Exploring System Architectures in AADL via Polychrony and SYNDEEx

Huafeng YU 1, Yue MA 1, Thierry GAUTIER (✉), Loïc BESNARD 2
Jean-Pierre TALPIN 1, Paul LE GUERNIC 1, Yves SOREL 3

1 INRIA Rennes - Bretagne Atlantique, 263, av. du Général Leclerc, 35042 Rennes, France
2 IRISA/CNRS, 263, av. du Général Leclerc, 35042 Rennes, France
3 INRIA Paris - Rocquencourt, Domaine de Voluceau, BP 105, 78153 Le Chesnay Cedex, France

© Higher Education Press and Springer-Verlag Berlin Heidelberg 2012

Abstract AADL (Architecture Analysis & Design Language) has been increasingly adopted in the design of embedded systems, and corresponding scheduling and formal verification have been well studied. However, little work takes code distribution and architecture exploration into account, particularly considering clock constraints, for distributed multi-processor systems. In this paper, we present an overview of our approach to handle these concerns, together with the associated toolchain, AADL-Polychrony-SYNDEEx. First, in order to avoid semantic ambiguities of AADL, the polychronous/multiclock semantics of AADL, based on a polychronous model of computation, is considered. Clock synthesis is then carried out in Polychrony, which bridges the gap between the polychronous semantics and the synchronous semantics of SYNDEEx. The same timing semantics is always preserved in order to ensure the correctness of the transformations between different formalisms. Code distribution and corresponding scheduling is carried out on the obtained SYNDEEx model in the last step, which enables the exploration of architectures originally specified in AADL. Our contribution provides a fast yet efficient architecture exploration approach for the design of distributed real-time and embedded systems. An avionic case study is used here to illustrate our approach.

Keywords Polychrony, SIGNAL, AADL, SYNDEEx, architecture exploration, modeling, timing analysis, scheduling, distribution

1 Introduction

The Architecture Analysis & Design Language (AADL) [1] is gradually adopted for high-level system co-modeling in embedded systems due to issues of system complexity, time to market, verification and validation, etc. It permits the fast yet expressive modeling of a system [2]. Early-phase analysis and validation can be rapidly performed [3–9]. AADL provides a fast design entry, however, there are still critical challenges, such as unambiguous semantics, architecture exploration, code distribution, timing analysis, or co-simulation. To address these issues, expressive formal models and complete toolchains are required.

Synchronous languages are dedicated to the trusted design of synchronous reactive embedded systems [10]. Thanks to their mathematical definition and the software tools that are founded on these definitions, synchronous languages are good candidates to address the above issues of AADL. Among these languages, the SIGNAL language stands out as it enables to describe systems with multicycle relations [11], and to support refinement [12]. Polychrony is a toolset based on the SIGNAL language, dedicated to multicycle synchronous program transformation and verification.

Various tools exist to help the implementation of applications onto distributed platforms composed of processors, specific integrated circuits, and communication media all together connected. Among them, SYNDEEx is a system level Computer Aided Design (CAD) software, based
on the Algorithm-Architecture Adequation (AAA) methodology [13]. It aids the designer to search, manually and/or automatically, for an optimized implementation of an embedded control application onto a distributed platform architecture, while satisfying real-time constraints. SynDEx provides interfaces to application description languages such as SIGNAL.

The AADL language is designed to be used with analysis tools. Our aim is to make Polychrony and SynDEx such tools in an integrated methodology.

In this paper, an overview of our approach is presented, including the formal timing modeling, timing analysis, clock synthesis, architecture exploration as well as the associated toolchain, called AADL-Polychrony-SynDEx. First, two polychronous/synchronous formalisms [10], Polychrony/SIGNAL [14, 15] and SynDEx [16], are adopted to provide support for formal timing modeling based on a polychronous Model of Computation (MoC) [11]. We revisit the timing semantics of AADL as multi-clocked, so that AADL components are modeled in this polychronous MoC. In this way, users are not suffered to find and/or build the fastest clock of the system, which distinguishes from [17], [4], [16], [18]. According to this principle, AADL models are transformed into SIGNAL models. To bridge the gap between the polychronous semantics of SIGNAL and synchronous semantics of SynDEx, clock synthesis in Polychrony [14], the design environment dedicated to SIGNAL, is applied. The translation from SIGNAL to SynDEx is integrated in Polychrony. Finally, SynDEx models are used to perform distribution, scheduling, and architecture exploration.

Main advantages of our approach are following ones: 1) a formal model is adopted to connect the three formalisms, and it helps to preserve the semantic coherence and correct code generation in the transformations; 2) the formal model and methods used in the transformations are transparent to AADL practitioners, and it is fast and efficient to have illustrative results for architecture exploration; 3) it provides the possibility for one of the three formalisms to take advantage of the functionalities provided by the other ones.

An AADL-Polychrony-SynDEx toolchain has been developed, which includes automatic model transformations between the three formalisms, considering both semantic and syntactic aspects. A tutorial avionic case study is used in this paper to show the effectiveness of our contribution. This compact yet typical and general case study has been developed in the framework of the CESAR project [19].

Outline. An overview of the tool architecture is given in Section 2. The three formalisms, Polychrony/SIGNAL, AADL, and SynDEx, are introduced in Section 3. Section 4 and Section 5 present our original contribution: the timing modeling, translations between the different formalisms, clock synthesis, architecture exploration. The whole work is illustrated by a case study in Section 6. Some related works are summarized in Section 7, and a conclusion is drawn in Section 8.

2 The scenario

Our proposed approach for architecture exploration of AADL high-level models is illustrated in Fig. 1. Three stages are presented in the design process, which include: high-level modeling in AADL, considering both architectural and behavioral aspects; model transformations, timing analysis and clock synthesis using Polychrony; and architecture exploration with the aid of SynDEx. The Polychrony design environment [14], associated with the SIGNAL language, provides a formal framework for back-end semantic-preserving transformations, scheduling, code generation, formal analysis and verification, architecture exploitation, and distribution [20]. The internal representation used for these purposes is a Data Control Graph (DCG), composed of a clock hierarchy and a conditioned precedence graph (see [20] for more details). The Polychrony toolset is an open-source modeling framework, DO-330 qualified (as verification tool, criteria 3); it is being integrated in the PolarSys toolset (Polarsys is an Industry Working Group of The Eclipse Foundation [21]).

The inherent formal model, associated transformations and tools used in this approach are transparent to system designers, i.e., AADL practitioners. The results of architecture exploration is graphically illustrated. This approach is thus fast and efficient to perform architecture exploration from a high-level modeling perspective.

The polychronous model is adopted to bridge between the three formalisms, i.e., AADL, SIGNAL and SynDEx. This
The AADL-Polychrony-SynDEx toolchain

3 The casting

3.1 The polychronous model and the SIGNAL kernel language

The semantics of SIGNAL is defined over a polychronous Model of Computation. A SIGNAL process defines a set of (partially) synchronized signals as the composition of equations. A signal \( x \) is a finite \( ((\exists n \in \mathbb{N})(x = (x_i)_{i \in \mathbb{N}})) \) or infinite \( (x = (x_i)_{i \in \mathbb{N}}) \) sequence of typed values in the data domain \( D_x \); the indices in the sequence represent logical discrete time instants. At each instant \( t \), a signal is either present and holds a value \( v \) in \( D_x \), absent and virtually holds an extra value denoted \#; or completed and never holds any actual or virtual value for all instants \( s \) such that \( t \leq s \). The set of instants at which a signal \( x \) is present is represented by its clock \( \hat{x} \). Two signals are synchronous iff they have the same clock. Clock constraints result from implicit constraints over signals and explicit constraints over clocks.

The semantics of the full language is deduced from the semantics of a kernel language, and from the SIGNAL definition of the extended features. A SIGNAL kernel language process is either a kernel equation \( x := f(x_1, \ldots, x_n) \) where \( f \) is a kernel function, or the composition \( PQ \) of two kernel processes \( P \) and \( Q \), or the binding \( P/x \) of the signal variable \( x \) to the kernel process \( P \). In this Section, we give a sketch of its functional part using data-flow models.

Semantic domains

For a set of values (a type) \( D \) we define its extended set \( D \# = D \cup \{\#\}) \, where \( \# \notin D \) is a special symbol used to denote the absence of a signal value. The semantics of SIGNAL is defined as least fixed point in domains: for a data domain \( D \), we consider the poset \( (D \# \cup \{\bullet, \perp\}, \leq) \) where:

- \((D \# \cup \{\bullet, \perp\}, \leq)\) is flat (i.e., \( x \leq y \Rightarrow x = y \));
- \(\bullet\) that denotes the presence of a signal is the infimum of \((D \cup \{\bullet, \perp\}, \leq)\); \(\bullet\) and \# are not comparable;
- \(\perp\) that denotes the absence of information is the infimum of \((D \# \cup \{\bullet, \perp\}, \leq)\).
We denote by $D^\omega = D^* \cup D^\omega$ the set of finite ($D^*$) and infinite ($D^\omega$) sequences of “values” in $D$. The empty sequence is denoted by $\epsilon$. All $n$-ary signal functions $f : D_1^\omega \times \ldots \times D_n^\omega \to D_{n+1}^\omega$ are defined using the following conventions: $s_1, \ldots, s_n$ are (possibly empty) signals in $D_i^\omega$, $v_1, \ldots, v_n$ are values in $D_i$ (cannot be $\#$), $x_1, \ldots, x_n$ are values in $D_{i#}$. As usual, $|s|$ is the length of $s$, $s_1.s_2$ is the concatenation of $s_1$ and $s_2$ (equal to $s_1$ if $s_1 \in D^\omega$).

