DMITRII USTIUGOV

About

Dmitrii Ustiugov is a PhD student at the University of Edinburgh, co-advised by Prof. Boris Grot (the University of Edinburgh) and Prof. Edouard Bugnion (EPFL). Dmitrii's research interests span across Computer Architecture and Computer Systems with a focus on virtual memory support for cloud and serverless computing.

Contact info

UK, Scotland, EH8 9AB, Edinburgh
10 Crichton Street, office 2.05

Website: homepages.inf.ed.ac.uk/s1373190
E-mail: dmitrii.ustiugov@ed.ac.uk

Research interests

Broad: Cloud computing, virtual memory, datacenter and rack-scale systems, emerging memory systems, security.
Current focus: Virtual memory support for serverless and cloud computing.

Education

Ph.D., Computer Science (2014 – ongoing) at **EPFL** (2014-2019) and the **University of Edinburgh** (from 2019).
Advisors: **Prof. Boris Grot** (University of Edinburgh) and **Prof. Edouard Bugnion** (EPFL).
- Topic: "Hardware and software support for virtual memory for tomorrow's cloud".

M.Sc. (with Honors) and **B.Sc.** (with Honors), **Applied Mathematics and Physics** (2008-2014) at **MIPT** (Moscow Institute of Physics and Technology)
Advisor: **Dr. Alexander Butuzov** (Intel/MIPT).
- B.Sc. thesis: "CPU performance analysis using critical path methods".
- M.Sc. thesis: "CPU power consumption analysis based on cycle-accurate microarchitecture simulation".

Current Research Projects

High-performance tail-aware serverless systems in collaboration with **EPFL, ETH, AWS, Microsoft Research, Arm, Huawei**.
- **vHive**: Open-source framework for serverless experimentation. github.com/ease-lab/vhive
- Cold start reduction with Record-and-Prefetch snapshots.
- Serverless-native communication fabric for data-intensive serverless applications.
- Characterizing architectural implications of extremely multi-tenant serverless deployments.
- Tail latency sources in production serverless deployments.

Architectural support for address translation in cloud.
- Accelerating page table walks with prefetch and by improving caching efficiency of page tables in CPU caches. github.com/ease-lab/PTEMagnet

Past projects

Bankrupt Covert Communication Channel, github.com/ease-lab/bankrupt
VISA: Vertically Integrated Server Architecture, parsa.epfl.ch/visa, joint work with Oracle Labs and ICS-FORTH.
Scale-Out NUMA, parsa.epfl.ch/sonuma/sonuma.html
CloudSuite: A Benchmark Suite for Cloud Services, cloudsuite.ch
Google Perfkit Benchmarker, www.perfkitbenchmarker.org
QFlex: Quick and Flexible Computer Architecture Simulation, qflex.epfl.ch

Teaching Assistance

Introduction to Computer Systems, University of Edinburgh, Fall 2019, 2020.
Extreme Computing, University of Edinburgh, Fall 2019.
Mathematical analysis (MAN), EPFL, Spring 2018.
Computer Architecture I, EPFL, Fall 2016, Fall 2017.
Probability and Statistics, EPFL, Spring 2016.
Introduction to Multiprocessor Architecture, EPFL, Fall 2015.

Student Research Projects Supervision

Shyam Jesalpura, Research Intern, BITS, 2021.
- Characterization of data-intensive serverless workloads.

Yuchen Niu, University of Edinburgh, 2020-2021.
- Implications of multi-tenancy on serverless hosts.

Theodor Amariucai, University of Edinburgh, 2020-2021.
- *Tail latency analysis for production and research serverless clouds.*

Plamen Petrov, University of Edinburgh, Summer 2019 – Fall 2020
- End-to-end serverless benchmarking framework.
- RDMA networks security.

Ivy Wang, University of Edinburgh, Fall 2019
- Software support for contiguous page tables allocation.

Sean Mullan, University of Edinburgh, Fall 2019 – Spring 2020
- Design exploration for TLB prefetching.

Lei Yan, MSc, EPFL/RWTH Aachen University, Spring 2018
- Design exploration for user-level cooperative scheduling for latency-critical cloud services.

Siddharth Gupta, PhD candidate, EPFL, Fall 2017
- Performance analysis of persistent memory systems on modern server CPUs.

Nikhil Gupta, EPFL/IIT Delhi, Summer 2016
- Building a robust infrastructure for QFlex (Flexus) simulation framework.

Virgile Neu, EPFL, Spring 2016
- Analyzing microprocessor front-end efficiency using performance counters.

Refereed Conference Publications

1. **D. Ustiugov**, P. Petrov, M. Kogias, E. Bugnion, and B. Grot, "Benchmarking, Analysis, and Optimization of Serverless Function Snapshots". *ASPLOS'21*. **Distinguished Artifact Award.**
2. A. Margaritov, **D. Ustiugov**, and B. Grot, "PTEMagnet: Fine-Grained Physical Memory Reservation for Faster Page Walks in Public Clouds". *ASPLOS'21*.
3. A. Margaritov, **D. Ustiugov**, E. Bugnion, and B. Grot, "Prefetched Address Translation". *MICRO'19*.
4. **D. Ustiugov**, A. Daglis, J. Picorel, M. Sutherland, E. Bugnion, B. Falsafi and D. Pnevmatikatos, "Design Guidelines for High-Performance SCM Hierarchies", *MEMSYS'18.*
5. M. Drumond, A. Daglis, N. Mirzadeh, **D. Ustiugov**, J. Picorel, B. Falsafi, B. Grot, and D. Pnevmatikatos, "The Mondrian Data Engine", *ISCA'17.*
6. A. Daglis, **D. Ustiugov**, S. Novakovic, E. Bugnion, B. Falsafi, and B. Grot, "SABRes: Atomic Object Reads for In-Memory Rack-Scale Computing", *MICRO'16*.

