# Soft-Error Tolerant Datapath Synthesis Based on Speculative Resource Sharing in Triple Algorithm Redundancy 

Junghoon Oh and Mineo Kaneko<br>School of Information Science<br>Japan Advanced Institute of Science and Technology<br>1-1 Asahidai, Nomi, Ishikawa, Japan 923-1292<br>\{junghoon.oh, mkaneko\}@jaist.ac.jp


#### Abstract

As semiconductor technologies have advanced, the reliability problem caused by soft-errors is becoming one of the serious issues in LSIs. In this paper, we propose a method to synthesize soft-error tolerant application-specific datapaths via high-level synthesis. The novel feature of our method is speculative resource sharing between the retry parts and the secondary parts for hardware/time overhead mitigation. A scheduling algorithm using a special priority function to maximize the speculative resource sharing is also an important feature of this study. We found that our method is more effective when a computation algorithm possesses higher parallelism and a smaller number of resources is available.


## I. Introduction

A soft-error is a transient fault which is triggered by cosmic rays induced neutron and alpha rays from radioactive contaminants in IC packing materials. The effect of soft-error is temporary but it can affect several spatial points simultaneously[1][2]. As the device size decreases, the reliability degradation caused by soft-errors has become one of the greatest issues in LSIs. Approaches to deal with soft-errors are roughly classified into the following three groups: (1) approaches on the device level such as the selection of IC packing materials and the improvement of well structures; (2) approaches on the circuit level such as a flip-flop (FF) with additional circuits for error detection, error recovery and error avoidance[3][4]; (3) approaches on the system level which include concurrent error detection (CED) and triple modular redundancy (TMR) [5]-[7].

Orailoglu and Karri[5] introduced a system which detects and corrects transient faults using CED, checkpointing and rollback. Wu and Karri[6] proposed a high-level synthesis of duplicated computation algorithms for fault detection considering partitioned data dependence of the algorithms. In [7], high-level synthesis for multi-cycle transient fault tolerant datapaths based on TMR was proposed. However, their study considered only a single fault
on a single functional unit. While single-bit upsets are principally a reliability concern in memory devices and systems, multi-bit upsets and multi-cell upsets have recently become a serious problem, as well[1]. The work[8] reports that TMR can be defeated by even 2-bit multi-cell upsets caused by a single soft-error.

In this paper, we propose a method to synthesize single soft-error tolerant application-specific datapaths via high-level synthesis. Triple algorithm redundancy is our starting point for designing soft-error tolerant datapaths. Concerning the tolerability against multiple-component errors caused by a single soft-error, in our design, the main parts and the secondary parts are used for error detection, and the retry parts are used for error recovery. Two most important and novel features of our method are (1) speculative resource sharing between the retry parts and the secondary parts for hardware/time overhead mitigation in compensation for the assumption of sufficiently low probability of soft-errors, and (2) an inherent priority function used in the scheduling algorithm to maximize the speculative resource sharing.

This paper is organized as follows: Section II explains preliminaries of our method. Section III shows requested features for single soft-error tolerance and speculative resource sharing. In Section IV, a scheduling and resource binding algorithm with the specific features shown in Section III is proposed. Section V shows some extensions for multi-cycle soft-error tolerance. Section VI shows experimental results, and Section VII concludes this paper.

## II. Preliminaries

## A. Error Detection and Error Correction Scheme

Our method is based on triple algorithm redundancy. Considering the tolerability against multiple-component errors caused by a single soft-error, and mitigating resource/time overhead, the first copies and the second copies are used for detecting errors, and the third copies are used as retries in our design. For example, the first copy $A^{(1)}$ and the second copy $A^{(2)}$ of a computation block $A$ are executed first, and their output data are compared
by the comparator ' $q$ '. If no error is detected, output data from $A^{(1)}$ is sent to the succeeding blocks. Only if a soft-error affects $A^{(1)}$ or/and $A^{(2)}$ and, as a result, an error is detected, the output data from $A^{(1)}$ and $A^{(2)}$ are abandoned and the third copy $A^{(3)}$ is executed as a retry. The retry data, then, is sent to succeeding computation blocks.

