Hyperreconfigurable Architectures: Reconfiguration Strategies and Costs

Sebastian Lange, Martin Middendorf

University of Leipzig
Germany

DFG Meeting 2008
1. Motivation

2. Partially Reconfigurable FPGAs

3. Stochastic Analysis of Partial Reconfiguration

4. 2-Level Reconfigurable Architectures

5. Heterogeneous Frame Sizes
**Motivation**

**Problem**
Extensive reconfiguration potential causes a great amount of reconfiguration data for applications with frequent use of dynamic reconfiguration.

**Observation**
Computations show phases with differing resource utilization ⇒ Time-variant demands for reconfiguration

**Basic Idea**
Dynamically adapt reconfiguration capabilities so that reconfiguration data contain ideally only information about resources that are needed

"Reconfiguration of Reconfigurability"
Motivation

Problem
Extensive reconfiguration potential causes a great amount of reconfiguration data for applications with frequent use of dynamic reconfiguration.

Observation
Computations show phases with differing resource utilization ⇒ Time-variant demands for reconfiguration

Basic Idea
Dynamically adapt reconfiguration capabilities so that reconfiguration data contain ideally only information about resources that are needed

"Reconfiguration of Reconfigurability"
Motivation

Problem
Extensive reconfiguration potential causes a great amount of reconfiguration data for applications with frequent use of dynamic reconfiguration.

Observation
Computations show phases with differing resource utilization ⇒ Time-variant demands for reconfiguration

Basic Idea
Dynamically adapt reconfiguration capabilities so that reconfiguration data contain ideally only information about resources that are needed

"Reconfiguration of Reconfigurability"
Partially Reconfigurable FPGAs

- FPGA consists of Logic Blocks, I/O cells and Interconnect
- Reconfiguration organized in frames
  - Contain configuration data for multiple, diverse components
  - Data size of frame uniform
  - Individually addressable

- Multiple frames with adjacent addresses reconfigured en bloc
- Overhead due to addresses and frame granularity

Frame Based Partial Reconfiguration

S. Lange, M. Middendorf (Uni Leipzig)
Description of FPGA

- Consists of set of $n$ reconfigurable units $X = \{x_1, \ldots, x_n\}$
- Units partitioned into ordered set of frames $F = \{f_1, \ldots, f_k\}$, $X = \bigcup_{i=1,\ldots,k} f_i$
- Frames do not overlap $f_i \cap f_j = \emptyset$ for $i \neq j$

Characterization of Algorithms

- Algorithm defines sequence of resource demands $c_1 \ldots c_n$
- Resource demand $c_i \subseteq X$ describes reconfigured units at $i$th reconfiguration step
- Partial reconfiguration of frame $f$ if some unit $x \in f$ required by $c_i$
Description of FPGA

- Consists of set of $n$ reconfigurable units $X = \{x_1, \ldots, x_n\}$
- Units partitioned into ordered set of frames $F = \{f_1, \ldots, f_k\}$, $X = \bigcup_{i=1,\ldots,k} f_i$
- Frames do not overlap $f_i \cap f_j = \emptyset$ for $i \neq j$

Characterization of Algorithms

- Algorithm defines sequence of resource demands $c_1 \ldots c_n$
- Resource demand $c_i \subseteq X$ describes reconfigured units at $i$th reconfiguration step
- Partial reconfiguration of frame $f$ if some unit $x \in f$ required by $c_i$
Reconfiguration Cost Model - Partial Reconfiguration

### Amount of State Information
- State of entire frame reconfigured if one unit in frame reconfigured
- \( \text{scost}(c) = \sum_{f \in F} \begin{cases} |f| & c \cap f \neq \emptyset \\ 0 & \text{otherwise} \end{cases} \)

### Addressing Costs
- For each block address of first frame, \#frames sent: \( \lceil \log_2 |F| \rceil \)
- No cost for padding frames
- \( \text{ocost}(c) = \sum_{i=1}^{\lfloor \log_2 |F| \rfloor} \begin{cases} 2 \cdot \lceil \log_2 |F| \rceil & c \cap f_i \neq \emptyset \land (c \cap f_{i-1} = \emptyset \lor i = 1) \\ 0 & \text{otherwise} \end{cases} \)

### Total Costs
- \( \text{cost} = \sum_{c \in S} (\text{scost}(c) + \text{ocost}(c)) \)
Reconfiguration Cost Model - Partial Reconfiguration

**Amount of State Information**

- State of entire frame reconfigured if one unit in frame reconfigured

\[ scost(c) = \sum_{f \in F} \begin{cases} |f| & c \cap f \neq \emptyset \\ 0 & \text{otherwise} \end{cases} \]

