An algorithm for automatically obtaining distributed and fault-tolerant static schedules
Alain Girault, Hamoudi Kalla, Mihaela Sighireanu, Yves Sorel

To cite this version:

Alain Girault, Hamoudi Kalla, Mihaela Sighireanu, Yves Sorel. An algorithm for automatically obtaining distributed and fault-tolerant static schedules. Jun 2003, pp.165-190. hal-00110453

HAL Id: hal-00110453
https://hal.archives-ouvertes.fr/hal-00110453
Submitted on 30 Oct 2006

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

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
An Algorithm for Automatically Obtaining Distributed and Fault-Tolerant Static Schedules

Alain Girault, Hamoudi Kalla
INRIA, 655 av. de l’Europe
3833 Saint-Ismier, Cedex - France
{Alain.Girault,Hamoudi.Kalla}@inrialpes.fr

Mihaela Sighireanu
LIAFA, Case 7014, 2 place Jussieu
75251 Paris, Cedex 05 - FRANCE
sighirea@liafa.jussieu.fr

Yves Sorel
INRIA, B.P.105
78153 Le Chesnay Cedex - FRANCE
Yves.Sorel@inria.fr

Abstract

Our goal is to automatically obtain a distributed and fault-tolerant embedded system: distributed because the system must run on a distributed architecture; fault-tolerant because the system is critical. Our starting point is a source algorithm, a target distributed architecture, some distribution constraints, some indications on the execution times of the algorithm operations on the processors of the target architecture, some indications on the communication times of the data-dependencies on the communication links of the target architecture, a number \(N_{pf}\) of fail-silent processor failures that the obtained system must tolerate, and finally some real-time constraints that the obtained system must satisfy. In this article, we present a scheduling heuristic which, given all these inputs, produces a fault-tolerant, distributed, and static scheduling of the algorithm on the architecture, with an indication whether or not the real-time constraints are satisfied. The algorithm we propose consist of a list scheduling heuristic based active replication strategy, that allows at least \(N_{pf} + 1\) replicas of an operation to be scheduled on different processors, which are run in parallel to tolerate at most \(N_{pf}\) failures. Due to the strategy used to schedule operations, simulation results show that the proposed heuristic improve the performance of our method, both in the absence and in the presence of failures.

Keywords: Fault Tolerance in Distributed and Real-Time Systems, Safety-Critical Systems, software implemented fault-tolerance, multi-component architectures, distribution heuristics.

1. Introduction

Embedded systems account for a major part of critical applications (space, aeronautics, nuclear...), as well as public domain applications (automotive, consumer electronics...). Their main features are:

- **critical real-time**: timing constraints which are not met may involve a system failure leading to a human, ecological, and/or financial disaster;
- **limited resources**: they rely on limited computing power and memory because of weight, encumbrance, energy consumption (e.g., autonomous vehicles), radiation resistance (e.g., nuclear or space), or price constraints (e.g., consumer electronics);
- **distributed and heterogeneous architecture**: they are often distributed to provide enough computing power and to keep sensors and actuators close to the computing sites.

Moreover, the following aspect, extremely important w.r.t. the target fields, must also be taken into account:

- **fault-tolerance**: an embedded system being intrinsically critical [20], it is essential to insure that its software is fault-tolerant; this in itself can even motivate its distribution; in such a case, at the very least, the loss of one computing site must not lead to the loss of the whole application.

The general domain of our research is that of distributed and fault-tolerant embedded systems. The target applications are critical embedded systems. Our ultimate goal is to produce automatically distributed and fault-tolerant code from a given specification of the desired system. In this paper,
we focus on a sub-problem, namely how to produce automatically a distributed and fault-tolerant static schedule of a given algorithm on a given distributed architecture.

Concretely, we are given as input a specification of the algorithm to be distributed (Alg), a specification of the target architecture (Arc), some distribution constraints (Dis), some real-time constraints (Rtc), and a number of processor failures (Npf). The goal is to find a static schedule of Alg on Arc, satisfying Dis, and tolerant to at most Npf processor failures, with an indication whether or not this schedule satisfies Rtc w.r.t. Exe. The global picture is shown in Figure 1. In this paper, we focus on the distribution algorithm.

