UNIVERSITATEA LUCIAN BLAGA DIN SIBIU

Școala doctorală de Științe Inginerești și Matematică
Domeniul de doctorat: Calculatoare și Tehnologia Informației

Teză de doctorat

Tehnici Anticipative și Predictive în Microprocesoarele Multicore

Autor:

Ing. Claudiu-Raul Buduleci

Conducător științific:

Prof. Dr. Ing. Remus Brad

SIBIU 2024

REZUMAT

Această teză oferă o analiză comprehensivă a cercetărilor și progreselor existente în diverse aspecte ale proiectării și optimizării microprocesoarelor, acoperind provocările principale și soluțiile inovatoare. Lucrarea începe cu introducerea limitării issue bottlenek, ce reprezintă o problemă fundamentală în arhitectura microprocesoarelor, care se referă la limitarea intrinsecă a numărului de instrucțiuni ce pot fi executate într-un ciclu de tact. Sunt analizate diferite strategii și mecanisme propuse în literatura de specialitate pentru a depăși aceste limitări și pentru a îmbunătăți performanța procesorului. În plus, se abordează aspectele critice ale securității și consistenței datelor în microprocesoarele cu posibilități de execuție speculativă. În continuare, se mai analizează diverse aspecte ale reutilizării dinamice a instrucțiunilor (DIR), predicției valorilor (VP) și scalării dinamice a tensiunii și frecvenței (DVFS), inclusiv fundamentele teoretice și implementările propuse. Pe lângă cele menționate anterior, sunt analizate cele mai moderne simulatoare de microprocesoare multi/many-core și benchmark-uri paralele, ce sunt utilizate pentru evaluarea și optimizarea acestor sisteme complexe.

O primă cercetare constă în augmentarea simulatorului Sniper, oferind posibilitatea de a accesa valorile operanzilor unui grup specific de instrucțiuni. Motivația pentru accesarea valorilor operanzilor este dată de faptul că programele software, în special aplicațiile grafice și multimedia, se caracterizează printr-un grad ridicat de redundanță, lucru care poate fi exploatat prin tehnici de tip DIR. Această lucrare se încadrează în categoria „arhitecturilor cu sursa deschisă”, pornind de la conceptul open-source, deoarece oferă cercetătorilor o metodologie extensibilă pentru citirea valorilor operanzilor instrucțiunilor. Cel mai important avantaj adăugat este unul de natură tehnică, cuprinzând detalii despre arhitectura și modificările aduse simulatorului. În plus, propune o metodologie de simulare pentru a studia gradul reutilizabilitate a anumitor tipuri de instrucțiuni dinamice. Studiul experimental este realizat pe suita de benchmark-uri Splash-2 folosind simulatorul modificat, variind următorii parametri: numărul de core-uri, nivelul de optimizare a compilatorului și dimensiunea istoricului operanzilor (numărul de perechi de operanzi stocate pentru fiecare instrucțiune). Rezultatele indică un potențial promițător de reutilizare, variind în medie de la 84% la 87% pentru instrucțiunile dinamice selectate, ducând la ideea implementării unui buffer de reutilizare (RB) asociativ într-un sistem multi/many-core. Nivelul de optimizare al compilatorului influențează gradul de reutilizabilitate și performanța. O estimare aproximativă a potențialului câștig în performanță (speedup) este de asemenea calculată, atingând un maxim de 17,5% și o medie de 3,6%.