**Addressing Costs**

- For each block address of first frame, #frames sent: \[\lceil \log_2 |F| \rceil\]
- No cost for padding frames

\[ ocost(c) = \sum_{i=1}^{|F|} \begin{cases} 2\lceil \log_2 |F| \rceil & c \cap f_i \neq \emptyset \land (c \cap f_{i-1} = \emptyset \lor i = 1) \\ 0 & \text{otherwise} \end{cases} \]

**Total Costs**

\[ cost = \sum_{c \in S} (scost(c) + ocost(c)) \]
### Amount of State Information

- State of entire frame reconfigured if one unit in frame reconfigured

\[
scost(c) = \sum_{f \in F} \begin{cases} 
|f| & c \cap f \neq \emptyset \\
0 & \text{otherwise}
\end{cases}
\]

### Addressing Costs

- For each block address of first frame, \#frames sent: \(\lceil \log_2 |F| \rceil\)
- No cost for padding frames

\[
ocost(c) = \sum_{i=1}^{\lceil \log_2 |F| \rceil} \begin{cases} 
2^{\lceil \log_2 |F| \rceil} & c \cap f_i \neq \emptyset \land (c \cap f_{i-1} = \emptyset \lor i = 1) \\
0 & \text{otherwise}
\end{cases}
\]

### Total Costs

\[
\text{cost} = \sum_{c \in S} (scost(c) + ocost(c))
\]
Stochastic Analysis of Partial Reconfiguration

Assumption

- System contains \( n \) reconfigurable units partitioned into \( k \) frames
- Each unit used with **same** probability \( p \)

Resulting Cost Model

- Probability of frame reconfiguration \( p_{on} = 1 - (1 - p)^{n/k} \)
- Expected amount of state information \( scost = (1 - (1 - p)^{n/k}) \cdot n \)
- Probability of frame beginning a block \( p_{pr} = (1 - (1 - p)^{n/k}) \cdot (1 - p)^{n/k} \)
- Expected amount of addressing cost \( ocost = 2[\log_2 k] \cdot p_{on}(1 + (k - 1) \cdot (1 - p)^{n/k}) \)
- Total reconfiguration cost \( cost = p_{on}(n + 2[\log_2 k](1 + (k - 1) \cdot (1 - p)^{n/k})) \)
## Assumption

- System contains $n$ reconfigurable units partitioned into $k$ frames
- Each unit used with **same** probability $p$

## Resulting Cost Model

- Probability of frame reconfiguration $p_{on} = 1 - (1 - p)^{n/k}$
- Expected amount of state information $s_{cost} = (1 - (1 - p)^{n/k}) \cdot n$
- Probability of frame beginning a block $p_{pr} = (1 - (1 - p)^{n/k}) \cdot (1 - p)^{n/k}$
- Expected amount of addressing cost $o_{cost} = 2 \lceil \log_2 k \rceil \cdot p_{on} (1 + (k - 1) \cdot (1 - p)^{n/k})$
- Total reconfiguration cost $c_{ost} = p_{on} (n + 2 \lceil \log_2 k \rceil (1 + (k - 1) \cdot (1 - p)^{n/k}))$
### Assumption
- System contains \( n \) reconfigurable units partitioned into \( k \) frames
- Each unit used with **same** probability \( p \)

### Resulting Cost Model
- Probability of frame reconfiguration \( p_{on} = 1 - (1 - p)^{n/k} \)
- Expected amount of state information \( scost = (1 - (1 - p)^{n/k}) \cdot n \)
- Probability of frame beginning a block \( p_{pr} = (1 - (1 - p)^{n/k}) \cdot (1 - p)^{n/k} \)
- Expected amount of addressing cost \( ocost = 2 \lceil \log_2 k \rceil \cdot p_{on}(1 + (k - 1) \cdot (1 - p)^{n/k}) \)
- Total reconfiguration cost \( cost = p_{on}(n + 2 \lceil \log_2 k \rceil(1 + (k - 1) \cdot (1 - p)^{n/k})) \)
Analysis for System with 10,000 Reconfigurable Units

- Expected cost of one reconfiguration
- Varied number of frames ($k$)
- Varied probability for reconfiguration ($p$)
- Addressing cost dominate amount of state information
- Discontinuities for special values of $k$
Observations:

- Partial reconfiguration useful if \( \lceil \log_2 k \rceil > n \) or if \( p < 1 - \left( 1 - \frac{n}{2k\lceil \log_2 k \rceil} \right)^{n/k} \) for \( \lceil \log_2 k \rceil > n \)
- Deducing formula for optimal number of frames hard
- Has been determined numerically
2-Level Reconfigurable Architectures