## B. Triple Algorithm Redundancy

We assume that a computation algorithm to be implemented is given by a pair $\left(G, D_{P}\right)$, where $G$ is a dependence graph and $D_{P}$ is the set of primary outputs. Let $O$ be the set of all operations in the original computation algorithm, let $D$ be the set of all variables in the original algorithm and let $E$ be the set of all dependencies between operations and variables. Then a given dependence graph $G$ can be described as $G=(O \cup D, E)$. To synthesize a soft-error tolerant datapath, the given graph $G$ is triplicated (three copies are denoted as $G^{(1)}, G^{(2)}$ and $G^{(3)}$ ), and then comparison-operations and dependencies related to the comparison-operations are inserted to form a resultant algorithm $\left(\tilde{G}, \tilde{D}_{P}\right)$ which is mapped on hardware and time domains by high-level synthesis in the end.

## C. Cone Partitioning/Comparison-Operation Insertion

We define "check variables" as variables to be compared for soft-error detection, and we will choose them selectively, not to all operation results. Dependence subgraphs separated by those check variables are named "cones". Now we let $D_{Q}(\subseteq D)$ be a set of the check variables and let $Q$ be a set of comparison-operations corresponding to the variables in $D_{Q}$.

Definition 1 A "cone" $c_{d}(\subseteq G)$ denotes a subgraph which is induced by tracing back from a check variable $d \in D_{Q}$ to primary inputs or other check variables. Similarly, triplicate copies $c_{d}^{(1)}$ (called "main-cone"), $c_{d}^{(2)}$ (called "second-cone") and $c_{d}^{(3)}$ (called "retry-cone") of $c_{d}$ are induced from the triplicate check variables $d^{(1)} \in G^{(1)}$, $d^{(2)} \in G^{(2)}$ and $d^{(3)} \in G^{(3)}$, respectively.
Definition 2 The set of main-cone $c_{d}^{(1)} \subseteq G^{(1)}$, secondcone $c_{d}^{(2)} \subseteq G^{(2)}$ and retry-cone $c_{d}^{(3)} \subseteq G^{(3)}$ which are induced from a check variable $d$ is called "stage".
Definition 3 The triplicate copies $o^{(1)}\left(\in c_{d}^{(1)}\right), \quad o^{(2)}(\in$ $\left.c_{d}^{(2)}\right)$ and $o^{(3)}\left(\in c_{d}^{(3)}\right)$ of an operation $o\left(\in c_{d}\right)$ are elements of the same stage, and they are called "series operations".

## III. Conditions for Single Soft-Error Tolerant Datapaths

## A. Fault/Error Model and Fault Tolerant Condition

In this study, we assume that a single soft-error can affect several spatial points simultaneously, and hence it

| $\square$ |
| :--- |
| $\square$ |
| Primary inputs |
| $\square$ |
| Check variables |
| Ordinary variables |
| Operations |
| Comparison |
| operations |
| $\square \because$ Main-cones |
| $\nabla \because \because$ Cones |


(b)

Fig. 1. A stage $s_{a}$ consists of $c_{a}^{(1)}, c_{a}^{(2)}$ and $c_{a}^{(3)}$; (a) An example of an original computation algorithm (b) An example of a cone-parti -tioned triplicate algorithm
causes multiple component errors, including errors on registers, functional units and other components at the same time. The proposed soft-error tolerant design is based on cone-level error masking, and relies on error detection by comparing the results of main-cone $c_{d}^{(1)}$ and second-cone $c_{d}^{(2)}$, and error correction by executing the retry-cone $c_{d}^{(3)}$ for each $d \in D_{Q}$.

Retry-cones are executed to recover from erroneous results of the corresponding main-cones or second-cones caused by soft-errors. If the execution of a retry-cone overlaps the executions of the main-cone, the second-cone or the comparison-operation in the same stage, multiple component errors due to a single soft-error may affect two cones or more (retry-cone and either main- or secondcone) in the same stage simultaneously. For this reason, the execution order of operations in each stage should be constrained as follows.

Condition 1 [Execution order in each stage] After a main-cone and the corresponding second-cone are executed, then error detection with a comparison operation is performed to check the results of two cones. After that, the corresponding retry-cone is executed only if the results differ.

In the CED-retry mechanism, only if an error is detected, then corresponding retry-cone is executed and its result is used immediately for the succeeding operations without error detection of the retry result. The validity of this treatment relies on the following assumption.
Assumption 1 The probability of the recurrence of softerrors in a short period is sufficiently low.

