Adaptivity in High-Performance Embedded Systems: a Reactive Control Model for Reliable and Flexible Design
Huafeng Yu, Abdoulaye Gamatié, Éric Rutten, Jean-Luc Dekeyser

To cite this version:

HAL Id: inria-00536883
https://hal.inria.fr/inria-00536883
Submitted on 18 Nov 2010

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.
Adaptivity in High-Performance Embedded Systems: a Reactive Control Model for Reliable and Flexible Design

Huafeng Yu\textsuperscript{1} and Abdoulaye Gamatié\textsuperscript{2} and Éric Rutten\textsuperscript{3} and Jean-Luc Dekeyser\textsuperscript{4}

\textsuperscript{1}INRIA Rennes/IRISA, Campus de Beaulieu, 263, avenue du Général Leclerc, 35042 Rennes, FRANCE
E-mail: huafeng.yu@inria.fr
\textsuperscript{2}CNRS/LIFL, INRIA Lille Nord Europe, 40 avenue Halley, 59650 Villeneuve d'Ascq, FRANCE,
E-mail: abdoulaye.gamatie@lifl.fr
\textsuperscript{3}INRIA Rhône-Alpes, 655 avenue de l’Europe, Montbonnot, 38334 Saint-Ismier cedex, FRANCE
E-mail: eric.rutten@inria.fr
\textsuperscript{4}USTL/LIFL, INRIA Lille Nord Europe, 40 avenue Halley, 59650 Villeneuve d’Ascq, FRANCE,
E-mail: jean-luc.dekeyser@lifl.fr

Abstract

System adaptivity is increasingly demanded in high-performance embedded systems, particularly in multimedia System-on-Chip (SoC), due to growing Quality of Service requirements. This paper presents a reactive control model that has been introduced in Gaspard, our framework dedicated to SoC hardware/software co-design. This model aims at expressing adaptivity as well as reconfigurability in systems performing data-intensive computations. It is generic enough to be used for description in the different parts of an embedded system, e.g. specification of how different data-intensive algorithms can be chosen according to some computation modes at the functional level; expression of how hardware components can be selected via the usage of a library of Intellectual Properties (IPs) according to execution performances. The transformation of this model towards synchronous languages is also presented, in order to allow an automatic code generation usable for formal verification, based of techniques such as model checking and controller synthesis as illustrated in the paper. This work, based on Model-Driven Engineering and the standard UML MARTE profile, has been implemented in Gaspard.

1 Introduction

With the popularization of mobile multimedia devices such as PDA, multimedia cellular phone and MP3 player, high-performance embedded systems (HPES’s) attract increasingly interests in both industry and academia. These systems feature intensive data processing, including multimedia video codecs, software-based radio and radar signal processing systems. These applications generally require high-performance computing resources, where parallel processing is a key feature. In addition, they concentrate on regular data partitioning, distribution and access.

1.1 Design complexity of High-performance SoCs

The previously mentioned applications are usually developed by using high-performance computing (HPC) programming languages that provide useful concepts to deal with parallel processing, in order to meet their real-time requirements. There exist several HPC programming languages and one of the most successful is High Performance Fortran (High Performance Fortran Forum, 1997). However, these languages tend to be very specific to users and lack features, which are increasingly required for the design of modern HPES’s: design of complex hardware topologies comprising multi-core architectures,
of hardware/software allocations, of intellectual property (IP) deployment enabling reusability, etc. In addition, they do not necessarily allow system verification, which is important for guaranteeing the reliability of the systems.

On the other hand, the system complexity issue is another major obstacle faced by System-on-Chip (SoC) designers. As computational power increases for SoCs, more functionality are expected to be integrated into these systems. Hence, more complex software applications and hardware architectures are involved. The resulting consequence leads to the disequilibrium in system design, particularly software design does not progress at the same pace as that of hardware. This has become a critical issue and has finally led to the productivity gap (Semiconductor Industry Assoc., 2004).

In order to address the above challenges, many efforts have been recently carried out to better address the SoC design (Sangiovanni-Vincentelli, 2007). Raising the levels of abstraction is considered as an effective solution to reduce the overall complexity in the design. High-level modeling approaches have been developed such as Model-Driven Engineering (MDE) (Object Management Group, 2007a). MDE also enables high-level system design (of both software and hardware) with the possibility of integrating heterogeneous components into the system. Furthermore, model transformations enable generation of executable code from high level models. MDE is also supported by large number of existing standards and tools, for instance, UML-based modeling and transformation tools (UML tool list, 2009).

1.2 Adaptivity in multimedia SoCs

System adaptivity specification as well as reconfigurable computing description are considered to be critical in current embedded systems design, particularly in multimedia SoC design. Such systems must be able to cope with end user requirements and environments. Being adaptive and/or reconfigurable to the environment is highly demanded due to current flexibility and Quality-of-Service (QoS) requirements. However, integration of adaptivity and reconfigurability into a system is expected not to affect system performance, especially regarding real-time constraints.

Mode-based control modeling, as one kind of behavioral specifications, plays an important role in multimedia systems, as it satisfies the following requirements that are becoming inevitable in mobile embedded systems: 1) mode changes specified in functionality, for example, color or monochrome mode for video effects; 2) mode changes due to resource constraints of targeted hardware/platforms, for instance switching from a high CPU load mode to a smaller one; or 3) mode changes due to other environmental criteria such as energy consumption constraint. Mode-based control integrates flexibility in the design. Consequently, it offers better QoS choices to designers/end users. Thanks to the above characteristics, mode-based control mechanisms are good candidates for the integration of adaptivity and/or reconfigurability into the high-performance systems design seamlessly. First, modes of data processing can be switched due to the requirements of users, platforms and environments. Secondly, data access, as a critical factor of performance, is kept unchanged so that mode changes have little influence on performance.