- Use 2 levels of reconfiguration
  - Upper level hypercontexts controls reconfigurability of frames
  - Frames not available in hypercontext bypassed during reconfiguration
- Upper level reconfiguration less frequently then lower level reconfiguration

Frame Based Partial Reconfiguration

- Components are controlled by SRAM cells
- Hypercontext SRAM cells
- Context SRAM cells
- Hyper/Reconfiguration Chain

Diagram showing 2-level reconfiguration with groups $g_1$ to $g_8$.
Reconfiguration Cost Model - 2 Level Reconfiguration

Amount of State Information

- Hypercontext dictates availability of frames

\[ scost(c) = \sum_{f \in F} \begin{cases} |f| & h \cap f \neq \emptyset \\ 0 & \text{otherwise} \end{cases} \]

Hyperreconfiguration Costs

- Cost for indirect addressing of frames
- Describes availability of each frame

\[ ocost(c_i) = \begin{cases} k & i = 1 \lor h(c_i) \neq h(c_{i-1}) \\ 0 & \text{otherwise} \end{cases} \]

Total Costs

\[ cost = \sum_{c \in S} (scost(c) + ocost(c)) \]
Reconfiguration Cost Model - 2 Level Reconfiguration

Amount of State Information

- Hypercontext dictates availability of frames

\[ scost(c) = \sum_{f \in F} \begin{cases} |f| & h \cap f \neq \emptyset \\ 0 & \text{otherwise} \end{cases} \]

Hyperreconfiguration Costs

- Cost for indirect addressing of frames
- Describes availability of each frame

\[ ocost(c_i) = \begin{cases} k & i = 1 \lor h(c_i) \neq h(c_{i-1}) \\ 0 & \text{otherwise} \end{cases} \]

Total Costs

\[ cost = \sum_{c \in S} (scost(c) + ocost(c)) \]
Amount of State Information

- Hypercontext dictates availability of frames

\[ scost(c) = \sum_{f \in F} \begin{cases} \vert f \vert & h \cap f \neq \emptyset \\ 0 & \text{otherwise} \end{cases} \]

Hyperreconfiguration Costs

- Cost for indirect addressing of frames
- Describes availability of each frame

\[ ocost(c_i) = \begin{cases} k & i = 1 \lor h(c_i) \neq h(c_{i-1}) \\ 0 & \text{otherwise} \end{cases} \]

Total Costs

\[ cost = \sum_{c \in S} (scost(c) + ocost(c)) \]
Experiment I

Homogeneous Frame Sizes

- 2 test cases: 8-bit Ripple Carry Adder and LED Decoder
- Implemented on a 1-dimensional FPGA (SHyRA)
- Differing numbers of frames, all equal size
- $p$ is average in stochastic model $\Rightarrow$ overestimates cost

8-Bit Adder

<table>
<thead>
<tr>
<th>Number of Frames</th>
<th>Reconfiguration information (bits)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td>Frame model (stochastic)</td>
</tr>
<tr>
<td></td>
<td>Frame model (real data)</td>
</tr>
<tr>
<td></td>
<td>2-level reconfigurable (I-HD-Switch)</td>
</tr>
<tr>
<td>5000</td>
<td>$k_{opt}$</td>
</tr>
<tr>
<td>10000</td>
<td></td>
</tr>
<tr>
<td>15000</td>
<td></td>
</tr>
<tr>
<td>20000</td>
<td></td>
</tr>
<tr>
<td>25000</td>
<td></td>
</tr>
<tr>
<td>30000</td>
<td></td>
</tr>
<tr>
<td>35000</td>
<td></td>
</tr>
</tbody>
</table>

LED Decoder

<table>
<thead>
<tr>
<th>Number of Frames</th>
<th>Reconfiguration information (bits)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5000</td>
<td>Frame model (stochastic)</td>
</tr>
<tr>
<td>10000</td>
<td>Frame model (real data)</td>
</tr>
<tr>
<td>2-level reconfigurable (I-HD-Switch)</td>
<td></td>
</tr>
<tr>
<td>15000</td>
<td></td>
</tr>
<tr>
<td>20000</td>
<td></td>
</tr>
<tr>
<td>25000</td>
<td></td>
</tr>
<tr>
<td>30000</td>
<td></td>
</tr>
<tr>
<td>35000</td>
<td></td>
</tr>
</tbody>
</table>

- 2-level outperforms frame model for large frame numbers!
Heterogeneous Frame Sizes

- In previous models all frames of equal size
- Usage pattern of specific frames differs, e.g. different routing capabilities, fixed location of components
- Different sizes of frames can save cost
- Problem: How choose good frame sizes?