Finding an algorithm that gives the best fault-tolerant schedule w.r.t. the execution times is a well-known NP-hard problem [10]. Instead, we provide a heuristic that gives an automatically a distributed fault-tolerant static schedule. For instance, Ramamritham requires that the execution cost of each sub-task is the same for each processor, and that the communication cost of each data-dependency is the same for each communication link [19], thereby assuming that the target architecture is homogeneous. Related approaches can be found in [4] (independent tasks and homogeneous architecture) and [18] (heterogeneous architecture but only one failure is tolerated).

Other researchers introduce some dynamicity. For instance, Caccamo and Buttazzo propose an on-line scheduling algorithm to tolerate task failures on a uniprocessor system [5], while Fohler proposes a mixed on-line and off-line scheduling algorithm to tolerate task failures in a multiprocessor system [9].

Finally, some researchers take into account much less restrictive assumptions, but they only achieve hand-made solutions, e.g., with specific communication protocols, voting mechanisms... See the vast literature on general fault-tolerance, for instance [17].

Like the other researchers belonging to the first group, we propose an automatic solution to the fault-tolerance distributed problem. The conjunction of the four following points makes our approach original:

1. We take into account the execution time of both the computation operations and the data communications to optimise the critical path of the obtained schedule.
2. Since we produce a static schedule, we are able to compute the expected completion date for any given operation or data communication, both in the presence and in the absence of failures. Therefore we are able to check the real-time constraints Rtc before the execution. If Rtc is not satisfied, we can give a warning to the designer, so that he can decide whether to add more hardware or to relax Rtc.
3. The given algorithm Alg can be designed with a high-level programming language based on a formal mathematical semantics. This is for instance the case of asynchronous languages, which are moreover well suited to the programming of embedded critical systems [15, 2]. The advantage is that Alg can be formally verified with model-checking and theorem proving tools, and...
therefore we can assume safely that it is free of design faults. The scheduling method we propose in this paper preserves this property.

4. Operations scheduled on the distributed architecture are guaranteed to complete if at most \( N_{pf} \) processors fail at any instant of time. There is no need for a complex failure detection mechanism, and in particular we do not need timeouts to detect the processor failures; there is no need for the processors to propagate the state of the faulty ones; and finally, due to the scheduling strategy used the time needed for handling a failure is minimal.

A different version of the method presented here has been published as an abstract in [12] and as a full version in a workshop [13]. It is different since it addresses distributed architectures consisting of several nodes connected to a single bus, while here we address more general distributed architectures since they can include point-to-point communication links (see Section 3.3). As a result, here the communications can be scheduled in parallel on the communication links, and the fault-tolerance is achieved with the software redundancy of both the computation operations and the data communications (see Section 4.1). In [12, 13] we used the time redundancy of the data communications. Also, we can cope with intermittent processor failures and we do not need to use timeouts to detect failures, which was not the case in [12, 13]. In conclusion, the method presented here is complementary and more general than the one presented in [12, 13].

There is another work involving some of the authors [8], where a totally different approach is taken: First, communication link failures are also taken into account, and second, the method presented involves building a basic schedule for each possible failure, and then merging these basic schedules to obtain a distributed fault-tolerant schedule. The method presented here is lighter, faster, and more efficient, but it only copes with processor failures.

The rest of the paper is organised as follows. Section 3 states our fault-tolerance problem, and presents the various models used by our method. Section 4 presents the proposed solution for providing fault-tolerance. Section 5 provide a correctness proof of the proposed algorithm. Simulation results are presented in Section 6. Finally, Section 7 concludes and proposes directions for future research.

3. Models

3.1. Failure Model

As said in the introduction, our goal is to find a static schedule of \( Alg \) on \( Arc \), satisfying \( Dis \), and tolerant to at most \( N_{pf} \) processor failures, with an indication whether or not this schedule satisfies \( Rtc \) w.r.t. \( Exec \). The failures considered are fail-silent processor failures (permanent as well as intermittent). By "tolerant" we mean that the obtained schedule must achieve “failure masking” [17]. More precisely, this will be done by means of error compensation, using software redundancy. The real-time constraints \( Rtc \) can be, for instance, a deadline for the completion date of the whole schedule. If the user wants to be more precise, he/she can specify a deadline on the completion date of a particular sub-task of the algorithm. The fact that the obtained schedule is static allows the computation of its completion date w.r.t. \( Exec \).