1.3 Contribution and outline

We present a high-level model allowing for the mixed description of data-intensive and control-oriented behaviors, as a solution to for the design of HPES’s. This model enables to:

- effectively handle issues such as adequate expression of inherent system parallelism both in functional specification and hardware architecture;
- express control behavior in order to characterize adaptivity and reconfigurability in a system in consideration of performance influence.

The control model is proposed in Gaspard, which is an MDE-based SoC co-design framework dedicated to HPES’s (Gamatié et al., 2010), and defined with the UML profile for MARTE (Object Management Group, 2008).
Adaptivity in High-Performance Embedded Systems

A transformation of our model towards synchronous reactive languages is also described, in order to obtain access to their formal validation technologies. It is formalized on the basis of an abstract syntax. Typical applications of our model for system behavior analysis and hardware synthesis are also briefly mentioned applications.

The rest of this paper is organized as follows. Some related works are discussed in Section 2. Then, an overview of Gaspard is provided in Section 3. Section 4 concentrates on the extension of Gaspard with a control model, which includes its abstract syntax, while Section 5 presents the transformation of this control model into synchronous reactive programs. Section 6 presents the implementation of the control in the framework of MDE. The application of this control model in functional specification and IP deployment are briefly described in Section 7. Finally, Section 8 gives the conclusion.

2 Related works

Programming languages for the specification of high-performance systems has already been broadly studied, for example, Message Passing Interface (MPI) (MPI Forum, 2007) and OpenMP (OpenMP API, 2008). More recent languages include StreamIt (Thies et al., 2002) and the DARPA high productivity languages Chapel (Callahan et al., 2004), Fortress (Allen et al., 2007) and X10 (Charles et al., 2005). Improving productivity is one of their main objectives, such as reducing specification complexity by providing suitable macro constructs. Furthermore, they consider specific architecture models. However, they are not well adapted for SoC design in which one needs to program specialized architectures.

Alpha (Wilde, 1994) and Array-OL (Boulet, 2008) (core formalism adopted in Gaspard), which manipulates polyhedra and arrays respectively, are interesting high level languages for the data-parallel formalism. Alpha particularly suits for the specification of systolic architectures while the aim of Gaspard is to cover embedded systems adopting a wide variety of architectures beyond systolic ones.

In several tools and development environments dedicated to dataflow applications, the multi-paradigm has been proposed to integrate languages in different styles, i.e., dataflow and some imperative features based on finite state machines. This approach benefits from their different expressivity: dataflow specification for numerical computation and state machines for control behavior. Some examples are given here:

- **SIMULINK and STATEFLOW** (The MathWorks, 2009): the former is used for the specification of block diagrams where some operators of STATEFLOW are used to specify the computation on dataflow. The results of the computation serve to control the system.

- **SCADE** (Esterel Technologies, 2009): in the SCADE environment, state machines are embedded and used to activate dataflow processing specified in Lustre (Halbwachs et al., 1991). Mode automata (Maraninchi and Rémond, 2003) and polychronous mode automata (Talpin et al., 2006) derive from the concept of combination of formalisms in synchronous languages. It extends the dataflow language Lustre and Signal respectively with certain imperative style, but without many modifications of language style and structure. Applications specified in SCADE or with mode automata and polychronous mode automata can be formally validated using tools associated with synchronous languages. Although this approach is very similar to our proposition, however, Lustre and Signal are limited in the expression of large parallel data processing.

- **PTOLEMY**: the composition of hierarchical finite state machines with some concurrency models has also been studied in (Girault et al., 1999). Among them, the composition of state machines with synchronous dataflow is similar to our proposition, and the latter can be considered as a special case applied on parallel data-intensive processing specified with Repetitive Structure Modeling (RSM) package of MARTE. The novelty is exhibited in the compositionality of state machines with data parallelism and data access performance conservation in case of mode changes.

In comparison with the above mentioned works, our work focuses on embedded systems with high-performance computing, where the way data-parallelism is dealt with is very important. We adopt mode automata where the mode concept helps to reduce performance loss via mode switch during data-intensive
computations: the interface of data processing is kept the same even if modes are changed, so data access are not changed allowing one to still benefit from the powerful way of expressing the data-parallelism in Gaspard.

3 High-performance system design within MARTE

Gaspard (Gamatié et al., 2010) is a MARTE compliant SoC design framework, that enables high-level modeling and automatic code generation by using UML graphical tools and technologies such as Papyrus\(^1\) and Eclipse Modeling Framework\(^2\) (EMF). Gaspard features hardware/software co-modeling using the MARTE profile, which enables to model *software applications*, *hardware architectures*, *their allocations* and *IP deployment* separately, but in a unique modeling environment. As high-level Gaspard models contain only domain-specific concepts, model transformations (implemented as Eclipse plugins) enable code generation for different execution platforms, such as synchronous domain for validation and analysis purposes (Gamatié et al., 2008b); or FPGA synthesis (Quadri et al., 2010). Thus technological concepts are introduced seamlessly in the intermediate and low level models.

The MARTE RSM is inspired from Gaspard. RSM is based on Array-OL (Boulet, 2007) that describes the potential parallelism in a system; and is dedicated to data-parallel multidimensional signal processing. Manipulated data in Gaspard and RSM are in the form of multidimensional arrays. RSM allows to describe the regularity of a system’s structure (composed of repetitions of structural components interconnected in a regular connection pattern) and topology in a compact manner. Gaspard adopts RSM for the specification of regular hardware architectures (such as multiprocessor architectures) and parallel processing applications. In addition, both data parallelism and task parallelism can be, via RSM, expressed with regard to a functional specification.

An example specified with RSM is shown in Figure 1. It expresses data-parallelism in *MonochromeMode*, used for the processing of a [320, 240]-image. Because the filter, called *MonoFilter*, only works on small [8, 8]-pixel subsets, it should be repeated 40×30 times to cover the whole image. In RSM [40, 30] is referred to as the shape of repetition space associated with MonoFilter.