On the other hand, triplicate data in three standard registers which have outputs of three cones are not reliable anymore under the assumption of multiple component errors caused by a single soft-error. Therefore,
in order to guarantee the correctness of the inputs to a retry-cone even if its main- and second-cone are affected by a single soft-error, we have decided to use specialized registers, such as BCDMR-ACFF[3], for input variables to each cone (they are primary inputs or check variables) and outputs of comparators.

## B. Speculative Resource Sharing

In order to mitigate hardware resource/time overhead, we propose "speculative resource sharing" as follows. In our treatment, operations in a retry-cone are not executed as long as no error is detected. It means that resources bound to operations in a retry-cone are in idle state if two results of corresponding main-cone and second-cone are identical. If the probability of the recurrence of softerrors in a short period is sufficiently low (Assumption 1), we can rebind the resources in idle state to other operations which have no dependency with the operations in the retry-cone. More specifically, operations in a retrycone can share resources speculatively with operations of second-cones in different stages.

From Assumption 1, while a retry-cone is running (an error caused by a soft-error in the main-cone or/and the second-cone of the running retry-cone is detected), the correctness of other main cones which are executed after the erroneous main- and second-cones of the running retry-cone is guaranteed. Based on this observation, resources can be managed more efficiently by speculative resource sharing between error detection parts and error correction parts.

In Fig. 2, if operations in $c_{m}^{(3)}$ and $c_{n}^{(2)}$ share resources by speculative resource sharing ( $m \neq n$ and there is no dependency between $c_{m}$ and $\left.c_{n}\right), c_{n}^{(1)}$ is unable to detect error when a soft-error affects some of $c_{m}^{(1)}, c_{m}^{(2)}$ and $q_{m}$, and $c_{m}^{(3)}$ is executed, since the execution of $c_{n}^{(2)}$ is abandoned by $c_{m}^{(3)}$. Hence, we need to carefully manage the execution of $c_{n}^{(1)}$ so that the soft-error which affects $c_{m}^{(1)}$, $c_{m}^{(2)}$ and/or $q_{m}$ may not affect $c_{n}^{(1)}$. Considering such behavior, we introduce the second condition.
Condition 2 [Speculative resource sharing] Only if the execution of a main-cone $c_{n}^{(1)}$ is scheduled later than the execution of a comparison operation $q_{m}$, speculative resource sharing between operations in $c_{m}^{(3)}$ and operations in a second-cone $c_{n}^{(2)}$ corresponding to $c_{n}^{(1)}$ is possible. ( $m \neq n$ )

## IV. Synthesis for Soft-Error Tolerant Datapath

In this section, we propose a scheduling and resource binding algorithm for single soft-error tolerant datapaths. The way how to select check variables for comparison and cone partitioning is an important factor for datapath optimization. However, in this paper, the set of check


Fig. 2. Scheduled DFG; An example of speculative resource sharing between operations in second-cones and retry-cones. $o_{i}^{(3)}$ and $o_{j}^{(2)}$ can share a resource speculatively. However, $o_{k}^{(3)}$ and $o_{i}^{(2)}$ cannot.
variables is assumed to be given as a part of an input description for high-level synthesis. Optimization of the selection of check variables is left as one of the important future works.

Our algorithm is based on the list scheduling but has an inherent priority function and an inherent resource counting considering speculative resource sharing.

In the following, we let $\sigma: O^{(1)} \cup O^{(2)} \cup O^{(3)} \cup Q \rightarrow \mathbb{N}$ be an operation schedule (a control step assignment), where $O^{(\ell)}, \ell=1,2,3$, is the set of operations in $G^{(\ell)}$.

## A. Schedule Constraints

1. Resource constraints for every resource type have to be satisfied at each control step cs. Let $A_{c s, r}$ be a set of operations which are scheduled at a control step $c s$ with a resource type $r$. The actual number of resources of type $r$ in use at step $c s$ is described as follows: The speculative resource sharing conditions are described in the next section B.
Number of resources of type $r$ in use at step $c s$
$=\left|A_{c s, r}\right|-$ "Number of operations which share
resources speculatively in $A_{c s, r} " / 2$
$\leq$ Number of allocated resources of type $r$
2. $c_{m}^{(3)}$, the retry-cone in a stage $m$, has to be executed after the comparison operation $q_{m}$ that compares the results of $c_{m}^{(1)}$ and $c_{m}^{(2)}$ under Condition 1.

