

ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ

ΤΜΗΜΑ ΕΠΙΣΤΗΜΗΣ ΥΠΟΛΟΓΙΣΤΩΝ

ΠΑΡΟΥΣΙΑΣΗ / ΕΞΕΤΑΣΗ ΜΕΤΑΠΤΥΧΙΑΚΗΣ ΕΡΓΑΣΙΑΣ

Μουσούρος Ορέστης

Μεταπτυχιακός Φοιτητής

Τμήμα Επιστήμης Υπολογιστών, Πανεπιστήμιο Κρήτης

Επόπτης Μεταπτυχιακής Εργασίας: Καθηγητής, Μ. Κατεβαίνης

Δρ. Ν. Χρυσός (επιβλέπων)

Τετάρτη, 13 Απριλίου 2022, ώρα 16:15 μ.μ.

Join Zoom Meeting

<https://zoom.us/j/94497047507>

“Επικοινωνία Ροών μεταξύ Νησιών Συνοχής RISC-V, με Πολιτικές Κρυφής Μνήμης Ανάγνωσης-Ακύρωσης και Εγγραφής-δια μέσου-Συνδυασμού”

ΠΕΡΙΛΗΨΗ

Τις τελευταίες δεκαετίες, η τεχνολογία έχει φτάσει ένα σημείο αργής κλιμάκωσης, κυρίως λόγω περιορισμών που οφείλονται στις αυξημένες ανάγκες κατανάλωσης ενέργειας, με επιπτώσεις όπως τη δυσκολία αύξησης της ταχύτητας ενός πυρήνα ή προσθήκης περισσότερων πυρήνων σε πολυπύρηνους επεξεργαστές. Επειδή υπάρχει ακόμα ανάγκη για αύξηση της απόδοσης, οι αρχιτέκτονες υπολογιστών έχουν στραφεί σε ενεργειακά αποδοτικούς επεξεργαστές, συμπεριλαμβανομένων ορισμένων που βασίζονται στην ανοιχτού κώδικα Αρχιτεκτονική Συνόλου Εντολών (Instruction Set Architecture - ISA) RISC-V, η οποία υπόσχεται ενεργειακή απόδοση, χαμηλό κόστος υλοποίησης και βελτιωμένη απόδοση σε πολυπύρηνους επεξεργαστές.

Η παρούσα εργασία συμβάλλει στη σχεδίαση και υλοποίηση μιας νέας προσέγγισης επικοινωνίας ροών μεταξύ επεξεργαστών που βρίσκονται σε διαφορετικά Νησιά Συνοχής (Coherence Islands) RISV-V. Παραδοσιακά, τα νησιά συνοχής επικοινωνούν μέσω δικτύων σε επίπεδο συστήματος, τα οποία βασίζονται σε διασυνδέσεις που χρησιμοποιούν είτε TCP/IP ή Απομακρυσμένες Άμεσες Προσπελάσεις Μνήμης (Remote Direct Memory Access - RDMA). Σε αυτές τις δομές, οι κόμβοι επικοινωνίας ανταλλάσσουν δεδομένα που βρίσκονται αποκλειστικά στις μνήμες τους, κάτι που αυξάνει τις χρονικές καθυστερήσεις και εξαντλεί κύκλους επεξεργασίας. Το RDMA βελτιώνει την επικοινωνία μεταξύ μνημών, προσφέροντας

μεταφορές δεδομένων οι οποίες εκκινούνται σε επίπεδο χρήστη, με μηδενικές αντιγραφές και μηδενικές επεξεργαστικές επιβαρύνσεις.