![Figure 1: A monochrome effect filter.](image-url)

The repeated MonoFilter runs in a repetition context, defined by the MonochromeMode. All repetition instances run in parallel. A Connector used in a repetition context is called *LinkTopology*. It adds a set of topological information to a connector. MonoFilter is connected to MonochromeMode via *Tiler* links. The repetitions of MonoFilter consume and produce identically shaped sub-arrays of pixels, which are respectively extracted from the input port *I_image* and stored in the output port *O_image*. These sub-arrays, referred to as patterns (shaped [8, 8] in the example), are tiled according to the Tilers, which are associated with each array (i.e. each edge in the graphical representation). A tiler extracts (resp. stores) tiles from (resp. in) an array based on some information: \(F\) a *fitting* matrix (how array elements fill the tiles), \(o\) the *origin* of the reference tile (for the reference repetition), and \(P\) a *paving* matrix (how the tiles cover arrays).

\(^1\)www.papyrusuml.org/
\(^2\)www.eclipse.org/emf/
The repetition space indicating the number of task instances is itself defined as a multidimensional array with a shape. Each dimension of the repetition space can be seen as a parallel loop and its shape space gives the bounds of the nested parallel loops. In Figure 1, the shape of repetition space is $[40, 30]$.

Given a tile, let its reference element denote its origin point from which all its other elements can be extracted. The fitting matrix is used to determine these elements. Their coordinates, represented by $e_i$, are built as the sum of the coordinates of the reference element and a linear combination of the fitting vectors, the whole modulo the size of the array (since arrays are toroidal) as follows:

$$\forall i, 0 \leq i < s_{\text{pattern}}, e_i = \text{ref} + F \times i \mod s_{\text{array}}$$

where $s_{\text{pattern}}$ is the shape of the pattern, $s_{\text{array}}$ is the shape of the array and $F$ is the fitting matrix.

For each repetition instance, the reference elements of the input and output tiles are needed to be specified. The reference elements of the reference repetition are given by the origin vector, $o$, of each tiler. The reference elements of the other repetitions are built relatively to this one. As above, their coordinates are built as a linear combination of the vectors of the paving matrix as follows:

$$\forall r, 0 \leq r < s_{\text{repetition}}, \text{ref}_r = o + P \times r \mod s_{\text{array}}$$

where $s_{\text{repetition}}$ is the shape of the repetition space, $P$ the paving matrix and $s_{\text{array}}$ the shape of the array.

Gaspard adopts a component-based approach. Repetitive components (MonochromeMode in Figure 1) are used to express data parallelism in an application. Sets of input and output patterns (scope of involved data in array structure) consumed and produced by the repetitions of the interior part (a data processing task). Thus a repetitive component provides the repetition context for its interior task. The repetitions are assumed to be independent and schedulable following any order, generally in parallel. In comparison, a hierarchical component contains several parts, and it allows to define complex functionality in a modular way and provides a structural aspect of the application. Specifically, task parallelism can be described using such a component. The shape of a pattern is described according to a tiler connector, which specifies the tiling of produced and consumed arrays. An inter-repetition dependency (IRD) is used to specify an acyclic dependency among the repetitions of the same component. Particularly, an IRD connector leads to the sequential execution of repetitions. A default link provides a default value for repetitions linked with an IRD, on condition that the source of dependency is absent.

4 Definition of the reactive control model

This paper presents is a control model for adaptive high-performance computing in the framework of Gaspard: basic concepts are described, and both parallel and hierarchical composition is formally defined as well as synchronous reactive semantics is integrated in order to benefit from existing formal validation and synthesis tools (Gamatié et al., 2008b; Gamatié et al., 2009). In addition to functional specification, this control model is also adopted in IP deployment for reconfigurable FPGAs, which is detailed in (Quadri et al., 2010).

4.1 Mode-based control modeling

Control behavior in Gaspard is expressed through a mode-switch based control model, which derives from mode automata (Maraninchi and Rémond, 2003). The notion of exclusion between modes helps to separate different computations in a modular way. As a result, programs are well structured and fault propagation from one mode to another is reduced. The control model is mainly composed of concepts such as Mode Switch and State Graphs.

4.1.1 Mode switch and modes

A mode switch contains at least one mode; and offers a switch functionality that chooses one mode to execute, among several alternative present modes (Labbani et al., 2005). The mode switch in Figure 2a illustrates such a component having a window with multiple tabs and interfaces. For instance, it has an $m$
The modes, $M_1, ..., M_n$, in the mode switch are identified by the mode values: $m_1, ..., m_n$. The switch between the modes $M_i$ is carried out according to the mode value received through the port $m$. Each mode can be hierarchical or elementary in nature. All modes have the same interface (i.e., $i_d$ and $o_d$ ports).

### 4.1.2 State graphs

A Gaspard state graph (Figure 2b), is similar to Statecharts (Harel, 1987) and its synchronous variant SyncCharts (Andre, 2004) and mode automata (Maraninchi and Rémond, 2003), which are used to model the system behavior using a state-based approach. A Gaspard state graph is a sextuplet $(Q, V_i, V_o, T, M, F)$ where:

- $Q$ is the set of states defined in the state graph;
- $V_i, V_o$ are the sets of input and output variables, respectively. $V_i$ and $V_o$ are disjoint, i.e., $V_i \cap V_o = \emptyset$. $V$ is the set of all variables, i.e., $V = V_i \cup V_o$;
- $T \subseteq Q \times C(V) \times Q$ is the set of transitions, labeled by conditions on the variables of $V$. The conditions $C$ are Boolean expressions on $V$;
- $M$ is the set of modes defined in the state graph;
- $F \subseteq Q \times M$ represent a set of surjective mappings between $Q$ and $M$.