$$
\begin{gather*}
\max \left(\bigcup_{\ell=1}^{2} \sigma\left(c_{m}^{(\ell)}\right)\right)<\sigma\left(q_{m}\right)<\min \left(\sigma\left(c_{m}^{(3)}\right)\right)  \tag{2}\\
\text { where } \sigma\left(c_{m}^{(\ell)}\right)=\left\{x \mid \sigma\left(o_{i}^{(\ell)}\right)=x, o_{i}^{(\ell)} \in c_{m}^{(\ell)}\right\}
\end{gather*}
$$

3. If $o_{i}$ is a predecessor of $o_{j}$,

- For operations $o_{i}^{(\ell)}, o_{j}^{(\ell)} \in c_{m}^{(\ell)}$, the successor $o_{j}^{(\ell)}$ has to be scheduled after the execution of $o_{i}^{(\ell)}$. $(\ell=1,2,3)$

$$
\begin{equation*}
\sigma\left(o_{i}^{(\ell)}\right)<\sigma\left(o_{j}^{(\ell)}\right) \tag{3}
\end{equation*}
$$

- For operations $o_{i}^{(\ell)} \in c_{m}^{(\ell)}, o_{j}^{(\ell)} \in c_{n}^{(\ell)}(m \neq n$ and $\ell=1,2,3)$, the successor $o_{j}^{(\ell)}$ has to be scheduled after the execution of $o_{i}^{(3)}$ whose output is the result of $c_{m}^{(3)}\left(\sigma\left(o_{i}^{(3)}\right)=\max \left(\sigma\left(c_{m}^{(3)}\right)\right)\right)$.

$$
\begin{equation*}
\sigma\left(o_{i}^{(3)}\right)<\sigma\left(o_{j}^{(\ell)}\right) \tag{4}
\end{equation*}
$$

B. Speculative Resource Sharing Conditions between Operations in Second-Cones and in Retry-Cones

When there is no dependency between $o_{i}$ and $o_{j}(i \neq j)$, $o_{i}^{(3)}$ and $o_{j}^{(2)}$ can be bound to the same resource under the following conditions (Fig. 2).

1. $o_{i}^{(3)}$ and $o_{j}^{(2)}$ belong to different stages.

$$
\begin{equation*}
o_{i}^{(3)} \in c_{m}^{(3)}, o_{j}^{(2)} \in c_{n}^{(2)} \quad(m \neq n) \tag{5}
\end{equation*}
$$

2. $o_{i}^{(3)}$ and $o_{j}^{(2)}$ are scheduled at the same control step.

$$
\begin{equation*}
\sigma\left(o_{i}^{(3)}\right)=\sigma\left(o_{j}^{(2)}\right) \tag{6}
\end{equation*}
$$

3. According to Condition 2,

$$
\begin{equation*}
\sigma\left(q_{m}\right)<\min \left(\sigma\left(c_{n}^{(1)}\right)\right) \tag{7}
\end{equation*}
$$

When the above conditions $(5) \sim(7)$ are satisfied, $o_{i}^{(3)}$ and $o_{j}^{(2)}$ can share a same resource speculatively.

## C. Proposed Scheduling Algorithm

Algorithm 1 shows the proposed algorithm which is to perform a list scheduling under given resource constraints. The following notations are employed to explain the proposed scheduling algorithm.

- $L_{c s, r}$ : a list of ready operations of resource type $r$ at control step $c s$
- $U_{c s, r}$ : a list of executing operations of resource type $r$ at control step $c s$
- $v_{r, k}$ : the $k$-th operation of resource type $r$
- res_inuse : the number of functional units in use

Scheduling proceeds from control step 1 toward the last control step as the original list scheduling does. In each control step $c s$, first, operations of type $r$ which are scheduled already and are executing at control step $c s$, are registered to $U_{c s, r}$. Ready operations which are not yet scheduled are registered to $L_{c s, r}$. After that, the function count_occupied_FU is called, which counts the number of presently occupied functional units (res_inuse) from $U_{c s, r}$, and then the function sched_with_speculative_share is called, which binds operations to control step cs considering the speculative resource sharing conditions (Section IV-B).

