

## ΤΜΗΜΑ ΕΠΙΣΤΗΜΗΣ ΥΠΟΛΟΓΙΣΤΩΝ

### ΠΑΡΟΥΣΙΑΣΗ / ΕΞΕΤΑΣΗ ΜΕΤΑΠΤΥΧΙΑΚΗΣ ΕΡΓΑΣΙΑΣ

Τζανάκης – Αρναουτάκης Λέανδρος  
Μεταπτυχιακός Φοιτητής

Τμήμα Επιστήμης Υπολογιστών, Πανεπιστήμιο Κρήτης  
Επόπτης Μεταπτ. Εργασίας: Καθηγητής, Μ. Κατεβαίνης

Πέμπτη, 28/02/2019, 17:30

Αίθουσα A121, Τμήμα Επιστήμης Υπολογιστών, Πανεπιστήμιο Κρήτης

**“ Ανάπτυξη Λογισμικού Βελτίωσης Της Ποιότητας Υπηρεσιών Σε Μικροελεγκτή Πραγματικού Χρόνου Για Χαμηλής Κατανάλωσης Λειτουργίες Άμεσης Προσπέλασης Μνήμης ”**

#### ΠΕΡΙΛΗΨΗ

Η υπολογιστική υψηλής απόδοσης (HPC) συμβάλλει καταλυτικά στην πρόοδο της επιστήμης καθώς και στην αύξηση του ανταγωνισμού στην βιομηχανία σε παγκόσμιο επίπεδο. Στις μέρες μας, η απόδοση των υπερυπολογιστών περιορίζεται από αυστηρές προϋποθέσεις ενεργειακής κατανάλωσης. Ωστόσο η χρήση διακομιστών χαμηλής κατανάλωσης σε συνδυασμό με επιταχυντές FPGA υψηλής ταχύτητας μπορούν να προσφέρουν μια εφικτή λύση για την αντιμετώπιση αυτής της πρόκλησης. Σε αυτή την κατεύθυνση, το έργο ExaNeSt το οποίο χρηματοδοτείται από την Ευρωπαϊκή Ένωση αναπτύσσει και παράγει ένα σύστημα που αποτελείται από επεξεργαστές αποδοτικής κατανάλωσης τεχνολογίας ARM, σε συνδυασμό με FPGAs. Στο σύστημά μας, χρησιμοποιούμε FPGAs για να υλοποιήσουμε μιας υψηλής ταχύτητας προσαρμοσμένη διασύνδεση η οποία επιτρέπει στους υπολογιστικούς κόμβους καθώς και στις γρήγορες μνήμες τους να επικοινωνούν μεταξύ τους. Αυτό δημιουργεί την ανάγκη για μια εξελιγμένη διεπαφή δικτύου η οποία γεφυρώνει τις εφαρμογές που τρέχουν στους πυρήνες ARM με το υλικό διασύνδεσης. Για μεγάλες μεταφορές δεδομένων μεταξύ μνημών, έχουμε υλοποιήσει μια προσαρμοσμένη πολυκάναλη μηχανή RDMA (Remote Direct Memory Access) με χαμηλή καθυστέρηση, η οποία επιτρέπει στις εφαρμογές να