Gaspard state graph is considered as a graphical representation of transition functions as discussed in (Gamatié et al., 2008a), hence there are no initial states defined. Each state is associated with some mode value specifications, and Transitions are conditioned by some events or Boolean expressions.
4.1.3 Composing mode switch and state graph

A component, whose behavior is exhibited by Gaspard state graphs, is called **state graph component**. A mode switch component indicates a component implementing a mode switch. Then a *macro* component can be used to compose state graph component and mode switch components together. The macro in Figure 2c illustrates one possible composition, i.e., other compositions are also possible. For instance, one state graph component can be associated with several mode switch components. In the macro, the state graph component produces a mode value (or a set of mode values) and sends it (them) to the mode switch component. The latter switches the modes accordingly.

4.2 Composition definitions

Composition of the control model is detailed in this paper due to the requirements of expressivity and verifiability. The parallel and hierarchical composition has been formally defined so that complex systems can be specified and system behavior is clear and verifiable by some formal verification tools. With the clear composition definition, it is also possible to abstract the control part of a system in order to make the verification more efficient.

4.2.1 Parallel composition

The parallel composition can be specified in two ways: composing state graphs directly or composing state graph components. The first one is similar to the parallel composition defined in SyncCharts and mode automata. The second one is considered as normal Gaspard component composition.

A set of state graphs (Figure 3) can be composed together. The composition of state graphs in this manner is considered as **parallel composition**. Figure 3 illustrates a parallel composition example. The two state graphs are placed in the same ellipse, but separated by a dash line.

![Parallel Composition Diagram](image)

Figure 3: An example of the parallel composition of state graphs.

Two state graphs $G^1_1: = (Q^1, V^1_i, V^1_o, T^1, M^1, F^1)$ and $G^2_2: = (Q^2, V^2_i, V^2_o, T^2, M^2, F^2)$ are considered here to illustrate the composition. The Composition operator is noted as: $\parallel$. The parallel composition is defined as:

$$
(Q^1, V^1_i, V^1_o, T^1, M^1, F^1) \parallel (Q^2, V^2_i, V^2_o, T^2, M^2, F^2) = ((Q^1 \times Q^2), (V^1_i \cup V^2_i)(V^1_o \cup V^2_o), (V^1_o \cup V^2_o), T, (M^1 \times M^2), F),
$$

where

$$
F = \{(q^1, q^2), (m^1, m^2) | (q^1, m^1) \in F^1 \wedge (q^2, m^2) \in F^2 \}
$$
and
\[ T = \{(q^1, q^2), C, (q'^1, q'^2)\} | (q^1, C^1, q'^1) \in T^1 \land (q^2, C^2, q'^2) \in T^2 \}, \]
where \( C \) is a new expression on \( C^1 \) and \( C^2 \): \( C = C^1 \) and \( C^2 \), and \( (V^1_o \cup V^2_o) \setminus (V^1_i \cup V^2_i) \) implies any output variable is not considered as an input variable in the composed graphs, hence it should be removed from the input variable set.

These state graphs in a parallel composition can be triggered to carry out transitions at the same time, upon the presence of the source states and events of both two state graphs. Moreover, the number of transitions fired is supposed to be always the same for these state graphs, i.e., the inputs of these state graphs have the same size.

As state graph components are considered as normal Gaspard components, the resulting component composed of several state graph components can be also considered as a normal Gaspard component. The composition of state graph components is illustrated with an example in Figure 4. State graph components can also be composed with other Gaspard components. But note that the interfaces of these state graph components should be coherent, i.e., it ensures the same transition rate of state graphs. In addition, state graph components can also be composed with standard Gaspard components, which can specify the control out of capacity of state graphs, e.g., some binary operations on events or conditions on numbers.

**Figure 4:** Parallel composition of state graph components in Gaspard.

### 4.2.2 Hierarchical composition

Hierarchical composition of state graphs is defined similar to that of SyncCharts and mode automata. A state in the state graph \( A \) can be refined by another state graph \( B \), where \( B \) is considered as sub graph of \( A \). Consider a state graph \( G = (Q, V_i, V_o, T, M, F) \) where \( Q = \{q_0, q_1, ..., q_n\} \) and a corresponding set of refining automata \( \{G_k\}_{k \in [0, n]} \), where \( G_k = (Q^k, V_i^k, V_o^k, T^k, M^k, F^k) \). The composition can be defined (inspired from (Maraninchi and Rémond, 2003)) as:

\[ G \triangleright \{G^k\}_{k \in [0, n]} = (Q', V'_i, V'_o, T', M', F') \]

where
\[ Q' = Q \triangleright \{Q^k\}_{k \in [0, n]} = \bigcup_{k=0}^{n} \{q^k \triangleright q^k_j \mid j \in [a, n^k]\} , \]
is refined by the state graph composed of states $S$. Figure 5 shows an example of the hierarchical composition of state graphs. In this example, the state

$$ V'_o = V_o \cup \bigcup_{k=0}^n V^k_o, $$

$$ V'_i = (V_i \cup \bigcup_{k=0}^n V^k_i) \setminus V'_o, $$

$$ T' = \{(q^k \triangleright q^k_1, C, q^d) \mid (q^k, C, q^d) \in T \land j_1 \in [0, n^k]\} \cup \{(q^k \triangleright q^k_1, (C \land \neg\bigvee_{(q^k, C, q^d) \in T} C_m), q^k \triangleright q^k_2) \mid (q^k_1, C, q^k_2) \in T^k \land j_2 \in [0, n^k]\}, $$

$$ M' = M \cup \bigcup_{k=0}^n M^k, $$

and

$$ F' = \{((q, q^1, \ldots, q^n), (m, m^1, \ldots, m^n)) \mid (q, m) \in F \land (q^1, m^1) \in F^1 \land \ldots \land (q^n, m^n) \in F^n\}. $$