3.2. Algorithm Model

The algorithm is modelled by a data-flow graph. Each vertex is an operation and each edge is a data-dependency. The algorithm is executed repeatedly for each input event from the sensors in order to compute the output events for actuators. We call each execution of the data-flow graph an iteration. This cyclic model exhibits the potential parallelism of the algorithm through the partial order associated to the graph. This model is commonly used for embedded systems and automatic control systems.

Operations of the graph can be either:

- a computation operation (comp): its inputs must precede its outputs; the outputs depend only on the input values; there is no internal state variable and no other side effect;

- a memory operation (mem): the data is held by a mem in sequential order between iterations; the output precedes the input, like a register in Boolean circuits;

- an external input/output operation (extio). Operations with no predecessor in the data flow graph (resp. no successor) are the external input interfaces (resp. output), handling the events produced by the sensors (resp. actuators). The extios are the only operations with side effects; however, we assume that two executions of a given input extio in the same iteration always produce the same output value.

Figure 2 is an example of algorithm graph, with nine operations: I and O are extios (resp. input and output), while A–G are comps. The data-dependencies between operations are depicted by arrows. For instance the data-dependency \( A \rightarrow B \) can correspond to the sending of some arithmetic result computed by A and needed by B.

3.3. Architecture Model

The architecture is modelled by a graph, where each vertex is a processor, and each edge is a communication link.
Classically, a processor is made of one computation unit, one local memory, and one or more communication units, each connected to one communication link. Communication units execute data transfers, called \texttt{comms}. The chosen communication mechanism is the send/receive [14], where the send operation is non-blocking and the receive operation blocks in the absence of data. Figure 2 is an example of architecture graph, with three processors and three point-to-point links.

![Figure 2. Example of (a) an algorithm graph \textit{Alg} (a); and (b) an architecture graph \textit{Are}](image)

3.4. Distribution Constraints, Execution Times, and Real-Time Constraints

For the operations, the execution times \(\text{\textit{exe}}\) consist of a table associating to each pair \((o, p)\) the execution time of the operation \(o\) on the processor \(p\), expressed in time units. Since the target architecture is heterogeneous, the execution times for a given operation can be distinct on each processor. Specifying the distribution constraints \(\text{\textit{dis}}\) involves associating the value “\(\infty\)” to certain pairs \((o, p)\), meaning that \(o\) cannot be executed on \(p\).

For the inter-processor communications, the execution times \(\text{\textit{exe}}\) consist of a table associating to each pair \((\text{data dependency}, \text{communication link})\) the value of the transmission time of this data dependency on this communication link, again expressed in time units.

![Table 1. Distributed constraints \textit{Dis} and execution times \textit{Exe} for operations](image)

Table 2. Execution times \(\text{\textit{exe}}\) for communications

Finally, the real-time constraints \(\text{\textit{RTC}}\) are also given in time units. They can be, for instance, a deadline for the completion date of the whole schedule. For our example, we will take \(\text{\textit{RTC}} = 16\), which means that the obtained static fault-tolerant distributed schedule must complete in less than 16 time units.

4. The Proposed Solution

In this Section we discuss some of the basic principles used in the proposed approach, followed by a description of our algorithm. The algorithm we propose is a list scheduling heuristic based active replication strategy [6], that allows at least \(N_{pf} + 1\) replicas of an operation to be scheduled on different processors, which are run in parallel to tolerate at most \(N_{pf}\) processors failures.

4.1. Algorithm Principle

The proposed solution uses the software redundancy of both \texttt{comps/mems/extios} and of \texttt{comms}. Each operation \(X\) of the algorithm graph is replicated on \(Rep\) different processors of the architecture graph, where \(Rep \geq N_{pf} + 1\). Each of these \(Rep\) replicas send their results in parallel to all the replicas of all the successor operations in the data-flow graph. Therefore, each operation will receive its set of inputs \(Rep\) times; as soon as it receives the first set, the operation is executed and ignores the later inputs. However, in some cases, the replica of an operation will only receive some of its inputs \textit{once}, through an intra-processor communication. For the sake of simplicity, suppose we have an operation \(X\) with only one input produced by its predecessor \(Y\) (see Figure 3(a)).

