# Reducing Conservatism in Response Time Analysis of Distributed Systems

Steven Bradley, William Henderson, David Kendall

*Abstract*— **Rate Monotonic Analysis (RMA) is a well-established technique for assessing schedulability of periodic and sporadic tasks which share a processor resource using fixed priority scheduling. An alternative approach to analysing such systems is to build a model which represents the behaviour of the system more dynamically, taking into account the dependency between the tasks.**

## I. Introduction

Since rate-monotonic scheduling (RMS) and rate-monotonic analysis (RMA) were first proposed [8], work has been carried out to extend the basic model of computation from a set of independent periodic tasks with fixed execution times sharing a single processor, and to extend the analysis from simple schedulability. Audsley et al [4] report on the development of the theory supporting fixed priority pre-emptive scheduling, including extensions to account for interdependence of tasks (through blocking) and the analysis of distributed systems. Analyses of end-to-end response times in distributed systems have been carried out using TDMA [9] and more recently with CAN [6] as the communication mechanism.

Although more recent work has taken account of extra delays incurred through interdependence of tasks, ensuring that the analysis remains conservative, one area which has not received so much attention (as noted by Audsley et al [4]) is to take advantage of the restrictions on possible execution paths brought about by interdependence. This has meant that predicted response times can be overly pessimistic. In this paper we aim to address the problem of pessimism by explicitly modelling precedence (ordering) relationships between tasks, and performing an exhaustive analysis of all possible execution paths through a system. The analysis is carried out be performing reachability analysis on a timed hybrid automaton [1].

The motivation for this work is demonstrated in section II through an example which exhibits pessimism under a standard static response-time analysis (RTA). In section III we introduce our system model, and define a simple language (PG) in which such systems can be expressed. We then define a translation from PG into timed hybrid automata in section IV. Section V explains how response time analysis is carried out, and briefly evaluates the results of the analysis. Finally, our conclusions, and the relationship to other work is presented in section VI.

Steven Bradley is with the Department of Computer Science, Durham University, South Road, Durham, DH1 3LE, UK; William Henderson, and David Kendall are with the Department of Computing, University of Northumbria at Newcastle, Ellison Place, Newcastle upon Tyne, NE1 8ST, UK

## II. Example of Pessimism

Our simple example consists of a distributed system with two processors connected by a single CAN-style bus, with non-destructive priority-based arbitration. On processor 1 there are two similar tasks, `Sender A` and `Sender B`. These tasks are periodically triggered, and they each read a sensor, the values of which need to be sent to corresponding receiver tasks, `Receiver A` and `Receiver B`, on processor 2. This is achieved by sending messages `Message A` and `Message B` via the bus. We shall assume that each of the tasks and messages has a fixed resource requirement of 200, so we have essentially two clone sub-systems, system A and system B, each of which has a periodically triggered sender task, which triggers a message upon completion, which in turn triggers a receiver task. The tasks and message of system A have higher priority than those of system B.

In a control situation where the receiver task acts upon the information provided by the sender task to provide feedback into the controlled system, a crucial factor in determining whether the system can be controlled in a stable way is the response time from the reading being taken to the control being applied. Therefore, as well as examining schedulability issues on the processors and bus, it is important to be able to predict the end-to-end response time from the start of the sender task to the end of the receiver task.

The analysis of the response time for subsystem A is reasonably straightforward, but subsystem B is more interesting. A standard response-time analysis adds the response times for each section, with each section having contributions from blocking and execution time. This gives

$$\begin{aligned} MaxBlocking(Sender_B) &= 200 \\ MaxBlocking(Message_B) &= 200 \\ MaxBlocking(Receiver_B) &= 200 \\ TotalExecution(B) &= 600 \\ TotalResponse(B) &= 1200 \end{aligned}$$

However, taking into account the precedence constraints, we can try to construct the actual worst case for the response time of subsystem B. This occurs when Sender B is released and just fails to complete before Sender A is released, which then pre-empts Sender B. Shortly after Sender A completes, Sender B completes and releases Message B. However, Message A has already been released, blocking Message B. Once Message A has completed, Receiver A is released and Message B has access to the bus. By the time that Message B has completed, Receiver A has completed, so Receiver B has access to processor 2 straight

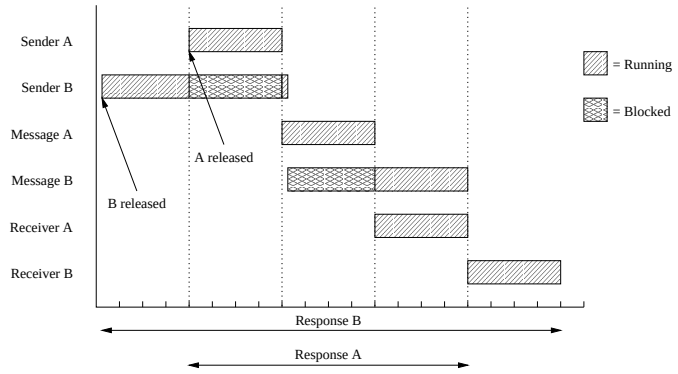


Fig. 1. Example of pessismism

away. This sequence of events is shown in figure 1. Because of the precedence relationships, it is impossible for competition to occur between subsystems A and B for all three of Sender, Message and Receiver if both systems use their maximum resource requirement (200) for all three, giving an actual worst-case end-to-end response time of 1000. This kind of pessimism has been noted experimentally [6].