## D. Scheduling Priority

First of all, we introduce priority and latency of a stage. Priority of a stage is defined as the smallest control step of ALAP schedule over all operations in the stage. Latency

```
Algorithm 1 Modified list scheduling algorithm
Require: \(N_{\max } \leftarrow \#\) of resource types
    \(c s \leftarrow 0\);
        while (until all operations are scheduled) do
            \(c s \leftarrow c s+1 ;\)
            for \(\left(r \leftarrow 1 ; r \leq N_{\max } ; r++\right.\) ) do
                res_inuse \(\leftarrow 0\);
            for ( \(k \leftarrow 1 ; k \leq \#\) of operations (type \(r\) ) ; \(k++\) ) do
                if ( \(v_{r, k}\) is already scheduled) then
                if (execution of \(v_{r, k}\) is not finish yet) then
                        Register \(v_{r, k}\) in \(U_{c s, r}\);
                end if
                else if (executions of all immediate predecessors of \(v_{r, k}\)
                    are already finished) then
                        Register \(v_{r, k}\) in \(L_{c s, r}\);
                end if
            end for
            res_inuse \(\leftarrow\) count_occupied_FU \(\left(U_{c s, r}\right)\);
            sched_with_speculative_share \(\left(c s, r\right.\), res_inuse, \(\left.L_{c s, r}\right)\);
        end for
    end while
```

of a stage is defined as execution time from the start to the end of the stage based on ALAP schedule.

$$
\begin{align*}
& \text { Priority of } s_{m}=\min \left(\bigcup_{\ell=1}^{2} \sigma_{A L A P}\left(c_{m}^{(\ell)}\right)\right)  \tag{8}\\
& \text { Latency of } s_{m} \\
& =\max \left(\sigma_{A L A P}\left(c_{m}^{(3)}\right)\right)-\min \left(\bigcup_{\ell=1}^{2} \sigma_{A L A P}\left(c_{m}^{(\ell)}\right)\right)+1 \tag{9}
\end{align*}
$$

The priority of an operation is determined by applying the following factors in this order (a tie in the first factor is broken by the second factor, a tie in the second factor is broken by the third factor, and so on).
(1) Operations which are in a stage having higher priority have higher priority. As a result, a stage which started earlier can finish earlier. If a stage which started earlier remains without being chosen and only operations in other stages are scheduled, operations in the remained stage can be a bottleneck on the entire schedule.
(2) Operations which are in a stage having smaller latency have higher priority, since a stage which is expected to finish earlier can include speculatively resource sharable operation pairs easier.
(3) Operations which have smaller ALAP schedules are given higher priority.
(4) Operations which are placed on a critical path are given higher priority.
(5) Among series operations, operations in second-cones are given higher priority for speculative sharing.
Using these factors, priority is given to every operation.

## E. Selecting Operations

Even if the priority of ready operations is fixed, selecting operations to be scheduled at the current control step is not straightforward, since we need to manage a complicated resource sharing between second-cones and retrycones. In order to choose speculatively resource sharable
pairs aggressively, we will use a bipartite graph $H$ with $X_{T}$ and $Y_{T}$ as its partite sets.

We define "vertex weight" $w(z) \in \mathbb{Z}_{+}$as priority of an operation $z$ and "edge weight" $w(e=\{x, y\}) \in \mathbb{Z}_{+}$as the sum of weights of two end vertices $x$ and $y$, that is, $w(e=\{x, y\})=w(x)+w(y)$.

- A bipartite graph $H=\left(X_{T} \cup Y_{T}, W\right)$ describes the following:
- A partite set $X_{T}$ is defined as $X_{T}=X \cup Y_{C}$, where $X=\left\{x \mid x \in L_{c s, r} \wedge x \in G^{(1)} \cup G^{(2)}\right\}$ and $Y_{C}$ is a set of auxiliary vertices with $\left|Y_{C}\right|=|Y|$.
- The other partite set $Y_{T}$ is defined as $Y_{T}=Y \cup X_{C}$, where $Y=\left\{y \mid x \in L_{c s, r} \wedge y \in G^{(3)}\right\}$ and $X_{C}$ is a set of auxiliary vertices with $\left|X_{C}\right|=|X|$.
- The edge set $W$ of $H$ is defined as follows.
$W=\left\{\left\{x, P_{X}(x)\right\} \mid x \in X\right\} \cup\left\{\left\{P_{X}(y), y\right\} \mid y \in Y\right\}$ $\cup\left\{\{x, y\} \mid x \in X \cap G^{(2)}\right.$ and $y \in Y$ can share a resource speculatively $\}$, where $P_{X}: X \rightarrow X_{C}$ and $P_{Y}: Y \rightarrow Y_{C}$ are arbitrary one-to-one mapping from $X$ to $X_{C}$ and from $Y$ to $Y_{C}$, respectively.
- We consider that weights of all auxiliary vertices are 0 , that is, $\forall z, w(z)=0$, where $z \in X_{C} \vee z \in Y_{C}$.