Consider the replica of \(X\) which is assigned to processor \(P\). Two cases can arise: either one replica of \(Y\) is also scheduled on \(P\), or all the replicas of \(Y\) are assigned to processors distinct from \(P\). In the first case, the \texttt{comms} from \(Y\) to
X will not be replicated and will be implemented as a single intra-processor communication (see Figure 3(b)). Indeed, the replicas of this \( \text{comm} \) would only be used if \( P \) failed, but in this case the replica of \( X \) assigned to \( P \) would not need this input. In the second case, the \( \text{comm} \) from \( Y \) to \( X \) will be replicated \( Npf + 1 \) times, each implemented as an inter-processor communication (see Figure 3(c)).

![Diagram](image)

Figure 3. (a) Algorithm sub-graph; (b) At least one replica of \( Y \) is on \( P \); (c) No replica of \( Y \) is on \( P \).

Since the communication cost between operations assigned to the same processor is considered to be negligible, replicating an operation more than \( Npf + 1 \) times reduces the global interprocessor communication overheads of the schedule. Consider the schedule of Figure 3(c): if \( Y \) is replicated on \( P \), the schedule length can be reduced, both in the presence and in the absence of failures, as shown in Figure 4.

### 4.2. Scheduling Heuristic

The heuristic implementing this solution is a greedy list scheduling [22], called the Fault-Tolerance Based Active Replication strategy (FTBAR) algorithm. We present the scheduling algorithm in macrosteps, the superscript number in parentheses refers to the step of the heuristic, e.g., \( O^{(n)} \)

Before describing the heuristic, we define the following notations which are used in the rest of this paper:

- \( O_{\text{cand}}^{(n)} \): The list of candidate operations, this list is built from the algorithm graph vertices. An operation is said to be a candidate if all its predecessors are already scheduled.
- \( O_{\text{sched}}^{(n)} \): The list of scheduled operations.
- \( \text{pred}(o_i) \): The set of predecessors of operation \( o_i \).
- \( \text{succ}(o_i) \): The set of successors of operation \( o_i \).
- \( R^{(n)} \): The critical path length.
- \( E_{\text{exc}}^{(n)}(o_i, p_j) \): The end execution time of operation \( o_i \) scheduled on processor \( p_j \).
- \( E_{\text{com}}^{(n)}(o_i, o_j) \): The end of data communication time from operation \( o_i \) to operation \( o_j \).
- \( S^{(n)}(o_i) \) is the latest start time from end of \( o_i \).
- \( S_{\text{best}}^{(n)}(o_i, p_i) \): The earliest time at which operation \( o_i \) can start execution on processor \( p_i \). It is computed as follows:
  \[
  S_{\text{best}}^{(n)}(o_i, p_i) = \max_{o_j \in \text{pred}(o_i)} \left\{ \min_{k=1}^{Npf+1} E_{\text{com}}^{(n)}(o_j^k, o_i) \right\}
  \]
  where \( o_j^k \) is the \( k \)-th replica of \( o_j \).
  If \( o_i \) and \( o_j \) are scheduled in the same processor \( p_i \) then
  \[
  E_{\text{com}}^{(n)}(o_j^k, o_i) = E_{\text{exc}}^{(n)}(o_j^k, p_i).
  \]
- \( S_{\text{worst}}^{(n)}(o_i, p_i) \): The earliest time at which operation \( o_i \) can start execution on processor \( p_i \), taking into account all the predecessors replicas. It is computed as follows:
  \[
  S_{\text{worst}}^{(n)}(o_i, p_i) = \max_{o_j \in \text{pred}(o_i)} \left\{ \max_{k=1}^{Npf+1} E_{\text{com}}^{(n)}(o_j^k, o_i) \right\}
  \]
  where \( o_j^k \) is the \( k \)-th replica of \( o_j \).
  If \( o_i \) and \( o_j \) are scheduled in the same processor \( p_i \) then
  \[
  E_{\text{com}}^{(n)}(o_j^k, o_i) = E_{\text{exc}}^{(n)}(o_j^k, p_i).
  \]

