Model Checking and Evaluating QoS of Batteries in MPSOC Dataflow Applications via Hybrid Automata

Waheed Ahmad, Marijn Jongerden, Mariëlle Stoelinga, Jaco van de Pol
University of Twente, The Netherlands
Email: {w.ahmad, m.r.jongerden, m.i.a.stoelinga, j.c.vandepol}@utwente.nl

Abstract—System lifetime is a major design constraint for battery-powered mobile embedded systems. The increasing gap between the energy demand of portable devices and their battery capacities is further limiting durability of mobile devices. Thus, the guarantees over Quality of Service (QoS) of battery-constrained devices under strict battery capacities are of primary interest for mobile embedded systems' manufacturers and stakeholders.

This paper presents a novel approach for deriving QoS of applications modelled as synchronous dataflow (SDF) graphs. We map these applications on heterogeneous multiprocessor platforms that are partitioned into Voltage and Frequency Islands, together with multiple kinetic battery models (KiBaMs). By modelling the whole system as hybrid automata, and applying model-checking, we evaluate, (1) system lifetime; and (2) minimum required initial battery capacities to achieve the desired application performance. We demonstrate that our approach shows a significant improvement in terms of scalability, as compared to a priced timed automata based KiBaM model. This approach also allows early detection of design errors via model checking.

I. INTRODUCTION

Mobile computing has experienced a major upswing over the last two decades. As a result, applications with increasing functionality and complexity are continuously implemented on mobile embedded devices such as smart phones, allowing these systems to operate independently. However, this trend has increased the energy consumption of mobile devices manifold. On the other hand, battery energy densities have not grown at the same rate over the years, thus leading to system lifetime as a major design constraint. In this paper, we define the lifetime as the time one can use the battery before it is empty.

Mobile embedded systems are often powered only by batteries that may or may not be charged regularly by an external power source. For example, in a military Software Defined Radio that is being operated in a desert or on a mountain where energy supplies are unreliable, the primary Quality of Service (QoS) concern is to determine the system lifetime. Similarly, a geostationary satellite with solar panels to charge on-board batteries, is recharged at a regular intervals of 12 hours when facing the sun. However, the satellites have strict limitations regarding mass and volume. In this case, the main QoS interest is to assess the battery sizes and weight that yield the performance criteria. In these cases, the evaluation of the QoS of battery-constrained mobile embedded systems has emerged as one of the most critical, challenging and essential concern for manufacturers, investors and users.

One can identify three QoS factors, and their relation with three different design choices, as given in Table I. First, the throughput of a system, defined as a measure of how many units of information a system can process in a given amount of time, has a direct impact on energy consumption which in turn influences the system lifetime. Secondly, the number of processors affects both the system lifetime, and manufacturing cost of the overall system. Lastly, the number of batteries relates not only to system lifetime and cost, but also to the mass and volume of a system. Therefore, this paper takes in account aforementioned design alternatives, with respect to system lifetime and minimum battery capacities.

This paper considers a very intuitive battery model termed Kinetic Battery Model (KiBaM) [16] as a representation of dynamic behaviour of a conventional rechargeable battery, see Figure 1. A KiBaM models the total charge in a battery as two separate tanks. One tank holds the charge which is immediately available to be consumed by the load. The other tank holds the charge which is chemically bound. It is because a chemical kinetics process is used as its basis, that the model is termed kinetic. Experimental studies show that the KiBaM provides a good approximation of the system lifetime across various battery types [13].

To reduce the power consumption, different system-level power management methods like Dynamic Power Management (switching to low power state) (DPM) [7] and Dynamic Voltage and Frequency Scaling (throttling processor frequency) (DVFS) [19] has gained significant value and success. The concept of voltage-frequency islands (VFIs) [11] further allows us to cluster a group of processors in such a way that each VFI runs on a common clock frequency/voltage. This achieves fine-grained system-level power management. In CMOS based processors, voltage/frequency scaling by a factor of $s$ causes the battery current to scale by $s^3$ [8]. Thus, the system lifetime depends hugely on the battery capacity and the level of the load current applied to it.

If we have multiple batteries in the system, another important factor contributing to overall lifetime is the usage pattern of the batteries, i.e., how batteries are scheduled. This leads to an important research problem of devising a battery-aware scheduling mechanism, where given a set of tasks, a set of resources to execute the tasks, and a given number of batteries, we are able to derive a battery-optimal schedule of tasks.
The charge stored in a battery is represented by a finite set of continuous variables in the KiBaM, making the behaviour of KiBaM hybrid. Synthesising optimal battery schedules for multiple batteries using existing analysis techniques for hybrid systems, is very expensive. Therefore, the state-of-the-art method in [14] discretises the KiBaM, and models it as priced timed automata (PTA) [6]. Furthermore, for a fixed load, this approach deploys the model-checker UPPAAL to search the whole state-space, and generating the optimal battery schedule. However, this method also does not solve the scalability problem. As increasing the initial battery capacities leads to searching the bigger state-space, this approach only allows to model limited battery capacities. Furthermore, this approach discretises the temporal dimension, which limits its accuracy.

We propose an alternative, novel approach based on Hybrid Automata (HA) [10]. These extend timed automata [4] (for the modelling of time-critical systems and time constraints) by continuous variables. HA can be analysed using UPPAAL [5], that supports both model-checking and highly scalable Monte Carlo simulations. In contrast to discretisation, as done in [14], we take into account the continuous variables of the KiBaM by modelling it as a hybrid automaton, which obviously makes it a more accurate model than PTA-based KiBaM. This approach enables us to utilise UPPAAL to employ highly scalable technique of Monte Carlo simulations to assess various QoS parameters, such as, system lifetime and adequate battery capacities. In this paper, we show that our approach scales better than the one presented in [14]. Furthermore, we also utilise UPPAAL for applying model checking to verify various user-defined properties. Thus, as opposed to other simulation based tools for hybrid systems, modelling as HA and using UPPAAL provides an additional benefit of model checking against state-based properties.

We use Synchronous Dataflow (SDF) [15] as a computational model. SDF provides a natural representation of real-time streaming and digital signal processing applications. In this paper, SDF graphs are used to represent software applications which are partitioned into tasks, with inter-task dependencies and their synchronisation properties.