Once $H$ is constructed, selecting an edge $e=\{x, y\}$ from $H$ means that two operations which are represented by the end vertices $x$ and $y$ of $e$ share a resource speculatively. If an end vertex of $e$ is an auxiliary vertex (a vertex in either $X_{C}$ or $Y_{C}$ ), it means that the operation which is represented by the other end vertex occupies a resource without speculative sharing. As a result, an operation selection considering speculative resource sharing can be considered as a problem to find a size-constrained maximum weight matching from $H$.

Algorithm 2 shows a detailed description of function schedule_with_speculative_share. After a graph $H$ is constructed, speculatively resource sharable operation pairs are chosen by a greedy selection procedure. More specifically, edges are selected in descending order of the edge weights within the number of available resources. Fig. 3 illustrates an example of operations selecting result by Algorithm 2.

## V. Multi-Cycle Soft-Error Tolerant Datapath

We can expand our schedule constraints and speculative resource sharing conditions so that they can manage multi-cycle soft-errors[7].

## A. Schedule Constraints

When we consider $k$-cycle soft-error tolerant datapath synthesis, the difference in schedule constraints is Equation (2) in Section IV-A.

- $c_{m}^{(3)}$, the retry-cone in a stage $m$, has to be executed $k$ steps or more after the completion of the comparison


Fig. 3. An example of operations selecting result produced by greedy selection procedure; Operations which have edge weight 10 and 15 are elements of $G^{(2)}$. Other operations in $X$ are elements of $G^{(1)}$. The number of available resources is 3 .

```
Algorithm 2
Algorithm of function schedule_with_speculative_share()
Require: cs \(\leftarrow\) Current control step
Require: \(r \leftarrow\) Currently considered resource type
Require: \(n_{r} \leftarrow\) Available number of resources of type \(r\)
Require: \(L_{c s, r} \leftarrow\) List of ready operations of type \(r\) at step \(c s\)
    Construct a bipartite graph \(H\) based on \(L_{c s, r}\);
    for \((i \leftarrow 0 ; i<\) Total number of \(W ; i++\) ) do
        if \(\left(n_{r}>0\right)\) then
            Select an edge \(e\) from \(W\) which has the largest edge weight;
            Schedule two end vertices of \(e\) at the current step cs;
            Remove the edge \(e\) and its end vertices from \(H\)
            \(n_{r} \leftarrow n_{r}-1 ;\)
        else
            break;
        end if
    end for
```

operation $q_{m}$ which compares the results of $c_{m}^{(1)}$ and $c_{m}^{(2)}$.

$$
\begin{array}{r}
\max \left(\bigcup_{\ell=1}^{2} \sigma\left(c_{m}^{(\ell)}\right)\right)+k-1<\sigma\left(q_{m}\right) \\
\sigma\left(q_{m}\right)+k-1<\min \left(\sigma\left(c_{m}^{(3)}\right)\right) \tag{11}
\end{array}
$$

B. Speculative Resource Sharing Conditions between Operations in Second-Cones and in Retry-Cones
These conditions, excluding Equation (7), are the same as the conditions in Section IV-B,

$$
\begin{equation*}
\sigma\left(q_{m}\right)+k-1<\min \left(\sigma\left(c_{n}^{(1)}\right)\right) \tag{12}
\end{equation*}
$$

## VI. Experimental Results

We have implemented the proposed scheduling algorithm as a computer program, and applied it to 16 -point fast Fourier transform (16FFT), 8-point inverse discrete cosine transform (8IDCT), 16-point FIR filter (16FIR), autoregressive filter (ARF), fifth-order elliptic wave digital filter (5EWDF), and inverse discrete cosine transform