$q^k \triangleright q^k_1$ denotes the state $q^k$ is refined by the state $q^k_1$. $T'$ has two kinds of transitions: $\{(q^k \triangleright q^k_1, C, q^d) \mid (q^k, C, q^d) \in T \land j_1 \in [0, n^k]\}$ implies the transitions of $G$, and $q^k$ and $q^d$ denotes the source and target state respectively. As $q^k$ is refined by states of $G^k$, $q^k \triangleright q^k_1$ is used instead of $q^k$. The target state is also a refined state of $q^d$, however, which state is entered is decided at run time. Hence, $q^d$ is used instead. $\{(q^k \triangleright q^k_1, (C \land \neg\bigvee_{(q^k, C, q^d) \in T} C_m), q^k \triangleright q^k_2) \mid (q^k_1, C, q^k_2) \in T^k \land j_2 \in [0, n^k]\}$ denotes the transitions in $G^k$, i.e., these transitions are fired when no transitions of $G^k$ are fired. This condition is expressed by: $\neg\bigvee_{(q^k, C, q^d) \in T} C_m$. $q^k_1$ and $q^k_2$ denote the source and target states of a transition in $G^k$. Figure 5 shows an example of the hierarchical composition of state graphs. In this example, the state $S_3$ is refined by the state graph composed of states $S_4$ and $S_5$, denoted by $S_3 \triangleright S_4$ and $S_3 \triangleright S_5$.

![Figure 5: An example of hierarchical composition of state graphs.](image)

Hierarchical composition using both state graph components and mode switch components can be also achieved. In this case, an state graph component acts as a mode in a mode switch component, thus it is only activated in this mode. The mode switch component therefore defines the collaboration between its corresponding state graph component and other state graph components acting as modes, where state graphs of the latter act as state refinements of state graphs of the former. An example is illustrated in Figure 6. The state graph associated with $SGT_1$ has three states: $S_1$, $S_2$ and $S_3$, which correspond to three modes $M_1$, $M_2$ and $M_3$ respectively. In the $MST$, $SGT_2$ is defined as the mode $M_3$, i.e., $SGT_2$ is activated only when $S_3$ is active. Hence the state graph of $SGT_2$ is considered as a refinement (sub state graph) of $SGT_1$’s $S_3$.

### 4.3 Reactive semantics

Basic Gaspard control constructs have been presented herein before, but its explicit semantics is still not given. We propose to integrate reactive semantics, inspired from mode automata (Maraninchi and Rémond, 2003), in order to confer certain expected properties related to safe design as well as enhance design correctness verifiability. The integration result of the control model is called Gaspard mode
automata. The basic structure of Gaspard mode automata is presented by a composition of state graph component and mode switch component, i.e., the Macro in Figure 2c. The state graph component in this macro structure acts as a state-based controller and the mode switch component achieves the mode switch function. Compared to mode automata, where computations are set in the states, the computations in Gaspard mode automata is placed in the mode switch component, i.e., outside of the Gaspard mode automata states. In Gaspard mode automata, a state graph component and its associated mode switch components are expected to be specified in the same repetition context in order to force the same executing cadence upon all the components.

There is also the incompatibility between specifications for parallelism in Gaspard and sequential trace semantics of automata. Hence, the parallel model is mapped onto a timed model through the time dimension defined in Gaspard. In addition, additional constructs (inter-repetition dependency and default link) are used to build a connecting link between the preceding and the following states of state graph components. More precisely, IRD specifications should be specified for the macro structure when it is placed in a repetition context. The reasons are as follows: the macro structure represents only one transition from source state on target state, whereas a Gaspard mode automata should have continuous transitions. Hence the macro should be repeated to enable multiple transitions. Thus the Gaspard mode automata can be built and executable.

### 4.4 A Gaspard mode automaton example

Figure 7 shows a Gaspard mode automata example, which can be transformed into synchronous mode automata eventually. MACRO is placed in a repetitive context, where each of its repetitions models one transition of mode automata. An IRD links the repetitions of MACRO and conveys the current state (sends the target state of one repetition as the source state to the next repetition) between these repetitions. The states and transitions of the automata are encapsulated in the SGT. The data computations inside the mode are set in the modes. SGT and its modes share the same repetition space, so they run at the same rate or clock. The detailed formal semantics of Gaspard mode automata can be found in (Gamatié et al., 2008a).
4.5 Abstract syntax of the control model

Models and their transformations sometimes are big in size and have divers technologies or domains involved, which leads to difficult understanding, maintenance and verification (Combemale et al., 2006; Mohagheghi and Dehlen, 2008; Chen et al., 2005). The extended control model and its transformation is first described using abstract syntax (Figure 8), as abstract syntax helps to concentrate on domain-specific concepts, models, and their transformations are therefore simplified and easy to be verified, particularly the coherence between the models involved in the transformation.

A Gaspard HTask (rule (r1) in Figure 8, where ( ) denotes a set; we call a Gaspard component as a task from the viewpoint of programming language here) consists of a set of Tasks, Deployment and Connections (r10), which connect the tasks. These tasks share common features (r2): a task_id, an interface (r3) and a Body (r5). Interface specifies input/output Ports (typed by i or o in rule (r3), and Port is defined in rule (r4)) from which each task receives and sends multidimensional arrays. Task has many types, including elementary, repetitive, hierarchical, mode switch, and state graph task. The type of a Task, is identified by the structure in the Body of the task. These task types are described as follows:

Figure 7: An example of Gaspard mode automata.

Figure 8: An abstract syntax extract of Gaspard concepts for functionality specification.
An elementary task \((r6)\) corresponds to an atomic computation block. Typically, it represents a function or an IP. Elementary task can be associated with an IP through Deployment.

A repetitive task \((r7)\) expresses data-parallelism in a task. The attribute \(r\) (in the rule \((r7)\)) denotes the repetition space for repetitions. In addition, patterns involved in each task repetition are defined via tilers \((r8)\). IRD can be also specified in a repetitive task \((r15)\), which describes the dependency between repetitions of the repetitive task. \(dep\_vector\) specifies which repetition relies on which repetitions and \(default\) gives a default value.

A hierarchical task \((r9)\) is actually a HTask. It is represented by a hierarchical acyclic task graph, in which each node consists of a task, and edges are labeled by arrays exchanged between task nodes.