The schedule pressure [21] is used as a cost function to select the best operation/processor pair. The schedule pressure noted by \( \sigma^{(n)}(o_i, p_j) \) tries to minimise the length of the critical path of the algorithm and to exploit the scheduling margin of each operation. It is computed for each processor \( p_j \in P \) (\( P \) is the processor’s set) and each operation \( o_i \in O^{(n)}_{\text{cand}} \) by using two functions:
1. The schedule-flexibility $SF$ is defined as:
   \[ SF^{(n)}(o_i, p_j) = R^{(n)} - S_{\text{worst}}^{(n)}(o_i, p_j) - S^{(n)}(o_i) \]

2. The schedule-penalty $SP$ is defined as:
   \[ SP^{(n)}(o_i, p_j) = R^{(n)} - R^{(n-1)} \]

With these two functions, the schedule pressure $\sigma$ is computed as follows:

\[
\sigma^{(n)}(o_i, p_j) = SP^{(n)}(o_i, p_j) - SF^{(n)}(o_i, p_j)
= S_{\text{worst}}^{(n)}(o_i, p_j) + S^{(n)}(o_i) - R^{(n-1)}
\]

The schedule pressure measures how much the scheduling of the operation lengthens the critical path of the algorithm. Therefore it introduces a priority between the operations to be scheduled. Note that, since all candidates operations at step $n$ have the same value $R^{(n-1)}$, it is not necessary to compute $R^{(n-1)}$.

The FTBAR fault-tolerance scheduling heuristic is formally described below:

**The FTBAR Algorithm:**

**begin**

Initialise the lists of candidate and scheduled operations:

- $O_{\text{cand}}^{(0)} := \{ o \in O \mid \text{pred}(o) = \emptyset \}$
- $O_{\text{sched}}^{(0)} := \emptyset$

**while** $O_{\text{cand}}^{(n)} \neq \emptyset$ **do**

1. Compute the schedule pressure for each operation $o_i$ of $O_{\text{cand}}^{(n)}$ on each processor $p_j$ using $S_{\text{worst}}^{(n)}$, and keep the first $Nfp + 1$ min results for each operation:
   \[
   \forall o_i \in O_{\text{cand}}^{(n)},
   \bigcup_{j=1}^{Nfp + 1} \sigma_{\text{worst}}^{(n)}(o_i, p_j) := \min_{p_j \in P} \sigma^{(n)}(o_i, p_j);
   \]

2. Select the best candidate operation $o$ such that:
   \[
   \sigma_{\text{urgent}}(o) := \max_{o_i \in O_{\text{cand}}^{(n)}} \bigcup_{j=1}^{Nfp + 1} \sigma_{\text{best}}^{(n)}(o_i, p_j);
   \]

3. Apply $\text{Minimise_start_time}$ for the best candidate operation $o$ on the first $Nfp + 1$ processors computed at 2;

4. Update the lists of candidate and scheduled operations:
   \[
   O_{\text{cand}}^{(n+1)} := O_{\text{cand}}^{(n)} \cup \{ o \};
   O_{\text{sched}}^{(n+1)} := O_{\text{sched}}^{(n)} \cup \{ o \} \cup \text{Succ}(o);
   \]
   with:
   \[
   \text{Succ}(o) = \{ o' \in \text{succ}(o) \mid \text{pred}(o') \subseteq O_{\text{sched}}^{(n)} \}
   \]

**end while**

**end**

Initially, $O_{\text{sched}}^{(0)}$ is empty and $O_{\text{cand}}^{(n)}$ is the list of operations without any predecessors. At the $n$-th step ($n \geq 1$), the list of already scheduled operations $O_{\text{sched}}^{(n)}$ is kept. Also, the list of candidate operations $O_{\text{cand}}^{(n)}$ is built from the algorithm graph vertices.

At each step $n$, one operation of the list $O_{\text{cand}}^{(n)}$ is selected to be scheduled. To select an operation, we select at the micro-step 3, for each operation $o_i$, the $Nfp + 1$ processors having the minimum schedule pressure. Then among those best pairs $(o_i, p_j)$, we select at the micro-step 2 the one having the maximum schedule pressure, i.e., the most urgent pair.