Given a non empty finite set of signal variables $A$, a function $b : A \to D^\omega$ that associates a sequence $b(x)$ with each variable of $A$ is named a behavior on $A$. The length $|b|$ of a behavior $b$ on $A$ is the length of the smallest sequence $b(a)$.

An event on $A$ is a behavior $b : A \to D^\#$. For a behavior $b$ on a set of signal variables $A$, an integer $i \leq |b|$, $b(i)$ denotes the event $e$ on $A$ such that $e(a) = (b(a))(i)$ for all $a \in A$. An event $e$ on $A$ is said to be empty iff $e(a) = \#$ for all $a \in A$.

The concatenation of signals is coordinatewise extended to tuples of signals. Two behaviors $b_1$, $b_2$ are stretch-equivalent iff they only differ on non-final empty events (see [11] for more details).

The Signal kernel functions

A Signal kernel function is a $n$-ary (with $n > 0$) function $f$ that is total, strict and continuous over domains [24] (w.r.t. prefix order) and that satisfies the following general rules:

- stretching: $f(\#.s_1, \ldots, \#.s_n) = \#.f(s_1, \ldots, s_n)$
- termination: $((\exists i \in 1..n)(s_i = \epsilon)) \Rightarrow f(s_1, \ldots, s_n) = \epsilon$

A $n$-ary function $f$ is synchronous iff it satisfies:

\[
\forall v_1, \ldots, v_n \in D_1, \ldots, D_n, x_1, \ldots, x_n \in D_{1#}, \ldots, D_{n#}, s_1, \ldots, s_n \in D_1^\omega, \ldots, D_n^\omega,
\]

\[
\{ - ((\exists i, j \in 1..n)(\exists v \in D_i)(x_i = \# \land x_j = v)) \Rightarrow \\
(f(x_1, s_1, \ldots, x_n, s_n) = \epsilon)
\}
\]

\[
- (\exists v \in D, s \in D^\omega)f(v_1.s_1, \ldots, v_n.s_n) = v.s
\]

Definition of stepwise extension kernel functions

Given $n > 0$ and a $n$-ary total function $f : D_1 \times \ldots \times D_n \to D_{n+1}$, the stepwise extension of $f$ denoted $F$ is the kernel synchronous function that satisfies:

\[
F(v_1.s_1, \ldots, v_n.s_n) = f(v_1, \ldots, v_n).F(s_1, \ldots, s_n)
\]

Usual infix notation is used for standard operators such as $=$, and, $+$, etc.

Definition of previous value kernel function

\[
delay : D_1 \times D_1^\omega \to D_1^\omega
\]

is the kernel synchronous (state-) function that satisfies:

\[
delay(v_1, v_2.s) = v_1.delay(v_2, s)
\]

The infix notation of $delay(v_1, s)$ is $s \$ init \ v_1$.

Definition of prioritized merge kernel function

\[
default : D_1^\omega \times D_1^\omega \to D_1^\omega
\]

is the kernel function that satisfies:

\[
default(v.s_1, x.s_2) = v.default(s_1, s_2)
\]

\[
default(\#, s_1, x.s_2) = x.default(s_1, s_2)
\]

The infix notation of $default(s_1, s_2)$ is $s_1 \ default s_2$.

Definition of Boolean sampling kernel function

Let $\mathbb{B} = \{ff, tt\}$ denote the set of Boolean values. When $D_1^\omega \times \mathbb{B} \to D_1^\omega$ is the kernel function that satisfies:

\[
- \text{ for } b \in \{#, ff\}, \ when(x.s_1, b.s_2) = #.when(s_1, s_2)
\]

\[
- \text{ when(x.s_1, tt.s_2) = x.when(s_1, s_2)
\]

The infix notation of $when(s_1, s_2)$ is $s_1 when s_2$. For a Boolean signal $s$, $when s$ is the unary notation of $when(s, s)$ that returns $tt$ iff $s = tt$, # otherwise.

Deterministic process

An equation is a pair $(x, E)$ denoted $x : = E$. An equation $x : = \# \# E$ associates with the variable $x$ the sequence resulting from the evaluation of the signal function $f$ denoted by $E$ (defined as a composition of kernel functions). If $A = \{x_1, \ldots, x_n\}$ (x $\notin A$) is the set of the free variables in $E$, the equation $x : = E$ denotes a process on $A$, i.e., a set of behaviors on $A \cup \{x\}$; a process is closed by stretch-equivalence (thanks to the stretching rule).

The parallel composition of processes $P_1|P_2$ defined on kernel equations is a process $P$ equal to a network of strict continuous functions interconnected on the basis of signal names, as usual in equations. When $P$ satisfies the Kahn conditions (no cycle, no double definition, ...), $P$ is a strict continuous function or Kahn Process Network (KPN) [25], defined as least upper bound satisfying the equations. This function satisfies the “stretching” and the “termination” rules. It may be or not synchronous. It is stretch-closed. The binding $P/x$ is a projection that makes local the variable $x$. In the “intersection” semantics of Signal [11], a process is the set of infinite behaviors accepted by the above “KPN” semantics.

Non deterministic process

A process with feedback or local variables may be non deterministic. The semantics of a non deterministic process can
be defined using Plotkin’s power-domain construction [26]. The input free equation \( x := x \$ \text{init} \ 0 \) is a typical example of a non deterministic process: \( x \) holds a sequence of constant value \( 0 \) separated by any number of \# (stretch-equivalence).

3.2 The full \texttt{Signal} language

Process models

A process model is a structuration feature illustrated by the counter modulo process (Listing 1): at each step, \( nb \) is the number modulo \( n \) (static parameter of the current occurrence of \texttt{val}).

\begin{verbatim}
process counterModulo = \{ type tau; integer n \}
   (? tau \ val \ integer nb) |
   (| nb := (0 when reset) \ default \ plus_un |
    reset := \ pre_nb = n-1 |
    plus_un := pre_nb + (1 when \^val) |
    pre_nb := nb \$ init \ (n-1) |
   )
where boolean reset; integer pre_nb, plus_un;
end
Listing 1  Definition of the \texttt{Signal} process counterModulo
\end{verbatim}

A process model is made of a name, a list of static parameters (formal type \texttt{tau} and constant value \texttt{n}), a list of signal inputs (the signal \texttt{val} of which occurrences are counted), a list of signal outputs (the signal \texttt{nb} that holds the number of each \texttt{val} occurrence modulo \texttt{n}), a body that contains a composition of equations and a list of local variables. An occurrence (such as \texttt{counterModulo(event, 60)(sec_event, sec_count)} in equation form, or \texttt{sec_count:=counterModulo(event, 60)(sec_event)} in function call form) of an instance of \texttt{counterModulo} is replaced by the body of \texttt{counterModulo} with required substitutions.

From non deterministic to deterministic process

The operator \texttt{var} (see [20]) is a typical example of an operator that is not a function. It is a \textit{derived} operator, the definition of which is that of \texttt{preMemoryCell} (Listing 2).

\begin{verbatim}
process preMemoryCell = \{ type tau; tau val \}
   (? tau input \ tau output) |
   (| mem := input \ default \ mem \$1 |
    output := mem when \^output |
   )
where tau mem init val;
end
Listing 2  Definition of the \texttt{Signal} process \texttt{preMemoryCell}
\end{verbatim}

The process model \texttt{preMemoryCell} is a mix of data-driven and demand-driven operator: \texttt{output := mem when \^output ("\^" is the operator that returns \texttt{tt} when its argument is present) outputs the current value when it is (implicitly) required. Thus \texttt{preMemoryCell} is not a function but a \textit{relation}. It is a \textit{polychronous} process that has two independent clocks, \texttt{input} and \texttt{output}. The clock \texttt{mem} is an upper bound of the clocks \texttt{input} and \texttt{output}.

Including the relation \texttt{preMemoryCell} (equivalent to the operator \texttt{var}) in a context for which the clock of the output is added as an input signal allows to build a deterministic process \texttt{memoryCell} (Listing 3).

\begin{verbatim}
process memoryCell = \{ type tau; tau val \}
   (? tau input; event clk_output |
    tau output) |
   (| mem := var input |
    output := mem when clk_output |
    mem =^ clk_output ^+ input |
   )
where tau mem init val;
end
Listing 3  Definition of the \texttt{Signal} process \texttt{memoryCell}
\end{verbatim}

The expression \texttt{mem =^ clk_output ^+ input} means that the clock of \texttt{mem} is the least upper bound of the clock \texttt{clk_output} and that of the \texttt{input} signal. The operators \texttt{=} and \texttt{=} denote respectively a relation and a function over clocks derived form \texttt{Signal} kernel features.