Our approach takes four ingredients: (1) a platform model that describes the specifics of the hardware, such as, VFI partitions, frequency levels and power usage per processor; (2) an SDF graph scheduler that maps the application tasks on the platform model in a static-order manner; (3) given number of batteries; and (4) a battery scheduler that defines the scheduling scheme of batteries. For given battery capacities and timing constraints, we compute system lifetime (SDF graph iterations). Similarly, for given application performance criteria, we determine the adequate battery capacities. This method facilitates system designers to evaluate aforementioned QoS factors for different design choices, such as, varying number of VFI, processors, and batteries. Furthermore, this method also allows system designers to detect subtle battery design errors in early phases via model checking.

In particular, our main contributions are as follows. (1) We utilise hybrid automata to model check and assess QoS of multiple KiBaMs for different design alternatives, without discretising time. (2) We consider realistic hardware platforms equipped with the novel energy management techniques, compared to the state-of-the-art [20]; (3) We analyse SDF graphs as input which are more versatile and allow more realistic data-dependencies than acyclic applications [9] [14] [20]; (4) We show that our approach allows better scalability than PTA-based discretised KiBaM [14]; (5) Our approach allows early detection of design errors via model checking.

**Paper organisation.** Section II reviews related work. Section III formalises the problem, and translation to HA using UPPAAL is illustrated in Section IV. Section V experimentally evaluates the QoS analysis. Finally, Section VII draws conclusions and outlines possible future research.

## II. RELATED WORK

An extensive survey paper [13] outlines the broad research work on various battery models. The work in [3] applies the combination of DVFS and DPM, on VFI-partitioned heterogeneous processors. However, this paper assumes energy source to be ideal. We solve this limitation by considering a realistic model of batteries.

The state-of-the-art methods in the realm of battery-aware scheduling for multiple batteries, are presented in [14] and [9]. The approach in [14], in comparison to ours, discretises time. This approach helps to find optimal battery schedules, but do not scale well because of the discretisation. The technique in [9] models KiBaMs as hybrid like us, and discretises time to search the state-space, leading to the better results than the work in [14]. But, due to the fact that the state-space grows larger with the number of batteries, the scalability of this approach also suffers. We, on the other hand, run Monte Carlo simulations, that allows us to avoid the state-space explosion. The analysis shows that the scalability of our approach is better than the technique in [14].

A more advanced technique that also utilises hybrid automata like us and [9], is presented in [20]. In this paper, the KiBaM provides energy to an uniprocessor. Unlike our method, this approach discusses a single battery case only. Another novel work in [12] extends KiBaMs with random initial SoC and load, without discretising time. In this way, probabilistic guarantees about the system lifetime can be provided. In comparison to our work, this technique is also confined to a single KiBaM only. Table II summarises different aforementioned KiBaM analysis methods.

To the best of our knowledge, there are no papers that

---

**Table I: Relation between Design Choices and QoS factors**

<table>
<thead>
<tr>
<th>Design Choices</th>
<th>QoS</th>
<th>System Lifetime</th>
<th>Cost</th>
<th>Volume and Mass</th>
</tr>
</thead>
<tbody>
<tr>
<td>Required Throughput</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Number of Processors</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Number of Batteries</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

**TABLE I: Relation between Design Choices and QoS factors**
analyse multiple KiBaMs without discretising time.

III. SYSTEM MODEL DEFINITION

We formalise battery model, SDF graphs, and hardware platform model in subsection III-A, III-B and III-C respectively.

A. Kinetic Battery Model

The kinetic battery model (KiBaM) [16] is a mathematical characterisation of state of charge of a battery. Key feature is that not all energy stored in a battery can be utilised at all times. To model this phenomenon, the total charge stored in a battery is divided into two ‘tanks’, the available charge and the bound charge, see Figure 1. Only the available charge can be consumed immediately by a load at the time-dependent rate \( i \), and thereby behaves similar to an ideal energy source. The bound charge is converted to available charge at a rate proportional to the height difference with the proportionality factor being the rate constant \( k \), and is available to be consumed. Thus, bound charge replenishes available charge, and this effect is termed recovery effect.

If the widths of the available and bound charge tanks are \( c \) and \( 1-c \) respectively, then the tanks are filled to heights \( h_a \) and \( h_b \), and the charges in both tanks are \( a = c h_a \) and \( b = (1-c) h_b \) respectively. Formally, the KiBaM is characterised by the following system of differential equations.

\[
\begin{align*}
\dot{a}(t) &= -i(t) + k(h_b - h_a) \\
\dot{b}(t) &= -k(h_b - h_a)
\end{align*}
\]  

The system starts in equilibrium, i.e. \( h_a = h_b \). With an initial capacity of \( C \), the initial conditions are \( a(0) = cC \) and \( b(0) = (1-c)C \). The battery is considered empty when \( a = h_a = 0 \), as it cannot supply charge any more at the given moment even though it may still contain bound charge. The system lifetime ends when all batteries are emptied.

Definition 1. A KiBaM system \( KS = (B, Cap) \) consists of a finite set of KiBaMs \( B = \{bat_1, \ldots, bat_m\} \), and a function \( Cap: B \rightarrow \mathbb{R}_{\geq 0} \) denoting the initial capacity of each \( bat \in B \).

In our case-studies, we consider batteries having the capacity of 1300 mAh, as used in the Samsung Galaxy Fame smartphones [1].

B. SDF Graphs

Typically, real-time streaming applications execute a set of periodic tasks, which consume and produce a fixed amount of data. Such applications are naturally modelled as SDF graphs [15]: a directed, connected graph in which tasks are represented by actors. Actors communicate with each other via streams of data elements, represented by tokens. The tokens are transported between the actors via edges. The execution of an actor is known as an (actor) firing.

Definition 2. An SDF graph is a tuple \( G = (A, D, Tok_0, \tau) \) where: \( A \) is a finite set of actors, \( D \subseteq A^2 \times \mathbb{N}^2 \) is a finite set of dependency edges, \( Tok_0 : D \rightarrow \mathbb{N} \) denotes distribution of initial tokens in each edge, and the execution time of each actor is given by \( \tau : A \rightarrow \mathbb{N}_{\geq 1} \).