The selected operation is implemented at the micro-step 3 on the $Nfp + 1$ processors computed at micro-step 2, and the $\text{comms}$ implied by this implementation are also implemented. At this micro-step the start time of the selected operation $o$ is reduced by replicating its predecessors using a procedure $\text{Minimise_start_time}$ proposed by Ahmad and al. in [1], which is formally described below:

**Minimise_start_time(o,p):**

**begin**

1. Determine earliest start time $S_{\text{worst}}^{(n)}(o,p)$;
2. if $S_{\text{worst}}^{(n)}(o,p)$ is undefined then quit because $o$ cannot be scheduled on $p$;
3. Find out the Latest Immediate Predecessor (LIP) of $o$;
4. Minimise the start time of this LIP by recursively calling $\text{Minimise_start_time(LIP);}$
5. Compute the new $S_{\text{worst}}^{(n)}(o,p)$;
6. if (new $S_{\text{worst}}^{(n)}(o,p) \geq S_{\text{worst}}^{(n)}(o,p)$)
7. then
   - Undo all the replications just performed in 1;
   - Schedule $o$ to $p$ at $S_{\text{best}}^{(n)}(o,p)$;
   - The $\text{comms}$ implied by (b) are also implemented here such that, each replica of $o$ receives data from each replica of these predecessors $o_i$ through parallel links;
8. else
   - Find out the new LIP of $o$ and repeat from 1;
**end**

For each pair (predecessor, operation replica), $\text{comms}$ are added in parallel links if and only if all the replicas of the predecessor are on different processors. If this is not the case, i.e., if there exists a replica of the predecessor on the same processor, no $\text{comms}$ is added (see Section 4.1 and Figure 3).

When a $\text{comms}$ is generated, it is assigned to the set of communication units bound to the communication medium connecting the processors executing the source and destination operations. At the end, all the $\text{comms}$ assigned to the same communication unit are statically scheduled. The $\text{comms}$ are thus totally ordered over each communication medium. Provided that the network preserves the integrity and the ordering of messages, this total order of the $\text{comms}$ guarantees that data will be transmitted correctly between processors. The obtained schedule also guarantees a deadlock free execution.

The strategy used to schedule operations ensures a minimum run-time overhead in the faulty system (a system presenting at least one failure) by using $S_{\text{worst}}^{(n)}(o,p)$ to give priority to operations and $S_{\text{best}}^{(n)}(o,p)$ to schedule operations.
4.3. An Example

We have implemented our fault-tolerant heuristic in the SYNDEX [21] tool, which is a tool for optimizing the implementation of real-time embedded applications on multi-component architecture.

We apply our heuristic to the example of Figure 2. The user requires the system to tolerate one permanent processor failure, i.e., $N_{pf} = 1$. The execution characteristics of each comp/mem/extio and comm are specified by the two tables of time units given in Section 3.4.

After the first two steps of our heuristic, we obtain the temporary schedule of Figure 5.

![Figure 5. Step 2](image)

In the next step, operation C is scheduled. Assigning C to P1, P2 and P3 gives an expected schedule pressure of 9.73, 10.53 and 9.23 respectively. But, if A, the LIP of C, is duplicated to P3, the schedule pressure of C can be reduced to 5.73, which means that the start time of C is also reduced. We therefore schedule a new replica of A on P3 and two replicas of C on P3 and P1, which minimizes the schedule pressure. As shown in Figure 6, operation A receives its inputs data twice from the replicas of I scheduled on P1 and P2, and the start time of A is the end of the earliest communication between (I, A) on {L1,3} and {L2,3}. We obtain therefore the temporary schedule of Figure 6.

![Figure 6. Step 3](image)

Similarly, operations B, D, E, F, G, and O are scheduled. At the end of our heuristic, we obtain the final schedule presented in Figure 7. Each operation of the algorithm graph is replicated at least twice and these replicas are assigned to different processors. More important, the real-time constraint is satisfied since the total time is 15.05 < $R_{tc}$.

