porting of the OS to several popular IoT platforms is not provided.

VII . CONCLUSION

We presented a rigorous model-based analysis for resource-constrained IoT applications using the BIP component framework. The whole approach was demonstrated through the systematic construction of faithful BIP models for RESTful service-based applications over nodes running the Contiki OS. As a proof of concept, our approach was applied to an application scenario with a client and multiple server nodes. Important functional and non-functional requirements were verified, which extend the analysis possibilities beyond those provided by the Contiki simulation tools. On the other hand, we showed that simulation results from our BIP model are very similar to those obtained by the Contiki’s Cooja simulator. Various aspects of service responsiveness were checked by state space exploration, whereas non-functional requirements concerning buffer utilization, collision rate and blocking time in the event queues were analyzed using statistical model checking. Finally, we proposed a fault injection approach for exploring the impact of various error-prone behaviors in the network communication.

For state space exploration, only a limited number of nodes and client/server processes is required that suffice to generate all the different execution traces at the node level. For being able to explore the performance aspects of an IoT application deployed on a large-scale system, we plan to develop an appropriate stochastic abstraction of the presented IoT system model. Such an abstraction can be applied only to the OS and protocol stack models, in order to avoid loosing event interleavings due to the application’s functional behavior. One possibility for the proposed abstractions is the use of the AAlergia learning algorithm [22]. Furthermore, we consider possible extensions in the IoT system model, in order to analyze various security risks related to the DTLS transport and the CoAP protocols [23] or the HTTPS URI scheme [24], as well as their overall impact on the system’s performance.

REFERENCES