Such an argument is difficult to construct convincingly by hand, even for a relatively simple example, but the consideration of the many different cases that may occur seems a good candidate for automation. To automate the analysis, however, we first need to declare our assumptions about the systems we analyse, and define the language we are going to use to describe the systems.

## III. PG: A Simple Language for Describing Distributed Systems

In a simple response time analysis, the precedence relationships between tasks are not made explicit, but the period of a task which is triggered by the completion of some other task is inherited from the triggering task. Our model of computation is very similar to that used in rate-monotonic style analyses, except that we explicitly model these precedence relationships. The most basic components in our model are *tasks*, which can be in one of three states: waiting (to be released), ready (to run, but not running) and running. Each task has a simple life cycle: it is released at some *trigger* event, after which it competes for some *resource* (e.g. access to processor) until completion. The time taken to complete will depend on its *bounded resource requirement* (e.g. processing time) and upon the competition for the resource. Competition is managed using fixed priority pre-emptive scheduling, although tasks may be declared as *nonpreemptible*, in which case they will retain access to the resource from their first access until their completion. A trigger event can either be the elapsing of a fixed period, or the completion of some other task. Communication is assumed to take place via a CAN-style bus with non-destructive priority-based arbitration. In this model, messages have the same properties as nonpreemtible tasks. Tasks with critical sections can be modelled as a series of tasks, with the critical sections labelled as nonpreemptible.

```
task senderA on processor1 needs [100,200]
  at priority 0 triggered by period 1300
task senderB on processor1 needs [100,200]
  at priority 1 triggered by period 1400

message messageA on can needs [100,200]
  at priority 0 nonpreemptible
    triggered by senderA on processor1
message messageB on can needs [100,200]
  at priority 1 nonpreemptible
    triggered by senderB on processor1

task receiverA on processor2 needs [100,200]
  at priority 0 triggered by messageA on can
task receiverB on processor2 needs [100,200]
  at priority 1 triggered by messageB on can
```

Fig. 2. Pessimism example in PG

The following assumptions are made:

- each task/message uses exactly one resource
- each task/message has exactly one trigger
- no task/message can be triggered by a task/message that it directly or indirectly triggers (there are no 'loops')
- bounded resource requirements account for any overheads (such as kernel activity, memory management etc)

Important differences between this model and a more standard response time analysis model [9] are that

- Precedence (i.e. triggering or ordering) relationships are made explicit, rather than assuming that all tasks/messages are periodic, with period inherited from their trigger. (Hence the name PG: precedence graph.)
- Jitter (i.e. possible delay of release after period has elapsed) is not included, as this is caused by the inheriting of periods from triggers: if an 'upstream' task or message does not complete in constant time, then its completion will not be purely periodic.

In summary, each task or message has a name, an associated resource, a bounded resource requirement, a priority, and a trigger. These are expressed using a very simple syntax: the example of section II can be expressed in PG as shown in figure 2.

## IV. Modelling Distributed Systems with Hybrid Automata

Before explaining how we translate system descriptions in PG into hybrid automata (section IV-B), we first briefly review the definition of hybrid automata (section IV-A). For a more detailed description (including a more formal definition) see the paper of Henzinger et al [7].

### A. Hybrid Automata

Timed hybrid automata are an extension of timed automata [2]. The basic notion is that of a finite state ma-

chine, extended with real-valued variables. These variables can be used to form enabling conditions for transitions (by comparing with a constant value), and to form invariant conditions for states. Variables can be reset when transitions take place, but cannot be assigned a value other than 0, and cannot be compared with each other. In a standard timed automaton all variables increase their value in line with the increase in global time, and are known as clocks. In a hybrid automaton, the rate at which a variable changes may vary with the state, but this rate will always be a natural number (an integer $n \geq 0$). For the purposes of this paper, we will only need to consider rates of 0 or 1. A hybrid automaton is defined by

- a set of states $S$
- a set of transitions between states $T \subset S \times S$
- a set of variables $V$

A *variable valuation* is a function $v : V \rightarrow Real$ which assigns a real value to each variable.

