Glitch-aware Variable Pipeline Optimization for CGRAs

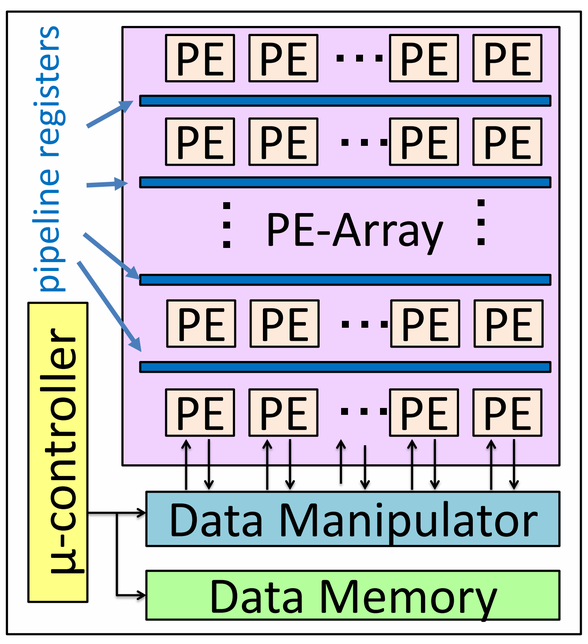
Takuya Kojima, Naoki Ando, Hayate Okuhara and Hideharu Amano, Keio University, Japan

Introduction

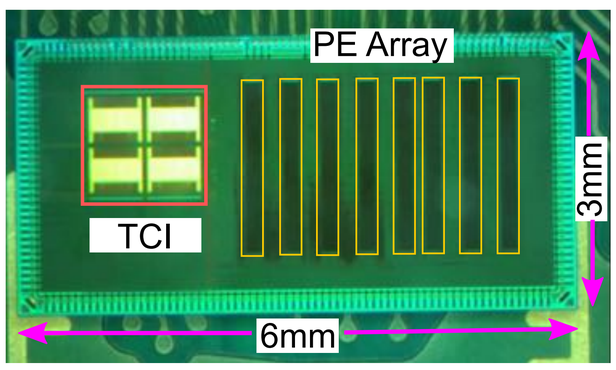
CGRA (Coarse-Grained Reconfigurable Array) is a type of platform proposed as accelerators for the forthcoming IoT and wearable computers. Especially, pipelined CGRAs can achieve high energy efficiency with control of their pipeline structure according to performance requirements. On the other hand, an increase of the dynamic power caused by glitch propagation can happen depending on the pipeline configuration. In this work, a dynamic power model considering glitch effects and an optimization method using it are proposed. Results of real chip measurements show that the optimized pipeline structures can achieve smaller energy consumption than fixed pipeline structures.

Pipelined CGRA

VPCMA (Variable Pipelined Cool Mega Array)



VPCMA diagram

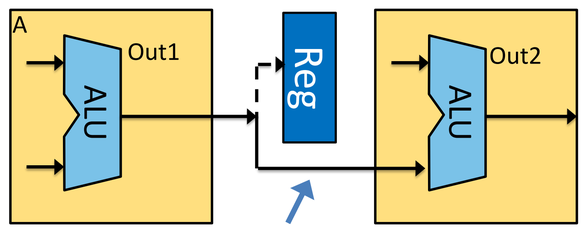


VPCMA real chip

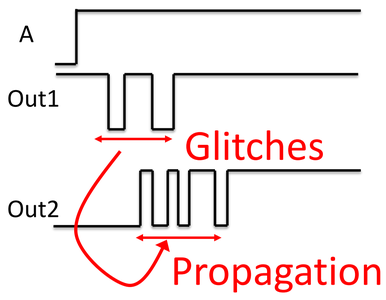
- ➢ PE (Processing Element)
  - No register file
  - No clock tree
- ➢ PE Array
  - 12 cols ×8 rows of PEs
  - Static configuration
  - 7 pipeline registers
- ➢ Variable Pipeline Structure
  - 2 modes of pipeline register
    1. Latch mode
    2. Bypass mode
  - Tradeoff between power and performance
- ➢ µ-controller
  - Data transfer control between data memory and PE array
- ➢ Real Chip of VPCMA
  - Operating experience
  - Achieving 2400 MOPS with 3.4mW
- ➢ Other pipelined CGRA
  - PipeRench, XPP, S5 Engine, EGRA, DT-CGRA

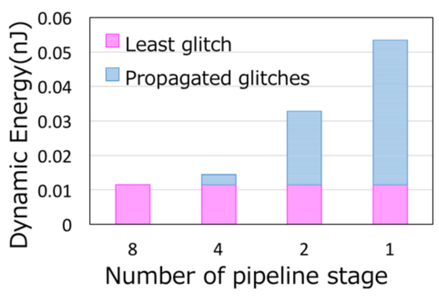
Negative impact of glitch propagation

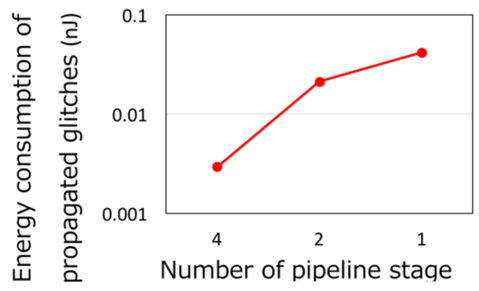
- ➢ PE unification with Bypassing registers
  - Feasible when the total delay time of PEs < clock cycle
  - Propagating undesirable switchings (glitches) to next PEs
  - Increase of PE dynamic power