Definition 3. Given an SDF graph \( G = (A, D, Tok_0, \tau) \), the sets of input and output edges of an actor \( a \in A \) are defined respectively as \( In(a) = \{(a', a, p, q) \in D | a' \in A, p, q \in \mathbb{N}\} \) and \( Out(a) = \{(a, b, p, q) \in D | b \in A, p, q \in \mathbb{N}\} \). The consumption and production rate of an edge \( e = (a, b, p, q) \in D \) are defined respectively as \( CR(e) = q \) and \( PR(e) = p \).

Informally, actor \( a \) can fire if each input edge \( (a', a, p, q) \in In(a) \) contains at least \( q \) tokens; firing actor \( a \) removes \( q \) tokens from the input edge \( (a', a, p, q) \). Firing lasts for \( \tau(a) \) time units and ends by producing \( p' \) tokens on each output edge \( (a, b, p', q') \in Out(a) \).

Example 1. Figure 2 shows the SDF graph of an MPEG-4 decoder [18]. The SDF graph contains five actors \( A = \{FD, VLD, IDC, RC, MC\} \), representing the tasks performed in MPEG-4 decoding. For example, the frame detector (FD) determines the number of macro blocks to decode. To decode a single frame, FD must process between 0 and 99 macroblocks, i.e., \( x \in \{0, 1, \ldots, 99\} \) in Figure 2.

Arrows between the actors depict the edges which hold tokens (dots) representing macroblocks. The execution time (ms) of the actors is represented by a number inside the actor nodes. The numbers near the source and destination of each edge are the production and consumption rates respectively.

To avoid unbounded accumulation of tokens in a certain edge, we require SDF graphs to be consistent.

Definition 4. A repetition vector of an SDF graph \( G = (A, D, Tok_0, \tau) \) is a function \( \gamma : A \rightarrow \mathbb{N}_0 \) such that for every edge \( (a, b, p, q) \in D \) from \( a \in A \) to \( b \in A \), the relation \( p \cdot \gamma(a) = q \cdot \gamma(b) \) exists. An SDF graph is consistent iff \( \gamma(a) > 0 \) for all \( a \in A \).

Definition 5. Let us consider an SDF graph \( G = (A, D, Tok_0, \tau) \) with a repetition vector \( \gamma \). An iteration of \( G \) is defined as a set of actor firings such that for each \( a \in A \), the set contains exactly \( \gamma(a) \) firings of actor \( a \). Thus, each actor fires according to \( \gamma \) in an iteration.

<table>
<thead>
<tr>
<th>Method</th>
<th>Without Discretisation</th>
<th>Multiple KiBaMs</th>
</tr>
</thead>
<tbody>
<tr>
<td>[9]</td>
<td>✗</td>
<td>✓</td>
</tr>
<tr>
<td>[14]</td>
<td>✓</td>
<td>✗</td>
</tr>
<tr>
<td>[20]</td>
<td>✗</td>
<td>✓</td>
</tr>
<tr>
<td>[12]</td>
<td>✓</td>
<td>✗</td>
</tr>
<tr>
<td>Our Method</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

TABLE II: Comparison among different battery analysis methods

![Fig. 1: Model of a KiBaM](image-url)
C. Platform Application Model

A Platform Application Model (PAM) models a multiprocessor platform where the application, modelled as SDF graph, is mapped on. Our PAM models supports several features, including (1) heterogeneity, i.e., actors can run on certain type of processors only; (2) a partitioning of the processors in voltage and frequency islands; (3) different frequency levels each processor can run on; (4) power consumed by a processor in a certain frequency, both when in use and when idle; (5) power and time overhead required to switch between frequency levels.

Definition 6. A platform application model (PAM) is a tuple $\mathcal{P} = (\Pi, \zeta, F, I_{\text{acc}}, I_{\text{idle}}, I_{\text{tr}}, T_{\text{tr}}, \tau_{\text{act}})$ consisting of:

- a finite set of processors $\Pi$ assuming that $\Pi = \{\pi_1, \ldots, \pi_n\}$ is partitioned into disjoint blocks $\Pi_1, \ldots, \Pi_k$ of voltage/frequency islands (VFIs),
- a function $\zeta : \Pi \rightarrow 2^A$ indicating which processors can handle which actors,
- a finite set of discrete frequency levels available to all processors denoted by $F = \{f_1, \ldots, f_m\}$,
- a function $I_{\text{acc}} : \Pi \times F \rightarrow \mathbb{N}$ denoting the operating load current, if the processor $\pi \in \Pi$ is running at a certain frequency level $f \in F$ in the working state,
- a function $I_{\text{idle}} : \Pi \times F \rightarrow \mathbb{N}$ denoting the idle load current, if the processor $\pi \in \Pi$ is running at a certain frequency level $f \in F$ in the idle state,
- a function $I_{\text{tr}} : \Pi \times F^2 \rightarrow \mathbb{N}$ expressing the transition load current, if processor $\pi \in \Pi$ switches the running frequency from $f \in F$ to another frequency $f' \in F$,
- a function $T_{\text{tr}} : \Pi \times F^2 \rightarrow \mathbb{N}$ expressing the time overhead of switching from one frequency level $f \in F$ to another $f' \in F$ for each processor $\pi \in \Pi$, and
- the valuation $\tau_{\text{act}} : A \times F \rightarrow \mathbb{N}_{\geq 1}$ defining the execution time $\tau_{\text{act}}$ of each actor $a \in A$ mapped on a processor at a certain frequency level $f \in F$.

Example 2. Exynos 4210 is a state-of-the-art processor used in high-end platforms such as Samsung Galaxy Note, SII etc, and has 5 DVFS levels [17]. Later in Section V, we consider Exynos 4210 as an use-case for experimental evaluation.

Definition 7. Given an SDF graph $G = (A, D, T_{\text{act}}, \tau)$, a static-order (SO) schedule is a function $\sigma : \Pi \times \mathbb{R} \rightarrow (A \times F) \cup (\perp \times F)$ that assigns to each processor $\pi \in \Pi$ over time, an ordered list of actors or idle slots to be executed at some frequency, where $\perp$ represents the idle slots.