Another example of a non deterministic process is the equation \( x := 0 \) that defines the constant signal \( x \) as being equal to \# or equal to \( 0 \). This equation is a shortcut for \( x := x \$ \text{init} \ 0 \).

A last example of non deterministic process is the equation \( x := E \) that defines \( x \) to be equal to \( E \) when \( E \) is present and undefined when \( E \) is \#. This equation is a shortcut for \( x := E \text{ default} \ x \). A signal \( x \) can be consistently defined by several equations \( x := E1, \ldots, x := En \) in a process, provided that for every pair of equations \( x := Ei, x := Ej \), when \( Ei \) and \( Ej \) are both present, they hold the same value. If \( E1, \ldots, En \) do not recursively refer to \( x \) and if they denote functions, then \( \{x := E1 \ldots |x := En\} \) is a deterministic process. These partial definitions are very useful in automata where the function that computes the value of a signal often depends on current state. The states being exclusive, the consistency property is satisfied.
Delay sensitivity

Deterministic processes are defined over data types extended with # symbol. The presence (or absence) of the occurrence of a signal can be tested in the context of a global program with a global clock by interface functions such as “present” in Esterel [10]. In the context of a distributed system, when an occurrence of a signal is “arrived”, the signal has not necessarily to be considered as present in the logical instant. Conversely, when the action associated with a logical time is started, some of the required occurrences may not be arrival. Not only because of the delay sensitivity, but also because of computation scheduling: for instance, when a process \( P_1 \) computes a signal \( x \) that is used in another process \( P_2 \) to compute a signal \( y \) that is used in \( P_1 \) at the same instant.

Processes are usually delay sensitive. In particular, a reactive process with more than one input is delay sensitive. Delay insensitive processes are flow functions, i.e., functions on pure signals that do not contain explicit occurrences of #. Clearly, stepwise extension functions and delay functions are delay insensitive. To get a delay insensitive sampling, one can associate with an equation \( \text{output} := \text{input when sample} \) a Boolean clock parameterization (\( C_{\text{input}}, C_{\text{sample}} \)) of the sampling equation as in flowSampling (Listing 4).

```
process flowSampling = {type tau}
  (? boolean C_input, C_sample;
    tau input; boolean sample;
    ! tau output)
  |
  (| output := input when sample
    | (| C_input := C_sample
      | sample := when C_sample));

Listing 4 Definition of delay insensitive sampling
```

The process model flowSampling denotes an endochronous process: it has a master clock (a tick) \( ^{\tau} C_{\text{input}} \) and \( ^{\tau} C_{\text{sample}} \), and the other clocks are (recursively) functions of this master clock and the input signal values. One can similarly define an endochronous extension of a default equation. A composition \( P \) of endochronous processes that satisfies the Kahn conditions is a strict continuous (flow) function and then \( P \) is delay insensitive. Note that a delay insensitive process may not be endochronous. Usually a KPN is not endochronous (it may require unbounded fifos to be implemented in a single thread).

Thanks to the properties of the Signal composition (commutativity, associativity, idempotence), the possibility of moving inner bindings out of a composition, and the properties of a powerful clock algebra, the so-called clock calculus can arrange a process \( P \) as a composition of endochronous components, adding required Boolean inputs as clock parameters.

Dependences

The equations of a Signal process induce guarded data-flow dependences. For instance in the equation \( z := x \text{ default } y, x \text{ precedes } z \) \( (x \rightarrow z) \) and \( y \) precedes \( z \) when \( x \) is not present \( ((y \rightarrow z) \text{ when } (\neg y \text{ or } \neg x)) \). Input/output guarded dependences of a process are computed in a path algebra. They are written by Polychrony as abstract interface properties of black box behaviors. Dependences can be made explicit by the user to enforce scheduling.

Specification

A process that has no output and that does not call external functions is a property process. For instance \( a \tableau{\neg\neg b} \) denotes the equality of clocks of \( a \) and \( b \), \( (a \rightarrow b) \text{ when } c \) can be written to specify that an input \( a \) precedes (is read before) an other input \( b \) when the Boolean input \( c \) is \( \text{true} \) (and then \( c \) implicitly precedes \( a \) and \( b \)).

Such a property process can be used as a constraint in the body of a process or as assumed or asserted properties in its interface.

Signal open features

An external process (a function, a subprogram, a method, ...) can be abstracted (referred to) as a specification of its interface.

A pragma is an annotation that is associated with a process for specific purpose. Pragmas are sorted by name related to an action (for instance code generation), a tool (for instance a model checker), etc. The set of pragma classes is open.

3.3 AADL

AADL [1] is a Society of Automotive Engineers (SAE) standard dedicated to design and analyze the architecture of performance-critical real-time systems. AADL describes the structure of an embedded application as an assembly of software components allocated on execution platform. The component interactions and the dynamic behavior can be described. The AADL standard specifies software and hardware components from an external perspective: AADL uses black box interfaces to describe non functional properties such as timing requirements or abstractions, fault and error behaviors, time and space partitioning, and safety and certification.
properties. Thus each component has a type, which represents the functional interface of the component and its externally observable attributes. Each type may be associated with zero, one or more implementation(s) that describe the contents of the component, as well as the connections between them.

In AADL, three main distinct component categories are provided:

- a software component is a process, a thread, a subprogram, or a data component;
- execution components model the hardware part of a system, including (virtual) processor, memory, device, and (virtual) bus components;
- composite component, named system, contains execution platform, application software or other composite components.

In this presentation we ignore other components such as thread groups that do not introduce new aspects in our translation scheme.

AADL provides mechanisms of exchange and control of data: message passing, event passing, synchronized access to shared components, thread scheduling protocols (periodic, aperiodic, sporadic, background), timing requirements, remote procedure calls. AADL components interact exclusively through defined interfaces. A component interface consists of directional flow through: data ports for unqueued data, event data ports for queued message data, event ports for asynchronous events (triggers for the dispatch of an aperiodic thread, initiators of mode switches, alarm communications), synchronous subprogram calls, explicit access to data components. Interactions among components are specified explicitly. For example, data communication (immediate or delayed) among components is specified through connection declarations.

Properties are specified to provide more information about model elements. For example, application components have properties that specify timing requirements such as period, worst-case execution time, deadlines, space requirements, arrival rates, and characteristics of data and event streams. In addition, properties identify the following elements: source code and data that implement the application component being modeled, constraints for binding threads to processors, source code, and data onto memory.

The behavior annex provides an extension to AADL core standard so that behavior specifications can be attached to AADL components. The behavior is described with a state transition system equipped with guards and actions.

Adapting the terminology of [1], we name complete application model an instance of a root system in which all components are recursively instantiated. A complete application model is *bound* if each thread is bound to a processor, each source text, data component and port is bound to a memory, each connection is bound to a bus. In this paper we are interested in complete application models for software architecture and bound complete application models. So a weak attention will be given to types and other abstraction or partial features.

In the following, an industrial case study of simplified doors and slides control system (SDSCS) in an avionic generic pilot application, proposed by Airbus in the frame of CESAR project, is used to illustrate the basic components of an AADL model. In this case study (Fig. 3), a typical safety critical system takes charge of the management of passenger doors. It includes different components modeling hardware and software, and allowing them to communicate and control the doors.

In the system *door_manager*, two subsystems, *door1* and *door2*, are managed by two processes, *doors_process1* and *doors_process2* (Fig. 3). The processor *CPIOM1* (resp. *CPIOM2*) is responsible for scheduling and executing threads in process *doors_process1* (resp. *doors_process2*). The devices, e.g., *LGS, DPS*, etc., interface with external environment of the system. All the communication between the devices and processors is through the bus *AFDX1*. 

![An overview of the SDCS case study in AADL](image-url)
3.4 SynDEx

SynDEx [27] is a system level CAD software based on the Algorithm-Architecture Adequation (AAA) methodology [13] which allows the designer to optimize the implementation of embedded control applications onto distributed platforms, while satisfying real-time constraints. This is a free-ware distributed free of charge [28]. Using graph models, the AAA methodology allows the designer to perform a functional specification that we call “algorithm”, a non-functional specification describing the real-time characteristics (Worst Case Execution Time (WCET), Worst Case Communication Time (WCCT), deadline, period, end-to-end latency, etc.) and the distributed platform (processors, specific integrated circuits and communication media) that we call “architecture”. Consequently, all the possible implementations of a given algorithm onto a given architecture are described in terms of graph transformations. An implementation consists in distributing and scheduling a given algorithm onto a given architecture. “Adequation” amounts to choose one optimized implementation among all the possible ones while taking into account the real-time characteristics. This is achieved automatically through off-line multiprocessor real-time schedulability analyses combined with optimization heuristics. The result of the adequation is a scheduling table. Finally, from the scheduling table our graph models allow to generate automatically, as an ultimate graph transformation, two types of codes: dedicated distributed real-time executives, or configuration of standard distributed real-time executives, e.g. Linux, Linux/RTAI, Windows/RTX, OSEK, etc., for processors, and structural VHDL for specific integrated circuits. In this paper we focus only on the specifications and on the optimized implementation. See [29] for details about the code generation.