Refereed Workshop Publications

1. **D. Ustiugov**, P. Petrov, M.R.S. Katebzadeh, and B. Grot, "Bankrupt Covert Channel: Turning Network Predictability into Vulnerability". *Workshop on Offensive Technologies at USENIX Security, 2020*.
2. A. Margaritov, **D. Ustiugov**, E. Bugnion, and B. Grot, "Virtual Address Translation via Learned Page Table Indexes". *In the Workshop on Machine Learning for Systems workshop at NeurIPS, 2018.*

Refereed Journal Publications

1. S. Novakovic, A. Daglis, **D. Ustiugov**, E. Bugnion, B. Falsafi, and B. Grot, "Mitigating Load Imbalance in Distributed Data Serving Through Rack-Scale Memory Pooling", *TOCS, 2019.*
2. M. Drumond, A. Daglis, N. Mirzadeh, **D. Ustiugov**, J. Picorel, B. Falsafi, B. Grot, and D. Pnevmatikatos, "Algorithm/Architecture Co-Design for Near-Memory Processing", *SIGOPS Operating Systems Review, 2018.*

Research and Invited Talks

Invited talks and lectures

- "Cloud Computing: Evolution, Technologies, Future", **invited lecture** at the Operating Systems course at the University of Edinburgh, 2021.
- "Benchmarking, Analysis, and Optimization of Serverless Function Snapshots", an **invited talk** at Amazon Web Services (Amazon Lambda and Firecracker teams), 2021.

- "Prefetched Address Translation", an **invited talk** at the 5th Computing Systems Day at NTUA, Athens, 2020.

Research talks

- "Benchmarking, Analysis, and Optimization of Serverless Function Snapshots", ASPLOS'21.
- "Bankrupt Covert Channel: Turning Network Predictability into Vulnerability", WOOT'20 co-located with USENIX Security.
- "Design Guidelines for High-Performance SCM Hierarchies", MEMSYS'18.
- "Design Guidelines for High-Performance SCM Hierarchies", Oracle, 2017.
- "Hardware Support for Remote Atomic Reads in Rack-Scale Systems", the EuroSys Doctoral Workshop (EuroDW) co-located with EuroSys, 2016.

Community Service and Professional Activity

Student ACM member since 2015, student USENIX member since 2020.
External reviewer for ISCA 2019, ATC 2019, and Computer Architecture Letters (CAL) 2019 and 2020.

Proposal Grants Writing Experience

- (Pending) Proposal to Huawei ($73,000): Efficient serverless applications through communication-aware function composition. PI: Boris Grot, 2021.
- Proposal funded by **Google** ($73,000): Accelerating Address Translation via a Learned Page Table Index. PI: Boris Grot, 2019.
- Proposal funded by **Oracle Labs** ($90,000): Scalable Memory Server Architecture for Disaggregated Memory Systems. PI: Virendra Marathe, Co-PI: Babak Falsafi, 2018.

Professional Experience

Research intern at **Oracle Labs**, April to June 2017
- Mentors: Dr. Virendra Marathe and Stephen Byan.
- Analysis of persistent memory applications and disaggregated memory systems.

Senior Lua/Perl developer at **IPONWEB**, December 2013 to August 2014
- Manager: Lev Leontiev.
- Developing high-load big data software platform for real-time bidding

Software simulation intern (full-time engineer since January 2012) at **Intel**, June 2010 to November 2013
- Mentors: Dr. Alexander Butuzov, Vladimir Gnatyuk.
- CPU performance/power/energy studies using software simulators and analyzer tools.

Awards and Fellowships

- **Distinguished Artifact Award** at ASPLOS'21.
- **The winner of 1-minute research pitch** competition at JOBS workshop co-located with MICRO'20.
- **Arm Center of Excellence fellowship** at the University of Edinburgh, 2019.

Other Publications

- B. Ziv B., G. Haber, L. Rumyantsev L., **D. Ustiugov**, "Use Cases for the Critical Path Analyzer Framework". *Software Professionals Conference (SWPC)*, 2013.
- A. Butuzov, O. Shimko, **D. Ustiugov**, "Fast and Easy Ways to Improve Software Development Teamwork Efficiency". *Software Professionals Conference (SWPC)*, 2013.
- K. Garifullin, **D. Ustiugov**, "Performance Simulation and Early Power Modeling of New HW/SW Co-Designed Architecture". *Software Professionals Conference (SWPC)*, 2012.
- N. Kosarev, O. Shimko, **D. Ustiugov**, "Critical path study tool for performance analysis of modern architectures". *Software Professionals Conference (SWPC)*, 2011.
- **D. Ustiugov**, "Survey of the state-of-the-art methods for dynamical power estimation and analysis conducted during early stages of CPU microarchitecture development". *MIPT conf., 2012*.
- **D. Ustiugov**, "CPU power estimation method using software cycle-accurate simulator". *MIPT conf., 2011*.
- **D. Ustiugov**, "Practical mobile banking with multi-factor authentication". *MIPT conf., 2011*.
- **D. Ustiugov**, "High-performance instruction cache for multithreaded architecture". *MIPT conf., 2010*.