Associated with each state in $S$ is

- an invariant function which takes a variable valuation and specifies whether it is possible to remain in the state with that valuation
- a rate of change for each variable. We will use the convention that *clocks*, with variable names based on $C$ (e.g. $C1, C_{period}$) always have a rate of 1, so do not need to be given with the state information. *Integrators.* with names based on $I$ will have a rate of 0 or 1, depending on the state.

Associated with each transition in $T$ is

- an enabling function which takes a variable valuation and specifies whether the transition is allowed with that valuation
- a reset function which specifies which variables are to be reset to 0 when the transition is taken.

### B. From PG to hybrid Automata

In order to translate from a distributed system described using PG to a hybrid automaton, we first of all need to construct the set of states. Each task or message can be in one of three conditions: waiting (to be released), ready, or running. (In the following, read 'task or message' for 'task'). The whole state space consists of all possible combinations of each of these conditions for each of the tasks, plus one (or more) state(s) corresponding to failure. Transitions between states correspond to the triggering or completion of tasks. If a task is released in a state where it is not already runnable, then a transition is made to the state where that task is runnable. The condition of each of the other tasks is not changed, unless the task that has been released has a higher priority than a preemptible running task which shares the resource. In this case, the preemptible task has its condition changed from running to ready. If a task is released in a state where it is already runnable, then this indicates a problem with schedulability, so a transition is made to a fail state. Invariants are added to states to ensure that some transition is made (either triggering or failure) once the period has elapsed.

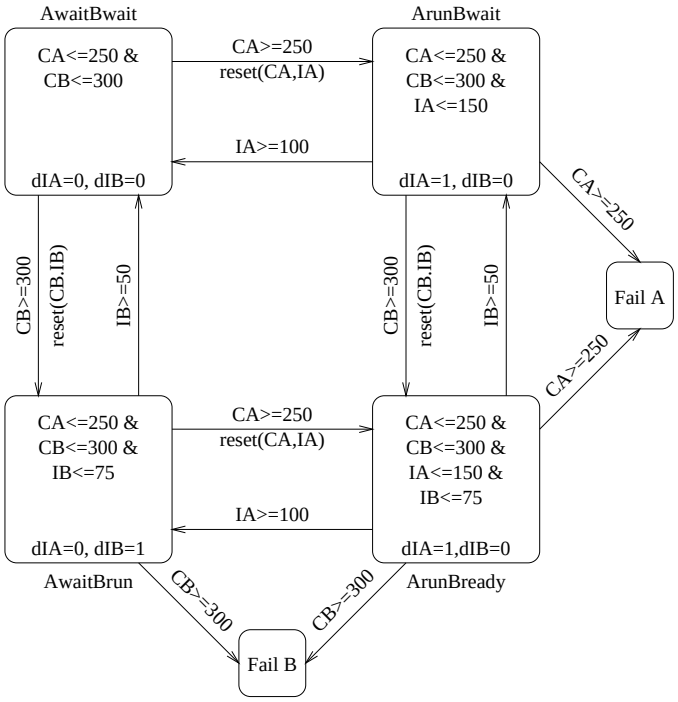


Fig. 3. Automaton for two periodic tasks

A clock variable is required for each periodically triggered task. This clock will be reset at the start of the period, and will be used to evaluate the enabling of transitions corresponding to the release of the task. An integrator variable will also be required for each task, to keep track of how much resource time has been used by that task. When a task is released, the integrator for that task is reset. In a state where a task is running, the corresponding integrator will have rate 1; otherwise the integrator will be 0. Transitions corresponding to task completions are enabled by the integrator for that task being greater than the minimum resource requirement. Invariants are added to ensure that completion transitions are taken at or before the maximum resource requirement. The destination state of a completion transition takes into account the change of condition of the completing task to waiting, and also the release of any tasks triggered by the completion. As before, if a triggered task is already in a runnable condition then the destination state will be a failure state. Figure 3 shows the timed automaton corresponding to the PG definition given below.

```
task A on processor needs [100,150] at priority 0
   triggered by period 250
task B on processor needs [50,75] at priority 1
   triggered by period 300
```

This example does not demonstrate all features of building the graph, as there is is only one resource, there are no non-preemptible tasks, and there are no completion triggers (as opposed to periodic triggers), as any example which covers all of these yields too large an automaton to present easily.