3.4.1 Algorithm model

If we want to use efficiently multiprocessor architectures offering some available parallelism, algorithms should be specified with at least as much potential parallelism as the available parallelism of the architecture. Moreover, since we want to be able to compare the implementation of an algorithm onto different architectures, the algorithm graph must be specified independently of any architecture graph.

Our algorithm model is an extension of the well-known data-flow model from Dennis [30]. It is a directed acyclic hyper-graph (DAG) that we call “conditioned factorized data dependence graph”, whose vertices are “operations” and hyper-edges are directed “data or control dependences” between operations. Hyper-edges are used to model data diffusion. The data dependences define a partial order on the execution of the operations, called “potential operation-parallelism”. Actually, two operations which are not data-dependent may be executed, in any order, on a unique processor or in parallel on two different processors. Each operation may be in turn described as a graph to allow hierarchical specification of algorithms. Therefore, a graph of operations is also an operation. Operations which are the leaves of the hierarchy are said “atomic” in the sense that it is not possible to distribute this kind of operation on more than one processor. The basic data-flow model was extended in three directions, firstly infinite (resp. finite) repetitions in order to take into account the reactive aspect [31] of real-time systems (resp. “potential data-parallelism” similar to loop or iteration in imperative languages for $i=1$ to $n$ do...), secondly “state” when data dependences are necessary between repetitions introducing cycles which must be avoided by specific vertices called “delays” (similar to $z^{-n}$ in automatic control), thirdly “conditioning” of an operation by a control dependence similar to conditional control structure in imperative languages (if...then...else...).

A conditioned operation is a hierarchical vertex which contains several sub-dataflow graphs. According to the value of its specific input called “condition” only one of the possible sub-dataflow graphs will be executed during the considered reaction. A finitely repeated operation is also a hierarchical vertex which contains $N$ times the same sub-dataflow graph. Conditioned operations may contain repeated operations, and vice-versa.

We denote by $G_{al} = (O, D)$ the conditioned factorized data dependence graph where $O$ is the set of vertices (operations) and $D \subseteq O \times O$ is the set of data dependences defining a partial order on the execution of the operations. As mentioned before, this graph is repeated infinitely such that every repetition corresponds to a reaction of the real-time system. During the functional specification, we do not consider non-functional specification, i.e., physical time used to define timing characteristics and architecture. Each reaction defines a logical time and according to the synchronous language principles presented previously, we do not care about physical time taken by the execution of operations which will be considered further on in the implementation model. Thus, every data dependence $d \in D$ is associated with a signal which is an infinite sequence of events taking values in a domain $D$. But we do not have the absent event (#) like in the Signal language and consequently all the signals have the same clock. This is the reason why SynDEx is considered “mono-clock”.
3.4.2 Architecture model

An architecture is composed of processors, possibly of different types, specific integrated circuits performing a unique function, and point-to-point or multi-point communication media. All these components are together connected. We propose an architecture model which is at an intermediate level between high level models neglecting details of the architecture and accurate low level models. The complexity of this model is sufficient to enable optimizations while it is not too fine leading to combinatorial explosions.

Our architecture model is an oriented graph, denoted by $G_{ar} = (V, E)$ where $V$ is the set of vertices, and $E$ the set of edges. $V$ corresponds to four kinds of Finite State Machines (FSM) called operator $(V_o)$, communicator $(V_c)$, memory $(V_m)$ and Bus/Mux/Demux with or without arbiter $(V_b)$: $V = V_o \cup V_c \cup V_m \cup V_b$, and $V_o \cap V_c \cap V_m \cap V_b = \emptyset$. Each edge $s \in E$ represents a connection between an input and an output of FSMs.

There are two types of memory vertices, RAM (Random Access Memory, $S_{RAM} \in S_M$) and SAM (Sequential Access Memory, $S_{SAM} \in S_M$). RAM memories are used to store operations of the algorithm graph, in this case we call them $RAM_P$ (program memory). When RAM memories store only data we call them $RAM_D$ (data memory). We call them $RAM_{DP}$ when they store both program and data. A RAM may be shared (i.e., connected to several operators and/or communicators), then it may be used for data communications. SAM memories are always shared since they are only used to communicate data using the message passing paradigm. In a SAM, data must be read in the same order as it has been written (as FIFO), all access is then said to be synchronized whereas in a RAM it is not synchronized since it is possible to read data independently of the order of the write operation. The synchronized, or not, properties are exploited during the code generation. SAM may be point-to-point or multi-point (bus), supporting or not broadcasting. Each memory is characterized by its size and its access bandwidth.

Each operator sequentially executes a finite subset of the algorithm operations stored in a $RAM_P$ (or $RAM_{DP}$) which must be connected to the operator. The WCET of an operation depends on the operator and the memory characteristics. An operation executed by an operator reads input data stored in a connected $RAM_D$ (or $RAM_{DP}$) and produces output data which is written in the $RAM_D$ (or $RAM_{DP}$) connected to it.

Each communicator sequentially executes communication operations stored in their connected $RAM_P$ (or $RAM_{DP}$). These operations transfer data from one memory ($SAM$, $RAM_P$, $RAM_{DP}$) connected to the communicator into another memory connected to the same communicator. The WCCT of a communication depends on the size of data to be transmitted but also depends on the available bandwidth computed from all parameters of the edge.

Bus/Mux/Demux (BMD) vertices are used to model the bus, the multiplexer and the de-multiplexer of an architecture. When a single sequencer (operator or communicator) requires to access more than one memory, a BMD vertex must be inserted between the sequencer and each memory. When a memory is shared by several sequencer vertices, a BMD including an arbiter (a BMDA) must be inserted between all sequencers and the shared memory.

In our model, a processor or a specific integrated circuit corresponds to an architecture subgraph made of one operator and optionally one or several communicator(s) and BMD(s). The architecture model is detailed in [32] as well as the set of connection rules for building a valid architecture graph. With respect to this model, the only difference between a processor and a specific integrated circuit is that the latter is only able to execute a unique operation whereas a processor may execute, sequentially, several operations. Therefore, only a unique operation may be distributed on a specific integrated circuit.

3.4.3 Implementation model

Given a pair of algorithm and architecture graphs, we transform the algorithm graph according to the architecture graph in order to obtain an implementation graph. This transformation corresponds to a distribution and a scheduling of the algorithm graph.

A distribution, also called allocation or partitioning, is modeled by the relation $dist$ applied to a pair of algorithm and architecture graphs. This produces a distributed algorithm graph $G’_{al}$ such that: $(G_{al}, G_{ar}) \xrightarrow{dist} (G’_{al}, (.))$. The distribution is obtained in three main steps. Firstly, a spatial allocation of the operations onto the operators, leading to inter-operator edges (data dependence between two operations belonging to two different operators). Secondly, each of these edges is replaced by a linear subgraph (sequence of edges and vertices). The new vertices of this subgraph are communication operations which are allocated to the communicators belonging to the different processors forming the route the data dependence has been allocated to. A route is a path in the architecture graph. Thirdly, for each operation allocated to an operator, new “memory allocation vertices” are added and allocated to one of the RAM memories connected to the operator.
A scheduling is modeled by the relation $sched$ applied to a pair $(G'_{al}, G_{ar})$ so that $(G'_{al}, G_{ar}) \xrightarrow{sched} (G''_{al}).$ For each operator (resp. communicator) the scheduling is a temporal allocation of the operations (resp. communication operations) allocated to this sequencer. This amounts to create new “precedence edges” without data transfer ($e_{p} \in E_{p}$) in order to transform the partial order associated with the operations allocated to an operator (resp. communication operations allocated to a communicator). This is necessary because operators (resp. communicators), which are FSMs, require a total execution order between the operations (resp. communication operations) allocated to them. This order must be compatible with the precedences involved by the data dependences of the algorithm graph.

From a pair of algorithm and architecture graphs we get the set of all the possible implementations by composing the two previous relations: $(G_{al}, G_{ar}) \xrightarrow{dist} (G'_{al})$ where $G''_{al} = (O \cup V_{alloc, o} \cup V_{alloc, p} \cup V_{c} \cup V_{i})$. The partial order of the algorithm graph $G_{al}$ has been transformed in an other partial order, thanks to the added precedence edges $E_{p}$, according to the available parallelism of the architecture graph. In [32] we prove that by construction the partial order of $G''_{al}$ includes the partial order of $G_{al}$.

For a given pair of algorithm and architecture graphs, there is a large but finite number of possible implementations, among which we need to select an optimized one, i.e., which satisfies real-time constraints and minimizes some cost function. This optimization problem, as most resource allocation problems, is equivalent to a “bin packing problem” known to be NP-hard, and then its set of solutions is tremendously huge for realistic applications. This is the reason why we use the term “optimized implementation” rather than “optimal implementation”. Consequently, we need to use a heuristic since exact algorithm, giving the optimal solution, should take too much time. The heuristic we chose is a fast and efficient “greedy list scheduling” algorithm, with a cost function based on the “critical path” and on the “schedule flexibility” of the implementation graph. It takes into account the execution WCET of operations and WCCT of communications. This heuristic can be improved by using “local searches” based on “backtracking” whereas greedy heuristics do not backtrack.