Σε αυτή την εργασία, σκοπεύουμε να προσφέρουμε επικοινωνία μεταξύ ενός πυρήνα κι ενός άλλου (απομακρυσμένου) κόμβου, ο οποίος μπορεί να είναι ένας πυρήνας ή μια μνήμη. Συγκεκριμένα, προτείνουμε μια καινούρια Κρυφή Μνήμη αποκλειστικά για την υποστήριξη επικοινωνίας ροών, η οποία βρίσκεται δίπλα από την Κρυφή Μνήμη Επιπέδου 1 (L1 Cache) του πυρήνα και χρησιμοποιεί την ίδια γρήγορη διεπαφή για επικοινωνία με αυτόν. Χωρίσαμε την Κρυφή Μνήμη Ροών σε δυο μέρη λογικής: α) του παραγωγού, όπου το εξερχόμενο μέρος διαχειρίζεται δεδομένα που αναχωρούν από τον κόμβο, και β) του καταναλωτή, όπου το εισερχόμενο μέρος διαχειρίζεται δεδομένα που καταφθάνουν στον κόμβο. Ουσιαστικά, στην προτεινόμενη δομή διαχείρισης ροών, αντί τα δεδομένα να μετακινούνται μεταξύ των κυρίων μνημών των κόμβων, τα δεδομένα τόσο του παραγωγού, όσο και του καταναλωτή, μπορούν να προσπελαστούν με καθυστέρηση όπως αυτής της L1 Cache. Για να βελτιώσουμε την απόδοση, επιλέξαμε οι πολιτικές της Κρυφής Μνήμης Ροών να βασίζονται στην αρχή μοναδικής-ανάγνωσης/μοναδικής-εγγραφής, ώστε να γίνεται άμεση ανακύκλωση του χώρου δεδομένων ροών στα οποία έχει υπάρξει ήδη πρόσβαση. Επιπλέον, ένας Προανακτητής (Prefetcher) ανακτά δεδομένα από τον (απομακρυσμένο) κόμβο πριν χρειαστούν, με αποτέλεσμα τη μείωση του κόστους στις προσβάσεις ανάγνωσης, ενώ οι προσβάσεις εγγραφής επωφελούνται από έναν Συνδυαστή Εγγραφών (Write-Combiner), ο οποίος συνδυάζει γειτονικά δεδομένα και τα στέλνει στον (απομακρυσμένο) κόμβο. Στην εργασία μας, οι προσβάσεις σε δεδομένα ροών αναγνωρίζονται από τις εικονικές διευθύνσεις των εντολών, χωρίς την ανάγκη επέκτασης του ISA.

Υλοποιήσαμε αυτό το σύστημα, με τη γλώσσα περιγραφής υλικού SystemVerilog, και το προσθέσαμε ως επέκταση του μονοπύρηνου RISC-V επεξεργαστή CVA6 (πρώην ARIANE). Τα Εισερχόμενα και Εξερχόμενα μέρη λογικής της Κρυφής Μνήμης Ροών χρησιμοποιούν το καθένα (4) πλαίσια εργασίας σε πραγματικό υλικό προκειμένου να υποστηρίξουμε εικονικοποίηση, και είναι άμεσα συνδεδεμένα με τη Μονάδα Αναγνώσεων/Εγγραφών (Load/Store Unit - LSU) του ARIANE. Επίσης, στα άκρα έχει υλοποιηθεί λογική επικοινωνίας, η οποία ζητά και στέλνει δεδομένα μέσω μιας διασύνδεσης AXI-4.

Η εργασία μας έχει υλοποιηθεί για τη Συστοιχία Επιτόπια Προγραμματιζόμενων Πυλών (Field Programmable Gate Array - FPGA) Zynq UltraScale+ MPSoC της Xilinx. Για το Εισερχόμενο μέρος λογικής, από πλευράς χώρου χρησιμοποιήθηκαν 16839 Προγραμματιζόμενες Πύλες (LUTs), 7506 Καταχωρητές και 8 Μνήμες Τυχαίας Προσπέλασης (BRAMs), λειτουργώντας στα 275 MHz, ενώ για το Εξερχόμενο μέρος λογικής, χρησιμοποιήθηκαν 23606 LUTs, 8615 Καταχωρητές και 8 BRAMs, λειτουργώντας στα 210 MHz.