Definition 8. The throughput for a static-order schedule of an SDF graph $G = (A, D, T_{\text{act}}, \tau)$ is the average number of graph iterations that are executed per time unit, measured over a sufficiently long period.

As discussed earlier, in case of more than one battery in the system, the batteries are chosen according to some schedule or scheduling policy. In most systems, the batteries are used sequentially, i.e., only when one battery is empty, the other is used [14]. However, as shown in [14], a specific scheduling scheme termed best-of-all achieves better system lifetime than other schemes. For the same reason, we consider best-of-all in this paper. In this scheduling scheme, after an SDF graph iteration finishes, (i.e., not during the execution of the iteration) the battery having the highest available charge is selected to provide energy for the next iteration.

IV. Translation to Hybrid Automata

A. Hybrid Automata

Timed automata are a popular and powerful formalism to model and analyse real-time systems [4]. TA are state-transition diagrams augmented with real-valued clocks, which can be used in enabling conditions for transitions and in state invariants that enforce deadlines.

Hybrid automata extend timed automata by continuous variables, which we use to model hybrid behaviour of the batteries. Let $X$ be a finite set of continuous variables. A variable valuation over $X$ is a mapping $\nu : X \rightarrow \mathbb{R}$, where $\mathbb{R}$ is the set of reals. We write $\mathbb{R}^X$ for the set of valuations over $X$. Valuations over $X$ evolve over time according to delay functions $F : \mathbb{R}_{\geq 0} \times \mathbb{R}^X \rightarrow \mathbb{R}^X$, where for a delay $d$ and valuation $\nu$, $F(d, \nu)$ provides the new valuation after a delay of $d$.

Definition 9. A hybrid automaton $\mathcal{H}$ is a tuple $(L, \text{Act}, X, E, F, \text{Inv}, l^0)$, where $L$ is a finite set of locations; $\text{Act}$ is a finite set of actions, co-actions and internal A-actions; $X$ is a finite set of continuous variables; $E$ is a finite set of edges of the form $(l, g, a, \varphi, l')$, where $l$ and $l'$ are locations, $g$ is a predicate on $\mathbb{R}^X$, action label $a \in \text{Act}$ and $\varphi$ is a binary relation on $\mathbb{R}^X$; $\text{Inv}$ assigns an invariant predicate $\text{Inv}(l)$ to any location $l$; for each location $l \in L$, $F(l)$ is a delay function; and $l^0 \in L$ is the initial location.

In particular, HA can be analysed by the tool UPPAAL, where each component of the system is described by an automaton whose clocks can evolve with various rates. These rates can be specified with, e.g., ODEs. We utilise UPPAAL engine to perform Monte Carlo simulations to estimate QoS, and to perform model checking.

B. Translation to Hybrid Automata

Our framework consists of separate models of KiBaMs, a KiBaM scheduler, an SDF graph scheduler, and the processor application model. In this way, we divide the problem of evaluating the QoS in terms of power source, tasks and resources. In this section, we describe the translation of an SDF graph scheduler along with a processor application model and KiBaMs to HA using UPPAAL.
Given an SDF graph \( G = (A, D, \text{Tok}_0, \tau) \) mapped on a processor application model \((\Pi, \varsigma, F, T_{tr}, \tau_{act})\) powered by a KiBaM system \( KS = (B, Cap, I_{occ}, I_{idle}, I_{tr}) \), we generate a parallel composition of HA:

\[
K_{sched}[[K_1] \ldots [K_m][K_{obs}][G_{sched1}] \ldots [G_{schedn}]
\]

\[
\Pi = \{\text{Processor1}, \ldots, \text{Processor}_n\}
\]

Here, the automaton \( K_{sched} \) models the scheduling scheme of batteries. This paper considers the best-of-all scheduling scheme, i.e., after every iteration, the battery with the highest available charge is chosen to serve for the next iteration. The HA \( K_{1}, \ldots, K_{m} \) model the batteries \( B = \{\text{bat}_1, \ldots, \text{bat}_m\} \).

The automaton \( K_{obs} \) keeps account of the end of system lifetime by counting the empty batteries. Similarly, the automaton \( G_{sched} \) implements the static-order firing of SDF actors on the processors. The HA \( \text{Processor}_1, \ldots, \text{Processor}_n \) model the processors \( \Pi = \{\pi_1, \ldots, \pi_n\} \). The SDF graph observer \( G_{obs} \) counts if each processor has fired all its mapped actors, according to its static-order schedule. Hence, this automaton determines when an iteration is finished. Note that the resulting hybrid automata is trivially extensible in the number of processors and batteries. Thus, the translation is, at least, composable with regards to the KiBaM system and PAM.

Figure 3 shows the interactions between the HA of different components. The interested readers can see some of the UPPAAL models in Appendix. Due to lack of space, we refer to [2] for more details on translation and UPPAAL models.

V. EXPERIMENTAL EVALUATION VIA MPEG-4 DECODER

We evaluate QoS factors by means of an example of the MPEG-4 decoder in Figure 2. We assume that MPEG-4 decoder is mapped on Exynos 4210 processors. The processors are provided energy by Samsung batteries. In our experiments, we use \( c = 1/6 \) and \( k = 2.2324 \times 10^{-4} \text{s}^{-1} \), as in [14]. The processors are available with two frequency levels (MHz) \( f_2 = 1400 \) and \( f_1 = 1032.7 \) [17]. Table III shows the idle and operating load currents of all processors at both frequencies. We evaluate the system lifetime in terms of completed number of video frames, with respect to various design choices discussed earlier, i.e., varying (1) frames per second (throughput); (2) number of processors; and (3) batteries. For these choices, we assess adequate battery capacities. Due to space limitation, we present adequate battery capacities only for one battery. The SO schedules considered in this section are given in [2].

