Abstract

An adaptive redundancy-based fault-handling approach exploiting the partial dynamic reconfiguration capability of SRAM-based FPGAs is evaluated. Fault signal in Signal Processing Systems is accomplished using a simplex hardware arrangement while a deterministic fault isolation scheme is employed, which neither requires test vectors nor suspends the computational throughput. The approach is validated by implementation of Discrete Cosine Transform (DCT) and Motion Estimation (ME) blocks for a H.263 video encoder benchmark in Xilinx Virtex-4 FPGA.

Introduction

• Reconfigurable Fabrics – enable novel adaptive recovery approaches – “Beyond Redundancy” [3]
• Fault Recovery – efficient use of reconfiguration resources – identifying faults in moving boundaries
• Resource Escalation [4] – enables a continuum of energy vs. quality tradeoffs

Resource Escalation

• Allocate Reconfigurable Slack (RS) depending upon input signal characteristics and area margin
• Time-multiplex the RS with different functions to compare their outputs with other active PEs
• (FaDReS) Fault Demotion using Reconfigurable Slack – Identified healthy RS is utilized to achieve diagnosis of all resources in the datapath. Requires deployment of high-priority function
• (PURE) Priority Using Resource Escalation – Identified healthy RS is utilized immediately for computation of highest-priority function – Focuses on availability and quality of throughput during fault recovery phase

Research Contributions

Oblivious Fault-Detection: Intrinsic measurement of applications’ health-metric using feedback loop → Simplex operation for most of mission. Desirable Fault-Isolation: System is kept online while concurrent error detection is performed using actual runtime inputs → No need for test vectors. Degraded Quality vs. Energy Consumption: Resources computing least priority functions can be reconfigured → Throughput is application-regulated.

Figure 1: Health-Metric-based Fault-Handling Strategy, motivating example showing image reconstruction with (a) fully functional DCT module, PSNR=35.21dB (b) faulty PE in DCT module which computes DC-coefficient (more significant to output quality than QPs computing AC coefficients), PSNR=7.07dB.

Figure 2: FPGA-based video encoder platform. “PSNR” is used as a health-metric in closed loop to regulate energy/resilience/quality.

Figure 3: Fault isolation (FI) and Recovery (FR) for 1D 8-point DCT. Here, PEi (active) and PEj (RS) are the faulty PEs which need to be identified and removed from the datapath. Initially, all resources (PEs) are deemed suspect (S). FaDReS starts by identifying a healthy RS (2nd iteration) and then proceeds to mark resources (PEs) as healthy (H) or faulty (F).

Figure 4: Fault Isolation and Recovery (FR) for H.263 video encoder block using different algorithms 1) FaDReS, 2) PURE. Fault-Handling Mechanism is triggered when there is a difference of 3% in PSNR (health-metric). Faults are injected in PE1 (DC-coefficient) and PE7 (AC-coefficient).

Table 1: Dynamic Energy Consumption of FaDReS during fault isolation for various fault rates in terms of number of faulty modules (Nf). Energy is calculated by product of power consumed during FI and latency of FI.

<table>
<thead>
<tr>
<th>Nf</th>
<th>E (Joules)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.91</td>
</tr>
<tr>
<td>2</td>
<td>1.27</td>
</tr>
<tr>
<td>3</td>
<td>1.57</td>
</tr>
<tr>
<td>4</td>
<td>1.81</td>
</tr>
<tr>
<td>5</td>
<td>1.99</td>
</tr>
<tr>
<td>6</td>
<td>2.16</td>
</tr>
<tr>
<td>7</td>
<td>2.18</td>
</tr>
</tbody>
</table>

Table 2: Number of RS created using resource predict algorithm which utilizes motion information while maintaining bit-rate within 3%.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Average Bit-rate (Mbps)</th>
<th>Avg. Increase in Bit-rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fault-free</td>
<td>3.75 0.0% (ref)</td>
<td></td>
</tr>
<tr>
<td>Faulty</td>
<td>8.17 15.4%</td>
<td>10.7%</td>
</tr>
</tbody>
</table>

Table 3: Bit-rate of encoded bitstreams for Encocana video sequence using various architectures.

Conclusion

Energy-Aware Fault-Handling

• A simplex configuration is shown to be sufficient for applications such as DCT when a health-metric such as PSNR is available
• Graceful Degradation during Diagnosis
• Degradation spanned a few frames, during which time a partial throughput is available, as an intrinsic provision of degraded mode

Contact

Rizwan Ashraf
Department of Electrical Engineering and Computer Science
University of Central Florida, Orlando, FL 32816
Email: rizwan.ashraf@knights.ucf.edu
Website: cal.ucf.edu