A mode switch task \((r11)\) achieves mode switch function as presented in Section 4.1. It is composed of tasks as mode, which have the same interface. The interface of the mode switch task is the union of the interface of its internal tasks and the mode port, i.e., \(p_m\).

A state graph task \((r12)\) is associated with Gaspard state graphs that provide mode values for corresponding mode switch tasks. It is composed of a set of states \((r13)\) and transitions \((r14)\). A reset flag indicates reset of current state \(s_c\) for a state graph. \(m_o\) denotes the output mode.

Deployment \((r16)\) indicates how to find and integrate an implementation, considered as an IP, of a specific elementary task. Each elementary task is associated with an IP, and \(depl\_info\) describes necessary information for the integration of the IP into the system.

5 Transformation into the synchronous model

In Gaspard, the proposed control has been applied on functional specification, IP deployment, etc. The corresponding model transformations are also under development. Here the transformation from Gaspard control integrated into functional specifications to synchronous languages (Benveniste et al., 2003) is illustrated. The latter is used for the formal validation of Gaspard models. The transformation is described with the help of the abstract syntax of both the Gaspard model and the synchronous equational model.

A code segment of the synchronous dataflow language Lustre (Figure 9) is used for the introduction of some basic concepts here. A node is a basic functionality unit in Lustre. Each node gives the same results with the same inputs because of its determinism. Nodes have modular declarations that enable their reuse. Each node has an interface (input at line (l1) and output at (l2)), local definition (l3), and equations (line (l5) and (l6)). Variables are called signals in Lustre. Equations are signal assignments. In these equations, there are possibly node invocations (l5) that are declared outside this node. Obviously, in Lustre, modularity and hierarchy are inbuilt. The composition of these equations, denoted by “;”, means their parallel execution w.r.t. data-dependencies. The node has the same meaning independently of the equation order.

```plaintext
node node_name (A1:int^4) returns(A3:int^4);         (11)
  var A2:int^4;                                          (12)
  let                                                 (13)
    A2 = a_function(A1);                               (15)
    A3 = A1 + A2;                                      (16)
tel                                                  (17)
```

Figure 9: A simple example of Lustre code.
5.1 Synchronous equations abstract syntax

This abstract syntax is constructed based on common aspects of synchronous languages, which is intended to model three synchronous dataflow languages such as Lustre, Signal, and Lucid synchrone. The syntax is illustrated in Figure 10:

\[
\begin{align*}
Module & ::= \{Node\} & (s1) \\
Node & ::= nodename ; Interface ; EqSystem & (s2) \\
Interface & ::= Interface^1 ; Interface^o & (s3) \\
Interface^1 & ::= \{SignalDeclaration\} & (s4) \\
Interface^o & ::= \{SignalDeclaration\} & (s5) \\
SignalDeclaration & ::= signal ; DataType & (s6) \\
DataType & ::= type ; shape & (s7) \\
EqSystem & ::= \{Equation\} \mid CaseEquations \mid Automata \mid extnodelink & (s8) \\
Equation & ::= EqLeft ; EqRight & (s9) \\
EqLeft & ::= null \mid signal & (s10) \\
EqRight & ::= signal \mid SignalDelay \mid Invocation & (s11) \\
Invocation & ::= nodename \mid \{signal\} & (s12) \\
CaseEquations & ::= case \mid \{(modevalue, Equation)\} & (s13) \\
SignalDelay & ::= signal \mid delayinstant \mid \{defaultvalue\} & (s14) \\
Automata & ::= aut_id ; S ; Tr ; si ; reset & (s15) \\
S & ::= \{(state_id \mid (state_id ; Automata))\} & (s16) \\
Tr & ::= \{(state_id ; label ; state_id)\} & (s17)
\end{align*}
\]

Figure 10: An abstract syntax extract of basic synchronous concepts.

Only a brief description of synchronous languages’ syntax is given here. A more detailed description can be found in (Yu, 2008). A Node is defined as a basic functionality. All the nodes in an application are declared in a module (s1). A node (s2) is composed of Interface and EqSystem. The Interface has two families: Interface^1 and Interface^o (s3). An EqSystem is the body of a node and defines the function of the node. An EqSystem (s8) can have at least one Equation, CaseEquations, Automata, or an extnodelink, which indicate four implementation types of the EqSystem. An Equation is either an assignment of a signal, SignalDelay or an Invocation of another node (s9, s10 and s11). Signal and SignalDelay are variables used in the program. An Invocation indicates a function call to another node defined in the module. A CaseEquations contains a case statement where equations are activated according to some condition such as mode values (s13). A SignalDelay is similar to the pre operator (Benveniste et al., 2003), which takes the value of the signal at the previous delayinstant instant (s14). delayinstant is defined as a positive integer. A defaultvalue is a default value of a signal when no value is provided for the previous instant. Automata (s15) have an S (set of states, s16), Tr (set of transitions, s17), an initial state si, and a reset flag. The reset implies a restart of the automata, i.e., the initial state si is taken as entering state. Finally an extnodelink indicates an external implementation of the node, i.e., a node is implemented by other languages.

5.2 Transformation between the two models

The first step of transformation is structural, and the second step involves semantic aspects. The correspondence in the transformation between Gaspard and synchronous concepts is indicated by \( \xrightarrow{T} \). In order to distinguish the concepts of Gaspard and the synchronous model, the number of the rule, in which the concept appears in the syntax, is also given in parentheses following the concepts. The rule numbers of Gaspard concepts begin with r, and those of synchronous model start with s.

5.2.1 Structural transformation

The synchronous Module (s1) is the container of all nodes. A Gaspard HTask (r1) is first translated into a Node (s2): \( HTask \xrightarrow{T} Node \). A Task (r2) is also translated into a Node: \( Task \xrightarrow{T} Node \). An Interface (r3)
can be translated into Interfaces (s3): Interface $\xrightarrow{T} Interface$. A Port (r4) in an Interface is a connection point of a Gaspard task, it is translated into a signal: Port $\xrightarrow{T} signal$.