A. Varying Frames per Second

Let us consider that we have 6 Exynos 4210 processors \( \Pi = \{\pi_1, \ldots, \pi_6\} \) served by two batteries \( B = \{\text{bat}_1, \text{bat}_2\} \). We consider different SO schedules with varying frames per second (fps) constraints. For those SO schedules, Figure 4 shows the total number of video frames per second completed. As we can see from Figure 4, at tighter performance constraints (throughput) when the idle time of processors is not sufficient to move to low power state, the batteries are drained more rapidly. Thus, we achieve fewer iterations. If we require fewer frames per second from an MPEG-4 decoder, then more iterations are acquired.

For the same SO schedules considered earlier, we compute the minimum battery capacities to complete 1000 video frames. The minimum required initial capacity \( Cap(\text{bat}_1) \) for the battery \( \text{bat}_1 \in B \) is shown in Figure 5. It can be seen from Figure 5 that if we relax the frames per second constraint, the minimum required capacity also decreases.

Nevertheless, if the video quality is enhanced from 125 to 200 fps, then the increase in required battery capacity is relatively small, equal to 84 mAh. However, the improvement in the video quality is considerable. Thus, higher performance can also be achieved at the expense of a small increase in the battery capacities, leading to high-performance systems with less mass and volume. Hence, this method allows us to obtain a Pareto front by sweeping throughput constraints, for a fixed number of processors and batteries.

B. Varying Number of Processors

In this case, we consider two batteries, and different SO schedules all of which yield 71 fps. Figure 6 shows the total number of video frames completed for varying number of processors. As we can see from Figure 6, for the same battery capacities, more processors achieve more or equal number of frames. The reason is that, if we reduce the number of processors, the same amount of work is done on fewer processors to attain same throughput, resulting in running on higher frequencies most of the time. Therefore, battery charge is consumed more rapidly, if the number of processors are
achieves low-cost and high-performance systems. than higher number of batteries with smaller capacities. This having fewer batteries with larger capacities is more beneficial capacities at a linear rate. Therefore, we can conclude that, increasing the number of batteries does not reduce the battery improves the attainable number of frames linearly. can be seen from Figure 8, increasing the number of batteries bat batteries, Figure 8 and 9 shows the total number of video further consider a SO schedule producing 71 fps. For varying C. Varying Number of Batteries Let us assume that we have 6 Exynos 4120 processors. We reduced. For the same SO schedules, Figure 7 shows the minimum required capacity $Cap(bat_1)$ to complete 1000 video frames. The results reiterate the earlier conclusions in Figure 6 that, to achieve the same throughput, fewer processors carrying out the work of same magnitude requires larger battery capacities. Hence, using this method, a system designer can estimate QoS for different design alternatives. For instance, in our running example, one can clearly see that we can achieve same throughput for 4 processors, as 5, without requiring extra capacities for batteries. Therefore, we may not need more processors in our platform, and reach a certain throughput with fewer number of processors, and same battery capacities, contributing to low-cost embedded systems with reduced mass and volume.

C. Varying Number of Batteries

Let us assume that we have 6 Exynos 4120 processors. We further consider a SO schedule producing 71 fps. For varying batteries, Figure 8 and 9 shows the total number of video frames completed, and minimum required initial capacity for $bat_1 \in B$ to complete 1000 video frames respectively. As it can be seen from Figure 8, increasing the number of batteries improves the attainable number of frames linearly. However, if we analyse the Figure 9, we can see that increasing the number of batteries does not reduce the battery capacities at a linear rate. Therefore, we can conclude that, having fewer batteries with larger capacities is more beneficial than higher number of batteries with smaller capacities. This achieves low-cost and high-performance systems.

D. Comparison with PTA-KiBaM

In this subsection, we compare the approach presented in this paper (HA-KiBaM) with the PTA-based approach (PTA-KiBaM) [14]. Let us take the example of an MPEG-4 decoder in Figure 2. We assume that there are two batteries $B = \{bat_1, bat_2\}$ in the system. We consider three arbitrary SO schedules, i.e., S1, S2 and S3. It is worth mentioning that the work in [14] evaluates the completed number of tasks, instead of iterations. However, as iterations are a key metric in SDF graphs, we compare both techniques in terms of completed number of iterations (video frames per second). Columns 3-8 in Table IV show the completed number of iterations and computation time (ms), calculated using both methods, against different battery capacities (mAh) in Columns 1-2. The experiments were run on a dual-core 2.8 GHz machine with 8 GB RAM. Table IV shows that HA-KiBaM achieves the same results as PTA-KiBaM except S2. The reason of not producing the same results in S2 is that PTA-KiBaM allows to change the active battery during the iteration. Whereas, we consider a specific scheduling scheme, where we change the battery after an iteration is finished.

However, the biggest advantage of HA-KiBaM is the scale of capacities it can handle. As Table IV shows, PTA-KiBaM can only handle very small battery capacities that are able to finish not more than one video frame. This makes PTA-KiBaM impracticable for modern-day systems, as opposed to our method that scales to much larger capacities (see Section V). Furthermore, PTA-KiBaM requires considerably longer computation time than HA-KiBaM. Please note that zero in Columns 1-2. The experiments were run on a dual-core 2.8 GHz machine with 8 GB RAM. Table IV shows that HA-KiBaM achieves the same results as PTA-KiBaM except S2. The reason of not producing the same results in S2 is that PTA-KiBaM allows to change the active battery during the iteration. Whereas, we consider a specific scheduling scheme, where we change the battery after an iteration is finished.

However, the biggest advantage of HA-KiBaM is the scale of capacities it can handle. As Table IV shows, PTA-KiBaM can only handle very small battery capacities that are able to finish not more than one video frame. This makes PTA-KiBaM impracticable for modern-day systems, as opposed to our method that scales to much larger capacities (see Section V). Furthermore, PTA-KiBaM requires considerably longer computation time than HA-KiBaM. Please note that zero in
its query language. We consider the case-study of an MPEG-4 decoder mapped on Exynos 4210 processors, and powered by a KiBaM system.

**Deadlock**

Checking deadlock freedom is achieved via the UPPAAL query (A[\ not deadlock]). This query allows us to check if a certain static-order schedule is deadlock free or not.