Προσομοιώσαμε την υλοποίησή μας προκειμένου 1) να επαληθεύσουμε τη λειτουργικότητα των ροών σε συνδυασμό με πυρήνες RISC-V και 2) να αξιολογήσουμε την απόδοσή της. Στις αξιολογήσεις μας, μεταφέρουμε δεδομένα ροών από και προς την κυρίως μνήμη του πυρήνα ARIANE, χρησιμοποιώντας πρώτα την παραδοσιακή ιεραρχία μνήμης και ύστερα την βελτιστοποιημένη Κρυφή Μνήμη Ροών. Τα αποτελέσματα παρουσιάζουν κέρδη απόδοσης χάρη στις πολιτικές βελτιστοποίησης ροών της υλοποίησή μας, αφού επιτυγχάνεται η σχεδόν πλήρης εξάλειψη των χρονικών καθυστερήσεων της διασύνδεσης του δικτύου στα ενδεικτικά προγράμματα συγκριτικής αξιολόγησης, χωρίς την υποστήριξη λειτουργικού συστήματος.

University of Crete

Computer Science Department

M.Sc. Thesis

Mousouros Orestis

**Master's Thesis Supervisor: Professor M. Katevenis
Dr. N. Chrysos(Thesis Co-Advisor)**

Wednesday, 13 April 2022, 16:15p.m.

Join Zoom Meeting

<https://zoom.us/j/94497047507>

**“Stream communication across RISC-V Coherence Islands, with
Read-Invalidate and Write-through-Combine Cache Policies”**

ABSTRACT

In the last decades, technology has reached a point of slow scaling, mainly due to limitations caused by the increasing amounts of power consumption. To gain performance speedup, hardware architects have turned to energy efficient processors, including some that are based on open-source RISC-V Instruction Set Architecture (ISA), which promise energy efficiency and high performance on multi-core chips.

This thesis contributes the design and implementation of a new approach for interprocessor stream communication across RISC-V Coherence Islands. Traditionally, the coherence islands use memory-to-memory communication over TCP/IP or Remote Direct Memory Access (RDMA) interconnections. Writing and reading data to and from memory at the endpoints heightens latency and depletes processor cycles. Instead, in our work, the communication confines itself between a core and another (remote) node, which can either be a core or a memory. In particular, we propose a new Streaming Cache that resides next to Level 1 Cache (L1 Cache) and uses the same fast interface for communication with the core. We split the Streaming Cache into two logical parts: a) the producer, an outgoing streaming cache that handles streaming data departing from the node; b) the consumer, an incoming streaming cache that handles streaming data arriving to the node. Effectively, in the proposed streaming framework, instead of moving data across the main memory of the end-points, data of both the producer and the consumer can be accessed with same latency as the L1 Cache. To improve performance, we use the read-once/store-once cache policies in the Streaming Cache, which immediately recycle the space of already accessed streaming data. Furthermore, a Prefetcher fetches data from the (remote) node before they are needed, thus reducing the

cost of read accesses, while the write accesses take advantage of a Write-Combiner, which combines neighboring data and sends them to the (remote) node. In our work, accesses to streaming data are recognized using virtual addresses without the need of extending ISA.

We implemented the proposed system in SystemVerilog, as an extension of the CVA6 (former ARIANE) single-core RISC-V CPU. We built the Incoming and Outgoing schemes of Streaming Cache, each with four (4) contexts (hardware streams) to support virtualization, and we tightly-coupled them with the Load/Store Unit (LSU) of the ARIANE. We also built a communication logic at the edges that sends/receives data over an AXI-4 interconnect.

We synthesized our design for Xilinx Zynq UltraScale+ MPSoC Field Programmable Gate Array (FPGA). The Incoming logic of our design utilizes 16839 Look-Up Tables (LUTs), 7506 Registers and 8 Block Random Access Memories (BRAMs), and operates at 275 MHz, while the Outgoing logic utilizes 23606 LUTs, 8615 Registers and 8 BRAMs, and operates at 210 MHz.

We performed behavioral simulations to our RTL design in order 1) to verify the streaming functionality when coupled with the RISC-V cores and 2) to evaluate its performance. In our preliminary evaluations, we stream data from/to main memory of the ARIANE core, first using the traditional memory hierarchy and second using our optimized streaming cache. The promising results underline the performance gains due to the stream-optimized cache policies of our design, by managing to almost completely eliminate the latency of network's interconnection in our indicative hand-made bare metal benchmarking programs.