![Figure 7. Final fault-tolerant schedule](image)

Figure 8 shows the schedule when P1 crashes. As expected, the data sent by all the comms toward the faulty processor P1 are discarded. The schedule corresponding to the subsequent iterations is the same except that all the operations scheduled on P1 as well as all the comms from and to P1 have disappeared. Finally, the real-time constraint is still satisfied since the total time is respectively 15.35, 15.05, 12.6 when P1, P2, or P3 fails at time 0.

![Figure 8. Timed execution when P1 crashes](image)

4.4. Analysis of the Example

To evaluate the overheads introduced by the fault-tolerance, let us consider the non fault-tolerant schedule produced for our example with a basic scheduling heuristic (for instance the one of SYNDEX). The schedule length generated by this heuristic is 10.7.
Neither this non fault-tolerant schedule nor the fault-tolerant schedule of Figure 7 are the best possible. Remember that finding the best schedule is an NP-hard problem; this is the reason why we have designed a heuristic scheduling algorithm. In this particular case, the fault-tolerance overheads are therefore 15.05 – 10.7 = 4.35.

In the fault-tolerant schedule, some communications take place although they are not necessary. On the other hand, the response time of the faulty system is minimized, since results are sent without waiting for any timeout (see Figure 8). For the same reason, the system supports the arrival of several failures during the same iteration since there are no risks that the sum of pending timeouts overtake the desired real-time constraints. In other words, we do not need to make any assumptions on the failure inter-arrival time.

This solution is appropriate to an architecture where the communication means are point-to-point links, which allow parallel communications to take place. For multi-point links, the overheads introduced by the replication of \texttt{comms} may be too high because of their serialization on a single link.

5. Runtime Behavior

In our heuristic, \(N_{pf}\) faults can be tolerated by scheduling \(N_{pf} + 1\) replicas for each operation on different processors. We assume in this paper that all values returned by the \(N_{pf} + 1\) replicas of any given input operation are identical in the same iteration. If no fault occurs, each of the \(N_{pf} + 1\) replicas of an operation receives its inputs in parallel from all the replicas of its predecessor operations in the data-flow graph; as soon as it receives the first set, the operation is executed and ignores the later \(N_{pf}\) inputs. If there are \(k\) permanent faults \((k \leq N_{pf})\), each replica of an operation scheduled on a non-faulty processor receives its inputs in parallel from all the replicas of its predecessors operations scheduled on a non-faulty processor; as soon as it receives the first set, the operation is executed and ignores the later inputs. Concerning the failure detection, there are two options:

1. Either we do not perform any failure detection, in which case, after a failure, the remaining processors will continue to send results to the faulty one. This will not help reducing the communication overheads, especially when the \texttt{comms} have to be serialized over a multi-point communication link. The advantage is that if a processor experiences an intermittent failure, then since it will continue to receive inputs from the healthy processors, it will be able to produce its results again when recovering from its intermittent failure.
2. Or we perform a failure detection procedure by knowing at what time each \texttt{comms} is supposed to happen, and by deciding accordingly that when a \texttt{comms} did not happen, then the sending processor is faulty. Each processor can therefore maintain an array of faulty processors and avoid further \texttt{comms} to the faulty processors in both the remaining of the transient iteration and the subsequent iterations. The drawback is that an intermittent failure cannot be recovered. Indeed, when a processor is detected to be faulty, the other healthy processors will update their array of faulty processors, and will not send any more data during the subsequent iterations. So even if this faulty processor comes back to life, it will not receive any inputs and will not be able to perform any computation. Therefore, in the subsequent iterations, it will fail to send any data on its communication links, and the other healthy processors will never be able to detect that it came back to life.

The same applies to failure detection mistakes.

The choice between these two options can be left to the user. It will depend on the intermittent failure rate of the application as well as on the topology and the bandwidth of the network.

6. Performance Evaluation

To evaluate our fault-tolerant scheduling heuristic, we have compared the performance of the proposed algorithm with the algorithm proposed by Hashimoto and al. in [16], called HBP (Height-Based Partitioning) which is the closest to FTBAR that we have found in the literature. Since, HBP assumes homogeneous systems and only use software redundancy of the algorithm’s operations, FTBAR is downgraded to these assumptions to make the comparison meaningful. The goal of our simulations is to compare the fault-tolerance overheads of HBP and FTBAR, both in the absence and in the presence of one processor failure.