**Parallel firings of actors**

We can check whether any actors can fire in parallel. For example, actors \( p \) and \( q \) mapped on the processors \( \pi_0 \) and \( \pi_1 \) respectively, can fire in parallel if the query \( E \langle SDFScheduler_0.activeActor = p \rangle \) and \( E \langle SDFScheduler_1.activeActor = q \rangle \) evaluates to true. In our experiment, \( p = MC \) and \( q = RC \). As these two actors cannot fire in parallel, the answer to this query turns out to be false.

**Same running frequency in a VFI**

We can also check safety properties such as, at a given time, all processors belonging to the same VFI should not run at the different frequency. For this purpose, we create a variable named “curr_freq” available to all processors, that keeps account of current running frequency of each processor. If we have two processors \( \pi_0 \) and \( \pi_1 \) in a same VFI, then we check the query \( \forall Processor_0.curr_freq = Processor_1.curr_freq \) to verify this property.

In the same way, functional correctness of properties related to KiBaMs can also be verified. However, to verify hybrid properties, UPPAAL offers statistical model checking instead of classical model checking, even though we do not have stochastic properties in our system. In the following, we demonstrate model checking of functional requirements of KiBaMs.

**Fair Scheduling**

We can also verify if only the best battery out of all batteries is selected after each iteration. Let us assume that we have two batteries, and an integer \( empty_count \) to count empty number of batteries. We run the query \( \Pr[\leq 1500] (\langle bound_1 - bound_0 \rangle > n \text{ and } empty_count < 2) \) that determines if difference between the bound charge of two batteries does not exceed more than a certain amount, and each battery gets a fair chance to recover its bound charge. UPPAAL answers that the probability for this query to hold is \([0.0, 0.0973938]\) with 0.95 confidence, which means that this property is not satisfied. In our experiments, \( n \) is 4.

**Active Number of Batteries**

Similarly, we can also check that no more than one battery should be active at any given time. Let us assume that we have two batteries, and boolean variables \( b0\_active \) and \( b1\_active \) is assigned to each battery respectively, to check if that battery is active. To verify this property, we use the query \( \Pr[\leq 1500] (\langle b0\_active \rangle \text{ and } b1\_active \rangle \text{ is true}). The probability for this property to hold is \([0.0, 0.0973938]\) with 0.95 confidence, which means that this property is not satisfied.

**VII. Conclusions**

With the growing gap between the energy demand and battery densities, yet compact methods for guaranteeing QoS of multiple KiBaMs are needed. We have presented a novel technique to predict system lifetime for SDF-modelled streaming applications, mapped on the processors equipped with energy reduction techniques and powered by multiple batteries. This provides us with a best trade-off between the throughput, the number of processors and batteries. The batteries are modelled as a hybrid system, which has the advantage of being accurate.

Future research direction is to explore the possibilities of battery-aware scheduling. We also plan to analyse preemptive scheduling, by having an observer automaton to record the elapsed execution time during the execution of an actor.

**Acknowledgement**

This research is supported by the EU FP7 project SENSATION (318490). The authors would like to thank reviewers for their valuable comments.

**References**

Otherwise, the guard condition on the second transition, i.e., 
\( \text{avail} \) \( \geq \) \( \text{bound} \), as shown in Figure 10a. processors and powered by KiBMs, in Figure 10. checking if the available charge of \( \text{bat}_y \) is, in \( K \) \( \text{active} \) \( \text{bat}_y \). For this purpose, after \( \text{fire} \), \( \text{EndFiring} \), \( \text{totalFirings} \), and \( \text{off} \), the HA \( K_{\text{sched}} \) has a variable: \( \text{active}_{\text{KiBM}} \) \( \text{id} \) that determines the currently active battery. For \( B = \{ \text{bat}_1, \text{bat}_2, \ldots, \text{bat}_m \} \), the initial value of \( \text{active}_{\text{KiBM}} \) \( \text{id} \) = 1, indicating that the battery \( \text{bat}_1 \) is the first to serve. For each battery \( \text{bat}_i \) \( \in \) \( B \) and \( \text{bat}_k \) \( \in \) \( B \), the transition set \( E \) have following transitions.

\begin{align*}
\text{avail}_k & \geq \text{avail}_i, \text{startNextIter}?, \text{active}_{\text{KiBM}} \text{id} = k & \text{avail}_k & \rightarrow \text{avail}_k \\
\text{avail}_k & > \text{avail}_k, \text{startNextIter}? & \text{active}_{\text{KiBM}} \text{id} = k & \rightarrow \text{avail}_k \\
\text{avail}_k & = \text{avail}_k, \text{startNextIter}?. \text{active}_{\text{KiBM}} \text{id} = k & \rightarrow \text{avail}_k \\
\text{avail}_k & < \text{avail}_i, \text{startNextIter}? & \text{active}_{\text{KiBM}} \text{id} = k & \rightarrow \text{avail}_k
\end{align*}

After each iteration finishes, the action \( \text{StartNextIter} \) synchronises with \( G_{\text{obs}} \) to start the new iteration. But, before the new iteration starts, the battery \( \text{bat}_y \) \( \in \) \( B \) with the highest available charge is determined using the guard conditions. This symbolises that only the battery having highest charge is going to serve for the next iteration, and all other batteries are going to stay idle. For the battery \( \text{bat}_i \) \( \in \) \( B \) and \( \text{bat}_k \) \( \in \) \( B \), the guard condition \( \text{avail}_k \geq \text{avail}_i \) on the first transition is checking if the available charge of \( \text{bat}_k \) \( \in \) \( B \) is greater than or equal to \( \text{bat}_i \) \( \in \) \( B \). If the guard condition turns out to be true, then \( \text{bat}_k \) \( \in \) \( B \) provides energy for the next iteration. Otherwise, the guard condition on the second transition, i.e., \( \text{avail}_i > \text{avail}_k \) is satisfied, and \( \text{bat}_i \) \( \in \) \( B \) stays as the active battery.

b) Hybrid Automata \( K_y \): The HA \( K_{\text{sched}} \) \( K_{\text{obs}} \) model the batteries \( B = \{ \text{bat}_1, \ldots, \text{bat}_m \} \) according to the description in Section III-A. The model of \( \text{bat}_y \) \( \in \) \( B \) is shown in Figure 10b. This automaton informs \( K_{\text{obs}} \), when the battery \( \text{bat}_y \) gets empty.