## V. Response Time Analysis

Having built an automaton which models the behaviour of the system we wish to analyse, we now discuss how to perform the analysis upon the model. If we are only interested in schedulability then we need to check whether there is any allowable sequence of events which leads to a failure state. However, we are interested in more than schedulability: as we stated earlier, end-to-end response times are of interest, particularly for distributed systems. We extend our language slightly to define any end-to-end properties we are interested in, using the syntax

> property name from start name on resource to end name on resource

Each property adds one further state and its associated transitions to the automaton. The state is to be used during reachability analysis, and is entered only when the property is violated. A new clock variable is added, and this clock is reset at the start of the first named task. All states which are on the trajectory of this end-to-end response have a transition to the fail state added (the assumption that each task has only one trigger is needed to unambiguously identify whether a state is on a given trajectory). Specific claims about end-to-end response times (e.g. that the response will occur within 1000 units) can be checked, by adding a within clause to the end of the property. In this case, the enabling condition of the fail transitions will be that the property clock has not exceeded the deadline set.

A more sophisticated analysis can be carried out by using the parameterisation mechanism within HyTech [7]. Instead of giving a literal constant, a declared parameter can be used instead. The constraint-solving engine which underlies the model-checker will then calculate the conditions on the parameter(s) under which reachability can occur. In our case, the set whose reachability we are interested in is the failure set, so by fixing the enabling condition on failure transitions to be bounded by a parameter, the model-checker will calculate the least value of the parameter for which failure will occur; in other words, the end-to-end response time. Note that because the property clock is reset at the release of the first named task, the response time recorded will be at most the minimum inter-release time (the period) for this task.

This technique has been tried out on a range of examples, from the simple example given in section II up to a system containing six tasks and four messages. As might be expected, the amount of pessimism identified increases with the length of chains of precedence, and is most marked for lower priority responses. In the largest problem looked at so far, the lower priority response is bounded by 3600 with a standard analysis, but our approach reduces this 1600. It must be noted, however, that these examples are all slightly contrived, in that they have fairly extensive chains of precedence, and that task and message resource bounds are evenly spread. In a system where the majority of a response time is taken up by a single task or message, our technique will not admit such gains.

## VI. Conclusions

We have argued that existing analysis techniques for response time can yield overly pessimistic results for distributed real-time systems with precedence constraints, and have presented a language (PG) for describing such systems. The translation of system definitions written in this language into hybrid automaton models has been discussed, along with the use of parameterised reachability analysis to derive end-to-end response times from the models. This technique has been applied to a range of examples with some positive results.

Related work has been carried out in two main areas:

- The adaptation of standard Response-Time Analysis (RTA) techniques to account for *offsets* which can be used to make the analysis less pessimistic. However, Audsley and Burns have shown that when task periods are co-prime, this technique fails [3].
- Corbett [5] has used timed automaton modelling of Ada tasking programs to carry out scheduling analysis. His work is related to ours, but is based on an explicit model of the scheduler, yielding less tractable (larger) models, and is only applied to uni-processor systems.

Further work is needed in the area of comparing the results we have obtained with those achieved in practice, and also with those obtained by standard and adapted RTA. There is also the possibility of adapting the approach to include more complex scheduling behaviour, such as that given by the priority ceiling protocol or with the use of dynamically assigned priorities (e.g. earliest-deadline first).

## References


[1] R Alur, C Courcebetis, N Halbwachs, T A Henzinger, P-H Ho, X Nicollin, A Olivero, and J Sifakis nad S Yovine. The algorithmic andalysis of hybrid systems. *Theoretical Computer Science*, 138:3–34, 1995.

[2] R Alur and D Dill. Automata for modeling real-time systems. In *17th International Colloquium on Automata, Languages and Programming (ICALP 90)*, number 443 in Lecture Notes in Computer Science. Springer, 1990.

[3] N C Audsley and A Burns. On fixed priority scheduling, offsets and co-prime task periods. *Information Processing Letters*, 67(2):65–70, July 1998.

[4] N C Audsley, A Burns, R I Davis, K W Tindell, and A J Wellings. Fixed priority pre-emptive scheduling: An historical perspective. *Real-Time Systems*, 8(2/3):173–198, March/May 1995.

[5] J C Corbett. Timing analysis of Ada tasking programs. *IEEE Transactions on Software Engineering*, 22(7):461–483, July 1996.

[6] W D Henderson. An holistic approach to performance prediction of distributed real-time can systems. In *5th International CAN Conference, San Jose*, 1998. To Appear.

[7] T A Henzinger, P-H Ho, and H Wong-Toi. A user guide to HYTECH. In E Brinksma, W R Cleaveland, K G Larsen, T Margaria, and B Steffen, editors, *Tool and Algorithms for the Construction and Analysis of Systems: (TACAS 95)*, volume 1019 of *Lecture Notes in Computer Science*, pages 41–71. Springer, 1995.

[8] C L Liu and J W Layland. Scheduling algorithms for multiprogramming in a hard real-time environment. *JACM*, 20(1):40–61, 1973.

[9] K Tindell and J Clark. Holistic schedulability analysis for distributed hard real-time systems. *Microprocessing and Microprogramming*, 40(2-3):117–134, April 1994.