Fig. 4. Experimental results by various computational algorithms. Each column represents allocated resources (C:Comparator, A:ALU, M:Multiplier). Every column in each graph has two grouping bars. Left bars represent scheduling results without speculative resource sharing and right bars represent proposed scheduling results with speculative resource sharing.
with column-wise decomposition (IDCT-c). In order to evaluate our scheduling algorithm, we also implemented a conventional list scheduling algorithm without speculative resource sharing.
Fig. 4 and Table I show experimental results. Each graph in Fig. 4 illustrates the achieved latencies with different specifications of resources. Table I shows the maximum improvement rates in latency between conventional list scheduling results without speculative resource sharing and proposed list scheduling results with speculative resource sharing. We can find from those results that, for every application algorithm, the latency is improved more or less by the speculative resource sharing. Also we found that the proposed method is more effective when a computation algorithm possesses higher parallelism (16FFT, IDCT-c and 8IDCT) and the improvement in latency is larger when a smaller number of resources is allocated. The reason is that there are more possibilities to share resources speculatively between retry-cones and secondcones and, in consequence, many chances for speculative resource sharing can conceal the extension of latency.

## VII. Conclusion

Concerning soft-error tolerant datapath synthesis based on triplication of an input computation algorithm via

TABLE I
Experimental results; latency improvement rate

| Computation <br> algorithm | Total \# of operations <br> (triplicate algorithm) | Max. improvement <br> rate [\%] |  |
| :---: | :---: | :---: | :---: |
| 16FFT | 328 | 32.3 |  |
| IDCT-c | 234 |  | 31.0 |
| 8DCT | 160 | 29.5 |  |
| ARF | 96 | 22 |  |
| 16FIRF | 80 | 20.7 |  |
| 5EWDF | 120 | 17.5 |  |

high-level synthesis, constraints for soft-error tolerance and a scheduling algorithm considering speculative resource sharing are proposed. Datapath circuits designed by our method tolerate multi-component and multi-cycle errors caused by single soft-errors. From the results of soft-error tolerant datapath synthesis experiments, we found that our speculative resource sharing achieves maximum $32.3 \%$ improvement in latency. Especially, our proposed method is more effective when an input computation algorithm possesses higher parallelism, and the number of allocated resource is relatively small. Optimization of cone partitioning and register binding are left as future works. Moreover, we needed to investigate that our method is applicable to more practical problems such as more large applications and CDFG with loop structures.

## References

[1] E. Ibe, H. Taniguchi, Y. Yahagi, K. Shimbo, and T. Toba, "Impact of scaling on neutron-induced soft error in srams from a 250 nm to a 22 nm design rule," Electron Devices, IEEE Transactions on, vol. 57, no. 7, pp. 1527-1538, July 2010.
[2] H. Fuketa, R. Harada, M. Hashimoto, and T. Onoye, "Measurement and analysis of alpha-particle-induced soft errors and multiple-cell upsets in 10 t subthreshold sram," Device and $M a$ terials Reliability, IEEE Transactions on, vol. 14, no. 1, pp. 463-470, March 2014.
[3] M. Masuda, K. Kubota, R. Yamamoto, J. Furuta, K. Kobayashi, and H. Onodera, "A 65 nm low-power adaptive-coupling redundant flip-flop," Nuclear Science, IEEE Transactions on, vol. 60, no. 4, pp. 2750-2755, Aug 2013.
[4] Y. Lin and M. Zwolinski, "Settoff: A fault tolerant flip-flop for building cost-efficient reliable systems," in On-Line Testing Symposium (IOLTS), 2012 IEEE 18th International, June 2012, pp. 7-12.
[5] A. Orailoglu and R. Karri, "Automatic synthesis of selfrecovering vlsi systems," Computers, IEEE Transactions on, vol. 45, no. 2, pp. 131-142, Feb 1996.
[6] K. Wu and R. Karri, "Fault secure datapath synthesis using hybrid time and hardware redundancy," Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on, vol. 23, no. 10, pp. 1476-1485, Oct 2004.
[7] T. Inoue, H. Henmi, Y. Yoshikawa, and H. Ichihara, "High-level synthesis for multi-cycle transient fault tolerant datapaths," in On-Line Testing Symposium (IOLTS), 2011 IEEE 17th International, July 2011, pp. 13-18.
[8] H. Quinn, K. Morgan, P. Graham, J. Krone, M. Caffrey, and K. Lundgreen, "Domain crossing errors: Limitations on single device triple-modular redundancy circuits in xilinx fpgas," Nu clear Science, IEEE Transactions on, vol. 54, no. 6, pp. 20372043, Dec 2007.