As mentioned before, real-time characteristics are associated with every operation. Some characteristics are dependent of the architecture like the WCET and the WCCT but other ones are not. A period and a deadline are characteristics associated with every operation, which are independent of the architecture. An operation temporally characterized like that is usually called a “real-time task” in the real-time scheduling community, as presented for example in the Liu & Layland’s seminal paper [33]. Consequently, we consider data-dependent task systems according to the algorithm graph.

The proposed heuristic is composed of three main steps: The first step performs an unrolling of the algorithm graph by duplicating every operation according to the ratio between its period and the Least Common Multiple (LCM) of all the operation periods.

The second step performs a multiprocessor real-time schedulability analysis assuming the inter-processor communications have no cost. This analysis is based on schedulability conditions that are different depending on whether the task system is preemptive or non preemptive.

The third step performs the actual scheduling by computing for every processor the start times of every operation that has been allocated to this processor, for computation as well as communication operations. Therefore, the communication costs are, now, taken into account. During this last step the response time of every operation, distributed and scheduled on a processor, is minimized according to a cost function which takes into account its schedule flexibility and the increase of the critical path when an operation has to receive data because it is data-dependent and has no cost. The cost function is detailed in [34]. Moreover, the cost of the communication is also minimized by choosing the shortest routes of the architecture graph.

We propose two versions of this heuristic. The first one considers non-preemptive tasks [35] and thus is best suited for safety critical applications but has a worse schedulability ratio than the other version which considers preemptive tasks [36]. Finally, the off-line scheduling heuristic produces a scheduling table that we call further on in the paper “adequate result”. This table is used to generate the real-time embedded distributed code.

4 From AADL to SIGNAL and clock synthesis

An AADL model describes the architecture and execution characteristics of an application system in terms of its constituent software and execution platform components and their interactions. An AADL feature description is mostly made of: i) properties that abstract the behavior of those components as well as their non functional characteristics; ii) interconnections; and iii) scheduling. The polychronous se-
The semantics of SIGNAL and its constraint programming style make it closer to AADL timing semantics and architectural style than other pure synchronous or asynchronous models.

Such characteristics depend on the hardware executing the software, where the timing properties and execution binding properties are associated. In this section, we mainly handle timing and binding properties of AADL components with regard to our polychronous MoC. Syntactic aspects in the transformation are only briefly described. The timing modeling of AADL applications in the framework of POLYCHRONY is based on these properties.

4.1 AADL to SIGNAL translation principles

The transformation from AADL to SIGNAL is a recursive function applied to an AADL component and its features.

A package, which represents the root of an AADL specification, is transformed into a SIGNAL module, the root of a SIGNAL program. Both allow to describe an application in a modular way.

Each AADL component and its interface is translated into a SIGNAL process and its corresponding interface. Assembly of AADL components, considering connections and binding, is represented by a parallel composition of corresponding translated SIGNAL processes.

- Data components are translated into signals that are present at every instant of the overall application. SIGNAL provides standard data types (Boolean, integer, char, string, ...), array, record and union (bundle) types. It also provides extern type declarations allowing to define abstract constants and standard operators ( = ) for those types.
- AADL data ports are considered as event data ports that are present on appropriate Dispatch. We consider an event port as an event data port that can be tt or #.
- AADL (remaining event data) ports are translated into SIGNAL processes that contain the required fifos (or memory cell), synchronizations, temporal and non functional properties, and possible local behavior. A "port" process is connected to its owning thread and to its environment owner (as described by the port connections).
- AADL processes and threads are translated into SIGNAL processes.
- AADL subprograms are considered as threads with specific Dispatch and input/output synchronization. When such a subprogram does not contain subprogram calls, it is declared as an extern SIGNAL process.
- AADL properties are translated into SIGNAL property processes (a set of constraints without output).

Each SIGNAL process resulting from translation following the above rules contains:

- an interface consisting of input/output signals translated from the features (ports) provided by the AADL component type;
- additional control signals depending on the component category (for instance Dispatch and Deadline for a thread);
- a body made of the composition of a process that defines the behavior of the AADL entity and a process that defines its non functional properties;
- a declaration area for processes resulting of the AADL local subcomponents translation and the above structuring processes.

4.2 Timing properties of components

AADL supports an input-compute-output model of communication and execution for threads and port-based communication (Fig. 4). The inputs of a thread received from other components are frozen at a specified point, represented by Input_Time property (by default the Dispatch time), during thread execution, and made available to the thread for access. From that point on, its content is not affected by the arrival of new values for the remainder of the current execution until an explicit request for input, e.g., the two new arrival values 2 and 3 (in Fig. 4) will not be processed until the next Input_Time. Similarly, the output is made available to other components at time specified by Output_Time property (for data ports by default at Complete or Deadline time depending on the associated port connection communication type).

![Fig. 4 A time model of the execution of a thread](image)

The key idea for modeling AADL computing latency and communication delay in SIGNAL is to keep the ideal view of instantaneous computations and communications, moving computing latency and communication delays to specific memory processes, which introduces delay and well suited synchronizations [37].
4.3 Example: translation of an in event data port

Event data ports are intended to message transmission. An event data port can have a queue associated with it. The default port queue size is 1 and can be changed by explicitly declaring a Queue_Size property association for the port. Queues will be serviced according to the Input_Time at order (fifo). For an in event data port, the items are frozen after the event data port is called. Two fifos are used: one for in event data port (write_flow) and one for the output event port called read_flow. The in FIFO is assigned to the port variable (implemented as constraint in the property part). This mechanism induces that the content of the port which is accessible to the application does not change during the execution of a dispatch even though the sender may send new values, and assures an input-compute-output model of thread execution.

To take into account the different events used to specify the semantics of the port, the translation of an in event data port \( xx \), with fifo as Queue_Processing_Protocol, is implemented as an instance of Signal process model \( xx_{\text{InEventDataPort}} \), composed of \( xx_{\text{InEventDataPort\_Behavior}}() \) and \( xx_{\text{InEventDataPort\_Property}}() \) subprocesses (Fig. 5).

![Fig. 5 In event data port translation](image)

The process \( xx_{\text{InEventDataPort\_Behavior}}() \) calls the \( \text{InEventDataPort\_Behavior}() \) process model, which is defined in a Signal library. Two fifos are used: in fifo that stores the receiving in event data (write_flow), and frozen fifo that is accessible by the thread (through read_flow). At Frozen_time_event (depending on the Input_Time, it is provided by the scheduling part), the actual items of the in fifo are frozen: certain items are moved to frozen fifo. The inputs arrived after the Frozen_time_event will be available at the next occurrence of Frozen_time_event. The signal Reference_time_event is the event sent when the port is read.

The properties associated with the port are implemented as Signal processes composed in the body of \( xx_{\text{InEventDataPort\_Property}}() \): the property values are provided as parameters, and a Signal process (called Input_time_property) is defined in a library to verify whether the constraints are satisfied using the input signals Frozen_time_event and Reference_time_event (in a simulation for example).

4.4 Example: translation of a thread

To further illustrate the principles of the translation, consider (Listing 5) a thread \( T \) implemented by \( T_{\text{RS}} \) with an input event data port called iport and an output event port called oport. A property Input_Time is assigned to the port iport: Input_Time \( \rightarrow \) (Time \( \Rightarrow \) Start; Offset \( \Rightarrow \) 10ms..15ms;). It specifies that the input is frozen at some amount of execution time from the beginning of execution (Start). The time is within the specified time range (10ms..15ms).

```
| thread T |
| features |
| iport: in event data port Integer |
| oport: out event port Integer |
| Input_Time => (Time => Start; Offset => 10 ms..15 ms.;); |
| end T |
```

Listing 5 Example of an AADL thread

The translation in Signal is given in Listing 6. The interface contains the input/output signals that represent the features provided by the thread declaration, and also some added control signals (top, ctl1, time1, ctl2, alarm). These added signals are provided by the scheduler (input ones) or sent to it (output ones). The body is composed of processes that represent the translation of the properties associated with the thread (\( T_{\text{RS\_Thread\_property}}() \) and the behavior of the thread \( T_{\text{RS\_Thread\_behavior}}() \), which may be described, for example, by some transition system), and the ports with their timing semantics \( (\text{iport\_InEventDataPort}(), \text{oport\_OutEventPort}()) \). The translation of the iport follows the principles described above. The translation of an output event port is quite similar. The properties of a port (default ones or specified ones) are also translated \( (\text{iport\_InEventDataPort\_Property}()) \) and \( \text{oport\_OutEventPort\_Property}() \): the port queue processing (size, protocol, overflow), and the input/output timing (Input time, Output time) are explicated.

```
process T_RS_Thread =
( ? integer iport:
  CTL1 ctl1:
  T\_TIME\_EVENT time1:
  event top:
  ! boolean oport:
  CTL2 ctl2:
  boolean Alarm;
) |
( % thread behavior %
  (l_iport, ctl1) := T_RS_Thread\_behavior()
  (l_iport, ctl1 Dispatch, Start, ctl1 Resume)
```
Restrictions. The current version of the AADL to SIGNAL translator does not yet implement all the AADL components. Among non-implemented components we distinguish:

- expected ones: mode, behavior annex, bus access;
- components that will not be implemented: thread group, subprogram group, virtual processor, virtual bus, flow;
- other ones that are partly translated (translated with some restrictions on properties such as Dequeue_Protocol restricted to one item, Queue_Processing_Protocol restricted to fifo).