Următoarea cercetare reprezintă o contribuție originală ce augmentează arhitectura multicore Intel Nehalem prin introducerea unui buffer de reutilizare (RB) asociativ, aplicat în mod selectiv asupra instrucțiunilor aritmetico-logice de mare latență. Arhitectura este simulată utilizând simulatorul Sniper, ce a fost adaptat pentru a putea estima consumul de energie, aria de integrare și temperatura

cipului, incluzând și modificările pentru adaptarea latenței, astfel încât să integreze și unitatea funcțională nou adăugată. Implementarea unui RB asociativ este o abordare nouă, împreună cu aplicabilitatea sa într-un microprocesor multicore, aplicat instrucțiunilor aritmetice cu latență mare, vizând creșterea performanței procesorului. În plus, s-a realizat și un proces manual de explorare a spațiului de proiectare, având ca și scop găsirea parametrilor optimi ai unității nou adăugate, în raport cu metricile de interes. Simulările pe benchmark-urile Splash-2 au arătat o rată medie de reutilizare de până la 33,27%, permițând o creștere maximă de performanță de 6,56%. În timp ce consumul de energie rămâne stabil, se poate observa, în medie, o reducere a temperaturii cipului cu 2,8 °C odată cu creșterea asociativității.

Un alt punct important al acestei teze constă în implementarea și evaluarea unui VP într-un sistem multicore, aplicat în mod selectiv asupra instrucțiunilor aritmetice cu latență mare. Obiectivul este de a creștere numărului de instrucțiuni ce pot fi executate într-un ciclu de tact**,** crescând astfel performanța sistemului. Simulatorul Sniper a fost utilizat pentru a augmenta arhitectura Intel Nehalem cu un VP și adaptat pentru a estima performanța, aria de integrare, consumul de energie și temperatura cipului pentru arhitectura îmbunătățită. Au fost realizate mai multe scenarii de simulare, în care au fost variați mai mulți parametri ai unității VP: numărul de intrări, asociativitatea precum și numărul de valori utilizate pentru predicție pentru fiecare instrucțiune. Prin creșterea lungimii istoricului, s-a măsurat, în medie, o creștere a performanței cu peste 3%, o reducere a temperaturii cipului de la 57,8 °C la 56,17 °C și un consum de energie mai mic în majoritatea cazurilor, comparativ cu configurația de bază. A fost realizată o comparație originală între tehnicile VP și DIR în condiții echitabile (pentru a exploata același grad de vecinătate al valorilor), evidențiind avantajele și dezavantajele fiecărei tehnici în contextul dat. Comparația a fost făcută variind numărul de core-uri din sistem.

A fost realizată și o analiză empirică a benchmark-urilor consacrate Splash-2 comparativ cu cea mai recentă versiune, Splash-4. S-a demonstrat că, într-o configurație cu 64 de nuclee, jumătate din benchmark-urile simulate ating temperaturi mult peste pragul critic de 105 °C, subliniind necesitatea unei evaluări multi-obiectiv din cel puțin următoarele perspective: consum de energie, performanță, temperatură și aria de integrare. În timpul analizei s-a observat că cu toate îmbunătățirile adăugate în noua versiune, core-urile petrec o mare parte din timp în stare de inactivitate, aproximativ 45% în medie. Acest lucru a fost exploatat prin implementarea unei tehnici predictive DVFS numită Simple Core State Predictor (SCSP). Scopul a fost reducerea consumului total de energie prin adaptarea predictivă a frecvenței și a voltajului la nivel de core, menținând în același timp performanța. Mai mult, tehnica SCSP, care operează cu informații abstracte la nivel de core, a fost aplicată în paralel cu tehnicile VP sau DIR, care se bazează pe informații la nivel de instrucțiune. Utilizând doar tehnica SCSP, s-a obținut o reducere de 9,95% a consumului de putere și o reducere de 10,54% a energiei consumate, menținând performanța. Prin combinarea SCSP cu tehnica VP, s-a obținut o creștere a performanței de 8,87%, reducând în același timp consumul de putere și consumul de energie cu 3,13% respectiv 8,48%

Lista lucrărilor publicate

Publicațiile realizate în timpul elaborării acestei teze sunt următoarele:

- **C. Buduleci**, A. Gellert, A. Florea, R. Chis, and R. Brad, "Multi-Objective Optimization of Speculative and Anticipative Multi-Core Architectures," in *Advanced Computer Architecture and Compilation for High-performance Embedded Systems*, Fiuggi, Italy: HiPEAC, 2020, pp. 11–14.
- **C. Buduleci**, A. Gellert, A. Florea, and A. Matei, "Extending Sniper with Support to Access Operand Values: A Case Study on Reusability Measurement," in *2022 23rd International Carpathian Control Conference (ICCC)*, Sinaia, Romania: IEEE, May 2022, pp. 70–75. doi: 10.1109/ICCC54292.2022.9805869.
- **C. Buduleci,** A. Gellert, and A. Florea, "Selective High-Latency Arithmetic Instruction Reuse in Multicore Processors," in *2023 27th International Conference on System Theory, Control and Computing (ICSTCC)*, Timisoara, Romania: IEEE, Oct. 2023, pp. 410–415. doi: 10.1109/ICSTCC59206.2023.10308483. **(Best Paper Award)**
- **C. Buduleci**, A. Gellert, A. Florea, and R. Brad, "Architectural and Technological Approaches for Efficient Energy Management in Multicore Processors," *Computers*, vol. 13, no. 4, p. 84, Mar. 2024, doi: 10.3390/computers13040084.
- **C. Buduleci**, A. Gellert, A. Florea, and R. Brad, "Improving Multicore Architectures by Selective Value Prediction of High-Latency Arithmetic Instructions," *Adv. Electr. Comp. Eng.*, vol. 24, no. 2, pp. 61–72, 2024, doi: 10.4316/AECE.2024.02007.

Publicații realizate anterior de către autor, ce sunt în strânsă legătură cu această teză:

- A. Florea, **C. Buduleci**, R. Chis, A. Gellert, and L. Vintan, "Enhancing the Sniper simulator with thermal measurement," in 2014 18th International Conference on System Theory, Control and Computing (ICSTCC), Sinaia: IEEE, Oct. 2014, pp. 31–36. doi: 10.1109/ICSTCC.2014.6982386.
- R. Chis, A. Florea, **C. Buduleci**, and L. Vintan, "Multi-Objective Optimization for an Enhanced Multi-Core SNIPER Simulator," Proceedings of The Romanian Academy, Series A: Mathematics, Physics, Technical Sciences, Information Science, vol. 19, no. 1, pp. 85–93, Mar. 2018.

Succesul acestei teze se datorează în mare măsură eforturilor remarcabile ale multor persoane care mi-au oferit inspirație, expertiză și resurse.

Aș dori să încep prin a-mi exprima cea mai profundă recunoștință față de coordonatorii mei, Prof. Dr. Ing. Remus Brad, Prof. Dr. Ing. Adrian Florea și Conf. Dr. Ing. Arpad Gellert, pentru mentoratul lor de neprețuit, sprijinul și îndrumarea oferite pe parcursul acestei călătorii. De asemenea, îi sunt recunoscător domnului Conf. Dr. Ing. Daniel Morariu pentru perspectivele sale valoroase și pentru că a fost membru al comisiei de îndrumare.

Într-o notă personală, sunt profund recunoscător familiei și prietenilor mei pentru dragostea, încurajarea și răbdarea lor neclintite pe parcursul acestui proces provocator. Un mulțumesc special soției mele, Mihaela, care a fost o sursă constantă de fericire și m-a sprijinit în atingerea acestui obiectiv. Sunt, de asemenea, profund recunoscător prietenului meu apropiat, Bogdan, pentru sprijinul și inspirația acordată de-a lungul anilor.

Cu o profundă recunoștință, aș dori să dedic această lucrare în memoria domnului Prof. Dr. Ing. Lucian Vințan, care a crezut în potențialul meu și m-a inspirat să pornesc în această călătorie provocatoare.

În final, aș dori să mulțumesc tuturor celor care au contribuit direct sau indirect la realizarea acestei lucrări.

CUPRINS