Title: ReMCA: A Reconfigurable Multi-Core Architecture for Full RNS Variant of BFV Homomorphic Evaluation

(TCAS I)

Preliminaries

The Textbook BFV Scheme

Parameter Setup

polynomial degree: 4096
the size of modulus q: 128 bit(product of four 32 bit primes)
the standard deviation of the Gaussian distribution to σ: 3.19
the size of the larger modulus Q to at least 288-bit
32-bit primes to construct the RNS for our implementation

Algorithm And Approach

Unified Low-Complexity NTT/INTT

""

It is a algorithm can control the butterfly unit to do NTT or INTT, decrease the complexity of memory access.

Algorithm And Approach

RNS Basis Extension

w:200px,

Architecture

Overall Architecture

w:150px,
Reconfigurable PE Array: perform NTT/INTT, the modular multplication. One
row or multiple rows of PE can be configured as a channel to perform the polynomial arithmetic operations on one RNS
base.
TF ROM Array: store the twiddle factor array
Data RAM Array: store the input polynomials, intermediate results and final results
Total of 40 PEs, in which each row corresponds to one channel and

Architecture

Reconfigurable PE Unit

w:150px,
Reconfigurable PE: INTT,NTT,MULT
Barrett modular multiplier: modular multiplier, modular reduction
1.PE not onlu supports the functions with the variable modulus, but also supports the summation of modular multiplication
2.By merging the multiplicative factor 1/2 into the twiddle factors, the reconfigurable PE eliminates the multiplication of 1/2 in the subtraction path and improves the performance of PE unit.

$$\frac{x}{2} = (2\lfloor \frac{x}{2} \rfloor + 1)\frac{q+1}{2} = \lfloor \frac{x}{2} \rfloor (q+1) + \frac{q+1}{2} = \lfloor \frac{x}{2} \rfloor + \frac{q+1}{2} (modq)$$

Architecture

Confilct-Free Memory Access for NTT/INTT

w:150px,
Bank is dual-port pattern(could select the bank read port based on PEs)
For the bit-reversal operation, could change the address mapping pattern to avoid timing-consuming or memory-consuming.

Unified Computing Model

Unified Hardware Architecture Mapping Model:

w:150px,
model 1: compute 32 bits modular mult of four contiguous integers in vector A_i nad vector B_i or four constants in parallel.
model 2: compute the summation of four products, while the inputs of four products are from four different vectors and four constants respectively
model 3: The NTT/INTT transforms are computed using Mode 3
model 4: compute the summation of four products followed by a rounding operation

Unified Data Memory Organization Model:

w:150px,
MEM consists of four memory banks, where each memory bank further contains four 1024-depth and 32-bit-width dual-port RAMs.
MEMA is used to store
the inputs/outputs and intermediates results of almost all functional units in homomorphic evaluation of RNS-BFV
except for the NTT and INTT.
MEMB is mainly used to store the inputs/outputs and intermediate results of NTT and INTT
w:150px,

Mapping Method And Execution Flow

The homomorphic multiplication of RNS-BFV includes four computing units: basis extension, ciphertext multiplica-tion, basis scaling and relinearization

Computing Units Mapping

Basis Extension Unit

avatar

Computing Units Mapping

Ciphertext Multiplication Unit

avatar

Computing Units Mapping

Basis Scaling Unit:

avatar

Execution Flow of RNS-BFV

avatar

IMPLEMENTATION RESULTS AND COMPARISONS

FPGA Implementation Result

Xilinx Vivado tool
Virtex-7 XC7VX1140T
synthesized the design and achieved 250MHz
frequency under the parameter set (N = 4096, log(q) = 128-bit, log(qi) = 32-bit)

avatar

avatar