6.1. Simulation Parameters

We have applied FTBAR and HBP heuristics to a set of random algorithm graphs with a wide range of parameters. A random algorithm graph is generated as follows: Given the number of operations \(N\), we randomly generate a set of levels with a random number of operations. Then, operations at a given level are randomly connected to operations at a higher level. The execution times of each operation are randomly selected from a uniform distribution with the mean equal to the chosen average execution time. Similarly, the communication times of each data dependency are randomly selected from a uniform distribution with the mean equal to the chosen average communication time.

For generating the complete set of algorithm graphs, we vary two parameters: \(N = 10, 20, \ldots, 80\), and the
communication-to-computation ratio, defined as the average communication time divided by the average computation time, $CCR = 0.1, 0.5, 1, 2, 5, 10$.

6.2. Performance Results and Analysis

We present in this Section the performance results on the fault-tolerance heuristic. We compute the fault-tolerance overhead in the following way:

$$\text{Overheads} = \frac{(FTSL) - (\text{non FTSL})}{(FTSL)} \times 100$$

where the non FTSL (Fault Tolerant Schedule Length) is produced by FTBAR with $N_{pf} = 0$.

We have plotted in Figures 9 and 10 the average fault-tolerance overheads (averaged over 60 random graphs) as a function of $N$ and $CCR$, both in the absence (Figure 9(a) and 10(a)) and in the presence of one processor failure (Figure 9(b) and 10(b); here we have computed the average overheads when each of the four processors fails, and plotted the max overheads over these four processors).

Figure 9 shows that average overheads increases with $N$. This is due to the active replication of all operations and communications. Figure 9 also shows that FTBAR perform better than HBP.

Figure 10 shows that, when the average communication time is strictly greater than the average execution time, the average overheads decrease. For $CCR \leq 1$, there is little difference between HBP and FTBAR. In contrast, for $CCR \geq 2$, FTBAR performs significantly better than HBP (by at least 20%). This is due to our schedule pressure which tries to minimize the length of the critical path.

7. Conclusion and Future Work

The literature about fault-tolerance of distributed and/or embedded real-time systems is very abundant. Yet, there are few attempts to combine fault-tolerance and automatic generation of distributed code for embedded systems.

In this paper, we have studied this problem and proposed a software implemented fault-tolerance solution.

We have proposed a new scheduling heuristic, called FTBAR (Fault-Tolerance Based Active Replication), that produces automatically a static distributed fault-tolerant schedule of a given algorithm on a given distributed architecture. Our solution is based on the software redundancy of both the computation operations and the communications. All replicated operations send their results but only the one which is received first by the destination processor is used; the other results are discarded. The implementation uses
scheduling heuristic for optimizing the critical path of the distributed algorithm obtained. It is best suited to architectures with point-to-point links. There are some communication overheads, but on the other hand, several failures in a row can be tolerated. Also, depending on the failure detection mechanism chosen, intermittent failures can be tolerated as well.

We have implemented our FTBAR heuristic within the SYNDEX tool. SYNDEX is able to generate automatically executable distributed code, by first producing a static distributed schedule of a given algorithm on a given distributed architecture, and then by generating a real-time distributed executive implementing this schedule. We have also implemented the HBP (Height-Based Partitioning) heuristic. Although HBP only considers homogeneous architectures and only tolerates one processor failure, it is the closest to our work that we have found in the literature. The experimental results shows that FTBAR performs better than HBP, both in the absence and in the presence of failures.

Currently, we are performing extensive benchmark testing of FTBAR on heterogeneous architectures. The first results show that the overheads increases with the number of failures $N/pf$.

Finally, our solution can only tolerate processor failures. We are currently working on new solutions to take communication link failures and reliability into account. We also plan to experiment our method on an electric autonomous vehicle, with a 5-processor distributed architecture.

Acknowledgments

The authors would like to thank Cătălin Dima, Thierry Grandpierre, Claudio Pinello, and David Powell for their helpful suggestions.

References