For each \( \text{bat}_y \) \( \in \) \( B \), the HA \( K_y \) is defined as, \( K_y = (L_y, Act_y, X_y, E_y, F_y, Inv_y, l^0_y) \) where \( L_y = \{ \text{Initial, Empty} \} \), and \( l^0_y = \{ \text{Initial} \} \). The automaton \( K_y \) contains two continuous variables \( X = \{ \text{avail}_{\text{bat}_y}, \text{bound}_{\text{bat}_y} \} \) to denote the available and bound charge in \( \text{bat}_y \) \( \in \) \( B \), respectively. There is an urgent broadcast action in \( K_y \), i.e, \( Act_y = \{ \text{empty}! \} \) to synchronise with \( K_{\text{obs}} \). The automaton \( K_y \) contains number of variables: a boolean variable \( on_y \) to determine if the battery has available charge left or whether it has run out of it; and a variable \( i_y \) to annotate the load current being consumed from \( \text{bat}_y \) \( \in \) \( B \). Initially, we have \( on_y = \text{true} \) and \( i_y = 0 \). The transition set \( E_y \) has only transition, given as follows.

\begin{align*}
\text{Initial} & \rightarrow \text{on}_y, \text{on}_y = \{ \text{true} \} \rightarrow \text{on}_y = \{ \text{false} \} \rightarrow \text{Empty}
\end{align*}

The above transition synchronises with \( K_{\text{obs}} \) over the urgent channel \( \text{empty}! \), and is taken if the available charge \( \text{avail}_y \) reaches or falls below zero, emphasising that the battery \( \text{bat}_y \) \( \in \) \( B \) is empty. As a result of this action, the value of \( on_y \) changes to false.

The initial location \( l^0_y \) uses equations (1) and (2) as a delay function. This represents that, as long as \( \text{bat}_y \) \( \in \) \( B \) is non-empty, the available and bound charge of \( K_y \) evolves according to equations (1) and (2) respectively.

c) Hybrid Automata \( G_{\text{sched}} \): The HA \( G_{\text{sched}} \) implement the static-order firing of SDF actors on the processors. For this purpose, after \( G_{\text{obs}} \) informs \( G_{\text{sched}} \) that an iterations has started, \( G_{\text{sched}} \) map actors on \( Processor \) according to the SO schedule of that processor. When all actors are fired according to the SO schedule on \( Processor \), \( G_{\text{sched}} \) inform \( G_{\text{obs}} \) back, indicating the end of current iteration. For a \( \pi_j \in \Pi \), Figure 10c presents the automaton \( G_{\text{sched}} \), with respect to our running example.

For each \( \pi_j \in \Pi \), \( G_{\text{sched}} \) is defined as, \( G_{\text{sched}} = (L_j, Act_j, X_j, E_j, F_j, Inv_j, l^0_j) \) where \( L_j = \{ \text{start, FireActor, EndFiring, totalFirings, off} \} \) and \( l^0_j = \{ \text{start} \} \). The HA \( G_{\text{sched}} \) contain three broadcast actions, i.e., \( Act_j = \{ \text{fire}, \text{end}, \text{startNextIter} \} \). The actions \( fire \) and \( end \) are parametrised with processor and action ids, and are used to synchronise with \( Processor \). The action \( \text{StartNextIter} \) synchronises with \( G_{\text{obs}} \). The actions \( fire \) and \( \text{StartNextIter} \) are the urgent actions. There are no clocks and invariants in \( G_{\text{sched}} \). There are no delay functions and continuous variables in \( G_{\text{sched}} \). The HA \( G_{\text{sched}} \) have a number of local variables: \( active_{\text{Actor}} \) that determines the active actor currently mapped on the processor \( \pi_j \) and \( s_j \) that determines the index of the active actor in the static-ordered list. Initially, \( active_{\text{Actor}} = 0 \) and \( s_j = 0 \). The HA \( G_{\text{sched}} \) also contain a parametrised variable \( \text{totalFirePerProc} \), that defines the total number of tasks in the SO schedule of the processor \( \pi_j \). Since these variables are local, we can abbreviate them by \( active_{\text{Actor}}, s \) and \( \text{totalFirePerProc} \) respectively. The transition set \( E_j \) has following transitions.
The following transition fetches the active actor according to the SO schedule for each processor \( \pi_j \), using the function \( \text{getReadyActor}(j) \). As a result of this transition, the value of \( s \) is incremented by 1, which means that the next actor in the SO schedule is fetched next time.

\[
\text{Start} \xrightarrow{\emptyset, \emptyset, \text{activeActor}:=\text{getReadyActor}(j) \land s++} \text{fireActor}
\]

- The following transition maps the fetched (active) actor, on the processor automaton \( \text{Processor}_j \), using the urgent channel \( \text{fire!} \).

\[
\text{fireActor} \xrightarrow{\emptyset, \text{fire}[j][\text{activeActor}]!, \emptyset} \text{endFiring}
\]

- In the following transition, the urgent action \( \text{end?} \) synchronises with the processor automaton \( \text{Processor}_j \). As a result, the processor automaton \( \text{Processor}_j \) informs the automaton \( G_{\text{sched}_j} \) that the firing of the active actor has finished.

\[
\text{endFiring} \xrightarrow{\emptyset, \text{end}[j][\text{activeActor}]?, \emptyset} \text{totalFirings}
\]

- The following transition checks if the SO schedule of a processor \( \pi_j \) is not fully executed, using the guard condition \( s < \text{totalFirePerProc} \). If this is the case, the following transition is taken, leading to the Start location where the next actor in the SO schedule is fetched.

\[
\text{totalFirings} \xrightarrow{s<\text{totalFirePerProc}, \emptyset, \emptyset} \text{Start}
\]

- If all actors in the SO schedule of a processor \( \pi_j \) are executed as checked by the guard condition \( s == \text{totalFirePerProc} \) on the following transition, the urgent channel \( \text{FiringFinished!} \) synchronises with the observer automaton \( G_{\text{obs}} \). In this way, \( G_{\text{sched}_j} \) informs \( G_{\text{obs}} \) that the processor \( \pi_j \) has executed all of the mapped actors in the current iteration. The variable \( s \) is also reset.