A Body (r5) represents the internal structure of a task. It can be considered as an Eqsystem (s8): Body $\xrightarrow{T} Eqsystem$. Five kinds of structures are involved in a body. A Structure$^e$ (r6) represents the structure of an elementary task, and the deployment will be used to integrate an IP for this task. A Structure$^r$ (r7) is translated into a set of equations (s9) (Yu et al., 2008b). A Structure$^h$ (r8) is a structure that has a compound task, i.e., HTask. A Structure$^{mt}$ (r11) corresponds to CaseEquations (s13): Structure$^{mt} \xrightarrow{T} CaseEquations$. A Structure$^{sg}$ (r12) can be translated into either CaseEquations (s13) or Automata (s15): Structure$^{sg} \xrightarrow{T} CaseEquations$ or Structure$^{sg} \xrightarrow{T} Automata$. Finally, Deployment (r10) is translated into an extnodelink (s8), which indicates the integration of an external implementation: Deployment $\xrightarrow{T} extnodelink$.

5.2.2 Transformation of behavioral aspects
In addition to structural aspects, transformation also involves behavioral aspects of Gaspard specifications. These aspects include the transformation of parallelism and control. The former, which includes elementary, repetitive, and hierarchical task transformation, has been presented in (Yu et al., 2008b). The latter, including mode switch and state graph tasks will be presented here.

IRD (r15), which enforces serialized execution and passes values between repetitions, is translated by SignalDelay (s14): IRD $\xrightarrow{T} SignalDelay$. A SignalDelay is used to convey the previous value of a signal. Depvector is translated by delayinstant: depvector $\xrightarrow{T} delayinstant$. These two concepts are used to indicate which previous value to take in Gaspard and synchronous program.

A Structure$^{mt}$ (r11) corresponds to CaseEquations (s13). Each pair of $(m_k, T_k)$, i.e., a mode, is translated into (modevalue, Equation) in a case statement. The modevalue is the condition and Equation is invoked when the corresponding condition is evaluated as true. The Equation represents the computation to carry out in a mode.

A Structure$^{sg}$ (r12) is translated into either CaseEquations (s13) or Automata (s15) according to target languages. The former translate state graphs into pure equations with case statements, while the latter is a direct transformation which keeps explicit structure of states and transitions.

6 Implementation within MDE
This section concentrates on the implementation of previously mentioned control modeling and transformation in the MDE framework, based on conceptual descriptions of models and their transformation.

6.1 Gaspard control metamodel
UML (Object Management Group, 2007b) is adopted in Gaspard to specify state graphs, which is actually a subset of UML State machines. It also helps to ensure the compatibility between Gaspard control with the MARTE profile. An extract of the Gaspard control metamodel is illustrated in Figure 11. This metamodel is proposed as a subset of the UML State machines metamodel in structure so that the compatibility between them simplifies the following transformation.

6.2 Model transformation
Based on the previously presented transformation on the foundation of abstract syntax, model transformation rules are built on OMG Query / Views / Transformations (Object Management Group, 2005) tools such as MOMOTE tool (MOdel to MOdel Transformation Engine) (INRIA DaRT team, 2009). Most part of the transformation from Gaspard to synchronous languages has been developed as Eclipse plugins (Yu et al., 2008b). An extension is expected to cover transformations to other platforms. Here, only an example is given in Figure 12 to illustrate the resulting automata (Lustre mode automata) obtained from Figure 2(b). Equations of data-parallel processing are not included in this example as Lustre mode automata is only used for model checking purpose.
Figure 11: An extract of Gaspard StateGraph, which is proposed according to the metamodel of UML state machines.

AUTOMATON auto
STATES
s1 init [ state = s1; mode = M1; ]
s2 [ state = s2; mode = M2; ]
s3 [ state = s2; mode = M3; ]
TRANS
FROM s1 TO s2 WITH [ e1 = true and e2 = true]
FROM s2 TO s1 WITH [ e1 = true ]
FROM s1 TO s3 WITH [ e3 = true ]
FROM s3 TO s1 WITH [ e1 = true ]
FROM s2 TO s3 WITH [ e1 = true and e3 = true]
FROM s3 TO s2 WITH [ e2 = true ]
PROCESS auto [in(e1, e2, e3), out(state, mode)]

Figure 12: An extract of a Lustre mode automaton obtained by transformation.

7 Some applications
Gaspard is well-fitted for applications with repetitive data-parallel computations, which include image processing, multimedia video codecs, software-based radio and radar signal processing, etc. A modern
cellular phone is then taken as a typical example in this paper, which have complex multimedia functionalities: camera, games, MP3 music, video. Let us focus on the video part. A global model view of this multimedia module is illustrated in Figure 13. The played video clips are obtained from different VideoSources, either on-line library or local storage. There are different display ImageStyle such as Black & White; Negative; Sepia; or Normal, meaning no effect. In addition, the Resolution of the video can be set to High, Medium and Low. Finally, the color can be in either Color or Monochrome for ColorEffect options. ImageStyle, Resolution and ColorEffect include both control and data processing parts.

Besides user commands, the video display modes are controlled by the system by the Controller, which validates mode change requests from other components according to current mode configuration and the available computational resources. This specific component is either coded manually, or generated automatically by discrete controller synthesis according to QoS requirements, including the status of energy level (EnergyStatus) and the communication quality (CommQuality). The former indicates the energy level according to events received from an energy monitor component, and the latter provides the communication quality level according to the energy level and the on-line transmission bandwidth of received data.

![Figure 13: A global model view of the multimedia functionality module.](image)

The above module has different configuration modes following which its components achieve algorithms for a suitable image display. Depending on the resource status, e.g., energy level, communication bandwidth, the display quality of an on-line video varies. Hence, each mode is associated with non-functional properties, which must be satisfied in order to display images at a good quality level. The modes defined in the components are characterized by quantitative attributes representing the following non-functional properties: energy consumption (E), communication quality required (CQ), computing resource consumption (CR) and memory consumption (M). These non-functional requirements are considered as instantaneous consumption of quantitative resources that may vary from one system reaction to another. The values associated locally with the modes are combined additionally when components are composed in parallel, so as to obtain global costs for the whole system from the local costs of its components.