4.5 Processor and its affine-scheduling

An AADL model is not complete and executable if the processor-level scheduling is not resolved. A scheduler is therefore required to be integrated so that a complete model is used for the subsequent validation, distribution and simulation. A scheduling based on affine clock systems [38] is thus developed for each AADL processor. A particular case of affine relations is affine sampling relation, expressed as \( y = \lfloor d \cdot t + \phi \rfloor \), of a reference discrete time \( x \) \((d, t, \phi \) are integers): \( y \) is a subsampling of positive phase \( \phi \) and strictly positive period \( d \) on \( x \). Affine clock relations yield an expressive calculus for the specification and the analysis of time-triggered systems. The scheduling based on the affine clocks can be easily and seamlessly connected to Polysyncrone for formal analysis.

4.6 Binding

For a complete system specification, the application component instances must be executed by the appropriate execution platform components. How to combine the components of a system to produce a physical system implementation is called binding in AADL.

A process is bound to the processor specified by the Actual_Processor_Binding property. Support for process/thread execution may be embedded in the processor hardware, or it may require software that implements processor functionality. Such software must be bound to a memory component that is accessible to the processor via the Actual_Memory_Binding property. The interactions among these execution platform components are enabled through a bus via the Actual_Connection_Binding property.

Binding properties are declared in the system implementation that contains in its containment hierarchy both the components to be bound and the execution platform components that are the target of the binding. This binding information, as well as timing properties, is reflected in the generated SIGNAL program by SIGNAL property processes representing these properties. In a further step, the SIGNAL program is analyzed and useful properties are kept as specific pragmas.

4.7 Timing analysis and clock synthesis

The previous translation from AADL to SIGNAL concentrates on the timing modeling of software architectures in AADL. This modeling enables a formal timing analysis and clock synthesis, based on the polychronous MoC.
Timing analysis

Timing analysis in this paper mainly refers to analyzing clock relations based on clock hierarchy. The clock hierarchy of a process is a component of its Data Control Graph (DCG). The DCG is made of a multigraph $G$ and a clock system $\Sigma$.

Description of $G$:
- A node $n$ in $G$ represents a polychronous process $\text{process}(n)$; it has input signals $\text{input}(n)$ and output signals $\text{output}(n)$ which are those of $\text{process}(n)$; it has a clock $\text{clock}(n)$ which is either a Boolean variable or a clock formula acting as a guard.
- A directed edge $(n_1, n_2, a, h)$ links a source node $n_1$ to a target node $n_2$; it is labeled by the name $a$ of the signal that is sent by $n_1$ and received by $n_2$; it has a clock $h$. The signal $a$ is transmitted when $h$ is $tt$. Two nodes $n_1$ and $n_2$ may be linked by several edges.

Description of $\Sigma$:
- Given $X$ the set of signal names and $B \subseteq X$ the set of Boolean signal names, an elementary formula is either $\hat{0}$ that stands for $\#$, $\hat{x}$ that denotes the clock of a signal $x$, $[b]$ that denotes the clock representing the $tt$ occurrences of a Boolean signal $b$, or $[-b]$ that denotes the clock representing the $ff$ occurrences of a Boolean signal $b$. The set of elementary formulas is denoted $EF$.
- The set of clock formulas $CF$ is the smallest set that satisfies:
  - $EF \subseteq CF$
  - for all formulas $f_1, f_2 \in CF$, $f_1 \hat{\circ} f_2$ and $f_1 \hat{\circ} f_2$ belong to $CF$.
- A clock equation is a class of equivalent clock formulas.
- $\Sigma$ is a forest (set of trees) of clock equations.
- The node $n$ of a tree is a clock equation; it contains the list of signals $\text{signal}(n)$ the clock of which is equal to this class.
- A tree is an endochronous process.
- Each subtree $S$ of a tree $T$ is itself an endochronous process; its clock is a (“non recursive”) function of the signals $\text{signal}(n)$ where $n$ is a node in $T$.

The following DCG levels are distinguished (Fig. 6) in the compiling stages of Signal programs:

- Every Signal process is represented by a general DCG as described above, the DCGPoly; its clock system $\Sigma$ can be normalized using:
  - elementary Boolean to event transformations such as $E \text{ when } (a \text{ and } b) \rightarrow (E \text{ when } a) \text{ when } b$,
  - transformations based on algebraic properties such as transitivity of partial order.
- When the forest $\Sigma$ of a process $P$ is a tree, $P$ is an endochronous process; its DCG belongs to the DCGEndo level. One can reduce a consistent forest to a tree by:
  - building a triangular system of equations equivalent to the original one, and
  - adding supplementary required parameters to the system as it is done to build memoryCell from preMemoryCell (Listings 2, 3).
- An endochronous process $P$ may be free of event type signals and binary event expressions $(\hat{\circ}, \hat{\circ})$. It can still use signals that are not always defined. Then $P$ is said to be a Boolean clocked process and its DCG belongs to the DCGBool level. It is always possible to transform a usual endochronous process into a Boolean clocked process that is equivalent: an event signal $h$ can be associated with a (possibly new) Boolean signal $c$ such that each occurrence of $h$ is replaced by $true$ when $c$ provided that $c$ is equal to $\#$ or $ff$ when $h$ is $\#$.
- A Boolean clocked process in which all Boolean clock signals are synchronous is named flat. Its associated DCG belongs to the DCGFlat level. The clock of these Boolean clock signals is the root of the clock hierarchy. The clock hierarchy has at most two levels: subclocks of the root represent the instants at which Boolean clocks are $tt$: Boolean clocks are used as guards for the processes associated with the nodes of the DCG. It is always
possible to transform a Boolean clocked process into an equivalent flat one. Quartz [39] and SynDEx belong to this class of processes.

Clock synthesis
In a clock tree with only one root, the simulation clock (the fastest clock in the system) is based on the root clock (i.e., the *tick*). In this case, the system is endochronous and it is possible to build the unique deterministic behavior in the code generation. But if there is no common root for all the trees, i.e., there is no fastest clock, the system is polychronous, and non deterministic concurrency is thus introduced. In an AADL multi-processor specification, it is generally hard to find the fastest clock, as each component may have its own activation clock or frequency. Code generation considering the deterministic behavior is therefore difficult, even impossible. To tackle this issue, independent clocks, particularly the root clocks in different trees, are required to be synchronized. This synchronization, called *endochronization*, can be performed in an ad-hoc way by the compiler, or in a specific way in a manual manner. More details can be found in [20]. A more general clock synchronization method via controller synthesis is also possible [40]. Endochronization leads to the transformation from the DCGPoly level to the DCGEndo level.

5 From SIGNAL to SynDEx and architecture exploration
Both SynDEx and SIGNAL belong to the family of synchronous/polychronous languages. However, there are still differences to consider for the translation. First, SIGNAL is based on a polychronous MoC, whereas SynDEx relies on the synchronous MoC. Clock synthesis is therefore required to endochronize multi-clocked specifications before the translation (presented in Section 4). Secondly, system representations of SIGNAL and SynDEx, based on graphs, are at different abstraction levels. Transformation between these levels is thus required. The DCGFlat level is appropriate for translating SIGNAL programs to SynDEx algorithm graphs. At this level, all the clocks have been expanded up to the most frequent clock, and state variables are also defined at this most frequent clock. Other signals can have a “don’t care” interpretation at the instants at which they are #. This corresponds to the SynDEx interpretation. Execution platform and real-time characteristics, preserved in SIGNAL pragmas, are translated into SynDEx architecture and constraints.

5.1 Syntactic translation rules
The correspondence between SIGNAL and SynDEx representations is defined in two aspects:

- the first one is related to the algorithm [41], including clock hierarchy, nodes and dependences. Clock hierarchy plays the semantic role in SIGNAL and serves as a structural *backbone* in the translation: each clock in the hierarchy is associated with a SynDEx algorithm graph. This algorithm combines the computations to be performed according to this clock, which is represented by a SynDEx condition. The translations of nodes and dependences are mainly syntactic, which will be detailed later.

- the second one is related to architecture, allocation constraints, temporal information: these information are preserved in *pragmas* of the original SIGNAL programs and will be finally translated into SynDEx *architecture* and *constraints* (presented in next subsection).

As this translation includes large syntactic details, only some key rules are briefly presented.

In the DCGFlat representation of a SIGNAL program, the clock hierarchy, which is a tree, contains at most two levels: the root of the tree (the *tick*), and the children level, which comprises less frequent clocks defined by the *tt* values of the Boolean clock signals. Let us denote these Booleans by $b_1, \ldots, b_n$, and the clock [$b_j$] denotes the instants at which $b_i$ is $tt$. These clocks are subclocks of the clock *tick*. Since during the SIGNAL compilation, each clock is associated with the subgraph containing the nodes that are defined at this clock, the translation strategy is the following: whenever a graph is associated with a clock, this graph is translated as an *algorithm* in SynDEx.