\[
\text{totalFirings} \xrightarrow{s==\text{totalFirePerProc}, \text{FiringFinished!}, s:=0} \text{allFired}
\]

- The following transition synchronises with the observer automaton \( G_{\text{obs}} \) on the urgent channel \( \text{StartNextIter?} \) to start executing the static-order schedule of the next iteration.

\[
\text{allFired} \xrightarrow{\emptyset, \text{StartNextIter?}, \emptyset} \text{Start}
\]

d) Hybrid Automata \( \text{Processor}_j \): Likewise, the HA \( \text{Processor}_1, \ldots, \text{Processor}_n \) model the processors \( \Pi = \{\pi_1, \ldots, \pi_n\} \), as shown in Figure 10d. For better visibility, Figure 10d shows the HA of \( \text{Processor}_j \), with respect to one actor only, i.e., \( FD \in A \).

For each \( \pi_j \in \Pi \), we define HA \( \text{Processor}_j = \)
The action fire[\pi]|\text{idle}_{fi}\) is enabled in the initial state Initial and leads to the location Idle_{fi}. Thus, fire[\pi]|\text{idle}_{fi}\) causes the processor \(\pi \in \Pi\) to go to Idle_{fi} at frequency level \(f_i \in F\), whenever the processor \(\pi \in \Pi\) is supposed to stay idle at \(f_i \in F\) in the static-order schedule. As the idle slots are divided into time slots of one time unit, each location InUse_{a,fi} has an invariant \(\text{Inv}_j(\text{InUse}_{a,fi})\leq 1\), the automaton can stay in InUse_{a,fi} for at most \(1\) time unit. The function selectBattery IdleFire_{\pi}\() consumes the charge from the active battery, i.e., \(I_{\text{idle}}(\pi, f_i, I)\). If \(x = 1\), the system has to leave Idle_{fi} at exactly one time unit, by taking the end[\pi]|\text{idle}_{fi}\) action.

For \(F = \{f_1, \ldots, f_m\}\) such that \(f_1 < f_2 < \ldots < f_i < f_m\), and \(\pi_j \in \Pi\), the transition set \(E_j\) has following transitions such that:

- Initial \(\emptyset, \text{fire}[\pi]|\text{f1}_{f2}\?; (x:=0) \land \text{selectBatteryTrFire}_{\text{f1}_{f2}}\)
- \(\text{Tr}_{f1}_{f2}\)
  \(x := T_{\mu}(\pi, f_1, f_2), \text{end}[\pi]|\text{f1}_{f2}!(i), \text{selectBatteryTrEnd}_{f1}_{f2}\)
- Initial \(\emptyset, \text{fire}[\pi]|\text{f1}_{f2}\?; (x:=0) \land \text{selectBatteryTrFire}_{f1}_{f2}\)
- \(\text{Tr}_{f1}_{f2}\)
  \(x := T_{\mu}(\pi, f_1, f_2), \text{end}[\pi]|\text{f1}_{f2}!(i), \text{selectBatteryTrEnd}_{f1}_{f2}\)
- Initial \(\emptyset, \text{fire}[\pi]|\text{ff}_{fm}\?; (x:=0) \land \text{selectBatteryTrFire}_{ff}_{fm}\)
- \(\text{Tr}_{ff}_{fm}\)
  \(x := T_{\mu}(\pi, f_m, f_m), \text{end}[\pi]|\text{ff}_{fm}!(i), \text{selectBatteryTrEnd}_{ff}_{fm}\)
- Initial \(\emptyset, \text{fire}[\pi]|\text{ff}_{fm}\?; (x:=0) \land \text{selectBatteryTrFire}_{ff}_{fm}\)
- \(\text{Tr}_{ff}_{fm}\)
  \(x := T_{\mu}(\pi, f_m, f_m), \text{end}[\pi]|\text{ff}_{fm}!(i), \text{selectBatteryTrEnd}_{ff}_{fm}\)
- Initial \(\emptyset, \text{fire}[\pi]|\text{ff}_{fm}\?; (x:=0) \land \text{selectBatteryTrFire}_{ff}_{fm}\)
- \(\text{Tr}_{ff}_{fm}\)
  \(x := T_{\mu}(\pi, f_m, f_m), \text{end}[\pi]|\text{ff}_{fm}!(i), \text{selectBatteryTrEnd}_{ff}_{fm}\)
- Initial \(\emptyset, \text{fire}[\pi]|\text{ff}_{fm}\?; (x:=0) \land \text{selectBatteryTrFire}_{ff}_{fm}\)
- \(\text{Tr}_{ff}_{fm}\)
  \(x := T_{\mu}(\pi, f_m, f_m), \text{end}[\pi]|\text{ff}_{fm}!(i), \text{selectBatteryTrEnd}_{ff}_{fm}\)
- Initial \(\emptyset, \text{fire}[\pi]|\text{ff}_{fm}\?; (x:=0) \land \text{selectBatteryTrFire}_{ff}_{fm}\)

The action fire[\pi]|\text{ff}_{fm}\) causes the processor \(\pi \in \Pi\) to incur the transition overhead, whenever the processor \(\pi \in \Pi\) is supposed to change the frequency \(f_i \in F\) to \(f_m \in F\) in the static-order schedule, using the function selectBattery TrFire_{ff}_{fm}, and so on.

**e) Hybrid Automaton \(G_{obs}\):** The automaton of SDF graph observer \(G_{obs}\) counts if each processor has fired its all mapped actors in an static-order schedule, observing the number of iterations finished. Figure 10e shows the HA model of \(G_{obs}\). The automaton \(G_{obs}\) has an integer variable Tot_Iter to count the number of finished iterations. Initially, Tot_Iter = 0.

After modelling the whole system, we run the following query, where bound is the time bound on running the simulation, and Tot_Iter is the variable representing the completed number of iterations. As a result, we get a plot, by which we determine the total number of iterations completed within bound time units. We use the same models and query to determine the adequate batteries’ capacities.

```
simulate 1[<= bound]|TOT_ITER
```