Bypassing







- ➢ Power due to glitch propagation
  - Accounting for up to 80% of total power
- ➢ Small pipeline stage
  = Many unified PEs
  - Higher than linear increase

➢ Optimization of pipeline structure with a glitch-aware power model is necessary
- Related work: glitch propagation models for FPGA [1][2][3]

Proposed method and evalutation

Model considering glitches

- ➢ Dynamic power of PE array : $P_{dyn} = E_{sw}\, S_{total} f$
  - E_{sw}: Energy consumption per a swithing , S_{total}: Total switching counts, f: Frequency

  → A model to evaluate the swithing count is needed
- Classification of switching
  1. Necessary switching for computation
  2. Glitches generated within a single PE

  } Depending on mapped operations
  3. Glitches due to propagation

     → Depending on the switching of previous PE
- ➢ $S_{total} = \sum_{i=0}^{n} \sum_{j=0}^{m} S_{PE}(i,j)$
  - n: number of rows, m: number of cols, i: Index of row, j: Index of col, $S_{PE}(i,j)$: switching count of a PE at i-th row and j-th col
- ➢ $S_{PE}(i,j) = S_{single}(op) + \beta\gamma^{length} \max_{dir} S_{prev}(dir)$

$\gamma = 0.2,\ \beta = 0.9$

AND (S_{single} =) $length = 2$

$S_{PE} = 5 + 0.9\times(0.2)^2\times\max\{28.6\} = 6.03$

MULT (S_{single}) $length = 1$

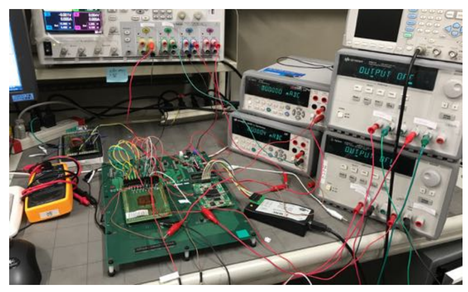
$S_{PE} = 25 + 0.9\times0.2\times\max\{20, 5\} = 28.6$

ADD (S_{single}) AND (S_{single} =)

- $S_{single}(op)$
  - Switching count of a single PE
- β, γ
  - Propagation factors
- $length$
  - Hop counts from an active pipeline register
- $S_{prev}(dir)$
  - S_{PE} of the predecessor PE

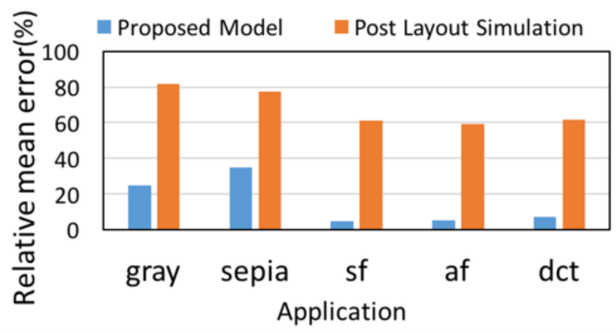
Accuracy of the proposed model

- ➢ Measuring the power of all pipeline structure
  - ➢ 5 kinds of test bench
- ➢ Obtaining parameters (E_{sw}, β, γ)
  - ➢ With least-square method



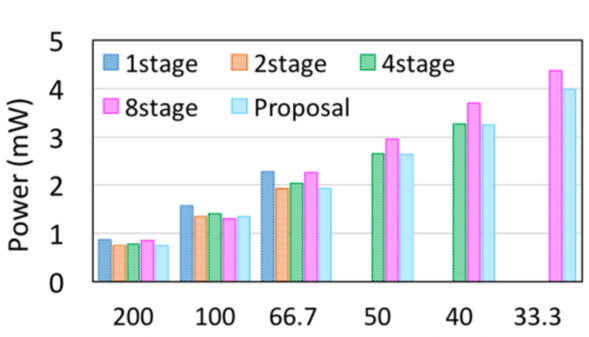
Obtained parameters

E_{sw} (pj/sw)	0.1117
β	0.053
γ	1.325



Power optimization

- ➢ Minimizing $P_{total} = \underbrace{E_{sw}\, S_{total} f}_{\text{proposed model}} + \underbrace{P_{reg,clk} \times N_{reg}}_{\text{power of registers}} + \underbrace{P_{leak}}_{\text{leakage}}$
- ➢ Constraints: Critical path delay ≤ Maximum allowed delay
- ➢ Compared to
  - ➢ Fixed pipeline structure
    - ➢ 1, 2, 4 and 8 stage
- ➢ Results
  - ➢ Smaller power consumption
  - ➢ Guarantee of operation



Reference: [1] C.Lei, *et al*. DAC'07, pp.318-323 [2] S.Cromar, *et al*. DAC'09, pp.838-843 [3] G.Altaf, *et al*. FPT'06, pp.349-352