Possible behaviors involving the above characteristics are, e.g., that the consumption of a resource must respect the bounds defined by its capacity. Therefore, if a new functionality is executed, then the other tasks that are already running should switch to lower consumption modes, possibly reducing their quality as well. Or, if the level of the battery goes down, then the control should switch task modes so...
Adaptivity in High-Performance Embedded Systems

that the lower energy capacity is respected. Such control strategies are defined by properties expressed in terms of the states and inputs of the system. The Sigali tool (Marchand et al., 2000) allows one to express Boolean properties on states and inputs \( (P : \mathbb{B}^n \rightarrow \mathbb{B}) \), and to build cost functions, associating numerical values (here, assumed to be integer, without loss of generality) with Boolean functions of states and inputs \( (f : \mathbb{B}^n \rightarrow \mathbb{N}) \). We essentially consider invariance, by specifying a subset of system states, defined by a Boolean property \( P \): \( P \) is invariant for the system if for all states in this subset, transitions from these states lead to states in the same subset. This invariance property of the system is noted \( \forall \Box P \): the Boolean property \( P \) is true at every instant of every trace of the system.

7.1 Model checking

Design validation through model checking of functional properties was first studied (Yu et al., 2008a). One of the examples is an exclusion relation between two modes from two different components. For instance, in order to avoid waste of resources, it can be useful to specify that the modes B&W (ImageStyle component) and Color (ColorEffect component) are never active at the same instant. This invariance property is denoted: \( \forall \Box (\text{B&W} \land \text{Color}) \). Nonfunctional properties, for instance, resources provided by the environment, platform or hardware, can also be checked. By associating each state of the automata with cost functions, a global cost function has been computed when composing these automata in parallel. This global function defines, for each resource, the sum of its cost functions computed from all possible combination of active states in the global automaton. A resource bound is then specified in the system. Finally, the reachability of certain states is checked under this configuration of resource costs. This example shows a fast verification in consideration of functional and/or nonfunctional properties in order to evaluate complex systems at a low cost in a fast manner, compared with other approaches that need to specify the system in a precise way (Yu et al., 2008a).

7.2 Discrete controller synthesis

Another example involves discrete controller synthesis (Gamatié et al., 2009): for a cost function where the global cost is defined by the sum of the local costs of the components, e.g. for memory footprint, there is a bound defined by the size of the memory. Thus, if we note by \( f_M = f_{IS} + f_{VS} + f_R + f_{CE} \) the cost function associating with each global state of the system, the memory usage in this state, we can enforce that this usage will always be bounded by the memory available (here, 90 units), by the invariance synthesis objective: \( \forall \Box (f_M \leq 90) \). The bound itself can vary in time: it is actually the case for the available energy. In this case, we add to our model an automaton which represents the environment, namely here the energy resource available. We associate a cost function \( f_{EA} \) with this automaton, associating with its states the energy quantity instantaneously available. Then, we can bound the energy consumption \( f_E \) by the available energy: \( \forall \Box (f_E \leq f_{EA}) \).

With these synthesis operations, different policies or strategies can be obtained automatically by changing the objectives, hence providing for separation of concerns and making the models easy to reuse.

The controller computed by Sigali is extracted, and co-simulated with the system with the SIGALSIMU tool. Figure 14 shows a particular simulation step, where the controller enforces the values of two controllable inputs so as to keep the properties satisfied. At this step, the system is in high energy, high resolution, and color state. We then simulate the discharge of the battery by the occurrence of the uncontrollable input event_energy_down. On the controllable inputs panel, the clearer inputs shown with ellipsis are those whose values have been forced by the controller. It is here the case of the input ctr_resolution, meaning that the controller has triggered the transition from high to medium resolution state.

7.3 Hardware synthesis

The control model has also been applied to move from high level MARTE specifications to reconfigurable architectures such as FPGAs, and specifically those supporting partial dynamic reconfiguration. The continuation of this work is in progress in the project FAMOUS, in which the OMG MARTE profile will
be extended into RecoMARTE, for the design of reconfigurable architectures, implemented on FPGAs. In addition, slight extensions have been made to the existing control/data flow concepts and the deployment level in our framework to integrate the partial dynamic reconfiguration aspects. An initial version of generating a complete IP core from model transformations has also been developed, along with a solution to avoid de-synchronization related to task parallelism in the modeled applications (Quadri et al., 2010). Currently the Xilinx based partial dynamic reconfiguration design flow is adopted due to its availability and extensible nature.

8 Conclusions

This paper presents a reactive control model and its transformation in a model-driven engineering-based SoC co-design framework for high-performance systems, which is compliant with the MARTE standard. The control model is based on mode automata in order to enable the specification of adaptivity for high-performance systems. Our contribution is the integration of composition and formal semantics into this model to enrich its expressivity and verifiability. Model transformation towards synchronous languages has also been studied in order to benefit from validation tools associated with these languages. Furthermore, the extended control is also integrated into Gaspard IP deployment for reconfigurable FPGAs, besides functional specifications. The model is illustrated with its abstract syntax, based on which a conceptual transformation is given. An implementation of the metamodel and model transformation in Eclipse has been carried out, and it is partially accomplished.

One perspective of this work is the application of the control model towards other target technologies such as SystemC. Reconfigurability is also a very interesting research topic in these platforms. Our control proposition is one of possible solutions, which take high-performance computing into account. Another perspective is related to the implementation of the control model, its transformation, and formal verification in the unique framework of model-driven engineering. The backstage technologies should
be made transparent to users and the only interface to users would be an integrated development and simulation environment, such as Eclipse.

References