παρακάμπτουν τον πυρήνα του λειτουργικού ώστε να αποφεύγεται το κόστος κλήσεων συστήματος καθώς και η παραδοσιακή επεξεργασία πρωτοκόλλων τύπου TCP/IP. Σε αυτήν την εργασία, αναπτύξαμε λογισμικό, χρησιμοποιώντας έναν ειδικό επεξεργαστή πραγματικού χρόνου, διάφορα στάδια του πρωτοκόλλου RDMA, συμπεριλαμβανομένης μιας καινοτόμας τμηματοποίησης του μεγέθους μεταφοράς RDMA σε μπλοκ, εφαρμογή χρονικών ορίων και αναμετάδοσης ανά μπλοκ, ποιότητα υπηρεσιών QoS, καθώς και έλεγχο ροής από άκρο σε άκρο και ένα νέο πρωτόκολλο για ειδοποιήσεις γρήγορης ολοκλήρωσης. Η νέα μηχανή RDMA υποστηρίζει επιλεκτικές και πολλαπλών διαδρομών αναμεταδόσεις ανά μπλοκ, οι οποίες προωθούν την υπερσύγχρονη τεχνολογία RDMA του InfiniBand πρωτοκόλλου. Η νέα RDMA επίσης, συμπεριλαμβανομένου του τμήματος συν-επεξεργαστή, το οποίο είναι το αποτέλεσμα αυτής της εργασίας, και το τμήμα υλικού που υλοποιήθηκε στο ITE, είναι πλέον πλήρως λειτουργική και χρησιμοποιήθηκε για την εκτέλεση πραγματικών εφαρμογών HPC στο πρωτότυπο ExaNeSt, το οποίο αποτελείται από δεκάδες διασυνδεδεμένων Ultrascale + MPSoCs. Με την εφαρμογή πολλών λειτουργιών επίπεδο μπλοκ και συναρτήσεων μεταφοράς χρησιμοποιώντας τον συν-επεξεργαστή, έχουμε μειώσει την πολυπλοκότητα και τον χρόνο ανάπτυξης πρωτοκόλλου RDMA, χωρίς αυτό να επηρεάσει το ρυθμό επεξεργασίας του. Σε αυτή τη διατριβή, αναφέρουμε τα χαρακτηριστικά που έχουμε εφαρμόσει στον συν-επεξεργαστή, το σκεπτικό πίσω από την επιλογή σχεδιασμού μας και τα αποτελέσματα της αξιολόγησης απόδοσης σε επίπεδο συστήματος.

**Tzanakis Arnaoutakis Leandros**  
**M.Sc. Thesis**

**Computer Science Department**  
**University of Crete**  
**Master's Thesis Supervisor: Professor, M. Katevenis**

**Thursday 28/02/2019, 17:30**  
**Room A121, Computer Science Dept., University of Crete**

**“Quality of Service Framework for Low Power RDMA Operations over Cortex R5 Real Time Microcontroller”**

### **ABSTRACT**

The High Performance Computing (HPC) contributes to the progress of science and the competitiveness of global industry. Nowadays, scaling the performance of

supercomputers is limited by strict power consumption constraints. Low-power servers tightly coupled with high-speed FPGA accelerators can offer a feasible solution to deal with this challenge. Along this direction, the ExaNeSt EU-funded project develops and prototypes a system composed of power-efficient ARM-based processors, tightly coupled with FPGAs. In our system, we leverage the FPGAs in order to implement a custom low-latency interconnect that will allow computing nodes to communicate with each other as well as with fast, non-volatile, in-node storage devices. This creates the need for a sophisticated network interface to bridge the processes that run on the ARM cores with the interconnection hardware. For bulk memory-to-memory transfers, we have developed a custom low-latency multi-channel Remote Direct Memory Access (RDMA) engine, which allows processes to bypass the kernel in order to avoid the overheads of system calls and of traditional TCP/IP protocol processing. In this thesis, we have implemented in software, using a special Real Time co-processor, several stages of the RDMA protocol, including a novel transfer segmentation into blocks, per-block timeouts and retransmissions, quality-of-service (QoS), as well as end-to-end flow control and a novel protocol for fast completion notifications. The new RDMA supports per-block multi-pathing and selective (block-level) retransmissions, which advance InfiniBand state-of-the-art RDMA. The new RDMA, including the co-processor part, which is the outcome of this thesis, and the hardware part implemented at FORTH, is now fully functional, and has been used to run real HPC applications on the ExaNeSt prototype, which consists of tens of interconnected Ultrascale+ MPSoCs. By implementing several block and transfer level functions using the co-processor, we have reduced the complexity and development time of the RDMA, without affecting its processing rate. In this thesis, we report the features that we have implemented in the co-processor, the rationale behind our design choice, and system-level performance evaluation results.