I-HD-FPGA Problem

- **Given:** Sequence of context requirements $S$ for a set of reconfigurable units $X$
- **Objective:** Find partition of $X$ into frames $f_1, \ldots, f_k$ such that cost are minimal

Partial reconfiguration with heterogeneous frame sizes
Solution of Problem

Frame Model

- I-HD-FPGA algorithm solves the problem
- Uses dynamic programming approach on number of frames and switches used
- Solvable in time $O(n^4 \cdot m)$ if #frames is known
- I-HD-FPGA problem solvable in time $O(\log_2 n \cdot n^4 \cdot m)$

2-level Reconfiguration

- Problem can be restated to finding groups of reconfigurable units showing similar behavior (I-HD-Switch)
- Was previously investigated (Journal of VLSI SPS '08)
- Results: NP-hard, but good heuristics exist
Experiment II

Heterogeneous Frame Sizes

- Size of frames cost optimal
- Optimal solution of Fixed-I-HD-FPGA problem for frame model
- Heuristic solution of Fixed-I-HD-Switch problem for 2-level reconfigurable architectures

8-Bit Adder

<table>
<thead>
<tr>
<th>Number of Frames</th>
<th>Frame model (homogeneous)</th>
<th>Frame model (heterogeneous)</th>
<th>2-level reconfigurable (homogeneous)</th>
<th>2-level reconfigurable (heterogeneous)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td>1.85</td>
<td>1.73</td>
<td>1.06</td>
<td>1.00</td>
</tr>
<tr>
<td>1500</td>
<td></td>
<td></td>
<td>1.06</td>
<td>1.00</td>
</tr>
<tr>
<td>2000</td>
<td></td>
<td></td>
<td>1.06</td>
<td>1.00</td>
</tr>
<tr>
<td>2500</td>
<td></td>
<td></td>
<td>1.06</td>
<td>1.00</td>
</tr>
<tr>
<td>3000</td>
<td></td>
<td></td>
<td>1.06</td>
<td>1.00</td>
</tr>
<tr>
<td>3500</td>
<td></td>
<td></td>
<td>1.06</td>
<td>1.00</td>
</tr>
<tr>
<td>4000</td>
<td></td>
<td></td>
<td>1.06</td>
<td>1.00</td>
</tr>
<tr>
<td>4500</td>
<td></td>
<td></td>
<td>1.06</td>
<td>1.00</td>
</tr>
</tbody>
</table>

LED Decoder

<table>
<thead>
<tr>
<th>Number of Frames</th>
<th>Frame model (homogeneous)</th>
<th>Frame model (heterogeneous)</th>
<th>2-level reconfigurable (homogeneous)</th>
<th>2-level reconfigurable (heterogeneous)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5000</td>
<td>1.89</td>
<td>1.71</td>
<td>1.15</td>
<td>1.00</td>
</tr>
<tr>
<td>10000</td>
<td></td>
<td></td>
<td>1.15</td>
<td>1.00</td>
</tr>
<tr>
<td>15000</td>
<td></td>
<td></td>
<td>1.15</td>
<td>1.00</td>
</tr>
<tr>
<td>20000</td>
<td></td>
<td></td>
<td>1.15</td>
<td>1.00</td>
</tr>
<tr>
<td>25000</td>
<td></td>
<td></td>
<td>1.15</td>
<td>1.00</td>
</tr>
<tr>
<td>30000</td>
<td></td>
<td></td>
<td>1.15</td>
<td>1.00</td>
</tr>
<tr>
<td>35000</td>
<td></td>
<td></td>
<td>1.15</td>
<td>1.00</td>
</tr>
</tbody>
</table>
Conclusion

- Two approaches to partially reconfigure FPGAs
  - Frame-based with direct addressing of frames
  - 2-level reconfiguration using indirect addressing
- Formal modeling of both approaches
- Stochastic analysis of Frame model
  - Limits on feasibility
  - Determination of optimal number of frames (and frame size)
- 2-level reconfiguration outperforms frame-based reconfiguration
  (on test cases, in particular for fine granularity)
- Introduced heterogeneous frame sizes for both approaches to reduce reconfiguration overhead
  - Algorithm for determining optimal frame sizes for frame-based model
  - Heuristic for determining frame sizes for 2-level reconfiguration model
- 2-level reconfiguration still outperforms frame-based reconfiguration (on test cases, in particular for fine granularity)
Thank you very much for your attention!