Translation of the clock hierarchy
Any graph $g$, associated with a clock that is a child of the *tick*, is translated into a SynDEx algorithm (which is a graph of operations), with a condition defined from the clock [$b_j$] of this graph. The translation of $g$ can be sketched as follows:

```
def algorithm P_bi:
  ? ... ! ... 
  conditions: bi = 1
```

The top level graph $P$ contains the nodes associated with the *tick*, as well as references to each of the graphs associated with the subclocks of *tick*. The translation of $P$ is a SynDEx algorithm $P$ with conditions: true.
The input and output signals (called ports in SynDex) are translated from signals communicating between clock graphs. Interface signals of a whole program are translated into sensors and actuators in SynDex.

Translation of nodes

A SynDex algorithm is mainly composed of four elements: ports, conditions, references, and dependences. Ports and conditions are generated along with the translation of the clock hierarchy, references and dependences are generated according to the attached nodes of the considered clock. In DCGFlat, the nodes can be categorized into:

- **Constants**. The constants are explicitly represented by references to constant vertices in SynDex.
- **Equations**. A general form of equations is considered: \( X := F(Y_1, \ldots, Y_n) \), where \( F \) represents SIGNAL n-ary operators. Only elementary expressions are considered as it is always possible to translate other expressions into a composition of elementary definitions. SynDex defines libraries to provide basic algorithm (atomic operation) declarations for SIGNAL elementary operators. Let \( X := F(Y_1, \ldots, Y_n) \) an operation of type \( T \), a corresponding SynDex algorithm named \( F \) with type \( T \) is declared in a predefined library. For example, the predefined algorithm corresponding to the SIGNAL default for integer type (as it is represented in the DCG) is defined as follows (dependences are discussed below):

  ```python
  def algorithm default:
    ? int i1; ? bool i2; ? int i3;
    ! int o;
    conditions: i2 = 1;
    references:
    dependences: strong_precedence_data i1 -> o;
    conditions: i2 = 0;
    references:
    dependences: strong_precedence_data i3 -> o;
    description: "x := (i1 when i2) default (i3 when (not i2))"
  ```

When translating such an equation, only a reference to \( F \) associated with \( T \) is used, \( T/F \).

Note that there is no need of predefined algorithm representing the when operator. Indeed, in the DCG representation, for a process \( X := E \) when \( b \), \( b \) represents the clock of \( X \) and the corresponding node is associated with the subclock \([b]\). This subclock is taken into account when translating the clock hierarchy, thus, in the algorithm corresponding to this subclock, \( X := E \) when \( b \) may be considered as \( X := E \).

- **Memorizations**. The translation of a memorization \( ZX := X $ \) is similar to the translation of the equations, except that it is a reference to a memory.

- **Partial definitions**. Let us consider the following partial definitions: \( X ::= E_1 \text{ when } b_1, \ldots, X ::= E_n \text{ when } b_n \). In the DCG representation, \( b_1, \ldots, b_n \) are clocks that are exclusive. When translating such partial definitions, the original graph is rewritten as follows:

  ```python
  X ::= (E_1 \text{ when } b_1) \text{ default } \ldots \text{ default } (E_n \text{ when } b_n)
  ```

- **Process calls**. The called process \( Q \) is defined explicitly as a SynDex algorithm. A call to the process \( Q \) is translated into a reference to the algorithm \( Q \).

- **External calls**. The translation of an external process is similar to the translation of any process, except that the body of the algorithm is empty since the process is not defined in SIGNAL.

Translation of dependences

After the references are declared for each node, dependences between them are set. The dependences (the edges of the SynDex graph) are built either from data-flow connections of the SIGNAL graph or from explicit precedence relations.

For each data-flow connection \( d \) from the \( m^{th} \) output of a node \( p_i \) to the \( n^{th} \) input of a node \( p_j \), the translation associates a strong precedence data with the corresponding translated vertices in SynDex:

\[
\text{strong_precedence_data p_k.o_m} \rightarrow p_l.i_n;
\]

(an example of such a dependence has been given for the predefined default).

Another kind of dependence is used to express precedence relations. A precedence \( E_i \rightarrow E_j \) in the DCG is translated directly as a corresponding precedence in SynDex.

5.2 Translating non-functional aspects

In the ASME2SSME transformation, execution platform, timing properties and binding specified in AADL are preserved in the pragma part of SIGNAL programs. In the translation from SIGNAL to SynDex, the information of execution platform, related to processing and communication units, are taken from the pragnas and translated into an architecture graph in SynDex. For example, the architecture of the SD-SCS case study is translated as follows (extract):

```python
def architecture IMA:
  operators:
    CPIOM CPIOM1;
    RDC RDC1;
```
medias:
AFDX AFDX1 no_broadcast;
connections:
CPIOM1.io AFDX1;
RDC1.io AFDX1;

CPIOM and RDC are types of computing units and AFDX is the type of bus. The computing units are connected via the bus.

Timing properties, such as period, deadline, computation time, are also taken from pragmas and finally set to corresponding SysDEEx nodes. Here is an example from SDSCS, where the computing time of AIRSPEED and CLL on RDC is 1000 units:

```
extra_durations_operator RDC:
    AIRSPEED = 1000;
    CLL = 1000;
```

The binding information of software and hardware are translated into groups and allocation constraints in SysDEEx, i.e., all the software components allocated on the same processing unit are put into the same group, and then an allocation constraint is specified in SysDEEx so that this group is allocated on the same processing unit. In the following example, on_ground and in_flight are declared in group g_rdc1, which is allocated onto the computing units, RDC1 or RDC2.

```
operation_group g_rdc1:
    ["aadl_door\on_ground,attach_all]
    ["aadl_door\in_flight,attach_all]
    absolute constraint:
    g_rdc1 on IMA.RDC1 IMA.RDC2;
```

5.3 Architecture exploration with SysDEEx

SysDEEx allows the designer to explore manually and/or automatically the design space solutions using schedulability analyses and optimization heuristics for distributing (allocating) and scheduling the application functions (operations) onto the processors while satifying real-time constraints and minimizing computing resources. The exploration results (multiprocessor scheduling table) show the real-time behavior of the functions allocated onto components of the architecture, i.e., processors, specific integrated circuits, and communication media. This approach conforms to the typical architecture exploration process [42, 43]. In addition, code is automatically generated as a dedicated real-time executive, or as a configuration file for a resident real-time operating system such as Linux, Linux/RTAI, Windows/RTX, Osek [44].

6 An avionic case study

SDSCS is a generic simplified version of the system that manages passenger doors on Airbus series aircrafts. As a safety-critical system, in addition to the fulfillment of safety objectives, high-level modeling and component-based development are also expected for fast and efficient design.

![Fig. 7 A simple example of the SDSCS architecture](image)

In this example, each passenger door has a software handler, which achieves the following tasks: 1) monitor door status via door sensors; 2) control flight lock via actuators; 3) manage the residual pressure of the cabin; 4) inhibit cabin pressurization if any external door is not closed, latched and locked. The four tasks are implemented with simple logic that determines the status of monitors, actuators, etc., according to the sensor readings. In addition to sensors and actuators, SDSCS is equipped with other hardware components, e.g., CPIOMs (Core Processing Input/Output Modules) and RDCs (Remote Data Concentrators) are connected via the AFDX (Aircraft Full DuplexX) network (Fig. 7). Sensors and actuators are also connected to RDCs and communicate with other systems via AFDX.

![Fig. 8 A partial view of the SysDEEx algorithm graph](image)

An overview of SDSCS modeled in AADL is shown in
Fig. 3. The whole system is presented as an AADL system. The two doors, modeled as subsystems, are controlled by two processes doors_process1 and doors_process2. Each process contains three threads to perform the management of doors. All the threads and devices are periodic and share the same period. The process doors_process1 (resp. doors_process2) is executed on a processor CPIOM1 (resp. CPIOM2). The bus AFDX connects the processors, CPIOM1 and CPIOM2, and devices that model the sensors and actuators, such as DPS, OCU, etc.

The case study was successfully transformed with the toolchain described in Fig. 2. From a .aadl file (AADL textual file), we first get .aaxl (AADL model file) with OS-ATE. Aided by high-level and low-level APIs, ASME2SSME is able to read the AADL models. While performing the model transformation towards the SSME model, a thread-level scheduler is also integrated. The SSME model is transformed into Signal code, from which static analysis can be performed. The Signal code can be transformed into .sdx file (SyndeX code), as well as other format for model checking [45] and simulation [46]. From the SyndeX tool, the .sdx file is read, and the adequation (optimized multiprocessor real-time scheduling) can be carried out. Fig. 8 illustrates a partial view of the SyndeX algorithm graph describing a partial order on the execution of the functions (vertices of the graph) through data dependences (edges of the graph). This algorithm graph is translated from the AADL functional part of SDS CS. Fig. 9 shows the SyndeX architecture graph, which is translated from the AADL architectural part of SDS CS. The architecture is composed of two processing units (CPIOM1 and CPIOM2), two concentrators (RDC1 and RDC2) and a communication media (AFDX1) of AFDX bus type, that are together connected.

Fig. 9 The SyndeX architecture graph

Fig. 10 illustrates a partial view of the adequation result (scheduling table) obtained with SyndeX: the algorithm (Fig. 8) is distributed and scheduled onto the architecture (Fig. 9). There are five columns in the figure, which represent the five components of the architecture. From left, the first column represents the bus AFDX1. The second and fourth columns represent RDC1 and RDC2 concentrators. The third and fifth columns represent CPIOM1 and CPIOM2 processors. For every column, the start time and the complete time of every operation (function) and every communication operation are given and materialized by a horizontal line. Thus, the processing time, flowing from top to bottom, of every operation is represented by a vertical box. In this case study, the algorithm has more than 150 nodes and the architecture has 5 nodes in SyndeX. The adequation takes about 15 minutes 35 seconds in average. With this toolchain, it is easy to change the configuration of the execution platform and binding. For example, the number of processing units can be changed, as well as the type of processing units and communication media. The influence of these changes has finally a repercussion in SyndeX in order to be exploited. In addition, generation of dedicated distributed real-time executives with optional real-time performance measurement is also possible for different processors and communication media [29]. The names of AADL components are always kept in all the transformations in order to enable traceability. Hence, our approach provides a fast yet efficient architecture exploration for the design of distributed real-time and embedded systems.

Fig. 10 A partial view of the SyndeX adequation result: a scheduling table

7 Related work

Many contributions on schedulability analysis and scheduling tools exist for AADL specifications, such as [3], [47], and [48]. The main reason to choose SyndeX in our work is that SyndeX and Signal have similar synchronous semantics; this proximity reduces the difficulties in the translation.
In addition, the scheduler generated by SynDEx can be integrated, in a manual manner currently, into the original SIGNAL programs for formal verification purpose.

AADL has been connected to many formal models for analysis and validation. The AADL2Sync project [17] and the Compass Approach [4] provide complete toolchains from modeling to validation, but they are generally based on the synchronous semantics, which is not close to AADL timing semantics. AADL2BIP [6] allows simulation of AADL models, as well as application of particular verification techniques, i.e., state exploration and component-based deadlock detection. The AADL2Fiacre project [49] mainly concentrates on model transformation and code generation, in other words, formal analysis and verification are performed externally with other tools and models. The previous projects do not address code distribution and real-time scheduling in a general sense. The Ocarina project [7] considers code generation for distributed real-time and embedded systems, but using formal models and semantics is not reported in the work.

8 Conclusion

In this paper, we present our proposed approach to address architecture exploration based on AADL, POLYCHRONY, and SynDEx. Software architecture is specified in AADL at a high level of abstraction; POLYCHRONY provides the polychronous model of computation, formal analysis, clock synthesis, and transformations to bridge between AADL and SynDEx; SynDEx is finally used for multiprocessor real-time schedulability analyses and resource optimizations, both allowing architecture explorations. With a reduced design cost of embedded systems, our approach and its associated toolchain make it possible to perform architecture exploration at the earliest design stage. An avionic case study is used to illustrate our approach.

In a longer term, as a perspective, the translation between SIGNAL and SynDEx is expected to be bidirectional, i.e., a particular distribution and scheduling determined by SynDEx is automatically synthesized in the SIGNAL programs for the purpose of formal verification, performance evaluation, and other analyses.

References

19. Cost-efficient methods and processes for safety relevant embedded sys-


20. L. Besnard, T. Gautier, P. Le Guernic, and J.-P. Talpin. Compilation of Poly-


24. S. Abramsky and A. Jung. Domain Theory. In S. Abramsky, D. Gab-


25. G. Kahn. The Semantics of a Simple Language for Parallel Program-

ming. Proceedings of the IFIP Congress 74, Stockholm, Sweden, Au-


26. G.D. Plotkin. A Powerdomain Construction. SIAM Journal on Com-


29. T. Grandpierre and Y. Sorel. From algorithm and architecture speci-


31. David Harel and Amir Pnueli. On the development of reactive sys-


33. C.L. Liu and J.W. Layland. Scheduling algorithms for multiprogram-


34. T. Grandpierre, C. Lavarenne, and Y. Sorel. Optimized rapid proto-

totyping for real-time embedded heterogeneous multiprocessors. In Proceed-

ings of 7th International Workshop on Hardware/Software Co-


35. O. Kermia and Y. Sorel. A rapid heuristic for scheduling non-

preemptive dependent periodic tasks onto multiprocessor. In Proceed-


36. F. Ndoye and Y. Sorel. Safety critical multiprocessor real-time schedul-


39. J. Brandt, M. Geminide, K. Schneider, S. Shukla, and J.-P. Talpin. Rep-

resentation of synchronous, asynchronous, and polychronous compo-


42. A.D. Pimentel, C. Erbas, and S. Polstra. A Systematic Approach to Ex-

ploring Embedded System Architectures at Multiple Abstraction Lev-


45. Y. Ma. Compositional Modeling of Globally Asynchronous Locally Synchronous (GALS) Architectures in a Polychronous Model of Compu-


49. B. Berthomieu, J.-P. Bodeveix, P. Farail, M. Filali, H. Garavel, P. Gau-

Huafeng Yu has been an expert research engineer within INRIA Rennes, France. His work is involved in timing analysis, formal verification, simulation, and synthesis of MARTE-based timed systems, AADL, and Simulink in the framework of several European projects, such as CESAR and OPEES. He received his PhD in computer science from Université des Sciences et Technologies de Lille (France) in 2008. He completed his Master’s study in systems and software at Université Joseph Fourier Grenoble 1 (France) in 2005. He is now working in Toyota InfoTechnology Center USA as a senior researcher. His main research interests include model-based systems engineering, automotive and aerospace engineering, embedded systems design, formal methods, and synchronous languages.

Yue Ma has been a post-doc fellow in IRISA / INRIA Rennes, France. She works on the modeling, temporal analysis, formal verification and simulation of Globally Asynchronous Locally Synchronous systems, especially AADL using Polychrony in the framework of European TopCased, CESAR and OPEES projects. She received her PhD degree in computer science from University of Rennes 1 (France) in 2010. She is now working in itemis France as a senior software architect. Her research interests include software engineering, embedded systems design, synchronous programming, AADL modeling and analysis, automotive engineering, such as AUTOSAR and EAST-ADL.

Thierry Gautier is researcher with INRIA. He received the graduate degree from the Institut National des Sciences Appliquées, Rennes, France, in 1980, and the PhD degree in computer science from Université de Rennes 1 in 1984. He is one of the designers of the Signal language, the polychronous model of computation and the Polychrony toolset. His main research interests lie in the safe design of complex embedded systems, including formal modeling, formal validation, and transformations of models to target architectures.

Jean-Pierre Talpin is Senior Researcher with INRIA and leads the project-team who develops the open-source Polychrony environment. He received his PhD from Université Paris VI Pierre et Marie Curie in 1993. He then was a research associate with the European Computer-Industry Research Centre in Munich before to join INRIA in 1995. Jean-Pierre edited two books with Elsevier and Springer, guest-edited more than ten special issues of ACM and IEEE scientific journals, and authored more than 20 journal articles and book chapters and 60 conference papers. He received the 2004 ACM Award for the most influential POPL paper, for his 2nd conference paper with Mads Tofte, and the 2012 LICS Test of Time Award, for his 1st conference paper with Pierre Jouvelot.

Loïc Besnard is currently a senior engineer at CNRS, France. He received his Ph.D. degree in computer science from Université de Rennes, France (1992). His research interests include the software reliability for the design of embedded systems: modeling, temporal analysis, formal verification, simulation, and synthesis of embedded systems. He is involved in the development of the Polychrony toolset based on the synchronous language Signal.

Paul Le Guernic graduated from Institut National des Sciences Appliquées de Rennes in 1974. He performed his Thèse de troisième cycle in Computer Science in 1976. From 1978 to 1984 he had a research position at INRIA. He is Directeur de Recherche in this institute since 1985. He has been head of the “Programming Environment for Real Time Applications” group, which has defined and developed the Signal language. His main current research interests include the development of theories, tools, and methods, for the design of real-time embedded heterogeneous systems. He is one of the architects of the Polychrony toolset.
Yves Sorel is Research Director at INRIA (National Institute for Research in Computer Science and Control), and scientific leader of the Rocquencourt’s team AOSTE (Models and Methods of Analysis and Optimization for Systems with Real-Time and Embedding Constraints). He is member of the Steering Committee of the Competitivity Cluster OCDS/SYSTEM@TIC Paris-Region. His main research topic concerns the analysis and the optimization of distributed real-time embedded systems. He is also the founder of the AAA methodology (Algorithm Architecture Adequation) and of SynDEx, a system level CAD software downloadable free of charge at: www.syndex.org. More details, including a publication list, can be found at: www-rocq.inria.fr/sorel/work.